KAUNO TECHNOLOGIJOS UNIVERSITETAS
INFORMATIKOS FAKULTETAS

Mašininio mokymosi metodai (P160B124)

Laboratorinių darbų ataskaita

Atliko:
IFF-1/4 gr. studentas
Mildaras Karvelis
2023 m. spalio 22 d.
Priėmė:
Doc. Iešmantas Tomas

KAUNAS 2023

TURINYS

1. 2.1. Lab task: linear regression

1.1. 1 task

Duomenų lentelėje yra 14 ypatybių
R kalboje tipai yra:

```
Date                 character
Rented_Bike_Count      integer
Hour                   integer
Temperature            numeric
Humidity               integer
Wind_speed             numeric
Visibility             integer
Dew_point_temperature  numeric
Solar_Radiation        numeric
Rainfall               numeric
Snowfall               numeric
Seasons              character
Holiday              character
Functioning_Day      character
```

Bike_test turi 1560 informacijos eilučių
Bike_train turi 7200 informacijos eilučių

1.2. 2 task

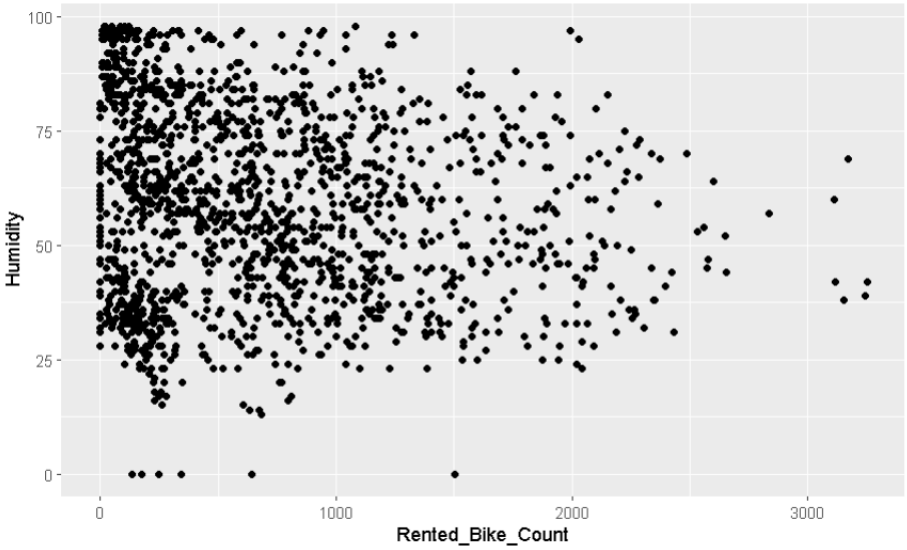


Diagrama parodo tai, jog esant normaliam drėgnumui (25%-70%) yra populiariausia nuomotis dviratį.

