140W, 109th St, #1E, New York
NY – 10025

RHEA GOEL

(425) 681-6534
rg2936@columbia.edu
LinkedIn

EDUCATION

Columbia University, New York **GPA: 3.6** **Aug 2014 – Present**
M.S. in Computer Science (*Machine Learning Track*) *Expected Graduation: Dec 2015*

IIIT Delhi, India **GPA: 9.1** **2010 – 2014**
B. Tech in Computer Science

SKILLS

LANGUAGES/WEB TECH	FRAMEWORKS/TOOLS	DATABASES/OS
JAVA, PYTHON, MATLAB, SQL, R, PHP, PERL, C, HTML, CSS	HADOOP, MAHOUT, AWS, STANFORD NLP API, UIMA, MALLET, WEKA	MYSQL, MONGODB, NEO4J, LINUX/UNIX, MAC OS X, WINDOWS

WORK EXPERIENCE

- **Graduate Teaching Assistant** **Sept 2014 - Present**
  - Computer Science Theory Fall'14, Spring'15 (Dr. Alfred Aho, Dr. Tal Malkin)
  - Discrete Mathematics Spring'15 (Dr. Ilia Vovsha)
- **NLP based Virtual Assistant for Banking and Insurance at Infosys Ltd., India** *(Intern)* **May 2013 - Jul 2013**
  Dr. Lokendra Shastri, G.M. Research, Infosys Ltd.
  - Conceptualized and developed a domain-specific, activity-driven virtual assistant for banking inquiries
  - Incorporated context-awareness for situated cognition, episodic information for personalization, and probabilistic voice response for realistic user experience
  - Achieved feature extraction using Infosys Semantic Extraction Engine (iSEE) and Stanford CoreNLP toolkit

PROJECTS

- **Social Power in INteractions (SPIN) (Apache UIMA Java Framework, Python, Perl)**
  WISR Research Group, Center for Computational Learning Systems, Columbia University
  - Analyzing the manifestations of administrative and influential power in online written interactions on Wikipedia Discussion forums. Used structural features (verbosity, thread structure), lexical features (tokenization, POS tagging) and deep NLP processing (Dialog Act Tagging and Overt Display of Power)
- **Yelp-er: Information Extraction on Yelp Database** *(JavaScript, Python, Java)*
  Big Data Analytics
  - Generated HeatMap to better visualize cuisine-specific search query results for the users
  - Performed topic modeling using Latent Dirichlet Allocation (LDA) on review text to identify business trends
  - Proposed a gamification model for Yelp to enhance user engagement
- **Celebrity Doppelgänger** *(MATLAB, Java, Python)*
  Computer Vision/Pattern Recognition
  - Built a tool to find celebrity look-alikes ranked by percentage similarity, by extracting Histogram of Gradients and Local Binary Patterns on the images in Public Figures (PubFig) database
- **SUVIDHA: Web-based Campus-wide Complaint Portal** *(Django, Java, MySQL)*
  Human Computer Interaction
  - Deployed an online complaint portal enabling campus residents to file housekeeping complaints/issues faced in the premises and aiding Facility Management Services in efficient redressal
- **Breast Cancer Diagnosis using Online SVM** *(MATLAB, Java, C)*
  Machine Learning
  - Implemented an Online 2v Support Vector Machine for breast cancer diagnosis based on real-valued cell attributes extracted from images of benign and malignant samples.
- **Movie Recommender System** *(Python, MATLAB)*
  Collaborative Filtering
  - Developed a series of evolving movie recommender systems working on the MovieLens database
  - Incorporated Potter and Koren-Bell estimates for user and item bias, variance weighting, neighborhood selection and rating normalization and applied matrix factorization methods.