Model Evaluation Report

Evaluation Metrics and Results

Naïve Bayes

- **Accuracy:** 0.67
- **Mean CV Accuracy:** 0.68
- **ROC-AUC:** 0.75
- **Macro Avg F1-Score:** 0.67
- **Cross-Validation Scores:** [0.6693, 0.6779, 0.6779, 0.6981, 0.6878]

Classification Report:

Metric	Class 0	Class 1	Macro Avg	Weighted Avg
Precision	0.63	0.73	0.68	0.68
Recall	0.76	0.60	0.68	0.67
F1-Score	0.69	0.66	0.67	0.67

Top 5 Most Discriminative Features:

Feature	Importance
Başvurular süreç	3.284
İçtihadın oluştuğu	2.967
Başvurunun ıçtihadın	2.966
Alana değerlendirilerek	2.958
İncelenmesine formu	2.955

Logistic Regression

- **Accuracy:** 0.78
- **Macro Avg F1-Score:** 0.78

Classification Report:

Metric	Class 0	Class 1	Macro Avg	Weighted Avg
Precision	0.74	0.82	0.78	0.78
Recall	0.82	0.74	0.78	0.78
F1-Score	0.78	0.78	0.78	0.78

Decision Tree

- **Accuracy:** 0.735
- **Mean CV Accuracy:** 0.731
- **ROC-AUC:** 0.832
- **Macro Avg F1-Score:** 0.73

Cross-Validation Scores: [0.7212, 0.7155, 0.7286, 0.7397, 0.7520]

Classification Report:

Metric	Class 0	Class 1	Macro Avg	Weighted Avg
Precision	0.79	0.70	0.75	0.75
Recall	0.59	0.86	0.73	0.74
F1-Score	0.68	0.77	0.73	0.73

Random Forest

- **Accuracy:** 0.78
- **Mean CV Accuracy:** 0.766
- **Macro Avg F1-Score:** 0.78

Cross-Validation Scores: [0.7406, 0.7686, 0.7733, 0.7854, 0.7618]

Test Set Performance:

Metric	Class 0	Class 1	Macro Avg	Weighted Avg
Precision	0.73	0.85	0.79	0.79
Recall	0.86	0.71	0.79	0.78
F1-Score	0.79	0.77	0.78	0.78

Turkish BERT

- **Accuracy:** 0.801
- **Macro Avg F1-Score:** 0.80

Classification Report:

Metric	Class 0	Class 1	Macro Avg	Weighted Avg
Precision	0.80	0.81	0.80	0.80
Recall	0.78	0.82	0.80	0.80
F1-Score	0.79	0.81	0.80	0.80

Comparison Between Models

1. **Turkish BERT** demonstrated the best overall performance, with the highest accuracy (80.1%) and consistent metrics, making it the most suitable for this task.
2. **Logistic Regression** and **Random Forest** achieved comparable results (accuracy of 78%), reflecting strong performance in simpler and ensemble-based models.
3. **Decision Tree** had a high ROC-AUC score (0.832) but showed slightly lower accuracy (73.5%), indicating possible overfitting.
4. **Naïve Bayes** showed the lowest performance metrics (67% accuracy), affected by its simplifying assumptions about feature independence.

Optimization Steps Taken

1. **Feature Engineering:**
    - Pre-processed textual data.
    - Identified the most discriminative features for Naïve Bayes.
2. **Hyperparameter Tuning:**
    - Optimized key parameters for Decision Tree and Random Forest models to improve cross-validation accuracy.
3. **Pre-trained Weights:**
    - Leveraged pre-trained Turkish BERT and fine-tuned it for the dataset to enhance predictive accuracy.
4. **Cross-validation:**
    - Applied cross-validation techniques to ensure robust and generalized model evaluation.