

DEEP LEARNING FOR TIME-SERIES FORECASTING

1. Time-Series Forecasting

Dự báo là bài toán quan trọng nhất của data mining. Để có một kết quả dự báo, các nhà nghiên cứu phải xây dựng được một mô hình dự báo phù hợp từ những dữ liệu thu thập được. Trong đó, dự báo chuỗi thời gian là việc sử dụng một mô hình để dự đoán các giá trị tương lai dựa trên các giá trị đã quan sát trước đó. Bài toán này quan sát thêm một yếu tố là thời gian, dự báo thông tin trong tương lai. Có nghĩa là dựa vào các thông tin đã biết (các giá trị trước đó và hiện tại), các nhà nghiên cứu sẽ xây dựng một số mô hình có khả năng dự báo trước các thông tin cần biết ở một thời điểm xác định ($t+1, t+2, \dots$). Tức là tại một thời điểm t nào đó, một giá trị sẽ được dự đoán là bao nhiêu để phục vụ một lợi ích nào đó.

Có nhiều phương pháp được đưa ra để giải quyết bài toán này như phân lớp, hồi quy, suy diễn mờ, ... Trong đó, một phương pháp gọi là deep learning sẽ giúp chúng ta giải quyết bài toán này một cách chính xác nhất có thể được áp dụng rộng rãi trong những năm gần đây.

2. Deep Learning

Deep Learning là một lĩnh vực nghiên cứu của Machine Learning. Deep Learning là một cách học với nhiều cấp độ được biểu diễn và rút trích để giúp nhận biết được dữ liệu như hình ảnh, âm thanh và văn bản. Deep learning là một neural network “rộng” và “sâu”. Vậy như thế nào là “rộng” và “sâu”. Với “rộng” là số input đầu vào lớn, thường là một ma trận với số chiều lớn, chủ yếu là số mẫu lớn, mỗi mẫu là một hàng của ma trận. Với “sâu” là có nhiều lớp neural ẩn (hơn hoặc bằng 2). Hay nói cách khác, Deep Learning là một neural network (1) bao gồm nhiều lớp hoặc giai đoạn của xử lý thông tin (2) phương pháp học có giám sát hoặc học không giám sát với số đặc trưng rất lớn. Deep learning là sự giao thoa giữa nhiều lĩnh vực nghiên cứu như mạng neural, AI, mô hình đồ họa, tối ưu, phát hiện mẫu và xử lý tín hiệu.

Deep Learning ra đời từ rất lâu nhưng có một thời gian dài không được các nhà khoa học nghiên cứu phát triển. Chỉ vài năm trở lại đây, Deep Learning mới thực sự hồi sinh

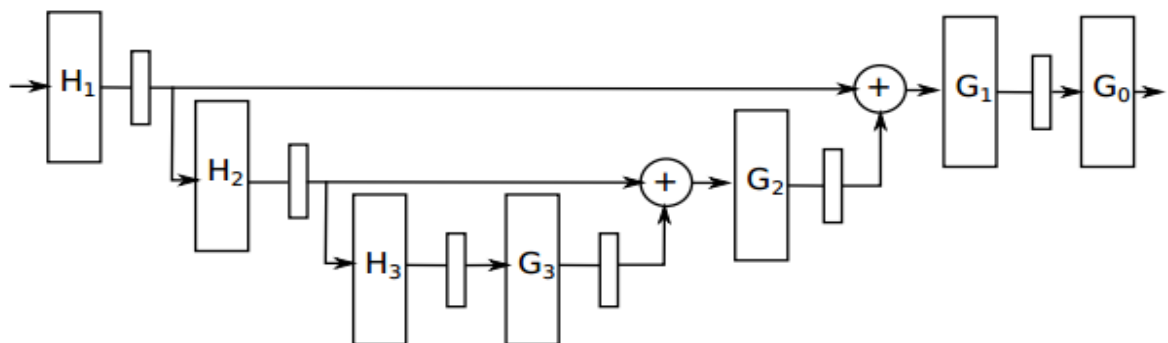
mạnh mẽ. Ba lý do quan trọng để lý giải sự hồi sinh của Deep Learning là (1) sự nâng cao khả năng xử lý của GPU (2) sự tăng lên nhanh chóng của dữ liệu có giá trị (3) sự phát triển của các nghiên cứu về kỹ thuật xử lý trong máy học. Nhờ những sự phát triển vượt bậc này mà kích hoạt các phương pháp Deep Learning để khám phá các vấn đề phức tạp, các hàm không tuyến tính, để học phân tán và phân cấp đặc trưng, rất hiệu quả trong phân lớp có dán nhãn hoặc không dán nhãn. Deep Learning có nhiều kiến trúc khác nhau như FNN, DBN, CNN,... Mỗi kiến trúc sẽ có một sự khác biệt về cách xây dựng, hàm kích hoạt, thuật toán lan truyền,...

