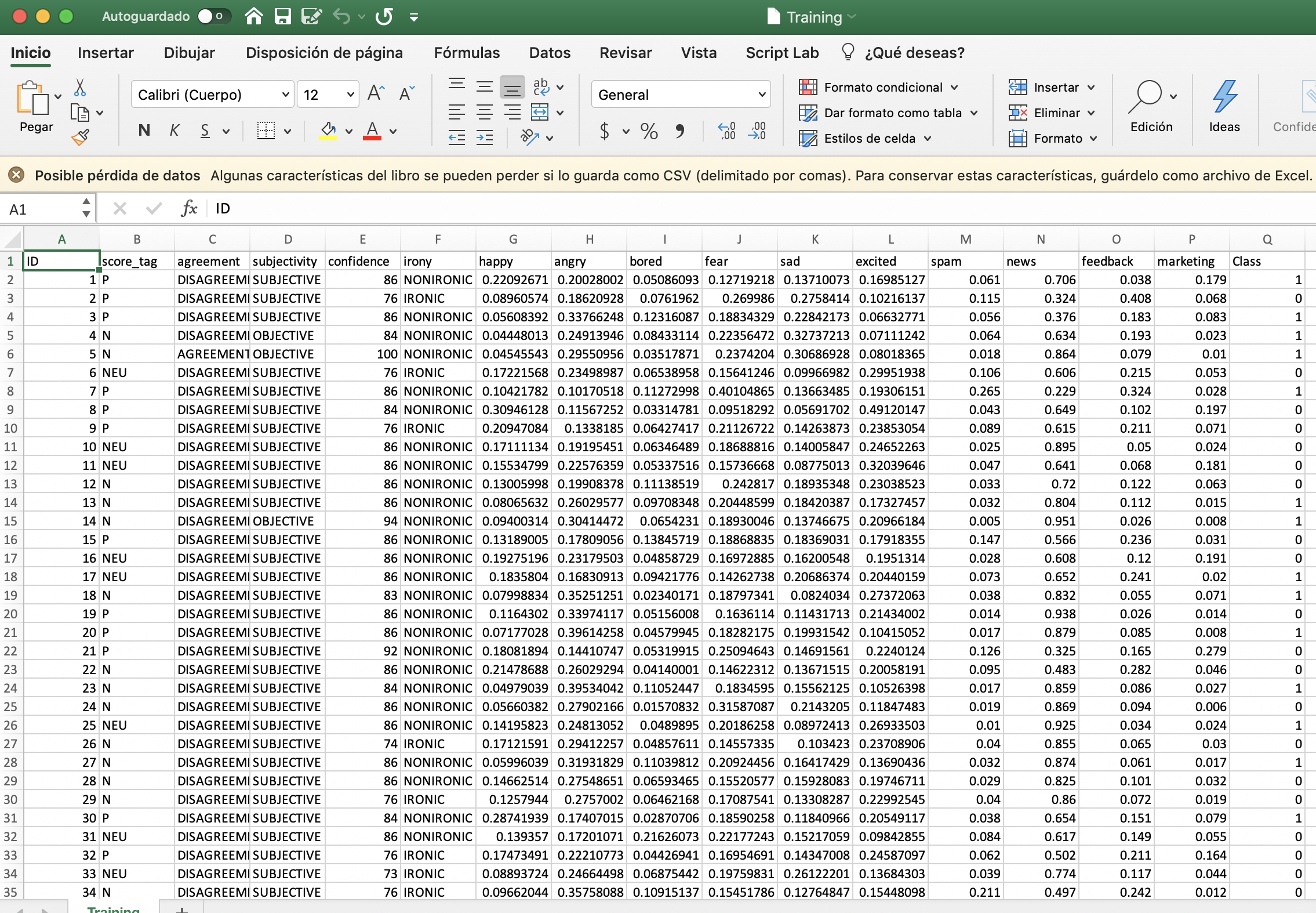
H9 – Testing a selected classifier

**Introduction**

In this homework, two classifiers of the type Logistic Regression were used: Linear and Multinomial. Both classifiers were used for checking which news are fake and which isn’t. With the help of WEKA, this task was easy to realize.

