ALON JACOVI

Curriculum Vitae

Givatayim, Israel ⋄ (+972) 050-9297995 ⋄ alonjacovi@gmail.com

EDUCATION

Ph.D in Natural Language Processing and Machine Learning *2019–Present*
Bar Ilan University
Advisor: Prof. Yoav Goldberg
Research Topics: Explainable Artificial Intelligence, Learning from Imperfect Supervision

M.Sc in Natural Language Processing and Machine Learning *2017–2019*
Bar Ilan University
Advisor: Prof. Yoav Goldberg
Final Grade: 95.85
Dissertation Topic: Understanding Convolutional Neural Networks for Text Classification

B.Sc in Computer Science *2014–2017*
Bar Ilan University
Graduated *cum laude* with final grade 95.2.
Final Project: Deep Reinforcement Learning Agent for the AI-BIRDS Angry Birds Competition

EXPERIENCE

Research Intern at the Allen Institue for Artificial Intelligence, USA *Fall 2020 (current)*
MOSAIC Team
Advisor: Dr. Swabha Swayamdipta
Topic: Contrastive Explanations of Neural Classifiers in NLP.

Student Researcher at IBM Research, Israel *2016–2020*

- Conversation and Language Team
  Work on improving graph-based virtual dialogue assistants, based on in-production conversation logs that were escalated to a human agent. Work published in KDD Converse 2020.
- Machine Learning Technologies Team
  Work on integration of non-differentiable systems in neural pipelines (e.g., knowledge bases and program interpreters in semantic parsing) for end-to-end learning. Work published in ICLR 2019.

Research Intern at RIKEN, Japan *Spring 2019*
Imperfect Information Learning Team
Advisors: Dr. Gang Niu and Prof. Masashi Sugiyama
Topic: Scalable Evaluation and Improvement of Document Set Expansion via Neural Positive-Unlabeled Learning.

PUBLICATIONS

Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI.
Alon Jacovi, Ana Marasović, Tim Miller, Yoav Goldberg.
In ACM FAccT 2021.

Scalable Evaluation and Improvement of Document Set Expansion via Neural Positive-Unlabeled Learning
Alon Jacovi, Gang Niu, Yoav Goldberg, Masashi Sugiyama.
In EACL 2021.

Aligning Faithful Interpretations with their Social Attribution.
Alon Jacovi, Yoav Goldberg.
In TACL 2020.

Amnesic Probing: Behavioral Explanations with Amnesic Counterfactuals.
Yanai Elazar, Shauli Ravfogel, Alon Jacovi, Yoav Goldberg.
In TACL 2020.

Exposing Shallow Heuristics of Relation Extraction Models with Challenge Data.
Shachar Rosenman, Alon Jacovi, Yoav Goldberg.
In EMNLP 2020.

Towards Faithfully Interpretable NLP Systems: How Should We Define and Evaluate Faithfulness?
Alon Jacovi, Yoav Goldberg.
In ACL 2020.

Improving Task-Oriented Dialogue Systems in Production with Conversation Logs.
Alon Jacovi*, Ori Bar El*, Ofer Lavi, David Boaz, David Amid, Inbal Ronen, Ateret Anaby-Tavor.
In KDD Converse @ KDD 2020.

Neural Network Gradient-based Learning of Black-box Function Interfaces.
Alon Jacovi*, Guy Hadash*, Einat Kermany*, Boaz Carmeli*, Ofer Lavi, George Kour, Jonathan Berant.
In ICLR 2019.

Learning and Understanding Different Categories of Sexism Using Convolutional Neural Network Filters. *(Extended Abstract)*
Sima Sharifirad, Alon Jacovi, Stan Matwin.
In Widening NLP @ ACL 2019.

Understanding Convolutional Neural Networks for Text Classification.
Alon Jacovi, Oren Sar Shalom, Yoav Goldberg.
In BlackboxNLP @ EMNLP 2018. *(oral presentation)*

PATENTS

Computerized dialog system improvements based on conversation data. *Filed, 2020*
Alon Jacovi, Ateret Anaby-Tavor, David Amid, David Boaz, Inbal Ronen, Ofer Lavi, Ori Bar El.
US Patent: P202001942 US01

ACADEMIC SERVICE

Reviewer: ACL 2020, HAMLETS 2020, EACL 2021, NAACL 2021.

Assistant Reviewer: IJCAI 2018, IJCAI 2019.

INVITED TALKS

Microsoft, Oct 2020: *Formalizing Properties of Interpretability in NLP.*

ACLi, Aug 2020: *Towards Faithfully Interpretable NLP Systems: How Should We Define and Evaluate Faithfulness?*

ISCOL, Sep 2018: *Understanding Convolutional Neural Networks for Text Classification.*

HONORS AND SCHOLARSHIPS

2019–2023	The President's Scholarship for Outstanding Doctoral Fellows
2016	The Miryam and Ezra Sofer Scholarship for Excellent Students
2016	Dean's Honors
2015	The Grace Shua and Jacob Ballas Scholarship
2015	Dean's Honors

LANGUAGES

Native	Hebrew
Fluent	English
Conversational	Japanese