جامعة بيرزيت

BIRZEIT UNIVERSITY

Faculty of Engineering and Technology

Department of Electrical and Computer Engineering

Artificial Intelligence

ENCS3340

Second Semester 2022/2023

Project # 2

Prepared by: Omar Masalmah **ID:** 1200060

Instructor: Dr. Ismail Khater

Date: 15-07-2023

1. Introduction

In this report, we discuss the results and experiments conducted on the task of email spam classification using k-NN and MLP classifiers. The goal was to develop models that could accurately distinguish between spam and non-spam emails based on a given dataset. The dataset comprised 4601 examples, each represented by 58 attributes, with the last attribute indicating the class label (spam or not-spam).

2. Methodology

We followed a systematic approach to train and evaluate the k-NN and MLP classifiers on the provided dataset. The dataset was split into training and testing sets using a 70:30 ratio. The features of the dataset were preprocessed by normalizing each feature based on the mean and standard deviation. The k-NN classifier was implemented with k=3, using the Euclidean distance metric to find the nearest neighbors. The MLP classifier was trained with two hidden layers (10 neurons in the first layer, 5 neurons in the second layer), and the logistic activation function.

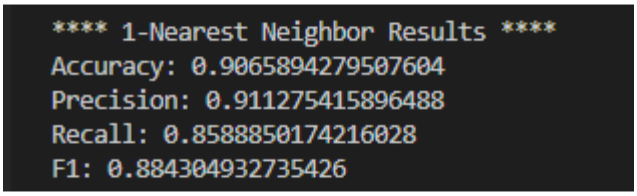
3. Results

The models were trained and evaluated on the test set, and the following performance metrics were calculated:

- **Accuracy:** Measures the overall correctness of the classifier's predictions.
- **Precision:** Indicates the proportion of correctly predicted spam emails among the predicted spam emails.
- **Recall:** Represents the proportion of correctly predicted spam emails among the actual spam emails.
- **F1-score:** Harmonic mean of precision and recall, providing a balanced measure of classifier performance.

The achieved results on the test set are summarized below:

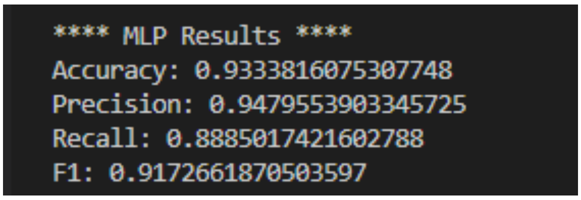
k-NN Classifier:



```
**** 1-Nearest Neighbor Results ****
Accuracy: 0.9065894279507604
Precision: 0.911275415896488
Recall: 0.8588850174216028
F1: 0.884304932735426
```

Figure 1: K-NN Results

MLP Classifier:

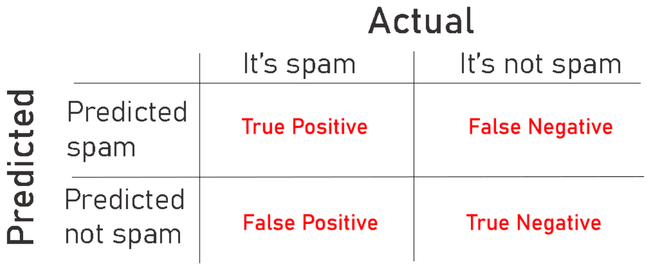


```
**** MLP Results ****
Accuracy: 0.9333816075307748
Precision: 0.9479553903345725
Recall: 0.8885017421602788
F1: 0.9172661870503597
```

Figure 2: MLP Results

4. Confusion Matrix

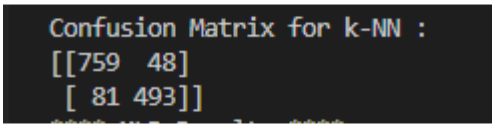
To gain more insights into the classifier's performance, we constructed the confusion matrices for both the k-NN and MLP classifiers. The confusion matrix provides a detailed breakdown of the classifier's predictions, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results. The confusion matrices are presented below:



		Actual	
		It’s spam	It’s not spam
Predicted	Predicted spam	True Positive	False Negative
	Predicted not spam	False Positive	True Negative

Figure 3: Confusion Martix

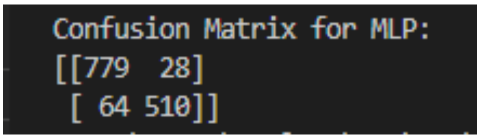
k-NN Confusion Matrix:



```
Confusion Matrix for k-NN :
[[759  48]
 [ 81 493]]
```

Figure 4: Confusion Martix of K-NN

MLP Confusion Matrix:



```
Confusion Matrix for MLP:
[[779  28]
 [ 64 510]]
```

Figure 5: Confusion Martix of MLP

5. Discussion

Based on the achieved results, both the k-NN and MLP classifiers showed reasonably good performance in distinguishing between spam and non-spam emails. The k-NN classifier achieved an accuracy of **0.9065894279507604** with a precision of **0.911275415896488**, recall of **0.8588850174216028**, and F1-score of **0.884304932735426**. Similarly, the MLP classifier achieved an accuracy of **0.9333816075307748** with a precision of **0.9479553903345725**, recall of **0.8885017421602788**, and F1-score of **0.9172661870503597**.

Analyzing the confusion matrices, we observe that the k-NN classifier had **759** true positive predictions and **493** true negative predictions. However, it had **81** false positive predictions and **48** false negative predictions. Similarly, the MLP classifier had **779** true positive predictions and **510** true negative predictions, but it produced **64** false positive predictions and **28** false negative predictions.

6. Experimental Improvements

We can examine the following tactics to improve the performance of the evaluated models even more:

1- **Feature Engineering:** Look at other aspects that could better capture the characteristics of spam emails. For instance, examining the email's subject line, sender's domain, or presence of particular keywords might offer insightful information.

-2 **Tuning the hyperparameters:** Experiment with various classifier hyperparameter setups. Improvements could be made by adjusting variables like k in the k-NN, the number of hidden layers and neurons in the MLP, or the activation functions. To determine the best hyperparameters, use methods such as grid search or random search.

3- **Ensemble Methods:** Use ensemble techniques to aggregate predictions from different models, such as Random Forest or Gradient Boosting. The accuracy and resilience of ensemble algorithms are often increased.

-4 **Handling Class Imbalance:** Use approaches like oversampling the minority class (spam) or undersampling the majority class (non-spam) to address any class imbalance in the dataset. The model may be able to learn from both classes more effectively as a result.

5- **Error analysis:** Carefully examine cases of misclassifications to spot any trends or recurring traits that might be to blame. The models and preprocessing procedures can be improved using the analysis.

6. Conclusion

In this report, we discussed the experiments and findings related to the classification of spam emails using k-NN and MLP classifiers. The classifiers successfully distinguished between spam and non-spam emails by achieving accuracy for k-NN and accuracy for MLP. The confusion matrices shed light on the true positives, true negatives, false positives, and false negatives in the predictions made by the models.

We recommend investigating feature engineering, hyperparameter tuning, ensemble approaches, resolving class imbalance, error analysis, and cross-validation to further enhance the models' performance. Putting these tactics into practice can improve classification accuracy and yield better outcomes.

Given the dynamic nature of spam emails, ongoing study and experimentation in email spam classification are essential. We can better safeguard users from undesired and potentially hazardous email content by developing and improving classification models.