

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

-----



**NGUYỄN TRỌNG THANH**

**XÂY DỰNG ỨNG DỤNG NGĂN CHẶN TIN NHẮN RÁC  
TRÊN THIẾT BỊ DI ĐỘNG**

**CHUYÊN NGÀNH : KHOA HỌC MÁY TÍNH**

**MÃ SỐ: 60.48.01.01**

**LUẬN VĂN THẠC SĨ KỸ THUẬT**

*(Theo định hướng ứng dụng)*

**NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. NGUYỄN TRỌNG ĐƯỜNG**

**HÀ NỘI - 2016**

Luận văn được hoàn thành tại:

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

Người hướng dẫn khoa học: TS. Nguyễn Trọng Đường

Phản biện 1: .....

Phản biện 2: .....

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 cấp thiết của đề tài

Sự phát triển của khoa học, kỹ thuật đã tạo ra nhiều công nghệ mới phục vụ hoạt động liên lạc, trao đổi thông tin, trong đó có SMS. Hiện nay, SMS được đánh giá là một trong những phương tiện liên lạc phổ biến, đơn giản và hiệu quả nhất. SMS cũng được coi là công cụ phục vụ đắc lực trong hoạt động kinh doanh, đặc biệt là quảng cáo.

Cũng như thư điện tử rác, tin nhắn rác là một trong những vấn nạn phát sinh bên cạnh những đặc điểm tích cực mà tin nhắn mang lại. Tại Việt Nam, tin nhắn rác đã và đang trở thành một vấn nạn lớn đối với thuê bao di động và diễn biến ngày một tinh vi và phức tạp. Chưa bao giờ việc phát tán những tin nhắn rác dễ dàng như hiện nay. Tin nhắn rác không chỉ có nội dung về mua bán, quảng cáo mà còn xuất hiện rất nhiều tin nhắn rác lừa đảo qua tin nhắn, dụ dỗ người dùng, thậm chí mang tính chất phản động, bôi nhọ danh dự người khác... Mỗi năm có hàng trăm triệu tin nhắn rác được phát tán đi. Chúng gây khó chịu cả về nội dung lẫn số lượng cho những người sử dụng điện thoại.

Theo định nghĩa về tin nhắn rác tại Luật Công nghệ thông tin thì tin nhắn rác là tin nhắn được gửi đến người nhận mà người nhận đó không mong muốn hoặc không có trách nhiệm phải tiếp nhận theo quy định của pháp luật. Điều này chứng tỏ vai trò rất quan trọng của các ứng dụng, phần mềm ngăn chặn tin nhắn rác dành cho người dùng. Vì vậy, bên cạnh các giải pháp ngăn chặn tin nhắn rác là từ phía doanh nghiệp di động, thì các phần mềm, ứng dụng cho phép người dùng chủ động ngăn chặn tin nhắn rác là rất quan trọng. Nếu như doanh nghiệp di động có thể ngăn chặn những tin nhắn với nội dung, tiêu chí nhất định, thì với các phần mềm, ứng dụng cài đặt trên điện thoại, người dùng có thể chủ động lựa chọn những nội dung tin nhắn mà họ không mong muốn, coi là rác và lựa chọn những nội dung mà họ muốn nhận được.

Hiện nay, điện thoại thông minh ngày càng phổ biến và trở thành một vật dụng không thể thiếu đối với nhiều người. Vì vậy, trong phạm vi luận văn thạc sỹ, chủ trì mong muốn xây dựng một phần mềm ngăn chặn tin nhắn rác trên điện thoại thông minh, đặc biệt là đối với tin nhắn rác tại Việt Nam theo cả hai hướng tự động và chủ động, giúp người dùng loại bỏ tin nhắn rác.

Xuất phát từ thực tế và mục tiêu như vậy, tác giả thực hiện đề tài luận văn có tên **“Xây dựng ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động”** để giải quyết vấn đề nêu trên.

## **Tổng quan về vấn đề nghiên cứu**

Sự phát triển của khoa học, kỹ thuật đã tạo ra nhiều công nghệ mới phục vụ hoạt động liên lạc, trao đổi thông tin, trong đó có SMS. Hiện nay, SMS được đánh giá là một trong những phương tiện liên lạc phổ biến, đơn giản và hiệu quả nhất. SMS cũng được coi là công cụ phục vụ đắc lực trong hoạt động kinh doanh, đặc biệt là quảng cáo.

Cũng như thư điện tử rác, tin nhắn rác là một trong những vấn nạn phát sinh bên cạnh những đặc điểm tích cực mà tin nhắn mang lại. Tại Việt Nam, tin nhắn rác đã và đang trở thành một vấn nạn lớn đối với thuê bao di động và diễn biến ngày một tinh vi và phức tạp. Chưa bao giờ việc phát tán những tin nhắn rác dễ dàng như hiện nay. Tin nhắn rác không chỉ có nội dung về mua bán, quảng cáo mà còn xuất hiện rất nhiều tin nhắn rác lừa đảo qua tin nhắn, dụ dỗ người dùng, thậm chí mang tính chất phản động, bôi nhọ danh dự người khác... Mỗi năm có hàng trăm triệu tin nhắn rác được phát tán đi. Chúng gây khó chịu cả về nội dung lẫn số lượng cho những người sử dụng điện thoại.

