WINNIE YEUNG

(404) 268-4204
dcvionwinnie@gmail.com
vionwinnie.github.io

Summary

Experienced in developing and implementing end-to-end machine learning models with 5+ year working experience.

Technical Skills

Programming Languages:
Python | R | SQL | Scala | Java

Tools: Spark | Hive | Tableau | Docker | Hadoop MapReduce | Flask | Bash | git

Modeling Techniques:
Transformer (Pytorch, Tensorflow)
Natural Language Processing (NLTK)
Anomaly Detection
Machine Learning (scikit-learn)
Time Series Forecasting

Education

Georgia Institute of Technology
Master of Science in Analytics
Graduated in Aug 2019 , GPA: 4.0

Coursework:
Deep Learning | Machine Learning | Database | Bayesian Statistics | Deterministic Optimization

Leadership

- **Divisional Director** of National Collegiate Table Tennis Association
- **Founder and editor-in-chief** of first multilingual magazine at Middlebury College with 12 editors

Interests

Table Tennis, Biodynamic Wine, Goodwill North Korea Ambassador

Work Experience

Visa — Foster City, CA
Data Scientist - Payment Systems Risk — Sep 2019-Present

Proof-of-concept and Machine Learning System Design:

- Partners with business stakeholders to understand business contexts to design and implement end-to-end a customized financial news engine that uses Named Entity Recognizer and Text Classification models (TF-IDF,SVM, Word2Vec, Transformer)
- Presented the prototype to senior leadership to secure funding for the next fiscal year to expand the project for wider coverage

Model Serving/ Deployment:

- Deployed to production server for an existing proof-of-concept Tensorflow risk model for ATM cashout, adding features that allow more time-sensitive monitoring of model serving and data quality with Splunk and bash scripts
- Refactored Hive data pipeline using Spark to reduce processing time by 70%
- Tools: Scala, Spark,Tensorflow, Python, Java, bash scripts, Hive, Kafka

Safely (Insurance for short-term rentals startup) — Atlanta, GA
Data Scientist — May 2019-Sep 2019

- Design a risk score model for short-term rentals owning full modeling cycle from feature engineering, modeling to visualization.
- Visualize model scores using d3.js and flask to illustrate score distribution to allow business users to interpret the score and the monetary implication
- Identify clusters of key issues in insurance claims using Latent Dirchlet Allocation (LDA) topic modeling techniques which speed up claims processing time by 30%

Fidelity Investments — Boston, MA/Hong Kong
Equity Business Associate — Jul 2015-Mar 2018

- Established a proprietary algorithm in sentiment analysis to extract trillions of items of data resulting in major buy and sell decisions of S&P key stock for portfolio managers. Out-of-sample predictions of past five quarters were proved consistent with stocks' key metrics
- Spearheaded a project with portfolio manager to collect million of research report of stocks as training set and create sentiment algorithm to predict news headline on stock performance with a 60% hit rate on weekly forward return
- Authored 7 research reports to provide actionable insights on 20 S&P large-cap stocks ahead of earnings announcements based on alternative data research. Two papers were voted as most-read research of the week by 60 portfolio managers.

Other Projects

- **Multivariate Anomaly Detection with PCA (Spring 2019)**
  Researched on applying principal component analysis and variable selection method to detect and diagnose change in covariance matrix for large-scale sensor data during factory production process