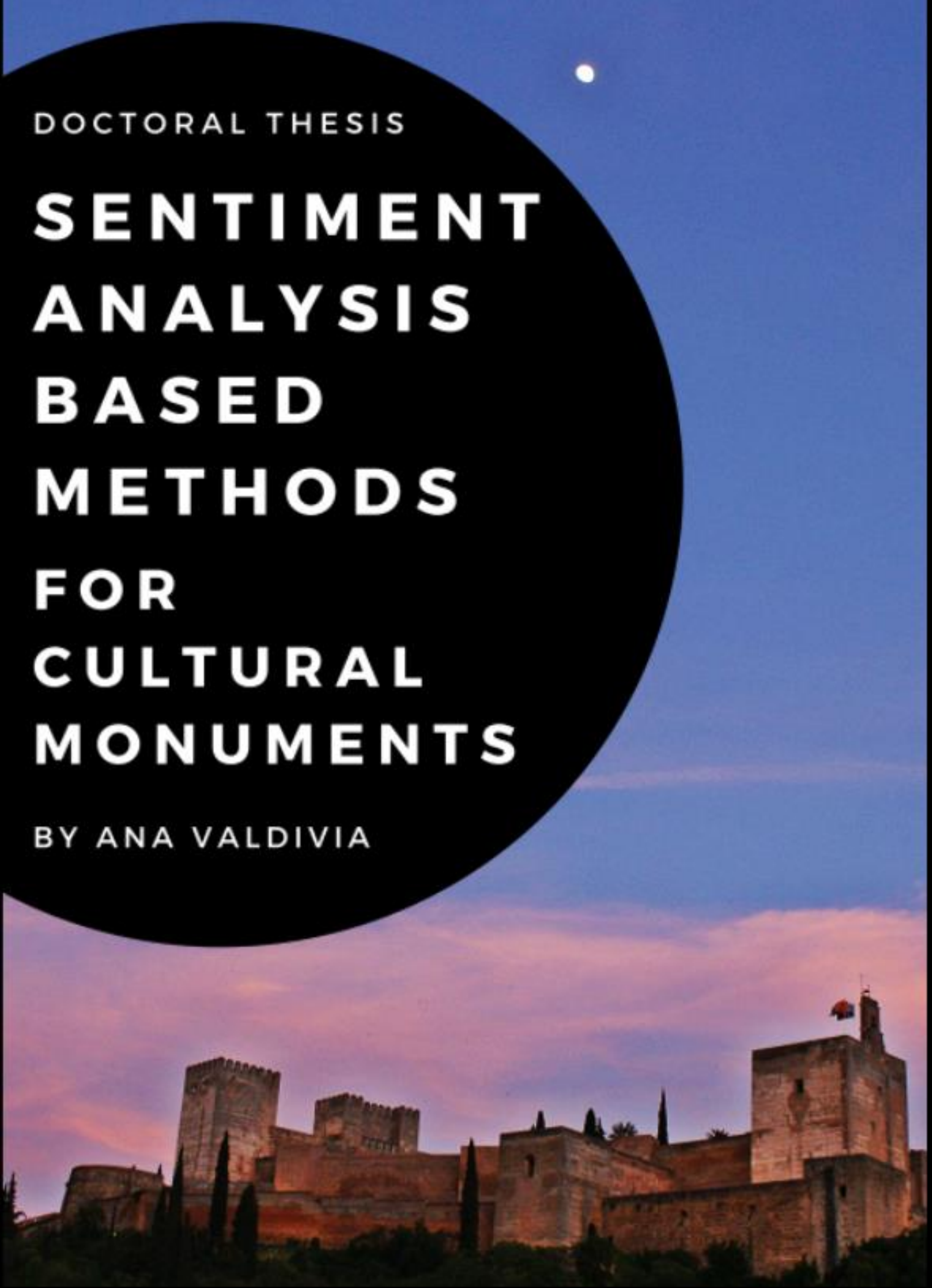


PhD Dissertation

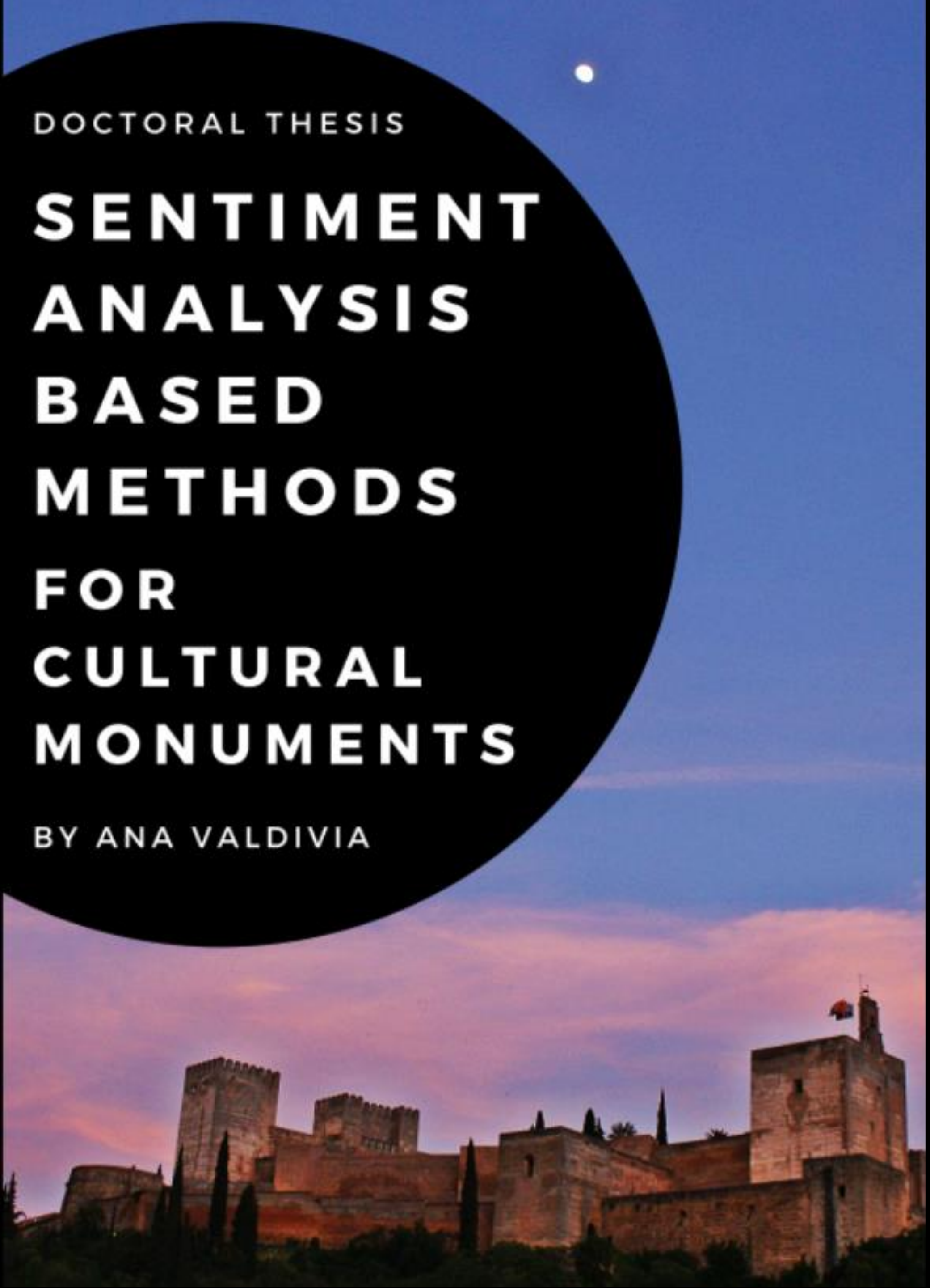
Supervisors:

Francisco Herrera & M. Victoria Luzón

1st February, 2019

UNIVERSIDAD DE GRANADA

OUTLINE



INTRODUCTION

UNIVERSIDAD DE GRANADA

INTRODUCTION

Basic concepts

Opinion

- It is an **expression** that is given in a **conscious way** due to an emotion.
- We can express an opinion through facial expressions or words.

Web 2.0

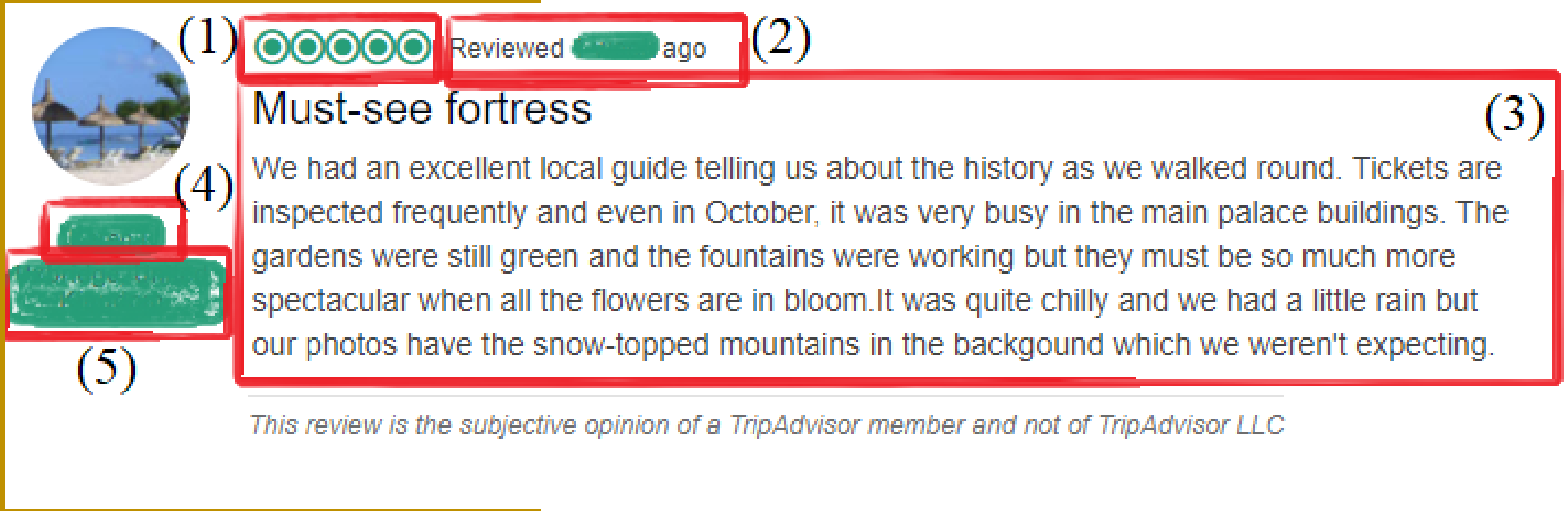
- A **second generation** in the development of the World Wide Web, conceived as a combination of concepts, trends, and technologies that **focus on user collaboration, sharing of user-generated content**, and **social networking**.

Source: op4g.com

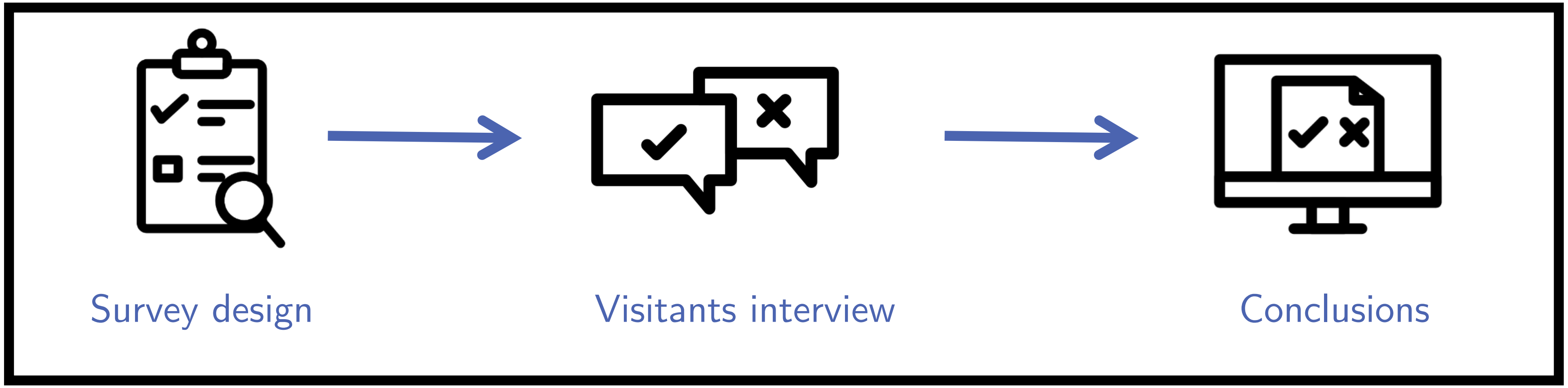
Basic concepts

TripAdvisor

- It is a **travel website company** that **shows hotel and restaurant reviews**, accommodation bookings and other travel-related content. It also includes interactive travel forums.



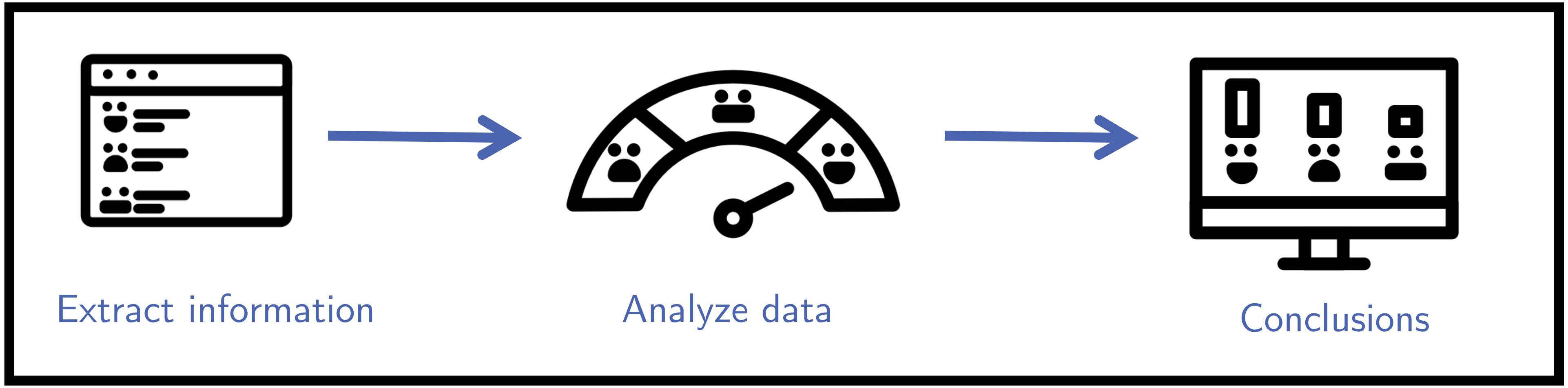
How cultural monuments analyze people’s opinion?



Survey’s drawbacks

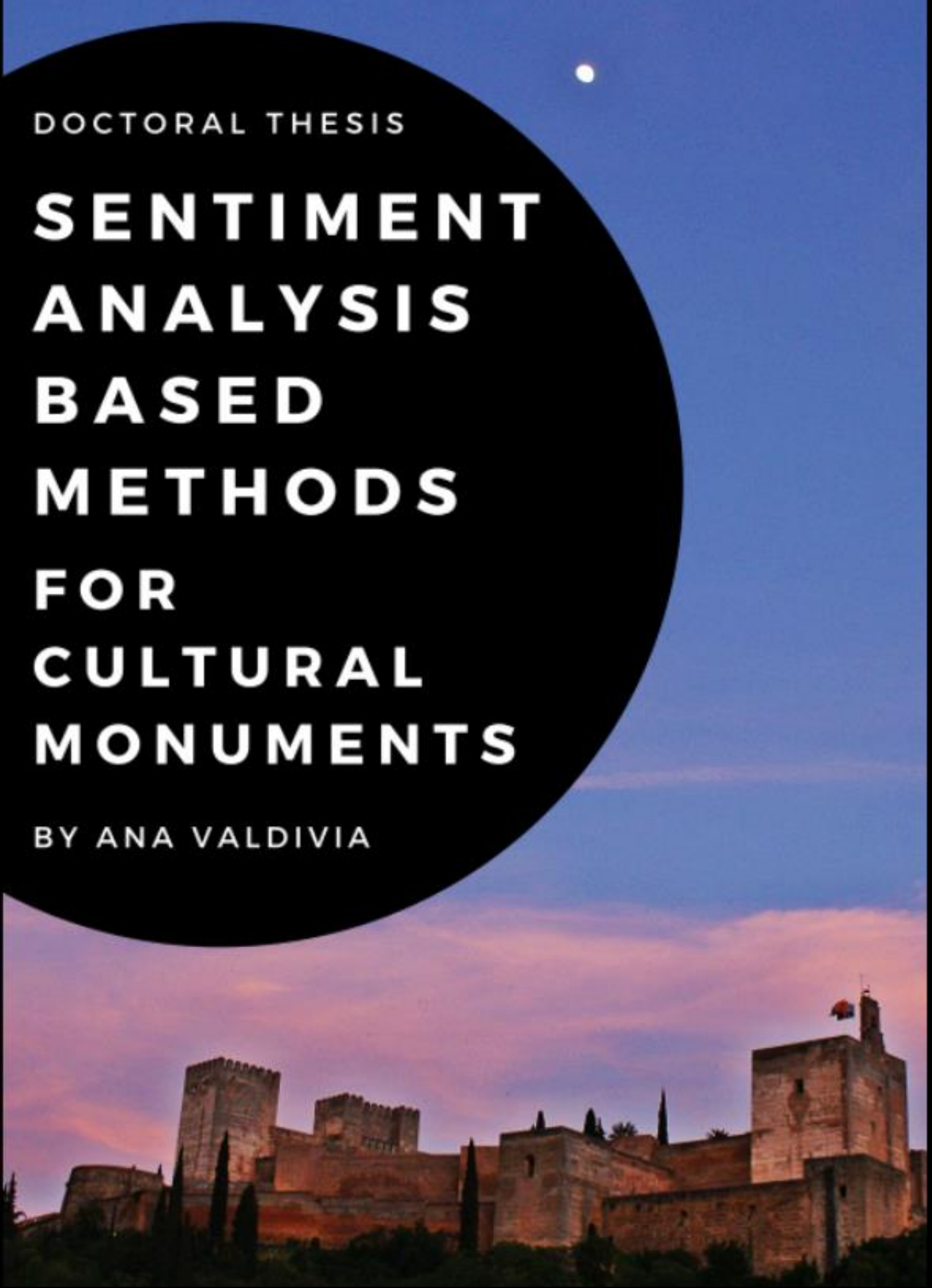
- What happens if we don’t ask key questions?
- What happens if people don’t want to answer questions?
- ...

How cultural monuments may analyze people's opinion?



Web 2.0 and social networks

Machine and Deep Learning
Sentiment Analysis



PRELIMINARIES

- Sentiment Analysis
- Sentiment Analysis Methods

UNIVERSIDAD DE GRANADA

P R E L I M I N A R I E S

Sentiment Analysis

Liu's definition

- It is the field of study that **analyzes people's opinions**, sentiments, appraisals, attitudes, and emotions towards **entities** and their **attributes** expressed in written text.

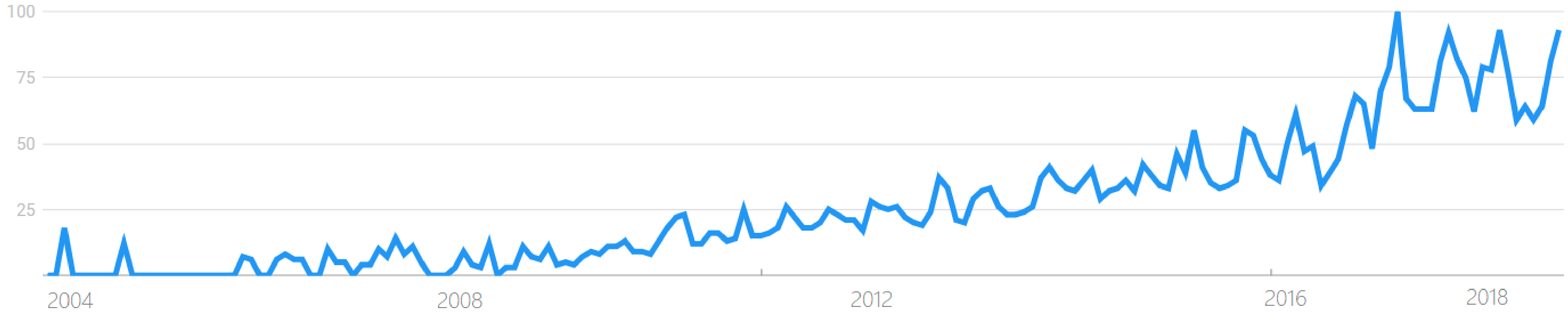
Cambria's definition

- The **set of computational techniques** for **extracting**, **classifying**, **understanding**, and **assessing** the **opinions** expressed in various on-line news sources, social media comments, and other user-generated contents.

