International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Data Mining Algorithm To Predict The Factors for Agricultural Development using Stochastic Model

P. Rajesh, M. Karthikeyan, R. Arulpavai

Abstract*: Data mining is a route of extract patterns from data and it is used to transforms unprocessed collections of data into information. In this paper proposes an efficient factor of agriculture development using different type of agriculture related datasets by adopting stochastic modeling and data mining approaches. Firstly, the novel model and algorithm is proposed to predict the growth of agriculture based on primary and secondary data using stochastic modeling approach. Numerical illustrations and various expected estimations also prove the proposed system approach.

Keywords: Agriculture, Data Mining, weather conditions and stochastic model.

I. INTRODUCTION

Data mining is a platform for retrieving hidden information using predefined data. It is used to retrieves different estimations using statistical model. Data mining include different techniques and algorithms to solve different unsolved problems [1]. Furthermore, in this paper a proposed novel stochastic model using data mining approach. It is used to gather lot of hiding information relating to the proof of agriculture development using rainfall, temperature, groundwater, fertilizers, etc.

In India 70% of the state populations are mixed up in agricultural activities and also one of the major revenue in India. Tamilnadu has occupied an area of 1.3 lakh sq. km with area of around 63 L.HA for plantation. Agriculture, with its related sectors, is the major source of livelihoods in growing country like India [2].

Stochastic models represent a situation where uncertainty is present. In other words, it's a model for a process that has some kind of unpredictability. The word stochastic comes from the Greek word inherit from stokhazesthai and its meaning to aim or guess or estimate. In the real word, doubt is a part of day by day life, so a stochastic model could strongly represent anything. The conflicting is a deterministic model, which predicts outcome with 100% confidence. Deterministic models for all time have a set of equations that describe the system inputs and outputs accurately. On the other hand, stochastic models will likely fabricate unusual results all the time the model is run. In "Stochastic" means random, so a "stochastic process" could more simple be called a random process [3].

Statistical model was used to estimate the groundwater dataset using precipitation as well as to predict the groundwater dataset with easily measurable climate date and also using regression model was established regression analysis were conducted to understand the relationship [4]. In any forming system taking to considering in to rainfall, water sources, region and production of crops. The authors to discover the level of concentration in paddy improvement using stochastic model [5]. Water table depth is an important map in environmental models' assessments. To develop an early water table prediction model for North Sinai, Egypt, different approaches consider namely GIS, remote sensing, simulation and stochastic methods. Stochastic (using time-series) modeling used to characterize the water table dynamics in terms of risk [6].

The author discuss with different labours and various types of agriculture workers and other important applicable details as input dataset. Observation of different years of government organization data is to be declared most of the agriculture related labours percentages are decreased year by year. Predicting the data and how to increase the agriculture labours involvement in future events using stochastic model approach [7]. Different stochastic models techniques and definitions explained related to agriculture and other social impact area has been clearly explained [8, 9].

II. PROPOSED SYSTEM

A stochastic model is one in which the epistemic uncertainties in the variables are taken into account. The uncertainties are those due to natural variation in the process being modeled. The variable is a quantity whose value changes in time series datasets. A discrete random variable is a variable whose values are obtained by counting. A continuous random variable is a variable which is used to whose values is obtained by measuring. A random variable is an important variable whose value is a numerical outcome of a number.

Manuscript published on 30 September 2019
* Correspondence Author
P. Rajesh*, Assistant Professor, PG Department of Computer Science, Government Arts and Science, C.Mutlur, Chidambaram, Tamilnadu, India. Email: rajeshdatamining@gmail.com
M. Karthikeyan, Assistant Professor, Division of Computer and Information Science, Annamalai University, Annamalainagar, Tamilnadu, India. Email: karthiaucse@gmail.com
R. Arulpavai, Statistical Inspector, Divisional Statistical Office, Department of Economics and Statistics, Mayiladuthurai, Tamilnadu, India. Email: pavaibsi@gmail.com

The proposed model using discrete random variable X, Y and Z has a countable number of possible values denoting the primary fields at the i^{th} decision epoch, i = 1, 2, … n. Y is a another discrete random variable using secondary fields. 'W' denoting the continuous random variable.

