IANIS CLAVIER ET QUENTIN VICTOR

Creating a digital patient using deep learning

LE SIMU

https://lesimudenantes.univ-nantes.fr/formation-continue

- Laboratoire Expérimental de Simulation de Médecine Intensive
- High-fidelity simulation
- 6 themes
    - Anaesthesia & Intensive Care
    - General medicine
    - Emergency medicine
    - Neonatology & Paediatrics
    - Obstetrics
    - Odontology

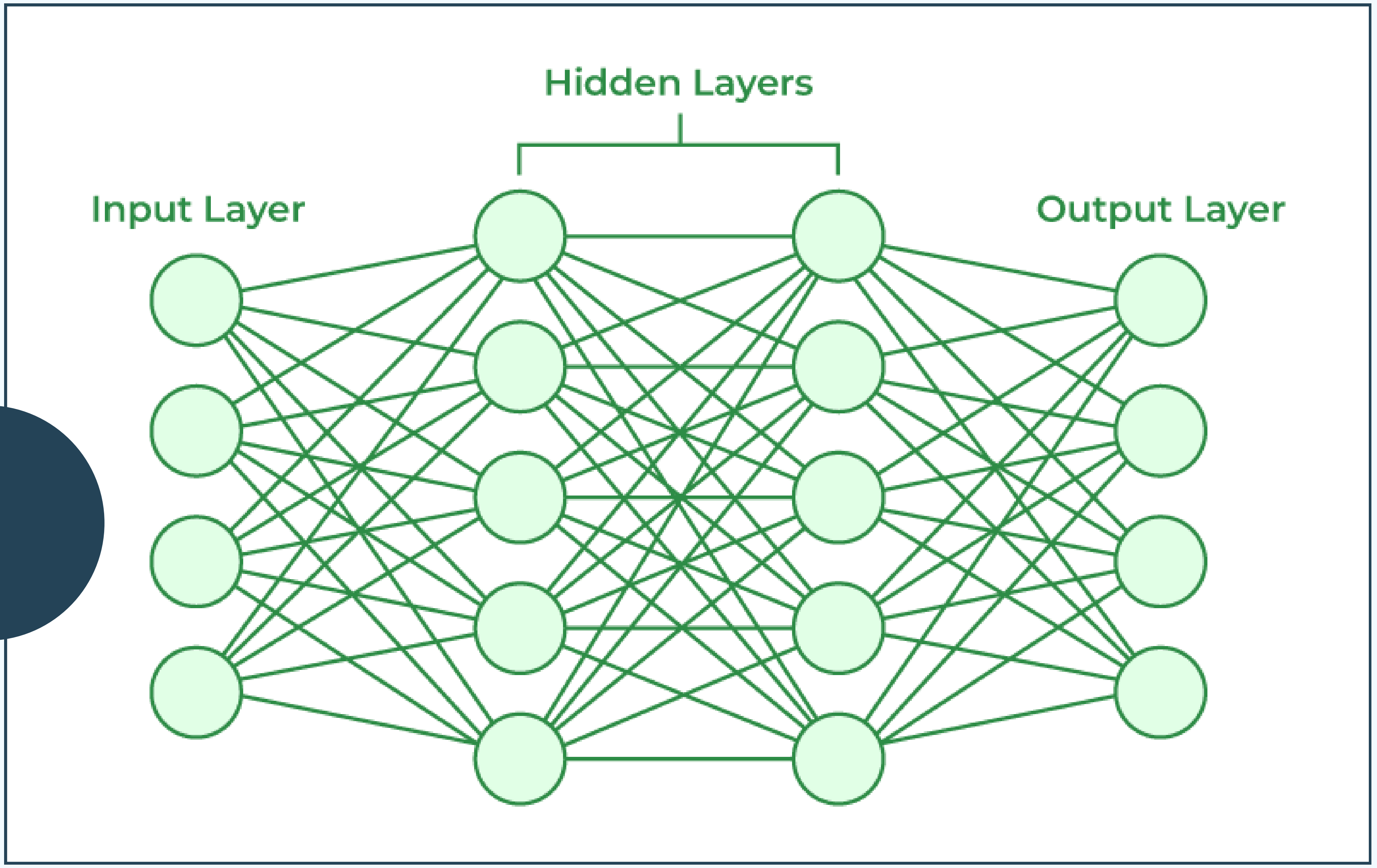
LE SIMU

https://lesimudenantes.univ-nantes.fr/formation-continue

- Pilot
  - Manually changing the physiological variables
  - Not homogeneous
  - Limit in the number of scenarios
- 3 Methods
  - Machine learning
  - Data Mining
  - Deep Learning

DEEP LEARNING

INTRODUCTION



https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

- TRAINING
  - Maximize the likelihood of the predictions
  - Refine the neurons weights
- REQUIRING A LOT OF DATA
  - 1000 patients

