Proceedings of the 3rd World Congress on Electrical Engineering and Computer Systems and Science (EECSS'17)
Rome, Italy – June 4 – 6, 2017
Paper No. ICBES 128
ISSN: 2369-811X
DOI: 10.11159/icbes17.128

An Adaptive Neuro-Fuzzy Inference System Approach to Neutrophil Prediction in Childhood Leukaemia

D. Avdic, M. S. Gallimore, C. Bingham
School of Engineering, University of Lincoln
Lincoln, UK
davdic@lincoln.ac.uk; mgallimore@lincoln.ac.uk; cbingham@lincoln.ac.uk

Abstract **–** Acute Lymphoblastic Leukaemia (ALL) is the most common form of cancer in childhood. Chemotherapy treatment for ALL involves three phases viz. induction, intensification and maintenance. The maintenance phase includes daily intake of the drug 6-Mercaptopurine that helps kill any remaining abnormal cells and prevent relapse. A well-known side effect of the drug, however, is a reduction in neutrophils, a type of white blood cell, leading to an increased risk of secondary infection, periods of hospitalisation and time off treatment. Blood counts are monitored on a weekly basis and drug dosages altered in an attempt to minimise this risk. However, there is currently no intelligent method of determining optimum dosages for individual patients and typically neutrophil counts will drop below an acceptable level every 6-8 week. This paper proposes a seven-day ahead neutrophil count prediction methodology based on a Takagu-Sugeno model and an Adaptive Neuro-Fuzzy Inference System (ANFIS) in order to support dosing decisions. The methodology is applied to data for the maintenance phase from one female patient, where it is shown to accurately predict neutrophil counts seven-days ahead with a mean squared error of 0.25. This is not only significant in terms of improving the outcomes for ALL patients but also has the potential to be applied to the treatment of other forms of cancer and other diseases where personalised dosing is important, e.g. in Schizophrenia treatment.

Keywords**:** Leukaemia, Neutrophil count prediction, Adaptive Neuro-Fuzzy Inference System, personalised drug dosing

1. Introduction

Acute lymphoblastic leukaemia is a cancer of a type of white blood cell (WBC) known as lymphoblasts. ALL manifests itself as an overproduction of lymphoblasts in the bone marrow inhibiting the production of normal healthy blood cells, i.e., red and white blood cells, and platelets. As a result, the human body loses the potential to fight infections that can ultimately lead to death [1]. ALL is mainly seen in children of 2-5 years of age [2], and represents around 80% of all leukaemia cases [3]. The need for prediction of the neutrophil counts in childhood Acute Lymphoblastic Leukaemia (ALL) has increased as the number of children with ALL grows worldwide (with 400 new cases per year in the UK alone [4]). The treatment of ALL is 2 years for girls and 3 years for boys and is delivered in three phases: induction, intensification and maintenance phase. In the induction phase, various drugs (e.g. vincristine, methotrexate and dexamethasone) are used for the initial remission. In the intensification phase, the aim is to kill any remaining abnormal cells, and in the maintenance phase, the most protracted phase, the aim is to kill any remaining cells and minimize the chance of relapse. This phase includes the administering of a chemotherapy drug called 6-mercaptopurine (6-MP), administered orally on a daily basis [5]. The aim is to keep neutrophil counts between $[1,1.5] \times 10^9$ neutrophils per litre of blood by altering weekly dosages of 6-MP. Typically the drug is administered in 100%, 75% or 50% of the initially calculated dosage (per kg body mass). When the count is lower than 0.5×10^9 patients are classed as neutropaenic, i.e. no effective immune system, at which point treatment ceases until the counts increase. Conversely, increasing drug dosages, where appropriate, can be beneficial to treatment success. However, dosing regimens are not always able to support those decisions with a satisfactory degree of accuracy. Typically, neutropenia results in regular hospitalizations due to secondary infections and/or unscheduled breaks in treatment, with 60% of ALL fatalities caused by such infections [6]. This paper proposes a novel neutrophil prediction methodology, based on a Takagu-Sugeno model and an Adaptive Neuro-Fuzzy Inference System (ANFIS) in order to support dosing decisions. The ability to predict neutrophil counts 7-days ahead in order to inform dosing regimens is significant as it allows medical professionals to maximise patient

outcomes, reduce the risk of secondary infections and associated risks, as well as minimise treatment costs and impact on patient families.

