

```
1 "C:\Users\My PC\AppData\Local\Programs\Python\Python36-32\python.exe" D:/pfe/Projets/pfe/apprentissage_HBR.py
2 Empty DataFrame
3 Columns: [classification_methode, path, validation_methode
, F_mesure]
4 Index: []
5
6
7
8 -----Decision Tree-----
9
10
11
12 dataset/HBR/taken/HBR_del.csv
13
14
15
16 count         97
17 unique        2
18 top           False
19 freq          51
20 Name: is_code_smell, dtype: object
21 [[15  0]
22 [ 2 13]]
23 the recall for this model is : 0.8666666666666667
24 TP 13
25 TN 15
26 FP 0
27 FN 2
28
29 -----Classification Report
-----
30
31
32
33
34
35
36
37
38
39
40
41
42
```

	precision	recall	f1-score	support
False	0.88	1.00	0.94	15
True	1.00	0.87	0.93	15
micro avg	0.93	0.93	0.93	30
macro avg	0.94	0.93	0.93	30
weighted avg	0.94	0.93	0.93	30

```
Precision = 1.0
Rappel= 0.8666666666666667
F_Mesure= 0.9285714285714286
Decision Tree dataset/HBR/taken/HBR_del.csv Validation_70/
```

```
42 30 0.9285714285714286
43
44 -----Using cross Validation
-----
45
46 -----k= 1 -----
47
48 F_Mesure= 1.0
49
50 -----k= 2 -----
51
52 F_Mesure= 1.0
53
54 -----k= 3 -----
55
56 F_Mesure= 1.0
57
58 -----k= 4 -----
59
60 F_Mesure= 1.0
61
62 -----k= 5 -----
63
64 F_Mesure= 1.0
65 F_Measures moyenne = 1.0
66
67
68
69 -----Decision Tree-----
70
71
72
73 dataset/HBR/taken/HBR_RandomUnderSampler.csv
74
75
76
77 count         92
78 unique        2
79 top           True
80 freq          46
81 Name: is_code_smell, dtype: object
82 [[16  0]
83 [ 0 12]]
84 the recall for this model is : 1.0
85 TP 12
```

```
86 TN 16
87 FP 0
88 FN 0
89
90 -----Classification Report
-----
91          precision    recall   f1-score   support
92
93      False       1.00     1.00      1.00      16
94      True        1.00     1.00      1.00      12
95
96  micro avg     1.00     1.00      1.00      28
97  macro avg     1.00     1.00      1.00      28
98 weighted avg   1.00     1.00      1.00      28
99
100 Precision = 1.0
101 Rappel= 1.0
102 F_Mesure= 1.0
103 Decision Tree dataset/HBR/taken/HBR_RandomUnderSampler.
     csv Validation_70/30 1.0
104
105 -----Using cross Validation
-----
106
107 -----k= 1 -----
108
109 F_Mesure= 1.0
110
111 -----k= 2 -----
112
113 F_Mesure= 0.88888888888889
114
115 -----k= 3 -----
116
117 F_Mesure= 1.0
118
119 -----k= 4 -----
120
121 F_Mesure= 1.0
122
123 -----k= 5 -----
124
125 F_Mesure= 1.0
126 F_Mesures moyenne = 0.9777777777777777
127
```

```
128
129
130 -----Decision Tree-----
131
132
133
134 dataset/HBR/taken/HBR_AllKNN.csv
135
136
137
138 count         95
139 unique        2
140 top          False
141 freq          49
142 Name: is_code_smell, dtype: object
143 [[15  0]
144 [ 0 14]]
145 the recall for this model is : 1.0
146 TP 14
147 TN 15
148 FP 0
149 FN 0
150
151 -----Classification Report
-----
152             precision    recall   f1-score   support
153
154      False       1.00     1.00      1.00      15
155      True        1.00     1.00      1.00      14
156
157  micro avg     1.00     1.00      1.00      29
158  macro avg     1.00     1.00      1.00      29
159 weighted avg   1.00     1.00      1.00      29
160
161 Precision =  1.0
162 Rappel=  1.0
163 F_Mesure= 1.0
164 Decision Tree dataset/HBR/taken/HBR_AllKNN.csv
165           Validation_70/30 1.0
166
167 -----Using cross Validation
-----
168 -----k= 1 -----
169
```

```
170 F_Mesure= 0.9473684210526316
171
172 -----k= 2 -----
173
174 F_Mesure= 1.0
175
176 -----k= 3 -----
177
178 F_Mesure= 1.0
179
180 -----k= 4 -----
181
182 F_Mesure= 1.0
183
184 -----k= 5 -----
185
186 F_Mesure= 1.0
187 F_Mesures moyenne = 0.9894736842105264
188
189
190
191 -----Decision Tree-----
192
193
194
195 dataset/HBR/taken/HBR_InstanceHardnessThreshold.csv
196
197
198
199 count         92
200 unique        2
201 top           True
202 freq          46
203 Name: is_code_smell, dtype: object
204 [[12  0]
205 [ 0 16]]
206 the recall for this model is : 1.0
207 TP 16
208 TN 12
209 FP 0
210 FN 0
211
212 -----Classification Report
-----
213             precision    recall   f1-score   support
```

```
214
215      False      1.00      1.00      1.00      12
216      True       1.00      1.00      1.00      16
217
218      micro avg    1.00      1.00      1.00      28
219      macro avg    1.00      1.00      1.00      28
220      weighted avg 1.00      1.00      1.00      28
221
222 Precision = 1.0
223 Rappel= 1.0
224 F_Mesure= 1.0
225 Decision Tree dataset/HBR/taken/
    HBR_InstanceHardnessThreshold.csv Validation_70/30 1.0
226
227 -----Using cross Validation
-----
228
229 -----k= 1 -----
230
231 F_Mesure= 1.0
232
233 -----k= 2 -----
234
235 F_Mesure= 0.9473684210526316
236
237 -----k= 3 -----
238
239 F_Mesure= 1.0
240
241 -----k= 4 -----
242
243 F_Mesure= 1.0
244
245 -----k= 5 -----
246
247 F_Mesure= 1.0
248 F_Mesures moyenne = 0.9894736842105264
249
250
251
252 -----Decision Tree-----
253
254
255
256 dataset/HBR/taken/HBR_NearMiss.csv
```

```
257
258
259
260 count         92
261 unique        2
262 top          True
263 freq          46
264 Name: is_code_smell, dtype: object
265 [[14  0]
266 [ 0 14]]
267 the recall for this model is : 1.0
268 TP 14
269 TN 14
270 FP 0
271 FN 0
272
273 -----Classification Report
-----
274             precision    recall   f1-score   support
275
276      False       1.00     1.00      1.00      14
277      True        1.00     1.00      1.00      14
278
279  micro avg     1.00     1.00      1.00      28
280  macro avg     1.00     1.00      1.00      28
281 weighted avg   1.00     1.00      1.00      28
282
283 Precision = 1.0
284 Rappel= 1.0
285 F_Mesure= 1.0
286 Decision Tree dataset/HBR/taken/HBR_NearMiss.csv
    Validation_70/30 1.0
287
288 -----Using cross Validation
-----
289
290 -----k= 1 -----
291
292 F_Mesure= 1.0
293
294 -----k= 2 -----
295
296 F_Mesure= 0.8888888888888889
297
298 -----k= 3 -----
```

```
299
300 F_Mesure= 1.0
301
302 -----k= 4 -----
303
304 F_Mesure= 1.0
305
306 -----k= 5 -----
307
308 F_Mesure= 1.0
309 F_Mesures moyenne = 0.9777777777777779
310
311
312
313 -----Decision Tree-----
314
315
316
317 dataset/HBR/taken/HBR_OneSidedSelection.csv
318
319
320
321 count         96
322 unique        2
323 top           False
324 freq          50
325 Name: is_code_smell, dtype: object
