THE GOOD, THE BAD & THE UGLY

SHAKESPEARE DATASET

Last Name: Rashad

First Name: Amr Mohamed Nazih Mohamed

ID: 991043

Aim Of The Project

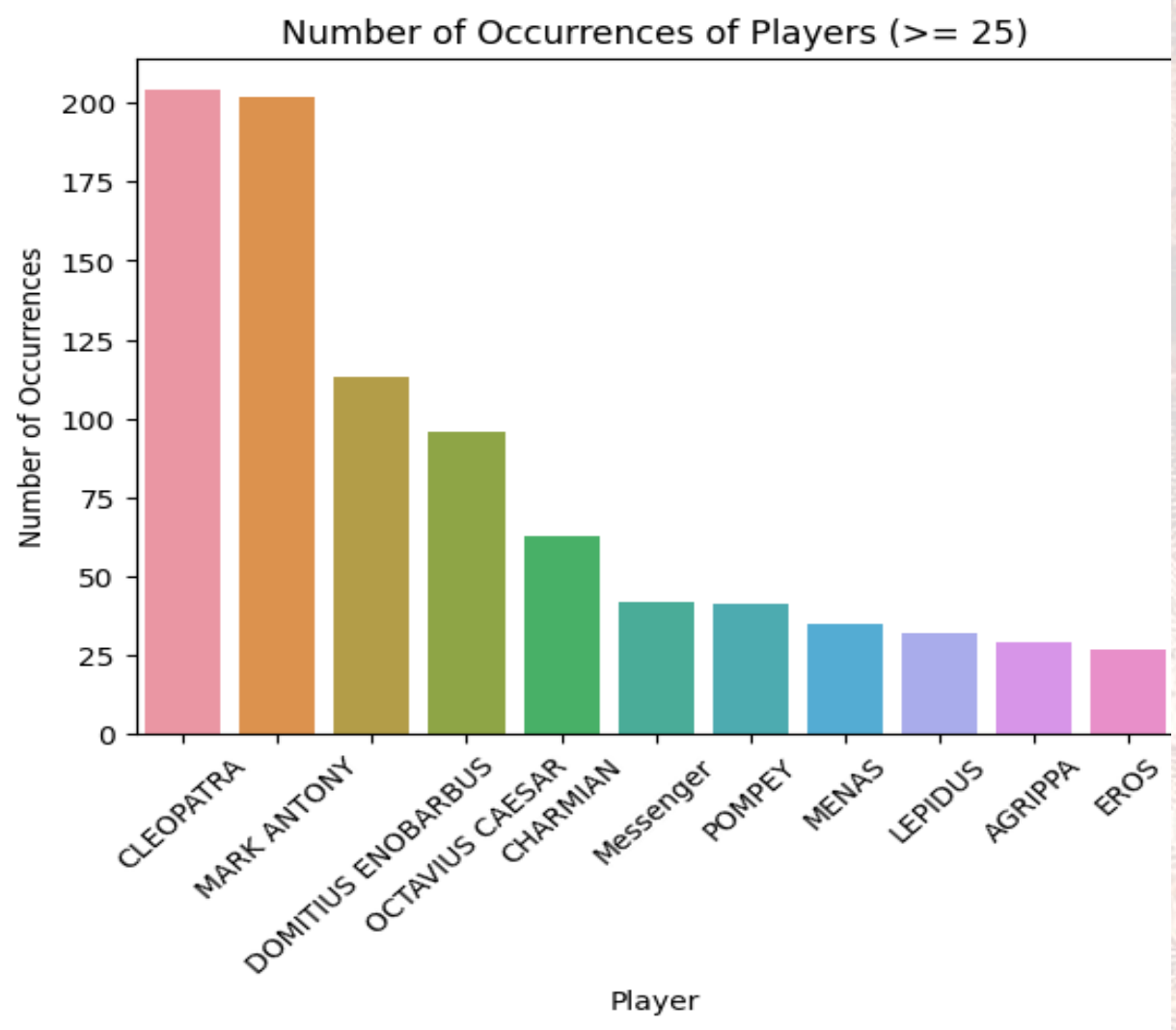
- The aim of this project is:
  - Perform emotion detection techniques on the given Shakespeare dataset.
  - Observe the emotional evolution of each character throughout the play.
  - Build an emotional profile for each character that would describe the various types of dominant emotions a character feels.
  - Representing character interactions through network graphs, visualizing the dominant emotions expressed during these interactions

