# Crowd Prediction System for Tourists

Abstract

Tourism Industry is one of the largest industries in the world with a global economic contribution of over 7.6 trillion USD in 2016 as mentioned on statista website. One of the major problems faced by tourist all over the world right now is that of crowd. Popular tourist destinations tend to get extremely crowded and tourists plan their trips without any insights. Research solutions till date propose results based on static data from previous years which is not cost effective. This model plans to provide a report predicting the crowd density of particular location with respect to the dates mentioned by the user. After taking a glance at the report the user can decide for himself, whether to proceed with the plan or pick up a new location. The model can also suggest similar locations as per the expectations of the user which may be less crowded during the time frame provided.
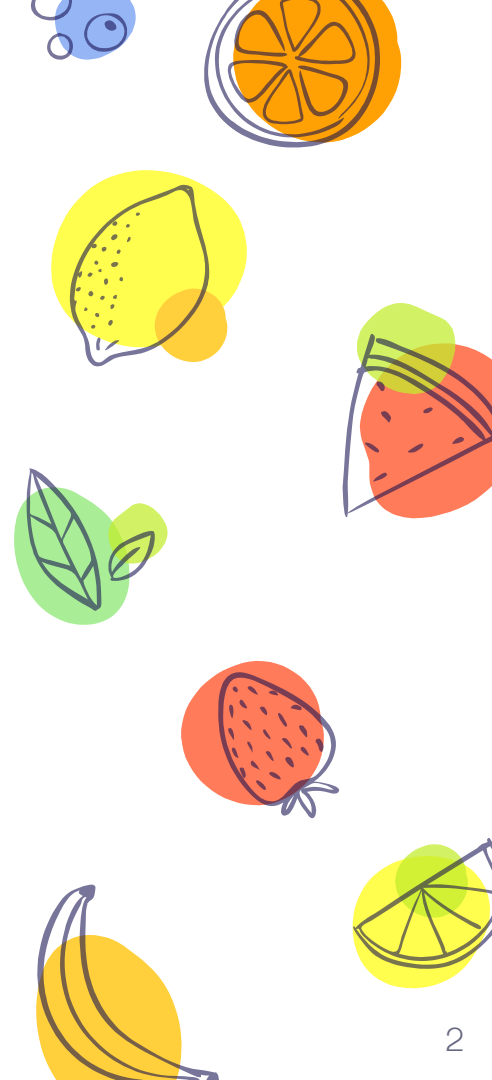
Introduction

US$ 7.6 Trillion

Contribution to the global
economy directly or indirectly.

Introduction

All over the globe, there has been a steady growth in the tourism industry where International tourist arrivals increased from 528 million USD in 2005 to a staggering 1.19 billion USD in 2015 [1]. In the past couple of decades, we witnessed an accelerated technological progress characterized not only by new innovations, but their application and diffusion also mattered.

The word 'Tourist' is coined for individuals traveling for recreation. Growing amount of tourists at famous locations can now be seen as an obstacle for recreation. With new technologies which are available right now, it is possible to predict the crowd density of any given location with respect to the dates provided. The average crowd density for a frame of 3 months will also be calculated with the dates provided serving as the median for the frame. Then both the values i.e average value and the predicted value will be compared for the report to be generated. Deviation of the predicted value and the average value will help us in predicting crowd density of the given time frame.

Introduction contd..

Crowd occurs due to regional exquisites, festivals and important events which attracts large crowd that demands immediate and instantaneous service. Understanding the crowd is a difficult task, therefore our data set will use information from Social Media, Hotels and Travel companies. Most of the websites nowadays share some of their information for public use which can be used through APIs (Application programming interface). Hashtags, location updates and mentions from social media platforms such as Facebook, Twitter and Instagram will considered as input sources. Availability and prices of respective rooms of hotels at the selected location. Flight, train fares and seat availability will also be considered. Averaging out all the listed data set values and then comparing with output values of the desired date, the model will try to predict the crowd density.

