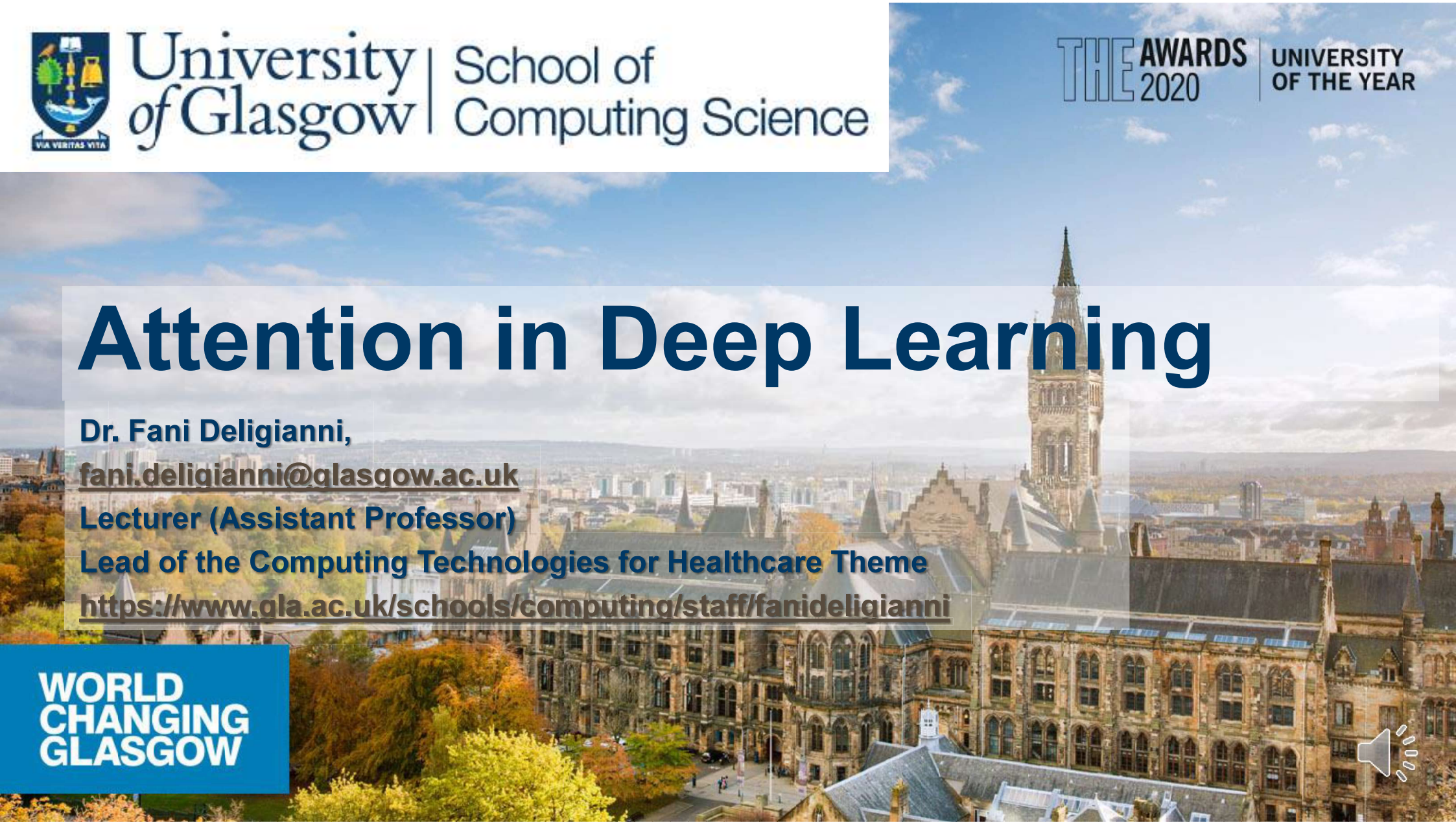


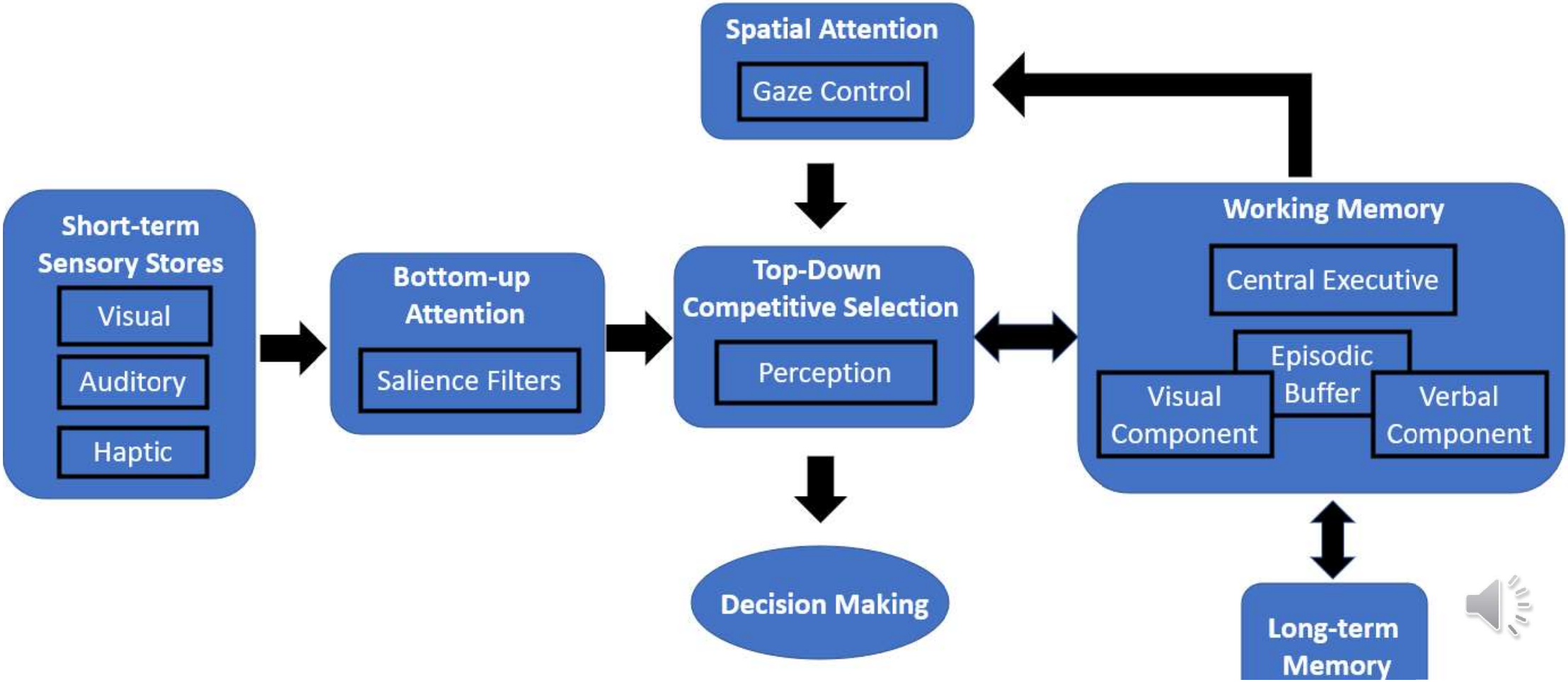
Attention in Deep Learning

Dr. Fani Deligianni,
fani.deligianni@glasgow.ac.uk
Lecturer (Assistant Professor)
Lead of the Computing Technologies for Healthcare Theme
https://www.gla.ac.uk/schools/computing/staff/fanideligianni