Theo định nghĩa về tin nhắn rác tại Luật Công nghệ thông tin thì tin nhắn rác là tin nhắn được gửi đến người nhận mà người nhận đó không mong muốn hoặc không có trách nhiệm phải tiếp nhận theo quy định của pháp luật. Điều này chứng tỏ vai trò rất quan trọng của các ứng dụng, phần mềm ngăn chặn tin nhắn rác dành cho người dùng. Vì vậy, bên cạnh các giải pháp ngăn chặn tin nhắn rác là từ phía doanh nghiệp di động, thì các phần mềm, ứng dụng cho phép người dùng chủ động ngăn chặn tin nhắn rác là rất quan trọng. Nếu như doanh nghiệp di động có thể ngăn chặn những tin nhắn với nội dung, tiêu chí nhất định, thì với các phần mềm, ứng dụng cài đặt trên điện thoại, người dùng có thể chủ động lựa chọn những nội dung tin nhắn mà họ không mong muốn, coi là rác và lựa chọn những nội dung mà họ muốn nhận được.

Hiện nay, điện thoại thông minh ngày càng phổ biến và trở thành một vật dụng không thể thiếu đối với nhiều người. Vì vậy, trong phạm vi luận văn thạc sỹ, chủ trì mong muốn xây dựng một phần mềm ngăn chặn tin nhắn rác trên điện thoại thông minh, đặc biệt là đối với tin nhắn rác tại Việt Nam theo cả hai hướng tự động và chủ động, giúp người dùng loại bỏ tin nhắn rác.

Xuất phát từ thực tế và mục tiêu như vậy, tác giả thực hiện đề tài luận văn có tên “Xây dựng ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động” để giải quyết vấn đề nêu trên.

## Mục đích, đối tượng, phạm vi và phương pháp nghiên cứu

Luận văn tập trung nghiên cứu, tìm hiểu các vấn đề cơ bản về tin nhắn rác tại Việt Nam và trên thế giới, các thuật toán học máy. Từ đó ứng dụng vào việc xây dựng ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động, cụ thể là trên hệ điều hành Android tại Việt Nam. Kết quả đạt được của luận văn là có những nghiên cứu lý thuyết, số liệu, tình hình về tin nhắn rác, các thuật toán học máy; xây dựng được ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động tại Việt Nam.

Luận văn nghiên cứu về các loại hình tin nhắn rác tại Việt Nam, từ đó, xây dựng nên ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động, cụ thể là thiết bị di động sử dụng hệ điều hành Android.

Luận văn kết hợp 2 phương pháp nghiên cứu. Phương pháp nghiên cứu lý thuyết: tổng hợp, thu thập, nghiên cứu tài liệu, ngôn ngữ lập trình, các công nghệ, giải pháp có liên quan đến ngăn chặn tin nhắn rác. Phương pháp nghiên cứu thực nghiệm: Phân tích được các yêu cầu của công việc, vận dụng các kết quả lý thuyết vào bộ dữ liệu cụ thể (dữ liệu tin nhắn rác) thu thập được để đánh giá và phân tích kết quả. Tập hợp, xây dựng được bộ dữ liệu mẫu để kiểm tra thử nghiệm, nhận xét, đánh giá được kết quả đạt được.

## Cấu trúc luận văn

Nội dung của luận văn được trình bày trong ba phần chính như sau:

### 1. Phần mở đầu

### 2. Phần nội dung: bao gồm ba chương

#### *Chương 1: Các giải pháp ngăn chặn tin nhắn rác*

Chương này đi sâu vào các giải pháp ngăn chặn tin nhắn rác theo các hướng, các giải pháp khác nhau. Giải pháp về mặt pháp lý: đánh giá các giải pháp hiện nay về mặt pháp lý ngăn chặn tin nhắn rác đến từ các cơ quan chức năng, các nhà mạng. Giải pháp về mặt kỹ thuật: các mô hình, các thuật toán ngăn chặn tin nhắn rác áp dụng vào bài toán phân loại và xử lý tin nhắn rác.

#### *Chương 2: Áp dụng thuật toán Naïve Bayes xây dựng bộ lọc tin nhắn rác*

Chương này trình bày cách tạo bộ lọc tin nhắn rác sử dụng thuật toán Naïve Bayes. Trong chương 2, luận văn sẽ trình bày mô hình thuật toán Naïve Bayes, các ưu nhược điểm của thuật toán, các bước xây dựng bộ lọc tin nhắn rác sử dụng thuật toán Naïve Bayes như lấy dữ liệu mẫu, huấn luyện bộ lọc.

***Chương 3: Xây dựng ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động***

Trong chương 3, luận văn sẽ đánh giá các tập dữ liệu, các kỹ thuật được dùng trong chương trình, xây dựng ứng dụng ngăn chặn tin nhắn rác trên hệ điều hành Android bằng thuật toán Naïve Bayes đã đưa ra ở chương 2, demo thực nghiệm và đánh giá thuật toán.

