Ali Najafi

+90 (543) 167-1440 / ali.najafi@sabanciuniv.edu / github.com/alinajafi1998 / najafi-ali.com

Education

- **Sabanci University**, Istanbul, Turkey, 2022 - Present
  M.S. in Computer Engineering, CGPA - 4/4
  Awarded fully funded scholarship

- **University of Tabriz**, Tabriz, Iran, 2017 - 2021
  B.S. in Computer Engineering, CGPA - 3.3/4

Publications

- **A Najafi**, A Gholipour-Shilabin, R Dehkharghani, A Mohammadpur-Fard, M Asgari-Chenaghlu. **ComStreamClust: A communicative multi-agent approach to text clustering in streaming data, Annals of Data Science**, 2022, **DOI: 10.1007/s40745-022-00426-4**

Experiences

- **Teaching Assistant - Sabanci University** - Istanbul, Turkey, 2/22 - Present
  - CS412 - Machine Learning
  - CS201 - Introduction to Computing - Fundamental C++ programming

- **Software Engineer - NOAY** - Tabriz, Iran, 6/21 - 8/21
  - Built automatic and reusable data processing modules for general use of scientists from difference backgrounds.
  - Developed modules using the Pandas library with a team of five developers.

- **Git Workshop - University of Tabriz** - Tabriz, Iran, 11/20 - 11/20
  - Taught Git and Github to a class of 30 freshmen

- **Machine Learning Engineer - RAyin Samaneh Arta (RASA)** - Tabriz, Iran, 6/20 - 8/20
  - Created a system to detect specific landmarks on dental images that are beneficial for dentists to evaluate patients' dental beauty.
  - Worked with a team of three developers, and used VGG-face as a pre-trained model to improve the system.
- **Android Developer - Khavaran Tosee** - Tabriz, Iran, 6/18 - 8/18
  - Made an Android application for internal use of the company using React Native and Java.

Projects

- **ComStreamClust, 2022**
  - Clustering and extracting trending topics from millions of tweets on streaming data.
  - Worked with a team of five engineers and researchers, used parallelization to assign the load on different cores to speed up the process, and later published the code and results as a paper.
  - Responsible for designing the system, partially developing the Python code, evaluating, and conducting analysis.

- **Trec2020DeepLearningTask - Document/Passage Ranking, 2022**
  - Developed a pipeline to retrieve the relevant documents/passages for the input queries (MSMARCO).
  - Utilized DocT5Query to expand the documents by utilizing the newly generated queries.
  - The expanded documents were retrieved using ElasticSearch(BM25).
  - T5 was applied to filter out retrieved documents from the BM25 algorithm.

- **Youtube Comments Sentiment Analysis, 2022**
  - Implemented a sentiment classifier for comments of youtube tutorial videos
  - Informal text preprocessing

- **NLPKickStart, 2022**
  - Developed a modular pipeline for preprocessing informal text data of online platforms such as Twitter.
  - This pipeline uses scikit-learn, NLTK, and emoji libraries.

- **Movie Recommender System, 2021**
  - Implemented a movie recommender system using item-based and user-based collaborative filtering with Python.
  - The classifier predicts the score a user might give to a movie on a 1-5 scale.
  - The model achieved a Manhattan distance of 0.781 for the test set, which is a 220% improvement over base models like linear regression.

- **Machine Translation, 2020**
  - Built an AI model for translating English to Italian.
  - Implemented a custom-made Transformer-based model from scratch using Tensorflow. In spite of the limited resources and the small size of the model, it achieved a BLEU score of 21.5.

- **Kaggle Competition (SIIM-ISIC Melanoma Classification), 2020**
  - Classified cancerous and healthy skin on extremely unbalanced data (anomaly detection).
  - Improved results by stacking and ensembleing different models, such as Random Forest and XGBoost.
  - Pre-trained models such as VGG19 and Xception were selected for transfer learning.

- **Silk Road Graph Analyzer, 2020**
  - Created a desktop application for graph analysis in a group of three students using Java and JavaFX.
  - Responsible for solving the Travelling Sales Man problem using AntColony and Dynamic Programming.

Skills

- **Languages:** Python (Proficient), C++
- **Frameworks**: Tensorflow, Keras, Pytorch, Transformers, Django
- **Other:** Git, MySql, MongoDB, Docker

Languages

- **Persian & Azerbaijani:** Bilingual
- **English:** Professional Working Proficiency (Academic IELTS 7)
- **Turkish**: Elementary