

SAHIBPREET SINGH

(437)7664913 | ss9334931@gmail.com |

 [Linkedin-profile](#) |  [Github-profile](#) |

Toronto, Ontario, M2N 5X7

OBJECTIVE

Machine Learning Engineer with 5+ years of experience building scalable ML systems, **RAG pipelines**, and production-grade AI applications across cloud platforms (**Azure, AWS**). Proven track record in leading high-impact projects, deploying ML models, and optimizing data pipelines. Adept in **Databricks, Spark, PyTorch, and Azure ML Studio**, with a strong commitment to ethical AI and business-driven solutions.

EXPERIENCE

• CGI

Gen AI Consultant

Aug 2024 – Present

Toronto, Canada

- Led a team of 3 to architect and deploy an **Agentic RAG** system using **Azure Promptflow** for Gen AI workflow orchestration, integrated with Azure Blob Storage for efficient document management, directly contributing to securing a High-Value project worth 700K USD for the organization.
- Designed and implemented a secure **ZERO-TRUST RAG System** for candidate opportunity matching using Semantic Matching, Azure Key Vault for secret management, and Azure Web Apps for scalable deployment. Improved internal recruitment efficiency of CGI by reducing the average opportunity matching time from 8.5 to 3.2 days.
- Integrated **Databricks** and **PySpark** for scalable ML model training and data transformation pipelines, improving the processing speed by 31% in GenAI production systems.

• AI Talentflow

Data Scientist

Mar 2024 – Jun 2024

Philadelphia, USA

- Supervised a team of 3 in competitor market research, analyzing **GenAI products** and ML architectures, resulting in optimized prompt engineering strategies and RAG implementation patterns for our product roadmap.
- Led the development and production deployment of **Resume-2-ResumeRAG**, a GenAI product that leveraging fine-tuned LLMs and custom RAG pipelines. Implemented secure API endpoints with token-based authentication, containerized the solution using Docker, and deployed on AWS ECS with auto-scaling, driving a **\$15k** revenue increase.

• Tatra Data

Data Scientist

Nov 2022 – Dec 2023

Delhi, India

- Engineered and executed a transformative project centred on deploying a **Text-to-SQL** solution with LLM technology, streamlined via **Streamlit**, and integrated with **Langchain** framework, leading to a notable **23%** upsurge in Monthly Active Users.
- Led a team of 5 to develop a centralized system to streamline the creation of **Contextual chatbots** powered by **RASA** for diverse counties in the United States.
- Developed and implemented an end-to-end machine learning model to analyze customer transactions. This initiative led to a significant **62% increase in accuracy**.

• ZS

Data Science Associate

July 2023 – December 2023

Delhi, India

- Supervised the development of a forecasting model to enhance company productivity and efficiency. This initiative successfully generated an impressive **\$325k** in additional annual revenue, highlighting the effectiveness of the model in driving tangible business outcomes.
- Worked closely with a leading global pharmaceutical industry leader **Pfizer** on a regression challenge to predict competition shortages. Utilizing advanced techniques like **DATAIKU**, **RabbitMQ**, **CELERY** and machine learning models like **SBERT** and **Roberta**.

• Tatra Data

Junior Data Scientist

Jan 2020 – October 2021

Delhi, India

- Led a team in developing a Contextual Chatbot for a Data Science portal, resulting in a significant revenue increase of **128%** in the first year of production.
- Designed and deployed a **Contextual chatbot** containerizing the application using **Docker**. Integrated the chatbot with RabbitMQ as the messaging system and utilized **Rasa X** for testing and optimizing its performance. The successful implementation of the chatbot significantly improved the user experience and contributed to revenue growth for the client.

SKILLS

Data Science: Langchain, LlamaIndex, Scikit-learn, Pytorch, Transformers, Databricks, PySpark, Keras, Pydantic, Pytest.

Cloud: Cloud: Azure Blob storage, Key-Vault, Promptflow, AWS S3, ECS, Lambda, AWS Sagemaker, Terraform, Azure ML Studio.

Databases: SQL Server, MySQL, MongoDB, CosmosDB, Azure AI Search

Agentic AI Frameworks: Pydantic, Flask, SmolAgents, Crew AI, Langgraph, Autogen

Graph Libraries: NetworkX, Neo4j

Others: Docker, Git (GitHub), Jira, Web Development, Software Development Life Cycle(SDLC), QA Testing, Microsoft Office 365 Tools, Machine Learning Development Lifecycle(MDLC)

EDUCATION

- **Postgraduate Certificate in Artificial Intelligence and Machine Learning** *Jan 2024 – Sep 2025*
(Lambton College, Toronto, Ontario, Canada)
 - Mentored 3 **Natural Language Processing Workshops** by GDSC as **Co-Tech Lead** on how to get started with NLP and **Rasa** Framework which impact **100+ students** at Lambton College.
- **Bachelor of Technology in Computer Science and Engineering** *Aug 2016 – Jun 2020*
(Punjabi University, Patiala, India)
 - Organized different technical workshops, as **Google Developer Student Club President** boosting coding culture and impacting more than 400+ students.
 - Coordinated 2 cohorts as a **Data Science Mentor** for **Kaggle**, resulting in a global impact on more than **50+ students** and helping them kickstart their Data Science and Conversational AI journey.

AWARDS

- **Second Prize** in **IEEE Hackathon** for Explainable AI
- **Best Mentor Award** in **Kaggle BIPOC** Mentorship Program