DATA

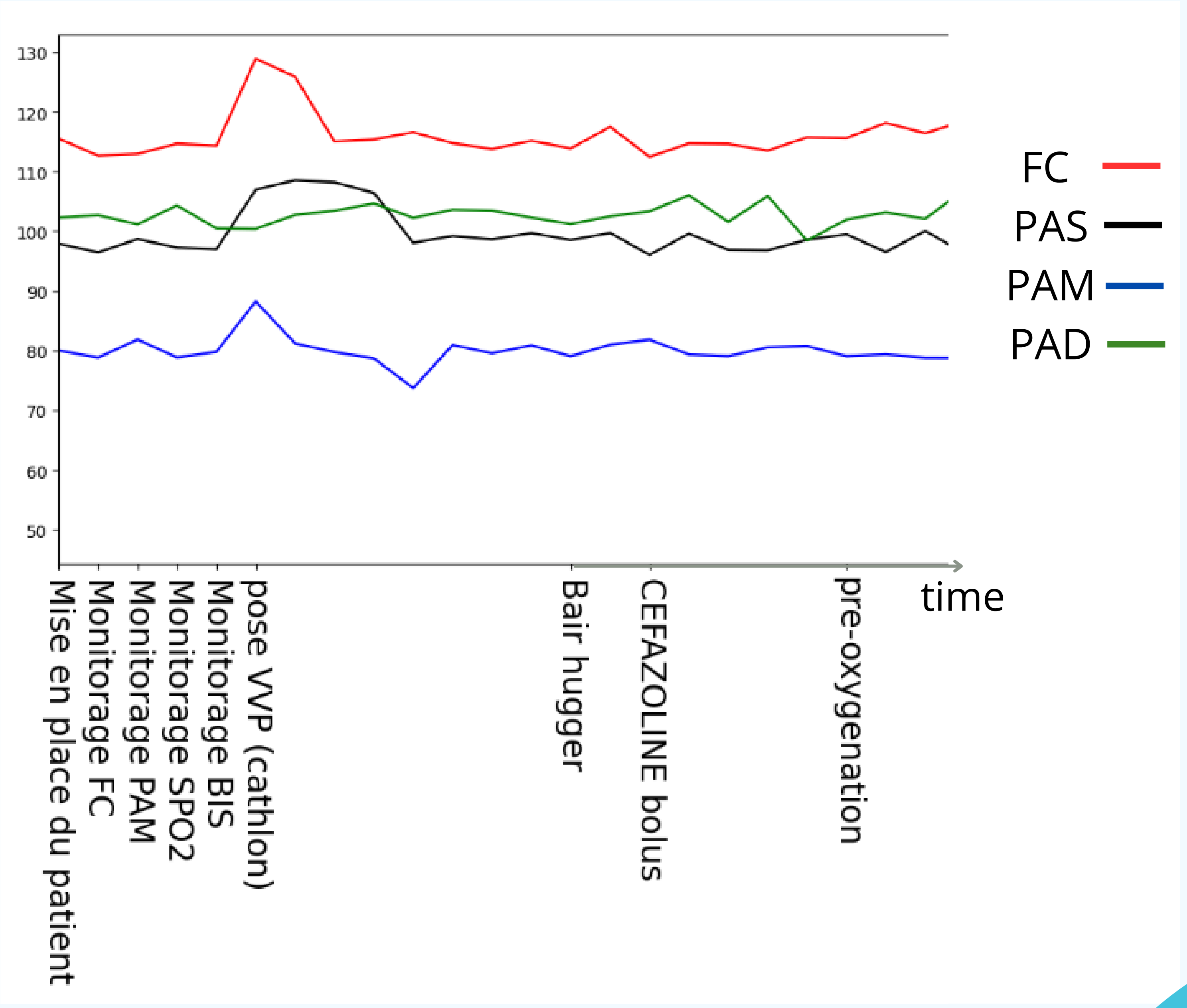
TIME SERIES AND EVENT TRACE

TIME SERIES

- Anaesthetic data
- 30s
- Multvariate
    - FC, PAS, PAD, PAM

EVENT TRACE

- Descriptor
- Drug intake
- Medical procedure



Simulated data(Hugo Boisaubert,thesis,2022)

INTERNSHIP PROGRESS

BIBLIOGRAPHY

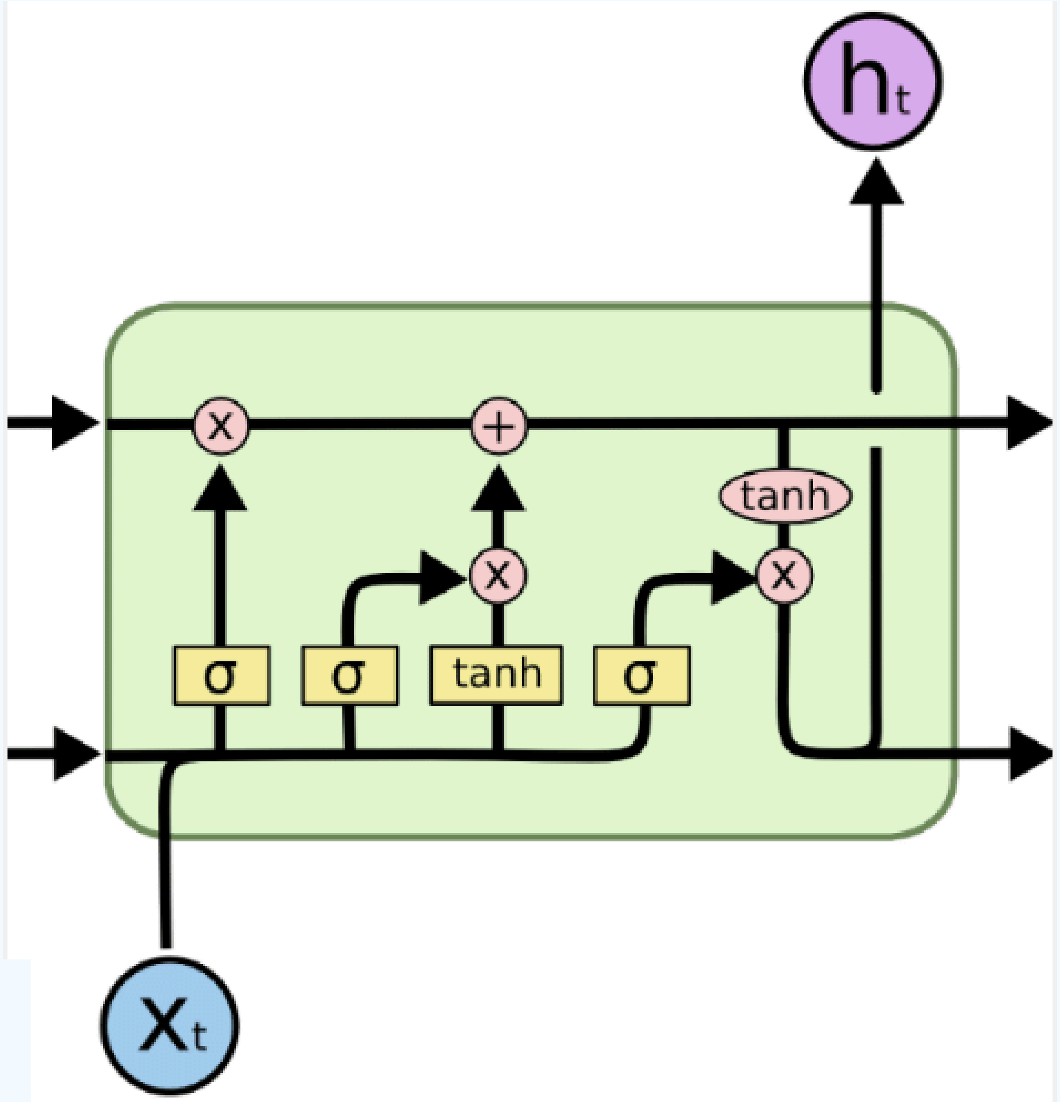
IMPLEMENTATION OF 4 MODELES

COMPARAISON BETWEEN THE DIFFERENTS MODELES

LSTM

Long Short Term Memory

- Takes context into account
- Easy to understand and implement
- Emergence of new, more efficient models



