

LOCATION BASED RECOMMENDATION SYSTEM FOR BUSINESS EXPANSION

A Survey Paper

¹Siddhant Mungad,²Sudhanshu Ranjan, ³Sudhin Pillai, ⁴Yash Yadav, ⁵Renuka Suryawanshi

¹Student, ²Student, ³Student, ⁴Student, ⁵Assistant Professor

¹School of Computer Engineering and Technology,

¹MIT World Peace University, Pune, India.

Abstract: When it comes to launching a new venture or developing an existing one, location is one of the most critical aspects to remember. The right location is important not only for attracting more consumers and increasing profits, but it can also decide whether a company can prosper based on factors such as the existence of competition. We plan to summarize current systems in this survey paper and, based on the research gaps, suggest a new recommendation framework that separates users into: A] Starting New Venture for an Established Business: Using the K-means clustering algorithm to propose locations, analyze the data of market locations with the type of sector where there is a high industry level and average customer ranking. B] Starting a New Business: Using Correlation Analysis between various business sectors to suggest locations using a ranking algorithm, find the best place to start a business where there is strong industry demand and little or very little supply.

Index Terms – Recommendation, business, location, correlation, weighted k-means, clustering, competition, ranking.

I. INTRODUCTION

The main focus of developing countries is aiming for development through infrastructure and increased economy by providing sustainable jobs and services throughout the country for better standards of living. For these developments to happen urbanization plays a vital role for developing countries because as the population density in urban areas continues to grow, the needs of the inhabitants also grows and should be met. These needs are fulfilled by businesses offering various services. But, the proper sustainability of business has become a major concern due to increased competition and overheads over a particular region. For a particular business to thrive in a region various factors come into play such as geographical, social and cultural aspects of the community [14]. Out of these geographical factors, Location is one of the most important aspects in sustainability because most of the attributes are directly or indirectly dependent on location.

Business location is one of the most important factors to be considered whether it is setting up a new business or optimizing an existing one. Selecting the right location is not only crucial for attraction of more customers and profitability, but can also determine whether the business can successfully thrive or not. We can see various business shops come and go within some time frame before diving into profit due to poor placement of shop location. Whenever a customer wants a service, he might look for locations to visit that offer his service of interest, for example, if a customer wishes to have food, they may look for locations with restaurants. This demand and supply of customers' needs for availing the services should be met for self-sufficiency whether it be a new business or existing business. Taking into consideration the above demands and finding the best suitable location for business, Location based recommender systems have been widely exploited by researchers. These systems aim to filter out information by returning venues to the user while considering spatial and personal preference of visited locations.

According to Location based recommender systems the location context problem can be majorly divided into three types i.e. Location analysis which aims on selecting the best location for any goal; Spatial analysis which extends to including spatial components and Point of interest recommender which makes use of check-in data to recommend location [6]. To tackle these problems various analytical approaches and algorithms have been used namely predictive, non-predictive, inference, optimization, but there is still room for improvement because there are few constraints that are left out. Also, presence of different industry types can impact the growth or loss of a single industry type which is not looked upon. In this paper we aim to provide a survey on Location based recommender systems and propose a new way to look upon recommendation by analyzing shop information by taking competition into picture and interlinking various industry types to provide location recommendation.

II. RELATED WORK

Location based recommender systems have a large dependency on check-in data made by the user on Location-based social networks [2]. Check-in data refers to records or history of the locations that a particular user has visited. The number of check-ins at a particular business venue reflects its ability to draw in people to avail services [16]. It also shows how popular a business is in a region giving a rough estimate of sustainability and profitability. Researchers analyze these check-ins from Facebook, Foursquare, Yelp, Kaggle etc. [4] to predict where the user will check-in next to provide recommendations of business location. However this information of check-in comes with a cold-start problem. It simply means that,

