

## **Avaliação de um ambiente virtual gamificado para auxiliar o ensino-aprendizagem de estudantes de medicina**

**Ademar F. de Sousa Neto<sup>1</sup>, Francisco M. Mendes Neto<sup>1</sup>,  
Rodrigo M. de Lima<sup>2</sup>, Salatiel D. Silva<sup>1</sup>, Everton J. de Oliveira<sup>1</sup>**

<sup>1</sup>Programa de Pós-Graduação em Ciência da Computação - Centro de Ciência Exatas e Naturais  
Universidade Federal Rural do Semi-Árido (UFERSA)  
BR 110 - KM 47 - 59625-900 - Mossoró/RN - Brasil

<sup>2</sup>Coordenação de Sistemas de Informação e Psicologia – Unidade de Ensino Dom Bosco (UNDB)  
Av. Colares Moreira, 443 - 65.075-441 - São Luís/MA - Brasil

{ademarneto,miltonmendes,salatiel.dantas,everton.jales}@ufersa.edu.br,

rodrigo.lima@undb.edu.br

**Abstract.** *In a health educational context, increasingly looking for tools that can assist the learning process of future health professionals. The use of games is usually believed to be an efficient means to try to gather students' attention, who are often distracted by an increasingly connected and interactive world. This paper proposes a 3D environment as a support tool to the training of medical students in medical diagnosis. Students analyze patients through symptoms informed by text and visible in their avatars. The usability of the tool was assessed through the Technology Acceptance Model methodology, with potential users, reaching very positive results, indicating a great potential to reach its objectives.*

**Resumo.** *No contexto da educação na saúde, cada vez mais se busca ferramentas que possam contribuir para facilitar o aprendizado dos futuros profissionais de saúde. O uso de jogos tem sido visto com bons olhos para tentar resgatar a atenção do aluno, muitas vezes distraído por um mundo cada vez mais conectado e interativo. O presente artigo propõe um ambiente 3D como ferramenta de apoio ao treinamento de estudantes de medicina no diagnóstico médico. Os estudantes analisam pacientes através de sintomas informados por intermédio de texto e visíveis em seus avatares. A usabilidade da ferramenta foi avaliada através da metodologia Technology Acceptance Model, com potenciais usuários, alcançando resultados bastante positivos, indicando um grande potencial para alcançar seus objetivos.*

### **1. Introdução**

Dados do Superior Tribunal de Justiça e os Tribunais de Justiça estaduais apontam o crescimento exponencial das discussões sobre erro médico, ao ano, a quantidade de ações judiciais que envolvem médicos aumenta 250% [Delduque 2017]. O Conselho Regional de Medicina do Estado de São Paulo (CREMESP) realizou um teste não obrigatório com alunos recém-formados em Medicina, no ano de 2017, com o objetivo de verificar os

conhecimentos adquiridos pelos estudantes no decorrer do curso. Os resultados foram preocupantes, pois 40% dos alunos foram reprovados, acertando menos de 60% da prova.

Dos mais de 2500 médicos que realizaram o teste, 60% não apresentam conhecimento suficiente sobre doenças parasitárias, 78% erraram o diagnóstico laboratorial de diabetes e 88% não souberam interpretar o resultado de um exame de mamografia [IG 2018]. Em 2016, o índice de reprovados foi de 57%, mesmo com uma diminuição do número de reprovados, os dados ainda são bastante preocupantes para o CREMESP, visto que os médicos irão lidar com vidas humanas, havendo risco para a própria população [IG 2018].

Com isso, o objetivo do artigo é apresentar um jogo sério, denominado DocTraining, que utiliza Realidade Virtual (RV) e um Sistema Multiagente (SMA) com Aprendizado de Máquina, além da sua avaliação, para auxiliar estudantes de medicina no processo de aprendizagem, possibilitando-os a treinar diagnósticos com pacientes virtuais. A RV pode ter várias definições, no entanto, a ideia principal é de uma interação do usuário com o computador em um mundo em 3D, gerando a sensação de presença tridimensional [Bryson 1996]. Bogoni e Pinho (2014) conceituam RV como uma tecnologia de interface humano-computador que possibilita a reprodução de situações do mundo real em ambientes criados em um computador, a fim de oferecer experiências e sensações presentes no mundo real.

