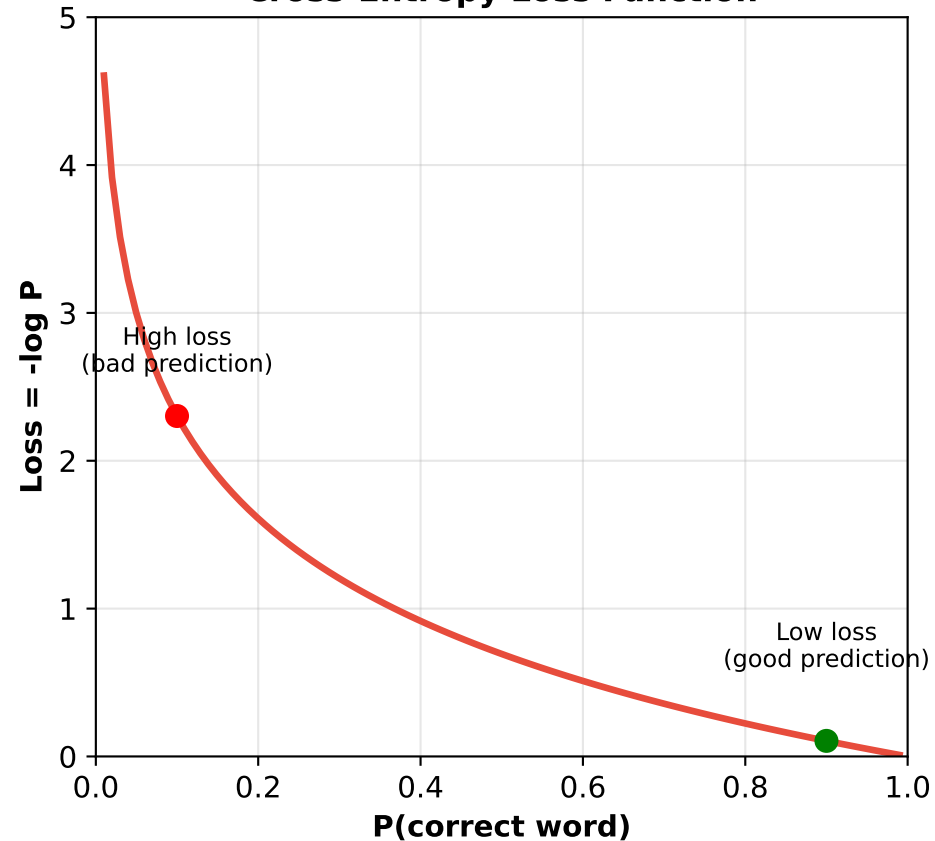


# Cross-Entropy Loss for Next Word Prediction

## Cross-Entropy Loss Function



## Training Loss Over Time

