Portfolio of projects

Vadym Hladyuk

Data visualisation

- Application for visualisation of portfolio performance and characteristics
- Dash python - Data Visualization Interfaces
- Interactive web application
- Visualization of portfolio evolution, portfolio metrics and other portfolio information

Analýza dat v tenisu

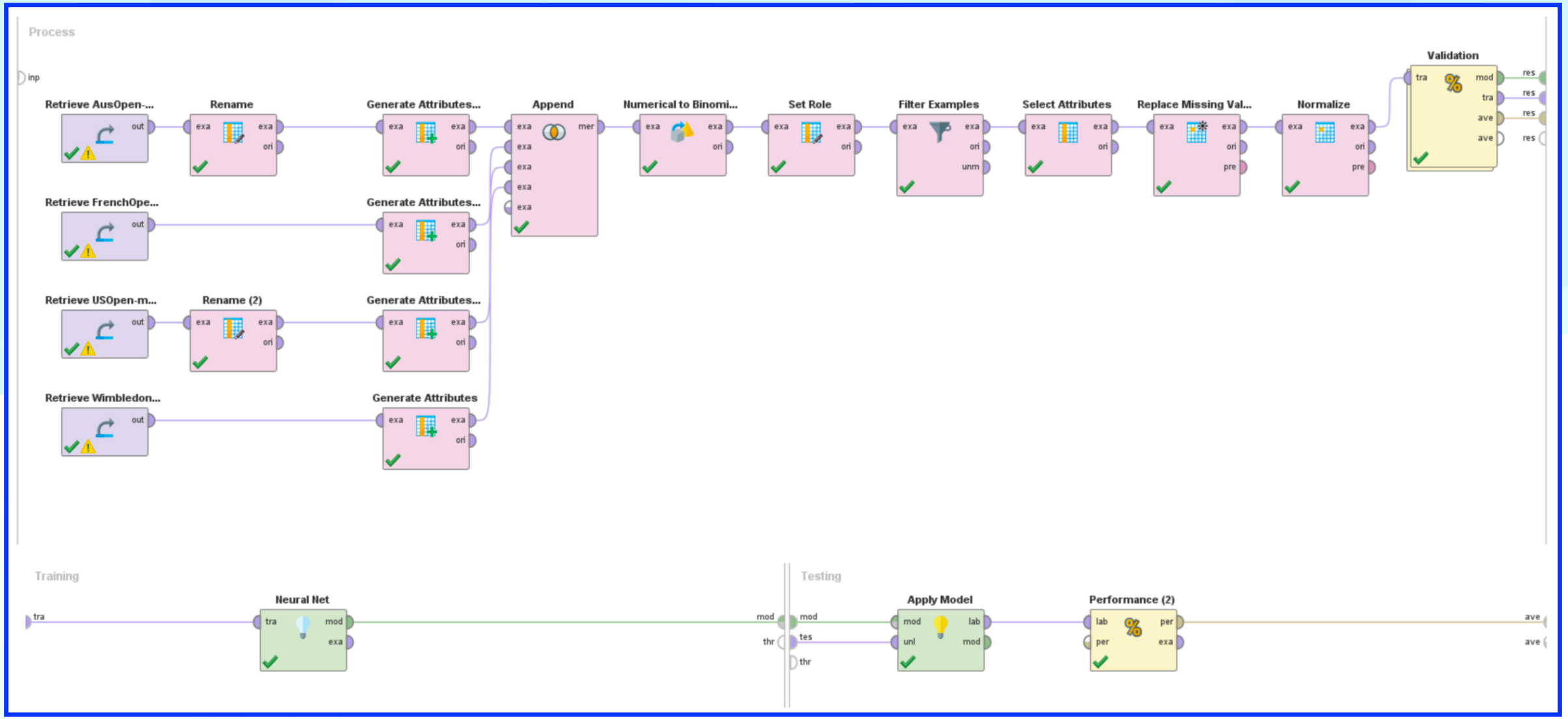
Predikce posledního setu a shlukování hráčů podle podání

- Školní projekt v programu Rapid Miner
- No-code platforma
- Dataset z tenisových turnajů (od údajů o podání až po počty vyhraných gamů)

