**NGHIÊN CỨU MỘT SỐ GIẢI THUẬT PHÂN LỚP VÀ ỨNG DỤNG VÀO KHAI PHÁ QUAN ĐIỂM TRONG MỘT WEBSITE ÂM NHẠC**

**Nguyễn Văn Giang**

*Khóa QH-2007-I/CQ , ngành Công nghệ Thông Tin*

**Tóm tắt khóa luận tốt nghiệp:**

Thời đại chúng ta đang sống được gọi là kỷ nguyên công nghệ, công nghệ phát triển một cách chóng mặt. Việc ta trao đổi thông tin, thu thập ý kiến hay mua bán trên mạng nói riêng thật sự phổ biến và đơn giản. Số lượng trang web tăng theo cấp số nhân. Ta có thể đưa ra ý kiến cá nhân mình về một chủ đề nào đó trên các trang blog, diễn đàn, mạng xã hội…và khi ta muốn mua gì hay nghe một bài hát nào đó ta có thể tham khảo phần bình luận của những người đã trải qua.Và vấn đề đặt ra là chúng ta có thể khai phá một cách tự động được gì từ những dữ liệu khổng lồ đó.

Khóa luận của tác giả đi vào việc nghiên cứu một số giải thuật phân lớp và ứng dụng vào khai phá quan điểm trong một website âm nhạc, thực nghiệm trên cơ sở dữ liệu comment của các người dùng về một bài hát. Mục tiêu hướng tới của tác giả là xây dựng mô hình phân lớp có độ chính xác cao để tìm ra các câu diễn tả các quan điểm, và có thể phân lớp tự động cho các tập dữ liệu mới để tìm ra các câu diễn tả quan điểm khen, chê, hay trung tính. Thực nghiệm của tác giả sử dụng công cụ Weka với ngôn ngữ phát triển là java. Kết quả đạt được vẫn còn nhiều hạn chế vì vấp phải khó khăn về việc xử lý ngôn ngữ tự nhiên, nhưng cũng đã đưa ra được các điểm mạnh và điểm yếu của các giải thuật phân lớp.

***Từ khóa:*** Phân lớp, Quan điểm