### Example assessment questions for model training and evaluation

Question

What is the total accuracy given the following confusion matrix? Please express your answer as a decimal rounded to 3 decimal places (For example, 68.25% would be 0.683).

|  | Predicted label | | | |
| --- | --- | --- | --- | --- |
| Actual  label |  | **A** | **B** | **C** |
| **A** | 183 | 5 | 7 |
| **B** | 0 | 199 | 11 |
| **C** | 10 | 23 | 162 |

Answer: 0.907

Explanation: Divide the total number of correct results (544) by the total number samples (600) to get an accuracy of 90.6667%, which we round to 0.907.

Question

What is the **total number of true positives** given the following confusion matrix? Your answer should be a positive, whole number (e.g. 123).

|  | Predicted label | | | |
| --- | --- | --- | --- | --- |
| Actual  label |  | **A** | **B** | **C** |
| **A** | 183 | 5 | 7 |
| **B** | 0 | 199 | 11 |
| **C** | 10 | 23 | 162 |

Answer: 544

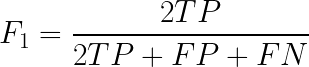
Explanation: Add up the cells on the diagonal of the matrix, where the actual label is the same as the predicted label (183 + 199 + 162 = 544).

Question

What is the **F1 Score of class C** given the following confusion matrix? Please express your answer as a decimal rounded to 3 decimal places (For example, 68.25% would be 0.683).

|  | Predicted label | | | |
| --- | --- | --- | --- | --- |
| Actual  label |  | **A** | **B** | **C** |
| **A** | 183 | 5 | 7 |
| **B** | 0 | 199 | 11 |
| **C** | 10 | 23 | 162 |

The F1 score is calculated by the following equation:



Where:

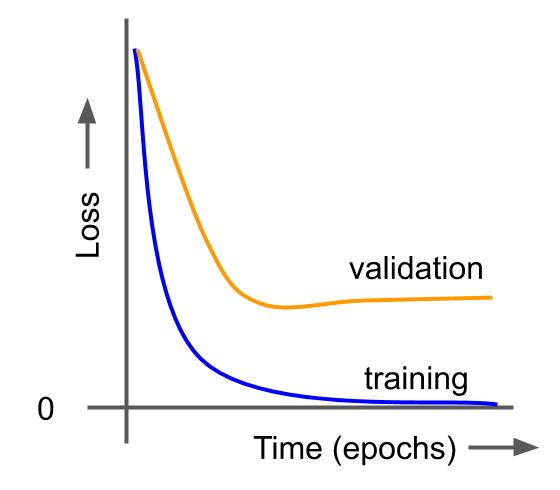
* True Positive (TP): Predicted positive matches actual positive
* True Negative (TN): Predicted negative matches actual negative
* False Positive (FP): Predicted positive does not match actual negative
* False Negative (FN): Predicted negative does not match actual positive

Answer: 0.864

Explanation: True positives (TP) for class C is 162, false positives (FP) for C is 18, and false negatives (FN) for C is 33. We plug those numbers into the F1 score formula to get 0.864.

Question

You plot the training and validation **loss** over the training period of the model, and you see the following. What can we conclude about the model?



1. **The model has overfit to the training data**
2. The model has a good fit with the training data
3. The model has underfit to the training data
4. The model’s input dimensions do not match the data’s dimensions

Question

Which of the following are good ways to try to reduce overfitting? (Select all that apply)

1. **Gather more data**
2. **Early stopping**
3. **Add regularization terms**
4. Use a more complex model