ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
Π.Μ.Σ
«Υπολογιστικές Μέθοδοι και Εφαρμογές»

ΠΑΡΟΥΣΙΑΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

“Design and development of a system based on short questions for retrieving relevant documents that express opinion.”

Κοζάρη Γεωργία, Μαθηματικός
Α.Μ: mai18068
Επόπτρια: Επίκουρη καθηγήτρια Κολωνιάρη Γεωργία

Οκτώβριος, 2019

CONTENTS

1. Introduction
2. Dissertation’s Purpose
3. Methodological consideration
4. Experiments
5. Results
6. Conclusion

1.

Introduction

Opinion mining and the enquiry of information retrieval systems

Opinion mining : Given a set of evaluative text documents D that contain opinions (or sentiments) about an entity (e.g. item/topic/person/product or service), opinion mining aims to extract aspects (e.g. properties or attributes) of the entity that have been commented on in each document d ∈ D and to determine whether the comments are positive, negative or neutral **(Bakhatawar and Farouque, 2012)**.

Levels of Opinion Mining :

1. Document Level
2. Sentence Level
3. Aspect-Feature Level

Applications of opinion mining :

- Marketing
- Business
- Politics
- Shopping
- Entertainment

Opinion mining and the enquiry of information retrieval systems

OPINION MINING

?

Identify opinion components and extract useful information of them.

INFORMATION RETRIEVAL

Information Retrieval (IR) :

- The most interesting part of IR is the new challenges and the motivation of researchers to look for intelligent information retrieval systems.
- These systems search and/or filter information automatically based on some higher level of understanding.

2.

Dissertation’ s purpose

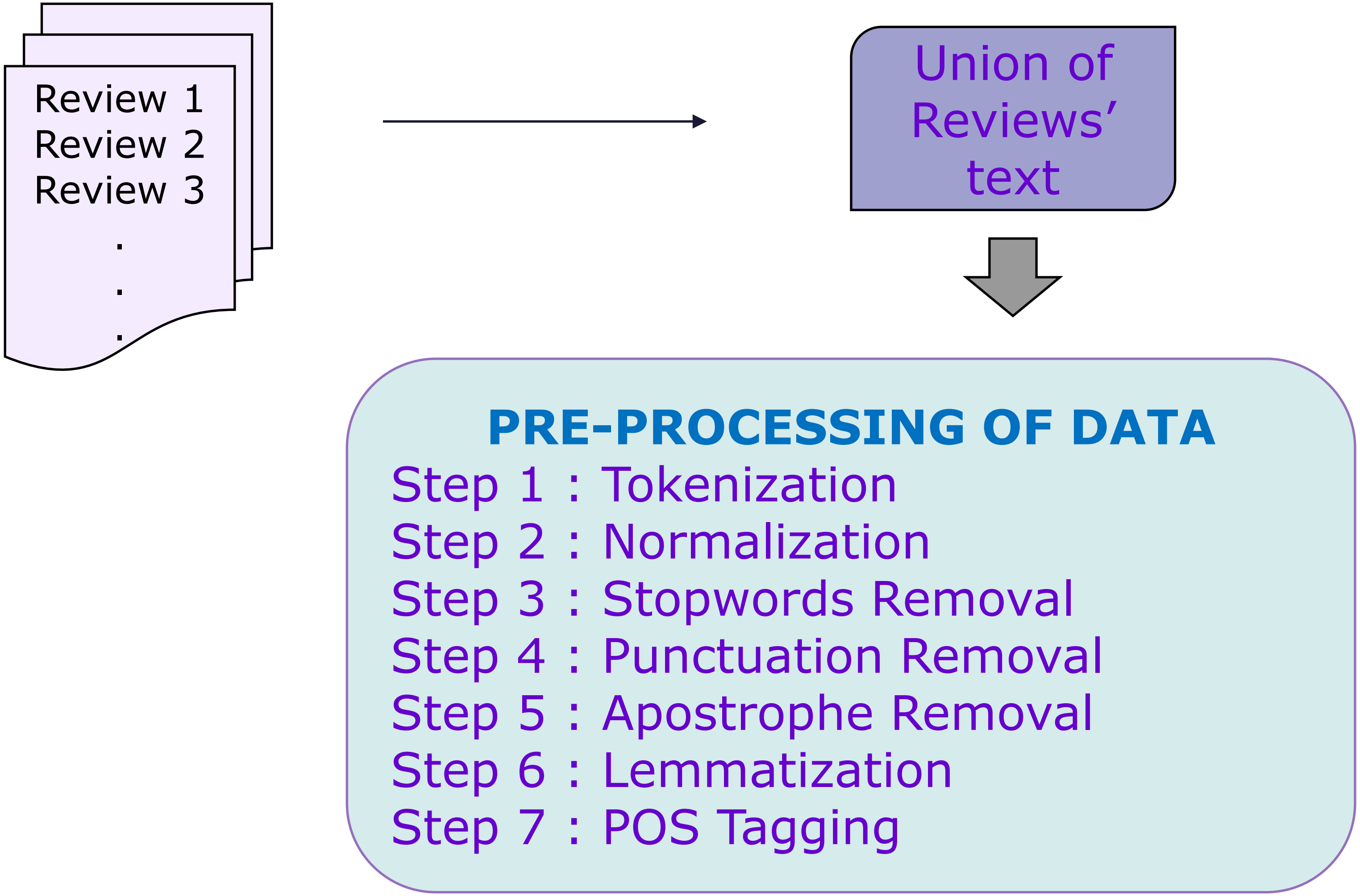
Dissertation’s purpose

- Design and development of an intelligent information retrieval system that will be based on short questions-keywords and its aim will be to retrieve relevant documents that express opinion.
- It will generate results, ranked by certain criteria (e.g. relevance) and corresponding to user's query.
- Expansion of the user's query in synonyms, hypernyms, and hyponyms by using thesauruses, while the calculation of term frequency-inverse document frequency score in order to find the most relevant documents were essential for the design and evaluation of the system.

3.

Methodological Considerations

3.1 Data and user's query Pre-processing



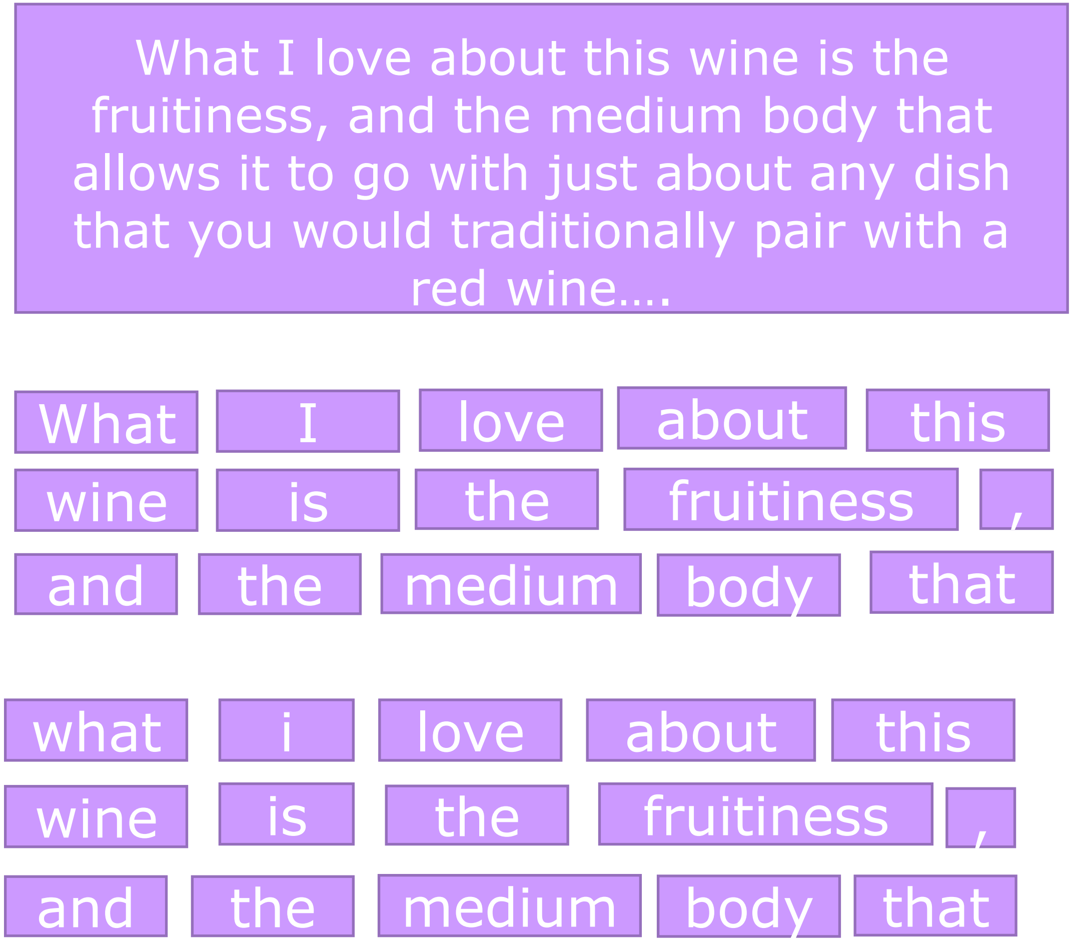
Review 1
Review 2
Review 3
.
.
.

Union of Reviews' text

PRE-PROCESSING OF DATA

Step 1 : Tokenization
Step 2 : Normalization
Step 3 : Stopwords Removal
Step 4 : Punctuation Removal
Step 5 : Apostrophe Removal
Step 6 : Lemmatization
Step 7 : POS Tagging

Data Pre-processing



EXAMPLE :

Union of reviews’ text →

What I love about this wine is the fruitiness, and the medium body that allows it to go with just about any dish that you would traditionally pair with a red wine….

