



Título: Robust Algorithms for Linear Regression and Locally Linear Embedding.

Data: 22/03/2017 Horário: 09h Local: Sala de Seminários - Bloco 952 - Campus do Pici

Resumo:

Na atualidade um grande volume de dados é produzido na nossa sociedade digital. Existe um crescente interesse em converter esses dados em informação útil e o aprendizado de máquinas tem um papel central nessa transformação de dados em conhecimento. Por outro lado, a probabilidade dos dados conterem outliers é muito alta para ignorar a importância dos algoritmos robustos. Para se familiarizar com isso, são estudados vários modelos de outliers. Neste trabalho, discutimos e analisamos vários estimadores robustos dentro do contexto dos modelos de regressão linear generalizados: são eles o M-Estimator, o S-Estimator, o MM-Estimator, o RANSAC e o Theil-Senestimator. A escolha dos estimadores é motivada pelo princípio de explorar algoritmos com distintos conceitos de funcionamento. Em particular os estimadores M, S e MM são baseados na modificação do critério de minimização dos mínimos quadrados, enquanto que o RANSAC se fundamenta em achar o menor subconjunto que permita garantir uma acurácia predefinida ao modelo. Por outro lado o Theil-Sen usa a mediana de modelos obtidos usando mínimos quadrados no processo de estimação. O desempenho dos estimadores em uma ampla gama de condições experimentais é comparado e analisado. Além do problema de regressão linear, considera-se o problema de redução da dimensionalidade. Especificamente, são tratados o Locally Linear Embedding, o Principal Component Analysis e outras abordagens robustas destes. É proposto um método denominado RALLE com a motivação de prover de robustez ao algoritmo de LLE. A ideia principal é usar vizinhanças de tamanhos variáveis para construir os pesos dos pontos; para fazer isto possível, o RAPCA é executado em cada grupo de vizinhos e os pontos sob risco são descartados da vizinhança correspondente. É feita uma avaliação do desempenho do LLE, do

RLLE e do RALLE sobre algumas bases de dados.

Banca:

- Prof. Dr. João Fernando Lima Alcântara (MDCC/UFC - Orientador)
- Prof. Dr. Francesco Corona (DC/UFC - Coorientador)
- Prof. Dr. Amauri Holanda de Souza Júnior (IFCE)
- Prof. Dr. João Paulo Pordeus Gomes (MDCC/UFC)