Data analysis in tennis

Prediction result of the last set (only 5th set)

- Converting input to correct data types and filling in missing values
- Filtering data
- Predicting the result of the fifth set
- Two models: Neural Network and SVM


Process
inp
Retrieve AusOpen-...
Rename
Generate Attributes...
Append
Numerical to Binomi...
Set Role
Filter Examples
Select Attributes
Replace Missing Val...
Normalize
Validation
Retrieve FrenchOpe...
Generate Attributes...
Retrieve USOpen-m...
Rename (2)
Generate Attributes...
Retrieve Wimbledon...
Generate Attributes
Training
Neural Net
Testing
Apply Model
Performance (2)

Data analysis in tennis

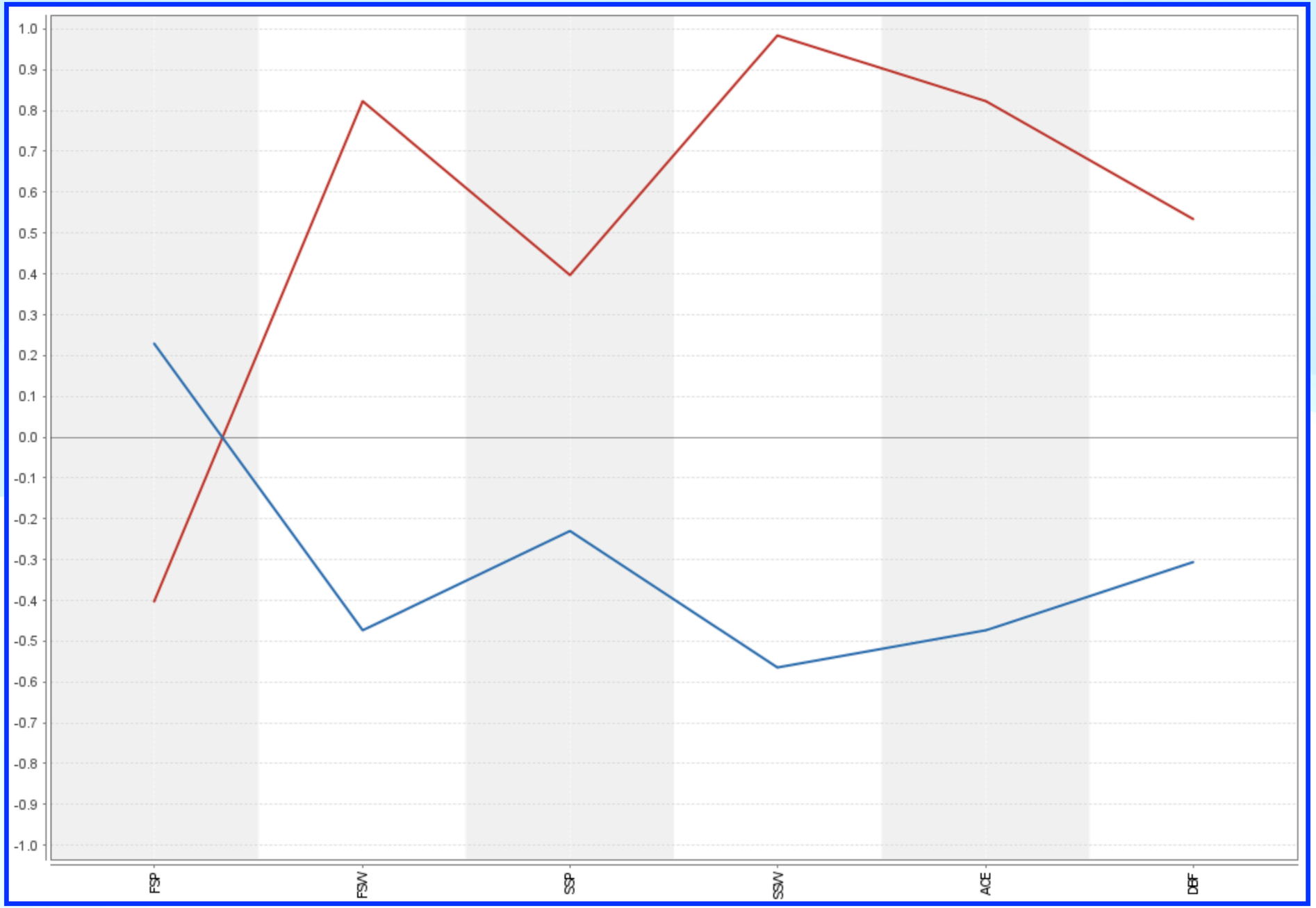
Prediction results

- Neural network: 65.38%
- SVM: 69.23%
- Limiting dataset size

Data analysis in tennis

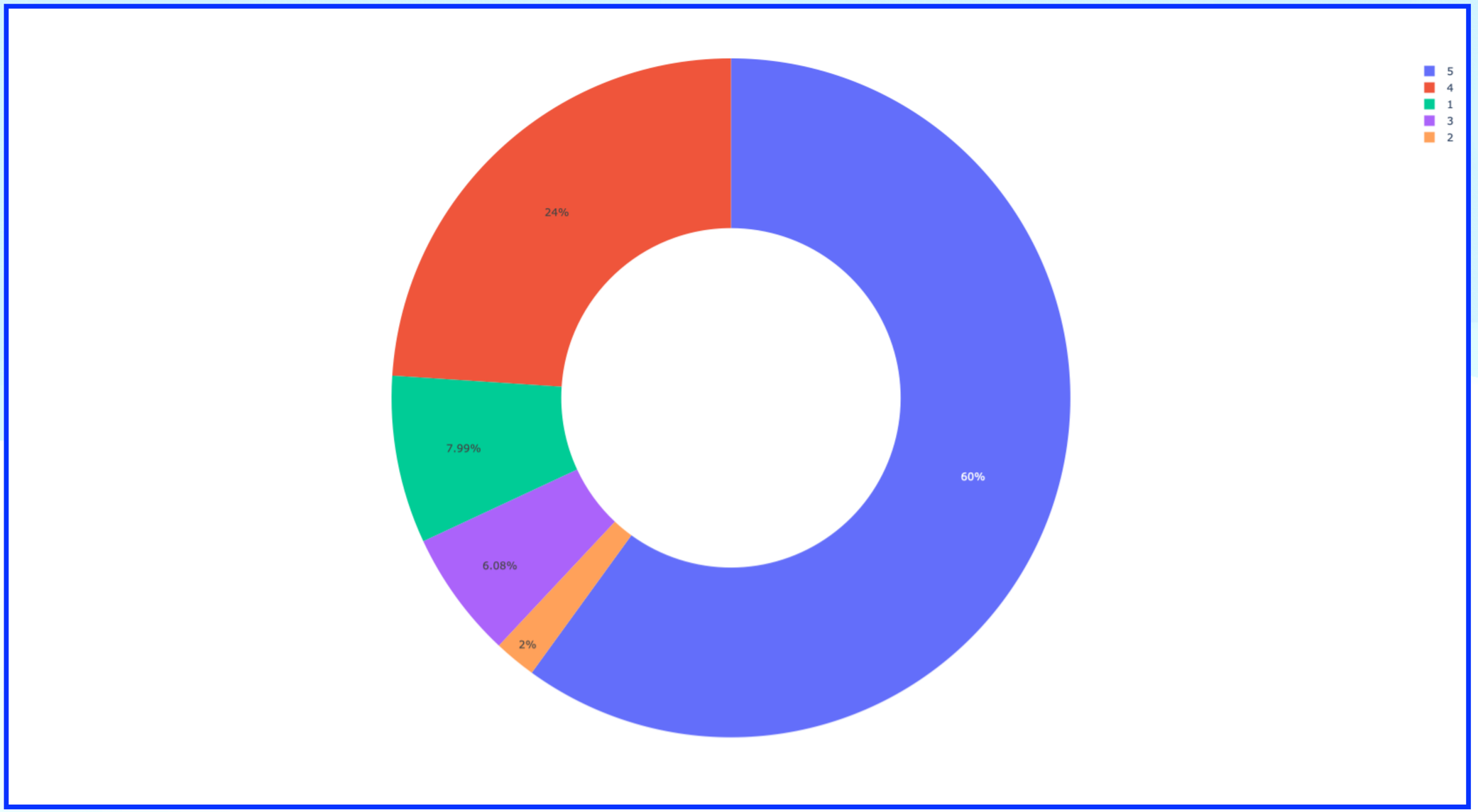
