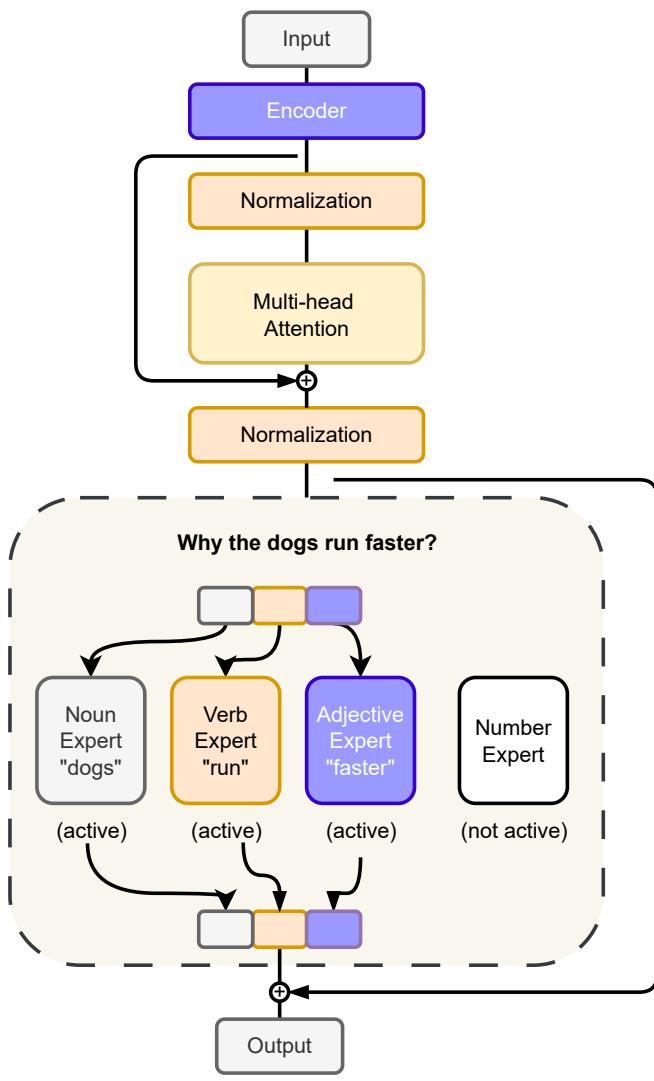
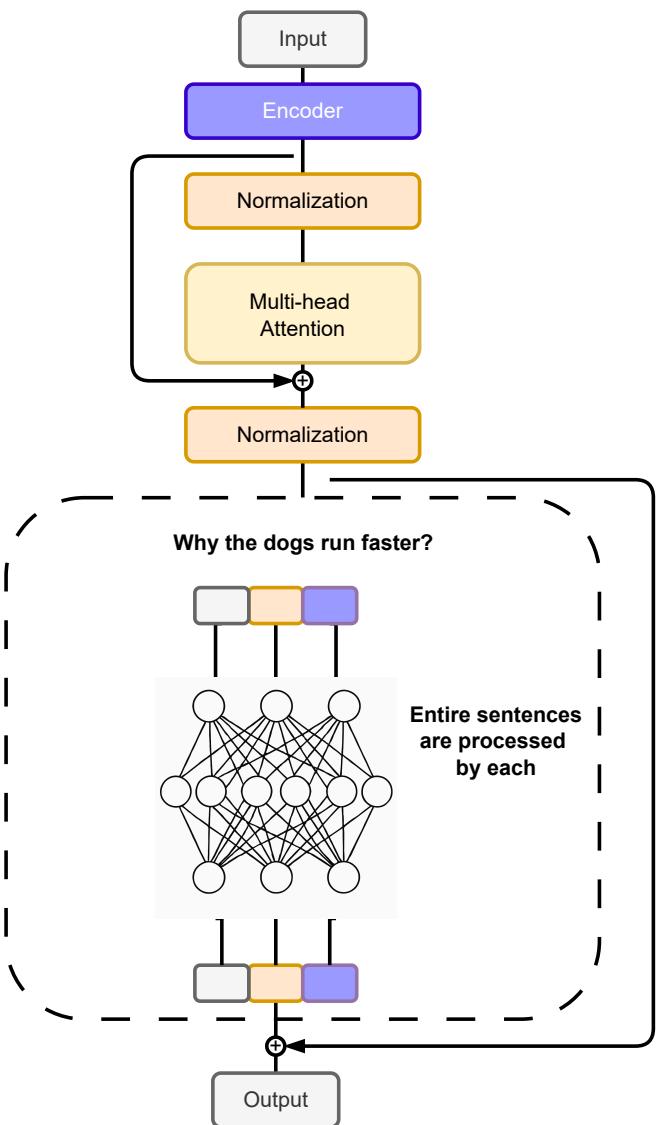


