Predicting the Effect of Violent Gameplaying to Violent Behavior Intention among Females using Tree Regression and AdaBoost Tree Regression

Maniah
Computer Science Department, BINUS Graduate Program-Doctor ofComputer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
Major of Informatics Management, Politeknik Pos Indonesia, Bandung, Indonesia 40151
maniah@binus.ac.id;
maniah@poltekpos.ac.id

Bahtiar Saleh Abbas
Computer Science Department, BINUS Graduate Program-Doctor of Computer
Science, Bina Nusantara University,
Jakarta, Indonesia 11480
bahtiars@binus.edu

Yaya Heryadi
Computer Science Department, BINUS Graduate Program-Doctor ofComputer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
yayaheryadi@binus.edu

Wayan Suparta
*Civil engineering Departement University of Technology Yokyakarta,*Yokyakarta Indonesia
drwaynesparta@gmail.com

Agung Trisetyarso
Computer Science Department, BINUS Graduate Program-Doctor ofComputer Science,
*Bina Nusantara University,*Jakarta, Indonesia 11480
atrisetyarso@binus.edu

Ford L. Gaol
Computer Science Department, BINUS Graduate Program-Doctor ofComputer Science,
*Bina Nusantara University,*Jakarta, Indonesia 11480
fgaol@binus.edu

Abstract—**The issue on the effect of violent video game to aggressive behavior has gained wide interest from various communities. This paper presents some results of predicting quantitative measure of aggressive behavior from variables that measure violent video game playing. Experiment results showed that Decision Tree Regression (DTR) and Adaptive Boosting Tree Regression (AB-DTR) models predicted aggressive behavior intentions with high accuracy. For predicting Hostile variable: DTR's training and testing RMSE (0.0, 0.0); AB-DTR's training and testing RMSE (0.08, 1.08). For predicting Instru variable: DTR's training and testing RMSE (0.0, 2.18); AB-DTR's training and testing RMSE (0.0, 3.30) respectively.**

Keywords—violent video game, aggressive behavior, games

I. INTRODUCTION

Violence is any behavior that has the potential to harm other people, such as behavior that causes damage to the structure or function of the body due to coercion or physical stress which leads to death [1]. Violent computer games are games that featuring violent acts to the game opponents in order to progress in or win the game. Aggression was defined by [1][2] as any behavior intended to harm another person who does not want to be harmed.

The issue on the effect of violent video game to aggressive behavior has gained wide interest from various communities including psychologist, educators, teachers, and parents. In the past two decades, a vast number of studies have tried to investigated correlation between exposure of young people to violent video game with aggressive cognition and behaviors [3][4][5][6][7][8]. In contrast to common believe that the effect of aggressive behavior due to video game exposure will last longer, the study by Ferguson [9] concluded that there were not enough evidence to support the claim about long-term effect of violent computer game to aggressive behavior.

Despite a plethora of report studies, to the best of our knowledge, little have been said on prediction model to predict aggressive behavior quantitatively as short-term effect from some measures of violent video game playing as predictor.

This study aims to investigate quantitative model to predict aggressive behavior from violent game playing constructs. In contrast to the study reported by [4] which only focused on correlation analysis between some constructs, the novelty of this study is proposing machine learning algorithms that learn from input data to estimate a target function that maped violent game playing variables to intention to do aggressive behaviors.
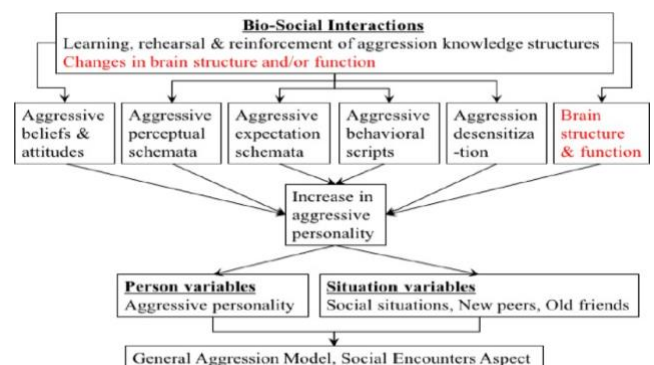

Fig. 1. Personality development model (Source: [2]).

