

Decision Tree Algorithm

Prajwal Vijaywargya - 2017B3A70954H
Siddhi Mahesh Burse - 2017B3A70972H
Parth Krishna Sharma - 2017B3A70907H

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib inline
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

In [2]: train_df = pd.read_csv('D:/BITS/DL/Assignments/Assignment1/mnist_train.csv')
test_df = pd.read_csv('D:/BITS/DL/Assignments/Assignment1/mnist_test.csv')
```

```
In [3]: x_train,y_train = train_df.iloc[:,1:].values, train_df.iloc[:,0].values
x_test,y_test = test_df.iloc[:,1:].values, test_df.iloc[:,0].values
```

```
In [4]: clf = DecisionTreeClassifier(criterion='entropy',min_samples_leaf=2,min_weight_fraction_leaf=0.0001,max
_depth=15,min_samples_split=4)
```

```
In [5]: clf.fit(x_train,y_train)
```

```
Out [5]: DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='entropy',
                                max_depth=15, max_features=None, max_leaf_nodes=None,
                                min_impurity_decrease=0.0, min_impurity_split=None,
                                min_samples_leaf=2, min_samples_split=4,
                                min_weight_fraction_leaf=0.0001, presort='deprecated',
                                random_state=None, splitter='best')
```

```
In [6]: prediction_test = clf.predict(x_test)
```

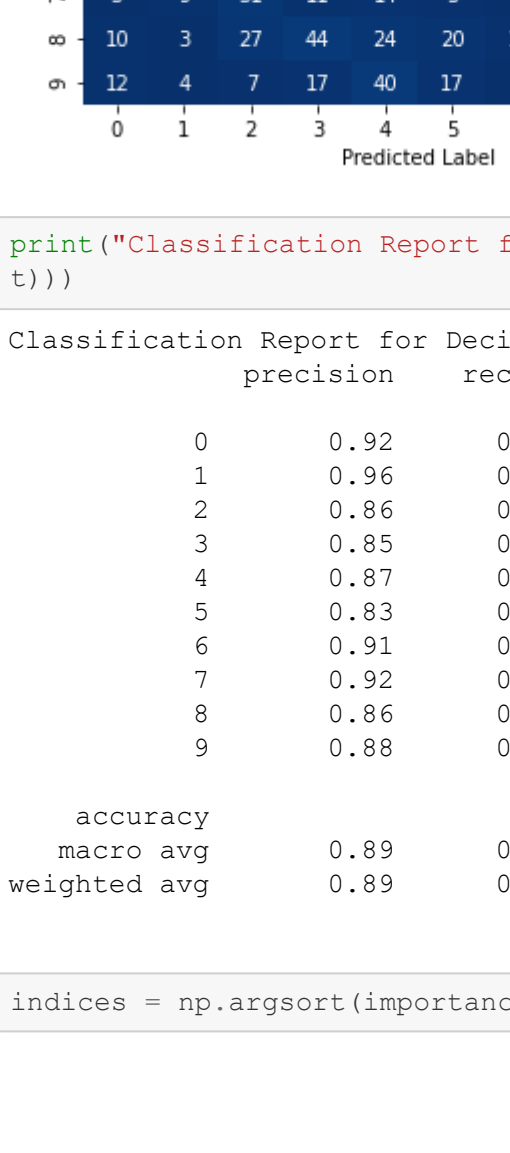
```
In [7]: acc = accuracy_score(y_test,prediction_test)
acc
```

```
Out [7]: 0.8874
```

```
In [8]: importances = clf.feature_importances_
ft = clf.feature importances
ft = ft*25500
plt.imshow(ft.reshape((28,28)))
plt.title("Feature Importances learnt via Decision Tree\ncriterion: Entropy,\nTest Accuracy={}%".format
(acc*100))
```

```
Out [8]: Text(0.5, 1.0, 'Feature Importances learnt via Decision Tree\ncriterion: Entropy,\nTest Accuracy=88.74%')
```

Feature Importances learnt via Decision Tree
criterion: Entropy,
Test Accuracy=88.74%



```
In [9]: sns.heatmap(confusion_matrix(y_test,prediction_test),cmap="Blues_r",fmt='',annot=True,cbar=False)
plt.xlabel("Correct Label")
plt.ylabel("Predicted Label")
plt.title("Confusion Matrix for Decision Tree\nModel Accuracy={}%".format(acc*100))
```

