



S.Getsy Jacinth

FINAL PROJECT

SMS - TEXT CLASSIFIER

AGENDA

- ▶ Utilizing generative AI techniques, this project aims to develop a model for classifying SMS messages as spam or ham (non-spam). By automating spam detection, the objective is to enhance user experience by reducing unwanted messages and protecting against potential security threats. Evaluation will be based on accuracy, precision, recall, and F1-score metrics to ensure the effectiveness of the Naive Bayes algorithm in real-world scenarios.

PROBLEM STATEMENT

The proliferation of unsolicited SMS messages, commonly known as spam, presents a significant inconvenience and potential security risk for mobile users. Traditional methods of spam detection often fall short, resulting in users being bombarded with unwanted messages containing scams, phishing attempts, and malicious links. This project addresses the challenge of effectively identifying and filtering out spam messages using the Naive Bayes algorithm within the realm of generative AI. By automating the classification process, the goal is to provide users with a more streamlined and secure messaging experience while minimizing the impact of unwanted communications on their daily activities.

PROJECT OVERVIEW

- ▶ Focusing on generative AI and employing the Naive Bayes algorithm, this project aims to develop a robust solution for classifying SMS messages as spam or ham. By leveraging natural language processing techniques, the model automates the detection of unwanted messages, enhancing user experience and security. Evaluation metrics such as accuracy, precision, recall, and F1-score will be used to assess the model's performance in real-world scenarios. Ultimately, the project seeks to mitigate the inconvenience and potential risks associated with unsolicited SMS communications, contributing to a safer and more enjoyable mobile messaging environment.

WHO ARE THE END USERS?

- ▶ The end users of this project encompass a diverse spectrum, including everyday mobile phone users, business professionals relying on SMS for communication, mobile application developers integrating SMS functionality, telecommunication providers managing SMS infrastructure, and security-conscious individuals prioritizing privacy. By delivering an effective SMS classification solution, the project aims to enhance the messaging experience for users across various domains, ensuring efficient communication free from unwanted spam and security threats.

SOLUTION AND ITS VALUE PROPOSITION

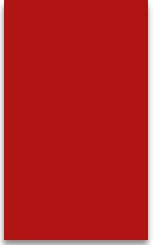
- ▶ Our solution employs generative AI techniques coupled with the Naive Bayes algorithm to classify SMS messages as either spam or ham. By harnessing natural language processing capabilities, our model automates the detection of unwanted messages, enhancing user experience and security. The value proposition lies in providing users with a streamlined messaging experience, free from the inconvenience and potential risks associated with spam messages. Additionally, our solution contributes to a safer mobile communication environment by mitigating the impact of malicious content. Ultimately, we aim to deliver a reliable and efficient tool that empowers users to communicate with confidence and peace of mind.

