Knowledge Representation in Big Data

1. Title page
2. Definition of Knowledge Representation:

Knowledge representation in big data refers to the process of information extraction and knowledge representation, via knowledge processing and analytics to visualization, and practical applications. It is used to represent knowledge concepts to organize the knowledge instances in a graph-based knowledge base. Knowledge representation is also used in data mining.

1. Importance of Knowledge Representation in Big Data:  
   The importance of knowledge representation in big data lies in its ability to represent real-world objects and their relationships in a form that is suitable for processing by an intelligent agent (software) which, through performing “inference steps” will eventuate to specific results. Correct representation of knowledge is considered the big first step in designing an algorithm which can extract it from the input dataset – especially in the case of Big Data. Knowledge representation is used to infuse data with scientific concepts and understanding in order to maximize its utility for furthering scientific insight.
2. Knowledge Representation Techniques:

There are several knowledge representation techniques used in big data. Some of them are:

* Semantic web-based knowledge representation
* Rule-based knowledge representation
* Logic-based knowledge representation
* Ontology-based knowledge representation

1. Semantic-based:

A semantic network is a graph whose nodes represent concepts and whose arcs represent relations between these concepts. They provide a structural representation of statements about a domain of interest. Semantic networks provide a means to abstract from natural language, representing the

knowledge that is captured in text in a form more suitable for computation.

1. Rule-based:

Another natural form of expressing knowledge in some domain of interest are rules that reflect the notion of consequence. Rules come in the form of IF-THEN-constructs and allow to express various kinds of complex statements. Rules can be found in logic programming systems, like the language Prolog, in deductive databases or in business rules systems. However, the intuitive reading with natural language phrases is not suitable for computation, and therefore such phrases are formalised to predicates and variables over objects

of the domain of interest.

1. Logic-based:

Logic-based representation allows to describe the domain of interest as consisting of objects, things that have individual identity, and to construct logical formulas around these objects formed by predicates, functions, variables and logical connectives. More complex restrictions that range over larger fragments of a network graph can be formulated in logic, where the intuitive graphical notation lacks expressivity.

1. Ontologies:

An ontology is a formal explicit specification of a shared conceptualisation of a domain of interest. Technically, the principal constituents of an ontology are concepts, relations and instances. We can think about it as data base scheme that defines relations between instances and concepts on different levels of abstraction.

1. Challenges of Knowledge Representation in Big Data:

* Scalability: As the amount of data increases, it becomes more difficult to represent and process the data.
* Diversity: Big data is often diversity, meaning that it comes from different sources and is in different formats.
* Complexity: Big data is often complex, meaning that it contains many different types of information that are difficult to represent and process.
* Incompleteness: Big data is often incomplete, meaning that it does not contain all the information that is needed to make accurate predictions or decisions.
* Uncertainty: Big data is often uncertain, meaning that there is a degree of uncertainty associated with the data.

1. Applications of Knowledge Representation in Big Data:

Knowledge representation has many applications in various fields such as artificial intelligence, robotics, machine learning, formal languages and other research areas. Knowledge representation is used to capture information about the world that can be used for solving complex problems. Some of the applications of knowledge representation are:

* E-commerce
* Manufacturing
* Telecommunications
* Robotics
* Natural language processing
* Semantic web

1. Thanks for attention!