Cancer Growth Prediction via Artificial Neural Networks

Rishi Pal, Pardeep Garg, Rajiv Chechi, Sandeep Kumar, Naresh Kumar
Electronics & Communication Engineering Department
Haryana College of Technology & Management, Kaithal, Haryana, India
er.rishi.p@gmail.com, garg.pardeep22@gmail.com, rajivchechi@gmail.com,
chechi.sandeep@gmail.com, naresh.chhoker@gmail.com

Abstract

Artificial Neural Network (ANN) abstracts the chemical events at synapses into simple transfer functions, usually sigmoid in shape. Through a variable topological connectedness, they are capable of learning, optimizing, recognizing patterns and other pseudo-cognitive functions. Most cells in the organism communicate by chemical signals of a variety of types, neurons being special because of their highly specialized anatomy. Seven primary sites for tumors are identified: 1.Buccal Mucosa 2. Tongue, 3.Retromolar Trigone 4. Floor of Mouth 5.Ventral Tongue 6. Oropharynx 7. Lower Alveolus. Ten secondary regions of lymph nodes metastasis are observed: five regions each on the same and opposite side (ipsilateral/contralateral) as the primary tumor site.: In this paper a study is carried out on oral squamous cell carcinoma using ANN and data analysis approach is explored with two ANN methods:- (1) a supervised multilayer feed forward back propagation (back-prop) method, and (2) an unsupervised self-organizing map (SOM) method. This experience provides insight into implementation of ANN and directions to future investigation. The results from back-prop are comparable to that using multidimensional scaling (MDS) with respect to prediction of lymph nodes that have highest percentage of being metastasized, while SOM requires further work to identify clustering for individual primary cancer as well as next level of lymph node metastases..

Keywords: ANN, Back Propagation, MDS, SOM, metastates..

1. Introduction

Knowledge of tumour growth is important in the planning and evaluation of screening programs, clinical trials and epidemiological studies. The successful treatment of most critical illnesses is based on early and accurate detection. The use of technology in medicine has changed the face of medical practice completely. But most of this technology is based on detecting the problem only after it has occurred. Even with the availability of superior computing and imaging capabilities, the concept of "*prevention is better than cure*" has not caught on in medical practice.

The human body is made up of many cells. Normally, the cells grow and die in a controlled manner. However, sometimes, cells will keep dividing in an uncontrolled manner, forming tumours. Most tumours do not invade surrounding tissues and are non-life threatening. They are deemed benign tumours. If the tumour does invade and destroy nearby tissues, it is classified as malignant tumour or cancer and might threaten the person's life. The occurrence of tumour is preceded by gradual change in

cell density in the neighborhood. If these signs are detected then it is possible to prognosticate the severe conditions that are likely to occur. In this case the tumour can be predicted well before it actually appears and appropriate counter measures can be taken at an earlier stage with a higher degree of success. The effectiveness of such a system is gauged by how early the changes are detected and the accuracy of the predictions.

Communication between cells is almost exclusively chemical. Many different forms have been observed [1-2]. Neural networks are among the most extremely studied examples known [3-4]. Since the time of McCollough and Pitts' pioneering work [3], the study of artificial neural networks as a set of models or abstractions of the actual biological networks has led to a plethora of astounding, significant results. More recently, Jerne has formulated a theory of immune networks which has also been developed very rapidly [5-6]. In each of these cases, successful studies have resulted from simplified models which emphasize a limited set of distinct properties selected from among the many networks displayed in the organism. This theme can be extended to almost any form of chemical communication between cells and should be fruitful if applied to cancer.

1.1 Stages of Cancer

The human body's numerous biological processes, such as cell growth, are extraordinarily well-controlled, for the most part. When cells undergo reproduction, there are four stages they must pass through: Gap 1, DNA synthesis, Gap 2, and mitosis.

