ISSN 1407-7493 COMPUTER SCIENCE DATORZINĀTNE 2008-7493

INFORMATION TECHNOLOGY AND MANAGEMENT SCIENCE
INFORMĀCIJAS TEHNOLOĢIJA UN VADĪBAS ZINĀTNE

APPLYING ANN ENSEMBLES TO MELANOMA-RELATED DIAGNOSTICS ON CANCER PATIENTS

Vilens Jumutcs, B.sc.ing., Mg.sc.ing student, Institute of Information Technology, Riga Technical University, Latvian Biomedical Research and Study Centre, Riga, Latvia, e-mail:jumutc@gmail.com

Pawel Zayakin, M.sc., PhD student, University of Latvia, Latvian Biomedical Research and Study Centre, Riga, Latvia, e-mail: pawel@biomed.lu.lv

Aija Linē, Dr.biol, Latvian Biomedical Research and Study Centre, Riga, Latvia, e-mail: aija@biomed.lu.lv

Keywords: artificial neural network, ensembles, antigens, K-Means, meta-model

1. Introduction

Artificial Neural Network (ANN) ensembles are being described in this paper with the aim of identifying new proper and more precise ways of melanoma-related diagnostics on cancer patients. ANN ensembles were chosen as the most convenient and closely reflecting provided data AI mechanism that sharpened preliminary imprecise ANNs on insufficient in number but effectively informative blood samples. In the provided data pools were included autoimmune antibodies' profiles (selected by immunoscreening of phage's library derived from melanoma-cancer tissue and acquired by microarray analysis) of patients of different melanoma stages as well as profiles of healthy control group. As the main intent of research effort was considered classification/prediction accuracy improvement of implemented model in comparison to simple multilayered ANN and regression classifier trees (for the sake of unbiased result was used Logistic Modelled Trees – LMT [1]).

As a result of research effort, firstly a new biologically robust methodology on selecting antigens with high expression level in cancer patients was introduced. Secondly a robust 2-layer meta-model was implemented consisting of K-Means clustering classifier on the first layer and ANN ensemble on the second. The evaluated model results were compared with simple Artificial Neural Network and LMT classifier outcomes on the same verification set and showed significant performance gain in comparison with simple ANN and increased stability in comparison with LMT classifier.

The proposed meta-model doesn't solve all kinds of problems related to insufficient in number training samples and hardly classified 20-25% of melanoma patients due to possibly missed on microarray chip specific for these patients antigens derived from phage-display library. But this model finally helps to minimize overfitting of initial ANN model, reduce mentioned noise across cancer patients and sustain more stable performance scores needed to effectively and more precisely apply bioinformatics tools for earlier cancer diagnosis. An example of DNA microarray can be seen in Figure 1.

2. Theory

2.1. Microarrays

The finding that patients with cancer produce autoantibodies against antigens in their tumours suggests that such autoantibodies could have diagnostic and prognostic value. New biomarkers, such as autoantibody signatures, may improve the early detection of cancer and enable a more clear understanding of processes in tumour tissues. With a phage-display library derived from melanoma-cancer tissue, we developed and used phage protein microarrays to analyze serum samples from 100 patients with melanoma cancer and 100 samples from healthy control group.

