Mudita Garg

(332) 201-4257 * muditagarg12@gmail.com * github.com/muditagarg * linkedin.com/mudita-garg * tableau.com/mudita.garg

EDUCATION

University Of North Texas, **Denton, Tx** *May 2022*
Master of Science, Business Analytics *GPA: 4.0*
Awards & Honors: Received Tuition Benefit Program (TBP) award of $5,740 for 2021-2022
Extracurricular Activities**:** Director of External Relations at Graduate Student Council, India Student Association
Guru Gobind Singh Indraprastha University, **New Delhi, India** *May 2020*
Bachelor in Technology, Information Technology *GPA: 3.8*

PROFESSIONAL EXPERIENCE

Research and Teaching Assistant, Advanced Data Analytics Dept., University of North Texas *January 2021-Present*
Constructed and evaluated a comprehensiveness Covid-19 knowledge graph
- Mentor students to create visually compelling dashboards with actionable insights that highlight trends in the data.
- Collaborated with team to extract entity and devise Q/A system based on Covid-19 research papers using SpaCy library.
- Extracted 20,000 triples and their relations from unstructured data using Open Information Extraction.

Machine Learning Intern, University of Calabria, Italy *October 2019-November 2019*
- Implemented data mining and pre-processing techniques in Healthcare and Informatics data.
- Summarized analysis on different machine learning models and compared the efficiency for disease prediction.
- Analyzed symptoms and diseases using patient data profiling and disease parameters by different algorithms in Python.

Data Analyst, University of Trás-os-Montes and Alto Douro, Portugal *May 2019-July 2019*
- Collected data to analyze and implement wireframing and prototyping for palliative care patients using Adobe XD.
- Improved usability by 40% by implementing user testing on the application for palliative care patients.
- Identified business and customer segments in the market for collaboration between universities in India and UTAD, Portugal.

PROJECTS

Chatbot using Google Cloud Platform (GCP): github.com/ChatBot *October 2021*
- Built a chatbot using Dialogflow in GCP for a ticket reservation company.
- Designed conversational flows for virtual agent.
- Added a phone gateway to your virtual agent.

Twitter Sentiment Analysis on Drug War: github.com/Drug-War *August 2021 - September 2021*
- Web scraping on tweets using Twitter API and analyzing 29 lakhs tweets for identifying the trends in data.
- Extracted keywords using YAKE and KeyBert and performed topic modelling using (Distil)BERT.
- Performed evolution analysis with year segmentation to analyze the timeline of the tweets about the drug war using BERT.

Supply chain management trend analysis: github.com/Trend-Analysis *June 2021*
- Gathered dataset by web scrapping trending keywords from websites using BeautifulSoup and NLTK in Python.
- Applied TF-IDF, bigram, and trigram to measure the frequency of words and performed EDA to remove unnecessary keywords.
- Analyzed the evolution of supply chain automation with a perspective of companies and researchers.

120 years of Olympics: tableau.com/120YearsofOlympics *May 2021*
- Led a team of 5 and designed an interactive coherent dashboard using filters, rank, table calculations, KPIs and parameters on the Olympics data.
- Applied data blending by using joins on related data from different data sources to create advanced visualizations.
- Created dual-axis advanced charts and pages in Tableau for better insights.

Electricity demand forecasting of Victoria, Australia: github.com/Electricity-Demand *March 2021*
- Performed time series forecasting using PROC ARIMA with an accuracy of 85% to analyze the impact of COVID-19.
- Applied Ljung-Box test and other tests to check the adequacy of the model at a 95% Confidence Interval.

Company bankruptcy prediction: github.com/Company-Bankruptcy *January 2021*
- Tested hypothesis on the model using logistic regression and selected nine variables out of 96 using stepwise regression.
- Increased the accuracy from 40% to 70% by using sampling to balance predictor variable, removing outliers, and normalizing the distribution curve.

TECHNICAL SKILLS

- **Data Visualization:** Tableau, Power BI, Python (Plotly, Dash, Flask)
- **Machine Learning:** Decision Trees, CNN, Supervised, Unsupervised and Deep Learning
- **Data Mining:** Classification, Association, Clustering, Linear, Logistic, Regression Analysis, KNN, K-means
- **NLP:** Keyword Extraction, Topic Modelling, Genism, Word2Vec, Bert, Yake, Text Mining
- **Tools:** SAS Enterprise Miner, SAS Studio, Google Colab, Jupyter Notebooks, MS Office (Word, Excel, PowerPoint)
- **Programming Languages:** Python (Pandas, NumPy, Scikit-learn, Keras), R, Java, C, C++, CSS, SAS, OOPS
- Statistical Modelling, Predictive Analytics, Time Series, ARIMA, Data Warehousing, Data Modelling, Dimensional Modelling, ETL, QlikView, Google Cloud Platforms (GCP), Astra, NOSQL database, DataWarehousing