Mounica Maddela

Contact Information	1823 Terry Avenue Seattle, WA 98101 USA	*Mobile:* (215) 290-6918 *E-mail:* mmaddela1005@gmail.com *Website:* https://mounicam.github.io/

Research Interests Natural Language Processing, Computational Linguistics, Machine Learning

Education

University of Pennsylvania, Philadelphia, Pennsylvania USA (2013- 2015)

Master of Science in Engineering in Computer and Information Science (CGPA - 3.64 / 4.00)

International Institute of Information Technology - Hyderabad, India (2009- 2013)

Bachelor of Technology(Honors) in Computer Science and Engineering (CGPA - 9.07 / 10.00)

Relevant Coursework Masters : Machine Translation(A+), Computational Linguistics(A), Machine Learning(A), Internet and Web Systems(A-)

Bachelors : Artificial Intelligence (A), Data Warehouse and Data Mining (A), Human Computer Interaction (A), Artificial Neural Networks(A), Statistical Methods in AI(A-), Optimization methods(B), Linear Algebra(B), Information Retrieval and Extraction (B-)

Technical Skills Programming Languages: Java, C++, C, MATLAB 2013a

Scripting and Other Languages: Python, JSP, PHP, Javascript, Bash/Shell scripting, HTML

Databases/Storage: MySQL, MongoDB, Berkeley DB, Amazon DynamoDB, Amazon S3

NLP and Data Mining Tools : Stanford CoreNLP, Stanford Topic Modelling Toolbox, MALLET, Scikit, NLTK, WEKA Version 3.7, MATLAB Statistics and Machine Learning Toolbox

Graphics and Visualization: OpenGL, QT4.7, MATLAB: Visualization and Image Processing

Cloud Computing Services: Amazon EC2, Amazon CloudSearch, Amazon Simple Queue Service

Development platforms: Eclipse, NetBeans

Operating Systems: Unix/Linux, Windows, Mac OS X.

Research Projects

Cross-Cultural Analysis using Twitter (Independent Study Project) **Spring 2014 - Spring 2015**
Advisor: Dr. Lyle Ungar
Goal of the project was to capture the different sources and interpretations of well-being across various cultures or countries. The idea was to capture the context of sentiment words and analyze their distribution across countries. The project was a part of World Well Being Project (WWBP).

For this project, implemented country filter to classify tweets, extracted n-grams highly correlating with positive/negative sentiment words and analyzed the results with word clouds. Also experimented with topics instead of n-grams and positive/negative categories of sentiment lexica rather than sentiment words. Assisted with similar experiments on Spanish tweets and Chinese microblogs form Weibo.

CROVHD (Honors Project and BTP - B.Tech Project) **Monsoon/Fall 2011 - Spring 2013**
Advisor: Dr. Kamalakar Karlapalem
Developed new visualization system called CROVHD (Concentric Rings of Visualization of High Dimensional Data) to visualize the high dimensional data as a 2-D representation. Extended this system to 3-D Cone visualization to visualize k-nearest neighbours.

Publications and Presentations

Mounica Maddela and Kamalakar Karlapalem, Visualizing Nearest Neighbours for Large High Dimensional Datasets, IEEE Symposium on Large Scale Data Analytics and Visualization, Seattle, October 2012. (Poster)

Poster on "CROVHD and CONE Visualization System", RnD showcase 2013, IIIT, Hyderabad

Poster on "Grasshopper query builder using natural language", Amazon eCFT Intern Science Fair 2014, Seattle

Professional Experience

Software Development Engineer I at Big Data Technologies, Amazon **June 2015 - present**

- Worked in a team of 2 to develop new search service for DataNet, our data management system.
- Mentored an intern to develop faceted search for DataNet.
- Refactored authorization process for DataNet which reduced the latency by 40%.
- Currently designing system to improve bulk SQL job monitoring in DataNet. Goal of the project is to track and visualize the state of the SQL dependencies graph at any point of time.

Software Development Intern at Big Data Technologies, Amazon **June 2014 - August 2014**
Developed natural language interface to help customers communicate with Grasshopper, our SQL query builder system. In Grasshopper user interface, the users can drag and drop query attributes and conditions, which are ultimately converted into SQL query. The goal of the project was to select these attributes or conditions using natural language rather than users drag-and-drop actions. For this project, extracted table/schema names and query conditions from the input natural language sentence and populated these entities in a pre-defined Grasshopper query template.

Application developer at Cancer Research Laboratory, UPENN **September 2013 - May 2014**
Developed web-based application using mySQL to monitor the orders and inventory of the laboratory.

