

JOHAN CRUYFF - CHATBOT

"Football is a game you play with your brain" - Johan Cruyff

— x — L'Oniva — x —

Brahim OUHAMMOU
Mohamed DALLEL
Rayen BEN ISMAIL



ABOUT THE JOHAN CRUYFF CHATBOT

Do you want to be the best manager in the world and have all possible trophies ?

The Johan Cruyff chatbot is a question answering chatbot that helps you build a high level competitive team by providing you with the tactics for each match.



"1% of chance, 99% of faith" - Neymar Jr





SERVICE US

PLAN

"I learned everything about life with a ball at my feet" - Ronaldinho



MACHINE LEARNING
APPROACH



QUESTION ANSWERING WITH_
BERT AND HUGGINGFACE

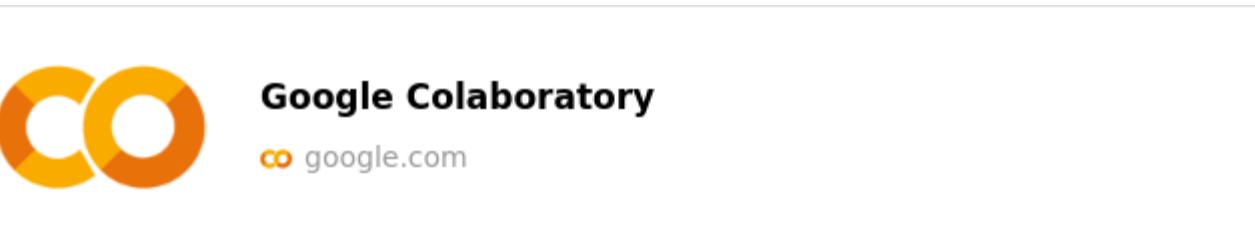


THE MACHINE LEARNING - APPROACH

2023



SVM CLASSIFICATION



CREATING DATA

TRAINING SENTENCES:
QUESTIONS ABOUT
FOOTBALL AND TACTICS

CLASSES:
GREETING, THE RIGHT
TACTIC TO CHOOSE,...

DATA PREPROCESSING

TOKENIZATION
LOWER CASING
STOP WORDS REMOVAL
STEMMING
LEMMATIZATION

MODEL TRAINING & EVALUATING

WE FIT THE MODEL TO
THE DATA AND THE
CHATBOT IS READY.

"If you do a little mistake, everyone kills you, this is normal, this is
life an I'm ready for it" CR7

CHATBOT_SVM.IPYNB - COLABORATORY (GOOGLE.COM)

LIMITATIONS

DATA

1. DATA REQUIREMENTS: MACHINE LEARNING ALGORITHMS REQUIRE LARGE AMOUNTS OF DATA TO TRAIN AND VALIDATE THE MODEL. HOWEVER, COLLECTING AND LABELING DATA FOR NLP TASKS CAN BE TIME-CONSUMING AND COSTLY.

LANGUAGE COMPLEXITY

1. LANGUAGE AMBIGUITY: HUMAN LANGUAGE IS COMPLEX AND OFTEN AMBIGUOUS, WITH MANY WORDS HAVING MULTIPLE MEANINGS DEPENDING ON CONTEXT. MACHINE LEARNING MODELS MAY STRUGGLE TO ACCURATELY CAPTURE THE NUANCE AND MEANING OF LANGUAGE.

LIMITED CONTEXTUAL UNDERSTANDING

1. LIMITED CONTEXTUAL UNDERSTANDING: MACHINE LEARNING MODELS OFTEN LACK THE ABILITY TO UNDERSTAND THE BROADER CONTEXT OF A SENTENCE OR DOCUMENT. THIS CAN LEAD TO MISINTERPRETATION OF THE MEANING OF A TEXT.

