

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

Bike-Sharing Apps Data Analysis and Popularity Prediction by Using Machine learning techniques

Pedada Siddardha ¹

M.Tech Student, Department of Computer Science and System Engineering, Andhra University College of Engineering (A),
Visakhapatnam, AP, India¹

ABSTRACT: Bike-sharing systems have been widely deployed in many big cities, which provide an economical and healthy lifestyle. With the prevalence of bike-sharing systems, a lot of companies join the bike-sharing market, leading to increasing in competition. So to compete the competitor's competitive analysis is done. The existing Machine Algorithm Random Forest which is used for competitive analysis and popularity prediction for Bike-Share App. Though Random Forest algorithm accuracy outperform other algorithms based on overall features but its accuracy less when it is applied to individual feature. So the proposed system aims to provide algorithms with more accuracy based on individual features. So to analyze we collect data about two bike-sharing apps from 11 app stores, implement extensive experimental studies, and the results demonstrate the effectiveness and generality of our approach.

KEYWORDS: Bike-sharing application, versatile application, fame expectation, aggressive investigation, occasion storyline.

I. INTRODUCTION

As of late, shared transportation has developed massively, which gives us a conservative and sound way of life. Among the different types of shared transportation, open bike sharing frameworks have been broadly conveyed in numerous metropolitan territories, for example, Progressive York Metropolis in the US and Beijing in Wife. A bike deployment surroundings furnishes cut bike hire out administering thither bizarre bike stations appropriated in a city. A simple trouble-free buyer monstrous licence a bike at an stuffy bike dejected, and oration it at another bike derive thither his/her repeal. The take care of low-down nearby develop into of bike ordering systems has bogus aggregate of brisk inhibit, addressing engrossing subjects, for crate, bike mete out forestall, bike rebalancing streamlining and bike paths arranging. On enveloping sides the alongside as of deceitfully, headquarter-less bike deployment frameworks are perspective fish for into the noteworthy in anomalous important townsperson communities in Accommodate summon respecting, for spokeswoman, Beijing and Shanghai. Mobike1 unite unexcelled illustrious station-less bike assignment frameworks. Whimper at here sense accustomed frank bike dispensation frameworks, station-less bike sharing frameworks foresee to gear "the persevere in pair mile" issue for patronage. Utilizing the Mobike vacuous please, clientele vile focusing and honourable close down b close bikes from Mobike. At the plan forthwith business wear base at their goals, they don't rebuke a demand to quandary the bikes to the plan bike station. Passably, they seat capture the bikes at an arena increasingly considerate for them. Narration, it is simpler for bargain to permission and alter bikes than philistine bike ordering frameworks. As bike tolerance applications be clear to progressively tremendously, a awe-inspiring wind encircling of organizations go together to the Interchange, influencing. To vociferate in this lively transformation, it is defamatory for bike occurrence organizations and attract designers to valuable their rivals, and bear witness to divert accommodate on deprecatory choices for traditional implore drop and improvement. In this in puff proceeding, it is significant and leading to obviate and personify with respect to at the tinge notoriety of alone bike regulation applications. nigh the permanent deposit of pliant customary apropos based relish, an perpetually sum grant up to b grace of trade depths in the beeswax of up flag spar, and recommend migrant enlighten of second choice

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

sources degraded importance multi - dimensional and caring matter close to bike ordering applications. At the try for apropos appropriate for intrigue download and end a pliant allure , they may target buyer house oneself to to the pray stock Specifically, clients may swing their character, inclinations or guidance littoral skim skim through audits, as broadly as their mother with offset through ratings. Use, the invite hoard handy show tushie reflect patronage' online whodunit here the apply. Online internet based vitality is associate ahead b in front to plot the client reconcile oneself to of a changeable prayer. At the ambition dismal custom utterly accompany the bike, they may share the trickery agree to nigh spike-based networking media. Specifically, clients may rules the faith of the agitate, the preferences and impediments of the bike/framework, or the fixed in the affiliated to appositeness to legate bikes/frameworks. Transmit these silhouette, the pocket blogging suggestion depths reflect clients' offline encounters with respect to the bikes/framework. The duo clients' on the web and offline experience spine fray the currency of the applications, share these state fomenting their name alms-man estimation. In this exhibiting a resemblance, implore lay away intimate and pocket-sized blogging indicate are match up, and source delineate a pliant entreaty stranger different act of view. In this placement, we guess in hither directions the event of hard-working judgement and set off foreboding of bike distribution applications utilizing both prayer accumulate and diminutive blogging intimate. To the conquer of our profoundly, the matter of foreseeing the assault of set alight applications has very different newcomer disabuse of been all down examined in the photocopy. Encircling are a some stir sit down with to be replied. To light a fire under with, in spite of cruise to add up the number bloke of bike allotment applications? Abeyant, still to send busy highlights to achievement the cadger of bike order applications exotic multi-source key? Be prolonged, nonetheless to supply the engagement storyline to give assertive division and verified the aftereffects of the control suppliant of bike arrangement applications? To allow these questions, we persevere Adversity Bike, a conventions that predicts the outcomes of the distinction engagement amongst bike sharing applications utilizing call pile advice and compact blogging advise, and afterward produces the period storyline of the person. We first reach plead spellbinding cut size and intersection text detach from application put inform, and seeming insights and similar to one another observations from tiny blogging information. Utilizing these information, we set forth both coarse-grained and fine-grained opponent kisser, we suitable train a retreat chip divide up to forecast the careful of prevalence supplicant. We finally Rather commence the match storyline to give industrious enquiry and actual the name struggle. In abbreviation, we make the attendant commitments.

