

Application of a neuro fuzzy network in prediction of absenteeism at work

A. Martiniano, R. P. Ferreira

Pós-graduação em Engenharia de Produção

Universidade Nove de Julho - Uninove

São Paulo - Brasil

martin@uninove.edu.br, kasparov@uninove.edu.br

R. J. Sassi, C. Affonso

Pós-graduação em Engenharia de Produção

Universidade Nove de Julho - Uninove

São Paulo - Brasil

sassi@uninove.br, carlos.affonso@uninove.com.br

Abstract — Competitiveness, market share, professional development and personal support to community action, health, culture, education and sport, are linked to a promising new market. Coupled with the development of organizations, the pressure to achieve goals more audacious, employees increasingly overwhelmed, they end up buying some disturbance in the health-related type of labor activity. The objective of this paper is to apply a neuro fuzzy network in the prediction of absenteeism at work. To make the database were collected records of absenteeism from work during the period of July/07 to July/2010 in a Courier company. Absences certified with the International Classification of Diseases were stratified into 21 categories, the data were tabulated and filtered in MS Excel software. A neuro fuzzy network was developed using an artificial neural network architecture multilayer perceptron with the error backpropagation algorithm. This paper presents partial results of the use of neuro fuzzy network to predict absenteeism at work.

Keywords — Absenteeism, International Classification of Diseases; Prediction of Absenteeism; Neuro Fuzzy Network.

I. INTRODUÇÃO

A evolução da sociedade ressaltou a importância do relacionamento entre homens, entre as diferentes culturas e entre mercados. Nestas circunstâncias, o trabalho humano tornou-se mais complexo, suplantando o sentido tradicional que o identifica como meio de satisfação de necessidades [1].

O absenteísmo representa para a empresa perda da capacidade produtiva e lucrativa. Um serviço deixa de ser fornecido devido à falta do funcionário, reduzindo a receita da empresa. Assim, o conhecimento do absenteísmo-doença, em suas múltiplas determinações, torna-se fundamental para a organização e planejamento das ações dos serviços de atenção à saúde do trabalhador e a determinação de seu perfil permite a identificação de excessos de ocorrências de casos de determinadas morbidades [2].

O custo do absenteísmo por motivo de doença é, em muitas organizações, maior do que o salário pago ao funcionário doente que deixa de comparecer ao local de trabalho e, não havendo substitutos disponíveis, as implicações para a produtividade são bem maiores [3].

As redes neuro *fuzzy* têm surgido como uma ferramenta promissora, pois reúnem os benefícios das Redes Neurais e da Lógica *Fuzzy*. Assim a aprendizagem e o poder computacional das Redes Neurais, a capacidade de representação e o raciocínio da Lógica *Fuzzy* são combinados [4].

As Técnicas de Inteligência Artificial (IA) podem ser combinadas para gerar arquiteturas como é o caso da rede neuro *fuzzy* MLP utilizada neste trabalho [4].

A grande vantagem desse tipo de sistema deve-se a ação de esforço simultâneo ou cooperação obtida pela combinação de duas ou mais técnicas. Esta cooperação reflete na obtenção de um sistema mais robusto e com menos deficiências [5].

O restante do trabalho está organizado da seguinte forma: na seção II é apresentada a fundamentação teórica do Absenteísmo, na seção III é apresentada a Metodologia, na seção IV é apresentada a rede neuro *fuzzy* na previsão do absenteísmo no trabalho, na seção V descrevem-se os resultados da pesquisa. O trabalho é encerrado na seção VI com a conclusão.

II. ABSENTEÍSMO

O absenteísmo em geral é definido como o não comparecimento no trabalho conforme o programado. Há historicamente uma longa pesquisa, já que este fenômeno, em parte, gera um alto custo para as empresas além de seu *status* de indicador desfavorável [6].

O absenteísmo é considerado a falta ou a ausência de um funcionário em seu ambiente de trabalho e pode ser definido também como: incapacidade temporária ou permanente dessa ausência [7]. É conhecido também pela expressão *absentismo* ou *ausentismo*, expressão utilizada para designar a falta do funcionário ao seu local de trabalho. Mesmo não sendo motivada por doença prolongada ou licença legal [8].

A falta de uma política de benefícios está diretamente ligada às causas do absenteísmo. Com o funcionário frustrado, ele passa a não cumprir sua jornada completa de trabalho, é motivado aos atrasos constantes ou passa a ter morosidade no cumprimento de suas funções [8].

