Optimal Traditional Versus Online Instructional Method Selection

Anthony Afful-Dadzie[a] and Theodore T. Allen[a]
[a]The Ohio State University, 1971 Neil Avenue, 210 Baker Systems, Columbus, Ohio 43210
Email: allen.515@osu.edu

Abstract

This paper discusses the selection of traditional versus online modes of education from the perspective of a profit seeking instructor and a cumulative grade maximizing student. Specifically, we consider a Markov Decision Process model for selecting the best instructional mode for improving academic performance of students assuming that (i) previous mode of teaching has no effect moving forward and (ii) all assignments are balanced so that transitions are independent of assignment. Our results are based on real data from a discrete event simulation course in which participants were randomly assigned to online or traditional model. The results show that, from the perspective of a student advisor, the best learning policy would assign students with superior academic performance to online or blended learning, and face-to-face type of instruction to average and below average students. On the contrary, from the perspective of a profit maximizing instructor, the best instructional policy places all students in an online course with the exception of extremely challenged students who would be switched to face-face after an initial period.

Introduction

A number of empirical studies have compared online course delivery to traditional face-to-face or in-person instruction. Online instruction typically takes place through web-based tools such as chat rooms, threaded discussion groups, Internet activities, videos or slides of course materials, and links to resources[1]. This method of instruction has the advantage of allowing students to work on their own schedule in different locations. In addition, it allows students to replay recorded lectures several times for thorough understanding unlike in traditional learning mode where time spent on instruction is limited. In traditional courses, professors and students meet in person at a designated location and interact with each other, creating a learning atmosphere that is sociable and also offer students the chance to seek direct help from instructors and colleague college students.

Many colleges have or are contemplating converting some of their traditional face-to-face courses into totally online or a hybrid of online and face-to-face instruction for several reasons. First, a considerable amount of recent research have concluded that students in online or a blend of online and traditional face-to-face learning courses perform as well as those in complete face-to-face courses when student performances are compared using pretest and post test scores and grades[2,3,4,5,6,7,8]. Online courses also compare favorably with traditional courses in terms of overall student satisfaction[9,10,11,12,13,14,15,16]. Second, online course delivery mode has undergone tremendous improvement and continues to be improved with current technology. It is now even possible for an instructor to record a full computer laboratory lecture on a personal computer using screen recording software such as Camtasia. The proliferation of laptops and hand-help portable devices such as iPad in recent times has also led to a situation where online lectures can be accessed anywhere without the need to be confined to a particular place. All these have combined to make online teaching considerably cheaper than its traditional face-to-face counterpart.

Yet, despite the numerous benefits online instruction offer and the evidence that it is as good as the traditional method of instruction with regards to academic performance, it is possible it may not be the best medium of instruction for some students especially those below the average performance level.

This paper analyzes the performance of students in an online and traditional face-to-face instruction within the framework of Markov Decision processes (MDP). MDP has been applied in various fields, from education [17] to healthcare[18]. Overall, MDP offers the ability to give policy prescriptions about actions for individuals in every possible state[19].

In Section 2, we introduce the MDP model for the instructional selection problem including a description of the states, actions and costs parameters of the model as well as the assumptions underpinning the model. Section 3 describes the case study data and how the model parameters were obtained. We also present two scenarios considered in the case study relating to whose perspective (student or faculty) the model is based on. In section 4 we present results of our analysis for both case studies. Conclusion and further work is then presented in section 5.

Model formulation

We consider an MDP problem in which a student is enrolled in a certain class for say, a semester or quarter period, and a planning horizon of T quizzes numbered $t = 1, 2, ..., T$ are to be given. For instance, there may be six quizzes for the class during the semester, each quiz offered every 2 weeks. In the subsequent subsections, we define the components of our MDP model.