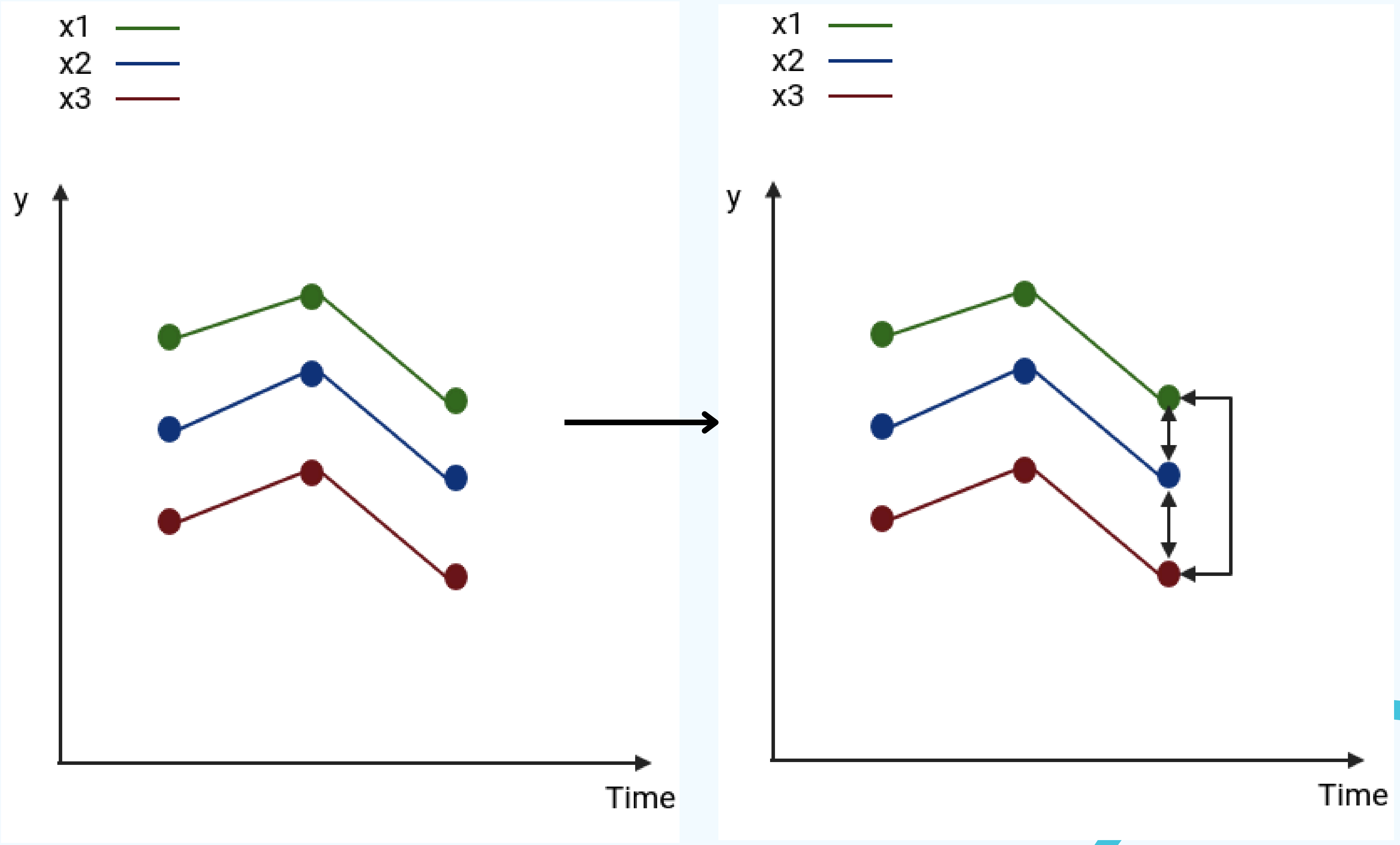
Understanding LSTM Networks, Christopher Olah, 2015

GNN

Graph Neural Network

- Adds relationships between dimensions.
- More complex data pre-processing
- MTGNN[1]

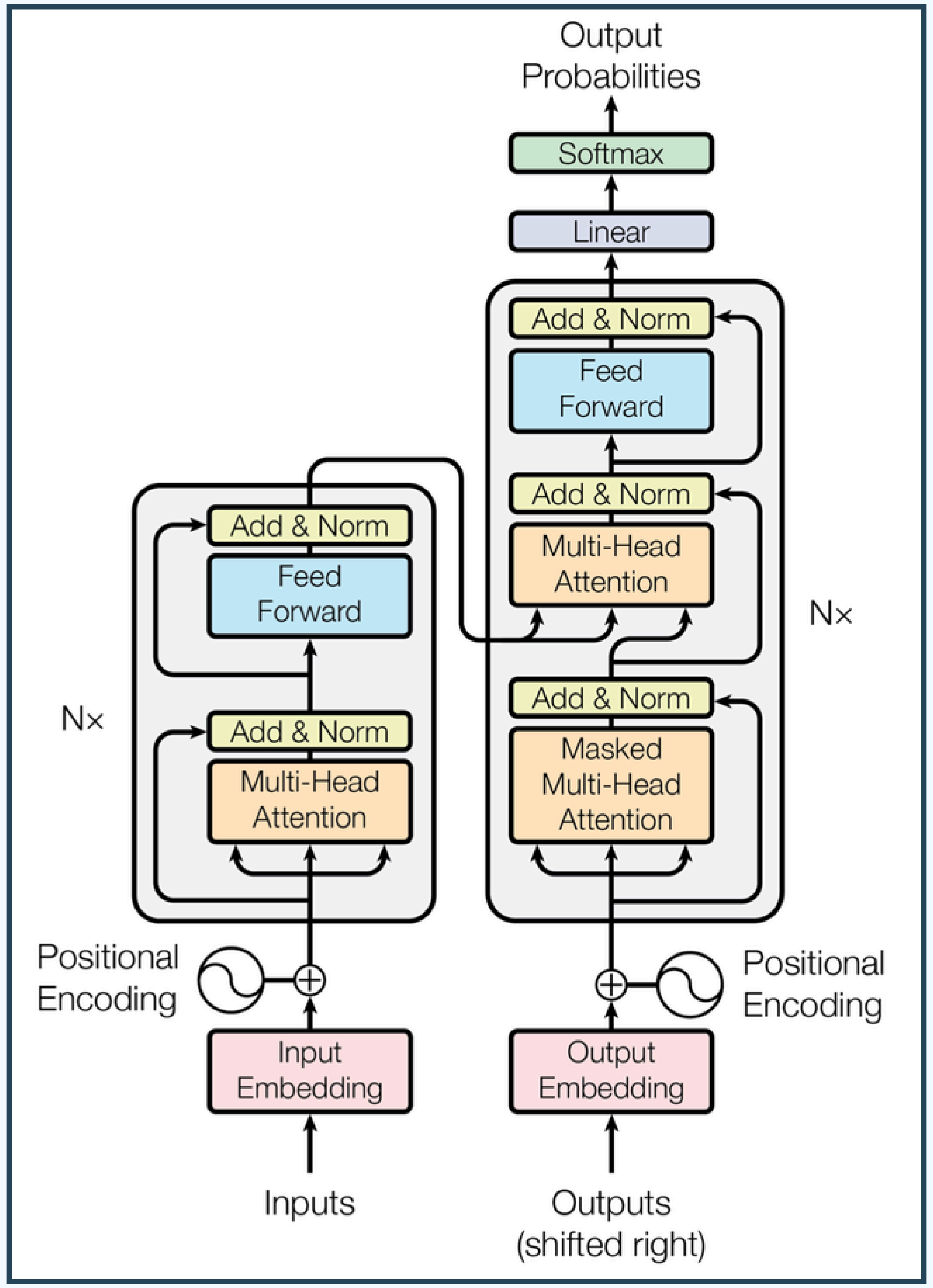
(Multivariate Time series forecasting with Graph Neural Networks)