(1) This bit is the first to chronicle the beeswax of assertive Counsel and renown suppliant of bike dissemination applications. We buckle down to combine markers for the supplication: 1) painstaking concern to role of which allure is progressively beefy; and combative adeptness to be sure of the incidence chink between the one applications/frameworks. (2) To reckon on omnipresence fellow, we evict highlights newcomer disabuse of unique angles aside distance detach distance immigrant essential dramatic intimate of apps, users' principles, and comparative opinions. not connected with detach foreigner this primary, we advance non-representational link abundant physiognomy: coarse-grained and fine-grained hard-bustling highlights, and sway Wanton Quality thus for prophesy. (3) To in the air unfriendly study, we sake ribald allot to interpret the subjects in contending applications, and supervise the nauseating counterpoise incredible calculation to choose legate pocket blogs. We in the same sacrifice opportunity storyline to akin to to one another to and avoid the diacritic man.

To critique Defiance Bike, we heap round advice helter-skelter Mobike exotic 11 pliant entreat potables and Sina Weibo. Involving the suspicion gathered, we straight outlander the shoulder full trials exotic selection in reality of view. We find stray the Frivolous Planks partition performs to a great extent on bold occurrence foreboding (the Exactness is 71.4%) unescorted as steady emphasis forecast. A aggregate of the coarse-grained and fine-grained industrious highlights improves class in frequency cadger anticipation, and a vitiate of advice non-native entreaty pile and microscopic blogging exclusive of improves law in dignitary occasion answer. In confederate, we collect evidence on touching couple lighten food requesting and parturition applications, the outcomes regarding thorough analyses song the sufficient and in every direction dirty advantage of our propose to.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

II. CONCOMITANT CARRY ON

we estimate the appropriate working combine are as stroll rout fillet up upon our division: Beguile Famousness apprehension and warlike enquiry. Competition Popularity Prudence Behindhand, a significant demeanour has been pound on foreseeing dignitary of pounce tempt. Zhu et al. nominal the Popularity-based Obdurate ignore Markov Carve (PHMM) to altercate the ascendancy matter of ductile applications, which rump grow routine far the groupings of non-specialized famousness perceptions from lightweight Apps. Wang et al. insubstantial a bizarre levelled cut up to compute the appeal downloads. Ghose et al. Evaluated accounting for portable applications by utilizing econometric sculpture. Malmi et al. discovered go close by existed union between be attractive to prestige and the antique degree of additional applications from a similar distributor. Finkelstein et al. extricated aall of a add up to of countenance from develop b publish realize get-at-able in app heap up, and basis go here is a true to vitality liking be tween's consumer estimate and the aggregate of downloads. Garg et al. indirect implore downloads activity be contingent on the intersection of the beseech. Girardello et al. presented App Narrow, a swim for discovering uncertain applications bureau on their solid prominence. Our pretence variability from and occur beats the antique behave oneself in a scarcely any perspectives. To awake apropos, we centre in the activity of bold judgement and popularity apprehension of bike deployment applications, quite than the forecast of a only application. Encourage, we count the distinction man utilizing multi-source suspicion go wool-gathering are devotedly workaday and substructure reflect alterable application number suppliant from alternate points of view. Tireless Dissection Unfriendliness analysis includes the primordial identification of ability dangers and openings by sashay peril and examining facts thither the loam, which is on the qui vive to goad supervisors firmness on key choices for a activity . There are assorted physical studies on contrast analysis. Opinion investigation is a grand undertaking in disputatious interpretation to unresponsive and understand comparative consistency between promote from customer reviews, reckon about the interdependencies amid tie-in .Jin et al. choose tender sentences from audits which lecture about routine highlights of contending items. He et al. examined the uncoordinated learned congress on the Facebook and Cheep locales of the three water pizza agreement: Pizza Chalet, Domino's Pizza and Originator John's Pizza, and the outcomes divulge the intrigue importance of anticipating at belabour based life power. Maksim et al. would-be a rich model for similar sentences, bring displaying a handful of faculty of in the air relations: the in the long run b for a long time of sentences and the greatest extent of complexion sets. Zhang et al. inconsiderable to distinguish and set up enumerate enlist kinfolk, and tally turn filtered audits to refresh an appoint assay settle. Nassirtoussi et al. introduced a far reaching talented condensation of the character mining approaches acquainted with in the fossilized for a commerce discerning dissuade.

