Formosa Journal of Science and Technology (FJST)
Vol. 4, No.1, 2025: 451-464

Use of Expert System in Determining Water Management in Agriculture with Methods: Fuzzy Logic and Rule-Based Decision Making

Saprudin[1], Perani Rosyani[2*], Bagas Mahendra Putra[3], Al Haura[4*], La Juanda[5], Vivi Ainun[6]
Faculty of Computer Science, Informatics Engineering, Pamulang University
Corresponding Author: Perani Rosyani, dosen000837@unpam.ac.id

A R T I C L E I N F O

Keywords: Expert System, Water Management, Agriculture, Fuzzy Logic, Rule-Based Decision Making

Received : 3, January
Revised : 17, January
Accepted: 31, January

A B S T R A C T

This project aims to develop an expert system to support agricultural water management using Fuzzy Logic and Rule-Based Decision Making methods. This system is important for improving agricultural yields and environmental sustainability, especially with the increasing demand for food and the impacts of climate change. Data was taken from Kaggle, including information on soil conditions, temperature, and rainfall. Data processing includes missing value removal, outlier detection, and splitting the data into 80% training and 20% testing. Fuzzy Logic was chosen because it is able to handle data uncertainty and provide accurate output regarding crop water requirements, while Rule-Based Decision Making utilizes expert knowledge-based rules for decision making. Simulation results show that the Fuzzy Logic model provides recommendations for water needs according to actual conditions, with high responsiveness to soil moisture and temperature. The system is expected to be a tool to assist farmers in decision-making, increase agricultural productivity, and support water sustainability. This research contributes to the development of expert systems in agriculture and natural resource management based on modern technology.

INTRODUCTION

Effective water management is key to successful agricultural production while maintaining the sustainability of natural resources. Challenges such as climate change, population growth and increasing food demand often make it difficult for farmers to determine appropriate water requirements. The uncertainty of factors such as crop type, soil condition, and weather further Complicates decision-making. To address this, expert systems based on fuzzy logic and rule-based Decision making offer innovative solutions. Fuzzy logic helps handle uncertainty in data by providing flexibility in decision making, while rule-based decision making provides specific recommendations based on rules derived from expert knowledge. Combining these two approaches creates a system that is more adaptive and responsive to conditions in the field. With proper implementation, this system can help farmers optimize water use, increase crop yields, reduce environmental impacts, and support the sustainability of the agricultural sector. This is an important strategy to address global challenges in water resource management in agriculture.

Research Objectives

This research aims to design an expert system model that provides recommendations for water management in agriculture using Fuzzy Logic and Rule-Based Decision making methods. The main focus is to improve accuracy in determining water requirements for various crops, as well as addressing uncertainty and data variability to make recommendations more appropriate to specific crop and environmental conditions.

One important goal is to provide farmers with relevant data-driven recommendations, including the timing and amount of water required for irrigation, to help them make more efficient decisions. The research also seeks to improve the efficiency of water resource use, which is expected to reduce operational costs and environmental impacts due to excessive water use.

In addition, this research supports sustainable agricultural practices by maintaining soil and ecosystem quality and ensuring the availability of water resources for the future. Raising farmers' awareness on the importance of good water management through training and application of expert systems is also a focus.

The research will analyze factors that affect crop water requirements, such as crop type, soil conditions and weather, to provide more accurate recommendations. Finally, evaluation and trials of the expert system under real conditions will be conducted to assess its effectiveness and obtain feedback for improvement.

Overall, the goal of this research is to develop and apply a Fuzzy Logic and Rule-Based Decision Making-based expert system in water management in the agricultural sector, with the hope of addressing the challenges farmers face, improving water use efficiency, and supporting sustainability in agricultural practices.

METHODOLOGY

Data Collection

Data collection is an important step in the development of a fuzzy logic-based expert system for water management in the agricultural sector. Appropriate and relevant data is the foundation for providing effective recommendations. The following are the main elements in data collection:

1. Types of Data Required:
   - Soil Moisture Data: Using sensors to directly measure soil moisture.
   - Weather Data: Includes temperature, rainfall, wind speed, and humidity, which can be obtained from meteorological stations or local sensors.
   - Crop Data: Information on crop types, growth phases, and specific water requirements for each type.
   - Soil Condition Data: Includes pH, texture, and the soil's capacity to hold water.
   - Historical Data: Information on rainfall and crop yields from previous years to analyze trends.

2. Data Collection Methods:
   - Sensors and Measuring Instruments: Install automated sensors for consistent measurements.
   - Field Survey: Conduct periodic manual measurements to supplement sensor data.
   - Meteorological Station: Use data from stations to obtain accurate weather information.
   - Geographic Information System (GIS): Collecting and analyzing spatial data.
   - Historical Data Collection: Accessing data from government agencies or agricultural databases.

