Bi Directional RNN

Best for NLP

is the one, which processes sequence (of text) in both Directions

1. Forward → past to future
2. Backward → Future to past

- A normal RNN, LSTM, or GRU only considers **past context** when predicting the current output.
- A BiRNN combines **past + future context**, making it more powerful for sequence tasks like **speech recognition, text translation, and named entity recognition (NER)**.

2 BiRNN Structure

- Two RNNs (can be simple RNN, LSTM, or GRU):
  - **Forward RNN** → processes sequence from $t_1 \rightarrow t_n$.
  - **Backward RNN** → processes sequence from $t_n \rightarrow t_1$.
- Their outputs are **concatenated or combined** at each time step.

So, the hidden state at time `t` is:

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$$

---

3 Why Better Than Simple RNN, LSTM, GRU?

- **Simple RNN:** Only remembers past context → loses important future info.
- **LSTM / GRU:** Better memory, but still **unidirectional** → only past influences current step.
- **BiRNN:** Uses **both past & future context** → richer understanding, especially in NLP.

Example: In the sentence *"He went to the bank to withdraw money"*,

- A unidirectional model might think *bank = river bank*.
- A BiRNN sees *"withdraw money"* later → correctly infers *bank = financial institution*. ✅

4 BiRNN vs Others — Interview-Ready Table

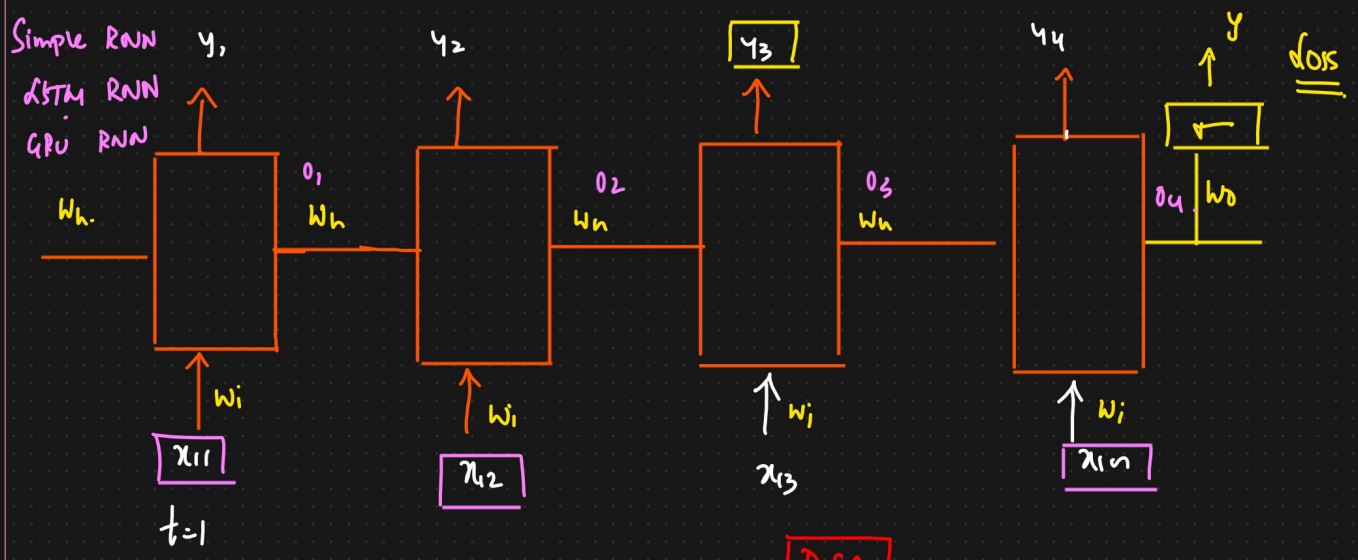
Feature	Simple RNN	LSTM/GRU	BiRNN
Direction	Forward only	Forward only	Forward + Backward
Context captured	Past	Past (better memory)	Past + Future
Complexity	Low	Higher	Highest (2x parameters)
Best use	Time series prediction	Long-sequence tasks	NLP, speech, context-heavy tasks

5 When to Use BiRNN

- Use **BiRNN** when both past and future context matter (e.g., text understanding, speech).
- Use **LSTM/GRU** when you only need **past context** (e.g., stock prediction, sensor data).