**Development & Results**

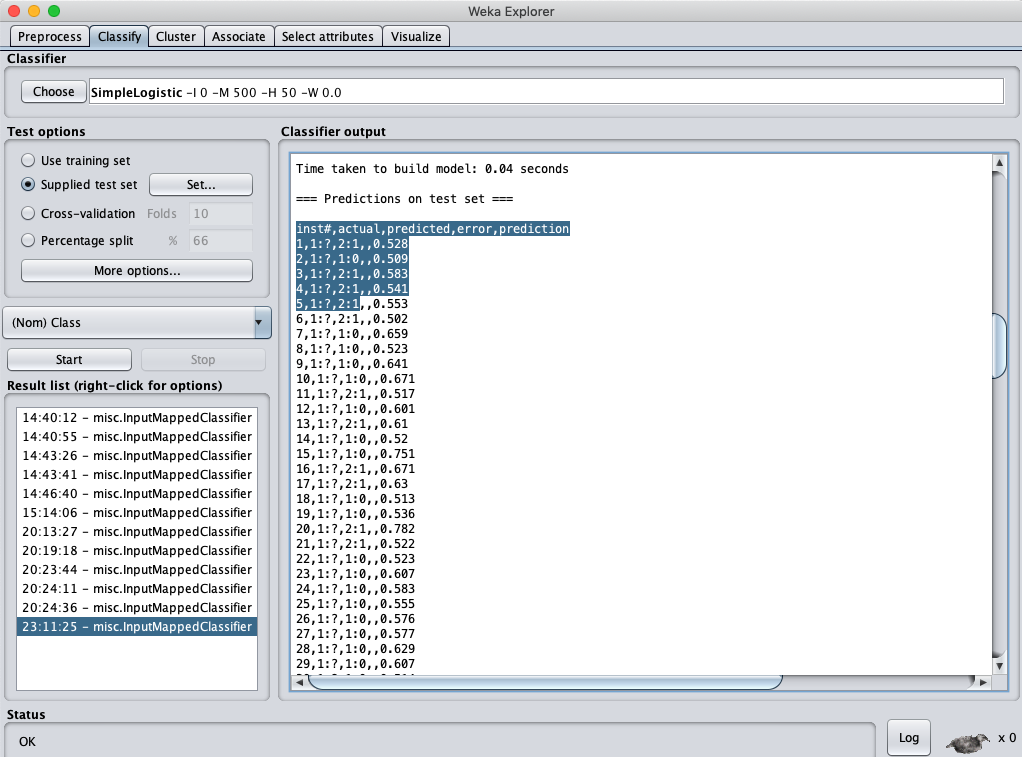
First of all, a new feature was extracted using the intent API of Parallel Dots, this new feature named Marketing was extracted for both the Training and Testing datasets.

