

Factors Impacting Prediction of DCU Car Parking

Anjana Aggarwal

16210381

School of Computing

MSc. Computing (Data Analytics)

Dublin City University

anjana.aggarwal3@mail.dcu.ie

Date of Submission: 22 August 2017

Disclaimer: An report submitted to Dublin City University, School of Computing for module CA685 Practicum, 2016/2017.

I understand that the University regards breaches of academic integrity and plagiarism as grave and serious.

I have read and understood the DCU Academic Integrity and Plagiarism Policy. I accept the penalties that may be imposed should I engage in practice or practices that breach this policy.

I have identified and included the source of all facts, ideas, opinions, viewpoints of others in the assignment references. Direct quotations, paraphrasing, discussion of ideas from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged and the sources cited are identified in the assignment references.

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work. By signing this form or by submitting this material online I confirm that this assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

By signing this form or by submitting material for assessment online I confirm that I have read and understood DCU Academic Integrity and Plagiarism Policy (available at: <http://www.dcu.ie/registry/examinations/index.shtml>)

Name: Anjana Aggarwal

Date: 22 August,2017

Factors Impacting Prediction of DCU Car Parking

Anjana Aggarwal
School of Computing
MSc. Computing (Data Analytics)
Dublin City University
anjana.aggarwal3@mail.dcu.ie

Abstract—Finding parking in DCU (Dublin City University) campuses in peak hours and on weekdays is a tedious task. According to the study, conducted by real estate department of DCU in past years which conclude that there are fewer parking spaces available compared to numbers of cars people used for traveling to DCU. This paper deals with the development of prediction model (by using appropriate probability distribution) by considering factors which impact car parking spaces. Car parking available spaces get impacted by many factors, for example, holidays, weekday and various time frames. The results have shown that factors like time frames, holidays, weekdays considerably impact available spaces in parking.

Keywords—*Prediction Model, Data Gathering, Polynomial Probability Distribution, Factors Impacting Parking, Correlation Matrix.*

I. INTRODUCTION

Currently DCU campuses (Glasnevin's Patrick's, ALL Hallows) have the total of 1837 parking spaces available which include reserved parking spaces.

The study conducted by real estate shows that the total number of people using cars for coming to DCU campus is more than the number of available parking spaces. So, due to this reason people have to use the residential parking spaces. In addition, according to study^[1], it usually takes 12 minutes to find a free car-parking space. It leads to the requirement of creating prediction model which can predict free parking spaces available at any time of day by considering the factors which impact

There are a number of applications and studies available which help in finding the free car parking spaces. These are usually based on user data or data gathered by sensors or by using prediction algorithms.

Graphs which shows the trend of parking according to different factors have been created using Tableau for data exploration. Python has been used for creating box plots and correlation matrix. Excel and python have

been used for manipulating data. The prediction model which follows polynomial distribution has been implemented using R.

In this study data analytics life cycle has been followed. The main aim of this paper is to build the prediction model and identifying the factors which impact car parking. The results also include how weekdays, time frames in a day^[2], holiday, examinations days have a significant impact on the prediction. Data has been gathered manually and by doing surveys. Three parking spaces Library, Crèche and Invent of Glasnevin Campus has been studied to develop the model. Furthermore, in coming sections will literature review will provide brief about existing solutions and data gathering methods. The paper will conclude with the Hypothesis, implementation methodology, and conclusion.

II. LITERATURE REVIEW

After studying existing literature, currently available solutions for finding free parking space can be divided in below categories: systems based on sensor data, systems based on data gathered by applications and systems based on parking data collected.

A. Existing solutions for prediction/finding of parking spaces

1. Systems based on sensor data

The sensors placed in parking spaces detect whether the parking system is empty or occupied and stores that in the database. This solution is offered by many companies^[3]. The data gathered by these systems are very precise. However, deploying these systems are very costly.

2. Systems based on data gathered by applications

Currently, many applications exist which reports free parking spaces. Data can be gathered by these

types of applications: Free parking spaces reported by the users or implicitly recognized by an application installed on the Smart phone of the users available. Example include: ParkSense^[4], ParkYa^[5].

