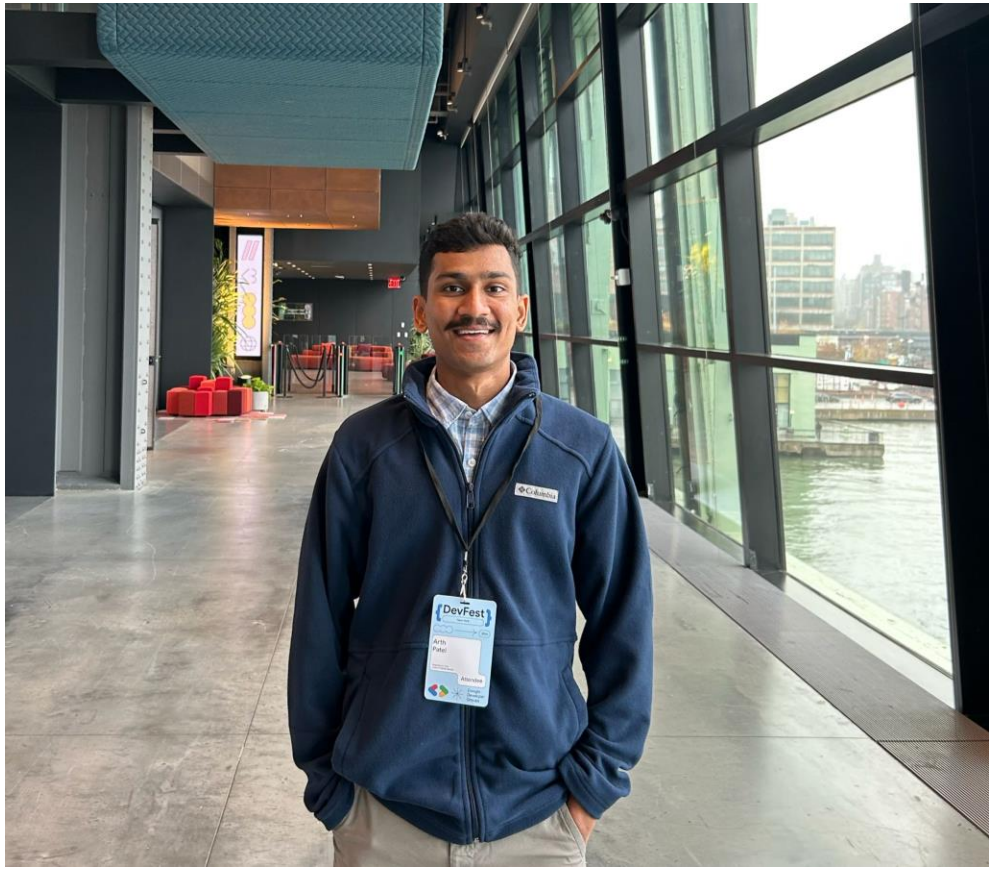
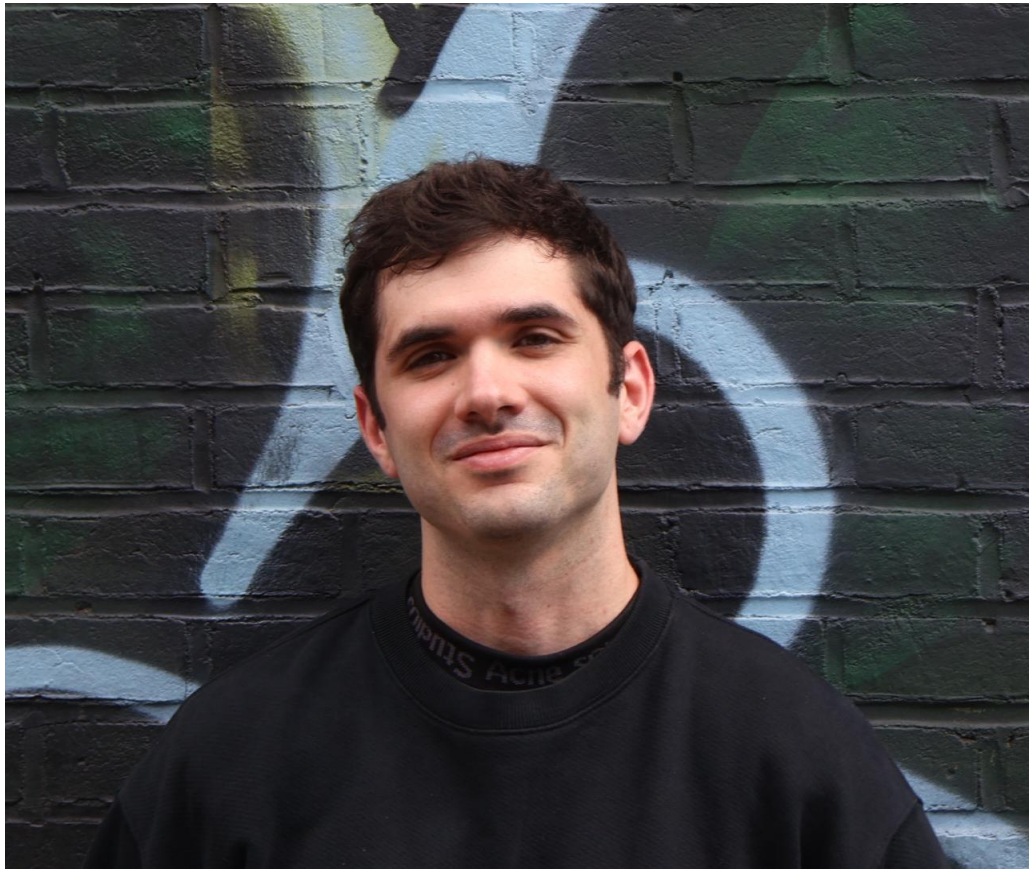
IT Architects
Credit Card Fraud Detection

