

# Siavash (Isaac) Barqi Janiar

✉ [siavashbarqi@gmail.com](mailto:siavashbarqi@gmail.com)   [in LinkedIn](#)   [Github](#)   [Google Scholar](#)   [SiavashBarqiJaniar.github.io](#)

---

## Experience

### Data Engineer | Docma

Jan. 2024 – Apr. 2025

- Designed and deployed full-scale **ETL pipelines** from ground up using **advanced Python coding**, including a **custom CLI tool**, **data preprocessing modules**, an **alpha-lag-decay forecasting model (implemented in R)**, and **export functionality**, resulting in **84.1% sales attribution prediction accuracy**.
- Designed and optimized Snowflake and PostgreSQL database architectures**, achieving **2.55× faster query execution** and a **25% reduction in compute resource usage** through **strategic indexing**, **table restructuring**, and **query optimization**.
- Led client-facing meetings** to deliver product demos, gather requirements, and communicate roadmap updates; ensuring 100% client feedback incorporation into engineering priorities and fostering **stronger stakeholder relationships**.
- Developed distributed data processing workflows using **Apache Spark and Spark MLlib**, including model training pipelines for large-scale **sales and churn prediction tasks**, achieving **over 80% accuracy in test environments**.
- Integrated Hadoop HDFS** for large-scale data storage and preprocessing as part of exploratory ML pipelines, enabling fault-tolerant and scalable handling of **multi-source datasets**.
- Automated 80% of recurring data workflows**, reducing manual effort by **12+ hours per week** enabling **self-service analytics** across teams and significantly accelerating reporting and analysis cycles, increasing team productivity and consistency across environments.
- Integrated data ingestion and export across five sources, **local files**, **integrated APIs**, **AWS S3**, **Google Cloud**, and **PostgreSQL databases**; enabling seamless, multi-source data synchronization and improving pipeline flexibility.

### Research Assistant | York University

Sep. 2021 – Feb. 2024

- Proposed a **transfer learning (TL)** method based on **feature extraction** to predict jamming patterns in a communication network. Reduced the time complexity of the primary model by **30 times**.
- Realized a comprehensive **eXplainable AI (XAI)** method comprising **pattern recognition** and **rule learning** for network security. Improved the **transparency** of the model compared to the benchmark explainable models by **17%**, whilst having a **32% less error rate**.
- Introduced an evaluation environment comparing the performance of **Recurrent Neural Network (RNN) structures** utilizing **Long Short-Term Memory (LSTM)** layers, achieving a **13% higher throughput rate** than **Convolutional Neural Network (CNN) structures** while having **x1.2 fewer parameters**.
- Implemented a **Transformer-based model** combined with **K-means clustering** to handle sparse data, using filtering and **PCA** as preprocessing steps, and achieved **88% accuracy** in annotating correct jamming pattern groups.

### Research Assistant | Amirkabir University

Feb. 2020 – Jan. 2021

- Realized an efficient **model-free reinforcement learning** MAC protocol for frequency resource allocation. Surpassed the benchmark protocol with nearly **60% better throughput**.
- Leveraged an **online actor-critic** algorithm for access problems in heterogeneous networks. Achieved **95% throughput** in the network marked as the **highest possible performance**.
- Optimized the resource allocation system in distributed computer networks with prioritized packets using ML/AI, which increased the throughput of the wireless system **by approximately 15%**.

