

---

**USO DE REDES NEURAS ARTIFICIAIS NA PREDIÇÃO DE RISCO DE EVASÃO ACADÊMICA**

**THE USE OF ARTIFICIAL NEURAL NETWORKS FOR PREDICTION OF THE RISK OF ACADEMIC DROPOUT**

Andriw Romeiro Superbe \*  
Ricardo Inacio Alvares e Silva \*\*

**RESUMO**

O abandono da vida acadêmica pode se dar por diversos motivos, como a retenção ou reprovação. Este é um dos problemas mais complexos e cruciais na educação, causando prejuízos sociais, econômicos, políticos, acadêmicos e financeiros. Tais prejuízos podem ser eliminados se for possível impedir a evasão dos alunos por meio da intervenção precoce por parte das instituições de ensino. Uma das diversas aplicações da Inteligência Artificial é a predição de acontecimentos com base na probabilidade estatística da ocorrência de determinados eventos pela filtragem de dados brutos coletados. Este trabalho propõe o uso de Redes Neurais Artificiais para a predição de risco de retenção acadêmica baseando-se em dados reais de desempenho de alunos de graduação em nível superior. Para esse fim, serão estudados os desempenhos dos alunos de um curso de graduação de uma faculdade localizada em Londrina, Paraná, buscando atingir eficácia acima do nível de 80% nas previsões, quando comparadas aos dados reais.

160

**Palavras-chave:** Redes neurais artificiais. Predição de risco. Desempenho acadêmico. Mineração de dados educacionais. Evasão acadêmica.

**ABSTRACT**

The abandonment of academic life can occur for several reasons, such as retention or failure. This is one of the most complex and crucial problems in education, causing social, economic, political, academic and financial damages. Such losses can be eliminated if it's possible to prevent student evasion through early intervention by the educational institutions. One of the several applications of Artificial Intelligence is the prediction of events based on the statistical probability of occurrence of these certain events by the filtering of raw data collected. This paper proposes the use of Artificial Neural Networks for the prediction of risk of academic retention based on real performance data of undergraduate students at a higher level. To that end, the students' performance in an undergraduate course at a university located in

---

\* Autor: Andriw Romeiro Superbe – E-mail: mandriw.m@gmail.com

\*\* Orientador: Ricardo Inacio Alvares e Silva – E-mail: ricardo.silva@unifil.br

Londrina-PR will be studied, aiming to reach an accuracy above the level of 80% in the predictions, when compared to real data.

**Keywords:** Artificial neural networks. Risk prediction. Academic performance. Educational data mining. Academic evasion.

## 1 INTRODUÇÃO

O uso de inteligência artificial tem crescido bastante em diversas áreas nos últimos anos, dentre elas a da educação. O sucesso de um estudante em qualquer nível de formação, do ensino fundamental ao superior, é de grande importância para as instituições de ensino, portanto muitas delas têm adotado métodos inovadores em busca de melhorar a experiência educacional e, conseqüentemente, aumentar a taxa de sucesso de seus alunos (SHEEHAN *et al.*, 2012). A Mineração de Dados Educacionais (MDE) é uma área emergente, que se preocupa em explorar os dados de desempenho acadêmico e, com base nas informações coletadas, encontrar as falhas e, por conseqüência, as melhores soluções para otimizar a tomada de decisões das instituições de ensino e seus administradores (LIVIERIS *et al.*, 2012; LAKKARAJU *et al.*, 2015; SHEEHAN *et al.*, 2012).

Após o tratamento dos dados coletados das próprias instituições sobre o desempenho de seus alunos, pode-se apontar quais deles precisam de mais atenção, e em quais matérias, para que seja possível evitar retenção ou reprovação do aluno, com a conseqüência de melhorar a taxa de aprendizado devido a maior atenção direcionada a ele. Alguns estudos buscam até mesmo entender as variáveis que afetam o aprendizado do aluno fora da instituição, considerando pontos como o tempo de leitura de um determinado texto relacionado a disciplina em uma plataforma web, o tempo de execução de uma atividade avaliativa, qual tema o aluno tem mais dificuldade em assimilar, tempo disponível para o estudo em casa, entre outros (ESSA *et al.*, 2012).

Condições ambientais podem interferir de formas bastante relevantes no aprendizado dos alunos e, por isso, encontrar meios de contorná-las pode ser uma das chaves para definir a melhor maneira de ajudar o aluno em sua jornada acadêmica. Porém, analisar as variáveis ambientais acaba por ser uma tarefa muito

complexa e imprecisa, pois exige acompanhamento da rotina de atividades do aluno fora do ambiente acadêmico, ou seja, em seu ambiente pessoal e reservado. Observações desse tipo podem causar imprecisão nos resultados devido às diferentes variáveis que podem compor a rotina de cada um dos diferentes alunos, além das questões éticas relacionadas à privacidade dos mesmos. Devido a tais complicações, este estudo não busca analisar esse tipo de variáveis, mas apenas os números diretamente ligados aos alunos dentro de suas respectivas instituições de ensino em determinado período de tempo.