O artigo está organizado da seguinte forma: na Seção 2 são apresentados os trabalhos relacionados. Na Seção 3 é abordado o ambiente Doctraining, na Seção 4 temos a avaliação de utilidade percebida e facilidade de uso percebida da ferramenta. Por fim, na Seção 5 temos as conclusões e trabalhos futuros.

## 2. Trabalhos Relacionados

A presente seção apresenta dois trabalhos relacionados a temática. Apesar de serem jogos sérios, que dão suporte ao aperfeiçoamento do estudante de medicina, eles não utilizam aprendizado de máquinas e sistemas multiagentes.

Ziebarth et al. (2014) desenvolveram um jogo sério para estudantes de medicina com suporte ao treinamento de entrevistas centralizadas em pacientes. O jogo utiliza conceitos de interpretação de papéis e conversas detalhadas, simulando uma consulta com um paciente real através de um paciente virtual. O objetivo da ferramenta é usar um jogo sério, ao invés de simulação, para melhorar a motivação dos estudantes em usar a aplicação, a fim de treinar suas habilidades.

Diehl et al. (2011) criaram um protótipo de jogo sério para treinamento de médicos e estudantes de medicina para gerenciamento de pacientes com *Diabetes Mellitus* (DM). O jogo simula uma série de casos clínicos de pacientes com doença, incorporando todas as principais características da terapia de insulina para DM no contexto de cuidado com a saúde. O jogo oferece ao jogador cenários clínicos que precisam de decisões dos alunos para diagnosticarem o melhor tratamento para os pacientes. Depois de cada decisão, os jogadores recebem *feedback* do sistema, comparando as decisões tomadas por estes com outras decisões armazenadas no banco de dados.

Os trabalhos apresentados auxiliam estudantes de medicina no aperfeiçoamento das suas habilidades, mas não foi encontrado trabalhos que utilizam algoritmo de apren-

dizado de máquina em um sistema multiagente integrado a um ambiente multijogador e multiplataforma. Partindo disso, pensou-se em criar um ambiente onde fosse possível o treinamento de alunos em casos clínicos e houvesse a possibilidade de analisar em quais situações os alunos possuem mais dificuldade.

### 3. Doctraining

A principal vantagem de simular situações de riscos em um ambiente virtual é que não serão gerados riscos a nenhum indivíduo, tendo em vista que tudo é simulado. O ambiente proposto neste artigo é tridimensional, gamificado e multijogador. A aplicação pode ser executada em plataformas móveis e pelo computador, através do navegador. A tela principal do ambiente é apresentada numa universidade virtual, onde há um hospital universitário, em que se encontram os consultórios para simular os diagnósticos.

Dentro da universidade virtual, fora do ambiente virtual do hospital, se encontram vários *Non Player Character's* (NPC's), personagens controlados por computadores que ofertam dicas e assuntos ligados à medicina, sendo somada uma pontuação a cada interação do jogador com eles. A Figura 1 mostra essa interação.

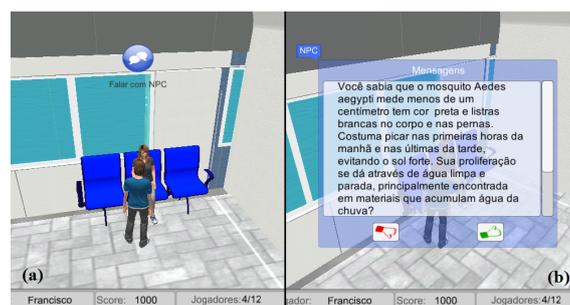


Figura 1. (a) Jogador se aproximando de NPC (b) Conteúdo educativo passado para o aluno

