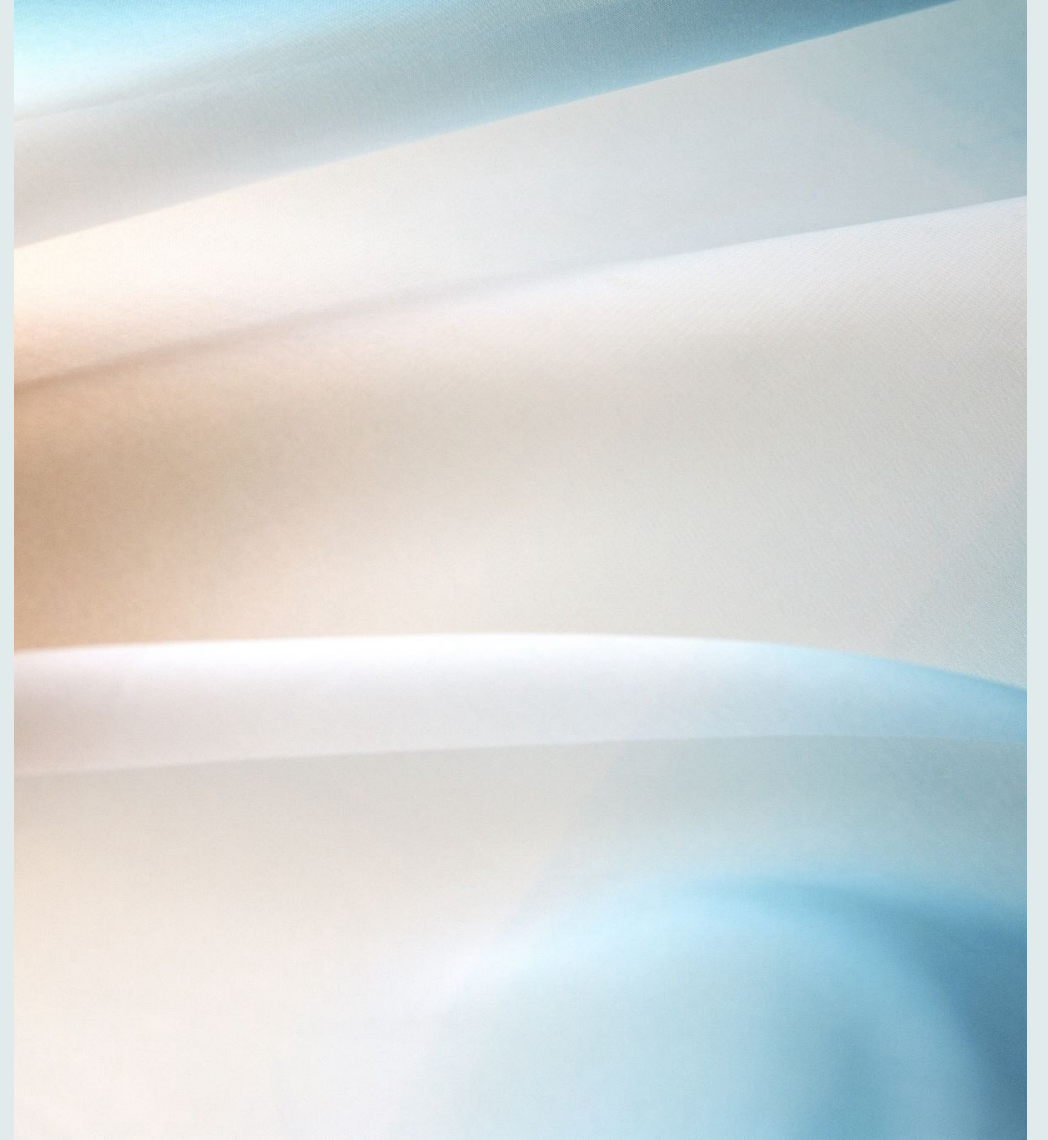


SPAM MESSAGE FILTER

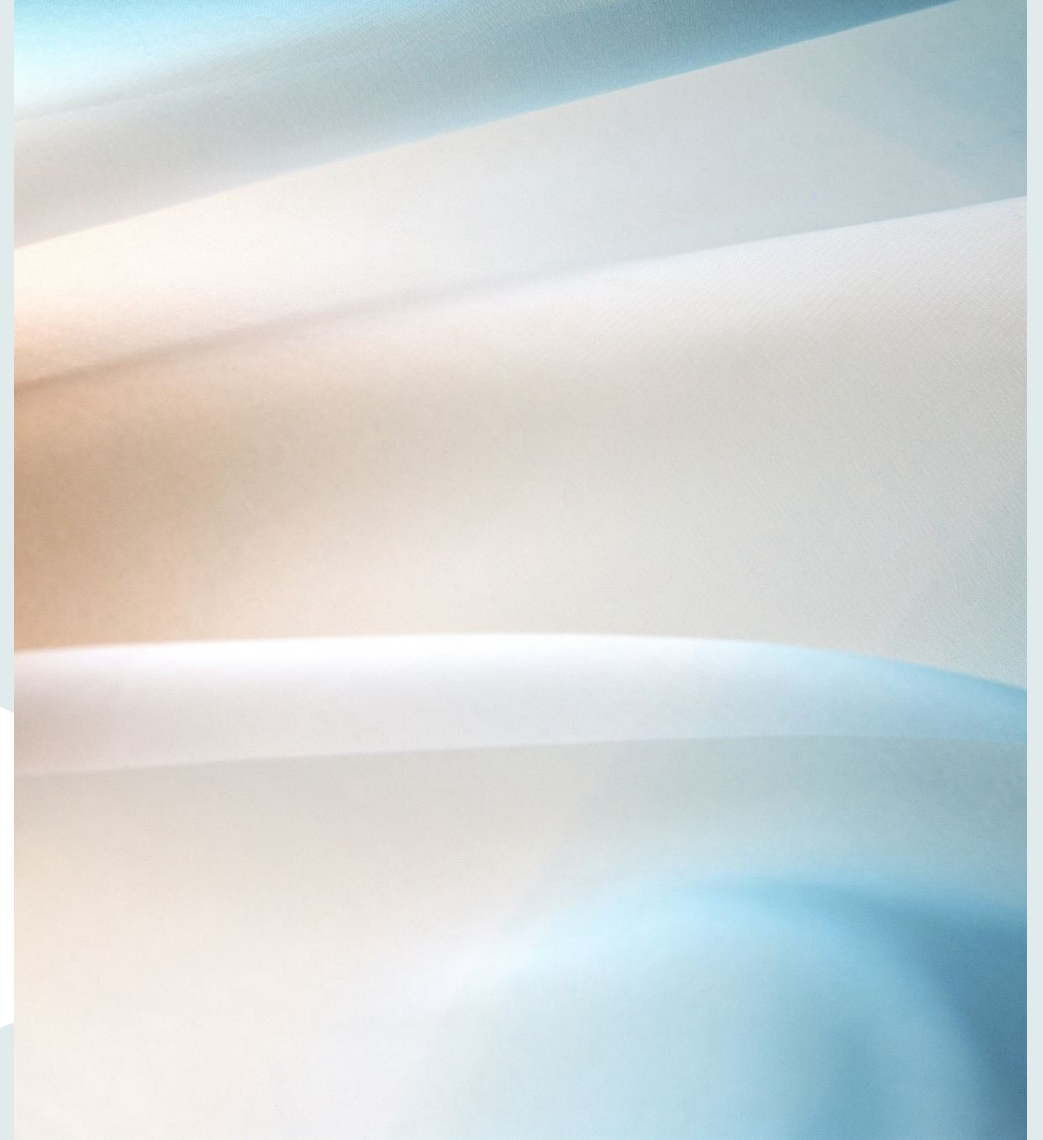
*Leveraging Machine Learning for
Safer Communication*

by Alperen Unal, 2024



INTRODUCTION

- **Spam SMS Problem:** Unsolicited messages ranging from harmless ads to malicious phishing attempts.
- **Importance of Spam Detection:** Essential for protecting personal and organizational security from privacy invasion and financial loss.
- **NLP and Machine Learning Solutions:** Utilizing text analysis and model training to automatically classify and filter spam SMS.



DATASET DESCRIPTION

- SMS Spam Collection Dataset
- <https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>

	v1		v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
0	ham	Go until jurong point, crazy.. Available only ...		NaN	NaN	NaN
1	ham	Ok lar... Joking wif u oni...		NaN	NaN	NaN
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...		NaN	NaN	NaN
3	ham	U dun say so early hor... U c already then say...		NaN	NaN	NaN
4	ham	Nah I don't think he goes to usf, he lives aro...		NaN	NaN	NaN

Spam message example:

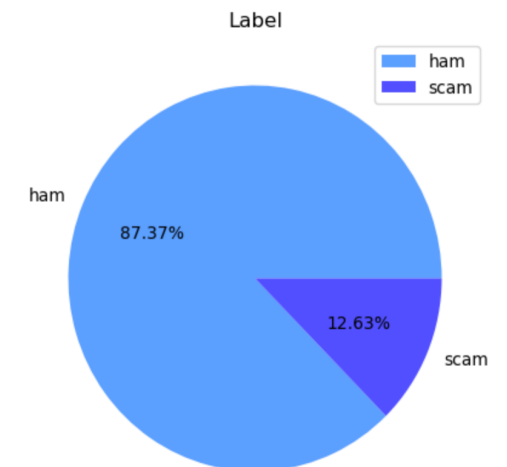
Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)T&C's apply 08452810075over18's

Ham message example:

Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...

	v1		v2
0	ham	Go until jurong point, crazy.. Available only ...	
1	ham	Ok lar... Joking wif u oni...	
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...	
3	ham	U dun say so early hor... U c already then say...	
4	ham	Nah I don't think he goes to usf, he lives aro...	
...
5567	spam	This is the 2nd time we have tried 2 contact u...	
5568	ham	Will _b going to esplanade fr home?	
5569	ham	Pity, * was in mood for that. So...any other s...	
5570	ham	The guy did some bitching but I acted like i'd...	
5571	ham	Rofl. Its true to its name	

5572 rows x 2 columns



DATA CLEANING

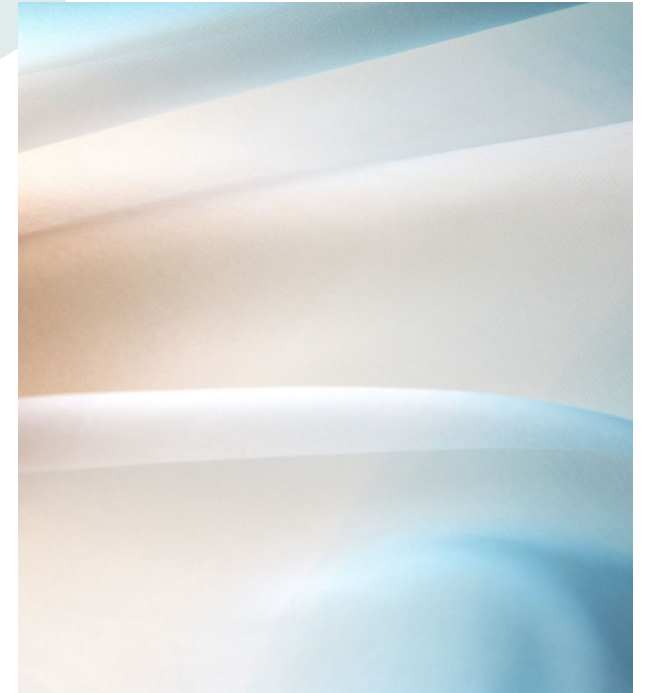
- *Removing irrelevant and null columns,*
- *Renaming features,*
- *Removing duplicated values,*
- *Removing Special Characters:* All non-textual characters such as punctuation marks and symbols.
- *Standardizing Case:* All text was converted to lowercase.
- *Converting Contractions:* Shortened versions of words or syllable were transformed into their fuller version.
- *Replacing URLs to 'URL', emails to 'EMAIL'.*

DATASET PREPROCESSING

- *Tokenization*: Braking down sentences.
- *Stemming*,
- *Lemmatization*,
- *Stop word analysis*
- *Creating processed text based on lemmatized words: Vectorizers and transformers require input as raw text strings rather than lists of tokens*

TECHNOLOGY STACK

- *Program Language*: Python 3.13
- *Jupyter Notebook*
- *Github*
- **Pandas** : Essential for data manipulation and cleaning.
- **Numpy**: scientific computing with Python.
- **NLTK**: Natural Language Toolkit.
- **Seaborn**: Python visualization library based on matplotlib.
- **Matplotlib**: 2D plotting library.
- **Sklearn**: Widely used for machine learning tasks including classification, regression, clustering, and dimensionality reduction.
- **Keras**: open-source neural network library written in Python.
- **Re**: For searching, matching, and splitting text based on patterns.
- **Contractions**: Used for converting contractions.



