

LSTM

LONG SHORT TERM MOMERY

หัวข้อ การนำเสนอ

LSTM คืออะไร?

1

การทำงานของ LSTM

2

- Architecture 3
- Forget 4
- Write 5
- Update cell state 6
- Read 8

Demo

10

Conclusion

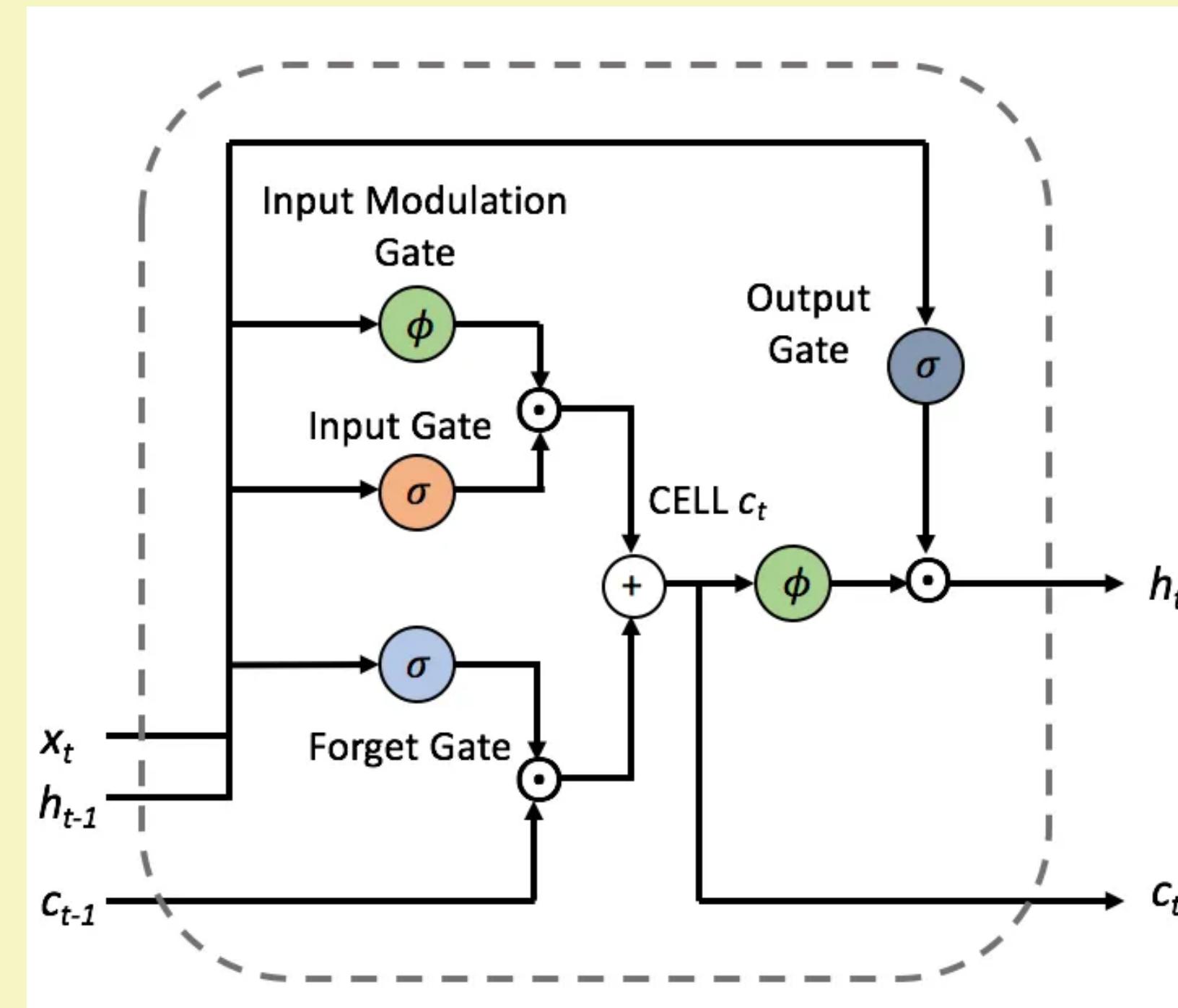
12

LSTM คืออะไร?

