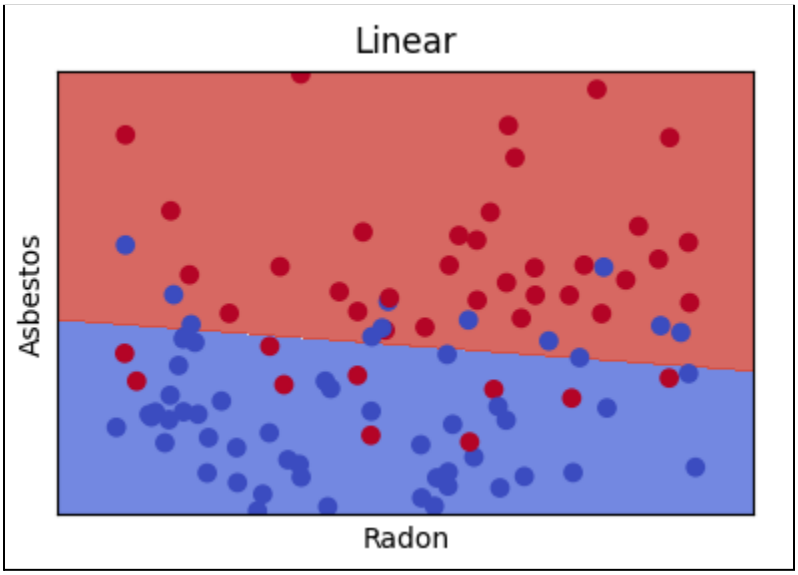
Lorenzo DeSimone
CS 4342 Machine Learning
Professor Jacob Whitehill
Homework 5

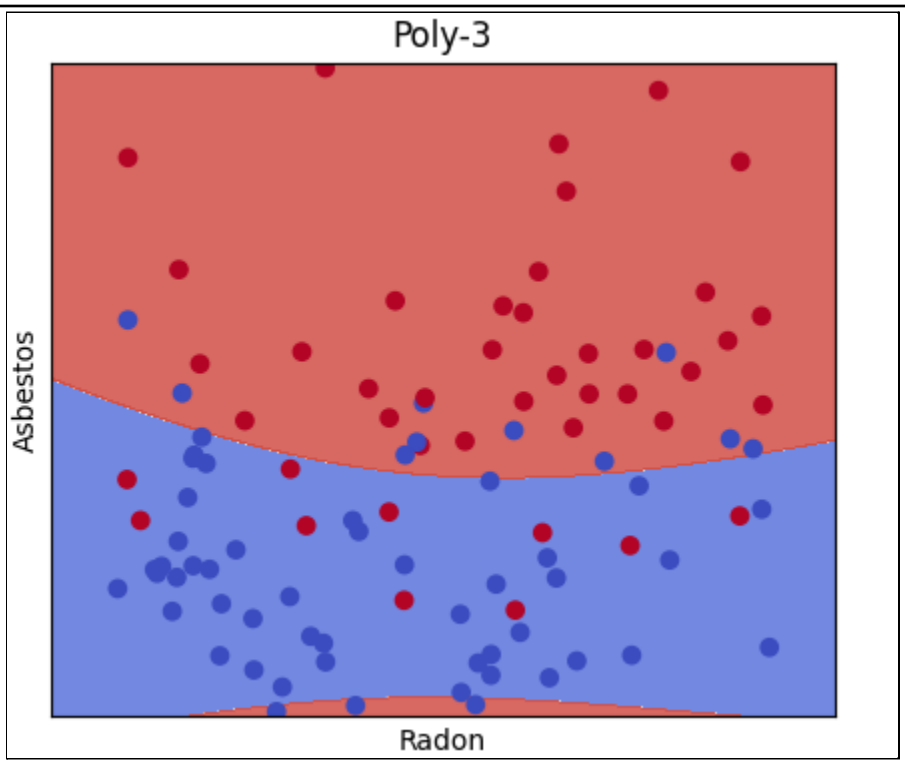
Linear SVM:
Plot (include it in your PDF) the trained machine's predictions on a 2-D grid of combinations of (radon, asbestos) values, where the color of each point in the scatter-plot indicates the predicted class label



