IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 6, June 2022

CPU Burst Time using Machine Learning Algorithm's

Suresh K[1], Sirisha B[2], Divya D S[3], Soumya[4], Sahana Patil[5]

Assistant Professor, Department of Computer Science and Engineering[1]

Students, Department of Computer Science and Engineering[2,3,4,5]

Rao Bahadhur Y. Mahabaleshwarappa Engineering College, Bellary, Karnataka, India

Abstract: *The amount of time required by a process for executing on the CPU is called CPU-Burst Time, on which implementation of some CPU-scheduling algorithms such as Shortest Job First (SJF) and Shortest-Remaining Time First (SRTF) are relied on. There are many vivid methods for prognosticating the CPU-Burst length, one among them is Machine Learning based approach to calculate the burst- time of the processes in the ready queue. This approach involves the selection of most appropriate attributes of the process by making use of selection techniques, and based on these attributes, the CPU-Burst for the process is calculated in the grid. The ML algorithms used for this are Linear Regression, K-Nearest Neighbors (KNN), and Decision Tree (DT). This approach results in higher efficiency as it obtained linear relationship between the process and CPU Burst-Time. When tested and evaluated the workload dataset. The efficiency of ML approach is high because KNN provides better results in terms of CC and RAE, and for the factors like time, space and burst-estimation, attributes selection techniques are used.*

Keywords: CPU Burst Time

I. INTRODUCTION

Scheduling algorithms are used in Operating Systems (OS) to grant processes, threads, and data flow access to system resources. With today's contemporary systems and multiprogramming, the need to run multiple processes at the same time has led to the development of scheduling algorithms. To determine which of the processes in the ready queue should be given to the CPU, various CPU scheduling techniques have been proposed. Those scheduling techniques take different approaches to choosing a process from the ready queue and allocating it to the processor. CPU scheduling strategies such as First Come First Serve (FCFS), Shortest Job First (SJF), Shortest Remaining Time First (SRTF), Priority Scheduling (PS), Round- Robin (RR), and others have been proposed in the literature. To use such scheduling techniques, such as SJF scheduling, it is necessary to know the length of the CPU burst in addition to the process. SJF prefers processes with short burst lengths over those with longer burst lengths to ensure that the processes with short burst times finish and exit the system as quickly as feasible, allowing the larger processes to complete and exit the system as soon as possible. The time spent executing on the CPU from when it was scheduled to when it is removed from the run queue is measured as the CPU burst length for a process. The necessity to know the length of the CPU burst is the main issue with SJF scheduling. Knowing the length of the next CPU burst for a process is a difficult operation that necessitates additional work and calculations.

II. LITREATURE SRUVEY

1. "Comparison Analysis of CPU Scheduling: FCFS, SJF and Round Robin" by Andysah Putera Utama Siahaan Universitas Pembangunan Panca Budi Jl. Jend. Gatot Subroto Km. 20122, Medan, Sumatera Utara, Indonesia, 4,5 Sei Sikambing. And this paper Proposed The calculation of three algorithms shows the different average waiting time .The FCFS is better for a small burst time. If the process is submitted to the processor at the same time, the SJF is preferred. Round Robin, the final method, is superior for adjusting the required average waiting time. The SJF or FCFS value will increase as the round robin quantum time progresses. All algorithms are good, but the speed of the process is determined by the amount of processing power available.

"A Machine Learning-Based Approach to Estimate the CPU-Burst Time for Processes in the Computational Grids" by Tarek Helmy, Sadam Al- Azani, Omar Bin- Obaidellah Department of Information and Computer Science,College of

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 6, June 2022

Computer Science and Engineering, King Fahd

2. University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia. ML techniques such as K-NN, SVM, ANN, and DT were used in this paper to estimate the CPU burst length for processes in the ready queue.. In the proposed approach, the attributes feature selection techniques to select the most significant attributes were applied on a grid workload named "GWA-T-4 AuverGrid". The findings of the experiments reveal that there is a strong linear link between process attributes and burst CPU time, and that K-NN outperforms alternative ML approaches.

3. Fuzzy- Grounded Scheduling Algorithm for Prediction of Next CPU- Burst Time to apply Shortest Process Next Abdolghader Pourali Amir Masoud Rahmani, we will present time series generalities and an algorithm used to read the processes prosecution time which acts statistically. also we will explain fuzzy systems By using of intelligent systems similar as fuzzy systems, it's possible to estimate a lot of time series including the CPU- burst.

