

Midocean University Student Name: **Alanoud Alnjrany**

Student ID: **261001412**

Course Name: **Research Methodology (AI800)**

Date: **March 12, 2026**

A Comparative Study of Naive Bayes and SVM for Sentiment Analysis in Movie Reviews

Abstract

In this study, I will evaluate how machine learning models perform when analyzing sentiment in movie reviews. Since digital evaluation is growing fast on platforms like IMDb, it is important to find automated ways to classify these reviews accurately. My work focuses on comparing two main algorithms: Naive Bayes and Support Vector Machine (SVM). So, I processed the data using Python, which included steps like Tokenization, Stop-words removal; and Stemming. I used metrics like Accuracy, Precision, and Recall to see which model works better. I expect that SVM might be more accurate with complex sentences, while Naive Bayes will be faster in terms of computation.

1. Interodaction

Sentiment Analysis is a key field in Natural Language Processing (NLP) used to understand and classify opinions in text. Today, industries rely on it to track customer feedback especially in the entertainment business. Since thousands of movies are reviewed every year, manual analysis is impossible. So machine Learning (ML) offers a great solution by automatically labeling reviews as positive or negative. In this assignment, I focus on two popular algorithms: Naive Bayes and Support Vector Machines (SVM) to predict sentiments with high accuracy. While there is a lot of progress in NLP, choosing the best algorithm is still tricky because some are faster while others are more precise. My goal is to compare these models to see which one works best for movie datasets.

2. Research Problem

In the field of text mining, choosing the right algorithm for sentiment analysis remains a significant challenge. My review of recent literature shows a clear division: some studies rely on heavy, complex models like BERT that require huge computing power (He et al., 2024), while others use basic methods that often fail to catch the nuances of movie reviews. This creates a gap for a solution that is both accurate and computationally efficient. My research addresses this problem by evaluating Naive Bayes and SVM on the IMDb dataset. The goal is to see if these more "lightweight" algorithms can still provide high-quality results without the need for expensive hardware or high memory usage.

3. Research Objectives

The main goal of my study is to test and compare how different machine learning algorithms work on movie reviews. The specific goals are:

- Set up and train both Naive Bayes and SVM models using the IMDb dataset.
- Test the accuracy of each model in correctly identifying positive and negative reviews.
- Compare the results using standard metrics like Precision, Recall, and F1-score.
- Measure the time each model takes to process the data to check for computational efficiency.

4. Research Questions

To guide this study, I have developed the following research questions :

- RQ1: Which algorithm Naive Bayes or SVM, achieves higher accuracy in classifying sentiments in movie reviews ?
- RQ2: How do these two models compare when using metrics like Precision and F1-score ?
- RQ3: Is there a significant difference in the time it takes for each model to process the IMDb dataset ?

5. Literature Review & Critical Analysis

While reviewing current research on movie sentiment analysis, I found several important trends and some gaps that my study can address. For example, Mulyawan et al. (2025) focused on how to convert IMDb text into numbers, but they mostly used traditional methods, which leaves room to explore more advanced approaches. On the other hand, He et al.

(2024) showed that BERT models are very powerful for understanding emotions, but they also mentioned that these models require a lot of computing power. My study tries to find a balance by testing Naive Bayes and SVM, which are much faster and work better on standard computers.

Furthermore, although Sutriawan et al. (2023) and Sharma et al. (2023) provided good benchmarks for movie algorithms, I noticed they didn't focus much on linguistic challenges like sarcasm. Similarly, the study by Ajhari and Azzam (2023) was limited to only one movie franchise, which means we still need to test these models on larger datasets like IMDb to see if they work generally. Finally, while Pavitha et al. (2022) explained why sentiment analysis is important

they work generally. Finally, while Pavitha et al. (2022)for recommendations, they didn't go deep into the technical performance of the models themselves. This is exactly what I plan to focus-on in my research for broader testing on larger datasets like IMDb. Finally, although Pavitha et al. (2022) highlighted the importance of sentiment analysis for recommendations, their work did not deeply explore the technical performance of the classification models themselves.

Literature Review Synthesis Matrix

study	Author & Year	Research Design	Methodology	Key Finding	Limitation
-------	---------------	-----------------	-------------	-------------	------------

1	.Mulyawan et al (2025)	/ Experimental Comparative	Text Vectorization comparison on IMDb	Identified best vectorization for .accuracy	Limited to traditional methods; didn't test deep learning.
2	(He et al. (2024	Quantitative / Model Building	BERT-based models for short comments	Advanced models handle nuances .better	High computational cost and resource .requirements
3	Sutriawan et al (2023)	Quantitative Evaluation	Algorithm performance benchmark	Established accuracy rankings for ML	Focused on genera text; lacks focus on specific genres
4	Ajhari & (2030) Azzam	/ Case Study Comparative	Cross-platform review analysis	Source of data affects model precision.	Small sample size limited to one film franchise.