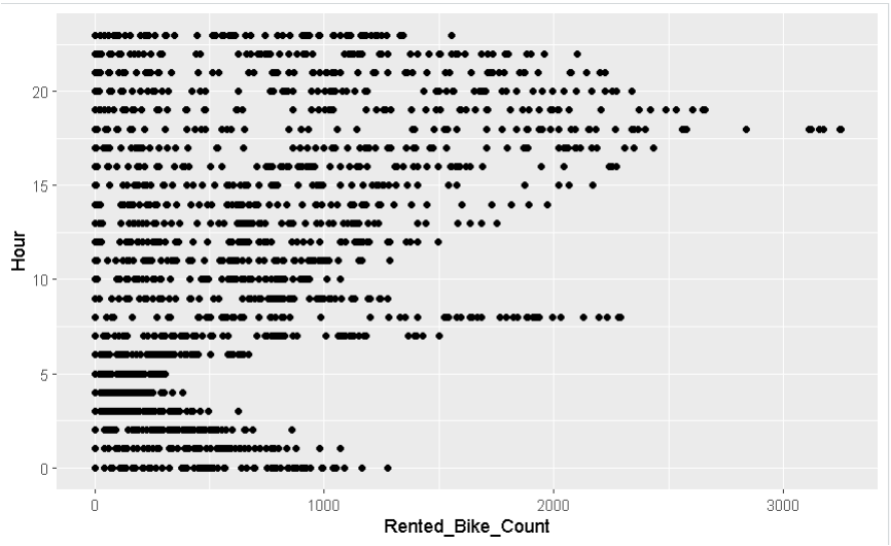
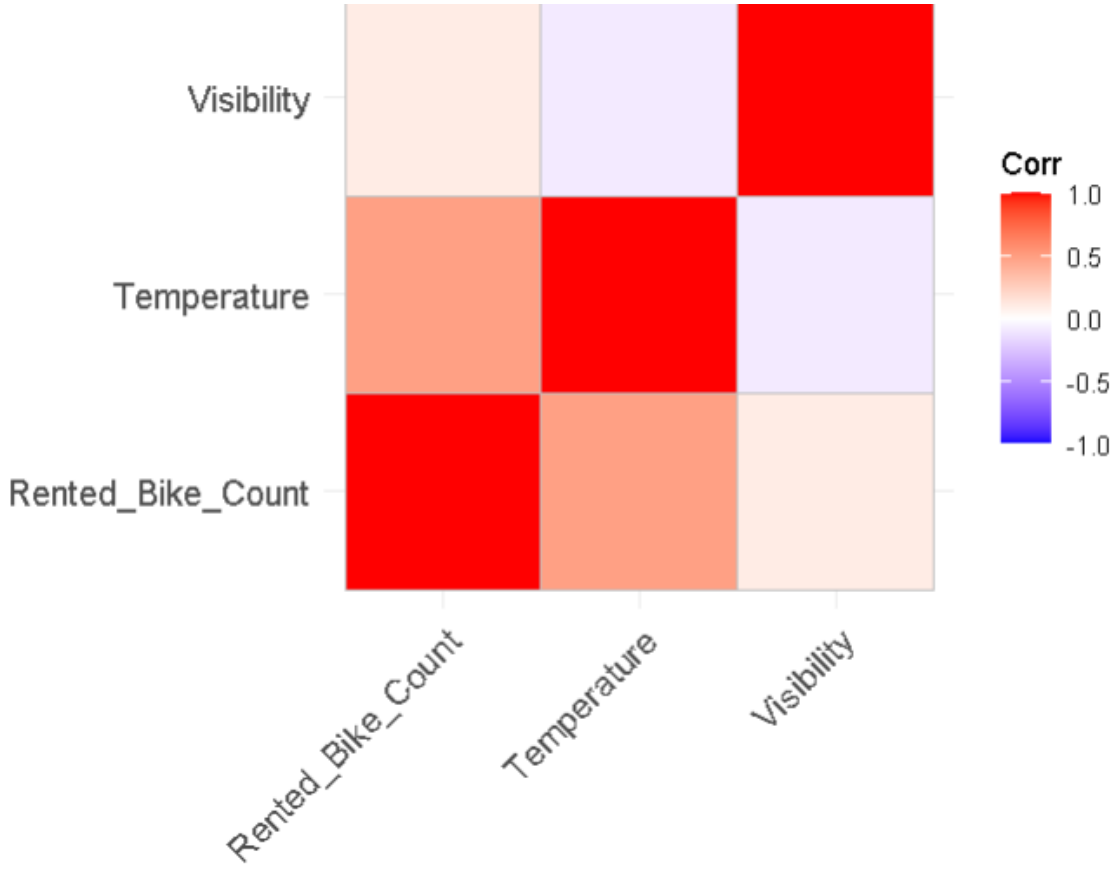
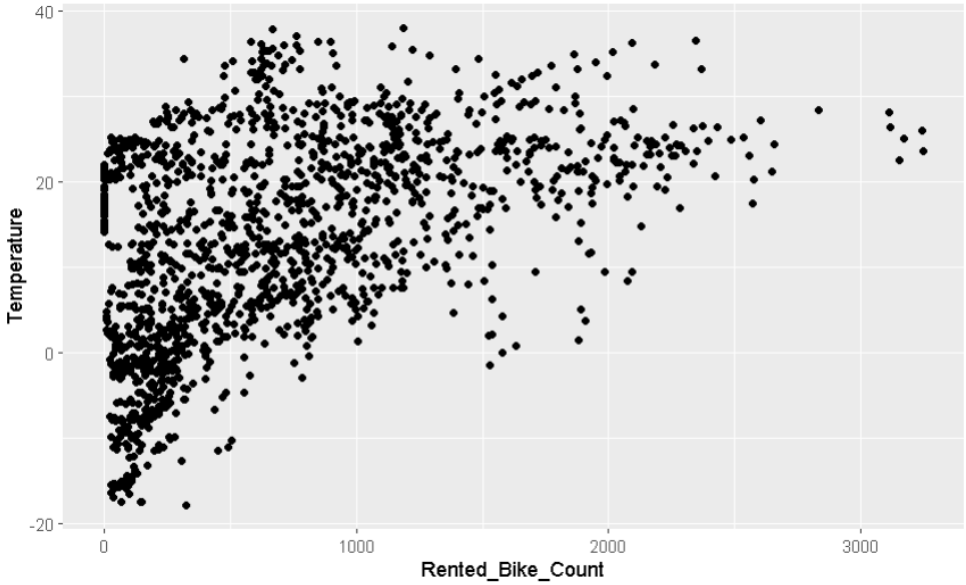


Diagrama parodo tai, jog labiausiai populiariausia nuomotis dviračius yra tarp 17 ir 22 valandos.