Clustering players by serve

- Data relating to submissions only
- Clustering method k-means
- 2 clusters


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
FSP
FSW
SSP
SSW
ACE
DBF

Sentiment analysis of product reviews

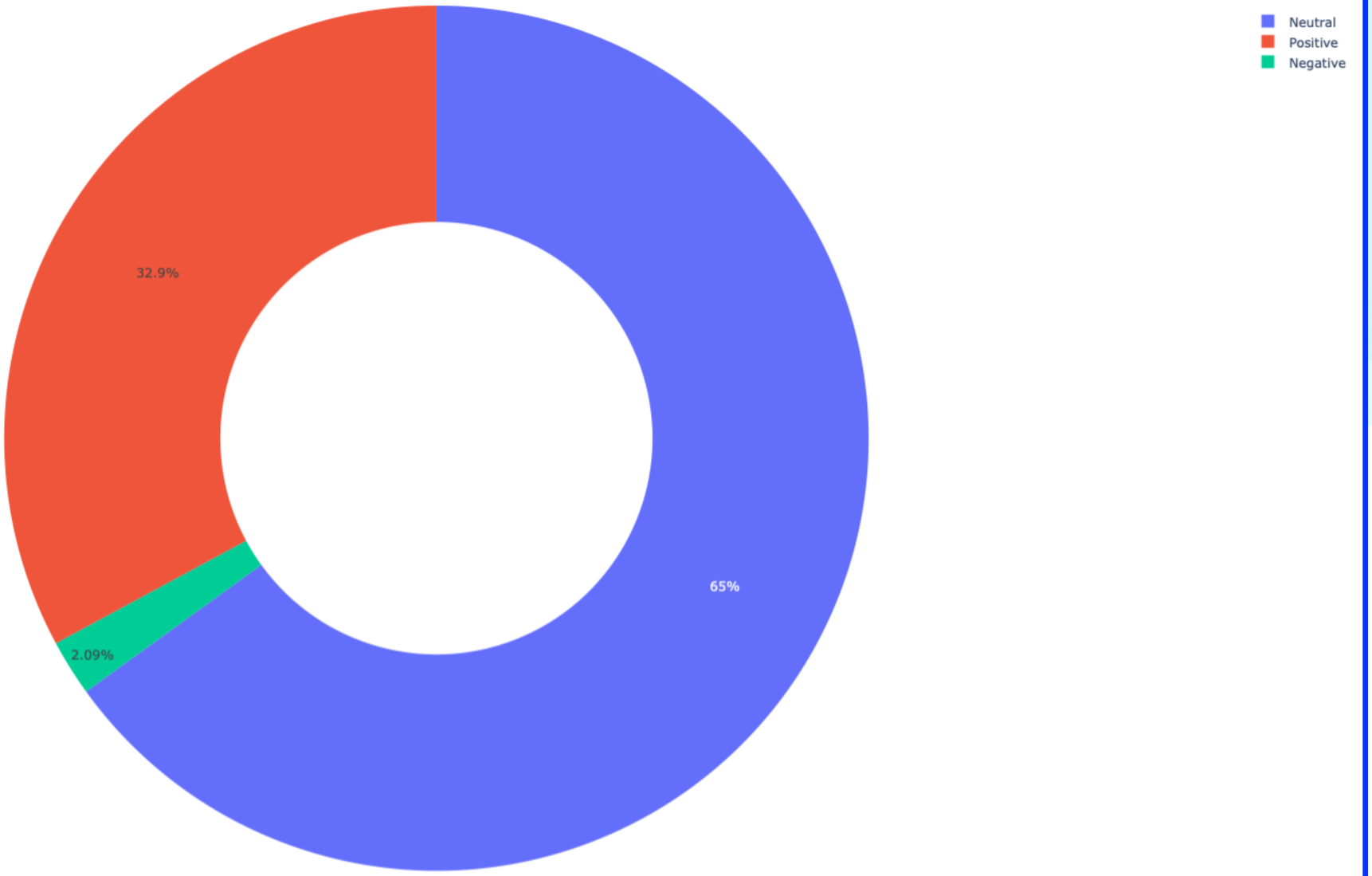
- Dataset (product, reviews and ratings)
- Review cleanup (stop words, stemming)
- Visualizing reviews using word cloud


