

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Nguyễn Tiến Hiệp

**XÂY DỰNG HỆ THỐNG HỖ TRỢ RA QUYẾT ĐỊNH HÒA
GIẢI, ĐỐI THOẠI TRONG CÁC TRANH CHẤP
HÔN NHÂN VÀ GIA ĐÌNH**

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

TÓM TẮT LUẬN VĂN THẠC SĨ

HÀ NỘI - 2020

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: **PGS. TS. Đỗ Trung Tuấn**
(Ghi rõ học hàm, học vị)

Phản biện 1: PGS.TS Nguyễn Mạnh Hùng ...

Phản biện 2: TS. Trần Minh Tân.....

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện
Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

Tính từ năm 2012 đến nay, số lượng các vụ án loại này đã tăng gấp hai lần Tòa án luôn trong tình trạng quá tải; nhiều vụ án dân sự, hành chính phải xét xử qua nhiều cấp trong nhiều năm; bản án, quyết định của Tòa có hiệu lực pháp luật nhưng chậm được thi hành đã ảnh hưởng đến quyền, lợi ích hợp pháp của các tổ chức, cá nhân, ảnh hưởng đến niềm tin của người dân đối với Tòa án.

Việc khai phá cơ sở dữ liệu về bản án, quyết định của Tòa án từ đó hỗ trợ các Hòa giải viên, Thẩm phán, lãnh đạo Tòa án có thể xem xét đánh giá các vụ việc sau khi thụ lý và trước khi xét xử, để từ đó có những định hướng hỗ trợ hòa giải, đối thoại có thể giúp các bên giải quyết, rút ngắn thời gian giải quyết vụ việc, tiết kiệm kinh phí của Nhà nước và các bên, hàn gắn những rạn nứt trong các quan hệ xã hội, góp phần xây dựng khối đoàn kết trong nhân dân; qua việc hòa giải, đối thoại, người tiến hành hòa giải, đối thoại còn có thể giải thích, nâng cao nhận thức pháp luật cho các bên, giúp việc thi hành thuận lợi. Xuất phát từ những nhu cầu thực tế trên và đó là những lý do học viên chọn đề tài “*Xây dựng hệ thống trợ giúp ra quyết định hòa giải, đối thoại trong các tranh chấp hôn nhân và gia đình*”.

Luận văn chia thành các chương.

- Chương 1 đề cập hệ thống trợ giúp quyết định, nhu cầu khai phá dữ liệu ra quyết định và các thuật toán cây phân loại ID3 và C4.5;
- Chương 2 thể hiện việc thực hiện phân loại nhờ cây quyết định, sử dụng thuật toán C4.5. Luận văn sử dụng cài đặt J48 trong phần mềm Weka;
- Chương 3 đề cập cơ sở dữ liệu về các án hôn nhân và sử dụng môi trường Visual C# để trợ giúp ra quyết định giải quyết vụ, việc hôn nhân gia đình. Hệ quản trị cơ sở dữ liệu là SQL SERVER.

Cuối luận văn là phần kết luận, tự đánh giá về các kết quả đã đạt được và phương hướng nghiên cứu tiếp theo.

CHƯƠNG 1.

KHAI PHÁ DỮ LIỆU VÀ CÁC HỆ THỐNG RA QUYẾT ĐỊNH

1.1. Tổng quan về khai phá dữ liệu

Khai phá dữ liệu là công việc trích rút tri thức một cách tự động và hiệu quả từ một khối lượng dữ liệu rất lớn. Tri thức đó thường ở dạng các mẫu có tính chất không tầm thường, không tường minh, chưa được biết đến và có tiềm năng mang lại lợi ích. Có một số nhà nghiên cứu còn gọi khai phá dữ liệu là phát hiện tri thức trong cơ sở dữ liệu. Ở đây chúng ta sẽ xem khai phá dữ liệu là cốt lõi của quá trình phát hiện tri thức.

1.1.2. Kiến trúc của hệ thống khai phá dữ liệu

Kiến trúc của một hệ thống khai phá dữ liệu điển hình được cho ở hình trên trong đó: (i) CSDL, kho dữ liệu hoặc các thông tin lưu trữ khác; (ii) Máy chủ CSDL hay máy chủ kho dữ liệu; (iii) Cơ sở tri thức; (iv) Máy khai phá dữ liệu; (v) Module đánh giá mẫu; (vi) Giao diện người dùng.

1.1.3 Các chức năng của khai phá dữ liệu

Nhìn chung các nhiệm vụ của một hệ khai phá dữ liệu có thể được phân chia thành hai loại: mô tả và dự đoán.

1. Công việc khai phá dữ liệu loại mô tả nhằm biểu thị các đặc điểm chung của dữ liệu có trong CSDL.
2. Công việc khai phá dữ liệu loại dự đoán nhằm thực hiện suy luận trên dữ liệu hiện tại để có thể đưa ra dự đoán.

1.1.4. Các phương pháp khai phá dữ liệu

Có nhiều phương pháp thực hiện việc khai phá dữ liệu theo có các loại công cụ chính sau: (i) Các phương pháp thống kê; (ii) Cây quyết định; (iii) Lập luận theo trường hợp; (iv) Tính toán nơ ron; (v) Các tác nhân thông minh; (vi) Các thuật toán di truyền; (vii) Các công cụ khác.

1.1.5. Đặc trưng hóa và phân biệt

Đặc trưng hóa là việc tổng kết các đặc điểm hay tính chất chung của một lớp dữ liệu đích. Dữ liệu đó tương đương với một lớp do người dùng đặc tả bằng một truy vấn CSDL.

1.1.5. Phân tích sự kết hợp

Phân tích sự kết hợp là việc khám phá ra các luật kết hợp trong một tập lớn dữ liệu. Các luật kết hợp thể hiện mối quan hệ giữa các giá trị thuộc tính mà ta nhận thấy được tự tần suất xuất hiện cùng với nhau. Các luật kết hợp được khám phá từ một tập lớn các bản ghi và những tập luật có ý nghĩa có thể giúp cho các nhà doanh nghiệp ra quyết định.

1.1.6. Phân lớp và dự đoán

Phân lớp là quá trình tìm một tập các mô hình (hoặc các chức năng) mô tả và phân biệt các lớp dữ liệu. Các mô hình này sẽ được sử dụng cho mục đích dự đoán về lớp của một đối tượng.

1.1.7. Phân cụm

Không giống như phân lớp và dự đoán, sự phân cụm sẽ phân tích các đối tượng dữ liệu khi chưa biết nhãn của lớp, nghĩa là nhãn của lớp không tồn tại trong quá trình huấn luyện dữ liệu. Phân cụm có thể được sử dụng để đưa ra những nhãn của lớp.

1.1.8. Phân tích phần tử ngoài cuộc

Một cơ sở dữ liệu có thể chứa các đối tượng dữ liệu không tuân theo mô hình dữ liệu. Những đối tượng như vậy được gọi là phần tử ngoài cuộc. Hầu hết các phương pháp khai phá dữ liệu đều coi các phần tử ngoài cuộc là nhiễu và loại bỏ chúng.

1.2. Khái niệm về hệ thống hỗ trợ ra quyết định

1.2.1. Quyết định

1.2.1.1. Khái niệm về quyết định

Theo Simon (1960); Costello & Zalkind 1963; Churchman (1968) đó là một lựa chọn về “đường lối hành động”, hay “chiến lược hành động” Fishburn (1964) dẫn đến một mục tiêu mong muốn” Churchman (1968).

“Một quá trình lựa chọn có ý thức giữa hai hay nhiều phương án để chọn ra một phương án tạo ra được kết quả mong muốn trong các điều kiện ràng buộc đã biết”.

1.2.1.2. Hiểu rõ thêm về ra quyết định

Việc đưa ra quyết định đối với một vấn đề xuất hiện trong khắp các lĩnh vực, hoạt động của đời sống mà đôi khi chúng ta không nhận ra. Từ những việc đơn giản

nghư chọn một bộ quần áo để đi dự tiệc cho đến các công việc lớn lao như phân bổ ngân sách vào các chương trình của quốc gia đều là các công việc đưa ra quyết định.

Vậy đưa ra quyết định chính là chọn ra trong các giải pháp khả thi một giải pháp mà theo người đưa ra quyết định là phù hợp nhất.

