

# STOCK NEWS SENTIMENT ANALYSIS



# วัตถุประสงค์

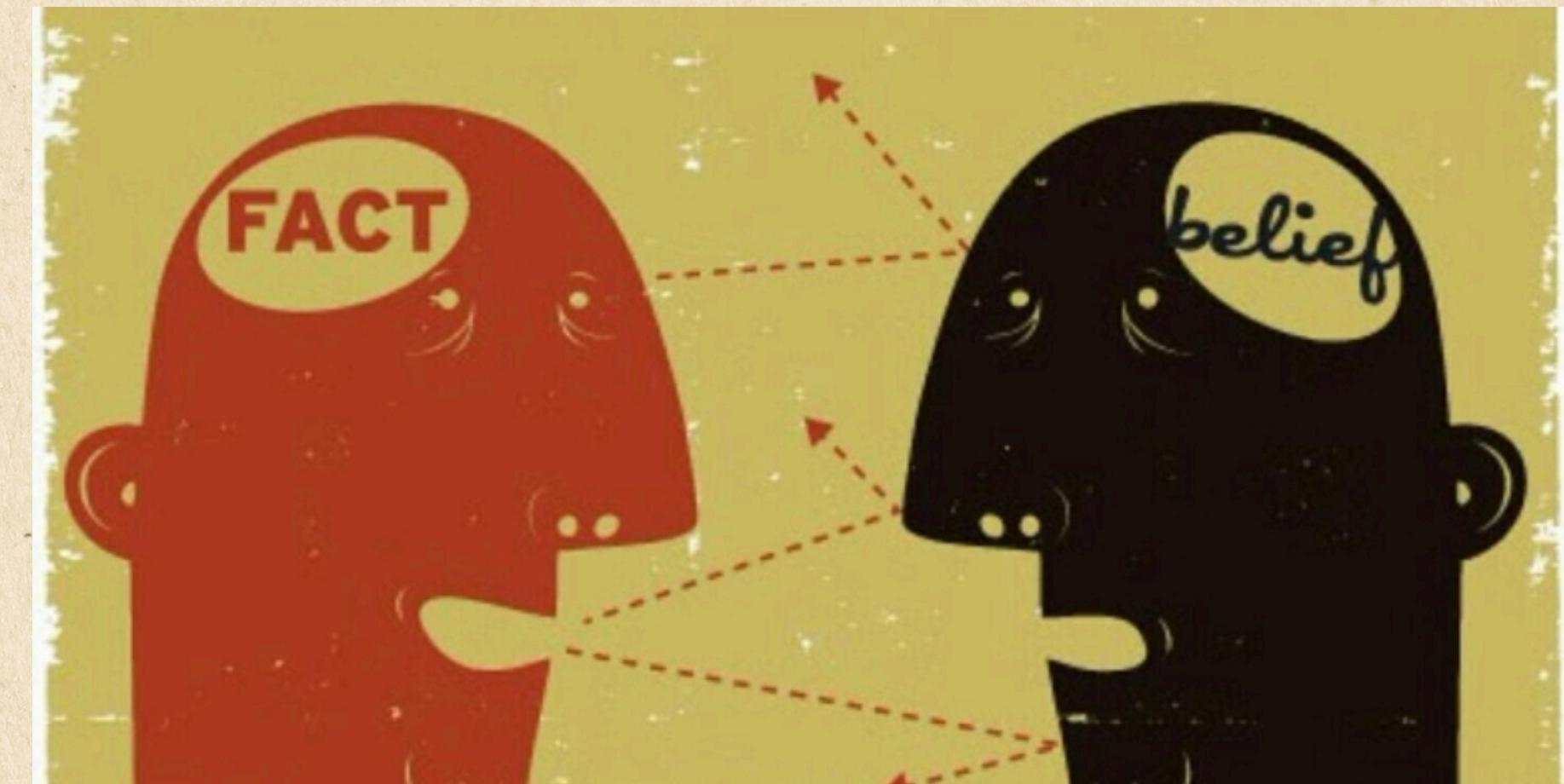
- ใช้ Machine Learning (Deep Learning) เพื่อ วิเคราะห์อารมณ์ของข้อความเกี่ยวกับหุ่น
- กำหนดว่าข้อความเป็น เชิงบวก หรือ เชิงลบ (ถ้าเป็นกลางก็อ้วว่าเป็นลบ)