The Laplace transform L(.) is simplification on a large class of functions. The inverse Laplace transform takes a function of a complex variable s and a function of a real variable time t. Given a simple mathematical description of an input or output to a system, the Laplace transform provides an alternative functional description that often simplifies the process of analyzing the behavior of the system, or in synthesizing a new system based on a set of specifications [10].

Laplace transformation from the time domain to the frequency domain transforms differential equations into algebraic equations and convolution into multiplication. It has many applications in the sciences and technology [11].

In mathematics convolution f_k (.) is an operation on two functions of f and g to produce a third function that expresses how the shape of one is modified by the other. The term convolution refers to both the result purpose and to the process it. Convolution is similar to cross-correlation [10].

For continuous functions, the cross-correlation operator is the adjoint of the convolution operator. Convolution has applications that include probability, statistics, computer vision, natural language processing, image and signal processing, engineering, and differential equations [12].

A convolution is an integral that expresses the amount of overlap of one function g as it is shifted over another function f. It therefore "blends" one function with another. For example, in synthesis imaging, the measured dirty map is a convolution of the "true" CLEAN map with the dirty beamistribution). The convolution is sometimes also known by its German name, faltung ("folding") [13].

Now the probability that the threshold level is not reached till't'.

S(t) = P[T > t] = P [The Total antigenic diversity due to ' k' contacts does not cross the threshold level and total due to 'k' contacts does not cross the threshold].

$$S(t) = P\left[\sum_{i=1}^{k} x_i < z_1 \cap \sum_{i=1}^{k} y_i < z_2\right]$$

$$= P\left[\sum_{i=1}^{k} x_i < z_1\right] P\left[\sum_{i=1}^{k} y_i < z_2\right]$$

= Pr [That there are k contacts in (0,t) and the total antigenic diversity does not cross the threshold and the virulence does not cross the threshold]

$$S(t) = \sum_{k=0}^{\infty}\left[F_k(t) - F_{k+1}(t)\right]\left[\int_0^{\infty} g_k(x)\overline{H(x)}\,dx\right]\left[\int_0^{\infty} q_k(y)\overline{M(y)}\,dy\right] \quad (1)$$

where $H(x) = 1 - \overline{H(x)}$ and $M(y) = 1 - \overline{M(y)}$

It is assumed that,

$Z_1 \sim \exp(\theta)$ and $H(x) = 1 - e^{-\theta x}$

$Z_2 \sim \exp(\lambda)$ and $M(y) = 1 - e^{-\lambda y}$

$\overline{H(x)} = e^{-\theta x}$ and $\overline{M(y)} = e^{-\lambda y}$

Hence

$$S(t) = \sum_{k=0}^{\infty}\left[F_k(t) - F_{k+1}(t)\right]\left[\int_0^{\infty} g_k(x) e^{-\theta x}\,dx\right]\left[\int_0^{\infty} q_k(y) e^{-\lambda y}\,dy\right] \quad (2)$$

$$S(t) = \sum_{k=0}^{\infty}\left[F_k(t) - F_{k+1}(t)\right]\left[g_k^*(\theta)\, q_k^*(\lambda)\right]$$

$$= \sum_{k=0}^{\infty}\left[F_k(t) - F_{k+1}(t)\right]\left[g^*(\theta)\, q^*(\lambda)\right]^k$$

$$= 1 - \left[1 - g^*(\theta)\, q^*(\lambda)\right]\sum_{k=1}^{\infty} F_k(t)\left[g^*(\theta)\, q^*(\lambda)\right]^{k-1}$$

L(t) = 1 – S(t)

$$= \left[1 - g^*(\theta)\, q^*(\lambda)\right]\sum_{k=1}^{\infty} F_k(t)\left[g^*(\theta)\, q^*(\lambda)\right]^{k-1} \quad (3)$$

Taking Laplace transform of both sides we have

$$l^*(s) = \left[1 - g^*(\theta)\, q^*(\lambda)\right]\sum_{k=1}^{\infty} f_k^*(s)\left[g^*(\theta)\, q^*(\lambda)\right]^{k-1} \quad (4)$$

we assume that $f(.) \sim \exp(\eta)$ and $f^*(s) = \frac{\eta}{\eta + s}$

$g(.) \sim \exp(\beta)$ and $g^*(\theta) = \frac{\beta}{\theta + \beta}$

$q(.) \sim \exp(c)$ and $q^*(\lambda) = \frac{\alpha}{\lambda + \alpha}$

we obtain the E(T) which means to finding the expected estimations based on different problems.