Following the study by [4], the reason for choosing female game player as the study respondents due to common believe that most of game players are males so that violent media have little impact on females game players.

Using the trained model resulted from this study and some evaluation score from some students or children on playing violent/non-violent computer games, teachers or parents can predict the level of aggressive behaviors that tends to happen as its effect to their students or children. The findings of this study; therefore, are usefully for game developers, teacher and parents for selecting the content of computer game which are entertaining or educating but cause less effect on aggressive behavior to their players, students, or children.

The rest of this paper is organized as follows. Chapter 2 briefly describes some previous works related to this study. Next, Chapter 3 explains research method. The study results and discussion describes in Chapter 4. Finally, Chapter 5 concludes the paper.

II. RELATED WORKS

A. The Effect of Playing Violent Game to Aggressive Behavior

As a basis for aggressive behavior study, [2] proposed a framework called General Aggression Model (GAM). This framework involved several factors such as: the role of social, cognitive, personality, developmental, and biological factors on aggression. A further study by [8] described how GAM can be used to explain aggressive behaviors as the effect of exposure to violence in the media including video games (see Fig. 1).

Another experimental study by [7] involving 471 primary and secondary schools students in Italy as participants showed some evidences that participants who prefer using violent video game: (1) tend to present more aggressiveness than participants who do not use violent video games and (2) tend to choose less adaptive strategy in developmental terms such as: distraction or avoidance coping.

The study by [10] involving 789 respondents concluded that: (1) loneliness and depression affect strongly aggression; and (2) in compared to loneliness and depression, aggression was the strongest determinant on game addiction.

B. Tree Regression and Adaptive Boosting Tree Regression

Tree-based regression models are machine learning models which have been widely used to address classification problems.

In general, regression trees is learned inductively using a divide and conquer greedy algorithm to partition training dataset recursively into smaller subsets.

The strength of the tree-based models [11] are: (1) the learned model can be used for understanding the interactions between the variables of the research domain and predicting the value of the target variable of new data, (2) simplicity and efficiency to deal with large dimension and large size of data, (3) no assumption about the function being approximated, (4) the regression-tree models are more interpretable.

However, The study by [11] concluded some drawbacks of the tree-based regression models namely: (1) the model estimation tends to be less reliable particularly in tree's lower levels if the size of training dataset is small, (2) the tree learning is highly affected by the training dataset, (3) the function approximation is highly non-smooth or contains discontinuities.

Following [10], splitting algorithm to learn a regression tree (decision tree regression) from data training can be described below.

Finding the best split of tree regression algorithm
1. Input: n_t cases, sum of their Y values (S_t), the variable X_v
2. Output : The best cut-point split on X_v
3. Sort the cases according to their value in X_v
4. $S_R = S_t, S_L = 0, n_R = n_t, n_L = 0$
5. BestTillNow = 0
6. for all instances i do
7. $\quad S_L = S_L + y_i$
8. $\quad S_R = S_R - y_i$
9. $\quad n_L = n_L + 1$
10. $\quad n_R = n_R - 1$
11. $\quad$ If $X_{i+1,v} > X_{i,v}$ then
12. $\quad\quad$ NewSplitValue $= \frac{S_L^2}{n_L} + \frac{S_R^2}{n_R}$
13. $\quad\quad$ If NewSplitValue > BestTillNow then
14. $\quad\quad\quad$ BestTillNow = NewSplitValue
15. $\quad\quad\quad$ BestCutPoint $= \frac{1}{2}(X_{i+1,v} + X_{i,v})$

Adaptive Boosting (AdaBoost) algorithm proposed by [12] is an ensemble technique or meta-algorithm to improve performance of machine learning algorithm, including tree regression. This algorithm works by combining weak learner algorithms into a weighted sum to form the boosted algorithm. In this research, AdaBoost was used to combine a number of tree regression to build a boosted tree regression. Following [13], AdaBoost algorithm to train AdaBoost Decision Regression Tree model in this study can be described below.

