Intro To ML – KNN

Amit Zulan – 207299033, Dvir Ben Zikri - 315409508

1. placeholder if needed.
   1. If therefore, .  
      By definition if is C-Lipschitz w.r.t. Euclidian distance then . Since both are in the sample which is sampled from , and has 0 bayes error meaning it’s deterministic, then we know and the same goes for . Therefore, we can deduce , and since . Combining everything together we get:
2. Dvir
3. 1. We can see from the fig. 2 that , since these thresholds would label wrong only 2 apples from the sample .  
      .
      * , the Euclidean distance formula from .
      * will be the ERM prediction rule selected from the above . It achieves 0 empirical error on the sample S, since all blue samples are closer than 6 units from , and all red sample are further away. This is not the only function that achieves 0 empirical error on the sample S. There is a range of thresholds that can achieve this.

Moreover, there is an extended hypothesis class that can also achieve 0 empirical error and might have a better prediction rule than the suggested from above.  
Choosing . This means there are more prediction rules to choose from but could reduce the approximation error. Real life reasons for this could be that the trunk is thicker or healthier in some axis.