Cardiff University
School of Computer Science and Informatics

Initial Plan

Credit Card Fraud Detection Using Novelty Detection Techniques.

CM3203 - One Semester Individual Project (40 Credits)

Author: Sara Essa Aldossary

Supervisor: Yuhua Li

Moderator: George Theodorakopoulos

4th February, 2019

Required Length: Maximum of 2000 words
Word count: 1546 words

Table of Contents:

Project Description

Background Information:

Nowadays, card fraud is increasing due to the prevalence of modern technology. As (Financial Fraud Action UK 2018) stated, the card fraud losses total was around 567 million in 2015, and it reached a 9% increase in 2016 with a loss of 618 million. Enterprises and public institutions must be well prepared to defend against such frauds that cause the loss of billions of dollars worldwide annually. Automatic systems to detect and prevent against card fraud are a significant tool in the financial industry's battle against card crime. However, it is not easy, or always possible to detect fraudulent patterns in transaction data by programmatic rules based analysis or inspection by fraud analysts; especially in large datasets.

Credit card usage is increasing year on year, and the credit card and card payments market plays a huge role in today's economy. Predicted statistics by (The UK Cards Association 2017) indicated that credit card payments will increase over the next decade. The predicted transactions are grown up to 3.7 billion in 2026, from 2.8 billion in 2016. Which consequently indicated that economic growth is one of the significant leads of the credit card volumes and values future. Moreover, credit card usage has well known advantages such as; simplicity of payment, access to credit, purchase guarantees and financial management to name a few, unfortunately these do not come without the risk of becoming a victim of fraudulent transactions. To reduce credit card fraud, many technologies can be used, such as rules based detection engines, Machine Learning and novelty detection techniques such as Clustering, Classification based and Nearest Neighbour (Miljković 2010). Effective technologies and methodologies that could detect frauds and illegal activities such as; money laundering, have been delivered and applied earlier successfully as detailed here: https://projecteuclid.org/download/pdf_1/euclid.ss/1042727940

It is important to remember during the project as a whole, that no autonomous system will be able to categorically define a transaction as fraud. The objective will be to highlight those transactions that have a high probability of being fraudulent based on some criteria, known or otherwise learnt.

Brief Description:

The aim of this project is to create a program that detects and identifies potentially fraudulent credit card transactions from a given dataset. The program will be trained with the given dataset, using standard Novelty Detection or Machine Learning techniques.

Project Aims and Objectives

This section will show in detail the main aims of the proposed project throughout the given time frame.

1. Establish a background in fraud detection novelty techniques and machine learning methodologies that need to be used in terms of achieving the solution.
   - Identify a software framework and tools needed to develop the engine for the project (e.g. programming language, tools, framework and techniques).
   - Gain an understanding of the software to be used throughout the solution development.
2. Implementing the Credit Card Fraud Detection engine.
   - Design a Reader for the input, a model for transaction data and card data.
   - Design a data store for transactions and cards.
   - Design the machine learning algorithm that classifies the data into potential fraud and otherwise by training the machine with the given data set.
   - Implement the algorithm using chosen framework and tools.
   - Validate the Implementation.
3. Documentation and evaluation on the established solution.
   - Test the program with different datasets to verify the solution works and trained properly.
   - Introduce any possible code enhancement and improvements on the solution.
   - Performance Analysis
4. Produce the final report, which includes the previous background research on the project.
   - Identify evaluation analysis and results from the implementation testing and improvements throughout the different stages.

Supervisor Roles - Yuhua Li

Arrange and attend regular meetings to check the progress of each stage throughout the semester, as well as discussing any concerns with the supervisor, receive guidance, advice and feedback. The meetings should be scheduled for 30 minutes each week.

Ethics

As discussed with the supervisor, most likely it is not required to have an ethical approval since the dataset does not contain any sensitive personal information and it is created for public usage. However, an ethical part will be considered while working on both, the project and the final report.

The dataset is given, and was collected on 2013 by European cardholders has been provided to test the developed program with. It has been collected and analysed during a research collaboration between Worldline and Machine Learning group (Machine Learning Group – ULB 2018). It shows that out of 284,807 transaction, 492 frauds have been detected and occurred in two days. Moreover, the source given stated that due to data confidentiality the original features and more background information about the data cannot be provided.

Work Plan

This work plan is created to manage the project stages, the final report, and to deliver an efficient project. Divided by week, and including regular meetings with the supervisor throughout the semester, excluding the Easter break. It details tasks and milestones.

