



DESENVOLVIMENTO DE UM SISTEMA PARA CLASSIFICAÇÃO DE PERFIS DE CLIENTES PARA SEGURADORAS DE CARRO

Rodrigo Mota Lacerda, Lucas Araújo Borges, João Paulo Pereira Nery, Maria Clara Fagundes Montalvão, Mariana Caroline Costa Quadros, Carla Eloisa Mendes Santos Oliveira, Rosilanny Soares Carvalho

Introdução

A aplicação proposta por este trabalho é um sistema programado na linguagem Java e que tem como objetivo principal classificar o perfil de usuários de seguros de carros quanto ao risco de sinistros ou não. O projeto aqui proposto não contempla todos parâmetros utilizados pelos demais softwares que fazem classificação de dados, mas se apresenta como um modelo simplificado que utiliza da Inteligência Artificial com Redes Perceptron de Múltiplas Camadas.

Realizada uma breve introdução sobre aquilo que a equipe propôs desenvolver, no decorrer deste artigo serão destacados importantes conceitos relacionados ao trabalho, bem como será explicitado detalhes referentes ao processo de desenvolvimento.

Elementos Fundamentais

Para o desenvolvimento deste trabalho foi necessário o conhecimento sobre os perfis considerados de risco pelas seguradoras de veículos, bem como se fez necessário a utilização e aplicação de Redes Perceptron de Múltiplas Camadas para a classificação de tais perfis. Visando enunciar os tópicos, serão lançadas as suas definições:

A. Seguro de carros

Segurar um carro consiste em minimizar perdas financeiras dos veículos. A contratação de um seguro tem como vantagens indenizações caso haja danos ao veículo e às pessoas, além assistência mecânica. Os danos que ocorrem ao bem segurado são conhecidos como sinistro.

Conforme a probabilidade de ocorrência de sinistros, podendo essa probabilidade variar com o tempo, o valor de um seguro é calculado com base na análise de diversos fatores relacionados ao veículo, ao(s) motorista(s) e ao proprietário. Esses fatores são:

- O perfil do usuário: histórico de sinistros, a idade, o sexo, ser casado(a), a experiência de habilitação;
- A utilização do veículo: os locais que o veículo fica exposto e os condutores;
- Local de circulação: o preço varia conforme as cidades;
- O veículo: o valor do veículo, os preços de reparo, índices de roubo e equipamentos de segurança;
- A cobertura do seguro: quais sinistros que o veículo estará assegurado.

Para este trabalho, foram utilizados os fatores de idade, sexo, estado civil, quantidade de filhos e utilização ou não do carro para ir ao trabalho. Outras definições serão expostas a seguir.

B. Inteligência Artificial

A inteligência artificial consiste em simular a capacidade humana de raciocinar, perceber e tomar decisões e resolver problemas. A Inteligência Artificial é obtida através da construção de redes neurais artificiais. Estas redes neurais são compostas de diversos neurônios que, conforme os padrões de entrada, são ativados ou não.

Um dos tipos de redes neurais é o Perceptron, que foi a primeira máquina criada para processamento de informação feita sobre o sistema de redes neurais

C. Multi-Layer Perceptron (MLP)

As unidades do tipo “Perceptron” foram criadas por Frank Rosenblatt em 1950. Este é um dos modelos de neurônios mais utilizados na atualidade. Ele é a base de diversos tipos de RNA com aprendizado supervisionado utilizando uma adaptação por correção de erros (usualmente baseada na descida da superfície de erro usando o gradiente).

O modelo do Perceptron de múltiplas camadas (MLP – Multi-Layer Perceptron) tornou-se muito conhecido e aplicado, sendo na maior parte das vezes associado a regra de aprendizado do Back-Propagation.

Este modelo possui as seguintes camadas:

- Uma camada de entrada: consiste em uma camada com os sinais de entrada (estímulo da rede). Esta camada não possui neurônios.



- Uma camada de saída: consiste em uma camada de neurônios que geram a saída da rede (resposta da rede a um estímulo).
- Camadas escondidas ou intermediárias: qualquer camada que se encontre entre a camada de entrada e saída. Não existe limites para quantidade de camadas escondidas, e também não é obrigatória a existência delas.