Sentiment Analysis

Google trends

- Evolution of interest of Sentiment Analysis.



Sentiment Analysis search trend in Google.

Tasks

- ❑ Subjectivity classification.
- ❑ **Polarity extraction and classification.**
- ❑ **Opinion summarization.**
- ❑ Opinion visualization.
- ❑ Sarcasm detection.
- ❑ **Entity/Aspect extraction.**
- ❑ …

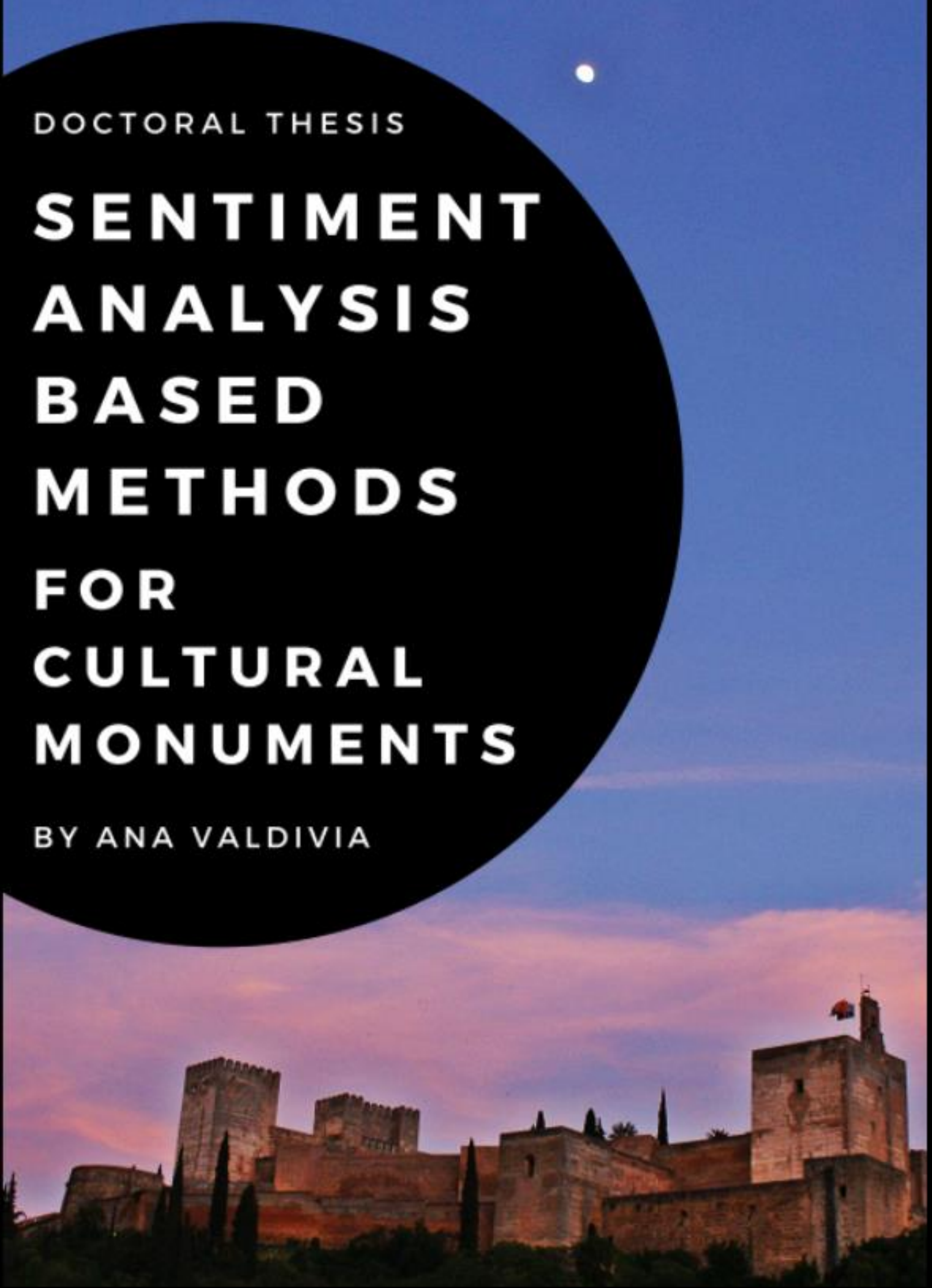
Sentiment Analysis

Polarity

- ❑ Label of the sentiment.
- ❑ Different categories:
  - {positive, neutral, negative}
  - {very positive, positive, neutral, negative, very negative}
  - {1, 2, 3, 4, 5}
  - [0, 1]

Levels of analysis

- ❑ Document
- ❑ Sentence
- ❑ Entity/Aspect



PRELIMINARIES

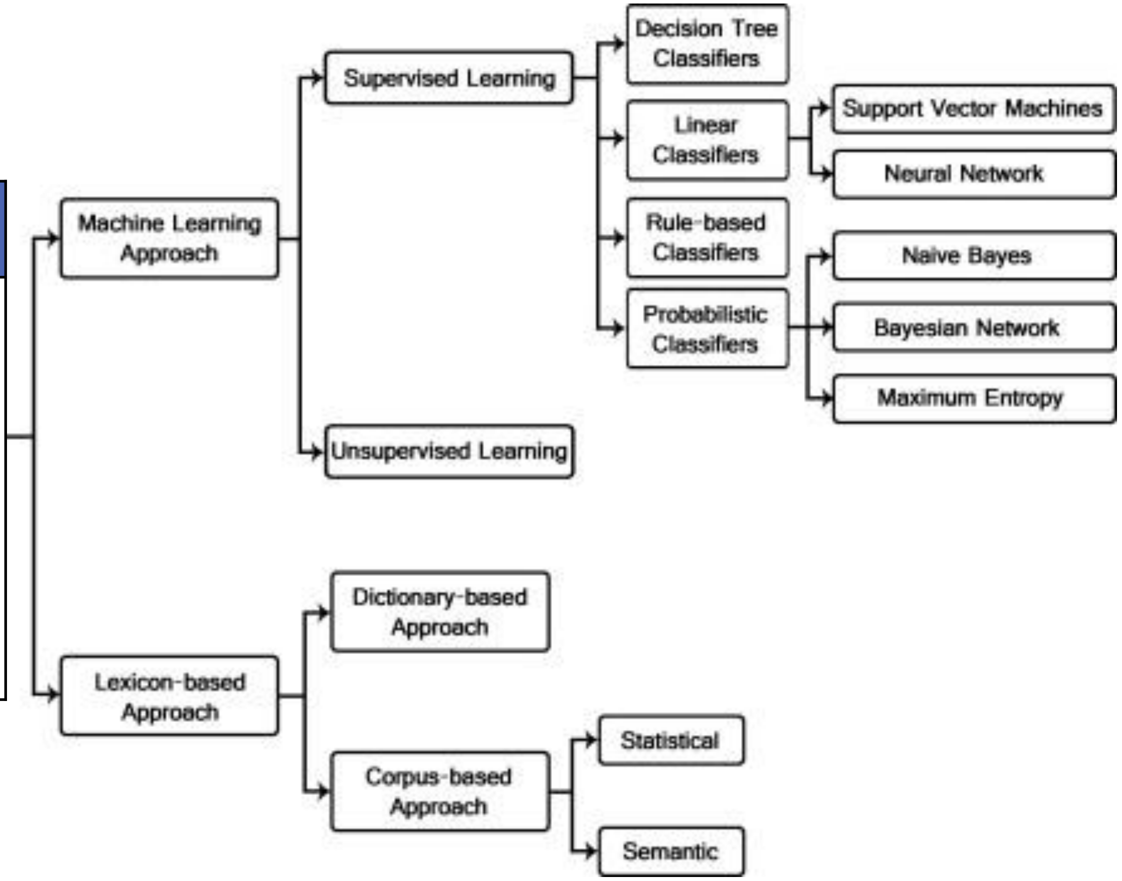
- Sentiment Analysis
- Sentiment Analysis Methods

UNIVERSIDAD DE GRANADA

Basic concepts

Sentiment Analysis Methods

- ❑ Algorithms that extract polarities from written text.
- ❑ Drawbacks:
  - Domain adaptation
  - Corpus creation
  - Labelling cost



Medhat, W., Hassan, A., & Korashy, H. (2014). **Sentiment analysis algorithms and applications: A survey**. *Ain Shams Engineering Journal*, *5*(4), 1093-1113

PRELIMINARIES

Machine Learning

Polarity Extraction Classification

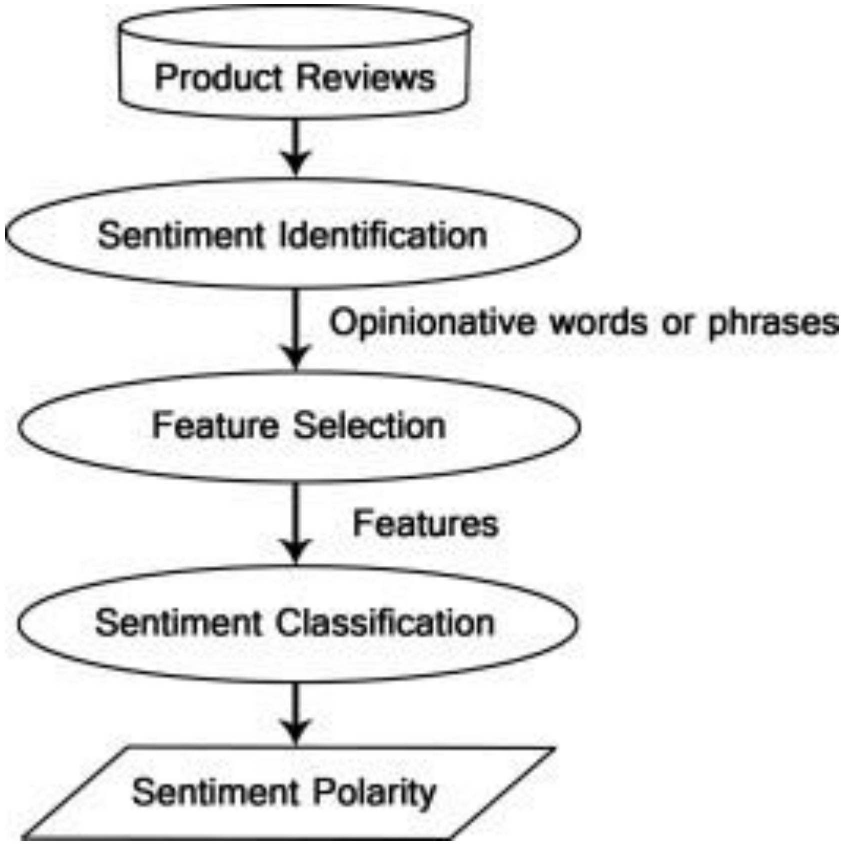


Fig. 5: Basic pipeline of ML-based approach.

Medhat, W., Hassan, A., & Korashy, H. (2014). **Sentiment analysis algorithms and applications: A survey**. *Ain Shams Engineering Journal*, *5*(4), 1093-1113

P R E L I M I N A R I E S

Machine Learning

Polarity Extraction Classification

Term-Document Matrix

$$
\begin{array}{c|cccc}
 & T_1 & T_2 & \dots & T_t \\
\hline
D_1 & w_{11} & w_{21} & \dots & w_{t1} \\
D_2 & w_{12} & w_{22} & \dots & w_{t2} \\
\vdots & \vdots & \vdots & & \vdots \\
\vdots & \vdots & \vdots & & \vdots \\
D_n & w_{1n} & w_{2n} & \dots & w_{tn}
\end{array}
$$

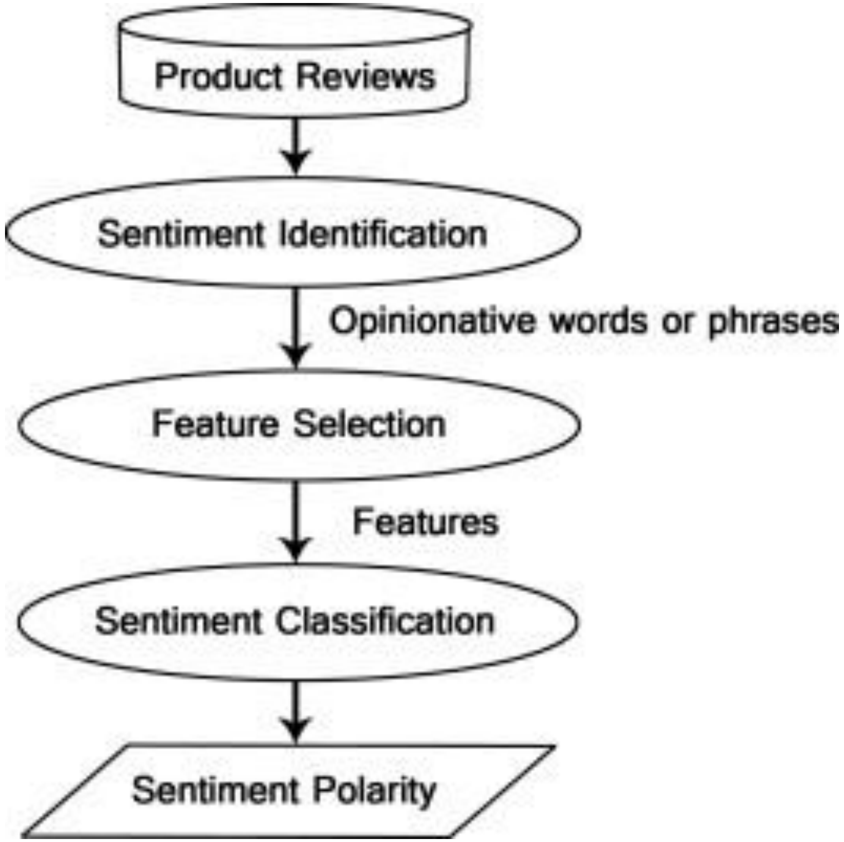


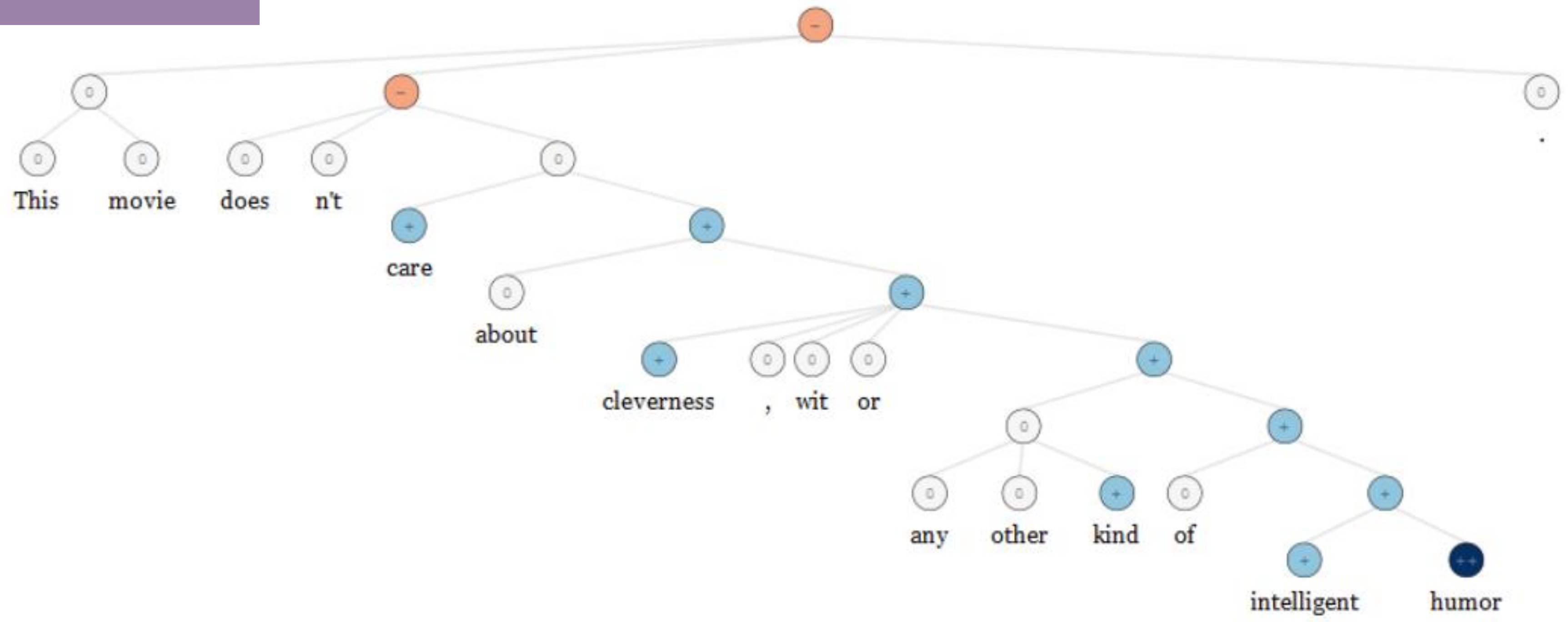
Fig. 5: Basic pipeline of ML-based approach.

Medhat, W., Hassan, A., & Korashy, H. (2014). **Sentiment analysis algorithms and applications: A survey**. *Ain Shams Engineering Journal*, *5*(4), 1093-1113

Examples: Machine Learning

Polarity Extraction Classification

CoreNLP

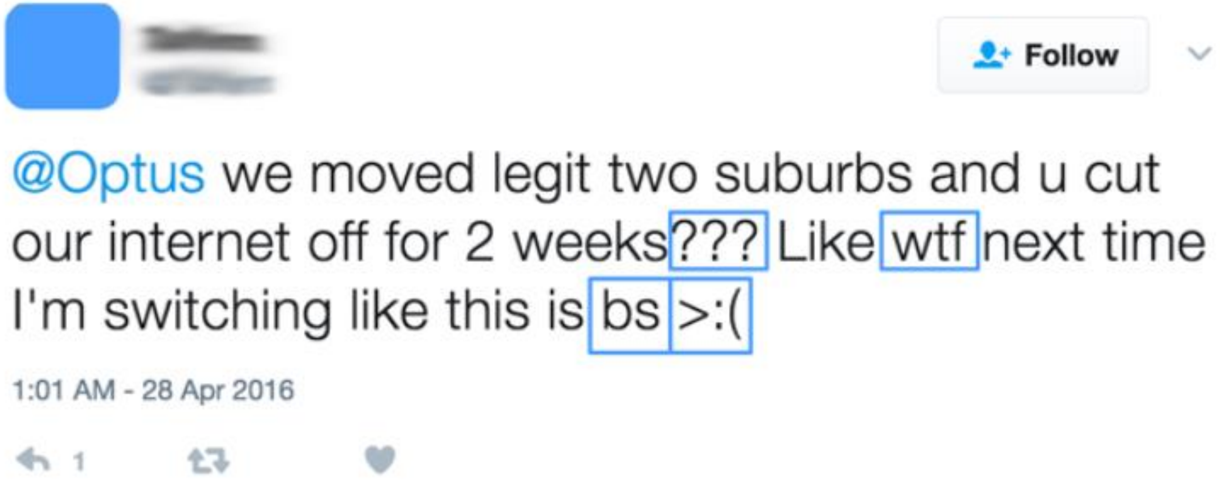


Manning, Christopher, et al. (2014). **The Stanford CoreNLP natural language processing toolkit.** *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*.

