



***INVESTIGAÇÃO DO PROBLEMA DE DESCOBERTA CAUSAL POR MEIO DA
DETERMINAÇÃO DE HIPERPLANOS DE SEPARAÇÃO***

Igor Correia da Silva¹, Eanes Torres Pereira ²

RESUMO

Grande parte das pesquisas em aprendizado de máquina atualmente se baseia em associações estatísticas entre variáveis, mas isso pode trazer resultados inesperados ao trabalharmos com correlações espúrias, pares de variáveis que possuem um alto grau de correlação, porém não apresentam nenhuma relação causal, sendo necessária a criação de mais métodos para a descoberta causal. Esta pesquisa investigou abordagens para determinar relações causais a partir de hiperplanos de separação, usando como ferramenta máquinas de vetores de suporte (Support Vector Machines, SVM), procurando evidências para aceitar ou rejeitar duas conjecturas relacionadas à descoberta causal. A partir dessa pesquisa, concluímos que não é possível definir a relação causal com base nos pesos dados a cada variável em uma predição usando SVMs, mas pudemos notar que o desempenho de uma predição está ligado à relação causal, mostrando potencial para filtrar parte dos pares de variáveis que não possuem relação causal.

Palavras-chave: Causalidade, Modelos Estruturados, Máquinas de Vetores de Suporte.

¹Aluno do curso de Ciência da Computação, Unidade Acadêmica de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: igor.correia.silva@ccc.ufcg.edu.br

²Doutor, Professor, UASC, UFCG, Campina Grande, PB, e-mail: eanes@computacao.ufcg.edu.br



***INVESTIGATION OF THE PROBLEM OF CAUSAL DISCOVERY BY
DETERMINING SEPARATION HYPERPLANS.***

ABSTRACT

Much of the current research in machine learning relies on statistical associations between variables, but this can lead to unexpected results when dealing with spurious correlations, which are pairs of variables that have a high degree of correlation but no causal relationship. Then, it is necessary to develop more methods for causal discovery. This research investigated approaches to determining causal relationships from separating hyperplanes, using Support Vector Machines (SVM) as a tool, seeking evidence to either accept or refute two conjectures related to causal discovery. From this research, we concluded that it is not possible to define causal relationships based on the weights assigned to each variable in a prediction using SVM. However, we observed that prediction performance is linked to causal relationships, showing potential for filtering out some pairs of variables that have no causal relationship.

Keywords: Causality, Structured Models, Support Vector Machine.