Systems based on data gathered by applications are preferred due to less operational and deployment cost. However large numbers of users are needed to generate sufficient data and gather information

3) Systems based on parking data collected

This report will be based on parking data collected by observing all DCU parking spaces. The factors which can influence prediction of available parking spaces are considered in this report are:

- Weekdays.
- Time frames on different weekdays and weekends.
- Holidays.
- Exam Study Week
- Examination week.

Above data has been used to develop prediction model by using probability distribution. On the basis of survey results, parking of Glasnevin Campus (Library, Creche, Invent) were selected for this study.

III. OBJECTIVES AND HYPOTHESIS

A. Objective

The main objective of the thesis is to build prediction model and find the factors which impact the car parking.

B. Hypothesis

H0: It is possible to design prediction function (by using probability distribution) by identifying the factors which impact car parking

H1: It is not possible to identify the factors which impact car parking and design prediction model according to those factors.

IV. DATA GATHERING

Data collection is the main pillar for research and analysis in data analytics. The main challenging task in this study was to collect the data. Data was not available for all the parking in DCU. For collecting the data below methods and sources were chosen:

A. Gathering Data Manually

Data has been gathered manually by observing three car parking (Library, Invent, Creche) for the period from 14 Feb,2017-29 May,2017 on different timings and days. Total 50 observations were collected on different days and times for all three parking. Data was then synced to DCU calendar manually and explored through Tableau. After studying data through initial findings, manually data was created for all the three parking. Below observations were made through data exploration on this data using Tableau:

- All parking are full on weekdays from 9 am-3 pm and behavior remains same on other weekdays.
- Time frames have the significant impact on parking.
- Holidays also impact parking spaces.
- Examination Week, Study Week and Reading week shows different trends and appears to impact parking trend.

B. Survey 2017

The survey was conducted in 2017 with the help of *Manager, Sustainability DCU* from Real Estate Department. 1,136 people responded to the DCU travel survey, which was conducted online between 3rd of April and the 28th of April 2017. This represents a response rate of 38% staff and 2% student based on 2,020 staff and 15,845 students. Below observations were made:

- Weekdays are the busiest times.
- Most of the people come to college between 8 am -10 am.

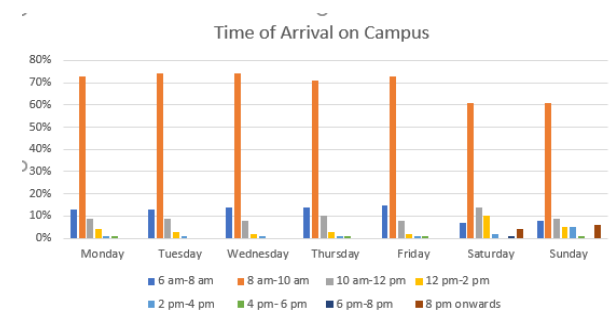


Fig 1. Arrival Time of Student and Staff

- Creche, Invent and Library are the most preferred parking after Multistorey parking in DCU.

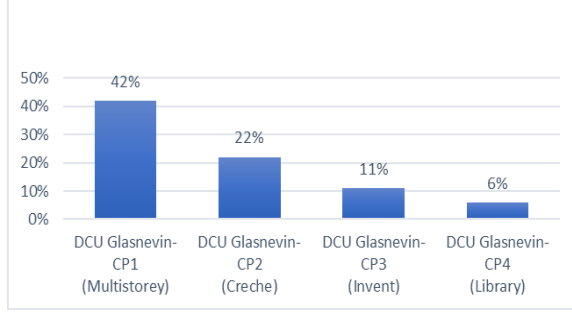


Fig 2. Preferred Parking

C. Survey 2013 and 2016

Summary report of surveys of years 2013 and 2016 were used for further analysis. Survey data of **2013** shows that **3036** people use Car for coming to DCU and in **2016** it is increased to **4889**. Although a total number of parking spaces available in DCU (all campuses) is 1837 and above study supports that it is difficult to find parking spaces in DCU.

