

Lab Assignment

BI

Joud Abdullah Aljehani

2111644

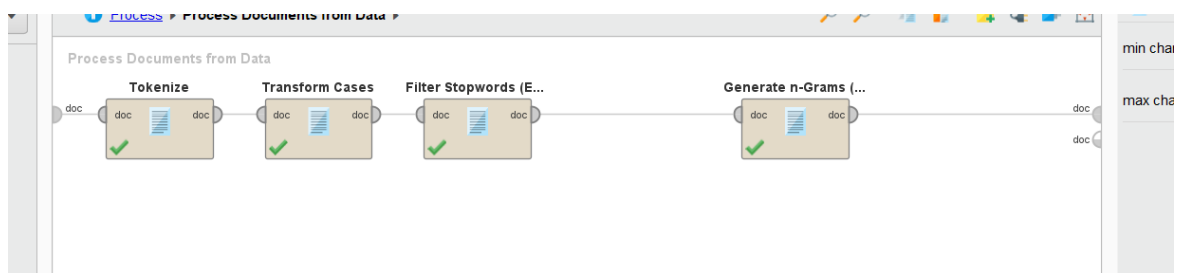
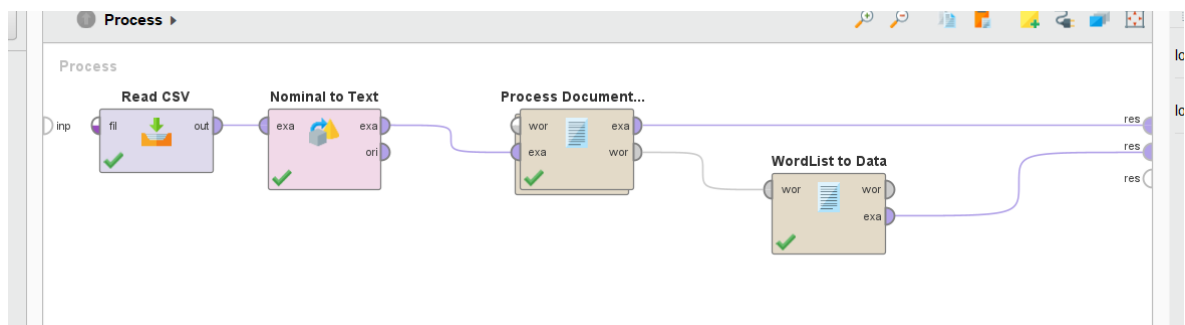
Part 1

A. List the word you think frequently appear in the dataset for positive sentiment, and negative

sentiment. List down FIVE words for positive sentiment, and FIVE words for negative

sentiment. [This is your own observation of the data. Do not run the occurrences model YET]

Example for positive sentiment	Example for negative sentiment
Love	hate
Great	horror
Family	worst
Life	bad
Happy	awful



B. Analyze the word frequency and check if the word you mention in your answer in question

(A) is what you see in this result. Write down how many occurrences of this word you see,

and how many times in appears in the data set (how many people are mentioning this

The word of positive sentiment	In reviews	Mention in total
Love	155	216
Great	241	339
Family	87	119
Life	170	258
Happy	37	40

The word of negative sentiment	In reviews	Mention in total
Hate	27	30
Horror	74	123
Bad	233	344
Awful	60	71
Worst	96	126

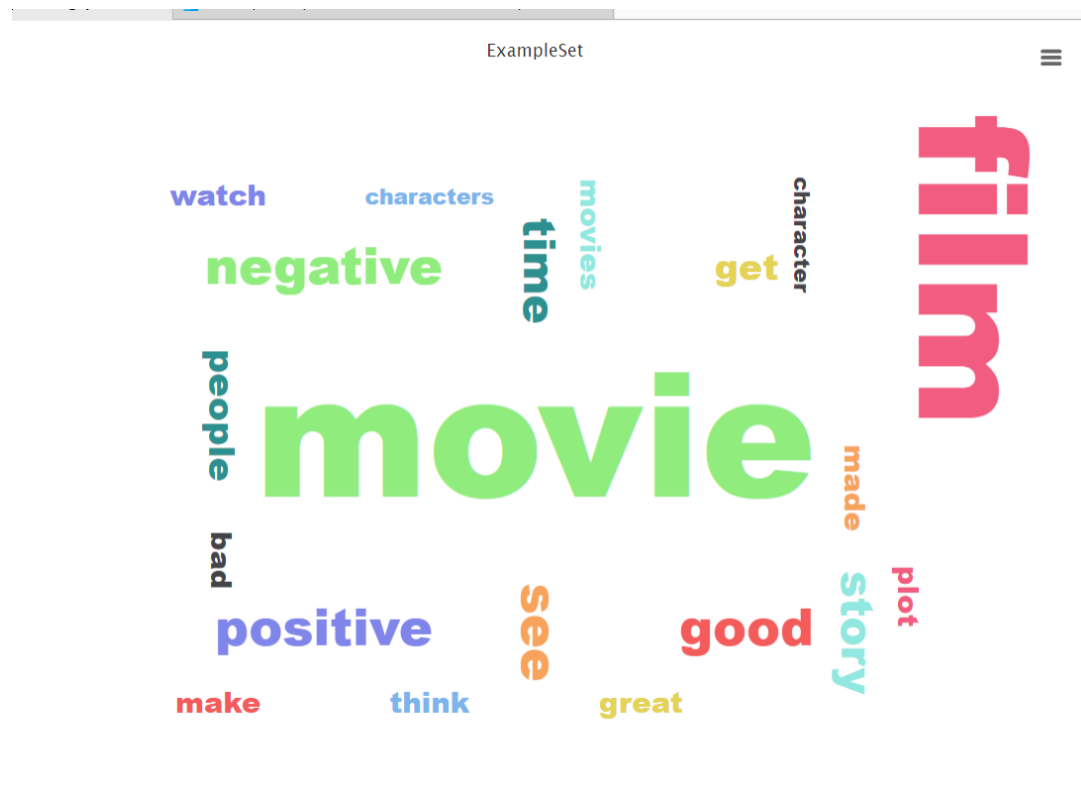
C. Screen shot your wordcloud. Does the wordcloud met your expectation of the word list?