326 [[15  0]
327 [ 0 14]]
328 the recall for this model is : 1.0
329 TP 14
330 TN 15
331 FP 0
332 FN 0
333
334 -----Classification Report
-----
335             precision    recall   f1-score   support
336
337      False       1.00      1.00      1.00       15
338      True        1.00      1.00      1.00       14
339
340  micro avg     1.00      1.00      1.00       29
341  macro avg     1.00      1.00      1.00       29
342 weighted avg   1.00      1.00      1.00       29
```

```
343
344 Precision = 1.0
345 Rappel= 1.0
346 F_Mesure= 1.0
347 Decision Tree dataset/HBR/taken/HBR_OneSidedSelection.csv
    Validation_70/30 1.0
348
349 -----Using cross Validation
-----
350
351 -----k= 1 -----
352
353 F_Mesure= 1.0
354
355 -----k= 2 -----
356
357 F_Mesure= 1.0
358
359 -----k= 3 -----
360
361 F_Mesure= 1.0
362
363 -----k= 4 -----
364
365 F_Mesure= 1.0
366
367 -----k= 5 -----
368
369 F_Mesure= 1.0
370 F_Mesures moyenne = 1.0
371
372
373
374 -----Decision Tree-----
375
376
377
378 dataset/HBR/taken/HBR_RandomUnderSampler_default.csv
379
380
381
382 count      92
383 unique     2
384 top        True
385 freq       46
```

```
386 Name: is_code_smell, dtype: object
387 [[14  0]
388 [ 0 14]]
389 the recall for this model is : 1.0
390 TP 14
391 TN 14
392 FP 0
393 FN 0
394
395 -----Classification Report
-----
396          precision    recall   f1-score   support
397
398      False       1.00     1.00      1.00      14
399      True        1.00     1.00      1.00      14
400
401  micro avg     1.00     1.00      1.00      28
402  macro avg     1.00     1.00      1.00      28
403 weighted avg   1.00     1.00      1.00      28
404
405 Precision =  1.0
406 Rappel=  1.0
407 F_Mesure= 1.0
408 Decision Tree dataset/HBR/taken/
      HBR_RandomUnderSampler_default.csv Validation_70/30 1.0
409
410 -----Using cross Validation
-----
411
412 -----k= 1 -----
413
414 F_Mesure= 1.0
415
416 -----k= 2 -----
417
418 F_Mesure= 0.8888888888888889
419
420 -----k= 3 -----
421
422 F_Mesure= 1.0
423
424 -----k= 4 -----
425
426 F_Mesure= 1.0
427
```

```
428 -----k= 5 -----
429
430 F_Mesure= 1.0
431 F_Mesures moyenne = 0.9777777777777779
432
433
434
435 -----Decision Tree-----
436
437
438
439 dataset/HBR/taken/HBR_TomekLinks.csv
440
441
442
443 count         96
444 unique        2
445 top          False
446 freq          50
447 Name: is_code_smell, dtype: object
448 [[16  0]
449 [ 0 13]]
450 the recall for this model is : 1.0
451 TP 13
452 TN 16
453 FP 0
454 FN 0
455
456 -----Classification Report
-----
457             precision    recall   f1-score   support
458
459      False       1.00     1.00      1.00      16
460      True        1.00     1.00      1.00      13
461
462  micro avg     1.00     1.00      1.00      29
463  macro avg     1.00     1.00      1.00      29
464 weighted avg   1.00     1.00      1.00      29
465
466 Precision =  1.0
467 Rappel=  1.0
468 F_Mesure= 1.0
469 Decision Tree dataset/HBR/taken/HBR_TomekLinks.csv
        Validation_70/30 1.0
470
```

```
471 -----Using cross Validation
472
473 -----k= 1 -----
474
475 F_Mesure= 1.0
476
477 -----k= 2 -----
478
479 F_Mesure= 1.0
480
481 -----k= 3 -----
482
483 F_Mesure= 1.0
484
485 -----k= 4 -----
486
487 F_Mesure= 1.0
488
489 -----k= 5 -----
490
491 F_Mesure= 1.0
492 F_Mesures moyenne = 1.0
493
494
495
496 -----Decision Tree-----
497
498
499
500 dataset/HBR/taken/HBR_CondensedNearestNeighbour.csv
501
502
503
504 count      49
505 unique     2
506 top        True
507 freq       46
508 Name: is_code_smell, dtype: object
509 [[ 1  0]
510 [ 0 14]]
511 the recall for this model is : 1.0
512 TP 14
513 TN 1
514 FP 0
```

```
515 FN 0
516
517 -----Classification Report
-----
518          precision    recall   f1-score   support
519
520      False       1.00     1.00      1.00       1
521      True        1.00     1.00      1.00      14
522
523  micro avg     1.00     1.00      1.00      15
524  macro avg     1.00     1.00      1.00      15
525 weighted avg   1.00     1.00      1.00      15
526
527 Precision =  1.0
528 Rappel=  1.0
529 F_Mesure= 1.0
530 Decision Tree dataset/HBR/taken/
      HBR_CondensedNearestNeighbour.csv Validation_70/30 1.0
531
532 -----Using cross Validation
-----
533
534 -----k= 1 -----
535
536 F_Mesure= 1.0
537
538 -----k= 2 -----
539
540
541 -----k= 3 -----
542
543
544 -----k= 4 -----
545
546 F_Mesure= 1.0
547
548 -----k= 5 -----
549
550 F_Mesure= 0.9411764705882353
551 F_Mesures moyenne = 0.9803921568627452
552
553
554
555 -----Random Forest-----
556
```

```
557
558
559 dataset/HBR/taken/HBR_del.csv
560
561
562
563 count         97
564 unique        2
565 top           False
566 freq          51
567 Name: is_code_smell, dtype: object
568 [[14  0]
569 [ 0 16]]
570 the recall for this model is : 1.0
571 TP 16
572 TN 14
573 FP 0
574 FN 0
575
576 -----Classification Report
-----
577
578
579      precision    recall   f1-score   support
580      False        1.00     1.00      1.00       14
581      True         1.00     1.00      1.00       16
582      micro avg   1.00     1.00      1.00       30
583      macro avg   1.00     1.00      1.00       30
584      weighted avg 1.00     1.00      1.00       30
585
586 Precision = 1.0
587 Rappel= 1.0
588 F_Mesure= 1.0
589 Random Forest dataset/HBR/taken/HBR_del.csv Validation_70
/30 1.0
590
591 -----Using cross Validation
-----
592
593 -----k= 1 -----
594
595 F_Mesure= 0.923076923076923
596
597 -----k= 2 -----
598
```

File - unknown

```
599 F_Mesure= 1.0
600
601 -----k= 3 -----
602
603 F_Mesure= 1.0
604
605 -----k= 4 -----
606
607 F_Mesure= 1.0
608
609 -----k= 5 -----
610
611 F_Mesure= 1.0
612 F_Mesures moyenne = 0.9846153846153847
613
614
615
616 -----Random Forest-----
617
618
619
620 dataset/HBR/taken/HBR_RandomUnderSampler.csv
621
622
623
624 count         92
625 unique        2
626 top           True
627 freq          46
628 Name: is_code_smell, dtype: object
629 [[14  0]
630 [ 0 14]]
631 the recall for this model is : 1.0
632 TP 14
633 TN 14
634 FP 0
635 FN 0
636
637 -----Classification Report
-----
638             precision    recall   f1-score   support
639
640      False       1.00      1.00      1.00       14
641      True        1.00      1.00      1.00       14
642
```

File - unknown