```
Out [9]: Text(0.5, 1.0, 'Confusion Matrix for Decision Tree\nModel Accuracy=88.74%')
```



```
In [10]: print("Classification Report for Decision Tree:\n").format(classification_report(y_test,prediction_test))
```

Classification Report for Decision Tree:				
	precision	recall	f1-score	support
0	0.92	0.93	0.93	980
1	0.96	0.97	0.97	1135
2	0.86	0.89	0.87	1032
3	0.85	0.85	0.85	1010
4	0.87	0.89	0.88	982
5	0.83	0.85	0.84	892
6	0.91	0.89	0.90	958
7	0.92	0.90	0.91	1028
8	0.86	0.82	0.84	974
9	0.86	0.86	0.87	1009
accuracy			0.89	10000
macro avg	0.89	0.89	0.89	10000
weighted avg	0.89	0.89	0.89	10000

```
In [11]: indices = np.argsort(importances)[::-1]
```



```
122: for i in range(1, len_shape[1]):
123:     print("Feature (", i, ") = {}".format(i+1, indices[F], importances[indices[F]]))
124:
1 Feature 461 : 0.1005771861662746
2 Feature 378 : 0.03971260146917099
3 Feature 401 : 0.0721154989787366
4 Feature 121 : 0.000149213523202164
5 Feature 597 : 0.035294919505361849
6 Feature 319 : 0.02547809417570173
7 Feature 155 : 0.000149213523202164
8 Feature 155 : 0.0232742105964107
9 Feature 376 : 0.02168575872188201
10 Feature 106 : 0.0001746754071280006
11 Feature 153 : 0.000149213523202164
12 Feature 271 : 0.0015783498817472343
13 Feature 514 : 0.01435680920304245
14 Feature 156 : 0.000149213523202164
15 Feature 515 : 0.01312016344633828
16 Feature 398 : 0.012145224673141193
17 Feature 123 : 0.000149213523202164
18 Feature 297 : 0.0107166254951924953
19 Feature 428 : 0.00963032336427893
20 Feature 185 : 0.0091599367170132
21 Feature 153 : 0.000149213523202164
22 Feature 657 : 0.0008063163835953811
23 Feature 294 : 0.0077380869323691
24 Feature 156 : 0.000149213523202164
25 Feature 267 : 0.00766210653236479
26 Feature 596 : 0.00071743775460348
27 Feature 174 : 0.000149213523202164
28 Feature 401 : 0.000149213523202164
29 Feature 487 : 0.000654311939915202
30 Feature 354 : 0.00615003654545412
31 Feature 156 : 0.000149213523202164
32 Feature 156 : 0.0055202383681823
33 Feature 327 : 0.005258228912766
34 Feature 156 : 0.000149213523202164
35 Feature 399 : 0.00497838598694417
36 Feature 300 : 0.000482859086495225
37 Feature 137 : 0.00476795572889959
38 Feature 153 : 0.000149213523202164
39 Feature 351 : 0.0047384885849104915
40 Feature 143 : 0.0041388717370578815
41 Feature 156 : 0.000149213523202164
42 Feature 458 : 0.003947534091642139
43 Feature 241 : 0.00373267133747403
44 Feature 123 : 0.000149213523202164
45 Feature 432 : 0.003545774311768565
46 Feature 247 : 0.0035295823767047857
47 Feature 185 : 0.0031022643795267
48 Feature 401 : 0.000149213523202164
49 Feature 487 : 0.000654311939915202
50 Feature 382 : 0.00316924031149772
51 Feature 156 : 0.000149213523202164
52 Feature 256 : 0.003042252931270926
53 Feature 404 : 0.00296921591387746
54 Feature 153 : 0.000149213523202164
55 Feature 401 : 0.000149213523202164
56 Feature 296 : 0.002679341857232007
57 Feature 367 : 0.0026148792422313032
58 Feature 156 : 0.000149213523202164
59 Feature 456 : 0.00255080450519805
60 Feature 491 : 0.0024958046216214913
61 Feature 156 : 0.000149213523202164
62 Feature 403 : 0.00243545425571405
63 Feature 516 : 0.00243852094172595
64 Feature 157 : 0.002382193752047956
65 Feature 657 : 0.0023783579646012
66 Feature 292 : 0.0022857945040019328
67 Feature 149 : 0.00227320219685501
68 Feature 156 : 0.000149213523202164
