Kushagra Seth

San Jose, CA

📱 +1 (669) 203 - 9966 | ✉ kushagraseth.1996@gmail.com | github.com/Kdotseth7 | in linkedin.com/in/kushagra-seth

Education

University of California Santa Cruz (UCSC)   *Santa Cruz, California*
Masters of Science (MS), Natural Language Processing (NLP)   *Sept 2022 - Dec 2023**
- **Tutor/Reader:** CSE13S(Computer Systems and C Programming)
- **Courses:** Introduction to NLP, Deep Learning for NLP, Transformers, Machine Translation

Guru Gobind Singh Indraprastha University (GGSIPU)   *Delhi, India*
BTech in Information Technology (IT)   *Aug 2015 - July 2019*
- **Courses:** C, Java, Python, Artificial Intelligence, Data Structures, DBMS, Operating Systems

Technical Skills

Programming   Python, C, C++, Scala, Java, JavaScript, React, SQL, Bash, Selenium (Java/Test-NG), Karate.
Libraries   PyTorch, Transformers, Datasets, Evaluate, Keras, Scikit-Learn, HuggingFace, Pandas, NumPy, NLTK, SpaCy, Matplotlib, Seaborn, Re.
Miscellaneous   Docker, Git, Jenkins, JIRA, Data Science, Machine Learning(ML), Deep Learning, Natural Language Processing(NLP), Artificial Intelligence(AI).

Work Experience

Software Engineer (Machine Learning)   *Gurugram, India*
Cvent   *June 2019 - July 2022*
- Built the product - **Lead Score and Engagement**. It required a multi-step design and integration with partner systems like Salesforce and Marketo using a mesh of microservices.
- Developed a **Content-Based Recommendation System** for the Cvent Platform to make session recommendations for event attendees.
- Led and implemented a project on a Deep learning-based **Sentiment Detection system** for Session and Exhibitor Reviews for the Cvent Platform using LSTM in PyTorch.
- Led a project on creating and maintaining fully automated CI/CD infrastructure using Jenkins for multiple production environments. It reduced the end-to-end deployment time by more than **50%**.

Data Science Intern   *Gurugram, India*
SS Supply Chain (3SC) Solutions   *June 2018 - July 2018*
- Designed and built a front-end dashboard for different Vendors and Retailers.
- Developed Deep learning **Regression** models using Keras for the delivery date and freight cost prediction.
- Implemented and compared the performance of Seq-to-Seq Models, namely **RNN**, **LSTM**, and **GRU**, on historical Vendor and Retailer data

Data Science Intern   *Gurugram, India*
Team Computers   *June 2017 - July 2017*
- Cleaned, filtered, and merged datasets for the past year and calculated Turn Around Time (TAT) for each contracted inquiry from the merged dataset.
- Developed a Regression Model based on the contracted inquiries and the vehicles delivered to predict Inquiries for next year and generated a cumulative dataset, i.e., the predicted number of vehicles Honda dealers will sell based on outstanding inquiries.
- Trained Regression Models using **Logistic Regression**, **Decision Trees**, **KNN**, and **SVM**.

Publications

CLIMATE CHANGE STANCE CLASSIFICATION using Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), RNN-LSTM & CNN-LSTM
  Kushagra Seth, S Sahil Patwal, Amita Misra, Amita Goel
    JETIR 6.5 (2019) pp. 35–46. 2019

Project Work

IOB Slot Tagging of Natural Language Utterances
University of California Santa Cruz (UCSC)
- Developed a Deep learning model using **GloVe Embeddings** and **LSTM** using PyTorch to tag IOB slots in natural language utterances.
- Improved performance using **Bi-LSTM** and using variations of Adam optimizer like AdamW and Adamax. Achieved an **F1 score** of **0.81** on the test dataset.

Cross-Lingual Projection using T5 and Fine-Tuning for SNLI using DistilCamemBERT
University of California Santa Cruz (UCSC)
- Implemented cross-lingual projection of the SNLI dataset from English to French using transformer-based **T5** model.
- Performed **fine-tuning** on the French dataset using state-of-the-art transformer-based **DistilCamemBERT** model from HuggingFace to improve the performance of the SNLI task. Achieved an **F1 score** of **0.87** on the test dataset.

Multi-Evidence Natural Language Inference(NLI) for Clinical Trial Data
University of California Santa Cruz (UCSC)
- Developed and implemented a **Classification** Model to predict labels for Premise-Hypothesis pairs of clinical data using Transformer models **BERT** and **deBERTa** using HuggingFace and PyTorch.
- Achieved an **F1 score** of **0.71** on the test dataset.