STEP 1 : TOKENIZATION

What	I	love	about	this
wine	is	the	fruitiness	,
and	the	medium	body	that

STEP 2 : NORMALIZATION

what	i	love	about	this
wine	is	the	fruitiness	,
and	the	medium	body	that

Data Pre-processing

STEP 3 : STOPWORDS REMOVAL

love | wine | fruitiness | ,

medium | body

STEP 4 : PUNCTUATION REMOVAL

love | wine | fruitiness

medium | body

STEP 5 : APOSTROPHE REMOVAL

don’t → dont

STEP 6 : LEMMATIZATION

dishes → dish

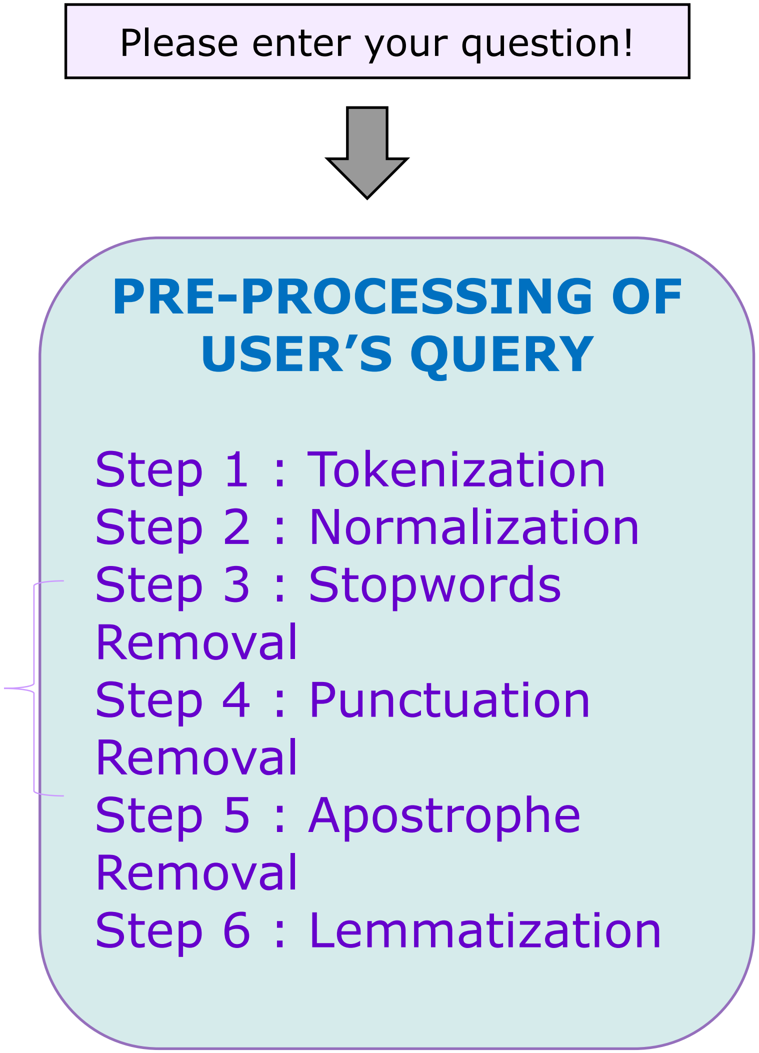
STEP 7 : POS TAGGING

love VB | wine NN | fruitiness JJ

traditionally RB | red JJ

Pre-processing of user’s query

USER’S QUERY



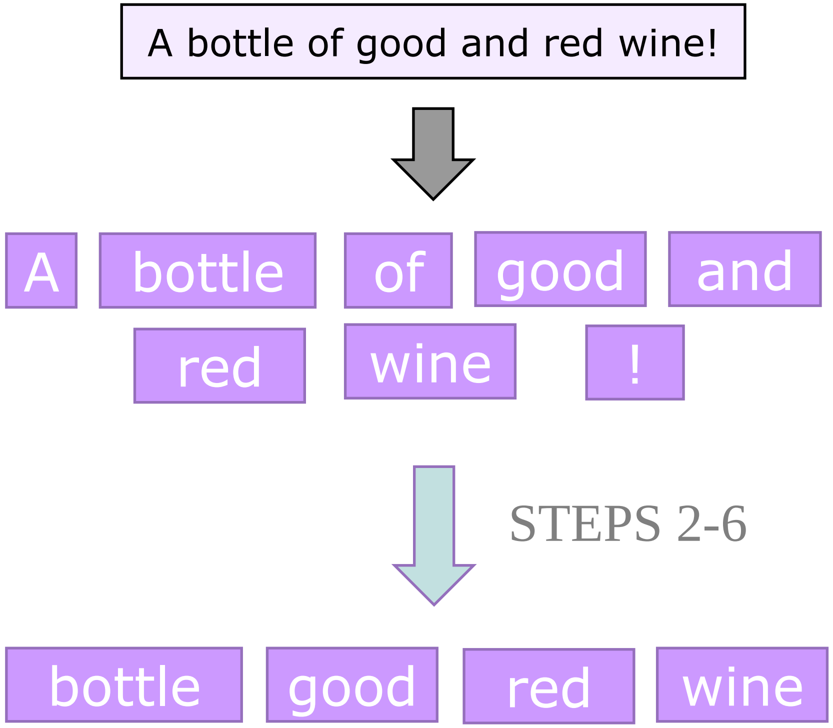
Please enter your question!

PRE-PROCESSING OF USER’S QUERY

Step 1 : Tokenization
Step 2 : Normalization
Step 3 : Stopwords Removal
Step 4 : Punctuation Removal
Step 5 : Apostrophe Removal
Step 6 : Lemmatization

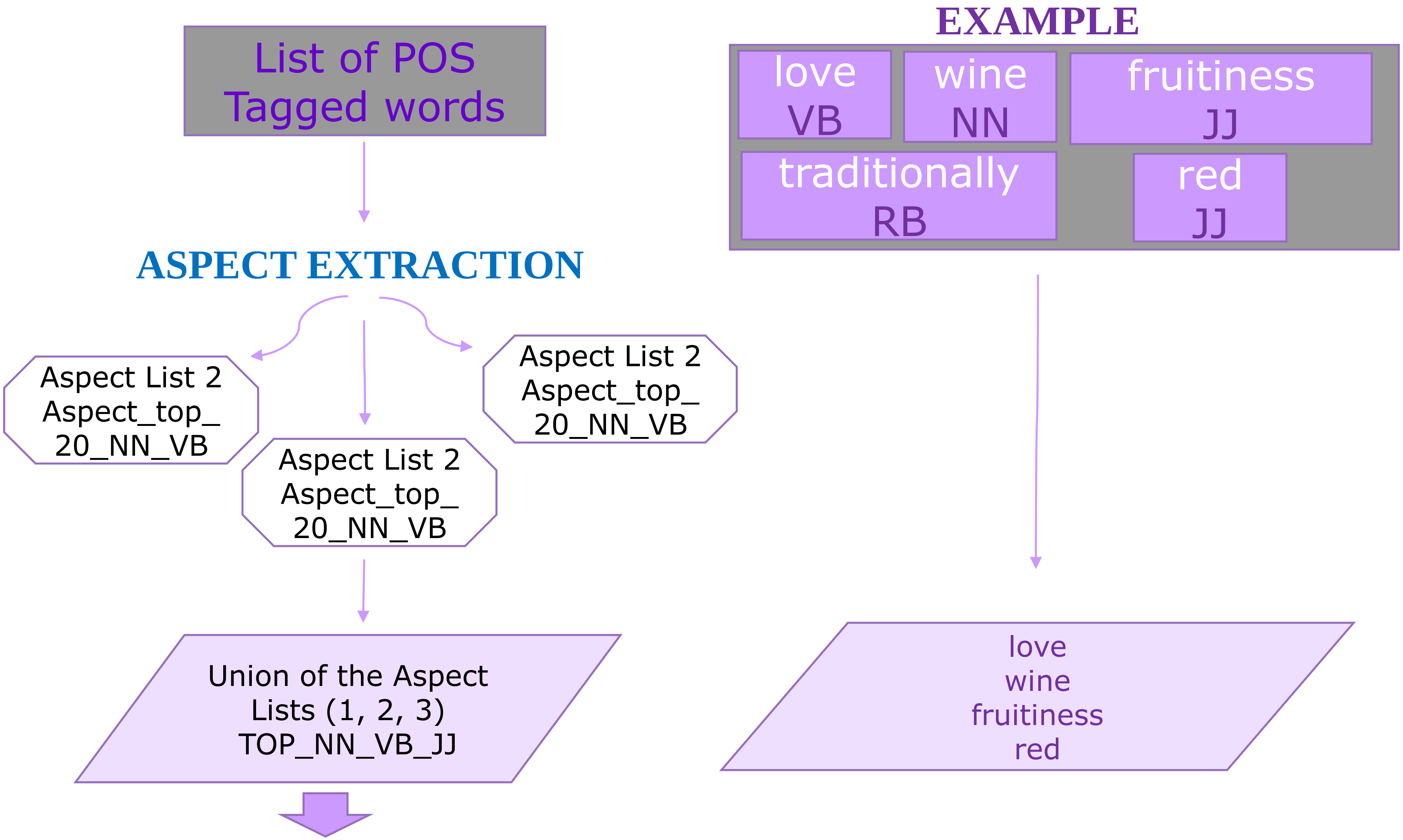
“ Noise “ Removal

EXAMPLE



A bottle of good and red wine!

