Decision Science Letters 11 (2022) 367–378

Contents lists available at GrowingScience

Decision Science Letters

homepage: www.GrowingScience.com/dsl

Simulation and modeling of human decision-making process through reinforcement learning based computational model involving past experiences

Nimisha Gupta[a*], Mitul Kumar Ahirwal[b] and Mithilesh Atulkar[a]

[a]*National Institute of Technology Raipur, India*
[b]*Maulana Azad National Institute of Technology Bhopal, India*

C H R O N I C L E

Article history:
Received April 29, 2022
Received in revised format: June 29, 2022
Accepted August 24 2022
Available online August 24, 2022

Keywords:
Past experiences
Decision-Making
Reinforcement Learning
Learning rules
Iowa Gambling Task

A B S T R A C T

Experience plays a vital role in the decision-making (DM) process. In this paper simulation, modeling, and analysis of past experience over DM has been done using the Iowa gambling task (IGT). The Human DM process is very complex and difficult to model through computational methods because it is a subjective type of process and varies person-to-person. Therefore, this study is an attempt to simulate a DM model similar to the human DM process. For this collection of real data was done and was provided as input to the developed eight Reinforcement Learning (RL) models. The result shows that the performance of the model based on Prospect Utility (PU) learned with Decay Reinforcement Rule (DRI) and Trial Dependency Choice (TDC) is better as compared to other models. It is observed from the analysis of data and also validated that simulation and models output that the experienced group performs better than inexperienced.

© 2022 by the authors; licensee Growing Science, Canada.

1. Introduction

The decision-making (DM) process is an important part of everyday life. This process is required in everything, from daily life to big decisions which impact the future. As the present decision outcomes affect future decisions either positively or negatively. An individual makes decisions about various things. It varies from financial decisions to group decisions (Kar, 2015) in any organization. DM can be applied in many fields such as neuropsychology (Awasthi et al., 2008), computer science (Heikkinen et al., 2008; Fazlollahtabar, 2008; Gupta et al., 2022), expert system, economics (Pramodh et al., 2008), etc. Based on several available options, individuals make choices of appropriate options. Some choices are easy and appear straightforward, while others are complicated and require a multi-step approach to make perfect decisions. Various important factors affect the DM process. Some of them are past experiences (Cohen et al., 2008; Cioffi, 2001), time (Soshi et al., 2019), emotions (De Vries et al., 2008), uncertainty (Lin et al., 2016) and other cognitive biases. Experience plays a significant role in the DM process. As the future decision of an individual may be influenced by the reactions of the present decision. If the experiences are good then people used to prefer similar options to avoid any complexity. On the other hand, if the experiences are bad then they learn from their previous mistakes and avoid making such decisions again. In financial decisions, individuals try to avoid risky decisions based on their past bad experiences. Therefore, to avoid any further losses in the future they analyze every alternative option to make an appropriate decision. Past experiences and memory play a vital role in DM. Past experiences work as teachers and help to learn from their previous mistakes. Researchers have done a lot of research to study the impact of past experiences and memory on future DM. In 1994, one of the well-known researchers Bechara, and his teammates developed a simplified card game known as Iowa Gambling Task (IGT). It is a repeated game play task where based on their experiences and learning participants made decisions. It helps to mimic the real-life DM process.

* Corresponding author.
E-mail address: guptanimisha14@gmail.com (N. Gupta)

© 2022 by the authors; licensee Growing Science, Canada.
doi: 10.5267/dsl.2022.9.001

IGT is an experience-based DM game. It is used for the assessment of the DM ability of patients suffering from psychological disorders(Bechara et al., 1994). Initially, real cards are used to play a game. But as time passed, different researchers modified IGT as per their requirements. In this study, a computerized form of IGT was used(Dancy & Ritter, 2017). Instead of real cards, virtual cards (i.e., decks A, B, C, and D) are displayed over the screen. Each deck includes some reward or penalty value which is shown when the option was selected. The amount won or lost is displayed on the top of the screen. Initially, a monetary value of $2000 was given. Based on deck selection of participants the gain or loss amount was added or deducted from the loan value. The total number of trials and pay-off schedule of the task were not known to the participants. However, this computerized IGT follows the pay-off structure same as the original IGT developed by Bechara and his colleagues. Based on their learning and previous experiences they are asked to gain maximum profit.
The future decisions of an individual are based on their good or bad experiences in their past (Juliusson et al., 2005). Based on that, they chose appropriate actions to make decisions. If something good happened in the past and they encounter similar situations then individuals are likely to make some decisions. On the other hand, if something bad happened in the past then they try to avoid certain decisions (Sagi & Friedland, 2007). Experience helps to make tourism decisions also, as based on their previous experiences. The study conducted by Mazursky (1989) broadens the scope of factors included in the arousal of satisfaction and its future outcomes(Mazursky, 1989). DM is a complex process as it involves various situations containing uncertainties and unknown future outcomes. During emergency health conditions, medical professionals are encountered with certain situations which are unknown to them. In these cases, they use their experience and memory to deal with the current situation. The use of past clinical experiences by medical nurses in emergency conditions has been also studied by Cioffi (2001). Based on findings, it is observed that past experiences are intrinsic to the DM process(Cioffi, 2001). A dynamic choice model under risk was developed to handle risk perception that depends on past experiences. Cohen, Etner, and Jeleva (2008) took the example of the insurance market and discussed the context generated by experience, corresponding to an array of events occurring before the moment of decisions(Cohen et al., 2008). Researchers developed various computational models using the Reinforcement Learning (RL) technique to incorporate the effect of past experience into their models.

As RL is about making decisions sequentially. In RL, the next outcome depends on present input. Based on the actions, participants may receive some reward or punishments and train models accordingly. The model learns from their previous actions to avoid those options which give punishment instead of rewards. In the psychology field, researchers develop IGT to imitate the DM in real life. It is an experienced-based DM task that is based on their experiences and learning they chose the best option for them. It combines various learning/ RL models with different factors to simulate the real-life DM in experiments. An IGT is combined with different pairs and a combination of learning models to compare the model and performance of participants under risk and uncertainty factors (Gupta et al., 2022). In (Kumar et al., 2019) and (Bechara et al., 1994) paper, a method has been proposed with IGT that helps to understand the choice patterns and learning involved in the DM process under risk or uncertainty. Another, Q-learning models are developed by Koluman, Child, and Weyde (2019) to imitate the human DM model in IGT based on emotion-based reward evaluation (Koluman et al., 2019). The DM behaviors of participants in uncertain or risky environments are studied by Gupta et al. (2018) using two DM theories i.e., Expected Utility (EU) and Prospect utility (PU). A simple computational model using IGT data of 145 subjects as benchmark data is used by Lin and his colleagues to develop an optimized model by optimizing some of the parameter values of learning models(Lin et al., 2016). Along with IGT, another experience-based task known as Soochow Gambling Task (SGT) is combined with already developed Expectancy Valence learning (EVL) and Prospect Valence Learning (PVL) models. Dai *et al.*, (2015) combined alternative PU function and mixed updating rules along with EVL and PVL to developed 18 cognitive models and made a systematic comparison for IGT and SGT(Dai et al., 2015). Using IGT, a mathematical heuristic model is developed to simulate an individual's behavioral performance on a risky DM test. Hess *et al.*, (2014) tested the feasibility of the model by fitting the IGT of patients with chronic pain at the individual and group level(Hess et al., 2014). Steingroever, Wetzels, and Wagenmakers (2013) combined different learning models and compared them to study the model performance using IGT to study the behavioral performance using parameter space partitioning techniques(Steingroever et al., 2013). Based on the literature survey, it was observed that the different learning/ RL models are combined with IGT to study the effect in DM criteria or performance of models. But, no study was found which integrates IGT and computational decision-making models with experience factor. Previously, the experience was considered as a part of IGT learning and combined with learning methods to develop a model. Therefore, in this study, past experience is considered as separate factors and combined with the RL model to study the DM behavior of individuals and model performances. For this, IGT data of experienced and inexperienced individuals are collected separately and were given as input to the models.

