Report

A data stream is a massive, continuous, and rapid sequence of data elements. As Data mining is considered as a process of discovering useful patterns beneath the data, also uses machine learning algorithms. There have been techniques which used computer programs to automatically extract models representing patterns from data and then check those models. This streaming information has the following characteristics. 1)The data arrives continuously from data streams, 2) No assumptions on data stream ordering can be made, 3) The length of the data stream is unbounded. From the well - established statistical and computational approaches the problems of mining data streams can be solved using the methodologies which uses: 1) Examining a subset of the whole data set or transforms the data to reduce the size, 2) Algorithms for efficient utilization of time and space.

Challenges: Streaming Data is Very Complex. Streaming data is particularly challenging to handle because it is continuously generated by an array of sources and devices and is delivered in a wide variety of formats. Business Wants Data, But IT Can't Keep Up. Mining big data streams faces three principal challenges: **volume, velocity, and volatility**. Volume and velocity require a high volume of data to be processed in limited time.
Advantages: Data streams allow an organization to process data in real-time, giving companies the ability to monitor all aspects of its business. The real-time nature of the monitoring allows management to react and respond to crisis events much quicker than any other data processing methods.
Disadvantages: 1) Cost, data mining involves lots of technology in use for the data collection process, 2) Security, Identity theft is a big issue when using data mining. 3) Privacy, when using data mining there are many privacy concerns raised, 4) Accuracy, 5) Technical Skills, 6) Information Misuse.
Applications: Analyzing data streams helps in applications like scientific applications, business, and astronomy, etc. Examples of data streams include computer network traffic, phone conversations, ATM transactions, web searches, and sensor data. Data stream mining can be considered a subfield of data mining, machine learning, and knowledge discovery.

Conclusion: The dissemination of data stream phenomenon has necessitated the development of stream mining algorithms. We can conclude that most of the current mining approaches use one passes mining algorithms and few of them even address the problem of drifting. The present techniques produce approximate results due to limited memory.