Literary Analysis of Homer's The Iliad using NLP Methods

Cultures, Civilizations, and Ideas Project

Selim Firat Yilmaz

Contents

1. The Iliad's Analysis

1.1 Objective

As a creative project, I plan to inspect The Iliad of Homer using Natural Language Processing (NLP) methods. NLP is a subfield of Artificial Intelligence which studies the computers' understanding of human language. Main goal of this project is to exhibit some techniques of artificial intelligence and visualization for better understanding of the text. After each of the results, I will answer the question "so what?". In other words, I will explain the reasoning behind numbers referring to the text.

1.2 Motivation and Relevance to the Humanities Course

Use of emerging technologies like artificial intelligence in the humanities topics would provide a bridge between ancient and contemporary worlds. Incorporating the applications of artificial intelligence would provide many benefits for both of the sides. For example, as we gain a better understanding of the classics, we are having an excellent application area. Also, I believe classics may be key to the artificial intelligence understanding of emotions. Such understanding would improve AI capabilities (which was my ENG 102 topic) as emotions promote our cognitive skills like face recognition (J. Megill, "Emotion, cognition and artificial intelligence", pp. 189–199)

Also, When I proposed the project, my main goal was to trace the text The Iliad in the context I understand, I think, I study, and I love. This project was very improving for me in terms of both Computer Science and Science of & Humanities. I realized that there are many university departments named "Digital Humanities", and journals like "Journal of Digital Humanities". Digital humanities field is tightly connected with artificial intelligence and visualization fields.

1.3 Syntactic Analysis

In this section, frequencies of words, frequencies of content words will be visualized and explained.

1.3.1 Most Frequent Words

Figure 1.1: Most Frequent Words' Frequency Plot

Figure 1.2: Words Cloud

Table 1.1: Table of the most frequent words

the	and	of	to	his	in	a	he	you	with	on	all
10569	5510	3542	3525	2959	2588	2515	1903	1860	1669	1384	1377

Although many of them are common in many writings, the dominance of male-oriented words like "his", "him", "he" is obvious. Male supremacy in Iliad shows itself in the numbers, too. Above all, it is very probable that The Iliad will not pass the Bechdel Test.

1.3.2 Most Frequent Content Words

Despite the fact that this still is syntactic analysis part, we are advancing towards the text's semantics. I extracted content words stripping out function words to start studying the meaning of the text.
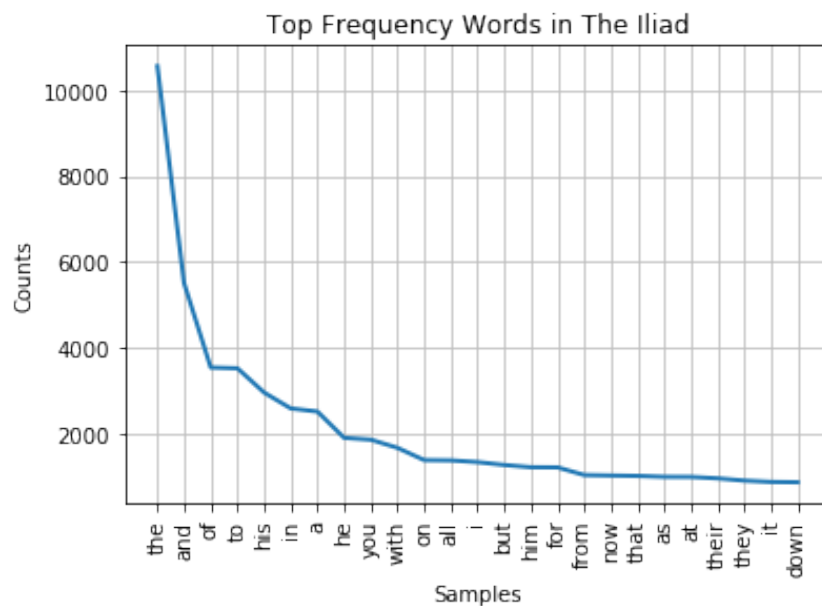

Figure 1.3: Most Frequent Content Words' Frequency Plot

Figure 1.4: Word Cloud

In this part, we obtain many of the character names and main themes. Due to the lack of enough time, I couldn't obtain and visualize them separately. Distribution of themes and characters would give many insights about the text.

Table 1.2: Table of Content Words

son	men	back	man	hector	zeus	achilles	trojans	god	ships	one	never	gods	heart
691	672	646	641	551	549	525	402	401	398	389	375	332	320

1.3.3   Repeating Phrases

Repeating Phrases are the most appearing consecutive groups of words in the text. n-gram means the group that contains these consecutive n words. Repeating phrases are listed below: In 2-gram table, all of them are functional words except the "son". In the many

Table 1.3: Repeating 2-grams (collocations)

in the	559
of the	553
to the	498
from the	346
on the	344
son of	276
all the	265
for the	227
at the	223

characters are also called through their father like "son of Atreus", "son of Aronus", "son of Zeus", "son of Peleus" and many more... This is so common that the word "son" appears as much as the functional words like "the".

