Institution: University of Surrey
Unit of Assessment: UOA 11 Computer Science and Informatics
Title of case study: **Financial Fraud Detection**

1. Summary of the impact (indicative maximum 100 words)

Payment card fraud is a significant cost to business, as well as being a route to funding of organised crime, drug smuggling and terrorism. Detection of fraud requires a technique that is both transparent and adaptive. We have used the Department of Computing's expertise in machine learning and rule induction to develop a scalable method of automated fraud detection that meets the industry's needs. This technique is now being commercialised by AI Corporation, with a contract for its use having been placed by the world's largest retailer. Contracts with major banks are currently under negotiation.

2. Underpinning research (indicative maximum 500 words)

The research at Surrey that underpins this case study employed a combination of genetic algorithms to induce association rules and Bayesian networks to identify the most likely explanation of an invalid record [1]. This theme of strategic use of combinations of techniques, rather than focusing on fine-tuning a specific technique, was subsequently applied to protein structure prediction [2]. This was enabling us to build up a body of experience on the use of machine learning approaches that (a) can be used to provide explanations as well as classification decisions, and (b) are effective in situations where the data is highly imbalanced.

This set a foundation for our work in Financial Fraud detection [3], which has a number of key requirements:

- Ability to process a large quantity of heterogeneous and noisy data;
- Support for fast decision making;
- Adaptability to changing patterns, being able to identify interesting new relationships dynamically;
- Overall the ability to detect and explain anomalous behaviour, where the patterns for that behaviour are extremely sparse in a sea of legitimate transactions.

Meeting these requirements needed a combination of symbolic and connectionist approaches. However, research into symbolic rule extraction has a tendency to be based on small-scale datasets (because these are more readily available, and tractable to work with). Our concern was that these techniques were typically unable to produce a small set of comprehensible rules when applied to the very large-scale data sets that we needed to be able to handle in the domain of payment card fraud. This need for the analysts to be able to inspect and understand the rationale for the decision boundaries in an automated system is key to its acceptability in this industry.

Our approach to rule extraction uses sensitivity analysis to avoid the exhaustive decision boundary searches of other rule extraction algorithms, and so is computationally efficient – a critically important requirement given the extremely high volume of card payment transactions that need to be processed. We had access to a real-world dataset of 60m transaction records, of which 4,000 were frauds (amounting to a value of €1m). The fraud analysts we were working with found that the rules generated by our approach were easy to understand, and identified both already known patterns as well as previously unknown patterns of fraud [4].

This combination of connectionist and symbolic reasoning lends itself to the application of the various optimisation strategies that are under development in the Department and this is an area of current activity [5].

Key Researchers and positions at Surrey University:

Prof. Paul Krause (2001 – Present)

Nick Ryman-Tubb (PhD Student, Jan 2010 – Present)

Prof. Yaochu Jin (2010 – Present)

3. References to the research (indicative maximum of six references)

[1] Pantziarka, P., Machine Learning and Data Validation, PhD Thesis, University of Surrey, 2005. Supervisor: Paul Krause

[2] Zhang F, Povey D. and Krause P.J., "Protein Attributes Microtuning System (PAMS): an effective tool to increase protein structure prediction by data purification", In proceeding of: Digital EcoSystems and Technologies Conference, 2007. DEST '07.

[3] Ryman-Tubb N.F. and d'Avila Garcez A., "SOAR – Sparse Oracle-based Adaptive Rule Extraction: Knowledge extraction from large-scale datasets to detect credit card fraud", Proc. IJCNN, Barcelona, Spain, July 2010.

[4] Ryman-Tubb N.F. and Krause P.J., "Neural Network Rule Extraction to Detect Credit Card Fraud", In: Engineering Applications of Neural Networks, IFIP Advances in Information and Communication Technology Volume 363, 2011, pp 101-110.

[5] Inden B., Jin Y., Haschke R., Ritter H. and Sendhoff B., "An examination of different fitness and novelty based selection methods for the evolution of neural networks", Soft Computing, pp. 1-15, 2013.

4. Details of the impact (indicative maximum 750 words)

Fraud is a serious and long-term threat to a peaceful and democratic society. A recent Europol (2012) report (source 1) estimates that payment card fraud brings organized crime groups in the EU an income of around €1.5B. Businesses use a range of methods to detect this, mostly based around the usage of an automated rules-based Fraud Management System (FMS). However, the generation of these rules is an expensive and time consuming task, and fails to address the fraud problem where the data and relationships change with time. The latter is typically the case, as credit card fraud is a highly organized crime with strategies being steadily adapted as criminals discover which forms of fraudulent transaction are being detected.

The AI Corporation, based in Guildford, Surrey is a leading provider of rule-based FMSs. AI Corporation works with major banking and retail customers around the world, who together process more than 20 billion payment card transactions a year. Indeed in the UK alone AI customers process 95% of acquiring card transactions. It was founded in 1998, but since that time the core technology that underpins its products has remained essentially unchanged. Consequently, by late 2012 its once rapid growth as a company had stagnated and there had been virtually no investment in new products or technology.

The AI Corporation was acquired at the beginning of 2013 by a team of investors who saw the growth potential of the company should finance be made available to update its products and services. The Department of Computing's research on payment card fraud detection is a central part of the new management's product roadmap. The core IP for a product based on this research lay with a spinout of the Department, Thoughtified. Consequently, the team of investors agreed to fund the acquisition of Thoughtified by AI Corporation in order to provide the latter with Thoughtified's capability in predictive analytics and visualization of big data to productise the Department of Computing's research in the automated detection of credit card fraud. This was phase II of the strategy to revitalize AI Corporation. This strategic relationship with the Department of Computing through Thoughtified was an important part of the investor's original acquisition decision.

We have also seen a significant upturn in AI Corp's business as a result of this work. As of July 2013 a £970k contract had been placed with AI Corporation by Shell UK for the new product. Further contracts are in an advanced state of negotiation with: Global Payments Inc (est. £500k); Barclays Bank (est. £250k); First Rand Bank (est. £100k).

5. Sources to corroborate the impact (indicative maximum of 10 references)

1. Payment Card Fraud in the European Union, https://www.europol.europa.eu/sites/default/files/1public_full_20_sept.pdf
2. Chairman/Investor, aiCorp (contact details provided)
3. CEO, aiCorp (contact details provided)
4. Global Cards Central Delivery Manager, Shell UK (contact details provided)
5. Sr. Product Manager/Risk and Fraud Systems, Global Payments Inc (contact details provided)