Mike Zhang

@ mikz@itu.dk · ⊕ jjzha.github.io · ⚪ jjzha · 🔗 /in/jjzha   Copenhagen, Denmark

INTERESTS

Natural Language Processing, Information Extraction, Active Learning, Distant Supervision, Weak Supervision, Domain Adaptation, Transfer Learning.

EDUCATION

IT University of Copenhagen, Copenhagen, Denmark
Ph.D., Computer Science, exp. 2023
Advisor: Prof. Barbara Plank

University of Groningen, Groningen, the Netherlands
M.A., Communication and Information Studies, 2020   *Cum Laude*
B.Sc., Information Science, 2018

PROFESSIONAL EXPERIENCE

IT University of Copenhagen, Copenhagen, Denmark
Ph.D. Fellow   **September, 2020 - Present**
Extracting information from multilingual job vacancy data with time-shift, as a result to create better better job matching. Research areas include Information Extraction, Active Learning, Distant Supervision, Transfer Learning, Domain Adaptation. Other duties include TA'ing.

Dataprovider.com, Groningen, the Netherlands
Data Engineer   **January, 2020 - August, 2020**
Maintained and implemented software to classify companies according to the SIC taxonomy codes. Generated domain names with sequence models. Software included Elasticsearch, Python, PyTorch, NGINX.

Dataprovider.com, Groningen, the Netherlands
Data Scientist Intern   **August, 2019 - December, 2019**
Conducted research on how to generate valid domain names from character representations in the Data Science team. Software included Python and PyTorch.

PUBLICATIONS
∗ = eq. contribution

K. N. Jensen∗ , **M. Zhang**∗, and B. Plank. De-identification of privacy-related entities in job postings. In Proceedings of the Nordic Conference on Computational Linguistics, 2021.

M. Zhang and A. Toral. The Effect of Translationese in Machine Translation Test Sets. In Proceedings of the Fourth Conference on Machine Translation, 2019.

M. Zhang, R. David, L. Graumans, and G. Timmerman. Grunn2019 at SemEval-2019 Task 5: Shared Task on Multilingual Detection of Hate. In Proceedings of the 13th International Workshop on Semantic Evaluation, 2019.

TEACHING

Teaching Assistant
IT University of Copenhagen
Natural Language Processing and Deep Learning (BSSEYEP1KU)   Spring, 2021
University of Groningen
Social Media (LIX017B05)   Spring, 2019
Learning from Data (LIX016M05)   Fall, 2019
Machine Learning (SOMINDW07)   Fall, 2019, 2020

SERVICES

Conference reviewing: CoNLL (2021), WMT (2019)

SKILLS   **Languages:** Native speaker of Dutch, near native in English, basic proficiency in both Mandarin and Spanish.
Technical Skills: Python (PyTorch, Scikit, NumPy, Pandas), Elasticsearch, Java, HTML, R.