Objectives

The primary objective is to create an automated system that will fetch information from different platforms such as social media, travel websites and then understand it to predict the crowd footfall at requested location on provided dates by the user. To attain this successfully, the system must meet the following goals.

1) To retrieve data dynamically from various platforms mentioned above to form the datasets for the system.

2) To understand the datasets and then calculate the crowd density at the requested location on provided dates by the user.

3) To provide multiple alternative locations if the expected crowd footfall is much more than the average footfall.
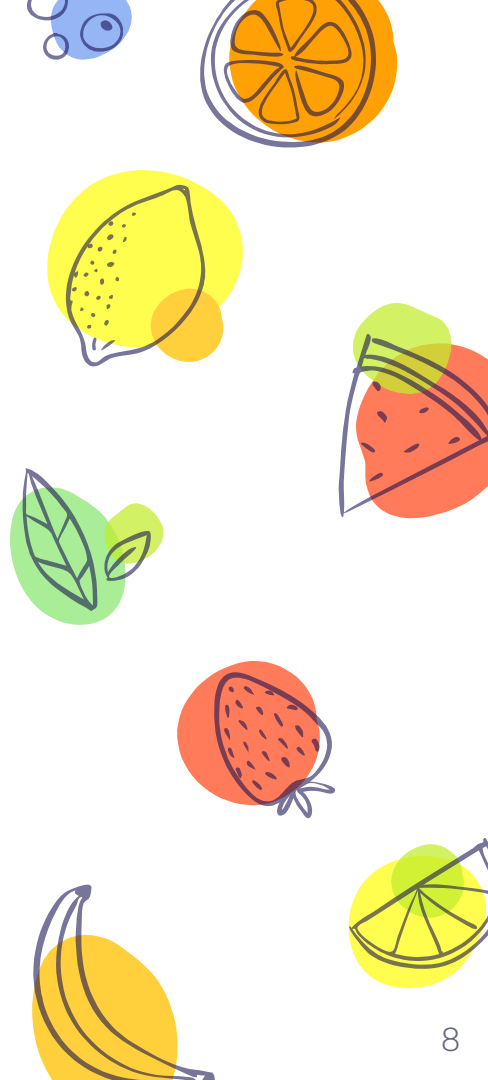
Literature Review Pt. 1

To generate personalized insights for the user, RS is important as it can help tourists save energy and time while planning for their trip.[1][2].

Now coming to social media, Twitter is a famous micro-blogging platform. On twitter people share tweets of maximum 140 characters. Generalizing all the tweets, we can further divide them broadly into two categories. First being the tweets where the user tweets about themselves and second where information is shared using the tweet[3].

This information shared is a very important source for collecting data and determining twitter sentiments.

Literature Review Pt. 2

Sentiment analysis is useful in commercial intelligence application environment and recommender systems [8], [9]. B. Liu (2012) defined sentiment analysis, which is also called opinion mining as the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [5].

Based on the definition, data can be then sub-divided into positive, negative or neutral.

Liebrecht et al. [4] discusses how sarcasm plays an important role of changing the polarity of a message and how it conveys a negative opinion using only positive words. Therefore, detection of sarcasm is important for the development of sentiment analysis systems

Literature Review Pt. 3

Maynard and Greenwood [6] have put forward a set of rules to determine whether a tweet is positive or negative when sarcasm is detected. Hashtags are used as labels on the internet. Hashtags gained popularity on twitter itself [7].

Maynard and Greenwood [6] discuss how hashtags contain much useful sentiment information, but how it is difficult to identify tags as they contain multiple words. They have developed an algorithm to extract the individual words from a single hashtag.

References Pt. 1

[1] A.Moreno, L.Sebasti and P.Vansteenwegen, "Recommender Systems in Tourism", 2015.

[2] Kevin Meehan, Tom Lunney, Kevin Curran, Aiden McCaughey, "Context-Aware Intelligent Recommendation System for Tourism", 2013.

