REGRESSION METRICS

REGRESSION

▶ Regression is one of the supervised learning techniques that is used for the prediction of continuous data. The regression technique is used to model relationships between dependent and independent variables and produce a line of best fit.

▶ There are different types of regression models. Linear regression, for example, fits a straight line to the data. This poster focuses on the metrics for linear regression.

WHAT ARE PERFORMANCE METRICS

▶ Performance metrics are measures used to evaluate the performance or effectiveness of a system, process, model, or any other entity. Here, we are going to focus on metrics that can be effectively used as loss functions.

MEAN ABSOLUTE ERROR

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Where: - n is the number of samples.
- y_i is the actual value of the i-th sample.
- $\hat{y}_i$ is the predicted value of the i-th sample.

ROOT MEAN SQUARE ERROR

$$\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}$$

Where: - n is the number of values in the dataset. - x_i represents each individual value in the dataset. The RMS is calculated by taking the square root of the average of the squared values in the dataset.

R2 SQUARED

The formula for R-squared is:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

Where: - SS_{res} is the sum of squared residuals - SS_{tot} is the total sum of squares.

▶ Ranges from 0 to 1: 0 indicates model explains none of variability. 1 indicates model explains all variability.

ADJUSTED R2 SQUARED

$$\text{Adjusted } R^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - p - 1}$$

Where: - R^2 is the R-squared value of the model. - n is the number of observations. - p is the number of predictors (independent variables) in the model.
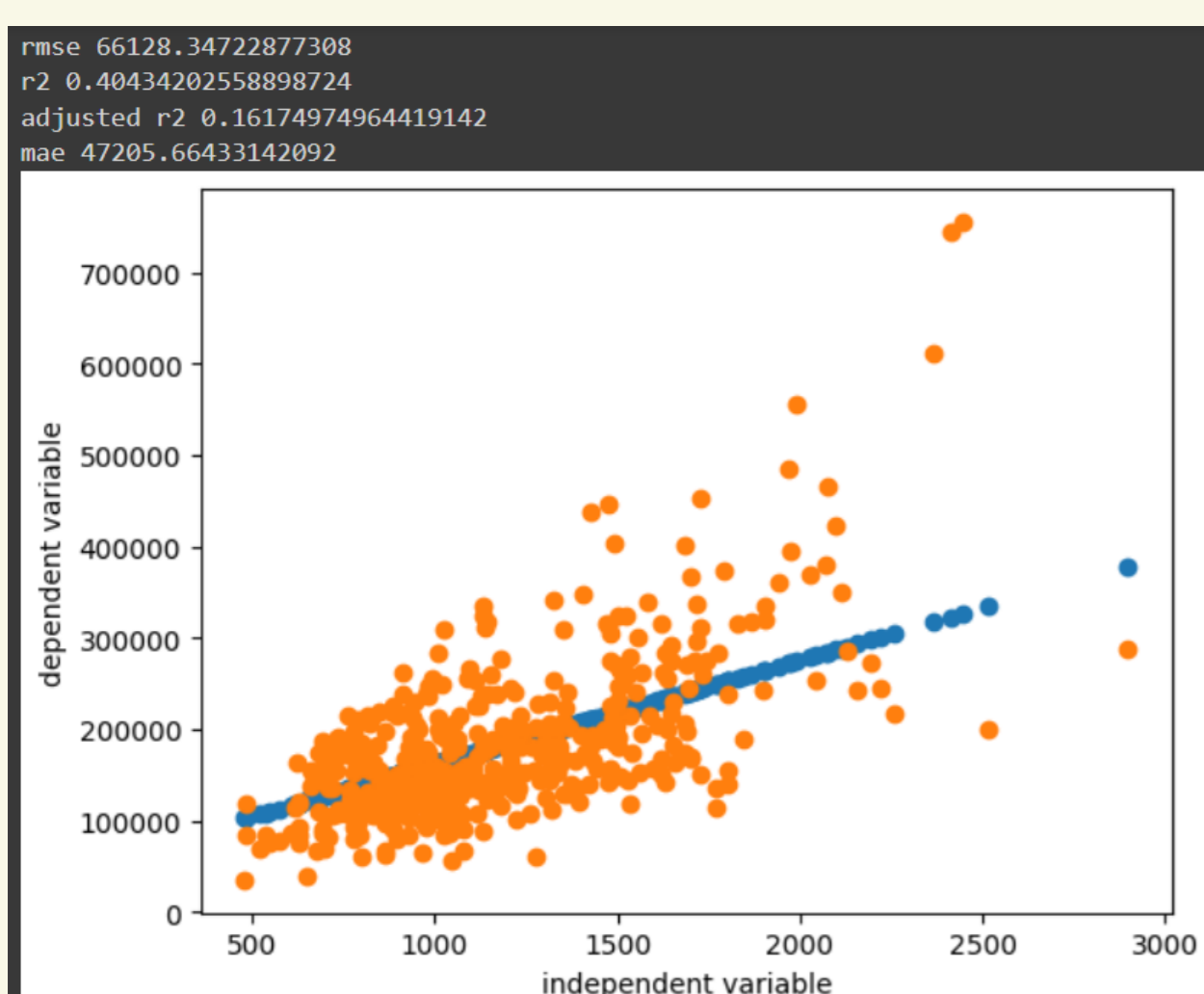
APPLICATION