Examples

Polarity
Extraction
Classification

Vader

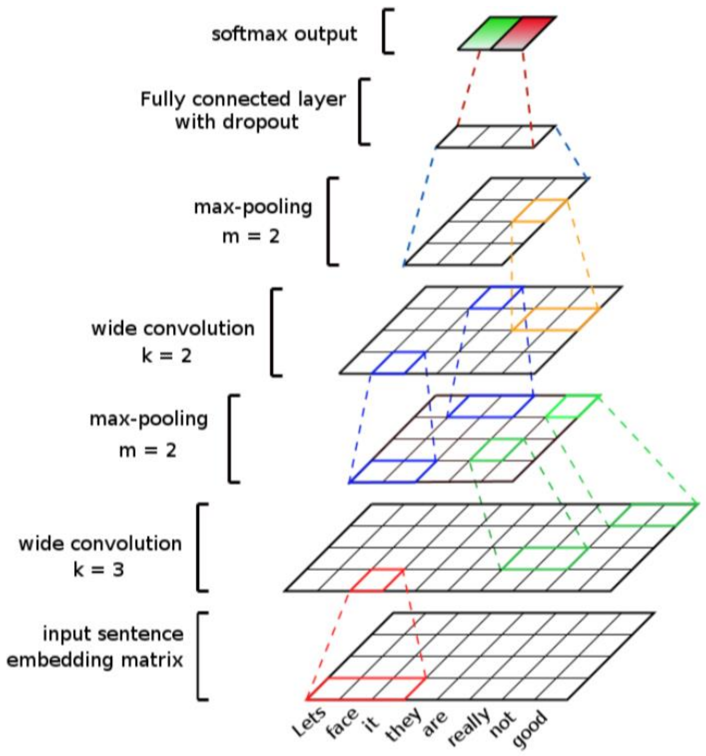


Gilbert, CJ Hutto Eric. (2014). **Vader: A parsimonious rule-based model for sentiment analysis of social media text.** *Eighth International Conference on Weblogs and Social Media (ICWSM-14). Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. vader. hutto. pdf.*

Deep Learning

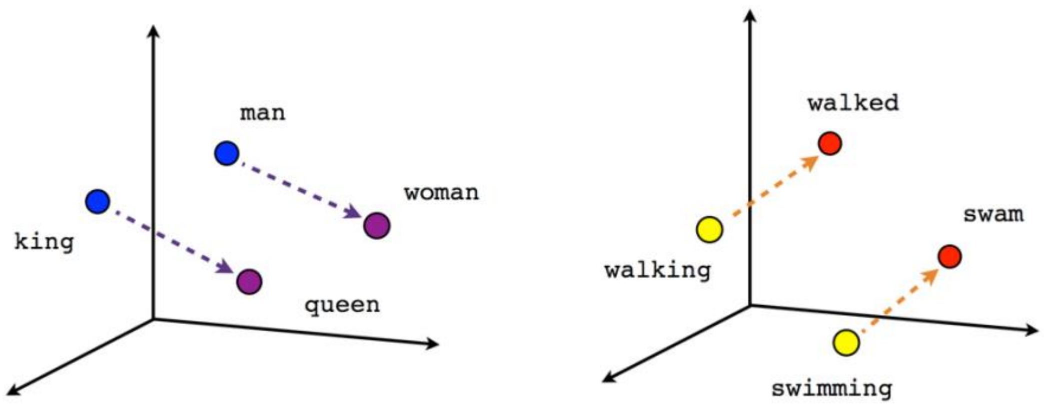
Aspect extraction

CNN



Poria, S., Cambria, E., & Gelbukh, A. (2016). **Aspect extraction for opinion mining with a deep convolutional neural network.** *Knowledge-Based Systems*, *108*, 42-49.

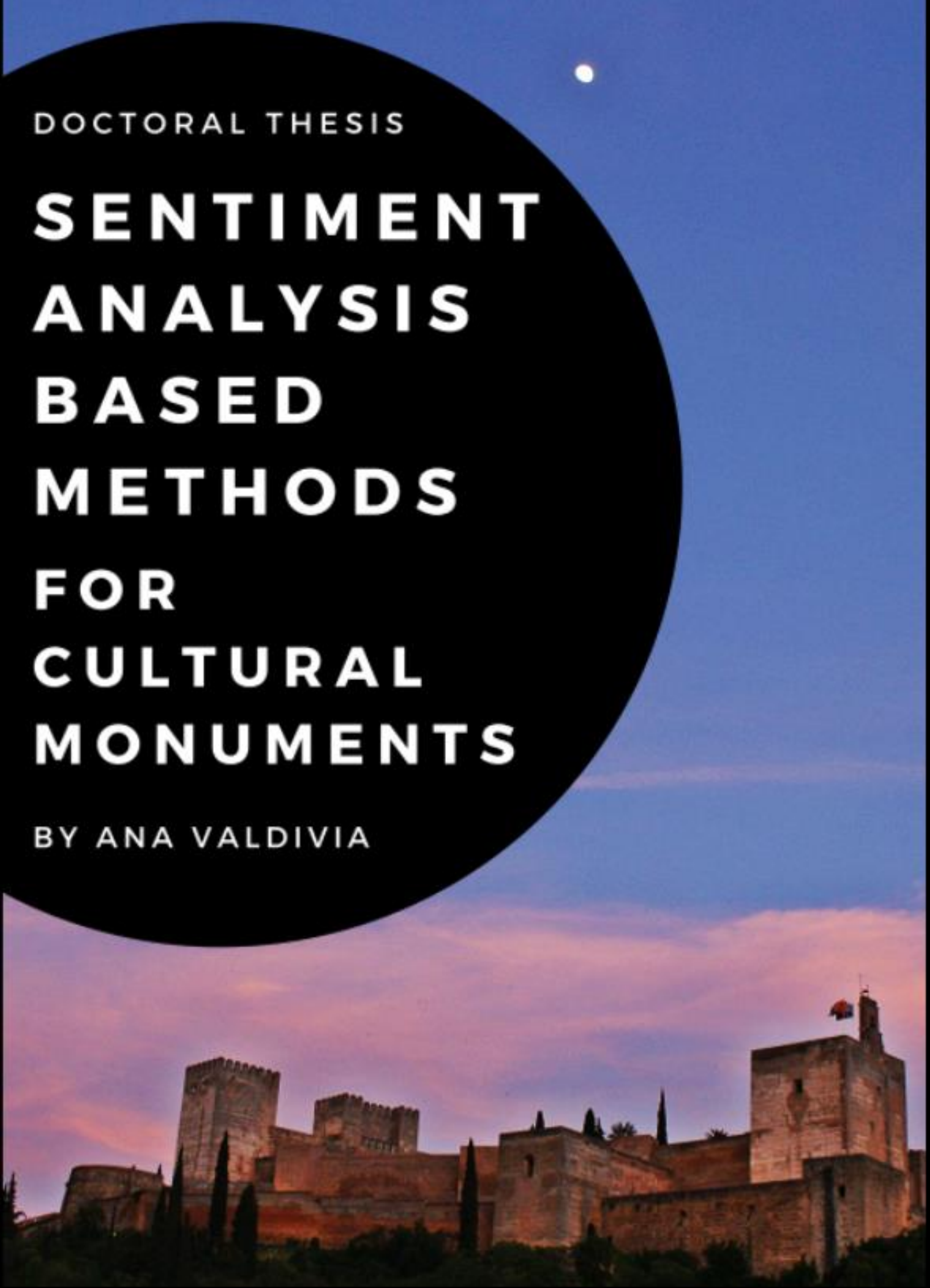
word2vec



Male-Female

Verb tense

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). **Distributed representations of words and phrases and their compositionality.** In *Advances in neural information processing systems* (pp. 3111-3119).



OBJECTIVES

UNIVERSIDAD
DE GRANADA

OBJECTIVES

1. To consider the advisability of TripAdvisor as a source for cultural monument reviews
2. To develop techniques to address the domain adaptation problem of SAMs
3. To address sentiment's inconsistencies detected on cultural monument reviews
4. To improve polarity classification by filtering neutral polarities
5. To develop a methodology that summarizes a large quantity of cultural reviews

OBJECTIVES

INCONSISTENCIES

1 To consider the advisability of TripAdvisor as a source for cultural monument reviews

2 To develop techniques to address the domain adaptation problem of SAMs

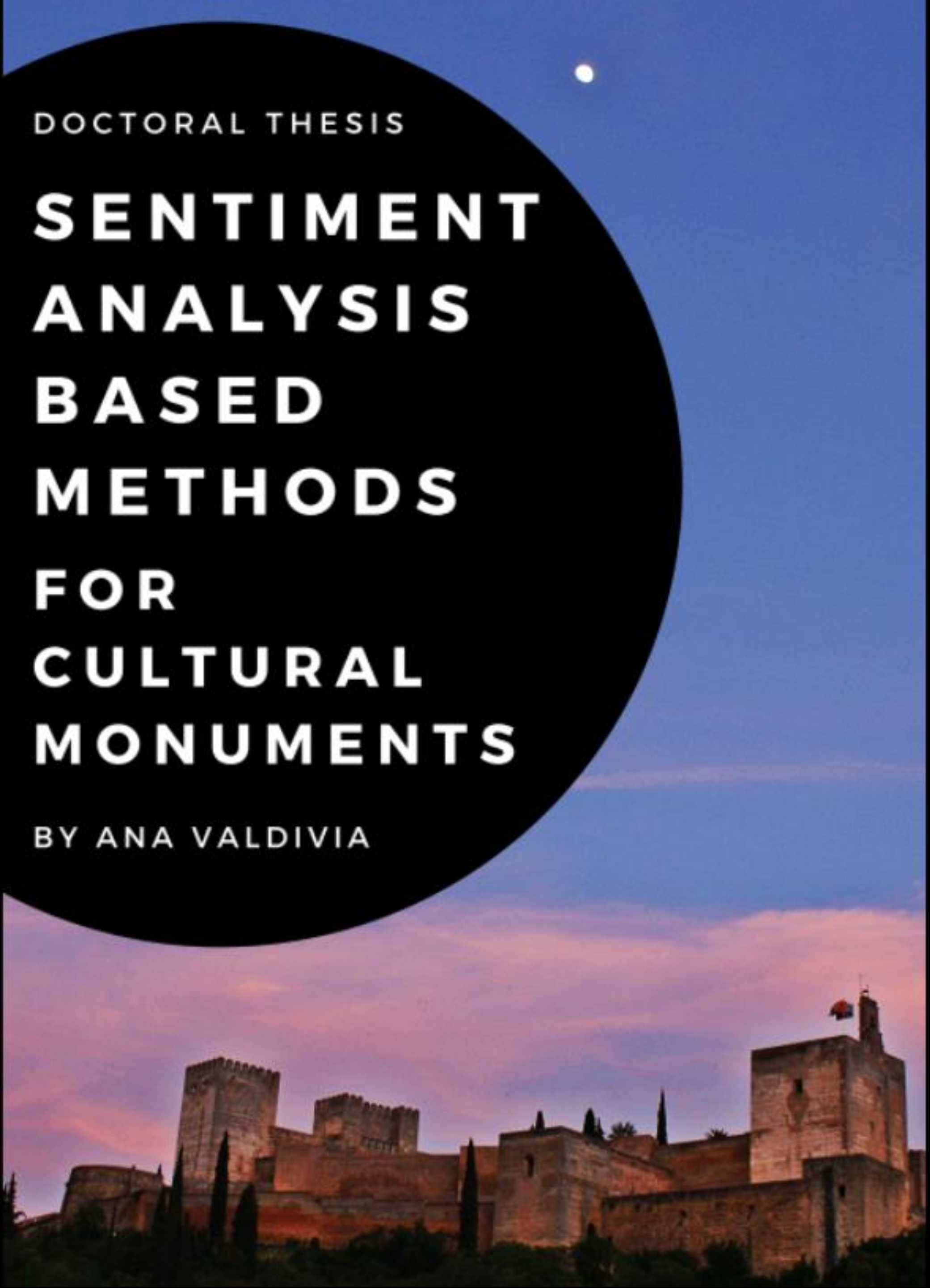
3 To address sentiment's inconsistencies detected on cultural monument reviews

NEUTRALITY

4 To improve polarity classification by filtering neutral polarities

OPINION SUMMARIZATION

5 To develop a methodology that summarizes a large quantity of cultural reviews



RESULTS

- The problem of inconsistencies
- Neutrality Detection
- Opinion Summarization

UNIVERSIDAD DE GRANADA

P U B L I C A T I O N S

The problem of inconsistencies

Motivation

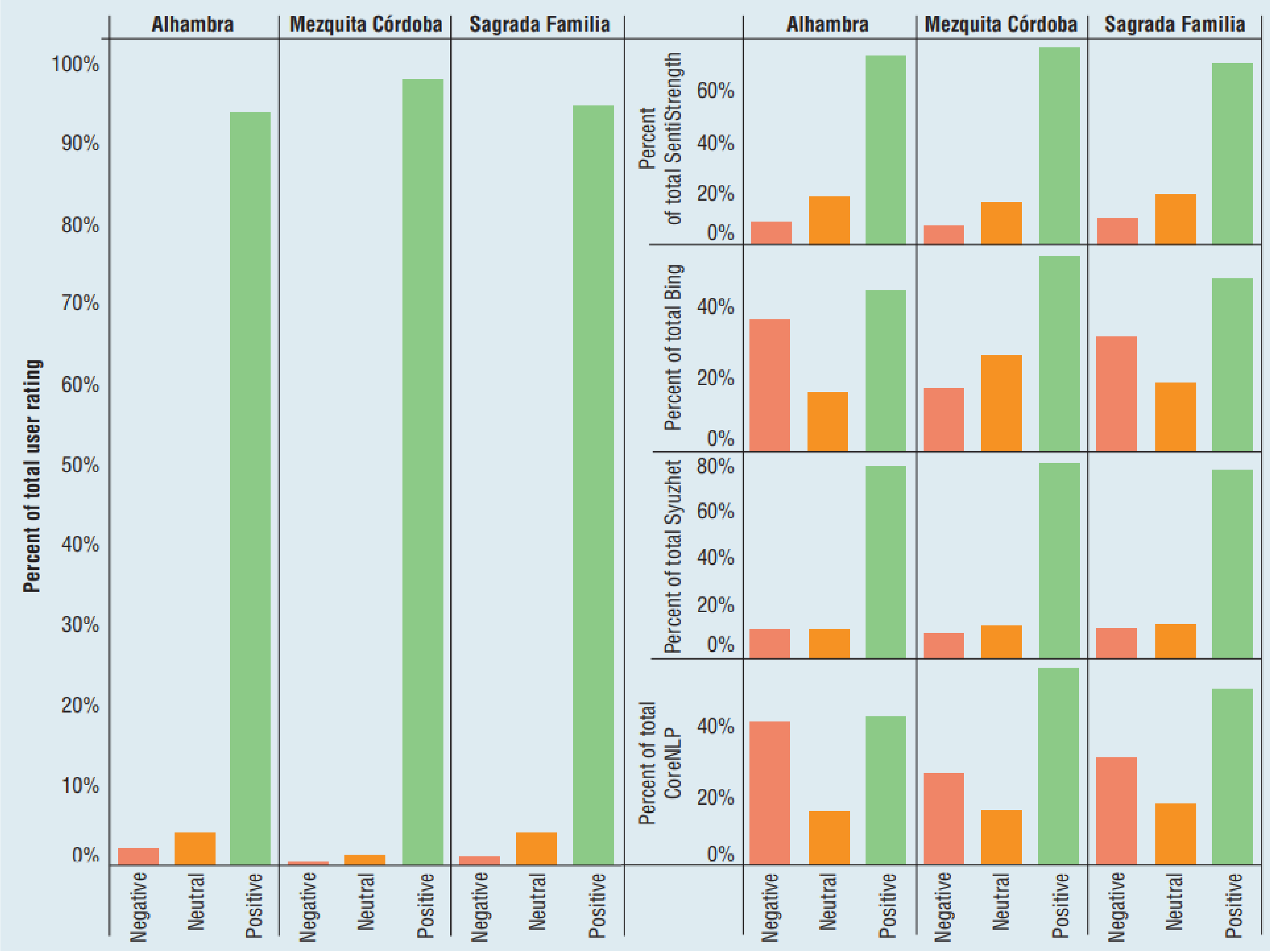
- TripAdvisor as a source for cultural heritage.
- Apply several off-the-shelf SAMs between Users polarities on TripAdvisor reviews.

Objectives

- Analyze the correlation between SAMs and Users polarities.
- Address SAM inconsistencies problem.

Fig. 6: TripAdvisor for cultural monuments.

The problem of inconsistencies

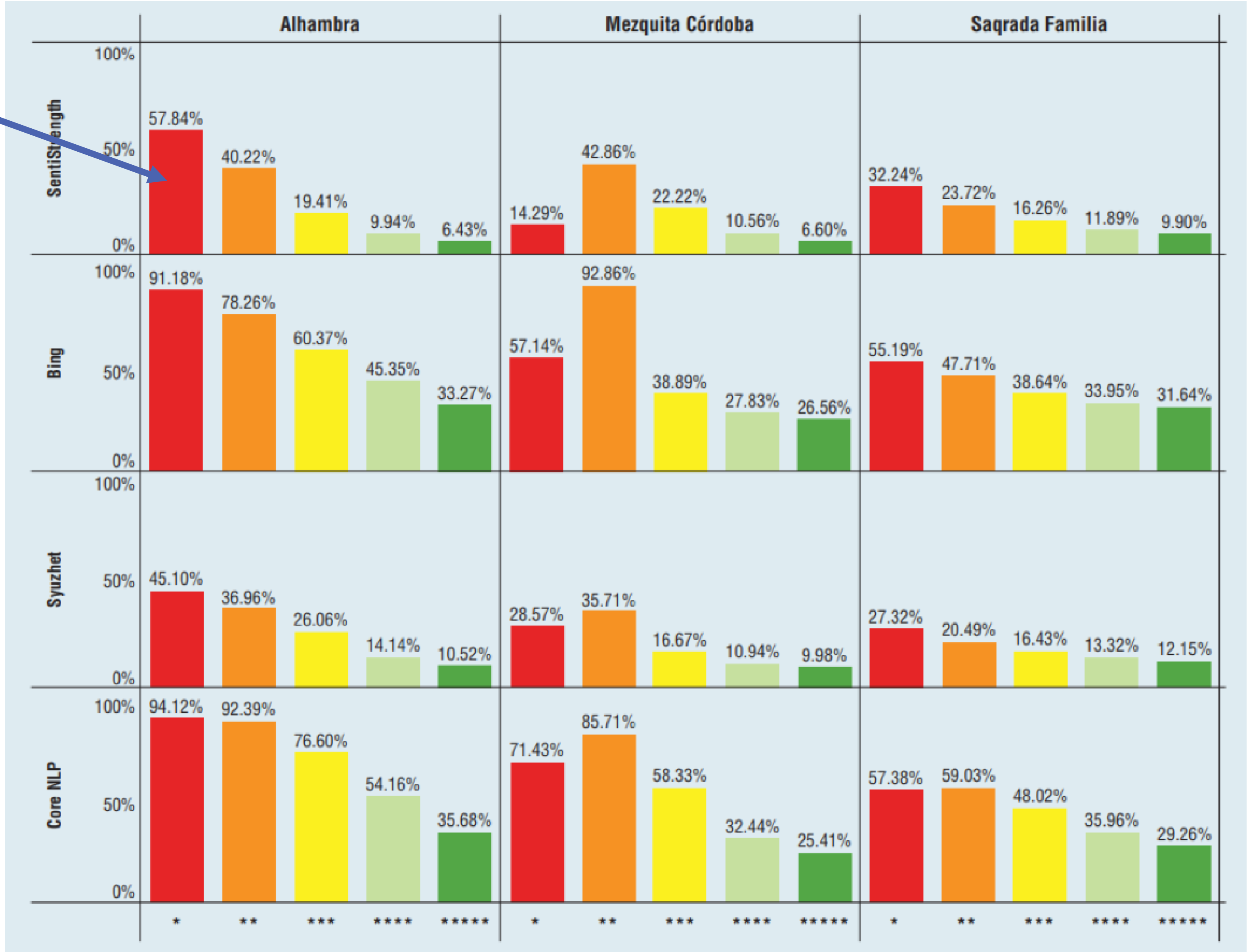


Detected inconsistencies.

The problem of inconsistencies

P U B L I C A T I O N S

SentiStrength detects as negative a 57.84 % of all reviews with one Bubble (*).



The problem of inconsistencies

	Reviews	Words	Sentences	Avg. # words	Avg. # sentences	Avg. User Polarity
Alhambra	7,217	676,398	35,867	93.72	4.97	4.69
Grand Canal	10,730	539,465	47,943	50.28	4.47	4.67
Mezquita de Córdoba	3,526	217,640	13,083	61.72	3.70	4.84
Pantheon	17,279	774,765	76,720	44.84	4.44	4.68
Sagrada Familia	34,558	2,220,719	136,181	64.26	3.94	4.72
Trevi Fountain	15,572	764,998	70,407	49.13	4.52	3.93

Table I: Summary of text properties of 6 datasets.