[3] Johan Bollen, Huina Mao, and Alberto Pepe., "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena", In International AAAI Conference on Weblogs and Social Media (ICWSM'11) 2011.

[4] C. C. Liebrecht, F. A. Kunneman, and A. P. J. van den Bosh, "The perfect solution for detecting sarcasm in tweets hashtagnot ", In Proc. WASSA, pp. 2937 , Jun. 2013.

[5] Bing Liu, "Sentiment Analysis and Opinion Mining", In Morgan and Claypool Publishers May 2012.

References Pt. 2

[6] D. Maynard and M. A. Greenwood, "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis", In Proc. 9th Int. Conf. Lang. Resour. Eval., pp. 42384243. May 2014.

[7] Godin, Frderic, Viktor Slavkovikj, Wesley De Neve, Benjamin Schrauwen, and Rik Van de Walle., "Using topic models for twitter hashtag recommendation", In Proceedings of the 22nd International Conference on World Wide Web, pp. 593-596. ACM, 2013.

[8] L. Terveen, W. Hill, B. Amento, D. McDonald, and J. Creter, "PHOAKS: A system for sharing recommendations", In Commun. ACM, vol. 40, no. 3, pp. 5962, 1997.

[9] J. Tatemura, "Virtual reviewers for collaborative exploration of movie reviews", In Proc. 5th Int. Conf. Intell. User Interfaces, 2000, pp. 272275. 2000.
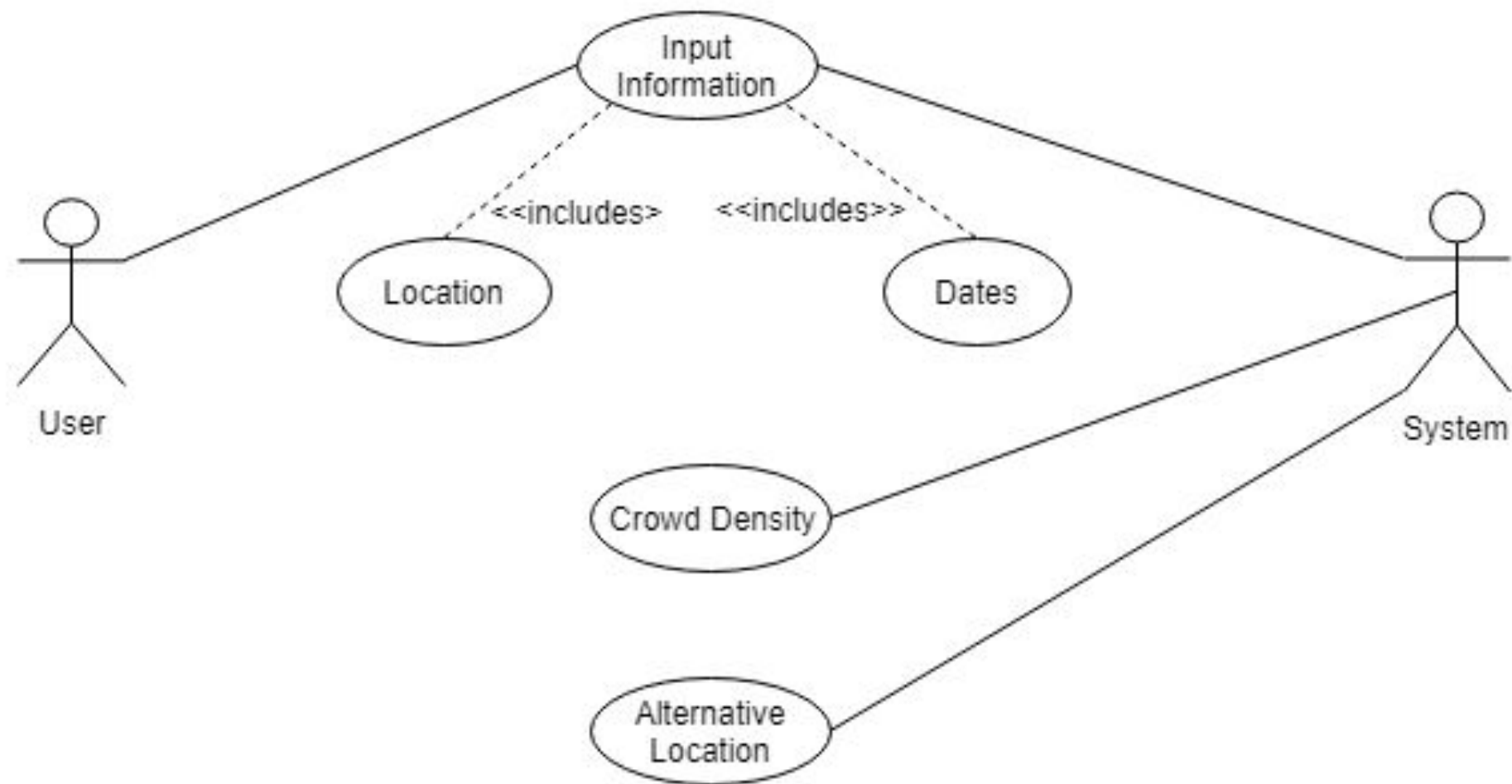
Problem Definition

