

SỞ GIÁO DỤC VÀ ĐÀO TẠO
THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG TRUNG HỌC LÝ THƯỜNG KIỆT
TỔ TIN HỌC



TÀI LIỆU HỌC TẬP ĐỀ CƯƠNG ÔN TẬP HKII KHỐI 12

Năm học 2025 - 2026

Tài liệu lưu hành nội bộ

CHỦ ĐỀ 4. GIẢI QUYẾT VẤN ĐỀ VỚI SỰ TRỢ GIÚP CỦA MÁY TÍNH

BÀI 16: ĐỊNH DẠNG KHUNG

A. LÝ THUYẾT

1. Phân loại phần tử khối và nội tuyến

- Các thẻ (hay phần tử) HTML được chia làm hai loại:
- + **Khối** (block level): Các phần tử khối thường bắt đầu từ đầu hàng và kéo dài suốt chiều rộng của trang web.
- + **Nội tuyến** (inline level): Các phần tử nội tuyến là các phần tử nhúng bên trong một phần tử khác.
- Mặc định các phần tử HTML sẽ thuộc một trong hai loại khối hoặc nội tuyến (Bảng 16.1).

Bảng 16.1. Phân loại phần tử CSS

Phần tử loại khối	Phần tử loại nội tuyến
h1 – h6, p, div, address, nav, article, section, aside, form, header, footer, table, hr, ol, ul, li, canvas	b, span, a, img, em, strong, sub, sup, var, samp, cite, dfn, kbd, pre, code, q, i, u, del, ins, mark, br, label, textarea, input, script

⇒ Có thể thay đổi loại phần tử HTML bằng thuộc tính **display**. Các giá trị của thuộc tính này bao gồm **block**, **inline**, **none**. Giá trị **none** sẽ làm ẩn (không hiển thị) phần tử này trên trang web.

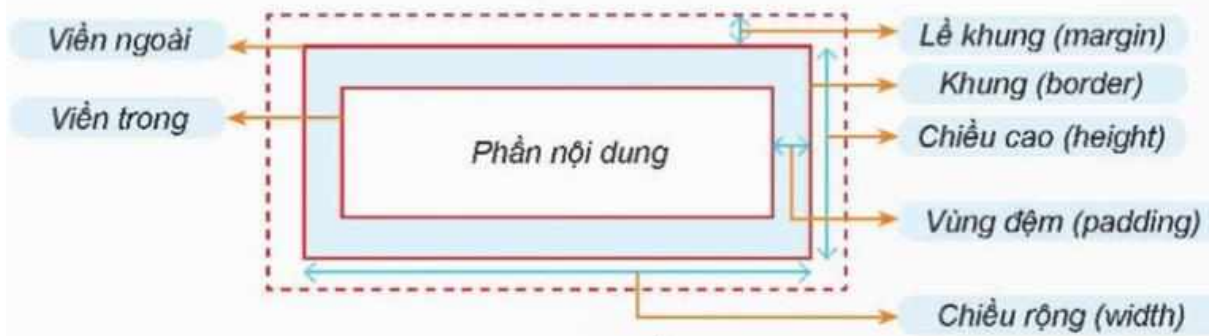
2. Thiết lập định dạng khung bằng css

- **Lưu ý:** Cần phân biệt hai loại phần tử HTML, phần tử khối và phần tử nội tuyến. Với phần tử dạng khối, các khung được xác định với đầy đủ tính chất, còn với các phần tử nội tuyến thì khung chỉ có thể thiết lập mà không có các thông số chiều cao, chiều rộng.

Thuộc tính	Ý nghĩa
width	Chiều rộng của khung. Thuộc tính này chỉ áp dụng cho phần tử dạng khối.
height	Chiều cao khung. Thuộc tính này chỉ áp dụng cho phần tử dạng khối.
padding	Vùng đệm, khoảng cách từ vùng text đến đường viền ngoài của khung.
margin	Lề khung, khoảng cách từ đường viền ngoài của khung đến văn bản xung quanh (nếu có).
border-color	Màu của viền khung.
border-width	Độ dày của đường viền khung.
border-style	Kiểu đường viền khung. Các giá trị có thể là: none, solid, dotted, dashed, double, inset, outset, ridge, groove.
border	Thuộc tính này có thể gán giá trị là đồng thời các thuộc tính border-width, border-style và border-color, ví dụ: {border: 2px solid red;}

Bảng 16.2. Các thuộc tính liên quan đến khung

Quan sát Hình 16.3 để biết các thông số chính của khung của phần tử HTML để có thể hiểu được cách thiết lập khung, viền bằng CSS.



Hình 16.3. Các thông số chính của khung

3. MỘT SỐ BỘ CHỌN ĐẶC BIỆT CỦA CSS

a) Thiết lập bộ chọn là **một lớp (class)** các phần tử có ý nghĩa gần giống nhau

- Trong thực tế, có thể có nhu cầu định dạng cho một nhóm phần tử có cùng ý nghĩa, ví dụ các đoạn văn bản có liên quan đến một sự kiện nào đó hoặc một số đoạn văn bản quan trọng cần nhấn mạnh.

⇒ Trong các trường hợp này, thiết lập bộ chọn lớp **class** cho các phần tử đó để có thể thiết lập định dạng chung.

- Cấu trúc chung của bộ chọn lớp:

```
.class {thuộc tính : giá trị;}
```

Bảng 16.3. Một số định dạng hay gặp với bộ chọn lớp

Bộ chọn	Ý nghĩa
.test	Áp dụng cho tất cả các phần tử có lớp tên là test.
.test.warning	Áp dụng cho tất cả các phần tử có tên lớp đồng thời là test và warning. Ví dụ: <p class "test warning">...</p>
p.test	Áp dụng cho tất cả các phần tử p có gắn lớp tên test.

b) Thiết lập bộ chọn riêng cho từng phần tử riêng biệt có **mã định danh id**

- Nếu muốn định dạng cho các phần tử riêng biệt thì cần thiết lập mã định danh **id** cho phần tử này.

- Trong một trang web, mỗi phần tử chỉ có một mã định danh id duy nhất.

- Cấu trúc chung của bộ chọn định danh:

```
#idname {thuộc tính : giá trị ;}
```

- Ví dụ một số mẫu định dạng id như sau:

```
#home {color : red;} /*chữ màu đỏ cho phần tử với id = "home"*/
p#home {font-size: 150%;} /*cỡ chữ 150% cho phần tử p có id = "home"*/
```

c) Thiết lập bộ **chọn thuộc tính CSS**

- Một tính chất quan trọng khác của CSS là có thể thiết lập bộ chọn là thuộc tính. Các định dạng này sẽ được thiết lập và áp dụng cho các phần tử nếu được gắn với thuộc tính cụ thể nào đó.

- Ví dụ:

```
[href] {border: 1px solid blue;}
/*áp dụng cho tất cả các phần tử có thuộc tính href, ví dụ các phần tử a.*/
a[target = "_blank"] {color : red; font-weight: bold;}
/*áp dụng chữ màu đỏ và in đậm cho tất cả các thẻ a với thuộc tính target =
"_blank" */
```

*** Lưu ý: Khi đặt tên cho id và class:**

- Tên của id và class phân biệt chữ in hoa, in thường.
- Tên bắt buộc phải có ít nhất một ký tự không là số, không bắt đầu bằng số, không chứa dấu cách và các ký tự đặc biệt khác.
- Một phần tử có thể thuộc nhiều lớp khác nhau. Để khai báo, chúng ta đặt các tên lớp cách nhau bởi dấu cách. Trong ví dụ sau phần tử p thuộc đồng thời ba lớp là “test”, “more” và “once”:

```
<p class = "test more once"> Đoạn chương trình cần kiểm tra. </p>
```

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Các thẻ (hay phần tử HTML) được chia thành những loại nào?

- | | |
|------------------------------|-----------------------|
| A. Khối và ngoại tuyến. | C. Khối và nội tuyến. |
| B. Nội tuyến và ngoại tuyến. | D. Trong và ngoài. |

Câu 2: Phần tử HTML nào sau đây thuộc loại nội tuyến?

- | | | | |
|------------|-----------|--------|-------------|
| A. script. | B. table. | C. li. | D. article. |
|------------|-----------|--------|-------------|

Câu 3: Thuộc tính nào sau đây chỉ áp dụng cho phần tử dạng khối?

- | | |
|-------------|------------------|
| A. margin. | C. border-style. |
| B. padding. | D. width. |

Câu 4: Phát biểu nào sau đây **sai**?

- A. Một phần tử có thể thuộc nhiều lớp khác nhau.
- B. Có thể thiết lập bộ chọn là thuộc tính trong CSS.
- C. Tên của id và class phân biệt chữ in hoa, in thường.
- D. Các thuộc tính liên quan đến khung của một phần tử HTML đều có tính kế thừa.

Câu 5: Thuộc tính border-style **không** nhận giá trị nào sau đây?

- | | | | |
|------------|------------|-----------|------------|
| A. Groove. | B. Shadow. | C. Ridge. | D. Dashed. |
|------------|------------|-----------|------------|

Câu 6: Em có thể thay đổi loại phần tử HTML bằng thuộc tính gì?

- | | | | |
|----------|-------------|----------|-------------|
| A. show. | B. display. | C. type. | D. section. |
|----------|-------------|----------|-------------|

Câu 7: Để định dạng khoảng cách từ vùng text đến đường viền ngoài của khung, em sử dụng thuộc tính nào?

- | | |
|-------------|------------------|
| A. margin. | C. padding. |
| B. spacing. | D. border-width. |

Câu 8: Cách thiết lập các mẫu định dạng với các phần tử có id tương ứng là

- | | |
|-------------------------------------|------------------------------------|
| A. .idname {thuộc tính: giá trị;}. | C. #idname {thuộc tính: giá trị;}. |
| B. id[name] {thuộc tính: giá trị;}. | D. idname {thuộc tính: giá trị;}. |

Câu 9: Cho mẫu CSS sau: h1 {padding: 5px 10px 3px 20px;}. Phát biểu nào sau đây **đúng**?

- | | |
|------------------------|--------------------------|
| A. padding-left = 3px. | C. padding-right = 10px. |
| B. padding-top = 20px. | D. padding-bottom = 5px. |

TRẮC NGHIỆM ĐÚNG SAI

Câu 10: Các thẻ (hay phần tử) HTML được chia làm hai loại, khối (block level) và nội tuyến (inline level).

- a) Các phần tử khối là các phần tử nhúng bên trong một phần tử khác.
- b) Một số phần tử nội tuyến là h1 – h6, p, div, b, a, label, input, table,...
- c) Chiều rộng của phần tử nội tuyến không phụ thuộc vào độ rộng của phần văn bản của phần tử này mà chỉ phụ thuộc vào chiều rộng của cửa sổ trình duyệt.
- d) Có thể thay đổi loại phần tử HTML bằng thuộc tính display.

Câu 11: Có thể thiết lập các mẫu định dạng với bộ chọn là class, ID hoặc thuộc tính.

- a) Cấu trúc chung của định dạng CSS liên quan đến lớp là .class {thuộc tính: giá trị;}.
- b) Tên của id và class không phân biệt chữ in hoa, in thường.
- c) Mỗi mã định danh (id) chỉ được gán duy nhất cho một phần tử HTML.
- d) Mẫu định danh này áp dụng cho các phần tử p là phần tử con của một phần tử div có mã định danh #price là div#price ~ p.

BÀI 17: CÁC MỨC ƯU TIÊN CỦA BỘ CHỌN

A. LÝ THUYẾT

1. Kiểu bộ chọn dạng pseudo-class và pseudo-element

a) Bộ chọn pseudo-class

- Pseudo-class (lớp giả) là khái niệm chỉ các trạng thái đặc biệt của phần tử HTML. Các trạng thái này không cần định nghĩa và mặc định được coi như các lớp có sẵn của CSS.
- Trong CSS, các lớp giả quy định viết sau dấu “:” theo cú pháp:

```
:pseudo-class {thuộc tính : giá trị ;}
```

Bảng 17.1. Một số lớp giả thường dùng

Bộ chọn	Ý nghĩa	Ví dụ
:link	Tất cả các liên kết khi chưa được kích hoạt.	Các liên kết (khi chưa kích hoạt) sẽ có màu xanh dương. a:link {color: blue;}
:visited	Tất cả các liên kết sau khi đã được kích hoạt một lần.	Các liên kết sau khi kích hoạt chuyển màu xám. a:visited {color: gray;}
:hover	Tất cả các phần tử, khi người dùng di chuyển con trỏ chuột lên đối tượng có id="home" sẽ hiển thị với cỡ chữ tăng lên 150%. trỏ chuột lên đối tượng.	#home:hover {font-size: 150%;}

⇒ Khai báo lớp giả để làm thay đổi các giá trị mặc định của đối tượng → Gây sự chú ý.

b) Bộ chọn kiểu pseudo-element

- Pseudo-element (phần tử giả) là khái niệm chỉ một phần (hoặc một thành phần) của các phần tử bình thường. Các phần này có thể coi là một phần tử giả và có thể thiết lập mẫu định dạng CSS.
- Quy định phần tử giả viết sau dấu “::” theo cú pháp:

```
::pseudo-element {thuộc tính : giá trị ;}
```

Bảng 17.2. Một số phần tử giả

Bộ chọn	Ý nghĩa	Ví dụ
<code>::first-line</code>	Dòng đầu tiên của đối tượng.	Dòng đầu tiên của các đoạn thuộc lớp test chuyển phong monospace. <code>p.test::first-line {font-family: monospace;}</code>
<code>::first-letter</code>	Kí tự đầu tiên của đối tượng.	Kí tự đầu tiên của đoạn có id="first" có màu đỏ và kích thước gấp đôi bình thường. <code>p#first::first-letter {font-size: 200%;}</code>
<code>::selection</code>	Phần được chọn (bằng cách kéo thả chuột trên màn hình) của đối tượng.	Vùng đang chọn bất kì sẽ chuyển màu nền xanh lá mạ. <code>::selection {background-color: lime;}</code>

⇒ Khai báo phần tử giả để thiết lập các mẫu định dạng đặc biệt cho phần tử giả đó khác so với mặc định.

