Q1

Micro-averaging and macro-averaging were the two methods that were used as for evaluation metrics of the implemented classifier.

Figure 1

```
train on train (train data used as test data):
accuracy:  0.7513661202185792
Macro-average Precision: 0.7528072273237285 Macro-average Recall: 0.7241518491180735 F1-score: 0.7382015580408111
Micro-average Precision: 0.7513661202185792 Micro-average Recall: 0.7513661202185792 F1-score: 0.7513661202185792


test:
accuracy:  0.7321428571428571
Macro-average Precision: 0.7059502923976608 Macro-average Recall: 0.7271031746031745 F1-score: 0.7163706177532024
Micro-average Precision: 0.7321428571428571 Micro-average Recall: 0.7321428571428571 F1-score: 0.7321428571428571
```

(Results before mean imputation)

Macro-averages focus on whether the classifier has classified each class correctly, while regarding the proportion of the classes in the instances. On the other hand, Micro-averages focuses on the overall performance of the classifier and does not pay much attention to the specifics of each class. In some cases where an equal amount of classes are in a dataset, micro-averages may be a sufficient method of providing a general insight on the performance of the classifier. However, in the pose dataset given, there are class imbalances where some classes have more instances than others in 'train.csv'. Hence taking the macro-average is useful as it operates as an indicator to show whether the classifier is correctly working in all classes. Also, for this classifier, since each instance is classified into a unique class, the accuracy is equal to the values of the micro-average metrics.

Q5

For this question, an additional mean imputation was performed in the preprocessing of the dataset to incorporate information about the missing values into the classifier. Three things were considered when doing the imputation. Firstly, since the numeric are implied to be Gaussian Distributed, it may be a safe assumption to allow a mean imputation on the dataset without skewing it. Secondly, since Naïve Bayes classification assumes independence between the attributes of a class, imputing the mean of the attributes will not affect the other attributes in that class. Hence the dataset may not become skewed.

Finally, considering the positioning of the poses, some poses inevitably have body parts constantly hidden from view. A raw observation of the dataset led to conclude that this may be true, where some attributes of a class have more missing values than other attributes of the same class (that body part is unseeable). From this, an assumption can be made that although hidden, the body parts may be in similar locations in every instance of a pose. In order to find out where the hidden body part is approximately, the mean of the non-missing values in that pose was used. The combination of the means of the attributes are extremely unique to each pose and could now act as a strong indicator when classifying the instances. Hence the accuracy of the Naïve Bayes classifier was expected to increase, which is reflected in the results below.

Figure 2

```
train on train (train data used as test data):
accuracy:  0.8989071038251366
Macro-average Precision: 0.9011067359560316 Macro-average Recall: 0.9016762383725847 F1-score: 0.9013913972108842
Micro-average Precision: 0.8989071038251366 Micro-average Recall: 0.8989071038251366 F1-score: 0.8989071038251366


test:
accuracy:  0.7678571428571429
Macro-average Precision: 0.7567136784783843 Macro-average Recall: 0.7523412698412699 F1-score: 0.7545211397459243
Micro-average Precision: 0.7678571428571429 Micro-average Recall: 0.7678571428571429 F1-score: 0.7678571428571429
```

(Results after mean imputation)