Exploration of Data Science Toolbox and Predictive Models to Improve SCM Via Reduction in Thefts and Fraud

Adam Canton1, Benjamin Goodwin2, Babatunde Olanipekun,3

1 Master of Science in Data Science, Southern Methodist University,  
Dallas, TX 75275 USA

{bgoodwin, cantona, bolanipekun}@smu.edu

**Abstract.** Modern Supply Chain processes and systems are extremely complex and rely on a high degree of automation and due to this reliance on automation and complexity theft and fraud are often prevalent throughout the process. We aim to build AI predictive models for financial supply chain management applications and databases. The intention of this research aims to outline an approach to reducing fraud and theft via machine learning and artificial intelligence techniques.

1 Introduction

1. The UNLV School of Public Health, in collaboration with the Southern Methodist University research team seek insights on financial supply chain management (FSCM) applications and databases as to the root causes behind fraud and theft of these tools. FSCM applications and databases are large, complex systems that cannot offer robust fraud protection in their current form and have the potential to hugely benefit from machine learning and artificial intelligence techniques when attenuated to the detection and prevention of fraud and theft.
2. Fraud and theft plague organizations both large and small and modern supply chains span the globe, rendering traceability a difficult task for the implementing organization. Due to the sheer amount of data generated via FSCM applications and databases, we propose implementation of artificial intelligence and machine learning algorithms to reduce this fraud and theft, with the intentions reducing organizational overhead and manpower required to combat such issues.
3. Fraud and theft are often difficult to detect and pursue, especially in the case of globalized FSCM applications and databases. However, the subject of fraud and theft is an often researched and discussed problem. Using machine learning techniques other research has proven fraud and theft are possible to detect but come with their own set of challenges. An often-discussed issue with fraud detection is its relative rarity when compared to legitimate transactions. Credit card companies and large organizations with widely spread assets have deeply studied such topics and have developed formidable and complex anti-fraud and theft systems.
4. Among the various issues associated with solving such a problem revolves around the access to data. Large organizations are often hesitant to hand over data involving known instances of fraud or theft. The central issue revolves around the size and complexity of FSCM applications and databases and a supposed lack of overall accountability.
5. The SMU data science team believes research in the fields of artificial intelligence and machine learning to combat fraud in FSCM applications and databases. Applying AI/ML techniques to FSCM applications and databases when implemented correctly can and will identify situations where fraud and theft occur.
   1. This research aims to use artificial intelligence to solve fraud and theft across FSCM applications and databases.

2 Literature Review

This section is on the background information to help to reader understand why this research’s method will work. (Need to do)

**FOR THE OUTLINE ONLY – You do an annotated bibliography (1 paragraph summary of each article that will be used in the paper).**

2.1 Theme

2.2 Citations -READ ME

The list of references is headed “References” and is not assigned a number. The list should be set in small print and placed at the end of your contribution, in front of the appendix, if one exists. Please do not insert a pagebreak before the list of references if the page is not completely filled. An example is given at the end of this information sheet. For citations in the text please use square brackets and consecutive numbers: [1], [2], [3], etc. Use **APA format** in the reference section. You can choose to either have it alphabetical order or order of which it is shown in the paper.

**Hypothesis at the end of your literature Review**

3 Methods

1. Data
2. A search for enough and accessible data presents challenges. As of time of writing we are still in discussions with various industry leads, the research team unsuccessfully contacted HIMSS and Wynn Resorts. Project advisor Chris Papesh and SMU team members continue to contact various organizations with the goal of securing data for analysis.
3. Once acquired, data and analysis will follow the typical data science pipeline of data acquisition, data processing, data integration, analytical modeling, and validation. The SMU team expects to use various ML/AI models to determine a best fit for the detection of fraud and theft.

4 Results

1. Through the research of FSCM applications and databases, the SMU data science team hopes to generalize the common occurrences of fraud and theft throughout these systems using AI/ML. The SMU team will determine “good” results using various statistical tests as well as participating in a continuous feedback loop with stakeholders who regularly handle and identify fraud and theft.

5 Discussion

1. Interpretations: Are the results sufficiently convincing in their ability to predict and detect fraud and theft? As a secondary objective, do the results encompass all the vast supply chain data?
2. Implications: If the models can accurately detect fraud or theft, will organizations utilize and trust this new tool?
3. What stood out as interesting/unique/unexpected?
4. Limitations
   1. What challenges occurred during analysis?
5. Ethics
6. Future Research
   1. Are there areas of research where others can pick up and go deeper?

6 Conclusion

2 paragraphs max on the overall findings and summary of the research.

Acknowledgments. The heading should be treated as a 3rd level heading and should not be assigned a number.

References

1. Acciani, C., Fucilli, V., & Sardaro, R. (2011). Data mining in real estate appraisal: A model tree and multivariate adaptive regression spline approach1. Aestimum, (58), 27-45.
2. Carlucci, M., Grigoriadis, E., Venanzoni, G., & Salvati, L. (2018). Crisis-driven changes in construction patterns: Evidence from building permits in a mediterranean city. Housing Studies, 33(8), 1151-1174. doi:10.1080/02673037.2017.1421910
3. Davanloo Tajbakhsh, S., Castillo, E. d., & Aybat, N. S. (2015). On convex optimization methods for fitting spatial statistical models to large data sets ProQuest Dissertations Publishing.
4. Development of a forecasting model to predict the downturn and upturn of a real estate market in the inland empire (2013). Universal-Publishers.com.
5. Dheeriya, P. (2009). Modeling volatility in california real estate prices. IUP Journal of Applied Economics, 8(1), 26-38.
6. Dokko, Y., & Edelstein, R. H. (1992). Towards a real estate land use modeling paradigm. (special issue on office markets). Journal of the American Real Estate & Urban Economics Association, 20(2), 199.

Appendix

Use if needed for additional information