Sahar Mirzapoor

Toronto, ON | 437-838-9530 | Sahar.Mirzapoor@georgebrown.ca

https://www.linkedin.com/in/sahar-mirzapoor/ | https://github.com/SaharM80

Objective

Aspiring Machine Learning Engineer currently enrolled in the Applied AI Solutions Development postgraduate program at George Brown College, seeking a position in machine learning or data science. I bring strong skills in data analysis, model development, deep learning, natural language processing, and computer vision, and I am eager to contribute to real-world AI solutions while gaining hands-on industry experience.

Highlights of Qualifications

- Developed **computer vision** deep learning models for **segmentation, classification, object detection**, and trained **autoencoders for anomaly detection**. In addition, used the **Diffusers library** to fine-tune **Stable Diffusion** models.
- Built **classical machine learning** models for **regression and classification** using algorithms like **ensemble methods** and **neural networks**, with full pipelines involving **preprocessing** and **feature engineering**.
- Trained **NLP models** for **classification, token classification, summarization**, and **question answering** using **LSTMs** and fine-tuned transformer models with the **HuggingFace Transformers library**.
- Experienced in **data management** with **MySQL** and **MongoDB**, including **ETL**, **schema/ERD design**, and **dashboard creation** using **Tableau**.
- Familiar with **deployment** using **Docker** and **AWS**, with experience in both **data science** and **software development** roles.
- Supported students as a **TA** for **Python** and **Java**, strengthening **communication** and **teaching skills**.

Technical Skills

Languages: Python, Java, SQL
Libraries & Frameworks: Scikit-learn, TensorFlow, PyTorch, Keras, FastAI, Hugging Face, OpenCV, Pandas, NumPy, Matplotlib, Plotly, NLTK, Librosa, PyCaret, FLAML
Tools & Platforms: Git, Docker, AWS (SageMaker, S3, EC2), Visual Studio Code, Tableau, Power BI, Excel
Databases: MySQL, MongoDB, SQL Server Management Studio, MySQL Workbench

Education & Certifications

Applied A.I. Solutions Development Postgraduate Program   **Jan 2025 – Dec 2025**
George Brown College – Toronto, ON (**GPA: 4.0/4.0**)
Relevant Coursework & Projects:

- **Online Shoppers Intention Prediction:** Trained multiple classification models to predict user purchase behavior using structured data by **preprocessing**, dealing with the class imbalance, **feature engineering** and **hyperparameter tuning.**
- **Old Book Illustrations Generator:** Fine-tuned **Stable Diffusion** using **LoRA** and **DreamBooth** to generate images in a vintage, old book illustration style, showcasing skills in generative AI and creative style transfer.
- **Dialogue Summarizer:** Trained an **encoder-decoder LSTM with attention** and fine-tuned **BART-base** for dialogue summarization. Deployed as a **Streamlit app** for real-time inference.
- **Movie Recommender System:** Built a hybrid movie recommendation engine by training an **SVD-based collaborative filtering** model, extracting feature embeddings from text for content-based filtering, and integrating a popularity-based ranking component.
- **Auto Feature Selector:** Built a Python tool to automate feature selection using statistical tests and model-based methods, with a voting system to rank the most informative features.

Bachelor of Science in Industrial Engineering   **Sep 2019 – Aug 2024**
Iran University of Science and Technology – Tehran, Iran (**GPA: 3.54/4.0**)
Thesis: Diagnosis and Management of Multiple Sclerosis Using Machine Learning

- Built **Convolutional Neural Network models** to detect lesions in MRI scans and predict disease outcomes in Multiple Sclerosis patients using both medical imaging and patient medical history data.

Sahar Mirzapoor

Toronto, ON | 437-838-9530 | Sahar.Mirzapoor@georgebrown.ca

https://www.linkedin.com/in/sahar-mirzapoor/ | https://github.com/SaharM80

Professional Experience

Machine Learning Engineer

DataCoLab – London, UK   *Jul 2023 – Feb 2024*

- **Consensus (Academic QA Chatbot):** Built and fine-tuned NLP models for tasks including question answering, text classification, and named entity recognition to power an academic research-focused chatbot.
- **iVnews (News Structuring System**): Developed NLP models to segment and differentiate sections within transcribed news videos, enhancing content structure and user experience on the platform.
- **Football360 (Comment Filtering Engine):** Led the end-to-end development of a comment-filtering system, handling data preprocessing, addressing class imbalance, and training models using techniques such as ULMFiT.
- **Educational Initiative (Excel & Power BI):** Created comprehensive course content for a humanitarian project, covering all major features of Excel and Power BI to support accessible data literacy education.

Software Developer

Sajayanegar – Tehran, Iran   *Dec 2021 – Jul 2022*

- **Backend Development**: Developed core HRM system features—such as employee calendars, work schedules, leave management, and support ticketing—using the Moqui framework with Groovy (a JVM-based language). Designed the backend architecture based on detailed ERDs to ensure scalable and maintainable data structures.
- **Frontend Development**: Built interactive and user-friendly frontend interfaces with React to support various HRM functionalities.
- **Process Automation**: Designed and implemented automated workflows—such as employee ticket issuance—using Camunda BPMN for streamlined business process management.

Undergraduate Teaching Assistant

Iran University of Science and Technology — **Tehran, Iran**   *Jul 2021 – May 2022*

- **Student Support**: Provided one-on-one assistance to students, helping them debug code and understand key programming concepts in Python and Java.
- **Assignment Grading**: Evaluated and graded programming assignments, offering constructive feedback to help students improve their coding skills and overall performance.

Personal Projects

- **Fire Detector:** Preprocessed images and trained **Faster R-CNN** and **Deformable DETR** models for fire detection.
- **MVTec Industrial Quality Control:** Developed anomaly detection and segmentation models for industrial quality inspection by fine-tuning **ResNet-34** and **Tiny-Swin** for semantic segmentation using **FastAI**. Trained **variational and convolutional autoencoders** for anomaly detection. Deployed via a **Gradio app**.
- **Customer Support QA Agent:** Fine-tuned **T5** for question answering, trained a **BiLSTM** and fine-tuned **DistilBERT** for question classification. Created a **multi-tool agent** using **LangGraph, LangChain, FAISS-based RAG**, and integrated **multi-context prompting** with external search APIs.
- **Industrial Boiler Forecasting:** Built a **FastAPI app** to predict steam temperature using **feature selection**, **statistical tests,** and trained **ARIMA, SARIMAX, XGBoost, Random Forest, LSTM variants**, and a **Temporal Fusion Transformer** with **PyTorch**.
- **Dog Image Generator**: Trained a **convolutional autoencoder** to generate and denoise dog images, and applied **dimensionality reduction techniques** (e.g., t-SNE, UMAP) for embedding visualization and interpretability.
- **Spotify Popularity Predictor**: Built and trained a deep learning regression model using **Keras** and **TensorFlow** to predict song popularity on Spotify based on **audio features** and **metadata**, including preprocessing and feature selection.