UMA REDE NEURAL PROBABILÍSTICA BASEADA EM MODELOS EVOLUTIVOS

Patrick Marques Ciarelli
Programa de Pós-Graduação em Engenharia Elétrica
Departamento de Engenharia Elétrica
Universidade Federal do Espírito Santo (UFES)
simply_pmc@yahoo.com

Elias Oliveira
Programa de Pós-Graduação em Informática
Departamento de Informática
Universidade Federal do Espírito Santo (UFES)
elias_de_oliveira@yahoo.com.br

Evandro O. T. Salles
Programa de Pós-Graduação em Engenharia Elétrica
Departamento de Engenharia Elétrica
Universidade Federal do Espírito Santo (UFES)
evandro@ele.ufes.br

ABSTRACT

Traditional techniques of machine learning that have off-line learning are commonly applied to become viable to the procedure of some data analysis. However, such techniques are inadequate or inefficient for data that are constantly expanding and/or changing over time.

Although an off-line model takes the advantage of to reach an optimized structure from a training data set, its performance may fall suddenly when a feature of the environment changes. Moreover, it needs to be always re-designed when new circumstances occur during the process.

Another drawback is that an off-line model considers that the whole data set is already available during the training phase. However, in many cases, the acquisition of sufficiently representative data of the problem is onerous, time consuming and, moreover, the data are usually available in small quantities over a period time.

A new paradigm in the field of computational intelligence based on building models from incremental algorithms and data flows has emerged about a decade ago. These models, called evolutionary models, not only minimize the problem of storing large amounts of data, but also offer important characteristics for modelling of non-linear adaptive processes. The main characteristics of the evolutionary models are the continuous learning, the self-organization and the adaptation to the unknown environments.

In this work, we propose a new technique of evolutionary model called Evolving Probabilistic Neural Network. This technique is based on Probabilistic Neural Network, Gaussian mixture model and Expectation Maximization. Whenever new data are available for training, which may occur in sequential mode, this neural network adjusts its architecture and its parameters to better adapt to the problem.

In order to evaluate our approach, we carried out a series of experiments with some data sets that are available on public domain. Beyond our technique, we also tested other neural networks for comparison, which are based on the same paradigm of evolutionary models. All

approaches were applied with zero initial knowledge, that is, none knowledge about the data sets was inserted into of the neural networks, including the number of classes. Their training was performed in incremental mode, that is, a sample at a time was classified then it was used for the training.

We have carried out a two-tailed paired t-test at 5% significance level on the accuracy between each pair of techniques, and our technique was statistically superior to other techniques. Furthermore, our approach has presented smaller time consumption and numbers of neurons than of the other techniques (in some cases they are many times smaller). These results indicate that our approach is promising.

KEYWORDS. Probabilistic Neural Network, incremental learning, Gaussian mixture model.

RESUMO

Técnicas tradicionais de aprendizado de máquina que possuem aprendizado *off-line* são comumente aplicadas para tornar viável o procedimento de análise de alguns dados. Todavia, tais técnicas são inadequadas ou ineficientes para dados que estão constantemente expandindo e/ou mudando ao longo do tempo.

Embora um modelo *off-line* tenha a vantagem de alcançar uma estrutura otimizada a partir de um conjunto de dados de treinamento, o desempenho dele pode cair repentinamente quando uma característica do ambiente muda. Além disso, ele precisa ser sempre reprojetado quando novas circunstâncias ocorrem durante o processo.

Outra desvantagem é que um modelo *off-line* considera que todo o conjunto de dados já esteja disponível durante a fase de treinamento. Contudo, em muitos casos, a aquisição de dados suficientemente representativos do problema é oneroso, consome tempo e, além disso, os dados são usualmente disponíveis em pequenas quantidades ao longo do tempo.

Um novo paradigma no campo de inteligência computacional baseado na construção de modelos a partir de algoritmos incrementais e fluxo de dados tem surgido há aproximadamente uma década atrás. Estes modelos, chamados de modelos evolutivos, não somente minimizam o problema de estocar grandes quantidades de dados, mas também oferecem características importantes para a modelagem de processos adaptativos não lineares. As principais características dos modelos evolutivos são o aprendizado contínuo, a auto-organização e a adaptação à ambientes desconhecidos.

Neste trabalho propomos uma nova técnica de modelo evolutivo chamada de Rede Neural Probabilística Evolutiva. Esta técnica é baseada na Rede Neural Probabilística, no modelo de misturas de gaussianas e na Maximização da Esperança. Sempre que novos dados estão disponíveis para treinamento, que pode ocorrer de forma sequencial, esta rede neural ajusta a sua arquitetura e seus parâmetros para uma melhor adaptação ao problema.

A fim de avaliar nossa abordagem, realizamos uma série de experimentos com algumas bases de dados que estão disponíveis em domínio público. Além de nossa técnica, testamos também outras redes neurais para comparação, que são baseadas no mesmo paradigma de modelos evolutivos. Todas as abordagens foram aplicadas com conhecimento inicial zero, isto é, nenhum conhecimento sobre as bases de dados foi inserido dentro das redes neurais, incluindo o número de classes. O treinamento deles foi realizado em modo incremental, isto é, uma amostra por vez era classificada e então ela era usada para o treinamento.

Realizamos um teste-t bicaudal pareado com 5% de nível de significância sobre a precisão entre cada par de técnicas, e nossa técnica foi estatisticamente superior às outras técnicas. Além disso, nossa abordagem apresentou tempo de processamento e número de neurônios menores do que das outras técnicas (em alguns casos são várias vezes menor). Estes resultados indicam que nossa abordagem é promissora.

PALAVRAS-CHAVE. Rede Neural Probabilística, aprendizado incremental, modelo de misturas gaussianas.