$$l^*(s) = \left[1 - g^*(\theta)\, q^*(\lambda)\sum_{k=1}^{\infty} \frac{\eta}{\eta + s}\left[g^*(\theta)\, q^*(\lambda)\right]\right] \quad (5)$$

$$E(T) = \left.\frac{-dl^*(s)}{ds}\right|_{s=0}$$

$$1 - g^*(\theta) q^*(\lambda) = 1 - \frac{\beta}{\beta + \theta}\cdot\frac{\alpha}{\alpha + \lambda}$$

$$= \frac{\beta\lambda + \theta\alpha + \theta\lambda}{\beta\alpha + \beta\lambda + \theta\alpha + \theta\lambda} \quad (6)$$

$$\left[g^*(\theta) q^*(\lambda)\right]^{k-1} = \left[\frac{\beta\alpha}{(\beta + \theta)(\alpha + \lambda)}\right] \quad (7)$$

$$E(T) = \frac{\beta\lambda + \theta\alpha + \theta\lambda}{\beta\alpha + \beta\lambda + \theta\alpha + \theta\lambda}\left[\frac{\beta\alpha}{(\beta + \theta)(\alpha + \lambda)}\right] \quad (8)$$

III. NUMERICAL ILLUSTRATIONS

Government of Tamil Nadu is pursuing systematic effort to increase the food grain production by resorting to numerous innovative methods. Relentless efforts taken by the Government for a speedy information dissemination of scientifically proven strategies and schemes, coupled with rapid technology adoption by the farmers led to a prodigious increase in food grain production over 100 L MT in 2010 to 2016 which building on the success. In table 1, the food grain production achieved by the Department of Agriculture, Government of Tamilnadu.

Retrieval Number: C4963098319/2019©BEIESP
DOI:10.35940/ijrte.C4963.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The following table taken from Department of Economic and Statistics, Department of Agriculture, ENVIS Centre, Tamilnadu State Council for Science, Ministry of Environment and Forests and Climate Change, India Meteorological Department (IMD), Government of India.
The dataset display time series data from 2010 to 2016, which is include food grains (L MT), rainfall (MM), temperature (Celsius), groundwater level (M) and fertilizers (tone.). In table-I, include different measurements of data then these type datasets not possible to apply to the proposed stochastic model (8).

Table-I: Actual time series data include Agri. Productions (L MT), rainfall (MM), temperature (Celsius), groundwater level (M) and fertilizers (Tones.)

Year	Agri. Production (L MT)	Rainfall (MM)	Ground Water Level (M)	Temp. (Celsius)	fertilizers (Tones.)
2010	126.67	937.80	13.20	34.2	32049
2011	124.75	1165.10	11.70	32.6	41799
2012	120.78	937.00	11.50	33.6	57902
2013	125.04	743.10	13.00	32.5	70758
2014	124.30	790.60	23.60	32.3	90974
2015	123.22	987.90	24.35	33.1	116393
2016	128.03	1138.80	21.80	33.4	143104

Data mining is a procedure of extracting associations and patterns from data; therefore, DM transforms unprocessed collections of data into information. It is well recognized that data can have dissimilar formats and can be stored through a variety of different storage models [15]. In a number of fields of machine intelligence, e.g. in texture clustering, image retrieval, speech recognition and clustering of databases, an object is often represented by a vector variable, namely the feature vector A. The collection of objects described by the same features is called a dataset. The goal of clustering is to discover similarities and patterns within a large dataset by splitting data into clusters (groups). Because it is assumed that the data are unlabelled, clustering is often considered as the most important unsupervised learning problem [16] [17]. The author discusses agricultural data with various factors that help to the growth of agriculture sectors using stochastic approach [18] [19].