3.2 Extraction of aspect words and expansion of them to thesauri



List of POS Tagged words

ASPECT EXTRACTION

Aspect List 2 Aspect_top_ 20_NN_VB

Aspect List 2 Aspect_top_ 20_NN_VB

Aspect List 2 Aspect_top_ 20_NN_VB

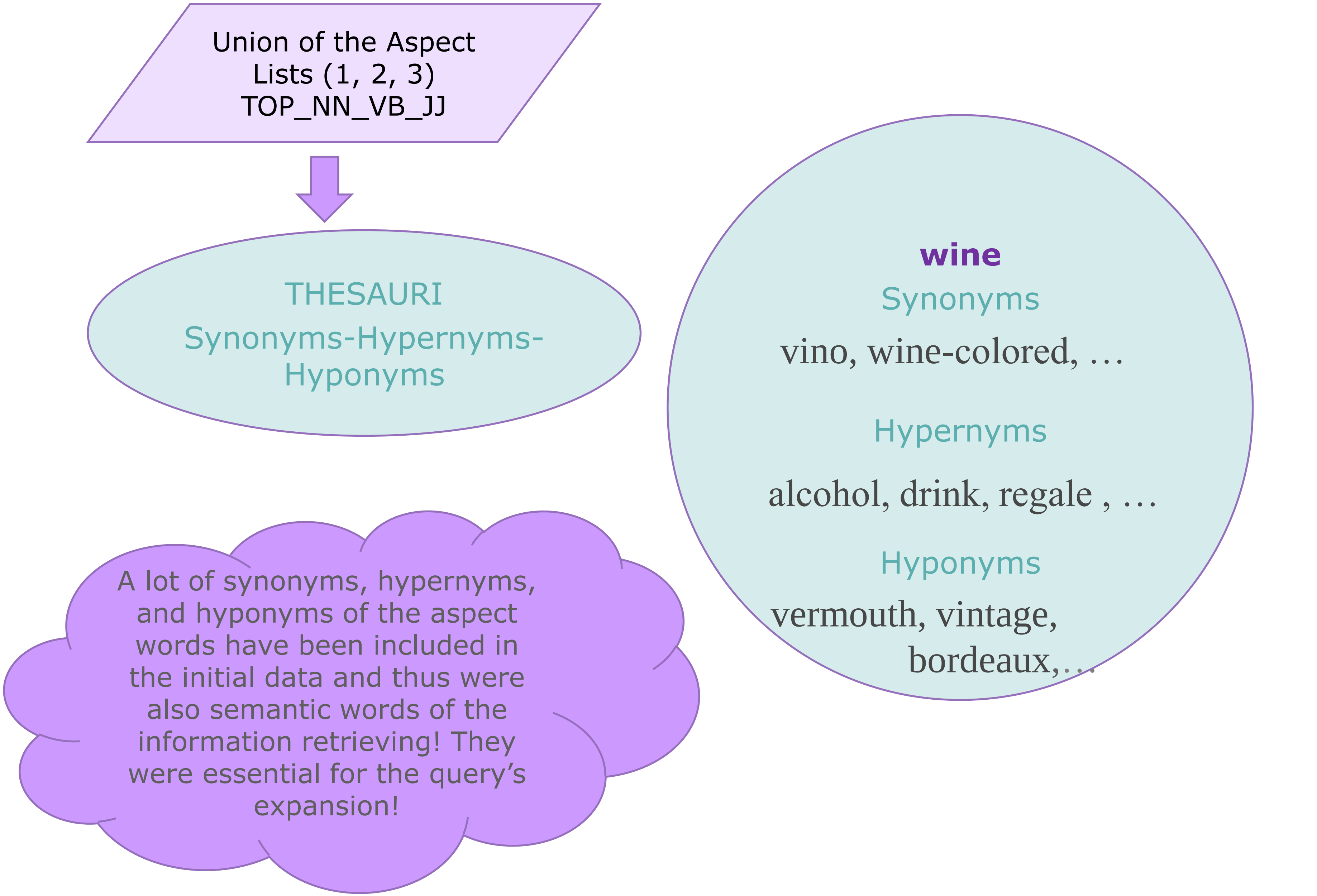
Union of the Aspect Lists (1, 2, 3) TOP_NN_VB_JJ

EXAMPLE

love VB	wine NN	fruitiness JJ
traditionally RB		red JJ

love
wine
fruitiness
red

Expansion of aspect words to thesauri



Union of the Aspect Lists (1, 2, 3) TOP_NN_VB_JJ

THESAURI

Synonyms-Hypernyms-Hyponyms

wine

Synonyms

vino, wine-colored, …

Hypernyms

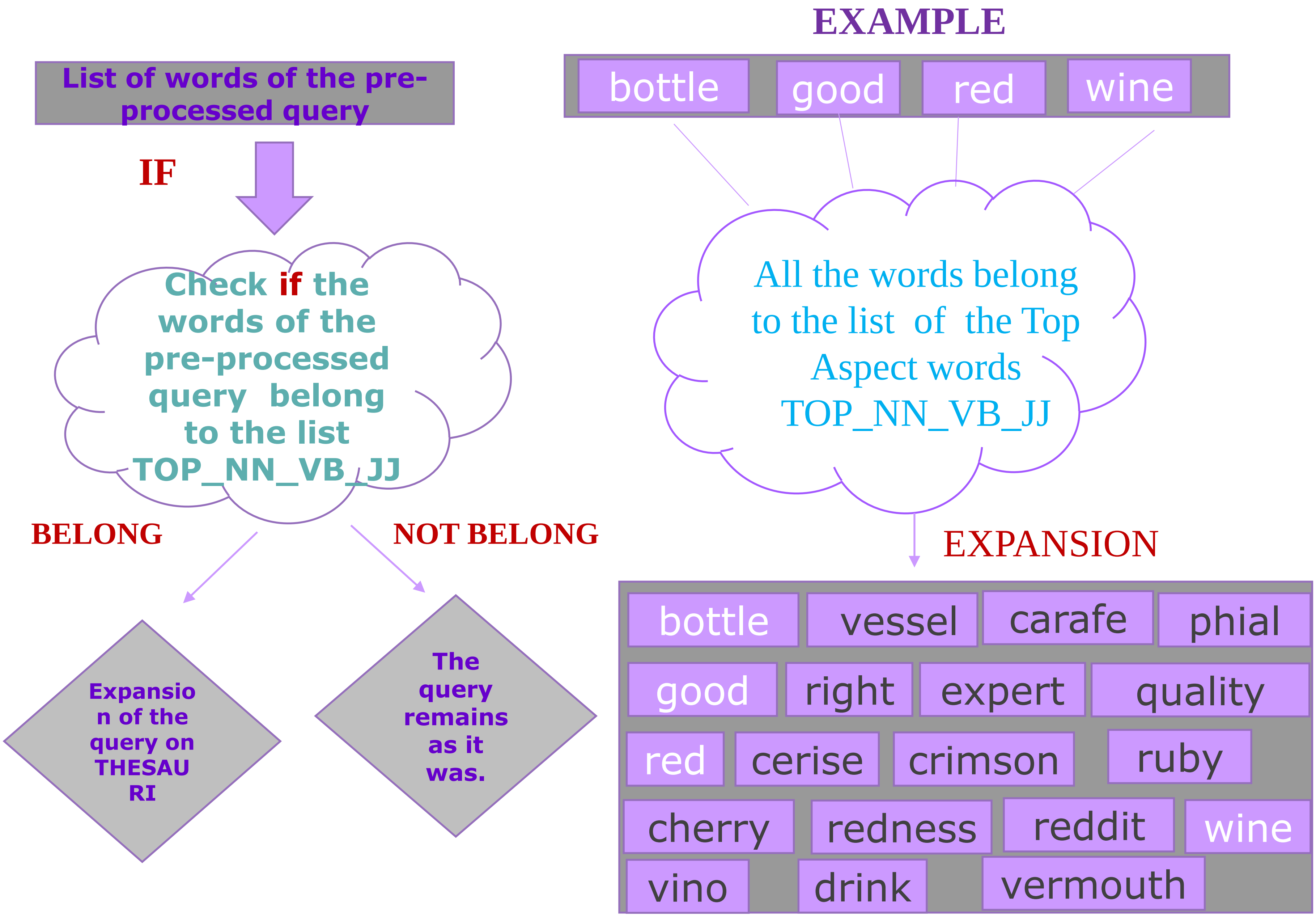
alcohol, drink, regale , …

Hyponyms

vermouth, vintage, bordeaux,…

A lot of synonyms, hypernyms, and hyponyms of the aspect words have been included in the initial data and thus were also semantic words of the information retrieving! They were essential for the query's expansion!

3.3 Query's expansion based on Thesauri



List of words of the pre-processed query

IF

Check **if** the words of the pre-processed query belong to the list TOP_NN_VB_JJ

BELONG

Expansion of the query on THESAURI

NOT BELONG

The query remains as it was.

