HOW FAR DO WE GET USING MACHINE LEARNING BLACK-BOXES?

Abstract
With several good research groups actively working in machine learning (ML) approaches, we have now the concept of self-containing machine learning solutions that oftentimes work out-of-the-box leading to the concept of ML black-boxes. Although it is important to have such black-boxes helping researchers to deal with several problems nowadays, it comes with an inherent problem increasingly more evident: we have observed that researchers and students are progressively relying on ML black-boxes and, usually, achieving results without knowing the machinery of the classifiers. In this regard, this paper discusses the use of machine learning black-boxes and poses the question of how far we can get using these out-of-the-box solutions instead of going deeper into the machinery of the classifiers. The paper focuses on three aspects of classifiers: (1) the way they compare examples in the feature space; (2) the impact of using features with variable dimensionality; and (3) the impact of using binary classifiers to solve a multi-class problem. We show how knowledge about the classifier's machinery can improve the results way beyond out-of-the-box machine learning solutions. (AU)

FAPESP's process:
10/05647-4 - Digital forensics: collection, organization, classification and analysis of digital evidences
Grantee:
Anderson de Rezende Rocha
Support type:
Research Grants - Young Investigators Grants

Author(s):
Total Authors: 3

Affiliation:
[1] Univ Campinas UNICAMP, Inst Comp, BR-13083852 Campinas, SP - Brazil
[3] Univ Campinas UNICAMP, Fac Technol, BR-13484332 Limeira, SP - Brazil
Total Affiliations: 3

Document type: Journal article Source: INTERNATIONAL JOURNAL OF PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE; v. 26, n. 2 MAR 2012. Web of Science Citations: 6