

## **CLASSIFICAÇÃO AUTOMÁTICA COM REDES NEURAIS ARTIFICIAIS DE EVENTOS DE ACORDO COM OS PERFIS DOS USUÁRIOS DO SCIENTIA.NET**

*Bruno Vicente Alves de Lima (Aluno Voluntário, ICV/UFPI), Vinicius Ponte Machado (Orientador, Depto. de Informática e Estatística – UFPI)*

### Introdução

A Internet é um dos maiores meios de se conseguir informação. Ao mesmo tempo em que permite o acesso das pessoas às informações nela disponíveis, gera um problema para controlar e organizar essas informações para que as pessoas utilizam a internet possam entrar em contato com informações que são relevantes para eles. A qualidade do conteúdo disponível na Internet está diretamente ligado ao nível de conhecimento de quem as publica.

Assim a disseminação do uso de Redes Sociais Online é reforçada na comunidade científica. O Congresso da Sociedade Brasileira de Computação do último ano teve comunidades oficiais em oito sistemas. Os organizadores suscitam a participação dos envolvidos para compartilhar informações, fatos e experiências relacionadas ao evento [1].

Neste contexto temos o Scientia.Net, é uma nova rede social on-line com tem como objetivo contribuir para a pesquisa científica. Um dos seus objetivos é disponibilizar automaticamente aos usuários da rede conteúdos relevantes de acordo com o seu perfil. Essa característica torna o Scientia.Net um agregador de conteúdo, estes, contidos em diversos lugares da internet, esses lugares podem ser fóruns, repositórios de artigos, sites, blogs e outras redes sociais online.

O Scientia.Net tem a capacidade de agregar informações acadêmicas, utilizando técnicas de Aprendizado de Máquina para classificar o conteúdo e apresentar de forma inteligente, informações sobre artigos e eventos que são compatíveis com os perfis dos usuários, além de agrupar usuários com perfis semelhantes.

Aprendizado de Máquina é uma área de Inteligência Artificial que lida com problemas de aprendizado computacional a fim de adquirir conhecimento de forma automática. Um sistema de aprendizado tem a função de analisar informações e generalizá-las, para a extração de novos conhecimentos. Para isso usa-se um programa de computador para automatizar o aprendizado. [2]

### Metodologia

A primeira etapa do plano de trabalho consiste em consolidar o conhecimento em algoritmos de aprendizagem de máquina através estudos sobre o tema. A próxima fase do trabalho foi implementar alguns algoritmos de Redes Neurais, para a implementação dos algoritmos foi utilizado os algoritmos implementados no WEKA.

Como o WEKA é uma aplicação em Java, foi criada uma aplicação nessa linguagem para importar os algoritmos do WEKA. Durante o processo de integração dos algoritmos do WEKA à uma aplicação Java, percebeu que seus algoritmos, mesmo em outra aplicação, utilizavam a base de dados em um tipo de arquivo especial do tipo ARFF. E como os dados utilizados no projeto estão armazenados em banco de dados MySQL, houve a necessidade de criar um algoritmo para converter as tabelas do MySQL em arquivos ARFF.

Como Scientia.Net é um site, houve a necessidade de transferir a aplicação para Web. Através de um Framework, o JSF 1.2, visando facilitar essa integração ao Scientia.Net. Com a

aplicação funcionando na Web, iniciou-se a implementação do site Scientia.Net. A criação do site Scientia.Net foi feita utilizando o componente de redes sociais do CMS (Content Management System) Joomla, o JomSocial. A Figura 1 abaixo mostra a tela de perfil de um usuário no primeiro protótipo do Scientia.Net.



Figura 1: Primeiro protótipo do Scientia.Net

A etapa seguinte, houve uma tentativa de integrar os algoritmos de Redes Neurais Artificiais ao Scientia.net utilizando o JSF, entretanto o resultado não foi satisfatório. Com isso, foi implementada separadamente a aplicação de classificação de eventos separadamente do Scientia.Net, tal aplicação é executada através de chamadas pelo sistema operacional instalado no servidor do Scientia.Net. O resultado dos Algoritmos são inseridos diretamente no Banco de Dados do Scientia.Net, assim o site pode acessar a esses resultados. Com a ajuda de alguns usuários, criou-se uma base de Dados com 60 eventos científicos nas áreas de Meio Ambiente, Inteligência Artificial, Engenharia de Software, Redes de Computadores, Biologia, Farmácia, Odontologia, Química, Física e Ciências Contábeis.