```
643     micro avg      1.00      1.00      1.00      28
644     macro avg      1.00      1.00      1.00      28
645 weighted avg      1.00      1.00      1.00      28
646
647 Precision = 1.0
648 Rappel= 1.0
649 F_Mesure= 1.0
650 Random Forest dataset/HBR/taken/HBR_RandomUnderSampler.
    csv Validation_70/30 1.0
651
652 -----Using cross Validation
-----
653
654 -----k= 1 -----
655
656 F_Mesure= 1.0
657
658 -----k= 2 -----
659
660 F_Mesure= 0.8888888888888889
661
662 -----k= 3 -----
663
664 F_Mesure= 1.0
665
666 -----k= 4 -----
667
668 F_Mesure= 1.0
669
670 -----k= 5 -----
671
672 F_Mesure= 1.0
673 F_Mesures moyenne = 0.9777777777777779
674
675
676
677 -----Random Forest-----
678
679
680
681 dataset/HBR/taken/HBR_AllKNN.csv
682
683
684
685 count          95
```

```
686 unique      2
687 top      False
688 freq      49
689 Name: is_code_smell, dtype: object
690 [[15  0]
691 [ 0 14]]
692 the recall for this model is : 1.0
693 TP 14
694 TN 15
695 FP 0
696 FN 0
697
698 -----Classification Report
-----
699          precision    recall   f1-score   support
700
701      False       1.00     1.00      1.00      15
702      True       1.00     1.00      1.00      14
703
704  micro avg       1.00     1.00      1.00      29
705  macro avg       1.00     1.00      1.00      29
706 weighted avg       1.00     1.00      1.00      29
707
708 Precision = 1.0
709 Rappel= 1.0
710 F_Mesure= 1.0
711 Random Forest dataset/HBR/taken/HBR_AllKNN.csv
    Validation_70/30 1.0
712
713 -----Using cross Validation
-----
714
715 -----k= 1 -----
716
717 F_Mesure= 1.0
718
719 -----k= 2 -----
720
721 F_Mesure= 1.0
722
723 -----k= 3 -----
724
725 F_Mesure= 1.0
726
727 -----k= 4 -----
```

```
728
729 F_Mesure= 1.0
730
731 -----k= 5 -----
732
733 F_Mesure= 1.0
734 F_Mesures moyenne = 1.0
735
736
737
738 -----Random Forest-----
739
740
741
742 dataset/HBR/taken/HBR_InstanceHardnessThreshold.csv
743
744
745
746 count         92
747 unique        2
748 top           True
749 freq          46
750 Name: is_code_smell, dtype: object
751 [[16  0]
752 [ 0 12]]
753 the recall for this model is : 1.0
754 TP 12
755 TN 16
756 FP 0
757 FN 0
758
759 -----Classification Report
-----
760
761
762      precision    recall   f1-score   support
763
764
765      False        1.00     1.00      1.00       16
766      True         1.00     1.00      1.00       12
767
768
769      micro avg    1.00     1.00      1.00       28
770      macro avg    1.00     1.00      1.00       28
771      weighted avg 1.00     1.00      1.00       28
772
773
774 Precision = 1.0
775 Rappel= 1.0
776 F_Mesure= 1.0
```

```
772 Random Forest dataset/HBR/taken/
    HBR_InstanceHardnessThreshold.csv Validation_70/30 1.0
773
774 -----Using cross Validation
-----
775
776 -----k= 1 -----
777
778 F_Mesure= 1.0
779
780 -----k= 2 -----
781
782 F_Mesure= 1.0
783
784 -----k= 3 -----
785
786 F_Mesure= 1.0
787
788 -----k= 4 -----
789
790 F_Mesure= 1.0
791
792 -----k= 5 -----
793
794 F_Mesure= 1.0
795 F_Mesures moyenne = 1.0
796
797
798
799 -----Random Forest-----
800
801
802
803 dataset/HBR/taken/HBR_NearMiss.csv
804
805
806
807 count         92
808 unique        2
809 top           True
810 freq          46
811 Name: is_code_smell, dtype: object
812 [[16  0]
813 [ 0 12]]
814 the recall for this model is : 1.0
```

```
815 TP 12
816 TN 16
817 FP 0
818 FN 0
819
820 -----Classification Report
-----
821             precision    recall   f1-score   support
822
823      False        1.00      1.00      1.00       16
824      True         1.00      1.00      1.00       12
825
826      micro avg     1.00      1.00      1.00       28
827      macro avg     1.00      1.00      1.00       28
828 weighted avg     1.00      1.00      1.00       28
829
830 Precision = 1.0
831 Rappel= 1.0
832 F_Mesure= 1.0
833 Random Forest dataset/HBR/taken/HBR_NearMiss.csv
      Validation_70/30 1.0
834
835 -----Using cross Validation
-----
836
837 -----k= 1 -----
838
839 F_Mesure= 1.0
840
841 -----k= 2 -----
842
843 F_Mesure= 0.88888888888889
844
845 -----k= 3 -----
846
847 F_Mesure= 1.0
848
849 -----k= 4 -----
850
851 F_Mesure= 0.875
852
853 -----k= 5 -----
854
855 F_Mesure= 1.0
856 F_Mesures moyenne = 0.9527777777777778
```

```
857
858
859
860 -----Random Forest-----
861
862
863
864 dataset/HBR/taken/HBR_OneSidedSelection.csv
865
866
867
868 count         96
869 unique        2
870 top          False
871 freq          50
872 Name: is_code_smell, dtype: object
873 [[17  0]
874 [ 0 12]]
875 the recall for this model is : 1.0
876 TP 12
877 TN 17
878 FP 0
879 FN 0
880
881 -----Classification Report
-----
882             precision    recall   f1-score   support
883
884      False       1.00     1.00      1.00      17
885      True        1.00     1.00      1.00      12
886
887  micro avg     1.00     1.00      1.00      29
888  macro avg     1.00     1.00      1.00      29
889 weighted avg   1.00     1.00      1.00      29
890
891 Precision =  1.0
892 Rappel=  1.0
893 F_Mesure= 1.0
894 Random Forest dataset/HBR/taken/HBR_OneSidedSelection.csv
     Validation _70/30 1.0
