Results narrative:

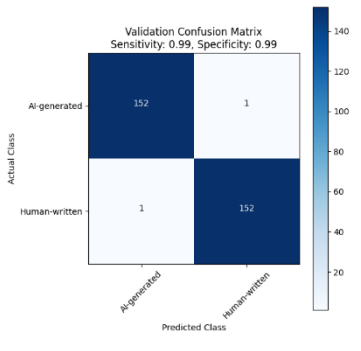
The following are the experiments running results in detail. For those interested in a deeper exploration of the data, the trained model weights derived from these experiments are available for download. These resources can be accessed via our dedicated GitHub repository at the following URL:

https://github.com/Hamed1Hamed/Arabic_AI_Detector/tree/main/Results.

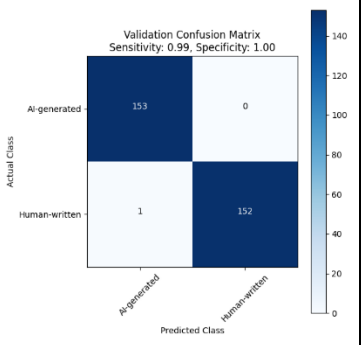
AraELECTRA-base-discriminator Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
1	Custom dataset			First Run on Dataset	We ran for 10 epochs. With enhancements on the learning rate by applying warmup phase. "learning_rate": 3.2e-05, "initial_learning_rate": 5e-6, Linear increase from initial to LR. Then applying cosine_annealing, where it smoothly decreases the learning rate according to a cosine function.

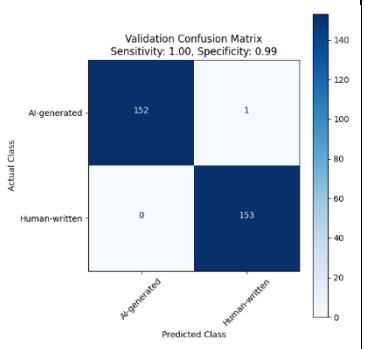
Confusion Matrix on Validation Set



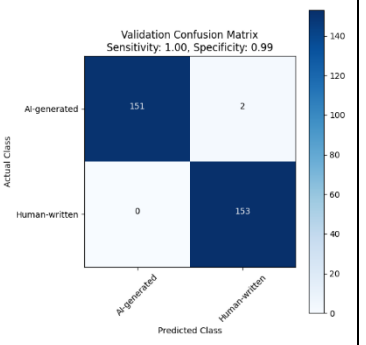
Epoch 1



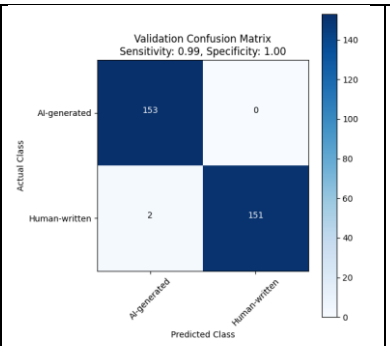
Epoch 2



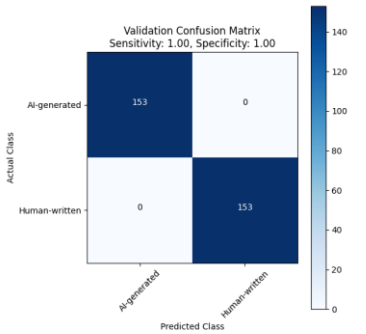
Epoch 3



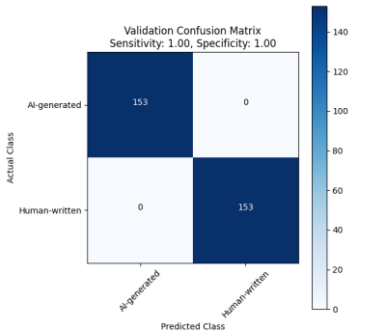
Epoch 4



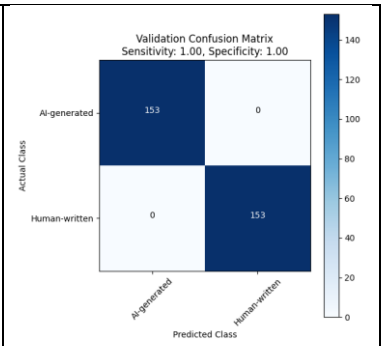
Epoch 5



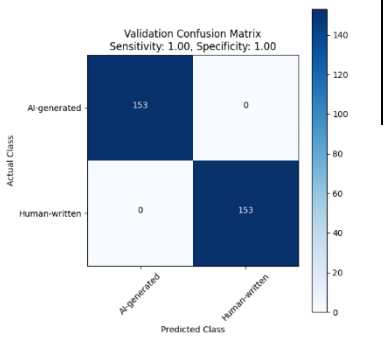
Epoch 6
Loss: 0.0002



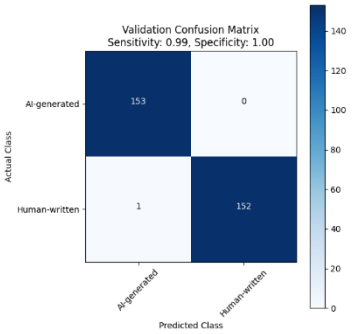
Epoch 7



Epoch 8

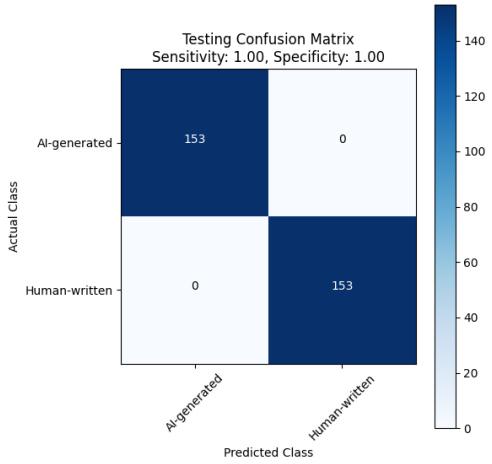


Epoch 9

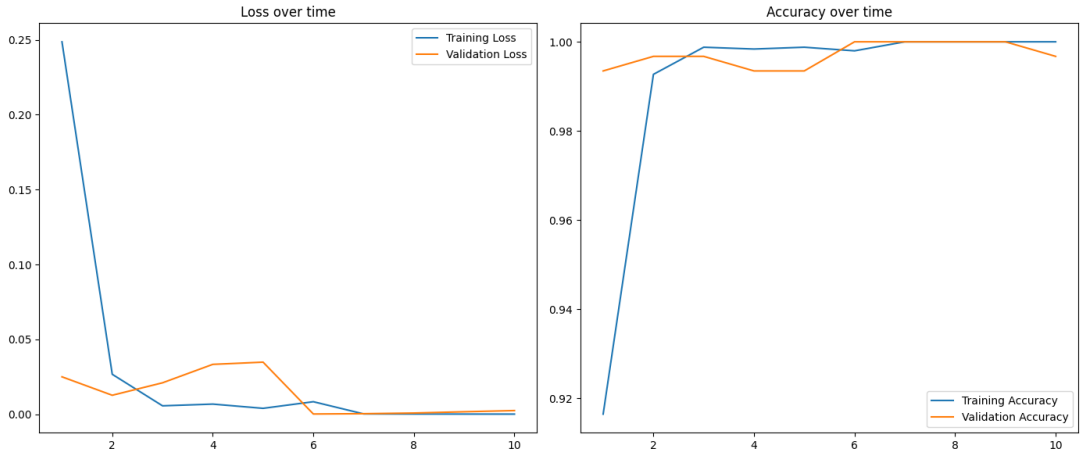


Epoch 10

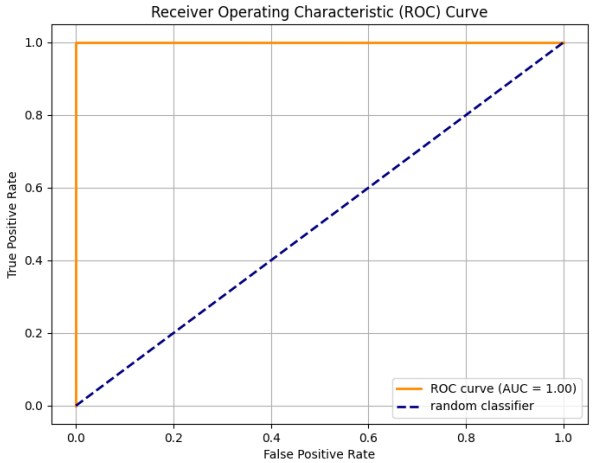
Confusion Matrix of Testing Set



Loss and Accuracy Over Time


Loss over time
Training Loss
Validation Loss
Accuracy over time
Training Accuracy
Validation Accuracy

