Machines and Algorithms
http://www.knovell.org/mna

Editorial

Volume 2 Issue 3

Editor: Dr. Farial Syed
Department of Computer Science, University of Regina, Saskatachewan, S4S 0A2, Canada

From the Editor

This issue of "Machines and Algorithms" is, significantly, a food for thought in a wide range of industries. No doubt, the Machine Learning (ML) algorithms are transforming problem-solving and decision-making across various sectors, including healthcare, finance, autonomous systems, and natural language processing. With this rapid advancement of development, the intricacy of issues in implementing, growing, and ethically managing new technologies also intensifies. This editorial examines the principal trends and emerging research domains that will characterize the forthcoming advancements, emphasizing practical applications and the essential demand for competitive methodologies.

The diversity of topics itself, has created an arena of opportunity for the ML enthusiasts reinforced by the authors' contributions in the field of health diagnostics, natural language processing depicting emotional impacts, utilizing internet of things in agriculture, smart billboards, and climate change with respect to the geographic area of Pakistan. These topics and contributions are worth an opinion.

Let us delve into the discussion of discussions; the paper titled, "Ensemble learning model for Classification of Hepatitis C Disease", introduces a hybrid ensemble model aimed at predicting and classifying data for patients with Hepatitis C Virus (HCV). The dataset is acquired from the UCI Machine Learning Repository which is a reputable source for authentic datasets and ensures the data validity and quality. While the study employs four principal classification algorithms; logistic regression, support vector machine (SVM), decision tree, and K-nearest neighbour (KNN) in the training phase, a hybrid ensemble model was developed employing the majority voting technique to integrate these base classifiers, with the objective of alleviating the deficiencies of individual methods. Though the model attained a significant classification accuracy of 94.07%, surpassing single-model methodologies, there is still room for bringing in strong players such as multilayered perceptions and competitive allies. Positively, the paper does recognize the future potential. In a nutshell, the achieved improved prediction accuracy is anticipated to serve as a significant asset for healthcare practitioners, enabling more accurate diagnosis and management of progressive disorders such as Hepatitis C, where early detection and intervention are essential.

Hybrid CNN-LSTM technique demonstrates significant enhancements in emotion categorization accuracy for Massive Open Online Courses (MOOC) reviews when juxtaposed with current literature, consistent with trends observed in recent research employing hybrid deep learning models. Similar studies utilizing CNN-LSTM models for sentiment analysis in social media and consumer reviews have demonstrated competitive outcomes, with accuracies often between 85% and 92%. Nevertheless, few have concentrated on the educational sector, where accurately capturing the nuances of student emotions can be especially difficult due to the varied motives and expectations of learners. Conventional models like solo LSTMs or CNNs frequently (seem to) fail to adequately capture both spatial and temporal dependencies in text, a limitation that the hybrid model overcomes more proficiently. The paper titled, "Emotion Prediction from Online Course Reviews by Using Deep Learning", achieved accuracy of 93.80% in this research exceeds numerous traditional deep learning methods and highlights the increasing significance of hybrid models in emotion analysis. The paper could utilize a more generalized form by incorporating some neutral expression representing a third orientation than "happy" and "sad". A contradictory/mixed emotion that may seem to offset each other, for example "disliked at first but eventually happy now". I believe this

emotional shift exists largely in reviews, and can improve the scope by stressing upon different data handling approaches.

The Internet of Things (IoT) is a developing framework aimed at interlinking various intelligent physical elements for cross-domain advancement. Several IoT-based frameworks have been developed to autonomously manage and monitor agricultural areas with low human involvement. The article, "Integrating the Web of Things in Agriculture: Trends, Challenges and Opportunities" provides a comprehensive analysis of the key components, emerging technologies, security concerns, problems, and future trends in the field of agriculture. Executing a survey on IoT in agriculture encounters numerous significant obstacles. Concerns around data privacy and security may render farmers reluctant to disclose critical information, while connectivity challenges in rural regions can hinder real-time data acquisition. The absence of standards and compatibility among diverse IoT systems hinders comparisons, while varying degrees of technological adoption by farmers, particularly in small or low-income areas, may distort outcomes. The paper in discussion has made a considerable effort to cover most of the aspects. I expect to see environmental variables and their influences on the functioning of IoT devices, and the impact of elevated cost of technology, hence limiting accessibility for more valuable data in future.

"Methodology for the Design and Implementation of Smart Billboards in Pakistan", focuses on the concept of smart billboards that use sensors, data analytics, and AI to deliver targeted and dynamic content. Typically, smart billboards adapt to audience demographics, weather, and traffic patterns in real time. These can analyze viewer participation, foot traffic, and vehicle information to customize ads, making marketing more interactive, efficient, and targeted. Though the paper in discussion has taken a different position of the seller's perspective, adding the potential advantages of a smart billboard may have changed the outcome and hence the acceptance or willingness numbers.

Considering the geographic regions, subject to floods, droughts, and rising temperatures could benefit from AI-based climate change prediction. These methods find climate factor correlations, improving extreme weather and long-term prediction. "Techniques Leveraging from Artificial Intelligence for the Prediction of Climate Change in Pakistan: A Systematic Literature Review", attempts on enabling AI-based predictive models to assist country's policymakers and communities prepare for and to mitigate climate change by analyzing risks to agriculture, water resources, and urban planning. Research questions like, which AI technologies, techniques, algorithms, and evaluation methods are currently employed in predicting climate change, what are the threats, techniques and recommended approaches related to the influence of climate change on agriculture and health have been identified.

To sum up, I must assert that this issue highlights the transformative potential of ML algorithms. The featured studies reflect the growing importance of hybrid models and advanced techniques in improving prediction accuracy. For example, a hybrid ensemble model for Hepatitis C diagnosis achieves high accuracy, while a CNN-LSTM model for emotion analysis in online course reviews demonstrates the advantages of combining deep learning techniques. IoT applications in agriculture face challenges such as data privacy, connectivity, and standardization, but offer promising solutions for autonomous monitoring and management. The concept of smart billboards, driven by AI and real-time data, illustrates the evolving nature of targeted marketing, while AI-based climate change prediction models in Pakistan present innovative ways to address environmental risks. Collectively, these papers emphasize the need for continuous research and development in AI and ML to overcome implementation challenges and unlock new opportunities across diverse fields.

I congratulate all the contributing authors, reviewers, and technical support staff for coming up with this quality research work as a result.

Machines and Algorithms
http://www.knovell.org/mna

Research Article

Ensemble learning model for Classification of Hepatitis C Disease

Sara Ashraf [1], Fatima Bukhari [1,*], Naeem Aslam[1] and Humera Batool Gill [3]

[1]Department of Computer Science, NFC-IET, Multan, 60000, Pakistan

[2] Institute of Computer Science & IT, The Women University, Multan, 60000, Pakistan

[*]Corresponding Author: Fatima Bukhari. Email: fbukhari43@gmail.com

Received: 07 June 2023; Revised: 27 June 2023; Accepted: 28 July 2023; Published: 31 October 2023

AID: 002-03-000026

Abstract: Supervised machine learning is gaining prominence in bioinformatics, particularly in the context of disease diagnosis. This discipline falls under the broader umbrella of artificial intelligence (AI). Hepatitis disease is a leading cause of death, with Hepatitis C being particularly concerning due to the absence of a vaccine. The transmission of Hepatitis C primarily occurs through blood transfusions, contaminated needles, and unsterilized medical instruments. Accurate diagnosis and prediction of Hepatitis C virus (HCV) infection are crucial for effective treatment of affected individuals. Traditional clinical approaches may lead to misdiagnosis in hepatitis cases. Machine learning technologies are enhancing the healthcare sector by improving the accuracy of disease diagnosis and prognosis. This research introduces a hybrid ensemble model aimed at predicting and classifying data related to HCV patients. The dataset utilized, known as HCV+data, is sourced from the UCI machine learning repository. Four classification algorithms such as logistic regression, support vector machine, decision tree, and K-nearest neighbour were employed in the training process. A hybrid ensemble model is created using the majority voting method to integrate various weak or base classification learners. Results demonstrate that the ensemble learning model achieves superior accuracy compared to single-learner machine learning algorithms, with a classification accuracy of 94.07% for hepatitis patients. This model is expected to assist healthcare professionals in accurately diagnosing complex and progressive diseases.

Keywords: Machine learning; Artificial Intelligence, Hepatitis; Ensemble model; Support vector machine Logistic regression; Decision Tree; K-nearest neighbour;

1. Introduction

Hepatitis is a medical condition characterized by inflammation of the liver. The liver plays a crucial role in cleansing the blood, digesting food, and protecting the body from infections. If the liver becomes damaged, it may not function properly. Hepatitis can be triggered by various factors, including excessive alcohol consumption, exposure to environmental chemicals or medications, and certain medical conditions [1].

Hepatitis A, B, and C are the three most common types of viral hepatitis. Hepatitis A has the highest incidence, followed by hepatitis B, and then hepatitis C. Hepatitis C is the most dangerous type of infection. Although there is a vaccine for hepatitis B, it remains unaffordable for many people with low incomes. Different antiviral medicines [2] are used to treat hepatitis C but currently, there is no vaccine for hepatitis

C [3] anywhere in the world. This makes it crucial to take proactive measures and disseminate information about the disease [1].

Patients with chronic hepatitis, which occurs when the virus remains in the body for more than six months, maybe prescribed antiviral medicine by their physicians. This is because chronic hepatitis can cause serious health complications. In the United States, the conditions that affect most people are stroke, diabetes, heart disease, cancer, hepatitis C, and osteoarthritis [4].

In the field of health informatics, the accurate prediction of chronic disease progression is absolutely essential. Chronic diseases can have long-lasting effects, even after treatment, making early and precise diagnosis crucial. This early detection not only leads to advances in disease prevention but also significantly improves the overall efficacy of therapy [4]. Machine learning, a specialized area of artificial intelligence, uses statistical models to make predictions and allows software to "learn" without explicit programming. To create accurate estimates of future output values, machine learning algorithms rely on historical data as input [5]. Feature selection is also critical for producing a more concise and critical description of data by eliminating redundant and unnecessary features.

1.1. Research Objectives

The death rate among Hepatitis patients has been increasing globally. There is a lack of resources and healthcare services for HCV patients, leading to Hepatitis becoming a chronic condition. Early detection and prediction of the disease can save many lives. Previous studies [6] have used single classifiers to classify diseases, but there is a risk of misclassification. To improve prediction accuracy, it is essential to build a model that can forecast diseases and enhance the efficiency of medical treatment. The goal of this study is to analyze different machine learning algorithms and develop a hybrid machine learning model to achieve higher accuracy compared to a single learning prediction model.

RO1: Classify the patients with Hepatitis C Disease by using supervised machine learning classifiers.

RO2: To Perform a comparison analysis between the supervised ML classification algorithms individually and ensemble ML model’s performance in terms of prediction accuracy.

RO3: Develop an ensemble Model to classify blood donors and hepatitis C virus infected patients in terms of classification accuracy.

1.2. Research Questions

RQ1: How can classification algorithms predict Hepatitis Disease?

RQ2: How does the performance of ensemble models stack up against traditional machine learning models (base Classification Algorithm) when it comes to predicting hepatitis

RQ3: How competently our purposed hybrid ensemble model will be fit in terms of accuracy?

This study's main contribution is to provide comprehensive information about the Hepatitis C disease, data preparation, and the prediction and classification of HCV patients using various machine learning algorithms. A hybrid ensemble model was developed to significantly improve classification accuracy. Data preprocessing, data cleansing, and univariate feature selection strategies were applied to achieve superior results. Four machine learning classifiers - Decision Tree, Support Vector Machine, Logistic Regression, and KNN - were rigorously tested on a publicly available dataset from the UCI machine learning repository. The performance of these algorithms was evaluated using a confusion matrix. Additionally, a hybrid ensemble model was developed by combining all the classifiers mentioned above, and its classification accuracy was extensively evaluated. The study presented a comprehensive performance comparison analysis of single algorithms and the hybrid ensemble model using a confusion matrix. The implementation of these classifiers involved using the Python language with sci-kit learn, Seaborn library, and Pandas profiling for finding the correlation between variables. The remaining sections of this paper are laid out as follows:In the related work section, we will review the relevant literature on the use of machine learning for Hepatitis disease. The methods and materials section will explain the methodology behind the

experiments, including all relevant data. In the results and discussion section, we will discuss the experimental results. Finally, in the last section, we will present the conclusion and discuss future directions."

2. Related work

It has been accounted for that hepatitis brings about millions of deaths every year. The Prognosis of hepatitis by traditional techniques is hectic work and needs, costly clinical examination [7].

Hepatitis, also known as hepatitis A and hepatitis B, is soreness of the liver typically brought on by a virus. Clinicians can determine whether or not a patient has hepatitis by examining a collection of datasets and using supervised data mining techniques [8].

In the past different machine learning models have been developed by different researchers.

In [9], the writer used a machine learning model, a support vector machine for hepatitis patient's diagnosis. The dataset was taken from, the UCI Machine learning repository having 155 patients record. The wrapper method is a feature selection technique that was used in this study for achieving higher accuracy. Performance of SVM Classifier, without feature selection and with feature selection was evaluated. SVM extract 74.5% performance accuracy after feature selection. While in the other research, data mining & machine learning techniques have been utilized for the prediction of hepatitis disease [10]. Different Machine learning models like KNN, NB, SVM, RF and Multi- Layer Perceptron had trained. The same dataset [9] of hepatitis patients has been used in this article. For finding out highly correlated features their proposed model used a feature selection procedure named Info-gain related to ranker search Method. Comparison of models had been evaluated on the ratio of following parameters F1-score, Recall, Precision and ROC graph. The accuracy achieved by SVM was 91.14% while RF was 92.41%.

In [11] numerous hepatitis disease diagnosis techniques have been discovered using data mining techniques. These methods were primarily created utilizing single- learning techniques. Additionally, these techniques prevent the facts from being learned collectively. Combining the outcomes of many predictors in classification problems can increase accuracy. This work aims to use the benefits of ensemble learning to come up with an accurate way to diagnose hepatitis. Researchers used groups of Neuro- Fuzzy Inference System, Self- Organizing Maps, and Non-linear Iterative Partial Least Squares to put all the data together and predict hepatitis disease. Data collection, used in this experiment was taken from UCI repository. They also use decision trees to determine which parts of the experimental dataset are the most important. They apply our methodology to a collection of data gathered from the actual world and then evaluate the outcomes in light of the most recent information from other research. After looking at the dataset, they found that their method works much better than the KNN, the ANFIS and the SVM. This method scored 93.06 percent accuracy.

In [12] researchers had been evaluated kinds of ML models. NB, KNN and SVM were used. These classifiers were utilized for the purpose of classification and prediction of data segmentation tools for the purpose of hepatitis illness detection and diagnosis. The dataset for this analysis was available at UCI ML repository. Implementation of these classifiers was done by using MATLAB software. For choosing the most accurate classifier matrices such as accuracy and mean square error were considered. The Naïve Bayes Algorithm was predicting better accuracy of 87% and low mean square error. While in another study [13] three-step approach was used to achieve the results. In the first part of the study, the 13 and 19 accessible dimensions in the datasets for heart disease and hepatitis illness are reduced using the C4.5 decision tree approach, which is part of the CBA software. In the second step, fuzzy pre-processing is used to add weights for heart disease and hepatitis datasets. The first step is to normalize the datasets within the range [0, 1]. These two phases happen at the same time, in the third step of the classification process. They looked at the classification accuracy, sensitivity, and specificity scores, as well as the confusion matrix, of the suggested method to see how well it worked. When the training and testing stages are split, the system can correctly classify 92.59 percent of heart disease datasets and 81.82 percent of hepatitis datasets.

In [14] researcher utilized data of Egyptian patients on liver fibrosis. This dataset has 1385 patients' records. In this dataset was accessed from the UCI repository. In this, different classifiers like KNN, support vector machine, RF, Naïve Bayes, Logistic Regression, Decision Tree and Gradient Boosting were used to sort the data. Additionally, Data balancing, SMOTE, and different feature selection methods were applied. Feature selection carries out in WEKA software. For attribute evaluation, different filter- based methods like Chi-Square, Info Gain, Gain Ratio, and Relief F were used. KNN consider the best model because it accomplished the higher ratio in the different matrices AUROC, Accuracy etc. The best accuracy achieved by the author's model is 94.40%. In this study [15], the author identifies the type and phase of hepatitis disease. They trained SVM and ANN to diagnosis. Various models like SVM, LVQ, GRNN and RBF were utilized in this examination. Dataset was collected from the 2 major hospitals in Mashhad, Iran, having 250 suspected persons. In this researchers differentiate what type of hepatitis has or the person was affected by hepatitis disease or not. The performance of each classifier was compared and analyzed that RBF performs more accurately as compared to other networks. In general accuracy of the diagnosis was approximately 96.4%.

W. Ahmad et al. A hybrid intelligent methodology was produced by combining (ANFIS) and an informative strategy. This methodology was proposed to diagnose a hepatitis disorder that can result in death. The dataset on hepatitis obtained from the ML repository UCI, Data was pre- processed to make it useable before mining. Following the completion of the pre-processing stage, the information gain methodology was applied to considerably cut down on the number of characteristics required to be computed After that, the selected elements were categorized using the ANFIS categorization system. Statistical methodologies were implemented so that the effectiveness of the proposed process could be evaluated. The proposed method achieved the most excellent overall scores in classification accuracy, specificity, and sensitivity analysis, with respective percentages of 95.24 percent, 91.7 percent, and 96.17 percent [7].

In [16] hybrid machine learning approach is used. Clustering and classification are two methods that are often used to start the first stage. These are examples of how the data can be pre-processed. Usually, the second step is based on the results of the first step. To do this, decision trees were used as a classification method, logistic regression as a clustering method, and neural networks as a clustering method. The results of an experiment performed on a dataset taken from the real-world show that the hybrid classification- on-classification technique performs at the highest level. The two-stage decision tree combination produced the best accuracy rate (99.73%) and the fewest Type I and Type II mistakes (0.22 percent and 0.43 percent). This study contributes by proving that hybrid machine learning algorithms outperform standalone ones.

In [17] Researchers used DT, LR, NB and SVM to compare and describe results. Based on the outcomes produced using the SciPy package and the Python language. Logical regression was found to be the method that was accurate 87.17% of the time, according to the study. On the other hand, the algorithm known as the Decision Tree has been shown to have an accuracy of 82.05 percent, making it the method with the second-highest level of precision. The Linear Support Vector Machine, which has the best accuracy at 76.92 percent, is the next algorithm after the Decision Tree Algorithm followed by the NB extracts with 76.92 percent accuracy. In another article [18], investigation of the liver fat of the 36,703 people who took part in the UK Biobank, the first step had been to create an ML technique that might permit precise quantification using abdominal MRI raw data. A Selection of 4,511 subjects whose liver fat had already been determined by Perspectum diagnostics was used to process. The datasets that were employed in the last testing had correlation values of 0.97 and 0.99, and the errors for the two stages were 0.50 and 0.41 percent. It was shown that a method that directly used imaging data rather than merely using clinical data to estimate the amount of fat in the liver was an approach that was much more accurate.

In this [19], included a total of 2235 CHB (chronic hepatitis B) patients. The endpoint was a lack of HBsAg detectability using ECL kits (also known as HBsAg seroclearance).106 CHB patients lost HBsAg antibodies. Dataset had been segmented, 1564 rows for training data and 671 rows as a testing data, with the training data accounting for 70 percent of the total. They developed a model based on 4 ML techniques, RF, LR, DT and XGBoost. AUCs for RF, LR, DT, and XGBoost near to 0.619, 0.829, 0.891 and confidence

interval: 0.677 to 0.683. XGBoost had the AUC overall. The final times for XGBoost, RF, DCT, and LR were all 0.97. For the comparison of previous ML literature ref to Table 1.

Table 1: Comparison ML related work

Ref	Year	Classifiers	Datasets	Findings	Limitations	Results
[20]	2023	Naive Bayes	142 records from UCI dataset	This article identifies age and medical history as important determinants impacting treatment decisions and patient outcomes, but also notes limitations in attribute independence for greater healthcare accuracy.	This study used a small dataset (142) instances with only one classifier, no algorithms comparison. No preprocessing details which resulted in lower predicted accuracy.	86.04%
[21]	2023	Logistic Regression, KNN RF SVM NB	155 instances UCI	This revealed the significance of missing value datasets and feature selection techniques in boosting classification model accuracy and reliability, potentially leading to enhanced decision-making in a range of domains.	A feature selection approach was not used in this investigation. Moreover, this article employed a smaller dataset. Limited hyper parameter optimization, potentially affecting model performance.	93.18 % highest accuracy
[22]	2022	Decision tree and KNN	44 instances	In this, the author uses the KNN algorithm for hepatitis C prediction which had been Optimizing enhance efficiency	This research had resource restrictions. Smaller number of samples had been used.	0.42%
[23]	2021	Support vector machine, NB	Covid-19 dataset	This article compares the accuracy of two machine learning approaches, SVM and Naïve Bayes, before and after selecting the features. The chi-square feature selection approach was implemented.	The accuracy of the SVM classifier had been reduced followed by these feature selection methods.	SVM-83.86% NB-87.09%
[24]	2021	NB, LR, DT, RF, KNN, SVM	Dataset from Kaggle	This research demonstrates the effectiveness of	Extremely unbalanced datasets were utilized. The primary drawback	Highest accuracy

	various machine learning algorithms in accurately predicting stroke based on distinct physiological factors.	of this research was that they utilized a textual dataset compared to a real-time brain imaging dataset.	NB-82%

3. Methodology

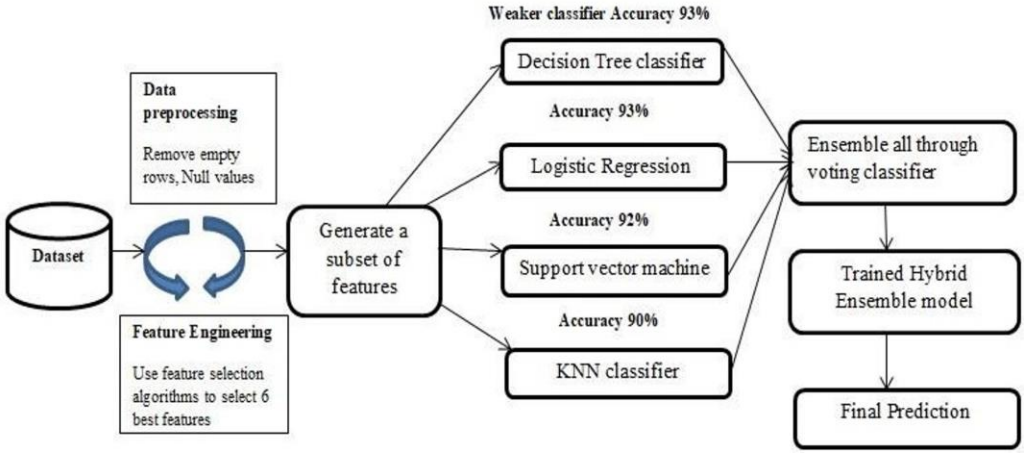


Figure 1: Methodology of the proposed model

This paper aims to develop a model for categorizing people suffering from hepatitis disease. To accurately forecast hepatitis disease, we utilized machine learning (ML) algorithms such as Support Vector Machine (SVM), Decision Tree, Logistic Regression, and K-Nearest Neighbors (KNN) classifiers. Before the classification task, we performed data preprocessing and data cleansing. Python language was used for implementing these classifiers. For feature selection, we employed the Univariate selection method and identified the top nine features from the dataset. We also created a correlation matrix heat map to analyze the Univariate feature extraction. Additionally, data visualization and statistical analysis were carried out on the dataset. Proper preprocessing of the hepatitis disease patient dataset was performed to ensure there were no missing values and noisy data. Subsequently, several machine learning models, as well as a hybrid ensemble, were trained including SVM, Decision Tree, Logistic Regression, and KNN classifiers, among others, for making predictions. The proposed model overview is illustrated in Figure 1.

The various parts of our proposed system include:

- Data Collection & Description
- Preprocessing
- Feature Engineering
- Algorithm Discussion
- The proposed ensemble Model Learning

3.1. Data collection & Description

We gathered this dataset from the UCI Dataset Repository; named HCV+data, this hepatitis disease dataset contains 13 hepatitis features (X (Patient ID/No), Category (binary 0,1), Age, Sex, ALB, ALP, ALT, AST, BIL, CHE, CHOL, CREA, GGT, PROT). This dataset [25] contains 615 samples of hepatitis disease patients.

Table 2: Description of dataset attributes

Features	Data Type	Description
Category	Binary (0,1)	Label
Age	Numerical	Attribute
Sex	Binary (0,1)	Attribute
Choline esterase (CHE)	Numerical	Attribute
Alkaline Phosphatase (ALP)	Numerical	Attribute
Alanine Transaminase (ALT)	Numerical	Attribute
Aspartate Aminotransferase (AST)	Numerical	Attribute
Bilirubin (BIL)	Numerical	Attribute
Albumin Blood (ALB)	Numerical	Attribute
Cholesterol (CHOL)	Numerical	Attribute
Creatine (CREA)	Numerical	Attribute
Gamma-glutamyl Transferase (GGT)	Numerical	Attribute

In Table 2, comprises various columns, including age and gamma-glutamyl transferase (GGT) as the independent variables, and the category column as the dependent variable. The data will undergo thorough analysis using machine learning techniques.

3.2. Preprocessing

In the preprocessing stage, which is an essential part of the procedure, we used an algorithm to exclude normal patients from the health examination data and include only patients with Hepatitis. We removed variables that have missing values and are planning to use an algorithm to fill in any missing values we find. We employed a variety of machine learning techniques and compared the accuracy rates of each of these algorithms.

3.3. Data splitting

We utilized the Scikit-learn library to divide the data, employing the train_test_split function. The data has been confidently split into 80:20 ratios.

3.4. Feature Selection/Engineering

When developing a reliable model, it is important to utilize various machine learning techniques to reduce the number of input variables. This is necessary due to the extensive nature of the hepatitis disease dataset. Doing so helps to decrease the computational requirements of the model and improve its functionality. We have illustrated the advancements made through the use of a correlation matrix and a Heat Map, as well as through the application of Univariate feature selection methods.

3.4.1. Univariate Selection

This test can be done mathematically or statistically to determine well-known characteristics that have the potential to have the most effective relationship toward performance variables. Specifically, the test is used to analyze which variables have the most effective relationship. The Select Best class found in the Scikit package has been used. This class selects a preset number of the most helpful qualities from a given dataset. A wide array of statistical tests, employing several methodologies, have been carried out by Select Best. After implementing the Univariate selection process, the following features have been extracted: GGT,

AST, BIL, ALT, CREA, ALP, ALB, CHE, and Age. Extracted features and their scores are presented in Figure 2 below.

	Specs	Score
10	GGT	11931.274568
5	AST	8179.105101
6	BIL	5707.235379
4	ALT	1861.780434
9	CREA	1700.214095
3	ALP	750.742309
2	ALB	110.200347
7	CHE	71.985583
0	Age	60.388018

Figure 2: Extracted Features

3.4.2. Correlation Matrix with Heat map

Correlation indicates how attributes are related to the target attribute and to each other. When the value of the target variable increases as the feature values increase, it's called a positive correlation. Conversely, a negative correlation exists when the target variable value decreases as the feature values decrease. A heat map simplifies the identification of the dataset's characteristics that are closely related to the target characteristic. We used the Seaborn Library to plot the associated characteristics on the heat map. For analyzing correlation with heat map refer to figure 3 below.

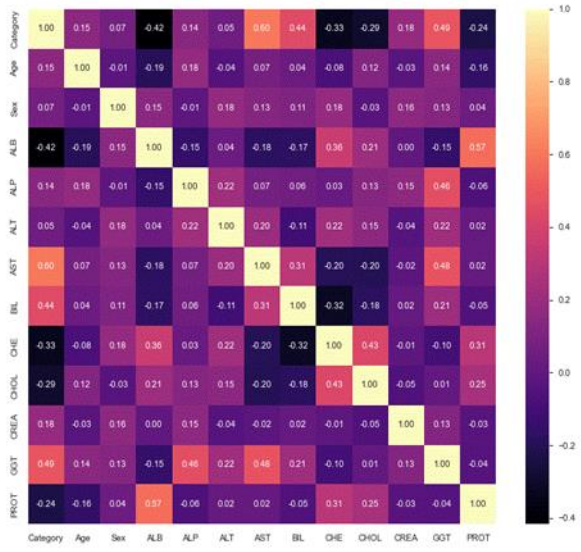


Figure 3: Correlation with Heat map

4. Algorithms Discussion

4.1. Support Vector Machine

In machine learning, Support Vector Machine (SVM) is used to identify a hyperplane in a given N-dimensional dataset. SVM helps to uniquely differentiate the data points in space and can be used to separate multiple classes by using multiple hyperplanes. The decision boundaries of these hyperplanes are used to classify data points of different classes. For our experiment, we utilized the Scikit library for Support Vector Machine and achieved an accuracy score of 92.37% using the linear kernel.

4.2. Decision Tree

In machine learning, a decision tree is a non-parametric algorithm used for classification and regression problems. It has a hierarchical tree structure with root nodes, internal nodes, and child nodes. The strategy used in the decision tree is divide and conquer. We utilized the Scikit library and the DecisionTreeClassifier header file for implementing the decision tree. Our model achieved an accuracy score of 93.22%.

4.3. Logistical Regression

In machine learning algorithms, Logistic Regression (LR) [26] is a linear model used for allocating records to multiple or binary classes with a discrete set. Logistic regression has been employed to solve classification problems such as distinguishing between Hepatitis patients and non-patients. It is also known as an algorithm for predictive analysis. We utilized the Scikit library to perform logistic regression and, with the assistance of the Logistic Regression header file, achieved an accuracy score of 93.22%.

4.4. K-Nearest Neighbors

The K-Nearest Neighbour classifier is a nonparametric instance-based algorithm [27]. This algorithm works based on supervised learning. In this algorithm, new cases are grouped based on similarity and the distance is measured using a distance matrix [28]. The algorithm identifies commonalities between previous and new datasets. K-NN algorithm stores all current data for one or multiple categories and classifies upcoming new records based on their similarity to a specific category. For K-Nearest Neighbors, we utilized the Sci kit library and Gaussian header file, achieving an accuracy of 90.68%.

5. What is Ensemble Learning?

Ensemble learning is a form of machine learning which brings together predictions from numerous algorithms to improve accuracy [29]. In machine learning, ensemble means combining homogeneous weak machine learning algorithms to make a strong predictor model. The implementation of ensemble models intends to lower conversion error. This approach optimizes model prediction efficiency when base models differ significantly and are independent [29]. Noise, variations, and bias are the most common causes of inconsistencies between actual and expected outcomes when using machine learning methods to estimate a target variable. Ensemble approaches use multiple algorithms for machine learning to provide accurate forecasts than a single classifier [30]. The four fundamental categories of ensemble learning algorithms (bagging, boosting, stacking, and voting) are essential for effective modeling of predictions. Here ensemble algorithms learn from a complete the training set or either part of training set [31]. Figure 4. Ref. to generalized overview of ensemble learning model.

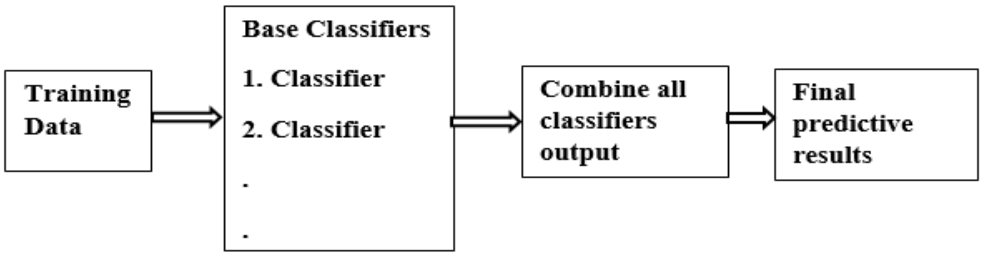


Figure 4: Generalized overview of ensemble learning model

- ***Bagging***

Bagging is known as Bootstrapping. Bootstrapping, similar to random replacement sampling, can help determine variance as well as bias in a dataset [32]. The term bagging is referring to the reduction in the prediction variance by making an extra dataset for learning with combined repetitions. This is done to generate multiple set of the original dataset.

- ***Boosting***

Boosting algorithms work on weighted average patterns to convert weak learner classifiers into strong ones. The term Boosting means to readjust the weights of all last classified dataset. In this process the original dataset is divided into different subsets to train the classifiers [33]

- ***Stacking***

Stacking is a learning-based technique that combines the fundamental outputs. When the ultimate selection of features is a linear framework, the stacking is commonly referred to as "model blending" or generally "blending"[33]. In stacking heterogeneous weak/base learners, trained and combined the results based on different weak models' predictions by using a voting classifier.

- ***Voting***

A voting classifier is an ensemble learning method, which aggregates the predictions from multiple independent models to generate final outcomes. In voting method major two types of voting exist hard voting & soft voting. The final prediction is generating by collective estimated probabilities from all base classifiers and select that class with the greatest average probability [34]. Voting approach in ensemble learning is one of the best approaches in previous studies [35].