Team 1: Arth Patel, Jay Patel, Tyler Helmrich, Xavier Jackson, Zenil Moradia

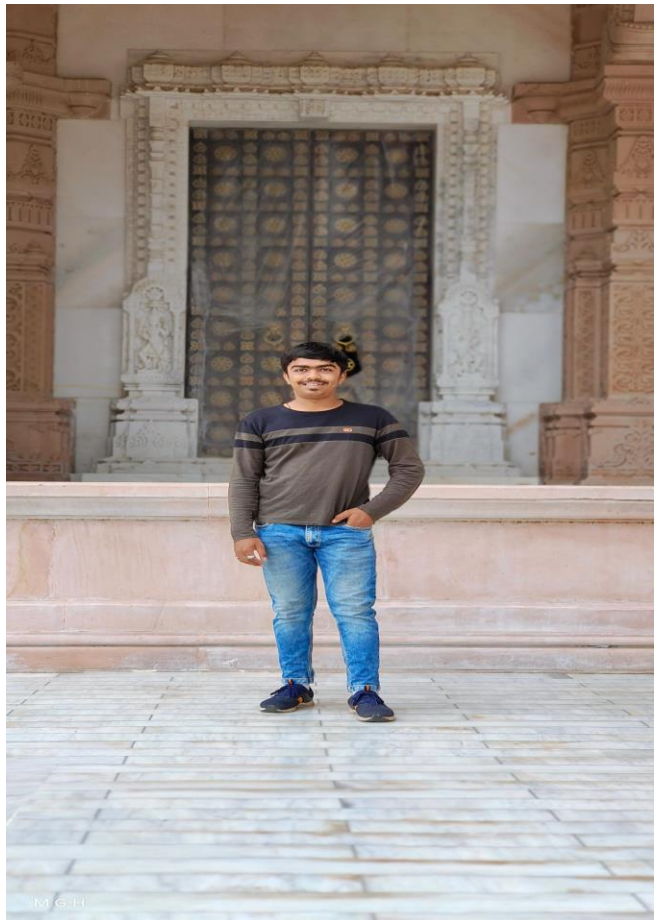
Team Members



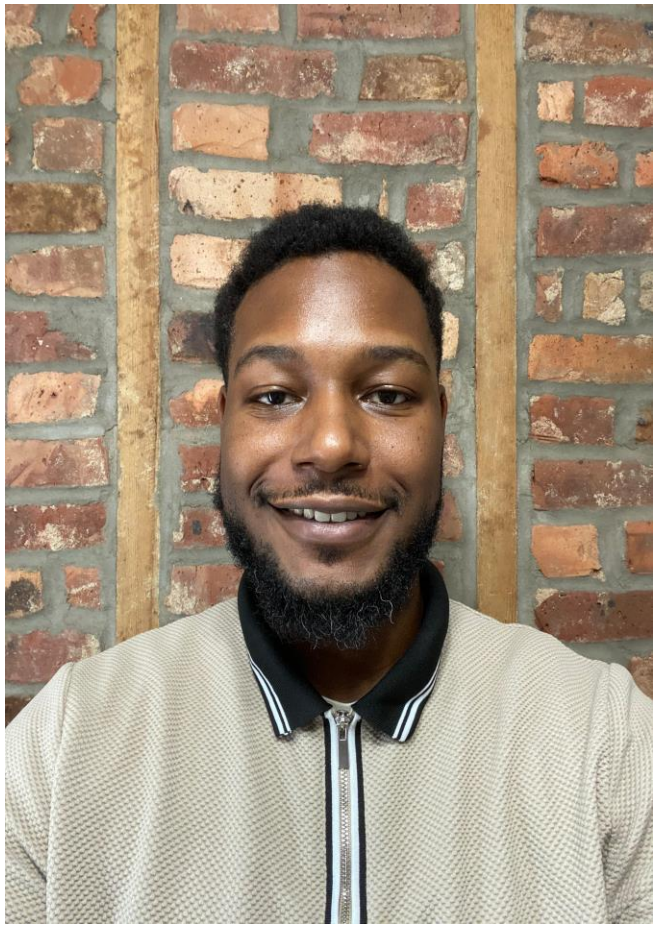
Mr Arth Patel



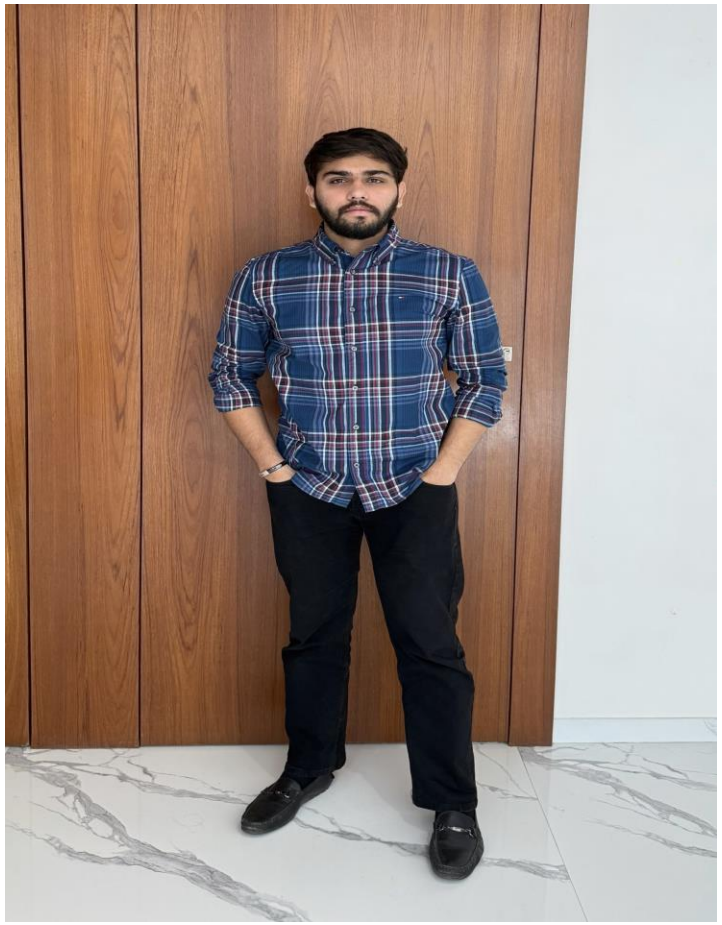
Mr Tyler Helmirch



Mr Jay Patel



Mr Xavier Jackon



Mr Zenil Moradia

Problem Statement

Credit card fraud remains the leading form of identity theft in the United States, with 2024 seeing over 449,000 complaints to the Federal Trade Commission and within the first quarter of 2025 adding another 151,000 cases. In the United States, 63% of cardholders have been targeted at least once, and more than half of these victims report experiencing repeated attacks.

This persistent threat not only inflicts hundreds of millions of dollars in annual losses but also erodes consumer confidence in digital payments. Existing detection systems struggle to keep up with the evolving tactics of scammers and the extreme imbalance between legitimate and fraudulent transactions. There is an urgent need for a robust, real-time solution that can accurately flag while minimizing false alarms.

