

BREAST CANCER PREDICTION USING MACHINE LEARNING

**Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc.Degree
programme in Statistics**

by

AISWARYA P S

Register No.CCAWMST001



**Department of Statistics
Christ College (Autonomous)
Irinjalakuda**

2024

CERTIFICATE

This is to certify that the project entitled 'BREAST CANCER PREDICTION USING MACHINE LEARNING', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by AISWARYA P S (CCAWMST001) during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.



Ms. Mary Priya

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda



Dr. Davis Antony M

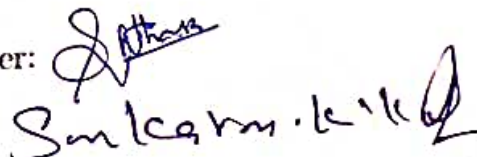
Head, Department of Statistics

(Self Financing)

Christ College (Autonomous)

Irinjalakuda

External Examiner:



Place: Irinjalakuda

Date: 29-06-2024



DECLARATION

I hereby declare that the matter embodied in the project entitled 'BREAST CANCER PREDICTION USING MACHINE LEARNING', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Ms. Mary Priya , Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-2024

Aiswarya

AISWARYA P S

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and help of several individuals who in various ways contributed and extended their valuable assistance during the preparation and completion of the study.

I would like to express my sincere gratitude to Ms. Mary Priya , Assistant Professor, Department of Statistics, Christ College (Autonomous), Irinjalakuda. Her invaluable guidance and encouragement have been instrumental in the successful completion of this project.

I would like to express my deep appreciation to Dr. Davis Antony M, Head of Department of Statistics, and my teachers for their continuous support and expert guidance.

I sincerely thank the librarian and non-teaching staff of Christ College (Autonomous), Irinjalakuda. I am very thankful to my classmates for their support and help.

I am deeply grateful to my parents for their unconditional love and support. Above all, I thank God for blessing me and helping me complete this work.

Irinjalakuda

Date: 29-06-2024

AISWARYA P S

Contents

1	Introduction	8
1.1	Risk Factors For Breast Cancer	9
1.2	Software Used	12
1.3	Objectives	12
2	Methodology	13
2.1	Exploratory Data Analysis	13
2.2	Logistic Regression	15
2.2.1	Types Of Logistic Regression	16
2.2.2	Assumptions Of Logistic Regression	16
2.3	Decision Tree Classifier	18
2.3.1	Types Of Decision Tree Classifier	18

2.3.2	Assumptions Of Decision Tree Classifier	20
2.4	Random Forest Classifier	22
2.4.1	Types Of Random Forest Classifier	22
2.4.2	Assumptions Of Random Forest Classifier	23
2.5	Confusion Matrix	25
2.5.1	True Positives (TP)	25
2.5.2	True Negatives (TN)	25
2.5.3	False Positives (FP)	26
2.5.4	False Negatives (FN)	26
2.5.5	Accuracy	26
2.5.6	Precision	27
2.5.7	Recall	27
2.5.8	F1 Score	27
3	Data Analysis	29
3.1	Data	29
3.2	Handling Missing Values	31

3.3	Visualization Of Output Variable	32
3.4	Correlation matrix	33
3.5	Splitting Dataset	34
3.6	Logistic Regression Model	35
3.7	Decision Tree Classification Model	36
3.8	Random Forest Classification Model	37
4	Conclusion	38
	References	39

List of Figures

3.1	Null Value	31
3.2	Malignant vs Benign Count Plot	32
3.3	Correlation Heatmap	33
3.4	Logistic Regression Model	35
3.5	Confusion Matrix-Logistic	35
3.6	Decision Tree Classification Model	36
3.7	Confusion Matrix- Decision Tree	36
3.8	Randsom Forest Classification Model	37
3.9	Confusion Matrix- Random Forest	37

A STUDY ON INDIAN AND GLOBAL STOCK MARKETS

**Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics**

by

ALWIN P P

Register No.CCAWMST002



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled 'A STUDY ON INDIAN AND GLOBAL STOCK MARKETS ', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **ALWIN P P (CCAWMST002)** during the period of his study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.

Linett George

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

Dr. Davis Antony M

Head , Department of Statistics

(Self Financing)

Christ College (Autonomous)

Irinjalakuda

External Examiner:

Place: Irinjalakuda

Date: 29/06/2024



DECLARATION

I hereby declare that the matter embodied in the project entitled ' A STUDY ON INDIAN AND GLOBAL STOCK MARKETS ', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Linett George, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29/06/2024



ALWIN P P

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Linett George , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date: 29/06/2024



ALWIN P P

Contents

1	Introduction	8
2	Methodology	13
2.1	Artificial Neural Network	13
2.1.1	Multi-layer Perceptron(MLP) Neural Network	15
2.2	Activation Function	17
2.3	Loss Function	17
2.4	Forward Propagation	18
2.5	Back Propagation	19
2.6	Hyperparameter Tuning	19
2.7	Pearson Correlation Coefficient	20
2.8	Scatter Plot	21

3 Data Analysis	22
3.1 Time Series Plots	22
3.2 Scatter Plots	26
3.3 Pearson Correlation Coefficient	29
3.4 Forecasting Using Multilayer Neural Network	32
3.4.1 Forecasting Of Sensex	32
3.4.2 Forecasting Of Nifty-50	35
3.4.3 Forecasting Of NYSE Composite Index	37
4 Conclusion	39
References	41

