

## Education

### Master of Science in Computer Science

Expected Dec 2014

- *The University of Texas at Dallas, Richardson, Texas*

GPA: 3.959/4

Thesis: Statistical Classification and String Manipulation Algorithms using MapReduce

### Bachelor of Technology in Electrical Engineering

August 2010

- *Indian Institute of Technology (IIT) Bombay, Mumbai, India*

GPA: 7.46/10

## Relevant Coursework & Programming Skills

- Design and Analysis of Computer Algorithms
- Artificial Intelligence
- Advanced Database: Big Data Analytics
- Database Design
- Machine Learning
- Natural Language Processing
- Discrete Mathematics
- Operating Systems Concepts

**Languages:** Python, Java, C++

**Operating Systems:** Linux, Windows

**Miscellaneous:** Hadoop, MapReduce, Pig Latin, Hive, Mahout, Cassandra, MySQL, Git, MATLAB

## Research Projects

- **Brown University** Providence, Rhode Island  
*Distributed memory abstraction* May 2013 - June 2013
  - Constructed a new system for **Big Data** Analytics using LLVM and Julia.
  - Performed in-memory computations on large clusters in a fault-tolerant manner.
- **Institut National De Recherche En Informatique Et En Automatique** Paris, France  
*Hybridization Methods for the analysis of Biomolecular Networks* May 2009 - July 2009
  - Proposed a new mathematical model to calculate *Violation Degree*
  - Composed a **MATLAB** library to implement this model
  - Increased the calculation speed of *Violation Degree* by about 30 times
  - Fashioned the library to guide the search for models satisfying a given specification.
- **The University of Texas at Dallas** Richardson, Texas  
*Calculating Edit Distance using MapReduce* Aug 2013 - Dec 2013
  - Investigated the parallelization of calculating edit distance for a large set of strings.
  - Devised and implemented algorithms tailored to MapReduce.
  - Analytically and experimentally confirmed their superiority over dynamic programming approaches.

## Academic Projects

- **Computational Framework for characterizing politeness** Aug 2013 - Dec 2013  
*Guide: Prof. Yang Liu & Prof. Latifur Khan*
  - Built politeness classifiers using different **machine learning** algorithms, filters and features.
  - Used politeness strategies to demonstrate the performance improvement when using **linguistically informed** classifiers.
  - Employed these classifiers to analyze the politeness of popular blogs.
- **Simulating Nine Men's Morris board game** Jan 2013 - May 2013  
*Guide: Prof. Haim Schweitzer*
  - Designed and developed this AI project in **Python** employing alpha-beta pruning algorithm.
  - **Winner** in the tournament for the game in the class (76 students).

- Wrote reliable and fast code and devised innovative static estimation algorithms.

- **Comparative analysis of approaches to sentiment analysis of tweets** Jan 2013 - May 2013  
*Guide: Prof. Latifur Khan*
  - Implemented twitter data processing in **Java** to predict the ratings of movies.
  - Employed standard lists of positive and negative words to classify each tweet.
  - Used the sentiment140 api for the classification using **MapReduce**.
  - Created a classification model with naive Bayes method in **Mahout** to classify tweets.
- **Database Design and Implementation Project** Aug 2012 - Dec 2012
  - Compiled a conceptual design (EER model) for a large custom City library database project.
  - Developed a relational schema in third normal form for this design.
  - Wrote **SQL** statements to create database, populate tables and solve challenging queries.
  - Assembled the database in Oracle and used a database state to verify the correctness of queries.
- **Date-Intensive Text Processing of White House Visitor Log** Jan 2013 - Feb 2013
  - Wrote programs in **Java** and applied Hadoop **MapReduce** chaining to derive statistics.

## Professional Experience

- **Research Assistant** Nov 2013 - Present  
*Data Management & Data Mining Lab, The University of Texas at Dallas* Richardson, Texas
- **Software Developer** Jan 2011 - June 2012  
*Future Bazaar* Mumbai, India
  - Head of the *Analytics and Business Intelligence* Team.
  - Designed an architecture to carry out ETL processes.
  - Built the E-commerce platforms for *futurebazaar.com* and *chaupaati.in* using **Django**.
  - Constructed modules for using api's of BI tools.
  - Conceptualized and implemented Order Life-cycle Management, a system to assign and regulate order-states to the order-items using **Python**.
  - Developed an interface for the sellers to track orders, sales summary and product reviews, upload product inventory and edit profile settings.
- **Math Lab Tutor** Jan 2013 - Present  
*The Student Success Center, The University of Texas at Dallas* Richardson, Texas
  - Tutor students at UT Dallas on an individual or group basis.
  - Assist in a variety of subject areas including math, physics and statistics.

## Honours

- All India Rank 94 in Indian Institute of Technology Joint Entrance Exam 2006 among 300,000 students.
- Selected from the state for the 4th Invitational World Youth Mathematics Intercity Competition.