Proceedings of the 3rd International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2025.23348

AI-GeoInfo Crop Recommendation Framework Using Decision Tree Classifier and Flask-based GeoAPIs

Hisham AbouGrad[1*], Manasa Yegamati[1]

[1]*School of Architecture, Computing and Engineering, University of East London, London, UK*
**Corresponding Author. Email: h.abougrad@uel.ac.uk*

Abstract: The crop recommendation AI-GeoInfo Framework is an advanced web application system that assists farmers and agricultural specialists in making informed crop selections based on soil and weather conditions. This framework combines geospatial data, machine learning, and web development technologies to generate tailored crop recommendations. At its core is a Decision Tree Classifier, a model chosen for its reliability and interpretability. The model is trained on comprehensive environmental data, including temperature, humidity, rainfall, soil pH, and composition with preprocessing handled by Python SimpleImputer library to address any missing values and enhancement. The web interface built using Flask-based GeoAPI, which enables users to input geographic coordinates. This identifies the nearest soil mapping unit using GeoPy for precise geodesic distance calculations by linking the input location with relevant soil data. This data is fed into the Decision Tree model, which generates an optimized list of crop recommendations based on the specific conditions of the location. The intuitive interface presents these crop recommendations make the AI-GeoInfo framework provide accurate soil and weather information to its users. The framework also allows easy updates and scalability by adapting diverse agricultural applications and responsive to advances in data and technology.

Keywords: Decision Tree Classifier, Python GeoPy Module, Crop Recommendation Machine Learning Model, Soil Composition, Flask-based GeoAPI Web Application.

1. Introduction

The crop recommendation artificial intelligence (AI) geographic information (AI-GeoInfo) framework represents a cutting-edge digital platform advancing sustainable agriculture through customized, geographically specific crop recommendations. Utilizing machine learning and geospatial analytics, this system enables farmers and agricultural experts to implement data-driven crop selections based on their land's distinct environmental parameters, including soil composition, pH levels, temperature variations, humidity conditions, and precipitation patterns.

The goal of the Crop Recommendation AI-GeoInfo Framework is to create and assess an intelligent system that provides precise, location-specific crop recommendations by fusing web technologies, machine learning, and geospatial data analytics [1]. The framework determines the best crops for certain regions by analysing a large dataset that includes soil composition, climate, and other important environmental parameters using a Decision Tree Classifier model [2]. Farmers and other agricultural professionals may obtain real-time, site-specific suggestions through an intuitive web interface thanks to the integration of a Flask-based GeoAPI [3]. The ultimate goal of this framework

DOI: 10.54254/2755-2721/2025.23348

is to improve agricultural sustainability, productivity, and resistance to environmental issues by bridging the gap between traditional farming practices and contemporary data-driven initiatives [4].

The framework's foundation rests on a Decision Tree Classifier, which processes comprehensive soil and climate datasets to determine optimal crop choices for specific locations [5]. Integration with Geo Weather APIs and the Harmonised World Soil Database (HWSD) provides real-time environmental data by enhancing recommendation accuracy. The Flask-based web interface incorporates Python’s GeoPy module functionality for geodesic calculations, which accurately identify relevant soil mapping units based on user-provided geographic coordinates. This approach minimises dependence on traditional crop selection practices, delivering evidence-based insights adaptable to diverse environmental conditions. The platform's accessibility makes it valuable for various stakeholders, farmers to policymakers. Its modular structure allows for seamless integration of emerging data sources and technologies, maintaining relevance as agricultural science advances. This initiative showcases how integrating machine learning with geospatial data through an accessible interface can enhance farming decisions, productivity, and sustainability [6].

2. Related work

In the twenty-first century, the agricultural industry confronts significant obstacles due to soil erosion, climate change, and the need for sustainable farming methods [7]. Conventional methods of crop selection, which frequently depend on broad guidelines, are unable to account for the environmental complexity of every site. This puts pressure on agricultural sustainability by resulting in inefficiencies in crop productivity and resource utilisation [8]. A possible remedy for these problems is precision agriculture, which makes crop suggestions based on data. To meet these demands, the Crop Recommendation AI-GeoInfo Framework was created, which uses geospatial analytics and machine learning to offer personalised crop selection suggestions [9]. This framework offers location-specific insights suited to the environmental circumstances of each field, with the goal of enhancing agricultural output and resource efficiency via the integration of cutting-edge technology [10].

