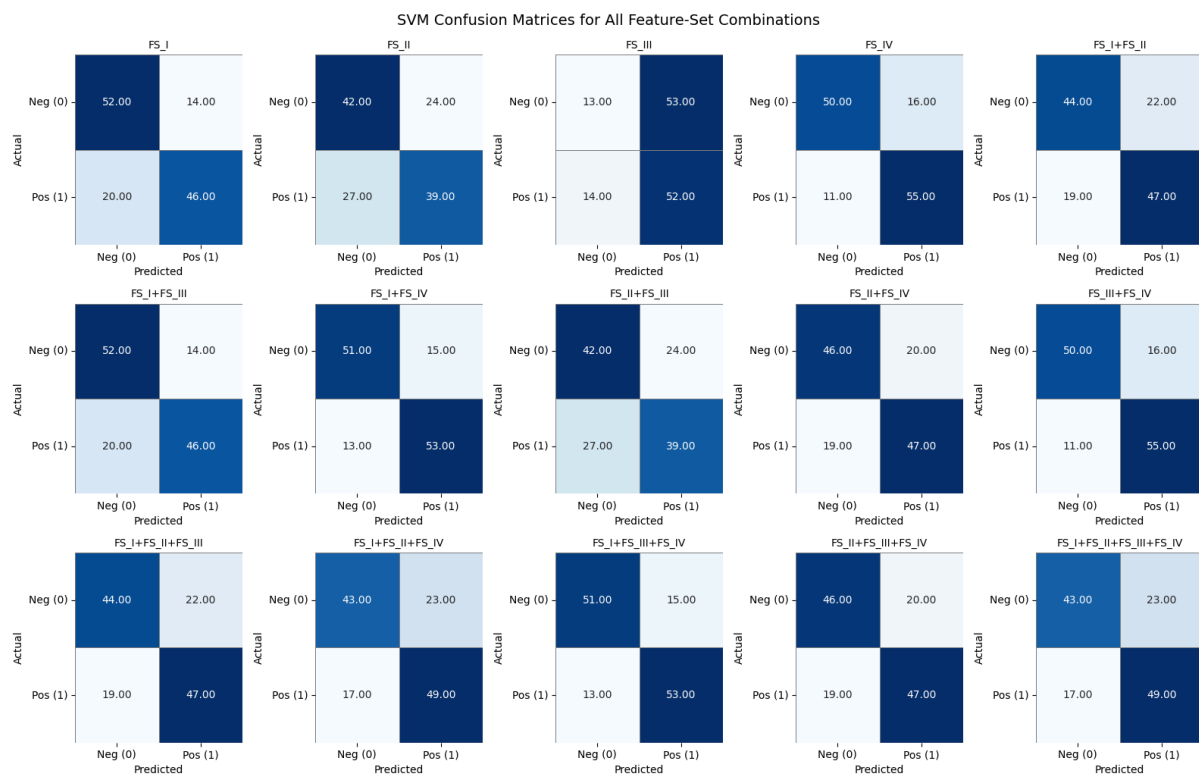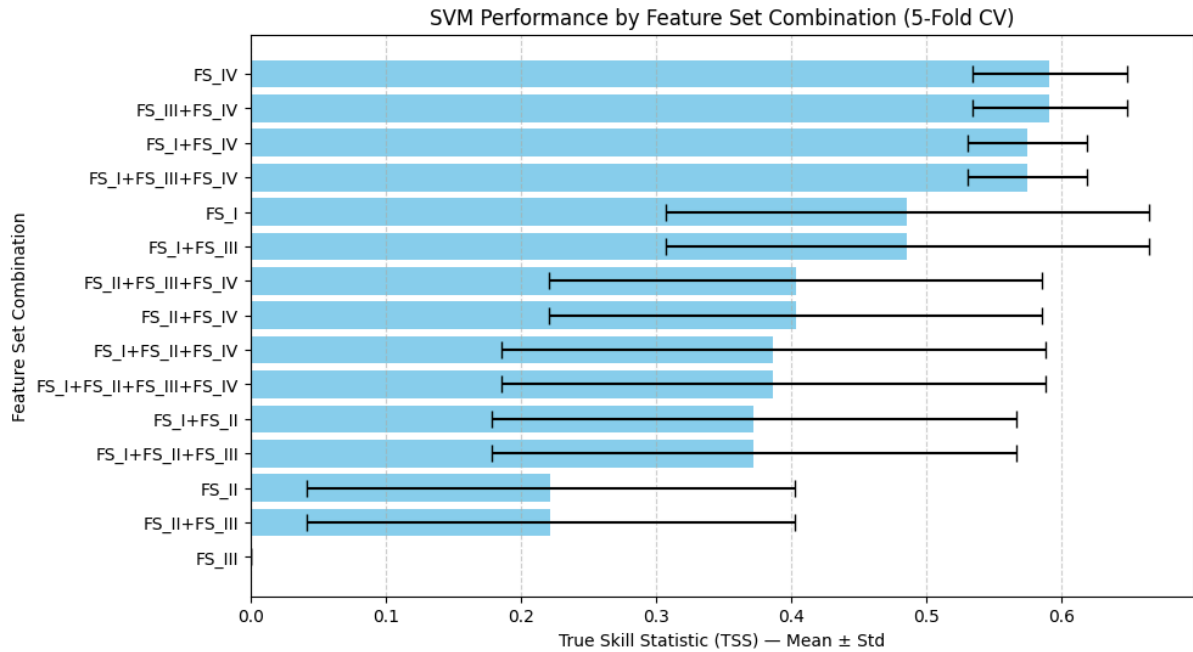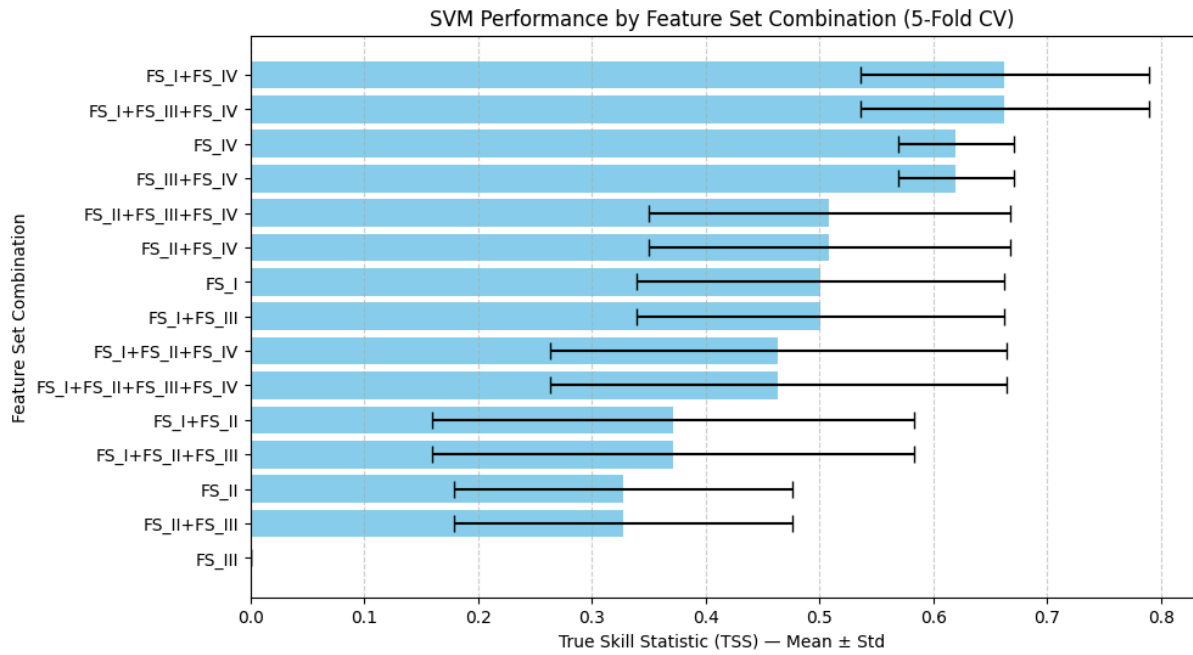1. Bar chart showing the performance of each model feature set combination across each fold
2. Confusion matrixes
3. Means and standard deviations

