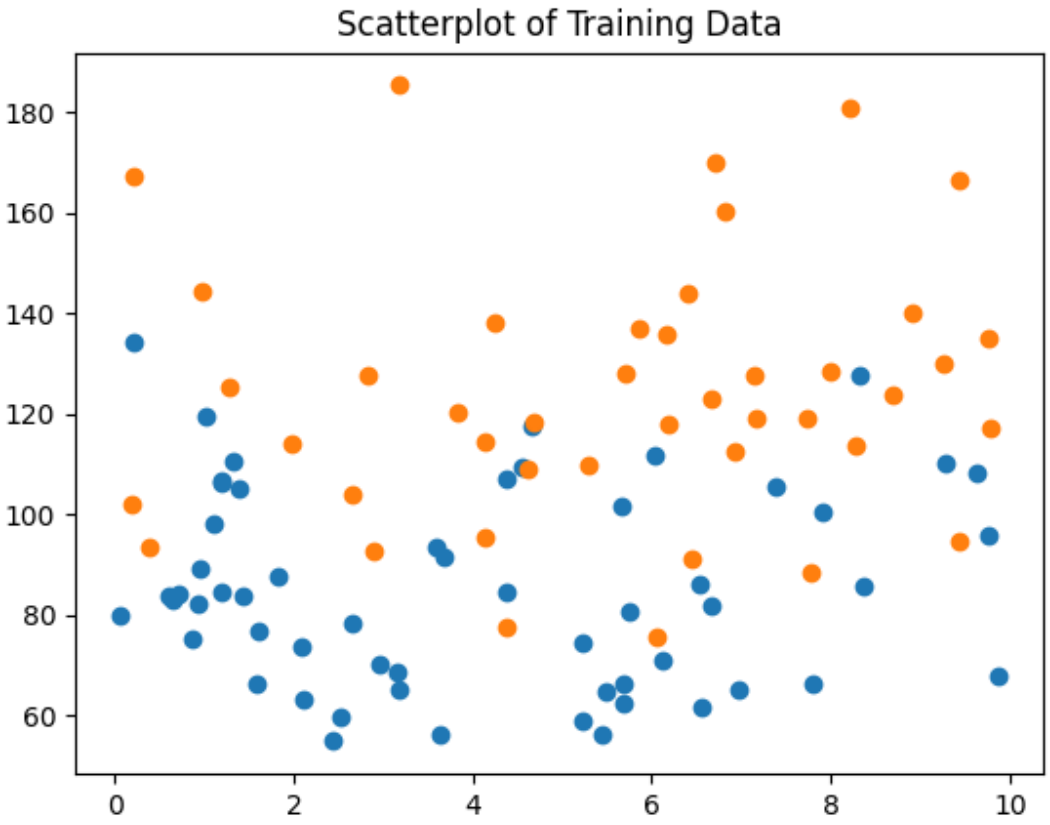
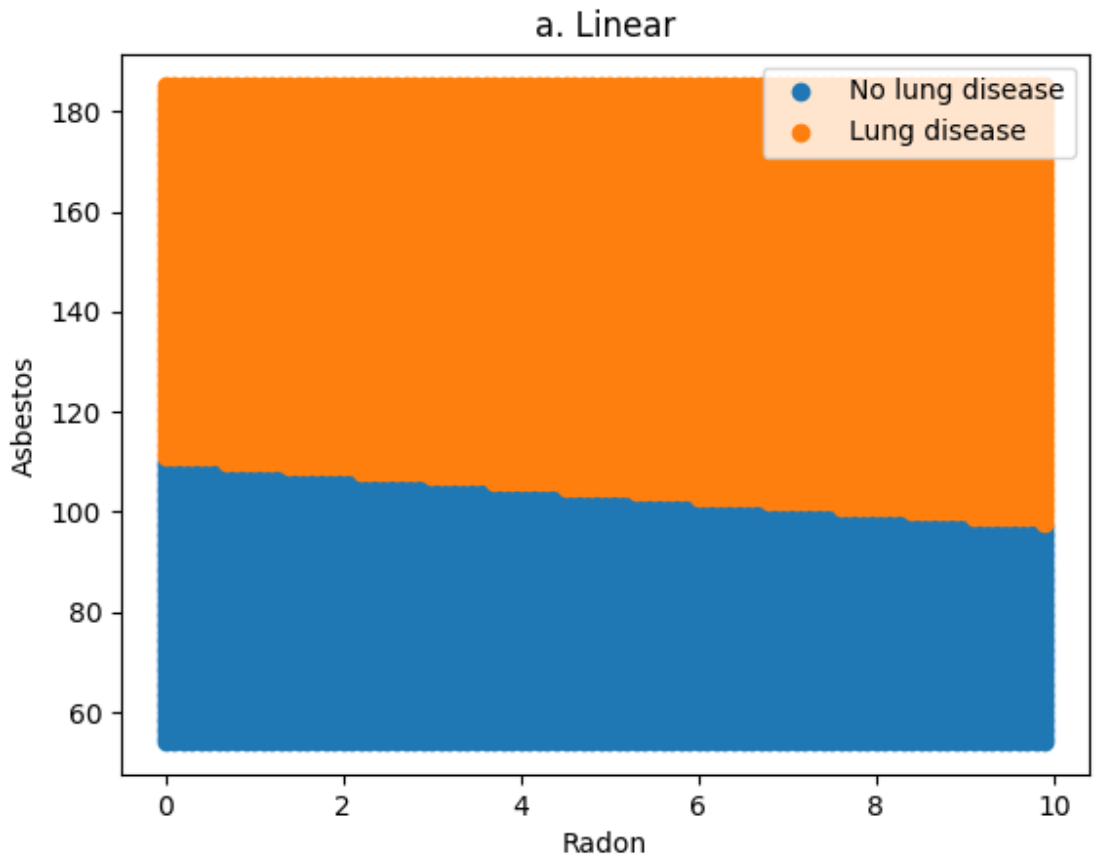
CS 4342 - Machine Learning
Homework - 5 SVMs
By Suryansh Goyal

