

# Alireza Mohammadshahi

Lausanne, Switzerland

✉ alireza.mohammadshahi@idiap.ch

🌐 www.idiap.ch/ amohammadshahi/

📧 alirezamshi

📱 Alireza-mohammadshahi

☎ +41 78 969 5566

## Education

### École Polytechnique Fédérale de Lausanne

Lausanne, Switzerland

DIRECT PH.D. IN COMMUNICATION AND COMPUTER SCIENCE (EDIC)

Sep 2018 - PRESENT

### IDIAP Research Institute

Martigny, Switzerland

RESEARCH ASSISTANT OF NATURAL LANGUAGE UNDERSTANDING GROUP

Feb 2019 - PRESENT

### Sharif University of Technology

Tehran, Iran

B.SC. IN ELECTRICAL ENGINEERING; COMMUNICATION, MINOR IN COMPUTER SCIENCE

Sep 2013 - Aug 2018

- GPA: 18.98/20
- last 4 semesters' GPA: 19.78/20 (4.0/4.0)

## Research Interest

Natural Language Processing

Machine learning and data science

Deep learning

Design of algorithms

## Honors & Awards

<b>Member</b> Natural language understanding group at IDIAP research institute	2019
<b>Offer</b> EDIC PhD program fellowship by EPFL	2018
<b>Offer</b> EE PhD program fellowship by UCLA and USC	2018
<b>Offer</b> EE Master program fellowship by UofT and UBC	2018
<b>Rank 18th</b> out of +250,000 undergraduate applicants in universities entrance exam for B.Sc. degree	2013
<b>Member</b> of national elite foundation of Iran	2013 - PRESENT
<b>Best Paper Award</b> Swiss Machine Learning Day	2019
<b>Offer</b> Google Travel Scholarship for 3rd Google NLP Summit	2019
<b>Offer</b> Amazon Travel Scholarship for the EMNLP conference	2019
<b>Rank 3rd</b> The first human-robot tandem race	2019
<b>Rank 1st</b> of kaggle EPFL Machine Learning competition among 201 teams	2018
<b>Finalist</b> of 10th ICC with proposing MyLinguist prototype (mylinguist.ch)	2021

## Volunteer Experience

<b>Member</b> of competitive programming association-PolyProg (polyprog.epfl.ch)	2019-PRESENT
<b>Member</b> of industrial committee in Enerwat (enerwat.sharif.ir)	2015

# Research Experiences

---

## Syntax-Aware Graph-to-Graph Transformer for Semantic Role Labelling

Dr. James Henderson

SUBMITTED TO EMNLP 2021

- The goal of semantic role labelling (SRL) is to recognise the predicate-argument structure of a sentence. Recent models have shown that syntactic information can enhance the SRL performance, but other syntax-agnostic approaches achieved reasonable performance. The best way to encode syntactic information for the SRL task is still an open question. In this paper, we propose the Syntax-aware Graph-to-Graph Transformer (SynG2G-Tr) architecture, which encodes the syntactic structure with a novel way to input graph relations as embeddings directly into the self-attention mechanism of Transformer. This approach adds a soft bias towards attention patterns that follow the syntactic structure but also allows the model to use this information to learn alternative patterns. We evaluate our model on both dependency-based and span-based SRL datasets, and outperform all previous syntax-aware and syntax-agnostic models in both in-domain and out-of-domain settings, on the CoNLL 2005 and CoNLL 2009 datasets. Our architecture is general and can be applied to encode any graph information for a desired downstream task.
- Link: <https://arxiv.org/abs/2104.07704>

