DIVYA NANDLAL SAHETYA

+1 (650)-231-9239 |sahetya@usc.edu | linkedin.com/in/divya-nandlal-sahetya/ | GitHub | Portfolio

EDUCATION

Master of Science in Electrical and Computer Engineering, University of Southern California
Linear Algebra, Probability, Machine Learning, Deep Learning, Algorithms, Cloud Computing
Bachelor of Engineering, Electronics and Communication, Sri Jayachamarajendra College of Engineering
Image Processing, Digital Signal Processing, Data Structures

Los Angeles, USA
August 2021-May 2023
Mysore, India
August 2015-May 2019

PATENTS AND PAPER PUBLICATION

System and method for dynamic translation of speech to Sign Language for Oral Health Education – IN Patent 201841039995, IJRASET

• Engineered an Automatic Speech Recognition system using CNN for dynamic translation to sign language for oral hygiene education.

System and method for Cleft Speech Training at home – IN Patent no. 202041045850

• Devised a machine learning system trained with MFCC features of speech samples for assisting partially speech disordered individuals to improve speech with interactive learning experience.

TECHNICAL SKILLS

- Programming/Scripting Languages: Python, R, Java, C++, C, MATLAB.
- AI Framework and tools: PyTorch, Keras, Tensorflow, PySpark, Pandas, Numpy, OpenAI, LLM, Langchain, HuggingFace, CNN, GANs, RNN, LSTM, Transformers, Object Detection SSD, RCNN, YOLO.
- Technologies: React JS, Flask, REST API, GraphQL.
- Data Skills: SQL (MySQL, PostgreSQL), NoSQL (MongoDB), Spark, Kafka, Hadoop, MapReduce, Hive, Tableau, Power BI.
- DevOps and Documentation tools: Kubernetes, Docker, Amazon Web Services, GCP, Git, JIRA, Confluence, Jenkins.
- **Domain Knowledge:** Probability & Statistics, Machine Learning, Deep Learning, Natural Language Processing, Software Engineering, Data Engineering, ETL Pipelining, Data Visualization.

WORK EXPERIENCE

Research Assistant Los Angeles, USA

Machine Learning for Medical Images - USC

Jan 2023-Present

- Proposed Novel Meta-Learning framework for few-shot multi-organ tumor segmentation (*Medical Segmentation Decathlon Challenge*), using dynamically weighted task subsampling and meta-update rules.
- Improved accuracy by 4% over state-of-the-art **Reptile** framework using **Pytorch** framework.

Research Assistant

Los Angeles, USA

Biomedical Imaging Group – USC

May 2022-August 2022

- Developed a deep learning approach to denoise human and mouse brain MRI images using auto-encoder architecture using Keras.
- Obtained a cleaner image compared to traditional image processing approach employed in the BrainSuite software.

Software Developer

Bangalore, India

Siemens Healthineers

January 2019-July 2021

- Log Sanitizer Constructed an end-to-end Python executable tool Log Sanitizer along with the team to encrypt Patient Health Information to comply with the HIPAA and deployed the **Docker** and **Kubernetes**.
- Parallel Processing Improved the processing speed of the tool to encrypt ~1 million amount of log files live from different Business units (B2B and B2C) by using multithreading and multiprocessing.
- **Python Libraries fixes** Enhanced the performance of conversion of EVTX and EVT files to XML in Python-Evtx and Python-Evt libraries by introducing multiprocessing.
- File Transfer Module Collaborated with other teams and developed a messaging module in Java to interact between two sister applications namely, the sanitization tool and file transfer using JMS Messaging Module.
- Storage Optimization Leveraged Amazon S3 buckets for efficient and scalable storage of large data volumes, minimizing duplication and optimizing data management during processing.
- Statistics Plugin Designed and implemented a real-time statistics plugin to generate performance metrics for the tool, enabling data-driven decision-making. Employed Power BI for seamless data visualization and analysis of the generated metrics, enhancing the understanding and interpretation of system performance.

ACADEMIC PROJECTS

Analysis of Supervised and Semi-Supervised Machine Learning for Cervical Cancer Diagnosis | Github

- Attained an accuracy of **82.67**% with Random Forest with Supervised Learning and attained an accuracy of 90.55% with Semi Supervised learning based on ML Learning theory.
- Designed a machine learning system to detect if a person has the risk of cancer based on the risk factors (categorical) data collected from patients and the results of four diagnosis tests namely Hinselmann, Schiller, Cytology, Biopsy.

Predicting Forest Fire in Algeria Using Machine Learning Techniques | Github

- Obtained a highest accuracy of 90% with a Logistic Regression classifier with standardized input and feature reduction using PCA.
- Constructed a comparative machine learning system to predict forest fires using various classifiers on a dataset containing numeric and categorical features.

American Sign Language Recognition using Deep Learning Techniques | Github

- Formulated an image recognizer using pre-trained ResNet (ResNet18, ResNet34, ResNet50) models in PyTorch.
- Deployed the model and obtained an accuracy of 96% 98% on the test set.
- Extrapolated the model to test on an external American Sign Language dataset and achieved a highest accuracy of 74% for ResNet50.