# Problem statement

---

- ต้องการ วิเคราะห์อารมณ์ของข้อความที่เกี่ยวข้องกับหุ่น
- ปัญหาที่มักพบเจอ:
  - บ่าวหุ่นมักใช้คำศัพท์เฉพาะ
  - ไม่มีพื้นฐานด้านการลงทุน
  - ตีความบ่าวผิดจากความเป็นจริง



# ขั้นตอนเตรียมข้อมูล

---

Step 1 โหลดข้อมูล: อ่านไฟล์ CSV

Step 2 ทำความสะอาดข้อมูล: ลบค่า NaN, ลบอักษรพิเศษ,  
แปลงข้อความให้เป็นตัวพิมพ์เล็ก

Step 3 Tokenization & Padding : แปลงข้อความเป็นตัวเลข  
และปรับขนาดของประโยคให้เท่ากัน

Step 4 ใช้ GloVe Embeddings เพื่อเพิ่มประสิทธิภาพการเรียนรู้ของโมเดล

Step 5 แบ่งข้อมูล: 80% ใช้สำหรับ Train 20% ใช้สำหรับ Test

---

Tokenization: แปลงข้อความเป็นตัวเลข (ใช้ Tokenizer ของ Keras)

Padding: ทำให้แต่ละประโยค มีความยาว 100 คำ เท่ากัน

Ex "หุนวันนี้ขึ้น 10%" → [345, 2, 980] (แปลงเป็นเวกเตอร์)