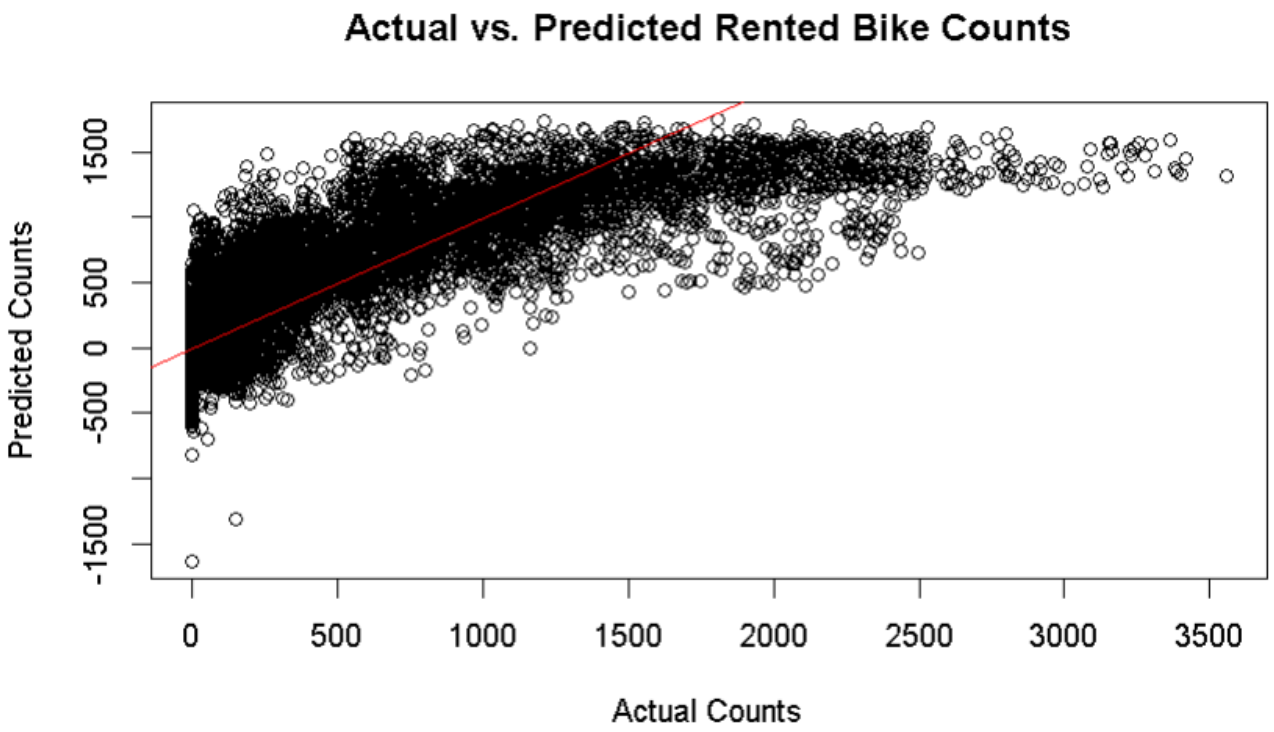
1.3. 3 task





Manyčiau, jog pašalinus vieną iš duomenų, galime prarasti nemažai informacijos, pvz. Pašalinsime „Rented_Bike_Count“, prarasime daugiau kaip 50% turimos informacijos.

1.4. 4 task



Spėtas kiekis yra daug didesnis nei tikrasis kiekis.

1.5. 5 task

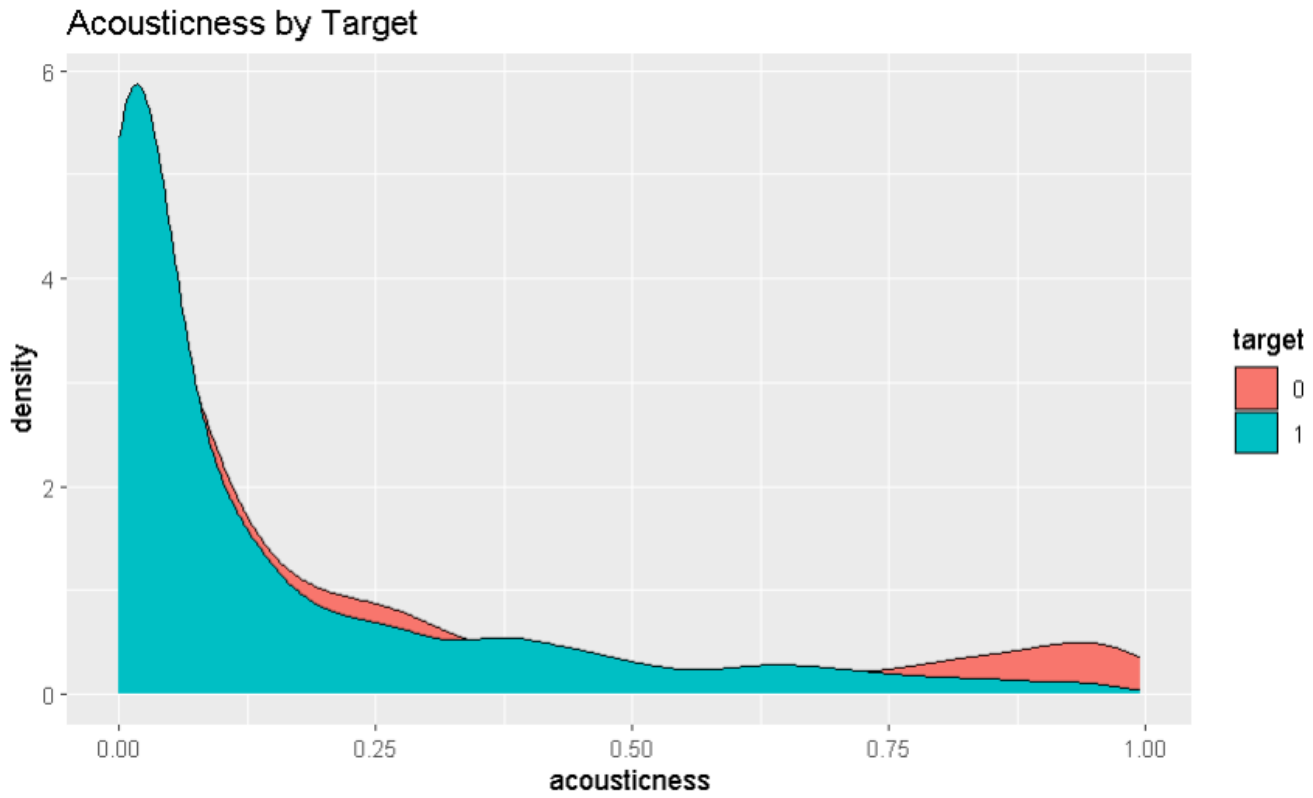
```
> cat( R-squared value: , rsq
R-squared value: 0.5561676>
```

