



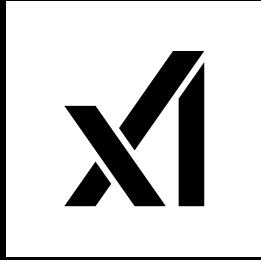
# GROK-1 X.AI

[ A - z ] E X P

March 17, 2024

Open release of Grok-1

We are releasing the weights and  
architecture of our 314 billion parameter  
Mixture-of-Experts model, Grok-1.



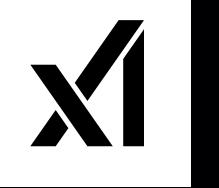
# ABOUT GROK-1 X.AI

X.AI released the base model weights and network architecture of Grok-1, a **314 billion parameter Mixture-of-Experts** model trained from scratch using **JAX**

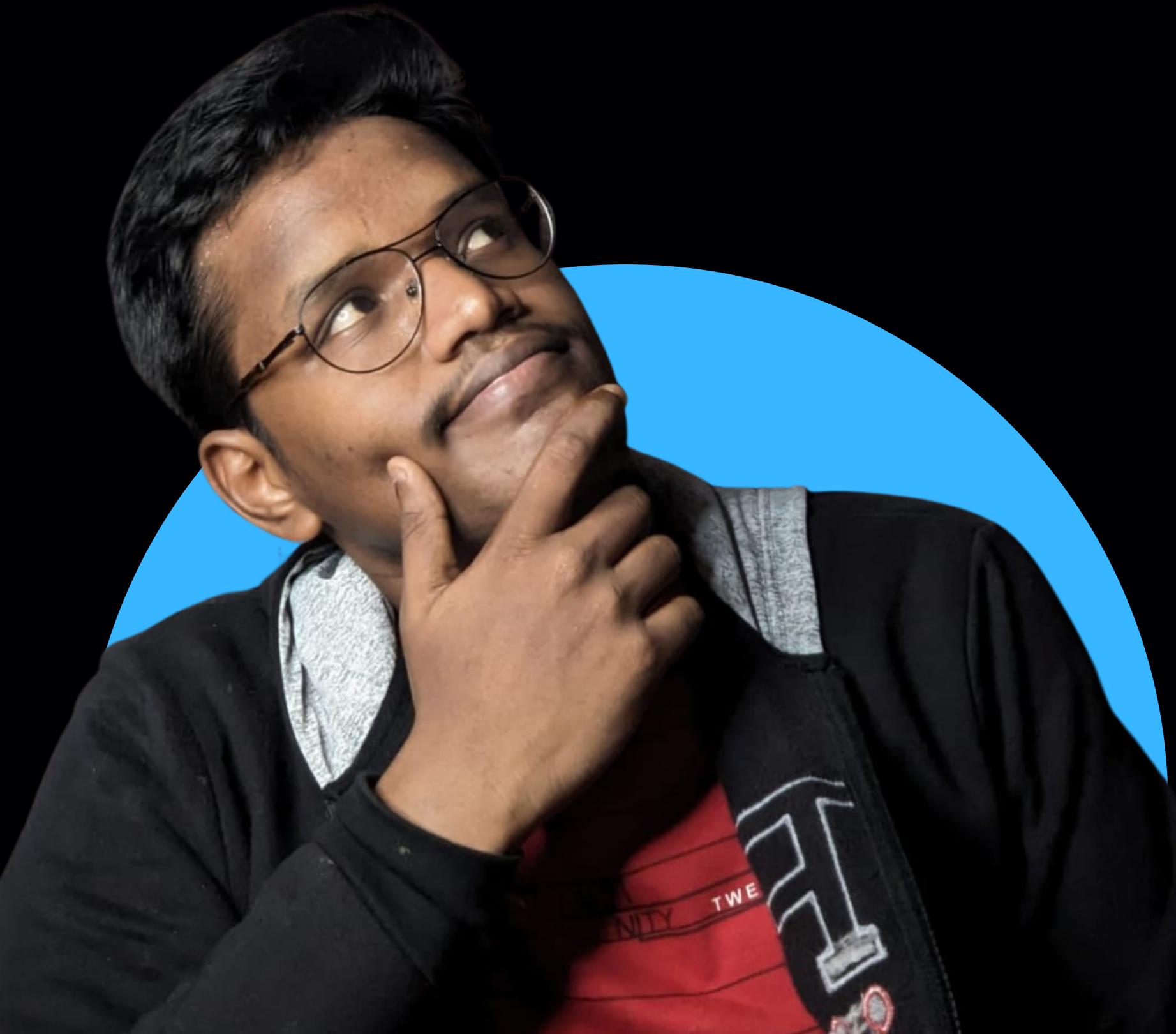
This is the **raw base model** checkpoint from the Grok-1 pre-training phase. This means that the model is **not fine-tuned** for any specific application

We are releasing the weights and the architecture under the **Apache 2.0 license**.





