**ACTIVITAT 1**

**Analitza el conjunt de dades segment-challenge.arff amb Weka, i després de**

**provar 3 models troba el millor per poder fer la predicció.**

**Genera un informe.docx mostrant comentant els resultats dels 3 models i**

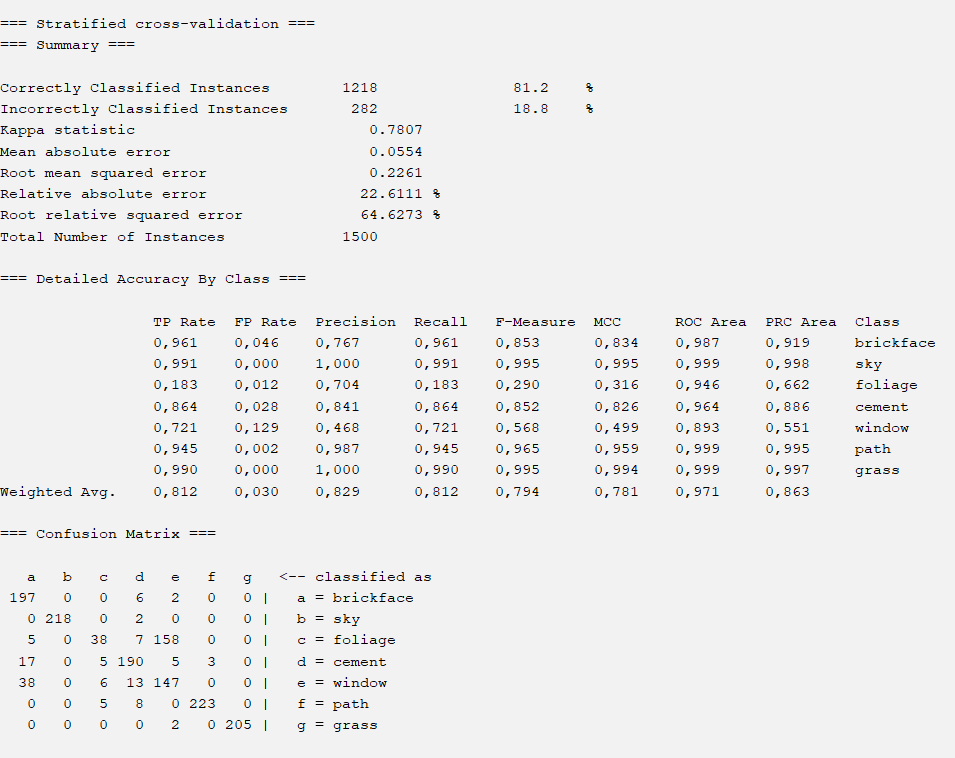
**justificant quin et sembla millor. Ha d’incloure algun gràfic que et sembli**

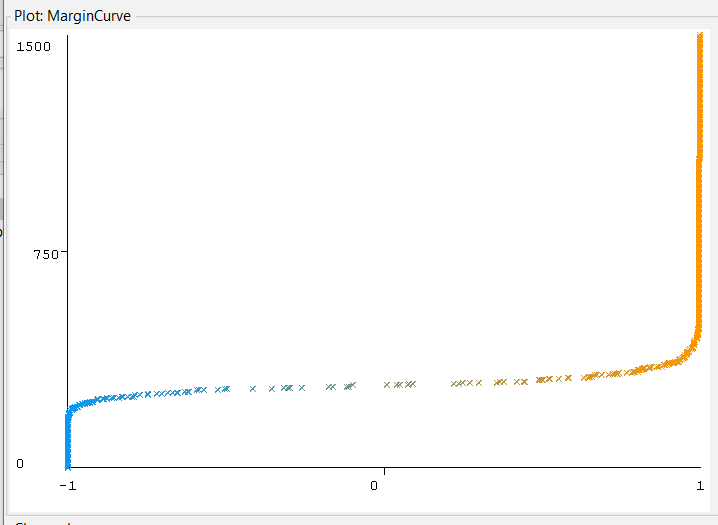
**significatiu.**

Els 3 models predictius seleccionats per fer la comparació han sigut el Naive Bayes, Random Tree i Multilayer Percepton.

