A New Approach to Prediction of Radiotherapy of Bladder Cancer Cells in Small Dataset Analysis

Gy-Yi Chao[1], Tung-I Tsai [2], Te-Jung Lu[3], Hung-Chang Hsu[4], Bo-Ying Bao[5], Wan-Yu Wu[3], Miao-Ting Lin[3], Te-Ling Lu[5,*]

[1] Department of Nursing, College of Medicine and Life Science, Chung-Hwa University of Medical Technology, Tainan, Taiwan
[2] Department of Information Management, Shu-Te University, Kaohsiung County, Taiwan
[3] Department of Medical Technology, College of Medicine and Life Science, Chung-Hwa University of Medical Technology, Tainan, Taiwan
[4] Department of Biomedical Engineering, Chimei Medical Center, Tainan, Taiwan
[5] School of Pharmacy, College of Pharmacy, China Medical University, Taichung, Taiwan

Corresponding author: Tel: +886 4 22053366 ext. 5703

E-mail address: lutl@mail.cmu.edu.tw

Keywords: *small sample size, artificial neural network, machine learning, molecular prediction, bladder cancer*

Abstract

Bladder cancer is a common urologic cancer. Radiotherapy plays an increasingly important role in treatment bladder cancer due to radiotherapy preserves normal bladder function. However, the five-year survival rate after radiotherapy for bladder cancer patients is 30-50%. Some biological proteins influence the outcome of radiotherapy. One or two specific proteins may not be sufficient to predict the effect of radiotherapy, analyzing multiple oncoproteins and tumor suppressor proteins may help the prediction. At present, no effective technique has been used to predict the outcome of radiotherapy by multiple protein expression file from a very limited number of patients. The bootstrap technique provides a new approach to improve the accuracy of prediction the outcome of radiotherapy in small dataset analysis. In this study, thirteen proteins in each cell line from individual patient were measured and then cell viability was determined after cells irradiated with 5, 10, 20, or 30 Gy of cobalt-60. The modeling results showed that when the number of training data increased, the learning accuracy of the prediction the outcome of radiotherapy was enhanced stably, from 55% to 85%. Using this technique to analyze the outcome of radiotherapy related to protein expression profile of individual cell line provides an example to help patients choosing radiotherapy for treatment.

1. Introduction

The improvement in surgical techniques for radical cystectomy is effective in preventing the development of new bladder tumors (Milosevic et al., 2007). Radiotherapy, using high-energy rays to destroy tumor cells, is the nonsurgical treatment in the attempt to save the bladder. A recent study shows that the outcome for radical cystectomy and radiotherapy treatment in invasive bladder cancer is similar. The five-year survival rates are 56.8% and 53.4% for radiotherapy patients and for surgery-treated patients, respectively. There is also no difference in radiotherapy patients and surgery-treated patients with 34% and 37.5% recurrence, respectively (Kotwal et al., 2008). Radiobiology has reinforced the important role in the treatment of bladder cancer. However, the cure rates with 5-year survival in the range of only 30–50% and some biologic factors that influence bladder cancer progression and radiation response should be evaluated.

Several proteins including activating oncoproteins and suppressing tumor suppressor proteins together contribute to therapy resistance and the low survival rates of bladder cancers. Oncoproteins involved in signal transduction cause tumor formation by uncontrolled cell cycle and cell growth. Cyclin D1 with cdc2 kinase positively regulate cell cycle progression. Cyclin D1 expression is significantly higher in low-stage, well differentiated bladder tumors (Tut et al., 2001). Overexpression of

EGFR, c-erbB3, c-erbB4, neu, retinoblastoma protein or bcl-2 shows poor response to radiotherapy (Colquhoun et al., 2006; Nix et al., 2005; Pollack et al., 1997). In the other hand, tumor suppressive proteins play an important role in inhibiting tumor formation. p16, a tumor suppressor gene, blocks the Cdc2 and cyclin-D complex (Kubo et al., 1999). Deletion of p16 gene has been reported in bladder tumors and bladder cell lines (Williamson et al., 1995). Lack of Bax expression was related to reduced patient's survival (Gonzalez-Campora et al., 2007). These results show that clusters of protein markers together better reflect the complexity of the underlying biological network. It is expected that establishing personalized protein expression profile will help to predict the outcome of radiotherapy of patients with bladder cancer.

To develop personalized radiotherapy of bladder cancer, a profile of thirteen protein expression related to resistance to radiotherapy was established. The expression intensity of thirteen proteins in each cell line was measured. Oncoproteins of MDR, Topo II, EGFR, Neu, c-ErbB-3, c-erbB-4, cyclin A, cyclin D1, Cdc2 and Bcl2 were scored with a positive value according to their expression intensity. Tumor suppressive proteins of Rb, P16 and Bax were scored with a negative value according to their expression intensity. The nine immortal bladder cancer cell lines were treated with 5, 10, 20, or 30 Gy of cobalt-60 (Co-60) and then determined the cell viability.

The cell resistance to radiotherapy will be influenced by combination of oncoprotein and tumor suppressive protein expression. A useful model of bootstrap technique makes these data to define the protein expression profile related to the outcome of radiotherapy.

The research is organized as follows: Section 2 briefly states the references of small sample set research; and Section 3 offers a numerical example and the detail process of the proposed method. Computational results and Conclusions are provided in Section 4.

2. References of small sample set research

To overcome a very small dataset, computational learning theory looks for some answers to machine learning problems concerning sample size, such as: how many training examples are needed to lead to a successful learning, how many reasonable computations are needed for successful learning, and what is the estimated misclassifying rate in learning. A probably approximately correct (PAC) concept has been developed to identify classes of hypotheses that can/cannot be learned from a polynomial number of training examples with a reasonable amount of computation (Anthony and Biggs, 1997). Furthermore, a sample size as small if the ratio of the number of training samples to the Vapnik-Chervonenkis dimensions (VC dimensions) of a learning machine function is less than 20 has been defined (Vladimir, 2000).

