Part 3 Results and Discussion

Evan Casper, Dallas Landry, and Parker Edwards

Dr. Jose Ramignoli

4/8/2022

A multilayer perceptron (MLP) is a type of neural network that has three layers of nodes in our case, an input layer, a hidden layer, and an output layer. The hidden layer takes inputs from the input layer that are weighted and uses them to compute some sort of activation function. Activation functions that were used for this experiment included sigmoids like tanh and logistic functions, and the rectifier linear unit (ReLU) function. Other hyperparameters that were varied included the number of neurons in the hidden layer, and the solver method being used to optimize the activation function.

A support vector machine (SVM) is another type of machine learning algorithm that works by attempting to find a hyperplane in a dimensional space equal to the number of dimensions in the dataset. For example, in a two-dimensional data set, the hyperplane in question would the line that best separates the data, and in a three-dimensional space, the hyperplane becomes a plane. The algorithm picks this hyperplane by maximizing the support vectors, which in 2-D space one can imagine as lines drawn from the closest data points to the hyperplane. By choosing the hyperplane with the longest support vectors, we choose the optimal hyperplane that best separates the data. While nonlinear hyperplanes can be used, our tests focused on linear for ease of analysis.

First, the known operation data and the actual data was split into testing and training sets. In order to keep an accurate point of comparison, all tests were done on the same splits. Since this is meant to be a fault classifier, both methods were trained on the known operation data in order to be able to identify when the actual data deviates. The criterion of comparison is the sklearn accuracy score, which takes a ratio of the number of times the predicted labels from the machine learning algorithm perfectly matched the true labels to the total amount of labels being evaluated, meaning this ranges from 0-1. While both SVM and MLP did not generate an accuracy score higher than .3, SVM slightly outperformed MLP on average with default hyperparameters. The next thing to investigate was tuning the hyperparameters to improve the accuracy score. First, the activation function in question was varied. Prior research suggested that the sigmoid functions are outdated, and the preferred activation function was the ReLU. Our work supported this, as ReLU gave a higher accuracy score compared to the others. The next hyperparameter investigated was the amount of neurons in the hidden layer, and our findings suggested 100 neurons was optimal among the number tested, however this began to give convergence errors. This was fixed by changing the solver method from “adam” to “lbfgs”, which research states usually has a faster convergence on smaller data sets.

In conclusion, both machine learning algorithms gave subpar results for when testing the accuracy of the model, however, the fine tuned ANN generally outperformed the SVM. The inaccuracy can possibly be attributed to the fact that the ANN is only has one hidden layer. Based on previous work and research, more layers generally improves the performance of models. For the SVM, it is possible that a nonlinear kernel could better separate the data, and future work should investigate this.