Anais do II Congresso de Ciência e Tecnologia da PUC Goiás
Disponível em: http://pucgoias.edu.br/ucg/prope/pesquisa/anais/2016
ISSN: 2177-3327

PARTICIONAMENTO DE CONJUNTO DE DADOS E SELEÇÃO DE VARIÁVEIS EM PROBLEMAS DE CALIBRAÇÃO MULTIVARIADA

ANDRÉ LUIZ ALVES, CLARIMAR JOSÉ COELHO
andre.luiz@pucgoias.edu.br

Objetivo: Desenvolvimento de método para a seleção de amostras e seleção de variáveis em conjunto de dados no contexto da calibração multivariada empregando o método do consenso de amostra aleatória. **Método:** Aplicação do algoritmo RANSAC – RANdom SAmple Consensus – para seleção de amostras (seu propósito é a identificação e remoção de pontos discrepantes) combinado com o APS – Algoritmo das Projeções Sucessivas – Successive Projections Algorithm – para a seleção de variáveis (seu propósito é atuar na presença de multicolinearidade). O método proposto será construído a partir de implementação destes algoritmos em uma linguagem de programação científica, cuja forma combinada será objeto de experimentação e a configuração de melhor desempenho será a adotada. Serão realizados processamentos com dados obtidos de experimentos relacionados à quimiometria e os resultados comparados, avaliados e relatados. Em seguida realizar a aplicação com dados obtidos de outros experimentos com o propósito de se avaliar sua capacidade de generalização. **Resultados:** Espera-se contribuir com o estado da arte em seleção de amostras e seleção de variáveis em calibração multivariada a partir da adaptação de um método robusto usado com sucesso em visão computacional. Estudos preliminares realizados no grupo de pesquisa/MEPROS onde será desenvolvimento esse projeto indicam que os resultados serão promissores. Pretende-se obter vantagens da seleção de amostras e seleção de variáveis simultaneamente, bem como explorar a robustez do método do consenso de amostra aleatória para a construção de modelos de calibração multivariada com boa capacidade preditiva. **Conclusão:** Como este trabalho encontra-se na fase inicial, conclusões ainda não há. Porém, alguns relatos científicos têm mostrado resultados positivos em aplicações de técnicas de seleção de amostras, enquanto outros trazem resultados também positivos em aplicações de técnicas de seleção de variáveis. A combinação de duas destas técnicas é uma nova e promissora forma de se atacar os problemas citados, que poderá ser trazer avanços e abrir novas perspectivas no campo da análise multivariada de dados para solução de problemas calibração.

Palavras-chave: Seleção de amostras. Calibração multivariada. Outliers.