Batch_1.arff
J48

Prueba 1

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
1
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): Sí
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): Sí
Número de palabras a conservar: 1000
J48:

Usar poda (Sí o No):
Sí
Confidence Factor:
0.25
Porcentaje de Split del conjunto (0.0 - 1.00):
0.70
```

```
Summary
======

Correctly Classified Instances         476               63.6364 %
Incorrectly Classified Instances       272               36.3636 %
Kappa statistic                          0.2683
Mean absolute error                      0.3664
Root mean squared error                  0.5866
Relative absolute error                 73.5939 %
Root relative squared error            117.5302 %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.593    0.325    0.617      0.593    0.605      0.268    0.644     0.576     0
                 0.675    0.407    0.652      0.675    0.663      0.268    0.644     0.634     1
Weighted Avg.    0.636    0.369    0.636      0.636    0.636      0.268    0.644     0.607


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 208 143 |   a = 0
 129 268 |   b = 1
```

*Usaremos minúsculas siempre porque obtienes mayor porcentaje

Prueba 2

```
Summary
======

Correctly Classified Instances         432               57.754  %
Incorrectly Classified Instances       316               42.246  %
Kappa statistic                          0.1505
Mean absolute error                      0.4175
Root mean squared error                  0.6233
Relative absolute error                 83.8721 %
Root relative squared error            124.8989 %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.536    0.385    0.551      0.536    0.543      0.151    0.598     0.537     0
                 0.615    0.464    0.600      0.615    0.607      0.151    0.598     0.600     1
Weighted Avg.    0.578    0.427    0.577      0.578    0.577      0.151    0.598     0.571


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 188 163 |   a = 0
 153 244 |   b = 1
```

Prueba 3

```
Summary
======

Correctly Classified Instances         480               64.1711 %
Incorrectly Classified Instances       268               35.8289 %
Kappa statistic                          0.2788
Mean absolute error                      0.3644
Root mean squared error                  0.5779
Relative absolute error                 73.2047 %
Root relative squared error            115.7909 %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.595    0.317    0.624      0.595    0.609      0.279    0.646     0.580     0
                 0.683    0.405    0.656      0.683    0.669      0.279    0.646     0.636     1
Weighted Avg.    0.642    0.364    0.641      0.642    0.641      0.279    0.646     0.610


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 209 142 |   a = 0
 126 271 |   b = 1
```

Prueba 4

```
Summary
======

Correctly Classified Instances         448               59.893  %
Incorrectly Classified Instances       300               40.107  %
Kappa statistic                          0.1911
Mean absolute error                      0.4059
Root mean squared error                  0.61
Relative absolute error                 81.536  %
Root relative squared error            122.233  %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.533    0.343    0.579      0.533    0.555      0.192    0.592     0.536     0
                 0.657    0.467    0.614      0.657    0.635      0.192    0.592     0.590     1
Weighted Avg.    0.599    0.409    0.598      0.599    0.597      0.192    0.592     0.565

Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 187 164 |   a = 0
 136 261 |   b = 1
```

Prueba 5

```
Summary
======

Correctly Classified Instances         468               62.5668 %
Incorrectly Classified Instances       280               37.4332 %
Kappa statistic                          0.2492
Mean absolute error                      0.3764
Root mean squared error                  0.5907
Relative absolute error                 75.605  %
Root relative squared error            118.3609 %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.610    0.360    0.599      0.610    0.605      0.249    0.628     0.562     0
                 0.640    0.390    0.650      0.640    0.645      0.249    0.628     0.621     1
Weighted Avg.    0.626    0.376    0.626      0.626    0.626      0.249    0.628     0.593

Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 214 137 |   a = 0
 143 254 |   b = 1
```

Prueba 6

```
Summary
======

Correctly Classified Instances         480               64.1711 %
Incorrectly Classified Instances       268               35.8289 %
Kappa statistic                          0.28
Mean absolute error                      0.3619
Root mean squared error                  0.5823
Relative absolute error                 72.6992 %
Root relative squared error            116.673  %
Total Number of Instances              748


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.610    0.330    0.620      0.610    0.615      0.280    0.651     0.592     0
                 0.670    0.390    0.660      0.670    0.665      0.280    0.651     0.634     1
Weighted Avg.    0.642    0.362    0.641      0.642    0.642      0.280    0.651     0.614


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 214 137 |   a = 0
 131 266 |   b = 1
```

Prueba 7

```
Summary
======

Correctly Classified Instances         502               59.1981 %
Incorrectly Classified Instances       346               40.8019 %
Kappa statistic                          0.1791
Mean absolute error                      0.4106
Root mean squared error                  0.6221
Relative absolute error                 82.5378 %
Root relative squared error            124.8544 %
Total Number of Instances              848


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.563    0.384    0.552      0.563    0.558      0.179    0.581     0.512     0
                 0.616    0.437    0.627      0.616    0.621      0.179    0.581     0.596     1
Weighted Avg.    0.592    0.413    0.593      0.592    0.592      0.179    0.581     0.558


