**Unit 3**

**COMPUTER APPLICATIONS**

**Starter 1:**

1 Supermarkets 🡪 Pricing, stock control, checking cash cards, checking sell-by dates, tracking customer buying habits….

2 Hospitals 🡪 database of equipment, accounts, staff records, drugs and supplies…

3 Airports 🡪 plotting aircraft movement –air traffic control, arrival and departure information displays, staff database…

4 Police Headquarters 🡪 database of crimes, criminals, missing people, staff rosters….

**Reading 2:**

1 The first unit records the time each vehicle passes.

2 It identifies each vehicle by its number plates using OCR software.

3 The second unit also records the time each vehicle passes.

4 The microprocessor calculates the time taken to travel between the units.

5 It relays the registration numbers of speeding vehicles to police headquarters.

6 A computer matches each vehicle with the DVLC database.

**Key 5:**

1 The time each vehicle passes is recorded by the first unit.

🡪 Thời gian mỗi xe đi qua được ghi nhận bởi đơn vị đầu tiên.

2 Each vehicle is identified by its number plates using OCR software.

🡪 Mỗi chiếc xe được xác định biển số của nó bằng cách sử dụng phần mềm OCR.

3 The information is relayed to the second unit.

🡪 Thông tin này được chuyển đến các đơn vị thứ hai.

4 The time each vehicle passes is also recorded by the second unit.

🡪 Thời gian mỗi xe đi qua cũng được ghi lại bởi các đơn vị thứ hai.

5 The time taken to travel between the units is calculated by the microprocessor.

🡪 Thời gian hành trình giữa các đơn vị được tính bằng bộ vi xử lý.

6 The registration numbers of speeding vehicles are relayed to police headquarters.

7 Each vehicle is matched with the DVLC database.

8 A letter is printed off to the vehicle owners using mailmerge.

**Key 6:**

1 -f The checkout operator scans the item.

2 -c The scanner reads the barcode.

3 -a The scanner converts the barcode into electrical pulses.

4 -g The scanner sends the pulses to the branch computer.

5 -i The branch computer searches the stock fine for a product matching the barcode EAN.

6 -d The branch computer records the sale of the product.

7 -b The branch computer sends the price and description of the product to the EPOS still.

8 -e The still shows the item and price.

9 -h The till prints the item and price on the paper receipt.

**PROBLEM – SOLVING**

Key 7:

ON –board diagnosis of faults, internet connections for passengers – for driver only if vehicle has stopped.

Systems to immobilize the vihicle if the driver is not recognized.

Automatic adjustment to suit individual drivers – seat hight, steering wheel position, mirrors …

**WRITING**

Key 9:

ATM

First the customer puts their card into the machine and enters their PIN number on the keypad and the amount they wish to withdraw. The ATM reads the information on the magnetic strip on the card. The strip contains the name of the account holder, their account number and details of the network it is linked to. This information is then sent to a central computer which hold information on many accounts. The customer’s PIN number is checked as well as the amount of money in their account. If their account has sufficient funds, the ATM is instructed to dispense the cash requested.

VISA

A customer uses a Visa card to pay for a $1.295 computer in Chicago. The store uses a swipe card reader to send details of the purchase and the customer’s card number to a processing center in New Jersey. The information may travel by satellite or phone lines. Because the purchase is above a certain limit, details are forwarded to a computer on Atlanta for closer scrutiny. Next the transaction is sent to San Mateo in California for processing. The California computer checks the card details in a database of card holder and finds that the card was issued by a Portland bank’s computer to see if the transaction request should be approved or denied. The response is sent back to the original store in approximately 15 seconds.

**SPEACIALIST**

**Key A:**

1🡪 Artificial Intelligence

2🡪 a decision trees

b neural networks

c clustering

3 🡪rules

4 🡪When data isn’t labelled in a way that is favourable to mining.

5 🡪All types of data storage, from large data warehouses to smaller desktop databases to flat files.