Phần tiếp theo chúng ta sẽ khảo sát một số công trình nghiên cứu áp dụng deep learning vào bài toán dự báo chuỗi thời gian.

3. Deep Learning for Time-Series Forecasting

Nhiều công trình nghiên cứu áp dụng deep learning vào time-series forecasting để hy vọng mang lại độ chính xác dự báo cao.

John Gamboa [1] đã trình bày một bài review về Deep Learning và một số ứng dụng trong phân tích chuỗi thời gian. Mục tiêu của tác giả là “khẳng định Deep Learning là một phương pháp có nhiều tiềm năng cho lĩnh vực phân tích chuỗi thời gian này”. Trong quá trình review, tác giả có nhắc đến một kiến mạng thuộc Deep Learning là UFCNN (Undecimated Fully Convolutional Neural Network) [8]. UFCNN có nhiều cải tiến về cách lấy mẫu dữ liệu (upsampling) và tổng hợp lại để suy ra kết luận (pooling).



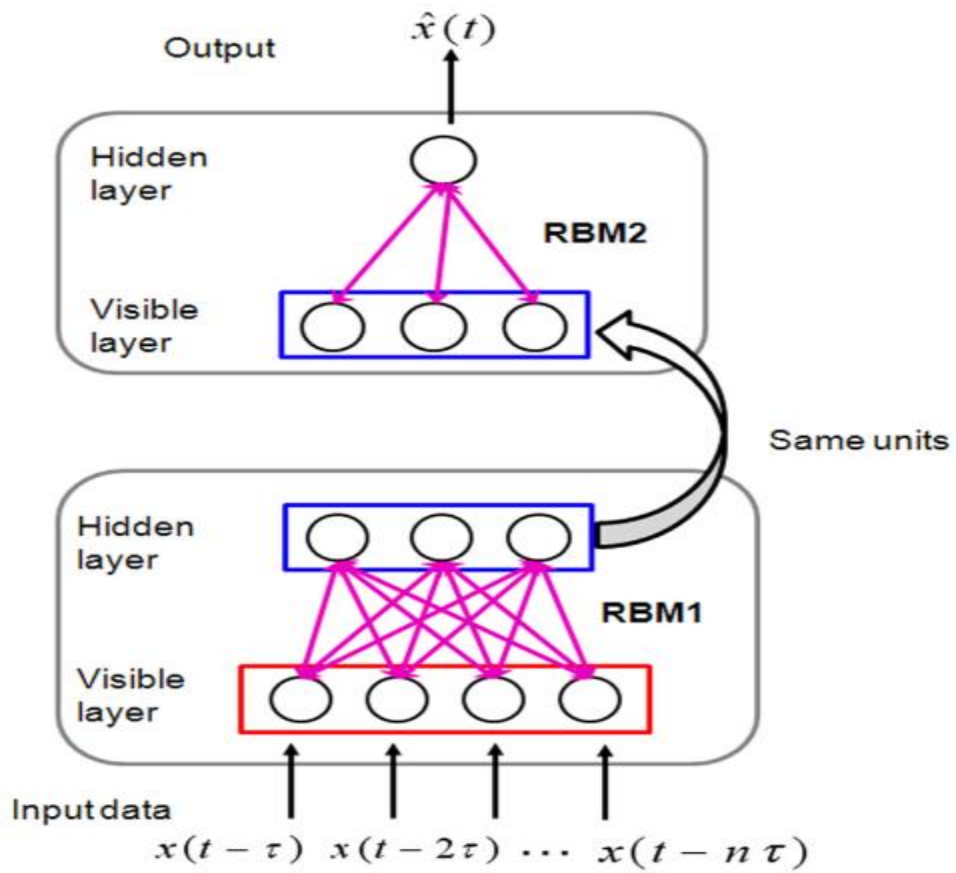
Hình 1. Kiến trúc UFCNN

Riêng UFCNN được thử nghiệm trong 3 thực nghiệm khác nhau để thấy sự hiệu quả của kiến trúc này. Thực nghiệm đầu tiên trên 2000 chuỗi huấn luyện, 50 chuỗi đánh giá, 50 chuỗi kiểm tra mô hình. Thực nghiệm thứ hai được tiến hành trên 2 tập datasets¹ là “MUSE” and “NOTTINGHAM”. Thực nghiệm thứ ba được tiến hành trên một tập dữ liệu trong kinh doanh². Trong các thực nghiệm trên, UFCNN đều cho kết quả tốt hơn các mô hình dự đoán đã có.

