

ĐẠI HỌC QUỐC GIA TP HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH

-----oOo-----



BÁO CÁO THỰC TẬP TỐT NGHIỆP

DỰ BÁO DÒNG TIỀN
BẰNG HỌC MÁY

Giảng viên hướng dẫn:

GS.TS. Cao Hoàng Trụ

Giảng viên phản biện:

PGS.TS. Dương Tuấn Anh

Sinh viên thực hiện:

Trần Như Văn - 51104177

Ngô Khắc Vũ - 51104300

Thành phố Hồ Chí Minh, 06/2015

MỤC LỤC

1.	GIỚI THIỆU ĐỀ TÀI.....	1
1.1.	Bối cảnh	1
1.2.	Mục tiêu	1
2.	CÁC CÔNG TRÌNH LIÊN QUAN	3
2.1.	Mô hình mô tả dòng tiền	4
2.2.	Mô hình dự báo dòng tiền	7
3.	KIẾN THỨC NỀN	10
3.1.	Báo cáo tài chính.....	10
3.2.	Dòng tiền doanh nghiệp	11
3.3.	Mạng nơon nhân tạo	15
4.	GIẢI PHÁP ĐỀ XUẤT.....	21
4.1.	Mô tả bài toán	21
4.2.	Phương pháp đề xuất.....	22
5.	THỰC NGHIỆM.....	26
5.1.	Tập dữ liệu	26
5.2.	Phương pháp đánh giá.....	26
6.	KẾ HOẠCH THỰC HIỆN.....	28
	TÀI LIỆU THAM KHẢO	29

1. GIỚI THIỆU ĐỀ TÀI

1.1. Bối cảnh

Dòng tiền có vai trò rất quan trọng đối với mọi doanh nghiệp. Giữ cho dòng tiền khỏe mạnh là điều cốt lõi giúp doanh nghiệp có thể tồn tại.

Điều đó có nghĩa là doanh nghiệp phải luôn luôn giữ cho dòng tiền của mình lưu thông liên tục, đồng thời vẫn giữ được lượng tiền mặt ổn định để chi trả cho các khoản nợ tới hạn. Trên lý thuyết doanh nghiệp làm ăn có lãi nhưng nếu không đủ lượng tiền mặt để chi trả cho các khoản nợ tới hạn, doanh nghiệp đó có thể bị phá sản[3][4].

Trên thực tế, doanh nghiệp có thể kiếm được bao nhiêu tiền trong tương lai không phải là vấn đề nếu như doanh nghiệp không có đủ tiền để duy trì hoạt động. Nhân viên không thể đợi đến khi khách hàng trả tiền rồi mới nhận lương. Người cho thuê nhà không quan tâm doanh nghiệp có thể trả gấp 10 lần vào năm sau, hay các nhà cung cấp nguyên liệu chắc chắn sẽ không bằng lòng khi thời gian trả nợ bị kéo dài.

Để giữ cho dòng tiền khỏe mạnh, các doanh nghiệp thường đưa ra chiến lược kinh doanh phù hợp. Dự báo dòng tiền tương lai là một trong những nguồn thông tin quan trọng nhất cho các doanh nghiệp để đưa ra chiến lược phù hợp. Do đó, cần phải áp dụng các phương pháp khoa học để cải thiện thông tin dự báo và chiến lược kinh doanh để giữ dòng tiền luôn vững chắc.

1.2. Mục tiêu

Dữ liệu kinh doanh được cho là tài sản quan trọng nhất mà một doanh nghiệp sở hữu. Dựa vào dữ liệu kinh doanh, các doanh nghiệp có thể đưa ra dự báo dòng tiền tương lai và có chiến lược kinh doanh phù hợp. Với lượng dữ liệu lớn và tốc độ tăng trưởng nhanh, một vấn đề được đặt ra là làm sao tận dụng tối đa thông tin để đưa ra dự báo dòng tiền hiệu quả nhất. Trong những thập kỷ gần đây, cùng với sự tiến bộ của khoa học máy tính, những giải thuật học máy được áp dụng khá thành công vào các dự báo được xem như một hướng giải quyết cho vấn đề này.

Trên thế giới hiện nay đã có nhiều công trình nghiên cứu được thực hiện, trong đó chủ yếu tập trung vào đánh giá hiệu suất dự báo dòng tiền tương lai hàng năm của doanh nghiệp bằng những phương pháp khác nhau hoặc đánh giá khả năng dự báo dòng tiền tương lai. Có rất ít nghiên cứu tập trung cải thiện hiệu suất dự báo dòng tiền tương lai.

Ngoài ra, hầu hết các nghiên cứu đều sử dụng dòng tiền trong hoạt động kinh doanh (Cash Flow Operating) của doanh nghiệp để dự báo. Bởi dòng tiền CFO thể hiện tính thanh khoản, thể hiện doanh nghiệp có đủ lượng tiền mặt để chi trả cho các nhu cầu hiện tại hay không, và như đã đề cập ở phần trước, doanh nghiệp sẽ bị phá sản nếu không thể chi trả các khoản nợ tới hạn.

Bởi vì tính quan trọng của việc dự báo dòng tiền trong tương lai, dựa trên những công trình nghiên cứu liên quan, mục tiêu của đề tài là nghiên cứu mô hình học máy để dự báo dòng tiền CFO tương lai. Đồng thời, có thể so sánh đánh giá với các phương pháp kinh điển nhằm tìm ra được mô hình dự báo phù hợp.

2. CÁC CÔNG TRÌNH LIÊN QUAN

Hiện nay, trên thế giới có rất nhiều các công trình nghiên cứu về dự báo dòng tiền tương lai, theo khảo sát dự báo dòng tiền được chia thành 2 bài toán con như sau: xây dựng mô hình mô tả dòng tiền và xây dựng mô hình dự đoán dòng tiền.

- **Xây dựng mô hình mô tả dòng tiền:** Mô hình mô tả dòng tiền trả lời cho câu hỏi “Những yếu tố nào ảnh hưởng đến dòng tiền tương lai?” dựa trên dữ liệu trong báo cáo tài chính của doanh nghiệp: dòng tiền tương lai, dòng tiền quá khứ, thu nhập, lợi nhuận, thuế doanh nghiệp, vốn chủ sở hữu, khấu hao. Bài toán này sử dụng các phương pháp thống kê như: phương sai, độ lệch chuẩn, hệ số tương quan và các mô hình thống kê để đánh giá độ tương quan giữa các yếu tố. Ngoài ra, các yếu tố được nghiên cứu không chỉ là con số trong dữ liệu báo cáo tài chính, mà còn là các yếu tố kinh tế như: kích thước công ty, ngành kinh tế. Một số nghiên cứu trong bài toán này như: Khả năng dự báo dòng tiền hoạt động kinh doanh của thu nhập – Shubita[9] (2013), Dự thu và dự báo dòng tiền tương lai tại Hong Kong – Yan[14] (2005). Đây là hướng tiếp cận truyền thống được thực hiện trên nhiều nghiên cứu và đem lại kết quả rất khả quan.
- **Xây dựng mô hình dự báo dòng tiền:** Mô hình dự báo dòng tiền trả lời cho câu hỏi “Giá trị dòng tiền tương lai là bao nhiêu?”. Đây là một bài toán chủ đạo trong hướng nghiên cứu dự báo dòng tiền. Bài toán sử dụng các mô hình hồi quy tuyến tính truyền thống hay mô hình khai phá dữ liệu với giải thuật học máy để đưa ra dự báo cho dòng tiền tương lai. Kết quả của bài toán mô hình mô tả được sử dụng như là đầu vào của mô hình dự báo, với các đầu vào là các yếu tố có ảnh hưởng đến đầu ra - dòng tiền tương lai. Tuy không phải là bài toán mới trong nghiên cứu dòng tiền, nhưng với việc áp dụng thành công mô hình dự báo mới như: mạng thần kinh nhân tạo, mô hình dự báo nhóm phi tuyến tính thì bài toán vẫn đang thu hút sự quan tâm lớn từ cộng đồng khoa học, các tổ chức kinh tế và doanh nghiệp. Theo Polat[2] (2003), một khảo sát thực hiện trong 500 công ty ở Mỹ, có đến 52% công ty sử dụng phương pháp mô phỏng máy tính cho việc dự báo dòng tiền tương lai để đưa ra quyết định đầu tư. Điều này cho thấy, tuy là một bài toán lâu đời nhưng nó vẫn mang tầm quan trọng trong kinh tế hiện nay.