5.1 Proposed Hybrid Ensemble Model

In the first step, we trained several machines learning algorithms, including Logistic Regression, Decision Tree, Support Vector Machine, and K-Nearest Neighbor (weak learners). We utilized different variations of each model to create a robust machine-learning algorithm. The term "hybrid ensemble" indicates that we used a combination of diverse weak machine learning algorithms. In Python, we utilized a majority voting classifier to generate variations of specific models. The proposed system improves upon existing models by combining base learners into a model that makes predictions through a majority vote using a hard majority classifier. For evaluation, we employed the majority Voting Classification process to consider the class mostly predicted by the weak learners as the final predicted class by the ensemble model. The algorithms' performance was assessed using popular evaluation metrics: accuracy, precision, recall, and F1 score.

6. Performance Evaluation through Confusion Matrix

Here we have used confusion matrix to present the performance of a weak machine learning algorithm.

1. (A) Accuracy

Factor, which determines the predictions accuracy, is known as algorithm accuracy.

$$\frac{TP+TN}{TP+FN+FP+TN} \tag{1}$$

In equation (1), TP, TN, FN,and FP show the number of TruePositive Negatives, True Negatives, False Negatives and False Positives.

2. (P) Precision

Factor, which used to find out the measure of classifier's [36], [37] the equation of precision is illustrated in Eq. (2)

$$\frac{TP}{TP+FP} \tag{2}$$

3. R) Recall

Factor, which used to find out the completeness and sensitivity of the algorithm, is known as recall of algorithm. The equation of Recall is illustrated in Eq. (3)

$$\frac{TP}{TP+FN} \tag{3}$$

4. F1-score

Factor which defined the precision and recall in the form of weightedaverage [36], [38]. The equation of F1-score is illustrated in Eq. (4)

$$\frac{2}{\frac{1}{Recall}+\frac{1}{Precision}} \tag{4}$$

7. Experimental Results

In this paper, different Machine learning algorithms such as support Vector Machine, Logistic Regression, Decision Tree, and KNN have been used.

5. Why uses these classifiers?

An ensemble model can be constructed by utilizing the power of two or more classifiers. Different Machine learning, data mining and deep learning classifiers can be used for classifying and predicting of diseases. The classifiers which are used in this study called as base/weak classifiers [35]. By combining these classifiers, we can make strong predictive model which can predict /classify data more efficiently as compared to single learning classifiers. There is a chance of miss classification if we use single learning technique, but by ensemble model there is a less chance of miss classification. Results demonstrate that SVM scored 92 %, Logistic Regression 93 %, Decision Tree 93 % and KNN scored 93% accuracy. Logistic Regression & Decision Tree achieved the same percentage of accuracy. The dataset is taken from the UCI ML repository and then preprocessing has been done on it. Divided the whole set of data: a training set and a testing set. A hybrid ensemble model for the Prediction & Classification of HCV patient's data is developed and implemented. This model extracted 94.07 % accuracy which was higher than all mentioned Machine learning algorithms. For implementing all these classifiers python languages with sci-kit learn, Seaborn library is used. Pandas profiling is used for finding a correlation between variables. The performance of all of these is evaluated through confusion matrix based on the following parameters: Accuracy, Precision, Recall and F1- Score.

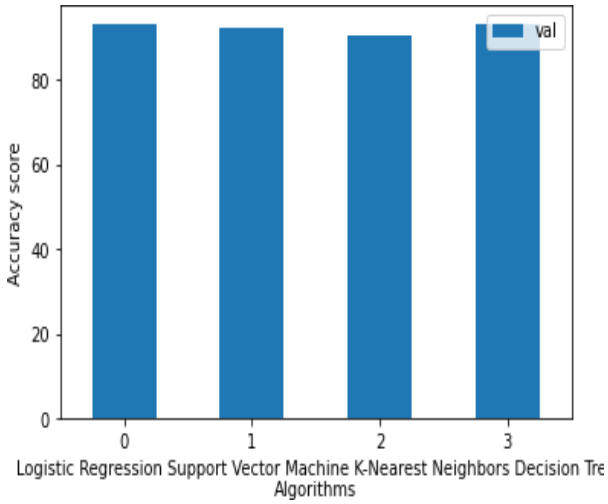


Figure 5: Graphical representations of all Algorithms

In fig. 5 graphical representation of machine learning algorithms are presented Support vector machine achieved 92% accuracy, Decision Tree 93%, Logistic Regression 93% and KNN 90%.

Table 3: Accuracy Comparison

Algorithms	Accuracy	Precision	Recall	F1-score
SVM	92%	0.92%	0.92%	0.92%
DT	93%	0.93%	0.93%	0.93%
LR	93%	0.90%	0.93%	0.92%
KNN	90%	0.89%	0.91%	0.89%
Proposed EM	94%	0.92%	0.94%	0.93%

Table 3 illustrate the accuracy comparison of all machine learning algorithms and the proposed hybrid ensemble model.

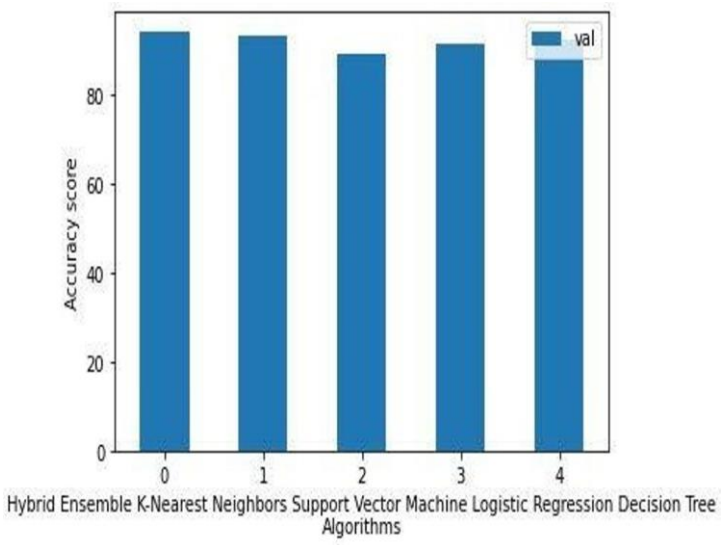


Figure 6: Graphical representations of Hybrid with of all algorithms

Lastly in Fig.6. classification comparison of proposed model and all individual machines learning models. Performances are presented in graphical form on the Hepatitis dataset. The prediction rate of the proposed hybrid ensemble model is 94.07% higher than all weak machine learning algorithms.

8. Conclusion

Healthcare sector has need to rapid improvement in disease diagnosis and prognosis. Detecting and diagnosing the deadly diseases at early stages is a big challenge now days. There is keen need to construct a model which can predict disease more accurately as compared to traditional treatment methods. The main focus of this research is to classify the hepatitis by analyzing different Machine learning classifiers. For this purpose, four machine learning algorithms are applied for the Classification of Hepatitis C patients. The dataset that has been used in this is publicly available at UCI. This research evaluates the classification algorithm's performance on hepatitis disease patients by using Python language and improves the accuracy. It discovers that the individual model is providing accuracy of up to 93%. The proposed ensemble model is including a Support vector machine, Decision Tree, Logistic regression and KNN has been developed. The proposed model extracted 94.07% accuracy. This predictive model will help Doctors and physicians in making an accurate identification of hepatitis disease patients. We conclude by our work and from the available literature, no model is completely accurate in all aspects. There are limitations to consider, such as the reliance on the quality of the dataset from UCI. This dataset may not include broad patient demographics or may contain incomplete data. Additionally, there is the potential neglect of interactions

between variables that could provide additional insights. There can be a chance of error and miss classification by outliers etc. Accuracies of algorithms may vary on different datasets due to the number of records, decision parameters and so on. Diversity of ensemble models is available that extract different levels of accuracy on the different types of data.so it is noted that no model is completely efficient and accurate in all aspect. However, the implications are significant, as this approach offers a more efficient and accurate diagnostic method for early detection of hepatitis C. This could improve patient outcomes through timely and targeted interventions and drive further innovation in the application of machine learning in medical diagnostics and healthcare.

9. Future Work

In future studies, there is still a need to improve the accuracy of this model. Furthermore, we will train this model in the prediction of various medical datasets. Dataset used in this study is small due to limitations on available resources. Moreover, we will also improve the model's performance by using a larger dataset. We will also implement deep learning and data mining algorithms.

References

[1] A. Orooji and F. Kermani, “Machine Learning Based Methods for Handling Imbalanced Data inHepatitis Diagnosis,” *Front. Heal. Informatics*, vol. 10, no. 1, p. 57, 2021, doi: 10.30699/fhi.v10i1.259.

[2] NHS.(n.d.).https://www.nhs.uk/conditions/hepatitisc/treatment/#:~:text=Hepatitis%20C%20medicines,for%208%20to%2012%20weeks.

[3] Hepatitis C,Mayo-clinic .(n.d.). https://www.mayoclinic.org/diseases-conditions/hepatitis-c/diagnosis-treatment/drc-20354284.

[4] M. Nilashi, H. Ahmadi, L. Shahmoradi, O. Ibrahim, and E. Akbari, “A predictive method for hepatitis disease diagnosis usingensembles of neuro-fuzzy technique,” *J. Infect. Public Health*, vol. 12, no.1, pp. 13–20, 2019, doi: 10.1016/j.jiph.2018.09.009.

[5] A. H. Roslina and A. Noraziah, “Prediction of hepatitis prognosis using support vectormachines and wrapper method,”*Proc. - 2010 7th Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD 2010*, vol.5, no. Fskd, pp. 2209–2211, 2010, doi: 10.1109/FSKD.2010.5569542

[6] H. Hartatik, M. B. Tamam, and A. Setyanto, "Prediction for diagnosing liver disease in patients using KNN and Naï ve Bayes algorithms," in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*, 2020: IEEE, pp. 1-5.

[7] W. Ahmad *et al.*, “Intelligent hepatitis diagnosis using adaptive neuro-fuzzy inference system and information gain method,” *Soft Comput.*, vol. 23, no. 21, pp. 10931– 10938, 2019, doi: 10.1007/s00500- 018-3643-6.

[8] M. Rouhani and M. M.Haghighi, “The diagnosis of hepatitis diseases by support vector machines and artificial neural networks,” *2009 Int. Assoc. Comput. Sci. Inf. Technol. Spring Conf. IACSIT-SC 2009*, pp.456–458,2009, doi: 10.1109/IACSIT-SC.2009.25.

[9] F. Penin, J. Dubuisson, F. A. Rey, D. Moradpour, and J. M. Pawlotsky, “Structural Biology ofHepatitis C Virus,” *Hepatology*, vol. 39, no. 1, pp. 5–19, 2004, doi: 10.1002/hep.20032.

[10] K. Ahammed, M. S. Satu, M.I. Khan, and M. Whaiduzzaman, “Predicting Infectious State ofHepatitis C Virus Affected Patient’s Applying Machine Learning Methods,” *2020 IEEE Reg. 10 Symp. TENSYMP 2020*, no. June, pp. 1371– 1374, 2020, doi:10.1109/TENSYMP50017.2020.9230464.

[11] M. Rouhani and M. M.Haghighi, “The diagnosis of hepatitis diseases by support vector machines and artificial neural networks,” *2009 Int. Assoc. Comput. Sci. Inf. Technol. Spring Conf. IACSIT-SC 2009*, pp.456–458,2009, doi: 10.1109/IACSIT-SC.2009.25.

[12] V. Vanitha and D. Akila,“Detection and Diagnosis of HepatitisVirus Infection Based on HumanBlood Smear Data in MachineLearning Segmentation Technique,”*2021 9th Int. Conf. Reliab. InfocomTechnol. Optim. (Trends Futur. Dir.ICRITO 2021*, pp. 1–5, 2021, doi:10.1109/ICRITO51393.2021.959648 2.

[13] X. Tian *et al.*, “Using machinelearning algorithms to predict hepatitis B surface antigenseroclearance,” *Comput. Math. Methods Med.*, vol. 2019, 2019, doi: 10.1155/2019/6915850.

[14] K. Polat and S. Güneş, “A hybrid approach to medical decision support systems: Combining featureselection, fuzzy weighted pre- processing and AIRS,” *Comput.Methods Programs Biomed.*, vol. 88, no. 2, pp. 164–174, 2007, doi: 10.1016/j.cmpb.2007.07.013.

[15] M. J. Nayeem, S. Rana, F.Alam, and M.A.Rahman,“Prediction of Hepatitis Disease Using K-Nearest Neighbors, NaiveBayes,Support Vector Machine,Multi-Layer Perceptron and RandomForest,” *2021 Int. Conf. Inf. Commun.Technol. Sustain. Dev. ICICT4SD2021 - Proc.*, pp. 280–284, 2021, doi:10.1109/ICICT4SD50815.2021.9397 013

[16] M. E. Haas *et al.*, “Machine learning enables new insights into genetic contributions to liver fat accumulation,” *Cell Genomics*, vol. 1,no. 3, p. 100066, 2021, doi: 10.1016/j.xgen.2021.100066.

[17] A. A. ABRO, E. TAŞCI, and A. UGUR, “A Stacking-based Ensemble Learning Method for Outlier Detection,” *Balk. J. Electr. Comput. Eng.*, vol. 8, no. 2, pp. 181– 185, 2020,doi:10.17694/bajece.679662.

[18] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.

[19] L. A. Bull, K. Worden, R. Fuentes, G. Manson, E. J. Cross, and N. Dervilis, “Outlier ensembles: A robust method for damage detection and unsupervised feature extraction from high-dimensional data,” J. Sound Vib., vol. 453, pp. 126–150, 2019, doi: 10.1016/j.jsv.2019.03.025.

[20] Yulhendri, Malabay, and Kartini, “Correlated Naïve Bayes Algorithm To Determine Healing Rate Of Hepatitis C Patients,” International Journal of Science, Technology & Management, vol. 4, no. 2, pp. 401–410, Mar. 2023.

[21] Sachdeva, R. K., Bathla., Rani, P., Solanki, V., and Ahuja, R., “A systematic method for diagnosis of hepatitis disease using machine learning,” *Innov Syst Softw Eng*, vol. 19, no. 3, pp. 71–80, Jan. 2023

[22] H. Mamdouh Farghaly, M. Y. Shams, and T. Abd El-Hafeez, "Hepatitis C Virus prediction based on machine learning

[23] Rosidin, S., Muljono, Shidik, G. F., Fanani, A. Z., Zami, F. A., and Purwanto, “Improvement with Chi Square Selection Feature using Supervised Machine Learning Approach on Covid-19 Data,” *International Seminar on Application for Technology of Information and Communication (iSemantic),* Oct. 2021, doi: 10.1109/iSemantic52711.2021.9573196.

[24] Sailasya, G., and Kumari, G. L. A., “Analyzing the Performance of Stroke Prediction using ML Classification Algorithms,” (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 12, no. 6, pp. 539–545, 2021.

[25] “UCI Machine LearningRepository: HCV data Data Set.” [Online]. Available: https://archive.ics.uci.edu/ml/datasets/HCV+data.

[26] Le Cessie S, Van Houwelingen JC (1992) Ridge estimators in logistic regression. J R Stat Soc Ser C Appl Stat 41:191–201

[27] Kumari R, Jose J (2011) Seizure detection in EEG using Biorthogonal wavelet and fuzzy KNN classifier. *Elixir Hum Physiol* 41:5766–5770

[28] Altay O, Ulas M (2018) Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children. In: 2018 6th International symposium on digital forensic and security (ISDFS). IEEE, pp 1–4

[29] Latha, C.B.C.; Jeeva, S.C. “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques”. *Inform. Med.* Unlocked 2019, 16, 100203. [CrossRef]

[30] Ali, R.; Hardie, R.C.; Narayanan, B.N.; De Silva, S. Deep learning ensemble methods for skin lesion analysis towards melanoma detection. In Proceedings of the 2019 IEEE National Aerospace and Electronics Conference

(NAECON), Dayton, OH, USA, 15–19 July 2019; pp. 311–316

[31] Tanuku, S.R.; Kumar, A.A.; Somaraju, S.R.; Dattuluri, R.; Reddy, M.V.K.; Jain, S. Liver Disease Prediction Using Ensemble Technique. In Proceedings of the 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 25–26 March 2022; pp. 1522–1525.

[32] Jongbo, O.A.; Adetunmbi, A.O.; Ogunrinde, R.B.; Badeji-Ajisafe, B. Development of an ensemble approach to chronic kidney disease diagnosis. *Sci. Afr. 2020, 8, e00456*.

[33] Igodan, E.C.; Thompson, A.F.-B.; Obe, O.; Owolafe, O. Erythemato "Squamous Disease Prediction using Ensemble Multi-Feature Selection Approach.: *Int. J. Comput. Sci. Inf. Secur*. IJCSIS 2022, 20, 95–106.

[34] Ashri, S.E.; El-Gayar, M.M.; El-Daydamony, E.M. HDPF: Heart Disease Prediction Framework Based on Hybrid Classifiers and Genetic Algorithm. IEEE Access 2021, 9, 146797–146809.

[35] Mahajan, P., Uddin, S., Hajati, F., & Moni, M. A. (2023, June). Ensemble learning for disease prediction: A review. In *Healthcare* (Vol. 11, No. 12, p. 1808). MDPI.

[36] T. I. Trishna, S. U. Emon, R. R. Ema,G. I. H. Sajal, S. Kundu, and T. Islam, "Detection of Hepatitis (A, B, C and E) Viruses Based on Random Forest, K-nearest and Naïve Bayes Classifier," *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–7, 2019, doi:10.1109/ICCCNT45670.2019.894445.

[37] M. Rizzetto, S. Hamid, and F. Negro,"The changing context ofhepatitis D," *J. Hepatol.*, vol. 74, no. 5, pp. 1200–1211, 2021, doi: 10.1016/j.jhep.2021.01.014.

[38] M. S. M. Serafim, V. S. dos SantosJúnior, J. C. Gertrudes, V.G.Maltarollo, and K. M. Honorio,"Machine learningtechniques applied to the drug design and discovery of new antivirals: a brief look over the past decade,"

Machines and Algorithms
http://www.knovell.org/mna

Research Article

Emotion Prediction from Online Course Reviews by Using Deep Learning

Muhammad Ahtesham Noor[1], Mustajeeb ur Rehman[2,*] and Muhammad Saddique[1]

[1]Department of Computer Science Bahria University Islamabad, 44790, Pakistan
[2]Department of Computing, National University of Modern Languages, Islamabad, 44790, Pakistan
[*]Corresponding Author: Mustajeeb ur Rehman Email: mustajib.rehman@numl.edu.pk
Received: 06 June 2023; Revised: 28 June 2023; Accepted: 31 July 2023; Published: 31 October 2023
AID: 002-03-000027

Abstract: Massive open online courses (MOOCs) emerged as a pivotal solution for distance learning during the COVID-19 pandemic, effectively breaking down barriers related to age, gender, and geography. This study focuses on developing a robust and precise emotion classification model using advanced deep learning techniques, specifically targeting the reviews on Coursera's online learning platform. Our research dives into key questions surrounding the performance of different deep learning models, particularly comparing the Long Short-Term Memory (LSTM) network with a hybrid model that combines Convolutional Neural Networks (CNN) and LSTM. We hypothesize that this hybrid approach not only enhances the predictive accuracy of emotion analysis but also outperforms traditional supervised learning methods. By analyzing a comprehensive dataset of 140K course reviews, we demonstrate that the hybrid CNN-LSTM model, when coupled with sophisticated word embedding techniques, achieves superior results, reaching a peak accuracy of 93.80%. This work underscores the potential of hybrid models in capturing the complexities of human emotions in educational content, offering valuable insights for improving online learning experiences.

Keywords: COVID-19; Emotion Classification; Deep Learning; LSTM; CNN-LSTM Hybrid Model; Coursera Reviews

1. Introduction

The world sees standards shift their daily activities in the ongoing coronavirus (COVID-19) pandemic the way we meet people, relate, operate businesses, shopping, and learn online. Such world disasters directly affect our lifestyle, however. Corona outbreaks keeping schools, restaurants, and borders open.

Emotions add an influential role in human communication. Emotions are more valuable than successful communication. A lot of necessary points are balanced learning in humans is based on emotions. Emotion analysis and Affected computing are essential for AI and other research fields in many situations and businesses, both large and small. Emotion Analysis is used to create automated analysis reviews and opinions from Texts and documents. Online learning platforms played a massive role during the pandemic. Many students use many online learning platform channels for learning. This research will emphasize discovering collective reactions to student reviews' expressed emotions and opinions to online learning platforms. It will highlight to be given analyzing students' responses towards the performance of online courses. Emotion Analysis is the study of analyzing people's opinions, appraisals, assessments, behaviors, and feelings toward objects of different kinds write in a text. Like, services, organizations, individuals' issues, events, topics, and features. Emotion analysis operates on every possible domain, like consumer

products, telecommunications, health care, e-commerce, education, ser- vice, financial, political campaigns, elections. When an Online Course Platform or institution wanted a community Emotions, it organized surveys and polls using online web platforms. It got reviews from these platforms, like forum discussions and blogs. From all this, they will get a lot of information about student opinions about the courses. Still, it is not easy to find valuable knowledge from that many student reviews.

Every site contains a massive amount of text that is not very easily translated reviewed. The reader will have to struggle to identify, refine the appropriate information. People are excited to know other people's Emotions for their profits, so they will ask friends and family when they need a view as per human psychology. Emotion analysis is applicable in organizations for effective decision-making. We need automated Emotion analysis. In 1999, the high-demanding Natural language processing studies (NLP). In previous years, web mining and opinion mining were useful. Emotion analysis grows up in every possible domain and industry that contains the text. Through sentimental analysis, convert the division of the text into groups determines the polarity of feelings.

Polarity characterizes a state as possessing opposing emotions. Polarity is positive, negative, neutral. Retrieves people's thoughts about the object known as Opinion Mining and describes the view expressed in a text extracted and analysis known as Emotion Analysis. Both methods are used to solve multi-step classification problems. There are three significant classification types of Aspect Level, Document Level, and Sentence Level. The document's polarity is based on a sentence or aspect; either is positive, neutral, or negative. The review data is considered the primary information unit in the Document-level analysis to classify the document's positive or negative opinion. Sentence level analysis is used to determine whether a sentence represents a standard positive or negative viewpoint. In many classifying applications, the texts are on sentence level, or the document level provides necessary detail. It does not mean that the positive view the reviewer writes the review is always opinionated positive. Likewise, a negatively reviewed document does not describe that the reviewer does not like all the entity's features. In a typical opinionated, the reviewer writes both negative, neutral, and positive opinions. However, most current approaches define the overall polarity, sentence, paragraph, or leaves not in detail. To get these hidden patterns, we need Aspect Based Sentiment Analysis. It identifies the aspects of entities. In mining and summarizing Analysis Aspect Based Sentiment is more complex. Aspect Based Sentiment Analysis systems have been made for various entities like Computers, Travel, Services, and Restaurants or Movie reviews in the past few years. Aspect Based Emotion Analysis systems use texts like (forum discussions and messages, product reviews, and comments) discussing Aspects like Phone selling reviews. The system retrieves the frequently discussed features. Aspects feature like Screen, Size, and Price of the entity and predict the sentiment that de- fines each aspect.

Table 1: Summary of the Previous Models used

Studies	SVM	NB	KNN	RF	DT	ANN	CNN	LSTM	Lexicon
[1]	-	-	-	-	-	-	-	-	T
[2]	-	-	-	-	-	-	-	-	T
[3]	-	-	-	-	-	-	-	T	-
[4]	T	-	-	-	-	-	-	-	-
[5]	-	T	-	-	-	-	-	-	-
[6]	T	-	-	-	-	-	-	-	-
[7]	-	-	-	-	-	-	-	-	T
[8]	-	T	-	-	-	-	-	-	-
[9]	T	-	-	-	-	-	-	-	-
[10]	T	-	-	-	-	-	-	-	-
[11]	T	-	-	-	-	-	-	-	-

[12]	T	-	-	-	-	-	-	-	-
[13]	T	T	-	T	-	-	-	-	-
[14]	T	-	-	-	-	-	-	-	-
[15]	T	-	-	-	-	-	-	-	-
[16]	T	T	-	T	-	-	-	-	-
[17]	-	-	-	-	-	-	-	-	T
[18]	-	-	-	-	-	-	-	-	T
[19]	-	-	-	-	-	-	-	-	T
[20]	-	T	-	-	-	-	-	-	-

[SVM] → Support Vector Machine [NB] → naive bayes

This section provides a synopsis of previous research in this area. The subject of inquiry. Adamopoulos [21] investigated user-generated content. online MOOC reviews to assess the effect of factors Course platforms and universities, for example, significantly impact student retention. Valakunde and Patwardhan conducted another report [15]. Students' performance summary comments were subjected to sentiment classification.

Nave Bayes used the frequency-inverse text frequency (TFIDF) approach, and support vector machines were used as machine learning algorithms. Wen et al [19]. Discussion boards for MOOCs. Students' dropout characteristics were identified using forum posts from three MOOCs in this scheme. In this model, the author collects posts from three MOOCs to classify student dropouts. The study reveals a strong relationship between the mood expressed in course forum feedback and MOOC completion rates. Altrabsheh et al [22] conducted another analysis. introduced a sentiment analysis focused on machine learning Methods for obtaining students' learning-related emotions from feedback reviews. In this approach, Student feedback is used to measure feelings about various courses, including leadership skills, database management, engineering, and Twitter information. Three traditional Ngram versions (namely unigram, bigram, and trigram). The Naive Bayes algorithm, as well as help vector machines in the classification process, the maximum entropy classifier, as well as the random forest algorithm, were used. Adinolfi and colleagues [23] The sentiment analysis system investigates student satisfaction on various online course sites, such as online course learning diaries and Twitter. Students and teachers' behavior has also been documented. Similarly, Bogdan [24] Sentiment research enhances course material and recognizes MOOC Courses' views. In another review, opinion analysis was an effective method for extracting views from students' comments on instructors' results [25]. For slant examination, the creator utilized assistance In the investigation, vector machines and an irregular woods calculation are talked about Moreno-Marcos et al [26] To reveal patterns in under- study action, characterize noticed results for vocabulary-based and AI put together conclusion classification calculations with respect to MOOC criticisms. Strategic relapse, support vector machines, choice trees, irregular timberland, and the Naive Bayes calculation are a portion of the regulated learning procedures utilized in this strategy. A random forest strategy was proposed by the study's findings. Data regarding the online success of higher education institutions drivers can also be extracted using text mining and sentiment analysis [11]. In the presented paradigm, research on topic modeling and profiling has extended to higher educational institutions Abdi et al [27]. Describe a multi-document, query-based, opinion-oriented summarizing method. To obtain sentiment orientation and personal knowledge, the suggested approach employs sentiment analysis. The summarizing module found related sentences from the user's questionnaire. Belbachir and Boughanem [28] To represent the questionnaire and text for sentiment analysis, language models from information retrieval were utilized. In addition, Al-Smadi et al. [29] Morphological, syntactic, and semantic characteristics were employed in the sentiment analysis process. Until recently, Bustillos et al [3] gave an intensive audit of AI and profound learning approaches for assessment mining in a shrewd learning climate. A few AI calculations (like Bernoulli Nave Bayes,

multinomial Nave Bayes, support vector machines, straight help vector machines, stochastic inclination plunge, and K-closest neighbor calculation) and profound learning models, (for example, convolutional neural organization and long transient memory) were utilized in this work, just as a few profound learning structures, (for example, convolutional neural organization and long momentary memory). With a classification accuracy of 88.26 percent, the highest prediction efficiency was attained using a profound learning-based engineering. Likewise, Cabada et al [30] On educational reviews, two deep learning architectures (convolutional neural network and long short-term memory) were used, yielding an arrangement precision of 84.32 percent. Nguyen and Nguyen [31] For emotion analysis in video comments, a bidirectional neural Ngam bidirectional LSTM word embedding architecture was presented. A word with semantic and social data in short and significant distance cycles has been communicated in the given framework. Lin et al [32] On student evaluations of education, researchers looked at the prediction yield of information-based and AI-based slant examination systems. The Study, Lo´pez et al [33], On educational resource networks, a system focused on opinion mining and semantic profiling was presented. As of late, Onan [34] scientists took a gander at the prescient accomplishment of conventional characterization calculations, outfit models, and profound learning calculations on understudy appraisals of instructing. Another examination discovered, Wang et al [18] introduced a half-breed profound learning-based plot for slant investigation dependent on convolutional neural organizations and long.

2. Research Methodology

Deep learning and machine learning are two types of learning. Models for sentiment analysis methods have been defined. To accumulate a book corpus on the MOOC stage, we scraped the reviews from Coursera. We got the review data from a well-known online course platform. We collect course reviews from many Courses and Fields. We get almost 1835 Course review data, such as algebra, accounting, aerospace engineering, computer science, agriculture, data science, and education. We have collected around 140k reviews from different courses. Course feedback can be managed and stored from an online learning platform when students review their enrolled courses. We are using a point scale, and the overall average score describes the standard that has been processed in preparation for a course assessment. Using the content scores assigned by the students, we produced a labeled corpus. The reviews classified 1 or 2 as negative, while scores of 4 or 5 were positive. We receive both positive and negative feedback following the marking procedure.

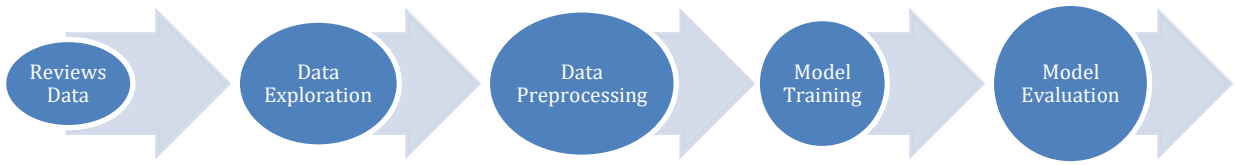


Figure 1: Workflow of study

2.1. Libraries

We are using TextBlob, nltk, plotty, Tensor-Flow. Keras genism, matplotlib, pandas to analyze Emotion In the second step, we import the dataset by using the panda's library.

We have a dataset that contains

- CourseId — Coursera course identifier
- Review — Customer Review Text
- Label — Customer Rating between 0 and 5

2.2. Data Exploration and Preprocessing

Adding sentiment score for review

- So, we will be doing the sentimental analysis for each review using TextBlob.
- TextBlob will assign a sentiment score to each review, ranging from -1 (negative sentiment) to 1 (positive sen- timent), with 0 being neutral.

Courses with the maximum number of reviews

Convert five classes into two classes

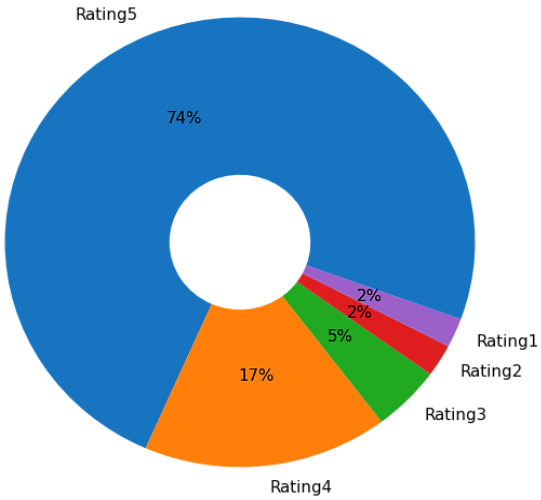


Figure 2: Percentage of Ratings

The following table 2 describes the emotions-oriented classes used to review the courses.

Table 2: Dataset converted with Emotion Orientation

Emotions Orientation	Course Review
Happy	Great, clear, concise explanations on everything. A collaborative effort with teams and specialists worldwide sharing their contributions to this body of knowledge. Particularly congratulate the effort to make this a well-rounded conversation both on the biological effects and economic, political, and communications side of this problem.
Sad	This course does not say anything about digitization which is the core subject of the digital wave.
Happy	Wonderful! Simple and clear language, good instructors, great stuff! Highly recommend!
Sad	This course contains no new material. It does not tell you anything, but rather displays well recognized information in an exciting manner. There are more productive ways to spend your time.

Because the main goal is to distinguish good and negative feedback, we divided the five-star rating system into two categories: (Happy = 1 and Sad = 0)

Figure 3: Common Words Used in Positive Reviews

2.3 Training Modal

2.3.1. Tokenize text data

We use the top 20000 unique terms due to computational costs. Tokenize the comments first, then turn them into sequences. I maintain 50 words to keep the number of words in each remark to a minimum.

2.3.2. Build neural network with LSTM

An embedding layer is the first layer in the network. The layer allows the system to extend each token into a larger vector, allowing the network to represent a word meaningfully. The layer accepts a first argument of 20000, which is the size of our vocabulary, and a second input parameter of 100, which is the dimension of the embeddings. The third argument is input length, which is set to 50 and represents the length of each comment sequence.

2.3.3. Train the network

There are about 1.6 million comments, and it takes a while to train the model in a MacBook Pro. To save time I have used only three epochs. GPU machines can be used to accelerate the training with more time steps. I split the whole datasets as 60per for training and 40per for validation.

2.3.4. Build Hybrid Neural Network Model with LSTM and CNN

The LSTM model performed admirably. However, training three epochs takes an eternity. Improving the network by adding a Convolutional layer is one technique to reduce training time. Image processing gives rise to Convolutional Neural Networks (CNN). They apply a filter to the data and generate a higher-level representation. They have been proven to perform very well for text, while lacking the sequence processing capacity of LSTMs.

