Figure 1.A: LIME samples around the boundary to find the delta between blue and red classes. From [10].

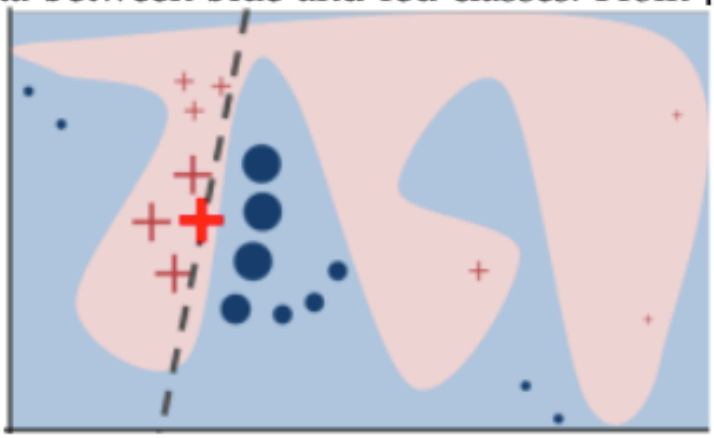

Figure 1.B: Feature importance as assessed by LIME. A **positive** weight means the feature encourages the classifier to predict the instance as a positive and vice versa for the **negative** weight. Larger weights indicate greater feature importance.

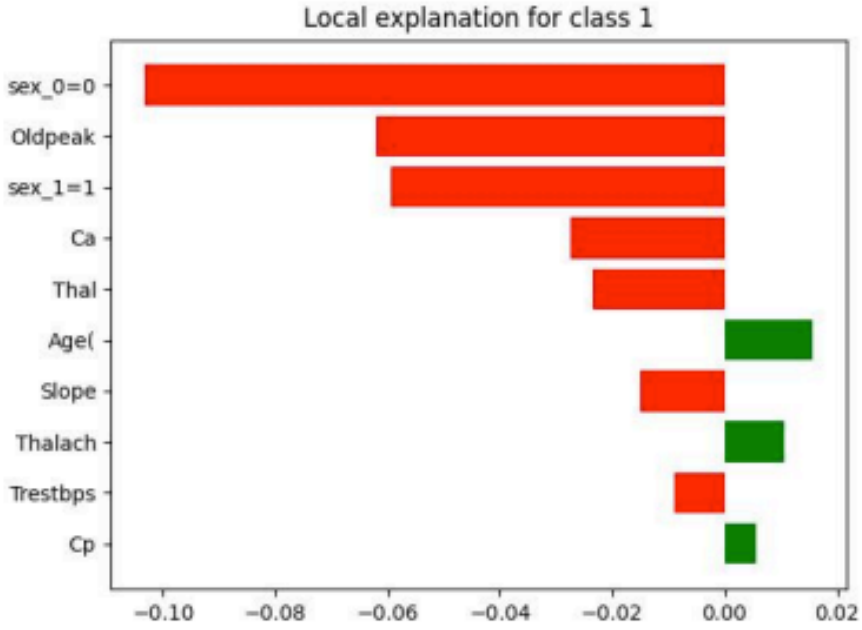

Figure 1.C: Blue/red = original/invented data [5].