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

WORK EXPERIENCE

Postdoctoral Research Associate *October 2019 - Present*
Natural Language and Information Processing Research Group, University of Cambridge
· Cloze Test Generation
· Multiword Expressions in Educational Applications

Postdoctoral Research Associate *October 2018 - September 2019*
Research Group in Computational Linguistics, University of Wolverhampton
· Question-answer generation
· Predicting item characteristics for Multiple Choice Question Answering

Research Assistant *2014 - 2018*
Research Group in Computational Linguistics, University of Wolverhampton

Researcher & Developer *2013 - 2014*
Computer Vision and Pattern Recognition Lab, Shiraz University

Instructor *Fall 2012 & Spring 2013*
Jahrom Universities, Iran
· **Discrete Mathematics**
· **Data Structures**
. **Artificial Intelligence**
. **Intro. to Formal Languages & Automata**

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

SELECTED PROJECTS

Multiword Expression (MWE) identification *July 2014 - Present*

- designed a multi-task learning system to identify MWEs (**ranked 1st** in Parseme 2020 shared task on "semi-supervised identification of verbal multiword expressions" [code]
- 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. [code]
- designed language-independent neural sequence labelling methods to identify MWEs (**ranked 1st** in Parseme 2018 shared task on "automatic identification of verbal MWEs" [code]
- re-examined standard evaluation for MWEs & proposed a more accurate cross-validation method for evaluating a classification approach to disambiguate idiomatic expressions [code]
- 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 [resource]

Irony Detection *October 2017 - January 2018*
- using sentiment and embedding features, system **ranked 3rd** in SemEval 2018 shared task 3 [code]

Semantic Discrimination *November 2017 January 2018*
- using SketchEngine, ConceptNet, embeddings & google Ngrams, system **ranked 4th** in SemEval 2018 shared task 10 [code]

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

Standing Committee *2020 - 2022*
SIGLEX-MWE section of the Special Interest Group on the Lexicon of the Association for Computational Linguistics (ACL)

Editorial Assistant *2014 - 2019*
Cambridge Journal of Natural Language Engineering

Organising Committee Member
- 17th Workshop on Multiword Expressions (MWE 2021) *6 August 2021*
- Summer School on Deep Learning in NLP (DLinNLP), Varna, Bulgaria *29-30 August 2019*
- EUROPHRAS 2017, London, UK *13-14 November 2017*
- Student Research Workshop, held in conjunction with EUROPHRAS 2017, London, UK.

Programme Committee Member & Reviewer
JNLE, EMNLP (2021, 2020), ACL (2021, 2020), EACL 2021, RANLP (2021, 2019, 2017, 2015), CoLing 2020, OxBridge 2020 WinAI, EUROPHRAS 2019, SemEval (2019 & 2018), SLSP 2017, CICLing 2017

TALKS AND LECTURES

Invited talk, University of Essex NLIP weekly seminar *13 May 2020*

Instructor, practical coding session on sequence tagging & multitask learning *30 August 2019*
Summer School on Deep Learning in NLP (DLinNLP), Varna, Bulgaria

Guest Lecturer, 7LN001 Python Programming, University of Wolverhampton *Fall 2016*

SELECTED PUBLICATIONS

MTLB-STRUCT @Parseme 2020: Capturing Unseen Multiword Expressions Using Multi-task Learning and Pre-trained Masked Language Models, S. Taslimipoor, S. Bahaadini, E. Kochmar. In *proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*, pp. 142148. 2020. (1st ranked system)

Verbal Multiword Expressions for Identification of Metaphor, O. Rohanian, M. Rei, S. Taslimipoor, L. A. Ha. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2890-2895. ACL, 2020.

Incorporating Multiword Expressions in Phrase Complexity Estimation, S. Gooding, S. Taslimipoor, E. Kochmar. In *Proceedings of the 1st Workshop on Tools and Resources to Empower People with REAding DIfficulties (READI)*, 2020.

What matters more: The size of the corpora or their quality? The case of automatic translation of multiword expressions using comparable corpora, R. Mitkov, S. Taslimipoor. Book chapter in Computational Phraseology. pp. 178187. John Benjamins, 2020.

Bridging the Gap: Attending to Discontinuity in Identification of Multiword Expressions, O. Rohanian, S. Taslimipoor, S. Kouchaki, L. A. Ha, R. Mitkov. In *proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*, 2019. (Equal collaboration with the first author)

Cross-lingual Transfer Learning and Multitask Learning for Capturing Multiword Expressions, S. Taslimipoor, O. Rohanian, L. A. Ha. In *proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, 2019.

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

SHOMA at Parseme Shared Task on Automatic Identification of VMWEs: Neural Multiword Expression Tagging with High Generalisation, S. Taslimipoor and O. Rohanian. In arXiv preprint 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.

WLV at SemEval-2018 Task 3: Dissecting Tweets in Search of Irony, O. Rohanian, S. Taslimipoor, R. Evans, R. Mitkov. In *SemEval*, 2018.

Bilingual Contexts from Comparable Corpora to Mine for Translations of Collocations, S. Taslimipoor, R. Mitkov, G. Corpas, A. Fazly. In *CICLing 2016. Lecture Notes in Computer Science, vol 9624, pp. 115-126.* Springer, 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.