

# **AI-Powered Transaction Classification System - An intelligent solution for automated bank transaction analysis and nominal code assignment using machine learning**

**Lead supervisor:** [Dr Jing Wang](#)

**Co-supervisor(s):** Luke Melville

**School/Institute:** [School of Computing and Digital Technologies](#)

**Contact for applicant queries:** Jing Wang [jing.wang@shu.ac.uk](mailto:jing.wang@shu.ac.uk)

## **Project summary**

In today's financial landscape, businesses struggle to efficiently manage and categorise banking transactions. The current process relies heavily on manual intervention through outsourcing firms using bookkeeping software and spreadsheets. While existing solutions offer basic automation through bank rules, they lack sophisticated context-aware capabilities, especially for complex business scenarios.

This limitation is evident in common scenarios like business travel expenses, where multiple transactions are interconnected - from flights to meals and transportation. Similarly, managing transactions for social media influencers presents unique challenges due to mixed business and personal expenses. New vendors or locations also require constant manual review due to the system's inability to understand context.

To address these challenges, we propose leveraging machine learning technologies to create an intelligent transaction classification system through:

1. Analysing transaction patterns and contextual relationships between expenses using natural language processing to extract features from structured and non-structured data.
2. Automatically assigning appropriate transaction tags (i.e., nominal code) such as "vehicle cost", "travel cost", "sale" and "direct expenses".
3. Learning from historical data to improve accuracy over time, adapting to new vendors and evolving spending patterns.

This solution aims to reduce manual processing time, minimise errors, and provide consistent transaction categorisation across client accounts.

In this project, the intern will gain experience in data analysis, financial principles, and project management, which ultimately will offer research experience for those who are considering applying for a Postgraduate Research (PGR) degree in the business and data science area.

They will learn data preprocessing, pattern recognition, bookkeeping principles, and transaction classification while developing project management skills through agile methodologies and documentation practices. This combination of technical, financial, and organisational competencies will provide valuable experience for future roles.

## **Specific skills and experience required for this project**

*Please also refer to the advert on our jobs pages for the person specification for these internships*

We are seeking an individual who can demonstrate the following capabilities and experience:

### **Technical Skills**

1. Strong programming skills in Python, with demonstrated experience in data analysis libraries (pandas, numpy) for handling large datasets and statistical computations
2. Solid understanding of machine learning frameworks (scikit-learn, TensorFlow, or PyTorch), including experience implementing supervised learning models and natural language processing techniques
3. Demonstrated experience with version control systems (Git), including branching strategies, pull requests, and collaborative development practices
4. Practical experience with software development, including knowledge of software engineering principles, agile methodologies, test-driven development practices, and the ability to write clean, maintainable, and well-documented code that follows industry best practices
5. Advanced proficiency in Excel file manipulation and automation, including experience with VBA macros, COM object model interactions, and Excel plugin/add-in development using frameworks like VSTO or Office.js, with demonstrated ability to create robust solutions for data processing, analysis, and user interface customisation within the Excel environment

### **Analytical Skills**

1. Strong problem-solving skills with the ability to break down complex challenges into manageable components and develop systematic solutions  
Experience in data cleaning, preprocessing, and feature engineering, with the ability to identify and handle outliers, missing data, and inconsistencies
2. Capability to interpret and analyse results, draw meaningful conclusions, and communicate findings effectively through visualisations and reports
3. Understanding of statistical concepts and their application in data analysis and machine learning
4. Critical thinking skills to evaluate different approaches and methodologies in solving research problems
5. Experience in conducting literature reviews and synthesising information from academic sources

## **Project location**

City Campus

Home working may be available

## **Project delivery**

This project can be delivered on a full-time or part-time basis, minimum 3 days per week.