UNIVERSITI PUTRA MALAYSIA

SINGLE AND MULTIPLE TIME–POINT ARTIFICIAL NEURAL NETWORKS MODELS FOR PREDICTING THE SURVIVAL OF GASTRIC CANCER PATIENTS

HAMID NILSAZ DEZFOULI

IPM 2016 15

SINGLE AND MULTIPLE TIME–POINT ARTIFICIAL NEURAL NETWORK MODELS FOR PREDICTING THE SURVIVAL OF GASTRIC CANCER PATIENTS

By

HAMID NIL SAZ DEZFOULI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

September 2016

DEDICATION

To my parents that I owe my life to them

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of requirement for the degree of Doctor of Philosophy

SINGLE AND MULTIPLE TIME–POINT ARTIFICIAL NEURAL NETWORK MODELS FOR PREDICTING THE SURVIVAL OF GASTRIC CANCER PATIENTS

By

HAMID NIL SAZ DEZFOULI

September 2016

Chairman : Mohd Rizam Abu Bakar, PhD
Institute : Mathematical Research

The extensive availability of recent computational models and data mining techniques for data analysis calls for researchers and practitioners in the medical field to opt for the most suitable strategies to confront clinical prediction problems. In many clinical research work, the main outcome under investigation is the time until an event occurs. Survival models are a collection of statistical procedures used to analyse data where the time until an event is of interest. Particularly the application of a data mining method known as 'neural networks' offers methodological and technical solutions to the problems of survival data analysis and prognostic model development.

In this context, artificial neural networks (ANN) have some advantages over conventional statistical tools, especially in the presence of complex prognostic relationships. ANN model applications for modeling the survival of gastric cancer patients have been highlighted in a number of studies but without a full account of censored survival data. The primary task under investigation in this thesis is to develop neural network methodologies for modeling gastric cancer survivability and fill the gap in the current literature by adopting strategies that directly incorporate censored observations in the process of constructing a neural network model. The dataset used in the study comprises of patients with confirmed gastric cancer who underwent surgery at the Cancer Registry Center of Taleghani Hospital, Tehran, Iran.

To achieve the research aims, single and multiple time-point ANN models are proposed. The first model is a single time-point ANN designed to predict the survival of patients at specific time points. The second is a multiple time-point model specifically designed to provide individualized survival predictions at different time points. Thus, an individual survival curve can be generated for a particular patient by plotting the survival probabilities produced by output units, which render the system more useful in clinical settings. The third model is a softmax ANN designed to estimate the unconditional

probability of death and predict the time period during which death is likely to occur for an individual patient. All models are extended to incorporate censored data. Employing the strategies for imputing the eventual outcome for censored patients has allowed all the available data to be used in developing an ANN predictor model. Several criteria are employed to validate the models. The research demonstrated how ANNs can be used in the survival analysis for predictive purposes without imposing any restricting assumptions. The proposed models provide accurate predictions of survival with high levels of sensitivity and specificity. Additionally, the sensitivity analysis provided information about the relative importance of each input variable in predicting the outcome. To sum up,

The ANN survival models presented in this thesis provide a framework for modelling survival data with censorship and facilitate individualized survival predictions. The findings will provide physicians and medical practitioners with information to improve gastric cancer prognosis and may assist in the selection of appropriate treatment plans for individual patients as well as efficient follow-up planning.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

MODEL RANGKAIAN NEURAL BUATAN TITIK MASA TUNGGAL DAN BERGANDA UNTUK MERAMAL MANDIRIAN PESAKIT KANSER GASTRIK

Oleh

HAMID NILSAZ DEZFOULI

September 2016

Pengerusi : Mohd Rizam Abu Bakar, PhD
Institut : Penyelidikan Matematik

Kemudahan secara meluas untuk mendapatkan model pengiraan dan teknik perlombongan data terkini untuk analisis data memerlukan penyelidik dan pengamal dalam bidang perubatan untuk memilih strategi yang paling sesuai dalam berhadapan dengan masalah ramalan klinikal. Dalam kebanyakan kerja-kerja penyelidikan klinikal, hasil utama dalam siasatan adalah masa sehingga berlakunya kejadian. Model mandirian atau *survival* adalah koleksi prosedur statistik yang digunakan untuk menganalisis data di mana masa sehingga peristiwa kejadian berlaku. Ini terutamanya aplikasi suatu kaedah perlombongan data yang dikenali sebagai 'rangkaian neural' yang menawarkan penyelesaian metodologi dan teknikal kepada masalah analisis data mandirian dan pembangunan model ramalan.

Dalam konteks ini, rangkaian neural tiruan (ANN) mempunyai beberapa kelebihan berbanding alat statistik konvensional, terutamanya dengan kehadiran hubungan ramalan yang kompleks. Aplikasi model ANN untuk model mandirian pesakit kanser perut telah ditonjolkan dalam beberapa kajian tetapi tanpa mengira sepenuhnya 'data mandirian' yang telah ditapis. Tugas utama dalam siasatan di dalam tesis ini adalah untuk membangunkan kaedah rangkaian neural untuk memodelkan mandirian kanser perut dan mengisi jurang dalam literatur semasa dengan menggunapakai strategi yang secara langsung menggabungkan tapisan pemerhatian dalam proses membina model rangkaian neural. Set data yang digunakan dalam kajian ini terdiri daripada pesakit yang disahkan mengidap kanser perut dan telah menjalani pembedahan di Pusat Pendaftaran Kanser, Hospital Taleghani, Tehran, Iran.

Untuk mencapai matlamat kajian, model titik masa tunggal dan berganda ANN telah dicadangkan. Model pertama adalah model ANN titik masa tunggal yang direka untuk meramalkan mandirian pesakit di titik masa tertentu. Yang kedua adalah model titik masa berganda yang direka khusus untuk menyediakan ramalan hidup individu di titik-

masa yang berbeza. Oleh itu, keluk mandirian individu boleh dihasilkan untuk pesakit tertentu dengan memplot kebarangkalian mandirian yang dihasilkan oleh unit output, yang menjadikan sistem tersebut lebih berguna dalam persekitaran klinikal. Model ketiga ialah softmax ANN yang direka untuk menganggarkan kebarangkalian tanpa syarat kematian dan meramal tempoh masa dalam mana kematian mungkin berlaku untuk individu pesakit. Semua model diperluas untuk merangkumi data yang telah ditapis. Dengan menggunakan strategi untuk mengandaikan keputusan akhir untuk pesakit yang telah ditapis, ia membenarkan semua data yang ada digunakan dalam membangunkan model peramal ANN. Beberapa kriteria yang digunakan adalah untuk menentukan keberkesanan model. Kajian menunjukkan bagaimana ANN boleh digunakan dalam analisis mandirian untuk tujuan ramalan tanpa mengenakan apa-apa andaian yang menjadi batasan. Model yang disarankan memberikan ramalan mandirian yang tepat dengan aras sensitiviti dan pengkhususan yang tinggi. Tambahan lagi, analisis sensitiviti memberi maklumat tentang kepentingan relatif setiap pembolehubah input dalam meramal hasilnya.

Kesimpulannya, model-model mandirian ANN yang dibentangkan dalam tesis ini menyediakan rangka kerja untuk data model mandirian yang ditapis dan membantu ramalan mandirian secara individu. Dalam konteks perubatan, maklumat ini amat berharga untuk kedua-dua doktor dan juga pesakit. Hasil kajian akan memberikan pakar-pakar dan pengamal perubatan dengan maklumat untuk meningkatkan prognosis kanser perut dan boleh membantu dalam pemilihan pelan rawatan untuk individu penyakit dan dengan perancangan susulan yang efisien.

ACKNOWLEDGEMENTS

In the Name of ALLAH the Most Merciful and Most Beneficent

I would like first and foremost to thank Allah for His endless blessings, mercy, and guidance. I want to express my deepest gratitude and sincerest appreciation to my supervisor, Associate Professor Dr. Hj. Mohd Rizam Abu Bakar, for his continuous support and guidance, who inspired me to become a better writer and a novice researcher. My thanks also go to my supervisory committee members, Associate Professor Dr.Jayanthi A/P Arasan, and Associate Professor Dr. Mohd Bakri Bin Adam, for their constructive questions, suggestions, and comments about this research. Special thanks go to Associate Professor Dr. Mohammad Amin Pourhoseingholi for the vivid discussions we have had about this research.

I would not have been able to complete this work without encouragement, sincere prayers, sympathy, understanding, and confidence of my beloved parents. Thank you for being such supportive parents, and for teaching me about the power of learning and knowledge. I thank them and I dedicate this work to them. Thanks are due to my colleagues and friends for their encouragement and practical support.

I would like to express deepest thanks to my friends, Amin Jula, Ali Mirsalehi, Roozbeh ALipour, Delaram Khansari, Ehsan Dehghanin, Ali Shekari and somehow helped me during school days. Thank you for supporting me, enlightening me, and most of all, you have made me what I am today. I love you, and thank God every day for you. I sincerely hope that this work would be beneficial to all who may need it.

