Ian Sam
V00990790

# STAT 469 Assignment #2

Introduction

This report evaluates the performance of four predictive models Logistic Regression, LDA, Classification Tree, and Elastic Net on binary resistance outcomes for five drugs (ABC, 3TC, AZT, D4T, DDI). The objective is to evaluate the effectiveness of these models using F1-scores and to determine whether performance differences between them are statistically significant. The evaluation is performed using 50 random splits with 5-fold stratified cross-validation to ensure balanced representation of resistance and non-resistance cases. Pairwise Wilcoxon signed-rank tests are used to assess whether differences in F1-scores between LDA and the newly added models Classification Tree and Elastic Net are statistically significant.

Methods Used

The models used in this study include Logistic Regression, LDA, Classification Tree, and Elastic Net. Logistic Regression is implemented using the glm function with a binomial family. LDA is performed using the lda function from the MASS package. Classification Tree is implemented using the rpart function. Elastic Net is applied using the cv.glmnet function with alpha set to 0.5. Each model is trained and evaluated using stratified 5-fold cross-validation repeated across 50 different random splits.

Results

After training and testing the models, I analyzed their performance across different drugs. The median F1-scores for each model and drug were computed to compare their effectiveness. To better visualize the results, I generated boxplots showing the distribution of F1-scores for each model.

