

Rede Neural Granular para Modelagem Evolutiva de Sistemas

Daniel F. Leite, Fernando Gomide

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (Unicamp)
Caixa Postal 6101, CEP 13083-970 – Campinas, SP, Brasil

{danf17, gomide}@dca.fee.unicamp.br

Abstract – In this paper, we propose a framework to granulate and model partially supervised drifting data streams. The modeling approach consists in evolving granular neural networks capable of processing nonstationary data streams from one-pass incremental algorithms. Evolving granular neural classifiers employ fuzzy hyperboxes and T-S neurons to granulate data and aggregate features then proving discriminant boundaries between classes. Additionally, classification results provide interpretable explanation about the model decision from IF-THEN statements. The associated learning algorithm allows model's structural and parametric adaptation whenever the environment changes. The algorithm needs no prior knowledge about statistical properties of data and classes, and can compute data originally numeric or in the form of confidence intervals. The experiments conducted give clues to the behavior of evolving granular neural networks in nonstationary environments. In particular, the approach has demonstrated to be robust against concept drift, and to be able to cope with missing classes.

Keywords – Evolving Intelligent Systems, Data Stream Mining, Concept Drift, Semi-supervised Learning.

1. Introdução

O aumento da disponibilidade de grandes quantidades de dados tem promovido uma busca por novos algoritmos *on-line* para aprendizagem a partir de fluxos de dados [1]-[8]. A mineração de fluxos de dados em ambientes dinâmicos não-estacionários tem criado problemas únicos e inspirado pesquisa na direção do desenvolvimento de abordagens construtivas para modelagem *on-line*.

Em ambientes não-estacionários, a distribuição estatística que gera os dados (médias, variâncias, estrutura de correlação) muda ao longo do tempo. Tais mudanças podem ser graduais ou abruptas, contractíveis ou expansíveis, determinísticas ou aleatórias, ou mesmo cíclicas, devido a sazonalidades. Modelos de aprendizagem *on-line* devem detectar prontamente estas mudanças.

Desafios envolvidos neste contexto compreendem ainda: **(i)** a impossibilidade de armazenar dados históricos. O modelo deve reter o conhecimento previamente adquirido e que ainda é relevante, enquanto usar apenas os dados mais recentes para adaptação; **(ii)** novos dados podem trazer novas características e novas classes. Estes clamam por adaptação estrutural do modelo; **(iii)** dados com ruído e valores perdidos são ocorrências comuns. Claramente, técnicas padrões de mineração de dados, que assumem estacionariedade e requerem múltiplos passos sobre bases de dados se tornam ineficazes neste contexto.

2. Proposta

Sugerimos para o problema da modelagem *on-line* de fluxos de dados não-estacionários uma abordagem baseada em redes neurais granulares evolutivas (eGNN) [6], [8]. eGNN é uma variante evolutiva de sistemas neuro-fuzzy capaz de lidar com ambientes dinâmicos. eGNN originou-se de nossa pesquisa recente em processamento de fluxos de dados sem a necessidade de re-treinamento de modelos. Neste artigo, focamos classificação parcialmente supervisionada.

Colocado de uma maneira geral, classificadores eGNN resumam o comportamento do sistema no espaço característico e as classes associadas a dados de treinamento em regras SENTENÇA. Para isto, eGNN usa hiper-caixas fuzzy para granular dados, e neurônios T-S [6] para agregar características. Seu algoritmo de aprendizagem incrementalmente adapta a estrutura e parâmetros de eGNN tão logo que o ambiente muda. Além disso, o algoritmo pode lidar com dados rotulados e não-rotulados simultaneamente usando um procedimento único. Ele pode processar dados originalmente numéricos ou granulares (na forma de intervalos de confiança). Na próxima seção, descrevemos as características gerais da modelagem eGNN. Na seção 4 avaliamos o comportamento de redes neurais eGNN quando estas são submetidas a mudanças de conceito e a diferentes proporções de dados rotulados. O trabalho é concluído com a Seção 5.

3. Redes Neural Granular Evolutiva

3.1 Introdução

O conceito de redes neurais granulares (GNN) foi inicialmente estabelecido em [9], enquanto que o de eGNN foi proposto em [6]. Ambas as abordagens enfatizam redes neurais artificiais capazes de processar dados originalmente numéricos ou granulares. Entretanto, eGNN focaliza aprendizagem incremental *on-line* a partir de fluxo de dados.

A aprendizagem em GNN e eGNN segue um princípio comum que envolve dois estágios conforme ilustrado na Fig. 1. Primeiro, grânulos de informação – intervalos ou conjuntos fuzzy – são construídos a partir de uma base de representação numérica. Em seguida, a aprendizagem – construção e refinamento – da rede neural é baseada nos grânulos de informação ao invés de nos dados originais. Assim, a rede neural não é exposta a todos os dados de treinamento, muito mais numerosos que os grânulos, e.g., quando não transportando novas informações, exemplos são descartados.