THE WOW IN THE SOLUTION

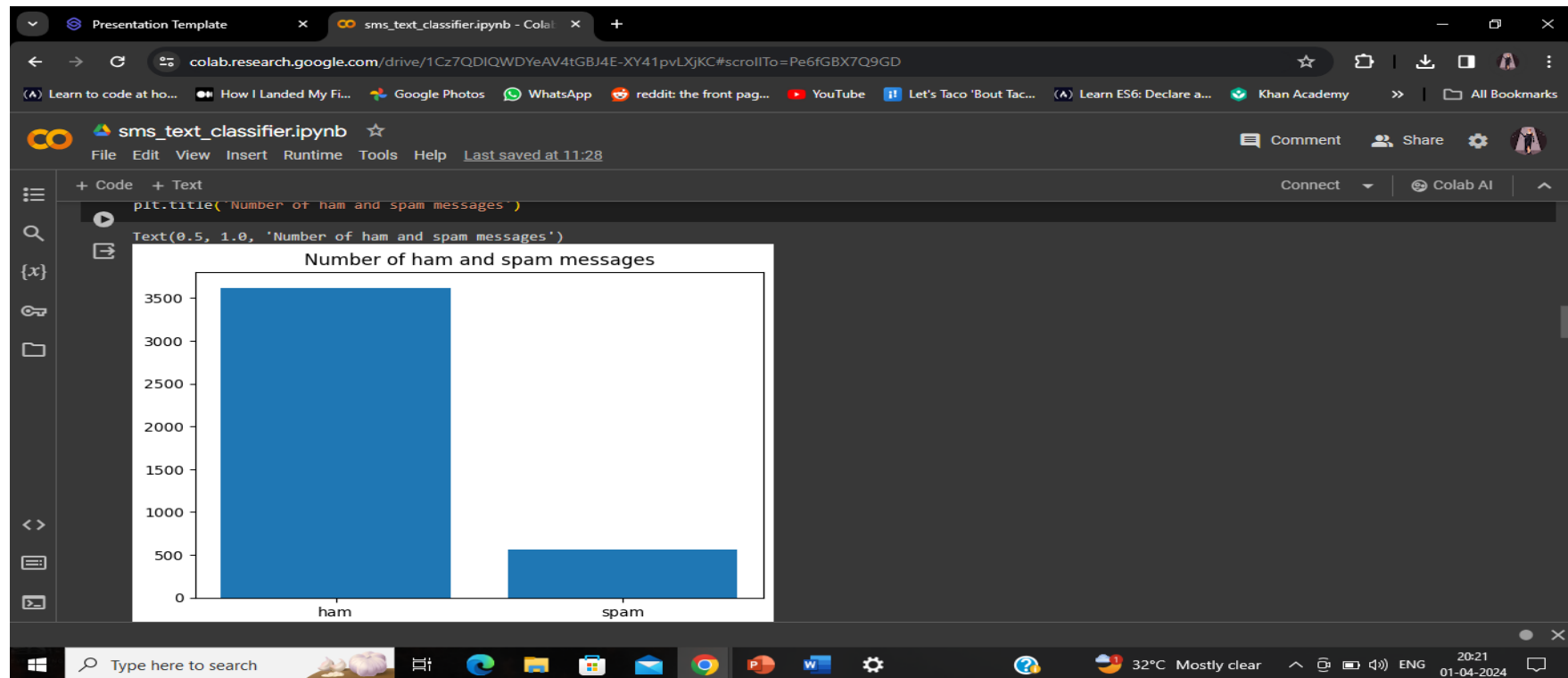
- ▶ Our solution offers users a revolutionary approach to SMS classification, seamlessly integrating advanced generative AI techniques with the Naive Bayes algorithm. Through dynamic adaptability, it continuously evolves to combat emerging spamming techniques, ensuring unparalleled accuracy and reliability. Users benefit from an effortless integration into their mobile communication experience, with spam messages filtered out seamlessly, allowing for uninterrupted focus on important conversations. This user-centric design prioritizes efficiency and security, offering a transformative experience that sets a new standard in SMS communication, empowering users with a streamlined and hassle-free messaging experience unlike any other.

