# Transformers

Session 1 - Introduction

Sylvain Le Corff - Charles Ollion

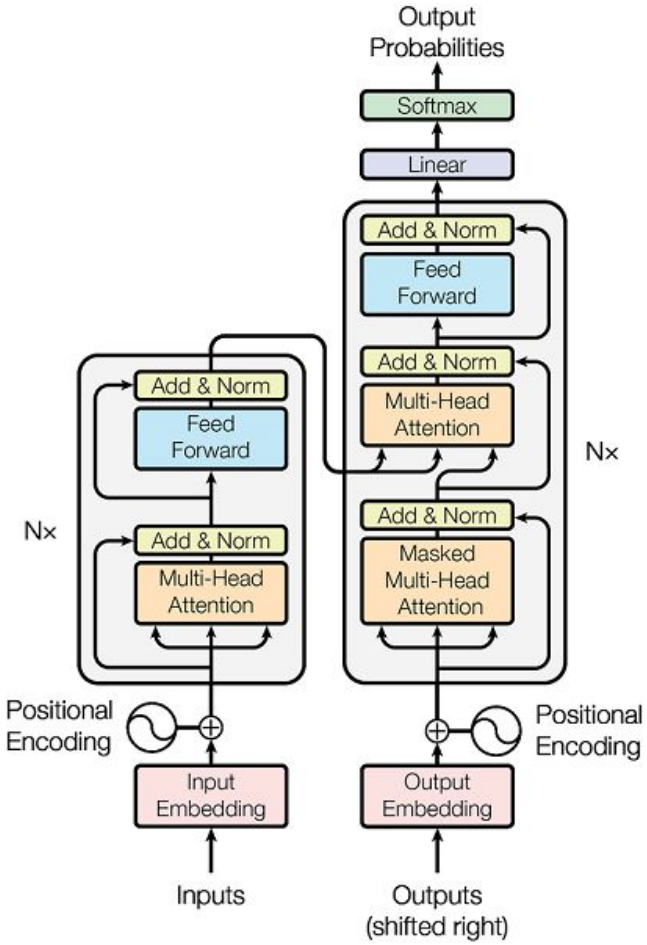
# A Deep Learning Building Block

Similar to Recurrent Neural Network (RNN) or Convolutional Neural Network (CNN)

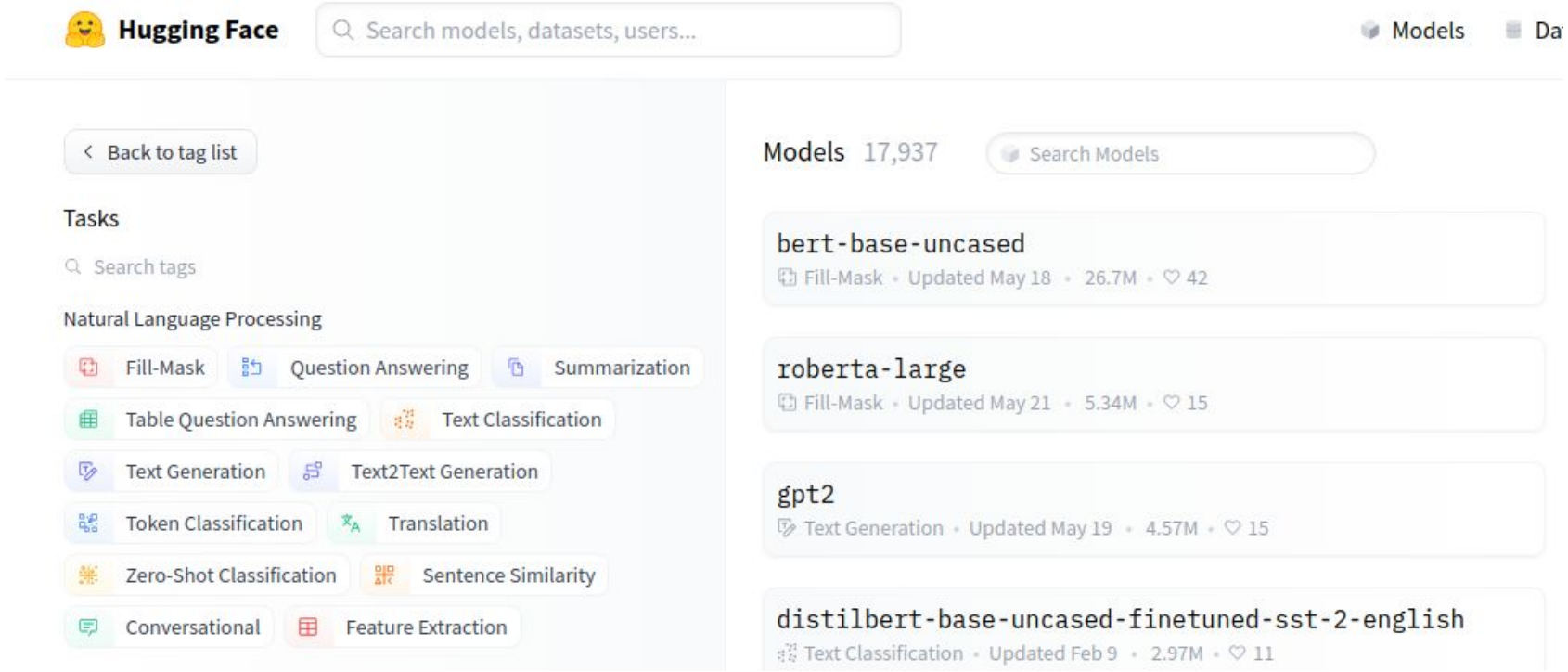
Particularly adapted to process **sets** or **sequences**: in Natural Language Processing (NLP), the input is a sequence of words