I certify that a Thesis Examination Committee has met on 5 September 2016 to conduct the final examination of Hamid Nilsaz Dezfouli on his thesis entitled "Single and Multiple Time-Point Artificial Neural Network Models for Predicting the Survival of Gastric Cancer Patients" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Habshah binti Midi, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Noor Akma binti Ibrahim, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Yong Zulina Zubairi, PhD
Associate Professor
University of Malaya
Malaysia
(External Examiner)

M. Ataharul Islam, PhD
Professor
East West University
Bangladesh
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 28 February 2017

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Mohd Rizam B Abu Bakar, PhD
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Chairman)

Jayanthi A/P Arasan, PhD
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Member)

Mohd Bakri Bin Adam, PhD
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Member)

Mohammad Amin Pourhoseingholi, PhD
Associate Professor
Shahid Beheshti University Iran
(Member)

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books,journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules, or any other materials as stated in the universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No: Hamid Nil Saz Dezfouli , GS27581

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature:	
Name of Chairman of Supervisory Committee:	Associate Professor Dr. Mohd Rizam B Abu Bakar
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Jayanthi A/P Arasan
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Mohd Bakri Bin Adam
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Mohammad Amin Pourhoseingholi

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT	Accelerated Failure Time
ANN	Artificial Neural Network
ARD	Automatic Relevance Determination
BP	Back-Propagation
CE	Cross Entropy
CV	Cross Validation
C&RT	Classification & Regression Trees
PDF	Probability Density Function
FFANN	Feed Forward Artificial Neural Network
FN	False Negative
FP	False Positive
K-M	Kaplan-Meier
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NN	Neural Network
NPV	Negative Predictive Value
PH	Proportional Hazards
PL	Partial Likelihood
PLANN	Partial Logistic Artificial Neural Network
PPV	Positive Predictive Value
Rprop	Resilient Back Propagation
S.V.M	Support Vector Machine
TN	True Negative
TP	True Positive

CHAPTER 1

INTRODUCTION

1.1 Overview

Machine learning, particularly Artificial Neural Networks (ANNs), provide valuable tools for the diagnosis, prognosis, and detection of certain outcomes in healthcare research (Figure1.1). ANNs also offer a potentially useful means of developing classification and prognostic models for cancer patients and have been used in clinical decision support (Marchevsky, 2007). On hearing the terms 'machine learning' and 'clinical decision support,' one might imagine a patient walking into a doctor's office, inputting their symptoms and vital values into a terminal, dispensing a blood drop, and allowing a computer to announce the most likely illness and a suitable course of treatment, but this is not the point at all. The goal is to assist clinicians to determine the prognoses of patients, and to offer treatment specifically suited for each individual (Kalderstam, 2015). Clinical decision support could mean several things, but the aspect which is the focus of this thesis is survival prediction. Many cancer studies concern cancer prognosis with the aim of predicting outcomes, such as life expectancy, survivability, progression, and tumor-drug sensitivity, subsequent to the diagnosis (Cruz and Wishart, 2006).

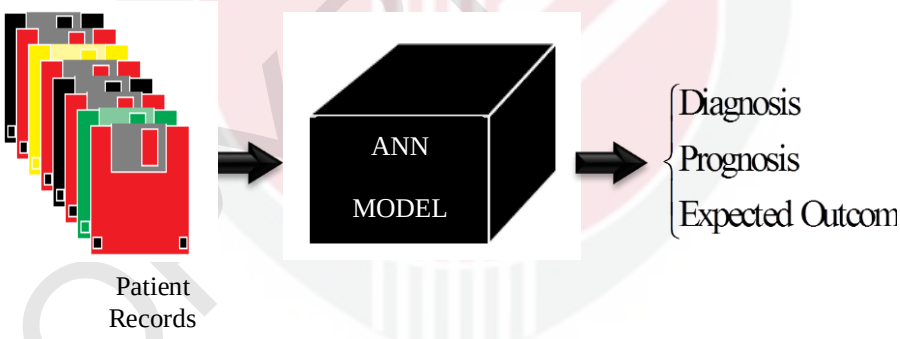

Figure 1.1 :ANN Applications in Health Care Research

Current statistical methods for survival analysis offer the possibility to model cancer survivability but require unrealistic assumptions about the survival time distribution or proportionality of hazard. As an alternative, the application of ANN models in survival analysis has drawn considerable interest (Bakker et al., 2004). In this regard, multilayer feed forward ANNs, also known as universal function approximators (Bishop, 2006), can overcome the proportionality and linearity constraints imposed by conventional survival analysis techniques. In addition, ANNs have the potential to provide more accurate prognosis models (Mani et al., 1999).

From a statistical viewpoint, the central challenge in developing an ANN model for cancer survivability is related to the presence of censoring (Crowther and Lambert, 2014). Censoring implies that for some patients, the time to the event of interest is not precisely known. This occurs either because a follow-up is no longer available to the patients for some reason, or simply because no significant event has yet occurred.

Such patients provide partial information about the disease survival characteristics, but are challenging to include in neural network models (Kalderstam, 2015). Disregarding censored patients not only results in disposing of a lot of information but also leads to a small and biased training dataset (Mani et al., 1999). In this regard, focus should be directed to developing ANN survival approaches that are able to handle censored observations. The current study investigates the feasibility of applying different ANN models to predict gastric cancer survivability, while the censored data are not omitted from the study. The results presented herein are from a retrospective study of patients with gastric cancer.

1.2 Background of the Study

Survival analysis is a group of statistical methods for data analysis, where the outcome variable is time until a certain event occurs (Lee and Wang, 2003). The event may be failure, injury, death, disease relapse, divorce, recovery or any other potential experience of interest to an individual. The time describes the distance from the beginning of follow-up until an event occurs. In survival analysis, the generic name for time is survival time (Clark et al., 2003). For the data set employed in this thesis, the endpoint of interest is typically death.

The specific difficulties concerning survival data analysis is that survival data are usually censored or incomplete in some way. Censoring occurs when incomplete information is available about some individuals' survival time. 'Right censored' data occurs when an individual experiences failure time after its final observed time. Right censoring means that the survival time is only known to exceed a certain value.

Censoring may arise in any of the following ways:

- A patient has not experienced the event by the end of the study period;
- A patient has lost to follow up at any time during the study period;
- A patient experienced a different event, for example he exited the study for reasons other than cancer-related death, making further follow-ups impossible (Clark et al., 2003).

It may also be unclear when the patients entered the study. For instance, for patients infected with HIV, the date of infection is usually unclear. Data from these patients are considered left censored. Moreover, event time data may also be interval censored. It means that the event time lies within an interval and is not precisely observed (Clark et al, 2003).

The data points in this thesis are considered to be right censored because the exact survival times of the patients are unknown, but it is known that each patient's time of death will occur after a specified time point. From a statistical perspective, standard methods used for survival analysis are valid only if the censoring is non-informative, e.g. random and uncorrelated with the true survival time (Cook, 2007). Non-informative censoring is assumed in the present study. In practical terms, it is assumed that censored patients should be as likely to subsequently die as the patients who remain in the study (Kleinbaum and Klein, 2005).

1.2.1 General Formulation

The distribution of survival times is usually described using the survival function, probability density function, and hazard function. Let T be a non-negative random variable that denotes the time elapsed from a particular starting time point to the occurrence of an event. A small t denotes any specific value for variable T. The probability density function is explained as the probability of failure in a small interval $(t, t+\Delta t)$ per unit time.

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < T + \Delta t)}{\Delta t} \tag{1.1}$$

The survival function, also called the survivor or survivorship function, is a basic quantity employed in survival analysis (Brandon et al., 2014). The survival function $S(t)$ is defined as the probability that an individual will survive beyond a specific time t. Since T is a continuous random variable, the survival function can be obtained by integrating it over the probability density function from time *t* to infinity

$$S(t) = P(T > t) = \int_{t}^{+\infty} f(x)dx \tag{1.2}$$

The hazard function $h(t)$ gives the conditional failure rate and is defined as the probability of failure during a very short interval $(t, t+\Delta t)$, assuming that the individual has survived until time t.

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t \mid T \ge t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{1.3}$$

The hazard function represents the instantaneous potential per unit time for an event to occur given that the patient has survived up to time *t* (Kleinbaum and Klein, 2005)

1.2.2 Methods of Survival Data Analysis

The survival probability can be estimated non-parametrically using the Kaplan-Meier (K-M), or product-limit estimator (Kaplan and Meier, 1958). The K-M estimator utilizes the information of a censored case until the patient becomes censored. Suppose the event times are ordered as $t_1, t_2, ..., t_n$. The general formula for the K-M survival probability at event time t_i is given by:

$$\hat{S}(t_i) = \hat{S}(t_{i-1})(1 - \frac{d_i}{n_i}) \tag{1.4}$$

where $t_0 = 0$ and $S(0) = 1$. Thus, the probability of being alive at time t_i is calculated from $\hat{S}(t_{i-1})$ the probability of being alive at t_{i-1}, d_j the number of events at time t_i, and n_i the total number of patients at risk just before time t_i.

Product-limit estimators perform well in describing the survival of a group of individuals. However, when predicting survival or comparing treatments in terms of survival, it is often rational to adjust for some patient-related factors known as covariates, which could potentially affect patient survival time. One drawback of the K-M method is that the K-M estimator only describes survival with respect to the factor under investigation and ignores the impact of any other covariates with potential effect on patient prognosis.

Parametric survival models are a premise for a particular form of survival distribution. For parametric models the functional form is completely specified except for the unknown parameters' values. The data are then used to estimates the model parameters that fully specify that distribution. The parameters are estimated using maximum likelihood estimation (MLE). The most commonly employed parametric models in survival data analysis are the exponential, Weibull, lognormal, and log-logistic, which get their names from the distribution that the survival times are assumed to follow (Kleinbaum and Klein, 2005).

Cox (1972) suggested the proportional hazard (PH) model for analyzing survival data. This model does not require knowledge of the underlying distribution. The Cox proportional hazard model is a multivariate regression method that describes the relationship between the occurrence of an event as explained by a hazard function and a set of explanatory variables $x = (x_1, x_2, ..., x_p)$. Mathematically, the Cox model is presented as:

$$h(t \mid x) = h_0(t) \exp(\beta^T x) \tag{1.5}$$

where $\beta = (\beta_1, \beta_2, ..., \beta_p)$ denote the coefficients of covariates. The hazard at time t depends on two quantities. The first quantity, $h_0(t)$, is called the baseline hazard function and it describes the risk for an individual with covariate vector $x = 0$. The second quantity is $\exp(\beta^T x)$, which is an exponential expression involving covariate x. The baseline hazard can be any complicated function of t and is estimated non-parametrically. The Cox proportional hazard model assumes that the hazard ratio of two patients with different covariate values is independent of time. The second assumption is that the impact of the covariates on the log risk scale is additive and linear. However, these assumptions are not applicable or not thoroughly checked in different clinical conditions where a Cox model is used to fit survival data (Jerez et al., 2005).

