



ESTUDO DE APRENDIZADO DE MÁQUINA PARA APLICAÇÃO EM PLATAFORMA INTERATIVA ACADÊMICA

Anderson Bartolomeu de Oliveira, Erick de Souza Fernandes, Rodolfo Lyu Shimotsu, Antonio Carlos da Cunha Migliano.

Universidade Braz Cubas, Av. Francisco Rodrigues Filho, 1233, Mogilar – 08773-380 – Mogi das Cruzes-SP, Brasil, anderson.bartholomeu.oliveira@gmail.com, erick.souza281355@gmail.com, lyu1989@hotmail.com, antonio.migliano@brazcubas.br.

Resumo - Este artigo apresenta o estudo de técnicas de aprendizagem de máquinas, mineração de dados e jogos sérios para criação de uma plataforma interativa acadêmica; possibilitando auxiliar os professores para uma melhor abordagem da matéria, facilitando o aprendizado dos alunos e coletando os dados individualmente. Para um resultado de maior nível de precisão à assimilação do conteúdo; tanto individualmente como no coletivo, A ideia principal é verificar em qual área se encontra com maior dificuldade e assim auxiliando o professor em sua metodologia de ensino, para ajudar os alunos que se encontram com uma maior dificuldade, possibilitando um nivelamento uniforme de todos em sala de aula.

Palavras-chave: Aprendizagem de máquina; mineração de dados; jogos sérios; plataforma acadêmica.

Área do Conhecimento: Ensino-Aprendizagem

INTRODUÇÃO

A rápida evolução da tecnologia de comunicação permitiu que empresas coletem e processem grandes quantidades de dados. Estima-se que até 2020, o número total de dispositivos conectados a internet esteja entre 25-50 bilhões de aparelhos. Para empresas de alta tecnologia a confiabilidade e a capacidade de processamento de dados é a chave de seu crescimento. Tal processamento de dados exige o emprego de algoritmos, tais como de Mineração de Dados (*Data Mining - DM*) e que empregam métodos de Aprendizagem de máquina (*Machine Learning - ML*), cujos conceitos surgiram do emprego da *Artificial Intelligence (AI)* (MAHDAVINEJAD, 2017). Na teoria de *games*, o modelo de *AI* ajudou a entender sua classificação atingir objetivos específicos. Esta habilidade é favorecida aplicando-se *DM* para se coletar e formular as estratégias que auxiliam o usuário para adquirir habilidades necessárias para a maior quantidade de dados possíveis, aumentando a precisão dos resultados (BUNKER, 2017). Por outro lado, desde 2006, as empresas de alta tecnologia têm manipulado grandes quantidades de dados, utilizando métodos de *ML* para incorporar inovações no desenvolvimento de seus produtos, adquirindo vantagens em relação à concorrência. Para corroborar este fato, a Tabela 1 apresenta o número de artigos publicados em periódicos e eventos científicos.

Um *game* pode ser capaz de desenvolver habilidades do usuário com uma série de desafios, de acordo com o comportamento do usuário (HERNÁNDEZ, 2017). Tais jogos são conhecidos como jogos sérios (*Serious Games - SG*), que tem como questão desafiadora no aprendizado, a padronização dos dados coletados (LAGUNA, 2017). No entanto, já existem evidências da eficácia de *SG* em processo de ensino, demonstrado na Tabela 1 ao se considerar as publicações referente ao tema (BUCHINGER, 2018).

O objetivo deste projeto é disponibilizar uma plataforma onde possam ser introduzidos conteúdos didáticos voltados à área da engenharia, operando como ferramenta de auxílio no aprendizado das disciplinas. Dessa forma, esta ferramenta será empregada na disciplina de Vetores e Geometria Analítica para testes de validação. Após a conclusão das atividades praticadas na plataforma, o professor responsável pela matéria irá avaliar o desempenho dos alunos. Espera-se que no final do curso, possa avaliar o grau de nivelamento entre os alunos.

Tabela 1 - Número de artigos publicados em periódicos eventos científicos considerando o tema “Data Mining, Machine Learning e Serious Games” nos últimos anos.

Ano	<i>Data Mining</i>	<i>Machine Learning</i>	<i>Serious Games</i>
2018	10493	9825	744
2017	17459	16379	1528
2016	15563	13744	1294
2015	14240	12361	1429

Fonte: <http://www.sciencedirect.com>, em (abril/2018).