Dataset Description

- The selected dataset is a csv file containing several of Shakespeare plays. Only one play is considered in this project which is: ‘Antony & Cleopatra’.
- The csv file is composed of the following columns:
  - Player (i.e., character)
  - Player Line
  - Play (as mentioned earlier the play name is ‘Antony & Cleopatra’)
  - ‘ActScenceLine’ (for example 1.2.3 means Act 1, Scene 2, Line 3)

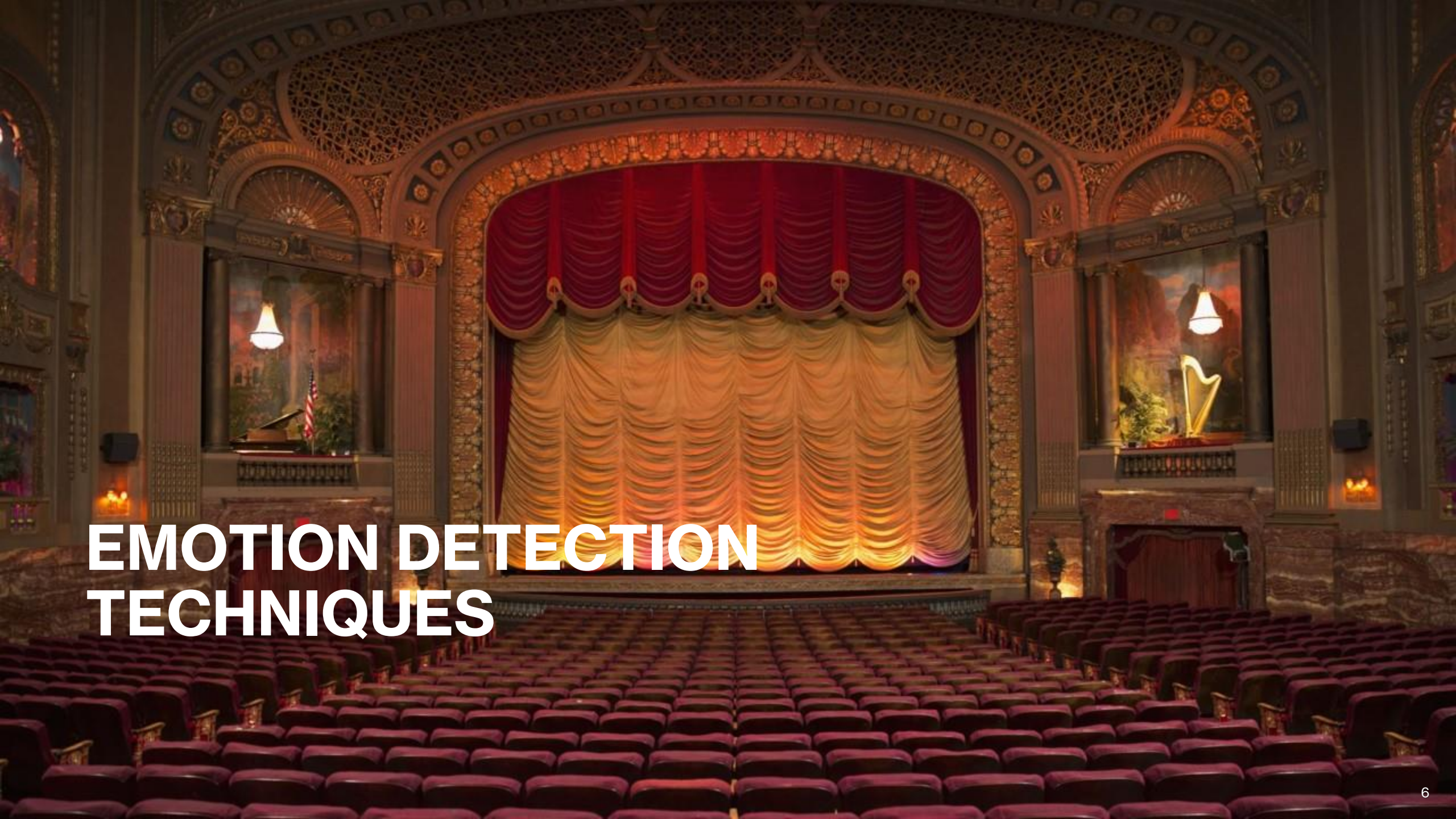
Emotion Detection in Text Mining:

- Emotion detection can be implemented in various ways to determine the sentiment behind a specific text.
- A simple implementation of this technique involves binary classification, where a text is classified as either having a positive or negative sentiment.
- Since we are working on theatrical play dataset, it should be appropriate to consider a more complex approach to describe the different emotions each character feels when reciting his/her lines and to create a more complex emotional profile for each character.
- That is why I have decided to use two different models to perform emotional detection on this dataset and see which one gives a more appropriate output.



Identifying The Main Characters:

- The main characters have been identified by introducing a threshold
- By Performing a bar plot, we can determine that we have 11 players that are considered as main characters.



EMOTION DETECTION TECHNIQUES

Text2Emotion:

- Text2Emotion is a python package that is used to classify textual data into different types of emotions.
- The output of this package is in the form of a dictionary, where each emotion is a key, and the value is between 0 and 1.
- There are 5 basic emotion categories represented in this package, which are the following: Happy, Angry, Sad, Surprise and Fear.
- We only consider the emotion with the highest value, which is 1.0, to be the dominant emotion that would represent the emotion of a given line. If there aren't any emotion represented with 1.0 as a value, then we would categorize the emotion of the given line as 'Neutral'.

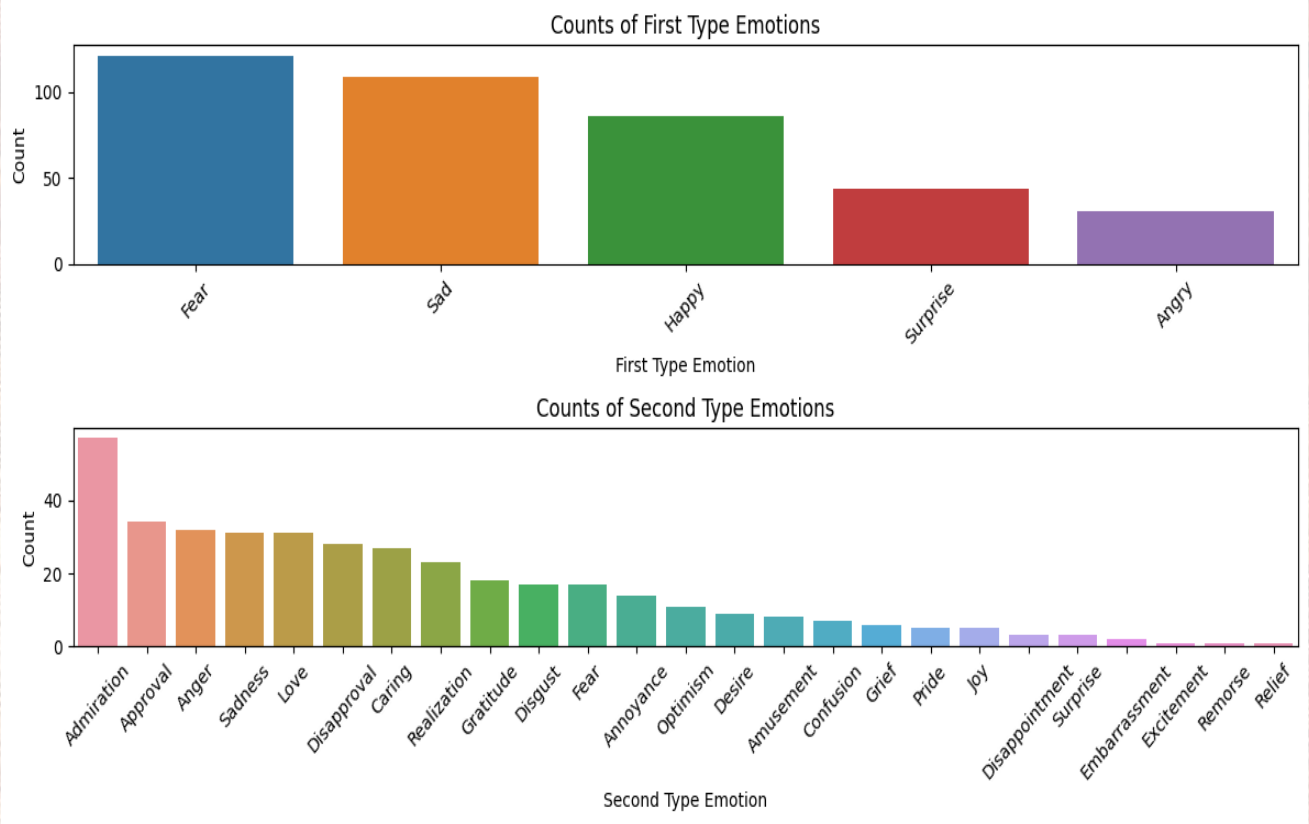
	Play	Player	PlayerLine	Act	Scene	Line	StartOfSequence	Happy	Angry	Surprise	Sad	Fear
0	Antony and Cleopatra	CLEOPATRA	If it be love indeed, tell me how much.	1	1	15	True	1.0	0.0	0.0	0.0	0.0
1	Antony and Cleopatra	MARK ANTONY	There's beggary in the love that can be reckon'd.	1	1	16	True	1.0	0.0	0.0	0.0	0.0
2	Antony and Cleopatra	CLEOPATRA	I'll set a bourn how far to be beloved.	1	1	17	True	0.0	0.0	0.0	0.0	1.0
3	Antony and Cleopatra	MARK ANTONY	Then must thou needs find out new heaven, new ...	1	1	18	True	0.0	0.0	1.0	0.0	0.0
4	Antony and Cleopatra	MARK ANTONY	Grates me: the sum.	1	1	20	True	0.0	1.0	0.0	0.0	0.0