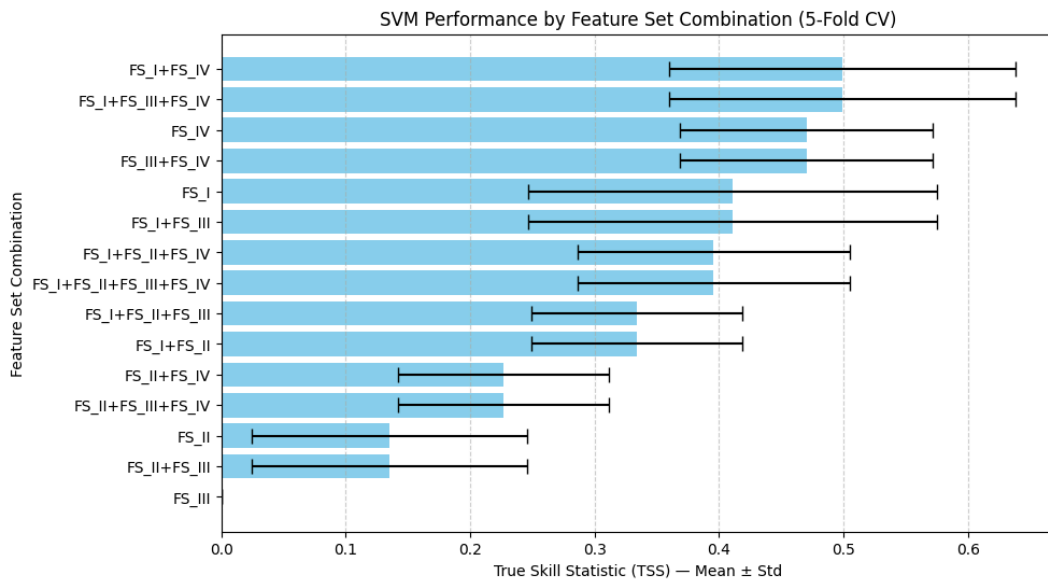**Dataset = 2020**

C = 1, gamma = 1 / n_features, kernel = linear

SVM Performance by Feature Set Combination (5-Fold CV)

SVM Confusion Matrices for All Feature-Set Combinations

C = 0.5, gamma = 1 / n_features, kernel = linear

SVM Performance by Feature Set Combination (5-Fold CV)

C = 1, gamma = 1 / (n_features * X.var()), kernel = poly

SVM Performance by Feature Set Combination (5-Fold CV)

C = 1, gamma = 1 / (n_features * X.var()), kernel = rbf

SVM Performance by Feature Set Combination (5-Fold CV)

