Leveraging Digital Image Processing and Machine Learning for Precision Farming and Loan Prediction

Gagan S[1], Mary Cherian[2] ,Ayushi Sah[3], Aditi Pandy[4], Anjali, Bhardwaj[5]

[1]Dayananda Sagar College of Engineering
[2]Dayananda Sagar College of Engineering
[3]Dayananda Sagar College of Engineering
[4]Dayananda Sagar College of Engineering
[5]Dayananda Sagar College of Engineering

Abstract: *The agricultural sector is the cornerstone of the economy of any country and also makes the country largely self-sufficient in staple food production, such as rice, wheat, and pulses. The crop production has seen a decline over the previous years due to climate changes and soil degradation. The lack of access to emerging technologies that can provide insightful information for precision farming and hence, increase productivity is a need that has to be addressed. This paper aims to help the Indian farmers who face numerous challenges, including soil variability, crop selection, revenue prediction, and access to financial resources by applying digital image processing and advanced machine learning models. It makes use of Digital Image processing for soil classification and CNN Model for crop recommendation. It also provides crop recommendations based on soil nutrients and pH levels with the help of a decision tree model. The ML Model for revenue prediction and loan eligibility assessments further assist farmers in financial planning. Additionally, to make the features easy to use and accessible, a mobile application is being developed.*

Keywords: Digital Image Processing, Machine Learning, CNN, Decision Tree, crop recommendation, yield prediction, revenue prediction, loan value (FOIR) calculation, chatbot, Precision farming.

1. Introduction

Agriculture is one of the oldest occupations amongst all civilizations all over the world. It is the main source of livelihood and food security for a major part of the population. Agriculture also helps to sustain and develop both secondary as well as tertiary sectors, making it the backbone of the economy. The 2020-21 Indian economic survey states, more than half of the Indian workforce is employed in agriculture. The agriculture sector has seen a decline in its contribution to the country's GDP which was 15% in the last fiscal year (FY23) as opposed to the 35% in 1990-91. Indian farmers have been relying on the observations made by their named eyes for the crop to plant and the fertilisers to use for best yield. The lack of availability of technology that can improve productivity needs a revolutionary change.

About 52% of the farming families in India were estimated to be indebted. At the nationwide level, about 60% of the unpaid loans were availed from institutional sources, including Government, Co-operative society and Banks. Farmers in India have been facing several issues like crop failure, changes in monsoon, lower yield, fluctuation in market rates etc. like this many problems are faced by the farmers in obtaining and repayment of agricultural loans.

The Indian agriculture sector requires adoption of precision farming leveraging digital image processing and machine learning technologies. Since, the current climate conditions are changing in ways that a farmer cannot predict, the adoption of cutting edge machine learning models can immensely improve the decision making as to what crop to plant and predicting an estimated revenue on the produce.

The proposed system takes an image of the soil as an input and its colour- oriented characteristics are analyzed using digital image processing. The raw image is processed and taken as an input to classify the type of soil using a CNN Model. Also, using the imputed nutrient values, pH levels and the humidity and climate conditions on the basis of location, the system provides crop recommendations using a Decision Tree model. Additionally, the system provides revenue prediction. In the mobile application the farmers can find a market place which acts as a community platform to lend and borrow equipment and sell produce at optimum market rates. The system also provides an AI enabled chatbot which can answer farming related questions such as best fertilizer for the crops, etc. Thereby this proposed work will assist the Indian farmers with precision farming and effective solutions for profitable production.

2. Related Work

Digital image processing is an effective tool for analyzing and extracting information from raw images. John Carlo Puno et.al [1] proposed a method which utilizes image processing and artificial neural network to determine the nutrients and pH level of soil. The study utilizes Soil Test Kit (STK) and Rapid Soil Testing (RST) to identify pH, nutrients (K, Zn, N, P, Ca, and Mg) levels in the soil accurately. Ricardo Alasco et.al [2] proposed SoilMATTic, which is an automated soil nutrient and pH level analyzer made using Arduino that implements digital image processing and artificial neural networks to provide faster and accurate soil analysis. The system uses image acquisition, image processing, and a training system to identify N, K, P, and pH levels in Philippine farmlands. The system achieved 96.67% accuracy.

