DETECTING SPAM MESSAGES: A COMPARATIVE ANALYSIS OF MACHINE LEARNING MODELS LEVERAGING NAIVE BAYES AND SUPPORT VECTOR MACHINES FOR ACCURATE SPAM DETECTION

Jasmine Kaur 4142340

Objective of the report

To compare machine learning models for spam detection

Dataset used

Dataset has been taken for kaggle

Models used

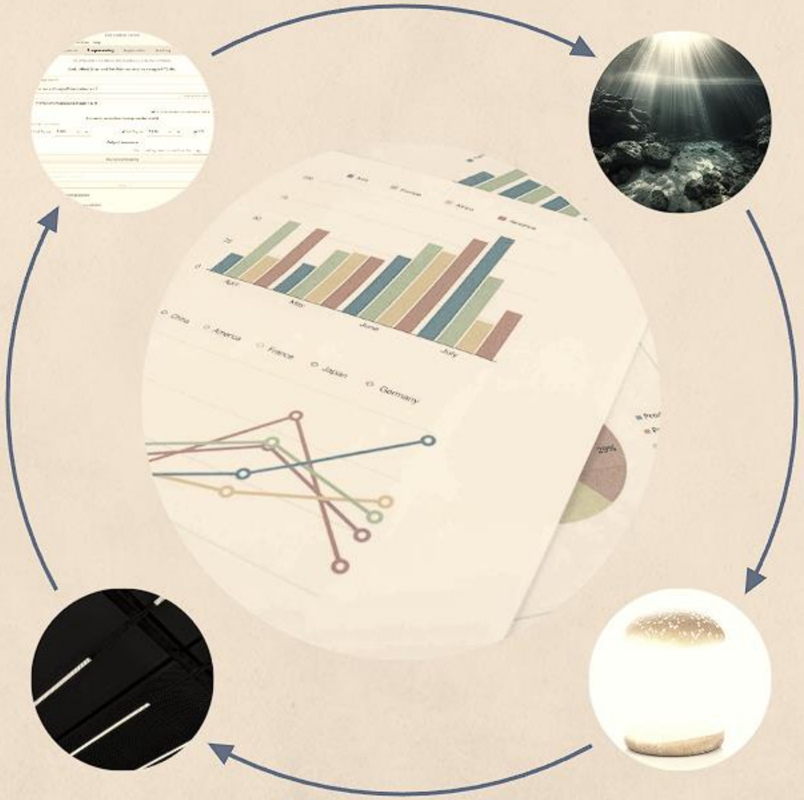
Naive Bayes and Support Vector Machines

Evaluation Metrics used

Accuracy, Precision, Recall, F1-Score

DATASET

Description of the Dataset



Preprocessing

Missing values: There were no missing values so it made the preprocessing task easier

Source

Publicly available spam message dataset

Features

The dataset had only two columns
one column had features of spam and not spam and
the other columns had messages related to them

Size

the data set had 5572 rows
in which 4825 were not spam messages and 747 were
spam messages so we tackled the imbalanced class
and oversampled the spam objects to 4825 as well

