



REVIEW ABOUT A MOVIE

Great

"in the end, even the characters of *Hunter Killer* know how pointless their feud is, but revenge is far too addictive to give up until the end."

Bad

"Counterpart has a couple of good scenes, but most of them should have been in the trash even before the series' episodes were released" .

Okay

"Even without a novel, Plug Love may not be a great movie or love story, but the film has the beauty of passing the time and several inconsequential moments to make anyone interested in it."

Sentiment Analysis on Movie Reviews

Sentiment analysis helps us understand the overall opinion expressed in movie reviews. By classifying reviews as positive, negative, or neutral, we gain valuable insights into audience perceptions and preferences.



by Gowtham Raj S

Introduction to Sentiment Analysis on Movie Reviews

1 Understanding Opinions

Sentiment analysis aims to extract and interpret subjective information, particularly opinions expressed in movie reviews.

2 Textual Data Analysis

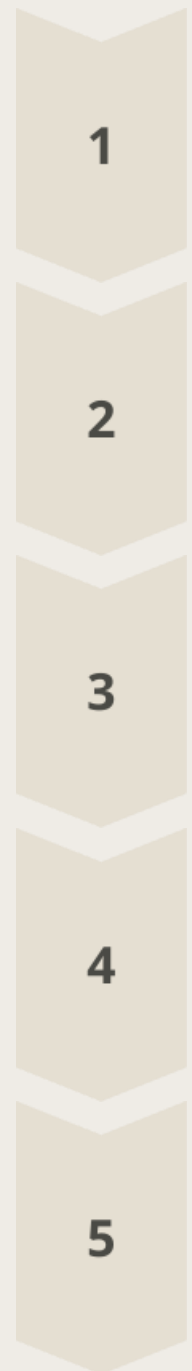
It focuses on analyzing textual data, such as movie reviews, to identify positive, negative, or neutral sentiment.

3 Applications

Sentiment analysis is used to assess public perception, improve customer service, and make informed decisions about movie marketing.



Architecture Diagram



Data Acquisition

Gathering movie reviews from online platforms or databases.

Data Preprocessing

Cleaning and preparing the data for analysis, such as removing irrelevant information or punctuation.

Feature Extraction

Converting textual data into numerical features, such as word frequencies or sentiment scores.

