i-Hack: India's Premier Hackathon
E-Summit'25

Problem Statement : Combating Spam Calls, Deepfake Fraud, and VKYC Exploitation in Financial Services.

Track : <i-Hack> Financial Security Track

Organizers: E-cell, IIT-Bombay

Team Name : Innovators
Team Members : 1) Ajinkya Wagh
2) Bhavesh Patil
3) Shivani Pawar

<i_hack>

CODE | CREATE | DISRUPT

Team Innovators

IDEA & SOLUTION APPROACH

Idea/Prototype

The financial industry faces growing threats from technologically advanced fraud schemes, including spam calls, deepfake-based identity fraud, and unauthorized VKYC manipulations. This project introduces a unified AI-driven platform to comprehensively detect and prevent fraud using cutting-edge AWS services and machine learning models.

Prototype :
The architecture integrates multiple AWS components for seamless fraud management:
•Data Storage:Amazon S3 for storing transcripts, video records, and transaction data.
•Serverless Execution:AWS Lambda for triggering real-time fraud checks.
•NLP and Language Analysis:AWS Transcribe and Comprehend.
•Face and Deepfake Detection:Amazon Recognition.
•Machine Learning Deployment:Sage Maker for training fraud-detection models.
•Alerts and Feedback:AWS SNS and Amplify to notify users and collect reports for iterative learning.

Solution Approach

1.Data Storage and Security:
  1.All data, including call logs, VKYC session videos, and financial transactions, is securely stored in Amazon S3 .
2.Preprocessing Pipeline:
  1.AWS Glue transforms raw data into structured formats for optimal analysis by downstream processes.
3.Spam Detection:
  1.Convert voice data to text with AWS Transcribe
  2.Apply AWS Comprehend for phishing keyword identification.
  3.Analyse metadata using AWS Lambda to detect suspicious patterns.
4.Transaction Monitoring:
  1.Machine learning models trained in Sage Makeranalyse patterns for transaction anomalies.

This expansion integrates detailed steps and technologies, ready for refinement into a presentation or report.

Team Innovators

INNOVATION & CONSTRUCTION

Innovation & Uniqueness

Holistic Fraud Management:Unlike single-purpose solutions, this system integrates call monitoring, deepfake detection, and transaction analysis

Proactive Fraud Alerts:Real-time alerts offer users immediate actionability.

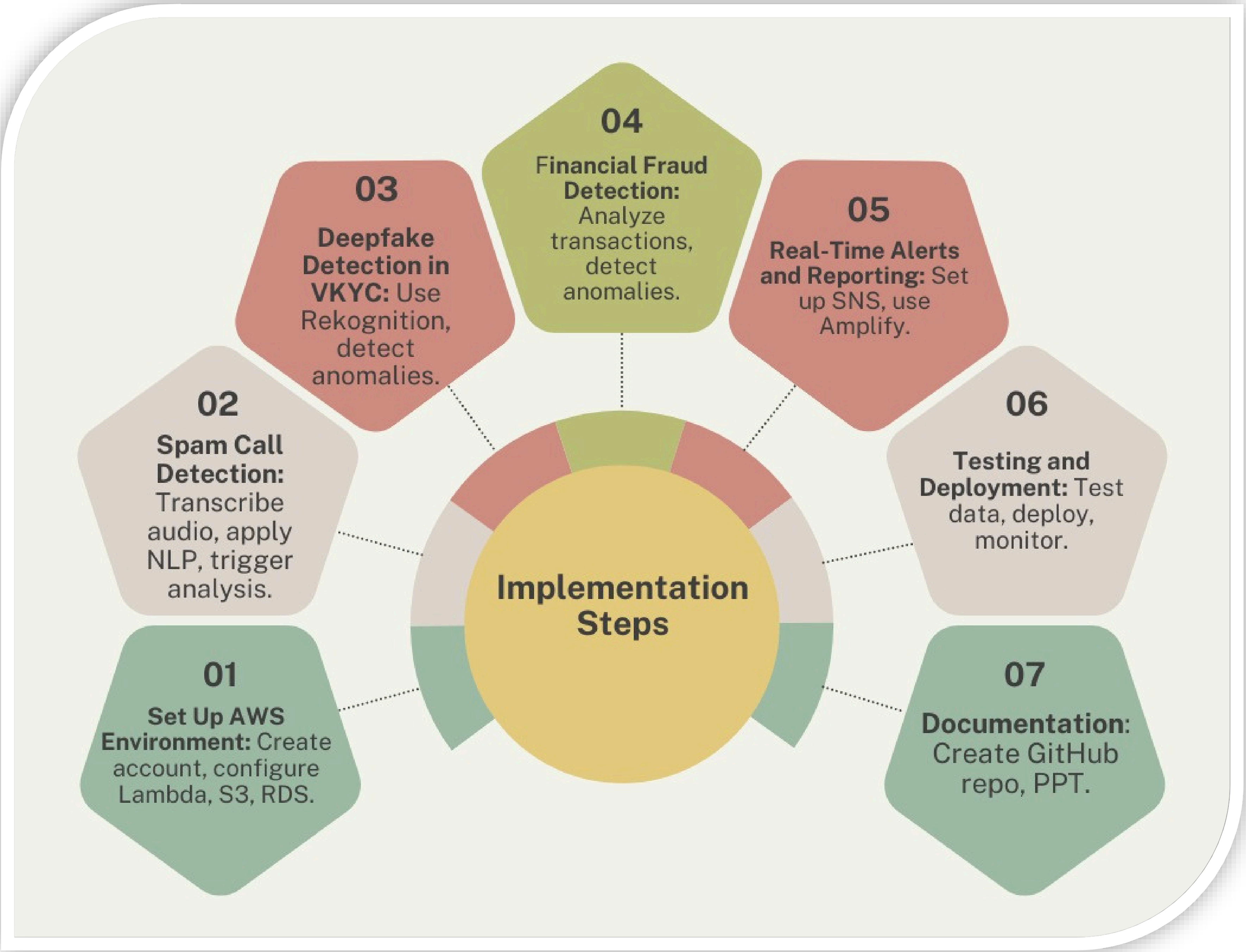
Cloud-Native Scalability:Built on AWS, ensuring resilience, reliability, and the ability to handle high transaction volumes.