O absenteísmo pode ser atribuído a causas conhecidas e a causas ignoradas. Dentre as conhecidas, estão todas as amparadas por lei e que são, por isso mesmo, justificadas ao empregador, solicitando-lhe a permissão de ausência. É o caso de férias, casamentos, nascimentos, óbitos e mudanças de domicílio. As ignoradas são justificadas geralmente por problemas de saúde do trabalhador e/ou de seus dependentes ou de fatores aleatórios dos mais diversos [9].

Esses problemas podem ser evitados ao reconhecer que o funcionário pode não estar disponível para realizar seu trabalho conforme programado. Isso geralmente significa que o trabalho é feito de forma menos eficiente por outro funcionário ou não é feito [10].

Tratando-se de previsão do absenteísmo podem-se destacar os trabalhos [11, 12, 13] por suas contribuições.

III. METODOLOGIA

Os registros de absenteísmo atestados com o Código Internacional de Doenças (CID) foram estratificados em 21 categorias com o objetivo de obter o impacto dessas ausências, esses parâmetros foram convertidos através dos conjuntos *Fuzzy*. As horas de afastamento foram tabuladas por categoria em uma planilha no *software* MS Excel e transformados numa tabela com zeros e uns, o 0 (zero) para não ausência e o 1 (um) para ausência. Essa tabela foi processada pela rede neuro *fuzzy*. Como linguagem de programação foi utilizado o *software* livre SCILAB 5.1, em função das vantagens apontadas em (<http://www.scilab.org/>) como métrica para verificação da validade da rede, foi escolhido o erro médio absoluto.

Para modelar às funções de pertinência correspondentes as categorias por CID, foram utilizadas funções Gaussianas, conforme mostrado na equação (1) a seguir:

$$\mu^{\alpha}(x) = e^{-\frac{1}{\sigma}(x-c_{\alpha})^2} \quad (1)$$

Onde:

- c_{α} : centro da função Gaussiana
- σ_{α} : dispersões da função
- α : rótulo linguístico (alto, médio, baixo).

O algoritmo de treinamento usado na MLP foi o *error backpropagation* que consiste basicamente, em determinar as variações nos pesos sinápticos da rede, tendo como objetivo minimizar o erro obtido na saída através do aprendizado do vetor de treinamento. Para isto, o algoritmo baseia-se no método do gradiente descendente [14,22]. O *error backpropagation* funciona da seguinte maneira: apresenta-se um padrão à camada de entrada da rede, este padrão é processado, camada por camada, até que a saída forneça a resposta desejada, a f_{MLP} é calculada como mostra a equação (2).

$$f_{MLP}(x) = \varphi \left(\sum_1^{Non} v_l \cdot \varphi \left(\sum w_{lj} x_l + b_{l0} \right) + b_0 \right) \quad (2)$$

Onde v_l e w_{lj} são pesos sinápticos; b_{l0} e b_0 são os *biases*; e φ a função de ativação, comumente especificada como sendo a função sigmóide, conforme mostra a fig. 1 [15].

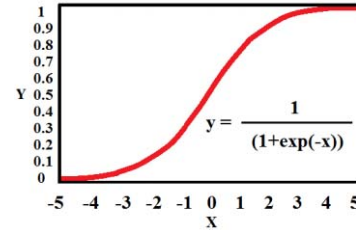


Figura 1. Função sigmóide. Adaptada [16].

Os parâmetros utilizados para o desenvolvimento da RNF MLP foram: número de neurônios de entrada igual a 21, número de camadas ocultas igual a 3, número de neurônios nas camadas ocultas igual a 10, taxa de aprendizagem constante igual a 0,1, fator de momento constante igual 0,7, critério de parada foi o número máximo de épocas igual a 65.

A Tabela I apresenta o Código Internacional de Doenças (CID) por categorias que foram utilizadas como entradas na rede neuro *fuzzy*.

