Ram Vegiraju

Data Scientist & Machine Learning Engineer

rav3sa@virginia.edu

7033577145

Ashburn, United States

ramvegiraju.github.io/Portfolio/

linkedin.com/in/ram-vegiraju-81272b162

github.com/RamVegiraju

medium.com/@rav3sa

SKILLS

PROGRAMMING: Python, R, Java, Javascript, HTML/CSS, Bash

DATA ANALYSIS: SAS, SQL, STATA, SPSS, MATLAB

DATA VISUALIZATION: Tableau, Power BI

ML: MLR, Logistic Regression, LDA, Lasso/Ridge Regression, KNN, Random Forest, SVM, K Means Clustering, PCA, Recommender Systems

Deep Learning: ANN, CNN, RNN/LSTM, Autoencoder, GANs

FRAMEWORKS/TOOLS: Tensorflow, Keras, PyTorch, Scikit-Learn, ReactJS, Flask, AWS

CERTIFICATES

AWS Cloud Practitioner (01/2020 - 01/2023)

CITI Certification (06/2019 - 06/2021)

EDUCATION

B.A Statistics (Biostatistics Concentration) & Mathematics, GPA: 3.5

University of Virginia

09/2017 - 05/2021 *Charlottesville, VA*

Courses:

- Machine Learning, Linear Algebra, Mathematical Statistics & Probability, Data Analysis with Python, Software Development Methods (Java), Data Science with R, Organic Chemistry, Statistics for Biologists
- UDEMY/Linkedln Learning: Python for Computer Vision OpenCV, NLP with Python, Python for Data Structures & Algorithms, Tensorflow 2.0: Deep Learning & Artificial Intelligence, The Complete SQL Bootcamp

WORK EXPERIENCE

R&D Engineer Intern

Amazon

06/2020 - 08/2020

Achievements/Tasks

- Objective: Created and presented a full-stack serverless application deployed on AWS CloudFormation that provides neural machine translation for client. Had weekly demos and presentations to demonstrate product to client.
- React/Javascript: Utilized AWS Amplify to create a front-end dashboard with authentication through AWS Cognito that processes client inputs and accesses a REST API created through Amazon API Gateway.
- Python/BackEnd: Developed AWS Lambda functions triggered off of API calls that query and store inputs in AWS S3 encrypted with AWS KMS keys. Translate inputs through accessing both AWS Translate and AWS Sagemaker Endpoint for custom model to return translated output to the front-end.
- NLP/ML: Preprocessing and tokenization of data through the use of NLTK and Transformers libraries. Developed, trained, and tuned a custom seq2seq model with encoder-decoder architecture on AWS SageMaker.

Data Science Intern

American Society of Clinical Oncology

06/2019 - 08/2019 *Alexandria, VA*

Achievements/Tasks

- Objective: Transferred data of patients and forms from over 110 registered sites from Syapse to a new platform in Rave EDC.
- SAS/SQL: Utilized Macro language and PROC SQL queries to develop a program that parses through Syapse patient and drug data to create forms ensuring successful migration into the new data platform.
- ML: Developed and conducted logistic regression and discriminant analysis testing to evaluate factors of re-enrollment of patients in TAPUR cohorts.
- Protocol Team: Communicated with the Data and Protocol Teams in agile meetings to understand the work-flow of patient and drug forms in the Healthcare Industry and earned a CITI certification.

PERSONAL PROJECTS

textSummarizer- NLP (06/2020)

- Objective: Created an extractive text summarizer application that also provides Named Entity Recognition.
- NLTK/Spacy: Developed a preprocessing data pipeline and worked with Displacy to return highlighted entities to user.
- Seq2Seq: Utilized a single layer bidirectional LSTM for encoder architecture along with attention mechanism for longer text inputs and a unidirectional LSTM for the decoder.
- Flask: Developed a web application to process user inputs and return entities along with summarized text.

COVID-19 Forecasting Dashboard- Time Series (03/2020)

- Objective: Developed an informative web application that forecasts and compares the growth of COVID-19 compared to other pandemics utilizing deep learning with Time-Series data.
- Javascript/HTML/CSS: Created a front-end dashboard to host ML model results and visualizations made through Plotly.
- Tensorflow/Keras: Developed a time-series model using a RNN/LSTM to forecast cases.
- SARIMA: Utilized Facebook's Prophet Library and Statsmodels package to develop a forecasting model for comparision of results.