

Architecture:

Spectrogram:

- Short time FFT takes with window length 445 and overlap 393
- 3x3 Convolutional layers(2D) were applied with max pooling to create abstract features
- Batch Normalization, ensures mean of 0 for consistent layer input
- DNN for classification based on VGG-16

Wave Form:

- Convolutional block(1D) applied to 8000x1 inputs
- Abstract features were input to an RNN network using LSTM
 - LSTMs allowed for long-term memory
- Leverage the sequential data to use the previous context to update parameters with memory
- Stacked for abstract features and longer memory
- Needed only 158094 parameters

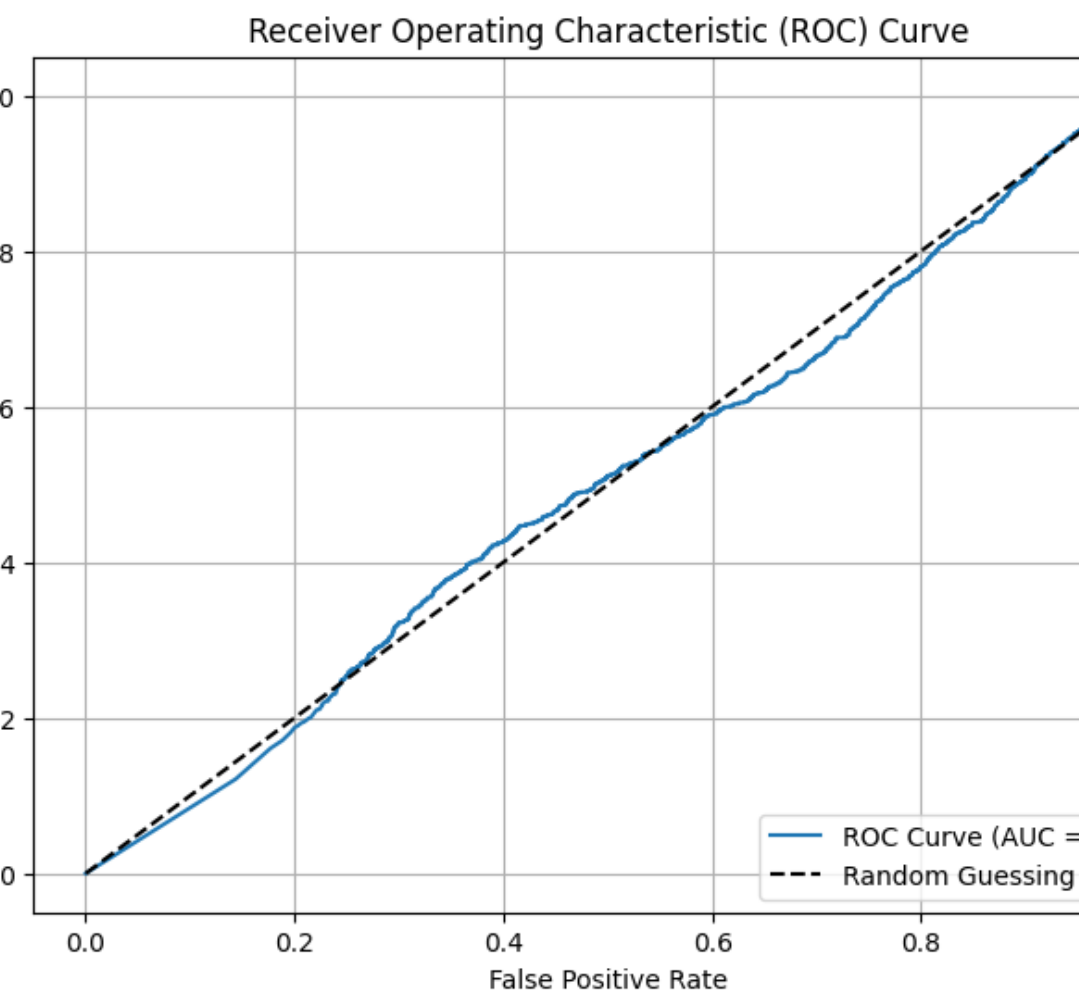
Training:

Optimizer: Adam, with an initial learning rate of (0.0005), allowed for fast convergence with an adaptive learning rate and momentum.

