Dhruv Kumar

(+1) 226-9783751 • ddhruvkr@gmail.com • Github • Website • LinkedIn

EDUCATION

University of Waterloo, Ontario, Canada — Sep 2018- Present
M. Math.(Thesis) in Computer Science — GPA: 3.88/4
Coursework- Deep Learning, Trust Modeling and Online Social Networks, Blockchains: Foundations and Applications, Privacy and Fairness in Data Science

Indian Institute of Information Technology, Allahabad, U.P., India — Jul 2012- Jun 2016
B. Tech.(Hons) in Information Technology — GPA: 8.68 / 10.00 (9.07 - Last 2 years)
Relevant Coursework- Information Retrieval, Artificial Intelligence, Natural Language Processing, Cognitive Process Modelling, Optimization Techniques, Data Mining, Probability & Statistics, Mathematics(I, II, III).

PUBLICATIONS

- "**D. Kumar**, L. Mou, L. Golab, O. Vechtomova, Iterative Edit-based Unsupervised Sentence Simplification", *In Proceedings of the 58th annual meeting of the Association for Computational Linguistics (ACL 2020)* - Long Paper
- "R. Cohen, R. Agarwal, **D. Kumar**, A. Parmentier, T. H. Leung, Sensitivity to risk profiles of users when developing AI systems", *33rd Canadian Conference on Artificial Intelligence (Canadian AI 2020)*
- "R. Agarwal* , **D. Kumar***, L. Golab, S. Keshav, Consentio: Managing Consent to Data Access using Permissioned Blockchains", *IEEE International Conference on Blockchain and Cryptocurrency (ICBC) 2020* - Full Paper
- "**D. Kumar**, R. Cohen, L. Golab, Online abuse detection: the value of preprocessing and neural attention models", in *10th workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA) at NAACL-HLT 2019*
- "N. Kushwaha, S. Mehrotra, R. Kalia, **D. Kumar**, O. P. Vyas, Inclusion of Semantic and Time-Variant Information Using Matrix Factorization Approach for Implicit Rating of Last.Fm Dataset," in *Springer Arabian Journal of Science and Engineering*, May 2016.

UNDER REVIEW / TECHNICAL REPORTS

- "O. Vechtomova, H. Bahuleyan, G. Sahu, **D. Kumar**, A. Ghabussi, V. John, Music conditioned lyrics generation (name changed)", Under Review
- "**D. Kumar**, Compressed Knowledge transfer via Factorization Models in Recommender Systems", Bachelor's Thesis, 2016

RESEARCH EXPERIENCE AND PROJECTS

University of Waterloo

Master's Thesis — Jul 2019-Present
- "Text simplification": Designed an unsupervised algorithm building on edit-based text generation techniques for sentence simplification. The model is highly adaptable and can achieve SARI and BLEU scores of **30.44** and **26.47** respectively on the Newsela dataset. Currently, adding neural paraphrasing module to the approach.
  [Supervisors: Prof. Lukasz Golab, Prof. Olga Vechtomova]

Course Project — Oct 2019-Dec 2019
- "Unsupervised neural models to anonymize personal attributes in human generated text": Text written by humans contains implicit linguistic information that can be used to identify our attributes such as gender, age and political leanings. Worked with the PASTEL and Yelp datasets to build an unsupervised style transfer model.
  [Supervisor: Prof. Xi He]

Research Assistant/Natural Language Processing Lab — Sep 2019-Present
- "Music conditioned lyrics generation": Designed bimodal neural network models to generate lines of lyrics using music audio clips. The model can a) learn an artist's musical style from audio clips and generate lines in the style of an artist, and (b) generate lines conditioned on any given short audio clip. This work is under review at a conference.
  [Supervisor: Prof. Olga Vechtomova]

Research Assistant/Sirius Blockchain Research Group — Jan 2019-Aug 2019
- Designed Consentio, a permissioned blockchain-based consent management system (CMS) that can handle up to **6000** transactions per second.
  [Supervisors: Prof. Lukasz Golab, Prof. Srinivasan Keshav]