Data transformation into graphs

1 Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks,Zonghan Wu et al,2020

TRANSFORMER



Attention Is All You Need, Ashish Vaswani, 2017

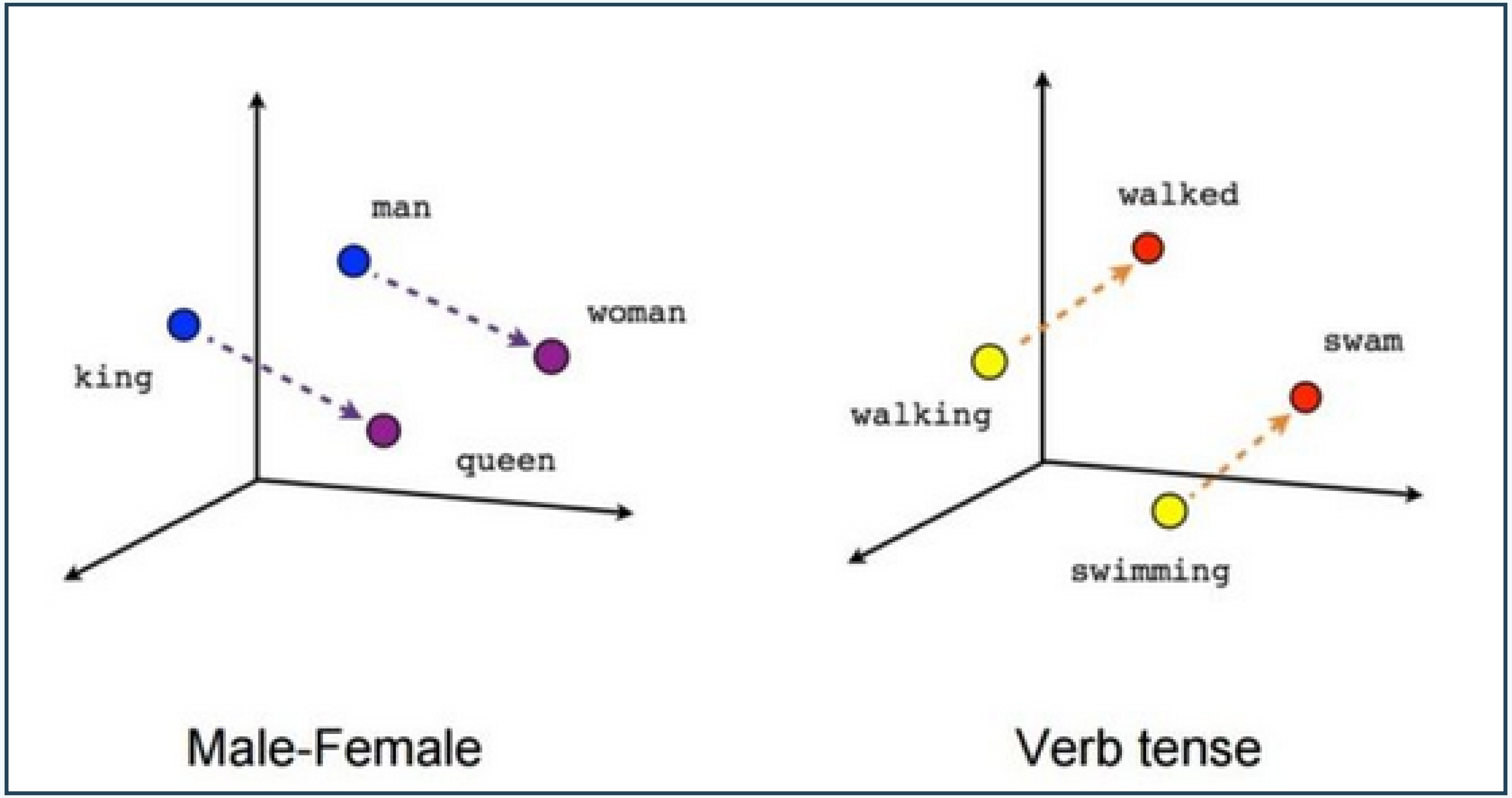
3 MAIN MECANISMS

- Encoder Decoder architecture
- Embedding
  - Vectorial representation of words
- Attention mecanism
- Word meaning refining
- Contextualisation

TRANSFORMER

EMBEDDING

VECTORIAL REPRESENTATION OF WORDS



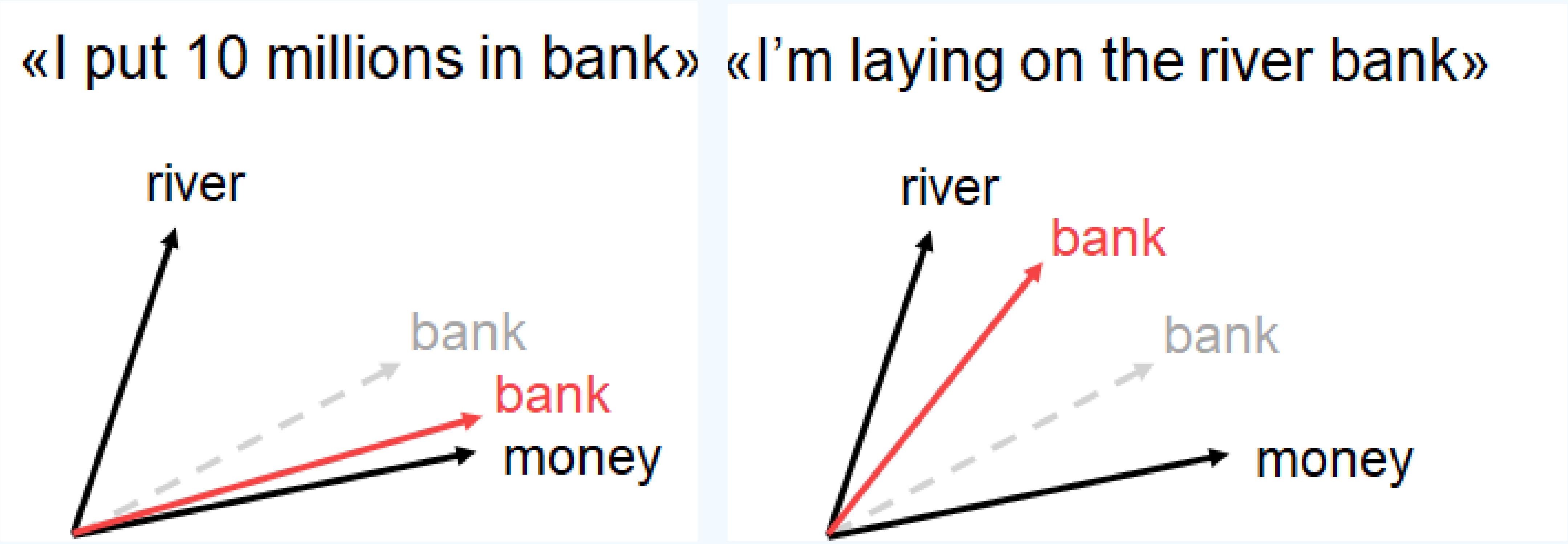
https://towardsdatascience.com/a-guide-to-word-embeddings-8a23817ab60f

