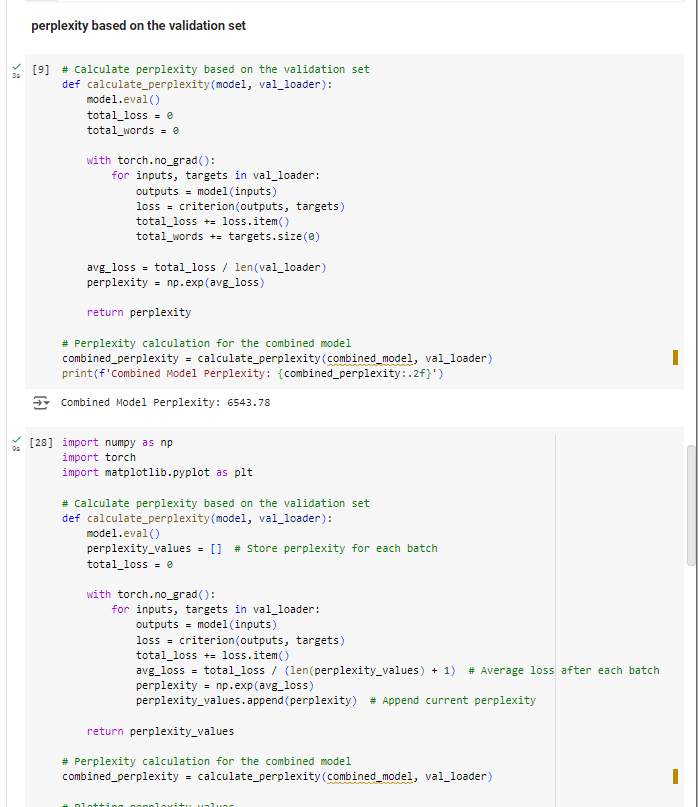
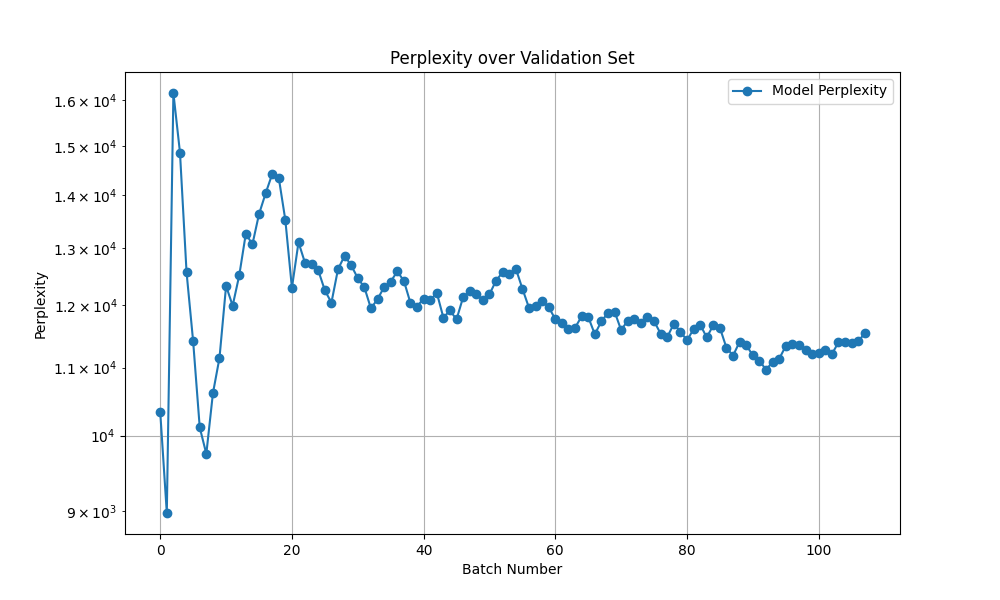
Perplexity Score



Output



The perplexity score plot over the validation set reveals significant insights into the model's performance. Initially, there is a marked spike in perplexity at batch 0, likely due to variability in predictions with the first batch of data. Following this, the perplexity steadily decreases, indicating that the model effectively learns to predict the validation data as training progresses. The values stabilize between 10,000 and 15,000 towards the end of the validation process, suggesting that the model captures the underlying probability distribution of the validation dataset well. Overall, the gradual decline in perplexity reflects the model's improved generalization capabilities, although ongoing monitoring is essential to detect any potential overfitting in future evaluations