Vinay Khatate

vinaykhatate7@gmail.com

+918291343475

Thane

mrvinaykhatate

vinaykhatate

github.com/VK2001s

Education

Masters of Science (Statistics),
University of Mumbai
2021 – 2023

Bachelors of Science (Statistics),
University of Mumbai
2018 – 2021

Certificates

- Currently Enrolled in "Google Data Analytics Professional Certificate" Coursera
- Basics of Exploratory Data Analysis Great Learning
- Fundamentals of Visualization with Tableau Coursera

Skills

Technical Skills

- **Programming**: Python
- **RDBMS**: MySQL
- **Visualization**: Matplotlib, Seaborn, Plotly
- **Statistical Analysis:** Regression Analysis, Hypothesis Testing
- **Version Control:** GitHub
- **BI Tools:** MS Excel, Tableau

Soft Skills
Communication, Decision Making, Teamwork & Collaboration, Critical Thinking

Organizations

National service scheme, *Leader*
2019 – 2020

National Service Scheme, *Volunteer*
2019 – 2021

Objective

Seeking a role in Data Science or Data Analysis, leveraging strong analytical skills and experience to drive data-driven decision-making.

Professional Experience

Data Scientist Intern, *Code Clause*
April 2023 – May 2023

- Aim - Churn Prediction in Telecom Industry using Logistic Regression
- Conducted in-depth analysis to identify key factors influencing customer churn and provided recommendations for retention strategies.
- Built machine learning models using Logistic Regression, applying cross-validation and hyperparameter tuning techniques.
- Language Used - Python

Data Analyst Intern, *Accenture Forage*
February 2022 – March 2023

- Assessed data quality and completeness to ensure accurate and reliable analysis.
- Targeted high-value customers based on customer demographics and attributes, resulting in improved marketing strategies.
- Utilized visualizations and data storytelling techniques to present actionable insights to stakeholders.
- Collaborated with cross-functional teams to drive data-driven decision-making.
- Tools Used - Excel, PowerPoint.

Projects

Spam Classifier on SMS Dataset
2023

- Objective - To build a prediction model that will accurately classify the text
- This project is a spam classifier that uses Natural Language Processing (NLP) techniques to distinguish between spam and non-spam (ham) messages.
- Predictive Model - Multinomial Naïve Bayes with 97% accuracy and 100% precision.
- Language Used - Python.

Statistical Analysis on Air Quality Dataset
2023

- Objective - To Identify Patterns of Air Pollutants in Overall India.
- Analyzed air quality data from multiple cities in India using data manipulation, visualization, and statistical analysis techniques like time series.
- Time Series Model Used - ARIMA, SARIMA
- Language Used - Python, R Studio