A Retelling of Bong Joon Ho's 'Parasite'

Ethan Chew

Abstract—As streaming subscription services disrupt the market, comes with it are an abundant amounts of movies and shows, begging to be analysed. These shows and movies can be treated as data covering an array of data types and by leveraging on it, it allows for potential understanding of films with applications in film recommendation algorithms. In this study, we look to retell 'Parasite', a film by Bong Joon Ho, utilising a text-based approach and framework which relies on visualisation techniques and human reasoning to extract information on character significance, interactions, relationships, plot points and underlying themes. We will see how the use of occurrence plots allows us to get a good initial view of our characters while interaction heatmaps and network analysis further emphasises our understanding and provides more context through character relationships. We will also see how, by aligning to the events of the film, splitting the script into two halves provides more granular insights. WordClouds will explore plot points and themes through a scene's setting and body, albeit inconclusive, while sentiment analysis will identify changes in sentiment when moving to the second half of the film and capture the general shift in sentiment over both halves.

1 Problem Statement

With the influx of movie and television subscription services from the likes of 'Netflix', 'Disney+', and 'Amazon Prime Video' over the past few years, not only did it disrupt the telecommunications and entertainment industry drastically but it also gave rise to significant amounts of data spanning a wide range of data types.

This study looks to retell Bong Joon Ho's film, 'Parasite', through visualisation techniques on text-based data with domain knowledge (i.e. film understanding) and human reasoning at the cornerstone of decision making. We look to present the essence of the film through character significance, plot points and underlying themes, and answer the following research questions along the way:

1. Are we able to get a representation of character significance and relationships through interaction plots and network analysis?
2. Are we able to recognise plot points and underlying themes through the use of WordClouds?
3. Does sentiment analysis provide an effective way of identifying the major plot twist and if it is able to capture the change in dynamics after it

Addressing the research questions increases the prospects of garnering more context of a film through its temporal sequence and underlying themes via a text-based approach, allowing for an additional potential dimension in film recommendation.

The remaining sections are structured accordingly. Section 2 discusses relevant works, while Section 3 provides an overview of the data and preparation techniques. Analysis, including the approach, process and results, are all covered in Section 4. Following this, Section 5, highlights reflection points and findings.

2 State of the Art

Though semantic analysis via visualisation techniques on 'Parasite' has not populated the research space, the flexible application of text analysis and natural language processing allows us to draw inspiration from and make references to similar works performed on other datasets.

Kurzhals et al. presented a summarisation and exploration approach to the analysis of movie content through the use of iterative visualisation and query-based techniques [1]. By leveraging on multiple datasets (i.e. video-based data from movie scenes and text-based data from movie scripts and subtitles), they were able to extract features in multiple dimensions (e.g. motion, similarity and occurrence tags) for semantic analysis. In addition, they made reference to works performed by Liu et al. and Chen et al., and stated the importance of four aspects (i.e. 'who', 'what', 'where', and 'when') in obtaining the full context of a movie scene [2], [3]. Despite this, they suggest the use of audio-based data as an additional dimension to support a wider range of queries and analysis tasks, especially with the challenges that come from the privacy surrounding movie scripts and the mismatch of scenes between scripts and videos.

On the contrary, works by Weng et al. proposes a different approach to movie analysis - one that strays away from the conventional approach of frame-based audio-visual analysis [4]. Instead, they utilise social networks and stress on the importance of character interactions throughout a film in understanding the context of it and treats the film as a society with macro and micro-communities. Changes in the communities allow them to subsequently segment the film accordingly. To capture character interactions, they employ facial recognition in identifying character appearances but recommend the adoption of audio data in tandem to account for any inaccuracies in facial recognition due to lighting conditions or poses, for example.

Surrounding a similar domain, Lee et al. explored the patterns in the narrative flow of movies using a combination of natural language processing, machine learning and visualisation techniques [5]. By solely utilising unstructured text-based data in the form of movie scripts and performing sentiment analysis, they clustered the movies (into six clusters) based on its sentiment. To assist in the visualisation of the sentiment of the films, they employed and considered various

smoothing techniques like Locally Weighted Scatterplot Smoothing ("LOWESS") and Savitzky Golay Filter. Nonetheless, they encourage the use of more granular analysis in the future to further understand properties (i.e. starting and ending sentiment) of their clusters.