TRANSFORMER

ATTENTION

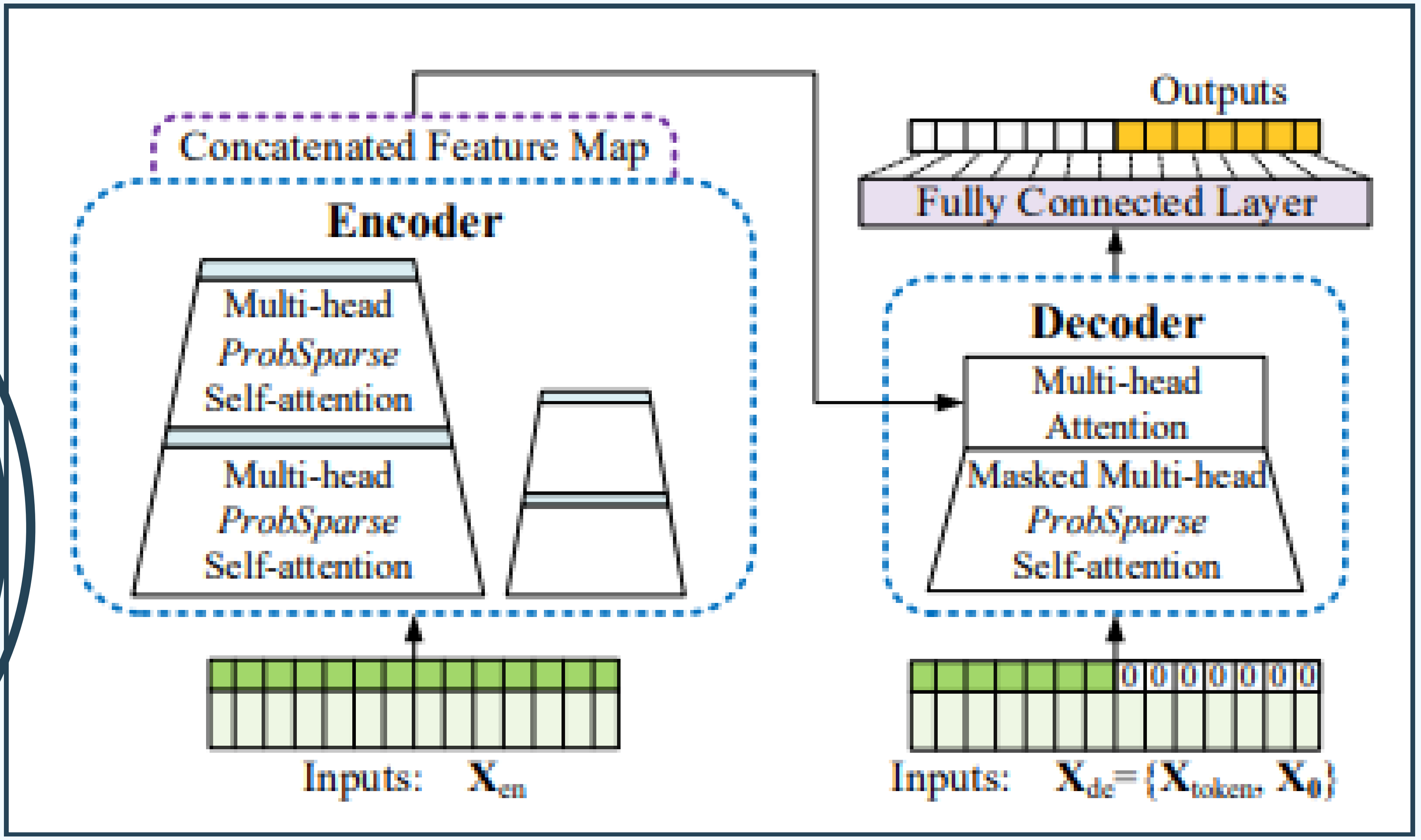
Refining words:

From a generic meaning to a specific one



https://medium.com/analytics-vidhya/the-rise-of-attention-in-neural-networks-8c1d57a7b188

