



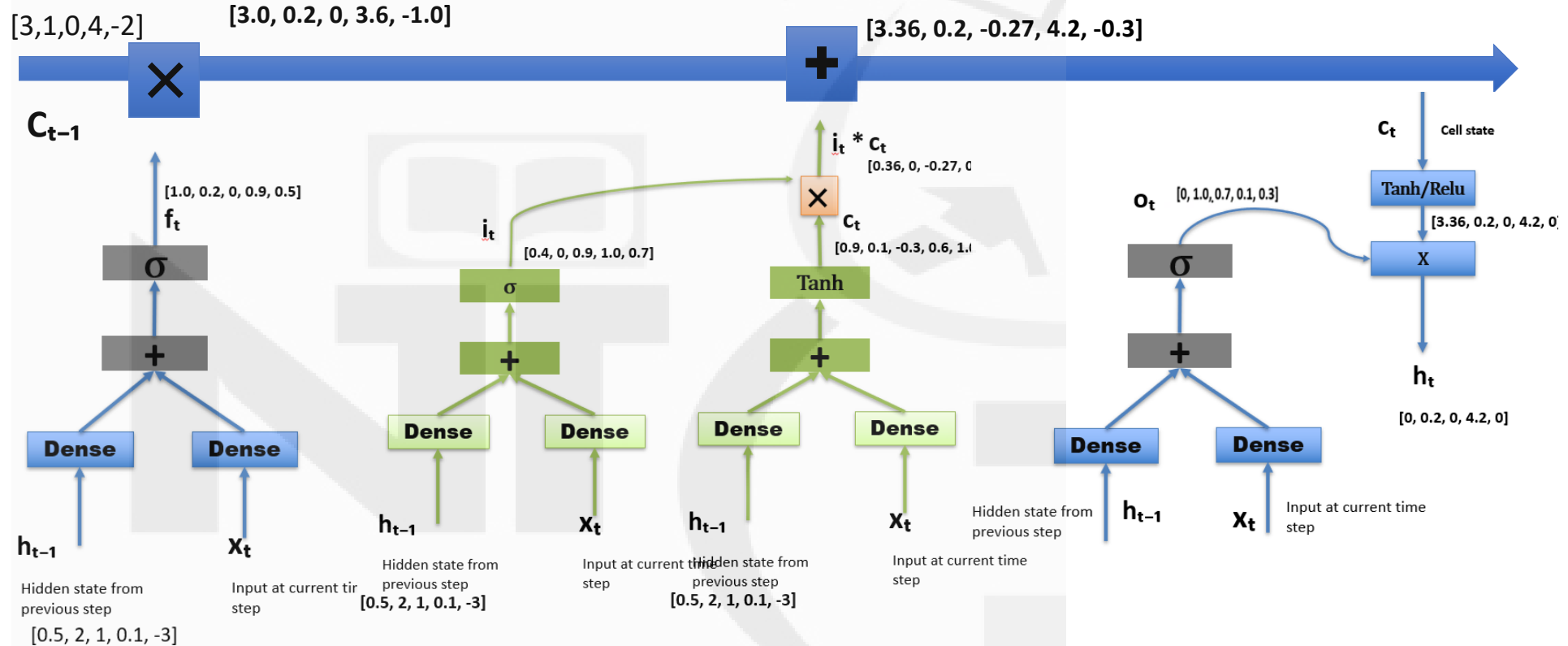
Gated Recurrent Unit

MUKESH KUMAR

GRU as a simplified alternative that:

- Retains the benefits of LSTM (handling vanishing gradients, learning dependencies),
- Has **fewer gates** and **fewer parameters**,
- Is **faster to train** and easier to implement.

