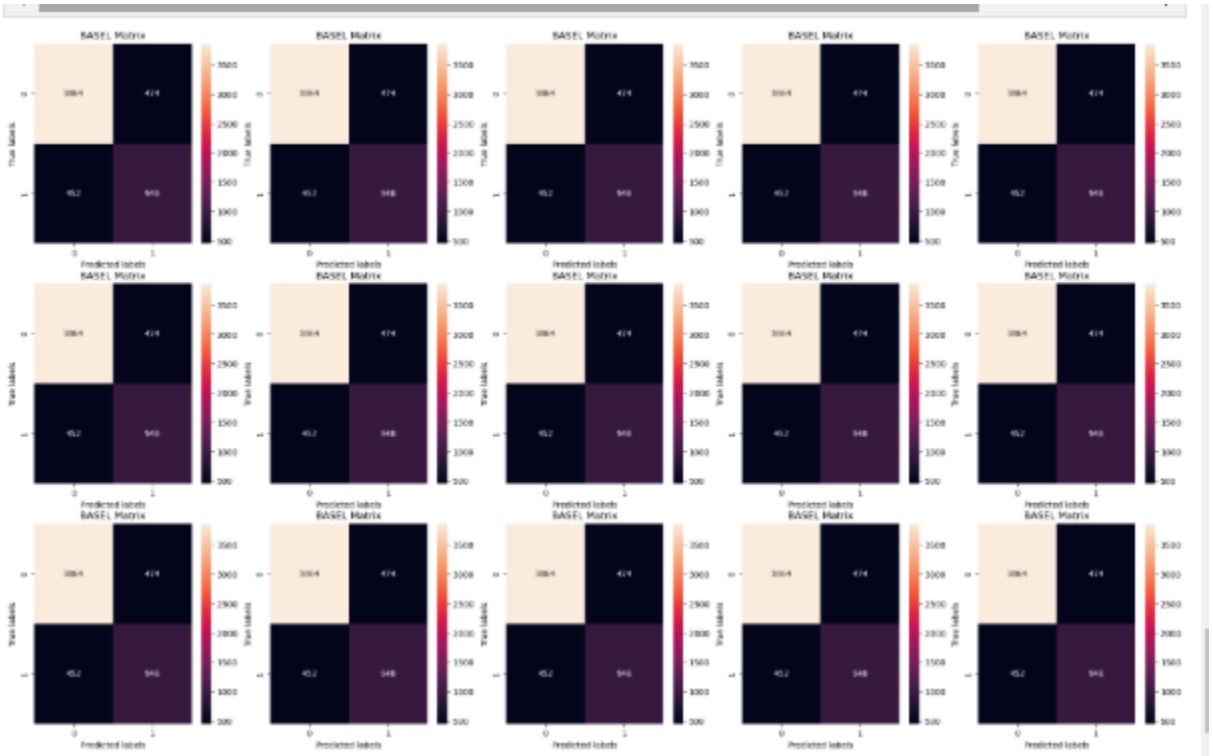
The accuracy of the decision tree was only 46 percent. I think the decision tree would definitely need to be pruned. I think the model is taking too many factors into consideration making it inaccurate. Removing branches may reduce the overfitting in the model, improving accuracy and run time. The testing confusion matrix confirms that there were plenty of incorrect classifications by the model.

