Anmol Singh

Experience

Aug 2019 - Dec 2021 | Research Assistant at **University of Waterloo**, Ontario
Data Systems Group
Advisor: Dr. Gordon Cormack
Research on efficient construction of test collections for the evaluation of information retrieval systems. Implemented a document review system which processed the entire ClueWeb12 dataset (730 million documents, 27.3 TB uncompressed) to find relevant documents using active learning and statistical sampling. Created and evaluated a statistical test collection for the TREC 2019 Medical Misinformation Track.

May 2017 - May 2019 | Software Engineer (Machine Learning) at **LinkedIn**, Bangalore
Relevance Team
Manager(s): Rushi Bhatt, Ram Madhavan
Worked on a team developing an end-to-end machine learning platform. Conceptualized and led the design and implementation of data ingestion, dataset versioning and human-in-the-loop annotation systems. Other projects include a Spark library to expose native ML tooling on Jupyter notebooks and a high performance in-memory key-value store to serve features at inference time.

May-July 2016 | Software Engineering Intern at **LinkedIn**, Bangalore
Relevance Team
Mentor: Dr. Sambuddha Roy
Project on prototyping a tool to increase manual labeling throughput by clustering images. Experimented with k-means and spectral clustering on AlexNet features. Additionally, contributed to literature reviews on active learning.

May-July 2015 | Summer Research Intern at **IIIT-Delhi**
Crowd Behaviour Test Bed
Mentor: Dr. Sachit Butail
Project on simulating and analyzing human behaviour in virtual crowds by tracking the movements of a person wearing a head mounted VR display. Responsible for the tracking part. Experience with Processing IDE and OpenNI API for the Microsoft Kinect.

Education

2019 - 2021 | M.Math (Thesis) in Computer Science
University of Waterloo, Ontario
Gpa: 91.4/100

2013 - 2017 | Bachelor of Technology in Computer Science and Engineering
IIIT-Delhi, New Delhi
Gpa: 8.76/10

Coursework and Skills

Selected Courses *UWaterloo*: Optimization for Data Science, Information Retrieval, Experimental Methods in HCI
IIIT-Delhi: Probabilistic Graphical Models, Machine Learning, GPU Computing, Linear Optimization, Randomized Algorithms, Artificial Intelligence, Modern Algorithm Design

Programming Skills *Languages*: Python, C/C++, Java, Scala

Tools and Libraries:
numpy, scikit-learn, pandas, *nix (Intermediate)
Django, Flask, Docker, Apache Spark, Apache Samza (Basic)

Selected Projects

May - Aug 2020 | Label-and-Learn for High-Recall Information Retrieval

Course Project, Experimental Methods in HCI | Mentor: Dr. Edward Lank
Task: Visualizations to assist document review for high-recall retrieval tasks. Implemented term highlighting, test set scatter plots and top model terms, with the intent of reducing the mental workload of human reviewers and getting accurate models faster. Experience in developing visualizations for text classifiers and conducting and analysing user studies.

Jan - April 2017 | Scalable Energy Breakdown

Independent Project | Mentor: Dr. Nipun Batra
Task: Building models to generate appliance level energy breakdown from just the monthly bill and static house features. Applied KNN, KNN with metric learning, matrix factorization and factorization machines. Experience in designing and implementing machine learning experiments and applying collaborative filtering techniques.

Feb - April 2016 | Parallel Real Time Ray-Tracing

Course Project, GPU Computing | Mentor: Dr. Ojaswa Sharma
Massively multi-threaded GPU implementation of a ray tracer in CUDA C. Implemented reflection, refraction and shadowing. Accelerated ray triangle intersection tests using a Uniform Grid data structure. Achieved 120x speed up over serial CPU implementation.

Sept - Dec 2015 | Scrabble Bot

Course Project, AI | Mentor: Dr. Sandip Aine
Implemented efficient algorithms to generate and score all possible plays given a state of a Scrabble board and a rack using a backtracking approach on tries. Implemented game tree search and Monte Carlo simulations to rank the highest scoring moves playable in a game state. Delivered a fully playable application written in Python.

Awards and Achievements

2019 - 2021 David R. Cheriton Graduate Scholarship, University of Waterloo
2019 - 2021 International Master's Award of Excellence, University of Waterloo
2016 ACM-ICPC Rank 17/391 teams, Amritapuri Regionals; 21/64 teams, India Finals
2015 - 2016 Dean's List, IIIT-Delhi

Positions of Responsibility

2019 - 2021 Teaching Assistant, CS246 (Object-Oriented Software Development)
Fall 2019, Winter 2020, Spring 2020, Fall 2020, Spring 2021
2015 - 2017 Admin, FooBar (Competitive Programming Club, IIIT-D)
2017 Head Teaching Assistant, Competitive Programming I (CSE200A)