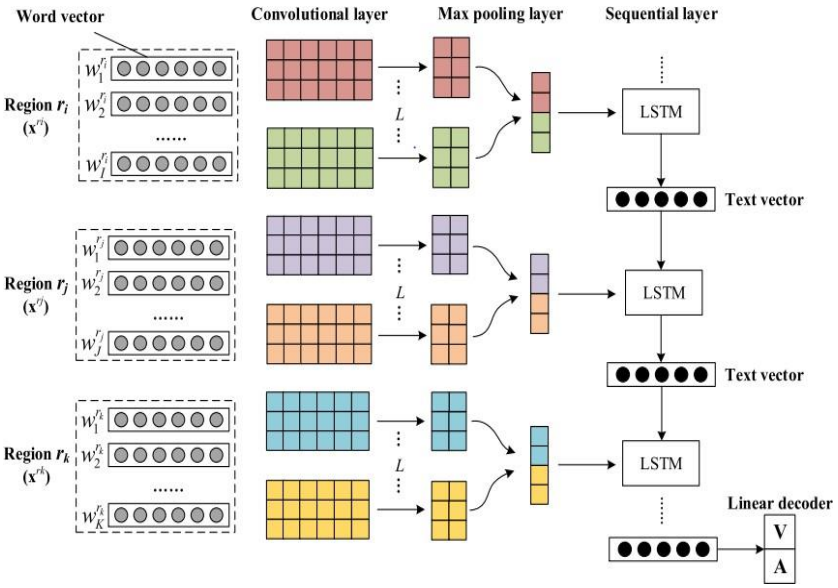


Figure 4: CNN-LSTM Model Working

3. Results

In both deep networks, we trained our data set on LSTM and CNN-LSTM Hybrid Modal. We used a supervised approach using labels on the below bar chart. Based on emotion classification, Result-LSTM and Result-CNN outperformed the baseline strategy on student reviews. LSTM obtains the best performance getting a score of 93.38% and the baseline is 93.30% Using CNN-LSTM Hybrid Model, Achieving the performance result, 93.80 % and the baseline is 91.05% performed better and improved the efficiency of Model. On our data set, CNN-LSTM Hybrid Approach performed better. CNN, which focuses only on partial information, and LSTM, which focuses only on global information in the context, both performed best among the contrast models we tested BiLSTM. On the other side, the focus was reversed on capturing the context's global information. However, it also ignored the comments' incomplete information. The CNN- LSTM model outperformed the baseline model by roughly 2per, since CNN was able to capture the properties of the incomplete data. and LSTM was used to complement the characteristics of the global information in the context, considerably improving the accuracy rate. Due to the present rapid creation of linguistic literature, common words have tended to be given new meanings, popular phrases have continued to be generated, and semantics might be read differently in various phrases, particularly in comment texts.

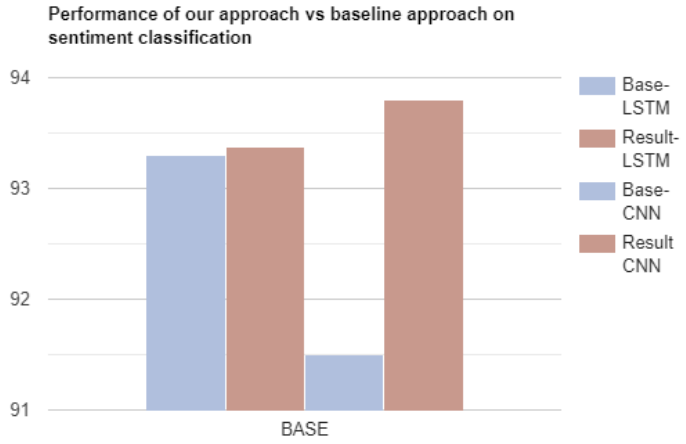


Figure 5: F1 Score percentage

4. Conclusions

Massive open online courses (MOOCs) are the cutting-edge alternatives to distance learning in the covid-(19) epidemic on these learning platforms, participants have not faced any problems like geography-related, race, age, gender barriers. Emotion analysis is applied to Data from an online learning site to collect feedback on a course, encouraging teachers to develop their teaching methods and students to have ac- cess to high-quality instructional materials. In this paper, we examined a corpus that included 140k MOOC evaluations using machine learning, deep learning techniques. In the first neural network modal, we used LSTM and a single embedding layer. In the second modal, an additional 1D convolutions layer was built to reduce training time on top of the LSTM layer. The network architecture is the same as in Model-2, but I use pre-trained glove 100-dimension word embedding. We consolidated two profound learning structures with two- word inserting plans (word2vec and GloVe) (i.e., Convolutions neural organization and long momentary memory by having these two models using a half breed design) in the deep learning-based method. This research aims to improve the accuracy and efficiency of Emotion classification with high predictive quality in the educational domain, MOOC feedback. The algorithms for predictive performance and deep learning have been written. This study has functional applications. According to the empirical research findings, deep learning architectures outperform conventionally supervised learning approaches. Long Short-term Memory Networks (LSTM) and CNN We accomplished the best prescient outcomes for deep learning architectures. The scientific research shows that the GloVe word embedding system produces better predictions. Choosing a suitable representation scheme is a key step in designing AI based estimation order plans. In this regard, the trial review provides detailed analytical evidence for various text representation systems, both supervised and unsupervised. Learning methods, Deep learning architectures for educational data processing, which could act as the field's benchmark methodological findings. We present the first large free online course review corpus, which could be helpful for future research.

Acknowledgement: I am deeply grateful to my parents for their unwavering support and encouragement throughout my studies. Their love and belief in me have been my greatest motivation. I would also like to express my sincere thanks to my teachers, for their invaluable guidance and mentorship. Their wisdom and patience have been instrumental in shaping this research.

Thank you all for your support and inspiration.

References

[1] Manyika, James. "Big data: The next frontier for innovation, competition, and productivity." *McKinsey Global Institute* 1 (2011).

[2] Sundmaeker, Harald, Patrick Guillemin, Peter Friess, and Sylvie Woelfflé. "Vision and challenges for realising the Internet of Things." *Cluster of European research projects on the internet of things, European Commision* 3, no. 3 (2010): 34-36.

[3] Yasin, Ahmad, Yosi Ben-Asher, and Avi Mendelson. "Deep-dive analysis of the data analytics workload in cloudsuite." In *2014 IEEE international symposium on workload characterization (IISWC)*, pp. 202-211. IEEE, 2014.

[4] Villars, Richard L., Carl W. Olofson, and Matthew Eastwood. "Big data: What it is and why you should care." *White paper, IDC* 14 (2011): 1-14.

[5] Capotosto, James. "Data opportunities and risks: the dynamic of public, personal, and commercial interest." *Journal of community safety and well-being* 2, no. 1 (2017): 18-21.

[6] Davenport, Thomas H., and Jill Dyché. "Big data in big companies." *International Institute for Analytics* 3, no. 1-31 (2013).

[7] Global, A., and O. Megatrends. "Big Data Analytics in Supply Chain: Hype or Here to Stay." *Big data analytics. Accenture company's report* (2014): 1-20.

[8] Provost, Foster, and Tom Fawcett. "Data science and its relationship to big data and data-driven decision making." *Big data* 1, no. 1 (2013): 51-59.

[9] Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. "Business intelligence and analytics: From big data to big impact." *MIS quarterly* (2012): 1165-1188.

[10] Memon, Mashooque Ahmed, Safeeullah Soomro, Awais Khan Jumani, and Muneer Ahmed Kartio. "Big data analytics and its applications." *arXiv preprint arXiv:1710.04135* (2017).

[11] Batty, Michael, Kay W. Axhausen, Fosca Giannotti, Alexei Pozdnoukhov, Armando Bazzani, Monica Wachowicz, Georgios Ouzounis, and Yuval Portugali. "Smart cities of the future." *The European Physical Journal Special Topics* 214 (2012): 481-518.

[12] Bryant, Randal, Randy H. Katz, and Edward D. Lazowska. "Big-data computing: creating revolutionary breakthroughs in commerce, science and society." (2008).

[13] Lucini, Filipe R., Flavio S. Fogliatto, Giovani JC da Silveira, Jeruza L. Neyeloff, Michel J. Anzanello, Ricardo S. Kuchenbecker, and Beatriz D. Schaan. "Text mining approach to predict hospital admissions using early medical records from the emergency department." International journal of medical informatics 100 (2017): 1-8.

[14] Khan, Zaheer, and Tim Vorley. "Big data text analytics: an enabler of knowledge management." *Journal of Knowledge Management* 21, no. 1 (2017): 18-34.

[15] Thet, Tun Thura, Jin-Cheon Na, and Christopher SG Khoo. "Aspect-based sentiment analysis of movie reviews on discussion boards." *Journal of information science* 36, no. 6 (2010): 823-848.

[16] Yu, Hong, and Vasileios Hatzivassiloglou. "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences." In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pp. 129-136. 2003.

[17] Piryani, Rajesh, Devaraj Madhavi, and Vivek Kumar Singh. "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015." *Information Processing & Management* 53, no. 1 (2017): 122-150.

[18] Qazi, Atika, Alireza Tamjidyamcholo, Ram Gopal Raj, Glenn Hardaker, and Craig Standing. "Assessing consumers' satisfaction and expectations through online opinions: Expectation and disconfirmation approach." *Computers in Human Behavior* 75 (2017): 450-460.

[19] Li, Fei, Meishan Zhang, Guohong Fu, Tao Qian, and Donghong Ji. "A Bi-LSTM-RNN model for relation classification using low-cost sequence features." *arXiv preprint arXiv:1608.07720* (2016).

[20] Chiu, Chaochang, Nan-Hsing Chiu, Re-Jiau Sung, and Pei-Yu Hsieh. "Opinion mining of hotel customer-generated contents in Chinese weblogs." *Current issues in tourism* 18, no. 5 (2015): 477-495.

[21] Kim, Erin Hea-Jin, Yoo Kyung Jeong, Yuyoung Kim, Keun Young Kang, and Min Song. "Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news." *Journal of Information Science* 42, no. 6 (2016): 763-781.

[22] Schumaker, Robert P., A. Tomasz Jarmoszko, and Chester S. Labedz Jr. "Predicting wins and spread in the Premier League using a sentiment analysis of twitter." *Decision Support Systems* 88 (2016): 76-84.

[23] Korkontzelos, Ioannis, Azadeh Nikfarjam, Matthew Shardlow, Abeed Sarker, Sophia Ananiadou, and Graciela H. Gonzalez. "Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts." *Journal of biomedical informatics* 62 (2016): 148-158.

[24] Lopez Barbosa, Rutilio Rodolfo, Salvador Sánchez-Alonso, and Miguel Angel Sicilia-Urban. "Evaluating hotels rating prediction based on sentiment analysis services." *Aslib Journal of Information Management* 67, no. 4 (2015): 392-407.

[25] Bagheri, Ayoub, Mohamad Saraee, and Franciska De Jong. "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews." *Knowledge-Based Systems* 52 (2013): 201-213.

[26] Hai, Zhen, Kuiyu Chang, Jung-Jae Kim, and Christopher C. Yang. "Identifying features in opinion mining via intrinsic and extrinsic domain relevance." *IEEE transactions on knowledge and data engineering* 26, no. 3 (2013): 623-634.

[27] Rodrigues, Ramon Gouveia, Rafael Marques das Dores, Celso G. Camilo-Junior, and Thierson Couto Rosa. "SentiHealth-Cancer: a sentiment analysis tool to help detecting mood of patients in online social networks." *International journal of medical informatics* 85, no. 1 (2016): 80-95.

[28] Yu, Yang, and Xiao Wang. "World Cup 2014 in the Twitter World: A big data analysis of sentiments in US sports fans' tweets." *Computers in Human Behavior* 48 (2015): 392-400.

[29] Ceron, Andrea, Luigi Curini, Stefano M. Iacus, and Giuseppe Porro. "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France." *New media & society* 16, no. 2 (2014): 340-358.

[30] Zhou, Qingqing, Rui Xia, and Chengzhi Zhang. "Online shopping behavior study based on multi-granularity opinion mining: China versus America." *Cognitive Computation* 8 (2016): 587-602.

[31] Paltoglou, Georgios, and Mike Thelwall. "Twitter, myspace, digg: Unsupervised sentiment analysis in social media." *ACM Transactions on Intelligent Systems and Technology (TIST)* 3, no. 4 (2012): 1-19.

[32] Saif, Hassan, Yulan He, Miriam Fernandez, and Harith Alani. "Contextual semantics for sentiment analysis of Twitter." *Information Processing & Management* 52, no. 1 (2016): 5-19.

[33] Lin, Yuming, Xiaoling Wang, You Li, and Aoying Zhou. "Integrating the optimal classifier set for sentiment analysis." *Social Network Analysis and Mining* 5 (2015): 1-13.

[34] Zavattaro, Staci M., P. Edward French, and Somya D. Mohanty. "A sentiment analysis of US local government tweets: The connection between tone and citizen involvement." *Government information quarterly* 32, no. 3 (2015): 333-341.

Machines and Algorithms
http://www.knovell.org/mna

Research Article

Integrating the Web of Things in Agriculture: Trends, Challenges and Opportunities

Faria Khan[1], Fatima Bukhari[1], Ratabah Mehras[1], Laiba Rehman[2] and Khadija Kanwal [2, *]

[1]Department of Computer Science, NFC Institute of Engineering & Technology, Multan, 60000, Pakistan
[2]Institute of Computer Science and Information Technology, The Women University, Multan, 60000, Pakistan
[*]Corresponding Author: Khadija Kanwal. Email: khadijakanwal.6022@wum.edu.pk
Received: 05 June 2023; Revised: 04 July 2023; Accepted: 1 August 2023; Published: 31 October 2023
AID: 002-03-000028

Abstract: The last few years have introduced the Web of Things (WoT) as a revolutionary technology in the agricultural field that provided new approaches to existing problems IoT applications face in agriculture. WoT promotes integration and cooperation among different devices, structures, and platforms and contributes to the evolution of other technologies. The objective of this research is to present the state of the art in the research on WoT based agriculture, its current problems, and future opportunities. To this end, this paper provides a systematic literature review (SLR) of the articles published from 2010 to 2020 based on the research type, approach, and application domain of the selected studies. Besides, this study gives a classification of WoT-based agricultural applications and offers the notion of WoT-based Smart Agriculture by presenting a conceptual model. Last but not the least, this SLR presents the emerging research areas and directions, issues for further research at the end of this paper.

Keywords: Web of Things (WOT); Agriculture; Smart Farming; Applications; Semantic web technology; Internet of things (IOT); Taxonomy; Systematic literature review (SLR);

1. Introduction

Technological development also influenced the people's life style along with shifting the old agricultural style to Smart Agriculture (SA) and Smart farming (SF) [1]. Technologies evolved to make agriculture a manageable and smart Farming or the Smart Farming [2]. Another term that has been adopted for smart agriculture is Precision Agriculture which is defined for managing the soil, water, weather conditions through ICT technologies [3]. PATs approached for maximum production and for decreasing the environmental causes [4]. The relevant information concerning the cost, the policy of long-term payback and the farms is provided by offering the training and technical assistance for the farmers [5]. The European Union has made it its political aim to foster agricultural productivity and sustainability through increased generation and diffusion of innovations that enhance the sector's overall competitiveness [6]. Precision agriculture could be a processing cycle whereby agricultural data is gathered for analysis, evaluation so that particular decisions can be made in managing the fields that were categorized under smart farm technologies (SFT) that includes GNSS [7] and mapping technologies. However, the idea of precision agriculture was inherent in the process or farmers' judgment even decades ago. Thus, the idea was called Precision Agriculture. A technical explanation for the selected agricultural industry was also provided in detail [8]. Internet of things (IOT) also paved a new opportunity to the agriculture field as it brought technological

advancement in each agricultural sector [9]. It is vital to understand that the evolvement of the Internet of things did not only increase numbers of devices but also the amount of data it produced and was expected to be 50 billion in 2020 [10]. The increase is not only in the quantity of devices but also in the of data they produce as estimated by CISO [11], 500 zett of data have been generated This rapid growth in devices and data has in a significant rise in the number of APIs. This research presents the Web of (WoT), emphasizing its main, the function of Web Technology (S), and its role in API design [12]. The Web of Things improves the efficiency and usability of the expanding smart ecosystem by incorporating web services into the creation of smart devices. These devices feature web capabilities, which allow for interaction and communication online [13]. As countless devices connect to the internet, a new Web of Things is taking shape, showcasing virtual representations of both physical and conceptual entities that become more accessible through web technologies [14].

Primary characteristics about three main technologies are defined in Figure 1 below:

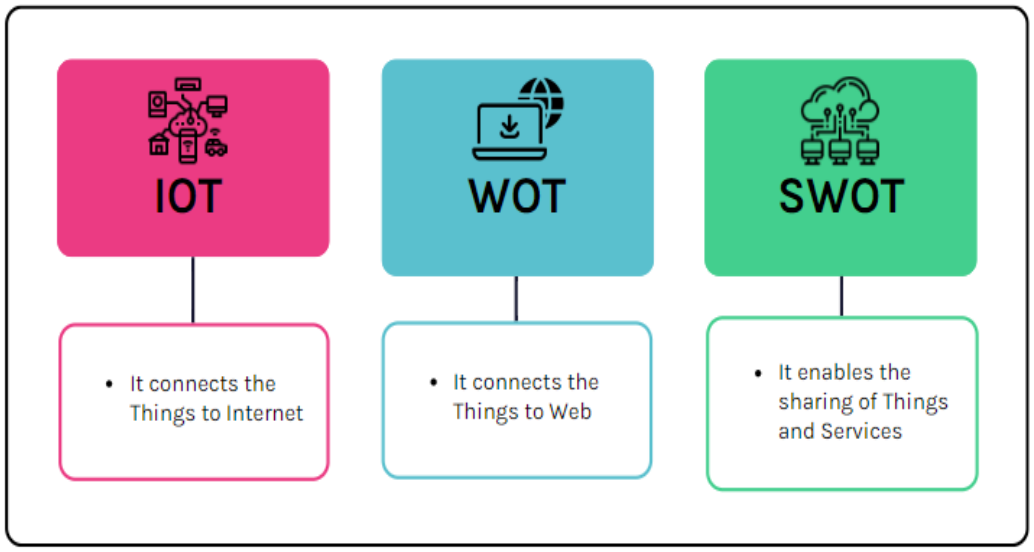


Figure 1: Primary idea of IOT, WOT, SWOT

The Web of Things (WoT) has created new research opportunities in agriculture, allowing innovative techniques such as the redesign of farmland and alternative farming practices that meet current market needs. This is made possible through the implementation of web-based decision-making systems [15]. One study [16] suggested a framework utilizing the Virtual Test Environment for Distributed Systems (VTEDS) to create smart sensor hubs with plug-and-play functionality. The system showcased remarkable efficiency, with intelligent sensors in Europe capable of self-registering and configuring remotely via a cloud service in South America in under three seconds, and delivering data to users in less than two seconds. Additionally, frameworks have been established to facilitate the shift from the Internet of Things (IoT) towards the Web of Things (WoT) [17]. Although the application of WoT has recently seen significant growth, many devices encounter various encryption issues. Semantic Web Technologies are regarded as the most effective means to address these challenges [18].

1.1. Integrated WOT Model

The agricultural sector is experiencing a swift rise in the adoption and incorporation of various new technologies. A range of systems, frameworks, infrastructures has been created that leverage IoT, wireless devices [19], machinery, and cameras for monitoring and managing crop activities. Multiple studies have

suggested infrastructure diagrams that depict solutions integrating Web of Things (WoT) services to improve interoperability and communication among devices and various levels of the models., [20], [21].

The model is shown in Fig. 2 below:

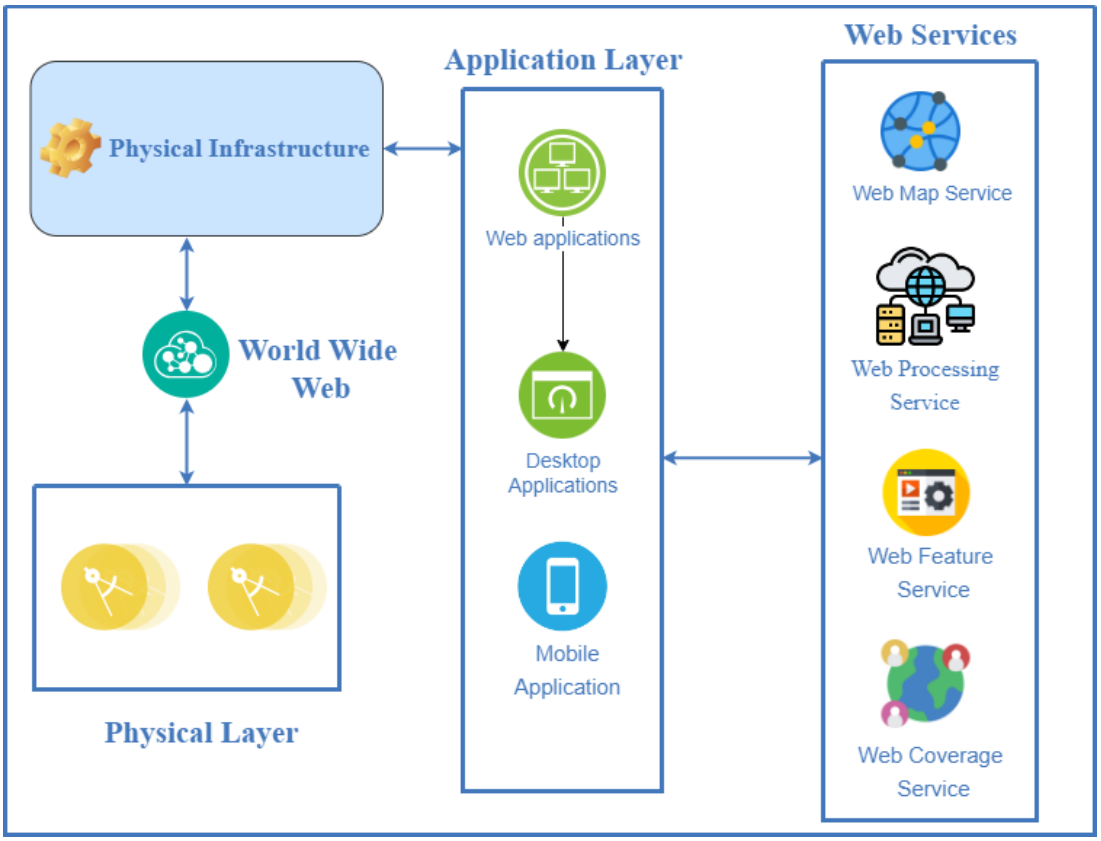


Figure 2: Open geospatial web physical model for PA

This model explains the integration of various web services that provide the interfaces for different analysis in the agriculture system by which the farmers can check, control and monitor the farms by using their mobile or web applications. A physical infrastructure layer added that provide the web services such as Web Map Service (WMS) for to provide the map of farms, Web Processing Service (WPS) to perform processes like to predict the level of water and production of crops. Web Feature Service (WFS) provide the deploying points of wireless sensors in soil as well as observe the weather state. Web Coverage Service (WCS) used to get the data collected from the farm's locations. All these services provide the interfaces for getting and sharing the agricultural data and enhanced the agricultural analysis and operations. Data can be published and share on different devices with the help of these web services and farmers can control and monitor the farms through web applications, Mobile applications and Desktop applications.

1.2. WOVT Model for Communication

In order to link IoT smart devices to the cloud for data exchange, storage, and communication purposes, another infrastructure incorporated the WOT technology. The model is shown in Figure 3 below.

This model is showing that a Web of Virtual Things layer is added between the IOT smart objects layer and Cloud layer for the purpose of interoperability, data management and abstraction to make the communication easy between diverse devices, protocols and cloud servers. This layer provides the standardization for storing and sharing the data with secure connection and abstraction feature hide the complex functions of diverse protocols and services of different clouds. This WoVT layer is used to

integrate the diverse devices and clouds services for their smooth communication purposes. In the last few years, Linked Open Data has merged multiple sorts of data, which makes it accessible and valuable in a variety of industries globally [22]. The exponential proliferation of data has accelerated the rapid growth of incorporation of data on the World Wide Web, which is presently moving toward the World-Wide Semantic Web (WWSW) [23]. The linked community of humans, gadgets, and services across the internet has generated significant real-time data in numerous forms, codes, and forms, lead to massive volumes of knowledge and operations [24]. Several algorithmic strategies have been created to effectively save and retrieve the data in real time in required forms, combining both current and unique technologies [25].

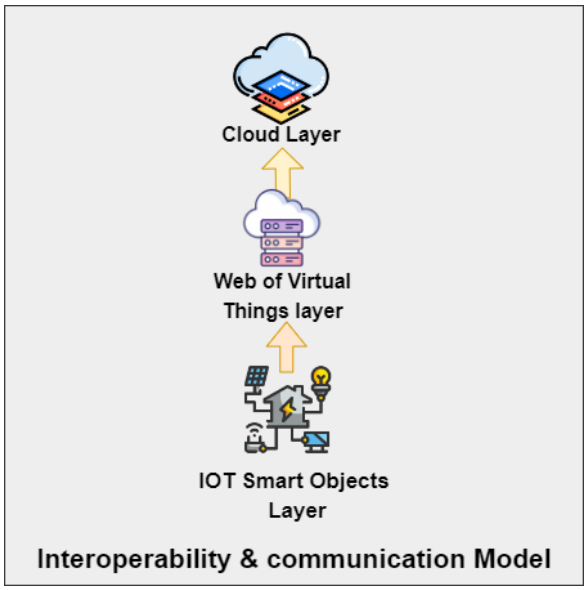


Figure 3: WoVT communication Model

Semantic web approaches improve the intelligence and efficacy of internet applications [26]. MIMOS developed a semantic technological foundation that improved and introduced new functionalities [27]. To help organizations publish agricultural data as well as models, the FAO developed AGROVOC, the world's biggest multilingual agricultural vocabulary [28]. As wireless connection has grown in popularity, smart devices and sensors have become indispensable for monitoring and controlling in real time in agriculture, changing traditional farming techniques into smart farming [29]. The use of big data, AI, and cloud technology have improved data quality and security, but they also have met with drawbacks such as high processing costs [30–31]. Furthermore, Automation of Aeronautical Vehicles (UAVs) have already been integrated with IoT to increase crop development and virus detection [32-33].

The integration of multiple technologies into smart farming has lowered fertilizer consumption while increasing efficiency and lowering costs [34]. Technologies such as 3D graphical visualizations for growth of crops and blockchain for transparent food supply networks are improving precision agriculture [35]. Blockchain technology, in particular, has gained importance for applications in precision agriculture by increasing transparency, decentralized governance, and trustworthiness. A study on combining blockchain and IoT in precision agriculture identified corresponding possibilities as well as challenges [36-37]. Several nations have created schemes for smart farming that include IoT, Blockchain, the Web of Things (WoT), web-based message protocols, and other technologies including MQTT, AMQP, DDS, REST HTTP, and WebSocket [38].Such technologies and procedures have been assessed for their efficacy, usefulness, and productivity, with recommendations for modifications to help the agricultural sector [39].

The study's primary objective is to provide a thorough literature review of previous research that has identified the value of WOT integration with agricultural applications in resolving issues with the handling, monitoring, and control of gadgets, software, and other interconnected technologies in the agricultural sector. The study looks into cutting-edge research that is being done to use WOT to solve problems. This

study is novel because it offers a taxonomy of the agricultural sectors where applications of WOT are used to address technological and technical issues in agriculture. Finally, the remaining portion included various sections that summarized this paper's contribution. The background of the WOT-agricultural sector is described in Section 2. The research method is presented in Section 3 and involves the use of defined research questions, inclusion and exclusion criteria, and a search string to locate directly relevant research articles in the WOT-agricultural field. The outcomes of the selected studies that were obtained by retrieving the data for conducting the research are defined in Section 4. A wot-based model about the smart agriculture, opened issues and encounters, research gaps, future directions for further research movements, and a description the threat to the validity study are all explained in Section 5. The paper's conclusion is given in the final section, number 6.

The sections of paper have shown in Figure. 4 below:

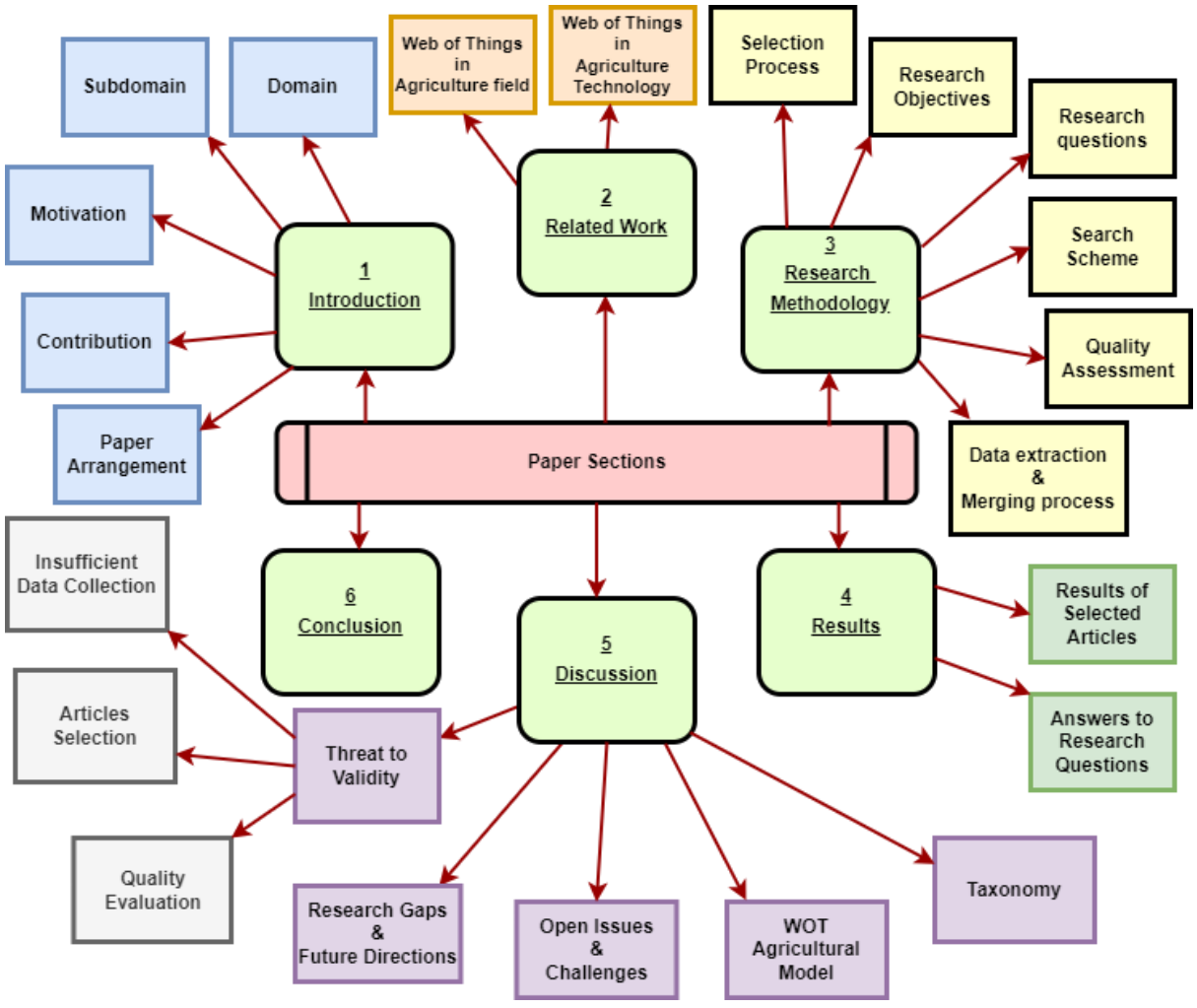


Figure 4: Paper Sections

2. Related Work

Web of Things has been a promising innovation for settling a few farming applications, frameworks reconciliation, their taking care of, checking, and dynamic issues. Incorporating effective precision features, data analysis, and remote sensors into the existing agricultural frameworks, infrastructures, and systems was WOT. In order to find solutions to the challenges that are arising, a number of studies have been carried out on the emergence of WOT in the field of agricultural technology. Subsequently, hardly any commitments have been made to assess the investigations that introduced the answers for the agrarian issues.