895
896 -----Using cross Validation
-----
897
898 -----k= 1 -----
```

```
899
900 F_Mesure= 1.0
901
902 -----k= 2 -----
903
904 F_Mesure= 1.0
905
906 -----k= 3 -----
907
908 F_Mesure= 0.9411764705882353
909
910 -----k= 4 -----
911
912 F_Mesure= 1.0
913
914 -----k= 5 -----
915
916 F_Mesure= 1.0
917 F_Mesures moyenne = 0.9882352941176471
918
919
920
921 -----Random Forest-----
922
923
924
925 dataset/HBR/taken/HBR_RandomUnderSampler_default.csv
926
927
928
929 count         92
930 unique        2
931 top           True
932 freq          46
933 Name: is_code_smell, dtype: object
934 [[14  0]
935 [ 0 14]]
936 the recall for this model is : 1.0
937 TP 14
938 TN 14
939 FP 0
940 FN 0
941
942 -----Classification Report
-----
```

	precision	recall	f1-score	support
943				
944				
945	False	1.00	1.00	1.00
946	True	1.00	1.00	1.00
947				
948	micro avg	1.00	1.00	1.00
949	macro avg	1.00	1.00	1.00
950	weighted avg	1.00	1.00	1.00
951				
952	Precision = 1.0			
953	Rappel= 1.0			
954	F_Mesure= 1.0			
955	Random Forest dataset/HBR/taken/ HBR_RandomUnderSampler_default.csv Validation_70/30 1.0			
956				
957	-----Using cross Validation			
958				
959	-----k= 1 -----			
960				
961	F_Mesure= 1.0			
962				
963	-----k= 2 -----			
964				
965	F_Mesure= 1.0			
966				
967	-----k= 3 -----			
968				
969	F_Mesure= 1.0			
970				
971	-----k= 4 -----			
972				
973	F_Mesure= 1.0			
974				
975	-----k= 5 -----			
976				
977	F_Mesure= 1.0			
978	F_Mesures moyenne = 1.0			
979				
980				
981				
982	-----Random Forest-----			
983				
984				
985				

```
986 dataset/HBR/taken/HBR_TomekLinks.csv
987
988
989
990 count          96
991 unique         2
992 top            False
993 freq           50
994 Name: is_code_smell, dtype: object
995 [[12  0]
996 [ 1 16]]
997 the recall for this model is : 0.9411764705882353
998 TP 16
999 TN 12
1000 FP 0
1001 FN 1
1002
1003 -----Classification Report
-----
1004              precision    recall   f1-score   support
1005
1006      False        0.92      1.00      0.96       12
1007      True         1.00      0.94      0.97       17
1008
1009  micro avg       0.97      0.97      0.97       29
1010  macro avg       0.96      0.97      0.96       29
1011 weighted avg     0.97      0.97      0.97       29
1012
1013 Precision = 1.0
1014 Rappel= 0.9411764705882353
1015 F_Mesure= 0.9696969696969697
1016 Random Forest dataset/HBR/taken/HBR_TomekLinks.csv
      Validation_70/30 0.9696969696969697
1017
1018 -----Using cross Validation
-----
1019
1020 -----k= 1 -----
1021
1022 F_Mesure= 0.9523809523809523
1023
1024 -----k= 2 -----
1025
1026 F_Mesure= 1.0
1027
```

File - unknown

```
1028 -----k= 3 -----
1029
1030 F_Mesure= 0.9600000000000001
1031
1032 -----k= 4 -----
1033
1034 F_Mesure= 1.0
1035
1036 -----k= 5 -----
1037
1038 F_Mesure= 1.0
1039 F_Mesures moyenne = 0.9824761904761905
1040
1041
1042
1043 -----Random Forest-----
1044
1045
1046
1047 dataset/HBR/taken/HBR_CondensedNearestNeighbour.csv
1048
1049
1050
1051 count      49
1052 unique     2
1053 top        True
1054 freq       46
1055 Name: is_code_smell, dtype: object
1056 [[ 0  1]
1057 [ 0 14]]
1058 the recall for this model is : 1.0
1059 TP 14
1060 TN 0
1061 FP 1
1062 FN 0
1063
1064 -----Classification Report
-----
1065          precision    recall   f1-score   support
1066
1067    False      0.00      0.00      0.00         1
1068    True      0.93      1.00      0.97        14
1069
1070  micro avg    0.93      0.93      0.93        15
1071  macro avg    0.47      0.50      0.48        15
```

```
1072 weighted avg          0.87      0.93      0.90      15
1073
1074 Precision = 0.9333333333333333
1075 Rappel= 1.0
1076 F_Mesure= 0.9655172413793104
1077 Random Forest dataset/HBR/taken/
    HBR_CondensedNearestNeighbour.csv Validation_70/30 0.
    9655172413793104
1078
1079 -----Using cross Validation
-----
1080
1081 -----k= 1 -----
1082
1083 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
    predicted samples.
1084     'precision', 'predicted', average, warn_for)
1085 F_Mesure= 1.0
1086
1087 -----k= 2 -----
1088
1089
1090 -----k= 3 -----
1091
1092
1093 -----k= 4 -----
1094
1095 F_Mesure= 0.9473684210526316
1096
1097 -----k= 5 -----
1098
1099 F_Mesure= 0.9411764705882353
1100 F_Mesures moyenne = 0.9628482972136223
1101
1102
1103
1104 -----Naive Bayes-----
1105
1106
1107
1108 dataset/HBR/taken/HBR_del.csv
1109
```

```
1110
1111
1112 count         97
1113 unique        2
1114 top          False
1115 freq          51
1116 Name: is_code_smell, dtype: object
1117 [[15  0]
1118 [ 1 14]]
1119 the recall for this model is : 0.9333333333333333
1120 TP 14
1121 TN 15
1122 FP 0
1123 FN 1
1124
1125 -----Classification Report
-----
1126              precision    recall   f1-score   support
1127
1128      False       0.94      1.00      0.97      15
1129      True        1.00      0.93      0.97      15
1130
1131  micro avg     0.97      0.97      0.97      30
1132  macro avg     0.97      0.97      0.97      30
1133 weighted avg   0.97      0.97      0.97      30
1134
