MEETING 8

FUTURE BIG DATA

A. LEARNING OBJECTIVES

In this meeting students will be able speak about data storage, understand reading and listening text about Big data used to predict the future.

B. DESCRIPTION OF MATERIAL

1. Compound Sentences

a subject and a verb is the simple sentence, the basic element that can express complete idea.

For example :

- ✓ I think I will take C++ course, or I will take designing course
- ✓ I really want to be programmer, but I still need to improve my skill on language programming.
- ✓ I have been spending three hours on typing, but the assignment is unfinished.
- ✓ Adam asked me to attend the robotic seminar in campus, and I said I would come.
- ✓ Do you want to take computer software, or computer hardware course?
- ✓ I was going to buy external hard disk, but I realized my money was not enough.
- ✓ I went to SQL class yesterday, and it was very interesting class.
- ✓ The teacher made the lesson interesting, and every body enjoyed it.
- ✓ Do you want to be a computer designer, or do you want to be computer programmer?
- ✓ I have studied language programming for 2 years, as a result I can make game for children.
- ✓ You are smart, however you need to be patient working with computer.
- ✓ My friend has finished her computer course. She works in financial office now.
- ✓ You might need to take certification before you apply job as computer analysis.
- ✓ you have to take C++ class or you can take Java language programming language class.

C. EXERCISES

1. Structure

1.1 Rewrite the sentences below and put the conjunction to make compound sentences.

1) A friend asked me to help her SQL lesson. I helped to finish her SQL project (so)

_____

_____

2) The Numerical Analysis was difficult for me. The lecturer was really good at explaining it (but)

_____

_____

3) The lecturer is explaining data processing system. Students are taking notes. (and)

_____

_____

4) Would you like to increase computer performance, would you like to reduce the quantity of information? (or)

_____

_____

5) Many students wants to be computer programmer. They don't like mathematic. (but)

_____

_____

_____

6) Ivana's computer memory was not big enough to store big data. she bought bigger memory. (so)

_____

_____

_____

2. Speaking:

2.1 Storage Capacity

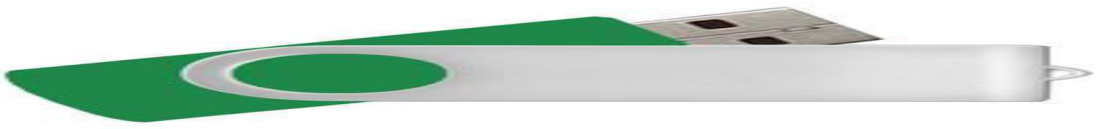
External disk, laptop, hand phone, hard drive, flash media, digital camera,

32 GB capacity **http://www.eyefi.com/**

4 GB RAM, DDR4, 128GB SSD storage **https://www.argos.co.uk/product/8729493**

RAM 6 GB internal, memory 128 GB

Toshiba hard drive 500GB



flash memory 32GB

See the capacity of storage above COMPARE WITH YOUR OWN DEVICE CAPACITIES and discuss the questions with your partner below! See example!

1) Do you know the capacity memory of each device above?
2) What is average capacity of hard drive? What is your laptop hard drive capacity?
3) What is the capacity of your hand phone memory?
4) What device has the biggest memory?
5) How much data can be stored in Blu-ray-disc?
6) What file need big storage?
7) What software application has the biggest size in your laptop?
8) what is the average size of application in smart phone?
9) What is the average size of Microsoft office file? JPEG file? Video file? And game file?
10) What file filled up your laptop? Why?

Example

Bobby Do you know the capacity memory of each device above?

Harry Each device has different capacity, a camera has 16 to 32 MB, a hard disk has 500 MB to 1 Terabyte, a flash disk can have 32 GB.

Bobby What is average capacity of hard drive? What is your laptop hard drive capacity?

Harry I think the average capacity of HDD is 500 GB.

Bobby What is the capacity of your hand phone memory?

Harry My phone has 3 GB memory

Bobby What device has the biggest memory?

Harry I think laptop has the biggest meory

Bobby How much data can be stored in Blu-ray-disc?

Harry Blu-ray-disc can store up to 200 GB

Bobby What file need big storage?

Harry Video file need big storage

Bobby What software application has the biggest size in your laptop?