**3. Phần kết luận**

# CHƯƠNG 1 - CÁC GIẢI PHÁP NGĂN CHẶN TIN NHẮN RÁC

## 1.1 Tổng quan về tin nhắn rác

### 1.1.1 Tin nhắn SMS

Tin nhắn SMS (Short Messaging Service hay Simple Message Service) là dịch vụ nhắn tin ngắn SMS cho phép gửi và nhận bản tin ngắn (tin nhắn) giữa các máy điện thoại di động cũng như giữa điện thoại di động và các thiết bị cung cấp thông tin khác (PC, PDA).

### 1.1.2 Tin nhắn SMS rác

Giống như thư điện tử rác, tin nhắn SMS rác là tin nhắn được gửi đến người nhận mà người nhận đó không mong muốn hoặc không có trách nhiệm phải tiếp nhận theo quy định của pháp luật.

## 1.2 Thực trạng tin nhắn rác tại Việt Nam

### 1.2.1 Thực trạng tin nhắn rác tại Việt Nam

Với sự bùng nổ của viễn thông công nghệ thông tin, con người có thể kết nối với nhau qua tích tắc dù cách xa nhau ngàn vạn cây số cũng như nắm bắt thông tin nhanh nhất để trở thành người tiêu dùng thông minh, dễ dàng tiếp cận với sản phẩm, dịch vụ của mình mong muốn. Nắm bắt được xu thế này, những năm gần đây, nhiều người đã sử dụng mạng viễn thông như một phương thức quảng cáo rẻ và hiệu quả trong việc đưa sản phẩm tới người tiêu dùng. Và trong bối cảnh ấy, tin rác ra đời tạo nên những bức xúc cho người dùng viễn thông di động. Ở lúc cao điểm, mỗi một sim di động trong 24 giờ có thể nhận tới cả chục tin nhắn rác quảng cáo dịch vụ bất động sản, sim số đẹp, chăn ga gối đệm, du lịch... Ngoài việc trở thành một phương thức quảng cáo giá rẻ ra, thì các tin nhắn rác còn có dấu hiệu lừa đảo như nội dung tin nhắn thông báo trúng thưởng, tặng nhạc chuông, hình ảnh, cài đặt GPRS nhưng khi chủ thuê bao di động nhấn tin đến đầu số theo hướng dẫn thì lập tức bị trừ tiền.

### 1.2.2 Các loại tin nhắn rác chính xuất hiện tại Việt Nam

Ở Việt Nam tin nhắn rác có thể phân chia thành một số loại như sau:

- Tin nhắn quảng cáo không hợp pháp
- Tin nhắn có tính chất dụ dỗ khách hàng
- Tin nhắn có tính chất lừa đảo
- Tin nhắn giả mạo

### **1.3 Các giải pháp ngăn chặn tin nhắn rác**

#### ***1.3.1 Giải pháp về mặt pháp lý***

Đây là các giải pháp được đưa ra bởi các cơ quan quản lý và các nhà mạng nhằm ngăn chặn sự phát tán tin nhắn rác. Tại Việt Nam, chính phủ đã ban hành Luật chống tin nhắn, thư rác theo Nghị định về chống thư rác (90/2008/NĐ-CP). Cục An toàn thông tin (Bộ Thông tin và Truyền thông) đề xuất giới hạn số tin nhắn gửi trong một khoảng thời gian nhất định. Đồng thời xử lý nghiêm minh theo Nghị định 174/2013/NĐ-CP của Chính phủ quy định về xử phạt vi phạm hành chính trong lĩnh vực viễn thông.

Các nhà mạng, báo chí cũng góp phần vào việc ngăn chặn phát tán tin nhắn rác ở Việt Nam.

#### ***1.3.2 Giải pháp về mặt kỹ thuật***

Sử dụng các giải pháp về mặt kỹ thuật giúp người dùng chủ động hơn trong việc ngăn chặn tin nhắn rác. Sử dụng các phương pháp như Blacklist, Graylist, sử dụng các Keyword thường xuất hiện trong các tin nhắn rác hay sử dụng các thuật toán học máy, giúp phân loại tin nhắn rác và tin nhắn hợp lệ.

### **1.4 Các phương pháp phân loại văn bản**

#### ***1.4.1 Phân loại văn bản***

Phân loại văn bản là sự phân loại không cấu trúc các tài liệu văn bản dựa trên một tập hợp của một hay nhiều loại văn bản đã được định nghĩa trước. Quá trình này thường được thực thi bằng một hệ thống tự động gán cho các tài liệu văn bản một loại nào đó.

#### ***1.4.2 Tiến trình phân loại văn bản***

Đưa ra một tập tài liệu mẫu D, cần được phân bổ thành một số loại tài liệu nhất định - mỗi tài liệu đó cần được gán cho một loại văn bản nào đó. Nhiệm vụ của chúng ta là tìm một hệ thống phân hoạch, mà nó sẽ cung cấp cho ta một nhãn y phù hợp cho một số tài liệu trong D vừa được đưa vào từ nguồn tài nguyên giống nhau như các văn bản mẫu.

Các bước trong tiến trình phân loại văn bản:

- Lựa chọn các đặc trưng văn bản
- Biểu diễn văn bản
- Học một bộ phân loại văn bản
- Tiến hành phân loại văn bản



### 1.4.3 Các phương pháp phân loại văn bản

#### 1.4.3.1 Nguyên mẫu

Nguyên mẫu (prototype) có thể là phương pháp đơn giản nhất được áp dụng trong phân loại văn bản. Ưu điểm của nguyên mẫu là có thể áp dụng cho mọi văn bản, tuy nhiên xác suất của phương pháp này không cao (lí do là có nhiều loại văn bản độ dài giống nhau nhưng nội dung khác hẳn nhau).