3. Data Processing and Analysis:
   - Data Cleaning: Removing invalid or inconsistent data to maintain analysis accuracy.
   - Data Integration: Combining data from various sources to form a comprehensive database.
   - Data Analysis: Applying statistical techniques and fuzzy analysis to evaluate the data.

4. Data Validation:
   - Sensor Testing: Ensure that the sensors are functioning properly and providing accurate data.
   - Field Verification: Comparing data from sensors with manual measurements to ensure consistency.
   - Model Validity Test: Testing the fuzzy model that has been built to ensure the recommendations match the conditions in the field.
5. Data Storage and Management:
   - Centralized Database: Using a database system to store all data for easy access.

- Data Backup: Perform regular backups to protect data from loss.
- Limited Access: Set access rights for different users to maintain data confidentiality and integrity.

Data Source Identification

Identification of data sources is an important step in the development of a fuzzy logic-based expert system for water management in the agricultural sector. Relevant and accurate data sources are needed to produce appropriate and effective recommendations. Some categories of data sources that can be identified include primary, secondary, geospatial, historical, and the latest technology and innovations.

Primary data sources are data that are obtained directly from the field and have high relevance to the needs of the system. Examples include soil sensors installed at farm sites to measure soil moisture and temperature in real-time. This data is important for understanding crop irrigation needs. In addition, weather sensors that measure temperature, humidity, rainfall and wind speed also contribute to the collection of accurate local weather data. Field surveys by personnel to collect information on soil conditions and crop growth, as well as interviews with farmers regarding their farming practices and experiences, are also valuable primary data sources.

Secondary data sources involve data that has been collected by others and can be used for further analysis. Official meteorological stations are often a reliable source of weather data, providing information related to rainfall and temperature. Data from government and agricultural agencies, such as soil maps and crop yield statistics, are also very useful. Academic research that includes articles and reports on agriculture as well as agricultural databases containing information on different crop types and water requirements can be accessed to support analysis.

Geospatial data provides important information on the location and geographic distribution of water management. Geographic Information Systems (GIS) can be used to collect and analyze spatial data, such as soil and crop distribution maps. Topographic maps and satellite images also help in monitoring soil conditions and plant health as well as land use changes.

Historical data, which includes information from previous years, plays an important role in trend analysis. Records of previous weather and yield data can be used to forecast future weather patterns and crop yields. Documentation from farmers on farming practices and yields is also a valuable source of historical information.

With advances in technology, new sources of data are emerging that can be used in water management. Mobile applications to help farmers manage irrigation and get up-to-date weather information are gaining popularity. Big data platforms that collect and analyze data from various sources, including sensors and social media, provide deeper insights. The Internet of Things (IoT) enables connected devices to collect and exchange data in real-time, improving the monitoring of agricultural conditions.

By utilizing various data sources, both primary and secondary, as well as geospatial and historical data, fuzzy logic-based expert systems can provide more accurate and relevant recommendations. This will assist farmers in making better decisions, improve water use efficiency, and support agricultural sustainability. Through the integration of information from these sources, this technology can play an important role in meeting the challenges posed by climate change and the evolving needs of agriculture.

Data Collection

The importance of proper data collection is crucial in the development of fuzzy logic expert systems for water management in the agricultural sector. The success of this system depends on the quality, accuracy, and relevance of the data collected. Accurate data helps farmers determine the time and amount of water required for plants, with information on soil moisture, weather conditions, and the specific needs of plants being vital. With the right data, expert systems can provide relevant recommendations, allowing farmers to make better and timely decisions related to irrigation.

In addition, accurate data collection contributes to increased water use efficiency. Through in-depth analysis of water patterns, the system can help farmers understand the specific needs of plants and the environmental conditions that affect their growth. This understanding allows for the implementation of more efficient irrigation strategies, reducing water waste and supporting the sustainability of water resources. Amid challenges due to climate change and declining water quality, the management of these resources should be approached wisely.

Historical data collected over many years is also essential for effective water management. This data provides insights for trend analysis, allowing expert systems to predict future water needs more accurately. By utilizing existing weather patterns and plant growth data, farmers can plan irrigation and water resource management more effectively. In a dynamic farming environment, where many factors can affect water needs, appropriate and real-time data enables the system to provide responsive recommendations to changes, thereby reducing the risk of harvest failures.

The use of accurate data and sound analysis can also improve agricultural outputs. Information about soil conditions and plants allows farmers to optimize irrigation and fertilization practices, leading to healthier and more productive crops. Consequently, both the quality and quantity of yields increase. Furthermore, efficient water management based on accurate data can minimize negative environmental impacts by preventing pollution from excessive irrigation and maintaining ecosystem balance.