4. Comparison Analysis of CPU Scheduling FCFS, SJF and Round Robin Average Waiting Time is a standard measure for giving credit to the scheduling algorithm. The computation of three algorithms shows the different average waiting time.

5. A relative Study of Scheduling Algorithms for Multiprogramming in Real- Time systems Then we use Throughput, Response Time, Waiting Time, appearance Time to calculate the process The Performance And effectiveness Of AnyKind Of Scheduling Algorithm Is Measured.

III. METHODOLOGY

1. Reading Dataset: The Process time of Different processes is already calculated and is in the form of the datasets with different attributes, so first we need to collect the dataset of the CPU burst time of different processes.
2. Preprocessing: Preprocessing of dataset includes two steps Removing unwanted data and normalizing the data, removing unwanted data means deleting the unnecessary attributes like JobStructure, UsedNetwork, UsedLocalDiskSpace, UsedResources, ReqPlatform.
3. Algorithms used are 1. K-NN algorithm, Decision Tree algorithm & Linear Regression.
4. K-NN algorithm: The K-Nearest Neighbor algorithm is based on the supervised learning technique and is one of the most basic machine learning algorithms.

For prediction, the K-Nearest Neighbors model must first be developed, and the data must then be trained and tested.
Importing the package K- Neighbors classifier from sklearn.neighbors makes the K-Nearest Neighbors model work.
Predict the response for test dataset.
Importing modules to find the measure correlation value between the prognosticated and test figures to see how nearly they're related.

5. Decision Tree algorithm The Decision Tree algorithm is part of the supervised literacy algorithms family. Unlike other supervised literacy algorithms, the decision tree fashion can be used to attack retrogression and bracket problems. A decision tree cuts down a dataset into lower and lower subsets while contemporaneously developing an associated decision tree. A tree with decision bumps and splint bumps is the end result.

A tree has numerous circumlocutions in real life, and it turns out that it has told a lot of machine literacy, similar as bracket and retrogression. In decision analysis, a decision tree can be used to visually and explicitly depict opinions and decision- timber.

Decision trees help you in importing your options Decision trees are excellent tools for helping you choose between several possibilities. They give a largely effective frame within which you can produce a list of possibilities and weigh the pros and cons of each.

Linear Regression Linear retrogression analysis is a statistical fashion for prognosticating the value of one variable grounded on the value of another. The dependent variable is the bone you are trying to prognosticate. You will use the independent variable to read the value of the other variable.

Grounded on the input features from the data fed into the system, retrogression algorithms read the affair values. The standard way is for the algorithm to produce a model grounded on the characteristics of training data and also use the model to read the value of new data.

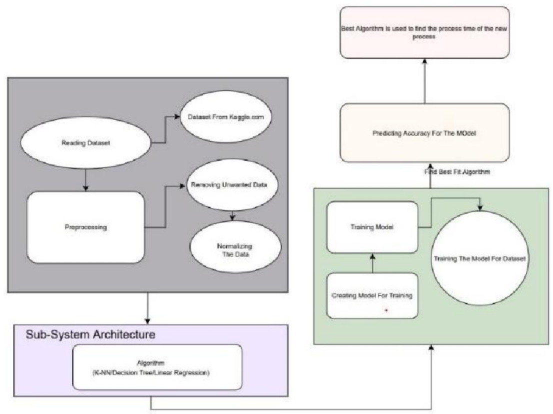
We can estimate the value of one variable using another variable using direct retrogression analysis.

6. For the three algorithms the model is created for training, also the created model is tested with respect to the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

dataset, and the delicacy and correlation of the three algorithm is set up.8. The best fit algorithm is found from the three algorithms by the result of correlation found after testing the models.

7. The best fit algorithm is used to find the burst time of the new processes using the attributes.



IV. RESULTS

After the successful testing of three algorithms we found the correlation of all the three algorithms, the correlation of the three algorithms are as follows.

- The correlation of k-NN algorithm is found to be 92%
- The correlation of Linear Regression algorithm is found to be 87%
- The correlation of the Decision Tree algorithm is found to be 94%

As the correlation of the Decision Tree algorithm is highest among the three algorithms with 94% correlation, so for the finding out the burst time of the new process the Decision Tree algorithm is found to be the best fit algorithm.