Regardless of the differences in types of dataset used between this study and other works, visualisation and text analysis techniques are often flexible and applicable to various domains and datasets. Besides, works from Kurzhals et al. also utilised text-based data and suggested extraction methods on movie scripts - of which can be applied to our study. Additionally, Weng et al. provides inspiration for character interaction analysis as part of understanding the film's context. Finally, works from Lee et al. offers valuable insights and comparisons on potential smoothing techniques for easier visualisation.

3 Properties of the Data

This study works solely with text-based data, namely the translated (from Korean) English script of 'Parasite' that was sourced from deadline.com, an online website for entertainment news. Given that the script is in a text-based pdf format, it will need to be worked into a parsable format.

However, a benefit of working with movie scripts is the inherit structure and standardised formatting rules that surrounds it. Here we employ the element extraction technique, as seen by Kurzhals et al. [1], with some alterations to the context of this study and script. The element extraction technique segments sections of the movie script into five main structural elements.

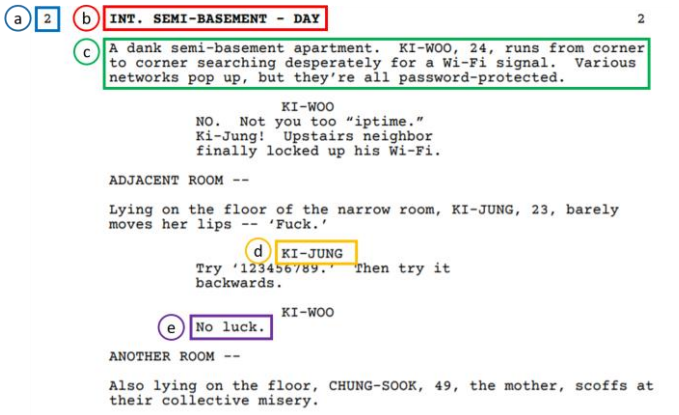

Fig. 1. Example of structural elements in a movie script

Fig. 1. above shows an example of structural elements that exist in sections of a movie script and these elements encapsulates the four aspects in capturing the context of a scene, mentioned by Chen et al., namely, the "who", "what", "where" and "when" [3]. To be precise:

a) Scene number represents the temporal flow of the film and highlights the "when" (i.e. when does a scene take place relative to the film?)

b) Scene setting describes the location and time of the scene and portrays the "where" (i.e. where is the scene taking place?)

c) Scene description provides context and narrative description and accentuates the "what" (i.e. what is unfolding in the scene?)

d) Character(s) within the scene which identifies the "who" (i.e. who is in the scene?)

e) A character's line or dialog within the scene which, with c), describes the "what" of a scene

From this, we manually segment the movie script into its structural elements accordingly. The resulting dataset is a corpus of 159 scenes with three columns: a column named "Scene Number" to store a), a second column named "Setting" to contain b), and a final third column named "Scene Body" to group c), d) and e). It should be noted here that c), d) and e) are grouped together to reduce the tediousness and intricacy of manually splitting those elements. Additionally, Kurzhals et al. noted on the complications that come with films without linear temporal order, like "Pulp Fiction" [1]. Luckily, 'Parasite' follows a linear temporal order with minimal flashbacks and thus, we maintain the temporal order as laid out in the script.

With text-based data, typical statistical visualisation techniques like boxplots or distribution plots may not be entirely applicable in assessing data quality. Furthermore, with movie scripts, each scene can be assumed to be integral to the overall narrative of the film and thus, outliers may not be applicable either. Nonetheless, for each scene, we compute the length of the setting and body of the scene to assess for any missing or erroneous values during the element extraction process. Plotting both on separate sets of axis allows us to ensure that there are no skips in scenes or scenes without a setting and/or a body.

