Comparison of Vehicle Silhouette Classification Models with Weka

Ryan M. Walker

Introduction

To better understand machine learning classification techniques, problems and applications, this report will compare four classification models using Weka. The dataset I have chosen to accomplish this was developed by JP Siebert at the Turing Institute in Glasgow, Scotland: “Vehicle Silhouettes.” My own interest in visual computation and analysis is the primary motivation for working with this data set. It is a clear representation of the detailed attributes needed to quantify visual objects for machine interpretation, pattern recognition, and categorization. This dataset was originally developed as a means of classifying three-dimensional objects using two-dimensional images of an object’s silhouette via supervised learning. This report will detail the process and results of classifying this dataset with four machine learning algorithms via Weka: Naïve Bayes, J48, IBk, and OneR. Through comparisons of their respective accuracy, confusion matrices, and error rates, it will be possible to show which of the four models performs best as they relate to the vehicle dataset.

Classification Process

This dataset specifically seeks to categorize a vehicle into one of four classes (i.e. Opel, Saab, Bus, or Van) using measurements obtained from silhouette images of models for these vehicles. The training dataset consists of 846 total instances (Opel - 212, Saab - 217, Bus - 218, Van - 199) of features calculated using a Hierarchical Image Processing System. Scale independent features where extrapolated from these images using classical moments-based and heuristic measurements. These features make up the first 18 attributes of the data set, with the last attribute representing the vehicle class:

1. Compactness
2. Circularity
3. Distance Circularity
4. Radius Ratio
5. PR Axis Aspect Ratio
6. Max Length Aspect Ratio
7. Scatter Ratio
8. Elongatedness
9. PR Axis Rectangularity
10. Max Length Rectangularity
11. Scaled Variance Major
12. Scaled Variance Minor
13. Scaled Radius of Gyration
14. Skewness About_Major
15. Skewness About_Minor
16. Kurtosis About_Major
17. Kurtosis About_Minor
18. Hollows Ratio
19. Class

The class attribute is of most interest to this dataset. The first 18 attributes will be used in the classification model to predict which of the fours categories a vehicle belongs to. To process data in Weka, it must first be cleaned and formatted properly for as a .CSV or .ARFF file. This dataset provided by the UCI Machine Learning Repository was formatted and ready to process in Weka prior to my download. However, I did need to discretize the first 18 attributes since they were real-valued, numerical attributes. To discretize the attributes, I used default discretize options selected under the supervise folder of the filter after the dataset was loaded into Weka. (Figure 1) After discretizing the dataset, I processed the data once for each of the four classifiers used in this report. (Figures 2) I ran Naïve Bayes, J48, IBk (k=3,5, & 7), and OneR on the data with default settings, cross validation (Folds 10) for this report. (Figure 3, see Appendix for Weka output) Each classifier will be evaluated by its accuracy, overall rate of error and mean absolute error, and its confusion matrix.

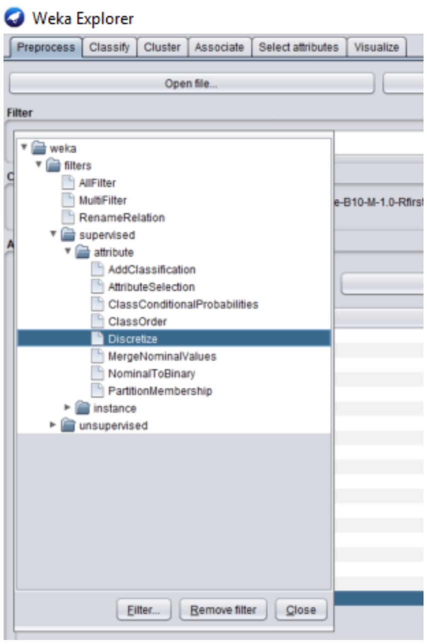

FIGURE 1 SHOWING DISCRETIZATION SELECTION FOR VEHICLE DATASET IN WEKA

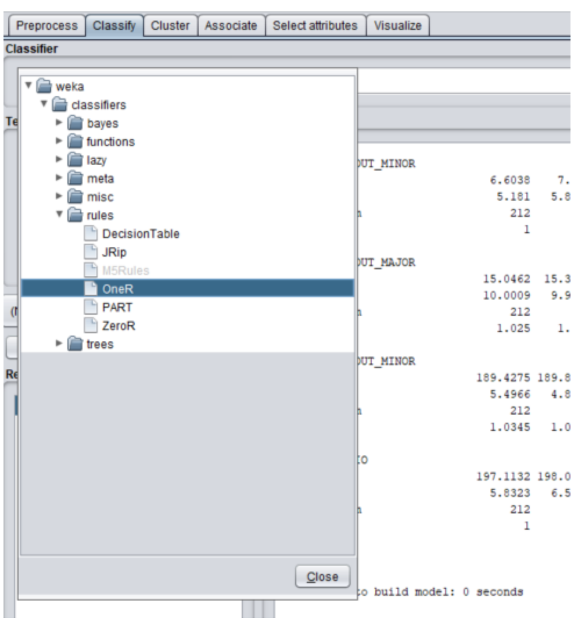

FIGURE 2 SHOWING CLASSIFIER SELECTION IN WEKA

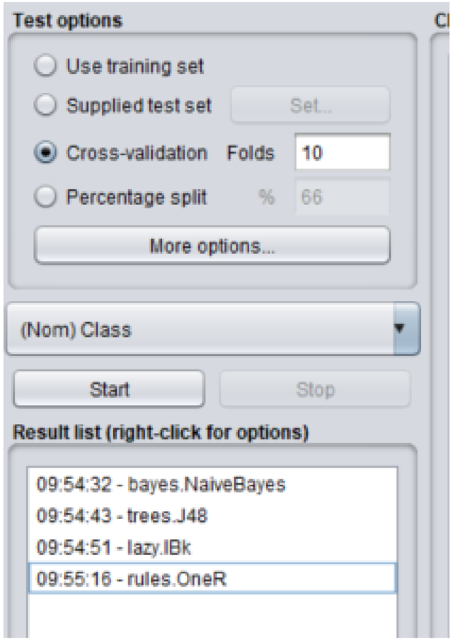

Figure 3 Showing test options selections for Weka classifiers

Weka Results

Naïve Bayes

The Naïve Bayes classifier is a probabilistic algorithm that uses Bayes' Theorem with independent, unrelated features. Naïve Bayes assumes these features do not depend on each other, yet all contribute the probability of belonging to a specific class. In the case of the vehicle dataset, although the features (e.g. the various measurements) of the models may depend on the existence of other features, they contribute to the resulting probability of a vehicle belonging to Opel, Saab, Bus, or Van independently of one another in the algorithm. (Obuandike, Isah, & Alhasan, 2015)

The Naïve Bayes classifier correctly classified 530 instances, at an accuracy rate of 62.6%, overall error rate of 37.4%, and mean absolute error 0.1897. (Table 1, Figure 4, Figure 5) The decision matrix shows the correct and incorrect classifications of each class in the dataset. (Table 2) As shown in the table and corresponding graph (Figure 6), classification accuracy for Opel was 34.9%, Naïve Bayes accuracy for "Van" is quite accurate at 97.5%. However, the distribution of the accuracy rate for each class brings the mean overall accuracy rate to 62.6%.

NAÏVE BAYES			
	Instances	**Rate (%)**	**Mean Absolute Error**
Correct	530	62.6%	0.1897
Incorrect	316	37.4%	

Table 1 Showing summary data for Naive Bayes

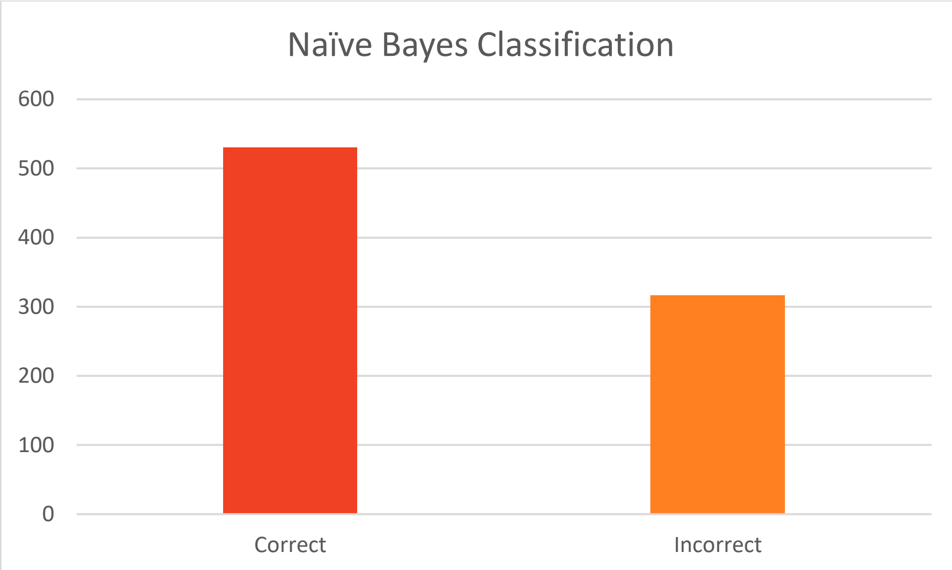

FIGURE 4 SHOWING CORRECTLY AND INCORRECTLY CLASSIFIED INSTANCES FOR NAÏVE BAYES

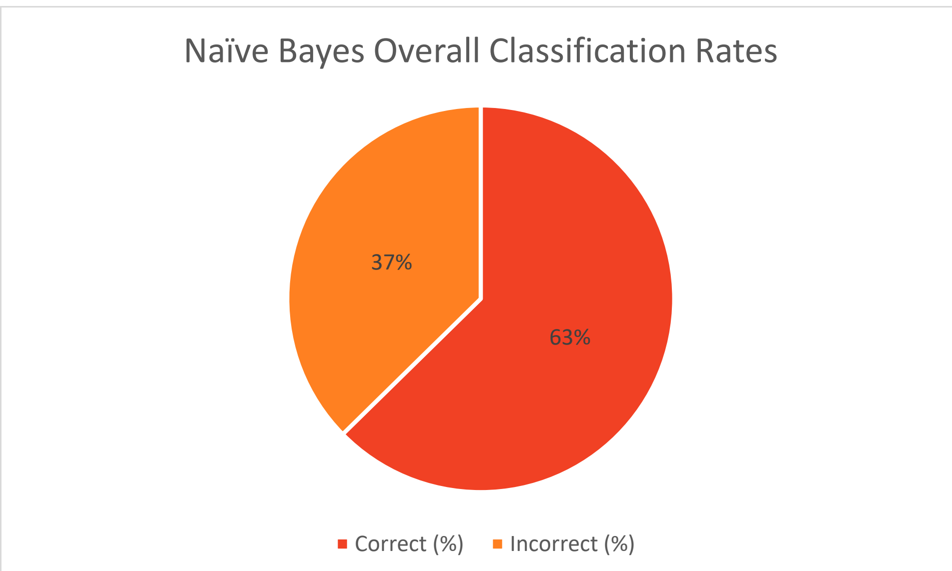

FIGURE 5 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF NAÏVE BAYES

NAÏVE BAYES						
CONFUSION MATRIX		PREDICTED				ACTUAL TOTAL
		OPEL	SAAB	BUS	VAN	
ACTUAL	OPEL	74	74	13	51	212
	SAAB	42	105	16	54	217
	BUS	29	23	157	9	218
	VAN	0	0	5	194	199
PREDICTED TOTAL		145	202	191	308	846

