Inverse Linear Regression in Machine Learning

Subhradeep Biswas

Received: 8 December 2016 Accepted: 2 January 2017 Published: 15 January 2017

Abstract

A linear regression machine learning model derives the linear relationship between single or multiple features (put in the x-axis of a co-ordinate plane) and a single response (put in the y-axis of the co-ordinate plane) for a given set of observations. The model then learns to predict the response for a set of new feature values using the derived relationship. However, the linear regression model does not have the flexibility to predict the feature values for a target response. The solution proposed in this paper can leverage the relationship derived by the linear regression model between multiple features and single response. Using the relationship, it can predict the feature values for a target response value. In the proposed solution, the model accepts the training data in two separate input datasets ? one contains the features in observations and the other contains the responses. After making the prediction on feature values for a queried response value, the model returns a two dimensional array of numbers. Each column of the output array contains the predicted values for a specific feature. Each row of the array contains different valid sets of feature values. Each set of feature values results the queried response value according to their linear relationship.

Index terms— inverse linear regression, reversed regression, machine learning.

The first column "TV" shows the amount of money in thousands of dollars spent on TV ads to advertise a single product. For example, in observation 200 (row number), \$232,100 was spent on TV ads. Similarly, \$8600 was spent on radio ads in the same observation. The "Sales" column represent the sales of the product being advertised in that observation in thousands of items. So, in observation 200, a quantity of 13,400 was sold. In this case, linear regression model can be useful to predict the sales based on amount of dollars spent on advertisement on different channel. The model uses "TV", "Radio" and "Newspaper" as the features and it predicts the sales as response. The model learns the linear relationship between the features and response in order to make the predictions. However, this model does not predict the advertisement expenses required to be put in different channels in order to reach a specific sales target.

The Inverse Linear Regression algorithm can fulfil such requirements by leveraging the relationship learnt by the linear regression model. The outcome of the model also suggests the relative weight of each feature and how well each feature contributes in order to reach the target response.

1 II. Proposed Algorithm a) Deriving Relationship

The relationship between features and response is learnt with the help of linear regression model. The linear regression model derives the weights of each feature (I.e. "TV", "Radio" and "newspaper") to calculate the response (I.e. "sales"). The formula representation of the relationship is given below: y (sales) $= w_0 + w_1 .x_1$ (TV ads) $+ w_2 .x_2$ (Radio ads) $+ w_3 .x_3$ (Newspaper ads)

Once the linear regression model is trained with the features and response data, the model returns the intercept (w_0) and the coefficients (w_1, w_2 and w_3).

2 b) Locating Nearest Features

The proposed solution locates the nearest observation that has the response value less than the queried response for which the feature values are to be predicted. For each queried response, the immediate lower 'value in response' (r) that was used while training the linear regression model is located. Using the index of the located training response(r), the corresponding feature values are obtained from the observations. 17 Calculating 2 nd set of features: The algorithm repeats the steps described above to calculates x 2 (radio ads) for x 1 and x 2 from nearest observation and y = queried sales value. The 2 nd set of predicted feature values [x 1 , Predicted_x 2 , x 3].

Calculating 3 rd set of features: The same steps are repeated to predict x 3 . The 3 rd set of predicted feature values [x 1 , x 2 , Predicted_x 3].

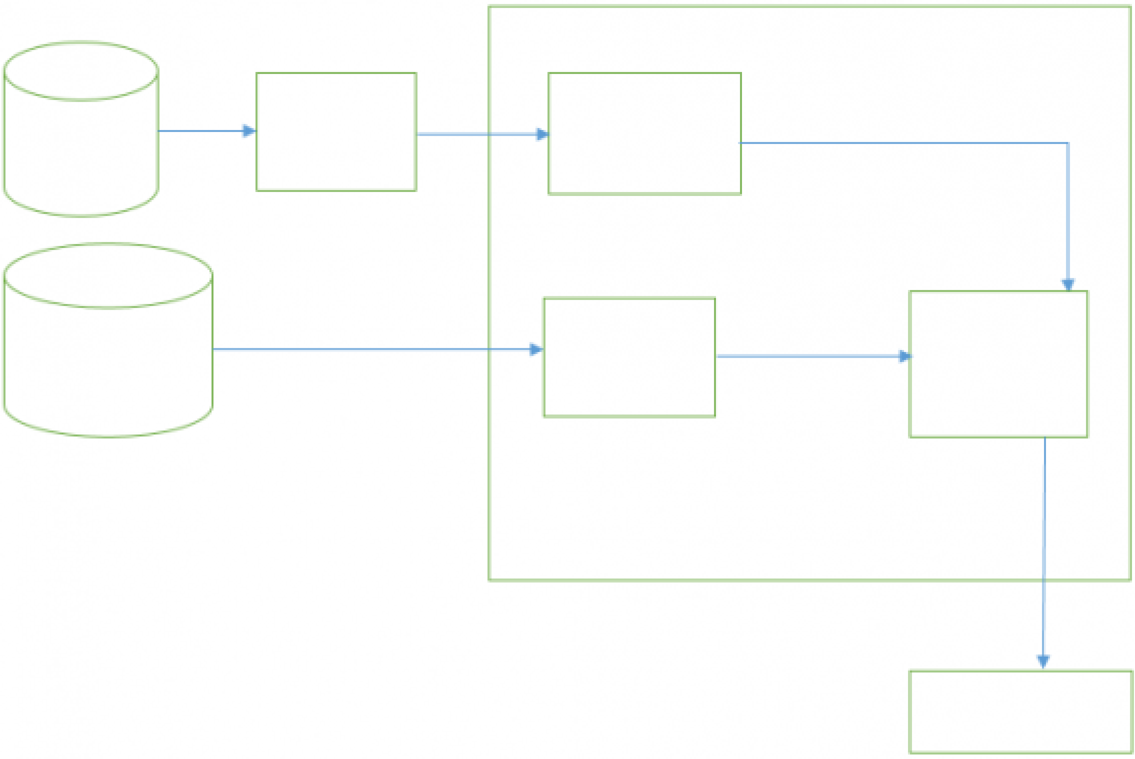
3 d) Output

The solution returns a two dimensional array. Columns of the array represent features. Each row of the array is a set of predicted feature values. The no. of columns in array= the no. of rows in array= the no. of features in the observation.

4 III.

5 Conclusion

The prediction accuracy by the proposed algorithm is as good as the accuracy of the linear regression model as the relationship determined by the linear regression model is leveraged by the new solution. Different sets of feature values provide the information about the effectiveness of each feature i.e. if the coefficient of a feature is relatively small, a larger amount is to be spent on that channel in order to get the target response. Thus, it provides the flexibility to the stakeholders to choose the appropriate mechanism to achieve the target.



1

Figure 1: Fig. 1 :

[Tibshirani] , Robert Tibshirani .

[Hastie ()] *Introduction to Statistical Learning*, Trevor Hastie . 2013.