The problem of inconsistencies

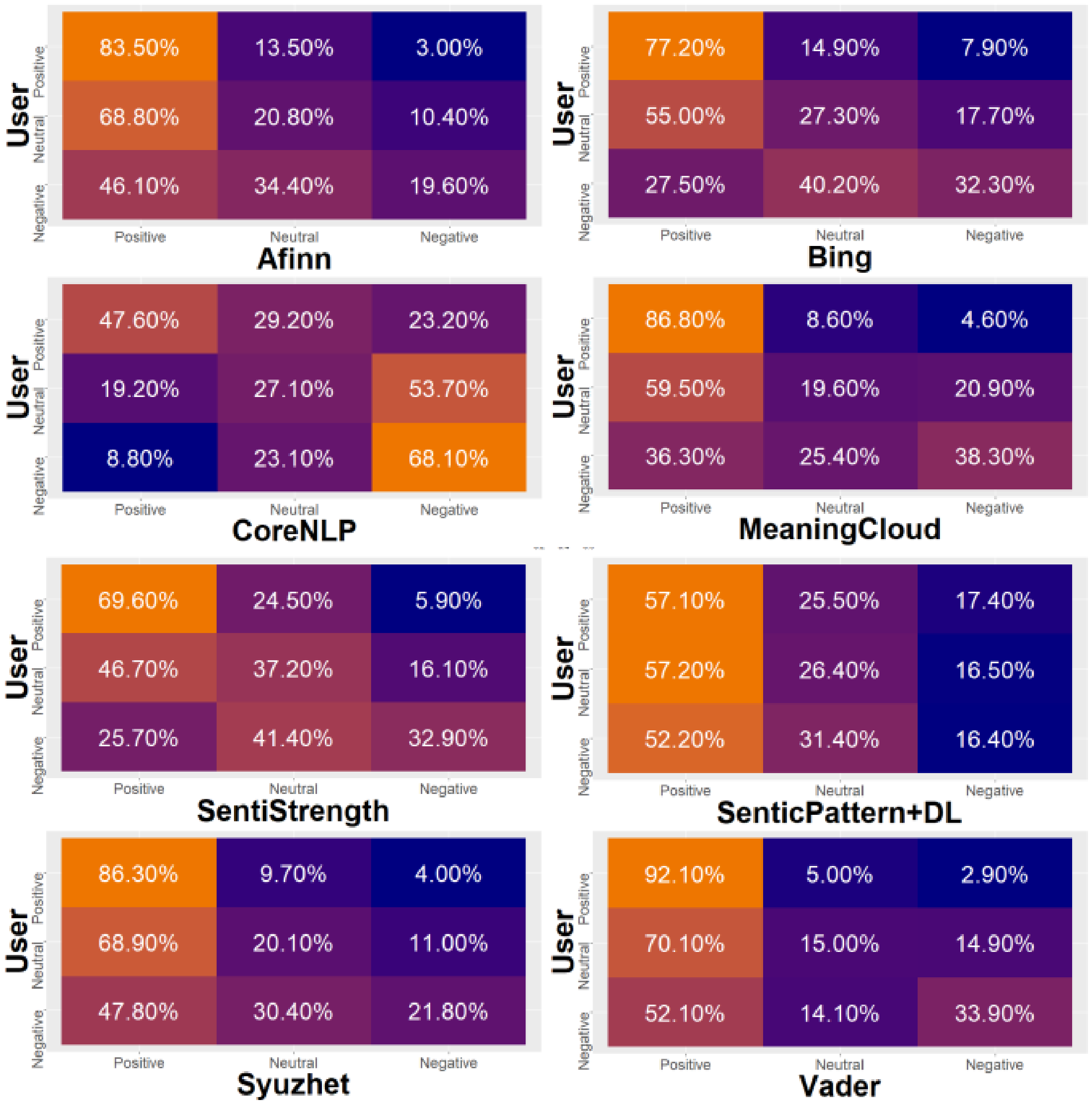


Table II: Correlation matrices between SAMs and User Sentiments.

The problem of inconsistencies

P U B L I C A T I O N S

Polarity Aggregation Model

- We create a new **polarity index** that takes into account both user and SAMs for **overcoming the inconsistency problem**.
- We propose an **aggregation model** guided by the **geometrical mean**, a variant including a parameter to control one variable influence.

$$f(x,y) = \sqrt{xy^{\beta}}$$

where:

- x is the **Normalized User Polarity** of the *i*th-opinion.
- y is the ***k*th-Normalized SAM Polarity** of the *i*th-opinion.
- β is the parameter to control the SAMs polarity influence.

The problem of inconsistencies

Polarity Aggregation Model

- We create a new **polarity index** that takes into account both user and SAMs for **overcoming the inconsistency problem**.
- We propose an **aggregation model** guided by the **geometrical mean**, a variant including a parameter to control one variable influence.

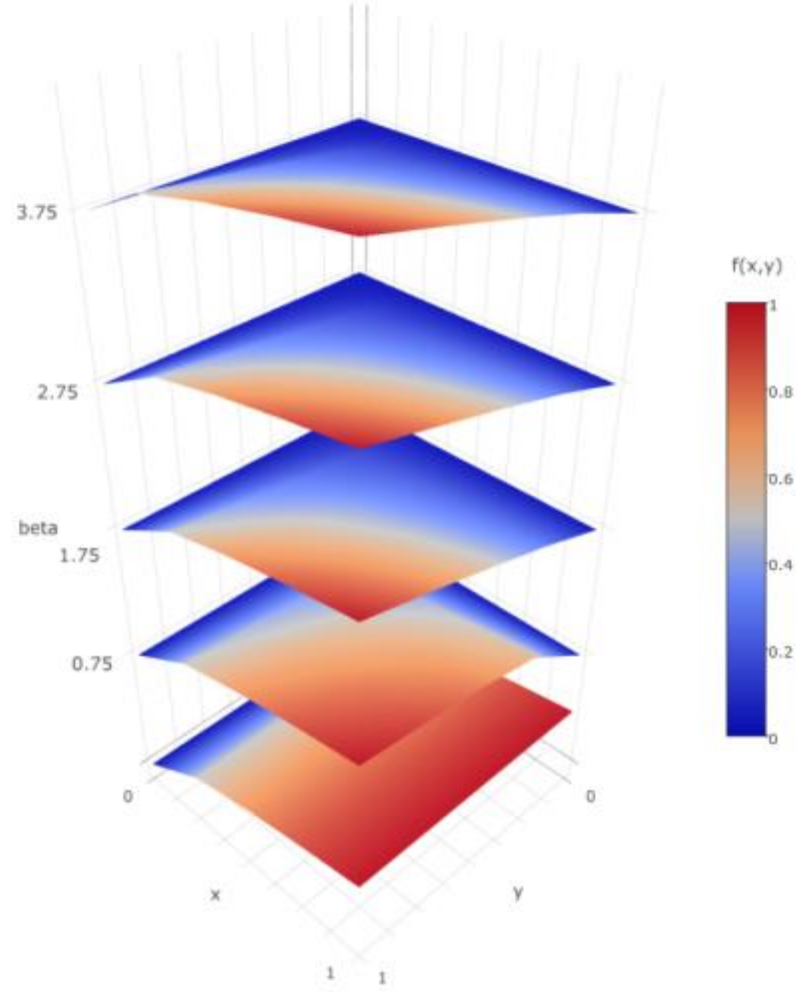


Fig. 9: Distribution of the Polarity Aggregation Model for different β values (0, 0.75, 1.75, 2.75 and 3.75). Bluer colors represent more negative aggregated polarities, more orange colors more positive aggregated polarities.

The problem of inconsistencies

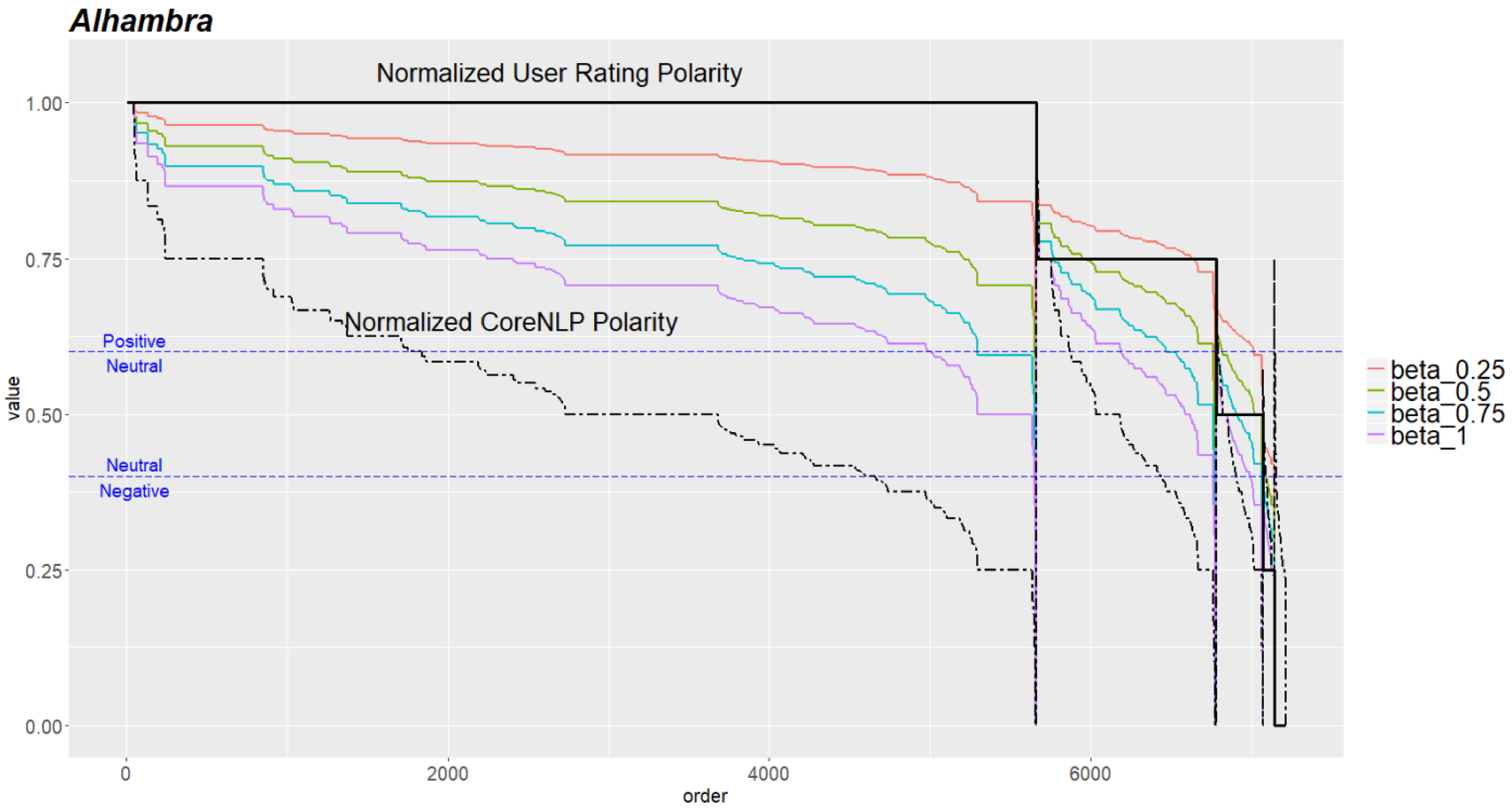
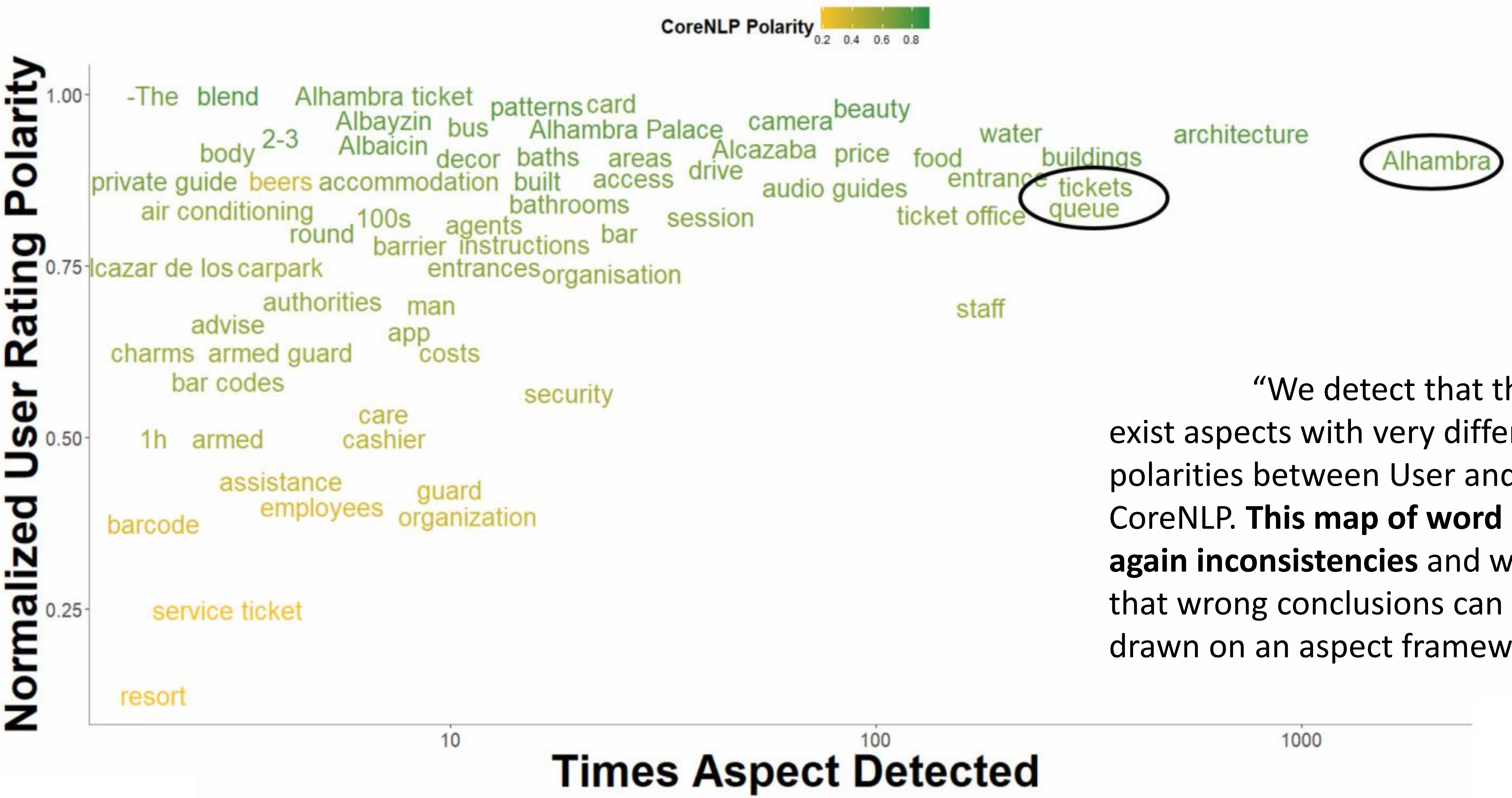


Fig. 10: **Different Polarity Aggregation Models taking account *beta*'s values**. Reviews are sorted on the *x* label in ascending order, from most positive (left) to most negative (right).

The problem of inconsistencies

P U B L I C A T I O N S



"We detect that there exist aspects with very different polarities between User and CoreNLP. **This map of word reflects again inconsistencies** and we show that wrong conclusions can be drawn on an aspect framework."

Fig. 11: This aspect map represents times that an aspect is detected (x axis) taking into account the value of User Polarity (y axis) and CoreNLP Polarity (color scale).

The problem of inconsistencies

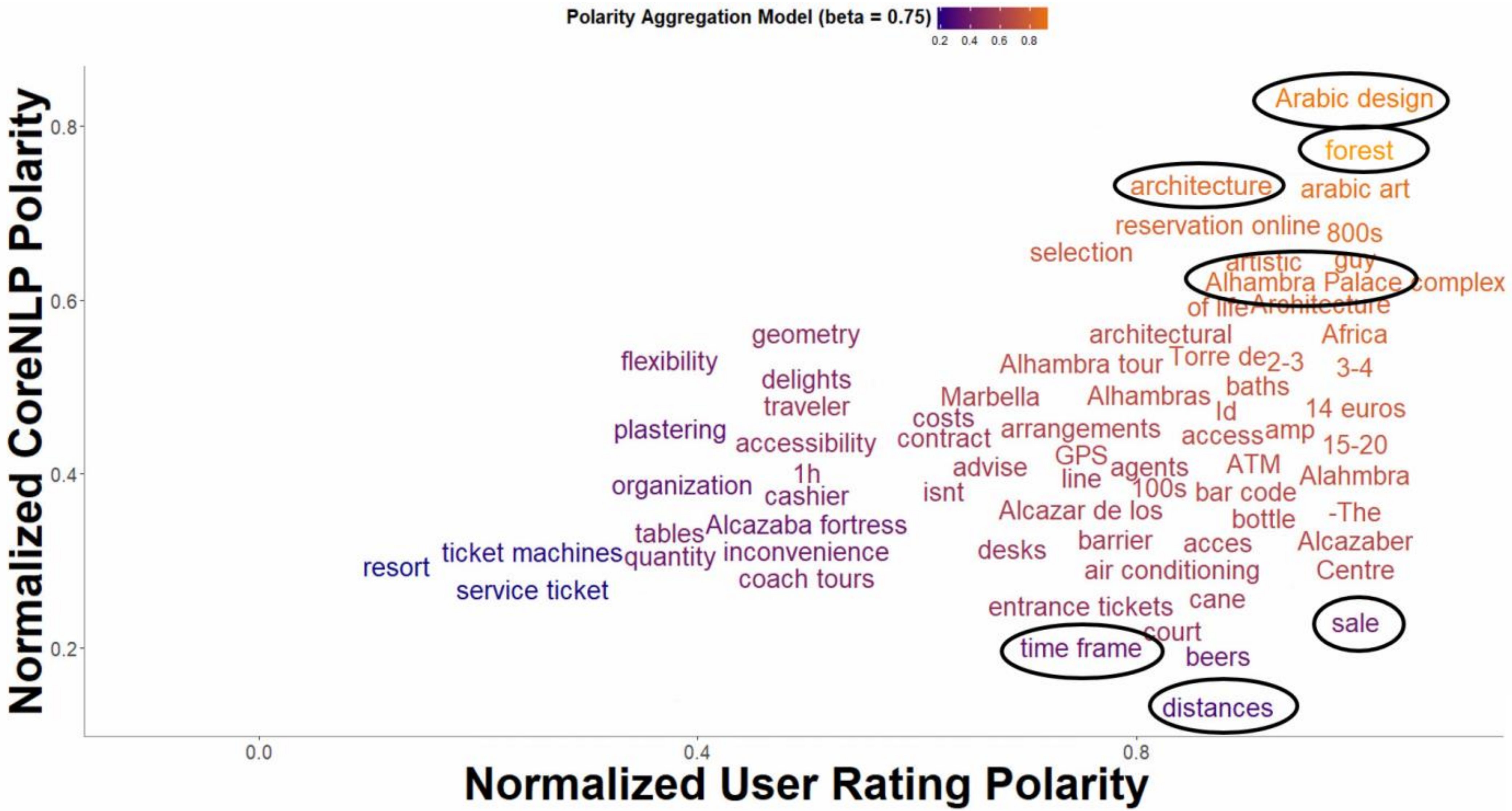


Fig. 12: This aspects map represents the mean of each polarity of each aspect.

PUBLICATIONS

The problem of inconsistencies

Aspect	Text	User	CoreNLP	$\beta = 0.75$
time frame	This place is amazing and should not be missed, no need to add to the thousands other good reviews written hear. I would like to write about my purchasing experience to possibly help someone out in getting this done the easiest way. Trying to get a ticket to see the Alhambra is a project you kind of have to study to know how to do so. I understand why many find it confusing and end up not getting it right. I can only recommend doing it the way I did, as it was simple as 1-2-3: 1. Go to Ticketmaster.es (the Spanish site) and search for tickets for the Alhambra. We got the cheapest best value ones- 15 euro for the general entrance, 2. Purchase tickets to either morning session (ends at 14) or afternoon session (starts at 14 ends at 18/20 depending on season). Know that you are allowed to be at the grounds within that **time frame** but that would be forced to exit, or not allowed in before/after your session. 3. Know that the specific time selected for your ticket indicates a 30 minute window for you to enter the Nasarid palace (but you can tour the rest of the grounds before or/and after visiting the palace) [...].	1	0.40	0.71
time frame	I tried to book a ticket for this place month in advance and my credit card was declined all the time. Even called the local ticket office and they couldn't help, so in desperation asked the hotel I stayed to try to get tickets-well. I think what they try to do is to discourage you to buy the 'cheap' 14 euro ticket and pay 35 or 50 euros for a guided tour-since you have to book a **time frame**. We thought that it will give you space to move around-certainly. It's not-hundreds of people lining up at every corner and rooms, so it's grossly overcrowded.	0.5	0	0

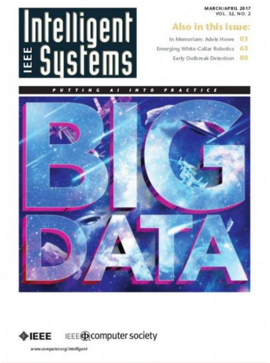
The problem of inconsistencies

Conclusions

- Inconsistencies are detected between SAMs and Users.
- This is because:
  - Humans do not use the same sentiment in every sentence.
  - Domain adaptation of SAMs.
- Polarity Aggregation Model solves the correlation problem.

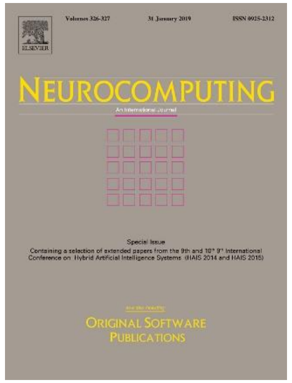
Fig. 13: TripAdvisor for cultural monuments.