Adaptive Boosting (AdaBoost) algorithm
1. Input: $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ where: $x_i \in X, y_i \in Y = \{-1, +1\}$
2. Output: Hypothesis $H(x)$
3. Initialize weights $D_1(i) = \frac{1}{m}$
4. For $t = 1, \dots, T$
5. $\quad$ Get weak hypothesis $h_t: X \to \{-1, +1\}$ such that $h_t = argmin_{h_j \in H} \epsilon_j = \sum_{i=1}^{m} D_t(i) [\![h_j(x_i) \neq y_i]\!]$
6. $\quad$ If $\epsilon_t \geq \frac{1}{2}$ Then STOP
7. $\quad$ Choose $\alpha_t = \frac{1}{2}(\frac{1-\epsilon_t}{\epsilon_t})$
8. $\quad$ Update $D_{t+1}(i) = \frac{D_t(i)}{Z_t} e^{-\alpha_t y_i h_t(x_i)}$ Where: Z_t be normalization factor.
9. Return $H(x) = sign(\sum_{i=1}^{T} \alpha_t h_t(x))$

III. RESEARCH METHOD

A. Dataset

Dataset for this research was a secondary dataset provided freely by Department of Psychology, Iowa State University (https://public.psych.iastate.edu/caa/classes/419/datasets.html). According to study report by [4], the dataset comprised of 91 samples. This dataset was used to investigate the effects of playing violent video games on aggressive behavior and intentions. Respondent in the data collection were all females.

According to [4], the initial process of data collection was assigning each respondent randomly to play one of two video games. The tested video games were: (1) "Oh No! More Lemmings" game as a sample of a nonviolent game, and (2) "Street Fighter II" game as a sample of violent game in which a player was assigned to control either a female or a male combatant. Violent behaviors were measured using the Taylor Competitive Reaction Time (TCRT) task combined with filling out a short questionnaire.

Following the study by [4], each respondents was requested to play the games in 25 trials. The independent variables of the predictor model were five game playing constructs namely: (1) aggr1 variable was the amount of punishment delivered to the opponent on trial 1; (2) aggrb1 variable was the average amount delivered in trials 2-9; (3) aggrb2 variable was the average amount delivered trials 10-17; (4) aggrb3 variable was the average amount delivered in trials 18-25; and (5) aggr variable was the average amount delivered across all 25 trials.

The dependent variables (response) of the predictor model were aggression variables: (1) hostile variable was a measure of respondent intention to hurt their opponent, and (2) instru variable was a measured respondent intention to interfere or control her opponent performance.

B. Regressor Model Training

In this study, two tree-based models were explored namely: Decision Tree Regression (DTR) and Adaptive Boosting Tree Regression (AB-DTR). Each of these models used to estimate target function that maps game playing construct to aggression variables. Therefore, in total, this studies investigated 4 (four) tree-based regression models.

DTR Model training was implemented using supervised training technique in which dataset was divided randomly into training dataset (80 percent) and testing dataset (20 percent). The depth of trees ranged from 2 to 100. Model performance of training and testing were measured using mean square error (mse):

$$rmse = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - o_i)^2} \qquad (1)$$

Where: n be data size, y_i be actual target value and o_i be predicted target value. The best depth (d) is the value of d whose rmse value is minimum.

Given d-value as a hyperparameter, AB-DTR was trained using number of trees ranging from 5 to 500 trees. The best number of trees is the value that minimize rmse of AB-DTR.

IV. RESULTS AND DISCUSSION

A. Correlation among Variables

Correlation matrix between two variables is summarized in Fig.2.

	AGGR1	AGGRB1	AGGRB2	AGGRB3	AGGR	HOSTILE	INSTRU
AGGR1	1.00	0.44	0.43	0.35	0.50	0.12	-0.03
AGGRB1	0.44	1.00	0.71	0.61	0.87	0.18	0.12
AGGRB2	0.43	0.71	1.00	0.73	0.91	0.36	0.11
AGGRB3	0.35	0.61	0.73	1.00	0.86	0.32	0.19
AGGR	0.50	0.87	0.91	0.86	1.00	0.31	0.14
HOSTILE	0.12	0.18	0.36	0.32	0.31	1.00	0.29
INSTRU	-0.03	0.12	0.11	0.19	0.14	0.29	1.00

Fig. 2. Spearman Correlation matrix between observed variables.

