

# Subham Sarkar

Data Scientist

Analytics professional with proven experience in Machine Learning, Deep Learning, NLP, Recommendation Systems and Computer Vision. Strong expertise in converting data to business achievements.

kingsubham27@gmail.com

7001363491

linkedin.com/in/subham-sarkar-4224aa147

github.com/SubhamIO

## WORK EXPERIENCE

### Data Scientist

Fractal Analytics - Bangalore

09/2021 - Present

Achievements/Tasks

- Developed **Entity Matching Algorithm** based on **Name and Transaction based matching** for **NA based trucking card company** to help standardize customers in different systems for C360.
- Developed **Time Series Forecasting Pipeline** for automatically **choosing the best model** and generate **Revenue forecasts** to understand changes to the portfolio and **raise Anomalies** where behavior is unexpected.
- BUSINESS IMPACT CREATED :**
- Identified cases of migration and created segmentation for these matches.
- 30% Improvement in entity resolution results of the MDM due to recommendations provided by Fractal.
- Algorithm also helped define cases of attrition and new customers as a by product.

### Data Scientist

Amantya Technologies - Gurgaon

10/2020 - 09/2021

AUTO ML PROJECT

- Worked on developing different features for the AutoML framework such as '**Forecasting module**', '**Compare different ML models**', '**Choose and Tune your ML model**', '**Retrain your ML model**' in the easiest possible way for the end user.
- Building **Web API's** or **REST API's** using **Flask** framework on **Python** and API testing in **Postman**.
- Deployment** or **Productionisation** of Machine Learning models on **AWS EC2** instance and , Input Data Storage : **MongoDB** and **Postgre DB** and Output Data and Model Storage : **AWS S3 Bucket**.
- Code CheckIns and Version Control : **Github, Bitbucket**.

## EDUCATION

### B-Tech in Computer Science and Engineering

Haldia Institute of Technology

08/2014 - 06/2018

CGPA : 8.73 / 10

### High School - CBSE (12th)

BSF Sr. Sec. Residential School

04/2012 - 06/2014

Percentage : 89.8%

### High School - ICSE (10th)

Good Shepherd School

02/2011 - 03/2012

Percentage : 85.14%

## SKILLS

Data Structures and Algorithms

Machine Learning

Deep Learning

NLP

Computer Vision

Statistical Modelling

Predictive Modelling

Flask

Docker

AWS Sagemaker

Python

Tensorflow

Keras

MongoDB

MySQL

Dataiku

REST API

Postman

Data Visualisation

Git

Tableau

Data Modelling

Big Data (PySpark)

## PERSONAL PROJECTS

Image Captioning using ATTENTION Mechanism

- Caption generation** is a challenging artificial intelligence problem where a textual description must be generated for a given photograph using **Encoder-Decoder models** and **Attention mechanism**.
- [Blog written on the topic.](#)
- [Github Link to the project.](#)

Netflix Movie Recommendation System

- Netflix** provided a lot of anonymous rating data, and a prediction accuracy bar that is 10% better than what Cinematch can do on the same training data set.
- Collaborative Filtering Based Recommendation.**
- [Github Link to the project.](#)

Amazon apparel Recommendation System

- To recommend **similar apparel** products/items in e-commerce.
- Content Based Recommendation.**
- [Github Link to the project.](#)

Stock Price Prediction using Deep Learning

- Predicting next n days prices using previous m days records.
- [Github Link to the project.](#)

## TUTORIAL AND VIDEOS BY ME

Approaching any NLP Problem

Different ways to tackle a NLP problem - From basic to SOTA Models . [Github Link](#)

PySpark in Action

How can anyone get started with PySpark MLLib easily. [Github Link](#)

Object Detection in Realtime Videos and Images using YOLO V3

[Youtube Link](#)

Build and Deploy a Machine Learning Flask Web API or REST API using DOCKER

[Youtube Link](#)

Face Recognition in Live Stream Video using VGG - 16 (Python,OpenCV)

[Youtube Link](#)