#### 1.4.3.2 Mô hình xác suất Naïve Bayes

Kĩ thuật phân hoạch của Naive Bayes dựa trên cơ sở định lí Bayes và đặc biệt phù hợp cho các trường hợp phân loại có kích thước đầu vào là lớn. Mặc dù Naive Bayes khá đơn giản nhưng nó có khả năng phân loại tốt hơn rất nhiều phương pháp phân hoạch phức tạp khác. Với mỗi loại văn bản, thuật toán Naive Bayes tính cho mỗi lớp văn bản một xác suất mà tài liệu cần phân hoạch có thể thuộc loại đó. Tài liệu đó sẽ được gán cho lớp văn bản nào có xác suất cao nhất.

#### 1.4.3.3 Phương pháp SVM (Support Vectors Machines)

SVM là một phương pháp tiếp cận gần đúng thường áp dụng để phân loại hai lớp văn bản. Phương pháp này được xác định dựa trên một không gian vector mà trong không gian này vấn đề phải tìm ra một “mặt quyết định” giữa hai lớp sao cho khoảng cách giữa các điểm dữ liệu giữa hai lớp này là lớn nhất. Nếu văn bản cần phân loại nằm về phía nào của mặt quyết định thì nó được phân vào loại văn bản đó. Phương pháp SVM thường áp dụng để phân loại hai lớp văn bản nhưng vẫn có thể áp dụng để phân loại nhiều lớp văn bản

#### Phương pháp cây quyết định (Decision Trees)

Cây quyết định là một trong những kĩ thuật học máy được nhiều người biết đến. Chúng được áp dụng rộng rãi và đa dạng của các lĩnh vực đặc biệt là các ứng dụng trong trí tuệ nhân tạo.

#### 1.4.3.4 Phương pháp mạng neuron (Neuron network)

Phân loại văn bản bằng mô hình mạng neuron là một cách thức phân loại mới được đề xuất. Cấu trúc và hoạt động của mạng neuron cơ bản dựa trên bộ não con người. Mạng neuron được áp dụng trong khá nhiều lĩnh vực khác nhau và có khả năng giải quyết các vấn đề phức tạp, phạm vi rộng.

## 1.5 Kết luận chương

Chương 1 đã đưa ra cái nhìn tổng quát về tin nhắn rác, khái niệm tin nhắn rác và thực trạng tin nhắn rác tại Việt Nam, đồng thời đưa ra được các phương pháp ngăn chặn tin nhắn rác. Để ngăn chặn tin nhắn rác, tác giả sử dụng thuật toán học máy Naïve Bayes để phân loại tin nhắn xem đâu là tin nhắn thường, đâu là tin nhắn rác. Chi tiết về thuật toán Naïve Bayes được đề cập tại Chương 2 của luận văn.

## CHƯƠNG 2 - ỨNG DỤNG THUẬT TOÁN NAÏVE BAYES XÂY DỰNG BỘ LỌC TIN NHẮN RÁC

### 2.1 Cơ sở lý thuyết

#### 2.1.1 Công thức xác suất có điều kiện

Xác suất điều kiện của biến cố A với điều kiện biến cố B đã xảy ra là một số không âm, ký hiệu là  $P(A|B)$  nó biểu thị khả năng xảy ra biến cố A trong tình huống biến cố B đã xảy ra.

$$P(A|B) = \frac{P(AB)}{P(B)}$$

- $P(A)$ : Xác suất của sự kiện A xảy ra.
- $P(B)$ : Xác suất của sự kiện B xảy ra.
- $P(B|A)$ : Xác suất (có điều kiện) của sự kiện B xảy ra, nếu biết rằng sự kiện A đã xảy ra.
- $P(A|B)$ : Xác suất (có điều kiện) của sự kiện A xảy ra, nếu biết rằng sự kiện B đã xảy ra.

#### 2.1.2 Công thức xác suất đầy đủ

Giả sử  $B_1, B_2, \dots, B_n$  là nhóm đầy đủ các biến cố. Xét biến cố A sao cho A xảy ra chỉ khi một trong các biến cố  $B_1, B_2, \dots, B_n$  xảy ra. Khi đó :

$$p(A) = \sum_{i=1}^n p(A|B_i) * p(B_i)$$

#### 2.1.3 Phát biểu thuật toán Naïve Bayes

Thuật toán Naïve Bayes dựa trên định lý Bayes được phát biểu như sau :

$$P(Y|X) = \frac{P(XY)}{P(X)} = \frac{P(X|Y)P(Y)}{P(X)}$$

Trong đó:

- Y đại diện một giả thuyết, giả thuyết này được suy luận khi có được chứng cứ mới X.
- $P(X)$  : xác suất X xảy ra.
- $P(Y)$  : xác suất Y xảy ra.
- $P(X|Y)$  : xác suất X xảy ra khi Y xảy ra (xác suất có điều kiện, khả năng của X khi Y đúng)

-  $P(Y|X)$  : xác suất  $Y$  xảy ra khi  $X$  xảy ra (xác suất có điều kiện, khả năng của  $Y$  khi  $X$  đúng)

## 2.2 Bài toán phân loại văn bản với Naïve Bayes

Naive Bayes classifier là một thuật ngữ trong xử lý số liệu thống kê Bayesian với một phân lớp xác suất dựa trên các ứng dụng định lý Bayes với giả định độc lập bền vững. Một thuật ngữ mô tả chi tiết cho những mô hình xác suất sẽ là “mô hình đặc trưng không phụ thuộc”.

