

SHIVA TASLIMIPOOR

Natural Language and Information Processing Research Group ◊ University of Cambridge

✉ sh.taslimi@gmail.com ◊ 🌐 shivaat.github.io ◊ 📄 https://github.com/shivaat

RESEARCH INTERESTS

Natural Language Processing, Multiword Expressions, Semantic Similarity, Figurative Language, Automatic Language Teaching and Assessment, Machine Learning and Deep Learning.

EDUCATION

PhD in Natural Language Processing, University of Wolverhampton, UK 2014-2018

. Thesis: Automatic Identification and Translation of Multiword Expressions

Supervisor: Prof. Ruslan Mitkov

M. Sc. in Artificial Intelligence, Shiraz University 2009-2012

. Thesis: Automatic Investigation of the Semantics of Persian Compound Verbs

Supervisor: Dr. Afsaneh Fazly

B. Sc. in Computer Science and Engineering, Shiraz University 2004-2008

. Thesis: Detecting Openness of Pistachio Nuts Using Image Classification Methods

supervisor: Prof. Zohreh Azimifar

TECHNICAL SKILLS

Programming Languages

Python, MATLAB, Java, C/C++ (prior experience)

NLP Tools

SpaCy, Gensim, NLTK, SketchEngine, WordNet

ML Tools

Keras (with Tensorflow backend), PyTorch, Scikit-learn

ML Expertise

Deep learning, representation learning, classification, regression, tagging, structured prediction and statistical analysis

WORK EXPERIENCE

Postdoctoral Research Associate

October 2019 - Present

Natural Language and Information Processing Research Group, University of Cambridge

Postdoctoral Research Associate

October 2018 - September 2019

Research Group in Computational Linguistics, University of Wolverhampton

Research Assistant

2014 - 2018

Research Group in Computational Linguistics, University of Wolverhampton

. Research towards PhD (refer to Section Selected Projects)

Researcher & Developer

2013 - 2014

Computer Vision and Pattern Recognition Lab, Shiraz University

. Stereopsis problems for traffic analysis project for the municipality of the city of Shiraz

Instructor

Fall 2012 & Spring 2013

Jahrom Universities, Iran

. **Discrete Mathematics**

. **Artificial Intelligence**

. **Data Structures**

. **Intro. to Formal Languages & Automata**

. **Supervisory of B.Sc projects**

. VSMs in Semantics, Document Categorization, Text Summerization, Spam Filtering, MT, etc.

SELECTED PROJECTS

Multiword Expression (MWE) identification

July 2014 - Present

- proposed a sequence labelling model that combined Graph Convolutional Neural Networks with Self-Attention improving state-of-the-art on a standard multi-lingual dataset
- designed language-independent neural sequence labelling methods to identify MWEs (implemented in **Keras** and **ranked 1st** in Parseme 2018 shared task on “automatic identification of verbal MWEs”)
- proposed a classification approach to disambiguate idiomatic expressions (using **gensim** & **sklearn**)
- re-examined standard evaluation for MWEs & proposed a more accurate cross-validation method
- utilised features derived from eye tracking data to improve MWE identification

Extracting Translation Equivalents (TE)

July 2014 - June 2018

- proposed an original semi-supervised **word2vec**-based method to extract TEs for MWEs
- developed and exploited comparable corpora for low resource languages

Irony Detection

October 2017 - January 2018

- developed irony detection system using sentiment and embedding features, which **ranked 3rd** in SemEval 2018 shared task 3 “Irony detection in English tweets”

Semantic Discrimination

November 2017 - January 2018

- developed a model using SketchEngine, ConceptNet, embeddings & google Ngrams, which **ranked 4th** (joint) in SemEval 2018 shared task 10 “Capturing Discriminative Attributes”

ACADEMIC ACHIEVEMENTS

Best paper award in RANLP 2017

Best paper award by a young researcher in Europhras2015

PARSEME grant for the 2nd PARSEME training school, Summer 2016

Ph.D. Scholarship from the University of Wolverhampton, UK, (2014-2017)

COMMUNITY SERVICE

Editorial Assistant

2014 - 2019

Cambridge Journal of Natural Language Engineering

Organising Committee Member

- Summer School on Deep Learning in NLP (DLinNLP), Varna, Bulgaria *29-30 August 2019*
 . I was also an instructor for a practical coding session on sequence tagging & multitask learning
- EUROPHRAS 2017, London, UK *13-14 November 2017*
- Student Research Workshop, held in conjunction with EUROPHRAS 2017, London, UK.

Programme Committee Member & Reviewer

ACL 2020, OxBridge 2020 WinAI, RANLP (2019, 2017 & 2015), EUROPHRAS 2019, JNLE, SemEval 2019 & 2018, SLSP 2017, CICLing 2017, AISP 2012

SELECTED PUBLICATIONS

Verbal Multiword Expressions for Identification of Metaphor, O. Rohanian, M. Rei, S. Taslimipoor, L. A. Ha, *accepted for publication in ACL*, 2020.

Incorporating Multiword Expressions in Phrase Complexity Estimation, S. Gooding, S. Taslimipoor, E. Kochmar, *accepted for publication in READI workshop at LREC*, 2020.

Bridging the Gap: Attending to Discontinuity in Identification of Multiword Expressions, O. Rohanian, S. Taslimipoor, S. Kouchaki, L. A. Ha, R. Mitkov, In *NAACL-HLT*, 2019. (Equal collaboration with the first author)

Identification of Multiword Expressions: A Fresh Look at Modelling and Evaluation, S. Taslimipoor, O. Rohanian, R. Mitkov, A. Fazly. In *Multiword expressions at length and in depth. Extended papers from the MWE 2017 workshop*, Language Science Press, 2018.

Wolves at SemEval-2018 Task 10: Semantic Discrimination based on Knowledge and Association, S. Taslimipoor, O. Rohanian, L. A. Ha, G. Corpas, R. Mitkov. In *SemEval*, 2018.

Investigating the Opacity of Verb-noun Multiword Expression Usages in Context, S. Taslimipoor, O. Rohanian, R. Mitkov, and A. Fazly, in *proceedings MWE@EACL*, 2017.

Bilingual Contexts from Comparable Corpora to Mine for Translations of Collocations, S. Taslimipoor, R. Mitkov, G. Corpas, A. Fazly, in *CICLing*, 2016.

Using Noun Similarity to Adapt an Acceptability Measure for Persian Light Verb Constructions, S. Taslimipoor, A. Fazly and A. Hamzeh: In *LREC'12*, 2012.