Farielberry Lab

Deep Learning for Genomic Prediction

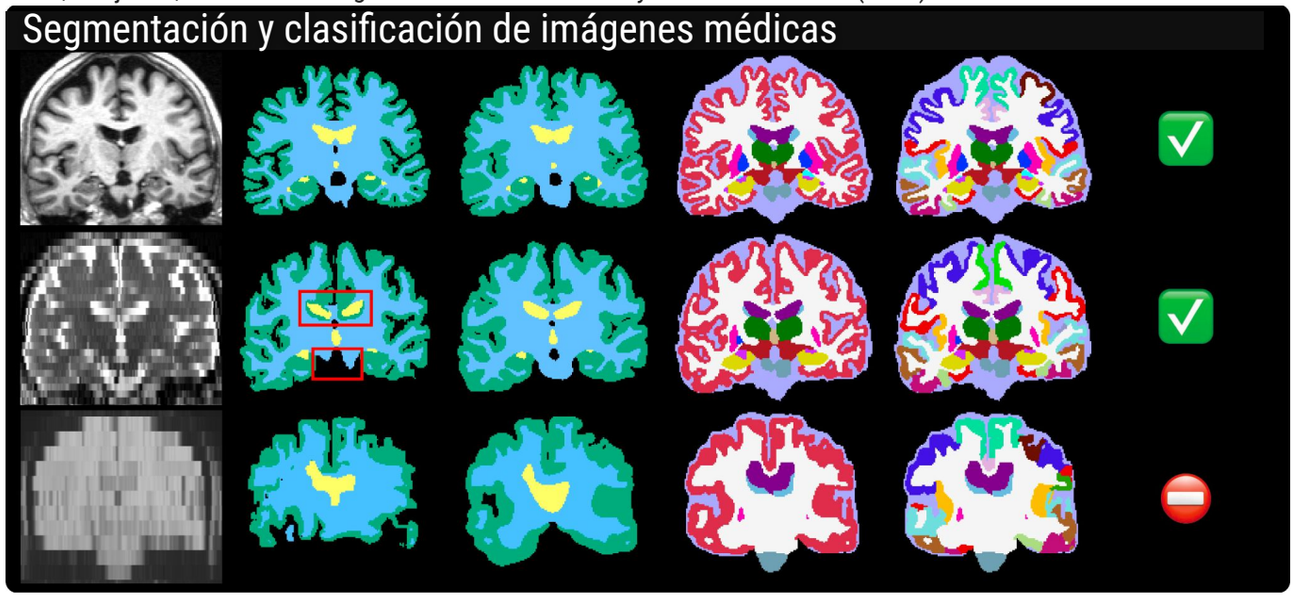
FACULTAD DE INGENIERÍA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Auto-Encoders (AE)
Generative Adversarial Networks (GAN)

Redes neuronales convolucionales

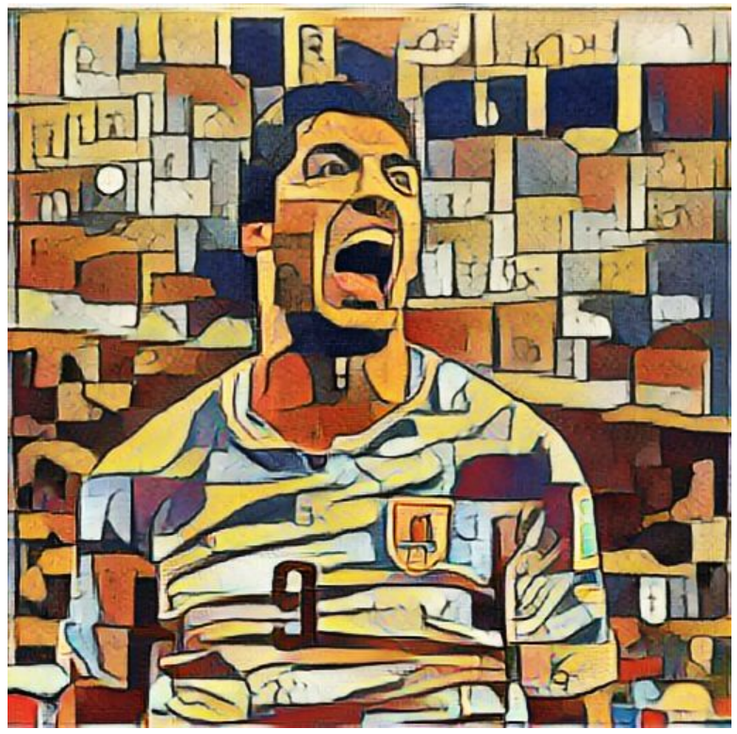
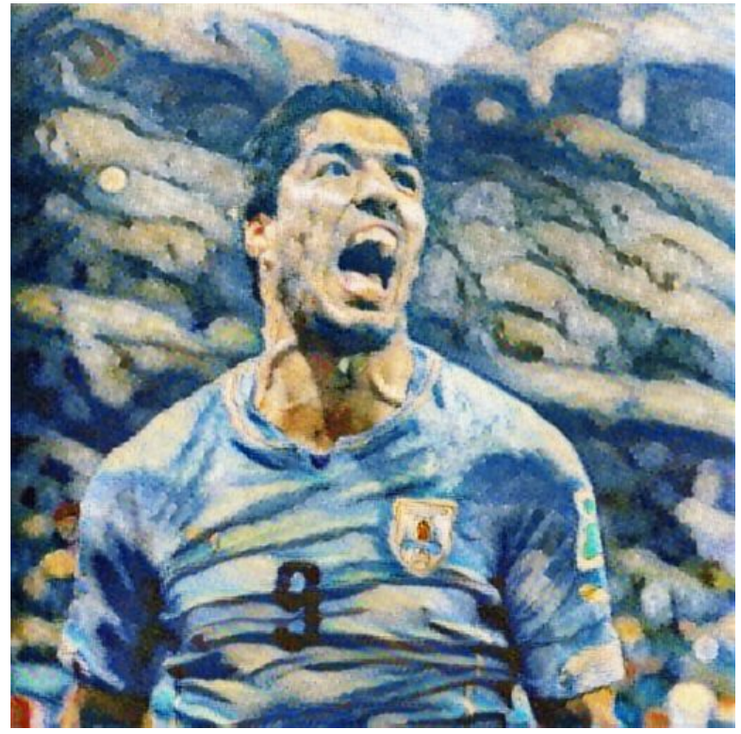
Billot, Benjamin, et al. *Proceedings of the National Academy of Sciences* 120.9 (2023)



Istanbul Traffic - YOLO COCO - Object Detection (youtube.com)

Artistic Style Transfer with Convolutional Neural Network, Manjeet Singh (medium.com)

