

Name: Chandeepea Janith

Index No : SKF2400091

Report on Model Performance Before and After Changes

Tokenization Changes

Previously, the model used character-level tokenization, resulting in a small vocabulary of 65 tokens and generating frequent misspellings. Switching to word-level tokenization increased the vocabulary to 25,670, allowing the model to generate coherent and correctly spelled sentences. This change significantly improved the quality of output, making it more understandable and contextually more accurate.

Activation Function:

Previously, the model utilized the ReLU activation function, which, despite being widely used, sometimes produced outputs that were somewhat robotic and lacked a natural flow. The ReLU function can lead to text that feels disjointed and less coherent. To address this, we replaced ReLU with the GeLU activation function. GeLU is known for its smoother and more natural behavior, particularly in advanced models like GPT. This transition has significantly enhanced the model's text generation capabilities, resulting in more fluid, human-like sentences with better flow and smoother transitions between words and phrases.

Performance Comparison

Before the changes, the model's output quality was poor due to character-level tokenization, this resulted in text that was often unclear and difficult to read. Also, the ReLU activation function sometimes makes disjointed sentences. The lack of effective word processing contributed to limited text coherence. After implementing word-level tokenization and replacing ReLU with GeLU, there was a considerable amount of improvement in output quality. The text became more coherent, readable, and more and most importantly we can clearly understand the intended meaning. The model now generates more natural and human-like text, with word-level tokenization enhancing clarity by properly handling complete words and the GeLU activation function improves sentence flow and readability.