Assignment4

Yuxiang Ren

1. The accuracy value of SVM with polynomial kernel is 0.123300257258.

   The accuracy value of SVM with Gaussian kernel is 0.366409408306.

   From the experiments, we can find that when SVM works with different kernels, different performances will outcome. Therefore, this result reminds us when we use SVM to implement classification tasks, it is very important to choose the most suitable kernel.

   In this problem, Gaussian kernel apparently outperforms Polynomial kernel based on the metric Accuracy defined by ourselves. From my perspective, this advantage may be acquired from that Gaussian kernel can map the samples to infinite dimensions. Besides, Polynomial kernel may achieve better performance if we tune the parameters used by it, because when we use this kernel, more parameters need to be set compared to Gaussian kernel, but in fact we didn't do this work.

2. The accuracy value of SVM is 0.957859120271.

   The accuracy value of KNN is 0.952014764688.

   The accuracy value of neural network is 0.887726853276.

   The accuracy value of ensemble majority voting is 0.957551522608.

From the above results, we can find that when combining multiple models with ensemble majority voting, the result may be worse than one of these models. The consequence may come from that different models have equal rights to decide on the final result. Truth may rest with the minority

from time to time. In this problem, SVM has the best performance which means it can make right decision in some samples where the others all commit mistakes. However, once based on the majority voting, the final results of these samples go wrong. Therefore, ensemble learning sometimes may affect the result badly, and which one of Bagging and Boosting should be used depends on different scenarios.

3. The SSE when k is 3 is 630.79715817164106.

The SSE when k is 5 is 426.52332314821922.

The SSE when k is 7 is 339.60392181129384.

With the rise of k, we can find the SSE keep decreasing. I think it's easy to understand, because it is equivalent to continuing to divide in the original cluster. Once we get smaller clusters, the elements inside are closer to the centers, and SSE will decrease. However, we can also find SSE can not be the only metric for measuring performance, because it's not the truth that the cluster smaller, the performance better.