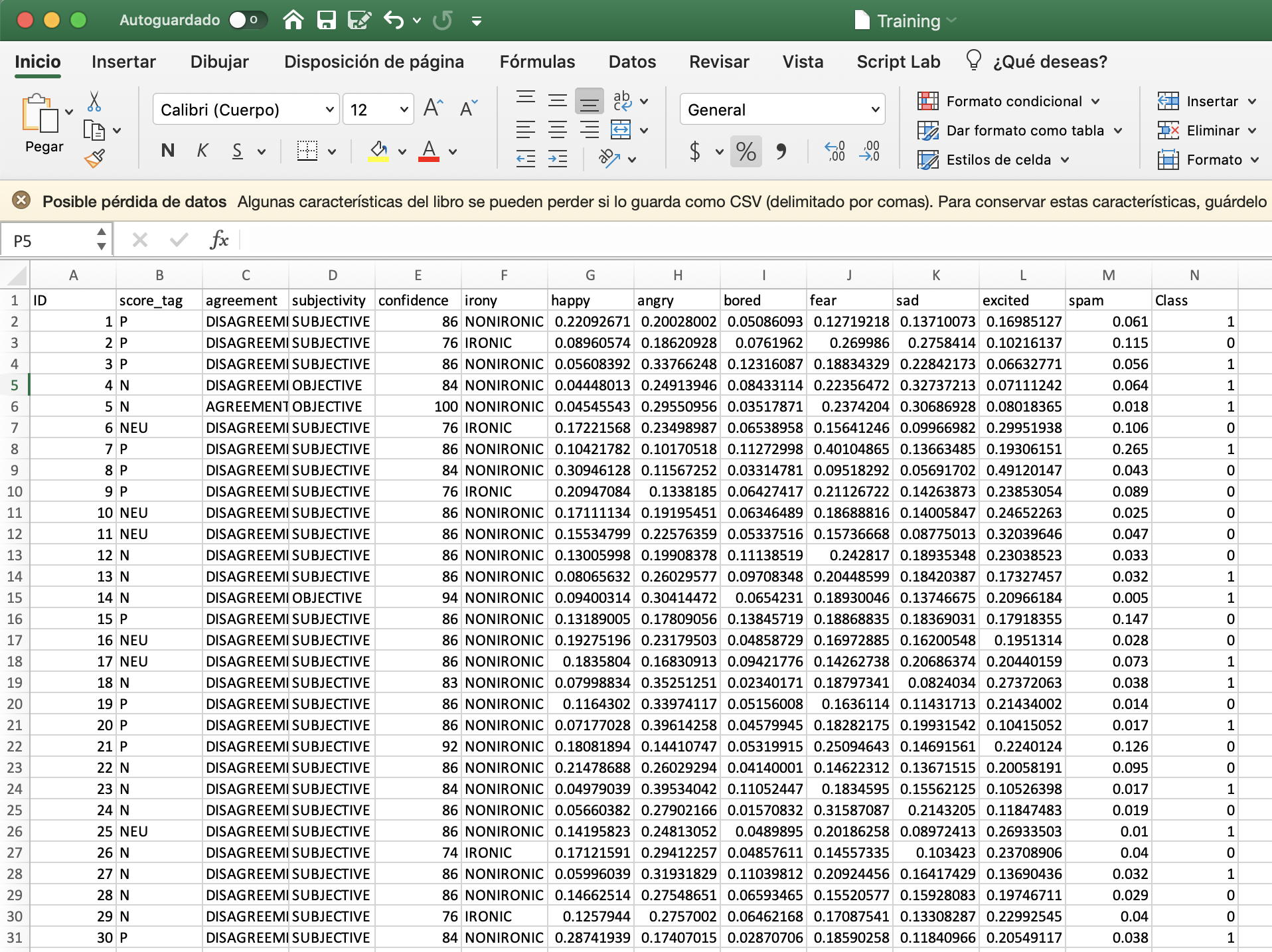
# H6 – Testing the bagging and adaboost methods

**Introduction**

Bagging and boosting methods are useful for dealing with noisy objects in machine learning projects. In WEKA, both methods are available and are going to be used in this homework for solving the Fake News problem.