However, these theories focus on general machine learning with a large number of training samples, which cannot be applied to practical cases with the small data set learning model.

Adding some artificial data to the system is one effective approach to increase learning accuracy. In Virtual Data Generation, mostly used in Pattern Recognition, used prior knowledge obtained from the given small training set to create virtual examples for improving recognition ability. In their method, from a given 3-D view of an object new views may be generated from any other angle through mathematical transformations. The new views generated are called virtual samples. With these virtual samples, a learning machine can verify an instance more precisely. It has been proved that the process of creating virtual samples is mathematically equivalent to incorporating prior knowledge (Niyogi et al., 1998).

Few closely related studies in the field of manufacturing are found in the literature: Li and Lin proposed the Functional Virtual Population (FVP) approach involving the use of a neural network in dynamic manufacturing environments that learn scheduling knowledge (Li and Lin, 2006). The FVP approach was the first method proposed for small data set learning for scheduling problems, and it was developed to expand the domain of the system attributes and generate a number of virtual samples for constructing the so-called early scheduling knowledge. However,

based on a trial-and-error procedure, the FVP approach requires many steps to complete the process.

In 2006, a unique data fuzzification technique, named mega-fuzzification, combined with a data trend estimation procedure to systematically expand the small data set obtained in the early stages of manufacturing (Li et al., 2006). In 1993, the Adaptive-Network-based Fuzzy Inference System (ANFIS) was applied to neuro-fuzzy learning (Jang, 1993). Although according to the results of ANFIS achieved by mega-fuzzification improved–the learning accuracy, the ANFIS is not commonly accepted in real world industries, and insensitive in small data set learning (Li et al., 2006).

Huang and Moraga combined the principle of information diffusion with a traditional neural network, called a diffusion-neural-network (DNN), for functional learning (Huang and Moraga, 2004). According to the results of their numerical experiments, the DNN improved the accuracy of the Backpropagation Neural Network (BPN). The information diffusion approach partially fills the information gaps caused by data incompleteness via applying fuzzy theories to derive new samples, but the research does not provide clear indications for determining the diffusion functions and diffusion coefficients. Besides, the symmetric diffusion technique sometimes over simplifies a generation of new samples, which could cause

over-estimation of the domain range. Either under-estimating or over-estimating the ranges would lead to reduced accuracy.

Therefore, in order to fully fill the information gaps, a technique called mega diffusion was substituted a sample set for diffusing samples one for one (Li et al., 2007). Furthermore, a data trend estimation concept is combined with the mega diffusion technique to avoid over-estimating. This technique, which combines mega diffusion and data trend estimation, was called mega-trend-diffusion (Li et al., 2007). Following mega-trend-diffusion, the production of virtual samples was proposed to improve the FMS scheduling accuracy. Unfortunately, in their research, the DNN is adopted to extract knowledge. The DNN has twice as many input factors as original ones, and this situation means the network has much more complex calculations than the ANN.

Ivănescu et al. proposed a procedure to solve the limited data problem in batch process industries. They assumed the job arrival moments obey a Poisson arrival process and utilized bootstrap procedure to generate another 250 the bootstrap jobs. According to their results, the procedure they proposed has improved the regression modeling performance (Ivănescu et al., 2006).

Tsai and Li utilized the bootstrap procedure once for each input factor and a real data set acquired from a Taiwanese manufacturer of multi-layer ceramic capacitors

(MLCC) was used to illustrate the effectiveness of the proposed procedure (Tsai and Li, 2008). Based on their research, the prediction accuracy was increased. This research adopts the procedure proposed by Tsai and Li and attends to improve the prediction accuracies of the effectiveness of Co-60 to bladder cancer cell lines.

3. The Detailed Processes

The bootstrap implies re-sampling a given data set with replacement and is used for measuring the accuracy of statistical estimates (Efron and Tibshirani, 1993). In this paper, we will attempt to use the bootstrap method to generate virtual samples and solve the learning problem using the data sheet (36 data in total) provided by a medical research center in Taiwan. At the beginning of this case study, we simulate a situation that when only 5 data are available. Thus, we use only 5 data for training the neural network and then use the rest of the data for validation. Following this, we will try other data scales (5 to 30, in increments of 5) to the training set each time.

To explain the procedure in detail, a total of 36 data are obtained from the medical research center, shown in Table 1. Among them, this research randomly selects a specific number (5 to 30, in increments of 5) of data as the training set from Table 1, and used the rest as the testing data for evaluating the average learning error rates of the ANN. Thus the experimental scales in this study are 5, 10, 15, 20, 25 and 30 training data. The following is an example of the process with 10 training data, and

the procedure is depicted in steps:

[Insert Table 1 here]

Step 1. Select 10 data randomly from Table 1 as the training data for training the ANN. The selected data set is listed in Table 2.

[Insert Table 2 here]

Step 2. Execute the bootstrap procedure for the data in Table 2 once for each input and output factors to acquire virtual samples. Repeating the procedure 100 times, we can acquire 100 virtual samples (the determination of the optimal number of virtual samples needs further study). The results of the values of each factor are given in Table 3.

[Insert Table 3 here]

Step 3. Apply the data in Table 2 and 3 together as the training data to train an ANN.

The resistance to radiotherapy value in Table 2 and 3 is the value assigned to the output node of the ANN; others are the inputs of the ANN. Although other researchers, such as Amirakian and Nishimura and Wang et al., provided algorithms suggesting ways to determine the number of hidden nodes and hidden layers, the number of hidden nodes and hidden layers are believed to differ case by case (Amirakian and Nishimura, 1994; Wang et al., 1994). In this study, the optimal structure of the ANN is

determined by the Evolutionary Optimizer tool of Pythia software. Pythia is a program for the development and design of Neural Networks and features Backpropagation Networks.

