

## NLTK Sentiment Analysis in Satire

### Project Overview

For this project, the code was designed to pull back Oscar Wilde's *The Importance of Being Earnest* and assess the accuracy with which the Natural Language Tool Kit (NLTK) package could classify the text. This particular text was chosen because NLTK's sentiment analysis capabilities are limited to determining a positive or negative tone of text, and Wilde's work is known for its satire, which is difficult to gauge in this way. It is also possible that satire itself could entirely invert the tone of text, further making Wilde's text an interesting sample.

### Implementation

To assess the sentiment of the text, the source information had to go through a series of steps, essentially pulling the information from Project Gutenberg, adapting it to be processed, and picking out a few samples. Though it is not included in the code, using outside software it was determined which samples were to be used for this, so they were not picked arbitrarily. There was some work to be done with the original text to remove the Project Gutenberg notes and other text not written by Wilde and additional code to adapt the play to a format where the character names are not stated before each of their lines. The heavy lifting is done by the NLTK package's Vader sentiment analysis.

When writing the code, it could have been organized without functions, but instead it was broken down into function that can easily be edited to work for other source texts. The source could be adapted to any other play and the only thing that would need to be changed are the start / end words and the character list. Organizing the text like this allows for the code

to be more easily read as well when edited. There was also opportunity to split the text manually based on punctuation, but for the sake of continuity the NLTK package's use, the NLTK method for tokenization was used instead of manual string methods.

## **Results**

For the text chosen, the samples all had some element of irony or satire. This was selected for the aforementioned reason of exploring the viability of sentiment analysis tools to properly gauge what is going on in a passage of text. The original point of this analysis was to see if the use of satire could be classified as negative, neutral, or positive. Looking at the results of the sentiment analysis, it can be seen that there is no clear pattern in any of the exhibits aside from the presence of a neutral tone. However, it is not prominent enough to predict satire based off the neutral tone.

The sensitivity analysis itself shows that the NLTK method fails to understand satire. This makes it clear that the NLTK sensitivity analysis takes all of its input literally. Exhibit 1 is a clear example of this, as saying "It is awfully hard work doing nothing" is clearly a joke, but words like "awfully" and "nothing" push the output to the negative. Satire is shown in Exhibit 2.0 as well, as it is a play on the character of the line's speaker, however, once again, NLTK shows it as negative.

Most interestingly, NLTK's tools show not just error with satire and comedy, but a disconnect between its method of analysis and human speech. Human speech is interpreted with context. The entirety of the book is seen with positive tone (Exhibit 5), which shows how the more context is provided in the input, the better the analysis. This can also be seen in how

both Exhibit 2.1 and 3.1 are considered fully neutral, but when included with the preceding sentences are given more depth. Exhibit 4's comedic value is only understood once finished with the whole book, so it is also not a correct value. The issue with this is that the code, even if it could properly detect sarcasm, would have to actively and dynamically discern which, if any, preceding lines are necessary to understanding the true tone of the line being evaluated.

## **Reflection**

This process accomplished what it set out to do. It uncovered the value and capabilities of sentiment analysis. The information in this process will be able to inform the use of sentiment analysis in the future. It also means that any use of sentiment analysis must be supervised. If analysis were to have been conducted on all sentences from the book, far too many would have to be contextualized with preceding sentences, the number of which would need to be determined on a per-case basis. The approach of using a handful of excerpts was useful, though maybe an additional source could have been used in order to verify the claims being made apply to all text inputs, and not just Wilde's style. While not necessary, this would reinforce and solidify the claims made about the capabilities of NLTK sentiment analysis.

