Data Mining

Student Name

Course Name

Professor Name

Due Date

Knowledge is derived as a result of transforming data and deducing information to portray the idea that one knows. However, knowledge is becoming an essential asset to learning institutions, enterprises, and government bodies in this new era. Increased use of technology has triggered the evolution of knowledge workers who mostly rely on information creation, distribution, and management. Data is growing in huge volumes every day, and there is a need to identify the significance of data by establishing the critical information and contrasting it with ‘noise’ instead of throwing it all away. In the beginning, data mining was thought of as an unethical process regarded as data fishing and spying. Still, the demand for analyzing patterns in data and extracting knowledge became prevalent like no other time in the past. Methods used to analyze data involve using a set of steps that carry out instructions on their own and are expected to produce an outcome after a series of steps. Data mining seeks to revolutionize how technology has grown by gaining more control of monitoring and even thinking for itself. Therefore, data mining is hype and an essential asset in today’s competitive world, whereby knowledge is becoming a widely accepted cultural and economic source of income.

Data mining can also be referred to as knowledge discovery when used in databases since knowledge is discovered in data. The methods used to discover knowledge include: data characterisation whereby data is analysed to produce a set of characteristics about it; Data discriminalisation whereby the characterized data is assessed to produce rules about the a specific model; Association analysis is a method that seeks to derive relationships between frequently occurring items in a database; Data classification groups data into various classes with predefined labels; prediction is done on a pattern of data and make expected outcomes; Clustering is the grouping of data into non defined classes and sorting it to appropriate labels; Outlier analysis is a method that identifies elements of data that are not able to be labelled so as to determine if they are noise or useful knowledge; Evolution analysis is a method that seek to mine data collected over a period of time and models it to an evolution trend; Deviation analysis is the method that seeks to analyse differences between measured and expected values so as to find cause of deviations.

Data mining is not a simple transformation but is a painstaking process of research and product development. It is supported by data collection tools that store large amounts of data in tandem with the development of algorithms in multiprocessing computers to derive information. The algorithm development is usually a complex procedure that requires the developer's high level of expertise and skill. The developed algorithms use techniques that analyze data in a meaningful way, including neural networks, decision trees, genetic algorithms, and rule induction. In conclusion, data mining is a way of modeling data in that data is set to fit a specific model and to use it to solve other similar scenarios.

References

Kendal, S., & Creen, M. (2021). *An Introduction to Knowledge Engineering*. Google Books. Retrieved from <https://books.google.nl/books?id=9Z5XRoy3TQ8C&pg=PA83&hl=nl&source=gbs_toc_r&cad=3#v=onepage&q&f=false>.