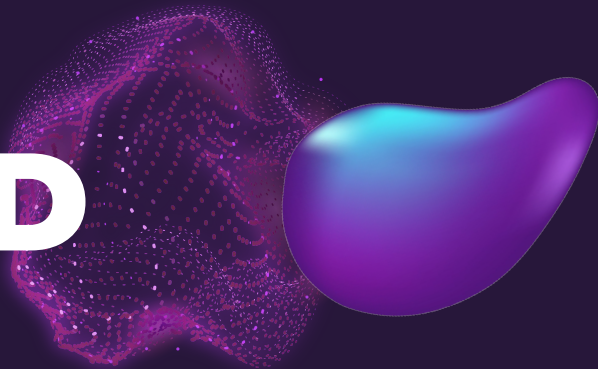


CREDIT CxARD FRAUD DETECTION

Machine Learning + AI

Emilio Pablo Neme



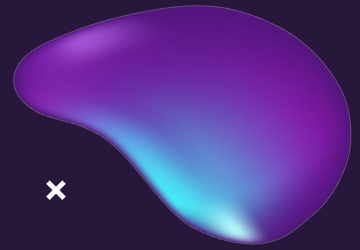


IDEA

This project focuses on detecting fraudulent credit card transactions using various machine learning techniques. The goal is to build and compare different classification models to accurately identify fraudulent transactions while minimizing false positives.



MODELS USED



01

KNN

02

Logistic Regression

03

Random Forest (ensemble)

