

ALEKHYA PUPPALA

apupp002@odu.edu

Contact :- +1 (757)777-1041

SKILLS:

Statistics : Hypothesis Testing , Probability Theory (Gaussian/ Bernoulli Distributions), Chi-Square, ANOVA

Python : Pandas, Numpy, Sklearn, Matplotlib, Seaborn, Spacy , PyTorch

Machine learning : linear models, decision tree, Ensemble models (Random forest, Ada boost, Gradient boost, XG Boost) , SVM, Naïve Bayes K Means Clustering , PCA .

Deep Learning : Convolution Neural Networks, Recurrent Neural Networks, Long Short Term Memory.

PROJECTS:

Problem Statement :- For a car manufacturing company , build a machine learning model to predict the time it takes to test a car given its several configurations such as 4WheelDrive, Added air suspension or a head-up display etc.

- 1.Performed feature selection using Chi-Square and ANOVA correlation techniques.
- 2.One-hot encoded the categorical inputs and scaled the continuous inputs.
- 3.To tackle the curse of dimensionality and multicollinearity , used Principal Component Analysis technique.
4. Trained the Random forest , Gradient Boost Regressor and performed hyper-parameter tuning using GridSearch CV.
- 5.Evaluated the performance of models using R-squared score and mean squared error.

Problem Statement :- For an e-commerce company, build a NLP model to classify the sentiment of a product review.

- 1.Tokenised each review using RegExpTokenizer.

- 2.Performed the Stop Word Removal step to filter the pure tokens/words.
- 3.Lemmatized the pure tokens using WordNetLemmatizer.
- 4.Built the Document Term matrix using Tf-idf vectorizer.
- 5.Trained Naïve Bayes Classifier
6. Evaluate the model using F1 score

EDUCATION :-

Bachelor of Technology – Osmania University (2016-2020)

EXPERIENCE :-

Worked as an intern at Blackcoffer.(August 2020 – December 2020)

Served as Trainer for data science at Sure Trust. (January 2020 – February 2021)

Worked as junior data scientist at Harsco Pvt Ltd. (January 2021- May 2022)