International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 7, 2025
https://www.theaspd.com/ijes.php

Intelligent Question Classification Through Machine Learning Techniques

Prashant Y Niranjan[1], Vijay S Rajpurohit[2], Ramesh Medar[3], Ranjana Battur[4], Ravi Kalkundri [5], Sagar S Talagatti[6], Sanjay Lote[7]
[1,2,4,5,6] KLS Gogte Institute of Technology, Belagavi, India
[3]Rajarambapu Institute of Technology, Ishwarpur
[7]Government Polytechnic Rabakavi –Banahatti, Karnataka.
prashant053@gmail.com[1], vijaysr2k@yahoo.com[2], rameshcs.git@gmail.com[3], rbbattur@git.edu[4], kalkundri.ravi05@gmail.com[5], sagartalagatti@gmail.com[6], sanjaylote2010@gmail.com[7]

Abstract
The agricultural sector plays a vital role in economic development and depends heavily on timely, accurate, and relevant information for effective decision-making. Farmers, agricultural scientists, and other stakeholders often require quick and reliable answers to queries related to crop management, soil health, pest control, weather conditions, and sustainable farming practices. In recent years, Question-Answering (QA) systems have gained significant attention as a promising technological solution to address this growing demand. These systems allow users to pose questions in natural language and receive precise, context-aware responses.
This research paper presents the conceptualization and development of a QA system specifically tailored for the agricultural domain. The proposed system harnesses the power of Natural Language Processing (NLP) and Machine Learning (ML) to classify questions based on their type. Accurate question classification is essential for retrieving relevant answers and improving the overall performance of the QA system. Several ML algorithms are explored implemented, and compared to determine the most effective model for this task.
Keywords: *Natural Language Processing Question Answering, Human Language Agricultural Practices Agricultural Sector*

INTRODUCTION (10 PT)

Objective of QAS to produce pertinent answers to a user's question [1]. The system must be able to understand the intention of the user, the context of the question, and the underlying meaning of the text. There are various reasons for planning and developing an agricultural quality assurance system. Some of these reasons include: 24 Access to information: Agriculture is an information-intensive field and access to exact information is essential to making informed decisions about crop management, pest control and soil fertility. QAS can provide quick and easy access to this data, which helps researchers; farmers and other stakeholders make knowledgeable decisions. Efficiency: QAS can provide fast access to data that can save time and resources, especially for farmers who do not have the resources to access and analyze large amounts of data. This efficiency can help increase the productivity of the agricultural sector [2]. 26 Accessibility: QAS can be designed to be accessible and usable even by people with limited technical knowledge. It ensures the relevant information is available to all agricultural stakeholders, including smallholders who may be located in remote or rural areas. 27 Integration with other technologies: QAS can be integrated with other technologies [3] such as mobile applications and precision farming tools to provide a more broadly to the challenges facing the agricultural sector. 28 Opportunities to address global food security: By 2050, the world population is hoping to reach 9.7 billion and the demand for food will increase by 50%. The agricultural industry needs to be more productive and efficient to meet this demand, and a quality assurance system can help address some of the challenges facing the industry [4]. QAS system uses machine learning techniques to gather and explore data from various sources such as agricultural databases, scientific publications and social media. In the previous work [6], The importance of automated question pairings has grown significantly, aiming to reduce the effort and time needed for manual question generation [5]. A method has been devised to automatically generate question-answer pairs for patient education, enabling the transformation of various patient education videos into corresponding question-answer pairs with accompanying streaming video.

In [7] When responding to a consumer's query[15], the program will identify an appropriate response. However, the computation time has tripled because the software must wait for the same set of questions to be submitted by multiple consumers. This is necessary as the software needs to offer simultaneous answers to multiple consumers on a regular basis. To address this, the 'PLANE' method has been developed to temporarily prioritize consumers selected first from the relevant query group. When a consumer poses multiple questions, the software will present responses in a hierarchical format, recommending the most relevant assessment response first. In [8], The objective of an Intelligent Question-Answering framework is to deliver precise responses to textual queries. Such frameworks heavily rely on text analysis, which often comes with a significant computational load [9]. Recently, Knowledge Graphs have emerged as an effective approach for these frameworks.

RESEARCH METHOD (10 PT)

A query processing module that finds the query type and the expected response type for certain semantic queries [10]. The proposed research is presented as follows A feature-aware gradient boosting model is proposed for the query processing module [11]. The model proposed is trained using the Kisan call center dataset, based on which the model predicts the expected reply to a specific query asked by farmer. In addition, the feature-aware gradient boosting model carrying out for query type detection is investigated. Here total of three categories such as vegetables, fruits and flowers are discussed. The classification task [12] is performed using various standard machine learning models. SGD achieves 88.03% perfectioness in comparison with other models and the proposed feature-aware gradient boosting model 94.1% correctness.

