IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

IMPROVING TAXI REVENUE

[1]Vishwas K V, [2]B P Sowmya,

[1]Research Scholar, Dept. of MCA, PES College of Engineering, Mandya, Karnataka.
[2]Assistant Professor, Dept. of MCA, PES College of Engineering, Mandya, Karnataka.

Abstract: **Recent years have witnessed the rapidly-growing business of ride-on-demand (RoD) services such as Uber, Lyft, and Didi. Unlike taxi services, these emerging transportation services use dynamic pricing to manipulate the supply and demand and to improve service responsiveness and quality. Despite this, on the drivers' side, dynamic pricing creates a new problem: how to seek passengers in order to earn more under the new pricing scheme. Seeking strategies have been studied extensively in traditional taxi service, but in RoD service such studies are still rare and require the consideration of more factors such as dynamic prices, the status of other transportation services, etc. We develop ROD-Revenue, aiming to mine the relationship between driver revenue and factors relevant to seeking strategies and to predict driver revenue given features extracted from multi-source urban data. We extract basic features from multiple datasets, including RoD service, taxi service, POI information, and the availability of public transportation services, and then construct composite features from basic features in a product form. The desired relationship is learned from a linear regression model with basic features and high-dimensional composite features. The linear model is chosen for its interpretability to quantitatively explain the desired relationship. Finally, we evaluate our model by predicting drivers' revenue.**

Index Terms **- Taxi Revenue Prediction, Forecasting, SVM, KNN, CNN Algorithm.**

I. INTRODUCTION

Provides us an economical and healthy lifestyle. among the various forms of shared transportation, public taxi service systems. as taxi service system become increasingly popular, a lot of companies join the market, leading to fierce competition.
several advancements has been made recently in the field of transportation system and taxis being a major part of urban transportation has seen a tremendous growth in the in recent years, shared transportation has grown tremendously, which past few years through the use of online cab services such as uber, ola etc. a lot of studies were made in the customer's perspective to give improved
customer experience. in this project, the focus is on improving performance and prediction the revenue generated from a driver. our approach to this problem was first to analyze and preprocess a collection of taxi records. we used a to understand which features best predict demand, and a reinforcement learning model to maximize an individual drivers income based on a set of summary statistics.

II. LITERATURE REVIEW

C Yanshen [1] The stable problem of taxi service industry have already been a focus around all the country.There are many complex elements in smooth development of the industry to be unstable, developing challenges have to be faced. Analysis on taxi service characters about patrolling operation, service for some of residents, ruled policy of government and belonging to society and so on, the service level and the quality of service are the eternal subjects in taxi industry development.

W Ya-jun [2] There is a direct relation between the government's control and taxi market's problem/conflict. The government's control results in the potential monopoly and more capital investment or speculation, then the investment and service chain has been prolonged and relevant contradiction has increased.

SA Zahra [3] Success in business ecosystems that include well-established companies and new ventures requires collaboration and competition, a task that demands strategic thinking to leverage a firm's resources and capabilities. Strategic thinking and the entrepreneurial activities in an ecosystem influence one another in a cycle that perpetuates and even sparks innovation.
N Bengtsson [4] Informal sectors in developing countries are often thought of as responses to rigid and cumbersome market regulations. In this paper I study informal trade as a first-best outcome. In the model I propose rigid regulations can be necessary to achieve efficiency even though they are always sidestepped. The key assumption is that the regulations define the trading parties' fall-back position in case the informal bargaining process breaks down. I set up a field experiment to test the model's mechanisms in the Cape Town market for metered taxis.
C Zhang [5] The goal of this paper is to evaluate the performance of public transit systems based on a combined evaluation method (CEM) consisting of information entropy theory and super efficiency data envelopment analysis (SE-DEA). Taking 13

transit operators in Yangtze Delta Region of China as the research object, we integrate the public transit industry regulations, transit operation and passenger requirements to construct an evaluation indicator system based on satisfaction and efficiency.

DZ Zhang [6] The main scheduling strategy method of existing taxi operation system focuses on the location and possible distribution of idling taxies, which leads to problems as low operation efficiency, high idling rate and long passenger waiting time arise. The taxi operation system encounters a great challenge without a comprehensive evaluation system.

III. PROPOSED METHODOLOGY

Predicting the taxi service demand can help rebalance the taxi services and improve the service quality of a taxi service sharing system. In this project, system takes more attention to those stations with higher taxi service demand. In the proposed system CNN algorithm has been used to assign different types of company into each cluster. The experimental results on the Taxi service system show the advantages of our approach to these above problems.We know that the data set has a complex texture which is a hard task to extract it accurately. Depending on the size, number of columns and value can be recognized and classified. Finally, after comparisons data set will be predicted.

