ISHA GOKHALE

(408) 930-1947 • ishagokhale@ucsb.edu

EDUCATION

University of California, Santa Barbara (UCSB)
Expected Graduation Date: December 2022
Bachelor of Science in Statistics and Data Science, Pursuing Technology Management Program Certification & Spatial Science Minor

Dean's Honors (Letters and Science) Winter 2020 and Winter 2021

Regression Analysis	Introduction to R and SQL	Multivariable Calculus
Geographic/Spatial Data Analysis	Intermediate C++	Probability and Statistics Theory
Stochastic Process	Intermediate Python	Linear Algebra
Bayesian Data Analysis	Time Series	

WORK EXPERIENCE

Data Analyst Intern, Clubfeast, Remote — January 2022 – March 2022

- Analyzed a dataset of 1000+ points to determine driver retention rare using SQL queries and data visualization through Tableau
- Collaborated with senior business analyst to determine location of high value customers and how to retain frequent user interaction with the app
- Formulated business action of offering discount/reward to high value customers to keep them engaged with the app
- Performed SQL queries on Snowflake to determine the effectiveness of repricing a time window on delivery app
- Provided daily updates on metrics to head of development and operations to determine structure of future projects
- Presented analysis to head of development and operations to determine what business actions to take

Student Research Assistant, UCSB Geography Department, Remote — September 2021-January 2022

- Collaborated alongside a team of two other students on project determining the effect of wildfires on covid mobility rates over a period of time using python
- Led the research of anomaly detection packages and implemented models, such as KNN for point wise anomaly detection on a time series data set to identify and analyze the significance of any potential outliers
- Utilized numpy, seaborn, and matplotlib libraries to clean data and select most important features
- Create data visualizations in the form of choropleth maps to gain a better visual of wildfire locations using geopandas library on shapefile data

Computer Science Instructor, Juni Learning, Remote — September 2021-Present

- Execute advanced computer science lesson plans ranging from basic data structures to applied programming over Zoom for students ages 8-18 while adapting to student needs and interests.
- Communicate with parents on a regular basis, providing updates on student progress and results from learning assessments.
- Maintain thorough records for each student, covering multiple courses and skill level progression.

PROJECTS

Time Series Forecasting Python Questions on Stack Overflow

- Obtained time series dataset of 100+ points representing number of Python questions asked on stack over a monthly time period of 10 years from 2009-2019 to forecast the growth of python as a language
- Applied time series techniques, such as differencing, log transform, identifying stationarity, analyzing AIC values, diagnostic checking the normality of residuals, and forecasting to predict the number of python questions asked on Stack over the next year

Sentiment Analysis Project — September 2020 - February 2021

- Created machine learning model to filter real news from fake news through the use of machine learning algorithms implemented through deep learning libraries such as scikit learn, pytorch, tensorflow, and torch text
- Led the development of a recurrent neural network applied to a dataset of 10,000+ text documents containing news in the form of tweets
- Implemented Tfidf vectorizer to convert text data to feature vectors and transformed our training and testing dataset accordingly
- Researched binary classifiers and applied a Naive Bayes classifier to our data, producing an accuracy score of 81%

Predicting Student's Math Scores Data Analysis — September 2021

- Obtained Kaggle dataset containing 8000+ elements on student demographics such as parents' education, income level and students' reading and writing test scores
- Self-taught and trained multiple machine learning models using regression to find and apply best model for the data
- Ridge regression model provided an accuracy score of 89%

SKILLS

Technical Skills: Python, R, SQL, Microsoft Word, Microsoft Excel, C++, Tableau, Snowflake

Languages: Fluent in English, Marathi; Conversational Proficiency in Hindi and Spanish