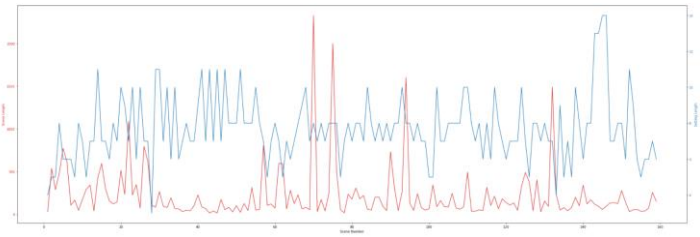

Fig. 2. Scene (red) and setting (blue) length over the film

4 Analysis

4.1 Approach

For the retelling of 'Parasite', we propose an analytical approach that puts visual analytical techniques and human reasoning at the forefront of decision making. An overview of the approach is shown in Fig. 3.

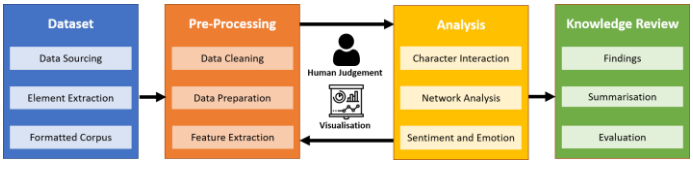

Fig. 3. Proposed analytical approach

The approach starts off with focus on the dataset where the data sourcing and element extraction technique was discussed

in the prior section, to output a corpus that has been formatted for further analysis and processing.

Specific to this approach, we propose an iterative process between data pre-processing and analysis. Ideally the data would be pre-processed with all the necessary requirements met prior to performing any analysis but in reality, vast amounts of iterations and intermediary results are analysed prior to the final outcome. It is vital here that, throughout the iterative process of intermediary results, we rely on human reasoning and leverage on visualisation techniques and outputs to assess the need of further analysis through other techniques or higher levels of granularity (e.g. demographic).

Nonetheless, with text-based data, tokenisation is essential for feature extraction, as will be seen in the subsequent section. That being said though, tokenisation can be computationally intensive, depending on the length and vocabulary of the text. Thus, to reduce the computational complexity, we will clean and prepare our corpus prior to any feature extraction. Here characters and words with little meaning are removed from our text and then transformed to a further simplified state through lemmatisation (i.e. the process of transforming a word to its root form, for example "running" will be transformed to "run"). The cleaned corpus then allows for a computationally easier feature extraction process for subsequent analysis on character occurrences, interactions, sentiment and plot points.

Here we will not only obtain sentiment and emotional scores for each scene, but we will also iterate over the movie script to identify when characters have been mentioned and who the characters are interacting with. We will treat an interaction between characters as consecutive dialog between characters or mentions of a character name from another character. To obtain an essence of context of the film, we will draw inspiration from works done by Weng et al. to adopt network analysis [4] and heatmaps to help better visualise the interactions between characters. Analysing certain emotions and overall sentiment throughout the script will allow us to capture the other dimension of context of the film by potentially identifying scenes with significantly high/low emotional/sentiment scores or even mapping iconic scenes to an emotional/sentiment score. In addition, we will explore the use of WordClouds in picking out potential plot points and underlying themes throughout the film.

Finally, we will look to summarise our findings throughout the study and critically evaluate it under the knowledge review section. It is important here that we, not only, identify any caveats to our study but also discuss areas for improvement for future work. Furthermore, reflecting back on our research questions and whether they have been answered will prove as a useful exercise.

4.2 Process

For some context, 'Parasite' tells a story of two families on complete opposite ends of the socioeconomic spectrum. While the Kim family is poverty stricken, the Park family is wealthy and shows disgust when being compared to average individuals. After landing an English tutor role with the Parks, Ki-Woo (son - Kim family) and family devise a plan of replacing each existing worker of the Parks and working under a pseudonym (i.e. fictitious name) and feeding off them (much like parasites) while hiding the knowledge of them being related. Midway through the film however, we are delivered a dramatic plot twist, with the reappearance of the previous caretaker, Mun-Kwang, and revealing the shelter of her husband, Kun-Sae, in a secret room at the basement of the Parks' home. What was once a comedic and light-hearted film now spirals into a suspenseful thriller with the eventual death of Mun-Kwang, Kun-Sae, Ki-Jung (daughter - Kim family) and Dong-Ik (father - Park family).

With this, we split our research questions into three tasks (one for each research question) and extract features tailored to each.