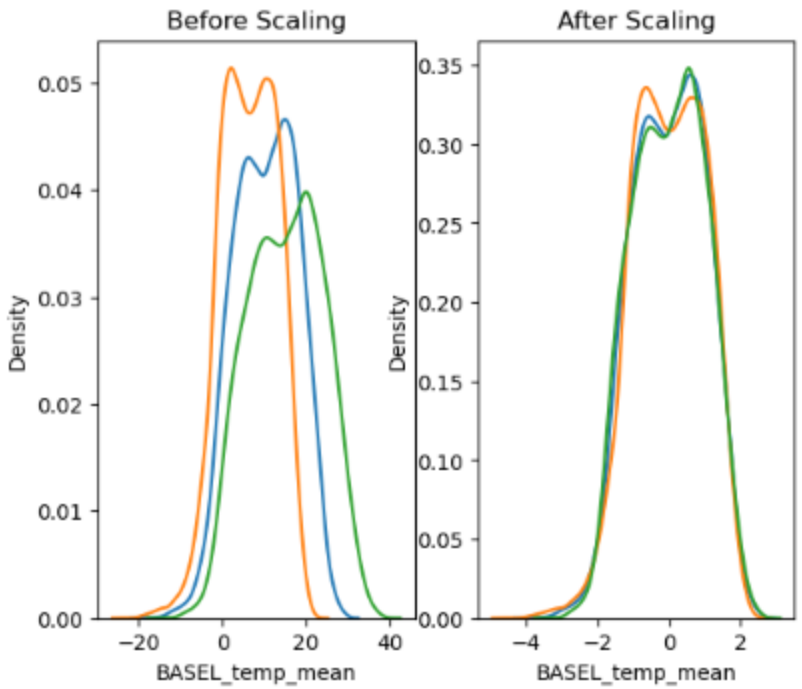
Training Data

Testing Data



ANN

I scaled the data and plotted the scaling differences for Basel temp mean, min, and max in the code. The before scaling chart makes sense to me because we are dealing with a minimum, maximum, and average value. Scaling makes the three values comparable.