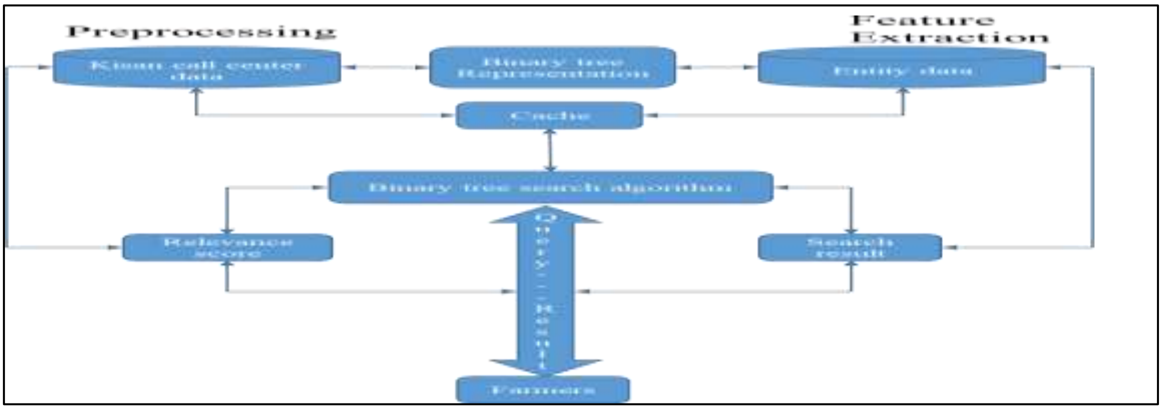

Figure 1. Architecture of QA System

A cutting-edge model known as the FAGB has undergone rigorous training using the Kisan call center dataset. This innovative model, named the Feature-Aware Gradient Boosting model outlined to cater to the unique information requirements [13] of users currently facing with daunting task of sifting through copious load of facts and information to facilitate well-informed decisions, improve productivity, and curtail expenses associated with crop cultivation.

Expanding on this, the Feature-Aware Gradient Boosting model has been meticulously crafted to recognize and prioritize the key features and attributes within the vast agricultural data [14] landscape. This feature- awareness is instrumental in streamlining the decision-making process for farmers. By discerning the most pertinent information, the model empowers farmers to make more accurate and timely choices, be it in crop selection, pest and disease management, irrigation scheduling, or soil nutrient optimization.

Furthermore, the model's potential extends beyond data-driven insights. It holds the promise of reducing the cognitive burden on farmers by providing interpretable and actionable recommendations. For example, it can suggest [15] optimal planting dates based on historical weather data or offer tailored advice on fertilizer application, thereby directly contributing to increased yields and cost savings.

In a broader context, this pioneering model represents a significant step forward in the agricultural technology landscape, bridging the gap between data-driven decision-making [16] and the practical requirements of stakeholders. Its adaptability and capability for blending into being farming practices position it as a valuable tool in the modern agricultural arsenal, helping farmers to not only navigate the sea of data but also thrive in an ever-evolving agricultural environment.

In the era of modern farming, the agricultural landscape [17] is filled with a plenty of data and information resources, making it an information-rich yet daunting realm for farmers to navigate. The sheer complexity and volume of available information [18] can often prove overwhelming. This is where Question and Answer Systems (QAS) emerge as indispensable tools. Custom-tailored to the specific needs of the agricultural sector, these QAS provide farmers with a good interface through which users pose the questions [19] and receive information finely tuned to their unique requirements. Farmers confront a multitude of challenges, including limited access to vital information, the intricacies of comprehending technical details, and resource constraints. Hence, the development of an effective agricultural [20] QAS should be rooted in the ambition to alleviate these challenges, ensuring that farmers can effortlessly access the resources they need.

Feature-Aware Gradient Boosting model specifically tailored for the agricultural Q&A system. We evaluate its performance based on accuracy, recall, and the F1 score.