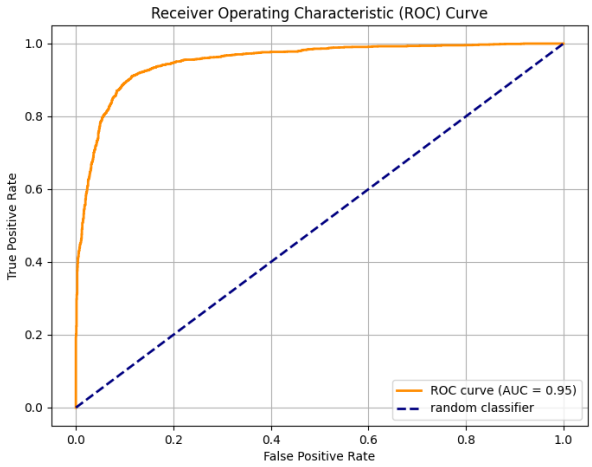
ROC-AUC


Receiver Operating Characteristic (ROC) Curve
True Positive Rate
False Positive Rate
ROC curve (AUC = 1.00)
random classifier

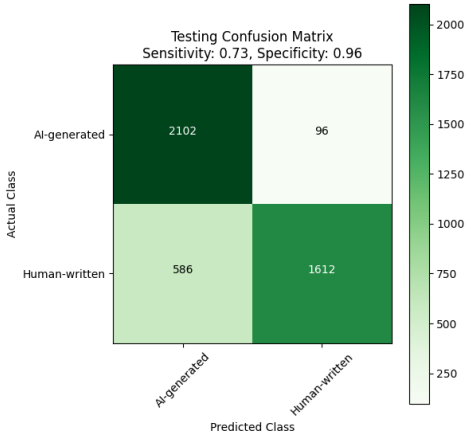
AraELECTRA-base-discriminator Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
2	-		Large dataset	Using Weights from Evaluation Index [1]	No training or validation. Evaluating directly using the best weights.

ROC-AUC

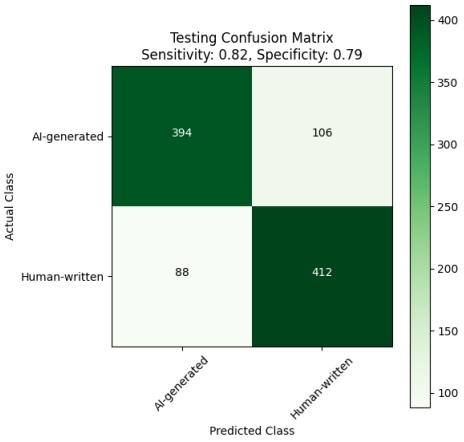
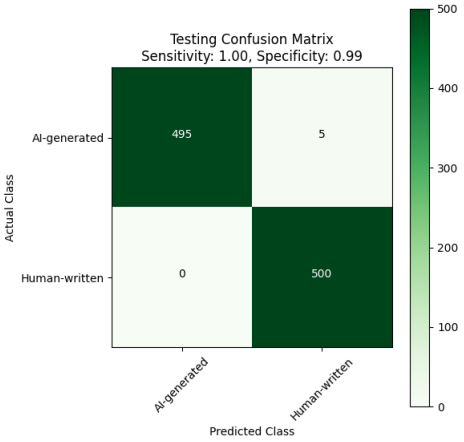
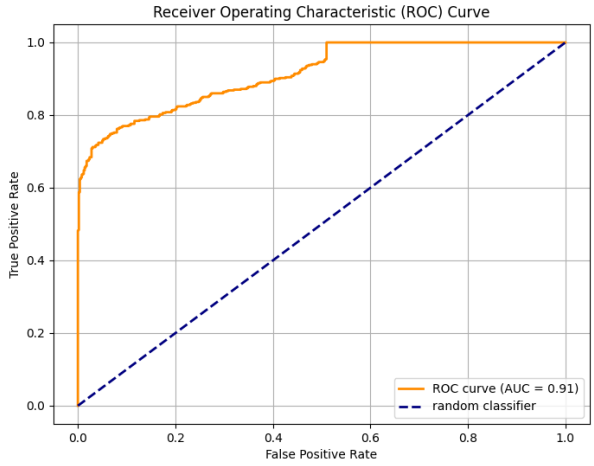
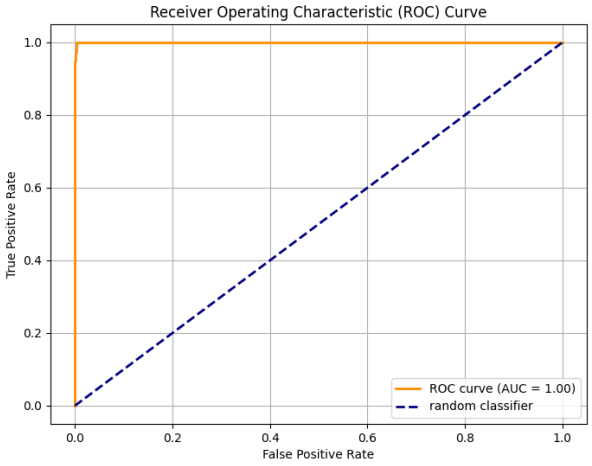


Confusion Matrix of Testing Set



AraELECTRA-base-discriminator Model

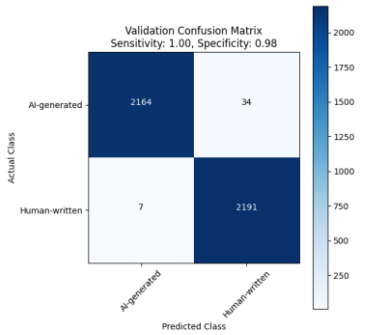
Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
3	-		AIRABIC	Using Weights from Evaluation Index [1]	No training or validation. Evaluating directly using the best weights.

Without Dediacritization Filter Layer	With Dediacritization Filter Layer
	
ROC-AUC	ROC-AUC
	

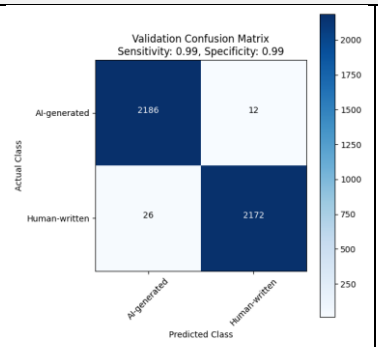
AraELECTRA-base-discriminator Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
4	Large dataset			First Run on Dataset	We ran for 10 epochs. With enhancements on the learning rate by applying warmup phase. "learning_rate": 3.2e-5, "initial_learning_rate": 5e-6, Linear increase from initial to LR. Then, cosine_annealing is applied, which smoothly decreases the learning rate according to a cosine function.

