# A NEW APPROACH BASED ON THE DETECTION OF OPINION BY SENTIWORDNET FOR AUTOMATIC TEXT SUMMARIES BY EXTRACTION

Reda Mohamed HAMOU[1], Mohamed Amine BOUDIA[2] and Abdelmalek AMINE[3]

Department of Computer, Dr. Moulay Tahar University Saida, Algeria
Laboratory Knowledge Management and Complex Data (GeCoDe Lab) Saida, Algeria
{hamoureda1, mamiamounti2, abd_amine1 3}@yahoo.fr

***ABSTRACT***

*In this paper, we propose a new approach based on the detection of opinion by the SentiWordNet for the production of text summarization by using the scoring extraction technique adapted to detecting of opinion. The texts are decomposed into sentences then represented by a vector of scores of opinion of this sentences. The summary will be done by elimination of sentences whose opinion is different from the original text. This difference is expressed by a threshold opinion. The following hypothesis: "textual units that do not share the same opinion of the text are ideas used for the development or comparison and their absences have no vocation to reach the semantics of the abstract" Has been verified by the statistical measure of Chi_2 which we used it to calculate a dependence between the unit textual and the text. Finally we found an opinion threshold interval which generate the optimal assessments.*

***KEYWORDS***

*Automatic Summary Extraction, Text Mining, Evaluation, Automatic Language Processing, F-Measure, correlation, ROUGE-SU (2), SentiWordNet, Opinion Mining.*

## 1. INTRODUCTION AND PROBLEMATIC

Currently, one of the major problems for computer scientists is access to the content of information, access itself or in other words the software and hardware infrastructure are no longer an obstacle, and the major problem is the exponential increase in the amount of textual information electronically. This requires the use of more specific tools i.e. access to the content of texts by rapid and effective means has become a necessary task.

A summary of a text is an effective way to represent the contents of the texts and allow quick access to their semantic content. The purpose of a summarization is to produce an abridged text covering most of the content from the source text.

Jan Zizka et al. (Eds) : ICAITA, SAI, CDKP, Signal, NCO - 2015
pp. 163–181, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.51514

Summary of text appears interesting for fast access to the content of textual information. A summary is a reissued the original text in smaller form that is done under the constraint of keeping the semantics of a document that is minimized entropy semantics. The purpose of this operation is to help the reader identify interesting information for him without having to read the entire document. The uses of automatic summaries aim to reduce the time to find the relevant documents or reduce treatment long texts by identifying the key information. The volume of electronic textual information is increasing, making access to information difficult. Producing a summary may facilitate access to information, but it is also a complex task because it requires language skills.

To do an automatic summarization, the current literature presents three approaches:

- Automatic Summarization by extraction
- Automatic Summarization by understanding
- Automatic Summarization by automatic classification

Another line of research that has gained momentum in recent years, in case the Opinion Mining or the fact of detecting opinion of a sentence, paragraph or text. Our job is to use detection methods to produce a summary opinion. We propose the hypothesis:

**"Textual units that do not share the same opinion of the text are ideas used for the development or comparison and their absences have no vocation to reach the semantics of the abstract"**

In this work we will generate a summary automatically by extraction approach, we will use the scoring technique where the score will calculate according to opinion by using a SentiWordNet

We will build a summary of the sentences that have an opinion similar to that of full text according to a threshold of opinion; our work will give an answer for the following question:

- Have our hypothesis been testable? If so, is it valid?
- What is the impact of opinion threshold on the quality of the summary?
- The opinion mining can he bring a plus for automatic summarization?

## 2. LITERATURE REVIEW

Automatically produce a summary is an idea that has emerged in the early 1950s, this is a branch of natural language processing (NLP). The first attempts made their apparition in 1950, the community has tried to implement simple approaches as extracting relevant sentences according to a scoring in order to arrive at a summary understandable and easily readable by a human. (Luhn, 1958) [1] (Edmundson, 1969)[2] proposed to identify lexical units carriers of semantics by manual analysis, the result is called extracts (in English), i.e. an extract is a summary built by phrases (considered as pertinent) original text. Their idea is to assign a weight to each sentence that represents its pertinence then extract either by a reduced rate, N sentences whose weight are greater, or a threshold scoring, tell that all the sentences with a score and greater than or equal to the threshold will be kept.

Other work focuses on automating the analysis for the detection of pertinent lexical units: several methods and variations have been proposed (Radev et al, 2001)[3]; (Radev et al, 2004)[4]; (Boudin and al , 2007)[5]; (Carbonell and al, 1998)[6].(Boudin and al, 2008)[7] we see that we can select two sentences pertinent but if it looks like, they have worked on eliminating redundancy by similarity measures sentences, they even prove their method detects the topic text.

The proposed approaches are intuitive, since the 1960s (Edmundson, 1969)[2] pro-poses to make an identification of keywords (based on a theme), followed by a job that involves a consideration of the position of sentences in the document. The MEAD2 system is the most popular nowadays, developed instead by Ramdev, he implemented the approach of this type (Radev et al., 2001)[3]. He identified the most salient words in each text, which he calls "centroid", and A extract is made by sen-tence who contain the greatest number of "centroid". Another Neo-cortex system was developed at the University of Avignon, which is based on the combination of screening measures, phrases to select the benefits covered by each of them (action selection), this system obtained very good results by evaluating the algorithm MMR (Maximal Marginal Relevance) of (Carbonell & Goldstein, 1998)[6].

Some years later other approaches were presented, these approaches are based on knowledge representation (Mani, 2004)[8], the thematic segmentation (Farzindar et al, 2004)[9] Or recognition of user profiles (author of full text) (Crispino & Couto, 2004)[10].

Noting some attempts to introduce deep linguistic systems in automatic summarization, most recent work was on the syntax tree by providing a method for comparing between syntactic trees, make the elimination of syntax trees or merge (Barzilay & McKeown, 2005)[11];. This attempt was followed by another job that offers alternative methods of syntactic compression based on theoretical and empirical linguistic properties (Yousfi-Monod, 2007)[12].

Most jobs are not interested in the opinion contained in the text, at the end of 2000 the community began to have more specialized query summaries, including dealing with the detection and analysis of the opinion. (Eyrich, et al, 2001)[13] propose a system diagram that has not been implemented: This system integrates a QR module and an analysis module opinion to make a summary of the responses to the opinion without changing the content.

Work on automatic summaries have neglected early analysis of the opinion. Analysis of opinion is divided into three main levels of subtasks: the first sub-task is to distinguish between a subjective texts and an objective text (Yu & Hatzivassiloglou, 2003) [14]. The second sub-task is to classify texts subjective positive or negative (Turney, 2002)[15]. The third level of refinement trying to determine the extent to which positive or negative texts (Wilson et al., 2004)[16]. The impulse given by campaigns such as TREC Blog task opinion since 2006 is undeniable (Pang & Lee, 2008)[17]. (Zhang et al, 2007)[18];(Dey & Haque, 2009)[19].