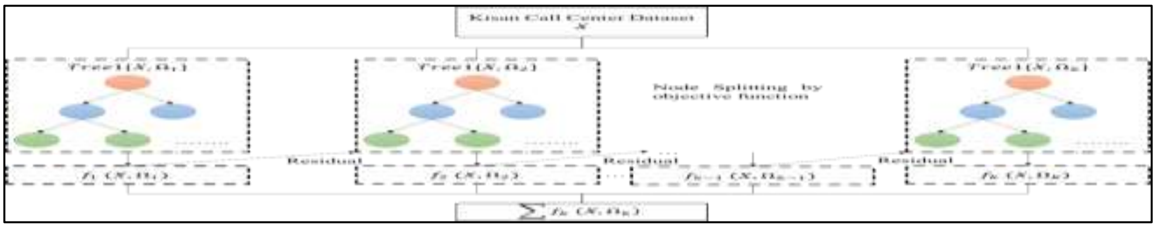

Figure 2. Boosting Method

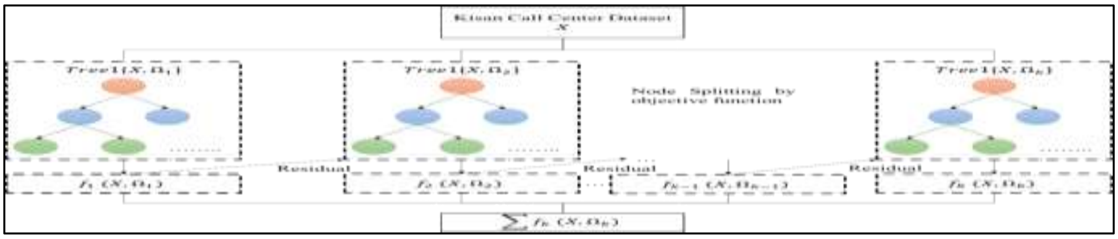

Figure 3. FAGB Boosting Model

XGBoost, ML method stands out as an ideal choice for agricultural question-answering systems [21]. This powerful tool that effectively combines several decision trees to make precise predictions. What sets XGBoost apart is its remarkable [22] aptitude for handling extensive and intricate datasets commonly found in agriculture. Whether dealing with numerical or categorical data, XGBoost demonstrates remarkable efficiency. Additionally, one of its standout features that provides key relevance insights, which allows users to discern the critical factors impacting crop production.

By integrating the model into an agricultural question-and-answer system, stakeholders can gain real-time retrieval of data [23] concerning crops, weather situations, and other elementswhich influences productivity of agricultural. This invaluable information empowers farmers and other industry players to make proper decisions regarding crop selection, cultivation practices. Consequently, this guides to the taking on to more fruitful farming practices, enhancing the overall health and prosperity of the farming sector.

Furthermore, the FAGB model [24] builds upon the strengths of XGBoost. It introduces a feature-aware gradient enhancement model, a novel approach to identifying and prioritizing the key element within an agricultural dataset. The model accounts for the complexities of data dimensions in the farming domain.

To delve into the technical details, let's consider the equation:

$E = \{(a_1, b_1), (a_2, b_2), \ldots, (a_m, b_m)\}$……………………………………………………………………… (2.1)

In this equation, 'j' represents the individual samples [25] within the dataset. For each sample 'j,' the corresponding 'b_j' denotes the binary class identifier, which could be {-1, 1}, reflecting the outcome for thatspecific sample. The n-dimensional vector 'a_j' encapsulates the observable properties of sample 'j.' Given the multitude of features, a dataset can become quite complex. When the sample dataset size is 'm,' which might contain fewer samples differentiating to the full dataset 'n,' accurately predicting the outcomes becomes challenging.

To direct this challenge, a predictive model 'G ' is constructed and analyzed using the below equation: $g: A \rightarrow B$…………………………………………………………………(2.2)

In this equation, 'A' symbolizes the independent characteristics within sample 'j,' while 'B' represents the predicted outcome. By leveraging this feature-aware gradient enhancement model, it becomes possible to make precise predictions and understand the crucial attributes influencing agricultural decisions. This, in turn, empowers farmers and stakeholders to make data-informed choices for better farming practices.

XGBoost classification algorithm creates a prediction model, with the objective of minimizing a specified objective function. The XGBoost technique, a widely adopted Gradient-Tree-Boosting (GTB) method,

is instrumental in addressing various classification problems. This is detailed by the following equation:
$\hat{Z}_j = G(Y_j) = \sum_{l=1}^{L} g_l(Y_j)$, where $g_l \in \alpha$ ....(2.3)
In the above equation:
Y_j constitute the 'j^{th}' sample in the dataset.
'L' signifies the size of the tree in classification process
$\hat{Z}_j$ reflects the classification outcome of the multi-label classification model with various dimensions.
The 'l^{th}' dimension indicates the probability that a specific class be included in the 'l^{th}' class.
α denotes the set of different decision trees for the given equation: $\alpha = \{g(y) = x_{t(y)}\}$............. (2.4)
In this context, all individual trees '$g(y)$' are consistent with one another concerning leaf mass 'x' and structural parameter 't.' The XGBoost classification model endeavors to minimize the loss parameter defined by the equation below:
$M(G) = \sum_j m(\hat{z}_j, z_j) \sum_l \beta(g_l)$ .... (2.5)
Here, '$\hat{z}_j$' represents to the predicted values, 'z_j' signifies the actual labels, and '$\beta(g_l)$' is connected with the individual classifiers. The objective is to minimize this loss parameter, resulting in an optimized classification model that can effectively categorize data into the desired classes. This approach, rooted in the principles of ensemble learning, is a robust framework for enhancing the accuracy and performance of agricultural question- answering systems, ultimately benefiting farmers and agricultural stakeholders in their process of taking a decision.
Equation (2.5) introduces the penalty term '$\beta(g_l)$,' which is represents the constants 'δ' and 'μ,' controlling the intricacy of the calculation. The 'U' parameter signifies the leaf size within the tree. This penalty term gains a important a role in the optimization process.
In the FAGB model, a weight functions given to train the data 'y,' with 'm' defining the selected sample. The model employs the negative log-likelihood loss function, represented by equation (2.7), to calculate the loss: $m(\hat{z}_j, z_j) = -\sum_k z(k) \log \tilde{z}(m) = -\log \tilde{z}(m)$ .... (2.7)
In this equation:
- '$z(k)$' denotes the 'k^{th}' feature of the sample.
- '$\tilde{z}(m)$' represents the '$\tilde{z}k^{th}$' dimensions of the sample's result.