5
4
1
3
2
24%
7.99%
6.08%
2%
60%

good
phone
best
nice
use
look
mobil
product
great
display
perform
work
awesom
camera
valu money
monitor
price
price rang
one
time
samsung
better
buy
go
love
amaz happi
set
even
good product
mi
super
fridg
perfect
feel
realm
day
cool
flipkart
design
batteri life
batteri backup
pictur qualiti
wash machin
refresh rate
nice phone
order
less
experi
devic
thank flipkart
colour
bright
wash
smooth
laptop
refriger
featur
charg
camera good
purchas
instal
batteri
dont
issu
much
fast
power
want
well
give
game
thing
front camera
good batteri
superb
qualiti
sound
thank
realli
overal
bad
enough
dont buy
hour
make
excel
screen
come
compar
sound machin
speaker
model
brand
first
decent
expect
low
got
ok
bought
lg
connect
recommend
display good
camera qualiti
qualiti good
price rang
tv

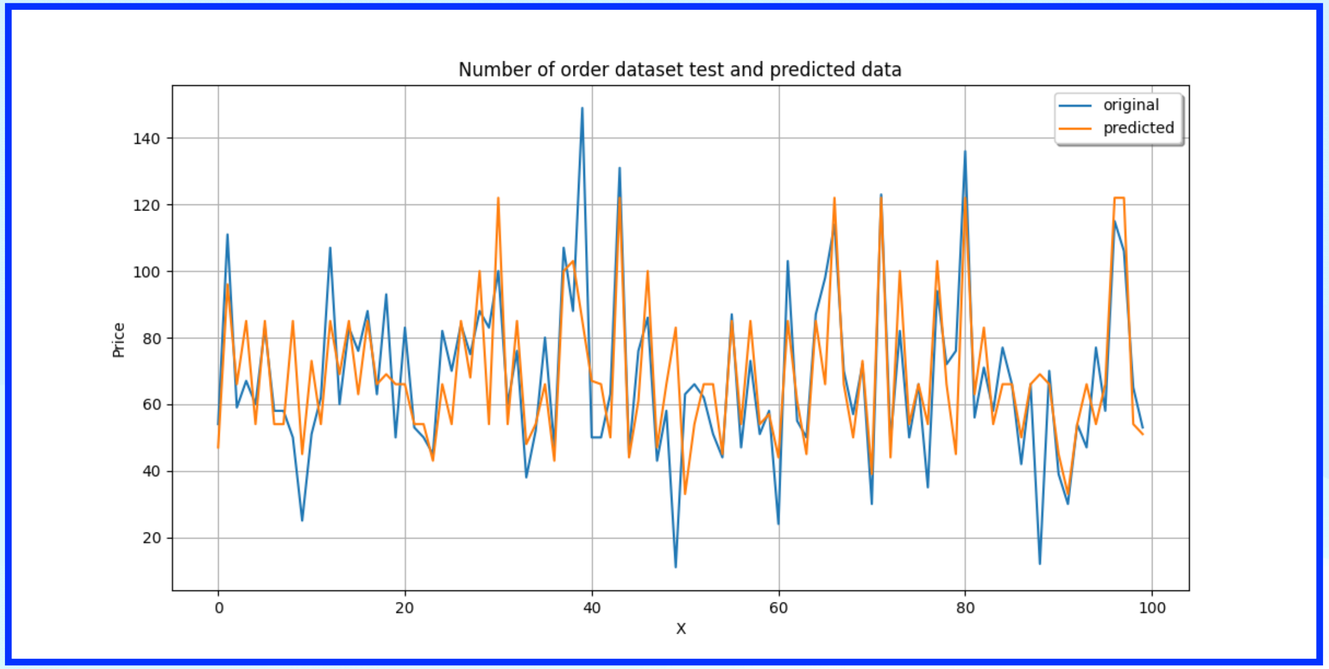
Sentiment analysis of product reviews

