

Aplicação de Técnicas de Aprendizagem de Máquina em Objetos de Aprendizagem baseado em Software: um Mapeamento Sistemático a partir das Publicações do SBIE

Clébia Alves Beserra – UAST - UFRPE, klebia.alves@gmail.com Cleyton
Carvalho da Trindade – UAST - UFRPE, cleyton13@yahoo.com.br Ellen
Poliana Ramos Souza – UAST - UFRPE, eprs@cin.ufpe.br
Cleyton Vanut Cordeiro de Magalhães – CIn - UFPE, cvcm@cin.ufpe.br
Ronnie Edson de Souza Santos – CIn - UFPE, ress@cin.ufpe.br

Resumo. Pesquisadores da área educacional buscam melhores formas de aplicação de práticas em sala de aula visando permitir uma melhor aquisição do conhecimento. Geralmente são propostas ferramentas que possam auxiliar no desenvolvimento desta área, com o auxílio de técnicas de desenvolvimento mais avançadas como é o caso da Aprendizagem de Máquina. Sendo assim, este trabalho teve como objetivo mapear de que forma as técnicas de aprendizagem de máquina são utilizadas nos objetos de aprendizagem utilizando a técnica de pesquisa de Mapeamento Sistemático. Os artigos analisados pertencem ao acervo do Simpósio Brasileiro de Informática na Educação (SBIE) abrangendo os anos de 2003 a 2012.

Palavras-chave: aprendizado de Máquina, objetos de aprendizagem, mapeamento sistemático.

Machine Learning Techniques for Learning Objects: A Mapping Study based on Publications of SBIE

Abstract. Researchers from educational area search better practices to apply in classrooms in order to improve the learning process. Generally, tools are proposed to contribute to this area, with the support of development techniques such as Machine Learning. This paper presents a mapping study that aims to map how learning machine techniques are used in learning objects. The articles analyzed are part of the collection of the Brazilian Symposium on Computer in Education (SBIE) between 2003 and 2012.

Keywords: machine learning, learning objects, mapping study.

1. Introdução

A partir do crescimento das pesquisas em informática, atualmente diversos objetos de aprendizagem são desenvolvidos para estimular o aprendizado e as necessidades dos alunos. Este tópico de pesquisa tem recebido bastante atenção dos pesquisadores da área, visto que estes sistemas buscam se adequar às necessidades de seus usuários e

assim, contribuem para o processo de ensino-aprendizagem (DETERS; OLDONI; FERNANDES, 2006).

A definição de um objeto de aprendizagem engloba “*qualquer entidade, digital ou não, que pode ser utilizada e reutilizada durante o processo de aprendizagem que utilize tecnologia. Tais objetos podem ter conteúdo hipermídia, conteúdo instrucional, outros objetos de aprendizagem e software de apoio*” (GAMA, 2007). Os objetos de aprendizagem podem ser desenvolvidos por grupos de alunos, professores, pesquisadores, empresas, profissionais autônomos, dentre outros, e atendem a diversas áreas e disciplinas, expondo diversos tipos de conteúdo, e de diversas maneiras. Estes recursos tendem a proporcionar a obtenção de conhecimento de maneira mais “divertida”, tornando as aulas mais atrativas.

A obtenção de resultados satisfatórios em sala de aula com o uso de objetos de aprendizagem também pode ser ampliada com o emprego de técnicas de Aprendizagem de Máquina, as quais se dedicam ao desenvolvimento de algoritmos e técnicas capazes de tornar uma máquina apta a tomar decisões sem a intervenção humana. A utilização de objetos de aprendizagem neste âmbito tende a aperfeiçoar o estilo do ensino educacional através, por exemplo, de softwares que “acompanhem” o nível de conhecimento do aprendiz, inserindo novos conteúdos à medida que este aprende.

De acordo com Monard e Baranauskas (2003), o aprendizado de máquina é uma área da inteligência artificial que tem como objetivo desenvolver técnicas computacionais sobre o aprendizado, além de possibilitar a criação de sistemas capazes de adquirir conhecimento de maneira automática. Cada sistema de aprendizado possui suas próprias características, não havendo, portanto um algoritmo universal, capaz de solucionar todos os problemas. De acordo com Gama (2007), os objetos de aprendizagem são classificados em: a) Objetos de Instrução: dedicam-se a auxiliar no processo de aprendizagem; b) Objetos de Colaboração: promovem a comunicação em ambientes de aprendizagem colaborativa; c) Objetos de Prática: visam a auto-aprendizagem, tem a interação como ponto forte, e; d) Objetos de Avaliação: desempenham o papel de conhecer o nível de conhecimento de um aprendiz.