Redes neuronales convolucionales

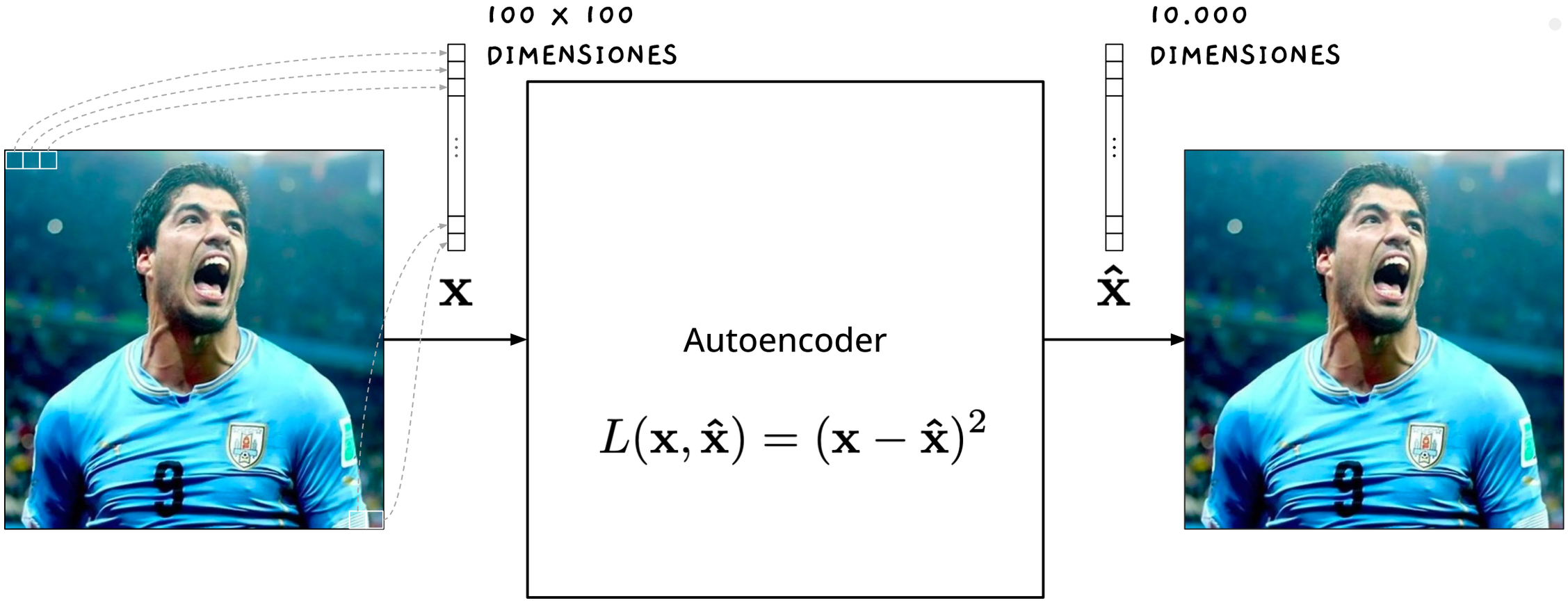





Hugging Face

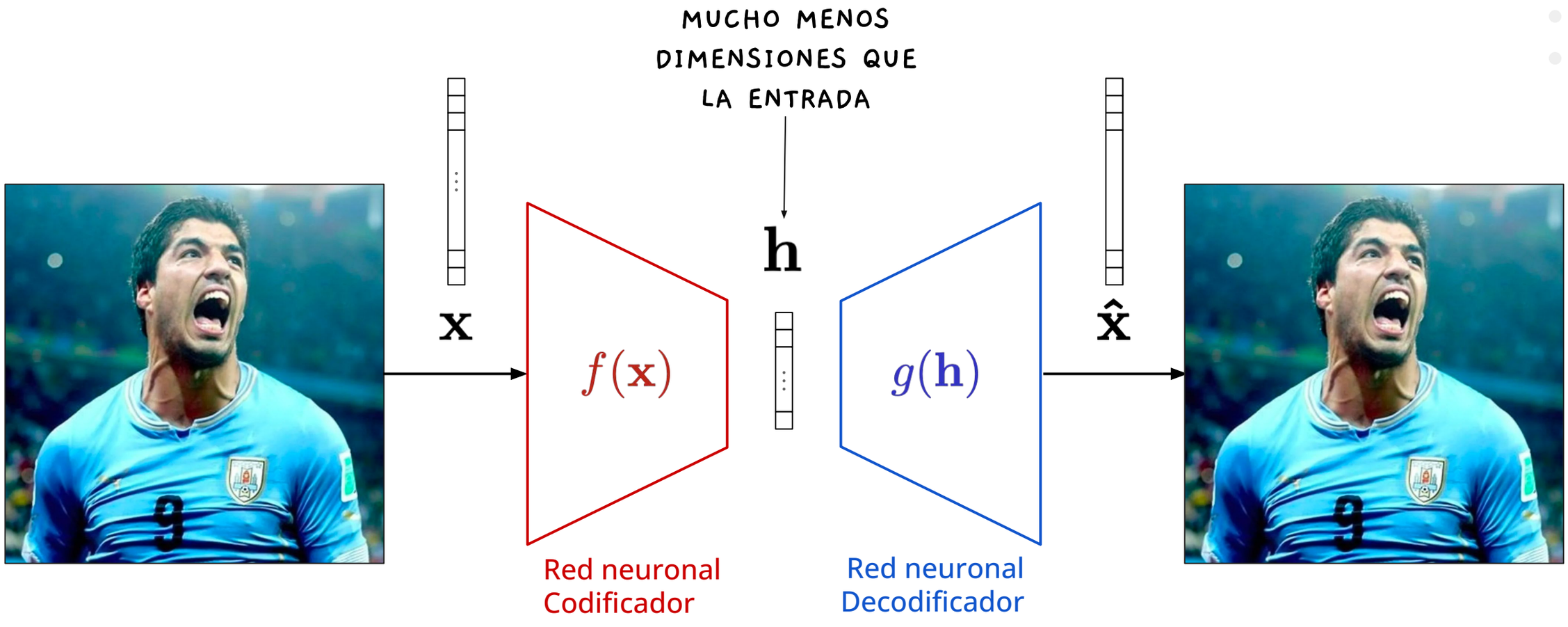
Autoencoders (AE)

'propio', 'de uno mismo' (autobiografía)

codificación: nueva representación eficiente

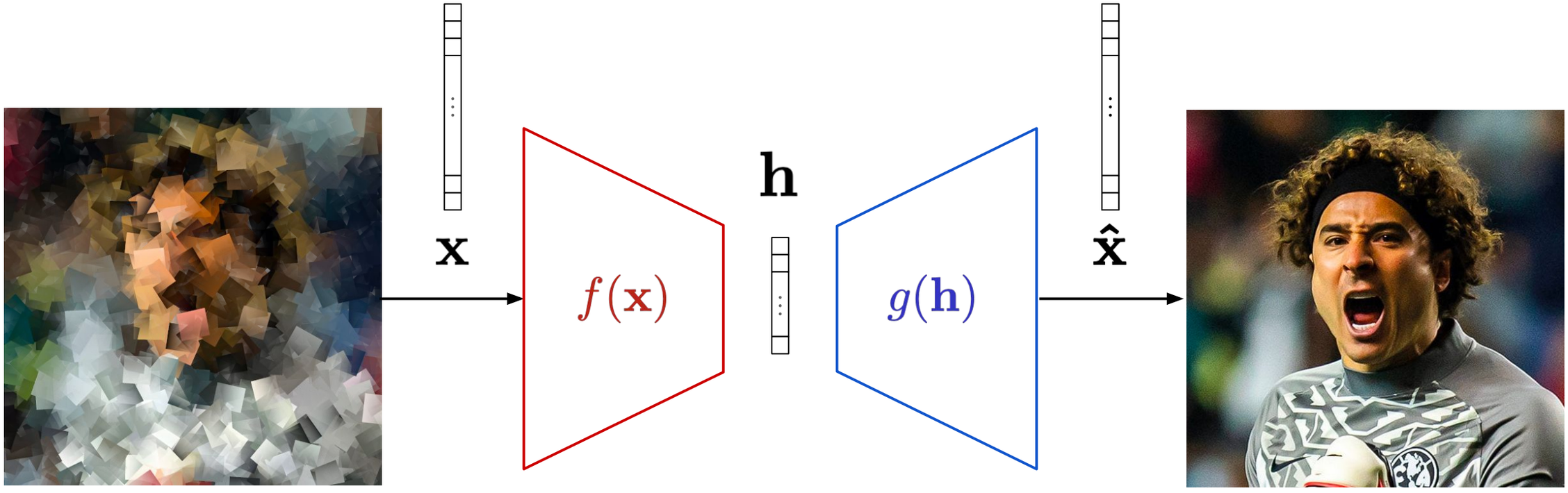


Autoencoders (AE)



$\mathbf{h}$ es un *resumen* de $\mathbf{x}$, describe lo *esencial* de la entrada dejando de lado *pequeñas perturbaciones*

Autoencoders (AE)



Autoencoders (AE)

Colorear imágenes

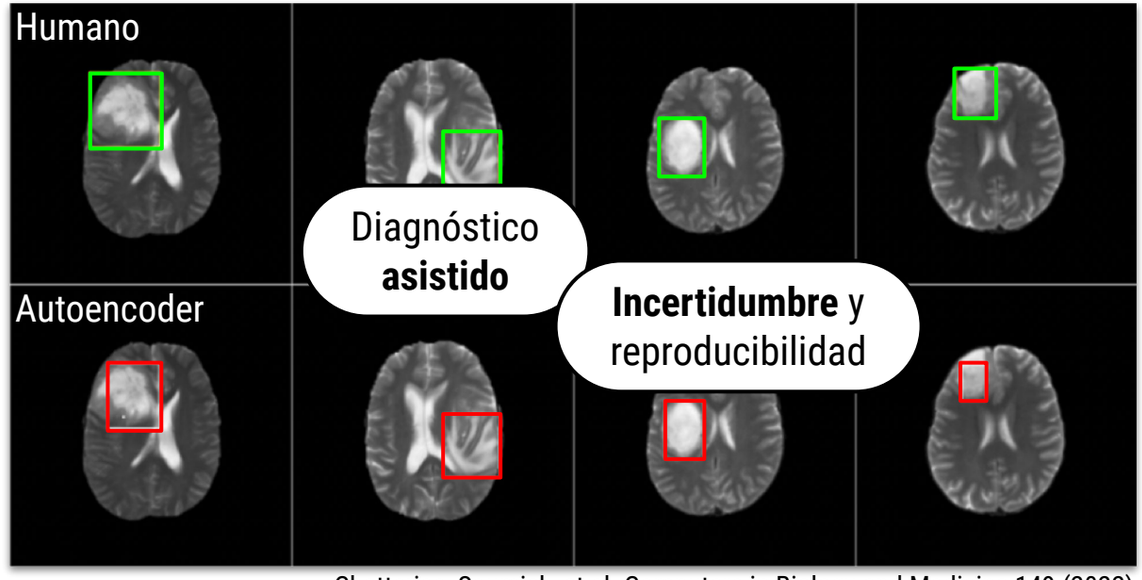
becominghuman.ai

Corregir iluminación

Dias Da Cruz, Steve, et al. *Sensors* 22.13 (2022)