Dynamic Model Updates:Integrates continuous learning from user feedback to adapt to evolving fraud techniques.

Behavioural Pattern Analysis:Goes beyond static data matching to analyse transaction behaviours and user interaction anomalies.

Comprehensive Security:Combines AI and cloud-native solutions to offer end-to-end security with minimal latency.

Flow of Construction Activities & ML Model Used :-



Implementation Steps

01 **Set Up AWS Environment:** Create account, configure Lambda, S3, RDS.

02 **Spam Call Detection:** Transcribe audio, apply NLP, trigger analysis.

03 **Deepfake Detection in VKYC:** Use Rekognition, detect anomalies.

04 **Financial Fraud Detection:** Analyze transactions, detect anomalies.

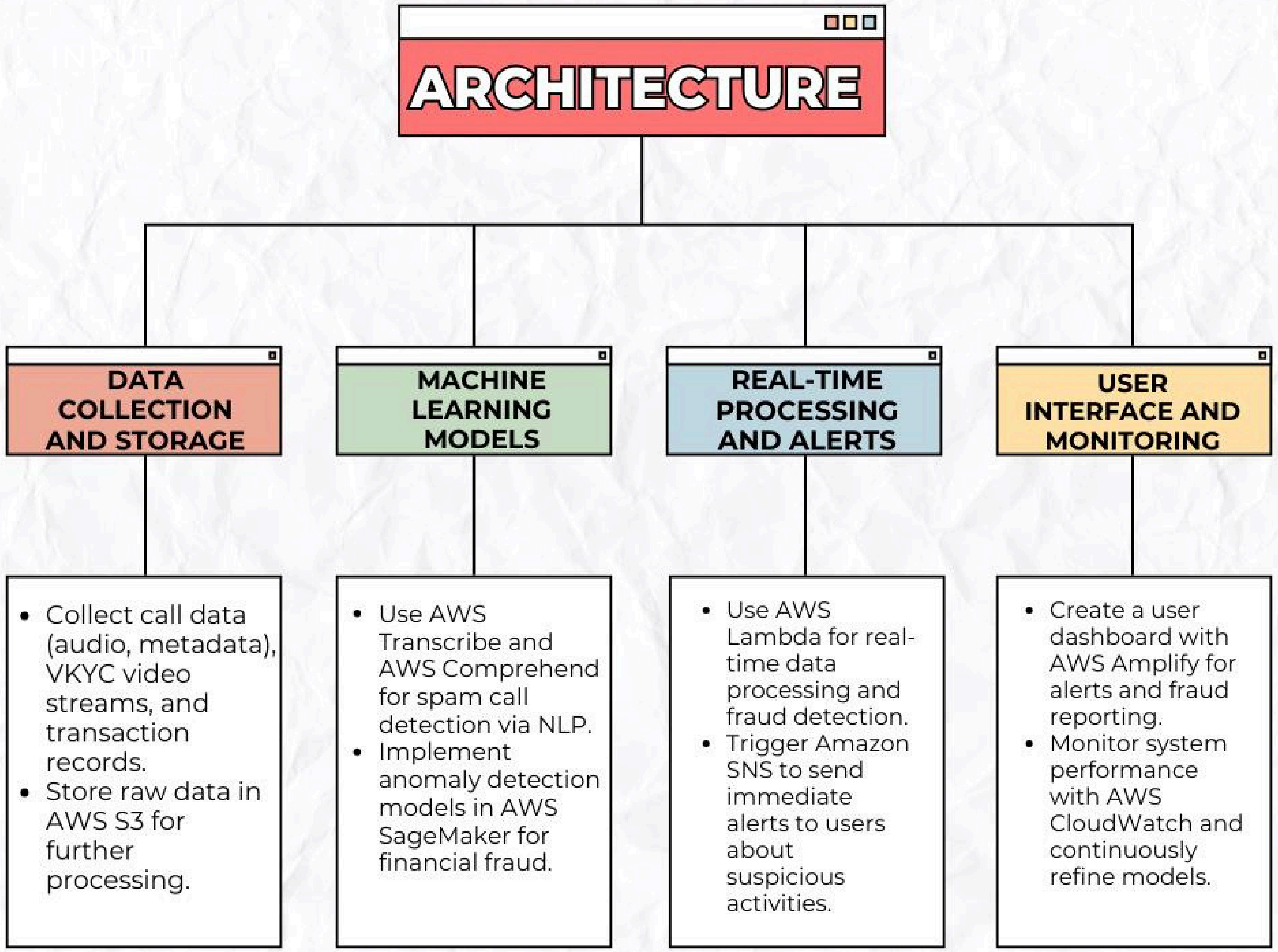
05 **Real-Time Alerts and Reporting:** Set up SNS, use Amplify.

06 **Testing and Deployment:** Test data, deploy, monitor.

07 **Documentation:** Create GitHub repo, PPT.

Team Innovators

TECHNICAL APPROACH



ARCHITECTURE

DATA COLLECTION AND STORAGE

- Collect call data (audio, metadata), VKYC video streams, and transaction records.
- Store raw data in AWS S3 for further processing.

MACHINE LEARNING MODELS

- Use AWS Transcribe and AWS Comprehend for spam call detection via NLP.
- Implement anomaly detection models in AWS SageMaker for financial fraud.

REAL-TIME PROCESSING AND ALERTS

- Use AWS Lambda for real-time data processing and fraud detection.
- Trigger Amazon SNS to send immediate alerts to users about suspicious activities.

USER INTERFACE AND MONITORING