D. Web Scraping Multistorey

Data has been gathered via web scraping through DublinCity website^[6] for Multistorey parking of City Centre from 24 Dec, 2016-26 Feb, 2017. Initially, plan was to check DCU Multistorey parking model with City Centre Multistorey. Due to unavailability of DCU Multistorey parking data the City Centre Multistorey parking data is available for future studies.

The Datasets, Surveys, Visualizations are available in the GitLab Repository.^[9]

V. DATA DESCRIPTION AND CREATION

After observing all data from above methods, data created manually for all three parking. Assumptions were made while creating data and it can impact results. Data which was collected from sensors or through other devices was available and had been used for this study can show different results as compared to result we will get from this study.

The time of the day is of great important factor impact parking space^[7]. The results outcome shows that holidays, weekdays and time of the day significantly impact parking space. Greengard^[8] works also has a comparable conclusion and support this study, his work argues that the traffic volume is different during weekdays, holidays making it difficult to find a parking space^[8]. Moreover, the time can be divided into five factors, which can impact the parking spaces: time of the day, day, month, bank holidays and vacation time.

Exploring manually collected by data, Surveys^[9], DCU calendar^[10] and previous works below factors were considered which impact car parking:

- Holidays (Banks/Public Holidays, Festival, Weekends, Vacations)
- Weekdays
- Time Frames
- Exam Study Week
- Examination Week.
- Reading Week

A. Creation of Data

Below steps were followed for creating Data for analysis purpose

- In the initial stage, data has been created on above factors for all three parking. Each file contains 1368 rows and below columns. Data has been created manually for months from Feb 2017 -May 2017 on the observations of data gathering methods.

Column Name	Data Description	Data Source Example
Date	Date on Particular Weekday of Months from Feb-May	For example :12-Feb-2017
Hour	Time of the Day	Hours of the day like 0,1,2
Holiday_Categories	0-Working Day 1-Holiday 2-Reading week 3-Exam study week 4-Examinations	Represent Holidays ,Working Day of DCU Calendar
Free Parking Spaces	Parking Occupancy	For example : 12,20
Free Space %	Calculated by using below formula: (Free Parking Spaces/Total Number of Parking Spaces) *100	For example :12%,0%
Occupancy(%)	For displaying information correctly below formula was applied: 100 -Free Space%	For example:100-12=88

Fig 3. Initial Data Description

- For creating prediction model and correlation matrix for every factors dataset was further divided using Pandas(python) method **get_dummies** to create dummy features of the nominal variable(Holiday). After splitting "**Holiday**" column according to factors below columns were finalized after renaming them:

Column Name	Data Description
Hour	Date on Particular Weekday of Months from Feb-May
Occupancy(%)	Showing Occupancy %
Working_Day	if 1 it is Working Day else 0
Holiday	if 1 it is Holiday else 0
Reading_Week	if 1 it is Reading Week else 0
Study_Week	if 1 it is Study Week else 0
Exam_Week	if 1 it is Exam Week else 0

Fig 4. Final Data Description

VI. IMPLEMENTATION AND EVALUATION OF HYPOTHESIS

A. Evaluation of Hypothesis

For evaluation of Hypothesis that identified factors considerably impact DCU car parking below mentioned parking have been used and findings have shown below:

1. Creche
2. Library
3. Invent

Below graph show trend of all three parking by considering all factors according to **Hour vs Hourly Parking Space % Trend**

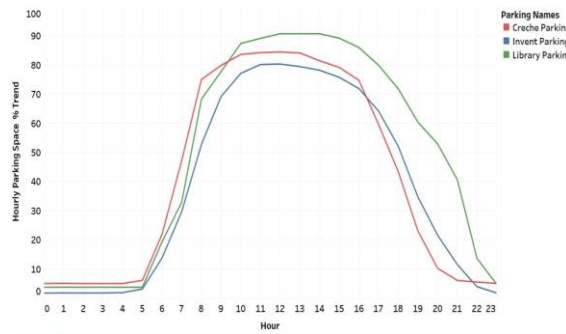


Fig 5. Parking Trend of all Parking (Hour vs Hourly Parking Space % Trend)

From above graph below observations can be drawn:

- All the parking follow slightly different trends.
- All parking follows the non-normal distribution.
- From 10 am to 4 PM all the parking almost remains full.
- Creche parking starts getting empty after 8 PM it is only 10% full.
- Invent also follow the same trend as Creche.
- Library Parking starts getting empty after 9 PM. We can conclude that as Library most of the days open until 9 PM when session and exams are in process people study and leave campus when the library is closed.