Why yes, or, why not?



The most common words in this result are made up of letters and word segments that don't match the criteria of the word list I previously created.

D. Run your model again, and make a new wordcloud. Screen shot your new wordcloud.



E. Modify your model, and make it look like the following. Explain what does connecting the

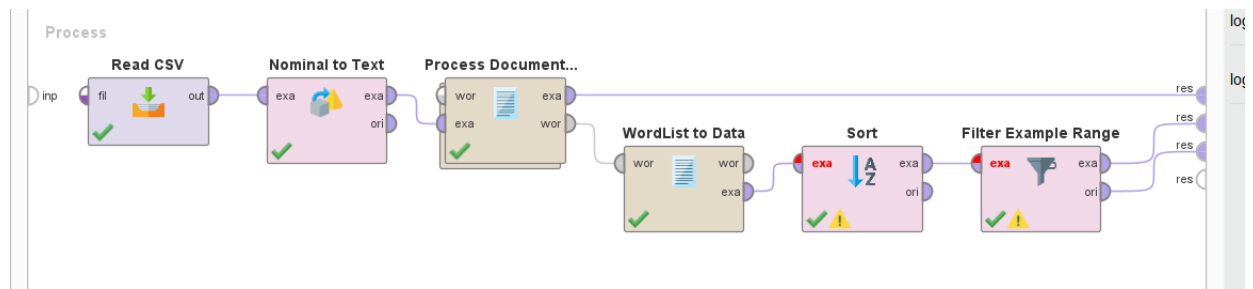
port 'ori' to 'res' do?

original data = ori

Through the ori port, our data that was provided as input is carried through without being altered.

This is often used to display the data in the Results Workspace or to reuse the same data in additional operators.

Result=Res Which give as the result of our modified data



Part 2

F. What is the accuracy of your model? Take the screen capture. Is this good? Suggest One

action you can do to improve the reading of your performance.

The performance is not poor, but it's also not very excellent. The performance of the kernel,

which affects the performance of the model itself, is significantly influenced by the

modifiable parameters. When we adjust the value C from 0 to 0.1, we will observe that,

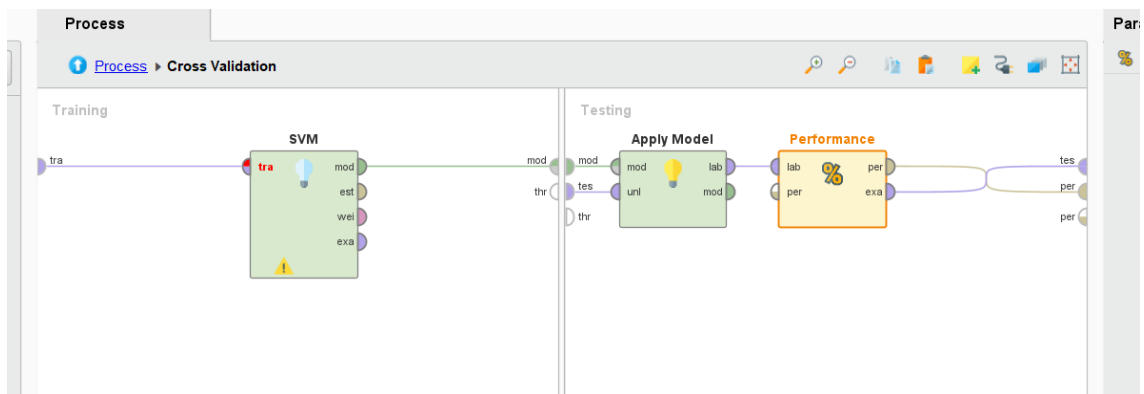
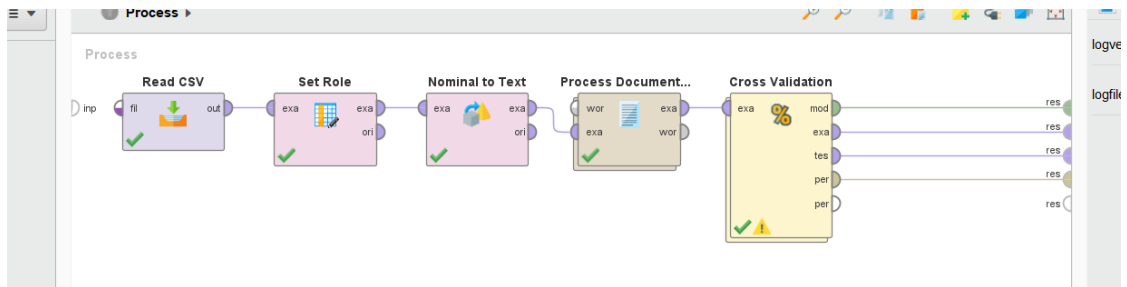
The model's accuracy increases to 76.28%. So, we can observe how making little

