Binh Vu

https://binh-vu.github.io
https://www.linkedin.com/in/binh-v-3828a16a/

Updated: October 1, 2019
binhvu@isi.edu
+1-213-269-9961

RESEARCH INTERESTS

Machine Learning, Semantic Web, Knowledge Graph Construction

EDUCATION

- **Ph.D. Student in Computer Science** Aug 2016 - Present
  University of Southern California, Computer Science Department, Los Angeles, CA
  Advisor: Professor Craig Knoblock
- **Bachelor of Engineering in Computer Science** Aug 2010 - Jan 2015
  HCMC University of Technology, Ho Chi Minh, Vietnam
  Thesis: *Wikipedia-based Entity Disambiguation using Deep Learning*, advisor Professor Tru Cao

RESEARCH EXPERIENCE

- **University of Southern California, Information Sciences Institute**, Marina del Rey, CA Aug 2016 - Present
  Graduate Research Assistant, *Center on Knowledge Graph*
  - **Learning Semantic Models of Data Sources** Oct 2017 - Present
    Learning semantic models that describe structured data sources using ontologies. The semantic models are used to automatically publish data to knowledge graphs.
  - **Identifying potential company hardware/software vulnerabilities** Oct 2016 - Sep 2017
    Auto-crawling online sources to retrieve expertise of employees of a company, then predicting the software and hardware used in the company. A list of vulnerabilities is obtained by linking the software and hardware to the CVE database.
- **Rakuten Inc.**, Tokyo, Japan Jun 2015 - May 2016
  Research Engineer, *Big Data Department*
  - **Fraud Detection in ID Hijacking and Payment**
    Developing a near real-time distributed streaming system using Apache Storm and Cassandra to analyzes time-series data for fraud detection. The system is designed to run models that use related historical events up to the past 60 days to flag fraudulent transactions within seconds.
    Implementing a machine learning model for payment fraud detection.
- **HCMC University of Technology**, Ho Chi Minh, Vietnam Jun 2014 - Dec 2015
  Undergraduate Research Assistant, *Computer Science Department*
  - **Wikipedia-based Entity Disambiguation using Deep Learning**
    Using autoencoders to extract latent features of entities in Wikipedia articles for the entity linking problem

KEY HORNORS

ISI Distinguished Top-Off Fellowship April 2016
Vietnam Education Foundation Fellowship to pursue Ph.D. degree in the U.S 2016
Outstanding Honor Student Award 2011 - 2014

SELECTED PUBLICATIONS

- **Binh Vu**, Craig Knoblock, and Jay Pujara. 2019. *Learning Semantic Models of Data Sources Using Probabilistic Graphical Models.* In The World Wide Web Conference, pp. 1944-1953.
- **Binh Vu**, Jay Pujara, and Craig Knoblock. 2019. *D-REPR: A Language for Describing and Mapping Diversely-Structured Data Sources to RDF.* In The Tenth International Conference on Knowledge Capture (K-CAP).

TECHNICAL SKILLS

- **Languages**: Python, Rust, Java, HTML, CSS, Javascript (Full-stack Web Developer)
- **Technologies**: Pytorch, Tensorflow, Docker, ElasticSearch, Databases (MySQL, Postgres, Redis, Cassandra), Apache Storm, Apache Hadoop, Apache Spark.
- **Operating Systems**: Linux, MacOS