**Development & Results**

First of all, a new feature named spam was extracted for the Testing and Training datasets using the Intent Analysis API of Parallel Dots.

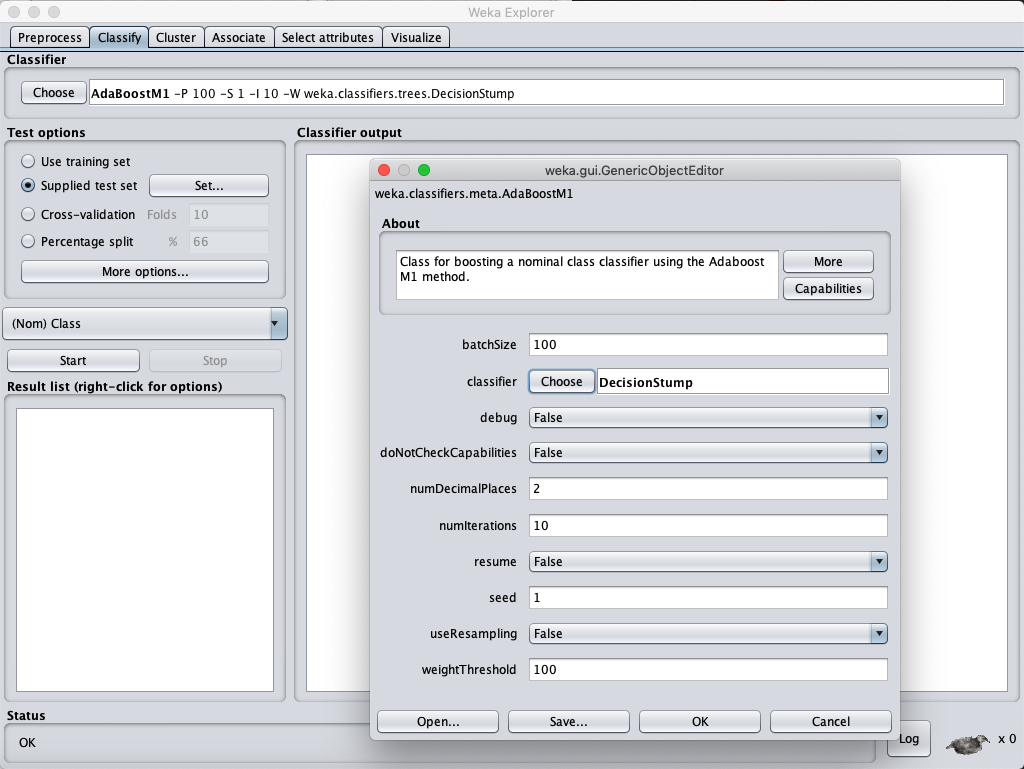


*Figure 1. Trainig.csv with the new spam feature.*

Then both Training.csv and Testing.csv were transformed into WEKA files.

