# COVID – 19 PANDEMIC DATA ANALYSIS

***A report of summer internship (2019-20)***

# BACHELOR OF TECHNOLOGY IN

**INFORMATION TECHNOLOGY /**

**COMPUTER SCIENCE AND INFORMATION TECHNOLOGY**

**Submitted by**

**Abhineet Sharma (1802913007)**

**2nd Year (Section : A)**

**Supervised by**

Prof. Navpreet Kaur



# DEPARTMENT OF INFORMATION TECHNOLOGY /

**DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY**

**KIET GROUP OF INSTITUTIONS, GHAZIABAD, UTTAR PARDESH**

**(Affiliated to Dr. A.P.J. Abdul Kalam Technical University, Lucknow, UP, India) Session 2019-20**

Certificate PART – A

1. Introduction
2. Project Overview
3. Implementation (with snap shots)
4. Conclusion
5. Daily Log PART – B
6. Summary of Paper-1
7. Summary of Paper-2
8. Summary of Paper-3
9. Summary of Paper-4
10. Summary of Paper-5

PART –C

1. Certificate of MOOC-1
2. Certificate of MOOC-2

# CERTIFICATE

This is to certify that the internship project report entitled, “**COVID 19 Pandemic Data Analysis”** submitted by **Abhineet Sharma** in the Department of Information Technology/ CSI of **KIET Group of Institutions, Ghaziabad**, affiliated to Dr. A. P. J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India, is a record of candidate summer internship work carried out by him / them under my supervision and guidance and is worthy of consideration for the same.

# Signature of Supervisor:

**Supervisor Name:** Prof. Navpreet Kaur

# Date: 29th July 2020

**PART - A**

INTRODUCTION

The **COVID‑19 pandemic**, also known as the **coronavirus pandemic**, is an ongoing global [pandemic](https://en.wikipedia.org/wiki/Pandemic) of [coronavirus disease 2019](https://en.wikipedia.org/wiki/Coronavirus_disease_2019) (COVID‑19), caused by [severe acute respiratory syndrome coronavirus 2](https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome_coronavirus_2) (SARS‑CoV‑2).The outbreak was first identified in December 2019 in [Wuhan](https://en.wikipedia.org/wiki/Wuhan), China.

The [World Health Organization](https://en.wikipedia.org/wiki/World_Health_Organization) declared the outbreak a [Public Health Emergency of International Concern](https://en.wikipedia.org/wiki/Public_Health_Emergency_of_International_Concern) on 30 January 2020 and a pandemic on 11 March. As of 29 July 2020, [more than 16.7 million cases](https://en.wikipedia.org/wiki/COVID-19_pandemic_cases) of COVID‑19 have been reported in [more than 188 countries and territories](https://en.wikipedia.org/wiki/COVID-19_pandemic_by_country_and_territory), resulting in [more than 660,000 deaths](https://en.wikipedia.org/wiki/COVID-19_pandemic_deaths); more than 9.76 million people have recovered.

In order to analyze all the factors that affect the outcome of this pandemic, we studied different datasets that were present online on sites like Kaggle.

After cleaning the data we used different libraries in RStudio to explore different aspects and to visualize the data in general.

The main objective was to understand how the pandemic unfolded in the entire world and then predict the total number of fatalities in different countries with different counter – measures and policies by using different features of the dataset.

This way the analysis could help us to understand the pandemic in more detail with actual comprehension of all the factors affecting the world.

PROJECT OVERVIEW

After downloading the datasets from Kaggle, we decided to implement 3 main regression algorithms to predict the fatalities in this diabolical pandemic.

1.Linear Regression Model

lm([target variable] ~ [predictor variables], data = [data source])

2.Support Vector Regression

#Regression with SVM

modelsvm = svm(Y~X,data)

3.Random Forests

randomForest(formula, data)

The goal was to determine the best suitable algorithm for predicting the total number of deaths.

IMPLEMENTATION

Linear Regression Model

library(ggplot2)

library(gganimate)

library(caTools)

library(ISLR)

library(ggthemes)

library(dplyr)

library(corrgram)

library(corrplot)

COVID<-read.csv("country\_wise\_latest.csv")

COVID

#Interactive plot

p <- ggplot(

COVID,

aes(x = COVID$Deaths, y=COVID$Active, size = COVID$Deaths, colour = COVID$Country.Region)

) +

geom\_point(show.legend = FALSE, alpha = 0.7) +

scale\_color\_viridis\_d() +

scale\_size(range = c(2, 12)) +

scale\_x\_log10() +

labs(x = "Deaths", y = "Active")

p+transition\_time(COVID$Recovered)+labs(title = "Recoveries: {frame\_time}")

#Corrgram and Corrplot

c.cols<-sapply(COVID,is.numeric)

cordata<-cor(COVID[,c.cols])

print(corrplot(cordata,method="color"))

corrgram(COVID)

final<-COVID[,-c(9,10,11,14)]

c.cols<-sapply(final,is.numeric)

cordata<-cor(final[,c.cols])

print(corrplot(cordata,method="color"))

corrgram(final)

# Linear Regression / Model Training

is.na(final)<-sapply(final,is.infinite)

na.omit(final)

set.seed(101)

p1<-sample.split(final$Deaths,SplitRatio = 0.7)

train<-subset(final,p1==TRUE)

test<-subset(final,p1==FALSE)

pmodel<-lm(Deaths~Confirmed+Active+Recovered,data=train)

summary(pmodel)

layout(matrix(c(1,2,3,4),2,2))

plot(pmodel)

#Residual Fitting

res<-residuals(pmodel)

res<-as.data.frame(res)

ggplot(res,aes(res))+geom\_histogram(fill="red")

#Prediction

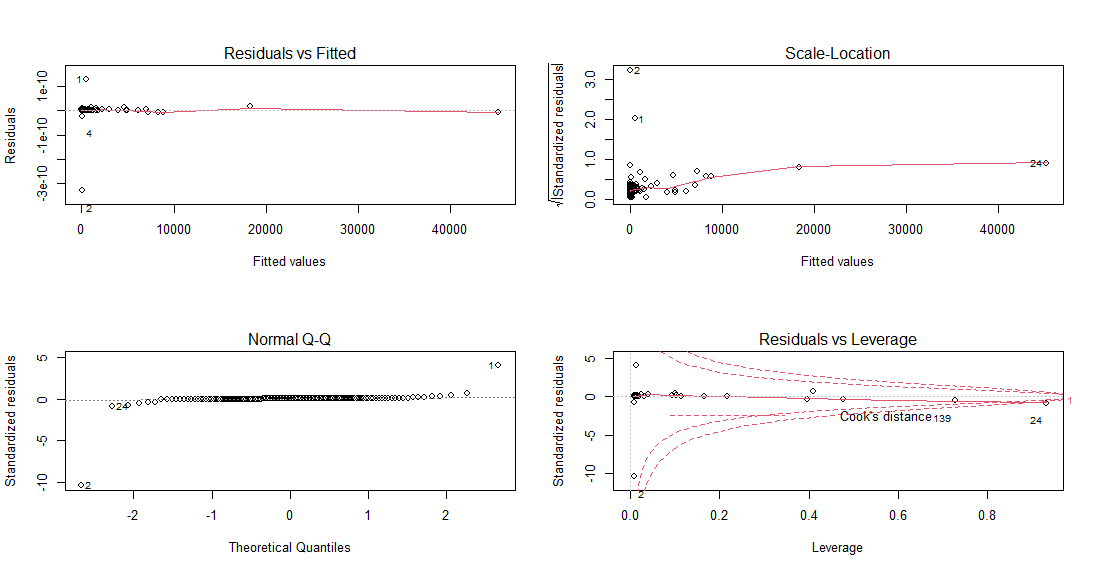
lp<-predict(pmodel,test)