States, Actions, and Transitions

An MDP model can be described in terms of a state space (S), an action set (A), a transition probability matrix (P), and a reward matrix (R). The set of states (S) describes the system. For example, the performance of a student in a particular class could be classified by a grading scheme from A to E. At any period, the current assignment grade of the student could take on any letter between A and E which defines the state. For our case study and for simplicity, we lump some of the letter grades together such that the system can be in only four states: letter grades (A and A-) as state 1, letter grades (B+, B, and B-) as state 2, letter grades (C+, C, and C-) as state 3, and letter grades (D to E) as state 4. In terms of percentages, 100%-90% corresponds to state 1, 89.9%-80% as state 2, 79.9%-70% as state 3, and 69.9%-0% as state 4.

The action set (A) is a list of the available options one of which could be selected to control the system. We have only two actions for our case study: online or blended learning (action 1) and traditional face-to-face or in-person learning (action 2). For example, an advisor, by looking at the current state of the performance of a student, could advise the student to either enroll in an online class or seek a face-to-face tutoring with an instructor. For each action $a \in A$ taken, a transition probability matrix with elements, P_{ij}^{a}, governs the transition of the system from state $i \in S$ to state $j \in S$. Thus, the transition probabilities describe the dynamics of the MDP model when a particular action is taken.

As an example, if a student's current performance is say a B- (state 2), and is advised to enroll in an online class, he may end up in any of the four states including state 2 in the next period. There is an expected cost R_{ij}^{a}, incurred if the system is in state $i \in S$ and transition to $j \in S$ after action $a \in A$ is taken. If cost is independent of which state $j \in S$ the system transition into after action $a \in A$ is taken, then R_{ij}^{a} can be represented by a vector R_{i}^{a}. For instruction, this assumption applies because we focus on the effect of the current grade on the final cumulative grade which does not depend on the previous state. In Section 3, we describe fully the probabilities and cost data for our model[19].

Thus, an MDP is a 4-tuple (S, A, P, R) with S as a set of the states, A as set of actions, P as an $m \times m$ matrix of element P_{ij}^{a} that governs the transition of the system from state i to state j when performing action $a \in A$, and R as an m dimensional vector with components R_{i}^{a} which gives the reward received when action a is performed in state i.

In our case study, because academic classes have a start and an end date, we focus on a discrete, finite-state, finite-action, finite horizon MDP. At each period, the finite horizon MDP model measures the performance of the system through an objective function cast in the form of a dynamic programming problem,

$$V_t(i_t) = \max_{a \in A}\{ R_{i_t}^{a} + \sum_{j \in S} P_{ij}^{a} V_{t+1}(i_{t+1})\} \tag{1}$$

where $V_t(i_t)$ is the maximum total expected reward attained starting in state i_t with $T-1$ decision periods remaining and $V_{t+1}(i_{t+1})$ is the expected remaining reward in period $t+1, t+2, \dots, T$. $R_{i_t}^{a}$ is the expected immediate reward in time t when taking action a.

The dynamic problem of equation (1) is typically solved using the well-known backward induction technique. By this technique, if R_{i}^{a} is the immediate reward for choosing action a when in state i, and P_{ij}^{a} as the probability that the system moves to state j next period given that the current state is i and action a is chosen, and $V_T(i) = g_T(i)$ is the terminal cost, then for $t = T-1, T-2, \dots, 1$ and for all i at each period,

$$\boldsymbol{V_t(i) = \max_{a \in A}\{ R_{i}^{a} + \sum_{j \in S} P_{ij}^{a} V_{t+1}(j)\}} \tag{2}$$

The optimal action to be taken when in state i at time t is then given as

$$a_t(i) = \mathbf{argmax}_{a \in A}\{ R_i^a + \sum_{j \in S} P_{ij}^a V_{t+1}(j)\} \quad (3)$$

We assume a terminal cost, $g_T(i) = 0$ for our case study.

3. Description of Instructional Method Selection Case Study and Data

Table 1 and Table 2 shows the data for our case study which consist of student performances in an instructional method comparison experiment in [1]. The experiment was conducted during the 2011 winter quarter at a public university in the mid-west. In the experimental study, third-year Industrial Engineering majors were enrolled in a Discrete Event Simulation Laboratory course that had a theory and a laboratory section. The theory section served as a compliment to the laboratory section by providing the theories underpinning the exercises carried out in the lab. The purpose of the study was to test the null hypothesis that face-face or traditional in-person instructional model is significantly better than a blended learning instructional model. Both instructional models required the same textbook and simulation modeling software package.

