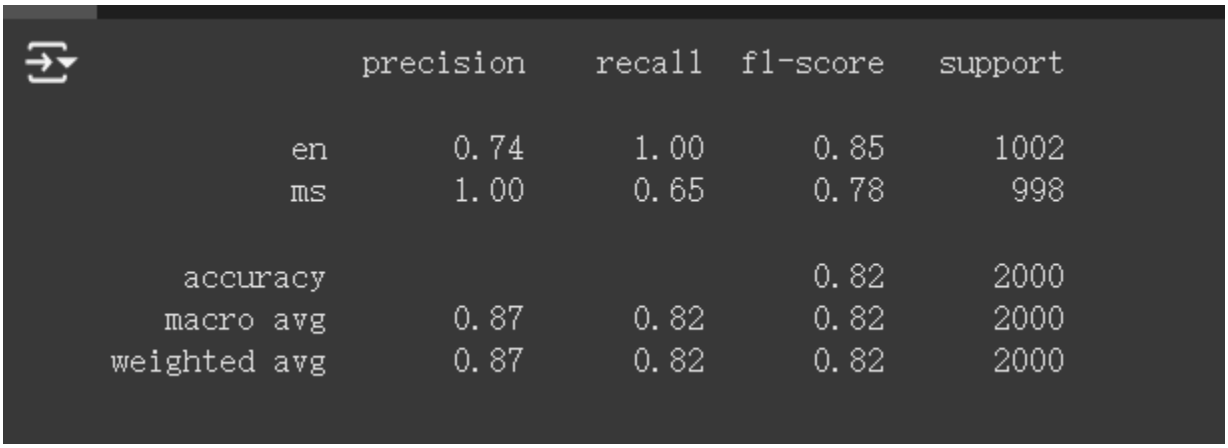
Classification Report



	precision	recall	f1-score	support
en	0.74	1.00	0.85	1002
ms	1.00	0.65	0.78	998
accuracy			0.82	2000
macro avg	0.87	0.82	0.82	2000
weighted avg	0.87	0.82	0.82	2000

Class: 'en' (English)

- **Precision (0.74)**: Out of all the samples predicted as 'en', 74% were actually 'en'. It measures how precise the predictions for this class are.
- **Recall (1.00)**: Out of all the actual 'en' samples, 100% were correctly predicted as 'en'. It measures the ability to capture all 'en' samples.
- **F1-Score (0.85)**: A harmonic mean of precision and recall, providing a balance between the two. Higher is better, and 0.85 indicates strong performance.
- **Support (1002)**: The number of true samples in the dataset belonging to the 'en' class.

Class: 'ms' (Malay)

- **Precision (1.00)**: Out of all the samples predicted as 'ms', 100% were actually 'ms'. This is perfect precision for 'ms'.
- **Recall (0.65)**: Out of all the actual 'ms' samples, only 65% were correctly predicted as 'ms'. This shows the model misses some 'ms' samples, leading to lower recall.
- **F1-Score (0.78)**: A balance between precision and recall. Though the precision is perfect, the lower recall brings down the F1-score.
- **Support (998)**: The number of true samples in the dataset belonging to the 'ms' class.

Accuracy (0.82):

- The overall percentage of correctly classified samples (both 'en' and 'ms').
- Here, 82% of the samples in the dataset were correctly classified.

Macro Avg (Precision: 0.87, Recall: 0.82, F1-Score: 0.82):

- These are the averages of precision, recall, and F1-score across both classes, calculated **without considering class imbalance**.
    - Precision: (0.74 + 1.00) / 2 = 0.87
    - Recall: (1.00 + 0.65) / 2 = 0.82
    - F1-Score: (0.85 + 0.78) / 2 = 0.82

Weighted Avg (Precision: 0.87, Recall: 0.82, F1-Score: 0.82):

- These are the averages of precision, recall, and F1-score, **weighted by the support (number of samples per class)**.
    - Weighted averages take into account that there are slightly more 'en' samples than 'ms' samples.