

Osservazioni cassificazioni MIL

a.daguanno1

May 2021

Chapter 1

CLASSE BINARIA MALWARE TRUSTED

In questa classificazione è su due valori **trusted** e **malware**.

1.1 istanze non bilanciate

Le istanze non sono bilanciate (**200** Trusted e **482** Malware).
I primi cinque risultati miglior hanno valore di precision e recall coincidente.

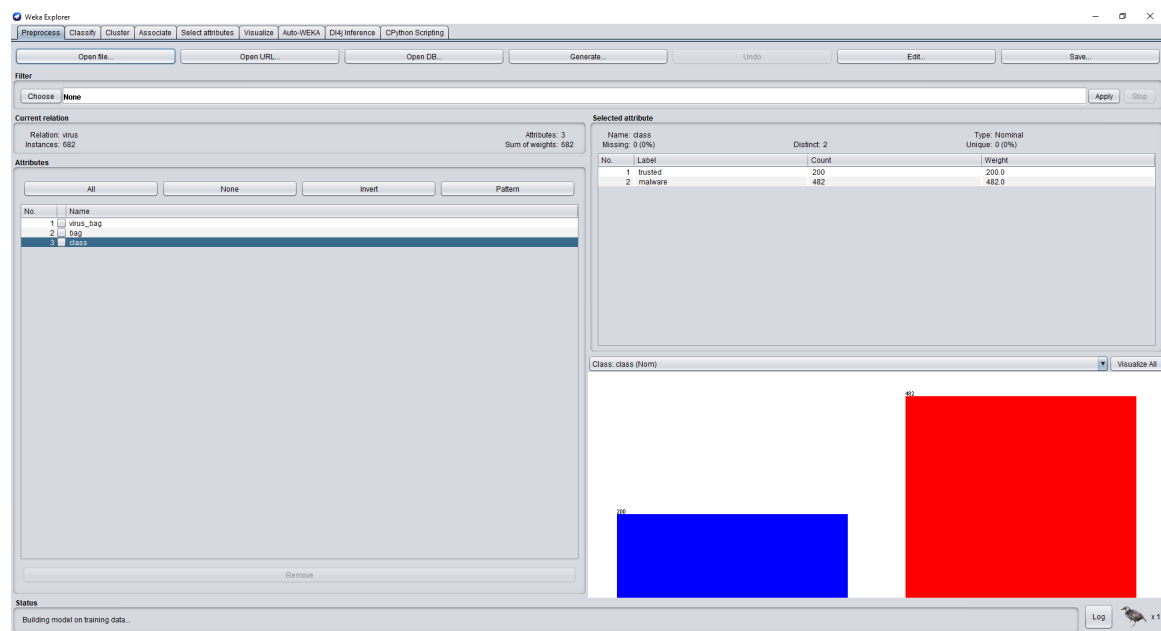


Figure 1.1: Istanze

ALG	PRECISION	RECALL
MIEMDD	0.999	0.999
MIDD	0.997	0.997
QUICK DD	0.996	0.966
TLC	0.993	0.993
MITI	0.981	0.981
MIRI	0.980	0.979
MILR	0.949	0.946
KNN	0.924	0.915
MDD	0.852	0.818
MISVM	0.806	0.733
TLD	0.79	0.299

(a) MIEDD

(b) MIDD

(c) QUICK DD

(d) TLC

(e) MITI

(f) MIRI

(g) MILR

(h) KNN

Figure 1.2: classificazione su classi binarie

1.2 istanze bilanciate

Ho bilanciato le istanze in modo da avere 200 trusted e 200 malware.