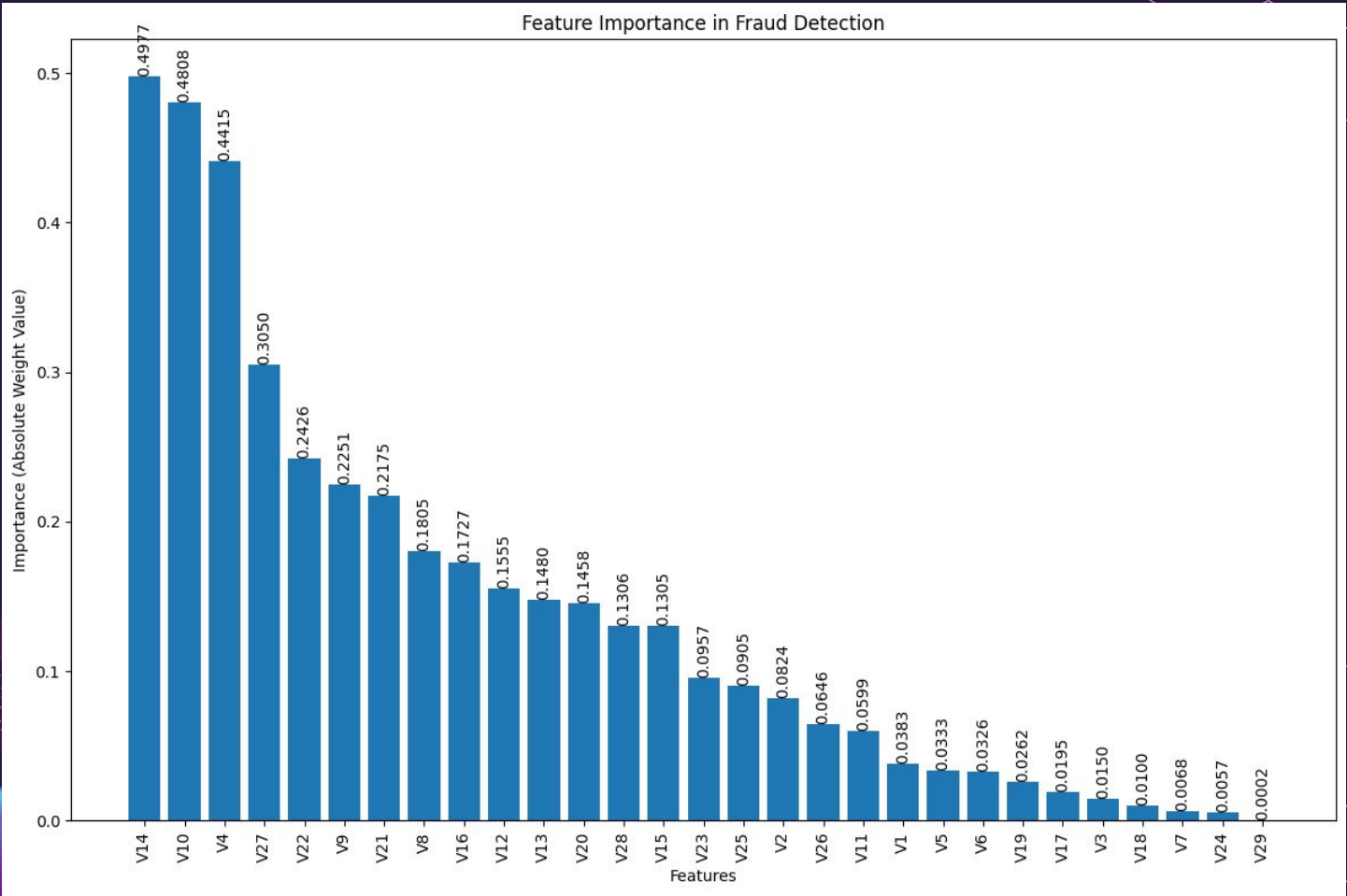
04

MLP



KAGGLE DATASET

- 28 features
- 284,807 transactions
- Numerical variables, result of a PCA transformation
- Fraud or Not fraud



