

RATNA VAMSI

Tenali , Guntur, Andhra pradesh
8520853106 | krvasi96@gmail.com

SUMMARY

Having 2+ years of extensive experience Analytics,Data science with machine learning and pattern recognition End - End data analytics project management experience which constitutes requirement elicitation data preparation and data processing building statistical models. Comfortable working with several frameworks including Pandas, Numpy, ScikitLearn, Keras, TensorFlow and NLP, CV Passionate about programming, open-source technologies, deep learning algorithms, and continuous self-learning.

EXPERIENCE

- **BEPEC SOLUTIONS**

07/2019 - PRESENT

DATA SCIENTIST

I have 2+ experience of involvement in conveying end to end information science ventures utilizing Python.Working with Python libraries like Pandas, NumPy, etc.Creating predictive models for AI and ML-based features.

Capability in measurements, machine learning and making convincing representations.Worked in multiple parts for Retail, Banking and Insurance for better improvement.Data understanding, Data perposoning, Modeling, Evaluation and Deployment.Experience on Data wrangling, Data preparation and Data Visualization python and PowerBI .Significant expository and critical thinking aptitudes alongside capacity to comprehend current business forms and execute successful solution.Worked with different business units within sales and marketing to understand information needs and provide effective data management and analysis solutions to analyze business environment

EDUCATION

- **Dr.Bhimrao Ambedkar university**

2018

B.Sc computer Science

TECHNOLOGY

- Statistics & Probabilities
- Data processing :- Pandas, Numpy, Scikit Learn
- Machine /Deep Learning frameworks :- Scikit_Learn, Keras , Tensorflow
- Computer Vision, OpenCv, image preprocessing
- Natural language processing:- NLTK,SpaCy, Textblob,OCR
- Data visualisations:- Matlab , Seaborn ,PowerBI
- Database :- MS SQL Server
- Deploy :- AWS SageMaker
- Program Language:- Python

- **VEHICLE INSURANCE CROSS SELL PREDICTION**

Customers of age between 30 to 60 are more likely to buy insurance. Customers with Driving License have higher chance of buying Insurance. Customers with Vehicle_Damage are likely to buy insurance. The variable such as Age, Previously_insured, Annual_premium are more affecting the target variable. Comparing ROC curve we can see that Random Forest model performs better. Because curves closer to the top-left corner, it indicates a better performance.

Technology:- Pandas, Numpy, Scikit learn

- **SENTIMENTAL ANALYSIS**

This project performed sentimental analysis based on opinion words (like good, bad)

As we are doing sentiment analysis, it is important to tell our model what is positive sentiment and what is a negative sentiment. In our rating column, we have ratings from 1 to 5. We can define 1 and 2 as bad reviews and 4 and 5 as good reviews. We need to define the input variable X and the output variable Y. X should be the 'all_features' column and y should be our 'sentiment' column. train_test_split automatically splits the data in 75/25 proportion. 75% for the training and 25% for the testing. CountVectorizer develops a vector of all the words in the string. Import CountVectorizer and fit both our training, testing data into it. We will use the Logistic Regression as this is a binary classification.

we used 1 for good reviews and 0 for a bad review

Technology:- NLTK, Spacy, scikit learn

- **LOAN AUTOMATIC PREDICTION**

We will choose decision tree based model.

We choose random forest. Because some of the features have null value, I will use either drop the records with null values or fill a value instead. I eventually choose to fill the missing feature value which appears the most often in each feature because of few training data. Predict if a loan will get approved or not. Prediction of loan prediction in USA based on various features.

TECHNOLOGIES:- Pandas, Numpy, Scikit learn