Both models also had the same grading rubric, covered the same material, and had identical syllabi and weekly schedules. In all, 53 students enrolled in the discrete event simulation course out of which 39 agreed to participate in the experiment. Of these 39 students, 13 attended the complete face-to-face instruction while the remaining 26 attended a blend of online and face-to-face type of instruction. One student from the face-to-face control group did not participate in two assignments and was not considered in the experimental data. The performance of students was assessed through six assignments that consisted of 3 homework, 2 quizzes and a project. There was a two weeks interval between each assignment. The study rejected the null hypothesis that traditional face-to-face instruction leads to a significantly better performance than a blended online and face-to-face instruction. The lesson from this and all past research comparing online and face-to-face instruction is that, if online or blended learning instructional method is cheaper to run and statistically achieves performances comparable to what a traditional face-to-face instruction will achieve, then online instruction should be preferred to traditional face-to-face instruction.

Table 3 shows the transition probabilities obtained from the data in Table 1 and 2. Grades of students on all six assignments and the number of state transitions were recorded. The transition probabilities were then estimated from the state transition data following the technique used in [20]. Figure 1 displays the transition probabilities for traditional in-person instruction shown in Table 3a.

Table 1. Grades on six assignments for students taught through a blend of online and face-to-face instruction in the discrete-event simulation course from the study in [1].

Instruction	HW1	HW2	HW3	Quiz 1	Quiz 2	Project
0nline	B	B-	A	A-	B+	B+
0nline	A	B+	A	A-	A	A
0nline	A	A	A	B	A	A
0nline	A	B	A	A	A-	A-
0nline	A	A	A	A	A	A
0nline	A-	A-	A	C+	B-	C+
0nline	A	A	A	C+	C+	B+
0nline	A	A	A-	B	B	A
0nline	A	A-	A	A	A-	A-
0nline	A	A	A	B	A-	B+
0nline	A-	B+	A	B+	B-	A
0nline	A	A	A	A-	A-	A
0nline	C+	A-	A	B	B	A-
0nline	A	A	A	B+	B	A-
0nline	A	A	A	A-	A	A-
0nline	A	A-	A	A-	A-	A
0nline	B	D	A-	B	B+	A
0nline	A	A	A	B-	B+	A
0nline	A	A	A	B	B	A
0nline	A	A	A-	A	B+	A-
0nline	A	A	A	A	A	A-
0nline	C-	E	D+	B	B-	A-
0nline	A	B	A-	B-	B	A-
0nline	A	A	A	B	A-	A-
0nline	C+	B+	A	B+	A	B+
0nline	A	A	A	B-	A	A-

Table 2. Grades on six assignments for students taught through a complete face-to-face instruction in the discrete-event simulation course from the study in [1].

Instruction	HW1	HW2	HW3	Quiz 1	Quiz 2	Project
In-person	A	A	A	C+	B	A-
In-person	A	A-	A-	B	B+	B+
In-person	A	A	A	A	A	A
In-person	A-	C-	E	A-	A-	A-
In-person	A	A	A	C	A-	A-
In-person	A	A-	B+	B	C	A-
In-person	A	B-	A	B	B+	A-
In-person	A	A	A	B-	A-	B
In-person	A	A	A	B+	A	A
In-person	A	A	A	A-	A-	A-
In-person	A	B	A-	C	C+	B+
In-person	A	B-	A	B+	B	A-

Conceivably, the data in Table 3 could be further divided so that we could have transition matrices and develop models specific to certain types of students. For example, we could consider under-represented groups. However, we only had a few students in those groups participating in our study.

Table 3. Data relating to observed grade transition probabilities for the two actions from the discrete event simulation class: (a) traditional instruction and (b) online instruction.

