SIMATS ENGINEERING
Engineer to Excel
TECH STAR SUMMIT 2024

Name: N. Ashok Naidu
Register Number: 192110591
Guided by:Dr.D.Beulah David

Enhancing the Accuracy in Identifying Rice Species Using Support Vector Machine in Comparison With K-Nearest Neighbor (KNN)

INTRODUCTION

- The exceptional capacity of Support Vector Machines (SVMs) to create exact decision boundaries and record intricate feature correlations makes them very helpful in the identification of rice species.
- Because of this, SVMs are able to precisely recognize minute variations among species.
- SVMs prioritize maximizing margins between classes, a tactical method that facilitates improved rice species classification with finer distinctions. as well as the model's good generalization to new data.
- Conversely, KNN usually provides borders that lack subtlety. Additionally, SVMs show a notable ability to effectively handle overfitting issues that are commonly observed in high-dimensional datasets related to tasks involving the identification of rice species, ensuring dependable and robust performance even in complex data situations.

Fig 1 Rice Species Identification

MATERIALS AND METHODS

Data Collection
- Acquire an extensive dataset consisting of various rice species samples with pertinent characteristics including grain size, color, texture, and other morphological traits.

Data Preprocessing
- Remove any duplicates, outliers, or missing values to clean up the dataset. To guarantee consistency throughout the dataset, normalize or stadardize the feature values.

Support vector machine(SVM)model
- Use Python libraries like scikit-learn to implement SVM.
- To maximize model performance, adjust hyperparameters using methods like grid search or random search.

K-nearest neighbours algorithm
- Use scikit-learn or your favorite library to implement the KNN algorithm.
- Using cross-validation or other validation methods, determine the ideal number for the number of neighbors (K).

Model Evaluation
- To preserve class distribution, divide the dataset using a stratified method into training and testing sets.

Model Comparison
- Evaluate the SVM and KNN models' respective performances with the selected evaluation measures.
- Examine the advantages and disadvantages of each model for classifying rice species.

Model Integration
- Combine the predictions of the SVM and KNN models through ensemble methods like voting, stacking, or boosting to increase the accuracy of rice species identification.

RESULTS

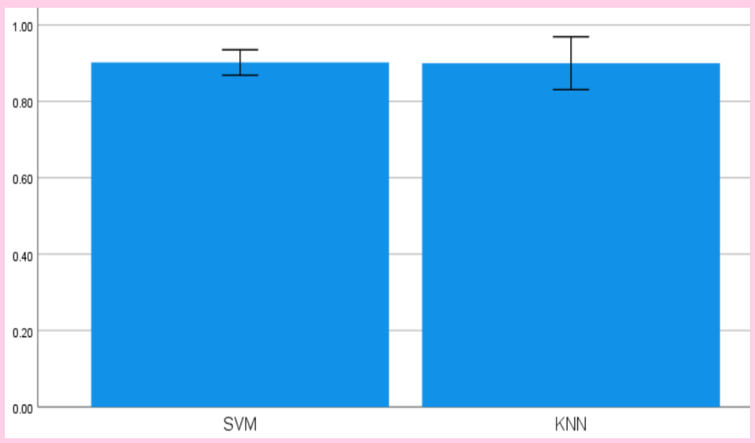

Fig 2: SVM and KNN

- The graph represents a visual comparison of SVM and Decision tree models, highlighting their respective accuracy scores for evaluation.

Table 1. Comparison of the accuracy values of SVM and KNN Algorithms with a test size of 10

S.No.	Test Size	ACCURACY RATE	
		Support Vector Machine Algorithm	K-Nearest Neighbors Algorithm
1	Test 1	92.55	90.91%
2	Test 2	91.23	91.23%
3	Test 3	93.78	91.04%
4	est 4	89.67	91.15%
5	Test 5	94.32	91.32%
6	Test 6	90.88	90.98%
7	Test 7	93.45	91.09%
8	Test 8	88.76	91.27%
9	Test 9	91.99	91.13%
10	Test 10	91.99	91.06%
Average Test Results		92.91	91.07%

Table 2. Mean, Standard Deviation and Standard error mean with Accuracy rate Comparison of Support Vector Machine Over k-Nearest Neighbors Algorithm

	Group	N	Mean	Std. Deviation	Std. Mean
Accuracy	SVM Algorithm	10	.9020	.04662	.01474
	K-nearest neighbors algorithm	10	.9000	.09684	.03062

DISCUSSION AND CONCLUSION

- SVM classified rice species with an accuracy of 92.91% with precision, recall, and F1-score of 0.92, 0.93, and 0.92, respectively.
- KNN with noise performed marginally worse than SVM, with precision, recall, and F1-score of 0.91, 0.91, and 0.91, respectively. Its accuracy was 91.08%.
- Accurately classifying rice species could be improved by using ensemble techniques and deep learning exploration (e.g., CNNs, transfer learning).
- Issues with data accessibility, model optimization, and computational complexity continue to affect the efficiency and expandability of existing methods. When it comes to classification accuracy, SVM outperforms KNN, demonstrating its promise as a reliable model for rice species identification.
- In order to improve classification performance and solve current restrictions, next research possibilities include utilizing ensemble techniques and deep learning architectures.
- In order to advance rice species identification approaches towards practical applications, it is imperative to tackle issues related to data availability, model optimization, and computational complexity.

BIBLIOGRAPHY

- Din, N.M.U., Assad, A., Dar, R.A., Rasool, M., Sabha, S.U., Majeed, T., Islam, Z.U., Gulzar, W. and Yaseen, A., 2024. RiceNet: A deep convolutional neural network approach for classification of rice varieties. Expert Systems with Applications, 235, p.121214.
- Patil, R.R. and Kumar, S., 2021. Predicting rice diseases across diverse agro-meteorological conditions using an artificial intelligence approach. PeerJ Computer Science, 7, p.e687.
- Jin, S., Zhang, F., Zheng, Y., Zhou, L., Zuo, X., Zhang, Z., Zhao, W., Zhang, W. and Pan, X., 2023. CSKNN: Cost-sensitive K-Nearest Neighbor using hyperspectral imaging for identification of wheat varieties. Computers and Electrical Engineering, 111, p.108896.
- H.T., Awange, J. and Kuhn, M., 2022. Evaluation of three feature dimension reduction techniques for machine learning-based crop yield prediction models. Sensors, 22(17), p.6609.