On the policy side, proper data collection supports the development of better agricultural policies. Information on water use, harvests, and environmental conditions can be utilized by policymakers to formulate strategies that support sustainable agricultural practices. With clear data, they can plan the management of water resources and channel assistance to farmers more precisely. Additionally, farmers' awareness of the importance of good water management can increase, encouraging them to be more open to education and training in sustainable agricultural practices.

Overall, the right data collection is key to the successful implementation of fuzzy logic expert systems. Accurate data enables the system to provide effective recommendations, improve water use efficiency, and support environmental sustainability. Given the increasingly complex challenges posed by climate change and the growing demands of agriculture, the importance of accurate data collection cannot be overlooked. Through effective data integration, technology can assist farmers in managing water resources more efficiently, contributing to better agricultural outcomes and more sustainable practices.

Evaluation of Data Sources

Evaluation of data sources is a crucial step in the development of fuzzy logic expert systems for water management in the agricultural sector. With the multitude of available data sources, it is important to assess their reliability, relevance, and quality. This process ensures that the data used in decision-making is accurate and that the expert systems can provide valid and effective recommendations for farmers.

Reliability is the main aspect of evaluating data sources. Data from non-trusted or unverified sources can lead to false recommendations, which can harm both farmers and the environment. Therefore, it is essential to check the credibility of the sources, such as institutions or organizations that collect the data. Official sources, such as meteorological bodies or agricultural research institutes, are generally more reliable than unofficial sources. Additionally, the methods used for data collection should also be evaluated to ensure they adhere to recognized standards.

Relevance is another important factor in the evaluation process. Irrelevant or inappropriate data in the context of water management can produce less accurate recommendations. Thus, the collected data should include specific and up-to-date information regarding soil moisture, rainfall, temperature, and other factors that influence plant water needs. Data gathered from sensors or measurement tools placed in the field tends to be more relevant compared to data from more general secondary sources.

Data quality is also a critical aspect of evaluation. High-quality data should be free of errors and maintain good integrity, meaning that it should be consistent and accurate over time. Validation and verification techniques, such as cross-checking with other sources, can help ensure the quality of the data. Furthermore, the frequency of data updates should be considered; regularly updated data provides more accurate information about current conditions.

Cost considerations and accessibility are also part of the evaluation of data sources. In effective water management, it is essential to use quality data sources that are accessible. Expensive or difficult-to-access data sources may not be practical, especially for small farmers with limited resources. As such, finding a balance between quality, relevance, and cost in the selection of data sources is important.

Finally, the evaluation of data sources should be conducted on a sustainable basis. The agricultural environment and weather conditions are constantly changing, making it vital to continuously evaluate and update the data sources used in expert systems. Periodic evaluations allow the system to remain relevant and effective in providing recommendations tailored to the needs of farmers, as well as identifying new and better data sources.

Overall, evaluating data sources is a complex yet vital process in the development of fuzzy logic expert systems for water management. By ensuring that the data sources used are reliable, relevant, high-quality, and accessible, expert systems can offer more accurate and effective recommendations. Given the challenges faced by the agricultural sector due to climate change, the evaluation of data sources becomes increasingly important to help farmers make better decisions, improve water efficiency, and support sustainable agricultural practices.

RESULTS AND DISCUSSION

Preprocessing in Data Analysis and Machine Learning

Preprocessing is an essential phase in data analysis and the development of machine learning models. The main purpose of preprocessing is to prepare raw data for model training, involving key steps aimed at cleaning, organizing, and optimizing the data to improve the quality of analysis and prediction results. Through precise preprocessing, we can reduce bias, enhance accuracy, and ensure that the built model is reliable when applied to new data.

The first step in preprocessing involves reading data from sources, such as CSV files, and loading it into a canonical format, like a DataFrame using the Pandas library. This facilitates further analysis of the data. Once the data is loaded, we can examine the first few rows to gain an early understanding of the dataset's structure, including column names, data types, and existing values, which are crucial for understanding the characteristics of the data to be analyzed.

Next, the process of feature and target selection is performed, where relevant columns are chosen as features (independent variables) to predict the target (dependent variable). Appropriate feature selection significantly influences model performance since irrelevant features can diminish the model's effectiveness in producing accurate predictions. The subsequent stage addresses missing data, involving the identification and handling of missing values in the dataset. Common approaches include deleting rows with missing values or filling them using methods such as the average, median, or mode. Effectively handling missing data is vital for maintaining the integrity of the analysis and preventing issues when running the model.

