Consumer behaviour analysis using User Generated Content

Problem statement

Now-a-days social media plays a key role in almost all the areas. Many kind of analysis can be done based on data obtained from user generated content in the form of reviews, comments related to some product or service. Customer behaviour analysis is one such field which is helpful to understand or predict the customer intentions and it can used for tune organizational directions. For example, by analysing the search patters (time spent/item, search frequency and return visits) of users on a website, an online merchant can have a good understanding the requirements, behaviour and intentions. So, the objective or problem statement here is to study the consumer behaviour (like hotel booking, online purchasing, travel behaviour etc...) from the user generated content in un-structured for using machine learning algorithm. An example can be taken in this context for predicting the online purchasing behaviour from user generated content (UGC).

Background

In the areas of e-Commerce, user generated data can be considered as a new mouth publicity for products or services. UGC can be positive or negative statements made by current or former customers and it is more trusted among the new customers. Clickstream data can be used to track the search behaviour of customers with the help of machine learning algorithms. Now-a-days search engines vastly use the machine learning algorithms like logistic regression, artificial neural network to predict the user activity. Apart from this, other deep neural architecture has more modelling strength which results in more accuracy and efficiency. Therefore, in general, online user generated content impact the behaviour of consumers to take purchasing decision in case of online purchasing.

Methodology

The architecture for customer behaviour analysis using UGC is shown in figure 1.

Step 1: Collection of relevant data

> Data is collected from reviews and comment obtained from different user about product or services.

Step 2: Pre-processing and transformation of data

> Transform the data into a form which can be used by a machine learning algorithm. Feature selection can be done by omitting certain features from the dataset which are not useful and detriment the performance of the learning model. Apart from this the data can be broken down into training, test set.

Step 3: Machine learning algorithm selection for training the model

> Any of the supervised machine learning algorithms like ANN or decision trees or deep neural architectures can be used.

Step 4: Testing and estimating the efficiency of the model.

Customer behaviour prediction is tested using the test dataset. Deep learning techniques can be used for training high accuracy prediction model which can be used to generate more accurate prediction in real time.

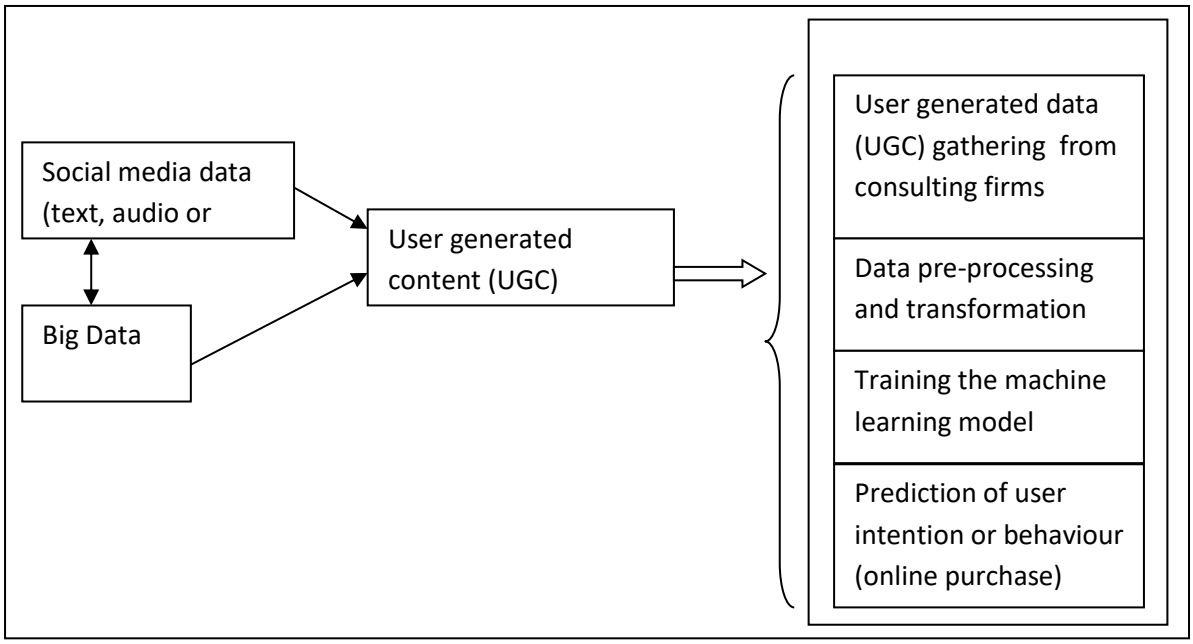

Figure 1. Architecture for customer behaviour analysis using user generated content

In the above-mentioned architecture, social media and big data are positively associated to UGC. UGC is associated to user behaviour (Purchasing behaviour, travel behaviour or hotel booking behaviour etc...). Big data is positively associated to customer purchase behaviour.

Experimental Design

Dataset
The most identified data source can be social media and big data through various websites in the form of reviews, comments created by several online users related to some product or services. This data can be obtained from consulting firms.

Evaluation Measures
Comparison of efficiency and accuracy of the propose model with other traditional method used for this kind of analysis.

Software and Hardware Requirements
Python programming and deep learning libraries as required.