Вопрос 1
Пока
нет ответа
Балл: 5,00

1. Match the Big Data vocabulary with the definitions.

1. Big Data

2. data volume

3. data velocity

4. data variety

5. big data analytics

a.   the process of collecting, organizing and analyzing large sets of data ('big data') to discover patterns and other useful information

b.   the speed of generation of data or how fast the data is generated and processed to meet the demands and the challenges which lie ahead in the path of growth and development

c.   the size of the data which determines the value and potential of the data under consideration and whether it can actually be considered as Big Data or not

d.   a massive volume of both structured and unstructured data that is so large that it's difficult to process using traditional database and software techniques

e.   one of the aspects of Big Data that means that the category to which Big Data belongs to is also a very essential fact that needs to be known by the data analysts

2. Watch this video and fill in the gaps with the words in the table.

recorded	quality	digital	Internet of things	data	Big Data	information	quantity

The _____ (1) of computer data generated on Planet Earth is growing exponentially for many reasons.

For a start retailers are building up vast databases of _____ (2) customer activity. Organizations

working in the logistics, financial services, healthcare and many other sectors are also now capturing more

_____ (3). The public use of social media is also creating vast quantities of _____ (4) material. As

vision recognition improves, it is additionally starting to become possible for computers to extract meaningful

information from still images and video.

As more smart objects go online, Big Data is also being generated by an expanding _____

(5). And finally, several areas of scientific advancement are starting to generate and rely on vast quantities of

data that were until very recently almost unimaginable.

Вопрос **3**

Пока
нет ответа

Балл:
10,00

3. Watch the video and choose the right answer: True, False or Not Stated.

1. Organizations working in many sectors such as financial services, the logistics, healthcare are now capturing less and less data.

○ a. False

○ b. True

○ c. Not Stated

2. Big Data is being generated by an expanding Internet of things.

○ a. True

○ b. False

○ c. Not Stated

3. Storing, interlinking and processing vast quantities of digital information offers tremendous possibilities for a wide range of activities which is understood by volume.

○ a. True

○ b. False

○ c. Not Stated

4. Velocity is a characteristic of Big Data, which means the types of data that many organizations are called on to process are becoming increasingly diverse and dense.

○ a. False

○ b. True

○ c. Not Stated

5. Variety is characterized by the rate at which data is flowing into most organizations which is increasing beyond the capacity of their IT systems to store and process.

○ a. False

○ b. True

○ c. Not Stated

6. Due to the great number of issues arising in the combination of three V's organizations are trying not to ignore large quantities of valuable information.

○ a. False

○ b. True

○ c. Not Stated

7. At present, many Big Data pioneers are deploying a Hadoop ecosystem alongside their legacy IT systems in order to allow them to combine old and new data in new ways.

○ a. True

○ b. False

○ c. Not Stated

8. Many organizations are unlikely to have the resources and expertise to implement their own Hadoop solutions. Fortunately they do not have to, as cloud solutions are already available.

○ a. True

○ b. False

○ c. Not Stated

9. As Netapp explain, Big Data developments are fundamentally about creating new IT systems that are more "systems of engagement" rather than just silos for data storage.

○ a. True

○ b. False

○ c. Not Stated

10. Using Big Data, we could start to run the world and allocate resources based on what we really need, not what we blindly guess people may in the near future demand. Or in other words, the more we can know and learn about human activities, the less we will need to go on producing and transporting goods to fill up retail outlets with things that people may not actually want.

○ a. True

○ b. False

○ c. Not Stated

Вопрос **4**
Пока
нет ответа
Балл:
10,00

4. Put the words in the correct form to fill in the gaps.

VOLUME, VELOCITY AND VARIETY

Big Data is often characterized using the "three Vs" of volume, velocity and variety.

Volume is Big Data's greatest challenge and as well as its greatest opportunity. This is because storing, interlinking and processing vast quantities of digital [_____] (1. inform) offers tremendous possibilities for a wide range of activities. These include predicting customer [_____] (2. behave), diagnosing disease, planning healthcare services, and modelling our climate. However, traditional [_____] (3. compute) solutions like relational databases are increasingly not capable of handling such tasks. Most [_____] (4. tradition) computer hardware solutions are also not scalable to Big Data proportions.

Big data velocity also raises a number of key issues. For a start, the rate at which data is flowing into most [_____] (5. organize) is increasing beyond the capacity of their IT systems to store and process. In addition, users [_____] (6. increase) want **streaming data** to be delivered to them in real time, and often on mobile devices. Online video, [_____] (7. locate) tracking, **augmented reality** and many other applications now rely on large quantities of such high velocity data streams, and for many companies delivering them is proving quite a challenge.

Finally, as already highlighted, Big Data is characterised by its variety, with the types of data that many organizations are called on to process becoming increasingly diverse and dense. Gone are the days when data centres only had to process documents, financial [_____] (8. act), stock records, and personnel files. Today, photographs, audio, video, 3D models, complex simulations and location data are all being piled in to many a corporate data silo. Many of these Big Data sources are also almost entirely [_____] (9. structure), and hence not easy to categorize, let alone process, with traditional computing techniques. All of this means that Big Data is in reality messy data, with a great deal of effort required in complex pre-processing and data cleansing before any [_____] (10. mean) analysis can be carried out.

Вопрос **5**

Пока
нет ответа

Балл: 5,00

5. Match the words and their definitions.

1. Data exhaust

2. Augmented reality (AR)

3. HDFS

4. The Apache™ Hadoop(r)

5. MapReduce

a.   the Hadoop Distributed File System (HDFS), which permits the high-bandwidth, cluster-based storage essential for Big Data computing.

b.   A data processing framework which is the second part of Hadoop.

c.   integrates the real world and cyberspace.

d.   A large proportion of the data that they gather is then not processed, with a significant quantity of useful information passing straight through them

e.   this project develops open-source software for reliable, scalable, distributed computing.

ПРЕДЫДУЩИЙ ЭЛЕМЕНТ КУРСА

◄ Class Work 2

Перейти на...

СЛЕДУЮЩИЙ ЭЛЕМЕНТ КУРСА

Check Point 2 ►

Инструкции по работе в системе