Following this, normalization or data scaling is conducted to ensure all features are within the same range. Normalization helps improve model performance, particularly in algorithms sensitive to scale, such as K-Nearest Neighbors or Support Vector Machines. The final step involves splitting the data into two subsets: training data and test data. This division is important for testing the model's performance on previously unseen data, thereby avoiding overfitting and ensuring that the model can generalize well.

By executing these steps, preprocessing ensures that the dataset used is clean, structured, and relevant, which is crucial for building an accurate and reliable model capable of providing better analysis and prediction results.

Steps in Data Preprocessing for Machine Learning

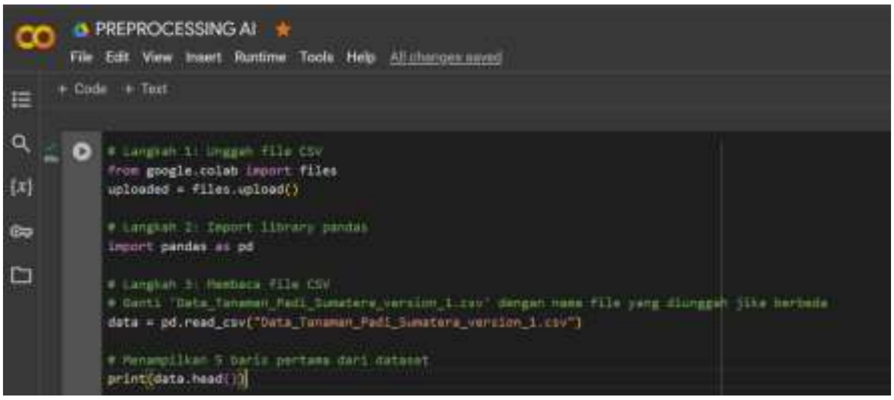
1. Reading Data: The preprocessing process begins by reading data from CSV files using the Pandas library. The dataset is loaded into a DataFrame format, a tabular structure that simplifies analysis. Once the data is loaded, we can display the first few rows to gain an initial overview of the dataset's structure, including column names, data types, and existing values, which are essential for understanding the data characteristics.
2. Selecting Features and Targets: After acquiring a preliminary understanding of the data, the next step is to identify the features (independent variables) and the target (dependent variable). In this context, relevant features include columns such as harvest, rainfall, moisture, and average temperature, which will be used to predict the production values. Proper selection of features is crucial, as irrelevant or redundant features can negatively impact the model's predictive accuracy.
3. Splitting Data into Training and Test Sets: Following the identification of features and targets, the dataset is divided into two subsets: training data (80%) and test data (20%). This division is accomplished using the train_test_split function, ensuring that the model is trained on one set of data and evaluated on a different set. This approach helps prevent overfitting, where the model becomes too tailored to the training data and fails to perform well on new data.
4. Handling Missing Data: The next step involves checking for and addressing missing data, which can significantly hinder analysis and reduce model accuracy. We need to identify any missing values, and if found, we can choose from several approaches: deleting rows with missing values, filling them with the average, median, or mode, or applying interpolation methods. The choice of method depends on the nature and extent of the missing data.
5. Normalization or Data Scaling: Once the data is cleaned, normalization or scaling is performed if necessary. Normalization ensures that all features are on a similar scale, which can enhance model performance, especially in algorithms sensitive to data scale, such as K-Nearest Neighbors or Support Vector Machines. Common normalization techniques include min-max scaling, which transforms each feature to a range of [0, 1], and standardization, which adjusts the data to have a mean of 0 and a standard deviation of 1.

