Training Time vs Vocabulary Size - RTX A6000