

Akarsh Bachu

Data Scientist

Skilled in developing end-end Data Science applications with 3.5+ years of experience. Achievements include developing a Merchant Location Cleaning application that involves the prediction of Brand and Location from messy bank transactions with an accuracy of 0.96, thereby generating revenue to our company and receiving a 5-star performance rating for the same.



bachu.akarsh@gmail.com

7729862404

Hyderabad, India

akarsh.netlify.app/

linkedin.com/in/akarshbachu

github.com/akarshbachu

WORK EXPERIENCE

Associate Data Scientist Cardlytics Inc.

05/2020 - Present

Cardlytics is into Purchase Intelligence, Based on transaction data from banks we generate offers and these offers are placed on the user's bank portal.

Description/Responsibilities

- **Predicting Brand** name from the given messy transaction data using **C-CNN** model, which gave an accuracy of **0.96** for **12k brands**.
- **Mapping** this transaction data to **store location** at various confidence levels like street, zip, and city-level with an accuracy of **0.85-0.95** based on the brand and information from the transaction using **cascading C-CNN models**.
- Developed this application with minimum guidance from the product team, application life cycle involves EDA, Pre-processing, building models, exposing this a service using FLASK with Unicorn WSGI HTTP server and deployed on Kubernetes cluster to handle multiple requests, which can **clean 24 Million transactions per day**.
- Automated train, test pipeline with a Cron job, Created dashboards in sumo logic and Datadog based on the application logs to monitor the performance of the application.

Jr. Data Scientist & Java Dev ADP India Pvt Ltd

08/2017 - 05/2020

The Product Provides Analytics of various other Products of ADP, Using Dashboards, KPI's(Key Performance Indicators) Benchmarks, and Explorers.

Description/Responsibilities

- Involved in all phases of Data Science and ML tasks like Data Preprocessing, EDA, Visualization, Feature Engineering, Developing Models, and Validation.
- **Recommending insights** based on the user interaction, Used Random Forest classifier to classify insights as liked or disliked i.e. converted recommendation problem to classification problem.
- Performed Machine Learning tasks to **predict the turnover of an employee**, achieved AUC of **0.88** using XGBoost.
- Creating KPIs (Key Performance Indicators) using Java and JPA Framework(myBatis) and developed API for exposing it.

EDUCATION

M.Sc in Data Science Liverpool John Moores University

07/2019 - 05/2021

Thesis

- Thesis on Character Level Convolutions for Text

SKILLS

Data Science

Machine Learning

NLP

Data Visualization

Modeling

Python

Java

SQL

Keras

Tensorflow

Flask

Docker

Kubernetes

Elastic Search

Probability & Statistics

HTML, CSS & Bootstrap

PERSONAL PROJECTS

StackOverflow Tag Predictor (10/2020)

- As part of the thesis, experimenting **Character Level CNN** model with **dynamic sized filters** to predict tags for the given question.

Data Augmentation with GAN (08/2020 - 08/2020)

- In the Credit Card fraud dataset there are only **0.017% of fraud transactions**, so **balanced** the data using **GAN** and performed classification using Random Forest which gave a **recall of 0.98 for Class 1**.

Stock Price Prediction (05/2019 - 05/2019)

- It's a Deep Learning Model, that predicts the stock price of (n+1)th day given n days data with features like Stock Metadata, the Volatility index, Nifty Index, Nasdaq Index, etc. using LSTM.

Donors Choose Application Screening (11/2018 - 12/2018)

- It's a Classification problem where the model classifies if the project proposed can be Approved or not, Achieved an AUC of **0.74** using **LightGBM**.

Netflix Movie Recommendation Engine (08/2018 - 08/2018)

- It's a Machine Learning Model that recommends the movies by analyzing user's previous history.

Personal Website (02/2018 - 02/2018)

- It's a Personal Portfolio created using HTML and Bootstrap, includes a Chat-bot (Virtual Representation of Me).

Heart Connoisseur (05/2016 - 06/2016)

- It's an IOT based alarming device built for detecting heart abnormalities that also notifies the pre-chosen contacts for help in the span of an SMS.

ACHIEVEMENTS

Star performer of Data Science team - Y20

Oracle Certified Java Programmer

Oracle Testing Id: OC1731881 (Java SE 8)

Hackerrank Algorithm Level 5

Received a Merit certificate from ICSI for presenting 'Heart Connoisseur'(IOT project)