2. Mức độ ưu tiên khi áp dụng css

** Thứ tự ưu tiên các mẫu định dạng trong CSS khi có nhiều hơn các mẫu cùng áp dụng cho một phần tử HTML*

Khi có nhiều mẫu định dạng có thể áp dụng cho một phần tử HTML nào đó trên trang web, CSS sẽ áp dụng định dạng theo thứ tự ưu tiên:

Bảng 17.3. Thứ tự (mức) ưu tiên của CSS

STT	CSS	Giải thích
1	!important	Các thuộc tính trong CSS với từ khoá !important sẽ có mức ưu tiên cao nhất.
2	CSS trực tiếp (inline CSS).	Các định dạng nằm ngay trong phần tử HTML với thuộc tính style .
3	CSS liên quan đến kích thước thiết bị (Media type).	Các định dạng loại này thường dùng để điều khiển cách hiển thị thông tin phụ thuộc vào kích thước màn hình của thiết bị. Ví dụ mẫu định dạng sau sẽ tăng kích thước chữ lên 150% nếu chiều ngang màn hình nhỏ hơn 600 px: <code>@media only screen and (max-width: 600px) { body {font-size: 150%; } }</code>
4	Trọng số CSS.	Mỗi định dạng CSS sẽ có trọng số (specificity) riêng của mình. Tại mức ưu tiên này, định dạng CSS có trọng số cao nhất sẽ được áp dụng.
5	Nguyên tắc thứ tự cuối cùng (Rule order).	Nếu có nhiều mẫu định dạng với cùng trọng số thì định dạng ở vị trí cuối cùng sẽ được áp dụng.
6	Kế thừa từ CSS cha.	Nếu không tìm thấy mẫu định dạng tương ứng thì sẽ lấy thông số định dạng CSS kế thừa từ phần tử cha.
7	Mặc định theo trình duyệt.	Nếu không có bất cứ định dạng CSS nào thì trình duyệt quyết định thể hiện nội dung mặc định.

⇒ Theo nguyên tắc trên, nếu có một dãy các mẫu định dạng CSS cùng có thể áp dụng cho một phần tử HTML thì tính kế thừa CSS và nguyên tắc thứ tự cuối cùng được xếp dưới trọng số CSS, tức là khi đó CSS sẽ tính trọng số các mẫu định dạng, cái nào có trọng số lớn hơn sẽ được ưu tiên áp dụng.

** Cách tính trọng số của một mẫu định dạng*

- Cách tính trọng số của CSS rất đơn giản dựa trên giá trị trọng số của từng thành phần của bộ chọn (selector) trong mẫu định dạng.
- Trọng số của mẫu định dạng sẽ được tính bằng tổng của các giá trị thành phần đó.

Bảng 17.4. Giá trị của các thành phần của bộ chọn

STT	Bộ chọn	Giá trị đóng góp trọng số
1	Mã định danh (ID).	100
2	Lớp, lớp giả, bộ chọn kiểu thuộc tính. (Class, pseudo-class, attribute selector).	10
3	Phần tử, phần tử giả (element, pseudo-element).	1
4	*	0

Bảng 17.5. Một số ví dụ tính trọng số

Bộ chọn	Trọng số	Giải thích
p > em	2	Bộ chọn có hai phần tử là p và em. ⇒ Trọng số bằng $1 + 1 = 2$.
.test #p11	110	Bộ chọn bao gồm 1 class và 1 id. ⇒ Trọng số bằng $10 + 100 = 110$.
p.test em.more	22	Bộ chọn có hai phần tử (p, em) và hai class (test, more). ⇒ Trọng số bằng $2 + 20 = 22$.
p > em#p123	102	Bộ chọn có hai phần tử (p, em) và 1 id. ⇒ Trọng số bằng $2 + 100 = 102$.

❖ Một số ví dụ về mức độ ưu tiên khi áp dụng CSS:

- Ví dụ: Các thuộc tính trong CSS với từ khóa **!important** sẽ có mức ưu tiên cao nhất.

```

<!DOCTYPE html>
<html>
  <head>
    <style>
      #myid {
        background-color: blue;
      }
      .myclass {
        background-color: gray;
      }
      p {
        background-color: pink !important;
      }
    </style>
  </head>
  <body>
    <p>This is some text in a paragraph.</p>
    <p class="myclass">This is some text in a paragraph.</p>
    <p id="myid">This is some text in a paragraph.</p>
  </body>
</html>

```

This is some text in a paragraph.

This is some text in a paragraph.

This is some text in a paragraph.

- Ví dụ: CSS trực tiếp (inline CSS)

```
<!DOCTYPE html>
<html>
  <body>
    <h1 style="color:blue;text-align:center;">This is a heading</h1>
    <p style="color:red;">This is a paragraph.</p>
  </body>
</html>
```

This is a heading

This is a paragraph.

- Ví dụ: Trọng số CSS:

```
<!DOCTYPE html>
<html>
  <head>
    <style>
      .favorite {
        color: red;
        font-weight: bold;
      }
      ul#summer-drinks li {
        font-weight: normal;
        font-size: 18px;
        color: green;
      }
    </style>
  </head>
  <body>
    <ul id="summer-drinks">
      <li class="favorite">Whiskey and Ginger Ale</li>
      <li>Wheat Beer</li>
      <li>Mint Julip</li>
    </ul>
  </body>
</html>
```

- Whiskey and Ginger Ale
- Wheat Beer
- Mint Julip

- Ví dụ: Nguyên tắc thứ tự cuối cùng (Rule order):

```
<!DOCTYPE html>
<html>
  <head>
    <style>
      p {color: green;}
      p {color: purple;}
    </style>
  </head>
  <body>
    <p>This is a paragraph</p>
  </body>
</html>
```

This is a paragraph

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Trong CSS, các lớp giả quy định được viết sau dấu gì?

- A. Dấu ":". B. Dấu "#". C. Dấu ".". D. Dấu "::".

Câu 2: Kiểu bộ chọn nào mô tả các thành phần (nhỏ hơn) của phần tử?

- A. pseudo-class. B. id.

C. pseudo-element.

D. class.

Câu 3: Trong CSS, quy định phần tử giả viết sau dấu gì?

A. Dấu ".".

B. Dấu "::".

C. Dấu ":".

D. Dấu "#".

Câu 4: Phát biểu nào sau đây **sai**?

A. Các lớp giả không cần định nghĩa và mặc định được coi như các lớp có sẵn của CSS.

B. Nếu không tìm thấy mẫu định dạng tương ứng thì sẽ lấy thông số định dạng CSS kế thừa từ phần tử cha.

C. Bộ chọn ::selection mô tả phần được chọn (bằng cách kéo thả chuột trên màn hình) của đối tượng.

D. Nếu có nhiều mẫu định dạng với cùng trọng số thì định dạng ở vị trí đầu tiên sẽ được áp dụng.

Câu 5: Em hãy sắp xếp theo thứ tự từ cao xuống thấp các mức độ ưu tiên khi áp dụng CSS sau:

1. Trọng số CSS.

2. Mặc định theo trình duyệt.

3. CSS liên quan đến kích thước thiết bị.

4. Kế thừa từ CSS cha.

A. 3, 1, 4, 2.

C. 1, 4, 2, 3.

B. 2, 3, 4, 1.

D. 3, 2, 1, 4.

Câu 6: Giá trị trọng số của bộ chọn ::selection là bao nhiêu?

A. 0.

B. 1.

C. 10.

D. 100.

Câu 7: Mẫu CSS định dạng các liên kết sau khi kích hoạt chuyển màu xanh dương là

A. a:link {color: blue;}.

C. a:hover {color: blue;}.

B. a:visited {color: blue;}.

D. a:clicked {color: blue;}.

Câu 8: Mẫu CSS thiết lập định dạng khi di chuyển con trỏ chuột lên đối tượng có id="title" sẽ hiển thị với kiểu chữ đậm là

A. #title:hover {font-style: bold;}.

C. title:hover {font-style: italics;}.

B. .title:hover {font-weight: bold;}.

D. #title:hover {font-weight: bold;}.

Câu 9: Mẫu CSS thiết lập định dạng dòng đầu tiên của các đoạn thuộc lớp description có kích thước gấp 3 lần cỡ chữ hiện thời của trình duyệt là

A. p#description::first-line {font-size: 3rem;}.

C. p.description::first-line {font-size: 3em;}.

B. p.description::first-line {font-size: 3pt;}.

D. #p.description::first-line {font-size: 3em;}.

Câu 10: Bộ chọn a:link ~ p#info có trọng số bằng bao nhiêu?

A. 22.

B. 130.

C. 112.

D. 220.

TRẮC NGHIỆM ĐÚNG SAI

Câu 11: CSS hỗ trợ thiết lập định dạng cho các lớp giả (pseudo-class) và phần tử giả (pseudo-element).

a) Trong CSS, các lớp giả quy định viết sau dấu "::".

b) Phần tử giả mô tả các thành phần (nhỏ hơn) của phần tử.

c) Bộ chọn ::selection mô tả phần được chọn (bằng cách kéo thả chuột trên màn hình) của đối tượng.

d) Để đổi màu chữ của một đoạn văn bản có mã định danh là note khi di chuyển chuột qua đoạn văn bản đó cần dùng định dạng CSS #note:visited {color: blue;}.

Câu 12: Khi có nhiều mẫu định dạng có thể áp dụng cho một phần tử HTML nào đó trên trang web, CSS sẽ áp dụng định dạng theo thứ tự ưu tiên.

a) Các thuộc tính trong CSS với kí hiệu * sẽ có mức ưu tiên cao nhất.

b) Quy tắc kế thừa từ CSS cha có thứ tự ưu tiên cao hơn quy tắc CSS trực tiếp (inline CSS).

- c) Nếu có nhiều mẫu định dạng CSS cùng mức ưu tiên áp dụng cho một phần tử HTML thì mẫu CSS nào có trọng số cao nhất sẽ được áp dụng.
- d) Bộ chọn .test > div#p123 có trọng số là 102.

CHỦ ĐỀ 6. MẠNG MÁY TÍNH VÀ INTERNET

BÀI 22: TÌM HIỂU THIẾT BỊ MẠNG

A. LÝ THUYẾT

1. Server:

- Thuật ngữ **server** có nguồn gốc từ “serve” nghĩa là phục vụ. Server có nghĩa là chủ thể cung cấp dịch vụ.
- Máy tính làm server thường được gọi là “máy chủ”. Máy tính yêu cầu và được cung cấp dịch vụ từ máy chủ được gọi là “máy khách” (client).
- Khái niệm: Server là một hệ thống phần mềm và phần cứng cung cấp các dịch vụ cho nhiều người dùng từ các máy tính khác gọi là máy khách.
- Các máy khách và máy chủ phải được kết nối với nhau qua mạng.
- Trong một mạng có thể không có máy chủ nhưng máy chủ thường nằm trong một mạng.
- Do cần đáp ứng yêu cầu từ nhiều máy khách, các server thường là các máy tính có cấu hình cao, tin cậy, an toàn bảo mật, hiệu suất cao và có khả năng mở rộng.
- Các máy chủ như file server, database server, web server không được coi là thành phần của mạng mà chỉ là thiết bị đầu cuối, sử dụng mạng để trao đổi thông tin.
- Chỉ những server cung cấp các dịch vụ quản trị mạng thì mới được coi như một thành phần của mạng.

2. Nhận diện và tìm hiểu tính năng kỹ thuật của các thiết bị kết nối

a) Hub và switch

- Về hình thức, khó phân biệt giữa switch, hub nếu không đọc các thông tin đi kèm thiết bị hoặc từ hồ sơ kỹ thuật.
 - Xung đột tín hiệu làm giảm hiệu quả truyền dữ liệu rất nhiều.
 - Miền xung đột (collision domain) là một phần của mạng gồm một nhóm máy tính được kết nối với nhau mà khi nhiều máy tính gửi tín hiệu đồng thời lên mạng sẽ gây ra xung đột tín hiệu.
- ⇒ Việc giảm quy mô của các miền xung đột đến một mức hợp lý là một việc cần làm trong thiết kế mạng.
- + Các máy tính nối vào cùng một hub sẽ thuộc về cùng một miền xung đột do hub chỉ là một bộ chia tín hiệu, cho tín hiệu lan toả từ một cổng ra tất cả các cổng khác.
 - + Các máy tính nối vào các cổng khác nhau của switch sẽ thuộc về các miền xung đột khác nhau do switch chỉ thiết lập kết nối tạm thời cổng của hai máy tính trong thời gian truyền.
- ⇒ Thay hub bằng switch sẽ giúp chia nhỏ các miền xung đột.
- ⇒ Đối với những mạng nhiều máy tính, dùng switch là thích hợp dù chi phí có cao hơn hub.
- Hub hay switch đều có một số tính năng quan trọng thường được ghi ở trên thiết bị:
 - + Là hub hay switch.
 - + Số cổng (nhiều cổng thì có thể kết nối được với nhiều thiết bị).
 - + Tốc độ truyền dữ liệu qua các cổng (100 Megabit/s, 1 Gigabit/s hay 10 Gigabit/s. Cổng Gigabit được hiểu là cổng có thể truyền với tốc độ từ một Gigabit/s trở lên).