Uma rede MLP é fortemente conectada com conexões feedforward. Ou seja, é uma rede em que as camadas estão organizadas em uma ordem e os neurônios de uma camada estimulam todos os neurônios da camada seguinte (fortemente conectada). Nenhum neurônio pode estimular um neurônio da mesma camada ou de camadas anteriores (feedforward).

Back-propagation (retro-propagação) é o algoritmo de treinamento de uma MLP mais famoso é o proposto por Rumelhart e McClelland em 1986.

Este algoritmo consiste em um algoritmo supervisionado estático (não auto organizável), ou seja, a arquitetura da rede deve ser previamente conhecida, e esta não é alterada durante o treinamento. Os únicos parâmetros alterados são os pesos da rede.

Metodologia

O Sistema de Classificação de Perfis de Clientes para Seguradoras de Carro foi desenvolvido por meio da IDE NET BEANS e utilizou-se da linguagem de programação Java para sua implementação. Tal sistema foi criado com o objetivo de colocar em prática conceitos relacionados à disciplina de Inteligência Artificial e, assim motivados, utilizou-se das redes MLP para a atividade de classificação de dados.

O sistema proposto classifica, mediante a entrada de dados fornecidos pelos usuários, os perfis dos clientes como tendo “perfil de risco” ou “não perfil de riscos”. São colhidos dados como nome, idade, sexo, estado civil, quantidade de filhos e utilização ou não do carro para ir ao trabalho. A partir dele, esses dados são manipulados de forma que possam ser processados pela rede desenvolvida.

Pela definição desta equipe, um perfil considerado de risco é aquele que compreende jovens e adultos, solteiros, que utilizam o carro para trabalhar e não possuem filhos. A partir da generalização dos dados, perfis que apresentam algumas destas características podem ser classificados como de risco, a apresentação de todas estas, obrigatoriamente os classifica perfil.

Os padrões de entrada e os resultados esperados são exibidos na tabela 1. As variáveis utilizadas para representação das características de sexo, estado civil, uso para trabalho e filho são booleanas. O resultado esperado, que é representado pela variável risco, também é booleano. Na variável sexo, o 1 representa masculino e 0, feminino. Já em idade, pessoas com até 21 são representados por 0, de 22 até 59, por 0.5 e, com 60 ou mais, é representado por 1. Nas demais variáveis, o 0 representa “não” e 1, “sim”.

O sistema desenvolvido é composto pelas seguintes classes PerfisMLP, Rede e consultaPerfil.

Na classe PerfisMLP apenas é criada uma nova instância da interface consultaPerfil e também é feita sua inicialização.

A classe consultaPerfil representa a interface gráfica do sistema. Nela se encontram os métodos responsáveis pela estruturação da interface, bem como de captar os dados preenchidos nela e *setar* respostas vindas de outras classes para que o usuário possa consultar. Nela também são definidos os padrões de treinamento para a rede para possibilitar o uso na aplicação, bem como são tratados os dados.

Já na classe Rede é onde se encontram as principais funções e onde ocorre a efetiva implementação da rede MLP. Nessa são definidos o número de neurônios presentes nas duas camadas da rede, bem como são inicializados os pesos sinápticos de cada neurônio. Nesta classe ainda se encontram as funções responsáveis pelo treinamento da rede, fazendo uso do algoritmo backpropagation, e também é feita, após a entrada de valores para teste, a classificação dos dados apresentados como entrada na interface. Como escolha da equipe deste trabalho, para a rede implementada são utilizadas duas camadas, sendo cada uma delas com 4 neurônios. A taxa de aprendizagem utilizada é 0.02 e a função de ativação utilizada pela rede é a tangente hiperbólica.

Após a classificação do perfil do cliente da empresa de seguro, também é exposto na interface do sistema qual o resultado apresentado para a classificação antes do treinamento da rede, bem como o número de ciclos utilizados para o treinamento da mesma.