COMPARISON

Transparent Models

KNN

Decisions based on proximity to neighbors

LogReg

Linear decision boundary based on feature weights

Black Box Models

MLP

Multiple layers of neurons and complex non-linear transformations

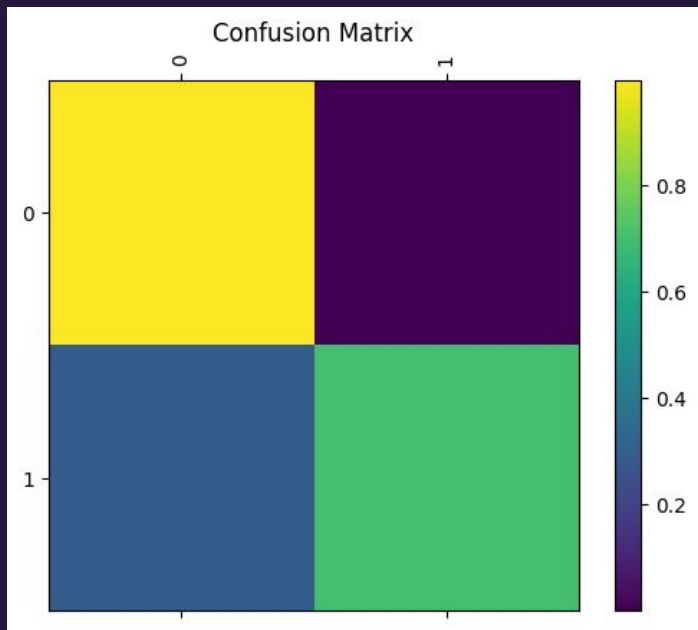
RandomForest

Combination of many decision trees and hard to trace exact decision path



KNN

x



71072 **TN**

6 **FP**

36 **FN**

87 **TP**

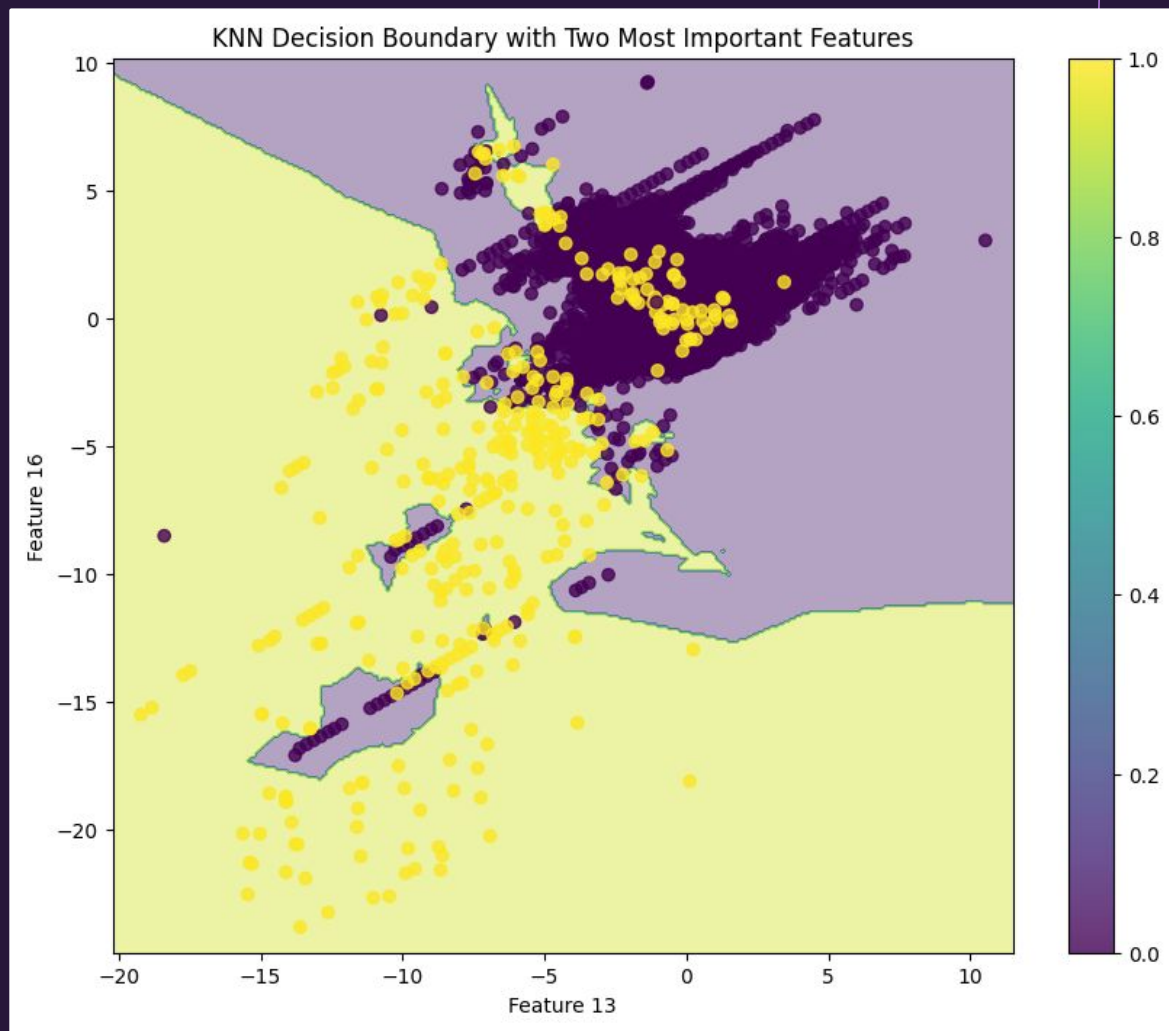
Accuracy: 0.99