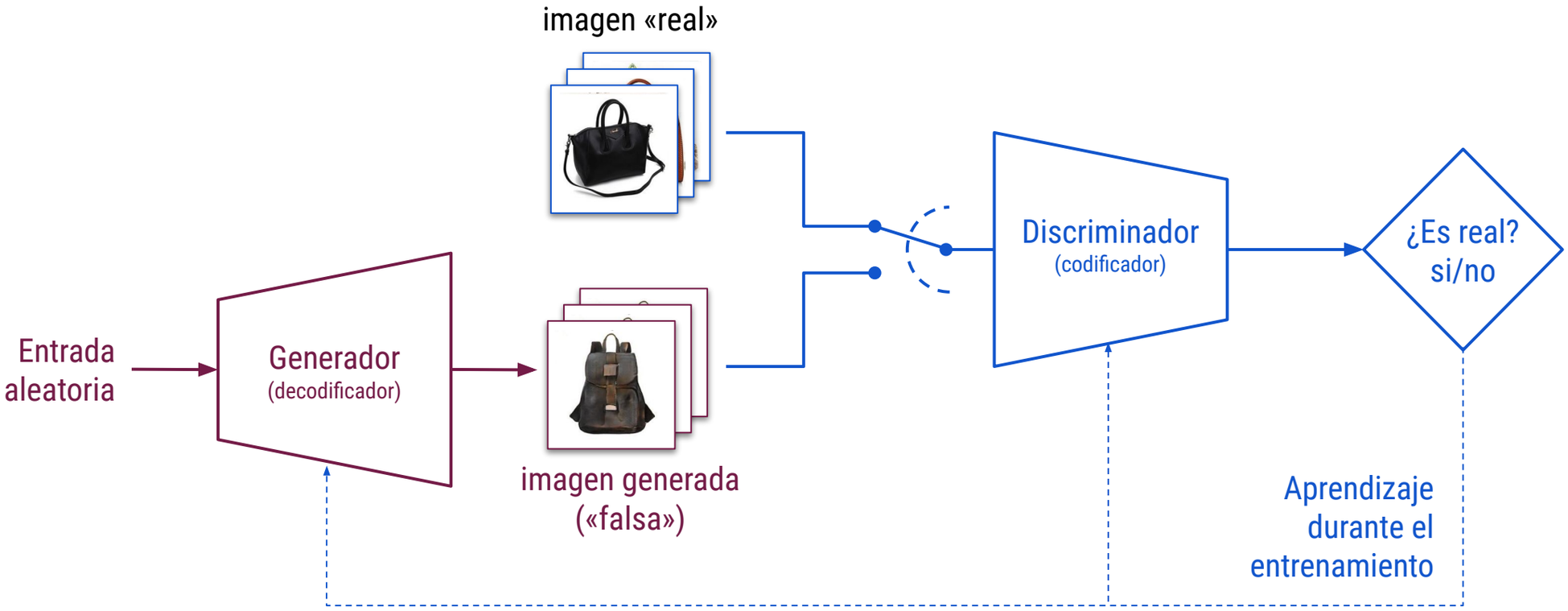
- Compresión de imágenes
- Eliminación de ruido en imágenes
- Sistema de recomendación
- Imputación de valores perdidos
- Generación de datos
- ...

Detectar *anomalías* en imágenes médicas



Chatterjee, Soumick, et al. Computers in Biology and Medicine 149 (2022)

Redes generativas adversarias (GAN)



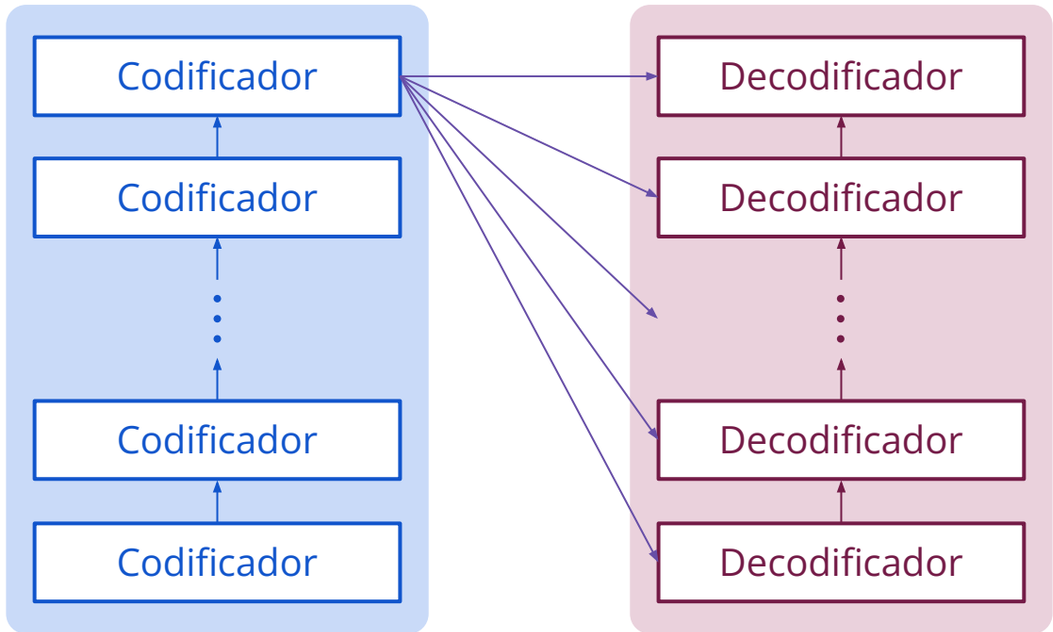
- La idea clave: la competencia mejora el desempeño.

Redes generativas adversarias (GAN)

Transformers

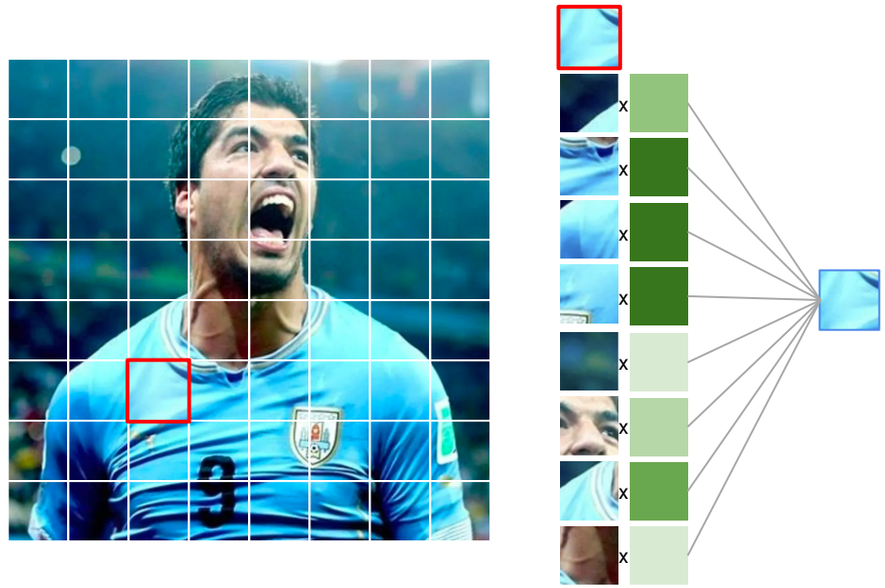
ChatGPT ⇒ **G**enerative **P**re-trained **T**ransformer

Arquitectura codificador-decodificador

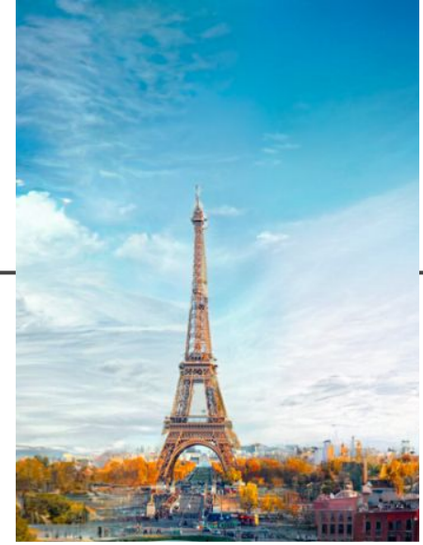


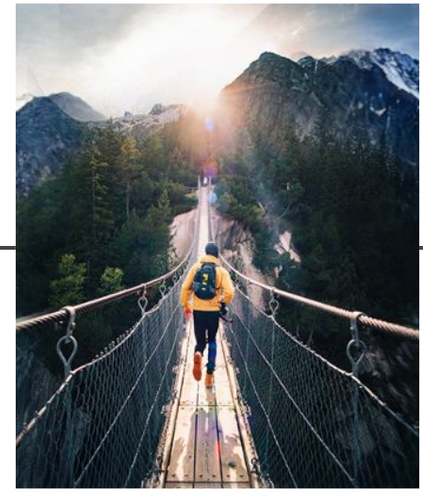
Auto-atención

Representar una parte de *algo* como una combinación de sus otras partes dada la *relevancia* entre ellas.



Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017) [~85.000 citas]





?

?

?

?

?

?

Queremos generar (sintetizar) la imagen de entrada (consulta) hacia los lados.

Chang, Huiwen, et al. "Maskgit: Masked generative image transformer." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

Transformers

¿Son «respuestas correctas»?