There are other techniques that have been applied in survival data analysis. Local fitting methods, spline function, tree-based methods, and artificial neural networks are the four major non-linear approaches used for analyzing survival data (Crowther and Lambert, 2014; Van Belle et al., 2011; Bakker et al., 2004; Ripley et al., 2004). Among all methods, neural networks are proven to integrate the consistency of conventional models and offer a relatively parsimonious framework compared to the other models (Biganzoli et al., 2003).

1.2.3 Neural Network Models in Survival Analysis

The artificial neural network is an approach based on a very simple and abstract version of how the biological neural networks in our brains function. Since ancient Greek times, there has been long-standing interest in understanding "intelligence" (Russell and Norvig, 1995). In 1943, McCulloch and Pitts devised a simple neuron model and described how neurons might work (Bishop, 2006). They also provided a model of a simple artificial neural network with electrical circuits. In 1949, Donald Hebb described several fundamental concepts of artificial neurons and their behavior.Neural network (NN) applications are deliberated into the three following categories (Jones, 2004):

- Predicting one or more outcomes from input data
- Classifying input data into one or more categories
- Statistical pattern recognition for uncovering patterns among a set of variables

Therefore, neural networks provide a framework for solving several prevalent traditional problems of prediction, classification and pattern recognition. In some situations, they also extend the range of problems that can be solved by conventional methods.

A neural network is comprised of a collection of neurons. Each neuron has a number of input and output connections to other neurons, called synapses. Biological neurons send signals to other neurons by "firing" electrical impulses via these connections. Multiple impulses from different neurons may amplify or dampen each other, depending on the synapses. A simplified biological neuron structure is illustrated on the left of Figure 1.2. Information enters the biological neuron via dendrites, and the cell body (soma) processes the information and expresses it via an axon.

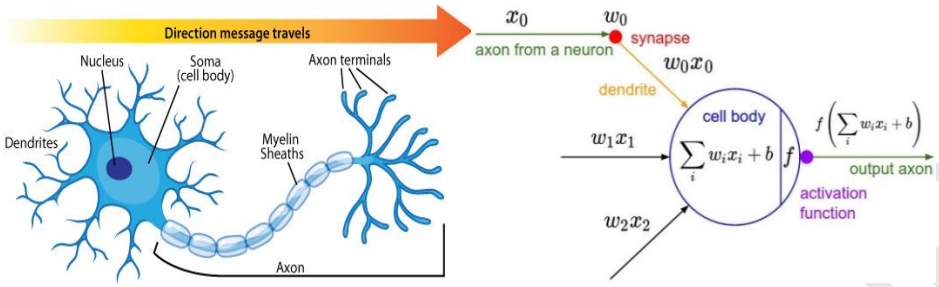

Figure 1.2 : Biological and Artificial Neuron Designs

A simplified model of an artificial neuron, called a perceptron (Rosenblatt, 1958), is illustrated on the right of Figure 1.2. The information enters the artificial neuron via inputs denoted by X_i and is multiplied by corresponding weights w_i. An artificial neuron also contains a bias, b, which is an external input for adjusting the net input of the activation function. The artificial neuron body sums the weighted inputs and biases, and then processes the sum with an activation function, f. The processed information then goes through the neuron output expressed by O. The output of a neuron can be formulated as:

$$o = f\left(b + \sum_i w_i x_i\right) \tag{1.6}$$

The activation function defines the neuron's output level for a given input. There are three general types of activation functions: identity, threshold and sigmoid transfer functions, which are described in Chapter 2. Among various ANN architectures, Multilayer perceptron (MLP) has been the most extensively used method for cancer prediction and prognosis (Ahmed, 2005; Schwarze et al., 2000)The effectiveness of MLPs in cancer diagnosis and prognosis has been evaluated using clinical, pathological and immunohistochemical data, which signify that MLP is a powerful cancer prediction technique.

MLP learning process is performed by adjusting the connection weights linking the layers. This is done by a training algorithm that adjusts the weights by minimizing the network error with respect to its weights. Back-propagation (BP) is the most widely implemented training algorithm for MLP training (An et al., 2015; Haykin et al., 2009)

Censoring in the data is the main reason it is difficult to use standard ANN techniques to model survival (Stajduhar and Dalbelo-Basic, 2010). Several approaches have been proposed for dealing with censored data in an ANN model. As such, various efforts have

been devoted to extending the Cox proportional hazard model. An ANN extension of the Cox proportional hazard model was proposed by Faraggi and Simon (1995). They developed the maximum likelihood neural network model for the general classification problem by replacing the linear predictor $\beta^T X$ in the Cox PH model with a non-linear network output. Other scholars (Bakker et al., 2004; Bakker and Heskes, 1999; Lisboa and Wong, 2001; Mariani et al., 1997) have also provided ANN models as extensions of the Cox proportional hazard model.

Hierarchical and modular ANN survival models have been introduced as well (Lapuerta et al., 1995; Ohno-Machado et al., 1995; Ohno-Machado, 1997). In these models, several neural networks were employed, where each neural network predicts the survival at a specific time point. In another form of ANN models, prognostic covariates are considered input variables, whereas the time to an event serves as the neural network output (Brown et al., 1997; Laurentiis et al., 1999; Jerez et al., 2003, Jerez et al., 2005; Zhu et al., 2013).

Closely related to this approach are the so-called "time coded models." In this ANN structure, time is added as a covariate and the model output indicates the event or no event at a given time (Ravdin and Clark, 1992; De Laurentiis and Ravdin, 1994; Boracchi, Biganzoli, and Marubini, 2001; Liestol and Andersen, 2002; Biganzoli et al., 2003; Biganzoli and Boracchi, 2009). The output of these models can be interpreted as conditional or cumulative probabilities depending on the structure considered for input data. These methods are explained in more in detail in Chapter 2.

1.3 Motivation

Predicting cancer outcome based on a set of prognostic variables has been a long-standing topic of interest in cancer studies. Among the different cancer types, gastric cancer accounts for considerable morbidity and mortality levels worldwide (McLean and El-Omar, 2014; Ferlay et al., 2013). The findings of the International Agency for Research on Cancer in 2012 ranked gastric cancer as the fifth most widespread malignancy in the world after lung, breast, colorectal, and prostate cancers. Furthermore, gastric cancer is the third leading cause of cancer death worldwide (McLean and El-Omar, 2014). Given these facts, gastric cancer represents one of the most preferred fields of investigation. Nonetheless, a comparison between the number of published papers and cancer incidence (Lisboa and Taktak, 2006) has revealed that there is a less-than-expected proportion of publications on gastric cancer, arguably signifying the need for greater consideration and investigation in this field.

Existing survival analysis methods, as described in Section 1.1, are usually utilized to explain data or determine the predictive value of different variables in the progression of disease rather than make predictions for individual patients. Parametric models are based on strict assumptions concerning the distribution of failure time and the effect of the covariates on the distribution parameters. These assumptions are mostly not feasible in applied situations (Eleuteri et al., 2007). In the domain of gastric cancer, the most commonly used method is the Cox proportional hazards model when the task is to define which variables influence survival. However, there are still a number of assumptions that

need to be assessed before applying the Cox model. Violations of the underlying assumptions invalidate the results (Bewick et al., 2004). Additionally, some of the assumptions are not carefully assessed in many different clinical conditions where the Cox model is applied for survival data analysis (George et al., 2014; Jerez et al., 2005).

To address these limitations, attention must be paid to the development of nonlinear models with less restrictive assumptions. Models based on ANN appear to be suited for this task and have been successfully employed in the field of medical diagnosis and prognosis. The usefulness of the ANN methodology is justified by the fact that ANNs do not assume a certain prior functional form and do not necessitate fulfilling the assumptions required by statistical techniques. In other words, the ANN approach is driven and confined by data in hand. The mathematical structure of neural networks enables them to analyze complex data with non-linear covariates, high-order interaction among covariates, and time-dependent covariates (Lisboa and Taktak, 2006). These characteristics have motivated us to develop prognostic models through survival analysis.

The application of ANN models to survival analysis also has implications from a biological point of view, as highlighted by Jerez et al. (2005) since the relationship among prognostic covariates and patients' outcomes is not necessarily linear in nature. Thus, the traditional methods of survival analysis, which rely utterly on the linear relationship among variables, may be inadequate.

Predicting the probability of survival for a patient can be very challenging for many diseases. Developing better clinical decision support systems for gastric cancer prognosis could decrease uncertainty in prognosis, allowing treatment to be focused on patients with the worst expected survival chances (Kalderstam, 2015). Neural networks can provide individualized survival predictions. In a medical context, such information is valuable for both clinicians and patients. As mentioned before, NNs assist clinicians to select appropriate treatments and plan follow-ups efficiently. Patients at high risk could be followed up more frequently than those at lower risk, such that valuable resources are channeled to those who need them the most. For patients, obtaining information about their prognosis is also valuable for planning their lives (Eleuteri et al., 2007).

1.4 Problem Statement

In recent years, considerable attention has been directed to the application of ANN-based methods for developing prognostic models in medicine. ANNs have been used in diagnosis, prognosis, and outcome prediction in numerous cancer research works. However, predicting the probability of survival or disease outcome for an individual patient remains a challenging task for many diseases.

