A DEEP LEARNING APPROACH FOR LSTM BASED COVID-19 FORECASTING SYSTEM

[1]K. Jothimani, [2]N. Gowsalya, [3]S. Logeshwaran, [4]E. Monika, [5]M. Vinothini

[1]Assistant Professor, Department of IT, Nandha Engineering College, Erode, Tamil Nadu, India

[2,3,4,5]UG Student, Department of IT, Nandha Engineering College, Erode, Tamil Nadu, India

Abstract: COVID-19 has proliferated over the earth, exposing mankind at risk. The assets of the world's most powerful economies are at stake due to the disease's high infectivity and contagiousness. The capacity of machine learning algorithms can estimate the amount of future COVID-19 cases, which is now considered a possible threat to civilization. Five conventional measuring models, notably LR, LASSO, SVM, ES, and LSTM, were utilised in this work to examine COVID-19's undermining variables. Each model contains three sorts of expectations: the number of newly contaminated cases, the number of passings, and the number of recoveries. However, it is hard to anticipate the patients' real outcomes. To address the issue, a suggested approach based on long transient memory (LSTM) forecasts the number of COVID-19 cases in the next 10 days as well as the impact of preventative measures such as social isolation and lockdown on COVID-19 spread.

Key words: Long Short Term Memory (LSTM), Lasso (least absolute shrinkage and selection operator), Support Vector Machine(SVM), Logistic Regression(LR), Covid-19 Prediction, Algorithms, Machine Learning, Deep Learning

Introduction:

Corona virus, the pandemic that is spreading around the world, has uncovered the weakness of human culture to extreme irresistible illnesses and the trouble of tackling this issue in an internationally interconnected complex framework. In a few of weeks, the Corona virus infected over 100 countries.

Corresponding Author: *K.Jothimani*

Department of IT, Nandha Engineering College, Erode

Tamil Nadu, India

Mail: kjothimani91@gmail.com

As an outcome, the entire human race ought team up to defeat the pestilence as well as sensibly mastermind to get back to work and creation as indicated by the real circumstance of every area and complete topographical danger assessment. Many endeavors have been directed to track down an appropriate and quick way of identifying tainted patients in a beginning phase. Guan et al discovered two-sided pneumonic parenchymal ground-glass and consolidative aspiratory opacities, at times with an adjusted morphology and a fringe lung conveyance, after performing chest CT sweeps of 21 individuals infected with COVID19 in China. As a result, COVID-19 conclusion can be treated as a picture division issue to eliminate the disease's core ingredients. The disease caused by the novel Covid, or Corona virus Disease 2019 (COVID-19), is rapidly spreading around the world. As of April 9, 2020, it had infected over 1,436,000 people in more than 200 countries and domains.

Covid infection 2019 (COVID-19) is an infectious respiratory and vascular illness caused by Covid 2, a severe respiratory illness (SARS-CoV-2). It was first discovered in Wuhan, China, and is now a global pandemic. Fever, hacking, tiredness, breathing difficulties, and loss of smell and taste are all common symptoms. Symptoms appear one to fourteen days following infection exposure. While the vast majority experience minor side effects, some people develop acute respiratory distress syndrome (ARDS), which can be accelerated by cytokine storms, multi-organ failure, septic shock, and blood clots [1-10].

Long-term harm to organs (specifically, the lungs and heart) has been observed, and there is worry about countless patients who have recovered from the intense phase of the illness but continue to experience a variety of side effects—known as long COVID—for a long time afterward, including severe tiredness, cognitive decline and other intellectual issues, second-rate fever, muscle weakness, and shortness of breath [11-16].