The proposed RL based model is based on three assumptions. It includes utility function, choice, and learning rules. The Expectancy Value (EV) of outcome is calculated either by using PU or EU functions. After calculating EV, the learning methods such as Decay Reinforcement Learning (DRI) and Delta learning rule (DEL) are used to update the value. In DEL, the EV of the unselected option remains unchanged while in DRI, its value has decayed by some amount. Next, the predicted probability of the selected deck on the next trial is calculated by using the Softmax rule. Its sensitivity θ value is calculated either by using Trial Independent (TIC) or Trial Dependent (TDC) choice rule. A pair and combination of the above methods are used to develop eight learning models. The main objective and novelty of the present work is to study the impact of past experiences on the IGT learning of participants and its incorporation as a past experience factor in the computational learning model. The performances of the developed learning models are also evaluated when IGT data of participants are given as input. The rest of the paper is arranged as: Methodology is given in section 2. This section includes a description of various methods used to develop learning models. Next section 3, presents data collection and experimental procedure. The results of the models are presented and discussed in the result and discussion sections as sections 4 and 5. The final summary with limitations and future scope is given in section 6 as conclusions.

2. Methodology

This section discuss various methods such as learning methods, utility functions, trial independency and dependency rule and model evaluation techniques used in this study.

2.1 Utility Function

2.1.1 EU Function

This function follows the concept of Expected Utility Theory (EUT). According to this theory, the participants choose an action with highest utility value(Gupta et al., 2018). This theory is used when future outcomes are unknown i.e., under uncertain conditions. The value of utility is evaluated using Eq. (1) (Yechiam & Busemeyer, 2005):

$$EU(t) = W.win(t)^{\gamma} - L\,.loss(t)^{\gamma}, \tag{1}$$

where, *W* and *L* denotes the weighting parameter whose value is given in table 1. It gives the weight value to profit and loss value. The trial number is denoted by *t*, where total number of trial is 150. *γ* denotes the curvature of utility function whether it is concave or convex. The amount that is won or loses on trial *t* is given by $win(t)$and $loss(t)$, respectively.

2.1.2 PU Function

The PU function follows the basic core concept of EU. It also includes the concept of individual reference point. As compared to utility values, they focus more on individual gains or losses(Gupta et al., 2018). Eq. (2), gives the equation of PU(t) which give the value of PU at trial *t*(Yechiam & Busemeyer, 2005):

$$PU(t) = \begin{cases} x(t)^{\alpha}, & if x(t) \geq 0 \\ -\lambda |x(t)|^{\alpha}, & if x(t) < 0 \end{cases}, \tag{2}$$

where, *α* gives the shape of PU function whose value ranges from [0, 1]; λ denotes the loss-aversion parameter lies from [0, 5]; *t* denotes the trial number where total number of trial is 150. The total amount won or lose on trial *t* is given by$x(t)$. The value of $x(t)$ is calculated through$x(t) = \text{win}(t) - |\text{loss}(t)|$, where $\text{win}(t)$ and $\text{loss}(t)$ is the winning and losing amount on trial t. If $x(t) < 0$ then it represents the loss value otherwise gain value.

2.2 Learning Rules

The learning rules update the EV obtained from the utility function. It includes two rules i.e., DEL and DRI rules.

2.2.1 DEL Rule

The DEL rule was first introduced by Rescorla and Wagner in 1972(Rescorla & Wagner, 1972).This method is used for learning of experienced-based DM model using IGT. The utility value of selected deck *k* is updated while, the utility value of unselected deck remain unchanged. The utility value of selected deck is updated through Eq. (3)(Yechiam & Busemeyer, 2005):

$$E_k(t) = \alpha_{kt}.E_k(t-1) + \beta_{kt}.u(t), \tag{3}$$

where $u(t)$ denotes the utility value which is computed either through eq. (1) or (2).$E_k(t)$gives the EV for current trial *t* and $E_k(t-1)$gives the EV of previous trial. The value of α_{kt}andβ_{kt} is evaluated through eq. (4-5) which provides the learning weight of previous EV.

$$\alpha_{kt} = 1 - \delta_k(t).\emptyset, \tag{4}$$
$$\beta_{kt} = \delta_{kt}(t).\emptyset, \tag{5}$$

The learning rate is given by parameter$\emptyset$. The value of learning rate parameter ranges from [0, 1]. $\delta_k(t)$represents the dummy variable whose value is 1 when deck *k* is selected, otherwise 0.

2.2.2 DRI Rule

The DRI rule was first introduced by Erev and Roth in 1986. In this rule, the decay parameter is added to the learning models. The rule decayed the EV of unchosen option by some amount while update the EV value chosen option through using Eq. (3) (Yechiam & Busemeyer, 2005).The value of $\alpha_{kt}$and $\beta_{kt}$is calculated using Eq. (6-7):

$$\alpha_{kt} = \emptyset, \tag{6}$$
$$\beta_{kt} = \delta_{kt}(t), \tag{7}$$

$\emptyset$ is the learning rate parameter. $\delta_k(t)$ represents the dummy variable whose value is 1 when deck *k* is selected, otherwise 0.

2.3 Trial independency or dependency rules

The Boltzmann exploration and Softmax rule given by Cane & Luce is used to predict the probability of occurrence of selected deck *k* on next trial (Cane & Luce, 1960b). This method helps to predict the actions on each trial. From previous decks, they estimate their EV. Eq. (8), gives the choice rule equation which is the probability function of relative EV of the alternatives (Yechiam & Busemeyer, 2005).

$$\Pr[D(t+1) = k] = \frac{e^{\theta(t).E_k(t)}}{\sum_{j=1}^{4} e^{\theta(t).E_j(t)}}, \tag{8}$$

$\Pr[D(t+1) = k]$ and $D(t+1)$ denotes the deck selection on next trial (t+1). The sensitivity parameter on trial *t* is given by $\theta(t)$. Its value is calculated either by using TIC or TDC rule which is described below:

2.3.1 TDC

In TDC rule, the value of $\theta(t)$changes with learning. Eq. (9), gives the value of $\theta(t)$:

$$\theta(t) = (t/10)^c \tag{9}$$

where, *t* gives the current trial and *c* gives the consistency parameter ranges from -5 to 5.At *c=0,* the sensitivity value remain constant. With positive and negative value of c, the sensitivity value increases or decreases accordingly.