TABLE 2 SHOWING CONFUSION MATRIX OF NAÏVE BAYES

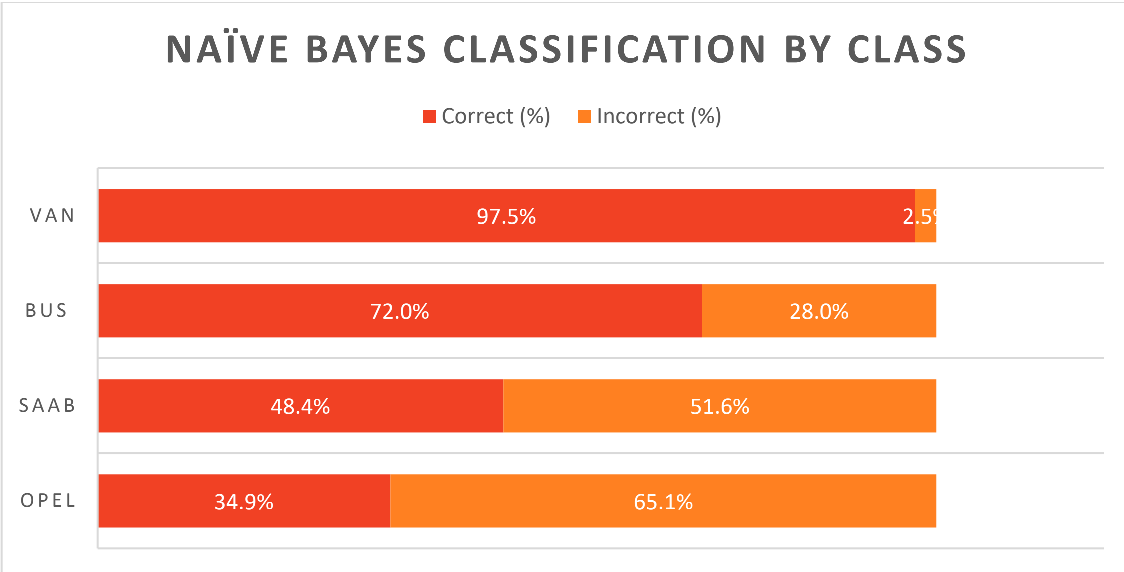

FIGURE 6 SHOWING NAÏVE BAYES RATES OF CLASSIFICATION RATES BY CLASS

J48

The J48 classifier is a decision tree classifier that builds its model using a tree structure with the following components: testing node, start node, terminal node, and branches. J48 builds the decision tree with a depth-first search process from the dataset using information gain to examine attributes and split data. The algorithm recursively executes on smaller subsets until all instances of a subset belong to the same class. The terminal node is then set in the decision tree to choose that class. J48 considers all possible tests that can split the dataset and selects a test with the best information gain. (Solanki, 2014)

The J48 classifier correctly classified 609 instances, at an accuracy rate of 72%, overall error rate of 28%, and mean absolute error 0.1553. (Table 3, Figure 7, Figure 8) The decision matrix shows the correct and incorrect classifications of each class in the dataset. (Table 4) Classification accuracy was highest for “Bus” at 90.4%, followed by “Van” at 85.4%. (Figure 9)

J48			
	Instances	**Rate (%)**	**Mean Absolute Error**
Correct	609	72.0%	0.1553
Incorrect	237	28.0%	

TABLE 3 SHOWING SUMMARY DATA FOR J48

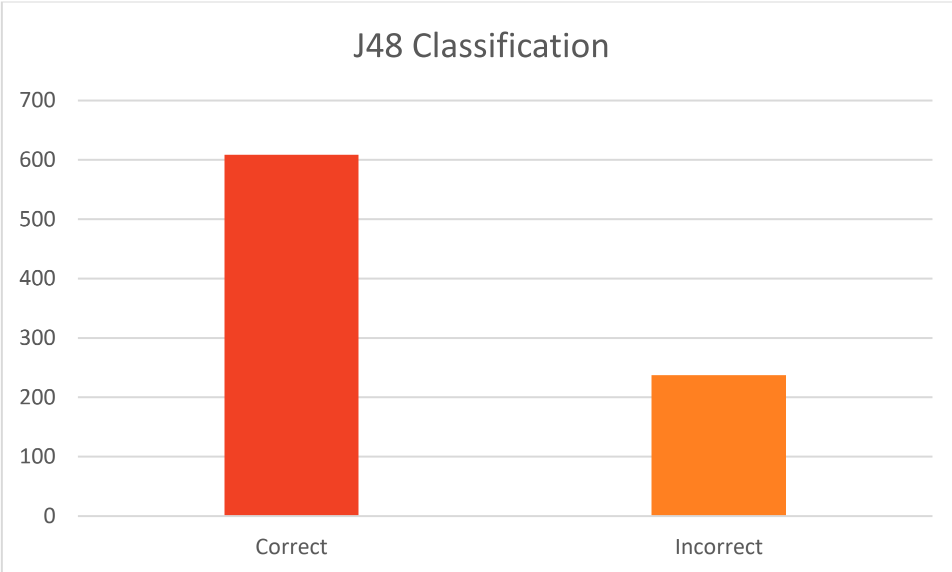

FIGURE 7 SHOWING CORRECTLY AND INCORRECTLY CLASSIFIED INSTANCES FOR J48

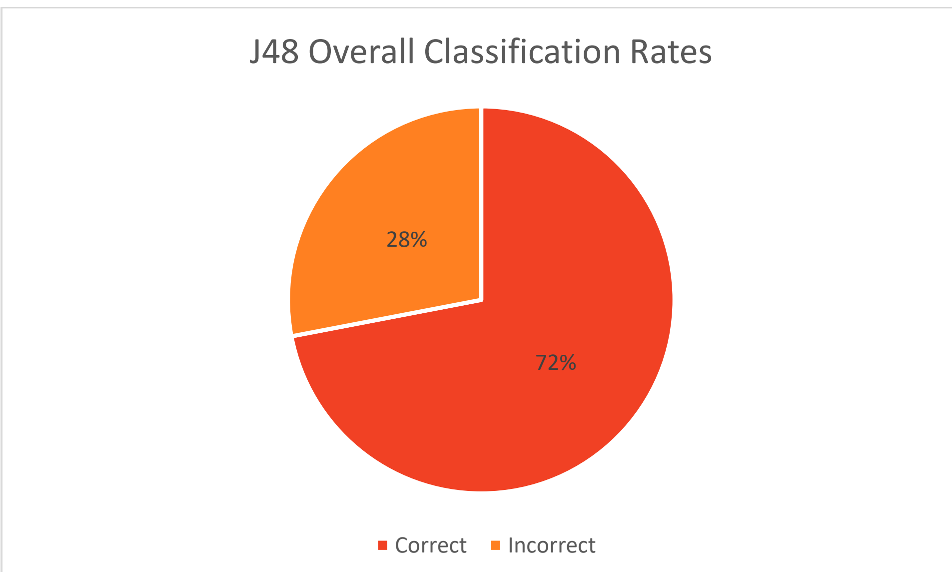

FIGURE 8 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF J48

J48						
CONFUSION MATRIX		**PREDICTED**				**ACTUAL TOTAL**
		OPEL	**SAAB**	**BUS**	**VAN**	
ACTUAL	**OPEL**	123	70	9	10	212
	SAAB	72	119	9	17	217
	BUS	9	6	197	6	218
	VAN	7	15	7	170	199
PREDICTED TOTAL		211	210	222	203	846

TABLE 4 SHOWING CONFUSION MATRIX FOR J48

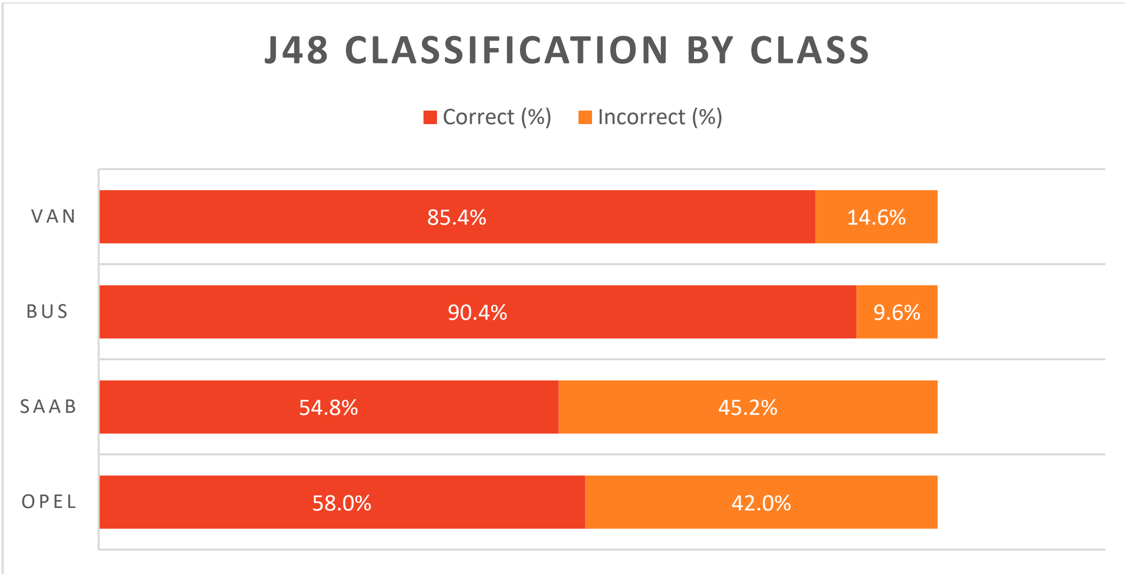

FIGURE 9 SHOWING J48 CLASSIFICATION RATES BY CLASS

IBk

The IBk classifier is an instance-based k-nearest neighbor classifier. This algorithm uses previous known data instances to predict unknown output data of a new data instance. A k-nearest neighbor classifier selects an appropriate value for k based on cross validation, or it is manually set as in this case with k=3,5, or 7. The algorithm then does distance weighting using a distance measure to find the training instance closest to a test instance, and then predicts the same class as the training instance using majority vote of the k-nearest neighbors (Mwagha, Muthoni, & Ochieg, 2014)

