

PREDIÇÃO DO GRUPO DE RISCO DE EVASÃO DISCENTE EM CURSOS SUPERIORES PRESENCIAIS UTILIZANDO UMA REDE NEURAL ARTMAP-FUZZY

Valquíria R. C. Martinho - vribeiro@terra.com.br

Instituto Federal de Mato Grosso - IFMT – Departamento de Eletro-Eletrônica

Rua Zulmira Canavarros, nº. 95 - Centro

CEP: 78000-000 - Cuiabá - MT

Clodoaldo Nunes - cncefet@gmail.com

Instituto Federal de Mato Grosso - IFMT – Departamento de Informática

Rua Zulmira Canavarros, nº. 95 - Centro

CEP: 78000-000 - Cuiabá - MT

Carlos Roberto Minussi - minussi@dee.feis.unesp.br

UNESP - Campus de Ilha Solteira – Departamento de Engenharia Elétrica

Av. Brasil 56, PO Box 31

CEP: 15385-000 - Ilha Solteira - SP

Resumo: *A evasão escolar é um dos problemas mais complexos e cruciais no âmbito da educação. Perpassa e aflige os vários níveis e modalidades de ensino. No que tange ao ensino superior, internacionalmente, o fenômeno é objeto de atenção e cuidado, no intuito de aumentar os índices de permanência e conclusão dos estudantes de graduação e minimizar os prejuízos sociais, econômicos, políticos, acadêmicos e financeiros causados a todos os envolvidos no processo educacional. Nesse contexto, é imprescindível o desenvolvimento de métodos e instrumentos eficientes e eficazes para predição, avaliação e acompanhamento de estudantes em risco de evasão, possibilitando o planejamento de ações proativas para o saneamento da situação. Assim sendo, este trabalho, tem por objetivo apresentar as potencialidades de um sistema inteligente, desenvolvido para a predição de grupos de risco de evasão discente, em cursos superiores presenciais, utilizando uma Rede Neural ARTMAP-Fuzzy, o que possibilita o aprendizado continuado do sistema. A pesquisa que corrobora este trabalho foi desenvolvida nos cursos de tecnologia do Instituto Federal de Mato Grosso, baseada na aquisição, compilação e transformação de dados a partir de dois bancos de dados, com registros acadêmicos e socioeconômicos dos estudantes. A consistência dos resultados, com índice de acerto do grupo evasivo em torno de 97% e acerto global superior a 76%, demonstra a confiabilidade, a acurácia e a eficiência do sistema. Destaca-se que esta proposta é arrojada e inovadora, pois pode-se identificar precocemente, ainda no primeiro semestre do curso, o grupo de estudantes em risco de evasão e suas possíveis causas.*

Palavras-chave: *Educação superior, Predição da evasão escolar, Proatividade, Sistema inteligente, Rede neural ARTMAP-Fuzzy.*

1. INTRODUÇÃO

Ao longo da história da Educação, o fenômeno da evasão escolar permeia os vários níveis e modalidades de ensino e tem gerado prejuízos sociais, econômicos, políticos, acadêmicos e financeiros a todos os envolvidos no processo educacional. É um problema complexo e

crucial (MARTINHO, 2012), advindo da superposição de fatores sociais, econômicos, culturais e acadêmicos, além de variáveis demográficas e atributos individuais que influenciam na decisão do estudante universitário de permanecer ou abandonar o curso.

Diante dessa situação, nos últimos anos, no Brasil, foram realizados monitoramentos, estudos empíricos, pesquisas científicas e levantamentos estatísticos sobre a evasão escolar, com vistas à adoção de medidas capazes de solucionar ou reduzir os altos índices existentes e, assim, evitar as perdas tanto sociais quanto financeiras.

As contribuições e ações resultantes desses processos podem ser identificadas sob três paradigmas: 1- contribuições teóricas, que visam compreender e identificar os fatores que contribuem e influenciam no fenômeno e no processo da evasão; 2- ações específicas de cada instituição de ensino superior (IES), centrada em sua realidade e especificidade; e, 3- ações governamentais, como o Programa REUNI - Reestruturação e Expansão das Universidades Federais (MEC/SESu, 2007) e o Termo de Acordo de Metas e Compromissos (TAMEC) (MEC/SETEC, 2009), estabelecidos, respectivamente, com as instituições federais de ensino superior (IFES) e os institutos federais (IF).

As ações governamentais se propõem a viabilizar mais recursos para as instituições federais de ensino, mediante o cumprimento de metas preestabelecidas. Na implantação do REUNI e do TAMEC, algumas das metas são estritamente relacionadas à evasão. Têm como objetivo elevar gradualmente a taxa de conclusão dos cursos de nível médio e graduação e, consequentemente, reduzir as taxas de evasão. Como, por exemplo, alcançar a meta mínima de 90% de eficiência (relação entre o número de alunos regularmente matriculados e o número de vagas de cada turma) e 80% de eficácia (relação entre o número de alunos concluintes e o número de vagas ofertadas no processo seletivo para cada uma dessas turmas) da instituição até o ano de 2016.

Mas, os estudos sobre evasão existentes no Brasil, ainda são incipientes e tímidos, se comparados com os países desenvolvidos onde os estudos e pesquisas sistematizados são inúmeros. Nos EUA, mais de 14 mil *sites* tratam dessa questão. Países como a Austrália e, também, os Estados Unidos têm abordado o tema de maneira científica e, por considerá-lo de alta relevância, criaram gerências ou diretorias para a gestão da retenção (PORTELA, 2013). O termo retenção, considerado mais objetivo e proativo, é largamente utilizado na literatura internacional para designar o conjunto de medidas destinadas a manter o estudante na instituição.

Neste trabalho, utilizar-se-á o termo evasão, considerada como a saída definitiva do aluno de seu curso de origem sem concluí-lo, conforme o relatório da Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras (MEC/SESu, 1996).