NAIVE BAYES MODEL

Performance and Evaluation

01 Training

80% of the dataset used for model training

02 Testing

20% of the dataset used for model evaluation

03 Accuracy

0.9811659192825112
that is 98%

04 Precision

Spam was 0.99
and not spam was 0.98

05 Recall

Spam was 0.87 and not spam was 1

06 F1-Score

Spam was 0.99 and for not spam was 0.93

SUPPORT VECTOR

Performance and Evaluation

01 Training

80% of the dataset used for model training

02 Testing

20% of the dataset used for model evaluation

03 Accuracy

0.979372197309417
that is 97%

04 Precision

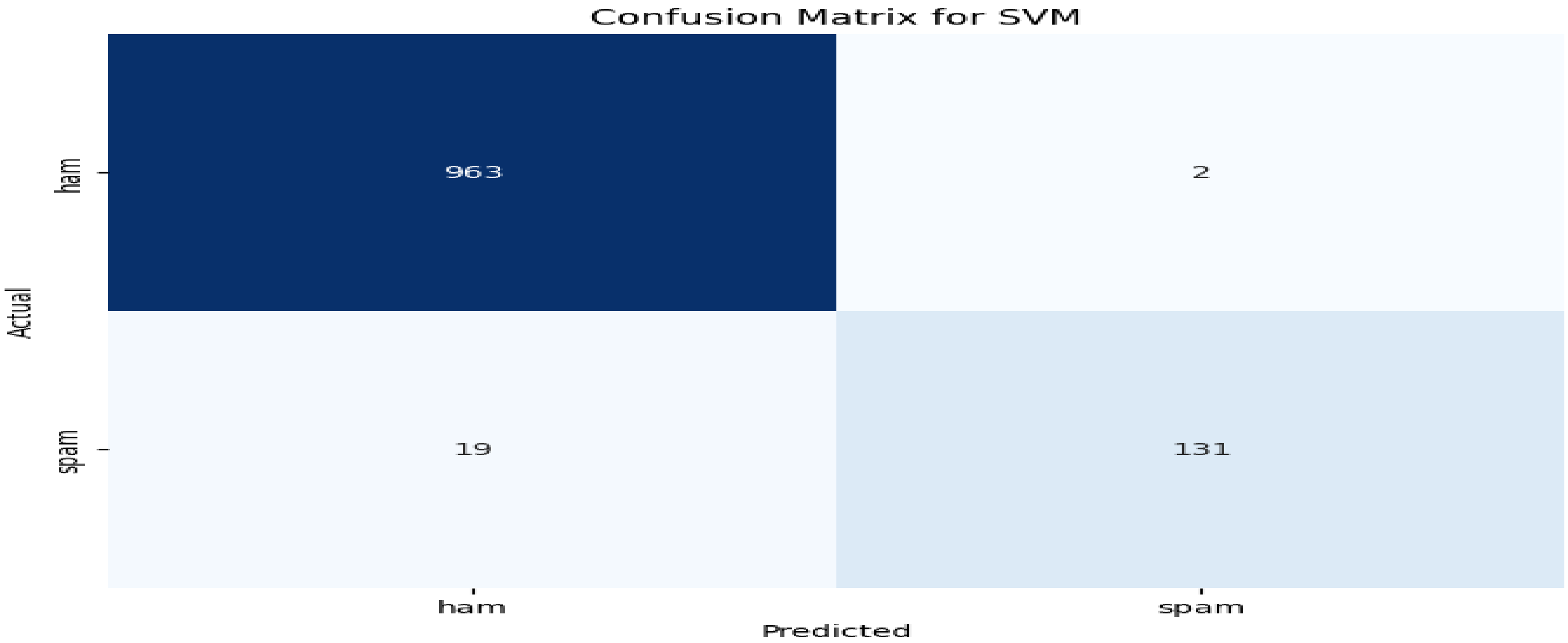
Spam was 0.98
and not spam was 0.98

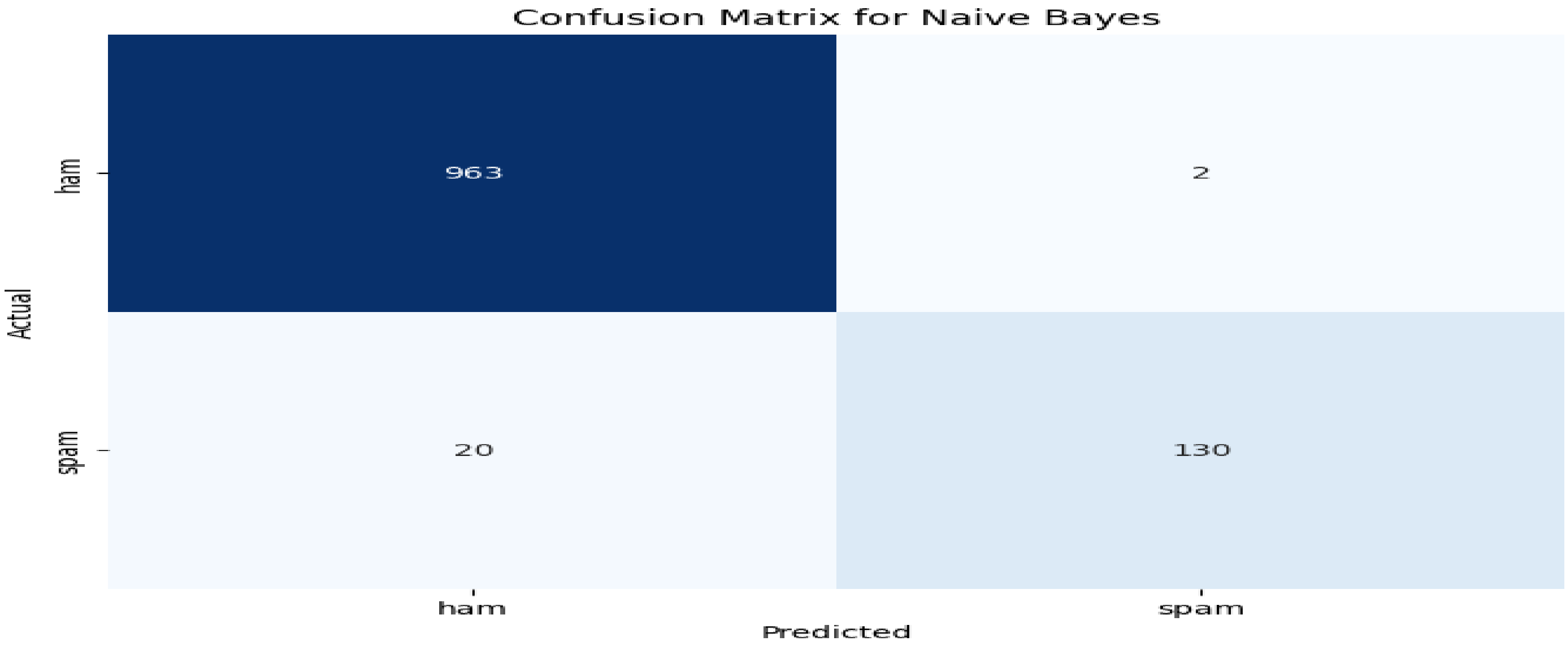
05 Recall

Spam was 0.87 and not spam was 1

06 F1-Score

Spam was 0.92 and for not spam was 0.99





TOP-LEFT (TRUE NEGATIVES): IN THIS CELL, WE HAVE 963 INSTANCES OF "HAM" (LEGITIMATE MESSAGES) THAT WERE CORRECTLY CLASSIFIED AS "HAM." THIS SHOWS THAT THE NAIVE BAYES MODEL ACCURATELY IDENTIFIED MOST LEGITIMATE MESSAGES.TOP-RIGHT (FALSE POSITIVES): IN THIS CELL, THERE