List of Figures

2.1	Single Layer Artificial Neural Network	14
2.2	Multi Layer Neural Network	16
3.1	Time Series Plot of Sensex	23
3.2	Time Series Plot of Nifty-50	24
3.3	Time Series Plot of NYSE Composite Index	25
3.4	Scatter Plot of Indian Stock Market Indices	26
3.5	Scatter Plot between Nifty-50 and NYSE Composite Indices .	27
3.6	Scatter Plot between Sensex and NYSE Composite Indices . .	28
3.7	Correlation Heatmap	30
3.8	Forecasted Closing Price of Sensex	33
3.9	Forecasted Closing Price of Nifty-50	35

REVIEW ON FRECHET DISTRIBUTION AND IT'S GENERALISATIONS

Dissertation report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

ARIFA A S

Register No.CCAWMST003



Department of Statistics

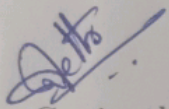
Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled 'REVIEW ON FRECHET DISTRIBUTION AND IT'S GENERALISATIONS', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **ARIFA A S(CCAWMST003)** during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.



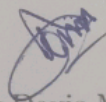
Geethu Gopinath

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda



Dr.Davis Mundassery

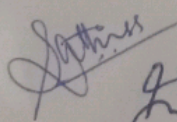
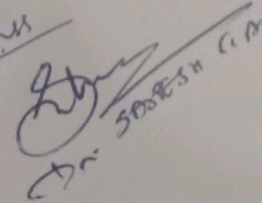
Head,Department of Statistics

(Self Financing)

Christ College (Autonomous)

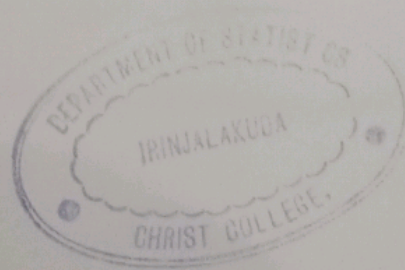
Irinjalakuda

External Examiner:



Dr. SURESH C. R.

Place: Irinjalakuda

Date: 29-06-2024



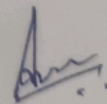
DECLARATION

I hereby declare that the matter embodied in the project entitled ' REVIEW ON FRECHET DISTRIBUTION AND IT'S GENERALISATIONS', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Dr. Davis Antony Mundassery, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2021-2023.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-2024



ARIFA A S

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Dr. Davis Antony Mundassery , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

ARIFA A S

Date: 29-06-2024

Contents

1	INTRODUCTION	8
1.1	Background and Motivation	8
1.2	Objective	9
2	EXTREME VALUE DISTRIBUTIONS(EVDs)	10
2.1	Types of Extreme Value Distributions	12
2.1.1	Gumbel Distribution (Type I)	12
2.1.2	Weibull Distribution (Type III)	13
2.1.3	Fréchet Distribution (Type II)	15
3	ODD FRECHET G FAMILY DISTRIBUTIONS	19
3.1	Introduction	19

3.1.1	Odds Ratio and Formation of the Odd Fréchet-G Family	20
3.1.2	PDF, CDF, and Hazard Function	20
3.2	Moments and Moment Generating Function of the Odd Fréchet Distribution	22
3.2.1	Moments of the Odd Fréchet Distribution	22
3.2.2	Moment Generating Function (mgf) of the Odd Fréchet Distribution	22
3.2.3	Order statistics	23
3.3	Special Cases of OFr-G Distribution	25
3.3.1	Odd Frechet-Weibull (OFr-W) Distribution	25
3.3.2	Odd Frechet-Lomax (OFr-L) Distribution	28
3.3.3	Odd Frechet-Gamma (OFr-Gam) Distribution	30
3.3.4	Odd Frechet-Nadarajah Haghighi(OF-NH)	33
4	EXTENDED ODD FRECHET G FAMILY DISTRIBUTIONS	36
4.1	Introduction	36
4.1.1	Cumulative Distribution Function (CDF)	37
4.1.2	Probability Density Function (PDF)	38

4.2	Properties of the Extended Odd Fréchet-G Family of Distributions	40
4.2.1	Moments	40
4.2.2	Moment Generating Function (MGF)	40
4.2.3	Order Statistics	41
4.2.4	Quantile Function	41
4.2.5	Generating Function	41
4.2.6	Order Statistics	42
4.3	Special Distributions of EOF-G family	43
4.3.1	EOF-Nadarajah-Haghighi (EOFNH) Distribution.	43
4.3.2	EOF-Weibull (EOFW) Distribution.	44
4.4	simulation and fitting of data	45
5	CONCLUSION	48
6	REFERENCE	50
7	APPENDIX	52

List of Figures

3.1	PDF of Weibull,OF-W	27
3.2	PDF of Gamma,OFr-Gamma	32
3.3	PDF of NH,OFr-NH	35
4.1	PDF of EOF-NH,EOF-W	45

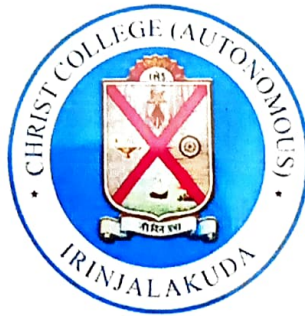
A STUDY ON E VALUE AND ITS SIGNIFICANCE

Dissertation report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

MALAVIKA P

Register No.CCAWMST005



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the dissertation entitled '**A STUDY ON E VALUE AND ITS SIGNIFICANCE**', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **MALAVIKA P(CCAWMST005)** during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.