69 Feature 405 : 0.00222532106329272
70 Feature 384 : 0.0022203797714281056
71 Feature 185 : 0.002172837612696584
72 Feature 431 : 0.002169515701758858
73 Feature 318 : 0.002159045020663236
74 Feature 347 : 0.0021342753976042
75 Feature 156 : 0.000149213523202164
76 Feature 384 : 0.0021259102911718716
77 Feature 214 : 0.002090872178717447
78 Feature 156 : 0.000149213523202164
79 Feature 658 : 0.002068940133890967
80 Feature 293 : 0.0020468081772430885
81 Feature 137 : 0.0020362921635252
82 Feature 247 : 0.0020384453285962059
83 Feature 216 : 0.0020365645032990416
84 Feature 353 : 0.0020340093949763
85 Feature 156 : 0.000149213523202164
86 Feature 212 : 0.001973536164877704
87 Feature 157 : 0.00194443284968176
88 Feature 156 : 0.000149213523202164
89 Feature 427 : 0.001902187994467654
90 Feature 708 : 0.0018799337158836144
91 Feature 127 : 0.001871769028817211
92 Feature 156 : 0.000149213523202164
93 Feature 328 : 0.0018239632379064221
94 Feature 155 : 0.00176609496247988
95 Feature 156 : 0.000149213523202164
96 Feature 180 : 0.0017559165020850018
97 Feature 629 : 0.0017390284894184654
98 Feature 156 : 0.000149213523202164
99 Feature 97 : 0.00166821024519031
100 Feature 577 : 0.00164592119335201
101 Feature 154 : 0.001644639243152025
102 Feature 619 : 0.0016431975379846012
103 Feature 684 : 0.0015486561831985143
104 Feature 598 : 0.001541990068284536
105 Feature 156 : 0.000149213523202164
106 Feature 436 : 0.0015077135424906086
107 Feature 322 : 0.001502177055040379
108 Feature 684 : 0.00149213523202164
109 Feature 247 : 0.00147053285962059
110 Feature 459 : 0.001448379035907015
111 Feature 518 : 0.0014400717304716448
112 Feature 156 : 0.000149213523202164
113 Feature 407 : 0.001438399557278681
114 Feature 244 : 0.001434243021653422
115 Feature 435 : 0.0014239946005990519
116 Feature 217 : 0.0013695921854432662
117 Feature 539 : 0.001347232978964597
118 Feature 247 : 0.00134053285962059
119 Feature 358 : 0.0013366573684350299
120 Feature 601 : 0.0013168062719285095
121 Feature 594 : 0.0013154884632835655
122 Feature 158 : 0.00124465952940045
123 Feature 320 : 0.0012390796018748667
124 Feature 156 : 0.000149213523202164
125 Feature 434 : 0.001205268086797036
126 Feature 567 : 0.001191541160739032
127 Feature 538 : 0.001173215046312993
128 Feature 619 : 0.001171025701920861
129 Feature 338 : 0.0011708601235371716
130 Feature 178 : 0.0011556720211346443
131 Feature 156 : 0.000149213523202164
132 Feature 570 : 0.0011183980819086044
133 Feature 570 : 0.0011183980819086044
134 Feature 247 : 0.001115851763484304
135 Feature 205 : 0.001115125845651991
136 Feature 247 : 0.00110759651788228
137 Feature 619 : 0.0011070260471189712
138 Feature 437 : 0.0010712375220204983
139 Feature 156 : 0.000149213523202164
140 Feature 185 : 0.0010518341184102085
141 Feature 546 : 0.0010409421697377106
142 Feature 156 : 0.000149213523202164
143 Feature 187 : 0.0010175409038678048
144 Feature 182 : 0.000998439633125611
145 Feature 543 : 0.000993071204400391
146 Feature 649 : 0.00098196763138358
147 Feature 457 : 0.0009428176898703322
148 Feature 521 : 0.00093813405589381
149 Feature 156 : 0.000149213523202164
150 Feature 489 : 0.00096094360267253
151 Feature 316 : 0.000903342147186314
152 Feature 634 : 0.000871532671932529
153 Feature 660 : 0.000862943765672225
154 Feature 545 : 0.00086227593500533