b) Router

- Có thể nhận diện router qua đặc trưng có cổng WAN để kết nối ra ngoài mạng cục bộ:
- + Các router dùng trong gia đình thường chỉ có một cổng WAN (có màu khác với cổng LAN, còn gọi là cổng Ethernet) và được ghi rõ “WAN” hoặc “Internet”.
- + Các router của các mạng lớn có thể có nhiều cổng WAN.
- Chức năng cơ bản nhất của router là chọn đường, phục vụ kết nối các mạng LAN với nhau. - Một số thông số kỹ thuật của router:
 - + Số cổng kết nối. Khác với hub/switch, router cần phân biệt cổng WAN và cổng LAN. Số cổng WAN là một thông số có ý nghĩa đối với router. Khi có nhiều cổng WAN, router mới có thể định tuyến thực sự và khả năng thiết lập kết nối dự phòng.
 - + Tốc độ truyền dữ liệu qua các cổng.
 - + Số lượng truy cập đồng thời.
 - + Một số tính năng khác thường được đảm bảo bởi phần mềm tích hợp trên router.
- Các router nhiều cổng, công suất lớn thường được dùng cho các tổ chức nhiều người dùng, cần truy cập Internet với khối lượng dữ liệu lớn, ổn định, an toàn cao như các trường đại học lớn, các tổ chức cung cấp dịch vụ trực tuyến cho nhiều người hay các công ty cung cấp dịch vụ mạng.
- Đối với các mạng nhỏ như mạng gia đình, mạng văn phòng, chỉ cần một router có một cổng WAN kết nối Internet. Không những thế, router có thể tích hợp với thiết bị thu phát Wi-Fi.

c) Repeater

- Khi sử dụng cáp xoắn, khoảng cách sử dụng có hiệu quả chỉ khoảng 100 m. Nếu truyền xa hơn, tín hiệu có thể bị yếu đi, bị biến dạng hay bị nhiễu đến mức không dùng được nữa.
- ⇒ Cần dùng một thiết bị gọi là “bộ lặp” (repeater) để sửa lại tín hiệu. Repeater nhận tín hiệu ở một đầu, chỉnh sửa rồi phát lại tín hiệu tốt về đầu bên kia.
- ⇒ Như vậy, có thể dùng repeater để mở rộng phạm vi địa lý của mạng cục bộ.

d) Bộ thu phát Wi-Fi

- Bộ thu phát Wi-Fi (còn gọi là điểm truy cập không dây) cho phép các thiết bị kết nối không dây vào mạng.
- Bộ thu phát Wi-Fi có thể được tích hợp với router hoặc độc lập.
- Một số thông số kỹ thuật quan trọng của bộ thu phát Wi-Fi:
 - + Băng tần hỗ trợ (tần số làm việc) thường tính theo GigaHz.
 - + Băng thông - tốc độ truyền tính theo Megabit/s hoặc Gigabit/s, phụ thuộc vào giao thức hỗ trợ.
 - + Khoảng cách hiệu quả (độ phủ) phụ thuộc vào công suất phát.
 - + Số lượng người dùng có thể truy cập đồng thời.
 - + Môi trường làm việc là trong nhà (indoor) hay ngoài trời (outdoor),...
- Một số ví dụ về bộ thu phát Wi-Fi:
 - + Bộ thu phát Wi-Fi như Hình 22.6 hỗ trợ băng tần kép, có thể cung cấp một băng thông 300 Mb/s trên băng tần 2,4 GHz và một băng thông 867 Mb/s trên băng tần 5 GHz, phục vụ được 25 người truy cập đồng thời trong phạm vi vài chục mét.
- ⇒ Thích hợp với quy mô gia đình hay một văn phòng của một cơ quan.
- + Bộ thu phát Wi-Fi trong Hình 22.7 hoạt động trên tần số 5 GHz hỗ trợ giao thức 802.11ac cho băng thông 5,2 Gigabit/s, có thể phục vụ cho 400 người truy cập đồng thời

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Phương án nào sau đây không phải là thông số kỹ thuật của bộ thu phát Wi-Fi?

- A. Băng tần hỗ trợ (tần số làm việc). thường tính theo GigaHz.
- B. Băng thông (tốc độ truyền).
- C. Khoảng cách hiệu quả (độ phủ) phụ thuộc vào công suất phát.
- D. Số cổng LAN.

Câu 2: Server là gì?

- A. Là một máy tính lớn, có hình dạng khác với máy tính để bàn thông thường như dạng tháp (tower), dạng phiến (blade) hay dạng chõng (rack).
- B. Là một hệ thống gồm phần cứng và phần mềm cung cấp một dịch vụ nào đó trên mạng máy tính.
- C. Là máy tính chạy trên mạng, dùng chung cho nhiều người.
- D. Là dịch vụ phần mềm dùng chung trên mạng máy tính.

Câu 3: DNS có chức năng gì?

- A. Thẩm định quyền hạn người dùng
- B. Lưu trữ dữ liệu
- C. Cung cấp dịch vụ quản trị mạng
- D. Xác định địa chỉ IP từ tên miền

Câu 4: Máy tính làm nhiệm vụ server thường được gọi là gì?

- A. Máy khách
- B. Máy chủ
- C. Máy tính cá nhân
- D. Máy tính xách tay

Câu 5: Máy chủ xác thực (authentication server) có vai trò gì?

- A. Cung cấp địa chỉ IP cho máy khách.
- B. Xác định danh tính và quyền truy cập của người dùng.
- C. Lưu trữ dữ liệu của người dùng.
- D. Quản lý các kết nối mạng.

Câu 6: Máy tính yêu cầu dịch vụ từ máy chủ được gọi là gì?

- A. Máy chủ
- B. Máy khách
- C. Máy tính cá nhân
- D. Máy tính xách tay

Câu 7: Server là một hệ thống bao gồm những gì?

- A. Phần mềm
- B. Phần cứng
- C. Cả phần mềm và phần cứng
- D. Không có phần nào cả

Câu 8: Máy chủ xác thực (authentication server) dùng để làm gì?

- A. Xác định địa chỉ IP từ tên miền
- B. Thẩm định quyền hạn người dùng
- C. Cung cấp dịch vụ quản trị mạng
- D. Lưu trữ dữ liệu

Câu 9: DNS là viết tắt của cụm từ nào?

- A. Domain Name System
- B. Digital Network Server
- C. Dynamic Network Solution
- D. Data Network Storage

Câu 10: Vì sao việc lựa chọn thiết bị mạng phù hợp lại quan trọng khi thiết kế mạng máy tính?

- A. Để tiết kiệm chi phí đầu tư.
- B. Để đảm bảo tính thẩm mỹ cho mạng.
- C. Để mạng hoạt động hiệu quả và đáp ứng nhu cầu sử dụng.
- D. Để mạng dễ dàng bảo trì và nâng cấp.

Câu 11: Vì sao switch được ưu tiên sử dụng hơn hub trong mạng nhiều máy tính?

- A. Switch có độ bền cao hơn hub.

- B. Switch tiết kiệm điện năng hơn hub.
- C. Switch giúp tăng cường bảo mật cho mạng.
- D. Switch giúp cải thiện hiệu quả truyền dữ liệu và giảm xung đột.

Câu 12: Repeater có vai trò gì trong mạng máy tính?

- A. Khuếch đại tín hiệu để mở rộng phạm vi của mạng.
- B. Kết nối các mạng LAN lại với nhau.
- C. Chia sẻ kết nối internet cho nhiều thiết bị.
- D. Bảo vệ mạng khỏi các tấn công từ bên ngoài.

Câu 13: Sự khác biệt cơ bản giữa hub và switch là gì?

- A. Hub có giá thành rẻ hơn switch.
- B. Hub dễ dàng cài đặt và sử dụng hơn switch.
- C. Hub chia sẻ băng thông cho tất cả các thiết bị, trong khi switch kết nối riêng biệt cho từng thiết bị.
- D. Hub chỉ có cổng RJ45, còn switch có cả cổng RJ45 và cổng quang.

TRẮC NGHIỆM ĐÚNG SAI

Câu 14: Chức năng chính của một router là gì?

- a) Kết nối các mạng con lại với nhau
- b) Tăng cường tín hiệu mạng
- c) Lọc các gói tin dữ liệu
- d) Cung cấp địa chỉ IP cho các thiết bị

Câu 15: Điểm khác biệt chính giữa repeater và bộ phát Wi-Fi là gì?

- a) Bộ phát Wi-Fi thường có tốc độ truyền dữ liệu cao hơn.
- b) Bộ phát Wi-Fi có thể kết nối với nhiều thiết bị hơn.
- c) Repeater không có khả năng tạo ra một mạng Wi-Fi mới.
- d) Repeater hoạt động dựa vào tín hiệu của bộ phát chính.

BÀI 23: ĐƯỜNG TRUYỀN MẠNG VÀ ỨNG DỤNG

A. LÝ THUYẾT

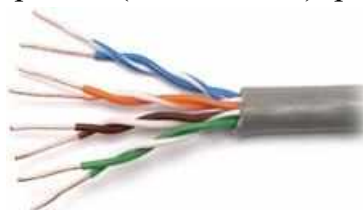
1. Đường truyền có dây:

- Có nhiều loại cáp truyền tín hiệu như cáp đồng truyền tín hiệu điện và cáp quang truyền tín hiệu ánh sáng.
- Cáp đồng cũng có nhiều loại như:
 - + Cáp đồng trục (Coaxial): được dùng phổ biến vào trước những năm 2000.



Hình 23.1. Cáp đồng trục

- + Cáp đôi dây xoắn gọi tắt là cáp xoắn (Twisted Pair): phổ biến nhất hiện nay.



Hình 23.2. Cáp xoắn không bọc (Unshielded Twisted Pair - UTP)

a) Cáp xoắn

- Cáp xoắn có bốn đôi dây xoắn với nhau (mỗi đôi dây được đánh dấu bằng một màu) giúp hạn chế ảnh hưởng của nhiễu từ môi trường xung quanh, giữ cho tín hiệu truyền qua cáp được ổn định và không bị biến đổi do nhiễu.
- Cáp sử dụng đầu nối và cổng RJ45.
- Cáp xoắn được dùng trong hầu hết các mạng cục bộ hiện nay.
- Chất lượng cáp xoắn khác nhau, dẫn đến hiệu suất truyền dữ liệu và chi phí cũng khác nhau.
- Người ta chia cáp thành một số loại (Category, viết tắt là CAT) theo các chuẩn truyền dữ liệu với các thông số về băng thông và khoảng cách truyền hiệu quả.

Ví dụ: CAT.4, CAT.5, CAT.6,...

- Việc lựa chọn sử dụng loại cáp nào cần phụ thuộc vào các tiêu chuẩn truyền thông trong mạng Ethernet.

Chuẩn	Loại cáp	Tốc độ	Khoảng cách truyền hiệu quả	Tên
10BASE5	Đồng trục	10 Mb/s	500 m	Thick Ethernet
10BASE2	Đồng trục	10 Mb/s	185 m	Thin Ethernet
10BASE-T	UTP	10 Mb/s	100 m	Ethernet Twist Pair
100BASE-TX	UTP	100 Mb/s	250 m	Fast Ethernet
100BASE-FX	Cáp quang	100 Mb/s	2000 m	
1000Base-SX	Cáp quang đa mode	Cỡ Gb/s	550 m	Gigabit Ethernet
1000BASE -LX	Cáp quang đơn hoặc đa mode	Cỡ Gb/s	550 m cáp đa mode 5000 m cáp đơn mode	
1000Base-CX	Cáp xoắn bọc kim STP	1-10 Gb/s	25 m	
1000Base-T	Cáp xoắn CAT 6.	1-10 Gb/s	100 m (CAT5, CAT6)	

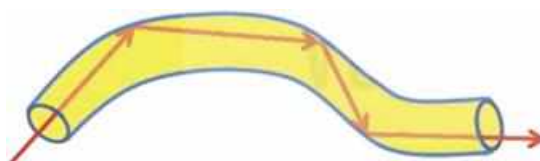
Bảng 23.1. Các chuẩn truyền thông trong mạng Ethernet

“Tốc độ truyền thông thường được đo bằng Megabit/s (Mb/s) hay Gigabit/s (Gb/s).”

Với cáp xoắn, khoảng cách truyền hiệu quả chỉ khoảng 100 m. Nếu cần mở rộng mạng, người ta có thể dùng các thiết bị chuyển tiếp dữ liệu như repeater hay switch để kéo dài thêm đường truyền, tuy nhiên có những giới hạn kỹ thuật về số lần chuyển tiếp.

b) Cáp quang

- Cáp quang là một phương tiện rẻ tiền, có thể truyền xa.
- Cáp quang là một ống sợi thủy tinh hay nhựa có đường kính rất nhỏ, mặt trong phản xạ toàn phần. Ngay cả khi cáp bị uốn cong, ánh sáng vẫn đi được thông suốt do tia sáng phản xạ trong lòng ống.



Hình 23.3. Ánh sáng truyền trong cáp quang

- Có hai loại cáp quang:
 - + Cáp đa mode (multimode).
 - + Cáp quang đơn mode (single mode – có đường kính rất nhỏ, truyền xa hơn).
- Cáp quang có nhiều ưu điểm như:

- + Tín hiệu truyền dẫn ổn định, không bị ảnh hưởng bởi nhiễu điện từ và các điều kiện ngoại cảnh khác.
 - + Do tín hiệu trong cáp quang ít suy hao hơn tín hiệu trong cáp điện nên có thể truyền xa hơn.
 - + Băng thông lớn gấp hàng trăm lần so với cách truyền thuê bao số trên mạng điện thoại ADSL.
 - + Gọn nhẹ, tiêu thụ năng lượng rất thấp và chi phí rẻ hơn so với cáp đồng.
 - + Bảo mật vì khó lấy tín hiệu trên đường truyền.
 - Cáp quang được dùng rất rộng rãi, đặc biệt với khoảng cách lớn:
 - + Việt Nam tham gia nhiều dự án cáp quang biển, nối các quốc gia với nhau.
 - + Đường truyền dữ liệu quốc gia, kết nối các vùng miền, các tỉnh chủ yếu là cáp quang.
 - + Các đường truyền dẫn tín hiệu Internet đến các gia đình trước đây dùng cáp ADSL trên mạng điện thoại hầu hết đã được thay thế bởi cáp quang.
 - Tuy nhiên, các máy tính trong mạng cục bộ sử dụng tín hiệu điện nên không thể dùng cáp quang thay thế cho cáp xoắn.
 - + Người ta chỉ dùng cáp quang trong một số trường hợp như:
 - + Nối các khu vực cách xa nhau của mạng cục bộ.
- Ví dụ: Các toà nhà trong một trường đại học, các khoa trong một bệnh viện,...
- + Kết nối trực tiếp các máy chủ với các dàn đĩa mạng (Network Attached Storage - NAS) để đảm bảo băng thông cao và ổn định.
 - + Khi dùng cáp quang, cần phải sử dụng các bộ chuyển đổi tín hiệu quang thành tín hiệu điện và ngược lại.