Trong các loại bài toán trên, mỗi loại có những mô hình đặc trưng khác nhau sẽ được trình bày sau đây:

2.1. Mô hình mô tả dòng tiền

Những hướng nghiên cứu của mô hình này đều tập trung vào việc phân tích dữ liệu báo cáo tài chính từ các doanh nghiệp để đưa ra mối liên quan giữa các yếu tố. Ngoài ra, có một số nghiên cứu khác là mô hình hóa tập dữ liệu bằng phương pháp phân cụm. Mục tiêu của các nghiên cứu này là tạo ra cụm dữ liệu đồng nhất trước khi áp dụng mô hình dự báo để cải thiện hiệu suất.

Liên quan đến hướng nghiên cứu về dự báo khả năng ảnh hưởng của các yếu tố đến dòng tiền tương lai, có thể kể đến: khả năng dự báo dòng tiền hoạt động kinh doanh của thu nhập (Shubita, 2013)[9], dự thu và dự báo dòng tiền tương lai (Yan, 2005)[14], khả năng của thu nhập và dòng tiền trong việc dự báo dòng tiền tương lai (Moeinaddin et al., 2012)[10]. Trong các nghiên cứu trên, nghiên cứu của Shubita[9] có tham khảo và sử dụng các yếu tố đầu vào của Yan[14] và Moeinaddin et al [10], ngoài ra, Shubita [9] cũng đề xuất sử dụng thêm một số yếu tố khác nên kết quả được đánh giá cao và có tính tổng hợp, vì vậy đề tài trình bày nghiên cứu này.

Về hướng mô hình hóa dữ liệu bao gồm các nghiên cứu: Kết hợp mạng nơron học có giám sát và không giám sát cho việc cải thiện dự báo dòng tiền (Smith et al., 2002)[11], nghiên cứu sử dụng mạng nơron cho việc dự báo dòng tiền hoạt động kinh doanh (Alborzi et al., 2013)[7]. Do 2 nghiên cứu trên khá tương đồng và đều sử dụng một mô hình mạng nơron học không giám sát nên đề tài sẽ trình bày mô hình đầu tiên.

Khả năng dự báo dòng tiền hoạt động kinh doanh của thu nhập (Shubita, 2013)[9]

Mục tiêu: Xem xét khả năng của thu nhập, khấu hao, vốn lưu động, dòng tiền hoạt động kinh doanh quá khứ và kích thước của công ty có thể dự báo dòng tiền hoạt động kinh doanh trong tương lai.

Tập dữ liệu: dữ liệu được lấy từ Sàn giao dịch chứng khoán Amman từ năm 2006 – 2011. 66 dữ liệu báo cáo tài chính của công ty được dùng với 362 mẫu (công ty/ năm).

Dữ liệu đầu vào: 4 biến đầu vào là thu nhập, thu nhập cộng khấu hao, vốn lưu động, dòng tiền hoạt động kinh doanh quá khứ. Ngoài ra, kích thước của công ty cũng được đưa vào như một đầu vào. Trong đó, kích thước của công ty là biến với 3 giá trị: nhỏ , trung bình và lớn.

Phương pháp: Nghiên cứu đề ra 4 mô hình giả thiết:

$$CFO_{it} = \alpha_0 + \alpha_1 EARN_{it-1} + e_{it}$$

$$CFO_{it} = \gamma_0 + \gamma_1 EDPR_{it-1} + e_{it}$$

$$CFO_{it} = \beta_0 + \beta_1 WCFO_{it-1} + e_{it}$$

$$CFO_{it} = \lambda_0 + \lambda_1 CFO_{it-1} + e_{it}$$

Trong đó:

$\alpha, \gamma, \beta, \lambda$ là các hệ số tương quan.

i: công ty

t: năm

CFO: dòng tiền hoạt động kinh doanh

EARN: thu nhập được định nghĩa là dòng thu nhập ròng sau thuế

EDPR: thu nhập hiện tại được định nghĩa là dòng thu nhập cộng với khấu hao

WCFO: vốn lưu động từ hoạt động kinh doanh.

Sau đó, một phương pháp hồi qui được sử dụng để tính toán các hệ số tương quan của 4 giả thiết và so sánh với độ tin cậy để đưa ra đánh giá cho các yếu tố.

Kết quả: Mô hình tính toán được các hệ số tương quan của các yếu tố đầu vào với dòng tiền hoạt động kinh doanh tương lai và đánh giá các yếu tố có ảnh hưởng đến khả năng dự báo dòng tiền. Qua đó, chỉ có yếu tố EARN, EDPR, CFO trong quá khứ, kích thước công ty là có khả năng dự báo CFO trong tương lai.

Đánh giá: Đây là nghiên cứu khá phổ biến trong bài toán này. Qua đó, ta thấy được không chỉ yếu tố cấu thành nên dòng tiền CFO mới ảnh hưởng đến nó, mà các yếu tố khác như: EARN, EDPR, kích thước công ty cũng có khả năng ảnh hưởng đến dự báo. Do đó, các yếu tố này được đề tài chọn để làm dữ liệu đầu vào cho bài toán.

Kết hợp mạng nơron học có giám sát và không giám sát cho việc cải thiện dự báo dòng tiền (Smith et al., 2002)[11]

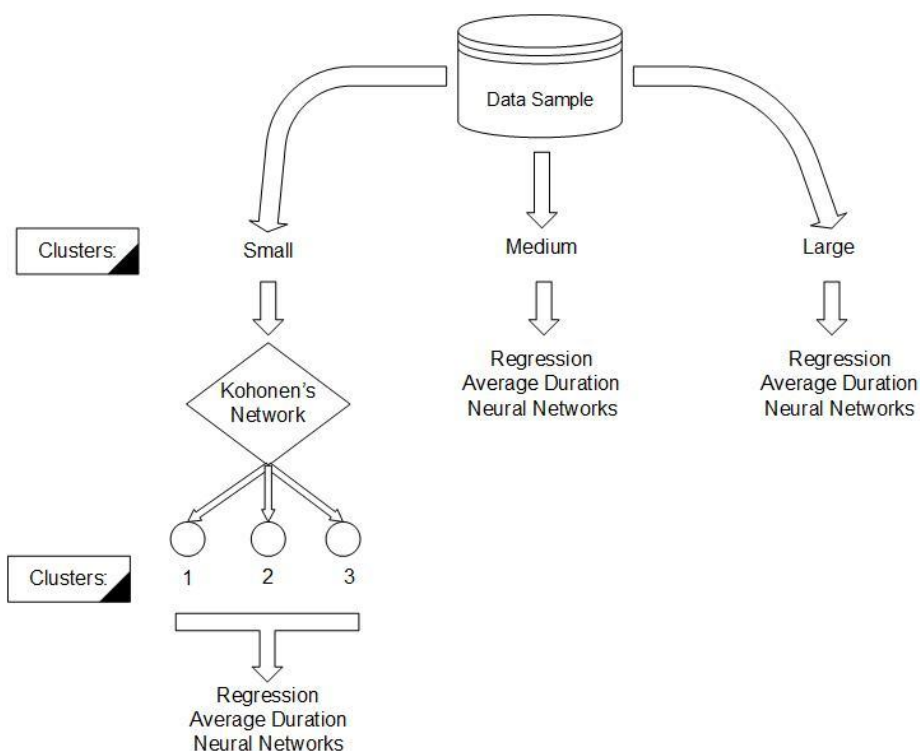
Mục tiêu: Dự báo dòng tiền hoạt động kinh doanh của một công ty theo ngày. Theo đó, vấn đề đặt ra là làm sao biết được ngày rút tiền của khách hàng để đảm bảo rằng quỹ của công ty cần có lượng tiền mặt đầy đủ cho việc chi trả khách hàng, tránh chi tiêu không hợp lý. Yếu tố đầu ra là lượng tiền khách hàng yêu cầu chi trả trong ngày sắp tới. Mô hình mô tả dòng tiền được

sử dụng ở đây là một mạng nơ-ron học không giám sát được sử dụng để phân cụm dữ liệu thành cụm đồng nhất trước khi áp dụng mô hình dự báo.

Tập dữ liệu: được lấy từ một công ty Úc trong khoảng thời gian 3 tháng từ 1 tháng 9 – 30 tháng 11 năm 1998. Với 10594 mẫu, trong đó mỗi mẫu tương ứng với một giao dịch rút tiền của khách hàng trong khoảng thời gian trên.