1.2.2. Quá trình ra quyết định

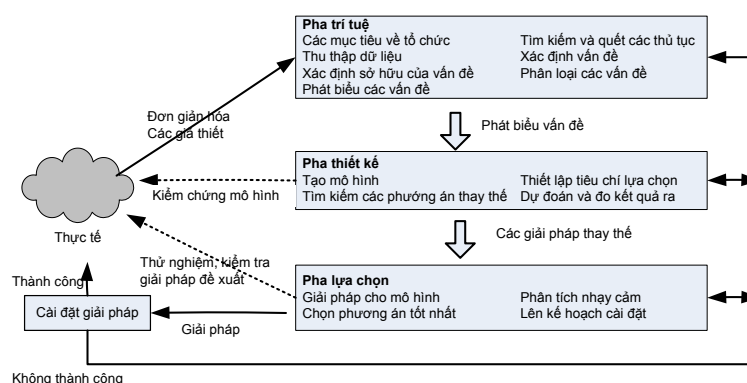
1.2.2.1. Phân loại quyết định

Có thể phân ra bốn loại quyết định như sau:

- Quyết định có cấu trúc (Structured Decision): Các quyết định mà người ra quyết định biết chắc chắn đúng. Ví dụ: Bài toán quyết định thưởng, phạt Nhân viên.
- Quyết định không có cấu trúc (NonStructured Decision): Các quyết định mà người ra quyết định biết là có nhiều câu trả lời gần đúng và không có cách nào để tìm ra câu trả lời chính xác nhất. Ví dụ: Bài toán quyết định chiến lược phát triển của Nhà trường.
- Quyết định đệ quy (Recurring Decision): Các quyết định lặp đi lặp lại.
- Quyết định không đệ quy (Nonrecurring Decision): Các quyết định không xảy ra thường xuyên.

1.2.2.2. Các giai đoạn của quá trình ra quyết định

Theo Simon, quá trình ra quyết định và quan hệ giữa chúng được giới thiệu ở hình dưới đây:



Hình 1.1. Các giai đoạn của quá trình ra quyết định

(Nguồn: “Hệ trợ giúp quyết định”, nxb. Đại học Quốc gia Hà Nội, 2016)

- Giai đoạn thứ nhất là nhận định (Intelligence): Tìm kiếm các tình huống dẫn

đến việc phải ra quyết định, nhận dạng các vấn đề, nhu cầu, cơ hội, rủi ro.

- Giai đoạn thứ hai là thiết kế (Design): Phân tích các hướng tiếp cận để giải quyết vấn đề, đáp ứng các nhu cầu, tận dụng các cơ hội, hạn chế các rủi ro.
- Giai đoạn thứ ba là lựa chọn (Choice): Cân nhắc và đánh giá từng giải pháp và chọn giải pháp tối ưu.
- Cuối cùng là tiến hành ra quyết định (Implementation): Thực hiện giải pháp được chọn, theo dõi kết quả và điều chỉnh khi thấy cần thiết.

1.2.2.3. Tìm kiếm và đánh giá các lựa chọn một phần rất quan trọng trong hỗ trợ ra quyết định.

Giai đoạn lựa chọn (Choice Phase) là giai đoạn quan trọng nhất của quá trình ra quyết định. Giai đoạn này bao gồm ba bước chính sau đây: (i): Tìm kiếm lựa chọn; (ii): Đánh giá lựa chọn; (iii): Giới thiệu lựa chọn.

1.2.3. Khái niệm hệ hỗ trợ quyết định

Trong thập niên 1970, Scott Norton đưa những khái niệm đầu tiên về hệ trợ giúp quyết định (Decision Support System, DSS). Ông định nghĩa “DSS là các hệ dựa trên máy tính, có tính tương tác, giúp các nhà ra quyết định dùng dữ liệu và mô hình để giải các bài toán phi cấu trúc, những bài toán mờ, phức tạp với lời giải không hoàn chỉnh”.

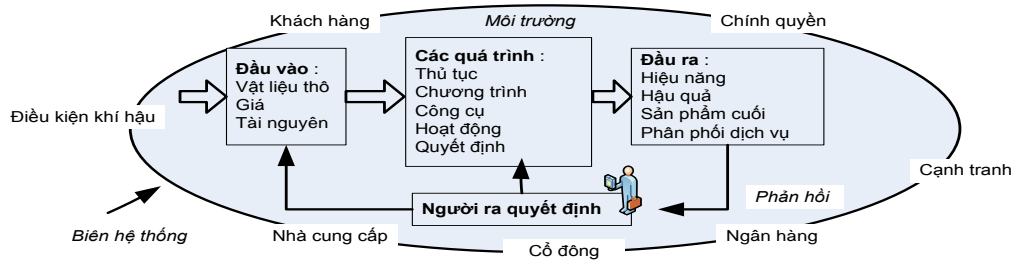
Cho đến nay chưa có một định nghĩa thống nhất về DSS. Tuy nhiên tất cả đều đồng ý mục đích cơ bản nhất của DSS là hỗ trợ và cải tiến việc ra quyết định.

1.3. Các thành phần của hệ thống ra quyết định

1.3.1. Các thành phần

Một hệ hỗ trợ quyết định gồm có ba thành phần chính:

- Quản lý dữ liệu;
- Quản lý mô hình;
- Quản lý giao diện người dùng.



Hình 1.2. Hệ thống ra quyết định và môi trường của nó

(Nguồn: “Hệ trợ giúp quyết định”, nxb. Đại học Quốc gia Hà Nội, 2016)

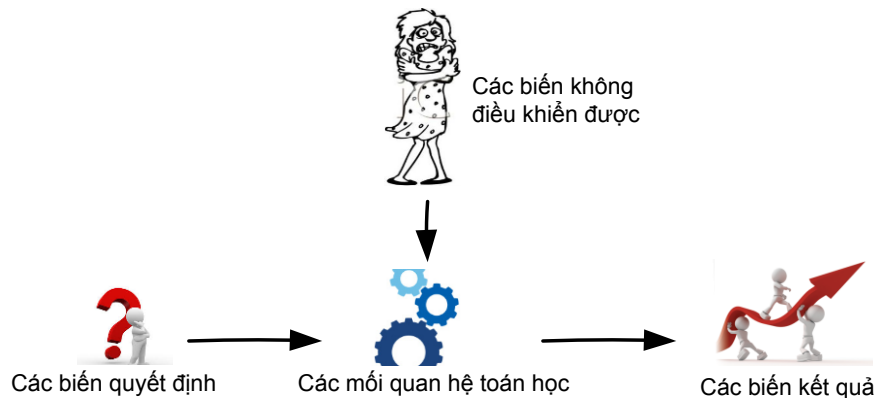
1.3.2. Mô hình ra quyết định

Một đặc trưng cơ bản của hệ hỗ trợ ra quyết định là phải có ít nhất một mô hình hỗ trợ ra quyết định. Việc chọn lựa và xây dựng mô hình nằm trong giai đoạn thứ 2 (Design Phase) của quá trình ra quyết định.

Mô hình là một khái quát hóa hay trừu tượng hóa các vấn đề thực tế thành các mô hình định tính hay định lượng. Đó là một quy trình kết hợp cả khoa học (sự chính xác, logic) và nghệ thuật (sự sáng tạo).

Một mô hình gồm ba thành phần cơ bản:

- Decision Variables: Đây là các lựa chọn xác định bởi người ra quyết định. Chẳng hạn trong bài toán quyết định thưởng phạt nhân viên ...



Hình 1.3. Cấu trúc chung của mô hình định lượng

(Nguồn: “Hệ trợ giúp quyết định”, nxb. Đại học Quốc gia Hà Nội, 2016)

- Uncontrollable Variables: Đây là các biến không nằm trong sự kiểm soát của người ra quyết định (bị tác động bởi các yếu tố bên ngoài). Chẳng hạn trong bài toán trên thì đây là...
- Result Variables: Đây là biến kết quả của mô hình. Chẳng hạn trong bài

toán trên đây...

1.4. Phân loại các hệ thống ra quyết định

Hệ hỗ trợ ra quyết định được phân loại dựa trên nhiều tiêu chí. Hiện nay, vẫn chưa có cách phân loại thống nhất. Sau đây là hai cách phổ biến nhất.