Neste sentido, o presente trabalho apresenta um mapeamento sistemático, com o intuito de identificar as técnicas de aprendizagem de máquinas utilizadas em objetos de aprendizagem (OA) baseados em software, bem como verificar a forma como essas técnicas estão sendo utilizadas. Os artigos analisados nesta pesquisa pertencem ao acervo do Simpósio Brasileiro de Informática na Educação (SBIE) abrangendo os anos de 2003 a 2012. Os resultados obtidos neste mapeamento sistemático indicaram quais aspectos dos objetos de aprendizagem baseado em software merecem atenção, de que maneira se dá a aplicação da aprendizagem de máquina nos mesmos e o avanço dos estudos na área.

Além desta introdução, o artigo está organizado nas seguintes seções: a seção 2 apresenta o método de pesquisa aplicado neste trabalho. A seção 3 discute os resultados e a seção 4 apresenta as conclusões e propostas de trabalhos futuros.

2. Metodologia

O método selecionado para esta pesquisa é um mapeamento sistemático (Kitchenham et al., 2004), através do qual busca-se obter dados e informações relevantes para o meio acadêmico em geral a partir da análise de publicações referentes ao tema pesquisado.

Arksey e O'Malley (2005), afirmam que um mapeamento sistemático busca identificar pesquisas relacionadas a um tópico específico para responder perguntas amplas, que se referem ao progresso da investigação e geralmente com o intuito de responder questões exploratórias. O processo para a execução deste mapeamento seguiu seis etapas, que podem ser observadas na figura 2, que também apresenta os resultados produzidos por cada etapa executada.

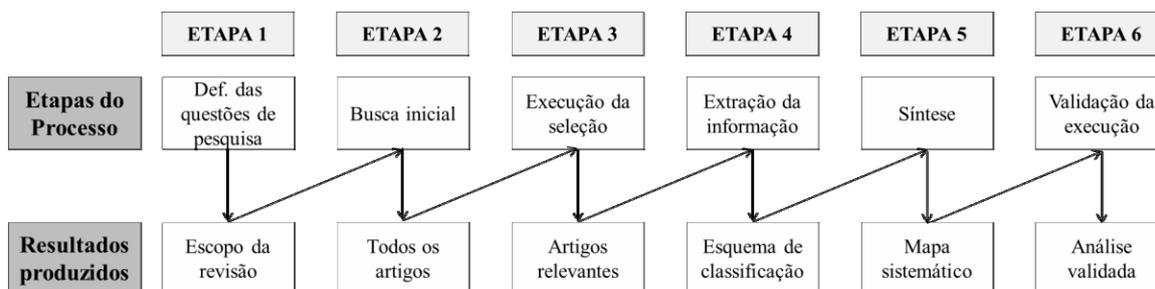


Figura 1. Etapas do Processo de Mapeamento Sistemático

A seguir são apresentadas as atividades que foram realizadas em cada etapa do mapeamento sistemático.

Etapa 1 – Definição das questões de pesquisa: Nesta etapa foi definido o protocolo que define os critérios de inclusão e exclusão e as questões de pesquisa. Os artigos selecionados pertencem ao acervo do SBIE, que foi escolhido por ser o principal fórum no Brasil, na área de educação e informática. As questões de pesquisa foram divididas em três facetas descritas a seguir:

- **Faceta histórica:** como se deu a evolução histórica das publicações em autores, organizações e quantidade no SBIE?
- **Faceta de técnicas:** que técnicas de aprendizagem de máquina são mais utilizadas?
- **Faceta de contribuição:** qual o tipo de contribuição dos objetos desenvolvidos ou propostos?

Etapa 2 – Busca inicial: A aquisição dos artigos foi realizada através do acesso ao portal com os anais do SBIE, do qual foram extraídos os artigos completos dos anos de 2003 a 2012.

