RENTAL BIKE SHARING DEMAND PREDICITON BY USING MACHINE LEARNING

D.USHARAJESWARI[1], K.MADHAN SAMPATH [2]
I. ASWANTH KUMAR [3], K. SNEHARIKA [4]
K. SANTOSH VAMSI KRISHNA [5]

1. Assistant Professor 2. Engineering student, 3.Engineering student
4. Engineering student, 5. Engineering student
DEPARTMENT OF COMPUTER SCIENCE,
SANKETIKA VIDYA PARISHAD ENGINEERING COLLEGE
Visakhapatnam, India

ABSTRACT

Bike-sharing services have proliferated in recent years. A number of things, including surrounding activities, road closures, and traffic regulations on campus, can have an impact on the use of bicycles. The study examined how usage of the Great Rides Bike Share program in Fargo, North Dakota, U.S.A. was affected by weather conditions (average temperature, total daily precipitation, average wind speed, and weather outlook), day of the week, holiday or weekday, month, and season. Using Bayesian techniques and decision trees, this study also aimed to forecast the Great Rides Bike Share program's rental demand in 2016. Additionally, the relative significance of the causative characteristics was evaluated. Decision trees were found to be an effective tool for forecasting the 2016 demand.

Keywords:- Bike-sharing, Decision trees, Great Rides Bike, traffic regulation.

1 INTRODUCTION

There are currently bike-sharing schemes in more than 500 cities in 49 nations. "Bike sharing has experienced the fastest growth of any mode of transport in the history of the planet," according to urban transport analyst Peter Midgley. By employing easily identifiable specialty bicycles with distinctive parts that would be of little value to a thief, tracking the locations of the cycles using GPS or radio frequency, and requiring credit card payment or smart card-based membership to check out bikes, modern bike-sharing systems have significantly reduced the theft and vandalism that hampered earlier programs. For the majority of systems Riders can pick up a bicycle that is locked to a clearly marked bike rack or electronic docking station for a brief ride (usually an hour or less) at no additional cost after paying a daily, weekly, monthly, or annual membership fee. They can then return the bicycle to any station in the system. To optimize the number of available bikes, riding for longer than the program's allotted time usually results in additional expenses. The following factors contribute to the growing popularity of bike-sharing programs:

• They increase transit use because of the additional bike transit trips, the better connectivity to other modes of transportation due to the first-mile/last-mile problem that bike-sharing helps solve, and the reduction in the number of personal vehicle trips.

• They lower greenhouse gas emissions and enhance public health. Since these bike-sharing programs are growing more and more well-known worldwide, it is crucial to examine these systems from many angles. My main focus is forecasting the demand for bike rentals in 2016 for the Fargo, North Dakota-based Great Rides Bike Share system. Fargo's Great Rides is a seasonal system with 101 bicycles and 11 stations. More bikes were used per capita in 2015—143,000 trips and an average of six to seven daily rides—than in New York, Washington, D.C., or Paris [3]. The program's integration with student IDs is primarily responsible for its success; at North Dakota State University (NDSU), the Great Rides seasonal pass is part of the required student activity costs. Success in the data-driven business environment of the mobility industry hinges on the ability to effectively predict client demand (Sohrabi et al., 2020; Wessel, 2020). These days, thanks to data technology, it's fairly simple to locate global corporations offering a range of services to clients based on demand prediction outcomes. Companies have succeeded in the short term because they were able to accurately forecast demand based on both internal and external factors.

2. LITERATURE SURVEY AND RELATED WORK

Over the past 45 years, bike share programs have existed in three generational forms (DeMaio, 2003; DeMaio & Gifford, 2004). The first generation of a bike sharing program was launched by Witt Fiesten in Amsterdam in the Netherlands on July 28, 1965. A typical bike type, painted white, was provided for public rental use. Anyone could use the program and, if it was placed near a destination for the next user, it could offer continuous use, but it did not function properly and ended within a few days (DeMaio, 2009). The second generation of the bike share program was implemented in Denmark in 1991 and 1993 (Nielsen, 1993). It was a small program consisting of twenty-six bike rental service types at four stations and later hosted a massive program in Copenhagen called City Bike in 1995. Copenhagen Bike was specifically built for practical purposes using hard rubber tires and wheels with advertising plates. It could even be returned with the payment and receipt of coin deposits at certain locations throughout the city. The program was formalized when it was managed by a station company and a non-profit organization but it still suffered from theft, which resulted in the creation of a next-generation bike-sharing program with improved tracking capabilities (DeMaio 2009). The third generation of the bike-sharing program was the Bike about program at the University of Portsmouth in England, which was implemented in 1996. Since University of Portsmouth students used electronic cards to rent bikes (Black et al., 1998), the bike share system became smart through various technological improvements, including electronic locking and communication systems, smart card and mobile phone access, and on-board computers (DeMario, 2009).