Mcm

Megha C M

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

Dr. Davis Antony M

Dr.Davis Antony M

Head

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

External Examiner:

Dr. Sathya P. B

Place: **IRINJALAKUDA**

Date: **29-06-2024**

DECLARATION

I hereby declare that the matter embodied in the dissertation entitled ' A STUDY ON E VALUE AND ITS SIGNIFICANCE ', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Megha C M, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this dissertation has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-2024



MALAVIKA P

ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Megha C M , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date:

MALAVIKA P

Contents

- 1 Introduction** **6**

- 2 A Prelude to E values** **9**
 - 2.1 Definition of E values 10
 - 2.2 Difference between e-value and p-value 13
 - 2.3 Merging of e values and their application in multiple testing
of hypothesis 15
 - 2.3.1 Advantage of applying merged e values in multiple hy-
pothesis testing 18

- 3 E TESTING** **21**
 - 3.1 General principles of e testing 22
 - 3.2 Non parametric e testing 24

3.2.1	Fisher type non parametric e test	24
3.3	Sign e test	26
3.3.1	Sign e test	26
4	FITTING OF SIGN TEST AND SIGN E TEST ON CHER-	
	NOBYL DATA	29
4.0.1	Fitting of sign test to Chernobyl data	29
4.0.2	Fitting of sign e test to Chernobyl data	30
5	CONCLUSION	31
6	References	37

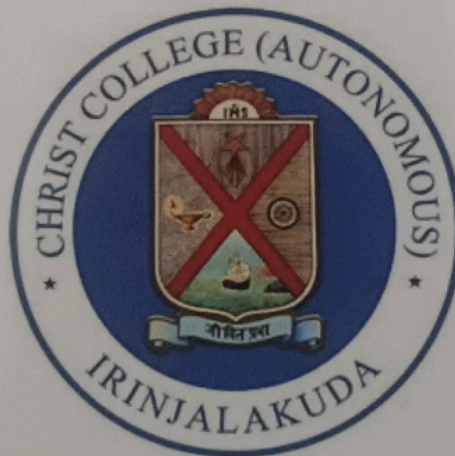
BIRCH CLUSTERING

Dissertation report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of the M.Sc Degree programme in
Statistics

by

NAMRUTHA T U

Register No.CCAWMST006



Department of Statistics
Christ College (Autonomous)
Irinjalakuda
2024

CERTIFICATE

This is to certify that the project entitled "BIRCH CLUSTERING " submitted to Department of Statistics in partial fulfilment of the requirement for the award of the M.Sc Degree programme in Statistics, is a bonafide record of original research work done by **NAMRUTHA T U (CCAWMST006)** during the period of her study in the Department of Statistics, Christ College (Autonomous), Irinjalakuda, under my supervision and guidance during the year 2022-2024

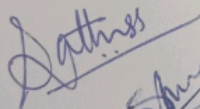
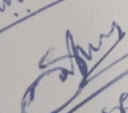
Mem
Mrs.MEGHA C M
Assistant Professor
Department of statistics
Christ College (Autonomous)
Irinjalakuda

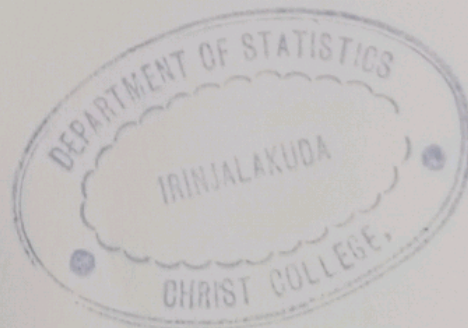

Dr.Davis Antony M
Head,
Department of Statistics(Self Financing)
Christ College (Autonomous)
Irinjalakuda

External Examiner :

Place : Irinjalakuda

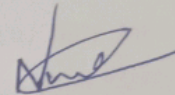
Date : 29/06/2024



Dr. S. D. K. S. R. C. S.



DECLARATION

I hereby declare that the project work entitled " **BIRCH CLUSTERING**" submitted to Department of Statistics, Christ College(Autonomous), Irinjalakuda in partial fulfillment of the requirement for the award of Master Degree of Science in Statistics is a record of original project work done by me during the period of my study in the Department of Statistics, Christ College(Autonomous), Irinjalakuda and under the supervision of Megha C M, Assistant Professor, Department of Statistics, Christ College(autonomous), Irinjalakuda



NAMRUTHA T U

Place : Irinjalakuda

Date : 29/06/2024

ACKNOWLEDGEMENT

First, there are no words to adequately acknowledge the wonderful grace that my Redeemer has given me. I greatly appreciate the inspiration; support and guidance of all those people who have been instrumental for making this project a success.

I express my deepest gratitude to my guide Mrs.Megha C M, Assistant Professor, Department of Statistics, Christ College(Autonomous), Irinjalakuda, who guided me faithfully through this entire project. Without her advice, support and guidance, it find difficult to complete this work. I take this opportunity to express my thanks to our beloved principal Fr. Dr. Jolly Andrews CMI, who gave me the golden opportunity to do this wonderful project on the topic "**BIRCH CLUSTERING**".

I mark my word of gratitude to Dr.Davis Mundassery, Head of the Department and all other teachers of the department for providing me the necessary facilities to complete this project on time.

I want to especially thank all the faculty of the library for providing various facilities for this project. Words cannot express the love and support I have received from my parents, whose encouragement has buoyed me up from the beginning till the end of this work.

