Shubham Prakash

Data Scientist

An independent and self motivated person with a command on several Machine Learning algorithm like **Linear Regression**, **Decesion Tree**, **Random Forest Classifier, Logistic Regression**, Clustering algorithm like **K-Means** and **Hierarchical Clustering**, Neural Networks like **ANN, CNN, RNN, LSTM,** and **Transformers**. Know different featuring engineering techniques for data cleaning, **outlier removing** and **dealing with uneven datasets**. Have full command on **Python** and **SQL** programming language. Have worked on cloud platform of **Ubunbtu 20.04** and also know **HTML CSS JS** to deploy a production ready model in a user friendly way

✉ shubhamprakash702@gmail.com

📱 8051350686

📍 Patna

in linkedin.com/in/shubhamDprakash

github.com/shubhamDprakash

SKILLS

Python | HTML

CSS | JavaScript

Machine Learning

Deep Learning | NLP

Linux

Financial Reporting Analysis

Full Stack Web Development

SQL

LANGUAGES

English
Full Professional Proficiency

Hindi
Native or Bilingual Proficiency

INTERESTS

Martial Arts - Karate

Drawing | An Otaku

Stocks | Investment

coding

EDUCATION

B.Tech (hons) Mechanical Engineering
Indian Institute of Technology (Indian School of Mines) Dhanbad
08/2016 - 07/2020   *8.29 GPA*

CFA Level-1
CFA Institute
08/2021 - Present   *90+ percentile*

PERSONAL PROJECTS

Banglore House Price Predictor (12/2021 - Present)
- www.banglorehousepricepredict.herokuapp.com
- This is an end-to -end ML project with dataset obtained from Kaggle and website as above
- First the data is cleaned with different featuring engineering technique and outliers are removed
- Then the modelling is done with different machine learning algorithm like **Linear Regression**, **Decision Tree**, **Lasso Regression** to get the best model using **GridSearchCV**. Linear Regression is given as the best model with **85% accuracy** which is used for website
- The website is made in Python Flask Server with HTML and CSS in the frontend

Stock Market Sentiment Analysis through news (12/2021 - Present)
- www.mynewsexpress.herokuapp.com
- This is a ML project which predict the sentiment of stock market as bull or bear through the current news headlines
- The dataset for training is obtained from kaggle containing **25 news headlines** and labels; and the the current news for prediction is obtained from newsapi.org
- First all 25 news headlines are converged into a single length of string and different NLP operations like **converting to lowercase**, **stemming**, **removing stopwords** from **nltk** library are performed for data cleaning
- Then different models are used to train and get the predictions like **LSTM** with **embedding layer** and **ONE-Hot-Representation** and **padding** for vectorization, **TF-IDF Vectorization** with **RandomForest Classifier**
- Currently working on **transformers state of the art algorithms** like **BERT** to increase the accuracy

Sudoku Solver With GUI Implication (12/2021 - Present)
- This is a Python program which solves the game of sudoku with backtracking algorithm which can be visualized in the real time with the help of **openCV** library

WORK EXPERIENCE

GET
L&T Limited

Vizag, Andhra Pradesh
Achievements/Tasks
- Worked in a project for Indian Navy with DRDO as a client as a **data analyst**
- Got the last one year DSR reports of different OEMs working and analysed the current non-fair work due to operations and delays due to L&T
- Formulated and implemented an action plan with the 80-20 rule to reduce such work and delays and increase the availability of assets
- Also made an excel macro to get the asset with most failure and its cause for Root cause analysis

POSITION OF RESPONSIBILITY

Class Representative of B.Tech. Mechanical Engineering 2020 (08/2017 - 06/2020)

Organized several events in college
like Skyline, Contraption in CONSETTO-2017; AmphiCart in ASME-CISMA-2017; BASANT-2018; UMEED-2018 for FFI

Worked for a NGO named FFI (01/2017 - 04/2020)
organized several Blood donation camps and taught children of slums