This tool executes genetic algorithm (GA) with crossover rate equals 0.2 and mutation rate equals 0.04. Initially, the original generation containing 50 randomly created networks and each network within this generation will be trained shortly and its fitness determined according to the parameters in “Goals to achieve”. The 10 fittest networks of the old generation are leaved as the parents of the next generation. It works persistently until it finds a network with a fitness of 100 or the 1000th generation.

The “Goals to achieve” is a setting to specify what the network should be optimized for. There are three goals possible:

1. Optimize for medium deviation (Ø deviation<)
2. Optimize for max. deviation within the pattern set (*deviation<)
3. Optimize for size (# neurons<=)

In this research, we use the default setting of Pythia software. That is, the medium deviation should be below 0.001, the max. deviation should be below 0.1 and the network size should be below or equal 100. Both checked goals will contribute equally to the overall fitness of an evolutionary created network.

After determining the topology of network, trains the network until 1000 repetitions or 300 seconds have passed with learn rate equals 0.5 as the default settings.

Step 4. Use the rest of the data in Table 1 as the testing data for the ANN to calculate the average error rates. The average error rate is defined as:

$$\frac{\sum_{i=1}^{n} \frac{\left|\text{resistance to radiotherapy}_i - \textit{output of network}_i\right|}{\text{resistance to radiotherapy}_i}}{n}$$

where n is the number of the samples in validation set and $i = 1, 2, \ldots, n$ and in this example n equals 26.

The resistance to radiotherapy and the *output of network* values are shown in Table 4, and the average error rate is 0.239478

[Insert Table 4 here]

Step 5. Repeat Steps 1 to 4 ten times and calculate the average error rate.

Step 6. Repeat Steps 1 to 5 with different scales of training data sets.

4. Computational results and conclusions

The computational results are compared with the results obtained using the primitive data, as represented in Table 5 and 6 and Figure 1 and 2.

[Insert Table 5 and 6 here]

[Insert Figure 1 and 2 here]

It is obviously although the average error rate of ANN using only primitive data

decreases as the number of training samples is increased, the standard deviation does not converge. That is, the ANN using only primitive data cannot build up a robust forecast neural network by using such rare pilot runs data. However, the proposed procedure of this study reveals lower and stable learning errors. Hence, when the data collected is insufficient, the procedure of this study works to make the forecast system better and more stable. The above results are encouraging, and as shown in Figure 1 and 2, when the training data set increases, the average error rate and standard deviation monotonically decrease.

From the p-value in Table 7, it is suggested that there are significant differences between ANN and the proposed procedure.

[Insert Table 7 here]

Analyze the cell sensitivity to cobalt-60 (Co-60) and protein expression profile of each cell line can be a useful forecast model to predict the radiotherapy outcome of bladder cancer. Several factors including tumor stage and age influence the radiotherapy outcome of bladder cancer. This model may offer the potential to improve cure rate and reduce adverse effects based on the protein expression profile of individual patient. The studies of small samples are few at present, and there is a lot of potential to seek better theories to obtain a higher rate of accuracy. This research may be applied in the determination of the effectiveness of radiotherapy in the

treatment of bladder cancer.

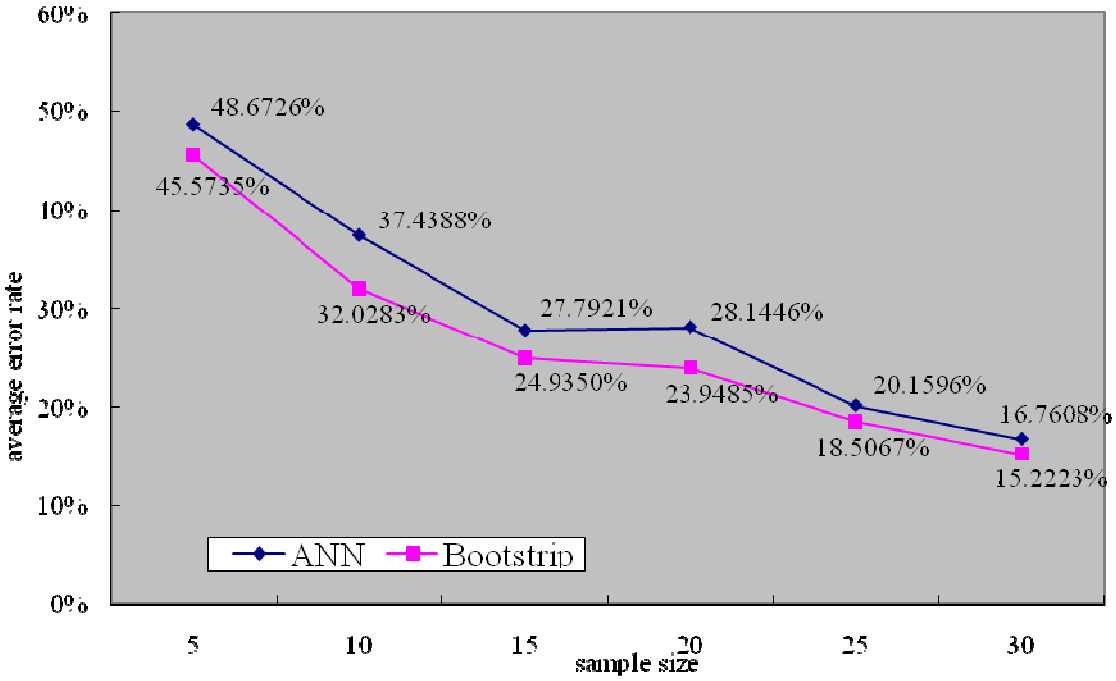

Figure 1. The computational results (average error rate)