On [20], research on the Web of Things was conducted. For monitoring purposes, both wired and wireless [19] sensors played a crucial role in Precision Agriculture. However, the interoperability of disparate sensors in systems became a problem, so a physical infrastructure with open geospatial service integration was offered for processing, disseminating, and incorporating the surveillance details on the web information servers. Many people have access to Precision Agricultural data thanks to a web sharing service. A Service Oriented Architecture (SOA) placement was made for the purpose of keeping and upgrading Precision Agriculture (PA) setups by joining distributed workable web services. Numerous web service configurations were planned and offered, but they lacked the ability to share information and incorporate with a variety of devices. In addition, numerous studies have attempted to resolve security issues arising from disparate networks, devices, and services. This study compared a number of previous infrastructure features, including SOA, which lacked support for sensors, process sharing, unknown models, and complexity issues. As a result, the proposed method now incorporates all of these solutions as well as a few more features. The physical, application, business, and sensor layers comprised the proposed SOA-based architecture. Sensors and devices for transmitting measurements of environmental factors are found in the physical layer. The application layer provided a means of communication between systems and devices. A SOS detecting layer in light of SWE utilized for dispersing and putting away the information over different detecting gadgets. The business layer is the more elevated layer for dissecting rural cycles as well as information. Workers can get measuring facts along with a great deal of information on the Web by using a basic observation, which offers statistics with absence of detecting systems to communicate with it. Using SOS, WCS, WMS, and WFS services, an advanced observing device offers the crucial information results. The following work offered the web enabled detectors plan and accomplished the detecting devices, procedures coordination, and ready to move, offer, and incorporate the information on Internet. However, there are still some features for multi-web facilities that needed in further work [40]. WebGIS and IOT, two technologies, were the focus of a study [41] on their use in precision agriculture. Subsequent to examining their placement, pros, and cons in China, a Framework (PAMS) installed in a chosen farmland by coordinating the WebGIS and IoT advances. It involved the IoT for having discernment precision capacity and WebGIS handled organization geological data effectively. There were four modules in this system: Agribusiness the board stage, spatial data foundation stage, and versatile client and IOT framework stage. This framework had the combination of a few high-level strategies, for example, IoT, WebGIS, Web and correspondence, Area Based Assistance, GPS, RS to gather, move as well as distribute information and facilitated the clients in checking and dealing with the creation. PMAS consists of six parts called: the informational structure, catalogue, local setups, WebGIS, productivity handling setup, and mobile client. By studying the available insights, circumstances, and farming conditions, the method could be upgraded, innovative modules can be adopted and make it useful in the Particular area. Another related study [42] was subject of additional investigation. Data can prepare to groundbreaking thoughts, ideas connected with the concerned disciplines. The associations have fabricated information-based frameworks that could be useful to the scientists and overall population to figure out the data, holes, and limits of their connected trains and could give groundbreaking thoughts and upcoming effort. Although agrarian research with information remained accessible on the internet, they were not arranged in software or catalogs. The informational frameworks have been utilized for exploring dismissed yields by using data for getting valuable outputs. Data directory strategy observed the guideline vocabularies and fostered libraries quite a while by the agricultural local area and these libraries facilitated in the improvement of meaningful items to give immediate inquiry responds to office. CFF UCKB (Harvests for the Future's Underutilized Yields Information Based Framework) gave path to agrarian scientists, associations, and people in general for instructive and research purposes. The plan contained the information gateway as the Internet interface, Information data set, and information toolbox for mining and tracking down limits through semantic items. Information was stored, shared, and exchanged in large part thanks to the knowledge-based systems and tools. This term was used in agriculture to refer to book systems, social forums, the internet, and knowledge networking systems. Agriinfo is an expert-developed web-based agricultural system, and the document management system AGRIS database can be found at agris.fao.org. An information the board framework (KMS) was one more innovation that aided to discovery of examination holes out of available works.

Knowledge-based systems, which use "if-then" question to answer questions, make up a large number of decision and expert systems [43]. Numerous web-based systems are utilized in the agricultural sector for various purposes [44, 33].

Linked-Data, which makes use of meta-data standards to create templates for sharing and publishing data online, is another method for providing data online. In order to connect the knowledge and data systems and provide the datasets necessary for research, the system made use of the same standards. Person to person communication has begun building the area explicit component for giving field-related data. The harvests interpersonal organization frameworks utilized this office to connected the ranchers, researchers, and analysts to speak with one another to takes care of the issues. The development of a collaborative research environment (CRE) for underdeveloped harvesting issues was yet another objective of utilizing semantic technologies. These were likewise utilized for creating archives and studies regarding meaningful towards devices and people. A different method of incorporating IT into integrated pest management was defined in a subsequent study. This Decision-making computing device carried multiple strategies to control all pest species to their best advantage. A tool that could forecast the pest based on the specific area was needed by agricultural experts and researchers. Thus, a forecasting the pest and web-based data framework with real-time functionality was introduced. This work resulted in the creation of software that utilized Web technologies for weather data storage and data mining. It downloaded files with the.txt extension and connected the network stations of the NOA. This software was upgraded during the subsequent phase to become a real-time pest forecasting system that made use of the internet. There were two phases to the software development process: On the first stage, a customized WebGrabber opened the NOA (National Observatory of Athens) climatic system to obtain weather data. On the second stage, a web interface was used to perform daily real-time population calculations from the web and store the data on a MYSQL local server. For forecasting and decision-making systems, an algorithm was developed. The first step of this blog is BIOFIX, which examines the progression of pest classes. A subsequent stage put away information. Tertiary stage assessed the heat status. Later a state monitors the level of temperature was implied. The next step was to calculate input for generating the results. It was simple to use, inexpensive, and virtually extensible thanks to its straightforward architecture [45].

The traditional and smart agriculture differences have shown in the Figure 5 and Figure 6 below.

Figure 5: Traditional Agriculture

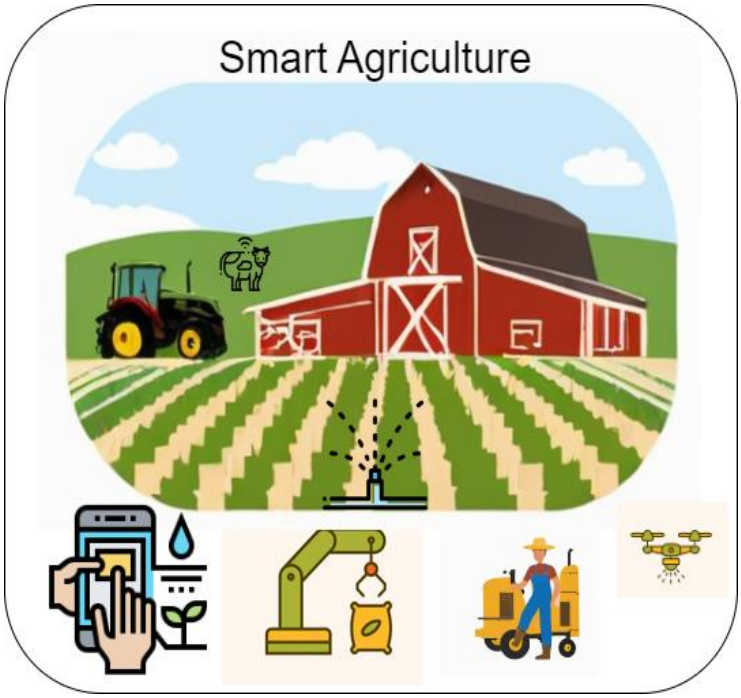


Figure 6: Smart Agriculture

"A Web-Based IoT Solution for Monitoring Data Using MQTT Protocol" [46] was the title of another study. An online engineering was proposed to track, screen, and examine continuous rural information. The primary objective concerned about observing tasks of IoT device of agrarian sector. Systems kept and shared the data from the sensors over the Internet. A platform of three-tiered, with various components such as the dashboard and storing part made up the Web application architecture. The data that was sent to the users from sensors and other monitoring devices was transferred through a protocol known as Message Queue Telemetry Transport (MQTT). Asynchronous communication, lower complexity, and lower power consumption were the main features of this protocol. Data was captured and transferred to servers before being given out to users. Yet, there is a chance of advancement in provincial region organizations. Further related work was done [47]. Giving crops the proper watering to obtain desired output is crucial. Weather patterns impacted the procedure for supplying water thus the ranchers required a framework aided to settling on great as well as effective choices. Constant cautions and independent direction expected to control the circumstances. By joining IoT and WOT, WSN (Wireless Sensor Networks) [19] served as an important and useful technology for farmers to effectively monitor and control various crop-related situations. Through the WOT, farmers were able to monitor changes in crop-related processes thanks to web services. exhibited a graphical representation, a real-time soil monitoring system, an SMS alert service, an irrigation management application, and an alerting setup for guessing the requirements. In order to make decisions more effectively in the future, this system can be improved further. A different study about farmers' knowledge requirements for crops' water quality revealed that their complicated farming practices were endangering the ecological space because of the low quality of their water. This could assist ranchers with improving their agrarian cycles as they knew about the water supply. Rather than giving attention to the ranchers, the water frameworks were created to target acquiring the water supply as indicated by the guidelines. For the purpose of providing farmers with real-time information, a user-centered 1622WQ web application was developed. There are some obstacles, such as a restricted net connection; 2) Unsatisfactory

data; 3) Within the program, technical difficulties discovered and addressed to ensure the provision of excellent assistance. Information about precipitation boundaries were additionally added. Using an open-source R package and the Shiny framework, the 1622WQ application prototype was developed in multiple versions. The interaction part has three Tabs: " Locn" displaying locations and allowing users for distinguishing; "Map" selecting located area; and "Data" displaying data position. By combining various technologies, the web-based application provided farmers with unified interacted place for a variety of initiatives for the delivery of water. The effort featured meaning of cooperation and combination of varied methodologies with advances and drove old techniques towards brilliant level [48]. A work about "Improving Crop and Farm Productivity," was conducted and inferred, output ought to be accomplished through exact ecological necessities. Be that as it may, the most common way of gathering the information was not productive because of their various old methods. For taking care of this issue, the development of IoT and different advances was utilized that incorporated the various IoT gadgets, remote sensors, and organizations, brilliant frameworks, observing cameras, and weather conditions stations. SmartFarmNet, a platform built on the Internet of Things, was shown created by incorporating the Semantic Web [26] Innovation that empowers the programmed assortment to information of necessary conditions, sifted through pointless information, accuracy towards improved yield execution. It could practically coordinate the different gadgets and sensors and put away the information to perform examination on them recommended the thoughts. The management of a large number of distinct networks and devices was the most difficult aspect of the SFN platform's development. The answer for this challenge was planning a typical Programming interface for every one of the coordinated gadgets and sensors for addressing their information through SWT. By combination of SWT, this stage had the option to perform constant examination on information, stretched out its reach to additional spaces, and could analyze the mistakes and execution of gadgets. SWT made it feasible to make use of semantic web standards. The real farming data used in the evaluation and experiments confirmed the platform's performance and ensured its scalability. The world's first infrastructure was this SmartFarmNet to support several systems and gadgets and provide services about data collection, storage, analysis, and crop performance forecasting. Another system presented advancements for upgraded highlights to help variety of gadgets, detecting devices, and continuous information delivery in agribusiness applications. The current systems were enhanced in terms of instant of processing data, analysis, and decision-making, but they only supported a limited number of areas and network grids. The suggested method maintained the reconciliation of a few stages with the assistance of semantic web innovations. It was able to interoperate heterogeneous devices, sensors, and networks and offered an accessible foundation on which pipeline processing applications could function, capable of interpreting massive amounts of facts and identify unidentified occurrences. On medium-to-large range farms, this framework performed well, according to evaluation. Additionally cleared the new way for the reconciliation of open norms and semantic web advances in future farming [50]. According to the findings of a study regarding WOT involvement [51], use of WOT requires improvements. This study gave an answer for transformation in the savvy climate using multi-reason options and reusable components. Relied upon the semantic advancements for getting data instantly, the arrangement was decided on account of the Brilliant Horticulture structure that utilized the ASAWoO standards. Also talked about how to make models from old sources of information. The ASAWoO project defined avatar, a component-based software for manipulating and handling. Because the WOT apps had a stepping-tool of operations, these avatars, which are dependent on the semantic design, could converse with one another and have unquestionable capabilities to integrate WOT applications for comparable objective achievement portrayed in a point-by-point structure in [52]. The WOT applications are now able to respond to questions about domain-independent adaptations thanks to the availability of a context adaptation process with multiple uses. The proposed method for achieving implementation and evaluation accuracy and performance was demonstrated using a Smart agriculture framework. The objectives regarding development of questions and answers were confirmed by outcomes. Additionally, the planned method was equated to methods for locating and overseeing relevant literature. Additionally, the WOT application established for gathering application management workflow requirements and contextual data. In the end, some perspectives on the work that should be done in the future for particular adjustments and transformations were established. An

earlier study [53] combined actual time Web technology into the current networked devices for farming in cities using a specialized Middleware architecture. Another layer was added by this paradigm in between the application and organization layers. The web framework continuously comprised of numerous principles and conventions. The fundamental capability of coordinating the online application was for collecting the information of each gadget through a one-of-a-kind separate identifier. This web innovation gave the capacity to information as well as introduced the information in graphical perception. Albeit this model was specific electronic proficient engineering there is as yet a need of improving it to store, examine and control the gigantic measure of information produced from now on. Combination of two technologies was described in a study [27]. As of late the Connected Open Information that coordinated the various types of information made it helpful for everybody on the planet in each field developed dramatically with heterogeneous organizations and sorts that prompted the advancement of information and reconciliation with the Internet as Overall Semantic Web (WWSW). The Assistance situated methods prepared for the advancement of semantic innovation to beat many difficulties connected with execution, proficiency, and accessibility of different administrations over the web world. Standards developed by the World Wide Web alliance made the it accessible to businesses, farmers, as well as financiers and businesspeople from all walks of life. The interconnected universe of individuals, gadgets, and administrations over the web produced a ton of continuous information in various structures, codes, designs, gigantic data, and cycles. Techniques for the semantic web [26] turned every single thing sensible as well as productive. Data from a variety of organizations was made available online for experiments and models, algorithms, and methods that were already in use. The semantic technology platform, developed by MIMOS, not only added new features but also enhanced existing ones. Joined Country's FAO fostered the biggest farming metaphysics on the planet as a multilingual horticultural jargon that gave numerous offices and assisted the associations with distributing their rural information and their models through AGROVOC. In more than 20 languages, AGROVOC had 40,000 ideas related to various fields like agriculture, fisheries, and forest, among others. It connected the other ontologies, such as chili, tomato, generic crop, corn, and so on. The four rural information models were distributed in MIMOS. The machines and frameworks can involve the two ontologies by involving connected information for looking through the assets. According to a study [54], semantic web technologies have been crucial in transforming unstructured data into usable form. As a supporting domain, SWTs were also used to tackle issues in the agricultural field. Semantic technology resources for the agricultural sector were developed by numerous large NGOs, including FAO. A review was deliberately finished to support future exploration on SWTs for issues in horticulture. This survey included a survey of current SWTs applications and a comprehensive evaluation on SWTs, procedures, and data interchange quality. The reviews of the articles for the survey came from conferences, books, and peer-reviewed journals. Processes in farming depended on crop types as well as on different factors like soil, water, climate, and natural circumstances. Semantic web technologies could be used to integrate data from time-based to time-varying means into a single standard structure or platform. This study looked at the agricultural semantic resources, which included taxonomies, controlled vocabularies, thesauri, and ontologies [55] with a lot of agricultural sub-domains. Two ways to deal with making the focused semantic assets for horticulture were depicted. One was to establish a brand-new one [56], and alternative included grouping the resources that were already in place [57]. The study also looked at and talked about agricultural-specific applications of semantic web technology. As there was a gigantic measure of SWTs assets that were determined for the farming space yet not many of them were utilized to tackle the issues. There were less works in the literature review that were related to semantic web technologies (SWTs) in agriculture, it’s to be a significant option for agricultural issues. Even though there was not a lot of research on this technology in the literature, it may open the door for future research on how SWTs can be used in agriculture. A research project was carried out that examined the Web of Things work in the agricultural sector [58]. The analysis was carried out in three stages: 1. wants and gauges of expert's area, 2. Production of agricultural machinery's movements, Status of existed achievements in the area. Agrarian field shifting towards modern one, thus analysts predicted a upcoming change in farming. Agriculture faced two primary obstacles in achieving the key solutions: changes in the weather and the environment's long-term safety. Agriculture benefited from IOT's fluctuating frequency. Through WOT, the data from a variety of sources

as well as IoT devices were merged to find solutions to various agricultural issues and make decisions for improved production. The fundamental components of smart agriculture and precision farming are WOT and IOT. Numerous US associations made the accomplishments in the horticulture field like AgJunction Inc., Monsanto and DuPont, Deere and Company, Raven Enterprises, and Trimble Route Ltd. In Europe, Future Web Program development in agro-industry gave 14 million euros to the development in this field [89]. Robot innovation likewise played out an immense job by presenting sensors, crop machineries as well as checking and handling gadgets that involved various innovations. The smart farms of the future are the result of all emerging technologies, protocols, and smart systems. This study inferred that the WOT is the principal character for a shrewd cultivating future, so the ranchers need to concoct game plans for their high efficiency and benefit as indicated by these innovations' thoughts. To furnish the ranchers with significant data for better direction, IT assumed a crucial part in the horticultural field. An exploration "Semantic Electronic Coordinated Farming Data System" was directed that proposed an online information model for social occasion, mining, coordinating the information from different places [59]. Web technologies were utilized in the integrated Agriculture Information Framework (IAIF) model for linking data as well as metadata for gathering information. A technique was likewise given to metaphysics to get the necessary information from various sources. IAIF likewise consolidated a dynamic module named PC based Conversation Emotionally supportive networks (CBDSS) to help the ranchers in better choices thoughts. The IAIF cosmology included three sections: Domain, Link, and Resource sections provide crops related details for great output. The XML parser and D2R systems were utilized to access data. The link subsection of cosmology connected the data sets and the Resource segment addressed the connecting sources on the WWW. This was all conceivable by web advances to associate the data sets for data and information gathering. It served as an extension of the previous investigation. A conversation was likewise held to connect the few different information assets for upgrade in this work. Thoughts about wireless involvement and device linking was fundamentally altered as a result of the involvement of the Web of Things. People taken advantage by combining WOT to food trade because they understood its value. Research presented the WOT coordination for information exchange as well as boosting production [60]. The Web of Things Supply Chain was utilized by numerous nations for the quality and safety of agricultural food items. The facility for sharing information about agricultural products in real time was made possible by the integration technology of RFID. The Internet of Things has benefited a lot to systems, devices, architectures, and applications in recent years, but it hasn't been able to work together because of different protocols, networks, and descriptions of objects. Regarding these integrations of IoT and WOT technologies, a study was carried out [21]. Numerous businesses attempted, but were unable, to develop standards to address issues arising from various platforms, networks, and objects. The investigation examined various IoT applications, projects, and services and offered an approach to handle these problems. There were three stages, each serving a distinct purpose. It had a server Web of Virtual Things installable between the cloud and Fog Layer of IoT, the middle layer, to solve compatibility issues. It likewise gave a point of interaction REST to gadgets to consolidate at the most minimal insight. There is no need for the other intelligent devices to wait for requests or to respond to questions. In a non-realistic device, it can express the received messages consistently in realistic way. Once the device reaches its real active state, it responds to queries and keeps the user informed of any responses. The examination and valuation task were carried out in order to evaluate the capacity, safety measures, and power of working. The outcomes proved WOT and IoT joining extremely beneficial. The role that WOT plays in agriculture was the subject of another study [61]. The combination of IoT improved the horticulture arena. As a result, enhancement also resulted in obstacles to and disclosure of agricultural data during data collection. This study presented a framework called "ASAWoO" to address this issue. It made it possible to control devices using WOT recommendations and computational methods to obtain the interconnections between different sensors. WOT aided gadgets to convey and connect with one another. To determine whether the proposed framework was effective, a four-month farm experiment was carried out. The emerging Internet of Things (IoT) and WOT provided outcomes of this deployment, and they have the potential to greatly benefit agriculture. In order to easily and freely obtain agricultural data, research was conducted [62]. As of late the efficient information opened up online with the expectation of complimentary use. The semantic web advanced development made the information accessible for public

use with liberated from cost office. Many nations' states involved these advances and gave free information to public. The notable states UK and USA likewise used the recent fads and gave helpful free information in many fields like in wellbeing, instructive and farming areas for boosting the economy. In addition to making the data freely accessible, combining web techniques with Linked Open Data made it possible to connect various sources and easily collect data. Yet, this information more often than not was not in the outfitted structure that made it difficult to involve it for examination and continuing purposes and furthermore made it hard to associate with different data sets. This study suggested using a Danish-language web framework called "RDF" for the Danish government to collect, process, and connect with various databases to address this issue. Also made it possible to connect agricultural data to organizations, making it easier for them to find answers to difficult questions. After acquiring the data, this model procedure analyzes it, cleans it, and connects it to various sources. Eventually, it was decided via testing and getting great outcomes. A report regarding handling of historical items was released [63]. In most recent years, organizations, public establishments, and neighborhood bunches have devoted creating interest to the character of good strategies for survival of old resources. Plans for controlling cultural property are entirely based on a lot of different standards. A strategy for selecting multiple standards was presented to rate beneficiation procedures with the goal of increasing cultural and financial well-being. A solitary programming of the A'WOT for assessment in helping the arrangement and work of control strategies of deserted foundation properties of culture. The AHP and SWOT techniques in this application helped with decision-making, resource management, problem-solving, and tourism-related resource management. The integration of all methods and the proper step-by-step arrangement of processing stages are necessary for this method to produce effective results. In the past, many concepts were proposed for constructing a distant study environment with the option of practical teaching for students in four subject areas such as science, engineering, mathematics, and technology. "Contribution to the Setting Up of a Remote Practical Work Platform for STEM:" was the subject of research. The Instance of Agribusiness" expressed that the overviews of regular and life sciences concentrate on regions showed that the subjects given to the understudies were not totally well-educates [64]. A pragmatic part ought to be added for better concentrate however the visit to the organic unique locale could be risky for people. By combining the capabilities of WOT and WebRTC, a multimedia server technology, this study also contributed to the advancement of distance learning by providing a virtual location for visiting and sharing various facilities. The work was mostly done in the agricultural field, but the results of the experiments could also apply to other fields. This introduced answer gives the office by which a few students and instructors could go to the field trips while different students could be worked with by a live transmission of that excursion and get the experience. The ease of use of WOT's APIs and standards, which made it possible for the various objects to communicate with one another through web language, was the main advantage. Rules from small level servers could also be used in devices. The Internet of Things (IoT), API, and Web Application Interface stages make up this framework. The simple web browser could be used to enable communication between multiple users and obtain data from devices that have already been deployed by implying this framework. Agricultural distance learning could be improved and implemented using the proposed framework. The internet, communication networks, and devices have all improved over the past few years, transforming the world into a global village and transforming everyday objects into smart objects into what is now known as the Internet of Things. A review "Brilliant Sensors from Ground to Cloud and Web Insight" was finished on Implanted Knowledge (EI) a coordinating examination field that deals with revealing the activities of everyone, things, complex shapes and furthermore figure out the secret characteristics of savvy gadgets [65]. Utilizing open-source agricultural data, the data extraction process resulted in the creation of useful ontologies and information resources for effective decision-making. The previously used agricultural data was thoughtfully analyzed in order to produce the new material for enhancement purposes. The past exploration record, properties, applications, structures, and EI research-related holes were checked on in this review and showed a clever water supply application. The primary objective was to develop intelligent management systems for agricultural processes and to simplify the threats and new directions of Intelligent Building Technologies. The semantic web advancements were utilized for the improved looking through process, distributing the information online in efficient structure and associating the gadgets for correspondence and

working. The astute rural information ought to be broke down by researcher's groups to really take a look at the unwavering quality to send in genuine ranches and horticultural regions by the nations for powerful creation. For safe data availability, international trusted standards are also required. despite the fact that numerous studies suggested that WOT and other technologies would emerge and find applications in the agricultural sector to address interoperability, performance, production, and management issues. There is very little research on it, and no study provided a taxonomy to provide a clear explanation or opportunities for future research in this area.

In any case, this work introduced a deliberate writing survey of ongoing pertinent examinations that were led to give the answers for the horticultural issues. The curiosity of this study is that we have introduced a scientific categorization of WOT applications to the farming space and a wot based horticulture model. Last but not least, open issues and directions for further work discussed.

3. Research Methodology

A systematic literature review, which will be the suggested methodology, will be carried out to gather, examine, and classify all current and planned research as well as the protocols, standards, and developing techniques [66].

Detailed procedure has shown in the Figure 7 below.

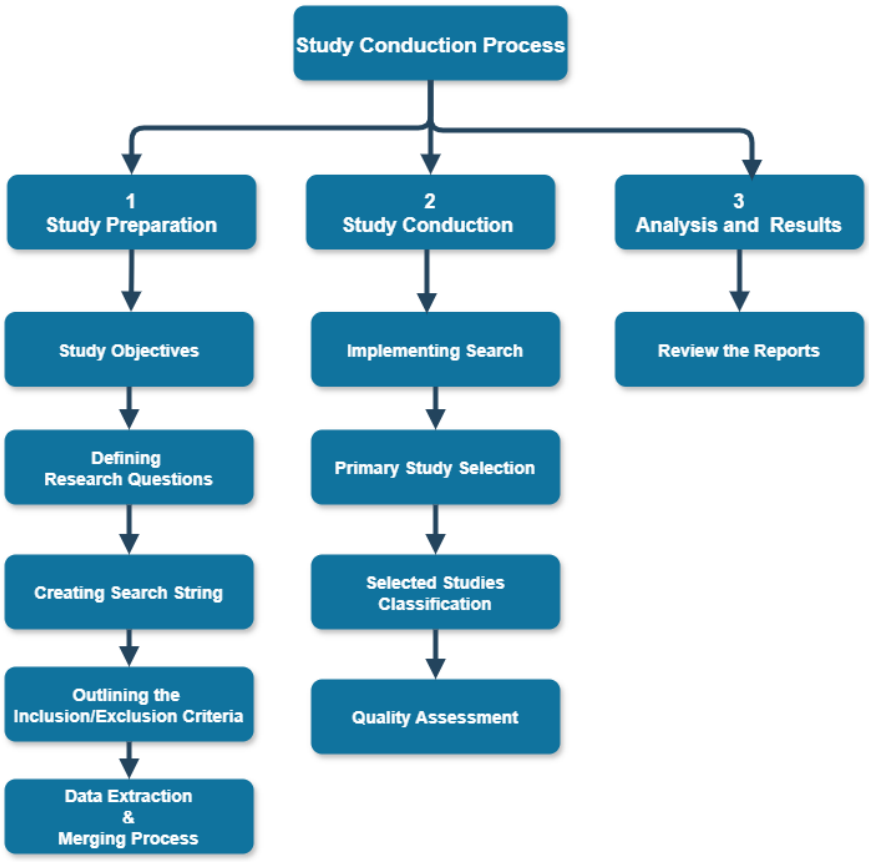


Figure 7: Systematic Literature Review Process

Three steps structure the process: preparation of the study, conducting the study, analysis, and results. The chosen papers will be categorized based on the application domains, research types, and research approaches. We will extract as well as classify studies based on their methodologies. Researchers will

discuss this SLR and provide recommendations for future directions [67]. The methodology that is being used has been suggested by [66].

3.1. Research Objectives

The main objectives related to the selected study are following:

- The purpose of this paper is to give an outline to the ongoing research on WOT technologies concerning the Agricultural industry.
- To acknowledge and reveal current and future work of WOT in Agriculture applications and shortcomings of work done so far.
- To provide an overview of movements in the research discipline.
- To map out publications in the related field of the study.

3.2. Research Questions

To have the overall view of the area of the interest identified with their corresponding motivations based on the selected study. The answers to the given questions will assist in giving an idea of the existing research trends, limitations in WOT Agricultural research and its potential future directions.

All the given questions and the motivation that will lead to the provision of answers to the given questions has been presented in form of table 1 as follows. For research questions, refer to Table 1.

Table 1: Research Questions

No.	Research Question	Motivation
RQ1	What are the primary target publication channels for WOT agricultural research?	Identify reputable sources for WOT agricultural research and future study.
RQ2	How has approaches occurrence for WOT agricultural articles changed throughout time?	To categorize publications in WOT-Agriculture Research as they change over time.
RQ3	How many kinds of WOT based agricultural studies are present?	To gain insights about the research approaches regarding WOT in the study.
RQ4	What are existing researches along with gaps in WOT agriculture research?	Understanding current researches might inform future research strategies and identify unresolved questions in agricultural research.
RQ5	What alternative approaches were offered to solve challenges in WOT agricultural research?	To identify the existing approaches offered in the WOT agriculture study.

3.3. Search Scheme

A searching string has been employed in order to gather relevant work in the given research area. Figure 8 below presents the strings combination.

For the purpose of the search, we used the relevant keywords relating to research area. In light of the inquiry, many databases were used to find relevant material for this investigation. In order to get the bibliometric studies, Google Scholar was used.

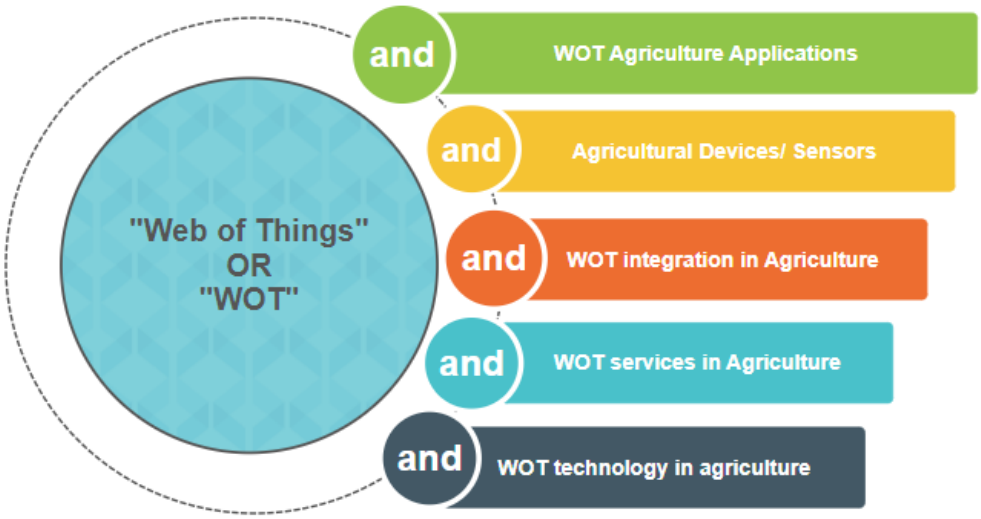


Figure 8: Search String

The search scheme is presented in the Figure 9 below.

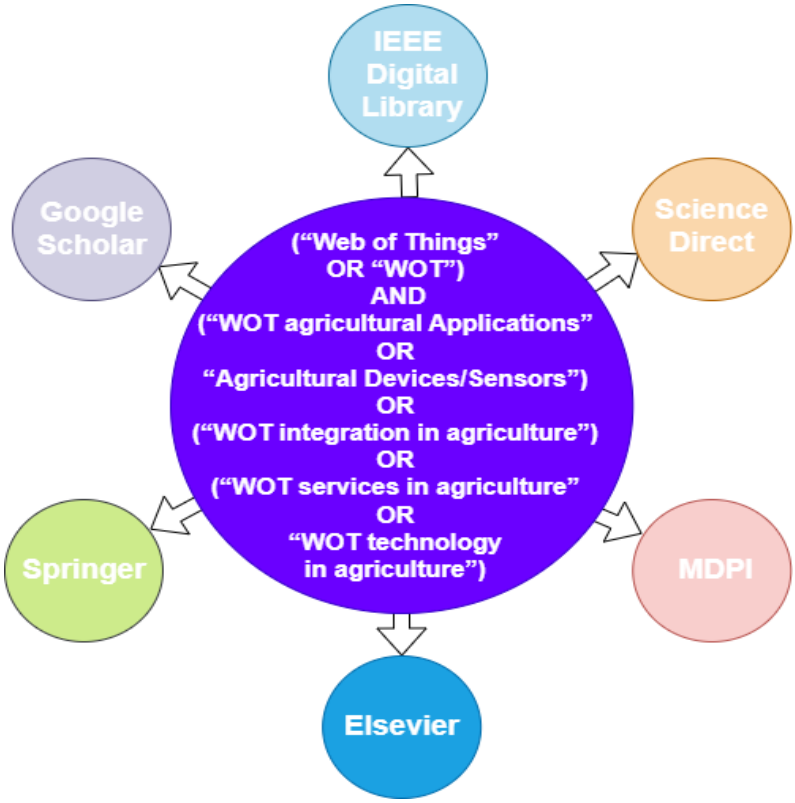


Figure 9: Searching Arrangement with Databases

3.4. Study Selection Process

In the selection criterion, an attempt was made to look for the most related work. Work published in the two different sources were counted only once the order in which the research was done. Whole material has been reviewed according to keywords. Secondly, we excluded identical and nonidentical topics. Thus, works were selected based on described conditions.

For criteria, refer to Table 2.

Table 2: Inclusion/Exclusion Criterion

Inclusion Criteria	Exclusion Criteria
IC-1 Articles presenting concepts and integration of WOT	EC-1 Articles that are not focused on WOT applications
IC-2 Articles that are focused on WOT applications and their implementation	EC-2 Articles not presenting new and emerging ideas
IC-3 Articles presenting WOT problems/goals	EC-3 Articles presenting general focus on WOT application integration-based model
IC-4 Articles presenting WOT standards/protocols/tools	EC-4 Articles not related to the search string
IC-5 Studies published in English Language	EC-5 Articles that are published before 2010

Outcomes of procedure has been explained through a diagram. For study selection process, refer to Figure 10.

3.5. Quality Assessment

One of the decisive steps in assessing goal is to raise the standard of elected papers. Commonly, the Quality Assessment (QA) is performed in Systematic Mapping Study and Systematic Literature Review. Using an SLR, an assessment was created to get level of related works that were chosen for this investigation [68].