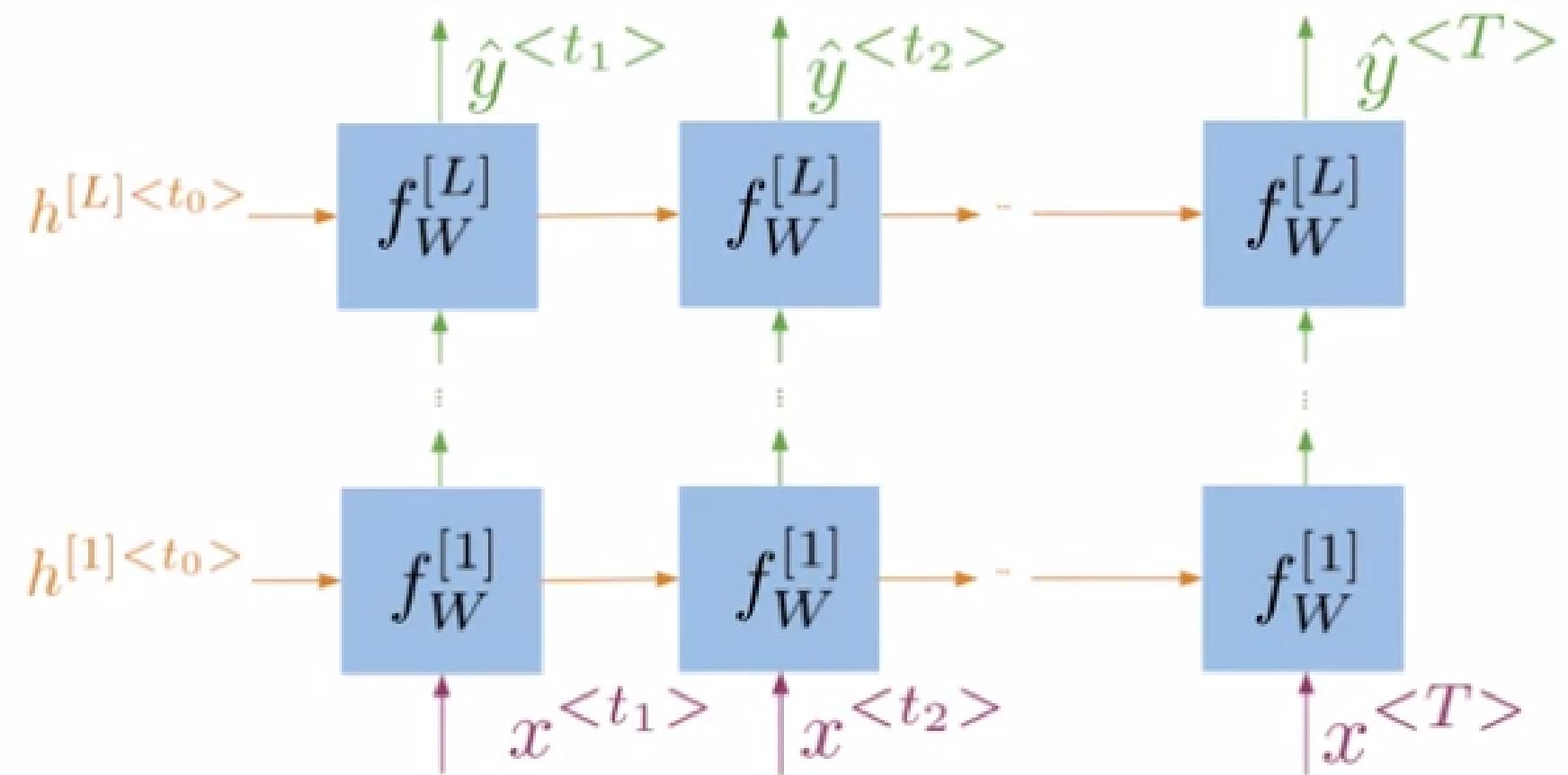
QUESTION ANSWERING WITH BERT AND HUGGINGFACE - APPROACH