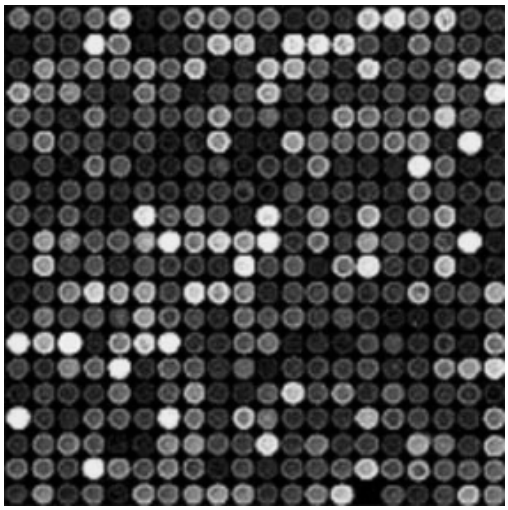

Figure 1. DNA microarray

We used a technique that combines phage-display technology with protein microarrays to identify and characterize new autoantibody-binding peptides derived from melanoma-cancer tissue. A similar approach has been used to identify selected antigens for the diagnosis of prostate cancer. T7 phage display cDNA expression library was constructed from melanoma-cancer tissue mRNA. Phages were used for the infection of *E. coli* and serum-reactive clones were detected by immunoscreening as in conventional SEREX technology. A set of 733 different serum-reactive phage clones was assembled, grown to high titres in *E. coli* cells and the lysates were used for the production of antigen microarrays. The phages were spotted in duplicates onto nitrocellulose – coated FAST slides (Whatman) using QArray ^{mini} compact arrayer with 150-micron solid pins (Genetix). Each array was tested for the reactivity with patients' sera that has been preabsorbed with *E. coli* and T7 phage lysate, and a mouse monoclonal antibody against T7 tail fiber protein. Serum reactivity was detected using Cy3-conjugated anti-human IgG secondary antibody and anti-T7 antibody reactivity was detected with Cy5-conjugated anti-mouse secondary antibody (Jackson ImmunoResearch Laboratories).

The arrays were scanned with aQuire two-laser confocal scanner (Genetix) and the results were recorded as TIFF files. Median of Cy3/ Cy5 ratios was calculated for each spot and averaged between replicates. In order to normalize the signal intensities among different arrays, an average Cy3/Cy5 ratio for all wild-type phage spots in an array was set to be equal to 1 and the values for the rest of the spots were re-calculated accordingly.

2.2. ANN ensembles

Artificial Neural Networks' ensembles are clearly ANN based models that are frequently mentioned in computer science and evolutionary programming (EP) as robust and efficient way of eliminating noise and other ANN topology related drawbacks that prevent ANNs of being one of the most wanted AI techniques nowadays. For instance, evolutionary artificial neural networks (EANNs) that deal with ANN ensembles refer to another special class of artificial neural networks in which evolution is another fundamental form of adaptation in addition to learning [1].

Basically the idea of assembling together differently trained and topologically shaped ANNs is not new. The main concept is to split initial ANN model into smaller compartments by segregating preliminary training data into smaller input arrays. This can be done by clustering initial input attributes or eliminating some redundant and noisy ones or developing smaller ANNs only for a subset of initial training data to emphasize some hardly classified samples and prevent overfitting. As a result of differentiation, we can get several ANNs instead of one solid model (see Figure 2).

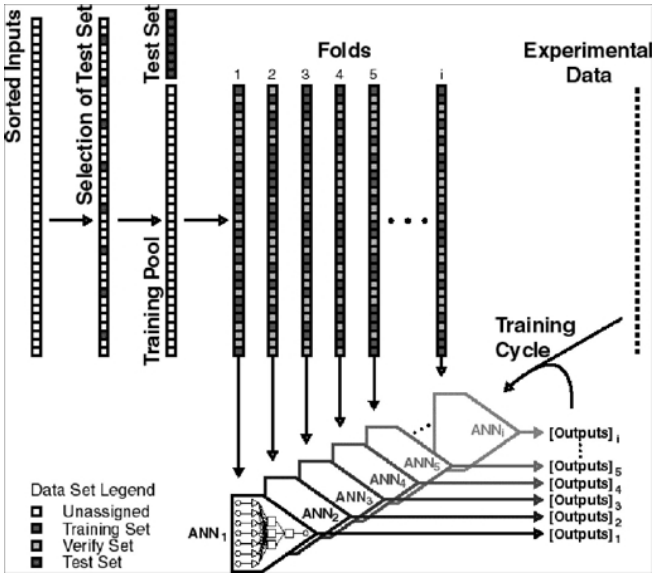

Figure 2. ANN ensemble

Explicit gain from this effort is initial attribute/sample demarcation by any a priori known function or congruence factor and noise reduction. Thus it is essential to find appropriate function or factor of interest that could help us to segregate initial input attributes and samples.

3. Methods and models

The applied theoretical approach assumes that modifying initial solid ANN model will benefit in several very practical ways. First of all we shall mention feasible improvement of performance score on hardly recognizable patterns that spread across test samples in maximum of few exemplars. Thus the exceptional utility of ANN ensemble is the possibility of building such model that incorporates knowledge about specific antibody patterns as well as associated with those patterns melanoma patients by performing learning procedures only on a subset of initial data (input arrays and training samples).

3.1. Antigen reduction

Antigen reduction is one of the most important and frequently mentioned aims in cancer diagnosis because of the extreme expensiveness of high through output microarray methods and impossibility to carry out such high-end laboratories in urban vicinity for the sake of prophylactics. Thus we have normalized all expression data against means of 20 wild type phage antigens, established serum reactivity threshold on the basis of control healthy group (HD) and picked out only those input attributes (antigens) that contained expression values exceeded maximum of 4 positive standard deviations from the mean value (SD was calculated on the basis of HD samples only) for melanoma patients and didn't contain those for HD samples.

