



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THESE DE DOCTORAT

Monsieur Kamal AMAROUCHE

soutiendra publiquement sa thèse de Doctorat en Informatique

Le Vendredi 17 Décembre 2021 à 16H au Grand amphi à l'ENSIAS

Intitulé de la thèse

CUSTOMER PRODUCT REVIEW SUMMARIZATION OVER TIME FOR COMPETITIVE INTELLIGENCE



Devant le Jury composé de :

Président :

Pr. Mohammed Abdou JANATI IDRISSE, PES, ENSIAS, Université Mohammed V de Rabat

Directeur de thèse :

Pr. Ismail KASSOU, PES, ENSIAS, Université Mohammed V de Rabat

Co-Directeur de thèse :

Pr. Houda BENBRAHIM, PES, ENSIAS, Université Mohammed V de Rabat

Rapporteurs :

Pr. Najima DAOUDI, PH, ESI, Rabat

Pr. Fatima-Zahra BELOUADHA, PES, EMI, Université Mohammed V de Rabat

Pr. Karim BAÏNA, PES, ENSIAS, Université Mohammed V de Rabat

Examineur :

Pr. Ahmed ZELLOU, PES, ENSIAS, Université Mohammed V de Rabat

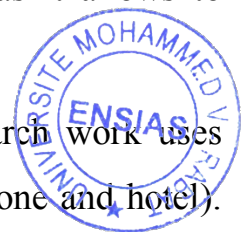
CUSTOMER PRODUCT REVIEW SUMMARIZATION OVER TIME FOR COMPETITIVE INTELLIGENCE

Abstract: Nowadays, Customer's product opinions can be widely found on the Web, such as in personal blogs, forums, and e-commerce websites. These opinions that contain important information about products became a new data source of competitive intelligence. On that account, these opinions need to be analyzed and summarized in order to help companies or customers make appropriate decisions in an effective way.

Previous studies have showed that users tend to have different interests regarding the features of products or services. Features, also called aspects, describe the characteristics of a product/service that users give opinions about, e.g., "Battery life" and "Image quality" for cell phones. A potential customer might be interested by the battery life of the cell phone, while another by the camera resolution. Nevertheless, most previous feature-based opinion summarization studies focus on summarizing sentiment distribution toward different features without taking into account the real advantages and disadvantages of a product clarify over time.

This thesis, proposes a new system for feature-based opinion summarization, which take into consideration "time" when reviews are expressed. The proposed system contains three main subtasks: feature extraction, summary generation, and product classification. The first subtask detects automatically the features of the product/service, being discussed. The second subtask generates a feature-based summarization that depends on the time when reviews are expressed. The main goal of this subtask is to generate a percentage of positivity and negativity around each feature in order to give more accurate and efficient information. The third subtask aggregates the outputs of each feature-based summary corresponding to a given product/service in a single value in order to classify it in its appropriate class (good, medium or bad product), this step is important as it allows to compare a product with its competitors.

To evaluate the proposed system for its reliability and validity, this research work uses datasets across two different domains in the customer reviews area (cell phone and hotel).



The results show the effectiveness of the proposed system at feature extraction, summary generation and classification product levels. The “domain-independent characteristic” is also, a strength of the proposed system.

Keywords: Feature extraction, Fuzzy logic, Competitive intelligence, Opinion mining, Feature- based Opinion Summarization, Sentiment analysis, SentiWordNet.

