Souradeep Sinha

▬ Education

2014–2016	**MS in Computer Science**	*Syracuse University*

GPA (out of 4) - 3.52

Relevant Coursework: Analytical Data Mining, Fundamentals of Data Mining, Design and Analysis of Algorithms, Cryptography

2010–2014	**BTech in Information Technology**	*West Bengal University of Technology*

GPA (out of 10) - 8.44

Relevant Coursework: Database Management Systems, Computer Networks, E-Commerce, Information and Coding Theory, Operations Research, Engineering Mathematics

▬ Professional Experience

Related Experience

2015–Present Graduate Researcher - Social Media Mining   *Data Lab, Syracuse University*

Advisor: Dr. Reza Zafarani, Area of Research: Data Science
- Implemented Bloom filter in Python to block spam accounts with 0.12% false positives
- Tested intrusion detection techniques using Weka's Naive Bayes, J48 Decision Tree and kNN classifiers as well as Liblinear's SVM classifier
- Investigated psychology and computer science research on deception and provided feature recommendation for analysing deception in social data
- Collected, cleaned, preprocessed data using Python, Selenium and Beautiful Soup

2013–2013 Data Analysis Intern   *Quikr,* India
- Investigated online advertisement data to identify consumer trends and buying patterns
- Developed statistical and predictive models that helped increase customer engagement time on premium ads by 14%
- Provided graphical recommendations to increase user bounce times

2013–2014 Database Lab Administrator   *Institute of Engineering and Management,* India
- Developed MySQL queries to effectively test core concepts of relational databases
- Designed course project with appropriate Visio schema diagrams and sample data to facilitate stored procedures, functions, triggers and error handling
- Collaborated and formulated a lesson plan aimed at increased student participation in labs

Additional Experience

2014–Present	**Late Night Manager**	*Student Centers and Programming Services,* Syracuse
2015–2015	**Summer Staff**	*Reservations and Building Operations,* Syracuse
2015–Present	**Event Staff**	*Events and Technical Services,* Syracuse
2010–2014	**Business Development Associate**	*In Good Taste,* India

Projects

α-Tweet: Retweet prediction with Naive Bayes   *Syracuse University*
- Designed multithreaded scraper bots to collect tweet history for users and their followers
- Used NLTK to perform word tokenization and TF-IDF filtering to clean data
- Modified Naive Bayes algorithm and Pearson's coefficient and identified negative correlation between tweet words and retweet possibility

Analysis of elasticity on AWS Cloud   *Syracuse University*
- Used Botocore to combine EC2, S3 and DynamoDB resources to store vectorized clusters
- Designed shell scripts to enable command line autologins to multiple EC2 instances
- Automated the procedure of data upload, download and performance monitoring with simple Python Boto3 API calls

Comparative study on dimensionality reduction techniques *Syracuse University*
- Compared Random Projection and PCA to reduce a synthetic dataset of 60 features
- Employed Agglomerative Hierarchical Analysis with Conditional Entropy and Normalized Mutual Information for cluster validation

ID3 Decision Tree with Python   *Syracuse University*
- Implemented a model capable of training and classifying nodes from any input dataset with varying parameters and splitting threshold values

Skills

{Languages}	Python, Java, R	{DB/rDBMS}	MySQL, MongoDB
{Big Data}	Hadoop, MapReduce, HiveQL	{AWS Stack}	EC2, S3, DynamoDB
{OS}	Linux, Unix	{Packages}	Pandas, Scipy, Urllib
{APIs}	NLTK, Tweepy, Botocore	{Other}	Shell Scripting, Multithreading

Leadership Roles

2015–Present   **Senator** *Graduate Student Organization* Syracuse University

2015–Present   **Executive Board Member** *Engineering and Comp Science GSO* Syracuse University

2014–Present   **Fellow** *Leadership and Cultural Program* Syracuse University

2014–Present   **Graduate Ambassador** *College of Enginering and Comp Science* Syracuse University

Certifications

- The Data Scientist's Toolbox - Johns Hopkins University
- Introduction to Hadoop and Mapreduce - Cloudera
- Data Wrangling with MongoDB - MongoDB

Interests

- Running(10K, 5K), Hiking, Soccer, Tennis
- Geometric Drawing - Fractals, Spirals, Pentagrams, Zentangles, 1-5 Point Perspectives
- Community Service - Child Education in India, Women Empowerment through Craft, Animal Welfare

<Insert Company Name>   January 21, 2016
<Insert Informal Address>

Dear Recruiter,

I would like to express my interest for the position of Data Scientist at <Insert Company Name>. I would also like to take this opportunity to introduce myself.

I am a data science researcher and computer science student, graduating in May 2016, and am currently working at Dr. Reza Zafarani's research laboratory at Syracuse University. My area of research is Social Media Mining and I have over two years of experience in handling data of more than a million records and sometimes terabytes in size. Apart from this, I have also worked as a Data Analysis intern at Quikr India as an undergraduate student, where I had the opportunity of diving deep into India's demographically-rich advertisement dataset, and came up with statistical models which helped the firm increase customer engagement times by 12

I also have experience in the field of technology as a Database Lab Administrator and also in the form of several academic projects for the last 6 years, where I worked on software development, computer networks, databases and ecommerce. To review my full professional experience, please visit https://www.linkedin.com/in/souradeepsinha

Given the skills required for this position, I believe that I will be able to fit in perfectly. I would look forward to your communication. Please feel free to reach me @ (315) 412 - 8844 or at ssinha04@syr.edu

Thanks,

Souradeep

Yours faithfully,

Souradeep Sinha