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 Decision Tree
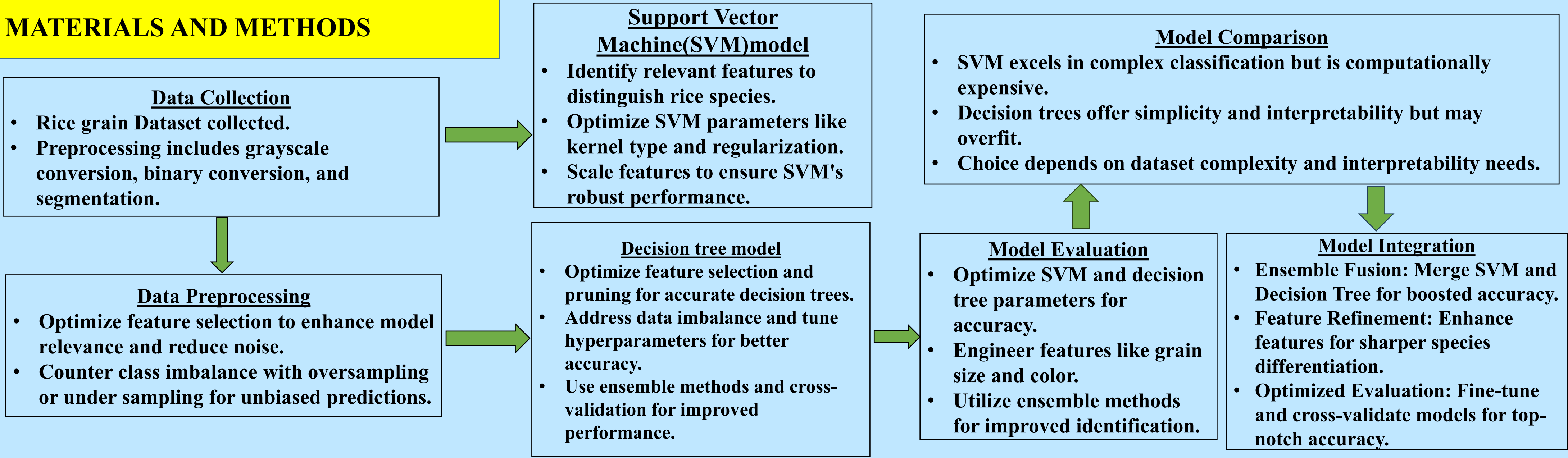
INTRODUCTION

➤ Support Vector Machines (SVMs) are particularly useful in the field of rice species identification because of their extraordinary ability to define precise decision boundaries and capture complex feature linkages. This allows SVMs to accurately identify small changes between species.

➤ SVMs emphasize the maximization of margins between classes, a strategic approach that helps the classification of rice species with finer distinctions and improve. As the model's ability to generalize well to unseen data. Decision Trees, on the other hand, frequently give borders deficient in nuance.

➤ Moreover, SVMs demonstrate a noteworthy capacity to efficiently address overfitting problems that are frequently seen in high-dimensional datasets associated with tasks involving the identification of rice species, guaranteeing dependable and resilient performance even in intricate data settings.

Fig 1 Rice species identification

MATERIALS AND METHODS

Data Collection
- Rice grain Dataset collected.
- Preprocessing includes grayscale conversion, binary conversion, and segmentation.

Data Preprocessing
- Optimize feature selection to enhance model relevance and reduce noise.
- Counter class imbalance with oversampling or under sampling for unbiased predictions.

Support Vector Machine(SVM)model
- Identify relevant features to distinguish rice species.
- Optimize SVM parameters like kernel type and regularization.
- Scale features to ensure SVM's robust performance.

Decision tree model
- Optimize feature selection and pruning for accurate decision trees.
- Address data imbalance and tune hyperparameters for better accuracy.
- Use ensemble methods and cross-validation for improved performance.

Model Evaluation
- Optimize SVM and decision tree parameters for accuracy.
- Engineer features like grain size and color.
- Utilize ensemble methods for improved identification.

Model Comparison
- SVM excels in complex classification but is computationally expensive.
- Decision trees offer simplicity and interpretability but may overfit.
- Choice depends on dataset complexity and interpretability needs.

