|  |  |  |
| --- | --- | --- |
| amblem | **ÇANKAYA UNIVERSITY**  **Computer Engineering Department**  **CENG 407–408 Senior Project (Innovative System Design and Development) Proposal Form** |  |

This form should be used for all CENG 407 – 408 Senior Project Proposals. A topic can be jointly proposed by faculty, company and/or student, signed by at least one of the partners.

**Part I. Project Proposer**

|  |  |  |  |
| --- | --- | --- | --- |
| **Names (supervisor, company, student) and organizations** | Supervisor: Murat SARAN  Co-advisor: Ulaş GÜLEÇ  Company: Turkish Standards Institution (tse).  Students: Merve TEMÜR and Naz BAYRAKTAR | **Mobile** |  |
| **E-Mail** |  |
| **Signature** |  |
|

**Part II. Project Information, to be completed by the proposer (Faculty, Student and/or Company)**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Starting Term** | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 2 | 0 | 1 | 5 | / | 2 | 0 | 1 | 6 | | ▢ Fall | ▢ Spring |
| **Title** | | Sentiment Analysis on Turkish Sentences | | |
| **Description (extra sheets can be added)** | | | | |
| Sentiment analysis has been an active research area for quite some time. Unfortunately, most works are specific to the English language. In this project, we present a framework for unsupervised sentiment analysis in Turkish text documents., we will get tweets and location information throughout Twitter ( is a social networking and microblogging service that allows users to post real time messages, called tweets.), and we will keep this information in a database.We will do sentiment analysis on Turkish sentences according to the tweet contents in the database.For example, positive / negative and neutral emotions, hashtags (Users usually use hashtags to mark topics. This is primarily done to increase the visibility of their tweets) and content (specific name, organization name, etc.) The aim of the project is to develop a new algorithm to do sentiment analysis on Turkish sentences In Turkish, most of the negation is done by the negation suffix in the predicate, which usually affects the whole sentence The added suffixes may change the polarity of words, for instance, adding (“-ma”) to the end of word will change the polarity of that word and so on. Also we will study effect of adjectives and effective words which have deferent mining’s. Moreover, Turkish has several letters that are missing in English (“ç”, “ı”, ”ğ”, “ş”, “ü”, ”ö”) and, in informal writing on the Web especially on Twitter, people tend to substitute these Turkish letters to the closest ASCII English letters which are(“c”, “i”, “g”, “s”, “u”, “o”). This creates a complication in identifying the words. So, we decided to create a sentiment analysis framework taking into account the above mentioned differences. For this purpose, there are some modules to correct the sentences of the input text that we decided to use in our project such as Sentence extractor, ASCII character converter, Morphological analyzer, Negation handler, and Polarity predictor. These modules could help us to analyze the Turkish sentences and create the algorithm about it. We will also have Microblog data like Twitter, on which users post real time reactions to and opinions about “everything”, poses newer and different challenges. Also, people may use some emoticons (These are facial expressions pictorially represented using punctuation and letters; they express the user’s mood) in their tweets, they may use tweets ending in positive emoticons like “:)” “:-)” as positive and negative emoticons like “:(” “:-(” as negative so we try to clear these at the same time catch the emoticons and classified they to categories (negative, positive and neutral) in Database. Also, we will take the location of the public tweets that we can access, also we will use CSS (Cascading Style Sheets) that helps us to design our user interface when we design the map of Turkey with the emotions on all of the cities in turkey. Our map will have colors that are divisible into three major colors (Red for sad mood, Green for happy mood, Yellow for neutral mood). As a result, we decided to design an algorithm to sentiment analyzes on Turkish sentences in Twitter using windows form application. | | | | |
| **Justification** | | | | |
| **Novelty** | The aim of the project is to develop a new algorithm to do sentiment analysis on Turkish sentences. In the scope of the project, Turkish tweets in the determined areas will be analyzed by using this algorithm to figure out the opinions of the users about the selected topics. | | | |
| **Complexity** | Although there are lots of studies, which can analyze English sentences, in the literature, there are not enough studies, which can analyze Turkish sentences, in the literature. The complexity of the project is to gain a new algorithm, which provides the researches to do sentiment analysis on Turkish sentences. To do this operation, the literature will be searched in detailed and the suitable algorithm will try to be found. After that, the found algorithm will be adapted for doing sentiment analysis on Turkish sentences. | | | |
| **Constraints:** economics, sustainability, environment, ethics, security, health, social and political issues, | Sometimes Twitter can be banned from the government so, our project will be constraint. | | | |
| **Risks involved** | We want to design our project for sentiment analysis on Turkish sentences, but unfortunately we may not have enough time to complete the algorithm that is needed for this project. | | | |