No Brasil, a educação superior passa por uma avaliação, através do Censo da Educação Superior, realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

Na Figura 1, tem-se os índices da evasão no ensino superior, no período de 2000-2011, obtidos pela análise do relatório do Censo da Educação Superior 2011 (MEC/SESu, 2013) e calculados conforme (CAPELATO, 2011), usando-se Equação (1).

No âmbito do Instituto Federal de Mato Grosso (IFMT), *locus* de desenvolvimento dessa pesquisa, foi realizado estudo estatístico para identificação e análise dos fatores determinantes da evasão escolar no Curso Superior de Tecnologia (CST) em Automação Industrial, no período de 2004-2010, onde se constatou que dos 389 estudantes matriculados no curso, 243 estão evadidos, representando um índice de 62,46% de evasão (BARBOSA, 2010). Este alto índice de evasão se constitui numa tarefa desafiadora, enquanto a meta é reduzi-lo para 10%.

Portanto, a análise atual dos índices da evasão escolar em nível de Brasil demonstra o preocupante estado de fadiga e fragilidade que se encontra a educação superior brasileira e, corrobora com a necessidade de se obter resultados satisfatórios, no sentido de mitigar tal

fenômeno e possibilitar uma educação de qualidade, impulsionando a Nação para um desenvolvimento socioeconômico sustentável, e quiçá, garantir a todo cidadão o direito, que lhe é de direito, o trabalho, a saúde, o lazer, a dignidade, a cidadania e, a tão sonhada felicidade (MARTINHO, 2012).

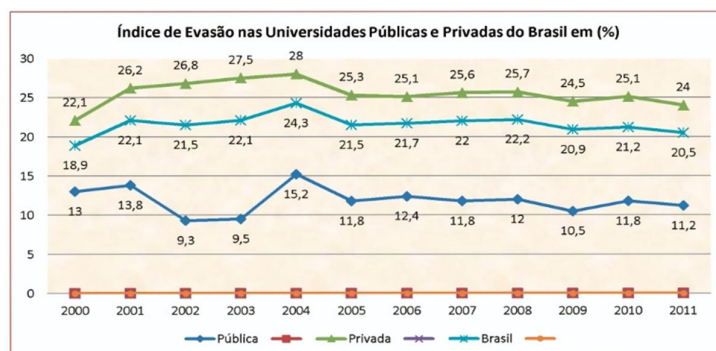


Figura 1 - Índice de Evasão no Brasil de 2000-2011

$$Tx \text{ Evasão (\%)} = \frac{(Mt_{AAnt} - Na_{Con}) - (Mt_{AAt} - Na_{Nov})}{(Mt_{AAnt} - Na_{Con})} \times 10 \quad (1)$$

sendo:

$Tx \text{ Evasão (\%)}$: Taxa de evasão percentual;

Mt_{AAnt} : Matrícula total do Ano Anterior;

Na_{Con} : Número de alunos Concluintes/graduados;

Mt_{AAt} : Matrícula total do Ano Atual;

Na_{Nov} : Número de alunos Novos/ingressantes.

Neste contexto, é imprescindível realizar estudos sistemáticos, observar os sinais de evasão iminente, desenvolver estratégias para favorecer a identificação precoce dos estudantes vulneráveis a evasão, e com isso possibilitar a aplicação de ações proativas no sentido de reverter as intenções de abandono. Os métodos preditores da evasão são factíveis de serem empregados de maneira coletiva. Porém, a ação em prol da permanência do estudante deve ser individualizada, atendendo suas necessidades específicas. Esta convicção vai ao encontro do que enfatiza a UNESCO (2004), de que a evasão pode constituir-se em fenômeno coletivo, mas é sempre num processo individual.

As variáveis envolvidas no processo da evasão guardam em si uma gama de especificidades inerentes aos diferentes níveis de ensino, cursos e instituições que se possa analisar, são complexas, subjetivas, não-lineares e inter-relacionadas. Uma possibilidade factível para representar situações de complexidade, não-linearidade e subjetividade, como o caso da evasão escolar, é a utilização de Redes Neurais Artificiais (RNAs) (HAYKIN, 2009). As Redes Neurais são sistemas computacionais que emulam o cérebro humano na interpretação e processamento das informações e, também, “aprendem” através de experiências, tornando-as capazes de fazer generalizações mais confiáveis. Isso permite o desenvolvimento de sistemas com capacidade para processar problemas intratáveis, lidar com variáveis não-lineares, imperfeitas e faltantes, interagir com dados ruidosos e, mesmo assim, apresentar rapidez de resposta e resultados desejados.

Na literatura, foram encontradas poucas referências pertinente a análise e predição da evasão escolar discente, em cursos presenciais do ensino superior, utilizando redes neurais artificiais. Os trabalhos mais próximos foram aqueles que empregam técnicas de mineração de dados (Data Mining) (SOUZA, 2008; JADRIĆ *et al.*, 2010; PAL, 2012) e Redes Neurais

(CRIPPS, 1996; HERZOG, 2006; KARAMOUZIS E VRETTOS, 2008), invariavelmente uma rede neural Perceptron multicamadas (MLP). Os resultados auferidos nestes trabalhos foram a identificação de padrões descritivos causais que levam os estudantes ao abandono da escola e a constatação da eficiência e desempenho das técnicas empregadas. No trabalho de (LYKOURENTZOU, 2009) é investigado o grupo de risco de estudantes com propensão à evasão em cursos e-learning. Utiliza-se a combinação de três técnicas de aprendizagem de máquina, entre elas a rede neural ARTMAP-*Fuzzy* e, (MUSTAFA, 2012) utiliza árvore de regressão e classificação para identificar a evasão discente, partindo dos dados da inscrição do estudante em um curso presencial. Ele conclui que os dados escolhidos conferem um baixo nível de precisão nos resultados desejados.