Precision: 0.93

Recall: 0.71

Specificity: 0.99

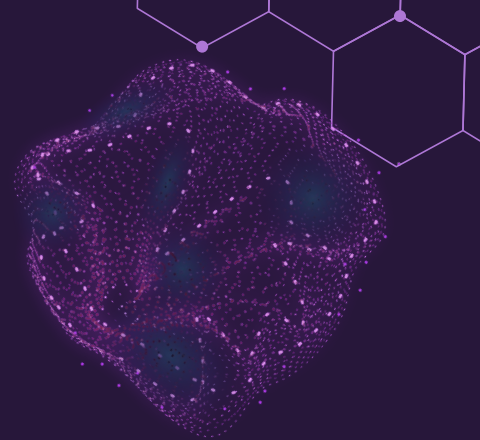
F1 Score: 0.81



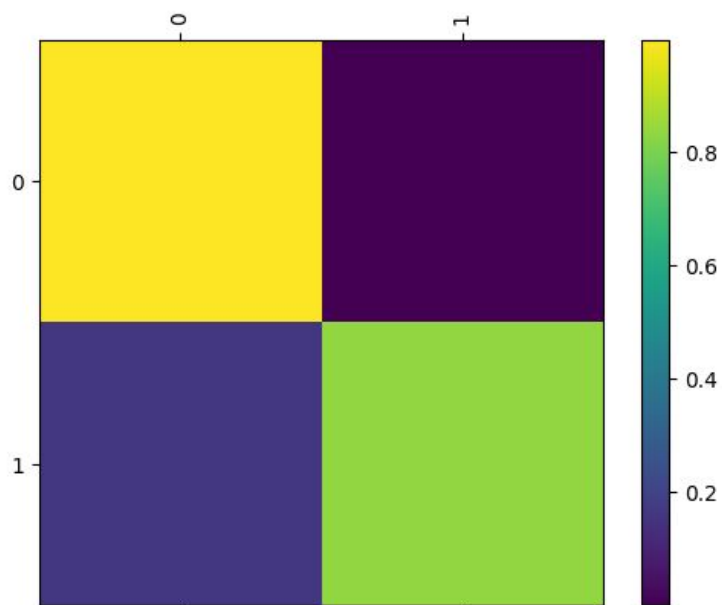


LogReg

x



Confusion Matrix



56835 **TN**

28 **FP**

16 **FN**

82 **TP**

Accuracy: 0.99