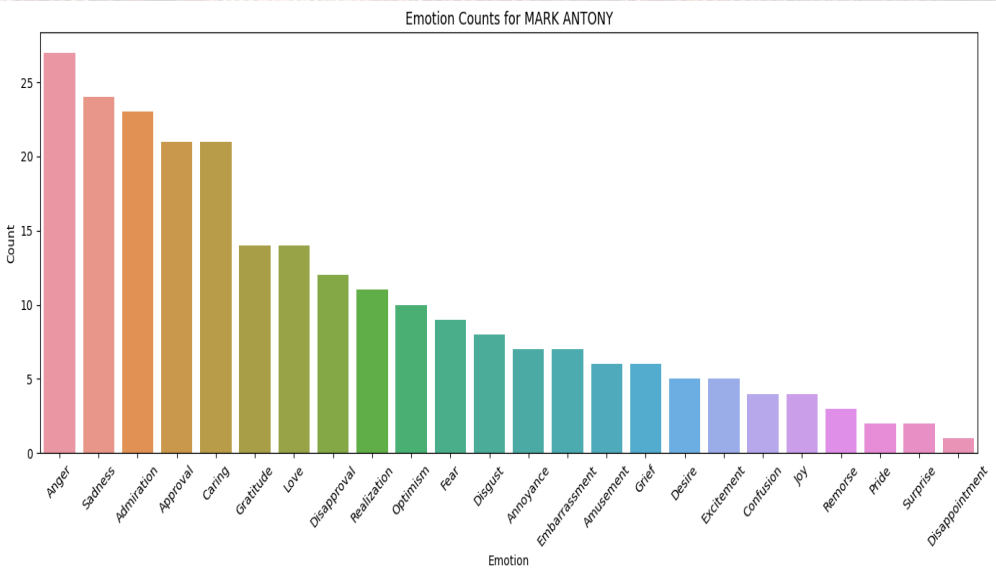
EmoRoBERTa:

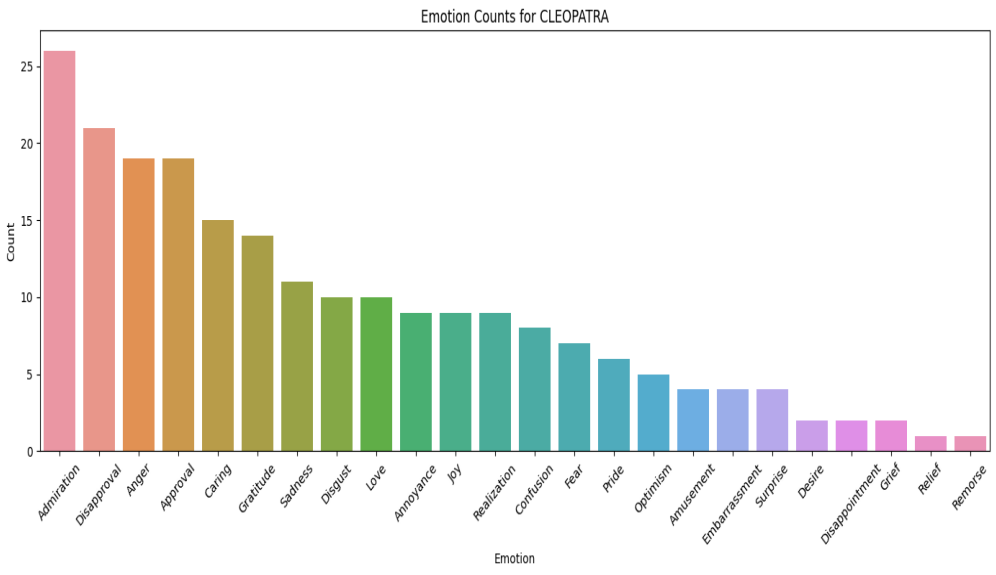
- EmoRoBERTa is a variant of the RoBERTa model designed for emotion recognition.
- EmoRoBERTa model is well-suited for certain applications that would require analyzing the emotions conveyed in textual data.
- There are various types of emotions represented by this model such as:
  - Anger
  - Admiration
  - Caring
  - Sadness
  - Love

Comparing Two Models:

- We can clearly see that this model offers a much wider range of emotions compared to the Text2Emotion package.
- That's why this model appears to be more suitable for implementation on textual data from a theatrical play, as it can provide a more detailed emotional profile for each character in the play.