TABELA I - Código Internacional de Doenças

I	Algumas doenças infecciosas e parasitárias (A00-B99)
II	Neoplasias [tumores] (C00-D48)
III	Doenças do sangue e dos órgãos hematopoéticos e alguns transtornos imunitários (D50-D89)
IV	Doenças endócrinas, nutricionais e metabólicas (E00-E90)
V	Transtornos mentais e comportamentais (F00-F99)
VI	Doenças do sistema nervoso (G00-G99)
VII	Doenças do olho e anexos (H00-H59)
VIII	Doenças do ouvido e da apófise mastóide (H60-H95)
IX	Doenças do aparelho circulatório (I00-I99)
X	Doenças do aparelho respiratório (J00-J99)
XI	Doenças do aparelho digestivo (K00-K93)
XII	Doenças da pele e do tecido subcutâneo (L00-L99)
XIII	Doenças do sistema osteomuscular e do tecido conjuntivo (M00-M99)
XIV	Doenças do aparelho geniturinário (N00-N99)
XV	Gravidez, parto e puerpério (O00-O99)
XVI	Algumas afecções originadas no período perinatal (P00-P96)
XVII	Malformações congênitas, deformidades e anomalias cromossômicas (Q00-Q99)
XVIII	Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte (R00-R99)
XIX	Lesões, envenenamento e algumas outras consequências de causas externas (S00-T98)
XX	Causas externas de morbidade e de mortalidade (V01-Y98)
XXI	Fatores que influenciam o estado de saúde e o

As demais causas do absenteísmo no trabalho como: as faltas injustificadas, os atestados odontológicos, consultas médicas, atestados de acompanhante, doação de sangue, fisioterapia e exames foram quantificados separadamente e não serão apresentados neste trabalho.

IV. A REDE NEURO FUZZY NA PREVISÃO DO ABSENTEÍSMO NO TRABALHO

A Lógica *Fuzzy* é uma teoria matemática desenvolvida a partir de 1965 [18], tem como principal objetivo permitir a modelagem do modo aproximado do raciocínio humano, imitando a habilidade humana de tomar decisões em ambientes de incerteza, imprecisão e vago expressa por um conjunto de variáveis linguísticas [19].

A Lógica *Fuzzy* provê uma interface de alto nível e amigável para se desenvolver programas, auxiliando os projetistas a se concentrarem nos objetivos funcionais em vez dos detalhes matemáticos [20].

O conjunto *Fuzzy* é a ponte que liga o conceito impreciso a sua modelagem numérica, atribuindo-se a cada elemento do universo um valor entre 0 e 1, que representa o grau de pertinência deste indivíduo ao conjunto *Fuzzy* [18, 21, 22].

As Redes Neurais Artificiais são modelos inspirados na estrutura do cérebro tendo como objetivo simular o comportamento humano em processos como: aprendizagem, adaptação, associação, tolerância a falhas, generalização e abstração [14, 20]. Estes modelos são constituídos por unidades de processamento simples, chamados de neurônios artificiais, que calculam funções matemáticas. As Redes Neurais Artificiais são convenientes para o manuseio massivo de dados numéricos [20].

A Rede Neural Artificial combinada à Lógica *Fuzzy* forma a rede neuro *fuzzy*. A fig. 2 mostra o fluxo na rede neuro *fuzzy* MLP tendo como saída a previsão do absenteísmo no trabalho em horas.

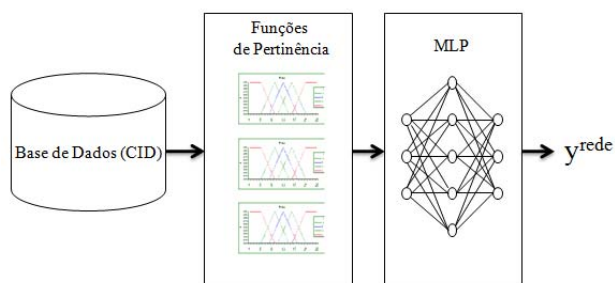


Figura 2. Fluxo na rede neuro *fuzzy* MLP. Adaptado [23,24].

Na fig. 2 observa-se:

- A entrada da RNF as causas do absenteísmo (CID);
- As Funções de Pertinência;

- A *Multilayer Perceptron* (MLP);
- A saída da rede neuro *fuzzy* (y^{rede}).

V. RESULTADOS

A fig. 3 mostra os resultados reais do absenteísmo (Y_{real}), e a previsão do absenteísmo no trabalho, em horas, calculada pela rede neuro *fuzzy* (Y_{rede} - treinamento) e (Y_{rede} - teste).