155 Feature 329 : 0.0008604712711258852
156 Feature 156 : 0.000149213523202164
157 Feature 510 : 0.00085205584002638
158 Feature 352 : 0.00084368505961555
159 Feature 156 : 0.000149213523202164
160 Feature 410 : 0.0008321607765363498
161 Feature 324 : 0.0008113368560203117
162 Feature 406 : 0.0008078175274985205
163 Feature 156 : 0.000149213523202164
164 Feature 613 : 0.0007619054195175124
165 Feature 127 : 0.00075939151345074
166 Feature 729 : 0.000758452106484224
167 Feature 468 : 0.000751934697405579
168 Feature 263 : 0.0007498751214650528
169 Feature 272 : 0.000746252976947832
170 Feature 268 : 0.0007436978836106149
171 Feature 152 : 0.0007334050397852868
172 Feature 150 : 0.0007311334552445099
173 Feature 247 : 0.0007296301549807015
174 Feature 356 : 0.000711072657794571
175 Feature 602 : 0.000713677373735871
176 Feature 156 : 0.000149213523202164
177 Feature 484 : 0.000670621017652629
178 Feature 125 : 0.0006510818421598735
179 Feature 108 : 0.0006316448815782842
180 Feature 247 : 0.000629439631054121
181 Feature 355 : 0.0006285296260592518
182 Feature 374 : 0.00062752467185369
183 Feature 759 : 0.0006254504047115898
184 Feature 624 : 0.000622088532427155
185 Feature 464 : 0.000610828450843337
186 Feature 156 : 0.000149213523202164
187 Feature 350 : 0.000573604064253225
188 Feature 429 : 0.0005681018095416924
189 Feature 261 : 0.0005661305275652021
190 Feature 513 : 0.00056553117975641
191 Feature 315 : 0.0005630282194180383
192 Feature 302 : 0.000558970224845637
193 Feature 739 : 0.00055839350429252
194 Feature 343 : 0.000534606719413778
195 Feature 603 : 0.000532397665462265
196 Feature 156 : 0.000149213523202164
197 Feature 378 : 0.0005207582814824337
198 Feature 126 : 0.000516297705536151
199 Feature 569 : 0.0005153227904646534
200 Feature 489 : 0.0005148328517885012
201 Feature 574 : 0.0005116557006293862
202 Feature 520 : 0.000507669384046368
203 Feature 549 : 0.00050580813250694
204 Feature 483 : 0.0004992918011514581
205 Feature 235 : 0.0004949271537361037
206 Feature 492 : 0.0004904692576952815
207 Feature 213 : 0.00047094142643598813
208 Feature 238 : 0.000468611879340568
209 Feature 190 : 0.000468206141782759
210 Feature 649 : 0.000467512082935887
211 Feature 220 : 0.000467271910325663
212 Feature 604 : 0.00046757664209903617
213 Feature 156 : 0.000149213523202164
214 Feature 299 : 0.00046172302667101347
215 Feature 207 : 0.00044713221281403275
216 Feature 409 : 0.0004467773313453805
217 Feature 233 : 0.00044579537986369
218 Feature 681 : 0.0004364313311673939
219 Feature 634 : 0.000436524571435377
220 Feature 169 : 0.0004342849767016
221 Feature 575 : 0.00042901361348541834
222 Feature 454 : 0.00041793317550689205
223 Feature 156 : 0.000149213523202164
224 Feature 410 : 0.00041074208110177245
225 Feature 494 : 0.0004097136370667303
226 Feature 573 : 0.00040620373024042916
227 Feature 390 : 0.00040538612538020285
228 Feature 511 : 0.0004032734913913545
229 Feature 385 : 0.00040281202219586334
230 Feature 156 : 0.000149213523202164
231 Feature 237 : 0.00039793318898601726
232 Feature 123 : 0.0003881307297708064
233 Feature 132 : 0.0003862168699193266
234 Feature 397 : 0.000384109913580237
235 Feature 689 : 0.000375286693073826
236 Feature 438 : 0.0003752320783629515
237 Feature 156 : 0.000149213523202164
238 Feature 400 : 0.00036574273522272
239 Feature 402 : 0.00036294030113354666