Table-II, shows, the normalization equation, which is used to convert different scales of data to uniform format using feature scaling (9). This movement is basic when dealing with the parameters using different units and sizes of data. In information handling, it is otherwise called data normalization and is generally performed the information preprocessing step is accustomed to carry all values into the range [0, 1]. This is additionally called unity-based standardization. This can be summed up to limit the scope of values in the dataset between any self-assertive point 'a' and 'b' and also assign (0.1 and 0.9) respectively.

$$X' = a + \frac{(X - X_{\min})(b - a)}{X_{\max} - X_{\min}} \quad (9)$$

The following estimations to proposed stochastic model (8) and normalization (9). In this approach, the get required inputs based on primary time series dataset (table-I) and normalized dataset (table-II) processes those data using the proposed model, finally deliver various expected estimation factors for increasing agriculture growth the sternness of the approach.

A. Pseudo Code for Normalization and Stochastic Model

```
BEGIN
Initialize the model parameters
SET n ← 5
SET β ← rainfall [937.80, 1165.10, 937.00, 743.10, 790.60, 987.90, 1138.80]
SET θ ← groundwater [13.20, 11.70, 11.50, 13.00, 23.60, 24.35, 21.80]
SET α ← temperature [34.2, 32.6, 33.6, 32.5, 32.3, 33.1, 33.4]
SET λ ← fertilizer [32049, 41799, 57902, 70758, 90974, 116393, 143104]
INPUT: Proposed stochastic model parameters (β, θ, α, λ)
OUTPUT: Expected factors that affect the agriculture growth
Generate control sequence using equation (8)
for i ← 1 to n do
    SET a ← 0.1 and SET b← 0.9
for j ← 1 to n do
 beta[j] ← a + ((β[j] − Min(β)) (b − a)) /(Max(β) − Min(β))
 RETURN beta
end for
for j← 1 to n do
 theta[j] ← a + ((θ[j] − Min(θ)) (b − a)) /(Max(θ) − Min(θ))
 RETURN theta
end for
for j← 1 to n do
 alpha[j] ← a + ((α[j] − Min(α)) (b − a)) /(Max(α) − Min(α))
 RETURN alpha
end for
for j← 1 to n do
 lambda[j] ← a + ((λ[j] − Min(λ)) (b − a)) /(Max(λ) − Min(λ))
     RETURN lambda
end for
Generate initial sequences using Equation (11)
SET count ← 1
SET x ← 1
while count < 6 do
    PART1[x] ← (beta[x] *lambda[x]) +
               (theta[x]*alpha[x]) + (theta[x]*lambda[x])
    PART2[x] ← (beta[x]*alpha[x]) + (beta[x]*lambda[x])
             + (theta[x]*alpha[x]) + (theta[x]*lambda[x])
    PART3[x] ← (beta[x]*alpha[x]) / ((beta[x] +
               theta[x])*(alpha[x] + lambda[x]))
    count+1 and x+1
End while
SET count ← 1 and x ← 1
while count < 6 do
     EAG[x] ← (PART1[x]/PART2[x])*PART3[x]
     RETURN EAG
     count+1 and x+1
End while
END
```

Table-II: Normalized data for Agri. Products (L MT), rainfall (MM), temperature (Celsius), groundwater level (M) and fertilizers (Tones.)

Year	Agri. Productions	Rainfall (α)	Ground water (β)	Temp. (λ)	Fertilizers (θ)
2010	0.7499	0.4691	0.2058	0.9000	0.1000
2011	0.5381	0.9000	0.1125	0.2263	0.1702
2012	0.1000	0.4676	0.1000	0.6474	0.2862
2013	0.5701	0.1000	0.1934	0.1842	0.3788
2014	0.4884	0.1900	0.8533	0.1000	0.5245
2015	0.3692	0.5641	0.9000	0.4368	0.7076
2016	0.9000	0.8501	0.7412	0.5632	0.9000

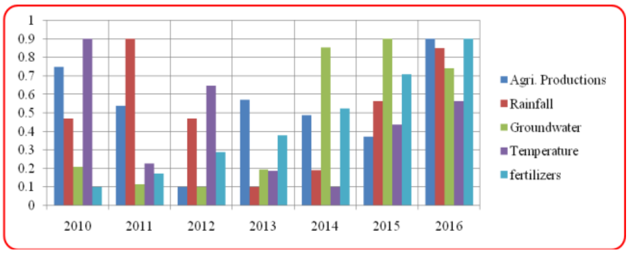

Fig. 1. Comparison of Normalized data for Agri. Products (L MT), rainfall (MM), temperature (Celsius), groundwater level (M) and fertilizers (Tones.)