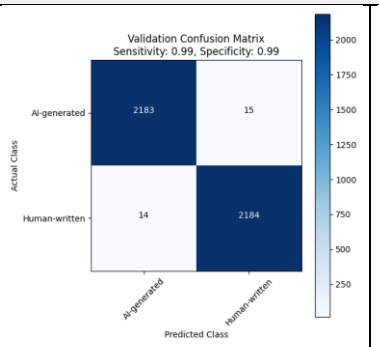
Confusion Matrix on Validation Set



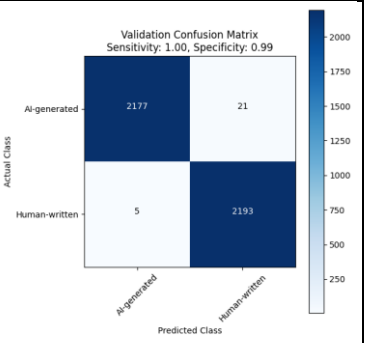
Epoch 1



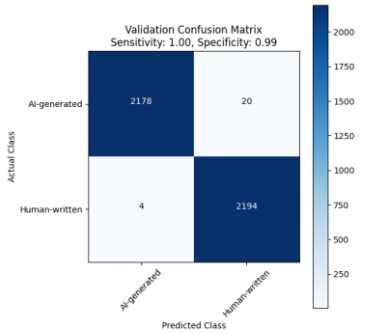
Epoch 2



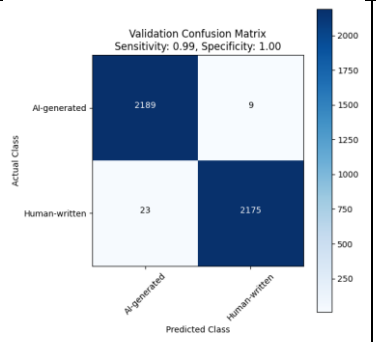
Epoch 3



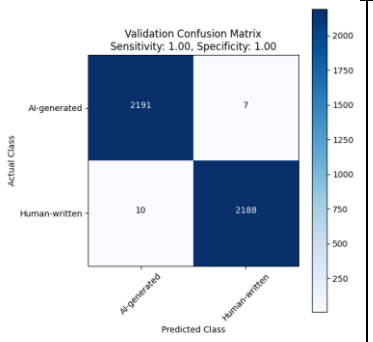
Epoch 4



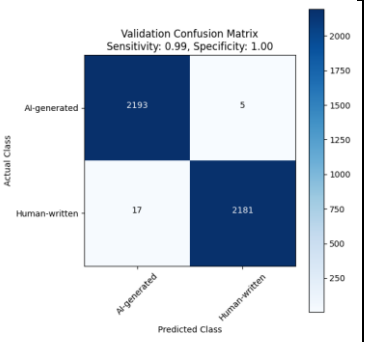
Epoch 5



Epoch 6

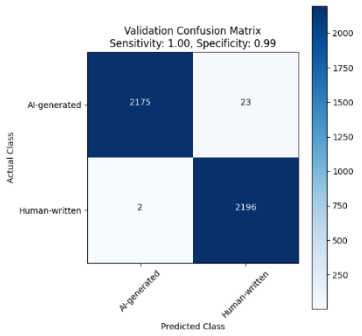


Epoch 7

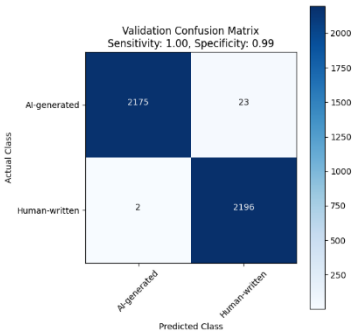


Epoch 8

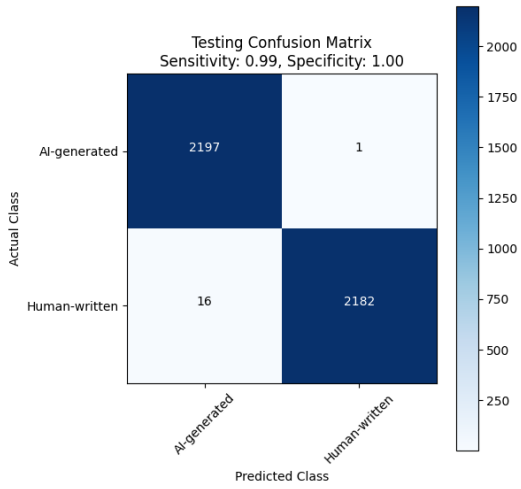
Confusion Matrix of Testing Set



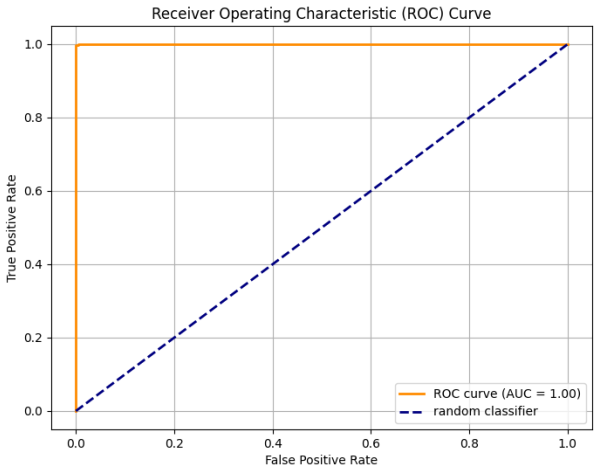
Epoch 9



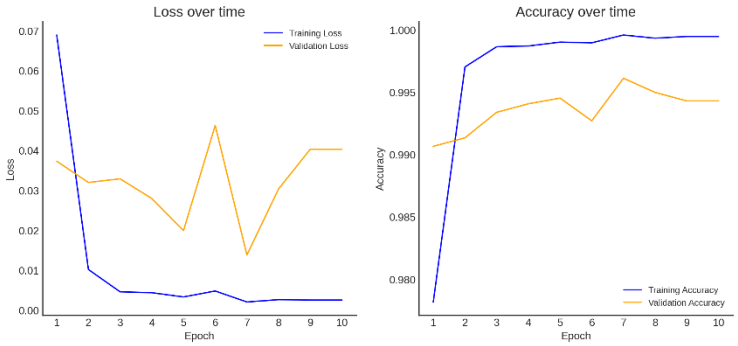
Epoch 10



ROC-AUC



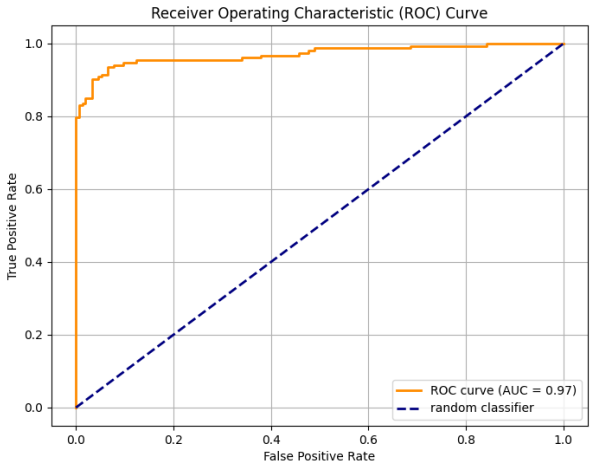
Loss and Accuracy Over Time

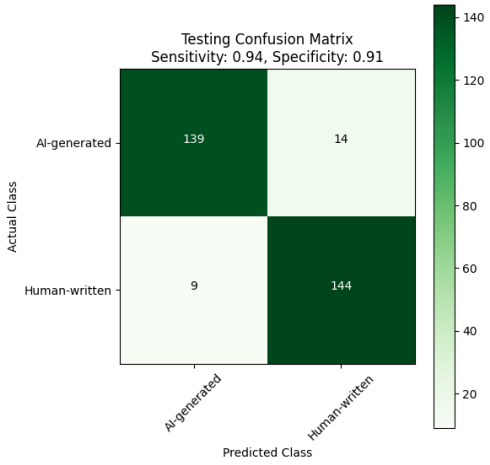


AraELECTRA-base-discriminator Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
5	-		Custom dataset	Using Weights from Evaluation Index [4]	No training or validation. Evaluating directly using the best weights.

