

Semantic Analysis on movie review data using Machine Learning Approach – Final Eval



Aakash Singh - 2021201087

Lalit Gupta - 2021201018

Sourabh Patidar - 2021201089

Problem Statement

01

Perform Semantic Analysis on movie review Dataset by IMDB using various machine learning and deep learning models and report the best model based on evaluation matrix on test data.

Methodology

- We had a dataset of 2000 movie review on IMDB.
- We had pre-processed the data and converted the text into multiple embeddings (i.e. Tf-Idf, Bert and Glove)
- We have divided the dataset into train and test set with test size = 0.3 as mentioned in paper.
- We have trained the classifier to predict positive or negative review on various machine learning model and deep learning model with multiple embeddings.
- We have evaluated the model based on accuracies, precision , recall , f1 -score and confusion matrix.

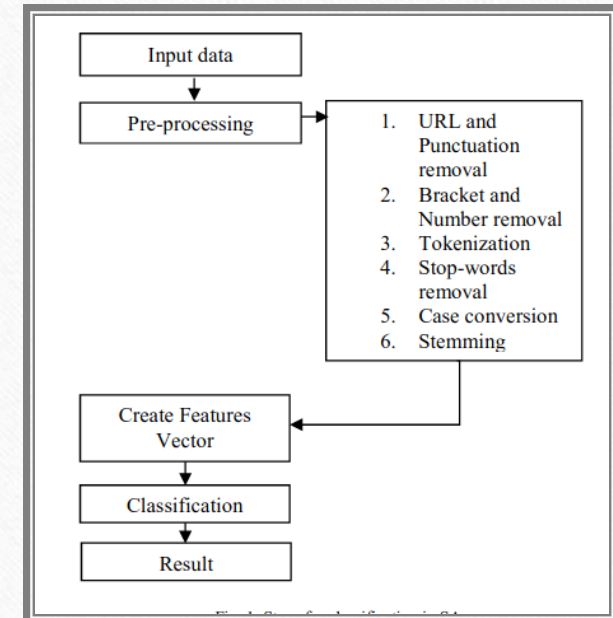
Understanding the Dataset

- We had dataset which contain 1999 movie reviews in Text column and corresponding sentiment "POS" or "NEG" in Sentiment Column of the dataset.
- We had perfectly balanced dataset with 1000 negative reviews and 999 positive reviews.

	URL	Text	Sentiment
0	http://www.imdb.com/title/tt0210075/usercomments	Girlfight follows a project dwelling New York ...	POS
1	http://www.imdb.com/title/tt0337640/usercomments	Hollywood North is an euphemism from the movie...	POS
2	http://www.imdb.com/title/tt0303549/usercomments	That '70s Show is definitely the funniest show...	POS
3	http://www.imdb.com/title/tt0716825/usercomments	9/10- 30 minutes of pure holiday terror. Okay,...	POS
4	http://www.imdb.com/title/tt0182225/usercomments	A series of random, seemingly insignificant th...	POS

Data preprocessing

- We have text data which we have to preprocessing it so that it can be used for creating embeddings.
- We have removed urls ,punctuation, brackets, numbers etc. for the text.
- We have tokenize the code and removed the stopwords.
- Convert the text into lowercase and applied stemming on the text.
- Also, the sentiment column has been converted to binary encoding(1 for positive review and 0 for negative review).



Data Splitting

- Convert the data into corresponding embeddings.
- Divided the dataset into train and test sets with test size = 0.3
- We had 1400 reviews to train the models and 600 reviews to test the models.
- We have created Tf-Idf, BERT and Glove embeddings of the given review dataset.

Fitting the models

- The research paper has implemented 5 different machine learning algorithms (MultinomialNB, BernoulliNB, SVM, Maximum Entropy and Decision Tree Classifier) and made comparison between them using various evaluation methods
- We have implemented all 5 models and in addition to that we have also implemented inbuilt MLP Classifier, MLP model and LSTM.
- We have used Tf-Idf embedding for the initial 5 models, BERT for inbuilt MLP Classifier and MLP model and Glove embeddings for LSTM model.

Evaluating the models

- We have evaluated model based on 5 parameters :
 - Accuracies
 - Precision
 - Recall
 - F1-score
 - Confusion Matrix
- All this values are calculated on test dataset which is 30% of whole dataset (600 reviews).