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 218 169 |   a = 0
 177 284 |   b = 1
```

Prueba 8

```
Summary
======

Correctly Classified Instances        1333               59.3764 %
Incorrectly Classified Instances       912               40.6236 %
Kappa statistic                          0.1828
Mean absolute error                      0.4131
Root mean squared error                  0.6189
Relative absolute error                 83.2974 %
Root relative squared error            123.6733 %
Total Number of Instances             2245


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.549    0.367    0.570      0.549    0.559      0.183    0.589     0.526     0
                 0.633    0.451    0.614      0.633    0.623      0.183    0.589     0.595     1
Weighted Avg.    0.594    0.411    0.593      0.594    0.593      0.183    0.589     0.563


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 579 475 |   a = 0
 437 754 |   b = 1
```

Naive Bayes:

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         570               76.2032 %
Incorrectly Classified Instances       178               23.7968 %
Kappa statistic                          0.5235
Mean absolute error                      0.2423
Root mean squared error                  0.4537
Relative absolute error                 48.6649 %
Root relative squared error             90.9174 %
Total Number of Instances              748
```

Prueba 2

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         574               76.738  %
Incorrectly Classified Instances       174               23.262  %
Kappa statistic                          0.5333
Mean absolute error                      0.2382
Root mean squared error                  0.4514
Relative absolute error                 47.851  %
Root relative squared error             90.4533 %
Total Number of Instances              748
```

Prueba 3

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): Si
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         526               70.3209 %
Incorrectly Classified Instances       222               29.6791 %
Kappa statistic                          0.404
Mean absolute error                      0.3016
Root mean squared error                  0.5345
Relative absolute error                 60.5805 %
Root relative squared error            107.0887 %
Total Number of Instances              748
```

Prueba 4

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): Sí
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         491               65.6417 %
Incorrectly Classified Instances       257               34.3583 %
Kappa statistic                          0.3092
Mean absolute error                      0.3486
Root mean squared error                  0.5782
Relative absolute error                 70.0242 %
Root relative squared error            115.8503 %
Total Number of Instances              748
```

Prueba 5

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): Sí
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         574               76.738  %
Incorrectly Classified Instances       174               23.262  %
Kappa statistic                          0.5333
Mean absolute error                      0.2382
Root mean squared error                  0.4514
Relative absolute error                 47.851  %
Root relative squared error             90.4533 %
Total Number of Instances              748
```

Prueba 6

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         578               77.2727 %
Incorrectly Classified Instances       170               22.7273 %
Kappa statistic                          0.5454
Mean absolute error                      0.235
Root mean squared error                  0.445
Relative absolute error                 47.2137 %
Root relative squared error             89.1567 %
Total Number of Instances              748
```

Prueba 7

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 10000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         566               75.6684 %
Incorrectly Classified Instances       182               24.3316 %
Kappa statistic                          0.5086
Mean absolute error                      0.2456
Root mean squared error                  0.4726
Relative absolute error                 49.3364 %
Root relative squared error             94.6881 %
Total Number of Instances              748
```

Prueba 8

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_1.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): NO
Número de palabras a conservar: 10000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         561               75      %
Incorrectly Classified Instances       187               25      %
Kappa statistic                          0.4945
Mean absolute error                      0.2498
Root mean squared error                  0.4758
Relative absolute error                 50.1729 %
Root relative squared error             95.3398 %
Total Number of Instances              748
```

Bayes Net

Prueba 1

```
Time taken to build model: 1.59 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.2 seconds

=== Summary ===

Correctly Classified Instances         565               66.6274 %
Incorrectly Classified Instances       283               33.3726 %
Kappa statistic                          0.3198
Mean absolute error                      0.4001
Root mean squared error                  0.4616
Relative absolute error                 80.4214 %
Root relative squared error             92.7106 %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.601    0.282    0.626      0.601    0.613      0.320    0.725     0.659     0
                 0.718    0.399    0.696      0.718    0.707      0.320    0.725     0.764     1
Weighted Avg.    0.666    0.348    0.665      0.666    0.665      0.320    0.725     0.718

=== Confusion Matrix ===

   a   b   <-- classified as
 224 149 |   a = 0
 134 341 |   b = 1
```

Prueba 2

```
Time taken to build model: 1.4 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.18 seconds

=== Summary ===

Correctly Classified Instances         574               67.6887 %
Incorrectly Classified Instances       274               32.3113 %
Kappa statistic                          0.3435
Mean absolute error                      0.3871
Root mean squared error                  0.4567
Relative absolute error                 77.8162 %
Root relative squared error             91.7138 %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.627    0.284    0.634      0.627    0.631      0.344    0.740     0.677     0
                 0.716    0.373    0.710      0.716    0.713      0.344    0.740     0.777     1
Weighted Avg.    0.677    0.334    0.677      0.677    0.677      0.344    0.740     0.733