The bike share program has grown slowly in the third generation, with one to two new bike share programs starting each year, and 1,500 bikes were implemented by JCDecaus in Lyon, France (Optimising Bike Sharing in Europe Cities, 2009). This was the largest of the third-generation bike-sharing programs. 15,000 members used bikes 6.5 times a day. Paris, the capital of France, also used the occasion to provide the program attention (Henley, 2005). Two years later, Paris launched its own bike-sharing program and grew in the city and suburban areas from about 7,000 to 23,600 bikes. This large-scale implementation has dramatically changed the flow of bike-sharing programs around the world. It has also been implemented outside Europe since 2008, with large numbers of programs in Brazil, Chile, China, New Zealand, South Korea, Taiwan, and the United States. As a result, there were about 60 bike-sharing programs in each country by the end of 2007 and there were about 60 third-generation bike programs worldwide (DeMaio, 2007). At the end of 2008, there were 92 Bike sharing programs (DeMaio, 2008). Bike-sharing programs have become more popular worldwide. Most bike-sharing programs are fourth-generation systems, which add demand response and multi-modal systems to third-generation systems (Parkes et al., 2013).

The cumulative demand for bike share programs has been increasing since 2007. There were an estimated 639 bike-sharing schemes operating in 53 countries located in nearly every region of the world in 2017, boasting a total of about 643,000 bikes. In the U.S., the total number of shared bike rides has exceeded 88 million since 2010. In 2016, the number of annual rides exceeded 28 million, which is equal to the annual number of rides on the Amtrak system. The five biggest bike share programs in the U.S. (Citi Bike in New York, Capital Bike share in greater Washington, D.C., Citi Bike in Miami, Divvy in Chicago, and Hub way in Greater Boston) contributed to 85% of all boke trips according to a report published by the National Association of City Transportation Officials (NACTO, 2017). This study used the capital bike share’s business case in the Washington, D.C. area as it was the first mover in bike share systems in the U.S. There are many different reasons of joining bike share programs in the Washington, D.C. area. Getting around quicker and more easily was the primary reason for joining the bike share program. It also indicates that people consider the bike share program an alternative travel options that costs less (LDA Consulting, 2017). Furthermore, Min et al. (2017) proposed an analysis methodology that could identify the meaning of the data using visualization. Locational characteristics of stops were found according to utilization rates, and bike use patterns were studied that vary according to time, day, and month. Meanwhile, the use pattern between stops was found through an analysis of travel route, and the purpose of use was identified through an analysis of the destination ratio for each stop. Based on these data, the development direction of public rental bike systems in Daejeon was presented. Kim et al. (2012) studied the effects of the weather on public bike demand based on public bike data from Goyang-si, Korea with temperature variables showing a decrease in bike use if temperatures fell significantly or rose above 23°C. Precipitation has a negative effect on bike use, falling about sixty percent for every 10cm increase in rainfall. In addition, the number of clouds in the sky also has a negative impact on the use of bikes. These results show a similar trend, although there are differences in values compared to overseas studies. Lee et al. (2011) built a bike demand estimation model with features composed of the number of students, excluding elementary school students, and the number of passenger cars

3 Implementation Study

While some have tried to duplicate the bike rental industry that is booming in Western countries in India, let's have a look at some data regarding the number of users of bike sharing programs. As of August 2014, over 600 towns worldwide have bike rental programs, with a fleet of about 500000 bicycles, mostly in Western countries, according to Wikipedia. Currently, two of the most widely used bike rental services worldwide are NextBike and Cogo BikeShare. When examining the Bike Share sector from an Indian standpoint, the country has not yet adopted the use of this new sector..

3.1 PROPOSED Methodology AND ALGORITHM

In general, it is crucial for the administrators of bike rental systems to know how many bikes will be required at each station. With this information, they can allocate the right amount of bikes at each station and determine whether additional bike stands are required at a specific location.Thus, we examine a variety of prediction algorithms in this research project, including Random Forest, Lasso, Ridge, and Linear regression. The goal of this research project is to determine which algorithm performs best when applied to the real-world problem of predicting bike rental demand. The current systems have not made adequate use of the analytics. In certain instances, there may be a large number of bikes in the station overall but fewer overall reservations.

.

4 METHODOLOGIES

4.1.Load Dataset:

Load data set using pandas read_csv() method. Here we will read the excel sheet data and store into a variable.

4.2.Split Data Set:

Split the data set to two types. One is train data test and another one is test data set.here we will remove missing values from the dataset.

4.3.Train data set:

Train data set will train our data set using fit method. 80% of data from dataset we use for training the algorithm.

4. 4.Test data set:

Test data set will test the data set using algorithm. 20% of data from dataset we use for testing the algorithm.