Chang, Huiwen, et al. "Maskgit: Masked generative image transformer." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

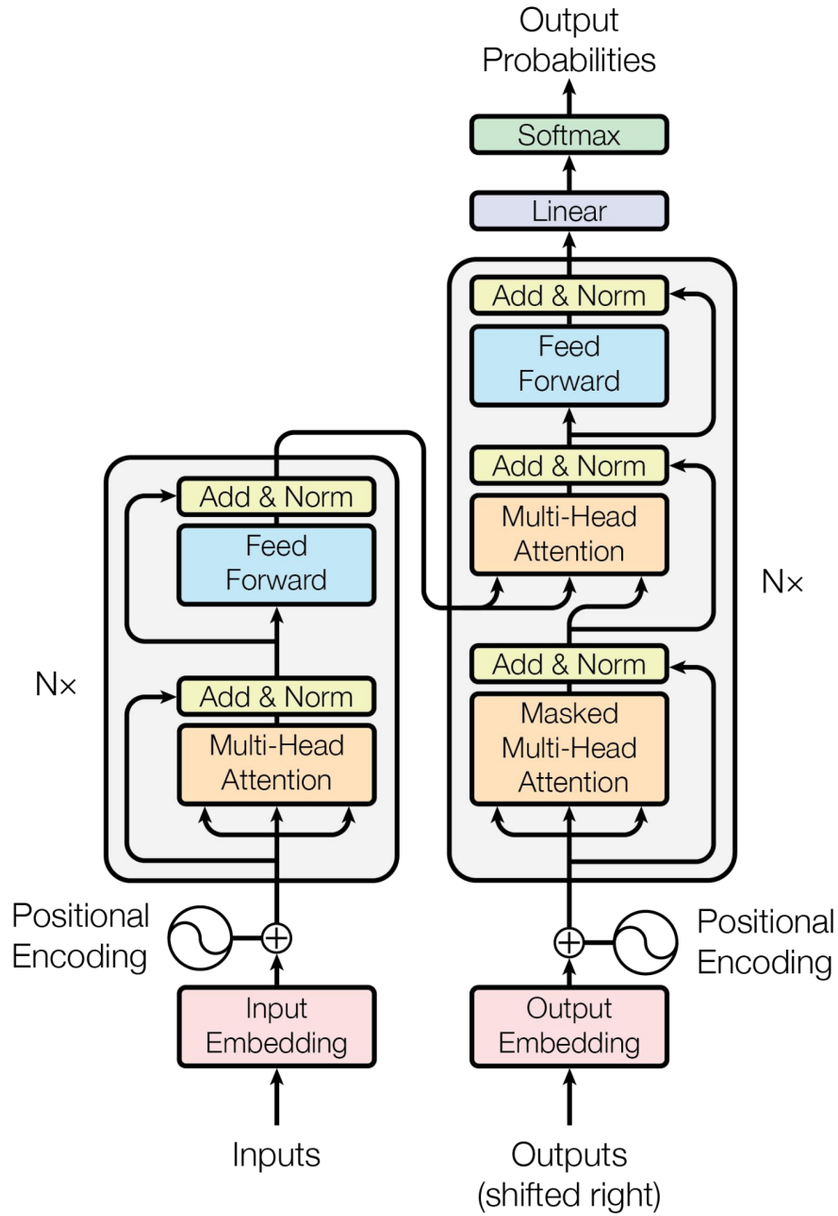
Transformers

Transformer model

- Can draw global dependencies between sequences.
- Originally created for Natural Language Processing (NLP).
- Bidirectional structure: Encoder - Decoder.
- Core block: Attention module.
- Some known applications:

175 billion parameters

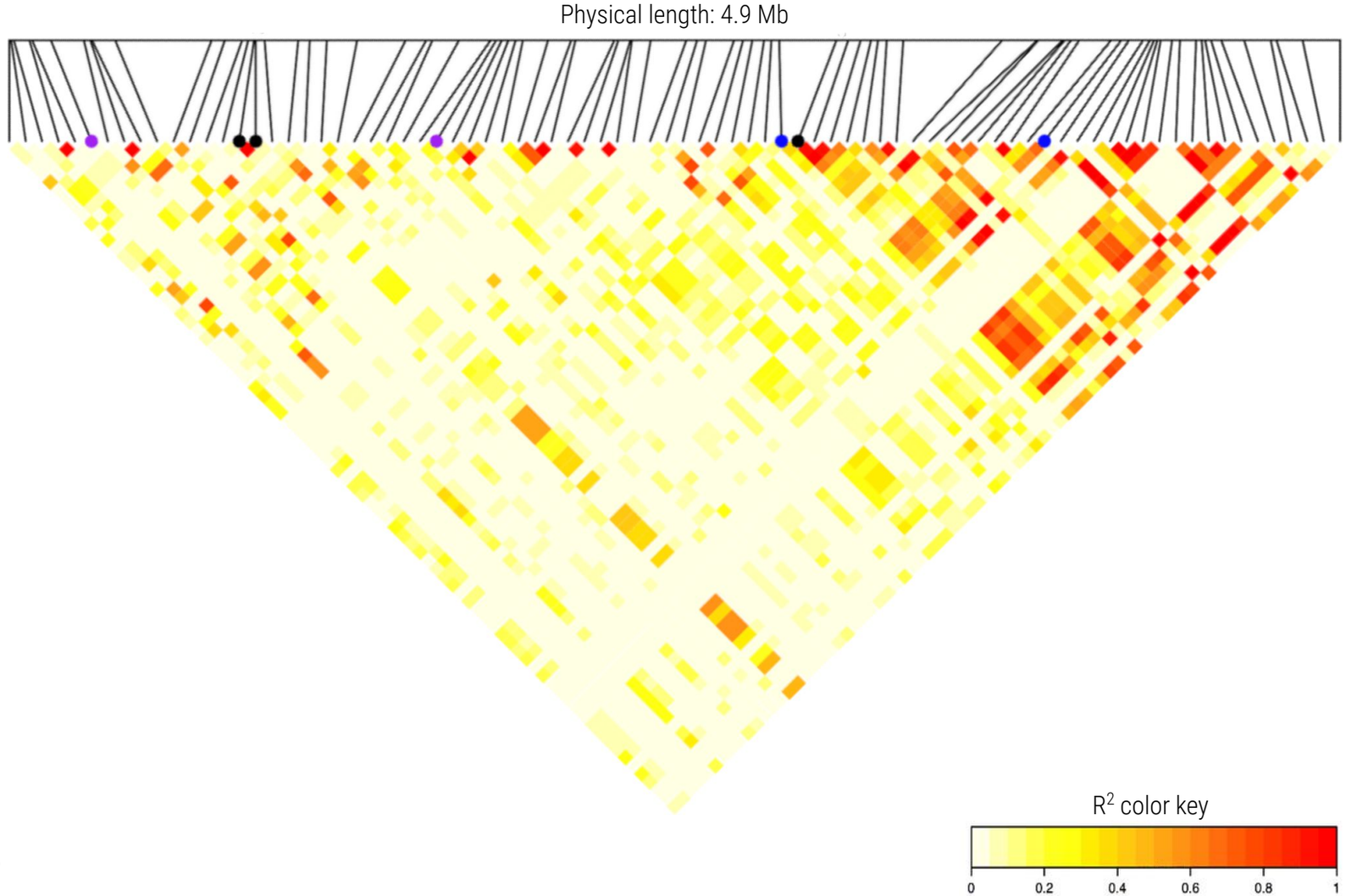
3.5 billion parameters



Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).

Why using Transformers in Genomic Prediction?

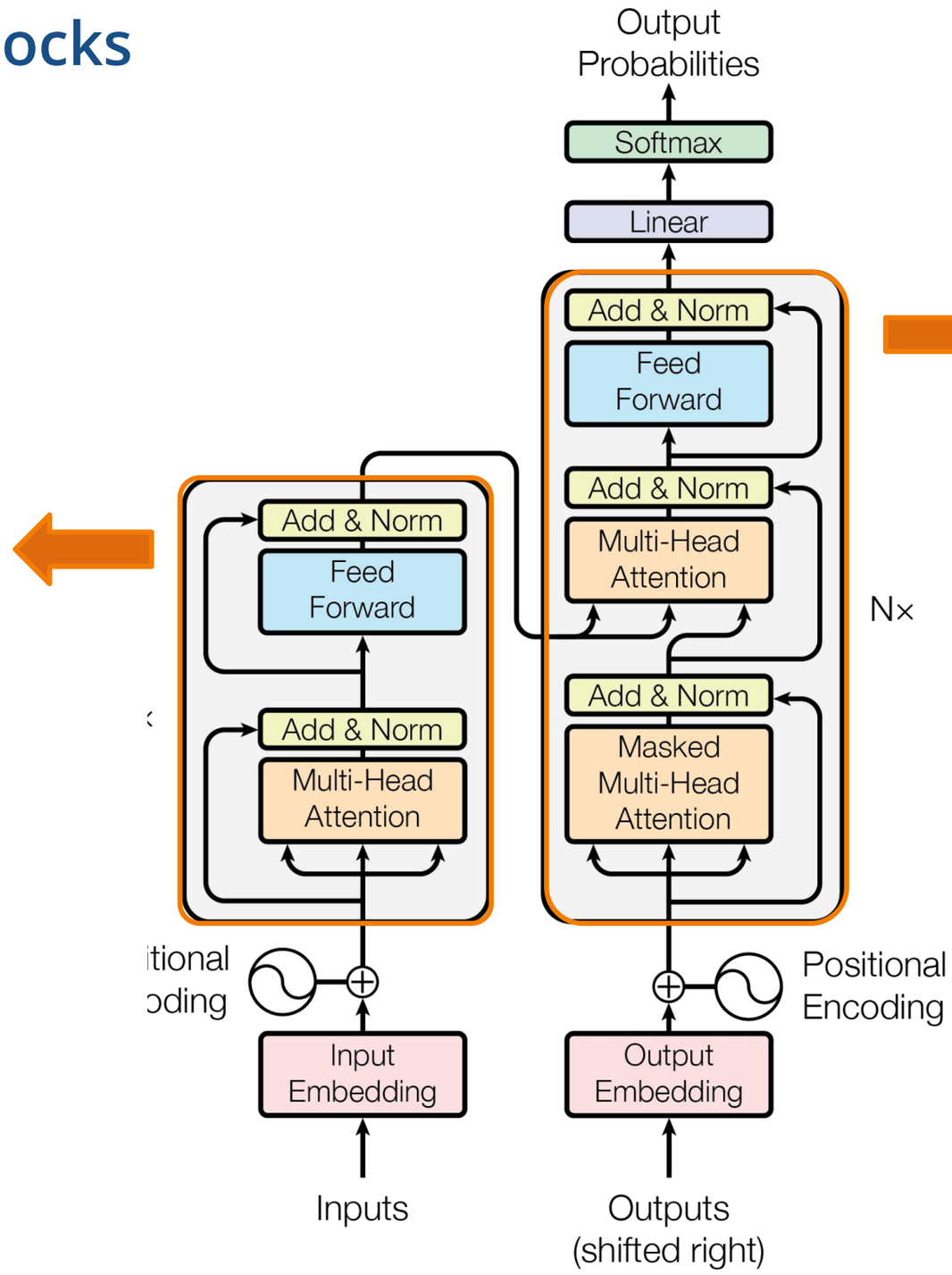
Pairwise Linkage Disequilibrium:



Chen et al. (2017). Genetics, 206(4), 1791-1806.