B. Predicting Hostile Variable

1) Finding the Most Optimum Depth of Tree Regression

From experimentation, it was found that in the range of tree depth from 2 to 100, the most optimum depth is $d = 44$ (see Fig. 3) with RMSE-training = 0 and RMSE-testing = 0. This value will be used as parameter of AB-DTR.

2) Performance of Tree Regression Training

Comparison between predicted and actual Hostile Value during training process can be visualized using Fig. 4. As can be seen in Fig. 4 most of predicted Hostile values were inline or very close to its actual values.
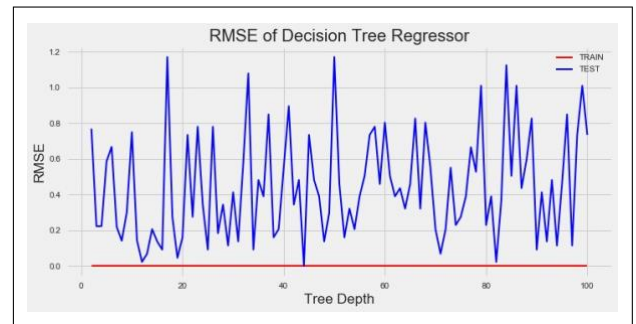

Fig. 3. RMSE-training value of decision tree regressor for tree depth $d \in [2,100]$ and Hostile target variable.

Fig. 4. Comparison between predicted and actual Hostile values using DTR from training process.

3) Finding the Most Optimum Number of Tree for Adaptive Boosting Tree Regression
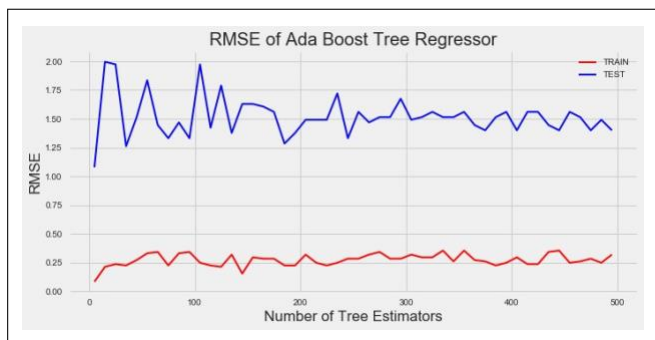
Fig. 5. RMSE value of AdaBoost decision tree regressor for [5,500] number of tress to predict Hostile target variable.

From experimentation, it was found that in the interval [5, 500], the most optimum number of trees for AB-DTR is 5 (see Fig. 5] with RMSE-training = 0.08 and RMSE-testing = 1.08.

4) Performance of Adaptive Boosting Tree Regression Training

Comparison between predicted and actual Hostile Value using AB-DTR as predictor model can be visualized Fig. 6 which showed that most of predicted Hostile values were very close to its actual values.
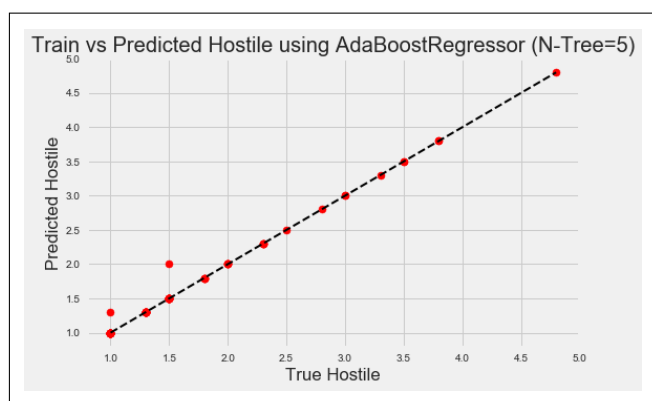

Fig. 6. Comparison between predicted and actual Hostile values using AB-DTR from training process.

