Carlos Augusto Amador Manilla
A01329447

Testing two pattern based classifiers

J48 Decision Tree

A j48 which is an implementation of a C4.5 decision tree was tested with different parameters.

j48simple10FKaggle.csv 2 days ago by Carlos A Applied J48 decision tree: default settings, 5 sentiment, 5 emotions	0.56231
j45simpleKaggle.csv 2 days ago by Carlos A Applied J48 decision tree: default settings, 5 sentiment, 5 emotions, 4 personality	0.60185

The result was better when the personality features were added to the data. With this in mind, I continue to test Random Forest which uses J48 decision trees.

Random forest

Most of Random Forest tests were performed using 5 sentiment, 5 emotion and 4 personality features. The runs were made changing parameters, adding more iterations, and changing the batch percentage and number of features picked to achieve diversity.

RFAllF1000IS2BP25-14F.csv 2 minutes ago by Carlos A Applied Random Forest: BSP 25, Seed 2, BS 100, MaxDepth unlimited, numFeatures All, 1000 iterations, seed 1: 5 sentiment 5 emotion 4 personality	0.61816
RFAllF1000IS2BP15-14F.csv 4 minutes ago by Carlos A Applied Random Forest: BSP 15, Seed 2, BS 100, MaxDepth unlimited, numFeatures All, 1000 iterations, seed 1: 5 sentiment 5 emotion 4 personality	0.61411
RFAllF1000IS2-14.csv 8 minutes ago by Carlos A Applied Random Forest: BSP 20, Seed 2, BS 100, MaxDepth unlimited, numFeatures All, 1000 iterations, seed 1: 5 sentiment 5 emotion 4 personality	0.62690
RF5F10PS10000I-14F.csv 2 days ago by Carlos A add submission details	0.61035
RFAllF1000I-14FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 25, BS 100, MaxDepth unlimited, numFeatures All, 1000 iterations, seed 1: 5 sentiment 5 emotion 4 personality	0.61835

Carlos Augusto Amador Manilla
A01329447

RF5F1000I-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures 5, 1000 iterations, seed 1: 5 sentiment 5 emotion	0.60324
RF1S100BS3NF-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures 3, 100 iterations, seed 1: 5 sentiment 5 emotion	0.60221
RF1S100BS2NF-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures 2, 100 iterations, seed 1: 5 sentiment 5 emotion	0.60991
RF1S100BS1NF-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures 1, 100 iterations, seed 1: 5 sentiment 5 emotion	0.60004
RF1S100BS5NF-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures 5, 100 iterations, seed 1: 5 sentiment 5 emotion	0.60221
RF1S100BS0NF-10FKaggle.csv 2 days ago by Carlos A Applied Random Forest: BSP 100, BS 100, MaxDepth unlimited, numFeatures all, 100 iterations, seed 1: 5 sentiment 5 emotion	0.59995

The best run had the following configuration: bagSize 20%, unlimited depth, all the features used in the trees and 1000 iterations. I think the reason is the bagSize enables diversity, however I expected better accuracy.

Carlos Augusto Amador Manilla
A01329447

FURIA

The Fuzzy Unordered Rule Induction Algorithm, tested with different parameters. However the results were not better than the ones obtained by RandomForest.

FURIA5Fo2Op-14F.csv 5 hours ago by Carlos A add submission details	0.58783
FURIA5Fo2Op-14F.csv 5 hours ago by Carlos A add submission details	Error
FURIA10Fo5Op-14F.csv 5 hours ago by Carlos A Applied FURIA: Product Tnorm, Rule stretching 10 folds, 2 optimization, seed 1: 5 sentiment, 5 emotions, 4 personality	0.60964
FURIA-14F.csv 5 hours ago by Carlos A Applied FURIA: Product Tnorm, Rule stretching 3 folds, 2 optimization, seed 1: 5 sentiment, 5 emotions, 4 personality	0.61171