

**TRƯỜNG ĐẠI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG HÀ NỘI**  
**KHOA CÔNG NGHỆ THÔNG TIN**

---



**KỲ THI KẾT THÚC HỌC PHẦN HỌC KÌ I**  
**NĂM HỌC 2020-2021**

**Đề bài tập lớn: Đề số 01**

**Sinh viên : Nguyễn Tấn Phát**  
**Mã sinh viên : 1911061044**  
**Lớp : DH9C5**  
**Tên học phần : Lý thuyết thông tin**  
**Giảng viên hướng dẫn: Nguyễn Đức Toàn**

***Hà Nội – 2021***

## Mục Lục

<b>LỜI MỞ ĐẦU .....</b>	<b>3</b>
<b>1. Sơ đồ chức năng và phương pháp xử lý thông tin. ....</b>	<b>4</b>
1.1. Sơ đồ chức năng. ....	4
1.2. Các phương pháp xử lý thông tin. ....	4
<b>2. Nghiên cứu về lý thuyết thông tin. ....</b>	<b>5</b>
2.1. Khái niệm thông tin. ....	5
2.2. Lượng thông tin. ....	5
2.3. Độ bất định và xác suất. ....	6
2.3.1. Xét ví dụ. ....	6
2.3.2. Kết luận. ....	6
2.4. Entropy và Entropy có điều kiện. ....	8
2.4.1. Tính chất thống kê của nguồn rời rạc và sự ra đời của khái niệm entropy. ....	8
2.4.2. Định nghĩa Entropy có điều kiện. ....	9
<b>3. Ví dụ minh họa. ....</b>	<b>9</b>
<b>TÀI LIỆU THAM KHẢO. ....</b>	<b>11</b>

## LỜI MỞ ĐẦU

Vũ trụ mà chúng ta đang sống thì tràn ngập với thông tin. Thông tin cung cấp một ngôn ngữ chung giúp vượt qua rào cản ngăn cách nhiều lĩnh vực riêng biệt: từ thơ của Shakespeare đến các bài báo khoa học của các nhà nghiên cứu trên Cornell ArXiv, từ bản in Đêm Đầy Sao của Van Gogh đến Bản Giao Hưởng Số 5 của Beethoven, từ ngôn ngữ lập trình đầu tiên Plankalkül đến các thuật toán học máy hiện đại nhất. Mọi thứ phải tuân theo các quy tắc của lý thuyết thông tin, bất kể chúng ở định dạng nào. Với lý thuyết thông tin, chúng ta có thể đo lường và so sánh lượng thông tin có trong các tín hiệu khác nhau. Trong phần này, chúng ta sẽ nghiên cứu các khái niệm cơ bản của lý thuyết thông tin và các ứng dụng của lý thuyết thông tin trong học máy.

Trước khi bắt đầu, chúng ta hãy phác thảo mối quan hệ giữa học máy và lý thuyết thông tin. Mục tiêu của học máy là trích xuất các đặc trưng đáng chú ý từ dữ liệu và đưa ra các dự đoán quan trọng. Mặt khác, lý thuyết thông tin nghiên cứu vấn đề mã hóa, giải mã, truyền và thao tác với thông tin. Do vậy, lý thuyết thông tin cung cấp ngôn ngữ cơ bản để thảo luận về việc xử lý thông tin trong các hệ thống học máy. Mát mát này có thể được trực tiếp rút ra từ các quan điểm từ góc nhìn lý thuyết thông tin.



## **2. Nghiên cứu về lý thuyết thông tin.**

### **2.1. Khái niệm thông tin.**

- Điều gì đã xác định (khẳng định được, đoán chắc được, không bấp bênh, ...) thì không có thông tin và người ta nói rằng lượng thông tin chứa trong điều ấy bằng không.
- Điều gì không xác định (bất định) thì điều đó có thông tin và lượng thông tin chứa trong nó khác không. Nếu ta càng không thể ngờ tới điều đó thì thông tin mà điều đó mang lại cho ta rất lớn.
- Tóm lại, ta thấy khái niệm thông tin gắn liền với sự bất định của đối tượng ta cần xét. Có sự bất định về một đối tượng nào đó thì những thông báo về đối tượng đó sẽ cho ta thông tin. Khi không có sự bất định thì sẽ không có thông tin về đối tượng đó. Như vậy, khái niệm thông tin chỉ là một cách diễn đạt khác đi của khái niệm sự bất định.

