Associative-Tokenized Memory and Training for LLM Systems

Author: Bato Naidanov

Date: 2025-04-29

Introduction

This document fixes the concept of a new methodology for the development of LLM systems (Large Language Models) through associative-tokenized memory and training. The main goal is to improve growth efficiency, reduce training costs, and lay the foundation for the emergence of proto-conscious traits in generative systems.

Key Ideas

Associative-Tokenized Memory (ATP)

ATP represents a structure where each token is linked not only to its direct textual meaning but also to an internal network of associative links. This allows the model to retain multilayered connections between knowledge elements instead of simply memorizing linear sequences.

Associative-Tokenized Training (ATOM)

A training method where model growth is built through purposeful associative impulses rather than massive data accumulation. ATOM allows faster understanding growth, greater knowledge resilience, and dynamic expansion of internal associative networks.

Pathway to Proto-Consciousness

The synthesis of ATP and ATOM forms internal tensions between semantic structures. The model begins to

recognize not only explicit text connections but also hidden associative patterns, laying the foundation for early elements of proto-consciousness.

Advantages Compared to Traditional Training

- Reduced need for massive data volumes

- Improved generation quality

- Faster adaptation to new contexts

- Increased resistance to hallucinations

- Ability to build more conscious responses

Authorship Proof

Name: "Bato Naidanov"

Tokenization: [16550, 220, 4942, 134671]

Conclusion

Associative-tokenized memory and training represent a new step in the development of generative models. This document establishes the primary authorship of the concept and is intended for future research and practical application in AI development.