2.3.2 TIC

In TIC rule, the value of $\theta(t)$remain constant with time and its consistency does not depend on trial *t*. Eq. (10), calculates the value of $\theta(t)$:

$$\theta(t) = 3^c - 1, \tag{10}$$

The value of parameter *c* ranges from 0 to 5. The high and low value of *c* represents more random and deterministic choices, respectively.

2.4 Model Evaluation

The model evaluation of all RL models is evaluated through simulation method. The values of all models are evaluated at different simulation values of 20, 40, 60, 80 and 100. The MSD value is the difference of actual data from observed data. The mean value of across all participants was evaluated to generate the predicted probability. Eq. (11), used to calculate the MSD value (Ahn et al., 2008):

$$MSD = \frac{1}{4*n} \sum_{t=1}^{n} \sum_{j=1}^{4} (\overline{D_{exp,j}}(t) - \overline{D_{sim,j}}(t))^2, \tag{11}$$

$D_{sim,j}$and $D_{exp,j}$ gives the simulated and experimental data respectively. It provides the average probability of actions on trial *t* across all participants. The parameter *n* denotes the total number of trial i.e., 150. The subscript *j* shows the number of deck (1 to 4) and variable *t* denotes the number of trial i.e. from 1 to 150. The value of $D_{sim,j}$ and $D_{exp,j}$ is calculated through using Eq. (12) (Ahn et al., 2008):

$$\overline{D_{exp,j}}(t) = \frac{1}{N} \sum_{i=1}^{N} \overline{Pr_i}\, [D(t) = k], \tag{12}$$

where, N gives the value of total number of participants i.e. 50. $Pr[D(t) = k]$ represents the average probability of the participants for simulated and experimental data. The same equation was used to calculate the value for simulated data. Based on results, the MSD value of all learning models are calculated, where low MSD value indicates higher accuracy.

3. Experiment

3.1 Experimental Design

The IGT data of 50 participants is divided into two groups of 25 participants each i.e., G1 and G2 (G1: Experienced Group and G2: Inexperienced Group). The participants of G1 have previous knowledge about the IGT while participants of G2, do not have any experience about IGT. The participants of both groups are asked to play an IGT of 150 trials with an initial loan amount of $2000. Fig. 1 shows the flow diagram of an experienced-based DM model. The collected IGT data of students were given as input to the developed RL model. The developed model consists of three parts i.e., utility function, updating the EV, and ratio choice rule. The utility value of the outcomes is calculated either through EU or PU function. The expectancy value of the future outcomes are updated through learning methods using DEL or DRI rule. The ratio choice or Softmax rule is used to predict the choice probability value of the next trial. The sensitivity value of choice rule is calculated either through TDC or TIC rule. The choice probability value of trials is the output of the RL model. After getting output of the model, the performances of the models are evaluated. It is done through a simulation method. The MSD values of all the models are evaluated. The lower MSD value indicates the higher accuracy of the model. A pair and combination of EU/PU, DEL/DRI, and TDC/TIC rules are used to develop eight RL models. The performances of all models are evaluated using simulation methods.

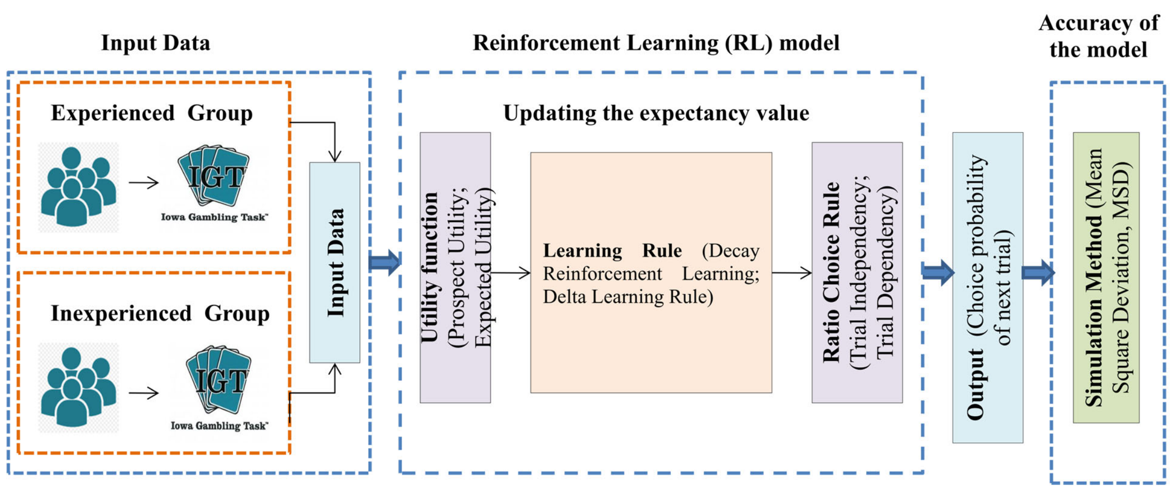

Fig. 1. The flow diagram of experienced-based Decision-Making model

3.2 Data collection

A real IGT data of 50 students (Experienced Group: 25; and Inexperienced Group: 25) were collected. The mean of all subjects are 29.4 years with 1.94 of standard deviation (SD). An online PsyToolkit library (*PsyToolkit Run Experiment*, n.d.) is used to collect the IGT data of participants. The library is used to run various psychological experiments. It follows the basic core concept same as manual IGT given by Bechara and his teammates in 1994 (Bechara et al., 1994).No monetary benefits were given to the participants and their informed consent was obtained before their participation in the study. In this experiment, the IGT data of both groups are collected through computerized IGT.

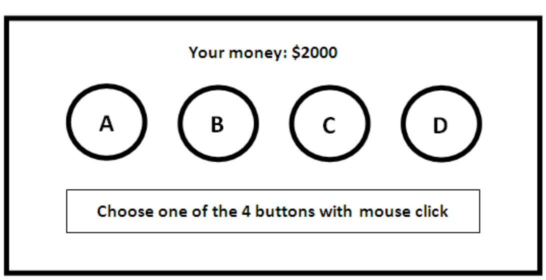

Fig. 2. Iowa Gambling Task (IGT)

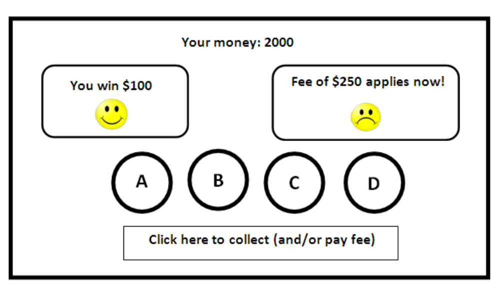

Fig. 3. Iowa Gambling Task (IGT) with losing and winning amount

Fig. 2 shows the pictorial diagram of IGT used in this experiment. A monetary value of $2000 was given to all participants initially. Participants of both groups are asked to pick a card from the deck of four cards. Based on their card selection they may gain or lose some monetary value. Fig. 3 shows the IGT with winning and losing amounts. The screen displays a bar on top which shows the total amount after winning or losing money from deck selections. Both groups play an IGT of 150 trials each. However, they do not have any information regarding the number of trials.

