

7a.020.UVA - Deep Learning and Adversarial Learning in Credit Card Fraud

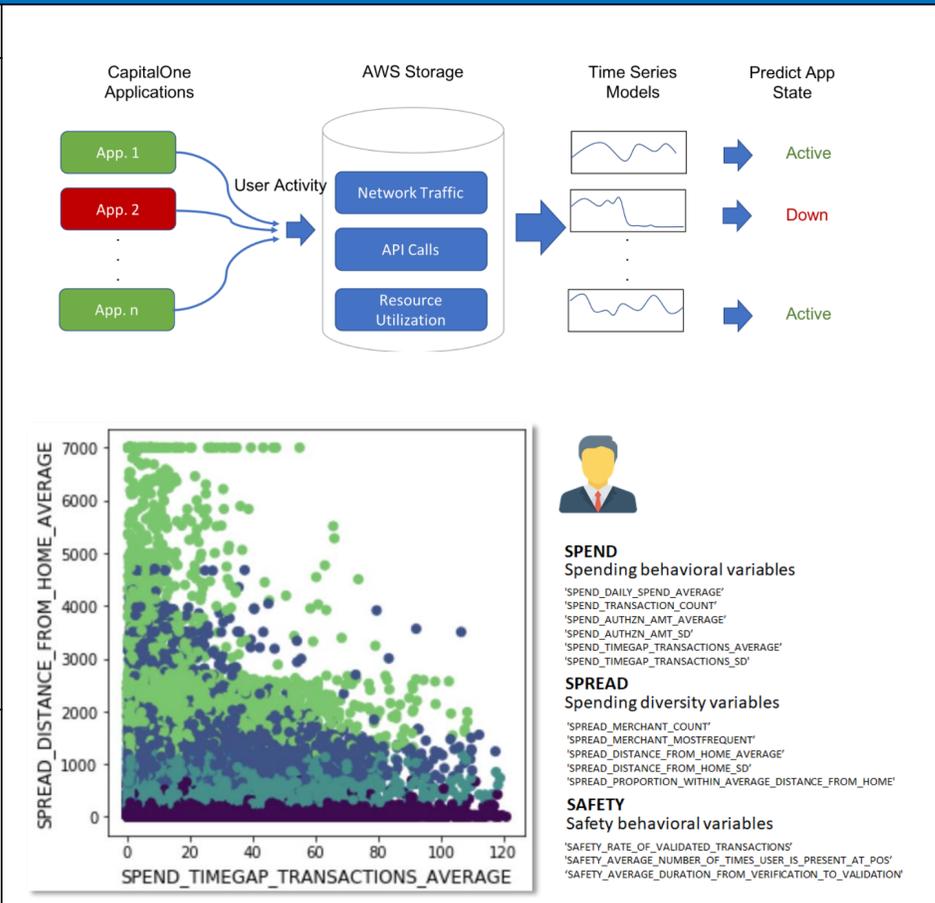
Peter Beling (PI), Stephen Adams (Co-PI), Brendan Abraham (Researcher), 2 DSI Capstone Teams
University of Virginia

Project Start: 8/1/2018	End Date: 7/30/2019	Project Budget: \$40K	Spent:
--------------------------------	----------------------------	------------------------------	---------------

Project Summary: In the United States in 2013 alone, credit card fraud cost companies almost \$7.1 billion dollars. Given these enormous costs, fraud detection and classification has become a very active area of research in machine learning and data mining domains. Although the power of machine learning techniques for fraud detection has greatly increased over the past decades, the incentives for fraudsters to circumvent and adapt to these classification algorithms has also grown. Effective fraud detection models must be able to adapt to behavioral changes on the part of the adversary, while maintaining high levels of accuracy and low levels of false positives. The credit card fraud team will investigate unsupervised methods with the goal of clustering customer types. The team will investigate how clusters can be used to improve fraud detection.

Capital One provides many different cloud-based applications for their customers. When an application goes down, it can lead to significant financial losses for both the company and customers. Each of these applications generates data such as network traffic, resource utilization statistics and API calls which are all stored in AWS. This data can be joined to form time series data to make useful inferences about the state of each application. The main objective is to build an application health-monitoring system that predicts when an application crashes based on the time series data to minimize downtime. Secondary objectives include catching as many crashes as possible while minimizing false positives.

Details of Progress/Achievements: The Capstone Teams have been formed and meet regularly. Both teams are still waiting on data from Capital One. The credit card fraud team is using a data set from previous projects to explore feature extraction and clustering methods. The network traffic team has performed a literature review and found some useful features to extract from network traffic. They have also found a third party network traffic dataset and plan on using it to develop the pipeline while we're waiting for the agreement to go through.



PROJECT DELIVERABLES

Deliverable	Achievements	Remaining To Do
Code for use at Capital One.	Both teams have begun scanning the literature. The fraud team is working with a dataset and exploring feature extraction and clustering techniques at an account level. The network traffic team is working with a third-party dataset.	Acquire new data sets from Capital One Perform analysis
Research summaries and presentations	Both teams will present their fall semester findings to the Capstone Class at the end of the semester.	Complete summaries and presentations
SIEDS papers	Work on the papers will begin in the spring	Write the paper