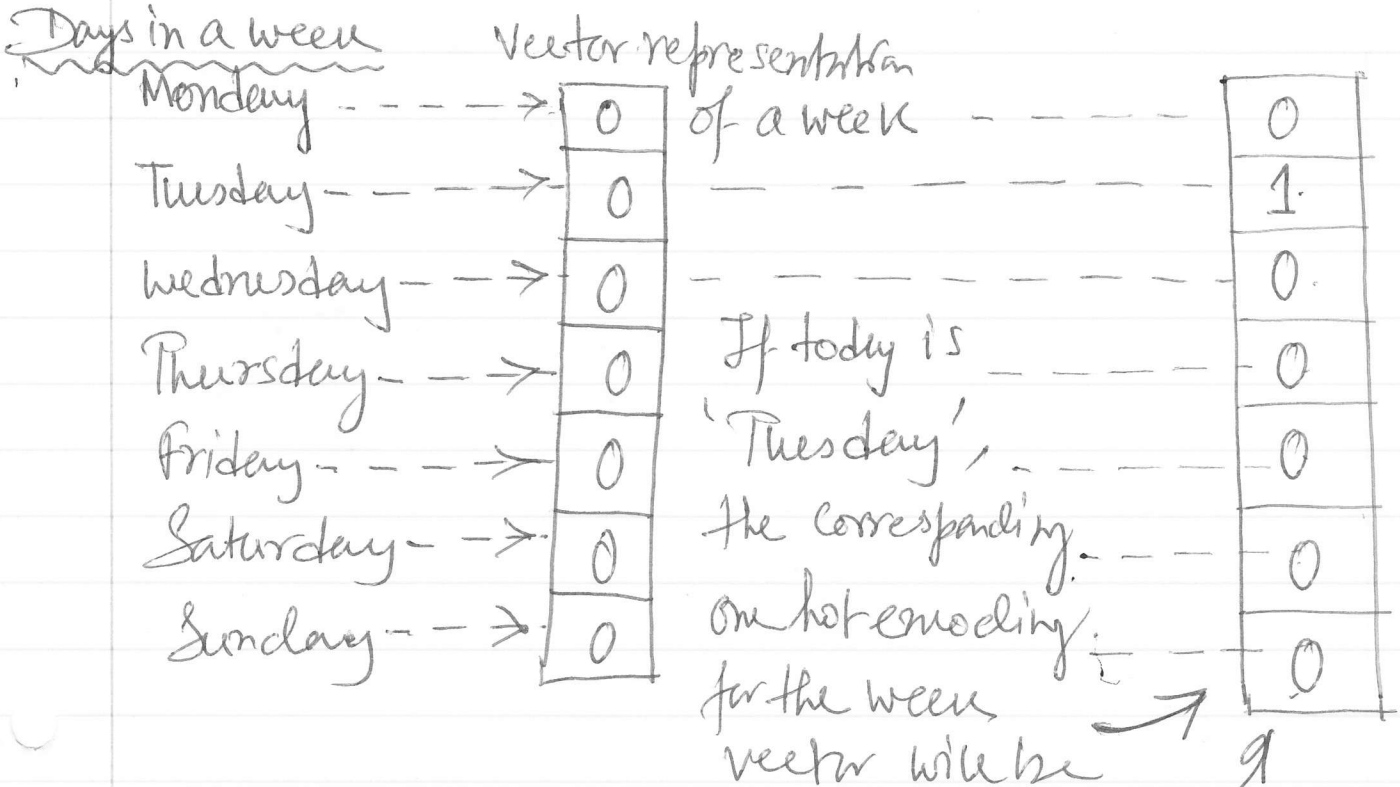
- Image Detection → Convolutional NN
  ↕
  Similar problems
- Sequence-to-Sequence (translation) → Recurrent NN
  - text generation/recognition
  - Speech detection/generation
  ↕
  Similar problems

  ↓
  particularly LSTM

One Hot encoding

In one hot encoding, all the values in a list is Zero ('0'), except the true value which is One ('1').

