

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

LƯU ĐỨC DŨNG

TIẾP CẬN PHÂN CỤM CHUỖI THỜI GIAN MỜ TRONG DỰ BÁO
TÍN DỤNG CHO NGÂN HÀNG ABBANK

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN, 2017

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



LƯU ĐỨC DŨNG

**TIẾP CẬN PHÂN CỤM CHUỖI THỜI GIAN MỜ TRONG DỰ BÁO
TÍN DỤNG CHO NGÂN HÀNG ABBANK**

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 01 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Người hướng dẫn khoa học : PGS. TS. Lê Bá Dũng

Thái Nguyên, 2017

MỤC LỤC

MỤC LỤC.....	3
DANH MỤC HÌNH ẢNH	5
DANH MỤC BẢNG BIỂU	6
MỞ ĐẦU.....	8
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT VỀ TẬP MỜ.....	11
1.1 Tập mờ	11
1.1.1 Định nghĩa tập mờ.....	11
1.1.2 Một số khái niệm cơ bản của tập mờ	13
1.1.3 Biểu diễn tập mờ	14
1.2 Các phép toán trên tập mờ và hệ luật mờ.....	15
1.2.1 Phần bù của một tập mờ.....	15
1.2.2 Phép hợp của các tập mờ.....	16
1.2.3 Phép giao của các tập mờ.....	16
1.2.4 Tích Descartes các tập mờ	17
1.2.5 Tính chất của các phép toán trên tập mờ	18
1.2.6 Hệ luật mờ	19
1.3 Lập luận xấp xỉ trong tập mờ	19
1.3.1 Logic mờ	19
1.3.2 Quan hệ mờ	20
1.3.2.1 Khái niệm về quan hệ rõ	20
1.3.2.2 Các quan hệ mờ.....	20
1.3.2.3 Các phép toán của quan hệ mờ	21
1.3.3 Suy luận xấp xỉ và suy diễn mờ	22
1.4 Số học mờ.....	23
1.4.1 Số mờ	23
1.4.1.1 Khái niệm số mờ	23
1.4.1.2 Dạng số mờ thường dùng.....	25
1.4.2 Biến ngôn ngữ và giá trị ngôn ngữ	25

1.5 Giải mờ.....	26
1.5.1 Phương pháp điểm cực đại.....	27
1.5.2 Phương pháp điểm trọng tâm.....	28
CHƯƠNG 2: KHÁI NIỆM VỀ CHUỖI THỜI GIAN MỜ VÀ CÁC PHƯƠNG PHÁP.....	30
2.1 Khái niệm về chuỗi thời gian mờ.....	30
2.1.1 Định nghĩa chuỗi thời gian mờ.....	30
2.1.2 Một số định nghĩa liên quan đến chuỗi thời gian mờ.....	31
2.2 Một số thuật toán dự báo trong mô hình chuỗi thời gian mờ.....	32
2.2.1 Thuật toán của Song & Chissom.....	32
2.2.2 Thuật toán của Chen.....	33
2.3 Một số phương pháp chia khoảng.....	34
2.3.1 Phương pháp độ dài dựa trên sự phân bố giá trị.....	35
2.3.2 Phương pháp độ dài dựa trên giá trị trung bình.....	35
2.4 Thuật toán mô hình dự báo dựa trên chuỗi thời gian mờ của Jens Rùni Poulsen (hay Jens Poulsen).....	36
2.5 Thuật toán phân cụm mờ - Thuật toán K-means.....	40
CHƯƠNG 3: DỰ BÁO TÍN DỤNG ỨNG DỤNG CHUỖI THỜI GIAN MỜ SỬ DỤNG KỸ THUẬT PHÂN CỤM.....	46
3.1 Ứng dụng phương pháp chuỗi thời gian mờ cải tiến cho dự báo tín dụng.....	47
3.2 Tiếp cận một phương pháp mới cho dự báo thời gian mờ.....	47
3.3 Đánh giá phương pháp tiếp cận.....	54
3.4 Kết luận.....	60
TÀI LIỆU THAM KHẢO.....	61

DANH MỤC HÌNH ẢNH

Hình 1.1. Hàm thuộc μ_{Ax} có mức chuyển đổi tuyến tính.....	12
Hình 1.2. Hàm thuộc của tập B.....	12
Hình 1.3. Miền xác định và miền tin cậy của tập mờ A	13
Hình 1.4. Biểu diễn tập mờ chiều cao.....	15
Hình 1.5. Tập bù A của tập mờ A	15
Hình 1.6. Hợp hai tập mờ có cùng tập nền	16
Hình 1.7. Giao hai tập mờ có cùng tập vũ trụ	17
Hình 1.8. Các loại hàm thành viên số mờ	24
Hình 1.9. Phân loại hàm thành viên số mờ	24
Hình 1.10. Số mờ hình thang	25
Hình 1.11. Số mờ hình tam giác	25
Hình 1.12. Những tập mờ thuộc biến ngôn ngữ nhiệt độ	26
Hình 1.13. Giải mờ bằng phương pháp điểm cực đại	28
Hình 1.14. Giải mờ bằng phương pháp điểm trọng tâm	29
Hình 2.1. Các thiết lập để xác định các ranh giới các cụm ban đầu	40
Hình 2.2. Tính toán trọng tâm của các cụm mới.....	41
Hình 2.3. Các bước thực hiện thuật toán K- means	42
Hình 2.4. Thuật toán K-means chi tiết.....	43
Hình 2.5. Ví dụ về một số hình dạng cụm dữ liệu được khám phá bởi K- means.....	45
Hình 3.1. Tín dụng tập mờ về các giá trị dự báo qua các năm	53
Hình 3.2. Giá trị dự báo của mô hình dự báo đề xuất so với giá trị dư nợ tín dụng thực tế.....	59