An iterative optimization process is applied to minimize the loss function for a given iteration 'u,' as described in equation (3.8):
$M^j = \sum_{j=1}^{o} m(\hat{z}_j^{(u-1)}\ g_u(y_j), z_j)\ \beta(g_u)$ .... (2.8)
In equation (2.8), 'g_u' is determined through an iterative process that greedily minimizes the loss. To further minimize this loss, equation (3.9) is used:
$M^u \cong \sum_{j=1}^{o} [m(\hat{z}_j^{(u-1)}\ z_j h_j g_j(y_j)\ 1/2\ i_j g_u^2(y_j)]\ \beta(g_u)$ .... (2.9)
In this equation:
- 'h_j' signifies the first gradient term of $m(\hat{z}_j^{(u-1)}\ z_j)$.
- 'i_j' represents the order of the second gradient $(\hat{z}_j^{(u-1)} z_j)$.

Equation (2.9) demonstrates the process for minimizing the tree 'g_u.' It's important to note that the choosing of features that impacts classification performance when using the standard XGBoost (XGB) algorithm.
In contrast, logistic regression (LR), as a linear method [26], given to classify and predict datasets. This predictive modeling method relies on regression and is grounded in probability. LR is often employed for binary data to determine outcomes based on one or more attributes. The choice of cost functions vary based on the regularization method employed, including L1 and L2 standard and elastic network regularization [28]. The regularization strength is defined by the regularization factor 'C,'
Additionally, the Multinomial NB classification algorithm is widely used in text classification applications. It leverages Bayes' theorem to estimate the probability that a sample exhibits properties associated with a particular class.
Support Vector Machine (SVM) is a versatile supervised learning technique that is effective for solving both classification and regression problems. SVM maps data points from a lower-dimen-sional space to a higher- dimensional space, enabling them to be linearly separated. A hyperplane is then built to provide as a classification boundary, effectively separating data points into different classes. SVM is known for its robustness and ability to handle complex data, making it a popular choice for various machine learning tasks. SVM objective function, considering 'n' data points, is defined by equation (2.13). It encompasses both normalization and a loss function. The primary hyperparameters governing SVM models are the normalization vector (ω) and the regularization factor (C). There is a delicate trade-off between penalties for inaccurate predictions and the margin of discrimination, which is regulated by

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 7, 2025
https://www.theaspd.com/ijes.php

the internal parameter 'C' initialized during SVM training. Additionally, SVM employs a kernel function 'f(x)' to assess the degree of similarity between pairs of data points ('x_i' and 'x_j') selected from a range of kernels. Common kernel types include the sigmoid, polynomial, linear, and radial basis function (RBF), with the gamma coefficient acting as a hyperparameter in the sigmoid, RBF, and polynomial kernels. The gamma parameter [29] determines the curvature of the decision boundaries.

Random forest is a favoured technique in machine learning for addressing classification and regression problems. It operates as an ensemble method, employing the outputs of multiple decision trees to reach a final prediction.

Decision trees are a versatile ML algorithm applicable to both regression and classification tasks. The model's structure forms a tree-like hierarchy, where each internal node represents a decision based on a function, each branch symbolizes a possible value of that function, and each leaf node signifies an expected outcome.

Equation (2.15) provides the mathematical representation of a decision tree
model: y = f(x)… ………………………………………………………………… (2.15)

Here, 'y' denotes the predicted result for a given data sample 'x.' A decision tree structure can be mapped into a series of if-else statements that collectively form a function, making it a valuable tool for solving various machine learning problems.

AdaBoost, an adaptive boosting technique, is a machine learning approach that leverages several weak classifiers and amalgamates them into a single robust classifier. AdaBoost is centered around the concept of training weak classifiers, each of which is trained on a modified version of the training dataset, with increased emphasis on samples that were significantly misclassified by previous weak classifiers. The process involves iteratively training a set of feeble classifiers, with each feeble classifier operating on a dataset containing more wrongly classified examples than the previous dataset.

SGD classifier, utilizes a one-to-all approach to combine multiple binary classifiers. SGD has gained popularity for its access to all data in each iteration, making it suitable for large datasets. Its underlying principle shares similarities with regression methods, making it relatively easy to comprehend and implement in practical applications. Regularization in the context of the loss function serves as a penalty that encourages model parameters to approach zero, with regularization [30] options such as L2 (squared Euclidean norm), L1 (absolute norm), or a combination of the two (elastic). One crucial hyperparameter in SGD is the regularization factor 'alpha,' which controls the extent of regularization and influences the learning rate. Properly configuring this hyperparameter is essential for obtaining reliable results.

The Kisan-Call-Center (KCC) database is a valuable resource for agricultural question-answering systems, collecting a set of questions and answers related to agriculture, soil management, disease and pest prevention, climate change, and government agricultural initiatives. It encompasses approximately 5,000 questions and answers in various Indian languages and developing algorithms for the agricultural sector. Researchers and developers exploring question-answering system technology [27]in agriculture can access this database from the Open Government Data policysupported by the Government of India, making it a critical resource for advancing agricultural technology and providing farmers support in the country.