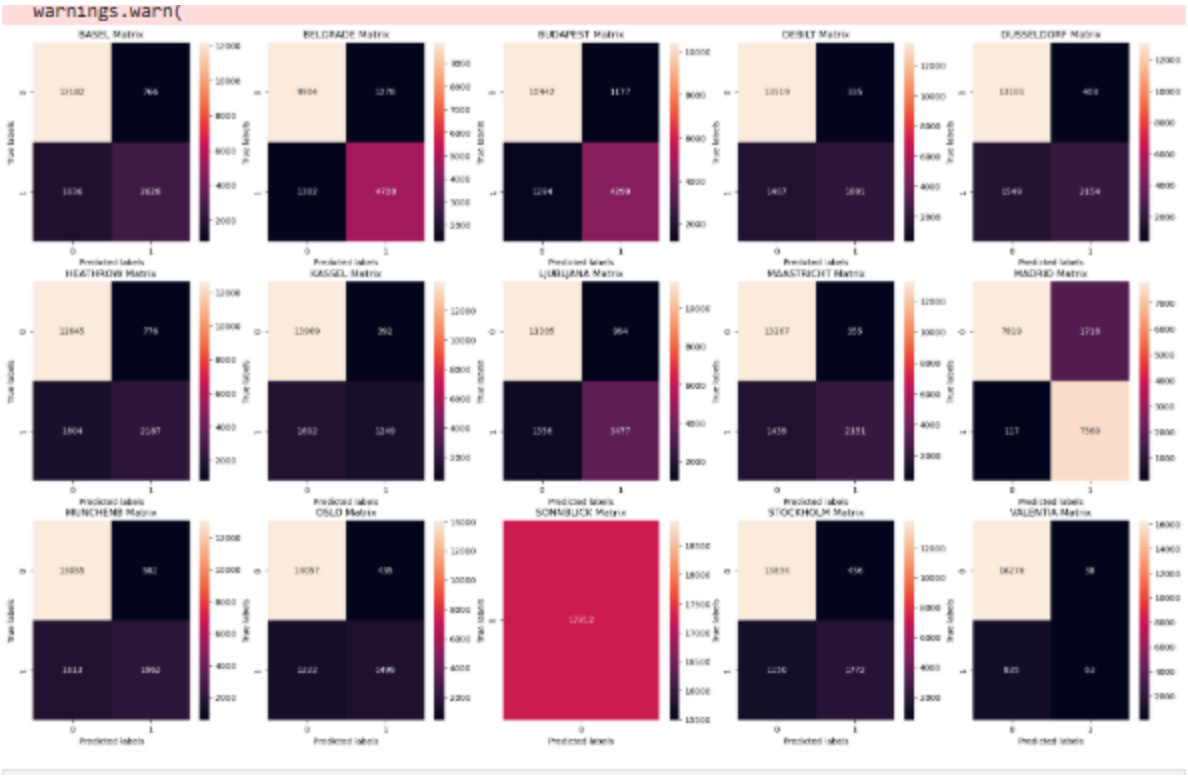


Hidden Layers	Nodes	Max Iterations	Tolerance	Train Accuracy	Test Accuracy
2	10, 5	500	.0001	46%	47%
3	10, 5, 5	1000	.0001	47%	47%
3	10, 10, 10	1000	.0001	48%	48%
3	15, 15, 15	500	.01	47%	47%

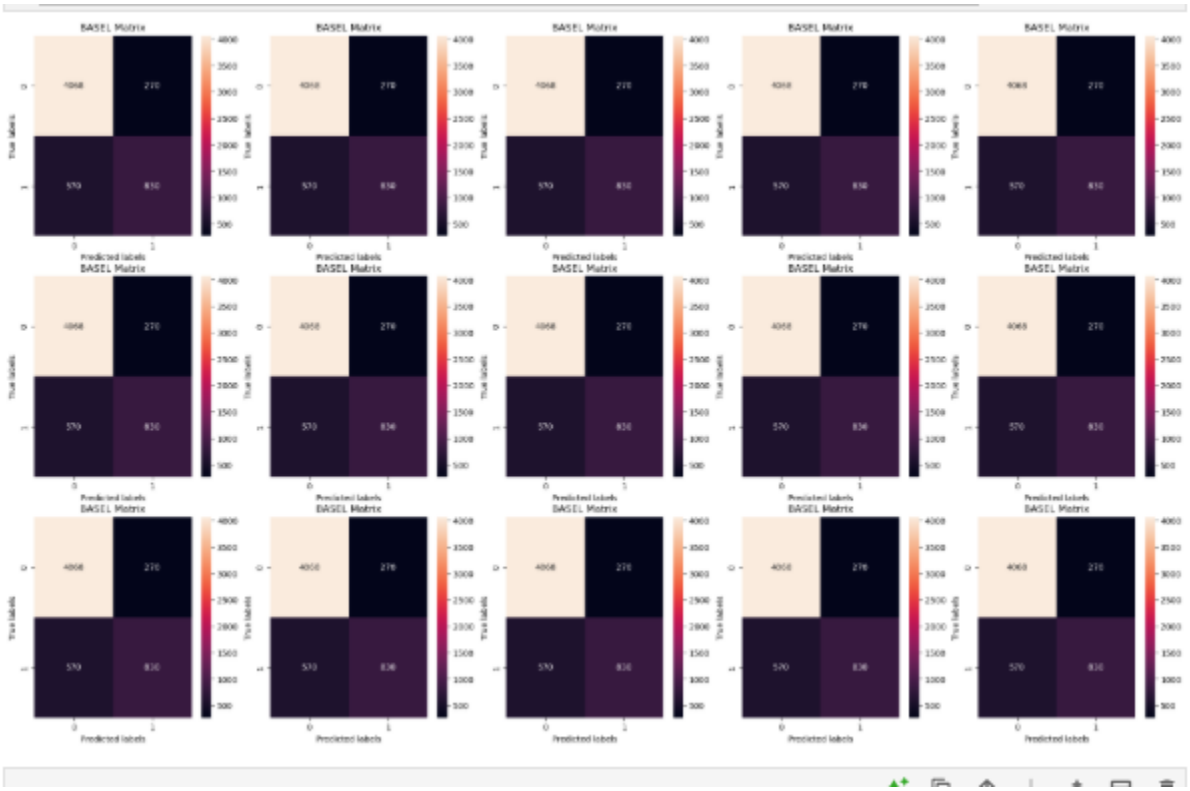
It does not appear that adjusting the hidden layers, nodes, max iterations or tolerance is improving the accuracy of this model, suggesting this may not be the best model to use in this situation.