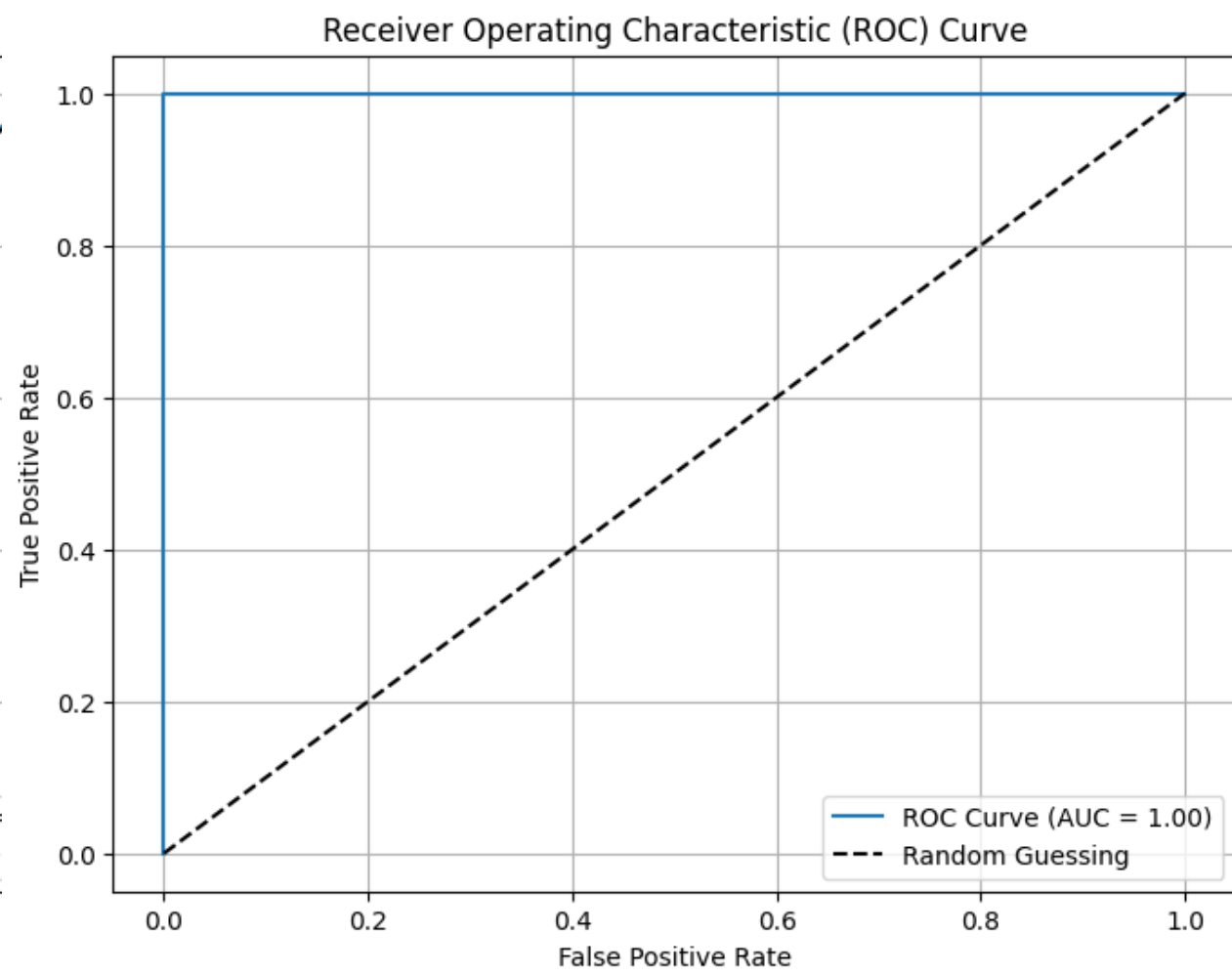
SMOTE: To compensate for gender class unbalance

- Random sample is taken from the minority class.
- k-nearest neighbors are calculated, and one is selected randomly.
- Line is drawn from the original sample to the selected neighbor in the feature space.
- The synthetic sample is then calculated as a convex combination

