Dat Quoc Nguyen

Research fellow, The University of Melbourne

Room 3.17, Doug McDonell building (Building 168)
School of Computing and Information Systems
Grattan Street, Parkville VIC 3010, Australia
✉ dqnguyen@unimelb.edu.au
🖥 https://people.eng.unimelb.edu.au/dqnguyen

Education

06/14–05/17   **Ph.D student**, *Macquarie University*, Australia.
- Principal supervisor: Mark Johnson; Associate supervisor: Mark Dras
- Thesis: MODELING TOPICS AND KNOWLEDGE BASES WITH VECTOR REPRESENTATIONS
- Thesis submitted for examination on 30/05/2017; Ph.D qualified on 20/09/2017; Ph.D conferred on 17/04/2018.

12/09–10/11   **M.Sc student**, *VNU University of Engineering and Technology*, Vietnam (hons).

09/05–06/09   **B.Sc student**, *VNU University of Engineering and Technology*, Vietnam (hons).

Experience

12/17–   **Research fellow**, *The University of Melbourne*, Australia.

Performing research in machine learning for natural language processing, with a recent focus on developing deep learning models for tagging, parsing, relation extraction and knowledge base completion.

07/17–08/17   **Research associate**, *Macquarie University*, Australia.

Studied how accuracy varies as a function of training data size for several NLP models and tasks (part-of-speech tagging, dependency parsing and machine translation).

07/09–05/14   **Researcher**, *VNU University of Engineering and Technology*, Vietnam.

Worked on projects of developing knowledge-based question answering and dialogue systems for Vietnamese; and on other projects which apply machine learning for natural language problems such as part-of-speech tagging, dependency parsing, text summarization and sentiment analysis.

07/09–05/14   **Teaching assistant**, *VNU University of Engineering and Technology*, Vietnam.

Prepared and conducted tutorials and laboratory sessions for the Bachelor degree on Information Technology including the following courses: Introduction to Artificial Intelligence, Data Structures and Algorithm, Programming Language (C/C++, Java) and Object-Oriented Programming.

Prepared assignment questions and marked assignments; prepared examination papers and supervised examinations; led discussion sessions.

08/12–11/12   **Research intern**, *Dublin City University*, Ireland.

Worked on a project of building a FAQ question answering system for mobile-phone forums.

11/11–05/12   **Research intern**, *Technische Universität Darmstadt*, Germany.

Worked on a project of analyzing discussion talk pages and edit activities over Wikipedia articles.

Awards & Honors

2016   ALTA 2016 student travel scholarship.

2016   ACL 2016 student volunteer award.

2016   Macquarie University PGRF travel fund.

2015   ALTA 2015 student travel scholarship.

2015   EMNLP 2015 student scholarship.

02/15–05/17   NRPA Top-Up scholarship, awarded by National ICT Australia (NICTA).

06/14–05/17   International Macquarie University Research Excellence Scholarship.

06/14–05/17   Australian government International Postgraduate Research Scholarship.

2013   ISWC 2013 conference travel grant from the Semantic Web Science Association.

2012   LaDaKT travel grant from the Winter camp on Language and Data/Knowledge Technologies, Italy.

2011   RANLP 2011 conference travel grant from the Vietnam's NAFOSTED foundation.

2011   Best master student award, VNU University of Engineering and Technology.

2010   TOSHIBA Foundation Scholarship for Excellent Master Students.

2009   Best undergraduate thesis award, VNU University of Engineering and Technology.

2009   First prize at the VNU University of Engineering and Technology's Undergraduate Research Conference.

2009   Consolation prize in Algebra at the National Mathematical Olympiad for Undergraduates, Vietnam.

Competitions

2018   Ranked 5th/43, Semeval 2018 Task 3 on Irony Detection in English Tweets.

2017   Top 10 semi-finalists (7th/469), CIKM 2017 AnalytiCup — Lazada Product Title Quality Challenge.

Academic Activities

Advising   **Student (co-)supervision**: Zenan Zhai's master thesis, The University of Melbourne (2018).

Examiner   Three master's theses, The University of Melbourne (2018); One master's thesis, The University of Melbourne (2019).

Service   **PC member**: NAACL (2018-2019), ACL (2019), EMNLP (2019), CoNLL Shared Task (2017-2018), ACL SRW workshop (2018-2019).

Journal reviewer: IEEE Transactions on Knowledge and Data Engineering (2018), BMC Bioinformatics (2019).

Talks

SLIDES   *Joint models for POS tagging and dependency parsing, and for NER and relation classification* – NLP seminar talk at the University of Melbourne (10/18) & talk at Elsevier Information Systems GmbH, Germany (10/18).

SLIDES   *A Convolutional Neural Network-based Model for Knowledge Base Completion* – NLP seminar talk at the University of Melbourne (04/18) & LTG seminar talk at Macquarie University (04/18).

SLIDES   *A rule-based framework for sequence labeling tasks and its application to Vietnamese NLP* – LTG seminar talk at Macquarie University (10/17) & NLP seminar talk at the University of Melbourne (01/18).

SLIDES   *Modeling multi-relational data from knowledge bases with embeddings* – PRaDA Seminar Series talk at Deakin University, Australia (12/16).

SLIDES   *Modeling topics and knowledge bases with embeddings* – FEAST Series talk at the COLI department, Saarland University, Germany (08/16) & talk at the UKP lab at Technische Universität Darmstadt, Germany (08/16) & talk at the Sydney NLP Meetup (11/16).

SLIDES   *Improving Topic Models with Latent Feature Word Representations* – talk at NICTA, Australia (11/15).

Computer Skills

ML toolkits   DyNet, PyTorch, Tensorflow, Keras, Scikit-learn, Weka

Languages   Python, Java, C/C++, R, PHP, SQL

Selected Software

For a full list of my research software, see: https://github.com/datquocnguyen

jointRE   Implementation of a neural network model for joint extraction of named entities and their semantic relations (ECIR 2019).

BioPosDep   A processing pipeline of tokenization, sentence segmentation, part-of-speech (POS) tagging and dependency parsing for biomedical texts (BMC Bioinformatics 2019).

VnCoreNLP   A Vietnamese NLP pipeline of word segmentation, POS tagging, named entity recognition and dependency parsing (NAACL 2018).

jPTDP   Implementations of neural network models for joint POS tagging and dependency parsing (CoNLL 2017–2018). jPTDP provides pre-trained joint models for 40+ languages.

STransE   Implementation of an embedding model of entities and relationships for knowledge base completion (NAACL 2016).

jLDADMM   A Java package for topic modeling on normal or short texts with LDA and DMM models.

LFTM   Implementations of latent feature topic models LF-LDA and LF-DMM, which extend the LDA and DMM topic models with word embeddings (TACL 2015).

RDRPOSTagger   A fast and accurate toolkit for POS and morphological tagging (EACL 2014).

Research Interests

NLP   Relation extraction, Knowledge base completion, Word segmentation, POS tagging, Named entity recognition, Parsing, Question answering, Dialogue system, Text summarization, Machine translation, Sentiment analysis

ML&IR   Deep learning, Topic modeling, Rule-based machine learning, Search personalization

Publications

Google Scholar: https://scholar.google.com/citations?hl=en&user=HVl7vyEAAAAJ