6 🡪 Refine the parameters, Uses other data analysis tools to examine the data

7🡪 1- Analysing Supreme Court decisions

2- Discovering patterns in health care

3- Pulling stories about competitors from newswires

4- Resolving bottlenecks in production processes

5- Analysing sequences in the human genetic makeup.

**Key B:**

1 A iii B iv C ii D i

2 A-F B-T C-T D-T E-F

3 Large amounts of data stored in data ***warehouses*** are often used for data ***mining***. The data is first ***cleansed*** to remove ***duplicate*** data and errors. The ***data*** is then analyzed using a tool such as ***artificial intelligence***. An analysis report is then analysed by an ***analyst*** who decides if the parameters need to be refined, other data ***analysis*** tools need to be used, or if the results need to be discarded because they are ***unusable***. The analyst passes the final results to the ***decision*** makers who decide on the ***appropriate*** action. (Lượng lớn dữ liệu được lưu trữ trong kho dữ liệu thường được sử dụng để khai thác dữ liệu. Các dữ liệu đầu tiên được làm sạch để loại bỏ trùng lặp dữ liệu và lỗi. Dữ liệu sau đó được phân tích bằng cách sử dụng một công cụ như trí tuệ nhân tạo. Báo cáo phân tích sau đó được phân tích bởi một nhà phân tích người quyết định nếu các thông số cần phải được tinh chế, công cụ phân tích dữ liệu khác cần phải được sử dụng, hoặc nếu kết quả cần phải được loại bỏ vì chúng không sử dụng được. Các nhà phân tích qua các kết quả cuối cùng để các nhà sản xuất quyết định quyết định về hành động thích hợp.)

**DATA MINING**

Data mining is simply filtering through large amounts of raw data for useful information that gives businesses a competitive edge. This information is made up of meaningful patterns and trends that are already in the data but were previously unseen.

*=> khai thác dữ liệu chỉ đơn giản là lọc qua một lượng lớn dữ liệu thô cung cấp những thông tin hữu ích giúp cho doanh nghiệp có một lợi thế trong cạnh tranh. thông tin này được tạo thành có ý nghĩ từ các mô hình và xu hướng đã có trong dữ liệu mà trước đây không nhìn thấy được.*

The most popular tool used when mining is artificial intelligence (AI), AI technologies try to work the way the human brain works, by making intelligent guesses, learning by example, and using deductive reasoning. Some of the more popular AI methods used in data mining include neural networks, clustering, and decision trees.

*=> công cụ phổ biến nhất được sử dụng khi khai thác là trí tuệ nhân tạo (AI), AI là công nghệ có thể làm việc theo cách mà bộ não con người hoạt động,bằng cách dự đoán thông minh, học tập bằng ví dụ, và sử dụng lập luận suy diễn. một số phương pháp phổ biến hơn AI được sử dụng trong khai thác dữ liệu bao gồm các lưới mạng, tách nhóm, và Cây quyết định.*

Neural networks look at the rules of using data, which are based on the connections found or on a sample set of data. As a result, the software continually analyses value and compares it to the other factors, and it compares these factors repeatedly until it finds patterns emerging. These patterns are known as rules, The software then looks for other patterns based on these rules or sends out an alarm when a trigger value is hit.

*=> Lưới mạng trông như các quy tắc của việc sử dụng dữ liệu, mà là dựa trên các kết nối hoặc tìm thấy trên một mẫu thiết lập dữ liệu. kết quả là,các phần mềm liên tục phân tích giá trị và so sánh nó với các yếu tố khác, và nó so sánh những yếu tố này lặp đi lặp lại cho đến khi nó tìm thấy những mẫu nổi bật. Các mẫu được gọi là quy tắc, các phần mềm sau đó tìm kiếm các mô hình khác dựa trên những quy tắc hoặc gửi ra báo động khi một giá trị kích hoạt được thực hiện.*

Clustering divides data into groups based on similar features or limited data ranges. Clusters arc used when data isn’t labelled in a way that is favourable to mining. For instance, an insurance company that wants to find instances of fraud wouldn‘t have its records labelled as fraudulent or not fraudulent. But after analysing patterns within cluster; the mining software can start to figure out the rules that point to which claims are likely to be false.

