I = Input

STM = Short-Term Memory

LTM = Long-Term Memory

RemL: Long-Term to remember

PerL: % Potential Memory to remember

PotL: Potential Long-Term Memory

PerS: % Potential Memory to remember

PotS: Potential Long-Term Memory

RemL = STM1 \* W1 + I \* W2 + B1

PerL = STM1 \* W3 + I \* W4 + B2

PotL = STM1 \* W5 + I \* W6 + B3

PerS = STM1 \* W7 + I \* W8 + B4

PotS = LTM3

STM1 = STM

LTM1 = LTM

LTM2 = LTM1 \* sigmoid(RemL)

LTM3 = LTM2 + sigmoid(PerL) \* tanh(PotL)

STM2 = sigmoid(PerS) \* tanh(PotS)

STM1 = STM

LTM1 = LTM

LTM2 = LTM1 \* sigmoid(STM1 \* W1 + I \* W2 + B1)

LTM3 = LTM2 + sigmoid(STM1 \* W3 + I \* W4 + B2)

\* tanh(STM1 \* W5 + I \* W6 + B3)

STM2 = sigmoid(STM1 \* W7 + I \* W8 + B4) \* tanh(LTM3)

Loss = STM2 - expected