Theo [4] có tất cả năm loại hệ hỗ trợ ra quyết định:

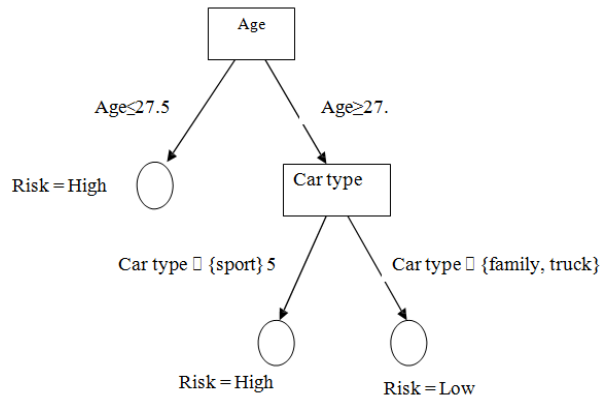
1. Hướng giao tiếp (Communication, Driven DSS);
2. Hướng dữ liệu (Data, Driven DSS);
3. Hướng tài liệu (Document, Driven DSS);
4. Hướng tri thức (Knowledge, Driven DSS);
5. Hướng mô hình (Model, Driven DSS).

1.5. Cây quyết định

1.5.1 Khái niệm

Cây quyết định là một cấu trúc ra quyết định có dạng cây. Cây quyết định nhận đầu vào là một bộ giá trị thuộc tính mô tả một đối tượng hay một tình huống và trả về một giá trị rời rạc. Mỗi bộ thuộc tính đầu vào được gọi là một mẫu hay một ví dụ, đầu ra gọi là loại hay nhãn phân loại. Thuộc tính đầu vào còn được gọi là đặc trưng và có thể nhận giá trị rời rạc hoặc liên tục. Để cho đơn giản, trước tiên ta sẽ xem xét thuộc tính rời rạc, sau đó sẽ mở rộng cho trường hợp thuộc tính nhận giá trị liên tục.

Cây quyết định được biểu diễn dưới dạng một cấu trúc cây (xem ví dụ trên hình 1.4). Mỗi nút trung gian, tức là nút không phải nút lá, tương ứng với phép kiểm tra một thuộc tính.



Hình 1.4. Ví dụ về cây quyết định

Trong cây quyết định:

- Gốc: là node trên cùng của cây
- Nút trong: biểu diễn một kiểm tra trên một thuộc tính đơn (hình chữ nhật)
- Nhánh: biểu diễn các kết quả của kiểm tra trên nút trong (mũi tên)
- Nút lá: biểu diễn lớp hay sự phân phối lớp (hình tròn)

1.5.3. Đánh giá cây quyết định trong lĩnh vực khai phá dữ liệu

1.5.3.1 Sức mạnh của cây quyết định

Cây quyết định có 5 sức mạnh chính sau [12]:

- Khả năng sinh ra các quy tắc hiệu được.
- Khả năng thực thi trong những lĩnh vực hướng quy tắc.
- Dễ dàng tính toán trong khi phân lớp.
- Khả năng xử lý với cả thuộc tính liên tục và thuộc tính rời rạc.
- Thể hiện rõ ràng những thuộc tính tốt nhất.

1.5.3.2. Điểm yếu của cây quyết định

Dù có những sức mạnh nổi bật trên, cây quyết định vẫn không tránh khỏi có những điểm yếu. Đó là cây quyết định không thích hợp lắm với những bài toán với mục tiêu là dự đoán giá trị của thuộc tính liên tục như thu nhập, huyết áp hay lãi xuất ngân hàng, ... Cây quyết định cũng khó giải quyết với những dữ liệu thời gian liên tục nếu không bỏ ra nhiều công sức cho việc đặt ra sự biểu diễn dữ liệu theo các mẫu liên tục.

- Dễ xảy ra lỗi khi có quá nhiều lớp.
- Chi phí tính toán đắt để đào tạo.

1.5.3.3 Xây dựng cây quyết định

Quá trình xây dựng cây quyết định gồm hai giai đoạn:

- *Giai đoạn thứ nhất phát triển cây quyết định*
- *Giai đoạn thứ hai cắt, tỉa bớt các cành nhánh trên cây quyết định.*

1.5.3.4. Giải thuật xây dựng cây quyết định

Trước khi sử dụng cây quyết định, ta cần xây dựng hay “học” cây quyết định từ dữ liệu huấn luyện. Có nhiều thuật toán khác nhau được đề xuất và sử dụng để học cây quyết định từ dữ liệu, trong đó đa số dựa trên nguyên tắc chung là xây dựng cây theo kiểu tìm kiếm tham lam, chia để trị, đệ qui từ cây đơn giản tới cây phức tạp hơn.

1.6. Các thuật toán cây quyết định

1.6.1. Thuật toán ID3

Là thuật toán được phát triển năm 1986 bởi Ross Quinlan. Tên đầy đủ là Iterative Dichotomiser 3 (Cây phân đôi lặp).

Mục đích là tạo ra cây nhiều nhánh, tại mỗi nút của cây xác định một đặc trưng phân loại mang Information Gain lớn nhất đối với các mục tiêu phân loại. Cây sẽ được tạo ra với số nút nhiều nhất có thể, sau đó là quy trình Pruning (cắt tỉa) cây để tăng khả năng phân loại và tổng quát hóa những dữ liệu chưa có trong cây.

1.6.2. Thuật toán C4.5

Là phiên bản cải tiến của thuật toán ID3. Loại bỏ điều kiện hạn chế của ID3: các đặc trưng phải là đặc trưng có thể phân loại bằng cách xác định một thuộc tính rời rạc (dựa trên các biến số) sao cho thuộc tính này có thể chia giá trị thuộc tính liên tục thành tập các khoảng rời rạc. C4.5 có những đặc điểm khác với các thuật toán khác, đó là: cơ chế chọn thuộc tính để kiểm tra tại mỗi node. C4.5 dựa vào nghiên cứu tối ưu hóa, và sự lựa chọn cách phân chia mà có độ đo lựa chọn thuộc tính đạt giá trị cực đại. ai độ đo được sử dụng trong C4.5 là Information Gain và Gain Ratio.

1.7. Kết luận

Chương 1 của luận văn đã trình bày hệ thống trợ giúp quyết định DSS và công việc khai phá dữ liệu.

Cây quyết định được sử dụng để xây dựng một kế hoạch nhằm đạt được mục tiêu mong muốn đó là hỗ trợ ra quyết định.

Đối với C4.5 cơ chế sinh cây quyết định hiệu quả và chặt chẽ bằng việc sử

dùng độ đo lựa chọn thuộc tính tốt nhất là information-gain. Các cơ chế xử lý với giá trị lỗi, thiếu và chống “quá vừa” dữ liệu của C4.5 cùng với cơ chế cắt tỉa cây đã tạo nên sức mạnh của C4.5.

Ngoài ra, mô hình phân lớp C4.5 còn có phần chuyển đổi từ cây quyết định sang luật dạng if-then, làm tăng độ chính xác và tính dễ hiểu của kết quả phân lớp. Đây là tiện ích rất có ý nghĩa đối với người sử dụng.

CHƯƠNG 2.

THỬ NGHIỆM HỆ THỐNG TRỢ GIÚP RA QUYẾT ĐỊNH HÒA GIẢI, XÉT XỬ

Chương này trình bày việc sử dụng phần mềm Weka để thực hiện thuật toán phân lớp dữ liệu án hôn nhân. Lựa chọn sử dụng thuật toán J48 vì thuật toán J48 được Weka cài đặt trên thuật toán C4.5.

2.1. Chuẩn bị dữ liệu

Dữ liệu đầu vào được cung cấp dưới dạng tệp excel với 265 bản án, quyết định đã có hiệu lực của Tòa án là dữ liệu mẫu của Tòa án nhân dân thành phố Hà Nội về số theo dõi giải quyết các vụ, việc về hôn nhân gia đình.

