Detecting influencer on twitter across different genres

Springboard Capstone 2 project proposal

Problem

Word-of-mouth (WOD) business strategy has been regarded as an important mechanism by which information about new products or brands or topics can reach large populations. Social media helps people to connect with other people anywhere and let them share their opinions regarding any matter. For the WOD marketing, companies need to find the social media influencers to spread information about their new products or brand more effectively. Twitter is one of the most popular microblogging platforms allowed to publicly discuss various topics or products using tweets. In this project, I propose building a supervised machine learning model to detect the Twitter influencer based on tweets across different categories. For this project I will be using tweets from three different hashtags: #fashion, #fitness and #travel tweets. Number of followers of Twitter users provide the labels for training ML models to predict the influencer based on their tweets regarding a particular product or brand.

Clients

Social media companies like Twitter, Instagram, YouTube etc can use my machine learning model to predict the influencer based on their social media account information.

Data Source

For this project, I will scrape the tweets from Twitter.com for three different hashtags: #fashion, #fitness and #travel. The dataset contains information about user_name, user_description, location, following (the number people the user is following), number of followers, total_tweets (the total tweets of the user), user_create_date, tweet_create_date, retweet_count and text (tweet).

Project Outline

In this project I will predict the Twitter influencer based on their tweets and number of followers by building supervised machine learning models. I will predict the influencers for three different hashtags tweets.

Deliverables:

I. Code as jupyter notebook:

For Data Acquisition, Data Wrangling, Data Exploration & Analysis, Machine Learning Model Development

II. Report on the project

III. Presentation on the project