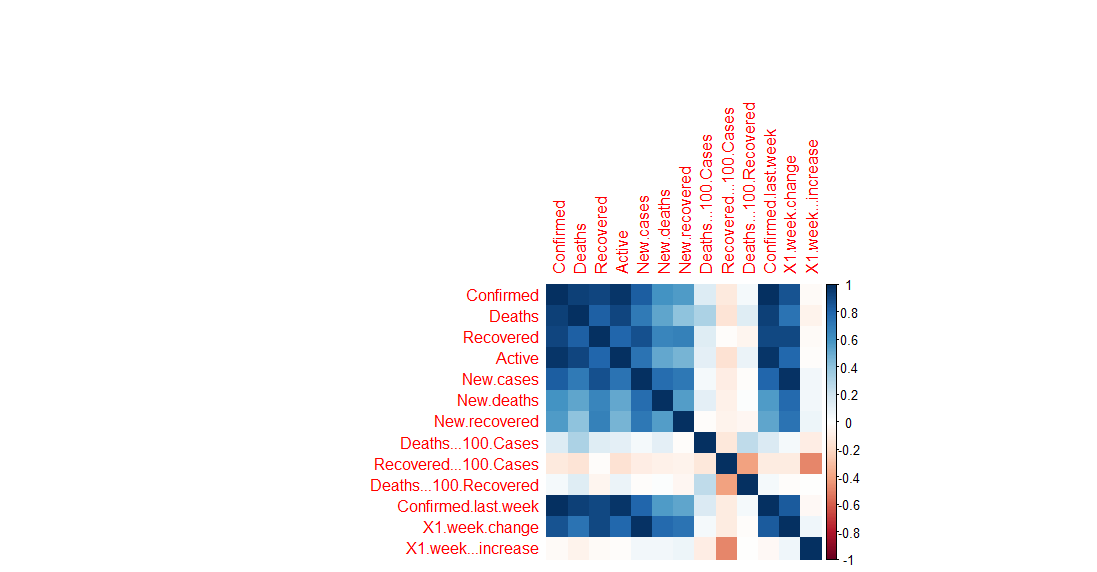
print(lp)

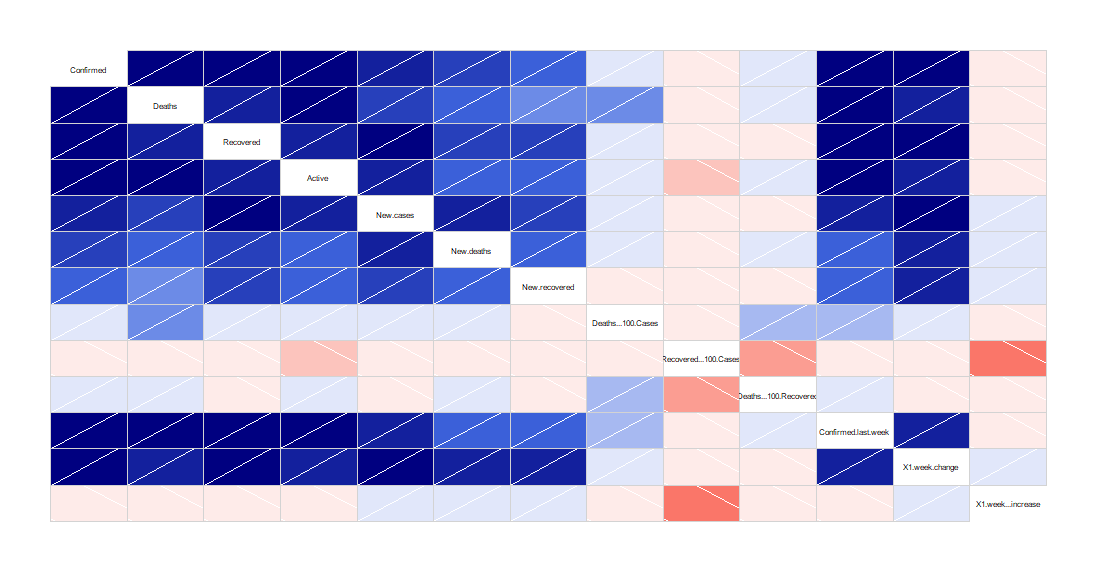
lres<-cbind(lp,test$Deaths)

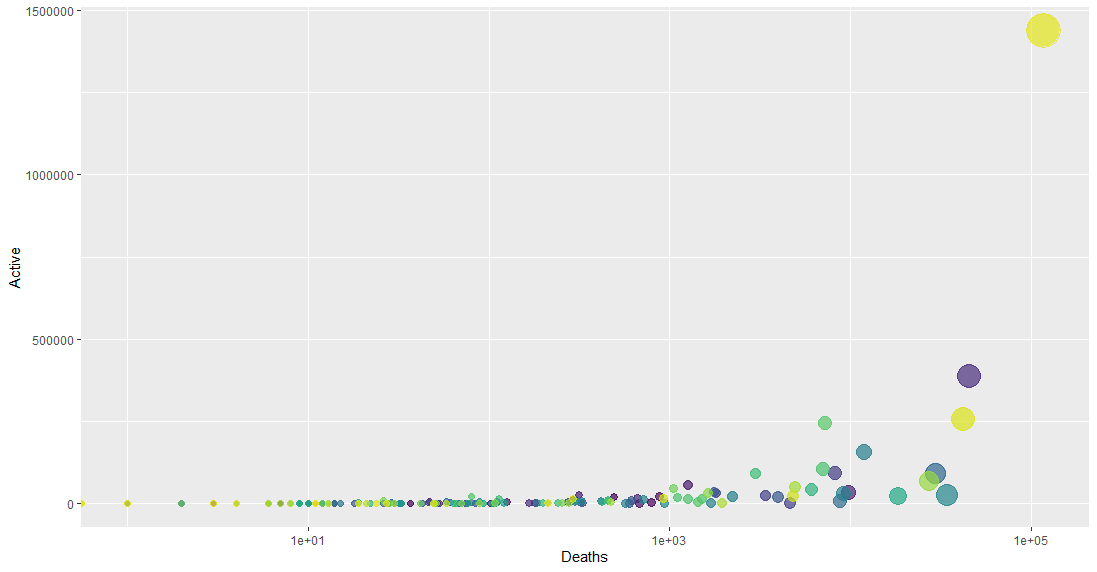
colnames(lres)<-c("Predicted","Actual")

print(lres)