After that individual parking were studied to identify the factors which impact car parking at different times and for evaluation of null Hypothesis(H_0).

1. Creche

Creche parking is one of the preferable and busiest parking of DCU due to its proximity with Helix. It contains 235 spaces. Study and analysis of Creche parking data has provided meaningful insights. Below methods were chosen to do data exploration and find factors which impact car parking:

a) Boxplot

Boxplot was created to see the distribution of data by taking Hour and Occupancy in consideration. Below diagram shows data follow a non-normal distribution.

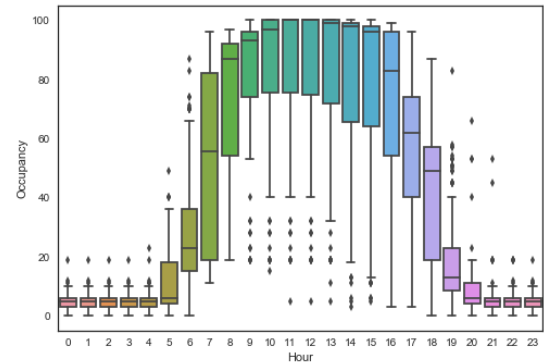


Fig 6. Distribution of Creche Parking (Hour vs Occupancy (%))

b) Trend Graph

Trend graph was used to display parking at different times with different factors. The graph has been generated by using all factors, Hours and average on Occupancy (%).

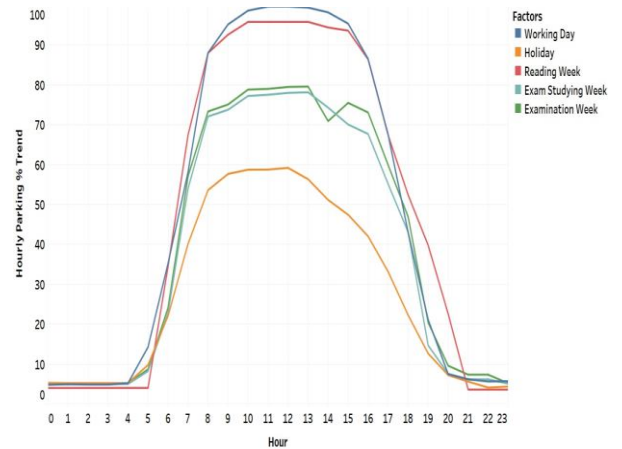


Fig 7. Factors Impact Creche Car Parking (Hour vs Hourly Parking (%) Trend)

From above graph below observations can be drawn:

- On working days, parking remains 100% occupied from 10 AM-4 PM
- Holidays impact the parking as many spaces are available during that time.
- During Reading week parking follows different trend than on working days. Parking starts getting empty after 9 PM. It might be because student study longer hours in library
- Examination and Exam study week also show significant impact on parking spaces.

c) Correlation Matrix

For further strengthening our Hypothesis we found the correlation among all factors and Occupancy (%) by creating correlation matrix. It strengthens our above analysis from the graph:

- Parking spaces are correlated with time with a correlation of 0.08.
- Holidays correlated with Parking spaces with high correlation of 0.2
- Parking spaces are inversely correlated with working days with a correlation of -0.2.
- Reading week is significantly related with parking spaces with a correlation of 0.04.
- Study week and Examination week doesn't have that much impact on the behavior of parking space trend.