****

*Figure 1. Training.csv with the new marketing feature.*

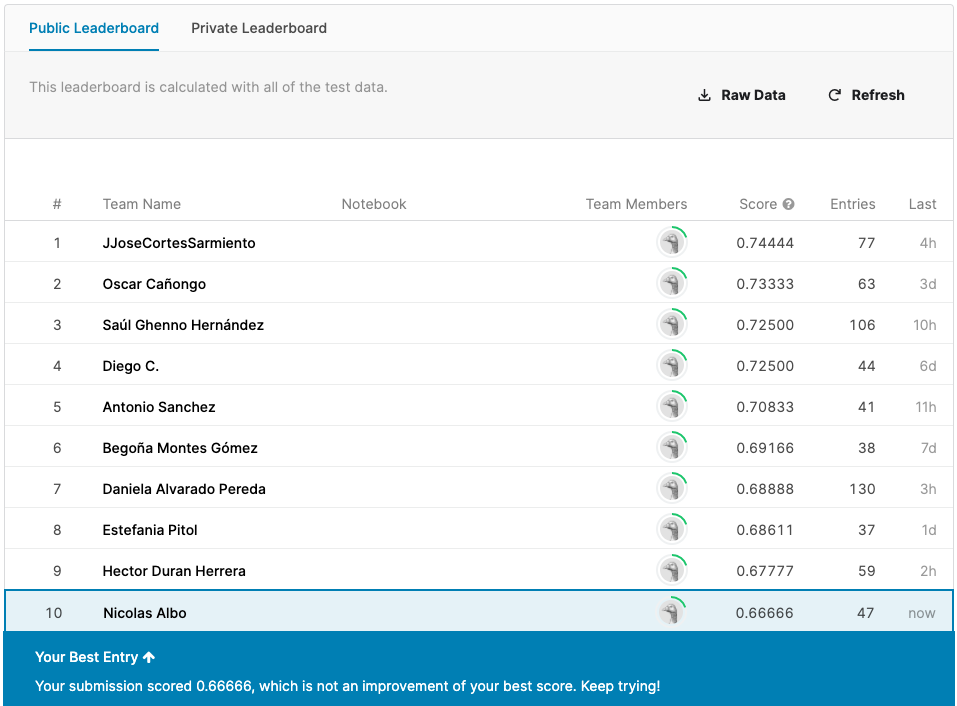
*Linear Logistic Regression*

The first classifier used in this homework was a Linear Logistic Regression Classifier. For using this classifier, the option *SimpleLogistic* was selected in WEKA. Then, the following results were obtained.

**

*Figure 2. WEKA results obtained from SimpleLogistic classifier.*

After that, by using the Bash Script named *Filter.sh*, we cut the unnecessary columns and obtain the format necessary for uploading our results to Kaggle.

