A3 No. and Name	Team members (name & role)	Stakeholders (role & department)	Company objective
Group 9	1. Lella Gopal, Samyukth Lalith 2. 3. 4.	1. AI & ML Coordinator, Conestoga College 2. Potential Client name(s) 3. Conestoga College stakeholder 4.	AI-Driven Synthetic Identity Detection

Team Leader (name & 'phone ext)	Start date & planned duration
	Q1 2024 – Q4 2024

Synthetic Identity Fraud Detection Using Ethical AI

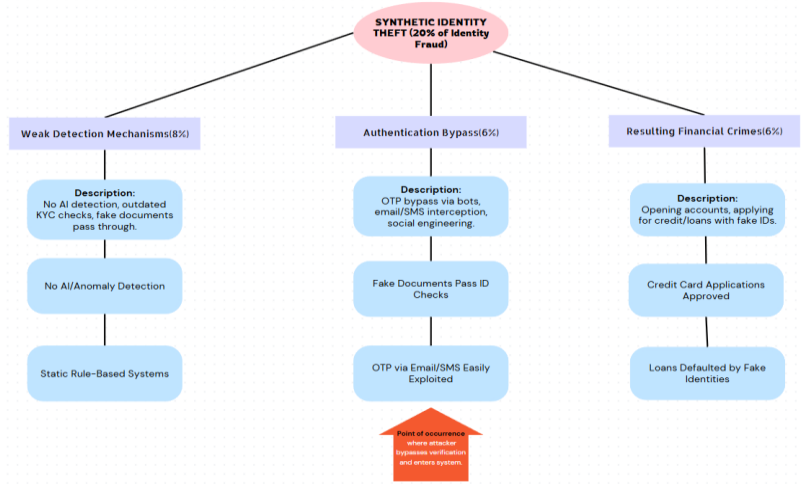
1. Clarify the problem

Ideal Situation: Reduce Synthetic Identity Theft to **below 10%** of all identity fraud-related financial cases by 2024.

Current Situation: As of 2023, Synthetic Identity Theft contributes to approximately 20% of identity fraud-related financial losses.

Gap: The current contribution of Synthetic Identity Theft is 20%, and the target is to reduce it to below 10%, resulting in a gap of 10 percentage points. This reduction will be addressed using AI-driven anomaly detection and explainable AI (XAI) solutions that support ethical, scalable, and transparent fraud detection systems.

2. Breakdown the problem



3. Set the Target

1	**Reduce the contribution of Synthetic Identity Theft from 20% to below 10%** of all identity fraud cases by the end of 2024.
2	**Achieve at least 85% detection accuracy** using AI anomaly detection and XAI models, reducing false positives by 30%.
3	Cut OTP bypass incidents by 50% through implementation of advanced authentication layers (e.g., app-based MFA).
4	Train and educate 1,000+ employees and users on identifying synthetic ID patterns and fraud risks by Q4 2024.

4. Analyse the Root Cause

WHY?	**Why is synthetic ID theft occurring?** Mixing real and fake data bypasses traditional ID verification
WHY?	**Why are these attacks increasing?** Weak KYC/AML systems and static fraud detection.
WHY?	**Why are authentication mechanisms weak?** Email and SMS OTPs are bypassed easily, no AI detection
WHY?	**Why don't companies detect synthetic IDs?** Lack of AI and data-link analysis for real-time detection
WHY?	**Why is AI not widely adopted?** High cost, complexity, and lack of explainability
Root Cause:	The root cause of synthetic identity fraud is the lack of AI-based real-time detection at the point of OTP and document verification, which allows attackers to bypass identity checks and execute fraud using synthetic profiles.

5. Develop Countermeasures

	Criteria >	Prevention of Financial & Data Loss	Protection of Employees & Customers	Reduces IT Workload & Costs	Builds customer Trust & Compliance	Reduces security response time.	Overall Score (/100):	Ranking:	Potential Problems:
Options	1 Real-Time AI Anomaly Detection at Verification Points	30	29	11	15	10	95	1	Requires access to large and diverse datasets for training; initial development costs can be high.
	2 AI-Powered Behavioral Biometrics (Typing, Mouse)	27	24	20	8	9	88	2	None!
	3 Graph-Based Authentication Security	24	18	16	11	9	78	3	Requires integration of multi-source data, may face resistance from organizations due to complexity.
	4 Explainable AI (XAI) for Transparent Fraud Decisions	24	30	18	8	5	85	4	None!
	5 Cybersecurity Education on Synthetic ID Risks	18	30	10	7	5	70	5	Public outreach can be slow and costly; depends on user engagement and willingness to learn.