Recommendation to new users can't be given because there's no information available of its existing check-ins. Also, Recommendation for places of a new place or region that a user wants to visit can't be given due to the same issue. Researchers are constantly addressing this problem, and new methodologies have been established to address it. Hongzhin and Ben Chui in [9] proposed that the new place or region problem can be given by using the activity history of other people or local inhabitants of that new city which can act as check-in for new users. It also solved the problem of data sparsity by exploiting the contextual in-formation and linking spatial items together by correlating users. But, the query processing tends to be slow if the ranking score of the recommendation list does not cross a threshold score. Also, Kerson and Salaiwarakal [7] used a kindred approach for cold start by using Friend's check in data to analyze ongoing activities and updates. However, these approaches have the drawback of limiting check-in to only certain media outlets that serve a large pool of correlated users, such as Facebook, twitter, etc and they appear to ignore the possibility that consumer habits change over time, causing the check in data to indirectly affect the recommendation. Another method to tackle cold start problem was done in [10], using predictive feature-based regression models that make use of all available information about users and items, such as user demographics and item content features. However this methodology limits itself to large scale data sets.

The most improved solution to the cold-start problem was given in [13], where Behavior-based Location Recommendation model was used, in which prediction is given as a result of past behavior's while also keeping the behavior of similar users into account. Furthermore, suitable locations are predicted based on user-specific and behavior-specific spatial models. This model recommends locations to the users based on the every-day activities which are effective for most cases. But, when users visit a new venue, though, using day-to-day behavior model to make suggestions will not be as accurate.

While using check in data for recommending places is a common approach, another way to work out the recommendation can be by exploiting geographic data. The geographic data contains physical features of the pre-existing business such as latitude, longitude, number of check-ins etc. Using geographical data avoids one of the disadvantages of relying exclusively on check-in data, which is that it lacks a location's spatial characteristics, which is a vital attribute to consider. Based on the fact, geographical information has been intensively used in location recommendations. One such paper [3] simply considers the physical location of major US cities to recommend business categories to its districts. The algorithm firstly divides the complete regions into business districts by using clustering algorithms and then decides the best business category for the cluster by measuring the success rates of existing businesses. It also stated that K-means clustering is best suitable for these kinds of scenarios. While the earlier paper [3] uses clustering algorithms as its base for recommendation, [8] is to make use of classification algorithms. It divides the region into zones and ranks these zones as per business profile given to the system through classification. Classification algorithms like SVM and random forest classifier are used to compute the matching score between a business profile and a zone ID. It also states that random forest outperforms all other classification algorithms. Similarly in light of geographical aspect, another paper [12] proposes use of three main features that are the user characteristics, geographical distance and location popularity to give the final recommendations and this is achieved by a model called the LC-G-P scheme. It is a scheme which uses Louvain algorithm for effectively identifying the communities and its result is then used by the content based collaborative filtering (CCF) method. The above scheme for community detection is comprised of two main stages, constructing the user similarity networks based on the check ins data and detecting the communities based on the geographical distance and the location popularity. Although the proposed system outperformed many existing systems in terms of accuracy, there is a limitation on predicting the user's exact location when calculating geographical distance because the method used to achieve this has a low accuracy. Considering the kindred aspect [17] addresses the vacant shops that are present in the city and an algorithm is proposed which recommends suitable places to open a new business based on two factors; the flow of people and social. It takes into account the topological properties, traffic of a given place, people in the street to recommend a location. But the authors had difficulty to determine how many specific shops are needed in each neighborhood.

The Literature above works only on design recommendation systems based upon properties like, i.e., user characteristics, user check ins [2,4,7,9,10,16], spatial information [13], location popularity [2], geographical factors [3,8,12,17], etc. In fact, there exist many other factors such as, the users search data [3] is only considered while recommending places and presence of competition [2] is not considered while recommending locations. Correlation [1] between different business industries can be a factor for business which is not looked upon. Availability and rent of properties at business locations, tax, legal issues [1] are not considered while recommending locations. Taking into consideration all above factors, a new approach for location based recommendation system is proposed.

III. PROPOSED SYSTEM

We propose a system which is divided into three parts namely dataset, pre-processing and the business success recommendation model. The dataset part is involved with managing dataset which includes making calls to google places API with input as location canter and output as nearby shops in json format. The contents of the returned json are then fed to the pre-processing part which involves cleaning the json and storing it in a csv file. This cleaning involves handling missing values, selecting specific variables and data transformation mainly drill down operation on the address field of json. This is then stored in a csv file. This data set after preprocessing is then fed to the business success recommendation model.