- Create a user dashboard with AWS Amplify for alerts and fraud reporting.
- Monitor system performance with AWS CloudWatch and continuously refine models.

TECH STACK

python™ · Amazon SageMaker · matplotlib · aws · {JSON} · scikit learn · Boto 3 · Flask · pytest · pandas

METHODOLOGIES:

NLP for Spam Detection : AWS Transcribe converts audio to text, and AWS Comprehend analyzes text for phishing keywords.. Deepfake Detection: Amazon Recognition verifies facial features, and custom models on Sage Maker detect anomalies in voice and facial patterns. Transaction Analysis: Data pipelines with AWS Glue prepare financial data, and Sage Maker anomaly detection algorithms identify suspicious transactions. Real-Time Alerts: AWS SNS provides notifications, while AWS Amplify offers a user-friendly interface for reporting fraud. Risk-Based Authentication: This methodology dynamically adjusts the level of authentication required based on the perceived risk of a transaction or user action.

Team Innovators

FEASIBILITY AND VIABILITY

Feasibility Analysis

TechnologicalFeasibility:Utilizes proven AWS services, ensuring robust AI/ML capabilities without extensive infrastructure management.
Operational Feasibility: Serverless architecture reduces maintenance; user-friendly interfaces allow easy fraud reporting.
Economic Feasibility: Pay-per-use AWS pricing makes the system cost-effective for various scalesofoperation.
Legal and Regulatory Feasibility: The system complies with industry standards and regulations related to data privacy such as GDPR, and financial data protection laws.

