

Universidade de Pernambuco

Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Tese de Doutorado

Área: Inteligência Computacional

Título: Sistemas Híbridos baseados Ensemble dinâmicos para aproximação de séries temporais e seleção de características.

Orientador – Roberta Andrade de A. Fagundes (roberta.fagundes@upe.br)

Co-orientador - João Fausto Lorenzato (fausto.lorenzato@upe.br)

Descrição

A inteligência computacional [1] é uma técnica tem como objetivo fazer as máquinas resolverem problemas por meio da observação das pessoas ou apenas por meio do estudo de seus próprios métodos como por exemplo, *ensemble regression*, redes neurais, árvores de decisão, algoritmos genéticos, *extreme learning machine* e etc. A previsão é o principal propósito dos modelos de regressão. O objetivo dos modelos de regressão é construir uma função, a partir de um conjunto de variáveis independentes, para estimar o valor das variáveis dependentes [2]. O objetivo do *ensemble regression* [4] é combinar vários modelos para melhorar a precisão da previsão em problemas de aprendizagem com uma variável alvo numérica. O processo de *ensemble learning* pode ser dividido em três fases: a fase de geração, a fase de poda e a fase de integração. Discutimos diferentes abordagens para cada uma dessas fases que são capazes de lidar com o problema da regressão, categorizando-as em termos de suas características relevantes e vinculando-as a contribuições de diferentes campos.

A seleção de conjunto visa escolher um subconjunto de um conjunto de modelos considerados os melhores preditores para uma instância de teste x . Os métodos de seleção podem ser estáticos, o que significa que eles selecionam um subconjunto fixo do conjunto original para todas instâncias de teste. Uma abordagem alternativa determinaria a subconjunto de especialistas dinamicamente de acordo com a entrada atual instâncias de teste, isso é conhecido como seleção dinâmica de conjunto (DES) também chamado baseado em instância [4]. A razão de usar abordagens DES é que diferentes modelos têm diferentes áreas e expertise no espaço de instâncias. Portanto, abordagens estáticas que são forçados a selecionar um subconjunto fixo antes de ver uma instância não rotulada podem ter uma desvantagem teórica em comparação aos dinâmicos [4].

Sistemas híbridos [3], que combinam estatísticas e técnicas de aprendizado de máquina (ML) usando modelagem residual (previsão de erros), têm se destacado na literatura devido a sua precisão e capacidade de prever séries temporais com diferentes características. Portanto, a seleção, especificação e treinamento de um modelo de ML prever os resíduos são tarefas caras e desafiadoras, pois problemas, como subajuste, superajuste e especificação incorreta, podem levar a um sistema com baixa precisão ou até mesmo deteriorar a linearidade previsão da série temporal. Este trabalho propõe um sistema híbrido, que emprega um algoritmo de seleção dinâmica (DS) modificado para decidir: o modelo de ML adequado e posteriormente, aumentar a precisão da série temporal prevista a partir da combinação linear. Visando, reduzir a incerteza do modelo ML seleção e evitar a deterioração da previsão da série temporal.

Referências Bibliográficas

1. NORVIG P. e RUSSELL S. Inteligência Artificial, 3ª Edição, 2013.
2. MONTGOMERY D.C.; PECK, E.A. e VINING, G. G. Introduction t Linear Regression Analysis, Wisley-Interscience, 2006.
3. J. F. L. de Oliveira, E. G. Silva and P. S. G. de Mattos Neto, "A Hybrid System Based on Dynamic Selection for Time Series Forecasting," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2021.3051384.
4. D. Boulegane, A. Bifet, H. Elghazel and G. Madhusudan, "Streaming Time Series Forecasting using Multi-Target Regression with Dynamic Ensemble Selection," 2020 IEEE International Conference on Big Data (Big Data), 2020, pp. 2170-2179, doi: 10.1109/BigData50022.2020.9378264.
5. M. Breskvar and S. Džeroski, "Multi-Target Regression Rules With Random Output Selections," in IEEE Access, vol. 9, pp. 10509-10522, 2021, doi: 10.1109/ACCESS.2021.3051185.

