ANKIT BAGUSETTY

(312) 918-9335 • bvsankitds@gmail.com • linkedin.com/in/ankitbvs • github.com/Ankit0207 • Portfolio/AnkitBvs • Chicago, IL 60607

SUMMARY

Highly-determined Data Science graduate having 3+ years of experience using ML, text mining, and deep learning algorithms to solve challenging problems. Received "**Exceeds Expectations Award**" for developing Automated analytical tool that resolved issues 40% faster and reduced the incident tickets by 33%. Strong accomplishment of building unique data science applications, now aspiring to bring actionable solutions to real-time industry problems.

EDUCATION

University of Illinois at Chicago (UIC), Chicago, IL **[Aug 2019 – Dec 2020]**
Master of Science in Business Analytics – Specialization in Data Science

Chaitanya Engineering College, India **[Aug 2012 – May 2016]**
Bachelors in Electronics and Communication Engineering

TECHNICAL SKILLS

Computer Programming : R, R Shiny, Python (Tkinter, PySpark, PyTorch, TensorFlow, Keras), SQL, Java, HTML, CSS, JavaScript, SAP ABAP, Git
Software Tools : Big Data(Hadoop, Spark, Map Reduce, Hive, HBase), MS Excel, IBM TWS, SAP, MySQL, Alteryx, R, Stata, Scikit learn
Analytics/Visualization : Statistics, Machine Learning, Advanced Text Mining (NLP), Deep Learning, Tableau, Power BI, HealthCare and Marketing Analytics
ML Models Deployment : Amazon Web Services (AWS), Kubeflow, Microsoft Azure ML, Amazon SageMaker, R Shiny, Tkinter, TabPy server

PROFESSIONAL EXPERIENCE

DATA SCIENCE INTERN, *AutomizeApps, Germany* **[Apr 2020 – Aug 2020]**

Sentiment Analysis of Product Reviews *[Pandas | NLTK | Keras | TensorFlow]*

- Implemented a rule-based classification and a Deep Learning LSTM model for sentiment analysis of 4 different languages i.e. English, Spanish, German, and French. The rule-based model outperformed the LSTM model achieving an accuracy of 92%.

Twitter Topic Modeling Application *[Python | CorEx | Microsoft Azure ML Deployment]*

- Developed an application that applies topic modeling on the tweets of a subject using unsupervised LDA, and semi-supervised CorEx. Obtained 10 different latent topics and their sentiment over 14 weeks. Deployed the model in production using Azure ML deployment.

ML RESEARCH ASSISTANT, *University of Illinois at Chicago, USA* **[Apr 2020 – Aug 2020]**

Prediction of Diabetes from Lumiata Claims dataset *[Alteryx | Python | Tableau]*

- Streamlined ETL operations on a semi-structured Claims dataset and analyzed the important variables i.e. LOINC, ICD 10 and, CPT codes. Balanced the data using stratified sampling and predicted the probability of diabetes for patients using SVC model achieving a recall of 89%.

Identification of Leader words from any local meaning *[Flare LSTM | Glove Embedding | Keras]*

- Optimized Python script to transform unstructured data from Merriam Webster dictionary to structured format a strong rule-based association. Modeled a Universal Leader words identifier using Bi-directional LSTM which takes Glove Embedding and POS tags attaining a recall of 96%.

BUSINESS ANALYST, *CGI, India* **[Jun 2016 – Jun 2019]**

- Performed statistical analysis of FI and SD data and suggested tangible solutions to stakeholders by visualizing the KPI metrics in Tableau. *[Tableau]*
- Streamlined workflows and built applications in SAP ABAP business modules deploying the applications using the IBM TWS tool. *[SAP ABAP | IBM TWS]*
- Optimized complex SQL Queries for faster data processing which improved the speed and efficiency of applications by 28% *[SQL | Oracle database]*
- **Key achievement**: Received "**Pat on the Back Award**" for spearheading the whole team and managing adhoc requests during challenging time.

REAL-TIME APPLICATIONS DEVELOPED - DATA SCIENCE

All in One Data Cleaning App *[R Shiny | Statistical Tests]*

- Developed an Automated App which performs 15 most important data cleaning features and reduces the overall data cleaning time by nearly 30-40%.

English News to Telugu Translator *[Keras | TensorFlow | Flask | AWS Orchestration]*

- Designed a language translator application using Flask and LSTM. Orchestrated the application on the AWS server by placing it inside a containerized docker and pushed its image to the AWS Linux EC2 instance making it globally accessible on AWS server.

Chronic Kidney Disease Prediction *[R Programming | R Shiny | ggplot2]*

- Created an interactive screening tool that predicts the risk of having a chronic Kidney disease using the Logistic regression model attaining 97% recall.

REGEX Harry Potter Themed Machine *[Pandas | NLTK |Tkinter]*

- Designed an interesting REGEX application using python Tkinter consisting of all the basic to most advanced REGEX Operations in one place.

KEY DATA SCIENCE PROJECTS

Depression Prediction on Instagram Users *[Pandas | Numpy | REGEX | Matplotlib]*

- Identified individuals at risk of depression by extracting text features, HSV and Face-count values from Instagram posts. Performed Un-supervised LDA and COREX and Modeled SVC to predict the probability of depression improving accuracy from 70% to 94.5% at 92% precision.

Automated Image Captioning System *[PyTorch | RESNET-152 | Amazon SageMaker]*

- Built an Image Captioning system using Encoder CNN and Decoder LSTM with an attention mechanism to generate relevant captions for any input image. Analyzed the results using BLEU 3, 4 metrics, and deployed the model in real-time using Amazon SageMaker.

Amazon Product Recommendation System *[Apache Spark | ALS | KNN | Spark RDD]*

- Implemented Spark RRD on 5,074,160 reviews and built a product recommendation system using ALS Collaborative Filtering obtaining 0.91 RMSE.

Walmart Sales Forecasting *[Python | ARIMA | Time Series | Tableau | TabPy]*

- Predicted the department wise sales for 45 Walmart stores modeling the effects of markdowns on holidays using the ARIMA and Holt-Winters model.

SkyWest Competitor Visualization *[Python |Excel |Tableau]*

- Created interactive dashboards to visually represent KPIs for analysis of operations at SkyWest Airlines. Captured the key metrics that affect the performance of SkyWest airlines compared to the rest of the US airlines.