Phân loại Bayes đơn giản (Naïve Bayes) - một phương pháp phân lớp dựa vào xác suất điển hình nhất trong khai thác dữ liệu và tri thức, được sử dụng rộng rãi trong lĩnh vực máy học và nhiều lĩnh vực khác như trong các công cụ tìm kiếm, phân loại văn bản, lọc tin nhắn rác. Phân loại Bayes đơn giản là trường hợp riêng của kỹ thuật học máy Bayes, trong đó các giả thiết về độc lập xác suất được sử dụng để đơn giản hóa việc tính xác suất.

Phân loại Bayes đơn giản sử dụng trong trường hợp mỗi ví dụ được cho bằng tập các thuộc tính  $\langle x_1, x_2, \dots, x_n \rangle$  và cần xác định nhãn phân loại  $y$ ,  $y$  có thể nhận giá trị từ một tập nhãn hữu hạn  $C$ .

Trong giai đoạn huấn luyện, dữ liệu huấn luyện được cung cấp dưới dạng các mẫu  $\langle x_i, y_i \rangle$ . Sau khi huấn luyện xong, bộ phân loại cần dự đoán nhãn cho mẫu mới  $x$ .

Theo lý thuyết học Bayes, nhãn phân loại được xác định bằng cách tính xác suất điều kiện của nhãn khi quan sát thấy tổ hợp giá trị thuộc tính  $\langle x_1, x_2, \dots, x_n \rangle$ . Thuộc tính được chọn, ký hiệu  $c_{MAP}$  là thuộc tính có xác suất điều kiện cao nhất (MAP là viết tắt của maximum a posterior), tức là:

$$y = c_{MAP} = \operatorname{argmax}_{c_j \in C} P(c_j | x_1, x_2, \dots, x_n)$$

Sử dụng quy tắc Bayes, biểu thức trên được viết lại như sau:

$$c_{MAP} = \operatorname{argmax}_{c_j \in C} \frac{P(x_1, x_2, \dots, x_n | c_j) P(c_j)}{P(x_1, x_2, \dots, x_n)} = \operatorname{argmax}_{c_j \in C} P(x_1, x_2, \dots, x_n | c_j) P(c_j)$$

Hai thành phần trong biểu thức trên được tính từ dữ liệu huấn luyện. Giá trị  $P(c_j)$  được tính bằng tần suất quan sát thấy nhãn  $c_j$  trên tập huấn luyện, tức là bằng số mẫu có nhãn là  $c_j$  chia cho tổng số mẫu. Việc tính  $P(x_1, x_2, \dots, x_n | c_j)$  khó khăn hơn nhiều. Vấn đề là số tổ hợp giá trị của  $n$  thuộc tính cùng với nhãn phân loại là rất lớn khi  $n$  lớn. Để tính xác suất này được chính xác, mỗi tổ hợp giá trị thuộc tính phải xuất hiện cùng nhãn phân loại đủ nhiều, trong khi số mẫu huấn luyện thường không đủ lớn.

Với giả thiết về tính độc lập xác suất có điều kiện, có thể viết:

$$P(x_1, x_2, \dots, x_n | c_j) = P(x_1 | c_j) P(x_2 | c_j) \dots P(x_n | c_j)$$

tức là xác suất đồng thời quan sát thấy các thuộc tính bằng tích xác suất điều kiện của từng thuộc tính riêng lẻ. Thay vào biểu thức ở trên, ta được bộ phân loại Bayes đơn giản (có đầu ra ký hiệu là  $c_{NB}$ ) như sau:

$$c_{NB} = \arg \max_{c_j \in C} P(c_j) \prod_i P(x_i | c_j)$$

Trong đó,  $P(x_i | c_j)$  được tính từ dữ liệu huấn luyện bằng số lần xi xuất hiện cùng với  $c_j$  chia cho số lần  $c_j$  xuất hiện. Việc tính xác suất này đòi hỏi ít dữ liệu hơn nhiều so với tính  $P(x_1, x_2, \dots, x_n | c_j)$ .

Quá trình học Bayes đơn giản là quá trình tính các xác suất  $P(c_j)$  và các xác suất điều kiện  $P(x_i | c_j)$  bằng cách đếm trên tập dữ liệu.

