

# 1.Result:

## 1.1 Review Cleaning (Text processing results)

	rating	recommended	text	Clean_text
0	4	1	Absolutely wonderful - silky and sexy and comfy...	absolutely wonderful silky sexy comfortable
1	5	1	Love this dress! it's sooo pretty. i happened to find store im ...	love dress sooo pretty happened find store im ...
2	3	0	I had such high hopes for this dress and really wanted to work initially o...	high hope dress really wanted work initially o...
3	5	1	I love, love, love this jumpsuit. it's fun, flirty, fabulous ev...	love love love jumpsuit fun flirty fabulous ev...
4	5	1	This shirt is very flattering to all due to the adjustable front tie perf...	shirt flattering due adjustable front tie perf...

Fig 1 – Text Processing Results

## 1.2 Positive-Negative Polarity Calculation:

	rating	recommended	text	Clean_text	class_by_rating	polarity_by_reviews	sentiment_score	sentiment_by_review	label_sent
1	5	1	Love this dress! it's sooo pretty. i happened to find store im ...	love dress sooo pretty happened find store im ...	Positive	{'neg': 0.0, 'neu': 0.503, 'pos': 0.497, 'comp...}	0.9710	Positive	1
3	5	1	I love, love, love this jumpsuit. it's fun, flirty, fabulous ev...	love love love jumpsuit fun flirty fabulous ev...	Positive	{'neg': 0.179, 'neu': 0.183, 'pos': 0.638, 'co...}	0.9437	Positive	1
4	5	1	This shirt is very flattering to all due to the adjustable front tie perf...	shirt flattering due adjustable front tie perf...	Positive	{'neg': 0.0, 'neu': 0.494, 'pos': 0.506, 'comp...}	0.9062	Positive	1
5	2	0	I love tracy reese dresses, but this one is no...	love tracy reese dress one petite foot tall us...	Negative	{'neg': 0.0, 'neu': 0.732, 'pos': 0.268, 'comp...}	0.9153	Positive	1
6	5	1	I aded this in my basket at hte last mintue see would look lik...	aded basket hte last mintue see would look lik...	Positive	{'neg': 0.049, 'neu': 0.824, 'pos': 0.127, 'co...}	0.6361	Positive	1

Fig 2 – Positive Negative Categorization Results

```
rev_df['sentiment_by_review'].value_counts()
```

```
Positive    13434
Negative    1476
Name: sentiment_by_review, dtype: int64
```

Fig 3 – Positive Negative Review Count

### 1.3 Classification result

Method / Model		Recom.	No	Yes	Accuracy	Support	Precision	Recall	f1 Score
Using count vectorizer									
MULTINOMIAL NAIVE BAYES	Actual	0	580	87	0.9475	667	0.80	0.87	0.83
		1	148	3658	0.9475	3806	0.98	0.96	0.97
SUPPORT VECTOR MACHINES	Actual	0	526	141	0.9479	667	0.85	0.79	0.82
		1	92	3714	0.9479	3806	0.96	0.98	0.97
RANDOM FOREST CLASSIFIER	Actual	0	319	348	0.9148	667	0.91	0.48	0.63
		1	33	3773	0.9148	3806	0.92	0.99	0.95
NEURAL NETWORK MODEL	Actual	0	521	146	0.9448	667	0.84	0.78	0.81
		1	101	3705	0.9448	3806	0.96	0.97	0.97
ADA BOOSTING CLASSIFIER	Actual	0	398	269	0.9191	667	0.81	0.60	0.69
		1	93	3713	0.9191	3806	0.93	0.98	0.95
KNN	Actual	0	39	628	0.8556	667	0.68	0.06	0.11
		1	18	3788	0.8556	3806	0.86	1.00	0.92

Table 1 - Classification Report

### 1.4 Charts & result

Fig 7 – Classification Chart – Bag of Words (ML models)