Đưa tập dữ liệu vào xử lý: Là bảng tính excel như sau:

Mã vụ việc	Ngày thụ lý	Bảng thông tin vụ án										Thẩm phán	Quyết định
		Nguyên đơn (vợ/chồng)				Bị đơn (vợ/chồng)				Thông tin con chung	Quan hệ pháp luật		
		Họ và tên	Năm sinh	Tuổi	Địa chỉ	Họ và tên	Năm sinh	Tuổi	Địa chỉ				
1	19/05/2016	Phạm Thị Cẩm Dung	1990	26	791/5/47/42 Bến Phú Định,	Phạm Thắng Lợi	1989	27	Phong 4 C3 Tô 21 Ngõ 124	có	(4) Do nghiên cứu, cơ bạc, ma túy (TN)	Bùi Thị Thanh Phương	(2) Xét xử (XX)
2	06/05/2016	Nguyễn Thanh Bình	1987	29	P512 C1 Thanh Xuân Bắc,	Nguyễn Duy Pháp	1987	29	P512 C1 Thanh Xuân Bắc,	có	(8) Bạo lực gia đình (BLGD)	Bùi Thị Thiếc	(2) Xét xử (XX)
3	16/05/2016	Đỗ Thị Nhung	1992	24	Thôn Tân An, xã Thụy	Nguyễn Đình Sơn	1992	24	Tổ 20, phường phú lương,	có	(1) Mâu thuẫn gia đình (MTGD)	CHU MINH SANG	(1) Hòa giải (HG)
4	23/05/2016	Lê Thị Uyên	1992	24	Tổ 8 khu phố dân Mộ Lao, Hà	Nguyễn Tiến Văn	1989	27	Tổ 8 khu phố dân Mộ Lao,	có	(1) Mâu thuẫn gia đình (MTGD)	DIỆP LÊ QUYNH ANH	(2) Xét xử (XX)
5	16/03/2016	Thân Thị Thảo	1989	27	số 59 ngách 57/88 ngõ 57	Nguyễn Ngọc Tuấn	1984	32	số 59 ngách 57/88 ngõ 57	có	(3) Ngoại tình (NT)	Danh Đới	(1) Hòa giải (HG)
6	10/05/2016	Nguyễn Thị Hào	1992	24	số 20 ngõ 111 đường Trung	Đào Tiến Long	1989	27	số 20 ngõ 111 đường Trung	có	(1) Mâu thuẫn gia đình (MTGD)	Dương Hồng Phương	(1) Hòa giải (HG)
7	20/09/2016	Phan Thị Hồng Nhung	1991	25	TDP Phú Đô, phường Phú Đô,	Hoàng Bá Tấn	1991	25	Ngã Sơn, Cầu Khê, Phú Thọ	có	(4) Do nghiên cứu, cơ bạc, ma túy (TN)	Dương Thị Huyền Trang	(2) Xét xử (XX)
8	02/03/2016	Đỗ Thị Hương	1973	43	xóm Chọ, phường Đại Mỗ,	Trần Đức Quyền	1959	57	Cao Sơn, Tiến Phương,	có	(1) Mâu thuẫn gia đình (MTGD)	Gián Thị Dung	(2) Xét xử (XX)
9	15/10/2016	Nguyễn Quỳnh Anh	1985	31	P304 nhà 2 tập thể Bộ Thương	Phạm Ngọc Anh	1977	39	P304 nhà A 2 Tập thể Bộ	có	(7) Mâu thuẫn kinh tế (MTKT)	HOANG THI XUÂN	(2) Xét xử (XX)
10	07/06/2016	Phạm Anh Hồng	1981	35	P9K8 ngõ 43 đường Trương	Đặng Minh Quang	1977	39	Số 2 phường Dịch Vọng Hậu	có	(7) Mâu thuẫn kinh tế (MTKT)	HUYNH THUY NGÂN	(2) Xét xử (XX)
11	20/06/2016	Trịnh Thị Hải Duyên	1982	34	Tổ 10 Đại Từ	Nguyễn Khắc Tín	1981	35	Tổ 10 Đại Từ	có	(7) Mâu thuẫn kinh tế (MTKT)	Hà Hoa Thiên	(2) Xét xử (XX)
12	19/08/2016	Nguyễn Thị Minh Thủy	1989	27	P114 tập thể Thủ y phường	Hoàng Quốc Lạc	1985	31	Tổ 5C	có	(7) Mâu thuẫn kinh tế (MTKT)	Hà Quỳnh Nga	(1) Hòa giải (HG)
13	15/08/2016	Kim Bảo Giang	1985	31	P506 N21 khu đô thị Pháp	Lê Vũ Long	1978	38	Tổ 2 p. Mai Động	có	(1) Mâu thuẫn gia đình (MTGD)	HÔNGCÓC TRƯỞNG	(2) Xét xử (XX)
14	12/07/2016	Hoàng Thị Xuân Thu	1978	38	1A/Q6 trường Đinh , Nguyễn	Phạm Hoàng Anh	1976	40	1A/Q6 Trường Đinh , Nguyễn	có	(3) Ngoại tình (NT)	HỒ THỊ KIM OANH	(1) Hòa giải (HG)
15	28/07/2016	Trương Thị Thu Hiền	1982	34	Số 2 tổ 17 ngách 351/108/16	Trần Đình Hai	1977	39	Số 2 tổ 17 ngách 351/108/16	có	(1) Mâu thuẫn gia đình (MTGD)	HỒ THỊ KIỀU TRANG	(1) Hòa giải (HG)
16	12/08/2016	Tạ Phương Thảo	1989	27	P609N10 Khu Đồng Tàu	Đỗ Văn Lợi	1985	31	P609N10 Khu Đồng Tàu	có	(3) Ngoại tình (NT)	Hồ Quỳnh Trâm	(1) Hòa giải (HG)
17	10/05/2016	Hà Thị Hà	1993	23	Tổ 22 phường Hoàng Văn	Lê Xuân Đạt	1982	34	Tổ 22 phường Hoàng Văn	Không	(5) Bệnh tật, không có con (BT)	Hồ Thị Thu Thảo	(1) Hòa giải (HG)
18	31/08/2016	Nguyễn Thị Lý	1978	38	Số 89 ngõ 259 tổ 49	Dương Văn Việt	1964	52	Số 89 ngõ 259 tổ 49	có	(1) Mâu thuẫn gia đình (MTGD)	Hồ Thị Yên Khanh	(1) Hòa giải (HG)
19	29/10/2016	Lê Thị Tinh	1975	41	P712 N3 Bán Đảo Linh Đàm	Nguyễn Văn Thành	1974	42	P712 N3 Bán Đảo Linh Đàm	có	(1) Mâu thuẫn gia đình (MTGD)	LÊ KHÁNH LINH	(2) Xét xử (XX)
20	03/11/2016	Phạm Thị Hà	1981	35	411 C10 Tân Mai	Nguyễn Trung Kiên	1979	37	411 C10 Tân Mai	có	(4) Do nghiên cứu, cơ bạc, ma túy (TN)	LÊ NGỌC TÚ QUYÊN	(2) Xét xử (XX)
21	28/09/2016	Minh Thị Lê Chính	1976	40	164 Trường Đinh, P.Trương	Đặng Hữu Nhuận	1968	48	số 9/69/255 Linh Nam	có	(1) Mâu thuẫn gia đình (MTGD)	LÊ THỊ HUƠNG	(1) Hòa giải (HG)
22	08/10/2016	Hoàng Thị Hải Hà	1992	24	A9 tập thể xí nghiệp cung ứng	Nguyễn Công Hà	1988	28	A9 tập thể xí nghiệp cung ứng	có	(4) Do nghiên cứu, cơ bạc, ma túy (TN)	LÊ THỊ THANH VÂN	(1) Hòa giải (HG)
23	20/10/2016	Nguyễn Thị Chính	1989	27	Số 16B tập thể Thô Nương	Dương Quốc Việt	1985	31	P101B C29	có	(4) Do nghiên cứu, cơ bạc, ma túy (TN)	LÊ VĂN TUẤN	(2) Xét xử (XX)

Hình 2.1 Dữ liệu số theo dõi các vụ việc hôn nhân gia đình

(Nguồn: Tòa án nhân dân thành phố Hà Nội)

Trong qui trình khai phá dữ liệu, công việc xử lý dữ liệu trước khi đưa vào các mô hình là rất cần thiết, bước này làm cho dữ liệu có được ban đầu qua thu thập dữ liệu (gọi là dữ liệu gốc ordinal data) có thể áp dụng được (thích hợp) với các mô hình khai phá dữ liệu (data mining model) cụ thể.