Opinion Mining is an area that has attracted many researchers which resulted in several works. There are two types of approaches for detecting opinion: Approaches based on corpus (Corpus-based Approach) (Hatzivassiloglou and McKeown, 1997)[20] (Wiebe, 2000) [21] (Kanayama and Nasukawa, 2006) [22] (Esuli and Sebastiani 2006) [23]; and (Qiu et al, 2009) [24], others based on a dictionary (Dictionary-based Approach) dictionary (Hu and Liu, 2004) [25], (Kim and Hovy, 2004) [26], (Kamps et al, 2004) [27], (Esuli and Sebastiani , 2005)[28], (Andreevskaya and Bergler, 2006) [29], and (Bouchlaghem et al, 2010) [30].

This is the second approach that will be used in this article, (Wu and Liu, 2004) use the adjectives for the detection of opinions. They manually build a list of adjectives they use to predict the orientation of the sentence and use WordNet to supply the list synonyms and antonyms of adjectives whose polarity is known. In (Liu and al, 2007)[31], the authors count the number of occurrences of each entity in the "Pros" expressing a positive opinion and "Cons" that negative opinions. In (Zhang and Liu, 2011) [32], the authors showed that the noun and noun phrases can also enclose opinions, They count the number of positive and negative sentences for each feature of the product using the lexicon of opinion prepared by (Ding and al, 2008)[33].Strength (intensity) of opinion is also required; Indeed, subjectivity is expressed in different ways; "Good battery" is different from "great battery" and "excellent battery." (Pang and Lee, 2008)[17] focus on detecting the strength of opinion using the techniques of boosting, rule learning and support vector regression. (Pang and Lee, 2008) [17]and (Turney, 2002)[15] classify documents as "thumbs up" or "thumbs down" according to the opinion they convey. However, (Pang and Lee, 2005)[34] exploit machine learning techniques to give a score from 1 to 5 on passages opinions. While (Esuli and Sebastiani 2006)[28] construct the SentiWordNet that a dictionary of general opinion; currently in its third version SentiWordNet 3.0; SentiWordNet can be defined as a lexical resource designed specifically for use by application of detecting opinions and feelings. SentiWordNet is the result of an annotation of all synsets of WordNet so that it assigns to each word (synset) an opinion score. SentiWordNet contains 1000 synsets which makes it very small compared with WordNet, besides the 1000 synsets SentiWordNet automatically ignores all other inputs. Another weakness is that several synsets are not carrying opinion[35].

we can not pretend that our opinion detection for the production of summary text goes beyond the assessment of degrees of positivity or negativity, we must shed light on recent efforts to introduce more linguistic and discourse approaches (taking into account the modality of the speaker) in this accompli by (Asher and al., 2008)[36].

As for the evaluation of abstracts is a crucial problem, emphasize the contribution (Goulet, 2007)[37] that goes beyond the coverage of n-grams and offers a terminology adapted to French. In recent years, large-scale assessments, independent designers systems have emerged and several evaluation measures have been proposed. As regards the assessment of automatic summary, two evaluation campaigns have already been conducted by the U.S. DARPA (Defense Advanced Research Projects Agency). The first, entitled SUMMAC, ran from 1996 to 1998 under the TIPSTER (Lin and al., 2003) [38] program, and the second, entitled DUC (Document Understanding Conferences) (noting that France still lagging behind several countries in all science especially: Computer science), (Das and al., 2007) [39] followed from 2000 to 2007. Since 2008 it is the Text Analysis Conference Such measures may be applied to the distribution of units in the summaries of P systems and those of reference Q. The method was evaluated by Lin et al. (2006) on the corpus DUC'02 for tasks mono and multi-document summary. A good correlation was found between measures of divergence and the two rankings obtained with ROUGE and coverage. (Louis & Nenkova ,2009) [39] went further and, proposed to compare the distribution of words in the complete documents with the words in automatic summaries to infer an evaluation measure based on the content.

# 3. OUR PROPOSED APPROACH

Our approach is based on the identification of opinions textual units (phrases, clauses, sentences, paragraphs), the identification of opinion original text, finally extracting textual units that share the same opinion that the original text

We start from the hypothesis mentioned in previous section.

Recall that the opinion is an expression of the feelings of a person towards an enti-ty or an aspect of the entity (Liu, 2010). An entity may be a product, a person, event, organization or topic.

SentiWordNet is used to filter the word bearer of opinion first, then we will use the score returned by SentiWordNet like so:

```
    If (score_opinion (term i) <0) then the opinion of (term i)
is negative, else opinion is positive
```

Our approach will follow the following steps:

## 3.1 Pretreatment

*Simple cleaning:* Empty words will not be deleted, because the method for automatic summarization by extraction is based on extracting the most informative sentences without change and because the final result is a text (abstract) : if any words will be deleted without information on their morphosyntactic and semantic impact in sentences, you can get a text summary inconsistent.

And for this cleaning will be limited to delete emoticons and replace spaces with _ and remove the special characters that cannot fit in French or English literature (#, \, [,]............)

*Choice of term:* for automatic summarization by extraction we will need two representations:

- Bag of words representation.
- Bag sentence representation.

The two representations are introduced in the context of vector model:

The first representation is to transform the text into a vector vi ( $w_1$ , $w_2$ , .... , $w_{|T|}$ ) where T is the number of all the words appearing at least once in the text. The weight $w_k$ indicates the occurrence of the word $t_k$ in the document.

The representation is to transform the text into a vector $V^1_i$ ( $w'_1$ , $w'_2$ , .... , $w'_{|R|}$ ) where R is the number of all the phrases that appear at least once in the text. The $w'_k$ weight indicates the occurrence of $t_k$ sentence in the document.

And finally a matrix of occurrence sentences * will generate a word from the two previous representations, the size of the matrix is :

- The number of words in the text * the number of words in the text,
- The weight pik represents the number of occurrences of the word k in sentence i;

## 3.2 Detecting opinion by SentiWordNet:

The "sentence-term" matrix is reduced to a "sentence - carrier term of opinion " matrix filtering the term vector vi by SentiWordNet, no-existing terms in the dictionary opinion will be eliminated.

At the end of this step, a matrix M of size nxp where n is equal to the number of phrases and p is equal to the number of term carrier opinion. $M_{ij}$ indicates the occurrence of the word (opinion holder) j in sentence i

$$O_{ij} = M_{ij} * score\ (j) \qquad (1)$$

The score (j) is the score obtained by the SentiWordNet for the term j

### 3.3 Construction of Summary

Weighting: Once the matrix "Phrase- carrier term of opinion" is ready, we calculate the score of phrases as well as the score of text in order to proceed to the final step.

The opinion score for textual units (sentences, paragraph or text) is equal to average the score of the holders of opinion obtained by the SentiWordNet terms.

So the score of opinion of each sentence will be calculated as follows:

$$Score_{phrase}[i] = \frac{\sum_{i=0}^{n} O_{ij}}{\sum_{i=0}^{n} M_{ij}} \qquad (2)$$

such that n = number of carrier term of opinion in the textual unit

Finally, we identified the opinion of text that is the average of opinion score of phrase that the compound:

$$Score_{texte} = \frac{\sum_{i=0}^{R} Score_{phrase}[i]}{R} \qquad (3)$$

As size R of vector V' (number of sentences)

*Summary Final*: "The suggested procedure claims on the principle that high-frequency words in a document are important words" [Luhn 1958]. In our case we will adapt this quote as follows: "The suggested procedure claims on the principle that sentences that share the same opinion that the document (text) are important phrases" ie phrases that do not share the same opinion with text without phrases that have been used by the author to develop an idea or comparison is equivalent to saying that we can eliminate them without causing a large entropy of sense.