Hidden Layers	Nodes	Max Iterations	Tolerance	Train Accuracy	Test Accuracy
2	10, 5	500	.0001	46%	47%

Train

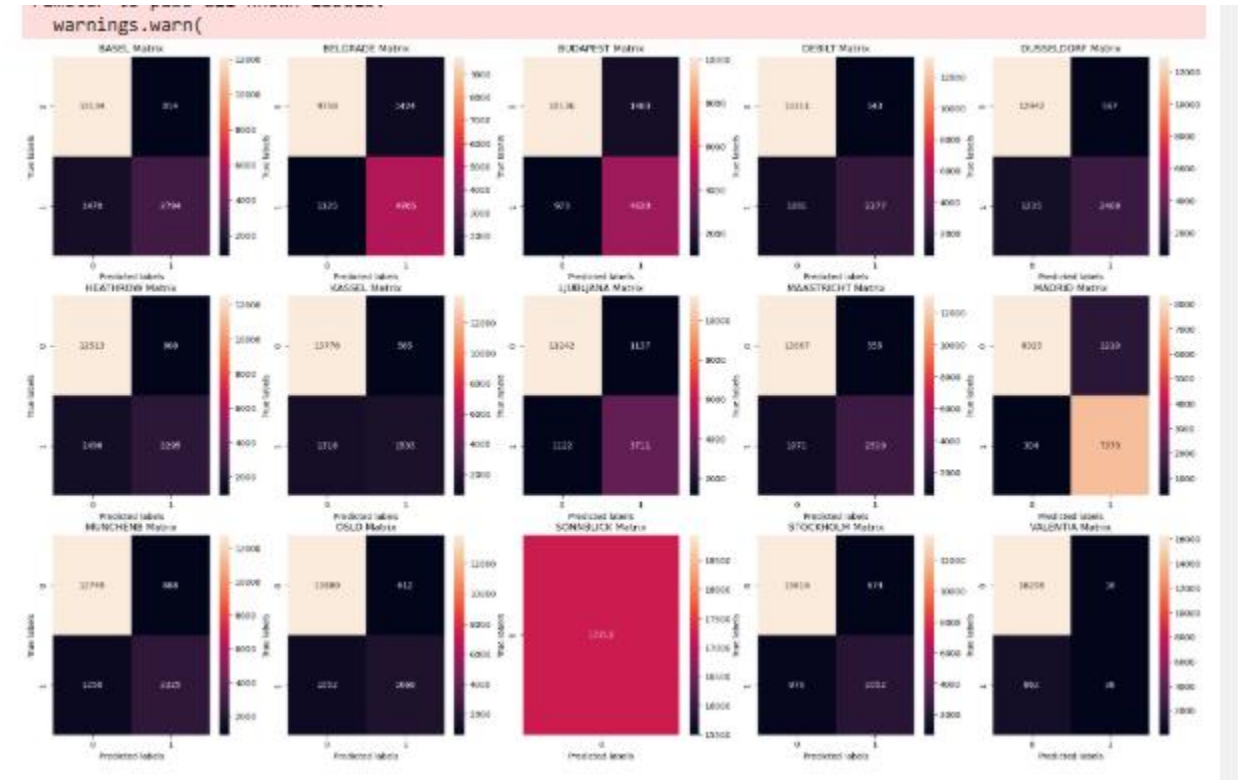

warnings.warn(
BASEL Matrix
BELGRADE Matrix
BUDAPEST Matrix
DEBILT Matrix
DUSSELDORF Matrix
HEATHROW Matrix
KASSEL Matrix
LJUBLJANA Matrix
MAASTRICHT Matrix
MADRID Matrix
MUNCHEN Matrix
OSLO Matrix
SONNBLICK Matrix
STOCKHOLM Matrix
VALENTIA Matrix

