

INTRODUCTION

1.1 OVERVIEW

Over recent years, as the popularity of mobile phone devices has increased, Short Message Service (SMS) has grown into a multi-billion dollar industry. At the same time, reduction in the cost of messaging services has resulted in growth in unsolicited commercial advertisements (spams) being sent to mobile phones. Due to Spam SMS, Mobile service providers suffer from some sort of financial problems as well as it reduces calling time for users. Unfortunately, if the user accesses such Spam SMS, they may face the problem of virus or malware. When SMS arrives at mobile it will disturb mobile user privacy and concentration. It may lead to frustration for the user. So, Spam SMS is one of the major issues in the wireless communication world and it grows day by day. To avoid such Spam SMS, people use white and black list of numbers. But this technique is not adequate to completely avoid Spam SMS. To tackle this problem, it is needful to use a smarter technique which correctly identifies Spam SMS. Natural language processing technique is useful for Spam SMS identification. It analyses text content and finds patterns which are used to identify Spam and Non-Spam SMS.

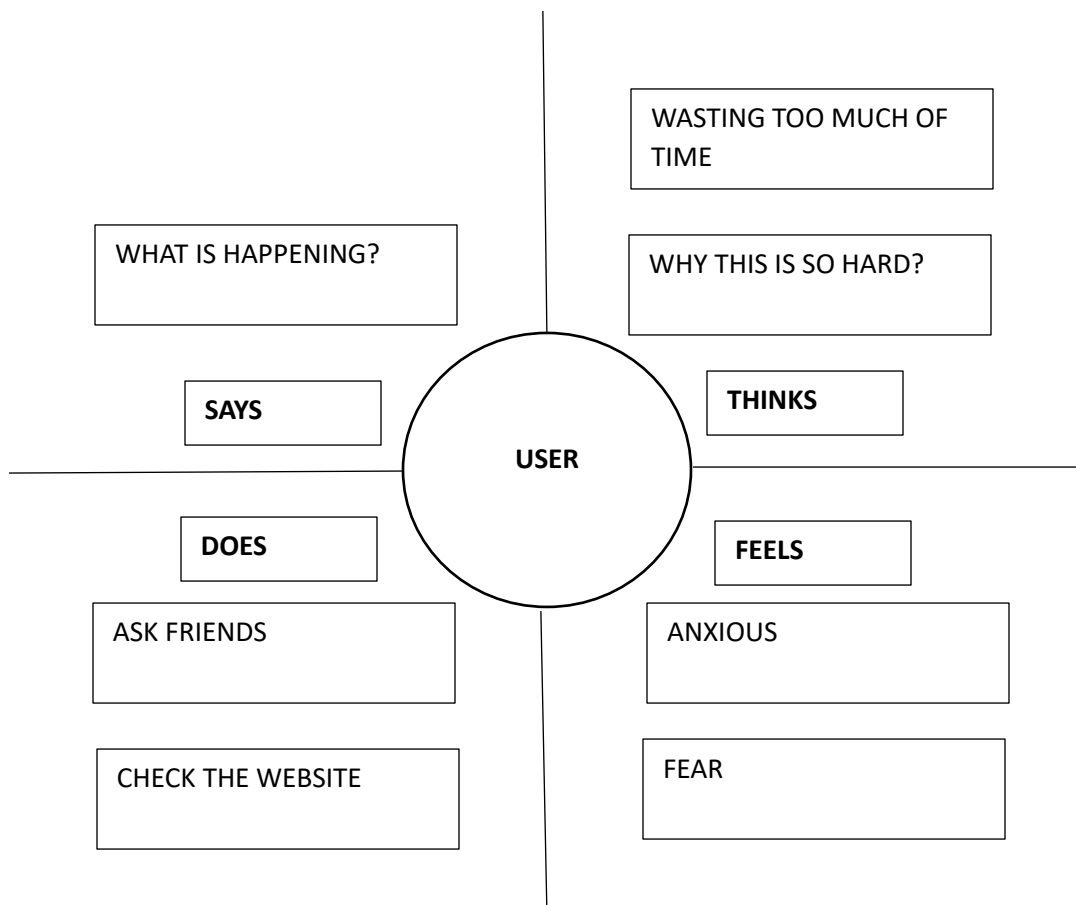
1.2 PURPOSE

Machine learning algorithms use statistical models to classify data. In the case of spam detection, a trained machine learning model must be able to determine whether the sequence of words found in an SMS are closer to those found in spam SMSs or safe ones. An SMS spam filter is a tool used in SMS hosting software that churns out unsolicited, unwanted, and virus-infested SMSs and keeps such SMSs off of the user's inbox. This protects the user from any potential cyber threat and facilitates smooth communications and workflow.