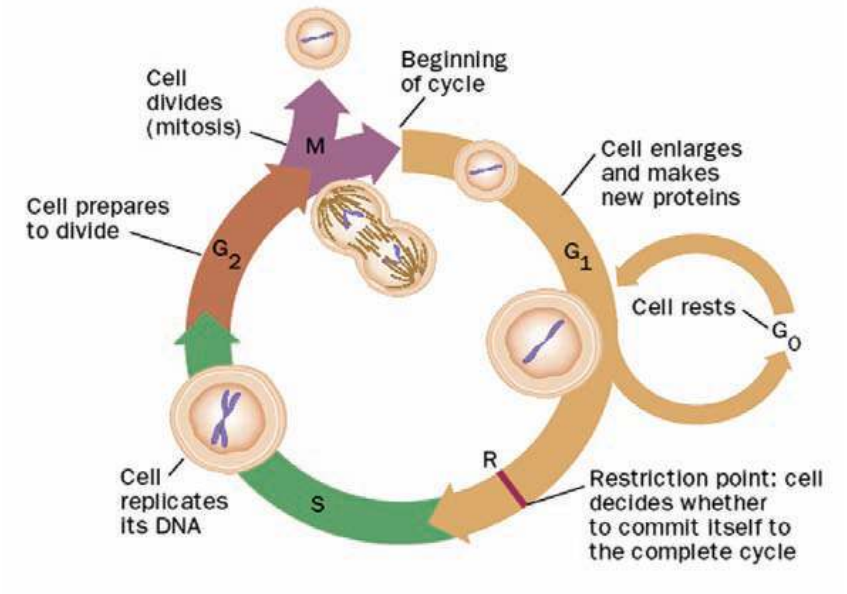

Figure 1. Cell Cycle Clock

A normal cell life cycle is often depicted as a so-called cell cycle clock. The clock is composed of four stages. In the first stage (G1) the cell increases in size and prepares to copy the DNA contained in its nucleus. The copying is performed in the synthesis (S) stage. In this phase, all the chromosomes of the cell are replicated and the cell then enters the G2 stage. During this stage, the cell prepares for mitosis (M stage), where the mother cell divides into two daughter cells, each containing a complete set of chromosomes. The two daughter cells immediately enter G1 and will either follow the cell cycle clock or temporarily or permanently stop cycling. At the restriction point (R) which occurs in the late G1 phase, a molecular "switch" determines whether the cell continues the cell cycle, enters the G0 state (senescence), or undergoes apoptosis. Two classes of genes play major roles in triggering cancer. Cells in the human body only can survive when their area code where the cell is compatible with the code of the surrounding cells and the extracellular matrix. As a consequence normal cells leaving their cell structure will die within a short time by apoptosis. This property is called anchorage dependence. In cancer cells, some proteins produced by oncogenes trick the nucleus by making it seem that the cell is attached while in reality it is not. This makes cancer cells anchorage independent and allows them to migrate freely throughout the human body via the blood circulation system.For modelling the cancer growth there are three stages.

1.2 Avascular Growth

Solid tumours arise from an alteration of a cell's genetic material causing it to respond differently to the host's growth regulators and this leads to the uncontrolled growth of these cells. The early tumour cells, like the neighboring normal cells, can obtain adequate nourishment (such as glucose and oxygen) from the existing vasculature, and hence the proliferation rate of these cells is regular, producing a growing spheroid of cells expanding at a near exponential rate. As the tumour grows the amount of nutrient that manages to diffuse to the centre decreases due to consumption by the outer cells and eventually the central cells become so deficient that their proliferative rate is reduced and cell division may indeed cease, with these cells becoming quiescent. These quiescent cells are still viable and they can recover on the restoration of sufficient nutrient. This reduction of proliferation within the spheroid retards its growth. As the tumour grows further, the continued absence of nutrients in the central regions will cause the cells there to die, forming a region of dead cells known as a necrotic core. As the spheroid continues to develop, the rim of adequately nourished viable cells at the surface becomes roughly constant in size, leading to a phase of near linear growth. Eventually, after a period of a few weeks, the combined action of necrotic disintegration, accumulation of waste products, mitotic inhibitory factors and cell shedding reduces the rate of growth of *in vitro* tumours, leading to a spheroid of a maximal (saturation) size.