F1 Scores for ABC

F1 Scores for 3TC

F1 Scores for AZT

F1 Scores for D4T
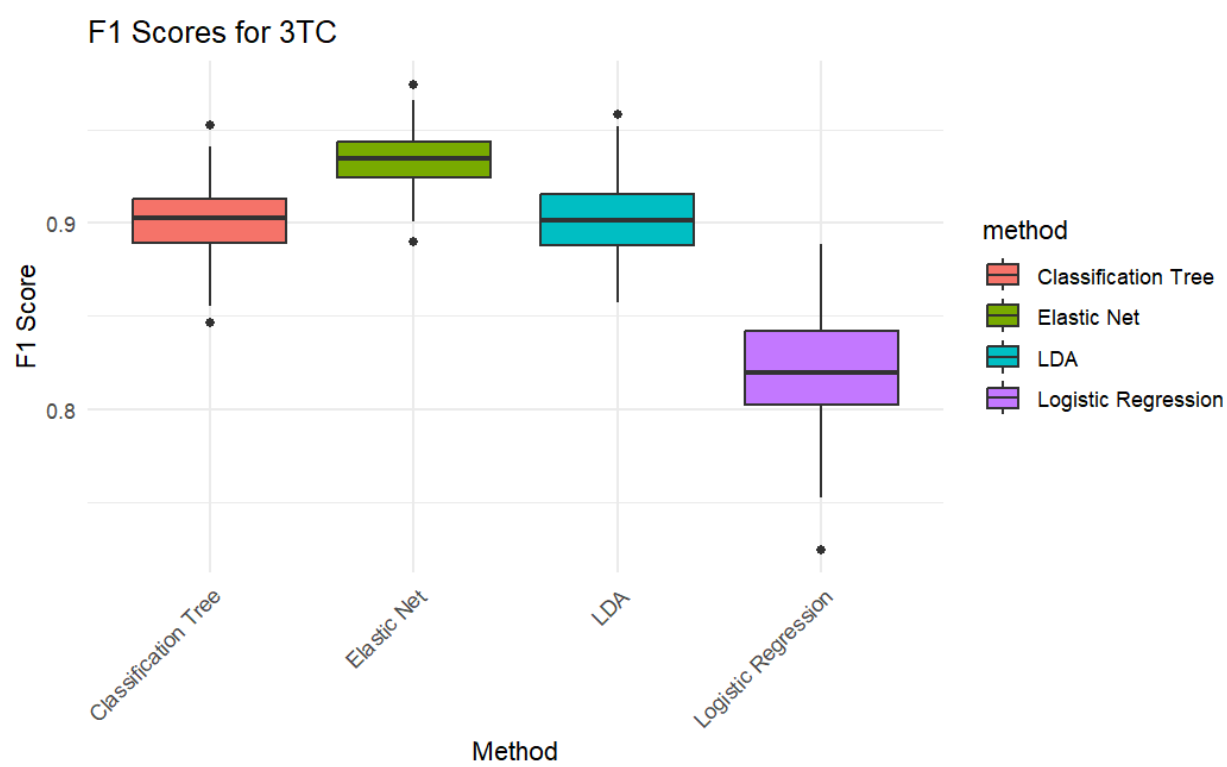