# ARCHITECTURE INFO



*There are changes in every aspect of the model from position encoding to attention. But is it novel and interesting? Lets see one by one*

- Neural Hacks with Vasanth

# ARCHITECTURAL INFORMATION

01

## Activation Function

Approximate **GeLU** activation function is used which also supports **activation sharding**

02

## Position Embedding

**Rotary Position Embedding** is used

03

## Layer Normalization

Unconventionally **4x LayerNorm** is employed using **RMS Norm**

04

## Attention

Uses **Grouped Query Attention** with **48 Query** Attention heads and **8 Key Value** Heads. Scaled with  **$30 * \tanh(30)$**

05

## MOE Related

Slightly different than usual. Router with **8 Experts** - **2 experts per token.**

**Total params 314B, active params 86B**

06

## Other important information

Bias - QKV with bias, O no bias, MLP no bias

Tokenizer - **Sentence Piece Tokenizer**

Vocab Size - **131072**, Max Length -**8192**

Embed\_Dim - **6144**, Num\_Layers - **64**

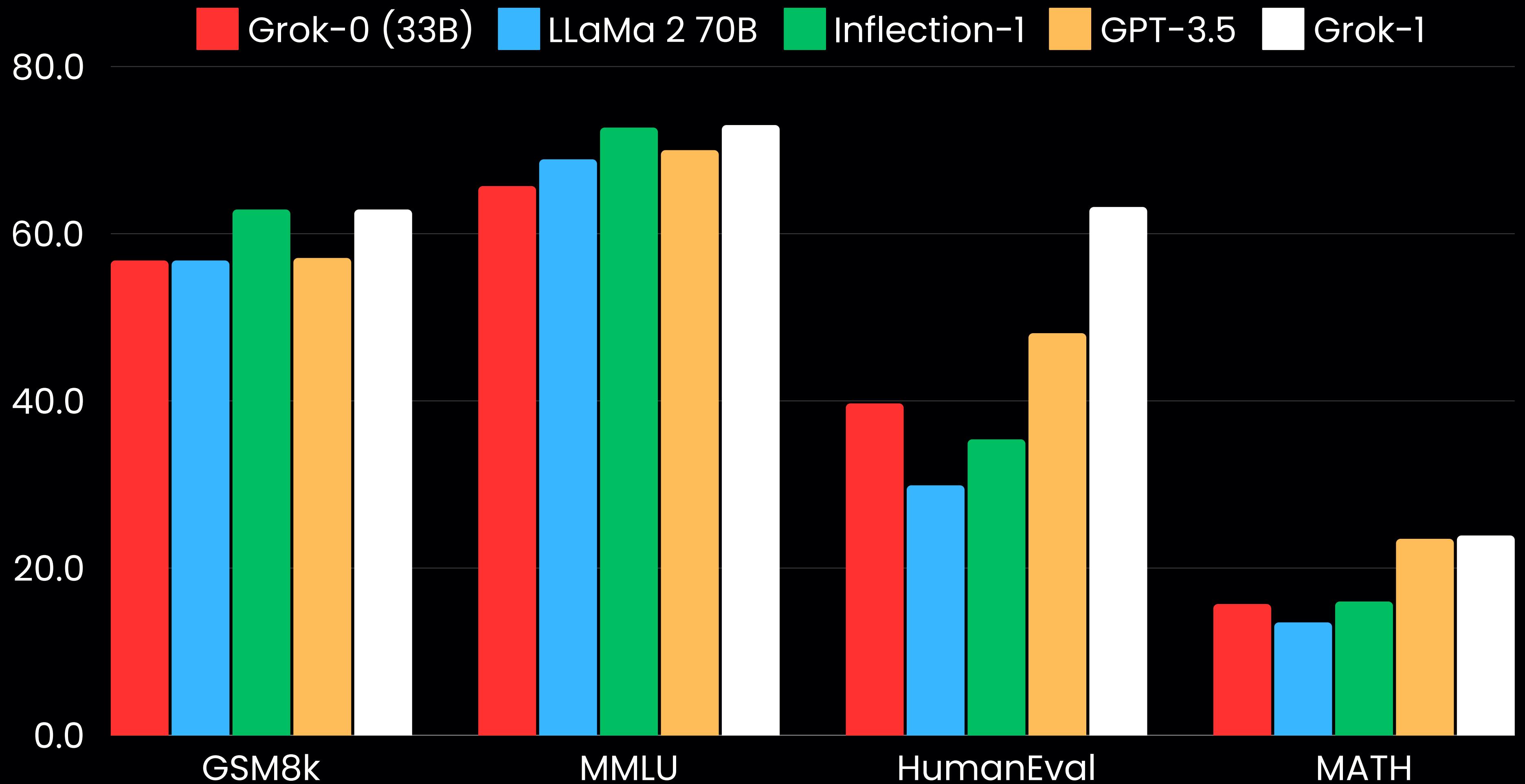


# BENCHMARK

## EXPLAINED

March 17, 2024

We are releasing the weights and  
architecture of our 314 billion parameter  
Mixture-of-Experts model, Grok-1.





# CODE

# WALKTHROUGH

March 17, 2024

Release of Grok-1

We are releasing the weights and  
architecture of our 24 billion parameter  
Mixture-of-Experts model, Grok-1.