1135 Precision = 1.0
1136 Rappel= 0.9333333333333333
1137 F_Mesure= 0.9655172413793104
1138 Naive Bayes dataset/HBR/taken/HBR_del.csv Validation_70/
      30 0.9655172413793104
1139
1140 -----Using cross Validation
-----
1141
1142 -----k= 1 -----
1143
1144 F_Mesure= 0.9600000000000001
1145
1146 -----k= 2 -----
1147
1148 F_Mesure= 1.0
1149
1150 -----k= 3 -----
1151
```

File - unknown

```
1152 F_Mesure= 1.0
1153
1154 -----k= 4 -----
1155
1156 F_Mesure= 0.9090909090909091
1157
1158 -----k= 5 -----
1159
1160 F_Mesure= 1.0
1161 F_Mesures moyenne = 0.9738181818181818
1162
1163
1164
1165 -----Naive Bayes-----
1166
1167
1168
1169 dataset/HBR/taken/HBR_RandomUnderSampler.csv
1170
1171
1172
1173 count      92
1174 unique     2
1175 top        True
1176 freq       46
1177 Name: is_code_smell, dtype: object
1178 [[18  0]
1179 [ 0 10]]
1180 the recall for this model is : 1.0
1181 TP 10
1182 TN 18
1183 FP 0
1184 FN 0
1185
1186 -----Classification Report
-----
1187          precision    recall   f1-score   support
1188
1189      False      1.00      1.00      1.00       18
1190      True      1.00      1.00      1.00       10
1191
1192  micro avg      1.00      1.00      1.00       28
1193  macro avg      1.00      1.00      1.00       28
1194 weighted avg      1.00      1.00      1.00       28
1195
```

```
1196 Precision = 1.0
1197 Rappel= 1.0
1198 F_Mesure= 1.0
1199 Naive Bayes dataset/HBR/taken/HBR_RandomUnderSampler.csv
    Validation_70/30 1.0
1200
1201 -----Using cross Validation
-----
1202
1203 -----k= 1 -----
1204
1205 F_Mesure= 1.0
1206
1207 -----k= 2 -----
1208
1209 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1210 F_Mesure= 0.9473684210526316
1211
1212     max_open_warning, RuntimeWarning)
1213 -----k= 3 -----
1214
1215 F_Mesure= 1.0
1216
1217 -----k= 4 -----
1218
1219 F_Mesure= 1.0
1220
1221 -----k= 5 -----
1222
1223 F_Mesure= 0.9333333333333333
1224 F_Mesures moyenne = 0.9761403508771931
1225
1226
1227
1228 -----Naive Bayes-----
1229
1230
1231
```

```
1232 dataset/HBR/taken/HBR_AllKNN.csv
1233
1234
1235
1236 count         95
1237 unique        2
1238 top           False
1239 freq          49
1240 Name: is_code_smell, dtype: object
1241 [[19  0]
1242 [ 0 10]]
1243 the recall for this model is : 1.0
1244 TP 10
1245 TN 19
1246 FP 0
1247 FN 0
1248
1249 -----Classification Report
-----
1250              precision    recall   f1-score   support
1251
1252      False       1.00     1.00      1.00      19
1253      True        1.00     1.00      1.00      10
1254
1255  micro avg     1.00     1.00      1.00      29
1256  macro avg     1.00     1.00      1.00      29
1257 weighted avg   1.00     1.00      1.00      29
1258
1259 Precision = 1.0
1260 Rappel= 1.0
1261 F_Mesure= 1.0
1262 Naive Bayes dataset/HBR/taken/HBR_AllKNN.csv
      Validation_70/30 1.0
1263
1264 -----Using cross Validation
-----
1265
1266 -----k= 1 -----
1267
1268 F_Mesure= 1.0
1269
1270 -----k= 2 -----
1271
1272 F_Mesure= 1.0
1273
```

```
1274 -----k= 3 -----
1275
1276 F_Mesure= 1.0
1277
1278 -----k= 4 -----
1279
1280 F_Mesure= 1.0
1281
1282 -----k= 5 -----
1283
1284 F_Mesure= 1.0
1285 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1286     max_open_warning, RuntimeWarning)
1287 F_Mesures moyenne = 1.0
1288
1289
1290
1291 -----Naive Bayes-----
1292
1293
1294
1295 dataset/HBR/taken/HBR_InstanceHardnessThreshold.csv
1296
1297
1298
1299 count      92
1300 unique     2
1301 top        True
1302 freq       46
1303 Name: is_code_smell, dtype: object
1304 [[14  0]
1305 [ 0 14]]
1306 the recall for this model is : 1.0
1307 TP 14
1308 TN 14
1309 FP 0
1310 FN 0
1311
```

```
1312 -----Classification Report
-----
1313             precision    recall   f1-score   support
1314
1315       False        1.00      1.00      1.00       14
1316       True        1.00      1.00      1.00       14
1317
1318   micro avg     1.00      1.00      1.00       28
1319   macro avg     1.00      1.00      1.00       28
1320 weighted avg    1.00      1.00      1.00       28
1321
1322 Precision =  1.0
1323 Rappel= 1.0
1324 F_Mesure= 1.0
1325 Naive Bayes dataset/HBR/taken/
      HBR_InstanceHardnessThreshold.csv Validation_70/30 1.0
1326
1327 -----Using cross Validation
-----
1328
1329 -----k= 1 -----
1330
1331 F_Mesure= 1.0
1332
1333 -----k= 2 -----
1334
1335 F_Mesure= 1.0
1336
1337 -----k= 3 -----
1338
1339 F_Mesure= 1.0
1340
1341 -----k= 4 -----
1342
1343 F_Mesure= 1.0
1344
1345 -----k= 5 -----
1346
1347 F_Mesure= 1.0
1348 F_Mesures moyenne = 1.0
1349
1350
1351
1352 -----Naive Bayes-----
1353
```

```
1354  
1355  
1356 dataset/HBR/taken/HBR_NearMiss.csv  
1357  
1358  
1359  
1360 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32  
    \lib\site-packages\matplotlib\pyplot.py:537:  
      RuntimeWarning: More than 20 figures have been opened.  
      Figures created through the pyplot interface (`  
      matplotlib.pyplot.figure`) are retained until explicitly  
      closed and may consume too much memory. (To control  
      this warning, see the rcParam `figure.max_open_warning  
      `).  