METODOLOGIA

A pesquisa deste trabalho é classificada como exploratória por se tratar de um novo tipo de plataforma de ensino baseado em técnicas de aprendizado de máquina junto a técnicas interativas, onde o desenvolvimento foi seguido pelas seguintes etapas: pesquisa bibliográfica, análise de projeto, desenvolvimento do algoritmo, desenvolvimento de interface.

Inicialmente foi realizado as pesquisas bibliográficas sobre AM (Aprendizado de máquina), Mineração de Dados, Ambientes Virtuais de Aprendizagem, Sistemas Acadêmicos e Agentes Pedagógicos para o estudo. Após estudos e entendimento dos conceitos referente ao tema, foram reunidos os conteúdos mais significativos para realizar a análise do projeto e a definição dos meios e ferramentas necessárias que deu início ao desenvolvimento do algoritmo.

Para a elaboração do algoritmo no intuito de desenvolver um modelo de game interativos acadêmicos infundido com AM (Aprendizagem de Máquina), usando a ideia primordial de compreender e executar ações inteligentes para ajudar o usuário em seu reforço acadêmico, foi desenvolvida técnicas de *e-learning* que são métodos de ensino a distância utilizando mídias digitais. Utilizando ferramentas de modelagem de algoritmos e interfaces gráficas e assim confeccionado um cenário para reforço acadêmico, em formato de games usando técnicas de AM sendo uma delas IA (Inteligência Artificial), também utilizando a linguagem de programação *CSharp*.

RESULTADOS E DISCUSSÕES

Uma as principais soluções e desenvolvimento para games sendo as principais citadas por Karlsoon (2005, apud Galdino, 2007) são:

Máquinas de Estado Finito: técnica utilizada para determinar e definir o comportamento dos elementos em um game. Formada por series de eventos, que podem ou não sofrem transação seguindo regras pré-estabelecidas, representando ações que os personagens podem fazer.

Path-Finding: Refere-se a opções de movimentação que poderão ser exercidas no cenário do jogo. Padrões de Movimentos: São as peculiaridades que determinam o comportamento em movimento com os objetos do cenário. Sistemas baseados em regras: conglomerados parâmetros e regras definidas, para realização de uma ação de saída de um objeto ou personagem.

Lógica Fuzzy: representação de tomadas de decisões humanas, analisando sentenças de forma incerta. Observando que um game pode ser considerado software de tempo real ou seja necessita de tomadas de decisão rápidas e coerentes, para o seguimento do seu sistema, analisando as necessidades que fazem um bom game acadêmico no melhor ambiente e com uma junção com IA e analisando o estudo de Galdino do que jogador procura.

“Os jogadores sempre buscam um jogo não muito fácil, mas nem tão difícil ao ponto de ser impossível de se completar o objetivo, eles buscam um jogo onde predomine o equilíbrio de dificuldade. Para se chegar ao equilíbrio é necessário utilizar a IA, pois é ela que proporciona inteligência aos NPC’s (*nonplayer characters*, personagens do jogo que não são controlados pelo humano). A IA fica responsável de possibilitar aos NPC’s a inteligência que somente os humanos possuem proporcionando ao jogador inúmeras situações diferentes no jogo, o que deixa o jogo mais real e interessante para o jogador (Galdino, 2007).”

Juntando tanto as ideias de Galdino com as técnicas Karlsoon pode ser desenvolvida uma plataforma interativa contendo games para reforços acadêmicos.

O desenvolvimento do projeto foi elaborado com fluxogramas e diagrama de classe e métodos sendo demonstrado na Figura 1 e Figura 2, onde mostra o processo de execução do mine game de derivadas.

Figura 1 - Fluxograma.

