

```
1 "C:\Users\My PC\AppData\Local\Programs\Python\Python36-32\  
2 python.exe" D:/pfe/Projets/pfe/apprentissage_BLOB.py  
3  
4  
5 -----Decision Tree-----  
6  
7  
8  
9 dataset/BLOB/taken/BLOB_del.csv  
10  
11  
12  
13 count      16435  
14 unique      2  
15 top        False  
16 freq       16033  
17 Name: is_code_smell, dtype: object  
18 [[4820    0]  
19 [    0 111]]  
20 the recall for this model is : 1.0  
21 TP 111  
22 TN 4820  
23 FP 0  
24 FN 0  
25  
26 -----Classification Report-----  
27  
28  
29          precision    recall   f1-score   support  
30      False        1.00     1.00     1.00     4820  
31      True        1.00     1.00     1.00      111  
32      micro avg     1.00     1.00     1.00     4931  
33      macro avg     1.00     1.00     1.00     4931  
34 weighted avg     1.00     1.00     1.00     4931  
35  
36 Precision = 1.0  
37 Rappel= 1.0  
38 F_Mesure= 1.0  
39  
40 -----Using cross Validation-----  
41  
42 -----k= 1 -----
```

```
43
44 F_Mesure= 1.0
45
46 -----k= 2 -----
47
48 F_Mesure= 1.0
49
50 -----k= 3 -----
51
52 F_Mesure= 1.0
53
54 -----k= 4 -----
55
56 F_Mesure= 1.0
57
58 -----k= 5 -----
59
60 F_Mesure= 1.0
61 F_Mesures moyenne = 1.0
62
63
64
65 -----Decision Tree-----
66
67
68
69 dataset/BLOB/taken/BLOB_RandomUnderSampler.csv
70
71
72
73 count      804
74 unique      2
75 top        True
76 freq       402
77 Name: is_code_smell, dtype: object
78 [[117    0]
79 [ 1 124]]
80 the recall for this model is : 0.992
81 TP 124
82 TN 117
83 FP 0
84 FN 1
85
86 -----Classification Report
-----
```

	precision	recall	f1-score	support
87				
88				
89	False	0.99	1.00	1.00
90	True	1.00	0.99	1.00
91				
92	micro avg	1.00	1.00	1.00
93	macro avg	1.00	1.00	1.00
94	weighted avg	1.00	1.00	1.00
95				
96	Precision = 1.0			
97	Rappel= 0.992			
98	F_Mesure= 0.9959839357429718			
99				
100	-----Using cross Validation			
101				
102	-----k= 1 -----			
103				
104	F_Mesure= 1.0			
105				
106	-----k= 2 -----			
107				
108	F_Mesure= 0.9933774834437086			
109				
110	-----k= 3 -----			
111				
112	F_Mesure= 1.0			
113				
114	-----k= 4 -----			
115				
116	F_Mesure= 1.0			
117				
118	-----k= 5 -----			
119				
120	F_Mesure= 1.0			
121	F_Mesures moyenne = 0.9986754966887418			
122				
123				
124				
125	-----Decision Tree-----			
126				
127				
128				
129	dataset/BLOB/taken/BLOB_AllKNN.csv			
130				

```
131
132
133 count      16435
134 unique       2
135 top        False
136 freq       16033
137 Name: is_code_smell, dtype: object
138 [[4801    0]
139 [    0   130]]
140 the recall for this model is : 1.0
141 TP 130
142 TN 4801
143 FP 0
144 FN 0
145
146 -----Classification Report
-----
147          precision    recall   f1-score   support
148
149      False       1.00     1.00      1.00      4801
150      True       1.00     1.00      1.00      130
151
152  micro avg     1.00     1.00      1.00      4931
153  macro avg     1.00     1.00      1.00      4931
154 weighted avg   1.00     1.00      1.00      4931
155
156 Precision =  1.0
157 Rappel=  1.0
158 F_Mesure= 1.0
159
160 -----Using cross Validation
-----
161
162 -----k=  1 -----
163
164 F_Mesure= 1.0
165
166 -----k=  2 -----
167
168 F_Mesure= 1.0
169
170 -----k=  3 -----
171
172 F_Mesure= 1.0
173
```

```
174 -----k= 4 -----
175
176 F_Mesure= 1.0
177
178 -----k= 5 -----
179
180 F_Mesure= 1.0
181 F_Measures moyenne = 1.0
182
183
184
185 -----Decision Tree-----
186
187
188
189 dataset/BLOB/taken/BLOB_InstanceHardnessThreshold.csv
190
191
192
193 count      16435
194 unique      2
195 top        False
196 freq       16033
197 Name: is_code_smell, dtype: object
198 [[4796    0]
199 [    0   135]]
200 the recall for this model is : 1.0
201 TP 135
202 TN 4796
203 FP 0
204 FN 0
205
206 -----Classification Report
-----
207
208
209          precision    recall   f1-score   support
210
211
212      False        1.00     1.00      1.00      4796
213      True         1.00     1.00      1.00      135
214
215
216      micro avg     1.00     1.00      1.00      4931
217      macro avg     1.00     1.00      1.00      4931
218      weighted avg   1.00     1.00      1.00      4931
219
220
221 Precision =  1.0
222 Rappel=  1.0
```

```
218 F_Mesure= 1.0
219
220 -----Using cross Validation
-----
221
222 -----k= 1 -----
223
224 F_Mesure= 1.0
225
226 -----k= 2 -----
227
228 F_Mesure= 1.0
229
230 -----k= 3 -----
231
232 F_Mesure= 1.0
233
234 -----k= 4 -----
235
236 F_Mesure= 0.9887640449438202
237
238 -----k= 5 -----
239
240 F_Mesure= 1.0
241 F_Mesures moyenne = 0.997752808988764
242
243
244
245 -----Decision Tree-----
246
247
248
249 dataset/BLOB/taken/BLOB_NearMiss.csv
250
251
252
253 count      804
254 unique     2
255 top        True
256 freq       402
257 Name: is_code_smell, dtype: object
258 [[122    0]
259 [ 1 119]]