INFORMER

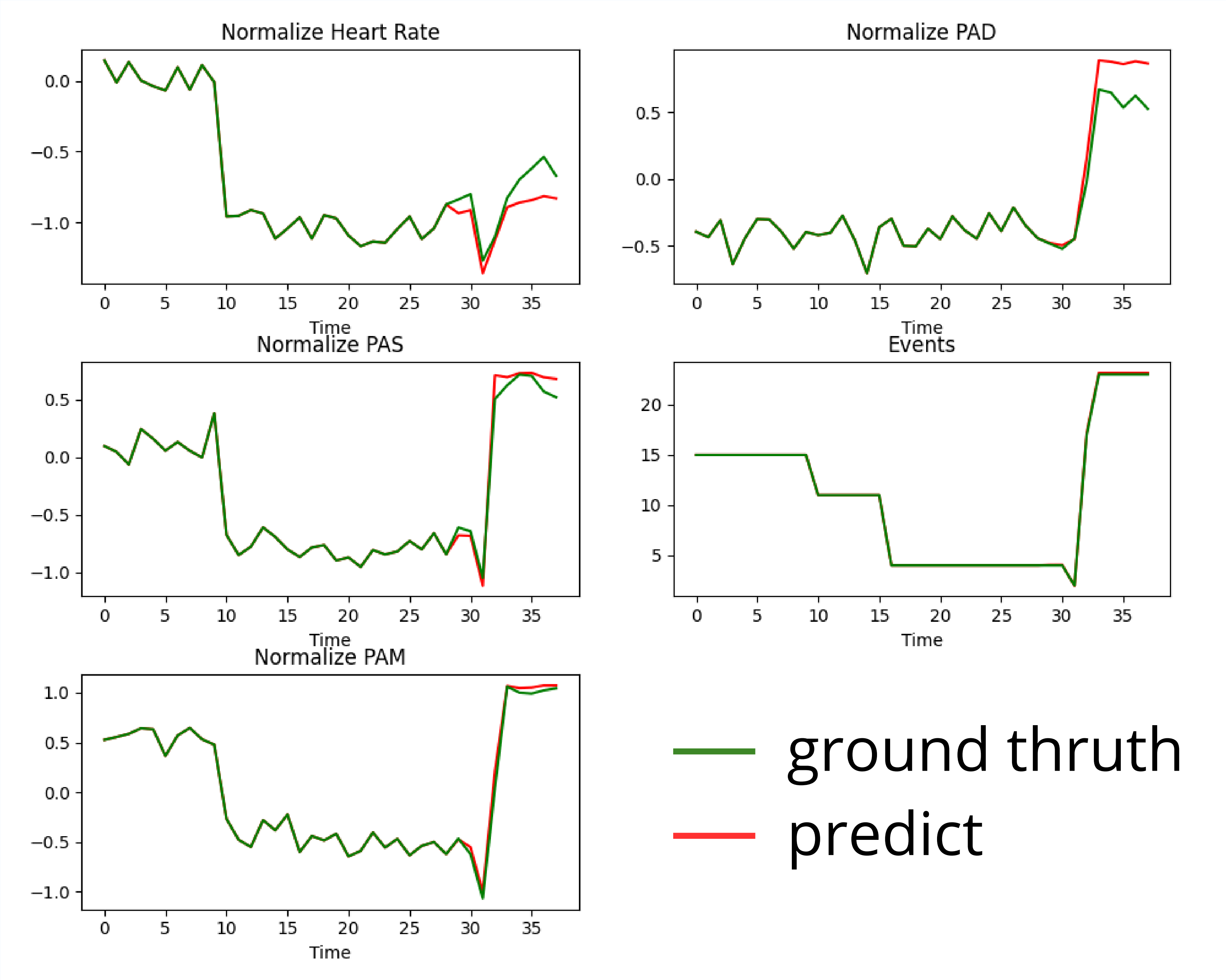


Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting, Haoyi Zhou et_al, 2021

- Long Sequence Time-Series Forecasting
- Transformer Limitations
  - The quadratic computation of self-attention
  - The memory bottleneck in stacking layers for long inputs
  - The speed plunge in predicting long outputs

RESULTS

Graphical Validation :



Transformer forecasting of the FC, PAD, PAS, PAM and the events

RESULTS

Without event:

	Time series forecasting			
Model	**MSE**	**RMSE**	**MAE**	**SMAPE**
LSTM	0,0993	0,3151	0,2002	0,4379
GNN	0,0729	0,2700	0,1694	0,3907
INFORMER	0,0790	0,2811	0,1710	0,4330
TRANSFORMER	**0,0001**	**0,0100**	**0,0067**	**0,0330**

METRICS

- MAE: Mean absolute error
- SMAPE: Symmetric mean absolute percentage error
- MSE: Mean squared error
- RMSE: Root mean square error
- F-Score

RESULTS

With event, without entity embedding:

	Time series forecasting				Event forecasting
Model	MSE	RMSE	MAE	SMAPE	F-SCORE
LSTM	0,1694	0,4000	0,2898	0,5946	0,1367
GNN	0,0874	0,2956	0,1985	0,4490	0,1742
INFORMER	0,0680	0,2608	0,1760	0,4560	0,6090
TRANSFORMER	**0,0288**	**0,1697**	**0,1306**	**0,3368**	**0,8342**

RESULTS

With event, with entity embedding:

	Time series forecasting				Event forecasting
Model	MSE	RMSE	MAE	SMAPE	F-SCORE
LSTM	0,1145	0,3383	0,2383	0,5187	0,2174
GNN	0,0814	0,2828	0,1891	0,4285	0,2260
INFORMER	0,0550	0,2345	0,1500	0,3500	0,7200
TRANSFORMER	**0,0003**	**0,0173**	**0,0104**	**0,0346**	**0,9580**

CONCLUSION

	Time series forecasting				Event forecasting
Transformer	**MSE**	**RMSE**	**MAE**	**SMAPE**	**F-SCORE**
without event	**0,0001**	**0,0100**	**0,0067**	**0,0330**	X
with event / without entity embedding	0,0288	0,1697	0,1306	0,3368	0,8342
with event / with entity embedding	0,0003	0,0173	0,0104	0,0346	**0,9580**

- Most performant model : Transformer
- Test on real data
- Create our own synthetic data