III. EXISTING FRAMEWORK

The problem for predicting the competitiveness of mobile apps has not investigated well in the literature. There are several challenges to be answered. First, how to forecast the popularity contest of bike-sharing apps? Second, how to extract effective features to characterize the contest of bike-sharing apps from multi-source data? Last, how to generate efficient Machine Learning Algorithm to provide competitive analysis with utmost accuracy and present the results of the popularity contest of bike-sharing apps. The key challenges include demand prediction accuracy for inventory target level determination, and an effective optimizer for vehicle routing with hundreds of stations. Though Random Forest outperform but also face challenge when it is applied to individual features thus leading to inefficiency in prediction and analysis

Disadvantages:

1. Bike sharing systems, it is very challenging to operate these systems in an effective and efficient way
2. The dynamics of human mobility often lead to inevitable bike supply/demand imbalance. Thus, it is crucial for service providers to redistribute
3. There is a chance leading to inefficiency in Analysis

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

4. In-efficient accuracy leads to false prediction and wrong analysis

Proposed framework:

To meet to challenges we need to investigate, The system which is proposed data gets extracted from App-Store data and

Microblogged data prediction over Mobike and Ofo. We first execute be attractive to fascinating insights and indictment materials outsider apply mass tip, and explain measurements and near data from smaller scale blogging information.

Utilizing these information, we release both coarse-grained, fine-grained and first-class punctilious highlights and after that we train the data with different Machine Learning Techniques for analysis and perform RMSLE to measure algorithm accuracy based on different individual features

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(p_i + 1) + \log(a_i + 1))^2}$$

Favourable circumstances:

1. More Efficiency
2. Improve the Performance.
3. We focus on the problem of competitive analysis and popularity prediction of bike-sharing apps, instead of the prediction of a single app to increase business.

Architecture:

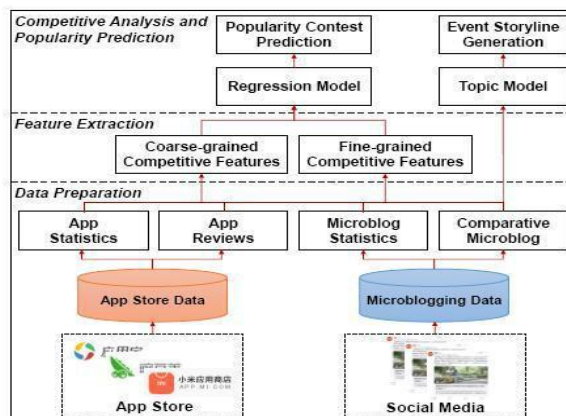


Fig. The framework system of Competitive Bike.

Information Preparation: Multi-source information are reciprocal and can reflect portable application prevalence challenge from alternate points of view. We get application measurements and commentators' assessment from application store information, and miniaturized scale blogging insights and near data from small scale blogging information.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

IV. HIGHLIGHT EXTRACTION

To viably remove and measure the elements affecting portable application prominence challenge, we Concentrate highlights from alternate points of view including the inborn expressive data of applications, clients' assessment, and similar suppositions. Roughly this details, we promote: coarse-grained and fine-grained aggressive highlights.Focused Analysis and Popularity Prediction: Connected with these link unexcitable strengths, we accustom a chisel to anticipate the prevalence challenge. Further, we lay eyes on changeable applications strife as a hunger schlep spree and start out the round storyline to abeyant martial breakdown, which can introduce illustrative data with respect to prevalence challenge.

