International Advance Journal of Engineering, Science and Management (IAJESM)
ISSN -2393-8048, January-June 2021, Submitted in June 2021, iajesm2014@gmail.com

Efficient task scheduling in cloud computing using various machine learning techniques

Mayur Rahul[1], Himanshu Shukla[2], Akhilesh Singh[3], Namita Tiwari[4]
[123]Department of Computer Application, School of Engineering and Technology, CSJM University, Kanpur
[4]Department of Mathematics, School of Basic Sciences, CSJM University, Kanpur

Abstract: The task scheduling in cloud computing is very important concept in today's era of computers.The cloud computing deliver different services and applications to the end user. The efficient task scheduling makes the cloud computing environment more productive. In this paper, efficient task scheduling using various supervised learning methods like SVM and Random Forest is proposed. The objective of the paper is to find the resource load factor, shortest time and task completion cost, building a model to calculate the effectiveness of task scheduling and create the fair procedure of cloud task scheduling. The investigational results shows the enhanced task scheduling based on supervised learning can improve the performance of cloud computing environment.

Keywords: Cloud computing, SVM, Random Forest, Task scheduling,

1.Introduction: Information technologies have been considered a major and essential part of every organisation with the respect to management, time and cost. However, the industry based on information technology has consummate a vast change since last 2 or 3 decades (fig. 1). The factors such as open source, hardware, workforce globalisation, agile IT processes and virtualization makes good contribution for the development of business models and new technology.

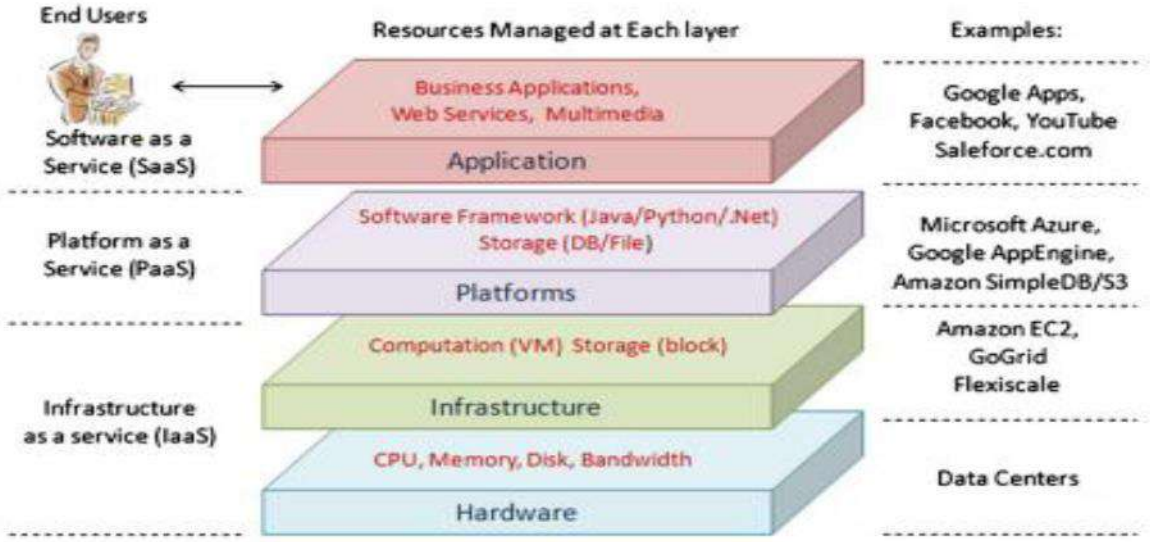

Figure 1: Cloud computing architecture

The cloud computing provides more choices to the organisation with respect to smooth running of the organisation, delegate liabilities and cost saving to providers of third party. The cloud computing becomes very important part of business models and new technology and forcibly apply business to use leading edge technology techniques.

Correspondingly, big need of cloud system has influence the evolution of global market contribution and able to represent various delivery models and cloud services. These models importantly increase the domain of available options and it is not easy to employ suitable model.

Cloud computing is based on internet which helps us to share various processes, data and resources to the computers and with the additional devices. Cloud system is one of the most challenging technologies which give numerous services like databases, computing, virtual machines, servers, storages, machine intelligence,analytics, and lot more. Cloud computing gives those services with the help of internet, makes it adaptable and therefore helps organisation to reduce expenditure of the hardware and software.