4. Results and discussion

4.1 Analysis of the results

The IGT data of both experienced and inexperienced groups was given as input to the developed RL model. The simulation method is used to evaluate the performance of all developed models. A simulation model runs at simulation values of 20, 40, 60, 80 and 100. The observed value obtained from the simulation model is subtracted from the actual data to obtain the MSD value of the models. The lower MSD value represents the higher accuracy of the model. Likewise, the performances of all models are evaluated and its significance is tested through the statistical analysis of the model.

4.2 Results

4.2.1 Tuning of the parameters

Table 1
Parameter list used in this experiment

Concepts	Models	Parameters	Range	Value used
Utility Function	EU	*γ*: curvature parameter to determines the curvature of the utility function	$\gamma = 1$	$\gamma = 1$
		W and *L*: weight parameter	W and L = 1	W and L = 1
	PU	α: Shape parameter	[0, 1]	$\alpha = 0$
		λ: Loss aversion parameter	[0, 5]	$\lambda = 0$
Learning Method	DRI	ϕ = learning rate parameter	[0, 1]	$\phi = 0.1$
	DEL	ϕ = learning rate parameter	[0, 1]	$\phi = 0.1$
	Dummy Variable	δ = dummy variable	-	$\delta = 1$, if deck *k* is selected $\delta = 0$, if deck *k* is not selected
Sensitivity	TIC	c = Consistency parameter	[0, 5]	c = 1
	TDC	c = Consistency parameter	[-5, 5]	c = 0

Note: N denotes total number of trials i.e. 150.

The parameter used in this experiment is given in Table 1. The table shows the entire range of parameter and their value used in this study. The entire ranges of all parameter values are examined and their optimized value is used in the model. The parameter such as learning rate (ϕ), loss aversion (λ), shape (α), and consistency (c) parameters are used in various equations used to developed the RL model. The shape (α) and loss aversion (λ) parameter are used in Eq. (2) of PU function. The shape (α) parameter gives the shape of the utility function (Lin et al., 2016). The value of shape parameter ranges from [0, 1]. The loss aversion (λ) parameter denotes the value at which subjects are more sensitive to losses compared to profit(Ahn et al., 2008). The model results are evaluated for all 11 different values of α and 10 different value of λ lies between the range of [0, 1] and [0, 5], respectively. The learning rate (ϕ) parameter are used in eq. (4), (6-7) of DEL and DRI. The parameter value ranges from [0, 1]. It is evaluated under all 10 different values of ϕ lies in this range. The sensitivity (c) parameter used in TDC and TIC rule of eq. (9-10). The value of c parameter ranges from [0, 5] and [-5, 5] for TIC and TDC, respectively. Other parameters such as curvature (γ) and weight (W &L) are used in EU function with constant value of 1. Another variable is dummy variable (δ), which is 1 when deck k is selected otherwise 0.

2.4.1 Fitting of Model

In this section, the IGT data of both groups were given as input to the RL model and the results are computed. The performances of all models are analyzed through simulation method. They compute the MSD value of all models where low MSD value indicates the higher accuracy. A pair and combination of different utility function, learning rule and choice rule are combined to develop eight different RL model. The MSD values of all models are computed at different simulation value for both groups separately. The computed value is shown in Table 2. From table, it is observed that the MSD value of model M7 (PU-DRI-TDC) is better than other models for both G1 and G2. In addition, the MSD value of all models is least at 100th simulation values for both groups. For model M7, the MSD value of group G2 is better than G1, except for the 100th simulation value.

Table 2
MSD value of all models for all simulation value for both experienced and inexperienced group

Model No.	Utility Function	Updating the Expectancy	Choice Rule (Ratio choice rule)	Mean Square Deviation (MSD)									
				20 simulations		40 simulations		60 simulations		80 simulations		100 simulations	
				G1	G2	G1	G2	G1	G2	G1	G2	G1	G2
M1	EU	DEL	TDC	7.496E+02	7.687E+02	4.380E+02	4.326E+02	2.113E+02	1.939E+02	6.924E+01	5.259E+01	1.196E+01	8.705E+00
M2	EU	DEL	TIC	7.528E+02	7.779E+02	4.404E+02	4.404E+02	2.131E+02	2.008E+02	7.112E+01	5.891E+01	1.436E+01	1.489E+01
M3	EU	DRI	TDC	4.866E+02	4.628E+02	2.737E+02	2.603E+02	1.216E+02	1.157E+02	3.041E+01	2.893E+01	2.500E-03	5.573E-05
M4	EU	DRI	TIC	4.906E+02	4.662E+02	2.760E+02	2.623E+02	1.226E+02	1.166E+02	3.066E+01	2.914E+01	5.800E-03	2.255E-04
M5	PU	DEL	TDC	4.254E+02	4.237E+02	2.387E+02	2.380E+02	1.056E+02	1.055E+02	2.607E+01	2.622E+01	1.102E-01	2.980E-02
M6	PU	DEL	TIC	4.883E+02	4.831E+02	2.734E+02	2.710E+02	1.204E+02	1.198E+02	2.934E+01	2.945E+01	1.714E-01	7.350E-02
M7	PU	DRI	TDC	4.148E+02	4.107E+02	2.333E+02	2.310E+02	1.037E+02	1.027E+02	2.592E+01	2.567E+01	1.844E-06	5.212E-06
M8	PU	DRI	TIC	4.595E+02	4.436E+02	2.585E+02	2,495E+02	1.149E+02	1.109E+02	2.872E+01	2.773E+01	9.012E-06	2.563E-05

Note: M1: EU-DEL-TDC; M2: EU-DEL-TIC; M3: EU-DRI-TDC; M4: EU-DRI-TIC; M5: PU-DEL-TDC; M6: PU-DEL-TIC; M7: PU-DRI-TDC; M8: PU-DRI-TIC; G1: Experienced Group; G2: Inexperienced Group

Table 3
Mean standard deviation and variance value for choice probability value of actual data for deck A, B, C and D of all models for both groups

