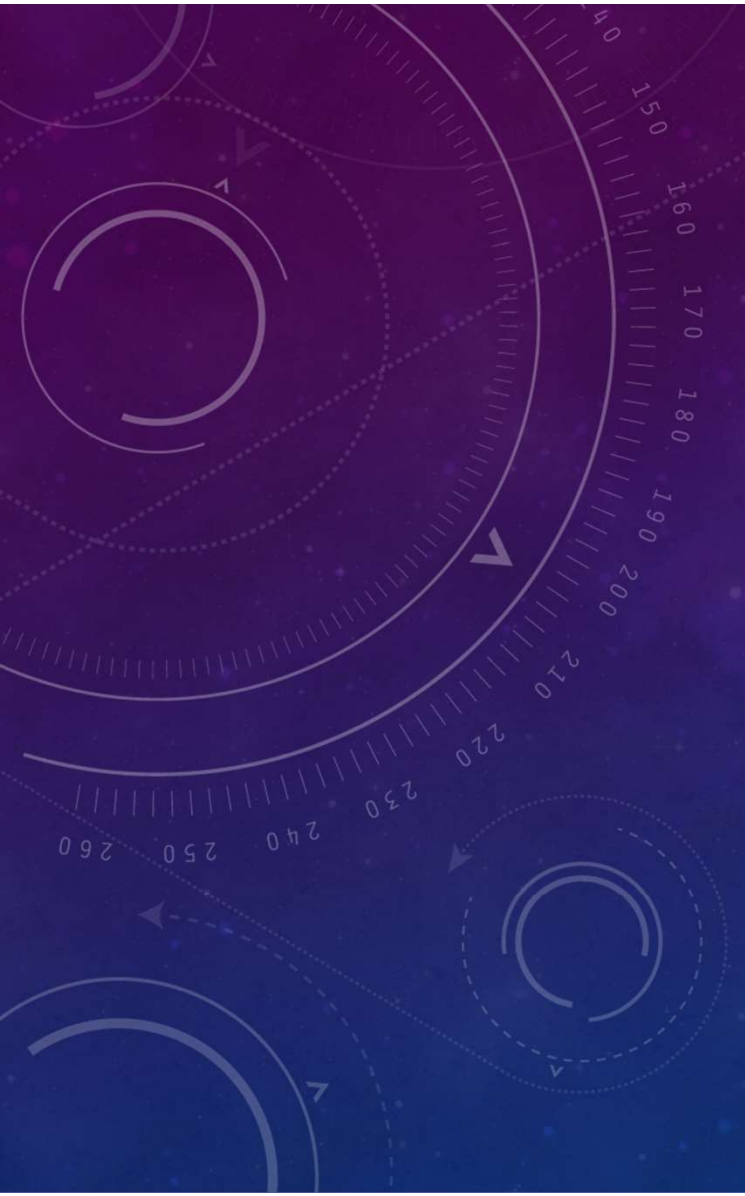


MACHINE LEARNING WITH .NET

Asmaliza Ahzan @ Emma

asmaliza@iverson.com.my

MATRIX FACTORIZATION MODEL

CHAPTER 7

MATRIX FACTORIZATION

- An unsupervised learning algorithm.
- A recommendation task enables producing a list of recommended products or services.
- ML.NET uses Matrix factorization (MF), a collaborative filtering algorithm for recommendations when you have historical product rating data in your catalog.
- For example, you have historical movie rating data for your users and want to recommend other movies they are likely to watch next.

USE CASES FOR MATRIX FACTORIZATIONS

- Music recommendations
- Product recommendations
- Movie recommendations
- Book recommendations

DIVING INTO THE MATRIX FACTORIZATION TRAINER

- The matrix factorization trainer is the only traditional trainer found in ML.NET currently.
- The matrix factorization trainer requires both normalization of the values and caching.
- To utilize matrix factorization in ML.NET, the Microsoft.ML.Recommender NuGet package is required if you are creating the project from scratch.