A crowd density predicting model which will understand information provided by the datasets and then predict the crowd footfall at the requested location on provided dates by the user. If it gets too crowded then the model will suggest similar locations which will have less crowd but same characteristics
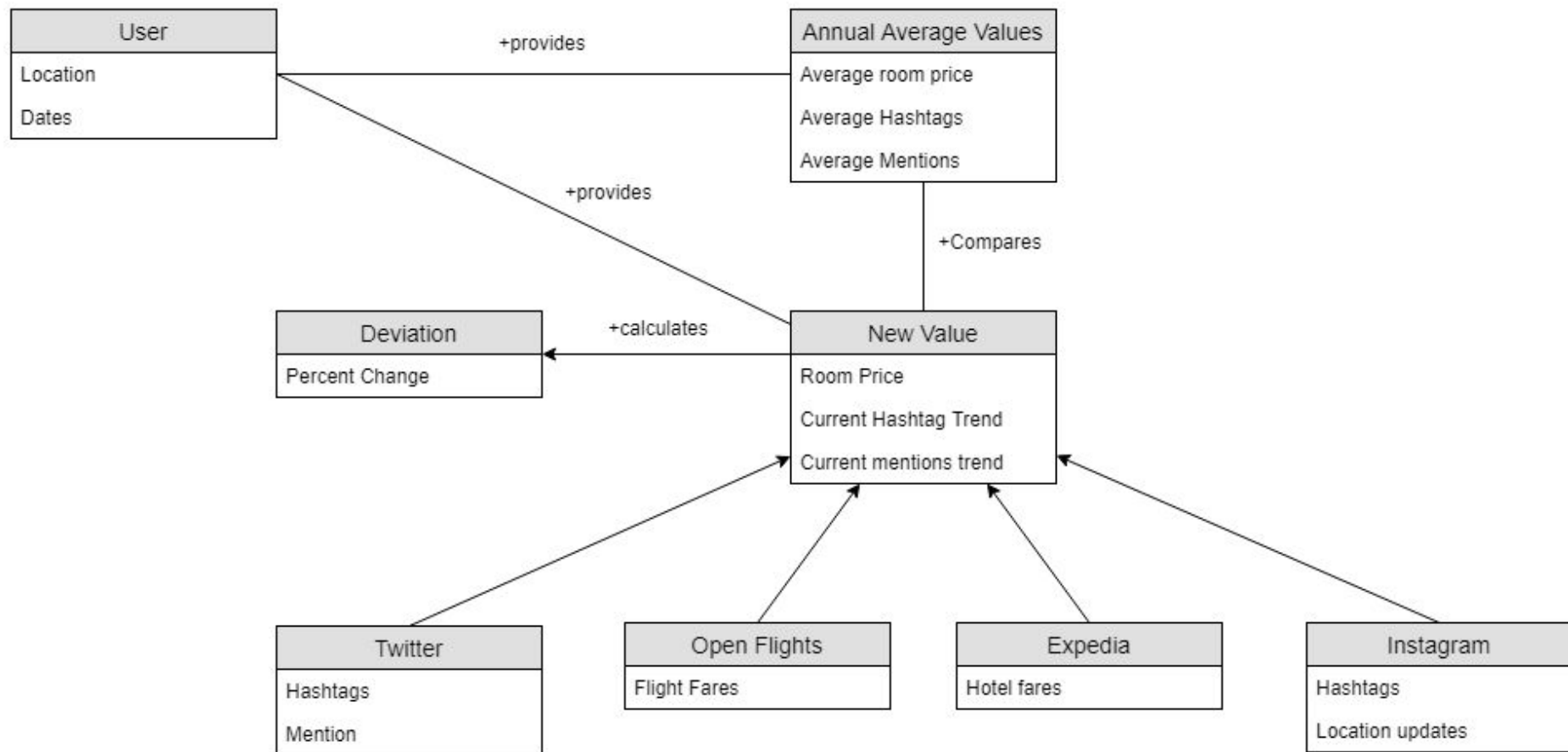
Use Case Diagram