Hement Kumar Sharma et.al [3] proposed using digital image processing techniques to classify and analyze different types of soil in Udaipur, Rajasthan,India. The paper explores the three levels of digital image processing: low level, medium level, and high level. It involves applying SVM classification and regression, with features such as filtration, enhancement, segmentation, and classification.

Kulkarni Varsha et.al [4] presented a way for soil classification, crop selection, and fertilizer requirements based on soil series. They collected data on aspects including pH, moisture, humidity, temperature, rainfall, and soil composition and preprocessed the data according to requirements. Machine learning algorithm, Support Vector Machines (SVM) is used to train the model and predict suitable crops and fertilizer requirements. The results obtained from it show higher accuracy in crop yield prediction compared to existing methods. S.Bhanumath et.al [5] proposed data mining techniques and ML algorithms to analyze various attributes such as location, soil type, pH value, and nutrient levels to predict crop yield. The authors compared the accuracy of two algorithms, Random Forest and Back propagation. The data indicates that the Random Forest algorithm outperforms Back propagation in terms of error rate. The study also emphasizes the need for considering factors like climate conditions, evapotranspiration, and previous crop information in yield prediction.

Shilpa Mangesh Pande et.al [6] proposed the use of ML for crop recommendation and yield prediction to recommend the most profitable crop for farmers. The system takes inputs such as location, type of soil, and other parameters. The Random Forest algorithm showed the highest accuracy of 95%. The system also suggests the best time to use fertilizers based on rainfall predictions. G. Mariammal et.al [7] presented a novel approach called modified recursive feature elimination (MRFE) for crop prediction based on soil and environmental characteristics. This study involves FS methods for enhanced accuracy in predicting crops. Various classification approaches including k nearest neighbour, naïve bayes classifier, decision trees, SVM, random forests, and bagging are used in assessing the performance of the MRFE method. The results show that MRFE gives better outcomes than other FS methods with a 95 percent success rate.

The preceding research often relied on complex climatic and agricultural factors, making it difficult for farmers to understand. In contrast, Potnuru Sai Nishant et.al [8] simplifies the prediction process by using basic parameters that farmers can easily input. The authors collected data from the Indian Government Repository, which included attributes such as location, crop, season, and production. They compared regression techniques like Lasso, ENet, and Kernel Ridge, along with Stacking Regression, to improve prediction accuracy. The output indicates that the stacked regression approach significantly improved the performance of yield predictions. M.Kalimuthu et.al [9] proposed using the Naive Bayes algorithm, to estimate crop production in India. The decisions about which crops to sow are made based on temperature, humidity, and moisture content. The researchers collected preceding and present data, including weather conditions and soil parameters, and used the Naive Bayes algorithm to train the model using a training data set. This method attained an accuracy of 97% in predicting crop yield. Ramesh Medar et.al [10] proposed various algorithms such as Naive Bayes, K-Nearest Neighbour, Artificial Neural Networks, and multiple linear regressions to analyze and predict crop yield rates. The authors have taken into consideration attributes like market price, production rate, and government policies in crop selection. The outcomes of their research include the ability to select suitable crops for specific regions and seasons, and the evaluation of accuracy between different methods.

Rishi Gupta et.al[11] proposed a method which focuses on weather-based crop prediction in India using big data analytics. The authors propose a system that utilizes the MapReduce framework and k-means clustering to analyze and process large datasets related to temperature, rainfall, soil, seed type, and crop production. The outcomes include the recognition of the relationship between crops, weather conditions, and region, and the ability to predict yield of crops with high accuracy.

G. Thapaswini et.al [12] presented a methodology for predicting crop price using ML techniques, decision tree and neuro-evolutionary algorithms after a thorough analysis of other algorithms. The study also suggests the future development of an automated price recommendation system using genetic algorithms.

