ISSN: 0974-5823 Vol. 7 (Special Issue, Jan.-Mar. 2022)

International Journal of Mechanical Engineering

Predicting Online Game Disorder Caused for Students using K-Nearest Neighbor Algorithm Compared with Support Vector Machine Algorithm to Improve Accuracy

V. Sunil[1], T. Rajesh Kumar[2]

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

ABSTRACT

Aim: Novel prediction of the online game disorder level of the students using K-Nearest Neighbors algorithm compared with the Support Vector Machine approach with improved accuracy. **Materials and Methods**: The online game disorder level of the student's analysis is done by machine learning algorithms such as K-Nearest Neighbors and Support Vector Machine approaches. Here pretest power analysis was carried out at 80% and the sample size for one group is 20 and the total sample size is 40, to attain the significance value is 0.001. **Result:** Implemented a novel prediction of the online games disorders for students using K-Nearest Neighbors algorithm and Support Vector Machine algorithm with the accuracy of 96.10% and 96.15% respectively. There is a statistical 2-tailed significant difference in accuracy for two algorithms is 0.001 ($p<0.05$) by performing independent samples t-tests. **Conclusion:** The disorder among students is predicted due to online games using the Support Vector Machine algorithm appears to produce better accuracy compared with the K-Nearest Neighbors algorithm. SPSS analysis helps to depict the reliability of data where time is a dependent variable and size is an independent variable.

Keywords: Online game, Disorder level, Novel Prediction, Machine learning, K-Nearest Neighbors Algorithm, Support Vector Machine Algorithm

INTRODUCTION

Nowadays online gaming is very familiar. The world has pushed to the situation that no youngsters are without using cell phones and in which they are always playing online games. Some of the internet gaming disorders include preoccupation with gaming. They spend more time on this type of gaming. This problematic condition is defined as the competitive utilization of video games which results in useful impairment to an individual capacity for function in many life domains over a prolonged period of duration (Zampelas and Micha 2015; Kim and Kim 2010). These disorders can be analyzed when an individual engages in gaming activity at the cost of satisfying the daily duties.Previously our team has a rich experience in working on various research projects across multiple disciplines(Ezhilarasan et al. 2021; Balachandar et al. 2020; Muthukrishnan et al. 2020;

Kavarthapu and Gurumoorthy 2021; Sarode et al. 2021; Hannah R et al. 2021; Sekar, Nallaswamy, and Lakshmanan 2020; Appavu et al. 2021; Menon et al. 2020; Gopalakrishnan et al. 2020; Arun Prakash et al. 2020)

In existing developed systems there is a lack of prediction of students' online gaming disorder performance (Lai 2021). The student's online gaming disorder performance analysis and novel prediction can be done by using machine learning algorithms such as Support Vector Machine to perform significantly better than the K-Nearest Neighbors Algorithm.

MATERIALS AND METHODS

This investigation was conducted in the Institute of Computer science and Engineering section of Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences. In this current research work, two groups were considered, one group refers to K-Nearest Neighbors and the other group refers to the Support Vector Machine analysis process. Pretest power analysis was carried out at 80% and sample size for one group is 20 and the total sample size is 40, the threshold value is 0.001(Chimidt and Filho 2020). SPSS analysis is carried out with the level of significance value $p<0.05$. Using G*power tool version 3.1, sample size for each group (RBD and LSD) is 62 with actual power 0.9234. Mean and standard deviation has been taken from previous literature for sample size calculation.

The Student's online gaming details datasets are downloaded from the Kaggle website. The dataset consists of 250 rows and 6 columns. Dataset is the combination of degree of study time for goal object materials(STG), the degree of repetition number of users for goal object materials(SCG), the degree of study time of the user for related objects with goal objects(STR), the online activates of a user for connected objects with goal objects(LPR), the online gaming performance of the student for goal objects(PEG), the knowledge level of student(UNS).

K-Nearest Neighbors Algorithm

K-Nearest Neighbors is a kind of supervised type learning approach which is applied for classification-based problems on labeled type data. It executes based on guessing each data point value near to another one that comes under the similar (Salmani, Desale, and Anthony 2021). This work K-Nearest Neighbors approach is applied for forecasting an individual's stage of anxiety level and depression level based on the classification outcome.

