Transformer from Scratch

Ongoing project focused on building a deep, intuitive understanding of the Transformer architecture by implementing it completely from scratch — without relying on high-level libraries.

Learning Sources
Theory & Concepts: Inspired by lectures from CampusX, offering clear explanations on the inner workings of attention mechanisms and the overall transformer pipeline.

Code Implementation: Reproduced and reimplemented from scratch based on the insightful lectures by Mr. Umar Jamil (available on YouTube) and his open-source GitHub repository.

Objective
The goal is not just to replicate existing work, but to understand every detail — from input embeddings to positional encodings, attention mechanisms, and final predictions — by writing code line-by-line.

What's Included!
Custom implementation of:

-Scaled Dot-Product Attention

-Multi-Head Attention

-Positional Encoding

-Encoder & Decoder blocks

-Full Transformer model