By Rosetta (Zidian Guo)
interpretersmemory.org

Neural Networks

Labeled Task-specific Training Data

- Trained on smaller, **task-specific** datasets
- Often **manually labeled** (e.g., sentence pairs for translation)

Input → output Interaction

- Each input is processed **independently**, without memory of previous interactions
- The user **doesn't "converse"** with it

Example Use Cases

- Earlier versions of Google Translate
- Traditional speech-to-text engines

Transformers

Massive Unstructured Training Data

- Trained on vast unstructured data (e.g., web text, documents)
- Largely **self-supervised** (no need for manual labels)

Input ↔ response Interaction

- Conversational or **dynamic interaction**: input ↔ response
- Can handle follow-up prompts and context

Example Use Cases

- ChatGPT and other LLMs
- Some self-driving systems that process and "reason" across modalities

Transformers are Built on Neural Networks

The Relationship:

- Transformers are a type of neural network—but with **a new architecture** that changed **how models process data.**

Think of it like:

- Neural Networks are the **foundation**, and Transformers are an advanced branch that opened up new possibilities.