

Isha Rajesh Angadi

(240) 426-9777 | ishaa@umd.edu

<https://www.linkedin.com/in/isha19/> | https://github.com/Ishaangadi/isha_resume

A summary of the major projects completed in the Masters in Data Science and Analytics at the University of Maryland, College Park is presented below.

Workout Analysis with Visual Pose Correction

(December 2022)

- The aim of the project is to create a human pose estimation and workout analyzer that assists in correcting user poses during their workouts, thus, reducing long-term injuries. The project pipeline will be inspired from *Real Time Indoor Workout Analysis Using Machine Learning & Computer Vision* (A. Nagarkoti et. al)
- The team used an amalgamation of techniques such as 2D pose estimation, optical flow, DTW (Dynamic Time Warping), and affine transformation. The project aims to build upon these techniques to provide a more computationally efficient workout analysis.
- A video recording of a live demonstration was created where the trainer skeleton is overlaid over the user skeleton to correct the user's exercise form.
- **Technologies Used:** Python, Python Libraries, Data Science and Visualization

Manhattan Real Estate Tool

(December 2022)

- The Manhattan Real Estate Tool is a Property Tech or PropTech data-driven tool created to assist real estate professionals to visualize and analyze spatial data in a way that is meaningful and useful. The tool can also be used for market analysis and target marketing.
- The project pipeline consists of using Geocode APIs to add latitude and longitude from the given attributes in the dataset, data cleaning, analysis, and creation of an R Shiny Application.
- An R shiny Application was created to explore the Manhattan Real Estate Dataset using Data Visualization, Selection, and Modeling. K-means clustering was used to generate property clusters based on the client's preferences.
- **Technologies Used:** R Programming, R shiny, Geocode API, Data Visualization and Analysis

UCI Wine Data Exploration Application

(November 2022)

- Clean, Analyze, Visualize, and create interactive components using Wine Data set.
- Multiple tabs were created within the app
 - Tab 1: an interactive table with a search bar is present to actively look through the data
 - Tab 2: an interactive histogram with a search bar is present to actively look the attributes of the data
 - Tab 3: a correlation plot to explore the correlation between continuous variable, a linear regression model for Magnesium and confidence intervals to explore both sections present
- **Technologies Used:** R Programming, R shiny

Reinvent the Wheel 2.0 Hackathon Held by Torqata

(North Carolina, November 2022)

- Challenge #1: Forecasting Scrap Tires (*How many tires of a given size_code will be collected by a given Distribution Center every day?*)

A stacked LSTM (long short-term memory networks) model architecture was used on exploring different features including — cost, tire size, width, retail price, day of week, day of month, and month.

- **Technologies Used:** Python, Python Libraries, Data Science and Visualization
- Challenge #2: Non-linear Optimization (*How to efficiently move the scrap tires downstream to the recycler?*)

Using Python SciPy function minimization with SLSQP (Sequential Least Squares Programming) — the given data was converted into graph format to obtain the weights of edges to solve the optimization problem.

- **Technologies Used:** Python, Python Libraries, Data Science and Visualization

Influencer Analysis of Congress People Twitter Data

(May 2022)

- Project's objective was to provide factual information to help people decide on their political affiliations and actions by giving them a tool to match them with the views of the appropriate congressperson. Tweets from members of Congress are vast, intricate databases that are updated frequently. The velocity of the tweets was approximately 26,000 tweets (in 10 days). Processing large amounts of unstructured, low-density Twitter data was the key challenge of the project.
- The twitter data was better understood and interpreted using the topic modeling technique. The team manually created a model of pre-assigned topic categories after scraping the data from Twitter using the Python script "snsrape". Following the topic modeling, feature extraction with BERT was used for influencer profiling of the congressperson's tweet.
- **Technologies Used:** Python, Python Libraries, Data Science and Visualization

Project Teal: Create predictions for ovarian cancer

(December 2021)

- The goal is to build a model that accurately predicts malignant or benign tumors. The Base Model was based on a research paper (*Using machine learning to predict ovarian cancer by Lu, et al.*), which analyzed data to find out if someone is at serious risk of Ovarian Cancer based on their 49 biomarkers and non-biomarkers. The team sought to improve the accuracy of the base model by tuning various hyperparameters.
- The procedure for the project consisted of handling missing values, feature selection, implementing a decision tree, and comparative analysis using logistic regression and ROMA.
- Upon utilizing stratified K-cross validation with shuffling — our team achieved 82.5% testing accuracy (with the criterias MRMR, 8 features, and entropy). When compared with the stratified K-cross validation without shuffling, our team's test accuracy with the same criteria increases by 16.7% (from 65.8% to 82.5%).
- **Technologies Used:** Python, Python Libraries, Data Analysis and Visualization