A prominent analytical feature in most survival analysis studies pertains to censoring, where survival times are not precisely determined for some patients. This may happen during the follow-up period when some patients leave the study for various reasons like

accident death. Others may survive the study without cancer recurring or death. What complicates matters is that the standard structure of a neural network does not allow direct modeling of censored survival data. Thus, the presence of this certain characteristic of medical data makes it difficult to use ANN methods. This implies that merging the survival analysis theory with ANN methodology requires introducing some strategies to overcome those difficulties and deal with censored data. One approach to handling this is to simply discard all censored patients and train the ANN model directly with the remaining non-censored data, but this would introduce bias in the model.

A lot of research has been conducted on the application of ANN structures in gastric cancer survival prediction (Zhu et al., 2013; Amiri et al., 2013; Gohari et al., 2011; Biglarian et al., 2011). Nevertheless, no work has been done on constructing different ANN strategies for modeling censored survival data. A review of some prominent studies in the field of gastric cancer revealed that scarce attention has been given to the problem of censoring in these studies. A literature review of ANN models applied for gastric cancer prognosis demonstrated that in many studies, censored patients have been excluded from the dataset or no clear strategies have been addressed for dealing with censoring (Zhu et al., 2013; Amiri et al., 2008, 2013; Gohari et al., 2011; Biglarian et al., 2011; Baghestani et al., 2009; Qiu et al., 2009).

This gap encouraged the researcher to focus on the development of specified ANN models for modeling gastric cancer survivability and making personalized survival predictions in the presence of censored data. This can be done by either modifying the standard way a dataset is presented to a standard neural network or by proposing some strategies that directly incorporate the censored observations in a neural network model. This would assist physicians and medical scientists to improve the clinical care and management of gastric cancer.

1.5 Objectives of the Study

The primary aim of this study is to make a new contribution to the development of single and multiple time-point ANNs for gastric cancer prognosis in terms of predicting the outcomes of patients. The current study also intends to fill the gap in comparative studies by developing strategies that directly incorporate censored observations in the process of constructing a survival analysis neural network model. More specifically, the objectives are listed as follows:

- To propose a specifically designed single time-point ANN method for modeling gastric cancer survivability, which is efficient in predicting the outcome at specific time points

- To propose multiple time-point ANN models that are able to predict the probability of survival in different time periods and generate an individualized survival curve for every patient in the data set

- To develop an ANN model that are able to estimate unconditional probability of death at different intervals and predict the time period during which a death is likely to occur for an individual patient

- To propose strategies for incorporating censored data in constructing ANN-based survival models rather than simply excluding censored data

- To determine the predictive value of different variables for the progression of gastric cancer and to define an effective combination of prognostic covariates and tumor markers for predicting the outcome of gastric cancer patients
- To analyze the data using the KM and Cox models and compare the ANN results with standard methods

1.6 Scope of the Study

The focus of this thesis is on developing ANNs for modeling gastric cancer survival. More specifically, the aim is to develop ANN-based models that are able to incorporate censored survival data instead of merely excluding these data from the study. Survival analysis can be viewed as a classification problem by establishing meaningful intervals of time according to a particular situation. Focus will be restricted to constructing ANN classification models that predict the probability of events occurring during one or more fixed time intervals. In other words, survival analysis is considered as a classification problem in this thesis. The endpoint of interest in this study is typically death. In this case, the outputs of the ANN classification models correspond to predetermined intervals of time and a prognostic estimate will be produced for each interval.

Real datasets of gastric cancer survival with a significant number of cases are not easy to obtain. Although the use of artificial data sets facilitates control of data, the results might not be generalizable to real data sets, which contain noise, are incomplete and have few cases (Ohno-Machado, 1997). The use of real data sets provides more useable and convincing information. The data set in the present study was derived from a retrospective study on patients with confirmed gastric cancer who underwent total or subtotal gastrectomy at the Cancer Registry Center of Taleghani Hospital, Tehran, Iran. All patients were diagnosed by endoscopy and/or biopsy and their disease was confirmed. In our data set, censored observations and missing data are frequently presented and hard to control. However, a good survival predictor must be able to deal with these obstacles.

1.7 Organization of the Remainder of the Study

The overall organization of this thesis is as follows: After the introduction, chapter 2 is dedicated to a review of the statistical and ANN methods previously employed for survival analysis. A literature of the most widely used ANN structures in gastric cancer prognosis along with the advantages and disadvantages of these methods is also detailed. A general introduction of gastric cancer, its risk factors, and stages of gastric cancer is also described.

Chapter 3 introduces a single time-point ANN model for modeling the survival of gastric cancer patients that is suitable for prediction at specific time points. The method of imputing the outcome to censored patients is also described. This chapter discusses the data set characteristics and prognostic variables used in the study. The final section of this chapter includes an analysis of the importance of prognostic covariates to demonstrate the degree of significance of each covariate for predicting the survival of patients.

Chapter 4 presents the proposed multiple time-point ANN model, which predicts the probability of survival at different time intervals for a patient with gastric cancer. This model generates a set of survival probabilities across all time periods and can predict any time-specific survival rate. Thus, for a particular patient, an individual survival curve can be generated by plotting the survival probabilities predicted by output units.

The ANN model presented in chapter 5 predicts how long after surgery a patient is expected to die. The model was designed to estimate unconditional probability of death in time intervals of less than one year, one to two years, two to three years, three to four years, four to five years, and greater than 5 years. More specifically, we have modeled the unconditional probability of death using a softmax ANN model.

Chapter 6 presents the analysis of gastric cancer dataset using Cox model and Kaplan-Meier method. The rest of the chapter compares and contrasts the results obtained by Kaplan-Meier and Cox proportional hazard analysis with those obtained by the Cox model.

Finally, Chapter 7 contains a brief summary of the main findings and the thesis conclusions, the consequences of the findings from this study and some directions for further research in the future

REFERENCES

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Ahmed, F. E. (2005). Artificial neural networks for diagnosis and survival prediction in colon cancer. *Molecular Cancer*, 4(1): 29.

Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine*, 11(2): 47–58.

Amiri, Z., Mohammad, K., Mahmoudi, M., Parsaeian, M., & Zeraati, H. (2013). Assessing the effect of quantitative and qualitative predictors on gastric cancer individuals survival using hierarchical artificial neural network models. *Iranian Red Crescent Medical Journal*, 15(1): 42–8.

Amiri, Z., Mohammad, K., Mahmoudi, M., Zeraati, H., & Fotouhi, A. (2008). Assessment of gastric cancer survival: using an artificial hierarchical neural network. *Pak J Biol Sci*, 11(8): 1076–1084.

An, D., Kim, N. H., & Choi, J.-H. (2015). Statistical aspects in neural network for the purpose of prognostics. *Journal of Mechanical Science and Technology*, 29(4): 1369–1375.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The Annals of Statistics*, 10(4): 1100–1120.

Anderson, J. A., & Rosenfeld, E. (2000). *Talking nets: An oral history of neural networks*. Cambridge, Massachusetts: MIT Press.

Andersson, T. M. L., Dickman, P. W., Eloranta, S., & Lambert, P. C. (2011). Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models, *BMC Medical Research Methodology*, 11(1): 96.

Antoniou, A., & Lu, W.-S. (2007). *Practical optimization: algorithms and engineering applications*. British Columbia, Canada: Springer Science & Business Media.

Ardoino, I., Ambrogi, F., Bajdik, C., Lisboa, P. J., Biganzoli, E. M., & Boracchi, P. (2010). Flexible parametric modelling of the hazard function in breast cancer studies. In *Proceedings of the International Joint Conference on Neural Networks*, Barcelona, Spain, July 2010.

Arsene, C. T. C., & Lisboa, P. J. (2012). Bayesian Neural Network Applied in Medical Survival Analysis of Primary Biliary Cirrhosis. *2012 UKSim 14th International Conference on Computer Modelling and Simulation*, Emmanuel College, Cambridge, United Kingdom.

Arsene, C. T. C., Lisboa, P. J., & Biganzoli, E. (2011). Model Selection with PLANN-CR-ARD. In J. Cabestany, I. Rojas, & G. Joya, *Advances in Computational*

Intelligence (pp. 210–219). Berlin, Heidelberge: Springer.

Arsene, C. T. C., & Lisboa, P. J. G. (2007). Artificial Neural Networks Used in the Survival Analysis of Breast Cancer Patients : A Node-Negative Study. In A. F. G. Taktak, & A. C. Fisher (Eds.), *Outcome Prediction in Cancer*, (pp. 191-239). Oxford, UK: Elsevier.

Baesens, B., Van Gestel, T., Stepanova, M., Van den Poel, D., & Vanthienen, J. (2005). Neural network survival analysis for personal loan data. *Journal of the Operational Research Society*, 56(9): 1089–1098.

Baghestani, A. R., Hajizadeh, E., & Fatemi, S. R. (2009). Bayesian analysis for survival of patients with gastric cancer in Iran. *Asian Pacific Journal of Cancer Prevention: APJCP*, 10(5): 823–826.

Bakker, B., & Heskes, T. (1999). A neural-Bayesian approach to survival analysis. *The 9th International Conference on Artificial Neural Networks: ICANN*, Edinburgh: IET.

Bakker, B., Heskes, T., Neijt, J., & Kappen, B. (2004). Improving Cox survival analysis with a neural-Bayesian approach. *Statistics in Medicine*, 23(19): 2989–3012.

Bartosch-Härlid, A., Andersson, B., Aho, U., Nilsson, J., & Andersson, R. (2008). Artificial neural networks in pancreatic disease. *British Journal of Surgery*, 95(7): 817–826.

Barwad, A., Dey, P., & Susheilia, S. (2012). Artificial neural network in diagnosis of metastatic carcinoma in effusion cytology. *Cytometry Part B: Clinical Cytometry*, 82(2): 107–111.

Baxt, W. G. (1990). Use of an artificial neural network for data analysis in clinical decision-making: the diagnosis of acute coronary occlusion. *Neural Computation*, 2(4): 480–489.