Week 1:

- Initial plan first draft writing.
- Meet the supervisor for further clarification of the project objectives and aims, and feedback on the initial plan draft.
- Finalise the initial report, considering the supervisor's feedback.

   Milestone 1:
   - Submit the initial plan (4th February 2019).

Week 2-4:

- Background Research.
- Start the implementation
- Meet the supervisor and discuss the initial steps and change if needed.

Week 5-7:

- Implementation and Development
- Meet the supervisor to discuss progress.
- Start first draft of final report.

   Milestone 2:

o Final report first draft.

Week 8-11:
- Implement and test data
- Write the report
- Meet the supervisor to discuss progress.

        Milestone 3:
        o Completed software and test data.

Week 12-13:
- Analyse and complete report
- Meet the supervisor to discuss final report.

        Milestone 4:
        o Second draft of final report.

Week 14-15:
- Finalise, proof read the final report.

        Milestone 5:
        o Submitting the Final Report.

Viva Week:
- VIVA demonstration preparation.

        Milestone 6:
        o Project Completion.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Initial Plan	■	●													
Background Research		■													
Start Implementation		■	■	■											
Implementation & Development					■	■	■								
Final report – 1st draft					■	■	■	●							
Completed software & test data								■	■	■	●				
Analyses												■	■		
Final report - 2nd draft														●	
Submit the report															●
VIVA presentation preparation															■

Gantt Chart

The Gantt chart above indicates the work plan graphically, along with the estimated time that each task may take to be completed. The yellow circles represent the project's milestones.

Milestones and Deliverables

The work plan stated number of deliverables and milestones within the project given time, the most significant ones will be mentioned in details in this section.

- Initial plan report submission – 4/2/2019
- First draft of some current and future work in the final report done during the implementation stages, as stated in the work plan, such as;
    - o Background research.
    - o Approach and implementation.
    - o Results and evaluation.
    - o Enhancement.
- The completion of the implemented Software.
- Second Final Report Draft.
    - o Completing the sections that have not been done yet, after completing the implementation.
- Final report submission – 10/5/2019
- Viva Presentation, to complete the project.

Risk Plan

The following risk plan indicates the possible risks that might occur throughout the project and affect the work plan.

Risk	Risk Level [Low, Medium, High]	Possibility of an event to occur	Solution
Data loss	High	Unlikely	Ensure having backups, that are updated regularly, of the project contents in different servers or drives.
Sickness	Low	Likely	The workload is fairly divided along with the time given for the project. So, if any sickness happened few days would be easy to make up later.
Being unable to submit the work by deadline in case of any encountered issues.	High	Somewhat Likely	Enough time range is given before the deadline 10th May, 2019, to resolve any encountered issues in the software.
Being behind the work plan.	Medium	Somewhat Likely	Trying to catch up with the tasks and do it week by week from the beginning, and if in case the week's tasks were more than expected, hence, not completed, a time from another week could be use.

Conclusion

To conclude, this report indicated the proposed final project description, background research, aims and objectives to reach the final solution, and the work plan following by a Gantt chart that the tasks and milestones for this project. Supervisor roles, and an ethical consideration as well as risk plan have also been discussed.

References

1- Bolton, R.J. and Hand, D.J. 2002. Statistical Fraud Detection: A Review. *Statistical Science* 17(3), pp. 235-255. Available at: https://projecteuclid.org/download/pdf_1/euclid.ss/1042727940 [Accessed: 2nd Feb, 2019].

2- Financial Fraud Action UK. 2017. *Fraud the Facts*. Available at: https://www.financialfraudaction.org.uk/fraudfacts17/assets/fraud_the_facts.pdf [Accessed: 1st Feb, 2019].

3- Machine Learning Group – ULB. 2018. *Credit Card Fraud Detection*. Available at: https://www.kaggle.com/mlg-ulb/creditcardfraud/home [Accessed at: 31 Jan, 2019].

4- Miljković, D. 2010. *Review of Novelty Detection Methods*. Available at: https://www.researchgate.net/publication/261424710_Review_of_novelty_detection_methods [Accessed: 2nd Feb, 2019].

5- The UK Cards Association. 2017. *UK Card Payment Summary 2017*. Available at: http://www.theukcardsassociation.org.uk/wm_documents/UK%20Card%20Payments%202017%20-%20Summary%20FINAL.pdf [Accessed: 2nd Feb, 2019].