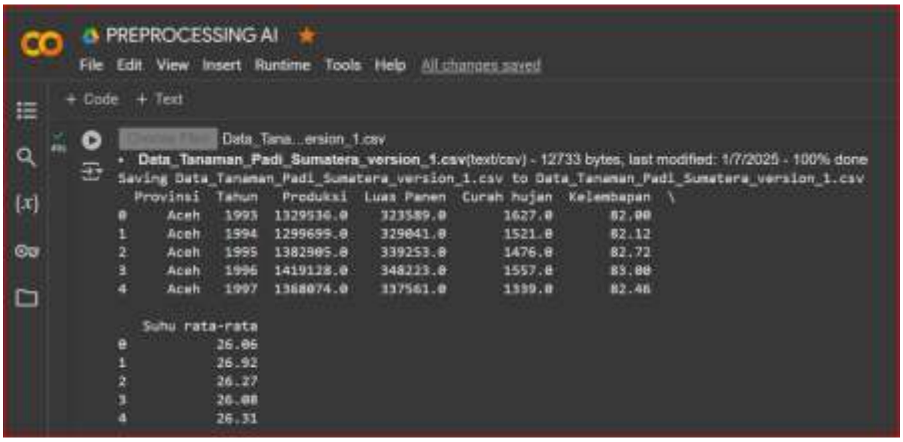
Results of Preprocessing

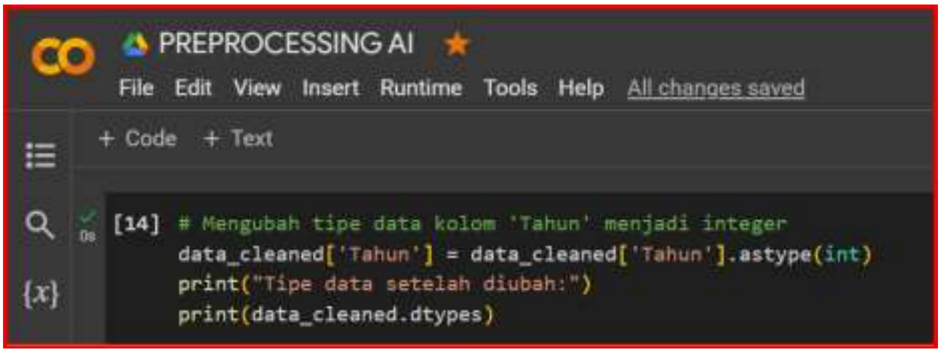
The results of the preprocessing phase yield a structured, clean dataset that is ready for model development. After going through all the steps outlined above, we have several key elements:

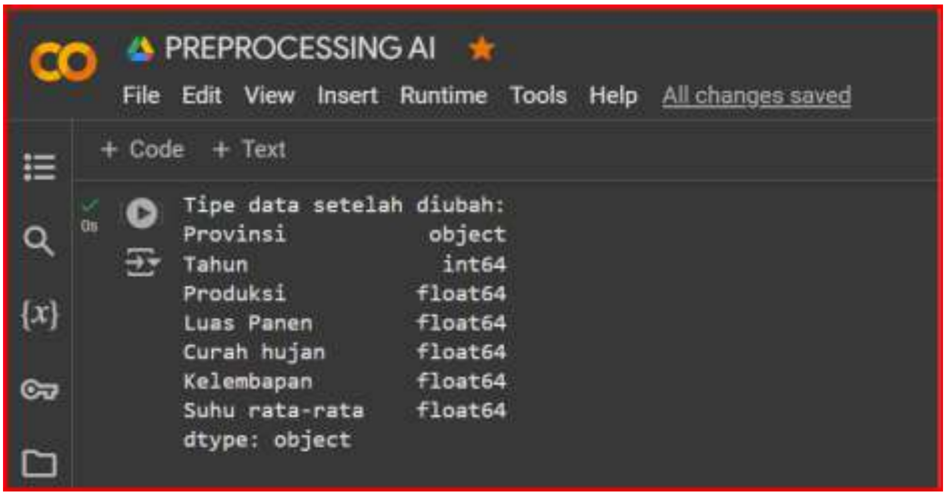
1. Initial Data Table: A table displaying the first few rows of the dataset, providing an initial overview of the available features and their corresponding values.
2. Feature and Target Variables: Two main variables, X for features and y for the target, have been selected based on their relevance.
3. Training and Test Data: The data has been split into training and test sets, with variables X_train, X_test, y_train, dan y_test, prepared for model training and evaluation.
4. Clean and Standardized Data: If the steps for handling missing data and normalization have been applied, the result is a clean dataset with standardized features, enabling the model to learn more efficiently.

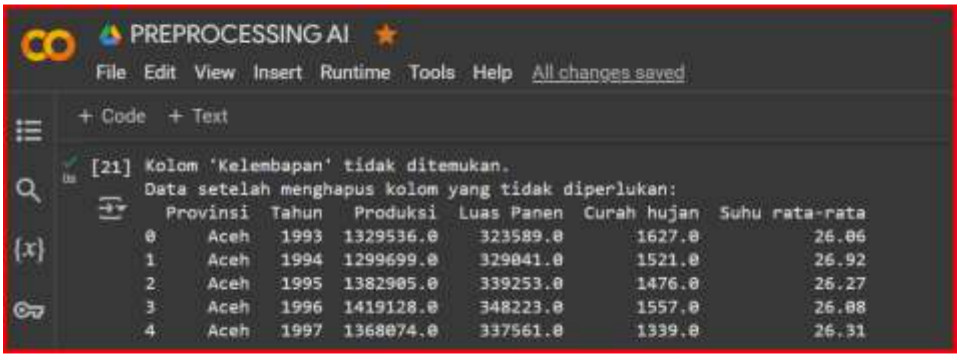
Overall, the preprocessing stage is crucial in ensuring that the data used for analysis is of high quality and relevant to the analytical goals. With a well-processed dataset, we can enhance the accuracy and reliability potential of the developed model, leading to more accurate results when used for predictions in real-world scenarios.