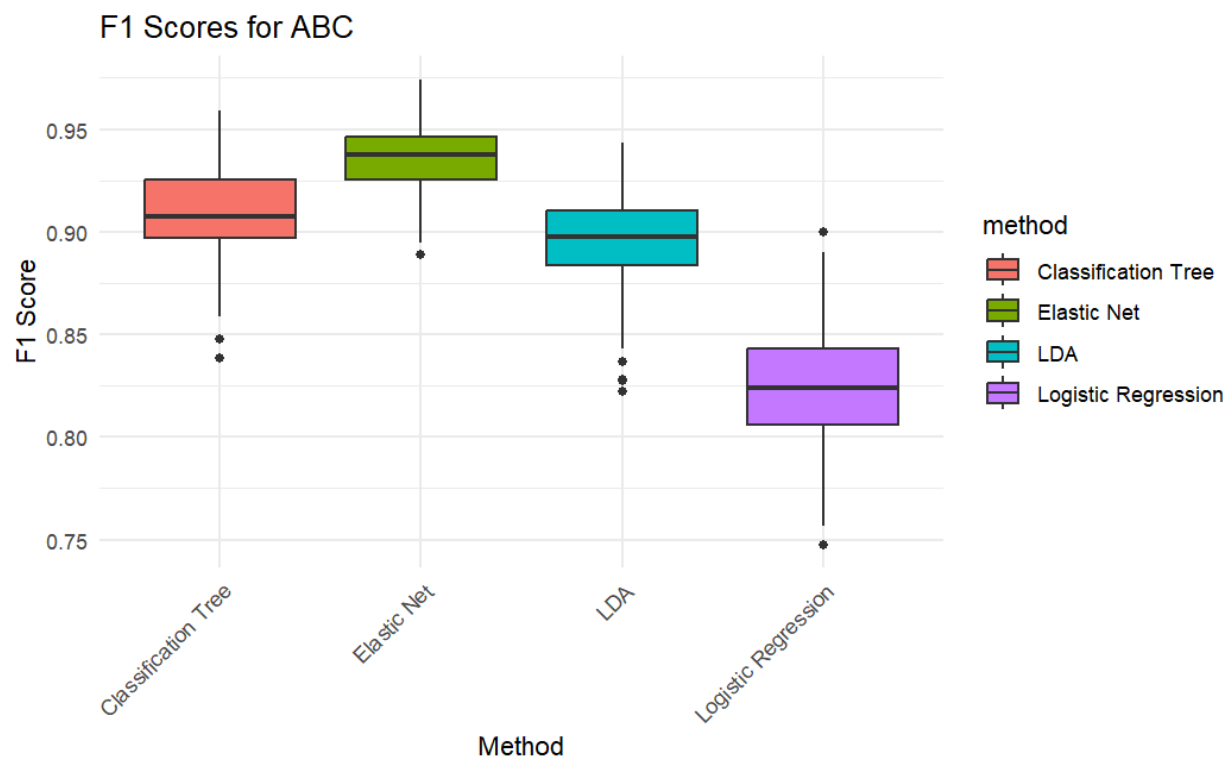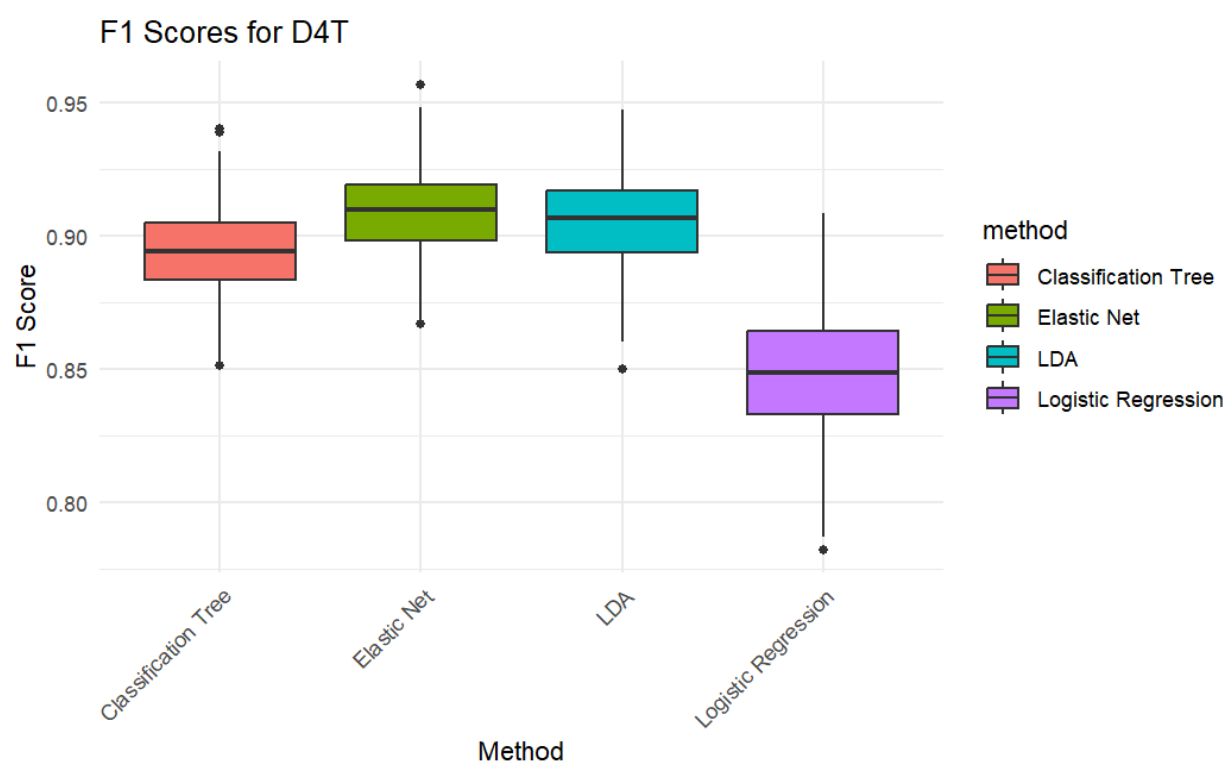