THANH DAC TRAN

CONTACT
INFORMATION

2240 E 3300 S #12,
Salt Lake, UT, 84109

Mobile: 435-294-7480
E-mail: tdtran@wpi.edu
Homepage: https://thanhdtran.github.io

RESEARCH
INTERESTS

Recommendation systems, deep learning, text mining, social media mining.

EDUCATION

Worcester Polytechnic Institute, Worcester, MA, U.S (**2017-Fall 2020**)
- Doctor of Philosophy, Data Science, Fall 2017.
- Advisor: Dr. Kyumin Lee.
- GPA: 4.0

Utah State University, Logan, Utah, U.S (**2015-2017**)
- GPA: 4.0

Hanoi University of Science and Technology (HUST), Hanoi, Vietnam (**2006-2011**)
- B.E., Computer Science, July 2011.
- Thesis: Establish a framework to solve automatically ontology matching problem (**Best undergraduate Thesis**).

HONORS AND
AWARDS

⋆ Yahoo Research Best presentation award for all interns, Yahoo Research 2020.

⋆ WPI Academic Excellence Award.

⋆ WPI Outstanding Graduate Research Innovation Exchange award, 2020 ($250).

⋆ Travel Grant: SIGIR 2020 (€250), WWW 2019 ($1,000), SIGIR 2019 ($1,000), SIGIR 2018 ($1,000), WPI 2018 ($1,000), USU 2016 ($300).

⋆ 3rd prize in annual coding competition at USU, 2016.

⋆ Fellowship award in Vietnam Education Foundation (VEF) Fellowship Program, 2015.

⋆ Top 5 social mobile applications award, Mobile Innovation Challenge for East Asia, May 2013.

⋆ Best undergraduate thesis award, HUST Vietnam 2011.

PUBLICATIONS

15. **Thanh Tran**, Yifan Hu, Changwei Hu, Kevin Yen, Fei Tan, Kyumin Lee, Se Rim Park, "HABERTOR: An Efficient and Effective Deep Hatespeech Detector ", (**EMNLP '20, full, AR=22.4%**).

14. **Thanh Tran**, Di You, Kyumin Lee, *Quaternion-based self-Attentive Long Short-term User Preference Encoding for Recommendation*, (**CIKM '20, full, AR=21%**).

13. **Thanh Tran**, Yifan Hu, Changwei Hu, Kevin Yen, Fei tan, Kequian Li, *Efficient Hyperparameter Optimization for Neural Networks*, (**Yahoo TechPulse' 20, poster**).

12. **Thanh Tran**, Yifan Hu, Changwei Hu, Kevin Yen, Fei Tan, Se Rim Park, "Language Model for Hate-speech Classification in Yahoo News and Finance", **Yahoo TechPulse' 19, full**.

11. **Thanh Tran**, Renee Sweeney, and Kyumin Lee '***Adversarially** Mahalanobis distance based Attentive Song Recommender for automatic playlist continuation*', **SIGIR '19, full, AR=20%**.

10. **Thanh Tran**, Xinyue Liu, Kyumin Lee and Xiangnan Kong "*Signed Distance-based **Deep Memory Recommender** *", **WWW '19, full, AR=18%**.

9. **Thanh Tran**, Kyumin Lee, Yiming Liao and Dongwon Lee "*Regularizing **Matrix Factorization** with user and item embeddings for Recommendation*", **CIKM '18, full, AR=17%**.

8. **Thanh Tran**, Kyumin Lee, Nguyen Vo, Hongkyu Choi "*Identifying On-time Reward Delivery Projects with Estimating Delivery Duration in a Crowdfunding Platform*", **ASONAM '17, full, AR=19%**.

7. **Thanh Tran**, Kyumin Lee, "*Characteristics of On-time and Late Reward Delivery Projects*", **short, ICWSM 2017**.

6. Nguyen Vo, Kyumin Lee, and **Thanh Tran** "*MRAttractor: Detecting Communities from Large-Scale Graphs*", **IEEE Big Data '17, full, AR=18%**.

5. Yiming Liao, **Thanh Tran**, Dongwon Lee and Kyumin Lee "*Understanding Backing Patterns in Online Crowdfunding Communities*", **Websci '17, full, AR=24%**.

4. Nguyen Vo, Kyumin Lee, Cheng Cao, **Thanh Tran** and Hongkyu Choi "*Revealing and Detecting Malicious Retweeter Groups*", **ASONAM 17, short**.

3. **Thanh Tran**, Kyumin Lee, *"Understanding Citizen Reactions and Ebola-Related Information Propagation on Social Media"*, **ASONAM '16, short**.

2. Prudhvi Ratna Badri Satya, Kyumin Lee, Dongwon Lee, **Thanh Tran** and Jiasheng Zhang, *"Uncovering Fake Likers in Online Social Networks"*, **CIKM '16, short**.