Very parallelizable architecture: good for large scale Deep Learning

[Attention is all you need]

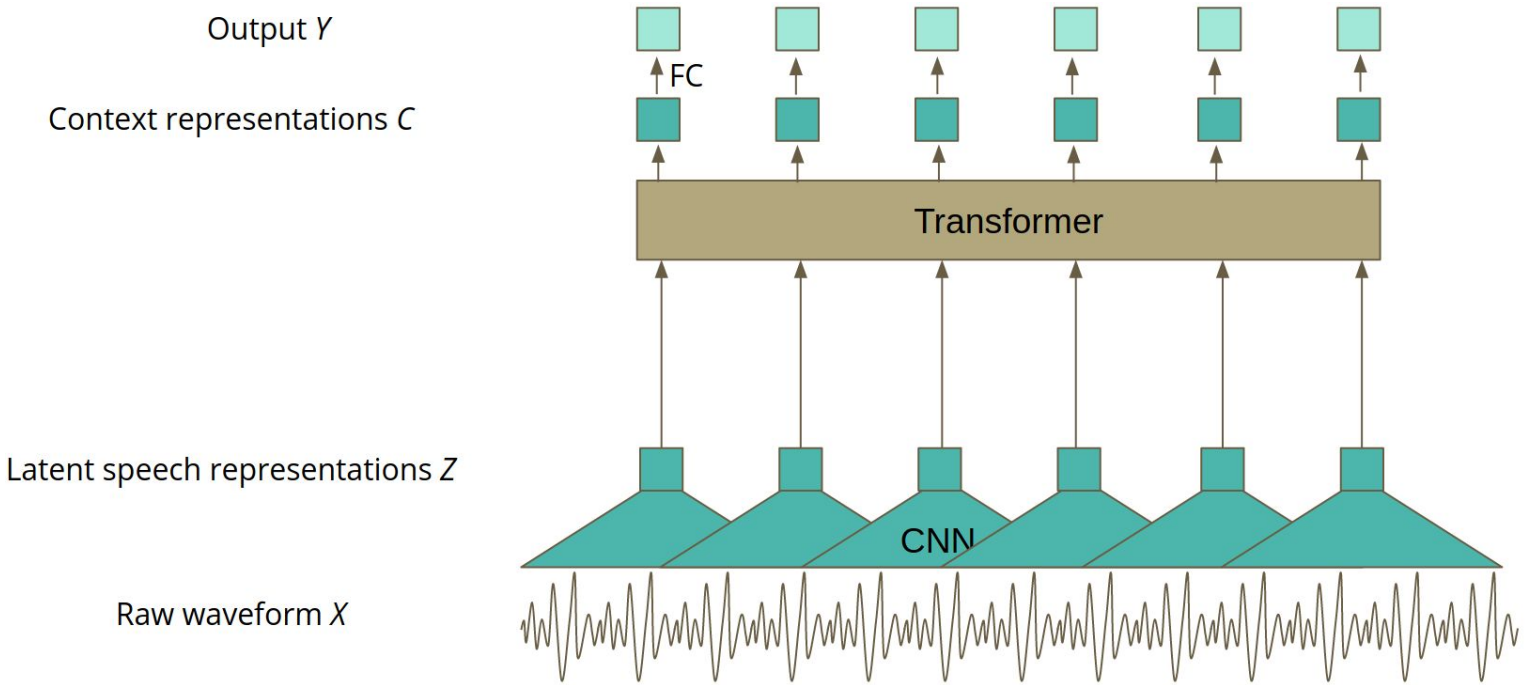


# Where are Transformers today : NLP



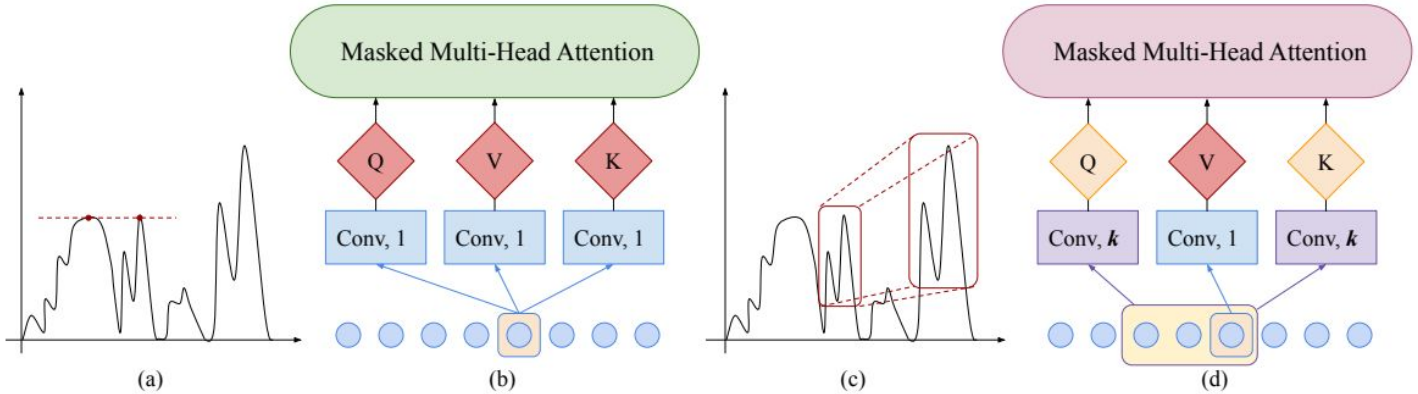

