Project Report – Model Selection and Comparative Analysis

Name: G S S SURYA PRAKASH
Student ID: PES2UG23CS192
Course Name: Machine Learning
Submission Date: 01-09-2025

1. Introduction

The objective of this lab was to implement and compare manual hyperparameter tuning with scikit-learn's GridSearchCV. The project explored how grid search and cross-validation help in selecting optimal models. Multiple classification algorithms were tested, and results were analyzed using metrics such as Accuracy, Precision, Recall, F1-Score, and ROC AUC. Both manual and automated approaches were applied to various datasets, and results were compared to understand trade-offs.

2. Dataset Description

We used one dataset: HR Attrition Dataset

- Instances: 1470
- Features: 46 employee-related features
- Target: Whether the employee left (Attrition: Yes/No)

3. Methodology

- Hyperparameter Tuning:
  Systematic search over hyperparameter values to optimize model performance.
- Grid Search:
  Tests all parameter combinations exhaustively.
- K-Fold Cross-Validation:
  Splits data into *k* folds to ensure robust evaluation (used 5-fold).

Pipeline Used:

- StandardScaler → normalize features
- SelectKBest → feature selection
- Classifier → Decision Tree, KNN, Logistic Regression

Implementation Approaches:

- Part 1 (Manual): Explicit grid search with nested cross-validation and manual selection of best parameters.
- Part 2 (scikit-learn): Used GridSearchCV for automation with n_jobs=-1 for faster parallel computation.

4. Results and Analysis

Manual Implementation – Best Models:

- Decision Tree → Accuracy: 0.796, AUC: 0.653
- KNN → Accuracy: 0.819, AUC: 0.724
- Logistic Regression → Accuracy: 0.880, AUC: 0.818
- Voting Classifier → Accuracy: 0.844, AUC: 0.791

Scikit-learn Implementation – Best Models:

- Decision Tree → Accuracy: 0.796, AUC: 0.653
- KNN → Accuracy: 0.819, AUC: 0.724
- Logistic Regression → Accuracy: 0.880, AUC: 0.818
- Voting Classifier → Accuracy: 0.841, AUC: 0.791

Analysis: Results were nearly identical between manual and built-in approaches, with only minor differences due to implementation details and randomness in splits.

Visualizations:

- ROC Curves confirmed Logistic Regression as the strongest individual classifier.
- Confusion Matrices showed it balanced precision and recall better than others.

Best Models Across Datasets

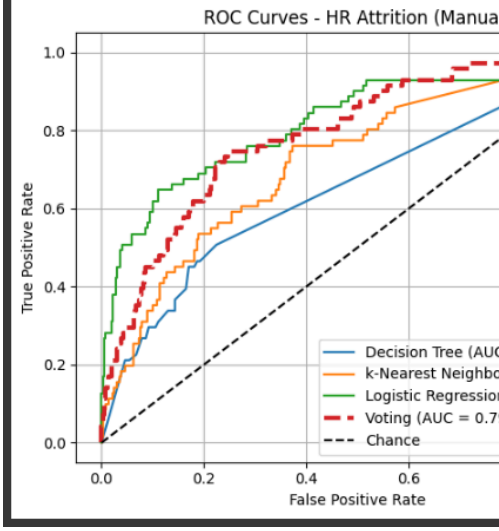
- HR Attrition: Logistic Regression again outperformed others with highest AUC.