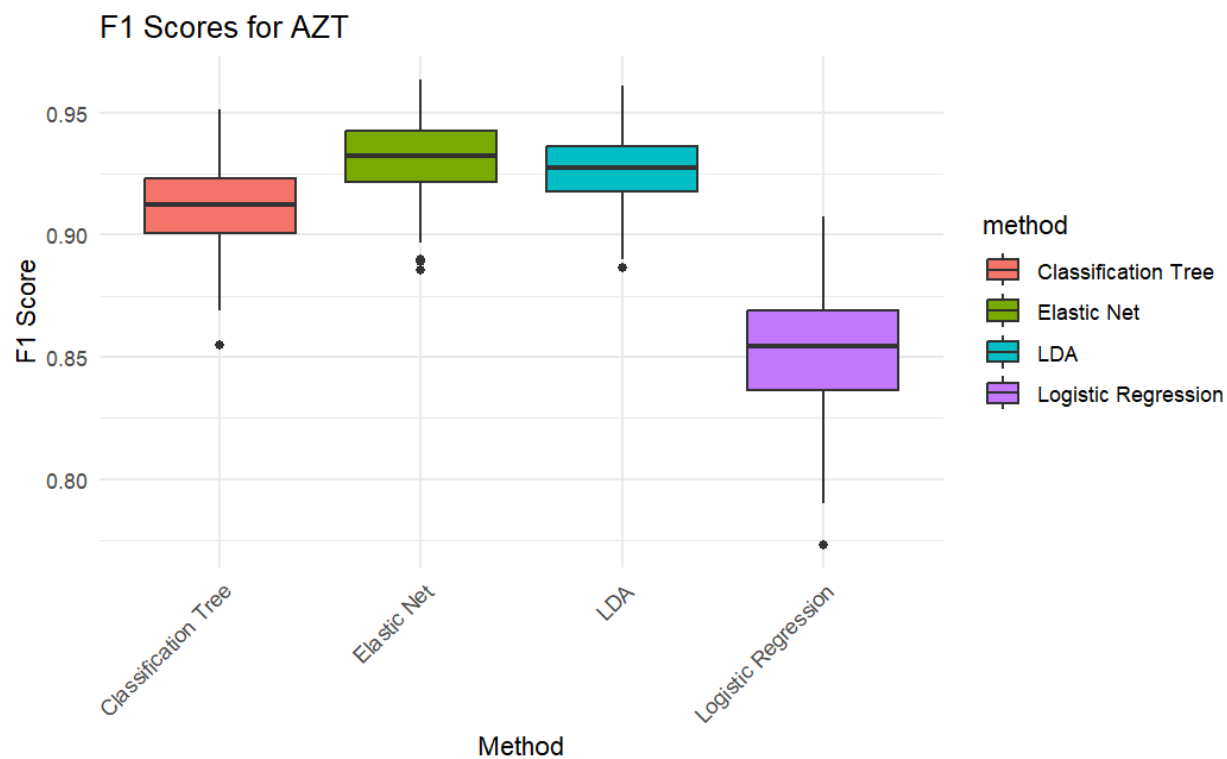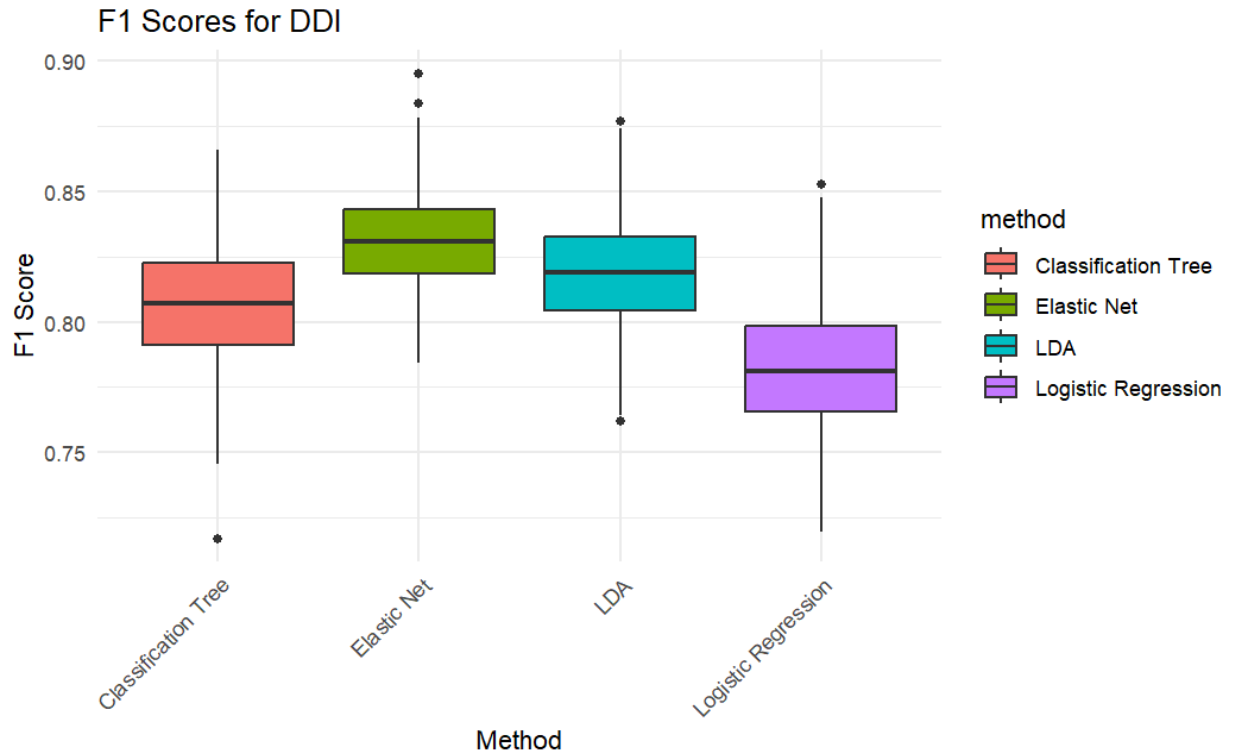
F1 Scores for DDI