Table 1.4: Repeating 3-grams

lord of the	59
the lord of	54
the god of	53
down to the	48
out of the	44
back to the	43
son of cronus	42
of all the	41
the old man	41

In 3-grams, again we see that many words are male oriented like "lord", "son", "man".

In the rest of n-grams, we see many repetitive scenes between different songs in the book. Also, we obtain how some of the characters are called.

Table 1.5: Repeating 4-grams

of the war cry	28
lord of the war	27
the house of death	23
the men of troy	21
to the house of	19
the swift runner achilles	18
the lord of men	18
lord of men agamemnon	18
the lord of the	18

Table 1.6: Repeating 5-grams

to the house of death	17
down to the house of	12
the lord of men agamemnon	12
the lord of the war	12
of zeus whose shield is	11
the father of men and	11
father of men and gods	11
diomedes lord of the war	9
marshal atrides lord of men	8

Table 1.7: Repeating 6-grams

the lord of the war cry	12
the father of men and gods	11
diomedes lord of the war cry	9
and the strong force of fate	7
of zeus whose shield is thunder	6
the old man of the sea	6
when they had put aside desire	6
they had put aside desire for	6
had put aside desire for food	6

1.4 Semantic Analysis

1.4.1 Battle Analysis

"O my son, my sorrow, why did I ever bear you?
    All I bore was doom...
    Doomed to short life, you have so little time" (Homer 1.492-496).

Even the legend Achilles is doomed to short life in just the beginning of the book. Who would expect not many deaths in the rest?
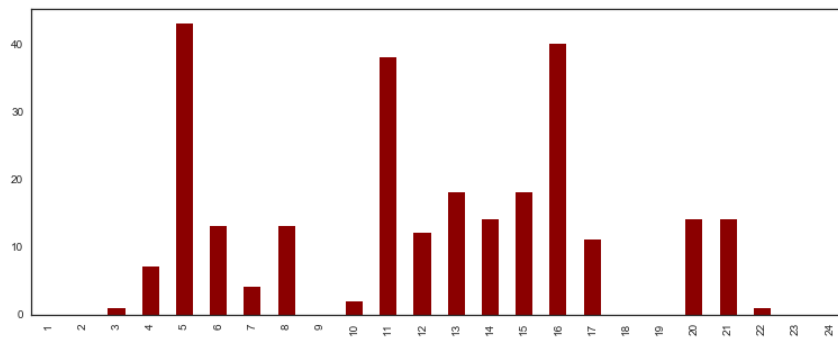

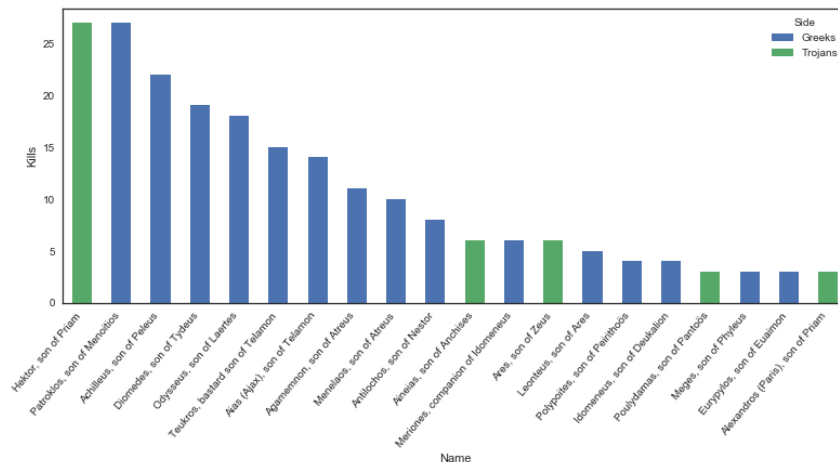Figure 1.5: Plot of the Number of Peoples killed for each Song.

Figure 1.6: Kill Statistics for different characters in The Iliad.

The surprising fact is that there is a tie between Hector and Patroclus in terms of the number of people killed by them. This might be intended by the author Homer.

1.4.2 Sentiment Analysis

This is the section where modern artificial intelligence techniques is used in addition to the natural language processing.