Transformer's blocks

Encoder: extract semantics or representations from the data.

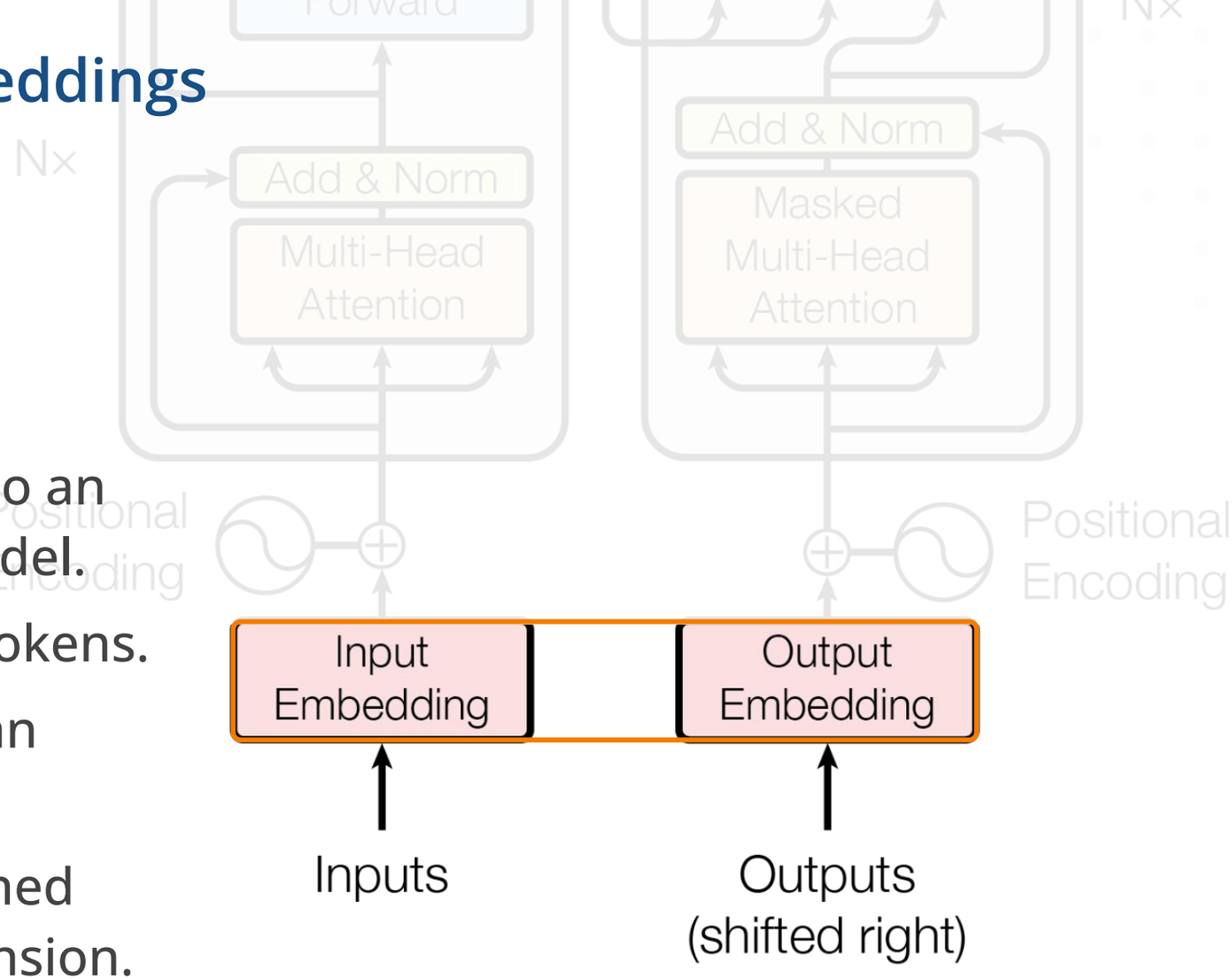


Decoder: used for generating new data (i.e.: create new sentences, translating the input, etc.)

Transformer's blocks: Embeddings

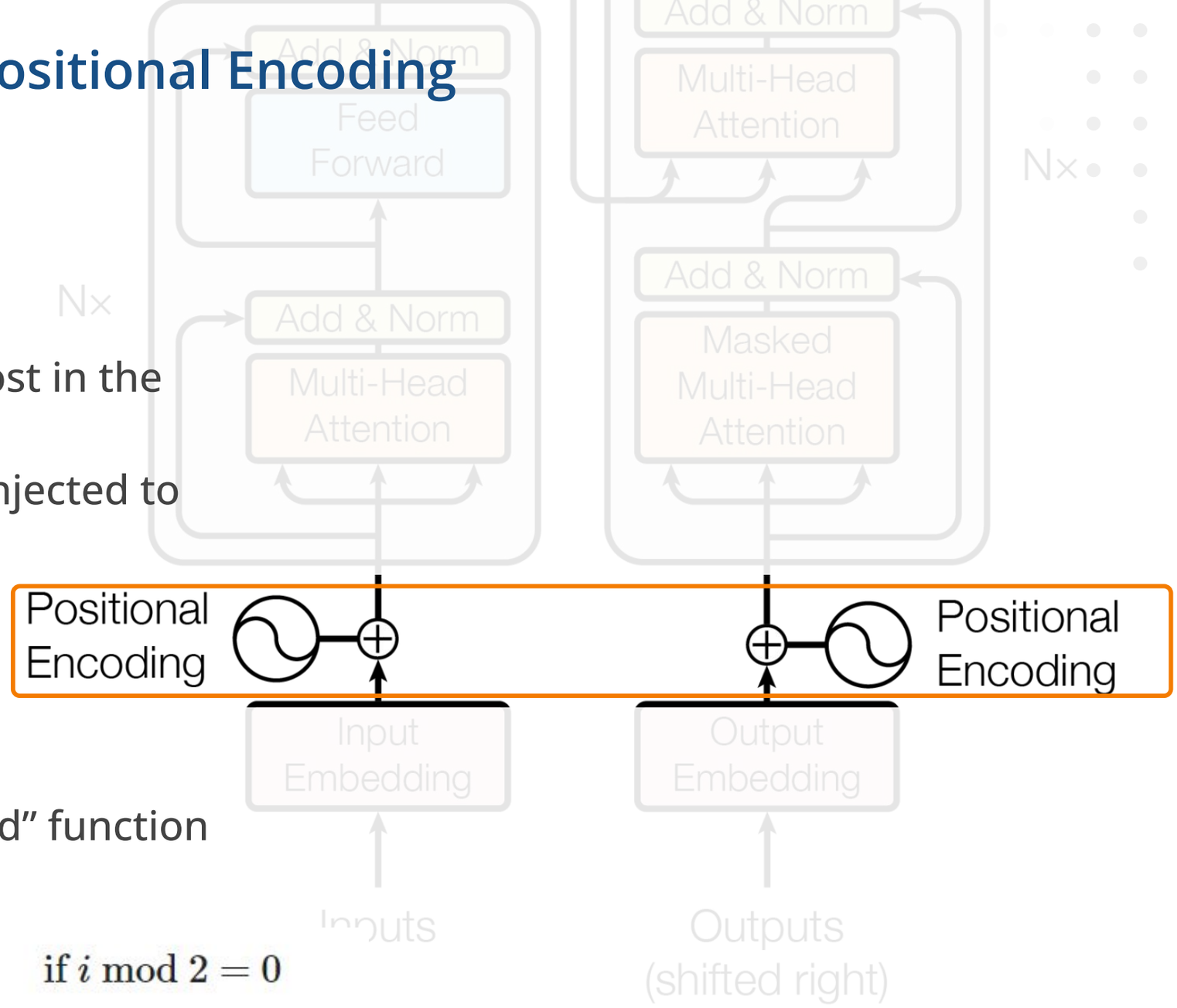
Transforms the input data into an acceptable format for the model.

- Each word is turned into tokens.
- Each token is mapped to an index of a look up table.
- Each entry maps to a learned embedding of d_{model} dimension.