---

## Skills

**Programming Languages:** SQL, Python, C/C++, Java, R, MATLAB, JavaScript, Node.js, C#

**Frameworks:** Power BI, Snowflake, Data Bricks, AWS, GCP, Pandas, Docker, Sci-kit learn, PostgreSQL, Dash, Matplotlib, Plotly, Psycopg2, AWS Wrangler, Apache Kafka, Spark (Spark MLlib), Hadoop, Regex, Boto3, PyTorch, Flask, Postman, Git

---

Selected Projects

Healthcare and Insurance Policy Analytics in Ontario — Power BI / Python [\[🔗 Link\]](#)

- Preprocessed and cleansed **Ontario accident reports data**, enabling a comprehensive analysis that revealed **policy gaps** and identified **areas of employer non-compliance** through advanced data visualization techniques.
- Utilized Scikit-learn to build **predictive models** for forecasting workplace injuries and accidents based on features such as **location, cause, and job type**.
- Performed in-depth data analysis to uncover **trends in high-risk occupations** and proposed data-driven **policy recommendations** to improve workplace safety.
- **Presented findings to senior stakeholders through interactive dashboards** built with **Power BI, Plotly, Dash, and Matplotlib**, highlighting key **insights**, identifying **data gaps**, and providing **actionable recommendations** to support evidence-based decision making.

Real-Time Stock Market Data Pipeline — Python, C++, SQL, Dash, Vue.js [\[🔗 Link\]](#)

- Implemented a **real-time data analysis pipeline** to ingest, transform, and load stock exchange data using **advanced Python and thread programming**, supporting continuous updates and ensuring data consistency across components.
- Used **Apache Kafka** to build a scalable data streaming architecture, improving system reliability and enabling efficient processing of real-time stock market data.
- Optimized **PostgreSQL** for analytical workloads and containerized the environment using **Docker**, leveraging indexed time-series partitions, materialized views, and query plan tuning to reduce data retrieval time.
- Developed a **RESTful API** using **Flask** and **Vue.js** to deliver real-time stock market predictions. Integrated **Apache ECharts** to enable users to interact with **over 10 types of dynamic charts and data visualizations**.

RAG-based Chatbot — LangChain / FAISS [\[🔗 Link\]](#)

- Built a **RAG-based chatbot** leveraging **LangChain** and **OpenAI language models** to provide context-aware, retrieval-augmented responses from custom knowledge sources.
- Implemented semantic search using a **FAISS retriever** with a **lightweight embedding model** for efficient vector-based document retrieval.
- Preprocessed documents using **recursive text splitter** for optimal chunking and embedded them into a **Chroma vectorstore** for fast and scalable retrieval.
- Deployed the chatbot using **Streamlit**, delivering an interactive and production-ready interface for end users.

Recent Publications [\[🔗 Link\]](#)

- S. B. Janiar, P. Wang, "Intelligent Anti-jamming based on Deep Reinforcement Learning and Transfer Learning," *IEEE Transactions on Vehicular Technology*, 2023.
- S. B. Janiar, Xian Lu, P. Wang, "Explainable Reinforcement Learning for Wireless Security at the Physical Layer: A Survey," *IEEE Transactions on Wireless Communications*, 2022.
- Barqi Janiar S, Pourahmadi V, "Deep-reinforcement learning for fair distributed dynamic spectrum access in priority buffered heterogeneous wireless networks," *IET Commun.* 2021;19. <https://doi.org/10.1049/cmu2.12098>
- S. B. Janiar, P. Wang, "A transfer learning approach based on integrated feature extractor for anti-jamming in wireless networks," *IEEE PIMRC, Toronto*, 2023.
- S. B. Janiar and V. Pourahmadi, "Deep-Reinforcement Learning for Fair Distributed Dynamic Spectrum Access in Wireless Networks," *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, 2021, pp. 1-4, doi: 10.1109/CCNC49032.2021.9369536.
- S. B. Janiar, A. Eckford, "The Theory and Applications of Coded Modulation in Digital Communications: A Survey", *York University*, Dec. 2021.

Education

York University   Master of Science in Computer Science	Sep. 2021 - May 2023
Field of Study: AI and Machine Learning	GPA: 8.8/9
Amirkabir University   Bachelor of Science in Electrical Engineering	Sep. 2016 - Jun. 2021
Focus: AI and Machine Learning	GPA: 3.84/4