Hide menu

Case Study

Reading: Regression Analysis Case Study - Demo
2h

Reading: Regression Analysis Case Study
2h

Discussion Prompt: Regression Analysis Exploration Exercise
2h

Graded Assignment: Self Reflection
1h

Self Reflection

Review Learning Objectives

Assignment details

Due
Mar 24, 11:59 PM IST

Attempts
Unlimited

Your grade
You haven't submitted this yet. We keep your highest score.

--

Like Dislike Report an issue

Instructions

Resume

1. Reflecting on the case study, what was the most challenging aspect of applying regression analysis to solve the real-world problem? How did you overcome this challenge, and what did you learn from it? **1 point**

Similar to the classification case study, here also preprocessing is a major challenge for regression analysis too. For example, we need to get rid of the features not carrying much information (e.g., car_ID, CarName etc.) and then for some regression models (e.g., SVR) we need to convert the categorical features to OHE (one-hot-encoding), prior to training the model.

Your answer cannot be more than 10000 characters.

2. Describe a situation where you encountered technical problems while applying regression analysis to the case study data. How did you troubleshoot and resolve these issues to ensure accurate results? **1 point**

Application of z-score normalization as a preprocessing step turns out to be very crucial for this dataset. For example, with hyperparameter tuning using GridSearchCV(), both the SVR and KernelRidgeRegession models had a negative R2 score when trained with un-normalized data, but as soon as StandardScaler() preprocessing was applied, the R2 became positive.

Your answer cannot be more than 10000 characters.

3. Reflect on the interpretation of the regression model results. How did you derive actionable insights from the model outcomes to make data-driven decisions for the real-world problem? **1 point**

By running OLS from scipy.statsmodel we can see that the features enginesize, boreratio, stroke, compressionratio are statistically significant and they have large coefficient magnitude, hence they are important predictor variables for the response variable price. We can also see that the simple OLS model has over 90% R2 and adjusted-R2 values. When we use the ensemble model RandomForestRegressor with 80-20 validation, we get 91% score on the held-out test dataset, we again observe that the predictors enginesize, boreratio and highwaympg have high feature importance values.

Your answer cannot be more than 10000 characters.

4. How did the application of regression analysis in the case study scenario enhance your critical thinking and problem-solving skills? Provide specific examples of how regression analysis aided you in making informed decisions [Practice this question as if you were in an interview!] **1 point**

Understanding the data and trying different preprocessing techniques to prepare the data ready for training machine learning models had been a key indispensable step needed at the very outset for the regression analysis. Starting from the simplest regression model to more complex ones to improve the performance and generalizability of the models was another important lesson learnt from this analysis. At every step when a model performs poorly, it needs a closer look to the data and the model and think for the reason (e.g., hyperparameter tuning, regularization etc.)

Your answer cannot be more than 10000 characters.

5. What were the most valuable lessons you learned from completing the case study? How do you plan to apply these insights to further develop your regression analysis skills and grow as a data analyst? **1 point**

Preprocessing and feature transformation (as per the requirement) turn out to be the most essential step prior to building machine learning models. Then every time we build a new model, we need to compare the performance of the model with the existing ones in terms of its generalizability (on the held-out dataset) and also the speed of training / prediction.

Your answer cannot be more than 10000 characters.

Coursera Honor Code Learn more

I, **SANDIPAN DEY**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Submit **Save draft**

Last saved on Feb 23, 2:02 AM IST

Like Dislike Report an issue