2. Đường truyền không dây:

- Việc truyền dữ liệu không dây ngày càng phổ biến.
- Lợi ích rất rõ ràng của phương thức này là không cần nối cáp nhưng đòi hỏi các thiết bị tham gia mạng phải có khả năng thu phát không dây.
- Truyền không dây mã hoá dữ liệu trên sóng vô tuyến điện tần số cao.
- Một số loại hình mạng không dây thông dụng:
 - + Mạng vệ tinh.
 - + Mạng thông tin di động toàn cầu GSM.
 - + Mạng Wi-Fi.
 - + Mạng sử dụng Bluetooth.
 - + Mạng sử dụng kết nối trường gần NFC.

a) Mạng vệ tinh

- *Cách hoạt động:* Các vệ tinh có những bộ thu phát tín hiệu. Các phương tiện dưới mặt đất có thể sử dụng các anten vệ tinh hoặc các cảm biến để thu tín hiệu vệ tinh.
- *Ưu điểm:* Vùng phủ sóng của mạng vệ tinh rất rộng.
- *Ứng dụng:*
 - + Hệ thống định vị toàn cầu gồm các vệ tinh liên tục phát sóng xuống mặt đất giúp các thiết bị định vị xác định tọa độ đã mang lại nhiều ứng dụng hữu ích.

Ví dụ: Việc tìm đường ngày nay rất dễ dàng. Nhiều phương tiện có thể tự lái nhờ được dẫn đường tự động qua hệ thống định vị toàn cầu.

- + Vệ tinh được sử dụng để kết nối Internet.

Ví dụ: Dự án Starlink dùng hàng nghìn vệ tinh quỹ đạo thấp kết nối Internet qua các trạm thu phát trên mặt đất đã trở thành hiện thực. Starlink mang lại cơ hội sử dụng Internet với chi phí thấp cho những nơi chưa có điều kiện thi công cáp như ở sa mạc, rừng sâu, đỉnh núi cao,...



Hình 23.5. Một anten mặt đất của Starlink

b) Mạng thông tin di động toàn cầu GSM

- Cách hoạt động:

- + Mỗi mạng thông tin di động toàn cầu (Global System for Mobile Communications – GSM) có nhiều trạm thu phát (Base Transceiver Station – BTS).
- + Các thiết bị di động sẽ tìm kiếm và kết nối với trạm thu phát sóng gần nhất để kết nối vào mạng.
- + Các trạm BTS sẽ chuyển tiếp tín hiệu cho nhau để chuyển dữ liệu giữa các thiết bị đầu cuối.



Hình 23.6. Các trạm BTS

- Mạng GSM phát triển qua nhiều thế hệ:

- + Mạng thế hệ thứ hai (2G): Chỉ cung cấp được dịch vụ nghe gọi và tin nhắn ngắn SMS.
 - + Mạng thế hệ thứ ba (3G):
 - + Có thể truyền dữ liệu số nói chung, cho phép gửi thư điện tử, truy cập Internet.
 - + Sử dụng các dịch vụ định vị toàn cầu.
 - + Truyền, nhận dữ liệu âm thanh, hình ảnh chất lượng cao cho cả thuê bao cố định và thuê bao đang di chuyển.
 - + Tốc độ trao đổi dữ liệu có thể đạt tới 40 Mb/s.
 - + Mạng thế hệ thứ tư (4G): Có thể truyền tải dữ liệu với tốc độ tối đa tới 1,5 Gb/s.
 - + Mạng thế hệ thứ năm (5G): Đạt tốc độ tới 10 Gb/s, độ trễ rất thấp, có thể hỗ trợ số lượng thiết bị kết nối lớn hơn rất nhiều so với mạng 4G nên rất phù hợp với các ứng dụng IoT.
- Tác động:* Mạng thông tin di động toàn cầu đã mở đường cho Internet di động, đưa Internet đến từng người dân qua thiết bị di động. Có thể nói GSM đã thúc đẩy tin học hoá xã hội lên một mức rất cao.

Ví dụ: Người ta có thể tương tác với nhau và sử dụng các hệ thống thông tin toàn cầu từ những thiết bị di động rất gọn nhẹ.

c) Mạng Wi-Fi

- Điểm truy cập không dây:

- + Rất phổ biến, cũng được gọi là bộ thu phát Wi-Fi.
- + Cho phép kết nối vào mạng cục bộ hay Internet rất đơn giản, giảm thiểu việc dùng các thiết bị kết nối và lắp đặt dây cáp với điều kiện thiết bị đầu cuối phải có khả năng kết nối Wi-Fi:
- + Các máy tính xách tay, máy tính bảng, điện thoại thông minh đã sẵn có khả năng này.
- + Máy tính để bàn muốn kết nối Wi-Fi thì cần lắp thêm mô đun Wi-Fi dưới dạng một bảng mạch mở rộng.

- Sóng Wi-Fi phổ biến dùng tần số cao như 2.4 GHz, 5 GHz và 60 GHz giúp chuyển tải dữ liệu nhanh.

- Có nhiều chuẩn trong họ giao thức IEEE 802.11 (chuẩn kết nối Wi-Fi) sử dụng tần số khác nhau với tốc độ truyền dữ liệu khác nhau. Khi thiết kế mạng có Wi-Fi chúng ta cần tính đến chuẩn nào sẽ được sử dụng.

Ví dụ:

- + Chuẩn 802.11b có tốc độ 11 Mb/s.
- + Chuẩn 802.11a và 802.11g có tốc độ 54 Mb/s.
- + Chuẩn 802.11n có tốc độ 450 Mb/s.
- + Chuẩn 802.11ac tốc độ tối đa đạt đến 1,3 Gb/s.
- + Chuẩn 802.11ad sử dụng dải tần 60 GHz có thể cho tốc độ tối đa tới 4,6 Gb/s.

d) Bluetooth

- Bluetooth là một loại hình mạng có tốc độ khoảng 1 Mb/s trong một phạm vi bán kính khoảng 10 m, rất thích hợp để kết nối các thiết bị cá nhân và đồ gia dụng.

- Bluetooth thường chỉ kết nối hai thiết bị với nhau.

- Bluetooth tiện lợi hơn rất nhiều so với cách kết nối qua cáp.

- Một số ví dụ sử dụng kết nối bluetooth:

- + Kết nối máy tính hay điện thoại di động với loa hay tai nghe không dây.
- + Truyền dữ liệu giữa các máy tính cá nhân hay điện thoại di động.
- + Kết nối không dây máy tính với thiết bị ngoại vi như chuột, bàn phím và máy in.
- + Thay thế các giao tiếp nối tiếp dùng dây cáp truyền thống giữa các thiết bị đo, thiết bị định vị dùng GPS, thiết bị y tế, máy quét mã vạch,...

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Đường truyền mạng là

A. Dây dẫn tín hiệu điện hoặc quang.

B. Môi trường truyền dẫn sóng điện từ (ánh sáng cũng là một dạng sóng điện từ).

C. Một môi trường vật lý được sử dụng cùng với một cách mã hóa dữ liệu trên sóng điện từ để chuyển tải tín hiệu số.

D. Phương tiện truyền dữ liệu.

Câu 2: Phương án nào sau đây **không** phải là đường truyền mạng thông dụng?

A. Sóng mang GSM.

C. Sóng siêu âm.

B. Sóng mang Wi-Fi.

D. Sóng mang Bluetooth.

Câu 3: Chuẩn truyền không dây nào sau đây có tốc độ Gigabit?

A. Chuẩn 802.11b.

C. Chuẩn 802.11n.

B. Chuẩn 802.11g.

D. Chuẩn 802.11ac.

Câu 4: Loại cáp đồng nào được sử dụng phổ biến nhất hiện nay trong mạng cục bộ?

- A. Cáp đồng trục (Coaxial)
- B. Cáp đôi dây xoắn (Twisted Pair)
- C. Cả hai loại cáp trên
- D. Không có loại nào

Câu 5: Cáp xoắn có bao nhiêu đôi dây?

- A. 2
- B. 4
- C. 6
- D. 8

Câu 6: Đầu nối nào được sử dụng cho cáp xoắn?

- A. RJ11
- B. RJ45
- C. USB
- D. HDMI

Câu 7: AT là viết tắt của từ gì và dùng để chỉ điều gì?

- A. Cable
- B. Category
- C. Connection
- D. Communication

Câu 8: Mạng không dây nào sử dụng vệ tinh để phát và thu tín hiệu?

- A. Mạng Wi-Fi
- B. Mạng Bluetooth
- C. Mạng vệ tinh
- D. Mạng GSM

Câu 9: Tên viết tắt GSM là gì?

- A. Global System Mobile
- B. Global System of Mobile communication
- C. Global System for Mobile communication
- D. Global System for Mobile

Câu 10: Mạng GSM 2G chỉ hỗ trợ dịch vụ nào?

- A. Nghe gọi và SMS
- B. Truy cập internet
- C. Gửi email
- D. Định vị toàn cầu

Câu 11: Vì sao cáp đồng được sử dụng phổ biến trong mạng cục bộ?

- A. Giá thành rẻ, dễ dàng lắp đặt và bảo trì.
- B. Tốc độ truyền dẫn cao, đáp ứng mọi nhu cầu sử dụng.
- C. Khoảng cách truyền dẫn xa, phù hợp với mọi loại địa hình.
- D. Khả năng chống nhiễu tốt, đảm bảo tín hiệu ổn định.

Câu 12: Loại cáp xoắn nào sau đây có tốc độ truyền dẫn cao nhất?

- A. CAT.4
- B. CAT.5
- C. CAT.6
- D. CAT.7

Câu 13: Ưu điểm chính của mạng vệ tinh là gì?

- A. Tốc độ cao
- B. Vùng phủ sóng rộng
- C. Giá thành rẻ
- D. Dễ dàng lắp đặt

Câu 14: Chuẩn Wi-Fi nào có tốc độ truyền dữ liệu cao nhất?

- A. 802.11b
- B. 802.11g
- C. 802.11ac
- D. 802.11ad

TRẮC NGHIỆM ĐÚNG SAI

Câu 15: Những chuẩn sau đây là thuộc loại Gigabit Ethernet cho khoảng cách truyền trên 500 m

- a) 1000Base-SX.
- b) 1000BASE-LX.
- c) 1000Base-CX.
- d) 1000Base-T.

Câu 16: Bluetooth - công nghệ kết nối không dây tiện lợi.

- a) Bluetooth là một loại mạng không dây có tốc độ truyền dữ liệu chậm và phạm vi hoạt động rộng.
- b) Bluetooth thường được sử dụng để kết nối các thiết bị cá nhân và đồ gia dụng.
- c) Kết nối Bluetooth tiện lợi hơn so với kết nối qua cáp.
- d) Bluetooth không thể được sử dụng để truyền dữ liệu giữa máy tính cá nhân hoặc điện thoại.

BÀI 24: SƠ BỘ VỀ THIẾT KẾ MẠNG

A. LÝ THUYẾT

Việc thiết kế mạng được thực hiện qua các bước sau:

1. Khảo sát và phân tích hiện trạng, nhu cầu ứng dụng, đặc điểm nơi triển khai mạng.
2. Thiết kế logic, xác định cấu trúc kết nối, mô hình tương tác trong mạng, giao thức mạng được sử dụng.
3. Thiết kế kỹ thuật (mức vật lý), chọn chủng loại thiết bị theo cấu trúc kết nối và chọn điểm đặt thiết bị, xác định tính năng của thiết bị và cáp nối.
4. Lựa chọn hệ điều hành mạng. Lưu ý rằng việc xác định các ứng dụng sẽ cài đặt trên mạng không phải là công việc thiết kế mạng, nhưng có ảnh hưởng tới việc lựa chọn mô hình mạng và tính năng của các thiết bị mạng để đáp ứng yêu cầu của các ứng dụng.

CHỦ ĐỀ 7. GIẢI QUYẾT VẤN ĐỀ VỚI SỰ TRỢ GIÚP CỦA MÁY TÍNH

BÀI 25: LÀM QUEN VỚI HỌC MÁY

A. LÝ THUYẾT

1. Tìm hiểu sơ lược về Học máy:

- Lịch sử Học máy bắt nguồn từ công trình *Máy tính và thông minh* của Alan Turing
- *Khái niệm*: Học máy là một lĩnh vực của AI tập trung vào việc phát triển các thuật toán và mô hình cho phép máy tính tự học và cải thiện từ dữ liệu để đưa ra dự đoán hoặc quyết định dựa trên dữ liệu mà không cần lập trình rõ ràng.
- Hai điểm mấu chốt trong khái niệm Học máy: “máy tính tự học từ dữ liệu” và “không cần lập trình rõ ràng”.
- Một số ví dụ của Học máy:
 - + Dự báo thời tiết: Máy tính sử dụng dữ liệu thời tiết lịch sử để học cách dự báo thời tiết mà không cần lập trình cụ thể các quy tắc dự báo.
 - + Phân tích ngôn ngữ tự nhiên (NLP): Máy tính tự học cách dịch ngôn ngữ, phân tích cảm xúc và xử lý ngôn ngữ tự nhiên từ dữ liệu văn bản lớn.
 - + Chẩn đoán y khoa: Hệ thống học máy phân tích dữ liệu lâm sàng và hình ảnh y tế để hỗ trợ chẩn đoán và phát hiện sớm bệnh tật.
 - + Tối ưu hoá chuỗi cung ứng: Máy tính học cách tối ưu hoá quản lý kho và dự báo nhu cầu sản phẩm từ dữ liệu lịch sử và xu hướng hiện tại.
 - + Nhận dạng giọng nói: Hệ thống học máy nhận dạng và chuyển đổi giọng nói thành văn bản dựa trên hàng ngàn giờ ghi âm.

