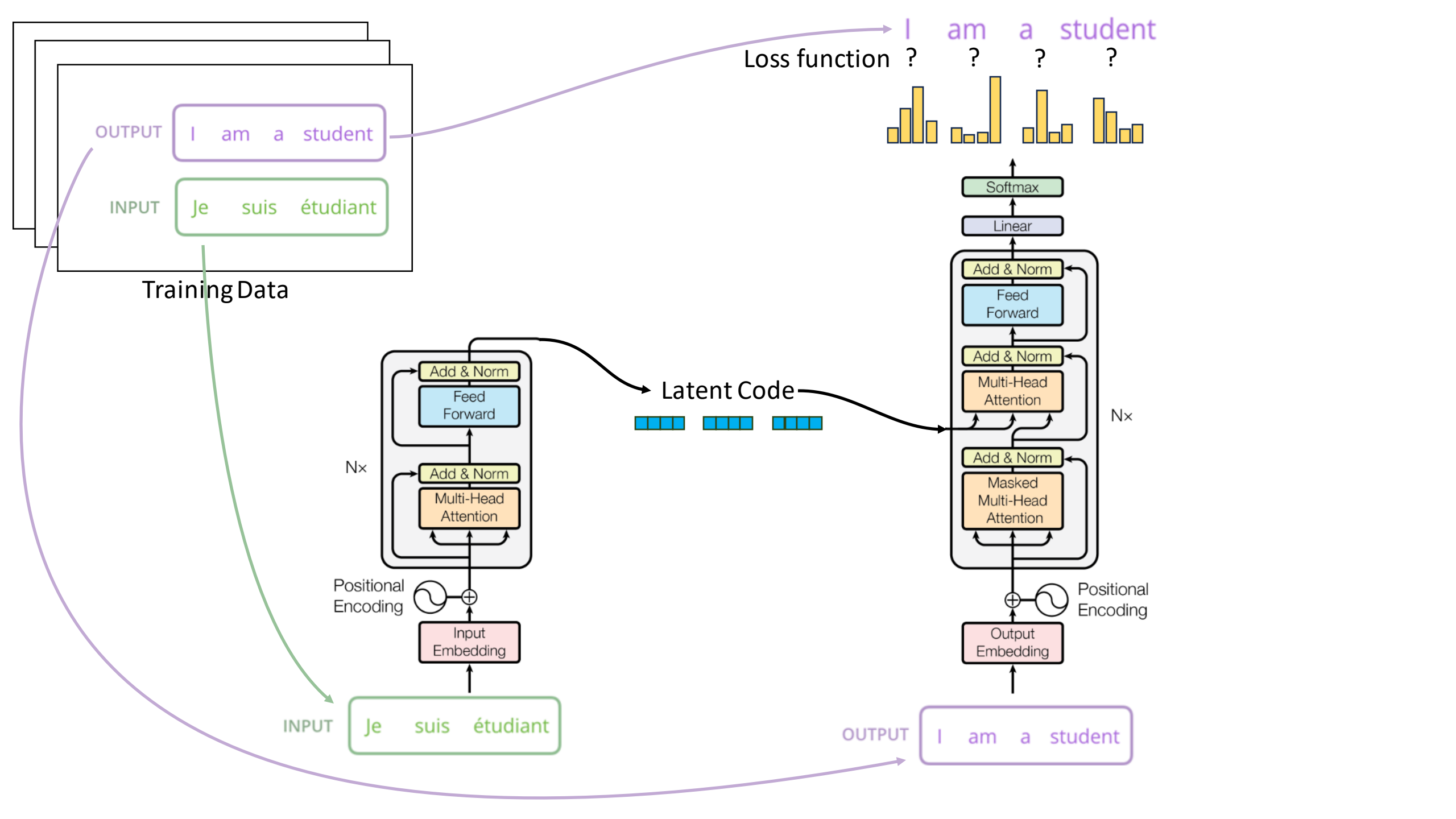
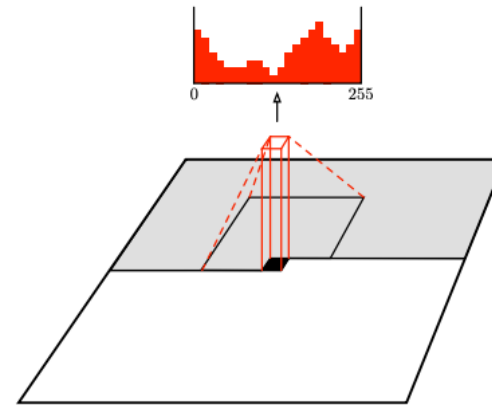
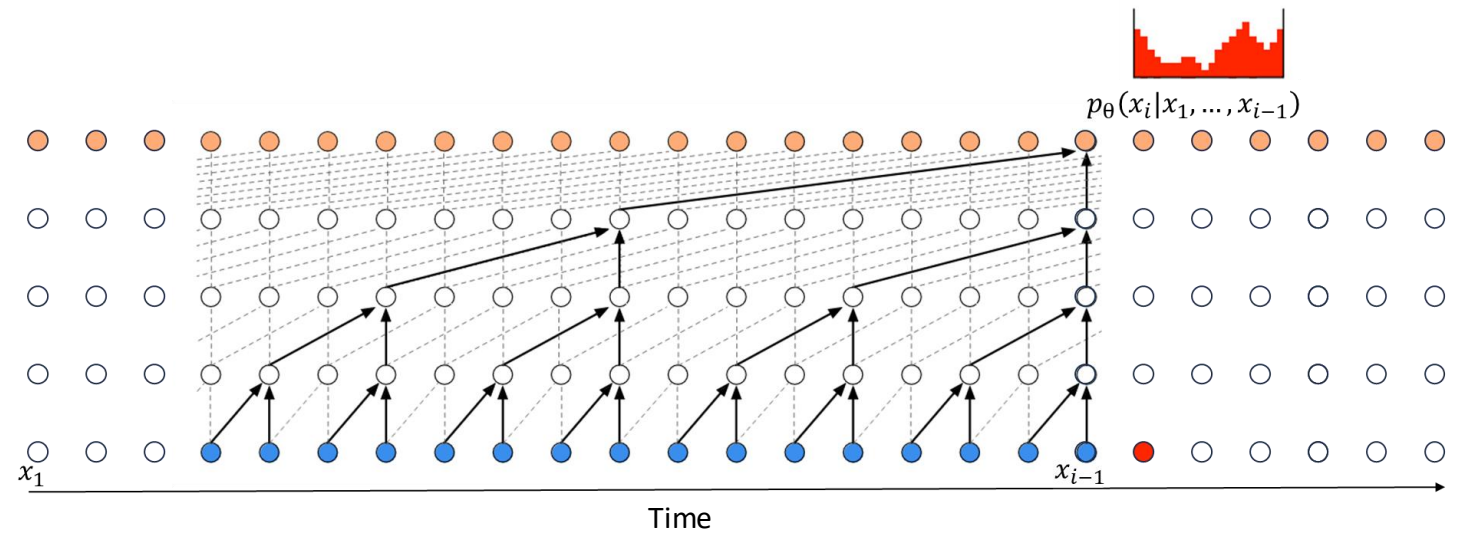
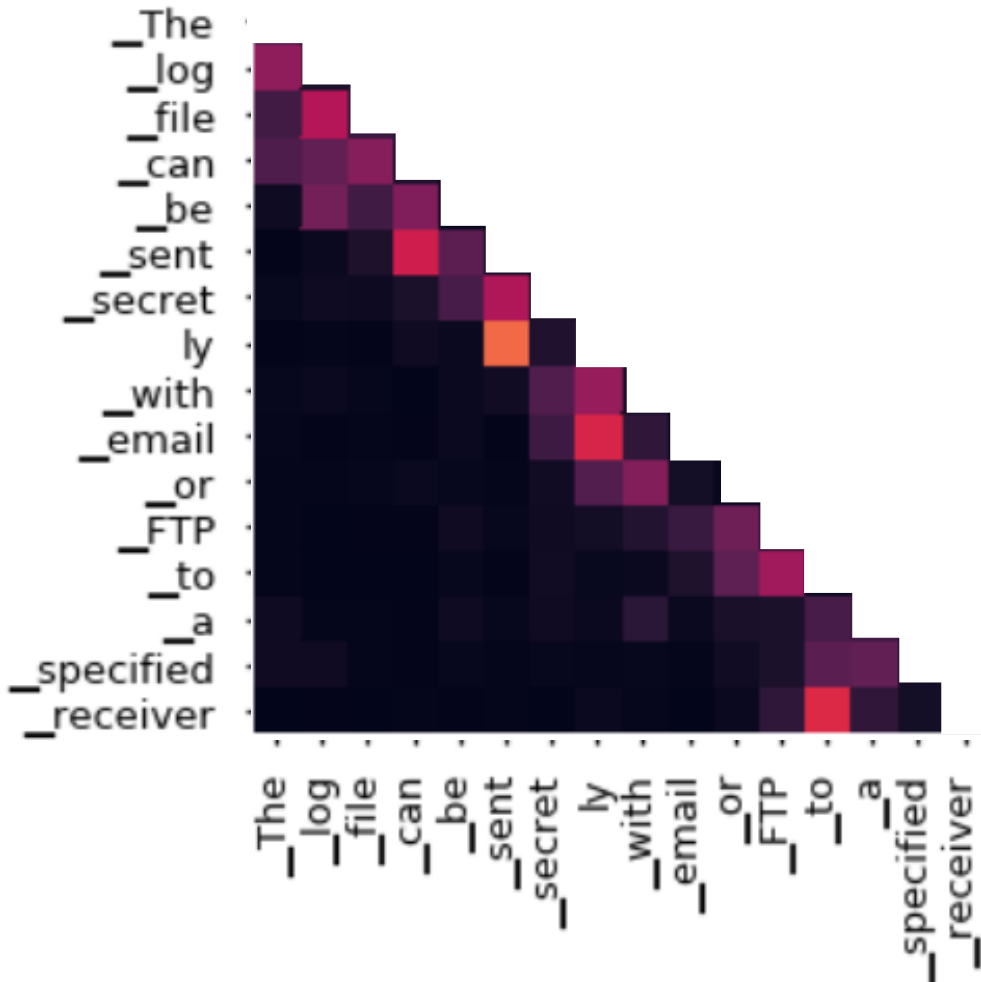


Neural Computation

Training



Masked Attention



Autoregressive Model:

- Output: dist. of next item
- Input: only previous items