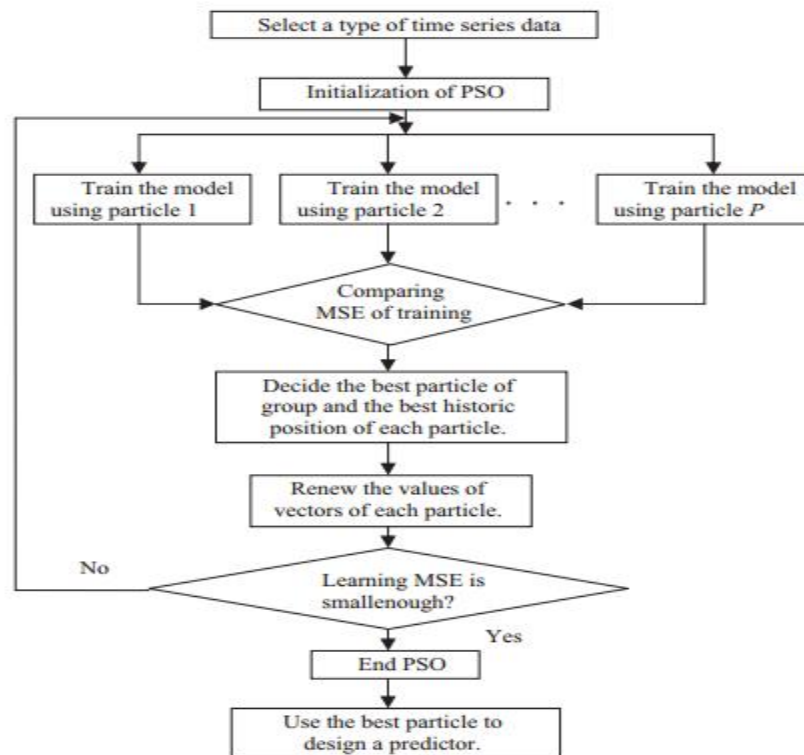
Các kỹ thuật multi-layer perceptron và ANN đã được sử dụng cho time-series forecasting từ những năm 1980s. Takashi Kuremoto et.al. [3] đề xuất một Deep Belief Network (DBN) 3 lớp – là một cải tiến của mạng DBN của Hinton, kết hợp với restricted Boltzmann machines (RBM) để tạo ra một mô hình dự đoán chuỗi thời gian hiệu quả. Mô hình này input là một ma trận với dữ liệu được sắp xếp theo một thời điểm nhất định. Sau đó, một RBM được sử dụng để tạo ra một lớp ẩn khác và lớp ẩn này chính là input của một RBM khác. Và chính RBM này đóng vai trò tính toán và tạo ra output dự báo cuối cùng.

¹ <http://www-etud.iro.umontreal.ca/>

² <http://www.circulumvite.com/home/trading-competition>



Hình 2. Kiến trúc DBN kết hợp RBMs



Hình 3. Quá trình tối ưu mô hình

Các tác giả thực nghiệm trên dữ liệu “The CATS benchmark” và phương pháp này đạt độ chính xác không cao hơn một số phương pháp khác nhưng được đánh giá là rất có tiềm năng để dự báo trong tương lai [3].