## Recursive Non-Autoregressive Graph-to-Graph Transformer for Dependency Parsing with Iterative Refinement

Dr. James Henderson

ACCEPTED TO TACL JOURNAL, PRESENTED TO EACL 2021

- We propose the Recursive Non-autoregressive Graph-to-graph Transformer architecture (RNG-Tr) for the iterative refinement of arbitrary graphs through the recursive application of a non-autoregressive Graph-to-Graph Transformer and apply it to syntactic dependency parsing. The Graph-to-Graph Transformer architecture of mohammadshahi2019graphtograph has previously been used for autoregressive graph prediction, but here we use it to predict all edges of the graph independently, conditioned on a previous prediction of the same graph. We demonstrate the power and effectiveness of RNG-Tr on several dependency corpora, using a refinement model pre-trained with BERT. We also introduce Dependency BERT (DepBERT), a non-recursive parser similar to our refinement model. RNG-Tr is able to improve the accuracy of a variety of initial parsers on 13 languages from the Universal Dependencies Treebanks and the English and Chinese Penn Treebanks, even improving over the new state-of-the-art results achieved by DepBERT, significantly improving the state-of-the-art for all corpora tested.
- Link: [https://direct.mit.edu/tacl/article/doi/10.1162/tacl\\_a\\_00358/97778/Recursive-Non-Autoregressive-Graph-to-Graph](https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00358/97778/Recursive-Non-Autoregressive-Graph-to-Graph)

## Graph-to-Graph Transformer for Transition-based Dependency Parsing

Dr. James Henderson

ACCEPTED TO EMNLP 2020

- We propose the Graph2Graph Transformer architecture for conditioning on and predicting arbitrary graphs, and apply it to the challenging task of transition-based dependency parsing. After proposing a novel Transformer model of transition-based dependency parsing, we show that the proposed mechanisms for graph input and graph output result in significant improvements over this strong baseline, especially with BERT pre-training. Both the novel Transformer and the Graph2Graph Transformer parsers significantly outperform the state-of-the-art in transition-based dependency parsing on both English Penn Treebank, and 11 languages of Universal Dependencies Treebanks. Graph2Graph Transformer can be integrated with many previous structured prediction methods, making it easy to apply to a wide range of NLP tasks.
- Link: <https://www.aclweb.org/anthology/2020.findings-emnlp.294/>

## Aligning Multilingual Word Embeddings for Cross-Modal Retrieval Task

Prof. Aberer & Dr. Remi Lebre

ACCEPTED TO FEVER AND LANTERN AT EMNLP 2019

- In this paper, we propose a new approach to learn multimodal multilingual embeddings for matching images and their relevant captions in two languages. We combine two existing objective functions to make images and captions close in a joint embedding space while adapting the alignment of word embeddings between existing languages in our model. We show that our approach enables better generalization, achieving state-of-the-art performance in text-to-image and image-to-text retrieval task, and caption-caption similarity task. Two multimodal multilingual datasets are used for evaluation: Multi30k with German and English captions and Microsoft-COCO with English and Japanese captions.
- The dataset is Multi30K: Multilingual English-German Image Descriptions which contains five captions per image in both German and English languages that are not the translation of each other.
- The results has been shown that we can almost keep the alignment while making the images and captions closer, and the result is between freezing the word embedding, and not having translators to align word embeddings
- As an great application, I have run the model on Tamedia photo archiv (French+German) dataset which is a news dataset that images and their description are weakly related together.
- Link: <https://www.aclweb.org/anthology/D19-6605/>