Many algorithms have been used in medical support such as Support Vector Machines (SVMs) for anti-diabetic drug failure prediction by Kang, S. *et al.* [7] and an application of Artificial Neural Networks (ANNs) to predict the likelihood of patient response to the clozapine drug during the treatment of schizophrenia by Lin, C. *et al.* [8]. However, to the authors' knowledge, there is no reported work into neutrophil prediction in leukaemia patients. One technique, Adaptive Neuro-Fuzzy Inference System (ANFIS), has been applied in many different prediction systems; such as weather prediction [9], engineering [10] and also some medical applications. For example, Mathur *et al.* [11] apply ANFIS to the prediction of skin temperature in lower limb prostheses, and Guler *et al*. [12] for the detection of electrocardiographic changes in patients with partial epilepsy.

2. Methodology

A pre-processing technique, Principal Component Analysis (PCA), is used to find the number of inputs that are most relevant to neutrophil count prediction. Data is interpolated in order to provide sufficient data points for model development. The seven-day ahead prediction is made and error calculated based on comparison with actual outputs. A clinical dataset from one female ALL patient is been used for this underlying study, which included 15 inputs. Raw data of neutrophil counts from the maintenance phase before the interpolation is shown in Fig 1.

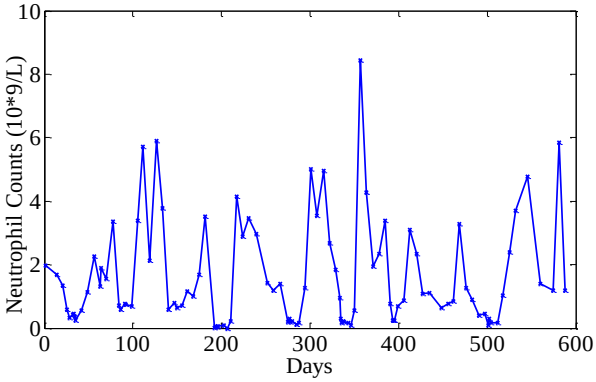

Fig. 1: Neutrophil counts during the maintenance phase.

For this study, blood measurements are taken on a weekly basis and interpolation (using linear interpolation) is used to provide sufficient data to formulate a predictive model. Principal Component Analysis (PCA) is used to remove unwanted characteristics in the data sets (eg. noise) whilst keeping the dominant underlying characteristics, according to (1),

$$S = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^T(x_i - \mu) \tag{1}$$

where $x_i \in X, \mu$ is the sample mean and N is the number of samples, so that,

ANFIS was first proposed by Jang [13] in the 1990s. It is a multilayer feed-forward network that combines the qualities of a fuzzy-inference system and artificial neural network (ANN) to develop models based upon a fuzzy inference system with optimized rules and membership functions. The Takagi-Sugeno (T-S) fuzzy system is a powerful tool for modelling and control in complex systems and has been proven to be an efficient solution in non-linear system modelling [14]. An adaptive neuro-fuzzy interface system (ANFIS) is based on T-S fuzzy interface system. ANFIS is a hybrid algorithm that combines the high level reasoning capability of Fuzzy Inference Systems (FIS), and the high level power of pattern recognition of Neural Networks (NN). The model was developed using MATLAB R2012b with the Fuzzy Logic Toolbox. The ANFIS architecture consists of 5 layers with every layer performing a special action. Layer 1 contains the data inputs. Layer 2 is the 'If' layer, that converts any inputs to fuzzy inputs and generates the membership function (MF) for each of the inputs. The MF used in this study is *gbellmf*. The output of each node in this layer is given by (2),