LSTM CELL



Forget Gate

Controls things to remember & forget from memory

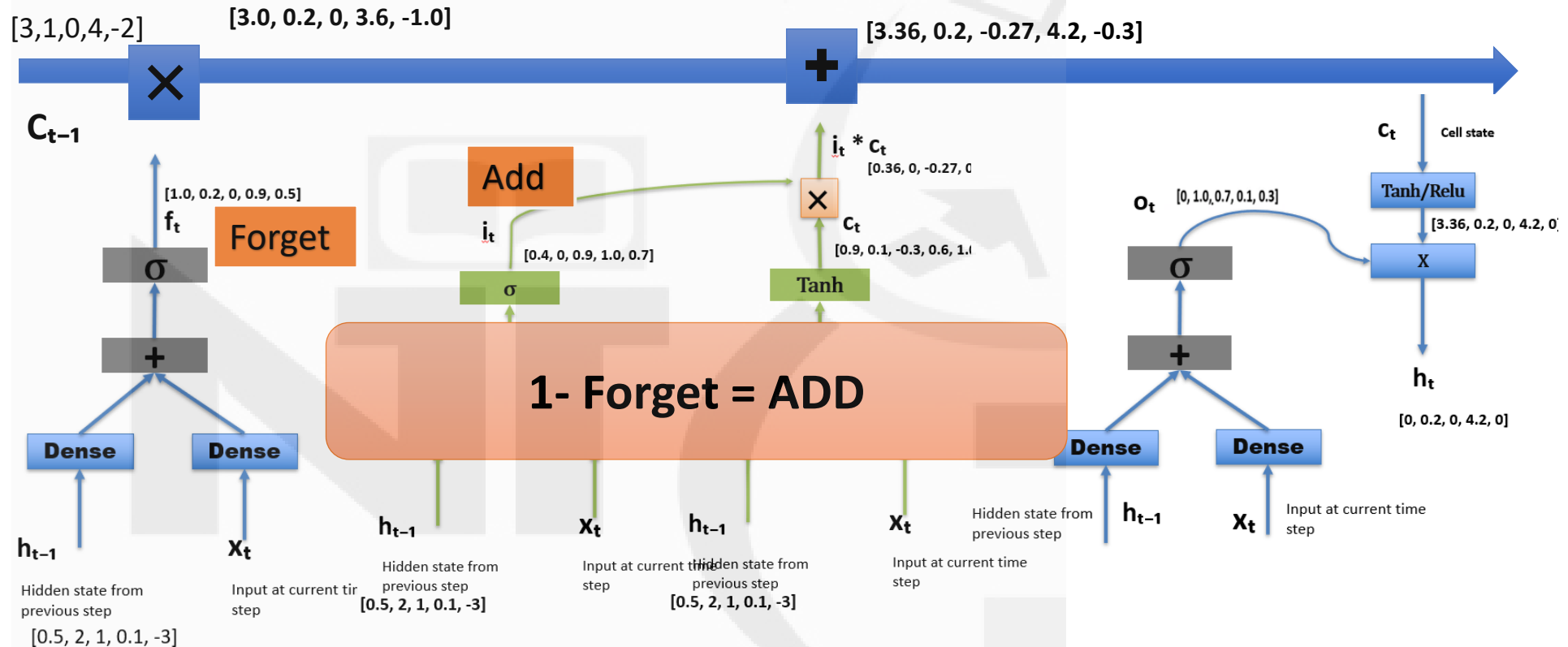
Input Gate

Controls what to add/update to memory

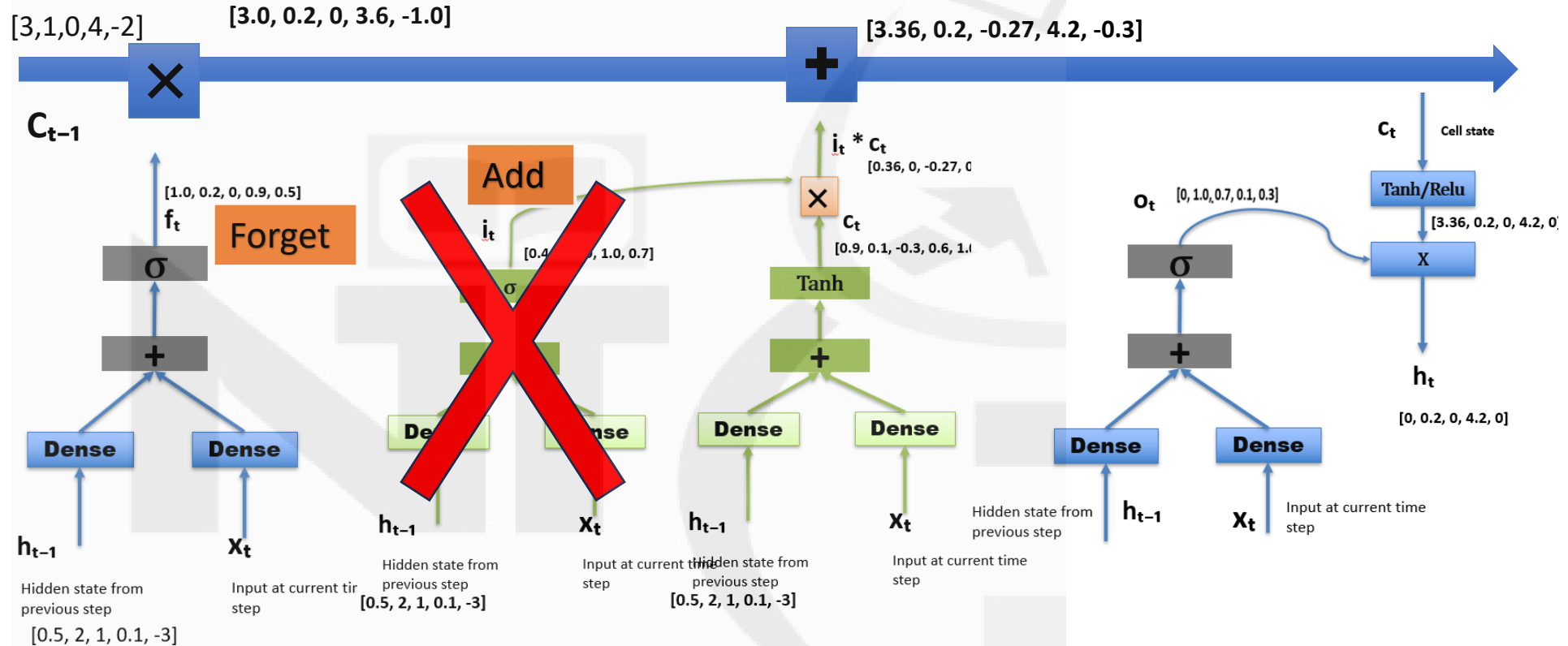
Output Gate