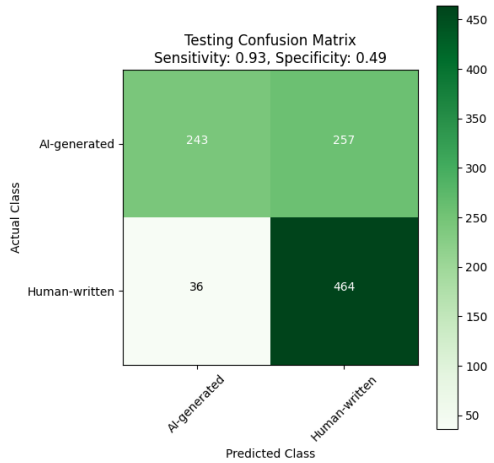
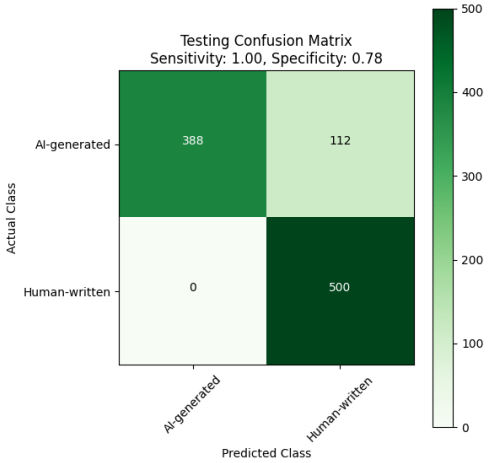
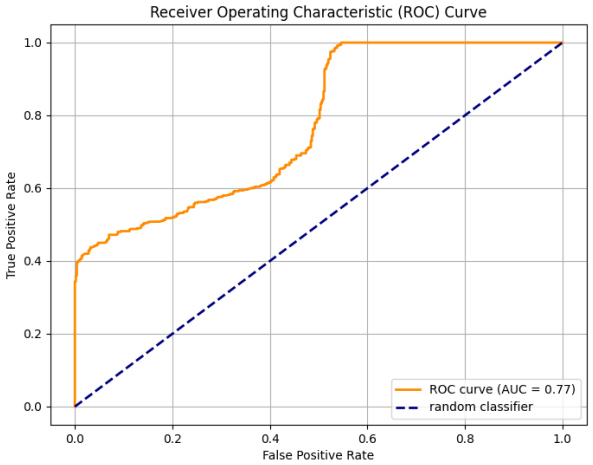
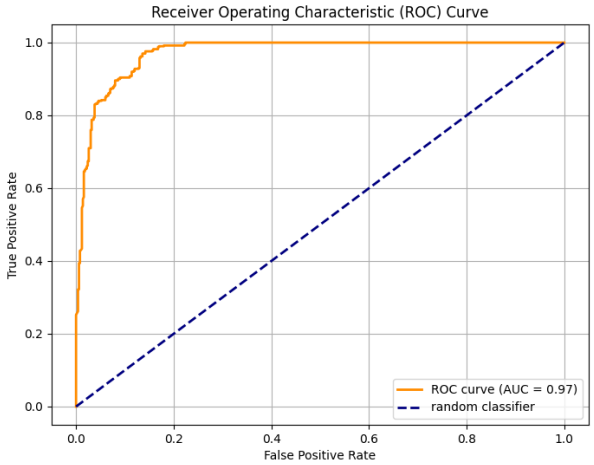
ROC-AUC	Confusion Matrix of Testing Set





AraELECTRA-base-discriminator Model

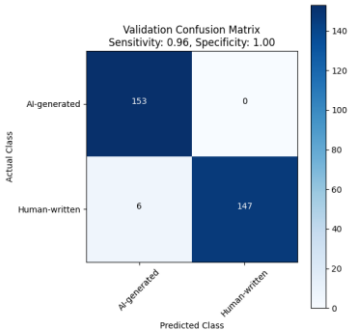
Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
6	-		AIRABIC	Using Weights from Evaluation Index [4]	No training or validation. Evaluating directly using the best weights.

Without Dediacritization Filter Layer	With Dediacritization Filter Layer
	
ROC-AUC	ROC-AUC
	

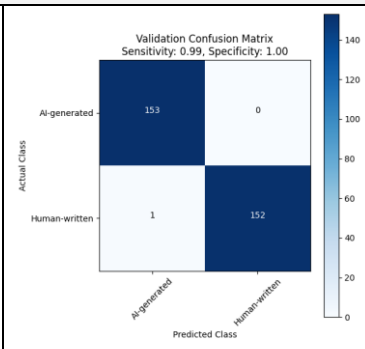
XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
7	Custom dataset			First Run on Dataset	We ran for 10 epochs. Batch size 32 with enhancements on the learning rate by applying warmup phase. "learning_rate": 3.2e-05, "initial_learning_rate": 5e-6, Linear increase from initial to LR. Then, apply cosine_annealing, which smoothly decreases the learning rate according to a cosine function.

