Arun Krishna Vajjala

www.arunkv.com | Cell: 703-303-6284 | akrishn@gmu.edu

EDUCATION

George Mason University — **Fairfax, VA**

Ph.D. Student Computer Science (2021 - present)
M.S Computer Science Machine Learning (2021)
B.S. Computer Science (2020)
Courses: *Data Mining, AI, Mining Massive Datasets using Map Reduce, Graph Algorithms, Algorithms, Software Development*

EXPERIENCE

Alcon — **Fort Worth, TX**

Research & Design Machine Learning Intern (Summer 2021)

- Worked on a Multi-Disciplinary team that worked to prototype a surgical voice assistant
- I worked to create the voice assistant's wake word using a Tensorflow Convolutional Neural Network.
- Got a wake word model working with 80% accuracy with limited data.
- Worked with AWS S3 and Sagemaker for all storage and computation.

George Mason University. — **Fairfax, VA**

Undergraduate Research Assistant (Summer 2020)

- Used machine learning to cluster air pollution data to extract different air pollution contributors
- Created a regressive prediction model to predict COVID-19 mortality rates based on air pollution.
- Worked with the health departments to create meaningful data interpretation on traffic related air pollution

International Software Systems Inc. (ISSI) — **Greenbelt, MD**

Software Engineering Intern (Summer 2019)

- Created an Application for Hospitals to track patient data and give doctors ability to communicate with patients
- Developed APIs to track calls, requirements, and patient to doctor communication. Used Node Js.
- Used MongoDB for Development Environment and deployed onto AWS

PROJECTS

Machine Learning & Data Mining

- **Classifying Severity of Diabetic Retinopathy from Images** (11/2020)
  - Performed a multi-level classification on images of retinas. The classification labels were 0-5 to depict the severity of Diabetic Retinopathy (5 being most severe). Used TensorFlow and Spark Databricks deep learning libraries to classify 16gb of images on AWS EMR and EC2. (97% accuracy)
- **Predicting NYC Taxi Time using Enhanced Decision Tree Regressor** (09/2020)
  - Predicted the duration of Taxi trips in NYS. Built a decision tree regressor with linear regression models in the leaf nodes to improve regression accuracy. The Enhanced Decision Tree Regressor's RMSE was 75% more accurate than a normal Decision tree regressor and predicted more accurate Trip durations. Built using Spark and Hadoop Distributed File System and 13gb of data on AWS using EMR and EC2.
- **COVID-19 Mortality rate in correlation to Pollution** (7/2020)
  - Regression model to find the correlation between historically air polluted cities and its impact on mortality rates within the city from COVID-19. Used US cities and discovered that there is a 6% higher chance of succumbing to the coronavirus when in a city with historically higher air pollutants.
- **University Professor to student Matching** (05/2020)
  - This software was designed to use machine learning to match students to their ideal professors. Some students are better at test taking, others may be project oriented, others love to interact with their professors more. Based on a user's preferences, a professor from their university would be recommended as the best fit for a student's learning style, therefore giving them an opportunity for a successful semester.
- **Topic Modeling using Covid-19 Data** (10/2020)
  - The goal for the project is to implement topic modeling to find the topic composition and membership of the given data. used Topic modeling with LDA (Latent Dirichlet Allocation) to find 10 topics. Found the 30 most prevalent words in each topic within a collection of 26gb of research papers. Ran on AWS EMR and EC2

Component Based Software

- **Student Survey Application** (11/2020)
  - Worked with the university and created a survey for students to fill out university information. Implemented an Angular front end with a spring back end. Deployed onto EC2 using Jenkins CI/CD pipeline. Dockerized the application and hosted it on a Kubernetes cluster for scalability. Used apache Kafka to manage microservices.

SKILLS & LANGUAGES

Skills: Java, Python, Kubernetes, Jenkins, Apache Spark, Hadoop, Docker, DevOps, AWS EC2 S3, C, R, Unix OS
Interests: Photography, Golf, Basketball, Football (NFL), Formula 1, and Tennis