Training Data

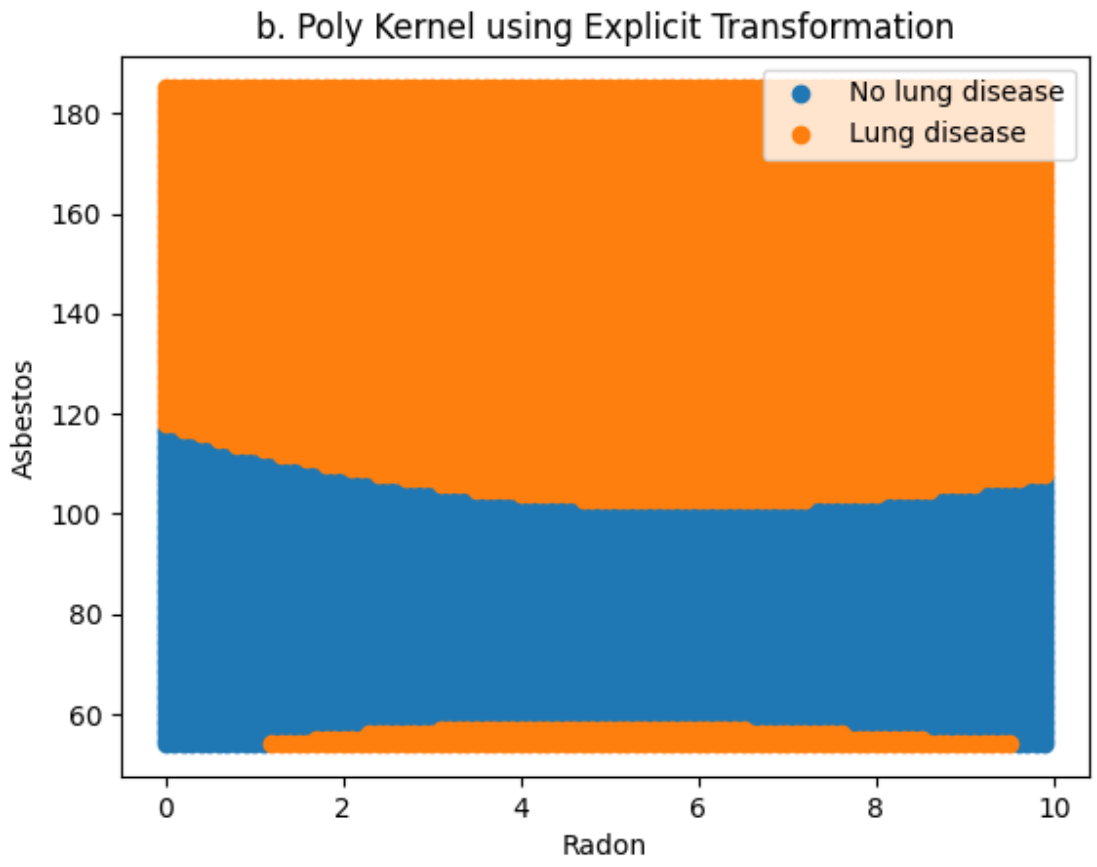


Part A - Linear SVM

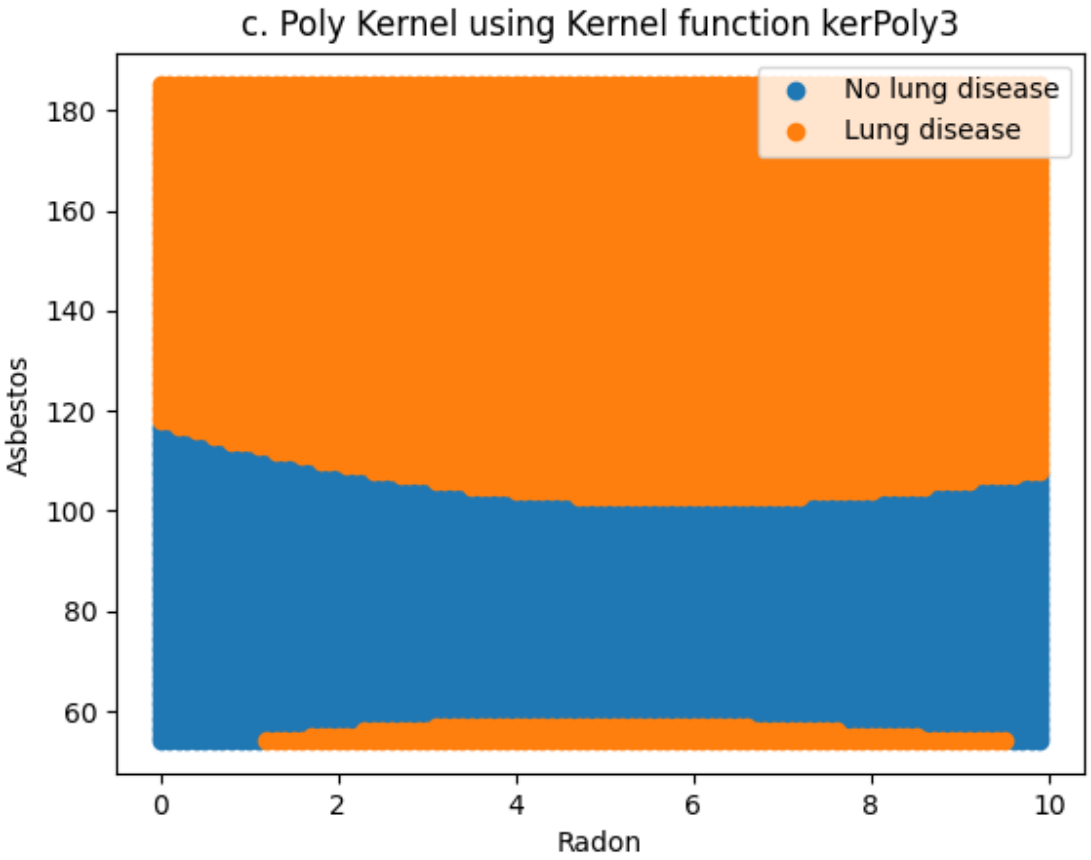


Part B
Poly-SVM using Explicit transformation

Formula for phi(x):
phi([r,a]) = [1, √3*r , √3*a, √6*r*a, √3*r*r, √3*a*a, √3*r*r*a, √3*r*a*a, r*r*r, a*a*a]

