NIKHIL KHANDELWAL

40 Newport Pkwy, Jersey City NJ (716) 479-4072
nikhandelwal.github.io nikhilkh@buffalo.edu

EDUCATION

University at Buffalo, The State University of New York (SUNY) **GPA: 3.80**
Masters of Computer Science, May 2015

Sinhgad College of Engineering, University of Pune, India
Bachelors of Technology, Information Technology. May 2011

COURSEWORK

Analysis of Algorithms, Information Retrieval, Distributed Systems, Machine Learning, Operating System, Data Intensive Computing

SKILLS

Keywords: Object-Oriented Programming, Java, Android, Apache Lucene, Apache Solr, Apache Hadoop
Languages: Java, JavaScript, Python , Perl
Framework and Tools: NLTK, Moses, Spring MVC, Android, MapReduce.

EXPERIENCE

FactSet Research Systems Inc., NY – NLP Engineer ***since* February 2015**

PROJECTS

Primary Company Tagging for Financial Article **(Perl, Python)**

- Developing a system that tags the primary companies in financial news feed.
- Using a hybrid approach of using the news vendor provided meta data and tagging the primary entities using a CRF based named entity recognition (NER) model.

Bidder/ Target identification in Merge& Acquisition Article **(Java, Python)**

- Identifying bidder and target in a merger and acquisition news article. An extended goal is to extract the amount of money.
- Trying two parallel approaches, using the Stanford NER library to train our own model and also using a python CRF library (crfsuite) to try and choose the approach with better results.

Language Translation using Statistical Machine Translation **(Python, Moses)**

- Built a language translation engine similar to google translate, for Greek, based on principles of statistical machine translation.
- The cleaning, tokenization, stemming and other language processing is done by using nltk and other NLP libraries in python.
- Using a Perl tool, Moses to train a model to the language translation. Achieved a BLEU scores of 39, for financial domain documents.

Accenture Pvt. Ltd, Bangalore, India - *Software Engineer* **July 2011 – July 2013**

PROJECTS

Middleware module for Mobile Number Portability, October 2012 – July 2013 **(TIBCO BW)**

- Developed the system to synchronize a MNP request and the subsequent notifications within multiple end systems.
- Also helped design the HTTP request protocols for processing outgoing or incoming MNP requests.

ACADEMIC PROJECTS

NLP based Question Answering System, *Fall 2013* **(Java, J2EE, Apache Solr, JavaScript)**

- Generated a back end from the Wikipedia corpus using Apache Solr and implemented a front end using J2EE technologies.
- The system could handle free text queries. Query processing was done on the queries using the Stanford NLP model.
- Advanced features like spell check and auto complete were also implemented using AJAX and JavaScript.
- Added 7 systems calls to the UNIX kernel which could be used by user level programs to manipulate these locks.

Simple Amazon Dynamo - Replicated Key Value Storage *Spring 2014* **(Android, Distributed Systems, Multithreading)**

- Designed and implemented a simplified version of Amazon Dynamo.
- Implemented Quorum replication with replication degree 3 and versioned all objects to differentiate between stale copies.
- Incorporated a functionality for handling node failures and ensured that when any node recovers/rejoins it should be updated with most recent copies.

Data Analysis using MapReduce and Hadoop DFS *Spring 2014* **(Java, MapReduce, Hadoop)**

- Designed and implemented the MR work flows to extract various information from the real time data of Twitter (approx 1GB). For instance: simple word count, # tag counts, @ counts etc.
- Designed and implemented Map and Reduce functionality for parallel breadth first search algorithm.
- Implemented K-means clustering using Hadoop MapReduce.

Handwritten Digits Classification *Spring 2013* **(MATLAB)**

- Classifying handwritten digits by implementing four machine learning techniques **a) Neural Networks, b) K-Nearest Neighbors, c)Logistic Regression, d) Support Vector Machines** and compare their performance.
- Also use validation set to tune hyper-parameters for Neural Network and choose appropriate value k for k-NN.