1.3 Vascularization

In order to grow beyond the diffusion-limited state, tumours have to have a blood supply. This is achieved by secreting a TAF, tumour angiogenesis factor. The TAF diffuses across the tissue between the tumour and blood vessel and activates angiogenesis, new blood vessel formation. Blood and lymphatic vessels develop in a parallel, but independent manner, and together form the circulatory system allowing the passage of fluid and delivering molecules within the body.

1.4 Metastasis

Cancer can begin in any organ or tissue of the body, and is usually named for the part of the body in which it starts. In the case of mesothelioma, it begins in the linings of either the lung, the abdominal cavity or, in rare cases, the heart, hence the terms pleural mesothelioma, peritoneal mesothelioma and pericardial mesothelioma. These original tumours are known as "primary tumours"

Metastasis refers to spread of the cancer, and occurs when cancer cells break away from the primary tumour and enter the bloodstream or the lymphatic system where they are carried to other parts of the body. When a new tumour forms at a site away from the primary site, it is known as a "metastatic tumour". Cancer cells can spread anywhere in the body, but frequently spread to the lymph nodes near the primary tumour. This is called "lymph node involvement" or "regional disease". Cancer that spreads to other organs or to lymph nodes far from the primary tumour site is called "metastatic", or "distant disease".

Many studies have been conducted in recognizing the pattern of lymph node metastases [2]. However, their findings are reported in a number of different ways, such as in text or in tabular fashion that make it difficult to compare and summarize the results. Thus it would be beneficial to generate visualization of the resulting data from the studies. The benefits are especially great for large or high dimensional data. Several methods are used in visualizing high dimensional data. Out of these one which is more traditionally used method is:- Multidimensional Scaling (MDS) with which the previous research on the same data was conducted. The key idea of MDS is to produce low dimensional visualization of the data so that the distance of data in the new space approximates the distance of data in the original space. It has been shown that MDS can summarize disease progression and produce a map of the metastases patterns in cancer [3,7]. While MDS demonstrates that it can summarize the cancer data, it faces difficulties in detecting the intrinsic dimension of the data, and to discover complicated nonlinear structure. Since artificial neural networks (ANN) has been successfully applied in bio-medical research and medicine decision support, a great deal of interest exists in trying to overcome MDS' drawbacks by analyzing and visualizing data using ANN. This paper attempts cancer metastasis prediction with a few ANN methods. The organization of the paper is as follows. In section 2 the tumor behavior and growth factors are discussed. Section 3 deals with the basic structure of

network used in the model. Section 4 shows the results using multilayer feed forward network and self organizing map. In section 5 the future work is given. Finally conclusion is drawn in section 6.

2. The Biological Problem

Tumors are generally heterogeneous populations of cells. Tumor cells are able to communicate with each other by chemical signals. The autocrine hypothesis suggests that some tumor cells secrete and respond to their own growth factors via specific membrane receptors. The response to these factors is an acceleration of the mitotic rate. Cells also secrete factors which inhibit growth and the factors also act through membrane receptors. Often stable tumor populations seem to be destabilized by chemical or surgical deletion of subpopulations of cells [7-8]. The model will attempt to reproduce many of these observations with a minimum of molecular detail and specificity. In particular, the cells will be genotypically equivalent and no mutations will be considered. This is clearly simplification to the extreme.

3. The Model

The basic structure of the model resembles a Multilayer network with the following significant modifications:

1. All weights are positive. The outputs of the cells can range from -0.5 to 0.5.
2. There are two outputs from each cell. One is stimulatory to growth, the other inhibitory. These can be viewed as two coupled Back-prop networks operating with the same weights, topology and values.
3. Each cell's receptors are also sensitive to levels of the signalling compounds in the vicinity and modify their affinity for the signal substances accordingly.
4. The transfer functions for calculating the outputs are more complex.
5. Random variation in the calculated outputs is constantly introduced. This random signal variation is a small percentage of the output, but significant.
6. Cells divide and die as a function of the stimuli received.
7. The weights of the connections between cells are designed to approximate the diffusion distance between cells and its alteration due to tumor growth in three dimensions.

