

CHAPTER 1 : INTRODUCTION

The Marvel Cinematic Universe (MCU) franchise has become a global sensation. A series of movies created by Marvel Studios has touched the hearts of millions of people worldwide. Die-hard fans are avidly waiting for the next movie to expand the interconnected fictional world of the MCU. One of the latest anticipated MCU movies released at the end of 2021 is Spider-Man: No Way Home. This movie sets to end the MCU's Spider-Man movie trilogy. It has garnered high praise from the fans who watched it. Many people have reviewed it to be higher than other MCU movies, as shown in Figure 1.

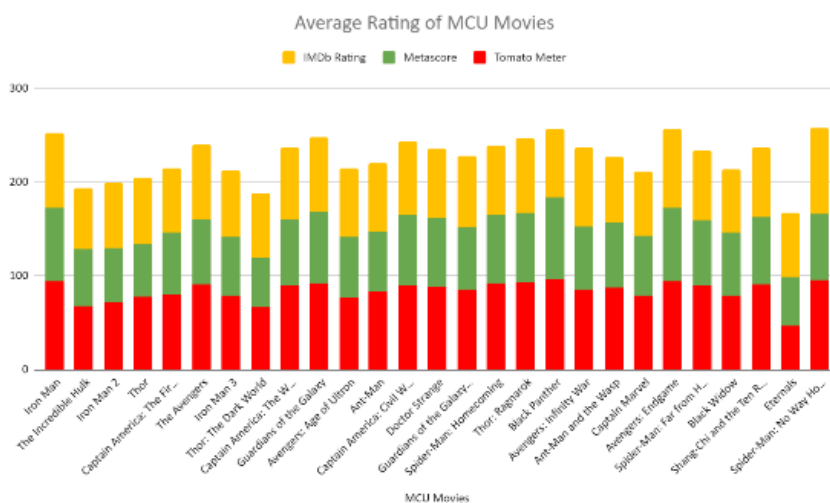


Figure 1. Average Movie Ratings of the current MCU films

When people review a movie, it isn't always classified as a perfect 10 or a solid 1. It mostly lands between the two numbers as they describe what they like or dislike about said movie. It is then read by other people and affects their judgment on watching the movie, based on the positive and negative sentiments. This is where the study of sentiment analysis comes in. Sentiment analysis is the process of determining a text-based dataset to be positive or negative [1]– [4]. By conducting sentiment analysis on a movie review, people will have an easier time understanding the review's overall sentiment and movie studios will receive the sentiments as feedback to improve their future movies and shows.

There are multiple classification methods in sentiment analysis. One of the more popular methods is using the Naïve-Bayes classifier. Researchers have used this method to analyze movie reviews [5]. Others have improved their methods by adding complex pre-processing methods to increase accuracy [1], [2]. This method is also applicable in product reviews [6]. The other popular method is the Support Vector Machine (SVM) classifier. This method has been proved compatible with product reviews [7], [8]. One research uses frequency-inverse with SVM classifier on movie reviews [3] and another compares it with Naïve-Bayes and random forest method [9]. Other methods used in sentiment analysis are using Decision Tree and Stochastic Gradient Descent classifier. Applications of

the decision tree have been used to analyze anti-LGBT tweets on Twitter [10], e-commerce reviews on Google Play Store [11], and Twitter [12]. Applications of stochastic gradient descent are on covid-19 tweets [13], e-commerce tweets [14], and global terrorist attack incident reports [15].

The purpose of this research is to perform a sentiment analysis on the reviews of Spider-Man: No Way Home with the discussed methods above. Later, this work will present a comparison between the four different approaches to classification methods. This comparison will show the ability of several methods (in probability and matrix-vector based) in the sentiment classification.