Season	Sector	Category	Crop	QueryTyp	QueryText	KccAns
RABI	AGRICULT	Others	Others	Weather	asking about Tiruvannamalai weather report today	Recommended for Tiruvannamalai weather report today have very light ra
RABI	ANIMAL H	Animal	Pig	Animal Br	asking about need for pig farming training	recommended for contact with veterinary department in tiruvannamalai, p
RABI	AGRICULT	Cereals	Paddy (Dh	Sowing Ti	Asking about Paddy navarai season of planting and var	Recommended for Navarai(Dec -Jan)ADT 36, ADT 39, ASD 16, ASD 18, MDU 5
RABI	AGRICULT	Others	Others	Weather	asking about tiruvannamalai weather report	recommended for today have light rain
RABI	AGRICULT	Others	Others	Weather	Asking about Weather report for Thiruvannamalai	Recommended for today have light rainfall.
RABI	AGRICULT	Others	Others	Governme	asking about Drought relief fund information	Recommended contact to block level agriculture office
RABI	AGRICULT	Cereals	Paddy (Dh	Crop Insu	Asking about Pradhan Mantri Fasal Bima Yojana insura	Recommended for Pradhan Mantri Fasal Bima Yojana insurance going on N
RABI	AGRICULT	Others	Others	Weather	asking about Thiruvannamalai today weather Report	Recommended for Thiruvannamalai today weather Report : Light Rain fall (
RABI	AGRICULT	Cereals	Paddy (Dh	Varieties	Asking about paddy varieties	recommended for ADT 36, ADT 39, ASD 16, ASD 18, MDU 5, CO 47,CORH 3, A
RABI	AGRICULT	Cereals	Paddy (Dh	Varieties	Asking about paddy varieties	recommended for ADT 36, ADT 39, ASD 16, ASD 18, MDU 5, CO 47,CORH 3, A
RABI	AGRICULT	Cereals	Paddy (Dh	Cultural P	asking about navarai paddy varieties	recommended for ADT 36, ADT 39, ASD 16, ASD 18, MDU 5, CO 47,CORH 3, A
RABI	AGRICULT	Oilseeds	Groundnu	Plant Prot	Asking about root rot management in Groundnut	Recommended for pseudomonas fluoroscens 1 kg mixed with 20 kg of FY
NA	AGRICULT	Others	Others	Weather	Asking about Weather report for Thiruvannamalai	Recommended for today have light rainfall.
RABI	AGRICULT	Cereals	Paddy (Dh	Varieties	Asking about suitable paddy varieties for Navarai seas	Recommended for paddy varieties ADT 36, ADT 39, ASD 16, ASD 18, MDU 5,
NA	AGRICULT	Others	Others	Weather	Asking about Weather report for Thiruvannamalai	Recommended for today have light rainfall.
RABI	AGRICULT	Others	Others	Weather	asking about weather detail in Tiruvannamalai	recommended for have moderate rain in Tiruvannamalai
RABI	AGRICULT	Others	Others	Weather	asking about Thiruvannamalai today weather Report	Recommended for Thiruvannamalai today weather Report : moderate Rain
RABI	AGRICULT	Others	Others	Weather	Asking about Weather report for Thiruvannamalai dist	Recommended for Today and tomorrow having cloudy weather and Possib
RABI	AGRICULT	Cereals	Paddy (Dh	Governme	Asking about Agriculture department phone number	Recommended for phone number:04173 â€" 223477
RABI	HORTICUL	Vegetable	Leafy Veg	Cultural P	Asking about leafy vegetables bio fertilizers applicatio	Recommended for apply Pseudomonas fluorescens 1 kg or Trichoderma vi
RABI	HORTICUL	Vegetable	Drum Stic	Seeds and	Asking about planting season for Drum stick	Recommended for season July â€" October

Figure 4. Data set of Kisan call center

RESULTS AND DISCUSSION (10 PT)

Performance

TruePositive

TruePositive + FalseNegative

Below equation determines the model's F1 − score.