=> Clustering chia dữ liệu thành các nhóm dựa trên các tính năng tương tự hoặc các phạm vi dữ liệu hạn chế. cụm hồ quang được sử dụng khi dữ liệu không được dán nhãn một cách thuận lợi để khai thác. Ví dụ, một công ty bảo hiểm mà muốn tìm thấy trường hợp gian lận, hồ sơ không được dán nhãn là gian lận hay không gian lận. nhưng sau khi phân tích mẫu trong cụm; phần mềm khai thác có thể bắt đầu để tìm ra các quy tắc mà chỉ để mà tuyên bố có thể sẽ là sai**.**

decision trees, like clusters, separate the data into subsets and then analyse the subsets to divide them into further subsets, and so on(for a few more levels). The final subsets are then small enough that the mining process can find interesting patterns and relationships within the data

*=> Cây quyết định, giống như cụm, tách dữ liệu thành các tập con và sau đó phân tích các tập con để phân chia chúng thành các tập con hơn nữa, và như vậy (ở một mức độ ít hơn). Các tập con cuối cùng đủ nhỏ để quá trình khai thác có thể tìm thấy các mô hình và các mối quan hệ trong dữ liệu*

Once the data to be mined is identified, it should be cleansed. cleansing data frees it from duplicate information and erroneous data. next, the data should be stored in a uniform format within relevant categories or fields, mining tools 45/ can work with all types of data storage, from large data warehouses to smaller desktop databases to flat files. Data warehouses and data marts are storage methods that involve archiving large amounts of data in a way that makes it easy to access when necessary.

*=> Một khi dữ liệu được khai thác được xác định, nó phải được làm sạch. làm sạch dữ liệu giải phóng nó từ bản sao thông tin và dữ liệu có sai sót. tiếp theo, các dữ liệu sẽ được lưu trữ trong một định dạng thống nhất trong các hạng mục hoặc các lĩnh vực có liên quan,*

*các công cụ khai thác có thể làm việc với tất cả các loại lưu trữ dữ liệu, từ kho dữ liệu lớn đến cơ sở dữ liệu máy tính để bàn nhỏ đến các tập tin. Kho dữ liệu và cửa hàng dữ liệu là phương pháp lưu trữ lượng dữ liệu lớn và dễ dàng truy cập khi cần thiết.*

When the process is complete, the mining software generates a report. An analyst goes over the report to see if further work needs to be done, such as refining parameters, using other data analysis tools to examine the data, or even scrapping the data if it's unusable. if no further work is required, the report proceeds to the decision makers for appropriate action

*=> Khi quá trình hoàn tất, phần mềm khai thác tạo ra một báo cáo. Một nhà phân tích xem qua báo cáo để xem nếu tiếp tục cần phải làm việc gì, chẳng hạn như các thông số tinh chỉnh, sử dụng các công cụ phân tích dữ liệu khác để kiểm tra dữ liệu, hoặc thậm chí tháo bỏ các dữ liệu nếu nó không sử dụng được. Nếu tiếp tục là không cần thiết, thì báo cáo tiền thu được để các nhà sản xuất quyết định cho hành động thích hợp.*

The power of data mining is being used for many purposes, such as analysing Supreme Court decisions, discovering patterns in health care, pulling stories about competitors from newswires, resolving bottlenecks in production processes, and analysing sequences in the human genetic makeup. there really is no limit to the type of business or area of study where data mining can be beneficial.

*=> Sức mạnh của khai thác dữ liệu đang được sử dụng cho nhiều mục đích, chẳng hạn như phân tích quyết định của Tòa án Tối cao,trình bày mô hình chăm sóc sức khỏe, giải quyết các vướng mắc trong quá trình sản xuất, và phân tích trình tự trong các mặt di truyền của con người. Lợi ích của khai thác dữ liệu thực sự không có giới hạn các loại hình kinh doanh hoặc ngành học.*