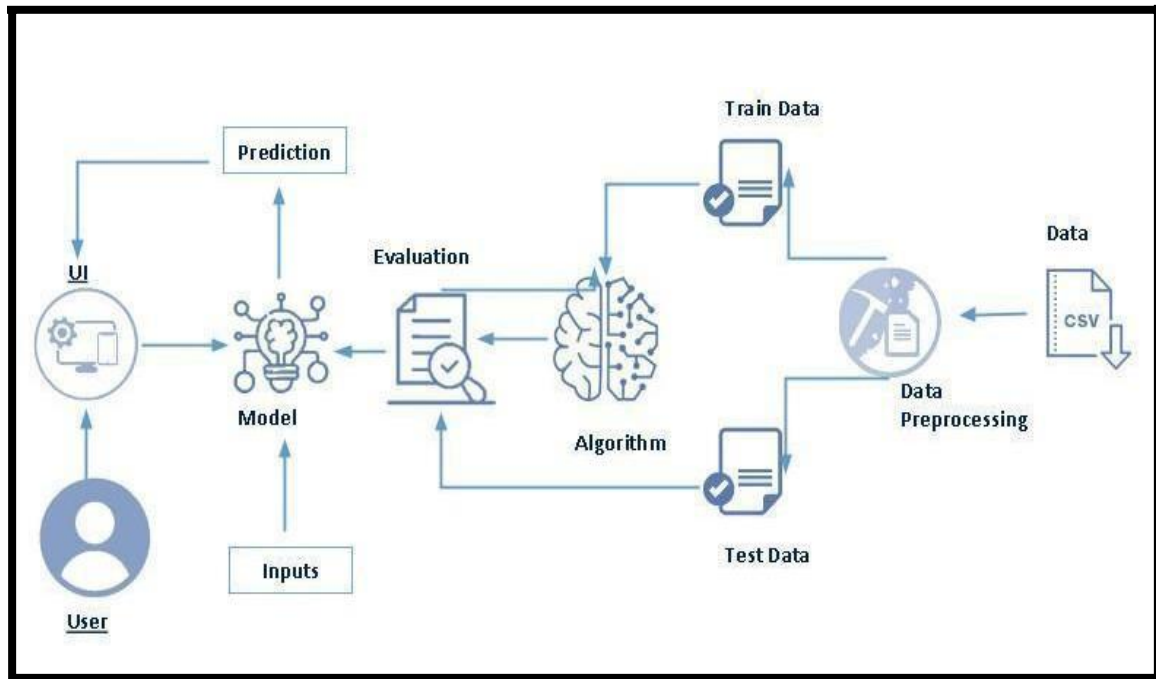
PROBLEM DEFINITION AND DESIGN THINKING

2.1 EMPATHY MAP

An empathy map is a square divided into four quadrants with the user or client in the middle. Each of the four quadrants comprises a category that helps us delve into the mind of the user. The four empathy map quadrants look at what the user says, thinks, feels, and does.

