**Title: A Comparative Analysis of Machine Learning Models for Classification Tasks**

**Abstract** This report presents a comparative evaluation of various machine learning models applied to a classification problem. The models analyzed include Decision Trees, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Random Forest, XGBoost, Logistic Regression, and Neural Networks. Each model's performance is assessed using key evaluation metrics such as accuracy, precision, recall, F1 score, and area under the curve (AUC). The results highlight the strengths and weaknesses of each model, providing insights into their applicability for different classification tasks.

**Introduction** Machine learning plays a critical role in solving classification problems across various domains. Selecting the best model requires evaluating multiple models using performance metrics that reflect real-world applicability. This study compares traditional machine learning models such as Decision Trees, Logistic Regression, and KNN with more advanced models like Random Forest, XGBoost, and Neural Networks. The objective is to identify the most effective model for our dataset based on multiple evaluation criteria.

**Materials and Methods** The dataset used for this study consists of labeled instances divided into training and testing subsets. The models were trained and tested on the same dataset to ensure a fair comparison. The metrics used for evaluation include:

* **Accuracy:** Measures the proportion of correctly classified instances.
* **Precision:** Evaluates the proportion of true positives among predicted positives.
* **Recall:** Determines the proportion of true positives among actual positives.
* **F1 Score:** Harmonic mean of precision and recall.
* **AUC (Area Under Curve):** Represents the ability of the model to distinguish between classes.

The machine learning models were implemented using Python with the Scikit-learn and TensorFlow libraries. Each model was tuned for optimal performance before evaluation.

**Results** The table below summarizes the performance of each model based on the computed metrics.

| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **AUC** |
| --- | --- | --- | --- | --- | --- |
| Decision Trees | 0.9816 | 0.9786 | 0.9765 | 0.9775 | 0.9808 |
| KNN | 0.9886 | 0.9848 | 0.9875 | 0.9861 | 0.9994 |
| SVM | 0.9944 | 0.9957 | 0.9905 | 0.9931 | N/A |
| Random Forest | 0.9901 | 0.9872 | 0.9887 | 0.9879 | 0.9997 |
| XGBoost | 0.9931 | 0.9900 | 0.9933 | 0.9916 | 0.9998 |
| Logistic Regression | 0.6085 | 0.0 | 0.0 | 0.0 | 0.4983 |
| Neural Networks | 0.9971 | 0.9933 | 0.9994 | 0.9963 | 0.9975 |

**A graph of a graph showing different colored bars

AI-generated content may be incorrect.**

**A red and blue graph

AI-generated content may be incorrect.**

**Discussion of Results**

From the results, Neural Networks emerged as the best-performing model, achieving the highest accuracy, recall, and F1 score. This indicates its ability to learn complex patterns in the dataset. XGBoost and Random Forest also demonstrated strong performance, with high accuracy and AUC values, making them reliable choices for classification tasks.

KNN and SVM also exhibited strong predictive performance, with SVM showing particularly high precision. However, one SVM variation performed poorly, likely due to improper hyperparameter selection. Decision Trees performed well but were slightly outperformed by ensemble-based models, which leverage multiple decision trees to improve generalization.

Logistic Regression, on the other hand, performed significantly worse than all other models. The zero precision, recall, and F1 score indicate that it failed to correctly classify positive instances. This suggests that the dataset is not linearly separable, making Logistic Regression an unsuitable choice for this classification task.

**Conclusion** The evaluation of different machine learning models provided valuable insights into their effectiveness for classification tasks. Neural Networks emerged as the best-performing model in terms of accuracy and overall metric performance. XGBoost and Random Forest also demonstrated high effectiveness, making them suitable for complex classification problems. Simpler models like Logistic Regression performed poorly, indicating their limitations for this dataset. These findings suggest that for high-dimensional and complex classification problems, ensemble methods and deep learning models should be preferred over traditional approaches.

Future work may involve testing additional models, tuning hyperparameters more extensively, and applying different feature engineering techniques to enhance performance further.