Giannis Karamanolakis

Department of Computer Science, Columbia University
New York, NY 10027
Office: Mudd 416, Data Science Institute (map)

WWW: https://gkaramanolakis.github.io
E-mail: gkaraman@cs.columbia.edu
Phone: +1 (917) 331-8844

Research Interests

My research has been focusing on **minimally-supervised learning** frameworks for addressing real-world text mining applications without expensive data annotation. My research goal is to develop tools that can **assist domain experts** in teaching and learning from machines via efficient, accurate, and expressive interaction beyond standard labeling.

Relevant topics: Machine Learning, Information Extraction, Natural Language Processing, Deep Learning, Active Learning, Semi/Weakly-Supervised Learning, Cross-Lingual Learning, Interactive Machine Teaching, Multimodal Learning.

Professional and Research Experience

5. **Graduate Research Assistant**   09/2017 - now
Columbia University, Department of Computer Science, New York, NY, United States.
Advisors: Luis Gravano, Daniel Hsu

   - Developing minimally-supervised learning algorithms for text mining (LLD'19, WNUT'19, EMNLP'19/20, LOUHI'20).
   - Deploying systems for "Information Extraction from Social Media for Actionable Inferences in Public Health" (NSF).

4. **Research Scientist Intern**   05/2020 - 08/2020
Microsoft Research AI, Redmond, WA, United States.
Supervisors: Ahmed Awadallah, Subho Mukherjee, Guoqing Zheng (Language & Information Technologies team).

   - Developed ASTRA, a semi-supervised learning framework for self-training with expert-defined rules (NAACL '21 paper).
   - Designed a benchmark on semi/weakly-supervised learning for language tasks (under submission).

3. **Applied Scientist / Machine Learning Scientist Intern**   05/2019 - 12/2019
Amazon.com, Seattle, WA, United States.
Supervisor: Xin Luna Dong (Product Graph team).

   - Developed TXtract, a neural network for large-scale knowledge extraction from product descriptions (ACL '20 paper).
   - Improved product understanding through 15% higher coverage in knowledge graph construction (KDD '20 paper).

2. **Machine Learning Engineer**   04/2016 - 08/2017
Behavioral Signals, Los Angeles, United States (worked remotely from Athens, Greece).

   - Trained statistical models for recognizing core emotions and key performance indicators from conversational speech.

1. **Research Intern**   11/2015 - 2/2016
NCSR "Demokritos", Computational Intelligence Laboratory, Athens, Greece.
Supervisors: Theodoros Giannakopoulos, Stavros Perantonis.

Education

2. **Ph.D. in Computer Science**   09/2017 - now
Columbia University, New York, NY, United States.

   - **Advisors:** Prof. Luis Gravano, Prof. Daniel Hsu. **Current GPA:** 4.26/4.00.

1. **M.Eng. in Electrical & Computer Engineering**   09/2011 - 07/2017
National Technical University of Athens, Athens, Greece.

   - **Advisor:** Prof. Alexandros Potamianos. **GPA:** 9.26/10.00 (top 2%).
   - **Thesis:** "Grounding Word Semantic Representations to Acoustic and Visual Perceptual Modalities".

Selected Publications

10. **Giannis Karamanolakis**, Subhabrata Mukherjee, Guoqing Zheng, and Ahmed Hassan Awadallah. **Self-Training with Weak Supervision**. In *NAACL*, 2021 (Oral).

9. **Giannis Karamanolakis**, Daniel Hsu, and Luis Gravano. **Cross-Lingual Text Classification with Minimal Resources By Transferring a Sparse Teacher**. In *Findings of EMNLP*, 2020.

8. **Giannis Karamanolakis**, Jun Ma, and Xin Luna Dong. **TXtract: Taxonomy-Aware Knowledge Extraction for Thousands of Product Categories**. In *ACL*, 2020 (Oral).

7. Xin Luna Dong, ..., **Giannis Karamanolakis**, ..., Christos Faloutsos, Andrew McCallum, and Jiawei Han. **AutoKnow: Self-Driving Knowledge Collection for Products of Thousands of Types**. In *KDD*, 2020 (Oral).

6. **Giannis Karamanolakis**, Daniel Hsu, and Luis Gravano. **Leveraging Just a Few Keywords for Fine-Grained Aspect Detection Through Weakly Supervised Co-Training**. In *EMNLP*, 2019 (Oral).

5. **Giannis Karamanolakis**, Daniel Hsu, and Luis Gravano. **Weakly Supervised Attention Networks for Fine-Grained Opinion Mining and Public Health**. In *EMNLP Workshop on Noisy User-generated Text*, 2019 (Oral).

4. **Giannis Karamanolakis**, Daniel Hsu, and Luis Gravano. **Training Neural Networks for Aspect Extraction Using Descriptive Keywords Only**. In *ICLR 2nd Workshop on Learning from Limited Labeled Data*, 2019.

3. Katy Ilonka Gero, **Giannis Karamanolakis**, and Lydia Chilton. **Transfer Learning for Style-Specific Text Generation**. In *NeurIPS Workshop on Machine Learning for Creativity and Design*, 2018.

2. **Giannis Karamanolakis**, Kevin Raji Cherian, Ananth Ravi Narayan, Jie Yuan, Da Tang, and Tony Jebara. **Item Recommendation with Variational Autoencoders and Heterogenous Priors**. In *RecSys Workshop on Deep Learning for Recommender Systems*, 2018 (Oral).

1. **Giannis Karamanolakis**, Elias Iosif, Athanasia Zlatintsi, Aggelos Pikrakis, and Alexandros Potamianos. **Audio-Based Distributional Representations of Meaning Using a Fusion of Feature Encodings**. In *INTERSPEECH*, 2016 (Oral).

Selected Fellowships, Honors, and Awards

7. **Leventis Foundation Scholarship**, New York, US — 08/2019, 08/2020, 08/2021

6. **Gerondelis Foundation Scholarship**, New York, US — 06/2018

5. **Joseph F. Traub Fellowship**, Columbia University, New York, US — 04/2018

4. **Addressing the Opioid Epidemic, Design Challenge Award**, Columbia University, New York, US — 12/2017

3. **Columbia Engineering Presidential Fellowship**, New York, US — 08/2017 - 05/2022

2. **Eurobank EFG Scholarship**, Athens, Greece — 06/2012

1. **Hellenic Mathematic Society "Euclides" & "Thales" Awards**, Sitia, Greece — 02/2009, 03/2010

Technical Skills and Academic Service

- **Programming:** Python (PyTorch, Tensorflow, Keras, Pyro), C++, Java, MATLAB, R, C.

- **Agile Development and Project Management:** Git, JIRA, Jenkins, Scrum, Kanban.

- **Reviewer Activity:** ICML '21 (expert reviewer), NeurIPS '19-21 (top 10% reviewer), ACL Rolling Review '21, ACL '19/21, EMNLP '21, Machine Learning Journal '21, TrueFact '20, TKDD journal '19, WWW '19, IEEE-TETCI '18, CSL Journal '17.

- **Mentoring:** Designed research projects for 12 students (8 undergraduate, 4 graduate); 5 published papers.