

# 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: 87 / 100

Coursework- Deep Learning, Trust Modeling and Online Social Networks, Blockchains: Foundations and Applications, Privacy and Fairness in Data Science (Ongoing)

**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**, 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 Berlin Heidelberg - Arabian Journal of Science and Engineering*, May 2016.

## UNDER REVIEW / PREPRINT / MANUSCRIPTS

- **“R. Agarwal\* , D. Kumar\***, L. Golab, S. Keshav, Consentio: Managing Consent to Data Access using Permissioned Blockchains”, arXiv, 2019 (Under Review)
- **“D. Kumar**, Compressed Knowledge transfer via Factorization Models in Recommender Systems”, Bachelor’s Thesis, 2016

## RESEARCH EXPERIENCE

**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.  
[Supervisors: Prof. Lukasz Golab, Prof. Olga Vechtomova]

Course Project

Oct 2019-Present

- **“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. Working with the PASTEL dataset that contains 39,778 sentences spanning 7 such attributes to build a unsupervised style transfer model.  
[Supervisor: Prof. Xi He]

Research Assistant/Natural Language Processing Lab

Sep 2019-Present

- **“Music conditioned lyrics generation”**: Designed a Variational autoencoder (VAE) that generates appropriate lyrics conditioned on feedback from a multimodal classifier. A modified beam search algorithm controls the effect of the classifier.  
[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]

|                                 |   |
|---------------------------------|---|
|                                 | <b>Universität Paderborn</b> , Paderborn, Germany         Jan 2016-Jun 2016<br>Studentische Hilfskraft (Student RA) at Analytic Information Systems and Business Intelligence Lab (AIS-BI) <ul style="list-style-type: none"> <li>▪ 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 <b>0.836</b> as compared to <b>0.853</b> when using a Joint Matrix Factorization method on the MovieLens 1M dataset. [Supervisors: Dr. Artus Krohn-Grimberghe, Prof. Ratna Sanyal]</li> <li>▪ Face recognition in videos: Built a web application on top of Microsoft's Project Oxford. [Supervisor: Dr. Artus Krohn-Grimberghe]</li> </ul> |
|                                 | <b>Indian Institute of Information Technology</b> , Allahabad, India         Jul 2015-Dec 2015<br>Research Assistant <ul style="list-style-type: none"> <li>▪ "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]</li> </ul>   |
| TEACHING ASSISTANT EXPERIENCE   | <ul style="list-style-type: none"> <li>▪ CS 115: Introduction to Computer Science, Fall 18, University of Waterloo.</li> <li>▪ CS 230: Introduction to Computers and Computer Systems, Winter 19, University of Waterloo.</li> <li>▪ CS 231: Algorithmic Problem Solving, Spring 19, University of Waterloo.</li> <li>▪ CS 241: Foundations of Sequential Programs, Fall 19, University of Waterloo.</li> <li>▪ CS 489/698: Topics in Computer Science, Neural Networks, Winter 20, University of Waterloo.</li> </ul>  |
| SERVICE                         | <ul style="list-style-type: none"> <li>▪ <b>Program Committee / Reviewer</b> for NeurIPS2019 AISG Workshop</li> </ul>   |
| INDUSTRY EXPERIENCE             | <b>Arcesium(DE Shaw Group)</b> , Hyderabad, India         Jul 2016-May 2018<br>Software Engineer <ul style="list-style-type: none"> <li>▪ Worked as a software engineer, designing Fund and Investor Accounting solutions for Hedge Funds.</li> <li>▪ Member of the team to onboard J.P.Morgan onto the platform.</li> </ul> <b>Citigroup</b> , Pune, India<br>Software Engineering Intern         May 2015-Jul 2015 <ul style="list-style-type: none"> <li>▪ Trading Controls application for Equities Team using MEAN stack and d3. Received a full-time offer.</li> </ul>  |
| SKILLS                          | <b>Languages-</b> C, C++, Python, Java, Javascript, HTML, PHP and SQL.<br><b>Libraries-</b> Pytorch, Tensorflow, Keras, Gym, Scikit-Learn, Pandas, and d3.js<br><b>Frameworks-</b> AngularJS, Node.js, MongoDB, Flask, Spring, MyBatis and Struts   |
| ACHIEVEMENTS & EXTRA-CURRICULAR | <ul style="list-style-type: none"> <li>▪ Got Graduate Studies Research Travel Assistantship to attend the NAACL-HLT Conference, 2019.</li> <li>▪ Accepted in the Deep Learning and Reinforcement Learning Summer School, 2019 (less than 25% acceptance rate).</li> <li>▪ Received International Masters Student Award and University of Waterloo Entrance Scholarship for graduate studies.</li> <li>▪ Stood in the top 0.5% in the All India Engineering Entrance Examination 2012.</li> <li>▪ Worked as the Events Head of the annual cultural cum technical festival <i>Effervescence</i> 2014.</li> <li>▪ Represented the college band as a drummer.</li> </ul>  |