4. Results

Simulations of two basic types are performed. One is designed to explore the variety of population behaviors by using random starting conditions with a single original cell. The second is to look at the effect of deleting subpopulations at various times during the tumors growth process (Subpopulations are defined interms of levels of secretion of the chemical factors. Six subpopulations are arbitrarily defined on the basis of quantitative differences in [l] secretion levels).

Method 1 – Supervised Back-Prop ANN: Two experiments are conducted with Back-prop using Easy NN-plus software. Experiment #1A aims to classify which secondary site each of the primary tumors will metastasize. Experiment #1B takes the secondary sites as input to the network, and aims to classify which is the primary tumor it originates from. Parameters to be adjusted to arrive at optimal network architecture are the learning rate, momentum, target error, number of iterations, and validation rules. The best architecture for Experiment #1A contains 7 input nodes (7 primary tumor sites), two hidden layers with 4 and 5 nodes in each layer, and 10 output nodes (10 secondary targeted nodes regions). The best architecture for Experiment #1B is 10-6-5-7 as shown in figure 2.

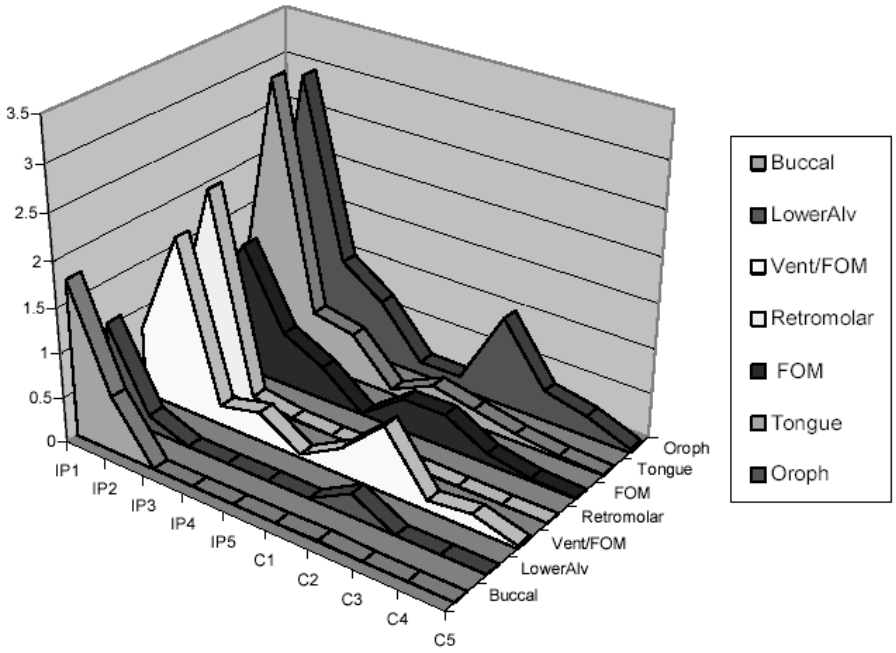

Figure 2. Growth of single cell

Method 2 – Unsupervised SOM: The second method employs an unsupervised algorithm called the self-organizing maps (SOM), with experiments conducted using ANN software tool box in the MATLAB® 7. The training process of SOM describes a topology-preserving mapping from a high dimensional input space onto a two-dimensional output space where patterns that are similar in terms of the input space are mapped to geographically close locations in the output space. Experiment #2A uses clustering of output data for a given primary tumor. Experiment #2A considers data only in an individual primary tumor, while Experiment #2B accepts all 130 cases. The first SOM architecture consists of 1 input node, and 10 output nodes, while the second SOM architecture consists of 10 input and 70 output nodes. Figure 3 shows the growth of cancerous tumour by taking growth factor (a) and proliferation rate (K) as 0.35 & 40 respectively.

