1. Introduction

Big Data is the term used to describe the enormous volume of organised, semi-structured, and unstructured data that is produced at an unprecedented scale and velocity from numerous sources. This data may originate from digital platforms including social media, IoT devices, sensors, and others. To extract useful insights and make wise judgements, big data analysis demands sophisticated tools and processes that can handle the volume, velocity, diversity, veracity, and value of the data. Big Data's ultimate objective is to transform raw data into usable insights that can enhance procedures, spot opportunities, and spur innovation.

1.1 5 Vs of Big Data:

Volume: The volume of data being generated is indicated by this term. There is a tremendous amount of data being produced due to the rising number of devices and the rising volume of data they produce. This volume is still increasing exponentially.

Velocity: This describes how quickly data is created, processed, and analysed. In order to get insights and make wise decisions, data must be processed quickly as it is produced at a rapid rate. Tools for effective and quick data processing are required.

Variety: This is the range of generated data kinds, such as structured, semi-structured, and unstructured data. The information can come from a variety of sources, including social media, sensors, and other IoT devices, and it can take the form of text, photographs, videos, or audio.

Veracity: This is used to describe the precision and dependability of the data being produced. Big Data is frequently imperfect, noisy, or inconsistent, therefore it's crucial to make sure the information being utilised for analysis is correct and trustworthy.

Value: This is the usefulness of the conclusions drawn from the data. Big Data's overarching objective is to uncover insightful knowledge that can assist businesses in making wise decisions, streamlining procedures, and gaining a competitive edge.