We first clean and prepare our data by removing punctuation, numbers and convert the text to lowercase. Subsequently, we remove stopwords to remove words with little significant meaning and add the words "int", "ext" and "int/ext" to our list of stopwords. These words are commonly used when describing the setting of a scene to outline whether it is shot in the interior, exterior of a place, or a mix of both. Additionally, we employ lemmatisation, using nltk's 'WordNetLemmatizer' to further reduce the computational complexity by simplifying our data.

Task 1: For each scene, we split the body of the scene into individual words (tokens) through tokenisation, using nltk's 'word_tokenize' method. Iterating over our tokens, we identify when and how often individual characters appear in each scene - either through lines of dialog or mentions by other characters. Plotting each character's appearance across the film, coloured by the frequency, will give us a good initial view of character significance.

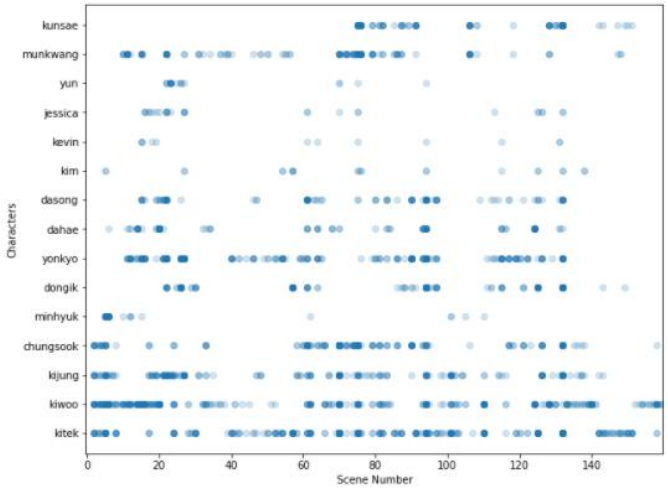

Fig. 4. Character appearance plot

With Fig. 4., we get an initial view on the main characters through the frequency and timing of their appearances. We see the Kim family: Ki-Woo, Ki-Jung, Ki-Tek and Chung-Sook, all appearing in the opening scene with consistent showing throughout the film. Conversely, a rough view of the Park family can be identified through the timing of their appearance in the film. Characters like Yon-Kyo and son, Da-Song, share similar appearance times, with daughter, Da-Hae, also

appearing at similar but less frequent times. We can also identify minor characters like Min-Hyuk (friend to Ki-Woo) or Yun (original driver of Dong-Ik) who appears early on and only reappears through passing remarks in the latter half. Throughout the latter half, we spot the sudden reappearance and importance of Mun-Kwang, and introduction of Kun-Sae; another character who plays a pivotal role. Towards the end, the film focusses on Ki-Woo and Ki-Tek with little mention of the remaining cast.

Though the appearance plot supplies us with valuable insights, one downfall is the inability to capture the interactions between characters. So, we determine interactions between characters as consecutive dialog between characters or mentions of them from other characters and plot an interaction heatmap between the characters coloured by the number of interactions.

