

Aprendizado de máquina incremental com o algoritmo de aprendizado da rede ART 2-A: uma biblioteca de software

Luciano José Senger

e-mail: ljsenger@uepg.br

Algoritmos computacionais que implementam técnicas de aprendizado de máquina têm como objetivos encontrar e descrever padrões a partir dos dados obtidos do ambiente. O aprendizado pode ser realizado de diferentes formas, de acordo com um paradigma. Exemplos de paradigmas são: simbólico, estatístico, baseado em instâncias, conexionista e genético. A tarefa de aprender através desses paradigmas consiste em escolher ou adaptar os parâmetros de representação do modelo. Independentemente do paradigma, os algoritmos de aprendizado têm a tarefa principal de aprender um modelo a partir do ambiente e manter esse modelo consistente de modo a atingir os objetivos de sua aplicação. Explicar os dados e fazer previsões são objetivos comuns da aplicação desses algoritmos.

A família de redes neurais ART (*Adaptive Resonance Theory*) consiste em arquiteturas que aprendem em tempo real códigos de representação estáveis em resposta a uma seqüência arbitrária de padrões de entrada [1]. Nas arquiteturas ART, o aprendizado é tratado como uma ação dinâmica e incremental, de forma que a rede possa continuamente adaptar-se aos novos padrões de entrada. O aprendizado nas redes ART é não-supervisionado, de forma que a rede consegue aprender tendo como entrada padrões não rotulados, ou seja, sem empregar um mecanismo supervisor externo, ao contrário de redes com aprendizado supervisionado, em que a rede recebe um conjunto de treinamento previamente classificado e rotulado. Sistemas biológicos, nas fases iniciais de desenvolvimento da visão e audição, são exemplos de aprendizado não-supervisionado [3]. Nessa forma de aprendizado, a rede tem a habilidade de formar representações internas para codificar as entradas através de um conjunto de unidades de saída ou representação. Além disso, as arquiteturas ART são projetadas de forma que o usuário possa controlar o grau de similaridade entre os padrões agrupados na mesma unidade de saída. Esse controle permite que a rede seja mais ou menos sensível às diferenças existentes entre os padrões de entrada, e consiga gerar mais ou menos grupos em resposta a esse controle.

Este trabalho apresenta uma biblioteca de software, construída através de classes C++, na qual são implementados os algoritmos de aprendizado e de classificação de uma arquitetura de redes neurais baseada na teoria da ressonância adaptativa. Tal implementação é baseada na versão algorítmica da rede ART, chamada de ART 2-A [2]. Essa biblioteca constitui um arcaçouço para a construção de modelos não supervisionados e de aprendizado dinâmico e incremental, fornecendo uma boa relação de estabilidade e plasticidade ao paradigma conexionista. A implementação compreende o algoritmo de treinamento básico, assim como um algoritmo de rotulação baseado em matrizes de significância, conforme descrito em [5]. O algoritmo de rotulação permite a identificação automática dos atributos mais relevantes para cada categoria identificada pela rede neural, baseando-se nos protótipos criados pela rede ART 2-A. No software, são também incluídas as implementações das medidas de avaliação de desempenho e de adequação automática do parâmetro de vigilância, definidos por [4]. As implementações do algoritmo de aprendizado, do algoritmo de rotulação e dos mecanismos de avaliação do aprendizado são distribuídas ¹ como um software livre, através da licença GPL.

Referências

- [1] G. A. Carpenter and S. Grossberg. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21(3):77–88, Março 1988.
- [2] G. A. Carpenter, S. Grossberg, and D. B. Rosen. ART 2-A: An Adaptive Resonance Algorithm for Rapid Category Learning and Recognition. *Neural Networks*, 4:494–504, 1991.

¹<http://www.ljsenger.net>

- [3] A. de P. Braga, T. B. Ludermir, and A. C. P. L. F. Carvalho. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC, 2000.
- [4] L. J. Senger. *Escalonamento de processos: uma abordagem dinâmica e incremental para a exploração de características de aplicações paralelas*. Tese (doutorado), ICMC/USP, São Carlos, SP, Brasil, Fevereiro 2005.
- [5] L. J. Senger, R.F. Mello, M. J. Santana, and R. H. C. Santana. An on-line approach for classifying and extracting application behavior on linux. In *High Performance Computing: Paradigm and Infrastructure*. John Wiley and Sons Inc., 2005.