Etapa 3 – Execução da seleção: Com base nos critérios de inclusão e exclusão, definidos no protocolo, os artigos foram excluídos do estudo quando não indicavam a utilização de técnicas de aprendizagem de máquina ou caso não se tratasse de um objeto de aprendizagem; caso contrário, os artigos foram incluídos. Para tanto, foi necessária a leitura em parte ou total dos mesmos para comprovar, ou não, a sua inclusão.

Etapa 4 – Extração de dados: Nesta etapa foram extraídas as informações relevantes do trabalho como o tipo de técnica de aprendizagem aplicada no estudo e nome de autores e instituições, sendo organizadas em planilhas para uma melhor manipulação dos dados;

Etapa 5 – Síntese: A partir dos resultados extraídos na fase anterior, foi realizada a criação de tabelas e gráficos que apresentam os dados extraídos na pesquisa.

Etapa 6 – Validação da execução: a execução desta pesquisa incluiu a presença de outros autores para validar a análise previamente estabelecida.

3. Resultados e Discussões

A busca total dos estudos primários resultou em um total de 759 artigos completos, dentre os anos de 2003 e 2012. Por falta de espaço, os artigos selecionados para esta pesquisa não foram acrescentados como apêndice. Após a análise do conteúdo destes estudos, o conjunto de 38 artigos relacionados com o objetivo da pesquisa foram utilizados para extrair as informações.

Faceta Histórica. Como se deu a evolução histórica das publicações, autores e organizações e quantidade no SBIE? A faceta histórica permitiu verificar que há uma frequência de estudos publicados sobre aprendizagem de máquina em objetos de aprendizagem a cada ano, com média de 3,8 artigos nos últimos dez anos, sendo que maior parte destes se concentraram em universidades da região sudeste.

Q1. Quais são os autores que mais publicam sobre a utilização de aprendizagem de máquina em objetos de aprendizagem?

Ao todo, 116 autores publicaram estudos sobre o tópico. A tabela 2 destaca os 17 autores que publicaram pelo menos 2 artigos referentes à utilização de técnicas de aprendizagem de máquina e objetos de aprendizagem, e os 99 autores restantes publicaram artigos abordando estes tópicos apenas uma vez.

Tabela 1. Autores que mais publicam sobre o uso de aprendizagem de máquina em OA.

AUTORES	ARTIGOS	AUTORES	ARTIGOS	AUTORES	ARTIGOS
Weber Martins	6	Ig Bittencourt	2	Fernando Mendes de Azevedo	2
Lauro E. G. Nalini	5	Viviane M. Gomes	2	Marta Costa Rosatelli	2
Eloi Luiz Favero	3	Márcia G. de Oliveira	2	Alexandre Direne	2
Evandro Costa	3	Elias Oliveira	2	Roberto L. de Oliveira Júnior	2
Cecília D. Flores	2	Andréia G. S. Brito	2	Ahmed A. A. Esmin	2
Andréia Tereza Riccio Barbosa	2	Leonardo C. Brito	2		

Q2. Quais são as organizações mais ativas em participações no SBIE?

A tabela 3 apresenta as 13 instituições com pelo menos 2 publicações sobre o tópico no SBIE aos longo dos últimos 10 anos.

Tabela 2. Instituições que mais publicam no SBIE nos últimos 10 anos

INSTITUIÇÕES	ARTIGOS	INSTITUIÇÕES	ARTIGOS	INSTITUIÇÕES	ARTIGOS
UFG	7	UNISANTOS	2	UFAL	3
UFSC	4	UFPR	2	UNISINOS	2
UFRGS	3	UCG	2	ITA	2
UFPA	3	UFES	2	UFSCar	2
UFLA	2				

Pode-se perceber que há uma correlação entre as organizações e autores mais ativos em participações neste evento. É importante também destacar que a totalidade de

publicações de uma organização pode ser menor que a soma de publicações por autor da mesma, já que é comum haver uma mesma publicação com diferentes autores da mesma instituição. A partir destes dados também se pode verificar a distribuição regional quanto à participação de autores e instituições (incluídos nesta pesquisa) que publicaram no SBIE, como se pode ver na figura 3.1:

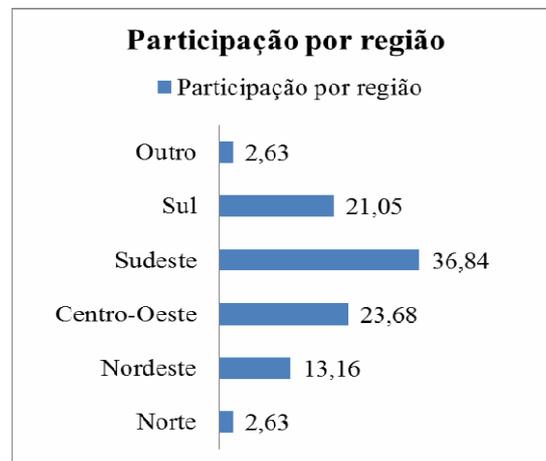


Figura 2. Distribuição regional de artigos incluídos

Q3. Como se deu a evolução da quantidade de artigos publicados no SBIE?

A Figura 3 mostra a evolução da quantidade de artigos completos publicados entre 2003 e 2012. Percebe-se um aumento na quantidade de artigos a partir do ano de 2008, sendo que a maior concentração ocorreu, de forma significativa, no ano 2009, a partir daí houve um pequeno declínio, que se manteve estável nos anos seguintes.

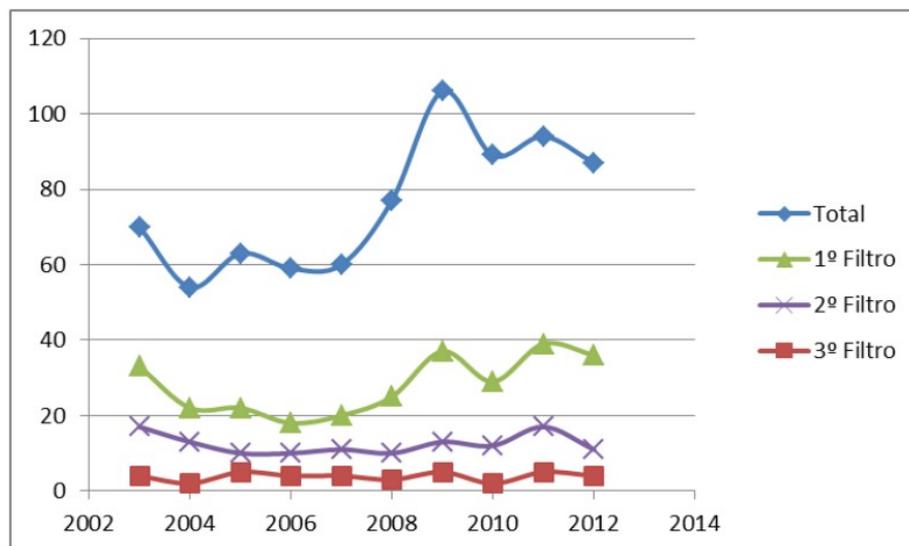


Figura 3. Evolução da quantidade de artigos publicados entre 2003 e 2012

A quantidade de artigos incluídos nesta pesquisa não mostrou ser proporcional a quantidade de artigos publicados em seus respectivos anos. A utilização de técnicas de aprendizagem de máquina em objetos de aprendizagem não apresentaram um progresso contínuo com o passar dos anos. Apesar disso, em todos os anos, houve a presença de pelo menos 2 estudos relacionados a tal tema.

Faceta de Contribuição. Como se deu o desenvolvimento dos objetos quanto à classificação dos objetos de aprendizagem?

Q4. Qual o tipo de contribuição dos objetos desenvolvidos ou propostos?

Como se percebe na figura 5, as técnicas de aprendizagem de máquina são mais aplicadas em objetos de aprendizagem do tipo instrucional, o qual se dedica a auxiliar no processo de aprendizagem. Vale ressaltar que, em alguns casos, um mesmo software foi classificado em mais de um tipo de objeto.

No caso do objeto *instrucional*, os mesmos eram providos de diferentes mídias, como áudio, vídeo e imagens, sendo estas ferramentas, métodos propostos para a contribuição de um modelo de aprendizado mais eficiente. Objetos do tipo *avaliativo*, e *colaborativo* foram a segunda maior ocorrência dentre os estudos classificados.