EXAMPLE

bottle good red wine

All the words belong to the list of the Top Aspect words TOP_NN_VB_JJ

EXPANSION

bottle vessel carafe phial

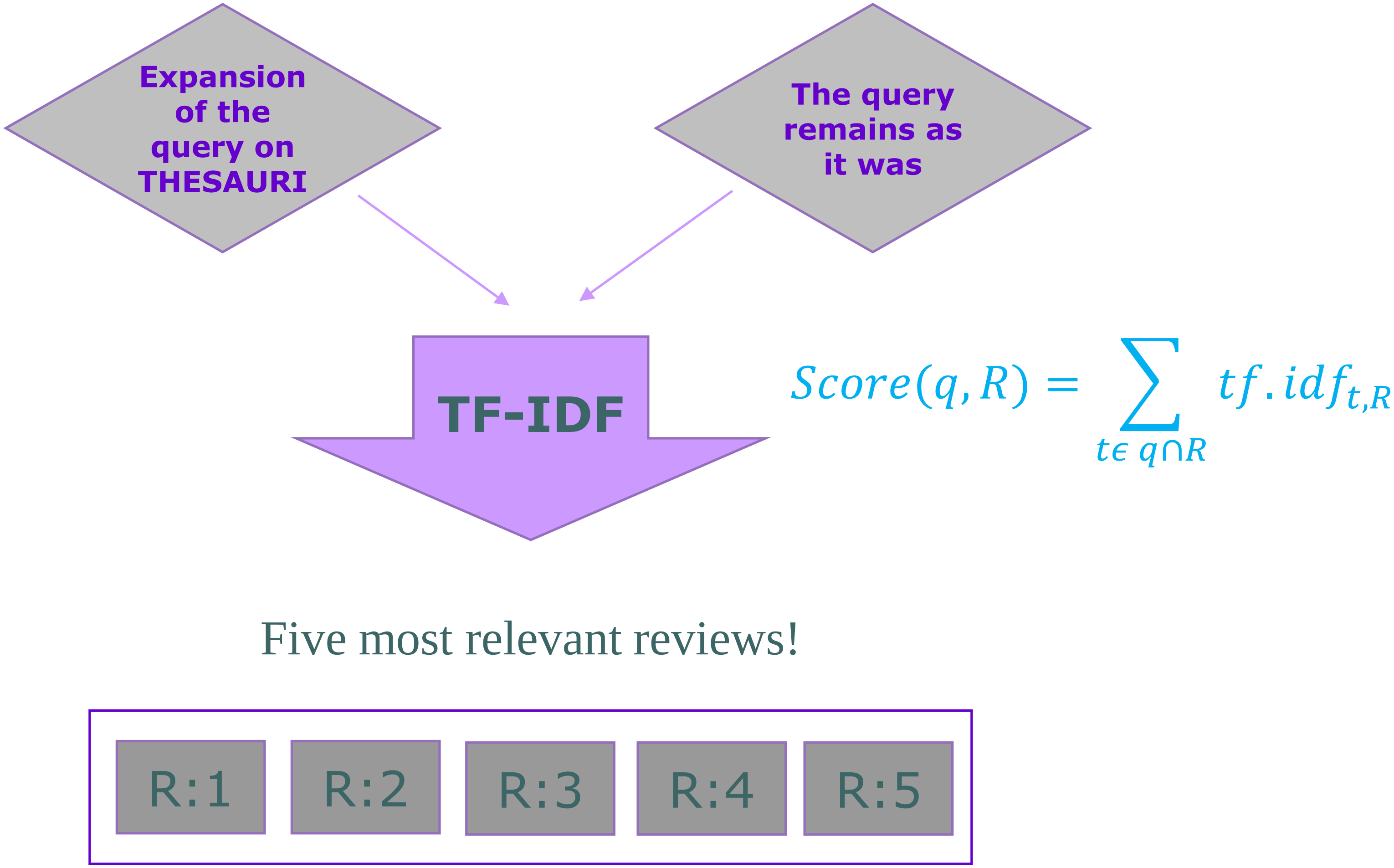
good right expert quality

red cerise crimson ruby

cherry redness reddit wine

vino drink vermouth

3.4 Connection between user’s query and the most relevant documents



Expansion of the query on THESAURI

The query remains as it was

TF-IDF

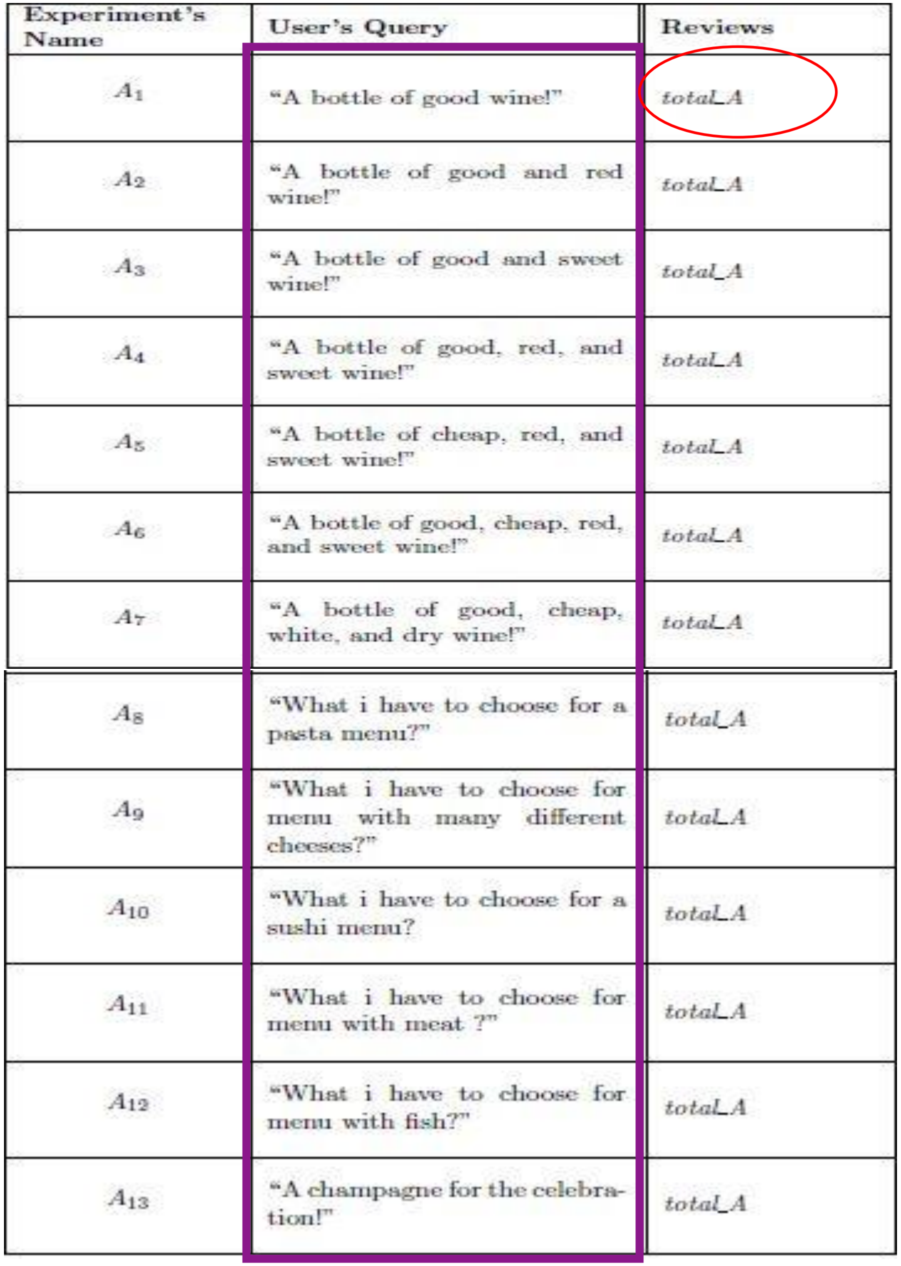
$$Score(q,R) = \sum_{t\epsilon\ q\cap R} tf.idf_{t,R}$$

Five most relevant reviews!

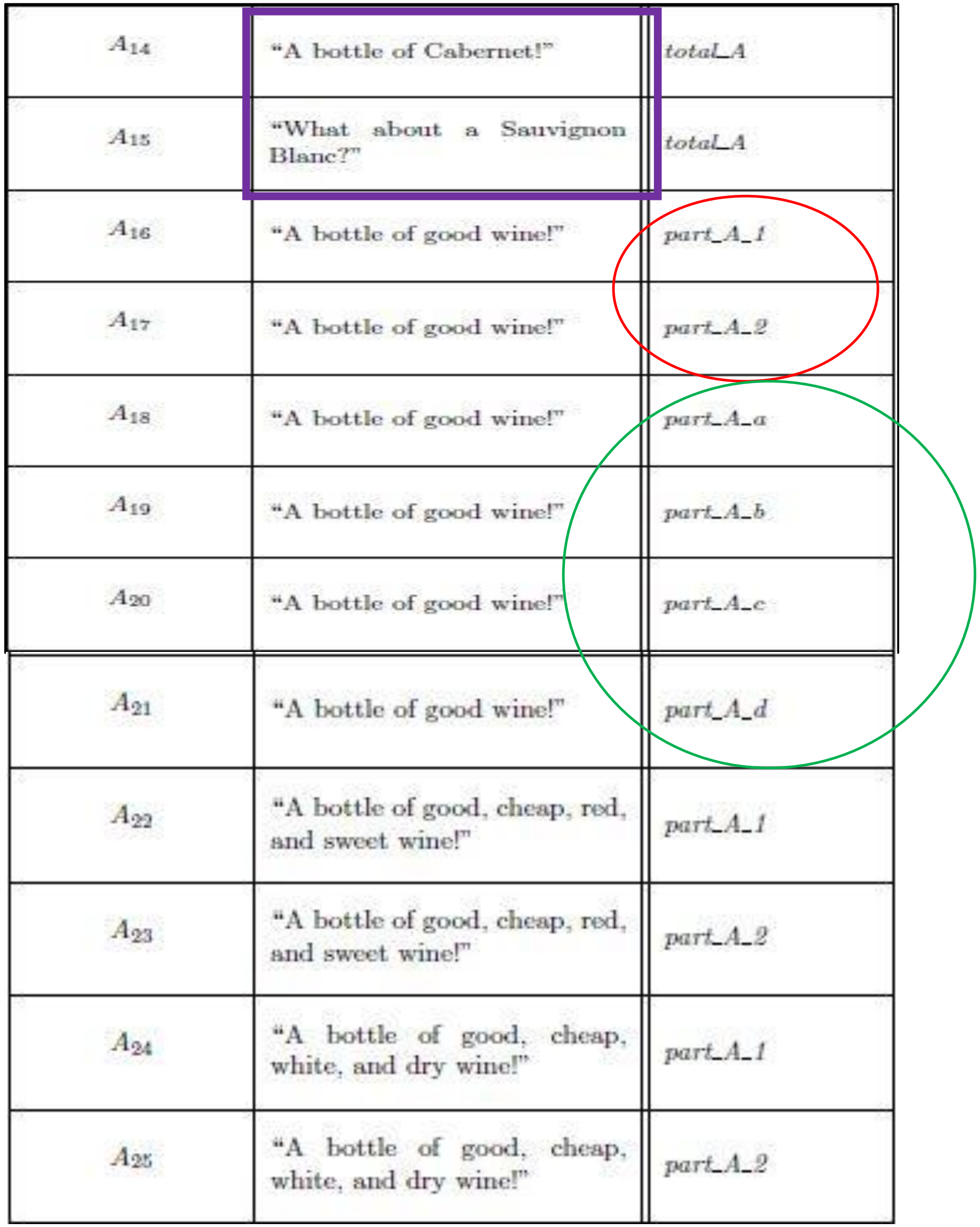
R:1 R:2 R:3 R:4 R:5

4.