=== Confusion Matrix ===

   a   b   <-- classified as
 234 139 |   a = 0
 135 340 |   b = 1
```

Prueba 3

```
Time taken to build model: 1.16 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.19 seconds

=== Summary ===

Correctly Classified Instances         565               66.6274 %
Incorrectly Classified Instances       283               33.3726 %
Kappa statistic                          0.319
Mean absolute error                      0.4113
Root mean squared error                  0.4619
Relative absolute error                 82.669  %
Root relative squared error             92.761  %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.595    0.278    0.627      0.595    0.611      0.319    0.716     0.652     0
                 0.722    0.405    0.694      0.722    0.708      0.319    0.716     0.740     1
Weighted Avg.    0.666    0.349    0.665      0.666    0.665      0.319    0.716     0.701

=== Confusion Matrix ===

   a   b   <-- classified as
 222 151 |   a = 0
 132 343 |   b = 1
```

Prueba 4

```
Time taken to build model: 1.45 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.18 seconds

=== Summary ===

Correctly Classified Instances         559               65.9198 %
Incorrectly Classified Instances       289               34.0802 %
Kappa statistic                          0.3054
Mean absolute error                      0.4095
Root mean squared error                  0.4623
Relative absolute error                 82.3102 %
Root relative squared error             92.8561 %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.592    0.288    0.617      0.592    0.605      0.306    0.718     0.658     0
                 0.712    0.408    0.690      0.712    0.701      0.306    0.718     0.747     1
Weighted Avg.    0.659    0.355    0.658      0.659    0.658      0.306    0.718     0.708

=== Confusion Matrix ===

   a   b   <-- classified as
 221 152 |   a = 0
 137 338 |   b = 1
```

Prueba 5

```
Time taken to build model: 1.73 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.19 seconds

=== Summary ===

Correctly Classified Instances         566               66.7453 %
Incorrectly Classified Instances       282               33.2547 %
Kappa statistic                          0.3275
Mean absolute error                      0.3871
Root mean squared error                  0.4566
Relative absolute error                 77.819  %
Root relative squared error             91.7052 %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.638    0.309    0.618      0.638    0.628      0.328    0.743     0.687     0
               0.691    0.362    0.708      0.691    0.699      0.328    0.743     0.784     1
Weighted Avg.  0.667    0.339    0.669      0.667    0.668      0.328    0.743     0.741

=== Confusion Matrix ===

   a   b   <-- classified as
 238 135 |   a = 0
 147 328 |   b = 1
```

Prueba 6

```
Time taken to build model: 1.5 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.2 seconds

=== Summary ===

Correctly Classified Instances         565               66.6274 %
Incorrectly Classified Instances       283               33.3726 %
Kappa statistic                          0.3198
Mean absolute error                      0.4001
Root mean squared error                  0.4616
Relative absolute error                 80.4214 %
Root relative squared error             92.7106 %
Total Number of Instances              848

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.601    0.282    0.626      0.601    0.613      0.320    0.725     0.659     0
               0.718    0.399    0.696      0.718    0.707      0.320    0.725     0.764     1
Weighted Avg.  0.666    0.348    0.665      0.666    0.665      0.320    0.725     0.718

=== Confusion Matrix ===

   a   b   <-- classified as
 224 149 |   a = 0
 134 341 |   b = 1
```

Prueba 9

```
Time taken to build model: 1.33 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.61 seconds

=== Summary ===

Correctly Classified Instances        1900               76.1523 %
Incorrectly Classified Instances       595               23.8477 %
Kappa statistic                          0.5198
Mean absolute error                      0.3131
Root mean squared error                  0.4092
Relative absolute error                 62.9149 %
Root relative squared error             82.0394 %
Total Number of Instances              2495

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.728    0.209    0.752      0.728    0.740      0.520    0.833     0.802     0
               0.791    0.272    0.769      0.791    0.780      0.520    0.833     0.852     1
Weighted Avg.  0.762    0.243    0.761      0.762    0.761      0.520    0.833     0.829

=== Confusion Matrix ===

    a    b   <-- classified as
  846  316 |    a = 0
  279 1054 |    b = 1
```

Batch_2.arff

J48

Prueba 1

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
1
Sí
Sí
stopwords.txt
Sí
1000
Sí
0.25
0.70
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
Transformar todo a minúsculas (Sí o No): Usar frecuencias (Sí o No): Ruta al archivo de Stopwords (puede ser un archivo vacío): Usar relevancia

Usar poda (Sí o No):
Confidence Factor:
Porcentaje de Split del conjunto (0.0 - 1.00):
```

```
Summary
======

Correctly Classified Instances         398               66.3333 %
Incorrectly Classified Instances       202               33.6667 %
Kappa statistic                          0.3248
Mean absolute error                      0.336
Root mean squared error                  0.5578
Relative absolute error                 67.1262 %
Root relative squared error            111.406  %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.642    0.317    0.647      0.642    0.644      0.325    0.665     0.604     0
                 0.683    0.358    0.678      0.683    0.680      0.325    0.665     0.640     1
Weighted Avg.    0.663    0.339    0.663      0.663    0.663      0.325    0.665     0.623


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 183 102 |   a = 0
 100 215 |   b = 1
```