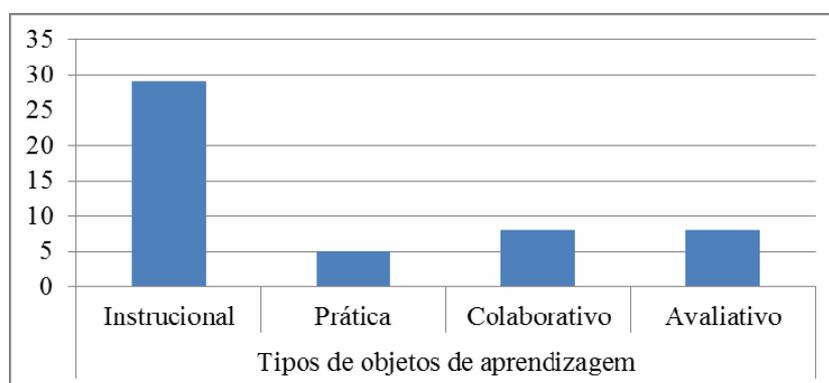


Figura 4. Objetos de aprendizagem mais aplicados nos estudos inclusos

Objetos de aprendizagem do tipo *prática* visam realizar virtualmente, situações que talvez não fossem possíveis de reproduzir em ambientes reais como em salas de aula. Esta foi a menor ocorrência entre os tipos de OAs identificados nesta pesquisa. Este resultado identifica a importância que se tem dado ao desenvolvimento de objetos de aprendizagem que automatize e possibilite um maior dinamismo nas atividades, que eram antes executadas apenas a lápis e papel. Desta forma, a resolução de exercícios propostos em sala de aula utilizando meios tecnológicos, bem como a exibição de vídeos, por exemplo, tem sido vistos como uma forma viável de favorecer aprendizes e acrescer o seu conhecimento.

Faceta de Técnicas. Que técnicas de aprendizagem de máquina são mais utilizadas?

Q5. Quais técnicas de aprendizagem de máquina são mais aplicadas nos objetos de aprendizagem?

Como se pode perceber, na Figura 6, todas as dez técnicas de aprendizagem de máquina fizeram parte de pelo menos 1 estudo de acordo com os critérios de inclusão para esta pesquisa. As que mais se destacaram foram o *clustering*, RNA e aprendizado bayesiano; ao passo que rede semântica e algoritmos evolutivos foram aplicadas, cada uma, em apenas um estudo.

As técnicas de árvores de decisão, k-NN e RBC foram identificadas em 9% dos trabalhos inclusos. Pode-se dizer que em comparação com as outras técnicas, estas tiveram uma participação mediana, incluindo o aprendizado por reforço que foi incluso em 11% dos trabalhos. Como se pode perceber ainda, a técnica de rede semântica, SVM

e algoritmos evolutivos tiveram uma baixa porcentagem quanto a utilização nos objetos de aprendizagem.

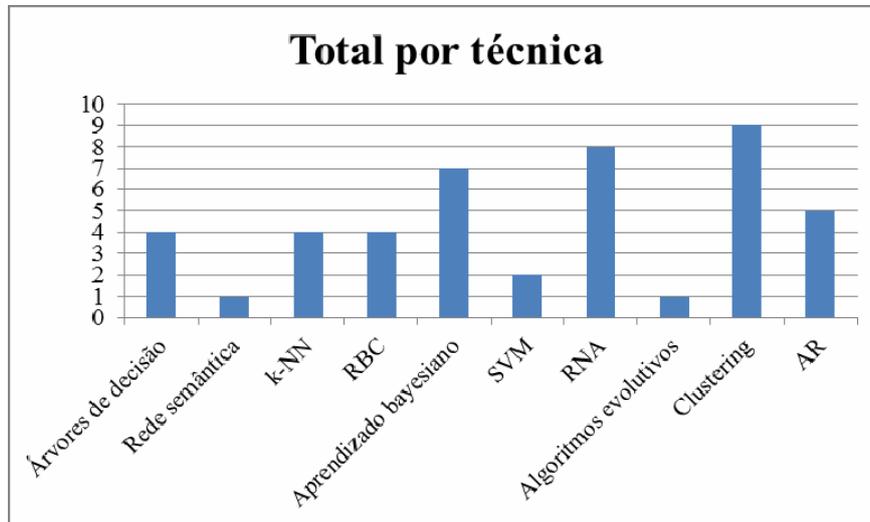


Figura 5. Utilização das técnicas entre 2003 e 2012

As técnicas que mais foram aplicadas e tiveram uma participação a partir de 16% foram o aprendizado bayesiano, as RNA e o *clustering*, sendo que houve uma maior aparição deste último. A boa participação destas talvez possa indicar que elas têm benefícios a oferecerem quanto a este tipo de emprego (objetos de aprendizagem).