Results

Method	U	V	X	Y
Multinomial NB	250	19	281	50
Bernoulli NB	259	34	266	41
SVM	268	44	256	32
Maximum Entropy	254	190	110	46
Decision Tree	247	66	234	53

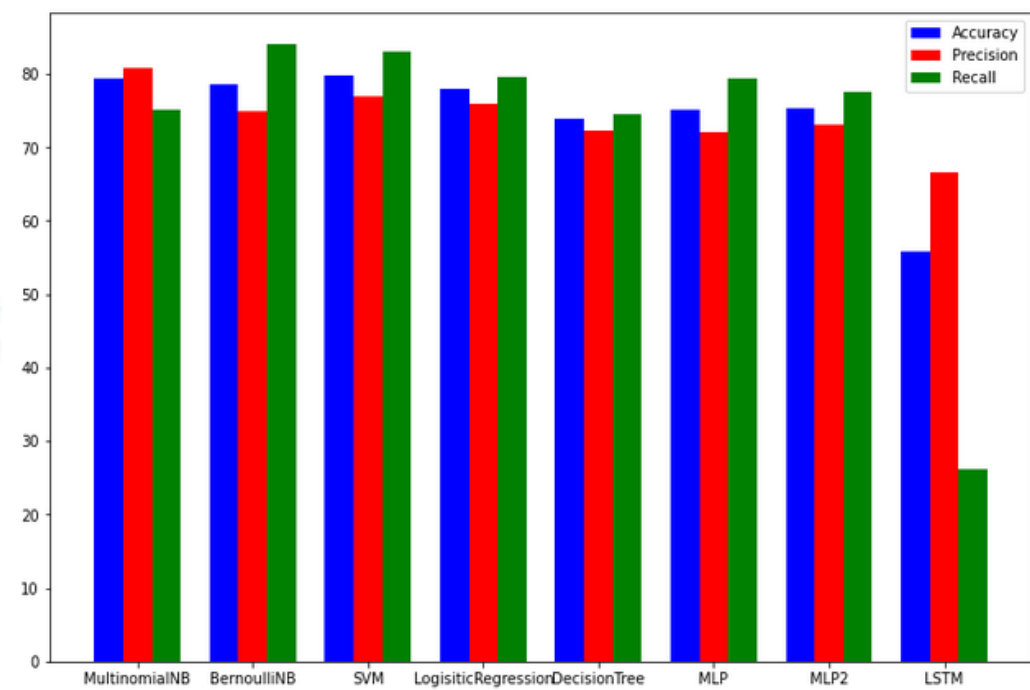
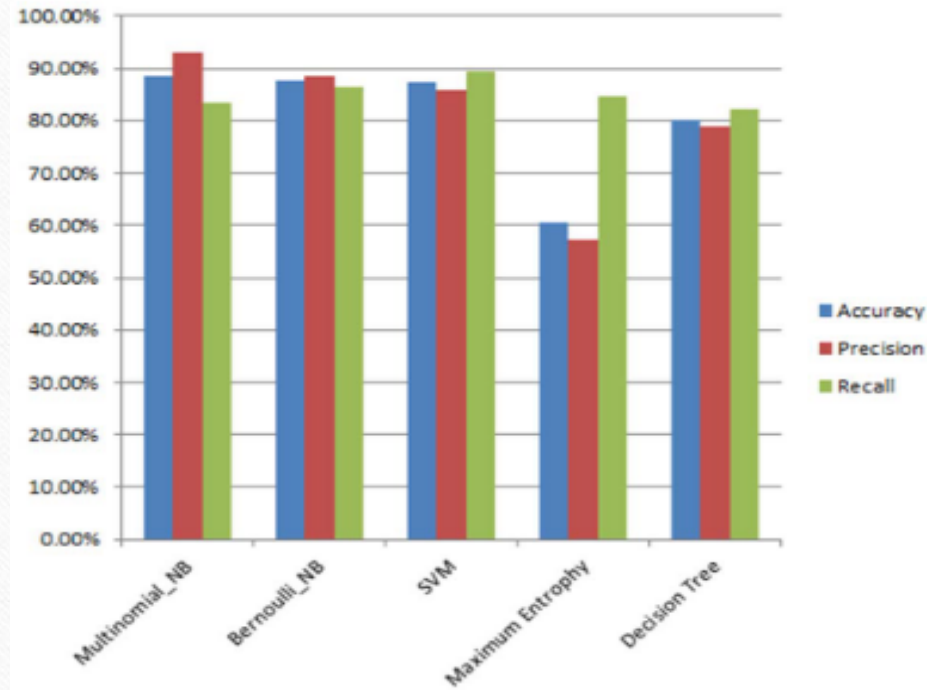
Method	U	V	X	Y
MultinomialNB	258	52	218	72
BernoulliNB	228	82	244	46
SVM	238	72	241	49
LogisticRegression	237	73	231	59
DecisionTree	227	83	216	74
MLP	221	89	230	60
MLP2	227	83	225	65
LSTM	255	40	80	225