Research Assistant — Oct 2018-Feb 2019

- Proposed a co-attention based neural model to detect implicit aggressiveness and incivility in online social networks. The model achieves a F1 score of **82.41**, **77.75** and **76.07** for the minority abuse class on the Wikipedia toxicity/attack/aggression datasets respectively.
  [Supervisors: Prof. Robin Cohen, Prof. Lukasz Golab]

Independent Project — Dec 2018-Jan 2019

- Implemented various state of the art deep learning models (e.g. Co-attention, Self-attention, Hierarchical attention) for text classification. The models were inspired by research conducted for different tasks in NLP.

Course Project — Oct 2018-Dec 2018

- Implemented the Bi-Directional Attention Flow (BiDAF) model from scratch over the SQuAD1.1 dataset.
  [Supervisors: Prof. Ali Ghodsi]

Universität Paderborn, Paderborn, Germany — Jan 2016-Jun 2016

Studentische Hilfskraft (Student RA) at Analytic Information Systems and Business Intelligence Lab (AIS-BI)

- Bachelor's Thesis: "Compressed Knowledge transfer via Factorization Models in Recommender Systems". Developed an algorithm to incorporate metadata in Factorization Machines, achieving the RMSE value of **0.836** as compared to **0.853** when using a Joint Matrix Factorization method on the MovieLens 1M dataset.
  [Supervisors: Dr. Artus Krohn-Grimberghe, Prof. Ratna Sanyal]
- Face recognition in videos: Built a web application on top of Microsoft's Project Oxford.
  [Supervisor: Dr. Artus Krohn-Grimberghe]

Indian Institute of Information Technology, Allahabad, India — Jul 2015-Dec 2015

Research Assistant

- "Analysis of Time-Aware and Semantic Feature Based Music Recommender System". Proposed a Joint Matrix Factorization algorithm for Music Recommendation that utilizes geographical and time-based tagging information of artists, in addition to implicit user feedback (user clicks) to provide better recommendations.
  [Supervisor: Dr. O.P.Vyas]

TEACHING ASSISTANT EXPERIENCE

- CS 115: Introduction to Computer Science, Fall 18, University of Waterloo.
- CS 230: Introduction to Computers and Computer Systems, Winter 19, University of Waterloo.
- CS 231: Algorithmic Problem Solving, Spring 19, University of Waterloo.
- CS 241: Foundations of Sequential Programs, Fall 19, University of Waterloo.
- CS 489/698: Topics in Computer Science, Neural Networks, Winter 20, University of Waterloo.

SERVICE

- **Program Committee / Reviewer** for NeurIPS2019 AISG Workshop

INDUSTRY EXPERIENCE

Arcesium(DE Shaw Group), Hyderabad, India — Jul 2016-May 2018

Software Engineer, Fund and Investor Accounting

- Enhanced the post-trade automation platform for funds operated by J.P. Morgan and D.E. Shaw.

Citigroup, Pune, India

Software Engineering Intern, Equities — May 2015-Jul 2015

- Designed the first prototype of the Trading Controls application. Declined the full-time offer.

SKILLS

Languages- C, C++, Python, Java, Javascript, HTML, PHP and SQL.
Libraries- Pytorch, Tensorflow, Keras, Gym, Scikit-Learn, Pandas, and d3.js
Frameworks- AngularJS, Node.js, MongoDB, Flask, Spring, MyBatis and Struts

ACHIEVEMENTS & EXTRA-CURRICULAR

- Got Graduate Studies Research Travel Assistantship to attend the NAACL-HLT Conference, 2019.
- Accepted in the Deep Learning and Reinforcement Learning Summer School, 2019 (less than 25% acceptance rate).
- Received International Masters Student Award and University of Waterloo Entrance Scholarship for graduate studies.
- Stood in the top 0.5% in the All India Engineering Entrance Examination 2012.
- Worked as the Events Head of the annual cultural cum technical festival *Effervescence* 2014.
- Represented the college band as a drummer.