Dữ liệu đầu vào: số thứ tự của giao dịch rút tiền, số lượng tiền rút, ngày giao dịch rút tiền được tạo, ngày giao dịch rút tiền được thực hiện, postcode – mã giao dịch của khách hàng theo khu vực, tình trạng giao dịch.

Phương pháp: Dữ liệu ban đầu được phân loại theo lượng tiền giao dịch với 3 mức độ là: nhỏ, trung bình và lớn. Tiếp theo, một mô hình mạng nơ-ron học không giám sát Kohonen's SOM được tạo ra và áp dụng phân cụm vào mỗi nhóm. Mỗi cụm sau đó được áp dụng một nơ-ron học giám sát để dự báo. Chi tiết về các bước xây dựng mạng nơ-ron học không giám sát SOM và hoạt động sẽ được trình bày chi tiết ở phần kiến thức nền. Hình 2.1 mô tả các kỹ thuật được áp dụng trong mô hình hóa dữ liệu.



Hình 2.1 Kỹ thuật áp dụng trong mô hình hóa dữ liệu

Kết quả: dữ liệu nhóm theo lượng tiền giao dịch (nhỏ, trung bình và lớn). Sau đó, mỗi nhóm được phân vào 3 cụm để tạo ra các cụm dữ liệu đồng nhất. Một so sánh được thực hiện giữa

việc áp dụng phân cụm trước khi dự báo và không phân cụm. Kết quả đưa ra rất khả quan rằng việc phân cụm làm cải thiện đáng kể độ chính xác của dự báo.

Đánh giá: Nghiên cứu đưa ra mô hình mạng nơron học không giám sát SOM để phân cụm dữ liệu. Áp dụng thành tựu của nghiên cứu trên, đề tài đề xuất sử dụng một mạng nơron học không giám sát cho việc phân cụm dữ liệu, để tạo cụm tương đồng nhằm tăng hiệu suất dự báo.

2.2. Mô hình dự báo dòng tiền

Trên thế giới đưa ra rất nhiều mô hình về dự báo dòng tiền như: mô hình mạng nơron, mô hình support vector machine, mô hình mạng nơron quần thể phi tuyến tính. Đề tài trình bày chi tiết về 2 nghiên cứu: nghiên cứu sử dụng mạng nơron cho việc dự báo dòng tiền hoạt động kinh doanh (Alborzi et al., 2013)[7], nghiên cứu so sánh các phương pháp trí tuệ nhân tạo cho việc dự báo dòng tiền (Hongjiu et al., 2012)[12].

Sử dụng mạng nơron cho việc dự báo dòng tiền hoạt động kinh doanh (Alborzi et al., 2013)[7].

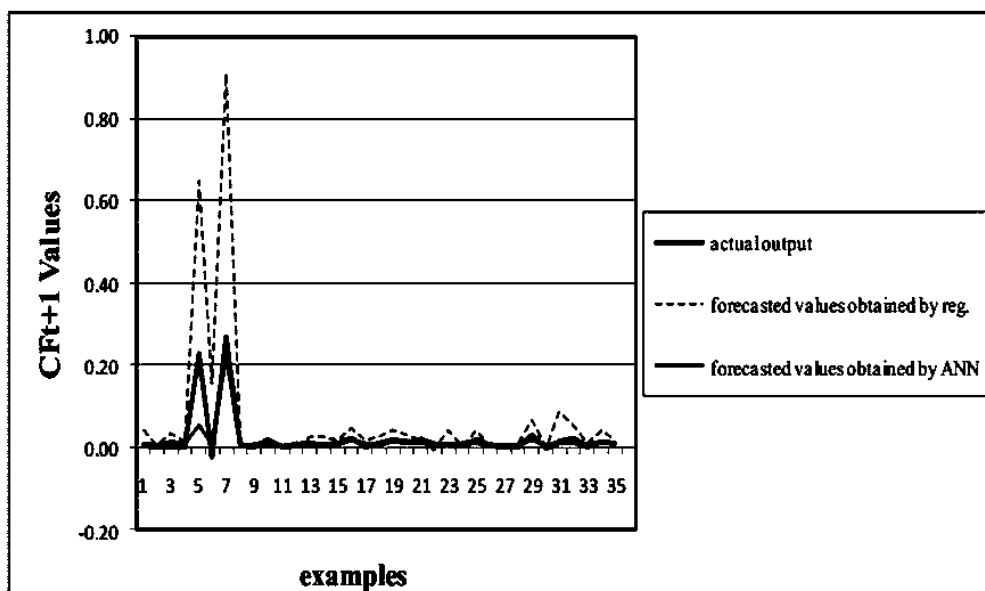
Mục tiêu: Đưa ra một nghiên cứu so sánh về hiệu suất dự báo của mô hình mạng nơron nhân tạo so với phương pháp hồi qui tuyến tính để dự báo dòng tiền hoạt động kinh doanh hằng năm của các công ty.

Tập dữ liệu: Được lấy từ sàn giao dịch chứng khoán Tehran (TSE). Trong đó, với 1008 mẫu (công ty/năm) bao gồm 126 công ty được lựa chọn trong 9 ngành kinh tế, trong khoảng thời gian 8 năm từ 3/2001 – 3/2009. Yêu cầu đặt ra là các báo cáo tài chính của công ty phải liên tục hằng năm. Trong báo cáo tài chính đảm bảo đủ 3 bảng báo cáo cần thiết: báo cáo thu nhập, bảng lưu chuyển tiền tệ và bảng cân đối tài sản.

Dữ liệu đầu vào: là các số liệu trong bảng báo cáo tài chính của công ty. Nghiên cứu này sử dụng 9 biến đầu vào trong bảng báo cáo lưu chuyển tiền tệ bao gồm: khoản phải thu, khoản phải nhận, thu nhập hoãn lại, chi phí trả trước, hàng tồn kho, chi phí khấu hao, dự trữ cá nhân, dòng tiền hoạt động kinh doanh.

Phương pháp: Một mạng nơron đa lớp truyền thẳng được xây dựng với lớp đầu vào 9 nút, 2 lớp ẩn với 9 nút tương ứng với 9 biến và lớp đầu ra 1 nút. Cấu trúc mạng nơron, lớp ẩn, số nút ẩn được chọn lựa sau quá trình thực nghiệm với các chỉ số học 0.1 và giá trị khởi tạo 0.7. Sau đó, giải thuật quay ngược được sử dụng cho mạng nơron để huấn luyện các mẫu. Ngoài ra,

trước khi áp dụng mạng nơron để dự báo, một mô hình mạng nơron Kohonen SOM được sử dụng để phân cụm các dữ liệu thành cụm đồng nhất nhằm cải thiện hiệu suất. Việc chuẩn hóa dữ liệu cũng được áp dụng nhằm giảm đi ảnh hưởng của yếu tố kích thước công ty. Hình 2.2 mô tả biểu đồ so sánh độ hiệu quả trong việc dự đoán dòng tiền bằng cách sử dụng phương pháp Neuron Network và phương pháp hồi quy thông thường.



Hình 2.2 So sánh phương pháp mạng nơron và hồi qui.

Kết quả: Một so sánh được thực hiện giữa 2 phương pháp sử dụng mạng nơron với phương pháp hồi qui. Kết luận rằng mạng nơron có hiệu suất dự báo cao hơn so với hồi qui. Nghiên cứu đánh giá việc áp dụng phân cụm giúp cải thiện hiệu suất dự báo.

Đánh giá: Nghiên cứu trên cho thấy mạng nơron đa lớp truyền thẳng như một mô hình dự báo phổ biến cho bài toán. Mô hình mạng nơron mang lại hiệu suất dự báo cao. Ngoài ra, nghiên cứu còn khẳng định lại việc phân cụm giúp cải thiện hiệu suất dự báo.

So sánh phương pháp trí tuệ nhân tạo cho việc dự báo dòng tiền (Hongjiu et al., 2012)[12].

Mục tiêu: So sánh các phương pháp Response surface (RSM), Back propagation neural network (BP), Radial basis function neural network (RBP), Support vector machine (SVM) trong việc dự báo dòng tiền tương lai.

Tập dữ liệu: lấy từ công ty ô tô trong thời gian từ 12/2002 – 3/2011, với 34 mẫu dữ liệu theo quý. Nghiên cứu này dựa trên tập dữ liệu khá ít (34 mẫu, ít hơn 300).