1361     max_open_warning, RuntimeWarning)  
1362 count         92  
1363 unique        2  
1364 top           True  
1365 freq          46  
1366 Name: is_code_smell, dtype: object  
1367 [[11  0]  
1368 [ 0 17]]  
1369 the recall for this model is : 1.0  
1370 TP 17  
1371 TN 11  
1372 FP 0  
1373 FN 0  
1374  
1375 -----Classification Report  
-----  
1376             precision    recall   f1-score   support  
1377  
1378     False       1.00      1.00      1.00       11  
1379     True        1.00      1.00      1.00       17  
1380  
1381     micro avg    1.00      1.00      1.00       28  
1382     macro avg    1.00      1.00      1.00       28  
1383 weighted avg   1.00      1.00      1.00       28  
1384  
1385 Precision = 1.0  
1386 Rappel= 1.0  
1387 F_Mesure= 1.0  
1388 Naive Bayes dataset/HBR/taken/HBR_NearMiss.csv  
      Validation_70/30 1.0  
1389
```

```
1390 -----Using cross Validation
-----
1391
1392 -----k= 1 -----
1393
1394 F_Mesure= 1.0
1395
1396 -----k= 2 -----
1397
1398 F_Mesure= 0.88888888888889
1399
1400 -----k= 3 -----
1401
1402 F_Mesure= 1.0
1403
1404 -----k= 4 -----
1405
1406 F_Mesure= 0.9411764705882353
1407
1408 -----k= 5 -----
1409
1410 F_Mesure= 1.0
1411 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`)
    are retained until explicitly closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1412     max_open_warning, RuntimeWarning)
1413 F_Mesures moyenne = 0.9660130718954247
1414
1415
1416
1417 -----Naive Bayes-----
1418
1419
1420
1421 dataset/HBR/taken/HBR_OneSidedSelection.csv
1422
1423
1424
1425 count          96
1426 unique          2
```

```
1427 top      False
1428 freq      50
1429 Name: is_code_smell, dtype: object
1430 [[18  0]
1431 [ 0 11]]
1432 the recall for this model is : 1.0
1433 TP 11
1434 TN 18
1435 FP 0
1436 FN 0
1437
1438 -----Classification Report
-----
1439          precision    recall   f1-score   support
1440
1441      False       1.00     1.00      1.00      18
1442      True        1.00     1.00      1.00      11
1443
1444  micro avg     1.00     1.00      1.00      29
1445  macro avg     1.00     1.00      1.00      29
1446 weighted avg   1.00     1.00      1.00      29
1447
1448 Precision = 1.0
1449 Rappel= 1.0
1450 F_Mesure= 1.0
1451 Naive Bayes dataset/HBR/taken/HBR_OneSidedSelection.csv
Validation_70/30 1.0
1452
1453 -----Using cross Validation
-----
1454
1455 -----k= 1 -----
1456
1457 F_Mesure= 1.0
1458
1459 -----k= 2 -----
1460
1461 F_Mesure= 1.0
1462
1463 -----k= 3 -----
1464
1465 F_Mesure= 1.0
1466
1467 -----k= 4 -----
1468
```

```
1469 F_Mesure= 1.0
1470
1471 -----k= 5 -----
1472
1473 F_Mesure= 1.0
1474 F_Measures moyenne = 1.0
1475
1476
1477
1478 -----Naive Bayes-----
1479
1480
1481
1482 dataset/HBR/taken/HBR_RandomUnderSampler_default.csv
1483
1484
1485
1486 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1487     max_open_warning, RuntimeWarning)
1488 count      92
1489 unique     2
1490 top        True
1491 freq       46
1492 Name: is_code_smell, dtype: object
1493 [[12  0]
1494 [ 0 16]]
1495 the recall for this model is : 1.0
1496 TP 16
1497 TN 12
1498 FP 0
1499 FN 0
1500
1501 -----Classification Report
-----
1502          precision    recall   f1-score   support
1503
1504      False       1.00      1.00      1.00       12
1505      True       1.00      1.00      1.00       16
```

```
1506
1507     micro avg      1.00      1.00      1.00      28
1508     macro avg      1.00      1.00      1.00      28
1509 weighted avg      1.00      1.00      1.00      28
1510
1511 Precision = 1.0
1512 Rappel= 1.0
1513 F_Mesure= 1.0
1514 Naive Bayes dataset/HBR/taken/
    HBR_RandomUnderSampler_default.csv Validation_70/30 1.0
1515
1516 -----Using cross Validation
-----
1517
1518 -----k= 1 -----
1519
1520 F_Mesure= 1.0
1521
1522 -----k= 2 -----
1523
1524 F_Mesure= 0.9473684210526316
1525
1526 -----k= 3 -----
1527
1528 F_Mesure= 1.0
1529
1530 -----k= 4 -----
1531
1532 F_Mesure= 1.0
1533
1534 -----k= 5 -----
1535
1536 F_Mesure= 1.0
1537 F_Mesures moyenne = 0.9894736842105264
1538
1539
1540
1541 -----Naive Bayes-----
1542
1543
1544
1545 dataset/HBR/taken/HBR_TomekLinks.csv
1546
1547
1548
```

```
1549 count         96
1550 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
RuntimeWarning: More than 20 figures have been opened.
Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).
1551 unique         2
1552     max_open_warning, RuntimeWarning)
1553 top        False
1554 freq        50
1555 Name: is_code_smell, dtype: object
1556 [[17  0]
1557 [ 0 12]]
1558 the recall for this model is : 1.0
1559 TP 12
1560 TN 17
1561 FP 0
1562 FN 0
1563
1564 -----Classification Report
-----
1565             precision    recall   f1-score   support
1566
1567      False       1.00     1.00      1.00      17
1568      True       1.00     1.00      1.00      12
1569
1570  micro avg     1.00     1.00      1.00      29
1571  macro avg     1.00     1.00      1.00      29
1572 weighted avg   1.00     1.00      1.00      29
1573
1574 Precision =  1.0
1575 Rappel=  1.0
1576 F_Mesure= 1.0
1577 Naive Bayes dataset/HBR/taken/HBR_TomekLinks.csv
    Validation_70/30 1.0
1578
1579 -----Using cross Validation
-----
1580
1581 -----k= 1 -----
1582
1583 F_Mesure= 1.0
```

```
1584
1585 -----k= 2 -----
1586
1587 F_Mesure= 1.0
1588
1589 -----k= 3 -----
1590
1591 F_Mesure= 1.0
1592
1593 -----k= 4 -----
1594
1595 F_Mesure= 1.0
1596
1597 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1598 -----k= 5 -----
1599     max_open_warning, RuntimeWarning)
1600
1601 F_Mesure= 1.0
1602 F_Mesures moyenne = 1.0
1603
1604
1605
1606 -----Naive Bayes-----
1607
1608
1609
1610 dataset/HBR/taken/HBR_CondensedNearestNeighbour.csv
1611
1612
1613
1614 count      49
1615 unique      2
1616 top        True
1617 freq       46
1618 Name: is_code_smell, dtype: object
1619 [[ 1  0]
1620 [ 2 12]]
1621 the recall for this model is : 0.8571428571428571
```

```
1622 TP 12
1623 TN 1
1624 FP 0
1625 FN 2
1626
1627 -----Classification Report
-----
1628          precision    recall   f1-score   support
1629
1630      False       0.33     1.00      0.50        1
1631      True        1.00     0.86      0.92       14
1632
1633  micro avg     0.87     0.87      0.87       15
1634  macro avg     0.67     0.93      0.71       15
1635 weighted avg   0.96     0.87      0.89       15
1636
1637 Precision = 1.0
1638 Rappel= 0.8571428571428571
1639 F_Mesure= 0.923076923076923
1640 Naive Bayes dataset/HBR/taken/
    HBR_CondensedNearestNeighbour.csv Validation_70/30 0.
    923076923076923
1641
1642 -----Using cross Validation
-----
1643
1644 -----k= 1 -----
1645
1646 F_Mesure= 0.9411764705882353
1647
1648 -----k= 2 -----
1649
1650
1651 -----k= 3 -----
1652
1653
1654 -----k= 4 -----
1655
1656 F_Mesure= 0.9473684210526316
1657
1658 -----k= 5 -----
1659
1660 F_Mesure= 0.9411764705882353
1661 F_Mesures moyenne = 0.9432404540763674
1662
```

```
1663
1664
1665 -----SVM-----
1666
1667
1668
1669 dataset/HBR/taken/HBR_del.csv
1670
1671
1672
1673 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`)
    are retained until explicitly closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1674     max_open_warning, RuntimeWarning)
1675 count         97
1676 unique        2
1677 top          False
1678 freq          51
1679 Name: is_code_smell, dtype: object
1680 [[16  0]
1681 [ 3 11]]
1682 the recall for this model is : 0.7857142857142857
1683 TP 11
1684 TN 16
1685 FP 0
1686 FN 3
1687
1688 -----Classification Report
-----
1689             precision      recall   f1-score   support
1690
1691       False        0.84      1.00      0.91        16
1692       True        1.00      0.79      0.88        14
1693
1694   micro avg       0.90      0.90      0.90        30
1695   macro avg       0.92      0.89      0.90        30
1696 weighted avg     0.92      0.90      0.90        30
1697
1698 Precision = 1.0
1699 Rappel= 0.7857142857142857
```

File - unknown

```
1700 F_Mesure= 0.88
1701 Naive Bayes dataset/HBR/taken/HBR_del.csv Validation_70/
    30 0.88
1702
1703 -----Using cross Validation
-----
1704
1705 -----k= 1 -----
1706
1707 F_Mesure= 0.9600000000000001
1708
1709 -----k= 2 -----
1710
1711 F_Mesure= 0.88888888888889
1712
1713 -----k= 3 -----
1714
1715 F_Mesure= 1.0
1716
1717 -----k= 4 -----
1718
1719 F_Mesure= 0.8
1720
1721 -----k= 5 -----
1722
1723 F_Mesure= 1.0
1724 F_Mesures moyenne = 0.9297777777777778
1725
1726
1727
1728 -----SVM-----
1729
1730
1731
1732 dataset/HBR/taken/HBR_RandomUnderSampler.csv
1733
1734
1735
1736 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`)
    are retained until explicitly closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning`
```

```
1736 `).