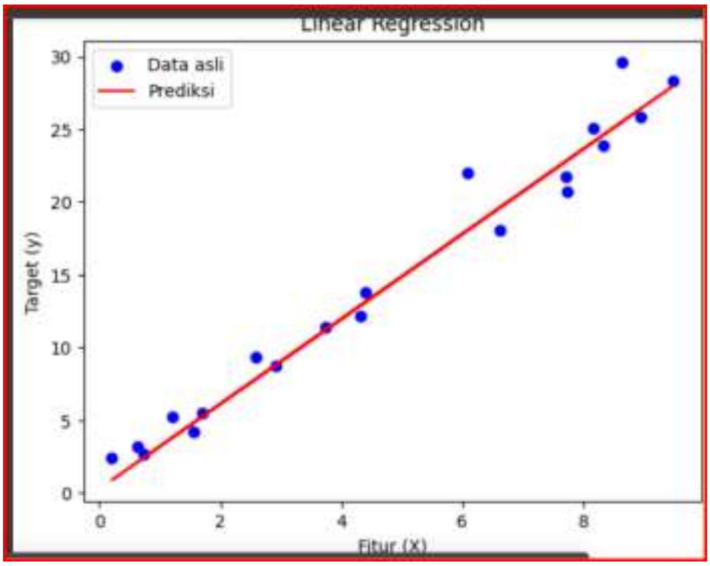

PREPROCESSING AI
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[14] # Mengubah tipe data kolom 'Tahun' menjadi integer
data_cleaned['Tahun'] = data_cleaned['Tahun'].astype(int)
print("Tipe data setelah diubah:")
print(data_cleaned.dtypes)

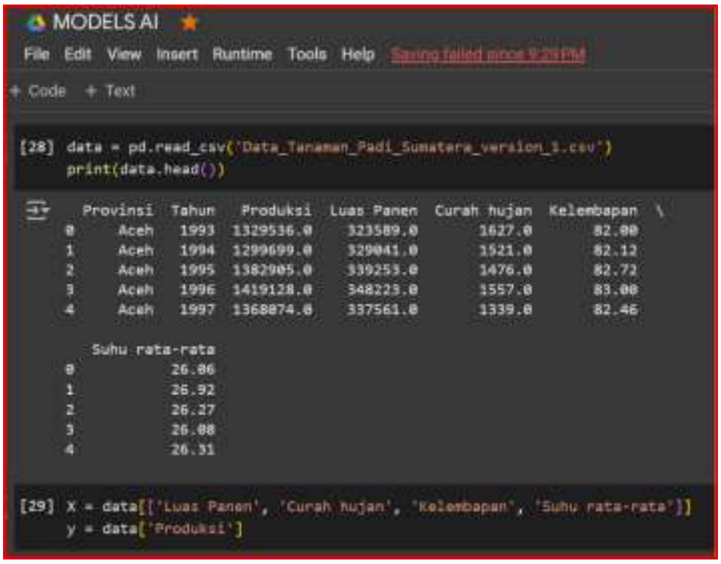

PREPROCESSING AI
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
Tipe data setelah diubah:
Provinsi object
Tahun int64
Produksi float64
Luas Panen float64
Curah hujan float64
Kelembapan float64
Suhu rata-rata float64
dtype: object

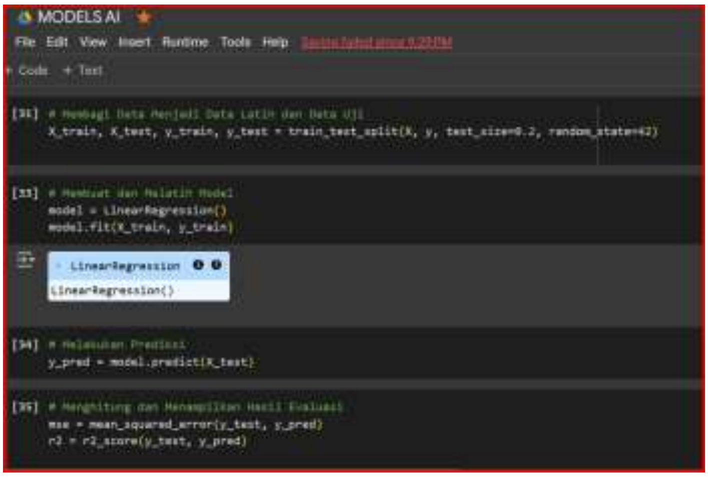

PREPROCESSING AI
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[21] Kolom 'Kelembapan' tidak ditemukan.
Data setelah menghapus kolom yang tidak diperlukan:
Provinsi Tahun Produksi Luas Panen Curah hujan Suhu rata-rata
0 Aceh 1993 1329536.0 323589.0 1627.0 26.06
1 Aceh 1994 1299699.0 329041.0 1521.0 26.92
2 Aceh 1995 1382905.0 339253.0 1476.0 26.27
3 Aceh 1996 1419128.0 348223.0 1557.0 26.08
4 Aceh 1997 1368074.0 337561.0 1339.0 26.31