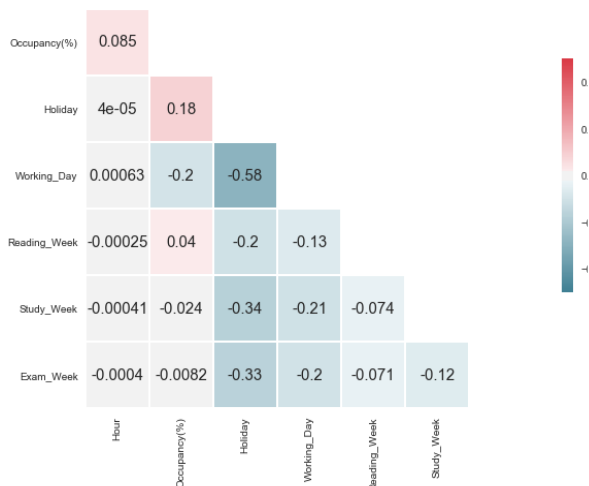


Fig 8. Correlation Matrix Creche Car Parking

2. Invent

Invent parking is second preferable parking of DCU after Creche. It contains 107 spaces. This was chosen for study due to its location near Invent, Library and Nursing Building. Below methods were chosen to do data exploration and find factors which impact car parking:

a) Boxplot

Boxplot was created to see the distribution of data by taking Hour and Occupancy in consideration. Below diagram shows our data follows a non-normal distribution.

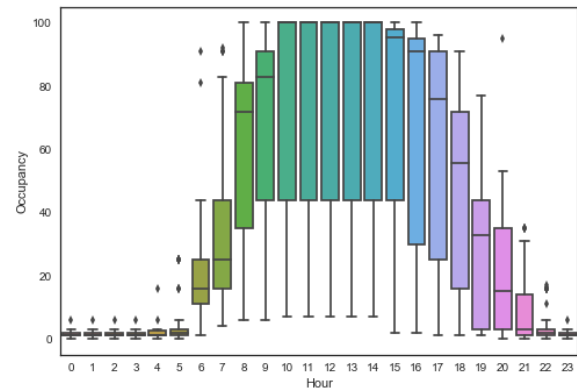


Fig 9. Distribution of Creche Parking (Hour vs Occupancy (%))

b) Trend Graph

Trend graph was used to display parking at different times with different factors. The graph has been generated by using all factors, Hours and average on Occupancy (%).

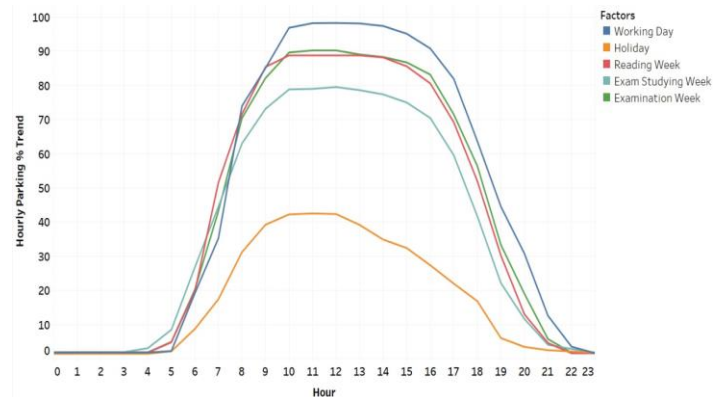


Fig 10. Factors Impact Invent Car Parking (Hour vs Hourly Parking (%) Trend)

From above graph below observations can be drawn:

- On working, days parking remains from 100% occupied from 10 AM-3 PM.
- Holidays impact the parking as many spaces are available during that time.
- Reading week and Examination week almost follow the same pattern.
- Exam study week is almost remaining 70% full of 9 AM – 4 PM and there are chances of finding parking spaces.

c) Correlation Matrix

For further strengthening our Hypothesis we found the correlation among all factors and Occupancy (%). It strengthens our above analysis from the graph:

- Parking spaces are correlated with time with a high correlation of 0.18.
- Holidays correlated with Parking spaces with high correlation of 0.24
- Parking spaces are inversely correlated with working with a correlation of -0.31.
- Study week and Examination week doesn't have that much impact on the behavior of parking space trend.