260 the recall for this model is : 0.9916666666666667
261 TP 119
```

```
262 TN 122
263 FP 0
264 FN 1
265
266 -----Classification Report
-----
267             precision    recall   f1-score   support
268
269     False        0.99      1.00      1.00      122
270     True         1.00      0.99      1.00      120
271
272     micro avg     1.00      1.00      1.00      242
273     macro avg     1.00      1.00      1.00      242
274 weighted avg     1.00      1.00      1.00      242
275
276 Precision = 1.0
277 Rappel= 0.9916666666666667
278 F_Mesure= 0.99581589958159
279
280 -----Using cross Validation
-----
281
282 -----k= 1 -----
283
284 F_Mesure= 1.0
285
286 -----k= 2 -----
287
288 F_Mesure= 0.9933774834437086
289
290 -----k= 3 -----
291
292 F_Mesure= 1.0
293
294 -----k= 4 -----
295
296 F_Mesure= 1.0
297
298 -----k= 5 -----
299
300 F_Mesure= 1.0
301 F_Mesures moyenne = 0.9986754966887418
302
303
304
```

```
305 -----Decision Tree-----
306
307
308
309 dataset/BLOB/taken/BLOB_OneSidedSelection.csv
310
311
312
313 count      578
314 unique      2
315 top        True
316 freq       402
317 Name: is_code_smell, dtype: object
318 [[ 60   0]
319 [  0 114]]
320 the recall for this model is : 1.0
321 TP 114
322 TN 60
323 FP 0
324 FN 0
325
326 -----Classification Report
-----
327
328
329          precision    recall  f1-score   support
330
331
332      False       1.00     1.00     1.00      60
333      True       1.00     1.00     1.00     114
334
335
336      micro avg    1.00     1.00     1.00     174
337      macro avg    1.00     1.00     1.00     174
338      weighted avg  1.00     1.00     1.00     174
339
340 -----Using cross Validation
-----
341
342 -----k= 1 -----
343
344 F_Mesure= 1.0
345
346 -----k= 2 -----
347
```

```
348 F_Mesure= 0.993939393939394
349
350 -----k= 3 -----
351
352 F_Mesure= 1.0
353
354 -----k= 4 -----
355
356 F_Mesure= 1.0
357
358 -----k= 5 -----
359
360 F_Mesure= 1.0
361 F_Mesures moyenne = 0.9987878787878788
362
363
364
365 -----Decision Tree-----
366
367
368
369 dataset/BLOB/taken/BLOB_RandomUnderSampler_default.csv
370
371
372
373 count      804
374 unique     2
375 top        True
376 freq       402
377 Name: is_code_smell, dtype: object
378 [[111  0]
379 [ 0 131]]
380 the recall for this model is : 1.0
381 TP 131
382 TN 111
383 FP 0
384 FN 0
385
386 -----Classification Report
-----
```

	precision	recall	f1-score	support
389 False	1.00	1.00	1.00	111
390 True	1.00	1.00	1.00	131

```
391
```

File - unknown

```
392     micro avg      1.00      1.00      1.00      242
393     macro avg      1.00      1.00      1.00      242
394 weighted avg      1.00      1.00      1.00      242
395
396 Precision = 1.0
397 Rappel= 1.0
398 F_Mesure= 1.0
399
400 -----Using cross Validation
-----
401
402 -----k= 1 -----
403
404 F_Mesure= 1.0
405
406 -----k= 2 -----
407
408 F_Mesure= 1.0
409
410 -----k= 3 -----
411
412 F_Mesure= 1.0
413
414 -----k= 4 -----
415
416 F_Mesure= 1.0
417
418 -----k= 5 -----
419
420 F_Mesure= 1.0
421 F_Mesures moyenne = 1.0
422
423
424
425 -----Decision Tree-----
426
427
428
429 dataset/BLOB/taken/BLOB_TomekLinks.csv
430
431
432
433 count      16435
434 unique      2
435 top        False
```

```
436 freq      16033
437 Name: is_code_smell, dtype: object
438 [[4823    0]
439 [    0 108]]
440 the recall for this model is : 1.0
441 TP 108
442 TN 4823
443 FP 0
444 FN 0
445
446 -----Classification Report
-----
447          precision    recall   f1-score   support
448
449      False       1.00     1.00      1.00      4823
450      True        1.00     1.00      1.00      108
451
452  micro avg       1.00     1.00      1.00      4931
453  macro avg       1.00     1.00      1.00      4931
454 weighted avg     1.00     1.00      1.00      4931
455
456 Precision = 1.0
457 Rappel= 1.0
458 F_Mesure= 1.0
459
460 -----Using cross Validation
-----
461
462 -----k= 1 -----
463
464 F_Mesure= 1.0
465
466 -----k= 2 -----
467
468 F_Mesure= 1.0
469
470 -----k= 3 -----
471
472 F_Mesure= 1.0
473
474 -----k= 4 -----
475
476 F_Mesure= 1.0
477
478 -----k= 5 -----
```

```
479
480 F_Mesure= 1.0
481 F_Mesures moyenne = 1.0
482
483
484
485 -----Random Forest-----
486
487
488
489 dataset/BLOB/taken/BLOB_del.csv
490
491
492
493 count      16435
494 unique       2
495 top        False
496 freq       16033
497 Name: is_code_smell, dtype: object
498 [[4831    0]
499 [    0   100]]
500 the recall for this model is : 1.0
501 TP 100
502 TN 4831
503 FP 0
504 FN 0
505
506 -----Classification Report
-----
507             precision    recall   f1-score   support
508
509      False       1.00     1.00     1.00      4831
510      True       1.00     1.00     1.00      100
511
512  micro avg     1.00     1.00     1.00      4931
513  macro avg     1.00     1.00     1.00      4931
514 weighted avg   1.00     1.00     1.00      4931