Controls what to read from memory

LSTM CELL



LSTM CELL



Forget Gate

Controls things to remember & forget from memory

Input Gate

Controls what to add/update to memory

Output Gate

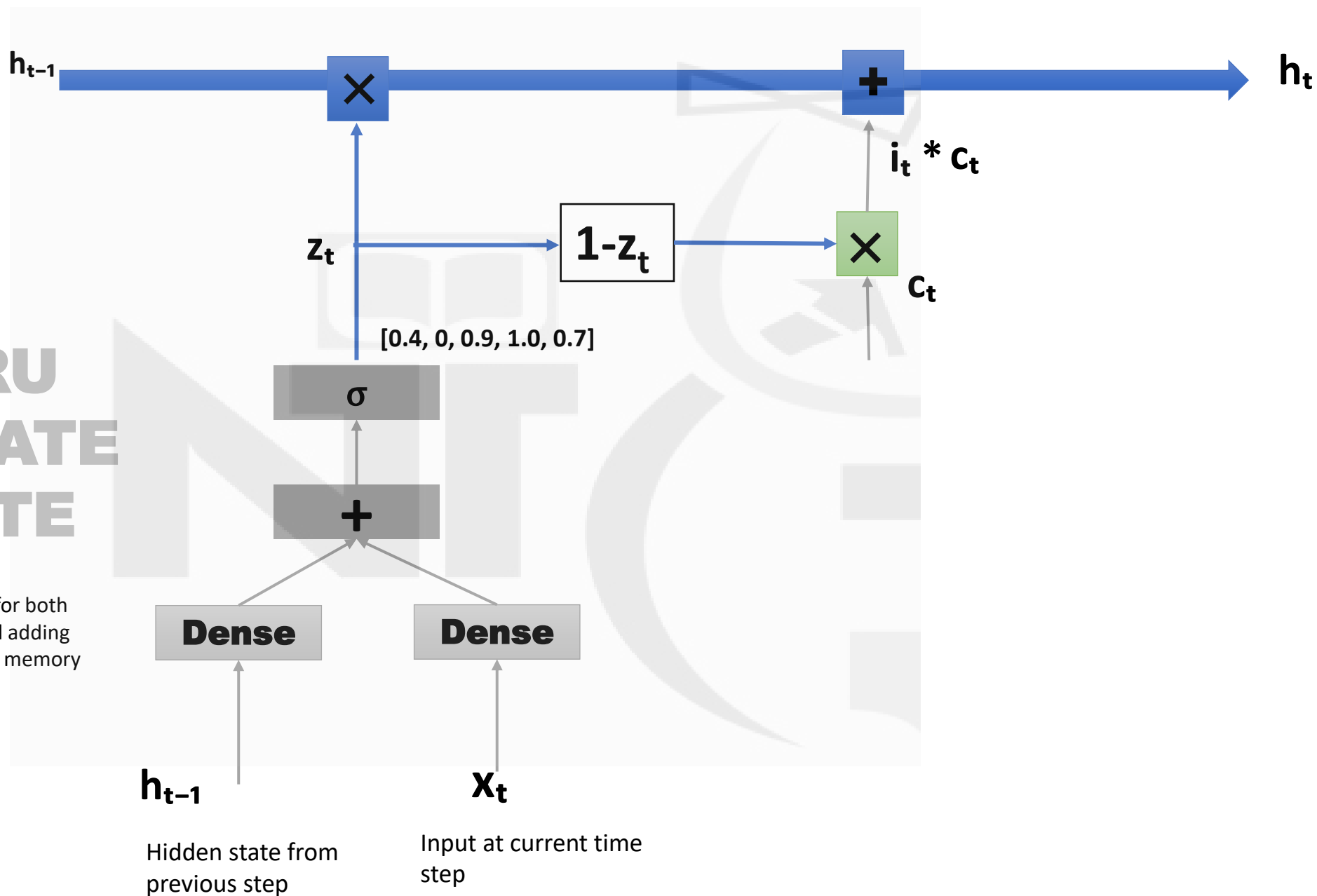
Controls what to read from memory

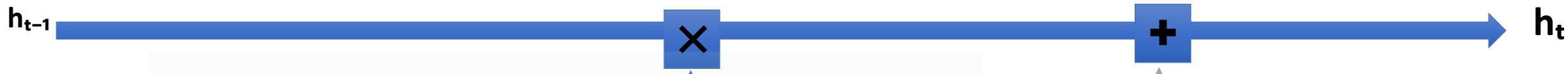
GRU VS LSTM

- GRU only has one memory h_t
- GRU doesn't have a add gate , it uses 1-forget for add
- This reduces 2 dense layers making is lighter than LSTM, needs less computational power