Recall × Precison F1 − Score = 2 × Recall + Precision

Accuracy

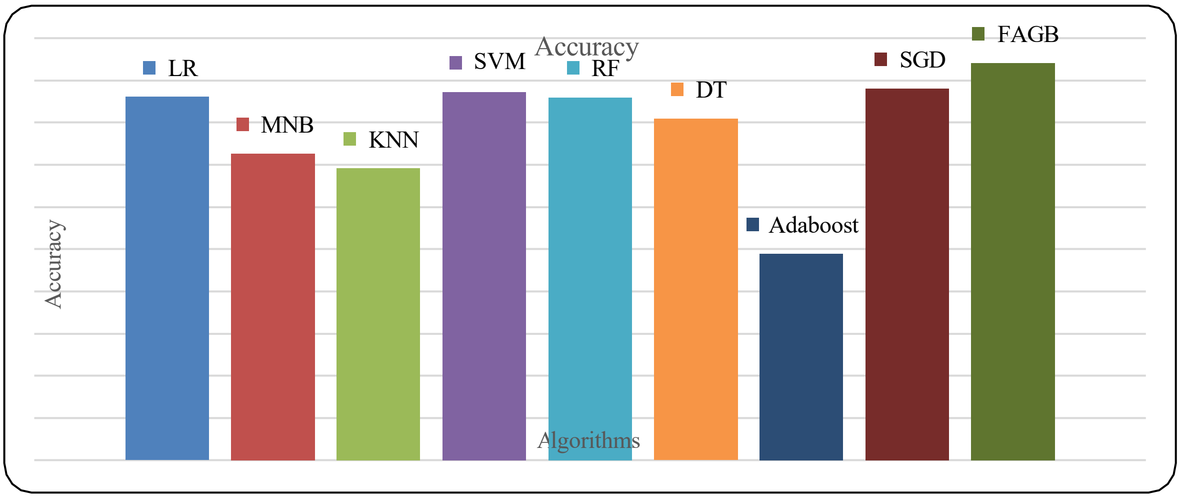

Figure 5. Accuracy

Three categories such as vegetables, fruits and flowers are discussed. Classification task is performed by various ML algorithms. Compared to all other algorithms I,e logistic regression, multinomial naïve bayes, K nearest neighbor , support vector machine , random forest, decision tree, adaboost, SGD , FAGB model proposed which produces the accuracy of 94.1%.

Precision

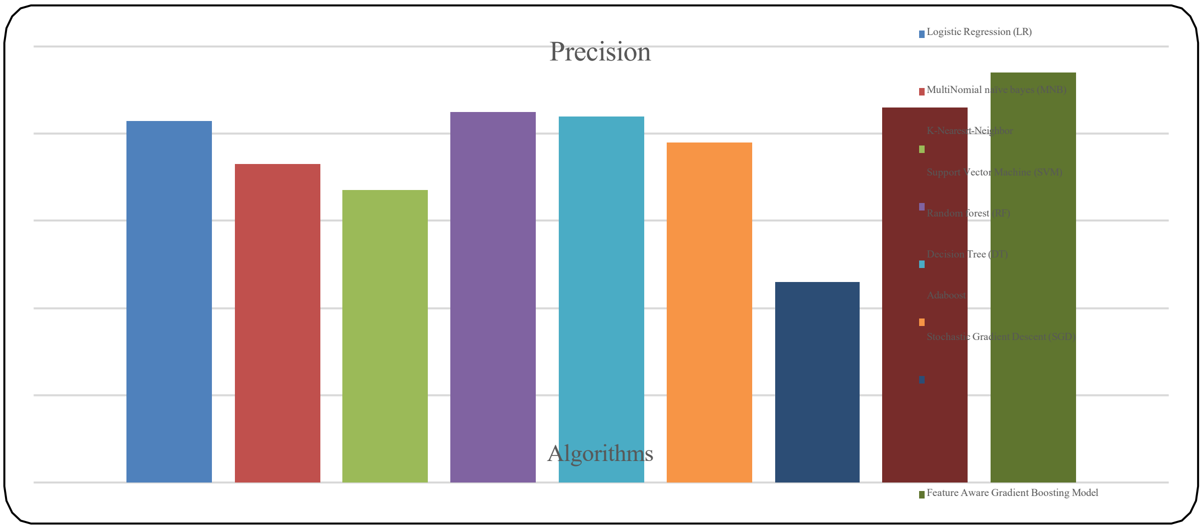

Figure 6. Precision

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 7, 2025
https://www.theaspd.com/ijes.php

Here a total of three categories such as vegetables, fruits and flowers are discussed. Compared to all other algorithms I,e logistic regression, multinomial naïve bayes, K nearest neighbour , support vector machine , random forest, decision tree, adaboost, SGD , the proposed FAGB model produces the accuracy of 0.94.
The classification task is performed using various standard machine learning models. SGD achieves a higher recall of 0.86, However, the proposed the proposed FAGB model produces the higher recovery value of 0.93.

F1Score

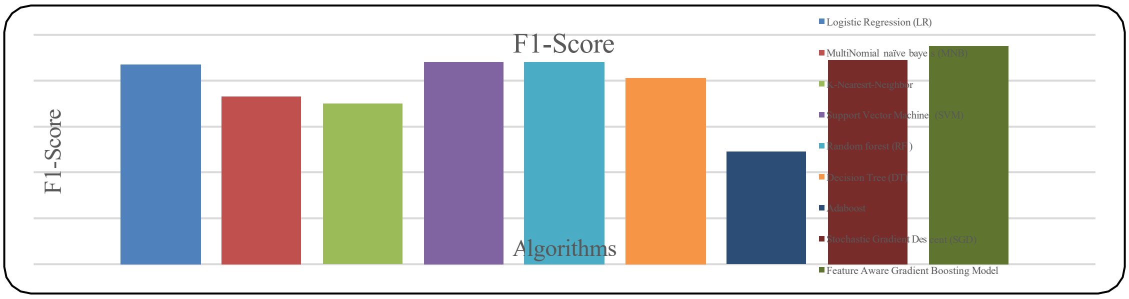

Figure 8. F1-Score

F1 score performance is evaluated to know the type of question. The task of classification is performed using various ML algorithms. Proposed FAGB model produces the higher F1 score of 0.95 compared with other existing models.
The table below illustrates a comparison of results obtained from various algorithms. For the question processing module, a Feature-Aware Gradient Boosting Model has been introduced. This model has been trained on the Kisan Call Center Dataset to predict the expected answer for queries posed by farmers.
Moreover, the performance of the Feature-Aware Gradient Boosting Model has been evaluated for identifying the type of question. Three classes, namely fruits, vegetables, and flowers, were considered for this task. Classification was carried out using different standard machine learning models. While SGD achieves a higher accuracy of 88.03% compared to other models, the proposed Feature-Aware Gradient Boosting Model achieves even higher accuracies, reaching 94.1%.