Poly-3 with φ:

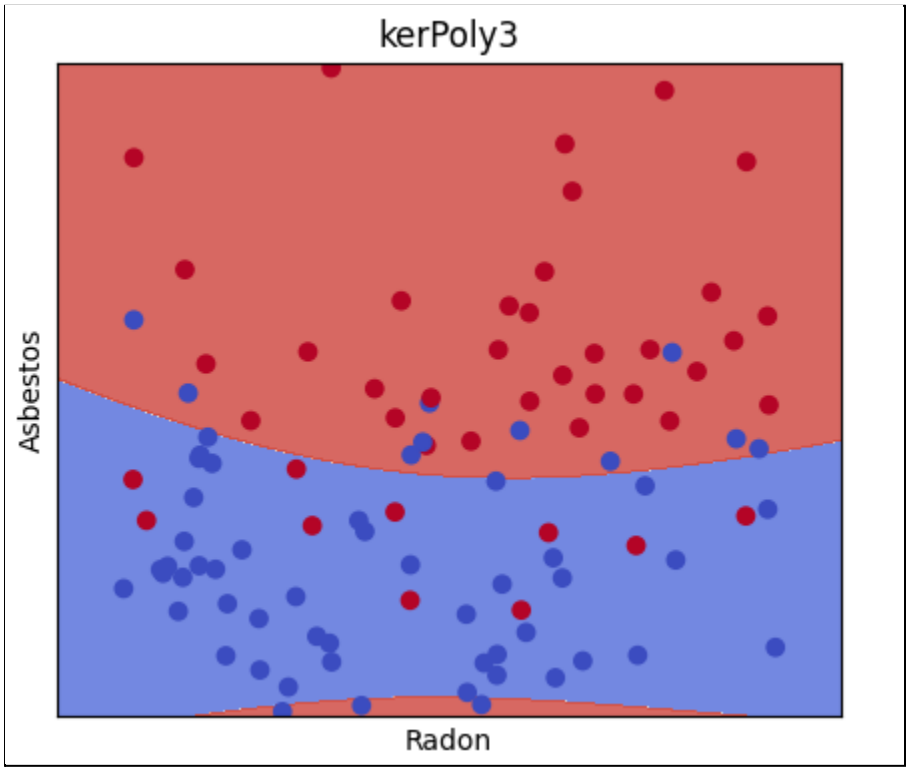
Write down the mathematical expression of your transformation in your PDF. (Hint: it should map from $\mathbb{R}^2$ to $\mathbb{R}^{10}$.)

Then apply this transformation to every example in the training set. Next, train an SVM on the transformed examples. (Since you are performing the transformation yourself, you should specify the linear kernel.) Finally, plot (include it in your PDF) the trained machine's predictions on a 2-D grid of combinations of the raw (radon, asbestos) values, like you did for part (a) above

