

Mohammad Javad Hosseini

Google Research, UK

Email: javadh@google.com

RESEARCH INTERESTS

- ◇ Natural Language Processing: Semantic Analysis, Knowledge Representation and Reasoning
- ◇ Machine Learning: Deep Learning, Structured Prediction, Graphical Models

EDUCATION

PhD, Informatics **2016-2020**

University of Edinburgh, UK

Affiliated with The Alan Turing Institute

- ◇ Supervisors: Mark Steedman (Principal) and Shay Cohen
- Thesis: Unsupervised Learning of Relational Entailment Graphs from Text

PhD student, Computer Science **2013-2016**

University of Washington, U.S. (Moved to Edinburgh due to family reasons. Obtained an MSc in 2015.)

GPA: 3.86/4

- ◇ Supervisor (2014-2016): Su-In Lee
 - Structured Sparsity for Gaussian Graphical Models
 - Structured Prediction for Detecting Multi-Word Expressions and Supersenses
- ◇ Supervisors (2013-2014): Oren Etzioni and Hanna Hajishirzi
 - NLP Methods for Solving 4th Grade Math Word Problems

MSc, Computer Software Engineering **2010-2012**

Sharif University of Technology, Iran

GPA: 19.45/20 (4/4)

- ◇ Supervisor: Hamid Beigy
 - Thesis: A Semi-Supervised Ensemble Learning Algorithm for Non-Stationary Data Streams Classification

BSc, Computer Software Engineering **2006-2010**

Sharif University of Technology, Iran

GPA: 18.76/20 (3.98/4)

WORK EXPERIENCE

- ◇ **Researcher at Google Research**, London, UK **June 2021-now**
Natural Language Understanding team
- ◇ **Postdoctoral Research Associate**, Edinburgh, UK **Oct 2020-June 2021**
School of Informatics, University of Edinburgh
Supervisor: Mark Steedman
Using Pre-Trained Language Models to Improve Entailment Graphs
- ◇ **Research Intern at Google**, Mountain View, U.S. **Sep 2019-Dec 2019**
Natural Language Understanding team: Coherent Resolution of Location Entities in Wikipedia Pages
- ◇ **Research Intern at Apple**, Cambridge, UK **Jun 2019-Sep 2019**
Siri Understanding team: Semantic Parsing for Conversational Language Understanding
- ◇ **Research Intern at Yahoo**, Sunnyvale, U.S. **Jun 2014-Sep 2014**
Ads & Data team: Modeling Users' Purchase Patterns based on Clickstream Data
- ◇ **R&D Engineer at Bayan**, Tehran, Iran **Mar 2012-Sep 2013**
Building a high precision web spam detector system and participating in the stemmer, language detection, and query clustering modules (part-time).

HONORS AND AWARDS

- ◇ Alan Turing Institute Doctoral Studentship (2016-2020)
- ◇ **1st Rank** in cumulative GPA among 70 BSc Computer Engineering students of the department, Class of 2010, Sharif University of Technology, Tehran, Iran.
- ◇ **2nd Rank (Silver Medal)**, National Scientific Olympiad in Computer Engineering, Summer 2010, Tehran, Iran