Test

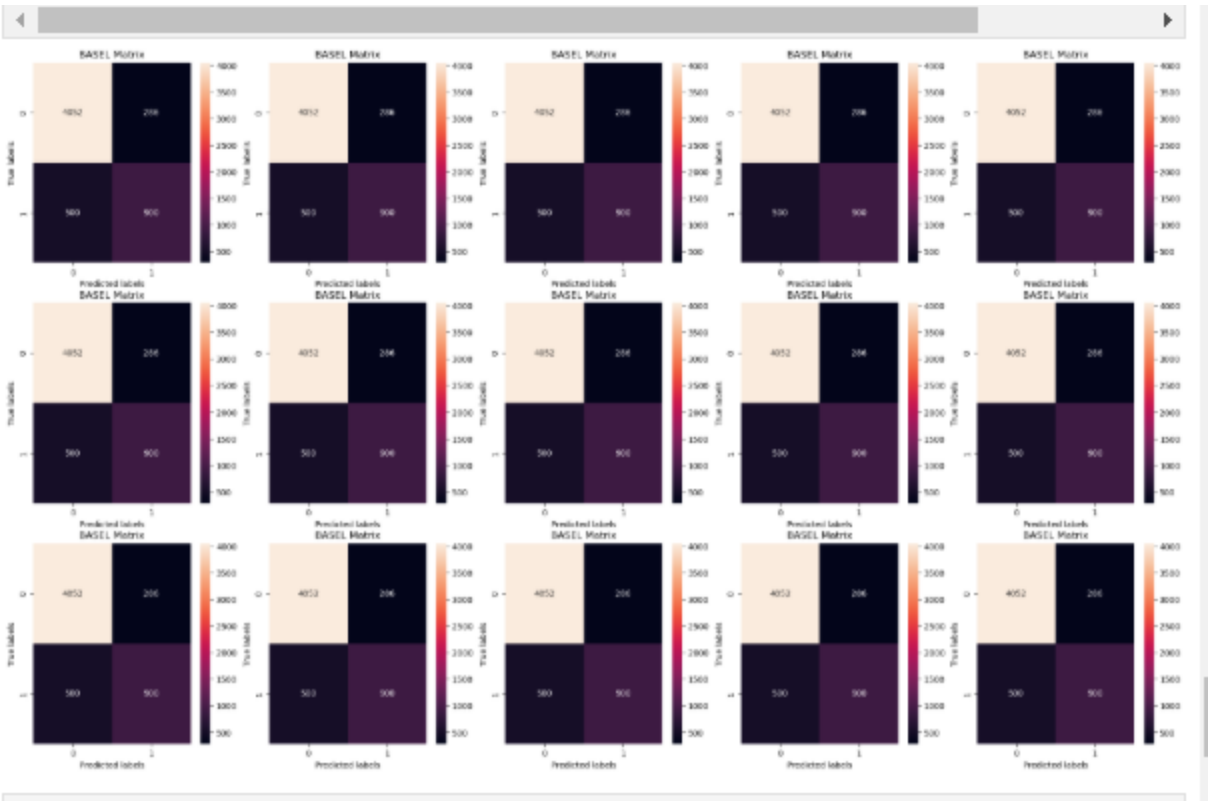

BASEL Matrix

Hidden Layers	Nodes	Max Iterations	Tolerance	Train Accuracy	Test Accuracy
3	10, 5, 5	1000	.0001	47%	47%

