Doctoral Consortium 2013
Dept. Lenguajes y Sistemas Informáticos UNED

17-19 June 2013

Monday 17 June

Salón de Actos, Facultad de Psicología, UNED

15.00-16.30: Invited talk

Eneko Agirre (Euskal Herriko Unibertsitatea/Universidad del País Vasco)

Semantic Textual Similarity: *SEM Shared Task in 2013

Semantic Textual Similarity (STS) measures the degree of semantic equivalence between two texts, on a scale from 0 to 5.

I will introduce the STS task as held in 2012 and 2013. In 2013, STS was chosen to be the *SEM conference shared task, and contained two subtasks. The core task similar to the 2012 task, based on pairs of sentences from news headlines, machine translation evaluation datasets and lexical resource glosses. The typed-similarity task is novel and involves pairs of cultural heritage items which are described with metadata like title, author or description.

Several types of similarity have been defined, including similar author, similar time period or similar location.

The annotation leveraged crowdsourcing, with relatively high inter-annotator correlation, ranging from 62% to 87%. In 2013 the core task attracted 34 participants corresponding to 89 runs, and the typed-similarity task attracted 6 teams corresponding to 14 runs.

Selectional Preferences for Semantic Role Classification

Given a sentence, semantic roles indicate "who" did "what" to "whom," "when" and "where". Given a predicate and its arguments (adjuncts) in a sentence, I will focus on the problem of identifying what is the role of each one. For instance, the roles of "in May" and "in Madrid" are different (temporal vs. location), even if the syntactic position and prepositions are the same. At present the best technology uses training

data, but given the sparse data available, we might find arguments that did not occur in the training data (sparseness problem). We mitigate this problem using models that integrate automatically learned selectional preferences, which encode what kind of entities verbs and prepositions expect in each role, e.g. the preposition "in" expects time expresions for the temporal role, and locations for the locations role. We explore a range of models based on WordNet and distributional-similarity. Furthermore, we demonstrate that the SRC task is better modeled by SP models centered on both verbs and prepositions, rather than verbs alone.

17.00-18.30: Invited talk

Maarten de Rijke (Universiteit van Amsterdam)

Metrics from Clicks

In recent years many models have been proposed that are aimed at predicting clicks of web search users. In addition, some information retrieval evaluation metrics have been built on top of a user model. In this this talk I bring these two directions together and propose a common approach to converting any click model into an evaluation metric. I then put the resulting model-based metrics as well as traditional metrics (like DCG or Precision) into a common evaluation framework and compare them along a number of dimensions.

One of the dimensions I am particularly interested in is the agreement between offline and online experimental outcomes. It is widely believed, especially in an industrial setting, that online A/B-testing and interleaving experiments are generally better at capturing system quality than offline measurements. I show that offline metrics that are based on click models are more strongly correlated with online experimental outcomes than traditional offline metrics, especially in situations where we have incomplete relevance judgements.

(This is based on joint work with Aleksandr Chuklin and Pavel Serdyukov.)

Tuesday 18 June

Salón de Actos, Facultad de Psicología, UNED

9.30-13.00: Workshop: Information Access Technologies for Online Reputation Management

Presentation

Social media is playing an increasingly important role in the communication strategy of organizations, creating both opportunities and pitfalls. On the one hand, social media can be used to interact with customers whilst, on the other hand, online opinions and discussions can have a profound impact on an organization's reputation. The growing acceptance of social media and the speed at which facts and opinions travel make them an integral part of a company's public relations strategy. Online reputation management (ORM) aims to monitor the online reputation of an organization, brand, or person. A key aspect of ORM is early detection of topics that may end up influencing the reputation of a given company, brand, or person and it is important to track stories that talk about issue(s) that affects the reputation.

While traditional reputation analysis was based mostly on manual analysis (clipping from media, surveys, etcetera), the distinguishing feature of online media lies in the ability of automatically processing, understanding, and aggregating potentially huge streams of facts and opinions about a company, brand, or individual (the "customer").

Information to be mined includes answers to questions such as:

- What is the general state of opinion about a customer in online media?

- What are the perceived strengths and weaknesses, as compared to its peers and/or competitors?

- How is the customer positioned with respect to its strategic market?

- Can incoming threats to a reputation be detected early enough to be neutralized before they effectively affect reputation?

In the workshop we will discuss the challenges faced by Information Access technologies, which are the current solutions and the prospects for the near future.

Chair: M. Felisa Verdejo (UNED)

Speakers:

(25-30 minutes maximum per speaker + general discussion)

9.30-10.00: ***Essential Aspects for Reputation Analysis***

Julio Villena (Daedalus)

Reputation analysis, as the process of tracking, investigating and reporting an entity's actions and other entities' opinions about those actions, covers different aspects of linguistic technologies, including automatic text classification to detect the topic(s) of the content and sentiment analysis to determine the positive/neutral/negative polarity of the information. In addition, other aspects must be also considered in an in-depth analysis such as named entity detection and disambiguation (to assign the polarity at entity level), subjectivity detection (to differentiate between objective facts and subjective opinions), relevance ranking (to assign the impact of the text on the reputation based on its topic, polarity, involved entities, author...), reputation alert detection (early identification of important issues that may have a substantial impact on the reputation), text normalization (to better deal with short and noisy texts from social media), language identification, and others. In my talk I will present an overview of all these essential aspects that are required to carry out a reputation analysis that fulfils the current market demands, describe the variety of state-of-the-art approaches that have been proposed by different research groups, discuss their strong and weak points and compare the results achieved in different open evaluation forums.