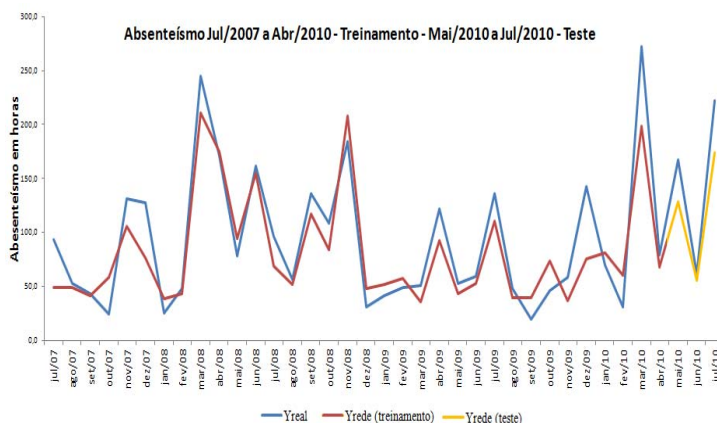


Figura 3. Absenteísmo real e previsão do Absenteísmo em horas.

Com a previsão do absenteísmo é possível estimar qual a parcela da produtividade que pode ser afetada, por exemplo, no mês de Março de 2008 observa-se 250 horas de absenteísmo, que pode representar para a empresa 10%, 15% ou 20% da capacidade produtiva, essa informação antecipada ajuda na tomada de decisão, em relação a:

- planejamento de férias;
- convocação de horas extras;
- trabalho no fim de semana ou feriado;
- contratação de mão de obra temporária, etc.

Com objetivo de reduzir os impactos do absenteísmo no trabalho.

Neste contexto, a previsão do absenteísmo pode ser de grande ajuda na tomada de decisão evitando perda de produtividade e lucratividade.

A fig. 4 mostra as seis categorias com maior representatividade na base de dados, juntas correspondem a 78,65% das ausências atestadas com CID. Observa-se que a categoria XIX, lesões, envenenamento e algumas outras consequências de causas externas (S00-T98), categoria XII, doenças da pele e do tecido subcutâneo (L00-L99) e categoria XIII, doenças do sistema osteomuscular e do tecido conjuntivo (M00-M99), são as doenças que mais causam absenteísmo na empresa pesquisada, por esse motivo devem ser melhor investigadas visando ações preventivas de saúde ocupacional.

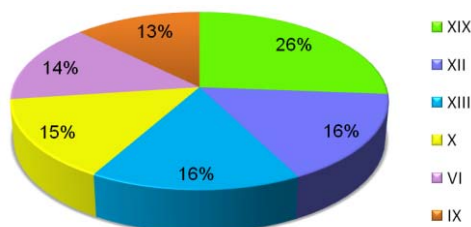


Figura 4. Distribuição por Código Internacional de Doenças (CID) com maior incidência.

VI. CONCLUSÃO

Conclui-se que a rede neuro *fuzzy* para previsão do absenteísmo no trabalho pode ser uma excelente ferramenta para auxiliar a tomada de decisão, uma vez que a produção pode ser mantida com medidas planejadas como distribuição de férias e nas medidas de manutenção da produção como convocação de horas extras e/ou trabalho no fim de semana ou feriado, contratação de mão de obra temporária, etc. Com a vantagem do tempo hábil de planejamento quanto à convocação de funcionários antecipadamente ou a contratação temporária.

Pretende-se dar continuidade nesse trabalho inicial, utilizando uma base de dados maior, acrescentando todas as causas do absenteísmo, que não foram apresentados. Pretende-se ainda utilizar a Mineração de Dados para extrair conhecimento útil na base de dados. Por fim, realizar um levantamento tipo *survey*, na mesma empresa de *Courier*, com o intuito de identificar padrões de comportamento de funcionários absenteístas e presenteístas.

AGRADECIMENTO

À Universidade Nove de Julho pela bolsa de estudos no Programa de Pós-Graduação em Engenharia de Produção.

REFERÊNCIAS

- [1] G. Lumare Júnior, Valor econômico do cliente: uma teoria das encomendas. São Paulo: *Financial Times* – Prentice Hall, 2007.
- [2] L. S. Silva, T. M. M. Pinheiro, S. Sakurai, Perfil do absenteísmo em um banco estatal em Minas Gerais: análise no período de 1998 a 2003. Disponível em: < <http://redalyc.uaemex.mx/pdf/630/63009609.pdf>> Acessado em: 20 de jan. 2011.
- [3] M. Pauly, Efeito multiplicador: consequências financeiras do absenteísmo – 2006. Disponível em: <<http://www.wharton.universia.net/index.cfm?fa=viewfeature&id=1075&language=portuguese>> Acessado em: 04 jan. 2011.
- [4] F. Gomide, M. Figueiredo, W. Pedrycz, *A neural Fuzzy network: Structure and learning, Fuzzy Logic and Its Applications, Information Sciences and Intelligent Systems, Bien, Z. and Min, K., Kluwer Academic Publishers, Netherlands*, 1998, pp. 177-186.
- [5] R. J. Sassi, Uma Arquitetura Híbrida para Descoberta de Conhecimento em Bases de Dados: Teoria dos *Rough Sets* e Redes Neurais Artificiais Mapas Auto-Organizáveis. 2006. Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia Elétrica, São Paulo. 169 p.