Problem Description

Project Name:	Machine Learning Model for Credit Card Fraud Detection
Team:	IT Architects
Project Description:	Current deterrents include rule-based fraud detection techniques, whereas our goal is to implement an AI-powered fraud-detection service that catches up to 80% of fraud attempts in real time while cutting false positives by 30%, so issuers lose less money and cardholders enjoy seamless, secure spending. This project aims to deliver an advanced machine-learning classifier for credit-card fraud detection, designed for issuers and cardholders frustrated by today's rule-based systems and high false-alarm rates. Leveraging SMOTE to rebalance data and XGBoost for precision, we'll identify fraudulent transactions before funds leave the account, reduce false positives on legitimate purchases, and trigger real-time alerts that let cardholders act immediately. Our mission is to replace rigid, hand-tuned rules with a data-driven model that adapts to evolving fraud patterns by minimizing losses, preserving reputations, and restoring trust in digital payments.
Benefit Outcomes:	Credit Card companies can expect a reduction in false alarms, monetary losses, and damage to their reputation. Credit Card holders can expect a reduction in unauthorized charges, fewer declined legitimate transactions, and greater peace of mind when using their cards as well as an increase in user satisfaction and engagement.
Github Link:	https://github.com/htmw/F2025-Team1/wiki

Persona 1: Sarah Miller

- Age: 28
- Occupation: Graphic Designer
- Sarah is an avid online shopper who loves finding unique items from small businesses. She uses her credit card for most of her purchases, appreciating the convenience and rewards. She's tech-savvy and expects seamless digital experiences. Recently, her card details were stolen from a third-party website, leading to several unauthorized small purchases. The process of getting a new card and updating her payment information on various subscriptions was a significant hassle and made her wary of future online transactions. She's looking for reassurance and security in her digital financial interactions.
- Goals:
  - Wants to shop online without constant worry about card fraud.
  - Needs quick and clear notifications for any suspicious activity on her accounts.
  - Hopes for an easy way to verify transactions or freeze her card directly from her phone if needed.
- Frustrations:
  - The time and effort involved in cancelling cards and updating payment details after fraud.
  - Feeling vulnerable and exposed when her financial information is compromised.
  - Receiving generic, unhelpful alerts that don't provide clear next steps.
- Needs from the Solution:
  - Real-time alerts that are easy to understand and act upon.
  - An intuitive mobile interface that allows her to quickly approve or deny a transaction and temporarily lock her card.
  - Minimal false positives on her legitimate purchases so she doesn't lose trust in the system.

Persona 2: Alex Chen

- Age: 32
- Occupation: Founder of "PetPals Boutique" | The E-commerce Startup Founder
- Alex recently launched "PetPals Boutique," an online store selling custom pet accessories. Business is booming, but he's starting to see a disturbing trend of chargebacks due to fraudulent orders. These chargebacks not only cost him money but also damage his store's reputation with payment processors. He's passionate about his business but has limited resources and can't afford a dedicated fraud team. He needs an automated, intelligent system that can protect his revenue and allow him to focus on growing PetPals Boutique, not fighting fraud.
- Goals:
  - To minimize chargebacks and financial losses from fraudulent transactions.
  - To protect his business's reputation with payment providers.
  - To have a simple, affordable, and effective fraud detection solution that integrates easily with his e-commerce platform.
- Frustrations:
  - The financial impact of fraudulent orders and chargeback fees.
  - The time-consuming process of manually reviewing suspicious orders.
  - Fear that his small business could be targeted by organized fraud rings, which he feels ill-equipped to handle.
- Needs from the Solution:
  - An easy-to-integrate API that works seamlessly with his e-commerce platform.
  - A dashboard that provides clear, actionable insights on potential fraud, without requiring a lot of technical expertise.
  - Automated decision-making that blocks high-risk transactions instantly and allows low-risk ones to proceed without delay.