Table-III: Expected Agriculture Growth using α, β, θ (increase) and λ (fixed)

Rainfall (α)	Ground water (β)	Temp. (λ)	Fertilizers (θ)	Expected Agri. Growth
0.1	0.1	0.1	0.1	0.1875
0.2	0.2	0.1	0.2	0.2222
0.3	0.3	0.1	0.3	0.2344
0.4	0.4	0.1	0.4	0.2400
0.5	0.5	0.1	0.5	0.2431

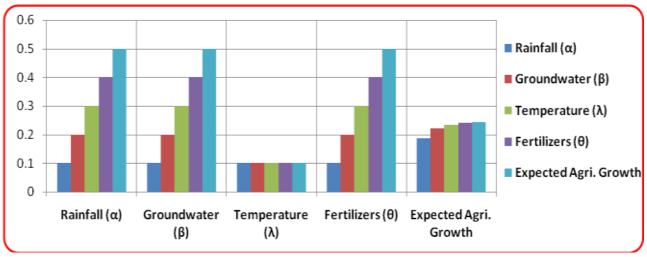

Fig. 2. Expected Agriculture Growth using α, β, θ (increase) and λ (fixed)

Table-IV: Expected Agriculture Growth using α, λ (increase) and β, θ (fixed)

Rainfall (α)	Ground water (β)	Temp. (λ)	Fertilizers (θ)	Expected Agri. Growth
0.1	0.1	0.1	0.1	0.1875
0.2	0.1	0.2	0.1	0.1728
0.3	0.1	0.3	0.1	0.1523
0.4	0.1	0.4	0.1	0.1344
0.5	0.1	0.5	0.1	0.1196

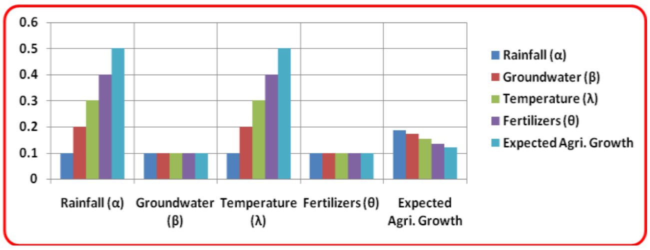

Fig. 3. Expected Agriculture Growth using α, λ (increase) and β, θ (fixed)

Table-V: Expected Agriculture Growth using α, β (decrease) and θ, λ (increase)

Rainfall (α)	Ground water (β)	Temp. (λ)	Fertilizers (θ)	Expected Agri. Growth
0.5	0.5	0.1	0.1	0.2522
0.4	0.4	0.2	0.2	0.2469
0.3	0.3	0.3	0.3	0.1875
0.2	0.2	0.4	0.4	0.0988
0.1	0.1	0.5	0.5	0.0270

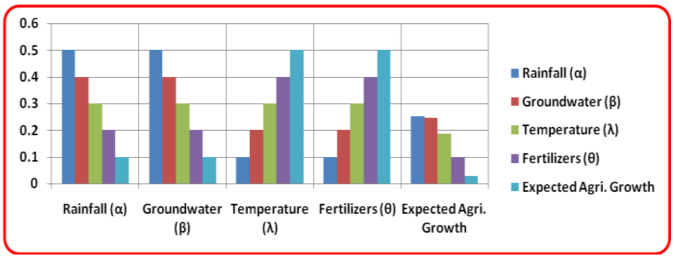

Fig. 4. Expected Agriculture Growth using α, λ (increase) and β, θ (fixed)

Table 6: Expected Agriculture Growth using α, λ, β (increase) and θ (decrease)

Rainfall (α)	Ground water (β)	Temp. (λ)	Fertilizers (θ)	Expected Agri. Growth
0.1	0.1	0.1	0.5	0.0764
0.2	0.2	0.2	0.4	0.1389
0.3	0.3	0.3	0.3	0.1875
0.4	0.4	0.4	0.2	0.2222
0.5	0.5	0.5	0.1	0.2430

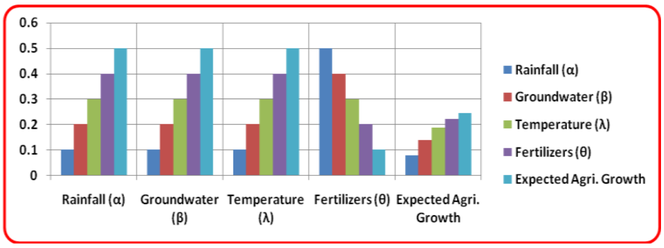

Fig. 5. Expected Agriculture Growth using α, λ (increase) and β, θ (fixed)

IV. RESULT AND DISCUSSION

Table-I shows the time series data from 2010 to 2015. The dataset include agriculture productions (L MT), rainfall (MM), groundwater (M), temperature (Celsius) and fertilizer (tones).