### **2.2. Lượng thông tin.**

- Trong lý luận ở trên, ta đã từng nói đến lượng thông tin và lượng thông tin lớn, lượng thông tin nhỏ mà không hề định nghĩa các danh từ đó. Dưới đây ta sẽ trả lời vấn đề đó.
- ở trên ta cũng đã nói: trước khi nhận tin thì độ bất định lớn nhất. Sau khi nhận tin (hiểu rõ hoặc hiểu một phần về đối tượng thì độ bất định giảm đến mức thấp nhất, có khi triệt hoàn toàn. Như vậy, có một sự chênh lệch giữa độ bất định trước khi nhận tin và độ bất định sau khi nhận tin. Sự chênh lệch đó là mức độ thủ tiêu độ bất định. Độ lớn, nhỏ của thông tin mang đến ta phụ thuộc trực tiếp vào mức chênh đó. Vậy:
  - + Lượng thông tin là mức độ bị thủ tiêu của độ bất định  $\Leftrightarrow$  Lượng thông tin = độ chênh của độ bất định trước và sau khi nhận tin = độ bất định trước khi nhận tin – độ bất định sau khi nhận tin (độ bất định tiên nghiệm – độ bất định hậu nghiệm).

## 2.3. Độ bất định và xác suất.

### 2.3.1. Xét ví dụ.

- Ta phải chọn một phần tử trong một tập nào đó. Phép chọn như thế (hoặc “chọn” hiểu theo nghĩa rộng: thử, tìm hiểu, điều tra, trinh sát, tình báo,...) bao giờ cũng có độ bất định.
- + Nếu tập chỉ có một phần tử thì ta chẳng phải chọn gì cả và như vậy không có độ bất định trong phép chọn đó.
- + Nếu tập có hai phần tử thì ta đã phải chọn. Như vậy, trong trường hợp này phép chọn có độ bất định. Nếu số phần tử của tập tăng thì độ bất định sẽ tăng
- + Các bước tiếp theo sẽ cho bởi bảng sau:

Số phần tử của tập	Độ bất định của phép chọn	Xác suất chọn một phần tử trong tập
1	0	1
2	$\neq 0$	$1/2$
3	$\neq 0$	$1/3$
.	$\neq 0$	.
.	.	.
.	.	.
n	.	$1/n$
.	$\neq 0$	.
.	.	.
.	.	.
$\infty$	.	$1/\infty = 0$
	$\infty$	

### 2.3.2. Kết luận.

- Bảng này cho thấy: độ bất định gắn liền với bản chất ngẫu nhiên của phép chọn, của biến cố.
- Độ bất định (Ký hiệu I) là hàm của số phần  $(x_k) = f(n)$ .

- Độ bất định có liên quan với xác suất chọn phần tử của tập  $\Rightarrow I(x_k) = E[p(x_k)]$  để tìm mối quan hệ giữa độ bất định I và xác suất chọn một phần tử x,  $(p(x_i))$  trong tập, ta xuất phát từ các tiêu đề sau:

- Theo suy nghĩ thông thường, độ bất định I phải thoả mãn:

$$+ I(x_k) \geq 0$$

$$+ p(x_k)=1 \rightarrow I(x_k) = E[p(x_k)] = E[1] = 0 \quad (3.1)$$

+ Tính cộng được:

Nếu  $x_k$  và  $x_i$  độc lập, thì:

$$E[p(x_k - x_i)] - E[p(x_k)p(x_i)] = E[p(x_k)] + E[p(x_i)]$$

Nếu  $x_k$  và  $x_i$  phụ thuộc thì:

$$E[p(x_k x_i)] = E[p(x_k)p(x_i/x_k)] = E[p(x_k)] + E[p(x_i/x_k)]$$

Đặt  $p(x_k)=p$  và  $p(x_i/x_k)=q$ , thì khi đó với mọi  $p, q$  ( $0 < p \leq 1, 0 < q \leq 1$ ), ta có:

$$E[p] + E[q] = E[pq] \quad (3.2)$$

Từ (3.2) ta có thể tìm được dạng hàm I(p). Lấy vi phân 2 vế của (3.2) theo p, ta có:

$$E'(p) = qE'(pq)$$

Nhân cả 2 vế của phương trình này với p và ký hiệu  $pq = r$ , ta có:

$$pE'(p) = rE'(r) \quad (3.3)$$

Biểu thức (3.3) đúng  $\forall p, r > 0$ . Nhưng điều này chỉ có thể có khi cả hai vế của (3.3)

bằng một hằng số k nào đó:

$$pE'(p) = rE'(r) = k = \text{const}$$

Từ đó chúng ta có phương trình vi phân  $pI'(p) = \text{const} = k$ , lấy tích phân phương trình này, ta tìm được:

$$E(p) = k \ln p + C \quad (3.4)$$

Kể đến điều kiện ban đầu (3.1), chúng ta có:

$$E(p) = k \ln p \quad (3.5)$$

Như vậy, ta có:

$$I(x_k) = k \ln[p(x_k)] \quad (3.6)$$

Hệ số tỷ lệ  $k$  trong (3.6) có thể chọn tùy ý, nó chỉ xác định hệ đơn vị đo của  $I(x, y)$ . Vì  $\ln[p(x_k)] \leq 0$  nên để  $I(x_k) \geq 0$  thì  $k < 0$ .

Nếu lấy  $k = -1$  thì:

$$I(x_k) = -\ln[p(x_k)] = \ln\left[\frac{1}{p(x_k)}\right] \quad (3.7)$$

Khi đó, đơn vị đo độ bất định sẽ là đơn vị tự nhiên, ký hiệu là nat.

Nếu lấy  $k = -\frac{1}{\ln 2}$  thì:

$$I(x_k) = -\frac{\ln p(x_k)}{\ln 2} = -\log_2 p(x_k) \quad (3.8)$$

Khi đó đơn vị đo độ bất định sẽ là đơn vị nhị phân, ký hiệu là bit (1 nat = 1,433 bit)

Một bit chính là độ bất định chứa trong một phân tử (biến cố của tập xác suất chọn (xuất hiện) bằng 1/2. Người ta thường sử dụng đơn vị [bit] do trong kỹ thuật tính và kỹ thuật liên lạc thường dùng các mã nhị phân.

Ngoài ra, người ta còn có thể sử dụng những đơn vị đo khác tùy theo cách chọn cơ sở của logarit. Vì vậy trong trường hợp tổng quát, ta có thể viết:

$$I(x_k) = -\log p(x_k) \quad (3.9)$$

## 2.4. Entropy và Entropy có điều kiện.

### 2.4.1. Tính chất thống kê của nguồn rời rạc và sự ra đời của khái niệm entropy.

- Trong mục trước, ta mới chỉ xét đến lượng thông tin về một biến cố (hay một tin) trong một tập các biến cố (hay tin) xung khắc, đồng xác suất.
- Thực tế tồn tại phổ biến loại tập các biến cố (hay nguồn tin, tập tin) xung khắc, không đồng xác suất. Tức là xác suất xuất hiện các biến cố khác nhau trong tập là khác nhau. Ta gọi sự khác nhau giữa các xác suất xuất hiện biến cố của tập (hay tin của nguồn rời rạc) là tính chất thống kê của nó.
- Trong một nguồn tin như thế, ngoài thông tin riêng của mỗi tin (hay dấu) của nó, người ta còn phải quan tâm đến thông tin trung bình của mỗi tin thuộc nguồn. Người ta còn gọi thông tin trung bình do mỗi dấu của nguồn mang lại là entropy.