Prueba 2

```
Summary
======

Correctly Classified Instances         398               66.3333 %
Incorrectly Classified Instances       202               33.6667 %
Kappa statistic                          0.3248
Mean absolute error                      0.336
Root mean squared error                  0.5578
Relative absolute error                 67.1262 %
Root relative squared error            111.406  %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.642    0.317    0.647      0.642    0.644      0.325    0.665     0.604     0
                 0.683    0.358    0.678      0.683    0.680      0.325    0.665     0.640     1
Weighted Avg.    0.663    0.339    0.663      0.663    0.663      0.325    0.665     0.623


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 183 102 |   a = 0
 100 215 |   b = 1
```

Prueba 3

```
Summary
======

Correctly Classified Instances         400               66.6667 %
Incorrectly Classified Instances       200               33.3333 %
Kappa statistic                          0.3341
Mean absolute error                      0.3316
Root mean squared error                  0.5547
Relative absolute error                 66.2505 %
Root relative squared error            110.7919 %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.688    0.352    0.638      0.688    0.662      0.335    0.680     0.608     0
                 0.648    0.312    0.696      0.648    0.671      0.335    0.680     0.664     1
Weighted Avg.    0.667    0.331    0.669      0.667    0.667      0.335    0.680     0.637


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 196  89 |   a = 0
 111 204 |   b = 1
```

Prueba 4

```
Summary
======

Correctly Classified Instances         395               65.8333 %
Incorrectly Classified Instances       205               34.1667 %
Kappa statistic                          0.316
Mean absolute error                      0.3434
Root mean squared error                  0.5662
Relative absolute error                 68.6061 %
Root relative squared error            113.092  %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.656    0.340    0.636      0.656    0.646      0.316    0.669     0.614     0
                 0.660    0.344    0.680      0.660    0.670      0.316    0.669     0.639     1
Weighted Avg.    0.658    0.342    0.659      0.658    0.659      0.316    0.669     0.627


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 187  98 |   a = 0
 107 208 |   b = 1
```

Prueba 5

```
Summary
======

Correctly Classified Instances         408               68      %
Incorrectly Classified Instances       192               32      %
Kappa statistic                          0.359
Mean absolute error                      0.322
Root mean squared error                  0.5472
Relative absolute error                 64.3373 %
Root relative squared error            109.3029 %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.674    0.314    0.660      0.674    0.667      0.359    0.690     0.638     0
                 0.686    0.326    0.699      0.686    0.692      0.359    0.690     0.656     1
Weighted Avg.    0.680    0.321    0.680      0.680    0.680      0.359    0.690     0.647


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 192  93 |   a = 0
  99 216 |   b = 1
```

Prueba 6

```
Summary
======

Correctly Classified Instances         381               63.5    %
Incorrectly Classified Instances       219               36.5    %
Kappa statistic                          0.269
Mean absolute error                      0.363
Root mean squared error                  0.5832
Relative absolute error                 72.5174 %
Root relative squared error            116.489  %
Total Number of Instances              600


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.628    0.359    0.613      0.628    0.620      0.269    0.652     0.602     0
                 0.641    0.372    0.656      0.641    0.648      0.269    0.652     0.621     1
Weighted Avg.    0.635    0.366    0.635      0.635    0.635      0.269    0.652     0.612


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 179 106 |   a = 0
 113 202 |   b = 1
```

Prueba 7

```
Summary
======

Correctly Classified Instances         453               66.6176 %
Incorrectly Classified Instances       227               33.3824 %
Kappa statistic                          0.3301
Mean absolute error                      0.3359
Root mean squared error                  0.5539
Relative absolute error                 67.1199 %
Root relative squared error            110.6318 %
Total Number of Instances              680


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.634    0.304    0.656      0.634    0.645      0.330    0.685     0.628     0
                 0.696    0.366    0.675      0.696    0.685      0.330    0.685     0.653     1
Weighted Avg.    0.666    0.337    0.666      0.666    0.666      0.330    0.685     0.641


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 206 119 |   a = 0
 108 247 |   b = 1
```

Prueba 8

```
Summary
======

Correctly Classified Instances        1000               55.5556 %
Incorrectly Classified Instances       800               44.4444 %
Kappa statistic                          0.1115
Mean absolute error                      0.4491
Root mean squared error                  0.6594
Relative absolute error                 89.7758 %
Root relative squared error            131.6004 %
Total Number of Instances              1800


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.593    0.481    0.549      0.593    0.570      0.112    0.539     0.522     0
                 0.519    0.407    0.564      0.519    0.540      0.112    0.539     0.522     1
Weighted Avg.    0.556    0.444    0.556      0.556    0.555      0.112    0.539     0.522


Confusion Matrix: false positives and false negatives
======

   a   b   <-- classified as
 530 364 |   a = 0
 436 470 |   b = 1
```

