**NAVIN KODAM**

**9730619443 |** [navinkodam1999@gmail.com](file:///C:\Users\navin.kodam\Downloads\navinkodam1999@gmail.com) **|** [Linkedin](https://www.linkedin.com/in/navin-kodam/)

**SUMMARY**

Data Scientist with 4+ years of experience in data analysis, specializing in data handling, feature engineering, and model tuning. Demonstrated ability to meet business requirements and a strong passion for advancing expertise in predictive modeling and analytics. Proficient in R, Python, and SQL with end-to-end conceptual knowledge **of Generative AI**, Eager to apply analytical skills to drive impactful data-driven solutions.

**SKILLS**

**Programming & Scripting:** SQL, Python, R, PySpark, SparkSQL

**ETL & Data Tools:** Dataiku, Tableau, Databricks, Alteryx

**ML Frameworks:** Scikit-learn, keras, Pandas, NumPy, Seaborn, Matplotlib, NLTK

**EXPERIENCE**

**Tiger Analytics, Hyderabad 03/2022 – Present**

**Data Scientist**

* Collaborated closely on the data side for a property leasing company operating across 3 continents, analyzing large datasets to extract actionable insights, resulting in a 15% increase in market coverage for price recommendations.
* Build Regression (**Linear Regression, GBM and Logistic regression**) Model for various markets displayed important KPI’s responsible for the price to the client and the leasing officer.
* Developed and maintained onboarding models and data pipelines on the Dataiku platform, ensuring data accuracy and performing end-to-end quality control.
* Collaborated with cross-functional teams to translate business requirements into data-driven solutions.
* Worked on internal projects involving **Market mix modeling** and **Forecasting model in CPG Retail**, driving data-driven decision-making processes.
* Currently working in the healthcare domain on developing multiple predictive models focused on different KRIs (Key Risk Indicators), supporting early risk identification and better operational insights.
* Platforms used: **R, python, SQL, Excel, Dataiku, Denodo, Snowflake.**

**Almabetter 07/2021 – 02/2022**

**Data Science Trainee**

* Developed and fine-tuned a multi-class classification model using Gradient Boosting, achieving the highest accuracy of 94%, by analyzing mobile technical specifications and implementing advanced feature engineering to handle missing values, outliers, and **multicollinearity** issues.
* Developed a Regression model using Polynomial Regression, Random Forest, and XGBoost to predict apartment energy usage, with preprocessing steps like scaling, outlier treatment, and missing value imputation. Applied Box plots, heatmaps, and VIF for feature analysis, and used Lasso, Ridge, and Grid Search CV for regularization and tuning, achieving an R-squared of 0.70 and Adjusted R-squared of 0.698 on test data.

**EDUCATION**

**BE Mechanical (Mumbai University) 06/2016 – 10/2020**

**CERTIFICATIONS**

* [Full Stack Data Science (Alma better)](https://drive.google.com/file/d/13BCAf5xhezgU3CwmnSxkhhP4nD7DPnZi/view)
* [Neural Networks and Deep Learning (Coursera)](https://drive.google.com/file/d/1avPaF5LW9wGDeWH-_R_XX48Tp4NOmkuJ/view)
* [MLE Foundation (Tiger Analytics)](https://drive.google.com/file/d/1Z0LL0G90bVfObFI7Pr4ugZ9gHkctQEYS/view)
* [Dataiku Core Designer (Dataiku Academy)](https://drive.google.com/file/d/1XPoMQdcL1iYQh49MWTrNSLfKsU5YUNa8/view)