515
516 Precision = 1.0
517 Rappel= 1.0
518 F_Mesure= 1.0
519
520 -----Using cross Validation
-----
```

```
522 -----k= 1 -----
523
524 F_Mesure= 1.0
525
526 -----k= 2 -----
527
528 F_Mesure= 1.0
529
530 -----k= 3 -----
531
532 F_Mesure= 0.9937888198757764
533
534 -----k= 4 -----
535
536 F_Mesure= 1.0
537
538 -----k= 5 -----
539
540 F_Mesure= 1.0
541 F_Mesures moyenne = 0.9987577639751553
542
543
544
545 -----Random Forest-----
546
547
548
549 dataset/BLOB/taken/BLOB_RandomUnderSampler.csv
550
551
552
553 count      804
554 unique      2
555 top        True
556 freq       402
557 Name: is_code_smell, dtype: object
558 [[126    0]
559 [ 1 115]]
560 the recall for this model is : 0.9913793103448276
561 TP 115
562 TN 126
563 FP 0
564 FN 1
565
566 -----Classification Report
```

```
566 -----
567             precision    recall   f1-score   support
568
569     False        0.99      1.00      1.00      126
570     True         1.00      0.99      1.00      116
571
572     micro avg     1.00      1.00      1.00      242
573     macro avg     1.00      1.00      1.00      242
574 weighted avg     1.00      1.00      1.00      242
575
576 Precision = 1.0
577 Rappel= 0.9913793103448276
578 F_Mesure= 0.9956709956709957
579
580 -----Using cross Validation
-----
581
582 -----k= 1 -----
583
584 F_Mesure= 1.0
585
586 -----k= 2 -----
587
588 F_Mesure= 0.9933774834437086
589
590 -----k= 3 -----
591
592 F_Mesure= 1.0
593
594 -----k= 4 -----
595
596 F_Mesure= 1.0
597
598 -----k= 5 -----
599
600 F_Mesure= 1.0
601 F_Mesures moyenne = 0.9986754966887418
602
603
604
605 -----Random Forest-----
606
607
608
609 dataset/BLOB/taken/BLOB_AllKNN.csv
```

```
610
611
612
613 count      16435
614 unique      2
615 top        False
616 freq       16033
617 Name: is_code_smell, dtype: object
618 [[4813    0]
619 [ 1  117]]
620 the recall for this model is : 0.9915254237288136
621 TP 117
622 TN 4813
623 FP 0
624 FN 1
625
626 -----Classification Report
-----
627             precision    recall   f1-score   support
628
629      False       1.00     1.00      1.00      4813
630      True        1.00     0.99      1.00      118
631
632  micro avg     1.00     1.00      1.00      4931
633  macro avg     1.00     1.00      1.00      4931
634 weighted avg   1.00     1.00      1.00      4931
635
636 Precision = 1.0
637 Rappel= 0.9915254237288136
638 F_Mesure= 0.9957446808510638
639
640 -----Using cross Validation
-----
641
642 -----k= 1 -----
643
644 F_Mesure= 1.0
645
646 -----k= 2 -----
647
648 F_Mesure= 1.0
649
650 -----k= 3 -----
651
652 F_Mesure= 1.0
```

```
653
654 -----k= 4 -----
655
656 F_Mesure= 0.9944134078212291
657
658 -----k= 5 -----
659
660 F_Mesure= 1.0
661 F_Mesures moyenne = 0.9988826815642458
662
663
664
665 -----Random Forest-----
666
667
668
669 dataset/BLOB/taken/BLOB_InstanceHardnessThreshold.csv
670
671
672
673 count      16435
674 unique      2
675 top        False
676 freq       16033
677 Name: is_code_smell, dtype: object
678 [[4810    0]
679 [    0 121]]
680 the recall for this model is : 1.0
681 TP 121
682 TN 4810
683 FP 0
684 FN 0
685
686 -----Classification Report
-----
687          precision    recall   f1-score   support
688
689      False      1.00      1.00      1.00      4810
690      True       1.00      1.00      1.00      121
691
692  micro avg     1.00      1.00      1.00      4931
693  macro avg     1.00      1.00      1.00      4931
694 weighted avg   1.00      1.00      1.00      4931
695
696 Precision = 1.0
```

```
697 Rappel= 1.0
698 F_Mesure= 1.0
699
700 -----Using cross Validation
-----k= 1 -----
701
702 -----k= 2 -----
703
704 F_Mesure= 1.0
705
706 -----k= 3 -----
707
708 F_Mesure= 1.0
709
710 -----k= 4 -----
711
712 F_Mesure= 1.0
713
714 -----k= 5 -----
715
716 F_Mesure= 0.9944134078212291
717
718 -----Random Forest-----
719
720 F_Mesure= 1.0
721 F_Mesures moyenne = 0.9988826815642458
722
723
724
725 dataset/BLOB/taken/BLOB_NearMiss.csv
726
727
728
729 dataset/BLOB/taken/BLOB_NearMiss.csv
730
731
732
733 count      804
734 unique      2
735 top        True
736 freq       402
737 Name: is_code_smell, dtype: object
738 [[126    0]
739 [  0 116]]
740 the recall for this model is : 1.0
```

```
741 TP 116
742 TN 126
743 FP 0
744 FN 0
745
746 -----Classification Report
-----
747             precision    recall   f1-score   support
748
749      False       1.00     1.00      1.00      126
750      True        1.00     1.00      1.00      116
751
752      micro avg    1.00     1.00      1.00      242
753      macro avg    1.00     1.00      1.00      242
754 weighted avg    1.00     1.00      1.00      242
755
756 Precision = 1.0
757 Rappel= 1.0
758 F_Mesure= 1.0
759
760 -----Using cross Validation
-----
761
762 -----k= 1 -----
763
764 F_Mesure= 1.0
765
766 -----k= 2 -----
767
768 F_Mesure= 0.9933774834437086