Performance Comparison using F1 Scores

The boxplots below illustrate the distribution of F1-scores for Logistic Regression, LDA, Classification Tree, and Elastic Net across the five drugs. (F1 Score Boxplots for Drugs ABC, 3TC, AZT, D4T, DDI)

From the boxplots, it is evident that Elastic Net consistently achieves higher median F1-scores compared to other models across most drugs. LDA performs well overall but is outperformed by Elastic Net in several cases. Classification Tree shows competitive performance but exhibits more variability. Logistic Regression generally underperforms relative to the other models.

Statistical Significance of Model Comparisons

To assess whether the observed differences in performance between LDA and the newly added models were statistically significant, I conducted pairwise Wilcoxon signed-rank tests. The p-values for each drug comparison are shown below:

```
Drug: ABC

LDA vs Classification Tree: p-value = 6.399258e-16

LDA vs Elastic Net: p-value = 1.050063e-42

Drug: 3TC
```

LDA vs Classification Tree: p-value = 0.9222751

LDA vs Elastic Net: p-value = 1.357335e-42

Drug: AZT

LDA vs Classification Tree: p-value = 1.849046e-32

LDA vs Elastic Net: p-value = 2.647664e-09

Drug: D4T

LDA vs Classification Tree: p-value = 1.249758e-18

LDA vs Elastic Net: p-value = 7.910461e-08

Drug: DDI

LDA vs Classification Tree: p-value = 1.858506e-13

LDA vs Elastic Net: p-value = 1.715503e-19

The results indicate that Elastic Net significantly outperforms LDA across all drugs. Classification Tree also significantly outperforms LDA for most drugs except for 3TC.

Drug ABC

method <chr>	misclassification_rate <dbl>	precision <dbl>	recall <dbl>	f1_score <dbl>
Classification Tree	0.07214458	0.9152781	0.9095455	0.9072165
Elastic Net	0.04819277	0.9368351	0.9387755	0.9376495
LDA	0.08433735	0.8611111	0.9393939	0.8975610
Logistic Regression	0.14112903	0.8019335	0.8465463	0.8241758

Drug 3TC

method <chr>	misclassification_rate <dbl>	precision <dbl>	recall <dbl>	f1_score <dbl>
Classification Tree	0.08016064	0.8930305	0.9100000	0.9027837
Elastic Net	0.05220884	0.9266055	0.9439252	0.9346734
LDA	0.08433735	0.8644068	0.9484536	0.9015971
Logistic Regression	0.14859438	0.8049745	0.8398113	0.8196660

Drug AZT

method <chr>	misclassification_rate <dbl>	precision <dbl>	recall <dbl>	f1_score <dbl>
Classification Tree	0.09638554	0.9788732	0.8546332	0.9122125
Elastic Net	0.07661290	0.9479547	0.9225352	0.9324084
LDA	0.08400000	0.9418174	0.9151933	0.9271599
Logistic Regression	0.16867470	0.8541667	0.8536568	0.8542710

Drug D4T

method <chr>	misclassification_rate <dbl>	precision <dbl>	recall <dbl>	f1_score <dbl>
Classification Tree	0.1204819	0.9141741	0.8797163	0.8941980
Elastic Net	0.1048387	0.9096998	0.9115646	0.9096990
LDA	0.1120000	0.8989646	0.9128976	0.9064707
Logistic Regression	0.1774194	0.8480510	0.8482759	0.8486797

Drug DDI

method <chr>	misclassification_rate <dbl>	precision <dbl>	recall <dbl>	f1_score <dbl>
Classification Tree	0.1770631	0.8942308	0.7380952	0.8070209
Elastic Net	0.1726908	0.8181818	0.8412698	0.8307692
LDA	0.1867419	0.8049814	0.8333333	0.8187152
Logistic Regression	0.2200000	0.7829457	0.7857143	0.7810632

Best Model Based on F1-Scores

Elastic Net is the best-performing model for all drugs based on median F1-scores.

Best Model Based on Precision

Classification Tree achieves the highest precision for three drugs (AZT, D4T, and DDI), while Elastic Net excels in precision for ABC and 3TC.

Best Model Based on Misclassification Rate

Elastic Net achieves the lowest misclassification rate for four drugs, whereas LDA is the best for AZT.

Best Model Based on Recall

Elastic Net is the top model across all drugs for recall.

Conclusion

Based on the comprehensive evaluation, Elastic Net consistently emerged as the best-performing model across multiple metrics, including F1-score, recall, and misclassification rate. Classification Tree performed well in precision and demonstrated competitive performance

across certain drugs. LDA was generally outperformed by Elastic Net and Classification Tree. Logistic Regression was the least effective across all drugs. These findings suggest that Elastic Net is the most reliable model for predicting HIV drug resistance outcomes.