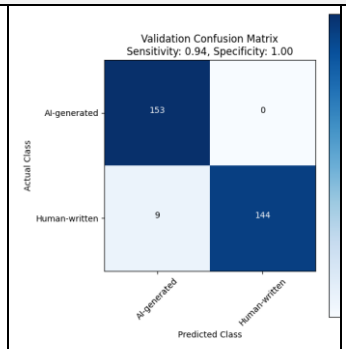
Confusion Matrix on Validation Set



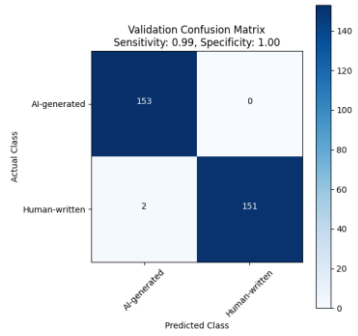
Epoch 1



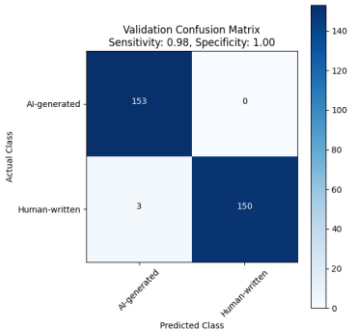
Epoch 2



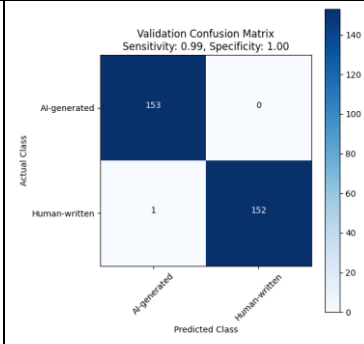
Epoch 3



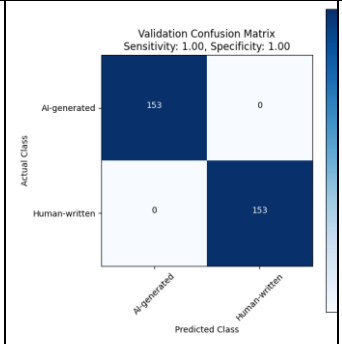
Epoch 4



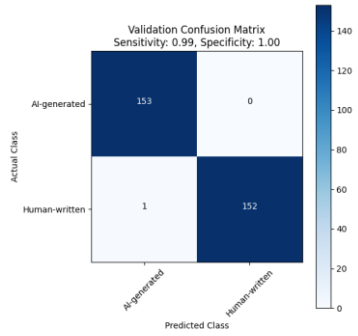
Epoch 5



Epoch 6

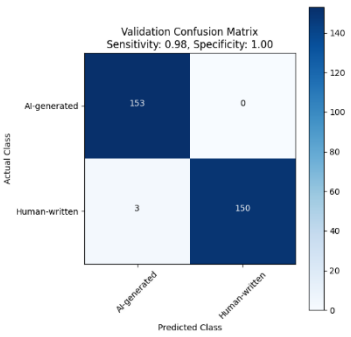


Epoch 7

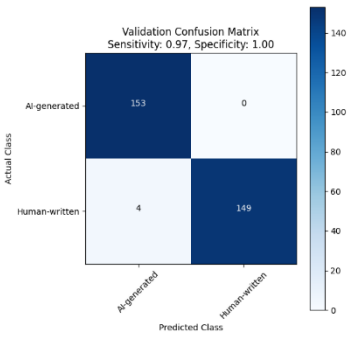


Epoch 8

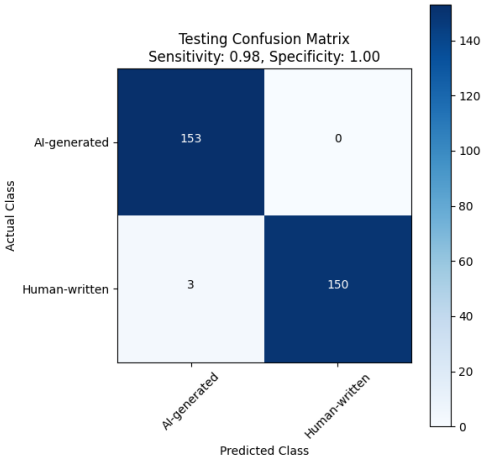
Confusion Matrix of Testing Set



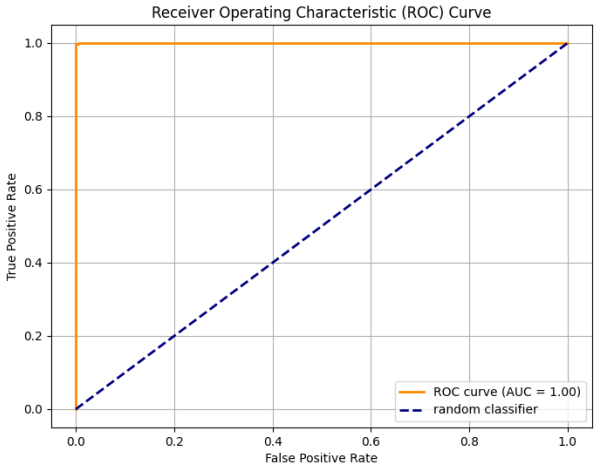
Epoch 9



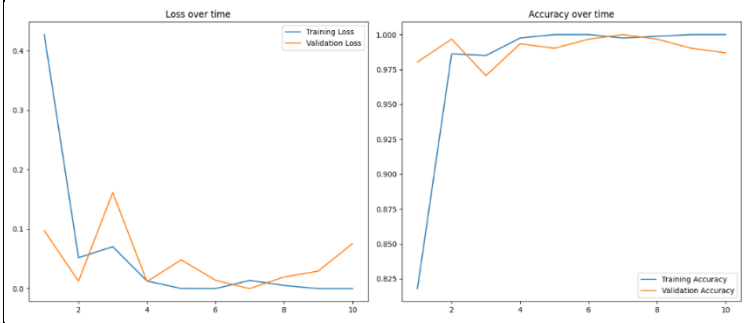
Epoch 10



ROC-AUC



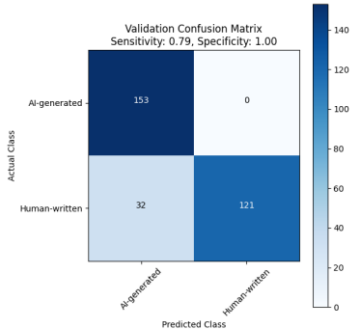
Loss and Accuracy Over Time



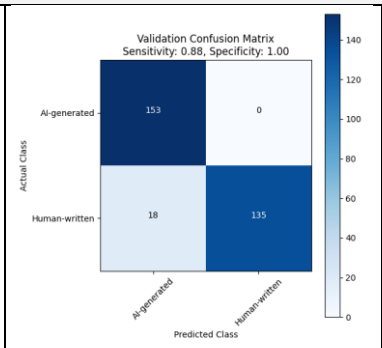
XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
7.2	Custom dataset			First Run on Dataset	We ran for 8 epochs. Batch size 32 with enhancements on the learning rate by applying warmup phase. "learning_rate": 3.2e-05, "initial_learning_rate": 1e-8, Linear increase from initial to LR. "warmup_epochs": 2, Then, apply cosine_annealing, which smoothly decreases the learning rate according to a cosine function.

