Shubendu Biswas

LinkedIn — Github — Kaggle — My Website

Email : biswasshubendu4@gmail.com
Mobile : +91-8968-686-3332

Professional Experience & Achievements

- **Machine Learning Engineer - Full time** April 2021 - Present
  Seasia Infotech Pvt. Ltd. *Mohali,IN*
  ROLES AND RESPONSIBILITIES
  - **Designed and developed POC of Mobile Screen Damage Assessment using CNN and DjangoAPI.** Scrapped data from various sources of damaged and non damaged mobile phones of over 3k+ images. Manually labelled data for 3 CNN based models to classify between phone or no phone, damaged or non-damaged then to classify severity of damage in high, low and moderate categories. Got validation accuracy of 80% + with VGG16 and ResNet50 models and tuned their weights. Deployed it on Linux Server using Django4 as a backend web server.
  - **Intent Sentiment Analysis** Used BERT models to develop the intent sentiment analysis. For intent classification used CLINIC150 data with 150 classes and used LSTM model with 83%+ accuracy on validation data. For sentiment classification used amazon reviews of 200K data points for each class with BERT model. Combine both models in a single pipeline and deployed on Linux server with flaskAPI.
  - **Sentiment Analytical Tool** Designed and deployed interactive data visualisation dashboard with Dash and plotly express. Deployed on Linux Server using Dash with Flask as a backend web server. Optimized the dashboard to work with large dataset with very low latency. Enhanced visualisation with various plots such as category wise sentiment analysis of positive, negative and neutral sentences.

- **Machine Learning Engineer - Internship** January 2021 - April 2021
  iNeuron Intelligence Pvt. Ltd. *Remote*
  - **Designed and developed Image Scrapper** Developed in flaskAPI using selenium, It can scrap n numbers of given images from google image.

- **WILL THEY CLAIM IT?** Mumbai
  Hackathon — Github — Certificate *Jan 2020*
  - Won first place in hackathon in machine learning on the insurance dataset conducted by Grey Atom.
  - Used popular supervised machine learning algorithms such as Random Forest, SVM, XgBoost, GBDT, Logistic Regression, Decision Tree, AdaBoost and Stacking with proper hyperparameter tuning to predict the false claim of any insurance.

Programming Skills

- **Languages**: Python, C++, SQL, NoSQL, HTML, CSS, JS
- **Technologies and Frameworks**: Keras, Tensorflow, NLP, Image Processing, OpenCV, Data Visualization, Git, Linux, Flask, Django4, GCP, AWS, Dash and Plotly, Probability/Statistics
- **Familiar with:** Hadoop, Big Data, Java, Spark, Unity

Technical Projects

- **AMAZON FASHION RECOMMENDATION ENGINE** Github
  Technologies used : Recommendation System, VGG16, CNN, Django3, Bootstrap4, Deployed in Heroku *Jan 2020*
  - 3 models used in this project. Model 1 is based on CNN, where it is extracting all the features of images and comparing it to other images. Finally comparing the euclidean distance between them.
  - Model 2 and 3 is based on BOW and TFIDF, where it is considering image's brand name, color, size and price and calculating the euclidean distance between them.
  - The whole pipeline is built in Django3 and deployed in heroku.

- **MAHINDRA FIRST CHOICE SERVICES** Github
  Technologies used : Data Analysis, Data Visualisation, Plotly, ARIMA, Time Series Analysis, Seaborn, Matplotlib *May 2020*
  - Identifying the ownership pattern of cars throughout the country. This also captures the problem wherein information regarding the spending patterns can be identified. A detailed analysis of car make and model, time and type of services etc vary with location. The servicing industry is local in nature, so this kind of analysis rendered some really interesting business insights.
  - Customer Lifetime value prediction: Revenue Forecast - Times Series Analysis with ARIMA.

- **YOUTUBE BOOKMARK PLAYLIST** Github
  Technologies used : Django3, PostgreSQL, AJAX, Bootstrap4, Deployed in Deployed in Heroku *August 2020*
  - A fully fledged website with login authentication and all CRUD operations in PostgreSQL and AJAX built in top of django3 framework
- **SENTIMENT ANALYSIS OF AMAZON FINE FOOD REVIEWS** Github
  Technologies used :SQL, NLP, Seaborn, Matplotlib, supervised and unsupervised techniques, SVD, tSNE, PCA *Jan 2020*
  - Preprocessing text data using NLP techniques, Used techniques such as word2vec to convert text data into vectors.
  - Used popular supervised machine learning algorithms such as KNN, Logistic Regression, SVM, Naive Bayes, Decision Tree, Random Forest, XgBoost.
  - Hyperparameter tuning of each and every algorithm using AUC ROC metric.
  - Applying unsupervised techniques such as SVD, PCA, T-SNE and K-means clustering.

Education

- **Full Stack Data Science Engineering** Mumbai,India
  Grey Atom School of Data Science *July 2019 - November 2020*
- **Bachelor Of Computer Applications (Hons.)** Jalandhar, India
  Lovely Professional University *July 2016 - June 2019*