Experiments



Experiment’s Name	User’s Query	Reviews
A_1	“A bottle of good wine!”	$total\_A$
A_2	“A bottle of good and red wine!”	$total\_A$
A_3	“A bottle of good and sweet wine!”	$total\_A$
A_4	“A bottle of good, red, and sweet wine!”	$total\_A$
A_5	“A bottle of cheap, red, and sweet wine!”	$total\_A$
A_6	“A bottle of good, cheap, red, and sweet wine!”	$total\_A$
A_7	“A bottle of good, cheap, white, and dry wine!”	$total\_A$
A_8	“What i have to choose for a pasta menu?”	$total\_A$
A_9	“What i have to choose for menu with many different cheeses?”	$total\_A$
A_{10}	“What i have to choose for a sushi menu?	$total\_A$
A_{11}	“What i have to choose for menu with meat ?”	$total\_A$
A_{12}	“What i have to choose for menu with fish?”	$total\_A$
A_{13}	“A champagne for the celebration!”	$total\_A$



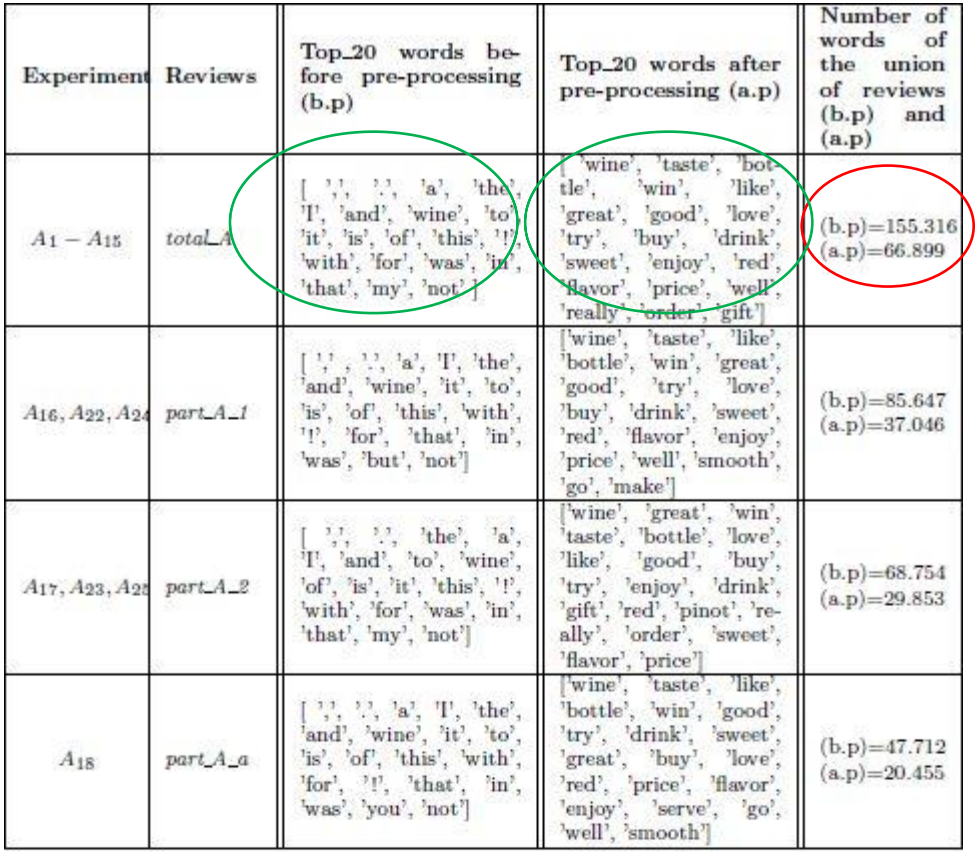
A_{14}	"A bottle of Cabernet!"	*total_A*
A_{15}	"What about a Sauvignon Blanc?"	*total_A*
A_{16}	"A bottle of good wine!"	*part_A_1*
A_{17}	"A bottle of good wine!"	*part_A_2*
A_{18}	"A bottle of good wine!"	*part_A_a*
A_{19}	"A bottle of good wine!"	*part_A_b*
A_{20}	"A bottle of good wine!"	*part_A_c*
A_{21}	"A bottle of good wine!"	*part_A_d*
A_{22}	"A bottle of good, cheap, red, and sweet wine!"	*part_A_1*
A_{23}	"A bottle of good, cheap, red, and sweet wine!"	*part_A_2*
A_{24}	"A bottle of good, cheap, white, and dry wine!"	*part_A_1*
A_{25}	"A bottle of good, cheap, white, and dry wine!"	*part_A_2*

Table 1 : Experiments of the dataset of wine reviews.

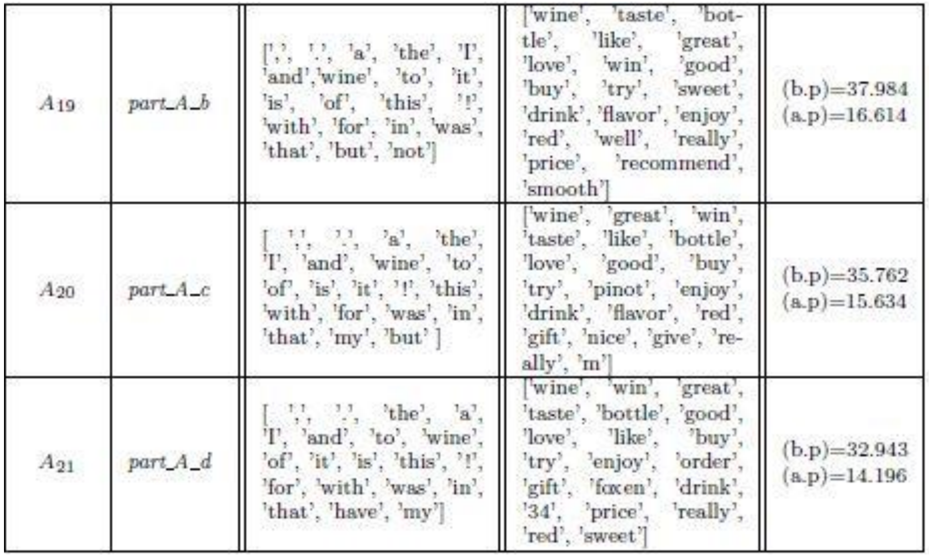
5.

Results

5.1 Results of data pre-processing



Experiment	Reviews	Top_20 words before pre-processing (b.p)	Top_20 words after pre-processing (a.p)	Number of words of the union of reviews (b.p) and (a.p)
$A_1 - A_{15}$	*total_A*	[’,’, ’.’, ’a’, ’the’, ’I’, ’and’, ’wine’, ’to’ ’it’, ’is’, ’of’, ’this’, ’!’, ’with’, ’for’, ’was’, ’in’, ’that’, ’my’, ’not’]	[’wine’, ’taste’, ’bottle’, ’win’, ’like’, ’great’, ’good’, ’love’, ’try’, ’buy’, ’drink’, ’sweet’, ’enjoy’, ’red’, ’flavor’, ’price’, ’well’, ’really’, ’order’, ’gift’]	(b.p)=155.316 (a.p)=66.899
A_{16}, A_{22}, A_{24}	*part_A_1*	[’,’ , ’.’, ’a’, ’I’, ’the’, ’and’, ’wine’, ’it’, ’to’, ’is’, ’of’, ’this’, ’with’, ’!’, ’for’, ’that’, ’in’, ’was’, ’but’, ’not’]	[’wine’, ’taste’, ’like’, ’bottle’, ’win’, ’great’, ’good’, ’try’, ’love’, ’buy’, ’drink’, ’sweet’, ’red’, ’flavor’, ’enjoy’, ’price’, ’well’, ’smooth’, ’go’, ’make’]	(b.p)=85.647 (a.p)=37.046
A_{17}, A_{23}, A_{25}	*part_A_2*	[’,’, ’.’, ’the’, ’a’, ’I’, ’and’, ’to’, ’wine’, ’of’, ’is’, ’it’, ’this’, ’!’, ’with’, ’for’, ’was’, ’in’, ’that’, ’my’, ’not’]	[’wine’, ’great’, ’win’, ’taste’, ’bottle’, ’love’, ’like’, ’good’, ’buy’, ’try’, ’enjoy’, ’drink’, ’gift’, ’red’, ’pinot’, ’really’, ’order’, ’sweet’, ’flavor’, ’price’]	(b.p)=68.754 (a.p)=29.853
A_{18}	*part_A_a*	[’,’, ’.’, ’a’, ’I’, ’the’, ’and’, ’wine’, ’it’, ’to’, ’is’, ’of’, ’this’, ’with’, ’for’, ’!’, ’that’, ’in’, ’was’, ’you’, ’not’]	[’wine’, ’taste’, ’like’, ’bottle’, ’win’, ’good’, ’try’, ’drink’, ’sweet’, ’great’, ’buy’, ’love’, ’red’, ’price’, ’flavor’, ’enjoy’, ’serve’, ’go’, ’well’, ’smooth’]	(b.p)=47.712 (a.p)=20.455



A_{19}	*part_A_b*	[',', '.', 'a', 'the', 'I', 'and','wine', 'to', 'it', 'is', 'of', 'this', '!', 'with', 'for', 'in', 'was', 'that', 'but', 'not']	['wine', 'taste', 'bottle', 'like', 'great', 'love', 'win', 'good', 'buy', 'try', 'sweet', 'drink', 'flavor', 'enjoy', 'red', 'well', 'really', 'price', 'recommend', 'smooth']	(b.p)=37.984 (a.p)=16.614
A_{20}	*part_A_c*	[',', '.', 'a', 'the', 'I', 'and', 'wine', 'to', 'of', 'is', 'it', '!', 'this', 'with', 'for', 'was', 'in', 'that', 'my', 'but']	['wine', 'great', 'win', 'taste', 'like', 'bottle', 'love', 'good', 'buy', 'try', 'pinot', 'enjoy', 'drink', 'flavor', 'red', 'gift', 'nice', 'give', 'really', 'm']	(b.p)=35.762 (a.p)=15.634
A_{21}	*part_A_d*	[',', '.', 'the', 'a', 'I', 'and', 'to', 'wine', 'of', 'it', 'is', 'this', '!', 'for', 'with', 'was', 'in', 'that', 'have', 'my']	['wine', 'win', 'great', 'taste', 'bottle', 'good', 'love', 'like', 'buy', 'try', 'enjoy', 'order', 'gift', 'foxen', 'drink', '34', 'price', 'really', 'red', 'sweet']	(b.p)=32.943 (a.p)=14.196

Table 2 : Results of the pre-processing of wine reviews.