6. Implement Countermeasure

7. Monitor Results & Process

8. Standardize & Share Success

https://github.com/Samyukth107/Synthetic-Identity-Theft-Detection

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Couldn’t upload the dataset to the git due to the size

Synthetic Identity Theft Detection using Ethical AI and Explainable AI

Abstract:

Synthetic Identity Theft (SIT) is an emerging digital threat where real and fake data are combined to create fraudulent identities. Traditional verification methods like KYC and SMS-based OTPs often fail to detect such fraud. This project proposes an AI-driven detection framework enhanced with Explainable AI (XAI) and behavioral biometrics to improve accuracy and transparency.

The goal is to reduce synthetic fraud by 70% and improve detection rates by 60% through a three-phase rollout: model development (Q1–Q2 2024), sector-specific deployment (Q3–Q4 2024), and national awareness campaigns (2025). Performance will be evaluated using fraud reduction, AI accuracy, explainability, and outreach success.

Problem Statement:

Synthetic identity theft, where real and fake data are combined to create fraudulent identities, accounts for roughly 20% of identity fraud losses in North America. Traditional fraud detection systems often fail to identify these sophisticated attacks due to weak KYC checks and outdated rules. This project addresses the need for intelligent fraud detection by proposing an AI-powered system supported by behavioral biometrics and Explainable AI (XAI), aiming to reduce synthetic ID fraud cases to below 10% by 2024.

Ideal State:

The ideal state is to reduce the contribution of synthetic identity fraud to below 10% of total identity fraud-related financial losses in North America by the end of 2024, using AI-based anomaly detection, behavioral biometrics, and explainable AI techniques to enhance fraud prevention and transparency.

Research hypothesis:

1. Introduction

Synthetic Identity Theft (SIT) is a rapidly growing form of identity fraud where attackers combine real and fake information to create new fraudulent identities. Unlike traditional identity theft, synthetic IDs are harder to detect because they mix authentic data like Social Security Numbers with fabricated names and addresses. This type of fraud has caused $1.8 billion in losses in the U.S. alone and accounts for 20% of total credit fraud losses.

Synthetic identities are used to open accounts, secure loans, and obtain credit cards, often going undetected for years. Cases of synthetic fraud have increased by 25% between 2018 and 2022, fueled by weak verification processes such as basic KYC checks and SMS or email-based OTPs.

To address this, AI-driven anomaly detection, along with Behavioral Biometrics and Graph-Based Authentication, provides a robust approach to identifying synthetic identities. Adding Explainable AI (XAI) ensures that AI systems are transparent and fair, addressing growing ethical concerns in fraud detection. This research proposes a combined AI and XAI-based system to effectively detect and mitigate synthetic identity fraud while maintaining ethical standards.

2. Literature Review

Synthetic Identity Theft leverages weaknesses in current KYC and authentication mechanisms (Trnka et al., 2022). Attackers exploit partial real information, bypassing static verification systems (Wilson, 2020). AI-based anomaly

detection has shown promise in identifying abnormal patterns in user data (Experian, 2023). Behavioural biometrics such as typing patterns and device analysis enhance user profiling (Sharma & Patel, 2023). Graph-based authentication helps visualize connections among data points to detect suspicious synthetic links (Mane & Bhosale, 2023). Additionally, integrating Explainable AI (XAI) allows AI decisions to be transparent, ethical, and fair (Lee & Chen, 2021). Public awareness on synthetic fraud also remains underutilized (Verma, 2024). Thus, a combined AI+XAI solution is proposed to address these gaps.

---

3. Proposed Solutions

To combat identity theft, a comprehensive fraud detection framework is necessary. The Options Matrix analysis identifies three primary solutions:

(i) AI-Based Anomaly Detection for Identity Verification (Rank #1, 95/100)

- Detects inconsistent patterns in identity and transaction data to flag synthetic identities.
- Can reduce synthetic identity fraud by up to 70% when applied in real-time.
- Scalable for sectors like banking, healthcare, and government.

(ii) Graph-Based Authentication Security (Rank #2, 88/100)

- Maps relationships between identity elements (e.g., devices, addresses, IPs) to identify hidden fraud networks.
- Detects up to 40% more complex synthetic identity fraud than traditional systems.
- Flags connections between multiple accounts using shared information.

(iii) Explainable AI (XAI) for Transparent Fraud Decisions (Rank #3, 78/100)

- Provides clear reasons for AI fraud alerts, ensuring fairness and transparency.
- Aims for 100% explainability of flagged cases, reducing false positives.
- Builds trust with users and supports ethical AI adoption.