CREATING A MATRIX FACTORIZATION APPLICATION

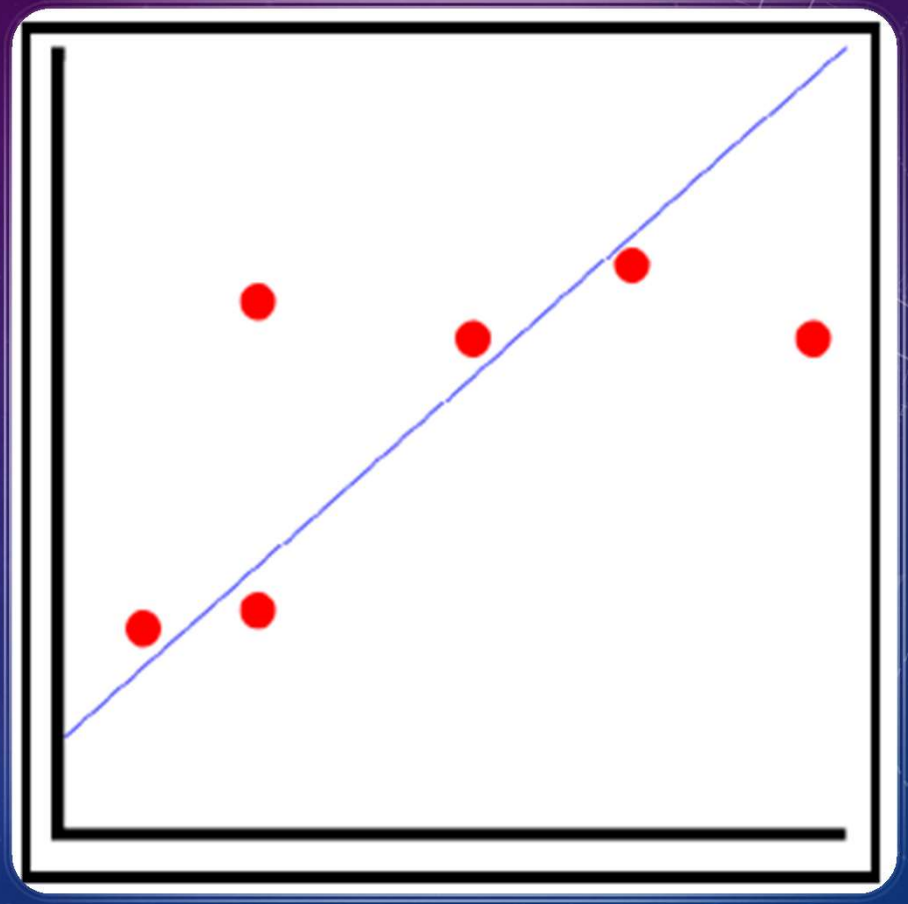
- The application we will be creating is for music prediction.
- Given a UserID, MusicID, and a rating, the algorithm will use that data to create recommendations.

LOSS FUNCTION

- The loss function is the function that computes the distance between the current output of the algorithm and the expected output.
- It's a method to evaluate how your algorithm models the data.
- It can be categorized into two groups. One for classification (discrete values, 0,1,2...) and the other for regression (continuous values).

MEAN SQUARED ERROR (MSE)

- Loss functions for regression.
- MSE is defined as the measure of the average of the squares of the errors.
- To put this simply, take the plot shown in the following screenshot:



MEAN SQUARED ERROR (MSE)

- The dots correlate to data points for our model, while the blue line is the prediction line.
- The distance between the red dots and the prediction line is the error.
- For MSE, the value is calculated based on these points and their distances to the line.
- From that value, the mean is calculated.
- For MSE, the smaller the value, the better the fitting, and the more accurate the predictions you will have with your model.
- MSE is best used to evaluate models when outliers are critical to the prediction output.