Naive Bayes

Prueba 1

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         514               85.6667 %
Incorrectly Classified Instances        86               14.3333 %
Kappa statistic                          0.7123
Mean absolute error                      0.1452
Root mean squared error                  0.3584
Relative absolute error                 29.0022 %
Root relative squared error             71.5839 %
Total Number of Instances              600
```

Prueba 2

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         514               85.6667 %
Incorrectly Classified Instances        86               14.3333 %
Kappa statistic                          0.7123
Mean absolute error                      0.1452
Root mean squared error                  0.3584
Relative absolute error                 29.0022 %
Root relative squared error             71.5839 %
Total Number of Instances              600
```

Prueba 3

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): Sí
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         425               70.8333 %
Incorrectly Classified Instances       175               29.1667 %
Kappa statistic                          0.4178
Mean absolute error                      0.2912
Root mean squared error                  0.5347
Relative absolute error                 58.1771 %
Root relative squared error            106.8039 %
Total Number of Instances              600
```

Prueba 4

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): Sí
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         425               70.8333 %
Incorrectly Classified Instances       175               29.1667 %
Kappa statistic                          0.4178
Mean absolute error                      0.2912
Root mean squared error                  0.5347
Relative absolute error                 58.1771 %
Root relative squared error            106.8039 %
Total Number of Instances              600
```

Prueba 5

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): Sí
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         514               85.6667 %
Incorrectly Classified Instances        86               14.3333 %
Kappa statistic                          0.7123
Mean absolute error                      0.1452
Root mean squared error                  0.3584
Relative absolute error                 29.0022 %
Root relative squared error             71.5839 %
Total Number of Instances              600
```

Prueba 6

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         511               85.1667 %
Incorrectly Classified Instances        89               14.8333 %
Kappa statistic                          0.7028
Mean absolute error                      0.1508
Root mean squared error                  0.3563
Relative absolute error                 30.129  %
Root relative squared error             71.1611 %
Total Number of Instances              600
```

Prueba 7

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 10000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         514               85.6667 %
Incorrectly Classified Instances        86               14.3333 %
Kappa statistic                          0.7112
Mean absolute error                      0.1443
Root mean squared error                  0.3649
Relative absolute error                 28.838  %
Root relative squared error             72.8844 %
Total Number of Instances              600
```

Prueba 8

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_2.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 10000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances         508               84.6667 %
Incorrectly Classified Instances        92               15.3333 %
Kappa statistic                          0.6911
Mean absolute error                      0.1504
Root mean squared error                  0.3723
Relative absolute error                 30.0549 %
Root relative squared error             74.3607 %
Total Number of Instances              600
```

Bayes Net

Prueba 1

```
Time taken to build model: 1.11 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.14 seconds

=== Summary ===

Correctly Classified Instances         468               78      %
Incorrectly Classified Instances       132               22      %
Kappa statistic                          0.56
Mean absolute error                      0.2565
Root mean squared error                  0.3923
Relative absolute error                 51.2895 %
Root relative squared error             78.4613 %
Total Number of Instances              600

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.781     0.221     0.776      0.781     0.779      0.560     0.870     0.874     0
               0.779     0.219     0.784      0.779     0.781      0.560     0.870     0.873     1
Weighted Avg.  0.780     0.220     0.780      0.780     0.780      0.560     0.870     0.873

=== Confusion Matrix ===

   a   b   <-- classified as
 232  65 |   a = 0
  67 236 |   b = 1
```

Prueba 2

```
Time taken to build model: 1.1 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.14 seconds

=== Summary ===

Correctly Classified Instances         468               78      %
Incorrectly Classified Instances       132               22      %
Kappa statistic                          0.56
Mean absolute error                      0.2565
Root mean squared error                  0.3923
Relative absolute error                 51.2895 %
Root relative squared error             78.4613 %
Total Number of Instances              600

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.781     0.221     0.776      0.781     0.779      0.560     0.870     0.874     0
               0.779     0.219     0.784      0.779     0.781      0.560     0.870     0.873     1
Weighted Avg.  0.780     0.220     0.780      0.780     0.780      0.560     0.870     0.873

=== Confusion Matrix ===

   a   b   <-- classified as
 232  65 |   a = 0
  67 236 |   b = 1
```

Prueba 3

```
Time taken to build model: 1.09 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.14 seconds

=== Summary ===

Correctly Classified Instances         463               77.1667 %
Incorrectly Classified Instances       137               22.8333 %
Kappa statistic                          0.5432
Mean absolute error                      0.266
Root mean squared error                  0.4
Relative absolute error                 53.1878 %
Root relative squared error             79.9902 %
Total Number of Instances              600

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.761    0.218    0.774      0.761    0.767      0.543    0.859     0.865     0
                 0.782    0.239    0.769      0.782    0.776      0.543    0.859     0.857     1
Weighted Avg.    0.772    0.229    0.772      0.772    0.772      0.543    0.859     0.861

=== Confusion Matrix ===

   a   b   <-- classified as
 226  71 |   a = 0
  66 237 |   b = 1
```

