



EFFECTIVE APPROACH VILOS JONE ALGORITHM COMPARE WITH LINEAR REGRESSION ALGORITHM IN SELF REGULATING ATTENDANCE SYSTEM

K Koushik Kishore¹, V. Karthick^{2*}

Article History: Received: 12.12.2022

Revised: 29.01.2023

Accepted: 15.03.2023

Abstract

Aim: To enhance efficient predictive analysis for self regulating attendance system using Viola jones algorithm compared with support Vector Machine

Materials and Methods: The study contains two groups i.e Viola algorithm is developed in the first group and Linear Regression Algorithm developed in the second group contains 104 samples. The sample size for Linear Regression is 52 and sampling technique is VJS (N=52) and G power (value=0.8)

Results: The performance has been improved in terms of accuracy for the viola boot algorithm with 79.80% while the has shown an accuracy of 86.10.%. The mean accuracy detection is $\pm 2SD$ and the significance value is 0.000 ($p < 0.05$) which shows the hypothesis is correct and it is carried out using an independent sample T test.

Conclusion: The final outcome of the 79.80% Viola Jones Algorithm is found to be significantly more accurate than the Linear Regression Algorithm 86.10%.

Keywords: Novel Effective framework, Attendance system, Viola Jones Algorithm, Linear Regression algorithm, Radio Frequency Identification.

¹Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.

^{2*}Department of Computer Science and Engineering, Saveetha School Of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 601205.

1. Introduction

Most organizations that operate on the basis of employee contributions, whether at the executive or operational levels, require full attendance before they can begin operations. As a result, the attendance system gives information that will assist them in planning if any students are absent. Attendance can be described as the act of being present at a specific location or event, such as being present at a party or at work. Almost all businesses keep track of their employees' attendance (Merwid 2021). There are several types of attendance systems that have been developed in earlier implementations, such as stamp cards or internet attendance systems. These implementations can nonetheless cause a slew of issues, such as making it impossible to maintain track of employee attendance and giving users inaccurate information (Dobrykh et al. 2022). The goal of this cloud-based radio frequency identification technology is to make the traditional method of recording attendance more efficient. Aside from that, it's used to create a real-time record. Despite this, these solutions can cause a plethora of problems, including making it impossible to keep track of staff attendance and providing users with erroneous information (Syafira 2017).

From the past five years there are about 750 articles from various sources such as Google Scholar, IEEE Xplore and Springer. The various techniques used are: To make the connection between the sensor and the raspberry pi work so that the data from the radio frequency identification sensor may be successfully sent (Kar, Saha, and Debbarma 2012). The final goal is to give the administrator internet access to the system (Shoewu and Olatinwo 2013). Time and attendance systems are used by businesses of all kinds to track when employees start and stop working, as well as the department where the work is done. In many workplaces, the punch card machine was commonly employed to store data about employee attendance (Yamalova, Churilov, and Medvedev 2015). However, it would be difficult to maintain and track employee attendance due to a lack of storage space and supply prices (Hagedorn et al. 2022).

Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; Kumar et al. 2022;

Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022). This paper has not fully involved in implementing security to automated attendance management system which is found to be the research gap. It will be advantageous to replace the punch card machine by using Radio Frequency Identification radio frequency identification, which is a form of wireless communication that employs radio waves to identify and track items. Radio frequency identification can be used to track employees' time attendance (Dobrykh et al. 2022). The technology makes it simple for the administrator to observe and track employee attendance, which is also saved in the cloud (Cadilhac et al. 2022). The system also has data recording capabilities and allows the administrator to group employees. Overall, this system is capable of managing daily attendance data via comprehensive reports stored in the cloud (Cadilhac et al. 2022). The aim of the novel effective framework is to detect attendance to alert the faculty from attendance through proxy.

2. Materials and Methods

This research paper was carried out in the Department of Artificial Intelligence Laboratory of Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences. This study involves 2 groups, group 1 is viola jones algorithm (84%) and group 2 is viola jones algorithm (91%). The total number of groups for this are two groups. Group one refers to the existing system, and Group two refers to the proposed system. The total number of Samples is 104, out of which 52 are the samples for the first group and the remaining 52 are used as samples for the second group. Size was calculated using previous study results in clinical.com by keeping threshold 0.05 G power 80%, confidence interval 95% and enrollment ratio as 1 calculated in clinical.com. This is an open source data repository (Golley, Jaivin, and Tomba 2017).