Ta có thể tóm tắt lại cách phân loại này như sau:

Dữ kiện cần có:

X: tập các thuộc tính  $\langle x_1, x_2, \dots, x_n \rangle$

Y: tập hữu hạn các nhãn  $\langle c_1, c_2, \dots, c_m \rangle$

Các thuộc tính  $\langle x_1, x_2, \dots, x_n \rangle$  độc lập xác suất đôi một với nhau

Theo định lý Bayes:

$$P(C_i | X) = \frac{P(X | C_i) P(C_i)}{P(X)}$$

Theo tính chất độc lập điều kiện:

$$P(X | C_i) = \prod_{k=1}^n P(x_k | C_i) = P(x_1 | C_i) * P(x_2 | C_i) * \dots * P(x_n | C_i)$$

Trong đó:

- $P(C_j | X)$  là xác suất thuộc phân lớp  $i$  khi biết trước mẫu  $X$ .
- $P(C_j)$  xác suất là phân lớp  $i$ .
- $P(x_k | C_j)$  xác suất thuộc tính thứ  $k$  mang giá trị  $x_k$  khi đã biết  $X$  thuộc phân lớp  $i$ .

Khi đó: Luật phân lớp cho các tài liệu mới  $X^{new} = \{x_1, x_2, \dots, x_n\}$  là:

$$\max_{C_i \in C} (P(C_i) \prod_{k=1}^n P(x_k | C_i))$$

Trong đó:

- $P(C_i)$ : được tính dựa trên tần suất xuất hiện tài liệu trong tập huấn luyện
- $P(x_k | C_i)$  được tính từ những tập thuộc tính đã được tính trong quá trình huấn luyện

## 2.3 Các ưu điểm của bộ lọc tin nhắn rác Naïve Bayes

Phương pháp Bayes nhận dạng một tin nhắn dựa vào các mô tả, nó nhận ra các từ khóa mà những từ đó có thể nhận dạng một tin nhắn rác (spam) và cũng nhận ra các từ chứng tỏ một tin nhắn hợp lệ (ham). Bộ lọc Bayes liên tục tự thích nghi bởi sự học tập từ các tin nhắn mới và các tin nhắn hợp lệ đến từ bên ngoài. Bộ lọc Bayes giải quyết và thích nghi với các công nghệ lọc tin nhắn rác kiểu mới.

Bộ lọc tin nhắn sử dụng thuật toán Naive Bayes cung cấp một chức năng lọc tin nhắn tự động. Trên cơ sở sử dụng các xác suất gần đúng để tính toán các khả năng một tin nhắn có thể là tin nhắn rác hay không. Sự tính toán này là quá trình tìm kiếm các từ thường xuất hiện trong các tin nhắn và so sánh chúng với tập mẫu. Thuật toán bắt đầu bằng việc học các nội dung của các tin nhắn hợp lệ và nội dung của những tin nhắn rác. Để sau đó khi nhận vào một tin nhắn mới, các thông tin có sẵn từ tập mẫu, các tiến trình tiền xử lý trước sẽ được áp dụng trên cơ sở nội dung của các tin nhắn.

## 2.4 Các bước xây dựng bộ lọc Naïve Bayes

Xác định rõ các đặc trưng sử dụng. Yêu cầu này chúng ta sẽ xem xét các tin nhắn và tìm các “từ” hoặc “nhóm từ” mà chúng là dấu hiệu của tin nhắn rác hay không phải tin nhắn rác, đây có thể coi là cơ sở dữ liệu cho bộ lọc. Đây là một phần quan trọng trong nhiệm vụ này và chúng ta có thể lặp lại một vài lần.

### 2.4.1 Lựa chọn các đặc trưng

Xác định rõ các đặc trưng sử dụng. Yêu cầu này chúng ta sẽ xem xét các tin nhắn và tìm các “từ” hoặc “nhóm từ” mà chúng là dấu hiệu của tin nhắn rác hay không phải tin nhắn rác, đây có thể coi là cơ sở dữ liệu cho bộ lọc. Đây là một phần quan trọng trong nhiệm vụ này và chúng ta có thể lặp lại một vài lần.

### 2.4.2 Biểu diễn các tin nhắn

Sau khi chọn được các đặc trưng văn bản chúng ta tiến hành biểu diễn các tin nhắn thành các vector mà không gian của nó là tập các từ hay cụm từ đã phân tích. Các cụm từ này đã xác định rõ số lần xuất hiện của chúng trong mỗi tin nhắn. Kết hợp các tin nhắn trong mỗi lớp tin nhắn ta được một vector biểu diễn lớp tin nhắn đó.

### 2.4.3 Xác định ngưỡng

Xác định rõ ngưỡng để loại bỏ tất cả các tin nhắn mà xác suất của chúng lớn hơn xác suất này.

#### ***2.4.4 Thử nghiệm hệ thống lọc tin nhắn rác hiệu quả trong thực tế***

### **2.5 Kết luận chương**

Chương 2 đã đưa ra cơ sở lý thuyết của thuật toán Naïve Bayes, những ưu điểm khi sử dụng thuật toán Naïve Bayes để phân loại văn bản. Đồng thời, chương 2 cũng nêu ra các thức để tạo một bộ lọc sử dụng thuật toán Naïve Bayes, đây là tiền đề để xây dựng nên ứng dụng ngăn chặn tin nhắn rác trên thiết bị di động mà luận văn đề cập ở chương 3.

## **CHƯƠNG 3 - XÂY DỰNG ỨNG DỤNG NGĂN CHẶN TIN NHẮN RÁC TRÊN THIẾT BỊ DI ĐỘNG**

### **3.1 Giới thiệu bài toán**

Các tính năng chính của ứng dụng như gửi, nhận tin nhắn như ứng dụng tin nhắn mặc định của máy; theo dõi các tin nhắn theo cuộc hội thoại; chặn tin nhắn rác theo nhiều phương pháp khác nhau.

