Koushik Ahmed Kushal

koushikkushal95@gmail.com | www.linkedin.com/in/koushikahamed1234/ | 01773481741 |Dhaka 1212

EDUCATION AND HONORS

American International University Bangladesh

- BSc in Computer Science (CSSE) *2020-May*
- Relevant Coursework: Machine Learning, Data Science Statistics, Big Data Analytics, Probability & Discrete Mathematics

The Duke of Edinburgh's International Award

- Bronze for DofE programme (**Bronze**)

PROFESSIONAL EXPERIENCE

Infosys Limited, Bengaluru, India
Data Science & Machine Learning Analyst, *June 2019 - Sept 2019*

- Utilized Python to implement unsupervised machine learning techniques for server and security logs on unstructured data, which clustered system activities , patterns, and operations.
- Processed 3.2Million samples of event log (security, audit, system) for analysis, that prevents risks through tracking (hacks, Unwanted login) failure audits and built a prolific system trend for the network.

Amal Foundation
Data Analyst & Backend Python Developer, *June 2020 - Nov 2021*

Proactively liaised with the design team and senior developer to enhance the efficiency and performance of the client site. Develop tools and keep superintends web trends on server and database. Analysis data on top of RFM criteria which is based on spendings and fund stat.

SKILLS

- **Programming Languages:** Python, Java, C, C++, C#, JavaScript, HTML, CSS, SQL,BigQuery
- **Big Data & Machine Learning:** Spark, Power BI, Python (eg. Tensorflow,scikit-learn, numpy, pandas, matplotlib,seaborn)
- **Data Science & Miscellaneous Technologies:** A/B testing, ETL, Data science pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, Time series, Hypothesis testing, OOP, APIs, Excel

RESEARCH

Underwater Object Detection and Localization (Computer Vision) *2018-Aug*

- It is challenging to produce an efficient object detection model with high precision and low processing time, there were many different networks have experimented in the traditional algorithm development process like HOG, CNN, SVM, R-CNN.
- Proposed YOLO-Lite model with Spatial Pyramid Pooling (SPP) vision-based execution through CPU achieved The precision of 75.80% with 9 fps.
- Google open images v6+ with 128X128 dim on top of YOLO v3-tiny architecture with 93(fps) with 80.69% precision

Textual Noise Reduction and Sentiment Analysis **(NLP)** *2019-May*

- This dataset consists of a few million Amazon customer reviews (input text) and star ratings (output labels) for learning how to train fastText for sentiment analysis.
- Tools like a bag of words (BOW), TF-IDF, NLTK, RegEx used and built some custom functions to Preprocess data.
- Tradition approaches with LightGBM that leads as a generalized classifier among in the pipeline -Xgboost, Gradient boosting, Random forest, Decision tree, and Naive Bayes with different combinations of hyperparameters.
- Artificial neural net that supports LSTM on top of BERT as a pretrained network to enhance predictive accuracy(93.06%).

PROJECTS & LEADERSHIP

➢ ***Sales Insight and End-to-End Predictive Analysis on Zomato***

- Here i get insight into why Zomato sales fall in different cities and distribution of ratings, price range, the average cost.
- EDA with cuisines has higher prices, food ratings which vary across services(delivery/reservation).
- Segregates customers Using order rate, frequency, recency, and get the optimal solution for raising the metric.

➢ ***Cohort Analysis and Customer Segmentation***

- Cluster customers on the basis of activeness and spendings the average visits.
- Exploratory analysis with time, frequency, and monitoring positive/negative activities.
- Prioritize customers on RFM criteria which lies in recency,frequncy, and monetary value of each customer.

➢ ***Siamese Network (geo image similarities)***

- This tool will possess a search capability whereby an analyst provides an image of interest and is presented with other images which are similar to it.
- This machine learning model should return the top K images that are most similar to this image based on a single image input.

➢ ***Recommender System (music)***

- Recommendation Systems are a type of information filtering systems as they improve the quality of search results and provides items that are more relevant to the search item or are realted to the search history of the user.
- We have come across few traditional approaches(content-based, collaborative and popularity.) before deep neural net for user pattern and interest detection based recommendation system.The precision is negligible in comparison(2.4%)

➢ ***Employee performance analysis with visualization(Power BI & Pandas on Jupyter)***

- Uncover the factors that lead to employee attrition and explore important questions such as ‘show me a breakdown of distance from home by job role and attrition’ or ‘compare average monthly income by education and attrition’.
- This is a fictional data set created by IBM data scientists (Unstructured data-3.2million samples).

➢ ***Radiologist level pneumonia detection and classification - (Coursera deeplearning certification)***

- We develop an algorithm that can detect pneumonia from chest X-rays at a level exceeding practicing radiologists. This algorithm, is a 121-layer convolutional neural network trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset, containing over 100,000 frontal-view X-ray images with 14 diseases.
- Four practicing academic radiologists annotate a test set, on which we compare the performance of this model to that of radiologists.

➢ ***Cassava Leaf Disease Analysis and Classification (handle class imbalance -reduce bias)***

- With the help of data science,we made it possible to identify common diseases so they can be treated.
- We introduced with a dataset of 21,367 labeled images collected during a regular survey in Uganda through farmers and annotated by experts at the National Crops Resources Research Institute (NaCRRI).
- Classify each cassava image into four disease categories or a fifth category indicating a healthy leaf.Deep cnn with 256X256 images with pretrained mobilenet gives highest accuracy in detecting diseases realtime.