Baxt, W. G. (1991). Use of an artificial neural network for the diagnosis of myocardial infarction. *Annals of Internal Medicine*, 115(11): 843–848.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2010). Neural Network Toolbox 7. *User's Guide*, Natick, MA: The *MathWorks* Inc.

Beigi, H. S. M., & Li, C. J. (1993). Learning algorithms for neural networks based on quasi-Newton methods with self-scaling. *Journal of Dynamic Systems, Measurement, and Control*, 115(1): 38–43.

Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: Current issues and guidelines. *International Journal of Medical Informatics*, 77(2): 81–97.

Bergmeir, C., & Benitez, J. M. (2012). Neural Networks in R Using the Stuttgart Neural Network Simulator: RSNNS. *Journal of Statistical Software*, 46(7): 1–26.

Bernstein, I. H. (2012). *Applied multivariate analysis*. New York: Springer Science &

Business Media.

Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 12: survival analysis. *Critical Care-London*, 8(5): 389–394.

Biganzoli, E., Boracchi, P., Coradini, D., Daidone, M. G., & Marubini, E. (2003). Prognosis in node-negative primary breast cancer: A neural network analysis of risk profiles using routinely assessed factors. *Annals of Oncology*, 14(10): 1484–1493.

Biganzoli, E., Boracchi, P., Mariani, L., & Marubini, E. (1998). Feed forward neural networks for the analysis of censored survival data: A partial logistic regression approach. *Statistics in Medicine*, 17(10): 1169–1186.

Biganzoli, E., Boracchi, P., & Marubini, E. (2002). A general framework for neural network models on censored survival data. *Neural Networks*, 15(2): 209–218.

Biganzoli, E. M., & Boracchi, P. (2009). The partial logistic artificial neural network (plann): a tool for the flexible modelling of censored survival data. Proceedings of *the 2009 international joint conference on Neural Networks*, USA: IEEE press.

Biglarian, A., Hajizadeh, E., Kazemnejad, A., & Zali, M. R. (2011). Application of Artificial Neural Network in Predicting the Survival Rate of Gastric Cancer Patients. *Iranian Journal of Public Health*, 40(2): 80–86.

Biglarian, A., Hajizadeh, E., Kazemnejad, A., & Zayeri, F. (2010). Determining of prognostic factors in gastric cancer patients using artificial neural networks. *Asian Pacific Journal of Cancer Prevention, : APJCP*, 11(2): 533–536.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Cambridge, UK: springer.

Boracchi, P., Biganzoli, E., & Marubini, E. (2001). Modelling cause-specific hazards with radial basis function artificial neural networks: application to 2233 breast cancer patients. *Statistics in Medicine*, 20(24): 3677–3694.

Bouhlila, D. S., & Sellaouti, F. (2013). Multiple imputation using chained equations for missing data in TIMSS: a case study. *Large-Scale Assessments in Education*, 1(1): 1–33.

Bourdès, V., Bonnevay, S., Lisboa, P., Defrance, R., Pérol, D., Chabaud, S., Bachelot, T., Gargi, T., Négrier, S. (2010). Comparison of Artificial Neural Network with Logistic Regression as Classification Models for Variable Selection for Prediction of Breast Cancer Patient Outcomes. *Advances in Artificial Neural Systems*, Volume 2010: 11.

Bozzetti, F., Marubini, E., Bonfanti, G., Miceli, R., Piano, C., Gennari, L., & Group, I. G. T. S. (1999). Subtotal versus total gastrectomy for gastric cancer: five-year survival rates in a multicenter randomized Italian trial. *Annals of Surgery*, 230(2): 170.

Bradburn, M. J., Clark, T. G., Love, S. B., & Altman, D. G. (2003). Survival analysis

part II: multivariate data analysis--an introduction to concepts and methods. *British Journal of Cancer*, 89(3): 431–6.

George B, Seals S, Aban I. Survival analysis and regression models. *J Nucl Cardiol*. 2014;21(4):686-694.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth Statistics/Probability, USA: Chapman and Hall/CRC.

Brodersen, K. H., Mathys, C., Chumbley, J. R., Daunizeau, J., Ong, C. S., Buhmann, J. M., & Stephan, K. E. (2012). Bayesian mixed-effects inference on classification performance in hierarchical data sets. *The Journal of Machine Learning Research*, 13(1): 3133–3176.

Brown, S. F., Branford, A. J., & Moran, W. (1997). On the use of artificial neural networks for the analysis of survival data. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, 8(5): 1071–1077.

Calik, M., Calik, I., Demirci, E., Altun, E., Gundogdu, B., Sipal, S., & Gundogdu, C. (2013). Goseki grade and tumour location influence survival of patients with gastric cancer. *Asian Pacific Journal of Cancer Prevention: APJCP*, 15(3): 1429–1434.

Carl-mcgrath, S., Ebert, M., & Röcken, C. (2007). Gastric adenocarcinoma : epidemiology, pathology and pathogenesis Review Article. *Cancer Therapy*, 5: 877–894.

Carrillo, H., Brodersen, K. H., & Castellanos, J. A. (2014). Probabilistic Performance Evaluation for Multiclass Classification Using the Posterior Balanced Accuracy. In *ROBOT2013: First Iberian Robotics Conference*. Universidad de Zaragoza, Spain: Springer.

Catalogna, M., Cohen, E., Fishman, S., Halpern, Z., Nevo, U., & Ben-Jacob, E. (2012). Artificial neural networks based controller for glucose monitoring during clamp test. *PloS One*, 7(8): e44587.

Cervellera, C., Macciò, D., & Muselli, M. (2008). Deterministic learning for maximum-likelihood estimation through neural networks. *IEEE Transactions on Neural Networks*, 19(8): 1456–1467.

Chi, C., Street, W. N., & Wolberg, W. H. (2007). Application of artificial neural network-based survival analysis on two breast cancer datasets. *Annual Symposium Proceedings / AMIA Symposium*, Volume 2007: 130-134.

Chien, C.-W., Lee, Y.-C., Ma, T., Lee, T.-S., Lin, Y.-C., Wang, W., & Lee, W.-J. (2007). The application of artificial neural networks and decision tree model in predicting post-operative complication for gastric cancer patients. *Hepato-Gastroenterology*, 55(84): 1140–1145.

Cho, Y., Lee, D. H., Oh, H. S., Seo, J. Y., Lee, D. H., Kim, N., … Park, Y. S. (2012). Higher prevalence of obesity in gastric cardia adenocarcinoma compared to gastric non-cardia adenocarcinoma. *Digestive Diseases and Sciences*, 57(10): 2687–2692.

Clark, T. G., Bradburn, M. J., Love, S. B., & Altman, D. G. (2003). Survival analysis part I: basic concepts and first analyses. *British Journal of Cancer*, 89(2): 232–238.

Colzani, E., Liljegren, A., Johansson, A. L. V, Adolfsson, J., Hellborg, H., Hall, P. F. L., & Czene, K. (2011). Prognosis of patients with breast cancer: causes of death and effects of time since diagnosis, age, and tumor characteristics. *Journal of Clinical Oncology*, 29: 4014-4021.

Cook, N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 115(7): 928–935.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34 (2): 187–220.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data*. USA: Chapman and Hall/CRC.

Crawley, M. J. (2007). *The R Book*. UK: John Wiley & Sons.

Crew, K. D., & Neugut, A. I. (2006). Epidemiology of gastric cancer. *World Journal of Gastroenterology: WJG*, 12(3): 354–362.

Crowther, M. J., & Lambert, P. C. (2014). A general framework for parametric survival analysis. *Statistics in Medicine*, 33(30): 5280–5297.

Cruz, J. A., & Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics*, 2: 59.

De Laurentiis, M., & Ravdin, P. M. (1994). A technique for using neural network analysis to perform survival analysis of censored data. *Cancer Letters*, 77(2): 127–138.

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine*, 34(2): 113–127.

Dhande, M. J. D., & Gulhane, S. M. (2013). Performance Evaluation of Artificial Neural Network Classifier Based On Receiver Operating Characteristic (ROC) Curves, *International Journal of Emerging Trends in Electrical and Electronics*, 3(2): 41–44.

Dolgobrodov, S. G., Moore, P., Marshall, R., Bittern, R., Steele, R. J. C., & Cuschieri, A. (2007). Artificial neural network: Predicted vs. observed survival in patients with colonic cancer. *Diseases of the Colon and Rectum*, 50(2): 184–191.

B. Dolhansky. (2013, September 27). Artificial Neural Networks : Linear Multiclass Classification. [Web log comment]. Retrived from http://briandolhansky.com

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, 35(5-6): 352–359.

Dreyfus, G. (2005). *Neural networks: methodology and applications*. Berlin, Heidelberg: Springer Science & Business Media.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, 72(359): 557–565.

Eleuteri, A., Aung, M. S. H., Taktak, A. F. G., Damato, B., & Lisboa, P. J. G. (2007). Continuous and discrete time survival analysis: neural network approaches. Proceedings from *the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS).* IEEE.

Er, O., Temurtas, F., & Tanrıkulu, A. Ç. (2010). Tuberculosis disease diagnosis using artificial neural networks. *Journal of Medical Systems*, 34(3): 299–302.

Faraggi, D., & Simon, R. (1995). A neural network model for survival data. *Statistics in Medicine*, 14(1): 73–82.

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W. W., Comber, H., … Bray, F. (2013). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *European Journal of Cancer*, 49(6): 1374–1403.

Fernandes, A. S., Jarman, I. H., Etchells, T., Fonseca, J. M., Biganzoli, E., Bajdik, C., & Lisboa, P. J. G. (2008). Missing Data Imputation in Longitudinal Cohort Studies: Application of PLANN-ARD in Breast Cancer Survival. Proceedings from *the seventh international conference on machine learning and applications (ICMLA)*. California, USA: IEEE.