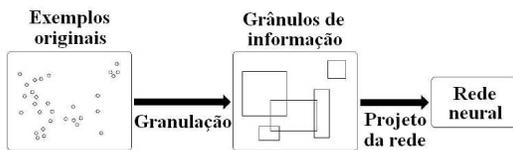


Figura 1. Projeto de redes neurais granulares

Fundamentalmente, modelos eGNN processam dados observando um fluxo somente uma vez. eGNN começa a aprender a partir de uma base de regras vazia e sem conhecimento prévio das propriedades estatísticas dos dados e classes. A abordagem consiste em formar limites discriminantes entre classes a partir da granulação do espaço característico usando hiper-caixas fuzzy. Adicionalmente, os resultados de classificação provêm explicações interpretáveis sobre a decisão do modelo a partir de proposições do tipo SE-ENTÃO. Em suma, entre as características principais de eGNN estão as seguintes, eGNN: ajusta sua estrutura e parâmetros para aprender um novo conceito, enquanto esquece o que não é mais relevante; lida com dados rotulados e não-rotulados utilizando um procedimento único; detecta mudanças no ambiente e lida com incerteza nos dados; possui habilidade não-linear de separação de classes; e desenvolve aprendizado “ao longo da vida” usando mecanismos construtivos *bottom-up* e destrutivos *top-down*.

3.2. Princípio de funcionamento

Redes eGNN aprendem a partir de um fluxo de dados $x^{[h]}$, $h = 1, 2, \dots$, onde os exemplos de treinamento podem ou não ser acompanhados de um rótulo de classe $C^{[h]}$. Cada grânulo de informação γ^i da coleção finita dos grânulos existentes $\gamma = \{\gamma^1, \dots, \gamma^l\}$ definido no espaço característico $X \subseteq \mathcal{R}^n$ é associado a uma classe C_k da coleção finita de classes $C = \{C_1, \dots, C_m\}$ em um espaço de saída $Y \subseteq \mathcal{X}$. eGNN associa o espaço característico e de saída usando grânulos (extraídos do fluxo de dados) e neurônios T-S nos passos de processamento.

A rede neural tem uma estrutura em cinco camadas como ilustrado na Fig. 2. A camada de entrada basicamente insere vetores característicos $x^{[h]} = (x_1, \dots, x_n)^{[h]}$, $h = 1, 2, \dots$, na rede neural; a camada evolutiva consiste de um conjunto de grânulos de informação $\gamma^i \forall i$ formado como um escopo do fluxo de dados. Sobreposição parcial de grânulos são permitidas; a camada de agregação contém os neurônios T-S, $Tsn^i \forall i$. Eles agregam características para gerar medidas de compatibilidade $o^i \forall i$ entre exemplos e grânulos; na camada de decisão, as medidas de compatibilidade são comparadas e a classe C_k associada ao grânulo γ^i que apresentou a maior compatibilidade para um dado exemplo é induzida na saída da rede como uma estimativa para $C^{[h]}$; e a camada de saída (que também evolui durante a aprendizagem), consistindo de rótulos de classes $C_k \forall k$.

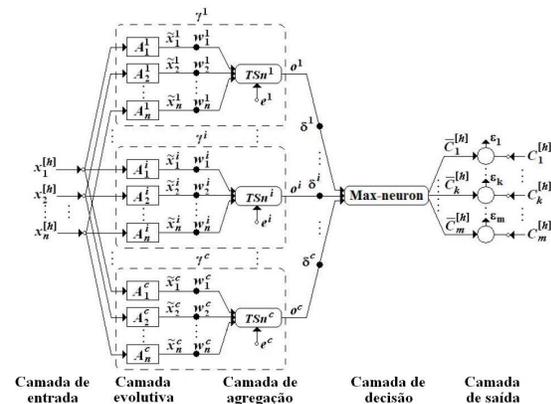


Figura 2. Modelo eGNN para classificação de fluxo de dados

Alternativas para o controle do crescimento estrutural de eGNN compreendem: controle do número de classes (usualmente empregado quando a quantidade de classes nos dados é conhecida); controle automático do número de grânulos na estrutura do modelo (usado quando

memória e tempo de processamento são requerimentos principais); ou nenhum dos citados. A evolução não controlada de grânulos e classes é justificada em ambientes imprevisíveis. A ênfase neste modo é colocada na operação do sistema. Claramente, todos os modos de controle também objetivam desempenho de classificação.