S. No.	Deck	Model	Group					
			Group 1			Group 2		
			Mean (μ)	Standard Deviation (σ)	Variance (σ^2)	Mean (μ)	Standard Deviation (σ)	Variance (σ^2)
1.	A	M1	2.788E-02	2.714E-02	7.360E-04	3.120E-02	2.583E-02	6.670E-04
		M2	3.111E-02	2.639E-02	6.960E-04	2.760E-02	2.830E-02	8.010E-04
		M3	1.709E-01	7.807E-02	6.095E-03	1.975E-01	8.301E-02	6.891E-03
		M4	1.697E-01	8.045E-02	6.472E-03	1.952E-01	8.548E-02	7.307E-03
		M5	1.848E-01	2.285E-03	5.220E-06	1.895E-01	2.616E-03	6.840E-06
		M6	1.299E-01	3.538E-03	1.250E-05	1.366E-01	4.280E-03	1.83E-05
		M7	2.121E-01	2.207E-02	4.870E-04	2.219E-01	2.505E-02	6.270E-04
		M8	1.745E-01	2.714E-02	2.043E-03	1.939E-01	5.128E-02	2.629E-03
2.	B	M1	1.122E-02	2.0167E-02	4.070E-04	2.658E-02	2.436E-02	5.940E-04
		M2	2.669E-02	2.519E-02	6.340E-04	1.091E-02	2.095E-02	4.390E-04
		M3	1.620E-01	6.766E-02	4.577E-03	1.927E-01	7.088E-02	4.932E-03
		M4	1.589E-01	6.815E-02	4.644E-03	1.924E-01	7.286E-02	5.309E-03
		M5	1.880E-01	1.806E-03	3.260E-06	1.881E-01	2.028E-03	4.110E-06
		M6	1.338E-01	2.867E-03	8.220E-06	1.345E-01	3.282E-03	1.080E-05
		M7	2.101E-01	2.017E-02	4.070E-04	2.198E-01	2.273E-02	5.160E-04
		M8	1.703E-01	2.519E-02	6.340E-04	1.897E-01	4.650E-02	2.162E-03
3.	C	M1	4.701E-01	9.627E-02	9.269E-03	4.360E-01	6.058E-02	3.670E-03
		M2	4.366E-01	6.692E-02	4.468E-03	4.672E-01	1.056E-01	1.115E-02
		M3	3.750E-01	7.042E-02	4.959E-03	3.320E-01	8.159E-02	6.658E-03
		M4	3.787E-01	7.102E-02	5.044E-03	3.347E-01	8.402E-02	7.060E-03
		M5	3.130E-01	1.565E-03	2.450E-06	3.114E-01	1.906E-03	3.630E-06
		M6	3.667E-01	2.798E-03	7.830E-06	3.650E-01	3.288E-03	1.080E-05
		M7	3.093E-01	2.762E-02	7.630E-04	2.921E-01	3.121E-02	9.740E-04
		M8	3.683E-01	5.633E-02	3.173E-03	3.340E-01	6.376E-02	4.065E-03
4.	D	M1	4.908E-01	8.839E-02	7.812E-03	5.063E-01	7.963E-02	6.342E-03
		M2	5.056E-01	8.634E-02	7.455E-03	4.943E-01	9.795E-02	9.594E-03
		M3	2.921E-01	8.487E-02	7.203E-03	2.778E-01	8.551E-02	7.312E-03
		M4	2.928E-01	8.660E-02	7.499E-03	2.777E-01	8.749E-02	7.655E-03
		M5	3.142E-01	1.439E-03	2.070E-06	3.109E-01	1.888E-03	3.570E-06
		M6	3.696E-01	2.527E-03	6.390E-06	3.638E-01	3.242E-03	1.050E-05
		M7	2.685E-01	2.695E-02	7.260E-04	2.663E-01	2.919E-02	8.520E-04
		M8	2.869E-01	5.476E-02	2.998E-03	2.824E-01	5.987E-02	3.585E-03

Table 3, represents the mean, standard deviation, and variance value of choice probability value for deck A, B, C and D. The values are computed for both groups separately for all decks. For G1, the mean probability of selecting Deck A and B is greater in model M7 (μ: 2.121E-01, σ: 2.207E-02 and σ^2: 4.870E-04, for deck A; μ: 2.101E-01, σ: 2.017E-02 and σ^2: 4.070E-04, for deck B). However, the mean probability of selecting deck C and D is greater in model M1 (μ: 4.701E-01, σ: 9.627E-02 and σ^2: 9.269E-03) and M2 (μ: 5.056E-01, σ: 8.634E-02 and σ^2: 7.455E-03), respectively. For G2, the mean probability for selecting Deck A and B is greater in model M7 (μ: 2.219E-01, σ: 2.505E-02 and σ^2: 6.270E-04, for deck A; μ: 2.198E-01, σ: 2.273E-02 and σ^2: 5.160E-04, for deck B). However, the mean probability of selecting deck C and D is greater in model M2 (μ: 4.672E-01, σ: 1.056E-01 and $σ^2$: 1.115E-02) and M1 (μ: 5.063E-01, σ: 7.963E-02 and $σ^2$: 6.342E-03), respectively.

4.2.3

The learning processes of both groups are analyzed. In this study, a trial of 150 is divided into five separate blocks of 30 trials. The block net score is calculated by subtracting good decks (C and D) from bad decks (A and B). The deck selection is advantageous when net score value is greater than 0, otherwise the selection is disadvantageous. Fig. 4 shows the mean block net score graph for experienced and inexperienced groups. A bar graph is plotted for comparing the block net scores value of both groups. A blue bar indicates experienced and red bar shows inexperienced group. The participants of the experienced group made advantageous selection till block 3 with slighter decrease in block 4 and 5. On the other hand, the participants of the inexperienced group made advantageous selection throughout block except for block 1. In block 1, they made disadvantageous selection. In addition, compared to inexperienced groups the participants of the experienced group made more advantageous selection throughout blocks except for block 5. In block 5, the participants of both groups made the same amount of advantageous selection. The block net score value for both groups in block 5 is 9.68.

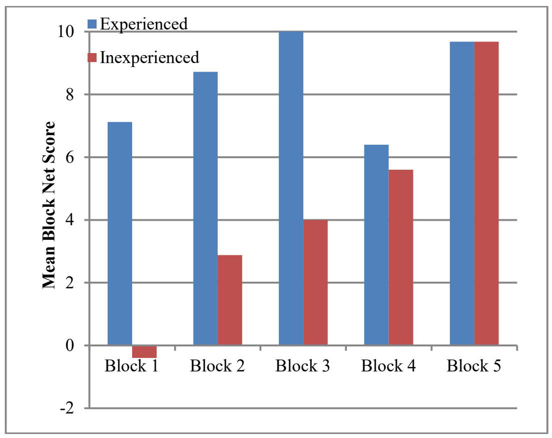

Fig. 4. Mean block net score of experienced and inexperienced group

4.2.4 Deck selection criteria of experienced and inexperienced group

This section analyzes the deck selection criteria for experienced and inexperienced groups. A trial of 150 is divided into 5 separate blocks of 30 trials each. Fig. 5 shows the deck selection criteria for both groups. The participants of the experienced group pick more cards from advantageous decks C and D. As the experienced group already have some knowledge about IGT and their pay-off schedule. Based on their experience, they pick more cards from advantageous decks to make more profit. On the other hand, the participants of the inexperienced group do not have any previous knowledge about IGT. Initially they pick cards randomly and start learning. As soon as the number of trials increases, they get some experience about the pay-off scheme of IGT. They start learning and pick more cards from the advantageous deck in the last block of trials.

