Vansh Shah

vanshshah815@gmail.com , 720 S College Mall Road Apt F4, Bloomington, IN 47401 ,+1(571)619-5077,
linkedin.com/in/vanshshah815 , github.com/vansh815

EDUCATION

Indiana University, Bloomington | M.S. in Data Science | GPA : 3.8/4 ***May, 2021***

Nirma University, India | B.TECH in Information Technology **|** GPA : 7.36/10 ***May, 2019***

Relevant Coursework: Computer Vision, Exploratory Data Analysis ,Big Data ,Machine Learning , Artificial Intelligence, Deep Learning, Statistics, Design and Analysis of Algorithms, Operating System, Computer Networks.

TECHNICAL SKILLS

LANGUAGES : Python, Java,R, C, JavaScript, HTML, CSS, XML, MATLAB.
DATABASE : MySQL,SQLyog
TOOLS : Spyder, Scilab, Android Studio, Arduino IDE, Spark, Colab, R Studio, Excel, Git.
ML & DL frameworks : Numpy, Pandas, Scikit-Learn, Keras, Tensorflow, Spark-mllib, OpenCv.

WORK EXPERIENCE

Infostretch Corporation | Data Science Intern **| Astute ml** **January, 2019 - June, 2019**

- ASTUTE is an AI testing suite which helps to optimize software testing life cycle by analyzing test cases and predicting useful information and performed various NLP methods and optimization techniques.
- Tagged the missing components using multi text classification and performed risk classification based on : “High”, “Moderate” , “Low”. Thereby achieving an accuracy of 82%.
- Learned and implemented tools and technology like pyspark, scala, PCA for dimensionality reduction, LSTM model, Naive Bayes and demonstrated daily and weekly progress.

Infostretch Corporation | Data Analytics Intern | **Data Analysis of Diabetic Patients** **May, 2018 - July, 2018**

- Implemented various concepts of supervised learning such as trees and regression algorithms.
- Learned and performed Exploratory data analysis and hypothesis testing.

Chetan Traders | Business Intern **May, 2015 - December, 2017**

- Developed a self- checkout inventory app for advertising and ease in placing orders for customers.
- Conducted and planned weekly and regional meetings for up to 25 people and worked in a team of 8 people to execute marketing campaigns.

ACADEMIC PROJECTS

Optical Music Recognition **December, 2019 - February, 2020**

- Developed an image recognition system which detects all musical notation and classifies the piano chords.
- Classified chords of the piano using Sobel operator and hamming distance. Implemented Hough transform to detect the upper staff and lower staff .
- Implemented the model using OpenCv , Pillow and implemented the model from scratch and dynamically scaled the test image.

Sentiment Analysis **August, 2019 - December,2019**

- Implemented sentiment analysis on different data-sets, from Amazon, Yelp and Imdb, using the Naive Bayes Algorithm with K-Fold. The mean accuracy reached over 80%. Both the maximum likelihood and MAP were compared for each of the data-sets.

Hyperparameter Tuning of GLM **August 2019- December,2019**

- Optimized the hyperparameters for GLMs including Logistic , Ordinal and Poisson regression.
- Used second-level maximum likelihood by approximating the posterior using Laplace approximation and observed 10 times decrease in time to find optimal hyperparameters.

Plant disease detection using CNN **July, 2018 - December, 2018**

- Developed a tool for classifying tomato diseases as healthy or diseased and if diseased, then classified them as : “early blight”,”septoria”,”bacterial spot”,“leaf mold”.
- Achieved accuracy of 99.73% using transfer learning using VGGNET and applied data augmentation for increasing the size of dataset thereby minimizing overfitting.
- Executed different optimization methods such as Relu, Adagrad , Adam for tuning of hyper-parameters to increase accuracy .

Movie Recommender System **January, 2018 - May, 2018**

- Developed a movie recommender system using different types of filters such as content, Collaborative and hybrid-based filtering.
- Implemented using Information retrieval System and Natural Language Processing and filtering techniques like tf-idf vectorisation for encoding.