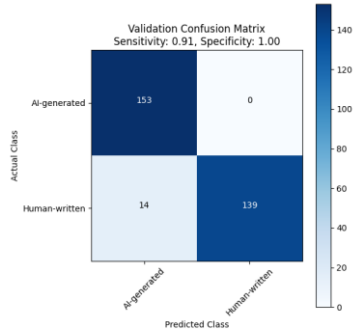
Confusion Matrix on Validation Set



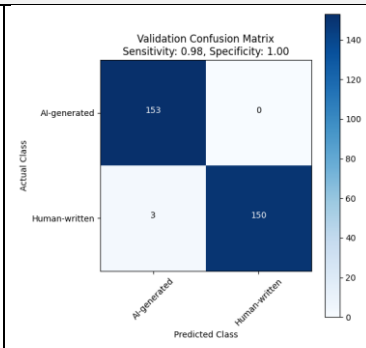
Epoch 1



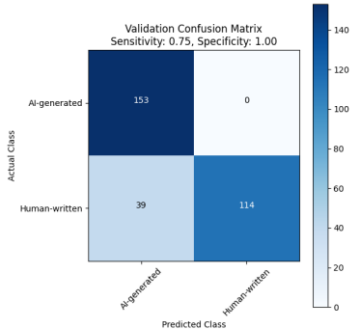
Epoch 2



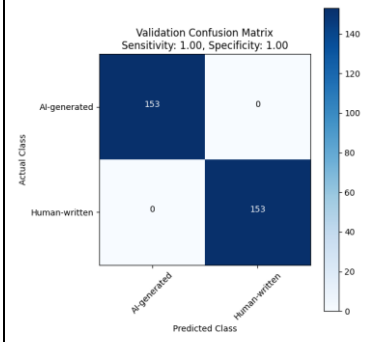
Epoch 3



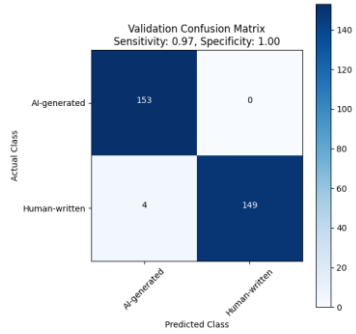
Epoch 4



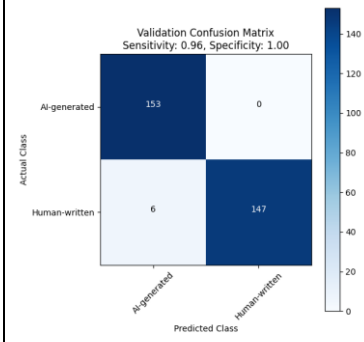
Epoch 5



Epoch 6
Loss: 0.0002

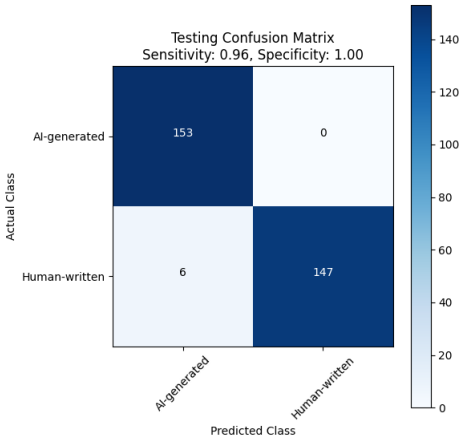


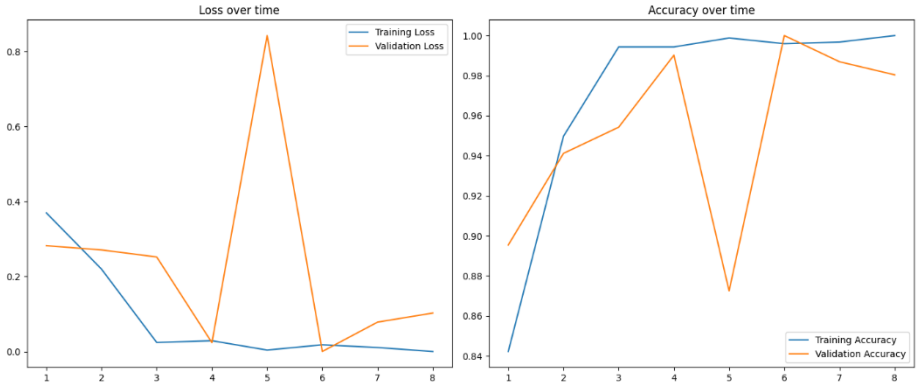
Epoch 7



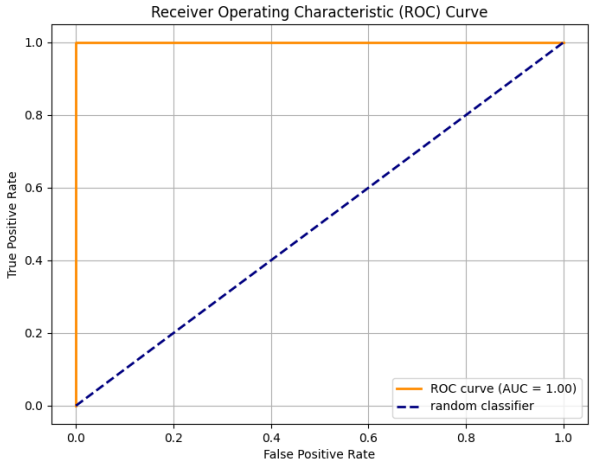
Epoch 8

Confusion Matrix of Testing Set & Loss and Accuracy Over Time


Testing Confusion Matrix
Sensitivity: 0.96, Specificity: 1.00
AI-generated
Human-written
153
0
6
147
Actual Class
Predicted Class
AI-generated
Human-written
140
120
100
80
60
40
20
0


Loss over time
Training Loss
Validation Loss
Accuracy over time
Training Accuracy
Validation Accuracy

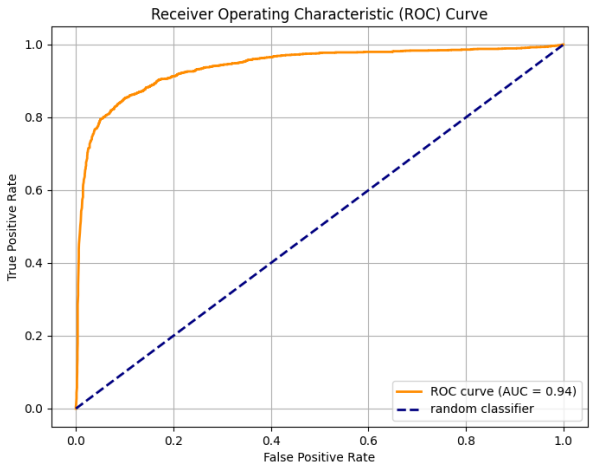
ROC-AUC

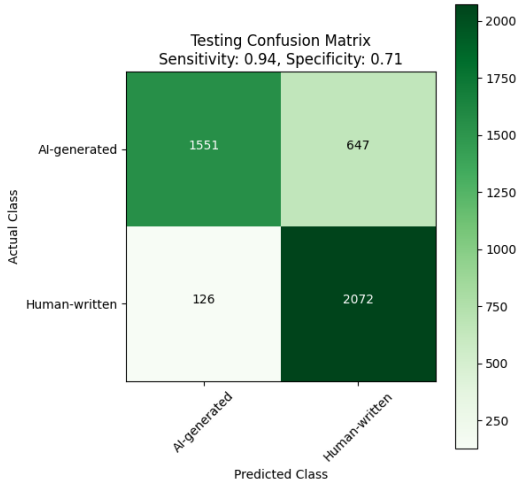

Receiver Operating Characteristic (ROC) Curve
True Positive Rate
False Positive Rate
ROC curve (AUC = 1.00)
random classifier

XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
8	-		Large dataset	Using Weights from Evaluation Index [7]	No training or validation. Evaluating directly using the best weights.

ROC-AUC	Confusion Matrix of Testing Set

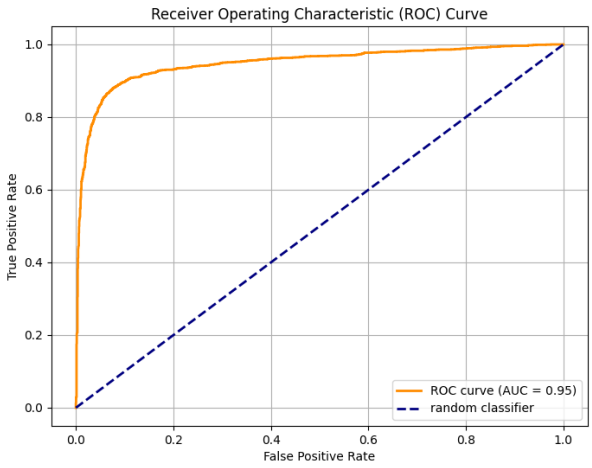


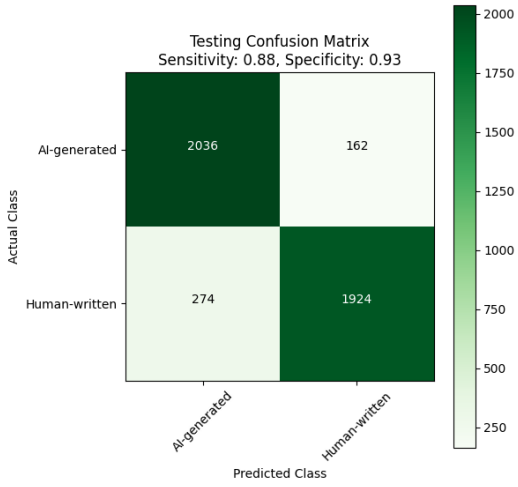


XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
8.2	-		Large dataset	Using Weights from Evaluation Index [7.2]	No training or validation. Evaluating directly using the best weights.

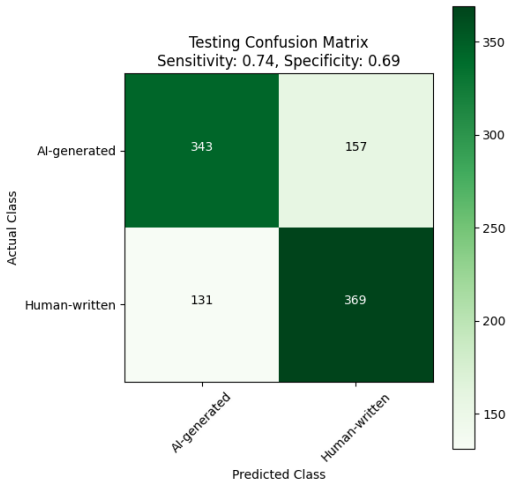
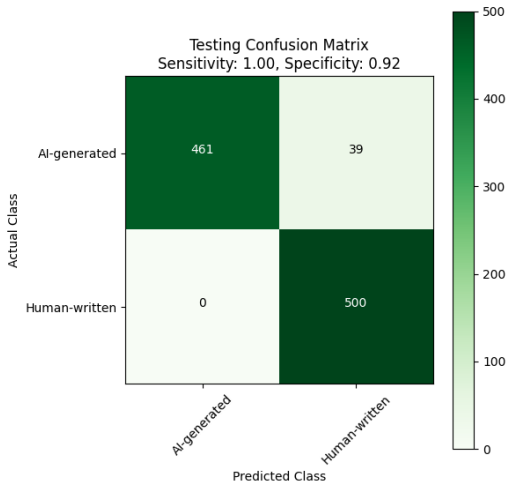
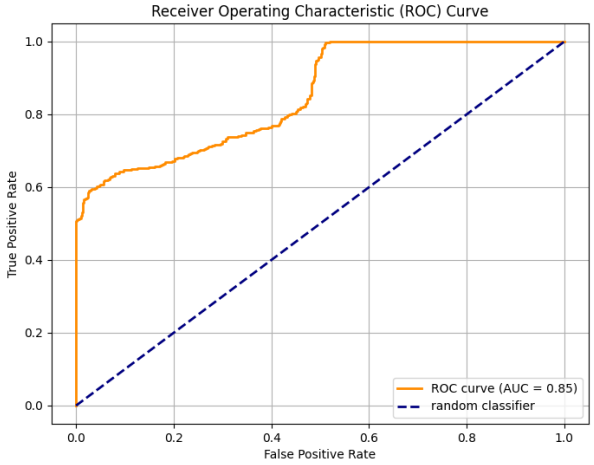
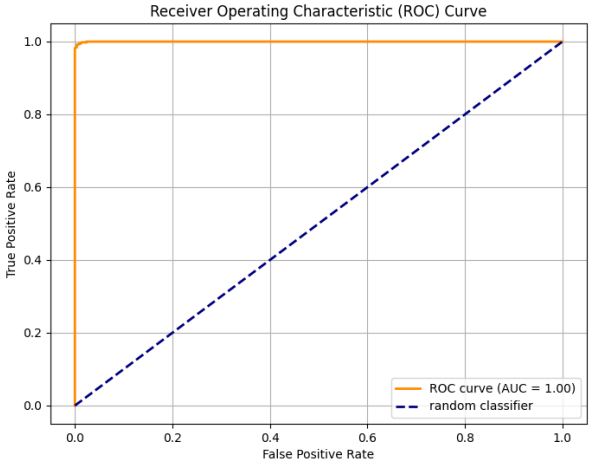
ROC-AUC	Confusion Matrix of Testing Set