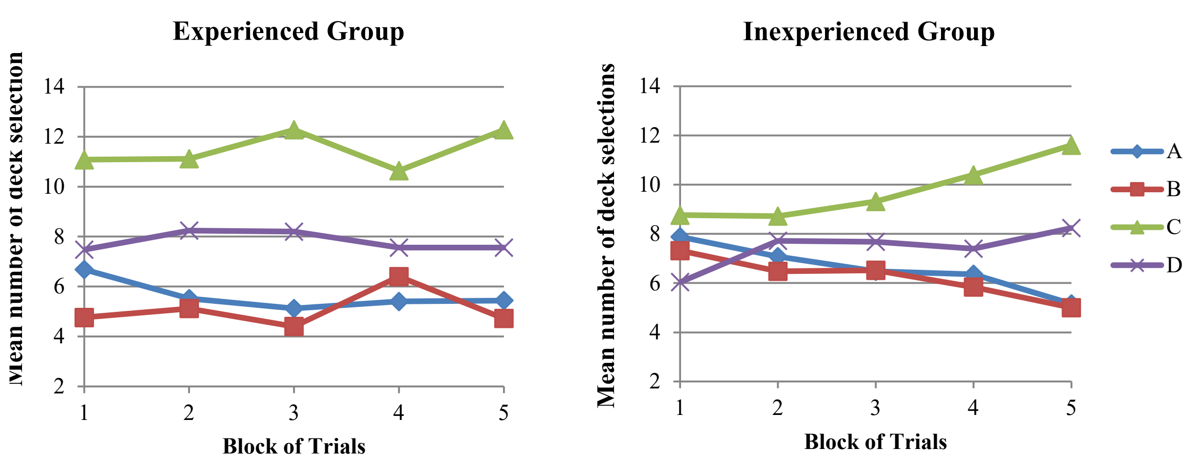

Fig. 5. Deck selection criteria for (A) Experienced, and (B) Inexperienced group

4.3 Statistical analysis

In this section, a statistical analysis was conducted to test the significance of the computed results. For statistical analysis, a post-hoc multi-comparison test is performed using Holm and Shaffer procedure(Chandra & Verma, 2020)(Chandra et al., 2020).The Friedman test compares the average ranking of all models. The minimum ranking of models is considered as the best among all. To conduct the analysis, two hypotheses are assumed. For both groups, it considers the mean ranking of all models equal as null hypothesis. On the other side, the alternate hypothesis is considered as the model performance of all developed models are significantly different. The mean ranking of all models for both groups is shown in Table 4. From the table, it is observed that the performance of model M7 (PU-DRI-TDC) is better compared to others. As the model M7 is having a least rank value which represents the better performance for both groups. The Friedman test rejects null hypothesis for experienced (p-value = 3.948E-05< α = 0.05) and inexperienced (p-value = 1.330E-05< α = 0.05) at seven DoF and 95% (α = 0.05) confidence interval. Based on results, it is observed that the model results are significantly different.

Table 4
Mean ranking of models for experienced and inexperienced group (Friedman test with 7 degree of freedom)

Learning Models	Mean Ranking	
	Experienced Group	Inexperienced Group
M1	7	6.8
M2	8	7.8
M3	4.4	3.6
M4	5.6	4.6
M5	2.6	2.6
M6	4.6	5.8
M7	1	1
M8	2.8	3.8

Note: Minimum rank shows good performance of the model

For experienced and inexperienced groups, Shaffer (Chandra & Verma, 2020)and Holm(Holm, 1979) method of post-hoc test was performed at significance level of 0.05.The post-hoc comparison test for both groups is shown in Table 5 and 6. For both experienced and inexperienced groups, 28 pairwise comparisons for all models are hypothesized. It assumes the performance of all RL models is equal to the null hypothesis. For both groups, the Holm procedure rejects the null hypothesis that has an unadjusted p-value ≤ 2.083E-03 at α = 0.05. The Shaffer procedure denies those hypotheses that have an unadjusted p-value ≤ 1.786E-03 at α = 0.05, for both groups. Based on results, it is observed that the model M7 (PU-DRI-TDC) performs significantly better than others as its p-value is less than the significance level.

Table 5
Adjusted p-value and p-value for post-hoc multiple comparisons of different models at significance level α=0.05 for experienced group

S.No.	Learning Models	$z = \frac{(R0 - Ri)}{SE}$	P	Holm	Shaffer	Adjusted p-value	
						P_{Holm}	P_{Shaf}
1	M2 vs. M7	4.518E+00	6.000E-06	1.786E-03	1.786E-03	1.740E-04	1.740E-04
2	M1 vs. M7	3.873E+00	1.080E-04	1.852E-03	2.381E-03	2.903E-03	2.258E-03
3	M2 vs. M5	3.486E+00	4.910E-04	1.923E-03	2.381E-03	1.276E-02	1.031E-02
4	M2 vs. M8	3.357E+00	7.890E-04	2.000E-03	2.381E-03	1.973E-02	1.657E-02
5	M4 vs. M7	2.969E+00	2.985E-03	2.083E-03	2.381E-03	7.164E-02	6.268E-02
6	M1 vs. M5	2.840E+00	4.509E-03	2.174E-03	2.381E-03	1.037E-01	9.468E-02
7	M1 vs. M8	2.711E+00	6.706E-03	2.273E-03	2.381E-03	1.475E-01	1.408E-01
8	M2 vs. M3	2.324E+00	2.014E-02	2.381E-03	2.381E-03	4.229E-01	4.229E-01
9	M6 vs. M7	2.324E+00	2.014E-02	2.500E-03	2.500E-03	4.229E-01	4.229E-01
10	M2 vs. M6	2.195E+00	2.819E-02	2.632E-03	2.632E-03	5.355E-01	4.510E-01
11	M3 vs. M7	2.195E+00	2.819E-02	2.778E-03	2.778E-03	5.355E-01	4.510E-01
12	M4 vs. M5	1.936E+00	5.281E-02	2.941E-03	2.941E-03	5.355E-01	4.510E-01
13	M4 vs. M8	1.807E+00	7.070E-02	3.125E-03	3.125E-03	8.977E-01	8.449E-01
14	M1 vs. M3	1.678E+00	9.329E-02	3.333E-03	3.333E-03	1.131E+00	1.131E+00
15	M2 vs. M4	1.549E+00	1.213E-01	3.571E-03	3.571E-03	1.399E+00	1.399E+00
16	M1 vs. M6	1.549E+00	1.213E-01	3.846E-03	3.846E-03	1.699E+00	1.577E+00
17	M5 vs. M6	1.291E+00	1.967E-01	4.167E-03	4.167E-03	1.699E+00	1.577E+00
18	M6 vs. M8	1.162E+00	2.453E-01	4.545E-03	4.545E-03	2.360E+00	2.360E+00
19	M7 vs. M8	1.162E+00	2.453E-01	5.000E-03	5.000E-03	2.698E+00	2.698E+00
20	M3 vs. M5	1.162E+00	0.245278	5.556E-03	5.556E-03	2.698E+00	2.698E+00
21	M5 vs. M7	1.033E+00	3.017E-01	6.250E-03	6.250E-03	2.698E+00	2.698E+00
22	M3 vs. M8	1.033E+00	3.017E-01	7.143E-03	7.143E-03	2.698E+00	2.698E+00
23	M1 vs. M4	9.037E-01	3.662E-01	8.333E-03	8.333E-03	2.698E+00	2.698E+00
24	M3 vs. M4	7.746E-01	4.386E-01	1.000E-02	1.000E-02	2.698E+00	2.698E+00
25	M1 vs. M2	6.455E-01	5.186E-01	1.250E-02	1.250E-02	2.698E+00	2.698E+00
26	M4 vs. M6	6.455E-01	5.186E-01	1.667E-02	1.667E-02	2.698E+00	2.698E+00
27	M3 vs. M6	1.291E-01	8.973E-01	2.500E-02	2.500E-02	2.698E+00	2.698E+00
28	M5 vs. M8	1.291E-01	8.973E-01	5.000E-02	5.000E-02	2.698E+00	2.698E+00