Tất cả các giá trị cụ thể và gán nhãn được thể hiện trong bảng sau:

Bảng 2.1. Biến số hóa dữ liệu “độ tuổi”

STT	Giá trị thuộc tính	Gán nhãn
1	Độ tuổi của nguyên đơn nhỏ hơn hoặc bằng 30.	tuoi_nguyen_don: <=30
2	Độ tuổi của nguyên đơn lớn hơn 30.	tuoi_nguyen_don: >30
3	Độ tuổi của bị đơn nhỏ hơn hoặc bằng 30.	tuoi_bi_don: <=30
4	Độ tuổi của bị đơn lớn hơn 30.	tuoi_bi_don: >30

Bảng 2.2. Biến số hóa dữ liệu “con chung”

STT	Giá trị thuộc tính	Gán nhãn
1	Có con chung	co_chung: co
2	Không có con chung	co_chung: khong

Bảng 2.3. Biến số hóa dữ liệu “độ lệch tuổi”

STT	Giá trị thuộc tính	Gán nhãn
1	Chênh lệch tuổi của nguyên đơn và bị đơn trong khoảng nhỏ hơn hoặc bằng 5.	do_lech_tuoi: <=5
2	Chênh lệch tuổi của nguyên đơn và bị đơn trong khoảng lớn hơn hoặc 5 và nhỏ hoặc bằng 10	do_lech_tuoi: 5_10
3	Chênh lệch tuổi của nguyên đơn và bị đơn lớn hơn 10.	do_lech_tuoi: >10

Bảng 2.4. Biến số hóa dữ liệu “quan hệ pháp luật”

STT	Giá trị thuộc tính	Gán nhãn
1	Mâu thuẫn gia đình.	quan_he_phap_luat: MTGD
2	Yếu tố nước ngoài	quan_he_phap_luat: NN
3	Ngoại tình	quan_he_phap_luat: NT
4	Cờ bạc, rượu chè, ma túy	quan_he_phap_luat: TNXH
5	Bệnh tật, không có con	quan_he_phap_luat: BT
6	Một người mất tích	quan_he_phap_luat: MT
7	Mâu thuẫn kinh tế	quan_he_phap_luat: MTKT
8	Bạo lực gia đình	quan_he_phap_luat: BLGD

Bảng 2.5. Ý nghĩa biến “quyết định”

STT	Giá trị thuộc tính	Gán nhãn
1	Hòa giải	quyet_dinh: Hoagiai
2	Xét xử	quyet_dinh: Xetxu

Sau khi được chuẩn hóa dựa trên các bảng mô tả trên có được dữ liệu đầu ra là tệp data_toaan.arff:

```

1 @relation cases
2
3 @attribute quan_he_phap_luat {MTGD,NN,NT,TNXH,BT,MT,MTKT,BLG
4 @attribute co_chung {khong,co}
5 @attribute tuoi_nguyen_don {<=30,>30}
6 @attribute tuoi_bi_don {<=30,>30}
7 @attribute do_lech_tuoi {<=5,5_10,>10}
8 @attribute quyet_dinh {Hoagiai,Xetxu}
9
10 @data
11 TNXH,co,<=30,<=30,<=5,Xetxu
12 BLGD,co,<=30,<=30,<=5,Xetxu
13 MTGD,co,<=30,<=30,<=5,Hoagiai
14 MTGD,co,<=30,<=30,<=5,Xetxu
15 NT,co,<=30,>30,<=5,Hoagiai
16 MTGD,co,<=30,<=30,<=5,Hoagiai
17 TNXH,co,<=30,<=30,<=5,Xetxu
18 MTGD,co,>30,>30,>10,Xetxu
19 MTKT,co,>30,>30,5_10,Xetxu
  
```

Hình 2.2 Dữ liệu sau chuẩn hóa

2.2. Sử dụng J48

a. Lựa chọn chỉ số **Information Gain**: Sau khi chạy thuật toán này với các lựa chọn trên ta thu được kết quả như sau:

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Choose: AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval" -S "weka.attributeSelection.Ranker-T-1.7876931348623157E308-N-1" -W weka.classifiers.trees.J48 --C 0.5-M 2

Test options

☒ Use training set
☐ Supplied test set
☐ Cross-validation Folds: 10
☐ Percentage split %: 65
 More options...

(Nom) quyet_dinh

Start Stop

Result list (right-click for options)

14:35:27 - meta:AttributeSelectedClassifier

Classifier output

Instances: 265
 Attributes: 6
 quan_he_phap_luat
 co_chung
 tuoi_nguyen_don
 tuoi_bi_don
 do_lech_tuoi
 quyet_dinh
 Test mode: evaluate on training data

=== Classifier model (full training set) ===

AttributeSelectedClassifier:

=== Attribute Selection on all input data ===

Search Method:
 Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 6 quyet_dinh):
 Information Gain Ranking Filter

Ranked attributes:
 0.2602 1 quan_he_phap_luat
 0.02204 3 tuoi_nguyen_don
 0.01021 5 do_lech_tuoi
 0.00511 2 co_chung
 0.00137 4 tuoi_bi_don

Selected attributes: 1,3,5,2,4 : 5

Header of reduced data:
 @relation "cases-weka.filters.unsupervised.attribute.Remove-V-R1,3,5,2,4,6"
 @attribute quan_he_phap_luat {MTGD,NN,NT,TNXH,BT,MT,MTKT,BLG}

Status
 OK

Log

Hình 2.3. Kết quả thực hiện với Information Gain

Đánh giá kết quả phân lớp với mẫu dữ liệu số lá của cây là 14, kích thước là của cây 20 và thời gian thực hiện 0.01s.

i. Các chỉ số phân lớp:

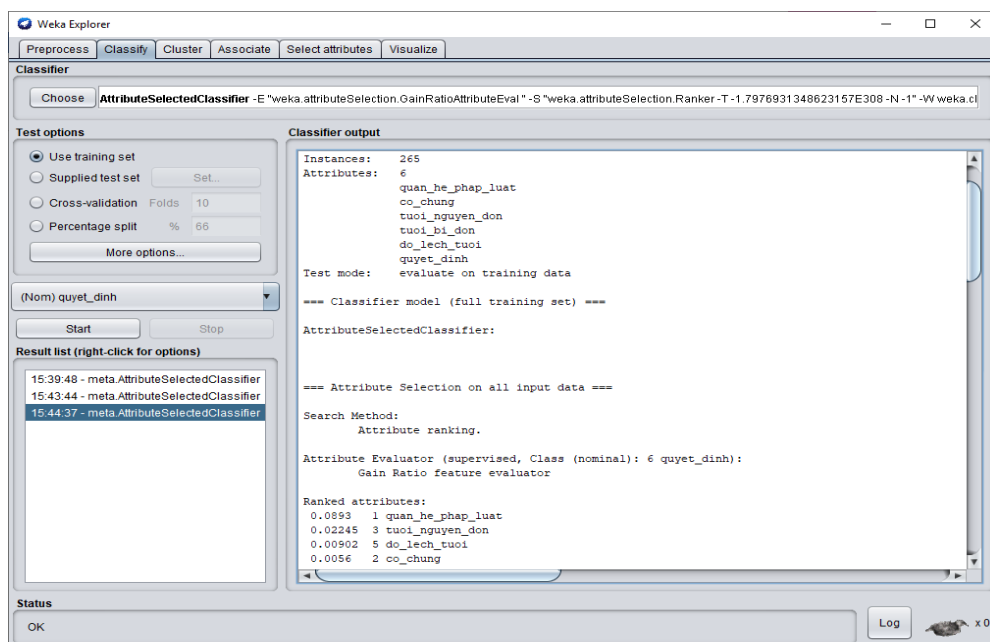
- Độ phân lớp chính xác của dữ liệu: Độ phân lớp đạt kết quả chính xác rất cao của thuật toán với 221/265 tương đương 83.3962 %;
- Độ phân lớp không chính xác của dữ liệu: Độ phân lớp đạt kết quả không chính xác của thuật toán 44/265 tương đương 16.6038 %.

ii. Đánh giá chỉ số Information Gain cho từng thuộc tính:

Bảng 2.6. Bảng xếp hạng chỉ số Information Gain

Xếp hạng	Thuộc tính đặc trưng	Chỉ số Information Gain
1	quan_he_phap_luat	0.26020
2	tuoi_nguyen_don	0.02206
3	do_lech_tuoi	0.01021
4	co_chung	0.00511
5	tuoi_bi_don	0.00137

b. Lựa chọn chỉ số ***Gain Ratio***: Sau khi chạy thuật toán này với các lựa chọn trên ta thu được kết quả như sau:



Hình 2.3. Kết quả thực hiện với Gain Ratio

Đánh giá kết quả phân lớp với mẫu dữ liệu số lá của cây là 14, kích thước lá của cây 20 và thời gian thực hiện 0.01s.

i. Các chỉ số phân lớp:

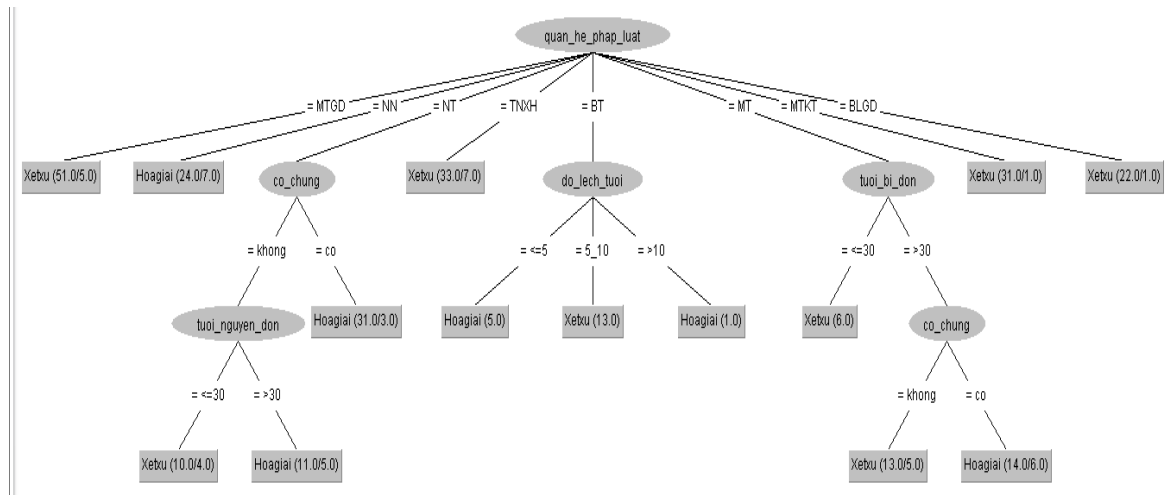
- Độ phân lớp chính xác của dữ liệu: Độ phân lớp đạt kết quả chính xác rất cao của thuật toán với 221/265 tương đương 83.3962 %;
- Độ phân lớp không chính xác của dữ liệu: Độ phân lớp đạt kết quả không chính xác của thuật toán 44/265 tương đương 16.6038 %.

ii. Đánh giá chỉ số Information Gain cho từng thuộc tính:

Bảng 2.7. Bảng xếp hạng chỉ số Information Gain

Xếp hạng	Thuộc tính đặc trưng	Chỉ số Information Gain
1	quan_he_phap_luat	0.0893
2	tuoi_nguyen_don	0.02245
3	do_lech_tuoi	0.00902
4	co_chung	0.0056
5	tuoi_bi_don	0.00164

Cả hai lựa chọn trên đều cho một kết quả là mô hình cây quyết định:



Hình 2.3. Kết quả thực hiện

Cả hai lựa chọn trên đều cho sinh ra cùng một cây và bộ luật như nhau:

1. Rule 1: IF quan_he_phap_luat = "MTGD" THEN quyet_dinh="Xetxu";
2. Rule 2: IF quan_he_phap_luat = "NN" THEN quyet_dinh="Hoagiail";

3. Rule 3: IF quan_he_phap_luat = "NT" AND co_chung = "khong" AND tuoi_nguyen_don="<=30" THEN quyet_dinh="Xetxu";
4. Rule 4: IF quan_he_phap_luat = "NT" AND co_chung = "khong" AND tuoi_nguyen_don=">30" THEN quyet_dinh="Hoagiai";
5. Rule 5: IF quan_he_phap_luat = "NT" AND co_chung = "co" THEN quyet_dinh="Hoagiai";
6. Rule 6: IF quan_he_phap_luat = "TNXH" THEN quyet_dinh="Xetxu";
7. Rule 7: IF quan_he_phap_luat = "BT" AND do_lech_tuoi="<=5" THEN quyet_dinh="Hoagiai";
8. Rule 8: IF quan_he_phap_luat = "BT" AND do_lech_tuoi="5_10" THEN quyet_dinh="Xetxu";
9. Rule 9: IF quan_he_phap_luat = "BT" AND do_lech_tuoi=">10" THEN quyet_dinh="Hoagiai";
10. Rule 10: IF quan_he_phap_luat = "MT" AND tuoi_bi_don="<=30" THEN quyet_dinh="Xetxu";
11. Rule 11: IF quan_he_phap_luat = "MT" AND tuoi_bi_don=">30" AND co_chung="khong" THEN quyet_dinh="Xetxu";
12. Rule 12: IF quan_he_phap_luat = "MT" AND tuoi_bi_don=">30" AND co_chung="co" THEN quyet_dinh="Hoagiai";
13. Rule 13: IF quan_he_phap_luat = "MTKT" THEN quyet_dinh="Xetxu";
14. Rule 14: IF quan_he_phap_luat = "BLGD" THEN quyet_dinh="Xetxu";

2.3. Kết luận

Từ dữ liệu đầu vào đánh giá độ hiệu quả phân lớp dữ liệu là rất cao với độ chính xác là 83.3962 %, không chính xác của thuật toán chỉ 16.6038 %.

Với việc có thể tự thay đổi quyết định và thêm quyết định thực tế vào cơ sở dữ liệu sẵn có, phần mềm sẽ không ngừng được phát triển và hoàn thiện nhằm đưa ra những quyết định chính xác, sát với thực tế hơn.

Chương trên đã thể hiện việc phân lớp các án hôn nhân với phần mềm Weka.

Với bộ luật được sinh là cơ sở để thực hiện xây dựng phần mềm hệ thống hỗ

trợ ra quyết định về hòa giải, đối thoại được trình bày tại chương 3 của luận văn.

CHƯƠNG 3.

XÂY DỰNG HỆ THỐNG HỖ TRỢ RA QUYẾT ĐỊNH VỀ CÁC TRANH CHẤP HÔN NHÂN VÀ GIA ĐÌNH

Chương này trình bày hiện trạng xử lý dữ liệu xử án hôn nhân và phân tích, thiết kế, xây dựng cơ sở dữ liệu trên SQL SERVER, nhằm trợ giúp quyết định xử án hôn nhân.

3.1. Nhu cầu về cơ sở dữ liệu các bản án hôn nhân gia đình

3.1.1. Nhu cầu về xây dựng cơ sở dữ liệu về các bản án, quyết định của Tòa án

Từ ngày 01/10/2018 đến ngày 30/9/2019, các Tòa án đã thụ lý 625.979 vụ việc, đã giải quyết được 500.361 vụ việc (đạt tỷ lệ 80%); so với năm 2018, số vụ việc đã thụ lý tăng 69.141 vụ (bằng 12,4%), đã giải quyết tăng 58.808 vụ (bằng 13,3%).

Qua khảo sát khó khăn thực tế là khi tiếp nhận các đơn khởi kiện cụ thể với lĩnh vực hôn nhân gia đình Thẩm phán được phân công xem xét cần nghiên cứu kỹ nội dung đơn và những thông tin nguyên đơn, bị đơn để quyết định xem vụ việc này có thể hòa giải hay khó hòa giải mà phải mở phiên tòa xét xử. Việc quyết định này cần nhanh chóng không ảnh hưởng đến quyền lợi của người dân, tuy nhiên với lượng vụ việc ngày càng lớn thì đây là áp lực không nhỏ.