DANH MỤC BẢNG BIỂU

Bảng 1.1. Bảng biểu diễn tập mờ A	13
Bảng 2.1. Cơ sở ánh xạ	35
Bảng 3.1. Số liệu dư nợ tín dụng của Ngân hàng ABBANK	48
Bảng 3.4. Số liệu tín dụng tập mờ về các giá trị dự báo qua các năm	53
Bảng 3.5. Kết quả dự báo	56
Bảng 3.6 Phân tích kết quả dự báo qua các sai số tiêu chuẩn	57
Bảng 3.7. Bảng so sánh kết quả dự báo giữa các mô hình	58
Bảng 3.8. Bảng các sai số dự báo của các mô hình	59
Bảng 3.9. Bảng so sánh các thước đo sai số của các mô hình	60

LỜI CAM ĐOAN

Tôi xin cam đoan đây là luận văn do tôi nghiên cứu và thực hiện.

Các thông số, bảng biểu và kết quả sử dụng trong luận văn là hoàn toàn có thật và chưa từng được công bố ở bất kỳ luận văn nào khác.

Thái Nguyên, ngày 14 tháng 04 năm 2017

Tác giả luận văn

Lưu Đức Dũng

MỞ ĐẦU

Nghiên cứu chuỗi thời gian nói chung, chuỗi thời gian mờ nói riêng đã và đang được nghiên cứu, ứng dụng mạnh mẽ và thành công ở nhiều lĩnh vực trong những năm gần đây [4, 5, 6, 7]. Các lớp bài toán của các lĩnh vực như dự báo tín dụng, thị trường chứng khoán, dự báo mô phỏng các hệ thống điều khiển... cũng có thể sử dụng và giải quyết theo các phương pháp truyền thống như phương pháp thống kê, quy hoạch tuyến tính, ... Phương pháp nghiên cứu chuỗi thời gian mờ được hình thành có nhiều khả năng vượt trội trong việc tuyến tính hóa, dự báo, phân tích, đánh giá dữ liệu, và áp dụng thành công cho một số lĩnh vực khoa học, kỹ thuật, kinh tế...

Nhiều mô hình chuỗi thời gian mờ (FTS) [6,7] đã được đề xuất trong tài liệu khoa học trong những thập kỷ qua. Trong số các mô hình chuỗi thời gian mờ chính xác nhất được tìm thấy trong tài liệu là những mô hình bậc cao. Tuy nhiên, ba vấn đề cơ bản cần phải được giải quyết liên quan đến các mô hình bậc cao. Đầu tiên, phương pháp dự báo hiện tại đã không thể để cung cấp tỷ lệ chính xác đạt yêu cầu cho đầu ra giải mờ (dự báo). Thứ hai, dữ liệu trở nên ít sử dụng khi tăng thứ bậc. Thứ ba, dự báo chính xác là nhạy cảm với các phân vùng khoảng được lựa chọn.

Nhằm giải quyết những vấn đề trên đề và với gợi ý của thầy hướng dẫn đề tài cho luận văn tốt nghiệp với tiêu đề là: *“Tiếp cận phân cụm chuỗi thời gian mờ trong dự báo tín dụng cho ngân hàng ABBANK”*.

Nội dung chính của luận văn là ứng dụng phương pháp chuỗi thời gian mờ cho dự báo tín dụng được trình bày trong 3 chương như sau:

Chương 1: Cơ sở lý thuyết về tập mờ

Chương 2: Khái niệm về chuỗi thời gian mờ và các phương pháp

Chương 3: Dự báo tín dụng ứng dụng chuỗi thời gian mờ sử dụng kỹ thuật phân cụm

TÀI LIỆU THAM KHẢO

Luận văn này được hoàn thành dưới sự hướng dẫn tận tình của thầy giáo **PGS. TS Lê Bá Dũng**, em xin đặc biệt bày tỏ lòng biết ơn chân thành của mình đối với thầy. Em cũng chân thành cảm ơn các thầy, cô giáo Viện Công nghệ thông tin, Trường Đại học Công nghệ Thông tin và Truyền thông - Đại học Thái Nguyên đã tham gia giảng dạy, giúp đỡ em trong suốt quá trình học tập nâng cao trình độ kiến thức.

Tuy nhiên vì điều kiện thời gian và khả năng có hạn nên luận văn không thể tránh khỏi những thiếu sót. Em kính mong các thầy cô giáo và các bạn đóng góp ý kiến để đề tài được hoàn thiện hơn.

1. Đối tượng và phạm vi nghiên cứu

1.1. Đối tượng

- Tập trung tìm hiểu, tiếp cận chuỗi thời gian mờ về lý thuyết, cấu trúc, phương pháp học và các hạn chế của nó.

- Sau đó áp dụng phương pháp để đánh giá để đạt được kết quả tối ưu cho dự báo về tín dụng.

1.2. Phạm vi nghiên cứu

- Tìm hiểu tổng quan về chuỗi thời gian mờ

- Tìm hiểu, so sánh các phương pháp về ứng dụng chuỗi thời gian mờ.

- Trình bày một phương pháp mới để dự báo tín dụng dựa trên kỹ thuật phân cụm K-means.

2. Hướng nghiên cứu của đề tài

- Nắm bắt các kiến thức cơ bản về phương pháp phân tích, đánh giá.

- Tiếp tục tìm hiểu sâu về các phương pháp khai phá dữ liệu.

- Một số mô hình mờ, trong đó ứng dụng cho việc dự báo tín dụng.

- Cài đặt thực nghiệm, ứng dụng vào một bài toán cụ thể trong thực tiễn.