Loss Curve

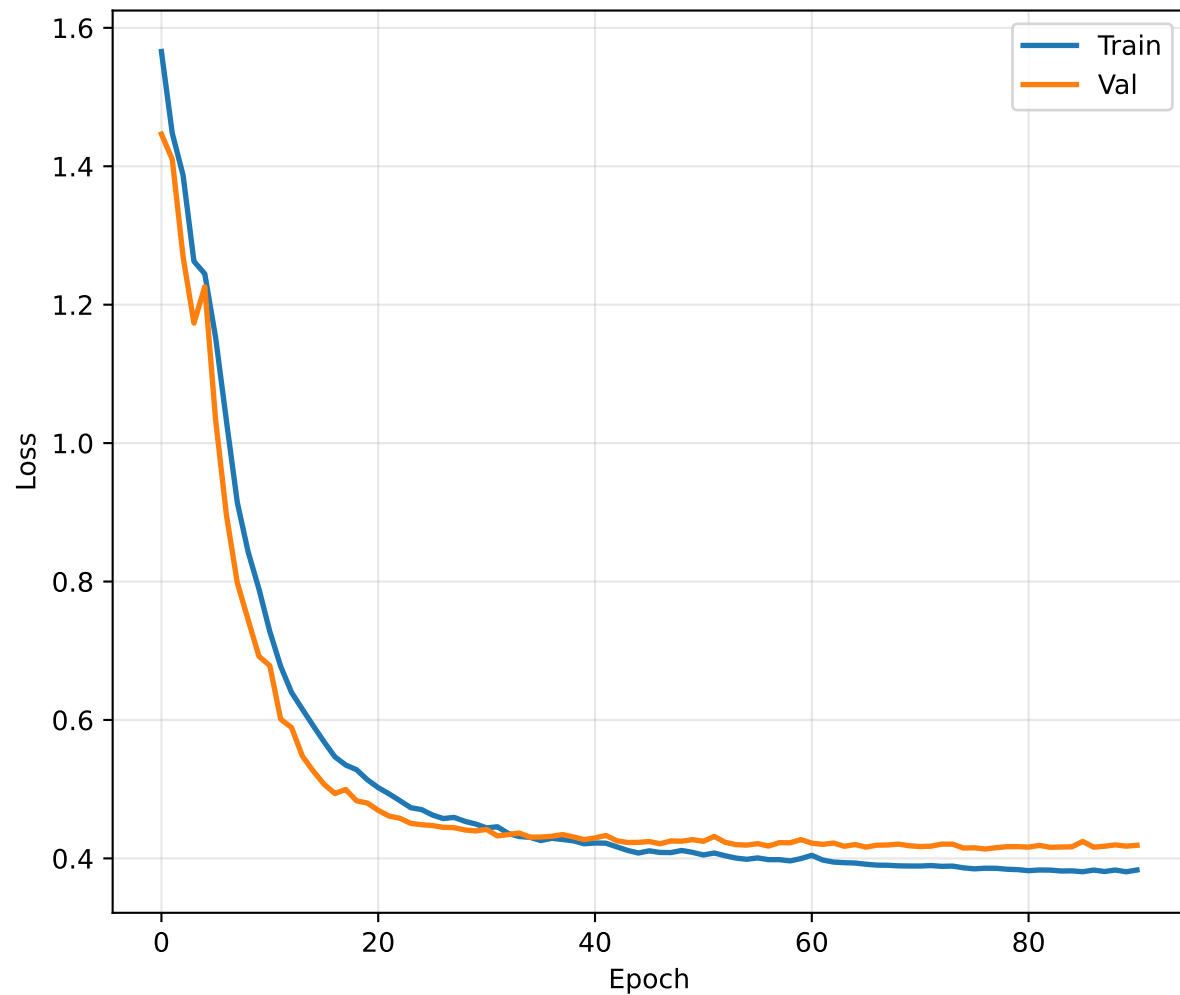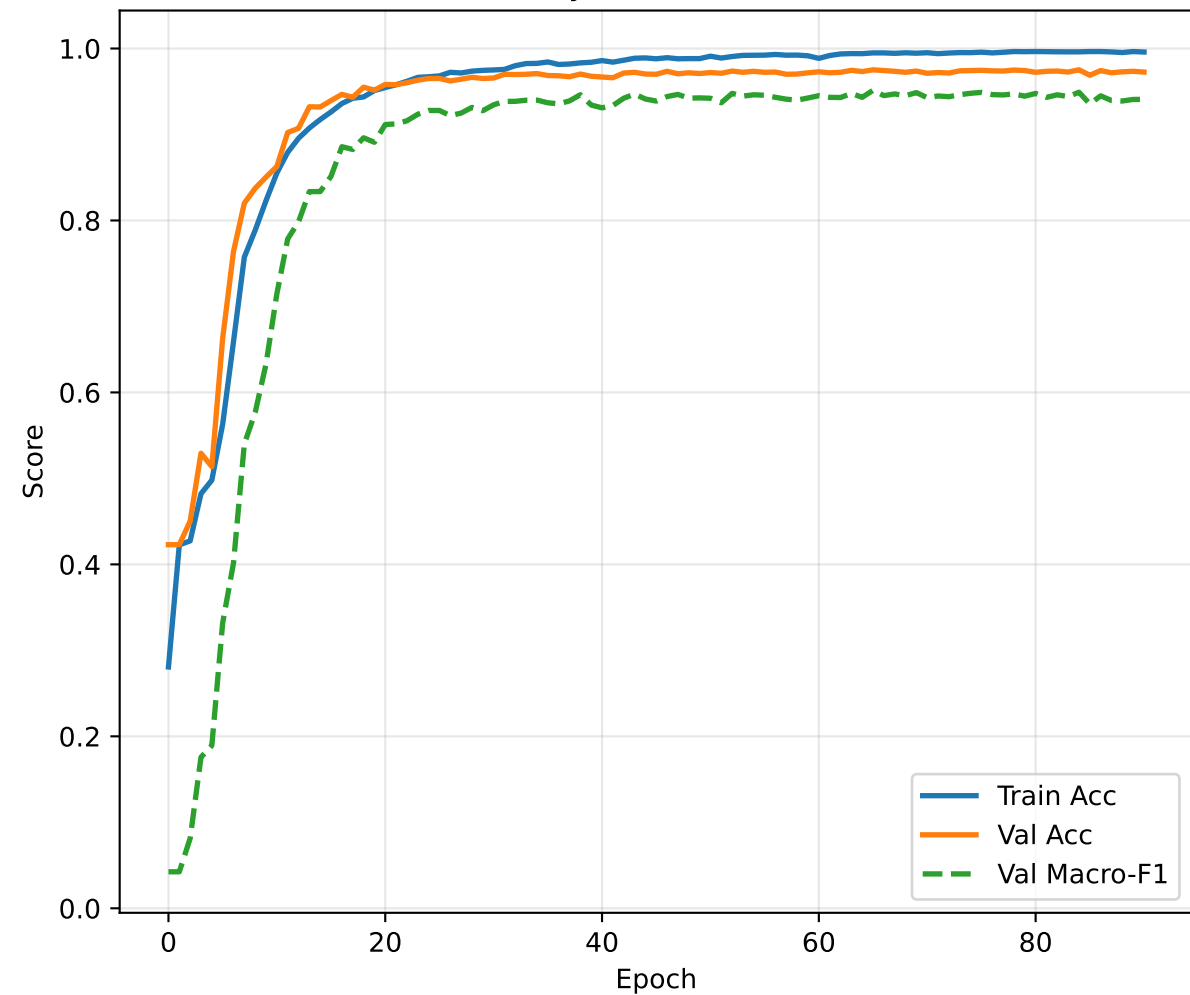