EXISTING SYSTEM

COVID 19 is currently considered a potential threat to humanity. In four standard prediction models, such as linear regression (left to right), at least complete summary and select operator, Support Vector Machine (SVM), have been used to predict COVID-19 threatening factors in this study. Predictions are made on each of the models, such as the number of new infections, the number of deaths, and the number of recurrences over the next 10 days. For the effects of the study it demonstrates a promising mechanism for the use of these methods in the current context of COVID 19 infection. Predictions are made on each of the models, such as the number of new infections, the number of deaths, and the number of recurrences over the next 10 days. For the effects of the study it demonstrates a promising mechanism for the use of these methods in the current context of COVID 19 infection. Despite the fact that COVID-19 appears to have little impact on children, various medical centers have been established to deal with COVID-19-related crises [17-23].

PROBLEM STATEMENT

COVID-19 is spread primarily through the air when people are in close proximity for lengthy periods of time, [a] predominantly through small droplets or aerosols, as an infected person breathes, coughs, sneezes, sings, or speaks. Transmission through fomites (infected surfaces) has not been shown conclusively. It can spread from asymptomatic (no symptoms) individuals as early as two days before infected people display symptoms (presymptomatic). In mild situations, people can be contagious for up to ten days, and in severe cases, up to two weeks. Real-time reverse transcription polymerase chain reaction (rRT-PCR) from a nasopharyngeal swab is the usual diagnosing procedure. Social distancing, quarantining, ventilation of indoor spaces, covering coughs and sneezes, hand washing, and keeping unclean hands away from the face are all preventive methods. To reduce the danger of transmission in public places, face masks or coverings have been proposed. COVID-19 vaccines and therapies have yet to be demonstrated, while several are in the works. Symptom management, supportive care, isolation, and experimental approaches are all part of the management process [24-30].

PROPOSED METHODOLOGY

Due to naturally distinguishing vital aspects from the preparation tests, taking care of the enactment from the past time venture as contribution for the current time step, and organizations self-associations, AI methods were viable for prediction [31-35]. We suggest that early crisis intervention efforts, such as obstructing, limiting individual development, and enhancing aid, had a crucial regulatory effect on the pandemic's initial spread, based on the findings of the model inquiry. The AI calculations LR, LASSO, SVM, ES, and LSTM are all used, and the best calculations are organized in the r-squared error and the changing r-squared error. Maintaining interest in various clinical assets to ensure that suspected patients can be investigated and treated without wasting time is an extremely viable anticipation and treatment method. Long transient memory (LSTM) of the plague were first fitted and broken down to illustrate the validity of the current numerical models. The findings were then used to fit and analyze COVID-19's situation. The expected outcomes of three different numerical models differ for different borders and locations. The findings acquired using the suggested technique for various parts (number of positive cases retrieved, number of cases, etc.) will be precise within a given range and will be a useful tool for management and health officials.

EXPONENTIAL SMOOTHING

Excellent smoothing is a method for smoothing time series data using dramatic window work as a guideline. While past sensations are weighted similarly in the basic moving normal, exceptional capacities are used to assign radically decreasing loads through time. It is a practical, intellectual, and efficiently applied methodology for providing assurance based on the client's prior assumptions, such as irregularity. For the examination of time-series data, dramatic smoothing is routinely used. Dramatic smoothing is one of the window capacities commonly used to smooth data in signal processing, acting as low-pass channels to minimize high-recurrence

noise. The foundation for this method was provided by Poisson's use of recursive dramatic window capabilities in convolutions in the nineteenth century, as well as Kolmogorov and Zurbenko's use of recursive moving midpoints in their disturbance research. There is no standardized way to choose the display style alpha. The analyst's judgement is occasionally utilized to identify an appropriate factor. A standard moving normal, on the other hand, permits a few cases to be missed with minimal data loss due to consistent weighting of tests inside the normal.