— x — L’Orme — x —



MOST USED DNN ARCHITECTURES IN NLP

DEEP RNN



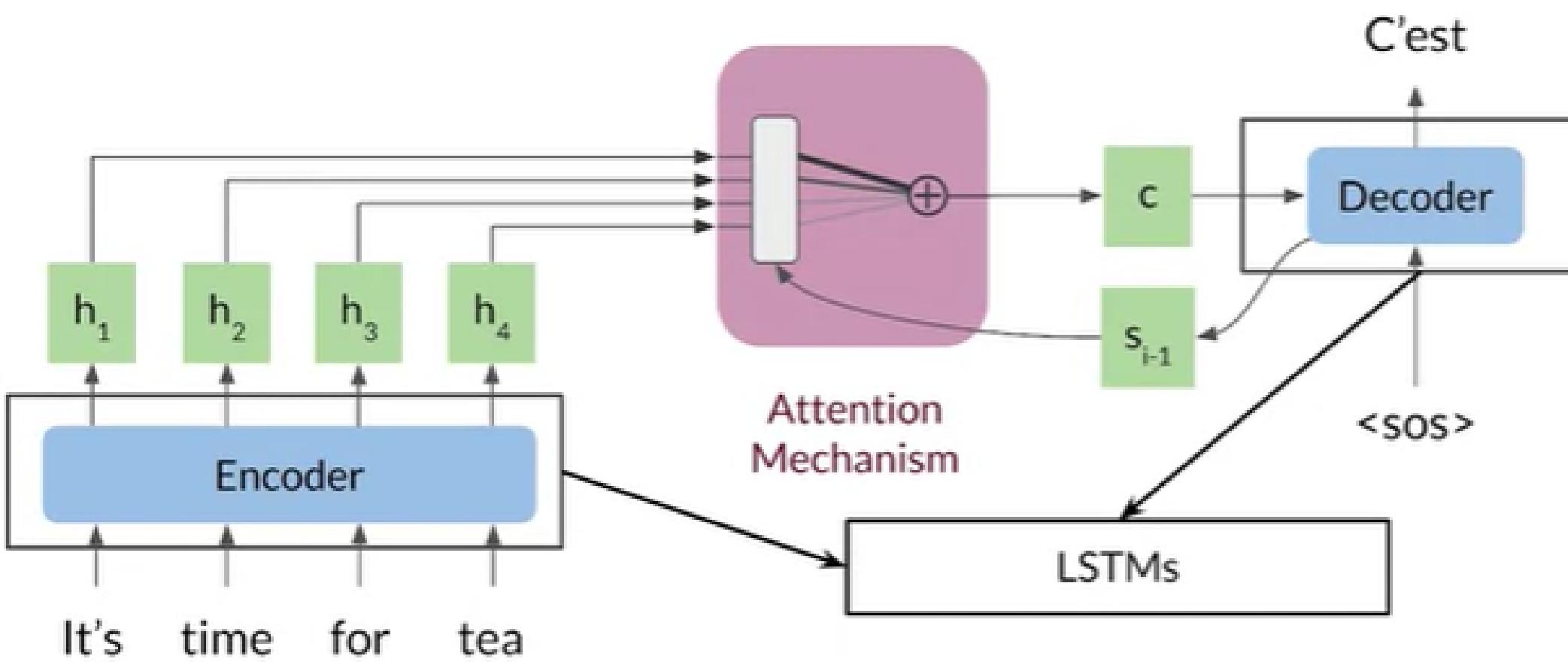
Loss of information

Vanishing gradient



MOST USED DNN ARCHITECTURES

ADDING ATTENTION



PROBLEMS WITH RNN

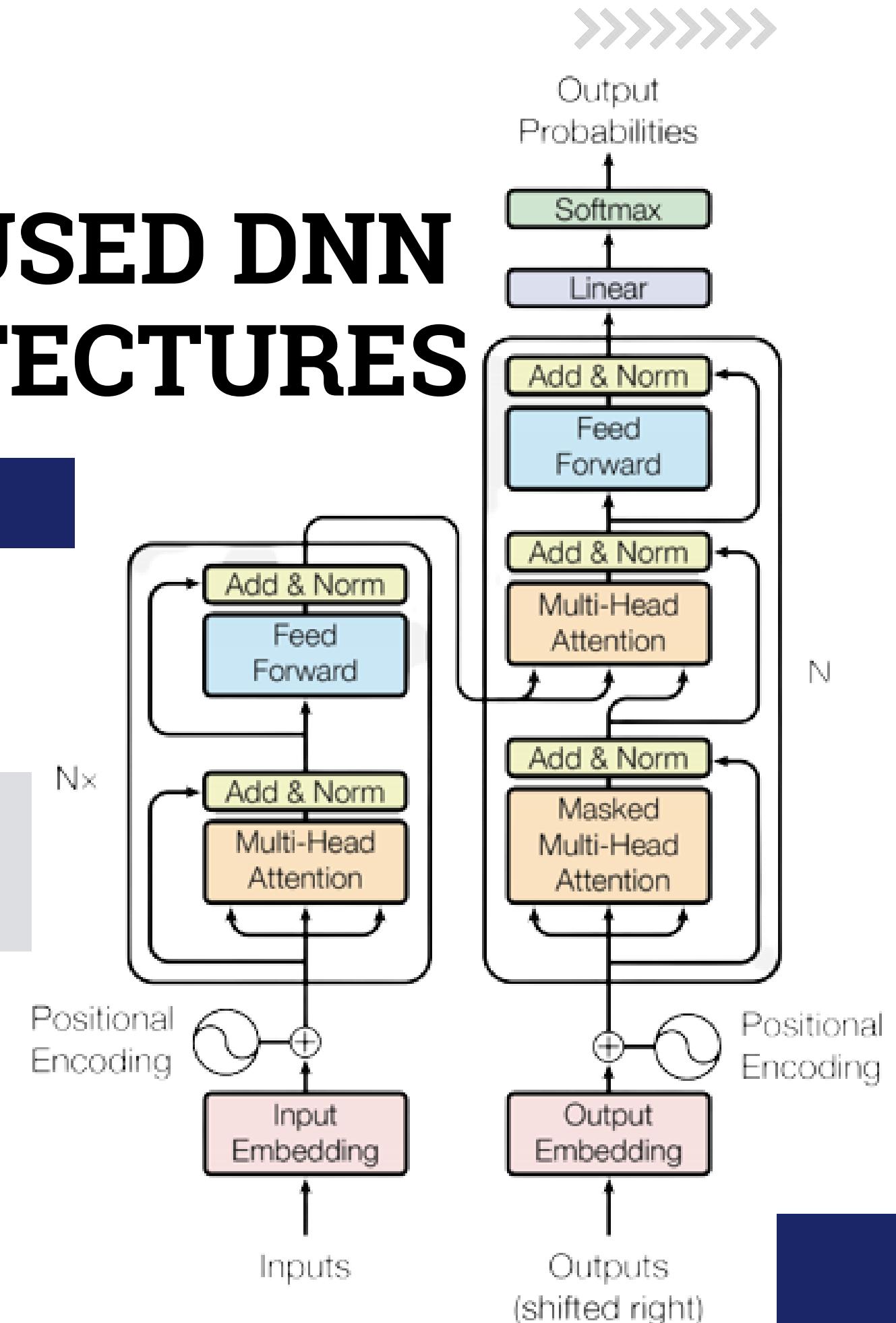
NO PARALLEL COMPUTING
LOSS OF INFO WHEN IT BECOMES TOO LONG



MOST USED DNN ARCHITECTURES

TRANSFORMERS