Model Integration
- Ensemble Fusion: Merge SVM and Decision Tree for boosted accuracy.
- Feature Refinement: Enhance features for sharper species differentiation.
- Optimized Evaluation: Fine-tune and cross-validate models for top-notch accuracy.
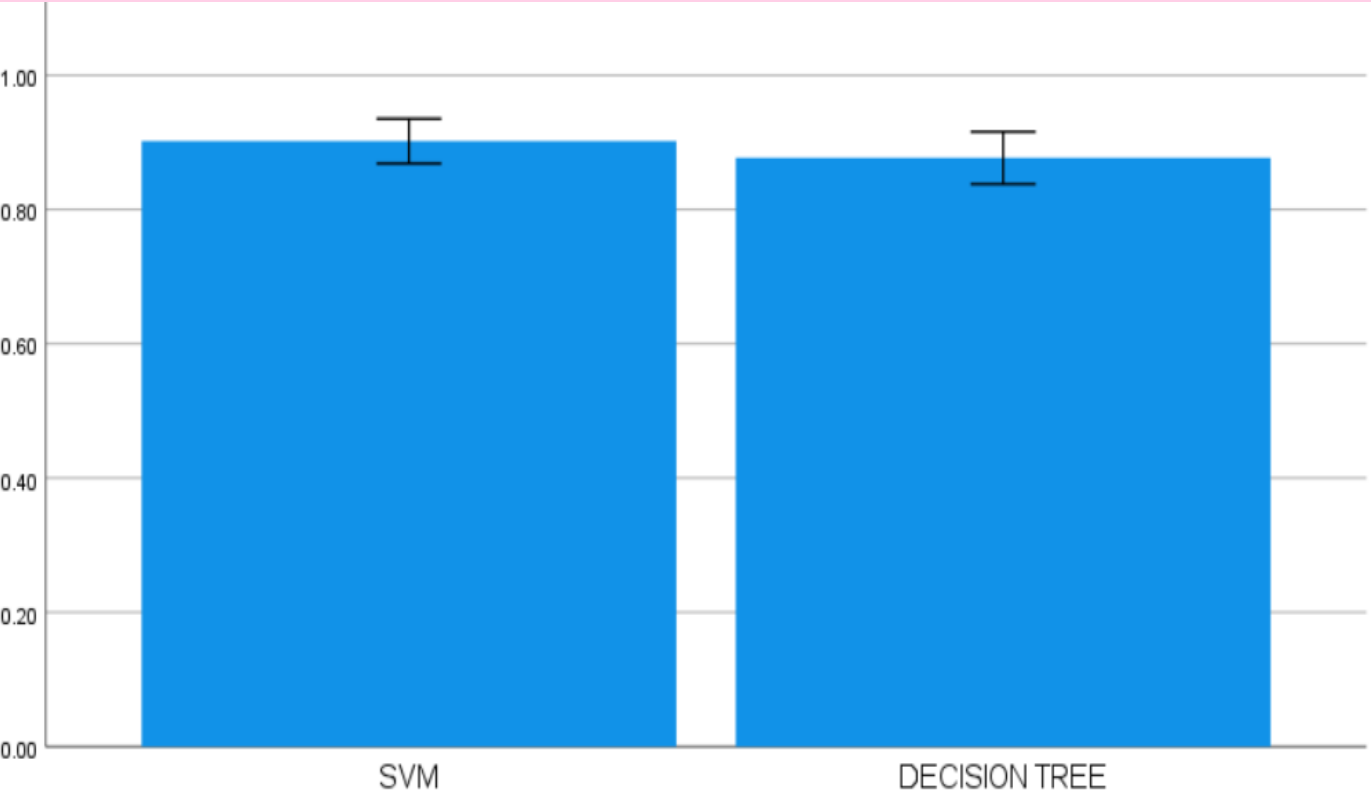
RESULTS

Fig 2: SVM and Decision Tree

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 DT Algorithms with a test size of 10

S.No	Test Size	ACCURACY RATE	
		Support Vector Machine Algorithm	Decision Tree Algorithm
1	Test 1	92.55	88.96
2	Test 2	91.23	86.83
3	Test 3	93.78	85.92
4	est 4	89.67	79.61
5	Test 5	94.32	87.91
6	Test 6	90.88	91.93
7	Test 7	93.45	90.07
8	Test 8	88.76	86.73
9	Test 9	91.99	89.86
10	Test 10	91.99	87.91
Average Test Results		92.91	87.79

Table 2. Mean, Standard Deviation and Standard error mean with Accuracy rate Comparison of Support Vector Machine over Decision Tree Algorithm

	Group	N	Mean	Std. Deviation	Std. Mean
Accuracy	SVM Algorithm	10	.9020	.04662	.01474
	Decision tree Algorithm	10	.8770	.05458	.01726

DISCUSSION AND CONCLUSION

➤ Considering an accuracy of 92.91% as opposed to 87.80%, SVM fared better than Decision Tree in the identification of rice species.

➤ For the "Cammeo" and "Osmancik" rice species, SVM demonstrated superior precision, recall, and F1-score, demonstrating its efficacy in classifying tasks.

➤ While offering respectable accuracy, Decision Tree's precision and recall for several rice species were lower than SVM's, which may have shown the tree's limitations in identifying intricate patterns in the data.

➤ SVM outperformed Decision Tree in a dataset with complex correlations between characteristics and rice species, which may have been attributed to its robustness in handling non-linear decision boundaries.

➤ Overall, SVM shows itself to be the most suitable model for precisely identifying rice species, providing superior classification outcomes and proving its applicability for improving accuracy in this domain.

BIBLIOGRAPHY

➤ Saxena, P., Priya, K., Goel, S., Aggarwal, P.K., Sinha, A. and Jain, P., 2022. Rice varieties classification using machine learning algorithms. Journal of Pharmaceutical Negative Results, pp.3762-3772.

➤ Alfred, R., Obit, J.H., Chin, C.P.Y., Haviluddin, H. and Lim, Y., 2021. Towards paddy rice smart farming: a review on big data, machine learning, and rice production tasks. Ieee Access, 9, pp.50358-50380.

➤ Patil, N.S., 2021. Identification of paddy leaf diseases using evolutionary and machine learning methods. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(2), pp.1672-1686.

➤ Son, N.T., Chen, C.F., Chen, C.R. and Minh, V.Q., 2018. Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. Geocarto international, 33(6), pp.587-601.