1. Paper strengthens the agricultural sector and WOT. Potential answers: Fully (+1), Partially (+0.5), and Not (+0).
2. Paper proved an identified concern within the agricultural field through WOT: Fully (+1), Partially (+0.5), and Not (+0).
3. The study's shortcomings and potential for more research are examined: Fully (+1), Partially (+0.5), and Not (+0).
4. Study was published in a well-known platform. Below is a potential response to this query regarding Conferences and Journals Rankings.

For Conferences:

- If CORE A (+1.5),
- If CORE B (+1),
- If CORE C (+0.5),
- If not ranked CORE (+0)

For Journals:

- If Q1 (+2)
- If Q2 (+1.5)
- If Q3 or Q4 (+1)
- If not ranked (+0)

At the end, the overall score of the study is computed from the individual scores of each question which ranges from 0-5.

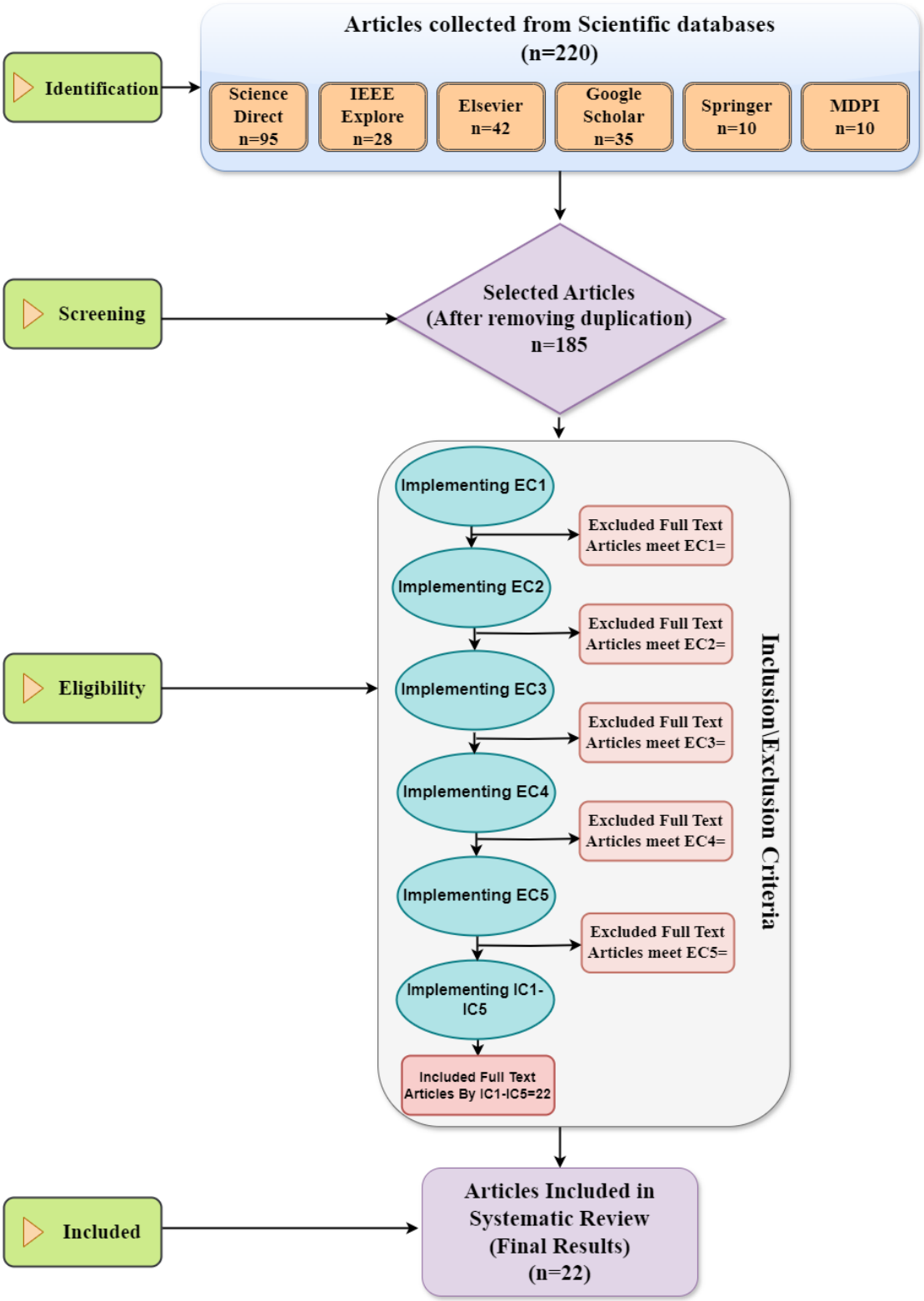


Figure 10: Study Selection Process

3.6. Data extraction and merging process

- RQ1. For answering this question (RQ), it is required to define a publishing channel.

- RQ2. Articles should be grouped by the year of their publication.
- RQ3. The categories that can be used to identify a study type [69]:

Solution proposal: Solutions to agricultural issues are provided including problem solving methods. it can be genuinely new or may be a development of an existing well-known strategy. The effectiveness and importance of the solution is illustrated through justification or by giving few examples only.

Conceptual Research: They were therefore able to simplify these concepts by seeing and examining the data that is currently available on the WOT apps There are no used.

Evaluation Research: Investigation and assessment are performed for the WOT-agriculture mechanisms as well as concerns recognizing problems in WOT-Agriculture applications.

Others: which include surveys, system, architecture, Development, experimental, reviews, performance analysis, and models.

- RQ4. Determine the current state of Web of Things (WOT) in agriculture industry is the study's main research question. Perhaps, one can possibly make sense by gathering concerned scholarly papers from different peer reviewed journal, identifying the research void, and is able to depict the trends of the research. This proposed SLR will help both new scholars and specialists to develop existing effort on WOT in agricultural resolutions
- RQ5. According to [69], a strategy can be divided into following groups:

System: It might be a suggestion that could offer solutions to monitor the agriculture, controlling and decision making using WOT.

Framework: There are particular theoretical or conceptual model developed for promulgating or achieving the interoperability and effectiveness of IoT systems in WOT.

Application: An idea that supports the function to continuously observe the agricultural processes, to administrate it and to informs the farmers or users about what is happening by utilizing web application.

Method: A structure proposed and a plan for the WOT services as a means to develop agricultural knowledge with step-by-step procedure.

Infrastructure: A managing setup must exist to track of farming via WOT services.

Architecture: Idea concerning the techniques of the agricultural sector.

Guideline: An example of a practice that can be used to find solutions for agriculture by WOT technologies is also a discussion.

Others: Platforms and analysis etc.

It is primarily involved with the ultimate studies that are grouped for each query; it recalls the important works along with ranking provides a graphic representation of the outcome of grouping.

4. Results

It gives conclusion of conducted study questions by elaborating on them. Out of them, the studies were selected by evaluation procedure and were sum up into contribution of WOT in Agriculture area.

4.1. Selected Studies Results

Initially all 220 identified research studies have been evaluated based on their titles, keywords and abstracts which resulted in the exclusion of 198 articles and inclusion of 22 articles only. The 22 papers were scrutinized effectively to respond to the research questions of this research investigation.

4.1.1. RQ1. What are the primary target publication channels for WOT agricultural research?

In the context of this Systematic Literature Review, the Publication channels are only those journals and conferences. Concerning the papers, all were published in the 12 journals (0. 54%) and in 10 conferences (0. 45%). Additionally, all studies utilized diverse publication areas.

For the publication channels, please refer to Figure 11 below.

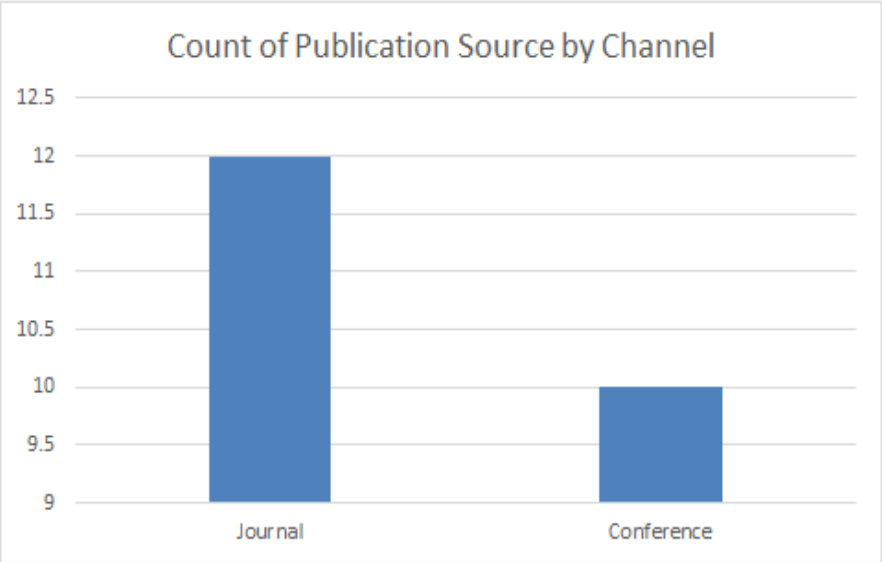


Figure 11: Publication Channels

4.1.2. RQ2. How has approaches occurrence for WOT agricultural articles changed throughout time?

In consideration of the papers selected for this study, the papers of focus were published in the period of 2010 to 2020. Over years different types of publications have exhibited the information in the figure. The figure below shows annual developments of WOT integration in agriculture. With the advent of IoT and other technologies, 2016 saw the greatest number of papers published, marking the beginning of the WOT era in all fields, including agriculture. Second busiest years are 2018 and 2019 where research was conceptual, proposals, and contributions to integrating WOT with internet of things systems and architectures.

For types of publications, please refer to Figure 12 below.

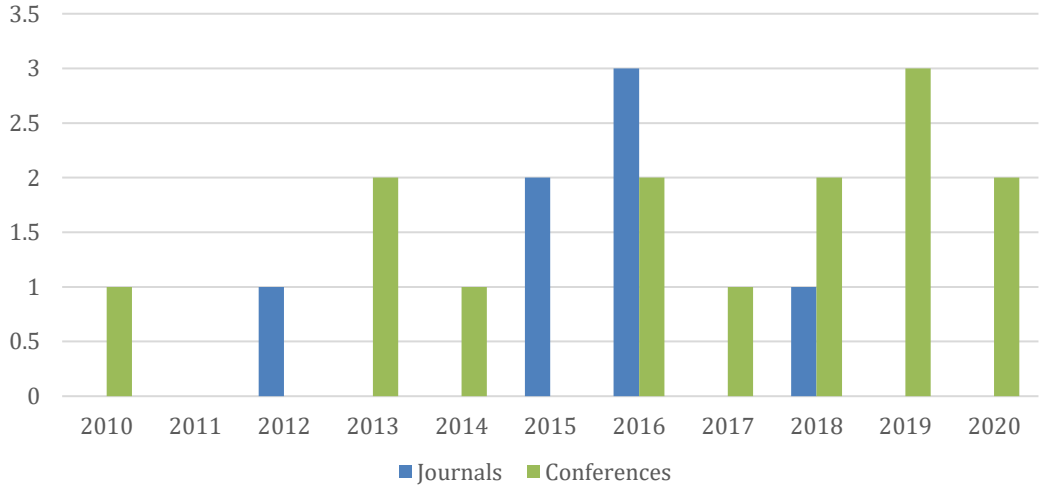


Figure 12: Publications over Years

The graph indicates that the mostly articles appeared in journals in 2011, with no paper being published during that year. In the field of agriculture, WOT was integrated in numerous ways, many of which involved web applications, systems, and frameworks. It is unlikely that this study, which was conducted in 2020,

will reveal the precise number of studies that were published in that year.

For trends in the studies by years, please refer to Figure 13 below.

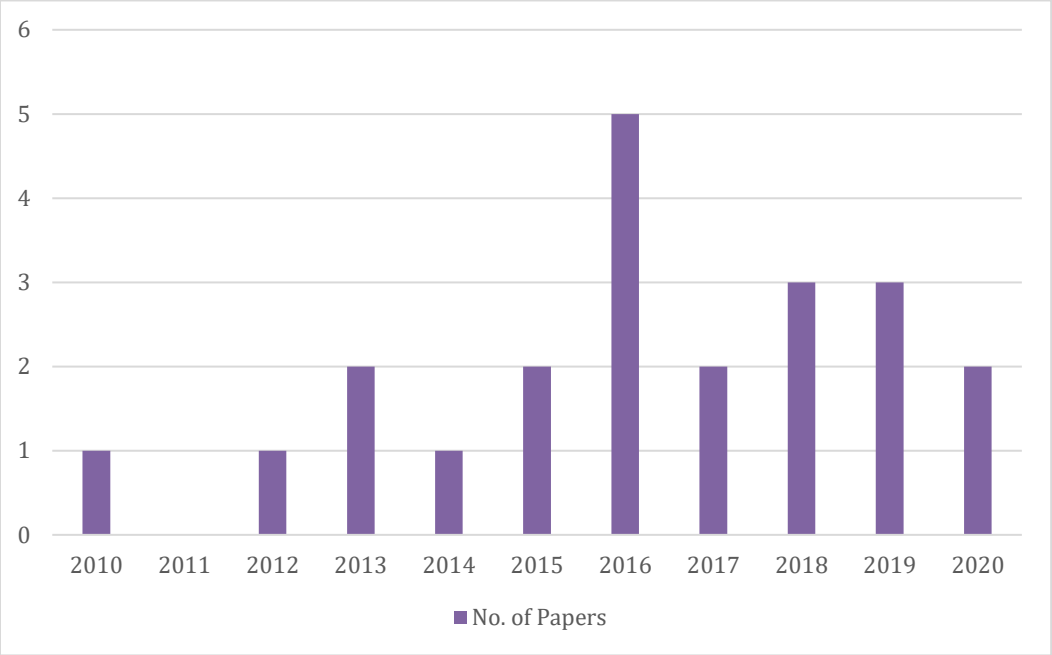


Figure 13: Trends in studies by year

4.1.3. RQ3. How many kinds of WOT based agricultural studies are present?

For presented methodical investigation, selected 22 articles were divided into nine categories of research, which included: Solution proposal (6 articles) (0.27%), Experimental research (3 articles) (0.14%), Proposed System (3 studies) (0.14%), Proposed Model (2 studies) (0.09%), Contributional Research (3 studies) (0.14%), Conceptual Research (2 studies) (0.09%) whereas only one study included survey and evaluation research, implementation and evaluation research. Figure 14 displays every type, and Figure 15 displays the outcomes of these types. The provided graphical representation demonstrates that the chosen studies (solution proposals) primarily address agricultural problems. Additionally, some experimental, practical, and assessment WOT-Agricultural solutions were presented. The author of this study [20] suggested using Service Oriented Architecture (SOA) to combine new and old infrastructure features in order to address the issue of sensor interoperability and integration into systems for monitoring and sharing. Webservices for sending, handling, and exchanging agricultural data over the Internet were included in new infrastructure, along with the concept of the sensor web. The author of this study [46] identified the challenges in obtaining and keeping track of real-time data and suggested a web enabled setup that makes MQTT protocol to track and monitor devices as well as analyze, store, and provide users with access to real-time data. An author looked at the issues with the water supply systems in terms of farmers' awareness and offered a single web enabled system where farmers could obtain up-to-date data on water quality [48]. The problems with the internet connection and data quality were also fixed by this application.

For research types, please refer to Figure 14 below.

Three tabs made up the application: the Locn tab displayed locations, the Map tab allowed users to choose their desired location, and the Data tab displayed location data. The author of this article [50] examined the problems with integrating heterogeneous platforms and devices and offered a framework for doing so by utilizing Semantic Web technologies. This framework would allow combination of multiple areas, devices, and set ups for streaming and examining actual data, identifying faults as well as coordinating various platforms and devices.

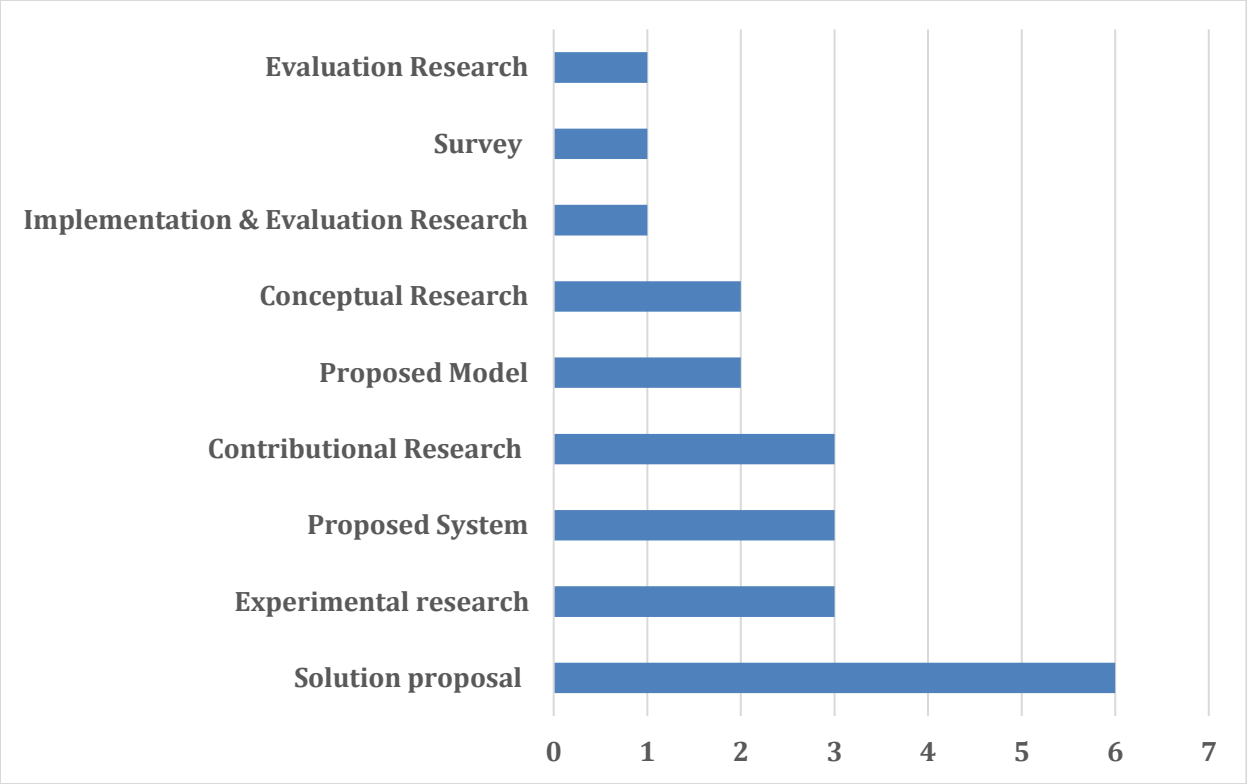


Figure 14: Research Types

For results of all research types, please refer to Figure 15 below.

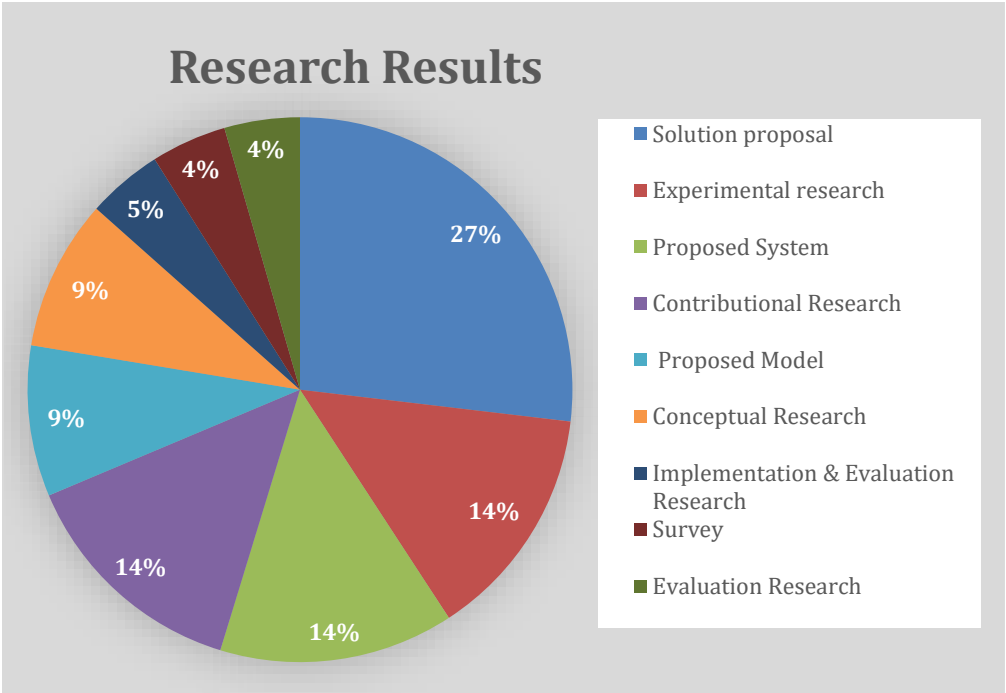


Figure 15: Research Types Results

4.1.4. RQ4. What are existing researches along with gaps in WOT agriculture research?

As displayed in the outcomes it is presumed that the vast majority of the examinations depend on the checking and the executive frameworks, Systems, and web applications for Rural area by coordinating WOT into few different innovations. Some of the cases discussed here include: In a work [53], the creator introduced a middleware design by coordinating Internet innovation in current IoT frameworks for saving information of each gadget by one-of-a-kind identifier independently and introduced the information in a graphical view. This study [45] introduced an irritation determining and data framework for indicated district bug expectation. Using WWW technologies, this software with a web interface offered functions for storing and mining data. The WebGrabber, which was used to obtain data and carry out the real-time calculation, was included in this system. Then, at that point, this framework improved the gauging and dynamic framework by making a calculation. This system was reusable, cost-effective, and virtually extensible. A technique of observation to help farmers manage and keep an eye on various crop-related states was suggested in a paper [47]. This framework created by coordinating the Remote sensors organizations, Web of Things advancements and highlights a ready framework to help constant observing of soil, water system conditions, graphical view, and dynamic in light of the yield circumstances. A review [63] presented in regards to protect verifiable culture resources. It can handle decision-making, resource management, problem detection, and problem-solving. Several managements, monitoring, control, and decision-making-related WOT-based solutions are presented. The vast majority of them coordinated the a few advances for productive execution and results yet some of them actually need accomplishing the objectives, for example, dealing with gigantic information sum, multi-space gadgets, organizations. The majority of works combine IOT involved agricultural systems and applications with WOT. To get over the shortcomings and restrictions of IoT-agriculture research, WOT technologies have arisen to provide immediate processing and agricultural activities. WOT services made it possible to collect, analyze, and make decisions based on real-time data. They also made it possible to control and manage the diverse infrastructures, sensors, and devices as well as crops. Yet, there are still a few holes that were acknowledged in this precise writing survey. The WOT made standards and infrastructures better, but secure international standards are still needed. Prior to applying the available agricultural data to real farms, it is important to examine it. A significant hole has been found in regards to instructive and research region in horticultural space that it needs far off functional stages for instructive purposes and just 1 review was tracked down on this subject. The fact that less research has been done in the WOT-Agricultural domain research area on SLR, SR, LR, and mapping studies is another gap. Along these lines, it needs further work.

The selection of articles is included in the Table along with specifics about their categorized outcomes, please refer to Table 4.

Table 4: Articles Classification

Ref.	**Quality Assessment**									
	P. year	**P. Channel**	**Research Types**	**Research Approaches**	**Application Domains**	**(1)**	**(2)**	**(3)**	**(4)**	**T. Score**
[20]	2015	Journal	Experimental Research	Infrastructure	Monitoring	1	1	1	2	5
[21]	2019	Journal	Conceptual Research	Framework	Interoperability	1	1	1	2	5
[27]	2012	Journal	Evaluation Research	Architecture	Analysis	1	1	0.5	2	4.5

[41]	2013	Conference	Solution proposal	Management system	Monitoring And Management	1	1	1	0	3
[42]	2015	Journal	Solution Proposal	Database System	Storing and Analysis	0.5	0.5	1	1	3
[45]	2013	Conference	Proposed System	Decision & Forecasting System	Storing and Analysis	1	1	1	0	3
[46]	2016	Conference	Proposed Model	Web Application	Monitoring	1	1	0.5	1.5	4
[47]	2017	Conference	Proposed Model	Monitoring System	Monitoring, Controlling and Decision making	1	1	1	0	3
[48]	2016	Journal	Proposed Solution	Web Application	Monitoring	1	1	0.5	2	4.5
[49]	2016	Journal	Implementation and Evaluation Research	Platform	Analysis	1	1	0.5	2	4.5
[50]	2016	Conference	Solution Proposal	Framework	Analysis	0.5	0.5	1	0	2
[51]	2017	Conference	Experimental Research	Platform	Storing and Analysis	1	1	1	0.5	3.5
[53]	2018	Conference	Proposed System	Architecture	Storing and Analysis	1	1	1	0	3
[54]	2019	Journal	Survey	Guideline	Analysis	1	0.5	1	2	4.5
[58]	2016	Journal	Contributional Research	Existing Aspects Analysis	Analysis	1	1	1	1	4
[59]	2010	Conference	Solution Proposal	Framework	Integration	1	1	1	0	3
[60]	2020	Journal	Conceptual Research	Guideline	Optimization	1	1	1	0	3
[61]	2019	Journal	Experimental Research	Framework	Monitoring and controlling	1	1	1	1.5	4.5
[62]	2014	Conference	Solution Proposal	Method	Publishing	1	1	0.5	0	2.5

[63]	2020	Journal	Solution Proposal	Application	Management	1		1	1.5	4.5
[64]	2018	Conference	Contributional Research	Platform	Education	1	1	1	0.5	3.5
[65]	2018	Journal	Contributional Research	Guideline	Analysis	1	1	1	1.5	4.5

4.1.5. RQ5. What alternative approaches were offered to solve challenges in WOT agricultural research?

In carefully chosen studies mostly are offered systems (4) (0.2%), Applications (4) (0.2%), Framework (4) (0.2%), guideline (3) (0.14%) and Platform (3) (0.14%). Whereas the remaining offered Architecture (2) (0.09%), Analysis (1) (0.045%), Method (1) (0.045%) and infrastructure (1) (0.045%).

These all has shown in the Figure 16 below.

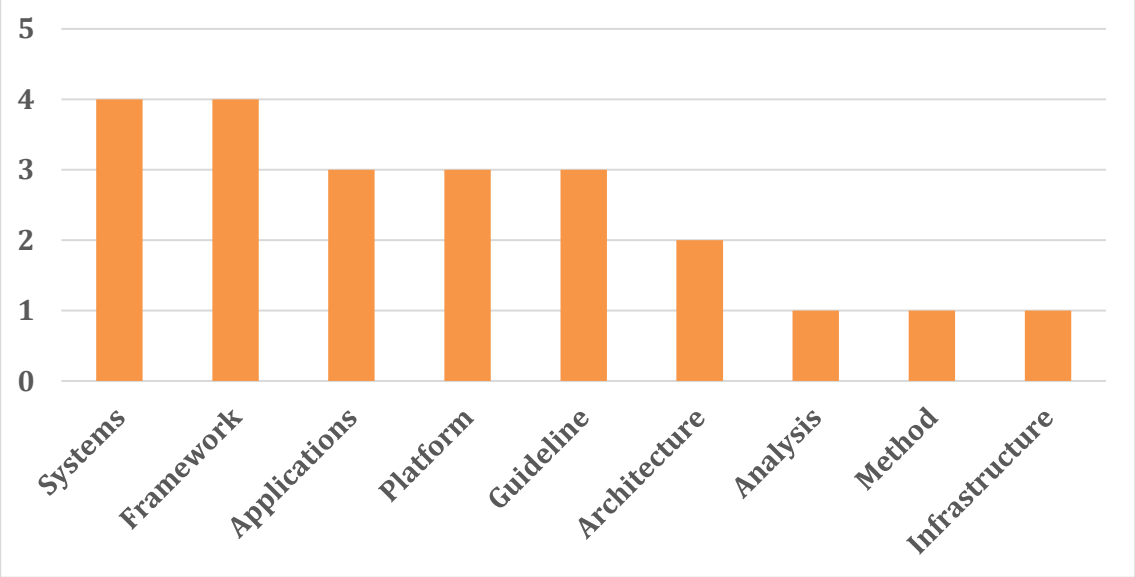


Figure 16: Research Approaches

All of the strategies were aimed at enhancing agricultural practices, equipment, and product attributes for successful farming. Some methods improved the previously offered solutions by incorporating the wot to transform traditional agriculture into smart and precision agriculture and make them affordable, dependable, and efficient agricultural applications. Also included is a table that summarizes the methods used in the chosen studies.

For the Overview of Research Approaches, please refer to the Table 5

Table 5: Detailed summary

Ref no.	Approach
[20]	In order to overcome the difficulty of interoperability, an open, geospatial web service-oriented physical infrastructure was suggested, which would be integrated with the PA monitoring systems and sensors. It gathers, organizes, and disseminates data from devices. Ten distinct sensors were placed in Wuhan, China to estimate the functionality of

	infrastructure. When the suggested infrastructure was compared to the current one, it was evident how well it could monitor processes.
[21]	Presented the integrated model with WOT into the supply chain workflow to achieve knowledge sharing and establish an effective supply chain across the networking using RFID to boost the production of agriculture. The presented research examined numerous IoT infrastructures, company projects, and IoT applications and proposed a conceptual model for integrating IoT and WOT techniques. A Web of Virtual Things (WoVT) set up was used as a Fog Layer to overcome collaboration issues, as well as a REST interface for integrating devices at the lower perceptual layer. The simulation procedure demonstrated that the combining the WOT and IoT can be quite valuable and productive.
[27]	This concept discussed the combination of Linked Open Data and Worldwide Semantic Web Technology. Multiple forms and data types resulted in data transformation and convergence of WWW under the name World Wide Semantic Web (WWSW). Methodologies opened emergence of semantic technology, which helped address numerous issues connected to the performance, efficiency, and access of multiple services on the internet.
[41]	By combining IOT and WebGIS, precision agriculture management system (PAMS) was created and implemented on the chosen farm. WebGIS processed the network geographical data efficiently, and IoT was used to improve perception accuracy. The platform for managing, spatial information infrastructure, the mobile client, and IOT infrastructure were its four stages. The data was collected, transferred, and published using Internet of Things (IoT), WebGIS and Location-Based Service (GPS), and RS methods for managing and monitoring the production.
[42]	In order to overcome the current obstacles, it investigated Agricultural Knowledge-Based Systems that employed Semantic Technologies. It was employed by scholars and people to investigate underutilized crops, offer fresh perspectives and ideas for future research, and make knowledge useful. Numerous ontologies were created to aid in the creation of semantic products. A lot of knowledge-based and expert systems, as well as web-based systems, employ "if-then" queries to get the answers to questions.
[45]	In this work, a software was utilized the World Wide Web technological advances for weather data mining and storage. The files with .txt extension were downloaded and the network stations of the NOA were connected. It was used for Actual time pest forecasting. It consists of two stages: first collects weather data by customized WebGrabber; the second stage calculates the population in real time and stores data on a local MySQL server. An algorithm was introduced for forecasting and decision-making. It was inexpensive, simple to use with an almost infinitely expandable architecture.
[46]	It was suggested to use a web-based configuration to track, observe, and evaluate the agricultural data in real time. The systems saved and shared the sensor data over the Web. Message Queue Telemetry Transport (MQTT)" protocol was used to transfer data and monitor gadgets. It had the characteristics of asynchronous communication, lower complexity, and lower power consumption.
[47]	A set up was developed for the purpose of monitoring and decision. WSNs, IoT, and WOT were employed to enhance the method of supervising and controlling crops conditions. A device by which water details can be estimated, an efficient application to handle the entire irrigation process, to identify the possible graphical view of the field and an option to monitor the real time soil conditions.