The final step is to select the phrases that have the same opinions as the text, for this we proposed this method:

**By neighborhood threshold of opinion of phrases:** we kept the phrase his degree of similarity of opinion between this phrase and the text is greater than or equal threshold neighborhood or opinion threshold.

*For each sentence k do*
*If (score_texte - threshold_opinion < threshold _phrase [k]) and (threshold _phrase [k] <score_texte + threshold _opinion)*
*Then selected the sentence k*

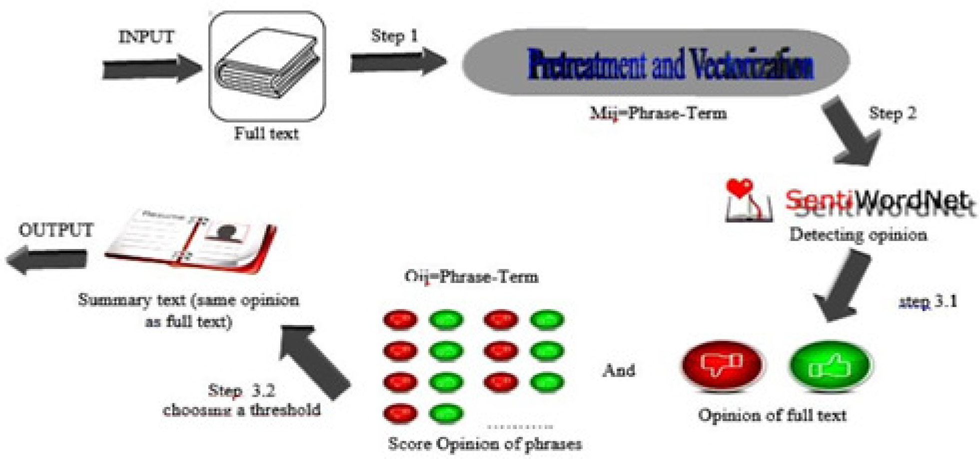


Fig. 1. Full process of the proposed approach

## 4. EXPERIMENTATION

To test our hypothesis already mentioned we use the Chi2 measure is a well-known statistical measure, it evaluates the lack of independence between a textual unit and a text. It uses the same concepts of co-occurrence word / text mutual information, but the difference lies on standardization, which makes them comparable terms. Measurement Chi_2 still loses relevance for infrequent terms.[41]

$$X^2\,(ut_k, text_i) = \frac{|T|.\,[P\,(ut_k, text_i).\,P\,(\overline{ut_k}, \overline{text_i}) -\, P\,(ut_k, text_i).\,P\,(\overline{ut_k}, text_i)]^2}{P(ut_k).\,P(\overline{ut_k}).\,P(text_i).\,P(\overline{text_i})} \quad (4)$$

The use of measurement Chi_2 determines the independence of sentences eliminated by detecting opinion with the original text. Chi_2 promotes the absence of terms and the most common and takes into consideration information from the text terms. A high value of the Chi-2 (k, i) reflects a dependency between the sentence k and the text i .

The second step of our experiment will begin to study is a robustness of summary, we use two evaluation method ROUGE-SU (2) (Recall-Oriented Understudy for Gisting Evaluation –Skip Unit) and F-measure.

## 4.1 Used corpus

Was used as the text corpus "Hurricane" in the English text (to use the SentiWordNet): the text contains a title and 20 sentences.

*Summaries reference*: we took three reference summary product successively by Summarizer CORTEX, Essentiel Summarizerand the third by a human expert.

| |
|---|
| **Cortex**<br><br>*Hurricaine Gilbert Swept towrd the Dominican Republic Sunday, and the Civil Defense alerted its heavily populated south coast to prepare for high winds, heavy rains, and high seas. The National Hurricaine Center in Miami reported its position at 2 a.m. Sunday at latitude 16.1 north, longitude 67.5 west, about 140 miles south of Ponce, Puerto Rico, and 200 miles southeast of Santo Domingo. The National Weather Service in San Juan, Puerto Rico, said Gilbert was moving westard at 15 mph with a "broad area of cloudiness and heavy weather" rotating around the center of the storm. Strong winds associated with the Gilbert brought coastal flooding, strong southeast winds, and up to feet to Puerto Rico's south coast.* |
| **Essential Summarizer**<br><br>*"There is no need for alarm," Civil Defense Director Eugenio Cabral said in a television alert shortly after midnight Saturday. Cabral said residents of the province of Barahona should closely follow Gilbert's movement. On Saturday, Hurricane Florence was downgraded to a tropical storm, and its remnants pushed inland from the U.S. Gulf Coast. Residents returned home, happy to find little damage from 90 mph winds and sheets of rain.* |
| **Human expert**<br><br>*Hurricane Gilbert is moving toward the Dominican Republic, where the residents if the south coast, especially the Barahona Province, have been alerted to prepare for heavy rain, and high winds and seas. Tropical storm Gilbert formed in the eastern Caribbean and became a hurricane on Saturday night. By 2am Sanday it was about 200 miles southeast of Santo Domingo and moving westward at 15 mph with of 75 mph. Flooding is expected in Puerto Rico and the Virgin Islands. The second hurricane of the season, Florence, is now over the southern United States and downgraded to a tropical storm.* |

## 4.2 Validation

To estimate the robustness of our summary, we calculate the correlation metric ROUGE-SU (2) (Lin 2004) which compares a candidate summary (automatically generated by the system to be evaluated) and a reference summary (created by human experts or other automatic summarization systems known). And another F-Measure metric that we have proposed in earlier work.

***The evaluation measure Recall - Oriented Understudy for Gisting Evaluation.***

The evaluation of abstracts can be done semi-automatically through measures of similarities computed between a candidate summary and one or more reference summaries. We evaluate the results of this work by the measure called Recall - Oriented Understudy for Gisting Evaluation

(ROUGE) proposed by (Lin, 2004) involving the differences between distributions of words. Heavily used in the DUC campaigns, these measures are increasingly considered standard by the community because of their strong correlation with manual ratings. Two variants of ROUGE will be developed; it is the measures used in the DUC campaigns.

***ROUGE (N)***

Measurement of return calculated on the co-occurrences of N-grams between a candidate summary $R_{can}$ and a set of reference summaries $R_{ref}$ Co-occurrences (N-gram) is the maximum number of co-occurrences N-grams and $R_{ref}$ in $R_{can}$ number and (N-grams ) to the number of N-grams appearing in the abstract

$$ROUGE\ (N) = \frac{\sum_{s \in \mathrm{R_{ref}}} \sum_{s \in \mathrm{R_{can}}} \mathrm{Co-occurences}\ (\mathrm{R_{ref}}, \mathrm{R_{can}}, \mathrm{N})}{\mathrm{Nbr-NGramme}\ (\mathrm{N})_{\mathrm{R_{ref}}}} (5)$$

***ROUGE-SU (M)***

Adaptation of ROUGE-2 using bigrams hole (skip units (SU)) maximum size M and counting unigrams.