⇒ Những ứng dụng này cho thấy sự linh hoạt và mạnh mẽ của Học máy trong việc tự học và thích ứng từ dữ liệu đa dạng.

- *Quy trình Học máy:*

Việc xây dựng các ứng dụng Học máy có thể chia thành 5 bước cơ bản như sau:



Hình 25.2. Quy trình Học máy

+ **Thu thập dữ liệu:** Tùy theo bài toán cần giải quyết, dữ liệu cần thiết để xây dựng mô hình Học máy có thể được lấy từ nhiều nguồn khác nhau, như các cơ sở dữ liệu, tệp tin hoặc thậm chí thông qua việc ghi chép trực tiếp.

+ **Chuẩn bị dữ liệu:** Thông thường, dữ liệu thu thập được không phù hợp để có thể sử dụng được ngay. Do vậy, cần thực hiện các thao tác chuẩn bị dữ liệu (còn được gọi là “làm sạch dữ liệu”) bao gồm:

- Loại bỏ dữ liệu nhiễu.
- Bổ sung các giá trị thiếu.
- Chuyển đổi dữ liệu sang định dạng phù hợp.
- Giảm kích thước dữ liệu (nếu cần).

⇒ **Lưu ý:**

✓ Đây là hai bước quan trọng, chiếm nhiều thời gian và công sức nhất của quá trình xây dựng ứng dụng Học máy.

✓ Hai bước này có thể phải thực hiện lặp đi lặp lại cho tới khi thu được bộ dữ liệu như mong muốn.

✓ Tập dữ liệu thu được thường được chia thành hai phần:

- ✓ Dữ liệu huấn luyện (thường chiếm khoảng 70% đến 80%) được dùng để huấn luyện mô hình.
- ✓ Dữ liệu kiểm thử được dùng để đánh giá mô hình.

+ Chọn thuật toán học máy phù hợp với loại bài toán và dữ liệu thu thập được:

- Các loại thuật toán này khá đa dạng như hồi quy tuyến tính, cây quyết định, mạng nơron,...
- Về mặt bản chất, thuật toán Học máy sử dụng các mô hình toán học để kết nối các đặc trưng và thông tin liên quan tới tập dữ liệu.

+ **Huấn luyện mô hình:**

• Thực hiện thuật toán học máy trên tập dữ liệu huấn luyện, giúp máy tính học cách phân biệt giữa các mẫu thuộc các lớp dữ liệu khác nhau.

- Kết quả của quá trình này sẽ là một mô hình Học máy để giải quyết một bài toán cụ thể.

+ **Đánh giá mô hình:**

• Áp dụng mô hình Học máy trên tập dữ liệu kiểm thử để đánh giá hiệu suất của mô hình trong việc dự đoán dữ liệu mới.

• Dựa trên kết quả đánh giá, mô hình có thể cần được cải thiện, bằng cách bổ sung thêm dữ liệu huấn luyện mới, điều chỉnh các tham số của thuật toán Học máy hoặc sử dụng các thuật toán Học máy khác.

⇒**Lưu ý:** Hai bước huấn luyện và đánh giá có thể được thực hiện lặp đi lặp lại cho tới khi thu được mô hình Học máy như mong muốn. Cuối cùng, sử dụng mô hình thu được để giải quyết vấn đề đặt ra, thực hiện dự đoán hay phân cụm trên dữ liệu mới.

2. Phân loại và vai trò của Học máy trong thực tế

a) Phân loại Học máy

- Trong Học máy, tập dữ liệu đầu vào gồm hai loại chính:

+ Dữ liệu có nhãn.

+ Dữ liệu không có nhãn.

- Dữ liệu được gắn kết với một nhãn hoặc một giá trị đích cụ thể được gọi là dữ liệu có nhãn, trường hợp ngược lại, là dữ liệu không có nhãn:

+ Nhãn hoặc giá trị đích của dữ liệu thường chỉ ra thông tin quan trọng về đối tượng, thuộc tính hoặc phân loại mà mẫu dữ liệu đó đại diện.

+ Việc gán nhãn dữ liệu thường được thực hiện bằng cách thủ công.

+ Dữ liệu có nhãn đóng vai trò rất quan trọng trong quá trình huấn luyện mô hình học máy, vì nó cung cấp thông tin cần thiết cho mô hình để học và đưa ra dự đoán chính xác trên các dữ liệu mới. Việc có dữ liệu được gán nhãn đúng và đa dạng là một yếu tố quyết định để xây dựng mô hình học máy hiệu quả và đáng tin cậy.

- Tương ứng với hai loại dữ liệu đầu vào là hai phương pháp học máy cơ bản:

+ *Học có giám sát:* Được sử dụng để giải quyết bài toán phân loại dữ liệu.

+ *Học không giám sát:* Được sử dụng để giải quyết bài toán phân cụm dữ liệu.

- *Học có giám sát*

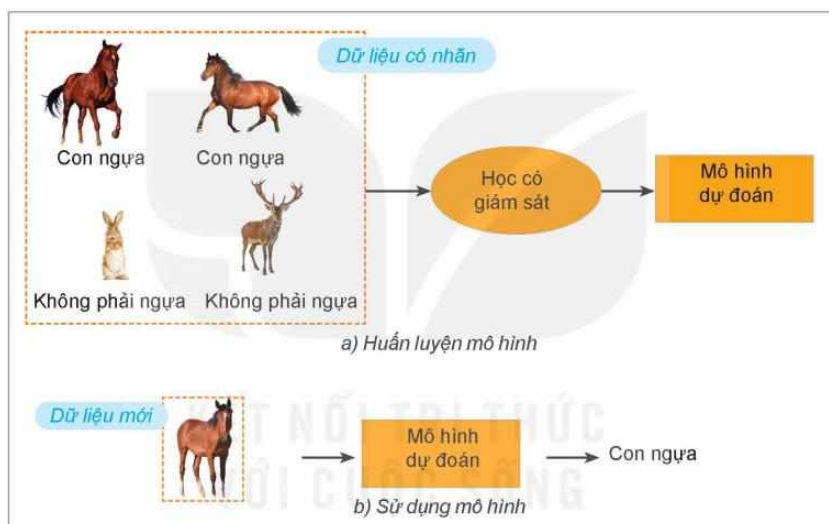
+ Khái niệm: Là phương pháp học máy trong đó tập dữ liệu đầu vào là dữ liệu đã được gán nhãn.

+ Cách hoạt động:

• Trên cơ sở được “học” từ dữ liệu có nhãn, máy tính có khả năng mô hình hoá mối quan hệ giữa dữ liệu đầu vào với đầu ra tương ứng (pha huấn luyện mô hình).

• Khi đưa một dữ liệu mới chưa biết vào, máy tính sẽ thực hiện việc xác định các đặc trưng dữ liệu, từ đó đưa ra phản hồi (dự đoán) dữ liệu đó cùng loại với dữ liệu nào được gán nhãn (pha sử dụng mô hình).

Ví dụ: Hình 25.3 mô tả một hệ thống học có giám sát, với dữ liệu là các hình ảnh được gán nhãn (Con ngựa hay Không phải ngựa), để xác định xem dữ liệu mới được đưa vào là một con ngựa hay là một loại động vật khác.



Hình 25.3. Hai pha của mô hình học có giám sát (phân loại dữ liệu)

+ **Ứng dụng:** Học có giám sát là phương pháp học máy được sử dụng rộng rãi nhất. Nó có nhiều ứng dụng trong thực tế như:

- Xây dựng bộ lọc thư rác.
- Nhận dạng hình ảnh.
- Nhận dạng chữ viết tay.
- Nhận dạng tiếng nói.

- *Học không giám sát*

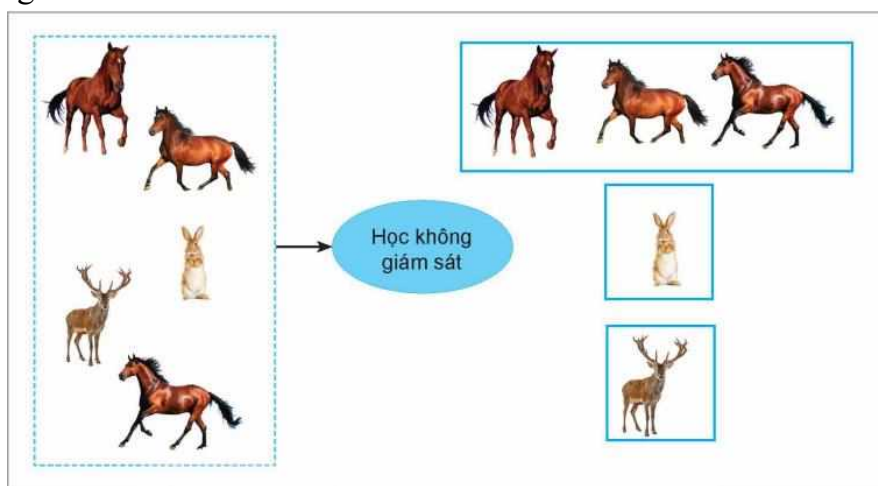
+ **Khái niệm:** Là phương pháp học máy sử dụng dữ liệu không có nhãn.

+ **Cách hoạt động:** Sử dụng thông tin về mối quan hệ tương tự hay khác biệt, cũng như dựa trên xác suất đồng xuất hiện của các đối tượng hoặc các biến có trong dữ liệu, các thuật toán và mô hình học trong phương pháp này sẽ thực hiện việc mô hình hoá cấu trúc hoặc mô tả các thông tin ẩn chứa trong dữ liệu.

+ **Ứng dụng:**

- Thường được ứng dụng để phân chia dữ liệu thành các nhóm dựa trên sự tương đồng của các mẫu dữ liệu.

Ví dụ: trong Hình 25.4, mô hình học không giám sát thực hiện việc phân nhóm các con vật dựa trên hình ảnh của chúng.



Hình 25.4. Mô hình học không giám sát (phân cụm dữ liệu)

- Một số bài toán khác có thể áp dụng học không giám sát:
- Xác định các phân khúc khách hàng dựa trên lịch sử mua hàng của họ.
- Phát hiện bất thường trong các giao dịch thẻ tín dụng để xác định gian lận.
- Xác định các chủ đề khác nhau hoặc xác định chủ đề chính được thảo luận trong một tập hợp các bài báo.

b) *Vai trò của Học máy*

- Học máy có vai trò quan trọng trong nhiều công việc và ứng dụng thực tế. Nó hỗ trợ khai phá các loại dữ liệu đa dạng, có quy mô lớn, bao gồm cả các dữ liệu không ngừng thay đổi theo thời gian, để trích xuất được những thông tin và tri thức hữu ích.

- Một số ví dụ về vai trò của Học máy:

+ *Lọc thư rác:*

- Học máy giúp xây dựng mô hình có khả năng phân loại thư điện tử là thư rác hoặc thư thường dựa trên các đặc điểm của thư gửi tới, như từ khoá, cấu trúc thư và nhiều yếu tố khác.
- Học máy giúp giảm thời gian và công sức của người dùng trong việc đánh dấu thư rác, đồng thời cải thiện hiệu suất lọc thư theo thời gian bằng cách học hỏi từ dữ liệu và cập nhật mô hình.

+ *Chẩn đoán bệnh:*

- Học máy sử dụng dữ liệu về tình trạng sức khỏe của bệnh nhân cùng kết quả xét nghiệm và các cơ sở dữ liệu bệnh lý khác để xây dựng mô hình chẩn đoán bệnh. Mô hình này còn có thể dự báo tình trạng sức khỏe và đề xuất phương án điều trị phù hợp cho bệnh nhân.

- Mô hình Học máy có thể học từ hàng ngàn lần chẩn đoán cho nhiều bệnh nhân khác nhau trước đó, giúp bác sĩ đưa ra quyết định dựa trên dữ liệu một cách chính xác và nhanh chóng hơn.

+ *Phân tích thị trường:*

- Học máy có thể phân tích dữ liệu thị trường từ nhiều nguồn khác nhau để xác định xu hướng, dự báo biến động giá cả, trợ giúp hình thành các chiến lược kinh doanh dựa trên các mô hình dự đoán.

- Học máy giúp người đầu tư và nhà kinh doanh hiểu rõ hơn về thị trường, tăng khả năng đưa ra quyết định đầu tư dựa trên thông tin và các phân tích kỹ thuật.

+ *Nhận dạng tiếng nói:* Học máy giúp xây dựng các mô hình âm thanh để biểu diễn những đặc trưng của tiếng nói, giúp máy tính có thể học và nhận dạng các biểu hiện âm thanh của từng đơn vị tiếng (phoneme), từ đó tạo ra biểu diễn số hoá của âm thanh. Những đặc điểm âm thanh cá nhân trong các mô hình âm thanh còn giúp cải thiện khả năng nhận dạng và phân biệt tiếng nói của những người nói khác nhau.

+ *Nhận dạng chữ viết:*

- Học máy giúp xây dựng mô hình hình học cho phép xác định hình dạng, kích thước, góc xoay của các kí tự trong hình ảnh chữ viết tay.

