#### **EVALUATING ACCURACY OF CLASSIFIER**

## 1) LOGISTIC REGRESSION (functions.Logistic)

```
=== Run information ===
Scheme:weka.classifiers.functions.Logistic -R 1.0E-8 -M -1
Relation: german_credit
Instances: 1000
Attributes: 21
       checking_status
       duration
       credit_history
       purpose
       credit_amount
       savings_status
       employment
       installment_commitment
       personal_status
       other_parties
       residence_since
       property_magnitude
       age
       other_payment_plans
       housing
       existing_credits
       job
       num_dependents
```

own\_telephone

foreign\_worker

class

Test mode:10-fold cross-validation

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

Logistic Regression with ridge parameter of 1.0E-8

Coefficients...

Class

Variable good

\_\_\_\_\_

checking\_status=<0 -0.778

checking\_status=0<=X<200 -0.4032

checking\_status=>=200 0.1877

checking\_status=no checking 0.9338

duration -0.0279

credit\_history=no credits/all paid -0.8129

credit\_history=all paid -0.9562

credit\_history=existing paid -0.2268

credit\_history=delayed previously 0.0403

credit\_history=critical/other existing credit 0.6229

purpose=new car -0.692

purpose=used car 0.9744

purpose=furniture/equipment 0.0996

purpose=radio/tv 0.1996

purpose=domestic appliance -0.1692

purpose=repairs -0.4756

purpose=education	-0.7283
purpose=vacation	0
purpose=retraining	1.3674
purpose=business	0.0481
purpose=other	0.7968
credit_amount	-0.0001
savings_status=<100	-0.4402
savings_status=100<=X<500	-0.0825
savings_status=500<=X<1000	-0.0641
savings_status=>=1000	0.8989
savings_status=no known savings	0.5065
employment=unemployed	-0.2934
employment=<1	-0.2265
employment=1<=X<4	-0.1106
employment=4<=X<7	0.5376
employment=>=7	-0.0168
installment_commitment	-0.3301
personal_status=male div/sep	-0.4923
personal_status=female div/dep/	mar -0.2168
personal_status=male single	0.3238
personal_status=male mar/wid	-0.1252
personal_status=female single	0
other_parties=none	-0.1798
other_parties=co applicant	-0.6158
other_parties=guarantor	0.7988
residence_since	-0.0048
property_magnitude=real estate	0.2572

property\_magnitude=life insurance -0.0242

property\_magnitude=car 0.0627

property\_magnitude=no known property -0.4732

age 0.0145

other\_payment\_plans=bank -0.3273

other\_payment\_plans=stores -0.2041

other\_payment\_plans=none 0.3191

housing=rent -0.3498

housing=own 0.0939

housing=for free 0.3341

existing\_credits -0.2721

job=unemp/unskilled non res 0.5095

job=unskilled resident -0.0265

job=skilled -0.0451

job=high qualif/self emp/mgmt 0.0301

num\_dependents -0.2647

own\_telephone 0.3

foreign\_worker 1.3922

Intercept 3.1983

Odds Ratios...

Class

Variable good

checking\_status=<0 0.4593

checking\_status=0<=X<200 0.6682

checking	status=>=200	1.2064

checking\_status=no checking 2.5443

duration 0.9725

credit\_history=no credits/all paid 0.4436

credit\_history=all paid 0.3843

credit\_history=existing paid 0.7971

credit\_history=delayed previously 1.0411

credit\_history=critical/other existing credit 1.8643

purpose=new car 0.5006

purpose=used car 2.6497

purpose=furniture/equipment 1.1047

purpose=radio/tv 1.2209

purpose=domestic appliance 0.8443

purpose=repairs 0.6215

purpose=education 0.4827

purpose=vacation 1

purpose=retraining 3.9251

purpose=business 1.0493

purpose=other 2.2184

credit\_amount 0.9999

savings\_status=<100 0.6439

savings\_status=500<=X<1000 0.9379

savings\_status=>=1000 2.4569

savings\_status=no known savings 1.6594

employment=unemployed 0.7457

employment=<1 0.7973

).8953
)