4.5.Predict data set:

Predict() method will predict the results. In this step we will predict the bike count

4.6 Linear regression:

in Linear Regression Method Algorithm we discussed about an algorithm for linear regression and procedure for least square method. In this article we are going to develop pseudocode for **Linear Regression Method** so that it will be easy while implementing this method using high level programming languages

4.6.1 Pseudocode for Linear Regression:

1. Start
2. Read Number of Data (n)
3. For i=1 to n:
Read Xi and Yi
Next i
4. Initialize:
sumX = 0
sumX2 = 0
sumY = 0
sumXY = 0
5. Calculate Required Sum
For i=1 to n:
sumX = sumX + Xi
sumX2 = sumX2 + Xi * Xi
sumY = sumY + Yi
sumXY = sumXY + Xi * Yi
Next i
6.Calculate Required Constant a and b of y = a+bx
b = (n * sumXY - sumX * sumY)/ (n*sumX2 - sumX * sumX)
a = (sumY - b*sumX)/n
7. Display value of a and b
8. Stop

4.6.2 Lasso regression:

S1 File. Pseudo code of fast Lasso regression using coordinate descent based on covariance updates

Input:

Problem: arg min$_\beta$

ipy: Inner product vector, $ipy_i = < y, X_{\cdot i} >$

ipx: Inner product matrix, $ipx_{ij} = < X_{\cdot i}, X_{\cdot j} >$

λ: Penalty parameter

N: Number of samples

Output:

$beta$: Regression parameter vector

function $FastLass(\ ipy, ipx, \lambda, N)$
$stop\_thr$ # Threshold for stopping iteration
$\leftarrow length\ \ ipy$
$beta \leftarrow 0$ with length p
$gc \leftarrow 0$ with length p # Gradient component vector
do
$difBeta_{max} \leftarrow 0$
for $j = 1 \rightarrow p$ **do**
$z \leftarrow \ ipy\ j - gc\ j \quad N + beta\ j$
$beta\_tmp \leftarrow$ max $0, z - \ -$ max $(0, -z - \lambda)$
$difBeta \leftarrow beta\_tmp - beta\ j$
$difabs \leftarrow abs\ difBeta$
if $difabs > 0$ then
$beta\ j \leftarrow beta\_tmp$
$gc \leftarrow gc + ip, j \times difBeta$ # Update gradient components
$difBetamax =$ max $difBetamax, difabs$
end if
end for
while $difBetamax \geq stop\_thr$
end do-while
return $beta$
end function

4.6.2 RANDOM FOREST REGRESSOR ALGORITHM:

Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees. Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap. An ensemble method is a technique that combines the predictions from multiple machine learning algorithms together to make more accurate predictions than any individual model. A model comprised of many models is called an ensemble model.

. Import Libraries: Import the necessary libraries, including scikit-learn's RandomForestRegressor.

Load Data: Load your dataset into X (features) and y (target variable).

Split Data: Split the dataset into training and testing sets using train_test_split.

Initialize Model: Initialize the RandomForestRegressor with desired hyperparameters. Here, n_estimators is set to 100, which is the number of trees in the forest.

Train Model: Fit the Random Forest Regressor to the training data.

Make Predictions: Use the trained model to make predictions on the test set.

Evaluate Model: Evaluate the model's performance using mean squared error (MSE) between the actual and predicted values.

Visualize Feature Importances (Optional): If desired, visualize the feature importances to understand which features are most influential in making predictions