- Những năm gần đây, sự phát triển của học sâu (một lĩnh vực của Học máy) cho phép học và trích xuất các đặc trưng phức tạp từ hình ảnh chữ viết tay, giúp cải thiện đáng kể khả năng nhận dạng chữ viết tay.

+ *Dịch tự động:*

- Học máy sử dụng dữ liệu về bản dịch và bản gốc trong các ngôn ngữ khác nhau để xây dựng mô hình dịch tự động. Mô hình này có khả năng dịch văn bản, tiếng nói từ ngôn ngữ này sang ngôn ngữ khác.

- Khả năng dịch tự động của máy tính giúp hạn chế rào cản ngôn ngữ trong giao tiếp, phát triển hợp tác và trao đổi thông tin mọi lĩnh vực, đặc biệt trong giáo dục, đào tạo và nghiên cứu khoa học.

Nhận xét: Trong các công việc trên, cũng như trong nhiều lĩnh vực khác, vai trò quan trọng của Học máy được thể hiện ở nhiều góc độ khác nhau:

- Giúp xử lý một lượng lớn dữ liệu trong thời gian thực một cách nhanh chóng và hiệu quả để xác định các mẫu và xu hướng quan trọng có trong dữ liệu.

- Tự động hoá các nhiệm vụ phức tạp mà trước đây cần sự can thiệp của con người.

- Do có khả năng học từ dữ liệu, Học máy có thể giúp các chuyên gia và các nhà nghiên cứu từng bước xây dựng và bổ sung tri thức.

- Nhờ khả năng không ngừng bổ sung dữ liệu và tự động cập nhật mô hình đã được huấn luyện, Học máy ngày càng có vai trò không thể thiếu trong các ứng dụng mà dữ liệu có quy mô và chủng loại đa dạng, không ngừng thay đổi theo thời gian, như sự xuất hiện các mẫu thư rác mới, các triệu chứng bệnh mới, hay các bản dịch ngôn ngữ mới,...

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Trong Học máy, dữ liệu có nhãn là

A. dữ liệu được gắn kết với một nhãn hoặc một giá trị đích cụ thể.

B. dữ liệu không liên quan đến mục tiêu học của máy tính.

- C. dữ liệu được thu thập từ nguồn ngoại vi.
- D. dữ liệu ngẫu nhiên không có hướng dẫn cụ thể.

Câu 2: Trong Học máy, dữ liệu có nhãn được sử dụng cho loại Học máy nào?

- A. Học có giám sát.
- B. Học không giám sát.
- C. Học tăng cường.
- D. Học máy kết hợp.

Câu 3: Khi sử dụng Học máy có giám sát, điều quan trọng nhất là:

- A. Cung cấp lượng lớn dữ liệu.
- B. Huấn luyện mô hình với dữ liệu có nhãn.
- C. Sử dụng thuật toán phức tạp.
- D. Thực hiện đánh giá mô hình thường xuyên.

Câu 4: Học máy là một lĩnh vực của:

- A. Trí tuệ nhân tạo
- B. Khoa học dữ liệu
- C. Mạng máy tính
- D. Phần mềm quản lý

Câu 5: Học máy là:

- A. Chương trình máy tính có khả năng đưa ra quyết định hay dự đoán dựa trên dữ liệu.
- B. Khả năng máy tính phân tích dữ liệu thu nhận được để đưa ra dự đoán hoặc quyết định dựa trên các quy tắc được xác định rõ ràng.
- C. Việc sử dụng các phương pháp và kỹ thuật cho phép máy tính học từ dữ liệu để đưa ra dự đoán hoặc quyết định mà không cần lập trình cụ thể.
- D. Chương trình máy tính có khả năng tự cải thiện hiệu suất thực hiện nhiệm vụ thông qua việc cập nhật các dữ liệu mới sau khi hoàn thành nhiệm vụ đó nhiều lần.

Câu 6: Mục đích của việc chuẩn bị dữ liệu trong Học máy là

- A. để tăng dung lượng dữ liệu có sẵn cho máy tính.
- B. để tạo ra dữ liệu mới từ nguồn dữ liệu hiện có.
- C. để loại bỏ dữ liệu nhiễu, bổ sung giá trị thiếu và chuyển đổi dữ liệu sang định dạng phù hợp.
- D. để máy tính có thể học dữ liệu một cách nhanh chóng.

Câu 7: Sử dụng mô hình Học máy trong thực tế phục vụ mục đích gì?

- A. Tạo ra dữ liệu mới không liên quan đến dữ liệu hiện có.
- B. Dự đoán hoặc phân cụm dữ liệu một cách ngẫu nhiên.
- C. Thực hiện dự đoán hoặc phân cụm trên dữ liệu mới.
- D. Phân tích dữ liệu mà không có mục đích cụ thể.

Câu 8: Trong Học máy có giám sát, đầu ra của mô hình thường được xác định như thế nào?

- A. Dựa trên mối quan hệ giữa dữ liệu đầu vào và nhãn.
- B. Dựa trên sự tương đồng của dữ liệu.
- C. Dựa trên xác suất xuất hiện của dữ liệu.
- D. Dựa trên cấu trúc dữ liệu ẩn.

Câu 9: Học không giám sát trong Học máy dựa trên loại dữ liệu

- A. không có nhãn.
- B. có nhãn.
- C. tổng hợp.
- D. được xác định trước.

Câu 10: Việc huấn luyện mô hình Học máy dựa trên

- A. trích xuất dữ liệu từ nguồn ngoại vi.
- B. dữ liệu huấn luyện được chọn lọc từ tập dữ liệu thu thập được.
- C. tạo ra mô hình toán học không dựa trên dữ liệu cụ thể.
- D. cập nhật thông tin từ dữ liệu mới nhận được liên tục.

Câu 11: Học máy không giám sát thường được sử dụng cho mục đích nào sau đây?

- A. Phân loại dữ liệu.
- B. Phân cụm dữ liệu.
- C. Dự đoán giá trị cụ thể.
- D. Ghi nhãn dữ liệu tự động.

Câu 12: Khi sử dụng Học máy để lọc thư điện tử, máy tính học cách phân loại thư rác và thư hợp lệ như thế nào?

- A. Máy tính được hướng dẫn từng bước cụ thể để nhận diện thư rác.
- B. Bằng cách nhận diện từ khóa cụ thể được lập trình trước trong nội dung thư.
- C. Bằng cách học từ tập dữ liệu ví dụ về thư rác, thư hợp lệ và tự xác định đặc điểm phân biệt chúng.
- D. Máy tính sử dụng các quy tắc cố định được thiết lập bởi người lập trình.

Câu 13: Một trong những ứng dụng của học máy trong y tế là:

- A. Dự đoán thời tiết
- B. Phát hiện bất thường trong dữ liệu sức khỏe
- C. Phân tích dữ liệu mạng
- D. Nhận diện thư rác

Câu 14: Tại sao cần chia dữ liệu Học máy thành hai phần: dữ liệu huấn luyện và dữ liệu kiểm tra?

- A. Để máy tính có thể học từ cả hai loại dữ liệu một cách cân bằng.
- B. Để đánh giá khách quan hiệu suất của mô hình học máy.
- C. Để tăng khối lượng dữ liệu mà máy tính có thể học.
- D. Để có thể dự đoán kết quả của mô hình trên dữ liệu thực tế.

TRẮC NGHIỆM ĐÚNG SAI

Câu 1: Học máy có thể hỗ trợ trong chẩn đoán bệnh

- a) Dự báo tình trạng sức khỏe.
- b) Đề xuất phương án điều trị phù hợp cho bệnh nhân.
- c) Thay thế bác sĩ thăm khám và đưa ra kết quả chẩn đoán.
- d) Hỗ trợ bác sĩ đưa ra quyết định dựa trên dữ liệu một cách chính xác và nhanh chóng hơn.

Câu 2: Việc xây dựng mô hình Học máy thường cần dữ liệu đa dạng vì

- a) Để mô hình có thể học và phân biệt chính xác giữa các mẫu dữ liệu khác nhau.
- b) Để tăng cơ hội đạt được kết quả chính xác trên mọi loại dữ liệu.
- c) Để giảm thiểu nguy cơ lỗi do dữ liệu không đủ.
- d) Để tăng dung lượng bộ nhớ mà mô hình cần xử lý.

BÀI 26: LÀM QUEN VỚI KHOA HỌC DỮ LIỆU

A. LÝ THUYẾT

1. Khái niệm và mục tiêu của khoa học dữ liệu

a) Khái niệm về Khoa học dữ liệu

- Khoa học dữ liệu là một lĩnh vực liên ngành, sử dụng các phương pháp khoa học, quy trình, thuật toán để khám phá tri thức từ dữ liệu, kết hợp những tri thức đó với tri thức chuyên ngành làm cơ sở cho những quyết định.

- Cụ thể, Khoa học dữ liệu sử dụng các phương pháp và công cụ của: khoa học máy tính, toán học và thống kê kết hợp với tri thức chuyên ngành để giúp tổ chức, cá nhân hiểu rõ hơn về dữ liệu mình sở hữu và tận dụng tri thức này để đưa ra những quyết định phù hợp.



Hình 26.1. Khoa học dữ liệu và các lĩnh vực liên quan

Trong đó:

+ *Khoa học máy tính* cung cấp các công cụ và kỹ thuật để xử lý, phân tích và khai phá dữ liệu. Các ngôn ngữ và thư viện lập trình cũng như Học máy, cùng với khả năng xử lý dữ liệu có quy mô và độ phức tạp khác nhau, đóng vai trò quan trọng trong việc phân tích, khai phá dữ liệu và xây dựng các mô hình dự đoán.

+ *Toán học và thống kê* cung cấp cơ sở cho các phương pháp phân tích và khai phá dữ liệu. Các thuật toán thống kê và toán học giúp kiểm tra giả thuyết, tính toán phân phối xác suất, xác định sự tương quan giữa các đại lượng trong dữ liệu,...

+ *Tri thức chuyên ngành* là tri thức của từng lĩnh vực, ví dụ kinh doanh, y tế, khoa học xã hội,... có vai trò quan trọng để hiểu ngữ cảnh và ý nghĩa của dữ liệu. Nó giúp các nhà Khoa học dữ liệu đánh giá được chất lượng và độ chính xác của dữ liệu, diễn giải được kết quả phân tích và khai phá dữ liệu theo cách có ý nghĩa phù hợp với lĩnh vực ứng dụng để đưa ra quyết định đúng đắn.

b) *Mục tiêu của Khoa học dữ liệu*

- *Mục tiêu chính*: Phân tích và khai phá dữ liệu để có được tri thức, vận dụng tri thức đó để giải quyết vấn đề và đưa ra các quyết định phù hợp.

- *Các mục tiêu cụ thể*:

+ *Tổ chức và quản lý dữ liệu*:

- Tập trung vào việc xây dựng, duy trì hệ thống tổ chức dữ liệu một cách khoa học để đảm bảo tính toàn vẹn, sẵn sàng và quản lý hiệu quả các nguồn dữ liệu.

- Đây là nhiệm vụ rất quan trọng để tạo ra cơ sở hạ tầng dữ liệu mạnh mẽ và linh hoạt, hỗ trợ quá trình phân tích và ra quyết định trong lĩnh vực Khoa học dữ liệu.

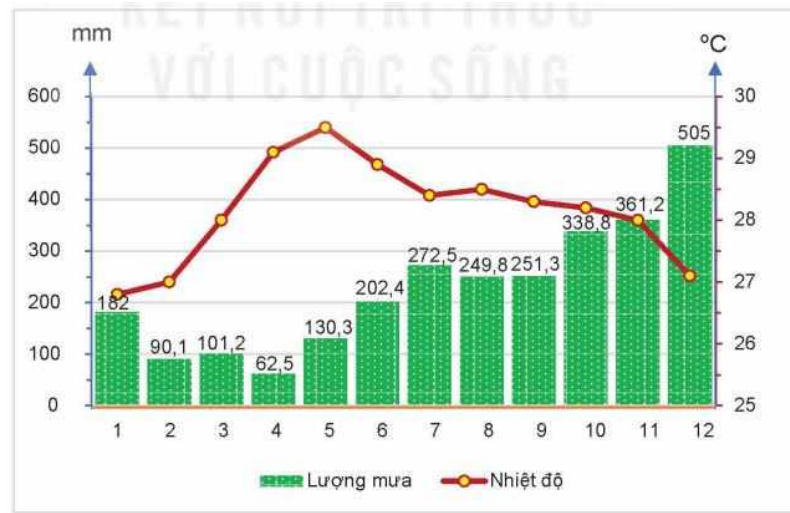
+ *Phân tích dữ liệu*:

- Nhằm hiểu rõ về nội dung, cấu trúc dữ liệu, xác định các đặc điểm quan trọng, nhận diện nhóm và xu hướng trong dữ liệu.

- Giúp tạo ra cái nhìn toàn diện về dữ liệu và hỗ trợ quá trình ra quyết định.

+ *Trực quan hoá dữ liệu* nhằm biểu diễn dữ liệu một cách trực quan, dễ hiểu bằng các sơ đồ, biểu đồ hay hình ảnh, giúp người dùng có được cái nhìn tổng quan về dữ liệu.

Ví dụ: Từ biểu đồ trong Hình 26.2 có thể dễ dàng suy ra được nhiệt độ và biên độ nhiệt cũng như tổng lượng mưa trung bình năm,...



Hình 26.2. Nhiệt độ và lượng mưa trung bình tháng của huyện đảo Trường Sa (Khánh Hoà)
(Số liệu: Trung tâm Thông tin và Dữ liệu khí tượng thuỷ văn)

+ *Tối ưu hoá quyết định* nhằm cải thiện quyết định dựa trên dữ liệu, bao gồm việc sử dụng các thuật toán tối ưu hoá để đưa ra quyết định tốt nhất dựa trên các ràng buộc và mục tiêu.