K = 3

With three nearest neighbors, the IBk classifier correctly classified 622 instances, at an accuracy rate of 73.5%, overall error rate of 26.5%, and mean absolute error 0.1553. (Table 5, Figure 10, Figure 11)

K = 5

With five nearest neighbors, the IBk classifier correctly classified 622 instances, at an accuracy rate of 73.5%, overall error rate of 26.5%, and mean absolute error 0.1603. (Table 5, Figure 10, Figure 12)

K = 7

With seven nearest neighbors, the IBk classifier correctly classified 611 instances, at an accuracy rate of 72.2%, overall error rate of 27.8%, and mean absolute error 0.1681. (Table 5, Figure 10, Figure 13)

Between the three nearest neighbor models, three and five nearest neighbors produces similar results. However, the mean absolute error with three nearest neighbors is slightly lower than five nearest neighbors. Given this, three nearest neighbors appears to be the best k-value for this data set. One would assume that the more neighbors calculated then the more accurate the predictions would be, but with too many neighbors data spread may cover too great of a distance and accuracy is reduced. On the other hand, data can be too noisy to produce accurate results is k is too small. (Nico, 2018) Thus it is important to find the best value for k. In this case, it appears this slightly better, but all are very similar. (Figure 10) Although seven nearest neighbors does improve the accuracy for “Van” and “Bus.” (Figure 16)

IBk									
	k=3	**k=5**	**k=7**	**k=3**	**k=5**	**k=7**	**k=3**	**k=5**	**k=7**
	Instances			**Rate (%)**			**Mean Absolute Error**		
Correct	622	622	611	73.5%	73.5%	72.2%	0.1553	0.1603	0.1681
Incorrect	224	224	235	26.5%	26.5%	27.8%			

TABLE 5 SHOWING SUMMARY DATA FOR IBK, K= 3,5, & 7

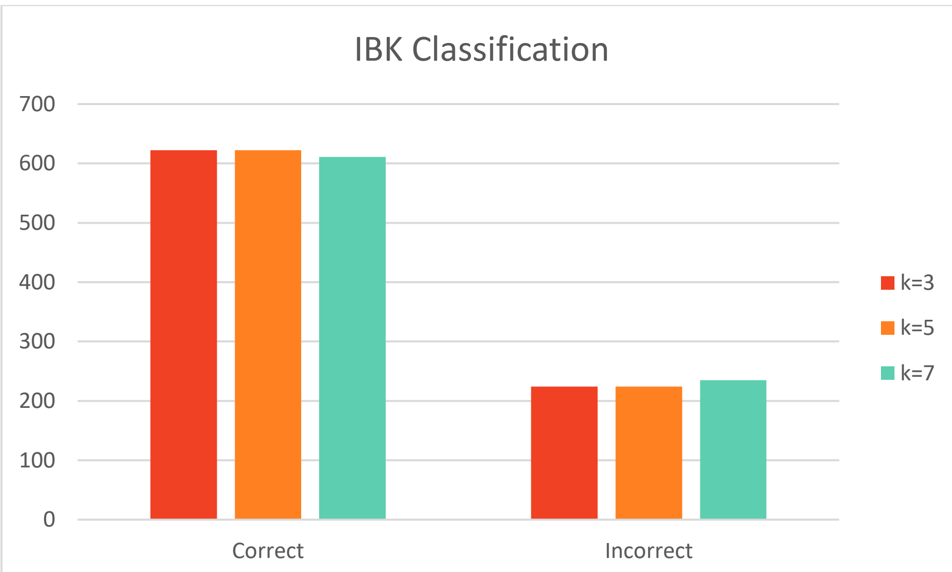

FIGURE 10 SHOWING CORRECTLY AND INCORRECTLY CLASSIFIED INSTANCE FOR IBK, K=3,5 &7

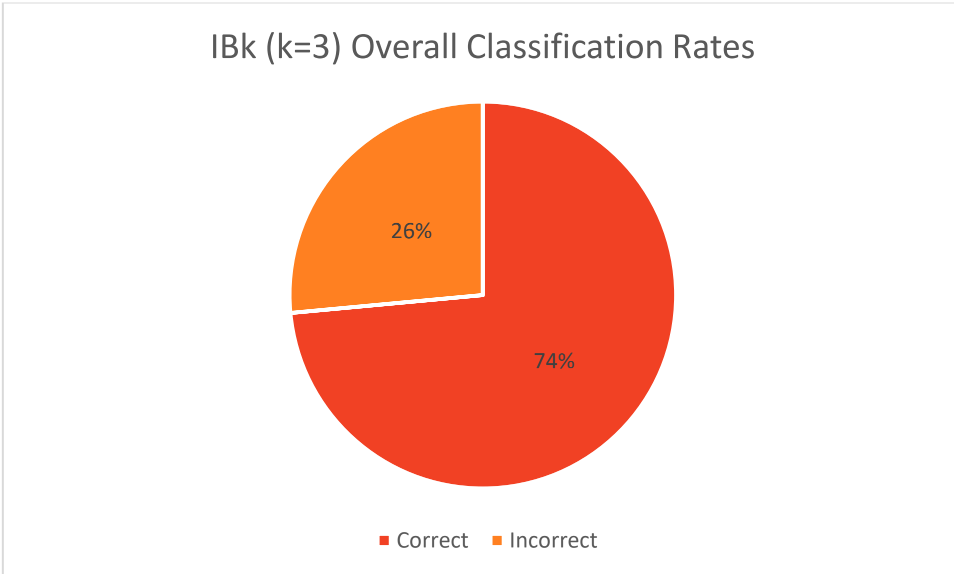

FIGURE 11 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF IBK, K=3

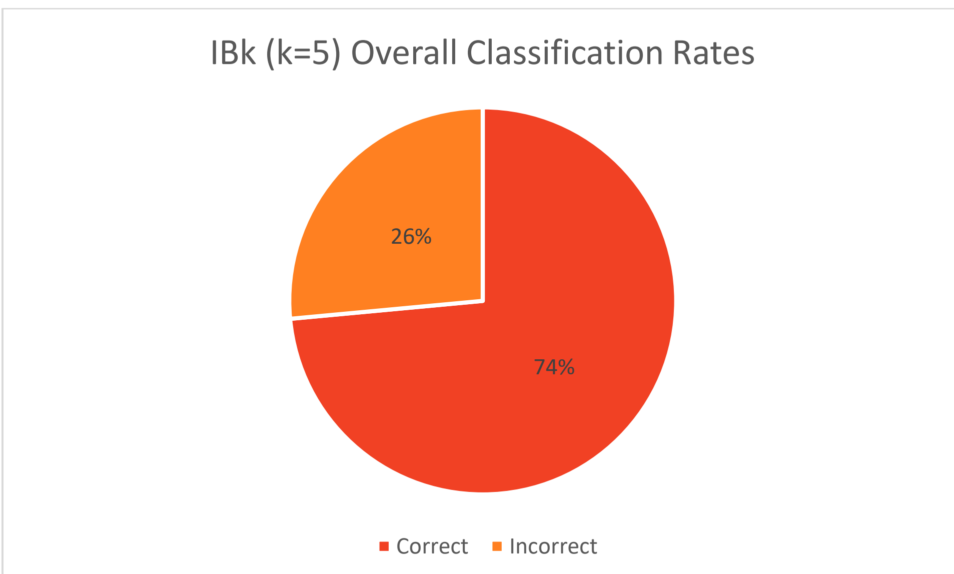

FIGURE 12 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF IBK, K=5

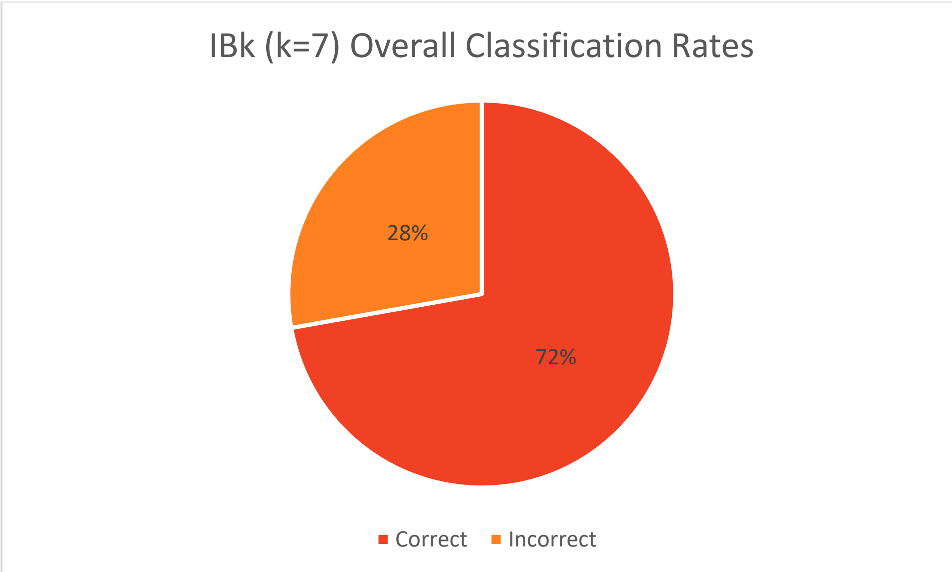

FIGURE 13 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF IBK, K=7

k=3						
CONFUSION MATRIX		PREDICTED				ACTUAL TOTAL
		OPEL	SAAB	BUS	VAN	
ACTUAL	OPEL	99	90	8	15	212
	SAAB	60	123	13	21	217
	BUS	3	4	211	0	218
	VAN	2	5	3	189	199
PREDICTED TOTAL		164	222	235	225	846

TABLE 6 SHOWING IBK, K=3 CONFUSION MATRIX

k=5						
CONFUSION MATRIX		PREDICTED				ACTUAL TOTAL
		OPEL	SAAB	BUS	VAN	
ACTUAL	OPEL	93	90	12	17	212
	SAAB	53	128	15	21	217
	BUS	3	4	211	0	218
	VAN	3	4	2	190	199
PREDICTED TOTAL		152	226	240	228	846

