

Harsh TRIVEDI

Email: harshjtrivedi94@gmail.com Website: <http://harshtrivedi.me> Github Handle: harshtrivedi

INTERESTS:	Natural Language Processing, Information Retrieval, Data Structures, Web development
LANGUAGES:	Ruby, Python, R, Java, C, Javascript and HTML, SQL
TOOLS & TECHNOLOGIES:	Ruby on Rails, MongoDB, Redis, PSQL, jQuery, AJAX, \LaTeX , Weka, Python NLTK and Scikit-learn, Git, AWS S3 and EC2

EDUCATION

2012-16	Dhirubhai Ambani Institute of Information & Communication Technology, Gujarat, India B. Tech in Information and Communication Technology CGPA: 8.23/10.0
2011-12	New Samarth, Gujarat, India - 12 th grade, HSc, Gujarat State Board Percentile: 93.63 GUJCET 2012: 95.32 percentile AIEEE 2012: Ranked 6186 / \approx 1.1 million students appeared
2009-10	M.K High School, Gujarat, India - 10 th grade, SSc, Gujarat State Board Percentage: 86.00

PUBLICATIONS

DEC'15	<i>"A New Approach to Syllabification of Words in Gujarati"</i> , Springer, LNAI 9468, pp. 631-639 (pdf) Harsh Trivedi , Aanal Patel, Prasenjit Majumder → Built a mathematical model for syllabification using probabilistic calculations in a systematic way. → Using CRF based approach achieved a 98.02% syllabic accuracy. → Paper accepted at "Mining Information and Knowledge Exploration (MIKE 2015)"
JUN-JUL'16	<i>"Noise Correction in Pairwise Document Preferences for Learning to Rank"</i> LNCS 9994, pp. 295-301 (pdf) Harsh Trivedi , Prasenjit Majumder → Learning to Rank is a machine learning approach to ranking. → Developed a way of correcting significant amount of errors in training data for Learning to Rank. → Up to 90% of errors could be corrected with high statistical difference. → Paper accepted in "Asia Information Retrieval Societies Conference (AIRS 2016)". → Received 439 USD ACM SIGIR Student Travel for presenting the paper.
JAN-JUN'16	<i>"Author Masking through Translation"</i> CEUR-WS Volume 1609 pp. 890-894 (pdf) Yashwant Keswani, Harsh Trivedi , Parth Mehta, Prasenjit Majumder → Author Masking is task of rewriting the document to obfuscate the stylometric identity of original author. → Built translation models with europarl parallel corpora using Moses & leveraged it for identity obfuscation. → Stood 2 nd in Author Masking Task of "Conference and Labs of the Evaluation Forum CLEF": PAN-2016.

WORK EXPERIENCE AND INTERNSHIP

DEC'13-JUL'14	Web Developer at Areysun Online Pvt. Limited (link) → Areysun is an IOS / Android app which allows people to share their expressions anonymously. → Worked on 'Ruby on Rails' on the backend with multiple data stores like Postgresql, Redis & MongoDB. → Was majorly responsible for the technical architecture and most of the server side code at Areysun.
MAY'15-JUL'15	Research assistant at IRLab - DAIICT, Gandhinagar → Developed a new statistical approach for syllabification of words in Gujarati yielding 98.02% accuracy. → Worked on government funded project Cross Lingual Information Access (CLIA). → Evaluated the performance of Cross Lingual Search Engine 'Sandhan'.

PROJECTS

CURRENT	Automatic Pseudo Relevance Label Generation for Learning to Rank → Developed a feature consensus based method for automatic training data generation for Learning to Rank. → Achieved better results with pseudo generated labels than human generated labels.
MAY-JUN'16	Syntax based Alignment for Paraphrase Generation (link) → Implemented: "Syntax-based alignment of multiple translations: Extracting paraphrases and generating new sentences". Demo available here . → It generates a word lattice (FSA) from semantically equivalent sentences that is used to extract / generate new lexical and syntactic paraphrase pairs.

- FEB-APR'16 **Diversification based Feature Selection Methods for Learning to Rank** ([link](#))
 → Implemented: "Exploiting Result Diversification Methods for Feature Selection in Learning to Rank".
 → Methods: Maximal Sum Dispersion, Modern Portfolio Theory, Maximal Marginal Relevance, Greedy Search Algorithm.
- JUL-AUG'15 **Theme based Clustering of Tweets** ([link](#))
 → Analyzed various ways of clustering tweets based on their themes.
 → Simulated the results of approaches proposed by related papers and provided comparisons on them.
 → Python NLTK and Scikit-learn, Ruby, R, Weka
- JAN-APR'15 **Class Connect** ([link](#))
 → Developed a web platform where students can share, manage and collaborate their academical resources.
 → Went through entire software development life cycle using Agile SDLC.
 → Ruby on Rails, Nginx, Remote Servers, Postgresql, Redis, HTML, Javascript, CSS
- SEP-OCT'14 **Planet Wars - Game** ([link](#))
 → Developed a real time single/multi player strategy game "Planet Wars".
 → Used client-server model and TCP/UDP Sockets for communication over LAN network.
 → Ruby, Shoes-GUI framework

EXTRA CURRICULAR ACTIVITIES

- ⇒ **Served an NGO (Tribhuvandas Foundation)** - Stayed in rural village and educated school students people about the harmful effects of tobacco consumption.
- ⇒ **Organizing team member of FIRE 2015/16 conference** - Developed and am maintaining it's new website ([link](#)) with easy content management system on admin interface.
- ⇒ **Member of GDG DAICT** - Contributed in organizing tech meetups and events as a member of Google Developer Group (GDG-DAICT) for 2 years.

PAST AWARDS AND ACTIVITIES

- ⇒ 1st prize in National as well as State competition of Mental Arithmetic (ALOHA) [2009]
- ⇒ 3rd prize in National drawing competition held by Department of Post , India [2008]
- ⇒ 1st prize in drawing competition held by Forest Department, Gujarat. [2007]
- ⇒ wrote a book of illustrated short stories for children in junior high school ([link](#)) [2007]