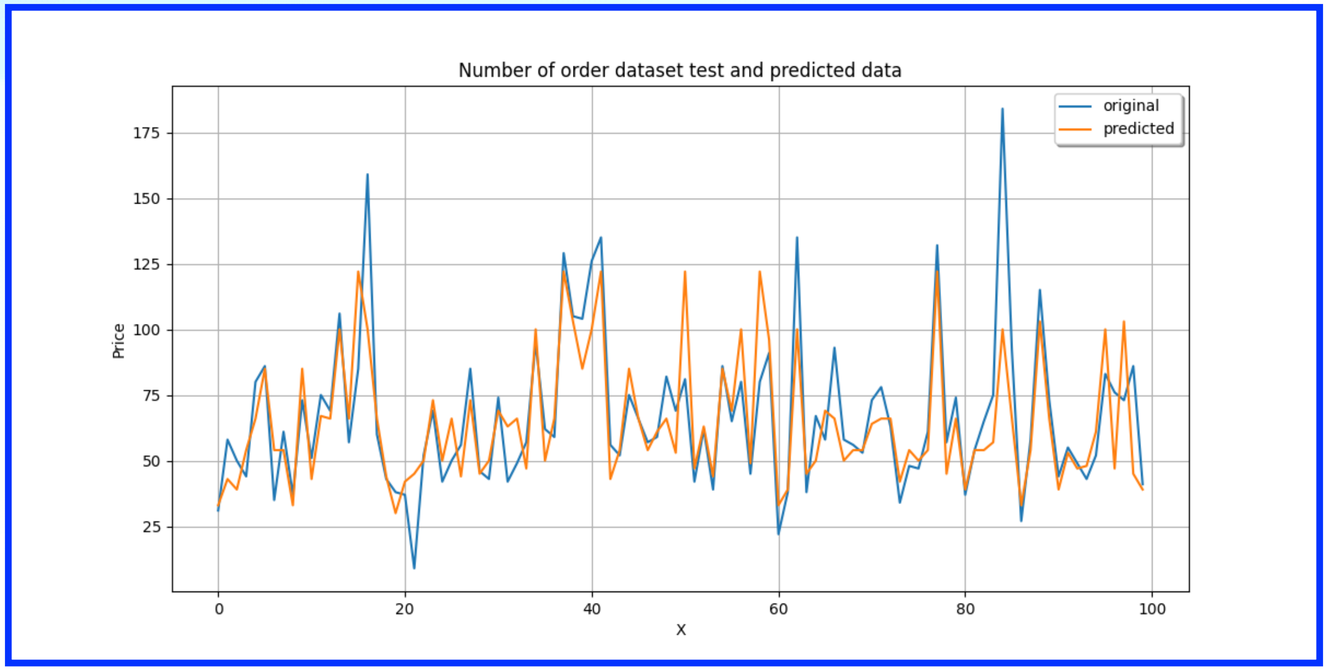
- Natural Language Toolkit
- Vader lexicon
- positive, neutral, negative and composite scores of each review


Neutral
Positive
Negative
32.9%
2.09%
65%

Prediction of the number of orders

- Dataset (Store_Type, Location_Type, Holiday, Discount and number of orders)
- Numeric values only
- To train the model, we use the light gradient boosting regression algorithm (LGBMR)
- On average, the number of orders varies by 13.33
- Root mean square deviation: 19.51


Number of order dataset test and predicted data
original
predicted
Price
X
0
20
40
60
80
100
20
40
60
80
100
120
140


Number of order dataset test and predicted data
original
predicted
Price
X
0
20
40
60
80
100
25
50
75
100
125
150
175

Click-through rate prediction

- Dataset (time on page, age, income group, daily internet usage, gender, click-through)
- Logistic regression
- Accuracy: 71.37%
- F1 score: 0.7074