NAMRUTHA T U

Contents

Introduction	5
1 PRELIMINARIES	9
1.1 Introduction	9
1.2 Clustering	9
1.2.1 The primary uses of clustering algorithms . . .	11
1.2.2 Dendrogram	13
2 CLUSTERING ALGORITHMS	16
2.1 K-Means Clustering	16
2.1.1 Working Of K Means Clustering	18
2.1.2 Limitations	19
2.2 Hierarchical Clustering	19

2.2.1	Hierarchical Agglomerative Clustering	20
2.2.2	Hierarchical Divisive clustering	21
2.2.3	Limitations	22
2.3	DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	22
2.3.1	Steps used in DBSCAN Algorithm	22
2.3.2	Limitations	23
2.4	Mean Shift	23
2.4.1	The process of mean-shift clustering algorithm can be summarized as follows	24
2.4.2	Limitations	25
2.5	Gaussian Mixture Models (GMM)	26
2.5.1	Limitations	26
2.5.2	The Clustering Algorithms challenges	27

3 BIRCH CLUSTERING 30

3.0.1	Key Concepts	32
3.0.2	Advantages	34

3.0.3	Limitations	36
3.0.4	Real-World Applications	38
4	METHODOLOGY	41
4.0.1	Steps in BIRCH Clustering	41
4.0.2	Basic Algorithm	42
4.0.3	Parameters of BIRCH Algorithm	44
4.0.4	Clustering Metrics in Machine Learning	45
5	PRACTICAL APPLICATION OF BIRCH CLUSTER- ING USING PYTHON	46
5.1	Generating data and applying birch clustering	47
5.2	Scatter Plot	54
6	REFERENCES	55

List of Figures

5.1	Scatter Plot of BIRCH CLUSTERING	54
-----	--	----

A STUDY ON MARKOV CHAIN MONTE CARLO METHODS

**Dissertation report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of the M.Sc Degree programme in
Statistics**

by

PARVATHY GOPALAKRISHNAN

Register No.CCAWMST007



**Department of Statistics
Christ College (Autonomous)**

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled "**A STUDY ON MARKOV CHAIN MONTE CARLO METHODS**" submitted to Department of Statistics in partial fulfilment of the requirement for the award of the M.Sc Degree programme in Statistics, is a bonafide record of original research work done by **Ms.PARVATHY GOPALAKRISHNAN (CCAWMST007)** during the period of her study in the Department of Statistics, Christ College (Autonomous), Irinjalakuda, under my supervision and guidance during the year 2022-2024

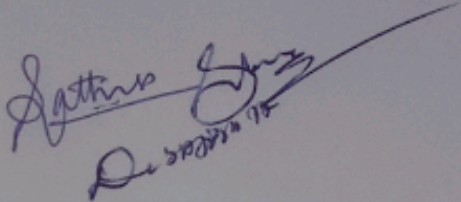

Mrs. Marypriva
Assistant Professor

Department of statistics
Christ College (Autonomous)
Irinjalakuda


Dr. Davis Antony M

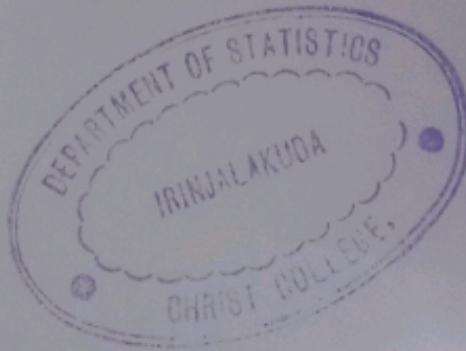
Head,
Department of statistics
Christ College (Autonomous)
Irinjalakuda

External Examiner :


Dr. Sathya

Place : Irinjalakuda

Date : 29-06-2024

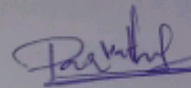


DECLARATION

I hereby declare that the project work entitled "A STUDY ON MARKOV CHAIN MONTE CARLO METHODS" submitted to the Department of Statistics, Christ College(Autonomous), Irinjalakuda in partial fulfilment of the requirement for the award of Master Degree of Science in Statistics is a record of original project work done by me during the period of my study in the Department of Statistics, Christ College(Autonomous), Irinjalakuda and under the supervision of Marypriya, Assistant Professor, Department of Statistics, Christ College(Autonomous), Irinjalakuda.

Place : Irinjalakuda

Date : 29-06-2024



PARVATHY GOPALAKRISHNAN

AKNOWLEDGEMENT

First, there are no words to adequately acknowledge the wonderful grace that my Redeemer has given me. I greatly appreciate the inspiration; support and guidance of all those people who have been instrumental for making this project a success.

I express my deepest gratitude to my guide Mrs. Marypriya, Assistant Professor, Department of Statistics, Christ College(Autonomous), Irinjalakuda, who guided me faithfully through this entire project. Without her advice, support and guidance, it find difficult to complete this work. I take this opportunity to express my thanks to our beloved principal Fr. Dr. Jolly Andrews CMI, who gave me the golden opportunity to do this wonderful project on the topic "**A STUDY ON MARKOV CHAIN MONTE CARLO METHODS**".

I mark my word of gratitude to Dr. Davis Antony M, Head of the Department and all other teachers of the department for providing me the necessary facilities to complete this project on time.

I want to especially thank all the faculty of the library for providing various facilities for this project. Words cannot express the love and support I have received from my parents, whose encouragement has buoyed me up from the beginning till the end of this work.