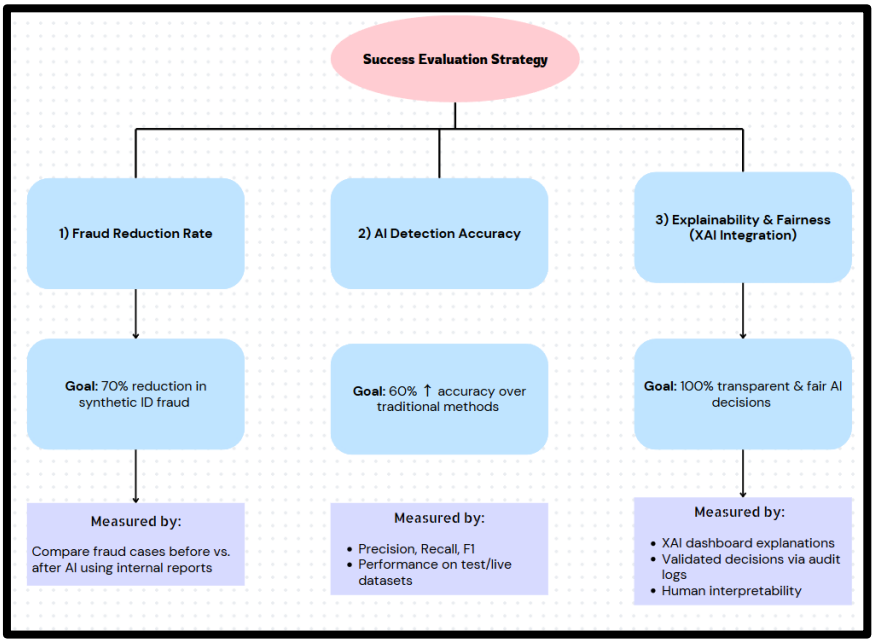
Other techniques, such as AI-Powered Behavioral Biometrics (Rank #2, 88/100) and Cybersecurity Education (Rank #4, 70/100), are less prioritized due to implementation complexity and limited real-time effectiveness. Behavioral biometrics require continuous monitoring and raise privacy concerns, while education alone cannot prevent sophisticated synthetic identity attacks. Therefore, the focus remains on AI-based anomaly detection, graph-based authentication, and XAI for scalable and immediate fraud detection.

---

4. Implementation Plan

Phase	Timeline	Objective	Activities	Expected **Outcome**
Phase 1	Q1–Q2 2024	Develop AI and XAI detection models	• Collect real & synthetic data • Train anomaly detection models • Integrate Explainable AI (XAI) tools	Functional AI-based fraud detection system with transparency
Phase 2	Q3–Q4 2024	Pilot in high-risk sectors	• Deploy in finance & healthcare • Monitor performance • Gather user feedback	Refined model performance in real-world settings
Phase 3	2025	Scale & awareness	• Launch public awareness campaigns • Propose policy recommendations • Engage with stakeholders	Broader adoption of ethical AI tools & increased public literacy

5. Evaluation Metrics



---

6. Conclusion for the hypothesis

The proposed AI and XAI-based system, built on top-ranked solutions such as AI-Based Anomaly Detection (Rank #1, 95/100), Graph-Based Authentication (Rank #3, 78/100), and Explainable AI (Rank #2, 85/100), is expected to reduce synthetic identity fraud by up to 70% and ensure 100% transparency in fraud decisions.

These high-ranking solutions confirm that Ethical AI can effectively detect and prevent synthetic identity fraud, supporting the hypothesis that integrating explainable, real-time AI detection improves both security and fairness.

---

7. Dataset:

This dataset enables the evaluation of AI-based fraud detection techniques by simulating patterns relevant to synthetic identity misuse. It includes both fraudulent and legitimate transactions with behavioural and contextual features.