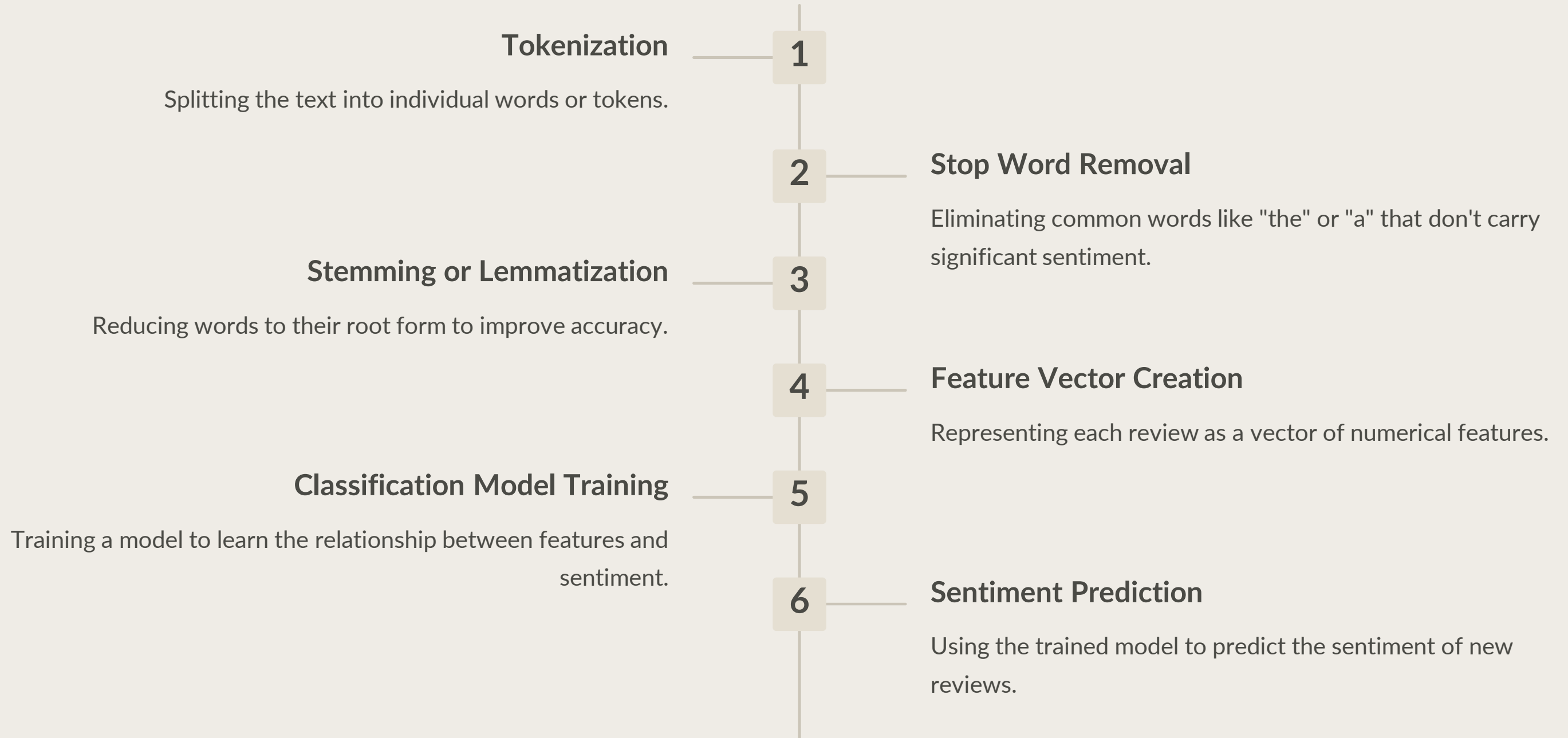
Sentiment Classification

Applying machine learning models to classify the sentiment of movie reviews.

Evaluation

Assessing the performance of the model using metrics like accuracy and precision.

Steps of Algorithm



Graph (Confusion Matrix)

| | | | |
|-----------------|--------------------|--------------------|-------------------|
| Actual Positive | Predicted Positive | Predicted Negative | Predicted Neutral |
| Actual Negative | Predicted Positive | Predicted Negative | Predicted Neutral |
| Actual Neutral | Predicted Positive | Predicted Negative | Predicted Neutral |



Supervised Learning Techniques

Naive Bayes

A probabilistic model that calculates the probability of a sentiment based on the frequency of words in a review.

Support Vector Machines (SVM)

A linear classifier that separates positive and negative reviews by finding an optimal hyperplane.