PARVATHY GOPALAKRISHNAN

Contents

Introduction	5
1 PRELIMINARIES	8
1.1 Markov chain	8
1.1.1 Definition	8
1.1.2 Application	9
1.2 Monte Carlo Method	10
1.2.1 Monte Carlo technique	11
1.2.2 Monte Carlo approximation	12
1.2.3 Monte Carlo integration	12
1.2.4 Algorithm	13
1.2.5 Advantages and Disadvantages	14
1.2.6 Application	15
2 MARKOV CHAIN MONTE CARLO METHOD	17

2.1	MCMC method	18
2.1.1	Working of MCMC method	20
2.2	Diagnostics Of MCMC Method	23
2.2.1	Principle behind MCMC diagnostics	23
2.2.2	Diagnostics techniques	24
2.3	Advantages and Disadvantages	27
2.4	Applications	28
2.5	Types of MCMC method	29
3	METROPOLIS - HASTING ALGORITHM	30
3.1	MH algorithm	31
3.1.1	Practical implementation	32
3.2	Advantages and Disadvantages	33
3.3	Applications	36
3.4	Variations of MH algorithm	36
4	GIBBS SAMPLING	38
4.1	Gibbs sampling	39
4.1.1	Working of Gibbs sampling	40
4.2	Advantages and Disadvantages	40
4.3	Applications:	43
4.4	Extensions	44

5	A SMALL GLANCE TO OTHER MCMC METHODS	46
	ODS	46
5.1	Hamiltonian Monte Carlo (HMC)	46
5.2	Slice Sampling	47
5.3	Gaussian Process MCMC (GP-MCMC)	47
6	PRACTICAL IMPLEMENTATION OF MH ALGORITHM AND GIBBS SAMPLING USING PYTHON	49
6.1	Sampling from bivariate normal distribution using Metropolis-Hasting algorithm	50
6.2	Sampling from bivariate normal distribution using Gibbs sampling algorithm	52
6.3	Scatter plots and Trace plots of MH and Gibbs algorithms	54
6.3.1	Scatter Plots	54
6.3.2	Trace Plots	56
	References	58

List of Figures

6.1	Scatter plot of MH algorithm	54
6.2	Scatter plot of Gibbs sampling	55
6.3	Trace plot of MH algorithm	56
6.4	Trace plot of Gibbs sampling	57

GOLD PRICE PREDICTION USING MACHINE LEARNING ALGORITHMS

**Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree**

programme in Statistics

by

PRANAV A P

Register No.CCAWMST008



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled '**GOLD PRICE PREDICTION USING MACHINE LEARNING ALGORITHMS**', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **PRANAV A P (CCAWMST008)** during the period of his study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.

Ms. Linett George

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

Dr. Davis Antony M

Head, Department of Statistics

(Self Financing)

Christ College (Autonomous)

Irinjalakuda

External Examiner:

Place: Irinjalakuda

Date: 29-06-24



DECLARATION

I hereby declare that the matter embodied in the project entitled ' GOLD PRICE PREDICTION USING MACHINE LEARNING ALGORITHMS ', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Linett George, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-24


PRANAV A P

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Linett George , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date:

PRANAV A P

Contents

1	Introduction	8
2	Methodology	12
2.1	Random Forest Regression	12
2.1.1	Data Preparation	13
2.1.2	Bootstrap Sampling	14
2.1.3	Tree Construction	15
2.1.4	Aggregating Predictions	16
2.1.5	Evaluation Metrics	17
2.2	Gradient Boosted Regression Trees (GBRT)	19
2.2.1	Data Preparation	19
2.2.2	Model Training	20

2.2.3	Boosting Process and Mathematical Foundations . . .	22
2.2.4	Evaluation Metrics	24
2.2.5	Feature Importance	25
2.3	K-Nearest Neighbors (KNN)	26
2.3.1	Data Preparation	26
2.3.2	Algorithm Mechanics	28
2.3.3	Prediction Making	29
2.3.4	Hyperparameter Tuning	30
2.3.5	Evaluation Metrics	31
3	Data Analysis	32
3.1	Dataset	32
3.2	Correlation Heatmap	34
3.3	Performance Matrices	36
3.4	Plot of Actual values vs Predicted Values	37
3.5	Plot of Residuals vs Predicted Values	39
4	Conclusion	41

References

List of Figures

3.1	Correlation Heatmap of the dataset	34
3.2	Plot of Actual values vs Predicted values	37
3.3	Plot of Actual values vs Predicted values	39

A STUDY ON NIFTY50 INDEX OF NATIONAL STOCK EXCHANGE

Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

RIYAN MOHAMMAD

Register No.CCAWMST009



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

A STUDY ON NIFTY50 INDEX OF NATIONAL STOCK EXCHANGE

Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

RIYAN MOHAMMAD

Register No.CCAWMST009



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled '**A STUDY ON NIFTY50 INDEX OF NATIONAL STOCK EXCHANGE**', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **RIYAN MOHAMMAD(CCAWMST009)** during the period of his study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.



Geethu Gopinath

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda



Dr. Davis Antony M

Head , Department of Statistics

(Self Financing)

Christ College (Autonomous)

Irinjalakuda

External Examiner:




Place: **IRINJALAKUDA**

Date: **29-06-2024**



DECLARATION

I hereby declare that the matter embodied in the project entitled ' A STUDY ON NIFTY50 INDEX OF NATIONAL STOCK EXCHANGE', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Geethu Gopinath, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-2024



RIYAN MOHAMMAD

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Geethu Gopinath , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date: 29-06-2024

RIYAN MOHAMMAD

Contents

1	INTRODUCTION	8
2	METHODOLOGY	10
2.1	Time Series Analysis	10
2.2	Stationary Process	11
2.3	Autoregressive Process	12
2.4	Moving Average Process	12
2.5	Auto Regressive Integrated Moving Average (ARIMA) Model	13
2.6	Autocorrelation Function (ACF)	15
2.7	Partial Autocorrelation Function (PACF)	16
2.8	Augmented Dickey-Fuller Test	17
2.9	Akaike Information Criterion (AIC)	17

