

# PRANJAL GUPTA

139A, Columbia St W, Waterloo N2L 3L2, ON, Canada | +1 (437) 234 1197 | pranjal.gupta@uwaterloo.ca

 <http://g31pranjal.github.io>  <https://github.com/g31pranjal>  <https://www.linkedin.com/in/g31pranjal/>

---

## OBJECTIVE

I am looking for full-time jobs as a Software developer preferably in research prototyping, systems design and prototyping, Database Systems, Graph-based systems and big-data management. I have an excellent breadth across domains of Computer Science and depth in Database and data-management systems.

## EDUCATION

**M. Math (thesis-tracked) in Computer Science, University of Waterloo, Canada** (September 2019 - Present)

Advisor *Prof. Semih Salihoglu*

- CGPA: 89.25/100
- Courses: Graph Databases: Growth, Advances and , Advanced Operating Systems, Advanced Distributed Systems, Machine learning for Data Cleaning.

**B.E. (Honours) in Computer Science, BITS Pilani, India** (August 2013 - May 2018)

Dual degree in Mathematics. Bachelor-thesis advisor *Prof. Poonam Goyal*

- CGPA: 9.06/10 (Passed with distinction)
- Relevant Courses: Graph Theory, Database Systems, Information Retrieval, Data Mining, Artificial Intelligence, Operating Systems, Algorithms, Analysis of Algorithms, Object-Oriented Programming and Computer Networks.

## EXPERIENCES

**Data Systems Group, Cheriton School of Computer Science, University of Waterloo, Canada** (September 2019 - Present)

Graduate Research Assistant

- Working on [Graphflow](#), a single-node in-memory property graph database system.
- I look at scaling graphflow on single-node environment by minimizing it's memory requirements without degrading performance. My research is on property storage data models, compression techniques for data and indices.

**Advanced Data Analytics and Parallel Technologies Lab, BITS Pilani, India** (January 2018 - July 2018)

Undergraduate Research Assistant

- Implemented a system detecting temporal events from a stream of tweets and creating a multi-level hierarchy of events.
- Implemented an abstractive summarization attention-based LSTM neural network model for generating synopsis and story out of the hierarchy of events.
- The work is submitted for review in IEEE Transactions on Computational Social Systems.

**Mitacs Globalink Research Internship at the University of Manitoba, Winnipeg, Canada** (May 2017 - July 2017)

Worked under the supervision of *Prof. Carson Kai-Sang Leung* in the [Database and Data Mining Lab](#)

- Developed an algorithm, SWITCH, that switched between classical algorithms for vertical frequent pattern-mining ECLAT, VIPER and DECLAT, with an objective to minimize the overall memory footprint of the process.
- Created relevant data structures and adopted ECLAT and VIPER algorithms for mining uncertain datasets, viz. uECLAT and uVIPER. Extended the SWITCH algorithm on the uncertain dataset.

**Zomato Media Pvt. Ltd, Gurgaon, India** (May 2016 - July 2016)

Intern in the Data Analytics team.

- Implemented Computational A/B Experimental framework and other statistical tools for evaluating metrics on the web and mobile applications for Zomato.
- Used probabilistic approaches for predicting network models on Zomato's user search data to discover overlapped communities of popular locations for food in prime Indian cities.

**Homi Bhabha Centre for Science Education, TIFR, Mumbai, India** (May 2015 - July 2015)

Worked under the guidance of *Dr. G. Nagarjuna* in the Knowledge Lab

- Worked on [GNU Gnowsys](#), an in-memory RDF Knowledge-base management system.
- Implemented Analytics, Data-visualization tools and RSS update feeds on MetaStudio, a studio-based learning platform, built on top of the Gnowsys kernel. [[🔗](#)]

## PUBLICATIONS

1. **Multilevel Event Detection, Storyline Generation Summarization for Tweet Streams.** Poonam Goyal, Prerna Kaushik, Pranjal Gupta, Dev Vashisth, Shavak Agarwal and Navneet Goyal.  
*IEEE Transactions on Computational Social Systems, May 2019*

## ACADEMIC PROJECTS

### Experimental evaluation of functions and MapReduce on serverless computing environment [🔗] (February 2019 - April 2019)

- Benchmarked performance metrics for functions running on serverless infrastructure using open-faas, Kubernetes and Docker containers. We studied system latency, I/O overhead and throughput, compute overhead and scalability.
- Implemented a simple MapReduce task on serverless infrastructure by designing an execution pipeline of triggers, coordinator functions, mappers and reducers. Studied variations by tweaking the parameters.