The problem of inconsistencies



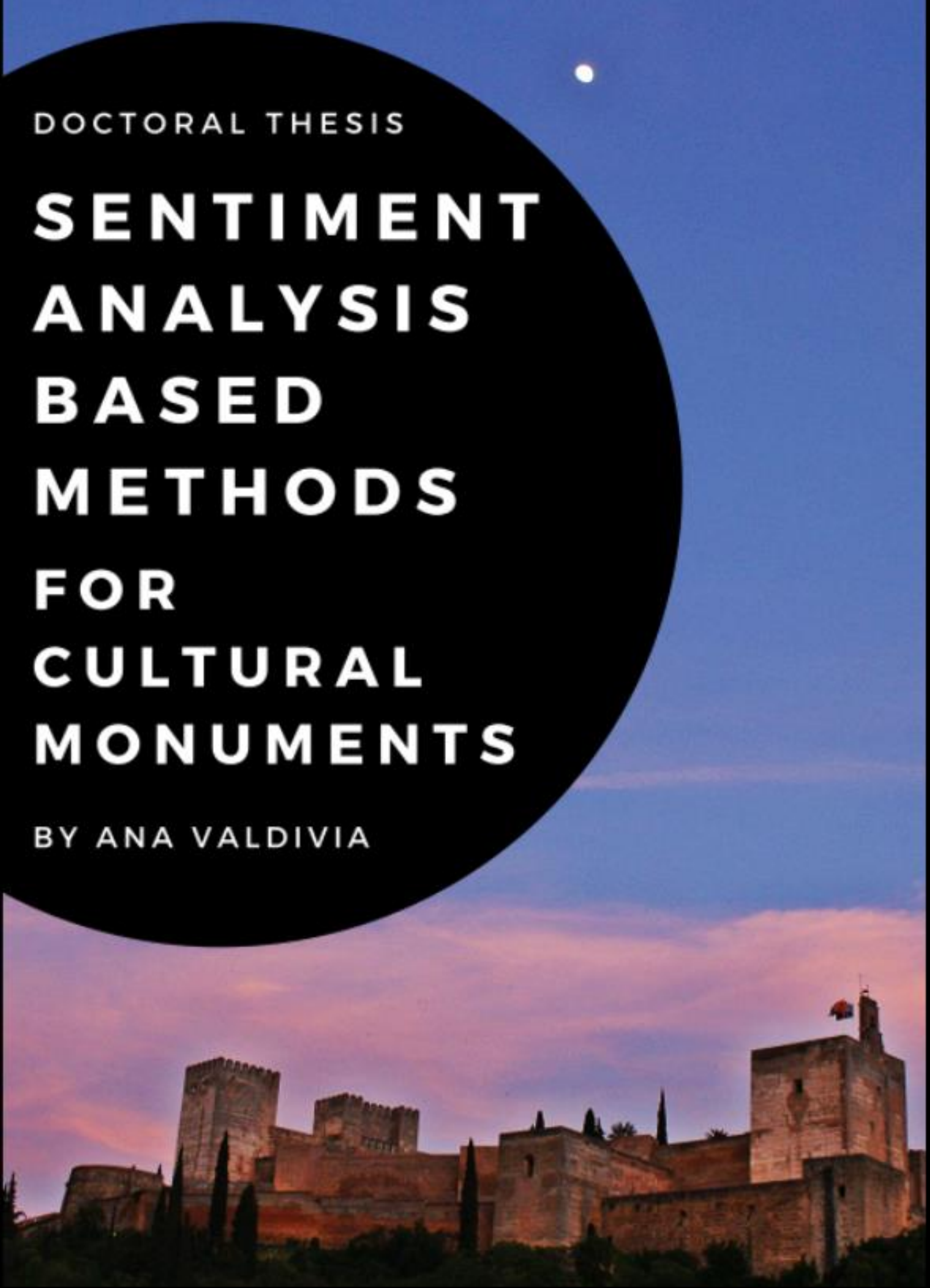
A. Valdivia, MV. Luzón, F. Herrera. **Sentiment analysis in tripadvisor**. IEEE Intelligent Systems 32 (4), 72-77 (2017)

- Status: **Published**.
- Impact Factor (JCR 2017): **2.596**



A. Valdivia, E. Hrabova, I. Chaturvedi, MV. Luzón, L. Troiano, E. Cambria, F. Herrera. **The inconsistencies on TripAdvisor reviews: a unified index between users and sentiment analysis methods**. Neurocomputing.

- Status: **Accepted**.
- Impact Factor (JCR 2017): **3.241**



RESULTS

- The problem of inconsistencies
- Neutrality Detection
- Opinion Summarization

UNIVERSIDAD DE GRANADA

Neutrality detection

Motivation

- ❑ Neutrality is widely considered as an ambiguous class in SA literature.
- ❑ Neutrality is also ignored in many SA applications.
- ❑ Low consensus on detecting neutral reviews.

Objective

- ❑ Enhance SA classification by detecting and filtering neutral reviews through consensus models.

Neutrality detection

RESULTS

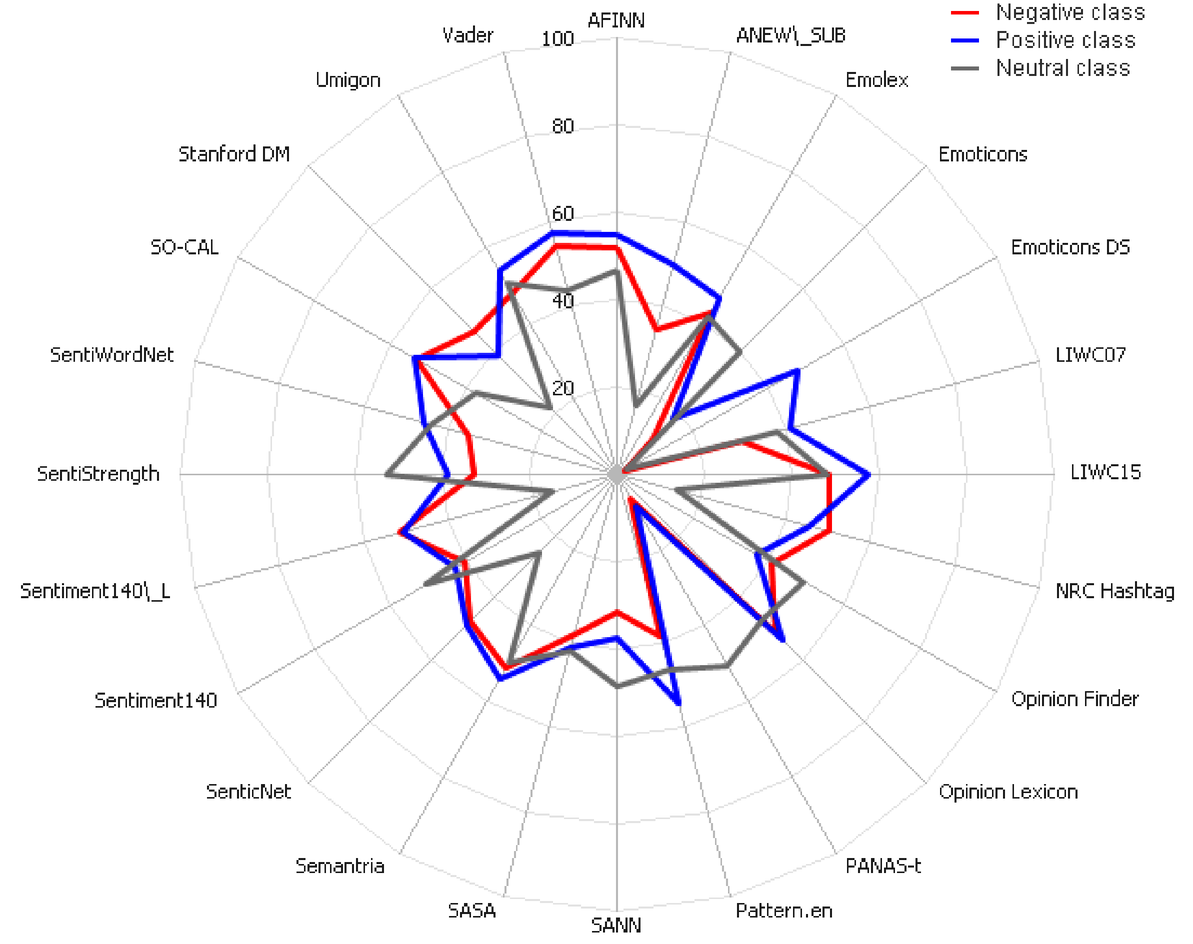


Figure 2 Average *F*1 score for each class. This figure presents the average *F*1 of positive and negative class and as we can see, methods use to achieve better prediction performance on positive messages.

Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A., & Benevenuto, F. (2016). **Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods.** *EPJ Data Science*, *5*(1), 1-29.

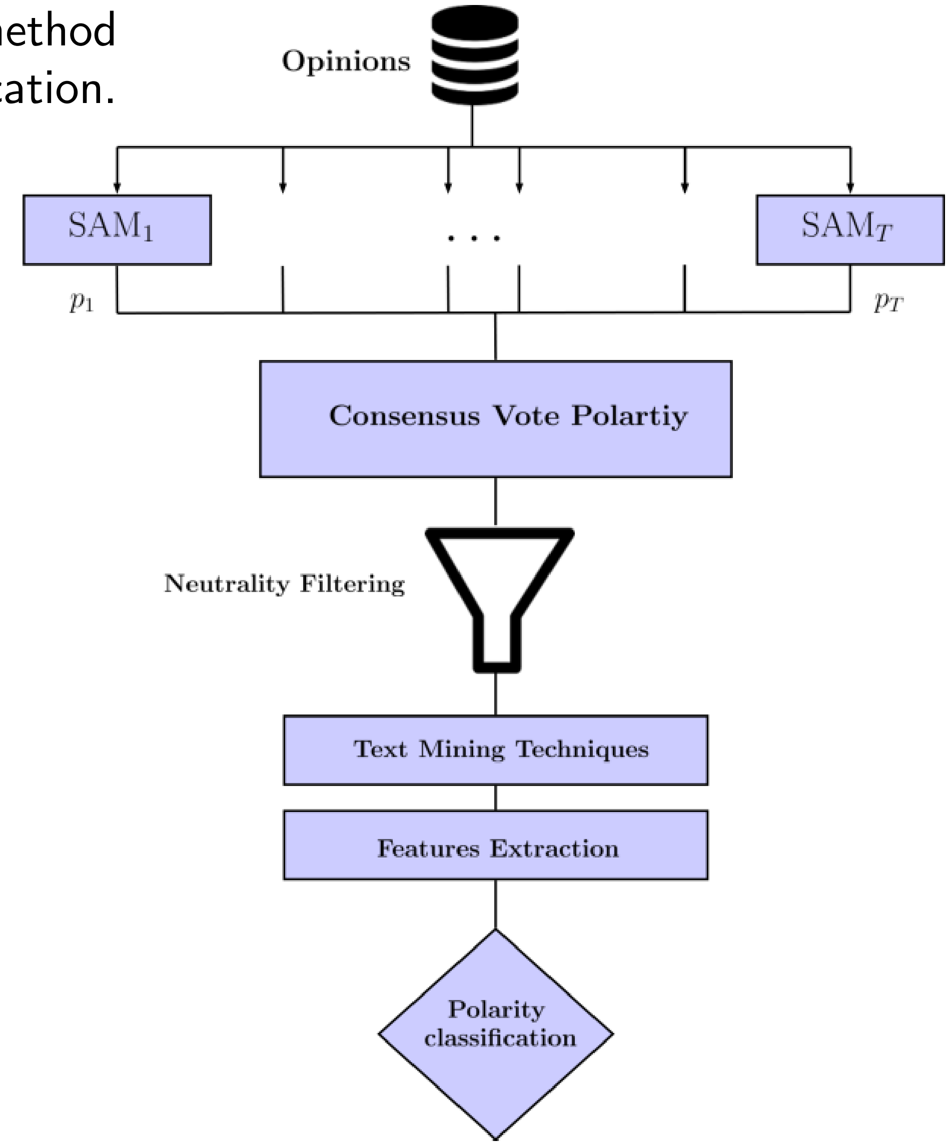
Neutrality detection

Table III: Number of Neutrality Consensus per Corpus

Corpus (500 reviews per Corpus)	AllAgree	AtLeastOneAgree
Amazon	3	404
ClintonTrump	9	433
Food	0	422
Cinema	3	377
Movies	0	357
RW	0	441
Ted	0	388
TA-Sagrada Familia	0	348
TA-Alhambra	0	369

Neutrality detection

Workflow of the proposed method for enhancing polarity classification.



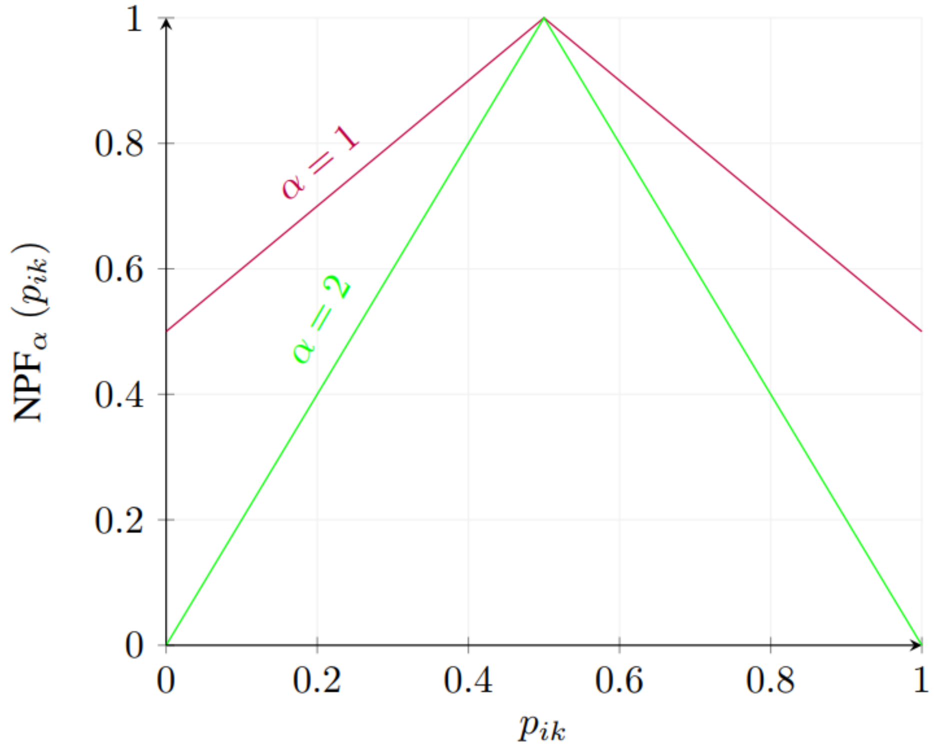
Neutrality detection

Neutrality Proximity Function (NPF)

The NPF is a function that measures the proximity of polarity p_{ik} to the neutrality, rising its absolute maximum when p_{ik} = 0.5.

We propose to use the following parametric function of NPF with $\alpha \in (0, 2]$:

$$\begin{aligned} NPF_\alpha : [0,1] &\rightarrow [0,1] \\ p_{ik} &\mapsto 1 - \alpha|p_{ik} - 0.5|, \alpha \in (0,2]. \end{aligned}$$



Representation of NPF for *α =1* and *α =2.*

Neutrality detection

8 aggregation functions

- **2 NPF-based aggregations**
  - ProN
  - ProNE
- **1 Average aggregation**
  - AvgN
- **5 Linguistic quantifiers-based aggregations**
  - Many As Possible
  - At Least Half
  - Most of
  - MinN
  - MaxN

Neutrality detection

Text Properties

Table 3: Summary of Quantitive Text Analysis of Datasets (Words and Sentences)

Corpus	NumWords	AVGNumWords	NumSentences	AVGNumSentences
Amazon	7,787	15.57	500	1.00
ClintonTrump	8,674	17.35	881	1.76
Food	40,775	81.55	2,512	5.02
Cinema	10,433	20.87	564	1.13
Movies	9,623	19.25	523	1.05
RW	34,871	69.74	2,337	4.67
Ted	8,971	17.94	502	1.00
TA-Sagrada Familia	30,520	61.04	2,033	4.07
TA-Alhambra	45,665	91.33	2,800	5.60

Neutrality detection

Number of neutral reviews by consensus

Table 4: Number of Neutral Instances of **Individual SAMs**, I=(0.4, 0.6)

Corpus	**Bing**	**CoreNLP**	**MC**	**Microsoft**	**SentiStr**	**VADER**
Amazon	115	105	186	144	251	193
ClintonTrump	277	221	115	93	228	130
Food	239	209	61	40	102	22
Cinema	283	34	92	52	152	110
Movies	124	42	124	86	174	150
RW	236	211	114	50	156	87
Ted	147	55	107	98	194	155
TA-Sagrada Familia	80	222	66	36	99	34
TA-Alhambra	99	239	54	22	84	16
Average	177.778	148.667	102.111	69.000	160.000	99.667

Table 5: Number of Neutral Instances of **Aggregation Models**, I=(0.4, 0.6)

Corpus	**MinN**	**MaxN**	**AvgN**	**ProN**	**ProNE**	**MAP-ProN**	**ALH-ProN**	**MO-ProN**
Amazon	3	404	221	245	287	183	182	174
ClintonTrump	9	433	207	228	298	163	225	149
Food	0	422	115	151	329	60	258	77
Cinema	3	377	143	167	239	103	156	106
Movies	0	357	199	220	289	175	168	157
RW	0	441	175	214	333	123	270	128
Ted	0	388	144	172	247	119	134	98
TA-Sagrada Familia	0	348	119	158	331	61	260	76
TA-Alhambra	0	369	96	132	307	41	281	61
Average	1.667	393.222	157.667	187.444	295.556	114.222	214.889	114.000

RESULTS

Results

Table 6: Test AUC for SVM models, **Individual SAMs**

Corpus	Bing	CoreNLP	MC	MSAzure	SentiStr	VADER	Average
Amazon	0.407	0.800	0.576	0.663	**0.851**	0.734	0.672
ClintonTrump	**0.878**	0.721	0.816	0.694	0.723	0.755	0.764
Food	**0.763**	0.647	0.500	0.613	0.469	0.413	0.567
Cinema	0.387	**0.628**	0.507	0.343	0.368	0.381	0.436
Movies	0.360	0.431	0.398	**0.546**	0.362	0.509	0.434
RW	0.697	**0.751**	0.505	0.559	0.682	0.596	0.632
Ted	**0.804**	0.260	0.263	0.341	0.215	0.346	0.371
TA-Sagrada Familia	**0.850**	0.656	0.714	0.651	0.690	0.724	0.714
TA-Alhambra	0.647	**0.895**	0.763	0.751	0.592	0.565	0.702
Average	**0.644**	0.643	0.560	0.573	0.550	0.558	0.588

Table 7: Test AUC for XGBOOST models, **Individual SAMs**

Corpus	Bing	CoreNLP	MC	MSAzure	SentiStr	VADER	Average
Amazon	0.400	0.265	0.434	0.613	**0.853**	0.690	0.542
ClintonTrump	**0.795**	0.735	0.767	0.719	0.711	0.731	0.743
Food	0.507	0.648	**0.692**	0.564	0.566	0.629	0.601
Cinema	0.485	0.387	0.518	0.367	0.438	**0.595**	0.465
Movies	0.529	0.471	0.437	**0.574**	0.415	0.409	0.472
RW	0.291	**0.710**	0.579	0.498	0.498	0.418	0.499
Ted	**0.366**	0.318	0.302	0.296	0.252	0.358	0.315
TA-Sagrada Familia	0.529	0.643	0.439	0.622	0.653	**0.746**	0.605
TA-Alhambra	0.678	0.870	**0.916**	0.592	0.501	0.541	0.683
Average	0.509	0.561	0.565	0.538	0.543	**0.569**	0.547

Results

Table 8: Test AUC for SVM models, **Aggregation Models**

Corpus	**MinN**	**AvgN**	**ProN**	**ProNE**	**MAP-ProN**	**ALH-ProN**	**MO-ProN**
Amazon	0.302	0.795	**0.853**	0.752	0.633	0.730	0.757
ClintonTrump	0.461	0.733	**0.870**	0.620	0.683	0.790	0.824
Food	0.456	0.407	0.302	0.500	**0.613**	0.609	0.609
Cinema	0.440	0.473	0.218	**0.740**	0.612	0.719	0.558
Movies	0.638	0.396	0.429	0.645	**0.645**	0.532	0.524
RW	0.349	0.375	0.631	0.652	0.713	**0.737**	0.530
Ted	0.305	0.215	0.230	0.286	0.713	0.708	**0.757**
TA-Sagrada Familia	0.535	**0.891**	0.693	0.693	0.875	0.776	0.629
TA-Alhambra	0.819	0.824	**0.941**	0.941	0.507	0.845	0.767
Average	0.478	0.568	0.574	0.648	0.666	**0.716**	0.662