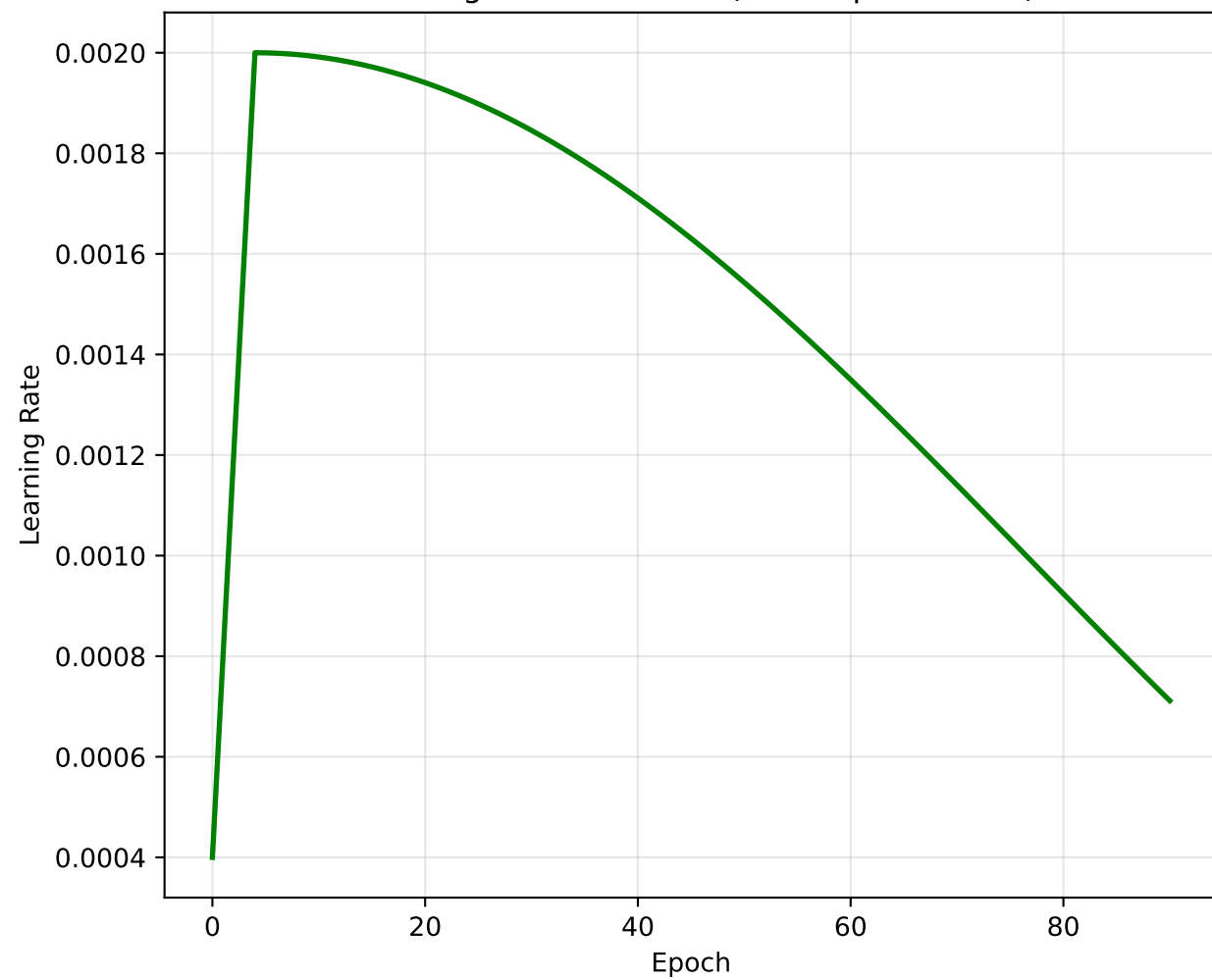**Accuracy / Macro-F1 Curve**

Learning Rate Schedule (Warmup + Cosine)

Confusion Matrix