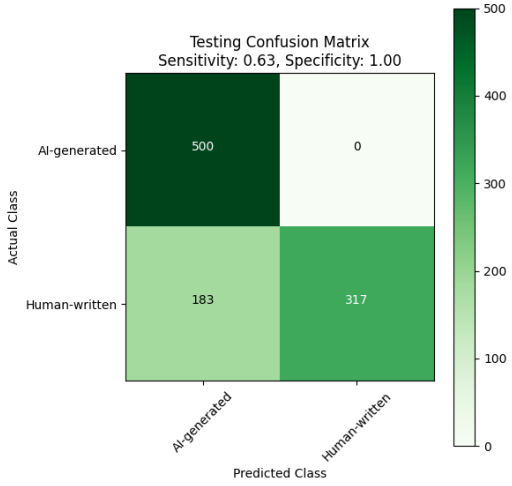
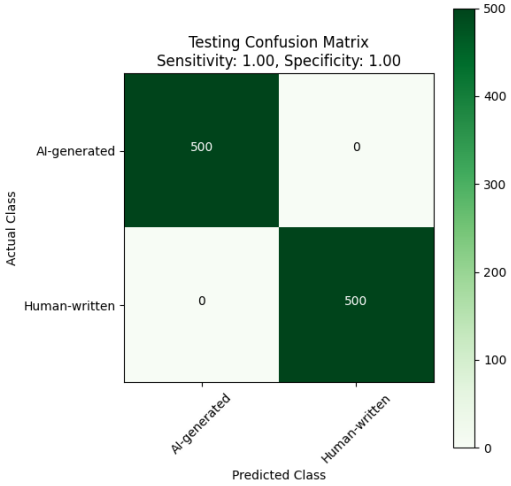
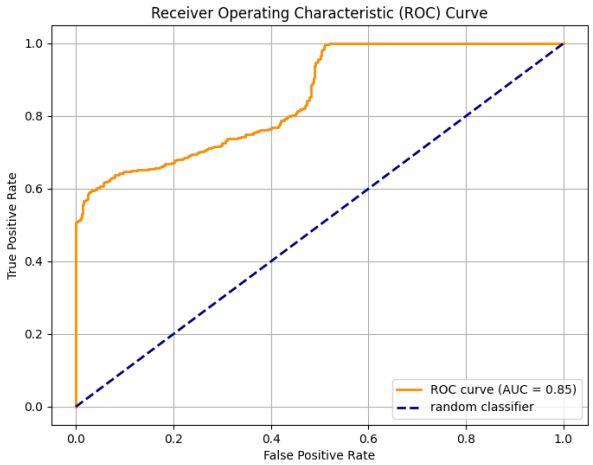
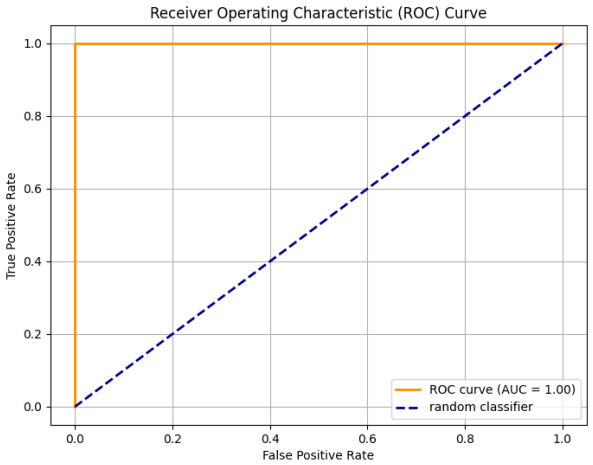
XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
9	-		AIRABIC	Using Weights from Evaluation Index [7]	No training or validation. Evaluating directly using the best weights.

Without Dediacritization Filter Layer	With Dediacritization Filter Layer
	
ROC-AUC	ROC-AUC
	

XLM-Roberta-base Model

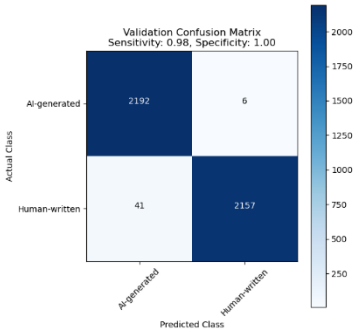
Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
9.2	-		AIRABIC	Using Weights from Evaluation Index [7.2]	No training or validation. Evaluating directly using the best weights.

Without Dediacritization Filter Layer	With Dediacritization Filter Layer
	
ROC-AUC	ROC-AUC
	

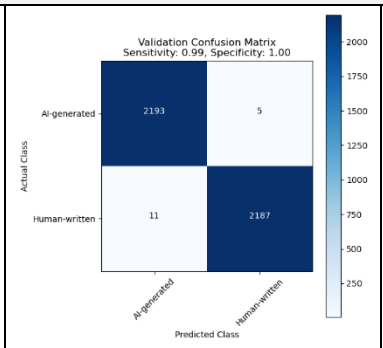
XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
10	Large dataset			First Run on Dataset	We ran for 10 epochs, batch size 64 with enhancements on the learning rate by applying warmup linear phase. "learning_rate": 3.2e-06, "initial_learning_rate": 5e-8, "warmup_epochs": 5, Then, apply cosine_annealing, which smoothly decreases the learning rate according to a cosine function.