Appendix

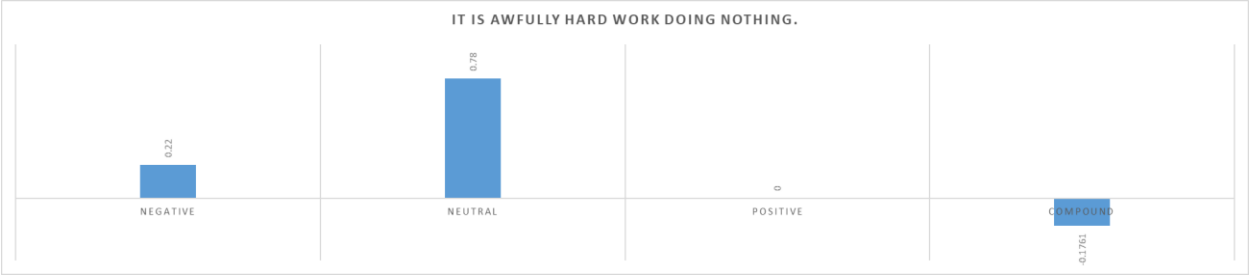


Exhibit 1

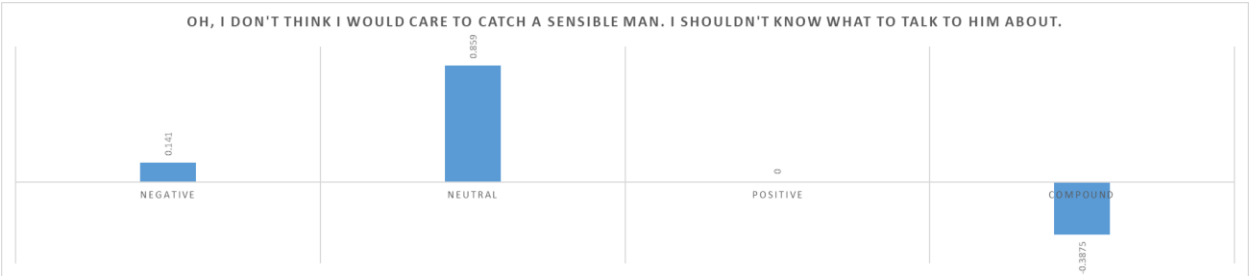


Exhibit 2.0

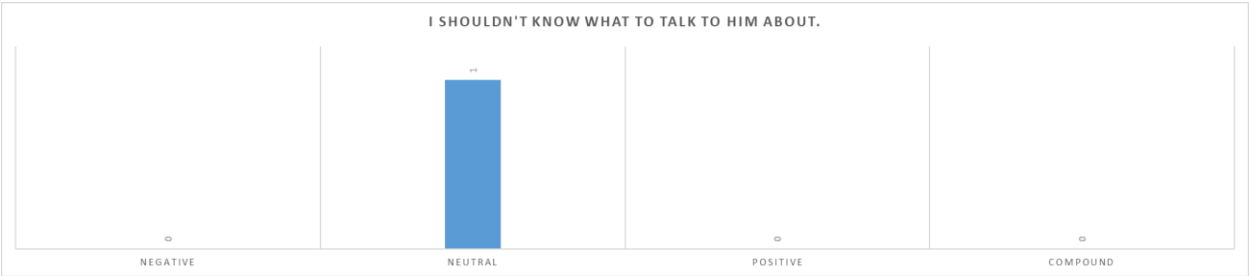


Exhibit 2.1

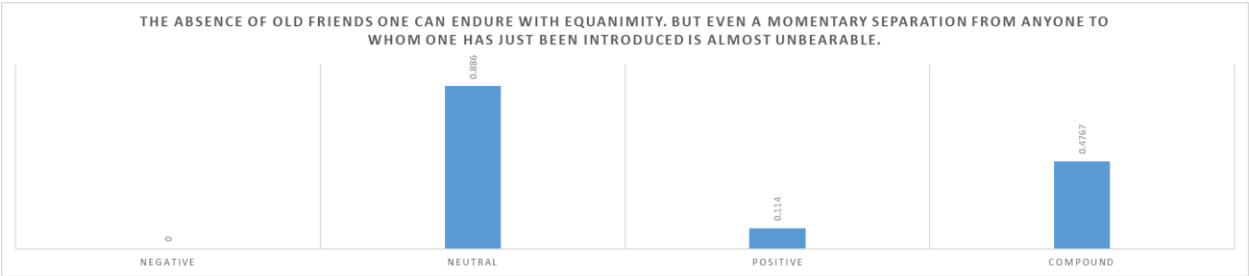


Exhibit 3.0

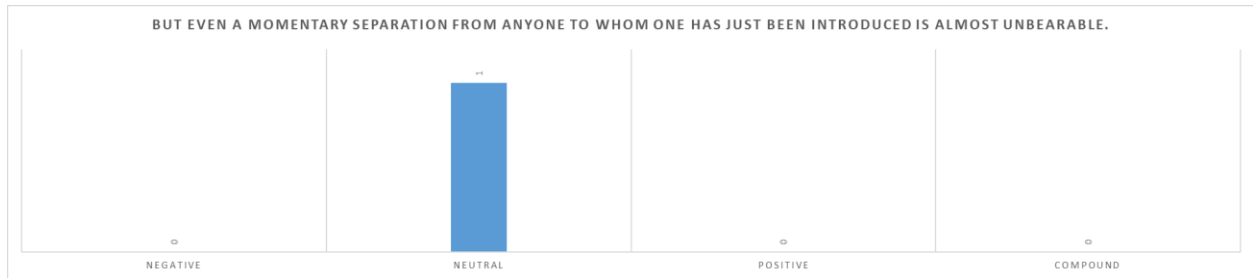


Exhibit 3.1

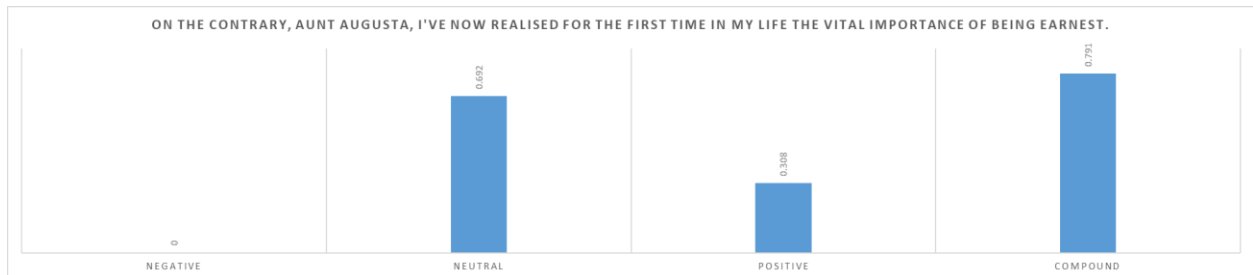


