

Trả lời câu hỏi

1. (NGUYEN HAI YEN 20195944) Các bạn giải thích giúp mình cách biểu thức hoạt động của Forget/Input/Output Gate và Cell State ạ

Trả lời:

- Forget Gate (Cổng quên): Cổng quyết định xem thông tin nào chúng ta sẽ cho phép đi qua ô trạng thái. Đầu tiên nó nhận đầu vào là 2 giá trị h_t và x_t và trả về một giá trị nằm trong khoảng 0 và 1 cho mỗi giá trị của ô trạng thái C_{t-1} . Nếu giá trị bằng 1 thể hiện ‘giữ toàn bộ thông tin’ và ngược lại bằng 0 thể hiện ‘bỏ qua toàn bộ chúng’.
- Input Gate (Cổng vào): Cổng vào giúp quyết định bao nhiêu lượng thông tin đầu vào sẽ ảnh hưởng đến trạng thái mới. Quyết định bằng cách thông qua đặc điểm của hàm sigmoid (đầu ra nằm trong khoảng $[0,1]$), như vậy khi một vector thông tin đi qua đây, nếu nhân với 0, vector sẽ bị triệt tiêu hoàn toàn. Nếu nhân với 1, hầu hết thông tin sẽ được giữ lại.
- Output Gate (Cổng ra): Cổng điều chỉnh lượng thông tin có thể ra ngoài y_t và lượng thông tin truyền tới trạng thái tiếp theo.
- Cell State: (Ô trạng thái) là một dạng băng chuyền chạy thẳng xuyên suốt toàn bộ chuỗi với chỉ một vài tương tác tuyến tính nhỏ giúp cho thông tin có thể truyền dọc theo đồ thị mạng nơ ron ổn định. Cell state sẽ mang thông tin nào quan trọng truyền đi xa hơn và sẽ được dùng khi cần.

2. (NGUYEN THI DUYEN 20195866) Cấu trúc mô hình LSTM của nhóm

bạn có bao nhiêu lớp, mỗi lớp có bao nhiêu nơ ron?

Trả lời : Cấu trúc LSTM có 4 lớp và chúng tương tác với nhau. Mỗi lớp có 32 nơron.

3. (NGUYEN THI DUYEN 20195866) **Hàm kích hoạt nhóm bạn sử dụng là hàm gì? Và tại sao**

Trả lời: Hàm kích hoạt trong mô hình LSTM là 'relu'. Mục đích của sử dụng Hàm kích hoạt tuyến tính được chỉnh lưu (ReLU) là cho phép mạng thần kinh học các phụ thuộc phi tuyến tính.

4. (NGUYEN NGOC QUANG 20185395) **Tại sao các bạn lại để tham số epoch là 200 ạ?**

Trả lời: Để số epoch 200 là tính đến trường hợp huấn luyện mô hình đã được trang bị "early stop" sẽ dừng huấn luyện mô hình khi đạt đủ điều kiện đặt ra, nên số 200 là lấy số lớn cũng có thể lấy 100 hoặc 300.

5. (NONG THI THUY 20195926) **Mô hình bạn bảo không tốt cho lắm, thì bạn định hiệu chỉnh các tham số mô hình của VARMAx và LSTM như nào để cải thiện hiệu suất dự báo?**

Trả lời: Với mô hình LSTM và VARMAX thì nhóm mình chưa có xét đến trường hợp mà các thuộc tính không tương quan với nhau hay độ tương quan thấp. Nên để cải thiện mô hình mình có thể tính đến trường hợp khi dự đoán một thuộc tính nào đó thì loại bớt các thuộc tính không tương quan để giảm nhiễu cho mô hình.

6. (HOANG MY GIA HUY 20195887) **nhóm bạn xây dựng tập train và test cho hồi quy tuyến tính như thế nào?**

Trả lời: Nhóm mình sử dụng hồi quy tuyến tính để xử lý missing data cho

từng cột bằng cách xây dựng

- Tập train: tất cả dòng có dữ liệu (notnull)
- Tập test: các dòng bị thiếu dữ liệu (isnull)

7. (HOANG MY GIA HUY 20195887) **nhóm bạn xử lý chuỗi không dừng cho mô hình VARMAX như thế nào?**

Trả lời: Với chuỗi không dừng thì có thể biến đổi sang dừng bằng cách lấy sai phân differencing.

8. (HOANG MY GIA HUY 20195887) **Biến ngoại lai cho mô hình VARMAX của nhóm bạn là gì?**

Trả lời: Có nhiều cách để xác định biến ngoại lai trong mô hình VARMAX như: kiểm tra giá trị tuyệt đối của các quan sát, kiểm tra khoảng cách mahalanobis, hay sử dụng đường cong roc.