Prueba 4

```
Time taken to build model: 1.07 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.14 seconds

=== Summary ===

Correctly Classified Instances         457               76.1667 %
Incorrectly Classified Instances       143               23.8333 %
Kappa statistic                          0.523
Mean absolute error                      0.2738
Root mean squared error                  0.4072
Relative absolute error                 54.7551 %
Root relative squared error             81.4338 %
Total Number of Instances              600

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.727    0.205    0.777      0.727    0.751      0.524    0.850     0.857     0
                 0.795    0.273    0.748      0.795    0.771      0.524    0.850     0.846     1
Weighted Avg.    0.762    0.239    0.763      0.762    0.761      0.524    0.850     0.852

=== Confusion Matrix ===

   a   b   <-- classified as
 216  81 |   a = 0
  62 241 |   b = 1
```

Prueba 5

```
Time taken to build model: 1.07 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.14 seconds

=== Summary ===

Correctly Classified Instances         457               76.1667 %
Incorrectly Classified Instances       143               23.8333 %
Kappa statistic                          0.523
Mean absolute error                      0.2738
Root mean squared error                  0.4072
Relative absolute error                 54.7551 %
Root relative squared error             81.4338 %
Total Number of Instances              600

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.727    0.205    0.777      0.727    0.751      0.524    0.850     0.857     0
                 0.795    0.273    0.748      0.795    0.771      0.524    0.850     0.846     1
Weighted Avg.    0.762    0.239    0.763      0.762    0.761      0.524    0.850     0.852

=== Confusion Matrix ===

   a   b   <-- classified as
 216  81 |   a = 0
  62 241 |   b = 1
```

Prueba 9

```
me taken to build model: 1.1 seconds

= Evaluation on training set ===

me taken to test model on training data: 0.46 seconds

= Summary ===

rrectly Classified Instances        1704               85.2    %
correctly Classified Instances       296               14.8    %
ppa statistic                          0.704
an absolute error                      0.1938
ot mean squared error                  0.334
lative absolute error                 38.7545 %
ot relative squared error             66.7976 %
tal Number of Instances             2000

= Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.860    0.156    0.846      0.860    0.853      0.704    0.924     0.922     0
                 0.844    0.140    0.858      0.844    0.851      0.704    0.924     0.925     1
ighted Avg.      0.852    0.148    0.852      0.852    0.852      0.704    0.924     0.924

= Confusion Matrix ===

  a   b   <-- classified as
50 140 |   a = 0
56 844 |   b = 1
```

Batch_3.arff

J48

Prueba 1

```
Summary
======

Correctly Classified Instances        5610               74.8    %
Incorrectly Classified Instances      1890               25.2    %
Kappa statistic                          0.496
Mean absolute error                      0.2677
Root mean squared error                  0.4803
Relative absolute error                 53.5408 %
Root relative squared error             96.057  %
Total Number of Instances             7500


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.725    0.229    0.760      0.725    0.742      0.496    0.745     0.694     1
                 0.771    0.275    0.738      0.771    0.754      0.496    0.745     0.696     0
Weighted Avg.    0.748    0.252    0.748      0.748    0.748      0.496    0.745     0.695


Confusion Matrix: false positives and false negatives
======

    a    b   <-- classified as
 2716 1030 |    a = 1
  860 2894 |    b = 0
```

Prueba 2

```
Summary
======

Correctly Classified Instances        5545               73.9333 %
Incorrectly Classified Instances      1955               26.0667 %
Kappa statistic                          0.4786
Mean absolute error                      0.2782
Root mean squared error                  0.4881
Relative absolute error                 55.6484 %
Root relative squared error             97.6298 %
Total Number of Instances             7500


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.723    0.244    0.747      0.723    0.735      0.479    0.732     0.684     1
                 0.756    0.277    0.732      0.756    0.744      0.479    0.732     0.672     0
Weighted Avg.    0.739    0.261    0.740      0.739    0.739      0.479    0.732     0.678


Confusion Matrix: false positives and false negatives
======

    a    b   <-- classified as
 2708 1038 |    a = 1
  917 2837 |    b = 0
```

Prueba 3

```
Summary
======

Correctly Classified Instances        5549               73.9867 %
Incorrectly Classified Instances      1951               26.0133 %
Kappa statistic                          0.4797
Mean absolute error                      0.2749
Root mean squared error                  0.4906
Relative absolute error                 54.9791 %
Root relative squared error             98.1206 %
Total Number of Instances             7500


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.726    0.247    0.746      0.726    0.736      0.480    0.728     0.671     1
                 0.753    0.274    0.734      0.753    0.744      0.480    0.728     0.674     0
Weighted Avg.    0.740    0.260    0.740      0.740    0.740      0.480    0.728     0.673


Confusion Matrix: false positives and false negatives
======

    a    b   <-- classified as
 2721 1025 |    a = 1
  926 2828 |    b = 0
```