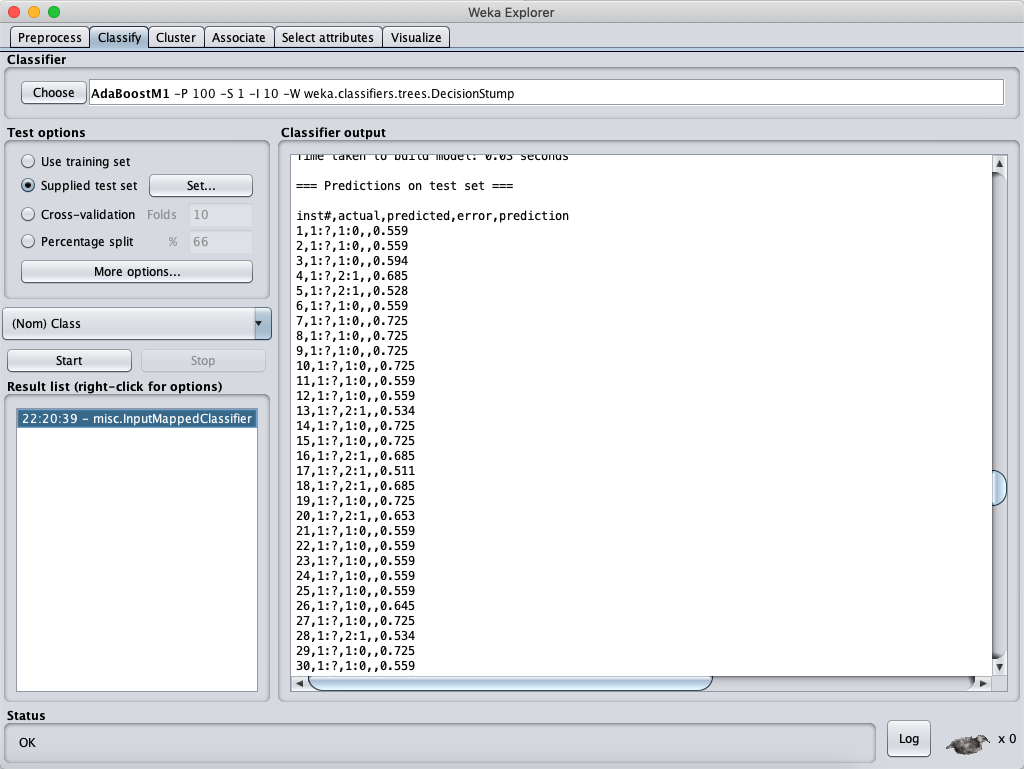
*AdaBoostM1 with DecisionStump*

For this method, in the Classify tab of WEKA, AdaBoostM1 is selected and then DecisionStump is picked as the classifier.



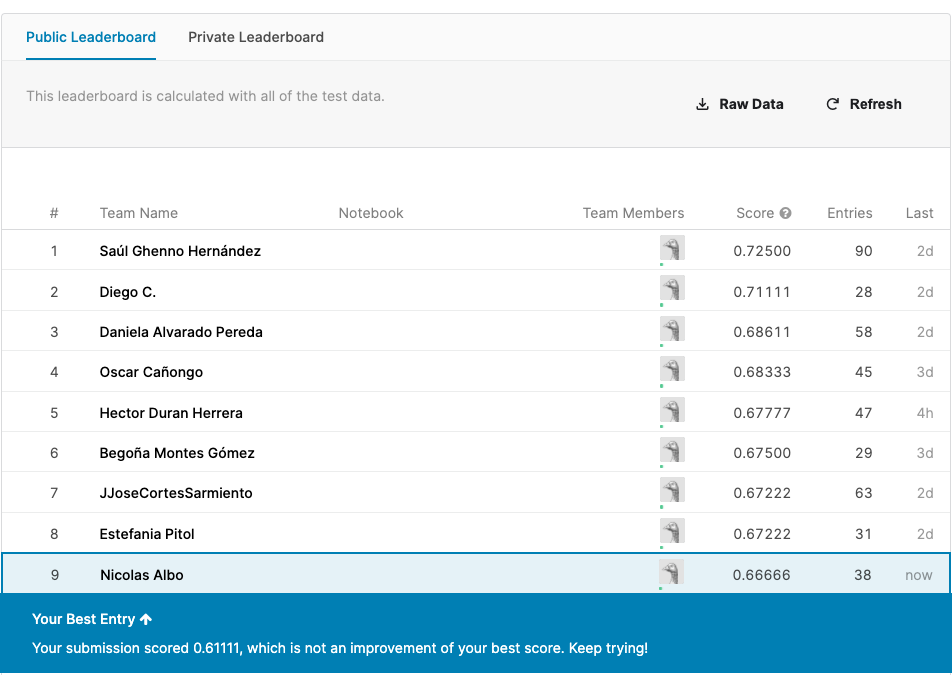
*Figure 2. AdaBoostM1 and DecisionStump in WEKA.*

The following results are obtained.



*Figure 3. Results of using AdaBoostM1 and DecisionStump in CSV format.*

Then, data is filtered and uploaded to Kaggle.