Ferri, C., Hernández-Orallo, J., & Salido, M. A. (2003). Volume under the ROC surface for multi-class problems. In N. Lavrac et al. (Eds), *Machine Learning: ECML 2003* (pp. 108–120). Berlin, Heidelberg: Springer.

Fletcher, R. (2013). *Practical methods of optimization*. Padstow, Cornwall: John Wiley & Sons.

Forman, D., & Burley, V. J. (2006). Gastric cancer: global pattern of the disease and an overview of environmental risk factors. *Best Practice & Research Clinical Gastroenterology*, 20(4): 633–649.

Fox, J. (2002). Cox Proportional-Hazards Regression for Survival Data. Appendix to *An R and S-PLUS Companion to Applied Regression*, 1-18.

K. Ganesan. (2015). Text Mining , Analytics & More. [Web log comment]. Retrived from http://www.text-analytics101.com

Garson, D. G. (1991). Interpreting neural network connection weights. *AI Expert*, 6(7): 47-51.

George, B., Seals, S., & Aban, I. (2014). Survival analysis and regression models. *Journal of Nuclear Cardiology*, 21(4): 686–694.

Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to

study the contribution of variables in artificial neural network models. *Ecological Modelling*, 160(3): 249–264.

Gill, P. E., Murray, W., & Wright, M. H. (1981). *Practical optimization*. UK: Emerald

Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3): 143–151.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3): 515–526.

Graupe, D. (2007). *Principles of artificial neural networks.* University of lllinois, Chicago, USA: World Scientific.

Grossi, E., & Buscema, M. (2007). Introduction to artificial neural networks. *European Journal of Gastroenterology & Hepatology*, *19*(12): 1046–1054.

Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *Neural Networks, IEEE Transactions on*, 5(6): 989–993.

Hagan, M.T., Demuth, H.B., and Beale, M. H. (1996). *Neural Network Design. Boston, USA*: PWS Publishing Company.

Hand, D. J., & Till, R. J. (2001). A simple generalization of the area under the ROC curve to multiple class classification problems. *Machine Learning*, 45: 171–186.

Harikumar, R., & Vinoth kumar, B. (2015). Performance analysis of neural networks for classification of medical images with wavelets as a feature extractor. *International Journal of Imaging Systems and Technology*, 25(1): 33–40.

Haykin, S. S., Haykin, S. S., Haykin, S. S., & Haykin, S. S. (2009). *Neural networks and learning machines*, Upper Saddle River: Pearson Education.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological approach*. New York: John Wiley & Sons.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. Hoboken, New Jersey: John Wiley & Sons.

Huang, X., Acero, A., Chelba, C., Deng, L., ….& Wang, Y. -Y. (2000). MiPad: A Next Generation PDA Prototype, in *International Conference on Spoken Language Processing*, International Speech Communication Association. Beijing, China.

Igel, C., Toussaint, M., & Weishui, W. (2005). Rprop Using the Natural Gradient. In D. H. Mache, J. Szabados, & M. G. de Bruin (Eds.), *Trends and applications in constructive approximation* (pp. 259-272). International Series of Numerical Mathematics, Berlin: Brikhauser Basel.

Jaimes, F., Farbiarz, J., Alvarez, D., & Martínez, C. (2005). Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Critical Care (London, England)*, 9(2): 150–156.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2): 69–90.

Jerez, J. M., Franco, L., Alba, E., Llombart-Cussac, a, Lluch, a, Ribelles, N., … Martín, M. (2005). Improvement of breast cancer relapse prediction in high risk intervals using artificial neural networks. *Breast Cancer Research and Treatment*, 94(3): 265–72.

Jerez-Aragonés, J. M., Gómez-Ruiz, J. A., Ramos-Jiménez, G., Muñoz-Pérez, J., & Alba-Conejo, E. (2003). A combined neural network and decision trees model for prognosis of breast cancer relapse. *Artificial Intelligence in Medicine*, 27(1): 45-63.

José M. Jerez, E. A. L. F. (2005). Artificial neural networks and prognosis in medicine. Survival analysis in breast cancer patients, *ESANN*, 91–102.

Joshi, R., Reeves, C., & Johnston, C. (2005). Probabilistic Artificial Neural Networks for Malignant Melanoma Prognosis. In B. Ribeiro, R. F. Albrecht, A. Dobnikar, D. W. Pearson, N. C. Steele (Eds), *Adaptive and natural computing algorithms* (pp. 425-428). New York: Springer Wien.

Kalbfleisch, J. D., & Prentice, R. L. (1973). Marginal likelihoods based on Cox's regression and life model. *Biometrika*, 60(2): 267–278.

Kalderstam, J. (2015). *Neural network approaches to survival analysis* (Doctoral Dissertation). Lund university.

Kamangar, F., Dores, G. M., & Anderson, W. F. (2006). Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *Journal of Clinical Oncology*, 24(14): 2137–2150.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282): 457–481.

Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support Vector Machines in R. *Journal of Statistical Software*, 15(9): 28.

Kim, J. G., Ryoo, B.-Y., Park, Y. H., Kim, B.-S., Kim, T.-Y., Im, Y.-H., & Kang, Y.-K. (2008). Prognostic factors for survival of patients with advanced gastric cancer treated with cisplatin-based chemotherapy. *Cancer Chemotherapy and Pharmacology*, 61(2): 301–7.

Kleinbaum, D. G., & Klein, M. (2005a). Competing risks survival analysis. In D. G. Kleinbaum, & M. Klein, *Survival Analysis: A Self-Learning Text* (pp. 391-461). New York: Springer.

Kleinbaum, D. G., & Klein, M. (2005b). *Statistics for Biology and Health*. New York: Springer

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis : A Self-Learning Text*. New

York, NY: Springer Science+Business Media, LLC.

Kuntz, G. R., & Hammond, J. (1994). Predicting outcomes after closed-head injury (CHI) using a cascade-correlation artificial neural network. *Crit Care Med*, 22(1): A25.

Kuruvilla, J., & Gunavathi, K. (2013). Lung cancer classification using neural networks for CT images. *Computer Methods and Programs in Biomedicine*, 113(1): 202–209.

Kycler, W., Teresiak, M., & Łoziński, C. (2006). Prognostic factors for patients with gastric cancer after surgical resection. *Reports of Practical Oncology & Radiotherapy*, 11(5): 235–246.

Lapuerta, P., Azen, S. P., & LaBree, L. (1995). Use of neural networks in predicting the risk of coronary artery disease. *Computers and Biomedical Research, an International Journal*, 28(1): 38–52.

Laurentiis, M. De, Placido, S. De, Bianco, A. R., Clark, G. M., & Ravdin, P. M. (1999). A Prognostic Model That Makes Quantitative Estimates of Probability of Relapse for Breast Cancer Patients. *Clinical Cancer Research*, 1: 4133–4139.

Laurentiis, M., & Ravdin, P. M. (1994). Survival analysis of censored data: Neural network analysis detection of complex interactions between variables. *Breast Cancer Research and Treatment*, 32(1): 113-118.

Lee, E. T., & Wang, J. (2003a). *Statistical methods for survival data analysis*. Hoboken, New Jersy: John Wiley & Sons.

Lee, E. T., & Wang, J. W. (2003b). Identification of Prognostic Factors Related to Survival Time: Cox Proportional Hazards Model. In E. T. Lee, & J. Wang, *Statistical Methods for Survival Data Analysis* (pp. 298–338). Hoboken, New Jersy: John Wiley & Sons, Inc.

Lee, H.-C. (1994). *Skin cancer diagnosis using hierarchical neural networks and fuzzy logic* (Master thesis). University of Missouri-Rolla, USA

Li, C., Oh, S. J., Kim, S., Hyung, W. J., Yan, M., Zhu, Z. G., & Noh, S. H. (2009). Risk factors of survival and surgical treatment for advanced gastric cancer with large tumor size. *Journal of Gastrointestinal Surgery*, 13(5): 881–885.

Liestbl, K., Andersen, P. K., & Andersen, U. (1994). Survival analysis and neural nets. *Statistics in Medicine*, 13(12): 1189–1200.

Liestøl, K., & Andersen, P. K. (2002). Updating of covariates and choice of time origin in survival analysis: problems with vaguely defined disease states. *Statistics in Medicine*, 21(23): 3701–14.

Lin, R. S., Horn, S. D., Hurdle, J. F., & Goldfarb-Rumyantzev, A. S. (2008). Single and multiple time-point prediction models in kidney transplant outcomes. *Journal of Biomedical Informatics*, 41(6): 944–952.

Lin, W., Sun, J., Chang, S., & Wu, P. (2014). Factors Predicting Survival of Patients with Gastric Cancer, *Asian Pacific Journal of Cancer Prevention: APJCP*, 15(14): 5835–5838.

Lisboa, P. J. G., Etchells, T. A., Jarman, I. H., Arsene, C. T. C., Aung, M. S. H., Eleuteri, A., … Biganzoli, E. (2009). Partial logistic artificial neural network for competing risks regularized with automatic relevance determination. *IEEE Transactions on Neural Networks*, *20*(9), 1403–1416.

Lisboa, P. J. G., Vellido, A., & Wong, H. (2000). Bias reduction in skewed binary classification with Bayesian neural networks. *Neural Networks*, *13*(4), 407–410.

Lisboa, P. J. G., & Wong, H. (2001). Are neural networks best used to help logistic regression? An example from breast cancer survival analysis. Proceedings from *IJCNN'01:The International Joint Conference on Neural Networks.* IEEE.

Lisboa, P. J., & Taktak, A. F. G. (2006). The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks*, 19(4): 408–415.

Little, A.S. (1952). Estimation of the t-year survival rate over a limited period of time. Human Biology 24, 87-116

Machado, L. O. (1996). *Medical Application of Artificial Network Connectionist Models of Survival*(Doctoral dissertation). Stanford University, California.