Prueba 4

```
Summary
======

Correctly Classified Instances        5526               73.68   %
Incorrectly Classified Instances      1974               26.32   %
Kappa statistic                          0.4736
Mean absolute error                      0.2765
Root mean squared error                  0.4946
Relative absolute error                 55.2929 %
Root relative squared error             98.926  %
Total Number of Instances             7500


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.729    0.255    0.740      0.729    0.734      0.474    0.718     0.672     1
                 0.745    0.271    0.733      0.745    0.739      0.474    0.718     0.662     0
Weighted Avg.    0.737    0.263    0.737      0.737    0.737      0.474    0.718     0.667


Confusion Matrix: false positives and false negatives
======

    a    b   <-- classified as
 2729 1017 |    a = 1
  957 2797 |    b = 0
```

Prueba 5

```
Summary
======

Correctly Classified Instances        5529               73.72   %
Incorrectly Classified Instances      1971               26.28   %
Kappa statistic                          0.4744
Mean absolute error                      0.2769
Root mean squared error                  0.4973
Relative absolute error                 55.3746 %
Root relative squared error             99.4541 %
Total Number of Instances             7500


Class Details
======

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.724    0.249    0.743      0.724    0.733      0.475    0.711     0.665     1
                 0.751    0.276    0.731      0.751    0.741      0.475    0.711     0.656     0
Weighted Avg.    0.737    0.263    0.737      0.737    0.737      0.475    0.711     0.661


Confusion Matrix: false positives and false negatives
======

    a    b   <-- classified as
 2711 1035 |    a = 1
  936 2818 |    b = 0
```

Naive Bayes

Prueba 1

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        6236               83.1467 %
Incorrectly Classified Instances      1264               16.8533 %
Kappa statistic                          0.6629
Mean absolute error                      0.1737
Root mean squared error                  0.3785
Relative absolute error                 34.7457 %
Root relative squared error             75.7084 %
Total Number of Instances             7500
```

Prueba 2

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        6217               82.8933 %
Incorrectly Classified Instances      1283               17.1067 %
Kappa statistic                          0.6579
Mean absolute error                      0.1789
Root mean squared error                  0.3827
Relative absolute error                 35.7874 %
Root relative squared error             76.5317 %
Total Number of Instances            7500
```

Prueba 3

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): Si
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        5814               77.52   %
Incorrectly Classified Instances      1686               22.48   %
Kappa statistic                          0.5504
Mean absolute error                      0.2261
Root mean squared error                  0.4588
Relative absolute error                 45.2247 %
Root relative squared error             91.7653 %
Total Number of Instances            7500
```

Prueba 4

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): Sí
Usar frecuencias (Sí o No): Sí
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        5858               78.1067 %
Incorrectly Classified Instances      1642               21.8933 %
Kappa statistic                          0.5621
Mean absolute error                      0.2212
Root mean squared error                  0.4544
Relative absolute error                 44.2367 %
Root relative squared error             90.8828 %
Total Number of Instances            7500
```

Prueba 5

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío): stopwords.txt
Usar relevancia de palabra en el documento (Sí o No): Si
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        6217               82.8933 %
Incorrectly Classified Instances      1283               17.1067 %
Kappa statistic                          0.6579
Mean absolute error                      0.1789
Root mean squared error                  0.3827
Relative absolute error                 35.7874 %
Root relative squared error             76.5317 %
Total Number of Instances             7500
```

Prueba 6

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 1000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        6148               81.9733 %
Incorrectly Classified Instances      1352               18.0267 %
Kappa statistic                          0.6395
Mean absolute error                      0.1859
Root mean squared error                  0.3964
Relative absolute error                 37.1851 %
Root relative squared error             79.2753 %
Total Number of Instances             7500
```

Prueba 7

```
run:
Ruta al archivo con extensión <<.arff>>: Batch_3.arff
Crearemos el filtro (StringToWordsVector personalizado)
Algoritmo a utilizar:
[1] J48
[2] Bayes Ingenuo
2
Transformar todo a minúsculas (Sí o No): No
Usar frecuencias (Sí o No): No
Ruta al archivo de Stopwords (puede ser un archivo vacío):
Usar relevancia de palabra en el documento (Sí o No): No
Número de palabras a conservar: 10000
Naive Bayes:

Porcentaje de Split del conjunto (0.0 - 1.00):
0.70

Correctly Classified Instances        6177               82.36   %
Incorrectly Classified Instances      1323               17.64   %
Kappa statistic                          0.6472
Mean absolute error                      0.1795
Root mean squared error                  0.39
Relative absolute error                 35.8987 %
Root relative squared error             77.9976 %
Total Number of Instances             7500
```