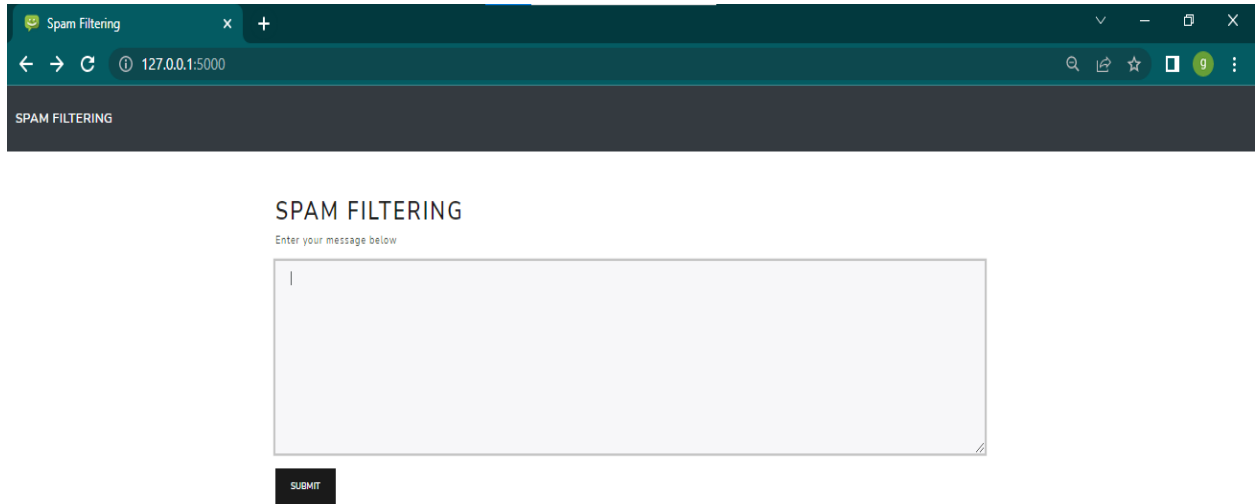


2.2 IDEATION MAP

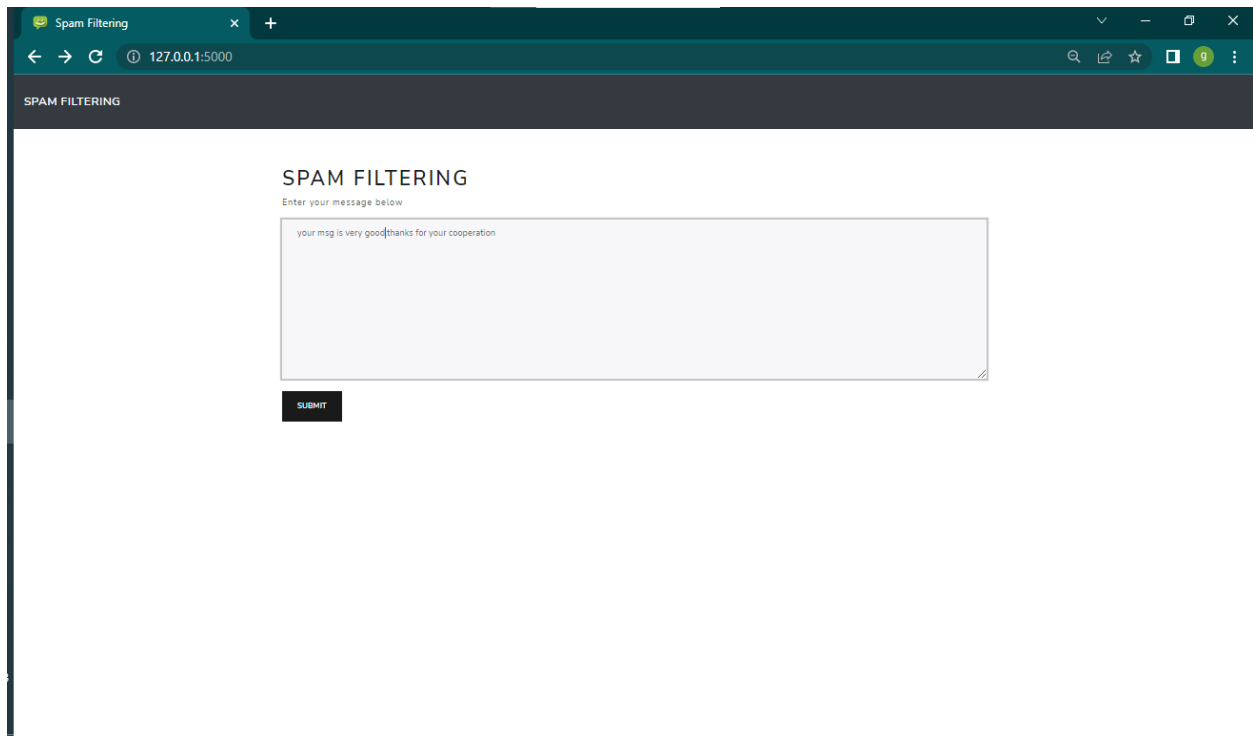


RESULT

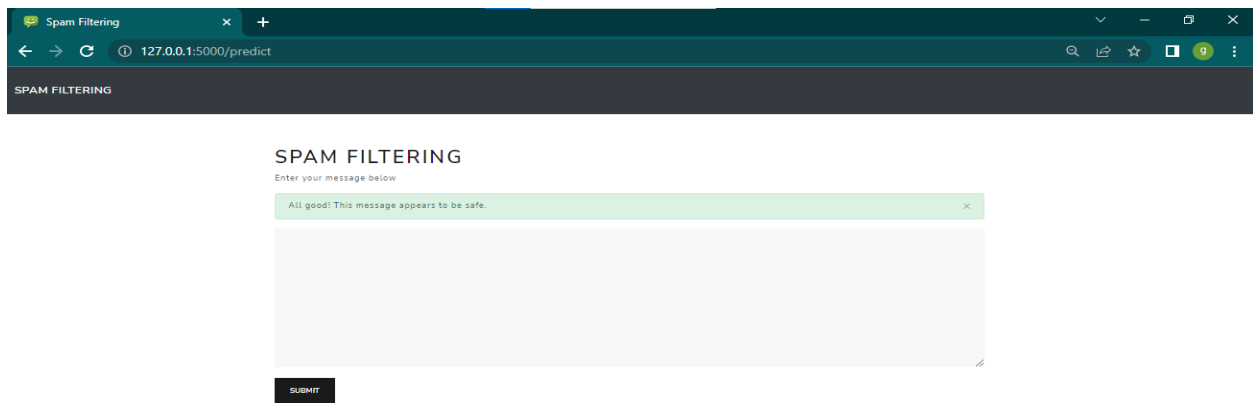
3.1 OUTPUT OF THE PROJECT



HOME PAGE



INPUT



OUTPUT (PREDICTION)

ADVANTAGES AND DISADVANTAGES

4.1 ADVANTAGES

- With the benefits of SMS spam filters, the security risk can be reduced since the user gets in hand the SMSs that have gone through various spam checks. Moreover, these SMS spam filters throw out malware, malicious, and virus-infested SMSs and protect user security.
- Having to filter out the 20% important SMSs out of the average 80% clutter does seem time-consuming. This can be of greater concern if these stats are put into an organization's SMSing communications. By filtering out the important SMSs and sending to the spam box the junk SMSs, an SMS spam filter saves time for the user and keeps the business communications going by streamlining the user inbox.
- Increased Productivity.

4.2 DISADVANTAGES

- Spam emails can be the source of a great amount of malware like viruses, Trojans, worms, and others which are specifically designed to disrupt or damage computer systems.
- Spam can lead to being shut off by Internet providers.
- Spamming has a low ROI.
- Spam mails can bring tangible profit only to their providers but they don't work as a means of email marketing. On the other hand, referring to an email marketing agency is likely to get everything right and meet one's marketing needs.

APPLICATIONS

5.1 APPLICATIONS

- This system can employed in all the mobile phones.
- Used in laptop.
- Used in Personal Computer.
- Used in emails.

CONCLUSION