[48]	The need of water in crops was explained to the farmers and then the A web application was developed for delivering the immediate details to the agriculturalists. Such factors as restricted internet access, poor data quality, and some general problems of the organization's functioning were detected and addressed. It offers the farmers an integrated data and communication solution. It also emphasized on the need and importance of synergy and/ or ensemble of heterogenous framework and tools for Smart Agriculture.
[49]	Studies were made concerning the increase in productivity of crops and farms through determination of the precise, required, or optimum weather and environmental conditions. Thus, appearance of IoT and other technologies like, IoT devices, wireless sensors, Mobiles, Smart systems, Monitoring cameras and Weather stations were included. The world's initial platform, Smart-FarmNet globally was developed using Semantic Web Technology to auto data gathering and virtually had integrated all the various devices and sensors, and the collected data was stored so as to perform various analyses. A base API to control and incorporate massive quantities of different devices and networks to exemplify their data with the help of SWT was developed.
[50]	A framework named Agri-IoT was created with the adoption of semantic web technologies to provide expanded functionality to support different devices, detectors, and actual data transmitting for evaluation purpose. It provides one computing unit with pipeline processing enabling applications to evaluate large amounts of data, identify unknown occurrences, interconnect heterogeneous devices, sensors, and networks. This framework performed exceptionally well on small to large farms.
[51]	According to the approach, WOT applications necessitated advancements in solutions for adapting them to common models for varied aims. A multi-functional solution with reusable settings was offered. A Smart agricultural framework highlights a suggested approach for obtaining accuracy and performance in both execution and evaluation. The results confirmed the desired outcomes.
[53]	A specific Middleware design was proposed for urban agriculture, integrating real-time Web technology into current IoT devices. The major role was to save the data of each detector unit with a distinct, independent identity through an online application with a visual image display.
[54]	A study was conducted with the purpose of encouraging further research on SWT applications for agricultural problem solutions, as SWTs played an important role in tackling issues in the agricultural region as a helping domain. It provides comprehensive analysis of existing SWT resources, methodologies, data trading standards. Two techniques for developing centralized semantic facilities to support agriculture were discussed. The study also explored and considered semantic technological web applications for agriculture.
[58]	This investigation examined the impacts of the Web of Things towards the agricultural industry. The analysis was conducted in three steps. 1. expectations and assumptions for area, 2. Actions of machine manufacturers, 3. Position of existing progress of sector. It indicated that WOT and IOT are the core elements for farming's high profitability and productivity.
[59]	Semantic Web-based knowledge paradigm presented for collecting, analyzing, and bringing together data from several sources for agricultural awareness and decision-making. In order to attain the concerned data and metadata for the extraction of information, web technologies were used. This work serves as an enhancement to the existing effort of giving farmers with knowledge to help them make better decisions about large-scale productivity using less resources and machinery to meet future needs.

[60]	It Presented the integrated model with WOT into the supply chain workflow to achieve knowledge sharing and establish an effective supply chain across the networking using RFID to boost the production of agriculture.
[61]	This highlighted WOT's contribution to agriculture area. Incorporation of IoT to equipment for environmental observance improved agriculture by bringing decision-making capabilities and transforming it into automated agriculture.
[62]	The research conducted here offered a system for obtaining and processing data, as well as interacting with other databases, for the Danish government in their own language using a web platform called "RDF". The recommended approach collects data, analyzes it, cleans it, and after that incorporates them into other data resources. A testing was undertaken, which yielded positive outcomes.
[63]	A'WOT integrated application was offered to upkeep the ancient cultural items at Agli'e Castle (Turin) that render decisions, manage items, identify issues to discover solutions.
[64]	It fulfilled the need for a remote practicable option for STEM courses and contributed to promote remote learning by combining the capabilities of WOT and WebRTC, a multi model service. By implementing it, communication between multiple users might be facilitated by allowing data to be retrieved from previously implanted devices.
[65]	This concentrates on a merged research area named Embedded Intelligence (EI). This study analyzed the past research history, properties, applications, structures, and gaps in EI research, and exhibited a proactive water supply application.

4.1.6. Quality Assessment Results

The level of each elected article is indicated in the table. Most of them have a highest score of around 50%, 45.5% have an average score, and 4.5% have the lowest point, as seen in figure. The obtained evaluation may be useful to WOT-Agricultural analysts and experts when selecting associated research.

For quality assessment, please refer to the fig 17 below.

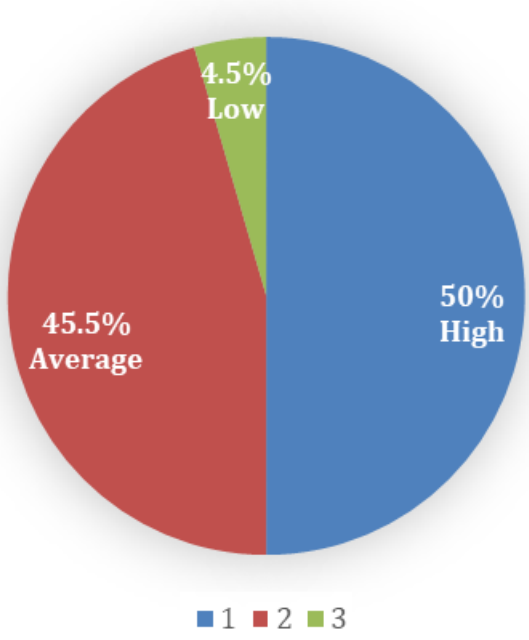


Figure 17: Article Ranking

For quality assessment score, refer to the Table 6.

Table 6: Quality assessment points

Ref.	T. Score	No.
[50]	2	1
[62]	2.5	1
[41] [42] [45] [47] [53] [59] [60]	3	7
[51] [64]	3.5	2
[46] [58]	4	2
[48] [49] [54] [27] [61] [63] [65]	4.5	7
[20] [21]	5	2

5. Discussion

This SLR extracted the works on the WOT emerged advances in agriculture research. The studies that were selected have been utilized for addressing the research questions posed in the methodology. Findings indicate that numerous studies were undertaken to identify solutions to agriculture sector difficulties.

5.1. Selected Studies Results

The principal rationale in directing this methodical writing survey was to explore the proceeding with research development of the agrarian innovation period by sifting through the 22 examinations from 220 examinations. After removing these papers in accordance with the guidelines provided. These are the most common outcomes: A number of application domains for agricultural research trends were identified through the review of selected studies on WOT involvement in agriculture. By offering improvements for complications with IoT agricultural systems, and integrated devices, the web of things paved way for smart agriculture. Web of things settled the issues of heterogeneous information, gadgets, principles, and stages. The readers are encouraged to conduct additional WOT-agricultural research by these application domains. The application areas have been grouped into eleven sorts: publishing, integration, education, interoperability, precision and decision-making, monitoring, control, management, storage, analysis, and optimization. Areas are additionally subdivided like air, water, weather gadgets, systems, monitor, controlling, managing sensors, and storing farm data, diseases, research, and environments situations, mining or analyzing agricultural data, production, and system performance and optimization as well as supply chain processes. These subdomains also include making predictions based on the data from farms and supporting decisions, combining networks, devices, and protocols for improved working, publishing datasets and experimental results, integrating the various IoT devices, networks, and sensors for efficient agriculture and providing learners with real-time information on study platforms and research implementation results. WOT and Internet of Things (IoT) combination have made it possible to break through the boundaries of previous technologies. The emergence of WOT and IOT will aid in the transition from conventional to smart agriculture. The classification table in the section devoted to the results of the research question served as the source for the application domains that are depicted in the taxonomy. The 22 selected research papers were examined in light of the WOT agricultural existing research trends in order to produce that classification table. The inclusion and exclusion criteria outlined in section 3 was used to evaluate these selected papers. This scientific classification will help in rousing for additional exploration around here.

The planned arrangement has shown in the Figure 17 below

Figure 17: Taxonomy of Application Domains

5.2. Wot Model for Agriculture

A web of things enabled sketch has been developed to demonstrate the WOT's involvement to the creation of a Smart Agricultural. It shows how WOT is used in several agricultural areas, including monitoring, controlling, and managing gadgets, machinery, yields and environment situations, and watering system by utilizing facilities of web. One of the most significant characteristics is the ability to provide real-time facilities and establish reliable connections. It provided a representation of WOT enabled advanced agriculture.

The Model has shown in the Figure 18 below.

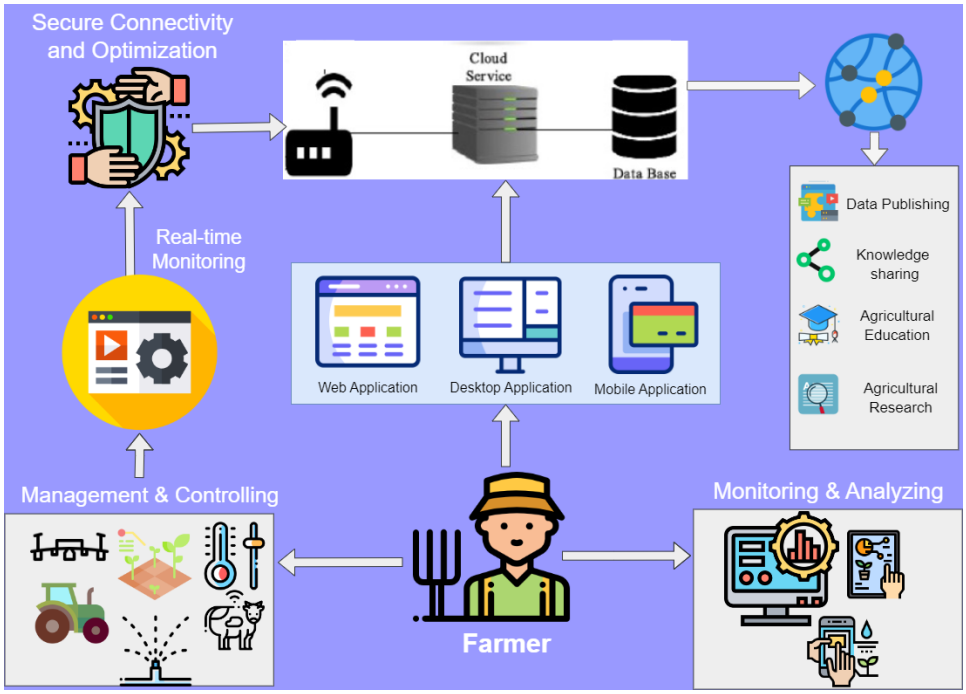


Figure 18: WOT Agricultural Model

5.3. Open Issues and Challenges

Numerous works have obtained the wot involvement and establishment of fresh approaches to the current applications and technologies issues in the agricultural area. WOT has emerged as the primary player to address the agricultural field's challenges. However, there are still unresolved obstacles and issues with WOT combination with farm applications, as indicated in the figure. The most widespread issues are security concerns related to the advent of new technologies at various layers in schemes, platforms as well physical structures, and online applications. There are still many threats to the credibility of agriculture statistics, cost concerns with adding WOT technology into hardware and software. International standards are required to provide effective security services.

The issues and challenges have shown in the Figure. 19 below.

Figure 19: Main Issues and Challenges

Farmers should be made aware of the benefits of using wot applications for more efficient farming and higher profits. There is a need of additional literature reviews, SLR, and establishment of WOT technology in agricultural industry. Farmers in remote areas have a poor understanding of technology. The refinement of WOT involvement for large-scale areas are needed. But there's a requirement for incentives to pursue additional study on existing and continuing areas of study.

5.4. Research Gap and Future Directions

The discussion was wrapped up with the presentation of numerous studies aimed at overcoming agricultural application difficulties through the use of Web of Things technologies. As a result, further comprehensive research into WOT solutions for agricultural domain challenges is still required. As previously noted, the WOT has proven to be the primary driver of agricultural progress. The majority of the studies provided existing and novel solutions in a non-systematic manner. As a result, thorough studies and literature evaluations of existing solutions research are required to aid in future research and find any flaws in the present and continuing study, as well as keys to those flaws. In addition, more emphasis should be placed on evaluation studies to assess the existing work associated to trends of WOT in agricultural industries in order give rise to smart automated agriculture.

The main research gaps and future directions are given below:

Research Gaps:

- Systematic Studies
- Literature Reviews
- Evaluation of existing approaches
- Development of new standards

5.5. Threat to Validity

This SLR may face various validity challenges, such as a poor screening of research articles, insufficient data collection, and an adequate quality evaluation of the studies that were selected.

5.5.1. Selection of Research Articles

Section III provides a detailed approach for selecting research articles. The portion provided comprehensive Inclusion/Exclusion criterion for screening research articles. Selected study years extend from 2010 to 2020 in order to determine the continual incorporation of WOT in the agricultural industry. However, it is still possible that some articles will be missing. The main cause for this possible outcome is the absence of articles published prior to 2010, as well as common interest papers. The second option might be a search string to retrieve related studies. Although the precise search string is explained in Section III, there is still the possibility that a few study publications will require additional exploring keywords. The search string mentioned in third Section has been adjusted multiple times to discover the most appropriate study articles, however a gap remains for the fresh term.

5.5.2. Insufficient Collection of Data

A different possible cause for uncertain findings is insufficient data gathering. The likelihood of this was reduced by evaluating the collected data three times.

5.5.3. Evaluation of Quality

Most important part in SLR is to test quality, as poor quality might lead to untrustworthy findings. Section III discusses a workable technique for ensuring the quality of the selected research articles.

6. Conclusion

A comprehensive literature review of previous studies in the Web of Things-based agriculture field is

described in this study. Based on the well-defined methodology, 22 studies were selected for this comprehensive survey. The main works were filtered out of the review in strict accordance with the established guidelines. Each and every piece of work was thoroughly examined for effectiveness. It came to the conclusion that WOT technology offered potential improvements in area of agriculture after conducting a comprehensive review of previous studies. By providing web-enabled services, improving existing systems, and assisting in the creation of online informative databases for farmers, educational institutions, and the general public, WOT astonishingly surpasses compatibility issues of IoT in agriculture industry. Taking into account what is happening, we have introduced a scientific classification of farming application spaces where the snare of things offered the types of assistance to conquer the current mechanical difficulties in the horticulture space. It depicts the agricultural sectors in which services were utilized in a variety of ways, including techniques, both ancient and new, to minimize complications. The process of integrating the web of things in the agricultural sector was outlined in a model of what-based smart agriculture. The researchers will benefit from this taxonomy and model in terms of gaining a clear understanding of the agricultural domains in which WOT has been and is being utilized, as well as in deciding on the preferred application areas in accordance with requirements. It will show the ways for additional new areas where WOT can be valuable in dealing with the cultivating and other horticulture tasks. The governments of various nations have WOT agriculture plans of their own and promote WOT-based research in agriculture. Determining the way these application areas are participating to motivate readers is one of the promising future directions in this study. Despite the fact that a lot of research has been done on WOT-based agriculture, not much of it has been presented as SLR, SMS, or SR. Further investigation in this field must be directed around here to inspire further exploration.

References

[1] I. Mat, M. R. Mohd Kassim, A. N. Harun, and I. M. Yusoff, “Smart agriculture using internet of things,” in 2018 *IEEE Conference on Open Systems (ICOS),* pp. 54–59, 2018.

[2] I. Charania and X. Li, “Smart farming: Agriculture’s shift from a labor intensive to technology native industry,” *Internet of Things*, vol. 9, p. 100142, 2020.

[3] J. Lindblom, C. Lundström, M. Ljung, and A. Jonsson, “Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies,” *Precision Agriculture*, vol. 18, no. 3, pp. 309–331, 2017.

[4] A. Balafoutis, B. Beck, S. Fountas, J. Vangeyte, T. V. d. Wal, I. Soto, M. Gómez-Barbero, A. Barnes, and V. Eory, “Precision agriculture technologies positively contributing to ghg emissions mitigation, farm productivity and economics,” *Sustainability*, vol. 9, no. 8, 2017.

[5] A. P. Barnes, I. Soto, V. Eory, B. Beck, A. Balafoutis, B. Sánchez, J. Vangeyte, S. Fountas, T. van derWal, and M. Gómez-Barbero, “Influencing factors and incentives on the intention to adopt precision agricultural technologies within arable farming systems,” *Environmental Science and Policy*, vol. 93, pp. 66–74, 2018.

[6] A. Knierim, M. Kernecker, K. Erdle, T. Kraus, F. Borges, and A. Wurbs “Smart farming technology innovations – insights and reflections from the german smart-akis hub,” *NJAS - Wageningen Journal of Life Sciences*, vol. 90-91, p. 100314, 2019.

[7] A. Marucci, A. Colantoni, I. Zambon, and G. Egidi, “Precision farming in hilly areas: The use of network rtk in gnss technology,” *Agriculture*, vol. 7, no. 7, 2017.

[8] A. T. Balafoutis, B. Beck, S. Fountas, Z. Tsiropoulos, J. Vangeyte, T. van der Wal, I. Soto- mbodas, M. Gómez-Barbero, and S. M. Pedersen, “Smart farming technologies–description, taxonomy and economic impact,” in *Precision Agriculture: Technology and Economic Perspectives*, pp. 21–77, Springer, 2017.

[9] M. S. Farooq, S. Riaz, A. Abid, T. Umer, and Y. B. Zikria, “Role of iot technology in agriculture: A systematic literature review,” *Electronics*, vol. 9, no. 2, 2020.

[10] D. Evans, “The internet of things: How the next evolution of the internet is changing everything,” *CISCO white paper*, vol. 1, no. 2011, pp. 1–11, 2011.

[11] C. V. N. Index, “Forecast and methodology, 2012–2017,” *White paper*, vol. 29, 2013.

[12] A. Mazayev, J. A. Martins, and N. Correia, "Semantic web thing architecture," in *2017 4th Experiment@ International Conference* (exp. at'17), pp. 43–46, IEEE, 2017.

[13] D. Guinard, V. Trifa, F. Mattern, and E.Wilde, "From the internet of things to the web of things: Resource-oriented architecture and best practices," in *Architecting the Internet of things*, pp. 97–129, Springer, 2011.

[14] D. Raggett, "The web of things: Challenges and opportunities," *Computer*, vol. 48, no. 5, pp. 26–32, 2015.

[15] E. Antonopoulou, S. Karetsos, M. Maliappis, and A. Sideridis, "Web and mobile technologies in a prototype dss for major field crops," *Computers and Electronics in Agriculture,* vol. 70, no. 2, pp. 292–301, 2010.

[16] D. L. Hernández-Rojas, T. M. Fernández-Caramés, P. Fraga-Lamas, and C. J. Escudero, "A plug-and-play human-centered virtual teds architecture for the web of things," *Sensors*, vol. 18, no. 7, p. 2052, 2018.

[17] M. Khan, B. N. Silva, and K. Han, "A web of things-based emerging sensor network architecture for smart control systems," *Sensors*, vol. 17, no. 2, 2017.

[18] A. Rhayem, M. B. A. Mhiri, and F. Gargouri, "Semantic web technologies for the internet of things: Systematic literature review," *Internet of Things*, p. 100206, 2020.

[19] F. Aznoli and N. J. Navimipour, "Deployment strategies in the wireless sensor networks: systematic literature review, classification, and current trends," *Wireless Personal Communications*, vol. 95, no. 2, pp. 819–846, 2017.

[20] N. Chen, X. Zhang, and C.Wang, "Integrated open geospatial web service enabled cyber-physical information infrastructure for precision agriculture monitoring," *Computers and Electronics in Agriculture*, vol. 111, pp. 78– 91, 2015.

[21] B. Negash, T. Westerlund, and H. Tenhunen, "Towards an interoperable internet of things through a web of virtual things at the fog layer," *Future Generation Computer Systems*, vol. 91, pp. 96–107, 2019.

[22] F. Bauer and M. Kaltenböck, "Linked open data: The essentials," *Edition mono/monochrom, Vienna*, vol. 710, 2011.

[23] T. Katayama, M. D. Wilkinson, G. Micklem, S. Kawashima, A. Yamaguchi, M. Nakao, Y. Yamamoto, S. Okamoto, K. Oouchida, H.-W. Chun, et al., "The 3rd dbcls biohackathon: improving life science data integration with semantic web technologies," *Journal of biomedical semantics*, vol. 4, no. 1, pp. 1–17, 2013.

[24] T. Heath and C. Bizer, "Linked data: Evolving the web into a global data space," *Synthesis lectures on the semantic web: theory and technology*, vol. 1, no. 1, pp. 1–136, 2011.

[25] O. Aziz, T. Anees, and E. Mehmood, "An efficient data access approach with queue and stack in optimized hybrid join," *IEEE Access*, pp. 1–1, 2021.

[26] H.-G. Kim, "Semantic web," 2003.

[27] D. Lukose et al., "World-wide semantic web of agriculture knowledge," *Journal of Integrative Agriculture*, vol. 11, no. 5, pp. 769–774, 2012.

[28] C. Caracciolo, A. Morshed, A. Stellato, G. Johannsen, Y. Jaques, and J. Keizer, "Thesaurus maintenance, alignment and publication as linked data: The agroovoc use case," in *Metadata and Semantic Research (E. García-Barriocanal, Z. Cebeci, M. C. Okur, and A. Öztürk, eds.), (Berlin, Heidelberg)*, pp. 489–499, Springer Berlin Heidelberg, 2011.

[29] A. Chehri, H. Chaibi, R. Saadane, N. Hakem, and M. Wahbi, "A framework of optimizing the deployment of iot for precision agriculture industry," *Procedia Computer Science*, vol. 176, pp. 2414–2422, 2020.

[30] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, "Big data in smart farming – a review," *Agricultural Systems*, vol. 153, pp. 69–80, 2017.

[31] A. R. de Araujo Zanella, E. da Silva, and L. C. P. Albini, "Security challenges to smart agriculture: Current state, key issues, and future directions," *Array*, p. 100048, 2020.

[32] H. S. Abdullahi, F. Mahieddine, and R. E. Sheriff, "Technology impact on agricultural productivity: A review of precision agriculture using unmanned aerial vehicles," in *Wireless and Satellite Systems (P. Pillai, Y. F. Hu, I. Otung, and G. Giambene, eds.), (Cham)*, pp. 388–400, Springer International Publishing, 2015.

[33] F. S. Khan, S. Razzaq, K. Irfan, F. Maqbool, A. Farid, I. Illahi, and T. U. Amin, "Dr. wheat: a web-based expert system for diagnosis of diseases and pests in pakistani wheat," in *Proceedings of the World Congress on Engineering*, vol. 1, pp. 2–4, Citeseer, 2008.

[34] A. Walter, R. Finger, R. Huber, and N. Buchmann, "Opinion: Smart farming is key to developing sustainable agriculture," *Proceedings of the National Academy of Sciences,* vol. 114, no. 24, pp. 6148–6150, 2017.

[35] A. D. Boursianis, M. S. Papadopoulou, P. Diamantoulakis, A. Liopa- Tsakalidi, P. Barouchas, G. Salahas, G. Karagiannidis, S. Wan, and S. K. Goudos, "Internet of things (iot) and agricultural unmanned aerial vehicles (uavs) in smart farming: A comprehensive review," *Internet of Things,* p. 100187, 2020.

[36] M. Torky and A. E. Hassanein, "Integrating blockchain and the internet of things in precision agriculture: Analysis, opportunities, and challenges," *Computers and Electronics in Agriculture,* p. 105476, 2020.

[37] A. Kamilaris, A. Fonts, and F. X. Prenafeta-Bold, "The rise of blockchain technology in agriculture and food supply chains," *Trends in Food Science Technology,* vol. 91, pp. 640–652, 2019.

[38] D. Pivoto, P. D. Waquil, E. Talamini, C. P. S. Finocchio, V. F. Dalla Corte, and G. de Vargas Mores, "Scientific development of smart farming technologies and their application in brazil," *Information processing in agriculture,* vol. 5, no. 1, pp. 21–32, 2018.

[39] D. Glaroudis, A. Iossifides, and P. Chatzimisios, "Survey, comparison and research challenges of iot application protocols for smart farming," *Computer Networks,* vol. 168, p. 107037, 2020.

[40] S. Blank, C. Bartolein, A. Meyer, R. Ostermeier, and O. Rostanin, "igreen: A ubiquitous dynamic network to enable manufacturer independent data exchange in future precision farming," *Computers and electronics in agriculture,* vol. 98, pp. 109–116, 2013.

[41] J. Ye, B. Chen, Q. Liu, and Y. Fang, "A precision agriculture management system based on internet of things and webgis," in *2013 21st International Conference on Geoinformatics,* pp. 1–5, IEEE, 2013.

[42] E. Jahanshiri and S.Walker, "Agricultural knowledge-based systems at the age of semantic technologies," *Inter. J. Know. Engin,* vol. 1, no. 1, pp. 64– 67, 2015.

[43] Wikipedia contributors, "Knowledge-based systems—Wikipedia, the free encyclopedia," 2021.

[44] C. Yialouris and A. Sideridis, "An expert system for tomato diseases," *Computers and electronics in agriculture,* vol. 14, no. 1, pp. 61–76, 1996.

[45] P. Damos, S. Karabatakis, et al., "Real time pest modeling through the world wide web: decision making from theory to praxis," *IOBC-WPRS Bulletin,* vol. 91, pp. 253–258, 2013.

[46] K. Grgic´, I. Špeh, and I. Hed¯i, "A web-based iot solution for monitoring data using mqtt protocol," in *2016 international conference on smart systems and technologies (SST),* pp. 249–253, IEEE, 2016.

[47] F. Karim, F. Karim, et al., "Monitoring system using web of things in precision agriculture," *Procedia Computer Science,* vol. 110, pp. 402–409, 2017.

[48] M. P. Vilas, P. J. Thorburn, S. Fielke, T.Webster, M. Mooij, J. S. Biggs, Y.- F. Zhang, A. Adham, A. Davis, B. Dungan, et al., "1622wq: *A web-based application to increase farmer awareness of the impact of agriculture on water quality," Environmental Modelling & Software,* vol. 132, p. 104816, 2020.

[49] P. P. Jayaraman, A. Yavari, D. Georgakopoulos, A. Morshed, and A. Zaslavsky, "Internet of things platform for smart farming: Experiences and lessons learnt," *Sensors,* vol. 16, no. 11, p. 1884, 2016.

[50] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu, and M. I. Ali, "Agri-iot: A semantic framework for internet of things-enabled smart farming applications," in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT),* pp. 442–447, IEEE, 2016.

[51] M. Terdjimi, L. Médini, M. Mrissa, and M. Maleshkova, "Multi-purpose adaptation in the web of things," in *International and Interdisciplinary Conference on Modeling and Using Context,* pp. 213–226, Springer, 2017.

[52] M. Mrissa, L. Médini, J.-P. Jamont, N. Le Sommer, and J. Laplace, "An avatar architecture for the web of things," *IEEE Internet Computing,* vol. 19, no. 2, pp. 30–38, 2015.

[53] A. Ordoñez-García, E. V. Núñez, M. Siller, and M. G. S. Cervantes, "Iot system for agriculture: Web technologies in real time with the middleware paradigm.," in *2018 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC),* pp. 1–4, IEEE, 2018.

[54] B. Drury, R. Fernandes, M.-F. Moura, and A. de Andrade Lopes, "A survey of semantic web technology for agriculture," *Information Processing in Agriculture,* vol. 6, no. 4, pp. 487–501, 2019.

[55] S. Staab, "Ontology engineering," 2009.

[56] W. Qian, T. Lan, and Z. Lijun, "Approach to ontology construction based on text mining," *New Zealand Journal of Agricultural Research,* vol. 50, no. 5, pp. 1383–1391, 2007.

[57] F. Amarger, J.-P. Chanet, O. Haemmerlé, N. Hernandez, and C. Roussey, "Skos sources transformations for ontology engineering: Agronomical taxonomy use case," in *Research Conference on Metadata and Semantics Research*, pp. 314–328, Springer, 2014.

[58] G. Donca et al., "Aspects of wot contribution to sustainable agricultural production.," *Analele Universit˘at, ii din Oradea, Fascicula: Protect,ia Mediului,* vol. 26, pp. 27–34, 2016.

[59] M. Shoaib and A. Basharat, "Semantic web based integrated agriculture information framework," in *2010 Second International Conference on Computer Research and Development*, pp. 285–289, IEEE, 2010.

[60] L. Liu and D. Ling, "Discussion on the optimization of web of things supply chain of agricultural products and information sharing based on rfid," in *Journal of Physics: Conference Series*, vol. 1578, p. 012114, IOP Publishing, 2020.

[61] L. Touseau and N. Le Sommer, "Contribution of the web of things and of the opportunistic computing to the smart agriculture: A practical experiment," *Future Internet*, vol. 11, no. 2, p. 33, 2019.

[62] A. B. Andersen, N. Gür, K. Hose, K. A. Jakobsen, and T. B. Pedersen, "Publishing danish agricultural government data as semantic web data," in *Joint International Semantic Technology Conference*, pp. 178–186, Springer, 2014.

[63] M. Bottero, C. D'Alpaos, and A. Marello, "An application of the a'wot analysis for the management of cultural heritage assets: The case of the historical farmhouses in the aglié castl (turin)," *Sustainability*, vol. 12, no. 3, p. 1071, 2020.

[64] B. M. Degboe, U. H. S. Boko, K. Gueye, and S. Ouya, "Contribution to the setting up of a remote practical work platform for stem: The case of agriculture," in *International Conference on e-Infrastructure and e-Services for Developing Countries*, pp. 88–97, Springer, 2018.

[65] W. Yong, L. Shuaishuai, L. Li, L. Minzan, L. Ming, K. Arvanitis, C. Georgieva, and N. Sigrimis, "Smart sensors from ground to cloud and web intelligence," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 31–38, 2018.

[66] O. Aziz, M. S. Farooq, A. Abid, R. Saher, and N. Aslam, "Research trends in enterprise service bus (esb) applications: a systematic mapping study," *IEEE Access*, vol. 8, pp. 31180–31197, 2020.

[67] I. Obaid, M. S. Farooq, and A. Abid, "Gamification for recruitment and job training: Model, taxonomy, and challenges," *IEEE Access*, vol. 8, pp. 65164–65178, 2020.

[68] M. S. Farooq, S. Riaz, A. Abid, T. Umer, and Y. B. Zikria, "Role of iot technology in agriculture: A systematic literature review," *Electronics*, vol. 9, no. 2, p. 319, 2020.

[69] Z. A. Barmi, A. H. Ebrahimi, and R. Feldt, "Alignment of requirements specification and testing: A systematic mapping study," in *2011 IEEE Fourth International Conference on Software Testing, Verification and Validation Workshops*, pp. 476–485, IEEE, 2011.

[70] E. Jahanshiri and S. Walker, "Agricultural knowledge-based systems at the age of semantic technologies," *International Journal of Knowledge Engineering*, vol. 1, pp. 64–67, 01 2015.

[71] P. Damos and S. Karabatak˙Is, "Real time pest modeling through the world wide web: decision making from theory to praxis.," *IOBC/WPRS Bulletin*, vol. 91, pp. 253–258, 2013.

[72] K. Grgic´, I. Špeh, and I. Hed¯i, "A web-based iot solution for monitoring data using mqtt protocol," in *2016 International Conference on Smart Systems and Technologies (SST)*, pp. 249–253, 2016.

[73] F. Karim, F. Karim, and A. frihida, "Monitoring system using web of things in precision agriculture," Procedia Computer Science, vol. 110, pp. 402–409, 2017. *14th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2017) / 12th International Conference on Future Networks and Communications (FNC 2017) /* Affiliated Workshops.

[74] M. P. Vilas and P. J, "1622wq: A web-based application to increase farmer awareness of the impact of agriculture on water quality," *Environmental Modelling Software*, vol. 132, p. 104816, 2020.

[75] P. P. Jayaraman, A. Yavari, D. Georgakopoulos, A. Morshed, and A. Zaslavsky, "Internet of things platform for smart farming: Experiences and lessons learnt," *Sensors*, vol. 16, no. 11, 2016.

[76] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu, and M. I. Ali, "Agri-iot: A semantic framework for internet of things-enabled smart farming applications," in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pp. 442–447, 2016.

[77] M. Terdjimi, L. Médini, M. Mrissa, and M. Maleshkova, "Multi-purpose adaptation in the web of things," in *Modeling and Using Context (P. Brézillon, R. Turner, and C. Penco, eds.), (Cham),* pp. 213–226, Springer International Publishing, 2017.

[78] A. Ordoñez-García, E. V. Núñez, M. Siller, and M. G. S. Cervantes, "Iot system for agriculture: Web technologies in real time with the middleware paradigm.," in *2018 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC),* pp. 1–4, 2018.

[79] B. Drury, R. Fernandes, M.-F. Moura, and A. de Andrade Lopes, "A survey of semantic web technology for agriculture," *Information Processing in Agriculture*, vol. 6, no. 4, pp. 487–501, 2019.

[80] G. Donca, "Aspects of wot contribution to sustainable agricultural production.," *Analele Universit˘at, ii din Oradea, Fascicula: Protect,ia Mediului,* vol. 26, pp. 27–34, 2016.

[81] M. Shoaib and A. Basharat, "Semantic web based integrated agriculture information framework," in *2010 Second International Conference on Computer Research and Development*, pp. 285–289, 2010.

[82] L. Liu and D. Ling, "Discussion on the optimization of web of things supply chain of agricultural products and information sharing based on RFID," *Journal of Physics: Conference Series,* vol. 1578, p. 012114, jul 2020.

[83] B. Negash, T. Westerlund, and H. Tenhunen, "Towards an interoperable internet of things through a web of virtual things at the fog layer," *Future Generation Computer Systems,* vol. 91, pp. 96–107, 2019.

[84] L. Touseau and N. Le Sommer, "Contribution of the web of things and of the opportunistic computing to the smart agriculture: A practical experiment," Future Internet, vol. 11, no. 2, 2019.

[85] A. B. Andersen, N. Gür, K. Hose, K. A. Jakobsen, and T. B. Pedersen, "Publishing danish agricultural government data as semantic web data," in *Semantic Technology (T. Supnithi, T. Yamaguchi, J. Z. Pan, V.Wuwongse, and M. Buranarach, eds.), (Cham)*, pp. 178–186, Springer International Publishing, 2015.

[86] M. Bottero, C. D'Alpaos, and A. Marello, "An application of the a'wot analysis for the management of cultural heritage assets: The case of the historical farmhouses in the aglié castle (turin)," *Sustainability*, vol. 12, no. 3, 2020.