employment=4<=X<7 1.7119

employment=>=7 0.9834

installment\_commitment 0.7189

personal\_status=male div/sep 0.6112

personal\_status=female div/dep/mar 0.8051

personal\_status=male single 1.3824

personal\_status=male mar/wid 0.8823

personal\_status=female single 1

other\_parties=none 0.8354

other\_parties=co applicant 0.5402

other\_parties=guarantor 2.2229

residence\_since 0.9952

property\_magnitude=real estate 1.2933

property\_magnitude=life insurance 0.9761

property\_magnitude=car 1.0647

property\_magnitude=no known property 0.623

age 1.0146

other\_payment\_plans=bank 0.7209

other\_payment\_plans=stores 0.8154

other\_payment\_plans=none 1.3758

housing=rent 0.7049

housing=own 1.0984

housing=for free 1.3967

existing\_credits 0.7618

job=unemp/unskilled non res 1.6645

job=unskilled resident 0.9738

job=skilled 0.9559

job=high qualif/self emp/mgmt 1.0306

num\_dependents 0.7675

own\_telephone 1.3499

foreign\_worker 4.0237

Time taken to build model: 0.06 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 752 75.2 %

Incorrectly Classified Instances 248 24.8 %

Kappa statistic 0.375

Mean absolute error 0.3098

Root mean squared error 0.4087

Relative absolute error 73.727 %

Root relative squared error 89.1751 %

**Total Number of Instances** 1000

## === Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.864 0.51 0.798 0.864 0.83 0.785 good

Weighted Avg. 0.752 0.398 0.741 0.752 0.744 0.785

```
=== Confusion Matrix ===
```

a b <-- classified as

605 95 | a = good

153 147 | b = bad

## 2) NAIVE BAYES ALGORITHM (bayes. NavieBayes)

=== Run information ===

Scheme:weka.classifiers.bayes.NaiveBayes

Relation: german\_credit

Instances: 1000

Attributes: 21

checking\_status

duration

credit\_history

purpose

 $credit\_amount$ 

savings\_status

employment

 $in stall ment\_commitment$ 

personal\_status

other\_parties

residence\_since

property\_magnitude

age

```
other_payment_plans
     housing
     existing_credits
     job
     num_dependents
     own_telephone
     foreign_worker
     class
Test mode:10-fold cross-validation
=== Classifier model (full training set) ===
Naive Bayes Classifier
               Class
Attribute
                   good
                         bad
               (0.7) (0.3)
checking_status
<0
               140.0 136.0
0<=X<200
                   165.0 106.0
        50.0 15.0
>=200
no checking
                   349.0 47.0
            704.0 304.0
[total]
duration
```