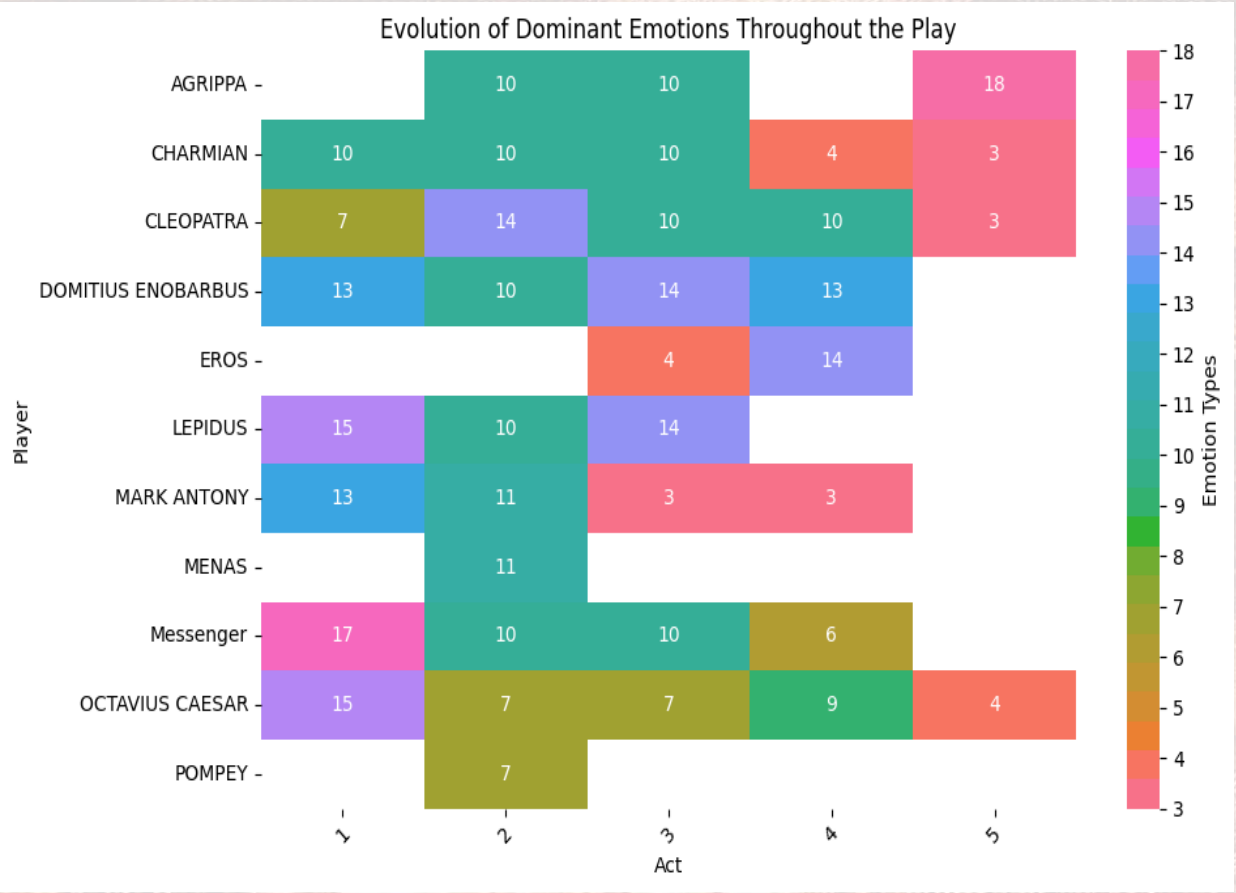


Character Emotional Profiles in the Play:

- Focusing on Cleopatra and Antony in the following plots:
- We can see that both characters express a variety of emotions, however some of the main emotions that are expressed are:
  - Admiration
  - Anger
  - Approval
  - Caring

Character Emotional Profiles in the Play (Cont'd):

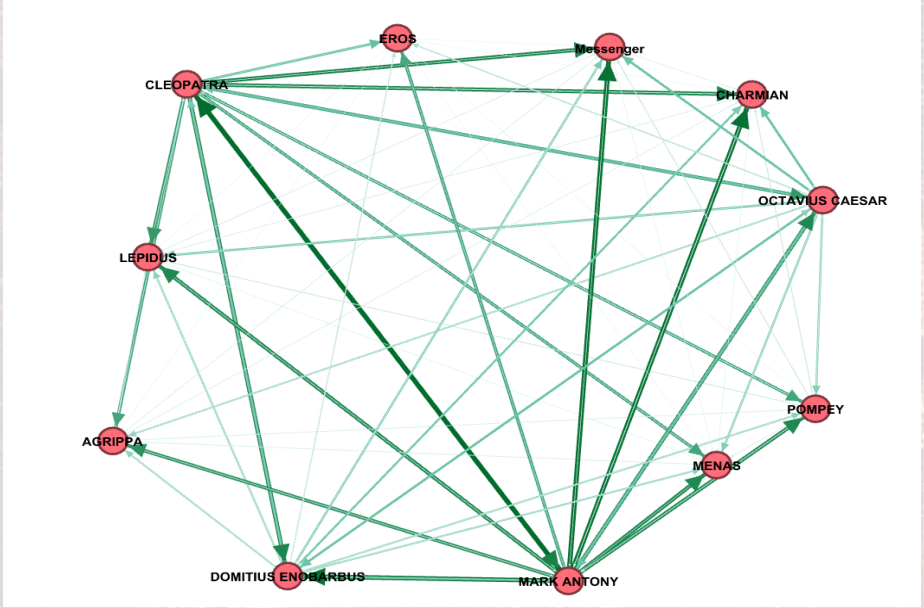
- Since the main characters express numerous types of emotions, I decided to investigate the dominant emotion of each character in each act and observe his/her emotional evaluation throughout the play.
- We can observe that 'Admiration' is the most frequently expressed emotion by all the main characters throughout the play. While this may not provide a comprehensive representation of each character's emotional profile, it does offer insight into the primary emotions they express throughout the entire play.
- Additionally, if we are familiar with the play's storyline, we can make further observations. For example, in the final act, when Antony's death is known, we can see that the characters in this act experience shock and anger in response to his demise.



