

Token Usage vs Calibration Error for Correct and Incorrect Predictions