Step 1: Calculate "$d(x, x_i)$" i =1, 2, ….., **n**; where **d** denotes the Euclidean distance between the points.

Step 2: Arrange the calculated **n** Euclidean distances in non-decreasing order.

Step 3: Let **k** be a +ve integer, take the first **k** distances from this sorted list.

Step 4: Find those **k**-points corresponding to these **k**-distances.

Step 5: Let $\mathbf{k}_i$ denotes the number of points belonging to the i^{th} class among **k** points i.e. $k \geq 0$

Step6: If $k_i > k_j \; \forall \; i \neq j$ then put x in class i.

Support Vector Machine Algorithm

Support Vector Machine algorithm is mainly in binary type classification due to its unique idea to identify hyperplanes to categorize two classes with the highest margin. The major benefit of the Support Vector Machine approach is its categorization nature to provide the solution for nonlinear type problems with the kernel method (Wei 2012).

Step 1: candidateSV = { closest pair from opposite classes }

Step 2: while there are violating points do

Find a violator

candidateSV = candidateSV S violator

Step 3: if any $\alpha p < 0$ due to addition of c to S then

candidateSV = candidateSV \ p

repeat till all such points are pruned

end if

end while

RBF method is used to decrease the hyperparameter computation. In this work, the following kernel method is used (Di et al. 2019).

$$x(x_i, x_j) = exp\ exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) - - - - - (1)$$

In equation (1) σ indicates the RBF function manipulated (Howlett and Jain 2013) by the concerned user.

The software tool used to assess the performance of the Support Vector Machine and K-Nearest Neighbors was in google colab and Jupiter notebook platform with python programming language. The hardware setup was Intel core i5 processor type with a RAM size of 8GB. The 64-bit system, OS, X64 based processor with an HDD of 917 GB is used to develop the model. The software setup includes the windows 10 OS (Operating system). The designed model was trained by a training dataset before testing the developed model. Once the model is getting trained, test the developed model using a test dataset. In this work training error is below the limit value then cancels the training process. Now given value from the test dataset can be given into a trained model to check the outcome. Once the learning convergence is reached the level then the result of testing is also verified.

Statistical Analysis

The analysis was completed using IBM SPSS tool version 21. It is one of the statistical software tools used for analyzing data. Independent variables project, team_exp, year_end etc and dependent variables are id, length, effort. For both proposed and previous algorithms, 10 iterations were completed with a maximum of 10-20 numbers of samples and for every iteration, the predicted exactness was noted for analyzing accuracy rate. SPSS analysis is carried out with the level of significance value $p<0.05$. With value obtained from the iterations, an Independent type Sample T-test was executed

RESULTS

In Table 1, Statistical Analysis of Mean, Standard Deviation and Standard Error of Accuracy of Support Vector Machine (96.15%) and K-Nearest Neighbors (96.10%) algorithms were represented. It was observed that for 10 samples in Support Vector Machine Algorithm mean is 86.565 with Standard deviations of 7.70556, which is significantly better than the K-Nearest Neighbors algorithm means of 85.7550 and Standard deviation of 7.81631.

Table 2 represents the Levene's Test for equality of variances and t-test for Equality of Means for Support vector machines and K-Nearest Neighbors algorithm. The 2-tailed significance value smaller than 0.002 ($p<0.05$) showed that our hypothesis holds good.

Figure 1 represents the mean accuracy of Support Vector Machine (96.15%)and K-Nearest Neighbors (96.10%) and Support Vector Machine performs significantly better than K-Nearest Neighbors algorithm($p=0.001$) with Mean Accuracy of detection $\pm$ 2 SD and 95% CI are specified.