Transformer's blocks: Positional Encoding

- The order of tokens is lost in the calculations.
- The sequence order is injected to each token.

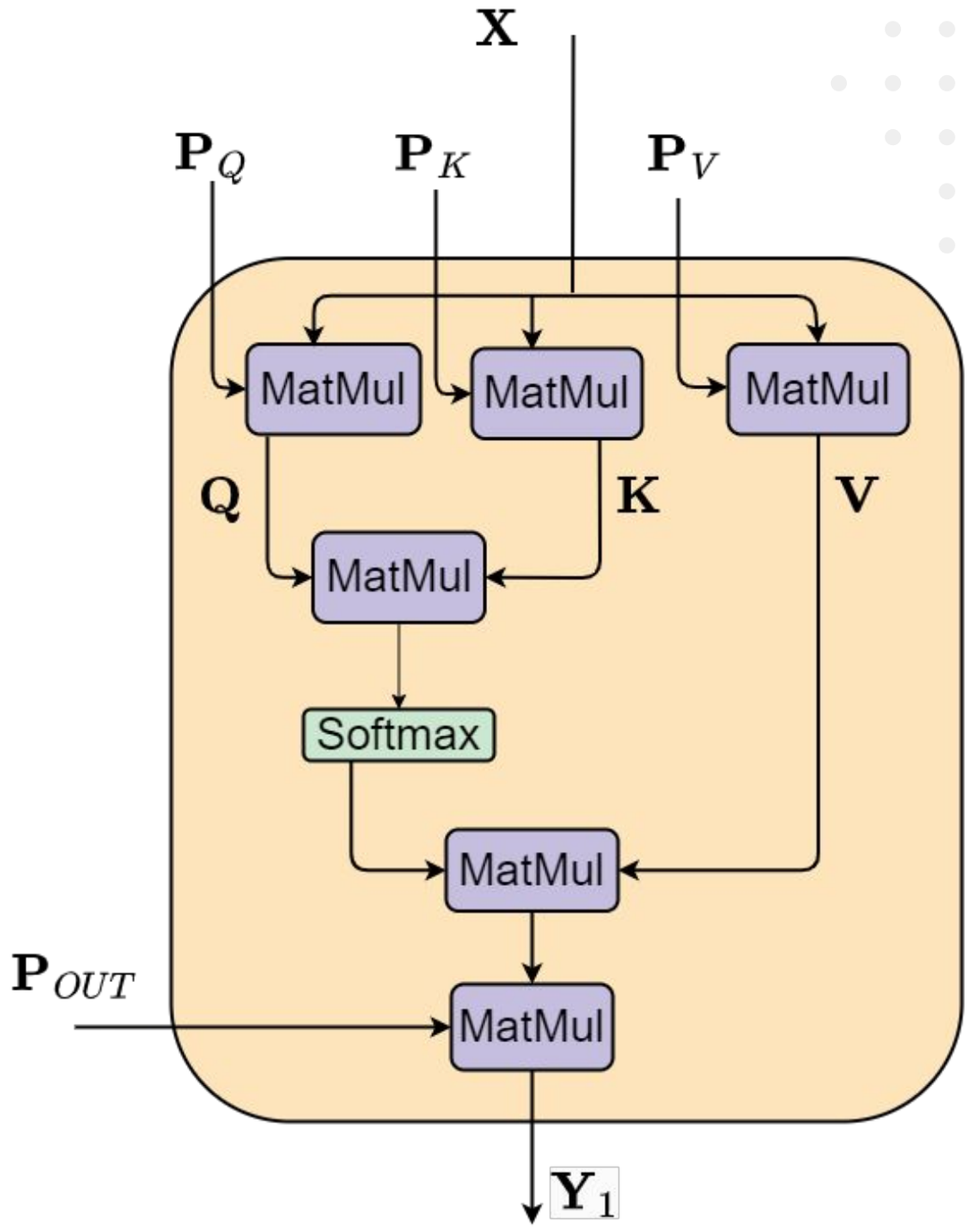


- "Attention is all you need" function for Positional Encoding:

$$P_D^i = \begin{cases} \sin\left(\frac{D}{10000^{i/dm}}\right), & \text{if } i \bmod 2 = 0 \\ \cos\left(\frac{D}{10000^{((i-1)/dm}}\right), & \text{otherwise} \end{cases}$$

Self-Attention

- Attention allows models to "focus" on different parts of an input; while considering the entire context.
- Self-attention is an attention mechanism where each vector of a given input sequence attends to the entire sequence.
- Three important matrices:
  - Query: asking for information.
  - Key: saying we have the information.
  - Value: getting the information.



Self-Attention

Compute Query, Key and Value:

d_{model}

$\mathbf{P}_Q =$

2	0	1
0	1	1

$\mathbf{P}_K =$

0	1	1
2	0	0

$\mathbf{P}_V =$

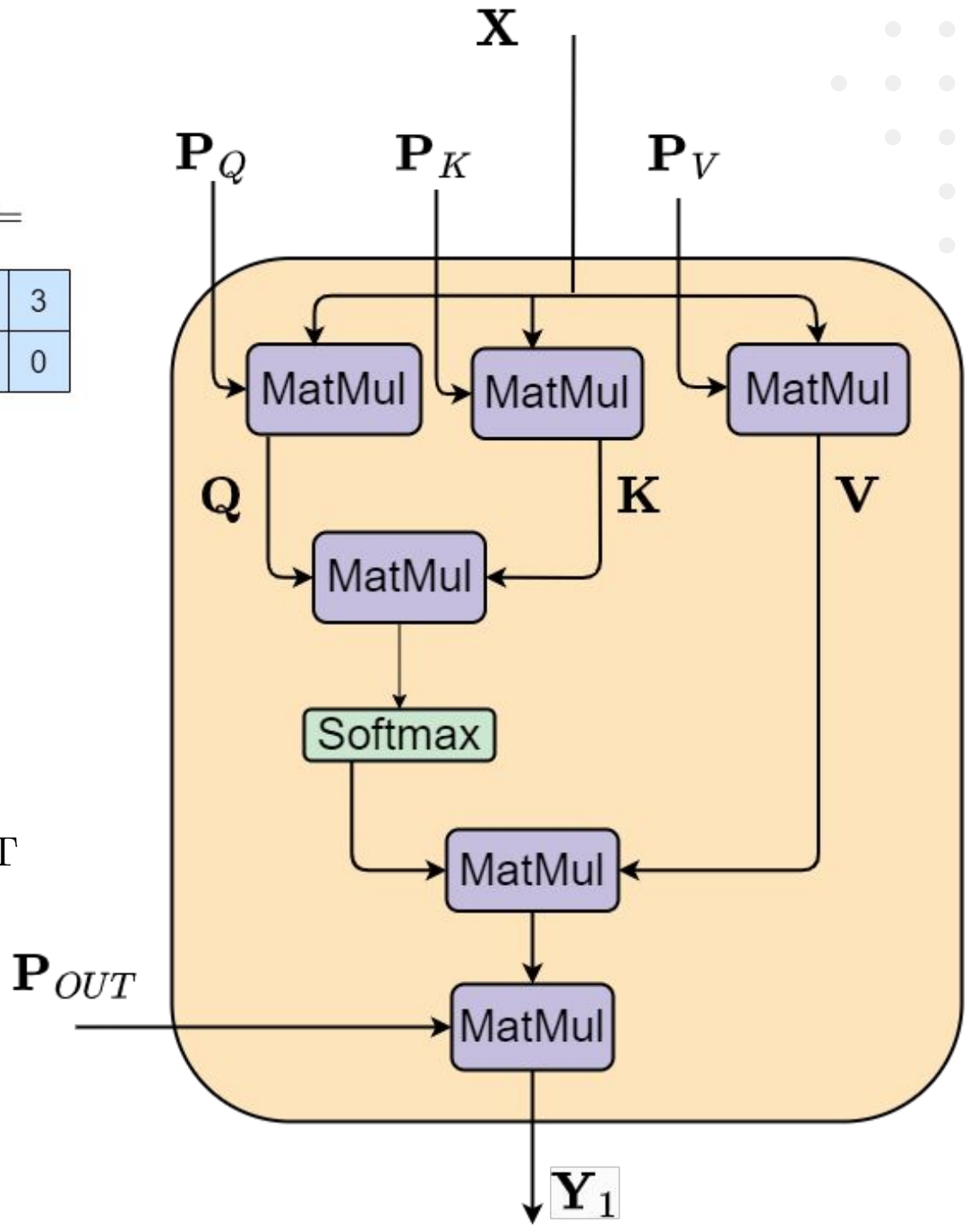
1	1	3
0	1	0

$$Q = P_Q X$$
$$K = P_K X$$
$$V = P_V X$$

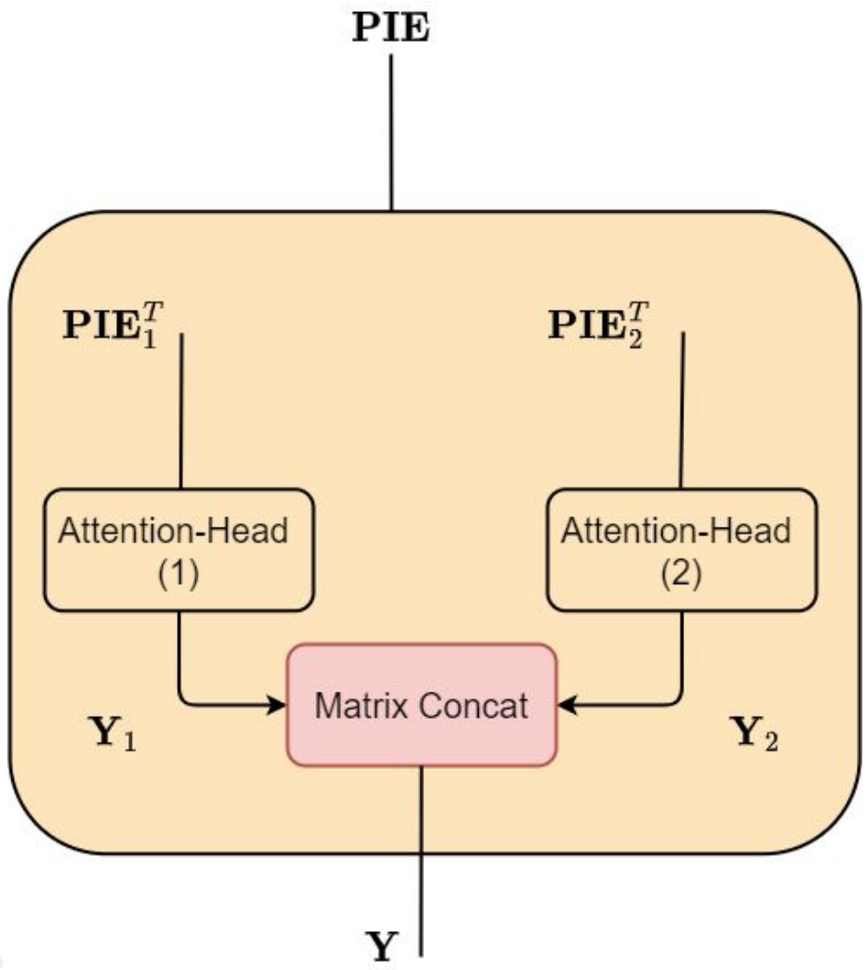
$$\mathbf{Y} = \text{Softmax}\left(\frac{\mathbf{Q} \times \mathbf{K}}{\sqrt{k}}\right) \times \mathbf{V} \times \mathbf{P}_{\text{OUT}}$$

Attention Score

The information we keep



Multi-Head Self-Attention



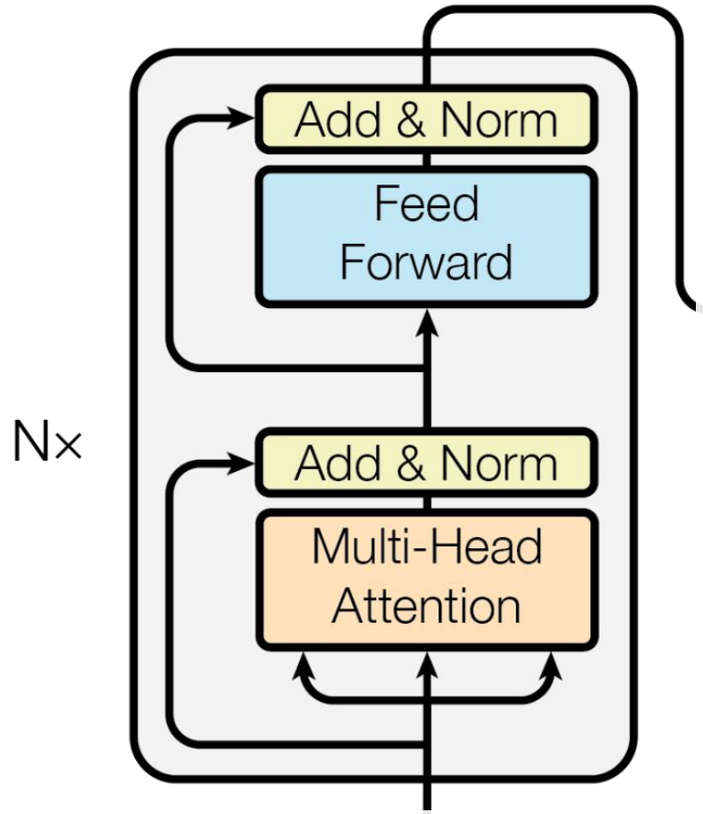
- Parallelization to optimize computation
- Separates each vector in k smaller vectors with dimension d_{model}/k
  k: number of heads
- Outputs the concatenation of the attention heads' outputs.

Transformer's blocks: Encoder

- Multi-Head Self-Attention to capture the dependencies of the input sequence with itself.
- Feed-Forward: then projecting these new token representations to a space that the next block can use more optimally

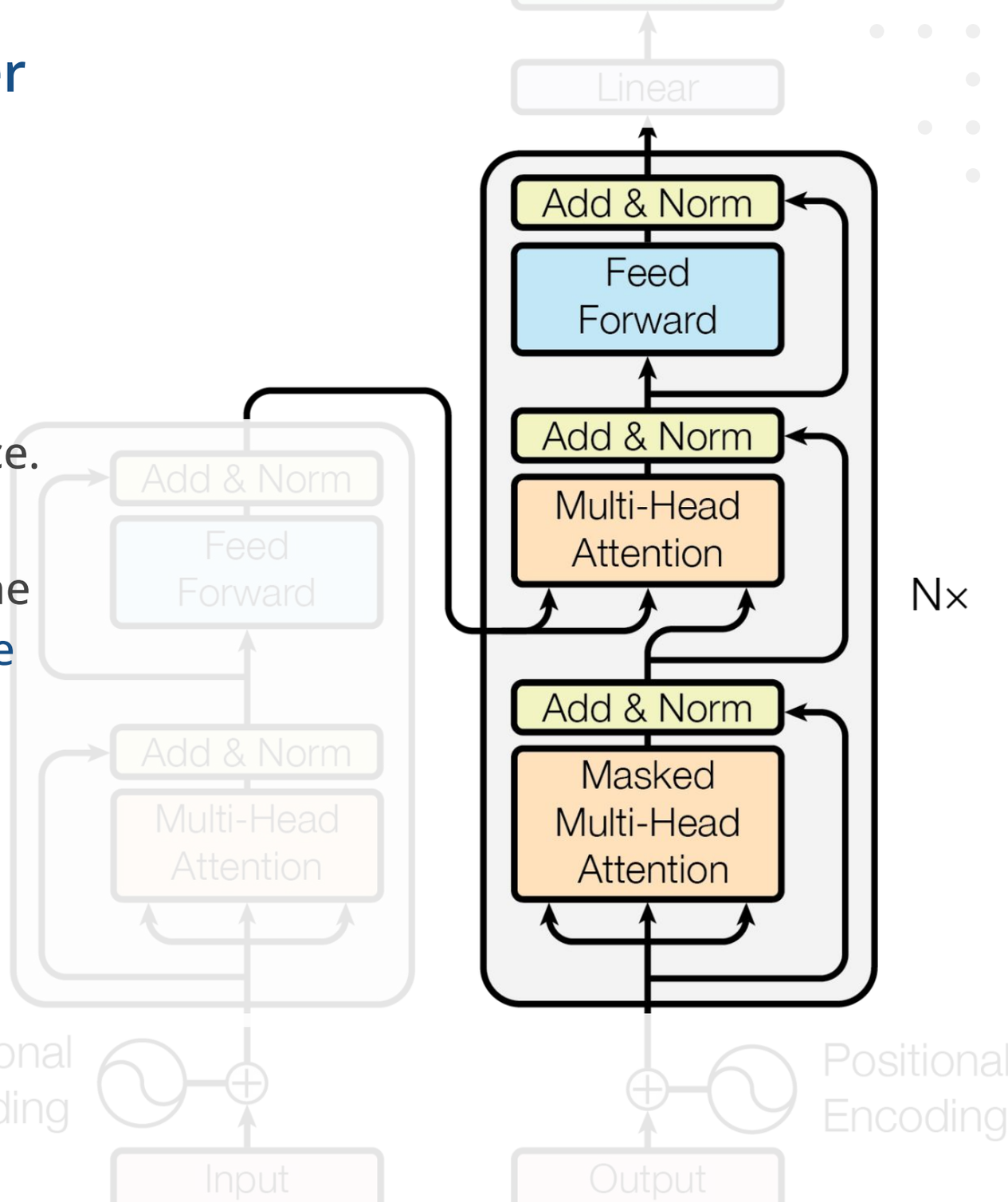
$$\mathrm{FFN}(x) = \max\left(0, xW_1 + b_1\right) W_2 + b_2$$

- The Add&Norm layer is a residual connection to prevent the vanishing gradients problem.