Dữ liệu đầu vào: CF dòng tiền trong quá khứ.

Phương pháp: 4 mô hình được xây dựng và huấn luyện với tập dữ liệu. Ngoài ra, phương pháp Sliding window còn được sử dụng để tìm ra kích thước tốt nhất cho việc dự báo của mỗi mô hình. Độ đo Mean of square error (MSE) được dùng cho quá trình huấn luyện và Average relative error (ARE) được dùng cho quá trình dự báo.

Kết quả: Trong 4 mô hình, SVM có hiệu suất dự báo cao nhất, theo sau là RBP, BP và RSM. Ngoài ra, BP là mô hình có quá trình huấn luyện tốt nhất, kế đến là SVM, RBP và RSM. Nhìn chung, SVM và RBF là 2 mô hình dự báo tốt hơn so với các mô hình còn lại trong cùng kích thước mẫu dữ liệu.

Đánh giá: Cùng với nghiên cứu của Shubita [9], nghiên cứu này đã đánh giá mô hình BP là tốt trong bài toán dự báo. Ngoài ra, BP còn có hiệu suất huấn luyện cao hơn so với mạng nơron RBP cùng loại. Học tập 2 nghiên cứu trên, đề tài đề xuất áp dụng mô hình mạng nơron đa lớp truyền thẳng và giải thuật lan truyền ngược cho mô hình dự báo dòng tiền.

Chi tiết 2 mạng nơron để áp dụng cho 2 mô hình dòng tiền được trình bày rõ ở phần sau.

3. KIẾN THỨC NỀN

3.1. Báo cáo tài chính

Trước khi tìm hiểu về dòng tiền trong doanh nghiệp, đề tài sẽ giới thiệu sơ lược về các bản báo cáo tài chính của doanh nghiệp, bao gồm bảng cân đối kế toán, báo cáo kết quả hoạt động kinh doanh và báo cáo lưu chuyển tiền tệ. Đây là những bản báo cáo tài chính quan trọng nhất của một doanh nghiệp.

Những báo cáo tài chính này cung cấp thông tin cho phép người đánh giá những thay đổi trong tài sản thuần của một doanh nghiệp, thể hiện tính thanh khoản, khả năng thanh toán và khả năng ảnh hưởng thời gian lưu chuyển tiền tệ của doanh nghiệp đó.

Bảng cân đối kế toán

Bảng cân đối kế toán là báo cáo tài chính tổng hợp, phản ánh tổng quát toàn bộ giá trị hiện có và nguồn hình thành tài sản của doanh nghiệp tại một thời điểm nhất định. Bảng cân đối kế toán sẽ được lập dựa trên những số liệu thực tế trong kinh doanh của doanh nghiệp.

Dưới đây là đồng nhất thức mô tả bảng cân đối kế toán:

$$\text{Tài sản} = \text{Nợ} + \text{Vốn chủ sở hữu}$$

Cấu trúc của bảng cân đối kế toán phụ thuộc vào bản chất của doanh nghiệp và cách thức điều hành của doanh nghiệp đó. Bảng cân đối kế toán gồm hai phần: nợ và vốn chủ sở hữu. Trong bảng cân đối kế toán, doanh nghiệp liệt kê các phần: tài sản lưu động, tài sản cố định, nợ ngắn hạn, nợ dài hạn và vốn chủ sở hữu.

Bảng cân đối kế toán đóng một vai trò quan trọng, bởi nó là tài liệu để nghiên cứu, đánh giá tổng quát tình hình và kết quả hoạt động sản xuất kinh doanh, trình độ sử dụng vốn của doanh nghiệp ở một thời điểm nhất định.

Thông qua bảng cân đối kế toán, doanh nghiệp có thể thấy rõ tình hình huy động nguồn vốn chủ sở hữu và nguồn vay nợ để mua sắm các loại tài sản, hoặc quan hệ giữa công nợ và khả năng thanh toán. Từ đó, doanh nghiệp có thể kiểm soát được tình trạng mất cân bằng nguồn vốn và có những biện pháp giải quyết kịp thời.

Báo cáo kết quả hoạt động kinh doanh

Báo cáo kết quả hoạt động kinh doanh thể hiện các kết quả của hoạt động kinh doanh trong một khoảng thời gian nhất định của doanh nghiệp.

Báo cáo kết quả hoạt động kinh doanh cho biết liệu hoạt động kinh doanh của doanh nghiệp có lãi hay không, nghĩa là thu nhập thuần dương hay âm. Ngoài ra, báo cáo kết quả hoạt động kinh doanh còn phản ánh lợi nhuận của doanh nghiệp trong một khoảng thời gian nhất định, ví dụ như trong một năm, đồng thời cho biết doanh nghiệp đó đã chi bao nhiêu tiền để sinh được tiền lời, từ đó các doanh nghiệp muốn đầu tư có thể xác định được tỷ lệ lợi nhuận trên doanh thu của doanh nghiệp đó.

Dưới đây là đồng nhất thức của báo cáo kết quả hoạt động kinh doanh:

$$\text{Doanh thu} - \text{Chi phí} = \text{Thu nhập thuần}$$

Như vậy, báo cáo kết quả hoạt động kinh doanh cũng là một báo cáo tài chính quan trọng giúp doanh nghiệp quản lý, cân đối dòng tiền ra vào, thể hiện tình hình tài chính của doanh nghiệp đó trong một khoảng thời gian. Đồng thời báo cáo kết quả hoạt động kinh doanh cũng là một trong những bản báo cáo được quan tâm nhất của doanh nghiệp khi các doanh nghiệp khác đầu tư vào.

Báo cáo lưu chuyển tiền tệ

Giống như bảng cân đối kế toán và báo cáo hoạt động kinh doanh, báo cáo lưu chuyển tiền tệ cũng là một trong những báo cáo tài chính quan trọng nhất của doanh nghiệp. Báo cáo lưu chuyển tiền tệ mô tả cách thức vận hành dòng tiền của doanh nghiệp, doanh nghiệp thu và chi bao nhiêu, vào mục đích gì đều nằm trong bản báo cáo này.

Các thành phần của báo cáo lưu chuyển tiền tệ bao gồm: dòng tiền từ hoạt động đầu tư, dòng tiền từ hoạt động tài chính và dòng tiền từ hoạt động kinh doanh. Đây cũng là bản báo cáo mà đề tài muốn hướng tới.

3.2. Dòng tiền doanh nghiệp

Dòng tiền từ hoạt động đầu tư

IASB định nghĩa dòng tiền từ hoạt động đầu tư của doanh nghiệp như sau: “*Dòng tiền từ hoạt động đầu tư (Cash flow from investing) bao gồm dòng tiền tạo ra từ việc thu thập và xử lý các tài sản dài hạn và các khoản đầu tư khác không bao gồm trong các khoản tương đương tiền.*”

Cụ thể, dòng tiền từ hoạt động đầu tư của doanh nghiệp (IASB, 2010)[5], gồm các thành phần sau:

- Tiền mua / bán tài sản, thiết bị, tài sản vô hình và các tài sản dài hạn khác (-)(+)
- Tiền mua / bán cổ phần hoặc công cụ nợ của doanh nghiệp và lợi ích khác trong liên doanh (-)(+)
- Tiền thanh toán / thu được từ các hợp đồng dài hạn, hợp đồng trong tương lai (-)(+)

Ghi chú: (+) cho dòng tiền vào, (-) dòng tiền ra.

Việc công bố riêng biệt lưu chuyển tiền tệ phát sinh từ hoạt động đầu tư là rất quan trọng bởi vì các lưu chuyển tiền tệ đại diện cho mức độ mà các chi phí này được thực hiện cho các nguồn lực nhằm tạo ra thu nhập trong tương lai và lưu chuyển tiền tệ.

Dòng tiền từ hoạt động tài chính

Đối với dòng tiền từ hoạt động tài chính của doanh nghiệp (Cash flow from financing), IASB định nghĩa rằng: “*Dòng tiền từ hoạt động tài chính là dòng tiền phát sinh từ các hoạt động dẫn đến thay đổi trong kích thước và thành phần vốn tự có và các khoản vay của doanh nghiệp.*”

Chi tiết được mô tả trong báo cáo IAS7 như sau:

- Các khoản tiền mua / thu được từ tiền phát hành cổ phiếu (-)(+)
- Tiền thu tiền từ phát hành trái phiếu, ghi chú, trái phiếu (+)
- Tiền mặt trả nợ / thu được từ các khoản vay (-)(+)

- Các khoản thanh toán bằng tiền bên thuê để giảm trách nhiệm nổi bật liên quan đến một hợp đồng thuê tài chính (-)

Ghi chú: (+) cho dòng tiền vào, (-) dòng tiền ra.

