

Feraidoon Mehri

✉ feraidoonmehri@gmail.com | 🌐 github.com/NightMachinery | 🔗 linkedin.com/in/feraidoon-mehri

EDUCATION

Sharif University of Technology

Master of Science (M.Sc.) in Artificial Intelligence

Tehran, Iran

9/2022 - ongoing

- Supervisors: Dr. Mahdieh Soleymani Baghshah and Dr. Mohammad Taher Pilehvar
- GPA: 19.25/20.0
- Courses:** (The grades are from 20.)
Natural Language Processing (20). *Information Theory & Coding* (18.5). *Signals and Systems* (20).

Sharif University of Technology

Bachelor of Science (B.Sc.) in Computer Science

Tehran, Iran

9/2016 - 7/2021

- GPA: 17.1/20.0
- GPA of the Last 64 Credits: 19.0/20.0
- Notable Courses:** (Graduate level courses are marked with ⁺. The grades are from 20.)
Statistics (20). *Probability* (19.8). *Applications of Stochastic Processes*⁺ (19). *Artificial Intelligence* (19.7).
Cryptography, Distributed Systems, and Blockchains⁺ (18.9). *Advanced Programming* (20). *Data Transfer and Networks*⁺ (19.3). *Theory of Computation & Complexity* (18.4). *Mathematical Logic* (18.5). *Systems Theory*⁺ (18.6). *Big Data Engineering* (19). *Analysis of Algorithms*⁺ (20). *Engineering Mathematics* (20).

Shahid Beheshti (National Organization for Development of Exceptional Talents)

Preuniversity in Mathematics & Physics

Sabzevar, Iran

9/2015 - 7/2016

- GPA: 20.0/20.0

Shahid Beheshti (National Organization for Development of Exceptional Talents)

High School Diploma in Mathematics & Physics

Sabzevar, Iran

9/2012 - 7/2015

- GPA: 19.9/20.0

HONORS AND AWARDS

- Ranked **second** (out of 16,703 students) in the Iranian national graduate entrance exam in *AI & Robotics* (and all other CE majors) of 2022 (**99.99 percentile**)
- Ranked **fourth** (out of 1,322 students) in the Iranian national graduate entrance exam in *Computer Science* of 2022 (**99.70 percentile**)
- Ranked **sixth** (out of 115,803 students) in the Iranian national undergraduate entrance exam in *Foreign Language Studies* of 2016 (**99.99 percentile**)
- Ranked 331st (out of 162,731 students) in the Iranian national undergraduate entrance exam in *Mathematics & Physics* of 2016 (**99.80 percentile**)
- Member of Iran's National Elites Foundation (INEF)

RESEARCH EXPERIENCE

Analysis and Interpretation of Transformer Models

Machine Learning Lab (MLL)

09/2022 - ongoing

Sharif University of Technology

Tehran Institute for Advanced Studies (TeIAS)

Under the supervision of Dr. Mohammad Taher Pilehvar and Dr. Mahdieh Soleymani Baghshah

- Decomposing transformer networks into attribution scores for each input token

Blockchain-Based Solutions to Privacy-Preserving Health Data

Under the supervision of Dr. Parviz Rashidi Khazaei

08/2022 - 09/2022

Urmia University of Technology

- health_blockchain** (Python): prototyped a blockchain for storing health data privately in a distributed manner

Manifold Learning and High-Dimensional Clustering

Sharif Optimization and Applications Laboratory (SOAL)

12/2021 - 09/2022

Sharif University of Technology

Under the supervision of Dr. Amir Daneshgar, Dr. Mohammad-Hadi Foroughmand, and Dr. Mojtaba Tefagh

- Solved the Optimizer 2022 challenges around manifold learning and clustering in high-dimensional data with outliers and noise
- Visualized the high-dimensional input data, the detected manifolds, clusters, convex hulls, outliers, and noise, which was critical in diagnosing many bugs
- Developed, tested, and calibrated the automatic judge (the autograder) of the Optimizer 2022 challenges
- Designed and tested the data generation algorithms for the Optimizer 2022 challenges
- Created a modular benchmark system for clustering algorithms that measures memory usage, execution time, and various accuracy metrics, with support for big (~100GB) data (Dask, RAPIDS, scikit-learn)
- Created guidelines for documenting and organizing libraries for the lab
- Mentored teammates in Git and Jupyter
- Supervised four undergraduate students

RESEARCH INTERESTS

- Deep Learning
 - Interpretability
 - Transformers
 - Natural Language Processing
 - * Large Language Models
 - Multimodal Learning
 - Computer Vision
 - AI Alignment

TEACHING EXPERIENCE

Miscellaneous Workshops

Clustering in Python from Scratch

Student Scientific Association (Hamband), Sharif University (Spring, 2022)