$$O_{1,i} = \mu_{A_i}(x) \qquad for\ i = 1,2.. \qquad O_{1,i} = \mu_{b_{i-2}}(x) \quad for\ i = 3,4.. \tag{2}$$

where, O_(1,i) is the membership grade of the input nodes x and y. In layer 3, each output generated from the second layer is normalized and nodes are fixed. Layer 4 converts all the fuzzy data into the original data type used. This layer contains fixed nodes that calculate the ratios of the firing strengths of the rules. Finally, the total of the all neurons in layer 4 form the output in layer 5. MFs are tuned using a hybrid algorithm that combines gradient descent and the least square methods. The hybrid learning algorithm is used here in order to identify variables in the ANFIS architecture. The hybrid learning algorithm is split into two parts. The forward pass applies the least-squares method to find the consequent parameters in Layer 4. In the backward pass, the error signals propagate backward and the premise variables are updated by gradient descent [13].

3. Results and Discussions

PCA identified important data which were are identified as white blood cell count, platelet count, 6-MP dosage and the neutrophil counts. These input parameters alone are taken as inputs into the ANFIS model, and neutrophil counts 7-days ahead are predicted. All results are significant according to a 95% confidence level. The Mean Squared Error (MSE) is used to validate system performance according to (3),

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(A - P)^2 \tag{3}$$

where, A is the actual data value, P is the predicted results and N is the number of predicted sample points in the study.

Once the membership function is determined using the hybrid-optimization technique, the system is ready for training. The training is progressed by using the first six months of maintenance phase. The process completes in 100 epochs, with a training error target of 0.005. Independent testing data is applied to the ANFIS model to predict 7 days ahead. Each week of subsequent data is then successively added to the training dataset and the following week's neutrophil count predicted, recursively, for a total of ten weeks. Fig. 2 shows the comparison between the actual and predicted counts. It is shown that the system accurately predicts the weekly counts and adapts itself to the new data, which verifies its reliability for future use in this field.

The MSE is found and the summery of the overall prediction results is shown in Table 1. MSE demonstrates the deviation between actual and predicted values. As can be seen, the average prediction error based on MSE is 0.25.

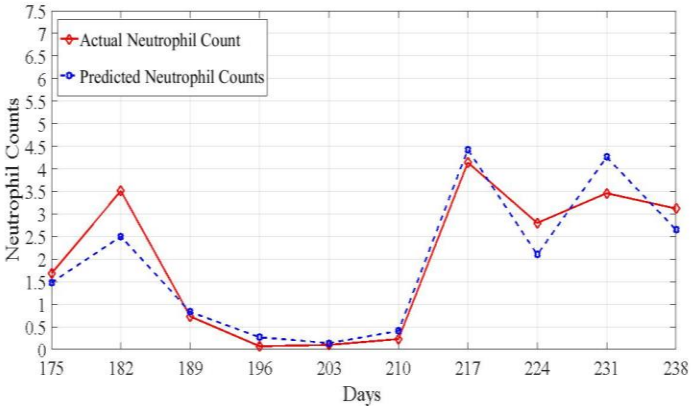

Fig. 2: Week ahead neutrophil count prediction using ANFIS.

Table 1: Error and CI in ANFIS neutrophil prediction.