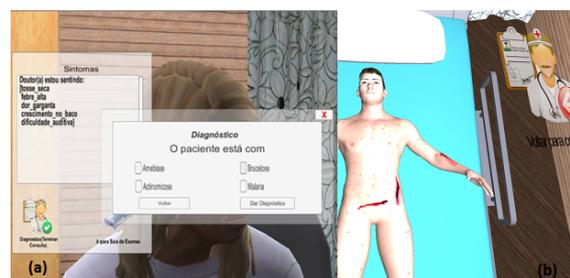


Figura 2. (a) Tela de seleção de diagnóstico (b) Sala de exames com paciente virtual

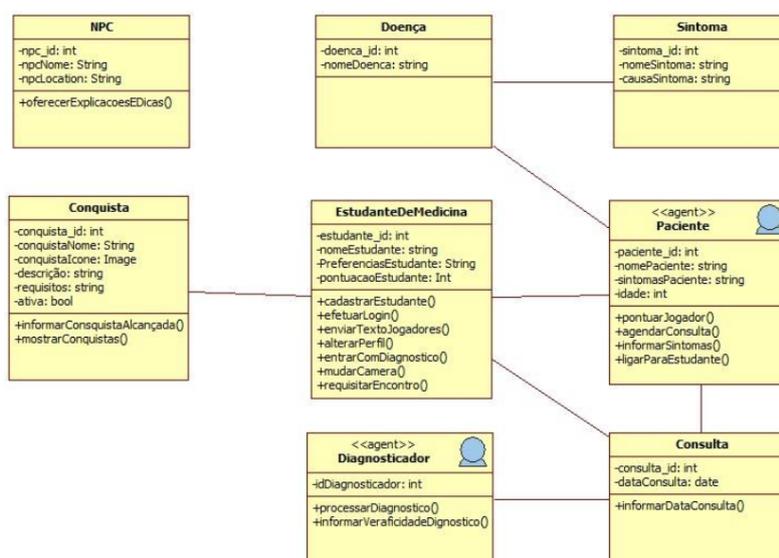
Nos consultórios, dentro do hospital, os pacientes são controlados pelo agente Paciente, que simula problemas de saúde através de amostras de sintomas cadastrados pelos profissionais de saúde. O jogador pode também escolher qual dificuldade deseja utilizar para definir os diagnósticos, quanto maior a dificuldade, menor o número de alternativas disponíveis para escolha e ainda, nos consultórios, há a opção de examinar o corpo do paciente. O usuário pode verificar se o paciente possui manchas no corpo, lesões, entre

outras anomalias. A Figura 2 (a) ilustra a tela de seleção de diagnósticos e a Figura 2 (b) mostra o paciente na sala de exames.

Após o estudante entrar com o diagnóstico do paciente, o sistema captura essa informação e a envia para o agente Diagnosticador. Esse agente possui a função de se comunicar com agente Paciente e, a partir dos sintomas informados por este, processar qual a probabilidade dos diagnósticos corretos.

### 3.1. Sistema Multiagente com Aprendizagem de Máquina

O sistema utiliza algoritmos de uma subárea da IA, o Aprendizado de Máquina, presente na API (*Application Programming Interface*) do WEKA (*Waikato Environment for Knowledge Analysis*). WEKA é um *software* livre, desenvolvido na Universidade de Waikato, na Nova Zelândia, que possui um conjunto de algoritmos de aprendizado de máquina [Kotthoff et al. 2017]. Com o objetivo de verificar a corretude do diagnóstico dado pelo estudante de medicina junto ao banco de dados, inserimos as funções do WEKA dentro do SMA. A Figura apresenta o modelo de recursos e objetos do SMA.



**Figura 3. Modelo de Recursos e Objetos baseado na metodologia MasCommon-Kads+**

Seguindo as normas de modelagem de Agentes de Software propostas pela metodologia MasCommonKads+ [Marçal 2010], foram criados alguns modelos da metodologia. O modelo de requisitos descreve os requisitos do sistema, dividindo-se por casos de uso e cenários da aplicação. Outro modelo utilizado foi o de organização. Esse modelo tem como objetivo demonstrar a estrutura organizacional de papéis e onde eles atuam. A estrutura da organização pode ser definida em: ponto a ponto, quando todos os papéis trabalham no mesmo nível de hierarquia; hierárquico, no qual alguns papéis em uma camada mais baixa são subordinados a outros; federado, quando todos os papéis trabalham no mesmo nível, mas existe um mediador; e, por fim, híbrida, quando a estrutura integra uma das estruturas citadas anteriormente [Marçal 2010]. A Figura 4 mostra o modelo de organização ponto a ponto do SMA.

