S. M. Amin Taheri G.

St. John's, Canada

📱 +1 709 219 7722 • ✉ amintg@mun.ca • ☗ amintaheri23.github.io

Experience

MITACS Research Intern   **St. John's, Canada**
Verafin Inc. (Nasdaq) ↗   *2023 Oct - 2024 Feb*
Under the supervision of **Dr. Terrence Tricco** ✉ and **Dr. John Hawkin** ✉:
- Addressed the challenge of financial fraud detection by harnessing the power of textual data and incorporating language modelling techniques such as Transformers and Word2Vec algorithms, improving **36.2%** with **Tiny LLAMA over One Hot encoding**.
- Investigated the effects of different categorical encodings on the performance of fraud detection models with **Stratified Cross-validation** tailored for **class imbalance**.
- Examined different machine learning models in fraud detection, such as **LightGBM**.
- Discovered that when there is little numerical data available, using the word embeddings of Transformers and Word2Vec can improve fraud detection.

Instructional Assistant (BAND LEVEL 7 - CUPE 1615)   **St. John's, Canada**
Memorial University of Newfoundland (MUN), Dept. of Computer Science   *2022 Nov - 2023 Apr*
- Responsible for two Comp1001 and two Comp1003 labs.
- Supervised Teacher Assistants during the labs.

Computer Vision Intern   **Tehran, Iran**
Shenasa-ai Company ↗   *2019 Oct - 2019 Dec*
Under supervision of **Alireza Akhavanpour** ✉:
- Developed a script for **Face Extraction in Video** with Retina Face and OpenCV, capable of running at 3x the speed of the video and extracting clips containing faces.
- Researched **Face Quality Assessment in Videos** and re-implemented a paper named "Score-based Face Quality Assessment" using LBP, HOG, and CNN feature extractors with an Auto Encoder.
- Presented on TDD with Python during the internship.

Education

M.Sc. in Computer Science   **St. John's, Canada**
Memorial University of Newfoundland (MUN) ↗   *2021 Sep - 2024 Apr*
Thesis: BacTermFinder: Bacteria-agnostic Comprehensive Terminator Finder using a CNN Ensemble ↗ under the supervision of **Dr. Lourdes Pena-Castillo** ✉, we worked on identifying bacteria terminator sequences from it's genomes using an ensemble of Convolutional Neural Nets (CNNs).
GPA : 93/100 - 4/4

Publications

Taheri Ghahfarokhi, SMA, & Peña-Castillo, L. (2024). BacTermFinder: A Comprehensive and General Bacterial Terminator Finder using a CNN Ensemble. *BioRxiv*, https://doi.org/10.1101/2024.07.05.602086

Honors and Awards

Fellow of the School of Graduate Studies.   **St. John's, Canada**
Memorial University of Newfoundland (MUN)   *2023 Oct*
This award was made in recognition of my continued academic excellence throughout my program

Outstanding Teaching Assistant Award   **St. John's, Canada**
Memorial University of Newfoundland (MUN), Dept. of Computer Science   *2023 Apr*

Selected Projects

IMDB Sentiment Classification with GPT2 ↗
- Fine-tuned **GPT2** with little data (around 500 sentences).
- Created a sentence embedding from averaging word sentences.
- Performs better than a logistic regression baseline.

Single Face Extractor ↗
- Under the co-supervision of **Alireza Akhavanpour** ✉ and Bahar Baradaran Eftekhari.
- Extracts video clips that contain only a single face from a video.
- Utilizes **OpenCV, Python, and Retina Face**.

Named Entity Recognition for Persian Text ↗
- Built and trained a model to recognize **Named Entities** in a sentence.
- Assigns a tag to each word in a sentence.
- A classic application for Natural Language Processing (NLP).

Extracurricular Activities

President (Head of Board of Directors)   **St. John's, Canada**
Computer Science Graduate Society of MUN ↗   *2022 Dec - 2024 Feb*
- Planning and executing an election for the new generation of the board members
- Holding weekly meetings to plan events

Graduate Students Representative   **St. John's, Canada**
Memorial University of Newfoundland (MUN) Dept. of Computer Science   *2022 Sep - 2023 Sep*
- Responsible for acting as a bridge between the department and the graduate student body

Skills

Programming: Skilled in Python3, C++ Familiar with Java, SQL, Bash, MATLAB, JavaScript

Databases: SQL, MySQL, SQL Server, SQLite

Could technologies: AWS SageMaker

Python Packages: Scikit-learn, Tensorflow, Keras, Pytorch, LGBM, XGBoost, Numpy, Pandas, Matplotlib, Seaborn, Jupyter, Django, Flask, Fast API

Social Skills: Leadership, Public Speaking, Team Working, Time Management

Other Skills: Linux (GNU Core Utilities, Text Processing, Networking, System Monitoring) LaTeX, Unified Modelling Language (UML), Git

Certificates

Deep Learning Specialization ↗
Andrew Ng, Deeplearning.ai by Coursera   *2020 Aug*
Through **five interconnected courses**, learners develop a profound knowledge of the hottest AI algorithms, mastering deep learning from its foundations (neural networks) to its industry applications (**Computer Vision, Natural Language Processing, Speech Recognition**, etc.)

Machine learning Tutorial ↗
Andrew Ng, Stanford online by Coursera   *2019 Nov*
Machine Learning Course tutorial by Stanford online in Coursera.com.

Deep Learning for NLP ↗
Reza Vasefi and Fateme Mashhadi, FanAsa Academy ↗   *2019 Aug*
Deep learning for NLP using Tensorflow, Keras in Persian Language