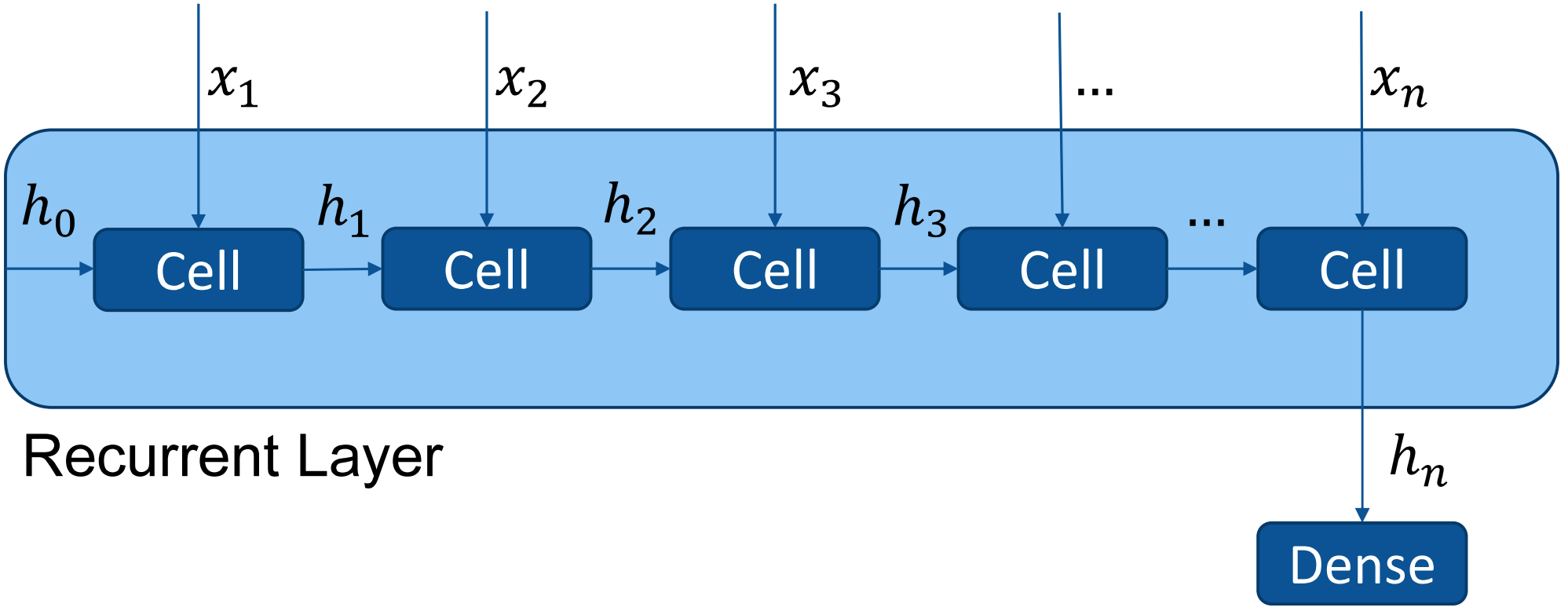
Attention in humans

Attention in humans

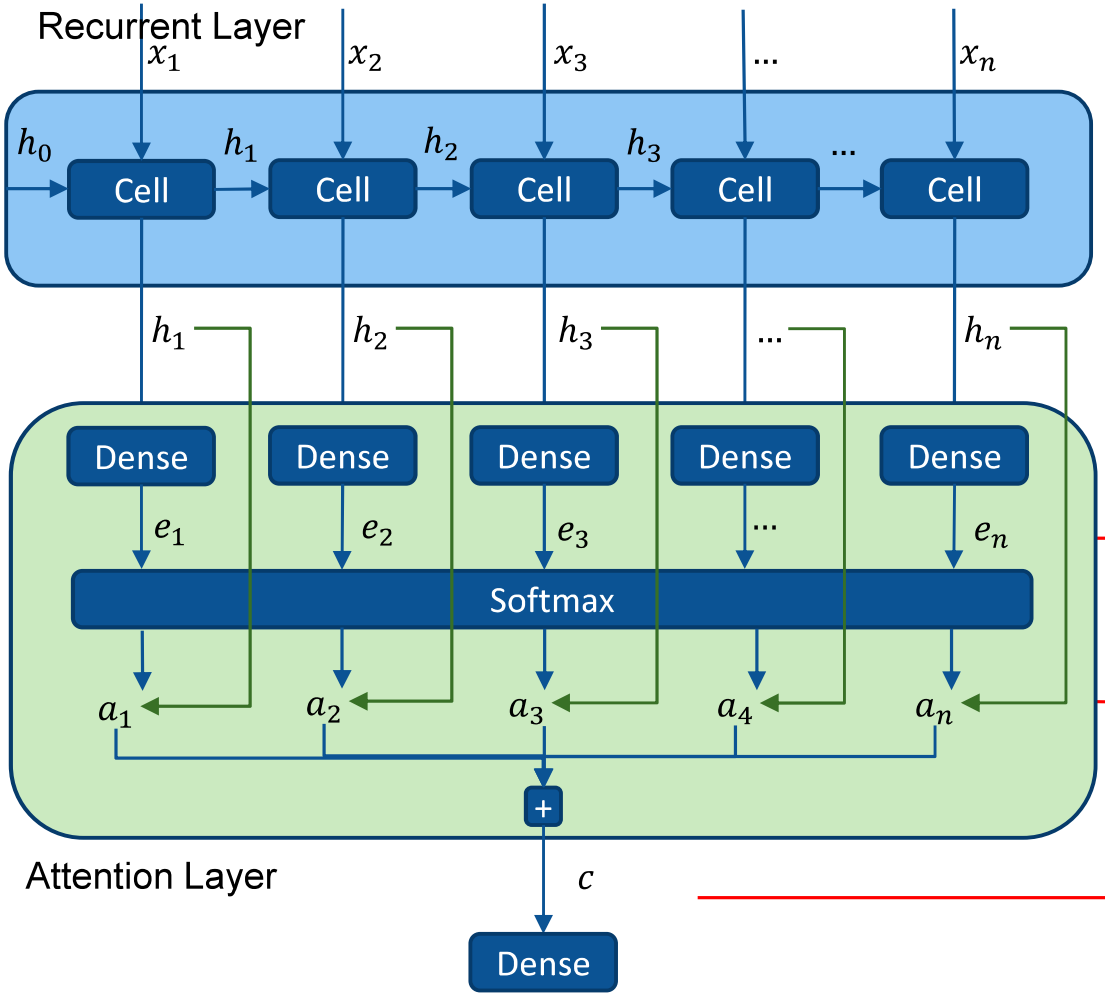
Attention as a cognitive function



Building attention in RNNs



Building attention in RNNs

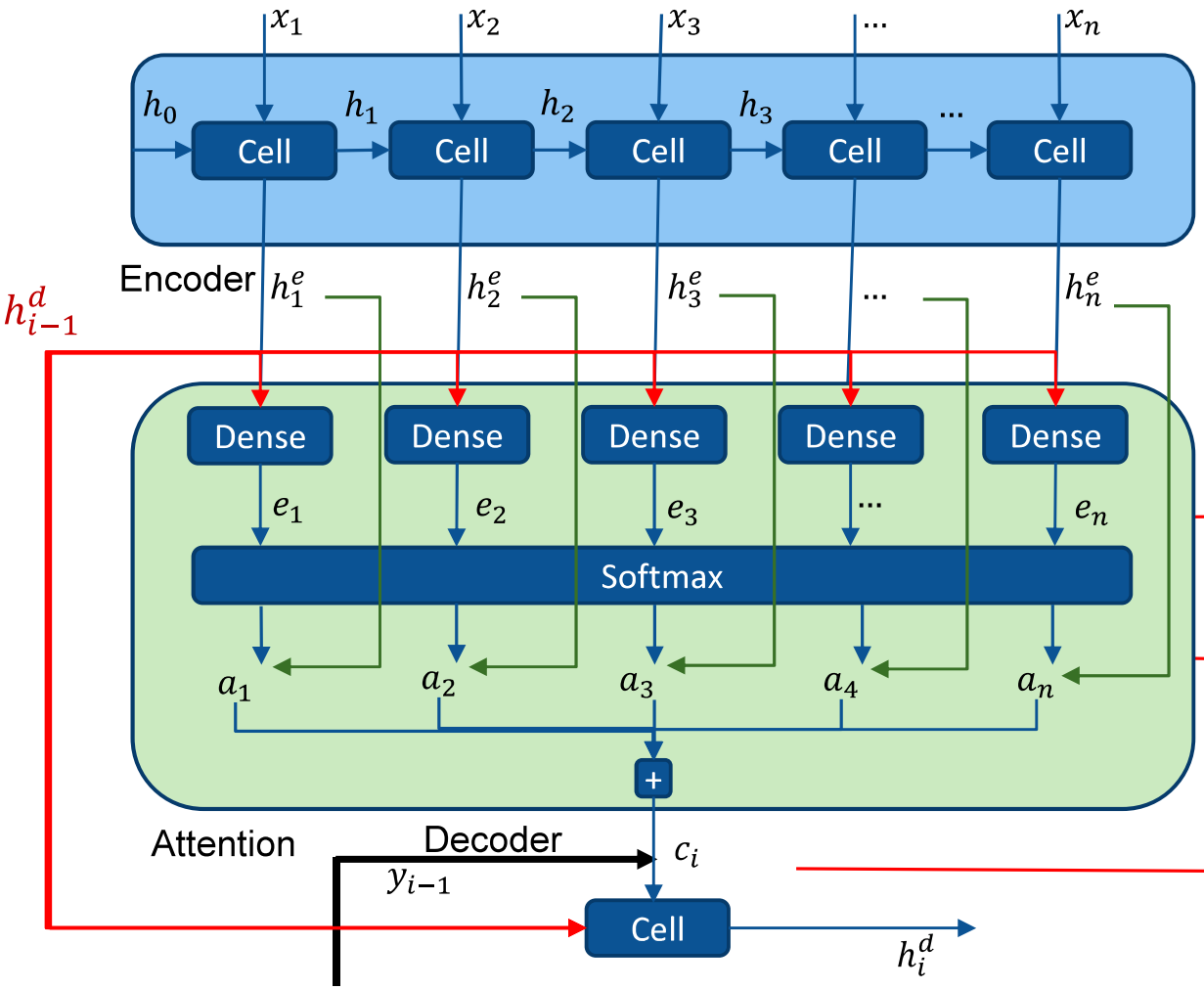


$$e_j = a(h_j) = \tanh(W * h_j)$$

$$\alpha_j = (softmax(e)_j) = \frac{\exp(e_j)}{\sum \exp(e_k)}$$

