ML_Haloween_price

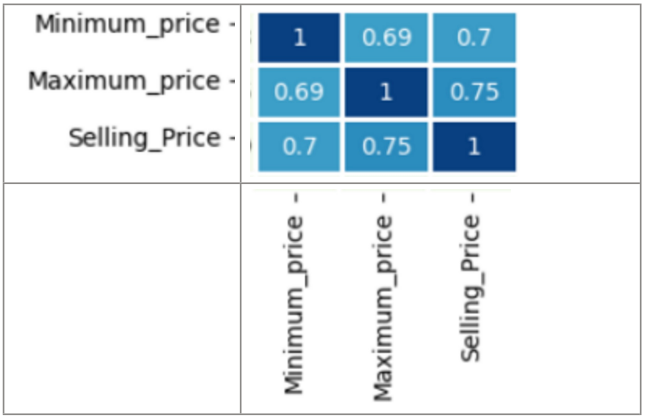
13 October 2024 17:21

Model: https://www.kaggle.com/code/shringivyas/halloweenstallproducts

Task: The goal was to predict the selling price of a product.

1. Data Preprocessing:
   - Missing values were handled by filling them with the mean.
   - One-hot encoding was applied to categorical variables (grade, loyalty_customer, Product_category).
   - Three columns were dropped, for them to have to relevance with selling price calculation (product_id, stall_no, customer_name, instock_date)

2. Exploratory Data Analysis:
   - A correlation matrix/heatmap was plotted to identify the relevant features for the model.



3. Modeling:
   - Data was transformed using `fit_transform`.
   - Polynomial regression and linear regression models were applied, with degree = 4.

4. Testing:
   - The same transformations and modeling were applied to the test file.

5. Performance:
   - The model achieved an R-squared score of 95%, indicating a high level of accuracy in predicting the selling price.

This process reflects strong data cleaning, feature engineering, and modeling steps, leading to excellent predictive performance.