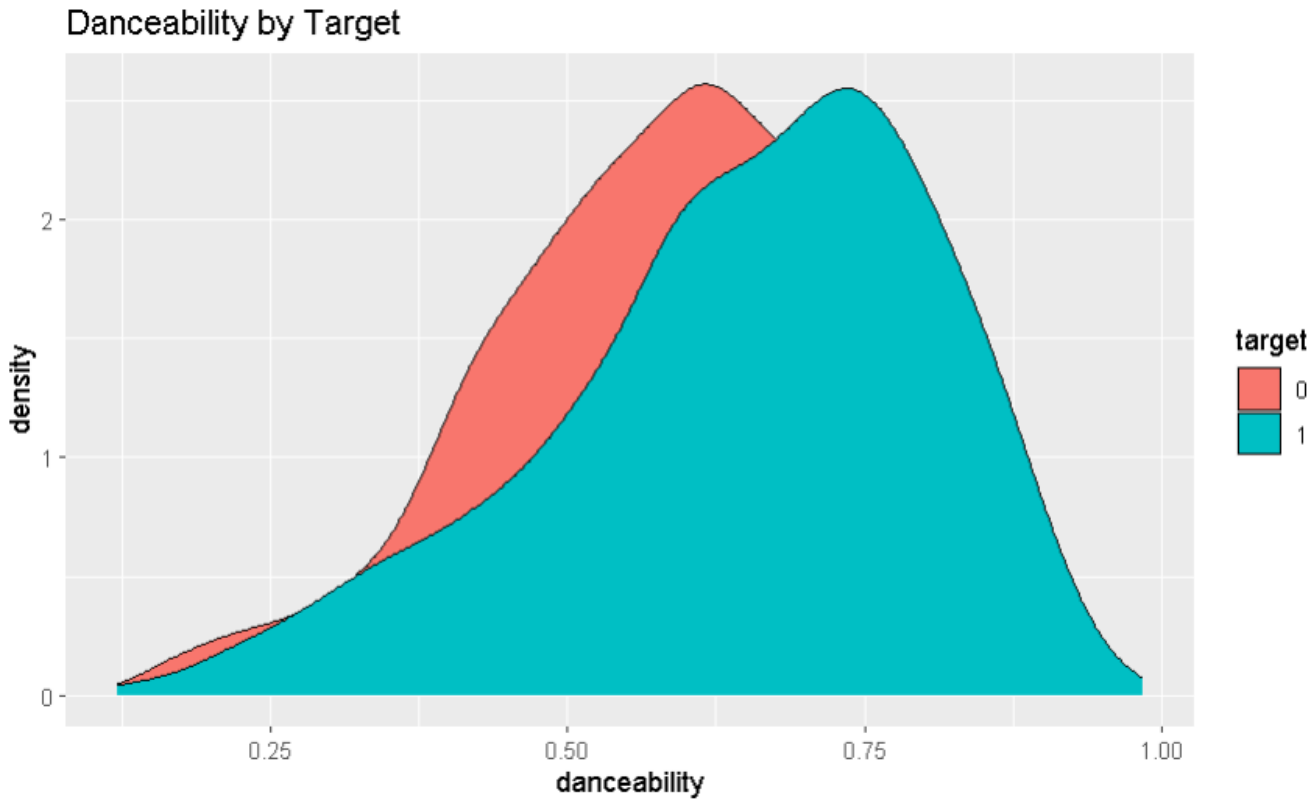
1.6. 6 task

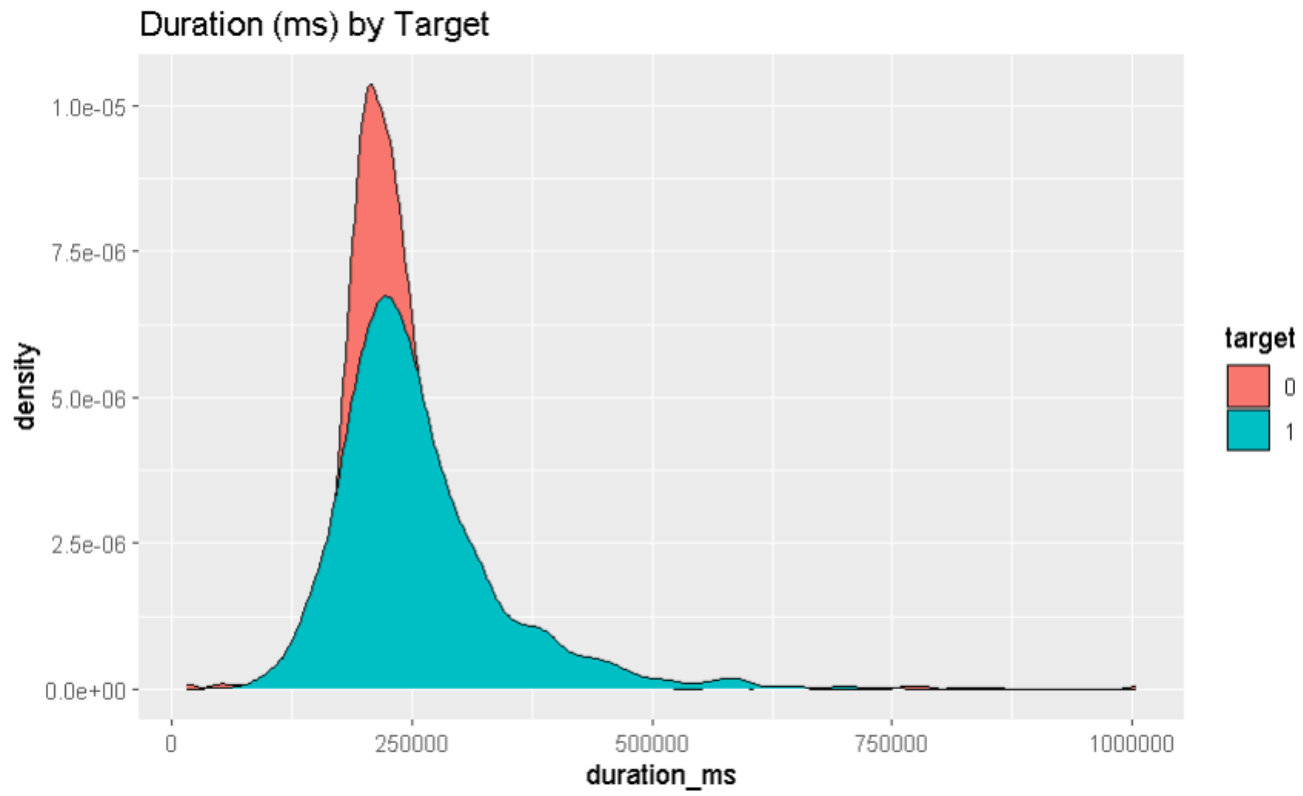
	Pair	Model	R2
1	Hour × Temperature	With Interaction	0.7930496
2	Hour × Temperature	Without Interaction	0.7653556
3	Hour × Humidity	With Interaction	0.7955672
4	Hour × Humidity	Without Interaction	0.7533521
5	Temperature × Humidity	With Interaction	0.7938326
6	Temperature × Humidity	Without Interaction	0.7609191
7	Humidity × Wind_speed	With Interaction	0.7929549
8	Humidity × Wind_speed	Without Interaction	0.7824659

