
Karpathy Inference for Llama-2

Ramesh Subramonian

December 20, 2025

Abstract

Step-by-step re-implementation of Karpathy code for llama2.

1 Utilities

1.1 rmsnorm

Arguments in Table ??

Argument	Type
o	float vector of length n
x	float vector of length n
w	float vector of length n
n	integer

Table 1: Arguments for rmsnorm