TABLE 7 SHOWING IBK, K=5 CONFUSION MATRIX

k=7						
CONFUSION MATRIX		PREDICTED				ACTUAL TOTAL
		OPEL	SAAB	BUS	VAN	
ACTUAL	OPEL	86	90	15	21	212
	SAAB	53	124	16	24	217
	BUS	4	2	210	2	218
	VAN	2	4	2	191	199
PREDICTED TOTAL		145	220	243	238	846

TABLE 8 SHOWING IBK, K=7 CONFUSION MATRIX

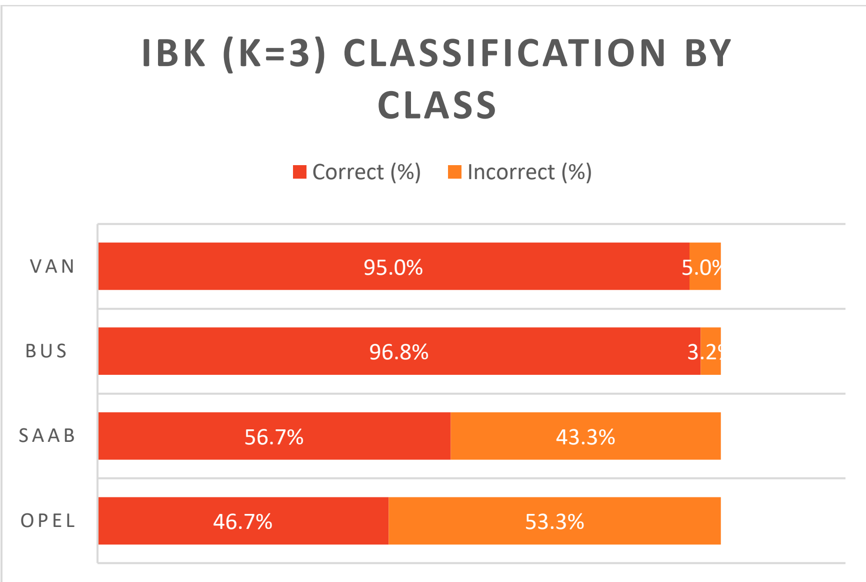

FIGURE 14 SHOWING IBK, K=3 CLASSIFICATION RATES BY CLASS

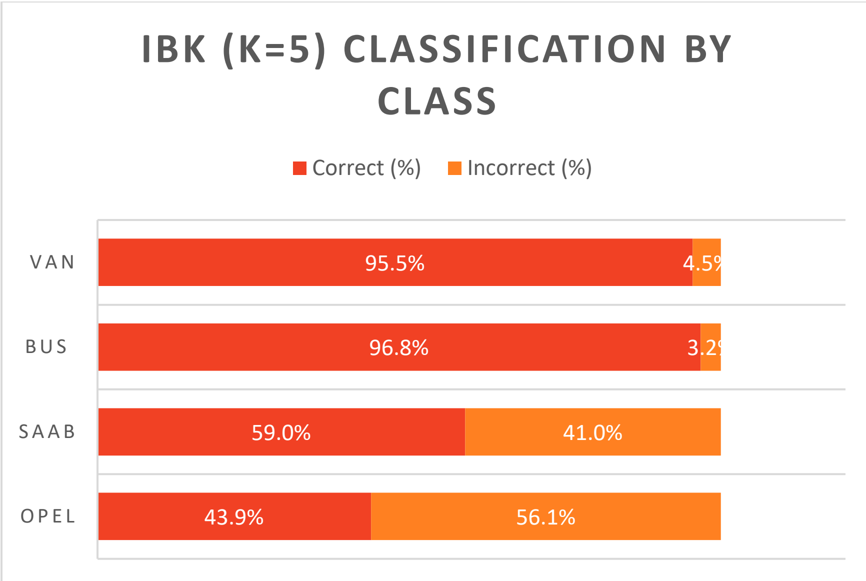

FIGURE 15 SHOWING IBK, K=5 CLASSIFICATION RATES BY CLASS

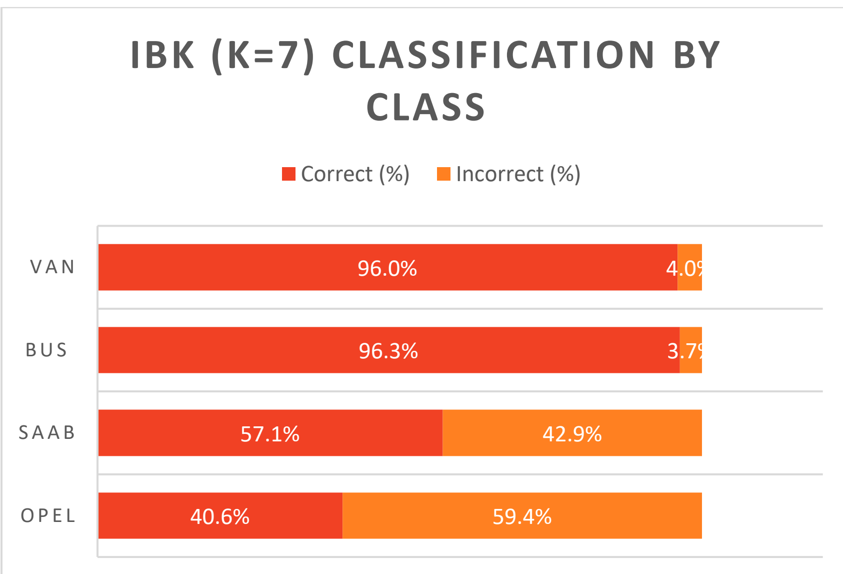

FIGURE 16 SHOWING IBK, K=7 CLASSIFICATION RATES BY CLASS

OneR

OneR stands for “One Rule.” For each attribute in the dataset, the value of each target class is counted, and the most frequent class is chosen as the predictor. Total errors of each predictor in the dataset are then calculated, and the predictor with the smallest total error is ultimately chosen as the “one rule.” OneR can only process nominal data, so as was done in this case, numeric data must be discretized before processing.

The OneR classifier correctly classified 466 instances, at an accuracy rate of 55.1%, overall error rate of 44.9%, and mean absolute error 0.2246. (Table 9, Figure 17, Figure 18) It can be observed from the confusion matrix that "Van" had the highest accuracy at 80.9%. (Table 10, Figure 19)

OneR			
	Instances	**Rate (%)**	**Mean Absolute Error**
Correct	466	55.1%	0.2246
Incorrect	380	44.9%	

TABLE 9 SHOWING SUMMARY DATA FOR ONER

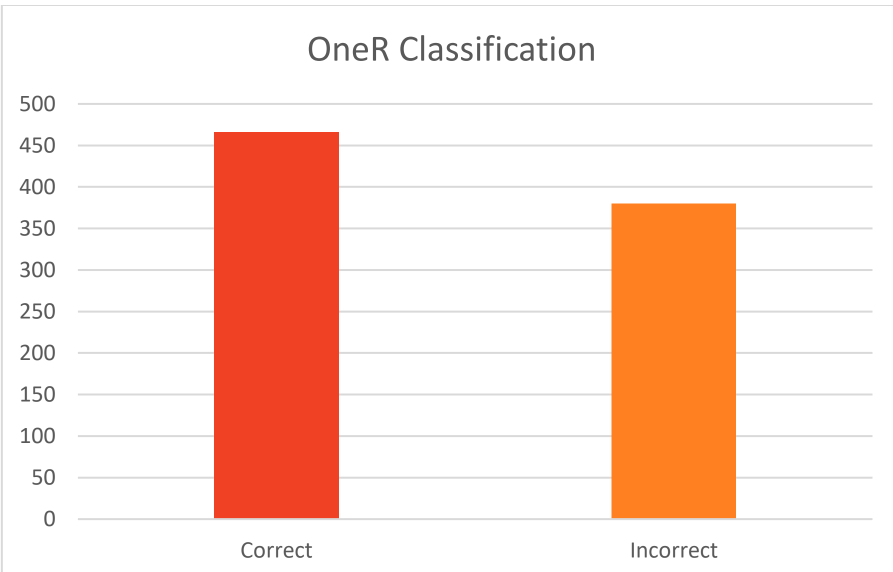

FIGURE 17 SHOWING CORRECTLY AND INCORRECTLY CLASSIFIED INSTANCE FOR ONER

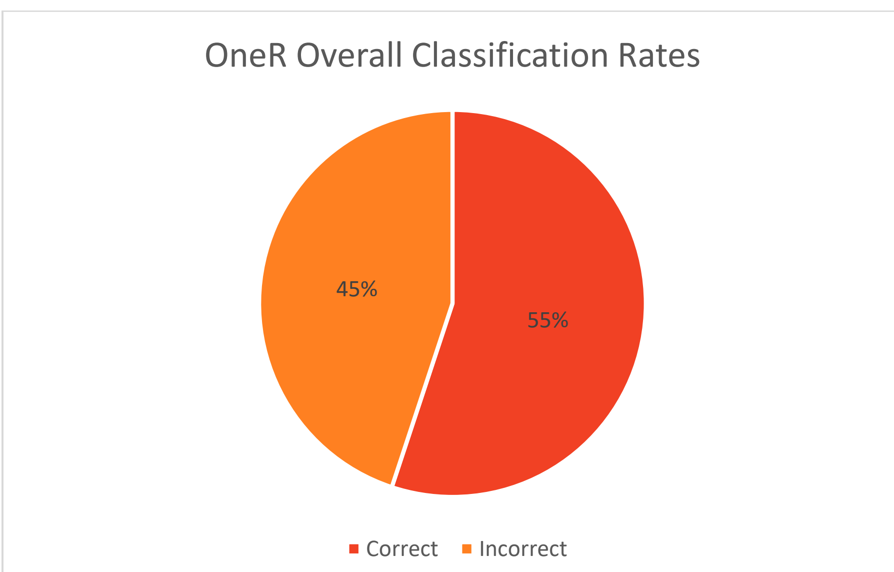

FIGURE 18 SHOWING RATE OF CORRECT (ACCURACY) AND RATE OF INCORRECT (OVERALL ERROR) OF ONER

OneR						
CONFUSION		PREDICTED				ACTUAL TOTAL
		OPEL	SAAB	BUS	VAN	
ACTUAL	OPEL	22	121	22	47	212
	SAAB	l0	142	20	55	217
	BUS	2	52	141	23	218
	VAN	0	0	38	161	199
PREDICTED TOTAL		24	315	221	286	846