Puneeth B. R at.al [13] proposed a methodology which utilizes ML models for loan prediction. The researchers evaluated the performance metrics of logistic regression, decision trees, random forest, and XGBoost models. The XGBoost model achieved the maximum accuracy of 81.17% and demonstrated high precision, recall, specificity, and F1-score. Such implementations can provide a quick and accessible way to select competent loan candidates. Youness Abakarim et.al [14] suggested the use of a real-time binary classification model for approval of loan credit using deep learning. The paper compares their proposed model with four benchmark algorithms: linear SVM regression, logistic regression, nonlinear auto-regression neural network, and a basic neural network approach. The results demonstrate how the deep learning model outperforms the benchmark algorithms in terms of accuracy, precision, and recall. The model is based on a deep neural network and classifies loan applicants as acceptable or not.

3. Proposed Methodology

The implementation of systems that utilize the capabilities of Machine Learning and Digital Image Processing can help farmers make informed decisions and maximize crop yield and revenue. Precision farming is a modern approach which utilizes technologies to optimize crop yields, minimize inputs and maximize resource utilization and that is what this system aims to achieve. The architecture of AgriLink comprises several interconnected components working together to provide a comprehensive solution for farmers. At its core is a backend system built using Node.js, which serves as the backbone for handling data processing, storage, and communication between various modules. This backend system interfaces with a MongoDB database, which stores crucial data such as soil characteristics, crop information, and user profiles.

On the frontend, AgriLink utilizes a React Native mobile application framework, offering a user-friendly interface accessible on smart phones. This mobile app serves as the primary interface for farmers to interact with the system, capturing soil images, receiving crop recommendations, and accessing marketplace features.

Additionally, AgriLink integrates FastAPI, a Python-based framework, to connect with machine learning models deployed for tasks such as soil classification and crop recommendation. These models leverage advanced algorithms, including Convolutional Neural Networks (CNN) and Decision Trees, to analyze soil images and environmental data, providing actionable insights for farmers.

Overall, the architecture of AgriLink is designed to seamlessly integrate diverse technologies, leveraging machine learning, image processing, and AI-driven chatbots to empower farmers with data-driven insights and facilitate efficient farming practices.

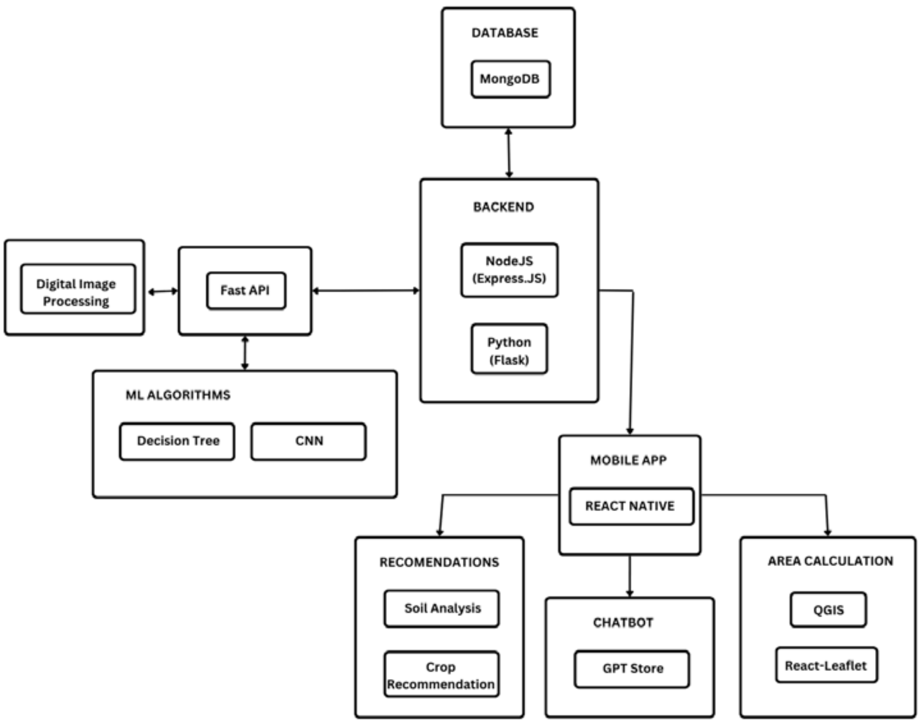

Figure 1: Agrilink System Architecture diagram

A. Soil classification using image processing

Digital image processing is essential for soil analysis. It extracts required data from images for further analysis. Initially, soil images are captured using digital cameras or other imaging devices. These images are then preprocessed to enhance quality, remove noise, and normalize lighting conditions, ensuring optimal analysis. Next, techniques such as thresholding, edge detection, and feature extraction, are applied to identify relevant soil characteristics, such as texture, colour, and structure from the image. These images serve as input data for a Convolutional Neural Network (CNN) model.

