UNIT-5

TEXT ANALYTICS:

Text analytics is a sub-set of Natural Language Processing (NLP) that aims to automate extraction and classification of actionable insights from unstructured text disguised as emails, tweets, chats, tickets, reviews, and survey responses scattered all over the internet.

Text analytics or text mining is multi-faceted and anchors NLP to gather and process text and other language data to deliver meaningful insights.

Maintain Consistency: Manual tasks are repetitive and tiring. Humans tend to make errors while performing such tasks – and, on top of everything else, performing such tasks is time-consuming. Cognitive biasing is another factor that hinders consistency in data analysis. Leveraging advanced algorithms like text analytics techniques enable performing quick and collective analysis rationally and provide reliable and consistent data.

Scalability: With text analytics techniques, enormous data across **social media**, emails, chats, websites, and documents can be structured and processed without difficulty, helping businesses improve efficiency with more information.

Real-time Analysis: Real-time data in today's world is a game-changer. Evaluating this information with text analytics allows businesses to detect and attend to urgent matters without delay. Applications of Text analytics enable monitoring and automated flagging of tweets, shares, likes, an spotting expressions and sentiments that convey urgency or negativity.

TECHNIQUES :

Text classification:

The process of assigning categories and tags to unstructured text through natural language processing. This helps to organize and structure textual data to streamline information management and deliver quick insights. Sentiment analysis is a form of text classification. Machine learning algorithms can assign tags such as "positive" or "negative" to different words in a conversation to determine how satisfied a customer is

Text extraction:

This is the act of extracting pieces of data from textual data. You could extract things like keywords (trending terms that appear most often), company names, prices, and more. Keyword extraction is often used for creating word clouds and insights into brand association. Marketing teams can use text extraction to find frequently discussed topics

Word frequency:

This is a technique used to measure the most frequently occurring words and phrases in specific conversations. For instance, you could use text analytics and word frequency to find out which features your customers mention most often when having a sales discussion with your team. Collocation and concordance can help to identify the words that usually occur at the same time and the context of those phrases

Named Entity Recognition:

Named Entity Recognition, or NER is a form of text analytics used for identifying named components like events, organisations, places, or people in unstructured text. The NER technology extracts the nouns from the text to identify the value of the nouns. For instance, search engines and recommendation tools use NER to help with information retrieval

Clustering:

Text clusters allow text analytics systems to understand and group large amounts of unstructured data. Though clusters aren't as accurate as classification algorithms, they are faster to implement. This makes it easier for smart algorithms to collect information and make predictions often without the use of supervised machine learning

APPLICATIONS:

- Social Media Listening
- Sales & Marketing
- Brand Monitoring
- Customer Service
- Business Intelligence
- Product Analytics
- Knowledge Management
- Email Filtering