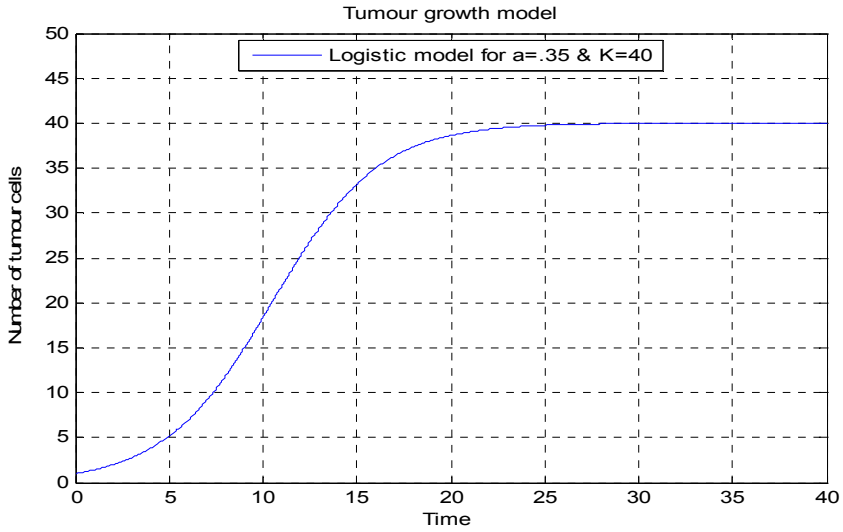

Figure 3. Growth of Cancerous Tumour

A. Overall trends in the growth pattern.

There were three basic patterns of behavior observed:

1. The population grew to a certain size and stabilized.
2. The population grew to a certain size, remained in an apparent stable pattern, but suddenly began to grow rapidly until the program's capacity was exceeded.
3. The population grew to a certain size, remained at that size for some time, then began to grow rapidly [3] only to finally die off.

B. Effects of deleting subpopulations

Both the time of deletion and identity of the subpopulation deleted are important in determining the outcome of any given deletion. These outcomes varied from a return to the growth pattern observed before the deletion to the extinction of the entire population. This spectrum of outcomes conforms to the variety of outcomes observed in real populations of tumor cells.

5. Future Work

The ultimate goal for modeling tumour growth is to gain a deeper understanding of their growth kinetics such that one can predict the effectiveness of any tumour treatment. By using the model, it is investigated that tumour growth as an emergent and self-organized behavior and this opens up new paradigms in tumour research and

clinical practice. Using this model, the effects of treating a certain portion of the tumour can be modeled easily and this is particularly useful in predicting the growth of cancer after some time.

An important task undertaken in this paper was to use the number of the network-hidden-layer neurons. In our study, the optimal number of the neurons depends on the applied training methods and takes values in the range from 5 to 7. This shows that a relatively small network is able to solve complex diagnostic problems. The number of the neurons will be optimized for each set of the training data by using some other optimizing techniques such as snake spray optimization etc.

6. Conclusion

As it is known that artificial neural network is successfully applied in biomedical science applications. It is concluded that ANN is capable of learning, optimizing, recognizing patterns and other pseudo- cognitive functions. The standard future cancer growth prediction methods reported in literature have not taken into consideration the concept of ANN. In this paper, both the aspects (i.e. cancer growth prediction via artificial neural networks) have been combined. As shown in figure 2, the curve drawn between time and number of cells for a given value of growth rate (a) & Proliferation rate (K), it is concluded that cancer cell in initial stage grows slowly. After some time it starts growing exponentially and after that it attains the metastases stage which is considered as incurable.