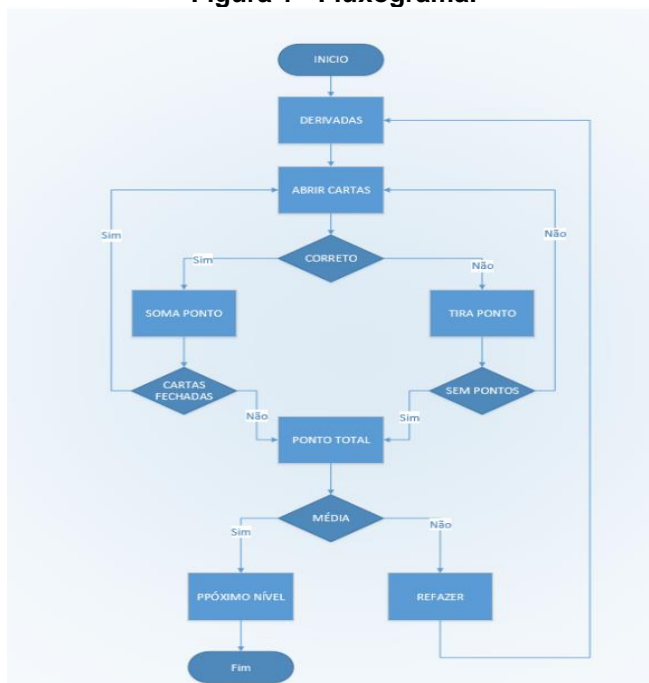
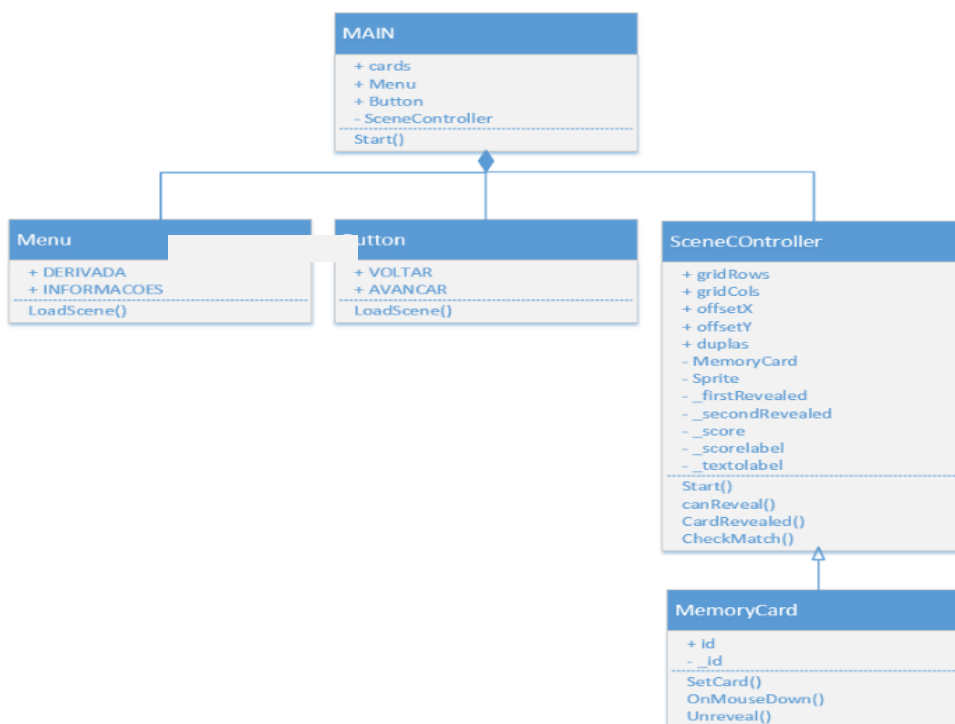


Figura 2 - Diagrama de classes.



Após a confecção do protótipo foram realizados teste de eficiência, funcionalidade e tempo de resposta; primeira fase a permitir correções e aprimoramento. A fase beta é a implementação em

ambiente controlado para observar a aceitação do público alvo e ter o feedback dos jogadores que testaram, tendo os dados necessário para futuras melhorias.

Os dados obtidos do teste permitiram visualizar que uma abordagem diferenciada teve resultados satisfatórios de aceitação, uso e interesse, sendo que mais de 90% dos usuários que testaram a plataforma em sua fase beta, deram seu *feedback* permitindo uma mineração de dados dos pontos a serem melhorados e corrigidos para que futuramente os jogadores tenham melhores interações com a plataforma de games interativos acadêmicos.

O resultado obtido neste projeto permite demonstrar que a implementação de machine learning em uma plataforma acadêmica ajuda a realçar a assimilação da informação passada em aula, onde o pedagogo em conjunto com a plataforma visando melhorar os resultados do aprendizado, obtendo em tempo real a evolução de cada aluno independente.

Na classificação inicial do problema, com mostra na Tabela 2 abaixo, podemos visualizar como foi coletado os dados de cada teste do usuário, onde o *dev_id* é a quantidade de partidas realizadas pelo usuário, *dev_clie* é o código do usuário que realizou o teste e *dev_lv1* a *dev_lv5* são os níveis do teste.