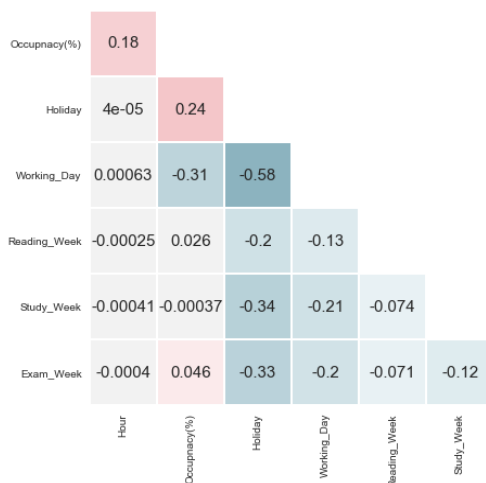


Fig 11. Correlation Matrix Invent Car Parking

3. Library

Library parking is the third busiest parking in DCU. It contains 100 spaces. Study and analysis of Creche parking data has provided meaningful insights. Below methods were chosen to do data exploration and find factors:

a) Boxplot

Boxplot was shown to see the distribution of data. Below diagram shows our data follows non-normal distribution

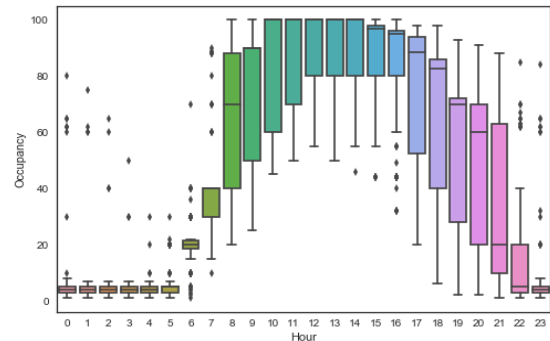


Fig 12. Distribution of Library Parking (Hour vs Occupancy (%))

b) Trend Graph

Trend graph was used to display parking at different times with different factors. The graph has been generated by using all factors, Hours and average on Occupancy (%).

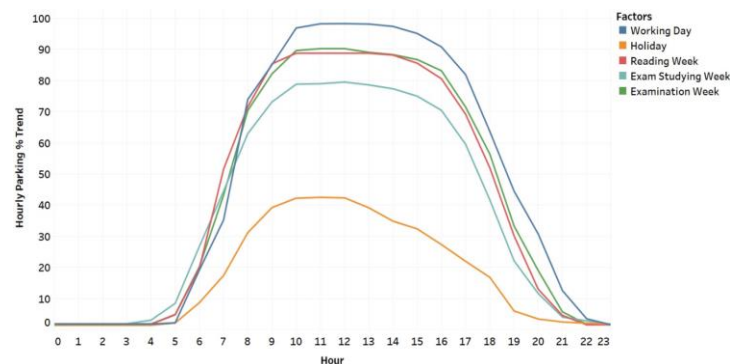


Fig 12. Factors Impact Library Car Parking (Hour vs Hourly Parking (%) Trend)

From above graph below observations can be drawn:

- On working days parking remains 100% occupied from 10 AM-2 PM.
- Holidays impact the parking and parking spaces are available.
- Library parking starts getting empty after 9 PM as compared to other parking.
- Examination and Reading week follow a similar pattern, we can interpret it is easy to find some places in parking around this time.
- Exam studying week shows significant impact on Library parking. It shows at peak time parking is 75% full. This pattern helps us

interpret that students prefer to study at home more than the library during exam time.

c) Correlation Matrix

For further strengthening our Hypothesis we found the correlation among all factors and Occupancy (%) by creating correlation matrix. It strengthens our above analysis from the graph:

- Parking spaces are correlated with time with a high correlation of 0.33.
- Holidays correlated with Parking spaces with a high correlation of 0.16.
- Parking spaces are inversely correlated with Working Day with a correlation of -0.25.
- Study week and Examination week have that much impact on the behavior of parking space trend and have a correlation of 0.044 and 0.041 respectively.