**F-Measure pour l'évaluation des résumés automatique par extraction**

We proposed in our previous work an adaptation of the F-measure for the validation of automatic summarization by extraction, since this technique is based on sentences to keep and delete else following some a philosophy (scoring, detection Thematic....), this can be considered as a two-class classification.

Our method is a hybrid between the two valuation methods: intrinsic and extrinsic.

We shall compare the applicant and the full text (summary) summary to identify textual units that have been kept and which have been deleted, then did the same operation between the reference summary and the full text (summary), and finally a comparison between the two summaries (candidate and reference) is performed to obtain the following confusion matrix.

Table 1. Confusion matrix summary automatic

| Confusion matrix summary automatic | | candidate summary | |
|---|---|---|---|
| | | Tu-K | Tu-D |
| reference summary | Tu-K | X | Y |
| | Tu-D | Z | W |

Tu-K: textual unit to keep

To-D: textual unit to delete

A Recall of "Tu-K" class is defined by the number of textual units kept in the candidate and reference summary (shared), divided by the number of units of text kept by the reference summary; in parallel calculates Recall of "Tu-D" class in the same way that is to say, the number of text units deleted in the candidate and reference summary (shared), divided by the number of units of text deleted reference summary

$$Rappel_{\mathrm{Tu-K}} = \frac{X}{X+Y} (6) \quad Rappel_{\mathrm{To-D}} = \frac{W}{W+Z} (7)$$

The precision of "Tu-K" class is defined by the number of textual units kept in the candidate and reference summary (shared), divided by the number of units of textual kept by the candidate summary; in parallel calculates the precision class "Tu-D" in the same way that is to say, the number of text units removed from the candidate and reference summary (shared), divided by the number of textual units deleted by the candidate summary.

$$Précision_{\text{Tu−K}} = \frac{X}{X+Z} \, (8) \qquad Précision_{\text{To−D}} = \frac{W}{W+Y} \, (9)$$

Since automatic summarization by extraction is a two-class classification thus:

$$R\text{appel} = \frac{\text{Rappel}_{\text{Ut toG}} + \text{Rappel}_{\text{Ut toS}}}{2} \, (10) \qquad \text{Précision} = \frac{\text{Précision}_{\text{Ut toG}} + \text{Précision}_{\text{Ut toS}}}{2} \, (11)$$

Finally combining precision and recall is calculated weighting to the F-Measure

$$F - Mesure = \frac{2 \, * (Précision \, * \, Rappel)}{(Précision + Rappel)} \, (12)$$

### 4.3 The algorithm description.

```
Begin
Pretreatment paper
Mij = integer array [1 .. number of sentence] [1 .. number of
terms]
Mij = vectorization (Word Bag, bag of sentences)
Oij = integer array [1 .. number of sentence] [1 .. number of
terms]
For each sentence i do begin
For each j terms to begin
```

$$O_{ij} = \, M_{ij} * score \, (j) (13)$$

```
End for
End for
Score _phrase = array real [1 .. number of sentence]
for each sentence k do begin
```

$$Score\_phrase \, [i] \, = \frac{\sum_{i=0}^{n} O_{ij}}{\sum_{i=0}^{n} M_{ij}} \, (14)$$

```
End for
```

$$Score\_texte \, = \frac{\sum_{i=0}^{R} Score_{phrase}[i]}{R} \, (15)$$

```
Threshold = défini_par _l'utilisateur
For each sentence do i start
If (score_texte-seuil_voisingane <score_phrase [i] <+ score_texte
seuil_voisingane)
then remember sentence [i] if not eliminated sentence [i]
End for
END
```

