Amirreza Naziri

MACHINE LEARNING ENGINEER

📱 +1 437 818 8378

✉ naziriamirreza@gmail.com

 https://github.com/Amir79Naziri

 https://www.linkedin.com/in/amirreza-naziri

🎓 Google scholar

Toronto, Ontario

Personal Profile

Machine Learning Engineer with 2+ years of experience specializing in developing AI-driven applications and conducting research in NLP, generative models, and deep learning. Proficient in Python, PyTorch, and TensorFlow with a strong track record in developing scalable solutions and contributing to open-source projects. Passionate about applying AI to real-world challenges and contributing to the broader AI ecosystem.

Education

York University

MSc in Computer Science   *Jan 2024 - Expected: Dec 2025*

- **Thesis:** Developing ML-based solutions for drug discovery (Single-Cell RNA) using generative AI models and fine-tuning LLMs for biological data interpretation.
- **Key Skills:** Machine Learning, Bioinformatics, Generative AI, LLMs, Python, PyTorch, AI systems, large language models.

Amirkabir University of Technology

BSc in Computer Engineering   *Sep 2018 - May 2023*

- **Related Courses:** Data Mining, Artificial Intelligence, Bioinformatics, AI systems, large language models, Algorithm Design, Applied Linear Algebra.
- **Thesis:** Implemented a BERT-based system for automated misspelling correction in Persian language processing.
- **Thesis Supervisor:** Prof. Hossein Zeinali
- **GPA:** **18.94** / 20

Publications

Exploring Visual Prompt Tuning for Demographic Adaptation in Foundation Models for Medical Imaging
   Artur Parkhimchyk; Amirreza Naziri; Laleh Seyyed-Kalantari
   Proceedings of the NeurIPS Workshop on Algorithmic Fairness in Medical Imaging (AFM), October, 2024

A Comprehensive Approach to Misspelling Correction with BERT and Levenshtein Distance
   Amirreza Naziri; Hossein Zeinali
   arXiv preprint arXiv:2407.17383 (2023). 2023

Related Experience

Vector Institute

Vector Faculty Affiliate Researcher   *Jan 2024 - Now*

- Developing ML-based algorithm for drug discovery focusing on enhancing prediction models using advanced techniques.
- Implementing generative models such as Diffusion models and VAEs to streamline the data synthesize process.
- Developed reference applications and prototype models that bridge AI research with real-world applications.
- Developed generative models and fine-tuned large language models for predictive analysis and commercialization opportunities.
- Applying Foundation models like scGPT for analyzing single-cell RNA sequences, aiming to improve biological data interpretation.
- Utilizing Python and PyTorch to design and test innovative ML workflows, contributing to the early stages of drug discovery.

Amirkabir University of Technology

Research Assistant   *Sep 2022 - Sep 2023*

- Preprocessed and handled large-scale, messy datasets to optimize model performance.
- Conducted research on automated misspelling correction using BERT, enhancing text accuracy for language tasks.
- Fine-tuned BERT models with HuggingFace API, optimizing performance metrics and reducing error rates in language models.
- Developed a Django-based platform to showcase AI capabilities, enhancing accessibility and usability of the model.
- Contributed to open-source libraries enhancing NLP applications in underrepresented languages.
- Utilized Python and PyTorch for NLP tasks, including model optimization and research dissemination through presentations.

Sharif DeepMine

Machine Learning Engineer (NLP Focus)   *July 2021 - Sep 2021*

- Developed and optimized text normalization tools using Python and NLTK, enhancing data preprocessing for NLP models.
- Created pipelines leveraging open-source NLP tools to address linguistic challenges in Persian text processing.
- Evaluated various NLP tools, including SpaCy and TextBlob, to assess their applicability for Persian text processing. itemApplied agile methodologies to enhance NLP pipeline development and model testing.
- Adhered to agile methodologies to streamline NLP model development and iterative testing.

Personal Projects

Image Generation using GANs

Source Code and More Details on GitHub

- Developed generative adversarial neural networks (GANs) using PyTorch and OpenCV to create Anime-style faces from image datasets.
- Enhanced the model's image generation capabilities through extensive training on MNIST and AnimeFace datasets.
- Showcased creative and technical skills in developing AI prototypes for real-world image synthesis challenges.
- **Skills:** Deep Learning, Computer Vision, PyTorch, OpenCV, Model Optimization.

Real-Time Twitter Data Sentiment Analysis

Source Code and More Details on GitHub

- Built a real-time sentiment analysis pipeline using Spark ML and Logistic Regression for tweet classification. Integrated SparkSQL Streaming for continuous data processing.
- Improved data classification speed and enabled dynamic analysis of social media trends.
- Focused on real-time data processing to address large-scale NLP challenges, emphasizing scalability and predictive accuracy.
- **Skills:** Big Data, NLP, PySpark, SparkSQL, Machine Learning, Real-Time Data Processing.

Diabetes Prediction using XGBOOST

Source Code and More Details on GitHub

- Developed a predictive model using XGBoost and Scikit-Learn to classify diabetes patients based on clinical data.
- Visualizing and Analyzing results using Pandas and Matplotlib.
- Improved prediction metrics with 75% accuracy, enhancing decision support in healthcare scenarios.
- Demonstrated expertise in predictive analytics using advanced ML techniques, bridging healthcare and AI solutions.
- **Skills:** Machine Learning, Python, XGBoost, Data Analysis, Model Validation.

More details and other projects are in GitHub.

Skills

Programming Languages	Python, R, JavaScript, SQL.
Machine Learning & AI	PyTorch, TensorFlow (Keras), Scikit-learn, HuggingFace, LLM fine-tuning, Django & Flask.
Big Data Technologies	Spark (PySpark, SparkSQL), Hadoop (MapReduce), Hive, Sqoop.
Data Analysis & Visualization	Pandas, NumPy, Matplotlib, Seaborn, PowerBI (DAX).
DevOps for ML	Docker, Linux, Shell Scripting (Bash).
Databases	MySQL, PostgreSQL, MongoDB.
OS and Virtualization	AWS Services, Linux, Shell (Bash), Windows10, Docker, VMware, VirtualBox, Cloud-Computing.
Miscellaneous	Agile methodologies, SLURM, AI software development, open-source contributions, AI commercialization.
Soft Skills	Time Management, Team Collaboration, Research & Presentation Skills, Problem-Solving.

Courses and Certificates

AWS Certified Cloud Practitioner

Amazon Web Services (AWS)

- Gained foundational knowledge of AWS Cloud concepts, core services, and architectural best practices.
- Understood AWS billing, pricing models, and support plans to optimize cloud cost management.
- Explored key security and compliance aspects within the AWS Cloud environment.
- Developed insights into how cloud impacts business operations and the value of cloud adoption.

Deep Learning specialization

Coursera

- Building neural network architectures such as Convolutional Neural Networks, Recurrent Neural Networks, LSTMs, Transformers.
- Learning how to make models better with strategies such as Dropout, BatchNorm, and Xavier/He initialization.
- Mastering theoretical concepts, and learning their industry applications using Python and TensorFlow.
- Tackling real-world cases such as speech recognition, music synthesis, chatbots, machine translation, natural language processing, and more.

Achievements

2023-2025	**Interdisciplinary AI Scholarship**, Awarded $20,000 Interdisciplinary AI Scholarship for groundbreaking contributions in AI research, particularly in generative modeling and AI applications in healthcare.
2023	**Top 5% Student Award**, Recognized among the top 5% of students in the Computer Engineering Department at Amirkabir University of Technology for outstanding academic achievements.