3.2. Meta-model

The introduced meta-model is a simple conjunction of 2 commonly used classifier models: K-Means clustering and ANN. The first layer composed of K-Means clusterer introduces N predefined clusters build from training data set over previously selected for classification task robust antigens (see Antigen reduction). Secondly we inspect newly created clusters and tag them as **true MEL** or **true HD** if they contain samples only from HD or melanoma related subsets. Otherwise in the second layer we create separate ANN model trained only on relevant to this cluster samples using all robust antigens.

The correctness of ANN outcome was interpreted by means of RMSE (root mean square error) for every predefined ANN output and real value. We have used binary output range – {0, 1} to train each ANN under supervised learning scheme and assigned "classified" and "misclassified" tags thus leveling up and down sensitivity and specificity scores only if RMSE <= 0.5 for "classified" samples and > 0.5 for "misclassified" ones. Overall model output is identified by the initial clusterer outcome and if needed one of the ANN ensemble compartments. If verified sample belongs to **true MEL** or **true HD** clusters then its class is identified in the first layer by selecting appropriate identity value from cluster (1 for "mel", 0 for "hd"); otherwise relevant to this cluster ANN is used to classify verification sample and desired class is calculated by the least distance to one of the binary outcomes predefined by the training stage.

The overall mathematical model can be written as follows:

$$k = class(y) = \begin{cases} f_T(i), if \ \arg\min_{i \in K} \{d(x,y) : x \in A_i\} \in TC \\ \arg\min_{j \in \{0,1\}} \left| f_{ANN}(i) - j \right|, if \ \arg\min_{i \in K} \{d(x,y) : x \in A_i\} \in NC \end{cases} \quad (1)$$

Where: **y** – verification sample, **i** – cluster index, **K** – enumeration of clusters, $\mathbf{A_i}$ – i-th K-Means cluster, **TC** – enumeration of true MEL and HD clusters, **NC** – enumeration of mixed type clusters, $\mathbf{f_t(i)}$ – identity value function for true MEL and HD clusters, $\mathbf{f_{ANN}(i)}$ – output of trained ANN for the i-th K-Means cluster, **j** – identity value;

Below a structural decomposition of this meta-model is shown (See Figure 3).

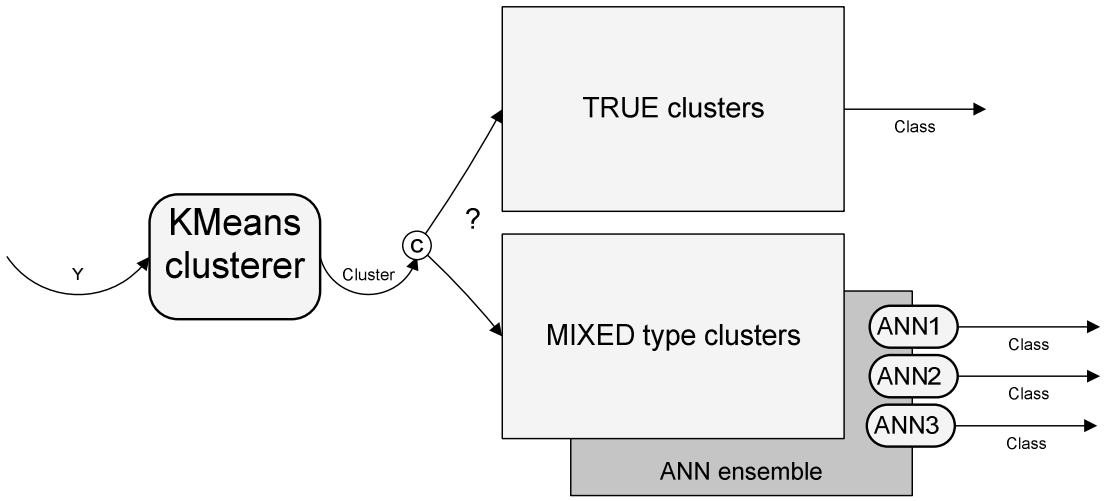

Figure 3. Structural decomposition of meta-model

4. Results

To make some crucial verifications of the proposed model we have compared its overall performance score with simple ANN model and one of regression tree classifiers – Logistic Model Trees [1]. Key measures of classification success were meant sensitivity (2), specificity (3), MCC – Mathews Correlation Coefficient (4) and classification failure score (5).

