

# Amirmehdi Sharifzad

San Francisco, California, United States ✉ amsharifzad@gmail.com ☎ 6502837845 📄 in/amirmehdi-sh

## SUMMARY

Dynamic Data Scientist with a proven track record of solving complex problems using analytical skills honed through rich experiences in software development, machine learning, and data science. Adept at translating data insights into actionable business strategies, with a focus on developing cutting-edge machine learning models and optimizing data-driven processes. Brings a unique blend of technical acumen, creative problem-solving, and a results-driven approach to drive impactful business outcomes.

## SKILLS

Languages: Python, SQL, Golang

Libraries/Tools: Spark, Scikit-Learn, Scipy, NLTK, PyTorch, Airflow, Kubernetes

## EXPERIENCE

### Data Scientist

Viaduct

September 2021 - Present, Menlo Park, CA

- Spearheaded the development and end-to-end deployment of survival models using **Python** to forecast vehicle population claims, enhancing accuracy by 12%.
- Managed data analysis, profiling, and cleansing for new customer data, leading to advanced analytics and custom dashboard creation in Metabase using **SQL**.
- Pioneered the design and development of a declarative feature platform for internal feature engineering, leveraging **Python**, **PySpark**, and **Kubernetes**. The platform automates computation, backfill, and orchestration of pipelines, significantly optimizing feature store management in Iceberg.
- Engineered and deployed an integrated suite of **NLP** projects, including a claim classification model, diagnostic code extraction pipeline, and embedding generation using a fine-tuned LLM on automotive terminology. This suite significantly enhanced claim processing efficiency, reducing processing time by 30%.
- Led the creation of an experimental chatbot using GPT-4, capable of interpreting user data queries, executing SQL queries or API calls, and presenting insights through data summarization or visualization.

### Machine Learning Research Intern

Huawei

May 2020 - March 2021, Montreal, Quebec

- Led a landmark research initiative for a multi-domain translation model, improving translation accuracy by 35% using enhanced FairSeq framework and **PyTorch**, resulting in a patented, industry-leading **Transformer** architecture.
- Key contributor in upgrading Huawei's translation app, integrating novel architecture, leading to a 14% increase in user satisfaction, backed by user feedback and engagement metrics.
- Collaborated in R&D for optimizing language processing algorithms, significantly enhancing real-time translation accuracy and user experience.

### Data Scientist Intern

Scribd

September 2024 - December 2019, San Francisco, CA

- Developed a Scala-based machine learning pipeline for search optimization, resulting in a 15% increase in search accuracy and a 25% reduction in response time.
- Implemented the FuzzyWuzzy string matching algorithm based on Levenshtein Distance in Scala, and developed a Gradient Boosting Regression model using **Spark ML** and **XGBoost**, showcasing advanced technical skills in feature engineering and model development.
- Achieved a 60% reduction in feature generation runtime by implementing the Fast KNN algorithm for approximating KNN graphs, significantly optimizing computational efficiency and processing speed.

## EDUCATION

### Bachelors of Computer Science

Minor in Statistics • University of Waterloo • Waterloo, Ontario • 2021

## PROJECTS

### doodle-ai

[https://github.com/Mehty/doodle\\_ai](https://github.com/Mehty/doodle_ai) • April 2019 - June 2019

- Developed and implemented the **NST algorithm** utilizing TensorFlow in Python to generate artistic images by blending the style of one image with the content of another image.