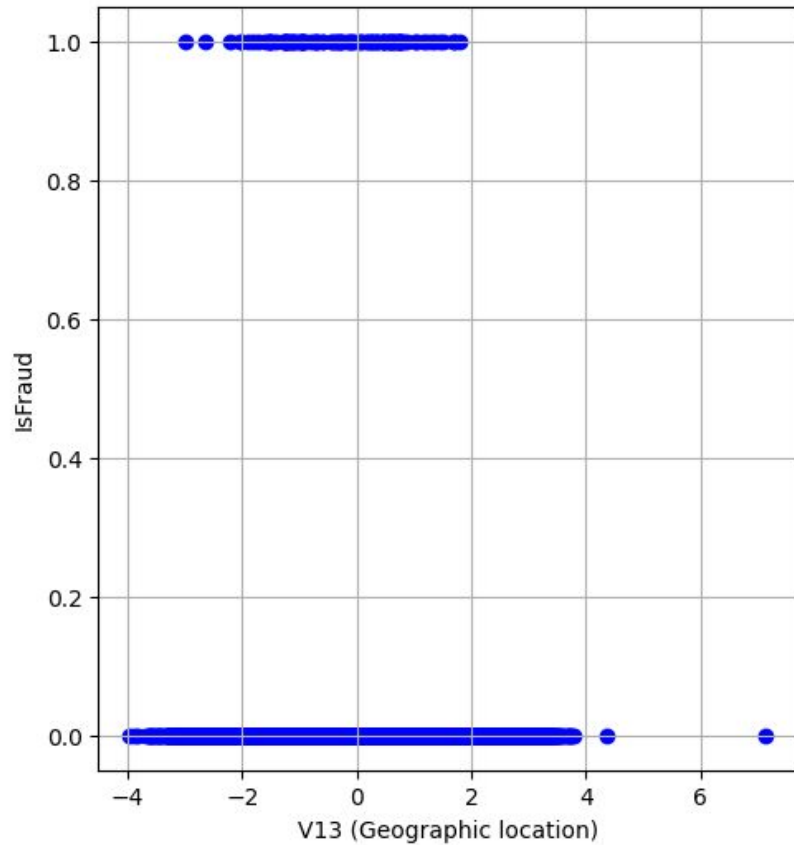
Precision: 0.75

Recall: 0.84

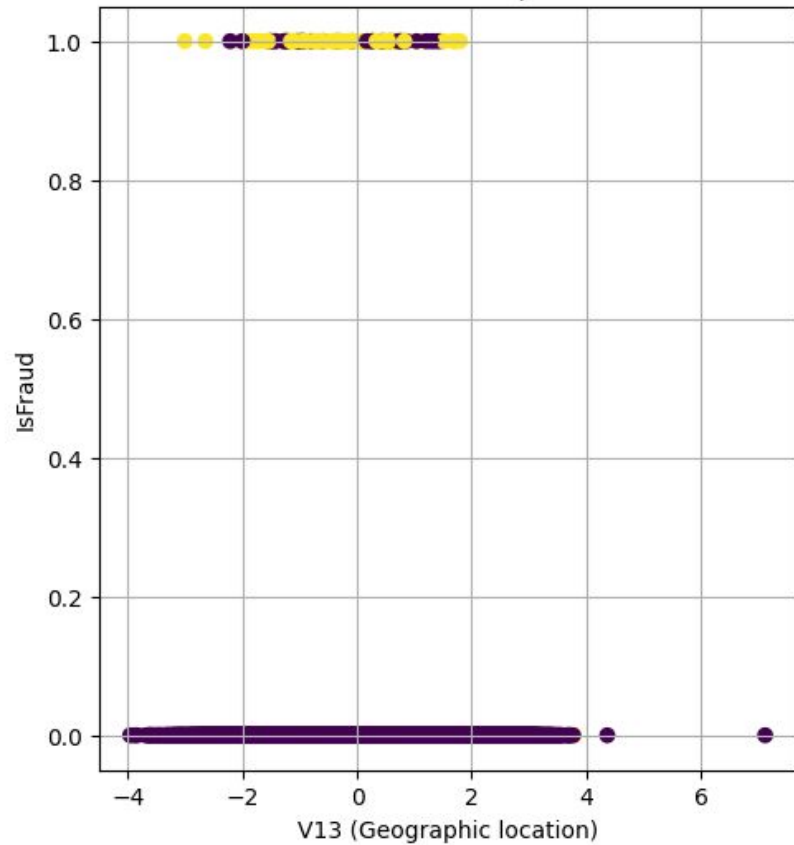
Specificity: 0.99

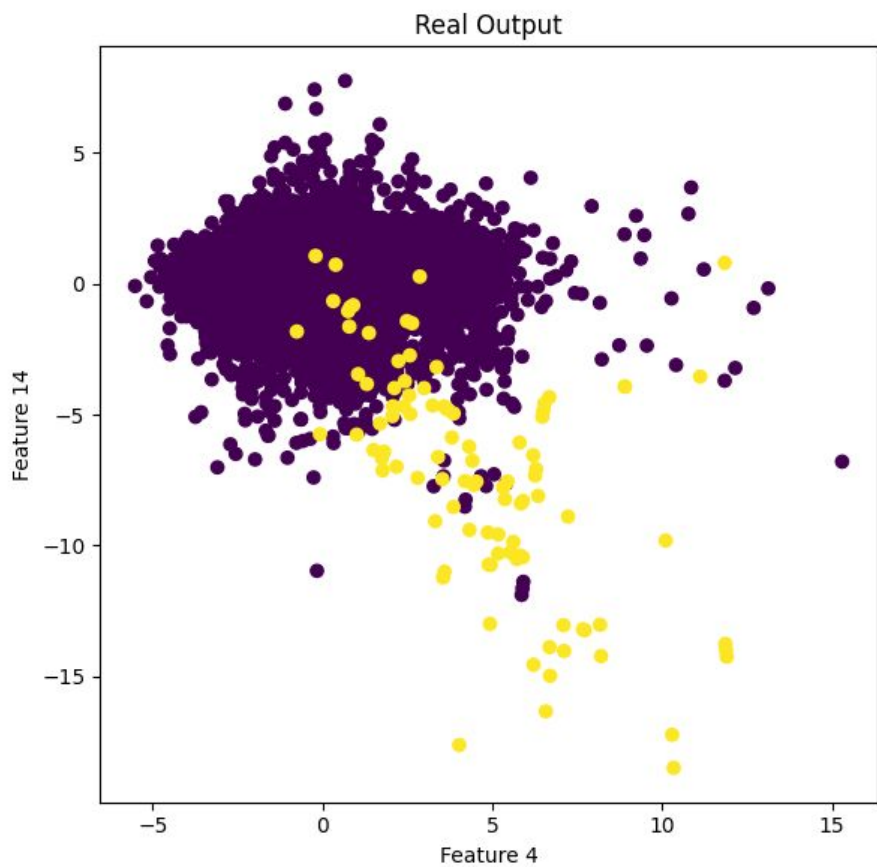
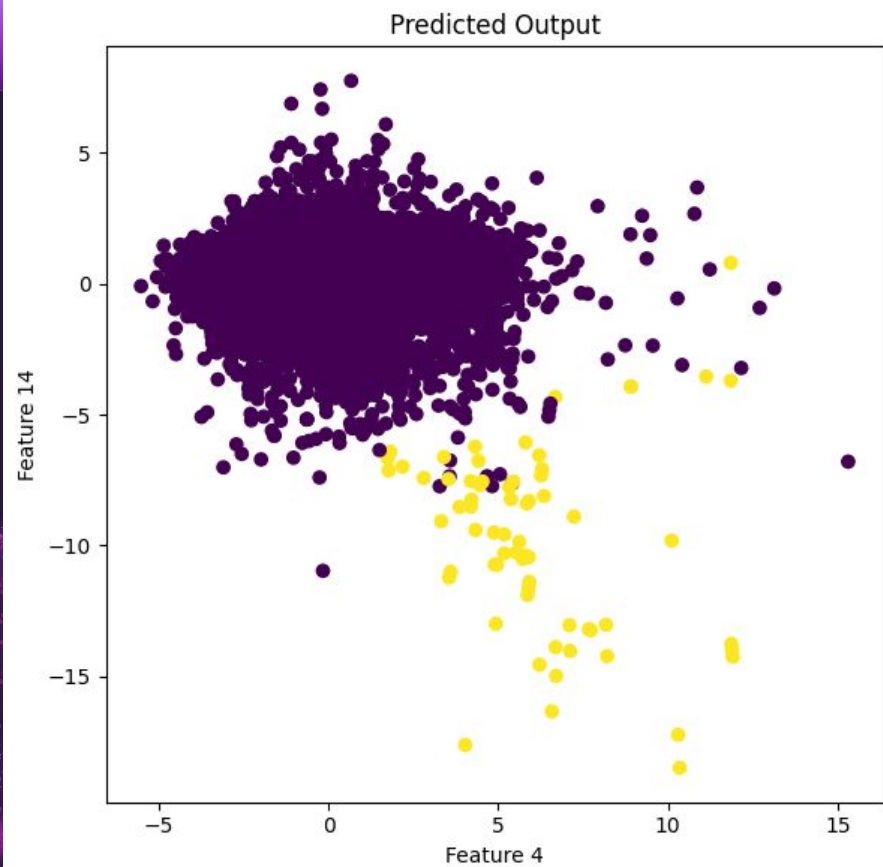
F1 Score: 0.79