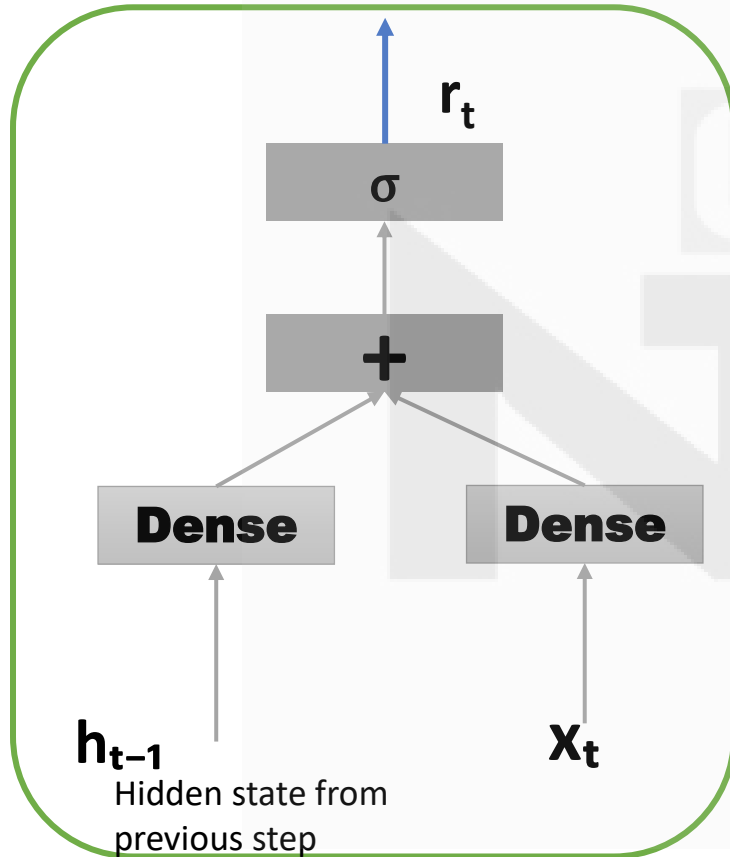
GRU UPDATE GATE

Responsible for both
removing and adding
information to memory



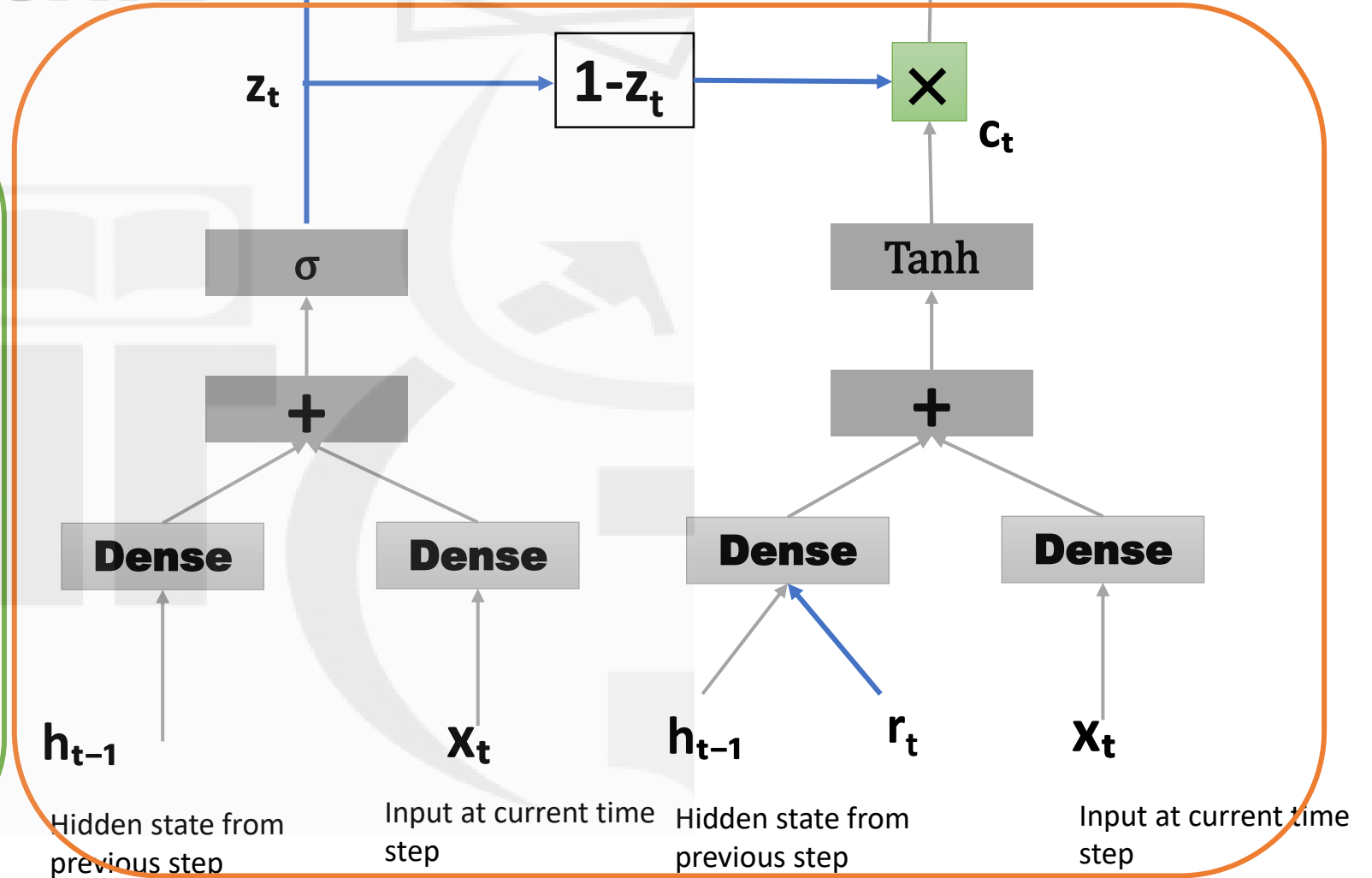


GRU UPDATE GATE



RESET GATE

Controls what things to use from long term memory to build add/update list



UPDATE GATE

Controls what things to forget and add/update to memory

Using GRU in Keras

```
model.add(tf.keras.layers.GRU(256))
```