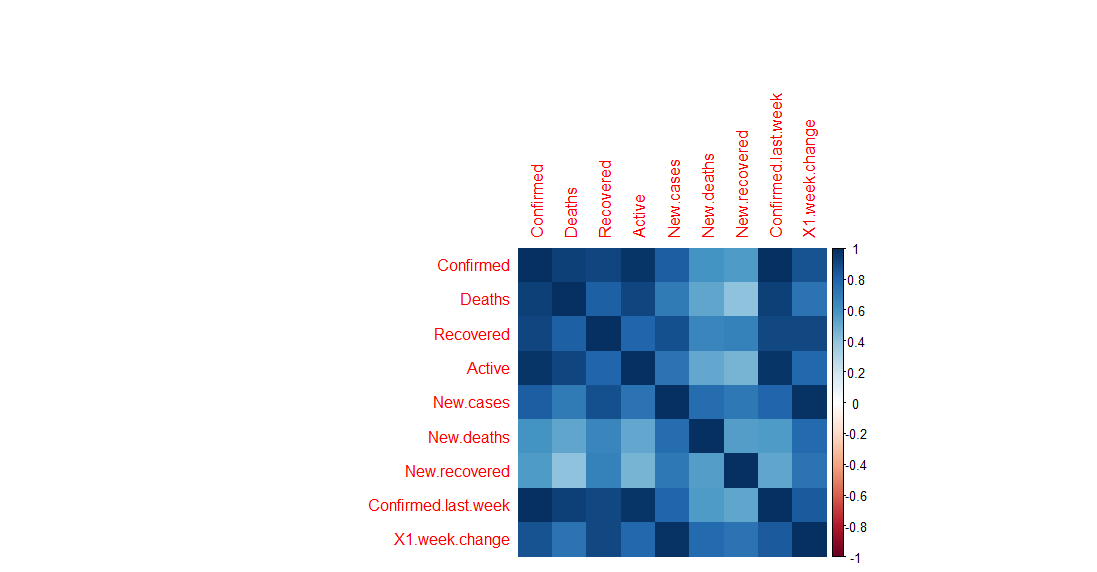


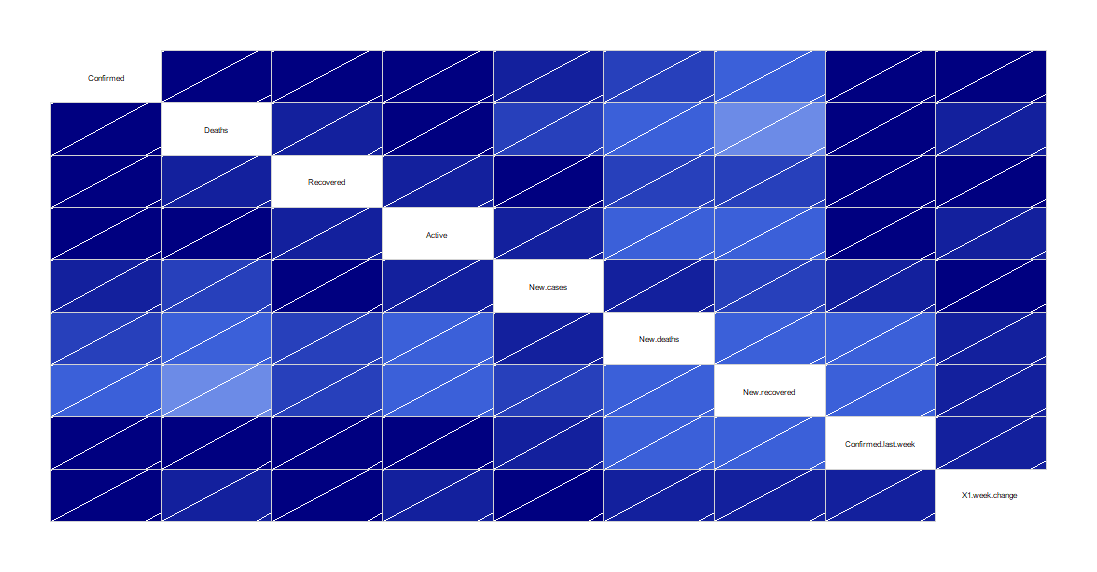






STRONG CORRELATION BETWEEN FEATURES





SVR (SUPPORT VECTOR REGRESSION) MODEL

# <Support Vector Regression>

# <Algorithm Comparision>

# Creating the model

# Loading Database

library(ISLR)

library(e1071)

library(dplyr)

library(mltools)

COVID<-read.csv("country\_wise\_latest.csv")

COVID

#COVID\_2<-COVID%>%select(Active,Deaths)

#COVID\_2

#plotting the linear Relationship

plot(COVID\_2,pch=16)

Lmodel<-lm(Deaths~.,COVID)

abline(Lmodel)

#</ / / Checking the Accuracy of the model / / />

PRED<-predict(Lmodel,COVID\_2)

points(COVID\_2$Active,PRED,col="blue",pch=4)

#Making the SVR Model

SVR<-svm(Deaths~Active,COVID\_2,kernel="linear",cost=10,scale=FALSE)

SVR<-svm(Deaths~.,data = COVID,kernel="linear",cost=10,scale = FALSE)

print(SVR)

plot(SVR,COVID)

PRED2<-predict(SVR,COVID)

points(COVID$Deaths,PRED2,col="red",pch=4)

# <Computing the Error>

error<-(COVID\_2$Deaths-PRED2)

svrPRED\_rmse<-rmse(error)

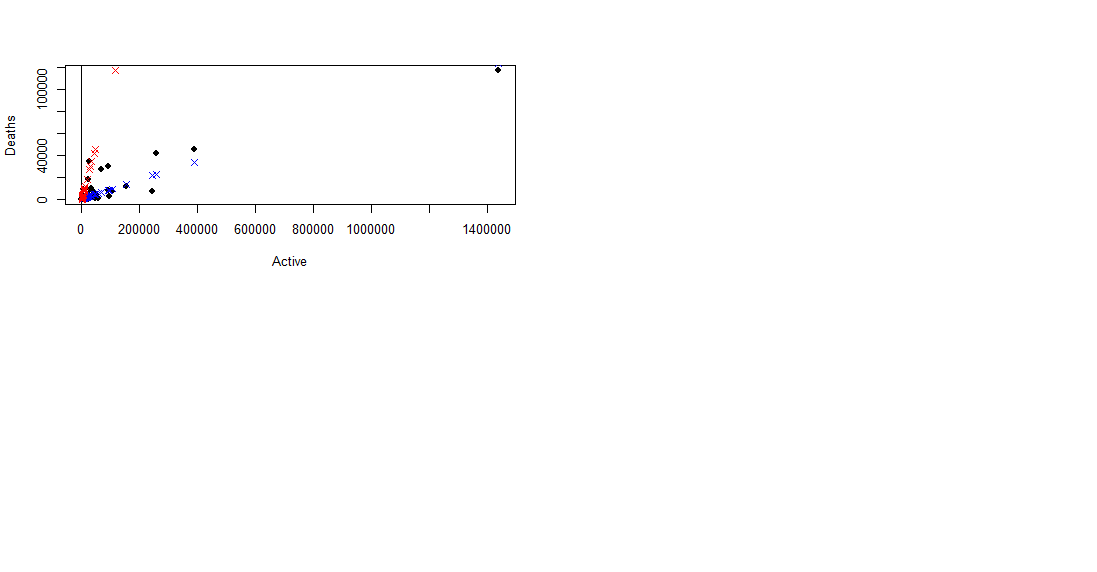
print(svrPRED\_rmse)

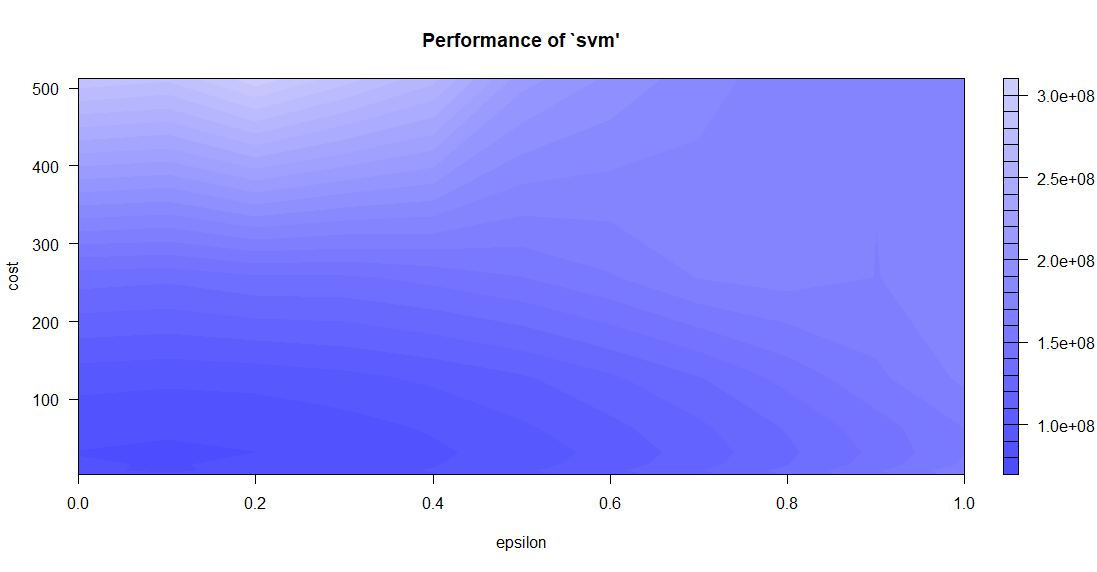
#Tuning the Result

SVRtune<-tune(svm,Deaths~Active,data=COVID\_2,ranges=list(epsilon=seq(0,1,0.1),cost=2^(2:9)))

print(SVRtune)

plot(SVRtune)





library(ISLR)

library(e1071)

library(dplyr)

library(ggplot2)

#COVID\_3<-COVID%>%select(Confirmed,Deaths,Recovered,Active,New.cases,New.deaths,New.recovered,Deaths...100.Cases,Recovered...100.Cases,Deaths...100.Recovered,Confirmed.last.week,X1.week.change,X1.week...increase)

#COVID\_3

is.na(check2)<-sapply(check2,is.infinite)

na.omit(check2)

#Model Building

model<-svm(Deaths~.,data=check2,kernel="rbf")

tune\_results<-tune(svm,Deaths~.,data=check2,ranges=list(cost=c(0.1,1,10),gamma=c(0.5,1,15)))

tunedvals<-tune\_results$best.model

predicted<-predict(tunedvals,check2)

tunedvals

results<-cbind(predicted,check2$Deaths)

colnames(results)<-c("Predicted","Actual")

results<-as.data.frame(results)

print(head(results))