Head Teaching Assistant

Computer Networks

Dr. Arshadi, Sharif University of Technology (Spring, 2023)

Principles of Computer Systems

Dr. Arshadi, Sharif University of Technology (Fall, 2022)

Teaching Assistant

Deep Learning

Dr. Soleymani, Sharif University of Technology (Spring, 2023)

Natural Language Processing

Dr. Pilehvar, Tehran Institute for Advanced Studies (Spring, 2023)

Artificial Intelligence

Dr. Rohban, Sharif University of Technology (Fall, 2022)

Big Data Engineering

Prof. Ostovari, Dr. Boomari, Prof. Taheri, Sharif University of Technology (Spring, 2022)

Advanced Programming

Dr. Fazli, Sharif University of Technology (Spring, 2022)

Digital Design (Logic Circuits)

Dr. Arshadi, Sharif University of Technology (Spring, 2022)

Computer Networks

Dr. Arshadi, Sharif University of Technology (Spring, 2022)

Principles of Computer Systems

Dr. Arshadi, Sharif University of Technology (Fall, 2021)

Advanced Programming

Dr. Boomari, Sharif University of Technology (Spring, 2018)

SKILLS

Programming Languages I Use Frequently (In Rough Order of Proficiency): Zsh (shell scripting), Python, Emacs

Programming Languages I Have Written Some Useful Things in (In Rough Order of Proficiency): Julia, Java, Common Lisp, Golang, Perl, Clojure, Scala, Kotlin, Racket, Lua, Javascript, Node.js, CSS, HTML, LaTeX, SQL, Nim, Rust, VB.NET, C#, C++

ML/Data Libraries: PyTorch, HuggingFace, Google's JAX, Google's Flax, Deepmind's Haiku, Deepmind's Optax, Flux.jl, scikit-learn, NVIDIA's RAPIDS, Dask, conda, numpy, numba, pandas, einops, networkx, spaCy, Matplotlib, seaborn, plotly

Backend Technologies: GNU/Linux, Docker, Caddy, Akka, Redis, FastAPI

Frontend Technologies: Svelte, Hugo, Docusaurus

Developer Tools: Git, tmux, Emacs, vim, VSCode, Jupyter, Google Colab

Other Technical Skills: web scraping, blockchains, distributed systems, regex, documentation writing and note taking, prompt engineering

NOTABLE OPENSOURCE PROJECTS

- Reported hundreds of bugs

Popular FOSS Projects I Have Contributed To

- **ugrep**: suggested many design improvements which *Dr. Robert van Engelen* liked and subsequently implemented
- **HuggingFace Datasets** (Python): added features
- **Emacs** (Elisp): added features
- **Doom Emacs** (Elisp): added features and fixed bugs
- **Flux.jl** (Julia): fixed mistakes in the documentation and wrote more documentation
- **Zsh**: reported a critical bug which was promptly fixed
- **tldr**: wrote documentation
- **fzf-tab** (Zsh): fixed bugs
- **quail** (Clojure): added features
- **learnxinyminutes.com**: fixed mistakes
- **bkmeans** (Python, scikit-learn): fixed bugs
- **Anime4KCPP** (C++): added macOS support

Academic Projects

- **stochastic** (Julia): an infectious disease model (a grad course project of mine), a Poisson picture redrawing filter, a colorful animator of a 2D ising model, and more
- **MLP From Scratch**: a simple trainable MLP using only numpy with support for batch axes
- **price_detector_fa** (Python, Hazm): extracts product/price/amount tuples from Persian text using rule-based methods
- **Cross-Lingual Transfer Learning From English to Persian With Zero-Shot/Few-Shot MAD-X** (Python, PyTorch, HuggingFace, adapter-transformers)
- **Char-RNN** (Python, PyTorch): a character-level language model using GRUs
- **fanfiction-classifier** (Python, JAX, Haiku, Optax): a character-level, variable-length text classifier using (optionally dilated) convolutional layers, dropout, layer normalization, and learning rate scheduling (runs on both TPUs and GPUs)
 - * A similar model in Julia's Flux.jl
- **twitter-scraper** (Python, Zsh, Docker): a fault-tolerant, distributed Twitter scraper which stores the data in Neo4j (a distributed graph database), plus a high-level CLI API for querying the data and load testing the system
- **distributed-prime-generator** (Scala, Akka, Docker): a fault-tolerant, distributed prime number generator using the actor model
- **Toy MIPS CPU**: an implementation of a simple, non-pipelining CPU using gate-level Verilog
- **reo** (Racket): a toy DSL