19.1766 24.8129

mean

std. dev. 10.9817 13.3608

weight sum 700 300

precision 2.125 2.125

credit\_history

no credits/all paid 16.0 26.0

all paid 22.0 29.0

existing paid 362.0 170.0

delayed previously 61.0 29.0

critical/other existing credit 244.0 51.0

[total] 705.0 305.0

purpose

new car 146.0 90.0

used car 87.0 18.0

furniture/equipment 124.0 59.0

radio/tv 219.0 63.0

domestic appliance 9.0 5.0

repairs 15.0 9.0

education 29.0 23.0

vacation 1.0 1.0

retraining 9.0 2.0

business 64.0 35.0

other 8.0 6.0

[total] 711.0 311.0

credit\_amount

mean 2985.6721 3938.1609

std. dev. 2399.7801 3529.4788

weight sum 700 300

precision 19.7543 19.7543

savings\_status

<100 387.0 218.0

100<=X<500 70.0 35.0

500<=X<1000 53.0 12.0

>=1000 43.0 7.0

no known savings 152.0 33.0

[total] 705.0 305.0

employment

unemployed 40.0 24.0

<1 103.0 71.0

1<=X<4 236.0 105.0

4<=X<7 136.0 40.0

>=7 190.0 65.0

[total] 705.0 305.0

 $installment\_commitment$ 

mean 2.92 3.0967

std. dev. 1.1273 1.0866

weight sum 700 300

precision 1 1

#### personal\_status

male div/sep 31.0 21.0

female div/dep/mar 202.0 110.0

male single 403.0 147.0

male mar/wid 68.0 26.0

female single 1.0 1.0

[total] 705.0 305.0

#### other\_parties

none 636.0 273.0

co applicant 24.0 19.0

guarantor 43.0 11.0

[total] 703.0 303.0

## residence\_since

mean 2.8429 2.85

std. dev. 1.1076 1.0928

weight sum 700 300

precision 1 1

## property\_magnitude

real estate 223.0 61.0

life insurance 162.0 72.0

car 231.0 103.0

no known property 88.0 68.0

[total] 704.0 304.0

age

mean 36.1723 33.9267

std. dev. 11.4005 11.259

weight sum 700 300

precision 1.0769 1.0769

## other\_payment\_plans

bank 83.0 58.0

stores 29.0 20.0

none 591.0 225.0

[total] 703.0 303.0

#### housing

rent 110.0 71.0

own 528.0 187.0

for free 65.0 45.0

[total] 703.0 303.0

#### existing\_credits

mean 1.4243 1.3667

std. dev. 0.5843 0.5588

weight sum 700 300

precision 1 1

### job

unemp/unskilled non res 16.0 8.0

unskilled resident 145.0 57.0

skilled 445.0 187.0

high qualif/self emp/mgmt 98.0 52.0

[total] 704.0 304.0

## num\_dependents

mean 1.1557 1.1533

std. dev. 0.3626 0.3603

weight sum 700 300

precision 1 1

#### own\_telephone

none 410.0 188.0

yes 292.0 114.0

[total] 702.0 302.0

## foreign\_worker

yes 668.0 297.0

no 34.0 5.0

[total] 702.0 302.0

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 754 75.4 %

Incorrectly Classified Instances 246 24.6 %

Kappa statistic 0.3813

Mean absolute error 0.2936

Root mean squared error 0.4201

Relative absolute error 69.8801 %

Root relative squared error 91.6718 %

Total Number of Instances 1000

## === Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.864 0.503 0.8 0.864 0.831 0.787 good

Weighted Avg. 0.754 0.393 0.743 0.754 0.746 0.787

#### === Confusion Matrix ===

a b <-- classified as

605 95 | a = good

151 149 | b = bad

#### 3)J48 ALGORITHM (trees.J48)

=== Run information ===

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: german\_credit

Instances: 1000

```
Attributes: 21
       checking_status
       duration
       credit_history
       purpose
       credit_amount
       savings_status
       employment
       installment\_commitment
       personal_status
       other_parties
       residence_since
       property_magnitude
       age
       other_payment_plans
       housing
       existing_credits
       job
       num_dependents
       own_telephone
       foreign_worker
       class
Test mode:10-fold cross-validation
=== Classifier model (full training set) ===
```

