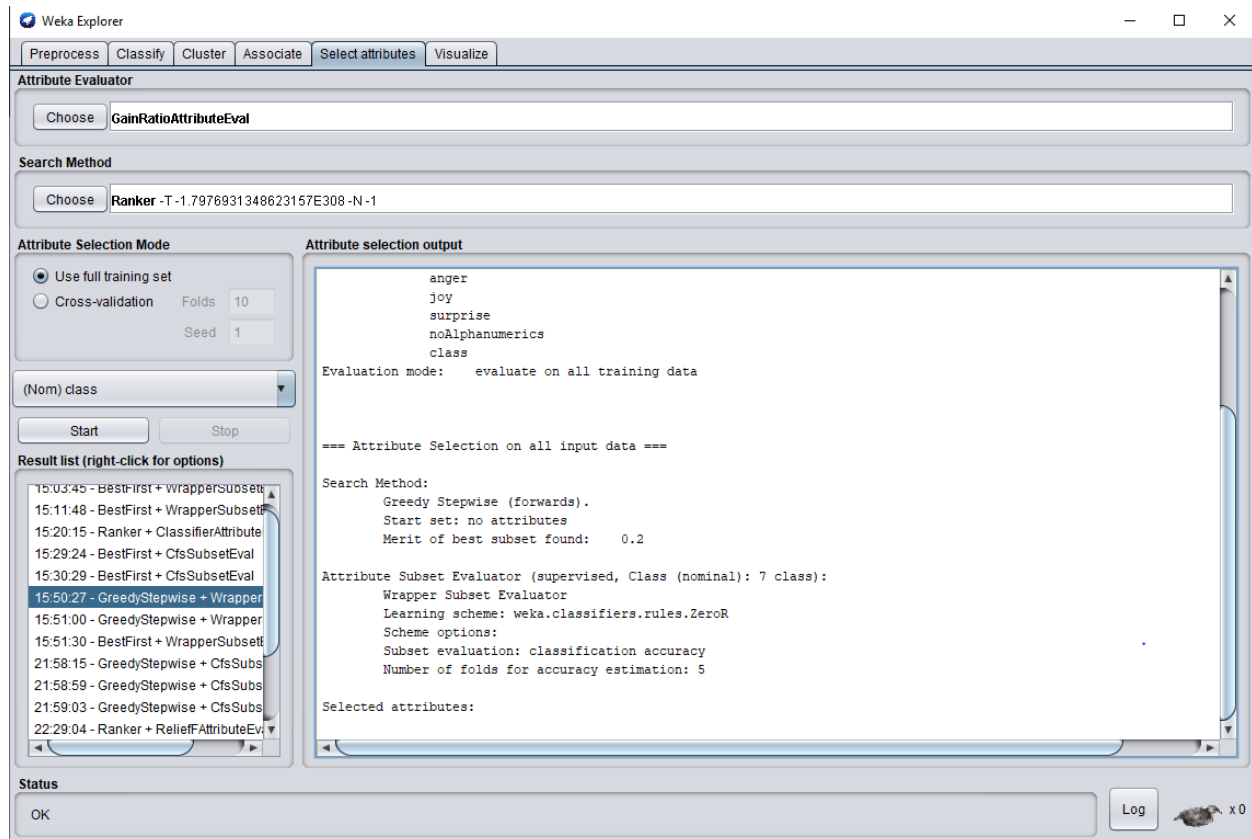


Reporte

Luis Daniel Medina Cazarez A01651070

For these submissions were use different select attributes methods, but most of them gave the same set of data. Some selector just gave an empty set, this is the case of greedy stepwise with wrapper subset eval:



Because most of the models gave the same attributes sets, for the first two submission were used the same attributes, but different classifications models. For the first submission was used misc input mapped classifier. With this I got 22.0% of accuracy:

Tarea81.csv 8 hours ago by Daniel add submission details	0.22000	<input type="checkbox"/>
--	---------	--------------------------

For the second one was used Decision Stump and I got an accuracy of 20.8%

Tarea82.csv an hour ago by Daniel add submission details	0.20800	<input type="checkbox"/>
--	---------	--------------------------

For the last attribute selection was use a Ranker with Gain Ratio Attribute.

Attribute Selection Mode

☒ Use full training set
 ☐ Cross-validation

Folds 10
 Seed 1

(Nom) class

Start

Stop

Result list (right-click for options)

15:50:27 - GreedyStepwise + wrapper
 15:51:00 - GreedyStepwise + Wrapper
 15:51:30 - BestFirst + WrapperSubse
 21:58:15 - GreedyStepwise + CfsSubs
 21:58:59 - GreedyStepwise + CfsSubs
 21:59:03 - GreedyStepwise + CfsSubs
 22:29:04 - Ranker + ReliefAttributeEv
 22:29:29 - Ranker + ReliefAttributeEv
 22:29:49 - GreedyStepwise + CfsSubs
 22:30:38 - BestFirst + ClassifierSubse
 22:30:48 - GreedyStepwise + Classifie
 22:31:03 - Ranker + GainRatioAttribute

Attribute selection output

Search Method:
 Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 15 class):
 Gain Ratio feature evaluator

Ranked attributes:

0.1552	7	joy
0.0935	1	score_tag
0.0785	6	anger
0.0718	10	surprise
0.0623	4	confidence
0.0568	2	agreement
0.046	14	misspelled
0.0366	11	noAlphanumerics
0.0358	13	word_number
0.0337	9	sadness
0.026	5	irony
0.0238	12	negation_words
0.0159	3	subjectivity
0	8	fear

Selected attributes: 7,1,6,10,4,2,14,11,13,9,5,12,3,8 : 14

For this, I deleted the last 3 attribute on the list (fear, subjectivity and negation_words). For the classification process was used a Random tree. The result was a little bit better than the previous ones, But still wasn't able to beat the submissions without the last 3 new attributes.

Tarea83.csv

an hour ago by Daniel

[add submission details](#)

0.22400



Conclusion:

None of the new attributes seems to improve any model, If we visualize the data seems like there is no distinction between classes, probably it is better choosing different attributes.

