

Anuj Saboo

778-862-8169 | saboo.anuj17@gmail.com | linkedin.com/anujsaboo | anuj-saboo.github.io
Vancouver, BC (Open to relocate)

SUMMARY

Meticulous professional with 4+ years of industry experience winning 8 performance awards designing analytical solutions.

Industries: Banking, Retail, Apparel, Edtech, Transportation

TECHNICAL SKILLS

Languages: Python, SQL, R, C++, HTML, CSS

Data Science Frameworks: Pandas, Scikit-learn, Spark, Airflow, Hadoop, Keras, Spacy, NLTK, Plotly, Power BI

Miscellaneous: Azure, AWS, Docker, MS SQL, SSIS, MySQL, PostgreSQL

Horizontals: Data Engineering, Predictive & Prescriptive Analytics, Classification, NLP, Recommendation Systems

EXPERIENCE

Data Scientist

May 2020 – Present

Transport Canada

Ottawa, Canada

- Developed ETL pipelines, performed cleaning, feature engineering and statistical analysis to segment data and detect patterns by applying dimensionality reduction, training clustering and anomaly detection algorithms
- Leveraged visualization and communication skills to explain risk indicators, analyze performance and cluster properties to answer challenging business problems to executives and stakeholders through interactive dashboards
- Led and designed an end-to-end NLP workflow to process transport incident emails, identify language and recommend similar incidents to detect emerging risks, patterns and inform predictions
- Developed and deployed a portal on Azure cloud to perform Social Media Analytics in real time that helps identify incidents, detect events and quantify regulatory benefits using Sentiment Analysis and analyze topics using Graph Networks improving response time by 90%

Senior Developer

Mar 2017 – July 2017

Mu Sigma

Bengaluru, India

- Led a team of 5 developing ETL workloads and designing dashboards to report KPI's for Microsoft Bing
- Worked as a single point of contact for infrastructure architecture design and performance tuning of MS BI services

Database Analyst/Developer

Aug 2014 – Mar 2017

Societe Generale

Bengaluru, India

- Driven diverse and complex analytical solutions using large scale data infrastructure using SQL and NoSQL data sources being the Subject Matter Expert leading the project transition from customer site at Paris
- Engineered a pipeline for data collection, preprocessing and regression modelling to monitor, measure and report infrastructure resource allocation through analytical dashboards
- Developed a text processing solution to process database and server logs leading to quantification of root causes, analysis of performance bottlenecks and automated service handling
- Programmed an in-house end to end data orchestration solution using SQL for data replication, reporting and high availability for 180 databases resulting in savings of \$200K per year for the organization

EDUCATION

Simon Fraser University - 4.14/4.33

Burnaby, BC

Master of Science in Computer Science, Big Data Specialization

Sep 2019 – Apr 2021

Manipal Institute of Technology - 8.09/10

Manipal, India

Bachelor of Engineering in Information Technology

July 2010 – May 2014

CERTIFICATIONS

- Natural Language Processing Specialization | deeplearning.ai | December 2020
- Advanced SQL Queries | Learning Tree | December 2020

PROJECTS

Flight Delay Prediction | R, Classification Algorithms, Paramter Hypertuning

Oct 2019 – Dec 2019

- Performed data cleaning, exploratory data analysis and developed a classification model to predict flight delays
- Evaluated and tuned machine learning models using algorithms such as Logistic Regression, Random Forest, SVM and XGBoost reporting causal factors for delays

Find Me Love | Recommendation Systems, Correlation, Clustering

Feb 2020 – Apr 2020

- Built a Recommendation System on Libemseti dataset to suggest dating profiles using cold and hot start strategies
- Applied Content based and Collaborative filtering methodologies such as Correlation, Nearest Neighbor etc.