240 Feature 156 : 0.000149213523202164
241 Feature 129 : 0.0003454215089029819
242 Feature 342 : 0.000344733480591251
243 Feature 482 : 0.000342702284432984
244 Feature 649 : 0.00034174564508259
245 Feature 582 : 0.0003339165339912495
246 Feature 260 : 0.0003333617720395895
247 Feature 739 : 0.000330594313242476
248 Feature 156 : 0.000149213523202164
249 Feature 688 : 0.0003286553695768854
250 Feature 156 : 0.000149213523202164
251 Feature 378 : 0.00032634216438569146
252 Feature 570 : 0.00032619363515735
253 Feature 370 : 0.0003175439676187797
254 Feature 330 : 0.000316857681304672
255 Feature 439 : 0.0003109708431193813
256 Feature 291 : 0.000310535841987066
257 Feature 499 : 0.0003045203121194689
258 Feature 630 : 0.0002992335040698016
259 Feature 156 : 0.000149213523202164
260 Feature 245 : 0.0002931109643698913
261 Feature 493 : 0.0002905228185601597
262 Feature 376 : 0.000287755826349794
263 Feature 709 : 0.000286525311867862
264 Feature 716 : 0.0002873516556310214
265 Feature 373 : 0.000284123356475517
266 Feature 739 : 0.000283728659203295
267 Feature 488 : 0.0002631931614444705
268 Feature 156 : 0.000149213523202164
269 Feature 687 : 0.000255200489168808
270 Feature 377 : 0.0002524682331879799
271 Feature 360 : 0.0002526522461097215
272 Feature 561 : 0.000248334175786257
272 Feature 576 : 0.00024818645735327095
273 Feature 151 : 0.0002381670987336016
274 Feature 156 : 0.000149213523202164
275 Feature 219 : 0.0002297339632684396
276 Feature 717 : 0.000225933473144498
277 Feature 363 : 0.00022096813752528975
278 Feature 597 : 0.0002184440546798192
279 Feature 497 : 0.00021424211230262
280 Feature 720 : 0.000207396735588243
281 Feature 539 : 0.000205785451929002
282 Feature 273 : 0.0002015879528069405
283 Feature 549 : 0.00019847380502129745
284 Feature 156 : 0.000149213523202164
285 Feature 664 : 0.00018386873291425916
286 Feature 677 : 0.0001838980661216595
287 Feature 568 : 0.00018247312168718193
288 Feature 659 : 0.000180453796850126
289 Feature 133 : 0.000176533150355218
290 Feature 548 : 0.00017709847361580336
291 Feature 156 : 0.000149213523202164
292 Feature 627 : 0.0001662439402285325
293 Feature 232 : 0.00016366382052065
294 Feature 150 : 0.000162831343397654
295 Feature 274 : 0.00015961390196279487
296 Feature 411 : 0.0001575673879432253
297 Feature 262 : 0.000154623450752575
298 Feature 561 : 0.00014858646176035778
299 Feature 301 : 0.00014768232816415893
300 Feature 635 : 0.00014760644737532788
301 Feature 181 : 0.000147391361832065
302 Feature 462 : 0.00014655966990464
303 Feature 611 : 0.00014516467541588382
304 Feature 639 : 0.00014474746252448198
305 Feature 459 : 0.0001444455253018702
306 Feature 130 : 0.0001405291759275455
307 Feature 419 : 0.0001371653661479607
308 Feature 509 : 0.000135135999155305
309 Feature 605 : 0.000130835566590511
310 Feature 500 : 0.0001291229093064502
311 Feature 465 : 0.000128104844945012
312 Feature 469 : 0.00012612534341893
313 Feature 606 : 0.000124005591333609
314 Feature 181 : 0.000122715729084077
315 Feature 671 : 0.00011737511857555
316 Feature 522 : 0.0001067959225612392
317 Feature 524 : 0.0001060804333910503
318 Feature 619 : 0.0001044353287502116
319 Feature 181 : 0.00010188843096817
320 Feature 341 : 9.826173815615303e-05
321 Feature 156 : 0.000149213523202164
322 Feature 259 : 9.6074342657131e-05