Neste trabalho, em específico, tem-se o objetivo de apresentar um sistema inteligente, utilizando uma rede neural ARTMAP-*Fuzzy* (CARPENTER *et al.*, 1992a; CARPENTER *et al.*, 1991; CARPENTER & GROSSBERG, 1992) capaz de realizar a predição do grupo de estudantes em risco de evasão, de maneira acurada e contínua, inferido de maneira fidedigna sobre as condições dos estudantes analisados quanto à evasão. A rede neural ARTMAP-*Fuzzy*, é uma das redes da família ART (*Adaptive Resonance Theory*) (CARPENTER & GROSSBERG, 1992). Possui uma arquitetura cujo treinamento é realizado de modo supervisionado e auto-organizável, com possibilidade de aprendizado continuado (CARPENTER *et al.* 1992b; MARCHIORI *et al.*, 2011). Ressalta-se que, para além das singularidades já descritas deste sistema, frente aos trabalhos científicos sobre a questão, a pujança e a audácia desta pesquisa encontra-se na possibilidade de usar somente os dados contidos no questionário de inscrição para identificar precocemente a iminente evasão.

Após a introdução, na Seção 2 apresenta-se a delimitação da pesquisa e a construção da base de dados. As redes neurais ART e ARTMAP-*Fuzzy* são abordadas na Seção 3. Na Seção 4 descreve-se o desenvolvimento do sistema neural ARTMAP-*Fuzzy* proposto para a predição da evasão discente. A implementação do sistema, os resultados e as análises das simulações são apresentados na Seção 5. As considerações mais relevantes deste estudo estão na Seção 6.

2. DELIMITAÇÃO DA PESQUISA E CONSTRUÇÃO DA BASE DE DADOS

Os cursos analisados estão inseridos no âmbito do Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso - IFMT, Campus Cuiabá.

A priori, o universo de interesse para esta pesquisa são os estudantes matriculados nos Cursos Superiores de Tecnologia (CST), de caráter presencial, pois o foco da investigação é o fenômeno da evasão escolar nos cursos presenciais do ensino superior. Os CST analisados são: Automação industrial e controle, Controle de obras e Sistemas para internet.

Os dados originais, brutos ou históricos utilizados nesta pesquisa, referentes às respostas dos questionários socioeconômicos, preenchidos pelos estudantes no ato da inscrição para o exame de seleção, estão armazenados no sistema gerenciador de dados Q-Seleção, que é a base de dados dos processos seletivos do IFMT - Campus Cuiabá.

O questionário socioeconômico é composto por 23 questões. As respostas de 13 delas são selecionadas como características preditoras para a análise da evasão, nessa investigação.

Além dos dados do questionário, foram consideradas outras características como distância entre residência e escola, o turno de estudo e percentual de frequência, coletadas no Q-Acadêmico (sistema gerenciador dos dados acadêmicos dos estudantes).

Para esta investigação, as características selecionadas como preditoras para análise e identificação da evasão são: sexo, idade, etnia, estado civil, quantidade de pessoas que residem em casa com o estudante, renda familiar, ter ou não computador em casa, escolaridade dos pais, escola de origem, autoavaliação como estudante, procedência, distância escola-residência, meio de transporte, situação de trabalho, turno de estudo e frequência.

Os dados coletados para análise compreende um período de sete anos, de 2004/2 a 2011/2.

O vetor de entrada da rede neural é composto por 16 parâmetros considerados como significativos para a predição da evasão escolar e a saída da rede constituída por duas classes, evasão e não-evasão. Os pares de vetores entrada-saída desejada são representados na codificação binária, sendo o vetor de entrada composto por 41bits e, as respostas previstas representadas por 1 bit.

Os nomes dos estudantes foram suprimidos para preservar a identidade dos mesmos, sendo na planilha final identificados com números (MARTINHO, 2012).

3. REDES NEURAS ART, ART-FUZZY E ARTMAP-FUZZY

Para inferir sobre os estudantes propensos à evasão, em que os dados são complexos, subjetivos, não-lineares e inter-relacionados, a técnica a ser usada é a rede neural ARTMAP-Fuzzy, uma rede da família ART (*Adaptive Resonance Theory*). São redes baseadas na teoria da ressonância adaptativa, auto-organizáveis, com aprendizagem supervisionada e possibilidade de manter um treinamento continuado. Esta ferramenta é capaz de lidar com dados complexos, não-lineares, imprecisos, incoerentes, ausentes ou ruidosos e, ainda, apresentar maior rapidez e precisão no resultado desejado.

3.1. Rede Neural ART Básica

Os sistemas de redes ART são capazes de resolver o dilema “estabilidade-plasticidade”. Eles podem se auto-organizar, em tempo real, produzindo categorias de reconhecimento estável, enquanto obtêm padrões de entrada além daqueles originalmente armazenados, ou seja, são capazes de aprender a se adaptar em um ambiente em mudança. Portanto, são plásticos, e, ao mesmo tempo, preservam seus conhecimentos aprendidos anteriormente, mantendo sua capacidade para aprender novos padrões, portanto são estáveis.

A Rede Neural ART1 é a mais simples das Redes ART. Ela é competitiva, auto-organizável, aprende de maneira não-supervisionada e aceita apenas entradas binárias.

A estrutura básica de uma rede neural ART, mostrada na Figura 2, consiste de dois subsistemas, de atenção e de orientação (CARPENTER & GROSSBERG, 1987), onde estão dispostos e interligados alguns elementos como: duas camadas de neurônios (F_1 e F_2) e seus pesos sinápticos (W_{ij} e V_{ji}), o módulo parâmetro de vigilância (ρ) e o módulo *reset*.