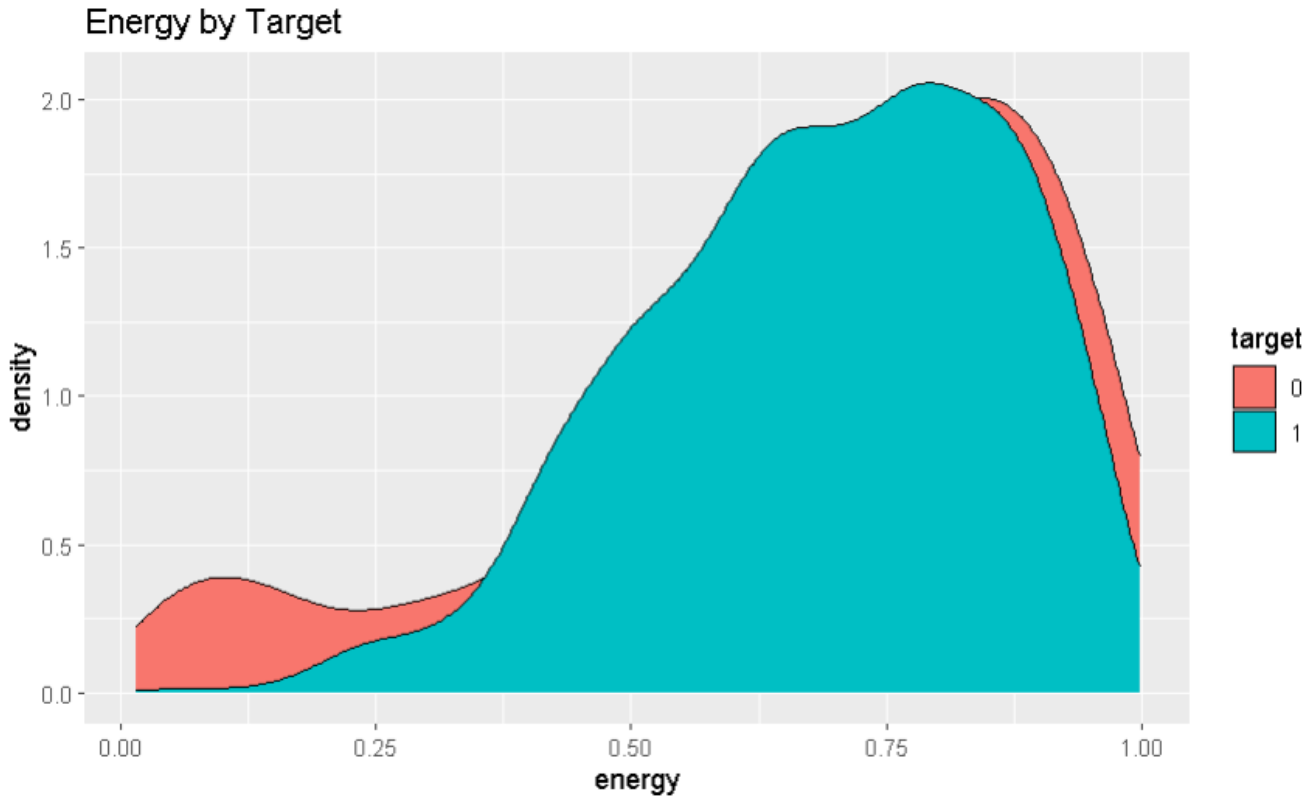
2. 3.1. Lab task: logistic regression

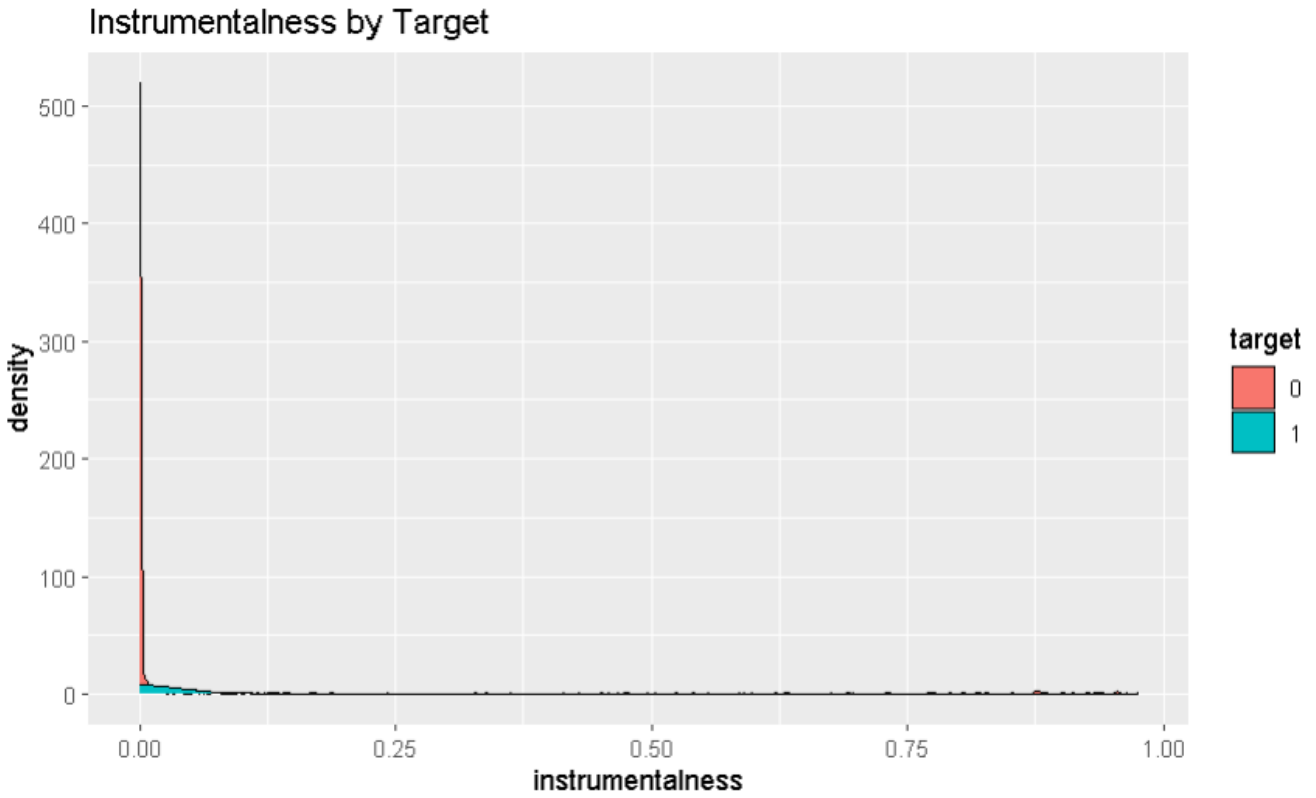
2.1. 2 task



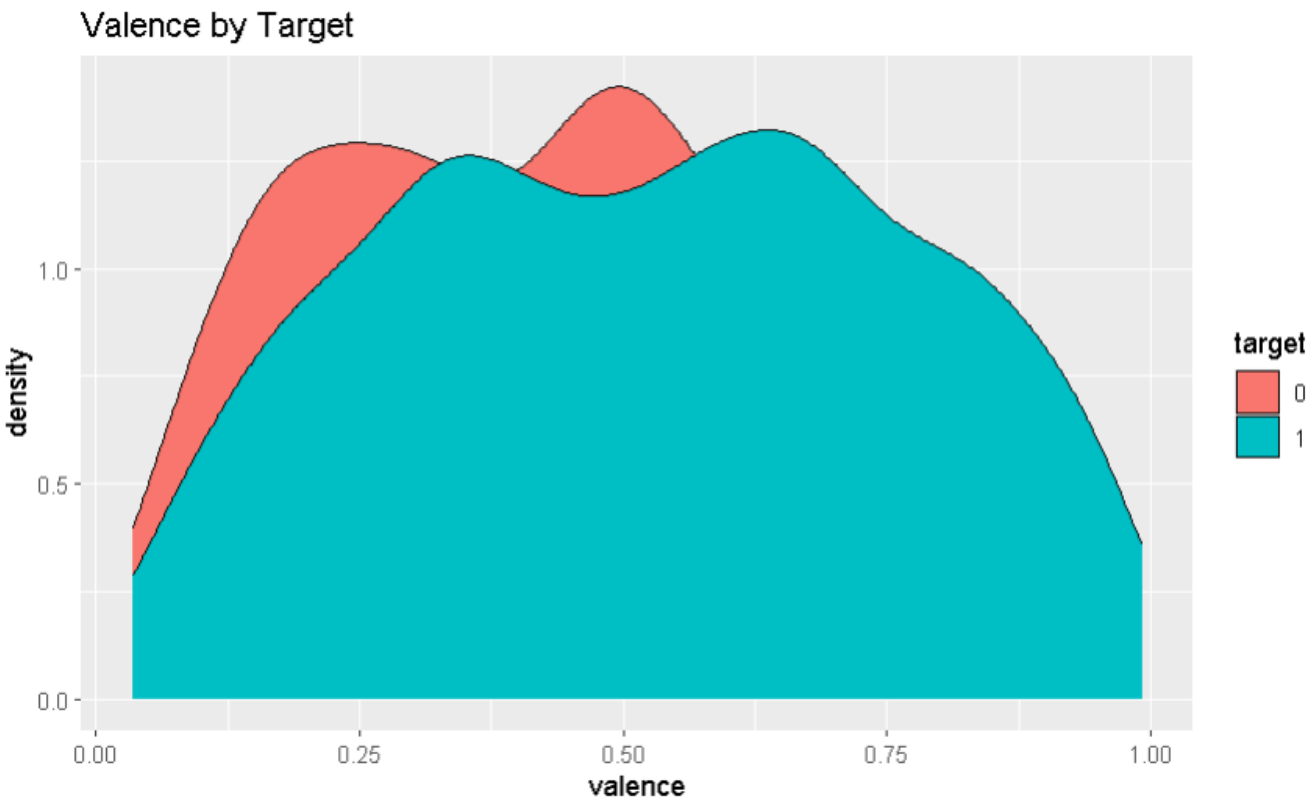