Design of COVID-19 Future Forecasting Process

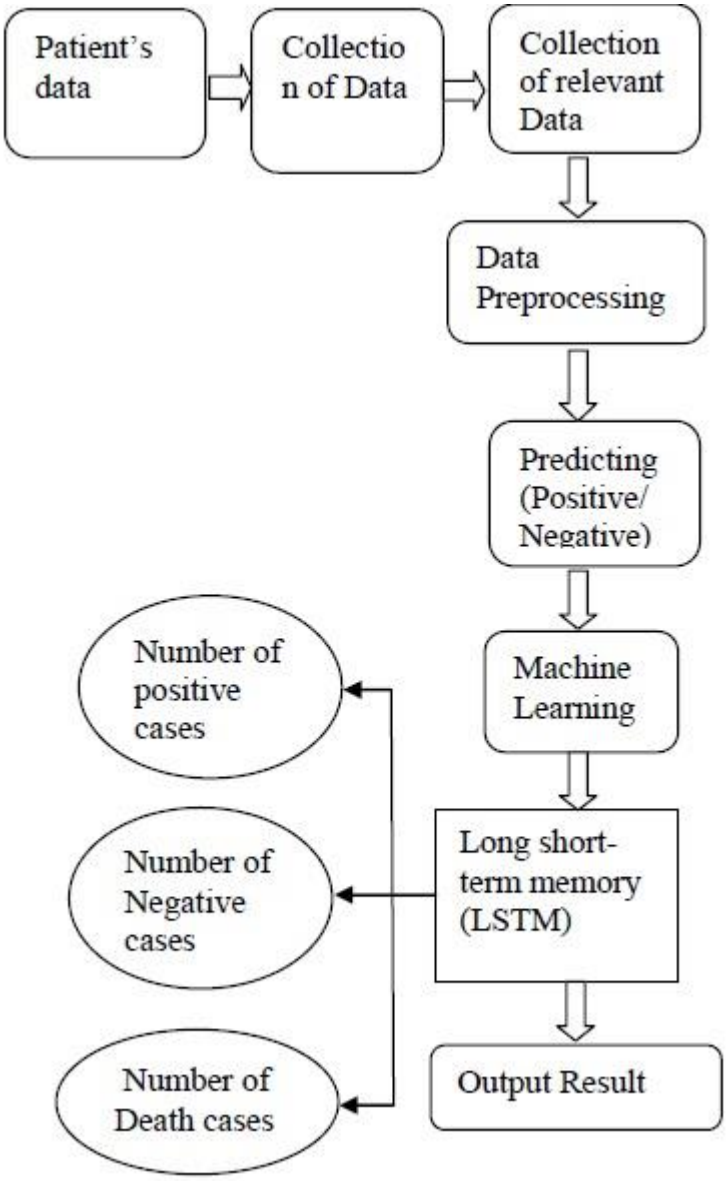

Figure 1 Overall Proposed System

To estimate future data as an element of past data, quantitative determining models are used. They are appropriate for use when previous mathematical information is available and it is reasonable to assume that some of the examples in the information may be trusted to carry on into the future. These strategies are typically used with short- or mid-range range options. Last period interest, basic and weighted N-period moving midpoints, uncomplicated remarkable smoothing, poison measure model based gauging, and multiplicative occasional lists are examples of quantitative deciding procedures. Previous research has shown that different methodologies can result in varying degrees of deciding precision. For example, GMDH neural organization performed typical anticipating computations like Single Exponential Smooth, Double Exponential Smooth, ARIMA, and back-proliferation neural organization in terms of gauging execution.

FUTURE FORECASTING

Estimating is the process of predicting future events based on data collected over a long period of time and, most commonly, by looking for patterns. An ordinary model could be used to evaluate a factor of interest at some point in the future. Expectation is a more general concept than comparison. Both can refer to formal factual procedures that use time series, cross-sectional, or longitudinal data, as well as less formal critical strategies. The terms "figure" and "determining" are sometimes reserved in hydrology for evaluations of attributes at specific future events, whereas the term "forecast" is used for more broad judgments.

