Dung Ngoc Thai

http://www.cs.umass.edu/~dthai
dthai@cs.umass.edu | (669)-233-3660

EDUCATION

UMASS, AMHERST
PH.D. COMPUTER SCIENCE
Sep 2016 - present

HCMUT
MS COMPUTER SCIENCE
Nov 2014 | HCM, VN
Cum. GPA: 3.9

SOCIALS

github.com/dungtn
linkedin.com/in/dung-thai

COURSES

Reinforcement Learning
by Philip Thomas

Machine Learning
by Srihdar Mahadevan

Deep Learning
by Erik Learned-Miller

SKILLS

Programming Languages
Python • Java • C++
Matlab • Julia

Frameworks
Theano • Tensorflow • Keras
CUDA • NLTK

REFERENCES

Prof. Andrew McCallum , UMass
mccallum@cs.umass.edu
Prof. Nam Thoai , HCMUT
nam@cse.hcmut.edu.vn
Prof. Vu Dinh , HCMUT
vudda@uit.edu.vn

SUMMARY

I'm a fourth year Ph.D. student at UMass Amherst, advised by Prof. Andrew McCallum . My research focuses on knowledge-informed language models and their applications for knowledge base completion and question answering. I'm intrigued by how models pre-trained with MLM objective perform so well on a wide range of tasks and would like to explore their capabilities when paired with semi-structured data such as knowledge graphs.

RESEARCH EXPERIENCES

INFORMATION EXTRACTION SYNTHESIS LAB | GRADUATE STUDENT
Sep 2016 – present | UMass Amherst, US
My ongoing research study how to enhance pre-trained language representations with information from knowledge graphs. Previously, I worked on training models on large-scale, distantly supervised data and modeling complex output constraints with Latent Conditional Random Field.

IBM | RESEARCH INTERN

Yorktown 2020	Compositional question answering over knowledge bases built a retrieval-based language model and learn a knowledge-informed encoder-decoder model for answering natural question given a knowledge base.

ADOBE INC. | RESEARCH INTERN

San Jose 2018	Question answering on semi-structured tables built a multi-heads attention model based on the Neural Programmer architecture.
San Jose 2017	Variational autoencoder for semi-supervised question answering learned an unsupervised question representation to improve the generalization of the supervised question answering model.

SELECTED PUBLICATIONS

EACL 2021	Bi-directional Entity to Text Attention for Knowledge Informed Text Representations Trapit Bansal, Dung Thai , Raghuveer Thirukovalluru, Andrew McCallum in submission to EACL
EACL 2021	Knowledge Informed Semantic Parsing for Conversational QA Dung Thai , Raghuveer Thirukovalluru, Mukund Sridhar, Shruti Chanumolu, Sankaranarayanan Ananthakrishnan, Andrew McCallum in submission to EACL
AKBC 2020	Using BibTeX to Automatically Generate Labeled Data for Citation Field Extraction Dung Thai , Zhiyang Xu, Nicholas Monath, Boris Veytsman, Andrew McCallum in Automated Knowledge Base Construction, Online
CoNLL 2018	Embedded-State Latent Conditional Random Fields for Sequence Labeling (Oral Presentation) Dung Thai , Sree H. Ramesh, Shikhar Murty, Luke Vilnis, Andrew McCallum in SIGNLL Conference on Computational Natural Language Learning, Belgium
ICML 2017	Low-rank hidden state embeddings for Viterbi sequence labeling Dung Thai , Shikhar Murty, Trapit Bansal, Luke Vilnis, David Belanger, Andrew McCallum in 1st DeepStruct Workshop, in 34^{th} International Conference on Machine Learning , Australia.