Resultados

O desenvolvimento do trabalho e principalmente o seu fim, expôs que o uso do aprendizado de máquina para a classificação dos perfis de usuários de seguradora de carros apresentara-se com atribuições de fácil implementação. Entretanto, para o desenvolvimento, necessita-se de conhecimento prévio e fundamentado



sobre os conceitos relacionados à Inteligência Artificial e, principalmente, sobre as Redes Perceptron de Múltiplas Camadas.

A aplicação se mostra como um produto simples, mas eficiente e propício a melhorias e adequações que possibilitem, por exemplo, a inserção de novos padrões de treinamento ou o aprendizado em tempo real. O modo como foi estruturada a aplicação também possibilita sua utilização para classificação de outros tipos de dados, sendo necessários apenas pequenos ajustes. A mesma possui uma interface também simplória e de fácil utilização, mas que de um modo geral fez com que a aplicação que cumprisse os objetivos propostos.

Conclusão

Ao fim deste trabalho conclui-se que o uso de aplicações que fazem uso de conceitos de Inteligência Artificial são de extrema utilidade para automatizar operações que são realizadas no cotidiano das pessoas. Para o propósito utilizado neste trabalho, o aprendizado de máquina mostrou-se como um grande auxiliar no processo de classificação de dados. A aplicação desenvolvida, mesmo apresentando-se como um modelo simplificado, proporcionou a esta equipe a utilização de experiências como transformação, redução e simplificação de dados, presentes, por exemplo, no cotidiano de quem utiliza-se da Mineração de Dados e também possibilitou o entendimento prático da utilização de conceitos próprios de redes MLP.

Referências

- [1] BITTENCOURT, Guilherme. **Inteligência artificial - ferramentas e teorias**. Editora: UFSC, 1998.
- [2] DICAS DE CARRO, 2015. **Seguro – Fatores que influenciam o Seguro dos Carros mais vendidos**. Disponível em: <<http://dicasdecarro.com.br/seguro-fatores-que-influenciam-o-seguro-dos-carros-mais-vendidos/>>. Acesso em: 24 de jun. 2015.
- [3] TECMUNDO, 2008. **O que é Inteligência Artificial?**. Disponível em: <<http://www.tecmundo.com.br/intel/1039-o-que-e-inteligencia-artificial-.htm>>. Acesso em: 24 de jun. 2015.

Tabela 1. Padrões de entrada das variáveis utilizadas para representação das características de sexo, estado civil, uso para trabalho e filho são booleanas.

Sexo	Idade	Estado Civil	Uso p/ trabalho	Filho	Risco
0	0	0	1	0	1
0	0.5	0	1	0	1
1	0.5	1	1	1	0
0	1	1	0	1	0
1	0	0	1	0	1
0	0	1	1	1	0
1	1	1	0	1	0



FEPEG

FÓRUM DE ENSINO,
PESQUISA, EXTENSÃO
E GESTÃO

TRABALHOS CIENTÍFICOS APRESENTAÇÕES ARTÍSTICAS E CULTURAIS DEBATES MINICURSOS E PALESTRAS

23 A 26 SETEMBRO DE 2015
Campus Universitário Professor Darcy Ribeiro

ISSN 1806-549X

A HUMANIZAÇÃO NA CIÊNCIA, TECNOLOGIA E INOVAÇÃO



The screenshot shows a web browser window with the title "Seguradora de Carros". The main content area is titled "Perfil do cliente" and contains a form with the following fields and options:

- Nome:** A text input field.
- Sexo:** Radio buttons for "Masculino" and "Feminino".
- Idade:** A text input field.
- Estado Civil:** Radio buttons for "Solteiro(a)" and "Casado(a)".
- Usa o carro para trabalhar?:** Radio buttons for "Sim" and "Não".
- Quantidade de filhos:** A text input field.
- Classificar:** A button located at the bottom right of the form area.

Below the form, there is a large empty rectangular box. At the bottom of the window, there are two buttons: "Limpar" and "Sair".

¹ Situação antes do treino
² Ciclos de treino

Figura 1. Interface do sistema desenvolvido. Fonte própria, 2015.