William Angell
Drew Verlee
David Hui
Dominic Fiore

Sentiment Analysis Project

1. Project overview
    a. What is sentiment analysis?
        i. Sentiment analysis is the process of assigning objective judgement values to a subjective piece of text
        ii. It's a part of natural language processing, a branch of artificial intelligence that deals with language
        iii. It's important in a world where the amount of produced information is overwhelming manual analysis processes
    b. Sentiment analysis breakdown
        i. Input: Piece of text (in this case, Amazon reviews and its rating)
        ii. Output: A gradient score indicating the positivity/negativity of the text
        iii. Examples
            1. Sentence: "Every day is a good day"
                a. Potential score: 1
            2. Sentence: "Colorless green ideas sleep furiously"
                a. Potential score: 0
            3. Sentences: "Very bad dog"
                a. Potential score: -2
        iv. Process
            1. Tokenization
            2. Lemmas
            3. Parsing/analysis
            4. Score
2. Project purpose, scope, objectives
    a. Analyzer will be integrated into a previously existing system, but our project is contained within itself: interfacing with other systems is not within scope
    b. Create a sentiment analyzer for Amazon reviews that will perform well (in this case, "well" is a metric decided by us)
    c. Create a utility that is well-designed and documented so that it is used by others within the academic community
3. Team organization
    a. Team Leader: Will Angell
        i. Planning the project so that it will be delivered by the due date
    b. Testing facilities: David Hui
        i. Tests deliverables against target scores/standards, assess the overall stability of the program
    c. Documentation: Dominic Fiore
        i. Reads/writes documentation, produces user manual
    d. Operations: Drew Verlee
        i. Handles servers and other resources

William Angell
Drew Verlee
David Hui
Dominic Fiore

4. Problem resolution policies
    a. Step one: Consult the group
        i. If one runs into a problem, the first step is always to consult other group members to see if they have a solution
    b. Step two: Consult an external source (professor, website, etc.)
        i. If there are any doubts within the group, one could look to an external source for answers
    c. Step three: Consult the client
        i. It's important to consult the client for issues that can't be resolved using the group/an external source
5. Project plan
    a. Scheduling
        i. First prototype: 9/26
        ii. Second prototype: 10/10
        iii. Third prototype: 10/24
        iv. Final prototype: 12/10
    b. Development strategy
        i. Rapidly develop first prototype
            1. Research projects are usually small in terms of code, but large in terms of the ideas used
            2. Sentiment analysis is a very easy process to describe, but it's difficult to master
        ii. Continually improve upon the first prototype to attempt to hit better accuracy/precision scores
            1. Scores are decided through manual group analysis of individual pieces of text
            2. Look for other ways of testing the analyzer
6. Configuration management plan
    a. Not very applicable to this project, however configurations will be managed via command-line options/flags
7. Technologies
    a. Python
        i. Cross-platform
        ii. Easy to learn/read
        iii. Popular in academia
        iv. Easy access to many natural language processing libraries
    b. Stanford Parser, NLTK, SentiWordnet