

SALONI POTDAR

WORK EXPERIENCE

MAY 2014 - PRESENT

IBM Watson, Littleton

Watson Core Algorithms Summer Intern

Worked on the question answering system as a part of the IBM Watson Team. Used distributional semantics to improve question classification, synonym generation and query expansion.

SEP 2013 - PRESENT

Carnegie Mellon University, Pittsburgh

Graduate Research Assistant

Advisor: Jamie Callan

Worked on text mining of social media sites that discuss chronic kidney diseases. Aim was to build a text mining solution using a search engine (Indri and Lucene API). Sentiment analysis, expert identification, topic modelling and machine learning techniques were used to develop tools which helped to explore the data. Also created a website-independent solution to remove redundant information from web pages.

DEC 2012 - JUL 2013

Tata Consultancy Services (TCS), Pune, India

Assistant System Engineer Trainee

Underwent training in C++ and Unix for the British Telecom project. Lead a team of 7 people in the projects at the training center which was highly appreciated by the training heads.

JUN 2011 - MAR 2012

PuneTech, Pune, India

Intern

Advisor: Dr. Navin Kabra & Dr. VV Phansalkar

Created a recommender system for the Indian stock market shares by mining financial websites and Twitter. This expert identification system ranks markets analysts and generates recommendations for trading shares using structured and unstructured textual data. The system theoretically gave higher returns than the SENSEX average over 3 months.

SOFTWARE SKILLS

COMPUTER LANGUAGES	Java, Python, C, C++, MATLAB, PHP
PLATFORMS & APIS	Hadoop, Amazon Web Services (AWS), Weka, Lucene, Indri, Scrapy Framework.
DATABASES	MySQL, Oracle (10g/11g).

PUBLICATIONS

- Identifying Student Leaders from MOOC Discussion Forums through Language Influence, Seungwhan Moon, Saloni Potdar and Lara Martin, accepted at EMNLP MOOC Workshop, 2014.

CERTIFICATIONS & HONORS

- Certificate in Cloud Computing, from the TCS Business Domain Academy.
- ILP Kudos Award, the highest honor during training at TCS.
- Best Outgoing Student 2012, Computer Engineering Department, RSCOE, University of Pune.
- Certificate of Leadership, National Entrepreneurship Network, India.

📍	5411, GHC, Carnegie Mellon University, Pittsburgh
☎	+1 412-245-6860
✉	salonipotdar@cmu.edu
🔗	cs.cmu.edu/~spotdar

EDUCATION

School of Computer Science, Carnegie Mellon University, Pittsburgh

MS in Intelligent Information Systems, Language Technologies Institute.

Aug 2013 - Dec 2014 (expected)

Cumulative QPA - 3.5

Courses: Machine Learning for Large Datasets, Machine Learning, Cloud Computing, Discourse Analysis, Search Engines and Text Mining, Algorithms for NLP.

University of Pune, India.

Bachelor of Engineering in Computers

Aug 2008 - Aug 2012

Grade: First Class with Distinction

GRADUATE PROJECTS

AUG 2014 - PRESENT

Java

Personalization in search & news recommendation

Modelling of user behavior to improve relevance of search results. News recommendation using contextual bandits.

MAR 2014 - MAY 2014

Java, AWS, Hadoop, MySQL

Twitter Analytics

Built a robust REST web service on the AWS cloud to analyze 250GB of twitter data.

FEB 2014 - APR 2014

Java, Hadoop, AWS, Pig

Machine Learning Algorithms for Large Datasets

Implemented Naive Bayes, Personalized Page Rank, Stochastic Gradient Descent in a memory constrained setting.

SEP 2013 - DEC 2013

Java and Weka API

Improving robustness of classifiers using ensemble of ensembles

Built and evaluated a robust classifier learner that performed well with minimal tuning. We used 24 development datasets and 20 evaluation datasets for this project.

Nov 2013

Java

Netflix Movie Recommender System

Used collaborative filtering to predict the ratings of movies, and evaluated them on a subset of the Netflix Prize dataset.

Nov 2013

MATLAB, SVM-Light

Learning to rank (LETOR)

Implemented regularized logistic regression (RLR) with gradient descent, and ran experiments on adapted RLR and Support Vector Machine for LETOR.

SEP 2013 - OCT 2013

Java, Lucene API

Implementation of Information Retrieval Models using Lucene API

Implemented IR models like the ranked and unranked boolean, BM25 and Indri. Also performed query expansion with pseudo relevance feedback.