#Outlier detection

Random Forests and Regression Tree Model

#Random Forest

library(rpart.plot)

library(rpart)

library(party)

library(dplyr)

library(rsample)

#Using the COVID-19 pandemic dataset

check2

#check2 contains the dataset

set.seed(123)

split<-initial\_split(check2,prop = 0.7)

check\_train<-training(split)

check\_test<-training(split)

#Decision Tree

tree<-rpart(formula=Deaths~.,method="anova",data=check\_train,control = list(minsplit = 10, maxdepth = 12, xval = 10))

rpart.plot(tree)

plotcp(tree)

tree2<-ctree(formula=Deaths~Confirmed+Recovered+Active,data=check\_train)

plot(tree2)

rpart.plot(tree)

#RandomForest

#Building Model using RandomForest

rf.model<-randomForest(Deaths~.,data=check2)

rf.model$ntree

rf.model$mse

summary(rf.model)

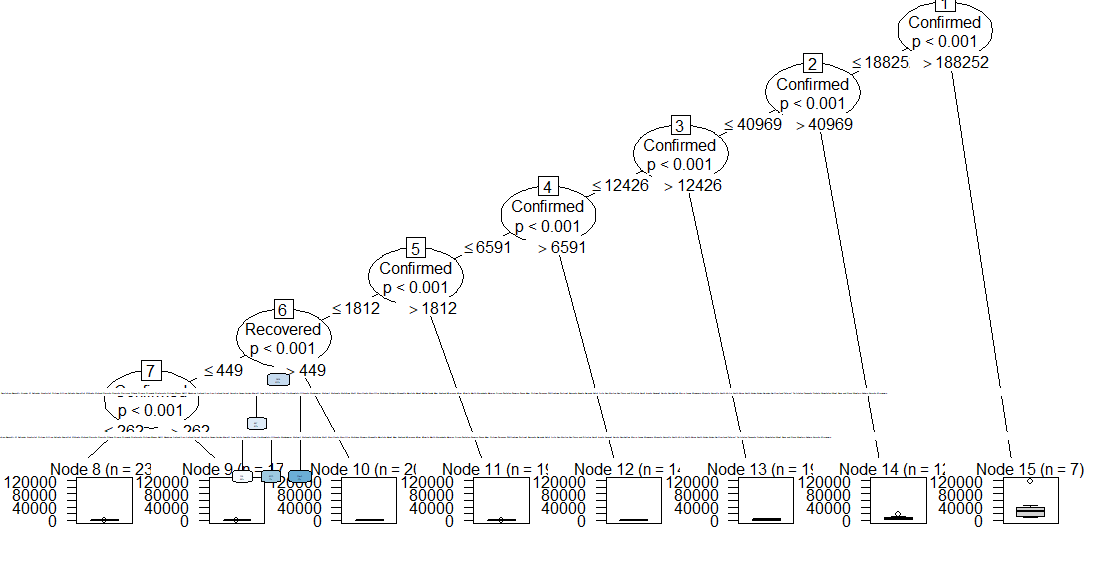
#Predicted V/S Actual

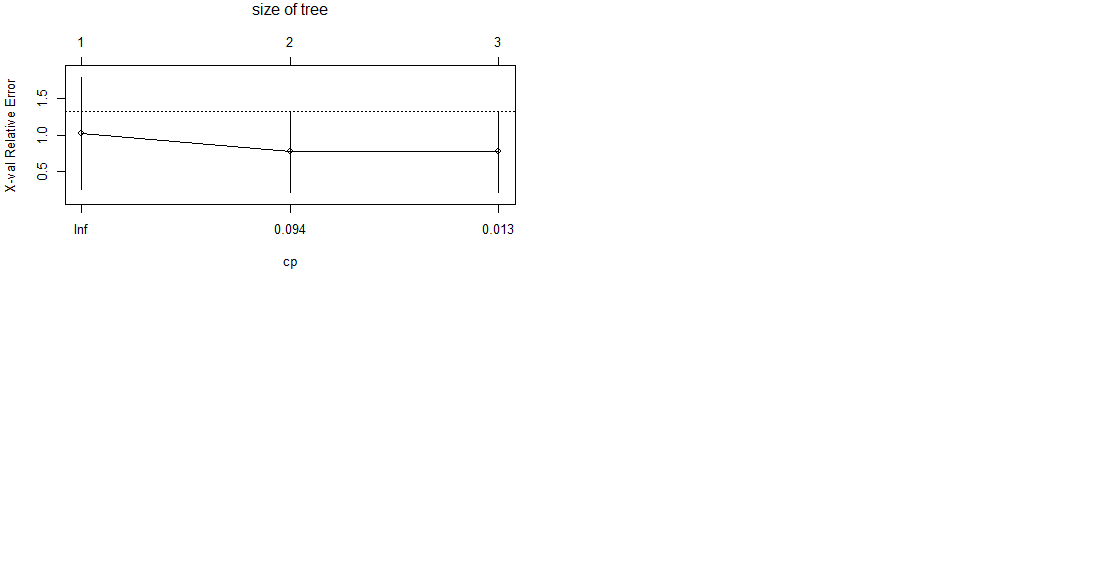
AvP<-cbind(rf.model$predicted,check2$Deaths)

colnames(AvP)<-c("Predicted","Actual")

print(head(AvP))

importance(rf.model)





CONCLUSION

LINEAR REGRESSION MODEL (Actual V/s Predicted)

Predicted Actual

11 356.785706 126

14 3325.918760 1262

17 2146.837232 9663

22 68.529850 168

25 -60.445555 3

26 70.658269 181

36 8494.896182 3383

38 1858.486154 1808

40 -34.193574 27

43 146.138121 46

48 535.670282 598

49 130.702719 43

SUPPORT VECTOR REGRESSION MODEL (Actual V/s Predicted)

Predicted Actual

1 186.691004 491

2 196.820497 37

3 105.337624 788

4 1106.632356 52

5 -640.202795 6

6 600.049630 3

7 275.121843 878

8 -583.056449 293

9 -294.887455 102

10 574.953403 681

11 -301.253117 126

12 -190.540470 11

RANDOM FOREST MODEL (Actual V/s Predicted)

Predicted Actual

1 9.997172e+02 491

2 7.092879e+01 37

3 2.915571e+02 788

4 4.217980e+01 52

5 6.652035e+00 6

6 2.411658e+00 3

7 1.568167e+03 878

8 4.229629e+02 293

9 3.328007e+02 102

10 7.403568e+02 681

11 2.702118e+02 126

12 5.248626e+00 11

**Random Forest Model has the best Accuracy.**

DAILY LOG

|  |  |
| --- | --- |
| **Name of Student** | Abhineet Sharma |
| **Roll No.** | 1802913007 |
| **Name of Course** | Data Science using R |
| **Date of Commencement of Training:** | 18th May 2020 |
| **Date of Completion of Training:** | 30th July 2020 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No.** | | **Please specify the learning of the day** | | **Date** | |
| 1 | | Introduction to Data Science | | 18.05.20 | |
| 2 | | Basics of Data Science, Bigdata and Data Analytics life cycle | | 19.05.20 | |
| 3 | | Data Types, variables and operators | | 21.05.20 | |
| 4 | | Decision Making | | 22.05.20 | |
| 5 | | Loops in R | | 25.05.20 | |
| 6 | | Loops in R | | 26.05.20 | |
| 7 | | Functions and Strings | | 28.05.20 | |
| 8 | | Vectors, Lists and Arrays | | 29.05.20 | |
| 9 | | Data Visualization and Exploration | | 01.06.20 | |
| 10 | | Importing External CSV and XML | | 02.06.20 | |
| 11 | | Handling missing values and detection of outliers | | 04.06.20 | |
| 12 | | Apply Function | | 05.06.20 | |
| 13 | | WhatsApp Text Analysis | | 15.06.20 | |
| 14 | | Natural Language Processing (Twitter Sentiment Analysis) | | 16.06.20 | |
| 15 | | Time Series Algorithm and decision trees | | 18.06.20 | |
| 16 | | Apriori Algorithm (Market Basket Analysis) | | 20.06.20 | |
| 17 | | Decision Trees and ID3 Algorithm | | 21.06.20 | |
| 18 | | K means Clustering and feature selection using RandomForests | | 22.06.20 | |
| 19 | | Logistic Regression | | 23.06.20 |
| 20 | | Multiple Linear Regression with ANOVA test | | 08.07.20 |
| 21 | | One-way ANOVA, two-way ANOVA, ANOVA with and without replicates | | 09.07.20 |
| 22 | | Extreme Gradient Boosting Algorithm | | 10.07.20 |
| 23 | | Extreme Gradient Boosting Algorithm | | 13.07.20 |
| 24 | | Naïve Bayes Theorem | | 14.07.20 |
| 25 | | Density Based Cluster Visualization | | 20.07.20 |
| 26 | | Support Vector Machines | | 21.07.20 |
| 27 | | Basics of RCloud | | 23.07.20 |
| 28 | | Goodness of fit and CHI SQUARE test | | 24.07.20 |
| 29 | | ARIMA model in time series algorithm and plotting ACF and PACF plots | | 27.07.20 |
| 30 | | Capturing the market annualized volatility using Apple Stock Price data | | 28.07.20 |