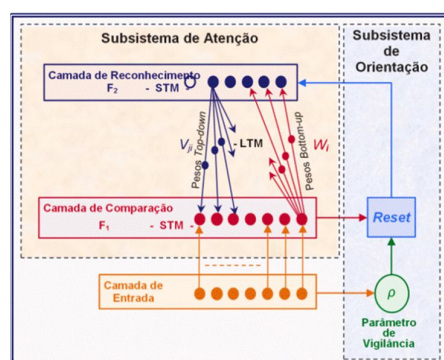


Figura 2 - Estrutura básica da rede neural ART

Os subsistemas de atenção e de orientação são complementares e se interagem para realizar o processamento dos padrões de entrada.

O subsistema de atenção através das camadas de comparação (F_1) e reconhecimento (F_2),

totalmente conectadas com os pesos *bottom-up* (W_{ij}) e *top-down* (V_{ji}), realiza o processamento de padrões de entrada conhecidos, em uma memória de curto prazo (STM), gerando respostas e representações internas mais acuradas desses padrões. O vetor padrão é a entrada para F_1 .

O subsistema de orientação inibe o subsistema de atenção quando um padrão desconhecido é apresentado à rede. O módulo *reset* controla a dinâmica do subsistema de atenção baseado no parâmetro de vigilância (ρ). Enquanto que, o parâmetro de vigilância determina o grau de similaridade entre os vetores padrão de entrada e os pesos sinápticos de F_1 e F_2 .

Os nós na camada de reconhecimento (F_2) representam as categorias ou grupos (*clusters*) formados.

A camada de comparação ou camada F_1 recebe o vetor de entrada e transfere-o para a camada de reconhecimento ou camada F_2 , buscando o seu melhor “casamento” ou combinação. O melhor “casamento” acontece quando encontra um único neurônio, cujo conjunto de pesos corresponde o mais próximo do vetor de entrada.

O subsistema de atenção indica a categoria vencedora, enquanto que o subsistema de orientação aceita a categoria indicada ou aciona a busca para a escolha de uma nova categoria.

Resumidamente, o processo de classificação de uma rede ART consiste em quatro etapas:

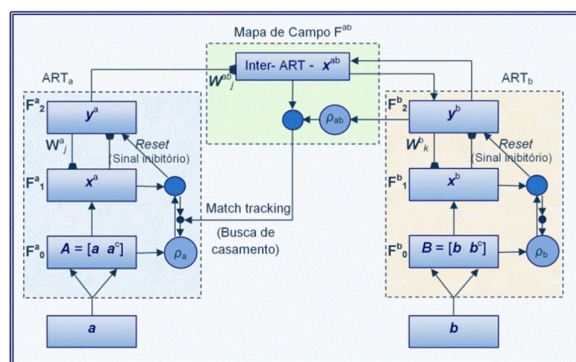
- Reconhecimento: reconhece os dos estímulos produzidos na camada F_2 e seleciona a categoria de maior valor após o cálculo da função escolha.
- Comparação: através do parâmetro de vigilância, testa a similaridade entre o vetor de entrada e o vetor protótipo, permitindo ou não a inclusão do padrão de entrada na categoria.
- Busca: para cada novo vetor de entrada, busca um neurônio na camada F_2 .
- Treinamento: o treinamento só começa após a conclusão do processo de busca.

3.2. Rede Neural ART-Fuzzy e ARTMAP-Fuzzy

A rede *ART-Fuzzy* utiliza a teoria dos conjuntos nebulosos (CARPENTER *et al.*, 1991), empregando o operador de mínimo (\wedge). E, realiza no pré-processamento dos vetores de entrada a normalização e a codificação por complemento de seus componentes.

O modelo *ARTMAP-Fuzzy* é baseado em operações da lógica *Fuzzy* (ZADEH, 1965) incorporadas no modelo ARTMAP. Neste modelo, também conhecido como ART preditiva, dois módulos ART estão interligados, através de um módulo inter-ART, denominado Mapa de Campo. Este módulo inter-ART dispõe de um mecanismo auto-regulador denominado *match tracking*, que busca casamentos entre as categorias, tendo como objetivo elevar o grau de generalização e diminuir o erro da rede. Ele forma associações preditivas entre as categorias dos módulos ART_a e ART_b e realiza um rastreamento e reconhecimento auto-organizado de categorias, em busca de casamento, em relação aos padrões de entrada.

A arquitetura da rede neural *ARTMAP-Fuzzy*, mostrada na Figura 3, foi concebida para a



realização de aprendizagem supervisionada em um ambiente ou conjunto de dados multidimensional. Quando a rede ARTMAP-Fuzzy é usada em uma situação problema de aprendizagem, ela é treinada até que consiga classificar corretamente todos os dados de treinamento. No modo supervisionado, os mapeamentos são aprendidos entre os vetores de entrada dos módulos “ART_a” e “ART_b”.

O desenvolvimento matemático e os algoritmos para o processamento de uma rede neural ART-Fuzzy e ARTMAP-Fuzzy, são encontrados, respectivamente, em (Carpenter *et al.*, 1991) e (Carpenter *et al.*, 1992a), e utilizados em (MARTINHO *et al.*, 2013).

4. SISTEMA NEURAL ARTMAP-FUZZY PROPOSTO PARA PREDIÇÃO DA EVASÃO

Nesta pesquisa, o modelo de sistema inteligente proposto foi desenvolvido utilizando uma rede neural ARTMAP-Fuzzy associada a um módulo da Lógica Fuzzy (ZADEH, 1965) e, em uma segunda simulação, um módulo que emprega a Teoria de Evidência de Dempster-Shafer - TDS (BARNETT, 1981), possibilitando identificar de maneira precoce, acurada e fidedigna o grupo de risco de evasão escolar discente, em cursos presenciais do ensino superior do IFMT, como uma base de dados extraída das informações no ato da inscrição do estudante para o exame de seleção.

