Amanpreet Singh

+1 631-312-2565

amanpreet.singh@stonybrook.edu

/amanpreet-singh-k

Education

- **Stony Brook University** — New York, U.S.A.
  Master of Science in Computer Science; GPA: 3.93 — *2019 – 2021*
  - **Courses**: Analysis of Algorithms, Natural Language Processing, Data Science, Probability & Statistics
  - **Teaching Assistant**: CSE 354 - Natural Language Processing, CSE 214 - Data Structures
- **University of Mumbai** — Mumbai, India
  Bachelor of Engineering in Information Technology; First Class with Distinction (72.9%) — *2011 – 2015*
  - **Courses**: Data Structures, Artificial Intelligence, Discrete Mathematics, Databases

Projects

- **Toxic Online Comments**: Toxicity classification in Wikipedia Comments and Twitter datasets by applying deep learning and transfer learning. Analyzed the results of sequential LSTM/GRU against state-of-the-art BERT models.
- **Chess Player Ratings**: Predicting the Elo rating of a chess player from the moves sequence. Efforts involved EDA and feature engineering using Pandas and Matplotlib; as well as modeling with Linear Regression and Random Forest.
- **Relationship Extraction**: A Bi-GRU model with Self Attention and Dropout layers to evaluate its performance on the SemEval-2010 Task 8 dataset.
- **Sentence Representations**: Multi-layered Deep Averaging Network(DAN) and GRU models implementation using TensorFlow and Keras API. The results are then evaluated on IMDB movie reviews for sentiment classification.
- **Skip-gram based Word2Vec**: A Word2Vec implementation with cross entropy loss and noise contrastive estimation as the objective functions. Evaluation is done on the semantic task of word analogies.
- **Physual(*Capstone*)**: A text to scene generation system for Physics word problems based on StanfordCore NLP, Java 3D and Blender Models.
- **Miscellaneous**: News Articles Classifier and Similar Document Clustering System based on TF-IDF using NLTK, Products Review Sentiments, Songs Recommendation, and Handwritten Digit Recognizer.

Publication(s)

- **Visualization of Mechanics Problems based on Natural Language Processing**: Summarized the research efforts and results of the aforementioned Capstone Project. (*International Journal of Computer Applications - April,2015*)

Experience

- **J.P. Morgan Chase & Co.** — Mumbai, India
  Senior Application Developer — *Feb 2018 – Aug 2019*
  - **NLP Query Service**: An interactive system to resolve user queries that uses a model trained on the CRF classifier from StanfordCore NLP and returns the nearest possible solution from an existing knowledge base.
  - **Trader Analytics**: Introduced statistical enhancements in the core application such as absolute and percent variance, market share and standard deviation of historical stock prices to aid in trading decisions.
  - **Real-Time Pricing**: Developed a component using Spring, JMS and TDD principles that approximates real-time market risk using live prices; and publishes out the result. It helped retire a legacy system saving the firm ∼$250k.

  Application Developer — *July 2015 – Jan 2018*
  - **Risk Management System**: Worked extensively on the core app used by traders for visualizing and hedging risk;
    1. Optimized the data feed using LMax Disruptor, a low latency Java queue for upto 20% faster processing.
    2. Process startup time improvement by 50% through the use of Multi-threading and Spring annotations.
  - **Market Data Source**: Framework for validating the functionality of a critical market data publishing app and reporting results using Java MXBeans and Apache POI. Reduced the manual testing effort by 90%.
  - **Quick-Deploy**: Streamlined application deployment, startup and monitoring on Unix production servers.
  - **MongoDB High availability**: Team point of contact for MongoDB. Built a mechanism to switch from replica set to standalone instance on the fly in case of a data center failure ensuring business continuity.

Technical Skills

- **Languages**: Python, Java, Unix Shell Scripting, SQL, MATLAB
- **Frameworks**: TensorFlow, Pandas, NumPy, Scikit-Learn, Keras, Matplotlib, Spring, Swagger
- **Databases**: Sybase ASE, MongoDB, MySQL