769
770 -----k= 3 -----
771
772 F_Mesure= 1.0
773
774 -----k= 4 -----
775
776 F_Mesure= 1.0
777
778 -----k= 5 -----
779
780 F_Mesure= 1.0
781 F_Mesures moyenne = 0.9986754966887418
782
783
```

```
784
785 -----Random Forest-----
786
787
788
789 dataset/BLOB/taken/BLOB_OneSidedSelection.csv
790
791
792
793 count      578
794 unique     2
795 top        True
796 freq       402
797 Name: is_code_smell, dtype: object
798 [[ 52   0]
799 [ 0 122]]
800 the recall for this model is : 1.0
801 TP 122
802 TN 52
803 FP 0
804 FN 0
805
806 -----Classification Report
-----
807          precision    recall   f1-score   support
808
809      False       1.00     1.00      1.00       52
810      True        1.00     1.00      1.00      122
811
812  micro avg     1.00     1.00      1.00      174
813  macro avg     1.00     1.00      1.00      174
814 weighted avg   1.00     1.00      1.00      174
815
816 Precision =  1.0
817 Rappel=  1.0
818 F_Mesure= 1.0
819
820 -----Using cross Validation
-----
821
822 -----k= 1 -----
823
824 F_Mesure= 1.0
825
826 -----k= 2 -----
```

```
827
828 F_Mesure= 0.993939393939394
829
830 -----k= 3 -----
831
832 F_Mesure= 1.0
833
834 -----k= 4 -----
835
836 F_Mesure= 1.0
837
838 -----k= 5 -----
839
840 F_Mesure= 1.0
841 F_Mesures moyenne = 0.9987878787878788
842
843
844
845 -----Random Forest-----
846
847
848
849 dataset/BLOB/taken/BLOB_RandomUnderSampler_default.csv
850
851
852
853 count      804
854 unique      2
855 top        True
856 freq       402
857 Name: is_code_smell, dtype: object
858 [[127    0]
859 [  0 115]]
860 the recall for this model is : 1.0
861 TP 115
862 TN 127
863 FP 0
864 FN 0
865
866 -----Classification Report
-----
867          precision    recall   f1-score   support
868
869      False       1.00      1.00      1.00       127
870      True       1.00      1.00      1.00       115
```

```
871
872     micro avg      1.00      1.00      1.00      242
873     macro avg      1.00      1.00      1.00      242
874 weighted avg      1.00      1.00      1.00      242
875
876 Precision = 1.0
877 Rappel= 1.0
878 F_Mesure= 1.0
879
880 -----Using cross Validation
-----
881
882 -----k= 1 -----
883
884 F_Mesure= 1.0
885
886 -----k= 2 -----
887
888 F_Mesure= 0.9933774834437086
889
890 -----k= 3 -----
891
892 F_Mesure= 1.0
893
894 -----k= 4 -----
895
896 F_Mesure= 1.0
897
898 -----k= 5 -----
899
900 F_Mesure= 1.0
901 F_Mesures moyenne = 0.9986754966887418
902
903
904
905 -----Random Forest-----
906
907
908
909 dataset/BLOB/taken/BLOB_TomekLinks.csv
910
911
912
913 count      16435
914 unique      2
```

```
915 top      False
916 freq     16033
917 Name: is_code_smell, dtype: object
918 [[4824    0]
919 [    0 107]]
920 the recall for this model is : 1.0
921 TP 107
922 TN 4824
923 FP 0
924 FN 0
925
926 -----Classification Report
-----
927          precision    recall   f1-score   support
928
929      False       1.00     1.00      1.00      4824
930      True        1.00     1.00      1.00      107
931
932  micro avg     1.00     1.00      1.00      4931
933  macro avg     1.00     1.00      1.00      4931
934 weighted avg   1.00     1.00      1.00      4931
935
936 Precision =  1.0
937 Rappel=  1.0
938 F_Mesure= 1.0
939
940 -----Using cross Validation
-----
941
942 -----k= 1 -----
943
944 F_Mesure= 1.0
945
946 -----k= 2 -----
947
948 F_Mesure= 1.0
949
950 -----k= 3 -----
951
952 F_Mesure= 0.9937888198757764
953
954 -----k= 4 -----
955
956 F_Mesure= 1.0
957
```

```
958 -----k= 5 -----
959
960 F_Mesure= 1.0
961 F_Mesures moyenne = 0.9987577639751553
962
963
964
965 -----Naive Bayes-----
966
967
968
969 dataset/BLOB/taken/BLOB_del.csv
970
971
972
973 count      16435
974 unique      2
975 top        False
976 freq       16033
977 Name: is_code_smell, dtype: object
978 [[4799    1]
979 [    0   131]]
980 the recall for this model is : 1.0
981 TP 131
982 TN 4799
983 FP 1
984 FN 0
985
986 -----Classification Report
-----
987          precision    recall  f1-score   support
988
989      False       1.00     1.00      1.00     4800
990      True       0.99     1.00      1.00      131
991
992  micro avg       1.00     1.00      1.00     4931
993  macro avg       1.00     1.00      1.00     4931
994 weighted avg      1.00     1.00      1.00     4931
995
996 Precision = 0.9924242424242424
997 Rappel= 1.0
998 F_Mesure= 0.9961977186311787
999
1000 -----Using cross Validation
-----
```

```
1001
1002 -----k= 1 -----
1003
1004 F_Mesure= 0.967032967032967
1005
1006 -----k= 2 -----
1007
1008 F_Mesure= 0.975
1009
1010 -----k= 3 -----
1011
1012 F_Mesure= 0.970059880239521
1013
1014 -----k= 4 -----
1015
1016 F_Mesure= 0.9836065573770492
1017
1018 -----k= 5 -----
1019
1020 F_Mesure= 0.9844559585492229
1021 F_Mesures moyenne = 0.9760310726397521
1022
1023
1024
1025 -----Naive Bayes-----
