

SUSHANT NARSALE

Apt # P33, 3501 St Paul Street, Baltimore, MD 21218. Phone: 201-310-5953

Email: sushant@jhu.edu Website: www.cs.jhu.edu/~snarsal

EDUCATION

- **Johns Hopkins University**, Master of Science in Engineering(Computer Science) May 2010
Relevant Coursework: Natural Language Processing, Information Extraction, Information Retrieval & Web Agents, Machine Learning, Parallel Programming.
- **Mumbai University**, India - Bachelor of Engineering (Computer Science) June 2006
Relevant Coursework: Algorithms, Theory of Computation, Data Structures.

RELEVANT SKILLS

- **Languages** C, Python, Java, Perl
- **Platforms** Linux, Windows, Hadoop
- **NLP Tools** Weka, SVM-Light, GIZA++, Moses, mallet-0.4, AT&T FSM tools

RESEARCH PROJECTS (More Details at: www.cs.jhu.edu/~snarsal/research)

Self Adjustable Boot Strapping for Expansion of Entity Sets for 150 NE categories

- Features like neighboring context words, dominant suffixes and prefixes, POS tags, capitalization, etc were used to identify new Named Entities (e.g. Names, Locations ...) from Wikipedia n-gram data.
- Developed a Self-Adjustable model that adjusts these parameters and features across 150 categories to obtain best possible entity sets for each individual category.

Research Assistant – DARPA GALE Project

- Member of JHU team working on System Combination for Machine Translation (MT).
- Built an SVM-rank based model for ranking the output of multiple MT systems for system combination.
- Implemented a log-linear model based on MERT for ranking of MT systems.
- Working on a perceptron based approximation algorithm for ranking of MT systems for combination.

Publication

- Dekang Lin, Kenneth Church, Heng Ji, Satoshi Sekine, David Yarowsky, Shane Bergsma, Kailash Patil, Emily Pitler, Rachel Lathbury, Vikram Rao, Kapil Dalwani and **Sushant Narsale**, “New Tools for Web-Scale N-grams.” In Submission to The Seventh International Conference on Language Resources and Evaluation (LREC 2010), Valleta, Malta.

NLP, IE, IR ACADEMIC PROJECTS (Details & Reports: www.cs.jhu.edu/~snarsal/projects)

Cross Lingual Information Retrieval

- Developed a Cross Lingual Information Retrieval System for English and Hindi Languages.
- Implemented a statistical generative bi-gram model for transliteration (using character alignments).
- Implemented a Hindi language Stemmer which removed 51 inflectional suffixes.
- Implemented a Dictionary based translation model and a TF-IDF based retrieval scheme.

Part Of Speech Tagging with HMM

- Implemented a Part of Speech tagger using Hidden Markov Model in Python.
- Used Baum-Welch algorithm with partially supervised training data to estimate HMM probabilities.

Decision-tree language model for English letters

- Implemented decision-tree as a 4-gram language model with cross validation.
- Generated an encoding of English letters into bit-strings using mutual information criterion.

Language Parser

- An English language parser for probabilistic context free grammar based on Earley's algorithm.
- Pruning techniques like agenda based best first strategy, left-cornering, inverted index were used.

Isolated Word recognizer

- Developed an isolated word recognizer for a small vocabulary system, based on Hidden Markov Model.

SYSTEMS PROJECTS (Details & Reports: www.cs.jhu.edu/~snarsal/projects)

Subgraph Isomorphism in Parallel environment using Map Reduce

- Given two input graphs the decision problem of finding if one graph or its subgraph is isomorphic to another graph was implemented.
- The serial algorithm for this process was then parallelized using Hadoop Map Reduce framework.

PROFESSIONAL WORK EXPERIENCE

Hexaware Technologies, India (July 2006 - July 2008)

Designation: Software Engineer

Worked for Leading European, South American and Asian Airlines on E-ticketing Application

Redesign of Revenue Reporting process

- Proposed and implemented changes in Revenue Reporting process used by leading European, South American and Asian airlines resulting in a more efficient and robust reporting process.
- Trained a batch of 10 new hires in USAS and FORTRAN.

Redesign of Nightly File Maintenance process

- Worked in a team that redesigned the Nightly File Maintenance process used by major airlines worldwide, the emphasis was on reducing time taken by batch jobs to complete reporting updates.
- These changes removed a critical performance bottleneck and time efficiency improved by 30%.