9. (TRAN THI HONG VAN 20195941) **Nhóm đang sử dụng hai mô hình dự báo cho đa chuỗi vậy trong quá trình dự đoán của mô hình giữa các chuỗi tương quan và các chuỗi không tương quan thì kết quả nhóm đánh giá dự đoán nào tốt hơn?**

Trả lời : Theo nhóm mình thì dự đoán chuỗi tương quan sẽ tốt hơn.

10. (NGUYEN THI XUAN HONG 20185364) **bạn sử dụng bn dữ liệu quá khứ để dự đoán 48h tiếp theo? và với số lượng dữ liệu quá khứ khác nhau bạn sử dụng liệu có ảnh hưởng đến kết quả b thực nghiệm k?**

Trả lời : Nhóm dùng 1000 điểm dữ liệu. Việc lấy 1000 điểm dữ liệu là do nhóm mình muốn thử với nhiều điểm dữ liệu thì liệu mô hình có được tốt.

Nhưng khi huấn luyện mô hình xong nhóm mình rút ra được việc lấy nhiều dữ liệu một phần làm cho mô hình bị giảm hiệu suất khi đưa ra dự đoán.

11. (NGUYEN LAN HUONG 20195884) **bạn xử lý dữ liệu đầu vào như nào để mô hình nó hội tụ?**

Trả lời: Nhóm mình chỉ dùng mô hình linear regression để dự đoán các giá trị thiếu. Còn lại không xử lý gì thêm với dữ liệu.

12. (LE VAN THAM 20195914) **Theo mình biết, đầu vào của của các bạn là đa biến, liệu nhiều biến đầu vào như vậy mô hình lstm có bị nhớ nhầm chuỗi này sang chuỗi khác hay không?**

Trả lời: LSTM là phương pháp dự báo chuỗi thời gian đa biến được xây dựng lên để áp dụng cho chuỗi thời gian đa biến, nên mình nghĩ nó sẽ không nhớ nhầm các chuỗi với nhau.

13. (VU VAN XUNG 20195943) **Tại sao khi dùng LSTM cho đa biến lại không khớp được mô hình?**

Trả lời: Theo nhóm mình thì vấn đề sử dụng LSTM chưa khớp được với mô hình là do dữ liệu vẫn chưa loại bớt các thuộc tính không tương quan, input shape đầu vào cho mô hình chưa thực sự phù hợp, các hyperparameter cũng chưa được chọn lựa và hiệu chỉnh một cách chính xác.

14. (NGUYEN THAI THINH 20195921) **Các bạn có thể dựa vào kết quả để so sánh 2 phương pháp này không? Ưu nhược điểm từng pp**

Trả lời : Dựa vào kết quả ta có thể thấy VARMAX là tốt hơn LSTM trong trường hợp xây dựng mô hình của nhóm. về ưu nhược điểm của từng phương pháp:

- Với LSTM thì nhóm xây dựng được mô hình cho kết quả test không

được khớp lắm với dữ liệu test nhưng cả MSE và RMSE đều thấp.

- Với VARMAX thì cho kết quả test độ chính xác cao và khớp với bộ test nhưng lại có MSE khá cao, thời gian huấn luyện lâu.

15. (BUI KHUONG DUY 20195864) **Về mô hình LSTM, tính giải thích đầu ra của mô hình là gì nhỉ, bạn bỏ qua mất phần đây**

Trả lời: Kết quả ở đầu ra sẽ dựa trên ô trạng thái, nhưng sẽ là một phiên bản được lọc. Đầu tiên, chúng ta chạy qua một tầng sigmoid nơi quyết định phần nào của ô trạng thái sẽ ở đầu ra. Sau đó, ô trạng thái được đưa qua hàm tanh (để chuyển giá trị về khoảng -1 và 1) và nhân nó với đầu ra của một cổng sigmoid, do đó chỉ trả ra phần mà chúng ta quyết định.

16. (NGUYEN VAN DAI 20195847) **Tại sao các bạn lại dùng 1000 điểm dữ liệu (tại sao lại dùng nhiều điểm dữ liệu như vậy - 1000 điểm chiếm đến 1/3 bộ dữ liệu rồi) liệu có ảnh hưởng đến hiệu suất của mô hình không**

Trả lời: Việc lấy 1000 điểm dữ liệu là do nhóm mình muốn thử xem việc lấy nhiều dữ liệu để mô hình thì liệu nó có chính xác. Nhưng khi huấn luyện mô hình xong nhóm mình rút ra được việc lấy nhiều dữ liệu một phần làm cho mô hình bị giảm hiệu suất.