Table 1 shows comparison table of the proposed model and other algorithms.

Methods	Validity (correctness)
Random Forest	86 %
Support Vector Machine	87.2 %
LR algorithm	86.01 %
Multi Nomial NB algorithm	72.61 %
KNN algorithm	69.1 %
Decision Tree	80.9 %
AdaBoost	48.8 %
SGD Classifier	88 %
FAGB (Feature Aware Gradient Boosting Method) (Proposed)	94.1%

1. CONCLUSION (10 PT)

Table 1. Shows that 48.8% accuracy produced by Adaboost algorithm, KNN algorithm and Multi Nomial NB algorithms accuracy were 69.16% and 72.61%, respectively. Additionally, 80.93 percent accuracy produced by the decision tree algorithm. In addition, accuracy produced by Random Forest and Logistic Regression are 86% and 86.01%. The SVM and SGD classifier algorithm produces better accuracy compared to the other previous methods and the proposed FAGB produced an accuracy of 94.1%. Thus, the proposed trait-aware gradient amplification model that gives exact solution to the particular questions of farmers.

REFERENCES (10 PT)

[1] Zaib, Munazza, Zhang, Wei, Sheng, Quan, Mahmood, Adnan, and Zhang, Yang. "Conversational Question Answering: A Survey." Knowledge and Information Systems, vol. 64, no. 12, 2022, pp. 3151–3195., https://doi.org/10.1007/s10115-022-01744-y.

[2] Ojokoh, Bolanle, and Emmanuel Adebisi. "A Review of Question Answering Systems." Journal of Web Engineering, vol. 17, no. 8, 2019, pp. 717–758., https://doi.org/10.13052/jwe1540-9589.1785.

[3] Adamopoulou, Eleni, and LefterisMoussiades. "Chatbots: History, Technology, and Applications." Machine Learning with Applications, vol. 2, 2020, p. 100006., https://doi.org/10.1016/j.mlwa.2020.100006.

[4] Huang, Zhen, Xu, Shiyi, Hu,Minghao, Wang, Xinyi, Qiu, Jinyan, Fu, Yongquan, Zhao, Yuncai, Peng, Yuxing, and Wang, Changjian. "Recent Trends in Deep Learning Based Open-Domain Textual Question Answering Systems." IEEE Access, vol. 8, 2020, pp. 94341– 94356., https://doi.org/10.1109/access.2020.2988903.

[5] Antoniou, Christina, and Nick Bassiliades. "A Survey on Semantic Question Answering Systems." The Knowledge Engineering Review, vol. 37, 2022, https://doi.org/10.1017/s0269888921000138.

[6] Bohr, Adam, and KavehMemarzadeh. "The Rise of Artificial Intelligence in Healthcare Applications." Artificial Intelligence in Healthcare, 2020, pp. 25–60., https://doi.org/10.1016/b978-0-12-818438-7.00002-2.

[7] Y. Ou, S. Chuang, W. Wang and J. Wang, "Automatic Multimedia-based Question-Answer Pairs Generation in Computer Assisted Healthy Education System," 2022 10th International Conference on Orange Technology (ICOT), Shanghai, China, 2022, pp. 01-04, doi: 10.1109/ICOT56925.2022.10008119.

[8] C. J. Awati, S. S. Shirgave and V. S. Raval, "Accurate Answers Selection and Expert Recommendation in Community Question Answers System," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 1171-1174, doi: 10.1109/ICICCS51141.2021.9432089.

[9] L. S. Nair and S. M. K, "Knowledge Graph based Question Answering System for Remote School Education," 2022 International Conference on Connected Systems & Intelligence (CSI), Trivandrum, India, 2022, pp. 1-5, doi: 10.1109/CSI54720.2022.9924128.

[10] Khurana, Diksha, Koli, Aditya, Khatter, Kiran, and Singh, Sukhdev. "Natural Language Processing: State of the Art, Current Trends and Challenges." Multimedia Tools and Applications, vol. 82, no. 3, 2022, pp. 3713–3744., https://doi.org/10.1007/s11042-022-13428-4.

[11] Sharma, Yashvardhan, and Sahil Gupta. "Deep Learning Approaches for Question Answering System." Procedia Computer Science, vol. 132, 2018, pp. 785–794., https://doi.org/10.1016/j.procs.2018.05.090.

[12] Wu, Jinmeng, Mu,Tingting,Thiyagalingam,Jeyan,andGoulermas, John. "Memory-Aware Attentive Control for Community Question Answering with Knowledge-Based Dual Refinement." *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, pp. 1–14., https://doi.org/10.1109/tsmc.2023.3234297.