## 4.4 Result

| threshold | Cortex | | | Essential Summarizer | | | Expert Humain | | |
|---|---|---|---|---|---|---|---|---|---|
| 0,00125 | VP=0 | FN=74 | P=0,2757 | VP=3 | FN=53 | P=0,5893 | VP=5 | FN=62 | P=0,7287 |
| | FP=6 | VN=91 | R=0,4690 | FP=3 | VN=112 | R=0,5137 | FP=1 | VN=103 | R=0,5325 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,3473 | 0,0 | | 0,5489 | | 0,00535 | 0,6153 | | 0,06172 |
| 0,0025 | VP=14 | FN=60 | P=0,6513 | VP=5 | FN=51 | P=0,4561 | VP=10 | FN=57 | P=0,5612 |
| | FP=6 | VN=91 | R=0,5634 | FP=15 | VN=100 | R=0,4794 | FP=10 | VN=94 | R=0,5265 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,6043 | 0,18918 | | 0,4674 | | 0,08928 | 0,5433 | | 0,12345 |
| 0,00375 | VP=16 | FN=58 | P=0,5236 | VP=8 | FN=48 | P=0,4424 | VP=13 | FN=54 | P=0,4960 |
| | FP=18 | VN=79 | R=0,5153 | FP=26 | VN=89 | R=0,4583 | FP=21 | VN=83 | R=0,4940 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,5194 | 0,21621 | | 0,4502 | | 0,14285 | 0,4950 | | 0,16049 |
| 0,005 to 0,0075 | VP=39 | FN=35 | P=0,6885 | VP=13 | FN=43 | P=0,4254 | VP=21 | FN=46 | P=0,4824 |
| | FP=18 | VN=79 | R=0,6707 | FP=44 | VN=71 | R=0,4247 | FP=36 | VN=68 | R=0,4836 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,6707 | 0,52702 | | 0,4251 | | 0,23214 | 0,4830 | | 0,25925 |
| 0,00875 | VP=41 | FN=33 | P=0,6473 | VP=14 | FN=24 | P=0,4025 | VP=22 | FN=45 | P=0,4478 |
| | FP=26 | VN=71 | R=0,6430 | FP=53 | VN=62 | R=0,3945 | FP=45 | VN=59 | R=0,4478 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,6451 | 0,55405 | | 0,3951 | | 0,25 | 0,4478 | | 0,27160 |
| 0,01 to 0,01125 | VP=46 | FN=28 | P=0,6780 | VP=15 | FN=41 | P=0,3970 | VP=23 | FN=44 | P=0,4375 |
| | FP=26 | VN=71 | R=0,6767 | FP=57 | VN=58 | R=0,3861 | FP=49 | VN=55 | R=0,4360 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,6776 | 0,62162 | | 0,3915 | | 0,26785 | 0,4367 | | 0,28395 |
| 0,0125 | VP=49 | FN=25 | P=0,6428 | VP=25 | FN=31 | P=0,4668 | VP=30 | FN=37 | P=0,4613 |
| | FP=36 | VN=61 | R=0,6455 | FP=60 | VN=55 | R=0,4623 | FP=55 | VN=49 | R=0,4594 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,6441 | 0,66216 | | 0,4685 | | 0,44642 | 0,4604 | | 0,37037 |
| 0,01375 | VP=49 | FN=25 | P=0,5934 | VP=25 | FN=31 | P=0,4276 | VP=30 | FN=37 | P=0,4144 |
| | FP=46 | VN=51 | R=0,5939 | FP=70 | VN=45 | R=0,4188 | FP=65 | VN=39 | R=0,4113 |
| | F-Measure | ROUGE-SU(2) | | F-Measure | | ROUGE-SU(2) | F-Measure | | ROUGE-SU(2) |
| | 0,5936 | 0,66216 | | 0,4232 | | 0,44642 | 0,4129 | | 0,37037 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0,015 | VP=50 | FN=24 | P=0,5612 | VP=26 | FN=30 | P=0,4011 | VP=38 | FN=29 | P=0,4662 |
| | FP=54 | VN=43 | R=0,5594 | FP=78 | VN=42 | R=0,3930 | FP=66 | VN=38 | R=0,4622 |
| | F-Measure 0,5603 | ROUGE-SU(2) 0,67567 | | F-Measure 0,3970 | ROUGE-SU(2) 0,46428 | | F-Measure 0,4662 | ROUGE-SU(2) 0,46913 | |
| 0,01625 to 0,0175 | VP=51 | FN=23 | P=0,5327 | VP=34 | FN=22 | P=0,4653 | VP=42 | FN=25 | P=0,4756 |
| | FP=61 | VN=36 | R=0,5301 | FP=78 | VN=37 | R=0,4644 | FP=70 | VN=34 | R=0,4768 |
| | F-Measure 0,5314 | ROUGE-SU(2) 0,68918 | | F-Measure 0,4648 | ROUGE-SU(2) 0,60714 | | F-Measure 0,4762 | ROUGE-SU(2) 0,51851 | |
| 0,01875 to 0,02 | VP=53 | FN=21 | P=0,4903 | VP=46 | FN=10 | P=0,5791 | VP=49 | FN=18 | P=0,5260 |
| | FP=61 | VN=36 | R=0,4921 | FP=78 | VN=37 | R=0,5715 | FP=75 | VN=29 | R=0,5050 |
| | F-Measure 0,4912 | ROUGE-SU(2) 0,71621 | | F-Measure 0,5753 | ROUGE-SU(2) 0,82142 | | F-Measure 0,5055 | ROUGE-SU(2) 0,60493 | |
| 0,02125 | VP=53 | FN=21 | P=0,4542 | VP=46 | FN=10 | P=0,5592 | VP=49 | FN=18 | P=0,4756 |
| | FP=76 | VN=21 | R=0,4663 | FP=83 | VN=32 | R=0,5489 | FP=80 | VN=24 | R=0,4812 |
| | F-Measure 0,4608 | ROUGE-SU(2) 0,71621 | | F-Measure 0,5545 | ROUGE-SU(2) 0,82142 | | F-Measure 0,4783 | ROUGE-SU(2) 0,60493 | |
| 0,0225 to 0,02625 | VP=64 | FN=10 | P=0,5672 | VP=47 | FN=9 | P=0,5226 | VP=57 | FN=10 | P=0,5422 |
| | FP=76 | VN=21 | R=0,5406 | FP=93 | VN=22 | R=0,5152 | FP=83 | VN=21 | R=0,5263 |
| | F-Measure 0,5536 | ROUGE-SU(2) 0,86486 | | F-Measure 0,5189 | ROUGE-SU(2) 0,83928 | | F-Measure 0,5341 | ROUGE-SU(2) 0,70370 | |
| 0,0275 to 0,03 | VP=64 | FN=10 | P=0,5191 | VP=47 | FN=9 | P=0,4809 | VP=59 | FN=8 | P=0,5420 |
| | FP=82 | VN=15 | R=0,5097 | FP=99 | VN=16 | R=0,4892 | FP=87 | VN=17 | R=0,5220 |
| | F-Measure 0,5144 | ROUGE-SU(2) 0,86486 | | F-Measure 0,4850 | ROUGE-SU(2) 0,83928 | | F-Measure 0,5318 | ROUGE-SU(2) 0,72839 | |

Table 2. Result Evaluation of summary produced (candidate) with ROUGE and F-measure using 3 reference summary Cortex, Essential Summaries, human expert (second part)

The above table includes an assessment summary with a different threshold of opinions comparing with all reference summary using the F - measure and ROUGE. The following table shows in a manner explicit the selected sentences (keep) (K) and the sentences deleted (D) at each threshold of opinion and gives the chi_2 value for each sentence with original text and the chi2 rate of each summary report by the original text.

The rate of chi-2 which is equal to the sum of value chi_2 sentence divided by retained by the number of all the phrases which constitutes the original text, it indicates the correctness of choice of phrases relative to their dependence original text.

| Phrase / Thres Hold | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | Chi_2 Summ4ry % | Redur ate % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | D | 0 | 100 |
| 0,00125 | D | D | D | D | D | D | D | D | D | D | D | D | D | D | K | D | D | D | D | D | D | 5,422 | 95,24 |
| 0,0025 | D | D | D | D | D | D | D | D | D | K | D | D | D | D | K | D | D | D | D | D | D | 9,836 | 90,48 |
| 0,00375 | D | D | D | D | D | D | D | D | D | K | D | D | D | D | K | D | D | D | D | D | K | 14,24 | 85,72 |
| 0,005 to 0,0075 | D | D | D | D | D | D | D | D | D | K | K | D | D | D | K | D | D | D | D | D | K | 18,01 | 80,95 |
| 0,00875 | D | D | D | D | D | D | D | D | D | K | K | D | D | D | K | K | D | D | D | D | K | 22,26 | 76,2 |
| 0,01 to 0,01125 | D | D | D | D | D | D | K | D | D | K | K | D | D | D | K | K | D | D | D | D | K | 27,17 | 71,43 |
| 0,0125 | D | D | D | K | D | D | K | D | K | K | K | D | K | D | K | K | D | K | D | D | K | 51,66 | 52,39 |
| 0,01375 | D | D | D | K | K | D | K | D | K | K | K | D | K | D | K | K | D | K | D | D | K | 55,59 | 47,62 |
| 0,015 | D | D | D | K | K | D | K | D | K | K | K | K | K | D | K | K | D | K | D | D | K | 59,90 | 42,86 |
| 0,01625 to 0,0175 | D | D | D | K | K | D | K | D | K | K | K | K | K | D | K | K | K | K | D | D | K | 64,33 | 38,1 |
| 0,01875 to 0,02 | D | D | K | K | K | D | K | D | K | K | K | K | K | D | K | K | K | K | D | D | K | 68,42 | 33,34 |
| 0,02125 | D | D | K | K | K | K | K | D | K | K | K | K | K | D | K | K | K | K | D | D | K | 73,01 | 28,58 |
| 0,0225 to 0,02625 | K | D | K | K | K | K | K | D | K | K | K | K | K | D | K | K | K | K | D | D | K | 76,83 | 23,81 |
| 0,0275 to 0,03 | K | D | K | K | K | K | K | D | K | K | K | K | K | D | K | K | K | K | D | K | K | 81,33 | 19,05 |
| >0,03 | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | K | 100 | 0 |
| Chi-2 | 0,44563291 | 0,5113027 | 0,47671441 | 0,54112348 | 0,45832587 | 0,53889372 | 0,57254609 | 0,68138226 | 0,71500946 | 0,51435078 | 0,43793806 | 0,50251734 | 0,83666115 | 0,47430067 | 0,63190694 | 0,49601393 | 0,51648899 | 0,76084162 | 0,50841062 | 0,52420163 | 0,51382485 | | |

Table 3. Distribution of sentences in summary product (candidate) for each threshold of opinion , his chi_2 rate and reduction rate, and Chi_2 value for each sentence

## 4.5 The best summary text

Hurricaine Gilbert Swept towrd the Dominican Republic Sunday, and the Civil Defense alerted its heavily populated south coast to prepare for high winds, heavy rains, and high seas.
The National Hurricaine Center in Miami reported its position at 2 a.m. Sunday at latitude 16.1 north, longitude 67.5 west, about 140 miles south of Ponce, Puerto Rico, and 200 miles southeast of Santo Domingo.
The National Weather Service in San Juan, Puerto Rico, said Gilbert was moving westard at 15 mph with a "broad area of cloudiness and heavy weather" rotating around the center of the storm.
Strong winds associated with the Gilbert brought coastal flooding, strong southeast winds, and up to 12 feet to Puerto Rico's south coast.