Persona 3: Marcus Johnson

- Age: 49
- Occupation: Head of Fraud Prevention at "SecureBank" | The Bank Security Officer
- Marcus leads the fraud prevention department at a mid-sized bank. His team handles thousands of credit card transactions daily, constantly battling sophisticated fraud schemes. With new tactics emerging all the time, his biggest challenge is staying ahead of the criminals while ensuring legitimate customer transactions are processed smoothly. He's looking for a powerful, scalable solution that can integrate with their existing systems, reduce false positives, and provide actionable intelligence to his team. He needs to improve detection rates, reduce operational costs associated with fraud investigation, and enhance customer trust.
- Goals:
  - To significantly reduce the bank's annual losses from credit card fraud.
  - To decrease the rate of false positives, improving customer experience and reducing operational overhead.
  - To implement a system that provides real-time, adaptive fraud detection capable of learning new patterns.
- Frustrations:
  - The constant arms race with fraudsters and the difficulty of keeping up with evolving tactics.
  - The pressure to maintain a high level of security without inconveniencing legitimate customers.
  - The sheer volume of data and the challenge of sifting through it to find genuine fraud.
- Needs from the Solution:
  - A highly scalable system that can process millions of transactions in real time.
  - Advanced machine learning capabilities that can adapt to new fraud patterns and minimize false positives.
  - Customizable rules and thresholds that allow his team to fine-tune the system's behavior.

Technologies

- Data Processing and Feature Engineering
  - **Pandas / NumPy** – Core tabular and numerical routines.
  - **Dask** – Scales Pandas workflows to larger-than-memory datasets with minimal code changes.
  - **Imbalanced-learn** – SMOTE, ADASYN, and other resampling techniques built atop scikit-learn APIs.
- Model Development and Experimentation
  - **Scikit-learn** – Standardized pipeline, preprocessing, and model-selection tools.
  - **XGBoost / LightGBM / CatBoost** – Gradient boosting implementations with GPU support for speed and advanced class-imbalance handling.
  - **Optuna / Hyperopt** – Efficient hyperparameter optimization via Bayesian search or Tree-structured Parzen Estimator.
  - **MLflow** – Track experiments, log parameters/metrics, and register model versions in one place.
- Visualization and End-User Interface
  - **Plotly / Dash** – Interactive EDA dashboards you can hook up to live model inference.
  - **Streamlit** – Superfast prototyping of apps – forms for inputting mock transactions, real-time results.
  - **Tableau / Power BI** – Drag-and-drop business dashboards for non-technical stakeholders, integrating CSV or database outputs.

Technologies Cont.

- Deployment
  - **FastAPI** – High-performance REST endpoints for batch or real-time scoring.
  - **Docker** – Containerizes entire pipeline (preprocessing → model → API) for consistency.
  - **Kubernetes** – Autoscale inference pods behind a load balancer; keeps service resilient.
  - **Terraform** – Infrastructure-as-code to spin up cloud resources reproducibly (EKS/GKE/AKS, databases, VPCs).
- Collaboration
  - **GitHub / GitLab** – Version control for code, issue tracking, and CI/CD integration with Actions or Runners.
  - **WikiPages**
- Data Storage & Logging
  - MongoDB

Algorithms

- Random forest
- Neural networks
- Sequence model
- Graph models like GNNs
- Graph features and Tree models
- Thresholding

Team Working Agreement

- Introduction
  - The following is the working agreement for the team "IT Architects". It defines how we plan to synchronize our work, timelines we will follow, professional expectations, and how we will handle group conflicts.
- Communication and Resource Sharing
  - WhatsApp will be our team's primary method of communication. Team members are expected to be responsive, and it will be considered unacceptable to go more than a day without responding to active conversation within WhatsApp without prior notice of unavailability.
  - For sharing documents either SharePoint or Google products will be acceptable. All code sharing will be done via GitHub.
- Work Division
  - Work will be divided during sprint planning and task assignment will be done based on interest and availability. No one will be expected to handle all tasks of a certain type or field of programming. All team members are welcome to work on all areas of the project.

Project Schedule

- Meetings
  - Sprint Planning
    - Sprint planning will be done at the start of the sprint where we will re-prioritize the product backlog, play planning poker to determine task difficulty, and assign tasks to group members.
  - Scrum Call
    - A 15 minute scrum call will be held daily at 7:30pm to realign the team on current project status. Prior notification is mandatory if a meeting will be missed.
  - Sprint Retrospectives
    - At the end of each sprint a sprint retrospective will be held to go over what did and did not go well over the course of the sprint.

Team Roles

- Scrum Master
  - Lead daily scrum calls and scrum retrospective meetings.

Team Member Name	Email
Tyler Helmrich	th4744n@pace.edu
Xavier Jackson	xj47172n@pace.edu
Jay Patel	jp57820n@pace.edu
Arth Patel	ap27525n@pace.edu
Zenil Moradia	zm70471n@pace.edu