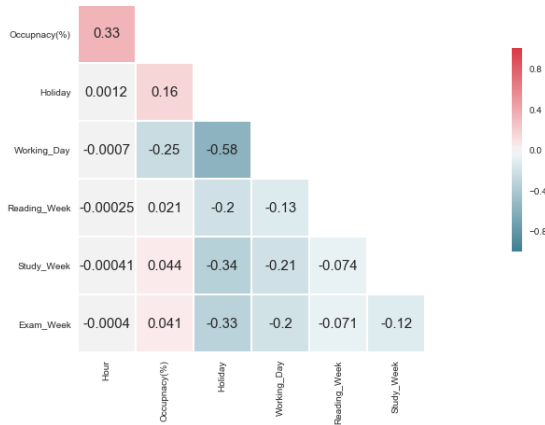


Fig 13. Correlation Matrix Invent Car Parking

C. Evaluation of Hypothesis

After studying all parking behavior with respect to different factors and above methods, it can be concluded that the initial Null Hypothesis(H0) was true. Parking spaces get impacted by various factors and due to this alternate Hypothesis (H1) can be rejected.

V. IMPLEMENTATION OF PREDICTION FUNCTION

The next step in this study was to develop prediction model (based on probability distribution) by considering one factor which impacts car parking.

In a final step, the parking model will be tested on other parking to check its accuracy and compatibility with other parking.

For developing the prediction model Working Day(factor) w.r.t Hour(Time) which considerably impact the parking spaces has been considered.

The parking are following non-normal distribution as shown previously. After exploring many non-normal distributions (like Beta, Gamma, Poisson, Exponential) the best accuracy was received with Polynomial Distribution with power 10 on “Hour” as independent variable and “Occupancy” as dependent variable using R.

The model has been trained by using Creche Car parking data and high R² value of 0.92 was achieved. The 90% of creche car parking data has been used to train the model and rest 10% has been used for testing. The developed model for working day has been tested then on other parking Invent and Library to test whether the model can be, for other parking. After testing model on Library, and Invent parking effective results were obtained.

For calculating accuracy Root Mean Square Errors (RMS errors) has been considered in this study

RMS Error

$$= \sqrt{\frac{1}{n} \sum_{i=1}^n (predicted_i - observed_i)^2}$$

The results have been shown in below table:

Table 1. RMSE and Accuracy

Parking Names	Creche	Invent	Library
RMSE	13.95	15.78	25.08
Accuracy (%)	86.05	84.21	74.91

The accuracy of Library car parking is less (74%) as compared to Creche and Invent car parking. The results are expected as Library opening and closing timings are different so this parking follows different trends as compared to other parking as shown previously.

Below figures shows distribution of Actual and Predicted values of all the parking (Creche, Invent, Library) which are tested using model developed for Working Day through Creche data.

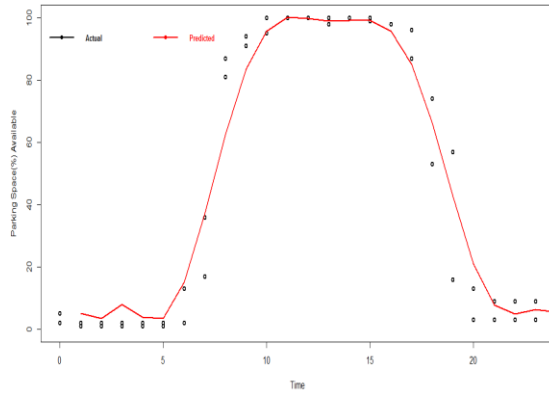


Fig 14. Prediction vs Actual Comparison of Creche Car Parking on Working Day (Time (Hours) vs Parking Space (%) Available)

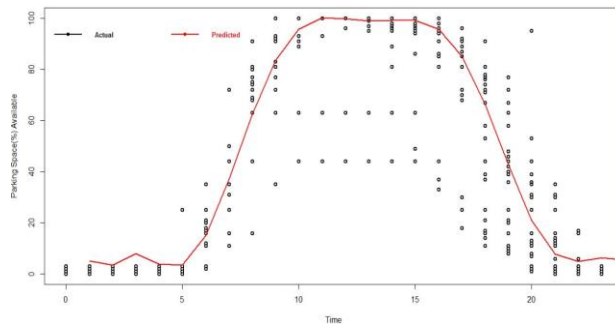


Fig 15. Prediction vs Actual Comparison of Invent Car Parking on Working Day (Time (Hours) vs Parking Space (%) Available)

