Sri Harsha Pamidi

3869 Miramar Street, #3840, La Jolla, CA - 92037
+1(858)241-4848 • spamidi@ucsd.edu • sri-harsha-pamidi • Sriharsha-pamidi

Education

University of California San Diego, *Master of Science(MS)* **Sep 2021 - Mar 2023**

- Major in Electrical and Computer Engineering, Machine Learning and Data Science Specialization - 3.93/4.00

Relevant Courses: *Statistical Learning I & II*, Neural Networks for Pattern Recognition*, Probability and Statistics for Data Science, Linear Algebra*, Recommender Systems and Web Mining, Programming for Data Analysis.* *(*) - ongoing*

Indian Institute of Technology Kharagpur, *Bachelor of Technology(Honors)* **July 2015 - May 2019**

- Major in Electrical Instrumentation Engineering - 8.72/10.00**,** Minor in Computer Science and Engineering - 9.12/10.00

Relevant Courses: *Deep Learning, Advanced Machine Learning, Speech and Natural Language Processing, Artificial Intelligence, Data Analytics, Algorithms-I, Programming and Data Structures.*

Work Experience

- **Senior Software Engineer, Machine Learning,** *Capillary Technologies, India.* **Jan 2021 - Jul 2021**
  - Constructed a framework to reduce latency by predicting the data that will be used at a particular time and pre-caching it.
  - Worked on migrating batch ETL to Delta Lake and implementing the Kappa architecture to support streaming data processing.
- **Software Engineer, Machine Learning,** *Capillary Technologies, India.* **Jun 2019 - Jan 2021**
  - Worked on in-house ETL Spark data pipelines for a multi-tenant analytics tool for offline events using configurable metadata.
  - Developed the backend of a query engine for self-serving analytics using star schema properties powered by Apache Spark.
  - Containerized microservices and migrated them to Kubernetes cluster for auto-scaling.
- **Deep Learning Intern,** *Real-time Human Emotion Detection, STEP, IIT Kharagpur.* **Jul 2018 - Sep 2018**
  - Designed a convolutional neural network that predicts real-time emotions from dynamic facial expressions and returns a confidence value across a set of emotions for each face in the image.
  - Implemented and deployed the model onto Raspberry Pi taking into account the memory and processor speed limitations.
- **Data Scientist Intern,** *Modeling Context in Word Embeddings, Envestnet Yodlee, India.* **May 2018 - Jul 2018**
  - Implemented a model that generates context-based word embeddings by replacing the context modeling of CBOW architecture with a neural model consisting of bidirectional LSTMs.
  - Improved the accuracy of company's existing sentence classification model by 2.1% by adding this to the embedding layer.

Selected Projects

- **Product Size Recommendation using Fit Semantics** **Oct 2021 - Dec 2021**
  - Explored various predictive algorithms to address the product fit problem in the highly imbalanced RentTheRunway Dataset.
  - Implemented fully connected NN model for multi-class fit prediction utilizing product attributes as well as semantics of fit feedback using GloVe word embeddings to improve the F1-score by 5.36% and accuracy by 1.42% over the baseline.
- **Emotional Analysis of Movies using Subtitles** **Oct 2021 - Dec 2021**
  - Performed exploratory data analysis on movie subtitles to analyze the role of emotional dialogues on movie IMDb ratings.
  - Utilized temporal emotion data to extract patterns and similarities in movies created by same crew or franchise.
- **Delineation of Hematoma and Estimation of Midline Shift** **Aug 2018 - Nov 2018**
  - Performed segmentation of hematoma on CT images and estimated the disturbed symmetry of the brain using midline shift.
  - Implemented an effective image thresholding method for preprocessing the scans to improve classifier performance.
  - Developed an algorithm to estimate the midline shift of the brain using the characteristics of mass segmented.
- **Synonym Set Extraction from Biomedical Literature** **Aug 2018 - Nov 2018**
  - Identified synonyms for Drugbank vocabulary entities in SPL dataset containing adverse reactions for over 200 drugs.
  - Developed a distributional semantics model to form target word vectors and used similarity measures to extract synonyms.
- **Electronic Tongue and Electronic Nose in Tea Industry** **Aug 2017 - Nov 2017**
  - Designed the instrumentation of E-tongue and E-nose for estimating tea quality in the Manufacturing Plant.
  - Estimated tea quality by various pattern matching techniques using electronic signatures of tea.

Technical Skills

- **Languages:** Python, Java, Scala, SQL, C/C++, R , MATLAB.
- **Machine Learning Libraries:** Numpy, Pandas, Scikit-Learn, Keras, Tensorflow, PyTorch, NLTK, OpenCV.
- **Big Data:** Apache Spark, MapReduce, Delta Lake, Hadoop, Amazon EMR.
- **Additional Tools:** Spring Framework, Neo4j, MongoDB, Query DSL, Docker, Kubernetes, Git, Jenkins, Intellij IDEA.