PART A

Q.no.1.

If a model is performing great on the training data but generalizes poorly to new instances, then it is overfitting to the training data. One solution could be training on more data so that it can effectively capture the overall pattern and not just of a localized region and the other solution could be regularization that controls the model's complexity by adding a penalty for higher terms and by possibly reducing the number of parameters used to make predictions.

Q.no.2

The value of N1 = 5 (Error needs to be found for each group one at a time and there are 5 groups)
The value of N2 = 160 (4 groups form the train set every single time and 4 * 40 = 160)
The value of N3 = 40 (1 group form the test set every single time, hence size 40)

Q.no.3

In all the answers below, a vaccine being effective is taken as positive and a vaccine being not effective is taken as negative.

a. Accuracy = $\frac{TP + TN}{TP + TN + FP + FN}$ = (970 + 15) / (970 + 15 + 10 + 25) = 0.9657 (96.57%)

b. Baseline Accuracy = $\frac{TP + TN}{TP + TN + FP + FN}$ = (995) / (1020) = 0.9754 (97.54%)

c. While comparing their accuracies, the classifier is found to be less useful than the majority-class base-line because it has less accuracy than blindly predicting all the classes as falling under the majority class (accuracy of majority-baseline class). However, a model might not want to be judged on the basis of accuracy alone especially when the classes are unbalanced. Some models might want to focus on precision or recall depending on its use cases so we need to compare these values as well to confirm if this model is completely less useful than the majority-class baseline or not.

d. Precision = $\frac{TP}{TP + FP}$ = (970) / (970 + 10) = 0.9898 (98.98%)

e. Recall = $\frac{TP}{TP + FN}$ = (970) / (970 + 25) = 0.9749 (97.49%)

f. F1 Score = $\frac{2 * Precision * Recall}{Precision + Recall}$ = (2 * 0.9898 * 0.9749) / (0.9898 + 0.9749) = 0.9823 (98.23%)