1026
1027
1028
1029 dataset/BLOB/taken/BLOB_RandomUnderSampler.csv
1030
1031
1032
1033 count      804
1034 unique      2
1035 top        True
1036 freq       402
1037 Name: is_code_smell, dtype: object
1038 [[121  0]
1039 [ 0 121]]
1040 the recall for this model is : 1.0
1041 TP 121
1042 TN 121
1043 FP 0
1044 FN 0
1045
```

```
1046 -----Classification Report
-----
1047             precision    recall   f1-score   support
1048
1049     False        1.00      1.00      1.00      121
1050     True         1.00      1.00      1.00      121
1051
1052     micro avg     1.00      1.00      1.00      242
1053     macro avg     1.00      1.00      1.00      242
1054 weighted avg     1.00      1.00      1.00      242
1055
1056 Precision =  1.0
1057 Rappel=  1.0
1058 F_Mesure= 1.0
1059
1060 -----Using cross Validation
-----
1061
1062 -----k= 1 -----
1063
1064 F_Mesure= 0.9941520467836257
1065
1066 -----k= 2 -----
1067
1068 F_Mesure= 0.9934640522875817
1069
1070 -----k= 3 -----
1071
1072 F_Mesure= 1.0
1073
1074 -----k= 4 -----
1075
1076 F_Mesure= 0.9938650306748467
1077
1078 -----k= 5 -----
1079
1080 F_Mesure= 1.0
1081 F_Mesures moyenne = 0.9962962259492109
1082
1083
1084
1085 -----Naive Bayes-----
1086
1087
1088
```

```
1089 dataset/BLOB/taken/BLOB_AllKNN.csv
1090
1091
1092
1093 count      16435
1094 unique      2
1095 top        False
1096 freq       16033
1097 Name: is_code_smell, dtype: object
1098 [[4811    10]
1099 [    0   110]]
1100 the recall for this model is : 1.0
1101 TP 110
1102 TN 4811
1103 FP 10
1104 FN 0
1105
1106 -----Classification Report
-----
1107          precision    recall   f1-score   support
1108
1109      False       1.00     1.00      1.00      4821
1110      True        0.92     1.00      0.96      110
1111
1112  micro avg     1.00     1.00      1.00      4931
1113  macro avg     0.96     1.00      0.98      4931
1114 weighted avg   1.00     1.00      1.00      4931
1115
1116 Precision =  0.9166666666666666
1117 Rappel= 1.0
1118 F_Mesure= 0.9565217391304348
1119
1120 -----Using cross Validation
-----
1121
1122 -----k= 1 -----
1123
1124 F_Mesure= 0.975609756097561
1125
1126 -----k= 2 -----
1127
1128 F_Mesure= 0.9866666666666666
1129
1130 -----k= 3 -----
1131
```

File - unknown

```
1132 F_Mesure= 0.9625
1133
1134 -----k= 4 -----
1135
1136 F_Mesure= 0.9782608695652174
1137
1138 -----k= 5 -----
1139
1140 F_Mesure= 0.9818181818181818
1141 F_Mesures moyenne = 0.9769710948295254
1142
1143
1144
1145 -----Naive Bayes-----
1146
1147
1148
1149 dataset/BLOB/taken/BLOB_InstanceHardnessThreshold.csv
1150
1151
1152
1153 count      16435
1154 unique      2
1155 top        False
1156 freq       16033
1157 Name: is_code_smell, dtype: object
1158 [[4813    6]
1159 [  0   112]]
1160 the recall for this model is : 1.0
1161 TP 112
1162 TN 4813
1163 FP 6
1164 FN 0
1165
1166 -----Classification Report
-----
1167          precision    recall   f1-score   support
1168
1169      False      1.00      1.00      1.00      4819
1170      True       0.95      1.00      0.97      112
1171
1172  micro avg     1.00      1.00      1.00      4931
1173  macro avg     0.97      1.00      0.99      4931
1174 weighted avg   1.00      1.00      1.00      4931
1175
```

```
1176 Precision = 0.9491525423728814
1177 Rappel= 1.0
1178 F_Mesure= 0.9739130434782608
1179
1180 -----Using cross Validation
-----
1181
1182 -----k= 1 -----
1183
1184 F_Mesure= 0.975609756097561
1185
1186 -----k= 2 -----
1187
1188 F_Mesure= 0.9866666666666666
1189
1190 -----k= 3 -----
1191
1192 F_Mesure= 0.9625
1193
1194 -----k= 4 -----
1195
1196 F_Mesure= 0.9782608695652174
1197
1198 -----k= 5 -----
1199
1200 F_Mesure= 0.9818181818181818
1201 F_Mesures moyenne = 0.9769710948295254
1202
1203
1204
1205 -----Naive Bayes-----
1206
1207
1208
1209 dataset/BLOB/taken/BLOB_NearMiss.csv
1210
1211
1212
1213 count      804
1214 unique      2
1215 top        True
1216 freq       402
1217 Name: is_code_smell, dtype: object
1218 [[128    0]
1219 [  0 114]]
```

```
1220 the recall for this model is : 1.0
1221 TP 114
1222 TN 128
1223 FP 0
1224 FN 0
1225
1226 -----Classification Report
-----
1227             precision    recall   f1-score   support
1228
1229      False        1.00     1.00      1.00      128
1230      True        1.00     1.00      1.00      114
1231
1232  micro avg     1.00     1.00      1.00      242
1233  macro avg     1.00     1.00      1.00      242
1234 weighted avg   1.00     1.00      1.00      242
1235
1236 Precision =  1.0
1237 Rappel=  1.0
1238 F_Mesure= 1.0
1239
1240 -----Using cross Validation
-----
1241
1242 -----k= 1 -----
1243
1244 F_Mesure= 1.0
1245
1246 -----k= 2 -----
1247
1248 F_Mesure= 1.0
1249
1250 -----k= 3 -----
1251
1252 F_Mesure= 1.0
1253
1254 -----k= 4 -----
1255
1256 F_Mesure= 1.0
1257
1258 -----k= 5 -----
1259
1260 F_Mesure= 1.0