Harry Microsoft window is the biggest size in my laptop.

Bobby what is the average size of application in smart phone?

Harry I think the average size of application is 20 MB

Bobby What is the average size of Microsoft office file? JPEG file? Video file? And game file?

Harry Microsoft office file about 300 kb, while game file up to 55 GB

Bobby What file filled up your laptop? Why?

Harry Most files in my laptop are videos, music, and document

3. Reading

3.1 Underline the compound sentence in the reading text below!

3.2 Read and record loud the reading text in your phone!

Big data used to predict the future

http://www.thomas-robert.fr/en/my-sum-up-of-big-data-concepts/

By University of Córdoba

Technology is taking giant leaps and bounds, and with it, the information with which society operates daily. Nevertheless, the volume of data needs to be organized, analyzed and crossed to predict certain patterns. This is one of the main functions of what is known as 'Big Data', the 21st century crystal ball capable of predicting the response to a specific medical treatment, the workings of a smart building and even the behavior of the Sun based on certain variables.

Researcher in the KIDS research group from the University of Cordoba's Department of Computer Science and Numerical Analysis were able to improve the models that predict several variables simultaneously based on the same set of input variables, thus reducing the size of data necessary for the forecast to be exact. One example of this is a method that predicts several parameters related to soil quality based on a set of variables such as crops planted, tillage and the use of pesticides. "When you are dealing with a large volume of data, there are two solutions. You either increase computer performance, which is very

expensive, or you reduce the quantity of information needed for the process to be done properly," says researcher Sebastian Ventura, one of the authors of the research article.

When building a predictive model there are two issues that need to be dealt with: the number of variables that come into play and the number of examples entered into the system for the most reliable results. With the idea that less is more, the study has been able to reduce the number of examples, by eliminating those that are redundant or "noisy," and that therefore do not contribute any useful information for the creation of a better predictive model.

As Oscar Reyes, the lead author of the research, points out "we have developed a technique that can tell you which set of examples you need so that the forecast is not only reliable but could even be better." In some databases, of the 18 that were analyzed, they were able to reduce the amount of information by 80% without affecting the predictive performance, meaning that less than half the original data was used. All of this, says Reyes, "means saving energy and money in the building of a model, as less computing power is required." In addition, it also means saving time, which is interesting for applications that work in real-time, since "it doesn't make sense for a model to take half an hour to run if you need a prediction every five minutes."

As pointed out by the authors of the research, these systems that predict several variables simultaneously (which could be related to one another), based on several variables known as multi output regression models,- are gaining more notable importance due to the wide range of applications that "could be analyzed under this paradigm of automatic learning," such as for example those related to healthcare, water quality, cooling systems for buildings and environmental studies.

https://www.sciencedaily.com/releases/2018/11/181109101444.htm

3.3 Answer the question below!

1) What is the main functions of Big Data?
2) What are the solutions dealing with a large volume of data?
3) Who is Oscar Reyes?
4) What is multi output regression models?
5) Give example of method that forecast several parameter connected to soil!
6) What are two problems that dealing with building predictive model?
7) how does multi output regression model work?

3.4 Mark the Correct for the correct statement and Mark incorrect for incorrect statement below!

1) Big data need to be managed to forecast specific patterns
   O correct
   O incorrect
2) The 21st century crystal ball isn't able to predict the response to a specific medical treatment.
   O correct
   O incorrect
3) It possible to predict behavior of the sun on specific varibles.
   O correct
   O incorrect
4) The research on Big data was done by the University of Cordoba's Departement of Computer science.
   O correct
   O incorrect
5) Models can be improved by numerical analysis
   O Correct
   O incorrect
6) The research on numerical analysis is undergone by a group of people from university of Cordoba.
   O correct
   O incorrect
7) Reducing the size of data is important so that the prediction will be exact.
   O correct
   O incorrect
8) several parameters prediction method that related soil can't be used for farmer to grow crops more efficient.
   O correct
   O incorrect
9) You don't need to increase your computer performance for big volume of data.
   O correct
   O incorrect