Cũng giống như việc công bố dòng tiền từ hoạt động đầu tư, việc công bố riêng biệt của các luồng tiền phát sinh từ hoạt động tài chính cũng rất quan trọng, bởi vì nó rất hữu ích trong việc dự đoán các khiếu nại về tiền trong tương lai bởi các nhà cung cấp vốn cho doanh nghiệp.

Dòng tiền từ hoạt động kinh doanh

“Dòng tiền từ hoạt động kinh doanh – (Cash Flow from Operating) bao gồm dòng tiền tạo ra từ hoạt động doanh thu chủ yếu của doanh nghiệp”.

“Trong đó, dòng tiền từ kinh doanh (CFO) là một chỉ số quan trọng để đánh giá khả năng trả nợ và giải quyết các nhu cầu kinh tế của mình”, (Jonathan Moreland, 1995)[6]. Cũng theo IASB, tiền phát sinh do hoạt động kinh doanh là yếu tố quan trọng cho các doanh nghiệp để trả nợ duy trì hoạt động của tổ chức, trả cổ tức và thực hiện đầu tư mà không cần đến nguồn tài chính bên ngoài.

Điều đó có nghĩa là chỉ số CFO của công ty cho biết công ty đó có đủ tiền để trả nợ hay không. Điều cốt lõi ở đây là công ty sẽ bị phá sản nếu không trả được các khoản nợ ngắn hạn mặc dù trên báo cáo tài chính công ty đó làm ăn có lãi. Do đó, đề tài chọn dòng tiền từ hoạt động kinh doanh (CFO) để dự báo.

Theo IASB, CFO bao gồm dòng tiền từ các hoạt động sau:

- Lợi nhuận trước lãi vay và thuế thu nhập (+)
- Cộng lại khấu hao (+)
- Cộng lại tăng hoặc giảm tài sản (-)(+)
- Các khoản phải thu trong hoạt động kinh doanh (+)
- Tăng giảm hàng tồn kho (-)(+)

- Các khoản phải trả trong hoạt động kinh doanh (-)
- Chi phí lãi vay (-)
- Lãi cộng dồn nhưng chưa thanh toán (-)
- Chênh lệch tỷ giá (-)(+)
- Lãi vay đã trả (-)
- Thuế thu nhập (-)

Ghi chú: (+) cho dòng tiền vào, (-) dòng tiền ra.

Hình 3.1 là mẫu ví dụ của IAS về dòng tiền trong hoạt động kinh doanh của doanh nghiệp.

Cash flows from operating activities	
Profit before taxation	3,350
Adjustments for:	
Depreciation	450
Foreign exchange loss	40
Investment income	(500)
Interest expense	400
	<hr/>
	3,740
Increase in trade and other receivables	(500)
Decrease in inventories	1,050
Decrease in trade payables	(1,740)
	<hr/>
Cash generated from operations	2,550
Interest paid	(270)
Income taxes paid	(900)
	<hr/>
Net cash from operating activities	1,380

Hình 3.1 Một ví dụ về dòng tiền trong hoạt động kinh doanh

Số tiền lưu chuyển phát sinh từ hoạt động kinh doanh là một chỉ số quan trọng, thể hiện của mức độ mà các hoạt động của các doanh nghiệp đã tạo ra tiền mặt dùng để trả vốn vay, duy trì khả năng hoạt động của doanh nghiệp, trả cổ tức và thực hiện đầu tư mới mà không trông cậy vào nguồn tài chính bên ngoài.

3.3. Mạng nơron nhân tạo

Mô hình mạng nơron nhân tạo là một công cụ hiệu quả để giải quyết các bài toán dự báo sau khi biết mối quan hệ giữa các biến đầu vào và đầu ra (Chan, 2000)[15]. Trong lĩnh vực dự báo tài chính, mạng nơron nhân tạo được áp dụng khá thành công. Theo Smith[1][13], mô hình mạng nơron được dùng để:

- Dự báo sự kiện trong tương lai dựa trên các mẫu quan sát từ quá trình huấn luyện.
- Phân loại dữ liệu vào các nhóm được xác định trước dựa trên các đặc điểm của các mẫu quan sát trong quá trình huấn luyện.
- Phân cụm dữ liệu huấn luyện vào trong các cụm tự nhiên dựa trên sự tương đồng của đặc điểm dữ liệu.

Trong khi có rất nhiều mô hình mạng nơron được phát triển trong hơn 2 thập kỷ qua, nhưng phổ biến nhất là 2 mô hình mạng nơron học có giám sát và học không giám sát. Mô hình mạng nơron học có giám sát được dùng để phân loại và dự báo mà tiêu biểu là mạng nơron nhân tạo đa lớp truyền thẳng (MFNN), còn mô hình mạng nơron học không giám sát dùng để phân cụm với mô hình bản đồ tự tổ chức (SOM).

Mạng nơron học có giám sát

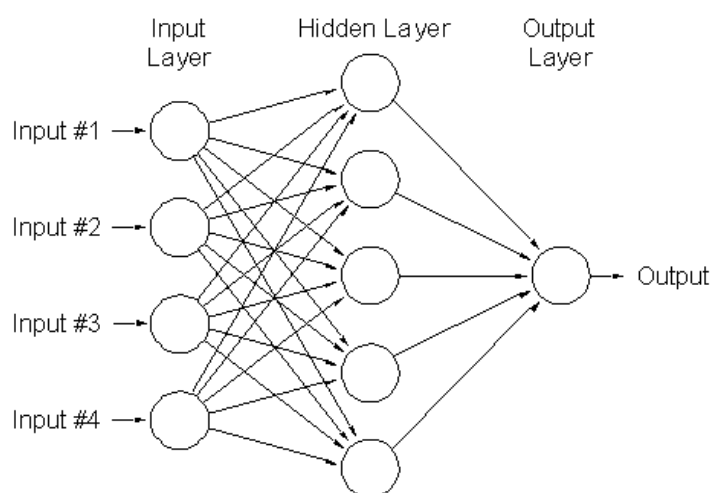
Theo nghiên cứu của Wong, et al [16], mạng nơron đa lớp truyền thẳng và giải thuật lan truyền ngược được sử dụng chiếm hơn 50% trong các nghiên cứu dự báo tài chính sử dụng mạng nơron. Đây là loại mạng nơron phổ biến được ứng dụng rộng rãi trong nhiều lĩnh vực liên quan đến doanh nghiệp và được đánh giá cao về hiệu suất huấn luyện. Do đó, đề tài chọn và trình bày mạng nơron đa lớp truyền thẳng và giải thuật lan truyền ngược tiêu biểu cho mô hình mạng này.

Mạng nơron đa lớp truyền thẳng

Kiến trúc:

Mạng nơron là kiểu mạng mô phỏng mạng lưới thần kinh của con người, bao gồm 3 lớp: lớp đầu vào, lớp ẩn và lớp đầu ra. Lớp đầu vào nhận dữ liệu đầu vào của bài toán, lớp ẩn tính toán dựa trên các trọng số w và giá trị bias, lớp đầu ra xuất giá trị dự báo. Trong đó, mỗi nơron có nhiệm vụ tính toán đầu ra dựa trên dữ liệu đầu vào mà nó nhận được. Cụ thể, đầu vào của một nơron là tổng đầu vào mà nó nhận được và đầu ra được tính toán dựa trên hàm sigmoidal với mỗi hàm phù hợp cho loại dữ liệu đầu vào cụ thể.

Ví dụ cụ thể với bài toán dự báo dòng tiền hoạt động kinh doanh: đầu vào là 4 biến dữ liệu bao gồm thu nhập, dòng tiền hoạt động kinh doanh hiện tại, khấu hao và thuế. Hình 3.2 mô tả mô hình mạng nơron cho bài toán trên.



Hình 3.2 Mô hình mạng Neuron Network với 4 đầu vào và 1 đầu ra.

Theo đó, 4 input tương ứng với 4 biến dữ liệu đầu vào, 1 output tương ứng với giá trị dòng tiền tương lai.