The huge decrease in the mood towards negative in song 5 is not surprising. As we seen above, song 5 is the bloodiest song in The Iliad.
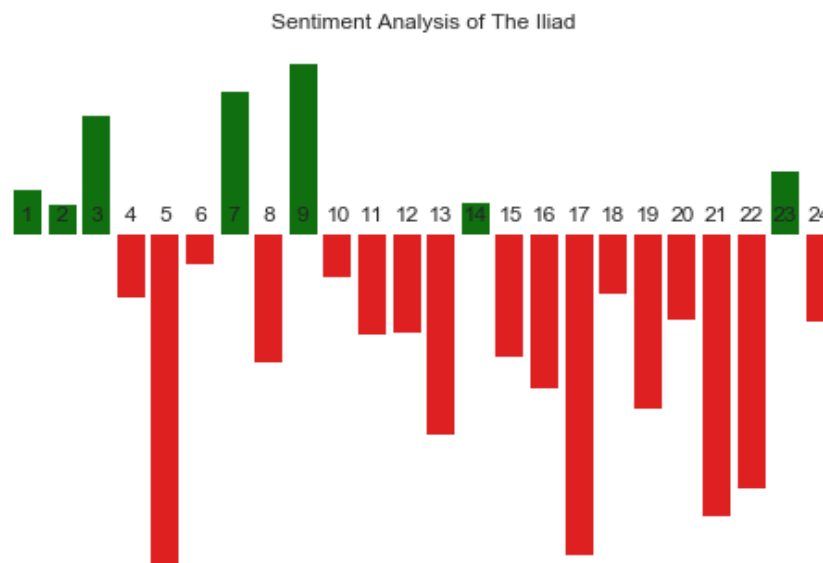
Figure 1.7: The green colored bars represent positive mood and red colored bars represent the negative mood for different books in The Illiad.

The decrease in the mood towards negative in son 25 is due to Hector's Burial. Mortals are doomed to suffer in the Iliad:

"Such is the way the gods spun life for unfortunate mortals,
That we live in unhappiness, but the gods themselves have no sorrows" (Homer, Iliad 24.525-526).

1.4.3 Similarities Between Songs

We can interpret this plot as the following: The books that are closer more similar compared to others. There are some clustering(grouping) of the books. First 12 books are clustered on bottom left of the graph as well as the last 12 are clustered on top right of the graph. There are 2 exceptions to these cases which is normal. Also, some (couldn't find the source again) claim that the last 10 books are written by another person, not Homer, so this kind of grouping might be an evidence to this claim. However, further research is required for this subject.
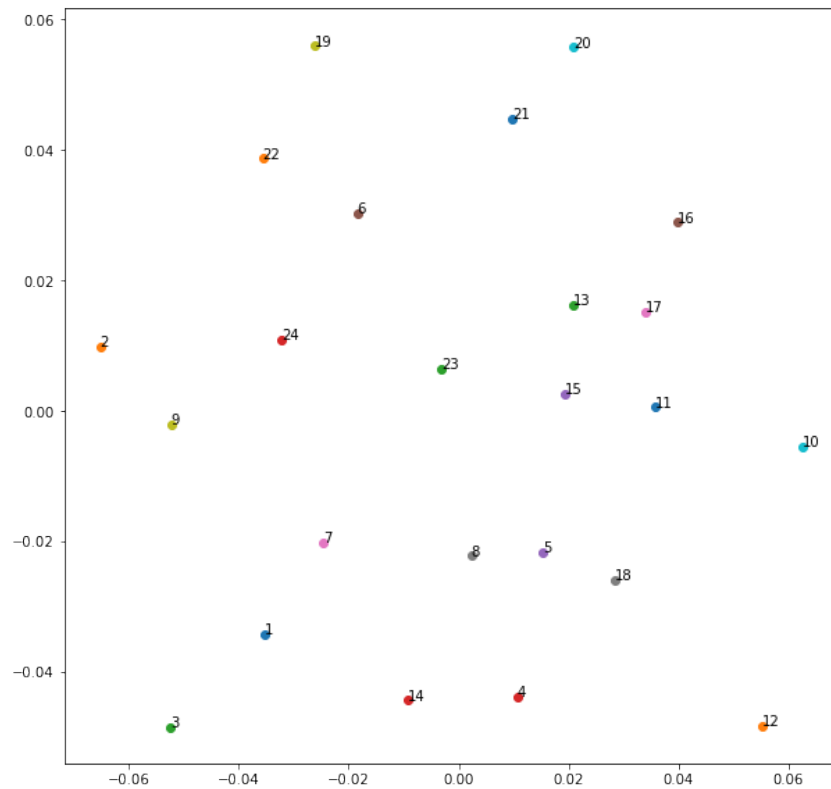
Figure 1.8: Similarities Between Songs in The Iliad

1.5   Conclusion

This project was intended to explain the reasons behind numbers of The Iliad. Inspiration for this project was "Things Hidden Behind The Numbers of İnce Memeds" (2015) which is the work of Professor Fazli Can in Bilkent University. However, unlike this work, the computer science related terms is used as least as possible since this project is going to be evaluated by Dr. Sjoerd Levelt in Bilkent University whose expertise is not computer science.