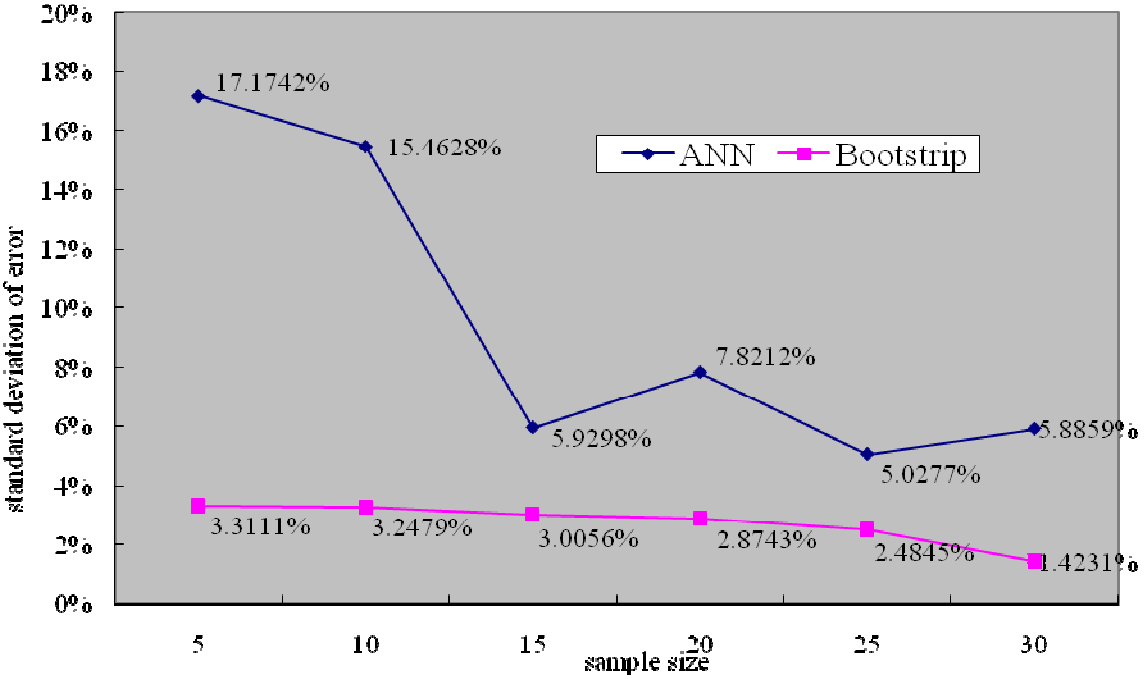

Figure 2. The computational results (standard deviation)

Table 1: The 36 data obtained from medical research center

No.	Cell line	MDR	Topo II	Rb	EGFR	Neu	c-ErbB-3	c-ErbB-4	Cyclin A	Cyclin D1	P16	Cdc 2	Bcl-2	Bax	Co-60 (Gy)	Resistance to radiotherapy
1	HT 1376	1	1	-0.5	5	0.5	6.5	6.5	0.7	0.1	-0.5	2.8	1.8	-0.3	5	97
2	HT 1376	1	1	-0.5	5	0.5	6.5	6.5	0.7	0.1	-0.5	2.8	1.8	-0.3	10	90
3	HT 1376	1	1	-0.5	5	0.5	6.5	6.5	0.7	0.1	-0.5	2.8	1.8	-0.3	20	84
4	HT 1376	1	1	-0.5	5	0.5	6.5	6.5	0.7	0.1	-0.5	2.8	1.8	-0.3	30	82
5	HT 1197	1	2	-1.3	3	0.3	3	2.5	0.1	1.3	-1.3	2.5	0.1	-0.8	5	92
6	HT 1197	1	2	-1.3	3	0.3	3	2.5	0.1	1.3	-1.3	2.5	0.1	-0.8	10	78
7	HT 1197	1	2	-1.3	3	0.3	3	2.5	0.1	1.3	-1.3	2.5	0.1	-0.8	20	72
8	HT 1197	1	2	-1.3	3	0.3	3	2.5	0.1	1.3	-1.3	2.5	0.1	-0.8	30	73
9	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	5	94
10	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	10	70
11	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	20	50
12	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	30	41
13	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	5	84
14	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	10	70
15	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	20	48
16	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	30	40
17	Sca-BER	1.2	0.8	-1.5	10	0.5	0.1	5	0.1	2.7	-0.1	2.5	0.1	-1.3	5	82
18	Sca-BER	1.2	0.8	-1.5	10	0.5	0.1	5	0.1	2.7	-0.1	2.5	0.1	-1.3	10	57
19	Sca-BER	1.2	0.8	-1.5	10	0.5	0.1	5	0.1	2.7	-0.1	2.5	0.1	-1.3	20	39
20	Sca-BER	1.2	0.8	-1.5	10	0.5	0.1	5	0.1	2.7	-0.1	2.5	0.1	-1.3	30	36
21	T24	0.3	1.3	-2.7	2.2	1	0.1	2.5	0.5	2	-0.1	2.5	1.3	-0.2	5	90
22	T24	0.3	1.3	-2.7	2.2	1	0.1	2.5	0.5	2	-0.1	2.5	1.3	-0.2	10	58
23	T24	0.3	1.3	-2.7	2.2	1	0.1	2.5	0.5	2	-0.1	2.5	1.3	-0.2	20	39
24	T24	0.3	1.3	-2.7	2.2	1	0.1	2.5	0.5	2	-0.1	2.5	1.3	-0.2	30	34
25	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	5	83
26	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	10	50
27	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	20	32
28	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	30	28
29	TSGH-8301	2	0.3	-0.3	9	3.5	4.5	6.5	0.1	0.1	-0.1	2.5	1.3	-1.5	5	60
30	TSGH-8301	2	0.3	-0.3	9	3.5	4.5	6.5	0.1	0.1	-0.1	2.5	1.3	-1.5	10	30
31	TSGH-8301	2	0.3	-0.3	9	3.5	4.5	6.5	0.1	0.1	-0.1	2.5	1.3	-1.5	20	29
32	TSGH-8301	2	0.3	-0.3	9	3.5	4.5	6.5	0.1	0.1	-0.1	2.5	1.3	-1.5	30	30
33	BFTC-905	1.8	0.1	-2.5	7	0.8	1	1	0.3	2.5	-0.1	2	0.1	-1.5	5	63
34	BFTC-905	1.8	0.1	-2.5	7	0.8	1	1	0.3	2.5	-0.1	2	0.1	-1.5	10	33
35	BFTC-905	1.8	0.1	-2.5	7	0.8	1	1	0.3	2.5	-0.1	2	0.1	-1.5	20	21
36	BFTC-905	1.8	0.1	-2.5	7	0.8	1	1	0.3	2.5	-0.1	2	0.1	-1.5	30	18

