Yi (Evers) Feng

Durham, NC | yi.feng-mids@duke.edu | (313) 420-8543 | Website | GitHub | LinkedIn

SUMMARY

Professional Data Scientist with experiences in HR and Product Analytics, A/B Testing, Machine Learning, and Cloud Computing.

SKILLS

Programming Languages: Python · R · MATLAB · SQL · SAS · Cypher · C++ · Java
Tools: GitHub · Spark · Scikit-Learn · Pandas · NumPy · TensorFlow · NLTK · Tableau · Neo4j · SPSS
Coursework: Causal Inference · Statistical Analysis · Probability · Natural Language Processing
Certifications: AWS Certified Cloud Practitioner [credentials] · Duke Innovation & Entrepreneurship Certificate

EDUCATION

Duke University   *Aug 2020 - May 2022*
- *M.S. in Interdisciplinary Data Science:* 3.91/4.0
    - Duke Interdisciplinary Product Management Club: *Founder & President*
    - Teaching Assistant for *Practical Data Science* [link]
    - Duke Student Founder Program: *Mentee*

University of California, Los Angeles   *Sep 2016 - Jun 2020*
- *B.A. in Sociology:* 3.92/4.0
- *B.S. in Cognitive Science:* 3.70/4.0   *Specialization in Computing:* 3.86/4.0

RESEARCH AND PROFESSIONAL EXPERIENCES

IEEE | Remote   *Present*
Capstone Project Manager & Data Scientist
- Develop an end-to-end pipeline on AWS to dynamically conduct topic modeling for IEEE's flagship product - IEEE *Xplore*.
- Establish measure of relevancy between concepts and papers, leveraging multiple large data sources and data lakes.
- Facilitate interdisciplinary field discovery by visualizing latent concept hierarchy within IEEE fields of study.

Truist Financial (formerly BB&T and SunTrust) | Durham, NC   *Jun - Aug 2021*
Data Science Intern
- Identified pain points in improving Talent Acquisition work experience by implementing LDA model on teammate survey data.
- Analyzed return-to-work public opinions in Tableau and created a data model with automated data ingestion in Python.
- Defined teammate persona profiling through combining Sentiment Analysis (*FLAIR*) and Random Forest model.

Duke Interdisciplinary Social Innovators | Durham, NC   *Nov 2020 - May 2021*
Project Innovator
- Provided technical problem categorization and data cleaning for the client - Engineering World Health.
- Informed client's decision-making in market expansion and product diversity by delivering a constructive market analysis.

Octi | Los Angeles, CA   *Jul - Sep 2019*
Software Behavior Analyst
- Pioneered a product launching plan on TikTok through detailed research, approved by the CEO.
- Supported core functionality of OctiGram by building a facial recognition processing tool using OpenCV. [code]

PROJECTS

CNN Integrated OCR on Street Sign Recognition [code]   *May 2021*
- Mobilized a team of five to deliver project milestones on time and attained high achievement as a team.
- Combined Convolutional Neural Network model in digit and character recognition to contribute to current OCR literature.

Serverless Data Engineering Pipeline [code]   *Sep 2020*
- Combined AWS Lambda Function and DynamoDB to extract game reviews from Steam, using SQS for scalable data storage.
- Applied Amazon Comprehend to conduct sentiment analysis on text data.

Humana-Mays Healthcare Competition | Top 50 Finalists [code]   *Sep 2020*
- Collaborated with a team of three and achieved top 15% nationwide on building a machine-learning-powered model to predict transportation difficulty from big data.
- Delivered a comprehensive report to interpret findings and offer business insights into resolving transportation difficulty.