Shubendu Biswas

LinkedIn — Github — Kaggle — My Website

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

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*

Programming Skills

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

Awards & Experience

- **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.
- **SENTIMENT ANALYSIS ON TWITTER** — Mumbai
  Hackathon — Github — Certificate — *March 2020*
  - Reached second place in hackathon on Twitter dataset among 12 challenging teams of professional.
  - Used NLP techniques to preprocess the data, also used word embeddings like word2vec and glove vectors to convert words into vectors.

Technical Projects

- **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.
- **NETFLIX MOVIE RECOMMENDATION SYSTEM** — Github
  Technologies used : Recommendation Sytems, Data Analysis, XGBoost, Surprise library, Matrix Factorisation, SVD — *August 2020*
  - 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.
  - Objectives: Predict the rating that a user would give to a movie that he has not yet rated and Minimize the difference between predicted and actual rating (RMSE and MAPE)
- **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.
- **DEEP SURVEILLANCE SYSTEM** — working
  Technologies used : Deep Learning, Computer Vision, Data Augmentation, ConvNets, VGG — *October 2020 - Present*
  - Building emotion detector using little VGG architecture with six classes.
  - Using Computer Vision detecting age and gender.
  - Integrating all these things into single Flask web app.