Arian S. Hosseni

No 53, Bagheri St, Resalat Sq,
Tehran, Iran

E-mail: arian.hosseini9@gmail.com
Home Page: http://arianhosseini.github.io

Education

◇ **Iran University of Science and Technology,** 2013 - Present

· B.Sc: Computer Engineering student, Software Major
Ranked **First** among the entire university computer engineering students.
Total GPA (6 Semesters) : 19.12/20 (4.0/4.0)
Selected Courses: Computational Artificial Intelligence, Introduction To Computer Vision and Digital Image Processing

◇ **Pre-University Degree in Physics and Mathematic,** 2012-2013
Saadat HighSchool and Pre-University, Fars.
Ranked First
GPA: 19.98/20

◇ **Diploma in Physics and Mathematic**, 2009-2012
Saadat HighSchool and Pre-University, Fars.
Ranked First
GPA: 19.67/20

Research Interests

· Deep Learning in Natural Language Processing
· Question Answering and Reading Comprehension
· Machine Translation
· Distributed Representations of Language Units
· Machine Learning / Deep Learning

Research & Work Experience

◇ **Intern Researcher in Deep Natural Language Processing** (June 2016 - Present),
Supervised by Hinrich Schuetze, **The Center for Information and Language Processing, Ludwig Maximilian University of Munich, Germany**.

· **Exploration of various deep neural networks for Question Answering**
Status: In progress
Sascha Rothe, Arian S. Hosseini, Hinrich Shuetze.

In this project, we are working on various deep models such as seq2seq models with different kinds of attention mechanisms, pointer networks and sequence rankers on the recently released dataset, Stanford Question Answering Dataset (SQuAD). Our end-to-end neural architecture encodes a passage based on a question and tries to find a boundary for answer in the passage. For more info check here. This project is implemented using Blocks and Theano.

◇ **Collaborating with Life Language Processing at UC Berkeley** (June 2016 - Present),
Life Language Processing: applications of deep language processing for processing of biological sequences (DNA, RNA, and proteins). These methods introduce a new perspective for bioinformatics to infer information about structure and function of biological sequences. I work with **Ehsan. Asgari**.

· **Predicting the specificities of DNA and RNA sequences using deep learning**
Status: In progress
E. Asgari, A. Hosseini

In this project we apply a deep convolutional neural network (CNN) over a sequence as inputs into a long short-term memory (LSTM). This task classifies whether or not there is a binding site for a particular TF of interest when given an input DNA sequence.

◇ **Research Assistant** (January 2016 - June 2016),
Supervised by Dr. Behrouz Minaei, **Machine Learning & Data Mining Lab, Computer Engineering Department, Iran University of Science and Technology**.

· **Alefba: Persian Optical Character Recognition System**
Status: To be submitted
A. Nourian, A. Hosseini, A. Mehdizade, B. Minaei.

Our model takes advantage of deep Convolutional layers for feature extraction, Bi-Directional LSTMs and Connectionist Temporal Classification (CTC) cost to label sequences. We also constructed a large dataset using several different fonts, and added noise to this artificial dataset in order to have a higher semblance to real noisy images. Trained on this dataset, our model outperforms state of the art Persian OCR models.

· **Iris: Persian Named Entity Tagger**
Status: Finished
A. Hosseini, B. Minaei.

In this project, firstly, different techniques, methods, design challenges, misconceptions and features which underline the development of an efficient and robust NER system are explored. Then, we developed Iris: the Farsi named entity recognizer agent, which is based on statistical distribution techniques and deep learning methods. Iris uses CRF on top of RNNs to label sequences of words considering Farsi words and sentences' syntax and structure. It is now actively integrated into/used in *Acas* (see below).

◇ **Natural Language Processing Projects**

· **Beatrix: Analysis of Persian Poetry using Language Models (2016)**
Status: Finished
A. Hosseini, B. Minaei.

In this project, we tried to create a language model that not only focuses on what is being written, but how it's written, which in poetry means that the model learns both orthographic and phonological features. This project was done for my computational AI course, and is implemented in Keras and Theano.

◇ **Work Experience**
Sayna Systems Co. (Summer 2015)
Mobile and web application developer

· **Acas**
Acas is a Social Network and News Analyser, using data mining, text similarity and graph discovery algorithms to analyse users' behaviours, and generate reports on mobile and web platforms. a joint work with V. Kharazi, A. Baghi.

◇ **Other Projects**

· **AbrIO (2016)**
A fast, soft real time, reliable and scalable message passing server as a service. AbrIO is developed in Scala, a scalable and functional programming language. This service is built on Akka, a framework (written in Scala) to build concurrent, resilient and distributed systems under Reactive manifesto. A joint work with K. Peymani, B. Heydari, H. Khalili, N. Mohammad Hasan.