2.9.1	Interpretation	18
2.10	Bayesian Information Criterion (BIC)	18
2.11	ARIMA Model Fitting	20
2.12	Forecasting	22
2.13	Correlation Analysis	22
2.14	Linear Regression Analysis	23
3	DATA ANALYSIS	24
3.1	Time series plot	24
3.2	Decomposition of Time Series	25
3.3	ACF and PACF plots	26
3.4	Test for stationarity	28
3.5	Model building	31
3.6	Residual Analysis	32
3.7	Forecasting	34
3.7.1	Forecasting table	34
3.8	Correlation analysis	36

3.9 Linear regression analysis	37
4 CONCLUSION	39
REFERENCE	41

List of Figures

3.1	Time Series Plot of Nifty50	25
3.2	Decomposition of Nifty50	26
3.3	ACF plot	27
3.4	PACF plot	27
3.5	Differencing plot	29
3.6	Differenced PACF plot	30
3.7	Residual analysis plot	32
3.8	Forecast	34
3.9	Scatter plot	37

THYROID DISORDER CLASSIFICATION IN WOMEN WITH
ABNORMAL UTERINE BLEEDING

Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

SAMIA A.A

Register No.CCAWMST010

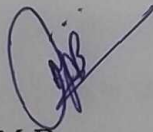


Department of Statistics
Christ College (Autonomous)
Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled '**THYROID DISORDER CLASSIFICATION IN WOMEN WITH ABNORMAL UTERINE BLEEDING**', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **SAMIA A.A(CCAWMST010)** during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.



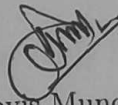
Jiji M.B

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda



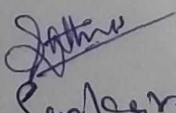
Dr.Davis Mundassery

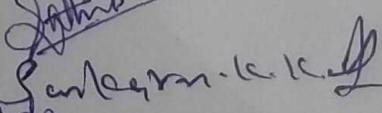
Head,Department of statistics

(Self Financing)

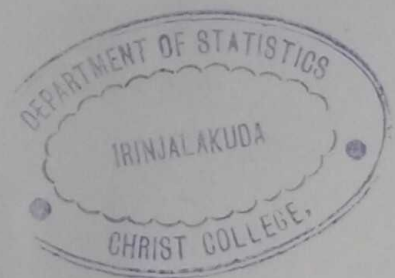
Christ College (Autonomous)

Irinjalakuda

External Examiner: 

Place: Irinjalakuda 

Date: 29-06-2024



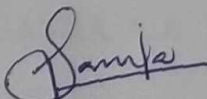
DECLARATION

I hereby declare that the matter embodied in the project entitled. 'THYROID DISORDER CLASSIFICATION IN WOMEN WITH ABNORMAL UTERINE BLEEDING', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Jiji M.B, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2021-2023.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29-06-2024



SAMIA A.A

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Jiji M.B , Assistant Professor, Department of Statistics, whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date: 29-06-2024

SAMIA A.A

Contents

1	INTRODUCTION	8
1.1	Introduction	8
1.2	Problem Statement	11
1.3	Objectives	11
1.4	Statistical tool used	12
1.4.1	Machine learning	12
1.5	Software used	13
1.5.1	Python	13
2	METHODOLOGY	14
2.1	Data Collection and Understanding	14
2.2	Exploratory Data Analysis	15

2.3	Classification	17
2.4	Data preprocessing for classification	18
2.4.1	Train-test splitting:	18
2.5	Model building for classification	19
2.5.1	Decision Tree classifier	19
2.5.2	Random Forest classifier	21
2.5.3	Naive Bayes Classifier	23
2.5.4	XGBoost(eXtreme Gradient Boosting) classifier	24
2.6	Comparison of the accuracy of the models for classification	26
2.6.1	Confusion matrix	26
2.6.2	Accuracy	27
2.6.3	Precision	28
2.6.4	Recall	28
2.6.5	F1-Score	29
2.7	Prediction using test data	29

3 DATA ANALYSIS 30

3.1	Dataset	30
3.2	Data cleaning	32
3.3	Exploratory Data Analysis	34
3.3.1	Visualization of distribution	34
3.4	Decision Tree Classifier	38
3.5	Random Forest Classifier	38
3.6	Naive Bayes Classifier	39
3.7	XG Boost Classifier	39
3.8	Comparison of various performance metrics	40
4	CONCLUSION	44
5	REFERENCES	46
6	APPENDICES	48

List of Figures

3.1	Dataset	30
3.2	Null values	32
3.3	duplicates	33
3.4	Description of data	33
3.5	Pie chart	34
3.6	Stacked barchart	35
3.7	Stacked barchart	35
3.8	Stacked barchart	36
3.9	Stacked barchart	36
3.10	Correlation Heatmap	37
3.11	User interface	43

A STUDY ON KUMARASWAMY DISTRIBUTION

**Dissertation report submitted to Christ College (Autonomous) in
partial fulfilment of the requirement for the award of the M.Sc
Degree programme in Statistics**

by

SELIN STELVIA RODRIGUES

Register No.CCAWMST011



**Department of Statistics
Christ College (Autonomous)**

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled "**A STUDY ON KUMARASWAMY DISTRIBUTION**" submitted to the Department of Statistics in partial fulfillment of the requirement for the award of the M.Sc Degree programme in Statistics, is a bonafide record of project work done by **Ms. SELIN STELVIA RODRIGUES (CCAWMST011)** during the period of her study in the Post-graduate and Research Department of Statistics, Christ College (Autonomous), Irinjalakuda, under my supervision and guidance during the year 2022-2024.