Feature 13 vs Class



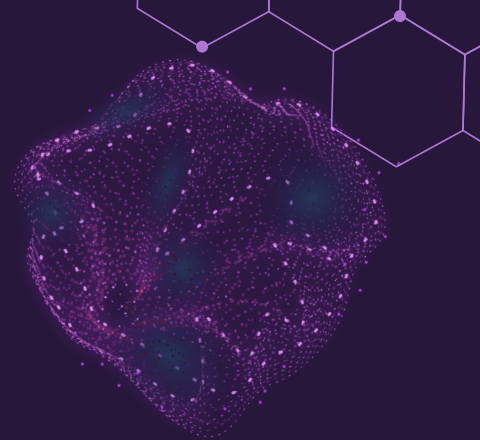
Predicted Output



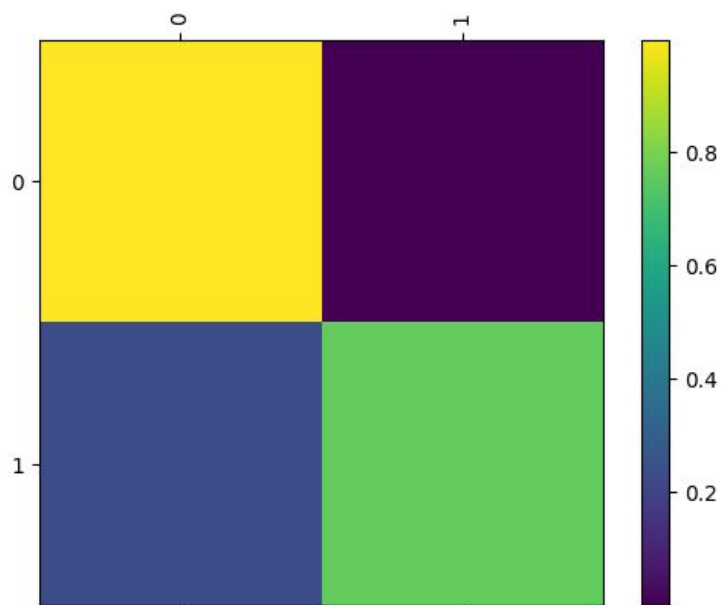




Random Forest_x



Confusion Matrix



56862 **TN**

2 **FP**

23 **FN**

75 **TP**

Accuracy: 0.99

Precision: 0.97

Recall: 0.77

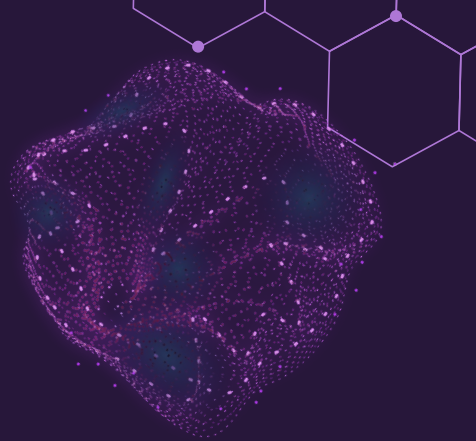
Specificity: 0.99

F1 Score: 0.86

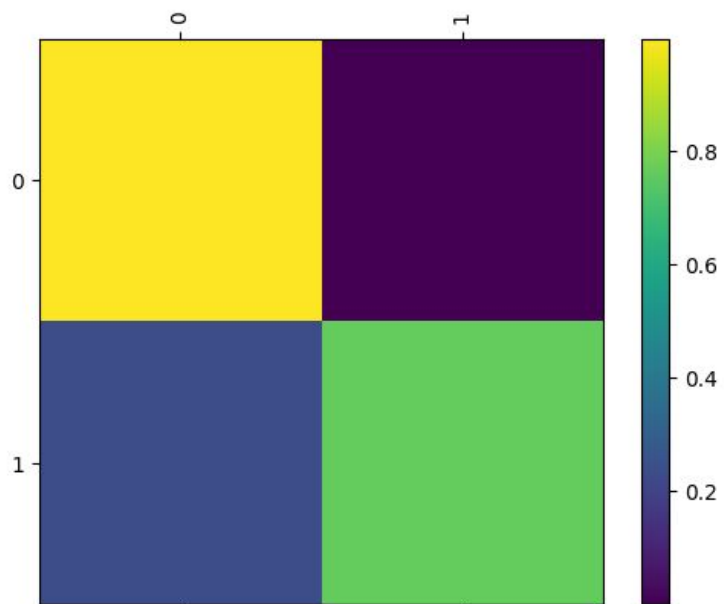


MLP

x



Confusion Matrix



56862 **TN**

2 **FP**

23 **FN**

75 **TP**

Accuracy: 0.99

Precision: 0.97

Recall: 0.77

Specificity: 0.99

F1 Score: 0.86

×

CONCLUSION