5	.Sharma et al (2023)	/ Quantitative Implementation	ML models for film feedback	Confirmed Python reliability for NLP.	Lacked comparison of execution speed and memory
6	.Pavitha et al (2022)	System Design / Integration	Sentiment analysis in recommendations	Improved Recommendation .quality	Sentiment analysis was not the primary technical focus

6. Methodology

For this study, I have chosen a **Quantitative Research Design** to effectively compare how different algorithms perform. My practical workflow is divided into the following steps :

6.1 Data Source

I will be using the IMDb Movie Reviews Dataset. I chose this specific dataset because it's a well-known benchmark in sentiment analysis, containing thousands of reviews already labeled as either Positive or Negative, which is perfect for training my models.

6.2 Data Cleaning & Pre-processing

To get the text ready, I'll use Python to perform several pre-processing steps. This includes Tokenization to split the sentences, and Stop-words removal to filter out common words like "the" or "is" that don't really help in identifying sentiment. I also plan to use Stemming or Lemmatization to bring words back to their base form, ensuring the models focus on the core meaning.

6.3 Model Implementation

I am going to implement two main algorithms using the Scikit-learn library:

- Naive Bayes (specifically MultinomialNB) : I selected this because it's famous for being fast and efficient with text data.
- Support Vector Machine (SVM): I chose this to see if its ability to handle high-dimensional data leads to better accuracy than Naive Bayes.

6.4 Measuring Results Finally

I will evaluate the results using a Confusion Matrix. My main focus will be on Accuracy (to see the overall correct guesses) and the F1-Score, which is important for checking the balance between precision and recall in my classifications.

Connection

Looking at the relationships between these studies, a clear consensus emerges. For example, both Sharma et al.(2023) and Sutriawan et al. (2023) strongly support the idea that Machine Learning is the most effective approach

for analyzing large-scale movie feedback. However, there is a shared focus between He et al. (2024) and Pavitha et al. (2022) on the practical impact of these models.

Conflicts between the findings

During my analysis, I identified a disagreement regarding model complexity. For example, He et al. (2024) argue that advanced models like BERT are necessary to achieve high accuracy in sentiment analysis. On the other hand, Mulyawan et al. (2025) found that traditional text vectorization methods can still deliver strong results while maintaining much faster processing speeds. This suggests that there isn't one perfect model for everything, and the 'best' model often depends on whether the priority is maximum precision or computational efficiency. My research will explore this further by testing if SVM and Naive Bayes can bridge this gap.

Synthesis & Research Gap:

From these papers, I realized that we need a model that is both accurate and fast. Most studies use very heavy models or very simple ones. My research will fill this gap by testing Naive Bayes and SVM on modern movie reviews to see if they can still give good results without needing too much computer memory.

7. Signification of the study

This research lies in its potential to improve how businesses and developers understand audience feedback. In the modern film industry, digital reviews are the primary source of information regarding a movie's success or failure.

The significance of this study can be summarized in the following points:

1. 1. Technical Contribution: By comparing Naive Bayes and SVM, this study helps identify the most efficient algorithm for text classification in movie datasets, which reduces the time and resources needed for data processing.
2. 2. Industrial Benefit: Film production companies and streaming platforms can use the findings to build better recommendation systems and understand viewer preferences more accurately.
3. 3. Academic Value: This research adds to the existing literature in Arabic and international NLP studies by providing a clear benchmark for sentiment analysis performance.

7. Scope and Limitation

To keep the research focused and effective, I have defined specific boundaries for this study:

- ***Dataset Scope:** The research is limited to English movie reviews from the IMDb dataset only.
- **Algorithms:** This study focuses only on comparing Naive Bayes and SVM. Other advanced models, like Neural Networks or Deep Learning, are not included in this particular assignment.
- **Sentiment Categories:** I am only performing binary classification (Positive vs. Negative). Neutral reviews or multi-level star ratings are outside the scope.
- **Technical Environment:** All tests and implementations are conducted using Python within a Jupyter Notebook, which is suitable for this type of academic comparison.

Platform: The implementation will be conducted using the Python programming language within a Jupyter Notebook environment

8. References

- **Ajhari, A., & Azzam, A. (2023).** The comparison of sentiment analysis of Moon Knight movie reviews between different platforms. Applied Information Technology Journal.
- **He, A., & Abisado, M. (2024).** Text sentiment analysis of Douban film short comments based on BERT-C. IEEE Access..
- **Mulyawan, R., Naparin, M., & et al. (2025).** Comparison of text vectorization methods for IMDB movie review sentiment analysis. Journal of Applied Informatics.
- **Pavitha, N., & Pungliya, V. (2022).** Movie recommendation and sentiment analysis using machine learning. GlobalTransitions Proceedings.
- **Sutriawan, A., et al. (2023).** Performance evaluation of classification algorithm for movie review sentiment analysis. International Journal of Advanced Computer Science and Applications.
- **Sharma, H., & Prasanna, S. (2023).** Sentimental analysis of movie reviews using machine learning. Journal of Computer Science and Technology