IS 1 INSTANCE OF A "HAM" MESSAGE THAT WAS INCORRECTLY CLASSIFIED AS "SPAM." IT'S A RELATIVELY LOW NUMBER, INDICATING THAT THE MODEL DIDN'T MAKE MANY FALSE POSITIVE ERRORS.BOTTOM-LEFT (FALSE NEGATIVES): HERE, WE SEE 20 INSTANCES OF "SPAM" MESSAGES THAT WERE INCORRECTLY CLASSIFIED AS "HAM." THESE ARE MESSAGES THAT THE MODEL SHOULD HAVE IDENTIFIED AS SPAM BUT DIDN'T. THIS IS A RELATIVELY LOW NUMBER OF FALSE NEGATIVES.BOTTOM-RIGHT (TRUE POSITIVES): IN THIS CELL, 130 INSTANCES OF "SPAM" MESSAGES WERE CORRECTLY CLASSIFIED AS "SPAM." THIS DEMONSTRATES THAT THE MODEL EFFECTIVELY DETECTED THE MAJORITY OF SPAM MESSAGES.OVERALL, THE NAIVE BAYES MODEL SHOWS A HIGH LEVEL OF ACCURACY, WITH ONLY A SMALL NUMBER OF FALSE POSITIVES AND FALSE NEGATIVES. HOWEVER, THERE IS SOME ROOM FOR IMPROVEMENT IN REDUCING THE NUMBER OF FALSE NEGATIVES, WHERE LEGITIMATE MESSAGES ARE INCORRECTLY CATEGORIZED AS SPAM.

COMPARISON

Comparison of Naive Bayes and SVM

Metric	Naive Bayes	Support Vector Machines
Accuracy	98%	97%
Precision	98%	98%
Recall	100%	100%
F1-Score	99%	99%