1261 F_Mesures moyenne = 1.0
1262
```

```
1263  
1264  
1265 -----Naive Bayes-----  
1266  
1267  
1268  
1269 dataset/BLOB/taken/BLOB_OneSidedSelection.csv  
1270  
1271  
1272  
1273 count      578  
1274 unique      2  
1275 top        True  
1276 freq       402  
1277 Name: is_code_smell, dtype: object  
1278 [[ 48    0]  
1279 [  0 126]]  
1280 the recall for this model is : 1.0  
1281 TP 126  
1282 TN 48  
1283 FP 0  
1284 FN 0  
1285  
1286 -----Classification Report  
-----  
1287          precision    recall   f1-score   support  
1288  
1289      False      1.00      1.00      1.00      48  
1290      True       1.00      1.00      1.00     126  
1291  
1292  micro avg     1.00      1.00      1.00     174  
1293  macro avg     1.00      1.00      1.00     174  
1294 weighted avg   1.00      1.00      1.00     174  
1295  
1296 Precision = 1.0  
1297 Rappel= 1.0  
1298 F_Mesure= 1.0  
1299  
1300 -----Using cross Validation  
-----  
1301  
1302 -----k= 1 -----  
1303  
1304 F_Mesure= 0.9938650306748467  
1305
```

```
1306 -----k= 2 -----
1307
1308 F_Mesure= 1.0
1309
1310 -----k= 3 -----
1311
1312 F_Mesure= 1.0
1313
1314 -----k= 4 -----
1315
1316 F_Mesure= 1.0
1317
1318 -----k= 5 -----
1319
1320 F_Mesure= 1.0
1321 F_Mesures moyenne = 0.9987730061349694
1322
1323
1324
1325 -----Naive Bayes-----
1326
1327
1328
1329 dataset/BLOB/taken/BLOB_RandomUnderSampler_default.csv
1330
1331
1332
1333 count      804
1334 unique      2
1335 top        True
1336 freq       402
1337 Name: is_code_smell, dtype: object
1338 [[119    1]
1339 [  0 122]]
1340 the recall for this model is : 1.0
1341 TP 122
1342 TN 119
1343 FP 1
1344 FN 0
1345
1346 -----Classification Report
-----
1347             precision      recall   f1-score   support
1348
1349     False          1.00      0.99      1.00      120
```

File - unknown

```
1350      True      0.99      1.00      1.00      122
1351
1352      micro avg      1.00      1.00      1.00      242
1353      macro avg      1.00      1.00      1.00      242
1354      weighted avg      1.00      1.00      1.00      242
1355
1356 Precision =  0.991869918699187
1357 Rappel=  1.0
1358 F_Mesure= 0.9959183673469388
1359
1360 -----Using cross Validation
-----
1361
1362 -----k= 1 -----
1363
1364 F_Mesure= 1.0
1365
1366 -----k= 2 -----
1367
1368 F_Mesure= 0.9934640522875817
1369
1370 -----k= 3 -----
1371
1372 F_Mesure= 0.980891719745223
1373
1374 -----k= 4 -----
1375
1376 F_Mesure= 1.0
1377
1378 -----k= 5 -----
1379
1380 F_Mesure= 1.0
1381 F_Mesures moyenne = 0.9948711544065609
1382
1383
1384
1385 -----Naive Bayes-----
1386
1387
1388
1389 dataset/BLOB/taken/BLOB_TomekLinks.csv
1390
1391
1392
1393 count      16435
```

```
1394 unique      2
1395 top      False
1396 freq     16033
1397 Name: is_code_smell, dtype: object
1398 [[4804    4]
1399 [    0   123]]
1400 the recall for this model is : 1.0
1401 TP 123
1402 TN 4804
1403 FP 4
1404 FN 0
1405
1406 -----Classification Report
-----
1407          precision    recall   f1-score   support
1408
1409      False       1.00     1.00      1.00      4808
1410      True        0.97     1.00      0.98      123
1411
1412  micro avg       1.00     1.00      1.00      4931
1413  macro avg       0.98     1.00      0.99      4931
1414 weighted avg     1.00     1.00      1.00      4931
1415
1416 Precision =  0.968503937007874
1417 Rappel=  1.0
1418 F_Mesure= 0.9839999999999999
1419
1420 -----Using cross Validation
-----
1421
1422 -----k= 1 -----
1423
1424 F_Mesure= 0.967032967032967
1425
1426 -----k= 2 -----
1427
1428 F_Mesure= 0.975
1429
1430 -----k= 3 -----
1431
1432 F_Mesure= 0.970059880239521
1433
1434 -----k= 4 -----
1435
1436 F_Mesure= 0.9836065573770492
```

```
1437
1438 -----k= 5 -----
1439
1440 F_Mesure= 0.9844559585492229
1441 F_Mesures moyenne = 0.9760310726397521
1442
1443
1444
1445 -----SVM-----
1446
1447
1448
1449 dataset/BLOB/taken/BLOB_del.csv
1450
1451
1452
1453 count      16435
1454 unique      2
1455 top        False
1456 freq       16033
1457 Name: is_code_smell, dtype: object
1458 [[4788    0]
1459 [ 13   130]]
1460 the recall for this model is : 0.9090909090909091
1461 TP 130
1462 TN 4788
1463 FP 0
1464 FN 13
1465
1466 -----Classification Report
-----
1467          precision    recall  f1-score   support
1468
1469    False       1.00     1.00      1.00      4788
1470    True        1.00     0.91      0.95      143
1471
1472  micro avg     1.00     1.00      1.00      4931
1473  macro avg     1.00     0.95      0.98      4931
1474 weighted avg   1.00     1.00      1.00      4931
1475
1476 Precision = 1.0
1477 Rappel= 0.9090909090909091
1478 F_Mesure= 0.9523809523809523
1479
1480 -----Using cross Validation
```

```
1480 -----
1481
1482 -----k= 1 -----
1483
1484 F_Mesure= 0.9885057471264368
1485
1486 -----k= 2 -----
1487
