Ngan Dong

Curriculum Vitae

PERSONAL DETAILS

Birth	July 24, 1988
Address	Knoevenagelweg 3, 30165, Hannover, Germany
Phone	+(49) 152-107-23190
Mail	dong@l3s.de
Website	http://l3s.de/~dong/
G. Scholar	https://scholar.google.com/citations?user=M4zbOcsAAAAJ

EDUCATION

PhD Computer Science 2019 - 2022

Leibniz University Hannover

Research Topic: "Deep learning for precision medicine"

MSc. Computer Science 2011-2013

Washington State University

Thesis: "Natural Language Generation from Graphs"

BSc. in Computer Science, Honor program 2006-2010

VNU University of Engineering and Technology

EXPERTISE

- More than 5 years of developing Machine Learning algorithms at industry scale for Recommendation systems and various Natural Language Processing tasks.
- Nearly 3 years working on Machine Learning models for Biological problems which focus on Data Integration and involve the use of Language modeling for Protein representation learning, Graph Representation Learning techniques, Multi-task learning, and Feature Selection.

WORK EXPERIENCE

Research Assistant 2019-Present

Leibniz University Hannover, Full-time

- Working on the PRESENt project[1] which aims at integrating clinical, biological, and big data research to advance our understanding of norovirus gastroenteritis.
- Developed state-of-the-art models for various bioinformatic problem which focus on multiple data sources integration, network analysis, multitask learning, graph representation learning, and feature selection.
- Developed a cochlear implant outcome prediction model from real patient data. The work involved heavy data preprocessing tasks, feature extraction, feature selection, and model development.
- Perform some others data retrieval, data pre-processing tasks and specialized analysis on

[1] http://www.translationsallianz.de/train-plattformen/train-projects/present/

the results retrieved from wet-lab experiments given by our biologist partners.

Research Engineer(Machine Learning) 2014-2019

FPT Technology Research Institute, Full-time

- Being one of the first ML researchers working on Natural Language Understanding module for FPT.AI[2] - one of the first and largest comprehensive AI platforms in Vietnam.
- Worked on (i) Recommendation systems for e-commerce websites and online news, (ii) Sentiment classification for electronic retailer, (iii) User segmentation, and (iv) online-news topic modeling

Software Developer 2013-2014

Citigo Joint stock Company., Full-time

Work as a full-stack developer on out-sourcing projects for Australian customers. Responsible for maintaining the existing systems, developing, testing and deploying new functionality on customers' production servers.

Research Assistant 2011-2013

Artificial Intelligence laboratory, WSU

- Develop GNLG - a Natural Language Generation (NLG) system for Resource Description Framework (RDF) Graphs.
- Build ontology models for cooking recipes and data from a defense-related project about people in a military- controlled area

Teaching Assistant 2010-2011

Computer Science Department, VNU University of Engineering and Technology

- Teaching Assistant for various bachelor level programming courses.

Research Assistant 2009-2010

Computer Science Department, VNU University of Engineering and Technology

- Research on POS taggers for Vietnamese using Conditional Random Fields and Hidden Markov Models.
- Develop a Vietnamese collocation extractor using different statistical methods

SKILLS

Languages	Vietnamese (mother tongue), English (fluent), German (basic)
Code related	Python, PyTorch, pytorch-geometric, sklearn, pandas, numpy, networkx

PUBLICATIONS

Ngan Dong, Johanna Schrader, Stefanie Mücke, Megha Khosla, "*A Message Passing framework with Multiple data integration for miRNA-Disease association prediction*", preprint at ResearchSquare, 2022.

Ngan Dong, Stefanie Mücke, Megha Khosla, "*MuCoMiD: A Multitask graph Convolutional Learning Framework for miRNA-Disease Association Prediction*", accepted to IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB), 2022.

Ngan Dong, Graham Brogden, Gisa Gerold, Megha Khosla, "*A multitask transfer learning framework for the prediction of virus-human protein-protein interactions*", BMC Bioinformatics, 2021.

Ngan Dong, Megha Khosla, "*A Multitask Convolutional Learning Framework for miRNA-Disease Association Prediction*", BIOKDD 2021.

Ngan Dong, Megha Khosla,, "*A multitask transfer learning framework for Novel virus-human protein interactions*", ICLR Workshop on AI for Public Health, 2021.

[2]https://fpt.ai/

Ngan Dong, Megha Khosla, "*Towards a consistent evaluation of miRNA-disease association prediction models.*", 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2020.

Ngan Dong, Megha Khosla. "*Revisiting Feature Selection with Data Complexity for Biomedicine.*" 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE).

Kim Anh Nguyen, **Ngan Dong**, Cam Tu Nguyen "*Attentive Neural Network for Named Entity Recognition in Vietnamese*", 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF). IEEE, 2019.

Ngan Dong, Larry Holder, "*Natural Language Generation from Graphs*", International Journal of Semantic Computing.Vol. 8, No. 3, pp. 335-384, 2014

ACADEMIC SERVICES

Teaching assistant for Machine Learning for Graphs course (2021)
Seminar supervisor for Artificial Intelligence course (2021)
Msc. students supervised/co-supervised

Johanna Schrader - Thesis title: Application of Graph Structure Learning in Biological Data Analysis (2021)

Luo Yi - Thesis title: At the Interface between Biomedical Research and Software engineering (2021)

HONOR AND AWARDS

August 2011 – May 2013: Vietnam Education Foundation (VEF) scholarship, which was funded by the US government.

July 2010: Scholarship for Digital Signal Processing Academy Summer School

Dec 2009: Outstanding female student in Information Technology
Scholarship from the Ministry of Information and Communication, Vietnam

Oct 2007: VNU University of Engineering and Technology Outstanding student

REFERENCES

Available upon request.