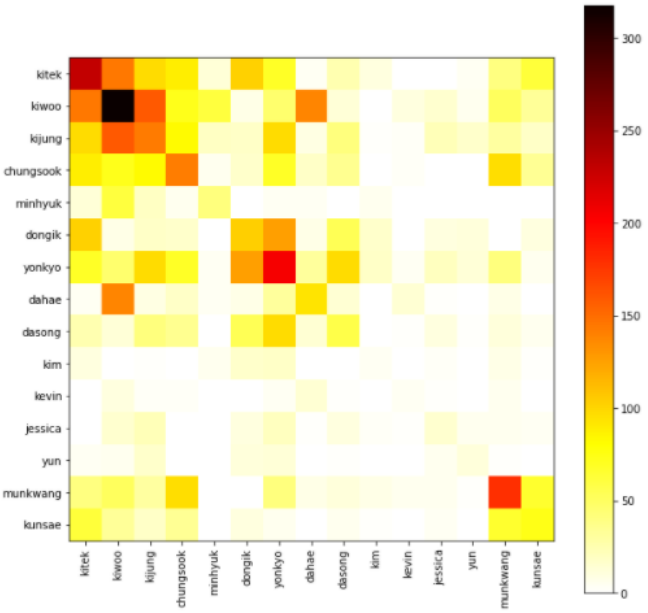

Fig. 5. Character interaction heatmap

Fig. 5. reinstates some of our learnings from Fig.4. The Kim family interacts in the top left corner of the grid frequently, cementing their relationship as a family. We start to see Ki-Woo's relationship with Da-Hae and Min-Hyuk, and Dong-Ik as a member of the Park family through his frequent interaction with his wife, Yon-Kyo. Interestingly, we get glimpses of the pseudonyms used by the Kim family. Through the interactions of Kim (Ki-Tek), Kevin (Ki-Woo) and Jessica (Ki-Jung), they mainly only interact with their real selves (i.e. themselves mentioning their own pseudonym), the Parks, and hardly ever amongst each other. The heatmap however, does not make the relationship between Mun-Kwang and Kun-Sae evidently clear.

With the heatmap plotted over the entire film, Mun-Kwang and Kun-Sae's relationship may not have had enough time to mature given that Kun-Sae only appears in the latter half. Thus, we will split the script into two halves: first and second half spanning scenes 1-70 and 71-159 respectively (i.e. before and after the plot twist). Drawing a network between the characters across both halves allow us to better visualise interactions between characters in the context of the events of each half.

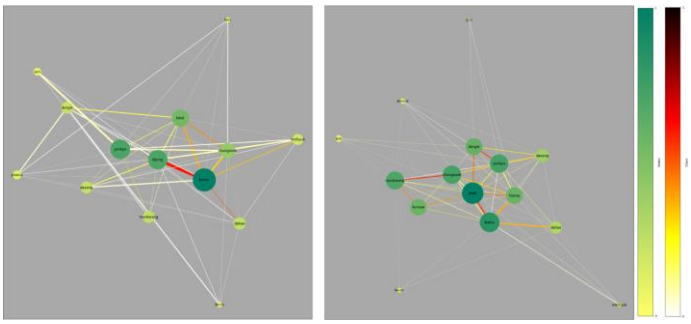

Fig. 6. Network of characters in the first (left) and second (right) halves of the film

Fig. 6. draws a character network over both halves, with the colour and size of the nodes and edges being proportional to the number of interactions. The network over the first half tells a similar story as earlier and makes the pseudonym relationship between the Kims and Parks more apparent. Kun-Sae does not appear in this network as he's not introduced until the second half. Going into the network of the second half, Kun-Sae and Mun-Kwang's relationship becomes more apparent with a decent amount of interaction between the two. Additionally, the darker edge between the two of them, accompanied with multiple lighter edges between the two characters and each member of the Kim family, potentially captures some of the rivalry between both sides.

Task 2: The setting and body of each scene are merged into its respective strings and parsed through a WordCloud to potentially identify plot points and themes. Similar to task 1, we draw and compare WordClouds over the entire script and both halves.

Fig. 7. Setting WordClouds over the entire script (left), first (middle) and second half (right)

Fig. 7. displays the most common bigrams in the setting of a scene over the entire script and both halves individually. Comparing the three, the majority of the film is largely set over a handful of settings: the semibasement (home of the Kims), mansion (home of the Parks) and secret basement (home of Kun-Sae). Looking at the first half, there is some emphasis on scenes set in the mansion and semibasement but no mention of the secret room or storage basement. Contrastingly, the second half introduces the secret room and storage basement where it's frequently visited.

Fig. 8. Scene body WordClouds over the entire script (left), first (middle) and second half (right)

The WordClouds of the body of the scene was recursively generated after removing additional words. Character names and common words between the three plots, like "look", "still", "one", were removed to not populate the WordClouds with redundant words.

Fig. 8. shows a relatively more eerie WordCloud when only considering the second half. Positive words like "good" no longer appears while words like "behind" and "dark" appear more frequently. Additionally, "stair" is frequently mentioned as a potential reference to the stairs that lead downwards to Kun-Sae's room. "Phone" is also a big plot point early on in the second half where Mun-Kwang threatens to expose the truth about the Kim family through photographic evidence. Though, it is unclear whether the frequent use of "phone" in the second half references that, as it is commonly used over the entire and first half of the film as well. Words like "line" and "smell" are common metaphors used throughout the film to portray the difference in socioeconomic status' with the tension reaching reach its peak in the latter half. These themes, though, are not picked up by any of the WordClouds.