The combination of web technologies, geographical data, and machine learning (ML) to enhance crop suitability forecasts based on soil and climatic data is examined in the literature review for the Crop Recommendation AI-GeoInfo Framework [11]. This research showing the benefits of algorithms like decision trees, random forests, and neural networks for forecasting crop yield, evaluating soil health, and controlling pests [12], the use of machine learning models—in particular, Decision Tree Classifiers—has shown great promise in optimising crop recommendations [3]. More sophisticated methods like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) have recently been used to analyse complex time-series data, allowing for more flexible crop recommendations that adapt to changing soil and weather conditions [4]. Localised crop recommendations that take into consideration spatial variability in environmental parameters [5] are made possible by the addition of geographic information system (GIS) and remote sensing datasets, which improves the accuracy of these frameworks [6].

The accessibility of these systems has been made possible by flask-based GeoAPIs, which offer a smooth interface for location-specific, real-time crop advise on mobile devices [7]. Flask GeoAPI integration capabilities enable developers to integrate crop suggestions and climate forecasts into an intuitive platform, enabling farmers and agronomists to make data-driven decisions [8]. Despite their great accuracy, ML-based frameworks nevertheless have drawbacks, especially when it comes to scalability, computing demands, and data integration [9]. Despite all these difficulties, research supports the possibility of integrating machine learning, geographic information systems, and web technologies in precision agriculture [10]. Future advancements in data processing and visualisation should boost crop recommendation systems effectiveness and usability [11].

DOI: 10.54254/2755-2721/2025.23348

3. Method

3.1. Data collection and description

The crop recommendation AI-GeoInfo Framework dataset combines real-time weather data (temperature, humidity, rainfall) from a Geo-WeatherAPI with key soil data (pH, clay, sand, silt content, and cation exchange capacity) from the Harmonised World Soil Database (HWSD), [5]. Accurate site-specific crop recommendations are made possible by the combination of static soil characteristics and dynamic meteorological weather composition input data. The dependability of the framework is increased by using historical crop data from Kaggle datasource to enhance the machine learning model training and testing [4].

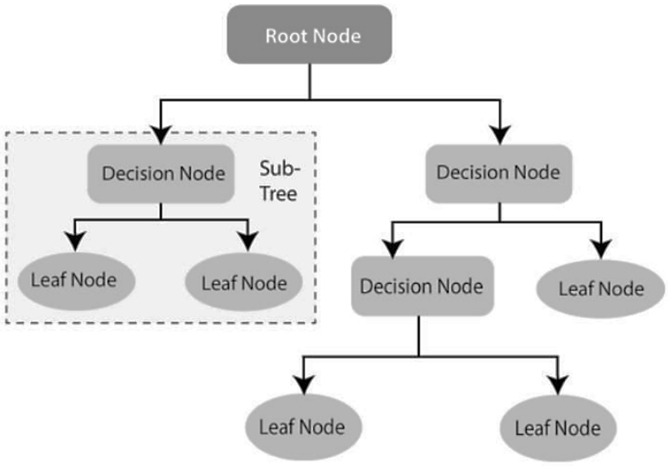

Figure 1: Decision tree model

3.2. Data pre-processing

Before the model is trained, the crop recommendation AI-GeoInfo Framework dataset is subjected to necessary preprocessing to guarantee quality and consistency [12]. To maintain equal scale across features, continuous variables are standardised, and the SimpleImputer is used to manage missing values in important characteristics such as soil pH, clay, sand, silt content, and cation exchange capacity using a mean or median approach [11]. To make categorical data compatible with machine learning algorithms, it is encoded into numerical form. After preprocessing, the dataset is divided into training and testing sets. Usually, this is done using an 80-20 split, in which 20% of the data is set aside for testing and 80% of the data is utilised to train the Decision Tree Classifier [3]. This method maintains sufficient data for an objective assessment of accuracy and prediction performance while fostering strong model development.