The dataset contains various measurements of data, it is not suitable for apply the proposed stochastic model equation. In table-II shows the normalized dataset, it is suitable for uniform data range from 0.1 to 0.9 and also to assign them in symbolic representation. The primary field agriculture productions, the rainfall values are assigned as ‘α’. Similarly, other field’s groundwater are assigned as ‘β’, temperature level is named as ‘λ’ and fertilizer named as ‘θ’. In this assumption is very useful for applying numerical values easily to the proposed model.

Numerical illustration of table-II, in 2016 occurs a maximum agriculture production at the same year rainfall, groundwater and fertilizer also increased in nature. In 2012 has low production which is presented groundwater, fertilizer and rainfall also decreased and also the temperature also high in this result depicted in fig. 1.

Further more in table-III, if rainfall, groundwater and fertilizer values as increased and the temperature value is fixed, in this case the expected agriculture growth also increased. The results show in fig. 2.

The result and discussion of table-III and figure-IV, the rainfall and temperature value increased equally at the same time groundwater and fertilizer performance is kept fixed as a minimum value. In this case the expected growth of agriculture also decreased.

Further more in table-V and fig. 4, if α and β values are decrease and λ, θ values are increase; in this nature the expected agriculture growth also decreased using the proposed stochastic equations. The result and discussion of table 6, the normalized parameter of α, λ and β values are increases year by year and at the same year θ value is decreased. In this case, the proposed stochastic model produced the expected agriculture growth is increases, the result shows in fig. 5.

V. CONCLUSION

The increasing factors of agriculture growth are one of the sustainable developments in nations like India. The use of different data mining and stochastic approaches the primary factors for increase agriculture growth is rainfall, groundwater, temperature and fertilizer, the same statement proved using proposed system. The proposed model convert into GUI in future and also easily predict this kind problems. This model is not only in the field of groundwater analysis and additionally in a more extensive setting to utilize other social effect zones.

REFERENCES

1. P. Rajesh, and M. Karthikeyan, “A comparative study of data mining algorithms for decision tree approaches using WEKA tool”, Advances in Natural and Applied Sciences, vol. 11(9), 2017, pp.230-243.
2. Wikipedia (2019, Sep, 05), *Economy of India*, Available:
3. https://en.wikipedia.org/wiki/Economy_of_India#cite_note-153
4. C. Peter (2017, Oct, 29). *Risk assessment professional who uses statistics.* Available: https://www.quora.com/An-example-of-stochastic-model.
5. S. Yan, S. Yu, Y. Wu, D. Pan, and J. Dong, “Understanding groundwater table using a statistical model”, Water Science and Engineering, vol. 11(1), 2018, pp. 1-7.
6. P. Rajesh, and M. Karthikeyan, “Prediction of agriculture growth and level of concentration in paddy - a stochastic data mining approach”, Advances in Intelligent Systems and Computing, vol. 750, 2019, pp. 127-139.
7. E. Sayed Omran, “A stochastic simulation model to early predict susceptible areas to water table level fluctuations in North Sinai, Egypt”, The Egyptian Journal of Remote Sensing and Space Science, vol. 19(2), 2016, pp. 235-257.
8. P. Rajesh, and M. Karthikeyan, “Predication of labour demand in agriculture based on comparative study of different data using data mining and stochastic approach”, International Journal of Engineering Science Invention, vol. 2, 2018, pp. 86-89.
9. D. J. Bartholomew, *The Stochastic model for social processes*, John Wiley and Sons, New York, 1982.
10. P. Mucherino, J. Papajorgji, and P. M. Pardalos, *Data mining in agriculture*. Springer Science & Business Media, 2009.
11. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers (2nd ed.)*, McGraw-Hill Companies, 1967.
12. Wikipedia (2019, Sep, 02), Laplace transform, Available: https://en.wikipedia.org/wiki/Laplace_transform
13. Wikipedia (2019, Aug, 27), Convolution, Available: https://en.wikipedia.org/wiki/Convolution
14. http://mathworld.wolfram.com/Convolution
15. J. Gurland, “Distribution of the Maximum of the Arithmetic Mean of Correlated Random Variables”, Annals of Mathematical Statistics, vol. 26, 1955, pp. 294-300.
16. M. M. Suarez-Alvarez, D. T. Pham, M. Y. Prostov and Y. I. Prostov, “Statistical approach to normalization of feature vectors and clustering of mixed datasets”, Proc. R. Soc. A, vol. 468, 2012, pp. 2630-2651.
17. Jain, A. K. and R. C. Dubes, Algorithms for clustering data. Englewood Cliffs, NJ: Prentice Hall, 1988.
18. K. J. Cios, W. Pedrycz, R. W. Swiniarski and L. A. Kurgan, *Data mining: a knowledge discovery approach*. Springer Nature Switzerland AG., 2007.
19. P. Rajesh and M. Karthikeyan, “Data assimilation of gross domestic product (gdp) in india using stochastic data mining approach”, Journal of Computational and Theoretical Nanoscience, vol 16(4), 2019, pp. 1478-1484.
20. P. Rajesh and M. Karthikeyan, “Data mining approaches to predict the factors that affect the agriculture growth using stochastic model”, International Journal of Computer Sciences and Engineering, vol. 7(4), 2019, pp. 18-23.