- Long Short-Term Memory (LSTM) เป็นหนึ่งใน Architecture ของโครงข่ายประสาทเทียมแบบ Recurrent neural network (RNN) ที่ออกแบบมาเพื่อแก้ปัญหา vanishing gradient ใน RNN

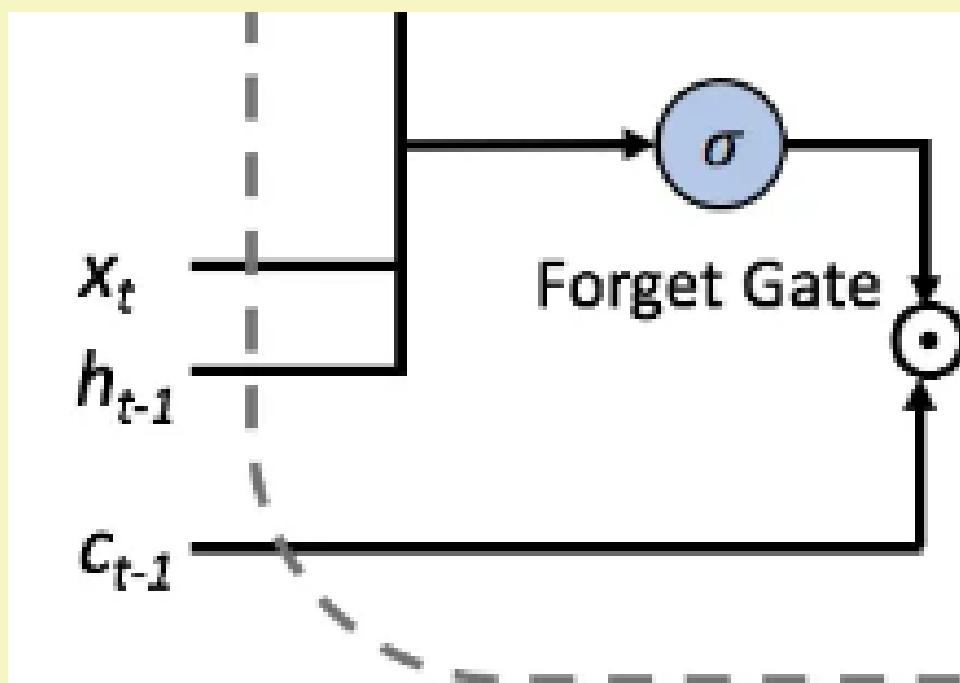
การทำงานของ LSTM

ARCHITECTURE



FORGET

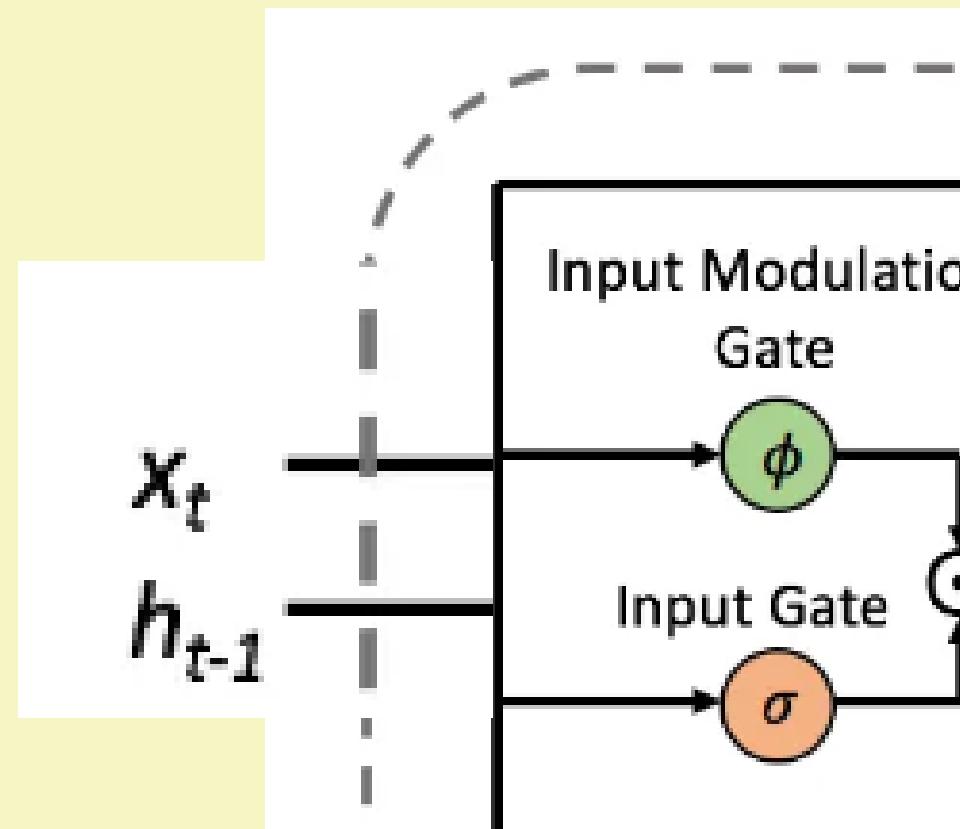
- การ FORGET คือการล้าง CELL STATE เดิมออกไป เคลียร์ข้อมูลเพื่อรับข้อมูลใหม่
- FORGET GATE ทำหน้าที่เป็นตัวตัดสินใจว่าจะลบ CELL STATE เดิมออก หรือจะยังเก็บ CELL STATE นี้ต่อไป
- ใช้ SIGMOID FUNCTION ในการตัดสินใจ



$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$$

WRITE

- INPUT GATE เป็นตัวตัดสินว่าจะอนุญาตให้อัพเดทหรือไม่
จะใช้ SIGMOID FUNCTION
- INPUT MODULATION GATE เป็นตัวจัดการค่าที่จะอัพเดท
จะใช้เป็น TANH FUNCTION