INPUT: TEXT
OUTPUT: NEXT WORD



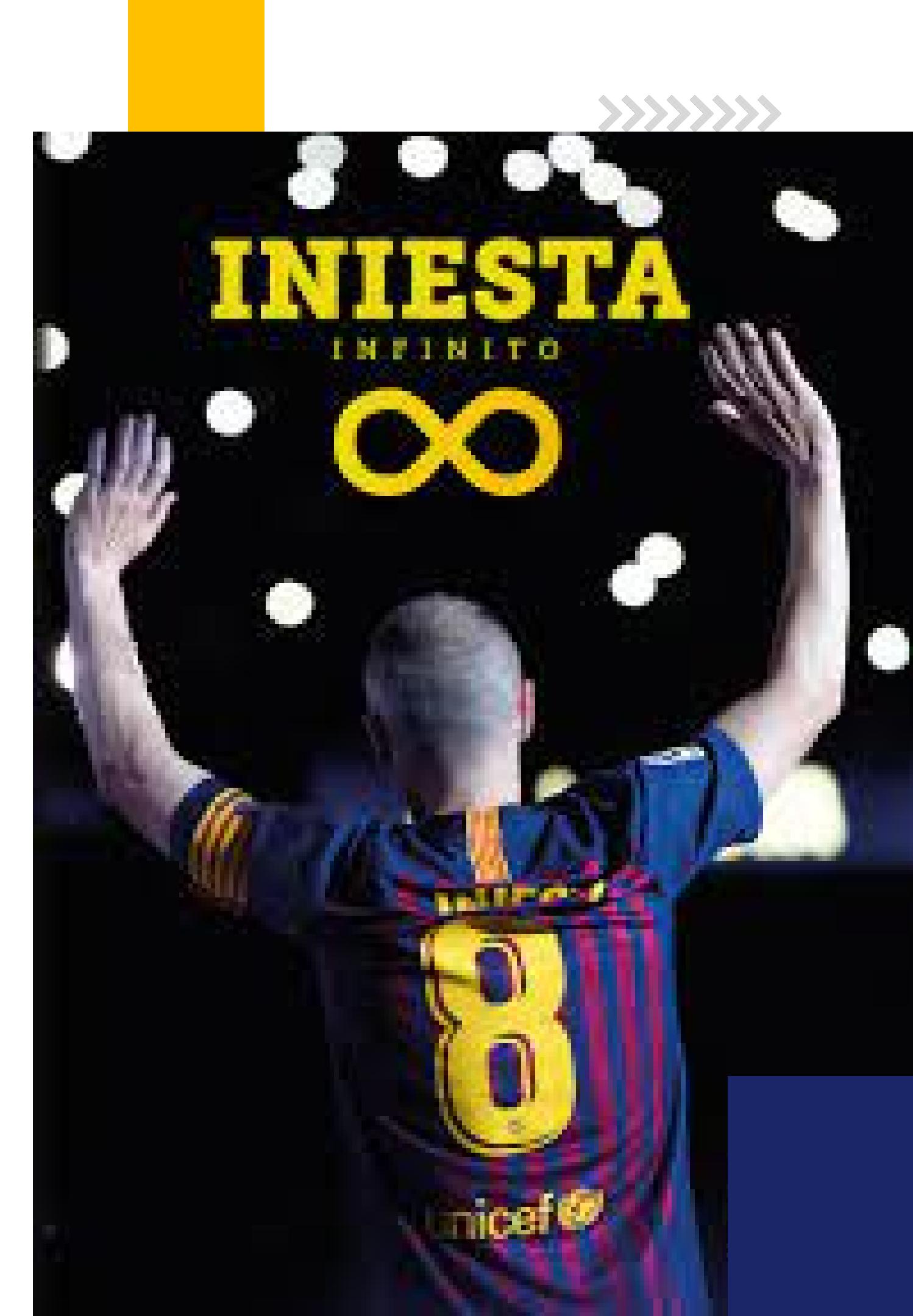
Attention Is All You Need

Ashish Vaswani*	Noam Shazeer*	Niki Parmar*	Jakob Uszkoreit*
Google Brain	Google Brain	Google Research	Google Research
avaswani@google.com	noam@google.com	nikip@google.com	usz@google.com
Llion Jones*	Aidan N. Gomez* †	Lukasz Kaiser*	
Google Research	University of Toronto	Google Brain	
llion@google.com	aidan@cs.toronto.edu	lukaszkaiser@google.com	
Illia Polosukhin* ‡			
illia.polosukhin@gmail.com			



BIDIRECTIONAL REPRESENTATIONS FROM TRANSFORMERS(BERT)

CHOOSEN ARCHITECTURE





OPPORTUNITY

PRE-TRAINING SELF-SUPERVISED

Original text

Thank you ~~for inviting~~ me to your party ~~last~~ week.

Inputs

Thank you <X> me to your party <Y> week.

