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

E Gyftodimos¹, L Moss¹, D Sleeman¹, A Welch²
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, neoadjuvant chemotherapy, machine learning.

Abstract
We discuss the use of machine learning algorithms to predict which breast cancer patients are likely to respond to (neoadjuvant) 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 cancer¹. Up to 25% of these patients have large (>3cm) tumours². For these patients, neoadjuvant chemotherapy is sometimes offered in an attempt to reduce the size of the tumour before surgery³,⁴. It is estimated that up to 25% of these patients do not respond to this chemotherapy², 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 neoadjuvant 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 observations⁷. 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 studies²,⁸. McDermott et al.⁸ investigated optimum times for imaging when using PET to predict response to neoadjuvant chemotherapy. By measuring the mean standard uptake value at the midpoint of neoadjuvant chemotherapy, they identified 77% of the low responding patients whilst identifying 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 naïve 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 naïve Bayes algorithm had ROC = 87.3% and SPC = 60.0%. Finally, in the setting MV2, naïve 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.