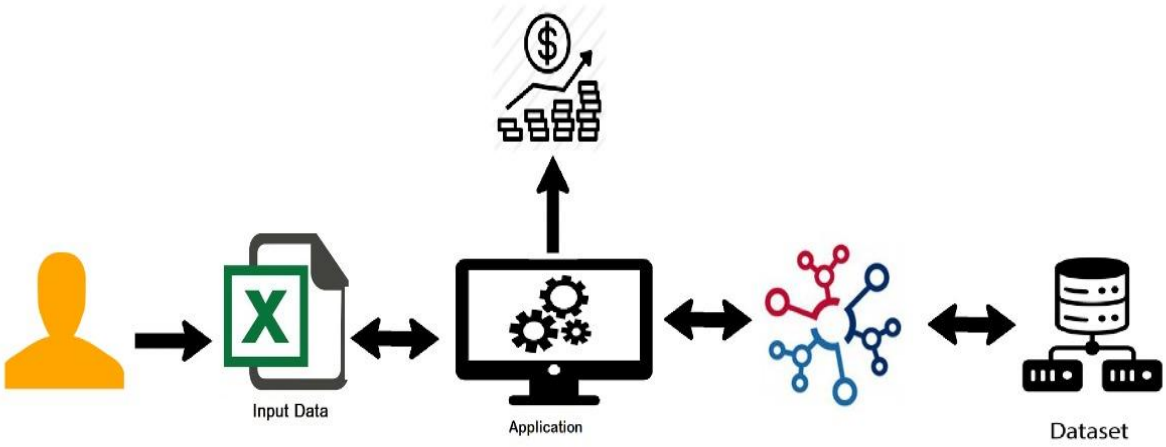


Fig (1): System Architecture

IV. WORKING OF KNN ALGORITHM

K-nearest neighbors (KNN) algorithm uses 'feature similarity' to predict the values of new data points which further means that the new data point will be assigned a value based on how closely it matches the points in the training set. We can understand its working with the help of following steps −41

Step 1 − For implementing any algorithm, we need dataset. So during the first step of KNN, we must load the training data from the exam taken data list and test data from the students currently being taking tests.
Test data set – Upload data based on user input dataset.
Train data set – collected data samples.
Step 2 − Next, we need to choose the value of K i.e. the nearest data points. K can be any integer.
Here k is taken from test data set, which acts like a centroid point.
Step 3 − For each point in the test data do the following −

- **3.1** − Calculate the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. The most commonly used method to calculate distance is Euclidean.
- **3.2** − Now, based on the distance value, sort the user exams preference data in ascending order.
- **3.3** − Next, it will choose the top K rows from the sorted array.
- **3.4** − Now, it will assign a class to the test point based on most frequent segments from the exam data of these rows.

Step 4 – End
In our project we compare k value with train data set using KNN methodology like Euclidean, Manhattan, where we cluster the values based on the nearest distance. Here we get nearest matches of the data preference to test data from train data set in turn we predict the content values from it.

V. WORKING OF CNN ALGORITHM.

When it comes to Machine Learning, Artificial Neural Networks perform really well. Artificial Neural Networks are used in various classification task like image, audio, words. Different types of Neural Networks are used for different purposes, for example for predicting the sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolution Neural Network.
Neural Network. In a regular Neural Network there are three types of layers:

1. **Input Layers:** It's the layer in which we give input to our model. The number of neurons in this layer is equal to total number of features in our data (number of pixels incase of an image).
2. **Hidden Layer:** The input from Input layer is then feed into the hidden layer. There can be many hidden layers depending upon our model and data size. Each hidden layers can have different numbers of neurons which are generally greater than the number of features. The output from each layer is computed by matrix multiplication of output of the previous layer with learnable weights of that layer and then by addition of learnable biases followed by activation function which makes the network nonlinear.
3. **Output Layer:** The output from the hidden layer is then fed into a logistic function like sigmoid or softmax which converts the output of each class into probability score of each class.

Layers used to build Convents

A convents is a sequence of layers, and every layer transforms one volume to another through differentiable function.

Types of layers:
Let's take an example by running a convents on of image of dimension 32 x 32 x 3.

1. **Input Layer:** This layer holds the raw input of image with width 32, height 32 and depth 3.
2. **Convolution Layer:** This layer computes the output volume by computing dot product between all filters and image patch. Suppose we use total 12 filters for this layer we'll get output volume of dimension 32 x 32 x 12.
3. **Activation Function Layer:** This layer will apply element wise activation function to the output of convolution layer. Some common activation functions are RELU: max(0, x), Sigmoid: 1/(1+e^-x), Tanh, Leaky RELU, etc. The volume remains unchanged hence output volume will have dimension 32 x 32 x 12.
4. **Pool Layer:** This layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents from over fitting. Two common types of pooling layers are **max pooling** and **average pooling**. If we use a max pool with 2 x 2 filters and stride 2, the resultant volume will be of dimension 16x16x12.

w can run on multiple CPUs and GPUs (with optional CUDA and SYCL extensions for general-purpose computing on graphics processing units). Tensor flow is available on 64-bit Linux, macOS, Windows, and mobile computing platforms including Android and iOS.