3.3. Aprendizagem em eGNN

Omitiremos as derivações e formulações referentes ao algoritmo de aprendizagem neste trabalho. Pedimos ao leitor que se refira a [6] ou [8] para uma abordagem completa. Basicamente, o algoritmo eGNN compreende os seguintes mecanismos: **(i)** atualização da granularidade; **(ii)** criação/refinamento de grânulos; **(iii)** monitoramento da matriz de distâncias; **(iv)** pré-rotulação de dados não-rotulados; **(v)** adaptação dos pesos da camada de agregação; **(vi)** poda de neurônios e conexões inativas; **(vii)** ajuste de elementos neutros; **(viii)** decisão de classificação; e **(ix)** detecção de *outliers*. Postos de forma sistemática, estes procedimentos formam o algoritmo de aprendizagem eGNN.

4. Resultados Experimentais

Nesta seção provemos resultados empíricos para estabelecer a efetividade da abordagem eGNN. Usamos dados simulados para manter controle sobre os tipos de mudanças nos conceitos e suas implicações na modelagem. Conduzimos dois experimentos com diferentes propósitos. Primeiro, consideramos duas funções Gaussianas parcialmente sobrepostas girando em torno de um ponto central. Cada Gaussiana representa uma classe. Buscamos encontrar um limite discriminante entre as classes tomando por base apenas os exemplos mais recentes. Fazendo isso, avaliamos a habilidade da modelagem eGNN em capturar mudanças graduais do conceito ao longo da evolução. O último experimento ilustra a importância do uso de todas as informações disponíveis no fluxo de dados para guiar o processo de classificação. Além disso, evidenciamos a habilidade de eGNN em classificar fluxos de dados não-estacionários parcialmente rotulados. Maiores detalhes dos experimentos e as parametrizações adotadas podem ser encontrados em [8].

4.1. Rotação das Gaussianas gêmeas

Neste experimento, duas Gaussianas parcialmente sobrepostas, centradas em (4,4) e

(6,6) com desvio padrão .8, giram gradualmente no sentido anti-horário até um ângulo de 90° em torno do ponto central (5,5) como mostra a Fig. 3. Queremos encontrar o limite discriminante entre as classes usando apenas os exemplos mais recentes aleatoriamente selecionados.

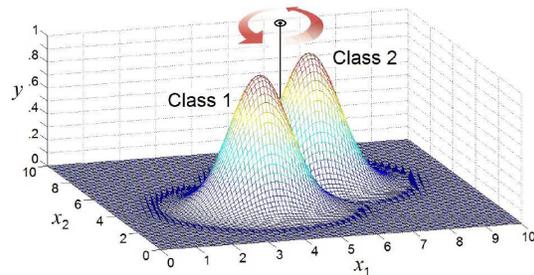


Figura 3. Problema das Gaussianas rotativas

A Fig. 4 mostra o limite de decisão e os últimos 200 exemplos em diferentes instantes da evolução: $h = 200$ (quando a mudança do conceito se inicia) e $h = 400$ (quando a mudança termina). Em $h = 200$, eGNN usa 5 grânulos (dois representando a classe 1 e três representando a classe 2) para modelar os dados. A rede neural obteve um desempenho de classificação correto/errado de 189/11 (94.5%) em $h = 200$. Após a rotação, em $h = 400$, a rede neural utiliza 5 grânulos (três para a classe 1 e o restante para a classe 2) e alcança um desempenho de 195/5 (97.5%).

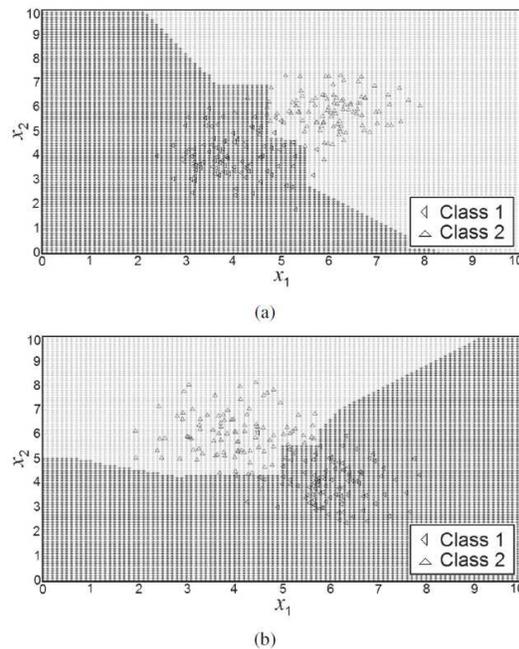


Figura 4. Limite de decisão e últimos 200 exemplos em (a) $h = 200$; e (b) $h = 400$

Enquanto hiper-caixas foram preferidas neste trabalho como representação dos grânulos de informação, nós evidenciamos neste experimento que tal consideração não restringe eGNN a lidar somente com distribuições de dados da mesma natureza (na forma de intervalos de confiança). Ao contrário, eGNN não precisa de informação prévia sobre os dados para capturar as mudanças de conceito ocorrendo no fluxo de dados.

