Jasneet Sabharwal

4007 Hamilton Hall, 8888 University Drive
Burnaby, BC, Canada - V5A 1S6
✆ *+1 778 886 4264*
✉ *jasneet.sabharwal@sfu.ca*
🖥 *www.sfu.ca/~jsabharw/*

Work Interest

Machine Learning, Natural Language Processing

Education

2013 – Present (Expected 2015) — **Master of Science, Computing Science (Thesis Track)**, *Simon Fraser University*, British Columbia, Canada.
- Affiliated to SFU Natural Language Laboratory
- Supervisors: Dr. Anoop Sarkar, Dr. Fred Popowich

2006 – 2010 — **Bachelor of Technology, Computer Science and Engineering**, *Guru Gobind Singh Indraprastha University*, New Delhi, India.

Publications

2012 — Apoorv Agarwal and Jasneet Singh Sabharwal, *"End-to-End Sentiment Analysis of Twitter Data"*, In Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data, COLING 2012, the 24th International Conference on Computational Linguistics.

Work Experience

November, 2014 – Present — **WorkSafeBC**, *Mitacs-Accelerate Graduate Research Intern*.
- Detecting trends in conversations using natural language processing and visualization.

September, 2013 – Present — **Natural Language Lab, Simon Fraser University**, *Research Assistant*.
- Visualization of historical human events using non-linear dimensionality reduction techniques.

May, 2014 – August, 2014 — **Maluuba Inc.**, *Natural Language Processing Development Intern*, English NLP & NLP Engine Team.
- *Feature Engineering*: New features for performing text classification tasks were developed to improve accuracies for various domains that were supported by the product. Two main types of features that were developed were Gazetteers (big lists) and Regular Expressions.
- *Feature Reduction*: Feature reduction was performed to identify features that would degrade the performance of the classifiers.
- Building and *fine tuning models for classification tasks*.
- Post-Processing: New feature in *time normalization* module was added to handle normalizing time to past.
- *Performance Improvements*: Modules in Python which were a bottleneck were converted into Java for improving performance of the engine.

May, 2012 – July, 2013 — **Germinait Solutions Pvt. Ltd.**, *Software Engineer*, Analysis Engine (Research Team).
The Analysis Engine's Research Team was responsible for developing textual analysis algorithms like named entity recognition, sentiment analysis, intention analysis, etc.
- *Designed a framework for realtime distributed analysis of big data*. The framework incorporated NoSQL databases (MongoDB, Redis & Cassandra), Apache Zookeeper (for monitoring and inter-algorithm communication). The framework supported inter-algorithm dependencies and new analysis algorithms could be plugged into the framework with minimal efforts.
- Led the efforts to *streamline the research process* of the newly formed Research Team.
- Developed NLP algorithm for performing *Sentiment Analysis of English text*. The new algorithm utilized machine learning techniques to boost the accuracy of the system and also required development of prior-polarity sentiment dictionary which contained polarity scores of English words present in Wordnet.
- Mentored the development of algorithms for identifying *Author Intentions from English texts*.

March, 2011 – May, 2012	**NextGen Invent Corporation**, *Software Engineer*, PureAnalyzer Team.

PureAnalyzer was a social media market research tool which analysed the public sentiments of a company, the company's products, their competitors, and their competitor products. The data was represented using various visualizations and reports.

- Led a team of 4 software engineers for development of Sentiment Analysis Engine, Web Crawler and other backend modules for the product.
- Designed backend framework and modules for the product.
- Research & Development of a *Document Level Sentiment Analysis Engine* for calculating the polarity of the attitude of a writer. The engine utilized *lexical level* and *shallow syntactic level* features. *Machine learning* algorithms were used to train the models and also *domain adaptation* was performed for better performance and generalization of the models to support multiple domains.
- Implemented Twitter Search and Streaming API.
- Terabyte sized database indexing and searching.
- Handled deployment and monitoring of product on 30+ cloud and dedicated servers.

Technical Skills

languages	Java, Python, R, Linux Shell Scripting	database	MySQL, MongoDB, Couchbase
natural language processing suites	Open NLP, GATE, Stanford NLP Tools, Wordnet, Natural Language Toolkit (NLTK)	machine learning suites	Support Vector Machines (LibSVM), Weka, Scikit Learn
data representation	JSON, XML	misc	Apache Hadoop, Apache Zookeeper, Apache Solr

Projects

2014	**Visualizing Wikipedia using t-SNE**.

It utilized t-Distributed Stochastic Neighbor Embedding using Barnes Hut Approximation for non-linear dimensionality reduction of high dimensional wikipedia text to 2 dimensions for visualization. It showed that using semantic role labels as features instead of bag of words features, the visualizations had well formed clusters.

2013	**Multi-Label Text Classification**.

Multi-label classification problem is that in which each example can be classified into multiple classes. It utilized Support Vector Machine algorithm to handle the multi-lable classification problem and showed promising results.

2010	**Face Recognition System**.

It was based on Eigenface generation by performing Principal Component Analysis on a large set of images of human faces and was implemented in MATLAB.

Volunteering

September, 2014 - Present	**Computing Science Graduate Students Association, Simon Fraser University**, *President*.
June, 2014 - Present	**Senate Graduate Studies Committee, Simon Fraser University**, *Graduate Student Member*.

Positions of Responsibility & Awards

October, 2012	Participated in International Coastal Clean-Up, 2012 held by Indian Maritime Foundation in Mumbai.
October, 2011	Received "Challenge Award" for my work in PureAnalyzer, a product of NextGen Invent Corporation.
August, 2011	Received "Achiever of the Month" Award for July 2011 for my work in NextGen Invent Corporation.
2009	Technical Team Head of Gates'09, GTBIT's Annual Techno-Cultural Fest.
2008	Organized Mopix Event, a Web Designing Event during Gates'08, GTBIT's Annual Techno-Cultural Fest.