A estrutura do sistema de predição proposto e a sequência de desenvolvimento deste trabalho estão representadas no fluxograma da Figura 4.

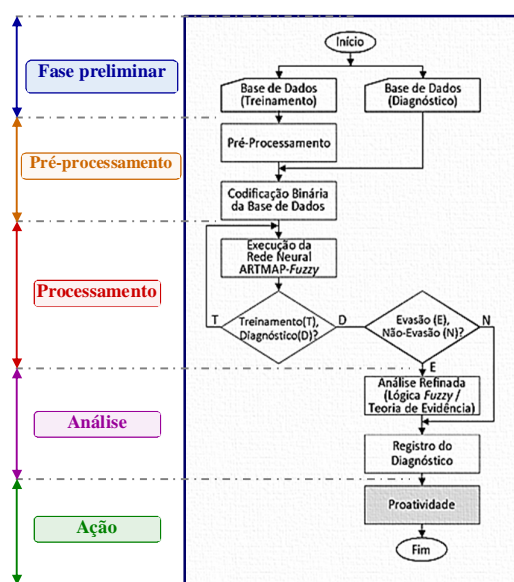


Figura 4 - Fluxograma do processo de desenvolvimento do sistema neural proposto.

Na fase preliminar delimita-se a pesquisa, define-se o método e a base de dados.

Em seguida, na fase de pré-processamento, realiza-se a “limpeza” e a filtragem dos dados brutos selecionados, a categorização das características dos estudantes e o tratamento das informações, semelhante ao proposto por (SILVEIRA, 2011). As bases de dados analógicas, pré-processadas, são codificadas para o código binário e convertidas em bases de dados binários, uma para o treinamento e outra para o diagnóstico da RNA. A concepção de um sistema neural essencialmente binário é consideravelmente vantajoso em relação à formulação da rede neural ARTMAP-Fuzzy híbrida (dados binários e analógicos). Desse modo, ela apresenta um comportamento mais eficiente (maior rapidez e melhor precisão qualidade das

respostas) e permite a extração do conhecimento de forma contínua, ou seja, através do treinamento continuado (MARCHIORI *et al.*, 2011), buscando uma melhor adaptação às condições da instituição e o aperfeiçoamento com o passar do tempo.

Após a binarização as informações são dispostas em linhas. Em cada linha do conjunto de informações estão contidas as características dos estudantes, que compõem o vetor de entrada (41 *bits*) da rede neural e representam uma amostra. A saída desejada da rede, também binária (1 *bit*), é formada pelo *status* do estudante (não evadiu (0) ou evadiu (1)) fornecido pela base de dados. Os vetores de entrada e saída desejada são apresentados às entradas dos módulos ART_a e ART_b, respectivamente, para o processamento das informações.

Concluído o processamento dos dados, chega-se à fase de interpretação, análise e avaliação dos resultados. Nessa fase, se na saída da rede neural a resposta em relação à evasão for negativa, com classe “0” nenhuma ação é adotada, apenas é realizado o registro da referida informação. Se a resposta de evasão for positiva, com classe “1”, essa informação passa pelo módulo *Fuzzy*, para uma melhor discriminação sobre a qualidade da informação (análise fina).

O final do processo é a fase da Ação. Diante dos resultados, faz-se um planejamento de ações individualizadas (ação proativa), para atender as necessidades de cada estudante identificado no grupo de risco, proporcionando condições para sua permanência na instituição. Ressalta-se que a proatividade é uma atividade que depende fortemente da atuação da instituição.

5. APLICAÇÃO E ANÁLISE DOS RESULTADOS

Nesse estudo, em específico, a entrada da rede ARTMAP-*Fuzzy* proposta é representada pelo vetor ***a*** (entrada do módulo ART_a) e sua respectiva saída desejada, na fase de treinamento, representada pelo vetor ***b*** (entrada do módulo ART_b), sendo estes descritos da seguinte maneira:

$$\mathbf{a} = [a_1 \ a_2 \ a_3 \ \dots \ a_{16}] \quad \text{e} \quad \mathbf{b} = [b] \quad , \quad \text{sendo: } b = \text{“0” ou “1”}$$

Os subvetores ***a***₁, ***a***₂, . . . ***a***₁₆ do vetor ***a*** (Tabela 1) são vetores linhas que contêm a representação binária das características dos estudantes. Cada bit corresponde a um componente do vetor associado.

Tabela 1- Composição dos Vetores de Entrada e Saída.

Características dos Subvetores de <i>a</i> e <i>y</i>				
Variáveis do Vetor de Entrada (<i>a</i>) da Rede	<i>Posição</i>	<i>Nome</i>	<i>Abrev.</i>	<i>Tam</i>
	<i>a</i> ₁	Sexo	Sx	1 <i>bit</i>
	<i>a</i> ₂	Faixa etária	FxEt	3 <i>bits</i>
	<i>a</i> ₃	Etnia	Etn	3 <i>bits</i>
	<i>a</i> ₄	Estado Civil	EstCiv	3 <i>bits</i>
	<i>a</i> ₅	Nº pessoas/casa	PesCs	3 <i>bits</i>
	<i>a</i> ₆	Renda Familiar	Rfam	3 <i>bits</i>
	<i>a</i> ₇	Tem Computador	Comp	1 <i>bit</i>
	<i>a</i> ₈	Escolaridade país	EscPai	3 <i>bits</i>
	<i>a</i> ₉	Escola de Origem	EscOri	3 <i>bits</i>
Características dos Subvetores de <i>a</i> e <i>y</i>				
Variáveis do Vetor Entrada (<i>a</i>)	<i>Posição</i>	<i>Nome</i>	<i>Abrev.</i>	<i>Tam</i>
	<i>a</i> ₁₀	Auto avaliação	AAval	3 <i>bits</i>
	<i>a</i> ₁₁	Procedência	Procd	1 <i>bit</i>
	<i>a</i> ₁₂	Dist. Esc.- Resid.	DistER	3 <i>bits</i>
	<i>a</i> ₁₃	Meio Transporte	Transp	3 <i>bits</i>
	<i>a</i> ₁₄	Situação Trabalho	Trab	3 <i>bits</i>
	<i>a</i> ₁₅	Turno de Estudo	Turno	2 <i>bits</i>
Saída (<i>y</i>)	<i>y</i>	Alunos/ Sala	Al/SI	3 <i>bits</i>
		Não-Evasão	NEv	1 <i>bit</i>
		Evasão	Ev	

