TML Assignment 1 Report

Team 3 - Hoang Nguyen Minh, Shreyansh Tripathi

May 28, 2025

1 Introduction

This report presents our approach and findings for Assignment 1 of the TML course. Our primary objective was to evaluate and compare the performance of multiple machine learning models and identify the most effective one for adversarial attack tasks.

2 Modeling Approach

To evaluate various machine learning models for attack prediction, we compared four approaches: MLP, XGBoost, CatBoost, and AdaBoost. The **CatBoost model achieved the highest ROC-AUC score of 0.6645**, outperforming the others and was therefore selected as the attacker model. Its success can be attributed to its **efficient handling of categorical features**, robust regularization strategies, and optimized tree structures. These characteristics enabled CatBoost to generalize well while avoiding overfitting, making it particularly suitable for this task.

XGBoost followed closely with a ROC-AUC of 0.6613, demonstrating strong generalization capabilities through its gradient boosting framework. However, unlike CatBoost, XGBoost requires more manual preprocessing for categorical variables, which may have slightly limited its performance. The **MLP**, serving as a baseline deep learning model, attained a ROC-AUC of 0.6474. While MLPs can capture non-linear patterns, they often underperform on tabular data unless extensively tuned or provided with large datasets.

Finally, **AdaBoost** scored the lowest with a ROC-AUC of 0.6440. As a simpler ensemble method, it may have struggled with complex data relationships and shown greater sensitivity to noise. Overall, the comparison highlights **CatBoost's superior performance**, driven by its ability to automatically handle categorical data and maintain strong generalization across samples.
We visualize these results in the following bar chart:

3 Conclusion

CatBoost achieved the highest ROC-AUC and was chosen as our primary attack model. The performance improvement, though incremental, highlights CatBoost's strength in handling complex feature interactions effectively.

Files in Repository

- `TML_Assignment_1.ipynb` - Main notebook
- `best_attack_model_catboost.pkl` - Saved CatBoost model
- `test.csv` - Provided test data
- `roc_auc_comparison.png` - Performance chart
- `README.md` - Summary and guide

GitHub repository link: **Link**

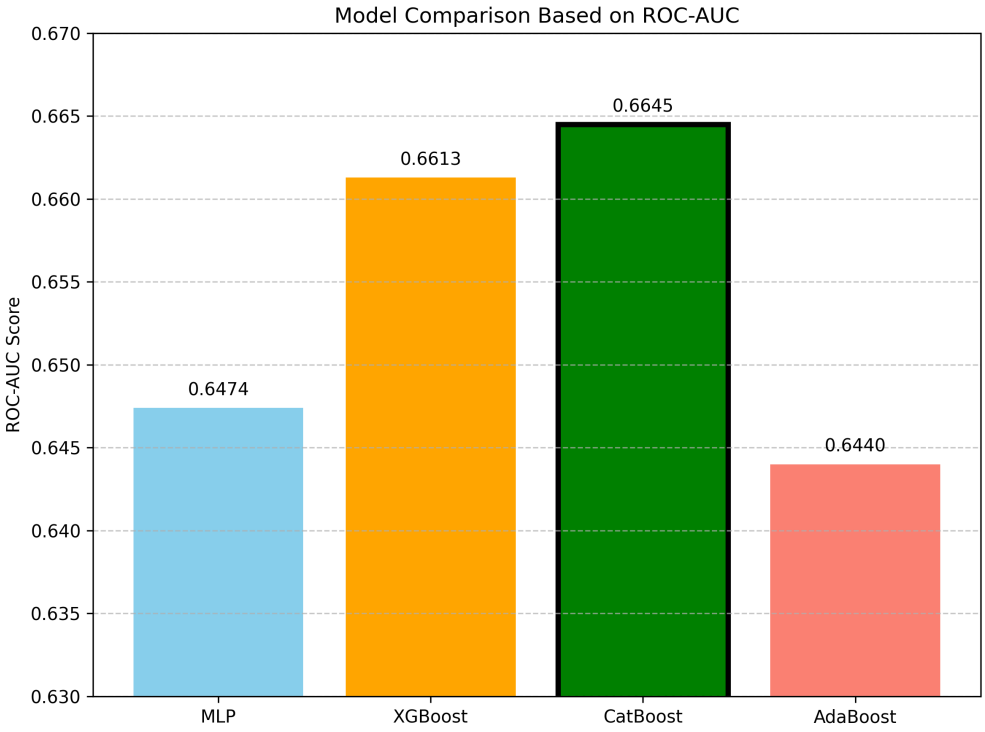

Figure 1: ROC-AUC comparison of evaluated models