▶ MAE: useful in housing price prediction, since it gives a straightforward difference in magnitude

▶ RMSE: in weather forecasts since it penalises outliers severely

▶ R2-score is used in fitting linear functions since it can tell us how good the data itself is, in ML, it can be used, for example, for stock price prediction

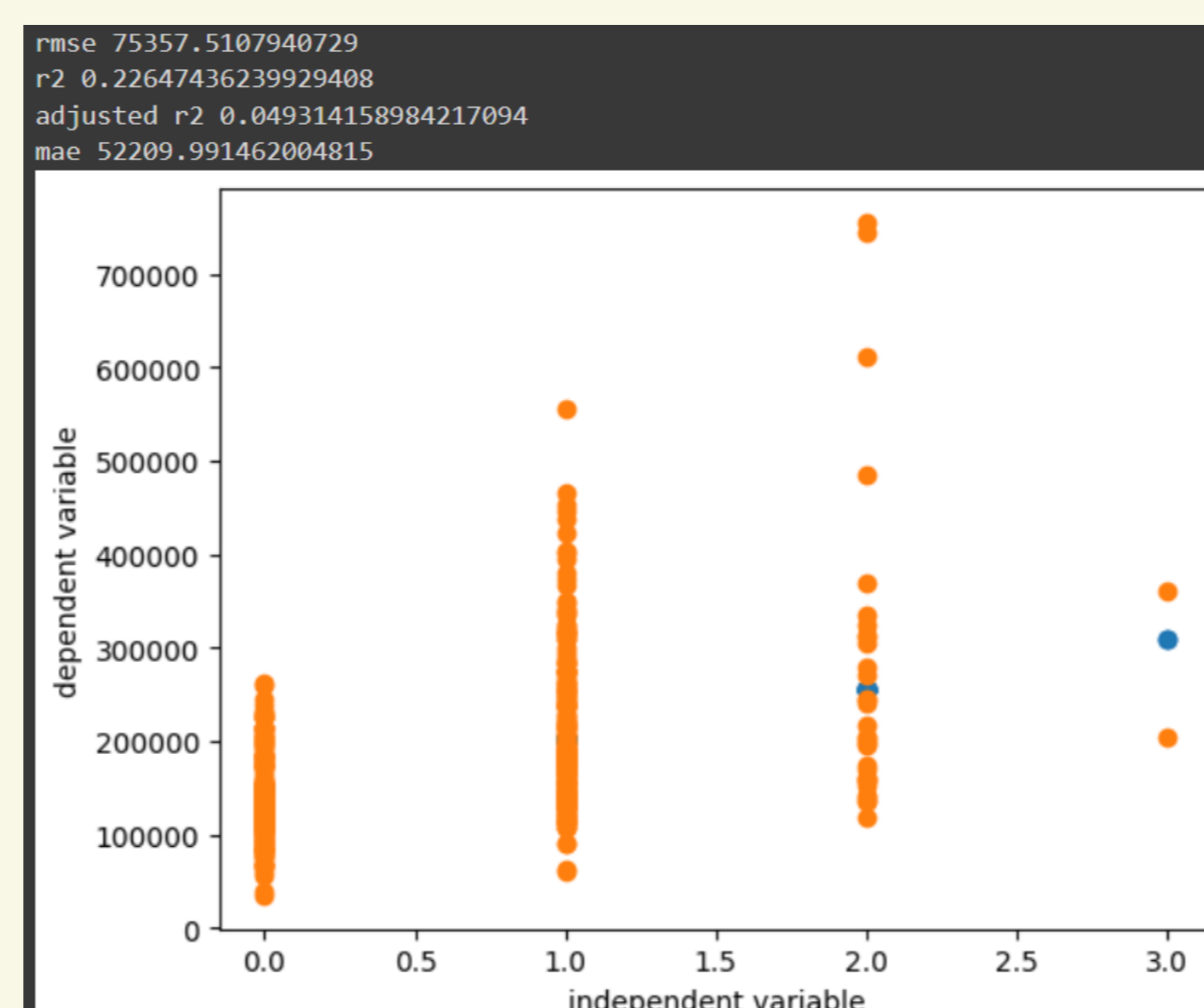▶ Adjusted R2-score penalises addition of unnecessary predictors