J48 pruned tree

-----

```
checking_status = <0
| foreign_worker = yes
| | duration <= 11
| | existing_credits <= 1
| | | property_magnitude = real estate: good (8.0/1.0)
| | | property_magnitude = car: good (2.0/1.0)
| | | property_magnitude = no known property: bad (3.0)
| | duration > 11
| | job = unemp/unskilled non res: bad (5.0/1.0)
| | job = unskilled resident
| | | | employment = unemployed: good (0.0)
| | | purpose = radio/tv
```

```
| | | purpose = domestic appliance: bad (1.0)
| | | purpose = retraining: good (1.0)
| | | purpose = business: good (3.0)
| | | | | | credit_history = all paid: bad (6.0)
| | | | | credit_history = existing paid
| | | | | | | property_magnitude = real estate
| | | | | | | age <= 26: bad (5.0)
| | | | | | | age > 26: good (2.0)
| | | | | | | | | credit_amount <= 1386: bad (3.0)
| | | | | | | | | | credit_amount > 1386: good (11.0/1.0)
| | | | | | | existing_credits > 1: bad (3.0)
```

```
| | | | | credit_history = delayed previously: bad (4.0)
| | | | | credit_history = critical/other existing credit: good (14.0/4.0)
| | | | | | credit_history = no credits/all paid: good (0.0)
| | | | | | credit_history = delayed previously: good (0.0)
| | | | | credit_history = critical/other existing credit: good (2.0)
| | | | | savings_status = no known savings
| | | | | | | own_telephone = none: bad (9.0/1.0)
| | | | | | own_telephone = yes: good (4.0/1.0)
| | | | | existing_credits > 1: good (2.0)
| | | other_parties = co applicant: bad (7.0/1.0)
| | | other_parties = guarantor: good (12.0/3.0)
| | job = high qualif/self emp/mgmt: good (30.0/8.0)
| foreign_worker = no: good (15.0/2.0)
checking_status = 0<=X<200
| credit_amount <= 9857
| | savings_status = <100
```

```
| | | | personal_status = female div/dep/mar
| | | | | purpose = new car: bad (5.0/1.0)
| | | | purpose = furniture/equipment
| | | | | duration <= 10: bad (3.0)
| | | | | duration > 10
| | | | | | duration <= 21: good (6.0/1.0)
| | | | | duration > 21: bad (2.0)
| | | | | purpose = radio/tv: good (8.0/2.0)
| | | | | purpose = domestic appliance: good (0.0)
| | | | | purpose = education: good (4.0/2.0)
| | | | | residence_since <= 2: good (3.0)
| | | | personal_status = male mar/wid
| | | | duration <= 10: good (6.0)
| | | | duration > 10: bad (10.0/3.0)
| | | | personal_status = female single: good (0.0)
| | other_parties = co applicant: good (2.0)
```

```
| | | purpose = furniture/equipment: good (0.0)
| | | purpose = domestic appliance: good (0.0)
| | | purpose = education: good (0.0)
| | | purpose = vacation: good (0.0)
| | | purpose = retraining: good (0.0)
| | | purpose = business: good (0.0)
| | savings_status = 100<=X<500
| | purpose = used car: good (3.0)
| | purpose = furniture/equipment: bad (4.0/1.0)
| | purpose = domestic appliance: good (0.0)
| | purpose = repairs: good (2.0)
| | purpose = education: good (0.0)
| | purpose = vacation: good (0.0)
| | purpose = retraining: good (0.0)
```

```
| | savings_status = 500<=X<1000: good (11.0/3.0)
| savings_status = >=1000: good (13.0/3.0)
| savings_status = no known savings: good (41.0/5.0)
| credit_amount > 9857: bad (20.0/3.0)
| checking_status = >=200: good (63.0/14.0)
| checking_status = no checking: good (394.0/46.0)
```

Number of Leaves: 103

Size of the tree: 140

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 705 70.5 %

Incorrectly Classified Instances 295 29.5 %

Kappa statistic 0.2467

Mean absolute error 0.3467

Root mean squared error 0.4796

Relative absolute error 82.5233 %

Root relative squared error 104.6565 %

**Total Number of Instances** 1000

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.84 0.61 0.763 0.84 0.799 0.639 good
0.39 0.16 0.511 0.39 0.442 0.639 bad
Weighted Avg. 0.705 0.475 0.687 0.705 0.692 0.639

=== Confusion Matrix ===

a b <-- classified as

588 112 | a = good

183 117 | b = bad



#### 4) K-NEAREST NEIGHBOUR (lazy.IBK)

=== Run information ===

Scheme:weka.classifiers.lazy.lBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: german\_credit

```
Instances: 1000
Attributes: 21
       checking_status
       duration
       credit_history
       purpose
       credit_amount
       savings_status
       employment
       installment\_commitment
       personal_status
       other_parties
       residence_since
       property_magnitude
       age
       other_payment_plans
       housing
       existing_credits
      job
      num_dependents
       own_telephone
       foreign_worker
       class
```

Test mode:10-fold cross-validation

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

#### **IB1** instance-based classifier

#### using 1 nearest neighbour(s) for classification

#### Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 720 72 %

Incorrectly Classified Instances 280 28 %

Kappa statistic 0.3243

Mean absolute error 0.2805

Root mean squared error 0.5286

Relative absolute error 66.7546 %

Root relative squared error 115.3422 %

Total Number of Instances 1000

#### === Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.51 0.19 0.535 0.51 0.522 0.66 bad

Weighted Avg. 0.72 0.4 0.716 0.72 0.718 0.66

=== Confusion Matrix ===

```
a b <-- classified as
567 133 | a = good
147 153 | b = bad
5)SMO ALGORITHM (functions.SMO)
=== Run information ===
Scheme:weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K
"weka.classifiers.functions.supportVector.PolyKernel -C 250007 -E 1.0"
Relation: german_credit
Instances: 1000
Attributes: 21
       checking_status
       duration
       credit_history
       purpose
       credit_amount
       savings_status
       employment
       installment\_commitment
       personal_status
       other_parties
       residence_since
       property_magnitude
       age
       other_payment_plans
       housing
       existing_credits
```