The Business Success Recommendation model based on user's business requests are divided into two categories, either established businesses or new businesses. The reason behind splitting the business into two parts and applying two different algorithms for the same is because each of the above two categories plan their business model differently to earn profit.

Established businesses who want to start a new venture want to choose the location where there is very high demand for their product. They are unconcerned about the rivalry in the area because they are proven, which means they have already reached a certain level of customer expectation and have a pre-existing audience. As a result, the top priority while choosing a location should be areas where their commodity is in high demand. They still have customers, and new customers can come as a result of strong word of mouth, allowing them to extend their company and gain better long-term profitability. The same idea can't be applied to new business, because they don't have an existing audience. Directly, placing their shops in regions having high commodity demand and many business providers would result in them getting crushed by the competition. Even if the product is highly capable it will not flourish under those circumstances. A better approach towards placing these businesses is by finding the correlation between different business industries to recommend location. For eg. : We can see that whenever a customer visits the movie theatre locations , there is a high probability of them visiting areas like restaurants, clothing shops, etc. Or An office premise, educational institutes being correlated to industries like stationary shops, canteens, day cares, recreational center's, etc. The audience of correlated businesses can become customers for new businesses, assisting them in their initial growth and setting their target audience. The analysis of the business success recommendation model is based on the type of business i.e For pre-established business weighted K-means clustering is used and for new business Correlation analysis is done which after analysis gives the location coordinates of the new shop to establish business.

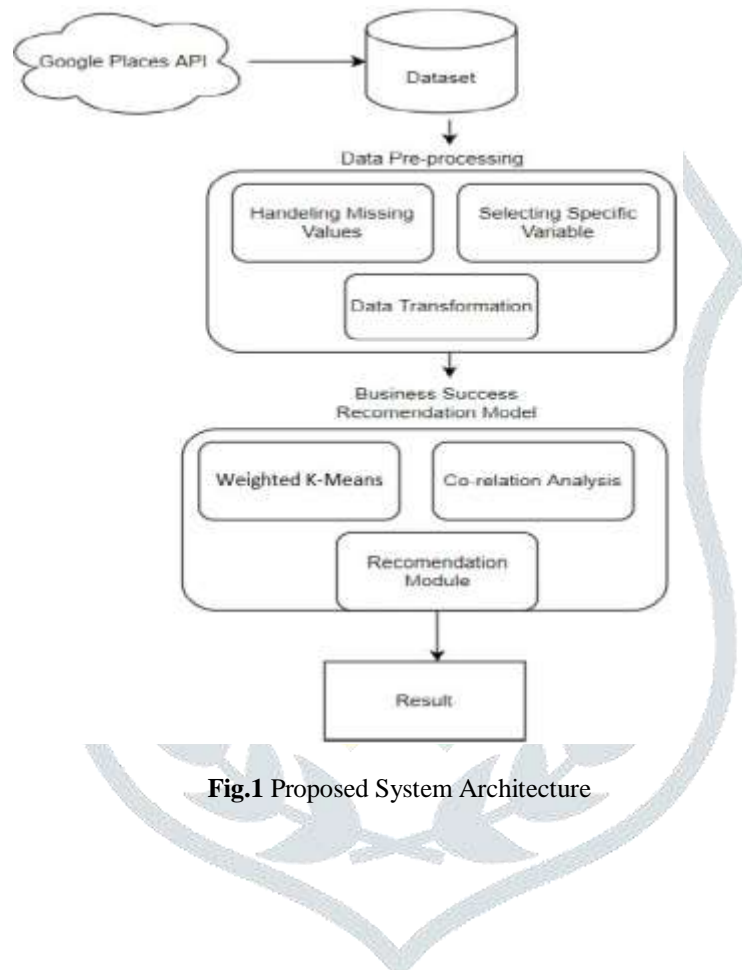


Fig.1 Proposed System Architecture

3.1 Algorithm for Established Business

The algorithm for established business is based on clustering algorithms. As seen in [3] K-means clustering outperforms other clustering techniques. But, each shop should have different weights while defining the cluster because different shops give out different demands and popularity. Having each data point (shops) the same weight means that clusters having more unpopular shops will give deceptivedemand than clusters having less shops but with high popularity. Hence, having weights to the data points is a must. So, a weighted K-means clustering algorithm will be used as the clustering on a filtered data frame with the business type of user requirement given as input for the system. The number of clusters will be defined by the elbow method. After cluster formation each cluster will be ranked based on the frequency count so to meet demand and the center of the most highly ranked cluster will be given out as the recommendation.