SVM Confusion Matrices for All Feature-Set Combinations

C = 5, gamma = 1 / (n_features * X.var()), kernel = rbf

SVM Performance by Feature Set Combination (5-Fold CV)
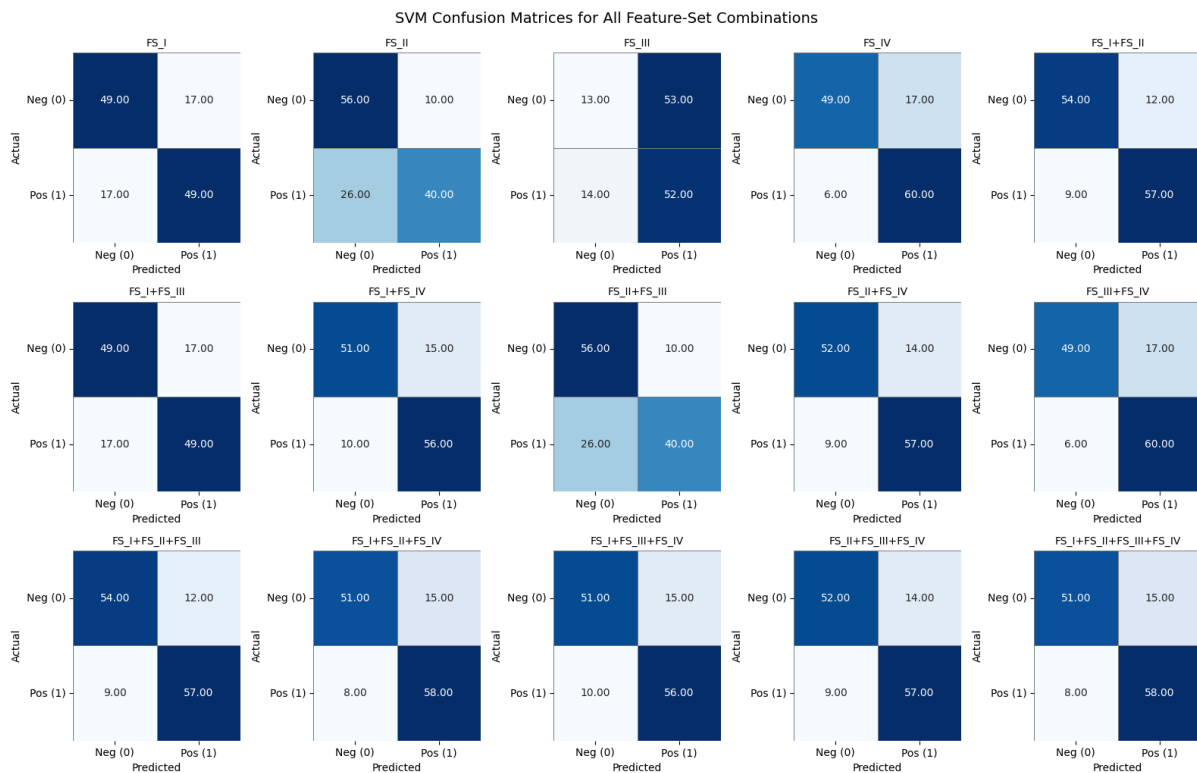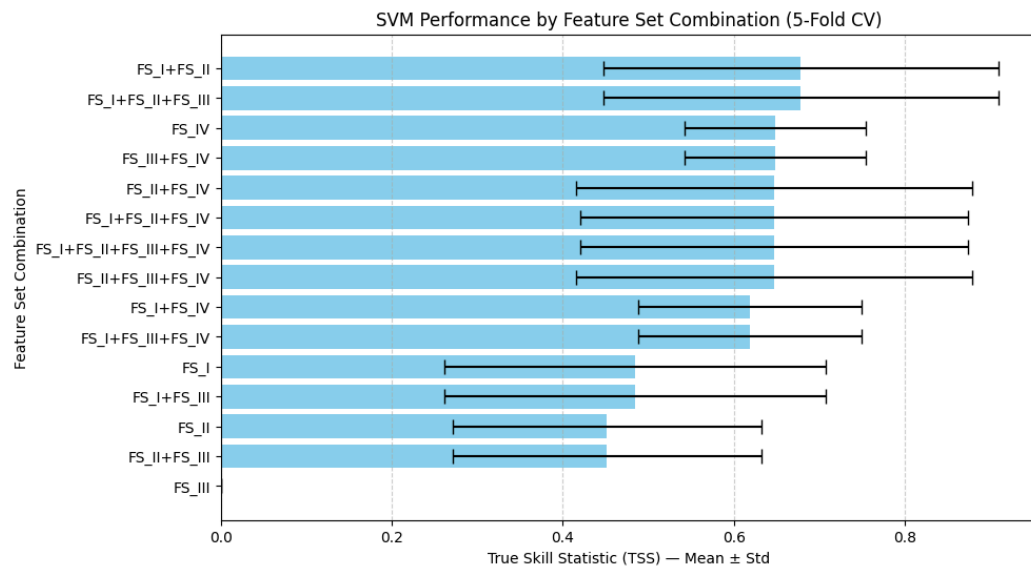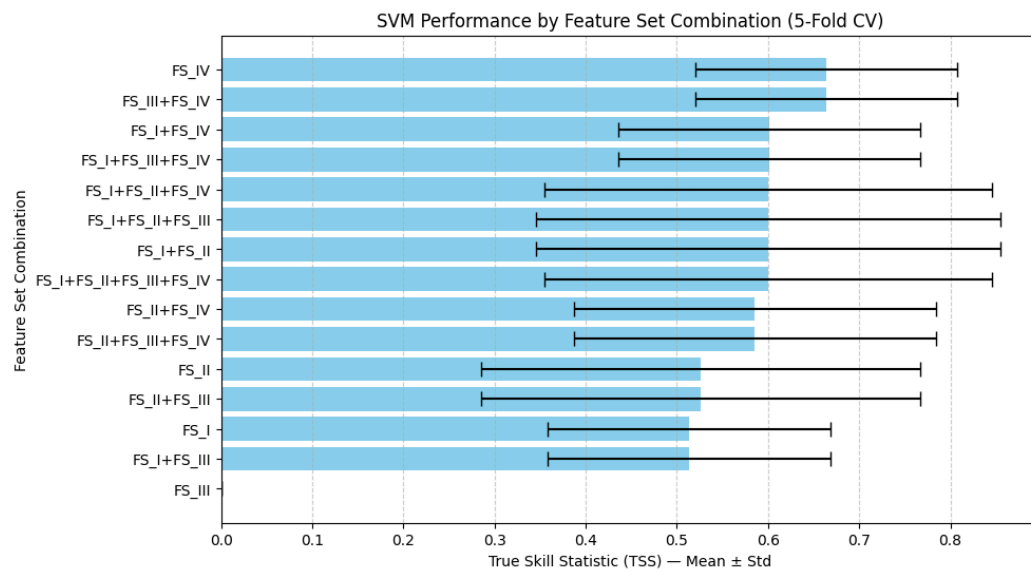
Dataset = 2010
C = 1, gamma = 1 / n_features, kernel = linear