GOOGLE COLAB LINK

Ipsita Rout.
Google colab notebook.
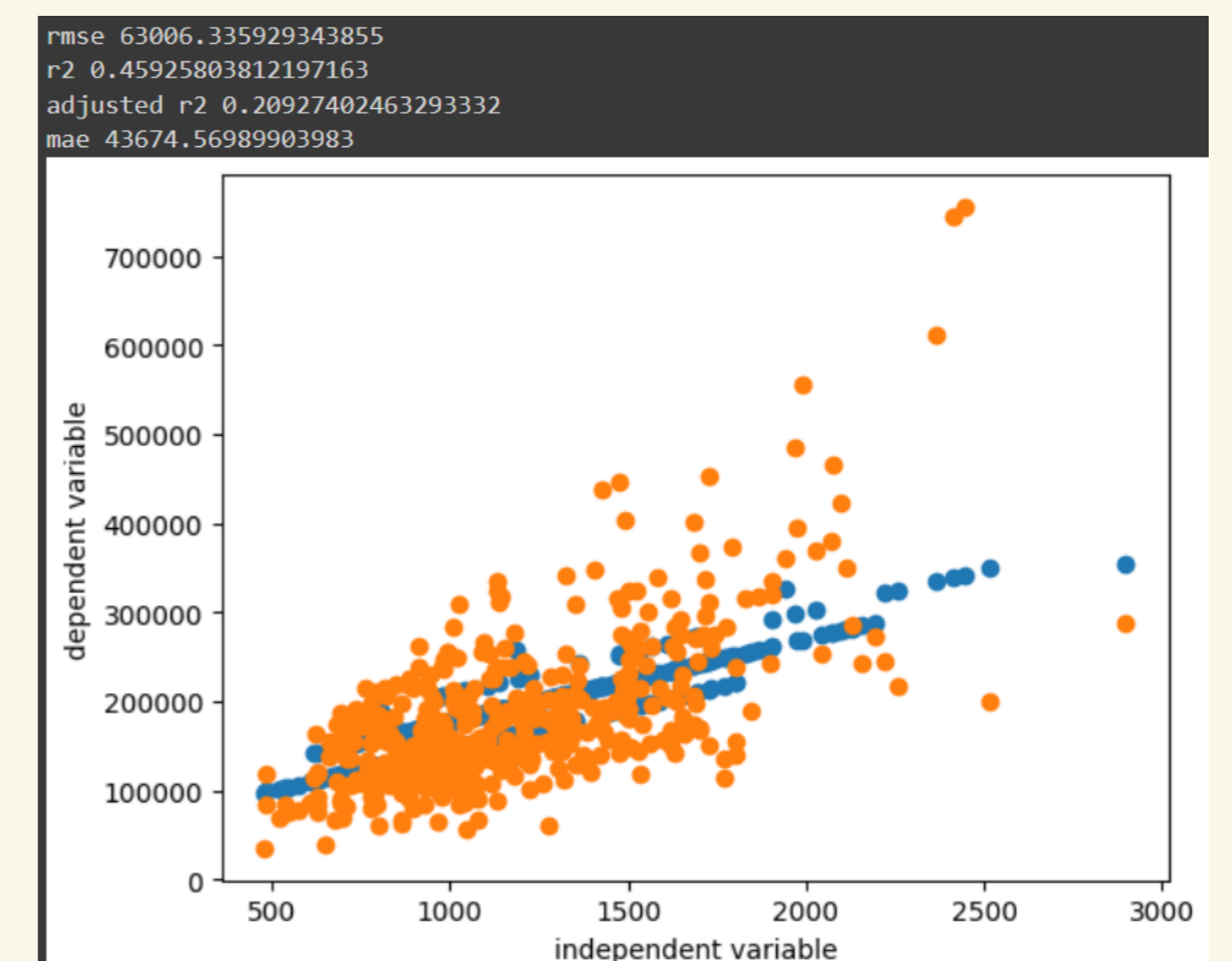Online, 2024.
Accessed on 1st April 2024.

AN EXAMPLE

(a) price vs first floor sq. ft.

(b) price vs no. of fireplaces

(c) price vs both features

Ipsita Rout

CS460 Poster Presentation

NATIONAL INSTITUTE OF SCIENCE EDUCATION AND RESEARCH