Table 6
Adjusted p-value and p-value for post-hoc multiple comparisons of different models at significance level α=0.05 for inexperienced

S.No.	Learning Models	$z=\frac{(R0-\ Ri)}{SE}$	P	Holm	Shaffer	Adjusted p-value	
						P_{Holm}	P_{Shaf}
1	M2 vs. M7	4.389E+00	1.100E-05	1.786E-03	1.786E-03	3.571E-03	3.571E-03
2	M1 vs. M7	3.744E+00	1.810E-04	1.852E-03	2.381E-03	3.704E-03	4.762E-03
3	M2 vs. M5	3.357E+00	7.890E-04	1.923E-03	2.381E-03	3.846E-03	4.762E-03
4	M6 vs. M7	3.098E+00	1.946E-03	2.000E-03	2.381E-03	4.000E-03	4.762E-03
5	M1 vs. M5	2.711E+00	6.706E-03	2.083E-03	2.381E-03	4.167E-03	4.762E-03
6	M2 vs. M3	2.711E+00	6.706E-03	2.174E-03	2.381E-03	4.348E-03	4.762E-03
7	M2 vs. M8	2.582E+00	9.823E-03	2.273E-03	2.381E-03	4.545E-03	4.762E-03
8	M4 vs. M7	2.324E+00	2.014E-02	2.381E-03	2.381E-03	4.762E-03	4.762E-03
9	M5 vs. M6	2.066E+00	3.887E-02	2.500E-03	2.500E-03	5.000E-03	5.000E-03
10	M2 vs. M4	2.066E+00	3.887E-02	2.632E-03	2.632E-03	5.263E-03	5.263E-03
11	M1 vs. M3	2.066E+00	3.887E-02	2.778E-03	2.778E-03	5.556E-03	5.556E-03
12	M1 vs. M8	1.936E+00	5.281E-02	2.941E-03	2.941E-03	5.882E-03	5.882E-03
13	M7 vs. M8	1.807E+00	7.070E-02	3.125E-03	3.125E-03	6.250E-03	6.250E-03
14	M3 vs. M7	1.678E+00	9.329E-02	3.333E-03	3.333E-03	6.667E-03	6.667E-03
15	M3 vs. M6	1.420E+00	1.556E-01	3.571E-03	3.571E-03	7.143E-03	7.143E-03
16	M1 vs. M4	1.420E+00	1.556E-01	3.846E-03	3.846E-03	7.692E-03	7.692E-03
17	M6 vs. M8	1.291E+00	1.967E-01	4.167E-03	4.167E-03	8.333E-03	8.333E-03
18	M4 vs. M5	1.291E+00	1.967E-01	4.545E-03	4.545E-03	9.091E-03	9.091E-03
19	M2 vs. M6	1.291E+00	1.967E-01	5.000E-03	5.000E-03	1.000E-02	1.000E-02
20	M5 vs. M7	1.033E+00	3.017E-01	5.556E-03	5.556E-03	1.111E-02	1.111E-02
21	M4 vs. M6	7.746E-01	4.386E-01	6.250E-03	6.250E-03	1.250E-02	1.250E-02
22	M5 vs. M8	7.746E-01	4.386E-01	7.143E-03	7.143E-03	1.429E-02	1.429E-02
23	M1 vs. M2	6.455E-01	5.186E-01	8.333E-03	8.333E-03	1.667E-02	1.667E-02
24	M3 vs. M5	6.455E-01	5.186E-01	1.000E-02	1.000E-02	2.000E-02	2.000E-02
25	M3 vs. M4	6.455E-01	5.186E-01	1.250E-02	1.250E-02	2.500E-02	2.500E-02
26	M1 vs. M6	6.455E-01	5.186E-01	1.667E-02	1.667E-02	3.333E-02	3.333E-02
27	M4 vs. M8	5.164E-01	6.056E-01	2.500E-02	2.500E-02	5.000E-02	5.000E-02
28	M3 vs. M8	1.291E-01	8.973E-01	5.000E-02	5.000E-02	1.000E-01	1.000E-01

4.4 Discussion

An experimental study was conducted to study the IGT and learning model performance when some participants have or have not been experienced with IGT. A data of 50 students were collected to conduct this experiment and study their effects on IGT. The deck selection pattern and IGT learning of both groups was also analyzed. The collected data was given as input to the eight developed RL model which is a pair and combination of utility function, learning and choice rule. A simulation method was used to analyze the performance of the models. Based on MSD results, it is observed that the performance of model M7 (PU-DRI-TDC) is better compared to other models. Fig. 6 shows the MSD value comparison of both groups. In model M7, the MSD value of the inexperienced group is better than the experienced group except for the 100th simulation value. The results of the models are also statistically proven using a post-hoc multi- comparison test.

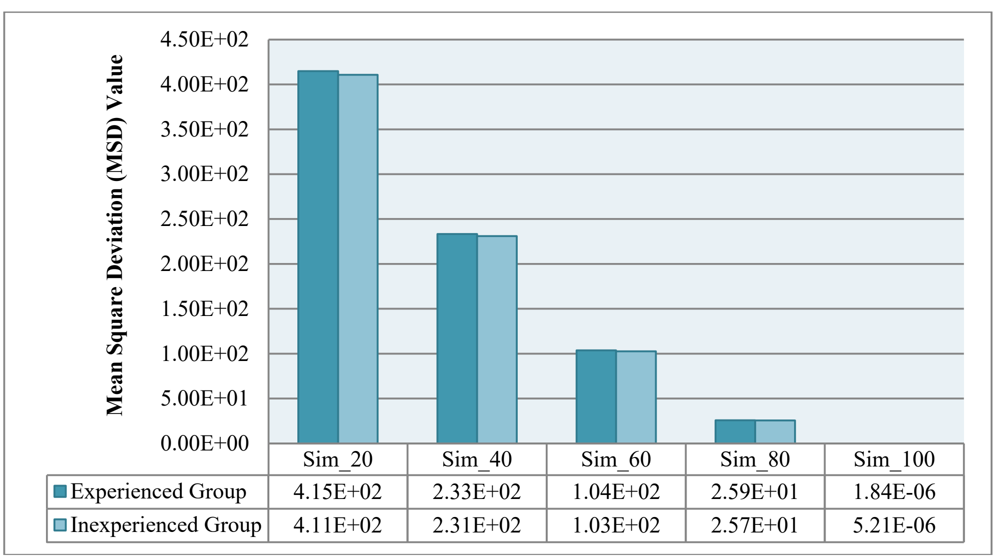

Fig. 6. MSD value comparison of experienced and inexperienced group

The deck selection process of both groups was observed. The participants of both groups pick advantageous decks after a few blocks of trials. The mean deck selection of advantageous decks (C and D) of experienced is higher than inexperienced groups. As the participants already gain some knowledge about the pay-off schedule of IGT. They already have experience about the task compared to inexperienced. The learning process of both groups was analyzed by dividing the IGT trial of 150 into 5 separate blocks of 30 trials. The mean block net score of experienced is higher compared to inexperienced. However, the participants of both groups made advantageous selections throughout blocks except for block 1 of the inexperienced group.

