Transformer Architecture Growth - Internship (Leo Burgund)

Stéphane Rivaud

April 2025

1 Introduction

Talk about the foundation of first order growth (Manon thesis and Theo report) qsdfmlsqdfqsdf qsdfmlkj

2 Modeling

$X \in \mathbb{R}^{s \times e}, W_Q \in \mathbb{R}^{e \times k}, Q = X W_Q \in \mathbb{R}^{s \times k}$

$$x = 2$$
$$x + 4$$
$$qsd$$

3 Algorithmic approach

4 Experimentation

5 Conclusion