

# JOHN BOAZ LEE

*Machine learning researcher working primarily on deep learning,  
sequence & graph modeling, data mining, & reinforcement learning.*

## Education

- 2015–2019 **Worcester Polytechnic Institute**, *PhD in Computer Science (GPA 4.0/4.0)*.  
Thesis: Deep Learning on Graph-structured Data  
Advisor: Xiangnan Kong
- 2010–2012 **University of the Philippines – Diliman**, *MS in Computer Science (GPA 4.0/4.0)*.  
Thesis: A Link Prediction Algorithm for Heterogeneous Bibliographic Information Networks  
Advisor: Henry Adorna
- 2004–2008 **Silliman University**, *BS in Computer Science (GPA 3.1/4.0)*.  
Thesis: Dumaguete e-Traveler – A Knowledge-based Decision Support System  
Advisor: Dave E. Marcial

## Experience

- 2020–present **Facebook**, *Research Scientist. CA, USA*.  
Working on sequence and graph-based ML. Launched four models within first year. First model doubled system recall, another increased top-line metric by ~2%, and a third boosted notification CTR by 2.3%.
- summer of **Facebook**, *Research Intern. CA, USA*.  
2019 Deployed a recurrent pipeline which generated unified person-level embeddings by leveraging interactions from across FB's family of apps. The embeddings increased the AUC-ROC of nine classification tasks.
- summer of **Adobe**, *Research Intern. CA, USA*.  
2018 Worked on a deep learning model for entity alignment which was deployed internally. Also produced two research publications and filed a patent application.
- summer of **Xerox PARC**, *Research Intern. CA, USA*.  
2017 Led a project on graph classification and collaborated on a second one on dynamic network representation learning. Published two research papers from these projects.
- 2015–2019 **Worcester Polytechnic Institute**, *Research Assistant. MA, USA*.  
Researched and developed deep learning solutions for relational data. Also assisted in the ff. courses during time as a TA: data mining, software security, computer networks, and theory of computation.
- 2012–2015 **Ateneo de Manila University**, *Instructor. Manila, Philippines*.  
Handled the ff. courses: algorithms and data structures, Java programming, C# programming, and design patterns. Also played a minor role in coaching the competitive programming team.
- fall of 2012 **Nara Institute of Science and Technology**, *Research Intern. Nara, Japan*.  
Proposed and implemented a method for automatic software patch reviewer recommendation. The research was published and won a "Best Presentation Paper" award.
- 2010–2012 **University of the Philippines – Diliman**, *Instructor. Manila, Philippines*.  
Handled the ff. courses: C programming, software engineering, algorithms and data structures. Courses were taught primarily to an audience with minimal computing background.
- 2008–2009 **Interprise Solutions**, *Software Engineer. Cebu, Philippines*.  
Designed and developed tools used in the application development framework of an ERP system.
- spring of **Neri and Hu Design and Research Office**, *Software Engineer Intern. Shanghai, China*.  
2006 Designed and implemented an in-house inventory management system.