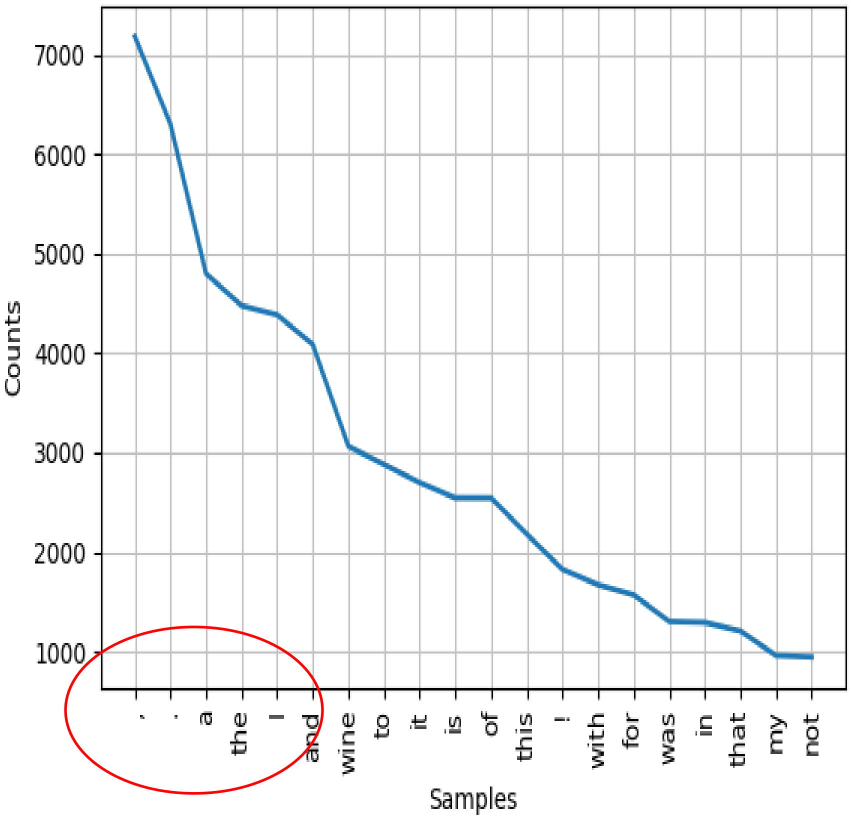


Figure 1 : Twenty most frequent words of the first fifteen experiments of wine reviews before preprocessing.

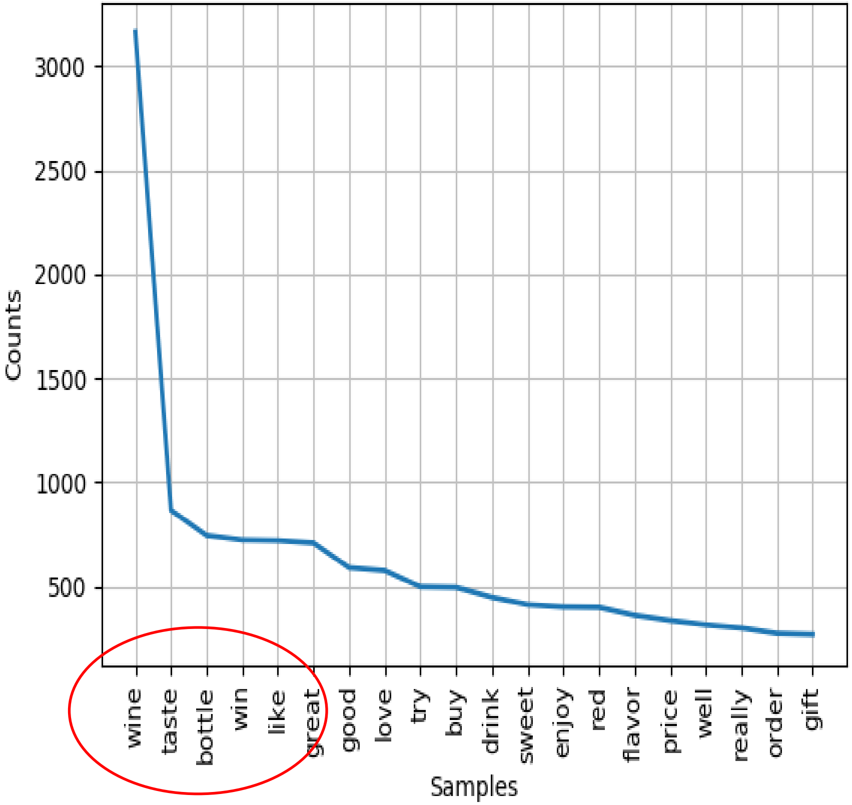
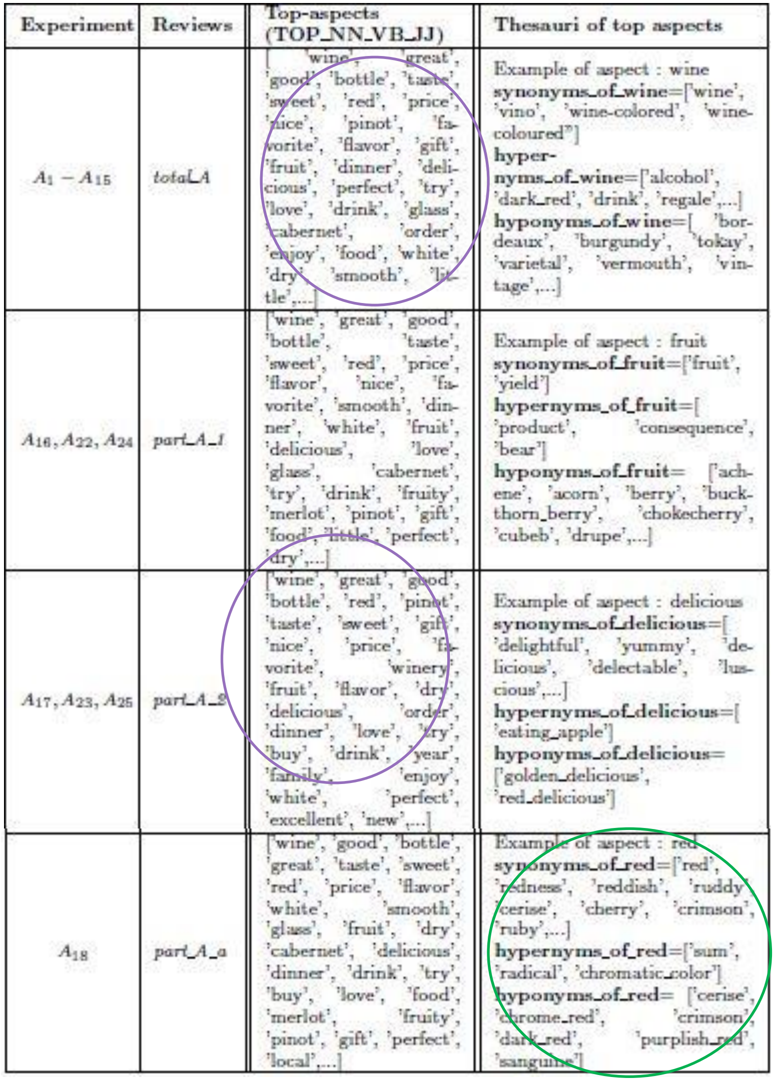
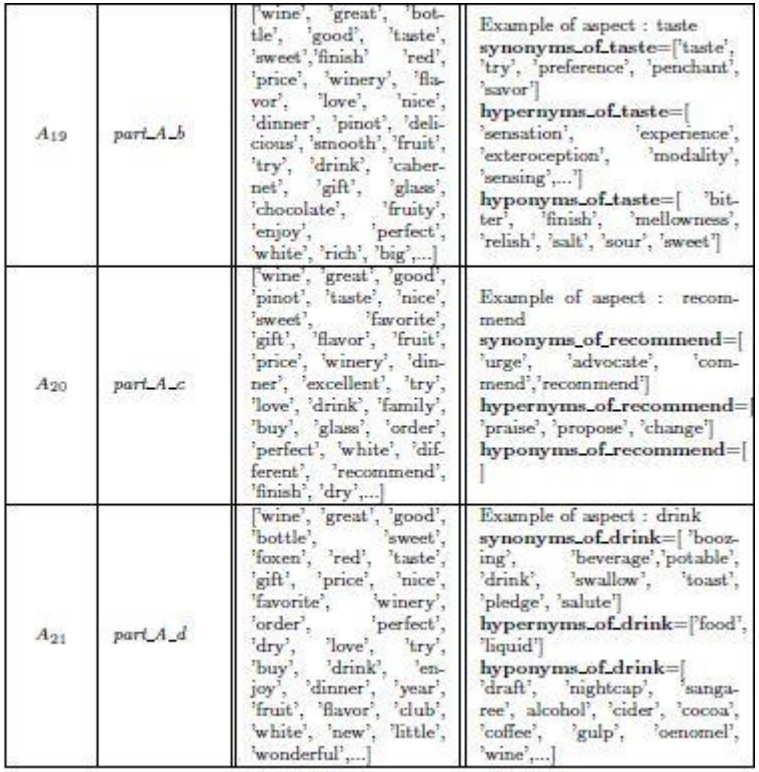


Figure 2 : Twenty most frequent words of the first fifteen experiments of wine reviews after preprocessing.

