# Krushna Chaure

Dedicated and results-driven **Data Scientist with 2 years of experience**. Proficient in leveraging data-driven insights to optimize business processes, drive growth, and enhance decision-making. Demonstrated expertise in applying advanced analytics, machine learning algorithms, and statistical modeling techniques to solve complex business challenges. Strong ability to communicate technical concepts to non-technical stakeholders and collaborate effectively with cross-functional teams. Recognized for delivering innovative solutions and driving impactful results for the organization.


chaurekrushna143@gmail.com | 9604247168 | linkedin.com/in/Krushna | github.com/Krushna


## WORK EXPERIENCE

### Ambroasian Research and Development pvt ltd

**Junior Data Scientist**

*08/2021 - Present* *Pune*

**Projects**

**Credit Card Default Prediction**

- Developed a **binary classification** model using algorithms - **Logistic Regression, Decision Tree, RF, XG Boost, and SVC** to predict credit Default.
- Conducted thorough **Exploratory Data Analysis (EDA)** utilizing univariate, bivariate, and multivariate analysis techniques, providing actionable insights for **feature selection and model building**. Constructed a correlation **heatmap** to identify inter-feature relationships, assisting in feature engineering and reducing **multicollinearity**
- Implemented **Synthetic Minority Over-sampling Technique (SMOTE**) to address class imbalance, significantly improving model performance on the minority class by generating synthetic observations.
- Achieved improvement in model accuracy through effective feature selection, SMOTE oversampling.
- Conducted model evaluation using key metrics such as **accuracy**, **precision**, **recall**, **F1-score**, leading to the selection of the best-performing model.
- Collaborated with cross-functional teams to gather domain expertise and ensure alignment of data science solutions with business objectives. Documented the entire model development process, EDA findings, feature engineering methods, and model performance results for knowledge sharing and future reference.
- **Received positive feedback from stakeholders** for translating technical findings into actionable business insights. **reducing credit default risk by building a robust predictive model.**

**NYC Taxi Trip Time Prediction**

- Developed a regression model using algorithms such as **Multiple Linear Regression, Decision Tree Regressor, XGBoost and Random Forest** to predict time it will take for a taxi to complete a trip in New York City.
- Conducted comprehensive Exploratory Data Analysis (EDA) to gain insights into the dataset, enabling informed feature selection and engineering decisions.
- Designed and established an efficient **data pipeline** to preprocess and analyze the data, ensuring smooth and consistent flow throughout the project lifecycle. Applied rigorous **data wrangling** techniques to clean and preprocess raw data, enhancing model robustness and reliability.
- Effectively **encoded** categorical columns using methods like **one-hot encoding**, enabling the incorporation of categorical variables into the predictive models. Employed **advanced outlier detection** and handling techniques to mitigate the impact of extreme data points on model performance. Ensured **data quality** and **model accuracy** by addressing **multicollinearity** among features, enhancing model interpretability and reducing overfitting risks.
- Collaborated with domain experts and stakeholders to ensure the model's alignment with practical requirements and its potential value for both taxi drivers and passengers. Documented the entire **model development** process, EDA findings, preprocessing strategies, and hyperparameter tuning outcomes for future reference and knowledge sharing. Effectively communicated complex technical findings and model insights to non-technical stakeholders, aiding decision-making and solution implementation.
- **Achieved an outstanding 97% R-squared score** through effective **feature engineering** and **hyperparameter tuning** using **XGBoost** model, showcasing the model's strong predictive capabilities**.**

**Netflix Movies and TV Shows Clustering**

- Implemented **K-means clustering** on description, genre, and cast to categorize 7.7K TV shows and movies into **6 different clusters.**
- Evaluated the optimal clusters using the **silhouette score**, **elbow method** and leveraged the visualization.
- **Hierarchical clustering** implementation gives a optimal **silhouette score at 12 clusters,** same as visualized using **dendrogram**.

## TECH STACK

**Languages**
Python, SQL

**Frameworks**
Scikit-Learn, Pandas, Numpy, Matplotlib, Seaborn, NLTK

**Platforms**
Jupyter Notebook, Google Colab, MS-Office, GitHub, Excel, Tableau, Google Sheets

**Professional Skills**
Data Science, Data Analytics, Business Analytics, ML Engineering, ML Modelling.

## COURSES

Full Stack Data Science (08/2021)
*Chankyauni, Pune*

## PUBLICATIONS

*Medium Blog*
**ML Use Cases in Finance, Banking & Insurance**
*27/03/2023*

*Medium Blog*
**Facebook Friend Suggestion Algorithm**
*26/03/2023*

## ACHIEVEMENTS

Earned Python and SQL Badge
*Hacker Rank*
Completed Python and SQL Exercises
*W3school*

Team **Leader** in **excellentia** 2019 which held in college, in which organize event successfully.

Played **Cricket in State level of Maharashtra** in School.

## EDUCATION

**BE | Mechanical Engineering**
Flora Institute of Technology, SPPU Pune
*2015 - 2019*,
*59.73%*

**DME | Diploma In Mechanical**
Shreeyash Polytechnic, MSBTE Mumbai
*2012 - 2015*,
*62%*

## INTERESTS

Yoga | Music | Travelling | Reading | Meditation | Playing