Future Generation Computer Systems 78 (2018) 59–76

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Kuaa: A unified framework for design, deployment, execution, and recommendation of machine learning experiments

Rafael de Oliveira Werneck [a,*], Waldir Rodrigues de Almeida [a], Bernardo Vecchia Stein [a], Daniel Vatanabe Pazinato [a], Pedro Ribeiro Mendes Júnior [a], Otávio Augusto Bizetto Penatti [a,b], Anderson Rocha [a], Ricardo da Silva Torres [a]

[a] RECOD Lab., Institute of Computing (IC), University of Campinas (Unicamp), Av. Albert Einstein, 1251, Campinas, SP, 13083-852, Brazil
[b] Advanced Technologies Group, SAMSUNG Research Institute, Av. Cambacica, 1200, Building 1, Campinas, SP, 13097-160, Brazil

HIGHLIGHTS

- A framework for designing and deploying machine-learning experiments.
- Standardized environment for exploratory analysis of machine-learning solutions.
- The modeling of a machine-learning experiment as a workflow.
- A framework capable of recommending machine-learning workflows to the user.
- Evaluation of four similarity measures and a learning-to-rank method for recommending workflows.

ARTICLE INFO

Article history:
Received 7 February 2017
Received in revised form 20 April 2017
Accepted 13 June 2017
Available online 18 July 2017

Keywords:
Machine learning
Science - experiments
Workflow
Recommendation systems (Information filtering)

ABSTRACT

In this work, we propose Kuaa, a workflow-based framework that can be used for designing, deploying, and executing machine learning experiments in an automated fashion. This framework is able to provide a standardized environment for exploratory analysis of machine learning solutions, as it supports the evaluation of feature descriptors, normalizers, classifiers, and fusion approaches in a wide range of tasks involving machine learning. Kuaa also is capable of providing users with the recommendation of machine-learning workflows. The use of recommendations allows users to identify, evaluate, and possibly reuse previously defined successful solutions. We propose the use of similarity measures (e.g., Jaccard, Sørensen, and Jaro–Winkler) and learning-to-rank methods (LRAR) in the implementation of the recommendation service. Experimental results show that Jaro–Winkler yields the highest effectiveness performance with comparable results to those observed for LRAR, presenting the best alternative machine learning experiments to the user. In both cases, the recommendations performed are very promising and the developed framework might help users in different daily exploratory machine learning tasks.

1. Introduction

With data deluge becoming ever more commonplace and pervasive due to spectacular advances in hardware and software acquisition technologies, it becomes imperative to properly process such data for extracting information that can lead to knowledge generation. This knowledge extraction process is usually performed by means of data mining and machine-learning methods, with the ultimate goal being the improvement of the decision-making process in a target application [1].

A typical machine-learning solution comprises several steps, including, for example, feature extraction and normalization methods, and the definition of appropriate classifiers. Since there is no *silver bullet* that solves all machine learning problems, each technique has its own pros and cons when designed for specific applications. In this sense, one common strategy adopted for developers of machine learning systems consists in performing exploratory analysis often relying on running several experiments with the objective of identifying which techniques are more appropriate for a given application.

* Corresponding author.
E-mail addresses: rafael.werneck@ic.unicamp.br (R.de O. Werneck), waldir.r.almeida@gmail.com (W.R. de Almeida), bernardovstein@gmail.com (B.V. Stein), danielvatanabe92@gmail.com (D.V. Pazinato), pedrormjunior@gmail.com (P.R.M. Júnior), o.penatti@samsung.com (O.A.B. Penatti), anderson.rocha@ic.unicamp.br (A. Rocha), rtorres@ic.unicamp.br (R.da S. Torres).

http://dx.doi.org/10.1016/j.future.2017.06.013
0167-739X/

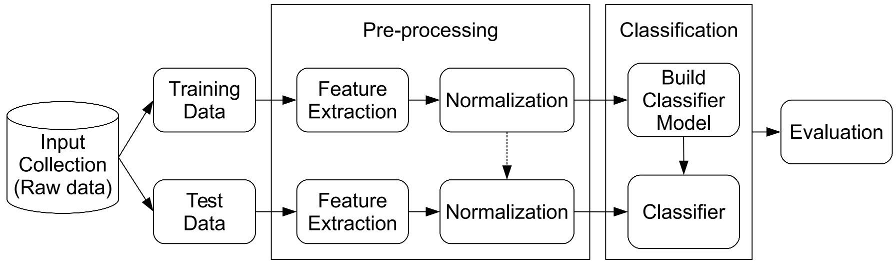

Fig. 1. Organization of a typical machine-learning experiment, composed of a collection input, a split of the collection into train and test sets, extraction of the feature descriptions of the collection, normalization of these features, classification of these features, and the evaluation of the classification results.

