

# Practice 2: Data Visualization and Performance Analysis

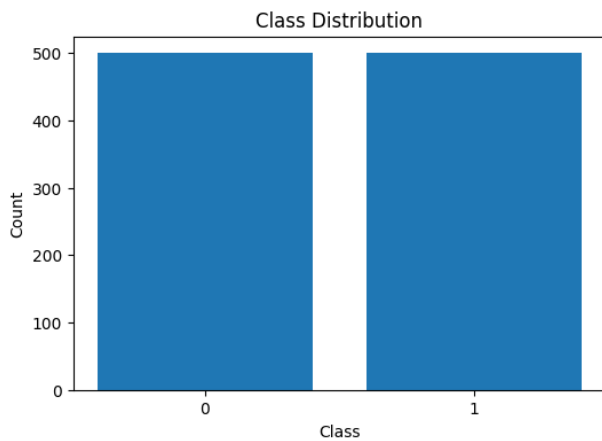
## Software Development Oriented to Machine Learning

Eneko Isturitz and Unai Lalana

### 1 Introduction

The aim of this practice is to explore the dataset and analyze the performance of the trained model through visualization. These visualizations help understand class distributions, sample examples, training performance, and model calibration.

### 2 Data Exploration



(a) Distribution of examples per class.



(b) Example images from each class.

Figure 1: Dataset exploration: class distribution and sample images.

### 3 Model Performance Analysis

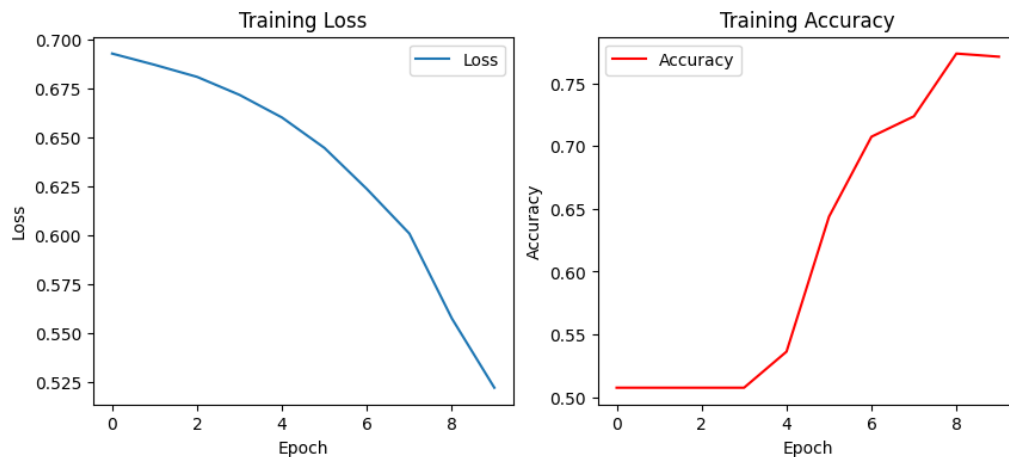


Figure 2: Training Loss and Accuracy over epochs.

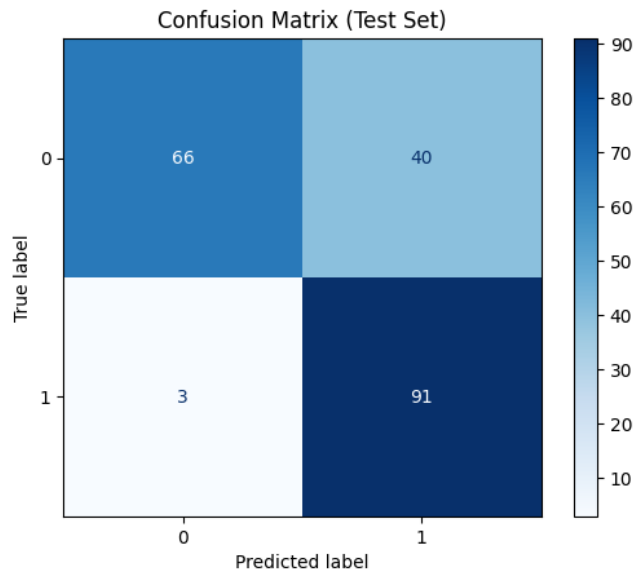


Figure 3: Confusion matrix showing model predictions vs. true labels.



Figure 4: Examples of images that the model classified with the highest error.

## 4 Model Calibration Curve

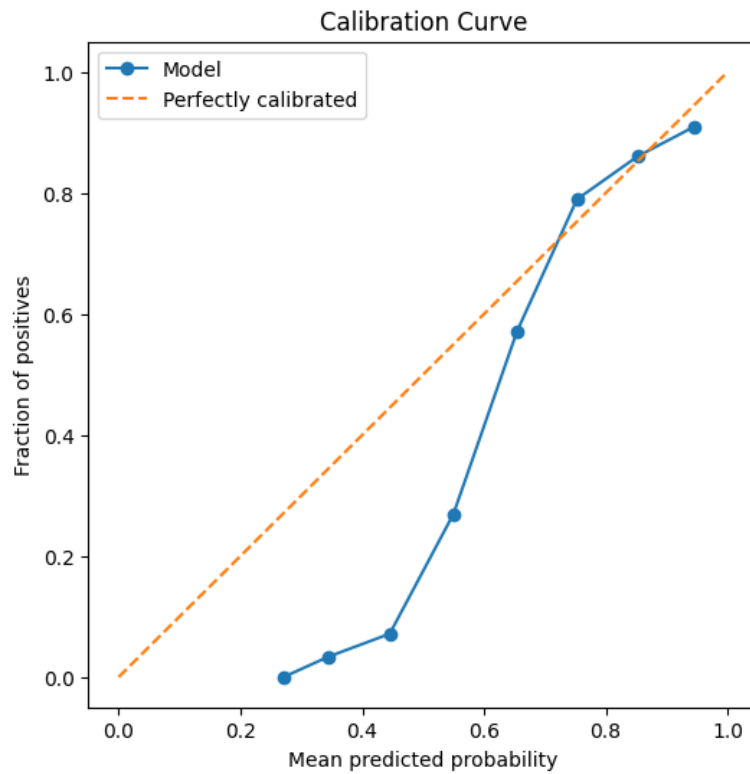


Figure 5: Calibration curve comparing predicted probabilities with actual outcomes.

## 5 Conclusions

We have a completely balanced subset of the dataset with classes 0 for real images and 1 for AI generated ones. On the other hand, our model achieves a solid performance even though it tends to predict more examples as class 1 (as seen in Figures 4 and 5)