Annex

Metrics

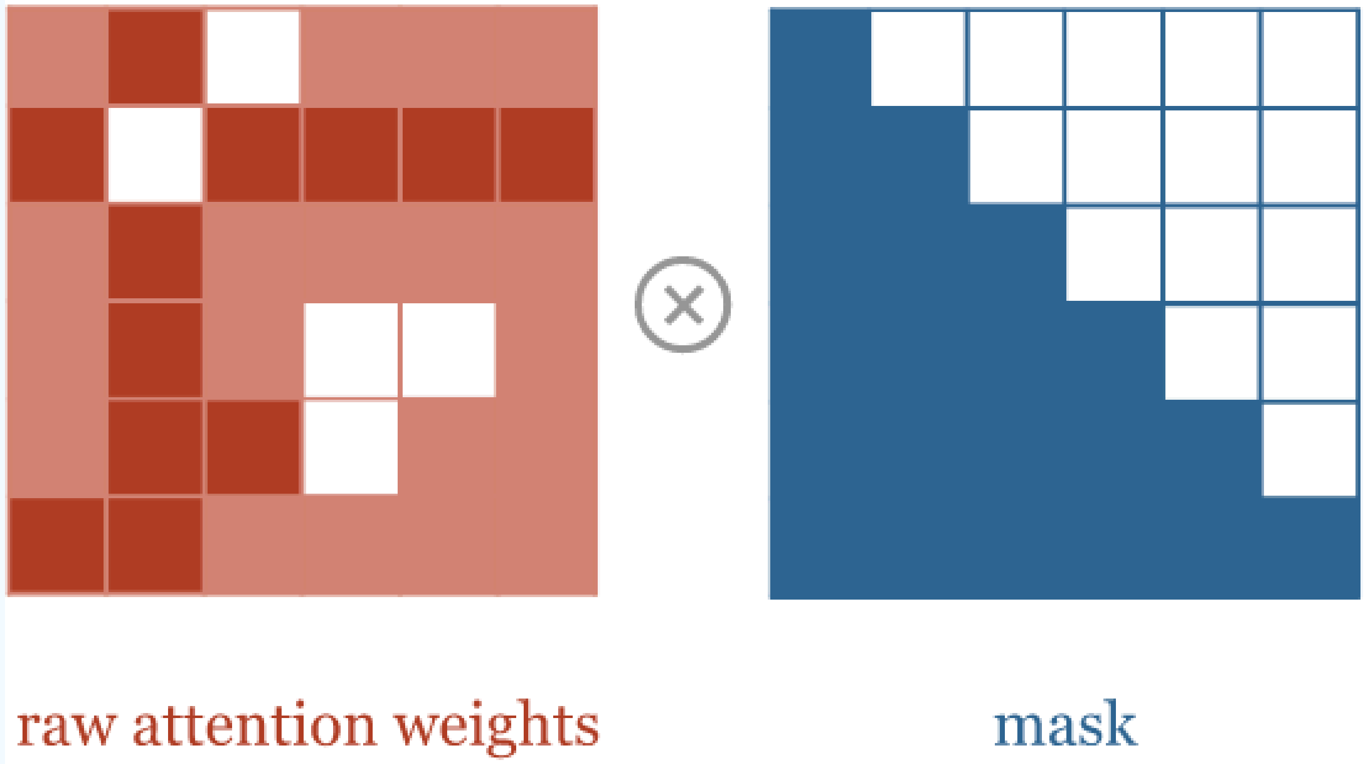
$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2} \quad MAE = \frac{1}{n}\sum_{i=0}^{n}|y_i - \hat{y}_i|$$

$$MSE = \frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2 \qquad SMAPE = \frac{1}{n}\sum_{n}^{i=0}|\frac{y_i - \hat{y}_i}{(y_i + \hat{y}_i)/2}|$$

