

Vaibhav Kumar

Language Technologies Institute

Education

- 2013 – 2018 **B.Tech in Computer Science and Master of Science in Computational Linguistics by Research**, International Institute of Information Technology, Hyderabad, India, **CGPA – 9.09/10**.
- 2018 – Now **Master of Language Technologies**, Carnegie Mellon University, Pittsburgh, USA.

Research Interests

Natural Language Processing: Information Retrieval, Text Mining and Analytics, Natural Language Understanding

Machine Learning: Deep Learning for Natural Language Processing

Publications

- 2018 **Dhruv Khattar, Vaibhav Kumar***, Manish Gupta, and Vasudeva Varma. "*Weave&Rec : A Word Embedding based 3-D Convolutional Network for News Recommendation*" In Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018.
- 2018 **Dhruv Khattar, Vaibhav Kumar***, Manish Gupta, and Vasudeva Varma. "*HRAM : A Hybrid Recurrent Attention Machine for News Recommendation*" In Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018.
- 2018 **Dhruv Khattar, Vaibhav Kumar***, Shashank Gupta, Manish Gupta, and Vasudeva Varma. "*RARE : A Recurrent Attentive Recommendation Engine for News Aggregators*" In Proceedings of the 27th ACM International Conference on Information and Knowledge Management Workshops. 2018.
- 2018 **Mrinal Dhar, Vaibhav Kumar, Manish Shrivastava**, "*Enabling Code-Mixed Translation: Parallel Corpus Creation and MT Augmentation Approach*" In Linguistic Resources for NLP organised as a part of COLING Workshops. 2018.
- 2018 **Vaibhav Kumar**, Mrinal Dhar, Dhruv Khattar, Yash Kumar Lal, Abhimanshu Mishra, Manish Shrivastava, Vasudeva Varma. "*SWDE: A Sub-Word And Document Embedding Based Engine for Clickbait Detection*" In Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval Workshops. 2018.
- 2018 **Vaibhav Kumar**, Dhruv Khattar, Siddhartha Gairola, Yash Kumar Lal and Vasudeva Varma. "*Identifying Clickbait: A Multi-Strategy Approach Using Neural Networks*" In Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval. 2018 (**Received Microsoft Research Travel Grant**).
- 2018 **Dhruv Khattar, Vaibhav Kumar***, Manish Gupta, and Vasudeva Varma. "*Neural Content-Collaborative Filtering for News Recommendation*" In Proceedings of Second International Workshop on Recent Trends in News Information Retrieval co-located with 40th European Conference on Information Retrieval. 2018.
- 2017 **Vaibhav Kumar**, Dhruv Khattar, Shashank Gupta, Manish Gupta, and Vasudeva Varma. "*User Profiling based Deep Neural Network for Temporal News Recommendation*" In Proceedings of 2017 IEEE 17th International Conference on Data Mining Workshops. 2017 (**Received Microsoft Research Travel Grant**).
- 2017 **Vaibhav Kumar**, Dhruv Khattar, Shashank Gupta and Vasudeva Varma. "*Word Semantics based 3D Convolutional Neural Networks for News Recommendation*" In Proceedings of 2017 IEEE 17th International Conference on Data Mining Workshops. 2017 (**Received Microsoft Research Travel Grant**).
- 2017 **Vaibhav Kumar**, Dhruv Khattar, and Vasudeva Varma. "*Leveraging Moderate User Data for News Recommendation*" In Proceedings of 2017 IEEE 17th International Conference on Data Mining Workshops. 2017 (**Received Microsoft Research Travel Grant**).

- 2017 **Vaibhav Kumar**, Dhruv Khattar, Shashank Gupta, Manish Gupta, and Vasudeva Varma. "Deep Neural Architecture for News Recommendation." In Working Notes of the 8th International Conference of the CLEF Initiative, Dublin, Ireland. CEUR Workshop Proceedings. 2017.

