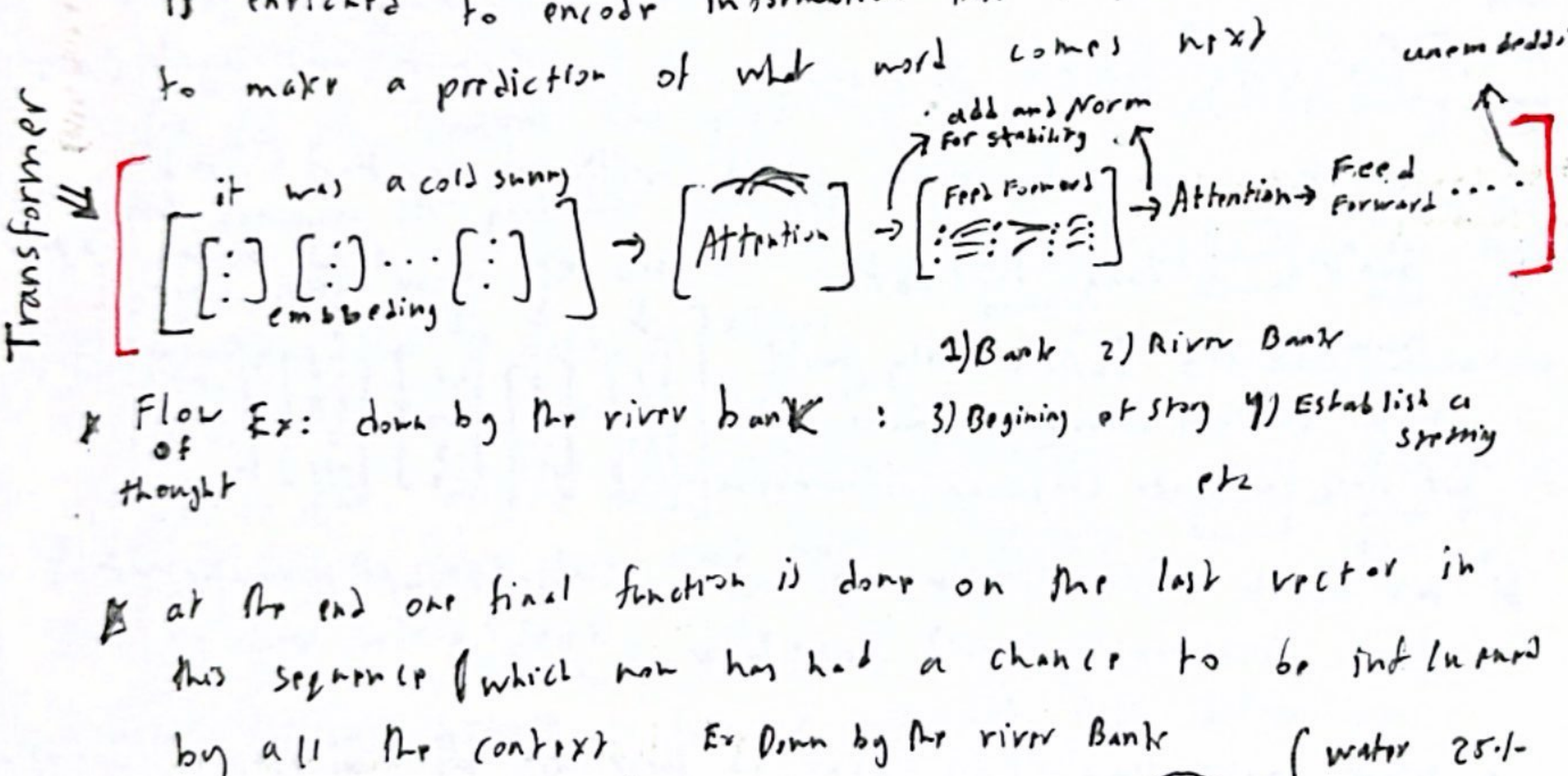
LLM intro cont.. (P-4)

- Apart from attention LLM's also have another type of operation called Feed Formed NN (MLP)

This gives the model extra capacity to store more patterns about language learned during training

Ex the fact MJ is a Basket Ball player (a fact) is stored in the Feed Forward layer

* all of this data repeatedly flows through many iterations of these two operations and as it does so the hope is that each list of numbers (word) is enriched to encode information that might be needed to make a prediction of what word comes next

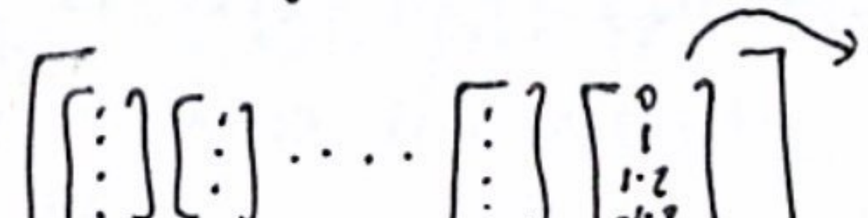


1) Bank 2) River Bank
3) Begining of story 4) Establish a setting etc

* Flow of thought Ex: down by the river bank :

* at the end one final function is done on the last vector in this sequence (which now has had a chance to be influenced by all the context and parameters of LLM) to produce a prediction of the next word

Ex Down by the river Bank



- once again the output is a probability of all posible next words

* Like NN why the model makes these predictions is probabilistic and hard to understand and impossible to understand the parameters.