4.2. Dados rotulados e não-rotulados

Aqui, o desempenho de eGNN é investigado considerando variações da proporção de dados não-rotulados entre 0% e 100%. Admitimos para isso o problema das Gaussianas rotativas e também um problema de surgimento repentino de uma nova classe no fluxo de dados [8]. A Fig. 5 ilustra o desempenho médio do algoritmo em 5 simulações para cada ponto do gráfico.

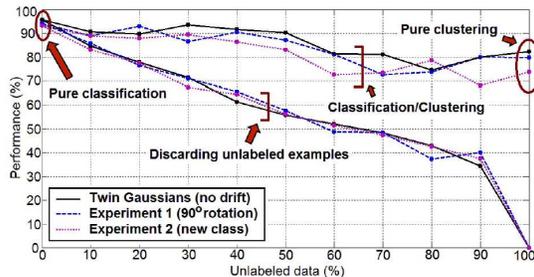


Figura 5. Desempenho de eGNN usando frações de dados não-rotulados

Referindo-se a Fig. 5, removemos sucessivamente a informação referente ao rótulo dos exemplos. A variação em desempenho se mostrou robusta ao número crescente de rótulos perdidos. A performance da rede eGNN com 50% de exemplos não-rotulados degradou ligeiramente em ambos os experimentos usando aprendizagem híbrida. Além disso, claros benefícios da combinação de dados rotulados e não-rotulados no treinamento foram noticiados. Os experimentos aproveitando toda a informação do fluxo de dados guiaram a classificação com sucesso, enquanto aqueles experimentos simplesmente descartando exemplos não-rotulados degradaram o poder de reconhecimento do modelo proporcionalmente a quantidade de dados descartados.

5. Conclusão

Uma abordagem híbrida baseada em redes neuro-fuzzy evolutivas para classificação e clusterização de fluxo de dados não-estacionários (especi-

ficamente, rede neural granular evolutiva) foi sugerida neste trabalho. Verificamos que a abordagem eGNN pode classificar dados sujeitos à mudanças de conceito em tempo real. Testamos a efetividade do algoritmo em dois experimentos compreendendo deslize gradual de conceito, e diferentes proporções de dados rotulados. Os resultados mostraram que o algoritmo é robusto a ambientes dinâmicos e também capaz de capturar não-estacionariedades. Trabalhos futuros na direção dos fundamentos da modelagem considerarão processamento de dados originalmente granulares; e aprendizagem participativa. Na direção de novas aplicações, admitiremos problemas de controle e previsão de séries.

Referências

- [1] P. Angelov; D. Filev. An approach to on-line identification of evolving Takagi-Sugeno models. *IEEE Trans. on SMC – Part B*, 34(1): 484-498, 2004.
- [2] B. Grabrys; A. Bargiela. General fuzzy min-max neural network for clustering and classification. *IEEE Trans. on Neural Networks*, 11(3): 769-783, 2000.
- [3] M. Muhlbaier; A. Topalis; R. Polikar. Learn ++.NC: Combining Ensemble of Classifiers With Dynamically Weighted Consult-and-Vote for Efficient Incremental Learning of New Classes. *IEEE Trans. on Neural Networks*, 20(1): 152-168, 2009.
- [4] S. Ozawa; S. Pang; N. Kasabov. Incremental Learning of Chunk Data for Online Pattern Classification Systems. *IEEE Trans. on Neural Networks*, 19(6): 1061-1074, 2008.
- [5] D. Leite; P. Costa Jr.; F. Gomide. Interval-based evolving modeling. *IEEE Workshop on Evolving and Self-Developing Intelligent Systems*, 1: 1-8, 2009.
- [6] D. Leite; P. Costa Jr.; F. Gomide. Evolving Granular Classification Neural Networks. *Int. Joint Conference on Neural Networks*, 1: 1736-1743, 2009.
- [7] D. Leite; P. Costa Jr.; F. Gomide. Granular Approach for Evolving System Modeling. *Lecture Notes in Computer Science (to appear)*, 2010.
- [8] D. Leite; P. Costa Jr.; F. Gomide. Evolving Granular Neural Networks for Semi-supervised Data Stream Classification. *World Congress on Computational Intelligence (submitted)*, 2010.
- [9] W. Pedrycz; W. Vukovich. Granular Neural Networks. *Neurocomputing*, 36: 205-224, 2001.