Before SMOTE

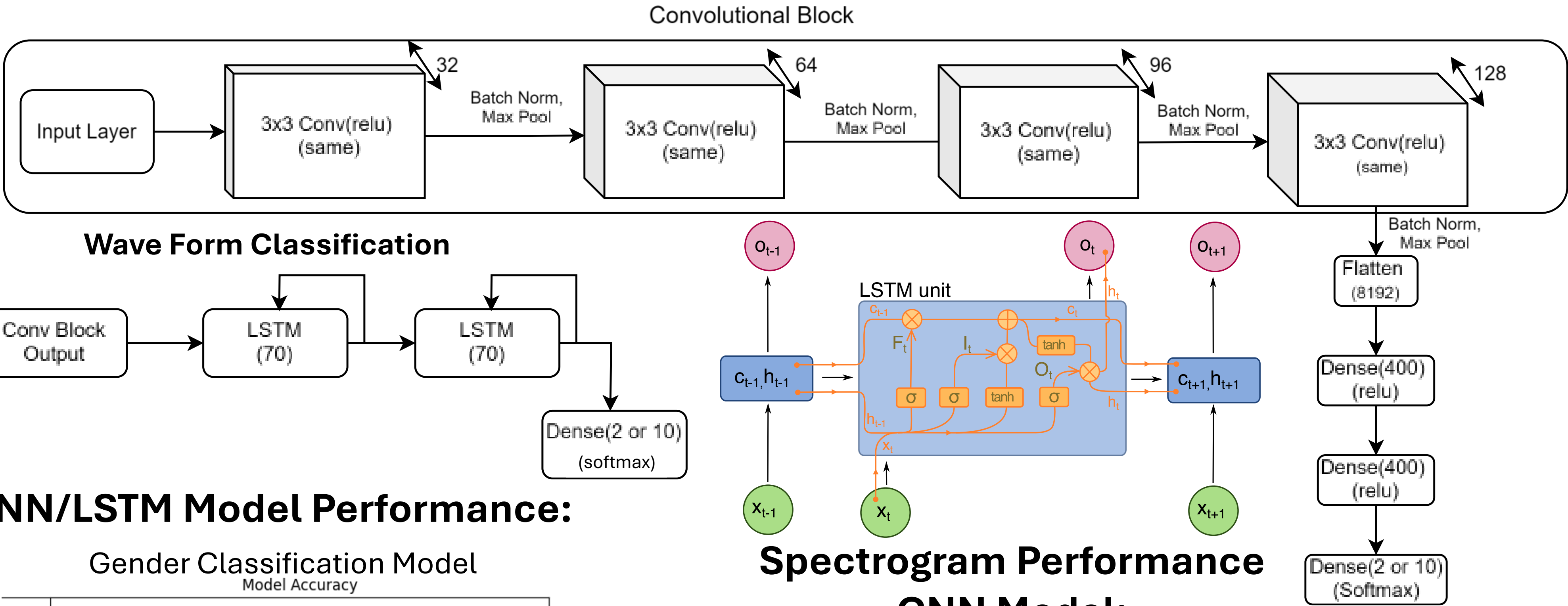


After SMOTE

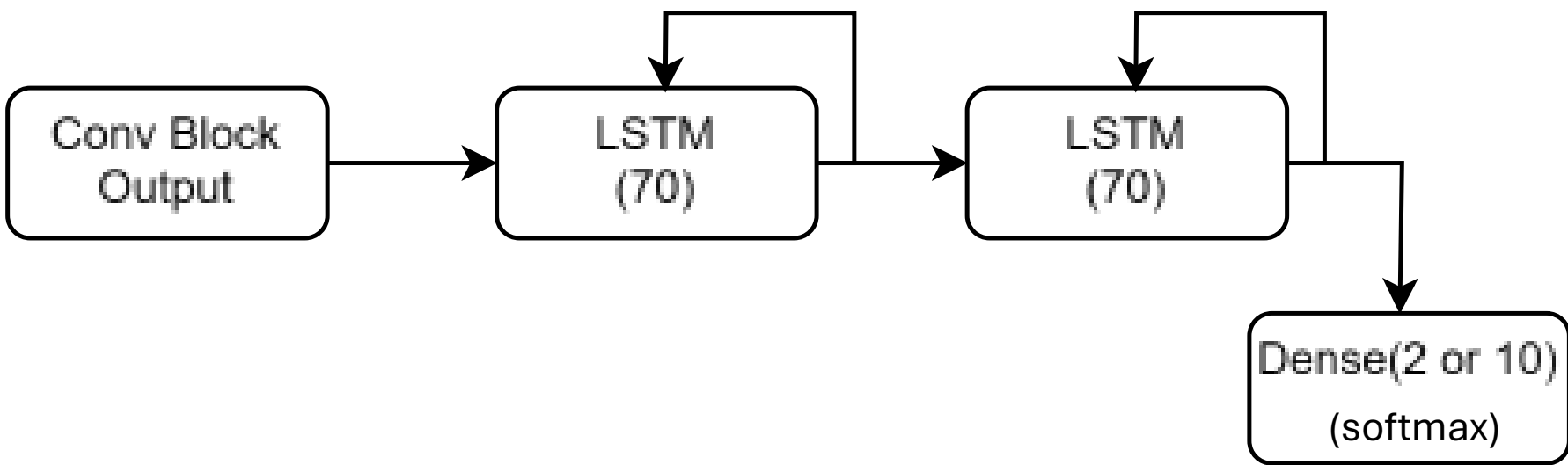


Architecture Block Diagram:

Spectrogram Classification

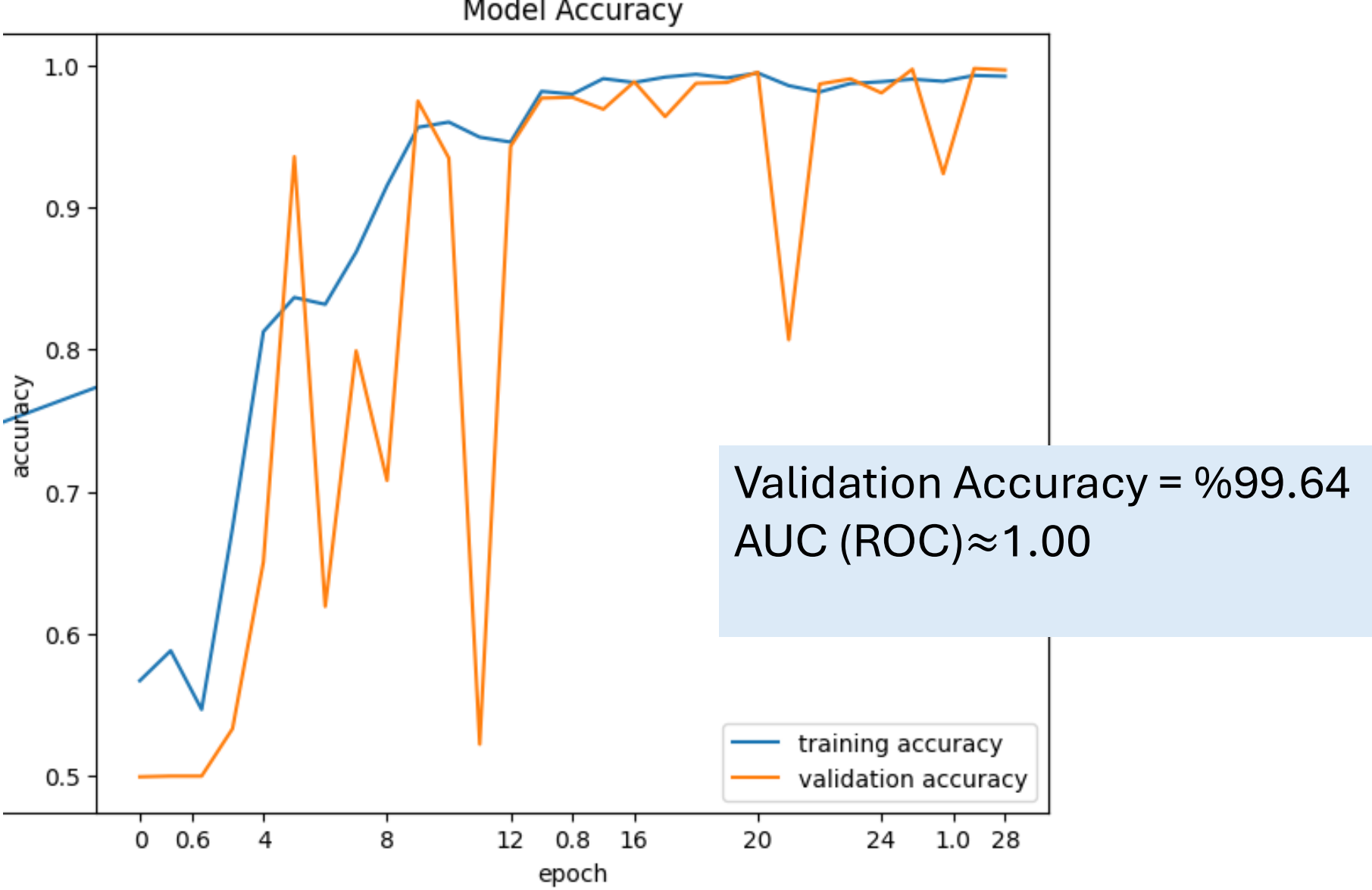


Wave Form Classification



CNN/LSTM Model Performance:

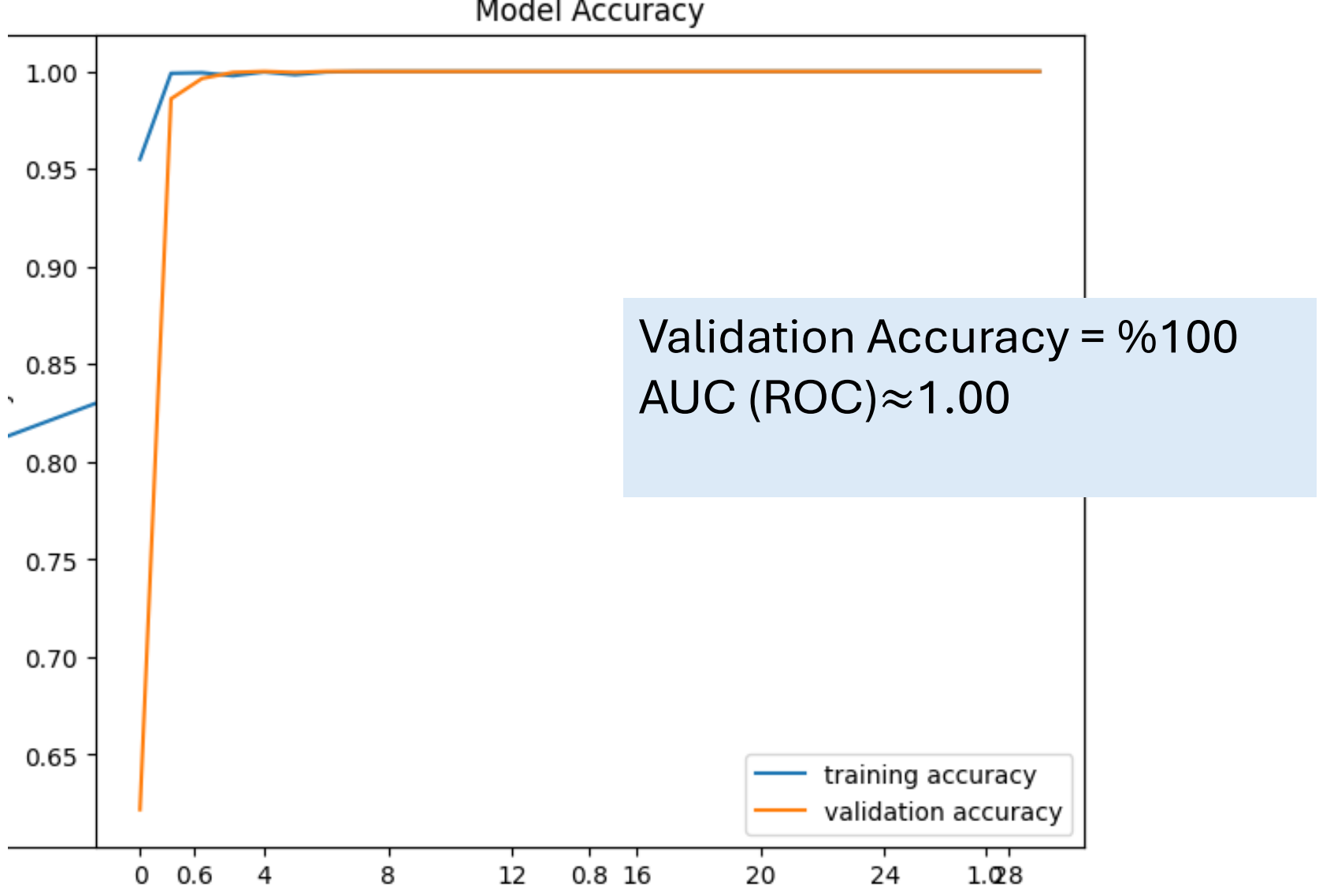
Gender Classification Model



