

Predição da ocorrência de sinistros em seguro de automóveis utilizando sistemas baseados em regras *fuzzy*

Cláudio E. L. da Silva ¹

ICSA/UNIFAL, Varginha, MG

Leandro Ferreira ²

ICSA/UNIFAL, Varginha, MG

Danilo Machado Pires ³

ICSA/UNIFAL, Varginha, MG

1 Introdução

A lógica *fuzzy*, baseada na teoria dos conjuntos *fuzzy*, foi proposta inicialmente por [4], com o objetivo de modelar situações de imprecisão ou incerteza, inerentes a comunicação humana. Um conjunto *fuzzy* A de um conjunto universo U é definido por uma função de pertinência μ_A representada por $\mu_A : U \rightarrow [0, 1]$, na qual $\mu_A(x) \in [0, 1]$ é o grau de pertinência com que o elemento x de U pertence ao conjunto *fuzzy* A .

A partir da teoria dos conjuntos *fuzzy*, sistemas baseados em regras *fuzzy* (SBRFs) foram propostos, sendo que de acordo com [1], tais sistemas podem ser divididos nas seguintes etapas: fuzzificação, base de regras, modulo de inferência e defuzzificação. Como exemplo, a base de regras pode ser do tipo “Se x é A e y é B então z é C ”, sendo que x e y são as variáveis de entrada do sistema, z é a variável de saída, e A , B e C são conjuntos *fuzzy* representados por funções de pertinência. Para modelar matematicamente a base de regras, o método de inferência de Mamdani pode ser utilizado, sendo que é considerada a existência de conjuntos *fuzzy* nas partes antecedentes e consequentes de cada regra. A etapa de fuzzificação consiste em transformar as variáveis de entrada em conjuntos *fuzzy*, sendo que a etapa de defuzzificação consiste em representar um conjunto *fuzzy* por um número real.

A lógica *fuzzy* tem potencial aplicação em diversas áreas do conhecimento. Em ciências atuariais, a lógica *fuzzy* é utilizada, por exemplo, na subscrição de seguros, estimativa de mortalidade e cálculo de prêmios. O objetivo do presente trabalho é a construção de SBRFs para avaliar a ocorrência de sinistro em seguro de automóveis, considerando o valor do veículo, tipo do veículo, idade do veículo, sexo do condutor, área de residência e idade do condutor.

2 Metodologia

Para a construção dos SBRFs, foi utilizado o banco de dados apresentado por [2], que se refere a informações sobre apólices de seguro colhidas entre os anos de 2004 e 2005, na região da Austrália. Foram utilizadas 5000 observações das quais 315 (6,3%) possuem pelo menos um sinistro, sendo

¹claudio.estevam@sou.unifal-mg.edu.br

²leandro.ferreira@unifal-mg.edu.br

³danilo.pires@unifal-mg.edu.br

que 80% das observações foram consideradas na etapa de treinamento dos SBRFs e 20% na etapa de teste. Os SBRFs propostos possuem 6 variáveis de entrada, sendo elas: valor do veículo, tipo do veículo, idade do veículo, sexo do condutor, área de residência e idade do condutor. Como variável de saída, tem-se a variável ocorrência de sinistro (1 - houve pelo menos um sinistro; 0 - não houve sinistro). Para tanto, foi utilizado o pacote *frbs* [3] do *software* R e o método de aprendizagem “FRBCS.CHI”, construindo funções de pertinência e o conjunto de regras a partir da partição dos domínios das variáveis. Diferentes SBRFs foram propostos, considerando funções de pertinência dos tipos triangular e trapezoidal, com diferentes quantidades de conjuntos *fuzzy*, e o método de inferência de Mamdani. Na etapa de defuzzificação, foi escolhido o método Centro de Gravidade.

3 Resultados

Os resultados das classificações referentes a ocorrência de sinistro, considerando as etapas de treinamento (TR) e teste (TE), estão apresentados na Tabela 1. Os sistemas S5TRI, S10TRI e S15TRI correspondem aos SBRFs com funções de pertinência triangulares com 5, 10 e 15 conjuntos *fuzzy*, respectivamente, e os sistemas S5TRA, S10TRA e S15TRA correspondem aos SBRFs com funções de pertinência trapezoidais com 5, 10 e 15 conjuntos *fuzzy*, respectivamente. As médias dos percentuais totais de acertos (mPTA) correspondem as médias de 10 simulações realizadas em cada sistema proposto.

Tabela 1: Resultados de classificação dos SBRFs propostos.

	S5TRI		S10TRI		S15TRI		S5TRA		S10TRA		S15TRA	
	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE
mPTA	0,93	0,92	0,94	0,93	0,94	0,92	0,90	0,88	0,93	0,89	0,94	0,90

Fonte: Elaboração própria do autor.

Os SBRFs utilizando função de pertinência triangular apresentaram médias dos percentuais totais de acertos maiores que os sistemas utilizando função de pertinência trapezoidal, exceto nos sistemas S15TRI e S15TRA, em que as médias foram iguais nas etapas de treinamento. Os SBRFs propostos são capazes de prever a ocorrência de sinistro, com assertividade suficientemente alta, servindo estes como ferramentas para auxiliar a seguradora na precificação de seguros de automóveis.

Agradecimentos

Agradecimento à PROBIC/UNIFAL pela concessão da bolsa.

Referências

- [1] Barros, L. C.; Bassanezi, R. C. *Tópicos de lógica fuzzy e biomatemática*. 3. ed. Campinas: UNICAMP/IMECC, 2015.
- [2] De Jong, P.; Heller, G. Z. *Generalized linear models for insurance data*. Cambridge University Press, Cambridge Books, 2008.
- [3] Riza, L. S. et al. *frbs: Fuzzy Rule-Based Systems for Classification and Regression in R*. *Journal of Statistical Software*, v. 65, n. 6, p. 1-30, 2015.
- [4] Zadeh, L. A. Fuzzy sets. *Information and Control*, v. 8, p. 338-353, 1965.