3.2 Algorithm for New Business

The algorithm for new business would be focused on determining the business's correlation. The assumption for correlation being, existing shops are already placed at location with their correlated business. To mine the correlation we will find the most popular shop in the region, then find the all neighboring shops [1] and their corresponding industry within a 500m radius to find the correlation between various industries. This must be done with five profitable shops in order to obtain knowledge about the business, which is related with one industry. The region with the highest frequency would

then see a high correlation. Following that, we can use Algorithm 1 to determine the high demand of associated industries and then propose locations.

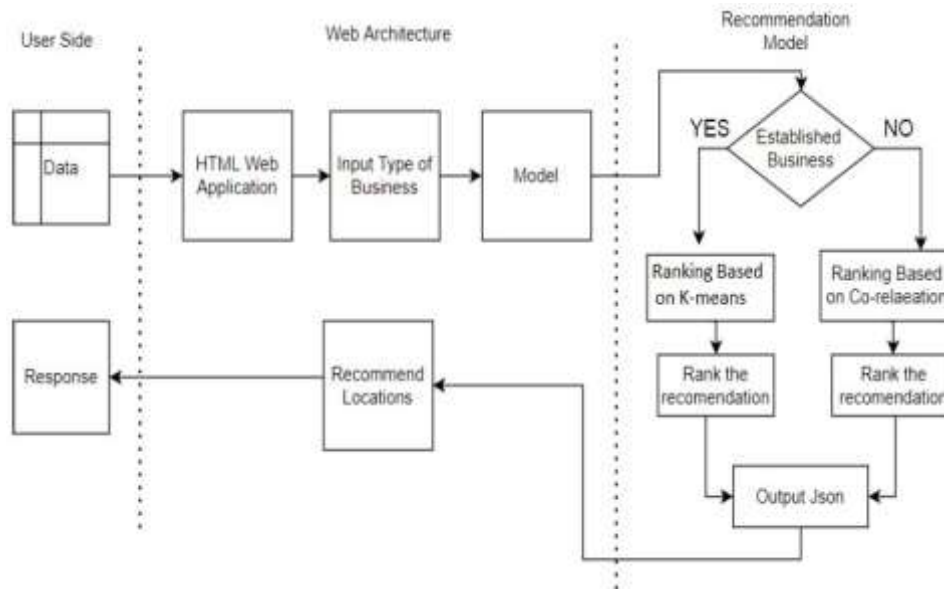


Fig.2 Proposed Flow Diagram

IV. CONCLUSION

Starting a company helps in many ways in developed countries where urbanization is significant, such as more jobs, a healthier infrastructure, increased GDP, more opportunities, and so on. When a customer needs a service, he will search for locations to visit that have those particular services. As a result, whether starting a new business or expanding an existing one, location becomes critical. In this paper we looked at different literature works to see what kinds of position background problems there are based on location context and how to solve them. We came up with a new approach after considering the shortcomings and recognizing the methods used in the literature works where considering users business requests, the new approach aims to recommend a suitable location where newly or pre-established businesses can start and bloom based upon factors such as industry competition, frequencies and their average user ratings. Though, we don't pretend to have given a complete solution for the problem taking into consideration all the factors such as availability and rent of properties at business locations, tax, legal issues but, we have indeed taken a step ahead in this field with a new approach which reduces risk and increases accuracy by recommendation location which could help in high expansion of business.