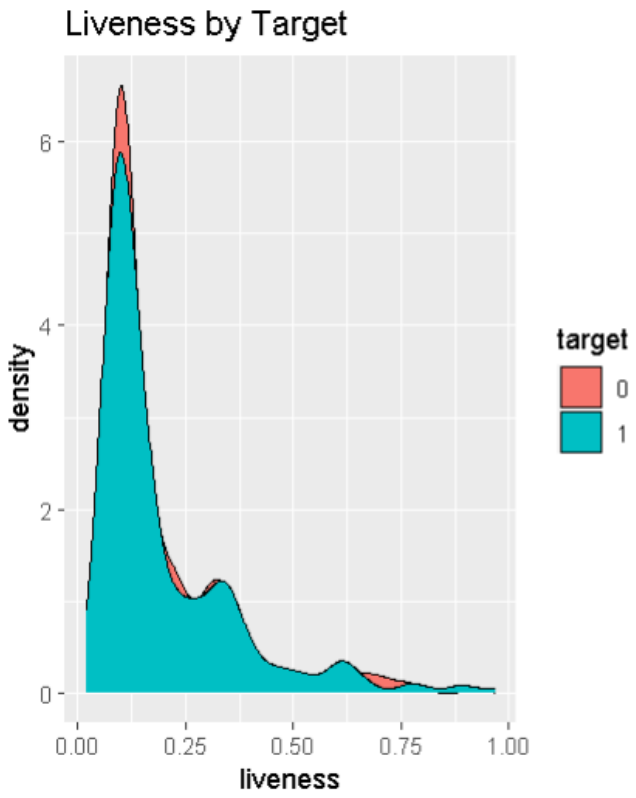

Duration (ms) by Target
density
1.0e-05
7.5e-06
5.0e-06
2.5e-06
0.0e+00
0
250000
500000
750000
1000000
