

<b>Student:</b>	Kusuma Mekathoti
<b>Student ID and email:</b>	D3747948@live.tees.ac.uk
<b>Course:</b>	MSc Computer science (with Advanced practice)
<b>Supervisor:</b>	
<b>Second Reader:</b>	

1. Project Title

Credit card Fraud detection

2. What problem are you trying to solve?

1. Detect fraudulent credit card transactions accurately to prevent financial losses and protect customers.
2. Handle imbalanced datasets where fraudulent cases are rare compared to legitimate transactions, ensuring high precision and recall.

3. Research Question.

1. How effective are machine learning models in detecting credit card fraud in highly imbalanced datasets, and what techniques can be employed to improve their performance?
2. What is the role of feature engineering and feature selection in enhancing the accuracy and interpretability of credit card fraud detection models?

4. Hypothesis.

**Advanced Machine Learning Models:** Machine learning models (e.g., Random Forest, XGBoost, Neural Networks) will outperform traditional rule-based systems in detecting credit card fraud due to their ability to learn complex patterns.

**Handling Imbalanced Datasets:** Techniques like SMOTE (Synthetic Minority Over-sampling Technique) will effectively address the challenge of imbalanced datasets, where fraudulent transactions are rare.

**Improved Accuracy:** The proposed approach will achieve higher accuracy in distinguishing between fraudulent and legitimate transactions.

**Enhanced Recall:** The model will detect a higher percentage of actual fraudulent transactions, reducing the risk of undetected fraud.

**Reduced False Positives:** The system will minimize false alarms (legitimate transactions flagged as fraud), improving customer experience.

**Superior Performance Metrics:** The integration of machine learning and imbalanced data techniques will result in better overall performance metrics (e.g., F1-score, ROC AUC) compared to traditional methods.

## 5. Objectives of the project.

1. **To investigate previous research** on credit card fraud detection techniques, including traditional rule-based systems and modern machine learning approaches.
2. **To analyse, design, and implement** a machine learning-based fraud detection system that effectively handles imbalanced datasets and improves detection accuracy.
3. **To explore the possibilities** of advanced techniques like SMOTE, ensemble learning, and feature engineering for enhancing fraud detection performance.
4. **To evaluate and compare** the performance of different machine learning models (e.g., Random Forest, Logistic Regression, Neural Networks) in detecting fraudulent transactions.

## 6. Expected end products.

### Comprehensive Report:

- A detailed dissertation or project report documenting the research process, methodology, findings, and conclusions. This will include:
- Literature review on credit card fraud detection techniques.
- Explanation of the implemented machine learning models and techniques.
- Analysis of results, including performance metrics and visualizations.
- Recommendations for future improvements.

### Journal/Conference Paper:

- A condensed version of the research findings, suitable for submission to a relevant journal or conference in the fields of machine learning, cybersecurity, or financial technology.

### Fraud Detection Software Prototype:

- A functional software application or prototype that demonstrates the real-time capabilities of the fraud detection system. This could include:
- A user-friendly interface for uploading transaction data and receiving fraud predictions.
- Integration of the trained machine learning model for real-time fraud detection.

### Results of Technical Experiments:

- Performance evaluation of various machine learning models (e.g., Random Forest, Logistic Regression, Neural Networks) on the credit card fraud dataset.
- Comparison of techniques for handling imbalanced datasets (e.g., SMOTE, undersampling, class weighting).

#### Evaluation of Novel Techniques:

- A critical assessment of the effectiveness of advanced techniques like feature engineering, ensemble learning, and anomaly detection in improving fraud detection accuracy.

#### Data Analytics Insights:

- Extensive analysis of the credit card fraud dataset, including visualizations of transaction patterns, feature importance, and fraud trends.
- Insights into how data analytics can be used to identify and mitigate fraud risks.

#### 7. Ethical Considerations.

Ethical considerations ensure compliance with legal and professional standards. If human participants are involved, informed consent, confidentiality, and the right to withdraw will be ensured. Data security measures, including encryption and restricted access, will protect personal information. Risks, including psychological harm or bias, will be minimized. Compliance with GDPR and institutional policies is a priority. A risk assessment will determine if ethical approval via ERM is required. If low risk, no approval is needed; otherwise, an application will be submitted. Ethical safeguards will be maintained throughout the project to ensure integrity and participant protection.

##### 7.1. Ethics Declaration:

	True	False
My project is entirely literature based and/or technical,	True	
My Project does not use any form of participants,	True	
My Project does not use external inputs (e.g. liaising with someone in industry),	True	
My Project does not require me to do work off campus (e.g. in a company),	True	
My Project does not use secondary data sets		False

#### 8. Justification for Masters level project.

This project is suitable for a master's level study due to its depth, complexity, and contribution to knowledge. The following points justify its appropriateness:

**Novelty and Innovation:** The project explores a novel technique, application, or research gap, either by introducing a new approach or by enhancing existing methods. This level of originality is essential for postgraduate research.

**Advanced Research and Methodology:** It involves sophisticated research methodologies such as experimental design, machine learning models, statistical analysis, or qualitative assessments, demonstrating a high level of academic rigor.

**Depth of Theoretical and Practical Knowledge:** The project requires a strong theoretical foundation combined with practical implementation, ensuring a well-rounded understanding of the subject matter. This balance aligns with master's level expectations.

**Critical Analysis and Evaluation:** The work is not just an implementation but includes a detailed evaluation of results, comparisons with existing approaches, and discussions on limitations and future improvements. This level of critical thinking is a key component of postgraduate work.

**Complexity and Technical Challenge:** The project demand a high degree of problem-solving, requiring the application of advanced concepts, tools, or technologies. This complexity distinguishes it from undergraduate-level work.

**Contribution to Academic or Industry Knowledge:** The project has the potential to contribute new insights, recommendations, or solutions relevant to academia, industry, or both. It may lead to a research paper, prototype, or practical application with real-world implications.

**Independent Research and Project Management:** Conducting this project requires self-directed learning, literature review, hypothesis testing, and project management skills, all of which are crucial for master's level research.