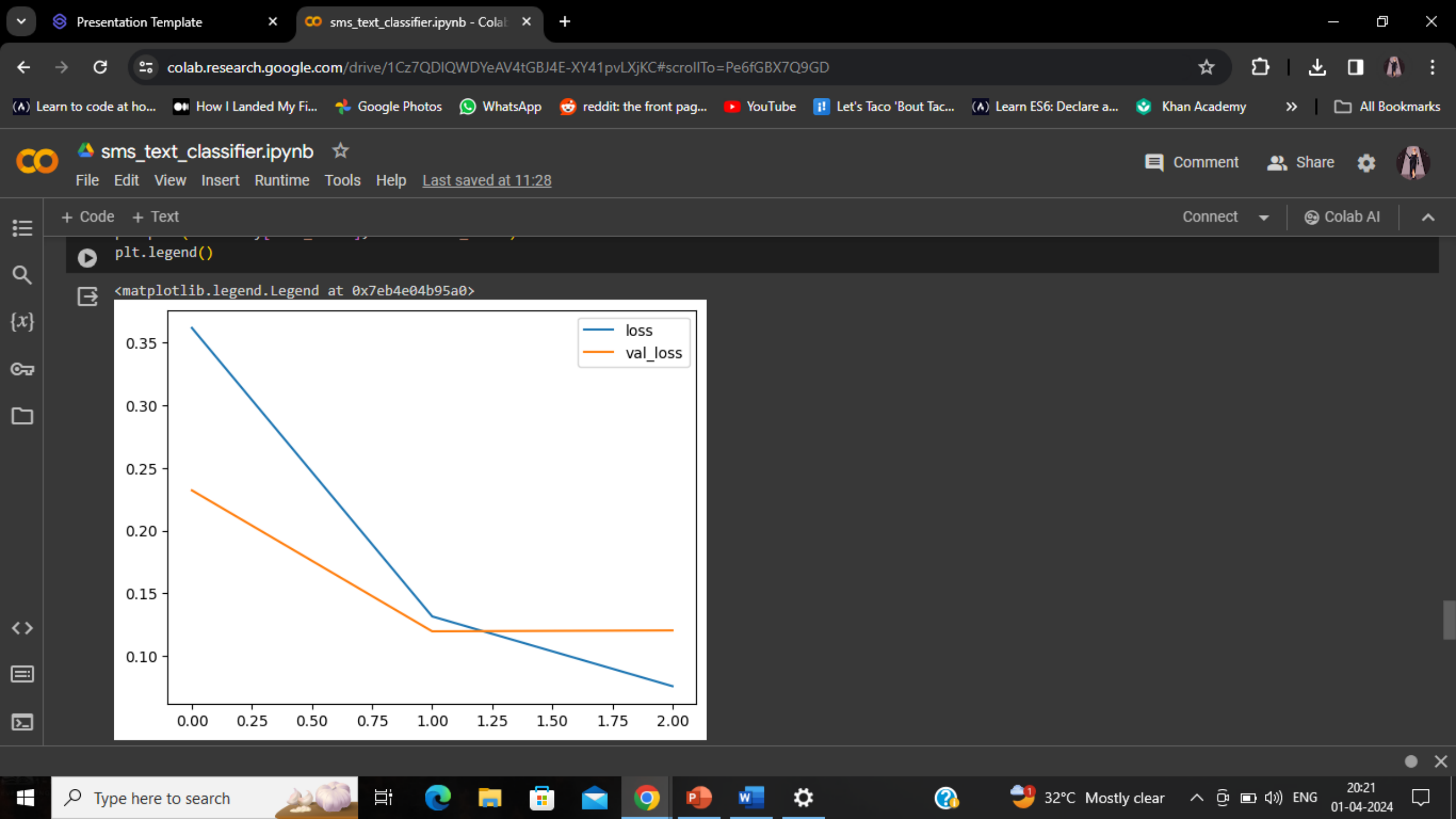
MODELLING

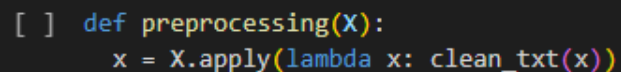
1. **Data Preprocessing:** The project initiates with data preprocessing, where raw SMS data undergoes cleaning processes such as removing punctuation and converting text to lowercase. This step ensures consistency and uniformity in the dataset.
2. **Feature Extraction:** Text data is transformed into numerical feature vectors using the TF-IDF vectorizer. This process converts textual information into a format suitable for machine learning algorithms, enabling further analysis.
3. **Algorithm Selection:** The Multinomial Naive Bayes algorithm is chosen for classification due to its effectiveness in text classification tasks. Its simplicity and ability to handle high-dimensional data make it suitable for this project.
4. **Model Training:** The training phase involves feeding preprocessed data into the Naive Bayes classifier to learn patterns and relationships between features and target labels (spam or ham). The model adjusts its parameters to optimize classification accuracy.

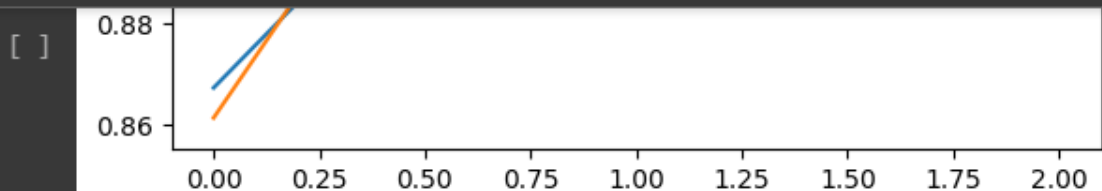
- 
5. **Model Evaluation:** The trained model is evaluated using metrics such as accuracy, precision, recall, and F1-score on a separate testing dataset. This step assesses the model's performance and generalization ability on unseen data.
 6. **Hyperparameter Tuning:** Fine-tuning of model parameters, such as smoothing techniques in Naive Bayes, may be performed to optimize performance further. This step involves experimentation to find the most suitable configuration for the model.
 7. **Cross-Validation:** Cross-validation techniques, such as k-fold cross-validation, may be employed to validate the model's robustness and mitigate overfitting.
 8. **Deployment:** Once the model demonstrates satisfactory performance, it can be deployed for real-world applications, where it classifies incoming SMS messages as spam or ham in real-time, contributing to improved user experience and security.

RESULTS









```
def preprocessing(X):
    x = X.apply(lambda x: clean_txt(x))
    x = t.texts_to_sequences(x)
    return sequence.pad_sequences(x, maxlen=max len)
```

```
s = model.evaluate(preprocessing(df_test['x']), y_test)
```

```
44/44 [=====] - 3s 68ms/step - loss: 0.0624 - accuracy: 0.9878
```

```
print('Loss: {:.3f}, Accuracy: {:.3f}'.format(s[0], s[1]))
```

Loss: 0.062, Accuracy: 0.988

CONCLUSION

- ▶ In conclusion, our project presents a robust solution for SMS classification, leveraging generative AI techniques and the Naive Bayes algorithm to effectively distinguish between spam and ham messages. Through rigorous data preprocessing, feature extraction, and model training, we have developed a highly accurate and efficient classification model. By prioritizing user-centric design and seamless integration, our solution enhances the messaging experience, ensuring users can communicate with confidence and security. With its transformative impact on mobile communication, our project sets a new standard for efficiency and reliability, empowering users with a streamlined and hassle-free messaging experience.



THANK YOU!