Results

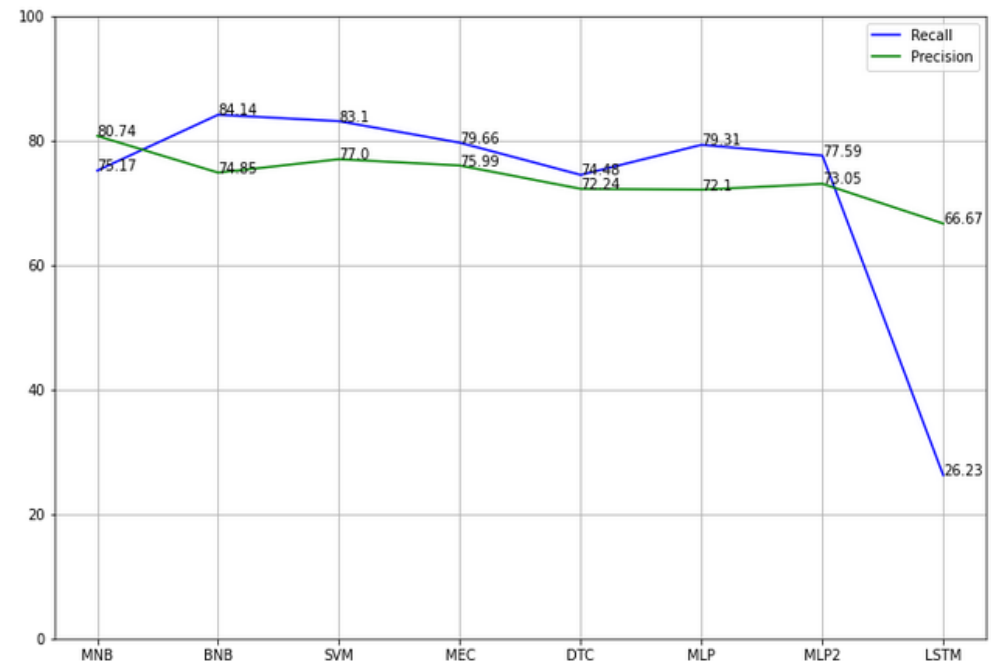
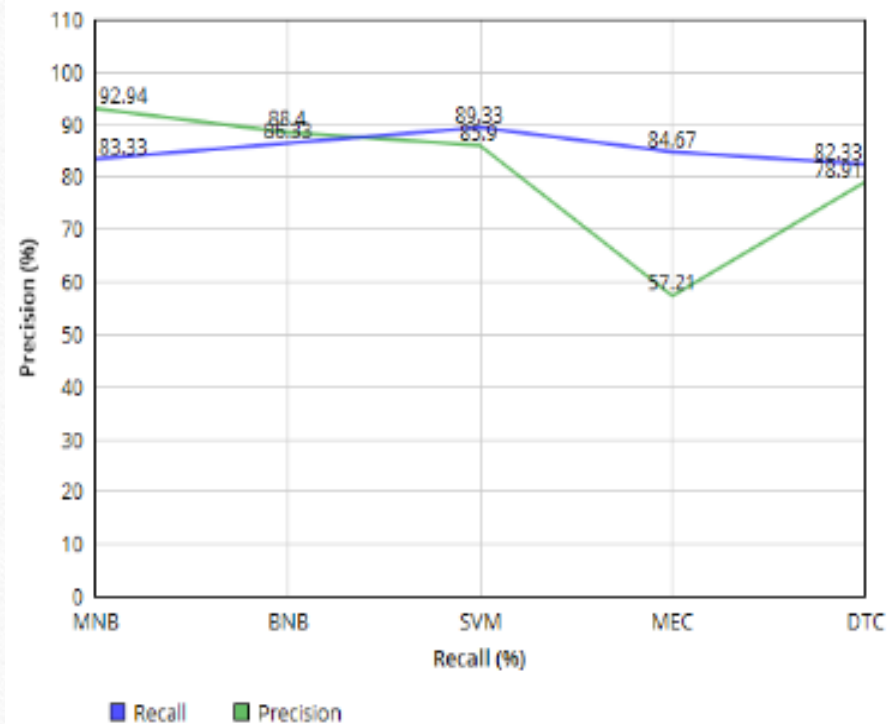
Method	Accuracy	Precision	Recall	F-score
Multinomial NB	88.50%	92.94%	83.33%	87.87%
Bernoulli NB	87.50%	88.40%	86.33%	87.35%
SVM	87.33%	85.90%	89.33%	87.58%
Maximum Entropy	60.67%	57.21%	84.67%	68.28%
Decision Tree	80.17%	78.91%	82.33%	80.58%

Method	Accuracy	Precision	Recall	F1-score
MultinomialNB	78.83 %	74.29 %	83.57 %	78.66 %
BernoulliNB	81.33 %	76.09 %	87.5 %	81.4 %
SVM	78.5 %	72.95 %	85.71 %	78.82 %
LogisticRegression	77.5 %	72.31 %	83.93 %	77.69 %
DecisionTree	71.33 %	66.77 %	76.79 %	71.43 %
MLP	76.67 %	74.68 %	78.77 %	76.67 %
MLP2	73.33 %	69.88 %	79.45 %	74.36 %
LSTM	55.83 %	59.35 %	41.64 %	48.94 %

Accuracy, Precision and Recall



Precision and Recall Comparison



Thank You