* Naive Bayes resultats

Test mode: 100-fold cross-validation

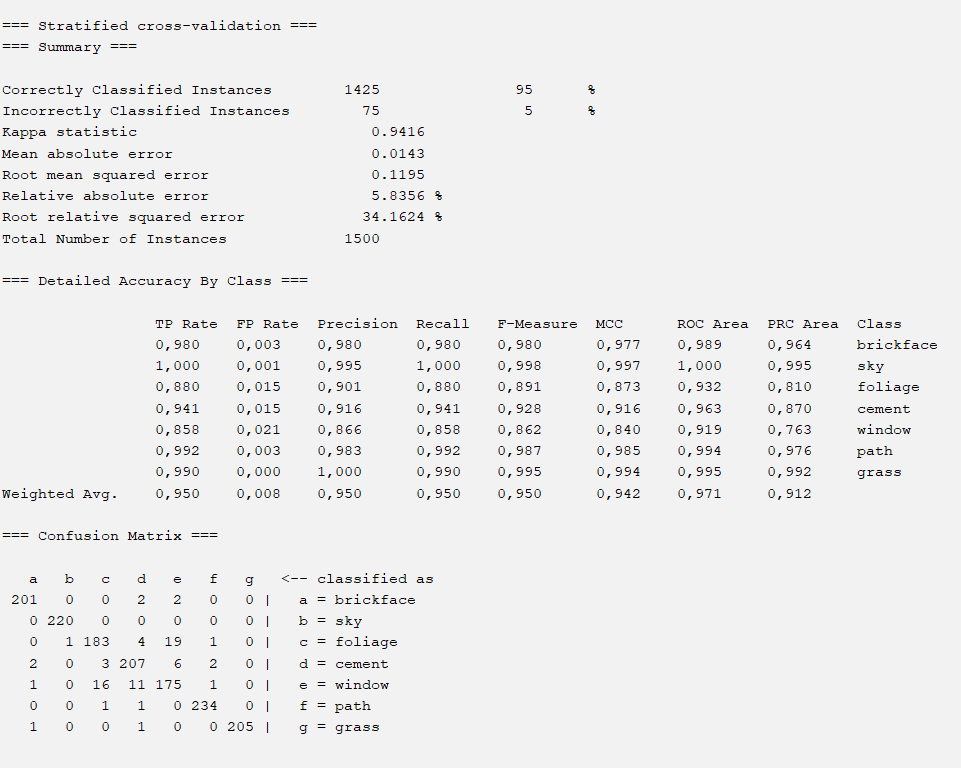


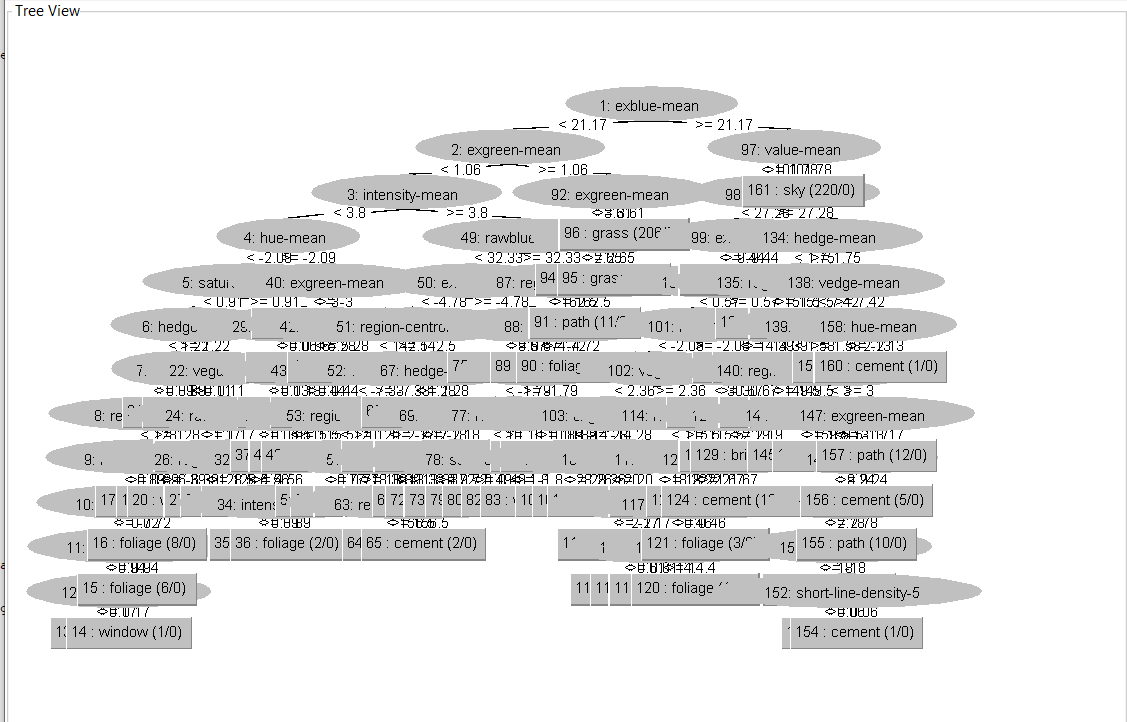


D'aquest model podem destacar que encara que el cost computacional ha sigut mínim, fixar-se quasi instantani el marge d'error és d'un 18.8% sent aquest molt elevat si el comparem amb els altres dos models que segueixen a continuació.

