

Aditya Kumar Ghosh

Certified Data Analyst

Certified Data analyst with a degree in Engineering, I have hands on experience in analyzing & interpreting data with good numerical accuracy. Skilled in Python, R & tableau, a solution-oriented mindset & effective communication skills make me stand out.

Adityaghoshsbkg@gmail.com ✉

7031030096 📞

Pune 📍

<https://www.linkedin.com/in/aditya-ghosh-955b5b161/> in

<https://github.com/adityasbg>

CERTIFICATION

Data Analyst, edWisor Oct 2019 - March 2020

- Learnt how to use scientific algorithms to extract insights from unstructured data using R and Python.
- Used predictive analytics and machine learning along with visualization tools like Tableau, predicting prices, sales & customer churn.
- Was mentored by working industry experts on technical skills along with Soft Skills training to enhance communication

PROJECTS

Pan OCR using YOLO v3 and Tesseract

- Created custom object detector to detect Name, Father Name, Pan number and Date of birth.
- Developed Algorithm to use Result Text generated by custom detector to crop image based on bounding box generated.
- Used google Tesseract to OCR cropped image to text Data.
- Created rest API to display Pan card information using Flask.

Kotak Bank Financial action tool

- Developed a Tool to extract fields from income tax return documents And populate fields to Financial action Tool.
- Used Bounding box and coordinate geometry to develop Algorithm to find label and its value and also map value year wise.
- Developed Algorithms to get tabular data from Amazon Tesseract Json file where csv file generated by Aws doesn't capture tabular information.

Santander Costumer Prediction

- Predicted customer transaction using prehistoric data. Used Machine learning Models like Logistic Regression and Random Forest XGBoost ,Decision Tree ,Naïve Bayes to obtain the outcome
- Used imbalanced data Precision-Recall curve as a metric

to evaluate the model in place of accuracy.

- Choose Model with high f1 score, low false negative rate, high ROC_AUC value. best model was XGBoost.

Bike Count Prediction

- The objective is to Predication of bike rental count on daily based on the environmental and seasonal settings.
- Used Decision tree Linear Regression and Random Forest on processed data to predict the target variable
- Lower values of MAPE and higher value of R-Squared Value indicate better fit. From the observation of all RMSE values, I concluded that the Linear Regression Model had a minimum value of MAPE

WORK EXPERIENCE

Java /Python --Engineer Persistent, Pune

Aug 2019 -

Present

- Performed Data analysis for Pune Municipal Corporation and Created Dashboard for slum Data visualisation using Plotly-Dash
- Worked on Pan ocr using Tesseract and YOLO v3 for object detection and Reduced cost incurred due to Aws Texterect.
- Worked on Kotak Bank Financial Action Tool to populate Income Tax data into template Which help them make credit Decisioning.

TECHNICAL SKILLS

Tableau R Machine Learning

Statistical Modelling Web scrapping

SKILLS THAT MAKE ME 'Stand Out'

Teamworker

Going the extra
mile

Solution Oriented mindset

EDUCATION

**Bachelor's in technology (Computer science and
technology)**

Siliguri institute of technology

2019

Higher secondary education

Tarar college 2015

ACHIEVEMENTS

2020 –Pulse Bidding wars Bronze winner

INTERESTS

Adventure

Artificial Intelligence

LANGUAGES

English



Hindi