1737 count      92
1738     max_open_warning, RuntimeWarning)
1739 unique      2
1740 top        True
1741 freq       46
1742 Name: is_code_smell, dtype: object
1743 [[14  1]
1744   [ 1 12]]
1745 the recall for this model is : 0.9230769230769231
1746 TP 12
1747 TN 14
1748 FP 1
1749 FN 1
1750
1751 -----Classification Report
-----
1752             precision    recall   f1-score   support
1753
1754     False        0.93      0.93      0.93      15
1755     True        0.92      0.92      0.92      13
1756
1757   micro avg     0.93      0.93      0.93      28
1758   macro avg     0.93      0.93      0.93      28
1759 weighted avg    0.93      0.93      0.93      28
1760
1761 Precision = 0.9230769230769231
1762 Rappel= 0.9230769230769231
1763 F_Mesure= 0.9230769230769231
1764 Naive Bayes dataset/HBR/taken/HBR_RandomUnderSampler.csv
    Validation_70/30 0.9230769230769231
1765
1766 -----Using cross Validation
-----
1767
1768 -----k= 1 -----
1769
1770 F_Mesure= 1.0
1771
1772 -----k= 2 -----
1773
1774 F_Mesure= 0.8888888888888889
1775
1776 -----k= 3 -----
1777
```

```
1778 F_Mesure= 0.888888888888889
1779
1780 -----k= 4 -----
1781
1782 F_Mesure= 0.933333333333333
1783
1784 -----k= 5 -----
1785
1786 F_Mesure= 0.933333333333333
1787 F_Mesures moyenne = 0.928888888888889
1788
1789
1790
1791 -----SVM-----
1792
1793
1794
1795 dataset/HBR/taken/HBR_AllKNN.csv
1796
1797
1798
1799 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1800     max_open_warning, RuntimeWarning)
1801 count         95
1802 unique        2
1803 top          False
1804 freq          49
1805 Name: is_code_smell, dtype: object
1806 [[14  0]
1807 [ 2 13]]
1808 the recall for this model is : 0.866666666666667
1809 TP 13
1810 TN 14
1811 FP 0
1812 FN 2
1813
1814 -----Classification Report
-----
```

File - unknown

	precision	recall	f1-score	support	
1815					
1816					
1817	False	0.88	1.00	0.93	14
1818	True	1.00	0.87	0.93	15
1819					
1820	micro avg	0.93	0.93	0.93	29
1821	macro avg	0.94	0.93	0.93	29
1822	weighted avg	0.94	0.93	0.93	29
1823					
1824	Precision = 1.0				
1825	Rappel= 0.8666666666666667				
1826	F_Mesure= 0.9285714285714286				
1827	Naive Bayes dataset/HBR/taken/HBR_AllKNN.csv				
	Validation_70/30 0.9285714285714286				
1828					
1829	-----Using cross Validation				

1830					
1831	-----k= 1 -----				
1832					
1833	F_Mesure= 0.9473684210526316				
1834					
1835	-----k= 2 -----				
1836					
1837	F_Mesure= 1.0				
1838					
1839	-----k= 3 -----				
1840					
1841	F_Mesure= 0.8750000000000001				
1842					
1843	-----k= 4 -----				
1844					
1845	F_Mesure= 1.0				
1846					
1847	-----k= 5 -----				
1848					
1849	F_Mesure= 0.923076923076923				
1850	F_Mesures moyenne = 0.949089068825911				
1851					
1852					
1853					
1854	-----SVM-----				
1855					
1856					
1857					

```
1858 dataset/HBR/taken/HBR_InstanceHardnessThreshold.csv
1859
1860
1861
1862 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
    `).
1863 count      92
1864     max_open_warning, RuntimeWarning)
1865 unique      2
1866 top         True
1867 freq        46
1868 Name: is_code_smell, dtype: object
1869 [[14  0]
1870 [ 0 14]]
1871 the recall for this model is : 1.0
1872 TP 14
1873 TN 14
1874 FP 0
1875 FN 0
1876
1877 -----Classification Report
-----
1878             precision    recall   f1-score   support
1879
1880     False       1.00     1.00      1.00      14
1881     True       1.00     1.00      1.00      14
1882
1883   micro avg    1.00     1.00      1.00      28
1884   macro avg    1.00     1.00      1.00      28
1885 weighted avg  1.00     1.00      1.00      28
1886
1887 Precision =  1.0
1888 Rappel= 1.0
1889 F_Mesure= 1.0
1890 Naive Bayes dataset/HBR/taken/
    HBR_InstanceHardnessThreshold.csv Validation_70/30 1.0
1891
1892 -----Using cross Validation
-----
```

```
1893
1894 -----k= 1 -----
1895
1896 F_Mesure= 1.0
1897
1898 -----k= 2 -----
1899
1900 F_Mesure= 0.9473684210526316
1901
1902 -----k= 3 -----
1903
1904 F_Mesure= 0.888888888888889
1905
1906 -----k= 4 -----
1907
1908 F_Mesure= 0.9333333333333333
1909
1910 -----k= 5 -----
1911
1912 F_Mesure= 1.0
1913 F_Mesures moyenne = 0.9539181286549707
1914
1915
1916
1917 -----SVM-----
1918
1919
1920
1921 dataset/HBR/taken/HBR_NearMiss.csv
1922
1923
1924
1925 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`  

        matplotlib.pyplot.figure`) are retained until explicitly
        closed and may consume too much memory. (To control
        this warning, see the rcParam `figure.max_open_warning
        `).
