ICS635 Homework 4: by Lambert Leong

Problem 1

Maximizing the conditional likelihood of the data.

$$\begin{aligned}\mathcal{L}(\mathcal{D}) = p_\theta(y_n|x_n) &= \prod_n^N (\hat{y_n})^{y_n}(1-\hat{y_n})^{1-y_n} \\ &= \sum_n^N (\hat{y_n})log(y_n) + (1-y_n)log(1-\hat{y_n})\end{aligned}$$

Problem 2

1. Sample size is large enough to do a 60%, 20%, 20% split for train, validation, and test.
2. The exact architecture or neural net model is ambiguous.
3. The loss function being optimized is not clear.
4. The exact hyper parameters optimized are not stated.
5. The evaluation method for the competition is unclear(i.e. are they using AUROC, AUPRC, or etc).
6. It seems as though the test set was not used. The individual does not report how the tuned model did on the test set.
7. It is unclear if the individual is first place on the leader board of if they acutally won the whole competition.
8. The final optimized parameters are not reported which makes it difficult to reproduce the work.
9. It is not clear what data set was used to train the final model that was submitted. Did the individual use just the model trained on the 80% training data or all the training data?
10. The final test set, mentioned in the last line, is ambiguous because it could be referring to the witheld dataset that is used to pick the winners or the final test set(10% of the training) that the individual set aside when he did his data split.

Problem 3

1. Centering data affects the principal components if it is calculated via singular value decomposition. However, if one is to compute the principal components by first computing the covariance matrix then centering should not affect prinicipal components.
2. Scaling data can help prevent one principal component explaining the majority of the variance. If the data is not scaled properly, one feature could dominate and might influence the axis of maximal variance.

3.

4. If the network is free of any nolinearities then it is convex. It will be e uivalent to plotting thd data in principal component space.

Problem 4

1.

2.

3.