

Interpretation Questions

Henrique Oliveira Baptista

1.

- A melhor distribuição entre bias e variância está na imagem C. A imagem A representa um modelo de regressão linear que sofre de underfitting, pois é muito simples e não captura a complexidade dos dados. Já a imagem B é um exemplo de overfitting, onde o modelo "aprende demais" sobre os dados de treino, mas não generaliza bem para novos dados. A imagem C é a mais equilibrada, pois a curva segue o padrão dos dados de forma razoável, sem exageros. Não é um modelo perfeito, mas reflete melhor a realidade, com boa generalização e resultados consistentes.
- The best balance between bias and variance is shown in image C. Image A represents a linear regression model suffering from underfitting, as it is too simple and fails to capture the complexity of the data. Image B is an example of overfitting, where the model "learns too much" about the training data but does not generalize well to new data. Image C is the most balanced, as the curve reasonably follows the data pattern without exaggeration. It is not a perfect model but reflects reality better, offering good generalization and consistent results.

2.

- O propósito do gráfico é mostrar a relação entre a Taxa de Verdadeiros Positivos (True Positive Rate) e a Taxa de Falsos Positivos (False Positive Rate), sendo conhecido como uma Curva ROC (Receiver Operating Characteristic).
- A linha tracejada no meio representa um modelo que faz previsões aleatórias, ou seja, sem capacidade de distinguir entre classes
- A curva que representa o "better fit" é a linha verde, pois indica um modelo com maior desempenho, já que tem maior área sob a curva (AUC), capturando melhor os acertos. Já a curva vermelha reflete um modelo com desempenho inferior, mais próximo de erros e classificações incorretas.

- A Curva ROC é um gráfico que nos permite entender melhor o desempenho do modelo, avaliando se ele consegue distinguir com eficiência a complexidade dos dados ou se apresenta resultados equivocados, como erros ou falsos positivos
- The purpose of the graph is to show the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR), commonly known as the ROC Curve (Receiver Operating Characteristic).
- The dashed line in the middle represents a model making random predictions, meaning it cannot distinguish between classes.
- The curve that represents the "better fit" is the green line, as it indicates a model with higher performance due to a greater Area Under the Curve (AUC), capturing more correct predictions. The red curve reflects a lower-performing model, closer to errors and incorrect classifications.
- The ROC Curve helps us better understand the model's performance, evaluating its ability to effectively distinguish data complexity or if it exhibits errors like false positives.

3.

- Eu não diria que o modelo tem um bom desempenho olhando para a avaliação do teste, porque a matriz de confusão mostra que ele não consegue acertar mais do que 50%. Mesmo com três classes, a performance é baixa.
- O fenômeno que aconteceu foi que, apesar dos gráficos de erro e acurácia indicarem progresso ao longo das épocas (com o erro diminuindo e a acurácia aumentando), o modelo ainda tem dificuldade em classificar corretamente
- Olhando para os gráficos sem levar em conta outras informações, é possível observar que o modelo apresenta uma boa progressão ao longo das épocas. Conforme a quantidade de épocas aumenta, a taxa de erro diminui de forma significativa, enquanto a acurácia melhora de maneira consistente. Isso

demonstra que, durante o treinamento, o modelo está aprendendo e se ajustando melhor aos dados de entrada.

No entanto, ao analisar a matriz de confusão da avaliação de teste, fica evidente que o modelo ainda enfrenta muitas dificuldades em classificar corretamente. Sua melhor pontuação foi de 0.5 para duas classes e 0.45 para a classe B, o que indica um desempenho bastante limitado. Isso reflete uma alta taxa de falsos positivos (FP) e falsos negativos (FN), além de erros recorrentes ou "alucinações" nas previsões. Dessa forma, apesar da aparente evolução mostrada nos gráficos, o modelo não pode ser considerado bom, pois não consegue generalizar bem para os dados de teste.

- I would not say the model performs well based on the test evaluation, as the confusion matrix shows it cannot achieve more than 50% accuracy. Even with three classes, its performance is low.
- What happened is that, although the error and accuracy graphs indicate progress over epochs (with error decreasing and accuracy increasing), the model still struggles to classify correctly.
- Looking at the graphs alone, it appears the model demonstrates good progression over the epochs. As the number of epochs increases, the error rate decreases significantly, and accuracy improves consistently. This suggests the model is learning and adjusting to the input data during training.

However, when analyzing the confusion matrix from the test evaluation, it becomes clear the model still struggles significantly to classify correctly. Its highest score was 0.5 for two classes and 0.45 for class B, indicating very limited performance. This reflects a high rate of false positives (FP) and false negatives (FN), as well as recurring errors or "hallucinations" in predictions. Therefore, despite the apparent progress shown in the graphs, the model cannot be considered good because it fails to generalize well to test data.