Predicting Response to Chemotherapy in Breast Cancer Patients using Machine Learning Techniques

E Gyftodimos[1], L Moss[1], D Sleeman[1], A Welch[2]

[1] Department of Computing Science, University of Aberdeen, Aberdeen AB24 3UE, UK.

[2] School of Medical Sciences, University of Aberdeen, UK.

Keywords: Positron emission tomography (PET), breast cancer monitoring, neoadjunctive chemotherapy, machine learning.

Abstract

We discuss the use of machine learning algorithms to predict which breast cancer patients are likely to respond to (neoadjunctive) chemotherapy. This group of 96 patients from the Aberdeen Royal Infirmary had their size of their tumours assessed by Positron Emission Tomography at various stages of their chemotherapy treatment; the study was attempting to predict the non-responders who would then be progressed more quickly to surgery. A variety of machine learning algorithms were used with this data set. Results indicate that machine learning methods outperform statistical approaches on the same data set.

Introduction

Each year, more than 41,000 people are newly diagnosed with breast cancer[1]. Up to 25% of these patients have large (>3cm) tumours[2]. For these patients, neoadjunctive chemotherapy is sometimes offered in an attempt to reduce the size of the tumour before surgery[3,4]. It is estimated that up to 25% of these patients do not respond to this chemotherapy[2], therefore it would be beneficial to predict which patients do not respond positively to this chemotherapy. Methods to detect the response of a breast cancer tumour to neoadjunctive chemotherapy include the use of Positron Emission Tomography (PET)[5,6] scans.

Machine learning, a subfield of artificial intelligence, involves the automatic construction of rules and models of a domain from previous observations[7]. In the learning phase classification algorithms accept as input a finite set of observations each of which is associated with a label (or class); the output of a classification algorithm is a rule set or model that accepts an unseen observation and predicts a value for its class.

The present research involves the application of various machine learning algorithms to data acquired and analysed in previous studies[2,8]. McDermott et al.[8] investigated optimum times for imaging when using PET to predict response to neoadjunctive chemotherapy. By measuring the mean standard uptake value at the midpoint of neoadjunctive chemotherapy, they identified 77% of the low responding patients whilst identifiying 100% of high responding patients, achieving an ROC area of 93%.

Experimental Methodology

The original study involved clinical data for 96 breast cancer patients. The patients underwent 6 or 8 cycles of chemotherapy treatment before having the tumour removed surgically. Medical imaging data on the tumour region were gathered for each patient using PET at four time points: before and after the first chemotherapy cycle, at the midpoint and at the endpoint of chemotherapy treatment prior to surgery.

The following data were available for each patient: age, pre-therapy body surface, pathological response (i.e. tumour shrinkage after chemotherapy), and for each of the 4 PET scans, the injected activity, image contrast, three different measurements of the activity in the tumour region, and the derived metabolic volume of the tumour. The initial data have been pre-processed in several ways before applying the various machine learning algorithms. Data discretisation using equal frequency intervals and subset feature selection resulted in significantly improved results compared to running the same algorithms on the initial data set. To be consistent with the previous statistical analysis of the data, only these patients with contrast values >5.0 in the pre-therapy scan were included in the study.

Prediction at different points of treatment: The data was split into three further versions with respect to the timings of the scans. One version contained all the information from the pre-therapy scan, the first scan after the start of chemotherapy, and the changes between the two scans. The second version contained the pre-therapy, first scan and midpoint scan information. The third version contained all the scan data.

Missing values: The previous analysis ignored records with missing values, i.e. patients who have missed one or more scans. Most of the machine learning algorithms we have tested are able to handle missing values. We have created two data sets, one which removes patients with missing data and the second which includes extrapolated data.

Classification: A number of classification algorithms were tested on the different versions of the data using the WEKA data mining software[9]. We have compared the performance of several decision tree algorithms, rule learning algorithms and Bayesian learning algorithms[10]. Stratified 10-fold cross validation was applied for evaluation. Two measures were used to assess the performance of a classification model, namely specificity rate at 100% sensitivity, and area under the ROC curve.