DISCUSSION

Based on the Kim et al.'s study it is to create a theoretical structure to supply realistic assistance to avoid and handle adult online gaming habits by designing a forecasting model via broad analysis of concern features (Kim and Kim 2010). In that they have collected data from 1,318 school students and analyzed them using the SPSS tool. Forecasting online game addiction using DT(Decision Tree) approach. In their research work, they have considered various factors like gender, school type, economical status, gaming location, amount of time spent for gaming, internet usage, occupation of the parents, supervising level, etc. The outcome of this research

recommends preventive measures for concerned groups in a regular way.Many current stories are evidence of the online gaming disorder level. Problematic type gaming prediction is a very critical process. Wu et al tried to forecast the online game playing time in the future. The outcome of this research work is used to assist people to manage a better balance between real-world activities and their virtual world. To forecast the playing time in the future here the authors change the game dataset into time-related data with the condition of problematic gaming. DL (Deep Learning) approaches are providing the solution for multiclass categorization issues (Wu and Carette 2020).

MOBA (Multiplayer Online Battle Arena) is one of the most familiar video games throughout the world. Many existing studies show the unnecessary commitment in online games can direct to IGD (Internet Gaming Disorder). These kinds of problems are directly interlinked with psychological issues like anxiety, impulsivity, etc. (Swati Aggarwal et al., 2019) propose a method to identify the self-esteem level of PUBG players, identify whether the people are suffering from internet gaming or not, and predict the psychological issues. Here the authors extract PUBG players' datasets from Asian Countries and use ML models to forecast the psychological levels. An entry-level experiment outcome demonstrates the prediction of disorder accuracy rate with 93.18%. The outcome of this research indicates a strong association is there between the PUBG game and the psychological disorder level (Swati Aggarwal et al., 2019).

The limitations of this current work are that the time complication was considerably more compared to an ideal state. This can be increased on future analysis and generality of the data which can significantly decrease process completion time. Many features can be considered for the Support Vector Machine in the future so that it can work proficiently and improve forecasting accuracy. Attributes like SCG, STG, LPR can increase the accuracy rate.

CONCLUSION

From these experiments, it is concluded that the Support Vector Machine algorithm appears to be better than the K-Nearest Neighbors algorithm. The online gaming student's disorder analysis was carried out by utilizing machine learning algorithms such as Support Vector Machine to perform significantly better than the K-Nearest Neighbors algorithm. Support Vector Machine proved with better accuracy of 96.15% than the K-Nearest Neighbors algorithm accuracy of 96.10%.

DECLARATIONS

Conflicts of Interest

No conflicts of interest in this manuscript.

Authors Contribution

Author VS was involved in data collection, data analysis, manuscript writing. Author TRK was involved in conceptualization, data validation and critical review of manuscript.

Acknowledgement

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:

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

1.Biozone Pvt. Ltd.Chennai.

2.Saveetha University

3.Saveetha Institute of Medical and Technical Sciences.

4.Saveetha School of Engineering.