7. References

[1] H. Rasmussen, "*Cell Communication in Health and disease",* New York, W. H. Freeman & Co., 1991.
[2] D. C. Mikulecky, "*A Comparison Between the Formal Description of Reaction and Neural Networks: A Network Thermodynamic Approach*", in *Biomedical Engineering: Opening New Doors,* D. *C.* Mikulecky and A. M. Clarke, eds., New York: New York University Press, pp 67-74, 1990.
[3]W. **S.** McCullough and W. H. Pitts, **"***A Logical Calculus of the Ideas Immanent in Nervous Activity*", *Bulletin of Mathematical Biophysics,* Chicago: University of Chicago Press, pp 115-133, 1943.
[4] C. Mead, *Analog VSLI and Neural Systems,* Reading: Addison- Wesley, 1989.
[5] N. K. Jerne, "*Towards a Network Theory of the Immune System*", *ANN. Immunol. (Inst. Pasteur) 125C,* pp. 373-389, 1974.
[6] A. S. Perelson, "*Towards a Realistic Model of the Immune System*", in *Theoretical Immunology,* part 11, A. **S.** Perelson, ed., Redwood City, Addison-Wesley, pp 377-4011988.
[7] J. Prideaux, "*A Systems Approach to Modeling and Simulating the Interactions between Cells in a Solid Tumor",* Master's Thesis, Virginia Commonwealth University, Richmond, VA, 1992.
[8] Gray, L, Woolgar, J., and Brown, J, "*A functional map of cervical metastases from oral squamous cell carcinoma*. *Acta Otolaryngologica,* 120:885-890, 2000.
[9] Karakiewicz, P.I., Shariat, S.F., Palapattu, G.S., Gilad, A.E., Lotan, Y., Rogers, C.G., Vazina, A., Gupta, A., Bastian, P.J., Perrotte, P., Sagalowsky, A.I., Schoenberg, M., Lerner, S.P.: Nomogram for predicting disease recurrence after radical cystectomy for Transitional Cell Carcinoma of the Bladder. J. Urol. 176, 1354–1362, 2006.

Authors

Rishi Pal born on March 04, 1982. After completing Bachelor of Engineering in the field of Electronics & Communication Engineering from HCTM, Kaithal, Haryana, India in 2006, he did his Master of Technology in the field of Electronics & Communication Engineering from NITTTR, Chandigarh, India. Presently, he is working as Lecturer in the Department of Electronics & Communication Engineering at Haryana College of Technology & Management, Kaithal, Haryana (India). His areas of interests include Artificial Neural Networks, Image Processing, Signal Processing etc.

Pardeep Garg is Lecturer in Electronics & Communication Engineering Department at Haryana College of Technology & Mangement, Kaithal, Haryana, India. He completed his M. Tech in Electronics and Communication Engineering from Jaypee Institute of Information Technology University, Noida, India in 2009. He completed B. Tech in Electronics & Communication Engineering from Haryana College of Technology & Mangement, Kaithal, affiliated from Kurukshetra University, Kurukshetra, India in 2004. His areas of interests include Artificial Neural Networks, Image Processing, Signal Processing and Wireless & Mobile Communication.

Rajiv Chechi is Associate Professor & Head in Electronics & Communication Engineering Department at Haryana College of Technology & Mangement, Kaithal, Haryana, India. Presently, he is associated with Thapar University, Patiala, India for his research work in the field of Wireless & Mobile Communication. He completed his M. Tech in Electronics and Communication Engineering from Punjab Technical University, Jalandhar, India in 2006. He completed B. Tech from Institute of Technology & Mangement, Gurgaon, affiliated from Maharishi Dayanand University, Rohtak, India in 2000. His areas of interests include Artificial Intelligence, Neural Networks, Wireless & Mobile Communication, Wireless Technology role in Medical Services, Image Processing and Signal Processing.

Sandeep Kumar is Lecturer in Electronics & Communication Engineering Department at Haryana College of Technology & Mangement, Kaithal, Haryana, India. He completed B. Tech in Electronics & Communication Engineering from Haryana College of Technology & Mangement, Kaithal, affiliated from Kurukshetra University, Kurukshetra, India in 2007. His areas of interests include Wireless & Mobile Communication and Image Processing.

Naresh kumar is Lecturer in Electronics & Communication Engineering Department at Haryana College of Technology & Mangement, Kaithal, Haryana, India. He is pursuing his M.Tech in Electronics & Communication Engineering from Haryana College of Technology & Mangement, Kaithal, affiliated from Kurukshetra University, Kurukshetra.He completed B. Tech in Electronics & Communication Engineering from Haryana College of Technology & Mangement, Kaithal, affiliated from Kurukshetra University, Kurukshetra, India in 2007. His areas of interests include Wireless & Mobile Communication and Image Processing.