Real Life Use Case of Regression

AGENDA OF THE VIDEO

Hello Guys, In this video, what we are going to be doing is continuing with one of our last discussion on “Regression Models in Tensorflow“, we are going to be taking up a real world problem and will try to solve that using Regression Methods in tensorflow. In particular, we will be doing:

- Acquiring the data.
- Exploring the data.
- Scaling the data.
- Making Features.
- Making the Sequential Model architecture for Regression.
- Training the model.
- Obtaining the Predictions from the model.
- Measuring the performance of the model.
- Some other real world problems of regression.

Acquiring the data

To acquire the data, we will be using the FREE yahoo API. This API gives us access to all of the stock data for all the companies registered in stock exchange. Particularly, it is going to return OpenTime, CloseTime, OpenPrice, ClosePrice, Volume traded.

To get this data, we will be using *pandas_datareader*.

Exploring the data

Now, that we have a data for a company for its OpenTime, CloseTime, OpenPrice, ClosePrice, Volume traded, we can have a quick look at this data and see how does the closing price changed throughout the time.

Let's do this really quick. We can use matplotlib for exploring the data visually.

Scaling the data

Scaling is very important in ML and it becomes even more important in neural networks. The reason that it becomes really important is because in neural networks, the number of weights that it learns is huge and sometimes because of unscaled data, it might happen that the neural network might get distracted because of these points and it might stray away from the optimum point.

Hence, it is always advisable that in neural networks, the data points needs to be scaled before it should be fed into the model.

We will be using MINMAXSCALER from sklearn to accomplish this task.

Making Features

How are we going to predict the closing price?

To do that, we will use Regression.

But, Regression needs some features. We cannot use dates in the feature set and also using opening price doesn't make any sense.

Since, the closing price is a time series data, we can predict today's price based on some past prices. How many?

Let say 60.

So, we will have a data which will have 60 features and one variable to predict.

Making Splits of training and testing

To check the performance of the model, we will use a seperate hold out dataset for testing.

We are going to be using 20% of the total samples for testing.

Making the Sequential Model Architecture

Building the Tensorflow model architecture.

```
Model = Sequential()

model.add( Dense() )
model.add( Dense() )
model.add( Dense() )
```

Specify, the necessary arguments to these functions and then compile the model by calling the function .compile of the model.

Training the model

Let's train the model by passing training dataset and calling the .train function of the model

Obtaining the predictions

Let us obtain the prediction on the hold out data.

But remember that these predictions will be scaled predictions and we need to use inverse transform of the MINMAXSCALER to get the right predictions.

Measuring the performance of the model

Let us now plot the training, testing data points with the predictions and let us see how close could we get.

From this step, you can go back and forth between changing the model configurations and checking the model performance to kind of optimize the mode.

Some other real world tasks in Regression

- House Price Prediction
- Predicting Electricity consumption
- Forecasting Sales
- Predicting Profit.
- Prediction of Energy Consumption.
- and many more...

Thank you