[87] B. M. Degboe, U. H. S. Boko, K. Gueye, and S. Ouya, "Contribution to the setting up of a remote practical work platform for stem: The case of agriculture," in *e-Infrastructure and e-Services for Developing Countries (G. Mendy, S. Ouya, I. Dioum, and O. Thiaré, eds.), (Cham),* pp. 88–97, Springer International Publishing, 2019.

[88] W. Yong, L. Shuaishuai, L. Li, L. Minzan, L. Ming, K. Arvanitis, C. Georgieva, and N. Sigrimis, "Smart sensors from ground to cloud and web intelligence," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 31–38, 2018.

[89] Verdouw, C. N., Vucic, N., Sundmaeker, H., & Beulens, A. J. (2014). Future internet as a driver for virtualization, connectivity and intelligence of agri-food supply chain networks. *International journal on food system dynamics*, *4*(4), 261-272.

Machines and Algorithms
http://www.knovell.org/mna

Review Article

Methodology for the Design and Implementation of Smart Billboards in Pakistan: (Case Study)

Fatima Riaz[1, *] and Rabea Saleem[2]

[1]Lecturer Computer Science, Higher Education Department, Multan, 60000, Pakistan
[2]Department of Computer Science, Ibadat International University Islamabad, Islamabad, 44000, Pakistan
[*]Corresponding Author: Fatima Riaz. Email: fatimariaz.fati@gmail.com
Received: 01 June 2023; Revised: 03 July 2023; Accepted: 15 August 2023; Published: 31 October 2023
AID: 002-03-000029

Abstract: Billboard's advertisement is an effective way to communicate with the audience by influencing the user mind and motivate them to purchase the respective items. Attention-Interest-Desire-Action (AIDA) model was initiated to target the audience through traditional advertisement (physical billboards) and the viewer's become willing to purchase the product or to avail the service. In result, the companies can earn more profits, but for these advertisements they have to pay a significant amount to the advertisement companies, which may be affordable for big brands but not by the small-scale companies. Moreover, they are placed for some fixed days even at those times when their audience are not frequent on the road, for example Candies (Sweets and Chocolate) advertisement can be best suited when the students are on road, Restaurant advertisement can be best suited after the office timing and family products advertisement can be best suited in evening. This highlights the main issue of observing the target audience with respect to their age groups and gender, and accordingly the registered companies can see the crowd behavior. The focus of this research is to automatically identify the audience and provide a platform to the registered users to bid for their advertisement on the specific time slots (the slots are hourly classified and a bid is for one week) by occupying the entire billboard or the segment of it. This may be useful for both advertising agencies and companies to promote their products in a sensible way. A questionnaire survey has been conducted on mall shopkeepers and customers to know their willingness for smart billboard that is justified by significant empirical results of this survey.

Keywords: Smart Billboards; Digital Billboards; Advertisement; Ambient Intelligence;

1. Introduction

Advertisement is an effective way of influencing the audience by keeping them aware, and retain, about the products, services or ideas. Billboards are a traditional way of advertisement; whose target audience is general public. Its appealing graphics work on the phenomena of Attention-Interest-Desire-Action (AIDA) model which motivate the viewers for the purchasing of product or availing the service. In result, companies increase their sales and earn more profits. Typically, billboard is an expensive way of advertisement, but due to its attention gaining advantage over other traditional methods e.g., television, mobile SMS, emails, or brochures, — as they can be easily avoided — the billboard advertisement is mostly adopted by big

brands. The local and small industries/companies — which are further refer as “Clients Billboard” (CB) in this research– are often working in tightrope of challenges and are not able to bear their expense, that’s way they are unable to avail the benefits of billboard advertisement and interact new customers. Which not only effect their sales/profit but also have an impact on country economy, as the growth of any country economy is depend on the productivity of their local industry. Thanks to the advancement in the area of Ambient Intelligence, image processing and smart technologies by which the billboard advertisement can be affordable to CB. The research in the area of technologically embedded billboard can also be referred as Billboard Digitalization or Smart Billboard (SB). To expedite the review and typesetting process, authors must follow the Microsoft Word template provided for preparing their manuscripts. This template must be strictly adhered to when formatting the manuscript for submission.

1.1. Research Gap and Contribution of Study

A surge of literature has paid attention to digital billboard advertisement for big or international brands but no study has focused to develop digital billboards for local manufacturers that operate at small scale. This study contributes to the literature by developing cost effective advertisement pattern of smart billboard that is suitable for local manufacturers because these manufacturers operate at small scale and are not well established so they cannot pay a handsome amount from their meagre profits for marketing of their products for a long period of time.

Although substantial literature has focused on billboard advertisement but no study has analyzed from the consumer perspective. This study also contributes to the literature by analyzing the feasibility of smart billboard advertisement pattern from consumer’s perspective after taking their opinion through statistical analysis of fata from field survey and considering human psychology. It is because traditional billboards become less attractive and thus affects human psychology severely as constantly watching the same advertisement molds consumer’s mind and makes them more loyal and attached to that specific product by ignoring other brands and they remain unaware of new products arrival.

1.2. Objective of the Study

The main objective of this research is to facilitate the local and small industries/companies by providing cost effective way of advertisement. For this, time and frame based slicing approach is adopting. Instead of full board advertisement, it can be displayed at some portion/slice of the SB and instead of full day display it can be displayed for some certain time. The motivation behind time slicing is that the Target Customers (TC) are not always frequent on the road. For the identification of TC image processing technique is used by which the age and gender of the pedestrians can be known. The occurrence frequencies of TC help the CB for the selection of appropriate time at which their advertisement can display, for example candies (Sweets and Chocolate) advertisement can be best suited when the students are on road, restaurant advertisement can be best suited after the office timing and family products advertisement can be best suited in evening.

2. Review of Related Work

Chen and Lin (2009) have presented an approach for the counting system of viewers. At first step they count the number of people by using face detection technique, then they filter the faces on the basis of those faces whose are watching towards the advertisement. Due to the low resolution of surveillance videos, sometimes, the identification of face portion is not practicable therefore they use extracted features from the torso region for the compensation of deficiency.

Shlomot and Beach (2013) have presented a methodology of displaying the advertisement on billboards according to user interest. For this, firstly they identify the passing cars, passengers in cars and items in car and then – by viewing/knowing the characteristics of these from their repository – the selected advertisement is displayed on the electronic billboards.

According to Khalil et al. (2023) all facets of advertising are undergoing a revolution due to technological advancements in recent decades, including Internet of Things (IoT) and Internet of Behavior (IoB) technologies. Out-of-home (OOH) advertising is no exception. The fact that OOH is currently unable to precisely target consumers based on their tastes and behaviors is one of its biggest issues. The research's objective is to examine how IoT and IoB might enhance OOH targeting capabilities before putting out a case study for potential future opportunities.

S.Sümer et al. (2022) have observed a fierce rivalry between the businesses along with the changes in customer preferences in recent years. The companies' decision-makers have begun designing new marketing strategies based on the requirements and expectations of their target markets. Businesses have begun to design their advertising strategies with a concentration on value. At this point, customized advertising is one of the most powerful means of developing the customer interest. To put it another way, customized marketing is the simplest approach to target particular audiences and markets. As a result, it grabs the interest of numerous companies. This indicates that, in comparison, targeted marketing initiatives attract greater customer attention. It is also true that it may positively affect the company's sales.

Naleer (2020) has developed software for productivity based advertising. For this purpose a camera's vision is utilized to recognize faces, gender, and age using artificial intelligence. (Processing Images). This study employs artificial intelligence to recognize faces, gender, and age through camera vision. Additionally, this research offers a technique that can determine an individual's gender, age range, and what shows are most likely to draw different people's attention to commercials. Robust Algorithm was utilized by the researcher to detect faces in color footage. In color photos, there is a robust algorithm for frontal face detection.

Pourhossein and Mahmoudabadi (2019) developed a smart urban street advertising pattern that will provide a meaningful link between the "planned messages" that smart billboards display and the "parameters received from the environment." The primary concept of the suggested pattern is to use a billboard equipped with magnetic, thermal, optical, and environmental sensors to sense various aspects. A questionnaire has been designed and is being used to gather audience opinions for the purpose of picking ads based on the parameters felt by smart billboards, with the aim of investigating the relationships that currently exist between the advertisement patterns. The acquired data has been evaluated using a sample size of 107 in order to assess the general idea of selecting adverts based on quantifiable factors. The study findings indicate a noteworthy correlation between environmental and traffic characteristics and the advertising contexts recommended by intelligent billboards. Thus, suggesting smart billboards that can identify traffic and environmental factors can enhance the effectiveness of advertisements to draw in viewers.

Shan and Roh (2018) have presented an analytical model in this research to describe the intricate correlations that exist between the discovery time and other BLE operational factors, including the number of advertisements, advertiser interval, scan interval, and so forth. We have also suggested a very efficient way to find the ideal advertisement interval to reduce the discovery time using the analytical model. It has been demonstrated that a significant reduction in the discovery time is possible with the suggested strategy.

Haval A. (2016) have reviewed the existing literature on "Face Behavior Recognition using Support Vector Machines". They have suggested new method for extracting the facial feature point extraction methodology and using it to recognize facial expressions from still images of two distinct sets of facial expressions. The five primary phases of this face behavior recognition system are as follows: pre-processing, feature extraction, feature selection, and classification.

Yang et al. (2016) have presented WIDER FACE (a sizable face detection dataset) by putting out a two-stage, multi-scale cascade framework that employs a divide and conquer tactic to address large-scale variances, they provide WIDER FACE. In this approach, a collection of convolutional networks with different input sizes are trained to handle a given range of scale faces.

Liu et al. (2004) have enhance the functionalities of SB by using wireless technologies. When a person see advertisement on SB, its identification by using data processing system (profile based) is made and

accordingly persons' interested products advertisement is played on the its mobile devices, PDAs or cell phones.

A parametric review of the above found literature is also presented in Table 1. All the parameters are self-explanatory and their values show the consideration of that activity in the work or not.

To expedite the review and typesetting process, authors must follow the Microsoft Word template provided for preparing their manuscripts. This template must be strictly adhered to when formatting the manuscript for submission.

Table 1: Parametric Analysis of the Found Literature.

Parameters	**Shlomot and Beach (2013)**	**Haritaoglu and Flickers (2002)**	**Nigam et al. (2016)**	**Smart Billboard**
Use of sensors	✓	✓	✓	✓
Digital / Electronic Billboards	✓	✓	✓	✓
Customer satisfaction	×	✓	×	✓
Targeted individual/ crowd of customers	Individual	Crowd of customers	Crowd of customers	✓
Changing Advertisement	✓	✓	✓	✓
Check effectiveness of companies Advertisement	×	✓	×	✓
Technique use	Camera and control unit to extract information	Silhouette based algorithm	Data mining and RFID tags	Face++ API
Customer Relationship Management	×	✓	×	✓
Cost effective	×	×	×	✓
Customer Behavior (based on gender)	×	×	×	✓
Gender and Age Detection				✓
Attractiveness	×	×	×	✓

3. Methodology

This section is further divided in two subsections. In the first section a survey about the feasibility of research is conducted from various stakeholders and in section 2 pattern of smart billboard has been designed.

3.1. Empirical Research Findings:

A survey is conducted for knowing the applicability constructing hypotheses to judge the interaction and interest in smart billboard of its stakeholders particularly shopkeepers and customers.

Considering the studies carried out in this area, several parameters are considered such as Customer Acceptance, Shopkeeper Acceptance, Customer Attractiveness, Shopkeeper Attractiveness, Smart Billboard, Traditional Billboard, Print Media, and Broadcast. Data is collected from shopkeepers and customers through questionnaire survey. A sample of size 100 in considered in case of shopkeepers and 150 in the case of customers. Responses have been collected through a five-score Likert scale (strongly agree, agree, no opinion, disagree, strongly disagree) is used for selecting advertising context in different areas and conditions. Consequently, a variety of hypotheses are tested to study the relationship between the above-mentioned parameters of this study.

Premises 1 concludes about the shopkeepers' response to SB, premise 2 concludes about customers' interest and preference of SB, whereas premises 3 shows the acceptance of SB over traditional billboards. These premises are analyzed using correlation coefficient (R), ANOVA table and regression techniques. These techniques incorporate statistical tests of significance for example F-test, t-test and coefficient of determination (R2) to accept or reject the null and alternative hypotheses. Moreover, the Variables used in premises are abbreviated as SAtt (shopkeepers' attractiveness), SAcc (shopkeepers' acceptance), Cust_Att (customers' attractiveness), Cust_Acc (customers' acceptance) and AD (Advertisement Method).

- **Premise 1:** The use of smart technologies in advertisement is attractive for shopkeeper.

Table 1.1 exhibits correlation coefficient (R) and coefficient of determination (R2). R exhibits 55% of linear association between these variables. R2 explains that 30.5% of variations in SAcc is caused by customization. Both indicators provide evidence that this model is valid.

Table 2: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error if the Estimate
1	.552[a]	.305	.298	.58613

a. Predictors: (Constant), Cust_M

Source: Author's estimates using SPSS 21

Ho: Billboard attractiveness does not affect the shopkeeper acceptance.

H1: Billboard attractiveness affects the shopkeeper acceptance.

Table 3: ANOVA Table

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	14.767	1	14.767	42.983	.000[c]
Residual	33.668	98	.344		
Total	48.435	99			

a. Dependent Variable: SAcc_M
b. Predictors: (Constant), SAtt_M
c. Significance level(a) = 5%

Source Author's estimates using SPSS21

Table 3 represents ANOVA analysis that incorporates F-test to check the overall significance of regression model. The probability value of F-test is 0.000 that is signifying the relationship between SAtt and SAcc.

Table 4: Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficient	t	Sig.
	B	Std. Error	Beta		
(Constant)	1.401	.345		4.061	.000[b]
SAtt_M	.557	.085	.552	6.556	.000[a]

a. Dependent Variable: SAcc_M
b. Significance level(a) = 5%
Source Author's estimates using SPSS21

Table 4 represents the coefficients obtained in regression model. Beta means that when shopkeepers are attracted from SB then they adopt SB technology 0.557 times on average. This coefficient is also significant as verified by its t-statistical probability.

- **Premise 2:** The attractiveness of smart technologies in advertisement is acceptable for customers.

Table 5: Model Summary

Model	R	R Square	Adjusted R square	Std. Error of the Estimate
1	.169[a]	.029	.022	.73015

a. Predictors: (Constant),Cust_Att_M
Source: Author's estimates using SPSS21

R signifies the 17% linear association between Cust_Acc and Cust_Att.

Null (Ho) and Alternative (H1) hypothesis in case premise 2 are following:

- H0: Billboard attractiveness does not determine the customers' acceptance.
- H1: Billboard attractiveness determines the customers' acceptance.

Table 6: ANOVA[a] Table

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	2.331	1	2.331	4.373	.038[c]
Residual	78.902	148	.533		
Total	81.233	149			

a. Dependent Variable: Cust_Acc_M
b. Predicators (Constant), Cust_Att_M
c. Significance level (α) = 5%

The significance of attractiveness of SB on customer's intention to adopt SB technology is analyzed by ANOVA table 6. The probability value of F-test is 0.038 that is less than 5% level of significance. So, it signifies the relationship between Cust_Att and Cust_Acc

Table 7: Regression Coefficient

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. error	beta		
Constant	3.024	.174	.533	17.344	.000[b]
Cust_Att_M	.091	.043	.169	2.091	.038[b]

a. Dependent Variable: Cust_Acc_M
b. Significance level (α) = 5%
Source : Author Estimate using SPSS21

Beta coefficient depicts that in response to one unit change in attractiveness customers adopt it 0.091 times. This relation is supported with t-statistic that is significant at 5% level of significance.

- **Premise 3:** The comparison between smart billboard traditional billboards.

Customer table 3 exhibits the frequency and percent distribute on of 4 modes of advertisement. The most preferred mode of advertisement is SB (57%) while print media is the least preferred mode (9%). Traditional billboard is preferred by 24% and broadcast mode of advertisement by 10% people.

Table 8: Advertisement Method

		Frequency	Percent	Valid Percent	Cumulative Percentage
Valid	Traditional Billboards	24	24.0	24.0	24.0
	Smart Billboards	57	57.0	57.0	81.0
	Print Media	9	9.0	9.0	90.0
	Broadcast	10	10.0	10.0	100.0
	Total	100	100.0	100.0	

Source: Authors Calculation using SSP21

Following Pie chart portrays the comparative preference of different modes of Advertisement by shopkeepers. SB has occupied the greatest slice of this pie as 57% people prefer it while print media has occupied the smallest slice 9%. The remaining area is distributed between traditional billboard (24%) and broadcast (10%).

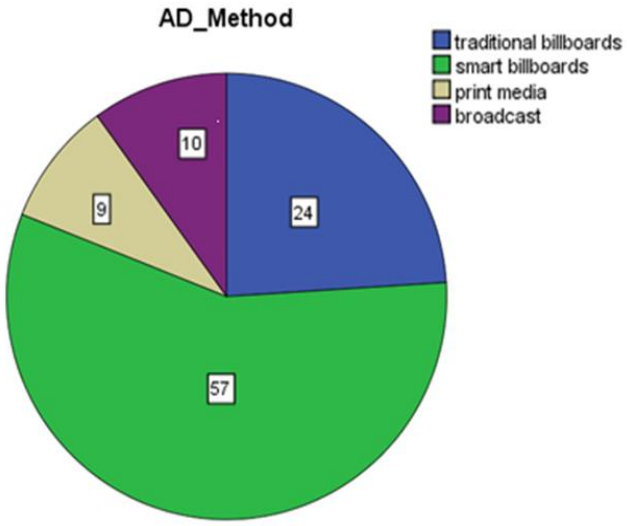


Figure 1: Pie Chart (Preferential distribution of Advertisement Method)

3.2. Designing the Pattern of Smart Billboard

The framework of our proposed research is shown in fig 3. In this system camera captures images that are stored in a storage file upon which face detection and feature analyzer API is applied. The results of this API are stored in database. After these results are fetched from database that can be viewed by stakeholders CB's and AC. Observing this CB's can give request for booking a time slot according to their

customer ratio related to their product. After processing the advertisement booking the advertisement will be shown on billboard.

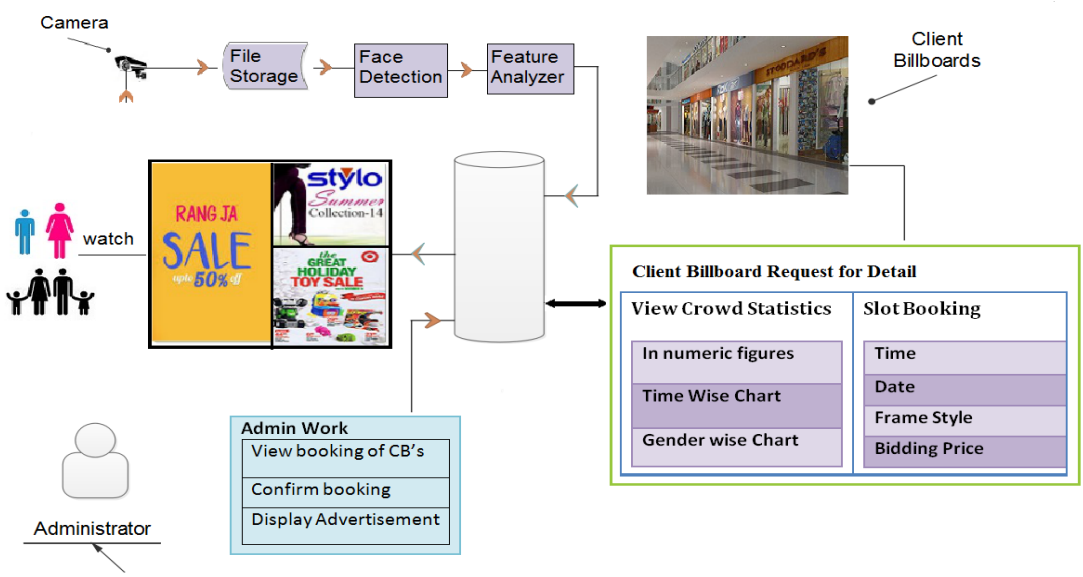


Figure 2: Framework of Smart Billboard

Figure 1 displays the general view of this research. Smart billboard runs advertisement for a specific limited time regarding customer behavior. In front an LCD has been displayed at mall entrance to run advertisement based on specific gender and age group of customers. The left side of this figure is portraying the control system where data about customers has been saved that is used by CB or AC for organizing advertisements.

Figure 3: Outlook of Research

3.2.1. Camera

There are (Closed Circuit Television) cameras in every shopping mall for security purpose. This system can use these existing CCTV cameras for face detection. This camera is capable of capturing video surveillance footage. This camera has IR facility of capturing video in low light and no light areas.

3.2.2. File Storage and Face Detection

This camera is attached with a personal computer. So that all captured images are stored in a storage file. After this apply face detection technique. For this purpose, use a Microsoft Azure face API. This estimates the number of persons and detects their faces in a crowd.

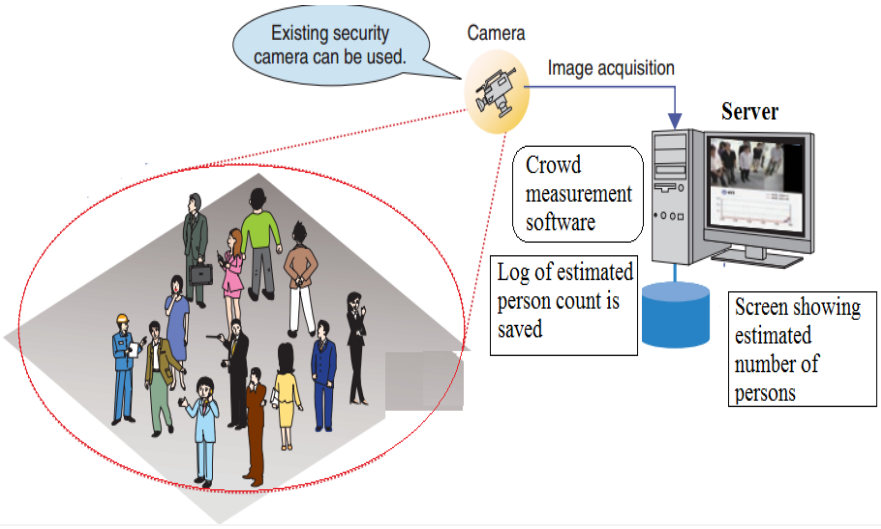


Figure 4: Crowd Measurement Technique

Figure 4 shows the procedure crowd measurement technique. In this technique the CCTV camera is capturing the images of incoming people to the mall and storing in personal computer storage file. API processes on the stored images and at the end the screen image of estimated results is shown.

3.2.3. Feature Analyzer:

The used API is capable of analyzing features of detected faces. This provides the facility to know the face attributes. The face attribute features available are: Age, Gender, Glasses and Smile.

Figure 5 shows the working of this feature analyzer component. API is processing on the images from file storage to scrutinize the feature based on age and gender and other attributes. At the end screen image of computer is displaying the estimated results of feature analyzer.

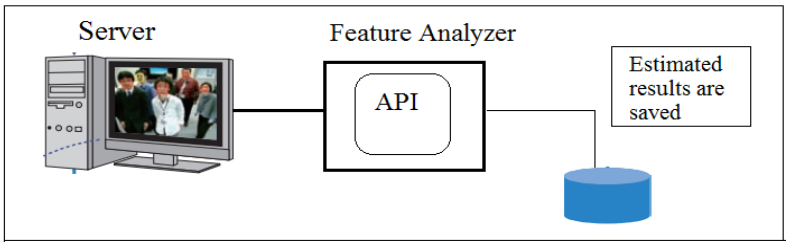


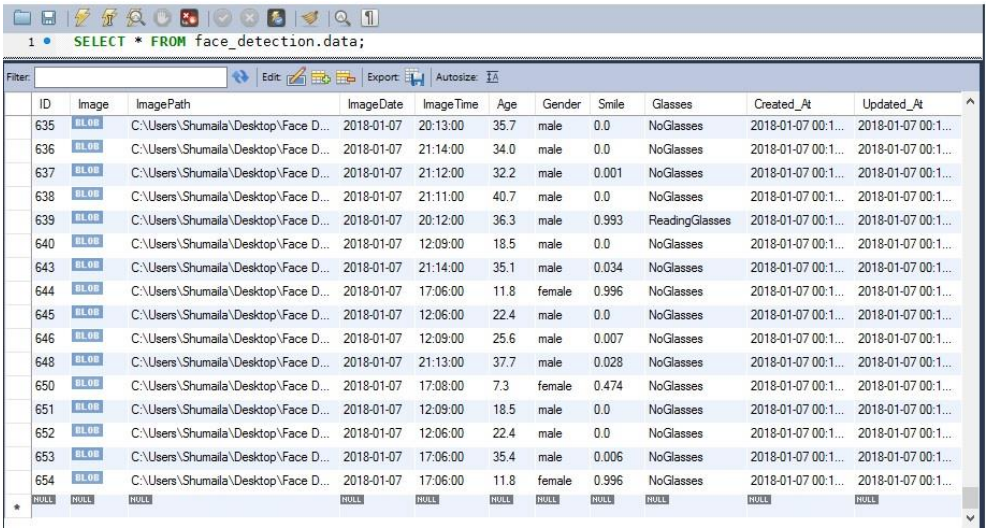
Figure 5: Face Identification

3.2.4. Database:

The processed data of API is saved in database. In this system database name is face_detection.db that has table titled as tbl_data. This table has columns that are ID, Image, ImagePath, ImageDate, ImageTime, Age, gender, Smile, glasses, Created_at, Updated_at. Image column is used to save images in BLOB (A Binary Large OBject) data type that is a collection of binary data stored as a single entity in a database management system. Blobs are typically images, audio or other multimedia objects, though sometimes binary executable code is stored as a BLOB.ImagePath column saves the storage information of image. ImageDate column stores the date at which a particular image was captured and stored in storage file. ImageTime stores the information of timing of images when it was captured form camera. Age column saves the age attribute of detected face (image) that is captured and shows its detail. Gender column saves the detail of gender attribute of captured images.

Figure 4 shows the view of table in which all detail of captured image is saved.

When run this task it will take automatically the images from this stored storage file and after applying the face detection and finding the feature attributes all data are saved in database



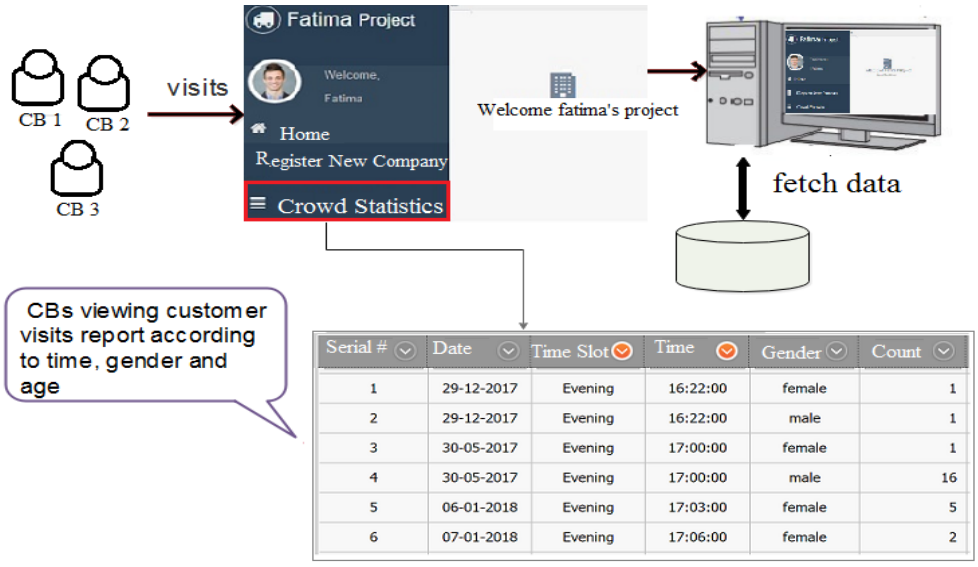
```
SELECT * FROM face_detection.data;
```

ID	Image	ImagePath	ImageDate	ImageTime	Age	Gender	Smile	Glasses	Created_At	Updated_At
635	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	20:13:00	35.7	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
636	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	21:14:00	34.0	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
637	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	21:12:00	32.2	male	0.001	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
638	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	21:11:00	40.7	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
639	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	20:12:00	36.3	male	0.993	ReadingGlasses	2018-01-07 00:1...	2018-01-07 00:1...
640	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	12:09:00	18.5	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
643	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	21:14:00	35.1	male	0.034	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
644	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	17:06:00	11.8	female	0.996	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
645	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	12:06:00	22.4	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
646	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	12:09:00	25.6	male	0.007	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
648	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	21:13:00	37.7	male	0.028	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
650	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	17:08:00	7.3	female	0.474	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
651	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	12:09:00	18.5	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
652	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	12:06:00	22.4	male	0.0	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
653	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	17:06:00	35.4	male	0.006	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
654	BLOB	C:\Users\Shumaila\Desktop\Face D...	2018-01-07	17:06:00	11.8	female	0.996	NoGlasses	2018-01-07 00:1...	2018-01-07 00:1...
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Figure 6: Overview of Database

- Website Visits:

The CBs (client billboards) visits the website to see detail about customers such as timing and other attributes of people along with gender and age. This system analyzes and shows the report in digits and chart form. This required information is fetched using SQL queries so that relevant data are accessed from database and CBs be able to watch the detail. So, they use this information to take decision regarding customer behavior based on gender and age groups. Figure 7 shows the detail of in which a website is maintain in ASP.NET



Serial #	Date	Time Slot	Time	Gender	Count
1	29-12-2017	Evening	16:22:00	female	1
2	29-12-2017	Evening	16:22:00	male	1
3	30-05-2017	Evening	17:00:00	female	1
4	30-05-2017	Evening	17:00:00	male	16
5	06-01-2018	Evening	17:03:00	female	5
6	07-01-2018	Evening	17:06:00	female	2

Figure 7: Website visit

3.2.5. Advertisement Booking

Figure 8 shows the Advertisement booking process. In this CBs visit the website and then according to detail of customer arrival (timing, gender and age information) they check the available slot. Then they select the required slot according to their product type. After this a new web page is available on which

they will enter their information and also their bid price for time slot. They also decide about the screen option either they want full or partial screen to run their advertisements and request for booking of that slot for their advertisement. Then our system checks the detail of all CBs and confirmation of booking is done of that CB (or CBs in case of partial screen selection by more than one CB) whose bid price is high. This system will be a facilitator study for customers, CBs and advertising agencies who will implement this system in malls.

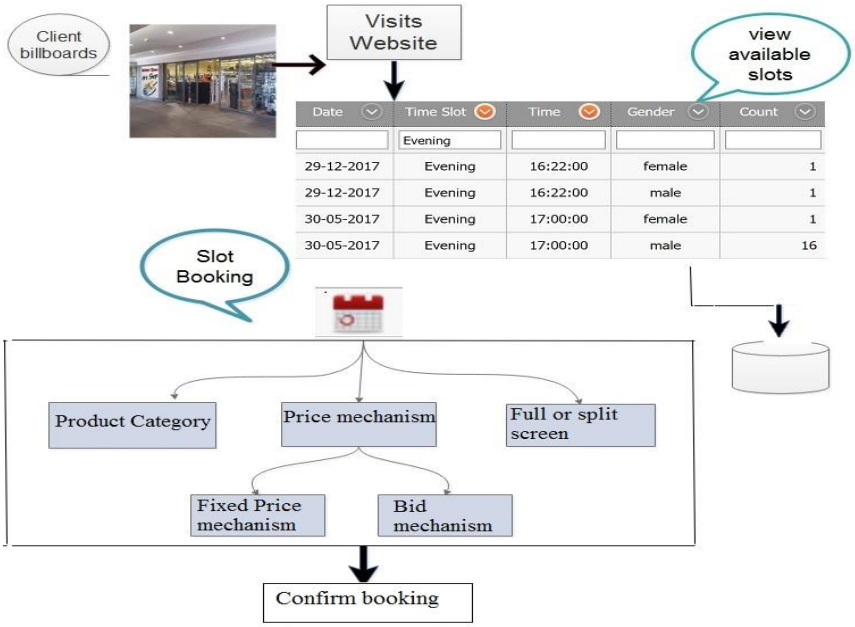


Figure 8: Overview of booking time slot

3.2.6. Confirmation of Booking and display Advertisement

Figure 9 shows that this system checks the detail of all CBs and administrator confirm the booking of that CB (or CBs in case of partial screen selection by more than one CB) whose bid price is high. This system will be a facilitator study for customers, CBs and advertising agencies who will implement this system in malls.

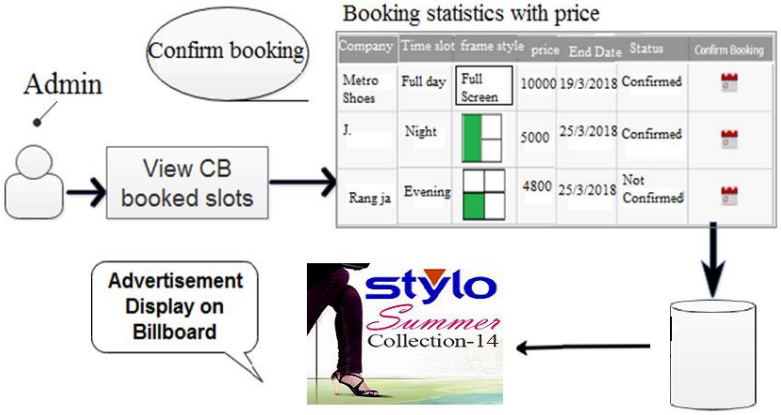


Figure 9: Booking Confirmation and Display of Advertisement

4. Conclusion:

This research proposes an advanced technology of billboard advertisement. A preliminary survey is conducted over a finite sample to come to know either proposed system is required or not. The proposed

system observes the customer behavior based on their gender and age group coming to mall and shows these results to CB's in statistical and charts form to enable them to decide to advertise their products according to customer behavior on SB.