1926     max_open_warning, RuntimeWarning)
1927 count         92
1928 unique        2
1929 top          True
1930 freq          46
```

File - unknown

```
1931 Name: is_code_smell, dtype: object
1932 [[16  0]
1933 [ 1 11]]
1934 the recall for this model is : 0.9166666666666666
1935 TP 11
1936 TN 16
1937 FP 0
1938 FN 1
1939
1940 -----Classification Report
-----
1941          precision    recall   f1-score   support
1942
1943      False       0.94     1.00      0.97      16
1944      True        1.00     0.92      0.96      12
1945
1946  micro avg     0.96     0.96      0.96      28
1947  macro avg     0.97     0.96      0.96      28
1948 weighted avg   0.97     0.96      0.96      28
1949
1950 Precision = 1.0
1951 Rappel= 0.9166666666666666
1952 F_Mesure= 0.9565217391304348
1953 Naive Bayes dataset/HBR/taken/HBR_NearMiss.csv
      Validation_70/30 0.9565217391304348
1954
1955 -----Using cross Validation
-----
1956
1957 -----k= 1 -----
1958
1959 F_Mesure= 1.0
1960
1961 -----k= 2 -----
1962
1963 F_Mesure= 0.8235294117647058
1964
1965 -----k= 3 -----
1966
1967 F_Mesure= 0.888888888888889
1968
1969 -----k= 4 -----
1970
1971 F_Mesure= 0.875
1972
```

```

1973 -----k= 5 -----
1974
1975 F_Mesure= 1.0
1976 F_Mesures moyenne = 0.917483660130719
1977 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
  \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
    this warning, see the rcParam `figure.max_open_warning
`).
1978     max_open_warning, RuntimeWarning)
1979
1980
1981
1982 -----SVM-----
1983
1984
1985
1986 dataset/HBR/taken/HBR_OneSidedSelection.csv
1987
1988
1989
1990 count         96
1991 unique        2
1992 top           False
1993 freq          50
1994 Name: is_code_smell, dtype: object
1995 [[14  0]
1996 [ 1 14]]
1997 the recall for this model is : 0.9333333333333333
1998 TP 14
1999 TN 14
2000 FP 0
2001 FN 1
2002
2003 -----Classification Report
-----
2004             precision      recall   f1-score   support
2005
2006      False       0.93      1.00      0.97       14
2007      True        1.00      0.93      0.97       15
2008
2009  micro avg       0.97      0.97      0.97       29

```

File - unknown

```
2010     macro avg      0.97      0.97      0.97      29
2011 weighted avg      0.97      0.97      0.97      29
2012
2013 Precision = 1.0
2014 Rappel= 0.9333333333333333
2015 F_Mesure= 0.9655172413793104
2016 Naive Bayes dataset/HBR/taken/HBR_OneSidedSelection.csv
      Validation_70/30 0.9655172413793104
2017
2018 -----Using cross Validation
      -----
2019
2020 -----k= 1 -----
2021
2022 F_Mesure= 1.0
2023
2024 -----k= 2 -----
2025
2026 F_Mesure= 0.9523809523809523
2027
2028 -----k= 3 -----
2029
2030 F_Mesure= 0.8571428571428571
2031
2032 -----k= 4 -----
2033
2034 F_Mesure= 1.0
2035
2036 -----k= 5 -----
2037
2038 F_Mesure= 0.923076923076923
2039 F_Mesures moyenne = 0.9465201465201465
2040
2041
2042
2043 -----SVM-----
2044
2045
2046
2047 dataset/HBR/taken/HBR_RandomUnderSampler_default.csv
2048
2049
2050
2051 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
      \lib\site-packages\matplotlib\pyplot.py:537:
```

```
2051 RuntimeWarning: More than 20 figures have been opened.  
    Figures created through the pyplot interface (`  
    matplotlib.pyplot.figure`) are retained until explicitly  
    closed and may consume too much memory. (To control  
    this warning, see the rcParam `figure.max_open_warning  
    `).  
2052     max_open_warning, RuntimeWarning)  
2053 count         92  
2054 unique        2  
2055 top          True  
2056 freq          46  
2057 Name: is_code_smell, dtype: object  
2058 [[ 9  0]  
2059 [ 6 13]]  
2060 the recall for this model is : 0.6842105263157895  
2061 TP 13  
2062 TN 9  
2063 FP 0  
2064 FN 6  
2065  
2066 -----Classification Report  
-----  
2067             precision      recall   f1-score   support  
2068  
2069     False       0.60      1.00      0.75       9  
2070     True        1.00      0.68      0.81      19  
2071  
2072   micro avg     0.79      0.79      0.79      28  
2073   macro avg     0.80      0.84      0.78      28  
2074 weighted avg    0.87      0.79      0.79      28  
2075  
2076 Precision = 1.0  
2077 Rappel= 0.6842105263157895  
2078 F_Mesure= 0.8125000000000001  
2079 Naive Bayes dataset/HBR/taken/  
    HBR_RandomUnderSampler_default.csv Validation_70/30 0.  
    8125000000000001  
2080  
2081 -----Using cross Validation  
-----  
2082  
2083 -----k= 1 -----  
2084  
2085 F_Mesure= 1.0  
2086
```

```
2087 -----k= 2 -----
2088
2089 F_Mesure= 0.9473684210526316
2090
2091 -----k= 3 -----
2092
2093 F_Mesure= 0.888888888888889
2094
2095 -----k= 4 -----
2096
2097 F_Mesure= 0.933333333333333
2098
2099 -----k= 5 -----
2100
2101 F_Mesure= 1.0
2102 F_Mesures moyenne = 0.9539181286549707
2103
2104
2105
2106 -----SVM-----
2107
2108
2109
2110 dataset/HBR/taken/HBR_TomekLinks.csv
2111
2112
2113
2114 count         96
2115 unique        2
2116 top          False
2117 freq          50
2118 Name: is_code_smell, dtype: object
2119 [[14  0]
2120 [ 3 12]]
2121 the recall for this model is : 0.8
2122 TP 12
2123 TN 14
2124 FP 0
2125 FN 3
2126 C:\Users\My PC\AppData\Local\Programs\Python\Python36-32
    \lib\site-packages\matplotlib\pyplot.py:537:
    RuntimeWarning: More than 20 figures have been opened.
    Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
    closed and may consume too much memory. (To control
```

```
2126 this warning, see the rcParam `figure.max_open_warning
   `).
2127     max_open_warning, RuntimeWarning)
2128
2129 -----Classification Report
-----
2130             precision      recall    f1-score   support
2131
2132     False          0.82      1.00      0.90       14
2133     True           1.00      0.80      0.89       15
2134
2135     micro avg      0.90      0.90      0.90       29
2136     macro avg      0.91      0.90      0.90       29
2137 weighted avg      0.91      0.90      0.90       29
2138
2139 Precision = 1.0
2140 Rappel= 0.8
2141 F_Mesure= 0.88888888888889
2142 Naive Bayes dataset/HBR/taken/HBR_TomekLinks.csv
    Validation_70/30 0.88888888888889
2143
2144 -----Using cross Validation
-----
2145
2146 -----k= 1 -----
2147
2148 F_Mesure= 1.0
2149
2150 -----k= 2 -----
2151
2152 F_Mesure= 0.9565217391304348
2153
2154 -----k= 3 -----
2155
2156 F_Mesure= 0.9600000000000001
2157
2158 -----k= 4 -----
2159
2160 F_Mesure= 0.7499999999999999
2161
2162 -----k= 5 -----
2163
2164 F_Mesure= 1.0
2165 F_Mesures moyenne = 0.933304347826087
2166
```

```
2167 Process finished with exit code 0
```

```
2168
```