adjustments .changes in parameters might have a big impact on the final outcome.

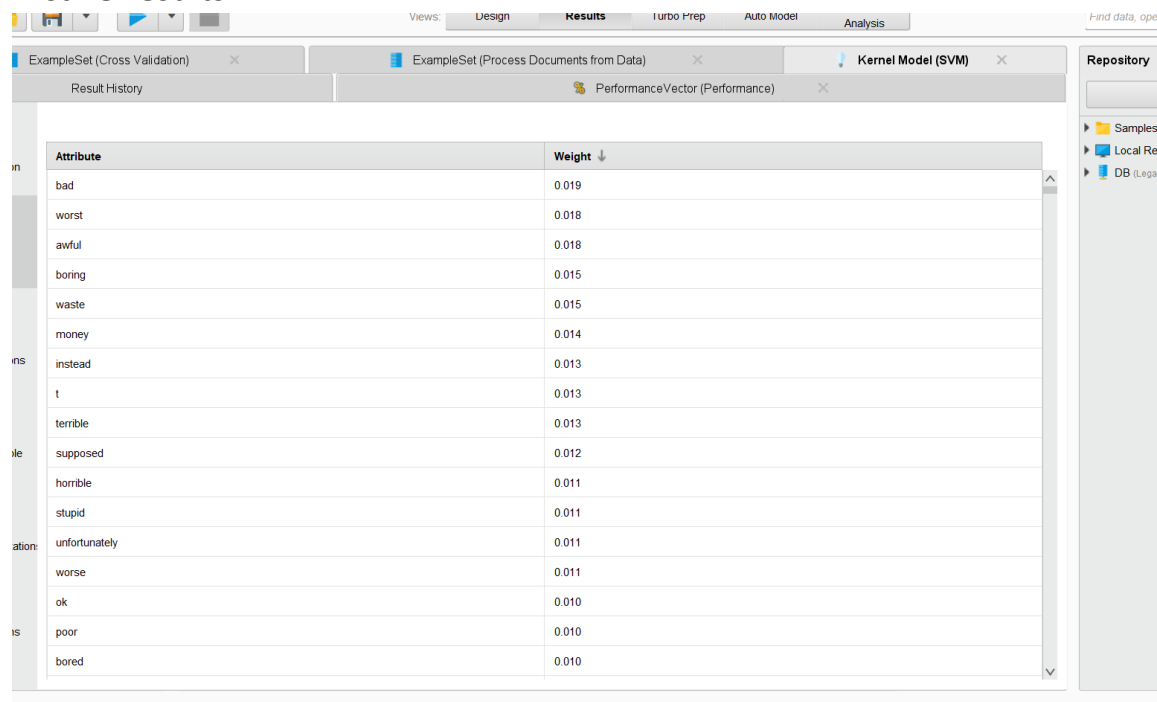
☒ Table View
 ☐ Plot View

accuracy: 75.87% +/- 2.98% (micro average: 75.87%)

	true positive	true negative	class precision
pred. positive	399	144	73.48%
pred. negative	92	343	78.85%
class recall	81.26%	70.43%	



G. (i) List down TEN words with negative sentiment. (ii) Take the screen shot of first 15 results.



Attribute	Weight ↓
bad	0.019
worst	0.018
awful	0.018
boring	0.015
waste	0.015
money	0.014
instead	0.013
t	0.013
terrible	0.013
supposed	0.012
horrible	0.011
stupid	0.011
unfortunately	0.011
worse	0.011
ok	0.010
poor	0.010
bored	0.010

1. bad
2. worst
3. awful
4. boring
5. waste
6. terrible
7. stupid
8. unfortunately
9. worse
10. Poor

H. (i) List down TEN words with positive sentiment. (ii) Take the screen shot of first 15 results.

1. great
2. excellent
3. wonderful
4. loved
5. brilliant
6. favorite
7. love
8. genius

9. entertaining

10.amazing

Attribute	Weight ↑
great	-0.015
excellent	-0.015
wonderful	-0.014
loved	-0.013
brilliant	-0.012
dvd	-0.012
favorite	-0.012
love	-0.012
young	-0.011
genius	-0.011
entertaining	-0.010
amazing	-0.010
definitely	-0.010
hilarious	-0.009
edge	-0.009
makes	-0.009
city	-0.009