Prueba 8

Bayes Net

Prueba 1

```
Time taken to build model: 33.65 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.53 seconds

=== Summary ===

Correctly Classified Instances        6320               84.2667 %
Incorrectly Classified Instances      1180               15.7333 %
Kappa statistic                          0.6854
Mean absolute error                      0.177
Root mean squared error                  0.3494
Relative absolute error                 35.3949 %
Root relative squared error             69.8823 %
Total Number of Instances             7500

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.878    0.192    0.819      0.878    0.847      0.687    0.922     0.921     1
                 0.808    0.122    0.870      0.808    0.838      0.687    0.922     0.923     0
Weighted Avg.    0.843    0.157    0.844      0.843    0.842      0.687    0.922     0.922

=== Confusion Matrix ===

    a    b   <-- classified as
 3274  456 |    a = 1
  724 3046 |    b = 0
```

Prueba 2

```
Time taken to build model: 34.4 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.63 seconds

=== Summary ===

Correctly Classified Instances        6282               83.76   %
Incorrectly Classified Instances      1218               16.24   %
Kappa statistic                          0.6753
Mean absolute error                      0.1801
Root mean squared error                  0.3513
Relative absolute error                 36.0263 %
Root relative squared error             70.2682 %
Total Number of Instances             7500

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.873    0.197    0.814      0.873    0.842      0.677    0.920     0.919     1
                 0.803    0.127    0.864      0.803    0.833      0.677    0.920     0.920     0
Weighted Avg.    0.838    0.162    0.839      0.838    0.837      0.677    0.920     0.920

=== Confusion Matrix ===

    a    b   <-- classified as
 3255  475 |    a = 1
  743 3027 |    b = 0
```

Prueba 3

```
Time taken to build model: 35.53 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.53 seconds

=== Summary ===

Correctly Classified Instances        6295               83.9333 %
Incorrectly Classified Instances      1205               16.0667 %
Kappa statistic                          0.6788
Mean absolute error                      0.1788
Root mean squared error                  0.3511
Relative absolute error                 35.7553 %
Root relative squared error             70.2221 %
Total Number of Instances             7500

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.875    0.195    0.816      0.875    0.844      0.681    0.921     0.919     1
                 0.805    0.125    0.866      0.805    0.834      0.681    0.921     0.921     0
Weighted Avg.    0.839    0.160    0.841      0.839    0.839      0.681    0.921     0.920

=== Confusion Matrix ===

    a    b   <-- classified as
 3262  468 |    a = 1
  737 3033 |    b = 0
```

Prueba 4

```
Time taken to build model: 38.62 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.79 seconds

=== Summary ===

Correctly Classified Instances        6097               81.2933 %
Incorrectly Classified Instances      1403               18.7067 %
Kappa statistic                          0.626
Mean absolute error                      0.1979
Root mean squared error                  0.3863
Relative absolute error                 39.5729 %
Root relative squared error             77.2607 %
Total Number of Instances             7500

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.852    0.226    0.789      0.852    0.819      0.628    0.900     0.899     1
                 0.774    0.148    0.841      0.774    0.806      0.628    0.900     0.899     0
Weighted Avg.    0.813    0.187    0.815      0.813    0.813      0.628    0.900     0.899

=== Confusion Matrix ===

    a    b   <-- classified as
 3179  551 |    a = 1
  852 2918 |    b = 0
```

Prueba 5

```
Time taken to build model: 41.38 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.74 seconds

=== Summary ===

Correctly Classified Instances        6132               81.76   %
Incorrectly Classified Instances      1368               18.24   %
Kappa statistic                          0.6354
Mean absolute error                      0.1936
Root mean squared error                  0.3827
Relative absolute error                 38.728  %
Root relative squared error             76.5342 %
Total Number of Instances             7500

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.861    0.225    0.791      0.861    0.824      0.638    0.904     0.904     1
                 0.775    0.139    0.849      0.775    0.810      0.638    0.904     0.903     0
Weighted Avg.    0.818    0.182    0.820      0.818    0.817      0.638    0.904     0.903

=== Confusion Matrix ===

    a    b   <-- classified as
 3210  520 |    a = 1
  848 2922 |    b = 0
```

Prueba 9

```
Time taken to build model: 36.06 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 5.32 seconds

=== Summary ===

Correctly Classified Instances       21104               84.416  %
Incorrectly Classified Instances      3896               15.584  %
Kappa statistic                          0.6883
Mean absolute error                      0.1713
Root mean squared error                  0.3467
Relative absolute error                 34.2542 %
Root relative squared error             69.3476 %
Total Number of Instances            25000

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.878    0.190    0.822      0.878    0.849      0.690    0.925     0.925     1
                 0.810    0.122    0.869      0.810    0.839      0.690    0.925     0.923     0
Weighted Avg.    0.844    0.156    0.846      0.844    0.844      0.690    0.925     0.924

=== Confusion Matrix ===

     a     b   <-- classified as
 10977  1523 |     a = 1
  2373 10127 |     b = 0
```