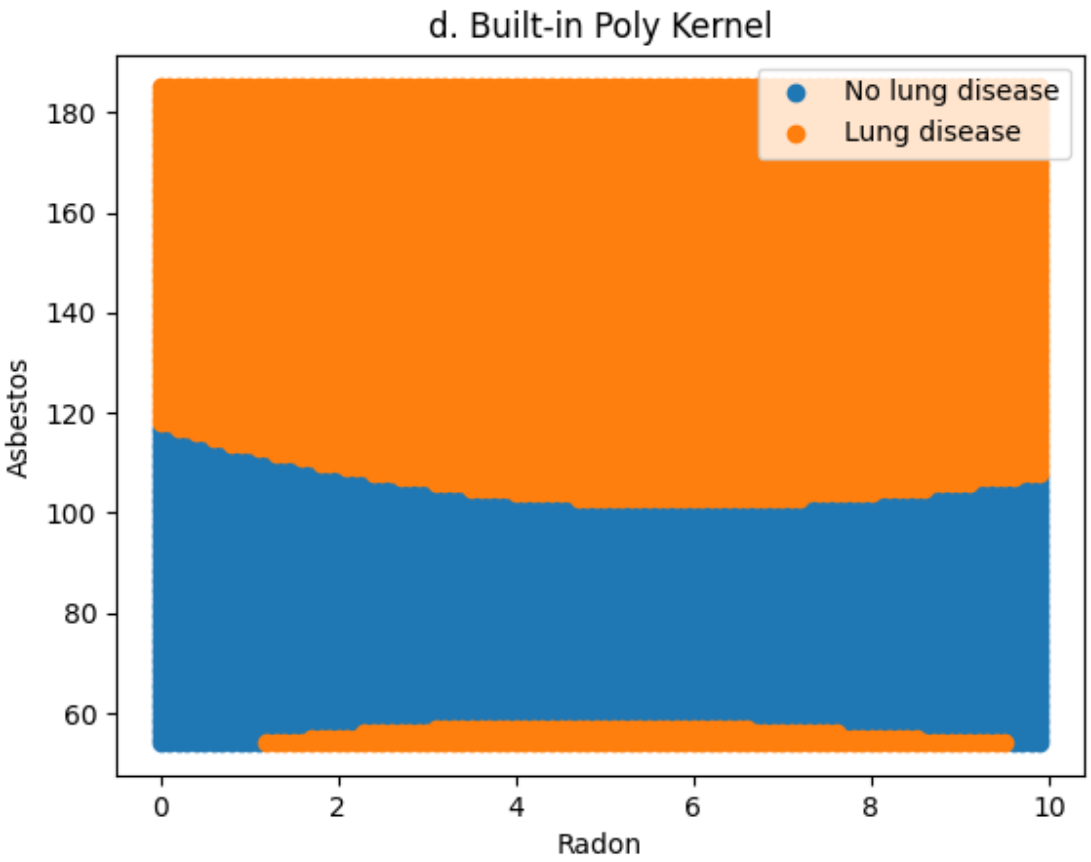


Part C
Poly-SVM using Kernel Trick



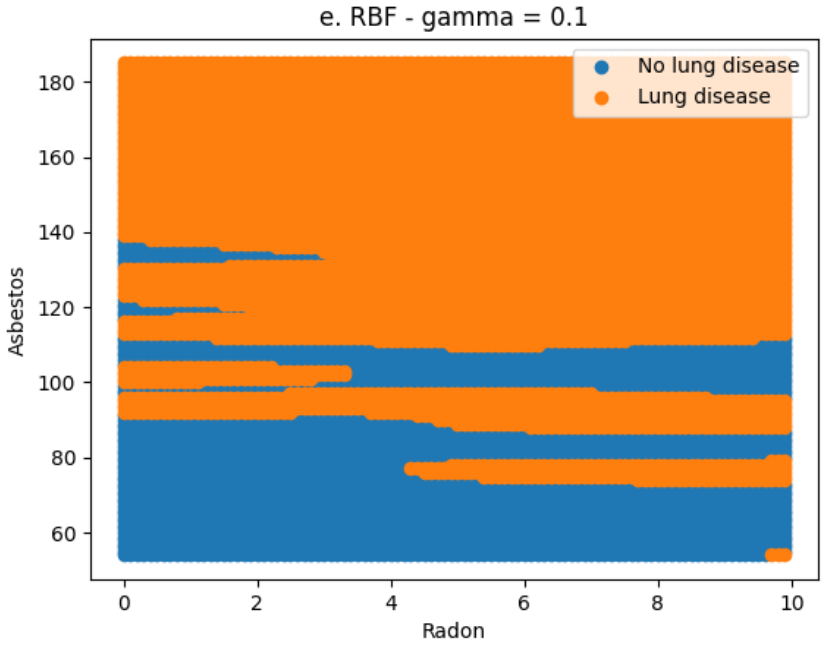
Part D

Poly-SVM using built-in “poly” kernel in scikit-learn

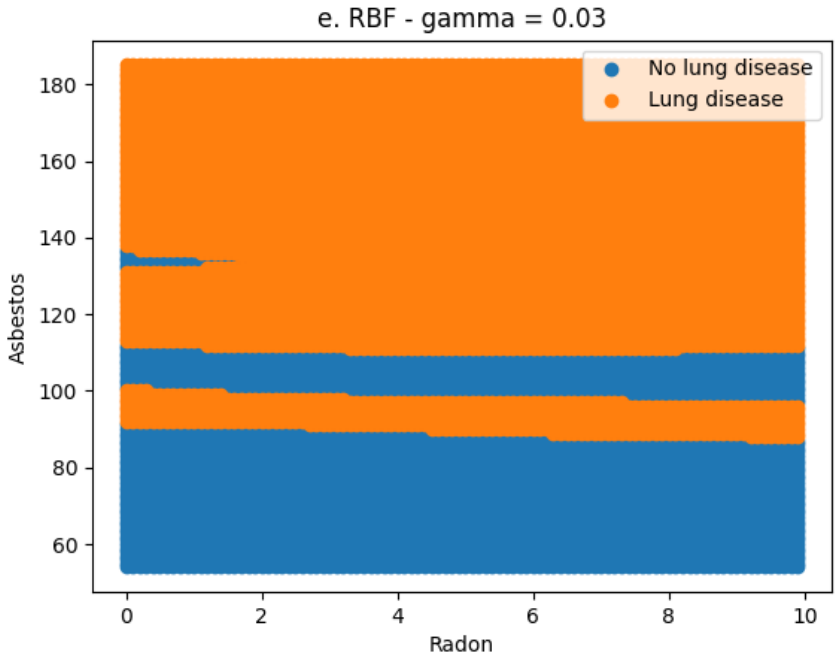


Part E
RBF-SVM using built-in RBF kernel in scikit-learn

1. **Gamma = 0.1**



2. **Gamma = 0.03**

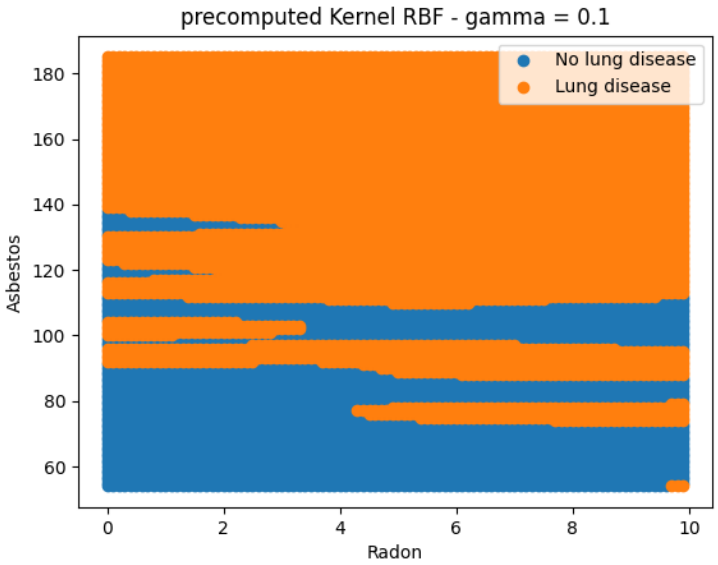


Here, the RBF with Gamma = 0.1 is more likely to overfit the data as it is predicting that exposure to minimal number of asbestos and high number of radons is going to cause