323 Feature 579 : 9.358978958989892e-05
324 Feature 652 : 9.23652028293439e-05
325 Feature 661 : 9.16148525510149e-05
326 Feature 98 : 9.104252927262319e-05
327 Feature 568 : 8.486271679160705e-05
328 Feature 156 : 0.000149213523202164
329 Feature 608 : 8.07006317634842e-05
330 Feature 228 : 7.79733123415233e-05
331 Feature 156 : 0.000149213523202164
332 Feature 567 : 7.7375133488578e-05
333 Feature 717 : 7.298489959392481e-05
334 Feature 712 : 7.11873064131002e-05
335 Feature 678 : 6.8322897532344e-05
336 Feature 509 : 6.434524175345378e-05
337 Feature 537 : 5.833015687412341e-05
338 Feature 715 : 5.19321564746476e-05
339 Feature 303 : 5.085178001547959e-05
340 Feature 706 : 4.698777342170635e-05
341 Feature 230 : 4.542514713740131e-05
342 Feature 651 : 4.255730111481394e-05
343 Feature 75 : 4.144134015800583e-05
344 Feature 425 : 3.70268689509823e-05
345 Feature 259 : 3.7107550578628e-05
346 Feature 563 : 3.148013561160946e-05
347 Feature 287 : 2.91110980884049e-05
348 Feature 602 : 2.90951691609194e-05
349 Feature 689 : 2.37894331309255e-05
350 Feature 286 : 2.2307276011617e-05
351 Feature 181 : 1.9373792495745914e-05
352 Feature 671 : 1.90134017555183607e-05
353 Feature 387 : 1.8603238061194635e-05
354 Feature 628 : 1.3313349140808286e-05
355 Feature 659 : 1.210190925528151e-05
356 Feature 496 : 8.65727263223689e-06
357 Feature 523 : 8.832209195365965e-06
358 Feature 222 : 0.0
359 Feature 227 : 0.0
360 Feature 221 : 0.0
361 Feature 200 : 0.0
362 Feature 159 : 0.0
363 Feature 202 : 0.0
364 Feature 250 : 0.0
365 Feature 229 : 0.0
366 Feature 249 : 0.0
367 Feature 248 : 0.0
368 Feature 228 : 0.0
369 Feature 671 : 0.0
370 Feature 224 : 0.0
371 Feature 229 : 0.0
372 Feature 189 : 0.0
373 Feature 199 : 0.0
374 Feature 49 : 0.0
375 Feature 156 : 0.000149213523202164
376 Feature 55 : 0.0
377 Feature 54 : 0.0
378 Feature 53 : 0.0
379 Feature 59 : 0.0
380 Feature 51 : 0.0
381 Feature 50 : 0.0
382 Feature 48 : 0.0
383 Feature 58 : 0.0
384 Feature 47 : 0.0
385 Feature 46 : 0.0
386 Feature 45 : 0.0
387 Feature 44 : 0.0
388 Feature 43 : 0.0
389 Feature 42 : 0.0
390 Feature 41 : 0.0
391 Feature 57 : 0.0
392 Feature 56 : 0.0
393 Feature 39 : 0.0
394 Feature 69 : 0.0
395 Feature 77 : 0.0
396 Feature 76 : 0.0
397 Feature 74 : 0.0
398 Feature 73 : 0.0
399 Feature 72 : 0.0
400 Feature 71 : 0.0
401 Feature 70 : 0.0
402 Feature 69 : 0.0
403 Feature 68 : 0.0
404 Feature 67 : 0.0
405 Feature 66 : 0.0
406 Feature 65 : 0.0
407 Feature 64 : 0.0
408 Feature 63 : 0.0
409 Feature 62 : 0.0
410 Feature 61 : 0.0
411 Feature 40 : 0.0
412 Feature 13 : 0.0
413 Feature 198 : 0.0
414 Feature 9 : 0.0
415 Feature 16 : 0.0
416 Feature 15 : 0.0
417 Feature 14 : 0.0
418 Feature 13 : 0.0
419 Feature 12 : 0.0
420 Feature 11 : 0.0
421 Feature 10 : 0.0
422 Feature 8 : 0.0
423 Feature 6 : 0.0
424 Feature 5 : 0.0
425 Feature 6 : 0.0
426 Feature 5 : 0.0
427 Feature 4 : 0.0
428 Feature 3 : 0.0
429 Feature 2 : 0.0
430 Feature 1 : 0.0
431 Feature 17 : 0.0
432 Feature 19 : 0.0
433 Feature 20 : 0.0
434 Feature 29 : 0.0
435 Feature 36 : 0.0
436 Feature 32 : 0.0
437 Feature 34 : 0.0
438 Feature 32 : 0.0
439 Feature 33 : 0.0
440 Feature 31 : 0.0
441 Feature 30 : 0.0
442 Feature 28 : 0.0
443 Feature 29 : 0.0
444 Feature 27 : 0.0
445 Feature 26 : 0.0
446 Feature 25 : 0.0
447 Feature 24 : 0.0
448 Feature 23 : 0.0
449 Feature 22 : 0.0
450 Feature 2
```