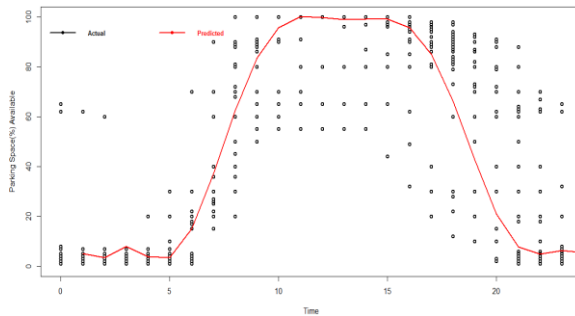


Fig 16. Prediction vs Actual Comparison of Library Car Parking on Working Day (Time (Hours) vs Parking Space (%) Available)

VI. CONCLUSION

The main aim of the paper was to identify the factors which impact prediction of DCU car parking and developing a prediction model which can be used for other parking. The study shows that Weekday, Holiday, Time frames, Exam Study Week, Reading week and Examination are the factors which impact car parking at DCU. The prediction model which was

developed for Working Day is significantly helpful in predicting parking spaces.

In future, this model can be studied further to improve the results and can be developed for further factors. Initially, in this study the plan was to develop the prediction model for Multistorey parking of DCU and use that model to Dublin, City Centre Multistorey Car parking. Due to this City Centre parking data was scraped from DublinCity website for the period of 25 Dec, 2017 - 26 Feb, 2017, which is available to use for future studies. Later due to unavailability of DCU Multistorey car parking data the model was not developed. In future, this scraped data can be used for development of prediction model for Multistorey car parking. This paper contribution serves as to provide a basis for developing efficient car parking prediction model for all types of parking.

ACKNOWLEDGEMENT

I am very grateful for the help, encouragement and guidance provided by my supervisors Dr. Suzzane Little and Dr. Marija Bezbradica in completing the thesis.

I would also like to thanks Samantha Fahy (Manager, Sustainability DCU) from Real Estate Department DCU for helping me by including my questions in Surveys for this study and providing data from previous Surveys.

REFERENCES

- [1] Le Fauconnier, A., & Gantelet, E. (2006). "The time looking for a parking space: strategies," associated nuisances and stakes of parking management in France. Paper presented at the Proceedings of the European Transport Conference (ETC) 2006, Strasbourg, France.
- [2] Christoph Pflügler¹, Thomas Köhn¹, Maximilian Schrieck¹, Manuel Wiesel¹ und Helmut Krcmar¹ Deliverables: "Predicting the Availability of Parking Spaces with Publicly Available Data". Siemens Mobility. (2015). Vision der intelligenten Stadt – Bewegung durch Innovation."
- [3] Retrieved from <http://www.mobility.siemens.com/mobility/global/de/nahverkehr/strassenverkehr/integrated-smart-parking-solution/seiten/integrated-smart-parking-solution.aspx>
- [4] Nawaz, S., Efstratiou, C., & Mascolo, C. (2013). "ParkSense: a smartphone based sensing system for on-street parking." Paper presented at the Proceedings of the 19th annual international conference on Mobile computing & networking, Miami, Florida, USA.
- [5] 2015, ParkYa, <https://parkya.com/>
- [6] Dublin Car Park Information, <http://www.dublincity.ie/dublintraffic/carparks.htm>
- [7] Zehe, W., Jinggang, Y., Jiangtao, L., & Xiuhua, Z. (2007). "Study on the Control Strategy of Parking Guidance System. Paper presented at the International

- Conference on Service Systems and Service Management, 2007”.
- [8] Greengard, S. (2015). Between the lines. Communications of the ACM, 58(6), 15-17.
- [9] GitLab Link DCU Code and Data,
<https://gitlab.computing.dcu.ie/aggarwa3/Factors-Impacting-Prediction-of-DCU-Car-Parking.git>
- [10] DCU Academic Calendar
2017,https://www4.dcu.ie/timetabling/ac_17_18.shtml