10.00-10.30: ***Why do enterprises worry about the Internet effect on their reputation? Is there an opportunity hidden behind that concern?***

Adolfo Corujo (Llorente y Cuenca)

In the next three years, spanish and latinoamerican companies will spend 2 billion USD trying to get valuable information from Internet. There are several reasons that explains this interest. One of them is the potential impact on reputation that conversations all through social networks have. However this effort can be helpless. Too much expressions, opinions, sentences, words in too different channels, websites, formats, platforms. At the end of the day, companies feel overwhelmed. And this is not only a technical problem. Professionals in consultancy firms, as Llorente & Cuenca, together with their peers in the communications and reputation departments in the big enterprises and scientists share their thoughts to understand which are the basis

of this monitoring. Along this next three years anyone who knows to comprehend, define and solve this puzzle will be in the best position to exploit this huge business opportunity.

<p style="text-align:center">10.30-11.00: Break</p>

11.00-11.30: *One Size Fits All? Entity Dependent Reputation Management*

Maarten de Rijke (Universiteit van Amsterdam)

In reputation management, knowing the impact a tweet has on the reputation of a brand or company is crucial. The reputation polarity of a tweet is a measure for how a tweet influences the reputation of a brand or company. We consider the task of determining the reputation polarity of a tweet. For this classification task, I propose a feature-based model based on three dimensions: the source of the tweet, the contents of the tweet and the reception of the tweet, i.e., how the tweet is being perceived. For evaluation purposes, I make use of the recently introduced RepLab data set. I contrast two training paradigms. The first is independent of the entity whose reputation is being determined, the second depends on the entity at stake, but which, on average, has over 90% fewer training samples per model. I will show that having less but entity-dependent training data is significantly more effective for predicting the reputation polarity of a tweet. The relative effectiveness of features is shown to depend on the training paradigm used.

(This is based on joint work with Hendrike Peetz.)

11.30-12.00: *Reputational polarity: challenges and opportunities*

Jorge Carrillo-de-Albornoz (UNED)

Reputational polarity is to do with determining whether a textual content has positive, negative or neutral implications for corporate reputation. The problem is related to sentiment analysis and opinion mining, but differs in some important points: first, what is being analyzed is not only opinions or subjective content, but also facts, and in particular, polar facts, i.e. objective information that might have negative or positive implications for a company's reputation; second, sponsored information and advertising may entail positive reputational polarity even in the absence of positive content; third, the mere mention of the company and its products names is considered as positive from a reputational perspective. Nowadays Twitter is one of the most important sources for reputational experts when analyzing corporate image and has been the focus of the RepLab 2012 evaluation campaign, where the aim was to compare classification systems trained to analyze

reputational polarity. The results of this campaign showed the complexity of this novel problem and the need for further specific research beyond sentiment analysis techniques.

12.00-12.30: *RepLab: An evaluation campaign for Online Reputation Management Systems*

Julio Gonzalo (UNED)

Born as a close collaboration between industry and academy, RepLab is an evaluation campaign devoted to Online Reputation Management problems from the perspective of Information Access technologies. In its two first years (2012 and 2013), RepLab is gathering a research & practice community around the topic, establishing well-defined technical tasks and challenges, creating suitable test collections and helping to establish a state-of-the-art know how of Information Access techniques for Online Reputation Management. The talk will summarize the progress made so far and will highlight the research opportunities in this area of increasing relevance from the perspective of Public Relations for companies, institutions and individuals.

12.30-13.00: Discussion

15.00-17.40: Doctoral Consortium

15.00-16.00: *Online Reputation Monitoring: Keyword-based Approaches for Filtering and Sub-Topic Detection in Microblog Streams*

Damiano Spina

16.00-16.40: *Textual content modelling*

Ángel Castellanos

16.40-17.00: Break

17.00-17.40: *Open Reading: Unsupervised Acquisition of Textual Knowledge for Inference and Enrichment*

Bernardo Cabaleiro

Wednesday 19 June

Room 1.03. ETSI Informática, UNED

9.30-13.30: Doctoral Consortium

9.30-10.30: ***Relational knowledge acquisition and temporal anchoring***

Guillermo Garrido

10.30-11.20: ***An Unsupervised Approach for Person Name Disambiguation in the Web***

Agustín Delgado

11.20-12.00 ***A new graph-based approach for Word Sense Induction in multilingual contexts.***

Andrés Duque

12.00-13.30: Committee Meeting