Consider the times when floods occur over a long period of time. Hazard and vulnerability are essential components of anticipating and expecting; demonstrating the level of vulnerability associated with conjectures is generally seen as excellent practice. Regardless, all of the data should be current in order for the number to be as accurate as possible. Occasionally, the data used to predict the variable of interest is itself gauged. Subjective judging procedures are abstract, based on the evaluation and judgement of buyers and experts; they are appropriate when previous knowledge is unavailable. They're usually used with transitional or long-reach options. Educated assessment and judgement, the Delphi strategy, statistical surveying, and documented life-cycle similarities are examples of subjective determining techniques.

ALGORITHMS USED:

LINEAR REGRESSION

One of its most basic and widely used Machine Learning techniques is linear regression. It is a statistical technique for performing predictive analysis. The linear regression algorithm reveals a linear relationship between a dependent (y) variable and one or more independent (y) variables, thus the name. Because linear regression displays a linear relationship, it determines how the value of the dependent variable changes as the value of the independent variable changes.

LASSO Regression

Lasso (least absolute shrinkage and selection operator; alternatively Lasso or LASSO) is a regression analysis approach in statistics and machine learning that does both variable selection and regularization to improve the predictability and interpretability of the produced predictive model.

SUPPORT VECTOR MACHINE

SVM stands for Support Vector Machine and is one of the most widely used Supervised Learning algorithms for Classification and Regression issues. It is primarily used in Machine Learning to solve classification problems. The goal of the SVM method is to discover the best line or decision boundary for categorizing n-dimensional space into classes so that subsequent data points can be easily placed in the right category. The ideal choice boundary is known as a hyperplane. SVM is used to select the extreme points/vectors that help build the hyperplane. The algorithm is termed a Support Vector Machine, and support vectors are the extreme examples.

ES ALGORITHM

Exponential smoothing is a univariate time series forecasting method that can be extended to data with a systematic trend or seasonal component. It's a powerful forecasting tool that can be used instead of the widely utilised Box-Jenkins ARIMA family of algorithms.

LSTM ALGORITHM

Long short-term memory networks, or LSTMs, are a type of Deep Learning network. It's a class of recurrent neural networks (RNNs) that can learn long-term dependencies, which is useful for solving sequence prediction issues. Apart from single data points like photos, LSTM has feedback connections, which means it can process the complete sequence of data. This is useful in speech recognition, machine translation, and other areas. The LSTM is a type of RNN that performs exceptionally well on a wide range of issues.

MODULES :DATA

The data includes the total number of confirmed cases, total number of passing, recently affirmed cases, and total number of relieved cases regions. We also used information from recent decisions in South Korea, Iran, and Italy, which includes information from actual warnings from various countries. All data comes from the daily case report, and the information is updated once a day.

ESTIMATION PROCESS

The Basic proliferation number varies dramatically in various control schemes, and it has a direct impact on the control force. Similarly, the infection's hatching period has a direct impact on the rate of transmission. These two lines should be examined. According to recent research,

Basic proliferation is uncontrolled. As a result, we chose the comparison range's valuation range. The scope of value for the regulated Basic proliferation number was chosen as [0, 1.5].

DATA-DRIVEN METHODS TO PREDICT COVID-19

When the main instance of COVID-19 was accounted for in India, 80 percent of the information was used for preparation and the remaining 20% was used for estimation and approval purposes. The subsequent plot showing the absolute number of affirmed cases, the noticed information is the information utilized for preparing purposes, official information (green line) demonstrates the authority information accessible and estimated information shows the figure of an all out number of affirmed cases. From this chart, it is seen that the estimated number of absolute affirmed positive cases intently coordinates with the accessible authority information.

DATA PRE-PROCESSING

Preprocessing is a technique for transforming raw data into a clean data set. The dataset is frequently fragmented, inconsistent, and lacking in specific practices or trends, and it will almost certainly contain multiple errors. Preprocessing data is a tried and true method for resolving such challenges.