(a) Transition Probabilities for In-Person

State	A, A-	B+, B, B-	C+, C, C-	D to E
A, A-	25/39	10/39	4/39	0/39
B+, B, B-	8/14	5/14	1/14	0/14
C+, C, C-	2/6	2/6	1/6	1/6
D to E	1/1	0/1	0/1	0/1

(b) Transition Probabilities For Online

State	A, A-	B+, B, B-	C+, C, C-	D to E
A, A-	64/87	21/87	2/87	0/87
B+, B, B-	21/34	11/34	1/34	1/34
C+, C, C-	1/6	3/6	1/6	1/6
D to E	1/3	1/3	0/3	1/3

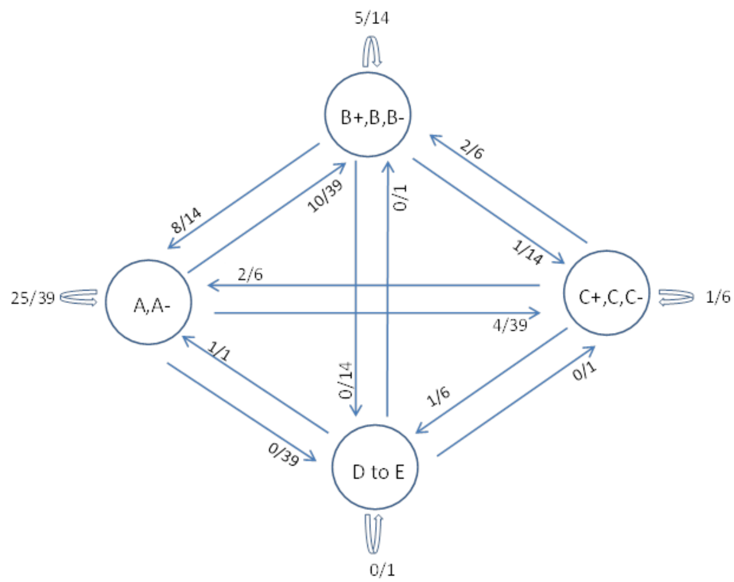

Figure 1. This diagram shows the actual transition probability estimates for students with various grades who consistently take the traditional in-person education for discrete event simulation.

Scenarios, Cost Data, and Assumptions

Two scenarios of the MDP model are considered. In the first scenario, the MDP is modeled from the perspective of a student or a student advisor whose objective is improvement in performance. At each period, the advisor looks up the grade of a student and advises the student to either enroll in an In-Person or Online method of instruction. The second scenario takes the view of an instructor who seeks to maximize profit from teaching by choosing the most cost effective method of instruction to improving the performance of the student.

Table 4 shows the reward matrix for scenario one for both instructional method. The model assumes a student attaches the same value to a grade irrespective of the method of instruction. Thus for instance, an improvement from B+ grade to A grade is valued the same by a student irrespective of whether the instructional method is In-Person or Online. Therefor the reward matrix is the same for both In-Person and Online instruction for scenario one. An arbitrarily average grade point of 95% is awarded to a student irrespective of his previous performance state if he attains an A or A- on the next exam. In a similar manner, a student attains 84.95%, 74.95% or 44.95% if found in state 2, state 3 or state 4 respectively. The model further assumes that the final grade of a student is cumulative of all grades on all scheduled quizzes within the semester and that all quizzes are equally weighted.

Table 4. Student advisor point of view reward matrix (in %) for both Online and In-Person instructional method.

State	A,A-	B+,B,B-	C+,C,C-	C,C-,D+,D,E
A,A-	95.00	84.95	74.95	44.95
B+,B,B-	95.00	84.95	74.95	44.95
C+,C,C-	95.00	84.95	74.95	44.95
D+,D,E	95.00	84.95	74.95	44.95

Table 5. Faculty point of view cost matrix (in $) for Online instructional method

State	A,A-	B+,B,B-	C+,C,C-	C,C-,D+,D,E
A,A-	120	100	80	60
B+,B,B-	120	100	80	60
C+,C,C-	120	100	80	60
D+,D,E	120	100	80	60