$$Sensitivit\ \ y = F_{sens} = \frac{TP}{TP + FN} \tag{2}$$

$$Specificit\ \ y = F_{spec} = \frac{TN}{TN + FP} \tag{3}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{4}$$

$$classifcat\ ion\ \_failure\ \ = \frac{\sum_{i}^{K} \begin{cases} 1, if\ \ldots F_{sens}^{i} \le \lambda \vee F_{spec}^{i} \le \lambda \\ 0, otherwise \end{cases}}{K} [\%] \tag{5}$$

Where **TP** – true positives, **TN** – true negative, **FP** – false positive, **FN** – false negative, **K** - number of iterations held, **λ** – classification failure bound (that means that all sensitivity or specificity scores below that bound will be considered failed and inadequate for classification under current verification set);

During experimental stage, different meta-model approbations were made and tuning efforts were carefully compared with LMT and simple ANN results.

4.1. Experiment 1

First of all a model with N = 50 (clusters on the first layer) was implemented and selected robust antigens were verified on 50 iterations with random verification sets. In order to get rid of biased performance scores exactly 10 random HD (control) samples and 10 random cancer patients were included in verification set. This procedure repeated each

iteration in order to guarantee probabilistic nature of forthcoming and unseen samples. Experimental results are given in Table 1.

Table 1

Performance scores of Experiment 1

Model	Averaged sensitivity	Averaged specificity	Averaged MCC	Classification failure, $\lambda = 0.5$ [%]
Meta-model	0.81999	0.67799	0.5700928	6
Simple ANN	0.47	0.74199	N/A	76
LMT	0.83199	0.66999	0.575242	20

4.2. Experiment 2

Secondly we examined the model with N= 100 and verified selected robust antigens on 30 iterations with random verification sets. The results are shown in Table 2.

Table 2

Performance scores of Experiment 2

Model	Averaged sensitivity	Averaged specificity	Averaged MCC	Classification failure, $\lambda = 0.5$ [%]
Meta-model	0.782499	0.735	0.585	6.67
Simple ANN	0.4925	0.6875	N/A	83.33
LMT	0.7675	0.655	0.51446	23.33

4.3. Experiment 3

To examine one of the key points of our paper – "antigen reduction" under melanoma related diagnosis we decided to test the proposed meta-model on all antigen data set. The model was implemented under N = 100, iteration count = 20. It is worth mentioning that simple ANN model and LMT classifier as well use only robust antigens as inputs in previous experiments and for the current one were adjusted as well to correspond in data sets to the proposed meta-model. The results can be seen in Table 3.

Table 3

Performance scores of Experiment 3

Model	Averaged sensitivity	Averaged specificity	Averaged MCC	Classification failure, $\lambda = 0.5$ [%]
Meta-model	0.88	0.715	0.64532	0
Simple ANN	0.43995	0.72	0.3354	85
LMT	0.87	0.67499	0.605	15

4.4. Extra scores

Finally we have calculated average misclassifications for every meta-model layer. This error is calculated for all samples included in verification set. Here are some values gained for Experiment 2: first layer average error – 11%, second layer average error – 14.25%. Here are some values gained for Experiment 3: first layer average error – 8.5%, second layer average error – 11.75%.

5. Conclusions

As can be seen from the experiments, the proposed meta-model works considerably better than simple ANN model and slightly better for specificity and sensitivity (recall) in comparison with LMT classifier. Although main performance scores for meta-model and LMT classifier differs only in 5-10% we can confidently claim that classification failure score (that was set to 0.5 due to feasible probability of selecting 10 control/melanoma samples from the verification set of 20 samples by the chance is exactly 50%) is much better for meta-model in all experiments.

This score (classification failure) is extremely important for correct, effective and more precise usage of bioinformatics tools for earlier cancer diagnosis. Finally, it crucially influences the way of further bioinformatics tool development for clinical and prophylactics purposes because the absence of permanent failures indicates safety and durability of proposed methods and models under whatever data and circumstances.

The other key intent of the paper is antigen reduction and effectiveness comparison to the “full-charged” meta-model and other classifiers. As can be seen, performance degradation is not significant and analyzing results of the second experiment we can conclude that performance scores tend to be more likely stabilized over 70%

Finally, let us discuss error distribution over 2 layers of meta-model. As can be seen from the gained statistics, the first layer is more precise and tending to incorporate more correct samples distribution for K-Means clusters. This can be explained by the tendency of simply classified samples to be grouped in one cluster.