Tabela 2 - Notas dos usuários coletados.

dev_id	dev_clie	dev_lv1	dev_lv2	dev_lv3	dev_lv4	dev_lv5
1	2	2	(NULL)	(NULL)	(NULL)	(NULL)
2	2	3	(NULL)	(NULL)	(NULL)	(NULL)
3	2	4	1	(NULL)	(NULL)	(NULL)
4	2	3	1	(NULL)	(NULL)	(NULL)
5	2	5	2	(NULL)	(NULL)	(NULL)
6	2	6	3	0	(NULL)	(NULL)
7	2	7	2	1	(NULL)	(NULL)
8	2	7	3	1	(NULL)	(NULL)
9	2	8	3	4	(NULL)	(NULL)
10	2	9	4	2	(NULL)	(NULL)

Seguindo na Tabela 3, *med* é a média de todos os dados por níveis. Em seguida o programa observa qual nível está com risco de ser ruim ou bom, em conjunto com os dados de *med*, foi calculado a *dist* que é a distribuição dos níveis que foram mais praticados, permitindo o programa saber qual nível necessita de mais treinamento do usuário.

Tabela 3 - Media da nota geral e Distribuição por nível.

	dev_id	dev_clie	lv1	lv2	lv3	lv4	lv5
med	1	2	6,083333	3,545455	3,375	0	0
dist	1	2	0,387097	0,354839	0,258065	0	0

Para realizar a classificação foi necessário calcular a porcentagem de maior acerto e menor acerto em seguida dividir usuarios por nível acima e abaixo da média, apurado esse dados foi possível calcular a entropia que são grau de impureza dos grupos, onde pode-se observar o grau de impureza em cada nível vide Tabela 4.

Tabela 4 - Porcentagem acima e abaixo da média e Entropia.

	dev_id	dec_clie	lv1	lv2	lv3	lv4	lv5
Acima	1	2	0,583333	0,272727	0,25	0	0
Abaixo	1	2	0,416667	0,727273	0,75	0	0
Entropia	1	2	0,979869	0,845351	0,811278	0	0

O grau de impureza é um dado essencial para ser usado no algoritmo de *Decision Tree* (Árvore de Decisão). Neste algoritmo os dados são divididos no maior número de grupos possíveis, até conseguir um grupo com um grau de impureza próximo de 0.

Tendo os dados coletados em testes aleatórios e os analisando, pode se medir a eficiência dos algoritmos empregados na plataforma e sua funcionalidade constatando que a fase Alpha teve um êxito de 97,5% de eficiência nos testes realizados.



CONCLUSÃO

Após a obtenção dos resultados, conclui-se que a fase beta teve um êxito considerável, podendo dar início as próximas fases e modelando a plataforma para interativo acadêmica, não fica só voltada para as áreas das engenharias, mas sim para uso acadêmico no contexto geral, ou seja, fica disponível em diferentes áreas e possibilitando ao professor ministrar sua aula de uma maneira inovadora, possibilitando uma abordagem diferenciada, explorando seu conteúdo com mais facilidade. Com o diferencial de obter os dados de cada alunos individual e analisar suas dificuldades.

A próxima fase terá incluso o *feedback* dos alunos que testaram a fase beta, isso gerará uma melhoria onde permita uma melhor interação para os usuários com um melhor resultado acadêmico.

REFERÊNCIAS

BUCHINGER, D.; HOUNSELL, M. da S. **Guidelines for designing and using collaborative-competitive serious games**, Computers & Education, v. 118, p. 133-149, ISSN 0360-1315, 2018.

BUNKER, R. P.; THABTAH, F. **A machine learning framework for sport result prediction**, Applied Computing and Informatics, 2017.

HERNÁNDEZ, J. A. C.; DUARTE, M. P.; DODERO, J. M. **Skill assessment in learning experiences based on serious games: A Systematic Mapping Study**, Computers & Education, v. 113, p. 42-60, ISSN 0360-1315, 2017.

LAGUNA, A. S.; ORTIZ, I. M.; HAAG, J. *et all.* **Applying standards to systematize learning analytics in serious games**, Computer Standards & Interfaces, v. 50, p. 116-123, ISSN 0920-5489, 2017.

MAHDAVINEJAD, M. S.; REZVAN, M.; BAREKATAIN, M. *et all.* **Machine learning for Internet of Things data analysis: A survey**, Digital Communications and Networks, 2017.