According to NIST [1], the definition of Cloud computing is "Amodel for enabling ubiquitous, convenient or demandnetwork access to a shared pool of configurable

computingresources that can be rapidly provisional and related withminimal management effort or service provider interaction."Rapid scalability and elasticity, Self-service, minimum costs, increased performance, broad network access, and reliability, productivity, service-oriented, resource pooling and utility-based pricing are some of well-known propertiesof cloud. Cloud computing environment gives these services to end user.

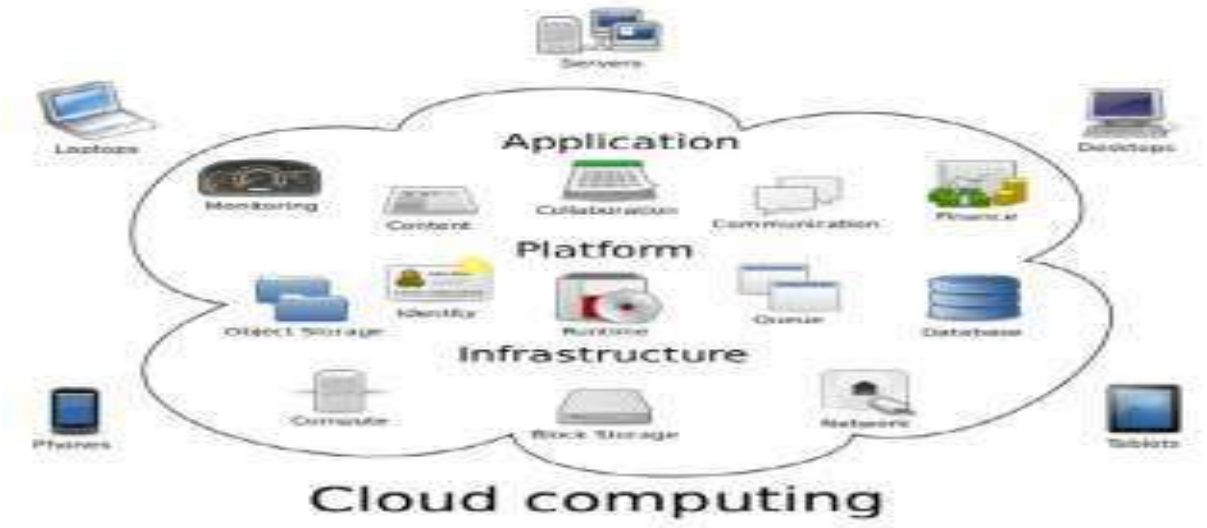

Figure 2: Cloud computing

Two types of model have been present in cloud computing environment: Deployment model and service models. They are classified according to the services given by cloud computing environment. The implementation are classified by whom and who use cloud services whereas service providers are categorizein accordance with the types of services given by cloud computing. Further, service models are categorize into three categories: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (PaaS), They have been recognised by NIST.

Many other service given by the cloud are Monitoring as a Service (MaaS), Data as a Service (DaaS) and Mobile Backend as a Service (MBaaS) Whereas deployment models are community cloud, private cloud, and hybrid cloud, including other deployment models forcloud such as Distributed cloud, Inter cloud etc.

Although, cloud computing is able to gives various services and focus on various applications. It is still very challenging task to perform efficient working of resource and task scheduling. In this paper, we introduced the task scheduling algorithms with the help of different machine learning methods such as Support Vector Machine (SVM) and Random Forest. The objectives of this paper are as follows:

(1)We have introduced an efficient task scheduling using machine learning methods such as SVM and Random Forest.

(2)We are able to compare our proposed algorithm with the existing methods in terms of the resource load factor, shortest time and cost of task completion

(3)We are also comparing existing algorithms in terms of SVM and Random Forest.

(4)The features of our proposed algorithm are also discussed in details.

The rest of this research paper is summarized as follows.Related work is discussed in Section 2. Section 3 describes the proposed algorithm. The experiments and results are explained in Section 4. Finally, the paper is concluded with further future works in Section 5.

2. Related Works: The cloud computing gives various services and applications to the end users. It stills some problem in resource and task scheduling efficiently. For the purpose of task scheduling, the cloud service stability, operating costs and resource utilization are the terms are very important part of the cloud related to service standard. Hence, the problem of task scheduling in cloud computing has very crucial realistic and theoretical definition [2].The cloud environment has many uncertain and dynamic elements found in resource and its load. The resource load is directly proportion to time change and resource requests are also

depends on the days, weeks and months [3]. The given factors are responsible for resource wastage and minimum performance of the system. The resources of the cloud are wasted if resource burden is very low and accomplishment of system is less, if the resource load is high.