Enzo Busseti và đồng sự [4] thực hiện một bài báo cáo tại đại học Stanford. Báo cáo tập trung cho bài toán dự đoán năng lượng tại Hòa Kỳ. Để xây dựng mô hình dự đoán này gồm 2 giai đoạn: giai đoạn tuyến tính, giai đoạn mạng neural. Trong bài báo cáo này, nhóm tác giả chủ yếu trình bày lại các công trình nghiên cứu đã có. Nhóm tác giả nêu cụ thể các vấn đề thách thức trong bài toán dự báo chuỗi thời gian như kích thước dữ liệu lớn, các vấn đề về xây dựng mô hình neural phức tạp, các thao tác lọc đặc trưng và giới thiệu thuật toán huấn luyện. Đóng góp chính của báo cáo này là cài đặt thực nghiệm một số mô hình dự báo chuỗi thời gian và phân tích chúng.

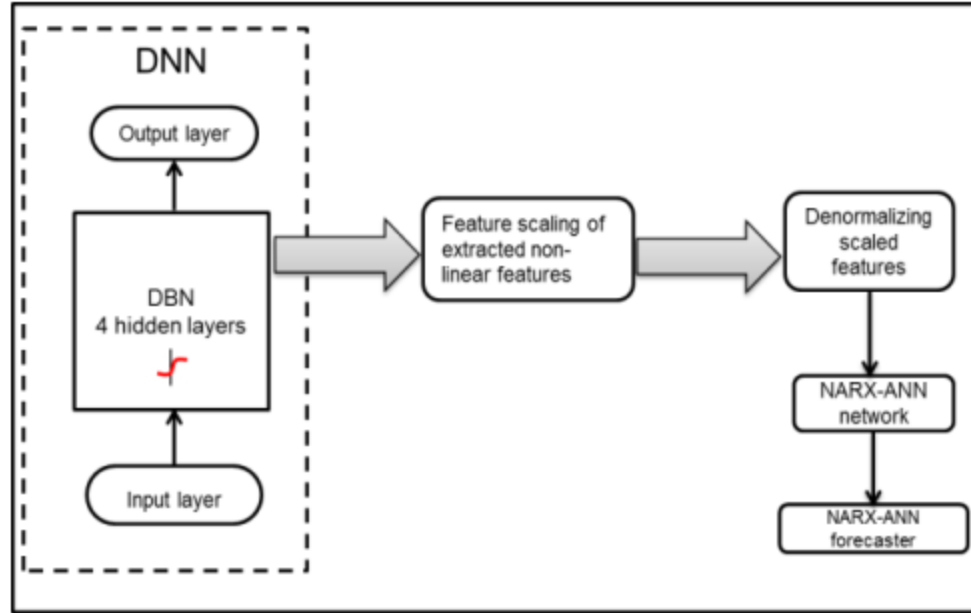
Xueheng Qiu [5] đề xuất phương pháp kết hợp hồi quy và deep learning cho bài toán dự báo. Đề xuất phương pháp ensemble cho deep learning. For regression and time series forecasting, the prediction results can be different when the number of epochs of back propagation training is changed. Therefore, we can combine all the outputs generated by FNNs trained with different number of epochs. Chi tiết gồm có 4 bước:

Bước 1: Huấn luyện DBN với input là ma trận đầu vào.

Bước 2: Duyệt qua toàn bộ dữ liệu input từ 100 đến 2000 lần. Mỗi lần cập nhật lại tham số và khởi tạo lại DBN.

