Di Jin

PhD Student

(617) 710-6221 | jindi15@mit.edu | jind11 | jindi11

Education

Massachusetts Institute of Technology — *Boston, MA*

PhD in Computer Science, GPA: 4.9/5.0 — *Sep. 2015-PRESENT*

- **Advisor:** Prof. Peter Szolovits
- **Research Focus:** Natural Language Processing (NLP); Clinical Text Mining; Bio-medical Image Processing; Clinical Time Series Analysis

Tsinghua University — *Beijing, China*

Bachelor in Precision Instrument & Economics, GPA: 3.85/4.0 — *Sep. 2011-June 2015*

- **Minor:** Computer Science
- **Distinguished Graduate Award** in both University & Beijing levels

Research Experience

Clinical Decision Making Group, CSAIL, MIT — *Boston, MA*

Research Assistant — *July 2017 - PRESENT*

- Mainly working on NLP, including joint sentence classification, relation extraction, sequence labeling, textual generation on style transfer, question answering, and dialogue state tracking
- Partial work on computer vision on bio-medical images such as super-resolution enhancement of pathological images
- Partial work on time series analysis, such as predicting the chronic disease risk progression based on patients' historical medical records

Laser Biomedical Research Center, MIT — *Boston, MA*

Research Assistant — *Sep. 2015 - June 2017*

- Utilized Full Convolutional Network to implement 2D phase image segmentation to separate individual cells
- Used convolutional neural network to classify various kinds of white blood cells based on 2D phase images as well as 3D refractive index maps

Ho–Systems Lab, UCLA — *Los Angeles, CA*

Research Assistant — *June 2014 - Sep. 2014*

- Created a portable compact wide field-of-view fluorescent microscopy controlled wireless to observe cell dynamics in incubators in real-time
- Used the device to measure cellular viability over time under drug treatment

Industry Experience

Amazon Alexa AI — *Sunnyvale, CA*

Applied Scientist Intern — *June 2019 - Aug. 2019*

- Conducted independent research on the multi-choice question answering (MCQA) task
- Proposed a multi-stage and multi-task transfer learning strategy to significantly improve the performance of large pre-trained models such as BERT on low-resource MCQA data
- Achieved new state-of-the-art performance on four benchmark datasets, surpassing previous baselines by at least 15%
- Best models achieved have shown great zero-shot performance on the dialogue state tracking task, thus being adopted in the production

Amazon Alexa — *Boston, MA*

Applied Scientist Intern — *June 2018 - Aug. 2018*

- Aimed to improve the named entity recognition (NER) model used in production, which is one of the three core components of the language understanding unit of Alexa devices
- Proposed a new deep learning based sequence labeling architecture
- Augmented the multi-domain joint F1 performance via multi-task learning by combining the intent classification and NER tasks together
- Achieved over 20% relative increase of F1 score across all domains compared with the currently used production model

Selected CS Publication

- Jin, Di and Peter Szolovits. "Hierarchical Neural Networks for Sequential Sentence Classification in Medical Scientific Abstracts." EMNLP-2018. </>
- Zhijing Jin*, Di Jin*, Jonas Mueller, Nicholas Matthews, and Enrico Santus. "IMaT: Unsupervised Text Attribute Transfer via Iterative Matching and Translation." EMNLP 2019. </>

