Chapter 5
Machine Learning Applications in Radiation Therapy

Hao H. Zhang
University of Maryland School of Medicine, USA

Robert R. Meyer
University of Wisconsin-Madison, USA

Leyuan Shi
University of Wisconsin-Madison, USA

Warren D. D'Souza
University of Maryland School of Medicine, USA

ABSTRACT

Cancer is one of the most complex diseases and one of the most effective treatments, radiation therapy, is also a complicated process. Informatics is becoming a critical tool for clinicians and scientists for improvements to the treatment and a better understanding of the disease. Computational techniques such as Machine Learning have been increasingly used in radiation therapy. As complex as cancer is, this book chapter shows that a machine learning technique has the ability to provide physicians information for better diagnostic, to obtain tumor location for more accurate treatment delivery, and to predict radiotherapy response so that personalized treatment can be developed.

INTRODUCTION

Cancer is a leading cause of death worldwide and can affect people at all ages. There are over one million cases of cancer diagnosed each year just in the United States, and many times that number in other countries. About 60% of US cancer patients are treated with radiation therapy, and increasingly complex radiation delivery procedures are being developed in order to improve treatment outcomes. A key goal of radiation therapy is to determine appropriate values for a large set of delivery parameters in order to ensure that as large a fraction as possible of the radiation that enters the patient is delivered to the tumor as opposed to depositing it in adjacent non-cancerous organs that can be damaged by radiation (the latter are termed organs-at-risk (OARs)). As complex as

DOI: 10.4018/978-1-4666-1833-6.ch005

cancer is, radiation therapy is also a complicated process. Informatics is becoming a critical tool for clinicians and scientists for improvements to the treatment and a better understanding of the disease. Computational techniques such as Machine Learning (ML) have been increasingly used in radiation therapy to help accurately localize the tumors in images, precisely target the radiation to the tumors, analyze treatment outcomes, and improve treatment quality and patient safety.

Machine learning tools are commonly used to extract implicit, previously unknown, and potentially useful information from data. This information, which is expressed in a comprehensible form, can be used for a variety of purposes. The idea is to build programs or models that sift through raw data automatically, seeking regularities or patterns. Strong patterns, if found, will likely generalize to yield accurate predictions on future data. Machine learning algorithms need to be robust enough to cope with imperfect data and to extract regularities that are inexact but useful. In order to achieve this, machine learning algorithms typically involve solving rigorous mathematical optimization (linear or non-linear) programs to obtain the coefficients for describing regression models or to derive rules, trees and networks for classification.

Recently, machine learning has gained great popularity in many aspects of cancer research, including tumor localization, prediction of radiotherapy response and image processing and pattern recognition. Regression methods are essential to any cancer data analysis which attempts to describe the relationship between a response variable (outcome) and any number of predictor variables (input features). Regression analysis helps us understand how the typical value of the outcome changes when any one of the predictor variables is varied, while the other predictor variables are held fixed. Most commonly used methods include linear regression and ordinary least squares regression, in which the regression function is defined in terms of a finite number of unknown coefficients that are estimated from the data.

Frequently in medical image analysis, situations involving discrete variables arise. In this circumstance, machine learning still plays an essential role, because objects such as lesions, cancer foci and organs in medical images cannot be modeled accurately by simple equations. Thus, it is natural that tasks in medical image analysis require essentially “learning from examples”. Logistic regression analysis extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical, that is, taking on two or more possible values. In cancer research, the goal of logistic regression analysis is to find the best fitting and most parsimonious, yet biologically reasonable model to describe the relationship between an outcome and a set of predictor or explanatory variables. But logistic regression requires many data points to ensure the stability of the model and has a disadvantage with respect to interpretability of the model in the face of multicollinearity.

One of the most popular uses of machine learning in medical image analysis is the classification of objects such as lesions into certain categories (e.g. abnormal or normal, lesions or non-lesions). This class of machine learning uses features (e.g. diameter, contrast, and circularity) extracted from segmented objects as information for classifying objects. Most commonly used techniques include artificial neural networks, support vector machines and decision trees. These methods involve solving an optimization problem in which the objective function has a measure of the errors in the model (e.g.: squared error) and may include a term that measures the complexity of the model (e.g.: norm of the weights of input features). An example of one such technique is the use of a sequential minimal optimization algorithm for “training” a support vector regression model, which employs a quadratic data fitting problem whose objective function is comprised of a weighted combination of two terms: the first term is a quadratic error

Related Content