Table 2: The 10 training data selected randomly from the total data set

No.	Cell line	MDR	Topo II	Rb	EGFR	Neu	c-ErbB-3	c-ErbB-4	Cyclin A	Cyclin D1	P16	Cdc 2	Bcl-2	Bax	Co-60 (Gy)	Resistance to radiotherapy
1	HT 1376	1	1	-0.5	5	0.5	6.5	6.5	0.7	0.1	-0.5	2.8	1.8	-0.3	5	97
9	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	5	94
12	TCC-SUP	0.1	1	-1	1.5	0.1	0.1	3	0.1	0.8	-2.3	2.8	0.1	-0.2	30	41
13	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	5	84
16	J82	1.5	3.5	-0.7	2	1	1.5	2.5	2	2.5	-2.3	2	2	-0.1	30	40
21	T24	0.3	1.3	-2.7	2.2	1	0.1	2.5	0.5	2	-0.1	2.5	1.3	-0.2	5	90
25	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	5	83
28	5637	1	1.3	-1	8.5	0.3	0.1	3.5	0.3	0.8	-2.8	2.5	1	-1.3	30	28
30	TSGH-8301	2	0.3	-0.3	9	3.5	4.5	6.5	0.1	0.1	-0.1	2.5	1.3	-1.5	10	30
36	BFTC-905	1.8	0.1	-2.5	7	0.8	1	1	0.3	2.5	-0.1	2	0.1	-1.5	30	18