As the forthcoming improvements of the proposed meta-model we are looking forward to some SVM (Support Vector Machines) algorithms on the first layer as more reliable and biologically relevant clustering/classification approach and expert (microbiologists’) ranking/notation of selected robust antigens for the second layer outcome more precise interpretation.

References

1. Niels Landwehr, Mark Hall, Eibe Frank (2005). Logistic Model Trees. Machine Learning. 95(1-2):161-205.
2. Xin Yao, Yong Liu: Ensemble structure of evolutionary artificial neural networks (1996). In Proc. of the 1996 IEEE Int'l Conf. on Evolutionary Computation (ICEC'96). URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.56.306. – Visit date July 2008.
3. Dan Pelleg, Andrew W. Moore: X-means: Extending K-means with Efficient Estimation of the Number of Clusters. In: Seventeenth International Conference on Machine Learning, 727-734, 2000.

Jumutcs Vilens, Zajakins Pāvels, Linē Aija. Mākslīgo neironu tīklu ansambļu izmantošana melanomas vēža diagnostikā

Dotajā darbā tiek apskatīti mākslīgo neironu tīklu ansambļi ar mērķi uzlabot melanomas vēža diagnostiku. Neironu tīklu ansambļi tika paņemti kā viens no efektīvākiem instrumentiem, kas labi atspoguļo diagnostikas uzdevumu un var uzlabot sākuma MNT klasifikācijas kļūdu uz nepietiekamiem daudzumā, bet tomēr informatīviem antigēnu paterniem. Izmantotajā datu kopā bija iekļauti autoimūnu antivielu profili (kas tika iegūti izmantojot fagu bibliotēkas immunoskriningu no melanomas vēža audiem un analizēti ar mikročipu metodi) no melanomas pacientiem no dažādām vēža stadijām, ka arī referatīvie paraugi no kontroles grupas. Kā galvenais pētījuma iemesls tika formulēta klasifikācijas/diagnostikas uzdevuma precizitātes palielināšana attiecībā pret parasta MNT modeļa un klasifikācijas regresijas kokiem. Kā pētījuma rezultāts, tika iegūts jaunais meta-modelis, kas sastāv no diviem slāņiem un kuru rezultāts tika apkopots gan pirmajā, gan otrajā slānī. Piedāvātais meta-modelis parādīja salīdzinoši labākus klasifikācijas rezultātus un lielāku stabilitāti uz atlasītām verifikācijas kopām

Jumutcs Vilens, Zayakin Pawel, Linē Aija. Applying ANN ensembles to melanoma-related diagnostics on cancer patients

Artificial Neural Network (ANN) ensembles are described in this paper with the aim of identifying new proper and more precise ways of melanoma-related diagnostics on cancer patients. ANN ensembles were chosen as the most convenient and closely reflecting provided data AI mechanism that sharpened preliminary imprecise ANNs on insufficient in number but effectively informative blood samples. In the provided data pools were included autoimmune antibodies' profiles (selected by immunoscreening of phage's library derived from melanoma-cancer tissue and acquired by microarray analysis) of patients of different melanoma stages as well as profiles of healthy control group. As the main intent of research effort was considered classification/prediction accuracy improvement of implemented model in comparison to simple multilayered ANN and regression classifier trees (for the sake of unbiased result was used Logistic Modeled Trees). As a result of the research, a novel meta-model with two classification layers and separated output interpretation, was implemented. The model reflected comparatively better performance scores and higher stability factors on selected verification sets.

Юмутц Вилен, Заякин Павел, Лине Айя. Использование ансамблей ИНС для диагностики меланомы

В данной работе рассмотрены ансамбли нейронных сетей в контексте диагностики и распознавания больных меланомой. Ансамбли нейронных сетей были выбраны как метод, наиболее точно отображающий предметную область, позволяющий настраивать изначально неточную нейронную сеть на недостаточных в количестве, но достаточно информативных экземплярах кровяной сыворотки. В работе использованы данные профилей автоиммунных антител больных меланомой на различных стадиях, а также - образцы контрольной группы здоровых людей. Главной целью данного исследования является улучшение точности диагностики пациентов больных меланомой в сравнении с простой ИНС и регрессионными классификационными деревьями. Как результат проведённых исследований, необходимо упомянуть полученную новую классификационную мета-модель, состоящую из двух классификационных слоёв – K-Means кластеризатора и ансамбля ИНС с разделённой интерпретацией выхода. Полученная модель показала сравнительно лучшие значения по классификации и фактору стабильности на одних и тех же проверочных данных, нежели простая ИНС и регрессионные классификационные деревья.