job

```
num_dependents
       own_telephone
       foreign_worker
       class
Test mode:10-fold cross-validation
=== Classifier model (full training set) ===
SMO
Kernel used:
 Linear Kernel: K(x,y) = \langle x,y \rangle
Classifier for classes: good, bad
BinarySMO
Machine linear: showing attribute weights, not support vectors.
    0.6805 * (normalized) checking_status=<0
     0.3347 * (normalized) checking_status=0<=X<200
    -0.4616 * (normalized) checking_status=>=200
    -0.5537 * (normalized) checking_status=no checking
     1.6987 * (normalized) duration
     0.5398 * (normalized) credit_history=no credits/all paid
     0.6015 * (normalized) credit_history=all paid
```

-0.109 \* (normalized) credit\_history=existing paid

- + -0.3182 \* (normalized) credit\_history=delayed previously
- + -0.7141 \* (normalized) credit\_history=critical/other existing credit
- + 0.5673 \* (normalized) purpose=new car
- + -0.5615 \* (normalized) purpose=used car
- + -0.1464 \* (normalized) purpose=furniture/equipment
- + -0.0798 \* (normalized) purpose=radio/tv
- + 0.5456 \* (normalized) purpose=domestic appliance
- + 0 \* (normalized) purpose=repairs
- + 0.4441 \* (normalized) purpose=education
- + -0.3951 \* (normalized) purpose=retraining
- + -0.0823 \* (normalized) purpose=business
- + -0.2919 \* (normalized) purpose=other
- + 1.1473 \* (normalized) credit\_amount
- + 0.4056 \* (normalized) savings\_status=<100
- + 0.115 \* (normalized) savings\_status=100<=X<500
- + 0.1378 \* (normalized) savings\_status=500<=X<1000
- + -0.3775 \* (normalized) savings\_status=>=1000
- + -0.2809 \* (normalized) savings\_status=no known savings
- + 0.2887 \* (normalized) employment=unemployed
- + 0.1663 \* (normalized) employment=<1
- + 0.0021 \* (normalized) employment=1<=X<4
- + -0.3348 \* (normalized) employment=4<=X<7
- + -0.1222 \* (normalized) employment=>=7
- + 0.6503 \* (normalized) installment commitment
- + 0.3335 \* (normalized) personal\_status=male div/sep
- + 0.1177 \* (normalized) personal\_status=female div/dep/mar
- + -0.3697 \* (normalized) personal\_status=male single

- + -0.0815 \* (normalized) personal\_status=male mar/wid
- + 0.0514 \* (normalized) other\_parties=none
- + 0.5697 \* (normalized) other\_parties=co applicant
- + -0.6211 \* (normalized) other\_parties=guarantor
- + -0.0001 \* (normalized) residence\_since
- + -0.2247 \* (normalized) property\_magnitude=real estate
- + -0.0544 \* (normalized) property\_magnitude=life insurance
- + -0.0795 \* (normalized) property\_magnitude=car
- + 0.3586 \* (normalized) property\_magnitude=no known property
- + -0.4191 \* (normalized) age
- + 0.0697 \* (normalized) other\_payment\_plans=bank
- + 0.159 \* (normalized) other\_payment\_plans=stores
- + -0.2287 \* (normalized) other\_payment\_plans=none
- + 0.3271 \* (normalized) housing=rent
- + -0.0702 \* (normalized) housing=own
- + -0.257 \* (normalized) housing=for free
- + 0.4503 \* (normalized) existing\_credits
- + -0.2026 \* (normalized) job=unemp/unskilled non res
- + 0.1501 \* (normalized) job=unskilled resident
- + 0.1027 \* (normalized) job=skilled
- + -0.0502 \* (normalized) job=high qualif/self emp/mgmt
- + 0.0198 \* (normalized) num\_dependents
- + -0.1394 \* (normalized) own\_telephone
- + -0.9888 \* (normalized) foreign worker
- 1.5398

Number of kernel evaluations: 436644 (90.558% cached)