SVM Performance by Feature Set Combination (5-Fold CV)

SVM Confusion Matrices for All Feature-Set Combinations

C = 5, gamma = 1 / (n_features * X.var()), kernel = rbf

SVM Performance by Feature Set Combination (5-Fold CV)

SVM Confusion Matrices for All Feature-Set Combinations

C = 1, gamma = 1 / (n_features * X.var()), kernel = poly

SVM Performance by Feature Set Combination (5-Fold CV)
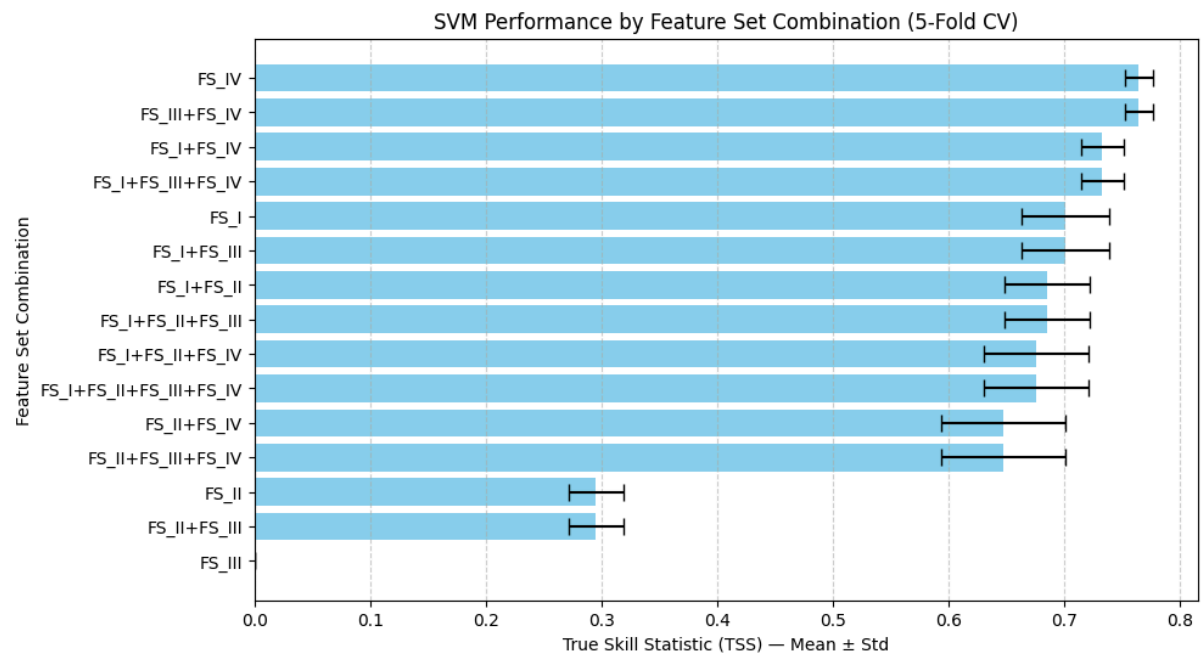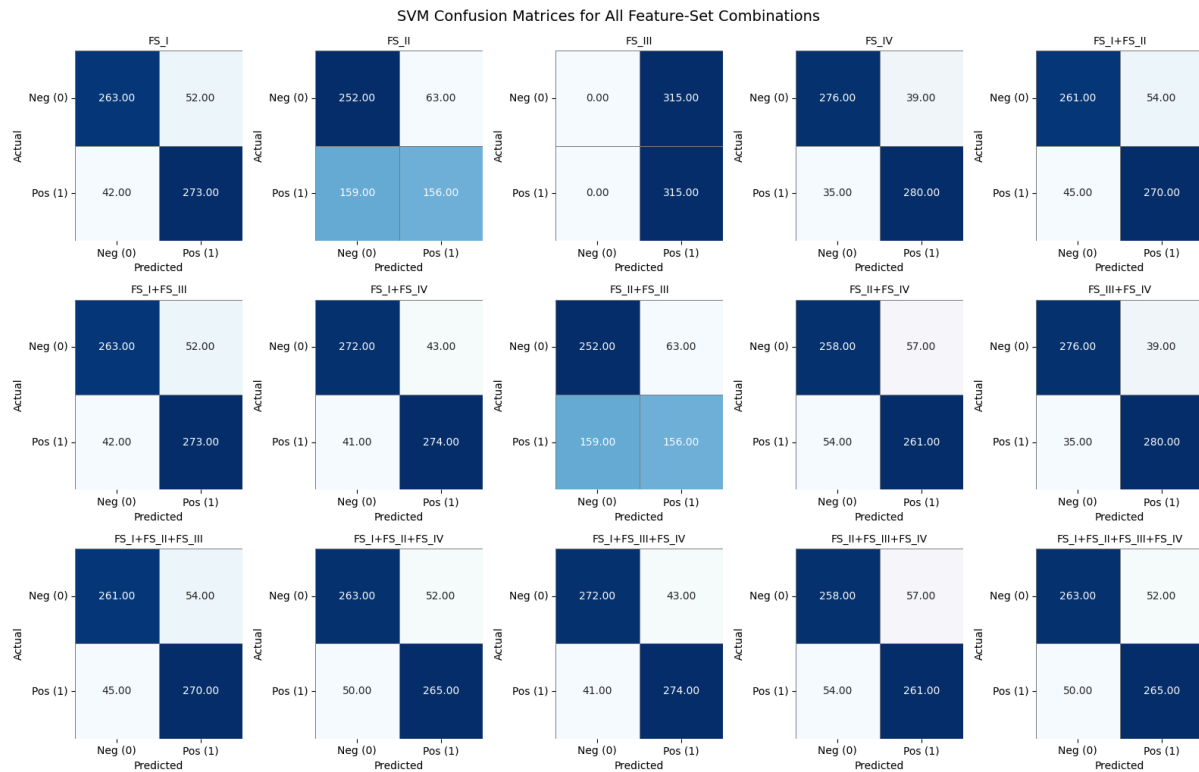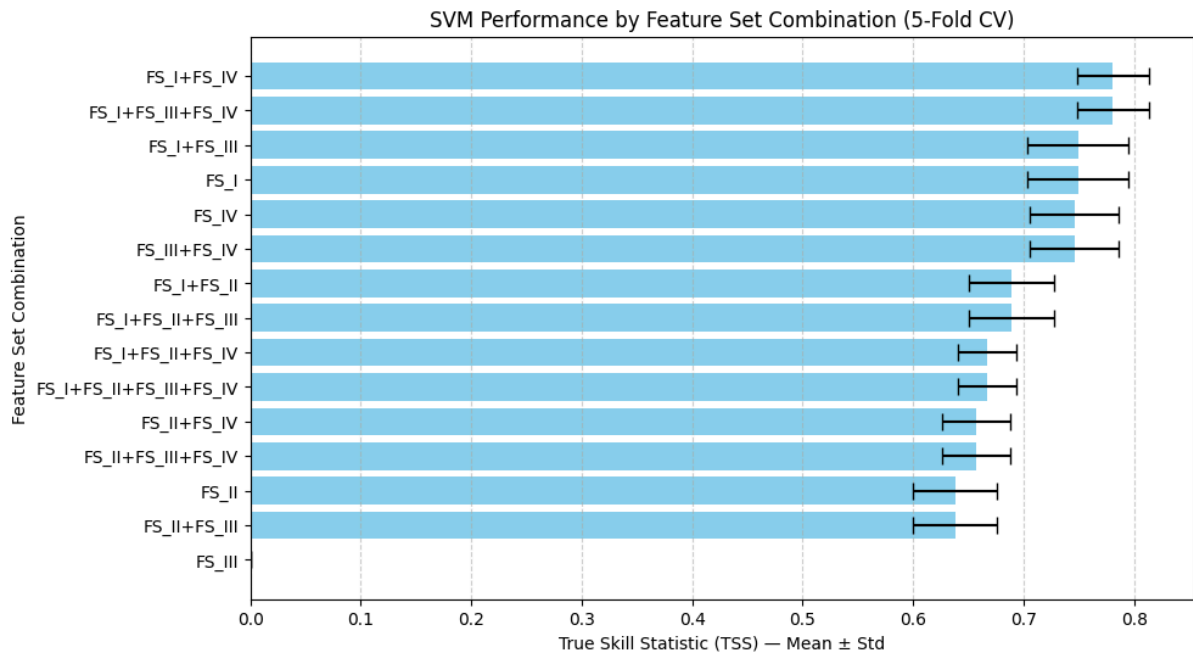