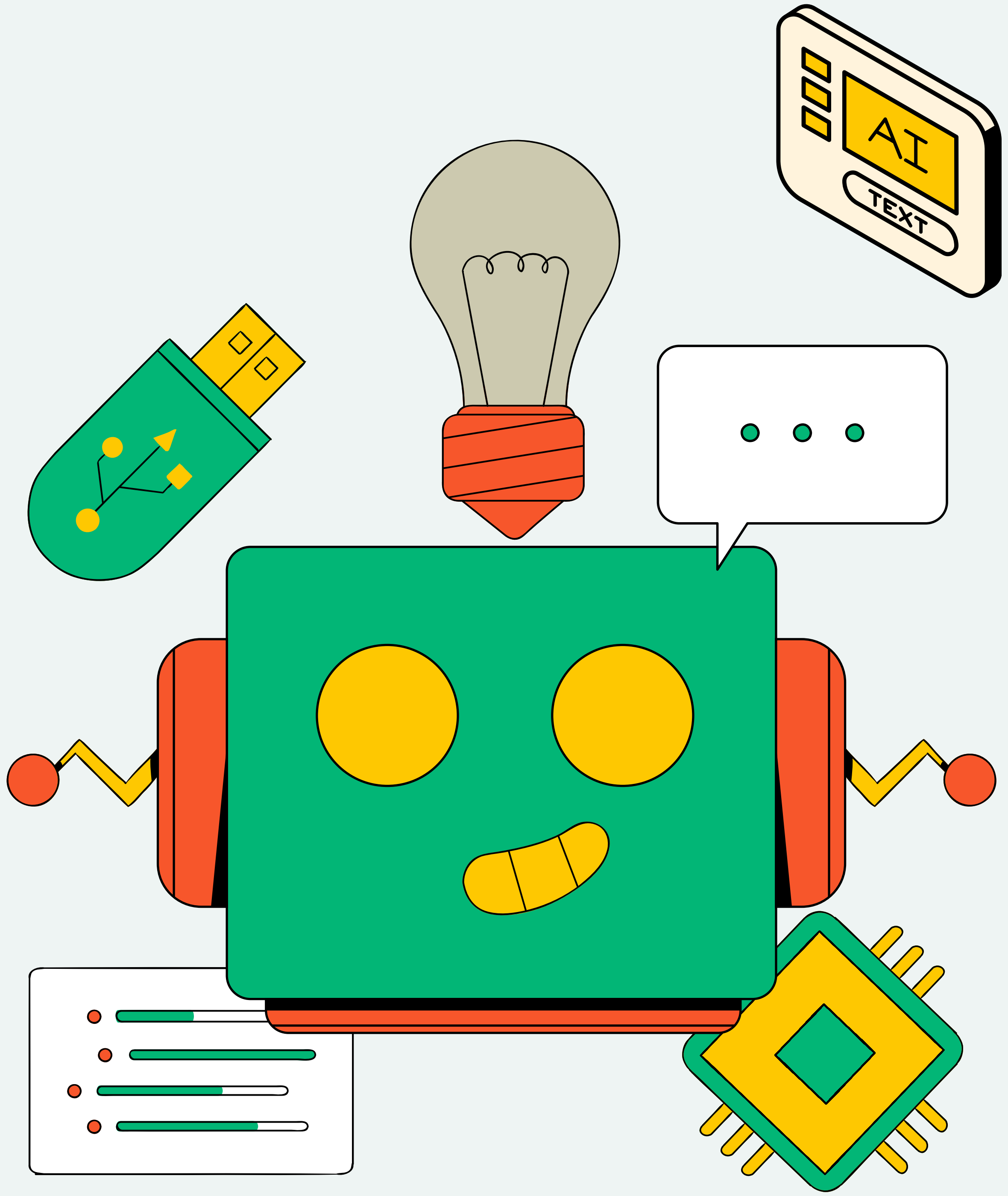
FRAUD NET - AI

DETECT. PREVENT. PROTECT

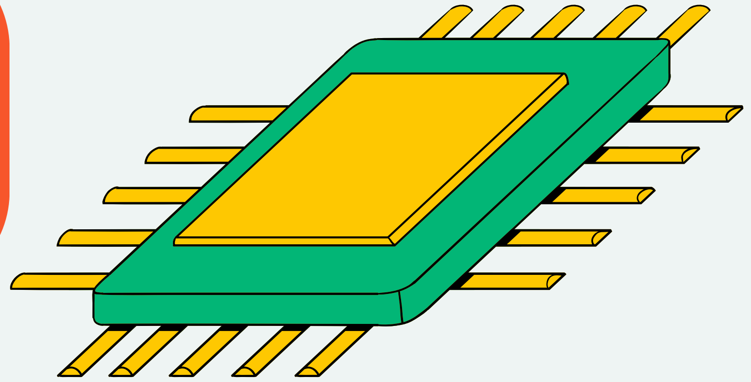


FRAUD NET AI

DETECT. PREVENT. PROTECT

PRESENTED BY:

D RAVI KIRAN



PRESENTATION OUTLINE

- Introduction
- Project Modules
- Datasets Used
- Key Concepts
- ML Workflow (Both Modules)
- Phishing Detection – ML Model
- Credit Card Detection – ML Model
- Accuracy Comparison
- Preparing for the Future
- Conclusion

INTRODUCTION

Fraud deception in networks refers to malicious activities like phishing and financial fraud that exploit systems and users.

With increasing online services, detecting these threats early using machine learning (ML) is vital.
This project aims to build two ML models:

- One for detecting Phishing URLs
- One for identifying Credit Card Fraudulent Transactions

Both models simulate a basic Network Intrusion Detection System (NIDS) approach.

WHAT IS PHISHING & CREDIT CARD FRAUD?

Phishing:

- A cyber-attack method where fake websites mimic legitimate ones to trick users into providing sensitive information (passwords, banking details, etc.).

Credit Card Fraud:

- Unauthorized use of a credit card to purchase goods or withdraw funds.

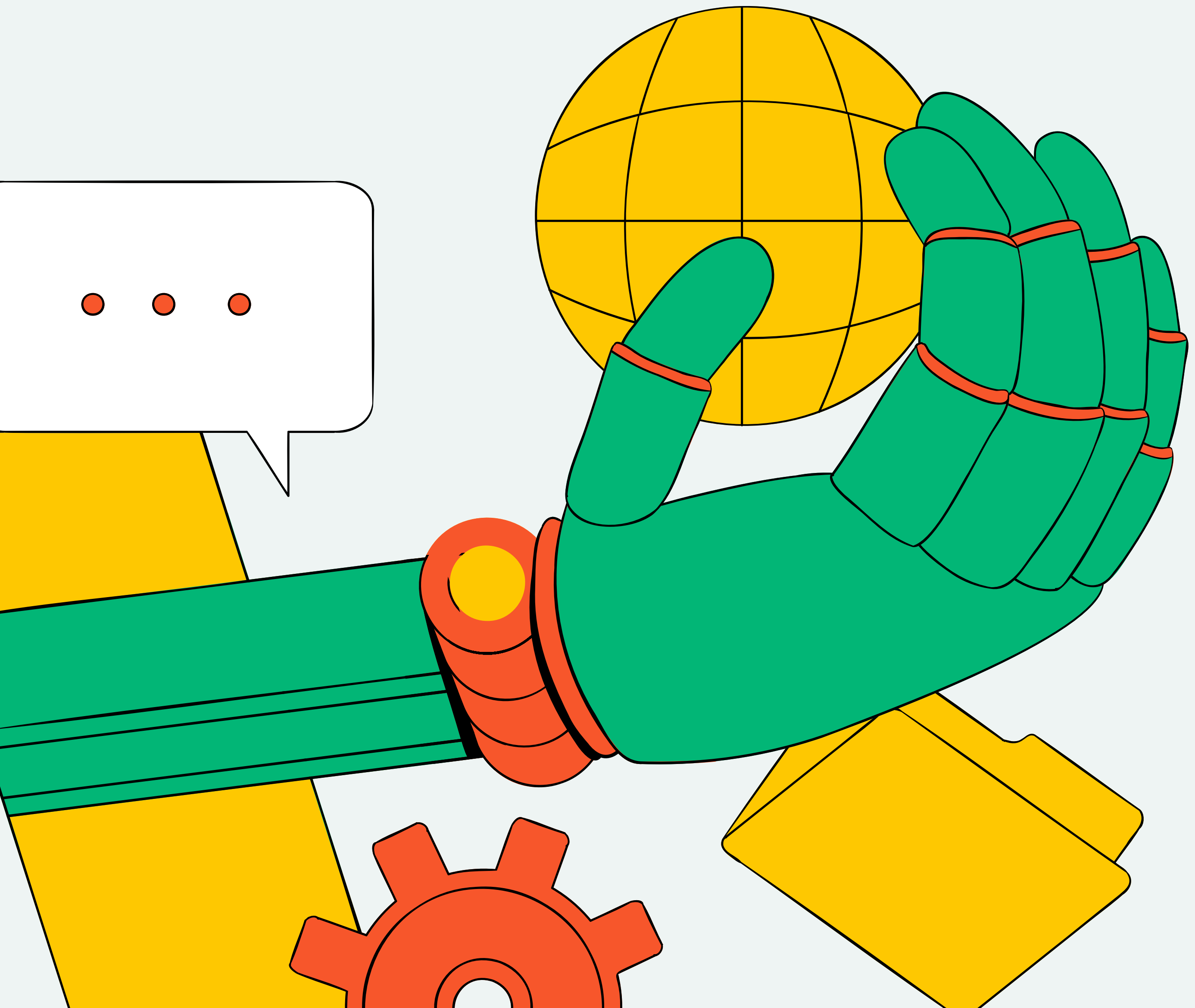
Both are major threats under the umbrella of network fraud.