PREDICTION OF ACCURACY

This method is excellent for using precognitive neural organizations or trademark data as disease event or non-occasion binomial affects. Different estimations' expected exactness might be used for a variety of objectives. They include the rate at which a typical (non-anticipated forecast accurately predicts affectability (non-irresistible illness), precision (anticipated level of anticipated pattern), positive predictive worth, negative predictive worth (effectively anticipated contamination rate is), and the proportion is Expected expectations are a percentage of the chance that the increment in the complete cycle exceeds the person's precision).

CLASSIFICATION

For each informative index point, the order procedure predicts the goal class. A risk factor can be associated with patients using the order approach by looking at their examples of infections.

EXPERIMENTAL SETUP

To create a framework for determining the number of instances affected by COVID-19 in the future using AI procedures. The dataset used for the study includes detailed reports on the number of late-tainted cases, recoveries, and passing in light of COVID-19 around the globe. As the fatality rate and confirmed cases continue to rise, the globe is becoming increasingly concerned. The number of people who could be affected by the COVID-19 pandemic in different countries isn't large. This evaluation is a project to determine the number of persons who may be affected, as well as the number of new sullied cases and deaths, as well as the predicted recoveries for the next 10 days.

The amount of as of late debased cases, the amount of passing, and the amount of recoveries were predicted using four AI models: LR, LASSO, SVM, ES, and LSTM. The plots of stated cases, passes, and recoveries on the bottom four sheets are followed by the plot of certifiable scenario gathered from the genuine data reports of the assessment's analyzing season on the fifth sheet. The figures illustrate that the ML models utilized in this evaluation were appropriate for the assessing task, pointing the way to a more comfortable study and future investigation of the nearby environment.

ALGORITHM	ACCURACY
LR	82
LASSO	78
SVM	52
ES	96
LSTM	98

Table 1 Comparative Algorithm with Accuracy Level

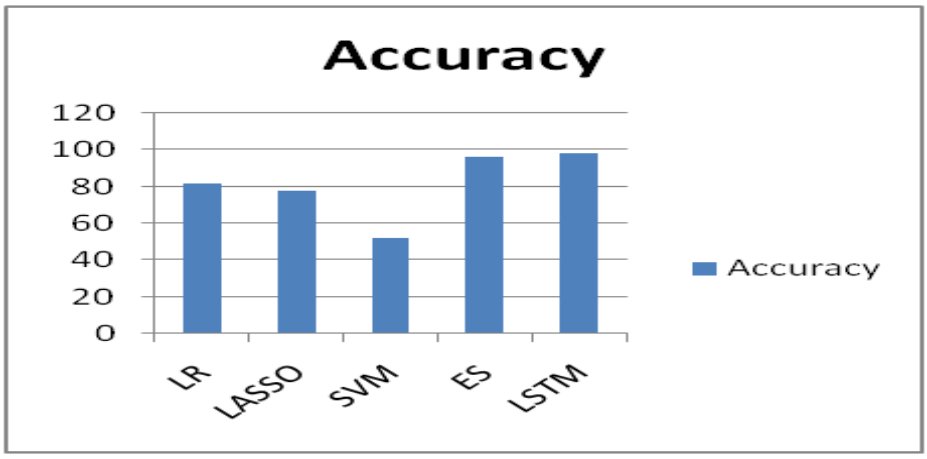

Figure 2 Proposed Accuracy Level

CONCLUSION

The number of possible COVID-19 positive cases in India for the next 10 days was estimated using an information-driven anticipating/assessment technique. The quantity of recuperated cases, long transient memory (LSTM) every day positive cases, and expired cases has likewise been assessed by utilizing and bend fitting. The impact of forestalling measures as friendly segregation and lockdown has additionally been seen which shows that by these preventive measures, the spread of the infection can be decreased essentially. Despite the fact that this technique frequently requires adequate information to assist it, in the early stages of pandemic transmission, this strategy can still be used to more precisely anticipate the pointers

of scourge transmission for the time being, in order to provide mediation control at all levels of the offices and strategy execution provides transient crisis avoidance programs. For varied boundaries and in various districts, the forecast outcomes of three distinct numerical models are diverse. The fitting influence of the Logistic model, on average, may be the greatest of the three models.

