Raj Tank | Toronto, ON

[LinkedIn](https://www.linkedin.com/in/raj-tank/)| +1-647-780-9445 | [rajtank2207@gmail.com](mailto:rajtank2207@gmail.com)

# SUMMARY OF QUALIFICATIONS

# I have over 2 years of experience in data analytics and NLP, specializing in designing, training, and deploying models for sentiment analysis, topic modeling, and chatbot development. Proficient in Python, SQL, TensorFlow, and Hugging Face, I work with LLMs and Generative AI tools to deliver innovative solutions. With expertise in cloud platforms like AWS and relational databases, I excel at turning business needs into actionable insights and driving impactful results.

# WORK EXPERIENCE

AI Writing Evaluator III |[**Outlier AI**](https://outlier.ai/) **, Canada**

Jan 2024 - Present

* Conducted a code review that enhanced the AI model’s training process, improving accuracy by 12%.
* Evaluated metrics in AI responses, leading to a 15% improvement in model precision.
* Identified and implemented key features, boosting model recall by 10% and significantly reducing false negatives.
* Enhanced model training effectiveness by integrating additional code, achieving a 20% reduction in training time.

Data Scientist Intern | [**Bit Space Development Ltd.**](https://www.bitspacedevelopment.com/)**, Canada** March 2023 – July 2023

* Improved a construction site safety model’s accuracy to 83% by implementing Yolov5 Deep learning model among 5 classes.
* Led a team of 5 people, enhancing model precision by 3% through image augmentation techniques.
* Implemented non-max suppression, reducing class overlap across 700 images to enhance model performance.
* Deployed construction site safety and PPE (Personal Protective Model) to NVDIA Jetson in Roboflow.

Data Scientist Intern | [**Danfoss**](https://www.danfoss.com/en-ca/)**, Canada**

September 2022 – December 2023

* Enhanced sales forecasting models, reducing the RMSE to 3.2%, improving the accuracy of sales predictions to 89%.
* Optimized feature selection, eliminating 9 out of 30 low-impact features, resulting in a 4% increase in model performance.
* Collaborated with team, and identified an 11% discrepancy between predicted and actual sales and implemented XGboost to reduce the gap by 3%.
* Conducted data cleaning by removing dummy and null values, resulting in the best model performance.

Data Scientist Intern | [**Fingertips Data Intelligence Solutions**](https://fingertips.co.in/), **India**

July 2021 – December 2021

* Optimized customer retention model on scrap data, improving silhouette score to 81% using K-Means Clustering.
* Executed the elbow method with the team to establish 3 optimal clusters (Low, Medium, High Spender.)
* Designed an intuitive Tableau dashboard, enhancing the understanding of the model's accuracy to non-technical clients.

# EDUCATION

**Masters In Analytics (Major: Applied Machine Intelligence)**

Northeastern University, Canada

**Bachelors in Mechanical**

L. D. College of Engineering, India

# RELEVANT ACADEMIC PROJECTS

**Customer Churn Prediction on Telecommunication Industry** March 2023

* Used XGBoost to balance the 'Churn' target variable, which helped reduce overfitting by 9%, making the model more reliable.
* Tuned hyperparameters for both XGBoost and Random Forest, boosting accuracy to 83% and improving the overall performance.
* Speeded up the model by removing 5 features that weren’t adding value, which also made the predictions more accurate.

## British Airways Customer Review Analysis July 2025

* Scraped 10,000 reviews using BeautifulSoup from the British Airways website and cleaned the text.
* Created a REST API to process the text data and deployed it on AWS Lambda, enabling efficient and scalable access to the sentiment analysis pipeline.
* Benchmarked VADER and RoBERTa models to label the data into three categories (Positive, Neutral, Negative).
* Attained an 87% accuracy rate by fine-tuning the pre-trained BERT-base-uncased model from Hugging Face and deploying the model on AWS SageMaker for production use.

# KEY SKILLS

* Programming Languages and Database: Python | R | SQL | MongoDB
* Reporting Tools: Power BI | Tableau | Looker | Matplotlib | ggplot2 | Seaborn
* Machine Learning Tools: Scikit-Learn | TensorFlow | Keras | PySpark
* Data Tools: NumPy | Pandas | dplyr | SciPy | Microsoft Office | Power Query
* Natural Language Processing: NLTK | Spacy | Hugging Face
* Soft Skills: KPI Measurement | Communication | Quantitative Modeling | Requirements Gathering

# COMMUNITY VOLUNTEERING

Community Contributor | [**Civic Tech Toronto**](https://civictech.ca/), **Canada** Nov 2024 - Present

* Provided AV support to ensure smooth events without technical interruptions.
* Participated in community-driven tech projects and contributed ideas to improve outcomes.
* Involved in creation of POLIS (AI powered sentiment gathering platform. More organic than surveys and less effort than focus groups.), and worked on opinion moderation using BERT.
* Helped attendees with technical issues and worked closely with the team to introduce innovative solutions.

Tech Support Specialist | [**Google Developer Group (GDG) Cloud Toronto**](https://developers.google.com/community/devfest)**, Canada** Nov 2024

* Welcomed attendees and managed the registration process for a seamless experience.
* Assisted speakers and workshop facilitators with technical setups, ensuring everything ran on schedule.
* Worked with the organizing team to address challenges and maintain a professional event environment.