My Own FOSS Projects

- **brish**: a thread-safe Python library using the metaprogramming API which lets the user embed and run Zsh code in Python via parallel processes, supporting safely interpolating Python variables into the Zsh code (downloaded **57k** times)
- **readability-cli** (Node.js): declutters and sanitizes scraped HTML using Mozilla Readability
- **betterborg** (Python): a pluggable Telegram (user)bot based on Telethon (forked from uniborg)
 - * Integration with my library brish, to allow using Telegram as, essentially, a terminal emulator, with support for safely giving users limited access to specific Unix commands
 - * A time and habit tracker that supports hierarchical activities, with an always learning DSL to submit events and request visualizations
 - * A declarative DSL based on JSON for Telegram's inline query UI
 - * A prototype of a declarative DSL based on JSON for Telegram's bot UI
- **JupyterGarden** (Python, FastAPI): an HTTP REST API to run code in Jupyter kernels, for languages with expensive startup costs
- **possiblycat** (Golang): cat with a timeout on waiting for the first byte of stdin
- **prefixer** (Golang): a modern alternative to GNU cut
- **jalalici** (Golang): a CLI utility for Jalali (Shamsi) dates
- **jalali-calendar-cli** (Python): a TUI Jalali (Shamsi) calendar (holiday data extracted using LLMs from official PDFs)

• My Megarepo of Scripts

- * Prompt crafting tools
- * Scraping tools (or API wrappers) for github.com, semanticscholar.org, arxiv.org, edu.sharif.edu, cw.sharif.ir, goodreads.com, reddit.com, tumblr.com, spotify.com, store.steampowered.com, nationalgeographic.com, bing.com, duckduckgo.com, fanfiction.net, kitsu.io, myanimelist.net, techcrunch.com, sounds-resource.com, lesswrong.com, messaged.com/tldr, web.archive.org, techmeme.com, news.ycombinator.com, sanjesh.org, patreon.com, ...
 - **web2audio**: creates an audiobook from a given set of URLs using pretrained deep TTS models
 - **t.me/techmemenews**
 - **t.me/tldrnewsletter**
 - **t.me/sharif_edu_diff**: notifications of changes in the scheduling, capacity, instructors, etc. of SUT's courses of the current semester
 - **sharif_course_list**: courses offered by SUT, in HTML and JSON, under git for archival purposes
 - **htmlmetadata** (Nim): a CLI app that extracts metadata (author, description, etc.) out of HTML
 - **ddg2json** (Python): parses scraped HTML of DuckDuckGo pages into JSON
 - **r_rational**: the subreddit r/rational archived in plain-text org-mode (good for, e.g., doing offline full-text searches)
- * A general Zsh function memoizer using Redis
- * A git-backed reminder system supporting natural language for setting the due time, recurrent reminders, and integration with Telegram, Google Calendar, and macOS/iOS (notifications, widgets, wallpapers)
- * A Redis-backed RSS manager with integrations for Telegram and Amazon Kindle (supports podcasts, as well)
- * A recurrent bill manager that parses my org-mode (plain-text) notes and presents me the bills likely to be due
- * A whole suite of index-less information retrieval CLI tools for searching IRC logs, reminders, contacts, function definition locations, oft-used directories, music, ...
 - A Perl-based custom grep tool for tree-shaped documents (org-mode), plus a TUI for viewing the results (built on top of Emacs)
- * A file tagging system using the file names as the database
- * A “thermostat” for display brightness based on the brightness of the current screen content
- * A clipboard manager integrated with Zsh and Emacs
- * Integration with Supercollider, to allow using real-time, stochastic generative audio for auditory notifications in CLI applications
- **rtl_reshaper.rs** (Rust): a CLI tool to reshape and reorder bi-directional, Arabic/Persian text for display in apps that do not support them natively
- **vcard-to-json** (Clojure): a CLI tool to convert vCard files into JSON
- **doom.d** (Elisp)
 - * An extension to the org-mode markup language to support visualizing and navigating annotated text for NLP purposes
 - * Integration with my library brish that allows embedding Zsh code in Elisp with support for safe interpolation of Elisp values
 - * A poor man's language-agnostic doctests using literate programming (via org-babel) and version control (via Git and ediff)

FOSS Projects I Developed When I Was a Child

2007-2012

- **Aero Form** (VB.NET): A subclass of Form, it allowed the user to extend the Aero effect of Windows Vista to the whole window (or a subset thereof). It was downloaded **60k** times on marketplace.visualstudio.com.
- **HyperAero Form** (VB.NET): The much-upgraded version of the above, it was my first lesson in marketing; while it was better than Aero Form for virtually all purposes, its not as catchy name doomed it to obscurity. It was downloaded **13k** times on marketplace.visualstudio.com.
- **Notify Msg** (VB.NET): Probably my first useful library, it allowed one to show desktop notification popups, supporting images and other goodies. It was downloaded **1.5k** times.
- **File Splitter** (VB.NET): A WPF-powered GUI for splitting and merging files.
- **Animation Maker** (VB.NET): Inspired by Windows Presentation Foundation's ease of creating animations, I built an animation library for Windows Forms using the reflection API (demo).