3.3. Model development

A comprehensive dataset that combined soil and real-time meteorological weather data was used to evaluate the algorithms (DecisionTreeClassifier, SVC, RFC, and KNN), and machine learning models are essential to the accuracy of the Crop Recommendation AI-GeoInfo Framework [2]. The model's performance was assessed using metrics including accuracy, precision, and recall to make sure the selected model offers trustworthy crop predictions tailored to a certain locale [5]. This method promotes farming that is both productive and sustainable by customising advice to the circumstances of each location [6].

DOI: 10.54254/2755-2721/2025.23348

3.3.1. Model training and selection

To find the most accurate model for crop recommendations, machine learning models were tested during the development of the crop recommendation AI-GeoInfo Framework. These models included K-Nearest Neighbours (KNN), Support Vector Classifier (SVC), Random Forest Classifier (RFC), and Decision Tree Classifier (DTC) [9]. Metrics like F1 score, recall, accuracy, and precision were used to assess each model. The Decision Tree Classifier proved to be the best performer following a thorough testing process. It optimises splits using Gini Impurity and excels at handling intricate, multi-dimensional agricultural data while maintaining accuracy and interpretability.

3.3.2. Model integration

The Decision Tree Classifier was chosen and then included into the Flask-based web application, which communicates with the GeoAPI to offer crop suggestions in real time that are location-specific. To customise suggestions for particular sites, the model maps user-input coordinates to the nearest soil mapping unit using geodesic distance computations. Frequently used in geographic applications, the geodesic distance formula is:

$$d = R.arc\cos(\sin(\emptyset_1)\sin(\emptyset_2) + \cos(\emptyset_1)\cos(\emptyset_2)\cos(\Delta\lambda)) \tag{1}$$

where: d is the distance, R is the Earth's radius, ϕ_1 and φ_2 are the latitudes of the two points, and $\Delta\lambda$ is the difference in longitudes.

4. Experimental results

4.1. Model performance

Four machine learning models were used by the system: K-Nearest Neighbours, Support Vector Classifier, Random Forest Classifier, and Decision Tree Classifier. These models were trained on a dataset that included weather information (temperature, humidity, rainfall) and soil properties (such as pH, clay, sand, silt, and Cation Exchange Capacity).

Table 1: Different models performance comparison

	Decision Tree	SVC	RFC	KNN
Accuracy	91%	89%	90%	85%
Precision	0.91	0.89	0.90	0.84
Recall	0.91	0.88	0.89	0.83
F1-Score	0.91	0.88	0.89	0.83

This table compares the performance of four models: Decision Tree, SVC, RFC, and KNN. The Decision Tree model achieves the highest accuracy 91% and scores across Precision, Recall, and F1-Score 0.91, making it the best choice for this application. SVC and RFC perform similarly, with accuracies of 89% and 90%, while KNN shows the lowest performance at 85% accuracy [3]. This indicates the Decision Tree as the most suitable model for crop recommendation.

