ML Junior Practical Test - Interpretation

1)

- I think distribution C has the best balance between bias and variance.
- My thoughts: Distribution A shows a linear fit (straight red line) that doesn't capture the slight curve in the data - this looks like high bias/underfitting. Distribution B has a super wiggly line that perfectly hits every data point but would probably perform terribly on new data - classic high variance/overfitting. Distribution C has a curved line that follows the general trend without being too sensitive to individual points - it captures the pattern without overfitting, showing good balance between bias and variance.

2)

- This graph is called a ROC curve (Receiver Operating Characteristic curve). It plots the True Positive Rate against the False Positive Rate and helps evaluate classification model performance.
- The dashed blue line represents a random classifier (50/50 chance) - it's what you'd get if your model was just guessing.
- The green curve represents a better fit compared to the red. The green curve reaches higher into the top-left corner, which means it achieves higher true positive rates with lower false positive rates. We want curves that maximize area under the curve (AUC), and the green one definitely has more area under it than the red one.
- My thoughts: When comparing ROC curves, we want ones that climb quickly toward the top-left corner. The green model does this - it gets high true positive rates even when keeping false positives low. The red model struggles until it allows many false positives, which isn't good for a classifier.

3)

- I don't think we can say the model has good performance in the test evaluation.
- The phenomenon that happened during test evaluation appears to be poor generalization, specifically the model is showing signs of overfitting. Looking at the confusion matrix, the model seems to misclassify many instances. For example, it correctly classifies only 50% of class C examples, while 40% of them are misclassified as class B.
- My thoughts: While the training graphs look promising (increasing accuracy and decreasing error over epochs), the confusion matrix tells a different story. The diagonal values (0.5, 0.45, 0.5) should ideally be much higher for a good model. This suggests the model learned the training data well but isn't generalizing properly to new data - a classic sign of overfitting. The model particularly struggles with distinguishing between classes B and C, where there's a lot of confusion between the two.