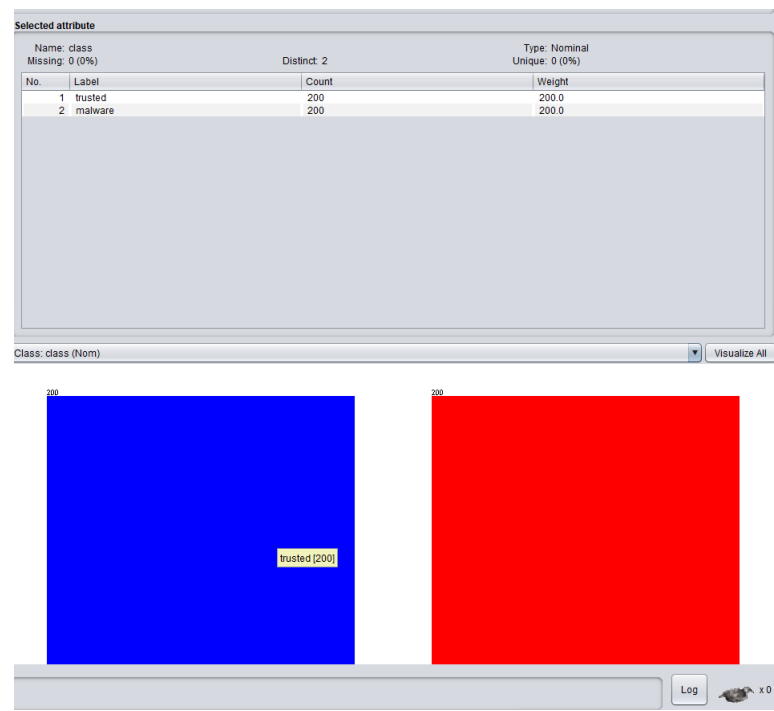


Figure 1.3: etichette trusted / malware bilanciate

ALG	PRECISION	RECALL
MIEMDD	1.000	1.000
MIDD	1.000	1.000
QUICK DD	0.995	0.995
TLC	0.990	0.990
MITI	0.978	0.978
MIRI	0.970	0.970
MILR	0.940	0.945
KNN	0.897	0.61
MDD	0.832	0.830

Chapter 2

CLASSE BROADCAST, EXTERNAL , SHARED

2.1 classificazione con 3 etichette

Rimosse tutte le istanze etichettate come "trusted" ed importato il file .arff
L'errore si ripresenta, i modelli di classificazione funzionanti sono solamente 4 tutta via solamente uno restituisce entrambi i valori di recall e precision il TLC. Precision = 0.779 Recall = 0.780

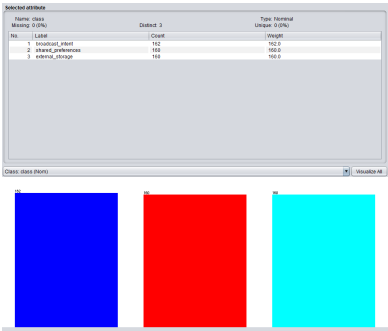


Figure 2.1: .arff con 3 etichette

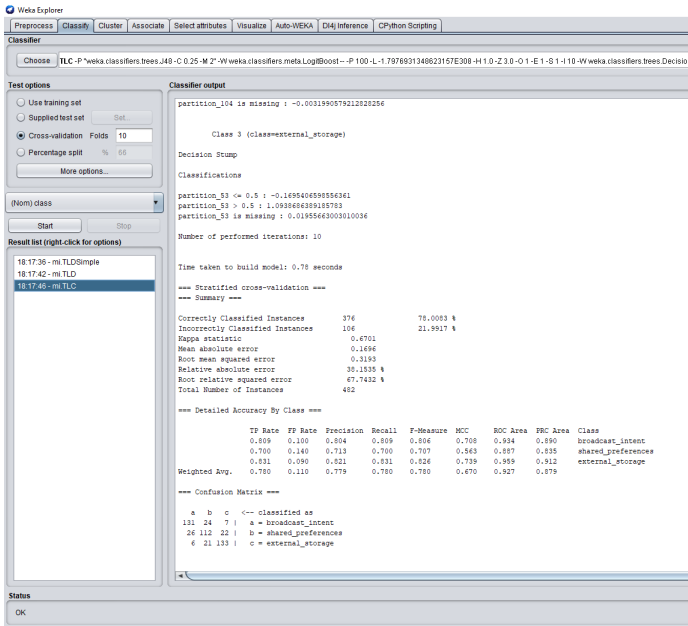


Figure 2.2: tlc - su 3 etichette