1488 F_Mesure= 0.9668874172185431
1489
1490 -----k= 3 -----
1491
1492 F_Mesure= 0.9615384615384615
1493
1494 -----k= 4 -----
1495
1496 F_Mesure= 0.9915966386554621
1497
1498 -----k= 5 -----
1499
1500 F_Mesure= 0.978494623655914
1501 F_Mesures moyenne = 0.9774045776389635
1502
1503
1504
1505 -----SVM-----
1506
1507
1508
1509 dataset/BLOB/taken/BLOB_RandomUnderSampler.csv
1510
1511
1512
1513 count      804
1514 unique      2
1515 top        True
1516 freq       402
1517 Name: is_code_smell, dtype: object
1518 [[120    0]
1519 [ 35   87]]
1520 the recall for this model is : 0.7131147540983607
1521 TP 87
1522 TN 120
1523 FP 0
1524 FN 35
```

```
1525
1526 -----Classification Report
-----
1527             precision    recall   f1-score   support
1528
1529     False        0.77      1.00      0.87      120
1530     True         1.00      0.71      0.83      122
1531
1532     micro avg     0.86      0.86      0.86      242
1533     macro avg     0.89      0.86      0.85      242
1534 weighted avg     0.89      0.86      0.85      242
1535
1536 Precision = 1.0
1537 Rappel= 0.7131147540983607
1538 F_Mesure= 0.8325358851674641
1539
1540 -----Using cross Validation
-----
1541
1542 -----k= 1 -----
1543
1544 F_Mesure= 0.9634146341463414
1545
1546 -----k= 2 -----
1547
1548 F_Mesure= 0.8571428571428571
1549
1550 -----k= 3 -----
1551
1552 F_Mesure= 0.8759124087591241
1553
1554 -----k= 4 -----
1555
1556 F_Mesure= 0.8671328671328672
1557
1558 -----k= 5 -----
1559
1560 F_Mesure= 0.9150326797385622
1561 F_Mesures moyenne = 0.8957270893839503
1562
1563
1564
1565 -----SVM-----
1566
1567
```

```
1568
1569 dataset/BLOB/taken/BLOB_AllKNN.csv
1570
1571
1572
1573 count      16435
1574 unique       2
1575 top        False
1576 freq       16033
1577 Name: is_code_smell, dtype: object
1578 [[4821    0]
1579 [    9   101]]
1580 the recall for this model is : 0.9181818181818182
1581 TP 101
1582 TN 4821
1583 FP 0
1584 FN 9
1585
1586 -----Classification Report
-----
1587             precision    recall   f1-score   support
1588
1589      False        1.00     1.00      1.00      4821
1590      True        1.00     0.92      0.96      110
1591
1592  micro avg     1.00     1.00      1.00      4931
1593  macro avg     1.00     0.96      0.98      4931
1594 weighted avg   1.00     1.00      1.00      4931
1595
1596 Precision =  1.0
1597 Rappel=  0.9181818181818182
1598 F_Mesure= 0.957345971563981
1599
1600 -----Using cross Validation
-----
1601
1602 -----k= 1 -----
1603
1604 F_Mesure= 0.9873417721518987
1605
1606 -----k= 2 -----
1607
1608 F_Mesure= 0.950354609929078
1609
1610 -----k= 3 -----
```

```
1611
1612 F_Mesure= 0.9801324503311257
1613
1614 -----k= 4 -----
1615
1616 F_Mesure= 0.9772727272727273
1617
1618 -----k= 5 -----
1619
1620 F_Mesure= 0.9875
1621 F_Mesures moyenne = 0.9765203119369659
1622
1623
1624
1625 -----SVM-----
1626
1627
1628
1629 dataset/BLOB/taken/BLOB_InstanceIdHardnessThreshold.csv
1630
1631
1632
1633 count      16435
1634 unique      2
1635 top        False
1636 freq       16033
1637 Name: is_code_smell, dtype: object
1638 [[4796    0]
1639 [ 12   123]]
1640 the recall for this model is : 0.9111111111111111
1641 TP 123
1642 TN 4796
1643 FP 0
1644 FN 12
1645
1646 -----Classification Report
-----
1647             precision      recall   f1-score   support
1648
1649      False      1.00      1.00      1.00      4796
1650      True      1.00      0.91      0.95      135
1651
1652  micro avg      1.00      1.00      1.00      4931
1653  macro avg      1.00      0.96      0.98      4931
1654 weighted avg      1.00      1.00      1.00      4931
```

```
1655
1656 Precision = 1.0
1657 Rappel= 0.9111111111111111
1658 F_Mesure= 0.9534883720930233
1659
1660 -----Using cross Validation
-----
1661
1662 -----k= 1 -----
1663
1664 F_Mesure= 0.9873417721518987
1665
1666 -----k= 2 -----
1667
1668 F_Mesure= 0.950354609929078
1669
1670 -----k= 3 -----
1671
1672 F_Mesure= 0.9801324503311257
1673
1674 -----k= 4 -----
1675
1676 F_Mesure= 0.9772727272727273
1677
1678 -----k= 5 -----
1679
1680 F_Mesure= 0.9875
1681 F_Mesures moyenne = 0.9765203119369659
1682
1683
1684
1685 -----SVM-----
1686
1687
1688
1689 dataset/BLOB/taken/BLOB_NearMiss.csv
1690
1691
1692
1693 count      804
1694 unique      2
1695 top        True
1696 freq       402
1697 Name: is_code_smell, dtype: object
1698 [[127    0]]
```

```
1699 [ 27 88]
1700 the recall for this model is : 0.7652173913043478
1701 TP 88
1702 TN 127
1703 FP 0
1704 FN 27
1705
1706 -----Classification Report
-----
1707 precision      recall   f1-score   support
1708
1709    False       0.82      1.00      0.90      127
1710    True        1.00      0.77      0.87      115
1711
1712    micro avg   0.89      0.89      0.89      242
1713    macro avg   0.91      0.88      0.89      242