Train



Test

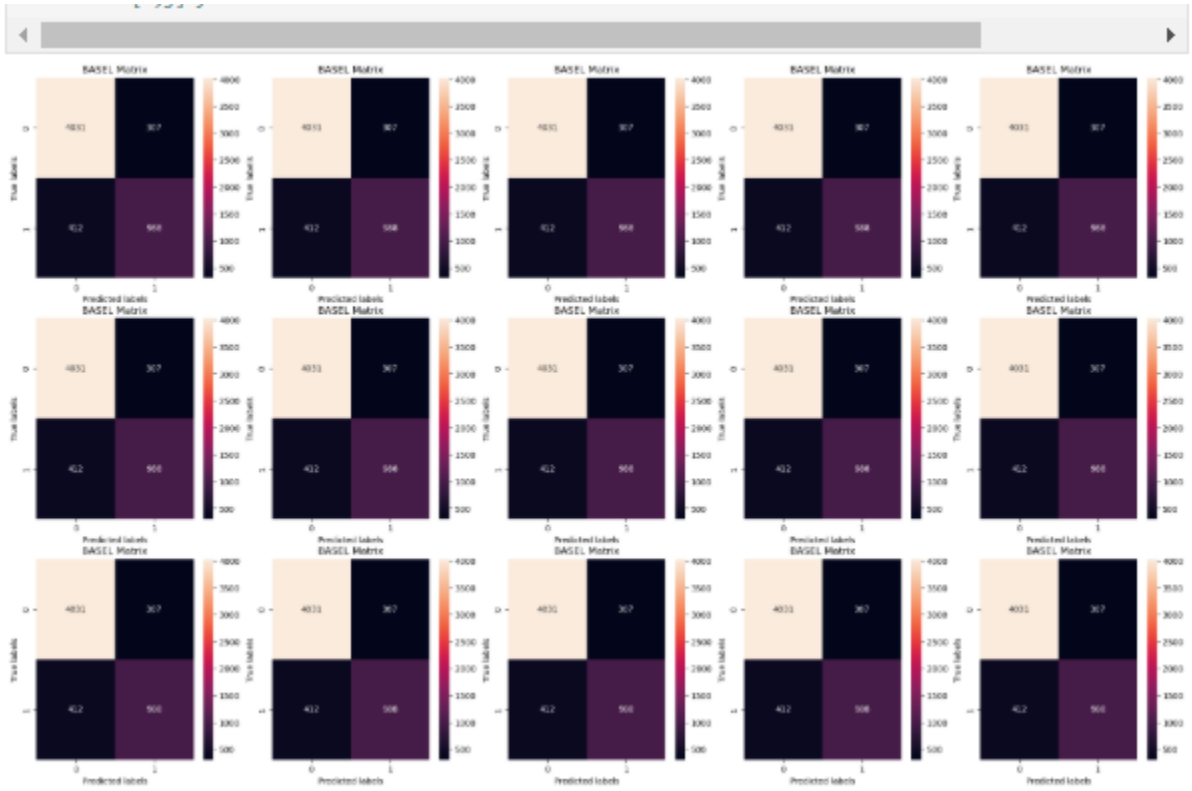


Hidden Layers	Nodes	Max Iterations	Tolerance	Train Accuracy	Test Accuracy
3	15, 15, 15	500	.01	48%	48%

Train


rameter to pass all known labels.
warnings.warn(
BASEL Matrix
BELGRADE Matrix
BUDAPEST Matrix
DUSSELDORF Matrix
HEATHROW Matrix
KASSEL Matrix
LJUBLJANA Matrix
MAASTRICHT Matrix
MADRID Matrix
MUNCHEN Matrix
OSLO Matrix
SONNBLICK Matrix
STOCKHOLM Matrix
VALENTIA Matrix
True labels
Predicted labels

Test


BASEL Matrix
True labels
Predicted labels

I would recommend that ClimateWins use the KNN model as its accuracy was between 83% and 100%. I believe this model with modifications would be the best for ClimateWins to use.