TABLE 10 SHOWING CONFUSION MATRIX FOR ONER

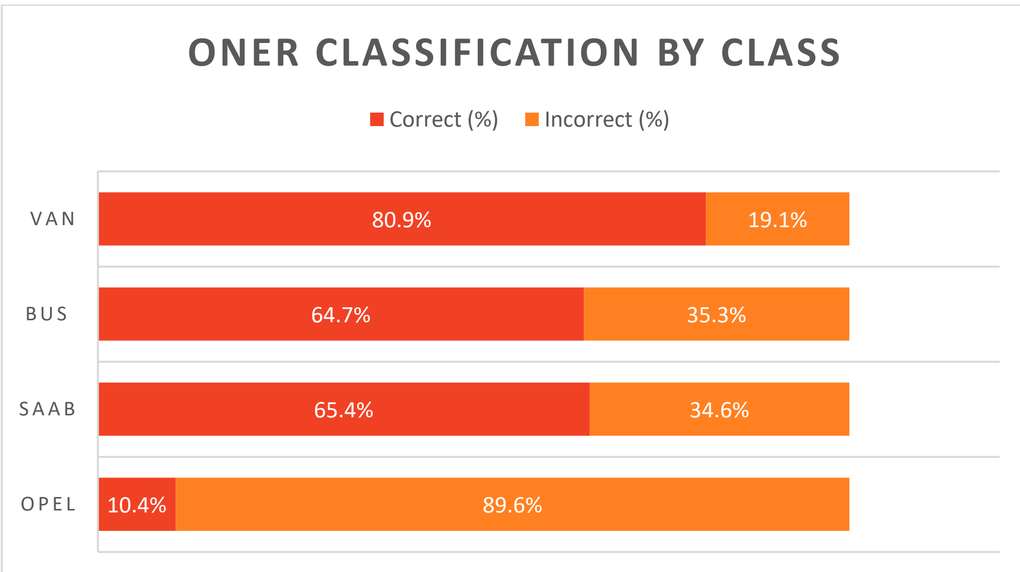

FIGURE 19 SHOWING ONER CLASSIFICATION RATES BY CLASS

Conclusion

IBk and J48 classifiers performed the best for the specified criteria of accuracy, overall error and mean absolute error rates of the four different models. IBk (k=3) slightly edges out J48 with its accuracy (Figure 20), but their mean absolute error is the same. (Figure 21) Of the four classes in the dataset, “Van” was classified with the most accuracy and “Opel” with the least accuracy over all models. Interestingly, J48 and IBk both classified “Bus” more accurately than “Van” while Naïve Bayes and OneR classified them moderately well. (Figure 22) “Bus” and “Van” being classified more accurately on those models is likely a contributing factor to them being more accurate overall than the other two models. Another surprise is the very low accuracy rate for OneR classification of “Opel” while other accuracy rates seem to be on par with other models. With the criteria of accuracy, overall error rate and mean absolute error rate considered for each classifier, it appears that IBk (k=3) is the best model of the four for this dataset.

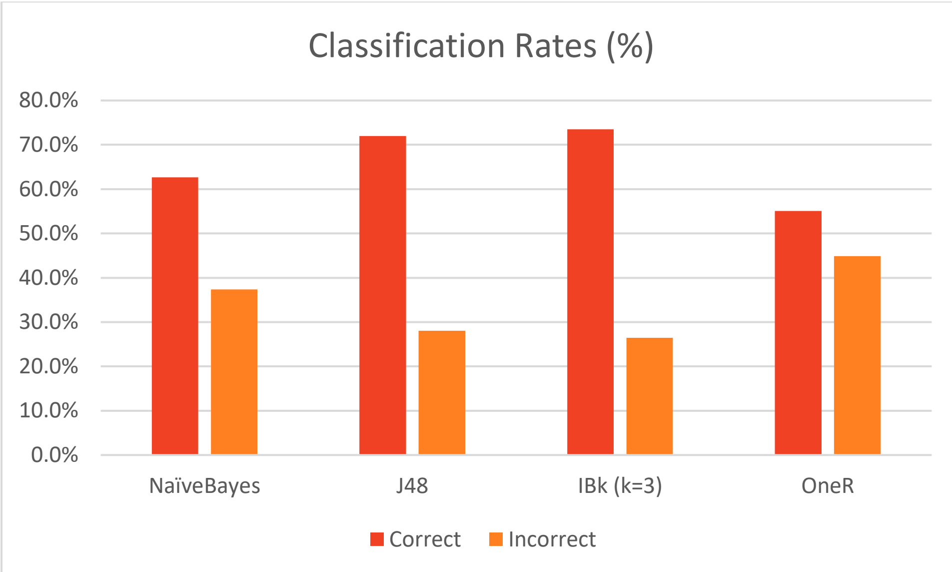

Figure 20 Showing overall classification rates for each model

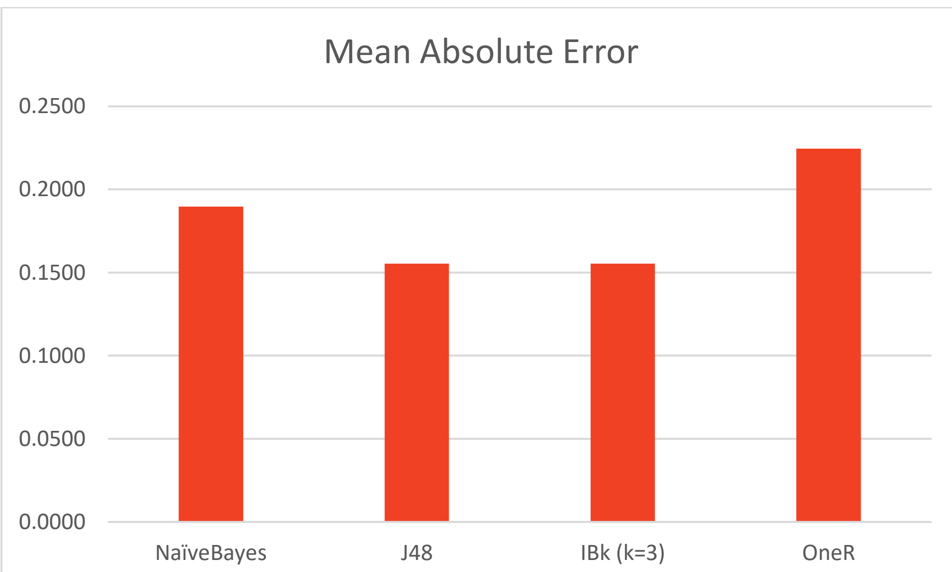

Figure 21 Showing mean absolute error for each model

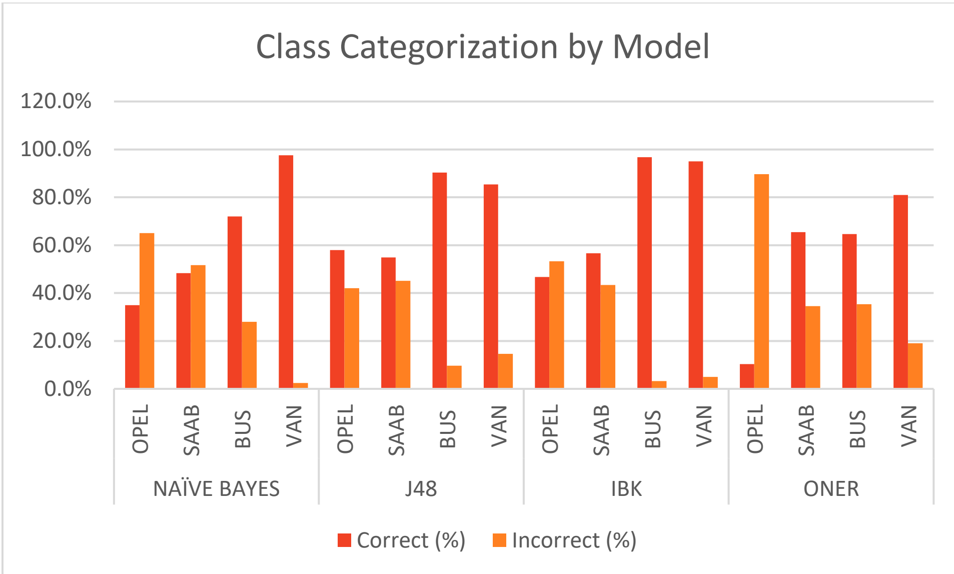

FIGURE 22 SHOWING SIDE-BY-SIDE CLASSIFICATION RESULTS FOR EACH MODEL

APPENDIX

WEKA OUTPUT

Shared Run Information:

```
Relation:     vehicle-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6
Instances:    846
Attributes:   19
              COMPACTNESS
              CIRCULARITY
              DISTANCE CIRCULARITY
              RADIUS RATIO
              PR.AXIS ASPECT RATIO
              MAX.LENGTH ASPECT RATIO
              SCATTER RATIO
              ELONGATEDNESS
              PR.AXIS RECTANGULARITY
              MAX.LENGTH RECTANGULARITY
              SCALED VARIANCE_MAJOR
              SCALED VARIANCE_MINOR
              SCALED RADIUS OF GYRATION
              SKEWNESS ABOUT_MAJOR
              SKEWNESS ABOUT_MINOR
              KURTOSIS ABOUT_MAJOR
              KURTOSIS ABOUT_MINOR
              HOLLOWS RATIO
              Class
Test mode:    10-fold cross-validation
```

Naïve Bayes:

```
=== Classifier model (full training set) ===

Naive Bayes Classifier

                            Class
Attribute                    opel   saab    bus    van
                           (0.25) (0.26) (0.26) (0.24)
========================================================
COMPACTNESS
  '(-inf-81.5]'              15.0    4.0   22.0    1.0
  '(81.5-87.5]'              29.0   31.0   77.0   44.0
  '(87.5-98.5]'              85.0   89.0   72.0  155.0
  '(98.5-103.5]'             52.0   25.0   25.0    3.0
  '(103.5-inf)'              36.0   73.0   27.0    1.0
  [total]                   217.0  222.0  223.0  204.0

CIRCULARITY
  '(-inf-40.5]'              65.0   75.0   29.0   76.0
  '(40.5-49.5]'              62.0   60.0  152.0  124.0
  '(49.5-54.5]'              47.0   75.0   23.0    2.0
  '(54.5-inf)'               42.0   11.0   18.0    1.0
  [total]                   216.0  221.0  222.0  203.0

DISTANCE CIRCULARITY
  '(-inf-64.5]'              15.0   25.0   12.0   48.0
  '(64.5-76.5]'              37.0   31.0  141.0   62.0
  '(76.5-92.5]'              52.0   55.0   36.0   92.0
  '(92.5-inf)'              112.0  110.0   33.0    1.0
  [total]                   216.0  221.0  222.0  203.0

RADIUS RATIO
  '(-inf-175.5]'             85.0   81.0  130.0  192.0
  '(175.5-234.5]'           129.0  138.0   86.0    4.0
  '(234.5-inf)'               1.0    1.0    5.0    6.0
  [total]                   215.0  220.0  221.0  202.0

PR.AXIS ASPECT RATIO
```

