Prasun Kumar

IIM CALCUTTA ·EX SENIOR DATA SCIENTIST· KAGGLE COMPETITOR · DATA SCIENCE BLOGGER

 8604650326 | [ prasunkba2020@email.iimcal.ac.in](mailto:adeshpande3@ucla.edu)

# **Education** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**IIM Calcutta**  *Calcutta*

POST GRADUATE DEGREE IN BUSINESS ANALYTICS  *2020-2022*

* Frameworks/Libraries Used: Tensorflow, Numpy, Pandas, Scikit Learn, Matplotlib
* Languages: Python, R
* Courses: Stochastic Processes and Applications, Inference, Computational Data Science, Database System

**IIT BHU (Varanasi)**  *Varanasi*

B. TECH IN CHEMICAL ENGINEERING – CGPA:8.95/10  *2014-2018*

# **Experience** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**SENIOR DATA SCIENTIST**

**Merilytics** *Hyderabad*

**Customer Clustering for Coffee Chain Client *Nov 2019 - Mar 2020***

* Grouped customers of a coffee chain client into 5 clusters based on behavioral and transactional data.
* Using clustering algorithms K-Means, DBSCAN, and GMM, helped client to reduce ad cost by 12% by targeted promotional campaigns.
* Built evaluation scripts to assess cluster quality based on internal indices such as BIC scores and Silhouette scores. Also devised a rigorous feature selection algorithm based on the Wrapper Method.

**Digital Platform for Reporting Industry Specific Trends to Clients *Apr 2020- July 2020***

* Created an automated module for collecting text data through APIs such as Spinn3r, Webhose etc and analyzing it using a Natural Language Processing module
* Built visualization dashboards showing industry specific trending topics which helped client to increase their revenue by 20% on YOY basis.

# Projects\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Home Credit Risk Modelling [Kaggle]

WRITTEN IN PYTHON *Nov 2020- Feb 2021*

* Scored 0.79 roc auc score and ranked among top 10% on the Kaggle leaderboard
* Handled highly skewed data by Oversampling using SMOTE and under sampling using Tomek Links.

**Big Mart Sales Prediction [AnalyticsVidhya Hackathon]**

WRITTEN IN PYTHON *February 2017 - April 2017*

* Ranked among the top 2% in the Analytics Vidhya competition.
* Achieved this feat by focusing on feature engineering, feature selection and hyperparametric tuning

# Achievements\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* Secured first place in Soccer Prediction Hackathon and won 35K cash prize
* Selected among 26 students from India for Charpak Scholarship
* Hospitality Head of 2-tiers team of associates & volunteers in Chem. Engg. Fest, IIT BHU
* Teaching Assistant for 3 consecutive semesters at IIT BHU