Quantitative And Qualitative Comparison of Binary And Multiclass Classification Models

Both models were trained on the same dataset: COCO dataset.

Binary Classification

Model One (Tensorflow and ViT)

(https://huggingface.co/aspends/coco_binary_classification)

- Binary Classification for Cat (0) and Horse (1):
    - Prediction for Cat: 0.947;
    - Prediction for Horse: 0.053;

- Evaluation Test Results:
    - Training Loss: 0.0361;
    - Validation Loss: 0.1211;
    - Training Accuracy: 0.96;
    - Epochs: 4

Model Two (Custom made CNN)

- Binary Classification for Cat (1) and Horse (0):
    - Accuracy: 0.6870;
    - Recall: 0.6620;
    - Precision: 0.6968;
    - F1-Score: 0.6790

Quantitative Comparison

Model One has a higher accuracy on the evaluation test compared to Model Two, indicating that it makes correct predictions more often. However, Model Two shows a relatively reasonable F1-Score, presenting a balance between precision and recall, whilst Model One does not provide this information. Model 1 achieves lower training and validation loss compared to Model Two, which suggests that it has learned the task more effectively during training. Model Two has a recall of 0.6620, indicating that it correctly identifies 66,20& of the true positive cases. It also has a precision of 0.6968, suggesting that it has a good balance

between true positives and false positives.

Qualitative Comparison

Model One provides probability scores for each class. In such a case, it has high confidence in classifying a given image. This suggests that Model One shows confidence in its predictions.

Model Two provides a more comprehensive set of performance metrics, which is crucial for understanding its performance.

Conclusion

If the primary goal is to maximize predictions' accuracy and confidence then it is preferred to use Model One.

If it is necessary to acquire a more balanced evaluation considering both precision and recall, then it is wise to proceed with Model Two.

Multiclass Classification

Model One (Tensorflow and ViT)
(https://huggingface.co/aspends/coco_multiclass_classification)

- Multiclass Classification for Horse, Cat, Zebra, Train:
    - Prediction for Horse: 0.880;
    - Prediction for Cat: 0.041;
    - Prediction for Train: 0.040;
    - Prediction for Zebra: 0.039;

- Evaluation Results:
    - Training Loss: 0.0932;
    - Validation Loss: 0.2218;
    - Training Accuracy: 0.9313;
    - Epochs: 4

Model Two (Custom Made CNN)

- Multiclass Classification for Horse, Cat, and Zebra:
    - Accuracy for Horse: 0.4990;
    - Accuracy for Cat: 0.5010;
    - Accuracy for Zebra: 0.6670;
    - Recall for Horse: 0.3320;
    - Recall for Cat: 0.3340;
    - Recall for Zebra. nan;
    - Precision for Horse: 0.4985;
    - Precision for Cat: 0.5015;
    - Precision for Zebra: 0;
    - F1-Score for Horse: 0.3986;
    - F1-Score for Cat: 0.4010;
    - F1-Score for Zebra: nan

Quantitative Comparison

Model One has a high overall accuracy, suggesting that it performs well on average for all classes, while Model 2 has a lower overall accuracy. However, Model Two's evaluation results include detailed metrics for each class.

Qualitative Comparison

Model One does not provide specific class-wise metrics, making it harder to assess its performance based on each class. However, Model One provides probability scores, indicating its confidence in each class prediction. The "nan" results in Model Two may indicate issues with data or model training.

Conclusion

Model One should be preferred if high accuracy and training performance are the user's primary considerations. If a more balanced evaluation considering multiple classes is

essential, Model Two provides a comprehensive set of metrics for each class, and should be used instead.

In both Binary Classification and Multiclass Classification, Model Two offers class-specific metrics, while Model One provides only probability scores. The interpretation of metrics may be a subject for change due to the number of classes and the task's requirements. In BC, the focus is on the positive and negative classes, while in MC; the emphasis is on each class's individual performance.

The differences between the models in BC and ML are similar, but the impact and interpretation may vary based on the characteristics of binary and multiclass classification objectives. Model One's probability scores provide class confidence, while Model Two's metrics offer detailed insights into class-specific performance.