· **Improving Context Switch Penalties in Linux Kernel Scheduling Mechanism (2015)**
In this project, I modified Linux scheduling mechanism, CFS (Completely Fair Scheduler), to add an amount of penalty to processes with high number of *context switches*. These modifications were made to affect the *vruntime* value (a virtual run-time value that is used for

scheduling in kernel) of a process. Data Structures in Kernel, such as linked lists and red-black trees were used in this project.

- **HeadBall (2014)**
  HeadBall is a single player 2D soccer game developed for Android devices using Unity game engine. This game has more than 1000 active users.

Awards and Honors

- ◇ Awarded **Outstanding Student** Certificate and Prize by Iran University of Science and Technology President – 2014 and 2015.
- ◇ Ranked **First** among the entire university computer engineering students – 2013 - present
- ◇ Ranked among the Top 1% Participant in Iran's National University Entrance Exam (+500000 Participants) – 2013
- ◇ Four years grant for undergraduate studies from the **Iranian National Foundation of Elites**.
- ◇ Online Course Statement of Accomplishment, **Machine Learning** Course by Prof. Andrew Ng, Stanford University through **Coursera**. – 2014
- ◇ Honored for outstanding performance at Computer Engineering Department Student Association (CESA). CESA is a student committee concerned with directing the department's extra-curriculum activities.

Skills

- ◇ **Programming and Scripting**: Python, Java, Scala, C,C++, Matlab, Octave, JavaScript, Swift
- ◇ **Web/DB Technologies**: HTML, CSS, Django, Flask, NodeJS, AngularJS, MYSQL, MongoDB, Redis
- ◇ **Tools and Frameworks**: Theano - TensorFlow - Blocks - Fuel - Keras
- ◇ **OS, Version Control and Issue Tracking**: Linux family - OS X - Git - Taiga - Jira
- ◇ **Hardware Description Languages/Micro-controller**: Verilog, AVR (Code-Vision)
- ◇ **Typesetting**: LaTeX, Markdown
- ◇ **Languages**: Persian (Native), English (Fluent)

  - **TOEFL iBT Score**: 111 (Reading: 27, Listening: 30, Speaking: 27, Writing: 27)

  - **GRE Scores**: Quantative: 169 , Verbal: 153 , Analytical Writing: 4

Teaching Experience and Talks

- ◇ **Teaching Assistant for Computer Architecture** (Fall 2015),
  instructor: Dr. Mohsen Soriani
- ◇ **Teaching Assistant for Discrete Mathematics** (Spring 2015),
  instructor: Dr. Hasan Naderi
- ◇ **Teaching Assistant for Advance Computer Programming** (Fall 2014),
  instructor: Dr. Adel Rahmani
- ◇ **Talk: "I want to talk to you", Language Models for poetry** (August 2016),
  The Center for Information and Language Processing, Ludwig Maximilian University of Munich, Germany
- ◇ **Workshop, PyCon 2016, Deep Learning for Natural Language Processing using Python** (Spring 2016),
  In this workshop the model in the paper Learning to Execute (a seq2seq model) was implemented using Keras. PyCon is the largest annual gathering for the community using and developing Python programming language.

References

◇ **Hinrich Schuetze**, Director, Center for Information and Language Processing, University of Munich (LMU). hs2016@cislmu.org

◇ **Ehsaneddin Asgari**, PhD Candidate in Applied Science and Technology, Life Language Processing and Deep Proteomics, University of California, Berkeley. asgari@berkeley.edu

◇ **Sascha Rothe**, Center for Information and Language Processing, University of Munich (LMU). sascha.rothe@live.de

◇ **Behrouz Minaei-Bidgoli**, Associate Professor, Computer Engineering Department, Iran University of Science and Technology (IUST). b_minaei@iust.ac.ir

◇ **Mohsen Soryani**, Associate Professor , Computer Engineering Department, Iran University of Science and Technology (IUST). soryani@iust.ac.ir

◇ **Hasan Naderi**, Assistant Professor , Computer Engineering Department, Iran University of Science and Technology (IUST). naderi@iust.ac.ir

◇ **Adel Rahmani**, Assistant Professor , Computer Engineering Department, Iran University of Science and Technology (IUST). Rahmani@iust.ac.ir

◇ **Sayna Systems Co**. http://systemnegar.ir

◇ **Ashkan Sadeghi**, Sayna Systems Company Head Manager. sadeghi@systemnegar.ir

◇ **Vahid Kharazi**, Developer and Shareholder, Sayna Systems Company. vahid@kharazi.net

◇ **Alireza Nourian**, Machine Learning and Data Mingin Lab, Computer Engineering Dept, Iran University of Science and Technology (IUST). az.nourian@gmail.com