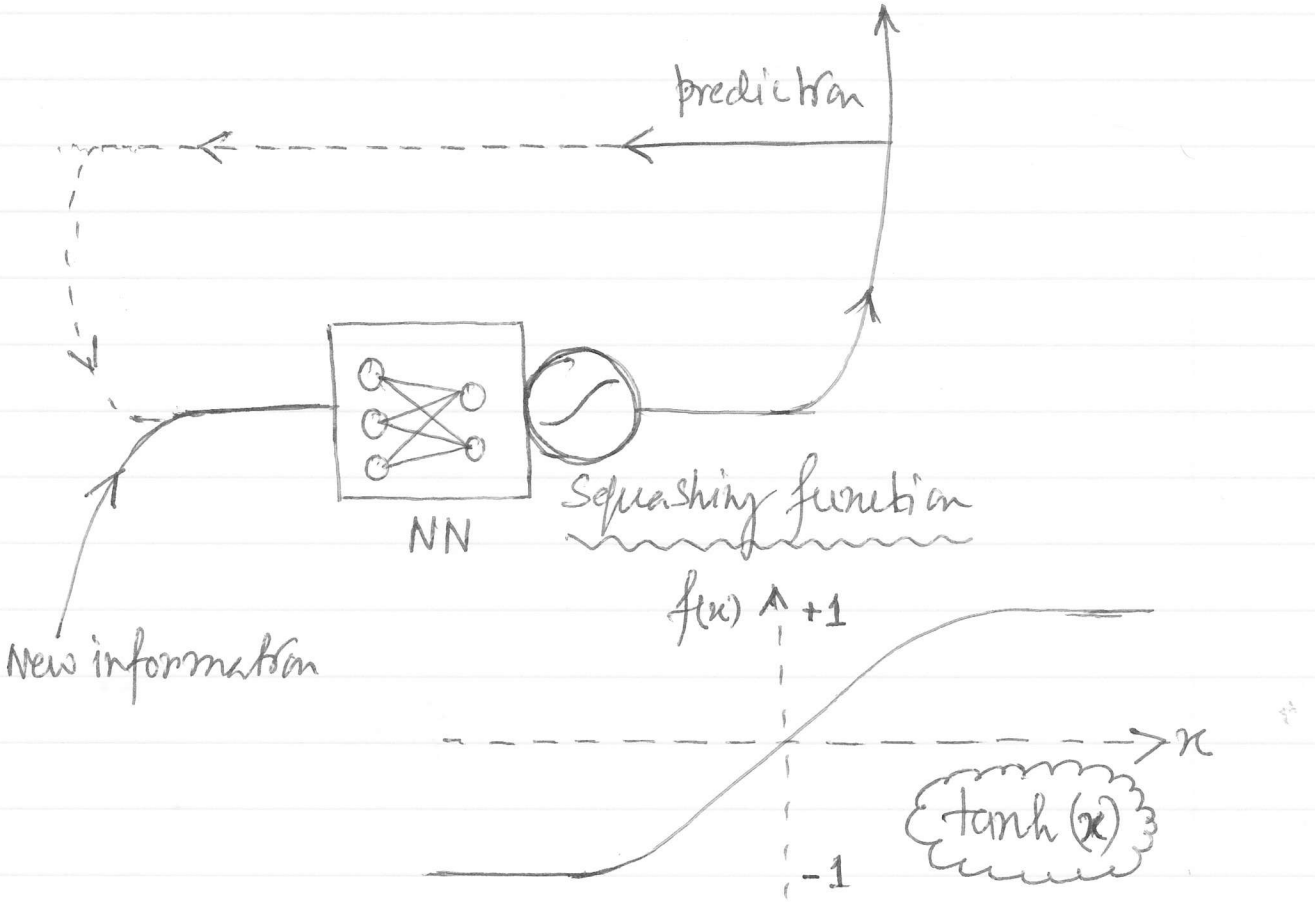


Also a binary representation of the data ①

In a feedforward network the outputs are independent of each others. The output never influences the input.

