Shruti Singh

Department of Computer Science & Engineering
Indian Institute of Technology, Gandhinagar
Gujarat, India

Email: singh_shruti@iitgn.ac.in
Links: Webpage, Github
Phone: +91 95371 60842

Education

Indian Institute of Technology, Gandhinagar 2019 - Present
Ph.D. Computer Science
Advisor: Mayank Singh

Dhirubhai Ambani Institute of Information & Communication Technology 2013 - 2017
Bachelor (Hons.) in Information and Communication Technology with Minor in Computational Science
GPA: 8.25/10.0

Central Board of Secondary Education 2013
Higher Secondary School - Kendriya Vidyalaya
Aggregate: 95.6%

Research Interests

Broad: Natural Language Processing, Information Retrieval, Artificial Intelligence
Specific: Information Extraction, Recommender Systems, Common Sense Reasoning

Research and Industry Experience

RAXTER.IO Feb, 2018 - June, 2019
Research Engineer

- Implemented a Q-Learning based literature recommendation system that adapts to the evolving research needs of the user with time.
- Implemented a system for information extraction from scientific literature, that extracts the problem statement, approach, evaluation techniques, results, and datasets for construction of scientific literature graph.

Sprinklr Inc. Jan, 2017 - Feb, 2018
Product Engineer

- Worked on near de-duplication of large-scale real-time incoming web data of the scale of 25 million documents/day, ensuring minimal latency.
- Implemented language detection module for close languages Malay and Indonesian.
- Setup the natural language pipeline for Asia Pacific languages on the Sprinklr Listening platform.

Rygbee Inc. May, 2016 - July, 2016
Research Intern

- Developed a recommendation system for scientific literature and MOOCs.
- Developed an in-house NLP pipeline for pre-processing and recommendation, used Elasticsearch for indexing papers.

Knowledge Discovery & Management Lab DA-IICT May, 2014 - Aug, 2015
Research Assistant (Advisor: Sourish Dasgupta)

- Worked on Numerical-Unit normalization to normalize numeric quantities written in natural language and occurrences of phrases representing magnitudes of physical quantities to standard SI units.
- Worked on an annotation module to compute similarity between sentences based on their SVO triples.
- Reconstructed the ancestral hierarchy of nouns from WordNet for efficient lookup, to facilitate similarity computation between nouns.

Technical Skills

Programming
Over 5000 lines: Java, Python
Over 1000 lines: C, C++

Technologies
Elasticsearch, MongoDB, MySQL, Apache Kafka, Apache Lucene, Flask, Docker, Git, CUDA, OpenMP, StanfordCoreNLP, Spacy, Wordnet, NLTK, OpenCV, Gensim, Sklearn, TensorFlow

Publications

NLPExplorer: Exploring the Universe of NLP Papers
Demonstrations, European Conference on Informataion Retrieval 2020
Monarch Parmar*, Naman Jain*, Pranjali Jain*, Jayakrishna Sahit*, Soham Pachpande*, Shruti Singh, and Mayank Singh

Academic Projects

- Citation Prediction - Tracing the Trajectory of Research Papers
  Proposed a novel method to predict citations on the basis of citation frequency distribution in paper sections and the citation intent.
- Query Translation for Cross Lingual Information Retrieval
  Devised a dictionary and word2vec based approach to translate search queries in Hindi, Gujarati, Bengali, and Telugu to English for supporting cross lingual information retrieval for english documents.

- Analyzing Patterns in Social Learning: StackOverflow and the Dota Community
  Performed network analysis of two online communities, StackOverflow and the Dota community, to analyze learning methodologies, namely learning through observation and learning through active participation with feedback. Analyzed the model to capture knowledge diffusion among users, and the transition of novices to experts.
- Character Recognition in Google Street View for Automated Geo-tagging using K-nearest neighbours
  Developed a web application that identifies characters from Google Street View images and automatically drops a pin on Google Maps with the image label.
- Parallelization of CYK Syntactic Parsing on GPU
  Parallelized CYK Syntactic Parsing to find the most likely syntactic parse tree of a sentence given a weighted context-free grammar learned from a tree bank. Achieved a speedup of $\sim$10.
- Cellular Automata Simulation to Model Forest Fire Spread
  A cellular automata simulation on a grid representing a forest, to study the spread of forest fires.

Awards & Responsibilities

- Certificate of Merit for being in top 1.5% students in the CBSE AISSCE 2013
- Dhirubhai Ambani SSC Merit Reward for outstanding performance at CBSE AISSCE 2011
- Core Committee Member - Programming Club DA-IICT
- Editor Entelechy - DA-IICT University Magazine