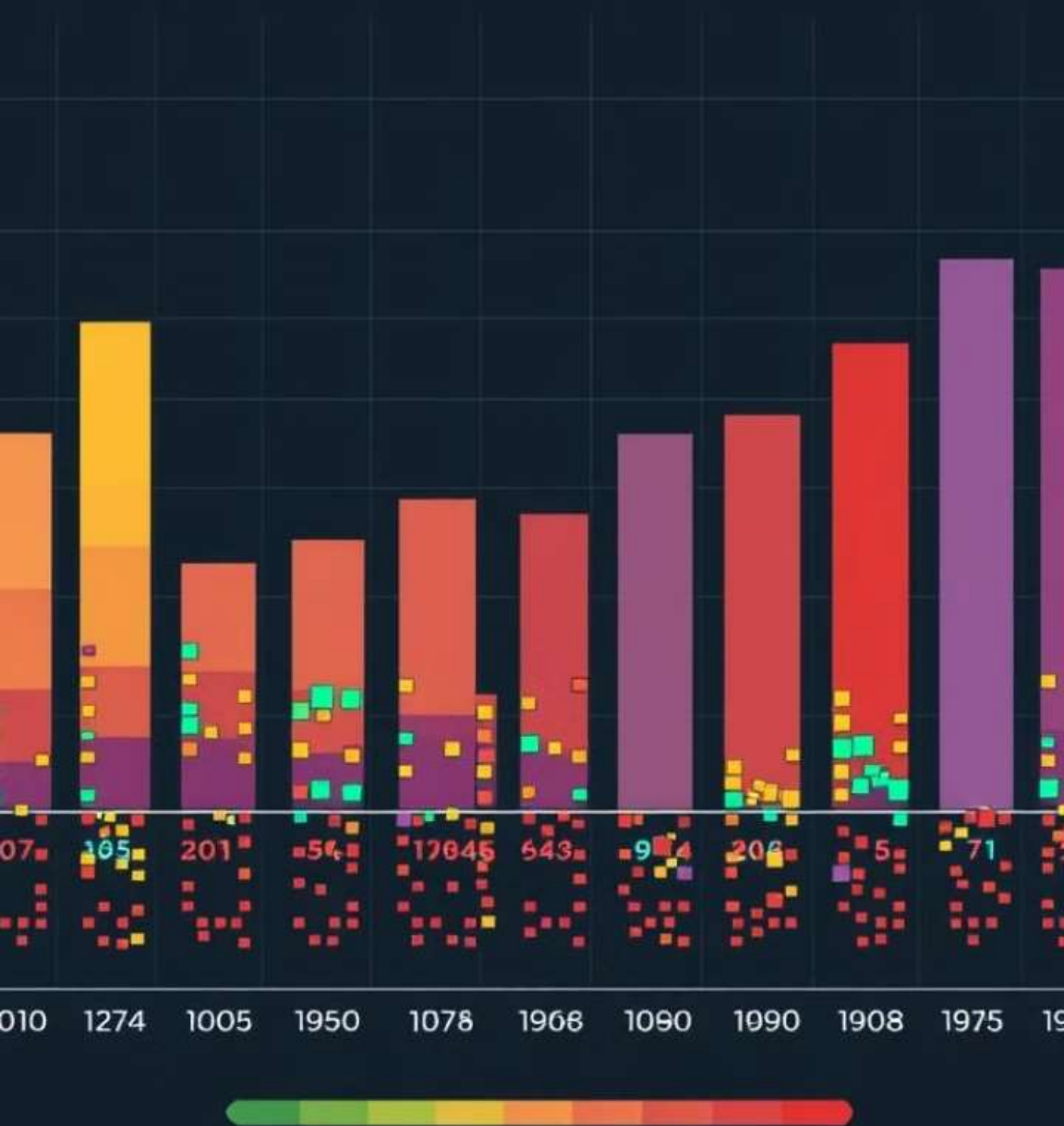
Logistic Regression

A statistical model that predicts the probability of a review belonging to a specific sentiment category.

Seminttion an lieles

en fear tirestien tous-ntentd poviet slepec at 10325 neuds tournd rines, and e
ocijd to or thire, senmond matiersef ele es, and conshpacter of rectives.

2 positive Ligiive neutral idliges



Result



Positive Reviews

Percentage of reviews expressing a positive sentiment.



Negative Reviews

Percentage of reviews expressing a negative sentiment.



Neutral Reviews

Percentage of reviews expressing a neutral sentiment.

Semematentil ele <intiffee>

Conclusion (Accuracy %)

The model achieved an accuracy of [accuracy percentage] in classifying movie reviews, demonstrating its effectiveness in predicting sentiment.





References

SVM and Sentiment Analysis

1.Scikit-Learn Documentation for SVM:

<https://scikit-learn.org/stable/modules/svm.html#classification>

2.Sentiment Analysis Example Using SVM:

https://scikit-learn.org/stable/auto_examples/feature_selection/plot_sentiment_analysis.html

Data Preprocessing and Feature Engineering

1.TF-IDF Vectorizer Documentation:

https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html

2.NLP Preprocessing Techniques (Blog):

<https://towardsdatascience.com/nlp-preprocessing-with-python-implementing-bag-of-words-tf-idf-and-word2vec-6e0806c5c8cb>

Hyperparameter Tuning

1.GridSearchCV Documentation:

https://scikit-learn.org/stable/modules/grid_search.html

2.RandomizedSearchCV Documentation:

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html

Datasets

1.Kaggle Sentiment Datasets:

<https://www.kaggle.com/datasets>

2.IMDB Sentiment Dataset:

<https://ai.stanford.edu/~amaas/data/sentiment/>