'(-inf-52.5]'	4.0	2.0	18.0	19.0
'(52.5-68.5]'	192.0	207.0	141.0	172.0
'(68.5-86.5]'	19.0	11.0	60.0	5.0
'(86.5-inf)'	1.0	1.0	3.0	7.0
[total]	216.0	221.0	222.0	203.0
MAX.LENGTH ASPECT RATIO				
'(-inf-7.5]'	59.0	66.0	192.0	61.0
'(7.5-8.5]'	33.0	27.0	23.0	34.0
'(8.5-16]'	123.0	127.0	1.0	100.0
'(16-inf)'	1.0	1.0	6.0	8.0
[total]	216.0	221.0	222.0	203.0
SCATTER RATIO				
'(-inf-140.5]'	33.0	31.0	1.0	100.0
'(140.5-154.5]'	28.0	32.0	128.0	42.0
'(154.5-163.5]'	16.0	18.0	18.0	60.0
'(163.5-230.5]'	139.0	140.0	51.0	1.0
'(230.5-inf)'	1.0	1.0	25.0	1.0
[total]	217.0	222.0	223.0	204.0
ELONGATEDNESS				
'(-inf-29.5]'	1.0	1.0	27.0	1.0
'(29.5-41.5]'	148.0	149.0	62.0	1.0
'(41.5-44.5]'	18.0	21.0	50.0	74.0
'(44.5-46.5]'	14.0	14.0	82.0	26.0
'(46.5-inf)'	36.0	37.0	2.0	102.0
[total]	217.0	222.0	223.0	204.0
PR.AXIS RECTANGULARITY				
'(-inf-18.5]'	34.0	32.0	6.0	104.0
'(18.5-19.5]'	30.0	33.0	133.0	47.0
'(19.5-20.5]'	23.0	27.0	19.0	51.0
'(20.5-25.5]'	129.0	129.0	40.0	1.0
'(25.5-inf)'	1.0	1.0	25.0	1.0
[total]	217.0	222.0	223.0	204.0
MAX.LENGTH RECTANGULARITY				
'(-inf-135.5]'	62.0	69.0	25.0	43.0
'(135.5-147.5]'	36.0	32.0	131.0	77.0
'(147.5-160.5]'	34.0	49.0	35.0	66.0
'(160.5-172.5]'	47.0	63.0	29.0	17.0
'(172.5-inf)'	38.0	9.0	3.0	1.0
[total]	217.0	222.0	223.0	204.0
SCALED VARIANCE_MAJOR				
'(-inf-165.5]'	36.0	41.0	5.0	105.0
'(165.5-180.5]'	25.0	27.0	126.0	90.0
'(180.5-242]'	154.0	152.0	62.0	5.0
'(242-inf)'	1.0	1.0	29.0	3.0
[total]	216.0	221.0	222.0	203.0
SCALED VARIANCE_MINOR				
'(-inf-298.5]'	35.0	35.0	1.0	100.0
'(298.5-347.5]'	23.0	21.0	116.0	39.0
'(347.5-389.5]'	14.0	22.0	24.0	63.0
'(389.5-581]'	52.0	57.0	41.0	1.0
'(581-721.5]'	71.0	87.0	13.0	1.0
'(721.5-761.5]'	23.0	1.0	3.0	1.0
'(761.5-inf)'	1.0	1.0	27.0	1.0
[total]	219.0	224.0	225.0	206.0
SCALED RADIUS OF GYRATION				
'(-inf-170.5]'	90.0	95.0	65.0	127.0
'(170.5-192.5]'	31.0	32.0	105.0	70.0
'(192.5-241.5]'	88.0	93.0	34.0	5.0
'(241.5-inf)'	7.0	1.0	18.0	1.0
[total]	216.0	221.0	222.0	203.0
SKEWNESS ABOUT_MAJOR				
'(-inf-64.5]'	31.0	36.0	1.0	30.0

```
  '(64.5-74.5]'              154.0  158.0   93.0  102.0
  '(74.5-inf)'                30.0   26.0  127.0   70.0
  [total]                    215.0  220.0  221.0  202.0

SKEWNESS ABOUT_MINOR
  '(-inf-11.5]'              172.0  159.0  214.0  167.0
  '(11.5-inf)'                42.0   60.0    6.0   34.0
  [total]                    214.0  219.0  220.0  201.0

KURTOSIS ABOUT_MAJOR
  '(-inf-17.5]'              123.0  127.0  188.0  179.0
  '(17.5-inf)'                91.0   92.0   32.0   22.0
  [total]                    214.0  219.0  220.0  201.0

KURTOSIS ABOUT_MINOR
  '(-inf-177.5]'               6.0    1.0    1.0    4.0
  '(177.5-181.5]'             13.0   18.0   70.0   13.0
  '(181.5-185.5]'             22.0   17.0   36.0   60.0
  '(185.5-191.5]'            112.0  117.0   34.0   59.0
  '(191.5-inf)'               64.0   69.0   82.0   68.0
  [total]                    217.0  222.0  223.0  204.0

HOLLOWS RATIO
  '(-inf-189.5]'              24.0   30.0  112.0   35.0
  '(189.5-inf)'              190.0  189.0  108.0  166.0
  [total]                    214.0  219.0  220.0  201.0


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         530               62.6478 %
Incorrectly Classified Instances       316               37.3522 %
Kappa statistic                          0.5033
Mean absolute error                      0.1897
Root mean squared error                  0.3797
Relative absolute error                 50.6045 %
Root relative squared error             87.6988 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.349    0.112    0.510      0.349    0.415      0.273    0.772     0.567
opel
                 0.484    0.154    0.520      0.484    0.501      0.338    0.769     0.542
saab
                 0.720    0.054    0.822      0.720    0.768      0.697    0.961     0.904
bus
                 0.975    0.176    0.630      0.975    0.765      0.704    0.951     0.832
van
Weighted Avg.    0.626    0.123    0.621      0.626    0.610      0.500    0.862     0.710

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
  74  74  13  51 |   a = opel
  42 105  16  54 |   b = saab
  29  23 157   9 |   c = bus
   0   0   5 194 |   d = van
```