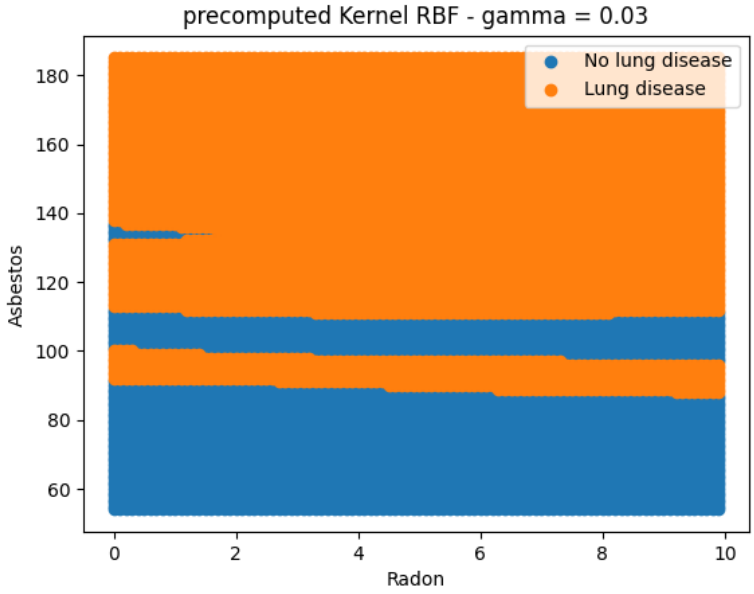
lung disease (which might've been a special case (not depending on asbestos and radons) in the dataset).

RBF-SVM using precomputed RBF kernel

1. **Gamma = 0.1**



2. **Gamma = 0.03**



As you can see, I have managed to get the exact same results for both kinds of RBF kernel (‘precomputed’ as well as ‘rbf’). Here also, we can conclude that gamma = 0.1 is more likely to overfit the data.