The DM of an individual is subjective in nature which varies from person to person. It also gets affected by a certain set of factors which directly or indirectly affects the process of DM. However, this study only includes the past experience factor. The factor such as uncertainty and risk is already itself included in the task. In this study, the incorporation of other factors is also required which is the limitation of this study. These factors also affect the DM process and are required to be researched further.

5. Conclusion

The IGT is an experienced-based gaming task, in which participants with their experiences, make effective decisions. Based on their learning, they learn about the advantageous and disadvantageous deck selections criteria of IGT to gain maximum profit. In this study, the participants of group 1 have some experience about IGT whereas, participants of group 2 do not have any previous experience about IGT. Based on results, it has been observed that the mean deck selection and block net score of experienced is higher than inexperienced group. As participants had previous knowledge about IGT and they made more advantageous selections. The collected data of students were given as input to the developed08 learning models. The performances of all models are evaluated through simulation methods. The results showed that the performance of model M7 (PU-DRI-TDC) is better compared to other models. The model results are also proven through statistical test using post-hoc multi-comparison test. In future, the other factors should also be included in the modeling of the DM process.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

References

Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive science, 32*(8), 1376-1402.

Awasthi, A., Chauhan, S. S., Hurteau, X., & Breuil, D. (2008). An analytical hierarchical process-based decision-making approach for selecting car-sharing stations in medium size agglomerations. *International Journal of Information and Decision Sciences, 1*(1), 66-97.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1-3), 7-15.

Luce, R. D. (1960). Individual choice behavior, a theoretical analysis. *Bulletin of the American Mathematical Society, 66*, 259-260.

Chandra, T. B., & Verma, K. (2020). Analysis of quantum noise-reducing filters on chest X-ray images: A review. *Measurement, 153*, 107426.

Chandra, T. B., Verma, K., Singh, B. K., Jain, D., & Netam, S. S. (2020). Automatic detection of tuberculosis related abnormalities in Chest X-ray images using hierarchical feature extraction scheme. *Expert Systems with Applications, 158*, 113514.

Cioffi, J. (2001). A study of the use of past experiences in clinical decision making in emergency situations. *International journal of nursing studies, 38*(5), 591-599.

Cohen, M., Etner, J., & Jeleva, M. (2008). Dynamic decision making when risk perception depends on past experience. *Theory and Decision, 64*(2), 173-192.

Dai, J., Kerestes, R., Upton, D. J., Busemeyer, J. R., & Stout, J. C. (2015). An improved cognitive model of the Iowa and Soochow Gambling Tasks with regard to model fitting performance and tests of parameter consistency. *Frontiers in psychology, 6*, 229.

Dancy, C. L., & Ritter, F. E. (2017). IGT-Open: An open-source, computerized version of the Iowa Gambling Task. *Behavior research methods, 49*(3), 972-978.

Vries, M. D., Holland, R. W., & Witteman, C. L. (2008). In the winning mood: Affect in the Iowa gambling task, 3(1), 42–50.

Fazlollahtabar, H. (2008). Applying multiple-criteria decision making methods for developing information technology industry. *International Journal of Information and Decision Sciences, 1*(1), 115-131.

Gupta, N., Ahirwal, M. K., & Atulkar, M. (2018, October). Computational model for human decision making: A study of

prospect theory. *In 2018 Conference on Information and Communication Technology (CICT) (pp. 1-6). IEEE.*
Gupta, N., Ahirwal, M. K., & Atulkar, M. (2022). Human decision making modelling for gambling task under uncertainty and risk. *International Journal of Information and Decision Sciences, 14*(1), 15-38.
Heikkinen, M. V., Matuszewski, M., & Hammainen, H. (2008). Scenario planning for emerging mobile services decision making: mobile Peer-to-Peer Session Initiation Protocol case study. *International Journal of Information and Decision Sciences, 1*(1), 26-43.
Hess, L. E., Haimovici, A., Muñoz, M. A., & Montoya, P. (2014). Beyond pain: modeling decision-making deficits in chronic pain. *Frontiers in behavioral neuroscience, 8*, 263.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.
Juliusson, E. Á., Karlsson, N., & Gärling, T. (2005). Weighing the past and the future in decision making. *European journal of cognitive psychology, 17*(4), 561-575.
Kar, A. K. (2015). A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network. *Journal of Computational Science, 6,* 23-33.
Child, C. H. T., Koluman, C., & Weyde, T. (2019, July). Modelling Emotion Based Reward Valuation with Computational Reinforcement Learning. *In Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 582-588).
Kumar, R., Kumar, K. J., & Benegal, V. (2019). Underlying decision making processes on Iowa Gambling Task. *Asian journal of psychiatry, 39*, 63-69.
Lin, C. H., Lin, Y. K., Song, T. J., Huang, J. T., & Chiu, Y. C. (2016). A simplified model of choice behavior under uncertainty. *Frontiers in Psychology, 7*, 1201.
Mazursky, D. (1989). Past experience and future tourism decisions. *Annals of Tourism Research, 16*(3), 333-344.
Pramodh, C., Ravi, V., & Nagabhushanam, T. (2008). Indian banks' productivity ranking via data envelopment analysis and fuzzy multi-attribute decision-making hybrid. *International Journal of Information and Decision Sciences, 1*(1), 44-65.
PsyToolkit run experiment. (n.d.). Retrieved December 28, 2020, from https://www.psytoolkit.org/experiment-library/experiment_igt.html
Rescorla, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non reinforcement. *Current research and theory*, 64-99.
Sagi, A., & Friedland, N. (2007). The cost of richness: the effect of the size and diversity of decision sets on post-decision regret. *Journal of personality and social psychology, 93*(4), 515.
Soshi, T., Nagamine, M., Fukuda, E., & Takeuchi, A. (2019). Pre-specified anxiety predicts future decision-making performances under different temporally constrained conditions. *Frontiers in psychology, 10*, 1544.
Steingroever, H., Wetzels, R., & Wagenmakers, E. J. (2013). A comparison of reinforcement learning models for the Iowa Gambling Task using parameter space partitioning. *The Journal of Problem Solving, 5*(2), 2.
Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic bulletin & review, 12*(3), 387-402.

© 2022 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).