

1. BLEU Score:

- As the beam size increases, the BLEU score generally improves, reaching higher values with larger beam sizes. This is because a larger beam size allows the model to consider more candidate sequences, potentially finding more accurate translations.
- However, after a certain point, the improvement in BLEU score diminishes. This indicates that beyond a certain beam size, the marginal gain in translation quality becomes minimal.

2. Processing Time:

- The processing time increases significantly with larger beam sizes. This is due to the computational complexity of maintaining and evaluating more candidate sequences during the translation process.
- The relationship between beam size and time appears to be exponential. Small increases in beam size lead to disproportionately large increases in processing time.

Personal Take on Optimal Beam Size

Given the trade-offs between BLEU score and processing time, the choice of beam size should balance translation quality with efficiency. Here are the considerations for selecting an optimal beam size:

- **Efficiency:** For applications requiring real-time or near-real-time translations, a smaller beam size might be preferable to ensure quick response times, even if it means a slight reduction in translation quality.
- **Quality:** For applications where translation accuracy is paramount, and processing time is less of a concern (e.g., offline translation of large documents), a larger beam size might be beneficial.

Based on the observed trends, a moderate beam size that offers a substantial improvement in BLEU score without excessively increasing processing time would be ideal. For instance, if the BLEU score improvements start to plateau beyond a beam size of 5 or 10, but the processing time increases significantly, choosing a beam size in that range would be a reasonable compromise.