Model Performance: Accuracy vs Tokens