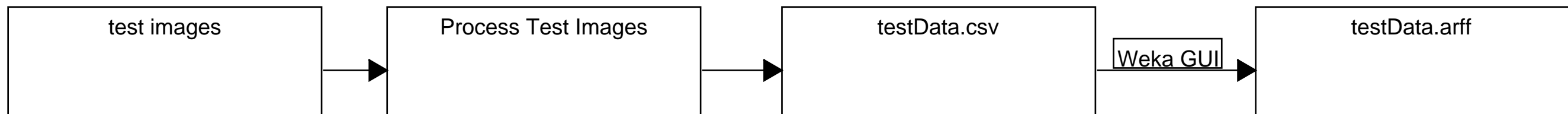
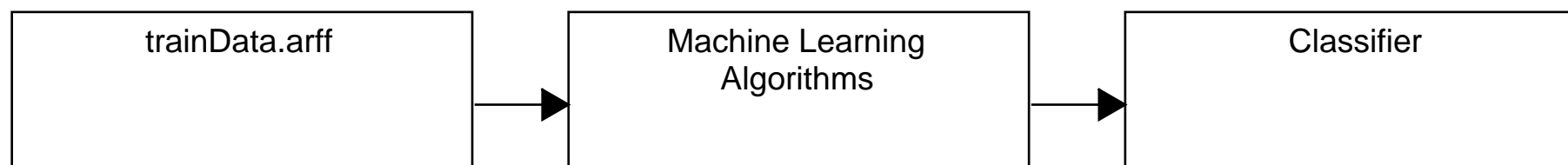


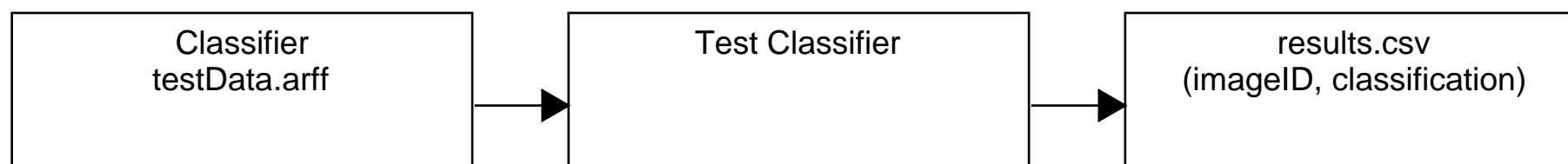
- Train images are in .jpeg format, though other formats are supported
- Train image labels is a .csv file with the columns (imageID, classification)
- Process Train Images converts each image into a comma-separated list of numerical "features" that represent that image
- trainData.csv is a file that contains the results of all of the image processing
- Using the Weka GUI, the trainData.csv file is converted to .arff format, which Weka understands when reading in data



- The above process is repeated for the test images
- The only difference is that no labels for the test data are provided, and each image is given a default label of -1



- The trainData.arff file, which contains all of the data generated from the train images, is read into memory and used to train a set of machine learning algorithms which Weka provides
- This produces a Classifier that can be used to provide a classification for an image that is not yet been labeled



- The Classifier is used to label all of the images data stored in the testData.arff file
- The results.csv file contains all of the images in the testData.arff file, and has two columns: imageID and the label that was calculated by the classifier