ANGEL UMANA, DATA SCIENTIST

Falls Church, VA | angel.d.umana@gmail.com | 703-859-2613
angelumana.com | github.com/ibidyouadu

EDUCATION

GEORGE MASON UNIVERSITY, FAIRFAX, VA
B.S. PHYSICS, B.S. MATH

SKILLS & ABILITIES

Python (pandas, numpy, matplotlib, dash, scikit-learn, TensorFlow, django, fastapi), R (tidyverse, shiny), HTML/CSS (Bootstrap), JavaScript, GitHub, git, shell, exploratory data analysis, data wrangling, data visualization, machine learning, SQL, web development, full stack development

EXPERIENCE

ASSOCIATE DATA SCIENTIST, RARE
SINCE JUNE 2021

I work as both a developer and analyst working with fisheries data.

- Wrote a tool that scans one of our data streams and flags potential outliers. Used Python and Postgres. The tool succeeded in identifying a bad outlier that inflated a key metric by over 200%.
- Development and maintenance of interactive dashboards written in R shiny and Python's dash. Developing these dashboards involves sourcing, cleaning, combining, transforming, and visualizing datasets.
- Developed website that enables users to view financial summaries from fisheries business. Written with django and html/css/js, deployed on DigitalOcean.
- Developing a tool that takes a natural language query and returns the data answering that query. It works by taking the natural language query, converting that to a SQL query using OpenAI's API, then sends that query to data.world, where our data is hosted.
- Ad-hoc data analysis for donor reports and grant proposals

STUDENT RESEARCHER, GEORGE MASON UNIVERSITY
FROM JULY 2019 TO MAY 2020

Explored the use of social media data in natural disaster assessment.

- Cleaned raw datasets, including 1 million tweets using regex and Python's NLTK library for tokenization and lemmatization.
- Exploratory data analysis w/ pandas, seaborn, scipy, statsmodels
- Machine learning w/ scikit-learn and TensorFlow
    - Decision tree regression to predict hurricane damages
    - Hierarchical, k-means clustering to identify power outage trends
    - Random forest and BERT to classify tweets
    - Open sourced code (github), documented work, presented findings, and co-authored a conference paper