Exhibit 4

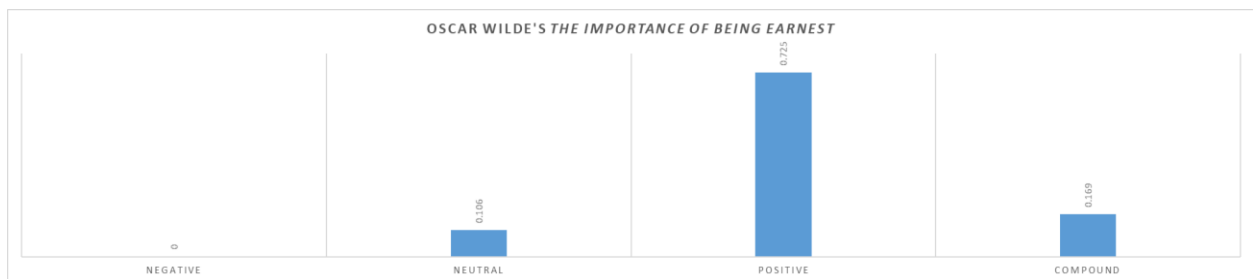


Exhibit 5

	Negative	Neutral	Positive	Compound
Exhibit 1	0.22	0.78	0	-0.1761
Exhibit 2.0	0.141	0.859	0	-0.3875
Exhibit 2.1	0	1	0	0
Exhibit 3.0	0	0.886	0.114	0.4767
Exhibit 3.1	0	1	0	0
Exhibit 4	0	0.692	0.308	0.791
Exhibit 5	0	0.106	0.725	0.169