Normalized error	**Error in Neutrophil prediction and 95% confidence intervals**		
MSE	***Lower***	***Mean***	***Upper***
0.25	0.21	0.4	0.59

4. Conclusion

This paper has proposed an intelligent technique to support medical professionals in the dosing of 6-MP during the maintenance phase of treatment for ALL. An Adaptive Neuro-Fuzzy Inference System (ANFIS) based approach is presented and validated using blood data from one female ALL patient for the full maintenance phase. PCA is applied to the full set of input parameters in order to identify and remove redundant inputs to minimize the risk of over-fitting and reduce the amount of training data necessary. The ANFIS model is used to predict the patient's neutrophil counts 7-days ahead and compared with actual counts. The Mean Squared Error (MSE) is used to validate prediction accuracy where it is shown to be capable of predicting counts with a MSE of 0.25. The ability to accurately predict neutrophil counts, and thus provide a means for more accurate drug dosing, is significant. This will minimise the risk of secondary infections and hence hospitalisation and time off treatment. This not only increases treatment success but also helps reduce associated treatment costs and the impact on patient families. The development of this methodology is not only significant in terms of improving the outcomes for ALL patients, it also has the potential to be applied to the treatment of other forms of cancer and other diseases where personalised dosing is important, e.g. in Schizophrenia treatment.

References

[1] S. Mishra, B. Majhi, P. Sa and L. Sharma, "Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection," *Biomedical Signal Processing and Control*, vol. 33, pp. 272-280, 2017.

[2] K. Kulkarni, R. Arora and R.K. Marwaha, "Survival outcome of childhood acute lymphoblastic leukemia in India: a resource-limited perspective of more than 40 years," *J. Pediatric Hematol./Oncol*., vol. 33, no. 6, p. 475-479, 2011.

[3] A. Radaelli, B. Laskin, J. M. Stephens, M. Botteman and C. Pashos, "A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL)," *Eur J Cancer Care (Engl*), vol. 14, no. 1, pp. 53-62, 2005.

[4] "Macmillan cancer support," 1 January 2013. [Online]. Available: http://www.macmillan.org.uk/cancerinformation/cancertypes/childrenscancers/typesofchildrenscancers/acutelymphoblasticleukaemia.aspx

[5] "Chemo care," [Online]. Available: http://chemocare.com/chemotherapy/drug-info/6-mp.aspx

[6] M. Asim, A. Zaidi, T. Ghafoor and Y. Qureshi, "Death analysis of childhood Acute Lymphoblastic Leukaemia; experience at Shaukat Khanum Memorial Cancer Hospital and Research Centre, Pakistan," vol. 61, no. 7, pp. 666-670, 2001.

[7] S. Kang, P. Kang, T. Ko, S. R. S. Cho and K. Yu, "An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction," *Expert Systems with Application*, vol. 42, no. 9, pp. 4265-4273, 2015.

[8] C. Lin, Y. Wang, J. Chen, Y. Liou, Y. Bai, I. Lai, T. chen, H. Chiu and Y. Li, "Artificial neural network prediction of clozapine response with combined pharmacogenetic and clinical data," *Computer methods and programs in biomedicine*, vol. 91, no. 2, pp. 91-99, 2007.

[9] A. H. Setyaningrum and P. M. Swarinata, "Weather Prediction Application Based on ANFIS (Adaptive Neural Fuzzy Inference System) Method In West Jakarta Region," *in 2014 International Conference on Cyber and IT Service Management (CITSM)*, 2014.

[10] F. Cabalar, A. Cevik and C. Gokceoglu, "Some applications of Adaptive Neuro-Fuzzy Inference System (ANFIS) in geotechnical engineering," *Computers and Geotechnics*, vol. 40, pp. 14-33, 2012.

[11] N. Mathur, I. Glesk and A. Buis, “Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses,” *Medical Engineering and Physics*, vol. 38, p. 1083-1089, 2016.

[12] I. Guler and E. D. Ubeylu, “Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction,” *Expert Systems with Applications,* vol. 27, pp. 323-330, 2004.

[13] J. Jang, “ANFIS: adaptive-network-based fuzzy inference system,” *IEEE Trans. Syst. Man Cybern*., vol. 23, pp. 665-685, 1993.

[14] R. Qi and M. A. Brdys, “Adaptive Fuzzy Modelling and Control for Discrete-Time Nonlinear Uncertain Systems,” *in American Control Conference,* Portland, 2005.