

Highlights of Qualifications

- Diversified data scientist with an M.Sc. in Computer Science (AI Specialization) and over 5 years of cross-sector experience spanning fields like banking, energy, government, and healthcare.
- Advanced proficiency in Natural Language Processing; adept at fine-tuning large and regular language models.
- Secured \$100 million (CAD) funding for AI initiatives at IRCC and designed pivotal systems at DBS Bank, including early warning mechanisms safeguarding over \$700 million (SGD).
- Expert in creating and maintaining ETL data pipelines; deep familiarity with tools like Hadoop, Hive, and Spark on cloud.
- Led significant optimization projects, realizing operational cost savings of over \$575,000 (USD) per turbine site.
- Published researcher in top conferences with accolades like the Best Employee Award at DBS Bank; holds certifications from renowned companies like Google, Microsoft, and IBM in data science and analytics.

Skills

Programming & Development: Python, R, Java, Shell

NLP: NLTK, spaCy, OpenAI API, BERT, GPT-4, Llama 2

Data Engineering: NumPy, Pandas, SQL, Hadoop, Spark

Data Visualization: Tableau, Qlik, Power BI, Excel

ML & AI: Scikit-Learn, TensorFlow, PyTorch, Transformers

Cloud & DevOps: AWS, Azure, GCP, Git, Docker, JIRA

Professional Experience

Machine Learning Associate, Vector Institute, Toronto, ON

Sept 2023 – Dec 2023

- Employed NLP techniques and fine-tune language models like BioBERT and ClinicalBERT to extract key insights from patient texts, such as chief complaints and standardized medical codes (SNOMED-CT and ICD-10).
- Utilized LLMs, including GPT-4 and Llama 2, to generate human-like synthetic training data.
- Developed an in-house LLM infrastructure on AWS, specifically fine-tuned on healthcare data, to tackle advanced medical reasoning tasks like disease detection, classification, and syndromic surveillance processes.

Research Associate in Natural Language Processing, IRCC & University of Windsor, Windsor, ON

Jan 2022 - Present

- Developed an AI recommendation system using a fine-tuned sBERT language model on Canadian immigration data aimed at aiding settlement agencies. This Salesforce based system provides personalized recommendations to caseworkers based on client questionnaires, facilitating support for newcomers to Canada.
- Secured significant funding of \$100 million (CAD) over the next 3 years from the Immigration, Refugees and Citizenship Canada (IRCC) for a pivotal project among 78 initiatives, exploring the potential of AI in crafting effective settlement strategies for newcomers.

Performance Analysis and Engineering Co-op, Liberty Utilities, Oakville, ON

May 2022 - Dec 2022

- Implemented analytics solutions with the performance and engineering teams, utilizing Python, advanced Excel, pivot tables, and Tableau to optimize wind turbine performance.
- Led aerodynamic studies on blade angle adjustments across 89 turbine sites, harnessing fluid dynamics simulations and realizing operational cost savings of over \$575,000 (USD) per site.
- Utilized anomaly detection techniques, like One-Class SVM and Isolation Forest, on diverse weather datasets extracted from various APIs, creating predictive models to pinpoint ice accumulation and abnormal turbine operations. Achieved a notable reduction in downtime, averting around 150 hours per turbine.
- Automated data mining for wind speed insights from multiple sources, employing APIs and web scraping with BeautifulSoup and Selenium, enhancing the understanding of wind patterns for informed wind turbine site selection.

Data Science Developer, DBS Bank, Hyderabad, India

Jul 2018 - Aug 2021

- Fine-tuned an NLP-driven chatbot using SVMs, spaCy, and NLTK, achieving 94% accuracy for Cantonese-speaking Hong Kong corporate clients, addressing 80% of inquiries and substantially reducing support calls.
- Designed an early warning system that detected over 95% of non-performing SME loans ahead of 3 months of credit stress, safeguarding more than \$700 million (SGD).
- Designed and maintained ETL data workflows in Hadoop, Hive, and Spark, employing Python, R, Java, SQL, and Jupyter on Cloudera.
- Implemented agile methodologies and ML-Ops, partnering with diverse teams, and utilizing Git, Subversion, JIRA, and Confluence for streamlined project management.
- Architected 8 Qlik dashboards aiding SME banking in Asia, advancing digitization, customer network analysis, loan oversight, and balance sheet strategies.

Education

Master of Science in Computer Science - Artificial Intelligence Stream - Co-op (M.Sc.) University of Windsor, Windsor, ON	Sept 2021 - Aug 2023
Bachelor of Technology in Computer Science & Engineering (B.Tech.) Jawaharlal Nehru Technological University, Hyderabad, India	Jun 2014 - May 2018

Projects and Professional Development

Quantile Regression Analysis of House Price Variables , University of Windsor, Windsor, ON 	Jan 2023 - Apr 2023
<ul style="list-style-type: none">Analyzed the relationship between different housing characteristics and their sale prices at various points of the sale price distribution using quantile regression techniques.Demonstrated that different housing characteristics exhibit varying significance levels at different quantiles of the sale price distribution.	
Novel Efficient Fake News Detection Method , University of Windsor, Windsor, ON 	Jan 2022 - Apr 2022
<ul style="list-style-type: none">Developed a novel method to detect fake news with an accuracy of over 92% on a dataset by leveraging the Singular Valued Decomposition (SVD) feature reduction method and the Long Short-Term Memory (LSTM) neural architecture, achieving a training time of fewer than 2 minutes.Conducted comparative studies with state-of-the-art fake news detection models on the same dataset.	
Customer Credit Default Prediction , University of Windsor, Windsor, ON 	Sept 2021 - Dec 2021
<ul style="list-style-type: none">Built a deep learning model with TensorFlow to classify customers based on the likelihood of credit card default, achieving an accuracy of 82.6% on randomized test data.Performed cross-validation for model tuning and optimization of parameters, resulting in an enhanced accuracy by approximately 2%.	
Dog Image Generation using GANs , University of Windsor, Windsor, ON 	Sept 2021 - Dec 2021
<ul style="list-style-type: none">Created a project to generate images of dogs using generative adversarial networks (GANs) and TensorFlow.Developed a model that generated accurate looking dog images with a resolution of 128x128 and could scale to larger resolutions based on training time and hardware capabilities.	

Certifications

Google Advanced Data Analytics , Professional Certificate 	Jun 2023
Microsoft Certified Azure Fundamentals , Professional Certificate 	Jun 2023
IBM Data Scientist , Professional Certificate 	May 2023
Outstanding Scholastic Achievement and Excellence , Golden Key International Honour Society 	Sept 2022
Data Analyst Nanodegree , Udacity 	May 2020
Data Scientist with Python Track , DataCamp 	May 2018

Publications

Anomaly Detection in Large Datasets: A Case Study in Loan Defaults , Thesis, University of Windsor 	Aug 2023
Hyperparameter Tuning for an Enhanced DDPG Framework , Paper, WIDECOM 2023 Conference 	Jul 2023
A Novel System Architecture for Anomaly Detection for Loan Defaults , Paper, DCAI 2023 Conference 	Jul 2023
Efficient Fake News Detection Method using Feature Reduction , Paper, WIDECOM 2022 Conference 	Mar 2022

Awards and Additional Experience

Judge , Youreka Windsor Research Symposium, Windsor, ON	Apr 2023
Treasurer , QIS Club, University of Windsor, Windsor, ON	May 2022
General Member , CKI International, University of Windsor, Windsor, ON	Oct 2021
Best Employee Award , DBS Bank, Hyderabad, India	Jun 2021
Rising Star Award , DBS Bank, Hyderabad, India	Aug 2019