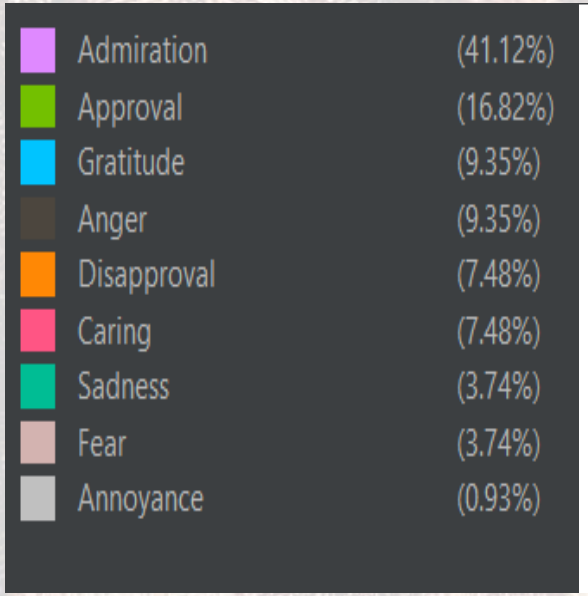
{'Love': 1, 'Grief': 2, 'Anger': 3, 'Caring': 4, 'Embarrassment': 5, 'Annoyance': 6, 'Disapproval': 7, 'Confusion': 8, 'Amusement': 9, 'Admiration': 10, 'Gratitude': 11, 'Optimism': 12, 'Sadness': 13, 'Approval': 14, 'Realization': 15, 'Disgust': 16, 'Fear': 17, 'Surprise': 18, 'Joy': 19, 'Remorse': 20, 'Desire': 21, 'Pride': 22, 'Disappointment': 23, 'Excitement': 24, 'Nervousness': 25, 'Relief': 26}

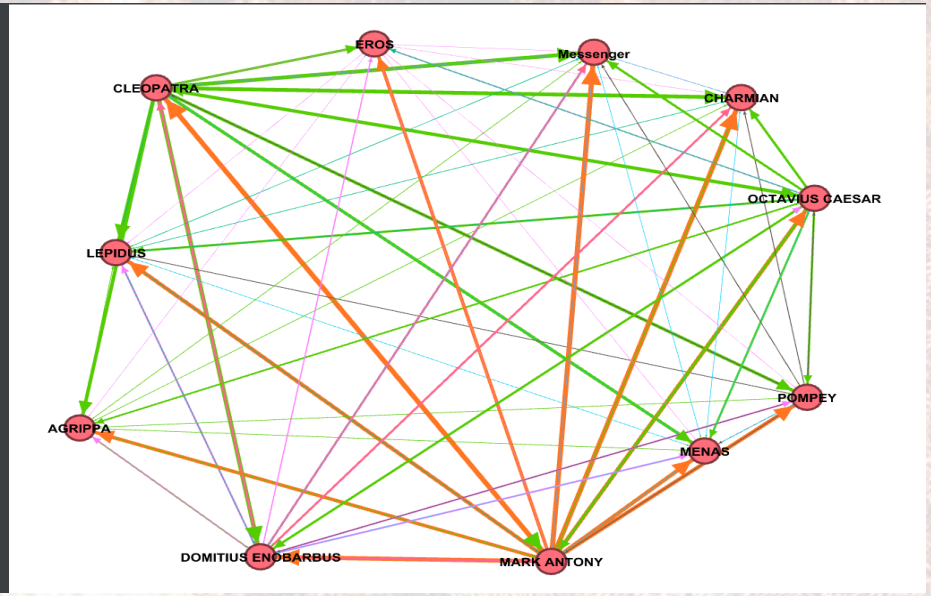
Analyzing Character Interactions and Emotions:

- To represent the emotion expressed by one character when interacting with another, I had to create a weighted directed graph using the Networkx library.



- The above graphs display the different dominant emotions each character expresses with another one and the level of interaction, for example we can see that Cleopatra and Antony share two edges where each one is dark green colored meaning, they both interact with one another a lot during the play, in this case they are together in a lot of scenes.





Conclusion & Final Thoughts:

- This project serves as a foundational exploration of emotion detection in textual data. It focuses on three key aspects:
  - Comparing the Text2Emotion and EmoRoBERTa models.
  - Utilizing the EmoRoBERTa model to depict the emotional profiles and evolution of characters throughout the play.
  - Representing character interactions through network graphs, visualizing the dominant emotions expressed during these interactions.
- Further improvements can be implemented to better represent the emotional profiles and the accuracy of interactions between characters.


THANK YOU.