5 RESULTS AND DISCUSSION

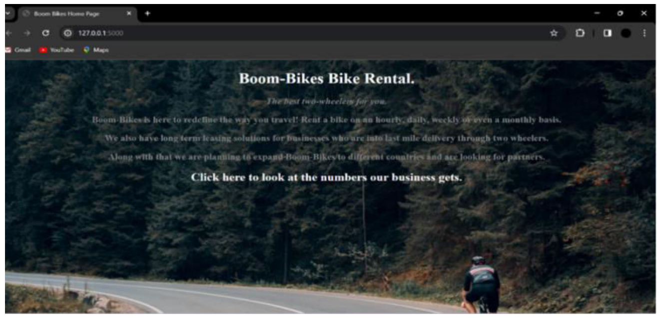

FIG 1:-home page

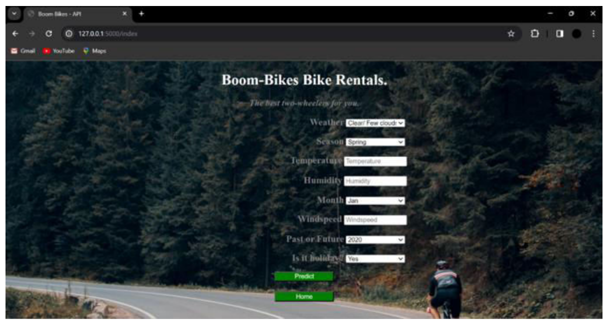

FIG 2 :-input page

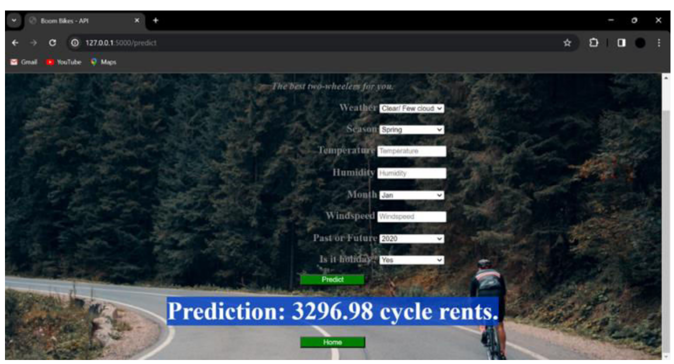

FIG 3 predicted output

6.CONCLUSION AND FUTURE SCOPE

This project's objective was to predict the demand for Fargo's Great Rides program's bike rentals for the 2016 riding season. Bicycle users' behavior can be influenced by a variety of things, such as an event at the Fargo Dome, a marathon nearby, road closures close to campus, or a temporary traffic policy. This project concentrated on predicting the daily demand for the bike-share program using available attributes for weather conditions (outlook), average temperature, average wind speed, total daily precipitation, workday/holiday, day of the week, month, and season. This was because it is impossible to take into account all the factors in a single study. The demand for the 2016 season was accurately predicted by the decision-tree model developed by J48, and this decision-tree structure's accuracy

7 REFERENCES

[1] J. Larsen. (April 25, 2013). Plan B Updates - 112: Bike-Sharing Programs Hit the Streets in Over 500 Cities Worldwide. Available: http://www.earth-policy.org/plan_b_updates/2013/update112

[2] P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," Journal of Public Transportation, vol. 12, no. 4, p. 3, 2009.

[3] A. Corbin. (May 19th, 2016). Why the country's best bike share might be in Fargo-Better-Bike-Share.Available: http://betterbikeshare.org/2016/05/19/countrys-best-bike-share-might-fargo/

[4] S. B. KOTSIANTIS, "Supervised Machine Learning: A Review of Classification Techniques.," in Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies, vol. 160, K. k. llias Maglogiannis, Manolis Wallace, John Soldatos, Ed. 1 ed. (Frontiers in Artificial Intelligence and Applications, Amsterdam: IOS Press, 2007, pp. 3-24.

[5] Kaggle Bike Sharing Demand. Available: https://www.kaggle.com/c/bike-sharing-demand/data

[6] R. Godavarthy, J. Mattson, and A. Taleqani, "Evaluation Study of Bike Share Program in Fargo, ND," Small Urban and Rural Livability Center2017.

[7] R. P. Godavarthy and A. Rahim Taleqani, "Winter bikesharing in US: User willingness, and operator's challenges and best practices," Sustainable Cities and Society, vol. 30, pp. 254-262, 4// 2017.

Bibliography

DAVULURY USHARAJESWARI: Project guide and has ten years of experience as teacher in colleges affiliated with Andhra University and JNTUK. she obtained her B.Tech in CSE from Regency Institute of Technology, which was affiliated to Pondicherry University and the M.Tech from Acharya Nagarjuna University. She is currently employed as an assistant professor in the Sanketika Vidya Parishad Engineering College's computer science department in Visakhapatnam. Artificial intelligence, machine learning, the Internet of Things, and data mining are among her primary areas of interest.

KARANAM MADHAN SAMPATH: Project Lead, is pursuing B.Tech in Computer Science & Engineering from Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, NAAC with A grade accredited, ISO Certified campus. With interest in programming , he has taken up an academic project on Rental Bike Sharing Demand Prediction By Using Machine Learning.

I. ASWANTH KUMAR: Project Member, is pursuing B.Tech in Computer Science & Engineering from Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, NAAC with A grade accredited, ISO Certified campus. With interest in programming , he has taken up an academic project on Rental Bike Sharing Demand Prediction By Using Machine Learning.

K. SNEHARIKA: Project Member, is pursuing B.Tech in Computer Science & Engineering from Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, NAAC with A grade accredited, ISO Certified campus. With interest in programming , she has taken up an academic project on Rental Bike Sharing Demand Prediction By Using Machine Learning.

K. SANTOSH VAMSI KRISHNA: Project Member, is pursuing B.Tech in Computer Science & Engineering from Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, NAAC with A grade accredited, ISO Certified campus. With interest in programming , he has taken up an academic project on Rental Bike Sharing Demand Prediction By Using Machine Learning.