REFERENCES

- [1] Kumar, A., Shankar, S., Kumar, P.K., Reddy, M., Reddy, T.G., Al-Ahmari, A., Abidi, M.H.: Location Based Business Recommendation Using Spatial Demand. In: Sustainability, vol. 12 (10), 4124. MDPI Open Access Journal (2020).
- [2] Janani, V., Balasubramanian, V., Sasikala, G., Vidhya, G., Kowsalya, T.: A Location Recommendation Based on user reviews using cart. In: IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), 1-7 (2019). <https://doi.org/10.1109/ICSCAN.2019.8878812>
- [3] Ferreira, V., Vallejo, A., Valdivia, P., Rebaza, J.V.: Exploiting Geographical Data to improve Recommender Systems for Business Opportunities in Urban Areas. In: The 8th Brazilian Conference on Intelligent Systems (BRACIS), 568-573 (2019). <https://doi.org/10.1109/BRACIS.2019.00105>
- [4] Eravci, B., Bulut, N., Etemoglu, C., Ferhatosmanoglu, H.: Location Recommendations for New Businesses using Check-in Data. In: IEEE 16th International Conference on Data Mining Workshops (ICDMW), 1110-1117 (2016). <https://doi.org/10.1109/ICDMW.2016.0160>
- [5] Larasatie, P., Setiowati, S.: From Fingerprint to Footprint: Using Point of Interest (POI) Recommendation System in Marketing Applications. In: The Asian Journal of Technology Management (AJTM), vol 12 (2), 118-131 (2019). <https://doi.org/10.12695/ajtm.2019.12.2.4>
- [6] Ferro-Díez, L.E., Villegas, N.M: Location Data Analytics in the Business Value Chain: A Systematic Literature Review. In: IEEE Access, vol. 8, 204639-204659 (2020). <https://doi.org/10.1109/ACCESS.2020.3036835>
- [7] Kesorn, K., Juraphanthong, W., Salaiwarakul, A.: Personalized Attraction Recommendation System for Tourists Through Check-In Data. In: IEEE Access, vol. 5, 26703 - 26721 (2017). <https://doi.org/10.1109/ACCESS.2017.2778293>
- [8] Lin(B), J., Oentaryo, R., Lim, E., Vu, C., Vu, A., Kwee, A., and Prasetyo, P.: A Business Zone Recommender

- System Based on Facebook and Urban Planning Data. In: European Conference on Information Retrieval, 641-647 (2016).
- [9] Yin, H., Sun, Y., Hu, Z., Chen, Z. : “LCARS: A Spatial Item Recommender System”. In: ACM Transactions on Information Systems 18, vol.32 (3),1-37 (2014). <https://doi.org/10.1145/2629461>
- [10] Park, S., Chu, W., Pairwise Preference Regression for Cold-start Recommendation. In :RecSys '09: Proceedings of the third ACM conference on Recommender systems,21-28 (2009). <https://doi.org/10.1145/1639714.1639720>
- [11] Su, X., and Khoshgoftaar, T.: A survey of collaborative filtering techniques. In: Advances in Artificial Intelligence, vol. 2009, 421425 (2009). <https://doi.org/10.1155/2009/421425>
- [12] Cai, W., Wang, Y., Lv, R., Jin, Q.: An efficient location recommendation scheme based on clustering and data fusion. In: Computers and Electrical Engineering vol. 77, pp. 289–299 (2019).<https://doi.org/10.1016/j.compeleceng.2019.06.006>
- [13] Rahimi, S. M., Wang, X., Far, B.: Behavior-based location recommendation on location-based social networks. In :Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 273–285 (2017).<https://doi.org/10.1007/s10707-019-00360-3>
- [14] Zhang, J. D., Chow, C. Y.,: GeoSoCa: exploiting geographical, social and categorical correlations for point-of-interest recommendations. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 443–452, Santiago, Chile (2015). <https://doi.org/10.1145/2766462.2767711>
- [15] Portugal, I., Alencar, P., Cowan, D.: The use of machine learning algorithms in recommender systems: A systematic review. In: Expert Systems with Applications, vol. 97, 205–227 (2018).<https://doi.org/10.1016/j.eswa.2017.12.020>
- [16] Zhao, S., King, I., Lyu, M. R., Zeng, J., Yuan, M.: Mining business opportunities from location-based social networks. In: Proc. of the 40th Int. ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR '17. ACM, 2017, 1037–1040 (2017). <https://doi.org/10.1145/3077136.3080712>
- [17] Grimaldi, D., Fernandez, V., Carrasco, C.: Heuristic for the localization of new shops based on business and social criteria. In: Technological Forecasting and Social Change, vol. 142, 249–257 (2019). <https://doi.org/10.1016/j.techfore.2018.07.034>

