ZACK STRATHE

zack.m.strathe@gmail.com | (724) 426 - 6503
Project Portfolio: zstrathe.github.io

Objective

Seeking full-time employment as a Data Scientist, where I seek to contribute toward innovation in developing actionable insights from data, and to continually hone my data science skillset by learning from experienced leaders in the field

Education

Kansas State University, Manhattan, KS / M.S. Data Analytics / Data Science track / In-Progress (August 2022 expected)

Selected Projects:

Satellite Image Feature Extraction & Classification Using PySpark

- Entirely using the Apache PySpark distributed computing framework in Python, evaluated methods of feature selection and compared performance of classification algorithms with a data set of 500,000 labeled satellite images, and implemented 10-fold cross validation with a paired t-test to validate evaluation results
- Deployed a Spark cluster on Amazon Web Services Elastic MapReduce to test functionality; utilized a Google Cloud Platform virtual machine with a 16-core vCPU and 64 GB RAM for development and evaluation
- Technologies used: Python, PySpark, MLlib, OpenCV, Numpy, GCP, AWS EMR

Training a Deep Reinforcement Learning Agent to Play Mario Bros

- Trained a deep reinforcement learning model to play the game Mario Bros, using OpenAI's Gym framework in Python, and evaluated methods of improving the trained Proximal Policy Optimization (PPO) model with modifications to the state-space, the action-space, and the reward function
- Developed and trained utilizing a Google Cloud Platform virtual machine with a 8-core vCPU and a Tesla T4 GPU
- Technologies used: Python, OpenAI Gym Retro, OpenAI Baselines, TensorFlow, GCP

Comparison of Deep Learning Text Generation Models Trained with Song Lyrics

- Trained unconditional text generation language models from a text corpus of song lyrics, utilizing recurrent neural networks (RNNs) and generative adversarial networks (GANs) in Python with the PyTorch deep learning framework, and evaluated text output by utilizing a combination of human scoring and a computed bilingual evaluation understudy (BLEU) score
- Technologies used: Python, PyTorch, TextBox module (GAN algorithms)

Statistical Analysis of Home Pricing with Linear Modeling in R

- Created a linear model in R to conduct statistical analysis of home pricing in Kansas City, MO, and evaluated the linear model compared to a more-complex generalized additive model (GAM) for predictive performance
- Technologies used: R, ggplot2

Business Capstone Data Visualization Project

- Coordinated student team in consulting with a client company to provide research findings into data visualization methods, and developed visualization dashboard implementations that provide enhanced insight into the status of their supply network
- Technologies used: Python, Plotly, PowerBI

Kansas State University, Manhattan, KS / B.S. Finance / Financial Management track / December 2012

Skills

Programming Languages:
Python, R, SQL

Data Analytics Tools:
Apache PySpark (distributed computing framework for Python), Pandas (data analysis Python package), NLTK (natural language processing Python package), Scikit-learn (machine learning Python package), Alteryx (data analytics platform)

Data Visualization Tools:
Tableau, Matplotlib (Python package), ggplot2 (R package)

Experience

Settlements Analyst II / Crestwood Midstream Partners LP, Kansas City, MO / April 2018 – January 2021

- Developed script in Python to automate process of entering expenses into cloud-based ERP system
- Produced monthly financial reconciliations for producer agency marketing agreements totaling approximately $200M annually

Settlements Analyst / Crestwood Midstream Partners LP, Kansas City, MO / August 2015 – March 2018

- Managed daily contract allocations and billing, and monitored accounting transactions for purchases and sales of LPGs for various marketing groups, and monitored marketing P&L statements to identify, analyze, and correct accounting errors

Billing Associate / HNTB Corporation, Kansas City, MO / October 2013 – August 2015

- Managed monthly billing process for approximately 75 architecture and civil engineering projects with varying fee structures

Certifications