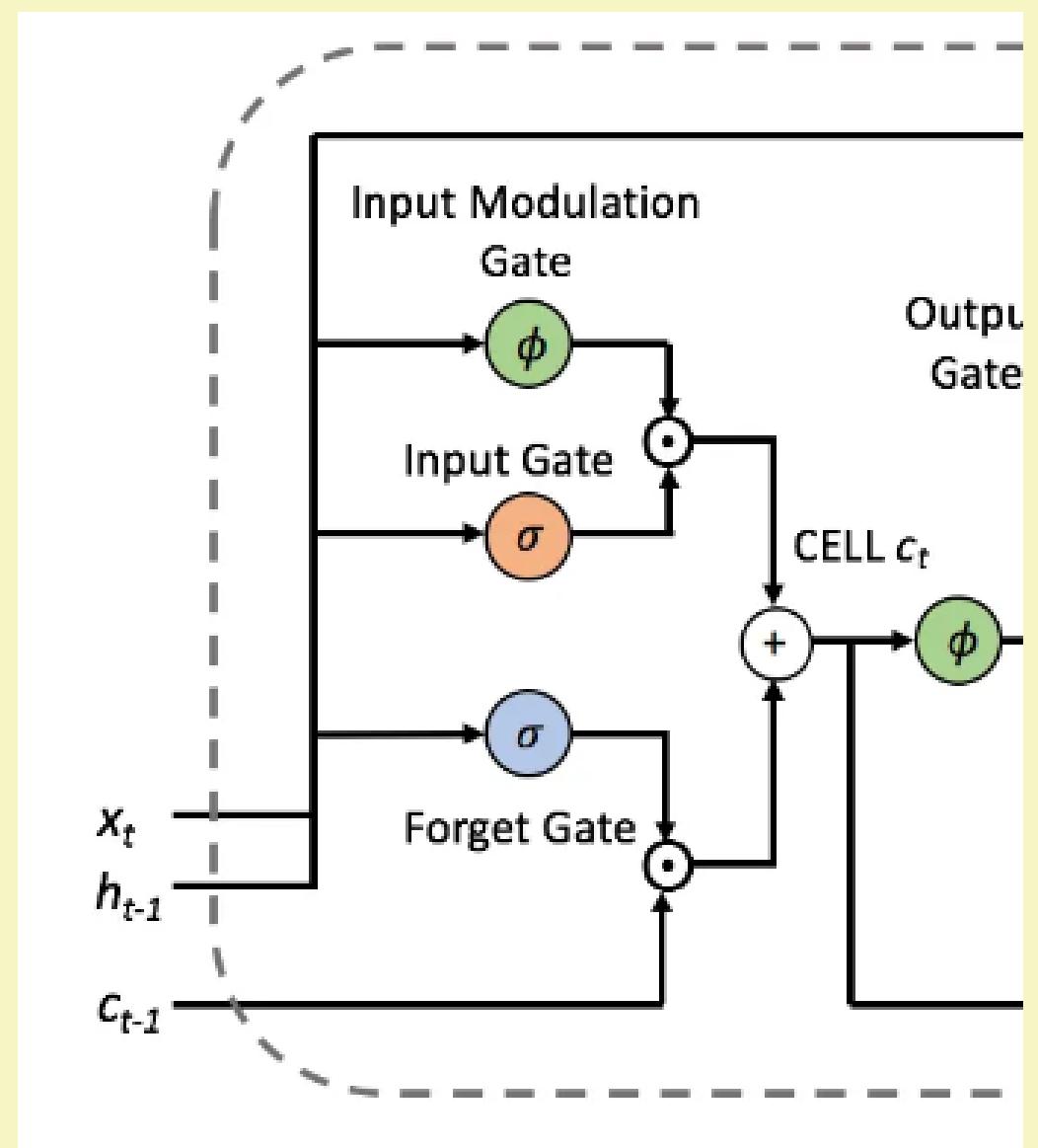


$$g_t = \tanh(W_{x^c}x_t + W_{h^c}h_{t-1} + b_c)$$

$$i_t = \sigma(W_{x^i}x_t + W_{h^i}h_{t-1} + b_i)$$

UPDATE CELL STATE

CELL STATE จะนำข้อมูลจาก FORGET GATE, INPUT GATE และ INPUT MODULATION GATE



$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$

UPDATE CELL STATE(ต่อ)

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$

ถ้า FORGET GATE ตัดสินใจให้ลบ CELL STATE
ค่า FT ก็จะมีค่าเท่ากับ 0 เราจึงไม่เอาค่า CT-1 มาประกอบการอัพเดท
แต่ถ้า FT มีค่าเป็น 1 เราจึงเอาค่า CT-1 มาใช้ในการคำนวณ

ถ้า INPUT GATE ตัดสินใจไม่อัพเดทค่า CELL STATE
ค่า IT ก็จะมีค่าเท่ากับ 0 เราจึงไม่เอาค่า GT มาประกอบการอัพเดท
แต่ถ้า IT มีค่าเป็น 1 เราจึงเอาค่า GT มาใช้ในการคำนวณ

READ

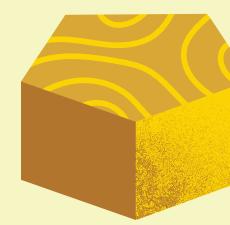
OUTPUT GATE ตัดสินใจว่าจะส่งค่า HT ต่อหรือไม่
โดยใช้ SIGMOID FUNCTION กับค่า HIDDEN STATE ตัว
ก่อนหน้า กับ INPUT DATA ในการตัดสินใจ

$$o_t = \sigma(W_{x^o}x_t + W_{h^o}h_{t-1} + b_o)$$

READ(ต่อ)