ARCHITECTURAL FLOW

1- Data Acquisition

2- Data Cleaning

3- Pre-processing

4- Feature Engineering

4.1- Stratify Splitting

4.2- N-grams

4.3- TF-IDF

4.4- BoW

5- Machine Learning Model Creations

5.1- Multinomial Naive Bayes for TF-IDF

5.1.1- CV

5.1.2- LOOCV (leave-one-out)

5.2- Multinomial Naive Bayes for BoW

5.2.1 -CV

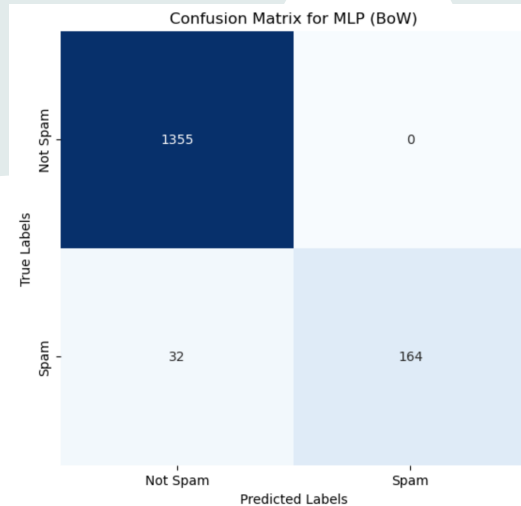
5.2.2- LOOCV

5.3- MLP for TF-IDF

5.4- MLP for Bow

6- Evaluation

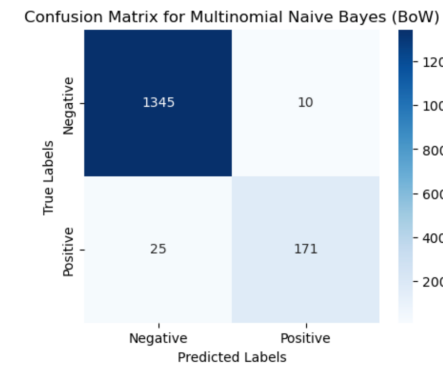
RESULTS AND OUTPUT



Precision: 1.0000
Recall: 0.8367
Accuracy: 0.9794
F1 Score: 0.9111

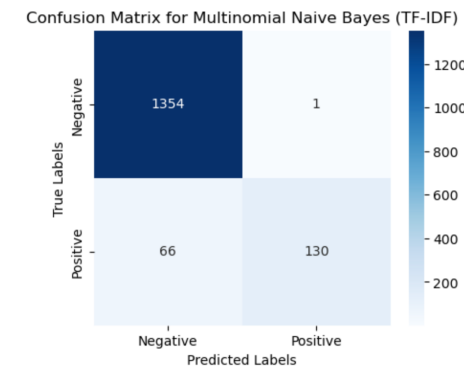
Accuracy: 0.9774339136041263

Classification Report:					
	precision	recall	f1-score	support	
0	0.98	0.99	0.99	1355	
1	0.94	0.87	0.91	196	
accuracy			0.98	1551	
macro avg	0.96	0.93	0.95	1551	
weighted avg	0.98	0.98	0.98	1551	

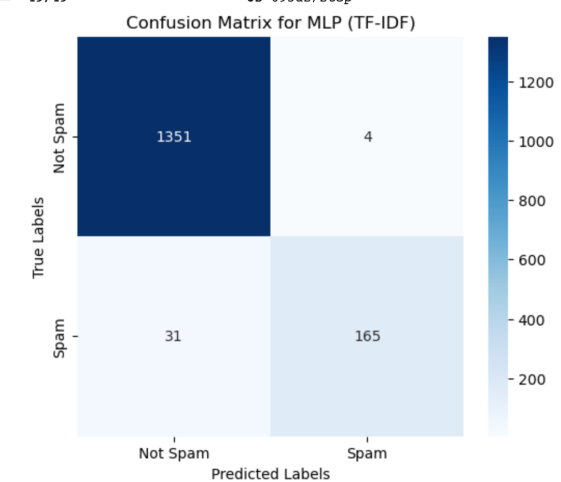


Accuracy: 0.9568020631850419

Classification Report:					
	precision	recall	f1-score	support	
0	0.95	1.00	0.98	1355	
1	0.99	0.66	0.80	196	
accuracy			0.96	1551	
macro avg	0.97	0.83	0.89	1551	
weighted avg	0.96	0.96	0.95	1551	



Precision: 0.9763
Recall: 0.8418
Accuracy: 0.9774
F1 Score: 0.9041



KEY LEARNINGS

- Data quality is paramount,
- Importance of Feature Engineering,
- Balancing Precision and Recall,
- Regular Updating and Retraining,
- NLP Techniques are Key,
- Evaluation Metrics are Crucial,

REFERENCES

- Dr. Akshi Kumar, lecture slides (2024)
- Dr. Akshi Kumar, lab activity notebooks (2024)