Supervisor Sign

(Prof. Navpreet Kaur)

**PART - B**

**RESEARCH PAPER 1:**

## Analysis of the Ebola Outbreak in 2014 and 2018 in West Africa and Congo by Using Artificial Adaptive Systems

**SUMMARY:**

The initial WHO epidemiologic investigation revealed the first (index) case was a 2-year-old child from Meliandou (a village in Gueckedou prefecture in Nzerekore region of southern Guinea), Longitude: −10.061, Latitude: 8.616. According to the initial investigation the child died on December 6, 2013 but a secondary investigation confirmed the death of a child at the end of December. After the death of the index case, the sister, mother, grandmother of the child and a nurse, became victims 2–5, died on December 29, 2013, December 13, 2013, January 1, 2014, and January 29, 2014, respectively. Evidently, the fourteenth patient, a health-care worker from Gueckedou, caused the spread of the virus to Macenta, Nzerekoré, and Kissidougou in February 2014. (Sylvain et al. [2014](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770), et al. 2014. Emergence of Zaire Ebola virus disease in Guinea. *The New England Journal of Medicine* 371 (15):1418–25. doi:10.1056/NEJMoa1404505. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0012&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.1056%2FNEJMoa1404505), [[PubMed]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0012&dbid=8&doi=10.1080%2F08839514.2020.1747770&key=24738640), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0012&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000342674800010), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2014&pages=1418-25&issue=15&author=B.+Sylvain&title=Emergence+of+Zaire+Ebola+virus+disease+in+Guinea))

The World Health Organization (WHO) was alerted of the Ebola outbreak in Guinea on March 23, 2014. The Ministry of Health (MoH) of Guinea notified the WHO of the rapidly spreading Ebola outbreak in the forested area south eastern of Guinea with a total of 49 cases including 29 deaths by March 22, 2014. The Ebola cases reported were in Guekedou (Longitude: −10.1336, Latitude: 8.5674), Macenta (Longitude: −9.4710, Latitude: 8.5435), Nzerekore (Longitude: −8.8238889, Latitude: 7.7472) and Kissidougou districts (Longitude:-10.1000, Latitude: 9.1833) (WHO Ebola Response Team [2014](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)WHO Ebola Response Team. 2014. Ebola Virus Disease in West Africa-The First 9 Months of the Epidemic and Projections. *The New England Journal of Medicine*371 (16). doi:10.1056/NEJMoa1411100. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0015&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.1056%2FNEJMoa1411100), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0015&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000342994700005), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2014&issue=16&author=WHO+Ebola+Response+Team&title=Ebola+Virus+Disease+in+West+Africa-The+First+9+Months+of+the+Epidemic+and+Projections)).

The recent (May 2018) Ebola outbreak was reported by WHO. According to the WHO reports the confirmed cases were in Bikoro and Iboka health zones, and in Wangata, one of the three health zones of Mbandaka (Ebola Virus Disease [2018](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Ebola Virus Disease, Democratic Republic of the Congo, External Situation Report 5, Heath Emergency Information and Risk Assessment, Date of Issue: 25 May 2018, Data as reported by 23 May 2018. <http://apps.who.int/iris/bitstream/handle/10665/272662/SITREP-EVD-DRC-20180525-eng.pdf>(Accessed: June 2018) [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2018&author=Ebola+Virus+Disease&title=Democratic+Republic+of+the+Congo%2C+External+Situation+Report+5%2C+Heath+Emergency+Information+and+Risk+Assessment%2C+Date+of+Issue%3A+25+May+2018%2C+Data+as+reported+by); Ebola virus disease-Democratic Republic of the Congo, Disease outbreak news [2018](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Ebola virus disease-Democratic Republic of the Congo, Disease outbreak news, 23May 2018, WHO, <http://www.who.int/csr/don/23-may-2018-ebola-drc/en/>(Accessed: June 2018) [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2018&author=Ebola+virus+disease-Democratic+Republic+of+the+Congo%2C+Disease+outbreak+news&title=WHO); Update on the Ebola outbreak in Democratic Republic of Congo, Medicines Sans Frontiers, Doctor without Borders, May 2018).

Different approaches have been used to study the Ebola outbreak such as artificial adaptive systems, a fixable spatiotemporal growth, and deterministic model, a Markov model to estimate the weekly average number of secondary cases and using *Mirador* to find the most informative clinical factors related to the Ebola virus Diseases (Buscema [2015](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M. 2015. Why Mathematical Computer Simulations Are the New Laboratory for Scientists. *Substance Use & Misuse* 50 (8–9):1068–78. doi:10.3109/10826084.2015.1012934. [[Taylor & Francis Online]](https://www.tandfonline.com/doi/10.3109/10826084.2015.1012934), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0001&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000361331800027), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2015&pages=1068-78&issue=8%E2%80%939&author=M.+Buscema&title=Why+Mathematical+Computer+Simulations+Are+the+New+Laboratory+for+Scientists); Clubri et al. [2016](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Clubri, A., T.Silver, T. Fradet, K. Retzepi, B. Fry, P. Sabeti, and T. S. Churcher Editor. 2016, Mar. Transforming Clinical Data into Actionable Prognosis Models: Machine-Learning Framework and Field-Deployable App to Predict Outcome of Ebola Patients. *PLoS Neglected Tropical Diseases* 10(3), 1-17. [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0006&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000373272500064), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2016&issue=3&author=A.+Clubri&author=T.+Silver&author=T.+Fradet&author=K.+Retzepi&author=B.+Fry&author=P.+Sabeti&author=T.+S.+Churcher&title=Transforming+Clinical+Data+into+Actionable+Prognosis+Models%3A+Machine-Learning+Framework+and+Field-Deployable+App+to+Predict+Outcome+of+Ebola+Patients); Santermans et al. [2016](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Santermans, E., E. Robesyn, T.Ganyani, B. Sudre, C. Faes, C.Quinten, B. W. Van, T. Haber, T.Kovac, R. F. Van, et al. 2016January 15. Spatiotemporal Evolution of Ebola Virus Disease at Sub-National Level during the 2014 West Africa Epidemic: Model Scrutiny and Data Meagreness.. *PloS One* 11 (1). doi:10.1371/journal.pone.0147172. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0010&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.1371%2Fjournal.pone.0147172), [[PubMed]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0010&dbid=8&doi=10.1080%2F08839514.2020.1747770&key=26771513), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0010&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000368628300069), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2016&issue=1&author=E.+Santermans&author=E.+Robesyn&author=T.+Ganyani&author=B.+Sudre&author=C.+Faes&author=C.+Quinten&author=B.+W.+Van&author=T.+Haber&author=T.+Kovac&author=R.+F.+Van&title=Spatiotemporal+Evolution+of+Ebola+Virus+Disease+at+Sub-National+Level+during+the+2014+West+Africa+Epidemic%3A+Model+Scrutiny+and+Data+Meagreness.); Wayne et al. [2015](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Wayne, T. A., L. W. Enanoria, F. Liu, D. Gao, S.Ackley, J. Scott, M. Deiner, E.Mwebaze, W. Ip, T. M. Lietman, et al. 2015. Editor, Evaluating Subcriticality during the Ebola Epidemic in West Africa. *PloS One*10 (10), 1-10. [[PubMed]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0014&dbid=8&doi=10.1080%2F08839514.2020.1747770&key=25775125), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0014&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000363028100058), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2015&issue=10&author=T.+A.+Wayne&author=L.+W.+Enanoria&author=F.+Liu&author=D.+Gao&author=S.+Ackley&author=J.+Scott&author=M.+Deiner&author=E.+Mwebaze&author=W.+Ip&author=T.+M.+Lietman&title=Editor%2C+Evaluating+Subcriticality+during+the+Ebola+Epidemic+in+West+Africa)).