## Publications

- [1] N. K. Ahmed, R. A. Rossi, **J. B. Lee**, T. L. Willke, R. Zhou, X. Kong, & H. Eldardiry. Role-based graph embeddings. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*. 2021. (to appear)
- [2] H. Yin, **J. B. Lee**, X. Kong, T. Hartvigsen, & S. Xie. Energy-efficient models for high-dimensional spike train classification using sparse spiking neural networks. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. 2021.
- [3] X. Dai, X. Kong, T. Guo, **J. B. Lee**, X. Liu, & C. Moore. Recurrent networks for guided multi-attention classification. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. 2020.
- [4] R. A. Rossi, D. Jin, S. Kim, N. K. Ahmed, D. Koutra, & **J. B. Lee**. On proximity and structural role-based embeddings in networks: Misconceptions, techniques, and applications. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 14, 5. 2020.
- [5] **J. B. Lee**, X. Kong, C. M. Moore, & N. K. Ahmed. Deep parametric model for discovering group-cohesive functional brain regions. In *Proc. of SIAM International Conference on Data Mining (SDM)*. 2020.
- [6] X. Dai, X. Kong, X. Liu, **J. B. Lee**, & C. Moore. Dual-attention recurrent networks for affine registration of neuroimaging data. In *Proc. of SIAM International Conference on Data Mining (SDM)*. 2020.
- [7] **J. B. Lee**, R. A. Rossi, S. Kim, N. K. Ahmed, & E. Koh. Attention models in graphs: A survey. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13, 6. 2019.
- [8] **J. B. Lee**, R. A. Rossi, X. Kong, S. Kim, E. Koh, & A. Rao. Graph convolutional networks with motif-based attention. In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)*. 2019.
- [9] **J. B. Lee** & X. Kong. Learning compact graph representations via an encoder-decoder network. *Applied Network Science (ANS)* 5, 1. 2019.
- [10] G. Nguyen, **J. B. Lee**, R. Rossi, N. Ahmed, E. Koh, & S. Kim. Dynamic network embeddings: From random walks to temporal random walks. In *Proc. of IEEE International Conference on Big Data (BigData)*. 2018.
- [11] **J. B. Lee**, R. A. Rossi, & X. Kong. Graph classification using structural attention. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. 2018.
- [12] G. Nguyen, **J. B. Lee**, R. Rossi, N. Ahmed, E. Koh, & S. Kim. Continuous-time dynamic network embedding. In *Comp. Proc. of International Conference on World Wide Web (WWW)*. 2018.
- [13] **J. B. Lee**, X. Kong, Y. Bao, & C. Moore. Identifying non-linear contrasting networks from time-series data: Application to brain network analysis. In *Proc. of SIAM International Conference on Data Mining (SDM)*. 2017.
- [14] **J. B. Lee**, A. Ihara, A. Monden, & K. Matsumoto. Patch reviewer recommendation in OSS projects. In *Proc. of Asia-Pacific Software Engineering Conference (APSEC)*. 2013.
- [15] **J. B. Lee**, M. Ybañez, M. M. de Leon, and M. R. E. Estuar. Understanding the behavior of Filipino Twitter users during disaster. *Journal on Computing (JoC)* 3, 2. 2013.
- [16] **J. B. Lee** & M. Ybañez. Characterizing behavior and features of participants and observers during disaster on Twitter. In *Proc. of International Conference on Computer Games Multimedia and Allied Technologies (CGAT)*. 2013.
- [17] **J. B. Lee** & H. Adorna. Link prediction in a modified heterogeneous bibliographic network. In *Proc. of International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 2012.

- [18] **J. B. Lee**, G. Cabunducan, R. Castillo, F. G. Cabarle, & J. A. Malinao. Uncovering the social dynamics of online elections. *Journal of Universal Computer Science (JUCS)* 18, 4. 2012.
- [19] J. A. Malinao, R. A. B. Juayong, E. R. F. Oquendo, R. M. U. Tadas, **J. B. Lee**, J. B. Clemente, M. S. Gaabucayan-Napalang, J. R. F. Regidor, & H. N. Adorna. A quantitative analysis-based algorithm for optimal data signature construction of traffic data sets. *Journal of Information Processing (JIP)* 20, 3. 2012.
- [20] G. Cabunducan, R. Castillo, & **J. B. Lee**. Voting behavior analysis in the election of Wikipedia admins. In *Proc. of International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 2011.
- [21] J. A. Malinao, R. A. B. Juayong, E. R. F. Oquendo, R. M. U. Tadas, **J. B. T. Lee**, J. B. Clemente, M. S. Gaabucayan-Napalang, J. R. F. Regidor, and H. N. Adorna. A quantitative analysis-based algorithm for optimal data signature construction of traffic data sets. In *Proc. of International Conference on Computers, Networks, Systems, and Industrial Engineering (CNSI)*. 2011.

## Professional service

PC member	CIKM ('20), BigData ('18, '20)
Reviewer	TKDD ('17, '21), TNNLS ('21), VLDB Journal ('20), TKDE ('17), & JUCS ('12)
Ext. reviewer	KDD ('16, '17, '18, '19), ICLR ('18, '19), IJCAI ('17, '19), ICDM ('19), SDM ('17, '18, '19), NeurIPS ('16, '18), ASONAM ('17, '18), ICDE ('17), AISTAT ('17), CIKM ('16), & AAAI ('16)

## Awards

2018	KDD Travel Award, ACM & NSF
2018	GRIE Best Poster Honorable Mention, WPI CS DEPT.
2013	Best Presentation Award, IWESEP
2012	Honors Graduate, PHI KAPPA PHI
2011	Best Poster Paper, ASONAM

## Computer skills

Programming	Python, Java, C#, C, $\LaTeX$ , & SQL (Presto, Spark)
Frameworks	PyTorch, TensorFlow, Airflow, scikit-learn, NumPy, SciPy, & Matplotlib
Tools	Jupyter Notebook, VS Code, GNU Screen, Virtualenv, Git, & Mercurial

## Languages

Native	English, Tagalog, & Cebuano
Intermediate	Mandarin

## References

A full list of references is available upon request.