REFERENCES

1. Appavu, Prabhu, Venkata Ramanan M, Jayaprabakar Jayaraman, and Harish Venu. 2021. "NOx Emission Reduction Techniques in Biodiesel-Fuelled CI Engine: A Review." *Australian Journal of Mechanical Engineering* 19 (2): 210–20.
2. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." *Biomass Conversion and Biorefinery*, August. https://doi.org/10.1007/s13399-020-00938-0.
3. Balachandar, Ramalingam, Logalakshmanan Baskaran, Ananthanarayanan Yuvaraj, Ramasundaram Thangaraj, Ramasamy Subbaiya, Balasubramani Ravindran, Soon Woong Chang, and Natchimuthu Karmegam. 2020. "Enriched Pressmud Vermicompost Production with Green Manure Plants Using Eudrilus Eugeniae." *Bioresource Technology* 299 (March): 122578.
4. Chimidt, Rafael Fabri, and Ernee Kozyreff Filho. 2020. "Análise Exploratória de Uma Década de Competições No Kaggle." https://doi.org/10.14488/enegep2020_tn_sto_344_1771_41218.
5. Di, Zonglin, Xiaoliang Gong, Jingyu Shi, Hosameldin O. A. Ahmed, and Asoke K. Nandi. 2019. "Internet Addiction Disorder Detection of Chinese College Students Using Several Personality Questionnaire Data and Support Vector Machine." *Addictive Behaviors Reports* 10 (December): 100200.
6. Ezhilarasan, Devaraj, Thangavelu Lakshmi, Manoharan Subha, Veeraiyan Deepak Nallasamy, and Subramanian Raghunandhakumar. 2021. "The Ambiguous Role of Sirtuins in Head and Neck Squamous Cell Carcinoma." *Oral Diseases*, February. https://doi.org/10.1111/odi.13798.
7. Gopalakrishnan, R., V. M. Sountharajan, A. Mohan, and M. Tholkapiyan. 2020. "The Strength and Durability of Fly Ash and Quarry Dust Light Weight Foam Concrete." *Materials Today: Proceedings* 22 (January): 1117–24.
8. Hannah R, Pratibha Ramani, WM Tilakaratne, Gheena Sukumaran, Abilasha Ramasubramanian, and Reshma Poothakulath Krishnan. 2021. "Author Response for 'Critical Appraisal of Different Triggering Pathways for the Pathobiology of Pemphigus vulgaris—A Review.'" Wiley. https://doi.org/10.1111/odi.13937/v2/response1.
9. Howlett, Robert J., and Lakhmi C. Jain. 2013. *Radial Basis Function Networks 2: New Advances in Design*. Physica.
10. Kavarthapu, Avinash, and Kaarthikeyan Gurumoorthy. 2021. "Linking Chronic Periodontitis and Oral Cancer: A Review." *Oral Oncology*, June, 105375.
11. Kim, Ki Sook, and Kyung Hee Kim. 2010. "[A prediction model for internet game addiction in adolescents: using a decision tree analysis]." *Journal of Korean Academy of Nursing* 40 (3): 378–88.
12. Lai, Carlo. 2021. "Review for 'Correlational Study between Online Friendship Network and Internet Game Disorder among University Students.'" https://doi.org/10.1002/brb3.2392/v1/review1.
13. Menon, Soumya, Happy Agarwal, S. Rajeshkumar, P. Jacquline Rosy, and Venkat Kumar Shanmugam. 2020. "Investigating the Antimicrobial Activities of the Biosynthesized Selenium Nanoparticles and Its Statistical Analysis." *BioNanoScience* 10 (1): 122–35.
14. Muthukrishnan, Sivaprakash, Haribabu Krishnaswamy, Sathish Thanikodi, Dinesh Sundaresan, and Vijayan Venkatraman. 2020. "Support Vector Machine for Modelling and Simulation of Heat Exchangers." *Thermal Science* 24 (1 Part B): 499–503.
15. Salmani, Sakina, Purvesh Desale, and Frank Anthony. 2021. "Analysis of Anxiety and Depression in Gaming Individuals." *2021 International Conference on Communication Information and Computing Technology (ICCICT)*. https://doi.org/10.1109/iccict50803.2021.9510061.
16. Sarode, Sachin C., Shailesh Gondivkar, Gargi S. Sarode, Amol Gadbail, and Monal Yuwanati. 2021. "Hybrid Oral Potentially Malignant Disorder: A Neglected Fact in Oral Submucous Fibrosis." *Oral Oncology*, June, 105390.
17. Sekar, Durairaj, Deepak Nallaswamy, and Ganesh Lakshmanan. 2020. "Decoding the Functional Role of Long Noncoding RNAs (lncRNAs) in Hypertension Progression." *Hypertension Research: Official Journal of the Japanese Society of Hypertension*.
18. Wei, Zhongyuan. 2012. *A SVM Approach in Forecasting the Moving Direction of Chinese Stock Indices*.
19. Wu, Qirui, and Jacques Carette. 2020. "Can Deep Learning Predict Problematic Gaming?" *2020 IEEE*

Conference on Games (CoG). https://doi.org/10.1109/cog47356.2020.9231887.
20. Zampelas, Antonis, and Renata Micha. 2015. *Antioxidants in Health and Disease*. CRC Press.

TABLES AND FIGURES

Table 1. Represents Statistical Analysis of Mean, Standard Deviation and Standard Error of Accuracy of Support Vector Machine (96.15%) and K-Nearest Neighbors (96.10%) algorithms.