The following table summarizes all the results mentioned in the two previous tables in case the table2 and table 3:

| | | 0,00125 | 0,0025 | 0,00375 | 0,005 to 0,0075 | 0,00875 | 0,01 to 0,01125 | 0,0125 | 0,01375 |
|---|---|---|---|---|---|---|---|---|---|
| ROUGE-SU(2) | cortex | 0,0 | 0,18918 | 0,21621 | 0,52702 | 0,55405 | 0,62162 | 0,66216 | 0,66216 |
| | Es.Sum. | 0,00535 | 0,08928 | 0,14285 | 0,23214 | 0,25 | 0,26785 | 0,44642 | 0,44642 |
| | Humain | 0,06172 | 0,12345 | 0,16049 | 0,25925 | 0,27160 | 0,28395 | 0,37037 | 0,37037 |
| Recall | Cortex | 0,4690 | 0,5634 | 0,5153 | 0,6707 | 0,6430 | 0,6767 | 0,6455 | 0,5939 |
| | Es,Sum, | 0,5137 | 0,4794 | 0,4583 | 0,4247 | 0,3945 | 0,3861 | 0,4623 | 0,4188 |
| | Humain | 0,5325 | 0,5265 | 0,4940 | 0,4836 | 0,4478 | 0,4360 | 0,4594 | 0,4133 |
| Precision | Cortex | 0,2757 | 0,6513 | 0,5236 | 0,6885 | 0,6473 | 0,6780 | 0,6428 | 0,5934 |
| | Es,Sum, | 0,5893 | 0,4561 | 0,4424 | 0,4254 | 0,4025 | 0,3970 | 0,4668 | 0,4276 |
| | Humain | 0,7287 | 0,5612 | 0,4960 | 0,4824 | 0,4478 | 0,4375 | 0,4613 | 0,4144 |
| F-Measure | Cortex | 0,3473 | 0,6043 | 0,5194 | 0,6707 | 0,6451 | 0,6776 | 0,6441 | 0,5936 |
| | Es,Sum, | 0,5489 | 0,4674 | 0,4502 | 0,4251 | 0,3951 | 0,3915 | 0,4685 | 0,4232 |
| | Humain | 0,6153 | 0,5433 | 0,4950 | 0,4830 | 0,4478 | 0,4367 | 0,4604 | 0,4129 |
| Chi2 rate summary | | 0 | 5,422 | 9,836 | 14,24 | 18,01 | 22,26 | 27,17 | 51,66 |
| Rate Reduction | | 100 | 95,24 | 90,48 | 85,72 | 80,95 | 76,2 | 71,43 | 52,39 |

this is the second part of Table above

| | | 0,015 | 0,01625 to 0,0175 | 0,01875 TO 0,02 | 0,02125 | 0,0225 to 0,02625 | 0,0275 to 0,03 |
|---|---|---|---|---|---|---|---|
| ROUGE-SU(2) | cortex | 0.68918 | 0.71621 | 0.71621 | 0.86486 | 0.86486 | 0,66216 |
| | Es.Sum. | 0.60714 | 0.82142 | 0.82142 | 0.83928 | 0.83928 | 0,44642 |
| | Humain | 0.51851 | 0.60493 | 0.60493 | 0.70370 | 0.72839 | 0,37037 |
| Recall | Cortex | 0.5301 | 0.4921 | 0.4663 | 0.5406 | 0.5097 | 0,6455 |
| | Es,Sum, | 0.4644 | 0.5715 | 0.5489 | 0.5152 | 0.4892 | 0,4623 |
| | Humain | 0.4768 | 0.5050 | 0.4812 | 0.5263 | 0.5220 | 0,4594 |
| Precision | Cortex | 0.5327 | 0.4903 | 0.4542 | 0.5672 | 0.5191 | 0,6428 |
| | Es,Sum, | 0.4653 | 0.5791 | 0.5592 | 0.5226 | 0.4809 | 0,4668 |
| | Humain | 0.4756 | 0.5260 | 0.4756 | 0.5422 | 0.5420 | 0,4613 |
| F-Measure | Cortex | 0.5314 | 0.4912 | 0.4608 | 0.5536 | 0.5144 | 0,6441 |
| | Es,Sum, | 0.4648 | 0.5733 | 0.5545 | 0.5189 | 0.4850 | 0,4685 |
| | Humain | 0.4762 | 0.5055 | 0.4783 | 0.5341 | 0.5318 | 0,4604 |
| Chi2 rate summary | | 59.90 | 64.33 | 68.42 | 73.01 | 76.83 | 81.33 |
| Rate Reduction | | 42.86 | 38.1 | 33.34 | 28.58 | 23.81 | 19.05 |

Table 4. Recapitulation of table 1 and 2, ROUGE-SU(2) vs F-Measure, Vs chi_2 rate summary Vs rate reduction

## 5. INTERPRETATION

We tested our approach with an incremental threshold 0.00125 to see the impact of threshold of opinion on the quality of the summary and to recommend range threshold that returns good results.

ROUGE is a intrinsic semi-automatic evaluation metric based on the number of co-occurrence between a candidate summary and one or more reference summaries divided by the size of the latter. Its weakness is that it is based on references summary and neglects the original text.

The value given by ROUGE for a summary with a negligible reduction rate is high. This high value is explained by the to the increased number of co-occurrence between the candidate summary and references summary.