# Selected Academic Projects

---

## Twittly Unsupervised Multilingual text classifier

Prof. Karl. Aberer & Martin Jaggi &  
Rüdiger Urbanke

### FINDING A MULTILINGUAL UNSUPERVISED CLASSIFIER ON SWISS SOCIAL MEDIA DATASET

- In this project, I have a social media dataset (twitter, Instagram and web) of people in Switzerland which don't have any labels. The main goal here is to classify it to three classes: 1. Soccer in Switzerland 2. Ice-Hockey in Switzerland 3. None.
- Since the dataset doesn't have any labels, I use some initial and simple signs to find soccer and ice-hockey classes, however, in the training process, I omit these signs, so the network should learn based on other words of a sentence. The data in training mode is the combination of initial labeled soccer and ice-hockey sentences with the same amount of unlabeled data from the corpus. Also, I add initial labels for soccer and ice-hockey not in Switzerland in "None" class label.
- Next, I use an active learning approach combined with semi-supervised algorithm. Here, I run the model for three different initialization. After the end of each training process, I pick texts that all my ensembles predict them with high probability and there are not any initial labels in these texts. Then, I ask the user to label them for the model. Finally, I add them to the model, and run the model again. Here, there is a significant and great amount of information in new texts, so the model can take the advantage of these new labeled data.
- The second goal of the project is to have a multilingual model which works on English, German and French texts which are most popular in Switzerland. So, for this goal, I train a language identifier on the corpus of three languages to find a model which detects the language of a sentence.

## EPFL machine learning competition

Prof. Martin Jaggi & Rüdiger  
Urbanke

### KAGGLE COMPETITION WITH 200 TEAMS

- This project is about my approach for the Higgs boson machine learning challenge. The Higgs boson is a fundamental particle and the final building block of the standard model of particle physics. This fundamental particle, which is also called God particle, was hypothesized many years ago, but its existence was only confirmed very recently at 2013 through experiments in the LHC at CERN. In this challenge, I should distinguish signal events from background events using 30 given features in each data sample. To put it in a nutshell, the training and test data include two types of features: primitive features and derived features which are derived by physicists to help us classify the data with a better accuracy. Moreover, since the ratio of signal to background events is really small, in the train set there are some simulated data using simulation in order to improve the final results precision and recall for the competitors. Besides, we are just allowed to use Numpy library for implementation.
- My approach for solving this problem involves feature expansion within the linear model framework (polynomial, trigonometric functions, etc.) and several new derived features, and methods for dealing with missing values. I also try categorical splitting. In the end, in order to improve the classification accuracy, instead of using a single neuron for training, I use a neural network with only one hidden layer to improve the quality of extracted features and my methods. At the end, I use some ensemble learning methods to improve the final result.

## B.Sc.Project

Prof. Babak Hosein Khalaj

### DEPLOYMENT OF A 5G CLOUD-NATIVE END-TO-END MOBILE NETWORK PLATFORM (OAI)

- This refers to a pioneer project on the deployment of an end-to-end LTE/5G test and development platform which is the first of its kind in Iran and has provided a unique opportunity for Mobile Virtual Network Operators (MVNOs), vendors and other customers like service providers to test and develop 4G/5G-related features at the Radio Access Network (RAN) as well as Evolved Packet Core (EPC) side.

## Mobile Network Projects

Prof. Babak Hossein Khalaj

### EXPERIMENTS ON OPENAIR PLATFORM

- Within my B.Sc. project, I have contributed to development of novel features of 4G/5G mobile communication networks like network slicing and RAN sharing based on open source network emulation software tools in C/C++.
- In this context, I have been involved in designing the Proof-of-Concept (PoC) demonstration platforms for Mobile Telecommunication Company of Iran (Iran's first and largest mobile operator) to evaluate over the network low-latency communications solutions for time-critical applications like traffic-safety and car controlling which were shown in Iran Telecom 2017
- The OpenAirInterface Software Alliance (OSA) is a non-profit consortium fostering a community of industrial as well as academic contributors for open source software and hardware development for the core network (EPC), access network and user equipment (EUTRAN) of 3GPP cellular networks

## Computer Structure and LAB

Prof. Saeid Bagheri

### FULL DESIGN OF A ACCUMULATOR

- A full engineering accumulator was designed, simulated and implemented on PCB

## Signal Processing LAB

Prof. Farokh Marvasti

### SIGNAL PROCESSING USING IMAT METHOD FOR SPARSE SIGNALS

- using Iterative Method with Adaptive Thresholding algorithm for sparse signal reconstruction