#### Time taken to build model: 0.11 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 751 75.1 %

Incorrectly Classified Instances 249 24.9 %

Kappa statistic 0.3654

Mean absolute error 0.249

Root mean squared error 0.499

Relative absolute error 59.2607 %

Root relative squared error 108.8905 %

**Total Number of Instances** 1000

#### === Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.871 0.53 0.793 0.871 0.83 0.671 good

Weighted Avg. 0.751 0.41 0.738 0.751 0.741 0.671

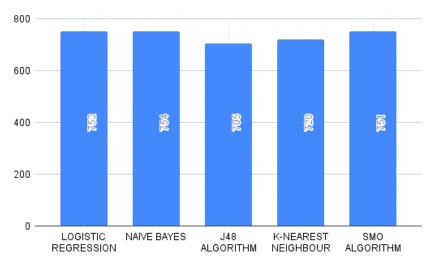
=== Confusion Matrix ===

a b <-- classified as

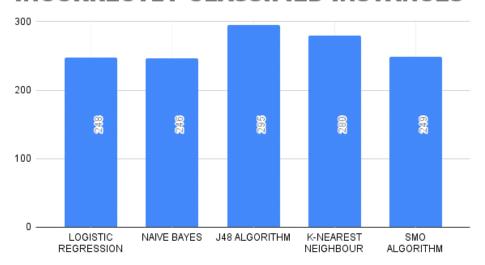
610 90 | a = good

159 141 | b = bad

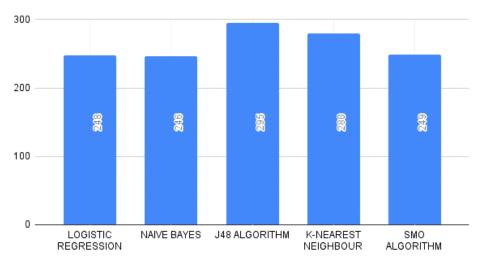
## **CORRECTLY CLASSIFIED INSTANCES**



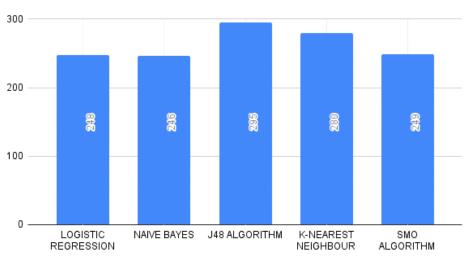
# **INCORRECTLY CLASSIFIED INSTANCES**



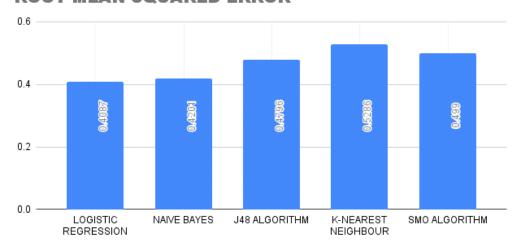
## **INCORRECTLY CLASSIFIED INSTANCES**



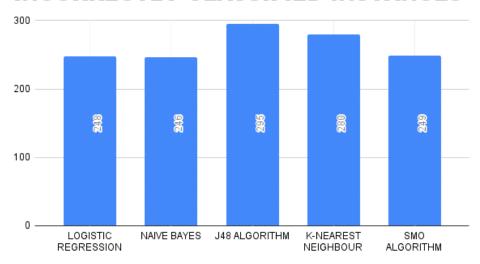
## **INCORRECTLY CLASSIFIED INSTANCES**

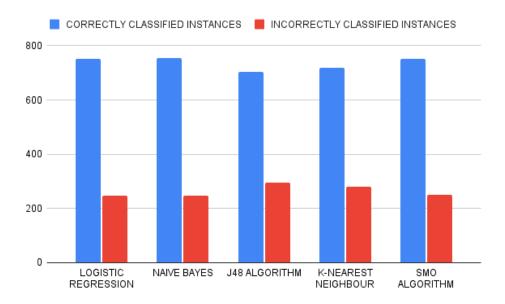


## **ROOT MEAN SQUARED ERROR**



# **INCORRECTLY CLASSIFIED INSTANCES**





# **INCORRECTLY CLASSIFIED INSTANCES**