J48:

```
=== Classifier model (full training set) ===

J48 pruned tree
------------------

SCALED VARIANCE_MAJOR = '(-inf-165.5]'
|   MAX.LENGTH RECTANGULARITY = '(-inf-135.5]'
|   |   PR.AXIS RECTANGULARITY = '(-inf-18.5]'
|   |   |   COMPACTNESS = '(-inf-81.5]': opel (11.0)
|   |   |   COMPACTNESS = '(81.5-87.5]'
|   |   |   |   KURTOSIS ABOUT_MINOR = '(-inf-177.5]': opel (1.0)
|   |   |   |   KURTOSIS ABOUT_MINOR = '(177.5-181.5]': saab (16.0/7.0)
|   |   |   |   KURTOSIS ABOUT_MINOR = '(181.5-185.5]'
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(-inf-64.5]': opel (0.0)
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(64.5-74.5]'
|   |   |   |   |   |   HOLLOWS RATIO = '(-inf-189.5]': saab (2.0)
|   |   |   |   |   |   HOLLOWS RATIO = '(189.5-inf)': opel (4.0)
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(74.5-inf)'
|   |   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(-inf-17.5]': van (3.0/1.0)
|   |   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(17.5-inf)': opel (2.0/1.0)
|   |   |   |   KURTOSIS ABOUT_MINOR = '(185.5-191.5]': saab (12.0/7.0)
|   |   |   |   KURTOSIS ABOUT_MINOR = '(191.5-inf)': van (9.0/1.0)
|   |   |   COMPACTNESS = '(87.5-98.5]'
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(-inf-7.5]'
|   |   |   |   |   DISTANCE CIRCULARITY = '(-inf-64.5]'
|   |   |   |   |   |   SKEWNESS ABOUT_MINOR = '(-inf-11.5]': van (10.0/2.0)
|   |   |   |   |   |   SKEWNESS ABOUT_MINOR = '(11.5-inf)': saab (6.0/2.0)
|   |   |   |   |   DISTANCE CIRCULARITY = '(64.5-76.5]': opel (12.0/6.0)
|   |   |   |   |   DISTANCE CIRCULARITY = '(76.5-92.5]': opel (1.0)
|   |   |   |   |   DISTANCE CIRCULARITY = '(92.5-inf)': van (0.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(7.5-8.5]': van (4.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(8.5-16]': van (2.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(16-inf)': van (0.0)
|   |   |   COMPACTNESS = '(98.5-103.5]': van (0.0)
|   |   |   COMPACTNESS = '(103.5-inf)': van (0.0)
|   |   PR.AXIS RECTANGULARITY = '(18.5-19.5]': saab (7.0/1.0)
|   |   PR.AXIS RECTANGULARITY = '(19.5-20.5]': van (0.0)
|   |   PR.AXIS RECTANGULARITY = '(20.5-25.5]': van (0.0)
|   |   PR.AXIS RECTANGULARITY = '(25.5-inf)': van (0.0)
|   MAX.LENGTH RECTANGULARITY = '(135.5-147.5]'
|   |   SCATTER RATIO = '(-inf-140.5]': van (58.0/6.0)
|   |   SCATTER RATIO = '(140.5-154.5]'
|   |   |   PR.AXIS RECTANGULARITY = '(-inf-18.5]': bus (5.0/1.0)
|   |   |   PR.AXIS RECTANGULARITY = '(18.5-19.5]'
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(-inf-7.5]': saab (3.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(7.5-8.5]': opel (2.0/1.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(8.5-16]': van (6.0)
|   |   |   |   MAX.LENGTH ASPECT RATIO = '(16-inf)': van (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(19.5-20.5]': van (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(20.5-25.5]': van (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(25.5-inf)': van (0.0)
|   |   SCATTER RATIO = '(154.5-163.5]': van (0.0)
|   |   SCATTER RATIO = '(163.5-230.5]': van (0.0)
|   |   SCATTER RATIO = '(230.5-inf)': van (0.0)
|   MAX.LENGTH RECTANGULARITY = '(147.5-160.5]': van (7.0)
|   MAX.LENGTH RECTANGULARITY = '(160.5-172.5]': van (0.0)
|   MAX.LENGTH RECTANGULARITY = '(172.5-inf)': van (0.0)
SCALED VARIANCE_MAJOR = '(165.5-180.5]'
|   MAX.LENGTH ASPECT RATIO = '(-inf-7.5]'
|   |   SCATTER RATIO = '(-inf-140.5]': van (5.0/1.0)
|   |   SCATTER RATIO = '(140.5-154.5]'
|   |   |   SKEWNESS ABOUT_MINOR = '(-inf-11.5]'
|   |   |   |   DISTANCE CIRCULARITY = '(-inf-64.5]': bus (9.0/1.0)
|   |   |   |   DISTANCE CIRCULARITY = '(64.5-76.5]'
|   |   |   |   |   CIRCULARITY = '(-inf-40.5]'
|   |   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(-inf-64.5]': opel (2.0)
|   |   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(64.5-74.5]': bus (13.0/1.0)
|   |   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(74.5-inf)': saab (3.0)
|   |   |   |   |   CIRCULARITY = '(40.5-49.5]': bus (74.0)
```

```
|   |   |   |   |   CIRCULARITY = '(49.5-54.5]': bus (0.0)
|   |   |   |   |   CIRCULARITY = '(54.5-inf)': bus (0.0)
|   |   |   |   DISTANCE CIRCULARITY = '(76.5-92.5]': opel (11.0/4.0)
|   |   |   |   DISTANCE CIRCULARITY = '(92.5-inf)': bus (0.0)
|   |   |   SKEWNESS ABOUT_MINOR = '(11.5-inf)'
|   |   |   |   SCALED RADIUS OF GYRATION = '(-inf-170.5]': opel (9.0/2.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(170.5-192.5]': saab (5.0/3.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(192.5-241.5]': opel (0.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(241.5-inf)': opel (0.0)
|   |   SCATTER RATIO = '(154.5-163.5]'
|   |   |   PR.AXIS ASPECT RATIO = '(-inf-52.5]': bus (0.0)
|   |   |   PR.AXIS ASPECT RATIO = '(52.5-68.5]': saab (3.0/1.0)
|   |   |   PR.AXIS ASPECT RATIO = '(68.5-86.5]': bus (7.0)
|   |   |   PR.AXIS ASPECT RATIO = '(86.5-inf)': bus (0.0)
|   |   SCATTER RATIO = '(163.5-230.5]': bus (0.0)
|   |   SCATTER RATIO = '(230.5-inf)': bus (0.0)
|   MAX.LENGTH ASPECT RATIO = '(7.5-8.5]'
|   |   HOLLOWS RATIO = '(-inf-189.5]': bus (18.0/1.0)
|   |   HOLLOWS RATIO = '(189.5-inf)'
|   |   |   PR.AXIS RECTANGULARITY = '(-inf-18.5]': van (5.0)
|   |   |   PR.AXIS RECTANGULARITY = '(18.5-19.5]'
|   |   |   |   PR.AXIS ASPECT RATIO = '(-inf-52.5]': saab (0.0)
|   |   |   |   PR.AXIS ASPECT RATIO = '(52.5-68.5]'
|   |   |   |   |   SCALED RADIUS OF GYRATION = '(-inf-170.5]'
|   |   |   |   |   |   ELONGATEDNESS = '(-inf-29.5]': saab (0.0)
|   |   |   |   |   |   ELONGATEDNESS = '(29.5-41.5]': saab (0.0)
|   |   |   |   |   |   ELONGATEDNESS = '(41.5-44.5]': saab (10.0/3.0)
|   |   |   |   |   |   ELONGATEDNESS = '(44.5-46.5]': opel (4.0/1.0)
|   |   |   |   |   |   ELONGATEDNESS = '(46.5-inf)': saab (0.0)
|   |   |   |   |   SCALED RADIUS OF GYRATION = '(170.5-192.5]': bus (2.0/1.0)
|   |   |   |   |   SCALED RADIUS OF GYRATION = '(192.5-241.5]': saab (0.0)
|   |   |   |   |   SCALED RADIUS OF GYRATION = '(241.5-inf)': saab (0.0)
|   |   |   |   PR.AXIS ASPECT RATIO = '(68.5-86.5]': bus (2.0)
|   |   |   |   PR.AXIS ASPECT RATIO = '(86.5-inf)': saab (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(19.5-20.5]': van (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(20.5-25.5]': van (0.0)
|   |   |   PR.AXIS RECTANGULARITY = '(25.5-inf)': van (0.0)
|   MAX.LENGTH ASPECT RATIO = '(8.5-16]'
|   |   SKEWNESS ABOUT_MAJOR = '(-inf-64.5]': saab (3.0)
|   |   SKEWNESS ABOUT_MAJOR = '(64.5-74.5]': van (47.0)
|   |   SKEWNESS ABOUT_MAJOR = '(74.5-inf)': van (30.0)
|   MAX.LENGTH ASPECT RATIO = '(16-inf)': bus (2.0)
SCALED VARIANCE_MAJOR = '(180.5-242]'
|   MAX.LENGTH ASPECT RATIO = '(-inf-7.5]'
|   |   SKEWNESS ABOUT_MINOR = '(-inf-11.5]'
|   |   |   PR.AXIS ASPECT RATIO = '(-inf-52.5]': bus (1.0)
|   |   |   PR.AXIS ASPECT RATIO = '(52.5-68.5]'
|   |   |   |   SKEWNESS ABOUT_MAJOR = '(-inf-64.5]': bus (0.0)
|   |   |   |   SKEWNESS ABOUT_MAJOR = '(64.5-74.5]'
|   |   |   |   |   KURTOSIS ABOUT_MINOR = '(-inf-177.5]': saab (0.0)
|   |   |   |   |   KURTOSIS ABOUT_MINOR = '(177.5-181.5]': saab (0.0)
|   |   |   |   |   KURTOSIS ABOUT_MINOR = '(181.5-185.5]': opel (1.0)
|   |   |   |   |   KURTOSIS ABOUT_MINOR = '(185.5-191.5]'
|   |   |   |   |   |   RADIUS RATIO = '(-inf-175.5]': saab (6.0)
|   |   |   |   |   |   RADIUS RATIO = '(175.5-234.5]'
|   |   |   |   |   |   |   CIRCULARITY = '(-inf-40.5]': opel (2.0)
|   |   |   |   |   |   |   CIRCULARITY = '(40.5-49.5]': saab (2.0)
|   |   |   |   |   |   |   CIRCULARITY = '(49.5-54.5]': opel (1.0)
|   |   |   |   |   |   |   CIRCULARITY = '(54.5-inf)': opel (0.0)
|   |   |   |   |   |   RADIUS RATIO = '(234.5-inf)': saab (0.0)
|   |   |   |   |   KURTOSIS ABOUT_MINOR = '(191.5-inf)'
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(-inf-18.5]': bus (0.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(18.5-19.5]': bus (1.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(19.5-20.5]'
|   |   |   |   |   |   |   DISTANCE CIRCULARITY = '(-inf-64.5]': opel (0.0)
|   |   |   |   |   |   |   DISTANCE CIRCULARITY = '(64.5-76.5]': bus (4.0/1.0)
|   |   |   |   |   |   |   DISTANCE CIRCULARITY = '(76.5-92.5]'
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(-inf-298.5]': saab (0.0)
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(298.5-347.5]': saab (0.0)
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(347.5-389.5]': saab (2.0)
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(389.5-581]': opel (3.0/1.0)
```

```
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(581-721.5]': saab (0.0)
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(721.5-761.5]': saab (0.0)
|   |   |   |   |   |   |   |   SCALED VARIANCE_MINOR = '(761.5-inf)': saab (0.0)
|   |   |   |   |   |   |   DISTANCE CIRCULARITY = '(92.5-inf)': opel (0.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(20.5-25.5]': bus (11.0/2.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(25.5-inf)': bus (0.0)
|   |   |   |   SKEWNESS ABOUT_MAJOR = '(74.5-inf)': bus (13.0)
|   |   |   PR.AXIS ASPECT RATIO = '(68.5-86.5]': bus (29.0)
|   |   |   PR.AXIS ASPECT RATIO = '(86.5-inf)': bus (0.0)
|   |   SKEWNESS ABOUT_MINOR = '(11.5-inf)'
|   |   |   PR.AXIS ASPECT RATIO = '(-inf-52.5]': opel (0.0)
|   |   |   PR.AXIS ASPECT RATIO = '(52.5-68.5]': opel (5.0)
|   |   |   PR.AXIS ASPECT RATIO = '(68.5-86.5]': bus (2.0)
|   |   |   PR.AXIS ASPECT RATIO = '(86.5-inf)': opel (0.0)
|   MAX.LENGTH ASPECT RATIO = '(7.5-8.5]': opel (36.0/15.0)
|   MAX.LENGTH ASPECT RATIO = '(8.5-16]'
|   |   SCALED VARIANCE_MINOR = '(-inf-298.5]': saab (0.0)
|   |   SCALED VARIANCE_MINOR = '(298.5-347.5]': saab (0.0)
|   |   SCALED VARIANCE_MINOR = '(347.5-389.5]': opel (1.0)
|   |   SCALED VARIANCE_MINOR = '(389.5-581]'
|   |   |   SCATTER RATIO = '(-inf-140.5]': saab (0.0)
|   |   |   SCATTER RATIO = '(140.5-154.5]': saab (0.0)
|   |   |   SCATTER RATIO = '(154.5-163.5]': saab (2.0)
|   |   |   SCATTER RATIO = '(163.5-230.5]'
|   |   |   |   SCALED RADIUS OF GYRATION = '(-inf-170.5]'
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(-inf-64.5]': saab (12.0/3.0)
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(64.5-74.5]'
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(-inf-18.5]': opel (0.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(18.5-19.5]': opel (0.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(19.5-20.5]': opel (2.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(20.5-25.5]'
|   |   |   |   |   |   |   CIRCULARITY = '(-inf-40.5]': saab (3.0)
|   |   |   |   |   |   |   CIRCULARITY = '(40.5-49.5]': opel (12.0/4.0)
|   |   |   |   |   |   |   CIRCULARITY = '(49.5-54.5]': opel (0.0)
|   |   |   |   |   |   |   CIRCULARITY = '(54.5-inf)': opel (0.0)
|   |   |   |   |   |   PR.AXIS RECTANGULARITY = '(25.5-inf)': opel (0.0)
|   |   |   |   |   SKEWNESS ABOUT_MAJOR = '(74.5-inf)': saab (0.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(170.5-192.5]'
|   |   |   |   |   SKEWNESS ABOUT_MINOR = '(-inf-11.5]'
|   |   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(-inf-17.5]': opel (3.0)
|   |   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(17.5-inf)': saab (3.0/1.0)
|   |   |   |   |   SKEWNESS ABOUT_MINOR = '(11.5-inf)': saab (6.0/2.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(192.5-241.5]'
|   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(-inf-17.5]'
|   |   |   |   |   |   CIRCULARITY = '(-inf-40.5]': saab (0.0)
|   |   |   |   |   |   CIRCULARITY = '(40.5-49.5]': opel (4.0/1.0)
|   |   |   |   |   |   CIRCULARITY = '(49.5-54.5]'
|   |   |   |   |   |   |   MAX.LENGTH RECTANGULARITY = '(-inf-135.5]': saab (0.0)
|   |   |   |   |   |   |   MAX.LENGTH RECTANGULARITY = '(135.5-147.5]': saab (0.0)
|   |   |   |   |   |   |   MAX.LENGTH RECTANGULARITY = '(147.5-160.5]': saab (7.0/1.0)
|   |   |   |   |   |   |   MAX.LENGTH RECTANGULARITY = '(160.5-172.5]'
|   |   |   |   |   |   |   |   KURTOSIS ABOUT_MINOR = '(-inf-177.5]': opel (0.0)
|   |   |   |   |   |   |   |   KURTOSIS ABOUT_MINOR = '(177.5-181.5]': opel (0.0)
|   |   |   |   |   |   |   |   KURTOSIS ABOUT_MINOR = '(181.5-185.5]': opel (0.0)
|   |   |   |   |   |   |   |   KURTOSIS ABOUT_MINOR = '(185.5-191.5]': opel (3.0)
|   |   |   |   |   |   |   |   KURTOSIS ABOUT_MINOR = '(191.5-inf)'
|   |   |   |   |   |   |   |   |   COMPACTNESS = '(-inf-81.5]': saab (0.0)
|   |   |   |   |   |   |   |   |   COMPACTNESS = '(81.5-87.5]': saab (0.0)
|   |   |   |   |   |   |   |   |   COMPACTNESS = '(87.5-98.5]': saab (0.0)
|   |   |   |   |   |   |   |   |   COMPACTNESS = '(98.5-103.5]': saab (2.0)
|   |   |   |   |   |   |   |   |   COMPACTNESS = '(103.5-inf)': opel (3.0/1.0)
|   |   |   |   |   |   |   MAX.LENGTH RECTANGULARITY = '(172.5-inf)': saab (0.0)
|   |   |   |   |   |   CIRCULARITY = '(54.5-inf)': saab (0.0)
|   |   |   |   |   KURTOSIS ABOUT_MAJOR = '(17.5-inf)': saab (3.0)
|   |   |   |   SCALED RADIUS OF GYRATION = '(241.5-inf)': opel (1.0)
|   |   |   SCATTER RATIO = '(230.5-inf)': saab (0.0)
|   |   SCALED VARIANCE_MINOR = '(581-721.5]'
|   |   |   COMPACTNESS = '(-inf-81.5]': saab (0.0)
|   |   |   COMPACTNESS = '(81.5-87.5]': saab (0.0)
|   |   |   COMPACTNESS = '(87.5-98.5]': opel (20.0/5.0)
|   |   |   COMPACTNESS = '(98.5-103.5]': opel (41.0/14.0)
```

```
|   |   |   COMPACTNESS = '(103.5-inf)'
|   |   |   |   CIRCULARITY = '(-inf-40.5]': saab (0.0)
|   |   |   |   CIRCULARITY = '(40.5-49.5]': saab (12.0/3.0)
|   |   |   |   CIRCULARITY = '(49.5-54.5]': saab (65.0/12.0)
|   |   |   |   CIRCULARITY = '(54.5-inf)': opel (16.0/5.0)
|   |   SCALED VARIANCE_MINOR = '(721.5-761.5]': opel (22.0)
|   |   SCALED VARIANCE_MINOR = '(761.5-inf)': saab (0.0)
|   MAX.LENGTH ASPECT RATIO = '(16-inf)'
|   |   HOLLOWS RATIO = '(-inf-189.5]': bus (3.0)
|   |   HOLLOWS RATIO = '(189.5-inf)': van (4.0)
SCALED VARIANCE_MAJOR = '(242-inf)'
|   HOLLOWS RATIO = '(-inf-189.5]': bus (28.0)
|   HOLLOWS RATIO = '(189.5-inf)': van (2.0)

Number of Leaves  :     164

Size of the tree :      220


Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         609               71.9858 %
Incorrectly Classified Instances       237               28.0142 %
Kappa statistic                          0.6264
Mean absolute error                      0.1553
Root mean squared error                  0.3191
Relative absolute error                 41.4218 %
Root relative squared error             73.7129 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.580    0.139    0.583      0.580    0.582      0.442    0.819     0.602
opel
                 0.548    0.145    0.567      0.548    0.557      0.408    0.814     0.564
saab
                 0.904    0.040    0.887      0.904    0.895      0.859    0.953     0.890
bus
                 0.854    0.051    0.837      0.854    0.846      0.798    0.956     0.850
van
Weighted Avg.    0.720    0.094    0.717      0.720    0.718      0.624    0.884     0.725

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
 123  70   9  10 |   a = opel
  72 119   9  17 |   b = saab
   9   6 197   6 |   c = bus
   7  15   7 170 |   d = van
```

