Predictive Modeling for Digital Marketing Campaign

Introduction:

In this report, I present the analysis and visualization of results for the given dataset. The dataset is focused on the digital marketing campaign using LinkedIn posts and their components. I conducted EDA on the dataset, pre-processed and cleaned the dataset as well as apply predictive modeling strategies on the processed data.

1. EDA and Pre-processing of dataset:

- In the first step I used the pandas dataframe for analysis of data. Analysis includes trends of performance metrics over time, relation between attributes and performance metrics as well as effective distributions wrt the attributes.
- For the pre-processing I used different techniques and step-by-step transforming the dataset.
- The missing values were handled by imputing mean in the case of numerical columns and mode in the case of categorical columns.
- The handling of outliers was done using IQR method and this data is then encoded using Label Encoder and Standard Scaler.
- Label Encoder is used for encoding categorical values and standard scaler for numerical values.
- The final processed dataset is then split using scikit-learn train-test split.

2. Predictive Modeling:

- The data is then used for prediction based on four-five techniques.
- The models or techniques used are - Linear Regression, Lasso Regression, Ridge Regression, Random Forest Regression and a Neural Network model.

- I also used GridSearch for hyperparameter tuning of the models and selection of best parameters.
- The models are evaluated on the metric of Mean Squared error and R2 score.
- On the basis of the experiments the best model for prediction turned out to be Random Forest followed closely by the Neural Network model.

3. Insight and Recommendations:

- Based on the results of the EDA and the predictive modeling some insights were made.
- Four attributes- dominant color, topic, type, number of faces play an important role in the predictive abilities of the models.
- It is recommended that during campaigning use the dominant colors as well as catchy topic and relevant material in the post.

4. Conclusion:

Following the step-by-step pipeline and methodology involves the use of many different techniques such as - EDA, pre-processing, outlier handling, visualization methods as well as machine learning models for predictive modeling. In this project, I have also demonstrated the relation between attributes and the performance metrics of a digital marketing campaign, Using the insights and methodology of this project can lead to better marketing strategies and paradigms.