É preciso perceber a dificuldade de cada aluno de forma individual para que a coordenação do curso, ou mesmo os professores, possam intervir de forma mais adequada para cada caso. É necessário que essa decisão seja tomada de forma rápida e precisa para que haja tempo de ajudar o aluno (LAKKARAJU *et al.*, 2015). Uma das abordagens existentes para auxiliar nessa questão é justamente o uso de Redes Neurais Artificiais para prever quais alunos podem acabar deixando a instituição.

162

O modelo de Rede Neural Artificial proposto por (WEI, 2011) obteve 79,1% de eficácia na predição do sucesso acadêmico da amostra de alunos selecionados para o experimento. O sistema de predição de risco desenvolvido por (MARTINHO *et al.*, 2013) utilizando a arquitetura de Rede Neural Fuzzy- ARTMAP chegou a 76% de eficácia nos experimentos realizados em seu artigo.

O objetivo deste trabalho é realizar o mapeamento das diversas técnicas já propostas no meio, a fim de auxiliar universidades pelo país no que corresponde à qualidade de ensino e suporte ao aluno, buscando compreender o cenário nacional e aperfeiçoar os modelos de predição de risco de retenção ou reprovação acadêmica por meio do uso de Redes Neurais Artificiais, que têm se mostrado muito promissoras na área de predição de risco (LIVIERIS *et al.*, 2012; PLAGGE, 2013; NURHUDA *et al.*, 2017; WEI, 2011; MARTINHO *et al.*, 2013; ARSAD *et al.*, 2013; KAZUMALI *et al.*, 2017).

Para validação da solução desenvolvida, os dados reais de desempenho acadêmico de alunos de determinado curso de graduação em uma faculdade da cidade de Londrina, Paraná, serão aplicados à Rede Neural Artificial estruturada com base na melhor arquitetura encontrada para o caso em questão. O nome do

curso e da instituição de ensino responsável pelo fornecimento dos dados serão mantidos no anonimato.

Serão registrados os dados obtidos por cada modelo implementado, comparando seus resultados com o cenário real, para que seja possível obter suas respectivas taxas de eficácia quando aplicados ao cenário em questão. Após selecionar o melhor modelo, objetiva-se sua otimização, buscando predições mais satisfatórias, caso os resultados obtidos já não cumpram esse característica.

## **2 MODELOS DE PREDIÇÃO APLICADOS À EDUCAÇÃO**

O abandono da vida acadêmica pode se dar por diversos motivos, como a retenção ou reprovação. Este é um dos problemas mais complexos e cruciais na educação, causando prejuízos sociais, econômicos, políticos, acadêmicos e financeiros (MARTINHO *et al.*, 2013). Na maioria dos casos, o problema não é percebido de forma precoce, o que dificulta a recuperação do aluno ao final do ano letivo e, mesmo quando o consegue, é subentendido que o aluno não tem total domínio da matéria.

O resultado deste trabalho pode ser usado na prática pelo corpo docente da instituição como ferramenta de supervisão da situação individual de cada um dos alunos do curso, tendo em vista diagnosticar precocemente os casos em que pode haver reprovação dos alunos. Detectando antecipadamente esses casos, a coordenação ou mesmo os professores do curso, podem dar mais atenção ao caso em questão e tomar providências, como averiguar quais causas provocam a dificuldade do aluno em determinadas disciplinas e, ao analisar o caso junto do aluno, encontrar a melhor solução possível para seu problema específico.

Dessa forma, tanto o aluno quanto a instituição se beneficiam, pois o aluno tem maior chance de concluir o curso com sucesso, assimilando as disciplinas de forma mais condizente com suas dificuldades, e a instituição aumenta seu número de alunos aprovados, ou seja, aumenta a qualidade do curso e, conseqüentemente, da própria universidade.

### **3 METODOLOGIA**

A metodologia utilizada neste trabalho se baseia nos modelos já concebidos e busca encontrar o que melhor se adapta ao cenário de estudo, tendo como próximo passo a busca pelas possíveis falhas na arquitetura e, conseqüentemente, pela melhor forma de otimizá-la, buscando nível de eficácia maior que 80%.

Para a criação de uma ferramenta de predição como a tratada nesta pesquisa, deve-se considerar que cada aluno pode sofrer influências externas diferentes de outros, como trabalho, relacionamento social e familiar e o tempo e ferramentas das quais o aluno dispõe para se dedicar ao curso fora da sala de aula, por exemplo. Por essa razão, o perfil de aluno alvo da pesquisa será definido para determinar as limitações do trabalho frente às variáveis de influência externa.