Với ngành Tòa án có thể hướng tới tương lai phát triển Tòa án thông minh dựa trên việc khai phá kho dữ liệu về các bản án, quyết định của Tòa án hỗ trợ công tác xem xét giải quyết các vụ việc trong đó có vụ việc về hôn nhân gia đình.

Cùng với việc đẩy mạnh cải cách hành chính và hướng tới giải quyết các vụ việc được nhanh chóng, cần xây dựng một cơ sở dữ liệu và phần mềm áp dụng khai thác dữ liệu về các bản án hỗ trợ ra các quyết định nhanh chóng kịp thời.

3.1.2. Hiện trạng dữ liệu về các bản án hôn nhân gia đình

Nội dung về các bản án hiện tại của ngành Tòa án được ghi các sổ theo dõi hoặc một số đơn vị ghi lại trên bảng tính, có nhiều đặc trưng. Tuy nhiên chưa có phần mềm quản lý lưu trữ tập trung cơ sở dữ liệu các bản án, quyết định của Tòa án. Nội dung này được gọi là sổ thụ lý và theo dõi thống kê kết quả giải quyết các vụ việc hôn nhân gia đình.

SỐ THỰ LÝ VÀ KẾT QUẢ GIẢI QUYẾT CÁC VỤ VIỆC HÔN NHÂN GIA ĐÌNH SƠ THẨM									
THỰ LÝ Số, ngày tháng năm	NGUYÊN ĐƠN HOẶC NGƯỜI YÊU CẦU Họ tên, địa chỉ, họ tên người đại diện, chức vụ, địa chỉ	BỊ ĐƠN HOẶC NGƯỜI LIÊN QUAN TRONG VIỆC HÔN NHÂN GIA ĐÌNH Họ tên, địa chỉ, họ tên người đại diện, chức vụ, địa chỉ	ĐƠN KHÔNG KIẾN CỬA CƠ QUAN, TỔ CHỨC	NGƯỜI CÒ QUYỀN LỢI, NGƯỜI BẢO LIÊN QUAN Họ tên, địa chỉ, Họ tên người đại diện, chức vụ, địa chỉ	HỌ TÊN NGƯỜI BẢO VỆ QUYỀN, LỢI ÍCH HỢP PHÁP CỦA ĐƯƠNG SỰ	QUAN HỆ PHÁP LUẬT KHI THỰ LÝ	ÁP DỤNG BIỆN PHÁP KHẨN CẤP TẠM TRÔI Số, ngày tháng năm	CHUYỂN HỒ SƠ VỤ VIỆC Số, ngày tháng năm và nơi chuyển	TAM ĐÌNH CHI Số, ngày tháng năm
1	2	3	4	5	6	7	8	9	10
155 02.10 2017	Nguyễn Hoàng Hậu 1992 Đ/c: T4 - 22 - 06 TimesCity, phường Minh Khai, quận Hai Bà Trung, Hà Nội	Terry John Arthur Moody 1987 Đ/c: 59 AvonGrove, Bletchley, Milton Keynes, MK37BL, UK			TB109/27.9.17 BL09641/29.9 2017 ST: 300.0004 Tp Lai Vinh Trung	Y/c công nhận thuận tình ly hôn			
156 02.10 2017	Nguyễn Thị Bích Hào 1958 Đ/c: Số 43, ngõ 70, phố Nguyễn An Ninh, tổ 20, quận Hoàng Mai, Hà Nội	Nguyễn Văn Hoà 1955 Đ/c: Đường Grusebska 39 - 1, phòng 6, nhà 2, Tp Odessa, CH Ukraina			TB122/27.9.17 BL09641/29.9 2017 ST: 300.0004 Tp Lai Vinh Trung	T/c hôn nhân và gia đình			

Hình 3.1. Thí dụ thông tin về một bản án

3.2. Phân tích bài toán về quản lý án hôn nhân

3.2.1. Thông tin nguyên đơn

Các thông tin cơ bản của bên Nguyên đơn: (i): Họ và tên; (ii): Giới tính: Nam/Nữ; (iii): Năm sinh; (iv): Tuổi: Tuổi của nguyên đơn khi nộp đơn ra Tòa; (v): Địa chỉ; (vi): Con chung.

3.2.2. Thông tin bị đơn

Các thông tin cơ bản liên quan đến bên Bị đơn: (i): Họ và tên; (ii): Giới tính: Nam/Nữ; (iii): Năm sinh; (iv): Tuổi: Tuổi của bị đơn khi nộp đơn ra Tòa; (v): Địa chỉ.

3.2.3. Thông tin quyết định

Với mỗi quyết định của toàn án, sẽ bao gồm các thông tin về ngày thụ lý bản án, thẩm phán phụ trách, kết quả của vụ việc và quan hệ pháp luật của hai phía nguyên đơn và bị đơn.

- Ngày thụ lý bản án: Ngày toàn án tiếp nhận thụ lý bản án từ phía nguyên đơn.
- Thẩm phán: Họ và tên thẩm phán phụ trách vụ việc.
- Kết quả: Kết quả của vụ việc được tòa án tuyên bố (hòa giải hay xét xử).
- Quan hệ pháp luật: Các nguyên nhân dẫn đến việc đệ đơn (i) Mâu thuẫn gia đình; (ii) Bạo lực gia đình; (iii) Yếu tố nước ngoài; (iv) Một người mất tích; (v) Ngoại tình; (vi) Bệnh tật, không có con; (vii) Mâu thuẫn kinh tế; (viii) Nghiện rượu bia, cờ bạc, ma túy.

3.3. Thiết kế cơ sở dữ liệu án hôn nhân gia đình

3.3.2. Thiết kế chi tiết các bảng dữ liệu

- Bảng *NguyenDon* có 7 thuộc tính.

	Column Name	Data Type	Allow Nulls
🔑	MaVuViec	int	<input type="checkbox"/>
	HoTen	nvarchar(100)	<input type="checkbox"/>
	GioiTinh	nvarchar(10)	<input type="checkbox"/>
	NamSinh	int	<input type="checkbox"/>
	Tuoi	int	<input type="checkbox"/>
	DiaChi	nvarchar(100)	<input type="checkbox"/>
▶	Conchung	nvarchar(50)	<input type="checkbox"/>

Hình 3.2. Lược đồ bảng nguyên đơn

- Bảng *BiDon* có 6 thuộc tính.

	Column Name	Data Type	Allow Nulls
🔑	MaVuViec	int	<input type="checkbox"/>
	HoTen	nvarchar(100)	<input type="checkbox"/>
	GioiTinh	nvarchar(10)	<input type="checkbox"/>
	NamSinh	int	<input type="checkbox"/>
	Tuoi	int	<input type="checkbox"/>
	DiaChi	nvarchar(100)	<input type="checkbox"/>

Hình 3.3. Lược đồ bảng bị đơn

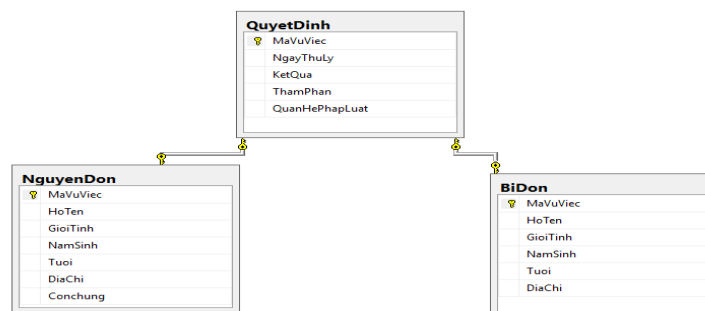
- Bảng *QuyetDinh* có 5 thuộc tính.

	Column Name	Data Type	Allow Nulls
🔑	MaVuViec	int	<input type="checkbox"/>
	NgayThuLy	date	<input type="checkbox"/>
	KetQua	nvarchar(100)	<input type="checkbox"/>
	ThamPhan	nvarchar(50)	<input type="checkbox"/>
	QuanHePhapLuat	nvarchar(100)	<input type="checkbox"/>

Hình 3.4. Lược đồ bảng quyết định

- Quan hệ giữa các bảng dữ liệu

Bảng **QuyếtDinh** có trường MaVuViec là khóa ngoại, tham chiếu đến trường MaVuViec của hai bảng **NguyenDon** và **BiDon** để lấy ra các thông tin của nguyên đơn và bị đơn.



