# **Linear SVM vs Perceptron Algorithm**

Krishna Kumar Veeraputhiran

Grand Canyon University

DSC 540 – O500 – Machine Learning for Data Science

Dr. Aiman Darwiche

September 1,2021

# **Linear SVM vs Perceptron Algorithm**

In this report we will see the difference between a Perceptron Algorithm and a Linear Support vector machine (SVM) algorithm. In this article we will discuss on what is Linear SVM and what cases can we apply Linear SVM algorithm and similarly we will discuss the same on Perceptron Algorithm.

**Perceptron Algorithm**

Perceptron Algorithm is a binary classification machine learning algorithm and it is considered to be the simples type of Neural Network model. To simplify, a Perceptron algorithm takes an input, aggregates it (weighted sum) along with bias. This is called activation. If the Activation is greater than the threshold value then it predicts the output as 1 or else 0. In simple terms *Activation = Weights\* Input + Bias*. According to Gopal (2019), the perceptron takes a real-valued inputs xj where j=1,..n, calculates the linear combination of these inputs *∑nj=1 wjxj = wTx*, then outputs +1 if greater than the threshold (-wo) else outputs -1. The perceptron algorithm validates each element for the decision function and if it could not classify the element, it adjusts the weight till it can be classified.

The Perceptron algorithm has two update rules,

**w = w+Δw = w+ η yixi**

**wo = wo + Δwo = wo + η yi R2**

In the above function **η** is the learning rate and R is called the radius of the training data. As mentioned previously the intuition for the update rule is that yi = y̅I .(meaning the data point is correctly classified and Δw and Δwo will become 0). If yi <> y̅I , then the update rule tries to correct the w and wo till the points are correctly classified. For instance if the perceptron outputs a data point to -1 while the target value is +1, then the update rule attempts to correct the misclassification by rotating the decision surface. One of the biggest advantage of the Perceptron algorithm is that is its simplicity. However there are some drawbacks to the algorithm. One being that the Perceptron does not try to optimize separation distance of the hyperplane (Maximize the Margin). Also Perceptron algorithm can be used to find a hyperplane only for a linearly separable data. In case of non linear data, the update function would try to classify the data points and hence get stuck in an infinite loop.

**Linear SVM**

Support Vector Machine (SVM) is a supervised machine learning that is widely used for classification. SVM provides options for classifying both Linear as well as Non Linearly separable dataset. For classifying the linear dataset, we can use the linear kernel function. In a linear SVM, a hyperplane is identified which separates the classes in an n-dimensional space. For the linear data, the SVM finds the linear equation,

**g(x) = wTx + w0**

The input vector xi can be assigned to class 1 if g(x) > 0 and to class 2 if g(x) < 0. The hyperplane which separates the data is given by,

**wTx + w0 =0**

The distance of the closest data point say class 1 to the hyperplane can be given as,

d1= g(xi) / ||w||

Similarly the distance of the closest data point say class 2 to the hyperplane can be given as,

d2= g(xk) / ||w||

Now for class 1 we know that g(xi) >= 1, and hence the classification output is +1. Similarly g(xk) <=1 and hence the classification output can be given as -1. Now the distance between the two closest data points from each class to the hyperplane is called a Margin. So the margin can be calculated as d1 – d2.

**d1 - d2 = g(xi) / ||w|| - g(xk) / ||w||**

**= 1/||w|| - (-1)/ ||w||**

**=2/||w||**

A linear SVM algorithm is optimized to maximize the margin which makes the classification more robust.

**Conclusion**

Both Linear SVM and Perceptron algorithm help in tackling the classification of linearly separable dataset. Perceptron algorithm is a type of neural network model and can also be termed as generalized SVM. One of the key differences between the Perceptron and the Linear SVM is that the Linear SVM tries to maximize the margin between the two closest data points of the two classifications whereas the Perceptron does not do any optimization. Also Linear SVM uses the kernel function to project the vectors in a higher dimension and then classifies whereas the Perceptron assumes the input dataset is linearly separable, due to this the update steps runs into an infinite loop. Due to the above reason I prefer using the Linear SVM over the Perceptron algorithm,

# **References**

Gopal, M. (2019). *Applied machine learning*. McGraw-Hill Education.

Brownlee, J. (2020, December 11). *Perceptron algorithm for classification in Python*. Machine Learning Mastery. https://machinelearningmastery.com/perceptron-algorithm-for-classification-in-python/.

Chandra, A. L. (2018, August 22). *Perceptron learning algorithm: A graphical explanation of why it works*. Medium. https://towardsdatascience.com/perceptron-learning-algorithm-d5db0deab975.