Text Mining Intern at SetuServ **June 2013 - July 2013**

- Captured insights from tweets posted during American Society of Clinical Oncology (ASCO) conference. Extracted the tweets discussing about cancer/drug/therapy and derived the topics discussed by clustering these tweets using k-means.
- Built a classifier using credit card transactions data to recognize the type of product based on product name, product description and web definition of the product.
- Validated the business type information given by the customers using Cross-Validation approach.

Teaching

Teaching Assistant for Internet and Web Systems (Spring 2015)

Teaching Assistant for Computational Linguistics (Fall 2014)

Teaching Assistant for Introduction to Databases (Monsoon 2012)

Honors and Awards

Research Award for undergraduate students at IIIT-H for 2011-2012

Deans Academic Award List-I for Semester V (Monsoon/Fall 2011)

Deans Academic Award List-II for Semester II (Spring 2010)

Deans Academic Award List-III for Semester VI (Spring 2012)

Deans Merit List for Semester I (Monsoon/Fall 2009)

Deans Merit List for Semester III (Monsoon/Fall 2010)

Deans Merit List for Semester IV (Spring 2011)

Selected Course Projects

Sentiment Analysis of restaurant reviews (Machine Learning)
Developed a system to predict the star rating of restaurant reviews given the review text and other review properties.The rating system consisted of three stages: data preprocessing to reduce the

dimensionality of text data from words to informative word clusters, individual labels prediction with an ensemble of classifiers and weighted voting of these labels for the final rating.

Automatic Multi-document Summarization (Computational Linguistics)
Developed multi-document extractive summarizer.

Crowdsourcing Translation (Machine Translation)
Developed a competitive task where the goal was to improve the accuracy of crowdsourced translations using translation and other metadata properties. Gathered data for the task, chose an evaluation metric to rank the solution for leaderboard, came up with the baseline for the task and experimented with different approaches to rank the quality of crowdsourced translations.

Language Research on Bhojpuri (Machine Translation)
Researched on the syntax, morphology, writing system and distribution of speakers for a low resource Indian language called Bhojpuri. Also gathered monolingual data for this language by crawling the Bhojpuri news sites and built a language identification system using this data.

Topic Modelling in Social Media (Information Retrieval and Extraction)
Implemented and experimented with Social Latent Dirichlet Allocation (Social LDA), an improvement over LDA to deal with social media data.

Blitz Search (Internet and Web Systems)
Developed a distributed search engine from scratch consisting of distributed crawler to crawl web, indexer to index the crawled pages, PageRanker component to calculate page-ranks of the URLs and search engine to extract matching documents along with ranking them.

Autism Prediction for High Risk Siblings (Biomedical Image Analysis)
Developed a classifier using brain-related biomarkers to predict if a high risk infant can develop Autism Spectrum Disorder (ASD). Extracted features from Diffusion Tensor image of brain, filtered the features using signal to noise ratio and used SVM for classification.

Term paper on Interior Point Methods (Optimization Methods)
Prepared a tutorial covering the Interior Point methods for Linear Programming. It also included some applications of Interior Point methods to Convex Programming problems.

Term paper on Comparison of Subspace Clustering Algorithms (Data Mining)
Compared two subspace clustering algorithms MAFIA (Merging of Adaptive Finite Intervals) and DOC (Density Based Optimal Projective Clustering) which involve adaptive grids.

Information Visualization and Data Analytics(Independent Study Project)
Studied the resources of a course called Information Visualization and its applications by Professor John Stasko of Georgia Institute of Technology and experimented with the visualization methods in the course.

Face Replacement (Computer Vision)
Developed a system to automatically recognize and replace faces in an image or video.

Workshops Seminars Hackathons Attended

Amazon Machine Learning Conference (AMLC-2015) held in Seattle from April 29th to May 1st 2016

Grad Cohort Workshop 2015 held in San Francisco, California, organized by Computing Research Association-Women from April 10th to April 11th 2015

Computer and Information Science(CIS) Colloquium Series held at University of Pennsylvania,

Philadelphia for Fall 2014 and Spring 2015

Amazon ML Hackend held in Seattle from July 18th to July 20th 2014

Workshop on Image and Speech Processing (WISP-2012) held at IIIT Hyderabad organized by C.R. Rao Advanced Institute of Mathematics, Hyderabad, 17th December 2012

International Conference on Management of Data (COMAD-2011) held at IIIT - Bangalore from 19th to 21st December 2011.

EVENTS ORGANIZED

Co-organizer for Amazon eCFT Intern Science Fair held in Seattle on July 3rd 2016

Volunteer for Girls Who Code workshop, held in Seattle on December 11th 2015, to educate middle school and high school girls about different types of STEM careers.