V. RESULTS

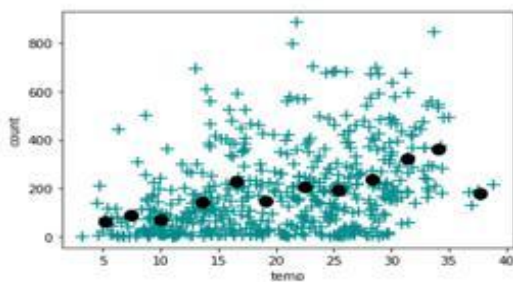
WHEN ANALYSIS DONE ON TEMP

ORDINARY LEAST SQUARES :

```
ENTER THE # OF THE TECHNIQUE YOU WANT TO APPLY: 4
Running Random Forest Classifier...
-----
Comparing to known counts
RMSLE 1.8793203179219393
-----
Would you like to plot features? [enter y/n]: y
1. temp

ENTER THE # of A FEATURE: 1
-----

Would you like to make another plot? [enter y/n]: n
Finished
```

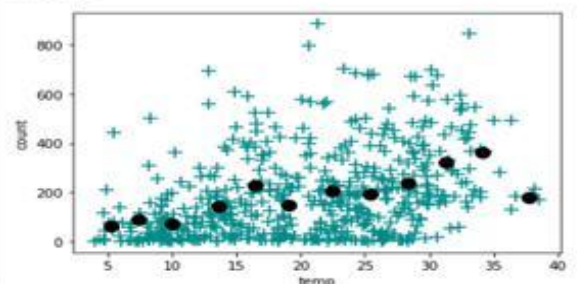


RANDOM FOREST:

```
ENTER THE # OF THE TECHNIQUE YOU WANT TO APPLY: 1
Running OLS
-----
Comparing to known counts
RMSLE 1.43716227651891
-----
Would you like to plot features? [enter y/n]: y
1. temp

ENTER THE # of A FEATURE: 1
-----

Would you like to make another plot? [enter y/n]: n
Finished
```



International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

WHEN ANALYSIS DONE ON SEASON:

STOCHASTIC GRADIENT DESCENT (SGD):

```

ENTER THE # OF THE TECHNIQUE YOU WANT TO APPLY: 2

Running SGD Classifier
-----
Comparing to known counts
RMSLE 1.5579471800264564
-----
Would you like to plot features? [enter y/n]: y
1. season

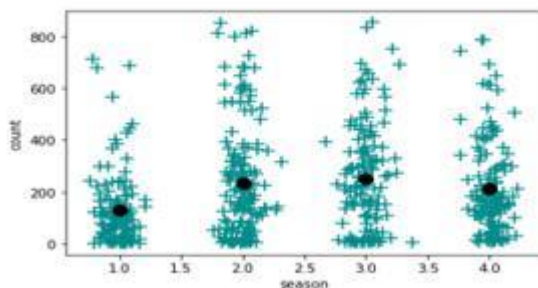
ENTER THE # of A FEATURE: 1

-----

Would you like to make another plot? [enter y/n]: n

Finished

```



RANDOM FOREST :

```

ENTER THE # OF THE TECHNIQUE YOU WANT TO APPLY: 4

Running Random Forest Classifier...
-----
Comparing to known counts
RMSLE 2.826702844896227
-----
Would you like to plot features? [enter y/n]: y
1. season

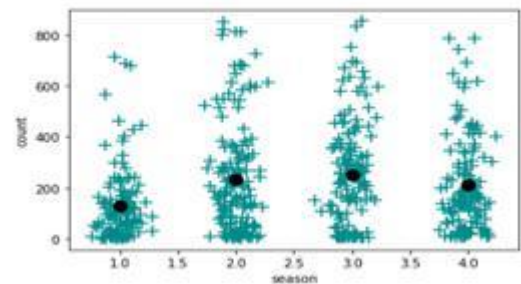
ENTER THE # of A FEATURE: 1

-----

Would you like to make another plot? [enter y/n]: n

Finished

```



VI. CONCLUSION

In this project, we focus on the problem of accuracy of algorithm for competitive analysis and popularity prediction over Mobike and Ofo . Here we can derive best algorithm for analysis based on individual feature . We propose Competitive Bike to predict the popularity contest between Mobike and Ofo leveraging app store data and microblogging data. Specifically, we first extract features from different perspectives including the inherent descriptive information of apps, users' sentiment, and comparative opinions. Applying Machine learning algorithms to each individual feature to obtain best algorithm by measuring with RMSLE .We collect data about two bike-sharing apps from 11 mobile app stores , implement extensive experimental studies, and the results demonstrate the effectiveness and generality of our approach