Ví dụ: Tối ưu hoá quy trình sản xuất để tối ưu hoá hiệu quả của dây chuyền sản xuất hay sản lượng, chất lượng sản phẩm,...

+ *Khám phá tri thức:*

- Để tìm ra các mối quan hệ ẩn chứa trong dữ liệu, xác định nguyên nhân và kết quả, tạo ra tri thức mới từ dữ liệu.

- Đây cũng là mục tiêu cụ thể cao nhất của Khoa học dữ liệu.

Ví dụ:

- Trong nghiên cứu dược phẩm, người ta có thể sử dụng dữ liệu bệnh nhân để tìm hiểu mối quan hệ giữa một loại thuốc và các phản ứng phụ, giúp họ hiểu rõ hơn về tác động của loại thuốc này đối với sức khoẻ của bệnh nhân.

- Nhiều trang web thương mại điện tử sử dụng dữ liệu lịch sử mua sắm của người dùng để dự đoán và đề xuất sản phẩm mà họ có thể quan tâm,...

Khoa học dữ liệu không chỉ nhằm mục đích thu thập và xử lý dữ liệu, mà quan trọng hơn là phân tích, hiểu biết và sử dụng dữ liệu để đưa ra quyết định hợp lý và hiệu quả trong nhiều lĩnh vực khác nhau.

- *Mối quan hệ giữa các mục tiêu:* Các mục tiêu của Khoa học dữ liệu có mối quan hệ chặt chẽ với nhau, hỗ trợ lẫn nhau để đạt được mục tiêu chính của Khoa học dữ liệu.

Ví dụ: Mối quan hệ giữa mục tiêu tổ chức và quản lý dữ liệu với mục tiêu phân tích dữ liệu. Dữ liệu cần được tổ chức và quản lý một cách khoa học thì mới có thể được phân tích một cách hiệu quả.

2. Một số thành tựu của khoa học dữ liệu

- Sự phát triển trong thời gian gần đây của Khoa học dữ liệu cũng như AI và Học máy không tách rời với sự ra đời và phát triển của dữ liệu lớn.

- *Thuật ngữ dữ liệu lớn* (Big Data) xuất hiện từ những năm 1990 và được sử dụng để mô tả quy mô lớn và phức tạp của dữ liệu, cả về khối lượng và sự đa dạng mà không thể xử lý bằng cách sử dụng phương pháp truyền thống.

- *Dữ liệu lớn thường được đặc trưng bởi “năm chữ V” (5V):*

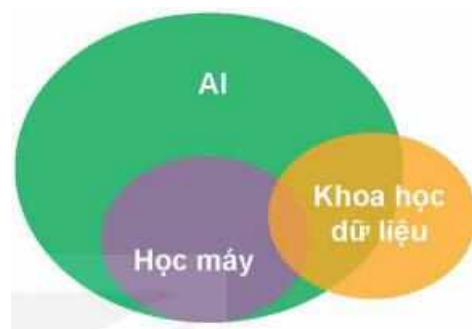
+ Volume (Khối lượng) đề cập đến kích thước lớn của các tập dữ liệu.

- + Velocity (Vận tốc) đề cập đến tốc độ mà dữ liệu đó được tạo ra và cần được phân tích.
- + Variety (Sự đa dạng) đề cập đến nhiều loại dữ liệu khác nhau, có thể ở dạng văn bản, âm thanh, video hoặc các dạng khác.
- + Value (Giá trị) đề cập đến tính hữu ích của dữ liệu.
- + Veracity (Tính xác thực) đề cập đến sự cần thiết phải đảm bảo tính xác thực của dữ liệu do dữ liệu lớn thường có nhiều nhiễu/sai số hoặc không chính xác trong dữ liệu.

- *Mối quan hệ giữa AI, Học máy và Khoa học dữ liệu:*

+ Không thể tận dụng một cách hiệu quả dữ liệu lớn nếu không tự động hoá quy trình xử lý, phân tích và khai phá. Khoa học dữ liệu cùng với AI và Học máy cung cấp các quy trình như vậy.

Việc phân tích và khám phá các tri thức hữu ích từ dữ liệu lớn có thể được coi là thành tựu và lợi ích chung lớn nhất mà Khoa học dữ liệu đem lại.



Hình 26.3. Mối quan hệ giữa AI, Học máy và Khoa học dữ liệu

Có thể thấy đây là mối quan hệ gắn bó và tương hỗ lẫn nhau giữa các lĩnh vực đang phát triển hết sức mạnh mẽ.

Nhiều thành tựu được coi là thành tựu chung của cả ba lĩnh vực, không ít ứng dụng thực tế được mô tả trong các tài liệu khác nhau như là ứng dụng của AI, của Học máy hay của Khoa học dữ liệu tùy theo bối cảnh mà chúng được nhắc tới.

- *Một số thành tựu cụ thể của Khoa học dữ liệu:*

+ Đổi mới quá trình ra quyết định – ra quyết định dựa trên dữ liệu góp phần tăng hiệu quả công việc:

- Thông qua việc phân tích và khai phá dữ liệu thu thập được, Khoa học dữ liệu có thể đưa ra những dự báo và phân tích xu hướng phát triển, từ đó giúp tổ chức, doanh nghiệp sớm chuẩn bị, sẵn sàng thích nghi với những thay đổi và đưa ra các quyết định kinh doanh phù hợp.

- Bằng cách sử dụng thuật toán học máy để phân tích và khai phá dữ liệu lớn về các giao dịch, ngân hàng và tổ chức tài chính có thể xác định những mẫu và điểm bất thường, từ đó xác định hoạt động gian lận, giúp ngăn ngừa tổn thất và cải thiện tính bảo mật tổng thể của hệ thống tài chính.

- Khoa học dữ liệu có khả năng hỗ trợ phân bổ tài nguyên hợp lý nhờ phân tích dữ liệu sử dụng tài nguyên, giúp các tổ chức tối ưu hoá việc phân bổ tài nguyên, giảm các nguy cơ lãng phí.

+ Tự động hoá và thúc đẩy quá trình đổi mới sáng tạo:

- Các mô hình học máy trong Khoa học dữ liệu có thể giúp tự động hoá những tác vụ lặp đi lặp lại và tốn thời gian, cho phép con người tập trung vào những công việc phức tạp và sáng tạo hơn.

Ví dụ: Các mô hình học máy trong Khoa học dữ liệu có thể giúp tự động hoá nhiều quy trình và công việc trong các lĩnh vực sản xuất, hậu cần (logistics), dịch vụ khách hàng, quản lý tài chính, giúp tiết kiệm thời gian và chi phí, tăng tính hiệu quả và độ chính xác.

• Do các công cụ và nền tảng Khoa học dữ liệu ngày càng trở nên dễ tiếp cận hơn, các tổ chức thuộc mọi lĩnh vực đều có thể vận dụng và hưởng lợi nhờ những khả năng của Khoa học dữ liệu.

Khoa học dữ liệu góp phần thúc đẩy quá trình đổi mới sáng tạo, tạo ra nhiều cơ hội mới cho các lĩnh vực khác nhau.

+ *Cá nhân hoá các dịch vụ, cải thiện trải nghiệm khách hàng:*

• Khoa học dữ liệu có thể hỗ trợ việc cung cấp các dịch vụ được cá nhân hoá, dựa trên việc phân tích các dữ liệu được thường xuyên cập nhật về khách hàng, giúp các doanh nghiệp có được những thông tin đầy đủ hơn về nhu cầu, sở thích và hành vi của họ để đưa ra được những giải pháp cải thiện trải nghiệm khách hàng, góp phần gia tăng doanh số.

• Các *hệ khuyến nghị* (còn được gọi là các hệ tư vấn) định hướng cá nhân hoá, được phát triển và ứng dụng rộng rãi để giới thiệu những sản phẩm hoặc nội dung mà khách hàng có thể quan tâm, đang là một trong các giải pháp kinh doanh hiệu quả.

Ví dụ: Trong lĩnh vực y tế, y học cá nhân hoá là một trong những thành tựu đáng lưu ý của Khoa học dữ liệu. Tiếp cận sử dụng Khoa học dữ liệu và Học máy, thông qua việc phân tích và khai phá các bộ dữ liệu lớn về thông tin di truyền và y tế liên quan, cho phép đưa ra phác đồ điều trị phù hợp với từng bệnh nhân, giúp nâng cao hiệu quả và kết quả chăm sóc sức khỏe cộng đồng.

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Khoa học dữ liệu là lĩnh vực kết hợp giữa các phương pháp nào?

- | | |
|--|---|
| A. Khoa học máy tính, toán học và thống kê | C. Lịch sử, xã hội học và nhân học |
| B. Kinh tế học, xã hội học và tâm lý học | D. Lịch sử, xã hội học và khoa học tự nhiên |

Câu 2: Mục tiêu chính của khoa học dữ liệu là gì?

- A. Thu thập dữ liệu
B. Phân tích và khai thác dữ liệu để tạo ra tri thức
C. Phát triển phần mềm quản lý dữ liệu
D. Tổ chức các sự kiện nghiên cứu dữ liệu

Câu 3: Khoa học dữ liệu có thể giúp tổ chức và cá nhân ra quyết định dựa trên:

- | | |
|------------------------|-----------------------------|
| A. Các sự kiện lịch sử | C. Thông tin từ mạng xã hội |
| B. Tri thức từ dữ liệu | D. Các lý thuyết học thuật |

Câu 4: Dữ liệu lớn được đặc trưng bởi bao nhiêu chữ V?

- | | | | |
|------|------|------|------|
| A. 4 | B. 3 | C. 6 | D. 5 |
|------|------|------|------|

Câu 5: Mối quan hệ giữa AI, Học máy và Khoa học dữ liệu là gì?

- | | |
|-------------------------------------|--|
| A. Chúng hoàn toàn độc lập với nhau | C. Chúng có mối quan hệ gắn bó và tương hỗ |
| B. Chúng không liên quan đến nhau | D. Chúng chỉ liên quan trong lĩnh vực y tế |

Câu 6: Khoa học dữ liệu giúp tổ chức và cá nhân làm gì?

- | | |
|--------------------------------|-----------------------------------|
| A. Tạo ra dữ liệu | C. Lưu trữ dữ liệu lâu dài |
| B. Ra quyết định chính xác hơn | D. Chỉ phân tích số liệu thống kê |

Câu 7: "Vận tốc" trong đặc điểm dữ liệu lớn ám chỉ điều gì?

- | | |
|---|--|
| A. Tốc độ tạo ra và cần phân tích dữ liệu | C. Khả năng di chuyển dữ liệu qua mạng |
| B. Sự tăng trưởng của cơ sở dữ liệu | D. Sự đa dạng của dữ liệu |

Câu 8: Một trong những mục tiêu của khoa học dữ liệu là:

- | | |
|---------------------------------|----------------------------------|
| A. Tìm hiểu về các loại dữ liệu | C. Thiết kế phần mềm cho dữ liệu |
| B. Khám phá tri thức từ dữ liệu | D. Quản lý mạng máy tính |

Câu 9: Trục quan hóa dữ liệu là gì trong khoa học dữ liệu?

- A. Lưu trữ dữ liệu vào các đĩa cứng
- B. Biểu diễn dữ liệu bằng sơ đồ, biểu đồ hoặc hình ảnh
- C. Dự đoán dữ liệu từ các mô hình
- D. Phân phối dữ liệu ra các nền tảng online

Câu 10: Khám phá tri thức trong khoa học dữ liệu giúp làm gì?

- A. Lưu trữ dữ liệu vào hệ thống phân phối
- B. Tìm ra mối quan hệ ẩn chứa và tạo ra tri thức mới
- C. Phân loại dữ liệu theo độ tuổi
- D. Xử lý dữ liệu trên các hệ thống đám mây

Câu 11: Một trong các mục tiêu quan trọng của khoa học dữ liệu là:

- A. Tối ưu hóa phần cứng máy tính
- B. Quản lý thông tin truyền thông
- C. Phân tích và trục quan hóa dữ liệu
- D. Thiết kế các thuật toán mã hóa dữ liệu

Câu 12: Trong khoa học dữ liệu giúp tối ưu hóa phân bổ tài nguyên và phát hiện hoạt động gian lận cần

- A. Mô hình học sâu
- B. Phân tích và khai thác dữ liệu
- C. Tối ưu hóa hệ thống phần cứng
- D. Lưu trữ dữ liệu hiệu quả

Câu 13: Trong các thành tựu của khoa học dữ liệu, tự động hoá công việc giúp ích gì?

- A. Giảm bớt sự sáng tạo trong công việc
- B. Tiết kiệm thời gian và chi phí
- C. Tăng khối lượng công việc thủ công
- D. Giảm chất lượng công việc

Câu 14: Phân tích dữ liệu trong khoa học dữ liệu giúp gì cho tổ chức?

- A. Xác định các nhóm và xu hướng trong dữ liệu
- B. Giảm thiểu lượng dữ liệu
- C. Tăng cường độ phức tạp của dữ liệu
- D. Phân phối dữ liệu ra bên ngoài

Câu 15: Cá nhân hóa dịch vụ trong khoa học dữ liệu có tác dụng gì?

- A. Cải thiện trải nghiệm khách hàng
- B. Giảm sự hài lòng của khách hàng
- C. Tăng giá trị sản phẩm
- D. Làm cho dịch vụ trở nên đồng nhất

Câu 16: Trong bối cảnh khoa học dữ liệu, "giá trị" của dữ liệu đề cập đến điều gì?

- A. Độ lớn của dữ liệu
- B. Tính hữu ích và ứng dụng của dữ liệu
- C. Thời gian xử lý dữ liệu
- D. Mức độ phức tạp của dữ liệu

Câu 17: Mục tiêu của việc tối ưu hóa quyết định trong khoa học dữ liệu là gì?