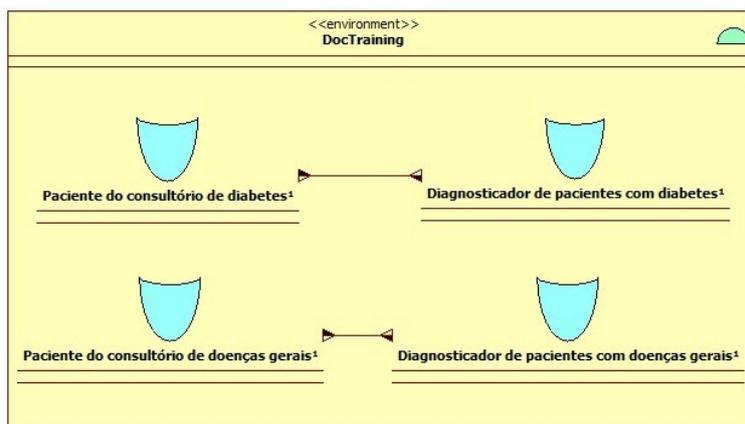


Figura 4. Modelo de organização

Já o modelo de interação consiste em descrever as interações entre agentes, usuários e classes do sistema, especificadas como diagramas de sequência de mensagem e de comunicação. A Figura 5 retrata um dos diagramas de sequência de mensagem utilizados no modelo de interação.

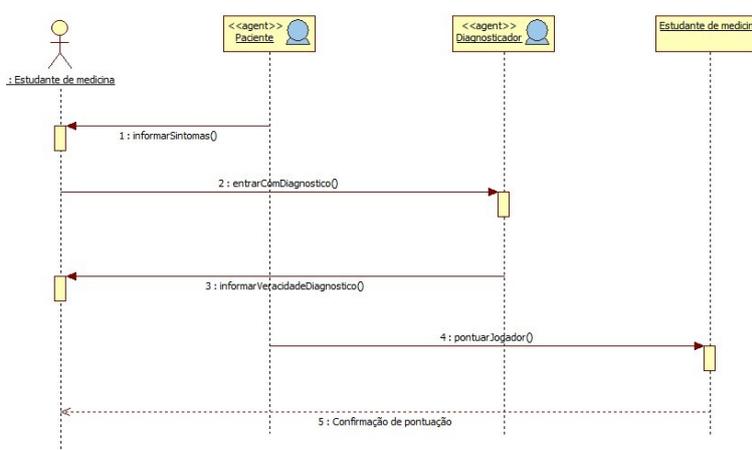


Figura 5. Diagrama de sequência de mensagem

Com o objetivo de treinar os agentes através dos algoritmos presentes no WEKA, o SMA determina, de acordo com os sintomas apresentados, qual doença o paciente possui. Para isso, o SMA conta com dois agentes:

- Paciente: este agente é responsável por capturar informações da base de consultas, acerca dos sintomas que o paciente virtual está sentindo e passá-las ao aluno, para que ele possa dar um diagnóstico;
- Diagnosticador: este agente é responsável por processar o diagnóstico dado pelo estudante de medicina, através de um classificador gerado pelo algoritmo de Aprendizado de Máquina, verificando se está correto.

## 4. Avaliação de Utilidade Percebida da Ferramenta

Como forma de avaliar a eficácia da solução proposta, foi realizado um experimento quantitativo no qual foi possível fazer uso da ferramenta com alunos, com supervisão dos professores da medicina de uma universidade federal.

### 4.1. Procedimentos

A realização do experimento se iniciou por meio de um seminário, apresentando o ambiente para 29 alunos. Os alunos utilizaram a ferramenta por 30 dias, explorando todas as funcionalidades da mesma. Por fim, para avaliação, foi aplicado um questionário seguindo a metodologia *Technology Acceptance Model* (TAM), para que os alunos avaliassem a ferramenta.