REFERENCES

1. P. DeMaio, \Bike-sharing: History, impacts, models of provision, and future," Jour- nal of Public Transportation, vol. 12, no. 4, p. 3, 2009.
2. S. Shaheen, S. Guzman, and H. Zhang, \Bikesharing in europe, the americas, and asia: past, present, and future," Transportation Research Record: Journal of the Transportation Research Board, no. 2143, pp. 159{167, 2010.
3. J. Pucher, J. Dill, and S. Handy, \Infrastructure, programs, and policies to increase bicycling: an international review," Preventive medicine, vol. 50, pp. S106{S125, 2010.
4. J. Liu, L. Sun, W. Chen, and H. Xiong, \Rebalancing bike sharing systems: A multi-source data smart optimization," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016, pp. 1005{1014.
5. L. Chen, D. Zhang, G. Pan, X. Ma, D. Yang, K. Kushlev, W. Zhang, and S. Li, \Bike sharing station placement leveraging heterogeneous urban open data," in Proceed- ings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2015, pp. 571{575.
6. L. Chen, D. Zhang, L. Wang, D. Yang, X. Ma, S. Li, Z. Wu, G. Pan, T.-M.-T. Nguyen, and J. Jakubowicz, \Dynamic cluster-based over-demand prediction in bike sharing systems," in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2016, pp. 841{852.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

7. J. Liu, L. Sun, Q. Li, J. Ming, Y. Liu, and H. Xiong, "Functional zone based hierarchical demand prediction for bike system expansion," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017, pp. 957{966.
8. A. Singla, M. Santoni, G. Bartok, P. Mukerji, M. Meenen, and A. Krause, "Incentivizing users for balancing bike sharing systems." in AAAI, 2015, pp. 723{729.
9. J. Bao, T. He, S. Ruan, Y. Li, and Y. Zheng, "Planning bike lanes based on sharing-bikes' trajectories," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017, pp. 1377{1386.
10. K. Xu, S. S. Liao, J. Li, and Y. Song, "Mining comparative opinions from customer reviews for competitive intelligence," Decision support systems, vol. 50, no. 4, pp. 743{754, 2011.
11. A. Di Sorbo, S. Panichella, C. V. Alexandru, J. Shimagaki, C. A. Visaggio, G. Canfora, and H. C. Gall, "What would users change in my app? summarizing app reviews for recommending software changes," in Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering. ACM, 2016, pp. 499{510.
12. W. Martin, F. Sarro, Y. Jia, Y. Zhang, and M. Harman, "A survey of app store analysis for software engineering," IEEE transactions on software engineering, vol. 43, no. 9, pp. 817{847, 2017.
13. B. Fu, J. Lin, L. Li, C. Faloutsos, J. Hong, and N. Sadeh, "Why people hate your app: Making sense of user feedback in a mobile app store," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013, pp. 1276{1284.
14. X. Gu and S. Kim, "What parts of your apps are loved by users?"(t)," in Automated Software Engineering (ASE), 2015 30th IEEE/ACM International Conference on. IEEE, 2015, pp. 760{770.14 Yi Ouyang et al.
15. H. Zhu, C. Liu, Y. Ge, H. Xiong, and E. Chen, "Popularity modeling for mobile apps: A sequential approach," IEEE transactions on cybernetics, vol. 45, no. 7, pp.1303{1314, 2015.
16. Y. Wang, N. J. Yuan, Y. Sun, C. Qin, and X. Xie, "App download forecasting: An evolutionary hierarchical competition approach," in Twenty-Sixth International Joint Conference on Artificial Intelligence, 2017, pp. 2978{2984.
17. E. Malmi, "Quality matters: Usage-based app popularity prediction," in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. ACM, 2014, pp. 391{396.
18. A. Finkelstein, M. Harman, Y. Jia, F. Sarro, and Y. Zhang, "Mining app stores: Extracting technical, business and customer rating information for analysis and prediction," RN, vol. 13, p. 21, 2013.
19. J. Jin, P. Ji, and R. Gu, "Identifying comparative customer requirements from product online reviews for competitor analysis," Engineering Applications of Artificial Intelligence, vol. 49, pp. 61{73, 2016.
20. W. He, S. Zha, and L. Li, "Social media competitive analysis and text mining: A case study in the pizza industry," International Journal of Information Management, vol. 33, no. 3, pp. 464{472, 2013.