```
                    main.py – loads the arff file
                       into a python object
```

1

2

3

4

partition.py – takes the data object as input and returns the same data partitioned into 10 parts

preprocess.py – takes the data as input and returns the same data preprocessed

classify.py – takes the data as an input and returns a list of classifications

evaluate.py – evaluator object that calculates the accuracy measures for each fold and stores the averages

Notes:

- preprocess.py replaces all missing values, and then picks a single attribute on which to perform entropy-based discretization on. It discretizes all other numeric attributes by separating the data into two evenly-sized bins
- the training and test sets are created by removing the first partition as the test set, joining the other 9 to form the training set, and then appending the test set back to the end of the partition list