HARI DEVANATHAN
DATA SCIENTIST

HD2ZM@VIRGINIA.EDU
hd2zm.github.io/
571-244-8781
hari-devanathan
hd2zm/Data-Science-Projects

EDUCATION

University of Virginia May 2015
Bachelors of Science Computer Science
Minor in Systems Engineering

SKILLS

PROGRAMMING: Python, SQL, NoSQL, Git
MACHINE LEARNING: Linear Regression, Logistic Regression, K-Nearest Neighbors, Decision Trees, Random Forest, Naive Bayes Models, Clustering, Principle Component Analysis, Natural Language Processing, Feature Engineering, Topic Modeling, Non-Negative Matrix Factorization (NMF), Term Frequency - Inverse Document Frequency (TF - IDF), ARIMA Time Series
DATA VISUALIZATION: Seaborn, Matplotlib, Tableau, Plotly
WEB DEVELOPMENT/SCRAPING: Javascript/Node.js, Flask, Selenium, Beautiful Soup
TOOLS AND PACKAGES: Pandas, NumPy, Skikit-learn, NLTK, SpaCy, Spark
DEPLOYMENT: AWS, Docker, Airflow

EXPERIENCE

METIS — Chicago, IL
Data Scientist — 2019 to 2019
- Developed five data science projects as part of a highly selective, ACCET accredited, 12-week immersive program.
- Focused on applying a wide array of machine learning and statistical modeling techniques.
- Performed data acquisition, cleaning, aggregation, exploratory data analysis, data visualization, feature engineering, and model creation.
- Communicated results and recommendations through a comprehensive presentation following each project.

GANNETT — Tysons, VA
Software Engineer II — 2017 to 2018
- Developed and deployed microservices for CMS platform, a tool for reporters to publish stories to USA Today.
- Collaborated with senior engineers on implementing CQRS and event sourcing for the architecture.
- Wrote UI features to aid journalists in creating content.

MICROSOFT — Seattle, WA
Software Engineer — 2015 to 2017
- Pitched, designed, and developed the Humanitarian Resource Dashboard admin feature to facilitate admin entry of different data input. Rewrote the backend using Azure Cosmos DB, and added frontend features using Angular 2.
- Optimized front end and middleware of Microsoft's SmartBuilding, an IoT Angular application for employees to book available conference rooms. Rewrote SignalR methods to include persistent connections and observation design pattern callbacks.
- Redesigned PartnerIncentive's existing website dashboard using Angular, and rewrote authentication using ADAL.

EPIC — Verona, WI
Software Development Intern — 2014 to 2014
- Implemented an iOS application that allows patients to instantly interact with next available doctors/nurses. Incorporated Microsoft's SignalR library to allow real-time functionality. Appended such project to Epic's existing MyChartMobile, an application that allows patients to view medical records, schedule appointments, etc.

SOFTEON — Reston, VA
Software Development Intern — 2013 to 2013
- Designed an Android test application on different data storage/transmission techniques (serialization, local database storage via SQLite) and analyzed which methods would improve application efficiency.

UNIVERSITY OF VIRGINIA - DEPARTMENT OF CHEMICAL ENGINEERING — Charlottesville, VA
Software Development Intern — 2012 to 2012
- Designed, implemented, and tested a Python tool for Intermol, an open source software package that converts between molecular dynamics file formats of various molecular simulation packages. Incorporated various data structures (hash maps and ordered sets) and OOP techniques to improve software efficiency.

DATA SCIENCE PROJECTS

ANALYZING AMAZON'S BOOK REVIEWS USING NLP
- Employed NTLK and SpaCy to conduct NLP techniques on book review data from 2005.
- Modeled review topics with TF-IDF and NMF.
- Utilized PCA and K-Means Clustering on topics and length of reviews to see patterns of low-quality reviews.

USED CAR COST PREDICTOR
- Created multithreaded web scraping tool that extracted car information from Autotrader.com.
- Performed feature engineering with correlation.
- Utilized Linear Regression and log transformation to find optimal model with minimal error.

MULTICLASS MUSIC GENRE CLASSIFIER
- Predicted whether the primary genre of a song is Rock, Electronic, or Hip-Hop.
- Created Random Forest Model to reach desired weighted f1 score.
- Tuned parameters to minimize false positives.
- Deployed model on a Dockerized Flask application, hosted on AWS.

BILLBOARD NUMBER 1 HIP HOP SONG RECOMMENDER SYSTEM
- Generated top 10 similar songs based on user hip hop song input.
- Web scrapped song lyrics using Spotify and Genius.
- Created key words out of lyrics using NLTK package Rake.
- Created cosine similarity matrix to find which lyrics/key words are most similar to each other.

MIRIAM'S KITCHEN DATA VISUALIZATION CHALLENGE
- Utilized Seaborn and hypothesis testing to show how impactful Miriam's Kitchen is as compared to other homeless shelters in regards to reducing homelessness.