The F-Measure is one of the most robust metric and most used for the evaluation of classification; The F-measure is a combination of Recall and precision. For our adaptation we added to the force F-Measures an extrinsic evaluation in the beginning, and continue with an intrinsic evaluation: So this is hybrid evaluation. For automatic summary reduced rate of reduction, F-measure gives better than ROUGE assessments because it takes account of the absence of term. But unlike ROUGE, evaluating a candidate summary with high reduction ratio summary can be distorted, because FALSE NEGATIVE FN and TRUE NEGATIVE TN is the maximum which will give good result in summary generally poor (highest reduction rate leads to an increase of entropy information)

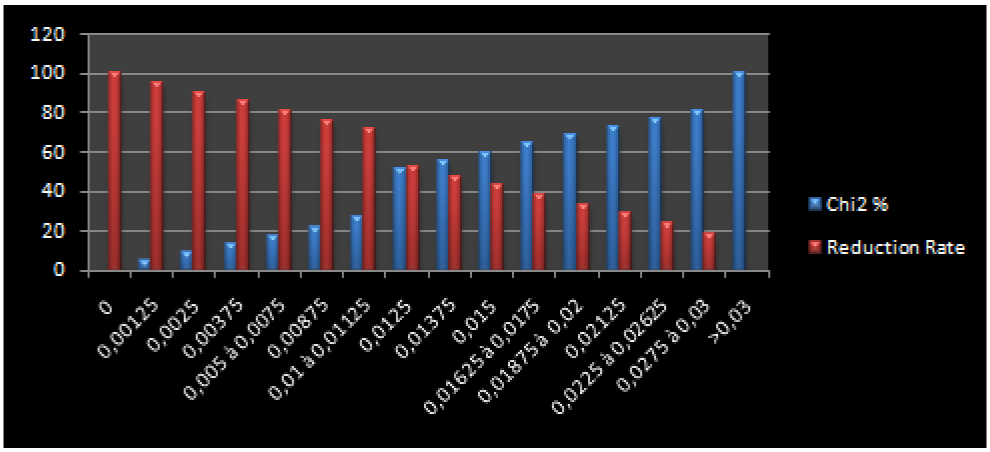


Fig. 2. Chi_2 rate candidate summary Vs reduction rate candidate summary

The precision indicated the purity of the candidate summary, while recall interprets the likeness of the candidate summary reference.

We can deduce from the above graph that the two variable: rate reduction rate Chi_2 text and have an inverse correlation, indicating that the increased number of selected sentences increases the independence original text. This indication is logic and expected, returning to table 2, we can see that in the made to retain P4, P 13 ,P18 sentences we increase the rate of chi-2 + by 25%, thanks to their strong dependence on text that is equal to 0.57 for the P4 and larger than 0.71 for P13 and P18

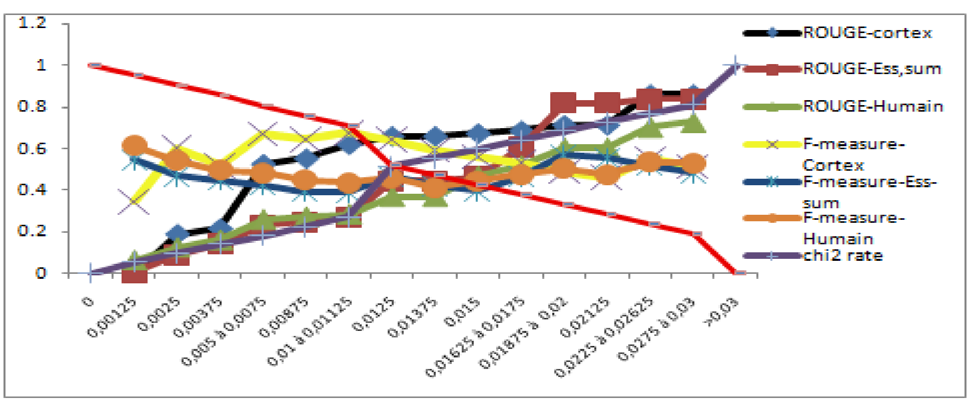


Fig. 3. ROUGE VS F-Measure (with 3 reference summary) vs chi2 rate vs reduction rate

We see the ROUGE index (black curve, red and green) is higher when the rate of reduction (red curve) is low; unlike rate chi-2 indicating a great loss to the sentence dependente with a orginale text, in this case assessment ROUGE is false, and this is the result of the high value of the co-occurrence between the candidate summary and the reference summary, the probability found that of co-occurrence between a long and a short text reference is more important than a short text summary with the same reference.

Seen on the graph that the F-measure overstates the automatic summarization which has a high rate of reduction, by against it does not overstate the summary is a low rate of reduction; is explained by the consideration of missing words (False Negative: textual untités deleted in the candidate resumé but retained by the abstract and negative reference True: textual untités deleted in the candidate and the reference summary resumé). This is a strong point of the F-Measure adapted to enable automatic extraction summaries, although it must be noted that its weakness against the very small summary and True Negative achieving the maximum value (all deleted sentences in the summary Reference will also be summarized in this summary has a high reduction ratio)

Finally, we can see all indexes used for the evaluation values are reached their optimal threshold set between 0.0125 and 0.0175, in this interval all evaluation value is good for the candidate summary.

We can see from Figure 3 and Table 4 that the selection of words that are less dependent originally the text does not improve Result, for example: between 0.0125 and 0.01375 threshold, the only difference is the selection of the fifth sentence in 0.01375 (it was not in the selected 0.0125 threshold), in Table 4 we seen Chi_2 the value of this sentence is low which is also readable on the graph not stagnation the ROUGE and a slight drop in F-measure for three reference summary. This confirms our hypothesis.

## 6. CONCLUSION AND PERSPECTIVE

In this article, we presented a new approach for the production of an automatic summary extraction based on the detection of conscience SentiWordNet.

First line, we proposed a hypothesis that will support this approach **"textual units that do not share the same opinion of the text are ideas used for the development or comparison and their absences have no vocation to reach the semantics of the abstract "**

The second line, we explain our approach to detecting and opinion proposed a flexible technique to choose the sentence that is near to the original text opinion poll threshold.

Given the results obtained, we have validated our hypothesis; and therefore this work can help solve one of the major problems of automatic summarization: the reduction of information entropy and conservation semantics.

Looking ahead, we will try to improve automatic summarization by extraction based on the detection of opinion by the application of technical and other conventional method such as detection thematic.

# 7. ANNEX

***Title : Hurricaine Gilbert***

Hurricaine Gilbert Swept towrd the Dominican Republic Sunday, and the Civil Defense alerted its heavily populated south coast to prepare for high winds, heavy rains, and high seas. The storm was approaching from the southeast with sustained winds of 75 mph gusting to 92 mph. "There is no need for alarm," Civil Defense Director Eugenio Cabral said in a television alert shortly after midnight Saturday. Cabral said residents of the province of Barahona should closely follow Gilbert's movement.

An estimated 100,000 people live in the province, including 70,000 in the city of Barahona, about 125 miles west of Santo Domingo. Tropical storm Gilbert formed in the eastern Carribean and strenghtened into a hurricaine Saturday night. The National Hurricaine Center in Miami reported its position at 2 a.m. Sunday at latitude 16.1 north, longitude 67.5 west, about 140 miles south of Ponce, Puerto Rico, and 200 miles southeast of Santo Domingo. The National Weather Service in San Juan, Puerto Rico, said Gilbert was moving westard at 15 mph with a "broad area of cloudiness and heavy weather" rotating around the center of the storm. The weather service issued a flash flood watch for Puerto Rico and the Virgin Islands until at least 6 p.m. Sunday. Strong winds associated with the Gilbert brought coastal flooding, strong southeast winds, and up to 12 feet to Puerto Rico's south coast. There were no reports on casualties. San Juan, on the north coast, had heavy rains and gusts Saturday, but they subsided during the night. On Saturday, Hurricane Florence was downgraded to a tropical storm, and its remnants pushed inland from the U.S. Gulf Coast. Residents returned home, happy to find little damage from 90 mph winds and sheets of rain. Florence, the sixth named storm of the 1988 Atlantic storm season, was the second hurricane. The first, Debby, reached minimal hurricane strength briefly before hitting the Mexican coast last month.