Table 9: Test AUC for XGBOOST models, **Aggregation Models**

Corpus	**MinN**	**AvgN**	**ProN**	**ProNE**	**MAP-ProN**	**ALH-ProN**	**MO-ProN**
Amazon	0.336	0.776	**0.819**	0.730	0.600	0.767	0.737
ClintonTrump	0.372	0.671	**0.827**	0.415	0.648	0.724	0.656
Food	0.671	0.234	0.344	0.324	**0.687**	0.637	0.538
Cinema	0.444	0.425	0.371	0.500	**0.657**	0.448	0.514
Movies	0.435	**0.598**	0.471	0.427	0.568	0.434	0.440
RW	0.315	0.375	0.640	0.708	0.691	0.704	**0.717**
Ted	0.346	0.215	0.264	0.323	0.701	**0.810**	0.686
TA-Sagrada Familia	0.499	**0.859**	0.667	0.667	0.721	0.756	0.551
TA-Alhambra	0.768	0.792	**0.932**	0.932	0.606	0.743	0.652
Average	0.465	0.549	0.593	0.559	0.653	**0.669**	0.610

RESULTS

Results SVM

	Best SAM	Best Aggregation
Amazon	0.851	**0.853**
ClintonTrump	**0.878**	0.870
Food	**0.763**	0.613
Cinema	0.628	**0.740**
Movies	0.546	**0.645**
RW	**0.751**	0.737
Ted	**0.804**	0.757
TA-SagradaFamilia	0.850	**0.891**
TA-Alhambra	0.895	**0.941**
Average	0.644	**0.716**

RESULTS

Results XGBOOST

	Best SAM	Best Aggregation
Amazon	**0.853**	0.819
ClintonTrump	0.795	**0.827**
Food	**0.795**	0.687
Cinema	0.595	**0.657**
Movies	0.574	**0.598**
RW	0.710	**0.717**
Ted	0.366	**0.810**
TA-SagradaFamilia	0.746	**0.859**
TA-Alhambra	0.916	**0.932**
Average	0.569	**0.669**

Neutrality detection

R E S U L T S

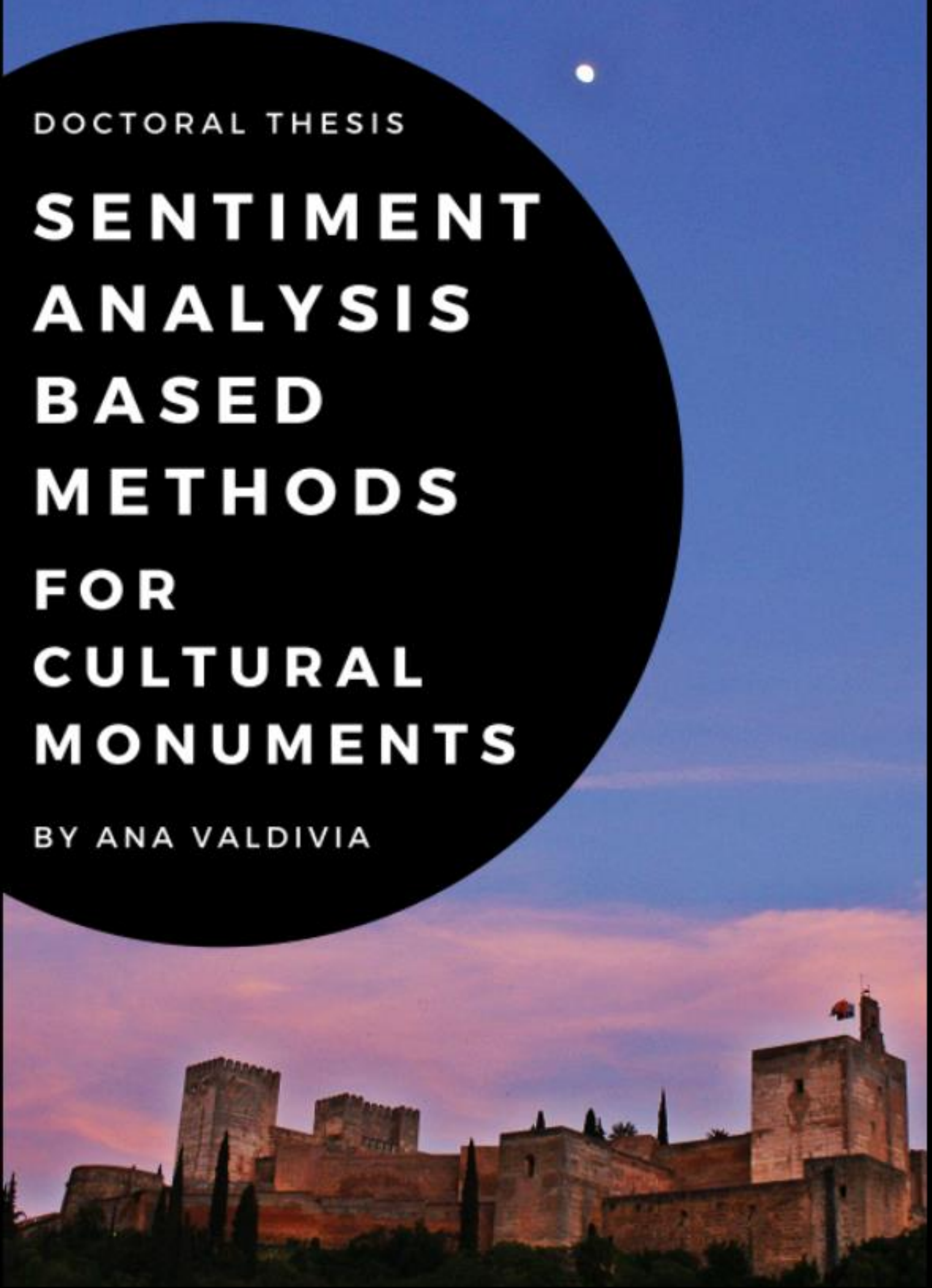
Conclusions

- This study have shown that **detecting neutrality** based on a consensus **improves classification precision**.
- The **ALH-ProN** model gets **the best** results on average.
- The **aggregation for detecting neutrality is** also **positive for neutrality detection** via polarity aggregation.

Neutrality detection

A. Valdivia, MV. Luzón, E. Cambria, F. Herrera. **Consensus vote models for detecting and filtering neutrality in sentiment analysis**. Information Fusion 44 126-135 (2018).

- Status: **Published**.
- Impact Factor (JCR 2017): **6.639**



RESULTS

- The problem of inconsistencies
- Neutrality Detection
- Opinion Summarization

UNIVERSIDAD DE GRANADA

R
E
S
U
L
T
S

Opinion Summarization

Motivation

- ❑ Need of summarization a large quantity of reviews.
- ❑ Improve the decision-making of cultural heritage operators.

Objective

- ❑ To depict negative opinions from cultural monuments in order to detect those features that need to be enhanced.
- ❑ We propose a novel methodology that extract aspects and summarize reviews in a rule-based visualization.

RESULTS

Opinion Summarization

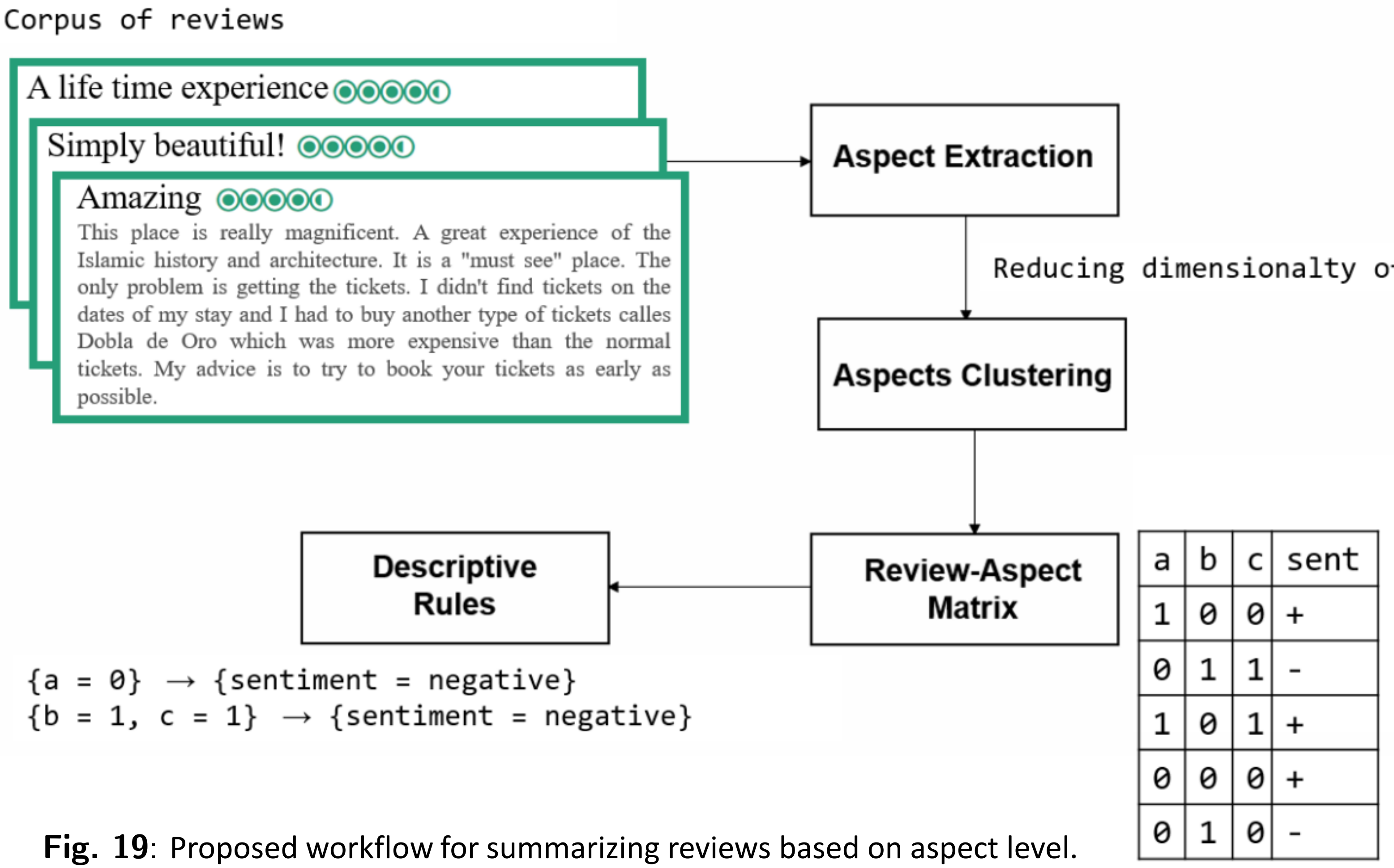
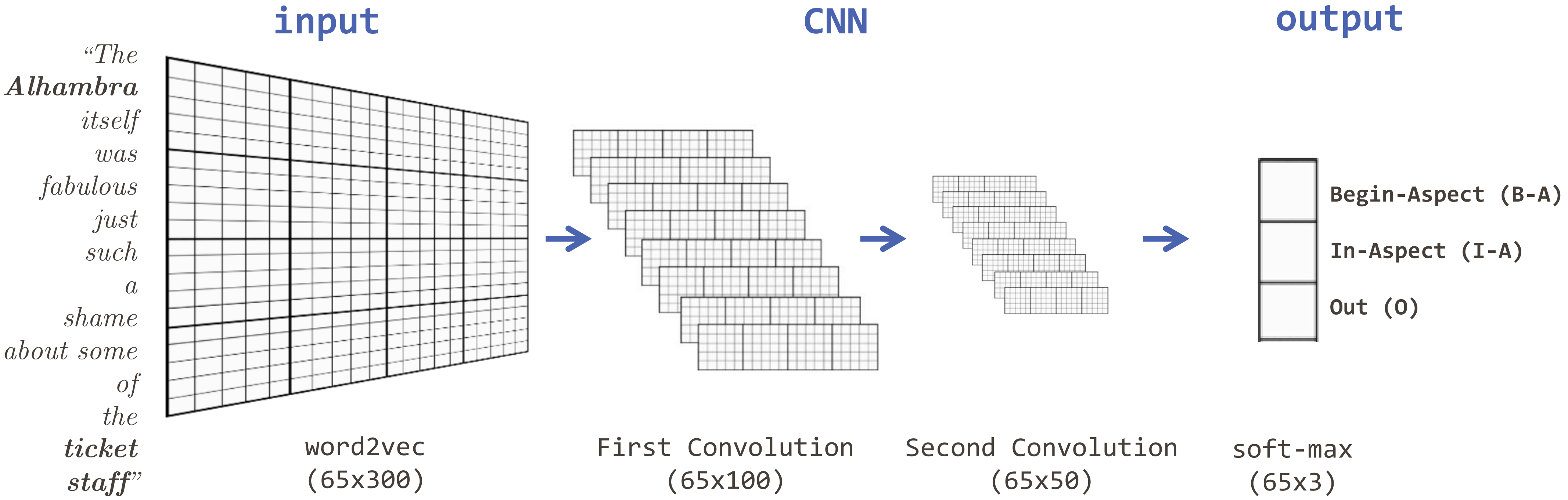


Fig. 19: Proposed workflow for summarizing reviews based on aspect level.

I. Aspect extraction

Aspect Based Sentiment Analysis (ABSA) focuses on **extracting aspects** and **entities** that have been evaluated in the reviews and **gives a more detailed information** about the purpose of the **opinion**.



RESULTS

II. Aspect Clustering

The great **diversity of language** implies:

1. **High dimensionality** of aspects.
2. Aspects with **similar meanings** have **different representations**.

ticket $\longrightarrow$ {*onsite ticket office, senior ticket, ticket area, garden ticket, ticket check points, ticket office, entry ticket, ticket seller, service ticket, machine ticket, ticket tip, ticket staff, ticket box, ticket master, ticket price, ticket process, ticket desks, ...*}.

word embeddings + k-means

Given a set of elements $\{w_1, \ldots, w_n\}$, *k-means aims to cluster* n *observations in* k *clusters* ($\{C_1, \ldots, C_k\}$), *minimizing the function:*

$$\arg\min_{C} \sum_{i=1}^{k} \sum_{w \in C_i} \|w - \mu_i\|^2, \tag{1}$$

where μ_i *is the mean of points in* C_i.

III. Descriptive Rules

Subgroup Discovery. It aims at discovering interesting rules fixing a class label.

Aspect – Review matrix

Staff	Wheel Chair	Ticket	sentiment
1	0	0	negative
0	1	0	negative
0	1	0	negative
1	0	0	positive
1	0	1	positive
0	1	0	negative
1	0	0	positive

III. Descriptive Rules

Subgroup Discovery. It aims at discovering interesting rules fixing a class label.

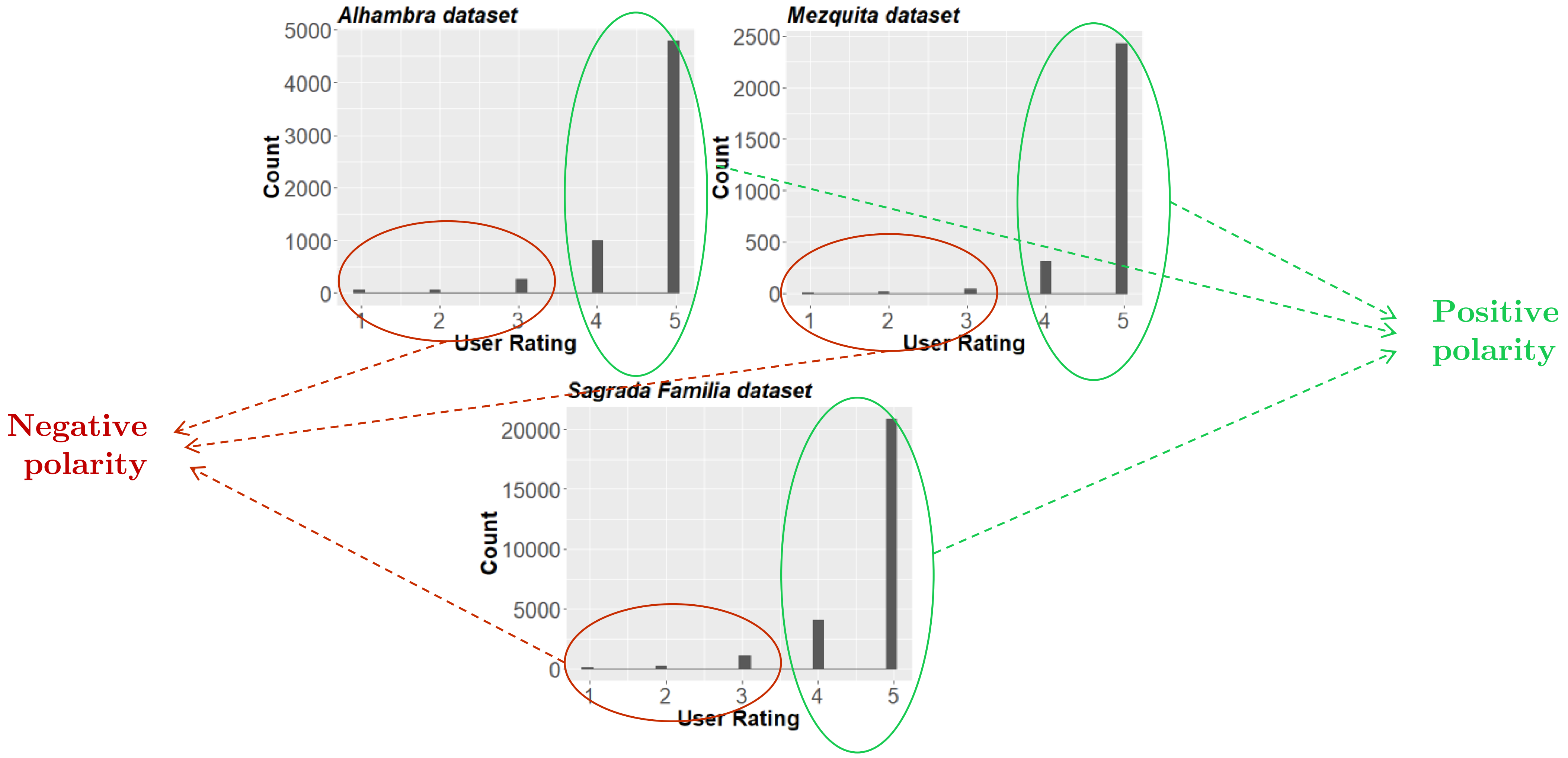
$$\mathrm{R}_1: \{\texttt{Staff = 1}\} \longrightarrow \{\texttt{Sentiment = negative}\},$$

$$\mathrm{R}_2: \{\texttt{Wheel Chair Accessible = 1, Staff = 0}\} \longrightarrow \{\texttt{Sentiment = positive}\},$$

$$\mathrm{R}_3: \{\texttt{Ticket = 1, Staff = 1}\} \longrightarrow \{\texttt{Sentiment = negative}\}.$$

RESULTS

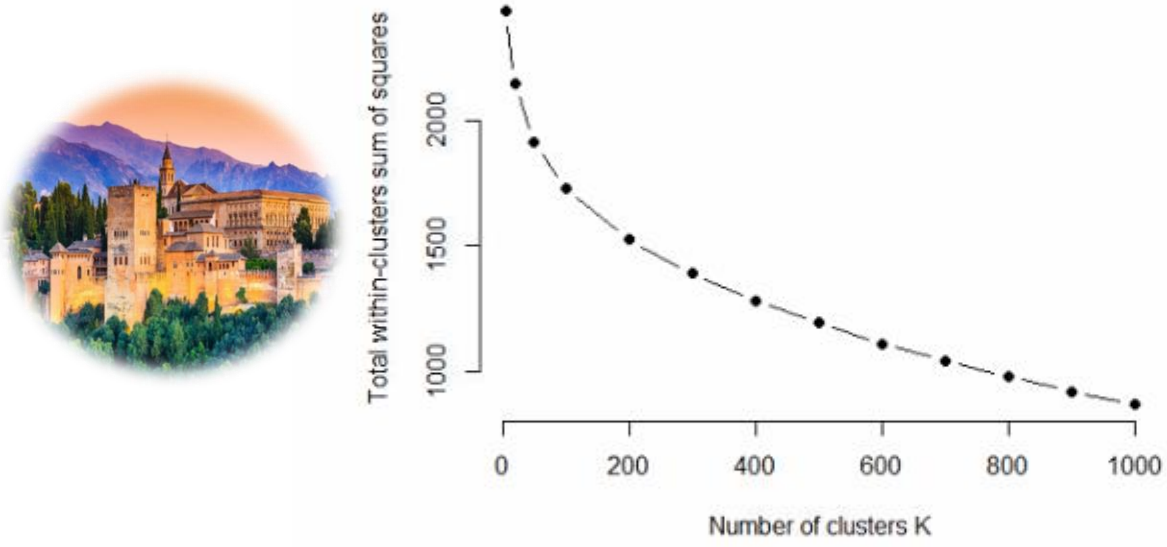
User Rating Distribution

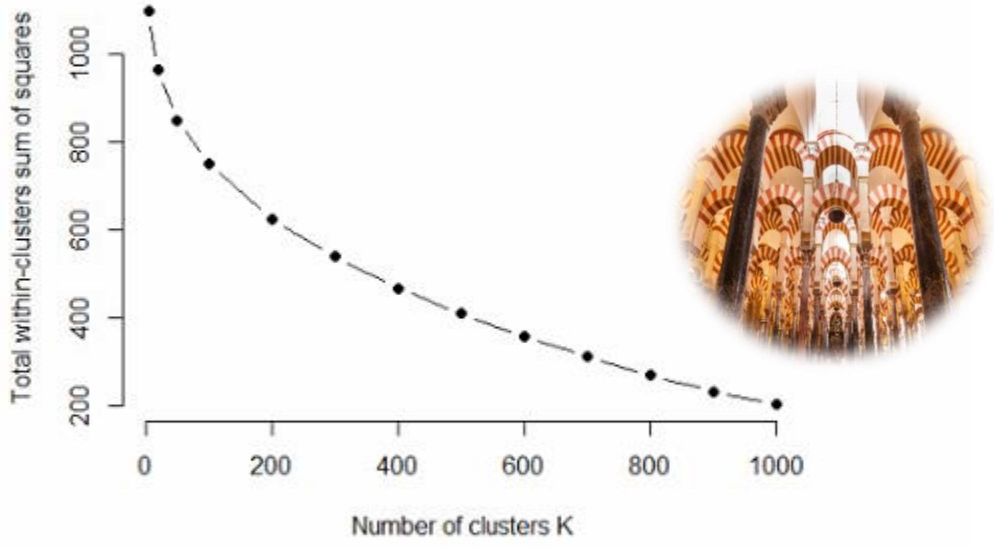


Distribution of User Rating in TripAdvisor reviews.