# Model

Embedding Layer

Input Dimension: 5000

Output Dimension: 100

Pretrain Model: GLoVe

Bidirectional LSTM

Units: 128

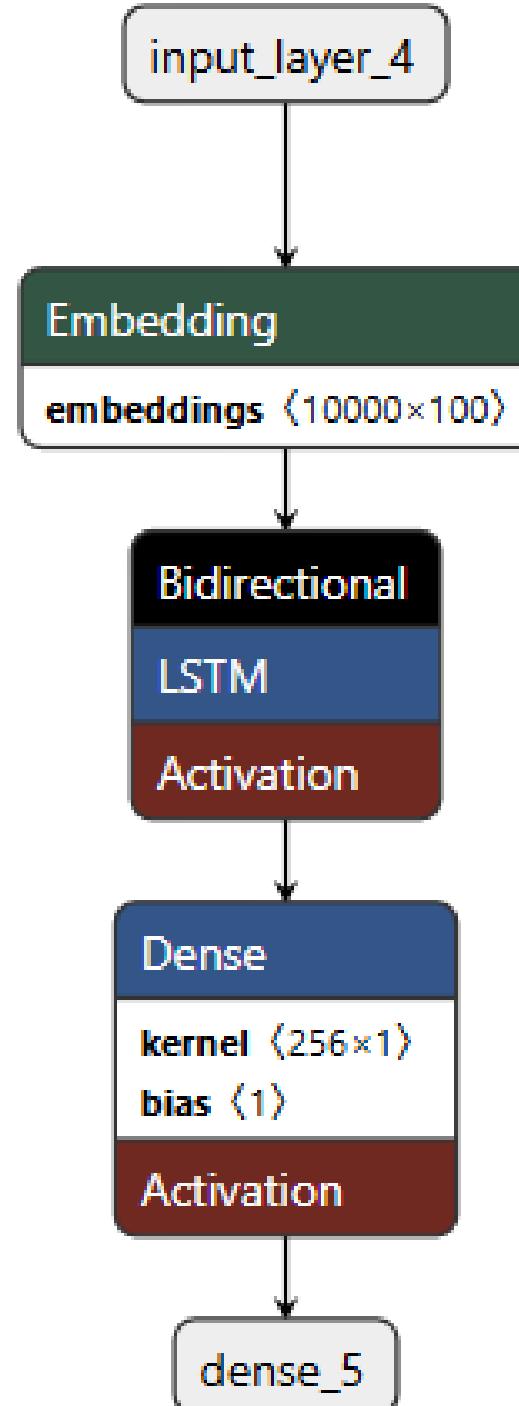
Dropout: 0.3

Recurrent Dropout: 0.3

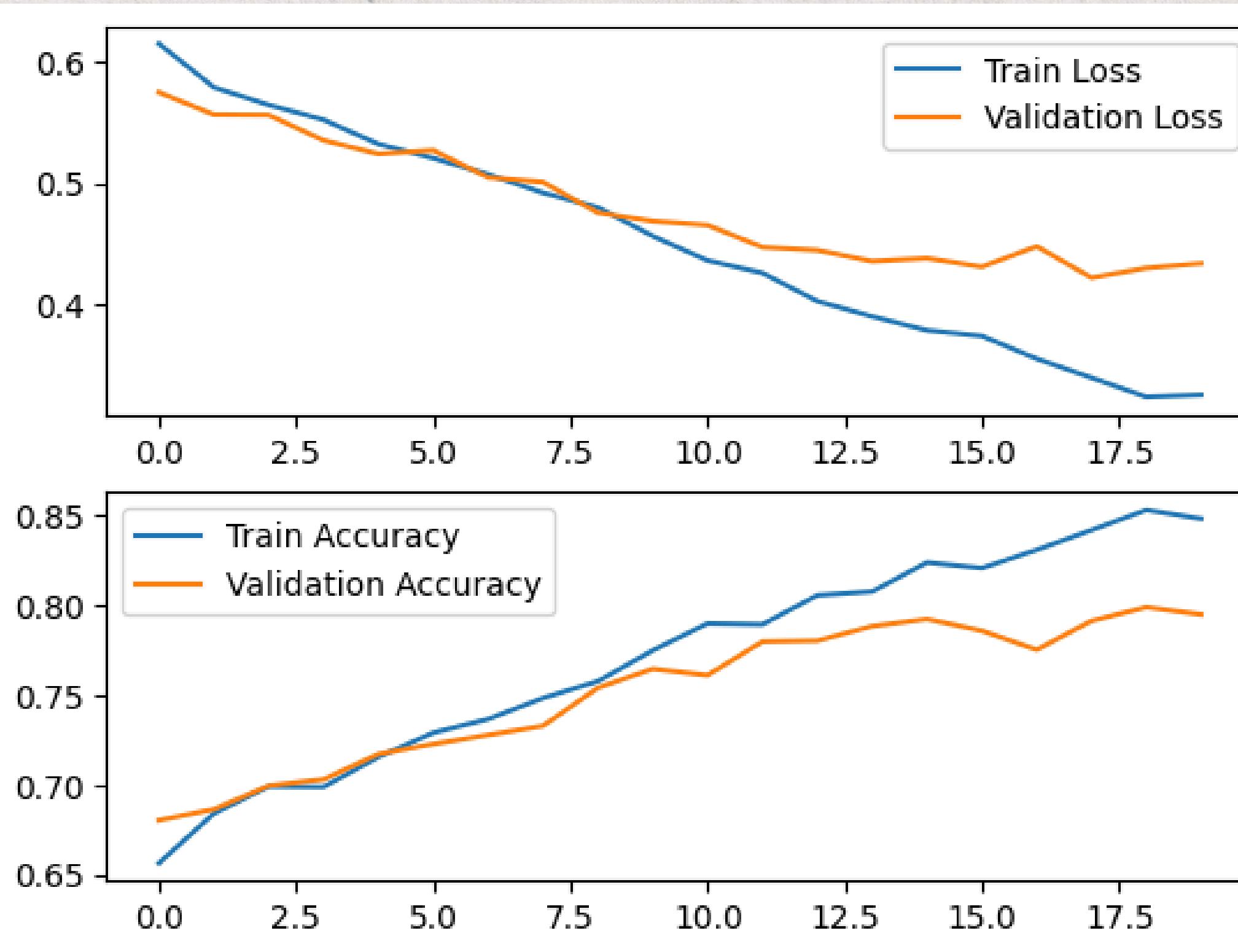
Dense

Units: 1

Activation: Sigmoid



# Train Validate Loss



## ผลการฝึกโมเดล

- Deep Learning : Bidirectional LSTM