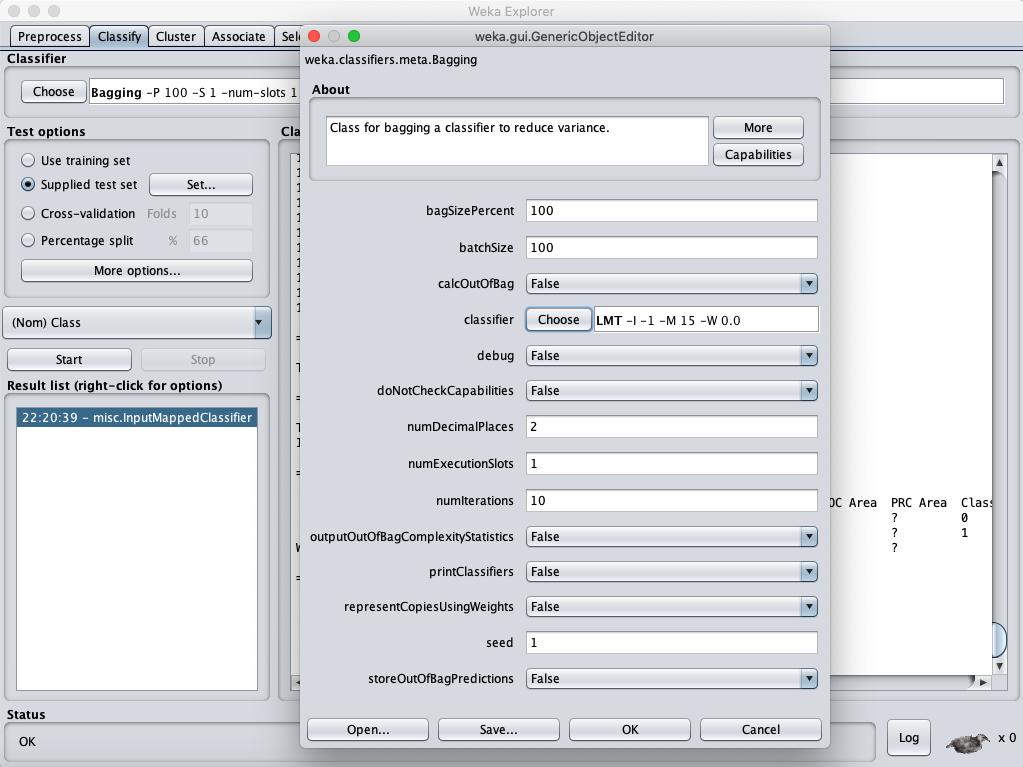


*Figure 4. Score in Kaggle obtained by the AdaBoostM1 method with DesicionStump classfier.*

Although it got a score higher than 0.60000, it couldn’t beat my Kaggle’s best entry, so maybe another method with another classifier will beat my high score.