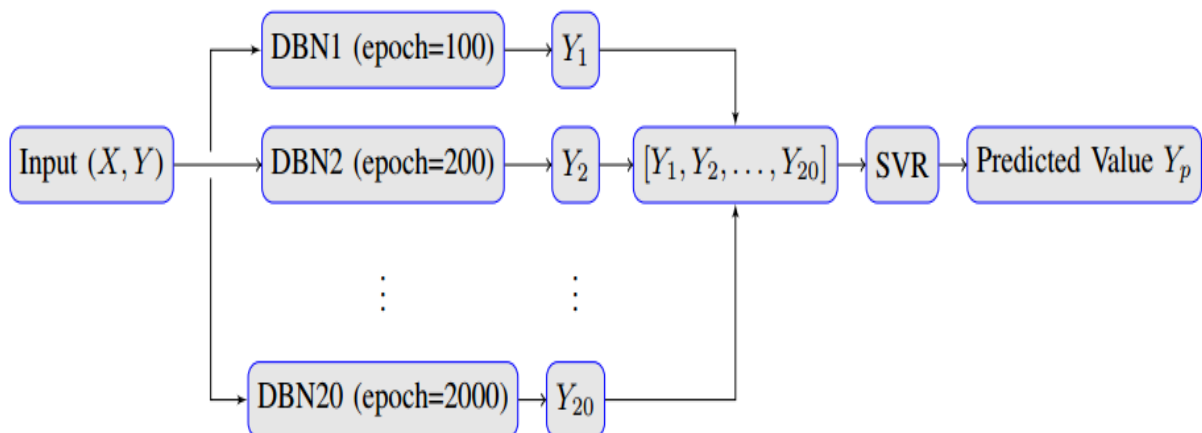
Bước 3: Đặt tất cả các output của mô hình huấn luyện vào một ma trận mới và sử dụng SVR (Support Vector Regression) để tìm ra mô hình output phù hợp nhất.

Bước 4: Đề xuất lấy mô hình có kết quả dự đoán chính xác nhất.



Hình 4. Mô hình ensemble cho bài toán dự báo chuỗi thời gian

Năm 2016, Sanam Narejo và Eros Pasero [7] đề xuất một phương pháp lai cho dự báo chuỗi thời gian sử dụng Deep Learning và Nonlinear Autoregressive neural network.



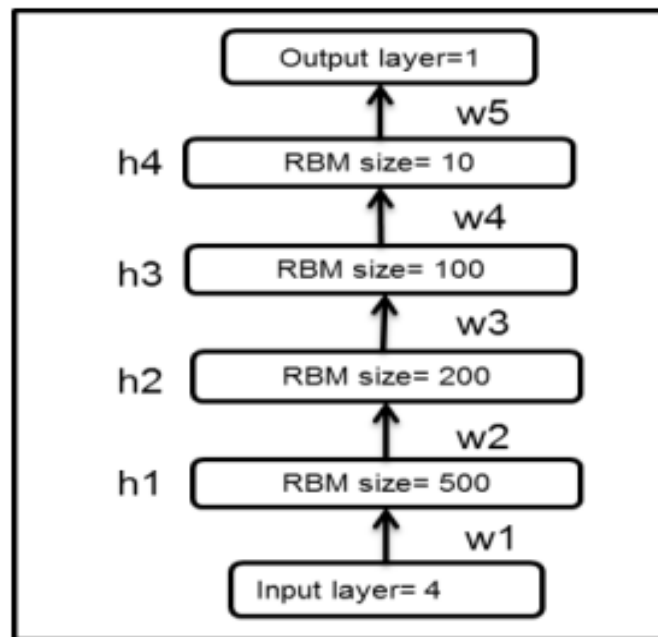
Hình 5. Phương pháp hybrid cho bài toán dự đoán chuỗi thời gian

Quá trình nghiên cứu được thực hiện bởi các bước sau:

Xử lý dữ liệu: Dữ liệu được thu thập từ trạm thời tiết Meteo của phòng thí nghiệm Neuronica tại Politecnico di Torino. Các mẫu được lấy lúc 11:45 ngày 4 tháng 10 năm

2010 đến lúc 09:15 ngày 7 tháng 8 năm 2015. Nhiệt độ được ghi lại 15 phút một lần. Dữ liệu đã được tiền xử lý bằng phương pháp nội suy tuyến tính và chuẩn hóa trong dãy (0,1). Dữ liệu đến tháng 3 năm 2014 được làm dữ liệu huấn luyện, phần còn lại dùng làm dữ liệu test.