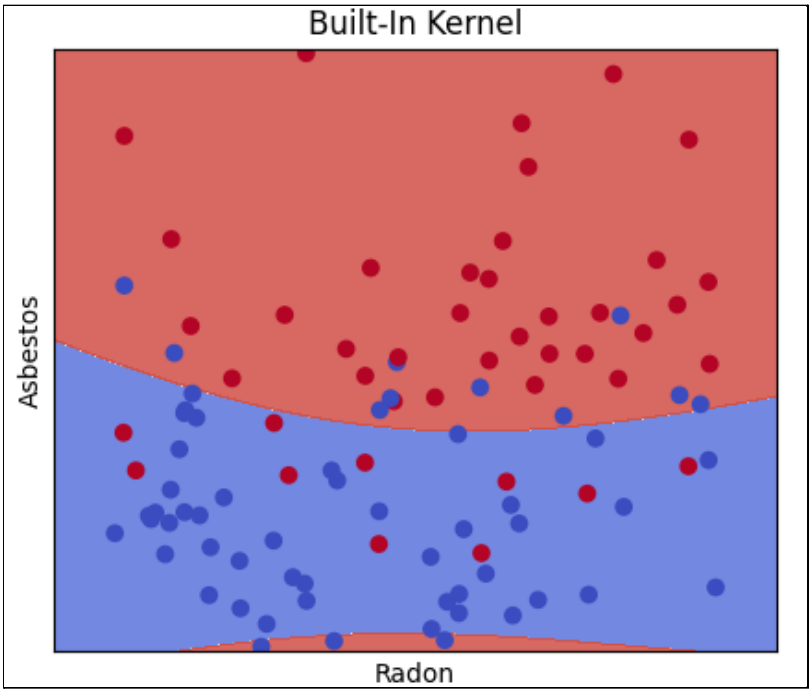


$$\phi(a, r) = [1, \sqrt{3}a, \sqrt{3}r, \sqrt{6}ar, \sqrt{3}a^2, \sqrt{3}r^2, \sqrt{3}a^2r, \sqrt{3}ar^2, a^3, r^3]$$

KerPoly3:
Using this procedure, create another prediction plot (include it in your PDF); it should look identical to part (b).

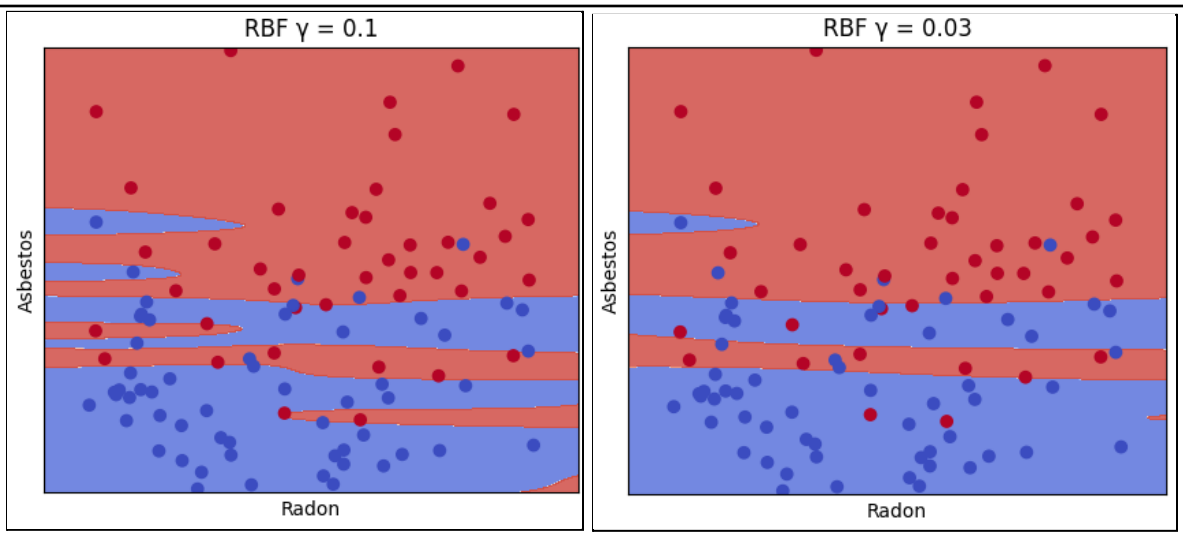


Built-In Kernel:
After training your SVM, create another prediction plot (include it in your PDF); it should look identical to parts (b) and (c).



RBF:

Describe in your PDF how (in qualitative terms) the prediction boundaries differ for the different values for bandwidth hyperparameter γ; say which one you think is more likely to overfit.



As γ was decreased from 0.1 to 0.03, the prediction boundaries became tighter around the actual data points on the scatter plot. A higher γ value led to the inclusion of a few more data points in the correct predictions. A lower γ value may have the tendency to overfit, as it contains the data points more tightly, and may stretch to fit in data points a lower γ value may not include.