Esse questionário foi elaborado com base na escala Likert e utilizando o modelo TAM. O TAM objetiva apresentar as intenções de uso do sistema [Davis et al. 1989]. Davis et al. (1989) diz que o TAM é fundamentado em utilidade percebida, a qual mede o grau em que uma pessoa acredita que o uso do sistema pode melhorar suas atividades, e facilidade de uso percebida, a qual mede o grau em que a pessoa acredita que o uso do sistema de informação é livre de esforço. Conforme Wainerman (1976), Likert afirma que o instrumento de medida proposto por ele pretende verificar o nível de concordância do sujeito com várias afirmações que expressem algo favorável ou desfavorável, em relação a um objeto psicológico. Nesta pesquisa, os sujeitos pesquisados são estimulados a escolher entre diversas opções, marcando aquela que mais se aproxima da sua atitude ou opinião. Atribuímos as seguintes afirmações na escala Likert: Discordância Total (DT); Discordância Parcial (DP); Discordância Leve (DL); Neutro (N); Concordância Leve (CL); Concordância Parcial (CP); Concordância Total (CT).

Os dados obtidos, primeiramente, foram analisados utilizando a proposta de Macnaughton, que considera a quantidade de concordantes (ConP) e discordantes (DisP). O ConP é calculado a partir da soma entre a quantidade de respostas totalmente concordantes e parcialmente concordantes (Equação 1). O DisP é a soma entre a quantidade de respostas totalmente discordantes e parcialmente discordantes (Equação 2) [Macnaughton 1996].

$$ConP = CT + CP + CL + \frac{N}{2} \quad (1) \quad DisP = DT + DP + DL + \frac{N}{2} \quad (2)$$

Com os valores ConP e DisP, Wilder Jr definiu uma formulação para definir o grau de concordância de cada proposição (GCP), que pode ser visualizada na Equação 3 [Wilder 1978].

$$GCP = 100 - \left( \frac{100}{\frac{ConP}{DisP} + 1} \right) \quad (3)$$

Por fim, a partir do resultado do GCP, Davis propôs uma tabela que indica, em linguagem natural, o quanto o sujeito concorda ou discorda em relação a determinada proposição [Davis et al. 1989]. A Tabela 1 apresenta a interpretação dos valores do GCP.

**Tabela 1. Interpretação de Valores do GCP**

<b>Valor de GCP</b>	<b>Frase adequada</b>
90 ou mais	Uma concordância muito forte
80 a + 89,99	Uma concordância substancial
70 a + 79,99	Uma concordância moderada
60 a + 69,99	Uma concordância baixa
50 a + 59,99	Uma concordância desprezível
40 a + 49,99	Uma discordância desprezível
30 a + 39,99	Uma discordância baixa
20 a + 29,99	Uma discordância moderada
10 a + 19,99	Uma discordância substancial
9,99 ou menos	Uma discordância muito forte

#### 4.2. Resultados Obtidos

O questionário levou em conta os dois fatores descritos anteriormente, a Utilidade Percebida, formada por 6 proposições, e a Facilidade de Uso Percebida, onde também é composta por 6 proposições. A Tabela 2 apresenta as questões definidas para esta pesquisa.

**Tabela 2. Proposições da Pesquisa**

<b>UTILIDADE PERCEBIDA</b>	
<b>Nº</b>	<b>QUESTÃO</b>
<b>1</b>	<i>A utilização do DocTraining é importante para o meu aprendizado.</i>
<b>2</b>	<i>Os pacientes virtuais do DocTraining oferecem sintomas verídicos aos da doença apresentada.</i>
<b>3</b>	<i>O DocTraining facilita meu entendimento de assuntos passados em sala de aula.</i>
<b>4</b>	<i>Usar o DocTraining agrega valor ao meu aprendizado em medicina.</i>
<b>5</b>	<i>As universidades deveriam utilizar o DocTraining para treinamento de alunos de curso de medicina.</i>
<b>6</b>	<i>Estou motivado a continuar usando o DocTraining.</i>
<b>FACILIDADE DE USO PERCEBIDA</b>	
<b>1</b>	<i>No DocTraining eu sempre sei onde estou e como chegar onde quero chegar.</i>
<b>2</b>	<i>O DocTraining tem uma interação compreensível e clara.</i>
<b>3</b>	<i>Os recursos de navegação são fáceis de encontrar.</i>
<b>4</b>	<i>O DocTraining possui uma boa jogabilidade.</i>
<b>5</b>	<i>Consigo utilizar o DocTraining sem auxílio de um instrutor.</i>
<b>6</b>	<i>Utilizar o DocTraining é agradável.</i>