Table 6. Faculty point of view cost matrix (in $) for In-Person instructional method

State	A,A-	B+,B	B-,C+	C,C-,D+,D,E
A,A-	100	80	60	40
B+,B,B-	100	80	60	40
C+,C,C-	100	80	60	40
D+,D,E	100	80	60	40

Table 5 and Table 6 also show the cost matrix for scenario two where the MDP model takes the view of an instructor who profits from offering classes. Unlike in scenario one, here an instructor is incentivized based on the final grade per student. Thus, the instructor chooses an instructional method that maximizes his incentive. For simplicity, we assume an instructor receives $120, $100, $80 or $60 per student based on the final grade being in state 1, state 2, state 3 or state 4 respectively for Online delivery. Online instruction is assumed to be $20 cheaper than face-to-face instruction such that incentives for face-to-face is $100, $80, $60 or $40 per student based on the final grade of a student being in state 1, state 2, state 3 or state 4 respectively as shown.

Results and Analysis

This section analyzes the optimal policy for the MDP model for both scenarios. We begin with scenario one where the MDP model takes the view of a student or a student advisor.

MDP from a student perspective

From a student point of view, we desire an instructional policy that maximizes the cumulative grade of the student over the course of a semester. Therefore the objective function of equation (1) can be modified to the form,

$$V_t(i_t) = \max_{a \in A}\{ \delta_t R_{i_t}^a + \sum_{j \in S} P_{ij}^a V_{t+1}(i_{t+1})\} \quad (3)$$

where δ_t is the weight of a particular assignment offered in period t. For instance, homework could be weighed differently than midterm or final exam. If all assignments are assumed equally weighted then equation (3) becomes,

$$V_t(i_t) = \max_{a \in A}\{ R_{i_t}^a / T + \sum_{j \in S} P_{ij}^a V_{t+1}(i_{t+1})\} \quad (4)$$

where T is the number of quizzes due to be taken within the semester. Table 7 below shows the results for both optimal policies and the expected cumulative grade for each period. It can be observed that if at the beginning of each period a student has an A grade, the optimal policy would place the student in an online class. For a grade between B+ and B-, the student will start with a face-to-face instruction for the first two periods and switch to online instruction for the rest of the semester. However, if the performance ranges between C+ and E, the optimal policy would enroll the student in a class with a face-to-face method of instruction throughout the semester. This is consistent with intuition, in that above average students are more likely to succeed with little guidance and would prefer the flexibility online instruction offers. The same cannot be said of below average students who may need more time with an instructor for more explanation and questioning. Table 7 also shows the expected cumulative grade during each period.

Table 7. Periodic Optimal Instructional policies and associated expected reward from a student advisor perspective. Here, the decision periods are effectively the dates of the assignments like homework #1 (HW1) and Quiz #1 (Quiz1).

Period		HW1	HW2	HW3	Quiz1	Quiz2	Project
Optimal Actions	A, A-	2	2	2	2	2	2
	B+, B, B-	1	1	2	2	2	2
	C+, C, C-	1	1	1	1	1	1
	D+ to E	1	1	1	1	1	1
Expected Rewards	A, A-	15.83	31.19	46.41	61.60	76.80	91.99
	B+, B, B-	14.16	29.15	44.30	59.49	74.68	89.88
	C+, C, C-	12.49	25.82	40.80	55.97	71.17	86.36
	D+ to E	7.49	23.33	38.68	53.90	69.09	84.29

MDP from an instructor's perspective

Let us assume that an instructor who receives incentive based on the performance of a student will seek to maximize his expected incentive by offering the least cost mode of instruction that will best improve the performance of the student. While the instructor would prefer the cheaper online mode of instruction, he will be mindful of fewer incentives that may result from poor performances of students. However, since it cost him more to offer a face-to-face instruction, he would be motivated to avoid face-to-face instruction if possible. Table 8 shows the results for the MDP model from the perspective of an instructor. The result is based on the objective function of equation (2).