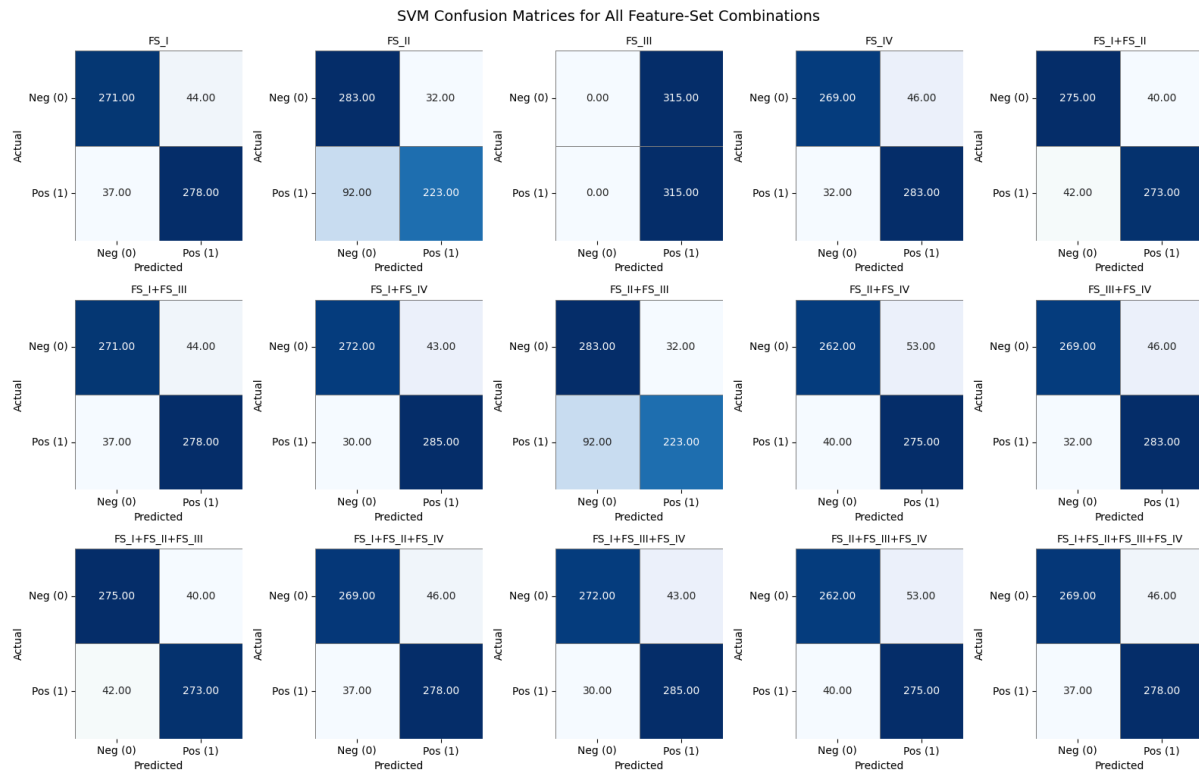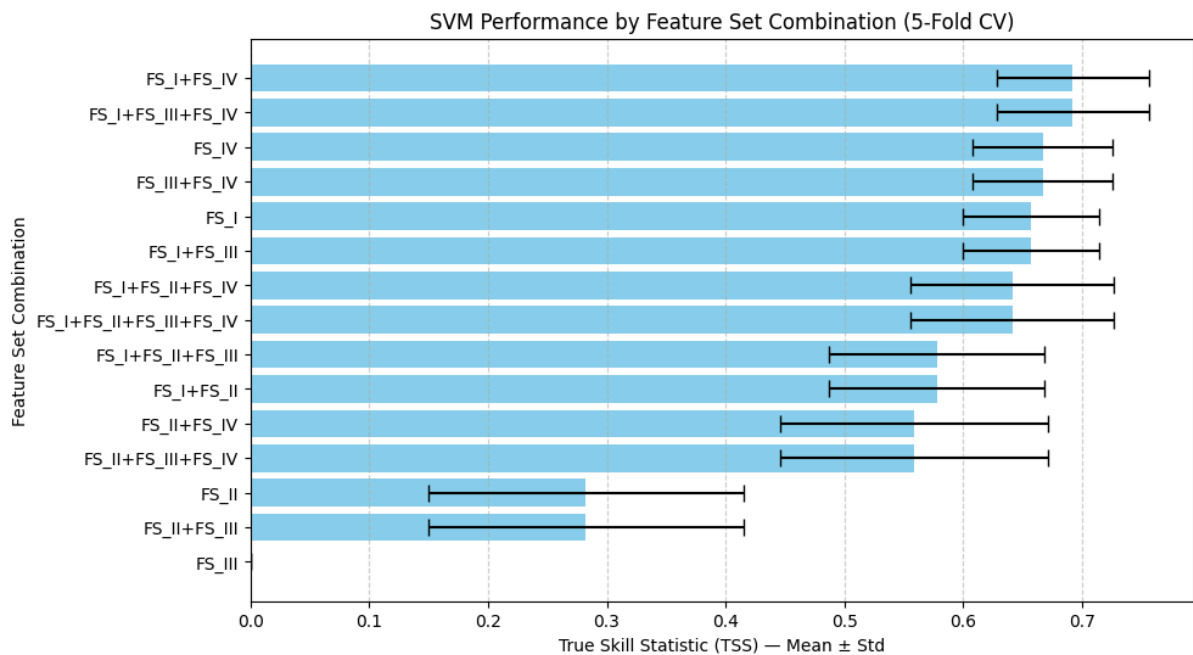
From visualizing all the feature sets I found the linear or rbf as the kernel works the best. Polynomial had very low accuracy. I also noted that 2010 dataset was more accurate than the 2020 dataset.

