Input Gate: What to Remember?

Example: "The dog was sleeping ..."

Inputs:

h_{t-1}: Previous output

x_t: Current word ("sleeping")

Input Gate

$i_t = \sigma(\dots)$

How much?

Candidate

$\tilde{C}_t = \tanh(\dots)$

What info?

Decision:

"dog" subject 90%

"sleeping" state 80%

New info to add:

$i_t * \tilde{C}_t$

Intuition: Remember "dog is sleeping" for future predictions