A Tabela 3, por sua vez, apresenta a quantidade de respostas contabilizada em cada proposição, com os valores de DisP e ConP, que somados devem resultar na quantidade total de respondentes, que são 29, e GCP resultantes.

**Tabela 3. Dados da Pesquisa**

<b>UTILIDADE PERCEBIDA</b>											
<b>QUESTÃO</b>	<b>DT</b>	<b>DP</b>	<b>DL</b>	<b>N</b>	<b>CL</b>	<b>CP</b>	<b>CT</b>	<b>QTR</b>	<b>DisP</b>	<b>ConP</b>	<b>GCP</b>
<b>1</b>	0	1	3	5	7	8	5	29	6,5	22,5	77,59
<b>2</b>	0	2	2	5	7	8	5	29	6,5	22,5	77,59
<b>3</b>	0	5	4	4	6	6	4	29	11,0	18,0	62,07
<b>4</b>	0	1	3	3	9	7	6	29	5,5	23,5	81,03
<b>5</b>	1	1	4	4	8	3	8	29	8,0	21,0	72,41
<b>6</b>	1	1	4	8	6	2	7	29	10,0	19,0	65,52
<b>FACILIDADE DE USO PERCEBIDA</b>											
<b>QUESTÃO</b>	<b>DT</b>	<b>DP</b>	<b>DL</b>	<b>N</b>	<b>CL</b>	<b>CP</b>	<b>CT</b>	<b>QTR</b>	<b>DisP</b>	<b>ConP</b>	<b>GCP</b>
<b>1</b>	4	5	3	3	3	4	7	29	13,5	15,5	53,45
<b>2</b>	1	0	5	5	3	7	8	29	8,5	20,5	70,69
<b>3</b>	1	1	2	6	2	7	10	29	7,0	22,0	75,86
<b>4</b>	2	2	6	4	4	5	6	29	12,0	17,0	58,62
<b>5</b>	1	0	1	1	4	5	17	29	2,5	26,5	91,38
<b>6</b>	1	1	4	2	8	5	8	29	7,0	22,0	75,86

Partindo da Tabela 3, na categoria Utilidade Percebida, as questões três e seis, que tratam da utilização para o entendimento dos assuntos passados em sala de aula e da motivação em continuar usando o Doctraining, mostrou-se com uma baixa concordância, entre os respondentes, ou seja, cerca de 60% dos alunos concordou de forma mais convincente que o DocTraining é útil para o entendimento dos assuntos. As questões um, dois e cinco, que levam em consideração a importância da utilização do ambiente no aprendizado, a veracidade dos sintomas das doenças apresentadas no ambiente e a utilização da mesma para treinar os alunos nas universidades, respectivamente, obtiveram uma concordância moderada. A questão quatro leva em consideração o valor do DocTraining no aprendizado em medicina e obteve uma concordância substancial. A partir dos resultados, pode afirmar que o DocTraining apresenta boa utilidade percebida pelos estudantes.

Em relação à Facilidade de Uso Percebida, as questões um e quatro obtiveram uma concordância desprezível, onde afirmam que é fácil se orientar na ferramenta e que tem uma boa jogabilidade, respectivamente. As questões dois, três e seis obtiveram concordância moderada, assim, parte dos alunos acharam a interação compreensível e clara, os recursos de navegação fáceis de encontrar e a ferramenta é muito agradável para utilização. Por fim, a questão cinco teve concordância muito forte, mostrando que é possível utilizar a ferramenta sem auxílio de um instrutor. Com base nos resultados, verifica-se que o DocTraining é de fácil utilização, sem necessidade de muito esforço.