The model used consists of several layers arranged in a Convolutional Neural Network (CNN). The input layer specifies the shape of the input images (150x150 pixels with 3 colour channels - RGB). The convolution layers perform feature extraction by applying filters to the input image. It uses the ReLU activation function for introducing non-linearity. The Pooling layers down-sample the feature maps, reducing their spatial dimensions while retaining the most important information. The Dropout layers help in regularization by randomly dropping a certain percentage of neurons during training, preventing over fitting. The Flatten layer flattens the multi-dimensional output of the previous layer into a one-dimensional vector. This prepares the data for input into the fully connected layers. The fully connected layers perform classification based on the features extracted by the convolutional layers. Finally batch normalization normalizes the activations of the previous layer, which helps in training deep neural networks more efficiently. Through the process of forward propagation and back propagation, the CNN iteratively adjusts its parameters during training to minimize prediction errors and improve accuracy. The soil classification aids in agricultural decision-making and crop management.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 150, 150, 32)	896
activation (Activation)	(None, 150, 150, 32)	0
max_pooling2d (MaxPooling2D)	(None, 50, 50, 32)	0
dropout (Dropout)	(None, 50, 50, 32)	0
conv2d_1 (Conv2D)	(None, 50, 50, 64)	32,832
activation_1 (Activation)	(None, 50, 50, 64)	0
conv2d_2 (Conv2D)	(None, 50, 50, 64)	65,600
activation_2 (Activation)	(None, 50, 50, 64)	0
max_pooling2d_1 (MaxPooling2D)	(None, 25, 25, 64)	0
dropout_1 (Dropout)	(None, 25, 25, 64)	0
conv2d_3 (Conv2D)	(None, 25, 25, 128)	131,200
activation_3 (Activation)	(None, 25, 25, 128)	0
conv2d_4 (Conv2D)	(None, 25, 25, 128)	262,272
activation_4 (Activation)	(None, 25, 25, 128)	0
max_pooling2d_2 (MaxPooling2D)	(None, 12, 12, 128)	0
dropout_2 (Dropout)	(None, 12, 12, 128)	0
flatten (Flatten)	(None, 18432)	0
dense (Dense)	(None, 1024)	18,875,392
batch_normalization (BatchNormalization)	(None, 1024)	4,096
dropout_3 (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 6)	6,150
activation_5 (Activation)	(None, 6)	0

Figure 2: CNN training model

B. Crop recommendation using Decision Trees

Decision Tree is a machine learning algorithm with versatile uses. In the context of crop recommendation, a Decision Tree algorithm utilizes various parameters such as soil nutrients, weather, and other relevant attributes to determine the most suitable crop for a given set of conditions.

In the proposed system, the algorithm analyses a dataset comprising attributes like potassium, nitrogen, phosphorus, moisture, temperature, rainfall, and other factors crucial for crop growth. The algorithm iteratively splits the dataset based on these attributes, creating decision nodes that effectively partition the data into subsets with similar characteristics. By evaluating splitting criteria at each node, the algorithm determines the most crucial features for recommending crops. Ultimately, the leaf nodes of the Decision Tree represent the recommended crops for specific conditions, enabling farmers with personalized and data-driven guidance for optimizing their agricultural practices. Through this approach, AgriLink leverages Decision Trees to offer tailored crop recommendations, thereby enhancing productivity and sustainability in farming endeavors. This crop recommendation system leverages decision trees to provide informed and personalized guidance to farmers, enhancing productivity and sustainability.