The dataset is used for observing both the algorithms (Datta, Datta, and Banerjee 2015). The dataset is checked and verified for any empty values. A minimum of 2GB ram is required to use and install all the necessary applications, and a minimum of i3 processor to run all the applications and processes simultaneously. A minimum of 50GB hard disk space is required to install the required software and files, and an internet connection is required to download and install all the necessary software and development environment to run this Novel Effective framework. Python programming language is used for the application (Zhang et al. 2009).. The version

of python used is 3.7, and the IDLE is used to run and execute the application.

Linear Regression Algorithm

Linear regression is a supervised learning machine learning algorithm. It carries out a regression task. Based on independent variables, regression models a goal prediction value. It is mostly used in forecasting and determining the link between variables. Different regression models differ in terms of the type of relationship they evaluate between dependent and independent variables and the amount of independent variables they use.

step1 : Start

Step 2: Navigate to the path of the dataset path.

Step 3: Determine the worth of each feature.

Step 4: Use linear regression to solve the problem.

Step 5 is where the dominant players are separated.

Step 6: Look for null values.

Step 7: Retrieving data from null values

Step 8: Stop

Viola jones Algorithm

The Viola-Jones face identification procedure uses simple feature values to classify photos. There are numerous advantages to using features rather than pixels. The most common reason for this is that features can be used to encode ad hoc domain knowledge, which is difficult to acquire with a little amount of training data. Another incentive to use features is that they are significantly faster than pixel-based systems because they are based on the operating system.

Step 1: Initialize the guess of the algorithm to predict the average value of the target.

Step 2: Choosing Haar-like attributes

Step 3: Construct an integral image

Step 4: Predict the pseudo residuals for further analysis

Step 5: Execute Gradient Boost training

Step 6: Make a second prediction and build the prediction using the second tree.

Step 7: The final prediction is equal to the average value initially computed plus the predicted

residual value at step one plus the predicted residual value at step two.

Statistical Analysis

The data for biometrics detection was gathered from a cloud, and also the system was tested with over 50 participants. The statistical software used for implementation in IBM SPSS version 21. Statistical package for the social sciences version 26 software tool was used for statistical analysis. An independent sample T-test was conducted for accuracy. The dependent variable for this test is the

image size and the viewing angle. Accuracy and number of objects in an image are the independent variables in this test. Standard deviation, standard mean errors were also calculated using the SPSS software tool. The significance values of proposed and existing algorithms contains group statistical values of proposed and existing algorithms.

3. Results

The comparative analysis of the existing and the proposed algorithms is carried out by taking the accuracy rate of detection for both the algorithms using the Novel Effective framework. The face detection accuracy for both the algorithms is taken and can be used for analysis and comparison. The proposed algorithm is said to be more accurate than the existing algorithm, and the accuracy rate of the proposed algorithm is found to be 79.80%. This can be observed in Table 1 that the dataset and various attributes that the dataset contains are used for input into the algorithm.

Table 2 the dataset is the independent and dependent variable. Table 2 the dependent variable and the geographical places of the incident are independent variables.

Table 3 shows the statistical mean and Group statistics of sample T-test for the existing algorithm and viola jones algorithm.

Table 4 shows the standard deviation and mean error of the self regulating attendance system and the correlation of the viola jones and the Linear Regression Algorithm.

Figure 1 represents the results are used as input into the statistical analysis tool and the graph is plotted using the values.

4. Discussion

The accuracy obtained from viola jones algorithm is 91% but LR resulted with an accuracy of 84%. Hence it is clear that viola jones algorithm is better than the LR. The values of the Novel Effective framework are analyzed statistically and the difference is found out by plotting the graph against the algorithms.

Opposite findings of this article are not found. Similar findings related to Viola Jones algorithm is significantly efficient in predicting the Image Enhancement of the user compared to the existing algorithm, (Miyashita, Sawahata, and Komine 2022)that is the Linear Regression Algorithm(Jingyuan et al. 2008). The dataset containing a large number of images is given as

input into both the algorithms, and the accuracy rate of prediction is obtained for the existing and the proposed algorithms (Rajput et al. 2022). These values obtained are used for analysis and comparison for efficiency.

The limitation in this model is that the accuracy of viola jones algorithm may get affected due to the inconsistent data and difficulty in getting the right datasets for analysis. Most of the data is simulated from nature which is far from reality in the Novel Effective framework. Effective data preprocessing techniques and the combination of viola jones with other machine learning algorithms such as LR and viola jones algorithm may give better accurate results in the future. In the future, real time data and instead of using the application for samples, the application can be developed by running the system through its tests with different biometrics. The disadvantage is that real-time applications require a large team. The sensor damage cannot be identified, which is one of the problems that affects this system (Miyashita, Sawahata, and Komine 2022).