C. Predicting Instru Variable

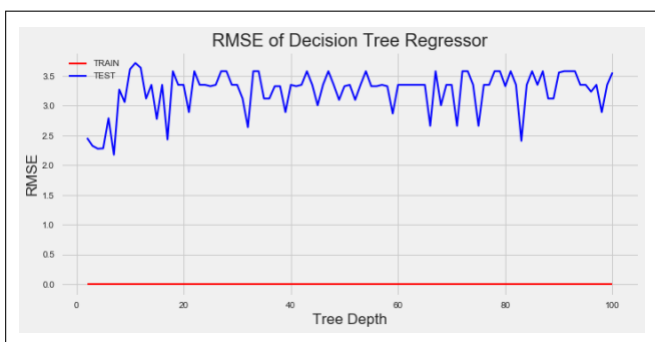1) Finding the Most Optimum Depth of Tree Regression

Fig. 7. RMSE value of decision tree regressor for tree depth $d \in [2,100]$ and Instru target variable.

From the experiments, it was found that from 2 to 100 tree depth, the most optimum depth is d=7 (see Fig. 7) with RMSE-

training = 0 and RMSE-testing = 2.18. This value will be used as parameter of AB-DTR.

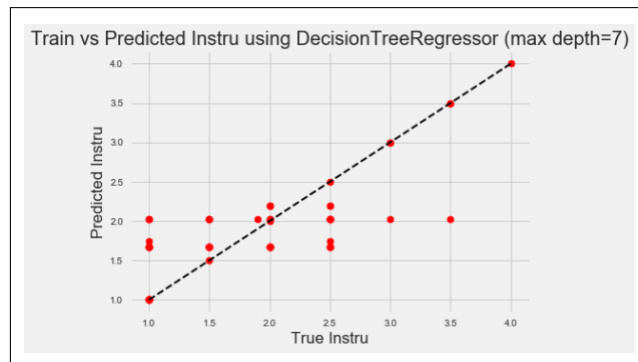2) Performance of Tree Regression Training

Fig. 8. Comparison between predicted and actual Instru values using DTR from training process.

Comparison between predicted and actual Instru values during training process can seen in Fig. 8 which showed that most of predicted Instru values were not very close to its actual values.

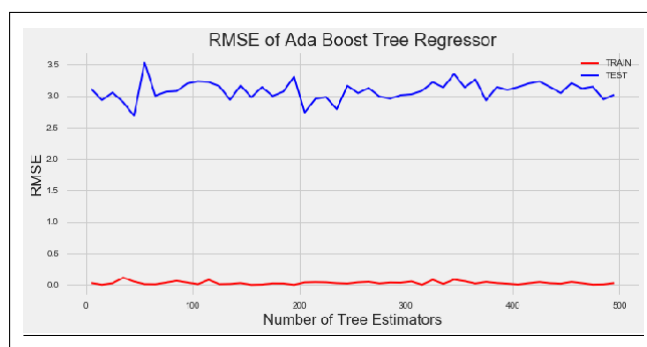3) Finding the Most Optimum Number of Tree for Adaptive Boosting Tree Regression

Fig. 9. RMSE value of AdaBoost decision tree regressor for [5,500] number of tress to predict Instru target variable.

The experiment found that in the interval [5, 500], the most optimum number of trees for AB-DTR is 195 with RMSE-training = 0 and RMSE-testing = 3.30 (see Fig. 10).

4) Performance of Adaptive Boosting Tree Regression Training

RMSE from both training and testing of both DTR and AB-DTR models to predict Hostile and Instru variables can be summarized in the following table.

TABLE I.   ROOT MEAN SQUARE ERROR OF MODEL TRAINING

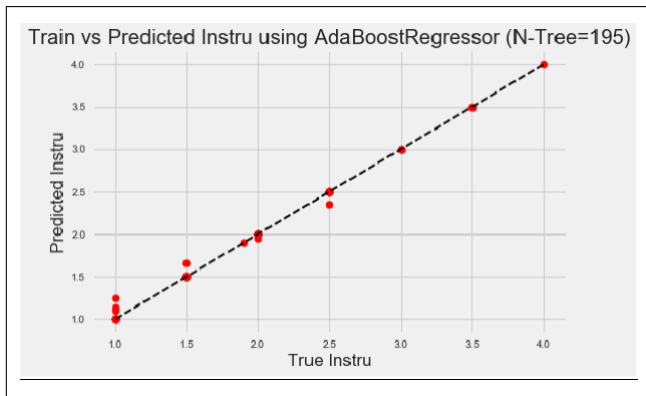| *Respons*	*Model*	*RMSE Training*	*RMSE Testing*
Hostile	Decision Tree Regression	0.00	0.00
	Adaptive Boosting Decision Tree Regression	0.08	1.08
Instru	Decision Tree Regression	0.00	2.18
	Adaptive Boosting Decision Tree Regression	0.00	3.30

