# SHREYA JAIN

Mobile : 646-954-1265 ◇ Email : sj2842@columbia.edu ◇ Github : shreyajn

## EDUCATION

**Columbia University, New York, NY** **Aug. 2017 – Dec. 2018**
**Master of Science in Computer Science, Specialization - Machine Learning**
Relevant Coursework – Machine Learning, Natural Language Processing, Deep Learning for Computer Vision

**Birla Institute of Technology and Science, India** **Aug. 2013 – Jul. 2017**
**Bachelor of Engineering (Honors) in Computer Science**
Merit scholarship awarded for all semesters for excellent academic performance.

## TECHNICAL SKILLS

| | |
|---|---|
| **Programming** | Python, R, SQL, Tensor Flow, Keras, Java, C++, C, HTML, CSS, Javascript, Torch-Lua |
| **Technologies** | Mac, Linux, Circle CI, AWS, Git, Latex, Docker, GCP, ROS, Postman |
| **Packages** | Pandas, Scikit-Learn, OpenCV, NLTK, spaCy, Flask, Matplotlib, Gensim, Scrapy |

## WORK EXPERIENCE

**Columbia University, New York, NY** **Sep 2018 – Dec 2018**
*Teaching Assistant, Natural Language Processing*

· Responsibilities include holding grading assignments, office hours and recitations for students.

**SAP Ariba, Palo Alto, CA** **Jun 2018 – Aug 2018**
*Machine Learning Intern - Search*

· Implemented deep learning models to predict the category of product from search query giving accuracy of 87%
· Designed branched CNN with Conceptnet Numberbatch and ELMO embeddings (Keras/Tensorflow)
· Developed a web application in Flask that takes about 70ms per inference to show predictions.
· Deployed multiple versions of the model in production using Tensorflow serving and GRPC (20ms/response).

**IIT Kanpur, Kanpur, IN** **Jun 2016 – Dec 2016**
*Research Intern*

· Developed Image captioning model using Convolutional and Recurrent neural networks using Torch-Lua
· Extended the Densecap model to correlate the captions of overlapping bounding boxes by predicting preposition and pass it through a LSTM to give enriched sentences as output.

## PROJECTS

**Snapsend** **Jan 2018 – May 2018**

· Developed a web app which enables sharing of images with users and downloadable in different formats and sizes
· Backend designed in Python (Flask) using MySQL as our database and front end designed in React.
· Dev, stagging and production separately maintained on GCP and were updated using CI/CD (Circle CI)

**Deep Fiction** **Jan 2018 – May 2018**

· Developed image captioning model to generate captions that captures the various aspects of the images verbally
· Implemented a RNN encoder decoder using BookCorpus dataset which maps each passage to skip thought vector
· Stylized the captions (preserve thought) to produce romantic stories for the image using this encoder decoder.

**NER API and Social Network Analysis (Graduate Research)** **Jan 2018 – May 2018**

· Developed NER API for History Lab to tag entities from raw data from SQL server eliminating manual annotation
· Implemented Matching algorithms to verify persons names in the NER tagged data from wiki or other sources
· Generated Social networks using NetworkX, Gephi and Stata using degree centrality and other measures

**Machine learning project - sentiment analysis of the US presidential candidates** **Nov 2015 – Dec 2015**

· Used twitter API to collect tweets and perform sentiment analysis on the US presidential candidates.
· Implemented ML algorithms to classify tweets as positive, negative or neutral using programming language R.

## PUBLICATIONS

**Machine Learning Model to Monitor the Progression of Parkinson's disease** **Apr 2016**

· Implemented predictive model to analyze the state of Parkinson's disease from the audio features of speech tests.
· Improving Accuracy in Noninvasive Telemonitoring of Progression of Parkinsons Disease using Two- Step Predictive Model was presented in EECEA, 2016. Later, it was published in IEEEXplore and Research Gate.