Após a definição de perfil alvo, os dados disponibilizados pela faculdade serão organizados em tipos, e a quantidade de tipos de dados relevantes selecionados ajudará a determinar a quantidade de neurônios que serão utilizados na camada de entrada. Os dados serão então tratados para serem apresentados à rede neural a ser treinada.

Pontos fundamentais em uma rede neural são seu tipo de arquitetura e a quantidade de camadas a serem utilizadas e sua topologia. As opções são diversas, portanto vários modelos serão testados e seus resultados serão comparados após o devido treinamento de cada uma delas.

Quando observada vantagem de um ou mais modelos sobre os outros, o próximo passo será tentar determinar quais características do modelo tornam seus resultados mais precisos. Com a identificação de tais características, pode ser possível aplicá-las aos outros modelos e, caso as atualizações aplicadas às demais redes surtam efeito, esse processo de análise e atualização das será executado de forma recursiva até que seus efeitos não sejam mais significativos.

## REFERÊNCIAS

ARSAD, P. M.; BUNIYAMIN, N.; MANAN, J. I. A. A neural network students' performance prediction model (nnsppm). In: 2013 IEEE INTERNATIONAL CONFERENCE ON SMART INSTRUMENTATION, MEASUREMENT AND APPLICATIONS (ICSIMA). 2013, Kuala Lumpur, Malaysia, **Proceedings...** Kuala Lumpur, Malaysia: IEEE, 2013. p. 1–5.

ESSA, A.; AYAD, H. Student success system: Risk analytics and data visualization using ensembles of predictive models. In: PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE. 2012, New York. **Proceedings...** New York: ACM, 2012. p. 158–161. Disponível em: <<http://doi.acm.org/10.1145/2330601.2330641>>. Acesso em: 18 jun. 2018.

KAZUMALI, E.; KALINGA, E. Neural network model for predicting students' achievement in blended courses at the university of dar es salaam. **International Journal of Artificial Intelligence and Applications**, Tanzania, v. 8, n.2, p. 23–35, 03 2017.

LAKKARAJU, H. et al. A machine learning framework to identify students at risk of adverse academic outcomes. In: PROCEEDINGS OF THE 21TH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING. 2015, New York. **Proceedings...** New York: ACM, 2015. p. 1909–1918. Disponível em: <<http://doi.acm.org/10.1145/2783258.2788620>>. Acesso em: 18 jun. 2018.

165

LIVIERIS, I. E.; DRAKOPOULOU, K.; PINTELAS, P. **Predicting students' performance using artificial neural networks**. In: Proceedings of the 8th Pan-Hellenic Conference "Information and Communication Technology in Education". Patras, Peloponnese, Greece: UNIÃO CIENTÍFICA HÉLICA DE TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO EM EDUCAÇÃO, 2012. Disponível em: <<http://www.etpe.gr/custom/pdf/etpe1938.pdf>>. Acesso em: 18 jun. 2018.

MARTINHO, V. R. C.; NUNES, C.; MINUSSI, C. R. Prediction of school dropout risk group using neural network. In: 2013 FEDERATED CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS. 2013, Poland. **Proceedings...** Poland: IEEE, 2013. p. 111–114.

NURHUDA, A.; ROSITA, D. Prediction student graduation on time using artificial neural network on data mining students stmik widya cipta dharma samarinda. In: PROCEEDINGS OF THE 2017 INTERNATIONAL CONFERENCE ON E-COMMERCE, E-BUSINESS AND E-GOVERNMENT. 2017, New York. **Proceedings...** New York: ACM, 2017. p. 86–89. Disponível em: <<http://doi.acm.org/10.1145/3108421.3108431>>. Acesso em: 18 jun. 2018.

PLAGGE, M. Using artificial neural networks to predict first-year traditional students second year retention rates. In: PROCEEDINGS OF THE 51ST ACM SOUTHEAST CONFERENCE. 2013, New York. **Proceedings...** New York: ACM, 2013. p. 1–5. Disponível em: <<http://doi.acm.org/10.1145/2498328.2500061>>. Acesso em: 18 jun. 2018.

SHEEHAN, M.; PARK, Y. pgpa: A personalized grade prediction tool to aid student success. In: PROCEEDINGS OF THE SIXTH ACM CONFERENCE ON RECOMMENDER SYSTEMS. 2012, New York. **Proceedings...** New York: ACM, 2012. p. 309–310. Disponível em: <<http://doi.acm.org/10.1145/2365952.2366027>>. Acesso em: 18 jun. 2018.

WEI, X. Student achievement prediction based on artificial neural network. In: 2011 INTERNATIONAL CONFERENCE ON INTERNET COMPUTING AND INFORMATION SERVICES. 2011, Hong Kong, China . **Proceedings...** Hong Kong, China: IEEE, 2011. p. 485–487.