Problem 3 answer:

1. Suggest which value of K you think captures sufficient information from the training dataset.

   I think the K=50 captures sufficient information, it looks has already capture the most information in the eigenvectors. But when k = 300, it will capture the most information.

2. Briefly explain the time vs accuracy trade-off for different values of K that you see.

   % First for a conclusion of my result
   %K=50:
   %-correct male accuracy of 0.521
   %-correct female accuracy of 0.601

   %time 0.010432 秒。

   %K=100:
   %-correct male accuracy of 0.521
   %-correct female accuracy of 0.604

   %time 0.001279 秒。

   %K=200:
   %-correct male accuracy of 0.521
   %-correct female accuracy of 0.603

   %time 0.001001 秒。

   %K=300:
   %-correct male accuracy of 0.521
   %-correct female accuracy of 0.604

   %time 0.000124 秒。

   We can see that by the K increase, the accuracy of male doesn't change a lot but the accuracy of female has a little increase, the running time will decrease a bit. Actually, the accuracy not changes a lot.

3. Argue about the time vs accuracy trade-off, which will be a lot more pronounced in this case.

   In the all result, we can see the result more obvious. As the K increase, the accuracy decrease, and the time also decrease. Therefore, it is more important to choose a suitable K to balance the accuracy and time during the training process.