TAKE AWAY FROM THE INITIAL MODEL

Key Takeaways

Both models perform well in spam detection

Both Naive Bayes and SVM achieve high accuracy and precision

Naive Bayes SVM outperforms in accuray

naive bayes initial model had better accuracy than the SVM Model

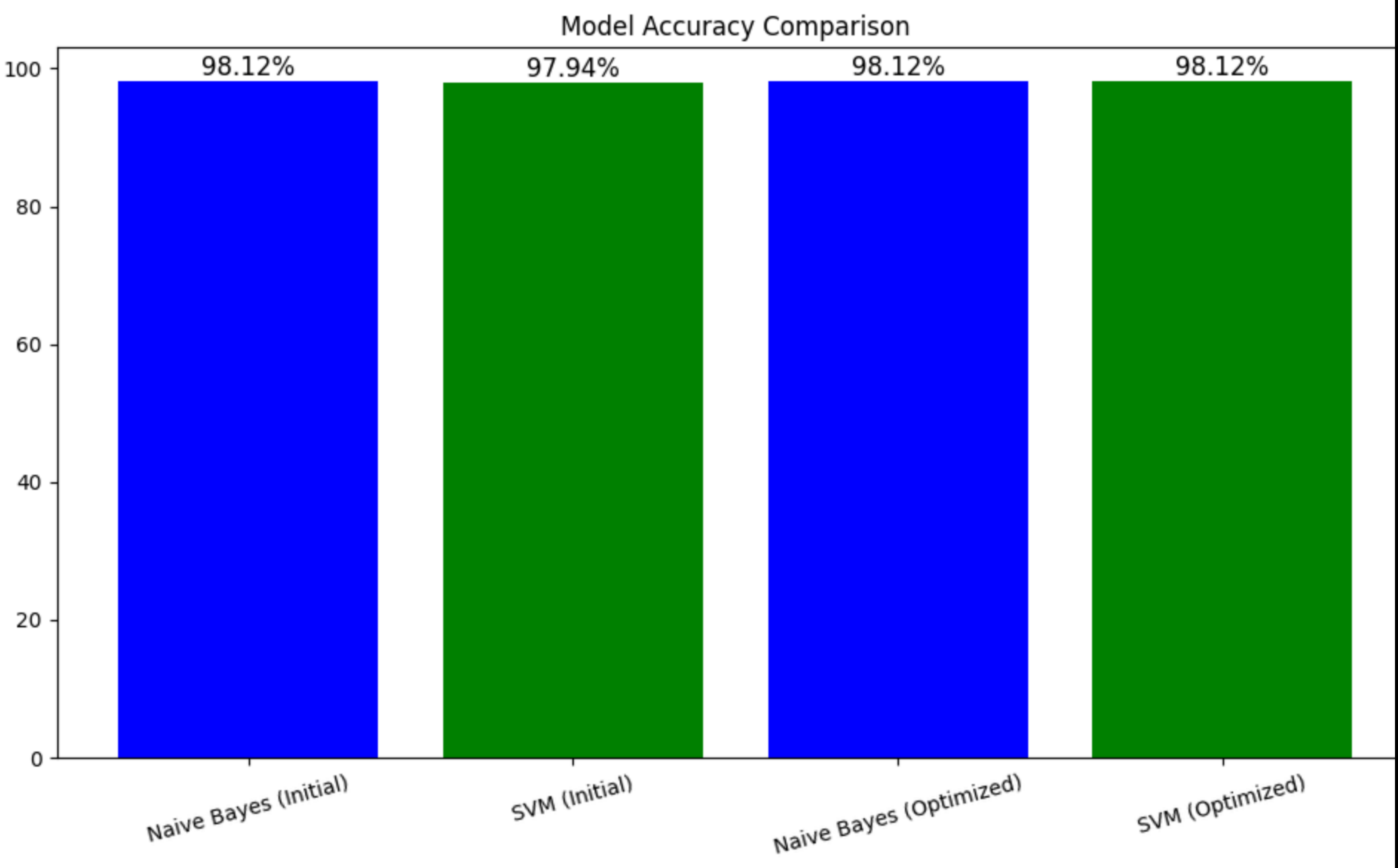
Considerations for Real-world Deployment

Scalability, computational resources, and model training time

COMPARISON AFTER TUINING THE MODELS

Comparison of Naive Bayes and SVM after hyperparameter tuining

Metric	Naive Bayes	Support Vector Machines
Accuracy	98%	98%
Precision	98%	98%
Recall	100%	100%
F1-Score	99%	99%


Model Accuracy Comparison
98.12%
97.94%
98.12%
98.12%
100
80
60
40
20
0
Naive Bayes (Initial)
SVM (Initial)
Naive Bayes (Optimized)
SVM (Optimized)

Both Naive Bayes and SVM can effectively fit the dataset, with Naive Bayes performing slightly better in this particular use case. It's essential to consider the trade-offs between model complexity, computational resources, and ease of use when selecting the appropriate algorithm for spam detection. In this context, Naive Bayes serves as a strong primary algorithm, and SVM as a secondary alternative.