Table 3: The 100 virtual sample values acquired in Step 3

No.	MDR	Topo II	Rb	EGFR	Neu	c-ErbB-3	c-ErbB-4	Cyclin A	Cyclin D1	P16	Cdc 2	Bcl-2	Bax	Co-60 (Gy)	Resistance to radiotherapy
1	1.5	1.3	-0.7	2.2	0.3	0.1	6.5	0.7	2	-2.3	2.8	2	-0.1	30	28
2	1.5	1	-2.5	9	0.5	6.5	3	0.5	0.8	-0.1	2.8	0.1	-0.1	30	41
3	0.3	1.3	-1	7	0.1	0.1	6.5	2	2.5	-2.3	2.5	1.8	-1.3	30	83
4	1.5	1.3	-1	2.2	0.3	1.5	2.5	0.1	0.1	-2.3	2.5	1	-0.1	5	40
5	0.1	1	-2.5	8.5	0.3	1.5	2.5	0.1	0.1	-0.5	2.5	1	-1.5	30	84
6	0.1	1	-0.7	2	1	0.1	6.5	0.3	0.8	-0.1	2.8	0.1	-0.2	5	84
7	0.1	1.3	-1	7	0.1	1.5	2.5	0.1	0.8	-2.8	2.8	0.1	-0.1	5	30
8	0.1	1.3	-1	9	0.1	0.1	2.5	0.7	2.5	-0.5	2.5	1	-0.3	30	40
9	0.3	3.5	-2.5	9	1	0.1	2.5	0.3	0.8	-2.8	2.8	2	-1.5	5	28
10	0.1	3.5	-0.3	1.5	0.1	4.5	3.5	0.1	2.5	-0.1	2	1	-1.3	5	97
11	1.5	0.1	-1	1.5	1	4.5	2.5	0.3	0.1	-0.1	2.5	2	-0.3	30	83
12	1.8	0.1	-2.7	5	1	1	3	2	2.5	-0.1	2	2	-0.2	5	41
13	1	1	-1	1.5	3.5	6.5	3.5	0.7	0.8	-0.1	2	1.3	-1.3	10	40
14	1	1.3	-1	7	0.1	0.1	3.5	0.1	0.8	-2.3	2.5	0.1	-1.5	5	97
15	0.1	0.1	-2.5	1.5	1	1.5	6.5	0.3	0.1	-0.5	2.5	2	-0.2	10	83
16	1.5	1.3	-0.7	2.2	0.3	1.5	2.5	0.3	0.1	-0.1	2	0.1	-0.1	30	84
17	0.1	0.3	-2.5	1.5	0.3	0.1	6.5	0.7	0.8	-2.3	2.8	1	-0.2	5	28
18	0.1	0.3	-1	7	0.3	0.1	3.5	2	2.5	-2.3	2	1.3	-1.3	5	97
19	2	1	-2.5	1.5	1	0.1	3	0.1	0.1	-2.3	2.5	1.3	-0.3	30	40
20	1.8	1	-1	2.2	1	1.5	2.5	0.1	2.5	-0.1	2.8	1.3	-0.3	5	97
21	0.1	3.5	-1	2.2	0.1	0.1	3.5	2	2.5	-2.8	2.5	2	-1.3	30	40
22	0.3	1.3	-2.5	7	1	4.5	1	0.3	2.5	-0.1	2.5	2	-0.2	30	90
23	1.5	0.3	-0.7	2.2	1	0.1	6.5	0.3	2.5	-2.3	2.8	1.3	-0.2	5	30
24	1	1.3	-1	7	0.8	0.1	6.5	0.3	0.8	-2.8	2	1	-0.1	30	94
25	0.3	1.3	-0.5	2	0.3	0.1	2.5	0.5	0.1	-0.1	2.5	0.1	-0.1	30	41
26	1	3.5	-0.5	8.5	1	1.5	1	0.1	0.8	-0.1	2.8	1.3	-1.5	5	97
27	1	0.1	-2.5	7	0.1	0.1	3	0.3	0.8	-2.8	2.5	2	-1.5	5	18
28	0.1	1.3	-2.5	1.5	0.1	6.5	3.5	2	0.8	-0.1	2.8	1.3	-1.5	5	18
29	1	3.5	-0.7	2.2	1	1	3	2	2	-2.3	2.5	1.8	-1.3	5	83
30	1	3.5	-1	2.2	0.1	0.1	3	2	2.5	-0.1	2.8	2	-0.2	30	83
31	1	0.3	-0.3	9	0.3	0.1	1	0.7	2.5	-2.8	2.8	0.1	-1.3	5	97
32	1	0.1	-0.5	2	0.3	0.1	3.5	0.5	2.5	-2.3	2.8	2	-0.1	10	30
33	2	1.3	-1	5	1	0.1	2.5	0.5	0.1	-0.1	2.8	1	-0.2	30	41
34	1.5	1	-2.5	8.5	0.5	0.1	1	2	0.8	-2.3	2.8	2	-1.5	5	84
35	1	1	-2.5	8.5	1	0.1	2.5	0.3	2.5	-2.3	2.8	0.1	-0.2	5	18
36	1	1.3	-1	2	1	1.5	3.5	2	2	-0.1	2.8	1	-1.5	5	28
37	1	1	-0.7	8.5	0.1	0.1	3.5	0.1	0.1	-2.3	2.5	1	-1.3	30	30
38	1	1	-2.7	2.2	1	6.5	2.5	0.5	2.5	-2.3	2.5	1	-0.2	30	94
39	0.3	1	-0.3	5	0.5	1.5	6.5	0.3	2.5	-2.8	2	1.8	-0.2	30	97
40	1	3.5	-2.7	1.5	3.5	4.5	6.5	0.3	0.1	-0.5	2.8	1	-0.1	10	97
41	1	1	-1	8.5	0.1	1.5	2.5	0.3	0.8	-2.3	2.5	1.8	-1.5	5	84
42	1	1.3	-0.7	8.5	0.8	1.5	2.5	0.3	2.5	-2.3	2.5	2	-0.3	5	84
43	0.3	3.5	-1	9	0.1	1.5	3.5	2	0.1	-2.3	2	2	-1.3	10	84
44	0.1	1.3	-0.7	1.5	1	0.1	2.5	0.1	0.1	-2.3	2.5	0.1	-0.2	5	90
45	1	3.5	-0.3	8.5	0.3	4.5	2.5	2	2	-2.3	2.5	1.8	-0.2	30	30
46	1.5	1	-1	2.2	1	0.1	6.5	0.1	0.8	-0.1	2	2	-0.3	5	30
47	1	3.5	-1	8.5	0.8	6.5	3.5	0.1	2.5	-2.3	2	1	-1.3	5	40
48	2	1.3	-0.7	1.5	1	1.5	6.5	0.1	2	-2.3	2	0.1	-0.1	5	41
49	1.5	1.3	-2.7	7	1	6.5	3.5	0.3	0.1	-2.8	2	1.3	-1.5	5	97
50	0.3	1	-2.7	5	1	1.5	3.5	2	0.1	-2.3	2.5	2	-0.3	30	97
51	1	1	-0.5	1.5	1	1	6.5	0.5	0.1	-0.1	2.8	1	-0.2	5	94
52	1	1.3	-0.7	7	0.8	6.5	6.5	0.3	0.1	-0.1	2	2	-1.5	10	94
53	2	1.3	-1	8.5	0.8	0.1	2.5	2	0.8	-2.8	2.5	1.3	-0.2	5	83
54	1	1.3	-1	1.5	1	1.5	3.5	0.3	2.5	-2.8	2.5	1	-0.1	5	97
55	1	0.3	-2.7	2.2	0.8	0.1	6.5	2	0.1	-0.5	2	1.3	-0.1	30	90
56	1.5	1.3	-1	2.2	0.1	0.1	6.5	0.5	0.8	-0.1	2	1	-0.2	30	28
57	1	1.3	-1	1.5	3.5	0.1	6.5	0.1	0.8	-2.3	2	2	-1.3	5	41
58	2	1	-2.7	7	0.3	0.1	2.5	2	0.8	-0.1	2	1.8	-0.2	5	40
59	1.5	0.1	-0.5	9	1	4.5	6.5	2	0.8	-2.3	2.8	0.1	-1.3	5	94
60	1.5	0.1	-0.7	9	0.1	0.1	6.5	2	0.8	-2.8	2.8	1	-1.5	30	90
61	0.1	0.3	-0.7	9	3.5	0.1	6.5	0.3	0.8	-2.8	2	1	-0.2	30	84
62	2	0.1	-1	2.2	3.5	0.1	2.5	0.3	0.1	-0.1	2	1.3	-0.2	30	41
63	0.3	3.5	-1	2	3.5	0.1	3.5	2	2.5	-2.3	2.8	2	-1.5	10	83
64	0.1	3.5	-2.7	2.2	0.8	0.1	2.5	2	0.1	-2.3	2.8	2	-1.5	5	84
65	2	0.1	-0.5	2	1	0.1	2.5	0.3	0.1	-0.1	2	0.1	-0.1	30	18
66	1.8	1	-2.7	8.5	0.5	0.1	3	2	0.8	-2.3	2.8	1.8	-0.2	5	28
67	0.3	1	-1	9	0.1	1	3	0.1	2.5	-2.3	2	0.1	-1.5	5	18
68	0.1	1.3	-1	2	1	0.1	2.5	2	0.8	-2.3	2.5	1.8	-1.3	10	41