$$c = \sum_j \alpha_j h_j$$

Attention in Autoencoders

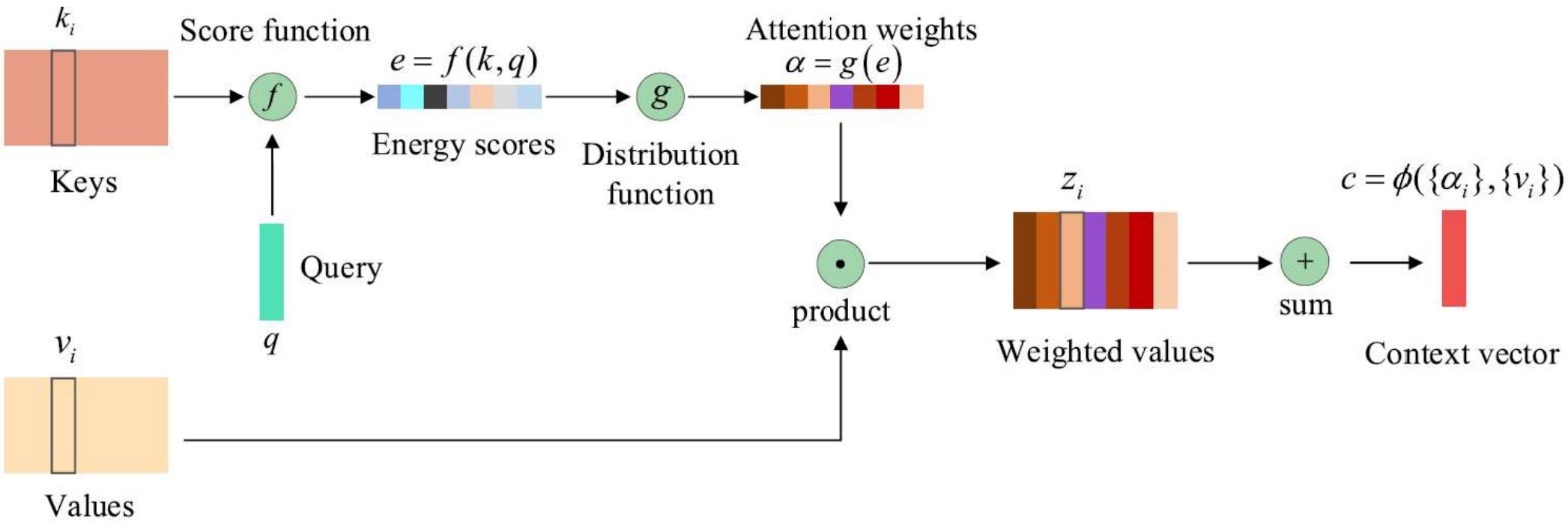


$$e_{ij} = a(h_j^e, h_i^d) = \tanh(W * h_j^e + U * h_{i-1}^d)$$

$$\alpha_{ij} = (softmax(e_i)_j) = \frac{\exp(e_{ij})}{\sum \exp(e_{ik})}$$

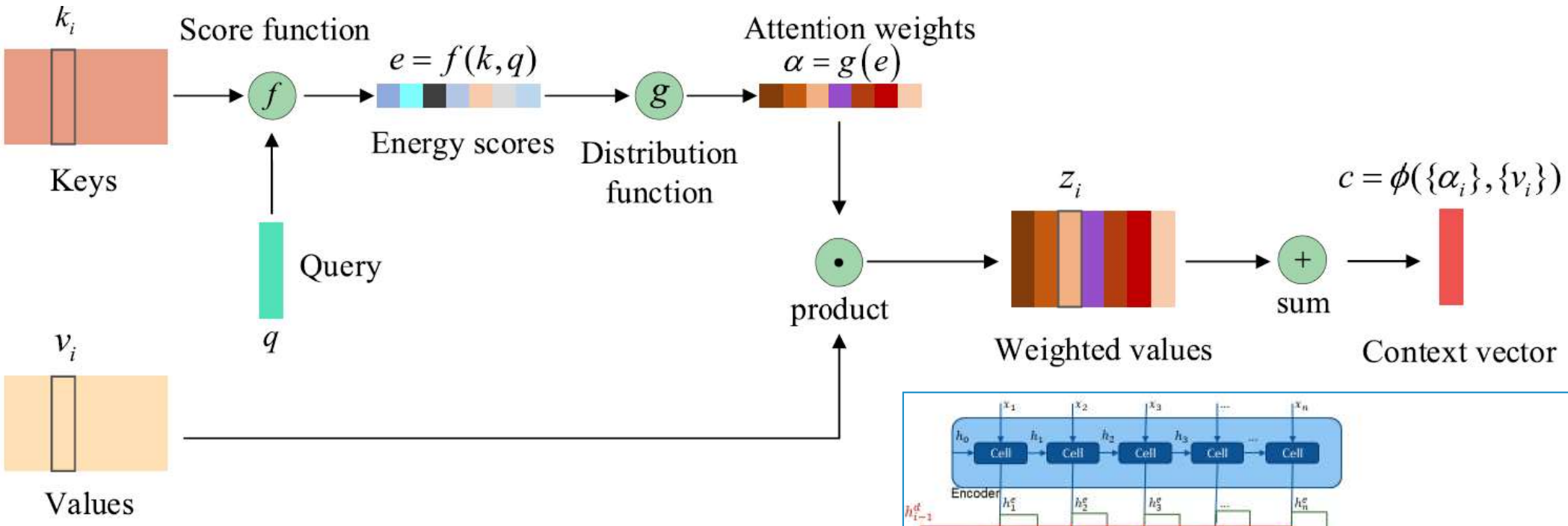
$$c_i = \sum_j \alpha_{ij} h_j^e$$

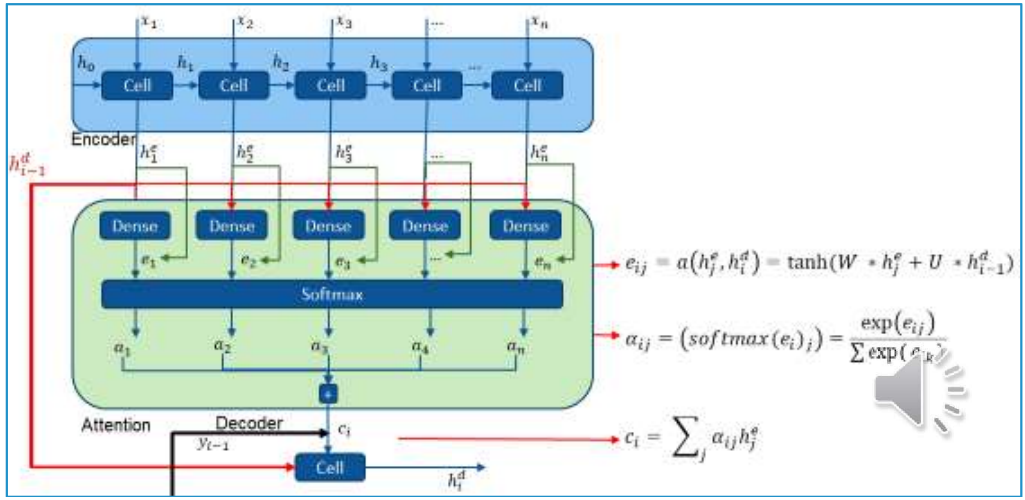
Generalised Model of Attention



Niu et al. A review on the attention mechanism of deep learning, Neurocomputing, 2021

Generalised Model of Attention





Niu et al. A review on the attention mechanism of deep learning, Neurocomputing, 2021

Summary

- Attention mechanisms is a significant breakthrough in deep learning
- Attention has been exploited to improve the performance of deep neural networks
- Attention can be also exploited to provide human understandable explanations

References

- Niu et al. A review on the attention mechanism of deep learning, Neurocomputing, 2021
- Foster, Generative Deep Learning – Teaching Machines to Paint, Write, Compose and Play, O'Reilly, 2019