Mrs. Jiji M.B

Assistant Professor

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

Dr. Davis Antony M

Head, Department of Statistics

Self Financing

Christ College(Autonomous)

Irinjalakuda

External Examiner :

Place : Irinjalakuda

Date : June 2024 - 29



DECLARATION

I hereby declare that the project work entitled "**A study on Kumaraswamy Distribution**" submitted to Christ College(Autonomous), Irinjalakuda in partial fulfilment of the requirement for the award of Master Degree of Science in Statistics is a record of research work done by me during the period of my study in the Postgraduate and Research Department of Statistics, Christ College(Autonomous), Irinjalakuda.

Place : Irinjalakuda

Date : June 2024 - 29


Selin Stelvia Rodrigues

ACKNOWLEDGEMENT

Firstly, words cannot express the immense grace bestowed upon me by my Redeemer. This dissertation has been made possible by the collaboration of numerous people. My sincere gratitude goes out to everyone who has contributed to the success of this project by providing inspiration, support, and guidance.

My sincere gratitude goes out to Ms.Jiji, Assistant Professor in the Postgraduate and Research Department of Mathematics, Christ College (Autonomous), Irinjalakuda, who diligently mentored me throughout this endeavor.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non-teaching staff of the Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Words cannot express the love and support I have received from my parents, whose encouragement has buoyed me up from the beginning till the end of this work.

Selin Stelvia Rodrigues

Contents

List of Figures	iv
Introduction	1
1 Kumaraswamy Distribution	2
1.1 Definiton	2
1.1.1 Probability Density Function	2
1.1.2 Cumulative Distribution Function	4
1.2 Need for Kumaraswamy distribution	5
1.2.1 Flexibiltiy	5
1.2.2 Bounded Data	5
1.2.3 Simple formulation	5
1.2.4 Generalisation	5
1.3 Applications	6
1.3.1 Hydrology	6
1.3.2 Finance	6
1.3.3 Engineering and Reliability	6
1.4 Central Tendency	7
1.4.1 Mean	7

Contents

1.4.2	Median	8
1.4.3	Mode	9
2	Characteristics	10
2.1	Moment Generating Function	10
2.2	Maximum Likelihood Estimation	14
2.3	Moments	16
2.3.1	Raw Moments	16
2.3.2	Central Moments	19
2.3.3	Variance	22
2.3.4	Skewness	24
2.3.5	Kurtosis	26
3	Statistical Measures	29
3.1	Survival Function	32
3.2	Hazard Function	34
3.3	Quantile Function	36
3.4	Quartile Deviation	37
4	Relationship between Beta Distribution	39
4.1	Introduction	39
4.2	Relation using PDF	40
4.3	Relation using CDF	41
4.4	Relation using Logistic Map	42
5	Generalization of Kumaraswamy Distribution	46
5.1	Conversion of distribution into Kw-G distribution	46

5.2	A few generalized distributions of Special Kw.	48
5.2.1	KW-Normal	48
5.2.2	Kw-Weibull	50
5.2.3	KW- Exponential	51
5.2.4	KW-Pareto	52
6	Fitting of Kumaraswamy Distribution	54
7	Conclusion	55
8	Appendix	56
8.1	Python code for fitting of kumaraswamy distribution	56
	References	59

List of Figures

1.1	PDF of Kumaraswamy Distribution	3
1.2	CDF of Kumaraswamy Distribution	4
2.1	MGF of Kumaraswamy Distribution	13
2.2	Skewness of Kumaraswamy Distribution	25
2.3	Kurtosis of Kumaraswamy Distribution	28
3.1	Survival function of Kumaraswamy Distribution	33
3.2	Hazard function of Kumaraswamy Distribution	35

A STATISTICAL STUDY ON WHEAT PRODUCTION IN INDIA

Project report submitted to Christ College (Autonomous)
in partial fulfilment for the award of the M.Sc. Degree
programme in Statistics

by

VISHNUPRIYA V

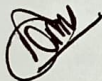
Register No. CCAWMST012



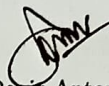
Department of Statistics
Christ College (Autonomous)
Irinjalakuda
2024

CERTIFICATE

This is to certify that the project entitled "A STATISTICAL STUDY ON WHEAT PRODUCTION IN INDIA", submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **VISHNUPRIYA V (CCAWMST012)** during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2023-2024

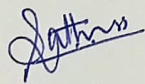


Dr.Davis Antony Mundassery
Head of Department
(Self Financing)
Department of Statistics
Christ College (Autonomous)
Irinjalakuda



Dr. Davis Antony Mundassery
Head of Department
(Self Financing)
Department of Statistics
Christ College (Autonomous)
Irinjalakuda

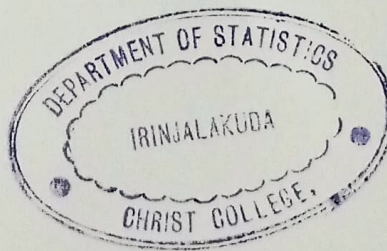
External Examiner:



Place : Irinjalakuda

Sankaran.K.K.

Date : 24/06/2024



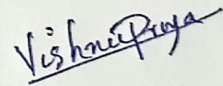
DECLARATION

I hereby declare that the matter embodied in the project entitled 'A STATISTICAL STUDY ON WHEAT PRODUCTION IN INDIA ', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Dr.Davis Antony Mundassery, Head of Department, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2023-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associateship, fellowship etc. of any other university or institution.

Irinjalakuda

Date:



VISHNUPRIYA V

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Dr.Davis Antony Mundassery, Head of Department, Department of Statistics,for his generous help,constructive criticism, scholarly guidance,valuable supervision and encouragement throughout the preparation of the project would not have been materialized.