### **3.2 Phân tích và xây dựng ứng dụng**

Bản chất của việc xây dựng bộ lọc tin nhắn rác chính là sử dụng thuật toán Naïve Bayes phân loại nội dung tin nhắn. Các tin nhắn sẽ được chia ra thành 2 loại: Tin nhắn spam (SPAM) và tin nhắn hợp lệ (HAM). Trong khuôn khổ của luận văn, dữ liệu huấn luyện với khoảng 200 bản ghi bao gồm các tin nhắn rác và tin nhắn hợp lệ đã được gán nhãn. Ngoài tập dữ liệu huấn luyện, ứng dụng còn có một tập dữ liệu test để kiểm tra đánh giá độ chính xác của thuật toán Naïve Bayes với khoảng 100 bản ghi.

Quá trình xử lý tin nhắn rác được xử lý ngầm. Khi có tin nhắn mới đến, nội dung tin nhắn sẽ được chạy qua bộ lọc tin nhắn rác. Nếu như bộ lọc phân loại tin nhắn đó là tin nhắn hợp lệ, ứng dụng sẽ thông báo cho người dùng biết có tin nhắn mới đến. Nếu tin nhắn đó được phân loại là tin nhắn rác, ứng dụng sẽ không thông báo đến người dùng.

### **3.3 Giao diện ứng dụng**

#### **3.3.1 *Giao diện danh sách tin nhắn***

Màn hình danh sách tin nhắn: Hiển thị toàn bộ các tin nhắn có trong máy. Các tin nhắn được đưa ra dưới dạng các cuộc hội thoại theo từng số điện thoại. Tin nhắn mới sẽ được đưa lên đầu để người sử dụng dễ dàng xem nội dung.

#### **3.3.2 *Giao diện chi tiết hội thoại***

Tại màn hình này, các tin nhắn trong cuộc hội thoại sẽ được hiển thị theo thời gian từ cũ đến mới, các tin nhắn mới đến sẽ được hiển thị ở dưới, tạo cho người dùng cảm giác như đang đối thoại.

Khi nhấn và giữ vào một tin nhắn bất kỳ sẽ có các tùy chọn



### 3.3.3 Giao diện danh sách Blacklist

Với những số điện thoại đã ở trong danh sách Blacklist, các số này sẽ được hiển thị trong màn hình danh sách Blacklist. Khi người dùng nhấn vào một số điện thoại, ứng dụng chuyển sang màn hình chi tiết cuộc hội thoại như ở màn hình danh sách tin nhắn.

### 3.3.4 Màn hình danh sách từ xấu

Màn hình danh sách từ xấu hiển thị những từ xấu mà người dùng định nghĩa. Các từ xấu này khi xuất hiện trong nội dung tin nhắn ngay lập tức sẽ được coi đó là tin nhắn rác. Người dùng có thể thêm, sửa, xóa các từ xấu này, mỗi từ cách nhau bởi dấu “;” (chấm phẩy)

## 3.4 Thực nghiệm và đánh giá thuật toán

Khởi chạy ứng dụng ngăn chặn tin nhắn rác, bộ lọc tin nhắn sẽ được khởi động theo. Toàn bộ quá trình nạp dữ liệu và huấn luyện dữ liệu được xử lý ngầm nhằm nâng cao hiệu quả của ứng dụng cũng như giúp trải nghiệm người dùng được tốt hơn.

Khi có tin nhắn đến, ứng dụng sẽ kiểm tra xem tin nhắn có phải là tin nhắn rác hay không. Trước tiên, tin nhắn đến sẽ được kiểm tra xem số điện thoại gửi đến có nằm trong danh sách Blacklist hay không. Nếu nằm trong danh sách này thì xác định đó là tin nhắn rác luôn. Nếu không nằm trong danh sách Blacklist, hệ thống sẽ kiểm tra nội dung tin nhắn có chứa các từ khóa của tin nhắn rác do người dùng tự định nghĩa hay không. Nếu có chứa một từ khóa trở lên thì xác định đó là tin nhắn rác. Sau cùng, nếu không chứa từ khóa tin nhắn rác, hệ thống sẽ kiểm tra nội dung tin nhắn bằng bộ lọc đã được khởi tạo ở trên để phân loại tin nhắn xem có phải là tin nhắn rác hay không.

Kết quả kiểm nghiệm cho thấy hệ thống có thể nhận diện đúng 94%, 6% nhận diện sai (nhận diện tin nhắn hợp lệ thành tin nhắn rác hoặc tin nhắn rác thành tin nhắn hợp lệ).

## 3.5 Kết luận chương

Chương 3 đã trình bày về các tính năng chính của ứng dụng ngăn chặn tin nhắn rác trên hệ điều hành Android, các kỹ thuật ngăn chặn tin nhắn rác được sử dụng trong ứng dụng. Thuật toán Naïve Bayes có độ chính xác khá cao khi phát hiện các tin nhắn là tin nhắn rác.