Declaration of Interest Statement: It is declared that authors have no potential conflict of interest.

References

[1] Chen, Duan-Yu, and Kuan-Yi Lin. "A novel viewer counter for digital billboards." In *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 653-656. IEEE, 2009.

[2] Khalil, Gehan Abdelalim, Maysoon Mohamed Qutp, and Mai Ali Nada. "Intelligent Billboards Targeted Advertising Systems." *Journal of Art, Design and Music* 2, no. 2 (2023): 5.

[3] Sümer, Selay Ilgaz, Çağatay Berke Erdaş, Emre Sümer, and Ahmet Paker. "A New Tool for Personalized Advertising in Shopping Malls: A Smart Billboard System." *Archives of Current Research International* 22, no. 8 (2022): 46-59.

[4] Mohamed Naleer, Haju Mohamed. "Human face recognition to target commercial on digital display via gender." (2020).

[5] Shan, Gaoyang, and Byeong-Hee Roh. "Advertisement interval to minimize discovery time of whole BLE advertisers." *IEEE Access* 6 (2018): 17817-17825.

[6] Pourhossein, Fatemeh, and Abbas Mahmoudabadi. "Original Paper Smart Urban Street Advertising Pattern Using Internet of Things Based on Environmental and Traffic Conditions."

[7] Yang, Shuo, Ping Luo, Chen-Change Loy, and Xiaoou Tang. "Wider face: A face detection benchmark." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5525-5533. 2016.

[8] Ahmed, Haval A., Tarik A. Rashid, and Ahmed T. Sadiq. "Face behavior recognition through support vector machines." *International Journal of Advanced Computer Science and Applications* 7, no. 1 (2016).

[9] Liu, Te-Kai, Yun-Wu Huang, and Jen-Yao Chung. "Interactive wireless electronic billboard." In *IEEE International Conference on Networking, Sensing and Control, 2004*, vol. 1, pp. 553-558. IEEE, 2004.

[10] Nigam, Saral, Shikha Asthana, and Punit Gupta. "IoT based intelligent billboard using data mining." In *2016 international conference on innovation and challenges in cyber security (ICICCS-INBUSH)*, pp. 107-110. IEEE, 2016.

[11] Kinebuchi, Tetsuya, Hiroyuki Arai, Isao Miyagawa, Shingo Ando, Kaori Kataoka, and Hideki Koike. "Image processing techniques for measuring advertising effectiveness of digital signage." *R&D on Digital Signage as an Advertising Medium* (2010).

[12] Shlomot, Eyal. "Smart Electronic Roadside Billboard." U.S. Patent Application 13/586,016, filed March 7, 2013.

[13] Azure AI Vision, Discover computer vision insights from image and video analysis with OCR and AI. *https://azure.microsoft.com/en-us/services/cognitive-services/face/*

[14] Cook, Diane J., Juan C. Augusto, and Vikramaditya R. Jakkula. "Ambient intelligence: Technologies, applications, and opportunities." *Pervasive and mobile computing* 5, no. 4 (2009): 277-298

Machines and Algorithms
http://www.knovell.org/mna

Review Article

Techniques Leveraging from Artificial Intelligence for the Prediction of Climate Change in Pakistan: A Systematic Literature Review

Sameen Fatima[1,*] and Aniqa Dilawari[1]

[1]Department of Computer Science & Information Technology, University of Home Economics, Lahore, 54000, Pakistan

[*]Corresponding Author: Sameen Fatima. Email: sameen.fatima@hotmail.com

Received: 13 July 2023; Revised: 10 August 2023; Accepted: 25 August 2023; Published: 31 October 2023

AID: 002-03-000030

Abstract: Forecasting weather is crucial for sustaining life and ecosystems. Throughout history, farmers have predicted weather patterns to optimize crop growth. Weather prediction can be useful for other more crucial factors as well, including flood and drought predictions. In today's era of technological advancement, machine learning, and deep learning have made significant strides across various domains, including weather forecasting. For safeguarding crops and alarming farmers for the protection of their fields, it is necessary to have the ability to provide precise and timely weather predictions. While traditional weather prediction has faced advancements in IoT and machine learning, such as motion detection, speech recognition, and computer vision, it also offers more accurate environmental change predictions. In Pakistan, rising temperatures, erratic weather patterns, and changing precipitation pose significant challenges for the population. Climate change and the emission of harmful industrial gases are accountable for altering rainfall patterns, subsequently impacting economic growth since Pakistan's major economy relies on it. At present, numerous advancements in weather prediction utilize artificial intelligence techniques resulting in notable benchmarks for researchers. This study explores diverse methods of deep learning, machine learning, and advanced IoT devices to enhance weather prediction through a systematic literature review.

Keywords: Weather Forecasting; Machine Learning; Deep Learning; Climate Change; IoT in Agriculture;

1. Introduction

Nowadays every country around the globe is affected by the most pressing issue of Climate change. As an agricultural country highly vulnerable to climate variability, Pakistan faces numerous socio-economic, health, and environmental risks due to climate change. For accurate future climate pattern prediction, these risks need to be effectively adapted and mitigated. This research paper aims to explore the use of artificial intelligence (AI) in climate change prediction specifically tailored to the context of Pakistan.

With its vast geographical diversity, Pakistan experiences various climate patterns, ranging from the arid regions of Balochistan to the monsoonal areas of Punjab and Sindh. Predicting climate change in such a complex system requires robust and sophisticated models that can analyze vast amounts of data. AI, with its capacity to process big data and identify complex patterns, presents a powerful tool for climate change prediction in Pakistan.

The study of climate change in Pakistan dates back several decades. However, traditional climate models have often faced limitations in accurately predicting the region's diverse climatic conditions. AI, on the other hand, offers a new approach that has the potential to overcome these limitations. By analyzing large volumes of climate data, AI algorithms can identify non-linear patterns and relationships that may have been overlooked by traditional models.

The identification of various sources of data by AI is one of its key features. Climate models traditionally rely on meteorological data such as temperature, precipitation, and wind patterns. However, AI can incorporate data from diverse sources such as satellite imagery, oceanic currents, and socio-economic factors, providing a more comprehensive understanding of climate change dynamics. This multi-dimensional approach enables more accurate and robust predictions. Furthermore, AI is capable of learning from historical data to improve its predictive capabilities. Machine learning algorithms can analyze the historical climate data of Pakistan and identify trends, patterns, and relationships that can be used to forecast future climate scenarios. As the algorithm continues to learn and adapt, its predictions become more precise and reliable.

Some challenges need to be addressed for climate change prediction in Pakistan using AI. Quality and data availability is one of the main concerns. Pakistan's climate data infrastructure faces various limitations, including sparse and unreliable data. AI models heavily rely on data, and without robust and high-quality inputs, the accuracy of predictions may be compromised. Therefore, efforts should be made to improve data collection systems and ensure data accuracy and accessibility.

Section 2 discusses the related work of this research. In section 3 the research methodology is discussed which includes research questions formulation, research process, inclusion & exclusion criteria, and article quality assessment. Section 4 presents the findings of the research. Section 5 provides recommendations and discussions. This research article is concluded in section 6.

2. Related Work

Predicting climate change in Pakistan using different approaches that involve advanced technologies to analyze historical data, simulate future scenarios, and enhance understanding of climate patterns. Several research papers have delved into the application of AI for climate change predictions in Pakistan but a literature review on this particular topic is not available. This literature survey is the first systematic review of climate change prediction and forecasting using AI. Hussain et al. [1] did a comprehensive literature review on the topic of climate change impacts, adaptation, and mitigation of environmental and natural calamities in Pakistan. This study evaluated the effects of climate change in various sectors such as agriculture, livestock, forestry, food, water, and energy security in Pakistan. Although this research does not focus on AI-related techniques. Its overall finding was the societal impacts of irregular weather patterns and other climate change effects in Pakistan. It highlights the need for climate change mitigation and adaptation practices and techniques in Pakistan, considering their economic, social, and environmental aspects. The paper emphasizes the importance of governmental interference and strict accountability of resources for sustainable development and the formulation of state-of-the-art climate policies in Pakistan.

In another article [2], the impact of El-Nino/Southern Oscillation (ENSO) on summer monsoon rainfall in southern parts of Pakistan is studied to deeply understand the characteristics of summer rainfall variation and factors causing severe water shortages in already drought-prone regions. This paper did not utilize any AI-related technique for the prediction. Another review article, T. A. Khan et al. [3] provided details on the engineering-based techniques (such as dam construction) and non-engineering-based techniques (such as Artificial Intelligence-based algorithms), have been used to investigate flash floods.it concluded that swarm intelligence weights optimization for multi-layer perceptron neural network configuration performed the best among all forecasting approaches and is recommended for future enhancement. This study only focused on the problem of flash floods. Other related papers other paper discusses calamities such as crop yield prediction [4], agriculture stock [5], rainfall [12], landslide [13], stream outflow [14], water quality [15] etc.

Overall literature review is not available but individual climate-related problems are addressed. In summary, these research papers collectively contribute to the growing body of knowledge on climate change predictions in Pakistan using artificial intelligence. The studies highlight the diverse applications of AI, ranging from rainfall and heatwave predictions to drought and flood assessments, with a focus on improving the resilience of critical sectors such as agriculture and urban planning.

3. Method

This study has been undertaken as a systematic literature review and its goal is to assess state-of-the-art research in the field of climate change in Pakistan. The systematic literature review method's steps are outlined here.

3.1. Research Questions

This study focuses on the following research questions:

RQ1: What are the existing AI tools, techniques, algorithms and evaluation methods used in the prediction of climate change in Pakistan?

RQ2: What are the hazards caused by climate change in Pakistan?

RQ3: What are the proposed methodologies on the impact of climate change on health in Pakistan?

RQ4: What are the proposed methodologies/ research available on the impact of climate change on agriculture in Pakistan?

3.2. Search Process

IEEE (Institute of Electrical and Electronics Engineers), Science Direct, and ACM (Association for Computing Machinery) digital library are the search engines that are used in this research. These search engines were selected because of the following reasons:

- Well-known professional organizations that publish and provide access to a wide range of scientific paper
- Renowned organizations in the fields of electrical engineering, computer science, and related disciplines.
- These organizations are highly regarded and considered to be of high quality.
- These platforms provide access to papers from researchers and practitioners around the world who may have conducted studies on climate change prediction in Pakistan.
- Advanced search options, allowing you to specify your search criteria.

Table 1: Number of articles against research query in each search engine.

Search Engine	Research Query	Papers
IEEE Xplorer	((("Document Title": "climate change" OR "Document Title":"climate" OR "Document Title": "weather" OR "Document Title": "drought" OR "Document Title": crop OR "Document Title": temperature OR "Document Title": flood) AND ("Document Title": prediction OR "Document Title": forecast OR "Document Title": evaluat*) AND ("Document Title": Pakistan))OR (("Abstract":"climate change" OR "Abstract":"climate" OR Abstract: "weather" OR "Abstract": "drought" OR "Abstract": crop OR "Abstract":temperature OR "Abstract": flood) AND ("Abstract": prediction OR "Abstract": forecast OR "Abstract": evaluat*) AND ("Abstract": Pakistan)))	72

ACM	[[Abstract: "climate change"] OR [Abstract: or "climate" or "weather"]] AND [Abstract: pakistan] AND [E-Publication Date: Past 5 years]	17
Science direct	"climate change" "climate" "pakistan" "Artificial Intelligence"	4

3.2.1. Searching through AI-Powered Research Tools

We used some AI-powered search engines to find the answers to our queries. We used one of the most relevant papers, Hussain et al. [1] as a seed paper for SCISPACE (AI Powered search engine). The search gave us the most relevant papers and their summary which showed that the awareness about this issue of climate change is still low and emphasized the need for a literature survey to Grasp the full magnitude of these peculiar challenges and develop ingenious tactics for fostering sustainable progress in Pakistan. It was also seen that there is a lack of artificial-related approaches to solving the problem as no such papers were found in the search.

Furthermore, we also searched for the keywords, "climate change Pakistan artificial intelligence machine learning IoT deep learning" in the same search engine. It showed that Artificial intelligence (AI) and machine learning (ML) approaches have significantly contributed to the modeling and prediction of hydrological processes, climate change, and earth systems. In the realm of weather forecasting, the integration of AI and ML techniques, coupled with advanced IoT devices, has proven beneficial for more precise and effective weather predictions but It does not specifically mention the use of artificial intelligence, IoT, or climate change in Pakistan.

We also used another AI-powered tool, "researchrabbitapp" to see if there was any relevant paper which has been overlooked. We again used Hussain et al. [1] as our seed paper as this was the most relevant literature survey found. The graph (Figure 1 & Figure 2) shows all the similar papers but most of the work was either not from Pakistan or did not use any AI/ML technique.

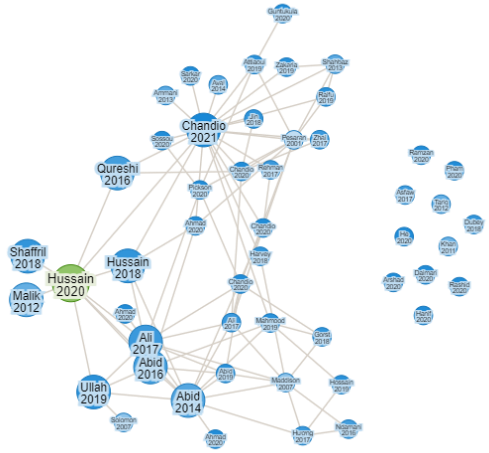


Figure 1: Similar work to the seed paper.

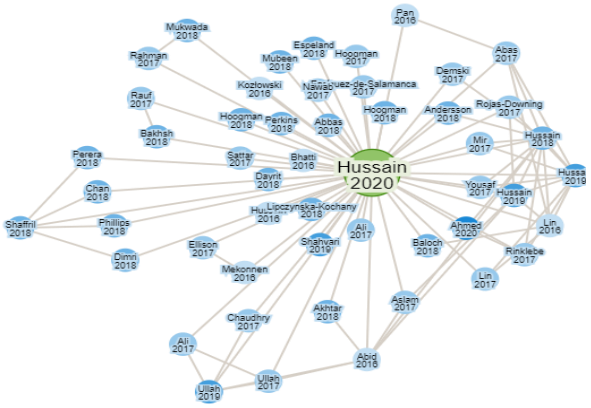


Figure 2: All reference articles of the seed paper

3.3. Inclusion and Exclusion Criteria

The query was designed to get all the results related to the topic therefore the keywords related to artificial intelligence, machine learning, etc were excluded. The research articles in the result of the query were manually checked by reading their titles and abstracts. The articles which were not discussed Artificial intelligence (AI) related techniques were excluded. The exclusion/ inclusion criteria are explained below:

1. This study captures the state-of-the-art work done in the last 5 years therefore the articles from 2018 to 2023 are included in this paper.
2. All articles that do not use AI, ML, DL, etc related techniques are excluded
3. Articles that have no relation with climate change, natural disasters such as floods, or the issues arising due to climate change are excluded
4. Research articles proposing solutions to better farming strategies, wheat price prediction, cotton & chili production, electricity load prediction, risk assessment, wind power generation, solar power production, and tea yield prediction are excluded
5. Papers that propose some forecasting models are included in this review

Papers selected after going through titles and abstracts. In IEEE Xplorer, out of 72, only 11 articles were found relevant to this study. The ACM showed a total of 17 articles from which only 7 were found relevant after reviewing their titles and abstracts. The search engine, science direct showed 4 articles from which 3 were selected.

4. Findings

4.1. Dataset Quality Assessment

In assessing the quality of the work, the first consideration involves whether a dataset is readily accessible. If not, alternative data collection methods are employed to gather the necessary information. These methods can include diverse data sources, such as scientific studies or sensor networks. The government's involvement in data provision underscores its commitment to facilitating research and decision-making processes.

Turning to the quality of the data, several factors come into play. The relevance of the data is crucial, ensuring that it aligns closely with the specific parameters being investigated. Accuracy is another pivotal factor, emphasizing the precision and correctness of the collected information. The assessment of data quality considers these factors collectively, recognizing the significance of robust data sources, relevance, accuracy, and integrity in the overall research framework. It can be seen in Table 2 that half of the studies used proper datasets for their work. Articles used data sources like government datasets including the Pakistan Meteorological Department while others collected data from sensors and IoT devices. Mostly was relevant to the research problem.

Table 2: Summary/ Evaluation of dataset quality and accessibility

Article	Dataset	Data Collection	Relevance	Accuracy	Tool	Transparency
(Hassan et al., 2021)	Yes	Statistical government publications by Pakistan Bureau of Statistic	yes	yes	R Studio	No specific details
(T. A. Khan et al., 2019)	Yes	Pakistan Meteorological Department (PMD)	Yes	Yes	No	No
(Munir, 2013)	No	data sources such as Shuttle Radar Topography Mission (SRTM) DEM, MODIS daily snow product MOD10A1, and spatially distributed NOAA precipitation data	Yes	Yes	No	No
(Bhutto et al., 2009)	No	Rainfall data from 16 stations, Nino 3.4 Index data from 1950 to 2000	Yes	Yes	No	No
(Mubashar et al., 2021)	Yes	Landslide Occurrence Data provided by NASA, TRMM, Shuttle Radar Topography Mission, and Global Land Data Assimilation System	Yes	Yes	No	No
(Kumari et al., 2023)	Yes	Humanitarian Data Exchange (HDX) and 55 years of rainfall data in open research repository	Yes	Yes	No	No
(Rahu et al., 2023)	No	Use of IoT devices to gather data	Yes	Yes	No	No
(Ayub et al., 2022)	Yes	Optical imagery obtained from the Sentinel-2 sensor	Yes	Yes	No	No
(F. Khan et al., 2015)	No	Collection of actual information regarding various barrages, river sizes, reservoirs,	Yes	Yes	No	No
(Ali et al., 2020)	No	Collection through sensors deployed at the riverbanks, sonar Sensor HC-SR04 and other relevant environmental parameters	Yes	Yes	No	No
(Gul et al., 2020)	No	Pakistan Meteorological Department (PMD)	Yes	Yes	No	No

(Saleem et al., 2021)	No	Available data sources and access to daily life observational patterns	Yes	-	No	No
(Jamshed et al., 2022)	Yes	GLOBAL FOREST WATCH platform	Yes	Yes	No	No
(N. Khan et al., 2020)	Yes	Pakistan Meteorological Department's gridded precipitation and temperature datasets (PGF datasets). National Centers for Environmental Prediction / National Center for Atmospheric Research (NCEP/NCAR) reanalysis datasets.	Yes	Yes	No	No

4.2. Methodology

The model selection process involves the careful consideration and choice of the appropriate mathematical or computational model to represent the system under study. The model can be further evaluated based on the use of historical data, its comparison with state-of-the-art models, sensitivity analysis, past Climate Behavior and Future Projections, and the availability of the features used in the process. The reviewed studies, Table 3, mostly use machine learning techniques including SVM and regression models and some of the models use artificial neural networks to solve the climate change problem. It was also observed that most of the studies used historical data to assess the performance of their model. Only one research paper [16], did sensitivity analysis which examines how variations or uncertainties in input parameters affect the model's output. Features refer to the variables or attributes used as input to the model. In the context of climate modeling, features could include various climatic parameters such as temperature, precipitation, atmospheric pressure, etc. The selection of relevant features is crucial for the model's accuracy, and it often involves domain expertise to determine which variables are most influential in representing the dynamics of the system under consideration. Most of the research articles explained the feature set used as input to the proposed model.

Table 3: Summary/ Evaluation of the proposed models

Article	**Model Selection**	**Model Evaluation**				**Features Provided**
		Historical	**Comparison**	**Sensitivity**	**Past/future implications**	
(Hassan et al., 2021)	Linear Regression Model	Yes	Root Mean Squared Error (RMSE), Linear Regression and Recurrent Neural Networks (RNNs)	No	Not specifically but selected factors depict past climate behavior	Yes
(T. A. Khan et al., 2019)	Support Vector Machine (SVM) and K-NN	Yes	Support Vector Machine (SVM) and K-NN classifiers	No	No	No

(Munir, 2013)	Snowmelt-Runoff Model (SRM)	Yes	Statistically, Nash-Sutcliffe Efficiency (NSE) R2 and average percentage volume difference Dv	No	Yes	Yes
(Bhutto et al., 2009)	No	Yes	16 stations' data of summer rainfall, sea surface temperature	No	No	No
(Mubashar et al., 2021)	Long Short Term Memory (LSTM) neural network	Yes	Artificial Neural Networks (ANNs)	No	No	Yes
(Kumari et al., 2023)	Regression-related machine learning algorithms	Yes	machine learning models, deep learning models, and statistical models	No	No	Yes
(Rahu et al., 2023)	regression models LSTM,, SVR, MLP, and NARNet	No	Compares the performance of different machine learning models	No	No	Yes
(Ayub et al., 2022)	Random Forest (RF)	Yes	Compared with other approaches in different regions	No	No	Yes
(F. Khan et al., 2015)	No	No	Other countries using wireless sensing techniques for flood prediction and monitoring	No	No	Data from sensor nodes
(Ali et al., 2020)	SP32 developmen t board	No	No	No	No	No
(Gul et al., 2020)	non-parametric statistical tests	Yes	on-parametric statistical tests, including the Mann-Kendall and Sen's slope	No	No	aridity index using precipitati on and reference evapotrans piration

(Saleem et al., 2021)	Water Evaluation and Planning system	No	Comparison with supply trends of housing society	Yes	Yes	Yes
(Jamshed et al., 2022)	long short-term memory network (LSTM)	Yes	root mean square error (RMSE)	No	No	Yes
(N. Khan et al., 2020)	Support Vector Machine (SVM) and Artificial Neural Networks (ANNs)	Yes	Support Vector Machine (SVM), Artificial Neural Networks (ANNs), and K-Nearest Neighbors (KNN)	No	No	Yes

4.3. Quality of the selected paper:

The credibility and impact of a research paper are often intricately tied to its publication details. Firstly, the journal name and its publication impact factor provide critical context regarding the paper's standing within the academic landscape, with a higher impact factor indicative of the journal's influence. Additionally, the peer-review process ensures scholarly rigor and authenticity, signifying that the paper has undergone critical evaluation by experts in the field before publication in a reputable journal. Assessing citations and references offers insights into the paper's scholarly influence and integration into academic conversations. Most of the research articles found on the desired problem were published in conferences and only a few journal articles were found. As the selected research articles were from the past 5 years therefore their number of citations was not huge.

Table 4: Summary of Research Articles Quality

Article	**Journal/ Conference**	**IF**	**No. of Citations**	**Article**	**Journal/ Conference**	**IF**	**No. of Citations**
(Hassan et al., 2021)	IEEE Consumer Electronics Magazine	-	3	(T. A. Khan et al., 2019)	IEEE 6th International Conference on Engineering Technologies	-	1
(Munir, 2013)	IEEE International Conference on Space Science and Communication	-	5	(Bhutto et al., 2009)	International Conference on Information Science and Engineering	-	5
(Mubashar et al., 2021)	International conference on	-	6	(Kumari et al., 2023)	International Conference on Computing,	-	0

	artificial intelligence				Mathematics and Engineering Technologies		
(Rahu et al., 2023)	IEEE Access,	-	3	(Ayub et al., 2022)	International Conference on Artificial Intelligence	-	4
(F. Khan et al., 2015)	IEEE SENSORS	-	17	(Ali et al., 2020)		-	7
(Gul et al., 2020)	2020 IEEE International Conference on Systems, Man, and Cybernetics	-	11	(Saleem et al., 2021)	Simulation journal	1.7	11
(Jamshed et al., 2022)	Physical Communication Journal	2.2	2	(N. Khan et al., 2020)	Advances in Water Resources	4.7	159

4.4. Transparency

The transparency of the article depicts how easy it is to replicate the work done. It was evaluated based on the availability of the provided dataset publicly and also the description of the code and tools used in the development of the proposed model. As it can be seen from Table 2 almost all research articles have not made their datasets, tools, and code available publicly.

5. Discussion

The research done on climate change about AI can be categorized in flood, landslide, crop yield prediction, river flow, and deforestation. Table 5 depicts the total number of papers against each category in the context of Pakistan. A limited number of research articles have addressed climate change-related issues, with a notable gap in addressing the particularly challenging problem of smog. Furthermore, there is a scarcity of research focusing on the application of artificial intelligence technologies in predicting climate-induced challenges such as landslides, deforestation, and river outflows. Efforts to explore and develop solutions in these critical areas remain relatively scarce.

Table 5: Categories of the Research Articles

Categories	**Total papers**	**Categories**	**Total Papers**
Flood	4	Landslide	1
Rainfall/drought	3	Water quality/health	2
Crops/yield prediction	2	Riverflow	1
Deforestation	1	Smog/ Air Quality	0

- RQ1: What are the existing AI tools, techniques, algorithms and evaluation methods used in the prediction of climate change in Pakistan?

AI plays a crucial role in climate change prediction. However, it is seen that the climate change-related problems in Pakistan are still a new field for researchers. In our survey following tools and techniques are mostly used:

1. Machine Learning Algorithms: Algorithms like Linear Regression, Random Forest, Support Vector Machines, and Neural Networks are used for climate modeling.
2. Remote Sensing: AI is employed to analyze satellite data and sensor data extracting valuable information for climate monitoring.
3. Deep Learning: Deep neural networks are used to solve complex problems.
4. Tools: Only one of the research articles mentioned the tool and framework they used in their research. The R studio was used.

Collaborative efforts and interdisciplinary approaches are crucial for comprehensive climate change studies in regions like Pakistan. Researchers can explore datasets and projects by organizations like NASA, and the Pakistan Meteorological Department or utilize platforms like Google Earth Engine. Moreover, the data that has already been collected, needs to be made public so that more innovative research can be performed for regions like Pakistan which are extremely affected by climate change.

- RQ2: What are the hazards caused by climate change?

Climate change has highly affected the countries like Pakistan. Climate change has been associated with various hazards in Pakistan, impacting the environment, economy, and overall well-being of the population. Extreme weather can cause floods, drought, glacier melting, crop failures, reduced river flow, the spread of diseases, deforestation, Air quality, and impact on energy production. The proposed artificial intelligence-related solutions are mostly on flood, rainfall, and drought. There is a need to invest in climate research and innovation to understand the merging climate risk and develop artificial intelligence-based predictive solutions. There is also a need to support the development of technologies and strategies for climate change adaptation and integration. Collaboration with international organizations/neighboring countries on climate change initiatives, data sharing, and joint research projects. The government of Pakistan should also facilitate researchers and scientists working on this particular challenge. Foster collaboration between government agencies, research institutions, and tech companies to share data and expertise. Encourage open-access platforms for sharing AI models and research findings.

- RQ3: What are the proposed methodologies on the impact of climate change on health in Pakistan?

In this survey, it was found that only a few researchers are working on the air quality and health issues arising due to climate change. There is a need to monitor air pollutants and analyze trends to link air quality variations with respiratory diseases. Evaluate water quality data, correlate it with climate variables, and study the impact on waterborne diseases, emphasizing preventive strategies.

- RQ4: What are the proposed methodologies/ research available on the impact of climate change on agriculture in Pakistan?

Research articles particularly focusing on agriculture-related problems due to climate change in Pakistan are very few. Only two articles were available which proposed some AI/ML techniques for this problem. They used a supervised machine learning framework that utilizes remote sensing data and machine learning algorithms to estimate wheat area and predict wheat yield and a Random Forest (RF) machine learning algorithm for wheat area estimation and generation of Land Use Land Cover maps. The other paper used Linear Regression Mode to forecast next year's crop consumption, enabling farmers to make informed decisions about their crops.

6. Conclusion

In conclusion, climate change prediction in Pakistan using AI presents a transformative opportunity to enhance our understanding of the region's complex climatic dynamics. By harnessing the power of AI, we can overcome the limitations of traditional climate models and generate more accurate, timely, and comprehensive climate change predictions. However, addressing challenges such as data availability and technical capacity is crucial to unlock the full potential of AI in climate change prediction. This research paper aims to delve deeper into these issues and explore the practical applications of AI for climate change prediction in Pakistan. Through this research, we contribute to the growing body of knowledge on climate

change adaptation and mitigation in the context of Pakistan and also highlight the issue of unavailability of data and research on the latest AI techniques.

References

[1] Hussain, Mudassar, Abdul Rahman Butt, Faiza Uzma, Rafay Ahmed, Samina Irshad, Abdul Rehman, and Balal Yousaf. "A comprehensive review of climate change impacts, adaptation, and mitigation on environmental and natural calamities in Pakistan." *Environmental monitoring and assessment* 192 (2020): 1-20.

[2] Bhutto, Abdul, Ming Wei, Yan-an Liu, and Nan Li. "Impact of ENSO on summer monsoon in southern parts of Pakistan." In *2009 First International Conference on Information Science and Engineering*, pp. 4903-4906. IEEE, 2009.

[3] Khan, Talha Ahmed, Muhammad Mansoor Alam, Zeeshan Shahid, and Mazliham Mohd Su'Ud. "Investigation of flash floods on early basis: A factual comprehensive review." *IEEE Access* 8 (2020): 19364-19380.

[4] Ayub, Maheen, Najeed Ahmed Khan, and Rana Zeeshan Haider. "Wheat Crop Field and Yield Prediction using Remote Sensing and Machine Learning." In *2022 2nd International Conference on Artificial Intelligence (ICAI)*, pp. 158-164. IEEE, 2022.

[5] Hassan, Najam Ul, Farrukh Zeeshan Khan, Hafsa Bibi, Nokhaiz Tariq Khan, Anand Nayyar, and Muhammad Bilal. "A decision support benchmark for forecasting the consumption of agriculture stocks." *IEEE Consumer Electronics Magazine* 10, no. 6 (2021): 45-52.

[6] Ali, Syed Ahmed, Fasih Ashfaq, Ehsan Nisar, Usama Azmat, and Jehan Zeb. "A prototype for flood warning and management system using mobile networks." In *2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pp. 326-331. IEEE, 2020.

[7] Jamshed, Muhammad Ali, Charalambos Theodorou, Tahera Kalsoom, Nadeem Anjum, Qammer H. Abbasi, and Masood Ur-Rehman. "Intelligent computing based forecasting of deforestation using fire alerts: A deep learning approach." *Physical Communication* 55 (2022): 101941.

[8] Khan, Feeza, Saira Memon, Imran Ali Jokhio, and Sana Hoor Jokhio. "Wireless sensor network based flood/drought forecasting system." In *2015 IEEE SENSORS*, pp. 1-4. IEEE, 2015.

[9] Khan, Najeebullah, D. A. Sachindra, Shamsuddin Shahid, Kamal Ahmed, Mohammed Sanusi Shiru, and Nadeem Nawaz. "Prediction of droughts over Pakistan using machine learning algorithms." *Advances in Water Resources* 139 (2020): 103562.

[10] Khan, Talha Ahmed, Muhammad Alam, Syed Faiz Ahmed, Zeeshan Shahid, and M. S. Mazliham. "A factual flash flood evaluation using SVM and K-NN." In *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, pp. 1-6. IEEE, 2019.

[11] Gul, Sajid, Jingli Ren, Yunlong Zhu, and Neal N. Xiong. "A systematic scheme for non-parametric spatiotemporal trend analysis about aridity index." In *2020 IEEE international conference on systems, man, and cybernetics (SMC)*, pp. 981-986. IEEE, 2020.

[12] Kumari, Sapna, Muhammad Owais Raza, and Arsha Kumari. "Performance Evaluation Of Machine Learning Algorithms For Rainfall Prediction Using Dimensionality Reduction Techniques." In *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1-6. IEEE, 2023.

[13] Mubashar, Mehreen, Gul Muhammad Khan, and Ramla Khan. "Landslide prediction using long short term memory (LSTM) neural network on time series data in Pakistan." In *2021 International conference on artificial intelligence (ICAI)*, pp. 175-181. IEEE, 2021.

[14] Munir, Muhammad Badar. "Climate change impact on flow discharge of Neelum river catchment using snowmelt runoff model." In *2013 IEEE International Conference on Space Science and Communication (IconSpace)*, pp. 350-355. IEEE, 2013.

[15] Rahu, Mushtaque Ahmed, Abdul Fattah Chandio, Khursheed Aurangzeb, Sarang Karim, Musaed Alhussein, and Muhammad Shahid Anwar. "Towards design of Internet of Things and machine learning-enabled frameworks for analysis and prediction of water quality." *IEEE Access* (2023).

[16] Saleem, Arfa, Imran Mahmood, Hessam Sarjoughian, Hasan Arshad Nasir, and Asad Waqar Malik. "A Water Evaluation and Planning-based framework for the long-term prediction of urban water demand and supply." *Simulation* 97, no. 5 (2021): 323-345.