* Random Tree resultats

Test mode: 100-fold cross-validation

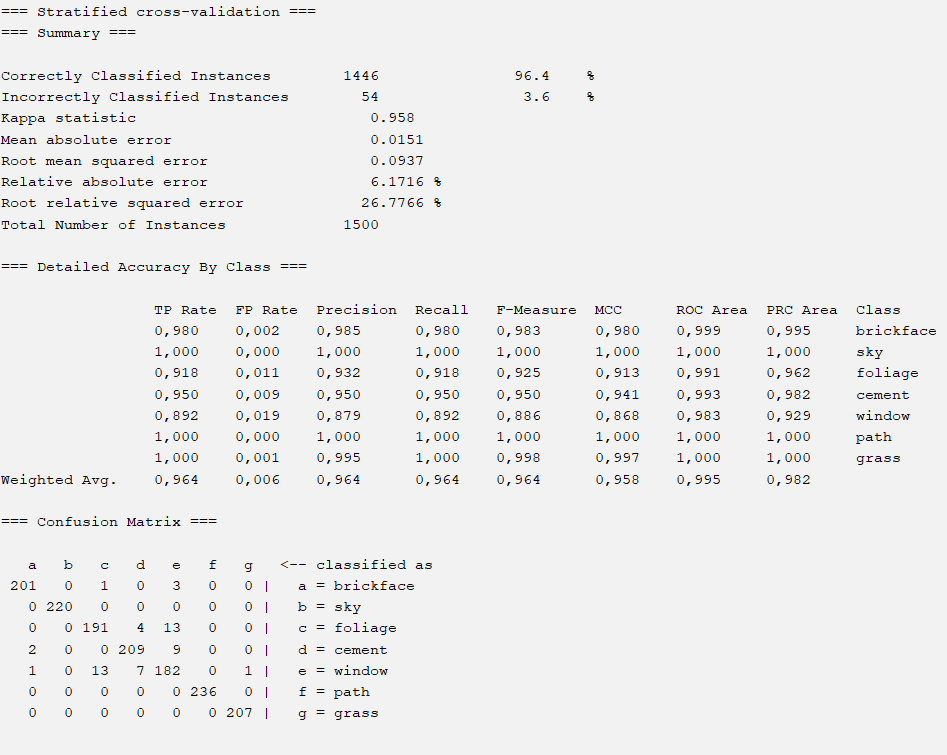


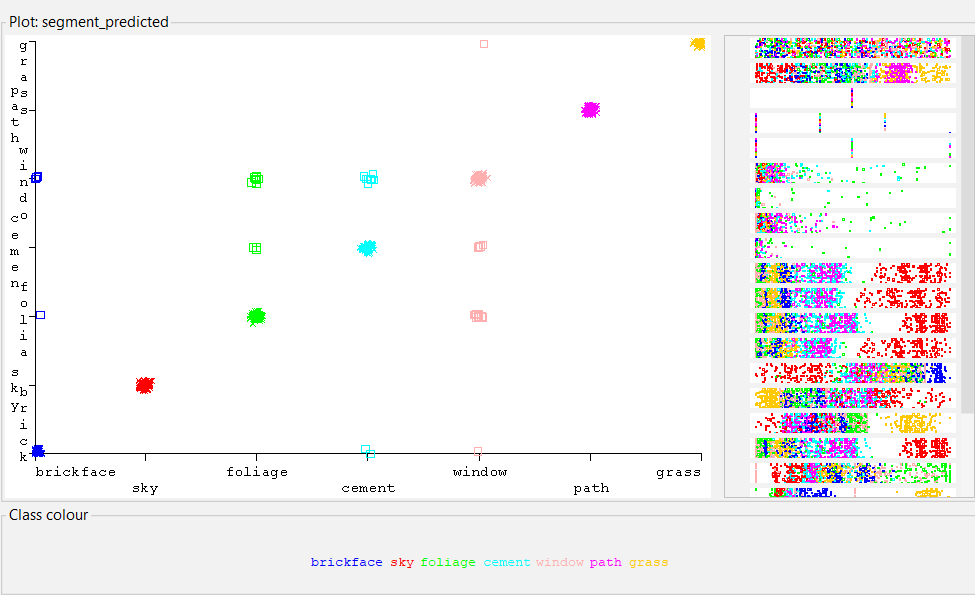


En aquest model observem que la classifica amb una tassa d'acert del 95% aplicant un cross-validation per evitar un possible overfitting veiem que tant la seva matriu de confusió com en els valors d'error són baixos en comparació amb el model de Bayes.

* Multilayer Perceptron resultats

Test mode: 100-fold cross-validation





Com es pot observar en aquest model s'ha millorat els resultats anteriors amb un acert del 96.4% , encara que el cost computacional en realitzar una cross-validation ha sigut molt més lenta que la dels altres dos models aplicats en aquesta prova. També s'ha millorat el valor de la root mean square error, el que indica que la dispersió dels errors és menor en comparació amb els anteriors.

Conclusions:

Dels models seleccionats si es vol decidir pel que sigui més precís escolliria el Multilayer Perceptron encara que aquesta elecció té una contra que és el cost computacional, ja que per assegurar-se d'evitar un overfitting del model tarda més temps.

En canvi, si es volgués tindre un model que tingués una ràtio d'acert elevat, però més àgil segons la prova realitzada optaria pel de Random Tree en aquest cas d'estudi.