MEAN ABSOLUTE ERROR (MAE)

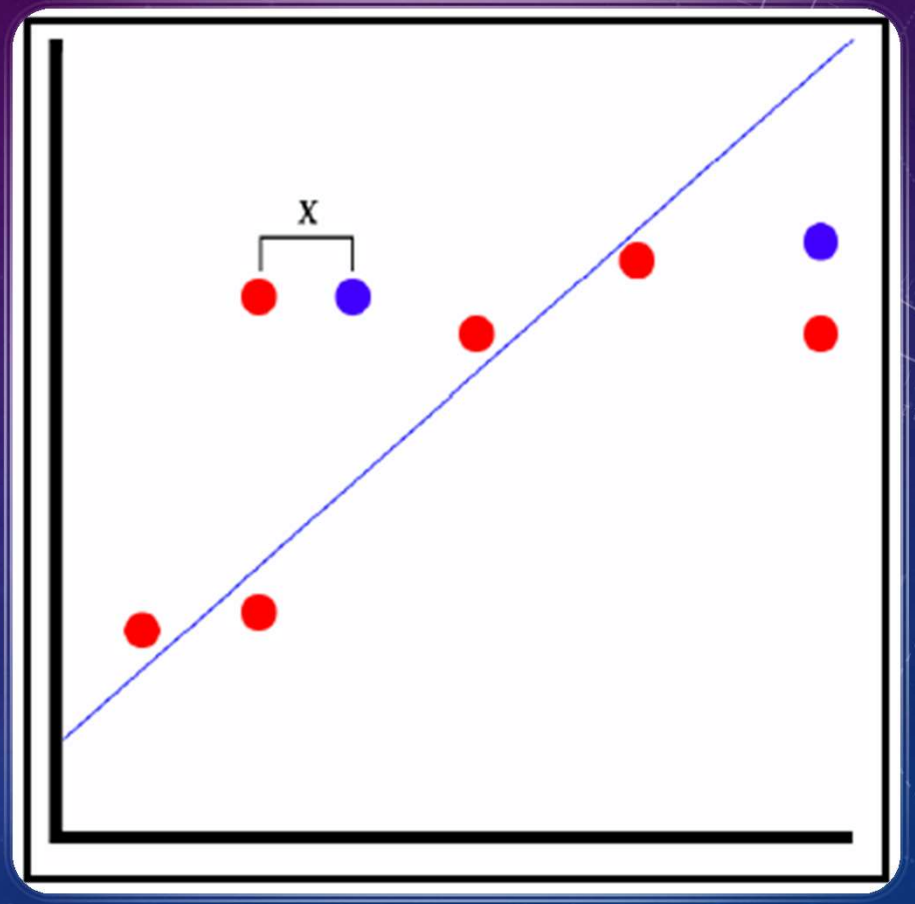
- Loss functions for regression.
- MAE is similar to MSE, with the critical difference being that it sums the distances between the points and the prediction lines, as opposed to computing the mean.
- It should be noted that MAE does not take into account directions in calculating the sum.
- For instance, if you had two data points equal distance from the line, one above and the other below, in effect this would be balanced out with a positive and negative value.
- In machine learning, this is referred to as Mean Bias Error (MBE).
- MAE is best used to evaluate models when outliers are considered simply anomalies, and shouldn't be counted in evaluating a model's performance.

R-SQUARED

- R-squared, also called the coefficient of determination, is another method of representing how well the prediction compares to the test set.
- R-squared is calculated by taking the difference between each predicted value and its corresponding actual value, squaring that difference, then summing the squares for each pair of points.
- R-squared values generally range between 0 and 1, represented as a floating-point value.
- A negative value can occur when the fitted model is evaluated to be worse than an average fit.
- However, a low number does not always reflect that the model is bad.
- Conversely, higher values aren't necessarily a sure sign of the model's performance, as this could be considered as overfitting of the model.

RMSE

- RMSE is arguably the easiest property to understand, given the previous methods.
- Take the plot shown in the following screenshot:



RMSE

- In the case of testing the model, as we did previously with the holdout set, the red dots are the actual values from the test set, while the blue dots are the predicted values.
- The X depicted is the distance between the predicted and actual values.
- RMSE simply takes a mean of all of those distances, squares that value, and then takes the square root.
- A value under 180 is generally considered a good model.

ACTIVITY

Build a movie recommender using matrix factorization with ML.Net

THE END