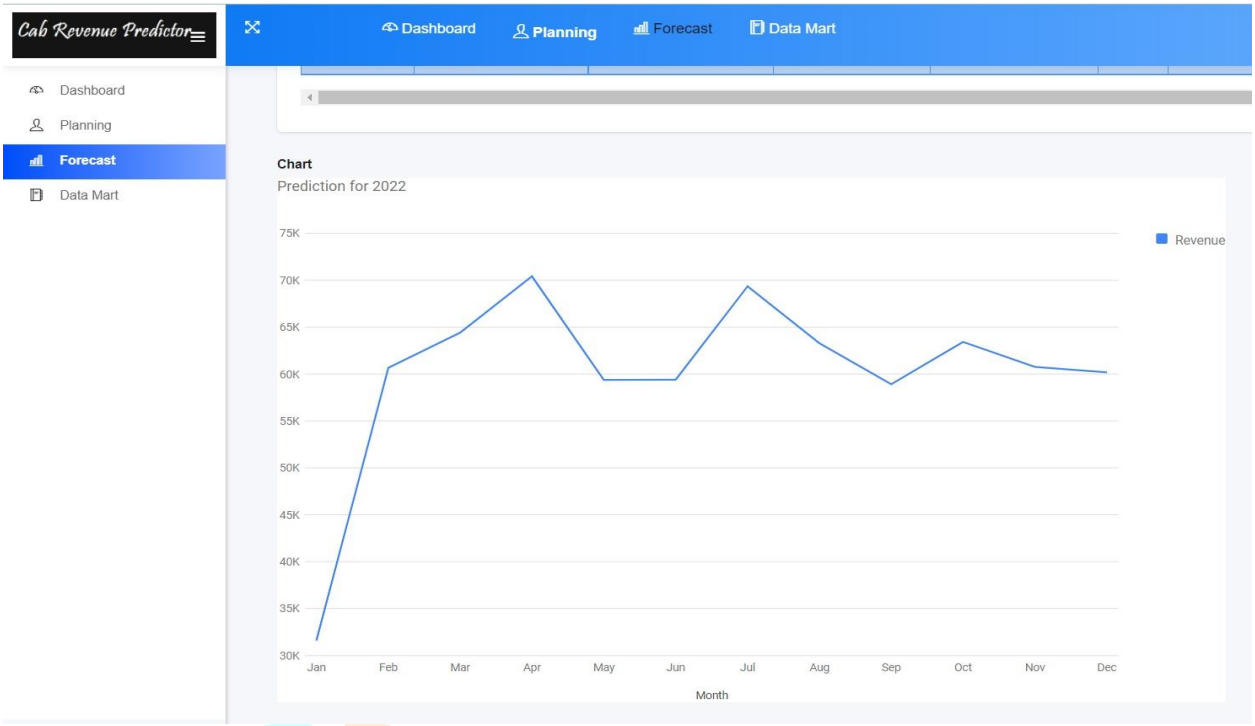
Its flexible architecture allows for the easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices.

VI. SVM

Support Vector Machine (SVM) is a machine learning algorithm approach for solving classification or regression issues. It is, however, mostly used to address classification problems. Each data item is modeled as follows in n-dimensional space (where n is the number of features), with the pixel value being the SVM algorithm's value for a certain position. Then we locate the hyper plane that best differentiates the 2 classes to complete classification.

VII. RESULTS AND DISCUSSION

In this project, system takes more attention to those stations with higher taxi service demand. In the proposed system CNN algorithm has been used to assign different types of company into each cluster. The experimental results on the Taxi service system show the advantages of our approach to these above problems. The application can be made to produce more accurate results. The application can also be deployed in android platform with a few researches about different technologies. On performing all these operation on a daily basis, the model trains itself and makes itself more efficient with every decision it makes. The model learns to predict the better place and route to be on at any given instant. This data can further be used by other taxi drivers or taxi company as reference so that they can implement it in their daily lives and thus a better functioning system is established for the cab drivers.



Fig(2): Prediction Representation.

VIII. CONCLUSION

The project talks about the enhancements that can be made in the field of taxi services so that it is user friendly and make more accurate predictions. In this project, an RL agent, with no knowledge of the environment or taxi demand scenario, is capable of obtaining revenue which is comparable to (and in some cases higher than) revenue earned by top 10 percentile of drivers.

REFERENCES

[1] Chen Y (2009) Analysis on current problems of taxi industry Stablizing and developing. Urban Stud.
[2] Wang Y (2009) The problems and consideration on the taxi industry under governmental regulation. Journal of Chongqing Technology & Business University
[3] Zahra SA, Nambisan S (2012) Entrepreneurship and strategic thinking in business ecosystems. Bus Horiz.
[4] Bengtsson N (2015) Efficient informal trade: theory and experimental evidence from the cape town taxi market. J Dev Econ.
[5] Schaller B (2007) Entry controls in taxi regulation: lessons from North American Experience. In: Transportation Research Board 86th Annual Meeting.
[6] Zhang DZ, Sun J, Peng ZR (2013) Urban taxi goodness index and system implementation of GIS. J Transp Syst Eng Inf Technol.