Three algorithms in this effort have been used to validate the final results. All of these algorithms have been developed by the Semeion research center. We used the Topological Weighted Centroid (TWC) algorithm to find the possible location of the outbreak’s origin and analyze the dynamics of disease spread. The TWC suite of algorithms is based on statistical thermodynamics. They have been tested in other cases such as the E-coli outbreak in Germany (Buscema et al., [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., M. Breda, and G. Catzola. 2009. The topological weighted centroid, and the semantic of the physical space – Theory. In *Artificial Adaptive Systems in Medicine*, ed. M. Buscema and E.Grossi, 69–78, Bentham Science Publisher: Bentham. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=69-78&author=M.+Buscema&author=M.+Breda&author=G.+Catzola&title=The+topological+weighted+centroid%2C+and+the+semantic+of+the+physical+space+%E2%80%93+Theory); Buscema et al. [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., E. Grossi, M.Breda, and T. Jefferson. 2009. Outbreaks source: A new mathematical approach to identify their possible location. *Physica A* 388:4736–62. doi:10.1016/j.physa.2009.07.034. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0003&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.1016%2Fj.physa.2009.07.034), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0003&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000270618500007), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=4736-62&author=M.+Buscema&author=E.+Grossi&author=M.+Breda&author=T.+Jefferson&title=Outbreaks+source%3A+A+new+mathematical+approach+to+identify+their+possible+location), [2013](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., E. Grossi, A.Bronstein, W. Lodwick, M. Asadi-Zeydabadi, R. Benzi, and F.Newman. 2013. A new algorithm for identifying possible epidemic sources with application to the German Escherichia Coli outbreak. *ISPRS International Journal Geo- Inf* 2:155–200. doi:10.3390/ijgi2010155. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0004&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.3390%2Fijgi2010155), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0004&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000209465600009), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2013&pages=155-200&author=M.+Buscema&author=E.+Grossi&author=A.+Bronstein&author=W.+Lodwick&author=M.+Asadi-Zeydabadi&author=R.+Benzi&author=F.+Newman&title=A+new+algorithm+for+identifying+possible+epidemic+sources+with+application+to+the+German+Escherichia+Coli+outbreak); Grossi, Buscema, and Jefferson [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Grossi, E., M.Buscema, and T. Jefferson. 2009. The topological weighted centroid, and the semantic of the physical space – Application. In *Artificial Adaptive Systems in Medicine*, ed. M. Buscema and E.Grossi, 79–89, Bentham Science Publisher: Bentham. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=79-89&author=E.+Grossi&author=M.+Buscema&author=T.+Jefferson&title=The+topological+weighted+centroid%2C+and+the+semantic+of+the+physical+space+%E2%80%93+Application)). The details and the new development in TWC algorithm with some applications also have been discussed in (Buscema, Massini, and Sacco [2017](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., G. Massini, and P. Sacco. 2017. The Topological Weighted Centroid (TWC): A topological approach to the time-space structure of epidemic and pseudo-epidemic processes. *Physica A* 492, 582-627. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2017&author=M.+Buscema&author=G.+Massini&author=P.+Sacco&title=The+Topological+Weighted+Centroid+%28TWC%29%3A+A+topological+approach+to+the+time-space+structure+of+epidemic+and+pseudo-epidemic+processes)). The Dynamic Naive Bayesian/Dynamic Networks Block (DNB) Algorithm was used to find the cause and effect of the Ebola outbreak and it is published here for the first time. The Selfie ANN has been applied also for the first time in paper. The main purpose of using different algorithms is to justify the results of using different methods and test them on the Ebola outbreak data. The core of our analysis is based on the TWC method. We have divided the data into eight sets and we have shown the method is stable. We have obtained the same prediction for the origin of the outbreak from all eight sets and the results were near the place that was reported by WHO as the origin of the Ebola outbreak. In summary, we validated our analyses by

* comparing the results from three different algorithms (TWC, DNB, and ANN Selfie),
* using different TWC algorithms (TWCα, TWCβ, and TWCϒ),
* applying TWCα for eight different sets and
* comparing with the results with the origin that was reported by WHO.
* According to information from the WHO, Ebola started in Guinea but spread out mostly to Liberia and Sierra Leone.

The Ebola dataset from the HealthMap included 83 locations (Table 2) from western Africa monitored for a period of 32 weeks (from March 4 to October 4, 2014). The number of new Ebola-infected persons was recorded every week for each city. Notice that the number of confirmed cases in seven locations is zero. We consider any error in the data as noise that can occur with real data.

## Topological Weighted Centroid Algorithm

The Topological Weighted Centroid, whose theory and successful applications have been discussed in detail in our previous publications (Buscema et al., [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., M. Breda, and G. Catzola. 2009. The topological weighted centroid, and the semantic of the physical space – Theory. In *Artificial Adaptive Systems in Medicine*, ed. M.Buscema and E. Grossi, 69–78, Bentham Science Publisher: Bentham. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=69-78&author=M.+Buscema&author=M.+Breda&author=G.+Catzola&title=The+topological+weighted+centroid%2C+and+the+semantic+of+the+physical+space+%E2%80%93+Theory); Buscema et al. [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., E. Grossi, M. Breda, and T.Jefferson. 2009. Outbreaks source: A new mathematical approach to identify their possible location. *Physica A*388:4736–62. doi:10.1016/j.physa.2009.07.034. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0003&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.1016%2Fj.physa.2009.07.034), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0003&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000270618500007), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=4736-62&author=M.+Buscema&author=E.+Grossi&author=M.+Breda&author=T.+Jefferson&title=Outbreaks+source%3A+A+new+mathematical+approach+to+identify+their+possible+location), [2013](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., E. Grossi, A.Bronstein, W. Lodwick, M. Asadi-Zeydabadi, R. Benzi, and F.Newman. 2013. A new algorithm for identifying possible epidemic sources with application to the German Escherichia Coli outbreak. *ISPRS International Journal Geo- Inf* 2:155–200. doi:10.3390/ijgi2010155. [[Crossref]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0004&dbid=16&doi=10.1080%2F08839514.2020.1747770&key=10.3390%2Fijgi2010155), [[Web of Science ®]](https://www.tandfonline.com/servlet/linkout?suffix=CIT0004&dbid=128&doi=10.1080%2F08839514.2020.1747770&key=000209465600009), [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2013&pages=155-200&author=M.+Buscema&author=E.+Grossi&author=A.+Bronstein&author=W.+Lodwick&author=M.+Asadi-Zeydabadi&author=R.+Benzi&author=F.+Newman&title=A+new+algorithm+for+identifying+possible+epidemic+sources+with+application+to+the+German+Escherichia+Coli+outbreak); Buscema, Massini, and Sacco [2017](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Buscema, M., G.Massini, and P. Sacco. 2017. The Topological Weighted Centroid (TWC): A topological approach to the time-space structure of epidemic and pseudo-epidemic processes. *Physica A* 492, 582-627. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2017&author=M.+Buscema&author=G.+Massini&author=P.+Sacco&title=The+Topological+Weighted+Centroid+%28TWC%29%3A+A+topological+approach+to+the+time-space+structure+of+epidemic+and+pseudo-epidemic+processes); Grossi, Buscema, and Jefferson [2009](https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1747770)Grossi, E., M. Buscema, and T. Jefferson. 2009. The topological weighted centroid, and the semantic of the physical space – Application. In *Artificial Adaptive Systems in Medicine*, ed. M.Buscema and E. Grossi, 79–89, Bentham Science Publisher: Bentham. [[Google Scholar]](http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=79-89&author=E.+Grossi&author=M.+Buscema&author=T.+Jefferson&title=The+topological+weighted+centroid%2C+and+the+semantic+of+the+physical+space+%E2%80%93+Application)), was used to predict the origin and to analyze the spread of the recent Ebola outbreak. In this paper, we discuss TWC briefly. The foundation of this artificial neural system algorithm is statistical mechanics. We define the possible states of the system in terms of the locations (longitude and latitude) of the outbreak. For each state, we use the Boltzmann factor to define the probability of the state and then we are able to define the free energy and entropy of the system. Using the free energy and entropy we can then configure the state space giving us the possible source(s). From the source we then predict the possible spread. We define the system similar to a two-dimensional physical system that is a combination of several point masses, charges, or heat sources. The system is a finite two-dimensional discrete space with some active points (the outbreak points similar to the heat sources or charges in a physical system). The effect of the outbreak on all points of the discrete space is given by some scalar fields similar to the electric potential or temperature field in a physical system. Here, in some of our algorithms, only the location of the outbreak points are used and in some in addition to the outbreak points, the coordinates of all points on the discrete space are used. We use the longitude and latitude of the points in the region of interest as the two-dimensional coordinates. Three different TWC algorithms, TWCα, TWCβ, and TWCγ were used. The main differences between them stem from different definitions of the state of the system.