Work Experience

- 2017 Aug-Nov **Teaching Assistant**, INFORMATION RETRIEVAL AND EXTRACTION, IIIT-H.
2017 Jan-Apr **Teaching Assistant**, STATISTICAL METHODS IN ARTIFICIAL INTELLIGENCE, IIIT-H.
2016 Aug-Nov **Teaching Assistant**, STATISTICAL METHODS IN ARTIFICIAL INTELLIGENCE, IIIT-H.
2016 May-July **Research Intern**, DEPARTMENT OF COMPUTER SCIENCE, CITY UNIVERSITY LONDON.
2015 May-July **Technical Intern**, VEOOZ.COM.

Projects

Research Projects

- 2017 **Non-Recurrent Models for Extracting Temporal Changes in Users Interests**, NLP/IR, ML.
Started exploring non-recurrent models for capturing temporal changes in users interests by using information based on word semantics. Came up with a setting of the 3D Convolutional Neural Network in order to tackle the problem.
- 2017 **Multi-strategy approach for Clickbait Identification**, NLP/IR, ML.
Came up with a hybrid approach which uses word level attention in combination with image features to classify clickbaits.
- 2017 **Recurrent Attention Model for News Recommendation**, NLP/IR, ML.
Came up with a recurrent network capable of handling temporal and diverse interests of users and generating recommendations. The method was also capable of handling user and item cold start problem respectively. Also, working on coming up with a hybrid variant of the model.
- 2016 **Leveraging Moderate User Data for News Recommendation**, NLP/IR, ML.
Came up with an Item-based Collaborative Filtering Approach using Markov Decision Process to generate recommendations. Also, came up with a semantic measure to capture diverse interests of users.
- 2016 **Deep Learning Algorithm for Sequence Labeling Tasks**, ML, NLP.
As a Research Intern at City University London, the project was to come up with a deep learning algorithm capable of assigning a categorical label to each member of a sequence of observed values. Worked on development of Recurrent Temporal Discriminative Restricted Boltzmann Machine.
- 2015 **Automatic Essay Grading**, MACHINE LEARNING.
An essay grader was built that would automatically score essays as close as possible to that which would be given by a human grader. Different features like POS tags, word diversity, title-text similarity were explored. Various methods like Support Vector Regression, Random Decision Forests were explored for building the grader.

Other Projects

- 2016 **Web Proxy with Caching Capabilities**, ADVANCED COMPUTER NETWORKS.
A Web Proxy capable of accepting HTTP Requests was implemented. Multiple Client Requests were handled concurrently. An LRU based cache was implemented for the same.
- 2016 **Calculating Sentence Similarity from Word Alignment**, NLP APPLICATIONS, Team of 3.
Word Similarity, contextual Similarity and dependency context were used for this purpose.
- 2016 **Text Processing Framework for Hindi**, INFORMATION RETRIEVAL AND EXTRACTION, Team of 3.
Framework included methods for tokenization, stop word detection, sentence breaking, POS tagging, keyword identification, entity recognition
- 2016 **Search Engine for Wikipedia**, INFORMATION RETRIEVAL AND EXTRACTION.
A search engine was made over a 40GB of Wikipedia dump.
- 2015 **Simulation of Key Management Protocol in WSN**, SYSTEMS AND NETWORK SECURITY.
Direct and path key establishment phases were simulated for a wireless sensor network.
- 2015 **File Sharing System**, SYSTEMS AND NETWORK SECURITY / NETWORKS.
A peer to peer file sharing system was implemented using RSA encryption scheme and SHA-1 in C.
- 2015 **ULTIMATE 9x9 Tic Tac Toe**, ARTIFICIAL INTELLIGENCE, Team of 2.
A program was designed which suggested the best possible move required to win the game of 9x9 tic tac toe.
- 2015 **Carron - A 2D board game**, GRAPHICS.
A multiplayer carron was built using OpenGL 2 which allowed different modes of game play.
- 2014 **Water Leakage Detection System**, SOFTWARE SYSTEMS, Team of 4.
Developed a software for detecting leaks in a network of pipes.

2014 **Hospital Database Application**, APPLICATION, Team of 2.

An efficient database was designed and developed for managing and administrating a hospital.

2013 **CloudHub**, APPLICATION, Team of 3.

The application allowed users to share photos, files, music and videos with people.

Programming Languages

C, C++, Python, PHP

Awards

2013-17 Deans Academic List