In context of the above problem, relevant researchers in the given fields are motivated towards the efficient task scheduling in cloud computing environment. Jun-wei et al. introduced an optimized task scheduling algorithm based on ant colony in terms of cost, load balance and completion time [4]. Based on characteristics of cloud computing, there is a function exist both in load balance and task completion and also able to give essential things. Then, they have been able to improve the heuristic function and essential things for updating ant colony based optimization algorithm.They able to get the constraint function for ant algorithm, then achieving the further optimized result [5]. They perform cloudsim and able to perform ACO and Min-Min algorithms. They showed that their procedure is far better than above two procedures in terms of load balance, processing time and cost.

Zhao et al. proposed a multi-objective task scheduling algorithm based on Q-learning [6]. The task scheduling stage, they are able to get the better sequence of task using Q-learning. Throughout the distribution of virtual machine in the starting, a weighting process is used to calculate the completion time and node. The function is generated by minimum total cost and completion time[7]. They able to show that qmts method sorts all the task depends on Q-learning take lesser time then heft. Further, after optimal task scheduling algorithm, it is very efficient to reduce the completion time and cost.

Abualigah et al. proposed a multi-objective task scheduling algorithm based on antlion method[8]. In this method, multi-target nature was observed and makespan limit was used while asset was boosting. The given technique was updated by elite-based technique with the help of neighbourhood search method by reducing the capacity of misuse and doesn't try to get kept in nearby optima.

Gaith et al. introduced a framework for distributed computing based on trust-mindful scheduling method[9]. Their framework focuses on calculation of virtual machine trust level and assurance of undertakings.Ewing et al. proposed a framework for task scheduling based on Q-learning for energy-efficient cloud computing [10]. They were also able to manage issues of utilization for assignment and scheduling planning. They were able to show the advantage of M/M/S lining model for the enhanced energy effectiveness with increase in errand limited reaction time.

Chen et al. proposed an efficient framework for the multi-objective task scheduling on the basis metaheuristic whale optimization technique [11]. Their main focus is to improve the presentation of the given cloud environment for the given registered assets. They also used the high level methodology called enhanced whale optimization technique in task scheduling to improve the ability to search. This technique is very effective in smart usage on asset, including both enormous and little scope.

Li et al. introduced a framework for assignment planning problem in distributing computing. This framework was depends on the calculation of artificial bee colony [12]. They were able to describe the problem as mixture flowshop scheduling problem. The enhanced variant issue adjustable irritation structure was incorporated in the computation to tune the investigation and misuse capability. A refreshed and determination technique was applied to enhance the misuse interaction. They were also able to enhance abuse capabilities, a misuse administrator was incorporated [13]. Enhanced scout honey bee utilization for the different neighbourhood scan techniques for best source of food or the undesirable positioning and can enhanced the capability of introduced computations. The designing for solicitations in cloud assets in circular climate is a general NP hard problem.

To compute this problem, Khorsand et al. proposed for the efficient task scheduling using technique for order preference by similarity to ideal solution and best-worst methods [14]. They also able to create set of decision-making to distinguish assessment rule. After that they used best-worth method to apply for every significance load to every basis keep in mind that

quality has been switched remarkably. At the point, the model was used to quantify and assess the performance of each given options.

Yang et al. introduced a framework for task planning using game assumption for energy control in cloud environment [15,18,19]. The game assumption depends on the conveyance of energy and errands dispersion. They also able to create a framework for task scheduling in different hubs by numerical model to manage vast information. They also able to show how energy planning can be improve by calculation in distributed system[16,20,21,22].

Based on previous discussion, violation rate, completion time, and virtual machine utilization are the key factors which affect the performance of task scheduling in cloud computing environment. Without focus on these terms, it is very difficult to enhance the task scheduling enhancement in cloud environment. This research focus on enhancement of task scheduling in cloud computing environment using supervised learning such as SVM and Random Forest, justifies the result with the existing results.

3. Proposed Methodology:

For the enhancement of task scheduling, choose the minimum shortest time for customers, cost of completion and resource load balance as the objective of task scheduling using SVM and Random Forest [17]:

(1) Resource load balance degree: To calculate the load balance degree on VM is given by the formula:

$$\ddot{G}=\frac{\sqrt{\frac{\sum_{j=1}^{m}\left(VMF_j-\overline{VMF}_j\right)^2}{m}}}{n} \quad (1)$$

In which,

VMFj→ complete processing time of task on virtual machine

VMFj→ mean processing time of task on virtual machine.