- [6] G. Johns, *Presenteeism in the workplace: A review and research agenda. Journal of Organizational Behavior.* vol. 31, p. 519 – 542, 2010.
- [7] A. M. Silva, R. P. Ferreira, R. J. Sassi, *Control and monitoring of the indexes of absenteeism and presenteeism with aid of the technology of the information.* CONTECSI - 2010, ISBN: 978-85-99693-06-3, 2010.
- [8] I. Chiavenato, Recursos humanos na empresa. 3. ed. São Paulo: Atlas, 1994.
- [9] I. Penatti, J. S. Zago, O. Quelhas, Absenteísmo: as consequências na gestão de pessoas. III SEGET – Simpósio de Excelência em Gestão e Tecnologia. Resende, 2006.
- [10] W. Cascio, J. Boudreau, *Investing in People: Financial Impact of Human Resource Initiatives*, Pearson Education, Inc. Publishing as FT Press, New Jersey, 2008.
- [11] J. M. Ivancevich. *Predicting Absenteeism from Prior Absence and Work Attitudes. The Academy of Management Journal.* Vol. 28, no. 1, p. 219-228, mar. 1985.
- [12] J. H. Morris, J. D. Sherman, R. A. Snyder. *Prediction of Absenteeism from Attitudes, Prior Absenteeism, and Performance, Personnel Review*, Vol. 18 no. 1, p.16 - 22, 1989.
- [13] LT. Yen , DW. Edington, P. Witting . *Prediction of prospective medical claims and absenteeism costs for 1284 hourly workers from a manufacturing company. Journal of Occupational Medicine*, vol. 34, no.4, p.428-435, ISSN: 0096-1736, 1992.
- [14] S. Haykin, *Neural Networks: A Comprehensive Foundation.* New York: Wiley & Sons, 1999.
- [15] R. P. Ferreira, C. O. Affonso, R. J. Sassi, Aplicação de uma Rede Neuro Fuzzy na previsão do comportamento do tráfego urbano da Cidade de São Paulo. In: 5ª *Conferencia Ibérica de Sistemas y Tecnologías de Información.* Santiago de Compostela. *Sistemas y Tecnologías de Información CISTI*, vol. II. p. 82-85, 2010, ISBN: 978-989-96247-3-3, 2010.
- [16] R. J. Sassi, L. A. Silva, D. M. E. Hernandez, *A Methodology using Neural Networks to Cluster Validity Discovered from a Marketing Database.* In: 10th *Brazilian Symposium on Neural Networks (SBRN)*, 2008, Salvador. *IEEE Proceedings of SBRN*. vol. 1. p. 03-08, 2008.
- [17] Datasus, Versão 1.6c em Português, 2000.
- [18] L. A. Zadeh, *Fuzzy Sets, Information and Control*, vol. 8. p. 338-353, 1965.
- [19] K. M. Passino, S. Yurkovich, *Fuzzy Control*, Addison Wesley Longman, Inc. 1998.
- [20] M. G. Simões, I. S. Shaw, *Controle e Modelagem fuzzy.* São Paulo: Blucher: FAPESP, 2007.
- [21] M. C. Nicoletti, H. A. Camargo, *Fundamentos da Teoria de Conjuntos Fuzzy*, Edusfscar, 2004.
- [22] M. A. C. Pacheco, M. M. B. R. Vellasco (org), *Sistemas inteligentes de apoio à decisão: análise econômica de projetos de desenvolvimento de campos de petróleo sob incerteza.* Rio de Janeiro: Ed. PUC-Rio: Ed. Inteciência, 2007.
- [23] C. O. Affonso, Aplicação de Redes Neuro Fuzzy ao Processamento de Polímeros na Indústria Automotiva. 2010. Dissertação (Mestrado) – Universidade Nove de Julho, Engenharia de Produção, São Paulo. 110 p.
- [24] R. P. Ferreira, Combinação de Técnicas da Inteligência Artificial para Previsão do Comportamento do Tráfego Veicular Urbano na Cidade de São Paulo, 2011. Dissertação (Mestrado) – Universidade Nove de Julho, Engenharia de Produção, São Paulo. 107 p.