[13] Rahmath K, Reji,ReghuRaj,P.C.,and P C,Rafeeque. "Malayalam Question Answering System Using Deep Learning Approaches." *IETE Journal of Research*, 2022, pp. 1–13., https://doi.org/10.1080/03772063.2022.2077846.

[14] Rahmath K, Reji,ReghuRaj,P.C.,and P C,Rafeeque. "Malayalam Question Answering System Using Deep Learning Approaches." *IETE Journal of Research*, 2022, pp. 1–13., https://doi.org/10.1080/03772063.2022.2077846.

[15] Pundge, AjitkumarMeshram, and C. NamrataMahender. "Evaluating Reasoning in Factoid Based Question Answering System by Using Machine Learning Approach." *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, 2018, https://doi.org/10.1109/cesys.2018.8724085.

[16] Golzari, Sh., Sanei, F., Saybani,M.R., Harifi,A., and Basir.M. "Question Classification in Question Answering System using Combination of Ensemble Classification and Feature Selection."*Journal of AI and Data Mining*, vol. 10, no. 1, 2022, pp. 15-24. https://doi.org/10.22044/jadm.2021.10016.2142.

[17] Wang, Peng, Yan,Yun,Si,Yingdong,Zhu,Gancheng,Zhan,Xiangping, Wang, Jun, and PanRunsheng. "Classification of Proactive Personality: Text Mining Based on Weibo Text and Short-Answer Questions Text." *IEEE Access*, vol. 8, 2020, pp. 97370– 97382., https://doi.org/10.1109/access.2020.2995905.

[18] Franco, Wellington, Viktor, Caio, Oliveira, Artur, Maia, Gilvan, Brayner, Angelo, Vidal, V., Carvalho, Fernando, and Pequeno, V. "Ontology-Based Question Answering Systems over Knowledge Bases: A Survey." *Proceedings of the 22nd International Conference* on Enterprise Information Systems, 2020, https://doi.org/10.5220/0009392205320539.

[19] Khongcharoen, Wirit, Saetia, Chanatip, Chalothorn, Tawunrat, and Buabthong, Pakpoom. "Question Answering over Knowledge Graphs for Thai Retail Banking Products." 2022 17th International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP), 2022, https://doi.org/10.1109/isai-nlp56921.2022.9960247.

[20] Sarker, Iqbal H. "AI-Based Modeling: Techniques, Applications and Research Issues towards Automation, Intelligent and Smart Systems." *SN Computer Science*, vol. 3, no. 2, 2022, https://doi.org/10.1007/s42979-022-01043-x.

[21] Arbaaeen, Ammar, and Asadullah Shah. "Ontology-Based Approach to Semantically Enhanced Question Answering for Closed Domain: A Review." *Information*, vol. 12, no. 5, 2021, p. 200., https://doi.org/10.3390/info12050200.

[22] Yunmar, RajifAgung, and I WayanWiprayogaWisesa. "Design of Ontology-Based Question Answering System for Incompleted Sentence Problem." *IOP Conference Series: Earth and Environmental Science*, vol. 258, 2019, p. 012032., https://doi.org/10.1088/1755- 1315/258/1/012032.

[23] Ou, Yang-Yen, Chuang, Shu-Wei, Wang, Wei-Chun, and Wang, Jhing-Fa. "Automatic Multimedia-Based Question-Answer Pairs Generation in Computer Assisted Healthy Education System." *2022 10th International Conference on Orange Technology (ICOT)*, 2022, https://doi.org/10.1109/icot56925.2022.10008119.

[24] Awati, Chetan J., Shirgave, Suresh, and Raval, Veera. "Accurate Answers Selection and Expert Recommendation in Community Question Answers System." *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2021, https://doi.org/10.1109/iciccs51141.2021.9432089.

[25] Nair, Lekshmi S, and Shivani M K. "Knowledge Graph Based Question Answering System for Remote School Education." *2022 International Conference on Connected Systems & Intelligence (CSI)*, 2022, https://doi.org/10.1109/csi54720.2022.9924128. Wang, Jindi, Sui, Haigang, Li, Yongcheng, and Hu, Lieyun. "A Semantics-Guided and Spatial-Aware Framework for Natural Resources Geo-Analytical Question Answering." *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, https://doi.org/10.1109/igarss46834.2022.9883649.

BIOGRAPHIES OF AUTHORS (10 PT)

BIOGRAPHIES OF AUTHORS

	Dr Prashant Y. Niranjan is an Assistant Professor in the Department of Computer Science and Engineering, KLS Gogte Institute of Technology, Belagavi, Karnataka, India. He completed his BE in Information Science and Engineering from Visvesaraya Technological University Belagavi in 2011, MTech in Computer Science and Engineering from Visvesaraya Technological University Belagavi in 2014 and PhD in Computer Science and Engineering from Visvesaraya Technological University Belagavi in 2023.He can be contacted at email: prashant053@gmail.com https://scholar.google.co.in/citations?user=ZMG-JOIAAAAJ&hl=enscopus https://www.scopus.com/authid/detail.uri?authorId=56619190000