6.1 CONCLUSION

In the last two decades, spam detection and filtration gained the attention of a sizeable research community. The reason for a lot of research in this area is its costly and massive effect in many situations like consumer behavior and fake reviews. This covers various machine learning techniques and models that the various researchers have proposed to detect and filter spam in emails and IoT platforms.

FUTURE SCOPE

7.1 FUTURE SCOPE

We plan to further carry out the machine learning algorithms to optimize and compare with different bio-inspired algorithms such as Firefly, Bee Colony and Ant Colony Optimization as researched in the previous sections. We could also explore the Deep learning Neural Network with PSO and GA by exploring different libraries such as Tensor Flow's DNN Classifier or similar.

APPENDIX

8.1 SOURCE CODE

```
from flask import Flask, render_template, request
```

```
import joblib
```

```
app = Flask(__name__)
```

```
model = joblib.load('models/saved/model.joblib')
```

```
encoder = joblib.load('models/saved/encoder.joblib')
```

```
@app.route('/')
```

```
def main():
```

```
    return render_template("index.html")
```

```
@app.route('/predict', methods=["POST"])
```

```
def predict():
```

```
    if request.method == "POST":
```

```
        message = request.form['submission']
```

```
        prediction = model.predict([message])
```

```
classification = encoder.inverse_transform(prediction)
```

```
    return render_template('index.html', message=message,  
classification=classification)
```

```
if __name__ == "__main__":
```

```
    app.run(debug=True)
```