Hình 3.5. Lược đồ cơ sở dữ liệu

3.4. Xây dựng hệ thống trợ giúp quyết định trong môi trường C#

3.4.1. Chức năng Trợ giúp ra quyết định

Chức năng chính của phần mềm là chức năng trợ giúp ra quyết định từ thông tin đầu vào của nguyên đơn và bị đơn được xây dựng trên bộ luật từ phần mềm Weka.

Hình 3.6. Giao diện chính

- Thực hiện chức năng “Trợ giúp ra quyết định”: Khi Thẩm phán nhập đầy đủ thông tin như hình trên và thực hiện “Trợ giúp ra quyết định” hệ thống đưa ra kết quả trợ giúp có kết quả là “Xét xử” như sau:

Hình 3.7. Kết quả trợ giúp ra quyết định

- Thực hiện chức năng “Lưu”: Thẩm phán xem kết quả trợ giúp và lưu kết quả trợ giúp và hệ thống phục vụ công tác lưu trữ cơ sở dữ liệu về các bản án và hỗ trợ tìm kiếm thống kê, giám sát thực hiện,...

Hình 3.8. Lưu kết quả trợ giúp ra quyết định

3.4.2. Chức năng tra cứu bản án, quyết định

- Chức năng này phục vụ việc tra cứu thông tin về các vụ việc, bản án, quyết định của Tòa án theo các thuộc tính thông tin đầu vào

Mã Vụ Việc	Ngày Thu Lý	Kết Quả	Thẩm Phán	Quan Hệ Pháp Luật	Nguyên Đơn	Giới Tính NE
1	19/05/2016	Xét xử	Bùi Thị Thanh Phương	Do nghiện rượu, cờ bạc, ma túy	Phạm Thị Cẩm Dung	Nữ
2	06/05/2016	Xét xử	Bùi Thị Thức	Bạo lực gia đình	Nguyễn Thanh Bình	Nữ
3	16/05/2016	Hòa giải	Chu Minh Sang	Mâu thuẫn gia đình	Đỗ Thị Nhung	Nữ
4	23/05/2016	Xét xử	Diệp Lê Quỳnh Anh	Mâu thuẫn gia đình	Lê Thị Uyên	Nữ
5	16/03/2016	Hòa giải	Danh Đồi	Ngoại tình	Thân Thị Thảo	Nữ
6	10/05/2016	Hòa giải	Dương Hồng Phương	Mâu thuẫn gia đình	Nguyễn Thị Hào	Nữ
7	20/09/2016	Xét xử	Dương Thị Huyền Trang	Do nghiện rượu, cờ bạc, ma túy	Phan Thị Hồng Nhung	Nữ
8	02/03/2016	Xét xử	Giản Thị Dung	Mâu thuẫn gia đình	Đỗ Thị Hương	Nữ
9	15/10/2016	Xét xử	Hoàng Thị Xuân	Mâu thuẫn kinh tế	Nguyễn Quỳnh Anh	Nữ

Hình 3.9. Màn hình tra cứu thông tin bản án, quyết định

3.3. Kết luận

Chương trên đã trình bày hệ thống truy cập dữ liệu và trợ giúp quyết định về án hôn nhân. Hệ thống sử dụng (i) hệ quản trị cơ sở dữ liệu SQL SERVER; (ii) môi trường Visual C#.

Thông tin từ cơ sở dữ liệu là hữu ích đối với việc tra cứu các án hôn nhân, cũng như đưa ra phương án dự thảo án mới, trước tình huống cần trợ giúp, tư vấn.

KẾT LUẬN

Những kết quả đạt được

Luận văn đã trình bày nhu cầu xây dựng và sử dụng hệ thống trợ giúp quyết định trong việc thụ lý, xem xét đưa ra quyết định về vụ, việc trong lĩnh vực hôn nhân và gia đình. Bản án, quyết định về hôn nhân gia đình được đề cập như đối tượng để thực hiện việc ra quyết định với vụ, việc về hôn nhân gia đình dựa trên khai phá dữ liệu về các bản án, quyết định của Tòa án.

Trong quá trình thực hiện và hoàn thiện luận văn với sự hướng dẫn của giáo viên hướng dẫn tôi đã thực hiện, hoàn thiện và đạt được một số kết quả như đã nêu trong mục tiêu của đề cương cụ thể như sau:

Chương đầu đã trình bày và hiểu được nội dung tổng quan, kiến trúc, các chức năng, phân tích lớp và dự đoán, phân cụm khai phá dữ liệu. Khái niệm về các hệ thống ra quyết định, phân loại các hệ thống ra quyết định, cây quyết định trong đó đã thực hiện nghiên cứu, phân tích đánh giá chung các thuật toán phân lớp dữ liệu là thuật toán ID3 và thuật toán C4.5. Qua đó đánh giá được thuật toán C4.5 là thuật toán xử lý đầy đủ các vấn đề phân lớp dữ liệu và phổ biến nhất, đây là cơ sở lựa chọn áp dụng thuật toán này vào việc xây dựng hệ thống hỗ trợ ra quyết định về hòa giải trong các tranh chấp về hôn nhân gia đình.

Chương thứ hai áp dụng được việc thử nghiệm phương pháp phân lớp bằng cây quyết định sử dụng thuật toán J48 trong bộ công cụ phần mềm WEKA, kết quả đưa ra được là cây quyết định và bộ luật phân lớp áp dụng xây dựng phần mềm hỗ trợ ra quyết định.

Chương cuối đã xây dựng cơ sở dữ liệu và phần mềm hỗ trợ ra quyết định mà luận văn đề xuất sử dụng trong công việc của Tòa án. Cơ sở dữ liệu sử dụng hệ quản trị cơ sở dữ liệu SQL SERVER. Việc đưa công nghệ thông tin và truyền thông vào công tác hỗ trợ giải quyết vụ việc trước hết cần hệ thống thông tin; mà cốt lõi của nó là cơ sở dữ liệu về án và cài đặt thuật toán phân lớp dữ liệu. Đối với luận văn, đó là cơ sở dữ liệu về các án hôn nhân gia đình và thuật toán phân lớp dữ liệu dựa trên cây quyết định. Những kết quả này có ý nghĩa khoa học thực tiễn và

có hướng gợi mở cho việc nghiên cứu tiếp theo cho các lĩnh vực tổ tụng khác của Tòa án như hình sự, dân sự, kinh doanh thương mại, lao động, hành chính. Là yếu tố hướng tới xây dựng Tòa án thông minh.

Hướng nghiên cứu phát triển của luận văn

Trong giới hạn của luận văn hệ thống mới xây dựng thử nghiệm trên bộ dữ liệu thực tế còn hạn chế từ sở thụ lý và kết quả giải quyết các vụ việc hôn nhân gia đình sơ thẩm của Tòa án nhân dân thành phố Hà Nội. Do luận văn đề xuất (i) hệ thống cơ sở dữ liệu; (ii) phân loại tự động các bản án hôn nhân gia đình, nên hướng tiếp theo sau luận văn của học viên sẽ là :

- Hoàn thiện và phát triển hệ thống cơ sở dữ liệu đã tạo ra trong quá trình làm luận văn tốt nghiệp;
- Hoàn thiện các chức năng của phần mềm hướng tới áp dụng thực tế;
- Bổ sung dữ liệu quá khứ về các án hôn nhân gia đình, dữ liệu của các đơn vị Tòa án khác trong ngành Tòa án;
- Nghiên cứu thêm các thuật toán khác như C5.0, SPURT;
- Nghiên cứu xây dựng hệ thống mở rộng áp dụng cho các lĩnh vực tổ tụng khác như hình sự, dân sự, hành chính, lao động, kinh doanh thương mại.

Thực tế trong công tác thụ lý giải quyết vụ, việc và xét xử còn nhiều yếu tố và cần phải nghiên cứu đòi hỏi tính chặt chẽ căn cứ theo các luật, bộ luật tổ tụng tại Việt Nam. Hệ thống phần mềm được xây dựng trong luận văn mang tính chất “hỗ trợ” ra quyết định trợ giúp cho công tác giải quyết vụ, việc ngày càng nhiều được nâng cao và hiệu quả hơn trong cải cách tư pháp tại Tòa án nhân dân.