## Resultados e Discussão

Após criar uma base fictícia com 60 eventos científicos, onde 40 foi utilizado para treinamento e 20 para teste, submeteu-se tais dados em um algoritmo de Rede Neural do WEKA para classificação. Colocando uma Rede Neural com 10 camadas ocultas e com um treinamento de 500 épocas, obteve-se um resultado de 100% de acertos por parte da Rede Neural. Após classificar os usuários pode-se apresentar aos usuários eventos científicos compatíveis com sua área de estudo. As Redes Neurais Artificiais demonstraram ser eficiente quanto a classificação de dados, possibilitando a classificação automática dos eventos.

A Figura 3 mostra a tela inicial da primeira versão do Scientia.Net.

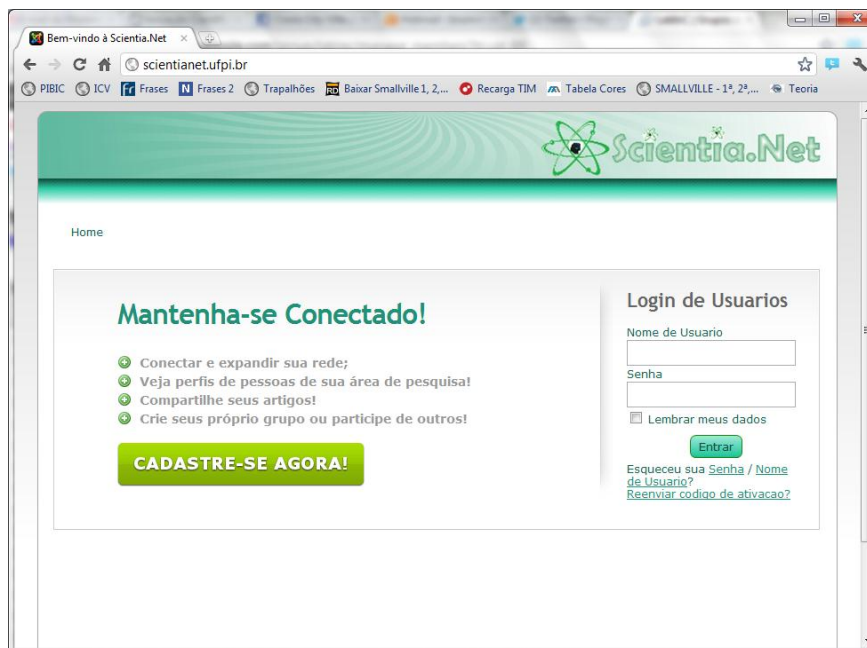


Figura 3: Tela inicial da primeira versão do Scientia.Net.

## Conclusões

O Scientia.Net, rede social acadêmica, tem como objetivo reunir pesquisadores nas mais diversas áreas do conhecimento afim de possibilitar a troca de informações entre eles, e isto feito de forma automática através de algoritmos de Aprendizagem de Máquina. Neste trabalho estão sendo utilizadas as Redes Neurais Artificiais.

As Redes Neurais Artificiais apresentaram um bom desempenho na classificação de exemplos utilizados para teste. Além disso, as RNAs são métodos seguros quando implementados corretamente, pois já existem e são estudados há muito tempo.

Para um desempenho eficiente, as Redes Neurais necessitam de treinamentos com uma enorme quantidade de dados, para que não haja problemas nesse trabalho em relação a isso, serão feitos treinamentos regularmente utilizando a base de dados do Scientia.Net. De forma a ampliar os estudos sobre o trabalho apresentado pretende-se ainda otimizar os algoritmos de aprendizado de máquina do Scientia.Net e utilizar outras fontes de usuários tais como Facebook, Orkut, Lattes.

## Referências Bibliográficas

- [1] SANTANA, V. F.; Melo-Solarte, D. S.; Neris, V. P. A.; Miranda, L. C.; Baranauskas, M. C. C. Redes Sociais Online: Desafios e Possibilidades para o Contexto Brasileiro. In: XXXVI Seminário Integrado de Software e Hardware (SEMISH) / XXIX Congresso da Sociedade Brasileira de Computação (CSBC), 2009, Bento Gonçalves. Anais do XXIX CSBC. Porto Alegre: SBÇ 2009. pp. 339-353.
- [2] MONARD, M.C.; BARANAUSKAS, J. A. Conceitos de Aprendizagem de Máquina. In: REZENDE, S.O. (Ed). Sistemas Inteligentes: fundamentos e aplicações. São Carlos: Manole, 2003. P. 89-114. Cap. 4.

**Palavras-chave:** Redes Sociais. Redes Neurais. WEKA.