I would like to give my sincere thanks to my teachers for the inspiration, encouragement and technical help they bestowed upon me.I am indebted to the faculty of the Department for sharing with me their knowledge base and for giving me a better perspective of the subject and for providing the necessary facilities during the span of my study

My sincere thanks are also due to Librarian and non-teaching staff of the Christ College(Autonomous) Irinjalakuda for their help and co-operations.Also,I register my heartfull thanks to my classmates for the co-operation and warmth I could enjoy from them.Last but not least,I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date:

VISHNUPRIYA V

CONTENTS

CHAPTER 1

1.1 INTRODUCTION.....	1
1.2 OBJECTIVE OF THE STUDY.....	3
1.3 DATA DESCRIPTION.....	3

CHAPTER 2

2.1 REGRESSION ANALYSIS.....	4
2.2 TIME SERIES ANALYSIS.....	6
2.3 TIME SERIES MODELS.....	12
2.4 DIAGNOSTIC CHECKING.....	16
2.5 MODEL IDENTIFICATION.....	18
2.6 FORECASTING	19

CHAPTER 3

3.1 DATA ANALYSIS.....	20
3.2 BEST MODEL SELECTION.....	30
3.3 CORRELATION.....	31
3.4 REGRESSION ANALYSIS.....	34

CHAPTER 4

4.1 CONCLUSION.....	35
---------------------	----

REFERENCES

Forecasting of Coffee Production in India

Project report submitted to Christ College (Autonomous) in partial
fulfilment of the requirement for the award of M.Sc. Degree
programme in Statistics

by

ROSMIYA JOSEPH

Register No.CCAWMST013



Department of Statistics

Christ College (Autonomous)

Irinjalakuda

2024

CERTIFICATE

This is to certify that the project entitled '**Forecasting of Coffee Production in India**', submitted to the Department of Statistics in Partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is a bonafied record of original research work done by **ROSMIYA JOSEPH(CCAWMST013)** during the period of her study in the Department of Statistics, Christ College (Autonomous) Irinjalakuda, Thrissur, under my supervision and guidance during the year 2022-2024.

Dr.Davis Antony M

Assistant Professor

Head, Department of Statistics (self financing)

Christ College (Autonomous)

Irinjalakuda

Dr. Davis Antony M

Course Coordinator

Department of Statistics

Christ College (Autonomous)

Irinjalakuda

External Examiner:

Place: Minjalarunda

Date: 29/06/24

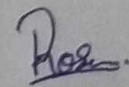
DECLARATION

I hereby declare that the matter embodied in the project entitled ' Forecasting of Coffee Production in India', submitted to the Department of Statistics in partial fulfillment of the requirements for the award of the Masters Degree in Statistics, is the result of my studies and this project has been composed by me under the Guidance and Supervision of Dr.Davis Mundassery, Assistant Professor, Department of Statistics, Christ College (Autonomous) Irinjalakuda, during 2022-2024.

I also declare that this project has not been previously formed the basis for the award of any degree, diploma, associate ship, fellowship etc. of any other university or institution.

Irinjalakuda

Date: 29/06/24



ROSMIYA JOSEPH

ACKNOWLEDGEMENT

This project would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of the study.

I would like to express my deepest gratitude to my guide Dr.Davis Antony M, Assistant Professor, Department of Statistics,whose guidance has been of immense help in successfully completing this report.

With great pleasure, I take this opportunity to express my sincere gratitude to my respected teachers, friends and non teaching staff of Department of Statistics, Christ College (Autonomous) Irinjalakuda, for their valuable advice and encouragement throughout my course.

Last but not least, I am indebted to my parents for their unconditional love and support.

Irinjalakuda

Date: 29/06/24

ROSMIYA JOSEPH

Contents

1	Introduction	9
2	Methodology	19
2.1	Time Series	19
2.2	Stationary Process	20
2.2.1	Weakly stationary	20
2.2.2	Strictly stationary	21
2.3	Auto Regressive processes (AR)	21
2.4	Moving Average (MA) process	22
2.5	Mixed auto regressive-moving average processes(ARMA) . . .	23
2.6	Auto Regressive Integrated Moving Average (ARIMA) model .	24
2.7	Auto Correlation Function(ACF)	25

2.8	Partial Auto Correlation Function (PACF)	25
2.9	Augmented Dickey-Fuller test	26
2.10	Akaike Information Criterion (AIC)	27
2.11	Normalised Bayesian Information Criterion(BIC)	28
2.12	Residual Analysis	29
2.13	Ljung-Box Test	30
2.14	Forecasting	31
3	Time Series Analysis	32
3.1	Time series plot	33
3.2	Decomposition of the Time Series	34
3.3	ACF and PACF plots	36
3.4	Augmented Dickey Fuller Test	37
3.5	Differencing plot	39
3.6	ARIMA Modelling using python	41
3.7	Residual Analysis	42
3.8	Forecasting	46

3.8.1 Forecasting table	47
4 Conclusion	48
5 References	50

List of Figures

3.1	Time Series Plot of Arabica and Robusta	33
3.2	Decomposition of Arabica coffee production	34
3.3	Decomposition of Robusta coffee production	35
3.4	ACF and PACF plot off Arabica production	36
3.5	ACF and PACF plot of Robusta production	37
3.6	Differencing plot of Arabica Production	39
3.7	Differencing plot of Robusta Production	39
3.8	ACF plot	40
3.9	PACF plot	40
3.10	Time series plot of residuals	42
3.11	ACF and PACF Plots of Residuals for Arabica	43

3.12 ACF and PACF Plots of Residuals for Robusta	43
3.13 Forecast	46