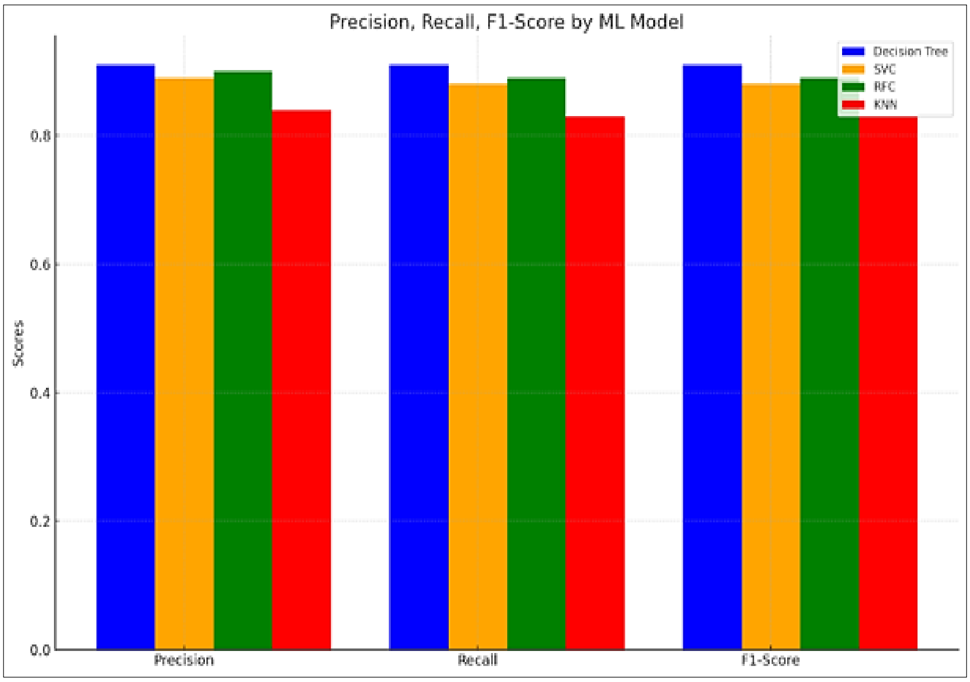

Figure 2: Precision, Recall, F1- Score of the proposed machine learning models

4.2. Crop recommendation

The results page offers users a transparent view of the recommended crop along with the key environmental factors influencing the suggestion. Displaying data such as temperature, humidity, and soil pH below the recommendation, the page provides a clear rationale for each crop choice. This user-friendly presentation, organized in a readable list format, helps build trust by showing exactly why specific crops are recommended for given conditions.

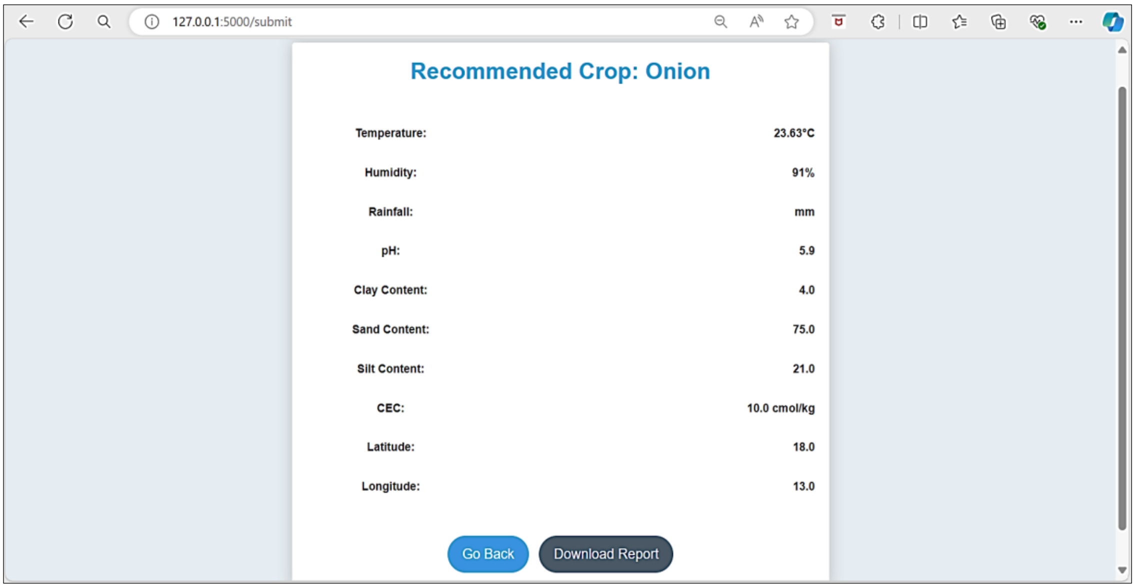

Figure 3: Recommended crop along with weather and soil details of coordinates

As shown in Figure 3, the crop recommendation AI-GeoInfo Framework's output interface has suggested Onion as the recommended crop based on environmental and soil parameters. Key factors

DOI: 10.54254/2755-2721/2025.23348

such as temperature 23.83°C, humidity 91%, soil pH 5.9, clay, sand, and silt content, as well as geographic coordinates, are displayed to support the recommendation. Users can navigate back or download a detailed report, making the interface user-friendly for farmers and agricultural experts seeking data-driven crop choices.