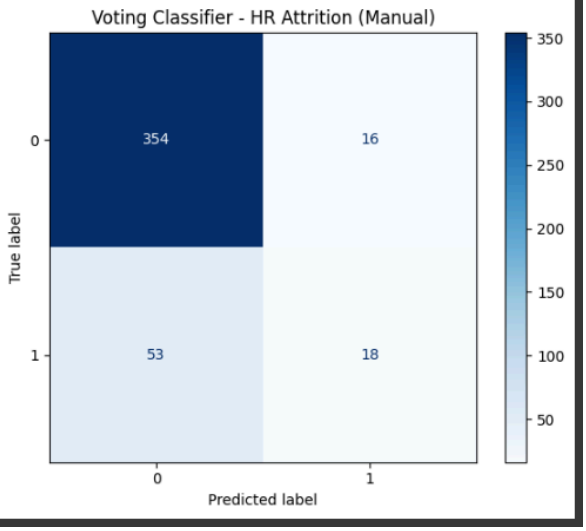
5. Screenshots

Manual Implementation:

```
============================================================
EVALUATING MANUAL MODELS FOR HR ATTRITION
============================================================

--- Individual Model Performance ---

Decision Tree:
  Accuracy: 0.7959
  Precision: 0.3492
  Recall: 0.3099
  F1-Score: 0.3284
  ROC AUC: 0.6528

k-Nearest Neighbors:
  Accuracy: 0.8186
  Precision: 0.3784
  Recall: 0.1972
  F1-Score: 0.2593
  ROC AUC: 0.7236

Logistic Regression:
  Accuracy: 0.8798
  Precision: 0.7368
  Recall: 0.3944
  F1-Score: 0.5138
  ROC AUC: 0.8177
```

```
Voting Classifier Performance:
  Accuracy: 0.8435, Precision: 0.5294
  Recall: 0.2535, F1: 0.3429, AUC: 0.7906
```

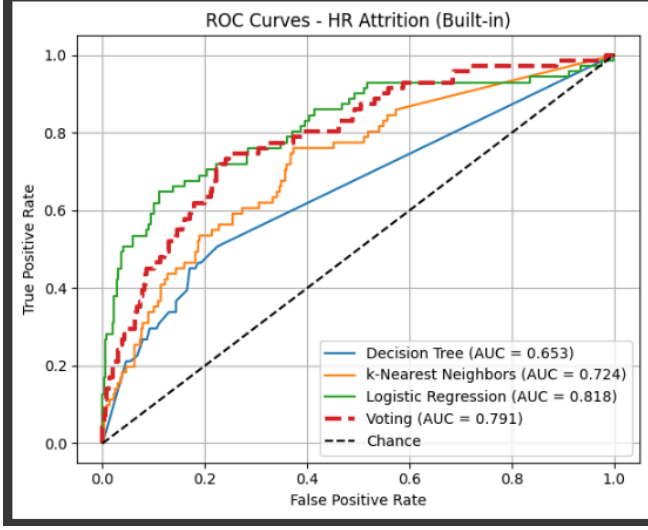


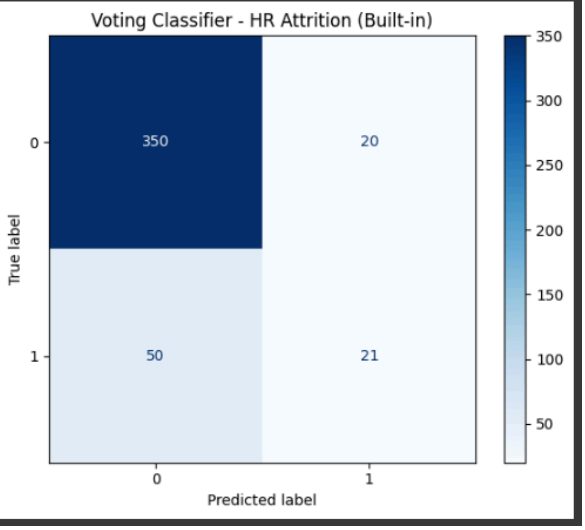


Scikit-Learn Implementation:

```
============================================================
EVALUATING BUILT-IN MODELS FOR HR ATTRITION
============================================================

--- Individual Model Performance ---

Decision Tree:
  Accuracy: 0.7959
  Precision: 0.3492
  Recall: 0.3099
  F1-Score: 0.3284
  ROC AUC: 0.6528

k-Nearest Neighbors:
  Accuracy: 0.8186
  Precision: 0.3784
  Recall: 0.1972
  F1-Score: 0.2593
  ROC AUC: 0.7236

Logistic Regression:
  Accuracy: 0.8798
  Precision: 0.7368
  Recall: 0.3944
  F1-Score: 0.5138
  ROC AUC: 0.8177
```

```
--- Built-in Voting Classifier ---
Voting Classifier Performance:
  Accuracy: 0.8413, Precision: 0.5122
  Recall: 0.2958, F1: 0.3750, AUC: 0.7906
```





6. Conclusion

This lab demonstrated the power of hyperparameter tuning and systematic model selection.

- Key Findings: Logistic Regression consistently outperformed others, while Voting Classifiers provided stability.
- Manual vs. scikit-learn: Manual tuning was more verbose but gave deeper insight into the process, whereas GridSearchCV simplified experimentation and was computationally efficient.
- Takeaway: Model performance depends heavily on proper hyperparameter tuning, and using library tools like scikit-learn makes the process practical without decreasing the accuracy.