## REFERENCES


[1] Luhn, H. P. (1958). The automatic creation of literature abstracts. IBM Journal of research and development, 2(2), 159-165.

[2] Edmundson, H. P. (1969). New methods in automatic extracting. Journal of the ACM (JACM), 16(2), 264-285.

[3] Radev, D. R., Blair-Goldensohn, S., & Zhang, Z. (2001). Experiments in single and multi-document summarization using MEAD. Ann Arbor, 1001, 48109.

[4] Erkan, G., & Radev, D. R. (2004). LexRank: Graph-based lexical centrality as salience in text summarization. Journal of Artificial Intelligence Research, 457-479.

[5] Boudin, F., & Moreno, J. M. T. (2007). NEO-CORTEX: a performant user-oriented multi-document summarization system. In Computational Linguistics and Intelligent Text Processing (pp. 551-562). Springer Berlin Heidelberg.

[6] Carbonell, J., & Goldstein, J. (1998, August). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 335-336). ACM.

[7] Boudin, F., & El-Bèze, M. (2008). A scalable MMR approach to sentence scoring for multi-document update summarization.

[8] Mani, I., Pustejovsky, J., & Sundheim, B. (2004). Introduction to the special issue on temporal information processing. ACM Transactions on Asian Language Information Processing (TALIP), 3(1), 1-10.

[9] Farzindar, A., Lapalme, G., & Desclés, J. P. (2004). Résumé de textes juridiques par identification de leur structure thématique. Traitement Automatique des Langues (TAL), Numéro spécial sur: Le résumé automatique de texte: solutions et perspectives, 45(1), 26.

[10] CRISPINO, G., & COUTO, J. (2004). Construction automatique de résumés: Une approche dynamique: Résumé automatique de textes. TAL. Traitement automatique des langues, 45(1), 95-120.

[11] Barzilay, R., & McKeown, K. R. (2005). Sentence fusion for multidocument news summarization. Computational Linguistics, 31(3), 297-328.

[12] Yousfi-Monod, M. (2007). Compression automatique ou semi-automatique de textes par élagage des constituants effaçables: une approche interactive et indépendante des corpus (Doctoral dissertation, Université Montpellier II-Sciences et Techniques du Languedoc).

[13] Eyrich, V. A., Martı-Renom, M. A., Przybylski, D., Madhusudhan, M. S., Fiser, A., Pazos, F., ... & Rost, B. (2001). EVA: continuous automatic evaluation of protein structure prediction servers. Bioinformatics, 17(12), 1242-1243.

[14] Yu, H., & Hatzivassiloglou, V. (2003, July). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In Proceedings of the 2003 conference on Empirical methods in natural language processing (pp. 129-136). Association for Computational Linguistics.

[15] Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 417-424). Association for Computational Linguistics.

[16] Wilson, T., Wiebe, J., & Hwa, R. (2004, July). Just how mad are you? Finding strong and weak opinion clauses. In aaai (Vol. 4, pp. 761-769).

[17] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1-135.

[18] Zhang, Y. C., Medo, M., Ren, J., Zhou, T., Li, T., & Yang, F. (2007). Recommendation model based on opinion diffusion. EPL (Europhysics Letters), 80(6), 68003.

[19] Dey, L., & Haque, S. K. (2009, July). Studying the effects of noisy text on text mining applications. In Proceedings of The Third Workshop on Analytics for Noisy Unstructured Text Data (pp. 107-114). ACM.[

[20] Hatzivassiloglou, V., & McKeown, K. R. (1997, July). Predicting the semantic orientation of adjectives. In Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics (pp. 174-181). Association for Computational Linguistics.

[21] Wiebe, J. (2000, July). Learning subjective adjectives from corpora. In AAAI/IAAI (pp. 735-740).

[22] Kanayama, H., & Nasukawa, T. (2006, July). Fully automatic lexicon expansion for domain-oriented sentiment analysis. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (pp. 355-363). Association for Computational Linguistics.

[23] Esuli, A., & Sebastiani, F. (2006, May). Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of LREC (Vol. 6, pp. 417-422).

[24] Qiu, H., Xue, L., Ji, G., Zhou, G., Huang, X., Qu, Y., & Gao, P. (2009). Enzyme-modified nanoporous gold-based electrochemical biosensors. Biosensors and Bioelectronics, 24(10), 3014-3018.

[25] Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177). ACM.

[26] Kim, S. M., & Hovy, E. (2004, August). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (p. 1367). Association for Computational Linguistics.

[27] Kamps, J., Marx, M., Mokken, R. J., & De Rijke, M. (2004, May). Using WordNet to Measure Semantic Orientations of Adjectives. In LREC (Vol. 4, pp. 1115-1118).

[28] Esuli, A., & Sebastiani, F. (2006, May). Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of LREC (Vol. 6, pp. 417-422).
[29] Ohana, B., & Tierney, B. (2009, October). Sentiment classification of reviews using SentiWordNet. In 9th. IT & T Conference (p. 13).
[30] Elkhlifi, A., Bouchlaghem, R., & Faiz, R. (2011, March). Opinion Extraction and Classification Based on Semantic Similarities. In FLAIRS Conference.
[31] Liu, J., Cao, Y., Lin, C. Y., Huang, Y., & Zhou, M. (2007, June). Low-Quality Product Review Detection in Opinion Summarization. In EMNLP-CoNLL (pp. 334-342).
[32] Zhang, L., & Liu, B. (2011, June). Identifying noun product features that imply opinions. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2 (pp. 575-580). Association for Computational Linguistics.
[33] Ding, X., Liu, B., & Yu, P. S. (2008, February). A holistic lexicon-based approach to opinion mining. In Proceedings of the 2008 International Conference on Web Search and Data Mining (pp. 231-240). ACM.
[34] Pang, B., & Lee, L. (2005, June). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (pp. 115-124). Association for Computational Linguistics.
[35] Baccianella, S., Esuli, A., & Sebastiani, F. (2010, May). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In LREC (Vol. 10, pp. 2200-2204).
[36] Asher, N., Benamara, F., & Mathieu, Y. Y. (2008). Distilling Opinion in Discourse: A Preliminary Study. In COLING (Posters) (pp. 7-10).
[37] Goulet, M. J. (2007). Terminologie et paramètres expérimentaux pour l'évaluation des résumés automatiques. Traitement Automatique des Langues, 48(1).
[38] Lin, C. Y., & Hovy, E. (2003, May). Automatic evaluation of summaries using n-gram co-occurrence statistics. In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1 (pp. 71-78). Association for Computational Linguistics.
[39] Das, D., & Martins, A. F. (2007). A survey on automatic text summarization. Literature Survey for the Language and Statistics II course at CMU, 4, 192-195.
[40] Louis, A., & Nenkova, A. (2009, August). Automatically evaluating content selection in summarization without human models. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1 (pp. 306-314). Association for Computational Linguistics.
[41] Moh'd A Mesleh, A. (2007). Chi square feature extraction based SVMs Arabic language text categorization system. Journal of Computer Science, 3(6), 430-435