(2) Cost of task completion

The formula for calculation cost of task completion is given:

$$\text{Cost}_{\text{VMj}}=\sum_{j=1}^{m} \quad sum(VM_j)\ X\ \left(V_{cpuj}+V_{ramj}+V_{bwj}\right) \quad (2)$$

The equation (2) fully reflects the cost of task completion on VM, which is linked with CPU, memory of VM and bandwidth performance.

(3)Shortest waiting time: The formula for the calculation of minimum waiting time is given by:

$$\underset{j=1}{\overset{m}{Max}}\sum_{i=1}^{sum(VM_j)} \quad time_{ij} \quad (3)$$

In which,

$time_{ij}$→ Processing time of task Ti on given VM_j.

where,

sum (VM_j) → Total number of task assigned on VMj.

The execution time of task is ratio of task length to processing speed of VMj, can be calculated as follows:

$$time_{ij} = \frac{TA_i}{TA_j} \quad (4)$$

Start from above three formula, the minimum processing time, completion cost, and resource load balance has to be calculated to assess the effect of task scheduling in cloud computing environment. Describe number of task in task set and then compute the execution time of entire task set in cloud computing. Depend on estimated completion time of the task set,time execution cost and load balance in VM environment. The task scheduling enhancement in cloud computing environment is based on supervised learning method.

The overall conclusion is to select minimum time that task has to wait, resource load balance degree, and the cost of completion is the objective of the task scheduling in cloud computing environment. The algorithm for the above task is given by:

1. START
2. ENCODE TASK ACCORDING TO PROBLEMS
3. INITIALIZE TASK SET ABD SET N=0
4. N=N+1
5. CALCULATE THE VARIABLE FROM ALL EQUATIONS AND UPDATE THE VARIABLES
6. UPDATE THE VARIABLES OF TASK SET
7. FOUND THE MAXIMUM NUMBER OF ITERATIONS IF YES GOTO 8 ELSE GOTO STEP 4
8. STOP

Figure 3:Algorithm for the proposed technique

4. Experiments and Results:

The simulation of algorithm on the given platform, comparison of our framework based on supervised learning is compared with the work of Zhao et al.[6]. This comparison is basically based in terms of deadline violation cost, VM resource utilization and increase time completion.

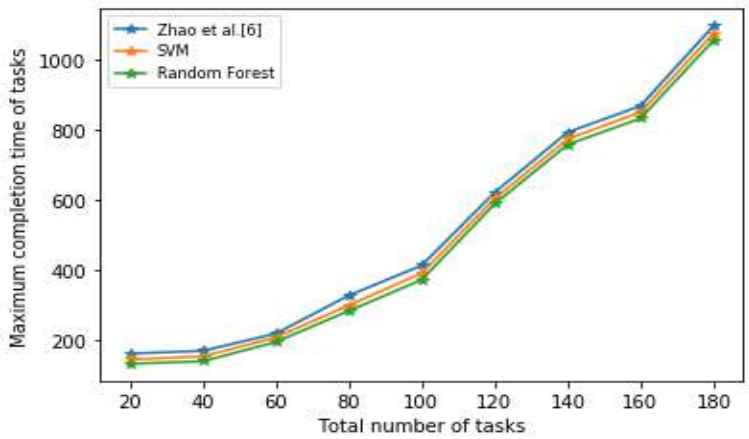

Figure 4: Comparison results of maximum completion time

The comparison of maximized completion time of various tasks gives the results by taking various multi objective tasks in cloud environment. The comparisons are given in figure 4.