REFERENCES

1. E. Chong and A. A. R. Alsaeedy (2020). 1–1 in IEEE Open Journal of Engineering in Medicine and Biology.
2. R. F. Sear, N. Velasquez, R. Leahy, R. Restrepo, N. J. Restrepo, S. El Oud, N. Gabriel,... Johnson, N. F. (2020). Using artificial intelligence to evaluate COVID-19 content in the online health assessment war. 1–1 in IEEE Access.
3. S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, G. Yang (2020). COVID-19 Infection Detection and Classification from CT Images with Poorly Supervised Deep Learning .
4. Zhang, Y., Li, Y., Yang, B., Zheng, X., and Chen, M., IEEE Access, 1–1. (2020). COVID-19 Hazard Assessment Using Multisource Data from a Geographical Perspective IEEE Access.
5. M. Abdel-Basset, R. Mohamed, M. Elhoseny, R. K. Chakrabortty, and M. Ryan (2020). A COVID-19 finding model that combines a more advanced marine hunter's calculation with a positioning-based variety reduction process. 1–1.

6. Petropoulos, F., and Makridakis, S., "Guaging the Novel Covid Coronavirus," PLOS ONE, vol. 15, no. 3, e0231236, 2020.
7. G. Grasselli, A. Pesenti, and M. Cecconi, Jama, 2020, "Basic consideration use for the Coronavirus flare-up in lombardy, Italy: early experience and forecast during a crisis reaction."
8. C. P. E. R. E. Novel et al., "The epidemiological features of a 2019 novel Covid infection (Coronavirus) incident in China," Zhonghualiuxingbingxuezazhi= Zhonghualiuxingbingxuezazhi, vol. 41, no. 2, p. 145, 2020.
9. International Journal of Forecasting, vol. 36, no. 1, pp. 178–185, 2020.
10. N. C. Mediaite. Harvard educator sounds caution on 'probable' Covid pandemic: 40% to 70% of world could be tainted for the current year. Gotten to on 2020.02.18. [Online]. Accessible: https://www.mediaite.com/news/harvardprofessor-sounds-alert on-possible Covid pandemic-40-to-70-ofworld-could-be-contaminated for this present year/.
11. Punarselvam, E., Suresh, P., & Parthasarathy, R. (2013). Segmentation of CT scan lumbar spine image using median filter and canny edge detection algorithm. Int J Comput Sci Eng, 5, 806-814.
12.Punarselvam, E., et al. "Segmentation of Lumbar spine image using Watershed Algorithm." International Journal of Engineering Research and Applications, ISSN (2013): 2248-9622
13. Punarselvam, E., et al. "Segmentation Analysis Techniques and Identifying Stress Ratio of Human Lumbar Spine Using ANSYS." Journal of Medical Imaging and Health Informatics 10.10 (2020): 2308-2315.

