# AAKASH BHATIA

1 Coed Lane, Stony Brook, NY, 11790 +1 (408)-668-5501 asbhatia@cs.stonybrook.edu aakashsbhatia
aakashsbhatia2 https://aakashsbhatia2.github.io/

## PROFESSIONAL EXPERIENCE

**Stony Brook University** **Stony Brook, NY**
*Graduate Researcher (NLP and ML) - Advised by Dr. Ritwik Banerjee* *Jan 2020 to present*

- Research: Study textual entailment and semantics in medical information.
- Master's Thesis: Introduction of a novel, crowd-sourced dataset to study textual entailment and semantics in medical news along with standard and deep learning models to classify this dataset.

**Ernst & Young LLP** **Mumbai, India**
*Risk Analytics - Information Technology Risk Consultant* *June 2016 to July 2019*

- Performed anomaly detection on information technology systems for clients in the Financial Services sector.
- Reduced weekly effort by over 75% by developing applications to automate analytics procedures in technology audits.
- Headed a team of 6 to deliver a risk analytics engagement for a multinational Investment Bank. My team was awarded the EY Spotlight Award for delivering exceptional quality on this engagement.

## SKILLS

**Programming Languages:** Python, SQL, JavaScript, C, C++, Java

**Data Science Libraries:** Pandas, Numpy, Scikit-Learn, NLTK, TensorFlow, PyTorch, d3.js

**Deep Learning Frameworks:** DAN, CNN, RNN/ LSTM, GRU, Transformers, BERT

**Technologies:** Docker, Apache Hadoop, Apache Spark, IBM DB2, Linux Operating System, Git

## ACADEMIC AND RESEARCH PROJECTS

**Authors sentiment prediction** **Stony Brook University**
*Research project: CSE 538 - Natural Language Processing* *Fall 2020*

- Problem statement: Infer the sentiment of an author towards the main entity in a news article.
- Developed a pipeline consisting of two BERT sequential classifiers to generate focused representations of an article by leveraging paragraph level information.
- Developed a modified deep averaging network using a BERT encoder and paragraph level averaged embeddings.
- Technologies Used: Python, PyTorch, Huggingface Transformers.

**Deep Learning models for Natural Language Processing** **Stony Brook University**
*Course projects: CSE 538 - Natural Language Processing* *Fall 2020*

- Developed a Deep Averaging Networks (DAN) and Gated Recurrent Units (GRU) to perform sentiment analysis. Studied the information learned at each layer using linear probes.
- Developed a neural-network based transition parsing (arc-standard algorithm) model with a custom cubic activation function. Studied the performance of cubic activation vs tanh and sigmoid activations.
- Developed a bi-directional GRU with custom attention layer to perform entity-relation extraction.
- Technologies Used: Python, TensorFlow

**Classification of Deceptive Hotel Reviews** **Stony Brook University**
*Research project: CSE 512 - Machine Learning* *Spring 2020*

- Re-implementation of an NLP Research paper to perform classification of truthful and deceptive reviews for the top 25 hotels in Chicago obtained from Trip Advisor. Classification was performed using Naive Bayes classifier and Support Vector Machines (SVM).
- Key Finding: SVM classifier using trigrams and tf-idf performed the best. Accuracy of 89.75% and F1 score of 89.94% was achieved.
- Technologies Used: Python, Scikit-Learn, NLTK

**COVID-19 Analytics using Hadoop and Spark** **Stony Brook University**
*Course project: CSE 532 - Theory of Database Systems* *Spring 2020*

- Studied the spread of COVID-19 from Dec '19 to Mar '20 using MapReduce.
- Statistics derived: No. of cases per million, No. of cases for a given time-period, No. of cases per country, city and zip-code.
- Technologies Used: Docker, Python, Java, Apache Hadoop, Apache Spark

## EDUCATION

**Stony Brook University** **Stony Brook, NY**
*M.S. Computer Science (Thesis) - Data Science and Engineering* *Expected Graduation: May 2021*
**Relevant Coursework:** Machine Learning, Natural Language Processing, Artificial Intelligence, Data Visualisation and Visual Analytics, Theory of Database Systems, Analysis of Algorithms

**University of Mumbai** **Mumbai, India**
*B.E. Computer Engineering* *Aug 2012 to May 2016*