The results indicate that for a profit maximizing instructor, the best policy will place all students in an online course with the exception of students with grades between D and E. Even for such students, the optimal policy is such that he will prefer enrolling them in an online class at the beginning before switching to face-to-face instruction. Intuitively, the instructor would prefer to begin with the cheaper instructional method until early results indicates that such a student would need extra help with a face-to-face instruction. Table 8 also shows the cumulative reward obtained by the instructor each period.

Table 8. Optimal Instructional policies and associated expected reward from a profit maximizing faculty perspective. Here, the decision periods are effectively the dates of the assignments like homework #1 (HW1) and Quiz #1 (Quiz1).

Period		**HW1**	**HW2**	**HW3**	**Quiz1**	**Quiz2**	**Project**
Optimal Actions	**A, A-**	2	2	2	2	2	2
	B+, B, B-	2	2	2	2	2	2
	C+, C, C-	2	2	2	2	2	2
	D+ to E	2	1	1	1	1	1
Expected Rewards	**A, A-**	120.00	234.25	347.14	459.68	572.17	684.67
	B+, B, B-	100.00	210.59	322.62	435.09	547.58	660.08
	C+, C, C-	80.00	173.33	279.89	391.52	503.93	616.42
	D+ to E	60.00	160.00	274.25	387.14	499.68	612.17

Conclusions

Some research comparing online and face-to-face instructional methods of teaching has concluded that there is no significant difference in average academic performance between online and traditional in-person mode of teaching1. This assertion generally leads to the conclusion that online style of teaching should be preferred over in-person instruction since it is relatively inexpensive to administer online classes. In this article, we presented two scenarios based on real data from a discrete event simulation course in which participants were randomly assigned to online or traditional model for assessing instructional preference. For both scenarios, we applied a Markov Decision Process (MDP) model to determine the optimal instructional method of teaching that should be followed to improve the academic grade performance of individual students in various grade states.

Our results for the specific course we studied show that the optimal instructional strategy from the perspective of a student or a student advisor who seeks to maximize a student’s expected final grade would place students with superior academic performance in an online class and offer face-to-face instruction for students with average or below average performance. Our results also show that from the perspective of a profit maximizing instructor, the optimal instructional policy would place all students in an online class with the exception of student far below the average performance level who will be offered a face-to-face type of instruction for the most part.

One advantage of using MDP over summative recommendations based on average performance is that results from MDP approach could be student specific. Specifically, MDP is able to offer instructional advice on an individual level, separating instructional decisions for students with superior performance from that for students with below average performance. Moreover, the model makes this decision taking into account the expected cost of each method of instruction.

Results from the MDP model based on real data from a single simulation class suggest the optimal instructional policy from a student point of view differs from that of an instructor. The student presumably is focused on grades and has already paid tuition. Therefore, there is minimal (if any) cost difference for online and traditional courses. Whereas, the instructor has costs with respect to teaching student in person that are not relevant for online students.

More specifically, the optimal policies from both points of view for students having current assignment grades ranging from A to B- are (for the most part) to take the online mode of education. However, there is disagreement as to which instructional method is optimal for students with relatively poor performance (i.e. grade C+, C, or C-). From an average student point of view (scenario 1), face-to-face instruction would be preferred over online instruction. In contrast, a profit seeking instructor would prefer otherwise because the expected cumulative reward are much higher from his perspective for traditional face-to-face instruction. For the students with the lowest grades, traditional instruction is prescribed by policies from both the student and the instructor points of view.

Our results suggest a number of opportunities for future research. First, Markov Decision Process models could be replicated for other courses, particularly in cases in which students truly have choice on a daily basis of which mode (traditional or online) to use. Second, the implications of our results suggest that traditional education might generally be relevant for students with poorer grades. These implications could be investigated and confirmed through experimental trials. Third, the level of data behind the recommendations could be investigated using the idea of sampling probabilities from models consistent with the available data. Through such approaches, the sufficiency of the existing data for developing recommendations could be evaluated.