Machin, D., Cheung, Y. B., & Parmar, M. K. B. (2006). *Survival analysis: a practical approach*. Chichester, UK: John Wiley & Sons

Mani, D., Drew, J., Betz, A., & Datta, P. (1999). Statistics and data mining techniques for lifetime value modeling. Proceedings from *the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: ACM press.

Marchevsky, A. M. (2007). The Use of Artificial Neural Networks for the Diagnosis and Estimation of Prognosis in Cancer Patients. In A. F. G. Taktak, & A. C. Fisher (Eds.), *Outcome Prediction in Cancer* (pp. 243–259). Oxford, UK: Elsevier.

Mariani, L., Coradini, D., Biganzoli, E., Boracchi, P., Marubini, E., Pilotti, S., … Rilke, F. (1997). Prognostic factors for metachronous contralateral breast cancer: A comparison of the linear Cox regression model and its artificial neural network extension. *Breast Cancer Research and Treatment*, 44(2): 167–178.

McClelland, J. L., Rumelhart, D. E., & Group, P. D. P. R. (1986). *Parallel distributed processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, USA: MIT press.

McLean, M. H., & El-Omar, E. M. (2014). Genetics of gastric cancer. *Nature Reviews Gastroenterology & Hepatology*, 11(11): 664–674.

Medsker, L., & Jain, L. C. (1999). *Recurrent neural networks: design and applications*. USA: CRC press.

Mehrabian, A. R., & Yousefi-Koma, A. (2007). Optimal positioning of piezoelectric actuators on a smart fin using bio-inspired algorithms. *Aerospace Science and Technology*, 11(2–3): 174–182.

Merry, A. H. H., Schouten, L. J., Goldbohm, R. A., & van den Brandt, P. A. (2007). Body mass index, height and risk of adenocarcinoma of the oesophagus and gastric cardia: a prospective cohort study. *Gut*, 53(11): 1503-1511.

Mofidi, R., Deans, C., Duff, M. D., de Beaux, A. C., & Paterson Brown, S. (2006). Prediction of survival from carcinoma of oesophagus and oesophago-gastric junction following surgical resection using an artificial neural network. *European Journal of Surgical Oncology : The Journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*, 32(5): 533–539.

Mojarad, S. A. (2012). *A reliable Network-Based Decision Support System for Breast Cancer Prediction* (Doctoral dissertation), Newcastle University.

Nasr, G. E., Badr, E. A., & Joun, C. (2002). Cross Entropy Error Function in Neural Networks : Forecasting Gasoline Demand. Proceedings from *FLAIRS conference*. Florida, USA: AAAI press.

Nelson, C. P., Lambert, P. C., Squire, I. B., & Jones, D. R. (2007). Flexible parametric models for relative survival, with application in coronary heart disease. *Statistics in Medicine*, 26(30): 5486-5498.

Neri, F., & Tirronen, V. (2010). Recent advances in differential evolution: a survey and experimental analysis. *Artificial Intelligence Review*, 33(1-2): 61–106.

Nielsen, S. F. (2003). Proper and improper multiple imputation. *International Statistical Review*, 71(3): 593–607.

O'leary, D. E. (1998). Using neural networks to predict corporate failure. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 7(3): 187–197.

Ohno-machado, L. (1997). A comparison of cox proportional hazards and artificial neural network models for medical prognosis. Computers in biology and medicine, 27(1): 55-65.

Ohno-Machado, L. (2001). Modeling medical prognosis: survival analysis techniques. *Journal of Biomedical Informatics*, 34(6): 428–39.

Ohno-Machado, L., & Musen, M. A. (1997). Modular neural networks for medical prognosis: Quantifying the benefits of combining neural networks for survival prediction. *Connection Science*, 9(1): 71–86.

Ohno-Machado, L., Walker, M. G., & Musen, M. A. (1995). Hierarchical neural networks for survival analysis, *Medinfo*, 8: 828–832.

Qiu F-M, Yu J-K, Chen Y-D, Jin Q-F, Sui M-H, Huang J. Mining novel biomarkers for prognosis of gastric cancer with serum proteomics. *J Exp Clin Cancer Res*.

2009;28:126. doi:10.1186/1756-9966-28-126.

Ou, G., & Murphey, Y. L. (2007). Multi-class pattern classification using neural networks. *Pattern Recognition*, 40(1): 4–18.

Oztekin, A., Delen, D., & Kong, Z. (James). (2009). Predicting the graft survival for heart-lung transplantation patients: An integrated data mining methodology. *International Journal of Medical Informatics*, 78(12): 84-96.

Paliwal, M., & Kumar, U. A.~~a.~~ (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36(1): 2–17.

Paoletti, X., Oba, K., Burzykowski, T., Michiels, S., Ohashi, Y., Pignon, J.-P., … Sasako, M. (2010). Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis. *JAMA: The Journal of the American Medical Association*, 303(17): 1729–1737.

Park, B., Shin, A., Park, S., Ko, K.-P., Ma, S., Lee, E.-H., … Yoo, K.-Y. (2011). Ecological study for refrigerator use, salt, vegetable, and fruit intakes, and gastric cancer. *Cancer Causes & Control*, 22(11): 1497–1502.

Parkin, D. M. (2004). International variation. *Oncogene*, 23(38): 6329–6340.

Patil, B. M., Joshi, R. C., Toshniwal, D., & Biradar, S. (2011). A new approach: role of data mining in prediction of survival of burn patients. *Journal of Medical Systems*, 35(6): 1531–42.

Pesonen, E., Eskelinen, M., & Juhola, M. (1998). Treatment of missing data values in a neural network based decision support system for acute abdominal pain. *Artificial Intelligence in Medicine*, 13(3): 139–146.

Pourhoseingholi, M. A., Moghimi-Dehkordi, B., Safaee, A., Hajizadeh, E., Solhpour, A., & Zali, M. R. (2009). Prognostic factors in gastric cancer using log-normal censored regression mode. *Indian Journal of Medical Research*, 129(3): 262–267.

Principe, J. C., Euliano, N. R., & Lefebvre, W. C. (1999). *Neural and adaptive systems: fundamentals through simulations with CD-ROM*. New York, USA: John Wiley & Sons, Inc.

Prinja, S., Gupta, N., & Verma, R. (2010). Censoring in clinical trials: review of survival analysis techniques. *Indian Journal of Community Medicine : Official Publication of Indian Association of Preventive & Social Medicine*, 35(2): 217–21.

Provost, F., & Domingos, P. (2000). *Well-trained PETs: Improving probability estimation trees*. Stern School of Business: New York University.

Puddu, P. E., & Menotti, A. (2012). Artificial neural networks versus proportional hazards Cox models to predict 45-year all-cause mortality in the Italian Rural areas of the Seven Countries Study. *BMC Medical Research Methodology*, 12(1): 100.

R, G. S., Vach, W., & Schumacher, M. (2000). On the misuses of artificial neural

networks for prognostic and diagnostic classification in oncology. *Statistics in Medicine*, 19(4): 541–561.

Rahman, M. M., & Davis, D. N. (2013). Addressing the Class Imbalance Problem in Medical Datasets. *International Journal of Machine Learning and Computing*, 3(2): 224–228.

Pereira, B. de B., & Rao, C. R. (2005). Survival analysis neural networks. *Learning and Nonlinear Models*, 3(2): 50–60.

Rao, N. (2010). Chapter-18 Survival Analysis, 5: 218–232.

Ravdin, P. M., & Clark, G. M. (1992). A practical application of neural network analysis for predicting outcome of individual breast cancer patients. *Breast Cancer Research and Treatment*, 22(3): 285–293.

Ravdin, P. M., Clark, G. M., Hilsenbeck, S. G., Owens, M. A., Vendely, P., Pandian, M. R., & McGuire, W. L. (1992). A demonstration that breast cancer recurrence can be predicted by neural network analysis. *Breast Cancer Research and Treatment*, 21(1): 47–53.

Riedmiller, M. (1994). Rprop-description and implementation details. University of Karlsruhe.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. University of Oxford, United Kingdom: Cambridge university press.

Ripley, R. M. (1998). Neural Network Models for Breast Cancer Prognosis. *Neural computing and applications*, 7(4): 367-375.

Ripley, R. M., Harris, a. L., & Tarassenko, L. (1998). Neural network models for breast cancer prognosis. *Neural Computing & Applications*, 7(4): 367–375.

Ripley, R. M., Harris, A. L., & Tarassenko, L. (2004). Non-linear survival analysis using neural networks. *Statistics in Medicine*, 23(5): 825–42.

Roberts, D. A., Holcomb, J. B., Parker, B. E., Sondeen, J. L., Pusateri, A. E., Brady, W. J., … Young, J. S. (2002). The use of polynomial neural networks for mortality prediction in uncontrolled venous and arterial hemorrhage. *The Journal of Trauma*, 52(1): 130–135.

Robin, A. X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, C., … Robin, M. X. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *B.M.C. Bioinformatics*, 12(1): 77.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6): 386.

Royston, P., & White, I. R. (2011). Multiple imputation by chained equations (MICE): implementation in Stata. *Journal of Statistical Software*, 45(4): 1–20.

Rubin, D. B. (1978). Multiple imputations in sample surveys-a phenomenological Bayesian approach to nonresponse. In *Proceedings of the survey research methods section, American Statistical Association*, 20–34.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Internal Representations by Error Propagation. In D. E. Rumelhart, & J. McClelland (Eds.), *Parallel Distributed Processing, Explorations in the Microstructure of Cognition* (pp.318-362), Cambridge, MA: MIT Press.

Russell, S., & Norvig, P. (1995). *Artificial intelligence: a modern approach*. Englewood Cliffs, New Jersey: Prentice Hall.