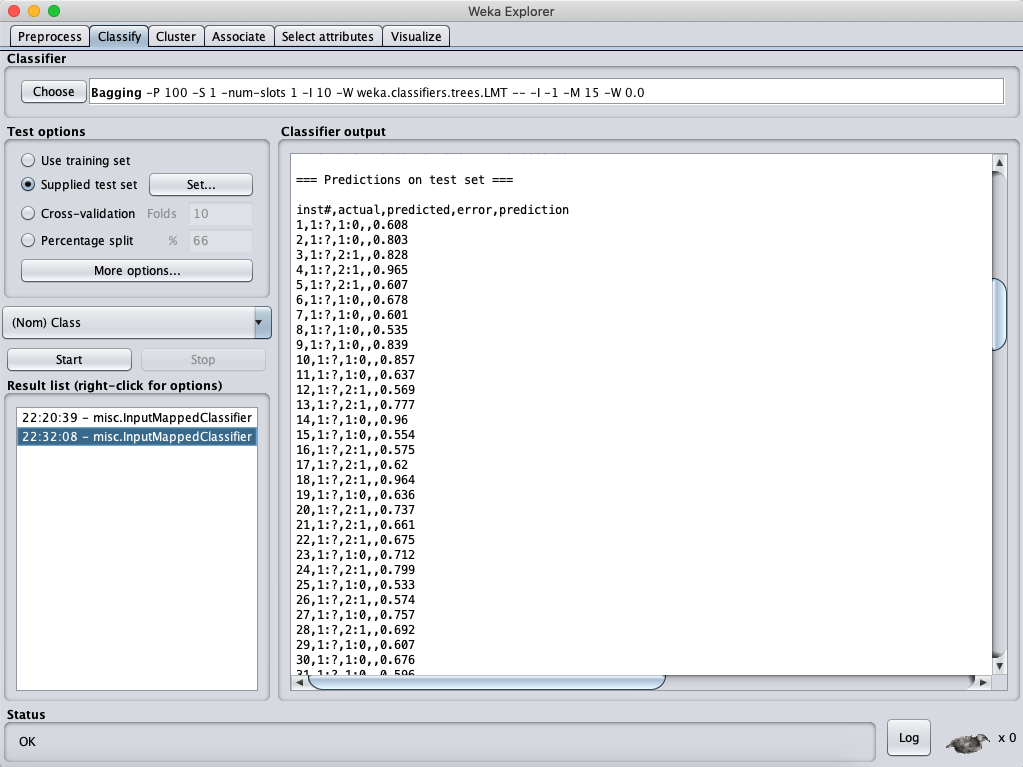
*Bagging with LMT*

For this method, in the Classify tab of WEKA, Bagging is selected and then LMTis picked as the classifier.

**

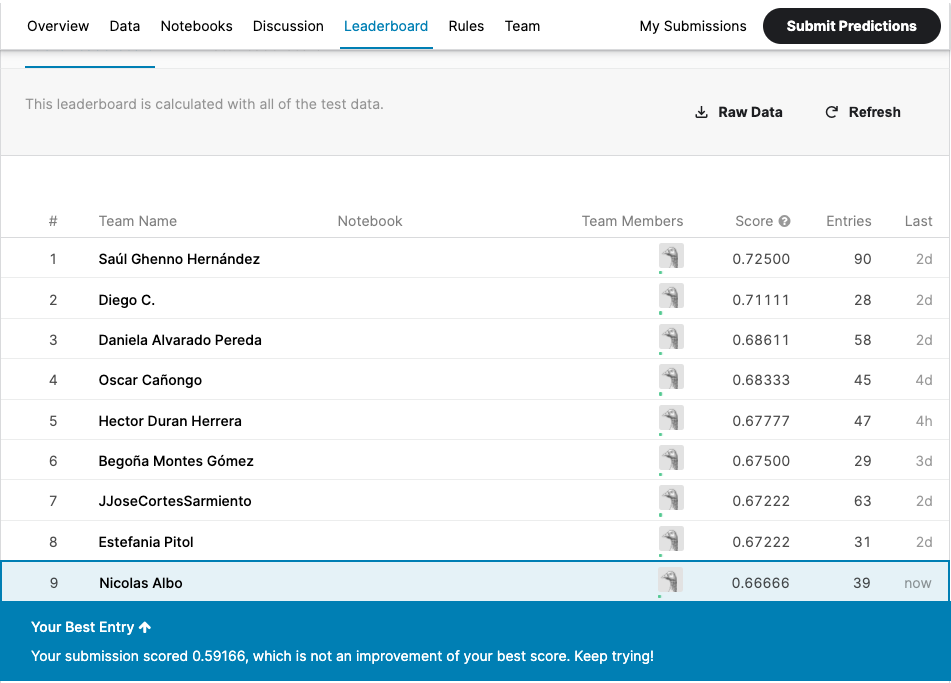
*Figure 5. Bagging and LMT in WEKA.*

The following results are obtained.

**

*Figure 6. Results of using Bagging and LMT in CSV format.*

Then, data is filtered and uploaded to Kaggle.

**

*Figure 7. Score in Kaggle obtained by the Bagging method with LMT classfier.*

In this case, the result could not even break the 0.60000 barrier and it was quite far from my best entry.

**Conclusion**

AdaBoostM1 method with DesicionStump classifier resulted more effective than Bagging method with LMT classifier. Maybe for this specific problem of detecting fake news, AdaBoostM1 is more effective than Bagging, but also the classfiers ifluence in the results.