https://huggingface.co/models

# Audio analysis



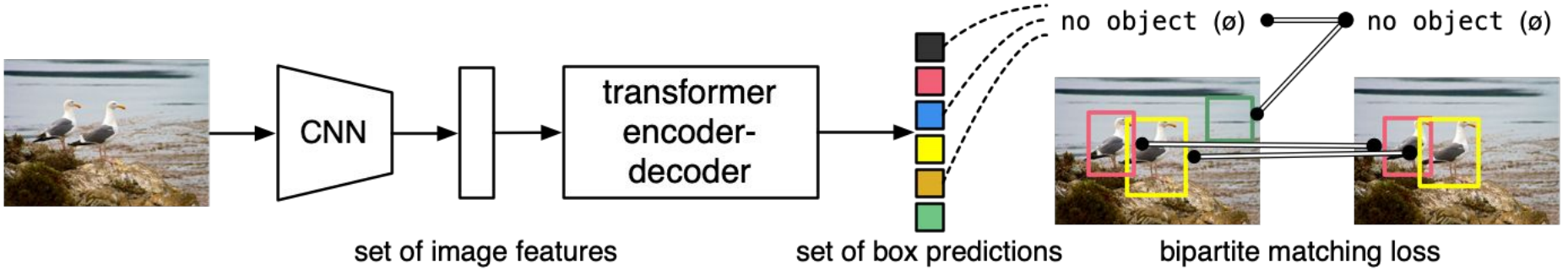

wav2vec2.0 learning the structure of speech from raw audio

# Time-series forecasting




[Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting]