PUBLICATIONS

- ◇ **Mohammad Javad Hosseini**, Shay B. Cohen, Mark Johnson, and Mark Steedman. Open-Domain Contextual Link Prediction and its Complementarity with Entailment Graphs. To Appear in *Findings of the Conference on Empirical Methods in Natural Language Processing (EMNLP Findings, Long Papers)*, 2021.
- ◇ Nick McKenna, Liane Guillou, **Mohammad Javad Hosseini**, Sander Bijl de Vroe, and Mark Steedman. Multivalent Entailment Graphs for Question Answering. To Appear in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP, Long Papers)*, 2021.
- ◇ Sayali Kulkarni, Shailee Jain, **Mohammad Javad Hosseini**, Jason Baldrige, Eugene Ie, and Li Zhang. Spatial Language Representation with Multi-Level Geocoding. *arXiv preprint arXiv:2008.09236*, 2020.
- ◇ Liane Guillou, Sander Bijl De Vroe, **Mohammad Javad Hosseini**, Mark Johnson, and Mark Steedman. Incorporating Temporal Information in Entailment Graph Mining. In *Proceedings of the Graph-based Methods for Natural Language Processing (TextGraphs)*, 2020.
- ◇ **Mohammad Javad Hosseini**, Shay B. Cohen, Mark Johnson, and Mark Steedman. Duality of Link Prediction and Entailment Graph Induction. In *Proceedings of Association for Computational Linguistics (ACL, Long Papers)*, 2019.
- ◇ **Mohammad Javad Hosseini**, Nathanael Chambers, Siva Reddy, Xavier R. Holt, Shay B. Cohen, Mark Johnson, and Mark Steedman. Learning Typed Entailment Graphs with Global Soft Constraints. *Transactions of the Association for Computational Linguistics (TACL)*, 2018.
- ◇ **Mohammad Javad Hosseini** and Su-In Lee. Learning Sparse Gaussian Graphical Models with Overlapping Blocks. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2016.
- ◇ **Mohammad Javad Hosseini**, Noah A. Smith, and Su-In Lee. UW-CSE: Detecting Multiword Expressions and Supersenses using Double-Chain Conditional Random Fields. In *Proceedings of the Semantic Evaluations Workshop, in conjunction with North American Chapter of the Association for Computational Linguistics (SemEval NAACL workshops)*, 2016.
- ◇ **Mohammad Javad Hosseini**, Su-In Lee. Learning Gaussian Graphical Models with Overlapping Blocks. In *Proceedings of the Machine Learning for Computational Biology Workshop, in conjunction with Advances in Neural Information Processing Systems (NeurIPS workshops)*, 2015.
- ◇ **Mohammad Javad Hosseini**, Hannaneh Hajishirzi, Oren Etzioni, Nate Kushman. Learning to Solve Arithmetic Word Problems with Verb Categorization. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP, Long Papers)*, 2014.

OLDER PUBLICATIONS

- ◇ Roohallah Alizadehsani, **Mohammad Javad Hosseini** et al. Non-invasive Detection of Coronary Artery Disease in High-risk Patients based on the Stenosis Prediction of Separate Coronary Arteries. *Computer Methods and Programs in BioMedicine*, 2018.
- ◇ **Mohammad Javad Hosseini**, Ameneh Gholipour, and Hamid Beigy. An Ensemble of Cluster Based Classifiers for Semi Supervised Classification of Non-stationary Data Streams. *Knowledge and Information Systems*, 2016.
- ◇ Roohallah Alizadehsani, Mohammad Hossein Zangoeei, **Mohammad Javad Hosseini** et al. Coronary Artery Disease Detection using Computational Intelligence Methods. *Knowledge-Based Systems*, 2016.

- ◇ Fereshte Khani, **Mohammad Javad Hosseini**, Ahmad Ali Abin, and Hamid Beigy. An Algorithm for Discovering Clusters of Different Densities or Shapes in Noisy Data Sets. In *Proceedings of ACM Symposium on Applied Computing*, 2013.
- ◇ Ameneh Gholipour, **Mohammad Javad Hosseini**, and Hamid Beigy. An Adaptive Regression Tree for Non-stationary Data Streams. In *Proceedings of ACM Symposium on Applied Computing*, 2013.
- ◇ Roohallah Alizadehsani, Jafar Habibi, **Mohammad Javad Hosseini** et al. A Data Mining Approach for Diagnosis of Coronary Artery Disease. *Computer Methods and Programs in BioMedicine*, 2013.
- ◇ **Mohammad Javad Hosseini**, Zahra Ahmadi, and Hamid Beigy. Using a Classifier Pool in Accuracy based Tracking of Recurring Concepts in Data Stream Classification. *Evolving Systems*, 2013.
- ◇ **Mohammad Javad Hosseini**, Zahra Ahmadi, and Hamid Beigy. New Management Operations on Classifiers Pool to Track Recurring Concepts. In *Proceedings of the International Conference on Data Warehousing and Knowl Discovery*, 2012.
- ◇ **Mohammad Javad Hosseini**, Zahra Ahmadi, and Hamid Beigy. Pool and Accuracy Based Stream Classification: A new ensemble algorithm on data stream classification using recurring concepts detection. In *Proceedings of the Workshop on Handling Concept Drift in Adaptive Information Systems, in conjunction with the 11th International Conference on Data Mining*, 2011.

SKILLS

- ◇ Programming Languages: Java, python, R, MATLAB, C++