1. **Dac-Thanh TRAN**, Duy-Hoa Ngo and Phan-Thuan DO, *"An information content based partitioning method for the anatomical ontology matching task"*, **SoICT 2012, full**.

MANUSCRIPTS

2. Lucas Vinh Tran, **Thanh Tran**, Shanshan Feng, Gao Cong and Xiaoli, *Wasserstein Distance based Metric Learning Chain for Recommender Systems*, under submission.

1. Yifan Hu, Changwei Hu, **Thanh Tran**, Tejaswi Kasturi, Elizabeth Joseph, Matt Gillingham, *What's in a Name? – Gender Classification of Names with Character Based Machine Learning Models*, under submission.

RESEARCH EXPERIENCE

Yahoo Research, NY, USA (June 2020 – August 2020):
- **Position:** Research Intern.
- **Mentor:** Dr. Yifan Hu.
- **Self-tuned Neural Network Project**: Hyper-parameters such as weight decay or regularization balance terms are crucial to the optimal performance of deep neural net-works. Unfortunately, they are hard to tune. Thus, we propose a method to self-tune hyper-parameters, which has a time complexity of only $O(h+m)$ with h is the number of hyper-parameters and m is the number of model's normal parameters, and does not increase the GPU memory. In particular, we show that our model training time increases at a maximum of **4 times** compared to the training time of the model with a fixed hyper-parameter set.

Yahoo Research, NY, USA (June 2019 – August 2019):
- **Position:** Research Intern.
- **Mentor:** Dr. Yifan Hu.
- **Hatespeech Prediction Project**: Pretrain from scratch an adversarial BERT model with only **26M** of parameters, and with a multi-source multi-head fine-tuning design to predict hate-speech comments on Yahoo News and Finance. The final model is **4∼5 times** faster, and uses less than 1/3 of the memory, while improving 1.3% of F1 score compared to **fine-tuned BERT-base**,
- **Gender Prediction Project:** Build a character-based BERT model and attentive character-based LSTM to predict the newly-registered user's gender based on their registered names.

Worcester Polytechnic Institute (August 2017 – now):
- **Position:** Research Assistant.
- **Advisor:** Prof. Kyumin Lee.
- **Project:** Recommendation systems.
- **Description:** Using deep neural networks to design personalized recommendation systems.

Utah State University (2015 – Aug 2017):
- **Position:** Research Assistant.
- **Advisor:** Prof. Kyumin Lee.
- **Project:** Crowdfunding mining and social media mining.
- **Description:** Perform feature engineering, then apply traditional machine learning algorithms to predict the successful rate of crowdfunding projects on crowdfunding platforms.

WORKING EXPERIENCE

Wala, Jsc (April, 2012 - October, 2014):
- **Company Description:** A startup company building a online social network for mobile phone.
- **Position:** Research Engineer
- **Project:** Distributed Storage System, Profile Server, Feed Server, Channel Server, Redis Proxy.
- **Detail:** Develop and maintain back-end systems, establish services for social actions

VNG Corporation (September, 2011 - March, 2012):
- **Company Description:** a leading corporation in popularizing social networks to Vietnamese.
- **Position:** Research Engineer
- **Project:** ZStore, Friend Suggestion, Feed Ranking
- **Detail:** Improved ZStore - a key/value distributed storage system; Developed Feed Ranking - a system of ranking feeds for users; Researched and implemented Friend Suggestion System.

ACADEMIC SERVICES

PC Member: AAAI 2021. **Reviewer:** WWW 2020, JVLDB (2019, 2018).

External Reviewer: 2019 (KDD, WWW, AAAI), 2018 (KDD, AAAI), 2017 (KDD, WWW, SAC, ECIR, ECIR), 2016 (ECIR, SWDM, CIKM, ASONAM, TIST, ICWSM), 2015 (TIST, IJCAI, TKDE, CIKM, TOIS, ICWSM).

ACTIVITIES

- Conference Volunteer: WWW 2019, SIGIR 2019, ODSC Boston 2018.
- Social activity: Trash Collection in Logan Canyon 2016, 2017; WPI freshmen welcome event 2018.

COURSEWORK

- **Machine Learning** (Deep Learning, Social Media Mining, Introduction to Data Science, Business Intelligence). **Statistics** (Stochastic Process in Electronic Systems, Statistical learning and data mining, Stat methods for big data, Linear Regression and Time Series). **Databases** (Big data management, Advanced Database Systems).

COMPUTER SKILLS

Programming: Python, Java, R.
Big data analytics frameworks: Hadoop, Spark.
Databases: Hadoop/Hbase, Pig; Relational Database: MySQL, SQLServer; NoSQL: Cassandra, Leveldb, Redis, Memcached, CouchDB.