IBk:

K = 3

```
=== Classifier model (full training set) ===

IB1 instance-based classifier
using 3 nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         622               73.5225 %
Incorrectly Classified Instances       224               26.4775 %
Kappa statistic                          0.647
Mean absolute error                      0.1553
Root mean squared error                  0.3019
Relative absolute error                 41.4415 %
Root relative squared error             69.7255 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.467    0.103    0.604      0.467    0.527      0.400    0.831     0.607
opel
                 0.567    0.157    0.554      0.567    0.560      0.406    0.822     0.588
saab
                 0.968    0.038    0.898      0.968    0.932      0.908    0.992     0.966
bus
                 0.950    0.056    0.840      0.950    0.892      0.858    0.987     0.956
van
Weighted Avg.    0.735    0.089    0.722      0.735    0.725      0.640    0.907     0.777

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
  99  90   8  15 |   a = opel
  60 123  13  21 |   b = saab
   3   4 211   0 |   c = bus
   2   5   3 189 |   d = van
```

K = 5

```
=== Classifier model (full training set) ===

IB1 instance-based classifier
using 5 nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         622               73.5225 %
Incorrectly Classified Instances       224               26.4775 %
Kappa statistic                          0.6469
Mean absolute error                      0.1603
Root mean squared error                  0.2942
Relative absolute error                 42.766  %
Root relative squared error             67.9659 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===
```

```
                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.439    0.093    0.612      0.439    0.511      0.390    0.846     0.644
opel
                 0.590    0.156    0.566      0.590    0.578      0.428    0.838     0.620
saab
                 0.968    0.046    0.879      0.968    0.921      0.894    0.991     0.969
bus
                 0.955    0.059    0.833      0.955    0.890      0.857    0.988     0.960
van
Weighted Avg.    0.735    0.089    0.721      0.735    0.723      0.640    0.915     0.796

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
  93  90  12  17 |   a = opel
  53 128  15  21 |   b = saab
   3   4 211   0 |   c = bus
   3   4   2 190 |   d = van
```

K = 7

```
=== Classifier model (full training set) ===

IB1 instance-based classifier
using 7 nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         611               72.2222 %
Incorrectly Classified Instances       235               27.7778 %
Kappa statistic                          0.6297
Mean absolute error                      0.1681
Root mean squared error                  0.297
Relative absolute error                 44.8453 %
Root relative squared error             68.5924 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.406    0.093    0.593      0.406    0.482      0.359    0.844     0.639
opel
                 0.571    0.153    0.564      0.571    0.568      0.417    0.836     0.624
saab
                 0.963    0.053    0.864      0.963    0.911      0.880    0.989     0.965
bus
                 0.960    0.073    0.803      0.960    0.874      0.837    0.986     0.957
van
Weighted Avg.    0.722    0.093    0.705      0.722    0.707      0.621    0.913     0.794

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
  86  90  15  21 |   a = opel
  53 124  16  24 |   b = saab
   4   2 210   2 |   c = bus
   2   4   2 191 |   d = van
```

OneR:

```
=== Classifier model (full training set) ===

SCALED VARIANCE_MINOR:
	'(-inf-298.5]'	-> van
	'(298.5-347.5]'	-> bus
```

```
       '(347.5-389.5]'       -> van
       '(389.5-581]'  -> saab
       '(581-721.5]'  -> saab
       '(721.5-761.5]'       -> opel
       '(761.5-inf)'  -> bus
(466/846 instances correct)


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         466               55.0827 %
Incorrectly Classified Instances       380               44.9173 %
Kappa statistic                          0.4015
Mean absolute error                      0.2246
Root mean squared error                  0.4739
Relative absolute error                 59.9144 %
Root relative squared error            109.4658 %
Total Number of Instances              846

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.104    0.003    0.917      0.104    0.186      0.263    0.550     0.320
opel
                 0.654    0.275    0.451      0.654    0.534      0.343    0.690     0.384
saab
                 0.647    0.127    0.638      0.647    0.642      0.517    0.760     0.504
bus
                 0.809    0.193    0.563      0.809    0.664      0.552    0.808     0.500
van
Weighted Avg.    0.551    0.150    0.642      0.551    0.505      0.417    0.701     0.426

=== Confusion Matrix ===

   a   b   c   d   <-- classified as
  22 121  22  47 |   a = opel
   0 142  20  55 |   b = saab
   2  52 141  23 |   c = bus
   0   0  38 161 |   d = van
```

References

Kibler, & Aha. (1991). Instance-based learning algorithms. *Machine Learning, vol. 6*, 37-66.

Mwagha, Muthoni, & Ochieg. (2014). Comparison of Nearest Neighbor (ibk), Regression by Discretization and Isotonic Regression Classification Algorithms for Precipitation Classes Prediction. *International Journal of Computer Applications, vol. 96, no. 21*, 0975-8887.

Nico, G. (2018, May 13). Retrieved from gerardnico.com/data_mining/knn: https://gerardnico.com/data_mining/knn

Obuandike, Isah, & Alhasan. (2015). Analytical Study of Some Selected Classification Algorithms in Weka Using Real Crime Data. *International Journal of Advanced Research in Artificial Intelligence, vol. 4, no. 12*.

Solanki. (2014). "Data Mining Techniques Using WEKA classification for Sickle Cell Disease". *International Journal of Computer Science and Information Technologies, vol. 5, no. 4*, 5857-5860.