### 2.4.2. Định nghĩa Entropy có điều kiện.

- Entropy có điều kiện về một trường tin này khi đã rõ một tin của trường tin kia được xác định bằng ký vọng của lượng thông tin riêng có điều kiện về  $a_k$  do một  $b_i$  mang lại:

$$\begin{aligned} H(A/b_i) &= M [I(a_i / b_i)] = \sum_{i=1}^x p(a_i / b_i) I(a_i / b_i) \\ &= - \sum_{i=1}^x p(a_i / b_i) \log p(a_i / b_i) \end{aligned}$$

- Ý nghĩa:  $H(A / b_i)$  là lượng thông tin tồn hao trung bình của mỗi tin ở đầu phát khi đầu thu đã thu được  $b_j$ .
- Tương tự:

$$H(B / a_i) = - \sum_{j=1}^x p(b_j / a_i) \log p(b_j / a_i)$$

- Ý nghĩa:  $H(B / a_i)$  là lượng thông tin riêng trung bình chứa trong mỗi tin ở đầu thu khi đầu phát đã phát đi một tin  $a_i$ .

### 3. Ví dụ minh họa.

Gọi X là tập hợp gồm 52 quân bài thì X gồm các  $x_i$ ; quân bài ( $i=1,2,3... 52$ ).

Khả năng chọn 1 trong số 52 quân bài là như nhau, xác suất là

$$p(x_i) = 1/52$$

Độ bất định chứa trong sự kiện rút 1 quân bài là:

$$I(x_i) = -\log p(x_i) = -\log 1/52 = \log 52 \text{ (bit)}$$

Khi B đặt câu hỏi cho A, sau mỗi lần hỏi, các câu trả lời có thể là:

- Đúng (xác suất nhận câu trả lời Đúng là  $p$ )
- Sai (xác suất nhận câu trả lời Sai là  $1-p$ )

Coi B là nguồn phát tín hiệu khi hỏi, thì B có thể được coi là nguồn rời rạc có 2 tín hiệu (Đúng, Sai)  $s = 2$ .

Lượng thông tin nhận được sau mỗi lần hỏi:

$$\begin{aligned} H(B) &= - \sum_{i=1}^2 p(b_i) \log p(b_i) \\ &= -p \log p - (1-p) \log (1-p) \end{aligned}$$

Để xác định được quân bài mà A đã rút, tổng lượng thông tin nhận được sau các lần hỏi phải bằng độ bất định của 1 lần rút quân bài.

Gọi số lần hỏi cần thiết là  $n$ , thì  $n = I(x_i)/H(B)$

Để số lần hỏi là nhỏ nhất thì  $H(B)$  phải đạt giá trị cực đại.

Do B được coi là nguồn rời rạc có  $s = 2$  tín hiệu (Đúng, Sai) nên:

$$H(B) = H(B)_{\max} = \log s = \log 2 = 1 \text{ (bit)}$$

$H(B)_{\max}$  đạt được khi  $p = 1 - p = 1/2$

Khi đó

$$n_{\min} = \frac{I(X_i)}{H(B)_{\max}} = \frac{\log 52}{1} = 6 \text{ lần}$$

Vậy B cần hỏi ít nhất là 6 lần cần để tìm được quân bài mà A đã rút

Thuật toán hỏi khi A đã rút ra quân 7 rô như sau:

Câu hỏi 1: Quân A rút ra là quân đỏ? Đúng

Câu hỏi 2: Quân A rút ra là quân cơ? Sai

Câu hỏi 3: Quân A rút ra có giá trị  $< 7$ ? Đúng

Câu hỏi 4: Quân A rút ra có giá trị  $< 3$ ? Sai

Câu hỏi 5: Quân A rút ra có giá trị  $< 5$ ? Sai

Câu hỏi 6: Quân A rút ra là 6 rô? Sai

Vậy quân A rút ra là 7 rô

## **TÀI LIỆU THAM KHẢO.**

1. GS.TS Nguyễn Bình, TS. Ngô Đức Thiện; Giáo trình Lý thuyết thông tin; Học viện công nghệ Bưu chính viễn thông.