C. Revenue Prediction and Loan Analysis

Fixed Obligation-to-Income Ratio (FOIR) assesses a farmer's debt servicing capacity relative to income. Lower FOIR implies healthier finances. Revenue prediction utilizes crop yield estimates, market rates, and input costs for informed decision-making. The farmland area calculation integrates the MapView component and markers from the react-native-maps library into the mobile application interface. Users mark the corners of their farmland by tapping on the map, with each tap representing a coordinate point. Current system utilizes 4 markers, but with a higher number of markers the accuracy in area calculation can be increased. The longitude and latitude of these points are captured and stored for further processing. The system utilizes a geometric algorithm, such as the shoelace formula or convex hull algorithm, to calculate the area enclosed by the marked coordinates. The algorithm iterates over the points, computing the area of the polygon formed by connecting them. This calculated area serves as an estimation of the farmland area. Additionally, the application retrieves the government-set rate per unit area (circle rates in Karnataka) and multiplies it by the calculated area to determine the value of the land. This system provides a convenient and efficient tool for accurately assessing the area and value of their farmland. This information along with an estimation of farmers revenue based on crop yield, the system does loan value analysis, aiding lenders in assessing borrowers' repayment capacity and farmers in financial planning. Integration of FOIR enhances risk management and supports farmers in accessing financial support.

D. Marketplace Integration

The system implements a mobile application with easy-to-use UI for the farmers to access all the facilities at one place. This app integrates a marketplace designed to facilitate equipment sharing, produce selling, and market rate discovery. Farmers can upload images and prices for resources such as tractors, tyres, fertilizers and also their crops, connected to the owner’s contact details on the App. Another farmer in need of these farming equipment can get in contact with the owner about buying or borrowing the same. This fosters community engagement and resource sharing, enabling farmers to collaborate, exchange knowledge, and support each other, ultimately enhancing efficiency and sustainability in agriculture.

E. Chatbot Integration using ChatGPT API

The farming-related chatbot integrates the ChatGPT API, allowing users to ask questions about fertilizers, pest management, and farming practices. Natural language processing techniques are employed to understand user queries and generate relevant responses. This enhances user engagement and provides valuable assistance in agricultural activities.

\

4. Results and discussions

The proposed system involves soil images, nutrient values and pH levels, weather and climate conditions data for accurately classifying soil type, recommending crops that would result in the best yield, estimating revenue and calculating the FOIR to help the farmer's financial situation. The chatbot adds another layer of insight into best practices and encourages precision, sustainable farming. The marketplace feature creates a community for the farmers to earn and understand the market value of their produce. This is a user-friendly mobile application that provides the users (farmers) with a one-stop solution for all their farming problems. Figure 3 depicts the user interface for the application.

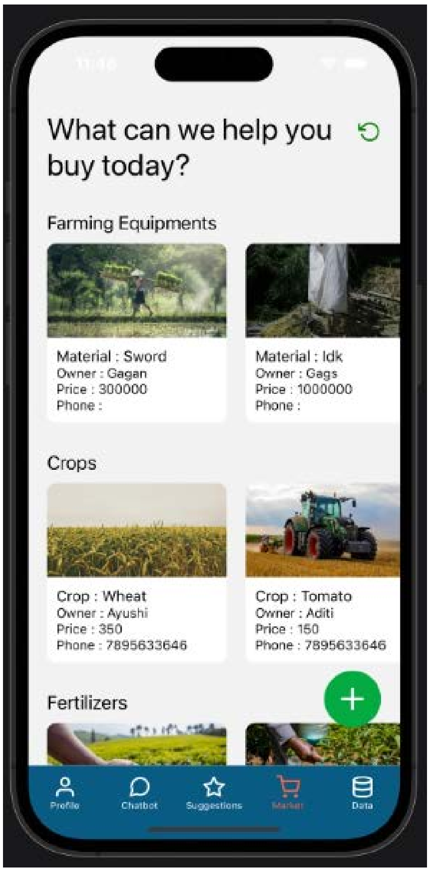

Figure 3: Marketplace items listing

5. Conclusion and Future Scope

This paper recognizes the well-documented challenges farmers face in optimizing their farming practices and accessing loans. It emphasizes the potential of advanced technologies like digital image processing, GIS, and Machine Learning in providing data-driven insights for sustainable crop rotation, yield prediction, and market value estimation, along with loan value prediction and FOIR calculations. It is a user-friendly mobile application that provides the users (farmers) with a one- stop solution for all their farming problems. The future work includes training the AI driven chatbot on more data to be able to provide extensive assistance on a wide range of issues such as identifying pest problems and suggesting solutions. We also intend to work on making the application more accessible by providing language options so the farmers can interact with the app in their regional language. The overarching goal is to empower Indian farmers, promoting self-sufficiency, efficiency, and profitability in their farming endeavors.