duration_ms
target
0
1

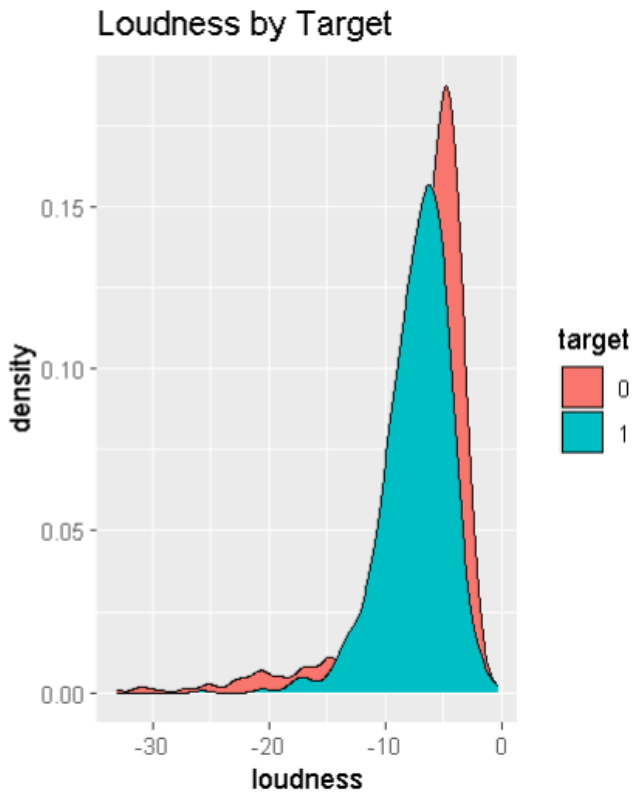

Energy by Target
density
2.0
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
energy
target
0
1