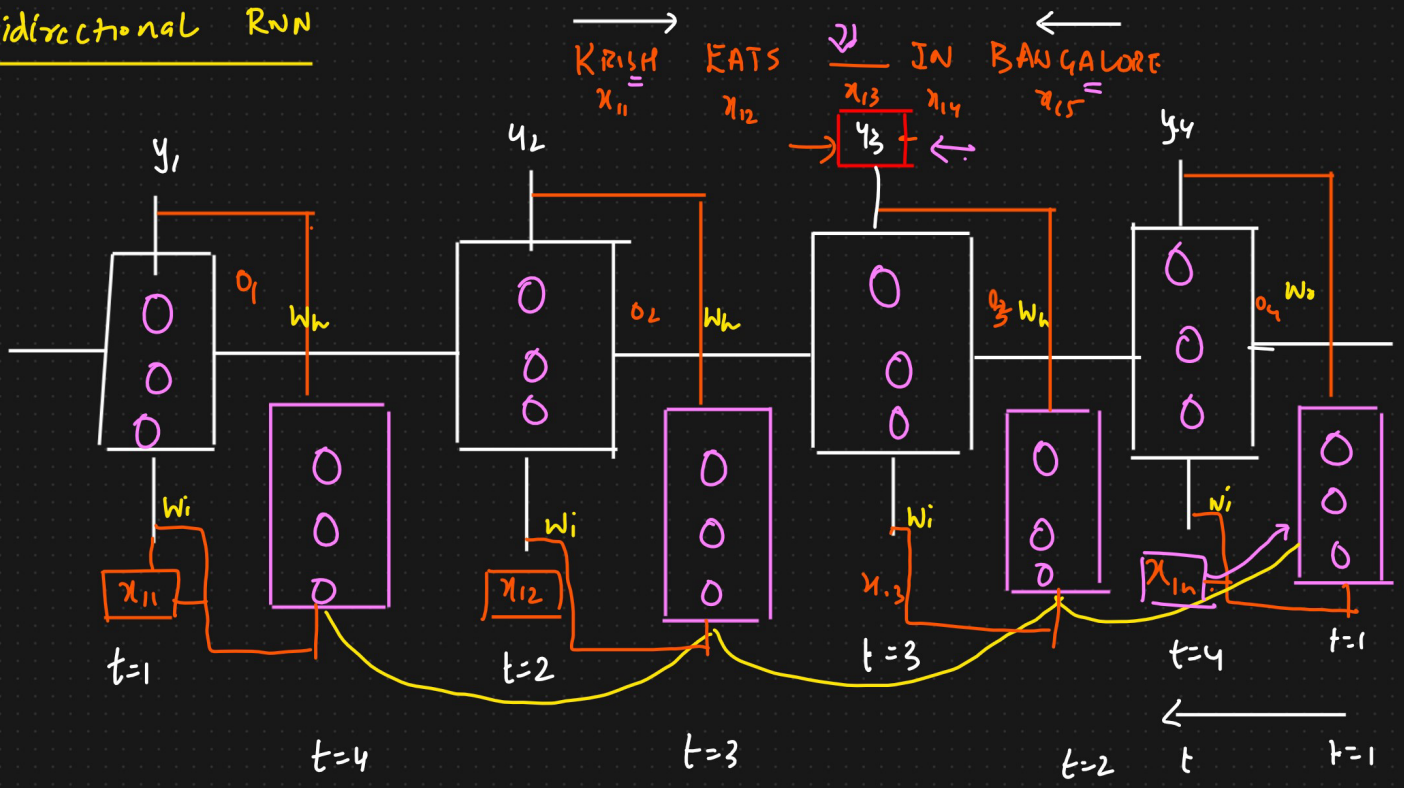
✅ **Quick One-Line Interview Answer:**

"A Bidirectional RNN processes a sequence in both forward and backward directions, allowing it to capture both past and future context, making it superior to standard RNNs, LSTMs, or GRUs in tasks like NLP and speech recognition where future context is as important as past context."



Text Example = KRISH EATS DOSA IN BANGALORE

KRISH EATS PIZZA IN PARIS



① Forward propogation → Equation ←