

CAREER SUMMARY

Data Scientist with 3 years of experience and a Master's degree in Data Science - Computational Linguistics from The University of British Columbia. Skilled in developing data applications using machine learning and Large Language Models. Proven ability to lead end-to-end projects, from conceptualization to deployment. Expertise in designing and implementing ETL pipelines and building data warehouse infrastructure. Experienced in developing Retrieval-Augmented Generation (RAG) solutions and fine-tuning Large Language Models. Strong background in fast-paced startup environments.

TECHNICAL SKILLS

Programming: Python, R, SQL, JavaScript, Bash Scripting, HTML, CSS

Data Engineering: PostgreSQL, MongoDB, Pgvector, Solr, Elasticsearch, Redshift, AWS (EC2, ECS, S3, Lambda), Data Orchestration (Prefect, Airflow), Github Actions, Docker

Machine Learning: scikit-learn, Pytorch, Large Language Models, XGBoost, LangChain, Llamaindex

Visualization and Business Intelligence: Power BI, Matplotlib, Seaborn, Plotly

Data Application: Flask, FastAPI, Streamlit

PROFESSIONAL EXPERIENCE

Independent Data Scientist | Self-employed – Vancouver

Dec 2024 – Present

- Developed **Iranian Newspaper Research Assistant**, an application that makes previously non-searchable scanned historical newspaper and document archives searchable through Retrieval-Augmented Generation (RAG).
- Collected and converted scanned archives into machine-readable text, improving accessibility for researchers and historians.
- Implemented full-text and semantic search, enabling intuitive document retrieval.
- Developed graph-based search, allowing users to explore relationships between documents through graph visualizations.
- Built a chatbot-style assistant for intuitive natural language queries, allowing easy exploration of the archives.

Data Scientist at Nu Data - Vancouver - Sustainability & AI Startup

Apr 2024 – Nov 2024

- Developed a custom Agentic Retrieval Augmented Generation (RAG) solution from scratch, providing users with reliable and vetted data, advanced analytics, and visualizations via an intuitive chatbot interface.
- Fine-tuned Large Language Models to improve performance in semantic search, question answering, and function calling.
- Built web scraping tools, automating large-scale data collection from 50+ sources.

Data Scientist at XYON Health - Vancouver - Pharmaceutical Manufacturing Startup

Sep 2023 – Mar 2024

- Sole data scientist responsible for designing and building a data warehouse, reducing reporting time by 70%.
- Designed and orchestrated scalable ETL pipelines integrating data from 11+ sources.
- Delivered actionable insights through connected dashboards for marketing, finance, and operations teams.

Data Scientist at BC Stats - Capstone Project- Vancouver

May 2023 – Jul 2023

- Developed a new semantic search system for BC Data Catalogue, improving search precision by 25% and recall by 40%.
- Developed topic modeling solution, providing users with interactive visualizations, improving BC Data Catalogue's user experience.
- Received two awards for project presentation and was featured on the UBC website.

Data Analyst at BoldZee - IT Services and IT Consulting - Tehran

Feb 2019 – Sep 2021

- Automated data scraping pipelines to extract structured and unstructured data from 40+ sources.
- Developed and optimized ETL pipelines, improving data processing efficiency by 20%.

EDUCATION

- Master of Data Science - Computational Linguistics**, University of British Columbia (GPA: A+)