|  |  |  |
| --- | --- | --- |
|  |  | Sreeneha SAMUDRALA  Student |
| Profile Aspiring student looking for exciting internship opportunities in data science, ML and/or full stack development Contact PHONE:  +91 7358292086  EMAIL:  neha.samud@gmail.com Skills - Python  - Java  - C / C++  - Scikit Learn  - Azure ML Studio  - Pandas and numPy packages  - EDA and Feature engineering  - Natural Language Processing  - PyTorch  - Spacy  - NLTK  - Transformers  - Computer Vision  - OpenCV package  - PIL package  - Image processing |  | EDUCATIONVIT Chennai 2019 - 2023  B. Tech in Electronics and Computer Engineering Area of Interest  * Data Science and Machine Learning * Full Stack development * Computer Vision  EXPERIENCE **Worked as a Data Science (NLP) Intern at Discite Analytics and AI for 4 months. [June 2022 – September 2022].**   * I worked on projects based on NLP and Computer Vision for clients. * A few examples are: * Text summarizations (abstractive, Extractive, Pegasus, CNN summarization models, HuggingFace Transformers, Spacy, PyTorch, BERT models etc.) * Essay grading (Criteria based like, grammar checking, vocabulary, spelling error and other NLP techniques), * Sentiment analysis (using traditional classification and text preprocessing and using GPT3 models), * automated post generation and quiz generator using various summarization models and OpenAI-GPT3 * Creating real-time Face and hair filter making using openCV and Augmented reality. * Working with Docker, Streamlit, HuggingFace Spaces and Google Colab for model Training, development, and deployment (MLOPs)   **Worked as an AI/ML intern at Appcilious pvt ltd for 2 months**  **(Dec 2021 – Jan 2022).**  Key Highlights: -   * worked on speech and voice processing projects, like voice cloning and text to speech models. * worked on Computer Vision image processing projects like lip-sync, background changing, video matting and face enhancement. * done research on the above topics to find the right useful packages and GitHub repositories. * trained a forward tacotron model for text-to-speech conversion on python * trained a model to lip-sync an image to a given audio from the tacotron model * worked on face enhancement by splitting video into frames * worked on changing the background to a custom video or photo background using CV techniques. * integrated the Lip-sync and background changing models using shell script in python * Developed an API on nodeJS to invoke the service to generate a custom lip-sync video  PROJECTS (GitHub: <https://github.com/neha04-12>) **Completed Projects:**   * **Handwritten Signature recognition and matching using OpenCV and CNN**   I have collected handwritten signature data, scanned and applied image preprocessing using Compute Vision techniques on them to format the entire dataset. Then applied image augmentation on the dataset. Trained a CNN to extract feature vectors for each signature and compare two of them to check their similarity. Lastly used decision trees and logistic regression for classification of train, validation and test set.   * **Digit Recognition Using Neural Networks and CNN**   This project is an implementation of neural networks from scratch with different cost functions, an activation function and back propagation algorithm written on my own. The two models with quadratic cost and Cross entropy cost give good accuracy on training as well as validation data. I have also implemented Neural Networks using TensorFlow with SoftMax function. I have even approached the problem statement with CNN which gave a pretty high accuracy.   * **Sentiment Analysis of Product reviews using natural language processing**   Built a classification model of the sentiments of reviews given by user using NLTK and skLearn   * **Speech generation/Prediction project:**   Built a model to predict the upcoming words, given a seed word as input, based on a text file data (Speech of a national Leader)   * **Semantic Textual Similarity:**   Finding the similarity metrics between two documents using cosine similarity and sentence transformers   * **Resume Parsing using NLP:**   Reading a list of resumes and extracting the educational qualification, year of study and institution name.   * **FAQ Virtual Agent:**   I Have developed an FAQ answering virtual agent using natural language processing, sentence transformers and cosine similarity of documents. I have collected my own data to train the model. I have used FAQs of Swiggy collected from Glassdoor.   * **Video Engagement Prediction**   This project was a solution to Analytics Vidhya Job-a-thon. I preprocessed the data, handles missing values, transformed data to categorical values (One hot encoding) to make it an easier problem statement. Lastly, I used Linear regression for training on train data and validation data for the model accuracy. Lastly predicting on test set.   * **Employee Churn Prediction**   This project was a solution to Analytics Vidhya Job-a-thon. Firstly, I performed univariate and multivariate analysis on the given dataset. And concluded the important features for the target variable. Next was preprocessing, and dealing with missing values. I performed one hot encoding and bucketed the numerical values to make it a categorical value making it easier for binary classification.   * **Systematic Momentum Investing: Back testing Results for  Indian Market**   This project is an implementation of the systematic momentum model to predict returns and built Back testing strategies from scratch and Machine Learning concepts like exponential regression for momentum calculation and standard deviation for volatility calculation.   * **Phishing website detection using ML techniques**   Collected data for phishing and non-phishing websites with their label. Extracted features for the URLs to make trainina and classification easier. Trained using Random Forest Classifiers, Neural Network Classifiers and Support Vector Classifiers   * **Implementation of Non-Repudiation in Financial messaging:**   Secured exchange of financial messages using asymmetric cryptography and digital signatures.   * **Future Sales prediction model linear regression and light GBM:**   Built a model to predict total sales for every product and store in the next month based of historic data with an accuracy of 80%   * **Loan prediction in Azure ML studio:**   Predicts Loan eligibility based on historical data of loan defaults. Extra-curricular Engagements  * Active member of the IEEE computer society during 2019-2020 * Active member of the Dance Club at VIT Chennai 2020-2021   - Workshop on ML and Data Structures by Coding Blocks, 2020  - TCS CodeVita cleared round 1, September 2019  - IEEE member of ML bootcamps, 2019-2020  - Cognitive computing 2-day certification course by external  Faculty, August 2019 |
|  |  | Courses And Certifications - Finlatics Business Analyst Experience program (Finlatics)  -SQL using MySQL and Database Design (Udemy)  -NLP with Python for Machine Learning Essential Training  (On LinkedIn Learning)  - Data Analysis in Python – University of Pennsylvania  (on Coursera)  - Udemy Full stack web development course, August 2021  - Coursera, How to Win a Data Science Competition: Learn  from Top Kagglers, July 2020  - Coursera, Cloud computing Basics, July 2020 Relevant courses  * Object Oriented Programming * Problem Solving in Python * Essentials of Machine Learning * Introduction to Computer Applications * Analysis of Data Structures, Algorithms and Graph theory * Introduction to Software engineering * Operating Systems * Data analytics and visualization * Computer Vision * Data mining |