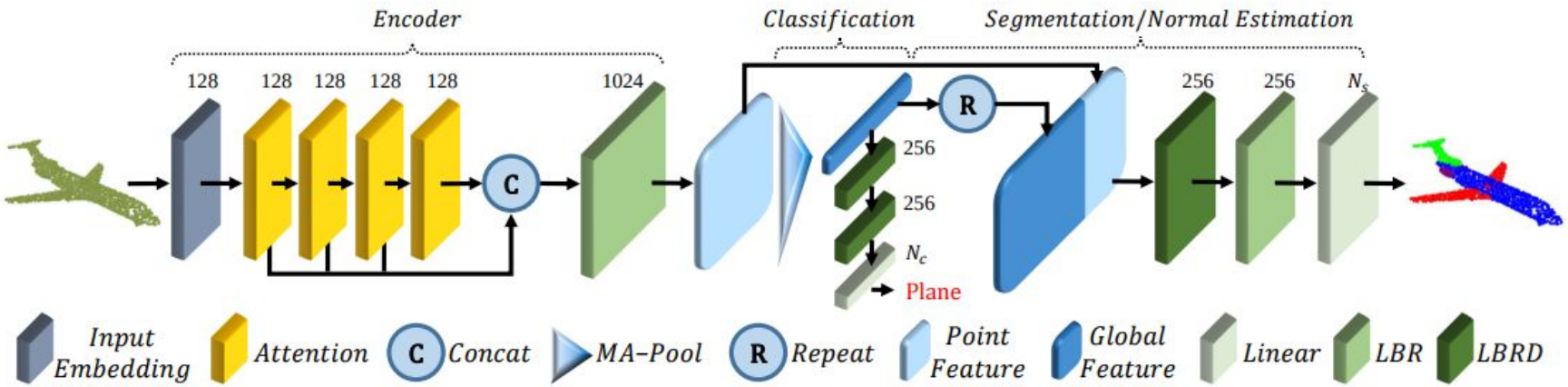
# Computer Vision




End to End Object Detection with Transformers

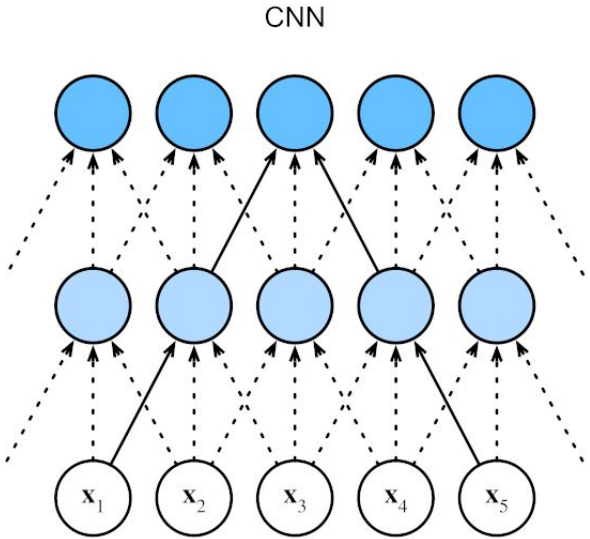
# Any set-structured problem

Input is a set of vectors, with an explicit structure (sequence, image) or implicit (list of objects, graph, point cloud)

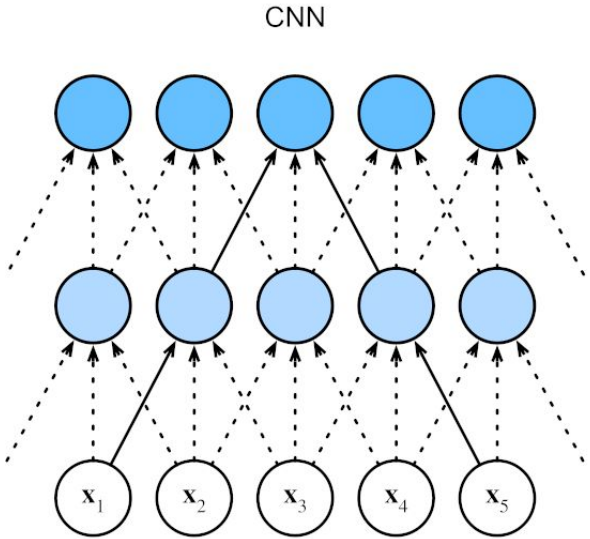


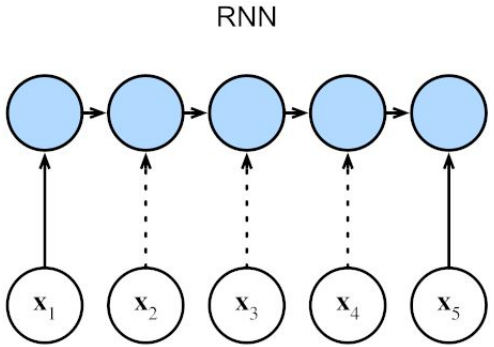

[PCT: Point cloud transformer]

# Main concept: Self-Attention



# Main concept: Self-Attention



# Main concept: Self-Attention