Q6. Qual a correlação das técnicas com cada um dos tipos de objetos de aprendizagem?

A tabela 4 apresenta as correlações entre as técnicas de aprendizagem de máquina e os tipos de objetos de aprendizagem:

Tabela 3. Correlação técnica de aprendizagem vs objeto de aprendizagem

TÉCNICAS	TIPOS DE OBJETOS DE APRENDIZAGEM			
	INSTRUCIONAL	PRÁTICA	COLABORATIVO	AVALIATIVO
Árvores de decisão	2	1	0	1
Rede semântica	1	0	0	0
KNN	2	0	1	1
RBC	2	1	1	1
Aprendizado bayesiano	6	1	1	0
SVM	0	0	2	0
RNA	7	0	1	0
AE	0	0	0	1
Clustering	5	1	2	3
AR	3	1	0	1

A existência dos tipos de aprendizagem de máquina nos objetos de aprendizagem não variou muito de acordo com os tipos de objetos, uma vez que a maioria delas se concentrou no do tipo instrucional. As técnicas que mais se destacaram quanto ao tipo deste objeto foram RNA, aprendizado bayesiano e *clustering* com 7, 6 e

5 artigos, respectivamente, estas são técnicas já bem conhecidas e difundidas na área da aprendizagem de máquina.

As RNA apesar de serem bastante associadas ao reconhecimento de padrões foram mais empregadas com o propósito de apresentar um conteúdo que se adequasse ao perfil de cada estudante individualmente, atentando, portanto, para o nível de conhecimento de cada um e provendo uma maneira de instrução mais específica, diferentemente de como ocorre em salas de aula onde o professor ministra as aulas visando o conjunto de alunos independentemente do nível de conhecimento de cada um.

O aprendizado bayesiano busca prever melhores resultados a partir de dados estatísticos, ou seja, se baseia na probabilidade para verificar a ocorrência de cada ação provável. Desta forma, um objeto do tipo instrucional poderia trabalhar em ambientes onde os resultados são incertos, de fato, pode-se perceber sua viabilidade uma vez que um objeto instrucional, diferentemente da prática, por exemplo, não contém um resultado previsto em alguns casos, antes estes são dados a partir da atuação do aprendiz.

A terceira técnica que mais se aplica ao caso dos objetos instrucionais é o *clustering*. O objetivo dela é agrupar objetos ou dados semelhantes; sendo assim, a utilização deste método se torna apropriado quando o conhecimento sobre a composição dos dados é pouca ou nenhuma, por este motivo, mais uma vez, pode-se afirmar que esta se adequa à finalidade de desenvolver um objeto instrucional.

Ainda com respeito ao *clustering*, este apresentou dois artigos dentro dos objetos colaborativos. Sabendo-se que a finalidade deste tipo de objeto é promover o aprendizado “em conjunto”, ou com a colaboração de outros participantes no mesmo sistema, percebeu-se como forma de melhoria para o aprendizado a criação de objetos que agrupem alunos com o mesmo nível de instrução para que o progresso no aprendizado ocorra de maneira igualitária e não irregular.

Quanto ao objeto avaliativo, o *clustering* aparece como a técnica mais empregada quando o objetivo é avaliar o nível de conhecimento dos aprendizes. A avaliação com esta técnica permite que o professor detenha do nível de conhecimento dos alunos e assim providencie melhorias que venham beneficiar toda a turma envolvida na aprendizagem dos mesmos conteúdos, uma vez que está engloba o nível de conhecimento destes, o professor poderá nivelar o mesmo por preencher as lacunas no aprendizado de alunos com o nível de conhecimento mais baixo.

Outras técnicas como a árvore de decisão, k-NN, RBC e aprendizado por reforço mantiveram uma quantidade equivalente em relação à presença dentre os estudos inclusos variando entre 5 e 4, porém, não tão significativa quando referente a um tipo de objeto de aprendizagem específico, embora sua maior parte tenha estado presente dentre os objetos instrucionais.