✓ Simple Models (Easier to Interpret)
KNN: Decisions based on closest neighbors

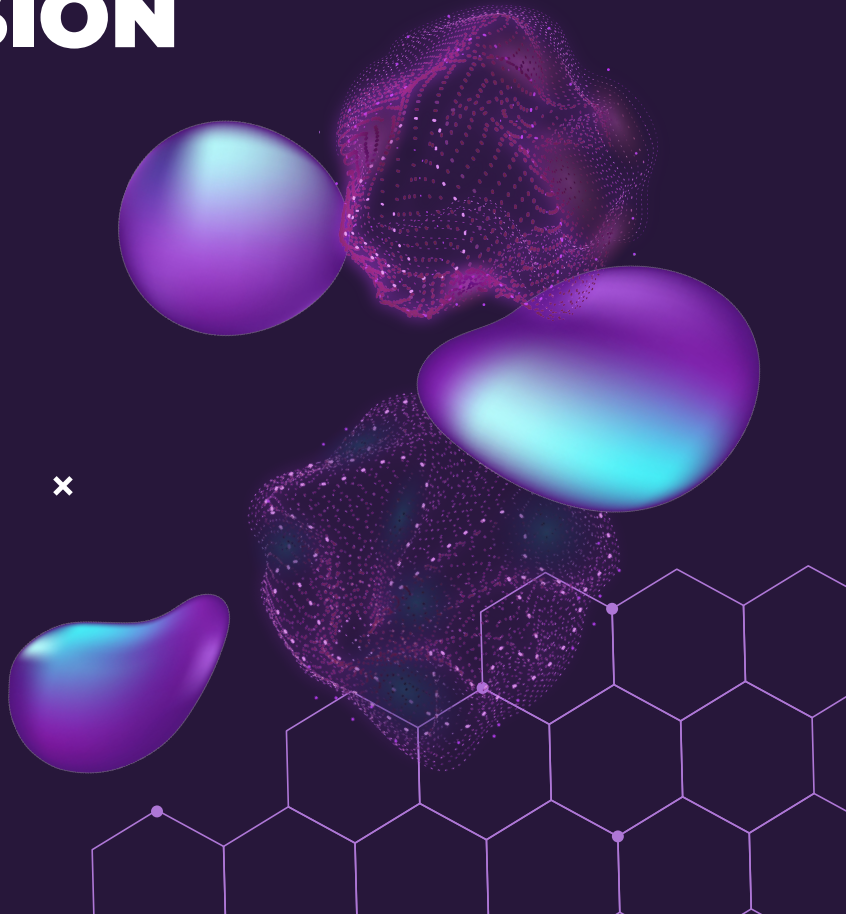
Logistic Regression: Linear decision boundary,
easy to explain

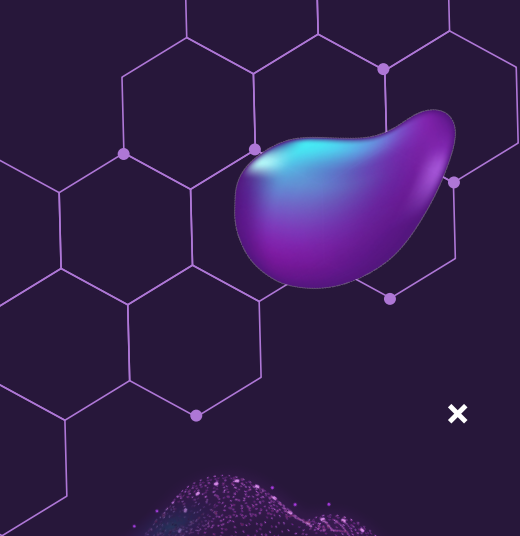
✗ Complex Models (Black-box)
Random Forest: Multiple decision trees, hard to
trace individual decisions

MLP: Multiple layers of neurons, complex
non-linear decisions

⚖ Trade-Off
Simple models = more explainable

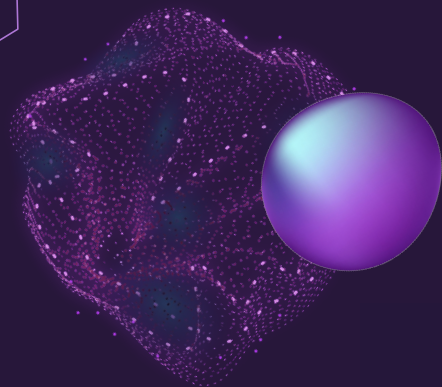
Complex models = higher accuracy, less
interpretability





x

THANKS!



+

x

