Abstract

In today's extremely developed world, each minute, individuals round the globe specific themselves via numerous platforms on the net. And in every minute, an enormous quantity of unstructured information is generated. This information is within the style of text that is gathered from forums and social media websites. Such information is termed as massive information. User opinions square measure associated with a good vary of topics like politics, latest gadgets and merchandise. Social Networking sites provides tremendous impetus for large information in mining people's opinion. Public API's catered by sites like Twitter provides North American nation with helpful information for studying writer's perspective in terms of a specific topic, product etc. To distinguish people's opinion, tweets square measure labeled into positive, negative or neutral indicators. This paper provides an efficient mechanism to perform opinion mining by coming up with a finish to finish pipeline with the assistance of Apache Flume, Apache HDFS, and Apache Hive. Here we proposed to develop a opinion Analysis mechanism to analyze the various polarity of opinions of Twitter users through their tweets in order to extract what they think.
Here we have used dictionary based approach for analysis for which we have implemented hive queries through which we can analyze these complex twitter data to check polarity of the tweets based on the polarity dictionary through which we can say that which tweets have negative opinion or positive opinion.

References


Index Terms

Computer Science  Information Sciences
Keywords

Opinion mining, hadoop, apache flume, hive, Dictionary based approach, bigdata.