Shruti Rijhwani

srijhwan@andrew.cmu.edu | shrutirijhwani.com

Education

July 2015	**Birla Institute of Technology and Science, Pilani**, India *Bachelor of Engineering (Honors) Computer Science* **Senior Thesis**: Fuzzy Matching for People Search and Neighbourhood Deduplication

Research

Jul '15–Jul '16	**Research Fellow, Microsoft Research**, Bangalore, India *Advisor: Monojit Choudhury* I worked on computational models for code-switching (CS), primarily using social media text content. **CS identification**: Developed a weakly-supervised and language-independent technique for word-level language identification of text data with no prior linguistic information. The model currently labels monolingual and CS Latin script data in 35 languages, including transliterated Indian languages. **Machine translation** of code-switched data. **Language and CS Preference** of Hindi-English bilinguals for sentiment and opinion expression on Twitter.
Jan '15–May '15	**Research Intern, Microsoft Research**, Bangalore, India *Team: Advanced Development Group* **People Search**: Developed retrieval and ranking techniques based on fuzzy matching of misspelled people names in e-mail search queries. The system is awaiting release as a Microsoft Garage product. **Neighborhood Deduplication**: Worked on detecting duplicates among unstructured and transliterated neighborhood address strings using fuzzy search, synonym learning and subset grouping. Bing teams are using the technique on datasets containing millions of Indian, US and German addresses.
May '14–Jul '14	**Bing Ads Intern, Microsoft India Development Center**, Bangalore, India *Advisor: Mehul Parsana, Principal Research Manager at Microsoft* **POS-tagging search queries** using weakly-labeled search logs and Conditional Random Fields. The model had 94.9% accuracy on the MS-251 dataset (9% higher than state-of-the-art). **Location entity identification** in Bing search queries using CRFs and iterative learning.

Publications

Koustav Rudra, Shruti Rijhwani, Rafiya Begum, Kalika Bali, Monojit Choudhury and Niloy Ganguly.
"Understanding Language Preference for Expression of Opinion and Sentiment: What do Hindi-English Speakers do on Twitter?".
Conference on Empirical Methods in Natural Language Processing (EMNLP). 2016. (To Appear)

Sunayana Sitaram, Sai Krishna Rallabandi, Shruti Rijhwani and Alan W Black.
"Experiments with Cross-lingual Systems for Synthesis of Code-Mixed Text".
9th ISCA Speech Synthesis Workshop. 2016. (To Appear)

Shruti Rijhwani, Royal Sequiera, Monojit Choudhury and Kalika Bali.
"Translating Code-Mixed Tweets: A Language Detection Based System". (Demonstration Track)
Workshop on Indian Language Data: Resources and Evaluation at LREC '16. 2016.

Shruti Rijhwani, Royal Sequiera, Kalika Bali and Monojit Choudhury.
"Unrestricted Word-Level Language Detection for Code-Mixed Text".
(Manuscript under review at *Transactions of the Association for Computational Linguistics (TACL)*)

Teaching

Dec '14–May '15	**Co-founder and Lecturer, getPlaced()** Co-founded getPlaced(), a group that conducted programming tutorials for undergraduates across India.
Aug '14–Dec '14	**Computer Programming Special Interest Group**, BITS Pilani, India Conducted lectures on algorithms with an audience of over 300 undergraduates.
Aug '13–Dec '13	**Teaching Assistant, Logic in Computer Science**, BITS Pilani, India

Programming Experience

May '15–Aug '15	**Google Summer of Code, MIT Media Lab** [GitHub] *Mentors: José Dominguez, Software Engineer at Red Hat and Prof. Hal Abelson, MIT* Worked on MIT's App Inventor, a project that is a beginner's introduction to programming and app creation. Built SVG-capture and Undo features for the Blocks Editor, a visual programming editor. Development and bug fixes involved hacking large amounts of code across multiple technologies.
Nov '13–Feb '14	**Lead Developer, Lost Child Art Studios**, Pilani, India *Web design and development startup.* I built four webapps end-to-end, using various web technologies.
May '13–Jul '13	**Web Design Intern, Tune Patrol Media Labs Pvt. Ltd.**, Bangalore, India *Tune Patrol is an online marketplace for independent Indian musicians to expand their audience.* Developed the user interface for music concerts and song playlist features for the product website.
May '12–Jul '12	**Software Engineering Intern, Tata Consultancy Services**, Bangalore, India *Multinational information technology consulting and solutions firm.* I built a database management system.
May '11–Jul '11	**Web Development Intern, InfoToros**, Bangalore, India *Startup that provides complete web solutions.* I designed and developed a client's business website.

Skills: C#, C++, Python, JavaScript, R, OpenCV, OpenGL, MATLAB.

Projects

Aug '13–Dec '13	**Implementation of Graph Theoretical Methods in Image Processing** *Advisor: Mukesh Kumar Rohil, Associate Prof. at BITS Pilani* Studied and implemented graph theory based algorithms for image thresholding, resizing and segmentation. Comparatively analyzed segmentation using Prim's, Kruskal's, Watershed and Random Walks algorithms.
Sep '14–Dec '14	**Event-based News Aggregation** *Advisor: Poonam Goyal, Associate Prof. at BITS Pilani* Built a web crawler to retrieve articles from news sources and used Hierarchical Agglomerative Clustering and the silhouette score metric to find articles discussing the same event.
Jan '14–Oct '14	**Cell Phone Microscope** Built a hardware attachment for a cellphone to magnify images to microscopic extent. Developed an image classifier to identify the presence of malarial cells in the magnified images. The project won an all-India social innovation competition and was incubated by the Sristi fund [link]
Jan '14–May '14	**Object Tracking using Segmentation** *Advisor: Sanjay Singh, Scientist at CEERI* Extended the CAMShift tracker in OpenCV to allow modifiable threshold values and noise reduction. Implemented a single object tracker based on image subtraction for frame rates up to 100 frames/second.
Aug '14–Dec '14	**"Postmodern"–The State of Contemporary Knowledge** *Advisor: Harikrishnan Nair, Assistant Prof. at BITS Pilani* Studied Jean-François Lyotard's *A Postmodern Condition*, which describes the changing state of knowledge with increasing computerization. I also investigated Massive Open Online Courses (MOOCs) and other contemporary forms of knowledge transfer in the context of Lyotard's philosophy.

Other Activities

Linguistic Olympiads
Volunteer, International Linguistic Olympiad 2016
Problem Tester and Volunteer, Panini Linguistic Olympiad 2016, 2015

Contributor at Poco Mix Maadi, a blog on code-mixing [Article: Code-Mixed Language Identification]

Leadership Roles
Co-ordinator and Editor, English Press during APOGEE 2013, an all India inter-collegiate technical symposium
Captain, BITS Pilani Swimming Team, 2013-2014

Extra-Curricular: Swimming, barefoot running, SCUBA diving (PADI Rescue Diver), Piano Performance (Grade 4 Trinity)