## **5. Conclusões e trabalhos futuros**

Este artigo apresentou o ambiente 3D multiagente DocTraining, com aprendizado de máquina, como ferramenta de auxílio a estudantes de medicina para treinamento de diagnósticos médicos com pacientes virtuais. O principal objetivo da ferramenta é a de mitigar erros médicos em consultas clínicas, bem como oferecer uma ferramenta gratuita para as universidades onde estudantes de medicina possam treinar esses diagnósticos de uma maneira mais realista e intuitiva.

O ambiente possui NPC's que oferecem dicas de saúde para os jogadores. Além disso, os consultórios virtuais possuem a opção de examinar o corpo dos pacientes, para verificar lesões, manchas, entre outras. Um técnica importante no ambiente é o aprendizado de máquina, que treina os agentes através dos seus algoritmos para que a cada execução ele esteja mais eficiente.

Os resultados deste trabalho foram obtidos, dos potenciais usuários (alunos de medicina), através de uma avaliação de usabilidade da ferramenta, por meio do método TAM. A pesquisa indicou que o ambiente é útil para os estudantes de medicina, especialmente no tocante ao valor agregado no aprendizado e na facilidade de utilização sem ajuda de um instrutor. Essa conclusão é embasada nas respostas dos usuários, com o GCP, tanto no Utilidade Percebida, quanto na Facilidade de Uso Percebida, resultando todos em concordância, variando apenas os níveis da mesma.

Como trabalhos futuros, espera-se utilizar comitês de classificação para verificação de melhoria na acurácia do cálculo de diagnóstico de amostras não rotuladas. A ideia é combinar múltiplos classificadores independentes para verificar melhorias significativas de veracidade das respostas encontradas, com o objetivo de sempre rotular amostras com o melhor grau de certeza.

## Referências

- Bogoni, T. N. and Pinho, M. S. (2014). Avaliação de um simulador háptico de realidade virtual para treinamento de endodontia. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 25, page 259.
- Bryson, S. (1996). Virtual reality in scientific visualization. *Communications of the ACM*, 39(5):62–71.
- Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8):982–1003.
- Delduque, M. C. (2017). O erro médico sob o olhar do judiciário: uma investigação no tribunal de justiça do distrito federal e territórios-[submetido aos cadernos ibero-americanos de direito sanitário]. *O ERRO MÉDICO SOB O OLHAR DO JUDICIÁRIO: UMA ANÁLISE DAS DECISÕES DO TRIBUNAL DE JUSTIÇA DO DISTRITO FEDERAL E TERRITÓRIOS*, page 52.
- Diehl, L. A., Lehmann, E., Souza, R. M., Alves, J. B., Esteves, R. Z., and Gordan, P. A. (2011). A serious game prototype for education of medical doctors and students on insulin management for treatment of diabetes mellitus. In *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–4.
- IG (2018). Maioria dos médicos recém-formados erra diagnóstico de diabetes, afirma cremesp.
- Kotthoff, L., Thornton, C., Hoos, H. H., Hutter, F., and Leyton-Brown, K. (2017). Auto-weka 2.0: Automatic model selection and hyperparameter optimization in weka. *The Journal of Machine Learning Research*, 18(1):826–830.

- Macnaughton, R. (1996). Numbers, scales, and qualitative research. *The Lancet*, 347(9008):1099–1100.
- Marçal, J. (2010). MAS-COMMONKADS+: Uma Extensão à Metodologia MAS COMMONKADS para Suporte ao Projeto Detalhado de Sistemas Multiagentes Racionais.
- Wainerman, C. (1976). *Escalas de medición en ciencias sociales*. Ediciones Nueva Visión.
- Wilder, J. W. (1978). *New concepts in technical trading systems*. Trend Research.
- Ziebarth, S., Kizina, A., Hoppe, H. U., and Dini, L. (2014). A serious game for training patient-centered medical interviews. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*, pages 213–217.