# Programming Languages & Applications

---

**General:** Microsoft Office, Linux, LaTeX

**Programming Languages:** MATLAB, OpenCV, C/C++, Python

**Applications:** Pytoch, Tensorflow, OpenAir, Altium Designer, Proteus, AutoCAD, Hspice, AVR

## Teaching Experience

---

**Deep Learning for Natural Language Processing** Dr. Henderson

**Natural Language Processing for MasterAI of IDIAP** Dr. Henderson

**Data Structure** Prof. Khazaei

**Computer Structure and LAB** Prof. Bagheri

**Engineering Mathematics** Prof. Aghajan

**Signals and Systems** Prof. Behroozi

**Logical Circuits** Prof. Mohammadzade

**Communication Systems** Prof. Behroozi

**Principles of Electrical Engineering** Prof. Fotowat-Ahmady

**Circuit Theory** Prof. Mirmohseni

**Analog Circuits** Prof. Mohammadzadeh

**Circuit Theory** Prof. Sharifkhani

**Principles of Electronics** Prof. Fakharzadeh

## Workshop & Conference Experience

---

<b>Presentation</b> European Chapter of the Association for Computational Linguistics (EACL)	2021
<b>Program Committee</b> TextGraphs:15th Workshop on Graph-Based NLP-NAACL	2021
<b>Program Committee</b> Association for Computational Linguistics(ACL)	2021
<b>Program Committee</b> European Chapter of the Association for Computational Linguistics (EACL)	2021
<b>Program Committee</b> American Journal of Computer Science and Technology	2020
<b>Program Committee</b> Swiss Machine Learning Days	2019
<b>Poster Presentation</b> Conference on Empirical Methods in Natural Language Processing	2019
<b>Poster Presentation</b> 3rd Google Natural Language Processing Summit at Google Zurich	2019
<b>Poster Presentation</b> Swiss Machine Learning Day	2019
<b>Member</b> the Research team of 5G cloud-native end-to-end platform (OpenAir) in MWTS	2017
<b>Member</b> the Research team of Mobile Telecommunication Company of Iran in Iran Telecom	2017
<b>Teaching Assistant</b> Robotics Workshop of 4th Sharif Open Robotic Competition (Sharif Cup)	2014

## Selected Courses

---

<b>Distributed Information Systems (Graduate)</b> Prof. Aberer	5.5/6.0
<b>Machine Learning EPFL (Graduate)</b> Prof. Martin Jaggi & Rudrigger Urbanke	5.5/6.0
<b>Machine Learning (Coursera)</b> Prof. Andrew NG	Auditing
<b>Data and Artificial Intelligence for transportation(Graduate)</b> Prof. Alahi & Sven Kreiss	5.75/6.0
<b>Design and analysis of Algorithms</b> Prof. Khazaei	20.0/20
<b>Design of Database</b> Prof. Boumeri	20.0/20
<b>Linear Algebra</b> Prof. Safdari	20.0/20
<b>Data Structure</b> Prof. Khazaei	20.0/20
<b>Stochastic Process</b> Prof. Mirsadeghi	19.5/20
<b>Statistics &amp; Applications</b> Prof. Mirsadeghi	20.0/20
<b>C &amp; C++ Programming</b> Prof. Jafari Nezhad	20.0/20
<b>Engineering Probability and Statistics</b> Prof. Mirhohseni	19.6/20
<b>Discrete Signal Processing</b> Prof. Shamsollahi	19.9/20
<b>Signals and Systems</b> Prof. Aghajan	20.0/20

## Language

---

**Persian:** Native

**English:** Fluent

**French:** Basic

## References

---

### **Dr. James Henderson**

HEAD OF NATURAL LANGUAGE UNDERSTANDING GROUP AT IDIAP

- Email: james.henderson@idiap.ch
- Tel: (+41) 27 72 17729
- Addr: Room 107.1, Centre du Parc, Rue Marconi 19, PO Box 592 ,CH - 1920 Martigny Switzerland

### **Prof. Karl Aberer**

PROFESSOR, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Email: karl.aberer@epfl.ch
- Tel: (+41) 21 69 34679
- Addr: EPFL IC IINFCOM LSIR , BC 108 (Bâtiment BC), Station 14, CH-1015 Lausanne