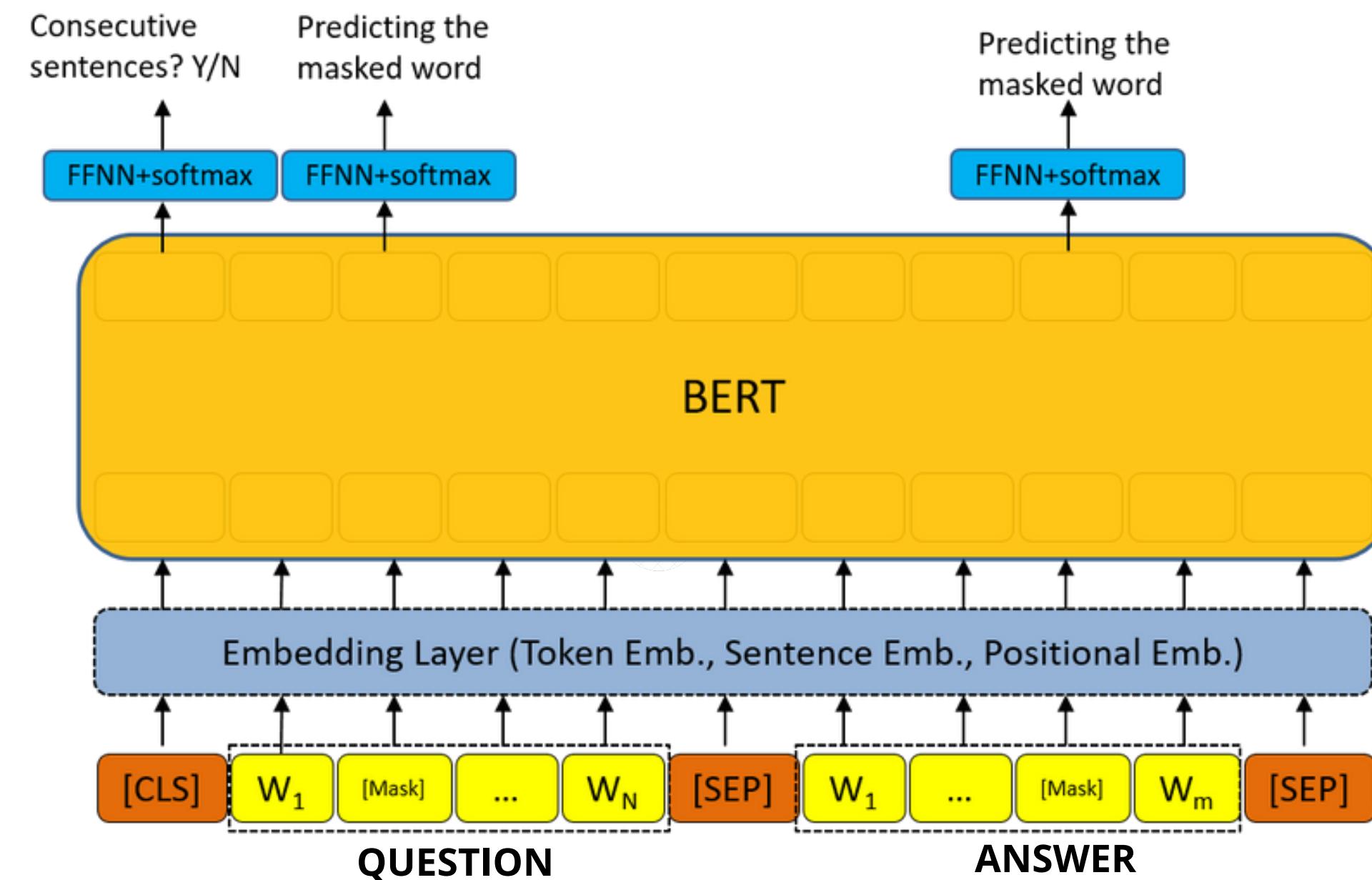
Targets

<X> for inviting <Y> last <Z>



OPPORTUNITY

FINE TUNING TO DOWNSTREAM TASK





OPPORTUNITY

IMPLEMENTATION DETAILS

🤗 Transformers



Google Colaboratory

google.com



THANK

YOU