R E S U L T S

Results





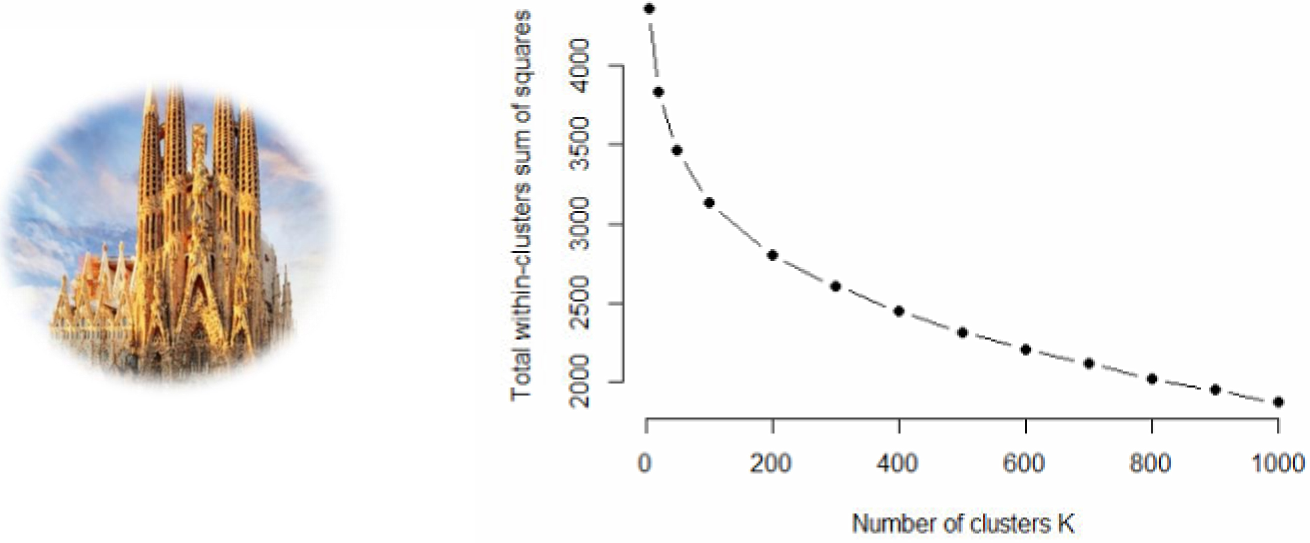


Fig. 20: The elbow method.

Results

Monument	Cluster Label	Cluster Content
Alhambra	BDA	staff, staff member, local staff, hotel staff, map staff, ground staff, male staff
Alhambra	GG	tickets, individual tickets, garden tickets, access tickets, internet tickets
Alhambra	BFD	gardens, garden, generalife gardens, gardens water features, garden ticket, beauty of the gardens, general life gardens, main garden, generalife garden
Mezquita	BJB	ceiling, floors, marble ceiling, walls, vaulted ceiling, marble floor, roof
Mezquita	EG	guard, security guard, security guard berating
Mezquita	GG	audio guide, audio guide available, audio guides, map audio guide, auto lingual guide, audio guide facility
Sagrada Familia	GI	lift, lift up amp, lift amp lift ride 65m, lift elevator, tower lift, lift down, towers lift up, lift down wait, lift service
Sagrada Familia	BAI	ticket online, tickets on-line online tickets, tickets online, entrance tickets online, online ticket, entrance ticket online, prepurchased online tickets, book your ticket online
Sagrada Familia	BFJ	shop, souvenir shop, gauds shop, bookshop, citys souvenir shops

Table 4 Examples of aspects grouped into clusters, with $k = 200$.

RESULTS

Results

Aspect	Rule	Cov	Sup	Conf	WRAcc
staff	{ BDA = 1 } $\longrightarrow$ { negative }	0.03	< 0.01	0.28	< 0.01
guard	{ BDG = 1 } $\longrightarrow$ { negative }	< 0.01	< 0.01	0.38	< 0.01
cashier	{ HH = 1 } $\longrightarrow$ { negative }	< 0.01	< 0.01	0.27	< 0.01
queue	{ BDI = 1 } $\longrightarrow$ { negative }	< 0.01	< 0.01	0.15	< 0.01
staff, price	{ BDA = 1, BGI = 1 } $\longrightarrow$ { negative }	< 0.01	< 0.01	0.48	< 0.01

Table 6 Most relevant rules of the Alhambra monument.

Aspect	Rule	Cov	Sup	Conf	WRAcc
mosque	{ BHF = 1 } $\longrightarrow$ { negative }	0.18	< 0.01	0.02	0
garden	{ BHB = 1 } $\longrightarrow$ { negative }	0.06	< 0.01	0.03	< 0.01
architecture	{ BEJ = 1 } $\longrightarrow$ { negative }	0.17	< 0.01	0.02	0
place	{ DJ = 1 } $\longrightarrow$ { negative }	0.06	< 0.01	0.03	0
arches	{ BIC = 1 } $\longrightarrow$ { negative }	0.07	< 0.01	0.02	0

Table 7 Most relevant rules of the Mezquita monument.

R E S U L T S

Results

Aspect	Rule	Cov	Sup	Conf	WRAcc
ceiling	{ `CJ = 1` } $\longrightarrow$ { `negative` }	< 0.01	< 0.01	0.1	< 0.01
natural	{ `HG = 1` } $\longrightarrow$ { `negative` }	0.04	< 0.01	0.07	< 0.01
entry	{ `CI = 1` } $\longrightarrow$ { `negative` }	0.01	< 0.01	0.07	< 0.01
queue	{ `BCD = 1` } $\longrightarrow$ { `negative` }	< 0.01	< 0.01	0.05	0
sagrada familia	{ `BEJ = 1` } $\longrightarrow$ { `negative` }	0.01	< 0.01	0.07	< 0.01

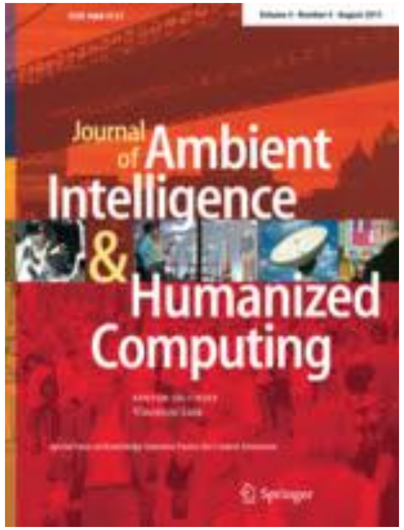
Table 8 Most relevant rules of the Sagrada Familia monument.

Opinion Summarization

Conclusions

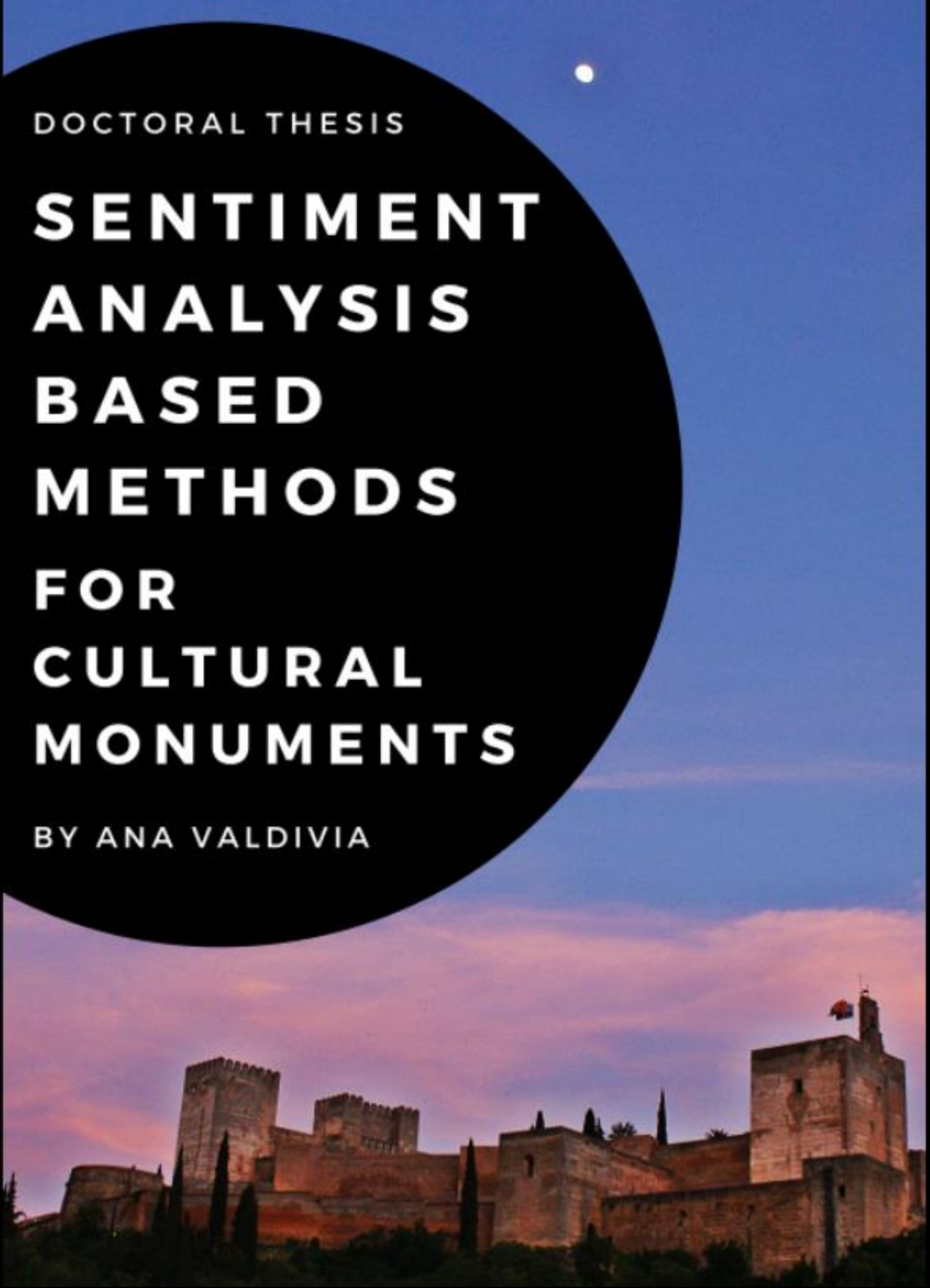
- The results shown that **the proposed methodology** is **effective** for **describing reviews**.
- Data sparsity and the unbalanced dataset **affect** to **measures of generality**.
- Using the **polarity of the review for all the aspects** may led to a **misinterpretation** of the text.

Opinion Summarization



Valdivia, A., Martínez-Cámara, E., Chaturvedi, I., Luzón, M. V., Cambria, E., Ong, Y. S., & Herrera, F. (2018). *What do people think about this monument?* **Understanding negative reviews via deep learning, clustering and descriptive rules.** *Journal of Ambient Intelligence and Humanized Computing*, 1-14.

- Status: **Published**.
- Impact Factor (JCR 2017): **1.423**



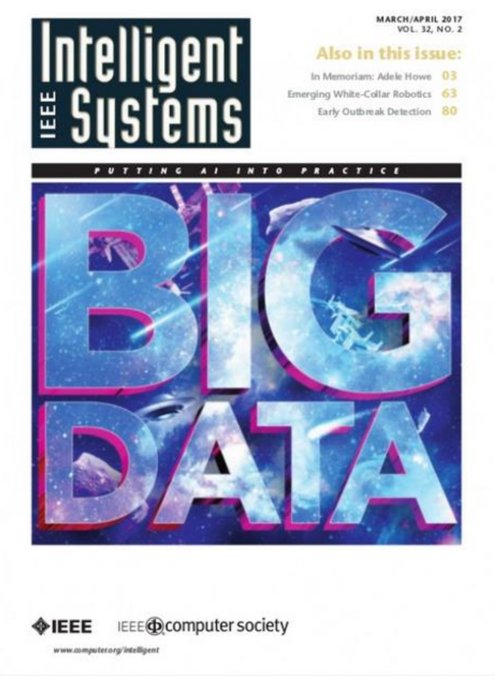
CONCLUSION

UNIVERSIDAD DE GRANADA

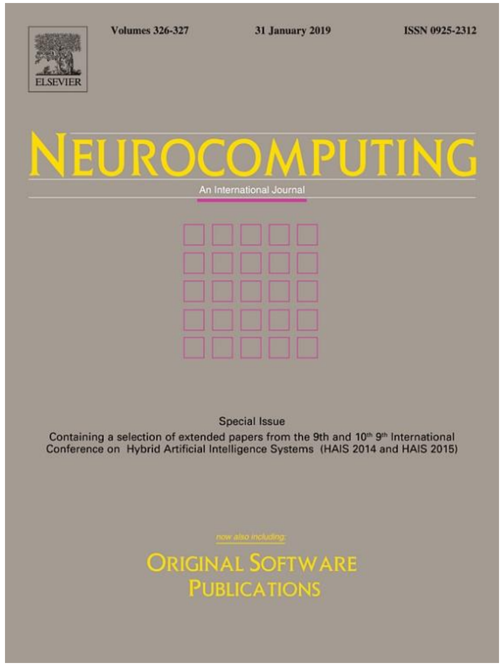
Conclusions

- **This thesis has presented a broad study about SAMs applied to cultural monument reviews using social media data.**
- **INCONSISTENCIES.** The Polarity Aggregation Method improves the incosistencies between SAMs and User polarities.
- **NEUTRALITY.** Neutral reviews are a key for enhancing polarity classification.
- **OPINION SUMMARIZATION.** Rule-based approaches are efficient for summarizing reviews.

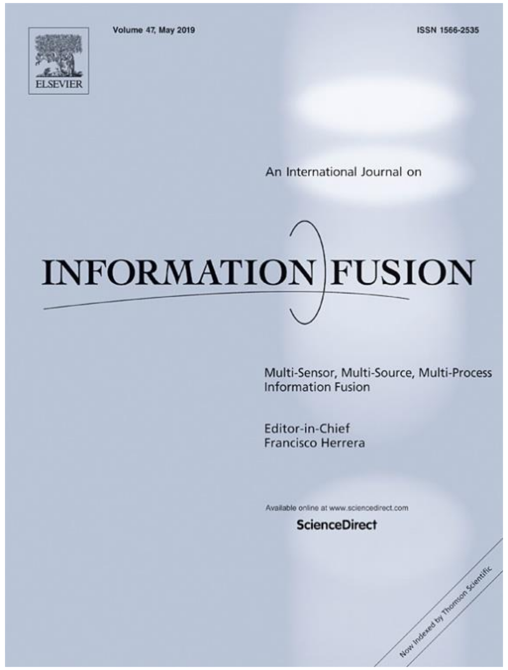
Results of this thesis



Sentiment Analysis in TripAdvisor (Q2)



Inconsistencies in TripAdvisor (Q1)



Neutrality Detection (Q1)

Opinion Summarization (Q3)

3 International Conferences

HAIS 2016 Hybrid Artificial Intelligence Systems

FUZZ-IEEE 2017 NAPLES, ITALY

CAEPIA 2018 Granada
XVIII Conference of the Spanish Association for Artificial Intelligence

Future work

- Evaluation studies between Sentiment Analysis and Surveys for cultural monuments.
- Creation of corpus for cultural heritage reviews.
- Extension of SA-based methods for other languages.
- Adaptation domain problem.

