Mubin Shaikh

Email: yaminmubin@gmail.com
Linkedin: linkedin.com/in/mubin-shaikh-datascientist/
Mobile: +91-8421681220
Github: https://github.com/Mubin170799

Education

- **North Maharashtra University**
  M.sc Applied Statistics; CGPA: 8.9 *SEPT 2020 - May 2022*
  Courses: *Python,Statistics, R, SAS, Time Series, Machine Learning*

Skills Summary

- **SKILLS**: Machine Learning, Python,Statistics,SQL,R language, Computer Vision,Tableau,Git,NLP,Deep Learning,Mongo dB,Apache Spark,Hadoop,Map Reduce,Kafka, Hive
- **Frameworks**: Scikit, NLTK, SpaCy, TensorFlow, Keras,Flask
- **Tools/IDE**: PyCharm, VScode, Jupyter Notebook, Databricks, Hortonworks Sandbox
- **Machine Learning**: Linear Regression, Logistic Regression, Decision Tree, Support Vector Machine, Naive Bayes, Unsupervised Learning, Ensemble Technique, Hyper parameter Tuning, Data Collection, Data cleaning, Data preprocessing, model Training.
- **Deep learning**: Ann,LeNet, Alex Net, VGG - 16, Google Net, Bert, ResNet , Computer vision, object detection

Experience

- **Data Science Intern** Remote
  Data Trained Pvt.ltd *Aug 2021 - Present*
  - **Working on projects for Machine Learning and doing tasks like:- Data Preprocessing, Model training, Model Selection, Prediction and Deployment for more than 10 Projects like - Insurance Claims- Fraud Detection, Loan Application Status Prediction and Customer Churn Analysis Projects.**:
  - : -2pt Expertise invalidating the data using EDA / ETL Techniques: Central Tendency, Dispersion, Quartile / Percentile, Standardization and Data Visualization
- **AI Engineer Intern** Remote
  Industry 4.0 Technology Consulting *OCT 2021 - DEC 2021*
  - **Worked on Supervised and Unsupervised Machine Learning algorithms with structure and unstructured data.**: Technical Skillset : Machine Learning, Deep Learning, Exploratory Data Analysis, Data Analysis, Data Visualization, Report Generation.
  - **Responsibilities- Study and transform data science prototypes. Design machine learning and deep learning systems. Develop clever algorithms and pragmatic solutions to our automation and optimization problems. Build high accuracy machine learning models that can learn and optimize performance from vast amount of data. Develop metrics to measure the outcome/impact of your introduced solutions. Work with other members to implement and integrate into our existing systems. Document and improve the solutions over time. Evaluate and identify new technologies for implementation. Communicate with our business and technical teams to understand the analytics requirements. Respond and follow up to incorporate feedback and draw new insights. Prioritize tasks to meet multiple deadlines.**:

Projects

- **Stock Market Analysis and predict the Stock price based on Reviews**: (Objective): extract information from online news feeds and then using the information so extracted to predict changes in stock price or volatilities. Hyperparameter tuning, Achieved more than 88percent of Accuracy, Deployed complete M.L pipeline on Heroku cloud platform
- **Real-Time Face Recognition(Computer Vision)**: The goal of this project is to detect and locate human faces, Performed MTCNN, Facenet, Face Embedding, Data Augmentation for the project, Developed this project using python, opencv Numpy, Used Tkinter GUI, This project has Real-time face detection with 90 Percent Accuracy
- **Worked on implementing an end to end Big data Insurance prediction Regression project(Environment: Pycharm, Mongodb stack - Kafka, PySpark, Mongodb, ML), Created Insurance prediction scripts for real-time using pyspark commands and Kafka-PySpark streaming and dumped data into mongodb, Developed data cleaning code using pyspark and pandas script, Design and development of Real-time data processing pipelines using Spark Structured streaming, Contribution on end to end delivery of data pipeline with Airflow orchestration to be ready for production deployment, Best ML Model - Random Forest Regressor**:

Publications

- **Connecting People: A Statistical Study of Telecom Service Providers - 07/2020**: International Journal For Research in Applied Science and Engineering Technology - To Understand the perception of customers with respect to service offered by the telecom service provider, Explore the services and products offered by the telecom service provider to individual customers, To study the customer satisfaction and understand their expectation

Languages

- **English** Professional Working Proficiency
  5pt
- **Hindi and Marathi** Native or Bilingual Proficiency
  5pt
- **Arabic and Urdu** Native or Bilingual Proficiency