## KẾT LUẬN

### 1. Kết quả đạt được

Luận văn đã đưa ra được cái nhìn tổng quan về tin nhắn và tin nhắn rác, thực trạng tình hình tin nhắn rác tại Việt Nam. Các giải pháp về mặt pháp lý và giải pháp về mặt kỹ thuật góp phần không nhỏ vào việc hạn chế sự phát tán của tin nhắn rác. Từ những giải pháp kỹ thuật, luận văn đã chỉ ra rằng sử dụng các thuật toán phân loại văn bản là cách tốt nhất để xem tin nhắn nào là tin nhắn rác. Và với những ưu điểm của thuật toán Naïve Bayes, luận văn đã lựa chọn thuật toán trên làm cơ sở để phát triển ứng dụng ngăn chặn tin nhắn rác trên điện thoại di động sử dụng hệ điều hành Android. Ứng dụng ngăn chặn tin nhắn rác phát triển với đầy đủ các tính năng cơ bản của một ứng dụng tin nhắn trên điện thoại di động như gửi, nhận tin nhắn, quản lý tin nhắn. Ngoài bộ lọc tin nhắn rác sử dụng thuật toán Naïve Bayes, ứng dụng ngăn chặn tin nhắn rác còn có thêm các cách ngăn chặn tin nhắn rác như thêm các số điện thoại vào danh sách Blacklist, các từ khóa xuất hiện trong tin nhắn rác. Khi đã xác định được tin nhắn rác, tin nhắn đến sẽ không thông báo cho người dùng, tránh ảnh hưởng đến người dùng. Sau đó người dùng vẫn có thể xem lại nội dung tin nhắn rác như thường. Bên cạnh đó, ứng dụng còn có chức năng gửi nội dung tin nhắn rác đến đầu số 456, đây là tổng đài tiếp nhận và xử lý tin nhắn rác của Bộ Truyền thông Thông tin. Với tỉ lệ phát hiện tin nhắn rác chính xác là trên 90%, ứng dụng khá hữu ích đối với những ai thường xuyên bị tin nhắn rác làm phiền.

### 2. Hạn chế

Do thời gian thực hiện luận văn không nhiều và hạn chế về mặt kiến thức, công nghệ nên luận văn vẫn còn một hạn chế.

Ứng dụng chưa xử lý được với ngôn ngữ Tiếng Việt. Hiện tại các dữ liệu đầu vào, tập dữ liệu huấn luyện và tập dữ liệu test đều được tiền xử lý loại bỏ dấu. Điều này dẫn đến việc tách các từ bị sai lệch đối với các từ ghép, các cụm từ trong tiếng Việt. Đây là một vấn đề khó khăn do bộ nhớ của các thiết bị di động có hạn, trong khi muốn tách từ tiếng Việt đòi hỏi phải có một bộ từ điển tiếng Việt với đầy đủ các từ đơn và từ ghép. Từ đó mới có thể nâng cao được tỉ lệ chính xác của thuật toán Naïve Bayes.

Số lượng bản ghi trong bộ dữ liệu huấn luyện và dữ liệu test ít và mang tính cá nhân. Hiện tại giao diện và các tùy chọn của ứng dụng sơ sài, chưa có những tính năng mới hấp dẫn.

### **3. Hướng phát triển**

Trong tương lai, ứng dụng ngăn chặn tin nhắn rác sẽ hỗ trợ tách từ tiếng Việt nhằm tăng tỉ lệ phát hiện tin nhắn rác. Tập dữ liệu huấn luyện sẽ phong phú và đa dạng hơn cả về số lượng lẫn chủ đề. Ứng dụng ngăn chặn tin nhắn rác có thể phát triển thêm cả ở các nền tảng di động khác như iOS hay Windowphone để phù hợp với nhiều đối tượng sử dụng. Việc phát hiện tin nhắn rác sẽ xử lý trên server, giúp giảm tải hiệu năng cho ứng dụng chạy trên điện thoại di động.

## TÀI LIỆU THAM KHẢO

### Tài liệu Tiếng Việt

- [1] Chính phủ (2008). *Nghị định 90/2008/NĐ-CP Về chống thư rác*
- [2] Chính phủ (2013). *Nghị định 174/2013/NĐ-CP Quy định xử phạt vi phạm hành chính trong lĩnh vực bưu chính, viễn thông, công nghệ thông tin và tần số vô tuyến điện*

### Tài liệu Tiếng Anh

- [3] GIF Software (2011), *Why Bayes filtering is the most effective anti-spam technology*
- [4] Arnulf B.A. Graf, Olivier Bousquet, Gunnar Ratsch, Bernhard Scholkopf (2008), *Prototype Classification: Insights from Machine Learning*, Max Planck Institute for Biological Cybernetics, Tübingen, Germany
- [5] Juniper Networks (2012), *2011 Mobile Threats Report*
- [6] S. M. Kamruzzaman, Ahmed Ryadh Hasan (2010), *Pattern Classification using Simplified Neural Networks with Pruning Algorithm*
- [7] Wei-Yin Loh (2010), *Classification and regression trees*
- [8] Yiming Yang, Xin Liu (1999), *A re- examination of text categorization method*

### Tài liệu Web

- [9] <https://wikipedia.org/wiki/SMS>, truy cập ngày 25 tháng 5 năm 2016