Fig. 10. Comparison between predicted and actual Instru values using AB-DTR from training process.

V. CONCLUSION

The issue on the effect of violent video game to aggressive behavior has gained wide interest from various communities thanks to the growing game industry. Many efforts have been reported to investigate relation between the construct of playing violent video games to aggressive behavior. These initiatives aims to to protect youngsters from negative effects of playing game. On the other hand, it also helps fostering game industries to produce games for greater goods. The experiment findings showed that Decision Tree Regression (DTR) and Adaptive Boosting Tree Regression (AB-DTR) models predicted aggressive behavior intentions with high accuracy. For predicting Hostile variable: DTR's training and testing RMSE (0.0, 0.0); AB-DTR's training and testing RMSE (0.08, 1.08). For predicting Instru variable: DTR's training and testing RMSE (0.0, 2.18); AB-DTR's training and testing RMSE (0.0, 3.30) respectively. The findings of this study are still preliminary studies of the effect of violent game playing to violent behavior intention of greater game player profiles for the purpose of anticipation and reducing the risk of violent behavior

REFERENCES

[1] B. J. Bushman and L. R. Huesmann, "Aggression," *Handb. Soc. Psychol.*, no. 734, pp. 833–863, 2010.

[2] C. a Anderson and B. J. Bushman, "Uman ggression," *Annu. Rev. Psychol.*, vol. 53, no. 1, pp. 27–51, 2002.

[3] B. C. A. Anderson and B. J. Bushman, "General Article EFFECTS OF VIOLENT VIDEO GAMES ON AGGRESSIVE BEHAVIOR , AGGRESSIVE COGNITION , AGGRESSIVE AFFECT , PHYSIOLOGICAL AROUSAL , AND PROSOCIAL BEHAVIOR : A Meta-Analytic Review of the Scientific Literature," vol. 12, no. 5, pp. 353–359, 2001.

[4] C. A. Anderson and C. R. Murphy, "Violent Video Games and Aggressive Behavior in Young Women," *Aggress. Behav.*, vol. 29, no. 5, pp. 423–429, 2003.

[5] J. S. Lemmens, B. J. Bushman, and E. A. Konijn, "The Appeal of Violent Video Games to Lower Educated Aggressive Adolescent Boys from Two Countries," *CyberPsychology Behav.*, vol. 9, no. 5, pp. 638–641, 2006.

[6] E. J. Kim, K. Namkoong, T. Ku, and S. J. Kim, "The relationship between online game addiction and aggression, self-control and narcissistic personality traits," *Eur. Psychiatry*, vol. 23, no. 3, pp. 212–218, 2008.

[7] L. Milani, E. Camisasca, S. C. S. Caravita, C. Ionio, S. Miragoli, and P. Di Blasio, "Violent Video Games and Children's Aggressive Behaviors," *SAGE Open*, vol. 5, no. 3, p. 215824401559942, 2015.

[8] J. J. Allen, C. A. Anderson, and B. J. Bushman, "The General Aggression Model," *Curr. Opin. Psychol.*, vol. 19, no. 2, pp. 75–80, 2018.

[9] P. M. Markey, "Finding the Middle Ground in Violent Video Game Research: Lessons From Ferguson (2015)," *Perspect. Psychol. Sci.*, vol. 10, no. 5, pp. 667–670, 2015.

[10] L. Torgo, "Inductive learning of tree-based regression models," *AI Commun.*, vol. 13, no. 2, pp. 137–138, 2000.

[11] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.

[12] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," vol. 139, pp. 23–37, 1995.