Potential Challenges & Risks

False Positives: Legitimate actions flagged as fraud leading to user frustration and potential loss of trust in the system.
Privacy Concerns: Handling sensitive user data securely such as personal identification, transaction records, and VKYC information, presents significant privacy challenges
Deepfake Sophistication: Rapid evolution of deepfake technology it becomes increasingly challenging to detect manipulated content.
Latency Issues: Delays in real-time fraud detection affecting user experience.

Overcoming Challenges

Enhanced ML Training: Use large, diverse datasets to improve detection accuracy.

Data Security: Implement strong encryption or AWS Identity and Access Management (IAM) controls.

Adaptive Models: Continuously update ML models to adapt to evolving fraud patterns.

Optimized Architecture: Use edge computing and caching strategies to reduce latency..

Revenue Sources:
Subscription Plans: Tiered pricing for individuals and businesses based on usage.
API Access Fees: Licensing fraud detection APIs to third-party applications.
Enterprise Solutions: Customized deployments for financial institutions.
Partnerships: Revenue-sharing models with telecom providersandbanks

Check Prototype:
bit.ly/40Hr2Si

Team Innovators

IMPACT AND BENEFITS

POTENTIAL IMPACTS

Target Audience:

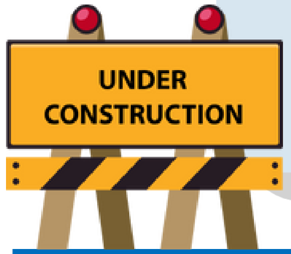
Fraud Reduction: Detects and prevents fraudulent activities like spam calls, deepfake fraud, and unauthorized transactions, significantly minimizing financial scams.

User Trust: Increases consumer confidence in digital financial services by providing proactive protection and real-time alerts against security threats.

Operational Efficiency: Automates fraud detection processes, reducing the need for manual monitoring and saving time and resources for financial institutions.

Social Safety and Security: Enhances financial literacy and safety for vulnerable populations, protecting them from scams and malicious actors.

Economic Stability: Reduces losses due to fraud, contributing to a more secure and stable financialecosystem.



BENEFITS

SOCIAL

Safer financial ecosystems for vulnerable populations.
Reduced exposure to scams and phishing attacks.

ECONOMIC

Minimized financial losses due to fraud.
Increased consumer confidence in digital banking.

ENVIRONMENTAL

Reduction in manual verification processes.
Lower carbon footprint from reduced physical paperwork and fewer in-person verifications.

Team Innovators

RESEARCH

Research Papers and Articles

1.Fraud Detection in Financial Transactions using Machine Learning Algorithms
Authors: S. Raj and V. Portia
*Published in*International Journal of Computer Applications
Link: [ResearchGate]()
Summary: This paper explores different machine learning techniques for fraud detection and evaluates their performance in identifying financial fraud. It highlights the use of decision trees, SVM, and neural networks.

2.Spam Call Detection Using Natural Language Processing
Authors: P. Gupta, M. Singh
Published in: International Journal of Engineering Science and Computing
Summary : This research presents NLP-based approaches to identify and filter spam calls based on textual data derived from audio.

3.Real-Time Fraud Detection Using AWS and Machine Learning
AWS Whitepaper
Link: Available from [AWS Documentation]()
Summary: Explains AWS services for implementing fraud detection solutions using Lambda, SageMaker, and other cloud services.

ExistingSystems

1.Truecaller (Spam Call Detection): Overview: Uses crowd- sourced data and machine learning for spam call detection. Limitations: Relies on user reports, making it less effective for newfraudschemes.

2.Call Control (Spam Call Detection): Overview: Blocks robocalls and spam using an AI-driven algorithm. Limitations: Dependent on community-driven spam lists, missing emerging fraud types.

3.Veriff (VKYC and Fraud Prevention): Overview: Offers identity verification with AI and machine learning for deepfake detection. Limitations: Vulnerable to advanced deepfakes and requiresimprovementsinvoice/behavioralanalysis.

4.PayPal (Financial Fraud Detection): Overview: Monitors transactions using machine learning to flag suspicious activities. Limitations: Relies on historical fraud patterns, which may not detectnewtactics.

5.AWS Fraud Detection (AWS Services for Financial Fraud): Overview: Uses services like Sage Maker, Lambda, and RDS for fraud detection. Limitations: Requires fine-tuning models to handleevolvingfraudtactics.