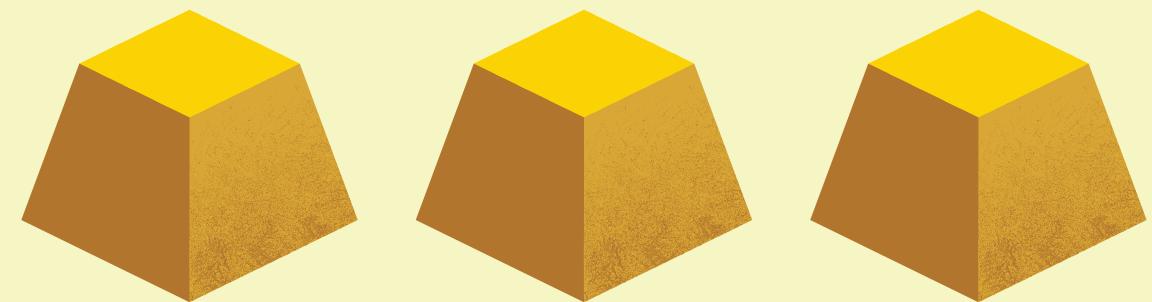
$$h_t = o_t \odot \tanh(c_t)$$

OUTPUT GATE ให้ตัว OT มีค่าเป็น 0 ค่าของ HT ก็จะมีค่าเป็น 0 เช่นกันหรือก็คือไม่มีการส่งค่าใดๆ ออกไป
ในขณะเดียวกัน ถ้า OT มีค่าเป็น 1 เรา ก็จะคำนวณค่า HT และส่งออกไป



DEMO

<https://www.kaggle.com/nattasitnts/lstm-v2/edit?fbclid=IwAR0GoZ6wb63c2MCpefBUWBzzVGgawFyr1w17bvGnsyioldX05PWo9yZtqQ0>



FIRST GOP DEBATE TWITTER SENTIMENT

- ใช้สำหรับศึกษาการทำงานของโมเดล LSTM จากรูปแบบข้อมูลที่เป็น TEXT ในการด้านการศึกษาและนำมาต่อยอดไปยังการทำข้อมูลในรูปแบบอื่นๆ
- ใช้ค้นคว้าหลักการทำงานและการปรับฝึกการปรับ PARAMETER เพื่อให้ได้ค่าความแม่นยำที่มากขึ้น
- เพื่อนำข้อมูลการ TRAINING มาปรับใช้ในด้านการความปลอดภัยใน คัดกรองคำเชิงลบและเชิงบวกจาก SOCIAL NETWORK

CONCLUSION

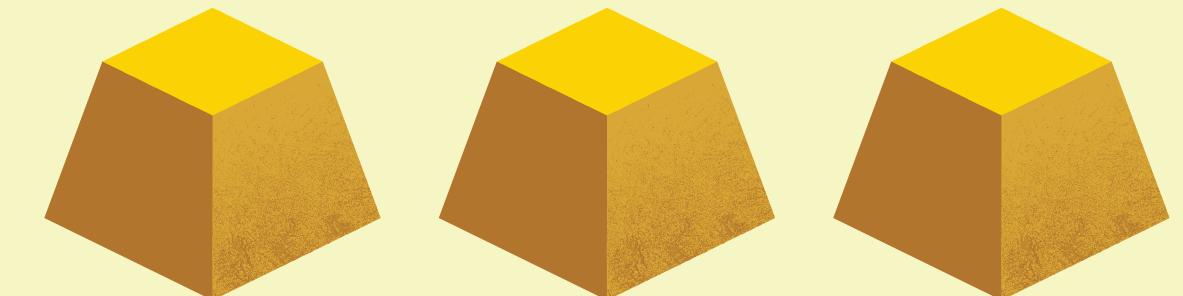
- LSTM มีประสิทธิภาพในการประมวลผลและทำนายข้อมูลแบบ Sequence
- ความสามารถในการเก็บรักษาข้อมูล
- เหมาะสำหรับงานที่เกี่ยวข้องกับการทำนายชุดข้อมูลลำดับต่อๆ กัน

สมาชิก

1.นาย ณัฐสิทธิ์ หมานลະงู 6410210104

2.นาย พาร์ดี เกปัน 6410210235

3.นาย สันเพชร แซ่ฟัง 6410210319



ຂອງគ្មាន