**DATA OF VIETNAMESE SENTENCES**

**EXTRACTING FEATURES**

**TAGGING WORDS**

**VIETNAMESE MOTIONAL DICTIONARY**

**PREPROCESSING DATA**

**CLASSIFYING BY SVM**

**THE TRAINING DATA**

**THE DATA HAS BEEN CLASSIFIED**

Extracting features is divided into two sections: The first is subjective classfication and the second is sentiment classfication.

* Module of extracting features to subjective classfication

***Input:*** tagger file và vietnamese emotional dictionary.

***Output:*** file was structured in vector format, each line is 01 feature vector.

**The applicable actions:**

For each sentence in the data, extract the following features:

*1) The amount of words in sentences.*

*2) The summary of emotinal value of adjectives.*

*3) The summary of emotinal value of adverbs.*

*4) The summary of emotinal value of nouns.*

*5) The summary of emotinal value of verbs.*

*6) The summary of emotinal value of sentence.*

*If this sentence is a condition sentence or a question, return 0.*

*However, return total of 4 features that is 2nd,3rd,4th,5th.*

Return feature vector

* Module of extracting features to sentiment classfication

***Input:*** tagger file và vietnamese emotional dictionary.

***Output:*** file was structured in vector format, each line is 01 feature vector.

**Các thao tác áp dụng:**

For each sentence in the data, extract the following features:

*1) The summary of emotinal value of adjectives.*

*2) The summary of emotinal value of adverbs.*

*3) The summary of emotinal value of nouns.*

*4) The summary of emotinal value of verbs.*

*5) The summary of emotinal value of sentence: total of 4 features that is 2nd,3rd,4th,5th.*

*6) The summary of emotinal value which depends on the intensification.*

*7) The summary of emotinal value which depends on the contrasting-linked words.*

*8) The summary of emotinal value which depends on the imperfect words.*

*9) The emotional value of a positive sentence.*

*10) The summary of emotinal value which depends on the negative words.*

Return feature vector

Compared with other researches related to our topic, our research has been features that were selected, were tested many times. **Also, other than the use of the combination of 5 different emotional dictionaries, we only use one the English emotional dictionary which was provided the best results to build the vietnamese emotional dictionary and add a lot synonyms equivalent and common words to this Vietnamese dictionary.** In addition, emotional value in the vietnamese emotional dictionary from -5 to +5 will help assess detail the emotional value of sentence. It helps improve accuracy in sentiment analysis for Vietnamese comments.