Task 3: With the tokens from *Task 1*, we obtain sentiment and emotion scores of each scene using 'TextBlob' and 'NRCLex' respectively. We aim to analyse the course of some key emotions (i.e. fear, anger, and disgust) throughout the film while sentiment scores provide a general outlook of polarity (i.e. how positive/negative scenes are). 'NRCLex' spans eleven emotions but for visibility and applicability, we will focus on the key emotions; ones that were experienced by the characters after the plot twist.

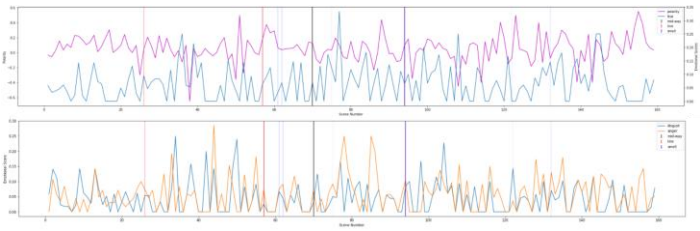

Fig. 9. Emotion and sentiment throughout the film

Fig. 9. plots the key emotions and sentiment throughout the film, with vertical markers indicating certain parts of the film. The black vertical line indicates where the film takes a twist, while red and blue vertical lines highlights the mention of a "line" between characters and the "smell" of characters (i.e. their place in the socioeconomic circle) respectively, coloured by the frequency of mentions.

Overall, the film starts relatively calm with fluctuations in emotions and sentiment midway through the first half when the Kims are executing their plan. Past the black line, general polarity decreases slightly with increases in fear, disgust and anger in the following scenes. The spike in fear and anger surrounds the reveal of Kun-Sae and the brawl that unfolds between the Kims, Mun-Kwang and Kun-Sae. Looking at the underlying themes though, sentiment and emotional scores fail to highlight the disgust of the Parks at the Kims' smell and the anger of Kims from the belittlement by the Parks in scene 94 (purple vertical line –frequent mentions of both "line" and "smell"). Here, overall sentiment is still positive while disgust and anger is minimal and same can be said for scene 57 (red vertical line). Nonetheless, sentiment and emotional scores may still prove useful to get a higher-level view on the overall sentiment. Specifically, fear and anger look to be higher on average in the second half as compared to the first half.

4.3 Results

From the frequency and timing of appearance, identifying a set of main and minor characters proved feasible whilst monitoring character significance throughout the film. What the occurrence plot lacked in character interactions was made up for in interaction heatmap and network analysis. A combination of both cemented our initial views and identified other relationships. Splitting the script proved to be useful, as networks were able to showcase relationships on a more contextual basis.

Some plot points were identified through analysing the setting of the scene. Changes in sentiment was also potentially present though it is unclear.

Through sentiment analysis, we saw changes in some key emotions after the plot twist but without prior knowledge to the film, it may not be evident without systemic changes in sentiment or emotion. It failed to capture the underlying theme and sentiment at certain scenes but captured the general sentiment instead.

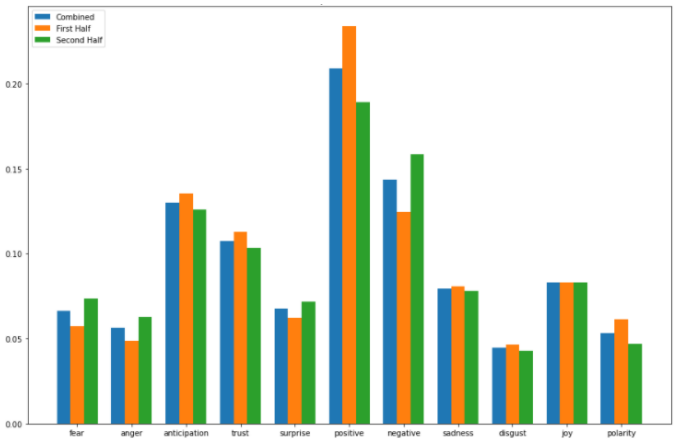

Fig. 10. Comparison of means of emotions and sentiment

Fig.10. shows an increase in amount of fear, anger and negativity going from the first to second half. Levels of disgust are similar throughout but this could be due to disgust from other sources. For a film which is praised for its imagery, supplementing the analysis with video data may be beneficial in capturing the theme and plot points.