69	1	1.3	-0.7	8.5	3.5	4.5	1	0.1	0.8	-0.1	2.8	0.1	-1.5	5	40
70	1	1	-0.7	8.5	1	0.1	3	0.3	2.5	-2.8	2	1	-0.3	30	83
71	0.1	0.1	-0.7	7	0.1	0.1	1	0.5	2.5	-0.1	2	2	-0.1	30	97
72	1.8	1	-0.7	5	0.1	1.5	6.5	0.3	2	-2.3	2.5	1.8	-1.5	5	18
73	0.1	1	-1	1.5	0.1	1.5	3	0.3	2.5	-0.1	2	2	-0.2	30	41
74	1	1	-0.7	2.2	1	6.5	6.5	0.1	0.8	-0.1	2.8	1	-1.5	10	94
75	1.5	1.3	-2.7	8.5	1	4.5	3.5	0.3	2.5	-0.1	2.5	1	-0.1	5	83
76	2	1.3	-1	2	0.8	0.1	6.5	0.1	0.1	-2.3	2.5	0.1	-1.5	5	97
77	1.8	1	-1	2.2	1	0.1	2.5	0.1	0.8	-0.1	2.5	1	-0.2	30	30
78	1.5	3.5	-0.7	8.5	0.1	0.1	6.5	2	0.8	-0.5	2	0.1	-1.5	30	90
79	1.8	1	-0.5	8.5	0.5	1.5	1	0.5	0.8	-2.3	2.5	2	-1.5	5	83
80	2	1	-1	2	0.8	0.1	1	0.1	2	-0.5	2.5	0.1	-1.3	30	84
81	2	1.3	-0.7	1.5	3.5	4.5	3	0.3	0.8	-2.3	2.5	0.1	-0.1	30	84
82	1	1	-2.5	9	1	6.5	3	0.3	0.8	-2.3	2.5	1.8	-1.3	5	28
83	1	0.3	-1	2.2	0.5	4.5	6.5	0.3	0.1	-2.3	2	1	-0.2	10	41
84	0.3	3.5	-0.7	2	0.5	1.5	3	2	0.1	-0.5	2	0.1	-0.1	30	97
85	1.8	1.3	-0.3	1.5	1	0.1	3	2	0.1	-2.3	2.8	1.3	-1.3	5	90
86	0.1	1	-0.5	9	0.5	0.1	2.5	0.3	0.8	-2.3	2.8	0.1	-1.3	5	83
87	1	3.5	-2.5	8.5	0.1	0.1	2.5	0.3	2.5	-0.1	2.8	1.8	-0.2	30	84
88	0.1	0.3	-1	1.5	1	4.5	6.5	0.5	0.1	-2.3	2.5	1	-0.2	5	41
89	0.1	0.3	-2.5	2	1	1	6.5	0.3	0.8	-0.1	2	2	-0.1	10	41
90	1.5	1.3	-0.5	1.5	0.8	1.5	1	2	2.5	-2.8	2.5	1	-0.3	5	94
91	0.1	0.3	-0.7	5	1	0.1	3.5	0.5	0.1	-2.3	2.8	2	-1.5	10	30
92	0.1	0.1	-1	2	1	0.1	3.5	0.7	2.5	-0.5	2.5	1.8	-1.3	5	40
93	1	3.5	-0.3	2.2	1	1.5	6.5	0.5	0.8	-0.5	2.8	2	-0.1	30	83
94	1	3.5	-2.7	2.2	1	6.5	3.5	0.1	0.8	-0.1	2.8	1	-0.1	5	41
95	2	1	-1	2	0.5	0.1	6.5	0.3	2.5	-0.1	2.5	0.1	-1.5	5	90
96	1.5	1.3	-1	9	0.3	1	3.5	0.3	0.8	-0.1	2	2	-0.1	5	30
97	1	1.3	-1	2.2	0.1	1.5	6.5	0.1	0.8	-0.1	2.5	0.1	-0.2	30	28
98	1	1	-2.5	8.5	0.1	0.1	2.5	2	0.1	-2.3	2	2	-1.3	5	30
99	1.5	0.3	-1	8.5	0.5	4.5	3.5	0.3	2.5	-0.1	2.5	1.8	-0.2	30	83
100	0.1	0.3	-1	8.5	1	0.1	3	0.1	0.8	-2.8	2.8	1	-0.2	30	18

Table 4: The 26 Resistance to radiotherapy and the *output of network* values

Resistance to radiotherapy	84	70	72	36	30	50	82
output	91.54935	84.62182	68.93678	22.52767	22.4529	62.90501	90.76537
Resistance to radiotherapy	60	39	92	32	50	48	63
output	40.21373	25.78362	89.06696	43.57823	64.70494	48.0058	22.40537
Resistance to radiotherapy	57	33	34	39	78	82	58
output	32.48688	21.88049	28.70464	48.82519	86.32889	37.5627	80.14252
Resistance to radiotherapy	29	90	73	70	21		
output	24.43233	91.80353	39.7655	67.35411	21.37013		

Table 5. The computational results (average error rate)

Size of training data set	5	10	15	20	25	30
ANN	48.6726%	37.4388%	27.7921%	28.1446%	20.1596%	16.7608%
Bootstrip	45.5735%	32.0283%	24.9350%	23.9485%	18.5067%	15.2223%

Table 6. The computational results (standard deviation)

Size of training data set	5	10	15	20	25	30
ANN	17.1742%	15.4628%	5.9298%	7.8212%	5.0277%	5.8859%
Bootstrip	3.3111%	3.2479%	3.0056%	2.8743%	2.4845%	1.4231%

Table 7. The computational results (p-value)

Size of training data set	5	10	15	20	25	30
p-value	1.86357E-05	3.83816E-05	0.027710176	0.003205308	0.02372162	0.000123237

References

Amirakian, B., & Nishimura, H. (1994). What size network is good for generalization of a specific task of interest? Neural Networks, *7*, 321-329.

