Mô hình ARIMA sử dụng những đặc tính thống kê để xây dựng lên mô hình. Mặc dù ARIMA khá linh hoạt, nó có thể dùng để biểu diễn cho nhiều loại chuỗi thời gian khác nhau: AR(p), MA(q), ARMA(p,q), ARIMA(p,d,q), SARIMA(p,q,d)(P,Q,D)S, nhưng hạn chế chính của nó là giả định về mối liên hệ tuyến tính trong chuỗi thời gian. Bởi vì giả định về sự tương quan tuyến tính giữa các giá trị trong chuỗi dữ liệu thời gian nên mô hình ARIMA không thể nhận biết được các thành phần không tuyến tính. Do đó dùng mô hình tuyến tính để dự đoán cho các bài toán thực tế có thể dẫn đến kết quả không thỏa đáng.

Mô hình nơron nhân tạo ANNs được xây dựng dựa trên mô phỏng bộ óc con người. Ưu điểm chính của mô hình này là khả năng mô hình hóa các chuỗi dữ liệu mà các giá trị của nó có mối liên hệ phi tuyến. Với ANNs ta không cần xây dựng một mô hình cụ thể, mà mô hình được thích nghi dần qua quá trình huấn luyện. Phương pháp này thích hợp đối với chuỗi dữ liệu mà ta chưa xác định được các lý thuyết để tìm mối liên hệ trong dữ liệu.

Mô hình lai sử dụng cả hai phương pháp:

* Trong thực tế, thường rất khó để xác định mối liên hệ trong một chuỗi thời gian và tuyến tính hay phi tuyến và phương phương pháp nào hữu hiệu hơn trong việc dự đoán. Do đó rất khó để chọn một mô hình thích hợp, nên ta thường thử nhiều phương pháp và chọn cái ra kết quả tốt nhất. Tuy nhiên, mô hình được chọn không nhất nhiết là mô hình tốt nhất cho dự báo tương lai do sự tồn tại các yếu tố tiềm ẩn: sai số lấy mẫu, mô hình chưa chắc chắn, cấu trúc thay đổi.
* Các chuỗi dữ liệu thực hiếm khi thuần tuyến tính hay thuần phi tuyến, mà chúng thường bao gồm cả hai thành phần. Trong trường hợp đó, một mình mô hình ARIMA cũng như ANNs không đủ tin cậy để mô hình và dự đoán bởi vì mô hình ARIMA không thể giải quyết được mối liên hệ phi tuyến, một mình mô hình ANNs cũng không thể giải quyết được cả hai mối liên hệ phi tuyến và tuyến tính tốt được. Do đó kết hợp cả hai mô hình sẽ ra được kết quả tốt.
* Hầu hết các tài liệu dự báo đều cho rằng không một phương pháp đơn lẻ nào là tốt trong mọi trường hợp. Điều đó là vì các bài toán thực tế rất phức tạp và mang nhiều đặc tính khác nhau, không một phương pháp đơn lẽ nào có thể mô phỏng các bài toán tốt như nhau.

1. Hiện thực mô hình

Ta giả sử rằng mỗi chuỗi thời gian gồm cả thành phần tuyến tính và thành phần phi tuyến:

trong đó:

L: biểu diễn thành phần tuyến tính của chuỗi thời gian

N: biểu diễn thành phần phi tuyến của chuỗi thời gian

Hai thành phần đó sẽ được ước lượng từ dữ liệu. Đầu tiên, ta dùng mô hình ARIMA để dự đoán thành phần tuyến tính của chuỗi thời gian, lúc đó phần lỗi của mô hình ARIMA sẽ chỉ chứa thành phần phi tuyến tính. Với là lỗi tại thời điểm t của mô hình tuyến tính ARIMA, ta có: trong đó là giá trị dự đoán từ mô hình ARIMA tại thời điểm t.

Sau khi mô hình hóa dữ liệu bằng ARIMA thì lỗi có thể vẫn còn mang tính phi tuyến. Mô hình hóa lỗi này bằng ANNs, ta có thể phát hiện ra được mối liên hệ phi tuyến. Với n giá trị đầu vào, mô hình ANNs cho phần lỗi là:

Trong đó là hàm phi tuyến được xác định bởi mô hình nơron và lỗi ngẫu nhiên.

Sau khi mô hình đã được xây dựng, giá trị dự đoán từ hai mô hình thành phần là:

Trong đó là giá trị dự đoán từ mô hình ARIMA, là giá trị dự đoán từ mô hình ANNs.

Time series data

Artificial Neural Network

Seasonal ARIMA model

Forecast

ARIMA forecast

Tóm lại, mô hình lai gồm hai bước chính để xây dựng. Ở bước đầu tiên, mô hình ARIMA dùng để phân tích thành phần tuyến tính của chuỗi dữ liệu. Ở bước thứ hai, mô hình ANNs dùng để phân tích thành phần lỗi từ mô hình ARIMA. Bởi vì ARIMA không thể nhận biết được cấu trúc phi tuyến của dữ liệu nên lỗi từ mô hình sẽ chứa thành phần phi tuyến. Kết quả từ mô hình mạng nơron được xem là giá trị dự đoán cho lỗi từ mô hình ARIMA. Mỗi thành phần mô hình sẽ giải quyết được những đặc tính khác nhau của dữ liệu, bằng cách kết hợp cả hai phương pháp, ta sẽ nâng cao hiệu suất mô phỏng và dự đoán.