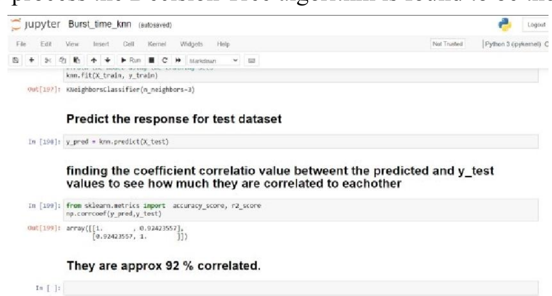


Figure 2: Correlation of K_NN algorithm

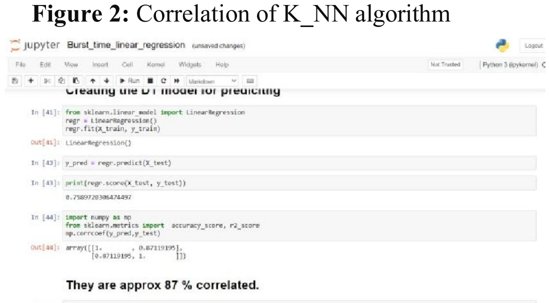


Figure 3: Correlation of Linear Regression algorithm

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 6, June 2022

Figure 4: Correlation of Decision Tree algorithm

V. CONCLUSION

The proposed method for calculating the cpu burst length for processes in the ready queue using several methodologies such as k-nearest neighbors, decision learning, linear regression. When compared to prior CPU scheduling methods such as SJF and SRTF, this method is more efficient.

FUTURE WORK

We will test the proposed approach to estimate the time quantum and load balancing in the CGs in future work.

REFERENCES

[1]. Silberschatz, Abraham, Peter B. Galvin, and Greg Gagne. Operating system concepts. Vol. 8. Wiley, 2013.

[2]. Mehmood Shah, Syed Nasir, Ahmad Kamil Bin Mahmood, and Alan Oxley. "Analysis and evaluation of grid scheduling algorithms using real workload traces." In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, pp. 234-239. ACM, 2010.

[3]. Fernández-Baca, David. "Allocating modules to processors in a distributed system." IEEE Transactions on 15, no. 11 (1989): 1427-1436.

[4]. Mahesh Kumar, M. R., B. Renuka Rajendra, C. K. Niranjan, and M. Sreenatha. "Prediction of length of the next CPU burst in SJF scheduling algorithm using dual simplex method." In *Current Trends in Engineering and Technology (ICCTET), 2014 2nd International Conference on*, pp. 248-252. IEEE, 2014.

[5]. Kasana, Harvir Singh, and Krishna Dev Kumar, eds. *Introductory operations research: theory and applications*. Springer Science & Business Media, 2004.

[6]. Pourali, Abdolghader, and Amir Masoud Rahmani. "A Fuzzy-based Scheduling Algorithm for prediction of next CPU- burst time to implement Shortest Process Next." In *Computer Science and Information Technology-Spring Conference, 2009, IACSITSC'09*, pp. 217-220, IEEE, 2009.

[7]. Kosko, Bart. *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*. Prentice-Hall, Inc., 1991.

[8]. Kohavi, R., and R. Kohavi. 1997. "Wrappers for Feature Subset Selection." Artificial Intelligence 97(1-2):273–324. Retrieved (http://linkinghub.elsevier.com/retrieve/pii/S0004370 29700043X).

[9]. Smith, Warren, Ian Foster, and Valerie Taylor. "Predicting application run times with historical information." *Journal of Parallel and Distributed Computing* 64, no. 9 (2004): 1007-1016.

[10]. Matsunaga, Andréa, and José AB Fortes. "On the use of machine learning to predict the time and resources consumed by applications." In *Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, pp. 495-504, IEEE Computer Society, 2010.

[11]. Tom Mitchell, *Machine Learning*, 1st Ed, pp: 52- 75, 154-183, 230-244, the Mc-Graw Hill Company. Inc. International Edition, 1997.

[12]. Shevade, S. K., S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy. 2000. "Improvements to the SMO Algorithm for SVM Regression." *IEEE Transactions on Neural Networks* 11(5):1188– 93.