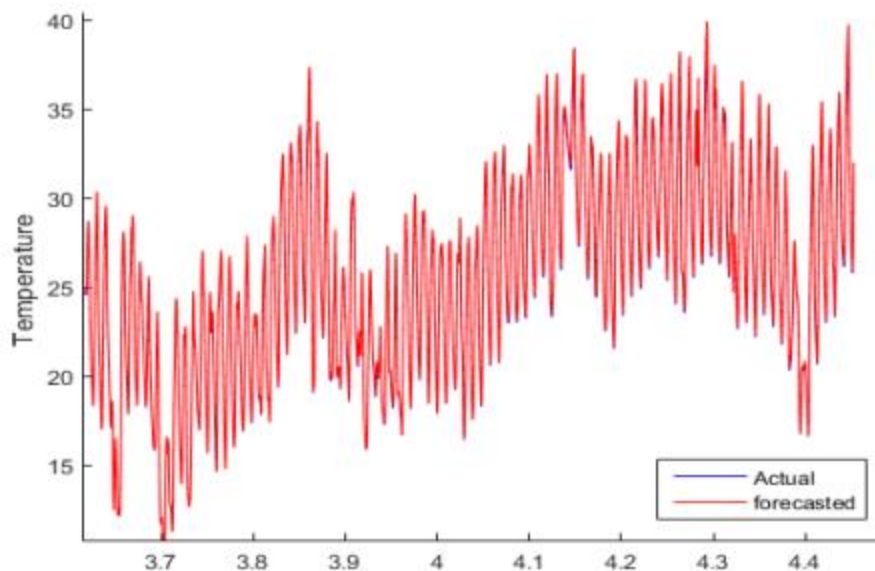
Thiết lập thực nghiệm: Xây dựng DNN với 6 lớp trong đó có 4 lớp ẩn. Số neural ở các lớp của DNN như sau 4-500-200-100-10-1.



Hình 6. Kiến trúc học đặc trưng của DNN

Dự báo: Sau khi học đặc trưng từ DNN, bước kết tiếp là xây dựng một mạng NARX ANN [7] để đưa ra kết quả dự báo nhiệt độ.

Kết quả thực nghiệm cho thấy độ lỗi RMSE của mô hình là 0.0010 trên tập huấn luyện, 0.0011 khi đánh giá, 0.0011 trên tập test. Những con số cho thấy kết quả dự đoán khá chính xác của mô hình được xây dựng.



Hình 7. Dự đoán nhiệt độ của mô hình đề xuất [7]

TÀI LIỆU THAM KHẢO

- [1] John Gamboa (2017), “*Deep Learning for Time-Series Analysis*”. CoRR, abs/1701.01887.
- [2] Anastasia Borovykh, Sander Bohte, Cornelis W. Oosterlee (2017), “*Conditional Time Series Forecasting with Convolutional Neural Networks*”. arXiv preprint arXiv:1703.04691.
- [3] Takashi Kuremoto, Shinsuke Kimura, Kunikazu Kobayashi, Masanao Obayashi (2014), “*Time series forecasting using a deep belief network with restricted Boltzmann machines*”. Neurocomputing, Volume 137, 5 August 2014, Pages 47-56.
- [4] Enzo Busseti, Ian Osband, Scott Wong (2012), “*Deep Learning for Time Series Modeling*”. CS 229 Final Project Report. Stanford University.
- [5] Xueheng Qiu, Le Zhang, Ye Ren, P. N. Suganthan (2014), “*Ensemble Deep Learning for Regression and Time Series Forecasting*”. In Proc. IEEE Symposium on Computational Intelligence and Ensemble Learning (CIEL’14), Orlando, US, Dec. 2014.
- [6] Zhicheng Cui, Wenlin Chen, Yixin Chen (2016). “*Multi-Scale Convolutional Neural Networks for Time Series Classification*”. arXiv preprint arXiv:1603.06995.

[7] Sanam Narejo, Eros Pasero (2016), “*A Hybrid Approach for Time series Forecasting using Deep Learning and Nonlinear Autoregressive Neural Network*”. INTELLI 2016 : The Fifth International Conference on Intelligent Systems and Applications (includes InManEnt 2016).

[8] Roni Mittelman (2015), “*Time-series modeling with undecimated fully convolutional neural networks*”. arXiv preprint arXiv:1508.00317.