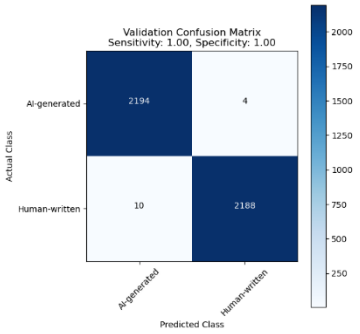
Confusion Matrix on Validation Set



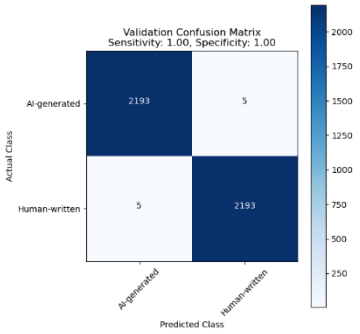
Epoch 1



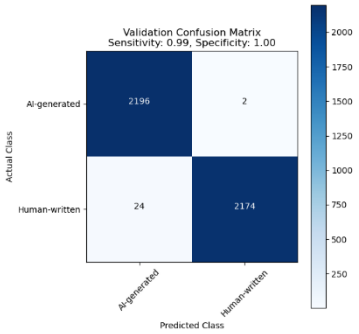
Epoch 2



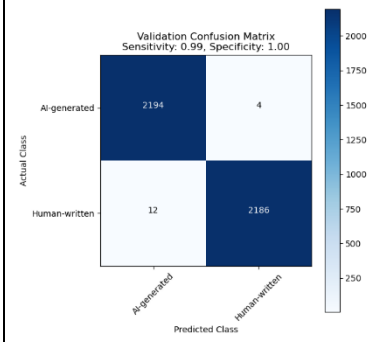
Epoch 3



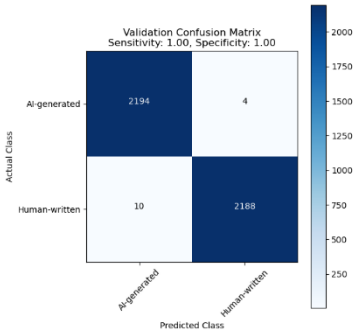
Epoch 4



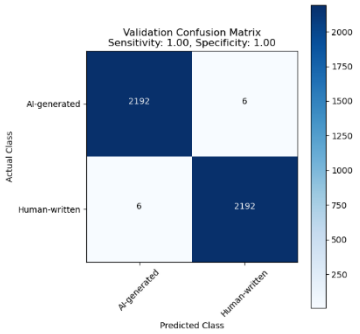
Epoch 5



Epoch 6



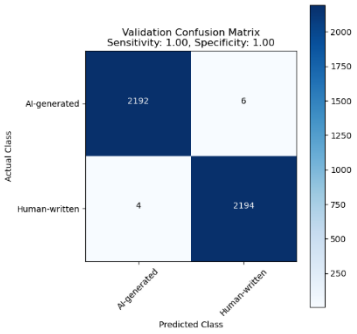
Epoch 7

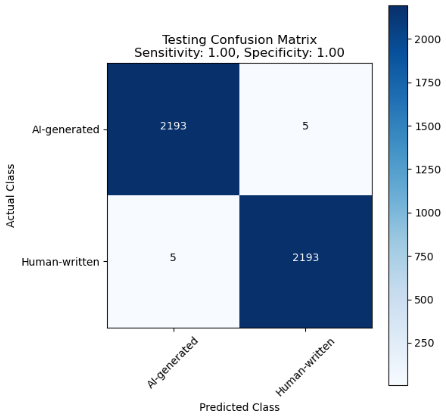


Epoch 8

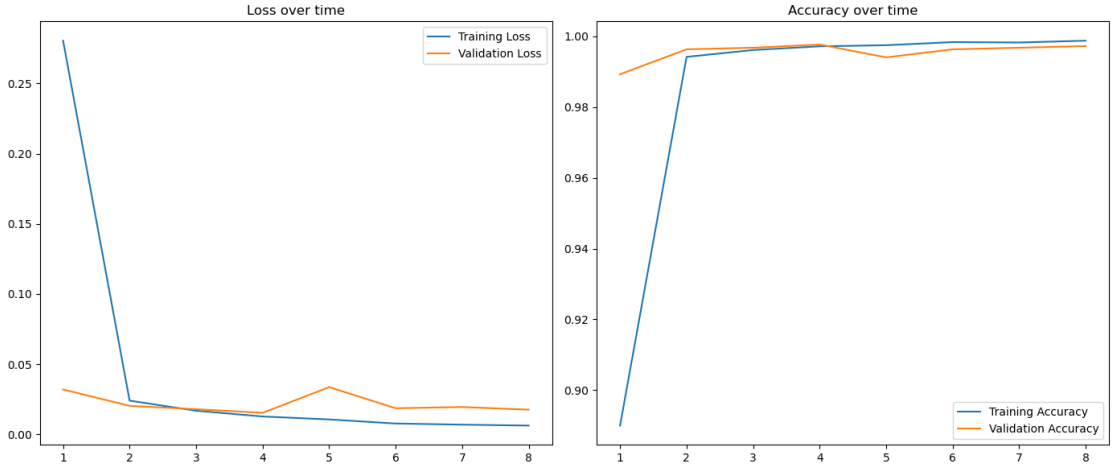
Confusion Matrix of Testing Set

Epoch 9

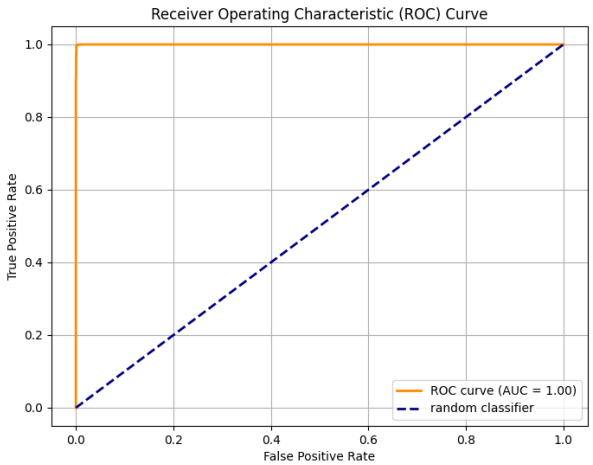




Loss and Accuracy Over Time



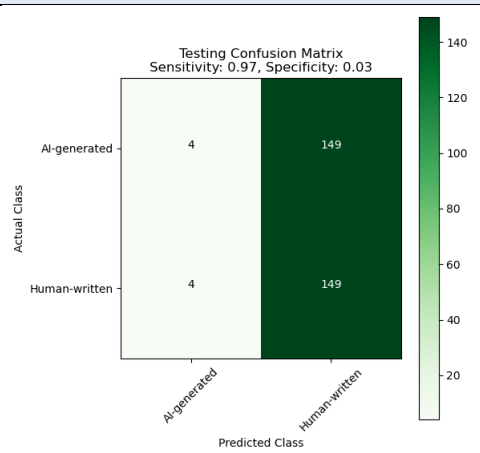
ROC-AUC

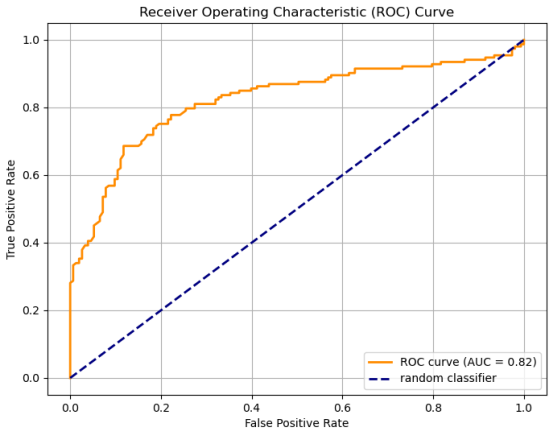


XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
11	-		Custom dataset	Using Weights from Evaluation Index [10]	No training or validation. Evaluating directly using the best weights.

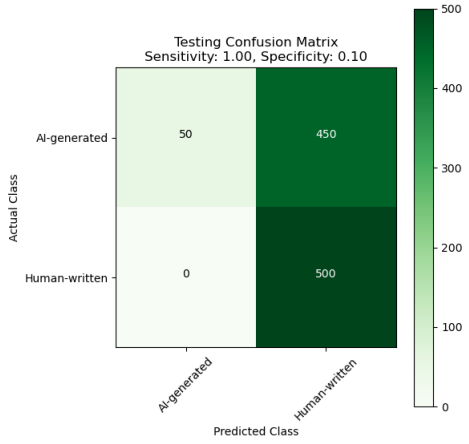
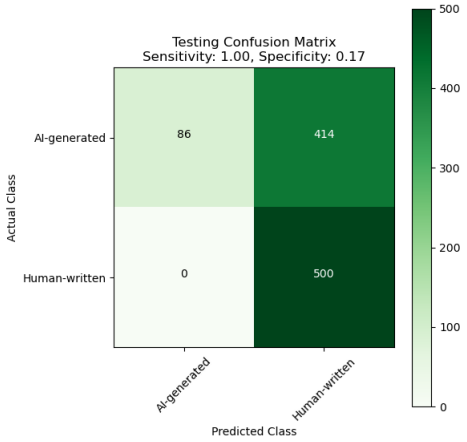
Confusion Matrix of Testing Set and ROC-AUC





XLM-Roberta-base Model

Evaluation Index	Trained	Validated	Tested	Experiment Weight Origin	Note
12	-		AIRABIC	Using Weights from Evaluation Index [10]	No training or validation. Evaluating directly using the best weights.

Without Dediacritization Filter Layer	With Dediacritization Filter Layer
	
ROC-AUC	ROC-AUC