10) High performance computer costs a lot.
    O correct
    O incorrect
11) To get reliable and better prediction, technique has been developed.
    O correct
    O incorrect
12) multi output regression models is also known as multi variable .
    O correct
    O incorrect
13) multi output regression models can be used for healthcare .
    O correct
    O incorrect
14) Sebastian venture is a researcher that had written article about the research.
    O correct
    O incorrect
15) Number of examples entered into the system is a problem that need to be handle in building predictive data.
    O correct
    O incorrect
16) So far the study hasn't be able to decrease the number of examples that are not beneficial for the predictive design.
    O correct
    O incorrect
17) Oscar Reyes also writes about the research.
    O correct
    O incorrect
18) The system of predictive model can predict several variables continually.
    O correct
    O incorrect
19) Application on predictive model can work real time.
    O correct
    O incorrect

4. Listening

4.1 Listen to Sultan Issa Hampton and mark the statement correct or incorrect.

1) According to the speaker word big data is old.
   O incorrect
   O correct
2) **T**he big data idea reached its time in early 2010
   O incorrect
   O correct
3) Doug Laney is an computer analyst .
   O incorrect
   O correct
4) **T**he size in a big data is management that gather data from various of sources such as business, social media and information from sensor or machine to machine data.
   O incorrect
   O correct
5) The speed of unprecedented data stream and have to deal with a specific time is called velocity in big data.
   O incorrect
   O correct
6) Variety in big data is Data that comes in only two types of formats.
   O incorrect
   O correct

4.2 Listen to the audio again and answer the question below!

1) The term of big data according to the speaker is
   a. New term c. modern term
   b. Old term d. the latest term
2) Collecting and storing massive information to be analyze …..
   a. Happened lately c. never has happened

b. Already exist long time ago d. still under developed

3) The concept of collecting big data was started….
   a. 20 years ago c. in 2020
   b. 12 years d. this year
4) Organization that gather data from different sources such as business transaction, information from machine to machine data and social media is called…
   a. Hadoop c. volume
   b. Variety d. data organization
5) The example technology that can store massive data that speaker mention is …….
   a. Velocity c. HDD
   b. Floppy disk d. hadoop
6) According to the speaker data stream in an unprecedented speed have to handle in a timely manner is called …
   a. Velocity c. volume
   b. Diskette d. hadoop
7) Data that come from all type of format such as numeric data in traditional database to text document that is not structured is called …..
   a. Variety c. volume
   b. Velocity d. data base
8) According the speaker that collecting big data from video, audio, email, financial transaction etc…..the example of
   a. Date base c. velocity
   b. Variety d. hadoop

5. Writing

5.1 Complete the sentences below!

1) Technology is taking improving so

   _____

2) People operate computer everyday

   _____

3) I want to increase my computer performance so

4) High capacity computer cost a lot but

5) You have to take courses on computer and

6) You have to back up your data or

7) When you are typing you have to save you data frequently and

8) Big data needs big memory and

9) Cloud Computing provides big storage and

10) Cloud computing can share resources and

11) Data base manager is important for financial institution

12) Because storage device needs be free from virus

13) Big data need to be organized

14) Researchers try to make much bigger data storage

15) Internet of thing will need nigger storage data

5.2 Translate the sentences below

1) The content of the data have to be analyzed, managed and crossed to forecast specific form.

2) The instance of a pattern that forecast few parameters connect to land quality based on a body of variable for instance tree planted, farmer plantation that used pesticide.

_____

_____

3) When constructing a prognostic pattern there were two trouble that need to be handled. The few variables that sometimes that we use and few example that get into the system for the best result.

_____

_____

4) Techniques have been improved that are able to you which set of example you need so that prediction is not only certain but also better.

_____

_____

5) It is not normal for a model if it takes 30 minutes' to run if you need a prediction each five minutes.

_____

_____

6) Those systems that forecast few variables synchronously based on few variables popular as multi output regression models.

_____

_____

6. key words

6.1 find the synonym and translate them

English	Synonym	Indonesian
behavior, certain, crops, deal, exact, forecast, increase, large, Leap, notable, organized, performance, pesticides, predict, predictive, properly, reduce, reliable, required, response, result, simultaneously, soil tillage, treatment, volume,		

D. BIBLIOGRAPHY

https://www.sciencedaily.com/releases/2018/11/181109101444.htm

http://www.thomas-robert.fr/en/my-sum-up-of-big-data-concepts/

http://www.eyefi.com/

https://www.argos.co.uk/product/8729493