**RESEARCH PAPER 2:**

ARTIFICIAL INTELLIGENCE IN PURCHASING

**SUMMARY:**

### Topic 1 – AI in the Purchasing Process

The first discussion table was concerned with the identification of possible AI solutions in the entire purchasing process. The results are presented in the short list which is displayed in Table 1. The columns “Points assigned by AI experts (in %)” and “Points assigned by PSM experts (in %)” display the relative amount of points assigned to the discussion points by the two expert groups respectively. While the original list included 22 discussion points, the short-list contains four ideas that account for 52% of the points assigned by the experts. For the first discussion table, the following ideas were prioritized:

1. ***Cost-optimizing engineering***: AI might help during the engineering stage or the early phase of the purchasing process to identify changes in the product that do not impair the functionality or the quality of the item under consideration but reduce the purchasing costs, e.g. through lower material costs or the avoidance of over-engineering.
2. ***Analysis of cost-breakdowns***: The cost-breakdowns provided by the suppliers could be evaluated systematically and intelligently to detect hidden cost drivers. Additionally, a database could be created that contains detailed information concerning the cost structure of each supplier.
3. ***Recognition of price patterns***: AI could identify patterns that reflect the development of commodity or material prices over time. As a result, optimal and anticyclical points of time for the sourcing process could be identified.
4. ***Analysis of negotiation behavior:*** AI could be used to analyze the individual negotiation behavior of each supplier to predict how the supplier will behave in the negotiation process.

### Topic 2 – AI in Mechanism Design-based Negotiations

The second discussion table aimed at identifying opportunities for AI to support mechanism design-based negotiations. The results of the discussion and the subsequent ratings are presented in Table 2, which is a short list, consisting of those three ideas (out of 13) that gained the most interest, i.e. 58% of the expert ratings.

1. ***Analysis of negotiation behavior:*** AI could be used to analyze the individual negotiation behavior of each supplier to predict how the supplier will behave in the negotiation process.
2. ***Simulation of negotiations:*** AI could analyze the negotiation design of an upcoming negotiation and make predictions on the expected outcomes as well as validity checks of the suggested negotiation rules.
3. ***Development of negotiation designs:*** AI autonomously develops negotiation designs that fit to the individual circumstances of each sourcing project and its market conditions.

### Topic 3 – Implementation of AI in Mechanism Design-based Negotiations

The third discussion table was concerned with the technical implementation of AI in the process of conducting a mechanism design-based negotiation. The short list (Table 3) consists of those five ideas (out of 15) that account for 52% of the expert ratings.

1. ***Simulation of negotiations:*** AI could analyze the predefined rules of an upcoming negotiation and make predictions of the expected outcomes as well as validity checks of the suggested negotiation design.
2. ***Expert systems:*** Expert systems that aim at imitating human knowledge and behavior could be developed. In the first step, these systems would still require human input and interaction. With an increasing amount of data available, the systems could become more intelligent through machine learning approaches and ultimately make the human input obsolete.
3. ***Information seeking across systems:*** AI could be capable of collecting meaningful data from the various IT systems of large buying organizations and to intelligently merge these data in a way that they facilitate the process of conducting mechanism design-based negotiations.
4. ***Heuristic mechanism design selection:*** If there are too many and partially conflicting goals and targets of a negotiation, the degrees of freedom might be too high in order to be able to develop one single solution. Heuristic selection systems, supported by AI, might compare the expected outcomes of each proposed negotiation design and then choose the most suitable one. This process could be repeated in various rounds and would result in the survival of the fittest design.
5. ***Goal definition:*** A precondition for designing negotiation rules is to define goals and premises a priori. So far, the complexity of the goals that can be taken into consideration is somewhat limited by human cognitive capacities. AI could make it possible to include a larger number of goals in the process of designing negotiation rules by providing the amount of cognitive capacity that is needed to do so.

**RESEARCH PAPER 3:**

ATTITUDE CONTROL: A key factor during the design of low – thrust propulsion for CubeSats.

**SUMMARY:**

Orbital capability is a decisive step forward for nanosatellites in general and CubeSats in particular. Although trajectory maneuvers and their implementation have been thoroughly studied for classical satellites, the high level of constraints on CubeSats in terms of mass, volume and power, makes the transition delicate. Orbit, attitude and power control subsystems available for this format limit too optimistic performance available in literature. To verify this hypothesis, we simulate trajectory maneuvers in Earth orbit with representative CubeSat hardware and software. A low-thrust trajectory solver based on classical orbital elements from the literature is adapted to our context. A home-made attitude control simulation tool is coupled to include both control and perturbative dynamics. Increases in time and propellant consumption of more than 100% are caused by thrust direction errors such as misalignments and attitude control limitations, sometimes leading to mission loss. These results highlight an important increase in complexity for the CubeSat format that is not covered by the usual approach. Such limitations should be considered from the very start of the design of a nanosatellite mission with trajectory modification requirements.

* Usual approach considers GNC and ADCS independently on CubeSats.
* Thrust disturbance torques are unavoidable and saturate reaction wheels (RW).
* RW desaturation either extend maneuver duration or increase propellant consumption.
* Mutual impacts between GNC and ADCS can lead to CubeSat loss.
* Design of CubeSat with propulsion must focus on ADCS before trajectory optimization.

**RESEARCH PAPER 4:**

A 3-Stage Machine Learning-Based Novel Object Grasping Methodology

**SUMMARY:**

The automatic grasping of objects previously unseen by a robotic system is a difficult task-of which there is currently no robust solution. The research presented in this article improves upon previous works that employ depth data and learning techniques to generate and select from a pool of hypothesised grasps by focusing on the pruning and selection process. In this work, a vision-based, sampling methodology that generates candidate grasps through a convolutional neural network is proposed. Each candidate grasp is assessed using scores derived from the candidate itself and other related input modalities-such as the centre of gravity of the object. The final selection is determined by a learning algorithm. To overcome human bias, objective measures of grasp performance are established that comprehensively measure the error introduced by the grasp trial itself. The proposed metrics are empirically demonstrated to quantify grasp quality, offer useful criteria for network training and provide better descriptive power than traditional measures of grasp outcome. Experimentation showed that the proposed methodology can generate a meaningful, final grasp within 1.3 seconds. Trials quantitatively demonstrate a small-object-in-isolation performance of 99%. For unknown objects, this equates to a 10% improvement relative to other similar methodologies. Testing also showed that grasp performance was improved by 5% when implementing the proposed metrics-compared to the baseline.

Autonomous novel object grasping and handling is a wide-ranging, high-impact field with many implications, especially within domestic and industrial application. Some instances where automatic grasping has been studied include an automated checkout robot [1], garbage sorting [2], cloth manipulation [3], bed making [4], dishwasher unloading [5], automated cooking [6], [7], service robotics [8]–[9][10][11], general household-related grasping [12]–[13][14], clutter clearing [15], [16] and stowing, picking and packing for warehouse automation [17]–[18][19][20]—which has gained significant traction since the 2017 Amazon Robotics Challenge [21].

