SUBRAMANYAM TUMARADA

Data Scientist / ML Engineer

Career Objective

To secure a challenging role as a Data Scientist in a reputable organization where I can apply my theoretical understandings and practical expertise in Machine Learning, Deep Learning and Exploratory Data Analysis to solve challenging business problems. My objective is to utilize my skills and apply my knowledge to promote data-driven solutions and contribute to the company's growth.

Contact

+91- 6303543289

anupsubbu98@gmail.com

https://github.com/anupsubbu98

Certifications

- ✓ Udemy certified ***Master SQL for Data Science***.
- ✓ Udemy certified ***Statistics for Data Science & Business Analysis***.
- ✓ Completed Machine Learning for Healthcare by MIT.
- ✓ ***5 star gold badge*** in SQL & Python on HackerRank coding platform.
- ✓ ***HackerRank certified*** Python (Basic) & SQL (Intermediate & Advanced) skill assessments.
- ✓ Awarded ***Associate of the Month*** for the month of July 2023 for delivering the critical payroll implementation within time.
- ✓ Recognized by the Finance team and manager for my contribution to the project.
- ✓ Published a paper on Water level indicator using IoT.
- ✓ Continuously learning, sharing my knowledge and engaging with tech communities on twitter.

Educational Qualifications

2015-2019 : B.Tech. in Computer Science Engineering, GMRIT, Rajam

- ✓ **Areas of Interest :** Computer Networks / Data Structures
- ✓ Active member of Computer Society of India (**CSI**).

Technical Skills

- ✓ Python
- ✓ SQL / Oracle
- ✓ Exploratory Data Analysis
- ✓ Statistical modeling & Machine Learning
- ✓ Artificial Neural Networks & Convolutional Neural Networks.
- ✓ NLP / Sequence to sequence learning with Neural Networks.
- ✓ Heroku / Git
- ✓ PeopleSoft Application Designer and PeopleCode.

Profile Synopsis

- ✓ Experience of 4 Years 3 Months in Oracle PeopleSoft Expense module as a developer and in designing Artificial Intelligence applications with Machine Learning techniques includes Regression, Classification, Feature Engineering, Supervised and Unsupervised algorithms and Data Analysis.
- ✓ Skilled in libraries such as Sklearn, Tensorflow, Numpy, Pandas, Matplotlib for Data Visualization.
- ✓ An introductory understadnig of LLMs and LangChain.
- ✓ Extensive knowledge of working with SQL databases like Oracle and MySQL.
- ✓ Additional skills of Python, Java & EDA.
- ✓ Also worked independently for deployment of ML Models in Heroku and AWS.
- ✓ Rich experience in L2 ticket handling, Process monitoring & Scheduling, handling production issues, developing test plans and perform testing in multiple environments.
- ✓ Well versed in manipulating and analyzing high dimensional data.
- ✓ Ability to work independently as well as in team environment.

Professional Experience

July 2019 to present : Cognizant Technology Solutions, Bengaluru

Responsibilities :

- ✓ Built a Lithium-ion battery end life Prediction Model for the Electric bus.
- ✓ Developed a predictive maintenance algorithm which reduced the asset downtime by 26%.
- ✓ Developed Payroll payment monitoring dashboard pages for LATAM and APAC countries.
- ✓ Implemented New Zealand and China Payroll systems.
- ✓ Worked on implementation of Remittance program for Cross Country payments for APAC countries.
- ✓ Involved in fixing hundreds of defects that were raised during daily pay cycle runs.
- ✓ Performed end to end testing for Argentina and Brazil payroll systems in multiple instances.
- ✓ Production support, developments, analyzing and resolving production issues using SQL queries.
- ✓ Created many PeopleSoft objects like fields, records, pages, views, app engines during code changes enhancements and major developments.
- ✓ Performed testing in multiple environments, project migrations in CAPI and documenting the developmentss.
- ✓ L2 ticket handling, Daily Process monitoring & scheduling.
- ✓ To collaborate with the team to meet the deliverables during the deadlines.

Projects

Summarization of Online Meeting Conversations using NLP & Attention mechanism : (*August 2023 – present)*

- ✓ This is an end to end project where we convert the meeting video/audio into text using a speech to text converter.
- ✓ For this purpose, we will make use of "**Whisper**" pre-trained model by **OpenAI** to generate the text from speech.
- ✓ The text thus generated will then be summarized using **seq2seq models** which make use of encoders, decoders, **BERT** and **attention mechanism**. We will be using 2 pre-trained models to perform these 2 tasks.
- ✓ Both these pre-trained models will be fine-tuned based on our dataset and additionally, we are looking at training the models in a QA manner (Question and Answer based training) which helps the model to learn faster.

Bangalore House Price Prediction using Linear Regression : *(May 2022 – Aug 2022)*

- ✓ It is an end to end House Price Prediction Project, we have used machine learning techniques to predict the price of houses.
- ✓ This project involves analyzing historical data to predict the sale price of a house in Bangalore city using Linear Regression based on various features such as location, size, number of bedrooms, and other relevant factors.
- ✓ Feature selection and engineering performed to improve the accuracy of the model.
- ✓ Selected suitable features for building a regression model and our trained model will predict the prices on our test data.
- ✓ The accuracy of the model will be validated by comparing predicted prices with actual prices, and once deemed accurate, the model will be deployed in a web application to predict house prices based on their features.

Github Project

Spam Classification using NLP : *(Aug 2022 – Sept 2022)*

- ✓ Developed a machine learning model for spam classification project using NLP techniques.
- ✓ Preprocessed the data by removing stop words, stemming the words, and converting the text into numerical features.
- ✓ Trained the model on a dataset of labeled emails using *Bag of Words* and *TF-IDF* techniques.
- ✓ Machine learning model was trained using various algorithms, including *Logistic Regression, SVM, Random Forest* etc.
- ✓ Evaluated the model using various performance metrics such as accuracy, precision, confusion matrix.
- ✓ Classified new emails as either spam or ham based on their content using the developed model.

Github Project

Cricket Data Analytics & Exploratory Data Analysis : *(Feb 2022 – May 2022)*

- ✓ Cricket Data Analytics project is made on IPL and International criket data till 2019.
- ✓ We can analyse the data of the matches and get some statistical insights to predict the performance of the players on that particular match.
- ✓ Performed Exploratory Data Analysis (EDA) to gain insights into patterns, relationships, and trends in data for data science projects.
- ✓ Leveraged EDA insights for developing predictive models or making informed decisions in business and research.
- ✓ Utilized EDA techniques to provide valuable insights for advancing the field of data science.

Github Project

Key results :

- ✓ Participated in Data Preprocessing Techniques in order to make data useful for creating Machine Learning models.
- ✓ Created pipelines which includes Data Collection, Feature Engineering, Feature Selection, Model Creation, Hyperparameter Tuning and Model Deployment.
- ✓ Built various regression and classification algorithms by using various Sklearn libraries such as Linear Regression, Logistic Regression, Decision Trees, SVM, Ensemble techniques.