5.2 Results of aspects’ extraction and expansion of them on thesauri.



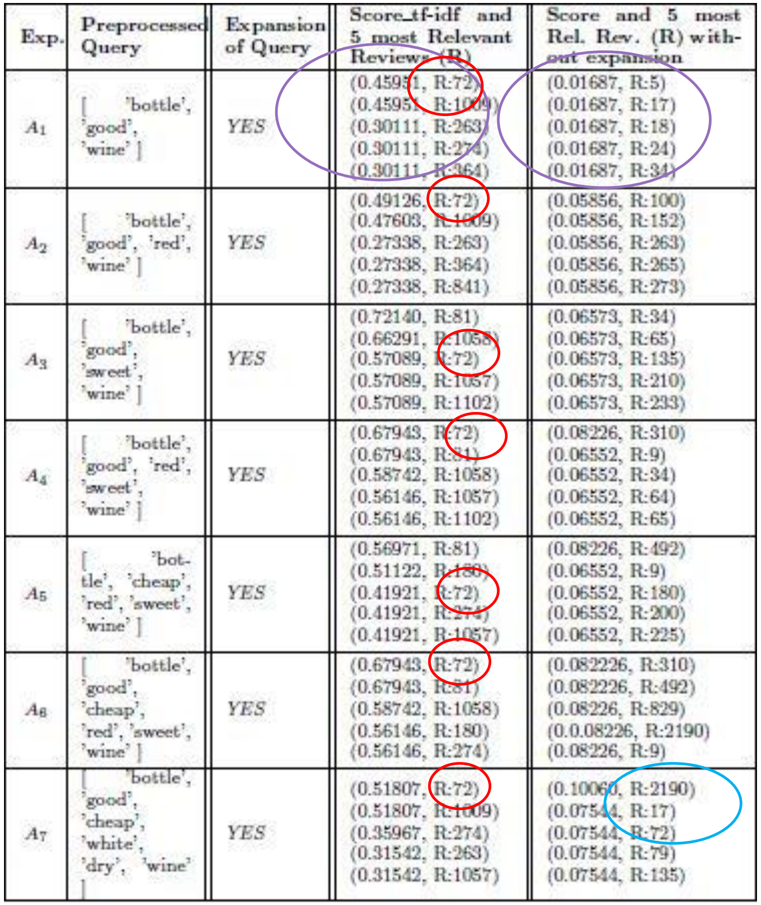
Experiment	Reviews	Top-aspects (TOP_NN_VB_JJ)	Thesauri of top aspects
$A_1 - A_{15}$	*total_A*	['wine’, 'great’, 'good’, 'bottle’, 'taste’, 'sweet’, 'red’, 'price’, 'nice’, 'pinot’, 'favorite’, 'flavor’, 'gift’, 'fruit’, 'dinner’, 'delicious’, 'perfect’, 'try’, 'love’, 'drink’, 'glass’, 'cabernet’, 'order’, 'enjoy’, 'food’, 'white’, 'dry’, 'smooth’, 'little’,...]	Example of aspect : wine **synonyms_of_wine**=['wine’, 'vino’, 'wine-colored’, 'wine-coloured”] **hypernyms_of_wine**=['alcohol’, 'dark_red’, 'drink’, 'regale’,...] **hyponyms_of_wine**=['bordeaux’, 'burgundy’, 'tokay’, 'varietal’, 'vermouth’, 'vintage’,...]
A_{16}, A_{22}, A_{24}	*part_A_1*	['wine’, 'great’, 'good’, 'bottle’, 'taste’, 'sweet’, 'red’, 'price’, 'flavor’, 'nice’, 'favorite’, 'smooth’, 'dinner’, 'white’, 'fruit’, 'delicious’, 'love’, 'glass’, 'cabernet’, 'try’, 'drink’, 'fruity’, 'merlot’, 'pinot’, 'gift’, 'food’, 'little’, 'perfect’, 'dry’,...]	Example of aspect : fruit **synonyms_of_fruit**=['fruit’, 'yield’] **hypernyms_of_fruit**=['product’, 'consequence’, 'bear’] **hyponyms_of_fruit**= ['achene’, 'acorn’, 'berry’, 'buckthorn_berry’, 'chokecherry’, 'cubeb’, 'drupe’,...]
A_{17}, A_{23}, A_{25}	*part_A_2*	['wine’, 'great’, 'good’, 'bottle’, 'red’, 'pinot’, 'taste’, 'sweet’, 'gift’, 'nice’, 'price’, 'favorite’, 'winery’, 'fruit’, 'flavor’, 'dry’, 'delicious’, 'order’, 'dinner’, 'love’, 'try’, 'buy’, 'drink’, 'year’, 'family’, 'enjoy’, 'white’, 'perfect’, 'excellent’, 'new’,...]	Example of aspect : delicious **synonyms_of_delicious**=['delightful’, 'yummy’, 'delicious’, 'delectable’, 'luscious’,...] **hypernyms_of_delicious**=['eating_apple’] **hyponyms_of_delicious**= ['golden_delicious’, 'red_delicious’]
A_{18}	*part_A_a*	['wine’, 'good’, 'bottle’, 'great’, 'taste’, 'sweet’, 'red’, 'price’, 'flavor’, 'white’, 'smooth’, 'glass’, 'fruit’, 'dry’, 'cabernet’, 'delicious’, 'dinner’, 'drink’, 'try’, 'buy’, 'love’, 'food’, 'merlot’, 'fruity’, 'pinot’, 'gift’, 'perfect’, 'local’,...]	Example of aspect : red **synonyms_of_red**=['red’, 'redness’, 'reddish’, 'ruddy’, 'cerise’, 'cherry’, 'crimson’, 'ruby’,...] **hypernyms_of_red**=['sum’, 'radical’, 'chromatic_color’] **hyponyms_of_red**= ['cerise’, 'chrome_red’, 'crimson’, 'dark_red’, 'purplish_red’, 'sanguine’]



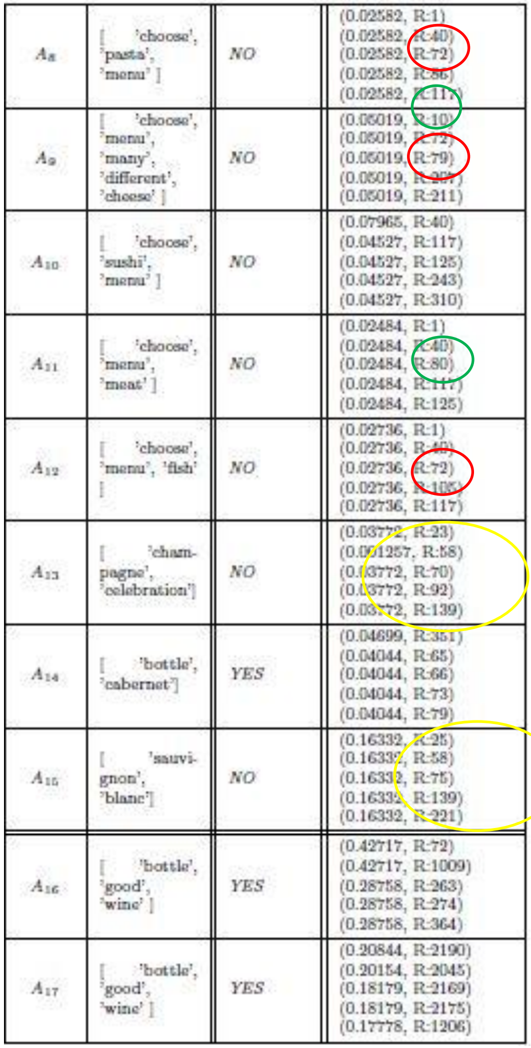
A_{19}	*part_A_b*	[’wine’, ’great’, ’bottle’, ’good’, ’taste’, ’sweet’,’finish’ ’red’, ’price’, ’winery’, ’flavor’, ’love’, ’nice’, ’dinner’, ’pinot’, ’delicious’, ’smooth’, ’fruit’, ’try’, ’drink’, ’cabernet’, ’gift’, ’glass’, ’chocolate’, ’fruity’, ’enjoy’, ’perfect’, ’white’, ’rich’, ’big’,...]	Example of aspect : taste **synonyms_of_taste**=[’taste’, ’try’, ’preference’, ’penchant’, ’savor’] **hypernyms_of_taste**=[’sensation’, ’experience’, ’exteroception’, ’modality’, ’sensing’,...’] **hyponyms_of_taste**=[’bitter’, ’finish’, ’mellowness’, ’relish’, ’salt’, ’sour’, ’sweet’]
A_{20}	*part_A_c*	[’wine’, ’great’, ’good’, ’pinot’, ’taste’, ’nice’, ’sweet’, ’favorite’, ’gift’, ’flavor’, ’fruit’, ’price’, ’winery’, ’dinner’, ’excellent’, ’try’, ’love’, ’drink’, ’family’, ’buy’, ’glass’, ’order’, ’perfect’, ’white’, ’different’, ’recommend’, ’finish’, ’dry’,...]	Example of aspect : recommend **synonyms_of_recommend**=[’urge’, ’advocate’, ’commend’,’recommend’] **hypernyms_of_recommend**=[’praise’, ’propose’, ’change’] **hyponyms_of_recommend**=[ ]
A_{21}	*part_A_d*	[’wine’, ’great’, ’good’, ’bottle’, ’sweet’, ’foxen’, ’red’, ’taste’, ’gift’, ’price’, ’nice’, ’favorite’, ’winery’, ’order’, ’perfect’, ’dry’, ’love’, ’try’, ’buy’, ’drink’, ’enjoy’, ’dinner’, ’year’, ’fruit’, ’flavor’, ’club’, ’white’, ’new’, ’little’, ’wonderful’,...]	Example of aspect : drink **synonyms_of_drink**=[’boozing’, ’beverage’,’potable’, ’drink’, ’swallow’, ’toast’, ’pledge’, ’salute’] **hypernyms_of_drink**=[’food’, ’liquid’] **hyponyms_of_drink**=[’draft’, ’nightcap’, ’sangaree’, alcohol’, ’cider’, ’cocoa’, ’coffee’, ’gulp’, ’oenomel’, ’wine’,...]