Samy, G. N., Ahmad, R., and Ismail, Z. (2012). Adopting and Adapting Medical Approach in Risk Management Process for Analysing Information Security Risk. In J. Emblemsvag (Ed.), *Risk Management for the Future - Theory and Cases* (pp.365-388). Rijeka, Kvarner: INTECH Open Access Publisher.

Schwarzer, G., Vach, W., & Schumacher, M. (2000). On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Statistics in Medicine*, *19*(4): 541–561.

Seshadri, R. a, Jayanand, S. B., & Ranganathan, R. (2011). Prognostic factors in patients with node-negative gastric cancer: an Indian experience. *World Journal of Surgical Oncology*, 9(1): 48.

Setiono, R., & Hui, L. C. K. (1995). Use of a quasi-Newton method in a feedforward neural network construction algorithm. *Neural Networks, IEEE Transactions on*, 6(1): 273–277.

Shaliz, C. R. (2013). Logistic Regression. In C. R. Shaliz, *Advanced Data Analysis from an Elementary Point of View* (pp.223–237). Preprint of book found at http://www. stat. cmu. edu/∼ cshalizi/ADAfaEPoV.

Shankaracharya, D. O., Samanta, S., & Vidyarthi, A. S. (2010). Computational intelligence in early diabetes diagnosis: a review. *The Review of Diabetic Studies: RDS*, 7(4): 252.

Siewert, J. R., Böttcher, K., Stein, H. J., & Roder, J. D. (1998). Relevant prognostic factors in gastric cancer: ten-year results of the German Gastric Cancer Study. *Annals of surgery*, 228(4): 449.

Silva, L. M., Marques de Sá, J., & Alexandre, L. A. (2008). Data classification with multilayer perceptrons using a generalized error function. *Neural Networks*, 21(9): 1302–1310.

Singh, R., & Mukhopadhyay, K. (2011). Survival analysis in clinical trials: Basics and must know areas. *Perspectives in Clinical Research*, 2(4): 145.

Siristatidis, C. S., Chrelias, C., Pouliakis, A., Katsimanis, E., & Kassanos, D. (2010). Artificial neural networks in gynaecological diseases: Current and potential future applications. *Medical Science Monitor*, 16(10): 231–236.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4): 427–437.

Spelt, L., Andersson, B., Nilsson, J., & Andersson, R. (2012). Prognostic models for outcome following liver resection for colorectal cancer metastases: A systematic review. *European Journal of Surgical Oncology (EJSO)*, 38(1): 16–24.

Stajduhar, I., & Dalbelo-Basić, B. (2010). Learning Bayesian networks from survival data using weighting censored instances. *Journal of Biomedical Informatics*, 43(4): 613–22.

Steevens, J., Schouten, L. J., Goldbohm, R. A., & Van Den Brandt, P. A. (2011). Vegetables and fruits consumption and risk of esophageal and gastric cancer subtypes in the Netherlands Cohort Study. *International Journal of Cancer*, 129(11): 2681–2693.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj*, 338: b2393.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., … Kattan, M. W. (2010). Assessing the Performance of Prediction Models. *Epidemiology*, 21(1): 128–138.

Street, W. N. (1998). A Neural Network Model fo Prognostic Prediction. In *Proceedings of the Fifteenth International Conference on Machine Learning*. San Francisco, USA: Morgan Kaufmann Publishers Inc.

Suárez, S., González, C., & Hernández, J. (2013). Design Methodology of an Equalizer for Unipolar Non Return to Zero Binary Signals in the Presence of Additive White Gaussian Noise Using a Time Delay Neural Network on a Field Programmable Gate Array. *Sensors*, 13(12): 16829–16850.

Swets, J., & Pickett, R. (1982). *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*. New York, USA: Academic Press.

van Belle, V., Pelckmans, K., van Huffel, S., & Suykens, J. A. K. (2011). Improved performance on high-dimensional survival data by application of survival-SVM. *Bioinformatics*, 27(1): 87–94.

Venables, W. N., & Ripley, B. D. (2013). *Modern applied statistics with S-PLUS*. New York, USA: Springer Science & Business Media.

Wang, X., Wei, W., Kong, F., Lian, C., Luo, W., Xiao, Q., & Xie, Y. (2012). Prognostic significance of Cdx2 immunohistochemical expression in gastric cancer: a meta-analysis of published literatures. *Journal of Experimental & Clinical Cancer Research : CR*, 31(1): 98.

Weibull, W. (1939). *A statistical theory of the strength of materials*. The Royal Swedish Institute of Engineering Research, Stockholm: Ingeniors Vetenskaps Akademiens.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained

equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4): 377–399.

Wu, Y., Giger, M. L., Doi, K., Vyborny, C. J., Schmidt, R. A., & Metz, C. E. (1993). Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology*, 187(1): 81–87.

Xue, X., Xie, X., Gunter, M., Rohan, T. E., Wassertheil-Smoller, S., Ho, G. Y. F., … Strickler, H. D. (2013). Testing the proportional hazards assumption in case-cohort analysis. *BMC Medical Research Methodology*, 13(1): 88.

Yamamura, Y., Nakajima, T., Ohta, K., Nashimoto, A., Arai, K., Hiratsuka, M., … Goto, M. (2002). Determining prognostic factors for gastric cancer using the regression tree method. *Gastric Cancer*, 5(4): 201–207.

Yang, Y. (2010). *Neural Network Survival Analysis* (Master dissertation). Universiteit Gent.

Yu, D., Yao, K., Su, H., Li, G., & Seide, F. (2013). KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition. Proceedings from *ICASSP: The International Conference on Acoustics, Speech and Signal Processing.* IEEE.

Zell, A., Mamier, G., Vogt, M., Mache, N., Hübner, R., & Döring, S. (2000). SNNS, Stuttgart Neural Network Simulator, User Manual, Version 4.2. Institute for Parallel and Distributed High Performance Systems, University of Stuttgart.

Zhang, G. P. (2003). *Neural networks in business forecasting*. Georgia state university, USA: IRM press.

Zhao, J., Liang, Q., Cheung, K.-F., Kang, W., Dong, Y., Lung, R. W.-M., … Yu, J. (2013). Somatostatin receptor 1, a novel EBV-associated CpG hypermethylated gene, contributes to the pathogenesis of EBV-associated gastric cancer. *British Journal of Cancer*, 108(12): 2557–64.

Zhu, H. P., Xia, X., Chuan, H. Y., Adnan, A., Liu, S. F., & Du, Y. K. (2011). Application of Weibull model for survival of patients with gastric cancer. *BMC Gastroenterology*, 11(1): 1.

Zhu, L., Luo, W., Su, M., Wei, H., Wei, J., Zhang, X., & Zou, C. (2013). Comparison between artificial neural network and Cox regression model in predicting the survival rate of gastric cancer patients. *Biomedical Reports*, 1(5): 757–760.

Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity , Specificity , Accuracy , Associated Confidence Interval and ROC Analysis with Practical SAS ® Implementations. In: NESUG Proceedings: Health Care and Life Sciences, Baltimore, Maryland.

BIODATA OF STUDENT

Hamid Nilsaz Dezfouli was born in the south of Iran, Ahwaz. He gained his primary education and secondary education at Ahwaz. He completed the bachelor degree in Statistics at Faculty of Mathematics and Computer Science at Shahid Chamran University in Ahwaz in 2003.He has experienced as a lecturer in statistics around 5 years. He continued his study in Master Degree again in Applied Statistics at Faculty of Mathematics and Computer Science at Shahid Chamran University and graduated in 2006. After his graduation, he got an offer to teach statistics at the university in Ahwaz. In 2010, he pursued his study at PhD level at the Institute for Mathematical Research (INSPEM) at University Putra Malaysia in the field of statistical modelling and forecasting.

LIST OF PUBLICATIONS

Nilsaz Dezfouli, H. Bakar, M. R. A, Pourhoseingholi, M. A, Arasan, J, Adam, M. B (2017). Improving Gastric Cancer Outcome Artificial Prediction Using Single Time Point Artificial Neural Networks Models, *Cancer Informatics* (**Accepted :** September 18, 2016).

Nilsaz-Dezfouli H. Bakar, M. R. A., Pourhoseingholi, M. A. (2016). An Artificial Neural Network Model for Outcome Prediction.in Gastric Cancer Patients, *J Fundam Appl Sci.*8(3s):1687-1698.

Nilsaz, H, Bakar, M. R. A. (2012). Feed Forward Neural Networks Models For Survival Analysis. Proceedings from *The International Conference on Statistics in Science, Business and Engineering* (pp.10-12). Langkawi, Malaysia: (IEEE Conference Publication indexed in Scopus)

Nilsaz, H, Bakar, M. R. A., Zabihipour, S. M. (2012). Neural network methods for censored survival data, *Seminar on Applications of Cutting-edge Statistical Methods in Research (pp.4-5).* Universiti Putra Malaysia.

Nilsaz, H., Bakar, M. R. A., Pourhoseingholi, M. A., Adam, M. B., Arasan, J. (2014). Survival Data Analysis Using a Classification Neural Networks Model. *The Regional Fundamental Science Congress(FSC).*19th to 20th August 2014 Kuala Lumpur, Selangor, Malaysia.

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : _____

TITLE OF THESIS / PROJECT REPORT :

SINGLE AND MULTIPLE TIME–POINT ARTIFICIAL NEURAL NETWORK MODELS FOR PREDICTING THE SURVIVAL OF GASTRIC CANCER PATIENTS

NAME OF STUDENT : HAMID NIL SAZ DEZFOULI

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☐	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐ **PATENT** Embargo from_____ until _____
(date) (date)

Approved by:

_____ _____

(Signature of Student) (Signature of Chairman of Supervisory Committee)
New IC No/ Passport No.: Name:

Date : Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]