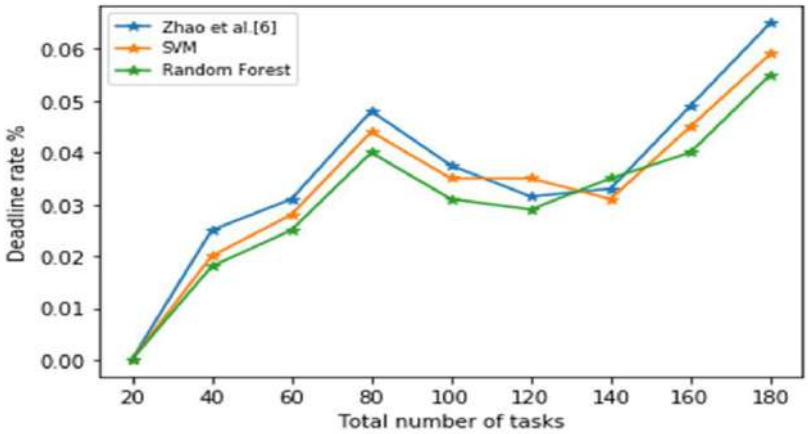

Figure 5: Comparison results of deadline rate

The figure shows that the inclusion of various MT results in longer completion time for all the three methods. The time taken by Zhao et al. is smaller than the SVM and Random Forestalso convergence rate of Zhao et al. is slower which makes the greater maximum completion time. The performance of our proposed work in terms of completion time is best after comparing with other given methods.

The comparisons of deadline violation rate of our proposed work with the SVM and Random Forest with the work of Zhao et al. is depicted in figure 5. The figure shows the increasing of tasks in cloud environment, our proposed method using SVM and Random Forest is able to maintain the deadline violation rate as compared to other methods.

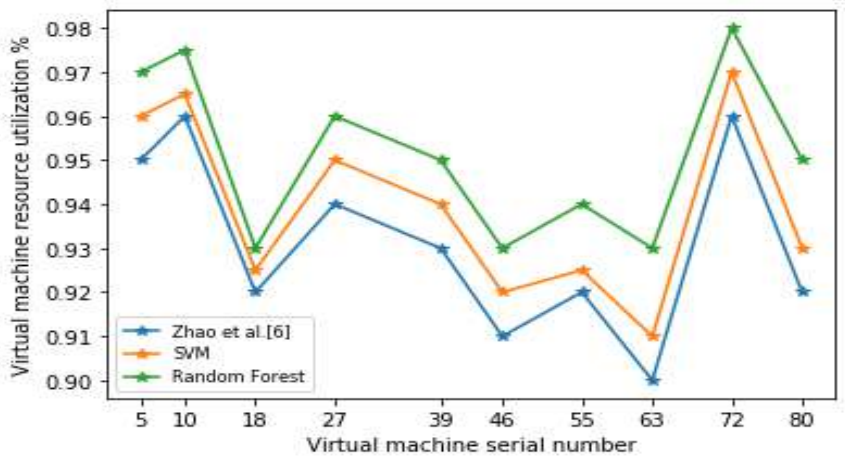

Figure 6: Comparison results of resource utilization

The quantity of multi-objective task is around 200, we have selected 12 task randomly and then calculate resource utilization factor of each method. The comparison between our proposed works with the existing work is shown in fig 6.

As depicted in figure 6, for the given quantity of tasks, results found from our proposed work with random forest are relatively good in terms of task scheduling. The figure shows that the rate of resource utilization in virtual machine using random forest is above 94%. Our proposed work using SVM use the concept of energy consumption and cost during task execution gets the rate of around 89% while Zhao et al uses the concept of deadline and energy consumption during task execution, get the rate around 84%.

4. Conclusion and future works:

The task scheduling based on SVM and random forest is proposed in this research. The main goal of our research is to enhance the performance of task scheduling. The objective function is created, which further used with SVM and Random Forest gives the results. The results shows the better performance of our proposed using Random forest work in terms of rate of deadline violation, improved completion time and enhanced virtual machine load utilization. The proposed work with Random forest is able to enhance the maximum scheduling waiting time and on increasing the number of tasks also lowers the violation deadline rate and resource allocation rate is very high. In future, we plan to add some more machine learning concepts to the task scheduling.

References:

[1] ChinthaguntaMukundha, K. Vidyamadhuri, “CloudComputing Models : A Survey” in Advances inComputational Sciences and Technology, ISSN0973-6107, Volume 10, Number 5, (2017).

[2] D.V. Kunhambu,. R V.S. Balan,” An efficient Multi-objective Particle Swarm Optimisation based Ranking system for cloud service selection”,IET communication,13(3),(2018),297-304.

[3] S.Z. Mirjalili. S. Mirjalili. S. Saremi. et a1, “Grasshopper Optimization Algorithm for multi-objective optimization problem”,Appl. Intelligence,48(4),(2018),805-820.