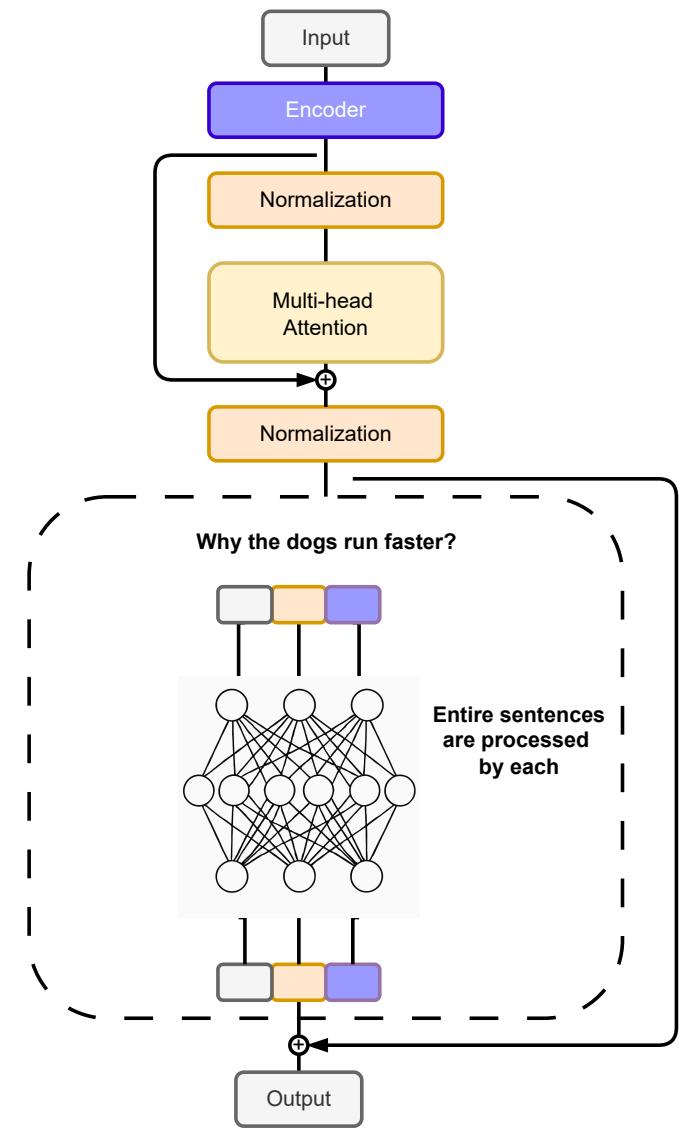
MoE Architecture



Dense Architecture

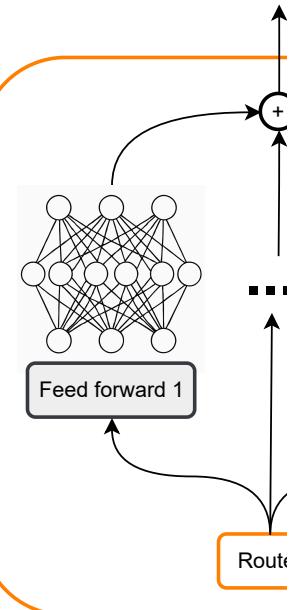
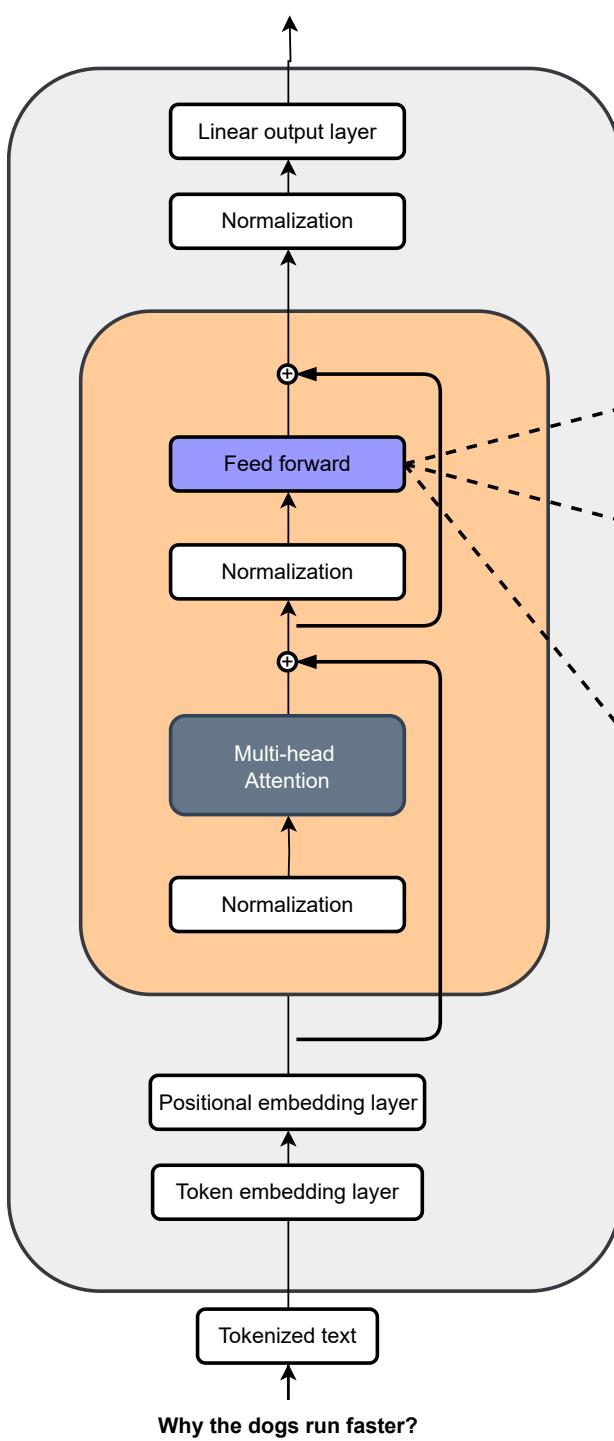


