Mounica Patnana

Email: mounica.patnana@gmail.com

Mobile:+91-9560939398

WORK EXPERIENCE

- **Data Engineer at Varahe Analytics Pvt Ltd**: June 2022 -Present
- **ENOP-CPVI Modules**: Designed and implemented multiple Python modules to systematically compute and analyze intricate political metrics, contributing to enhanced data-driven insights and strategic decision-making
- **Survey Central Project**: Engineered a robust data handling module capable of efficiently managing and processing large datasets for cleaning and validation.
  2. Developed Python modules to generate maps for tracking the location of survey agents leveraging**Geo-Pandas**
  3. Devised and implemented a collection of scientifically rigorous methodologies for grading booths and villages within a constituency. These methodologies serve as the foundation for conducting survey sampling, ensuring a robust and accurate representation of data. By leveraging these methods, the survey process is enriched with credibility and precision, yielding valuable insights for informed analysis
  4. Developed **Linear Regression model** to observe a correlation between the vote share of a political party and caste demographic in the Constituency
  5. Developed Python module for **vanilla normalization** using **iterative proportional fitting** to estimate the winner in election utilizing survey data conducted and achieved an accuracy of **85 percent** for **Karnataka,2023 AE**.
- **Print Electronic Media**: I spearheaded the design and implementation of a comprehensive **BigQuery** database architecture, tailored for the storage and management of daily print and electronic media data. This strategic initiative ensures the availability of up-to-date information, facilitating seamless report generation and timely delivery to clients for well-informed decision-making.
- **Leader Profiles**: Created and deployed a dynamic code base that fetches data from in-house databases and generates a pre-designed pdf of Union, State Ministers leveraging **ReportLab** library on an AWS server. Designed a user-friendly button functionality that triggers the PDF generation process alienating dependency.
- **Kamkhaya Survey Web Application**: Currently working on the development of **React** web application with a strong focus on presenting a diverse range of insightful survey reports to clients.
  1. Implementing robust session management and stringent user authentication and authorization mechanisms..
  2. Designed and implemented a robust **ETL** pipeline to seamlessly integrate and deliver reports within the web application
  3. Currently working on deployment of scripts on **AWS Lambda**
- **Chi Square Automatic Interaction Detection**: Currently, I am engaged in applying CHAID analysis to extract meaningful insights from survey data. By uncovering significant variable interactions, this analytical endeavor aims to reveal critical patterns and trends within the data. These insights will empower clients to make well-informed decisions based on a deeper understanding of the survey data

Education

- **Indian Institute of Technology Delhi** — New Delhi, India
  Dual Degree in Biochemical and Biotechnology; GPA: 7.1 — *July 2015 - Feb 2021*

Skills Summary

- **Skills**: Python, SQL, Git, AWS,React
- **Databases**: Postgresql,Mongodb,BigQuery
- **Frameworks**: Scikit Learn, Pandas, Matplotlib, Seaborn, Numpy, Tensorflow, Keras, Folium, Geo Pandas, Flask, React
- **Algorithms**: Linear and Logistic Regression,Random Forests, Support Vector Machines,KMeans Clustering KNN, Bagging and Boosting, Neural Networks,CNN

M'tech Thesis

- **Surface Complexation of Selenium Nanoparticles.**: Instructor (Prof Rohan Jain) Aug 2019 - Mar 2020
  - : Synthesized and Characterized selenium nanoparticles in the lab to conduct absorption experiments and to optimize Selenium Nano Particles for maximum absorption of heavy metals from synthesized wastewater.

Projects

- **Heart Attack Prediction (Machine Learning)**: Classified patient's risk of heart attack with 90 percent Recall using classification algorithms. Tech: Python, Scikit, Numpy, Pandas, Matplotlib, Seaborn, Logistic Regression, KNeighbours, ExtraTress, Decision Trees
- **Walmart Sales Prediction(Regression Model)**: Regression model for predicting Walmart sales . faster. Tech: Python, Sckit Learn, Pandas, Seaborn, Matplotlib, Linear Regression, Lasso and Ridge Regression, Random Forests.
- **Cervical Cancer Prediction(Machine Learning)**: Classified patients with a high risk of cervical cancer based on their health profiles with an accuracy of 99.1 percent using Random Forest Classifier. Tech: Python, Pandas, Numpy, Seaborn
- **Credit Card Fraud Detection(Machine Learning)**: Classified credit card fraud defaulters using ensemble techniques Tech: Python, Ensemble Techniques
- **KMeans Clustering for Customer Segmentation(Clustering)**: Clustered customers to know target customers to focus on by optimizing sales.Tech: Python, Scikit, Pandas, Numpy, Matplotlib, Seaborn
- **Digit Recognition using C.N.N**: Trained and evaluated a large CNN model for MNIST dataset using several convolutional layers, pooling layers, and fully connected layers to achieve an accuracy of 92 percent.Tech: Python, Sckit, Pandas, Numpy, Matplotlib,Seaborn