5 Critical reflection

This study presented a retelling of 'Parasite' through the combination of feature extraction and visual techniques like interaction heatmaps, networks and WordClouds, with human judgement resulting in more granular analysis through the split of the movie script. Working with text-based data, text and visual analysis methods could be drawn from other studies without the limitation of using a similar dataset.

Though the underlying themes were not apparent, a combination of occurrence plots, interaction heatmap and network analysis allowed us to identify the main and minor characters, their significance throughout the film and their relationships. In further support, clustering of the characters through k-means clustering, may be useful in drawing out groups or cliques. Iterating over various number of clusters, elbow or silhouette plots can assist the analyst in applying human reasoning to determine the optimal number of clusters.

Additionally, despite the use of pseudonyms by the Kim family, the script is formatted such that the line is assigned to that character and not their pseudonym. For example, if Ki-Woo is presenting himself as Kevin, the dialog is still allocated to Ki-Woo. With more time, formatting the text such that the pseudonym of a character is allocated the line instead will be useful in accentuating the type of relationship the Kims share with the Parks.

WordClouds did not perform spectacularly but had some utility in identifying some of the plot points and sentiment, albeit inconclusive. In the future, topic modelling with nltk's or gensim's Latent Drirchlet Allocation (LDA) could be explored to assess its performance in picking out underlying themes and plot points.

Sentiment analysis managed to capture the shift in dynamics after the plot twist but was not entirely reflective on the sentiment of certain scenes. This could be due to the sheer length of scenes, where a sentiment is outweighed by more frequent associations of other sentiments. Tokenising the scene into sentences and averaging the sentiment scores over the sentences may prove a better representation. Also, 'TextBlob' here was used for getting sentiment scores but nltk's 'VADER' was a consideration. Ultimately the sentiment scores from 'TextBlob' were relatively more stable and interpretable.

Working solely with the movie script, we may not get the entire story. With each scene, the dialog plays a part in selling the narrative but so do the imagery, camera work, lighting, music and presence of characters in the scene. Especially so with 'Parasite'; a film that is praised for its imagery. Moving forward, supplementing the movie script with image, video and audio-based data, in the form of storyboards, the movie itself and its audio, may prove to be beneficial in capturing the entire essence of the film. From it, facial expressions, tempo of the scene, changes in camera work and music, can all play a role in capturing a dimension of the film and scene. Furthermore, the movie script is translated from Korean to English. Thus, adding other types of data may help to account for text that potentially got lost in translation.

Table of word counts

Problem statement	250
State of the art	496
Properties of the data	481
Analysis: Approach	494
Analysis: Process	1507
Analysis: Results	206
Critical reflection	498

References

[1] K. Kurzhals, M. John, F. Heimerl, P. Kuznecov and D. Weiskopf. Visual Movie Analytics, *IEEE Transactions on Multimedia*, 18(11): 2149-2160, 2016.

[2] A. Liu, S. Tang, Y. Zhang, Y. Song, J. Li, and Z. Yang. A hierarchical framework for movie content analysis: Let computers watch films like humans, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* pp. 1-8, 2008.

[3] B. -W. Chen, J. -C. Wang and J. -F. Wang. A Novel Video Summarization Based on Mining the Story-Structure and Semantic Relations Among Concept Entities, *IEEE Transactions on Multimedia*, 11(2): 295-312, 2009.

[4] C. Weng, W. Chu and J. Wu. RoleNet: Movie Analysis from the Perspective of Social Networks, *IEEE Transactions on Multimedia*, 11(2): 256-271, 2009.

[5] S. Lee, H. Yu and Y. Cheong. Analyzing Movie Scripts as Unstructured Text, *2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService)*, pp. 249-254, 2017.