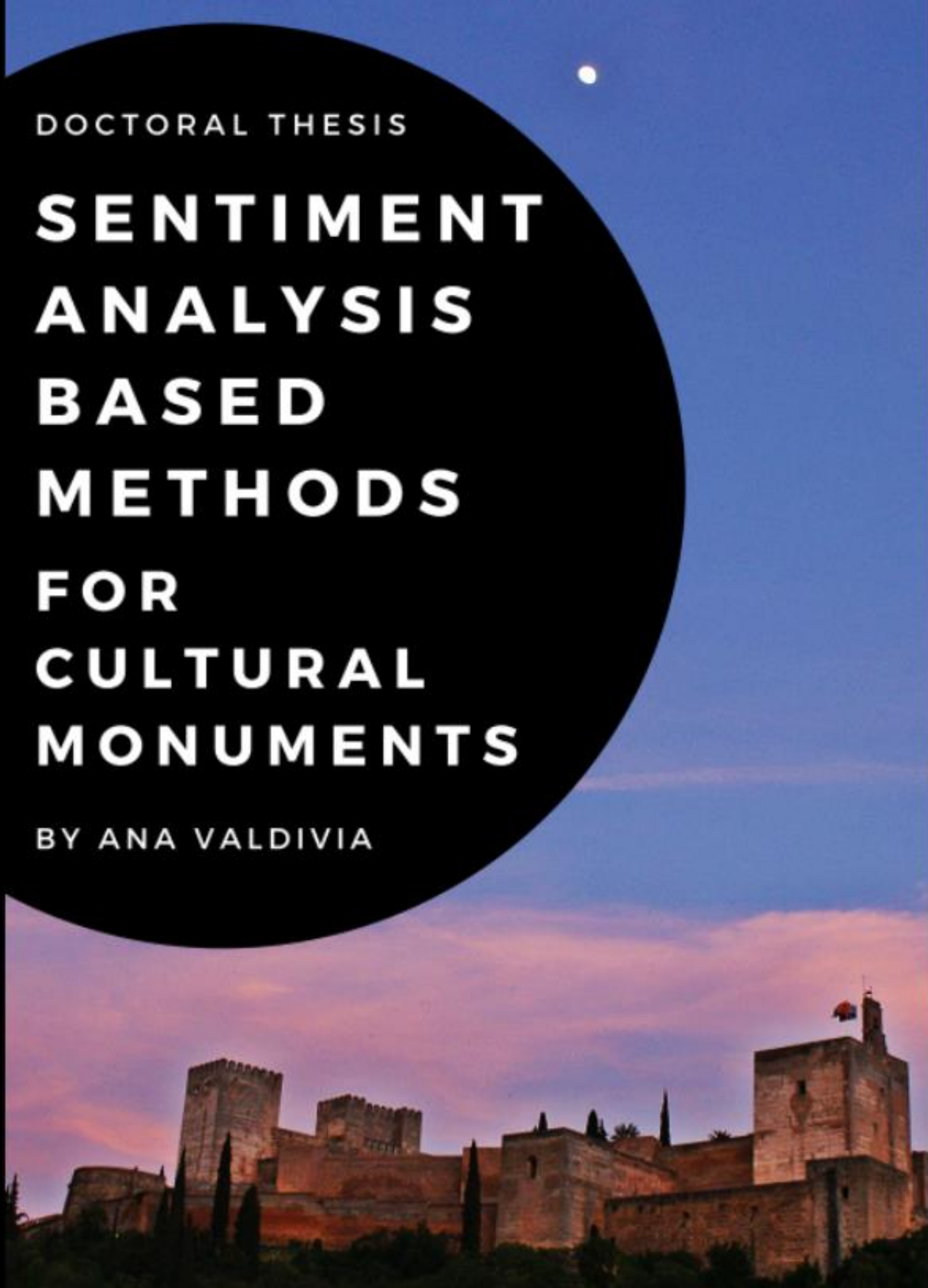
PhD Dissertation

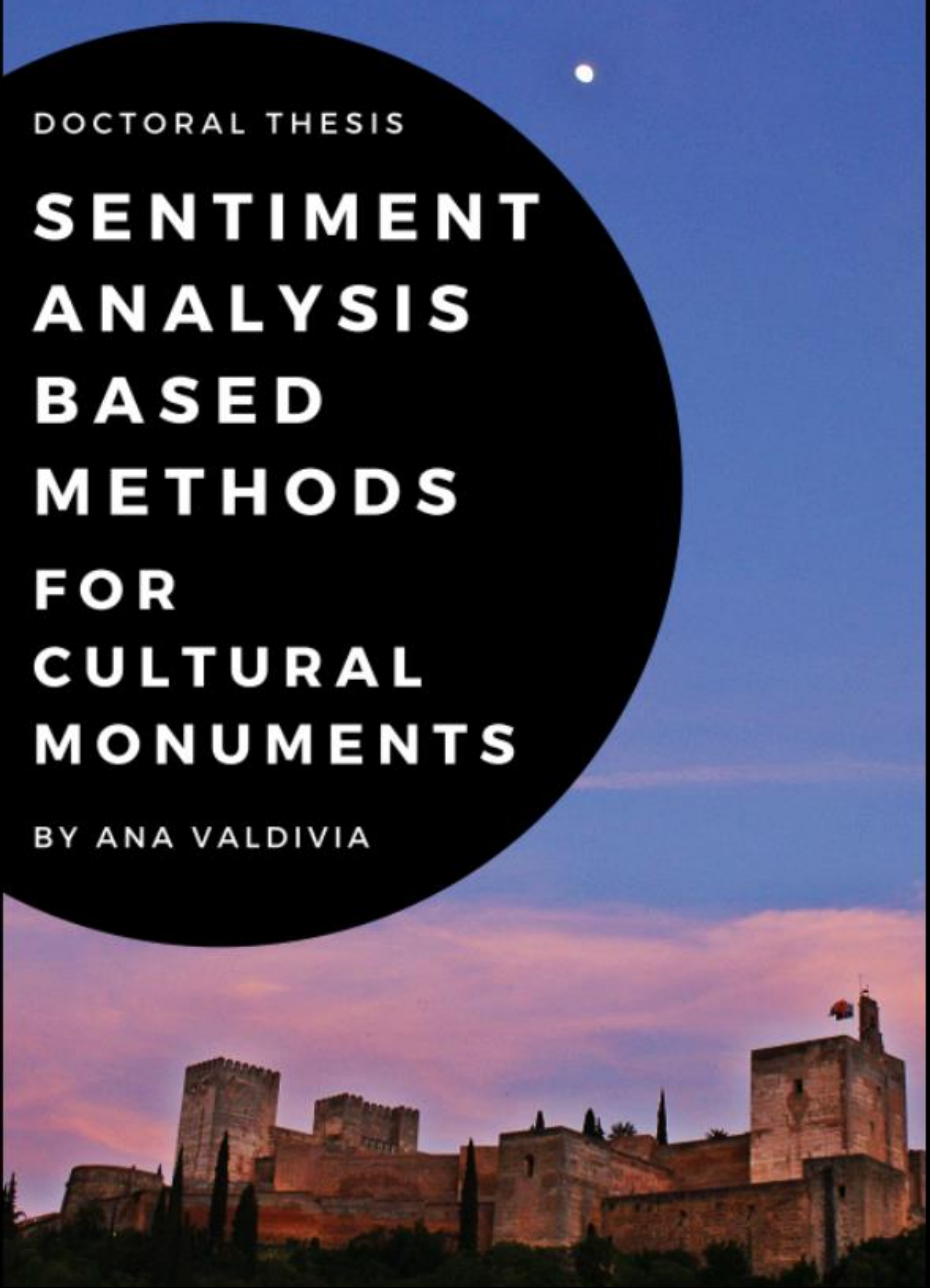
Supervisors:

Francisco Herrera & M. Victoria Luzón

1st February, 2019

UNIVERSIDAD DE GRANADA





APPENDIX

UNIVERSIDAD DE GRANADA

Examples

Table 1
Main characteristics of the SAMs used in our evaluation.

Name	Description	Approach	Output	Domain	Genre
Azure	Supervised ML method that works at document level. It was trained with a large corpus of reviews.	ML	$[0, 1] \in \mathbb{R}$	General	Reviews
Bing [17]	SO method that works at document level. It uses an opinion lexicon, and the polarity value of the document is the average polarity value of its sentences.	SO	$\{-1, 0, 1\}$	General	NA
CoreNLP [26]	Supervised ML method that works at document/sentence level. It is based on a Recursive Neural Tensor Network (RNTN), and classifies 5 intensity levels of opinion.	ML	$\{0, 1, 2, 3, 4\}$	Review	Movie reviews
MeaningCloud	SO method that works at document level. It is grounded in the pos-tags of the words and the use of an opinion lexicon, and it classifies 5 intensity levels of opinion.	SO	$\{0, 1, 2, 3, 4\}$	NA	NA
SentiStrength [43]	Supervised ML method grounded in features built upon opinion lexicons that work at document/sentence level. The output is the positive (Pos) and the negative (Neg) value of the input text.	SO & ML	Neg: $[-5, -1] \in \mathbb{Z}$; Pos: $[1, 5] \in \mathbb{Z}$	General	Social media
Syuzhet	SO method that works at document/sentence level. It uses an opinion lexicon. The document polarity value is the average of the polarity value of its sentences.	SO	$\{-1, 0, 1\}$	General	Novels
Vader [18]	SO method that works at document level and uses an opinion lexicon.	SO	$[-1, 1] \in \mathbb{R}$	General	Micro-blogging

López, M., Valdivia, A., Martínez-Cámara, E., Luzón, M. V., & Herrera, F. (2018). **E2SAM: Evolutionary Ensemble of Sentiment Analysis Methods for Domain Adaptation.** *Information Sciences*.

P R E L I M I N A R I E S

Opinion

Opinion

- An opinion is a 5-tuple contain the target of the opinion (entity), the attribute of the target at which the opinion is directed, the polarity, the opinion holder and the date when the opinion was emitted:

$$(e_i,\ a_{ij},\ s_{iijkl},\ h_k,\ t_l)$$

where:

- e_i is the i-th entity,
- a_{ij} is the j-th attribute of the entity,
- s_{iijkl} is the polarity of the opinion,
- h_k is the k-th opinion holder,
- t_l is the l-th time when the opinion was emitted.

Example

Lucy, 15/12/2018

I went there with some friends to work on a project and we loved it! We all ordered hot chocolate drinks and thought they were absolutely delicious! Very cool atmosphere with jazz music, but it is very dark if you wish to work with your laptop in there. The waitress was also very kind!

PRELIMINARIES

Opinion

Opinion

- An opinion is a 5-tuple contain the target of the opinion (entity), the attribute of the target at which the opinion is directed, the polarity, the opinion holder and the date when the opinion was emitted:

$$(e_i,\ a_{ij},\ s_{iijkl},\ h_k,\ t_l)$$

where:

- e_i is the i-th entity,
- a_{ij} is the j-th attribute of the entity,
- s_{iijkl} is the polarity of the opinion,
- h_k is the k-th opinion holder,
- t_l is the l-th time when the opinion was emitted.

Example

Lucy, 15/12/2018

I went there with some friends to work on a project and we loved it! We all ordered hot chocolate drinks and thought they were absolutely delicious! Very cool atmosphere with jazz music, but it is very dark if you wish to work with your laptop in there. The waitress was also very kind!

PRELIMINARIES

Opinion

Opinion

- An opinion is a 5-tuple contain the target of the opinion (entity), the attribute of the target at which the opinion is directed, the polarity, the opinion holder and the date when the opinion was emitted:

$$(e_i,\ a_{ij},\ s_{iijkl},\ h_k,\ t_l)$$

where:

- e_i is the i-th entity,
- a_{ij} is the j-th attribute of the entity,
- s_{iijkl} is the polarity of the opinion,
- h_k is the k-th opinion holder,
- t_l is the l-th time when the opinion was emitted.

Example

5-Tuple	Entity	Aspect	Sentiment	Opinion holder	Time
$(e_1, a_{11}, s_{1111}, h_1, t_1)$	e_1 = coffee shop	a_{11} = drinks	s_{1211} = positive	h_1 = Lucy	t_1 = 15/12/2018
$(e_1, a_{12}, s_{1211}, h_1, t_1)$	e_1 = coffee shop	a_{12} = drinks	s_{1211} = positive	h_1 = Lucy	t_1 = 15/12/2018
$(e_1, a_{13}, s_{1211}, h_1, t_1)$	e_1 = coffee shop	a_{13} = place	s_{1311} = brightness	h_1 = Lucy	t_1 = 15/12/2018
$(e_1, a_{14}, s_{1411}, h_1, t_1)$	e_1 = coffee shop	a_{14} = service	s_{1411} = positive	h_1 = Lucy	t_1 = 15/12/2018

Examples

Table 1
Main characteristics of the SAMs used in our evaluation.

Name	Description	Approach	Output	Domain	Genre
Azure	Supervised ML method that works at document level. It was trained with a large corpus of reviews.	ML	$[0, 1] \in \mathbb{R}$	General	Reviews
Bing [17]	SO method that works at document level. It uses an opinion lexicon, and the polarity value of the document is the average polarity value of its sentences.	SO	$\{-1, 0, 1\}$	General	NA
CoreNLP [26]	Supervised ML method that works at document/sentence level. It is based on a Recursive Neural Tensor Network (RNTN), and classifies 5 intensity levels of opinion.	ML	$\{0, 1, 2, 3, 4\}$	Review	Movie reviews
MeaningCloud	SO method that works at document level. It is grounded in the pos-tags of the words and the use of an opinion lexicon, and it classifies 5 intensity levels of opinion.	SO	$\{0, 1, 2, 3, 4\}$	NA	NA
SentiStrength [43]	Supervised ML method grounded in features built upon opinion lexicons that work at document/sentence level. The output is the positive (Pos) and the negative (Neg) value of the input text.	SO & ML	Neg: $[-5, -1] \in \mathbb{Z}$; Pos: $[1, 5] \in \mathbb{Z}$	General	Social media
Syuzhet	SO method that works at document/sentence level. It uses an opinion lexicon. The document polarity value is the average of the polarity value of its sentences.	SO	$\{-1, 0, 1\}$	General	Novels
Vader [18]	SO method that works at document level and uses an opinion lexicon.	SO	$[-1, 1] \in \mathbb{R}$	General	Micro-blogging

López, M., Valdivia, A., Martínez-Cámara, E., Luzón, M. V., & Herrera, F. (2018). **E2SAM: Evolutionary Ensemble of Sentiment Analysis Methods for Domain Adaptation.** *Information Sciences*.

Neutrality detection

2 NPF

Definition 2 *Pro-Neutrality Weight Based Model (ProN)*. The *ProN*, Φ_{ProN}, is an aggregation model for sentiment polarities which defines the weighting vector W guided by the NPF$_{\alpha=1}$, i.e., $w_{ik} = \text{NPF}_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$ and it is expressed such that:

$$\begin{aligned}\Phi_{ProN} &: [0,1]^T \rightarrow [0,1] \\ (p_{1k}, \cdots, p_{Tk}) &\mapsto \sum_{i=1}^{T} \frac{w_{ik}}{\sum_{i=1}^{T} w_{ik}} p_{ik}.\end{aligned}$$

Definition 3 *Pro-Neutrality Extreme Weight Based Model (ProNE)*. The *ProNE*, Φ_{ProNE}, is an aggregation model for sentiment polarities which defines the weighting vector W guided by the NPF$_{\alpha=2}$, i.e., $w_{ik} = \text{NPF}_{\alpha=2}(p_{ik})$ $= 1 - 2|p_{ik} - 0.5|$ and it is expressed such that:

$$\begin{aligned}\Phi_{ProNE} &: [0,1]^k \rightarrow [0,1] \\ (p_{1k}, \cdots, p_{Tk}) &\mapsto \sum_{h=1}^{T} \frac{w_{ik}}{\sum_{h=1}^{T} w_{ik}} p_{ik}.\end{aligned}$$

P U B L I C A T I O N S

1 AVG

Definition 4 *Average Based Model (AVG).* The *AVG* is an aggregation model for sentiment polarities which defines the weighting vector by $W = \frac{1}{T}$ and it is expressed such that:

$$\Phi_{AVG} \colon [0,1]^T \to [0,1]$$
$$(p_{1k}, \cdots, p_{Tk}) \mapsto \frac{1}{T} \sum_{h=1}^{T} p_{ik}.$$

Note that this model is equivalent to the *arithmetic mean* over the k polarities.

P U B L I C A T I O N S

5 LINGUISTIC QUANTIFIERS

Definition 5 *OWA Yager (1988); Chiclana et al. (2007).* An *OWA operator* of dimension n is a mapping $\phi : \mathbb{R}^n \to \mathbb{R}$ that has an associated weighting vector W such that $w_i \in [0,1]$, $\sum_{i=1}^{n} w_i = 1$, and is defined to aggregate a list of values $\{p_1, \ldots, p_n\}$ following this expression:

$$\phi(p_1, \ldots, p_n) = \sum_{i=1}^{n} w_i p_{\sigma(i)},$$

being $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ a permutation such that $p_{\sigma(i)} \geq p_{\sigma(i+1)}, \forall i = 1, \ldots, n-1$.

Definition 6 *IOWA Yager and Filev (1999); Chiclana et al. (2007).* An *IOWA operator* of dimension n is a mapping $\Psi : (\mathbb{R} \times \mathbb{R})^n \to \mathbb{R}$ that has an associated weighting vector W such that $w_i \in [0,1]$, $\sum_{i=1}^{n} w_i = 1$, and it is defined to aggregate the set of second arguments of a list of n 2-tuples:

$$\Psi(\langle u_1, p_1 \rangle, \ldots, \langle u_n, p_n \rangle) = \sum_{i=1}^{n} w_i p_{\sigma(i)},$$

being $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}, \forall i = 1, \ldots, n-1$.

The vector of values $U = (u_1, \ldots, u_n)$ is defined as the *order-inducing* vector and $(p_1, \ldots, p_n)$ as the *values of the argument variable*. In this way, the *order-inducing* reorders the *values of the argument variable* based on its magnitude.

P U B L I C A T I O N S

5 LINGUISTIC QUANTIFIERS

Linguistic quantifiers are widely used for modeling the concept of quantification to represent the fuzzy majority Pasi and Yager (2006). *At least half*, *Most of* and *Many as possible* are some examples of these quantifiers

$$Q_{(a,b)}(x) = \begin{cases} 0 & \text{if } 0 \leq x < a, \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b, \\ 1 & \text{if } b \leq x \leq 1 \end{cases}$$

The values that are used for the pair (a, b) are Kacprzyk (1986):

$$Q_{\textit{At least half}}(x) = Q_{(0,0.5)}(x)$$

$$Q_{\textit{Most of}}(x) = Q_{(0.3,0.8)}(x)$$

$$Q_{\textit{Many as possible}}(x) = Q_{(0.5,1)}(x)$$

5 LINGUISTIC QUANTIFIERS

Linguistic quantifiers are widely used for modeling the concept of quantification to represent the fuzzy majority Pasi and Yager (2006). *At least half*, *Most of* and *Many as possible* are some examples of these quantifiers

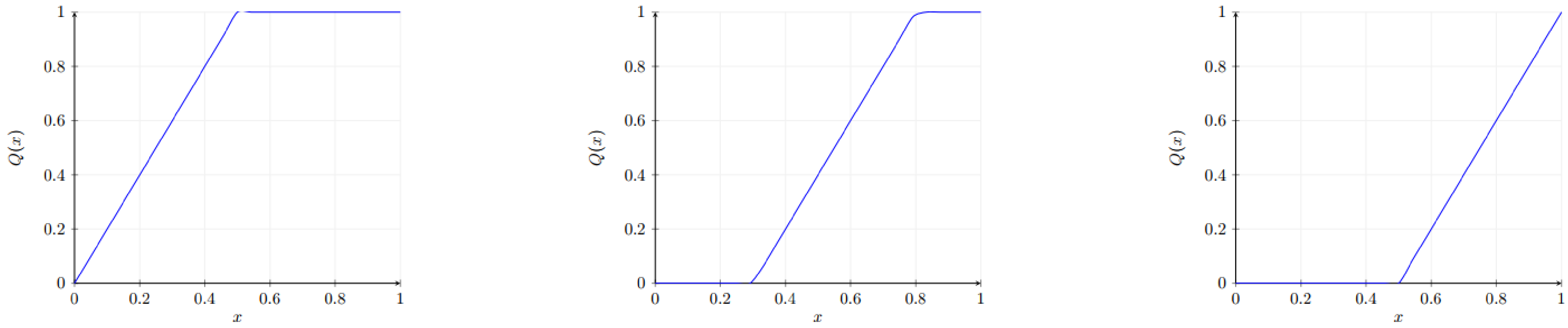


Fig. 17: Linguistic Quantifiers Represented as Fuzzy Sets: *At least half*, *Most of* and *Many as possible*, respectively.

5 LINGUISTIC QUANTIFIERS

Definition 7 *IOWA At Least Half Pro-Neutrality System Based (ALH-ProN).* The *IOWA ALH-ProN operator* of dimension T is a mapping $\Psi_{ALH-ProN} : ([0,1] \times [0,1])^T \rightarrow [0,1]$ that has an associated weighting vector W such that $w_i^{(0,0.5)}$ and it is defined to aggregate the set of second arguments of a list of T 2-tuples:

$$\Psi_{ALH-ProN}(\langle u_1, p_{1k}\rangle, \ldots, \langle u_T, p_{Tk}\rangle) = \sum_{i=1}^{T} w_i^{(0,0.5)} p_{\sigma(i)k},$$

being $\sigma : \{1,\ldots,T\} \rightarrow \{1,\ldots,T\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall i = 1,\ldots,T-1$, and $u_i = NPF_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$.

Definition 8 *IOWA Most Of Pro-Neutrality System Based (MO-ProN).* The *IOWA MO-ProN operator* of dimension T is a mapping $\Psi_{MO-ProN} : ([0,1] \times [0,1])^T \rightarrow [0,1]$ that has an associated weighting vector W such that $w_i^{(0.3,0.8)}$ and it is defined to aggregate the set of second arguments of a list of T 2-tuples:

$$\Psi_{MO-ProN}(\langle u_1, p_{1k}\rangle, \ldots, \langle u_T, p_{Tk}\rangle) = \sum_{i=1}^{T} w_i^{(0.3,0.8)} p_{\sigma(i)k},$$

being $\sigma : \{1,\ldots,T\} \rightarrow \{1,\ldots,T\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall i = 1,\ldots,T-1$, and $u_i = NPF_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$.

5 LINGUISTIC QUANTIFIERS

Definition 9 *IOWA Many As Possible Pro-Neutrality System Based (MAP-ProN).* The *IOWA MAP-ProN operator* of dimension T is a mapping $\Psi_{MAP-ProN} : ([0,1] \times [0,1])^T \rightarrow [0,1]$ that has an associated weighting vector W such that $w_i^{(0.5,1)}$ and it is defined to aggregate the set of second arguments of a list of T 2-tuples:

$$\Psi_{MAP-ProN}(\langle u_1, p_{1k}\rangle, \ldots, \langle u_T, p_{Tk}\rangle) = \sum_{i=1}^{T} w_i^{(0.5,1)} p_{\sigma(i)k},$$

being $\sigma : \{1,\ldots,T\} \rightarrow \{1,\ldots,T\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall i = 1,\ldots,T-1$, and $u_i = NPF_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$.

P U B L I C A T I O N S

5 LINGUISTIC QUANTIFIERS

Definition 10 *IOWA Minimum Neutrality (MinN).* The *IOWA MinN* of dimension T is a mapping $\Psi_{MinN} : ([0,1] \times [0,1])^T \to [0,1]$ that has an associated weighting vector W and it is defined to aggregate the set of second arguments of a list of T 2-tuples:

$$\Psi_{MinN}(\langle u_1, p_{1k}\rangle, \ldots, \langle u_T, p_{Tk}\rangle) = \sum_{i=1}^{T} w_i p_{\sigma(i)k},$$

being $\sigma : \{1,\ldots,T\} \to \{1,\ldots,T\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall h = 1,\ldots,T-1$, and $u_i = NPF_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$ with $w_T = 1$ and $w_i = 0$ for $\forall i = 1,\ldots,T-1$.

Definition 11 *IOWA Maximum Neutrality (MaxN).* The *IOWA MaxN* of dimension T is a mapping $\Psi_{MaxN} : ([0,1] \times [0,1])^T \to [0,1]$ that has an associated weighting vector W and it is defined to aggregate the set of second arguments of a list of T 2-tuples:

$$\Psi_{MaxN}(\langle u_1, p_{1k}\rangle, \ldots, \langle u_T, p_{Tk}\rangle) = \sum_{h=1}^{T} w_i p_{\sigma(i)k},$$

being $\sigma : \{1,\ldots,T\} \to \{1,\ldots,T\}$ a permutation such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall h = 1,\ldots,T-1$, and $u_i = NPF_{\alpha=1}(p_{ik}) = 1 - |p_{ik} - 0.5|$ with $w_1 = 1$ and $w_i = 0$ for $\forall h = 2,\ldots,T$.