Class Diagram

Summary

A new approach for predicting crowd density is proposed by using this prototype model, which could be used as an application to grow tourism by suggesting the appropriate locations for the right time frame. It uses dynamic values, rather than static values of data as source of information for generating datasets. Implementing machine learning technologies on these data sets, the model will be able to predict the crowd density. A subordinate function will execute if the crowd density surpasses the annual average values providing with an alternative location with same characteristics. Depending on the success rate of use cases, we propose two strategies, namely social media and travel fare prices.
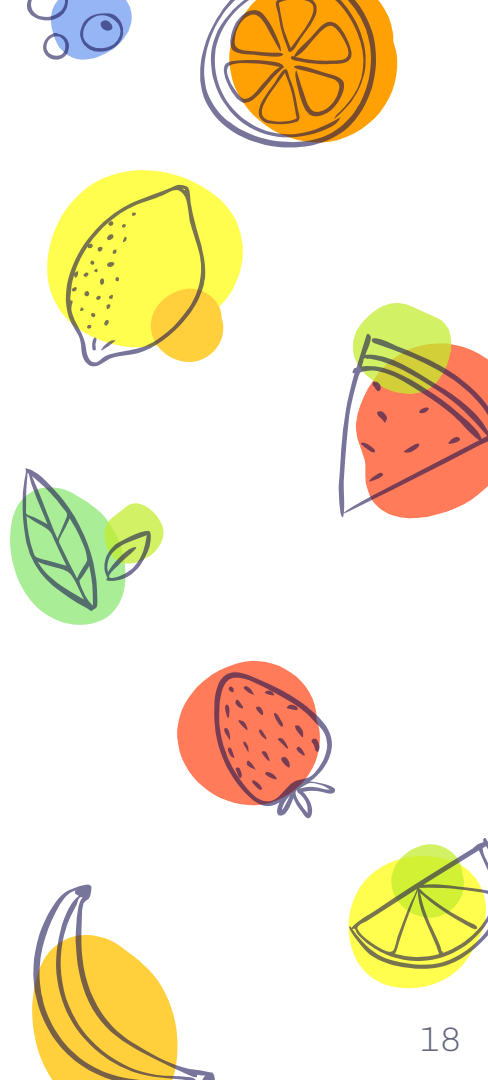
Made By
Neil Shringarpure
Kartik Kokane
Rohit Dhuri
Pratik Jogdand

Under the Guidance of
Prof. Sachin Takmare

Thanks!

Any questions?