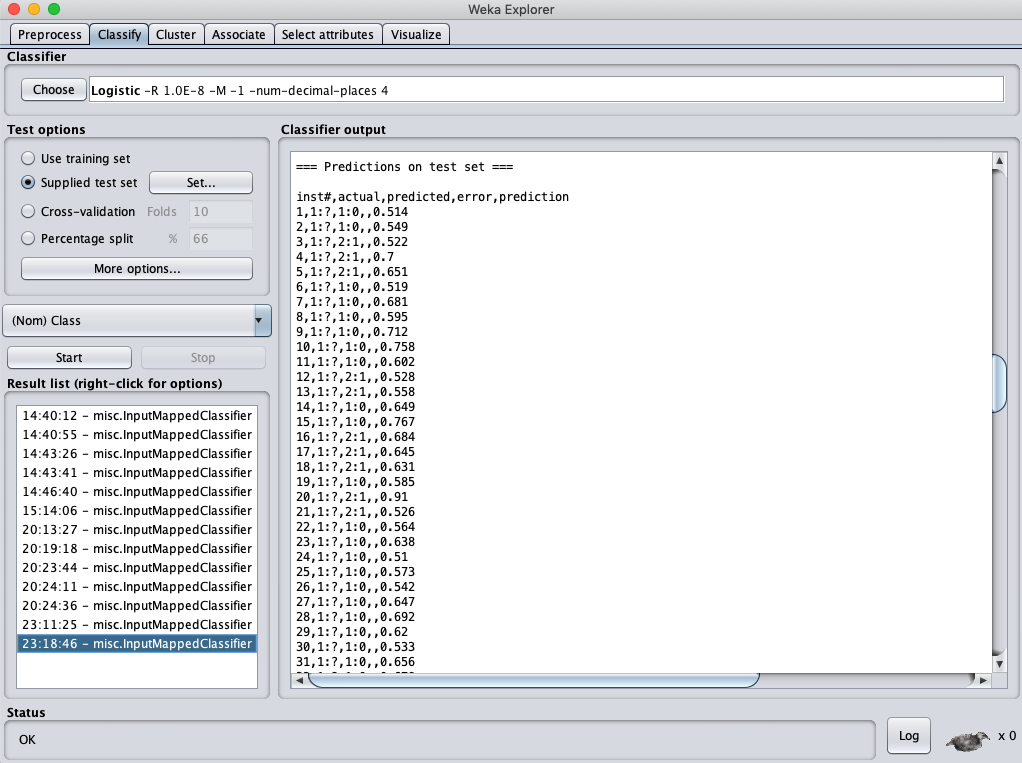
**

*Figure 3. Score obtained in Kaggle by the SimpleLogistic classifier.*

As it can be appreciated, the score obtained was good because it was the same score as the best score, therefore, it can be concluded that the classifier performed well.

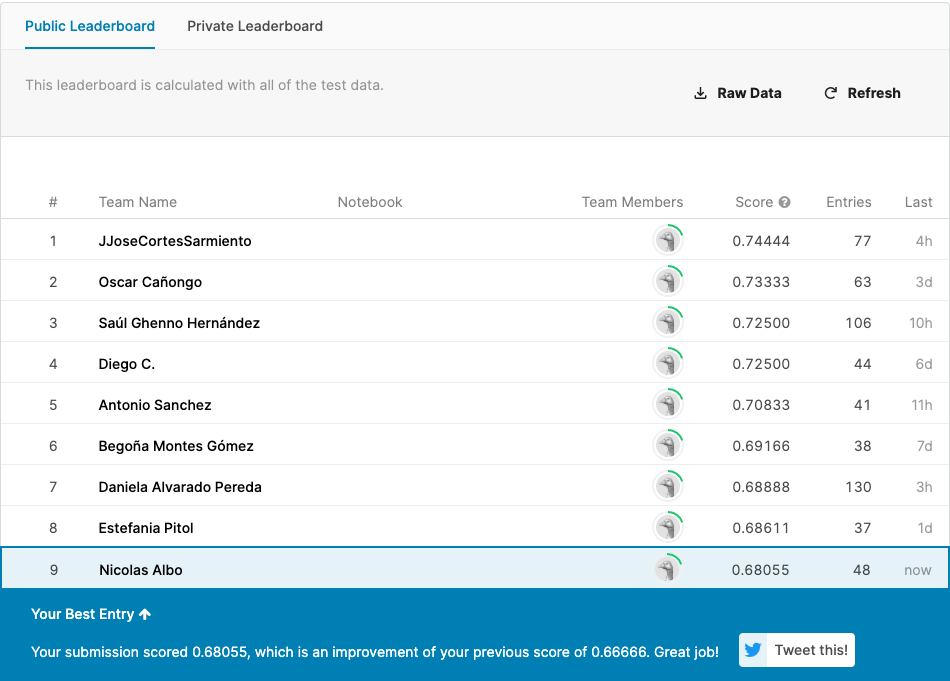
*Multinomial Logistic Regression*

The second classifier used was a Multinomial Logistic Regression Classifier. For using this classifier, the option *Logistic* was selected in WEKA. The following results were obtained.

**

*Figure 4. WEKA results obtained from Logistic classifier.*

After that, by using the Bash Script named *Filter.sh*, we cut the unnecessary columns and obtain the format necessary for uploading our results to Kaggle.

**

*Figure 5. Score obtained in Kaggle by the Logistic classifier.*

As it can be appreciated, this classifier performed better than the previous classifier used, but not only that, it also performed better than my best score, setting a new best score.

**Conclusion**

Logistic classifiers may be quite simple if compared with other type of classifiers like Decision Trees or Bayesian Nets, but for this specific problem, they obtained better results than the other type of classifiers mentioned above.