- Jin, Di, Gao, Shuyang, Kao, Jiun-Yu, Chung, Tagyoung, Hakkani-tur, Dilek. “MMM: Multi-stage Multi-task Learning for Multi-choice Reading Comprehension.” Accepted by AAAI-2020.
- Jin, Di, Zhijing Jin, Joey Tianyi Zhou and Peter Szolovits. “Is BERT Really Robust? A Strong Baseline for Natural Language Attack on Text Classification and Entailment.” Accepted by AAAI-2020.
- Jin, Di, Zhijing Jin, Joey Tianyi Zhou and Peter Szolovits. “Unsupervised Domain Adaptation for Neural Machine Translation via Back Translation.” Submitted to IJCAI-2020.
- Jin, Di, Zhijing Jin, Joey Tianyi Zhou, Lisa Orii and Peter Szolovits. “Hooks in the Headline: Learning to Generate Headlines with Controlled Styles.” ACL-2020.
- Jin, Di, Franck Dernoncourt, Elena Sergeeva, Matthew McDermott, and Geeticka Chauhan. ”MIT-MEDG at SemEval-2018 task 7: Semantic relation classification via convolution neural network.” In Proceedings of The 12th International Workshop on Semantic Evaluation, pp. 798-804. NAACL 2018.
- Zhou, Joey Tianyi, Hao Zhang, Di Jin, Hongyuan Zhu, Minglei Fang, Rick Siow Mong Goh and Kenneth K Kwok. “Dual Adversarial Neural Transfer for Low-Resource Named Entity Recognition.” ACL (2019).
- J. T. Zhou, H. Zhang, D. Jin and X. Peng, ”Dual Adversarial Transfer for Sequence Labeling,” in IEEE Transactions on Pattern Analysis and Machine Intelligence.
- Zhou, Joey Tianyi, Hao Zhang, Di Jin, Xi Peng, Yang Xiao and Zhiguo Cao. “RoSeq: Robust Sequence Labeling.” IEEE transactions on neural networks and learning systems (2019): n. pag.
- Zhang, Hao, Chunyu Fang, Xinlin Xie, Yicong Yang, Wei Mei, Di Jin, and Peng Fei. ”High-throughput, high-resolution deep learning microscopy based on registration-free generative adversarial network.” Biomedical optics express 10, no. 3 (2019): 1044-1063.
- Jin, Di, and Peter Szolovits. ”Advancing PICO Element Detection in Biomedical Text via Deep Neural Networks.” Bioinformatics (2020).
- Jin, Di and Peter Szolovits. “PICO Element Detection in Medical Text via Long Short-Term Memory Neural Networks.” Proceedings of the BioNLP 2018 workshop, ACL 2018.
- Emily Alsentzer, John Murphy, William Boag, Wei-Hung Weng, Di Jindi, Tristan Naumann, Matthew McDermott. “Publicly Available Clinical BERT Embeddings” Proceedings of the 2nd Clinical Natural Language Processing Workshop, ACL 2019.

Selected Non-CS Publication

- Jin, Di, Bin Deng, J. X. Li, W. Cai, L. Tu, J. Chen, Q. Wu, and W. H. Wang. ”A microfluidic device enabling high-efficiency single cell trapping.” Biomicrofluidics 9, no. 1 (2015): 014101.
- Jin, Di, Renjie Zhou, Zahid Yaqoob, and Peter TC So. ”Tomographic phase microscopy: principles and applications in bioimaging.” JOSA B 34, no. 5 (2017): B64-B77.
- Jin, Di, Dennis Wong, Junxiang Li, Zhang Luo, Yiran Guo, Bifeng Liu, Qiong Wu, Chih-Ming Ho, and Peng Fei. ”Compact wireless microscope for in-situ time course study of large scale cell dynamics within an incubator.” Scientific reports 5 (2015): 18483.
- Jin, Di, Yongjin Sung, Niyom Lue, Yang⊠Hyo Kim, Peter TC So, and Zahid Yaqoob. ”Large population cell characterization using quantitative phase cytometer.” Cytometry Part A 91, no. 5 (2017): 450-459.
- Jin, Di, Renjie Zhou, Zahid Yaqoob, and Peter TC So. ”Dynamic spatial filtering using a digital micromirror device for high-speed optical diffraction tomography.” Optics express 26, no. 1 (2018): 428-437.
- Hosseini, Poorya, Di Jin, Zahid Yaqoob, and Peter TC So. ”Single-shot dual-wavelength interferometric microscopy.” Methods 136 (2018).
- Zhou, Renjie, Di Jin, Poorya Hosseini, Vijay Raj Singh, Yang-hyo Kim, Cuifang Kuang, Ramachandra R. Dasari, Zahid Yaqoob, and Peter TC So. ”Modeling the depth-sectioning effect in reflection-mode dynamic speckle-field interferometric microscopy.” Optics express 25, no. 1 (2017): 130-143.

Honors

Fall 2015	Chee C. Tung (1966) Fellowship	*MIT*
2014	Academic Excellence Scholarship	*Tsinghua University*
2013	National Encouragement Scholarship	*Tsinghua University*
2012	Comprehensive Excellence Scholarship	*Tsinghua University*
2014	**2nd Prize,** Capital Entrepreneurship Competition	*Beijing*

Teaching Experience

CSAIL, MIT *Boston, MA*

Undergraduate Research Opportunities Program (UROP) Mentor *May. 2019 – PRESENT*

Was the primary supervisor and mentor of five MIT undergraduates for the undergraduate research program and one rotational PhD student. Helped them publish three papers in top-conferences.

GGU Graduate School Application Consulting *Boston, MA*

Course Lecturer *Sep. 2019 – Dec. 2019*

Served as a remote course lecturer, teaching on the topic of how to use NLP tools to detect fake news.