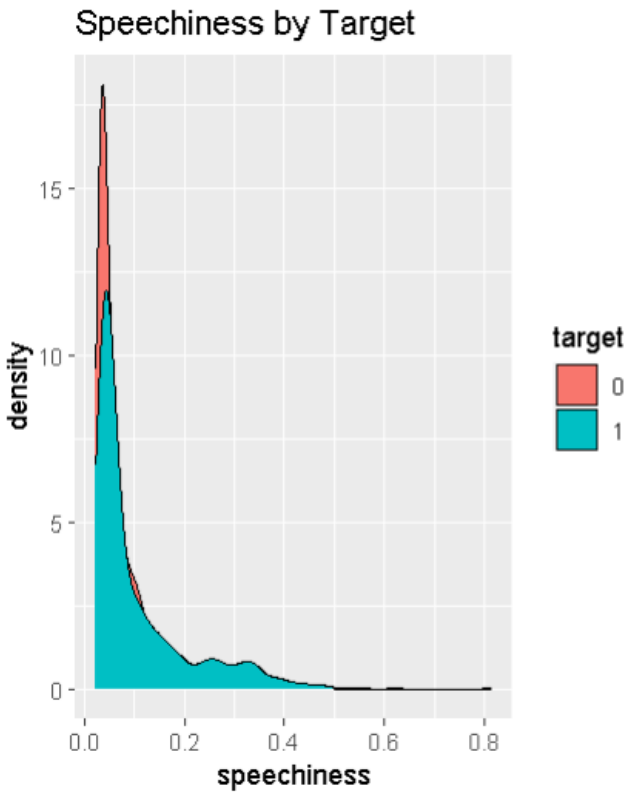

Instrumentalness by Target
density
500
400
300
200
100
0
0.00
0.25
0.50
0.75
1.00
instrumentalness
target
0
1

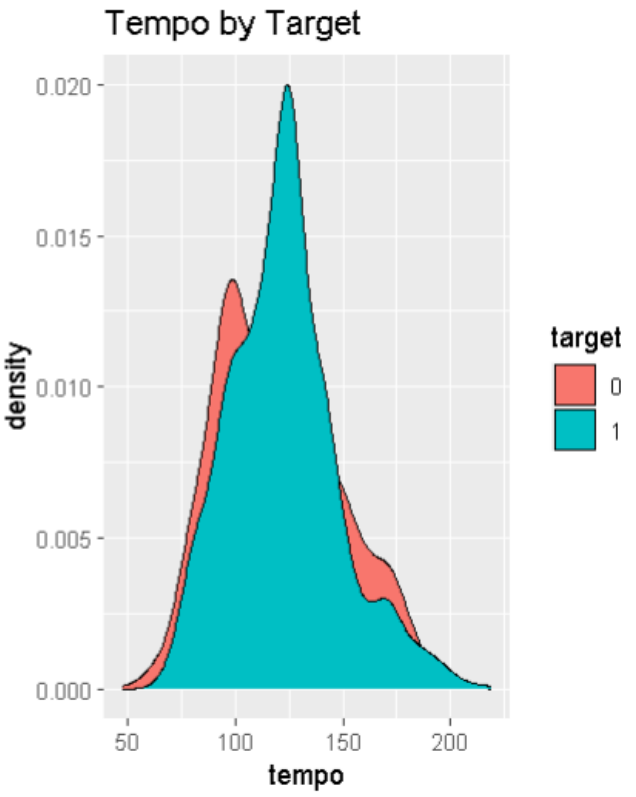
2.2. 3 task


Valence by Target
density
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
valence
target
0
1


Liveness by Target
density
liveness
target
0
1
0
2
4
6
0.00
0.25
0.50
0.75
1.00


Loudness by Target
density
loudness
target
0
1
0.00
0.05
0.10
0.15
-30
-20
-10
0


Speechiness by Target
density
speechiness
target
0
1
0
5
10
15
0.0
0.2
0.4
0.6
0.8


Tempo by Target
density
tempo
target
0
1
0.000
0.005
0.010
0.015
0.020
50
100
150
200

2.3. 4 task

2.4. 5 task

Mano manymu density ir duration yra geriausi faktoriai norint atskirti mėgstamą muziką nuo nepatinkančios, nes „Target“ kintamojo reikšmės labiausiai diferencijuoja šiose lentelėse.

2.5. 6 task

```
> # Display the results
> confusion_table  # Confusion table
      Predicted
Actual   0   1
     0 675 322
     1 355 665
> class_accuracies  # Class-specific accuracies
        0         1
0.6770311 0.6519608
> overall_accuracy  # Overall accuracy
[1] 0.664353
>
```

2.6. 7 task

```
Threshold: 0.3  Accuracy: 0.543877
Threshold: 0.4  Accuracy: 0.6197323
Threshold: 0.5  Accuracy: 0.664353
```

Geriau yra nuspėti, kuri muzika yra nemėgiama.

3. 4.1. Lab task: LDA and QDA classifiers

3.1. 1 task

3.2. 2 task

```
> print(class_accuracies)
         0          1          2          3          4          5          6          7
0.57627119 0.67741935 0.94573643 0.41032609 0.71394799 0.48148148 0.62230216 0.76221498
         8         10         11         12         13         14         15         16
0.23456790 0.37283237 0.48677249 0.30722892 0.28000000 0.56284153 0.45000000 0.29347826
        17         18         19         20         21         22         23         24
0.05528846 0.24495677 0.25105485 0.26229508 0.35658915 0.19617225 0.51479290 0.58641975
> 
```

3.3. 3 task

```
Accuracy with feature reduction factor 2 : 0.5050195

Accuracy with feature reduction factor 3 : 0.5528444
Accuracy with feature reduction factor 4 : 0.4875906
```

3.4. 4 and 5 task

```
        0         1         2         3         4         5         6         7         8
1.0000000 1.0000000 1.0000000 0.8401487 1.0000000 1.0000000 0.8601399 1.0000000 1.0000000
       10        11        12        13        14        15        16        17        18
0.9382716 1.0000000 1.0000000 0.5417722 0.9081081 1.0000000 0.1183261 0.2985075 0.4146789
       19        20        21        22        23        24
0.7754011 0.9939024 1.0000000 0.4322430 0.9702381 1.0000000
>
```

```
Accuracy with feature reduction factor 2 : 0.6483547
Accuracy with feature reduction factor 3 : 0.7023146
Accuracy with feature reduction factor 4 : 0.5599554
```