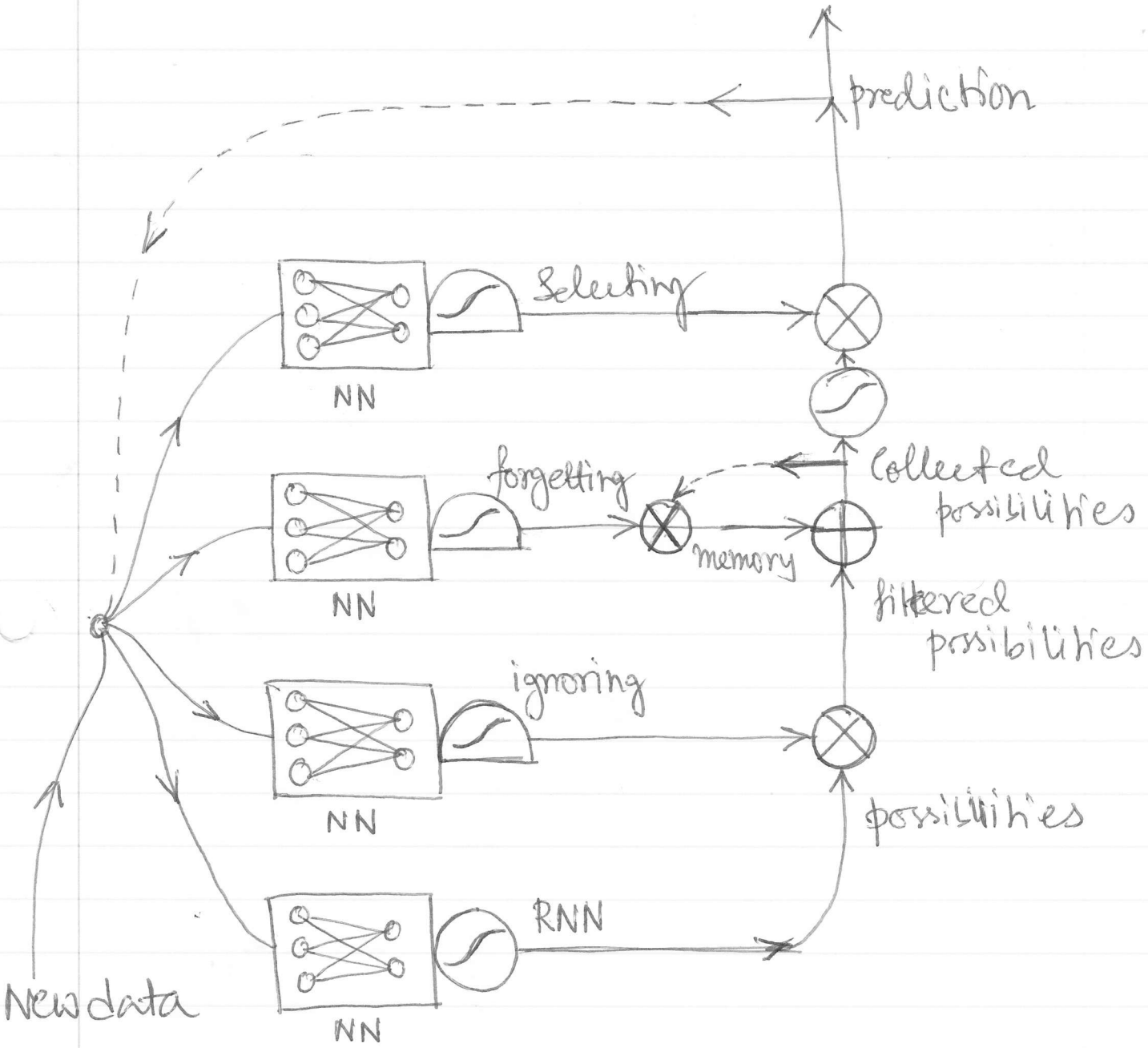
So, a feedforward network is not very useful for sequential data, e.g. text recognition, where words are related to each others.

Recurrent Neural Network



- A part of the prediction mixed with the new information to give the next and more improved prediction
- RNN can only use information from the previous step and no further back — limitation of RNN

Long Short Term Memory (LSTM)



- memory → we want to remember what happened many time steps ago
- Selecting → relearning Selected memory with predicting
- ignoring → ignores irrelavent memory to add to the prediction

- ⊕ plus junction: element-by-element addition

1	⊕	6	=	1+6	=	7
2		7		2+7		9
3		8		3+8		11

↑ equal size ↑

- ⊗ times junction: element-by-element multiplication

3	⊗	6	=	3×6	=	18
4		7		4×7		28
5		8		5×8		40

- Can be used as a gating

0·8	⊗	1.0	=	0·8	passes full signal — fully opened gate
0·8		0·5		0·4	passes half of the signal — gate half way opened
0·8		0·0		0·0	no signal — signal fully blocked

Signal — Switch/gate

-  : logistic (sigmoid) squashing function