Group	N	Mean	Std. Deviation	Std. Error Mean
Accuracy				
K-Nearest Neighbors	10	85.7550	7.81631	2.47173
Support Vector Machine	10	86.5650	7.70556	2.43671
Loss K-Nearest Neighbors	10	14.2450	7.81631	2.47173
Support Vector Machine	10	13.4350	7.70556	2.43671

Table 2. Represents Statistically Significant Difference in accuracy values between the algorithms. Support Vector Machines have higher accuracy. K-Nearest Neighbors Algorithm has the lowest accuracy.

	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 Difference	
								Lower	Upper
Accuracy									
Equal variances assumed	0.000	0.995	-0.223	18	0.818	-0.81000	3.47088	-8.10205	6.48205
Equal variances not assumed			-0.223	17.996	0.818	-0.81000	3.47088	-8.10216	6.48216
Loss									
Equal variances assumed	0.000	0.995	0.233	18	0.818	0.81000	3.47088	-6.48205	8.10205
Equal variances not assumed			0.223	17.996	0.818	0.81000	3.47088	-6.48216	8.10216

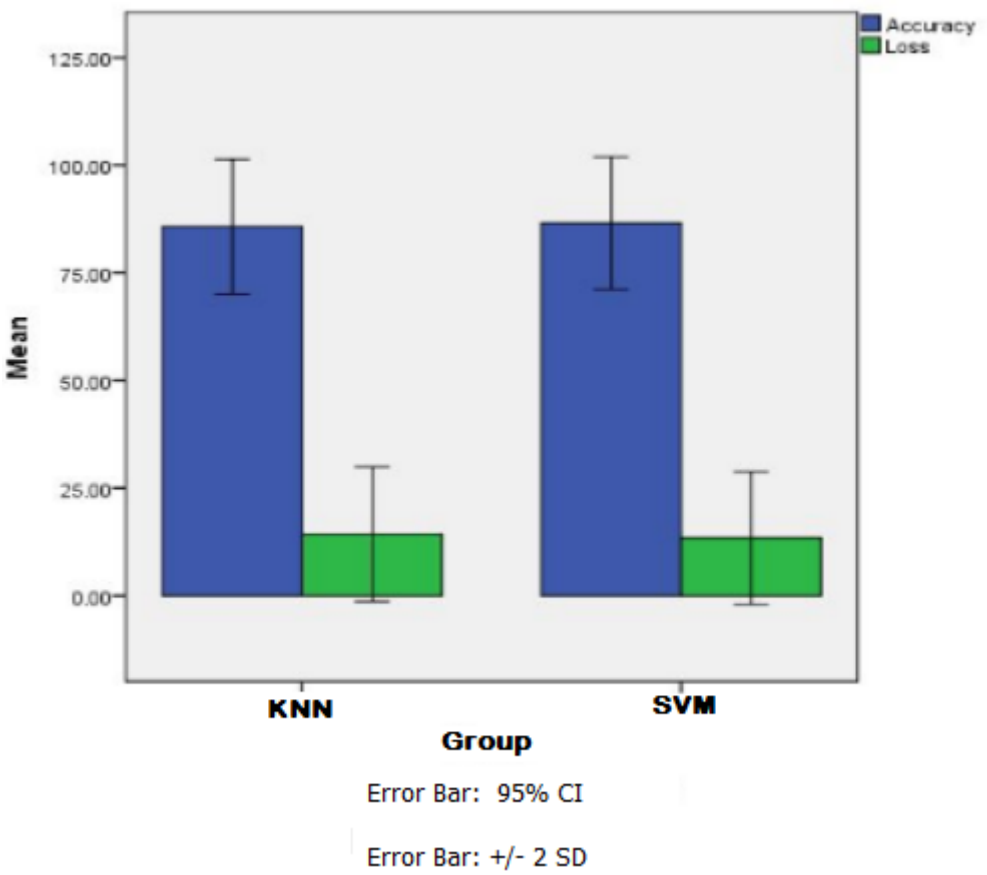

Fig. 1. The Graph concludes that the accuracy of Support Vector Machine (96.15%)and K-Nearest Neighbors (96.10%) is shown and Support Vector Machine performs significantly better than K-Nearest Neighbors algorithm(p=0.001). X-Axis: Support Vector Machine vs K-Nearest Neighbors Algorithm Y-Axis: Mean Accuracy of detection ± 2 SD.