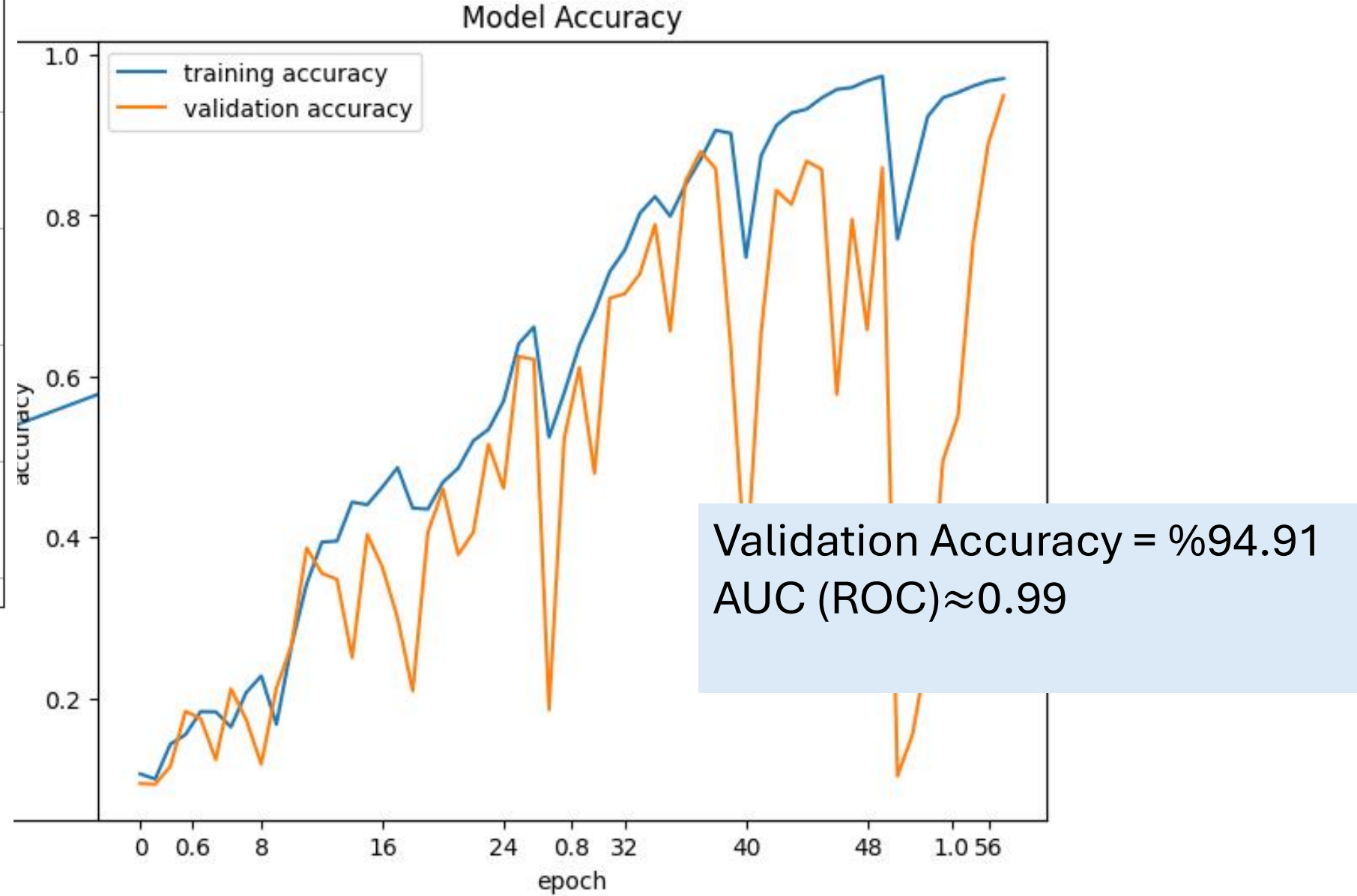
Spectrogram Performance

CNN Model:

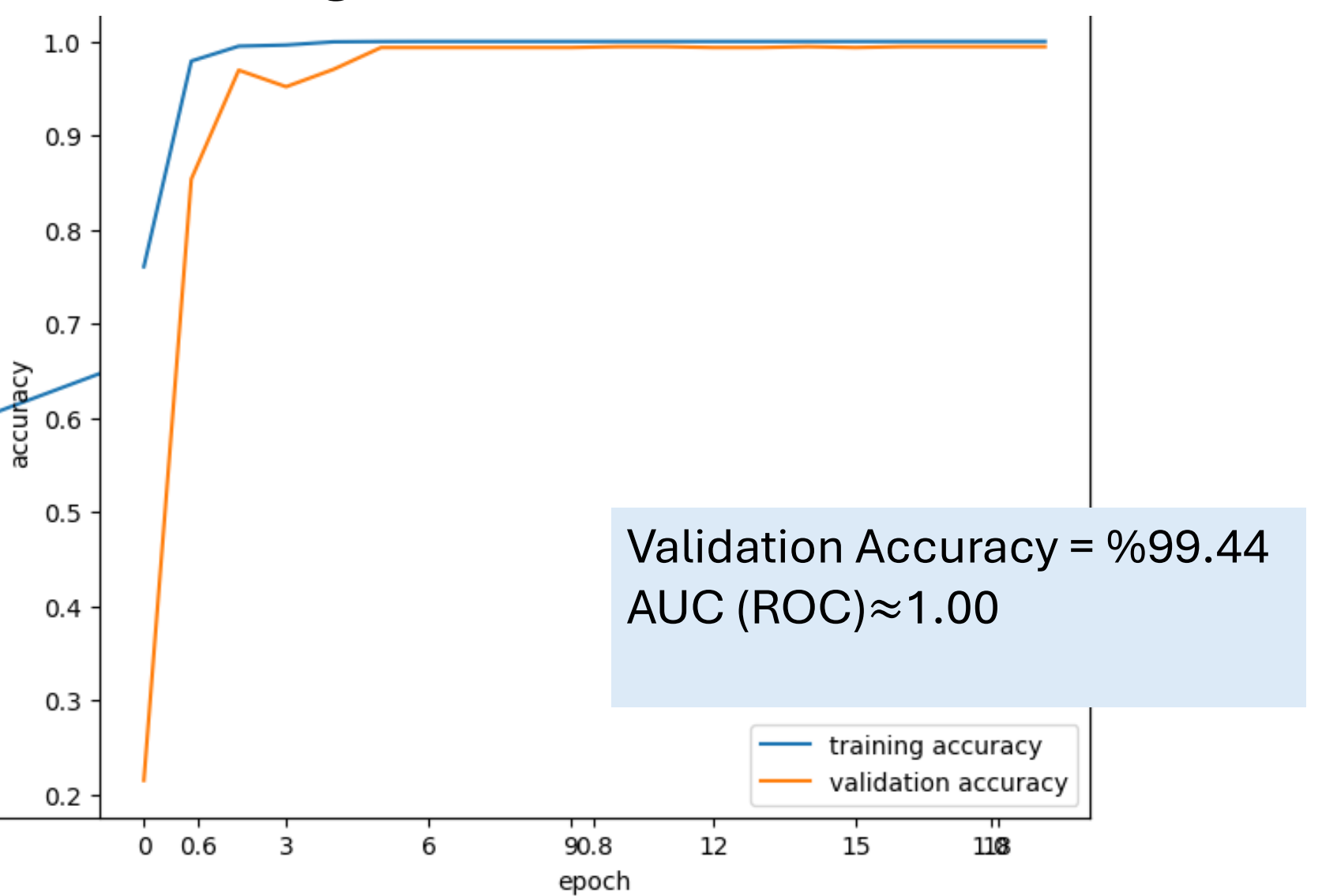
Gender Classification Model



Digit Classification Model



Digit Classification Model



References:

J. Brownlee, "Smote for imbalanced classification with python," MachineLearningMastery.com, <https://machinelearningmastery.com/smote-oversampling-for-imbalanced-classification/> (accessed Nov. 19, 2024).