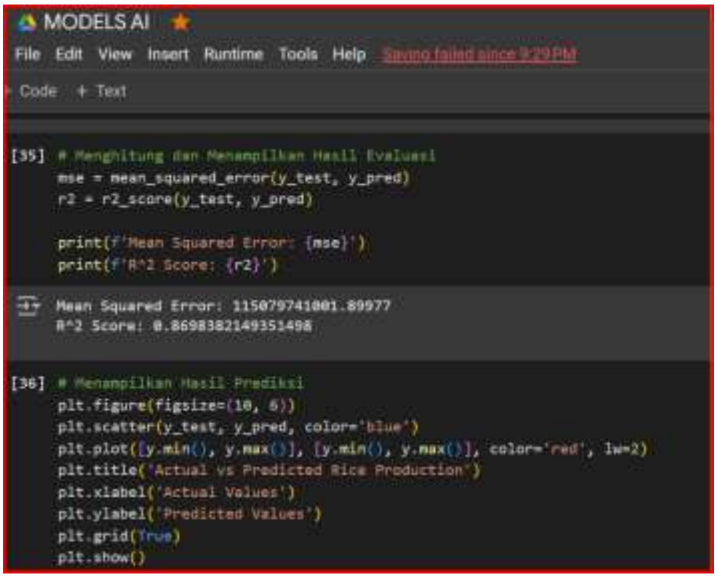

PREPROCESSING AI
File Edit View Insert Runtime Tools Help All changes saved
Code + Text
Share

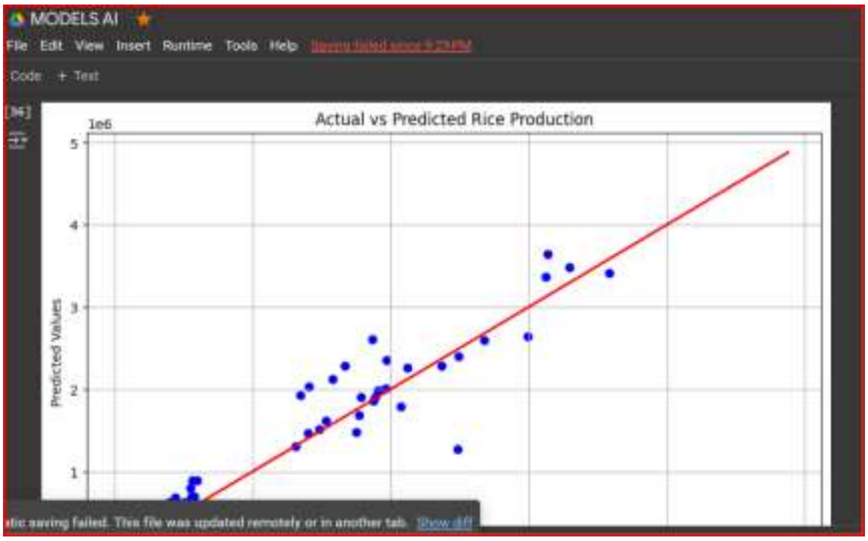

PREPROCESSING AI
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[23] # Menyimpan data yang sudah diproses ke file CSV baru
data_cleaned.to_csv('Data_Tanaman_Padi_Sumatera_Cleaned.csv', index=False)
print("Data yang sudah diproses telah disimpan.")
Data yang sudah diproses telah disimpan.


Linear Regression
Data asli
Prediksi
Target (y)
Fitur (X)
0
5
10
15
20
25
30
0
2
4
6
8


MODELS AI
File Edit View Insert Runtime Tools Help
+ Code + Text
[28] data = pd.read_csv('Data_Tanaman_Padi_Sumatera_version_1.csv')
print(data.head())
Provinsi Tahun Produksi Luas Panen Curah hujan Kelembapan \
0 Aceh 1993 1329536.0 323589.0 1627.0 82.00
1 Aceh 1994 1299699.0 329041.0 1521.0 82.12
2 Aceh 1995 1382905.0 339253.0 1476.0 82.72
3 Aceh 1996 1419128.0 348223.0 1557.0 83.00
4 Aceh 1997 1368074.0 337561.0 1339.0 82.46
Suhu rata-rata
0 26.06
1 26.92
2 26.27
3 26.08
4 26.31
[29] X = data[['Luas Panen', 'Curah hujan', 'Kelembapan', 'Suhu rata-rata']]
y = data['Produksi']