Uma árvore de decisão pode ser usada para a classificação de novos exemplos gerando, assim, novas regras. Justamente para esta finalidade foi utilizada a técnica, visando a classificação/identificação de dados e padrões dentro dos objetos do tipo instrucional, de prática e avaliativo. As técnicas estatísticas de k-NN e RBC também não demonstraram ter um significado maior para determinado tipo de objeto, o que não nos permitiria tirar conclusões quanto ao seu uso e correlação técnica-objeto.

A utilização do aprendizado por reforço pode indicar a busca por uma maneira mais eficiente de atuação em um ambiente educacional, visto que esta técnica atua a partir de resultados positivos dentro do sistema, pode-se dizer que esta é a técnica mais independente de todas pelo fato de não conter a priori nenhum dado de entrada pelos quais venha a se basear para prover reações e resultados se saída. Talvez o fato de esta demorar em prover resultados por meio do aprendizado do agente, tenha influenciado na não utilização em mais pesquisas.

4. Conclusão e Trabalhos Futuros

De um total de 759 trabalhos completos publicados entre 2003 e 2012 no SBIE, a busca manual para este mapeamento sistemático resultou em um total de 38 artigos que fazem uso de técnicas de aprendizagem de máquina em objetos de aprendizagem baseado em software. Com isto, foi possível responder a todas as questões de pesquisa definidas no protocolo de revisão, uma vez que se obteve uma quantidade suficiente de artigos, através dos quais foi possível fornecer as respostas para cada questão.

O desenvolvimento deste trabalho poderá servir como fonte de dados para futuras pesquisas a respeito da aprendizagem de máquina como parte de objetos de aprendizagem, indicando que aspectos destes sistemas merecem atenção, de que maneira se dá a aplicação da aprendizagem de máquina nestes sistemas e o avanço dos estudos na área.

A principal contribuição a ser destaca dentro desta pesquisa foram os resultados referentes às técnicas de aprendizado de máquina e sua relação com os objetos de aprendizagem, que poderão servir de base para futuras pesquisas, bem como, para o desenvolvimento de objetos ou outro software, levando em consideração os objetivos de cada técnica adotada nos artigos aqui vistos.

Este trabalho também poderá servir de base para que sejam feitas outras pesquisas, dado que não foram avaliadas todas as informações disponíveis e relevantes para a contribuição do meio acadêmico como, por exemplo, com respeito à quais tipos de pesquisas eram desenvolvidas. Outro tipo de questão que se pode levar em conta para trabalhos futuros seria ampliar a execução do mapeamento para todas as vinte e duas edições do SBIE, o que, provavelmente, permitiria ampliar a quantidade de estudos inclusos, bem como uma ratificação (ou porque não, uma “retificação”?) dos resultados obtidos neste trabalho.

Ainda há espaço para ampliar a pesquisa por incluir a análise de trabalhos em outros idiomas, como o inglês e espanhol, bem como outras fontes de pesquisa científica. Também será possível identificar de maneira mais profunda as técnicas e onde exatamente se encaixam suas aplicações.

Referências

- ARKSEY, H.; O'MALLEY, L. Scoping studies: towards a methodological framework. **International Journal of Social Research Methodology**, 8(1), 19-32, 2005
- SBIE. Comissão Especial de Informática na Educação da SBC. **Anais do SBIE**. Acesso em maio de 2012. Disponível em < <http://www.br-ie.org/index.php/anais/bie>>.
- DETERS, J. I.; OLDONI, A.; FERNANDES, A. M. R. **Bernardo - Agente Pedagógico do Sistema Tutor Inteligente Aplicado a Neurofisiologia**. In: SBIE – Simpósio Brasileiro de Informática na Educação, Brasília, DF, 2006.



- GAMA, C. L. G. **Método de Construção de Objetos De Aprendizagem Com Aplicação em Métodos Numéricos**. Tese (Métodos Numéricos em Engenharia). Universidade Federal do Paraná, Curitiba, 2007.
- KITCHENHAM, B.; DYBÅ, T.; JØRGENSEN, M. **Evidence-based Software Engineering**. In: ICSE - International Conference on Software Engineering, Washington DC, USA, pp 273 – 281, 2004.
- MONARD, M. C. AND BARANAUSKAS, J. A. **Conceitos de aprendizado de máquina**. In: REZENDE, S.O. Sistemas inteligentes: fundamentos e aplicações. Barueri, SP: Manole, cap.4, p.89-94, 2003.