The automated manipulation of objects previously unseen by a robotic system is an extremely difficult task, as a good grasp is related to object shape, size, material, weight-distribution, surface properties, friction coefficients and object deformability, and can be severely affected by sensing and actuation accuracy. Moreover, the relationship between these variables and a specific grasping strategy, robotic hardware and a gripper is not always clear.

Research in this field has been active for decades, yielding a colourful range of promising avenues—especially with the recent interest from well-known and well-resourced research institutions, such as Google and the Massachusetts Institute of Technology (MIT). With the attention of such institutions, we have seen unprecedented, large-scale dataset generation frameworks that allow for the training of complex, self-supervised neural networks [22]–[23][24][25]. Moreover, machine learning in general has become overwhelmingly represented within this area, with a great deal of work utilising RGB and/or depth input modalities [12], [26], [27]. Despite the success seen throughout literature, the automated manipulation of novel objects remains challenging and an active topic of research.

Over the past three decades consumers have increasingly been demanding a wider variety of goods in smaller batches, resulting in rapid changes in production technologies [28]–[29][30]. This trend reveals the importance for flexible, reconfigurable and automated production systems for future markets [31]–[32][33]—which is not usually considered by object manipulation literature. Although flexible robotic hardware is progressively becoming a popular topic, such as the dual-arm, scalable concept developed by ABB [34], [35], the adaptability of the related grasping methodology is not usually considered to the same degree. Fully manual assembly lines are still common in low-wage countries, particularly for manufacturers of consumer electronics, small appliances, toys, etc. Novel object manipulation methodologies that utilise object identification, accurate grasping location/orientation and a robust handling process play a crucial role in future automation and production lines. This paper presents a grasping methodology based on machine learning that aims to improve on research related to novel object detection, grasping and handling with autonomous robotic systems—further closing the gap between manual and fully automated production.

Automated grasping is typically posed as a search problem—find the location that will best facilitate handling of the object from a potential infinite number of grasps. The goal of our research is to sample some of these candidate locations and select a meaningful subset of grasp hypotheses, which may then be pruned based on quality metrics to select a suitable and reliable grasp for execution. In contrast to many current systems that utilise depth information [18], [20], [27], [36], 3D models of objects [37]–[38][39] or wrist-mounted sensors [12], [26], [40], our methodology operates on raw, monocular RGB observations of the scene.

This paper presents an approach that utilises machine learning and part-related information to find, grade and select suitable robotic grasping locations in 3D space and is an extension of our previous work [41]. In this paper, a selection-stage is established. The datasets used for training have also been considerably improved. Moreover, the proposed methodology is implemented on a physical robot and trials are conducted for validation. The method consists of three main stages, each coupled with a learning component. First, a classifier is trained to determine whether the object within a region of interest is known or unknown. For unknown objects, a small convolutional neural network (CNN) quickly classifies segments of the object through vision to identify potential grasping locations. A scoring network is then used to rank these locations and decide on the final grasping position. Input features for the scoring network are derived from the assessed grasp itself and other features related to the object or grasp location. Methodology performance is quantified as per literature and the proposed set of metrics. To evaluate the proposed methodology, experimentation and testing focus specifically on 2-fingered, parallel jaw gripping that uses force-closure within the scope of object-agnostic grasping, approached from an industrial perspective.

The performance evaluation of our system through physical trials demonstrates quantitatively that our approach can grasp small objects not seen during training 98.9% of the time—despite relying only on rudimentary sensing, such as an RGB webcam. Compared to the relevant literature, this constitutes an improvement of roughly 10%. An illumination-controlled imaging chamber and conveyor system was constructed for dataset generation and methodology testing. Objects are placed haphazardly at one end of the conveyor and grasped at the other end. The final system generated a grasp within 1.3 seconds, producing on average 83 viable grasps per object. New quantitative metrics that more accurately reflect the quality of a grasp have also been proposed. Trials revealed that the proposed metrics are capable of further improving grasp rates by 2.7% for unknown objects and 5.3% for known objects—compared to highest confidence selection.

This paper is organised as follows. Section 2 introduces current approaches to novel object grasping and describes the associated difficulties therein. Section 3 states the proposed metrics used to define and improve grasping performance. Sections 4 to 7 cover the proposed methodology, experimentation, analysis and future research. Finally, section 8presents some of the conclusions from this research.

**RESEARCH PAPER 5:**

## Mixing quantitative and qualitative methods for sustainable transportation in smart cities

**SUMMARY:**

In particular, we focus on the last mile segment of the supply chain, which is a complex system due to several critical factors (e.g., large-scale problems generated by a huge number of delivery points, uncertainty, multiple actors, contraction in the timing for decisions). These issues bring new challenges and complexities in urban transportation that becomes a system of systems, incorporating current structures and new and future business models (e.g., new delivery options and low-emission transportation modes). This makes necessary a holistic approach that looks at the system according to a global vision and considers the sustainability from economic, operational, environmental and social standpoints.

The main contribution of this Ph.D. research is a multi-disciplinary approach that is a mixture of qualitative and quantitative techniques by the research communities in Operation Research, Management, and Computer Science, within the City Logistics framework. To our knowledge, this type of approach is new in the literature. Thus, to illustrate the effectiveness of the proposed approach in supporting the decision-making processes, it is applied to two new planning problems in urban freight transportation. The first application concerns the integration of traditional transportation modes (i.e., vans) and vehicles with a low-environmental impact (i.e., cargo bikes). This study integrates a managerial analysis of these current business models describing the stakeholders’ profiles in terms of their needs and costs and revenues structures. Then, the integration of business and operational models is supported by a performance analysis of the traditional and green delivery options, based on the main variables that affect the last-mile logistics in urban areas (e.g., distance, delivery time). Finally, a quantitative analysis of the system is guaranteed by a Monte Carlo simulation, to extrapolate mixed-fleet policies (see G. Perboli and M. Rosano, “Parcel delivery in urban areas: Opportunities and threats for the mix of traditional and green business models”, *Transportation Research Part C: Emerging Technologies*, 99, 19–36, 2019).

We extend this analysis including new delivery options as the automated pick-up and delivery points “lockers”, which reflect the current practice in the market. To deal with this issue, we propose a new standard simulation-optimization framework for building instances and assess operational settings. According to the multi-disciplinary approach presented in this Ph.D. research, the framework generalizes many types of routing problems and allows to combine requirements from various stakeholders, and data gathered from different sources (e.g., behavioral and socio-economic data, city network, operational data). Generating new realistic instance sets, it mitigates the issues of current City Logistics solutions that are based on artificial data, making their exploratory capacity and transferability to the industry limited. To illustrate the usefulness of that framework, we apply it to address the dynamic and stochastic VRP with time windows problem under the context of the online urban freight distribution in the City of Turin (Italy).

The integration of different transportation modes and outsourcing practices necessitate complex negotiation of contracts between shipper and carrier. In particular, the second application concerns the tactical capacity planning problem that considers a shipper who seeks to secure transportation and warehousing capacity from a carrier. The medium-term nature of contracts requires to deal with the uncertainty, that in this thesis is expressed in terms of the demand of loads to be transported, the cost and availability of additional capacity if required, and the availability of contracted capacity, which is a novelty in the literature. To deal with this problem, we formulate a stochastic two-stage model and propose a meta-heuristic based on the Progressive-Hedging algorithm, presented in T.G. Crainic, G. Perboli and M. Rosano, “Simulation of intermodal freight transportation systems: a taxonomy”, *European Journal of Operational Research*, 270(2), 401–418, 2018. An extensive computational campaign concludes the analysis evaluating the impact of considering uncertainty in the first and last extremities of the supply chain and providing managerial insights.

Our results show that the integration of traditional and new business models is not a free lunch. Although the bikes and lockers can generate environmental gains, with more than 40% of CO2CO2 savings, their improper adoption reduces the operational efficiency of the traditional models, up to 80% in the worst case. In conclusion, to avoid this, the provided multi-disciplinary approach and managerial insights can support the decision-makers in the continuous optimization process and in extrapolating, and deploying effective industrial and public mixed-fleet policies for sustainable transportation in urban areas.

**PART - C**

**CERTIFICATE OF MOOC - 1**



**CERTIFICATE OF MOOC - II**

