Matthew Wang

Averages Matrix	sepal length	sepal width	petal length	petal width
Iris-setosa	5.006	3.418	1.464	0.244
Iris-versicolor	5.936	2.77	4.26	1.326
Iris-virginica	6.588	2.974	5.552	2.026

STDEV Matrix	sepal length	sepal width	petal length	petal width
Iris-setosa	0.352489687	0.381024398	0.173511159	0.107209503
Iris-versicolor	0.516171147	0.313798323	0.469910977	0.19775268
Iris-virginica	0.635879593	0.322496638	0.551894696	0.274650056

T Tests	sepal length	sepal width	petal length	petal width
setosa-versicolor	-10.52098627	9.282772556	-39.46866259	-34.01237859
setosa-virginica	-15.38619582	6.289384997	-49.96570336	-42.73822967
versicolor-virginica	-5.62916526	-3.20576075	-12.60377944	-14.62536705
Averages:	-10.51211578	4.122132267	-34.01271513	-30.45865844

Equation for T test:

t = (mean_1 - mean_2)/sqrt(variance_1/n_1 + variance_2/n_2)

Equation for Signal to Noise:

S = |mu1-mu2|/(sigma1+sigma2)

Signal to Noise Ratio	sepal length	sepal width	petal length	petal width
setosa-versicolor	1.070613481	0.932611989	4.345514151	3.547980897
setosa-virginica	1.600616319	0.631111192	5.635465955	4.666637143
versicolor-virginica	0.565947295	0.320606028	1.26442829	1.481786508
Averages:	1.079059032	0.628109736	3.748469465	3.232134849

The T-test is used to determine if two distributions are statistically similar or different. The T-test is used for pairs so the approach is to do each combination of classes to determine if they are statistically different. The result would be the a t-value that must be larger than the value of the table in order to reject the null hypothesis and show that the two distributions are statistically different. To determine which are the two best features, an average is taken of the t-values for each feature and top two highest averages are selected. From the results above, petal width and petal length should be selected.

The signal to noise ratio is used to determine how strong the signal is for a feature compared to the noise of the feature. A larger value is more desirable and shows that there is a stronger signal in the feature.
The value is calculated using the averages of the feature in each class and the standard deviation of the feature in each class. Since this is also a pair method, each combination pair is calculated with an average over all pairs.
The two features with the highest signal to noise ratios are selected.

From the results above, petal width and petal length should be selected, which also agree with the above t-test.