AAKASH BHATIA

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

WORK EXPERIENCE

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

- Research: Investigating misinformation in medical news using textual entailment and semantics.
- Master's Thesis:
  - Introducing a novel data-set with over 10,000 sentences to study textual entailment and semantics in medical information.
  - Developing a crowd-sourcing pipeline using Amazon Mechanical Turk (MTurk) to generate negative samples for classification.
  - Classifying this dataset using standard and deep learning-based ML models.
  - Performing a quantitative analysis of results to highlight gaps in the learning task.

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.
- Led the EY India team of 6 people to deliver a risk analytics project for a multinational investment bank. This was a global engagement spanning 5 countries. My team was awarded the EY Spotlight Award for delivering exceptional quality for this project.

SKILLS

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

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

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

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

ACADEMIC AND RESEARCH PROJECTS

Author's sentiment prediction **Stony Brook University - Fall 2020**

- Developed deep learning-based models using state-of-the-art BERT encoders to infer the sentiment of an author towards the main entity in a news article.
- This is known to be a difficult task. The best model obtained an F-1 score of 0.48. Analysis of the results showed that the model found it difficult to distinguish between author specific sentiments and the general sentiment of an article
- Technologies Used: Python, PyTorch, Huggingface Transformers.

Deep Learning models for Natural Language Processing **Stony Brook University - Fall 2020**

- Developed a neural-network based transition parsing (arc-standard algorithm) model with a custom cube activation function. The inputs to the model consisted of words, POS tags, and arc labels. Analysis of the results showed that the cubic activation captures the relation between the 3 inputs better than tanh and sigmoid non-linearities.
- 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 - Spring 2020**

- Re-implemented an NLP research paper to classify truthful and deceptive reviews obtained from Trip Advisor. Classification was performed using Naive Bayes classifier and Support Vector Machines (SVM).
- 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 - Spring 2020**

- Used MapReduce to generate statistics related to COVID -19 from Dec '19 to Mar '20. The statistics that were derived were - number of cases per million, number of cases for a given time-period and number 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*
CGPA: 3.42/4.0
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*
CGPA: 3.40/4.0