A saída da rede é representada pelo vetor atividade da camada F₂ (***y***) e fornece respostas na codificação binária com 1 *bit* (Tabela 1), sendo que o código “1” corresponde à evasão e o código “0” a não-evasão discente, assim definido: ***y*** = [*y*] (saída da rede ARTMAP-*Fuzzy*).

Na fase de validação e diagnóstico do sistema neural ARTMAP-*Fuzzy* foram utilizadas duas amostras, uma delas com 389 linhas e a outra com 499 linhas, ambas com 41 colunas, cerca de 30% do total das amostras do treinamento. Cada linha possui as características de um

estudante e representa os vetores padrão de entrada e sua respectiva saída desejada, no treinamento. Os dados das colunas de 1 a 41 representam os atributos correspondentes ao vetor \mathbf{a} , entrada do módulo ART_a. Na coluna 42 estão representadas as saídas desejadas, vetor \mathbf{b} (entrada ART_b) da rede neural ARTMAP-*Fuzzy*.

Os parâmetros utilizados no processamento dos dados estão especificados na Tabela 2.

Tabela 2 - Especificação dos parâmetros da rede neural ARTMAP-*Fuzzy*

Parâmetros e Valores de Referência		Parâmetros e Valores de Referência	
<i>Parâmetros</i>		<i>Parâmetros</i>	
Parâmetro de escolha ($\alpha > 0$)		Parâmetro de vigilância módulo ART _b ($p_b \in [0,1]$)	
Taxa de treinamento ($\beta \in [0,1]$)		Parâmetro de vigilância módulo inter-ART _a ($p_{ab} \in [0,1]$)	
Parâmetro de vigilância módulo ART _a ($p_a \in [0,1]$)		Parâmetro de vigilância no <i>match tracking</i> ($p_{amat} \in [0,1]$)	
Incremento do parâmetro de vigilância p_a (ϵ)			
	0,001		0,999
	1,0		0,7
	0,2		0,75
	0,05		

Após o treinamento da rede foram realizadas cinco simulações, com a base de dados para o diagnóstico, para a validação do modelo proposto, sendo que, em uma delas as amostras foram processadas de maneira naturalmente aleatória e nas demais de forma randômica.

Os resultados dos processamentos foram comparados e analisados, utilizando-se o “critério de votação” [CARPENTER *et al.*, 1992a], foi verificada a saída “0” ou “1” de maior incidência para cada uma das entradas. O resultado de maior incidência constituiu-se na saída da rede neural.

Posteriormente, confrontando a saída resultante da rede com a real situação de cada amostra do grupo de estudante analisado, foi possível averiguar a coincidência de evasão (“1”) e não evasão (“0”) entre as amostras processadas e a realidade.

Concluída a etapa do processamento através do sistema neural ARTMAP-*Fuzzy* e das análises pertinentes e necessárias à compreensão do comportamento da rede quanto à evasão e não evasão discente, os resultados foram compilados e, sucintamente, demonstrados nas Tabelas 3a e 3b.

Tabela 3 - Resultado do diagnóstico da evasão discente

a - evasão total: período de 2004-2010.

b - evasão 1º semestre: período de 2004-2011.

Diagnóstico Evasão Total 2004/2010	Saída Rede - Valores Quantitativos e Percentuais					
	Evasão		Não-Evasão		Total Amostras	
	<i>Qtd.</i>	%	<i>Qtd.</i>	%	<i>Qtd.</i>	%
Amostras	243	100	146	100	389	100
Acertos	231	95,1	135	92,5	366	94,1
Erros	12	4,9	11	7,5	23	5,9

Diagnóstico Evasão 1º sem. 2004/2011	Saída Rede – Valores Quantitativos e Percentuais					
	Evasão		Não-Evasão		Total Amostras	
	<i>Qtd.</i>	%	<i>Qtd.</i>	%	<i>Qtd.</i>	%
Amostras	90	100	409	100	499	100
Acertos	88	97,8	295	72,1	383	76,7
Erros	2	2,2	114	27,9	116	23,3

Os dados da Tabela 3b mostram os resultados da simulação B, que considerou apenas os estudantes evadidos no 1º semestre do curso, no período de 2004 a 2011. Obtendo um índice de acerto de 97,8% com relação aos estudantes evadidos e um percentual de acerto global de 76,7% (MARTINHO *et al.*, 2013).

A interpretação e análise dos dados resultantes da simulação A (Tabela 3a), em que todos os estudantes evadidos do curso, entre 2004 e 2010, foram considerados, corrobora os índices de acurácia obtidos na simulação B. Observa-se que o sistema proposto identificou 231 possibilidades de evasão e ignorou 12, de um total de 243 estudantes evadidos, obtendo um índice de acerto de 95,1%. Dentre as 146 amostras de estudantes não evadidos, 135 foram reconhecidas, acertando 92,5% dos casos. Nesta simulação, o sistema neural ARTMAP-*Fuzzy* alcançou um percentual de acerto global de 94,1%.

