Sanchit Agarwal

CONTACT INFORMATION

5819 Elwood Street Apt#6
Pittsburgh, PA 15232, USA

+1-412-499-2271
sagarwa1@cs.cmu.edu

EDUCATION

Carnegie Mellon University, School of Computer Science, Pittsburgh December 2016
Master of Science in Intelligent Information Systems **4.00/4.00**

Indian School of Mines (ISM), Dhanbad May 2013
Integrated Master of Technology in Mathematics and Computing **9.37/10.00**

SKILLS

Programming: C,C++,Java,Python,MATLAB,MySQL,Android development

Technologies: Eclipse, Microsoft Visual Studio, SVN, Perforce, Apache UIMA, Lucene

PROJECTS & INTERNSHIPS

Text Based Information Retrieval Sept 2015 – Dec 2015
Carnegie Mellon University, Pittsburgh, USA
Built Ranked Boolean, Okapi BM25 and Indri retrieval models in Java using Lucene search engine library. Implemented several query operators and added query expansion and learning-to-rank capabilities. Conducted experiments to analyze the performance of the search engine on ClueWeb09 dataset using metrics like Mean Average Precision (MAP) and Precision@N (P@N).

Analysis of fMRI data using Machine Learning Nov 2015 – Dec 2015
Carnegie Mellon University, Pittsburgh, USA
Performed multi-class kernelized Support Vector Machine (SVM) classsification to categorize the activity being performed by test subjects using their fMRI scans. In another task, used Principal Component Analysis (PCA) with ridge regression to predict missing voxels in fMRI scans of test subjects. Used holdout testing and cross-validation to test the model accuracy and prevent any over-fitting. Implementation was done in python using scikit-learn library.

Predicting star ratings from customer reviews Nov 2015
Carnegie Mellon University, Pittsburgh, USA
Built using python, a machine learning system from scratch to preprocess data, extract features (bag of words), and learn a multi-class classifier via logistic regression. Implemented stochastic gradient descent (SGD) and Batched SGD to train the logistic regression model. The work was done on a dataset from Yelp containing approximately 1 million reviews.

Movie Recommender System Oct 2015
Carnegie Mellon University, Pittsburgh, USA
Developed, in MATLAB, a collaborative filtering based recommender system for movies using a subset of Netflix Prize dataset. Implemented both memory-based and model-based approaches for rating prediction. Tried clustering techniques to address cold-start problem.

Clustering on News Articles Sept 2015
Carnegie Mellon University, Pittsburgh, USA
Build a Bipartite Clustering system, a type of reinforcement clustering that produces both document clusters and word clusters simultaneously, in MATLAB. Performed experiments on a subset of TDT4 dataset containing around 2000 documents. Evaluated the system's performance using Sum Of Cosine Similarity metric.

Fuzzy Transform and its Applications May 2012 – May 2013
Indian School of Mines, Dhanbad, India
Studied and worked on Fuzzy Transform for discrete and continuous functions in one and two variables. Implemented in MATLAB, an approach to solve ordinary differential equations with ordinary and fuzzy initial conditions using fuzzy transform. Investigated the approach by comparing it to the exact and Euler's solutions. Also wrote MATLAB code for lossy image compression and reconstruction using fuzzy transform.

Proofs of Retrievability (POR) in Cloud Storage June 2011 – July 2011
University of Calgary, Calgary, Canada
Implemented a cryptographic scheme (proof of retrievability) that had been proposed for verifiability of data storage in cloud computing environments. The implementation is in C++ and uses several open source libraries like GNU Multiple Precision Arithmetic library, OpenSSL and Pairing-Based Cryptography library.

Work Experience

Samsung India Electronics, Noida June 2013 – July 2015
Lead Engineer
Worked on the development of new features and code maintenance of Samsung's voice assistant mobile application, *SVoice*. Managed the java framework and underlying C libraries for speech recognition (to control apps like Music and Camera through voice) and speaker verification (for unlocking a mobile device through voice) features in Samsung smartphones. Also, looked into android code, was involved in commercialization of *SVoice* and OS upgrade projects for wide range of Samsung smartphones and tablets.

Selected CourseWork

Graduate: Machine Learning, Search Engines, Machine Learning for Text Mining

Undergraduate: Computer Organization, Data Structures, Algorithms Design & Analysis, Operating Systems, Software Engineering, Theory of Computation, DBMS, Probability and Statistics, Statistical Inference, Optimization Techniques, Mathematical Modeling & Calculus of Variations, Stochastic Process, Abstract Algebra, Linear Algebra, Numerical Methods, Fuzzy Set Theory and Applications

Honours and Awards

Awarded the **Best Student Shield** by **The President of India** at the convocation in May 2014

Selected twice for the University of Auckland International Summer Scholarship in 2011 and 2012

Selected for the University of Queensland Summer Research Scholarship in 2010-11

Topper of the college for the academic year 2009-2010

Recipient of the INSPIRE scholarship from the Department of Science and Technology, Govt. of India from 2008-2012

Placed in the National top 1.5% in the IIT-JEE 2008

Extra Curricular Activities

Won *Best Stock Price Data Hack* (Sponsored by FINRA) at PennApps XII, Sept. 2015

Placement Representative, Integrated M.Tech Mathematics & Computing, Batch 2k13

Hostel Prefect, *Amber Hostel* (3rd yr. hostel) for the academic session 2010-2011

Selected among top 150 Indian students to attend Level-0 session of Mathematical Training and Talent Search (MTTS), at Sambalpur University, Orissa from June 2009 - July 2009
(MTTS is a summer school organized by the National Board of Higher Mathematics (NBHM))

Hobbies: Reading