6. References

[1] J. C. Puno, E. Sybingco, E. Dadios, I. Valenzuela and J. Cuello, "Determination of soil nutrients and pH level using image processing and artificial neural network," *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, 2017, pp. 1-6, doi: 10.1109/HNICEM.2017.8269472.

[2] R. ALASCO et al., "SoilMATTic: Arduino-Based Automated Soil Nutrient and pH Level Analyzer using Digital Image Processing and Artificial Neural Network," *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM),* Baguio City, Philippines, 2018, pp. 1-5, doi: 10.1109/HNICEM.2018.8666264.

[3] H. K. Sharma and S. Kumar, "Soil Classification & Characterization Using Image Processing," *2018 Second International Conference on Computing Methodologies and Communication (ICCMC),* Erode, India, 2018, pp. 885-890, doi: 10.1109/ICCMC.2018.8488103.

[4] K. Varsha, P. G, S. N and V. J, "Soil Classification, Crop Selection and Prediction of Fertilizer based on Soil Series," *2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC),* Bengaluru, India, 2022, pp. 1174-1178, doi: 10.1109/IIHC55949.2022.10059636.

[5] S. Bhanumathi, M. Vineeth and N. Rohit, "Crop Yield Prediction and Efficient use of Fertilizers," *2019 International Conference on Communication and Signal Processing (ICCSP),* Chennai, India, 2019, pp. 0769-0773, doi: 10.1109/ICCSP.2019.8698087.

[6] S. M. Pande, P. K. Ramesh, A. Anmol, B. R. Aishwarya, K. Rohilla and K. Shaurya, "Crop Recommender System Using Machine Learning Approach," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC),* Erode, India, 2021, pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351.

[7] G. Mariammal, A. Suruliandi, S. P. Raja and E. Poongothai, "Prediction of Land Suitability for Crop Cultivation Based on Soil and Environmental Characteristics Using Modified Recursive Feature Elimination Technique With Various Classifiers," *in IEEE Transactions on Computational Social Systems, vol. 8,* no. 5, pp. 1132-1142, Oct. 2021, doi: 10.1109/TCSS.2021.3074534.

[8] P. S. Nishant, P. Sai Venkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," *2020 International Conference for Emerging Technology (INCET),* Belgaum, India, 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.

[9] M. Kalimuthu, P. Vaishnavi and M. Kishore, "Crop Prediction using Machine Learning," *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT),* Tirunelveli, India, 2020, pp. 926-932, doi: 10.1109/ICSSIT48917.2020.9214190.

[10] R. Medar, V. S. Rajpurohit and S. Shweta, "Crop Yield Prediction using Machine Learning Techniques," *2019 IEEE 5th International Conference for Convergence in Technology (I2CT),* Bombay, India, 2019, pp. 1-5, doi: 10.1109/I2CT45611.2019.9033611.

[11] R. Gupta et al., "WB-CPI: Weather Based Crop Prediction in India Using Big Data Analytics," *in IEEE Access*, *vol. 9*, pp. 137869-137885, 2021, doi: 10.1109/ACCESS.2021.3117247.

[12] G. Thapaswini and M. Gunasekaran, "A Methodology for Crop Price Prediction Using Machine Learning," *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC),* Tumkur, Karnataka, India, 2022, pp. 1-7, doi: 10.1109/ICMNWC56175.2022.10031852.

[13] P. B. R, A. K, A. Kumar, B. Rao, P. S. K and S. A. P, "An Approach to Predict Loan Eligibility using Machine Learning," *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE),* Karkala, India, 2022, pp. 23-28, doi: 10.1109/AIDE57180.2022.10059881.

[14] Y. Abakarim, M. Lahby and A. Attioui, "Towards An Efficient Real-time Approach To Loan Credit Approval Using Deep Learning," *2018 9th International Symposium on Signal, Image, Video and Communications (ISIVC),* Rabat, Morocco, 2018, pp. 306-313, doi: 10.1109/ISIVC.2018.8709173.