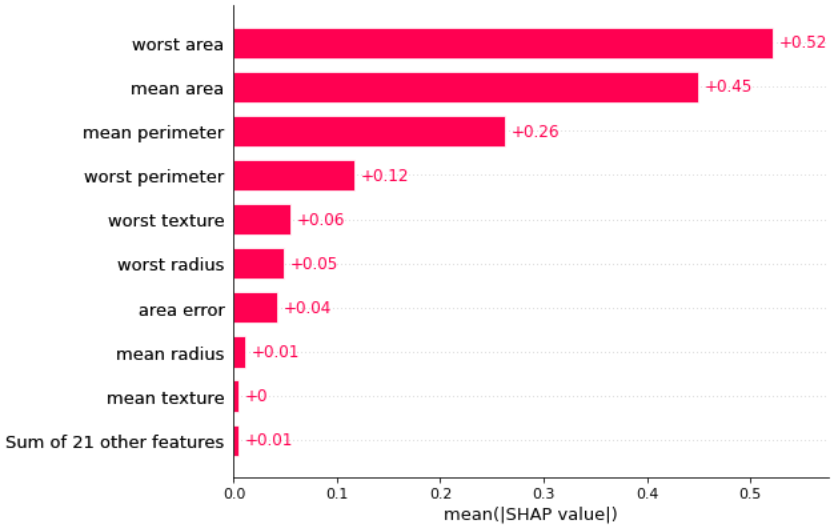
Vincent Purcell

Data Mining 2 – Summer 2021

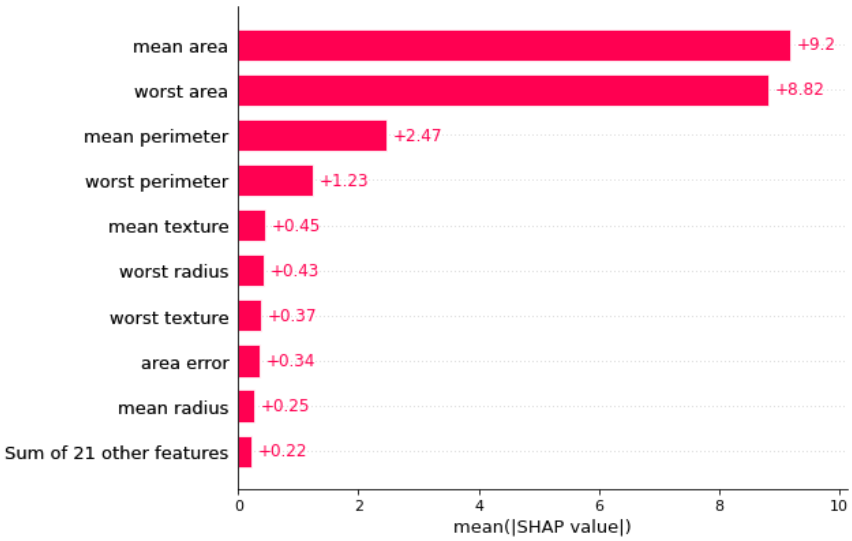
HW2 – Sensitivity Analysis

Are the most important features for the SVM the same as the NN?

For my SVM I used a polynomial kernel and switched it to regression the most important features were "worst area", "mean area", and "mean perimeter" as seen below.



For my NN I used a MLP regression model, and the most important features were also found to be the same 3 in a different order of importance, as seen below. Though the MLP model favored the mean and worst area features much more than the SVM model.



In reference to the Random Forest, the only similarity was the importance in the "worst area" feature.

Challenge Question

For the challenge question I decided to investigate alternative methods to sensitivity analysis.

I found an alternative method called Fourier amplitude sensitivity analysis. I found this particularly interesting due to my Computer/Electrical Engineering background and how I could not escape Fourier series and transforms in my courses.

Fourier amplitude sensitivity Test or FAST uses periodic sampling approach and a Fourier transformation to decompose the variance of a model output into partial variances. I found a package in SALib that implements an extended FAST found here:

https://salib.readthedocs.io/en/latest/api.html#fast-fourier-amplitude-sensitivity-test

I looked into trying to implement this with the models that I generated for this homework, but This package seems to be somewhat outdated, and the documentation is not descriptive. For example, this function takes a dictionary as an input and the description of what that is referencing is 'The problem definition.' I gave up on trying to run this sensitivity test.

I began looking more into this analysis method and found that generally it takes a multivariate function as the model and is able to analyze models with features that have imaginary numbers. This method certainly has its use cases and I found it interesting due to the relevancy with my past coursework in undergrad.