Table 3 : Results of the aspects' extraction and expansion of them on thesauri of the wine reviews.

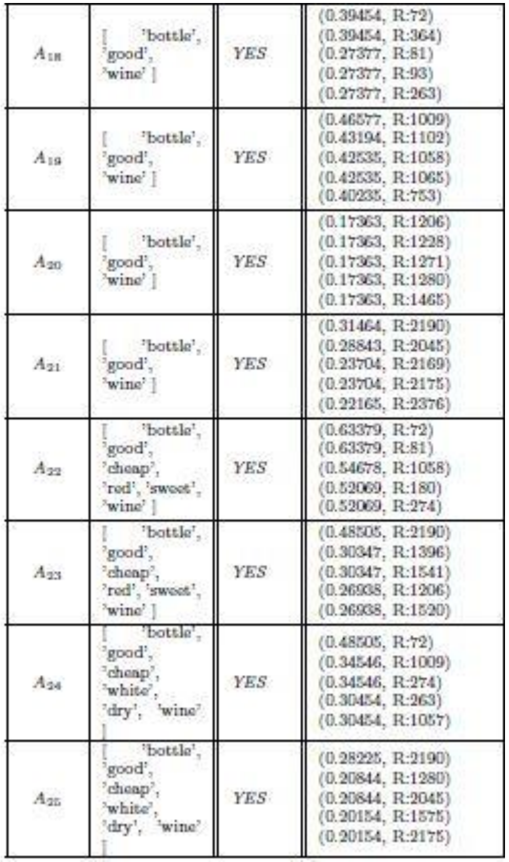
5.3 Results of the query system for the most relevant reviews.



Exp.	Preprocessed Query	Expansion of Query	Score_tf-idf and 5 most Relevant Reviews (R)	Score and 5 most Rel. Rev. (R) without expansion
A_1	['bottle', 'good', 'wine']	*YES*	(0.45951, R:72) (0.45951, R:1009) (0.30111, R:263) (0.30111, R:274) (0.30111, R:364)	(0.01687, R:5) (0.01687, R:17) (0.01687, R:18) (0.01687, R:24) (0.01687, R:34)
A_2	['bottle', 'good', 'red', 'wine']	*YES*	(0.49126, R:72) (0.47603, R:1009) (0.27338, R:263) (0.27338, R:364) (0.27338, R:841)	(0.05856, R:100) (0.05856, R:152) (0.05856, R:263) (0.05856, R:265) (0.05856, R:273)
A_3	['bottle', 'good', 'sweet', 'wine']	*YES*	(0.72140, R:81) (0.66291, R:1058) (0.57089, R:72) (0.57089, R:1057) (0.57089, R:1102)	(0.06573, R:34) (0.06573, R:65) (0.06573, R:135) (0.06573, R:210) (0.06573, R:233)
A_4	['bottle', 'good', 'red', 'sweet', 'wine']	*YES*	(0.67943, R:72) (0.67943, R:81) (0.58742, R:1058) (0.56146, R:1057) (0.56146, R:1102)	(0.08226, R:310) (0.06552, R:9) (0.06552, R:34) (0.06552, R:64) (0.06552, R:65)
A_5	['bottle', 'cheap', 'red', 'sweet', 'wine']	*YES*	(0.56971, R:81) (0.51122, R:180) (0.41921, R:72) (0.41921, R:274) (0.41921, R:1057)	(0.08226, R:492) (0.06552, R:9) (0.06552, R:180) (0.06552, R:200) (0.06552, R:225)
A_6	['bottle', 'good', 'cheap', 'red', 'sweet', 'wine']	*YES*	(0.67943, R:72) (0.67943, R:81) (0.58742, R:1058) (0.56146, R:180) (0.56146, R:274)	(0.082226, R:310) (0.082226, R:492) (0.08226, R:829) (0.0.08226, R:2190) (0.08226, R:9)
A_7	['bottle', 'good', 'cheap', 'white', 'dry', 'wine']	*YES*	(0.51807, R:72) (0.51807, R:1009) (0.35967, R:274) (0.31542, R:263) (0.31542, R:1057)	(0.10060, R:2190) (0.07544, R:17) (0.07544, R:72) (0.07544, R:79) (0.07544, R:135)



A_8	['choose', 'pasta', 'menu']	*NO*	(0.02582, R:1) (0.02582, R:40) (0.02582, R:72) (0.02582, R:86) (0.02582, R:117)
A_9	['choose', 'menu', 'many', 'different', 'cheese']	*NO*	(0.05019, R:10) (0.05019, R:72) (0.05019, R:79) (0.05019, R:207) (0.05019, R:211)
A_{10}	['choose', 'sushi', 'menu']	*NO*	(0.07965, R:40) (0.04527, R:117) (0.04527, R:125) (0.04527, R:243) (0.04527, R:310)
A_{11}	['choose', 'menu', 'meat']	*NO*	(0.02484, R:1) (0.02484, R:40) (0.02484, R:80) (0.02484, R:117) (0.02484, R:125)
A_{12}	['choose', 'menu', 'fish']	*NO*	(0.02736, R:1) (0.02736, R:40) (0.02736, R:72) (0.02736, R:105) (0.02736, R:117)
A_{13}	['champagne', 'celebration']	*NO*	(0.03772, R:23) (0.001257, R:58) (0.03772, R:70) (0.03772, R:92) (0.03772, R:139)
A_{14}	['bottle', 'cabernet']	*YES*	(0.04699, R:351) (0.04044, R:65) (0.04044, R:66) (0.04044, R:73) (0.04044, R:79)
A_{15}	['sauvignon', 'blanc']	*NO*	(0.16332, R:25) (0.16332, R:58) (0.16332, R:75) (0.16332, R:139) (0.16332, R:221)
A_{16}	['bottle', 'good', 'wine']	*YES*	(0.42717, R:72) (0.42717, R:1009) (0.28758, R:263) (0.28758, R:274) (0.28758, R:364)
A_{17}	['bottle', 'good', 'wine']	*YES*	(0.20844, R:2190) (0.20154, R:2045) (0.18179, R:2169) (0.18179, R:2175) (0.17778, R:1206)



A_{18}	['bottle', 'good', 'wine']	*YES*	(0.39454, R:72) (0.39454, R:364) (0.27377, R:81) (0.27377, R:93) (0.27377, R:263)
A_{19}	['bottle', 'good', 'wine']	*YES*	(0.46577, R:1009) (0.43194, R:1102) (0.42535, R:1058) (0.42535, R:1065) (0.40235, R:753)
A_{20}	['bottle', 'good', 'wine']	*YES*	(0.17363, R:1206) (0.17363, R:1228) (0.17363, R:1271) (0.17363, R:1280) (0.17363, R:1465)
A_{21}	['bottle', 'good', 'wine']	*YES*	(0.31464, R:2190) (0.28843, R:2045) (0.23704, R:2169) (0.23704, R:2175) (0.22165, R:2376)
A_{22}	['bottle', 'good', 'cheap', 'red', 'sweet', 'wine']	*YES*	(0.63379, R:72) (0.63379, R:81) (0.54678, R:1058) (0.52069, R:180) (0.52069, R:274)
A_{23}	['bottle', 'good', 'cheap', 'red', 'sweet', 'wine']	*YES*	(0.48505, R:2190) (0.30347, R:1396) (0.30347, R:1541) (0.26938, R:1206) (0.26938, R:1520)
A_{24}	['bottle', 'good', 'cheap', 'white', 'dry', 'wine']	*YES*	(0.48505, R:72) (0.34546, R:1009) (0.34546, R:274) (0.30454, R:263) (0.30454, R:1057)
A_{25}	['bottle', 'good', 'cheap', 'white', 'dry', 'wine']	*YES*	(0.28225, R:2190) (0.20844, R:1280) (0.20844, R:2045) (0.20154, R:1575) (0.20154, R:2175)

Table 4: Query system and the five most relevant reviews to user's query of the experiments of wine reviews.

6.

Conclusion

Conclusion-Future work

- Pre-processing of data and user's query contributed to the reduction of the volume of the data and to the transformation of questions in the latter of keywords.
- The extraction of aspect words was an essential procedure for the increase of the accuracy of our system, as most information and opinions are gathered on these words.
- The expansion of the queries to synonyms, hypernyms, or hyponyms of the aspect words turned out necessary, as many times an individual is searching for a product or an idea with specific aspects and expresses it on an equivalent way. Also in this way we can obtain more accurate results.

Future work : A good thought would be to take advantage of further elements of the reviews (e.g. “helpful” or “reviewTime”) in order to enhance our system with a method that will also generate results ranked by other criteria (e.g. reliability or date).

Σας ευχαριστώ πολύ για την προσοχή σας!