Kalyan Chatterjee

Data Scientist | Kolkata

Kalyanchatterjee84@gmail.com | 6290594393 | linkedin.com/kalyan-chatterjee | github.com/kalyanchatterjee84

Professional Experience

FlaktGroup India Private Limited

Graduate Engineer Trainee (MOG-Design)
08/2020 – 03/2022 | 1 year 7 Months

- **Analyzed** past project data and designed electrical components by doing Heat load Calculation for HVAC system in Ship increasing accuracy by **10%.**
- **Analyzed** past records forecasted estimated project cost with an accuracy of **75%.**
- Developed **procurement and project plan** decreasing lead time by **20%.**
- **Analyzed and Planned** and executed production plan and shorten the production time by **30%**

Skills

Python | MySQL | MongoDB | Web scrapping | Pandas | NumPy | Tableau| Scikit Learn | Machine Learning| Spacy | Matplotlib | Seaborn | Statistical Analysis| Azure Data Lake Zen-2 | Azure SQL | Hadoop Architecture | HTML |Flask

Language

English | Hindi | Bengali | German

Education

Kalyani Government Engineering College

Bachelor of Technology in Electrical Engineering
08/2016 - 07/2020 | Kalyani, West Bengal, India
CGPA: 8.43

Shantinagar Vidyamandir (HS)

Higher Secondary 06/2014 - 06/2016
HS (12th): 91.2%

Asansol Ramakrishna Mission High School

Secondary | 05/2004- 05|2014
Madhyamik (10th): 92.57%

Training & Certifications

Post Graduate Program on Advanced Data Science and Machine Learning

National Institute of Information Technology (NIIT) – Stack Route
03/2022 – 08/2022 | 6 Months

Descriptive and inferential statistics | Work with RDBMS | Data Analysis using Python | Preparing data for analysis | Visualization & Dashboards using Tableau | ML models - Forecasting and predictive analytics | Construct, define and validate the ML models | NLP for analysing textual data |Introduction to Deep Learning.

Programming for Everybody (Getting Started with Python)

University of Michigan (Coursera) – 2months
Python Syntax and Semantics | Use variables to store, retrieve and calculate information| Utilize core programming tools such as functions and loops

Python Data Structures

University of Michigan (Coursera) – 2 months
Data Structure | Store data as key/value pairs using Python dictionaries | Accomplish multi-step tasks like sorting or looping using tuple

Introduction to Big Data

University of California San Diego – 2 Months
Characteristics of Big Data and Dimensions of Scalability | The Hadoop Ecosystem | Hadoop Distributed File System| YARN

German Language Training - A1

Goethe Institut Indien – 3 Months

Projects

Data Science Capstone Project Using Airbnb Data for the City of Seattle: *NIIT STACKROUTE VERIFIED PROJECT (07/2022-08/2022) Team size: 4*
(keywords: Classification, Tableau, dashboard, Scikit learn, Clustering, Flask, Feature Engineering, Semantic Analysis,3-d Scatterplot, Plotly) (🔗)

- ***EDA and Data Visualization:*** *Analyzed different visualizations using **Tableau** and reached to some **interesting inferences** that increased our model building accuracy by **15%.***
- ***Host and Customer Segmentation: Clustered** all the hosts and customers each into **6** different categories using **Agglomerative Clustering** and came to a conclusion how the customer satisfaction is depending on host label.*
- ***Recommend House type:** Using **Binary Classifier** classified all the listings in two categories to narrow down the customer's searching. Also Build a recommender system with Flask API to recommend best properties as per customer requirement **reducing** customer's effort **by 20%.***

Predicting Grade of House and Properties Using Machine Learning: *NIIT STACKROUTE VERIFIED PROJECT (06/2022-07/2022)*
(keywords: Classification, KNN Classifier, Random Forest, Adaboost Classifier, Accuracy, confusion matrix, Interactive widgets) (🔗)

- *Developed a classification model using algorithms such as **KNN Classifier, Random Forest Classifier and Adaboost** to predict grade of a Property.*
- *Carried out hyperparameter tuning using **Grid Search** technique.*
- *Achieved overall **accuracy of 88.17%** with a **5-fold CV score** of 87.8%*

Predicting Sales of a Store Using Machine Learning: *A Random Forest regressor model to predict the sales of a Store for different parameters. (🔗)*

Furniture Sales Forecast Using Time Series Analysis: *Analyzed the trend and Seasonality of past records forecasted the future Sales with ARIMA model. (🔗)*

Clustering - Mall Customers Segmentation: *Analyzed Shopping pattern, spending pattern, shopping score, salary and clustered the customers into different categories using Agglomerative Clustering, DBSCAN and HDBSCAN. (🔗)*

Model Deployment:

- *Movie Review Sentiment Analyzer using **Flask** and **text blob** with an accuracy of **90%.** (🔗)*
- *Property Recommender for Airbnb using Flask, CSS and Random forest Classifier. (🔗)*

Coding Profile