Os resultados quantitativos e em percentuais do diagnóstico de evasão obtidos na

simulação A, podem ser percebidos, com maior clareza, nos gráficos da Figura 5.

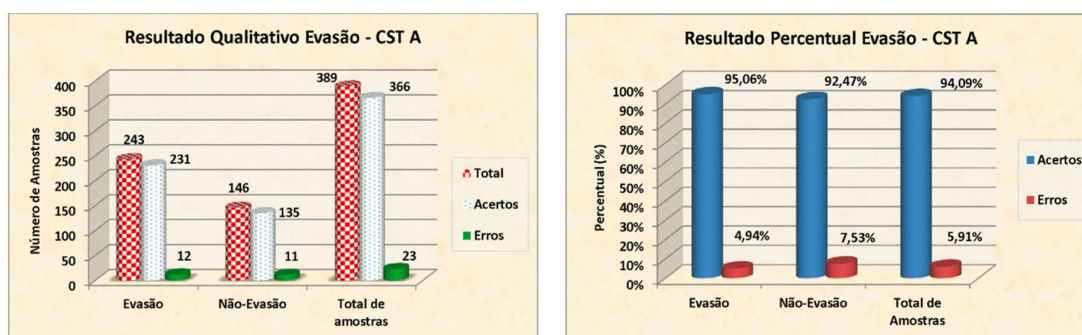


Figura 5 - Resultado quantitativo e percentual da predição da evasão total em CST- 2004 a 2010.

Considerando os testes realizados e a consistência dos resultados obtidos, pode-se inferir que o sistema inteligente, utilizando rede neural ARTMAP-*Fuzzy*, proposto para identificar os estudantes propensos à evasão, é um modelo com significativo grau de confiabilidade e exprime fidedignamente a situação que se encontra os estudantes analisados.

6. CONSIDERAÇÕES FINAIS

Este trabalho está alicerçado em uma pesquisa que tem como objetivo propor um método inovador para identificar, de maneira proativa, continuada e acurada os estudantes considerados do grupo de risco de evasão discente, em cursos superiores presenciais, com uma base de dados composta pelas informações da inscrição. Para esse propósito foi implementado um sistema neural ARTMAP-*Fuzzy* associado com um módulo *Fuzzy* e um módulo da Teoria de Evidência de Dempster-Shafer.

Com a análise dos resultados obtidos é possível inferir que o sistema proposto é apropriado, eficaz e com significativo grau de confiabilidade, alcançou um índice de acerto global entre 76% e 94% na identificação antecipada do grupo de risco evasivo.

E, ainda, o confronto dos resultados obtidos com o sistema neural ARTMAP-*Fuzzy*, adotado neste estudo, com aqueles descritos por (MUSTAFA *et al.*, 2012) com a aplicação de árvores de decisão e classificação (entre 28% e 38% de acertos) para identificar a evasão com base nos dados da inscrição dos estudantes, demonstra a qualidade e relevância do sistema neural proposto.

Neste estudo lida-se com a previsão de ações advindas de deliberações e decisões de seres humanos. Assim, reconhece-se as limitações da metodologia e suas possíveis falhas, haja vista que as previsões fogem ao completo determinismo sendo a evasão o resultado de um processo estocástico.

Portanto, diante dos resultados, evidencia-se que o do sistema neural ARTMAP-*Fuzzy* é uma poderosa, arrojada e inovadora ferramenta para a predição de grupos de risco de evasão discente, em cursos superiores presenciais, inserindo-se na lacuna existente nas produções da comunidade científica mundial, no que tange o tema em questão, contribuindo desta forma com algo que seja útil para a sociedade. Pois, com uma ação proativa, conseguir com que um estudante, com potencial de evadir, tenha sucesso é uma missão nobre, pois são sonhos que se realizam.