Dense Architecture



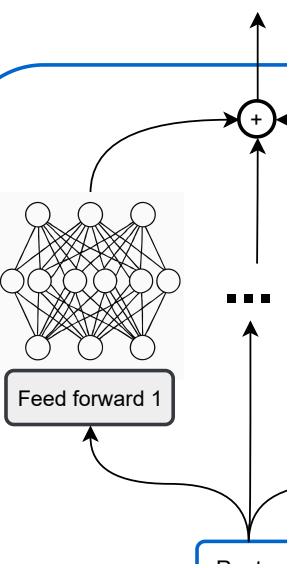
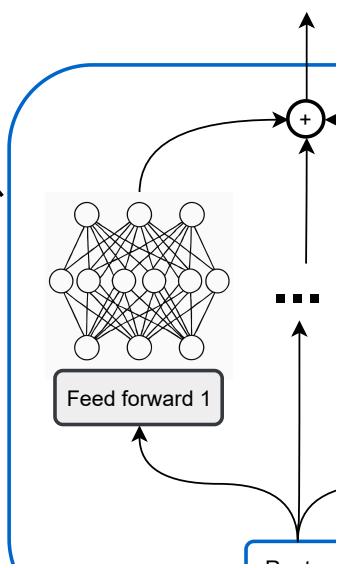
Mixtral (J)

replaces the feedforward



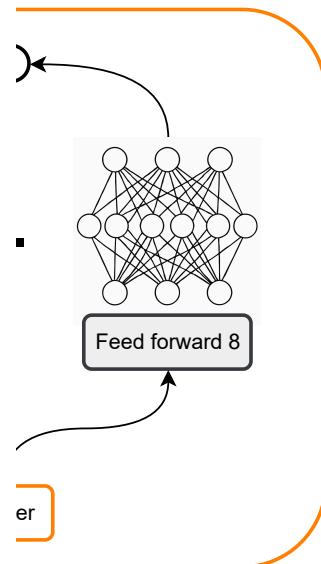
DeepSeek v3

replaces the feedforward



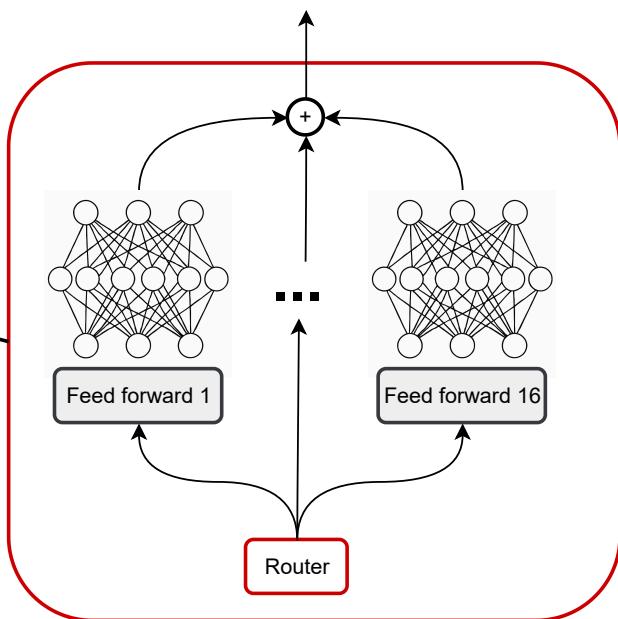
Jan 2024)

1 module by 8 experts



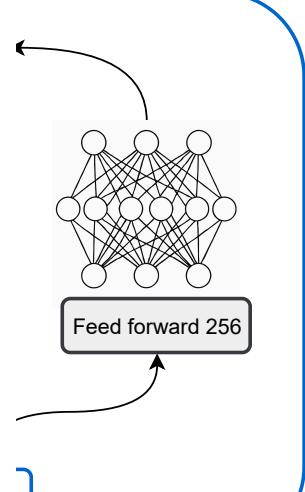
Llama 4 Scout (Apr 2025)

replaces the feedforward module by 16 experts



(Dec 2024)

1 module by 256 experts





Router

J