Project Proposal for NRMA Fraudulent Insurance Claims Detection System

PACE Team 14 (PM: Noorullah Khan)

# Project Title:

GenAI-Powered Fraud Detection System for NRMA Insurance Claims

# Project Overview:

We propose to develop a machine learning-powered solution to assist NRMA insurance agents and operators detect fraudulent insurance claims. While the primary focus will be on insurance claims, the system will also cover other insurance claim categories managed by NRMA. The final deliverable will feature an intuitive form linked to a robust AI model that provides real-time feedback on the validity of each claim.

# Objectives:

1. **Develop a user-friendly form**: Agents will input claim details through drop-downs and text boxes.
2. **Integrate with a Machine Learning model**: The form will be connected to an AI model that assesses the likelihood of fraud.
3. **Provide actionable feedback**: The model will generate a probability percentage indicating the likelihood of a claim being fraudulent, prompting further actions based on the result.
4. **Continuous improvement**: The AI model will self-update using new data to enhance accuracy over time, considering economic events and seasonal variations.

# Key Features:

**User Input Form**:

* Details to be entered include:
  + **Car Insurance Claims**:
    - Year, make, model of the car
    - Age and gender of the driver
    - Driving experience
    - Previous incidents
    - Credit score
    - Yearly access of the driver
    - Number of other adults and minors in the car
    - Type of accident and damage
    - Day/time and weather of the accident
    - Estimated repair cost
    - Fault determination (applicant or other party)
  + **Home Insurance Claims**:
    - Property type
    - Age of property
    - Homeowner's age, gender
    - Incident type (e.g., fire, theft, water damage)
    - Previous claims
    - Security measures in place
    - Property location
    - Estimated damage cost
  + **Health Insurance Claims**:
    - Patient age, gender
    - Medical history
    - Type of treatment
    - Hospital/clinic details
    - Previous claims
    - Treatment cost
    - Doctor's notes
    - Date of treatment
  + **Life Insurance Claims**:
    - Policyholder’s age, gender
    - Cause of death
    - Beneficiary details
    - Policy details
    - Medical history
    - Date of death

**The model’s capabilities:**

* **Data augmentation:** a fraud detection problem has traditionally been a classification problem. A significant challenge in this domain is the imbalance between fraudulent and legitimate claims, with fraudulent claims typically being much less frequent.To address this issue, we propose employing a Generative AI model that will learn the characteristics of fraudulent claims and generate synthetic data that closely mimics these characteristics. By augmenting our dataset with this synthetic data, we aim to enhance the model's ability to detect fraud and improve overall detection performance.
* **Claims analysis**: Given that claims data can come in various forms, including images, documents, and voice recordings, we propose employing a multimodal Generative AI. Assuming our AI team is equipped to design and implement a multimodal Generative AI capable of encoding diverse data types, we can leverage this technology to perform the following tasks:
  + Summarisation: this will greatly reduce the amount of texts a Claims Manager will have to read, whilst keeping the relevant information
  + Enhance data representation: integrate information from different modalities (e.g., text, images, audio) to create a comprehensive representation of claims, improving model understanding and detection capabilities.
  + Cross-modal validation: validate and cross-reference information across different modalities to identify inconsistencies or anomalies indicative of fraudulent activity.
* **Extra -> Analyse interactions between company and customer:** Assuming the implementation of a multimodal Generative AI model, we can leverage its capabilities to analyse phone calls and chats with customers to extract valuable insights. Specifically, the model can:
  + Analyse phone calls: evaluate various aspects of customer phone calls, including language, tone, and attitude. This analysis can provide critical insights into customer sentiment and behaviour.
  + Enhance chat interactions: examine chat logs to identify patterns and nuances in customer communication. This can help improve service quality by providing a deeper understanding of customer needs and preferences, leading to more effective and responsive support.

**Continuous Learning and Improvement**:

* Since the model will not directly classify claims as fraudulent or legitimate, the primary goal is to assist humans in their decision-making process. To optimise the model's performance and accuracy, incorporating human feedback will be crucial, provided that feedback is honest and constructive. This approach is similar to the Reinforcement Learning From Human Feedback (RLHF) methodology employed by OpenAI.
* In practice, whenever the model raises a flag or generates a comment that proves inaccurate or incorrect, users can provide immediate feedback. This feedback can then be integrated into the model to refine its algorithms and improve its accuracy over time. By continuously incorporating user insights, the model can better support human decision-making and enhance overall fraud detection capabilities.

# 

# Project Conclusion:

By leveraging GenAI and machine learning, NRMA can significantly enhance its ability to detect fraudulent insurance claims, resulting in reduced financial losses and increased operational efficiency. The implementation of this system aligns with business intelligence and analytics (BIA) Principles, which emphasise the importance of data driven decision making. This approach will not only protect NRMA from potential fraud but also contribute to improved profitability and productivity. Additionally, minimising fraudulent claims supports fair pricing and policyholder trust, strengthening NRMA’s market position.

We look forward to collaborating with EY and NRMA to bring this innovative solution to life.

# Contact Information:

For further details or queries, please contact:

* **Project Manager**: Noorullah Khan
* **Data Engineer**: Quoc Hung (Alan) Nguyen
* **Data Engineer**: Aasnayem Gazzali Chowdhury
* **Business Analyst**: Tashiya Vilathgamuwa
* **Business Analyst**: Ninuri Mahagoda

We appreciate your consideration and look forward to your feedback.