Mike Zhang

Website: jjzha.github.io
E-mail: mikz@itu.dk
github.com/jjzha

EDUCATION

IT University of Copenhagen   Copenhagen, DK
Ph.D. Natural Language Processing, supervisor: Prof. Barbara Plank   2020–Current

University of Groningen   Groningen, NL
M.A. Information Science, GPA: 8.4/10.0 (Cum Laude)   2018–2020

- Thesis: *Translationese in the WMT: The Effect of Translationese in Machine Translation Test Sets.*
  Graded with a 9.0/10.0, supervisor: Dr. Antonio Toral Ruiz

University of Groningen   Groningen, NL
B.S. Information Science, GPA: 7.6/10.0   2015–2018

- Thesis: *From Relations to Falsehood: Labeled Bilexical Dependencies for Deception Detection.*
  Graded with an 8.0/10.0, supervisor: Prof. Barbara Plank

- Exchange Semester at The Chinese University of Hong Kong, Hong Kong SAR during Spring 2018. GPA:
  8.0/10.0

EXPERIENCE

IT University of Copenhagen   Copenhagen, DK
PhD Fellow   Sep. 2020–Current

- *Multilingual Information Extraction for Job Post Analysis*: Information Extraction
  on multilingual job vacancy data over time, to create better better job matching.

- Keywords: *low-resource learning, cross-lingual learning, continual learning, (E/I)SCO*

Dataprovider.com   Groningen, NL
Data Science Intern, Data Engineer   Aug. 2019–Aug. 2020

- *Companies*: Classification of companies according to the SIC taxonomy codes.

- *Domain Discovery*: Generation of undiscovered existing domain names, by exploiting
  domain names already found.

- Keywords: *extreme multiclass classification, distant supervision, SIC, NAICS,
  ensemble learning, web domain, deep learning*

PUBLICATIONS

[1]   K. N. Jensen, **M. Zhang**, and B. Plank, "De-identification of privacy-related entities in job postings",
      NoDaLiDa 2021, p. 210, 2021.

[2]   **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, pp. 391–395.

[3]   **M. Zhang** and A. Toral, "The effect of translationese in machine translation test sets", in *Proceedings of
      the Fourth Conference on Machine Translation (Volume 1: Research Papers)*, 2019, pp. 73–81.

Teaching

Teaching Assistant at IT University of Copenhagen   Spring 2021
Introduction to Natural Language Processing and Deep Learning (BSSEYEP1KU)

Teaching Assistant at University of Groningen   Spring 2019, Fall 2019, Fall 2020
Social Media (LIX017B05)
Learning from Data (LIX016M05)
Machine Learning (SOMINDW07)