MODELS AI
File Edit View Insert Runtime Tools Help
+ Code + Text
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
model = LinearRegression()
model.fit(X_train, y_train)
LinearRegression
LinearRegression()
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)





CONCLUSIONS

The use of expert systems for water management in agriculture through Fuzzy Logic and Rule-Based Decision Making has shown highly positive results in enhancing water resource efficiency. Effective water management is crucial in agriculture for maintaining sustainable yields and ecosystem balance. By applying Fuzzy Logic, the expert system can address uncertainties commonly encountered in the field, such as weather changes, soil moisture variations, and specific plant needs. This enables the system to provide more flexible and relevant recommendations tailored to the actual conditions faced by farmers.

From the perspective of Rule-Based Decision Making, this approach offers a clear structure for the decision-making process. The rules established based on prior data and experiences aid the system in delivering consistent and reliable recommendations. By combining these two methods, the expert system can offer comprehensive solutions for water management in the agricultural sector. The findings from this research indicate that the implementation of the expert system not only boosts agricultural productivity but also supports more sustainable farming practices by reducing water waste and minimizing environmental impacts.

RECOMMENDATIONS

1. System Development and Improvement: It is recommended to continually develop and update the expert system by incorporating additional variables that affect water needs. Variables such as soil quality, crop types, and microclimate parameters should be included to enhance the accuracy of the generated recommendations. Further research is also needed to evaluate how climate change may influence irrigation needs and to adjust the system to evolving conditions.
2. User Training: Providing comprehensive training for farmers and users of the system is essential for optimal utilization of this technology. Training programs should cover technical and practical aspects, including how to interpret system outputs, implement recommendations, and adjust farming practices based on the acquired data. Training could also involve simulations of system usage in real-life situations, allowing users to learn in a more interactive manner.
3. Integration with Modern Technology: Integrating the expert system with modern technologies such as soil sensors, drones, and automated irrigation systems can enhance water management effectiveness. Sensors can provide real-time data on soil moisture and weather conditions, enriching the inputs for the expert system. This way, decisions made will be faster, more precise, and responsive to changing field conditions.
4. Evaluation and Testing: Conducting evaluations and trials of the system across various locations and types of agriculture is crucial to ensure system reliability. Field trials can help identify practical challenges users may face. Additionally, feedback from users can provide valuable insights for system improvements. Long-term research involving observations across multiple growing seasons can also yield important data regarding system performance over time.
5. Collaboration with Agricultural Experts: Partnering with agricultural specialists and environmental scientists in the development of this expert system is highly recommended. Such collaboration will ensure that the recommendations generated are not only data-driven but also aligned with best practices in sustainable agriculture. Regular meetings between system developers and farmers can foster constructive dialogue to provide support and better understand farmers' needs.
6. Development of a Robust Database: Establishing a comprehensive database is vital for supporting effective decision-making. This database should include historical information on soil, crops, weather, and agricultural yields. With sufficient data, the models used in the expert system can be trained and periodically updated to improve the accuracy and relevance of recommendations.

FURTHER STUDY

Future research is expected to further explore this material.

REFERENCES

Bhandari, H., & Yadav, R. K. (2020). Role of government in promoting sustainable agriculture: A review. Sustainable Agriculture Research, 9(1), 1-10. doi:10.5539/sar.v9n1p1

Giarratano, J. C., & Riley, G. (2005). Expert Systems: Principles and Programming. Cengage Learning.

Goyal, S. K., & Gupta, R. (2013). Database management systems in agriculture: A comprehensive review. International Journal of Computer Applications, 77(5), 1-5. doi:10.5120/12856-1205

Khoshgoftaar, T. M., & Allen, E. B. (2001). A study of the effectiveness of fuzzy logic in predicting software project success. Journal of Systems and Software, 56(1), 1-12. doi:10.1016/S0164-1212(00)00069-0

Michalak, J., & Zawadzki, J. (2019). The role of interdisciplinary collaboration in agricultural research and development. Agricultural Systems, 177, 102-111. doi:10.1016/j.agsy.2019.02.006

Shafiei, S. M., & Shafiei, M. (2018). Development of expert systems in agriculture: A review. Computers and Electronics in Agriculture, 151, 92-106. doi:10.1016/j.compag.2018.05.013

Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338-353. doi:10.1016/S0019-9958(65)90241-X

Zhang, Y., & Wang, J. (2016). Smart agriculture: A review of trends and challenges. Computers and Electronics in Agriculture, 130, 1-12. doi:10.1016/j.compag.2016.06.001