Module 2: Critical Thinking

Predicting Future Sales

Nolan Byrnes

CSC580 – Applying Machine Learning and Neural Networks - Capstone

Colorado State University – Global Campus

Dr. Isaac Gang

September 22nd, 2023

Module 2: Critical Thinking

With machine learning, we can predict the future sales that a product could have based off historical data. To predict the future video game sales, we can use a supervised learning approach, since the dataset contains input variables, as well as the output variable that we would like our model to predict after training. With the model having to predict the total revenue of a video game, the problem is a regression problem, which is “when the output variable is a real or continuous value, such as ‘salary’ or ‘weight’” (Shukla, 2023, para. 4). This paper goes over how I preprocessed the data, trained the model using supervised learning, and the analysis of the results of predicting the video game sales based on historical data.

When preprocessing the Training Data, we scaled the values to be in the range of 0 and 1. Brownlee warns that “unscaled input variables can result in a slow or unstable learning process” and “unscaled target variables on regression problems can result in exploding gradients causing the learning process to fail” (Brownlee, 2019, para. 3). Without scaling the input and target variables, the model could end up with large weight values, which could be unstable. To scale the train and test data sets, Sklearn’s MinMaxScaler was used, which transforms “features by scaling each feature to a given range” (Sklearn, n.d., para. 1).

After scaling the data, it is important to keep track of how the values were scaled. This is because when your model makes predictions, it would be returning the scaled values, which would need to be translated to its actual value to help us make sense of the predictions. For example, in figure 1 below, you can see that we scaled the total earnings values by multiplying by 0.0000036968, and then adding -0.116913.

Figure 1.

Scaling Total Earnings Values

A black and white text on a black background

Description automatically generated

Note. This figure displays how the total\_earnings values were scaled to be within the range of 0-1.

When we get predicted results from the model, we would perform the inverse to translate the predicted scaled value to the unscaled value we are looking for.

The Mean Squared Error (MSE) was used as the loss function for when the model was trained. The reason why we would want to use Mean Squared Error is because it “measures how close a regression line is to a set of data points” (Gupta, 2023, para. 2). From the first epoch, it started with a MSE of .0624, and went down to 0.000073844 at the final epoch on the training dataset, as displayed in figures 2 and 3 below. The trend of the MSE values decreasing as training progressed tells us that the model’s predictions were becoming more accurate to the training dataset between each epoch.

Figure 2.

Head of Epochs

A screenshot of a computer screen

Description automatically generated

Note. This figure displays the results of the first 10 epochs out of the total 50 that the model was trained on.

Figure 3.

Tail of Epochs

A screenshot of a computer

Description automatically generated

Note. This figure displays the results of the final 10 epochs out of the total 50 that the model was trained on.

After training the model, I tested the neural network against the test data set, and got a Mean Squared Error (MSE) of .00018 as shown in figure 4 below.

Figure 4.

MSE of Test Data Set

A black background with white text

Description automatically generated

Note. This figure displays the resulting MSE when using the test data set with the trained model.

A small MSE tells us that the data points are close to the regression line, therefore, the predictions that the model gave on the test data set were great predictions because the points on the regression line were close to the actuals. This also tells us that there were likely no outliers in the test data set, as having outliers within could skew the regression line, resulting in a higher MSE.

For a portable video game that is exclusive to the USA, a sequel, and suitable to kids that had a critic rating of .7, the trained model predicted the earnings to be a total of $228,854.90 as shown in figure 5 below. With the MSE of the test data set being .00018 on the test data set, the regression line was still close to the actual values, just as it was with the training data set.

Figure 5.

Prediction of Total Earnings

A black screen with white text

Description automatically generated

Note. This figure displays the total earnings prediction for a portable action video game that is exclusive to the USA, a sequel, suitable for kids, with a critics rating of 0.7, and a scaled unit price of 0.8.

**Conclusion**

Predicting the future video game total sales is a regression problem, since the output variable, total sales, is one that contains a continuous value. Using the Mean Squared Error is a better metric to use in this situation since helps to measure how close the regression line is to the actual data points. Scaling the data points in our data set is useful, as it can help the machine learning model become more stable and converge sooner. It is important to take note of how you scaled the values, so that they can be descaled back to the original values, so it makes more sense to us humans. The MSE that we received on the test data set was larger than with the training data set, however, it still had very low MSE, which told us that the points on the regression line were still close to the actuals.

**REFERENCES**

Brownlee, J. (2020, August 25). How to use data scaling improve deep learning model stability and performance. MachineLearningMastery.com. <https://machinelearningmastery.com/how-to-improve-neural-network-stability-and-modeling-performance-with-data-scaling/>

Gupta, A. (2021, December 15). Mean squared error : Overview, examples, concepts and more. Simplilearn.com. <https://www.simplilearn.com/tutorials/statistics-tutorial/mean-squared-error>

Shukla, S. (2023, January 10). Regression and classification | Supervised machine learning. GeeksforGeeks. [https://www.geeksforgeeks.org/regression-classification-supervised-machine-learning/#](https://www.geeksforgeeks.org/regression-classification-supervised-machine-learning/)

Sklearn. (n.d.). Sklearn.preprocessing.MinMaxScaler. scikit-learn. Retrieved September 13, 2023, from <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>