- A. Tối ưu hóa quy trình sản xuất và hoạt động
- B. Tạo ra dữ liệu từ các nguồn bên ngoài
- C. Thu thập dữ liệu từ các nguồn dữ liệu lớn
- D. Lưu trữ dữ liệu trong hệ thống an toàn

TRẮC NGHIỆM ĐÚNG SAI

Câu 18: Khoa học dữ liệu có những khả năng

- a) Giúp tự động hoá những tác vụ lặp đi lặp lại.
- b) Trực tiếp làm ra các sản phẩm mới.
- c) Cho phép con người tập trung vào những công việc phức tạp và sáng tạo hơn.
- d) Mọi lĩnh vực đều có thể vận dụng và hưởng lợi nhờ những khả năng của Khoa học dữ liệu.

Câu 19: Ứng dụng nào sau đây không liên quan đến khoa học dữ liệu:

- a) Tự động lái xe
- b) Tính toán diện tích hình tròn
- c) Dự báo thời tiết
- d) Gợi ý phim trên Netflix

BÀI 27: MÁY TÍNH VÀ KHOA HỌC DỮ LIỆU

A. LÝ THUYẾT

1. Tìm hiểu vai trò của máy tính trong Khoa học dữ liệu

- Sự phát triển của máy tính đóng vai trò quan trọng trong việc cung cấp nền tảng cho sự phát triển của Khoa học dữ liệu:

Năng lực tính toán mạnh mẽ, lưu trữ lớn, tốc độ xử lý cao, xử lý đa nhiệm và kết nối mạng là các yếu tố quan trọng giúp **nâng cao khả năng phân tích và hiểu biết dữ liệu lớn**.

Các thành tựu của phần cứng máy tính như: bộ xử lý đồ họa (Graphics Processing Unit – GPU), bộ xử lý tensor (Tensor Processing Unit – TPU), điện toán đám mây, máy tính lượng tử,... đã giúp Khoa học dữ liệu **phát triển nhanh chóng và đạt hiệu quả cao hơn trong việc phân tích xử lý dữ liệu lớn**.

Giai đoạn	Đặc điểm của máy tính	Khoa học dữ liệu
1945 – 1955	Máy tính không có hệ điều hành và được dùng để thực hiện các tính toán đơn giản.	Sử dụng các phương pháp thống kê để phân tích dữ liệu nhỏ.
1955 - 1965	Máy tính có hệ điều hành và xuất hiện ngôn ngữ lập trình.	Thống kê máy tính được áp dụng cho lượng dữ liệu lớn hơn. Máy tính được sử dụng để hỗ trợ, tính toán và phân tích dữ liệu.
1965 - 1980	Máy tính có khả năng xử lý đa nhiệm.	Sự tăng trưởng về số lượng và độ phức tạp của dữ liệu đòi hỏi sử dụng các kỹ thuật xử lý dữ liệu phương pháp thống kê nâng cao hơn.
1980 – 1990	Máy tính cá nhân trở nên phổ biến, Internet phát triển, bắt đầu lưu trữ dữ liệu trên mạng.	Sự phát triển của các phương pháp học máy. Khoa học dữ liệu tập trung vào xử lý và phân tích dữ liệu lớn.
1990 – 2000	Internet và thiết bị di động trở nên phổ biến, tạo ra lượng lớn dữ liệu từ các nguồn khác nhau.	Sự phát triển mạnh mẽ của Học máy và Khoa học dữ liệu, Dữ liệu lớn.
2000 – nay	Điện toán đám mây. Sự xuất hiện của GPU, TPU giúp tăng cường tốc độ và hiệu suất tính toán.	Học máy và Trí tuệ nhân tạo trở nên phổ biến, giúp tăng cường hiểu biết dữ liệu và tự động hóa. Các mô hình học máy tận dụng GPU, TPU để học và giải quyết các vấn đề phức tạp trong Khoa học dữ liệu.

Bảng 1. Đặc điểm của máy tính đối với sự phát triển của Khoa học dữ liệu qua các thời kỳ

- Hiện nay, các dự án liên quan đến Khoa học dữ liệu thường áp dụng một quy trình chung, được gọi là **quy trình Khoa học dữ liệu**:



Hình 1. Quy trình Khoa học dữ liệu

1 Xác định vấn đề: Đặt ra câu hỏi cần trả lời hoặc mục tiêu cụ thể cần giải quyết của dự án.

2 Thu thập dữ liệu: Tìm kiếm, thu thập hoặc tạo dữ liệu cần thiết để giải quyết vấn đề. Dữ liệu thu thập thường được chia làm hai phần:

- + Phần thứ nhất (gọi là dữ liệu huấn luyện) dùng cho mô hình học máy học.
- + Phần còn lại (gọi là dữ liệu kiểm tra) dùng để đánh giá hiệu suất của mô hình.

3 Chuẩn bị dữ liệu: Làm sạch dữ liệu, khám phá tri thức ẩn trong dữ liệu, trực quan hoá dữ liệu và chuẩn bị cho việc xây dựng mô hình học máy.

4 Xây dựng mô hình: Chọn và áp dụng mô hình học máy phù hợp với dự án.

5 Đánh giá: Đánh giá hiệu suất của mô hình với dữ liệu kiểm tra đã có ở bước thu thập dữ liệu.

6 Triển khai: Triển khai mô hình vào sử dụng trong môi trường thực tế.

Nhận xét: Máy tính có vai trò quan trọng trong mọi giai đoạn của quy trình khoa học dữ liệu:

- + Máy tính được sử dụng để lưu trữ, xử lý và phân tích dữ liệu.
- + Các mô hình học máy và khai phá dữ liệu được thực hiện trên các máy tính để tìm kiếm các mẫu và thông tin tiềm ẩn trong dữ liệu.
- + Máy tính cũng được sử dụng để tạo ra các mô hình dự đoán và phân loại.

Nội dung	Vai trò của máy tính	Khoa học dữ liệu
Lưu trữ và xử lý dữ liệu	Dung lượng lưu trữ của phần cứng ảnh hưởng đến khả năng xử lý dữ liệu lớn.	Khoa học dữ liệu yêu cầu lưu trữ hiệu quả để quản lý và truy xuất dữ liệu.
Tính toán và trực quan hoá	Phần cứng máy tính cung cấp sức mạnh tính toán nhanh chóng và hiệu quả.	Khoa học dữ liệu sử dụng tính toán để phân tích, xử lý và hiểu biết dữ liệu lớn.
Điện toán đám mây	Phần cứng mạng ảnh hưởng đến khả năng kết nối và truyền tải dữ liệu, hỗ trợ tính toán phân tán và triển khai trên nhiều máy cùng lúc.	Khoa học dữ liệu sử dụng Internet, điện toán đám mây để lưu trữ và quản lý khối lượng lớn dữ liệu, cho phép sử dụng nguồn lực tính toán mạnh mẽ để xử lý dữ liệu lớn.

Bảng 2. Vai trò của máy tính đối với Khoa học dữ liệu

2. Tính ưu việt trong việc sử dụng máy tính và thuật toán hiệu quả để xử lý dữ liệu lớn.

Tính ưu việt của việc sử dụng máy tính và các thuật toán hiệu quả trong xử lý khối dữ liệu lớn được thể hiện qua khả năng lưu trữ, xử lý, phân tích, khai phá dữ liệu ấy một cách nhanh chóng, nhất quán và hiệu quả.

Trong xử lý dữ liệu lớn, bên cạnh việc sử dụng máy tính, các thuật toán cũng được sử dụng để tăng cường khả năng xử lý:

- **Xử lý song song:** tận dụng sức mạnh của máy tính để giảm thời gian và chi phí trong quá trình xử lý dữ liệu.

- **Tự động hoá:** Thuật toán giúp tự động hoá nhiều quy trình xử lý dữ liệu, giảm bớt sự phụ thuộc vào các quá trình thủ công.

Ví dụ cụ thể: Hệ gene người là một chuỗi các nucleotide, kí hiệu là A, C, G, T mang thông tin di truyền quyết định đến hình dáng, sức khỏe, bệnh tật và thậm chí cả tính cách và Dự án Hệ gene

người (Human Genome Project - HGP) là một nỗ lực mang tính đột phá, nhằm xác lập hệ gene và giải mã bản thiết kế di truyền hoàn chỉnh của con người.

B. BÀI TẬP

TRẮC NGHIỆM

Câu 1: Giai đoạn đầu tiên trong quy trình Khoa học dữ liệu thường là gì?

- A. Phân tích dữ liệu.
- B. Xác định vấn đề
- C. Trực quan hóa dữ liệu.
- D. Báo cáo kết quả.

Câu 2: Đặc điểm nào sau đây không phải là ưu điểm của việc sử dụng máy tính trong xử lý dữ liệu lớn?

- A. Tốc độ xử lý nhanh.
- B. Khả năng lưu trữ dữ liệu lớn.
- C. Độ chính xác cao.
- D. Nâng cao khả năng bảo mật dữ liệu.

Câu 3: Việc sử dụng các quy trình tự động trong phân tích dữ liệu gen giúp cải thiện yếu tố nào sau đây?

- A. Độ chính xác của kết quả.
- B. Tốc độ in ấn.
- C. Độ nét của hình ảnh.
- D. Màu sắc của dữ liệu.

Câu 4: trong Khoa học dữ liệu, 'làm sạch dữ liệu' chủ yếu thuộc giai đoạn nào?

- A. Sau khi phân tích dữ liệu.
- B. Trước khi thu thập dữ liệu.
- C. Trong quá trình thu thập dữ liệu.
- D. Sau khi thu thập dữ liệu và trước khi phân tích dữ liệu.

Câu 5: Quy trình Khoa học dữ liệu gồm những bước nào?

- A. Xác định vấn đề, thu thập dữ liệu, chuẩn bị dữ liệu, xây dựng mô hình, đánh giá và triển khai
- B. Chỉ thu thập và xử lý dữ liệu
- C. Chỉ phân tích và báo cáo kết quả
- D. Chỉ khám phá tri thức

Câu 6: Máy tính và thuật toán hiệu quả giúp xử lý dữ liệu lớn như thế nào?

- A. Chỉ lưu trữ dữ liệu
- B. Giúp xử lý dữ liệu nhanh chóng, nhất quán và hiệu quả
- C. Chỉ phân tích dữ liệu nhỏ
- D. Chỉ giải thích thông tin di truyền

Câu 7: Công việc nào không phải là một phần của quy trình Khoa học dữ liệu?

- A. Khám phá tri thức.
- B. Phân tích dữ liệu.
- C. Tạo lập ngân sách.
- D. Triển khai mô hình.

Câu 8: Mục đích chính của việc sử dụng thuật toán hiệu quả trong xử lý dữ liệu lớn là gì?

- A. Tăng tốc độ phân tích dữ liệu.
- B. Thay đổi định dạng của dữ liệu.
- C. Tăng cường độ âm thanh khi xử lý dữ liệu.
- D. Làm giảm chất lượng của dữ liệu.

Câu 9: Máy tính giúp gì trong quy trình khoa học dữ liệu?

- A. Xử lý và lưu trữ dữ liệu, phân tích, trực quan hóa và tự động hóa
- B. Chỉ xử lý dữ liệu
- C. Chỉ lưu trữ dữ liệu
- D. Chỉ phân tích dữ liệu

Câu 10: Thuật toán trong dự án HGP giúp làm gì?

- A. Giảm bớt dữ liệu

- B. Giải thích thông tin di truyền và xác định gene
- C. Lưu trữ dữ liệu không cần phân tích
- D. Hạn chế khả năng phân tích dữ liệu

Câu 11: Điều nào sau đây không đúng về vai trò của máy tính trong Khoa học dữ liệu?

- A. Chỉ thực hiện các phân tích đơn giản.
- B. Tạo các biểu đồ và đồ thị trực quan.
- C. Xử lý và lưu trữ dữ liệu.
- D. Tự động hóa các nhiệm vụ lặp đi lặp lại.

Câu 12: Thuật toán hữu hiệu trong xử lý dữ liệu lớn giúp cải thiện khía cạnh nào của việc giải thích thông tin di truyền?

- A. Tạo ra mô hình vật lý của gen.
- B. Tạo ra hình ảnh 3D của gen.
- C. Phát triển trò chơi điện tử về gen.
- D. Xác định gen và các đặc tính của chúng.

Câu 13: Vai trò của máy tính trong Khoa học dữ liệu?

1. Chỉ thực hiện các phân tích đơn giản.
2. Tạo các biểu đồ và đồ thị trực quan.
3. Xử lý và lưu trữ dữ liệu.
4. Tự động hóa các nhiệm vụ lặp đi lặp lại.

Số phát biểu đúng là:

- A. 2.
- B. 1.
- C. 3.
- D. 4

Câu 14: Điện toán đám mây có thể đem lại lợi ích gì cho Khoa học dữ liệu?

- A. Cung cấp tài nguyên tính toán linh hoạt.
- B. Cung cấp giải pháp chống virus.
- C. Tạo ra dữ liệu mới.
- D. Thực hiện phân tích tâm lý con người.

TRẮC NGHIỆM ĐÚNG SAI

Câu 15: Xử lý song song trong Khoa học dữ liệu nói chung và trong Dự án gen người nói riêng (HGP) có ý nghĩa

- a) Tăng độ chính xác tính toán.
- b) Giảm thời gian xử lý dữ liệu lớn.
- c) Tăng cường bảo mật dữ liệu.
- d) Tăng khả năng xử lý dữ liệu.

Câu 16: Những khả năng giúp máy tính có vai trò quan trọng trong việc trực quan hóa dữ liệu

- a) Tạo các hình ảnh 3D.
- b) Cung cấp công cụ tạo biểu đồ và đồ thị.
- c) Thay thế nhà khoa học dữ liệu.
- d) Tự động hoàn toàn việc tạo nội dung các báo cáo.