5. Conclusion

The crop recommendation AI-GeoInfo Framework provides accurate, location-specific crop recommendations with 91% accuracy rate by integrating machine learning and geographical data. By combining soil characteristics (pH, clay, sand, silt content) with real-time meteorological weather data (temperature, humidity, rainfall), the system adjusts to the present situation and makes sure that recommendations are pertinent and supported by facts. The Decision Tree Classifier was chosen for crop suitability prediction because of its excellent performance in deciphering complicated environmental data. The framework, created with Flask, provides a scalable and secure online interface that makes it simple for farmers, researchers, and policymakers to enter geographical data and get crop recommendations that can be put into practice. This easy-to-use configuration guarantees data protection and privacy while facilitating accessibility for users with varying technical backgrounds. The method greatly improves agricultural decision-making, promoting resource optimisation and sustainable practices. Predictions might be further improved and the framework's applicability expanded with future improvements like using satellite images and experimenting with deep learning.

References

[1] *Banerjee, S., & Mondal, A. C. (2023). A region-wise weather data-based crop recommendation system using different machine learning algorithms. International Journal of Intelligent Systems and Applications in Engineering, 11(3). https://ijisae.org/index.php/IJISAE/article/view/3169*

[2] *Dahiphale, D., Shinde, P., Patil, K., & Dahiphale, V. (2023). Smart farming: Crop recommendation using machine learning with challenges and future ideas. TechRxiv. https://doi.org/10.36227/techrxiv.23504496.v1*

[3] *Kumar, A., & Singh, R. (2022). Soil analysis and crop recommendation using machine learning. Proceedings of th e 2022 International Conference on Smart Technologies and Agriculture (ICSTA), 45–50. https://ieeexplore.ieee.o rg/document/9725901*

[4] *Kasera, R. K. (2024). Comparative analysis of different machine learning based techniques for crop recommendation. In Advanced Computing, Machine Learning, Robotics and Internet Technologies, pp. 1–12. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-031-47221-3_1*

[5] *van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systema tic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.20 20.105709*

[6] *G., A. C., Karthick, S., & Samy, S. S. (2023). Predict the growth of crop yield using machine learning. Applied and Computational Engineering, 2, 276–282. https://doi.org/10.54254/2755-2721/2/20220639*

[7] *Gupta, Y., & Srivastava, G. (2024). Crop recommendation system for growing best suitable crop. International Journal of Science and Research Archive, 12(1), 2928–2936. https://doi.org/10.30574/ijsra.2024.12.1.1111*

[8] *Kiran, P. S., Abhinaya, G., Sruti, S., & Padhy, N. (2024). A machine learning-enabled system for crop recommendation. Engineering Proceedings, 67(1), 51. https://doi.org/10.3390/engproc2024067051*

[9] *AbouGrad, H., Chakhar, S., & Abubahia, A. (2023). Decision making by applying machine learning techniques to mitigate spam SMS attacks. In L. Troiano, A. Vaccaro, N. Kesswani, I. Díaz-Rodriguez, I. Brigui, & D. Pastor-Escuredo (Eds.), Key Digital Trends in Artificial Intelligence and Robotics: Proceedings of 4th International Conference on Deep Learning, Artificial Intelligence and Robotics (ICDLAIR) 2022, pp. 154–166. Springer, Cham. https://doi.org/10.1007/978-3-031-30396-8_14*

[10] *Senapaty, M. K., Ray, A., & Padhy, N. (2024). A decision support system for crop recommendation using machine learning classification algorithms. Agriculture, 14(8), 1256. https://doi.org/10.3390/agriculture14081256*

[11] *Mavi, H., Upadhyay, S. K., Srivastava, N., & Sharma, R. (2024). Crop recommendation system based on soil quali ty and environmental factors using machine learning. In Proceedings of the 2024 International Conference on Em*

DOI: 10.54254/2755-2721/2025.23348

erging Innovations and Advanced Computing (INNOCOMP), pp. 1–6. IEEE. https://ieeexplore.ieee.org/document/10664127

[12] *Abhinov, K., Saranya, K. S., Mahendra, M., & Babu, C. S. (2024). Soil-based crop recommendation system using machine learning. In Proceedings of the 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), pp. 1–6. IEEE. https://ieeexplore.ieee.org/document/10533537*