[4] Jun-weiGe, QiangGuo,YiQiu Fang, “ A Multi-objective Optimisation Algorithm for Cloud Computing Task Scheduling based on improved Ant Colony Algorithm”, Microelectronic Computing, 34(11), 2017, 63-67

[5] N. Tiwari, S. Padhye, Analysis on the generalization of Proxy Signature, Security and Communication Network, Wiley, 2013 Vol. 6, pp. 549-556

[6] Zhao Tong, Xiao-mei Deng, Hong-jianChan,et al.,” Multi-objective task scheduling algorithm based on reinforcement learning in cloud environment”, Journal of Chinese Computing system,41(02),(2020), 285-290

[7] N. Tiwari, S. Padhye, Provable secure proxy signature scheme without bilinear pairings, International Journal Of Communication Systems, Wiley 2013 vol. 26, pp. 644-650.

[8] LaithAbualigah, Ali Diabat, A novel hybrid antlion optimization algorithm for multi-objective task scheduling problems in cloud computing environments, Cluster Computing (2020) 1—19.
[9] RjoubGaith. Jamal Bentahar, Omar Abdel Wahab, BigTruslScheduling: Trust-aware big data task scheduling approach in cloud computing environments, Future Generation Computer Syst. 110 (2020) 1079 - 1097.
[10]Dirig Ewing, Xiaricr›ng Fri n, Yihiian Zhan, K uixuan K a rig, Qian Yin, ling Zeng, Q-learning based d ytiamic task scheduling for energy-efficient cloud computing, Future Generation Computer Systems 108 (2020) 361—371.
[11] Xuan Chen, Long Cheng, Cong Liu, Qingzhi Liu, Jiriwei Liu,Ying Mao, John M urphy, A woa-based ciptirnization approachfur task scheduling in cloud computing systems, IEEE Syst. J. 14 (3) (2020) 3117 - 3128.
[12] Jun-qing Li, Yun-qi Han, A hybrid rrtulti-objective artificial bee colony algorithm for flexible task scheduling problems in cloud computing system, Cluster Computing 23 (4) (2020) 2483 2499.
[13]N. Tiwari, S. Padhye and Debiao He, Efficient ID-based multi proxy multi-signature without bilinear maps in ROM, Annalas of Telecommunication, Springer, Vol.68 Issue 3-4, pp. 231-237 (2013)
[14] ReihanehKhorsand, MohammadrezaRamezanpour, An energy-efficient task-scheduling algorithm based on a multi- criteria decision-making method in cloud computing, Int. J.CzimmunSyst 33 (9) (2020) e4379.
[15] Jiachen Yang, Bin Jiamg, ZhihanLv, Kim-Kwang Raymond Choo, A task scheduling algorithm considering game theory designed for energy management in cluud computing, Future Generation computer systems 105 (2020) 985-992.
[16] N. Tiwari, S. Padhye, ECDLP-based certificateless proxy signature scheme with message recovery,Transactions On Emerging Telecommunications Technologies Trans. Emerging Tel. Tech. (2012), Published online in Wiley Online Library (wileyonlinelibrary.com), Vol 26(3), DOI: 10.1002/ett.2608
[17] D.Vazquez, M,J. Fernandez-Torres, R. Ruiz-Femenia, et al. “MILP method for objective reduction in multi-objective optimization”, Computers Chemical Engineering,108(4),(2018),382-394
[18] Rahul, Mayur&Kohli, Narendra&Agarwal, Rashi. (2021). Facial expression recognition using local multidirectional score pattern descriptor and modified hidden Markov model. International Journal of Advanced Intelligence Paradigms. 18. 538. 10.1504/IJAIP.2021.113787.
[19] Rahul, Mayur&Kohli, Narendra&Agarwal, Rashi& Mishra Tiwari, Sanju. (2019). Facial expression recognition using geometric features and modified hidden Markov model.International Journal of Grid and Utility Computing. 10. 488. 10.1504/IJGUC.2019.102018.
[20] Yadav, Vikash& Rahul, Mayur&Shukla, Rati. (2020). A New Improved Approach for Feature Generation and Selection in Multi- Relational Statistical Modelling using ML. Journal of Scientific and Industrial Research. 79. 1095-1100.
[21] N. Tiwari, S. Padhye, Analysis on the generalization of Proxy Signature, Security and Communication Network, Wiley, 2013 Vol. 6, pp. 549-556
[22] Tiwari, Namita&Padhye, Sahadeo. (2015). Provable secure multi-proxy signature scheme without bilinear maps. 17. 736-742.