Theo Lai et al, mạng nơron là một phương pháp học tập không ổn định, tức là những thay đổi nhỏ trong quá trình huấn luyện hoặc lựa chọn tham số có thể ảnh hưởng lớn đến kết quả dự báo. Nhiều nghiên cứu đã chỉ ra rằng, đối với một số vấn đề đơn giản, cấu trúc khác nhau của mạng (ví dụ: số nơron ẩn, lớp ẩn hay điều kiện ban đầu khác nhau) dẫn đến mô hình khác nhau của mạng đưa đến những kết quả sai lệch. Do đó, trong một bài toán cụ thể cần phải thử nghiệm nhiều mô hình mạng để tìm ra mạng nơron phù hợp nhất.

Giải thuật lan truyền ngược

Giải thuật lan truyền ngược được dùng trong mô hình mạng để tính toán cập nhật lại trọng số và tối thiểu hóa hàm lỗi E.

Trong đó, một số hàm lỗi E được sử dụng phổ biến cho mạng nơron là:

$$\text{Mean Squared Error (MSE)} = \frac{1}{n} \sum_{i=1}^n (t_i - z_i)^2$$

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - z_i)^2}$$

$$\text{Normalized Mean Squared Error (NMSE)} = \frac{\sum_{i=1}^n (t_i - z_i)^2}{\sum_{i=1}^n (t_i - \bar{z}_i)^2}$$

$$R^2 \text{ (Coefficient of determination)} = 1 - \frac{\sum_{i=1}^n (t_i - z_i)^2}{\sum_{i=1}^n (t_i - \bar{z}_i)^2}$$

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n} \sum_{i=1}^n |t_i - z_i|^2$$

$$\text{Mean Absolute Percentage Error (MAPE)} = \frac{100}{n} \times \sum_{i=1}^n |t_i - z_i|^2$$

Trong đó: n số lượng mẫu của tập dữ liệu

 t dòng tiền tương lai từ dữ liệu báo cáo

 z dòng tiền tương lai từ output của mạng

Cụ thể giải thuật gồm 3 bước chính:

- Tính toán output (Feed-forward computation)

Bao gồm việc tính toán đầu vào và đầu ra của các lớp ẩn và lớp ra.

$$Net_j = \sum_{i=1}^{n+1} w_{ji} x_i$$

$$y = f(Net_j)$$

Theo đó: Net_j là đầu vào của nơon j.

Y_j là đầu ra của nơon j.

F() là hàm tổng.

- Tính toán lan truyền ngược (Backpropagation computation)

Lỗi của mỗi nơon được tính toán như sau:

$$Error = E + \sum_{i=1}^n (d_n - z_n)^2$$

- Cập nhật trọng số (Weight updates)

Dựa vào giải thuật lan truyền ngược ta tính toán trọng số cho output như sau:

$$v_j = v_j + c\gamma(d - z)z(1 - z)y_j$$

Và cho các nơon ở lớp ẩn.

$$w_{ij} = w_{ij} + c\gamma^2 y_j (d - y_j) x_i (Error)$$

Với γ là hằng số học (learning constant) có giá trị rất nhỏ khoảng 0.01 hay 0.001 để đảm bảo lượng số cập nhật không quá lớn nhằm đạt điểm tối ưu.

Giải thuật này lặp lại 3 bước cho đến khi giá trị lỗi E thay đổi và nhỏ hơn ngưỡng quy định hoặc dừng khi đạt số lần lặp nhất định.

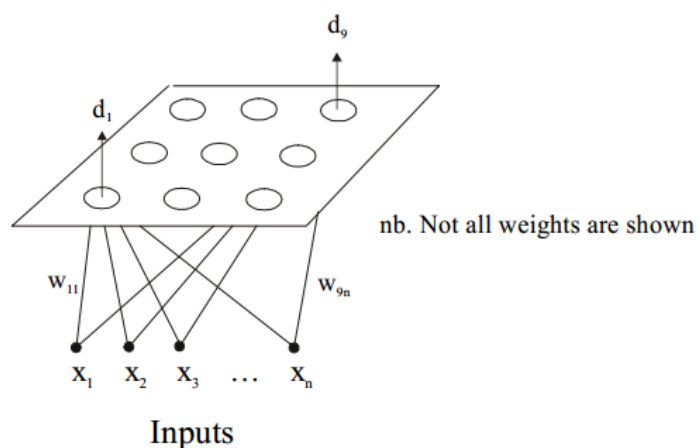
Mạng nơon học không giám sát

Mạng nơron học không giám được dùng cho việc phân cụm dữ liệu với nhiều mục đích trong phân tích và mô hình hóa dữ liệu. Nó cho phép chúng ta kiểm tra tập dữ liệu và phát hiện các mẫu ngoại lai và nhiễu. Ngoài ra, nó còn cho phép nhóm các dữ liệu có cấu trúc để tạo thành các cụm được mô hình hóa riêng biệt. Bản đồ tự tổ chức Kohonen (SOM) được xem như là mạng nơron học không giám sát nổi tiếng trong việc phân cụm. Đề tài sẽ trình bày mô hình bản đồ tự tổ chức SOM sau đây:

Bản đồ tự tổ chức

Kiến trúc:

Mạng nơron Kohonen chỉ bao gồm một lớp dữ liệu đầu vào và một lớp dữ liệu đầu ra, nó không chứa lớp ẩn. Số lượng nơron ở lớp đầu vào tương ứng với số biến trong mẫu dữ liệu. Số lượng nơron ở lớp đầu ra tương ứng với số cụm và phụ thuộc vào mục đích của việc phân cụm. Hình 3.3 mô tả mạng Neuron SOM với n đầu vào và 9 nơron đầu ra.



Hình 3.3 Mô hình mạng Neuron SOM với n đầu vào và 9 đầu ra.

Giải thuật học SOM:

Ý tưởng chính của thuật toán SOM là: khởi tạo mẫu ngẫu nhiên trong tập dữ liệu, sau đó tính toán trọng số của các nơron đầu ra với mẫu dữ liệu trên. So sánh trọng số của nơron và chọn ra nơron chiến thắng. Cập nhật lại trọng số của nơron ở vùng lân cận. Giải thuật kết thúc khi các sai số ở mức thấp hơn so với mục tiêu hoặc quá số lần lặp xác định.

Tương tự như mô hình MFNN trong phần trước, thuật toán SOM bao gồm 2 bước cơ bản sau:

- Tính toán nơon đầu ra (Calculating neural outputs)

Tính toán trọng số cho mỗi nơon:

$$d_j = \|x - w_j\| = \sqrt{\sum_{i=1}^n (x_i - w_{ij})^2}$$

- Cập nhật trọng số (Weight updating).

Cập nhật trọng số của nơon chiến thắng và nơon lân cận theo luật học:

$$w_{ji}(t+1) = w_{ij}(t) + c[x_i - w_{ji}(t)]$$

Trong đó: c giá trị học tập mà mỗi nơon nhận được khi cập nhật trọng số, theo đó với các nơon ở xa nơon chiến thắng hơn, giá trị này sẽ giảm theo hàm mũ xác định.

$$c = \alpha(t) \exp\left(\frac{-\|r_i - r_m\|}{\sigma^2(t)}\right)$$

Với $r_i - r_m$ là khoảng cách từ nơon i đến nơon chiến thắng.

4. GIẢI PHÁP ĐỀ XUẤT

4.1. Mô tả bài toán

Từ kiến thức thu được qua các công trình nghiên cứu đã tìm hiểu, nhóm đề xuất bài toán: “Xây dựng mô hình mô tả và dự báo dòng tiền hoạt động kinh doanh tương lai”. Các nghiên cứu về dự báo dòng tiền dựa vào các bản báo cáo tài chính của công ty để thực hiện. Ngoài ra, do Việt Nam là nước đang phát triển nên số lượng các công ty và thời gian hoạt động là không đủ để đề tài lấy mẫu. Do đó, đề tài chọn các công ty trên sàn giao dịch chứng khoán London để thực hiện thu thập dữ liệu. Tập dữ liệu của đề tài được trình bày chi tiết ở phần sau. Trước khi đi vào phương pháp thực hiện, đề tài trình bày rõ về dữ liệu đầu vào và kết quả đầu ra.

Dữ liệu đầu vào

Thời gian: Dữ liệu được thu thập hằng năm, trong khoảng 10 năm liên tục (2004 – 2014). Việc thu thập dữ liệu hằng năm là do niên độ kế toán của công ty là 1 năm. Trong đó, niên độ kế toán là khoảng thời gian mà công ty phải thực hiện báo cáo tài chính.