1. **Which feature combination worked best and which feature set was the worst?**

Across both datasets and all hyperparameter tuning. FSI + FSIV was usually the highest performing model. Sometimes FSI + FSII. the worst feature set was by far FSIII (but thats because the dataset is broken so disregarding that) standalone FSI or FSII i found at the bottom alot. The more complicated datasets combining multiple features were usually floating in the middle.

Dataset	Kernel	Best FS	TSS approx.	Worst FS	TSS approx.
2020	Linear	FSI + FSIV	0.68	FSII	0.33
2020	rbf	FSIV	0.68	FSI	0.51
2010	Linear	FSIV	0.76	FSII	0.29
2010	rbf	FSI + FSIV	0.76	FSII	0.64

(b) Does adding additional feature sets improve the TSS score? What do you observe?
From the bar graphs i observed that adding additional features usually improved the TSS score. Standalone features was usually lower ranked on the TSS bar graph. Adding alot of features usually hovered in the middle of the rankings somewhere. The most optimal rankings were combining 2 features. I.e. FSI + FSIV

(c) Which dataset led to a better TSS score (2010 or 2020)? Why?
The 2010 dataset clearly produced a way higher Tss score than the 2020 dataset. 2020 highest rankings were usually in the 0.6-0.7 margin while the 2010 dataset top ranking was in the 0.7-0.8 margins. This could be due to many factors such as the 2010 dataset could be more clean, less variability, less noise in the data. The data could be more clean compared to the new 2020 dataset.