References

1. Allen, T. T., Artis, S., Afful-Dadzie, A. Allam, Y. (2013), "Case Study Application of Blended Learning for Delivering Discrete Event Simulation Laboratories," Technical Report – The Ohio State University, Integrated Systems Engineering.
2. J.B. Arbaugh, M.R.Godfrey, M. Johnson, B. L. Pollack, B. Niendorf, and W. Wresch, "Research in online and blended learning in the business disciplines: Key findings and possible future directions,".*Internet and Higher Education*, 2009, Vol. 12, No. 2, pp. 71-87.
3. Dutton, J., Dutton M. (2005), "Characteristics and performance of students in an online section of business statistics," *Journal of Statistics Education,* Vol. 13, No. 3.
4. Schulman, A. H., Sims, R. L. (1999) "Learning in an Online Format versus an In-Class Format: An Experimental Study," *T.H.E. Journal (Technological Horizons in Education),* Vol. 26, No. 11, pp. 54-56.
5. Miller, B., Cohen, N. L., Beffa-Negrini P. (2001). "Factors for success in online and face-to-face instruction," *Academic Exchange Quarterly*, Vol. 5, No. 4, pp. 4-10.
6. Tucker, S. (2001). "Distance Education: Better, worse, or as good as traditional education?," *Online Journal of Distance Learning Administration*, Vol. 5, No. 5, np.

7. Yablon, Y. B., Katz ,Y.J. (2001). "Statistics Through the Medium of Internet: What Students Think and Achieve," *Academic Exchange Quarterly*, Vol. 5, 17-22.
8. Utts, J., Sommer, B., Acredolo, C., Maher, W. M., Matthews, R. H. (2003) "A study comparing traditional and hybrid internet-based instruction in introductory statistics classes," *Journal of Statistics*, Vol. 11, No. 3, np.
9. Moskal, P., Dzibuban, C. (2011). "A course is a course is a course: Factor invariance in student evaluation of online, blended and face-to-face learning environments," *Internet and Higher Education*, Vol. 14, No. 2, pp. 236-241.
10. Navarro, P., Shoemaker, J. (2000). "Performance and perception of distance learners in cyberspace," *The American Journal of Distance Education*, Vol. 14, No. 2, pp. 15-35.
11. Ashkeboussim, R. (2001). "A comparative analysis of learning experience in a traditional vs. virtual classroom setting,".*The MAHE Journal*, Vol. 24, pp. 5-21.
12. Brace-Govan, J., Clulow, V. (2001). "Comparing face-to-face with online: the learner's perspective," *Academic Exchange Quarterly*, Vol. 5, No. 4, pp. 112-118.
13. Gagne, M., Shepherd, M. (2001). "Distance Learning in Accounting :A comparison between a distance and a traditional graduate accounting class," *T.H.E. Journal*, Vol. 28, No. 9, pp. 58-64.
14. Maddux, C. D., Ewing-Taylor, J., Johnson, D. L. (Eds.), *Distance education: Issues and concerns*. New York: The Haworth Press, 2002.
15. Coomey, M., Stephenson, J. (2001). "Online learning: it is all about dialogue, involvement, support and control—according to the research," in: J. Stephenson (Ed.) *Teaching and learning online: pedagogies for new technologies*, Kogan Page, London, pp. 37-52.
16. Plum, D., Robinson, C. (2012). "In-Class Versus Interactive Television Instruction: A Study to Determine Comparability," *Journal of Quality Approaches in Higher Education*, Vol.3, No. 1, pp. 18-24.
17. Tetreault, J.R., Bohus, D., Litman, D.J. (2007). "Estimating the reliability of MDP policies: A confidence interval approach," *Proceedings of NAACL HLT 2007*, pp. 276–283.
18. Kim, Minsun, ArchisGhate, Phillips, M. H. (2009). "A Markov decision process approach to temporal modulation of dose fractions in radiation therapy planning." *Physics in medicine and biology*, Vol. 54,No. 14: 4455.
19. Puterman, M. L., (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, New York: John Wiley and Sons
20. Mannor, S., Simester S., Sun P., and Tsitsiklis, J.N. (2007). "Bias and variance approximation in value function estimates,"*Management Science*, 53, 308–322.