5. Conclusion

The research study found that the proposed novel viola jones algorithm is significantly efficient and accurate compared to the Linear Regression Algorithm. The accuracy of the prediction of the proposed algorithm is found to be 86.10%, and hence using the proposed algorithm gives better results compared to the existing algorithm that

Declarations

Conflicts of Interest

No conflict of interest in this manuscript

Author Contribution

Author KKK is involved in data collection, data analysis and manuscript writing. Author VK was involved in conceptualization, data validation and critical review of the manuscript.

Acknowledgements

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

Thankful for the following organizations for providing financial support that enabled us to complete the study.

1. Vee Eee Technologies Solution Pvt.Ltd.
2. Saveetha University

3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering.

6. References

- Cadilhac, Dominique A., Joosup Kim, Geoffrey Cloud, Craig S. Anderson, Emma K. Tod, Sibilah J. Breen, Steven Faux, et al. 2022. "Effect of the Coronavirus Disease 2019 Pandemic on the Quality of Stroke Care in Stroke Units and Alternative Wards: A National Comparative Analysis." *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association* 24 (1): 79–87.
- Datta, Asit Kumar, Madhura Datta, and Pradipta Kumar Banerjee. 2015. *Face Detection and Recognition: Theory and Practice*. CRC Press.
- Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.
- Dobrykh, Dmitry, Ildar Yusupov, Pavel Ginzburg, Alexey Slobozhanyuk, and Dmitry Filonov. 2022. "Self-Aligning Roly-Poly RFID Tag." *Scientific Reports* 12 (1): 2140.
- Golley, Jane, Linda Jaivin, and Luigi Tomba. 2017. *Control: China Story Yearbook 2016*. ANU Press.
- Hagedorn, Peter H., Jeffrey M. Brown, Amy Easton, Maria Pierdomenico, Kelli Jones, Richard E. Olson, Stephen E. Mercer, et al. 2022. "Acute Neurotoxicity of Antisense Oligonucleotides After Intracerebroventricular Injection Into Mouse Brain Can Be Predicted from Sequence Features." *Nucleic Acid Therapeutics*, February. <https://doi.org/10.1089/nat.2021.0071>.
- Jingyuan, Tang, Shi Yibing, Zhou Longfu, and Zhang Wei. 2008. "Analog Circuit Fault Diagnosis Using Ada Boost and SVM." 2008 International Conference on Communications, Circuits and Systems. <https://doi.org/10.1109/icccas.2008.4657978>.
- Kar, Nirmalya, Ashim Saha, and Mrinal Kanti Debbarma. 2012. *Biometric Based Attendance System: Your Face Is Your Identity*. LAP Lambert Academic Publishing.
- Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*.

- <https://doi.org/10.1016/j.fuel.2022.123594>.
- Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj, and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*.
<https://doi.org/10.1016/j.jwpe.2022.102712>.
- Merwid, Mary H. 2021. *Employees Attendance Logbook: Work Tracker Log , Attendance Register Journal , Organizer Notebook to Monitor Productivity , 120 Pages , 8. 5 X 11 In.*
- Miyashita, Yamato, Yasuhito Sawahata, and Kazuteru Komine. 2022. "Perceptual Assessment of Image and Depth Quality of Dynamically Depth-Compressed Scene for Automultiscopic 3D Display." *IEEE Transactions on Visualization and Computer Graphics* PP (February).
<https://doi.org/10.1109/TVCG.2022.3148419>.
- Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K₂TiF₆ and KBF₄ to Develop Nanoparticles Based TiB₂ Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.
- Rajput, Shyam Singh, Nafis Uddin Khan, Amit Kumar Singh, and Karm Veer Arya. 2022. *Digital Image Enhancement and Reconstruction*. Academic Press.
- Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022. "Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable Energy Technologies and Assessments*.
<https://doi.org/10.1016/j.seta.2022.102102>.
- Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.
- Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*.
<https://doi.org/10.1016/j.fuel.2022.123798>.
- Shoewu, O., and Segun O. Olatinwo. 2013. *Development of a Biometrics-Based Student Attendance System*. LAP Lambert Academic Publishing.
- Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.
- Sundaraman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*.
<https://doi.org/10.1016/j.scitotenv.2022.153897>.
- Syafira, Adinda Rizkita. 2017. "Sistem Deteksi Wajah Dengan Modifikasi Metode Viola Jones." *Emitor: Jurnal Teknik Elektro*.
<https://doi.org/10.23917/emitor.v17i1.5964>.
- Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*.
<https://doi.org/10.1002/er.7780>.
- Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.
- Yamalova, Diana, Alexander Churilov, and Alexander Medvedev. 2015. "Design Degrees of Freedom in a Hybrid Observer for a Continuous Plant under an Intrinsic Pulse-Modulated Feedback**A. Medvedev and D. Yamalova Are in Part Financed by the European Research Council, Advanced Grant 247035 (SysTEAM) and Grant 2012-3153 from the Swedish Research Council. D. Yamalova and A. Churilov Were Partly Supported by Saint Petersburg State University, Grant 6.38.230.2015. A. Churilov Was Partly Financed by the Russian Foundation for Basic Research, Grant 14-01-00107-A." *IFAC-PapersOnLine*.
<https://doi.org/10.1016/j.ifacol.2015.09.337>.
- Zhang, Taiping, Yuan Yan Tang, Bin Fang, Zhaowei Shang, and Xiaoyu Liu. 2009. "Face Recognition Under Varying Illumination

