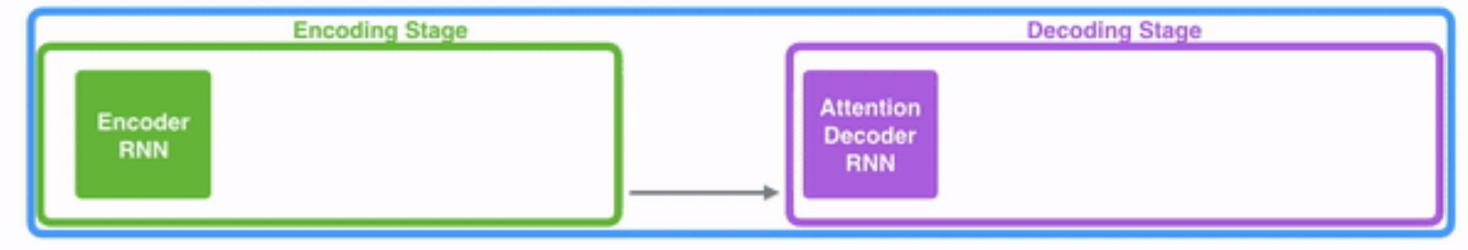


GPT3 175 billion weights Transformer

# **Neural Machine Translation**

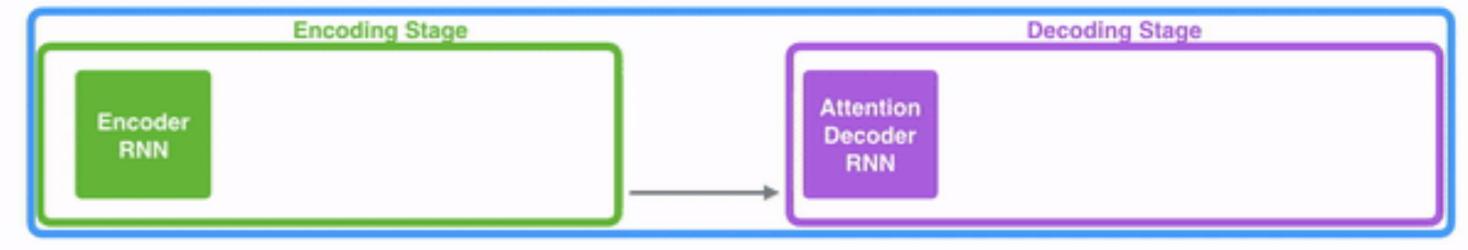
## SEQUENCE TO SEQUENCE MODEL WITH ATTENTION