### Discovering SHACL constraints on RDF Datasets[🔗] (February 2019 - May 2019)

Worked with [Prof. Ihab Ilyas](#)

- Worked on a running project on discovering constraints in RDF dataset using **SHApe Constraint Language**.
- My contributions include enhancing node feature discovery, optimizing search algorithm by pruning the constraint space, ranking discovered constraints by estimating their relevance and doing the empirical evaluations on the algorithm.

### Constructing Image captions from contextual information and features using Neural Nets (January 2017 - December 2017)

Worked as a Research Assistant under the supervision of [Prof. Poonam Goyal](#) in the [WISoC Lab](#).

- Designed a Recurrent neural network model based on the statistical probability for predicting the most appropriate caption for an image, using auxiliary contextual information; inspired by `encoder-decoder` machine translation models.

### Basic Compiler for a toy language [🔗] (February 2017 - April 2017)

- Developed a compiler for a C-like imperative toy language.
- Incorporated and integrated modules for lexical analysis, top-down parsing of LL(1) grammar, Abstract Syntax Tree construction, Symbol tree generation, Type checking, Semantic analysis, and NASM code generation.

### Multiplayer CLI-based game of Hearts over a TCP network [🔗] (March 2017 - April 2017)

- Developed a game of cards on the server-client paradigm that follows a round-robin flow of control over TCP protocol.
- Implemented a server module to listen for incoming connections and spawning a game-room per 4 client connections. Also, implemented the client module as a DFA, driven by commands from the game room. The client module interfaced to the user for I/O.

### Wikie: the retrieval system [🔗] (October 2016 - November 2016)

- Developed a search engine to retrieve and rank Wikipedia pages, based on the vector-space model.
- The ranking mechanism employed PageRank scoring (for measuring the intrinsic importance of a page) and Elo Ratings (for measuring user popularity of the page), to calculate relevance and importance of a webpage.

### Modelling of Natural Language sentences using SPN Graphs [🔗] (November 2016)

- Developed a scheme for parsing natural language sentences and represented them in a graph structure, called SPN Graphs, thereby forming an overall connected graph for a particular set of sentences. The scheme parsed sentences based on their construct comprising an agent, action and patient.
- Developed a question-answering system on top of the described model.

## AWARDS, ACHIEVEMENTS and SCHOLARSHIPS

- **2018:** Received International Masters Student Award and University of Waterloo's Graduate Scholarship for pursuing research-based Master's.
- **2018:** Received MITACS Graduate Research Fellowship for pursuing research in Canada.
- **2018:** Ranked 1st in the batch of 2013 of M.Sc.(Hons) Mathematics at BITS Pilani.
- **2017:** Received the award for `Best Student` in the batch of 2013 of M.Sc. (Hons) Mathematics.
- **2015:** Received 2nd prize at the technical festival of BITS Pilani for the project "Automated Personal Email Classifier" in the "Software Development - Machine Learning Applications" category.
- **2013:** Recipient of MCN Merit Scholarship at BITS Pilani for being in the top 4% of the merit list for all the 10 terms.

## TEACHING ASSISTANTSHIPS

- **Fall 2019, Summer 2019, Winter 2019:** CS 338 (Computer Application in Business: Databases) at UWaterloo.

- **Fall 2018:** CS 136 (Elementary Algorithm Design and Data Abstraction) at UWaterloo.
- **Fall 2017:** CS F415 (Data Mining) at BITS Pilani.
- **Fall 2016:** CS F213 (Object-oriented programming) at BITS Pilani.

## TECHNICAL SKILLS

**Programming Languages:** Java, Python, C/C++, Prolog, Scala.

**Databases:** SQL, Cypher - Graph database query language, Noe4j, Oracle 12c.

**Machine Learning Libraries:** Keras, TensorFlow, sci-kit learn, numpy and pandas.

**Distributed Frameworks:** Spark, Kubernetes, Docker, open-faas.

**Web Development:** HTML5/CSS3, JavaScript, PHP, Django Framework, jekyll.

**Testing:** gdb, valgrind, junit, mockito.

**Software/IDE:** IBM SPSS Modeller, Idea Java IDE, Microsoft Visual Studio Code, PyCharm