Using Gradientfaces.” IEEE Transactions on Image Processing. <https://doi.org/10.1109/tip.2009.2028255>.
Narayanasamy, S., Sundaram, V., Sundaram, T., & Vo, D. V. N. (2022). Biosorptive ascendency

of plant based biosorbents in removing hexavalent chromium from aqueous solutions– Insights into isotherm and kinetic studies. Environmental Research, 210, 112902.

Tables and Figures

Table 1. Iteration for the accuracy of Viola Jones algorithm and LR . The accuracy of method 1 is 79.80% and method 2 is 86.10%.

Iteration	Viola jones algorithm	LR
1	79.80	86.10
2	79.85	86.15
3	79.88	86.20
4	79.93	86.25
5	79.96	86.28
6	79.97	86.32
7	79.98	86.38
8	79.99	86.42
9	79.99	86.48
10	80.00	86.53

Tabel 2. Statistical Analysis of Mean, Standard Deviation, and Standard Error of accuracy of Viola Jones algorithm method and Linear Regression algorithm. There is a statistically significant difference in accuracy between the methods. The Viola Jones algorithm method has the highest accuracy (79.8%) and the Linear Regression algorithm has the least accuracy (86.1%).

Pair 1	N	Mean	Std. deviation	Std.Error Mean
VJA	10	79.9350	.06884	.2177
LR	10	86.3110	.14169	.4481

Table 3.The correlation of the existing and the proposed algorithm with p-value where N=52and significance value of .722 from the statistical analysis tool.

Pair 1	N	Correlation	Significance value
Vja and gradient boosting algorithm	20	.966	.722

Tabel 4: Comparison of Significance Level with value $p < 0.05$. Both Viola jones algorithm and decision tree have a confidence interval of 86.1% with the significance level of accuracy is < 0.05

		Levene's test for equality of variances.		T- test for equality of means						
		F	Sig.	t	df	Sig.(2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									Lower	Upper
ACCURACY	Equal variance assumed	1.36	.722	145.866	8	.000	6.37600	.04982	6.2714	6.48066
	Equal variances			145.866	7.849	.000	6.37600	.49820	6.26840	6.48360

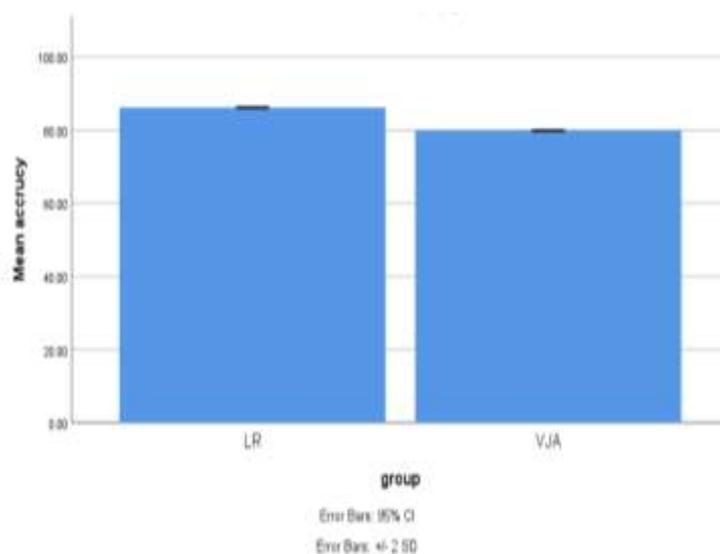


Fig. 1. Bar chart representation of the comparison of mean accuracy of the proposed and the existing algorithm. The accuracy of the prediction of the proposed algorithm is found to be 86.10% and the proposed algorithm gives better results compared to the existing algorithm that has accuracy of 79.10% the mean accuracy detection is ± 2 SD.