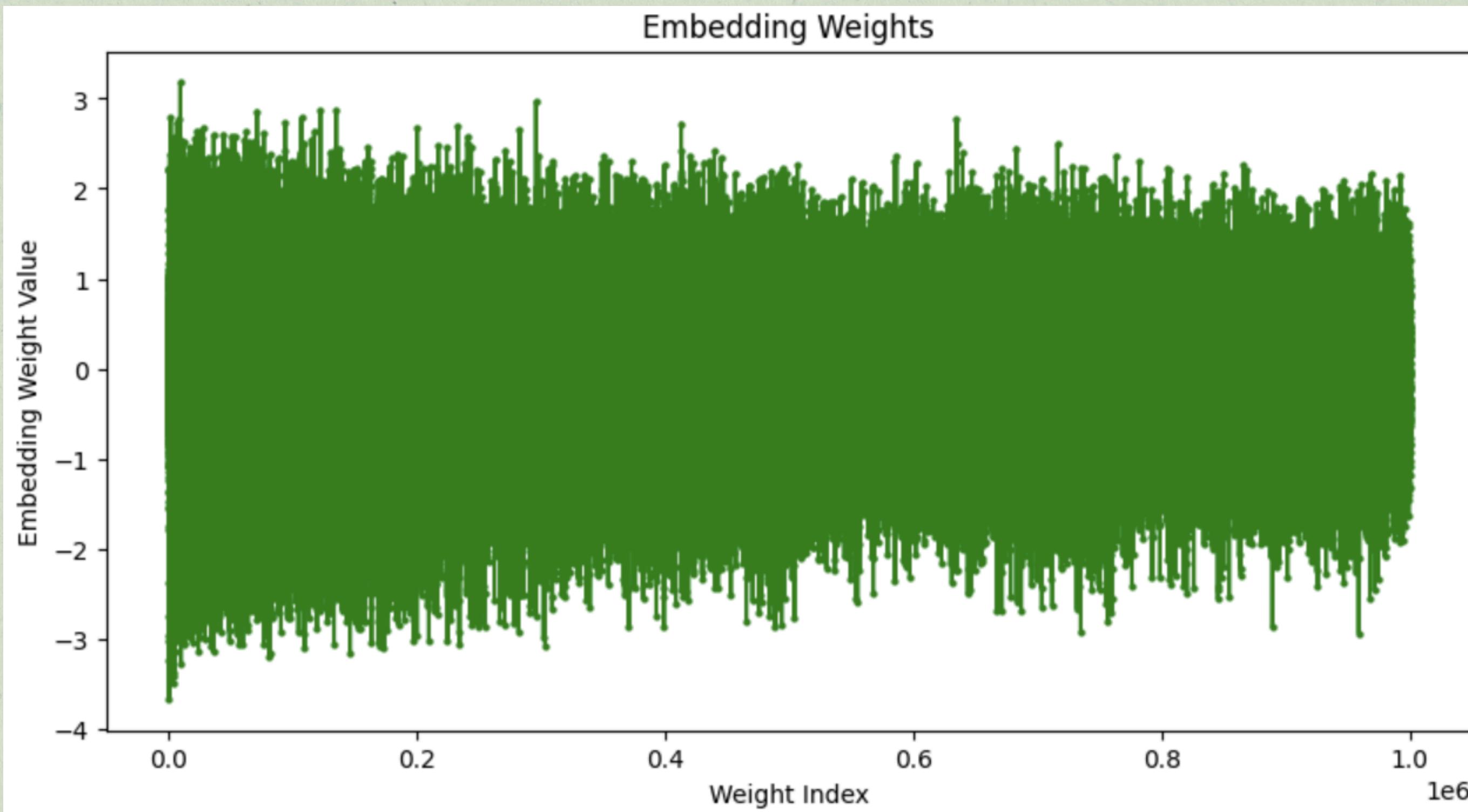
## Test Loss

- ค่าที่ได้คือประมาณ 0.4335

## Test Accuracy

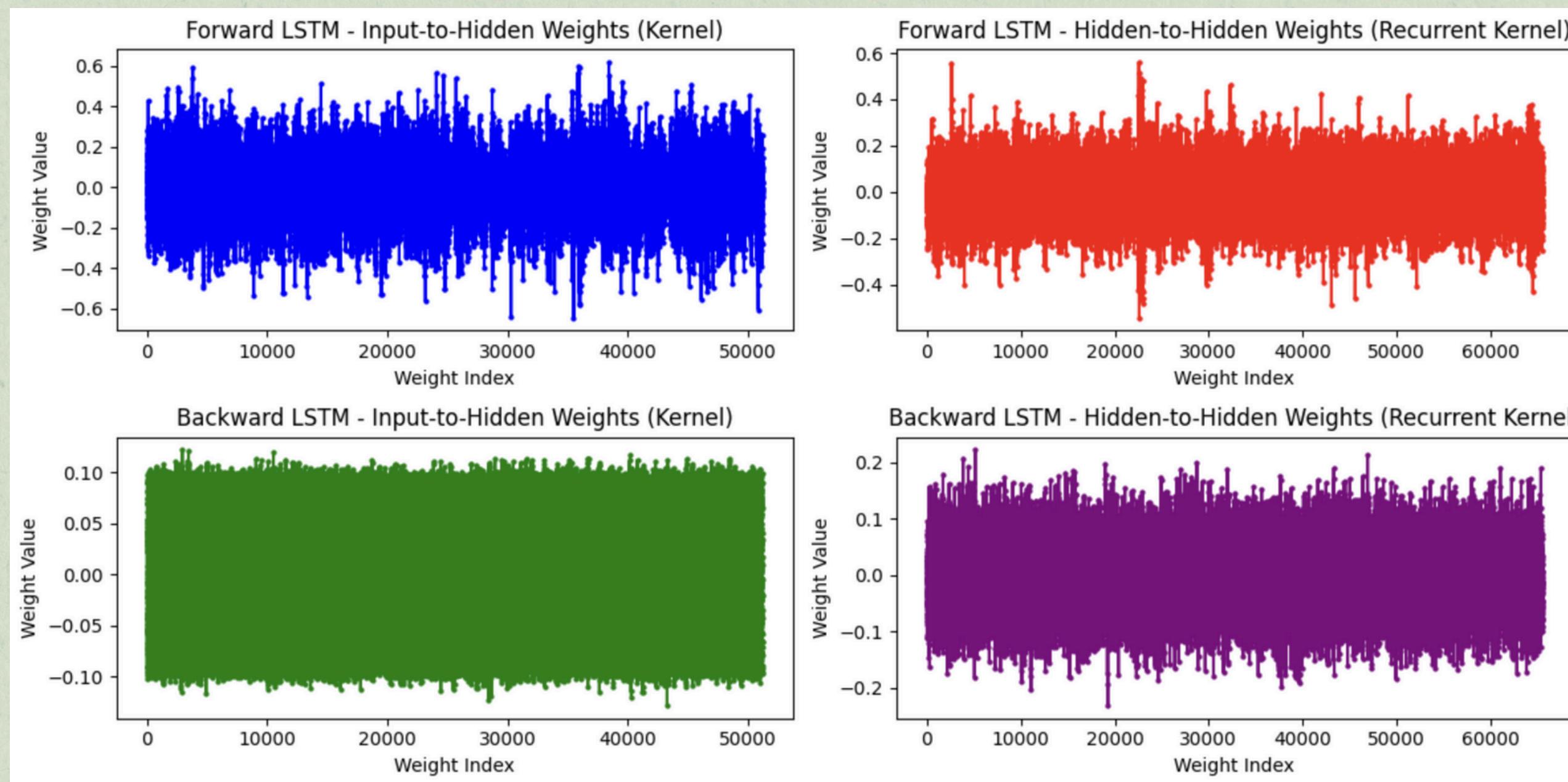
- ค่าที่ได้คือประมาณ 0.794

# กราฟค่า Weights



แสดงค่าค่วงน้ำหนัก (Weights) ที่ Embedding Layer เรียนรู้

# กราฟค่า Weights



กราฟของน้ำหนักจาก Bidirectional LSTM Layer ที่ใช้ในโมเดล

Forward LSTM - Input-to-Hidden Weights      **-0.6 ถึง 0.6**

Forward LSTM - Hidden-to-Hidden Weights      **-0.4 ถึง 0.4**

Backward LSTM - Input-to-Hidden Weights      **-0.1 ถึง 0.1**

Backward LSTM - Hidden-to-Hidden Weights      **-0.2 ถึง 0.2**

---

# THANK YOU FOR LISTENING!

---

Members



---

នកស រកា

នាកស នាការ