Thomas Dinh

Oct. 11, 2023

DSE 6211

Week 7

Exercises
Hint for all exercises: see the synchronous live session slides.
1) What is the main difference between unsupervised and supervised learning?

The main difference between unsupervised and supervised learning is that unsupervised does not use labeled data. This means that the model must find patterns and relationships in the data on its own.

2) Is centering required for PCA? Is scaling required for PCA? Explain your answers.

Centering and scaling are required for principal component analysis (PCA). Centering in PCA means subtracting the mean of each feature from all of the data points in the dataset. This ensures that the first principal component is aligned with the direction of maximum variance in the data, rather than being biased towards the mean of the data. Scaling is required for PCA because PCA is sensitive to the scale of the variables. If the variables have different scales, the features with larger scales will dominate the variance in the data, and the features with smaller scales will have less influence on the results of the PCA.

3) After running the code above, run the following code to obtain the PCA feature loadings. Copy and
paste the output. Which feature has the strongest influence on the first PC? Which feature has the
weakest influence on the first PC?

Feature 9 has the strongest influence on the first PC and feature 10 has the weakest influence on the first PC.

```{r}
pca_results$rotation
```

```
              PC1         PC2         PC3          PC4          PC5         PC6          PC7          PC8         PC9          PC10
 [1,]   0.14755942  0.31787120  0.36691173 -0.053292437 -0.64336343  0.52368912 -0.12999741  0.160582316  0.09411120  0.0031769146
 [2,]   0.48345100 -0.15747031 -0.10224937 -0.069634298 -0.02261135 -0.04769628 -0.65939099 -0.502368774  0.19030493 -0.0011527840
 [3,]  -0.10972402 -0.53176322  0.10016800 -0.385746152 -0.14791704 -0.06738729 -0.35865707  0.404631451 -0.45843092 -0.1319616758
 [4,]   0.11404381 -0.03647767  0.64842518  0.226436690  0.16082744  0.01206961  0.11982285 -0.415615073 -0.54847809  0.0013573994
 [5,]   0.08338397 -0.26409215  0.59370312  0.135078876  0.20287679 -0.23309579 -0.04680126  0.322321784  0.59245111 -0.0003246952
 [6,]   0.13452438  0.38324428  0.16119543 -0.524350159 -0.24112641 -0.66790212  0.15186852 -0.081211800 -0.01411429 -0.0024810807
 [7,]  -0.44154485  0.36402425  0.06143469  0.291417692  0.01257464 -0.19452123 -0.43643606 -0.005877226 -0.01603962 -0.5958700582
 [8,]   0.39349691  0.34505447 -0.09972823  0.361697053  0.14614613 -0.19870902 -0.26697536  0.489181523 -0.29935954  0.3525515653
 [9,]   0.58642704 -0.03461611 -0.11744627  0.008254542  0.08846697  0.08244647  0.30067375  0.173736228 -0.05116813 -0.7093745921
[10,]  -0.02660746  0.34736566  0.13583471 -0.532194283  0.63967426  0.36638771 -0.16703416  0.073296414  0.01583939  0.0007391287
```