14. Karthick, R., et al. “Overcome the challenges in bio-medical instruments using IOT–A review.”; Materials Today: Proceedings 45 (2021): 1614-1619.
15. Suresh, Helina Rajini, et al. “Suppression of four wave mixing effect in DWDM system.” Materials Today: Proceedings 45 (2021): 2707-2712.
16. Soundari, D. V., et al.”Enhancing network-on-chip performance by 32-bit RISC processor based on power and area efficiency.” Materials Today: Proceedings 45 (2021): 2713-2720.
17. Sabarish, P., et al., Investigation on performance of solar photovoltaic fed hybrid semi impedance source converters." Materials Today: Proceedings 45 (2021): 1597-1602.
18. Karthick, R., and M. Sundararajan. “SPIDER-based out-of-order execution scheme for HtMPSOC “; International Journal of Advanced Intelligence paradigms 19.1 (2021): 28-41.
19. Karthick, R., and P. Meenalochini. “Implementation of data cache block (DCB) in shared processor using field-programmable gate array (FPGA).” Journal of the National Science Foundation of Sri Lanka 48.4 (2020).
20. Sabarish, P., et al., “An Energy Efficient Microwave Based Wireless Solar Power Transmission System” IOP Conference Series: Materials Science and Engineering. Vol. 937.No. 1. IOP Publishing, 2020.
21. Vijayalakshmi, S., et al. ”Implementation of a new Bi-Directional Switch multilevel Inverter for the reduction of harmonics” IOP Conference Series: Materials Science and Engineering. Vol.937. No. 1. IOP Publishing, 2020.
22. Karthick, R., and M. Sundararajan. “Design and implementation of low power testing using advanced razor based processor” International Journal of Applied Engineering Research 12.17 (2017): 6384-6390. 23. Punarselvam, E., and P. Suresh. "Non-Linear Filtering Technique Used for Testing the Human Lumbar Spine FEA Model." Journal of medical systems 43, no. 2 (2019): 1-13.
24. Punarselvam, E., Hemalatha, E., Dhivahar, J., Gowtham, V., Hari, V., & Thamaraikannan, R. PREDICTING WIRELESS CHANNELS FOR ULTRA-RELIABLE LOW-LATENCY COMMUNICATIONS.
25. Punarselvam, E., et al. "Segmentation of Lumbar spine image using Watershed Algorithm." International Journal of Engineering Research and Applications, ISSN (2013): 2248-9622.
26. Punarselvam, E., and P. Suresh. "Edge detection of CT scan spine disc image using canny edge detection algorithm based on magnitude and edge length." 3rd International Conference on Trendz in Information Sciences & Computing (TISC2011). IEEE, 2011. (2021): 4991-5004.
27. Punarselvam, Dr E., and S. Gopi. "Effective and Efficient Traffic Scrutiny in Sweet Server with Data Privacy." International Journal on Applications in Information and Communication Engineering 5.2 (2019): 1-5
28. Punarselvam, E., and P. Suresh. "Investigation on human lumbar spine MRI image using finite element method and soft computing techniques." Cluster Computing 22, no. 6 (2019): 13591-13607.
29. Karthick, R., and M. Sundararajan. “A Reconfigurable Method for Time Correlated Mimo Channels with a Decision Feedback Receiver” International Journal of Applied Engineering Research 12.15 (2017): 5234-5241.

30. Karthick, R., and M. Sundararajan. “A novel 3-D-IC test architecture-a review” International Journal of Engineering and Technology (UAE) 7.1.1 (2018): 582-586.
31. Karthick, R., and M. Sundararajan.,”PSO based out-of-order (ooo) execution scheme for HTMPSOC” Journal of Advanced Research in Dynamical and Control Systems 9 (2017): 1969.
32. Punarselvam, Dr E., and S. Gopi. "Effective and Efficient Traffic Scrutiny in Sweet Server with Data Privacy." International Journal on Applications in Information and Communication Engineering 5.2 (2019): 1-5.
33. Punarselvam, E., et al. "Different loading condition and angle measurement of human lumbar spine MRI image using ANSYS." Journal of Ambient Intelligence and Humanized Computing 12.5
34. Punarselvam, E., and P. Suresh. "Non-Linear Filtering Technique Used for Testing the Human Lumbar Spine FEA Model." Journal of medical systems 43, no. 2 (2019): 1-13.
35. Punarselvam, E., Hemalatha, E., Dhivahar, J., Gowtham, V., Hari, V., & Thamaraikannan, R. PREDICTING WIRELESS CHANNELS FOR ULTRA-RELIABLE LOW-LATENCY COMMUNICATIONS.