Anthony, M., & Biggs, N. (1997). Computational Learning Theory. Cambridge University Press.

Colquhoun, A.J., Sundar, S., Rajjayabun, P.H., Griffiths, T.R., Symonds, R.P., and Mellon, J.K. (2006). Epidermal growth factor receptor status predicts local response to radical radiotherapy in muscle-invasive bladder cancer. Clin. Oncol. (R Coll Radiol) *18*, 702-709.

Efron, B., & Tibshirani, R.J. (1993). An Introduction to the Bootstrap. New York, Chapmen & Hall.

Gonzalez-Campora, R., Davalos-Casanova, G., Beato-Moreno, A., Garcia-Escudero, A., Pareja Megia, M.J., Montironi, R., and Lopez-Beltran, A. (2007). BCL-2, TP53 and BAX protein expression in superficial urothelial bladder carcinoma. Cancer Lett. *250*, 292-299.

Huang, C.F. (1997). Principle of information diffusion. Fuzzy Sets and Systems, *91*, 69-90.

Huang, C.F., & Moraga, C. (2004). A diffusion-neural-network for learning from small samples. International Journal of Approximate Reasoning, *35*, 137-161.

Ivănescu, V.C., Bertrand, J.W.M., Fransoo, J.C., & Kleijnen, J.P.C. (2006). Bootstrapping to solve the limited data problem in production control: an application in batch process industries. Journal of the Operational Research Society, *57*, 2-9.

Jang, J.S.R. (1993). ANFIS: Adaptive-Network-based Fuzzy Inference Systems. IEEE Transactions on Systems, Man and Cybernetics, *23*, 665-685.

Kotwal, S., Choudhury, A., Johnston, C., Paul, A.B., Whelan, P., and Kiltie, A.E. (2008). Similar treatment outcomes for radical cystectomy and radical radiotherapy in invasive bladder cancer treated at a United Kingdom specialist treatment center. Int. J. Radiat. Oncol. Biol. Phys. *70*, 456-463.

Kubo, A., Nakagawa, K., Varma, R.K., Conrad, N.K., Cheng, J.Q., Lee, W.C., Testa, J.R., Johnson, B.E., Kaye, F.J., and Kelley, M.J. (1999). The p16 status of tumor cell lines identifies small molecule inhibitors specific for cyclin-dependent kinase 4. Clin. Cancer Res. *5*, 4279-4286.

Li, D.C., & Lin, Y.S. (2006). Using virtual sample generation to build up management knowledge in the early manufacturing stages. European Journal of Operational Research, *175*, 413-434.

Li, D.C., Wu, C.S., Tsai, T.I., & Chang, F.M. (2006). Using Mega-Fuzzification and Data Trend Estimation in Small Data Set Learning for Early FMS Scheduling Knowledge. Computers & Operations Research, *33*, 1857-1869.

Li, D.C., Wu, C.S., Tsai, T.I., & Lin, Y.S. (2007). Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. Computers & Operations Research, *34*, 966-982.

Milosevic, M., Gospodarowicz, M., Zietman, A., Abbas, F., Haustermans, K., Moonen, L., Rodel, C., Schoenberg, M., and Shipley, W. (2007). Radiotherapy for bladder cancer. Urology *69*, 80-92.

Nix, P., Cawkwell, L., Patmore, H., Greenman, J., and Stafford, N. (2005). Bcl-2 expression predicts radiotherapy failure in laryngeal cancer. Br. J. Cancer *92*, 2185-2189.

Niyogi, P., Girosi, F., & Tomaso, P. (1998). Incorporating prior information in machine learning by creating virtual examples. Proceeding of the IEEE (pp. 275-298).

Pollack, A., Czerniak, B., Zagars, G.K., Hu, S.X., Wu, C.S., Dinney, C.P., Chyle, V., and Benedict, W.F. (1997). Retinoblastoma protein expression and radiation response in muscle-invasive bladder cancer. Int. J. Radiat. Oncol. Biol. Phys. *39*, 687-695.

Tsai, T.I., & Li, D.C. (2008). Utilize bootstrap in small data set learning for pilot run modeling of manufacturing systems. Expert Systems with Applications *35*, 293-1300.

Tut, V.M., Braithwaite, K.L., Angus, B., Neal, D.E., Lunec, J., and Mellon, J.K. (2001). Cyclin D1 expression in transitional cell carcinoma of the bladder: correlation with p53, waf1, pRb and Ki67. Br. J. Cancer *84*, 270-275.

Vladimir, N.V. (2000). The Nature of Statistical Learning Theory. New York, pringer.

Wang, Z.N., Dimassimo, C., Tham, M.T., & Morris, A.J. (1994). A procedure for determining the topology of multilayer feedforward neural networks. Neural Networks, *7*, 291-300.

Williamson, M.P., Elder, P.A., Shaw, M.E., Devlin, J., and Knowles, M.A. (1995). p16 (CDKN2) is a major deletion target at 9p21 in bladder cancer. Hum. Mol. Genet. *4*, 1569-1577.