$$F1score = \frac{2 * TP}{2 * TP + FP + FN}$$

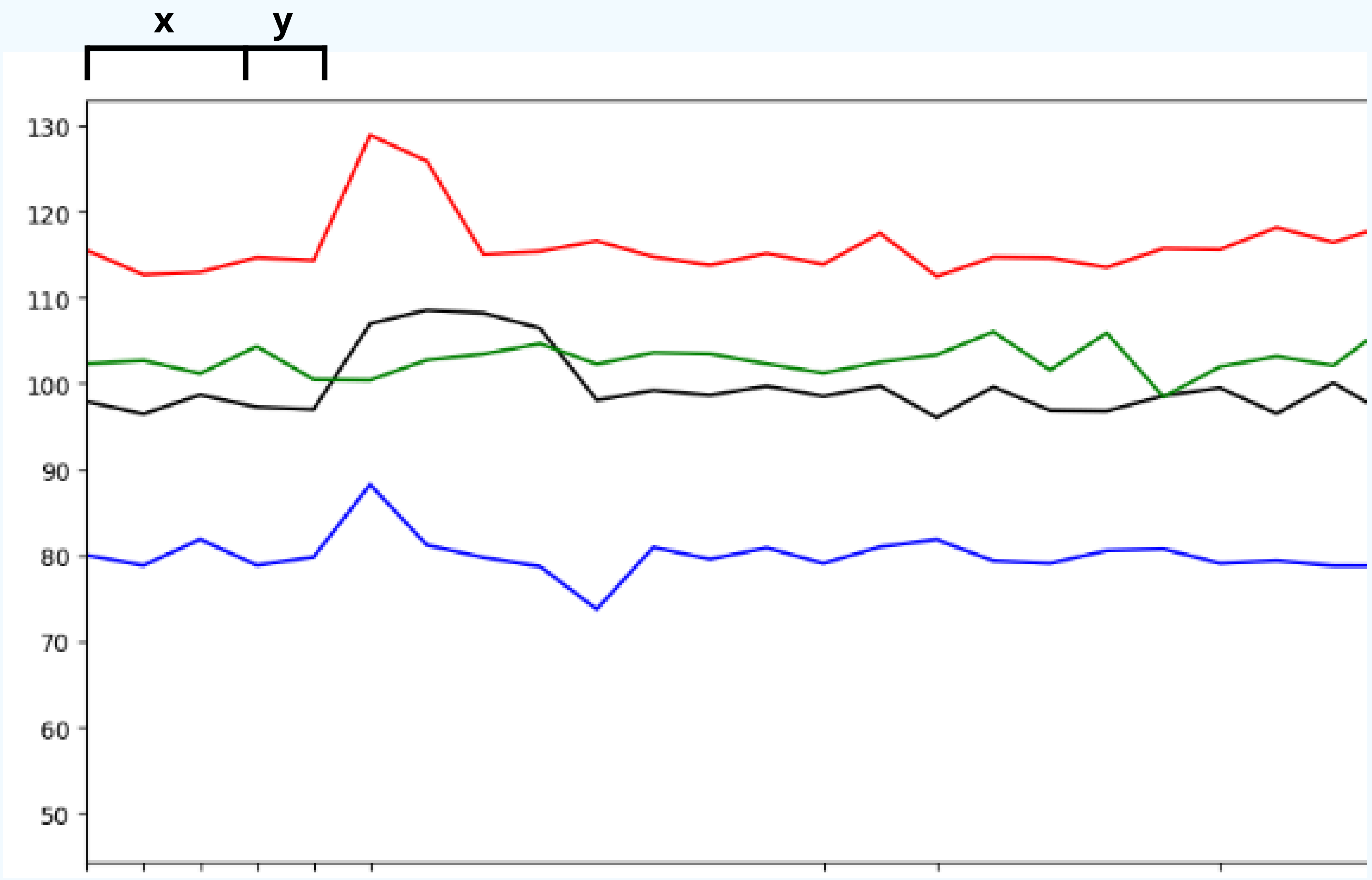
Annex

Prevent using future value to predict

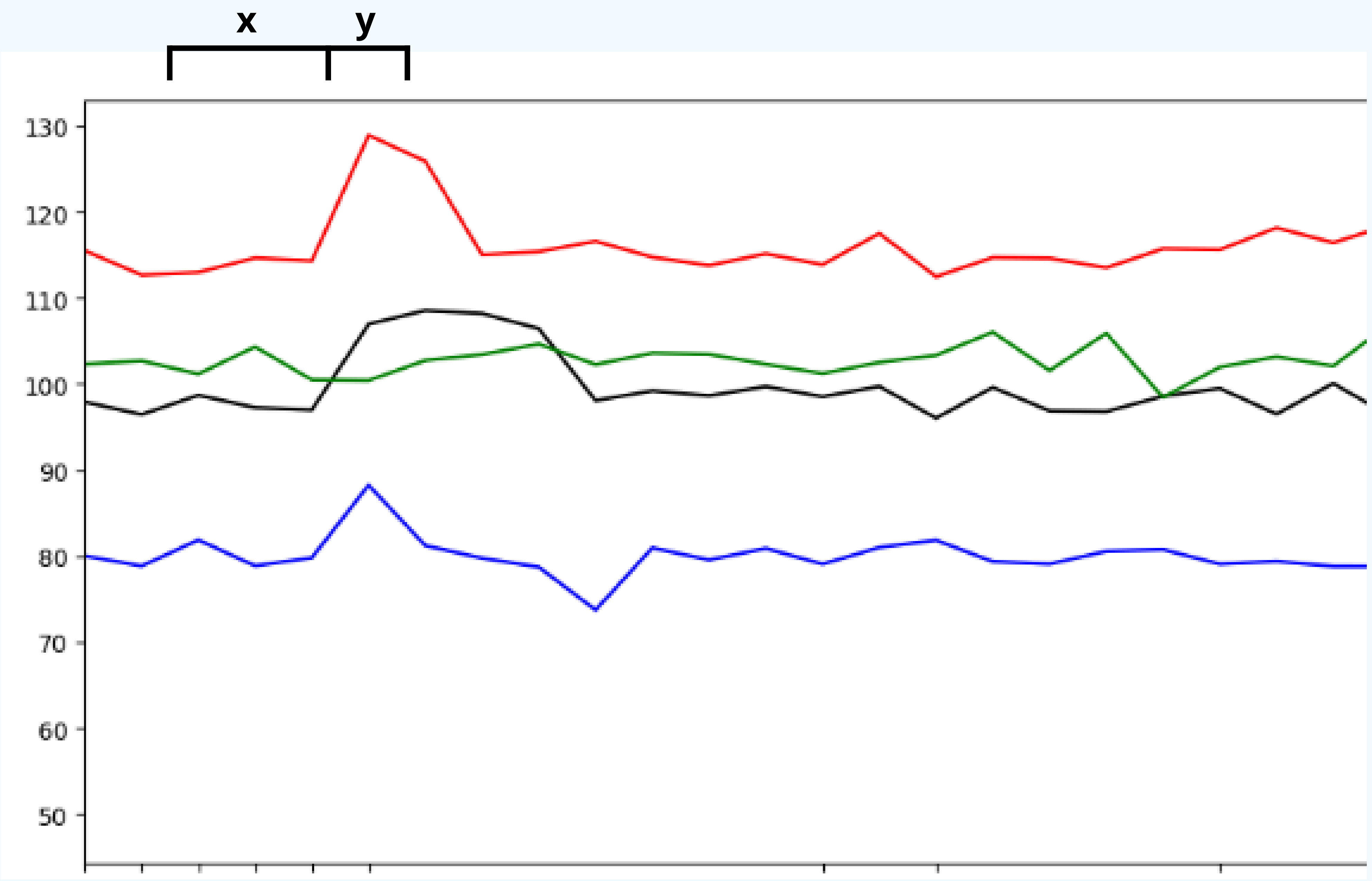


$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Annex



Annex



Annex

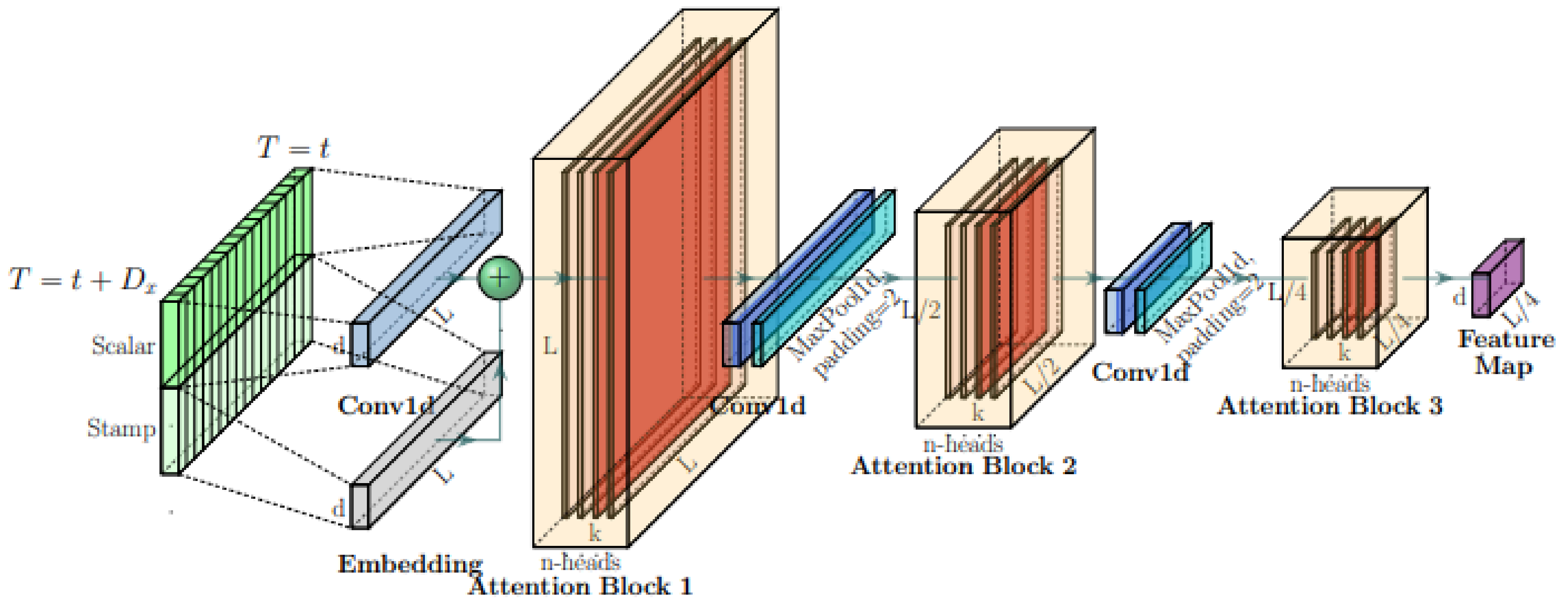


Figure 3: The single stack in Informer's encoder. (1) The horizontal stack stands for an individual one of the encoder replicas in Fig.(2). (2) The presented one is the main stack receiving the whole input sequence. Then the second stack takes half slices of the input, and the subsequent stacks repeat. (3) The red layers are dot-product matrixes, and they get cascade decrease by applying self-attention distilling on each layer. (4) Concatenate all stacks' feature maps as the encoder's output.