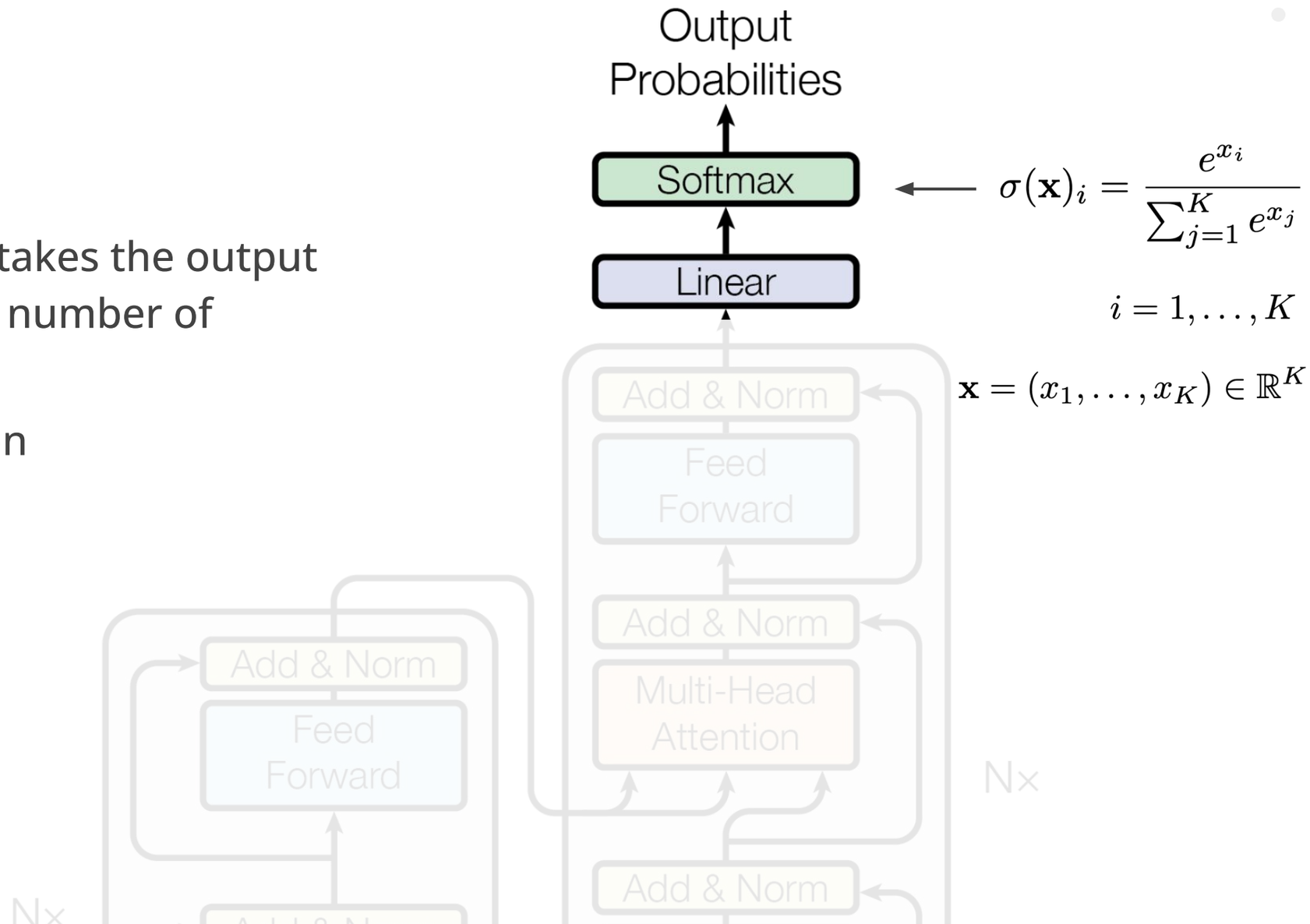
Transformer's blocks: Decoder

- Masked Multi-Head Attention: masks information of future positions of the input sequence.
- Cross Multi-Head Attention: incorporates information of the encoder output (Key and Value matrix) and the masked multi-head attention output (Query matrix).
- FFN and Add&Norm work as in the Encoder.



Transformer's blocks: Output Layer

- Linear layer that takes the output dimension to the number of possible outputs.
- Softmax activation
- MAP estimation

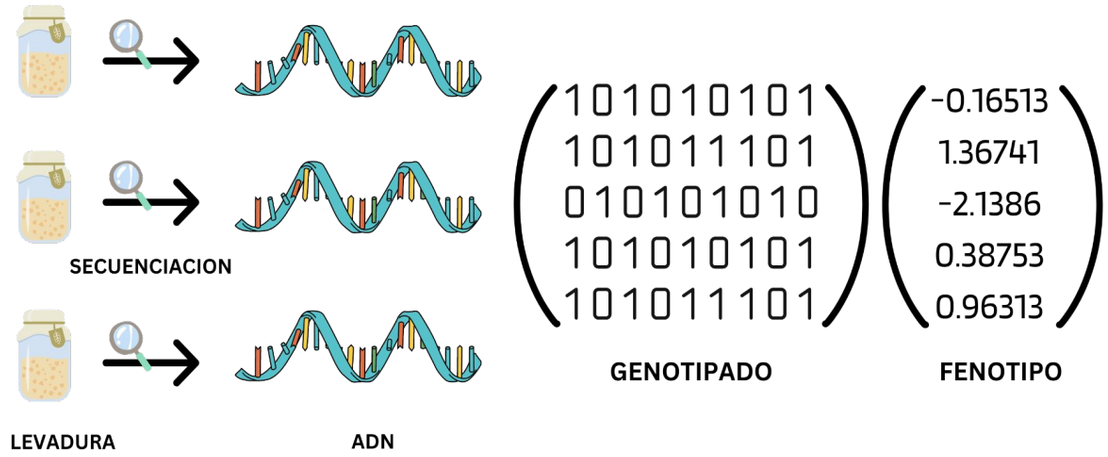


Practical

Yeast database:

- Phenotype: yeast growth in 48 different environments
- 1008 samples with 11623 SNPs

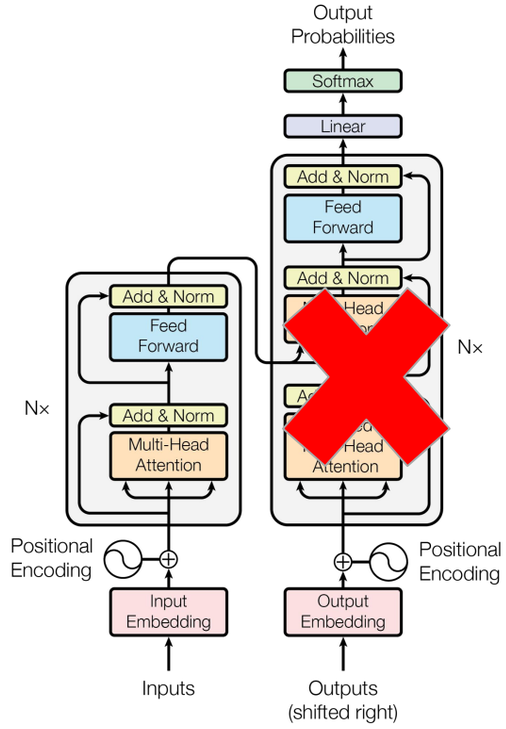
- Working with two environments: Lactate and Lactose.
- Single-trait: predict one phenotype per model
- Multi-trait: predict both phenotypes training one single model

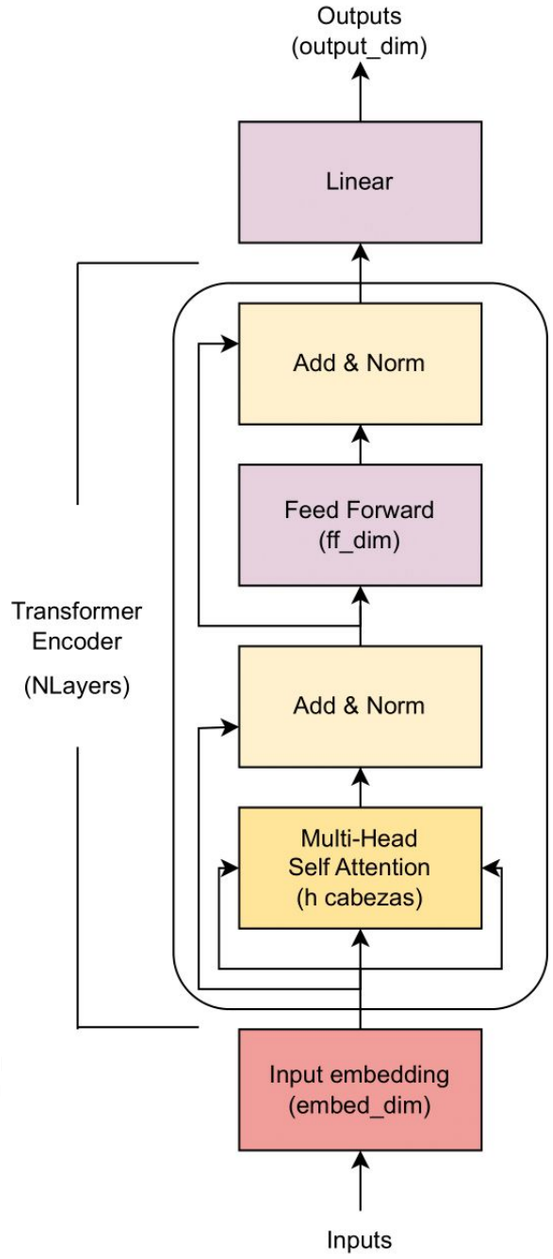


Practical

- GPTransformer (Jubair et al.,2021)
  - Barley

- No Decoder, we want to learn the semantics of the data.
- Linear layer substitutes the Embedding and Positional Encoding layers
- Linear output





Practical

- Input: batches of 8 individuals.
- Metrics:
  - MSE
  - Pearson Correlation Coefficient
- Output dimension 1 or 2 depending on single-trait o multi-trait prediction.