Several libraries and machine-learning frameworks have been proposed in the literature to support users in the process of defining the most appropriate methods for their applications. However, many frameworks have limitations including the lack of flexibility to include novel proposed descriptors and machine learning methods, and specially, the inability to reuse previous experiments and learn from them.

In this work, we address these issues by presenting Kuaa, a framework that can be used for designing, deploying, and executing machine learning experiments in an automated fashion. This framework is able to provide a standardized environment for exploratory analysis of machine-learning solutions, as it supports the evaluation of feature descriptors, normalizers, classifiers, and fusion approaches in a wide range of tasks involving machine learning. The Kuaa conceptual model relies on modeling machine-learning experiments as scientific workflows [2,3]. Workflow is the automation of a process, in which information is passed from one resource to another for action, according to a set of rules. The advantages of using workflows are that they are easily understandable, flexible, and reproducible, in which it is possible to redesign them and reproduce their results. The Kuaa's implementation relies on the use of *plugins*, which supports the incorporation of new machine-learning methods as workflow components into the framework, making it flexible to be used in different exploratory analysis.

We also empowered Kuaa with the capability of recommending machine-learning workflows. This service is useful even for experienced users, but specially for beginners in machine learning, as it may guide the user during the configuration of an experiment when facing new and challenging classification problems. The use of recommendations allows users to identify, evaluate, and possibly reuse previously defined successful machine-learning solutions. We also performed experiments in the recommendation system aiming at evaluating four similarity measures (Jaccard, Sørensen, Jaro–Winkler, and a TF–IDF-based measure) in order to define which one is more appropriate for ranking workflows. In addition, we performed experiments with the Learning to Rank using Association Rules (LRAR) method with the objective of comparing it with the methods that do not use any learning mechanism.

Along with this paper, we consider two scenarios in which Kuaa was used to support the identification of appropriate machine-learning solutions. The first one is related to the heart-views classification problem [4], which refers to the automatic recognition of heart view plane of 2D echocardiogram ultrasound images. The second one refers to the produce recognition problem [5,6], which refers to the automatic recognition of fruits and vegetables based on their visual properties. In both contexts, machine-learning-based system developers were interested in comparing several image descriptors (e.g., color, texture, and mid-level representations). By using Kuaa, researchers were able to include novel representations; reuse, based on recommendations, existing workflows previously defined in different contexts; and design, deploy, and perform experiments to assess the effectiveness of tested solutions.

This paper is organized as follows. Section 2 introduces background concepts and presents related work. Section 3 describes the proposed machine-learning framework. Case studies are presented in Section 4, in which we show the use of Kuaa. Section 4 also presents an overview of Kuaa's recommendation system and the experiments performed to evaluate different similarity measures used in the system. Finally, Section 5 presents our conclusions and proposes research directions for future work.

2. Related work and background concepts

This section introduces preliminary concepts related to our proposal, as well as related work on exploratory analysis using machine learning tools and workflow recommendation approaches.

2.1. Machine learning

Machine learning is the study of computational methods that extract useful knowledge from experience to improve performance of a target application [7]. Fig. 1 presents the typical six-step machine-learning classification experiment. Once the input collection is defined, the method selected for splitting it into train and test sets is executed and a feature descriptor is then employed to extract a feature vector from each object within the collection. If a normalization method is selected, the feature vectors of all data in the train and test sets are normalized accordingly. After that, a classification method is applied and the results are analyzed considering the chosen evaluation measure.

2.2. Exploratory analysis using machine learning tools

Several frameworks have been proposed to execute machine-learning experiments. Among the frameworks we may refer to PyML[1], Accord.NET[2], mlf [8], and Rattle [9]. PyML (see footnote 1) is an interactive object oriented framework for machine learning written in Python. The framework has implemented the most used classifiers, as Support Vector Machines [10] and Nearest Neighbor [11]. This framework allows combining classifiers and testing classifiers using a typical evaluation process (cross-validation, ROC curves). Accord.NET (see footnote 2) is a framework that provides several scientific computing related methods, such as machine learning, statistics, and computer vision, to the .NET environment. The machine learning framework for Mathematica [3] (mlf) [8] is a collection of machine-learning algorithms for intelligent data analysis, combining an optimized kernel with the manipulation, descriptive programming and graphical capabilities of Mathematica.

1 http://pyml.sourceforge.net/ (December 2016).
2 http://accord.googlecode.com (December 2016).
3 Mathematica is a registered trademark of Wolfram Research Inc. www.wolfram.com (December 2016).