Agradecimentos

Ao Instituto Federal de Mato Grosso pela disponibilização da base de dados dos sistemas Q-acadêmico e Q-seleção que fomentaram essa pesquisa, bem como, a CAPES pelo apoio financeiro.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARBOSA, J. O. Identificação e análise dos fatores determinantes da evasão escolar no CST em automação industrial no IFMT - 2004-2010. Relatório Interno, 40 p., 2010.
- BARNETT, J. A. Computational methods for a mathematical theory of evidence. Proceedings of the 7th International Joint Conference on Artificial intelligence. Vancouver, 1981.
- BERGER, Joseph B.; LYON, Susan C. Past to present: A historical look at retention. College student retention: Formula for student success, Westport: Ed. Praeger, p. 1-30, 2005.
- CAPELATO, R. O que os números revelam sobre o perfil da evasão no ensino superior brasileiro. Seminário ABMES: Evasão e retenção: problemas e soluções. Brasília: ABMES, 2011.
- CARPENTER, G. A.; GROSSBERG, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics and Image Processing, v.37,1987.
- CARPENTER, G. A.; GROSSBERG, S. A self-organizing neural network for supervised learning, recognition, and prediction. Communications Magazine, v. 30, n. 9, p. 38-49, 1992.
- CARPENTER, G. A.; GROSSBERG, S.; IIZUKA, K. Comparative performance measures of fuzzy ARTMAP, learned vector quantization, and back propagation for handwritten character recognition. International Joint Conference on Neural Networks, v.1, p.794-799, 1992b.
- CARPENTER, G. A.; GROSSBERG, S.; MARKUZON, N. et al. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE - Transactions on Neural Networks, v. 3, n. 5, p. 698-713, 1992a.
- CARPENTER, G. A.; GROSSBERG, S.; REYNOLDS, J. H. ARTMAP: Supervised real-time learning classification of nonstationary data by a self-organizing neural network. Neural Networks, 1991a.
- CARPENTER, G. A.; GROSSBERG, S.; ROSEN, D. B. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural Network, v.4, 1991.
- CRIPPS, A. Using artificial neural nets to predict academic performance. Proceedings of the Symposium on Applied Computing. Philadelphia, Pennsylvania, USA: ACM: p.33-37, 1996.
- HAYKIN, Simon S. Neural networks and learning machines. Prentice Hall, 2009. 906p, il.
- HERZOG, S. Estimating student retention and degree-completion time: Decision trees and neural networks vis-à-vis regression. New Directions for Institutional Research, n. 131, p. 17-33, 2006.
- INEP. Resumo Técnico: **Censo da Educação Superior 2011**. Disponível em: <http://download.inep.gov.br/educacao_superior/censo_superior/resumo_tecnico/resumo_tecnico_censo_educacao_superior_2011.pdf> Acesso em: 15.05.2013.
- JADRIĆ, M.; GARAČA, Ž.; ČUKUŠIĆ, M. Student dropout analysis with application of Data Mining methods. Journal of Contemporary Management Issues, v. 15, n. 1, p. 31-46, 2010.
- KARAMOUZIS, S. T.; VRETTOS, A. An artificial neural network for predicting student graduation outcomes. World Congress on Engineering and Computer Science, San Francisco: Newswood Limited, p. 991-994, 2008.
- LYKOURENTZOU, I.; GIANNOUKOS, I.; NIKOLOPOULOS, V. *et al.*. Dropout prediction in e-learning courses through the combination of machine learning techniques. Journal Computers & Education, v. 53, n. 3, p. 950-965, 2009.
- MARCHIORI, S. C.; SILVEIRA, M. C. G. da; MINUSSI, C. R. *et al.*. Neural network based on adaptive resonance theory with continuous training for multi-configuration transient stability analysis of electric power systems. Applied Soft Computing, v. 11, p. 706-715, 2011.
- MARTINHO, V. R. C.; NUNES, C.; MINUSSI, C. R. A new method for prediction of school dropout risk group using neural network Fuzzy. XV International Conference on Artificial Intelligence - ICAI'2013. Las Vegas, USA: Proceedings of the ICAI'13, 2013. (no prelo)
- MARTINHO, Valquíria R. de C. UNESP, Faculdade de Engenharia. Sistema inteligente para predição do grupo de risco de evasão discente, 2012. 58p, il. Exame de qualificação (Doutorado).
- MEC/SESu. **REUNI - Reestruturação e expansão das universidades federais**. Disponível em: <<http://portal.mec.gov.br/sesu/arquivos/pdf/diretrizesreuni.pdf>> Acesso em: 15 mai. 2013.
- MEC/SETEC. **Termo de acordo de metas e compromissos**. Disponível em:

<<http://xa.yimg.com/kq/groups/22865919/89930844/name/Termo>> Acesso em: 15 mai. 2013.

MEC-SESu. **Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras**. Disponível em: <<http://www.dominiopublico.gov.br/download/texto/me001613.pdf>> Acesso em: 15 mai. 2013.

MUSTAFA, M. N.; CHOWDHURY, L.; KAMAL, S. Students dropout prediction for intelligent system from tertiary level in developing country. International Conference on Informatics, Electronics & Vision - ICIEV'12. Dhaka, Bangladesh: IEEE, 2012.

PAL, S. Mining educational data using classification to decrease dropout rate of Students. International Journal of Multidisciplinary Sciences and Engineering, v. 3, n. 5, p. 35-39, 2012.

PORTELA, S. Evasão ou retenção?. ABMESeduca.com. São Paulo, 17 jan. 2013.

SILVEIRA, Maria do Carmo Gomes; UFMT, Departamento de Engenharia Elétrica. Sistema de gerenciamento de base de dados para aplicação em ambientes de sistemas elétricos de potência, 2011. Relatório de Estágio (Pós Doutorado).

SOUZA, Solange; UFRJ, Programas de Pós-Graduação de Engenharia. Evasão no ensino superior: um estudo utilizando a mineração de dados como ferramenta de gestão do conhecimento em um banco de dados referente à graduação de engenharia, 2008. 107p. Dissertação (Mestrado).

UNESCO. **Observatorio de la Educación Superior en América Latina y el Caribe**. Disponível em: <http://unesdoc.unesco.org/images/0014/001400/140087s.pdf>> Acesso em: 15 mai. 2013.

ZADEH, L. A. Fuzzy sets. Information and Control, v. 8, n. 3, p. 338-353, 1965.

PREDICTION OF STUDENT DROPOUT RISK GROUP IN HIGHER EDUCATION CLASSROOM USING FUZZY ARTMAP NEURAL NETWORK

Abstract: *School dropout is one of the most complex and crucial problems in the field of education. It permeates and afflicts the several levels and teaching modalities. In relation to higher education, internationally, the phenomenon is an object of attention and care, aiming to increase the retention and completion rates of the undergraduate students and minimize the social, economic, political, academic and financial damage caused to all involved in the educational process. In this context, it is fundamental to develop efficient and effective methods and instruments for prediction, assessment and monitoring of the students risk of dropping out, making the planning of proactive actions possible for the improvement of the situation. Thus, this study, aims to present the potentialities of an intelligent system, developed for the prediction of student dropout risk group in higher education classroom courses, using a Fuzzy-ARTMAP Neural Network, which makes the continued learning of the system possible. The research which supports this work was developed in the technology courses of the Federal Institute of Mato Grosso, based on the acquisition, compilation and processing of the data from two databases, with academic and socioeconomic records of the students. The consistency of the results, with a success rate of the dropout group around 97% and overall accuracy over 76%, shows the reliability, accuracy and efficiency of the system. It is noteworthy that this is a bold and innovative proposal, as the group of students at risk of dropping out and its possible causes can be identified as early as the first semester of the course.*

Key-words: *Higher education, Dropout prediction, Proactivity, Intelligent system, Fuzzy-ARTMAP neural network.*