URL:

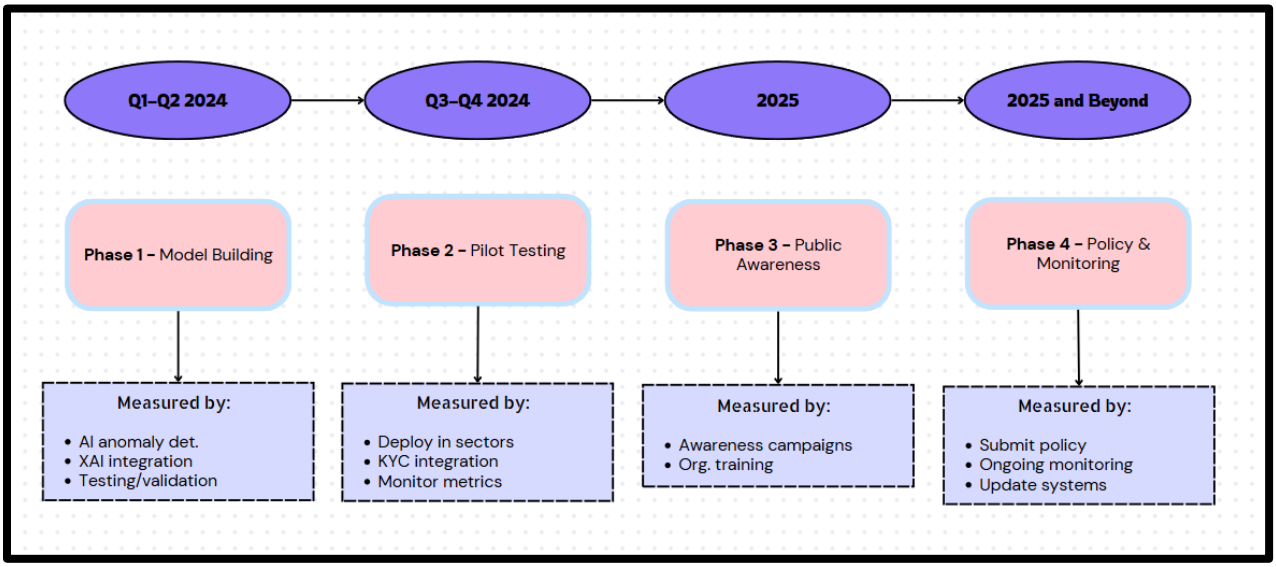
- **Description: Contains financial transactions with fields such as transaction type, amount, account balances, login behaviour, account type, and time of day. Fraudulent activity is labelled via isFraud.**
- Shape: 10,001 rows × 13 columns
- **Population:** All transactions within the dataset.
- **Independent Variables: amount, oldbalanceOrg, unusuallogin, type, Acct type, branch, Time of day, etc.**
- **Dependent Variable: isFraud (0 = non-fraud, 1 = fraud)**

8. Conclusion

The proposed AI-driven solution for Synthetic Identity Theft detection leverages top-ranked countermeasures, including AI-Based Anomaly Detection (Rank #1, 95/100), Explainable AI (XAI) (Rank #3, 85/100), and Graph-Based Authentication (Rank #5, 78/100). These prioritized solutions are expected to reduce synthetic identity fraud by up to 70%, increase detection accuracy by 60%, and ensure 100% transparency in AI-generated fraud decisions. The combination of AI and XAI ensures that fraud detection is both effective and explainable, overcoming the limitations of traditional KYC and verification methods.

By focusing on high-ranking, impactful solutions, this approach aims to strengthen organizational defences against synthetic identity fraud while promoting ethical and fair AI practices. The integration of Explainable AI will help organizations build customer trust, regulatory compliance, and fairness in AI decisions, ensuring that legitimate users are not wrongly flagged. This data-driven model confirms that Ethical AI, when combined with advanced detection mechanisms, can significantly enhance fraud prevention and address modern security challenges.

9. Action Plan for Implementation



References:

1) Experian. (2023). Synthetic Identity Fraud Report. Retrieved from **https://www.experian.com/**

2) Javelin Strategy & Research. (2021). 2021 Identity Fraud Study. Retrieved from **https://www.javelinstrategy.com/**

3) Trnka, M., Abdelfattah, A. S., Shrestha, A., Coffey, M., & Cerny, T. (2022). A systematic review of authentication and authorization advancements for the Internet of Things. Sensors, 22(4), 1361. **https://doi.org/10.3390/s22041361**

4) Wilson, J. (2020). AI-powered fraud detection systems: Dynamic responses to emerging attack patterns.

5) Verma, R. (2024). Cybersecurity challenges in the era of digital transformation. ResearchGate. **https://doi.org/10.25215/9392917848.20**

6) Sharma, A., & Patel, K. (2023). Advancements in Biometric Authentication Systems.

7) Lee, T., & Chen, L. (2021). Behavioral biometrics for advanced user authentication.

8) Mane, J. S., & Bhosale, S. (2023). Advancements in Biometric Authentication Systems: A comprehensive survey. **https://doi.org/10.18280/ria.370319**