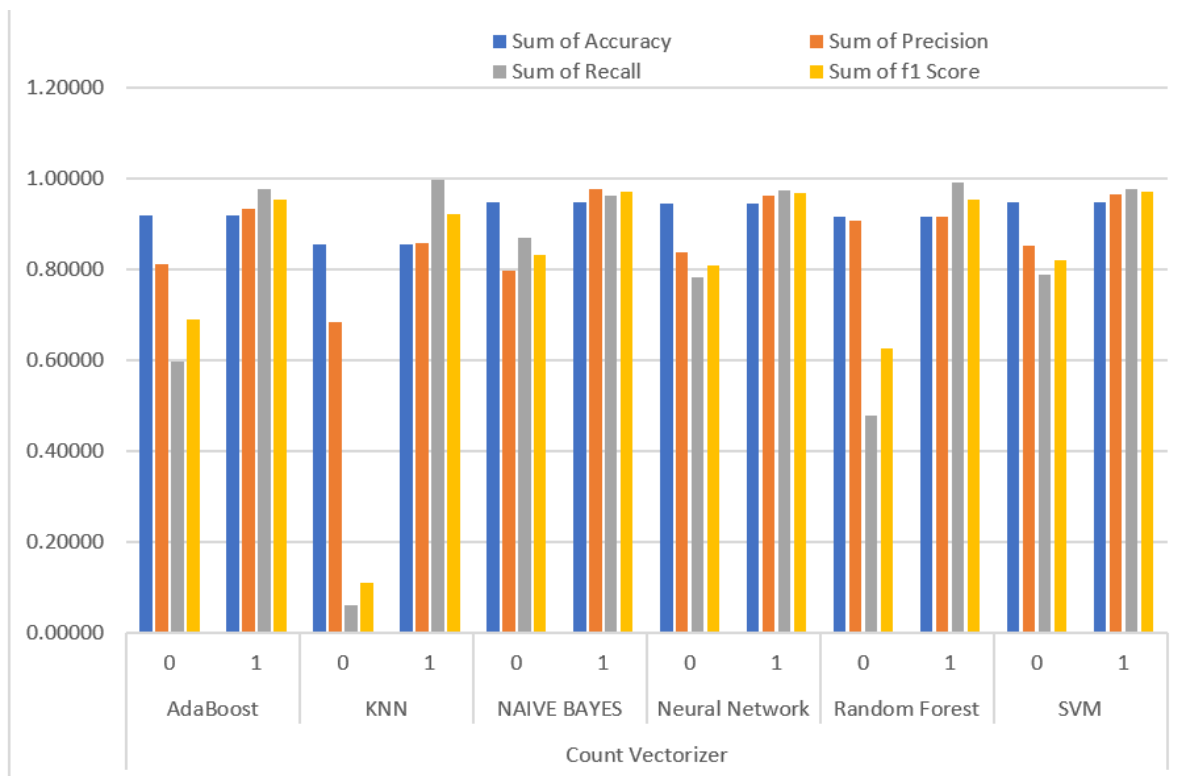


Fig 4 – Classification Chart – Count Vectorizer (ML models)

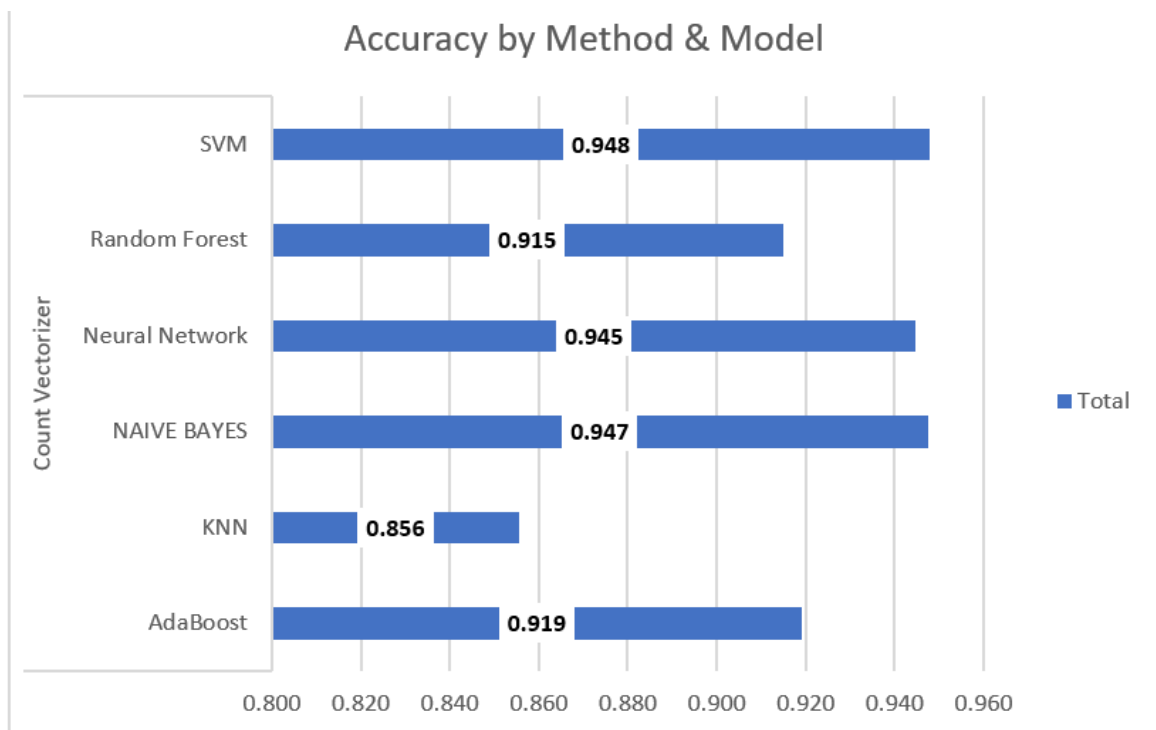


Fig 5 - Accuracy Comparison – Count Vectorizer (ML Models)

## 1.5 CONCLUSION:

- Sentiment of customer reviews were analyzed, and it was observed that there were 13434 positive reviews and 1476 negative reviews.
- Six different Machine Learning models were built and compared based on their accuracy and classification report.
- The comparative analysis clearly shows that Support Vector Machine (SVM) was the best classification model with the highest accuracy of 94.79%.