AUTHORS PROFILE

P. Rajesh is an Assistant Professor in PG Department of Computer Science, Government Arts College, Chidambaram, Tamilnadu, India. He completed his B.Sc. (Computer Science), from Madras University in 2001, M.C.A. from Annamalai University, in 2004 and M.Phil. from Annamalai University in 2008. Presently pursuing Ph.D. in Computer Science in Annamalai University and the area of specializations as Data Mining and Big Data analysis. Since, starting his teaching carrier as Guest Lecturer, Department of Computer Science, Government Arts College, C.Mutlur, Chidambaram from 2004 to 2007 and 2007 onwards working as an Assistant Professor/Programmer, Department of Computer and Information Science, Annamalai University. In February 2018 onwards working as Assistant Professor, PG Department of Computer Science, Government Arts College, C.Mutlur, Chidambaram (Totally 15 years of teaching experience). He is published two books and 8 research articles published in Scopus indexed journals. In 2007 to till participated more than 25 National and International seminars / conference / workshops and also presented more than 15 papers for various National and International level conferences. He is getting best research paper presentation award in 2018 in Annamalai University and he is the member of Indian Society of Genetic and Plant Breeding.

Dr. M. Karthikeyan working as Assistant professor in Division of Computer and Information Science, Annamalai University, India. He completed his M.Sc. [Computer Science] from Bharather University in 1993 and M.Phil [Computer Science] and Phd from Annamalai University in 2005 and 2014 respectively. His area of interest is Data Mining, Digital Image Processing, and Artificial Neural Networks. He is published more than 25 research papers published in national and international journals and also more than 25 papers presented in various international and national level seminars and conferences. Under his guidance more than 10 students completed the M.Phil. degree and presently 8 scholars pursuing Ph.D. under his guidance.

Dr. R. Arulpavai working as Statistical Inspector, Divisional Statistical Office, Department of Economics and Statistics, Mayiladuthurai. She is completed his B.Sc., from Annamalai University in 2000, M.Sc. (Statistics) with Distinction from Annamalai University, in 2002, M.Phil. (Statistics) with Distinction from Annamalai University in 2005 and completed her Ph.D. (Statistics) from Manonmaniyam Sundaranar University in 2015. Her area of research interest in manpower planning and stochastic model. Since, starting his teaching carrier as Guest Lecturer, Department of Statistics, Government Arts College, C.Mutlur, Chidambaram, India from 2004 to 2014 (Totally 10 years of teaching experience) and presently working as Statistical Inspector from 2014 to till date. She is published three books and 10 research articles published in various indexed journals. In 2004 to till participated more than 15 National and International seminars / conference / workshops and also presented more than 10 papers for various National and International level conferences.