Je suis étudiant

# **Neural Machine Translation**

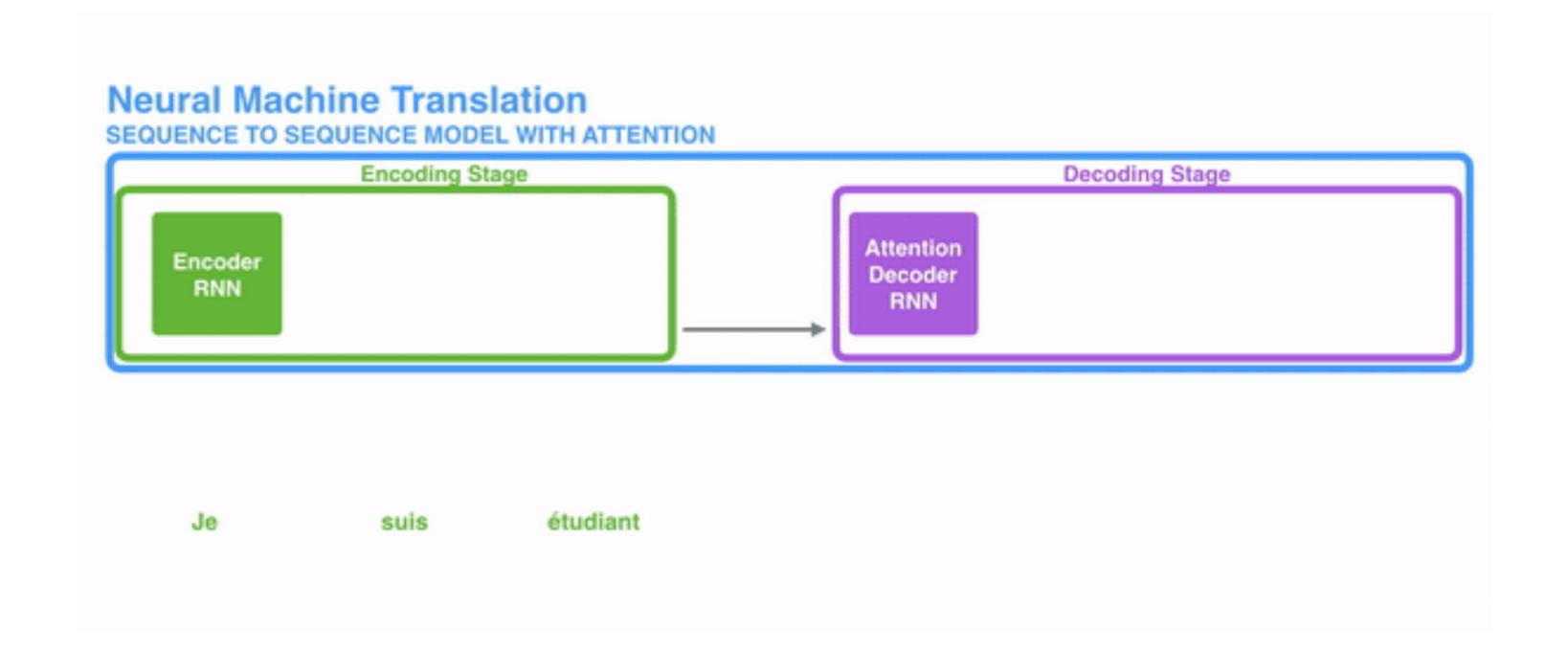
## SEQUENCE TO SEQUENCE MODEL WITH ATTENTION

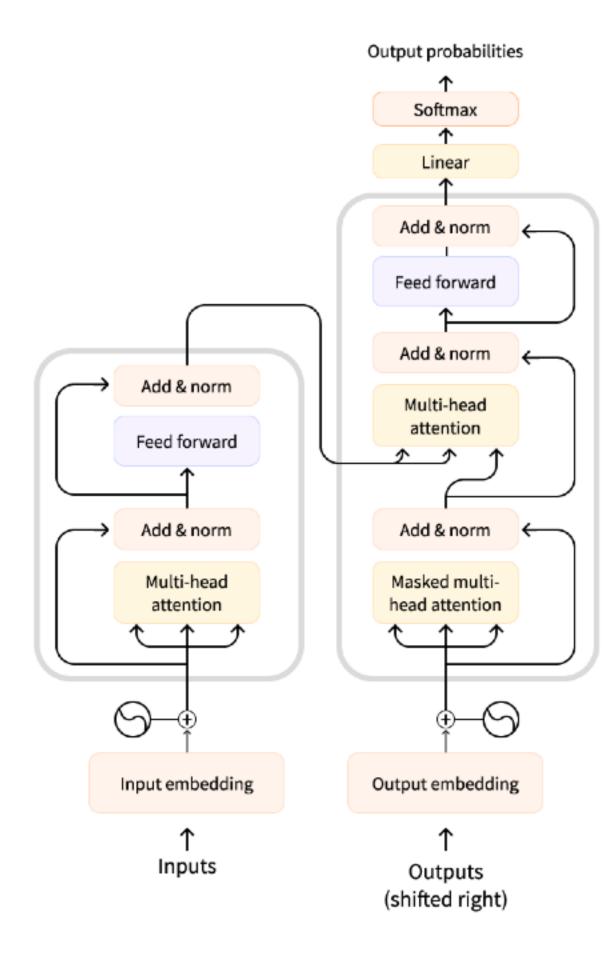


Je suis étudiant

# THE NETWORK

- GPT3 175 billion weights
- Transformer





# ATTENTION

What was wrong with previous approaches?