PROJECT MODULES

Phishing URL Detection: Uses machine learning to analyze URLs and determine if they are safe or phishing attempts.

Credit Card Fraud Detection: Uses transaction data to detect unusual or fraudulent patterns.

DATASETS USED



Phishing Websites Dataset (UCI Repository):

- It contains features extracted from URLs (like length, domain presence, use of '@', etc.).
- Binary labels: 1 = Phishing, 0 = Legitimate

Credit Card Fraud Dataset (Kaggle):

- Real anonymized data from European cardholders.
- 30 features: Time, Amount, V1 to V28 (PCA-transformed)
- Highly imbalanced: ~0.17% fraud transactions

KEY CONCEPTS

Feature Engineering: For phishing: Extract features from the structure of a URL.

For credit card: Use scaled transaction values and PCA components.

Class Imbalance Handling: Fraud data is rare → use techniques like class weights or resampling.

Classification Algorithm:
Random Forest Classifier Effective for tabular data and classification tasks.

ML WORKFLOW

- Data Preprocessing
- Feature Extraction
- Train – Test Split
- Model Training
- Model Evaluation
- Model Evaluation
- Prediction

ML MODEL

Phishing Detection – ML Model:

Input: URL (e.g., from user or website logs)

Feature Set: URL length, @ symbol, HTTPS presence, use of IP address, suspicious keywords

Model Used: Random Forest Classifier

Output: Classifies as Phishing or Legitimate

```
test_url = "http://samrhamburg.com/78gz11on"
print(predict_url(test_url, rfc))
```

```
Phishing
```

ML MODEL

Credit Card Detection – ML Model:

Input: Transaction features (V1–V5, Time, Amount)

Preprocessing:
  Normalize Time and Amount using StandardScaler

Model Used:
  Random Forest Classifier

Output:
  Classifies as Fraudulent or Genuine

```
sample_transaction = {
    'Time': 40660,
    'V1': -2.3122265423263,
    'V2': 1.95199201064158,
    'V3': -1.60985073222,
    'V4': 3.9979055875468,
    'V5': -0.522187864667764,
    'Amount': 0.00
}

print(predict_transaction(sample_transaction, rfc, sc))
```

```
Fraudulent
```

PREPARING FOR THE FUTURE

01

REAL-TIME DEPLOYMENT

Can be integrated into a web dashboard or fraud detection API

02

SCALABILITY

Extendable to mobile apps, banks, browsers

03

IMPROVEMENT AREAS

Try advanced models like XGBoost, Deep Learning (LSTM for time-sequences) and Use active learning with updated fraud data

CONCLUSION

PROJECT ACHIEVEMENTS

- Successfully developed a machine learning system for fraud detection.
- Implemented two modules..
- Achieved high accuracy in both models using real-world datasets.

SYSTEM CAPABILITIES

- Capable of performing real-time predictions.
- Combines both modules into a unified fraud deception detection system.
- Handles imbalanced data effectively with appropriate ML techniques.

FUTURE SCOPE

- Can be extended to integrate with web applications or APIs.
- Potential to enhance with deep learning models for improved accuracy.
- Scalable for broader network intrusion detection systems.

THANK YOU