1714 weighted avg  0.91      0.89      0.89      242
1715
1716 Precision = 1.0
1717 Rappel= 0.7652173913043478
1718 F_Mesure= 0.8669950738916257
1719
1720 -----Using cross Validation
-----
1721
1722 -----k= 1 -----
1723
1724 F_Mesure= 0.9506172839506173
1725
1726 -----k= 2 -----
1727
1728 F_Mesure= 0.8307692307692308
1729
1730 -----k= 3 -----
1731
1732 F_Mesure= 0.8592592592592593
1733
1734 -----k= 4 -----
1735
1736 F_Mesure= 0.8345323741007193
1737
1738 -----k= 5 -----
1739
1740 F_Mesure= 0.8551724137931034
1741 F_Mesures moyenne = 0.866070112374586
```

```
1742  
1743  
1744  
1745 -----SVM-----  
1746  
1747  
1748  
1749 dataset/BLOB/taken/BLOB_OneSidedSelection.csv  
1750  
1751  
1752  
1753 count      578  
1754 unique      2  
1755 top        True  
1756 freq       402  
1757 Name: is_code_smell, dtype: object  
1758 [[ 0  55]  
1759 [ 0 119]]  
1760 the recall for this model is : 1.0  
1761 TP 119  
1762 TN 0  
1763 FP 55  
1764 FN 0  
1765  
1766 -----Classification Report  
-----  
1767          precision    recall   f1-score   support  
1768  
1769      False        0.00     0.00     0.00      55  
1770      True        0.68     1.00     0.81     119  
1771  
1772  micro avg     0.68     0.68     0.68     174  
1773  macro avg     0.34     0.50     0.41     174  
1774 weighted avg   0.47     0.68     0.56     174  
1775  
1776 Precision =  0.6839080459770115  
1777 Rappel= 1.0  
1778 F_Mesure= 0.8122866894197952  
1779  
1780 -----Using cross Validation  
-----  
1781 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32  
    \lib\site-packages\sklearn\metrics\classification.py:  
    1143: UndefinedMetricWarning: Precision and F-score are  
    ill-defined and being set to 0.0 in labels with no
```

```
1781 predicted samples.
1782     'precision', 'predicted', average, warn_for)
1783
1784 -----k= 1 -----
1785
1786 F_Mesure= 0.8750000000000001
1787
1788 -----k= 2 -----
1789
1790 F_Mesure= 0.9878048780487805
1791
1792 -----k= 3 -----
1793
1794 F_Mesure= 0.961038961038961
1795
1796 -----k= 4 -----
1797
1798 F_Mesure= 0.7958115183246073
1799
1800 -----k= 5 -----
1801
1802 F_Mesure= 0.8817204301075269
1803 F_Mesures moyenne = 0.9002751575039751
1804
1805
1806
1807 -----SVM-----
1808
1809
1810
1811 dataset/BLOB/taken/BLOB_RandomUnderSampler_default.csv
1812
1813
1814
1815 count      804
1816 unique      2
1817 top        True
1818 freq       402
1819 Name: is_code_smell, dtype: object
1820 [[111    0]
1821 [ 34   97]]
1822 the recall for this model is : 0.7404580152671756
1823 TP 97
1824 TN 111
1825 FP 0
```

```
1826 FN 34
1827
1828 -----Classification Report
-----
1829             precision    recall   f1-score   support
1830
1831      False        0.77      1.00      0.87      111
1832      True         1.00      0.74      0.85      131
1833
1834  micro avg       0.86      0.86      0.86      242
1835  macro avg       0.88      0.87      0.86      242
1836 weighted avg     0.89      0.86      0.86      242
1837
1838 Precision = 1.0
1839 Rappel= 0.7404580152671756
1840 F_Mesure= 0.8508771929824561
1841
1842 -----Using cross Validation
-----
1843
1844 -----k= 1 -----
1845
1846 F_Mesure= 0.9634146341463414
1847
1848 -----k= 2 -----
1849
1850 F_Mesure= 0.8571428571428571
1851
1852 -----k= 3 -----
1853
1854 F_Mesure= 0.8759124087591241
1855
1856 -----k= 4 -----
1857
1858 F_Mesure= 0.8671328671328672
1859
1860 -----k= 5 -----
1861
1862 F_Mesure= 0.9150326797385622
1863 F_Mesures moyenne = 0.8957270893839503
1864
1865
1866
1867 -----SVM-----
1868
```

```
1869
1870
1871 dataset/BLOB/taken/BLOB_TomekLinks.csv
1872
1873
1874
1875 count      16435
1876 unique       2
1877 top        False
1878 freq       16033
1879 Name: is_code_smell, dtype: object
1880 [[4814    0]
1881 [    4   113]]
1882 the recall for this model is : 0.9658119658119658
1883 TP 113
1884 TN 4814
1885 FP 0
1886 FN 4
1887
1888 -----Classification Report
-----
1889          precision    recall   f1-score   support
1890
1891      False      1.00      1.00      1.00      4814
1892      True      1.00      0.97      0.98      117
1893
1894  micro avg      1.00      1.00      1.00      4931
1895  macro avg      1.00      0.98      0.99      4931
1896 weighted avg      1.00      1.00      1.00      4931
1897
1898 Precision = 1.0
1899 Rappel= 0.9658119658119658
1900 F_Mesure= 0.9826086956521739
1901
1902 -----Using cross Validation
-----
1903
1904 -----k= 1 -----
1905
1906 F_Mesure= 0.9885057471264368
1907
1908 -----k= 2 -----
1909
1910 F_Mesure= 0.9668874172185431
1911
```

```
1912 -----k= 3 -----
1913
1914 F_Mesure= 0.9615384615384615
1915
1916 -----k= 4 -----
1917
1918 F_Mesure= 0.9915966386554621
1919
1920 -----k= 5 -----
1921
1922 F_Mesure= 0.978494623655914
1923 F_Mesures moyenne = 0.9774045776389635
1924
1925 Process finished with exit code 0
1926
```