Results

We report the area under the ROC curve measure and the specificity rate at 100% sensitivity (SPC) for the top-scoring algorithms in our experiment, obtained on various settings. We refer to the experimental setting with removed missing values, pre-therapy and start of therapy data as NoMV1, and to the one including the midpoint of therapy data as NoMV2. The respective settings including records with missing values are referred to as MV1 and MV2.

In the setting NoMV1, the Bayesian network learning algorithm achieved an ROC = 87.9% and SPC = 70.6%. In the setting NoMV2, the naive Bayes algorithm scored an ROC = 96.4% and SPC = 96.0%. This setting is the one which has been considered in the earlier study and therefore provides a **direct** comparison of the performance of the different approaches. In the setting MV1, the tree-augmented naive Bayes algorithm had ROC = 87.3% and SPC = 60.0%. Finally, in the setting MV2, naive Bayes achieved ROC = 95.6% and SPC = 84.4%.

Conclusions

We have discussed the use of machine learning algorithms for analysing PET imaging to predict response to chemotherapy in breast cancer patients. We have evaluated several algorithms using real-world clinical data. Machine learning algorithms outperform previous methods applied to the same data. An additional advantage of these algorithms is that they are applicable to clinical cases where missing values occur. Furthermore, machine learning techniques were able to predict a significant proportion of patients with low response to chemotherapy before the second chemotherapy cycle, while previous statistical approaches achieved comparable results only when the midpoint data was included; earlier prediction is of course clinically important.

References

1. Cancer Research UK.

2. Smith IC, Welch AE, Hutcheon AW, Miller ID, Payne S, Chilcott F, Waikar S, Whitaker T, Ah-See AK, Eremin O, Heys SD, Gilbert FJ, Sharp PF, Positron Emission Tomography Using [18F]-Fluorodeoxy-D-Glucose to Predict the Pathologic Response of Breast Cancer to Primary Chemotherapy, Journal of Clinical Oncology, 18, pp. 1676-1688, 2000.

3. Bonadonna G, Valagussa P, Zucali R, et al., Primary Chemotherapy in surgically resectable breast cancer, CA Cancer J Clin, 45, pp. 227-243, 1995.

4. Fisher B, Brown A, Mamounas E, et al., Effect of preoperative chemotherapy on loco-regional disease in women with operable breast cancer: Findings from National Surgical Adjuvant Breast and Bowel Project B-18, J Clin Oncol, 15, pp. 2483-2493, 1997.

5. Gennari A, Donati S, Salvadori B, Giorgetti A, Salvadori PA, Sorace O, Puccini G, Pisani P, Poli M, Dani D, Landucci E, Mariani G, Conte PF, Role of 2-[18F]-fluorodeoxyglucose(FDG) positron emission tomography(PET) in the early assessment of response to chemotherapy in metastatic breast cancer patients, Clin Breast Cancer, pp. 156-61, 2000.

6. Jansson T, Westlin JE, Ahlstrom H, Lilja A, Langstrom B and Bergh J , Positron emission eomography studies in patients with locally advanced and/or metastatic breast cancer: a method for early therapy evaluation?, Journal of Clinical Oncology, 13, pp. 1470-1477, 1995.

7. Mitchell T, *Machine Learning,* McGraw Hill, 1997.

8. McDermott GM, Welch A, Staff RT, Gilbert FJ, Schweiger L, Semple SIK, Smith TAD, Hutcheon AW, Miller ID, Smith IC and Heys SD, Monitoring Primary Breast Cancer throughout Chemotherapy using FDG-PET, Breast Cancer Research and Treatment, 10.1007/s10549-006-9316-7, 2006.

9. http://www.cs.waikato.ac.nz/ml/weka/

10. Witten IH, Frank E, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2005.