Vị trí: các công ty quốc tế trên sàn giao dịch chứng khoán London[17].

Loại dữ liệu:

Dòng tiền

- Doanh thu hằng năm (đôla).
- Lợi nhuận hằng năm (đôla).
- Thuế doanh nghiệp (đôla).
- Khoản phải thu (đôla).
- Khoản phải trả (đôla).
- Chi phí sản xuất (đôla).
- Hàng tồn kho (đôla).
- Khấu hao (đôla).

Ngành kinh tế

- Ngành kinh tế của công ty. Ví dụ: Thiết bị điện tử, Vận chuyển hàng hóa, Hàng không vũ trụ, ...

Tài sản

- Tổng tài sản của công ty (đôla).
- Nợ phải trả bao gồm nợ ngân hàng và các công ty khác (đôla).
- Vốn chủ sở hữu (đôla).

Kết quả đầu ra

Dòng tiền hoạt động kinh doanh trong một năm tới của từng công ty.

4.2. Phương pháp đề xuất

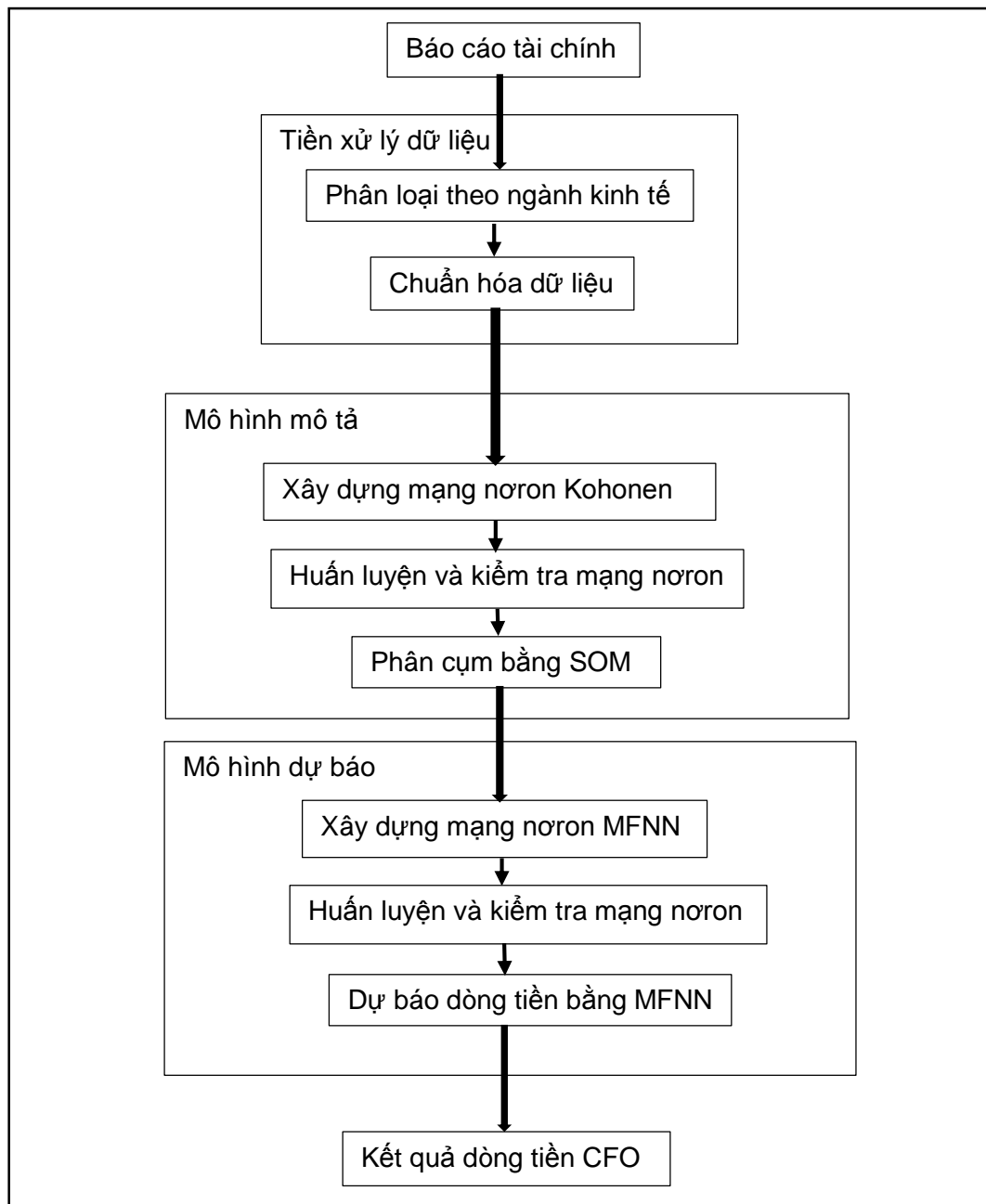
Đề tài đề xuất các bước thực hiện được biểu diễn như trong hình 4.1. Có 3 bước thực hiện theo tuần tự bao gồm: tiền xử lý dữ liệu, xây dựng mô hình mô tả và áp dụng phân cụm dữ liệu, xây dựng mô hình dự báo và áp dụng dự báo dòng tiền. Hình 4.1 mô tả tổng quan các bước thực hiện.

Tiền xử lý

Dữ liệu ở đây là bảng báo cáo tài chính của doanh nghiệp, trong đó bao gồm 2 bảng báo cáo con: báo cáo lưu chuyển tiền tệ, bảng cân đối tài sản và báo cáo thu nhập. Trong đó bảng báo cáo lưu chuyển tiền tệ đặc biệt quan trọng vì nó chứa thông tin của dòng tiền hoạt động kinh doanh bao gồm các dòng vào và dòng ra. Ngoài ra, trong 2 bảng báo cáo thu nhập và bảng cân đối tài sản chứa thông tin của yếu tố ảnh hưởng khả năng dự đoán dòng tiền và thông tin để phân cụm dữ liệu như: thu nhập, khấu hao, thuế doanh nghiệp, vốn chủ sở hữu, tài sản.

Tuy các báo cáo tài chính thống nhất về chuẩn mực, nhưng đối với mỗi công ty còn dùng tên gọi khác nhau cho cùng một thuộc tính. Ví dụ: Với thuộc tính Account receivable (Tài khoản thu), trong báo cáo tài chính của công ty SamSung mang tên Interest Received còn trong công ty Bowleven mang tên (Decrease)/increase in trade and other recievables. Do đó, cần xử lý lại tên thuộc tính trong báo cáo tài chính trước khi lấy dữ liệu.

Theo đề xuất, dữ liệu sau khi được tiền xử lý sẽ được phân loại thành các nhóm theo ngành nghề kinh tế để tạo thành các tập dữ liệu con đồng nhất. Việc này giúp cải thiện hiệu suất dự báo.



Hình 4.1 Tóm tắt các bước của giải pháp đề xuất

Ngoài ra, kích thước của công ty cũng ảnh hưởng đến việc dự báo. Vì vậy, thay vì phân loại công ty, chuẩn hóa dữ liệu được áp dụng để giảm thiểu mức độ ảnh hưởng của kích thước và đơn giản hóa.

$$CF'_t = \frac{(CF_t - CF_{t\min})}{(CF_{t\max} - CF_{t\min})}$$

Trong đó: $CF_{t\max}$ và $CF_{t\min}$ là 2 giá trị tối đa và tối thiểu của toàn bộ tập dữ liệu.

Xây dựng mô hình mô tả và áp dụng phân cụm

Theo các nghiên cứu trước đây của Lokmic, việc phân công ty vào những cụm đồng nhất tạo ra các cụm với các mẫu có đặc tính giống nhau, qua đó, mỗi cụm được áp dụng một MFNN riêng biệt để tăng độ chính xác của dự báo. Việc phân cụm được thực hiện bằng một mô hình mạng nơron học không giám sát. Theo đánh giá, mô hình Kohonen SOFM vượt trội hơn trong việc phân cụm. Do đó, đề tài sẽ áp dụng SOFM để phân cụm công ty trước khi dự báo.

