VinodKumar

Data Scientist

- Data Scientist with **3+ years** of experience executing data-driven solutions to increase efficiency and utility of internal data processing.
- Skilled in designing, building, and deploying data analysis systems for large data sets; creating algorithms to extract information from large data sets; establishing efficient, automated processes for model development, validation, implementation, and large-scale data analysis
- knowledge of a variety of machine learning techniques (clustering, decision tree learning, artificial neural networks, etc.) and their real-world advantages/drawbacks
- Knowledge of advanced statistical techniques and concepts (regression, properties of distributions, statistical tests and proper usage, etc.) and experience with applications; creating and using advanced machine learning algorithms and statistics: regression, simulation, scenario analysis, modeling, clustering, decision trees, neural networks, etc.

+91 9949494228 vinod.art44@gmail.com Hyderabad

KEY SKILLS

• Artificial Intelligence • Machine Learning Algorithms • Tensorflow • Deep Learning • Neural Networks (CNN & RNN)
• Unsupervised Learning• Natural Language Processing (NLP)
• Data and Quantitative Analysis • Predictive Modelling and Analysis • Statistics - Hypothesis Testing, Normal/Gaussian Distribution, Z-test, T-Test, Chi-Square, Anova Test
• Data Mining • Data Wrangling • Data Visualization
•Linear Regression, Logistic Regression, Decision Tree, Random Forest, Ada-Boost, XG-Boost, Support Vector Machine (SVM), Naïve Bayes , KNN (K-Nearest neighbor) ,PCA (Principal Component Analysis), K-means, Time Series.

TECHNICAL SKILLS

- **Tools/Languages**: Python, Flask, Power BI
- **Database:** MySQL
- **Cloud**: Amazon AWS
- **Packages:** Scikit-Learn, Numpy, SciPy, Pandas, NLTK, Matplotlib, Seaborn, Statsmodels, Jupyter Notebook, Anaconda

PROFESSIONAL EXPERIENCE

Data Analyst
Nov '18 - Present

Lince soft Solutions Pvt Ltd
Hyderabad

Data Collection
- Collecting relevant data from various sources like databases, etc.

Data Analysis & Visualization
- Processed structured/unstructured data & gained insights by leveraging predictive modelling, statistics.
- Created various charts in **Jupyter Notebook** using **Matplotlib** to perform a preliminary analysis on the collected data

Data Wrangling & Feature Engineering
- Handling missing values and outliers in the variables
- Converting categorical into numerical since most algorithms need numerical features.
- Scaling the features if necessary.
- Cleaned, merged and manipulated data-sets and conducted **feature engineering using Pandas**
- Splitting data into training and testing sets

Machine Learning - Model Selection and Evaluation

- Applied various **machine learning algorithms** to analyze the data and develop a model to predict the future
- Used some **metric or combination of metrics** to "measure" objective performance of model
- Tested the model against previously unseen data

Hyperparameter Tuning
- Tuning model parameters for improved performance
- Simple model hyperparameters may include: number of training steps, learning rate, initialization values and distribution, etc.

KEY PROJECTS

PROJECT 1: Prediction Of Compressive Strength Of Concrete.

Brief: The goal is to predict the compressive strength of concrete based on given attributes .
- Created a **Machine Learning Regression model** to analyze datasets containing attributes like cement, fly ash, water, blast furnace slag, coarse aggregate, fine aggregate, etc.

PROJECT 2: Telecom Churn Prediction

Brief: Make predictions which customer will switch to other telecom service providers, based on the relevant customer data.
- Created a **Machine Learning Classification model** to analyze datasets containing attributes like voice mail plan, total call duration, call charge, customer service calls, etc.

PROJECT 3: Diabetes Prediction

Brief: Make predictions based on the patient's characteristic data set to predict whether s/he is diabetic or not
- Created a **Machine Learning model** to analyze datasets containing attributes like glucose level, blood pressure, age, etc.

PROJECT 4: Twitter Analytics

Brief: Deploy **social media analytics** to measure, analyze & interpret interactions & associations between people, topics & ideas
- Captured a dataset via Live Twitter Streaming for further analysis
- Performed **Sentiment Analysis** on the tweets obtained for visualizing the conclusion

PROJECT 5: Air Passengers Forecasting

Brief: Analyze the data and apply **time series modelling** to forecast the number of expected bookings for each month
- Analyzed the dataset containing monthly totals of international airline passengers between 1949 to 1960
- Rendered **assistance** to the **management** for making informed decisions w.r.t staffing, hospitality and ticket pricing

CERTIFICATIONS

Post Graduate Program in Artificial Intelligence and Machine Learning

Edureka

- **Course Modules:**
  - Introduction to Data Science | Data Extraction, Wrangling & Visualization
  - Introduction to ML with Python | Supervised Learning I & II
  - Dimensionality Reduction | Reinforcement Learning | Time Series Analysis
  - Deep Learning | Artificial Neural Networks |Image Classification with CNN
  - Object detection with RCNN |Natural Language Processing

EDUCATION

- Bachelor of Technology in Electronics and Communication Engineering | JNTU Hyderabad| June'14