Xây dựng mô hình dự báo và áp dụng dự báo dòng tiền

Theo Lai, mạng nơron là phương pháp học tập không ổn định. Do đó, bằng cách tạo ra nhiều mạng nơron với cấu trúc khác nhau, sau đó, bằng phương pháp thực nghiệm chọn ra các mạng nơron tốt nhất. Việc áp dụng mạng nơron để dự báo bao gồm 2 bước: lựa chọn thuộc tính và huấn luyện, kiểm tra mạng nơron.

Lựa chọn thuộc tính

Independent variables (input): Bao gồm các thuộc tính trong bảng báo cáo tài chính.

- Accounts receivable
- Accounts payable
- Deferred Income
- Prepaid expenses
- Inventories
- Depreciation and amortization expense,
- Personnel privilege reserve (PR).

Theo nghiên cứu của Shubita [9], Moeinaddin et al [10], đã chỉ ra rằng thu nhập và dòng tiền từ quá khứ cũng ảnh hưởng đến khả năng dự báo dòng tiền tương lai. Do đó, đề tài đề xuất sử dụng thêm 2 biến: **EARNs** và **EDPR**. Trong đó EARNs được định nghĩa là dòng tiền ròng sau thuế và EDPR được định nghĩa là thu nhập hiện tại cộng với khấu hao.

Dependent Variables (Output): Dòng tiền từ hoạt động kinh doanh trong một năm tới – $CFO(t+1)$.

Huấn luyện và kiểm tra mạng nơron

Theo khảo sát các nghiên cứu, mạng nơron đa lớp truyền thẳng và giải thuật lan truyền ngược (Back Propagation Neural Network) là phương pháp mạng nơron có hiệu suất dự báo và huấn luyện tốt so với các phương pháp trí tuệ nhân tạo khác. Do đó, đề tài xây dựng các mạng nơron

dựa vào mô hình nơron trên. Các mạng nơron sẽ được huấn luyện và kiểm tra trên tập dữ liệu. Sau đó, bằng phương pháp đánh giá với 3 độ đo được đề xuất theo nghiên cứu của Alborzi et al [7], chọn ra được mạng nơron có kết quả tốt nhất. Mạng nơron được chọn được huấn luyện và kiểm tra lại trên tập dữ liệu đã phân cụm. Kết quả cuối cùng được so sánh với giá trị thực tế và một số phương pháp đánh giá được dùng để đánh giá mô hình dự báo. Các phương pháp đánh giá được trình bày trong mục 5.2.

5. THỰC NGHIỆM

5.1. Tập dữ liệu

Tập dữ liệu của đề tài lấy từ bảng báo cáo tài chính của các công ty niêm yết trên sàn giao dịch chứng khoán London. Theo khảo sát, đề tài sẽ sử dụng tập dữ liệu với mẫu khoảng 100 – 150 công ty, thuộc khoảng 5 – 8 nhóm ngành kinh tế trong vòng 10 năm trở lại đây. Số lượng báo cáo tài chính dự kiến khoảng 1000 – 1500 công ty/năm, đây là số lượng khá đủ cho việc huấn luyện và kiểm tra. Ngoài ra, các công ty khi lên sàn giao dịch chứng khoán London thì dữ liệu tài chính phải minh bạch và theo chuẩn mực của quốc tế nên đảm bảo được tính chính xác và hệ thống của dữ liệu.

Dữ liệu thu thập sẽ chia thành 5 phần trong đó: 4 phần dành cho việc huấn luyện mạng nơron và 1 phần cho kiểm tra.

5.2. Phương pháp đánh giá

Dựa vào những công trình nghiên cứu của Alborzi et al [7], độ đo sử dụng phổ biến là MSE (mean square error), RMSE (root mean square error) và MAE (mean absolute error) được tính bằng công thức.

$$MSE = \frac{1}{n} \sum_{i=1}^n (System.output - ground.truth)^2$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (System.output - ground.truth)^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |System.output - ground.truth|$$

Trong đó:

- System.output: Kết quả dự đoán dòng tiền tương lai của một công ty.

- Ground.truth: Kết quả dòng tiền tương lai được đánh giá thực tế của tập dữ liệu.
- N: Số mẫu của tập dữ liệu.
- MSE: Sai số trung bình bình phương.
- RMSE: Sai số tiêu chuẩn.
- MAE: Sai số trung bình tuyệt đối.

6. KẾ HOẠCH THỰC HIỆN

Trong giai đoạn thực tập tốt nghiệp, nhóm đã thực hiện được:

- Tìm hiểu tổng quan về dòng tiền và các bài toán liên quan.
- Đề xuất bài toán cụ thể.
- Tìm hiểu các phương pháp hiện thực đề tài.

Kế hoạch dự kiến thực hiện trong giai đoạn luận văn tốt nghiệp:

- Thu thập và tiền xử lý dữ liệu.
- Xây dựng mô hình mạng nơron Kohonen SOM và hiện thực việc phân cụm dữ liệu.
- Xây dựng mô hình mạng nơron đa lớp truyền thẳng MFNN và hiện thực việc dự báo dòng tiền.
- Đánh giá kết quả và so sánh với các phương pháp kinh điển của kinh tế.
- Tổng hợp và viết báo cáo.

TÀI LIỆU THAM KHẢO

- [1] Kate A.Smith, Jatinder N.D. Gupta: Neural networks in business: techniques and applications for the operations researcher. Computers & Operations Research 27, 2000, pp 1023-1044.
- [2] Polat, C. The expert opinions on cash flow forecasting practices: The case of British companies. Soleyman Demirel university, 2003.
- [3] Hwee N. G., Tiong R. L. K.: Model on cash flow forecasting and risk analysis for contracting firms. International Journal of Project Management, 20, 5, 2002, pp. 351-363.
- [4] Orpurt S. F., Zang Y.: Do Direct Cash Flow Disclosures Help Predict Future Operating Cash Flows and Earnings? Accounting Review, 84, 3, 2009, pp. 893-935.
- [5] The International Accounting Standards Board (IASB). International Accounting Standard 7: Statement of cash flows, 24, 3, 2010.
http://ec.europa.eu/internal_market/accounting/docs/consolidated/ias7_en.pdf
- [6] Jonathan Moreland, Individual Investor, 1995.
<http://www.ukessays.com/dissertation/literature-review/literature-review-of-cash-flow-in-finance-finance.php>.
- [7] Madmmod Alborzi, Fraydoon Rahnamay Roodposhti, Somaye sabet: Using neural network for forecasting operating cash flows. Interdisciplinary Journal of Contemporary Research in Business, vol 4, 1, 2013.
- [8] Lai , K.K., Yu L., Wang S.Y. & Huang, W. A novel nonlinear neural network ensemble model for financial time series forecasting. Lecture Notes in Computer Science 3991, 2006, pp. 790 – 793.
- [9] Mohammad Fawzi Shubita: The forecasting ability of earning and Operating cash flow. Interdisciplinary Journal of Contemporary Research in Business, vol 5, 7, 2013.
- [10] Mahmmmod Moeinaddin, Saeid Saeida Ardakani, Fatemeh Akhoondzadeh: Examination the ability of earing and cash flow in predicting future cash flows. Interdisciplinary Journal of Contemporary Research in Business, vol 5, 7, 2013.
- [11] Lokmic, L., Smith, K.A. (2000). Cash flow forecasting using supervised and unsupervised neural networks. Proceedings of the international joint conference on neural networks. Vol. 6, pp. 343 – 347.
- [12] Liu Hongjjiu, Robert Rieg, Hu Yanrong: Performance comparison of artificial intelligence methods for predicting cash flow, 4, 12, 2012.
- [13] Kate A.Smith, Jatinder N.D.Gupta (2002). Neural Network in Bussiness: Techniques and Applications Book. pp. 1 – 26.
- [14] Yan P.W. (2005). Accruals and the prediction of future cash flows in Hong Kong. Hong Kong Baptist University, www.lib-sca.hkbu.edu.hk/trsimage/hp/02004607.pdf

[15] Chan Man - Chung, Wong Chi-Cheong, Lam Chi-Chung (2000). Financial Time series forecasting by neural network using Conjugate Gradient Learning Algorithm and Multiple Linear Regression Weight initialization, Computing in economics and finance. No. 61.

[16] Wong, B. K., Jiang, L. and Lam, J. (2000). A bibliography of neural network business application research: 1994-1998. Computers and Operations Research, 27(11), 1045-1076.

[17] London Stock Exchange. List of all companies, 31, 5, 2015.
<http://www.londonstockexchange.com/statistics/companies-and-issuers/list-of-all-companies.xls>