T.R.N.C

TURKISH REPUBLIC OF NORTHERN CYPRUS
NEAR EAST UNIVERSITY

INSTITUTE OF HEALTH SCIENCES

Artificial Neural Networks for Prediction of High Risk Very Pre-term/Pre-mature Birth

SANDRA SENDI DASHAN

Master of Science in Biostatistics

Supervisor:

Prof. Dr. İlker ETİKAN

NICOSIA, 2020

T.R.N.C

TURKISH REPUBLIC OF NORTHERN CYPRUS

NEAR EAST UNIVERSITY

INSTITUTE OF HEALTH SCIENCES

Artificial Neural Networks for Prediction of High Risk Very Pre-term/Pre-mature Birth

SANDRA SENDI DASHAN

Master of Science in Biostatistics

Supervisor:

Prof. Dr. İlker ETİKAN

NICOSIA 2020

NEAR EAST UNIVERSITY

GRADUATE INSTITUTE OF HEALTH SCIENCES

THESIS APPROVAL CERTIFICATE

I hereby declare that this project was based on my research, findings and gathering of study tools in relation to artificial neural networks, pre term birth/ pre-mature birth. The information I was working with were from some of my original ideas and the ones that were not i made sure to put their references below and i also in declaration are mindful to state that I was careful in abiding by the rules and regulations stated by the thesis committee.

Thank you!!

ACKNOWLEDGEMENT

I would like to express my sincere gratitude towards some people that have been very helpful towards my time of study in NEAR EAST UNIVERSITY, BIOSTATISTICS till the 2019/ 2020 session. And now that i am drawing closer to the finish line. To the people that have been of great help towards the preparation of my thesis in giving me ideas, tips and many other points. I would first of all like to give my profound gratitude to God almighty without his grace this wouldn't have been possible.

Secondly, I would like to give thanks to my supervisor Prof. Dr. İlker ETİKAN he has been of great help by leading me in the right path of getting study materials and the advices. I would love to give thanks to my parents Mr. and Mrs. DASHAN, for leading me in the right path and for financial, social and emotional support.

Finally I would also like to thank my friends and lecturers in the department Adebayo and Asst. Prof. Dr. Özgür TOSUN which have been together with me along it has been a wonderful ride.

To God, family and friends………….

ABSTRACT

This aim of this study is to be able to predict high risk pre-term birth and very pre term birth through the use of an Artificial Neural Network (ANN) that will be trained and be able to predict if the output will be full term, very preterm and moderate/late preterm. This is an extensive study by which pre term birth is evaluated using various risk factors. Pre term delivery is a delivery that occurs between twenty weeks and before 37 weeks of pregnancy it is common and at a high rate in most developing countries in the world. Most times proper diagnosis of pre term birth is a luxury not too common in these countries. Pre term birth is one of the leading causes of peri natal mortality. Pre term birth is most times caused by a number of risk factors that cause spontaneous pre term birth or pre term rupture of the membranes (whose causes are mostly unknown) or induced pre term birth due to medical or non-medical reasons. Artificial neural networks (ANN) is used as an alternative or a complimentary technique of testing high risk maternal populations using clinical tests that cost a lot and are invasive. The data used in this study was obtained from the maternal and obstetrics records of the labor ward of ELITE IVF AND RESEARCH HOSPITAL, LEFKOSIA, NORTH CYPRUS. The data was gotten from the duration of 2015 to 2019. These information was obtained from 218 women which had delivered live births. 153 (70.2%) of the births were full term, 27 (12.4%) were moderate/late preterm and 38 (17.4%) were very preterm. Here labor is defined as regular uterine contractions with changes in cervix. Cases of indicated preterm birth were excluded and vaginal induced delivery was excluded so in other words all induced deliveries were excluded and caesarean deliveries whereby the woman went through labor was included. 7 independent variables which serve as our risk factors were used as the input, namely; in vitro fertilization (IVF), multiple birth (multbirth), gestational diabetes (gestdiabetes), hypertension, premature rupture of membranes (PROM), body mass index (BMI) classified age (ageclassified).

An ANN model was achieved which shows the relations and the magnitude of the relations between the input variables which are the risk factors and the output. We obtained a model that had an error of 16.4% incorrect predictions for the training set and 18.5% for the test set. The model was alo represented by an ROC curve to depict the accuracy of the test of which the AUC showed full term to have the highest accuracy. It also showed the bar chart of the normalized importance of the variable of which they have the order IVF, multiple birth, gestational diabetes, hypertension, PROM, BMI, age2.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PTB: Pre Term Birth

NICU: Neonatal Intensive Care Unit

SIDS: Sudden Infant Death Syndrome

IV: Intravenous

RBC: Red Blood Cells

PROM: Preterm Rupture Of Membranes

CDSS: Clinical Decision Support System

ANN: Artificial Neural Networks

ROP: Retinopathy of Prematurity

TPR: True Positive Rate

TNR: True Negative Rate

ROC: Receiver Operator Characteristics

AUC: Area Under The Curve

IPBC: International Preterm Birth Collaborative

GAPPS: Global Alliance to Prevent Prematurity and Stillbirth

RHR: Reproductive Health Research

CNN: Convolution Neural Network

CHAPTER 1

INTRODUCTION

Pre term birth has been an undeniably harsh reality for a lot of women and the society for a very long time. There are a lot of factors that lead to or cause pre-term birth, which will further be discussed. For infants worldwide pre-term birth is the most common cause of infant mortality (Liu et al, 2012). There are about 15 million pre-term births globally, and the chances of survival with the babies coming out with no complications is relatively low. Most preterm births leave survivors with breathing issues, brain bleeds or digestive problems as babies and in the long run we can start to notice problems of developmental delays. During the course of a pregnancy, a baby goes through very important growth and development from the beginning till even the final weeks (Rogers and Velten, 2011). When a baby is born prematurely, this affects their growth and though the complications associated with preterm birth varies depending on how early in the pregnancy the pre-term birth occurred the earlier a baby is born, the higher the chances of being faced with serious developmental issues and death.

There are a lot of risk factors associated with pre-term births, although the cause is mostly not really clear, now this risk factors are not a yardstick to pre-term birth, a woman can not show any of the signs of this risk factors and actually have pre-term birth (Martin et al, 2011). On the other hand this risk factors are highly effective in predicting pre-term birth. Some of this risk factors include history of previous premature birth, problems associated with uterus, cervix or placenta, smoking or usage of drugs, physical injuries or trauma and many more. This risk factors help in diagnostics or the prediction of a pre-term birth.

It is of great importance to be able to predict the risks of pre-term birth this is in order to be able to curb the high number of pre-term births happening worldwide, as this is a very important public health priority. We can curb pre-term birth by providing medical attention and access to pregnant women both before and between the pregnancy (WHO, 2007), predicting high risk preterm birth through the various risk factors and providing effective treatments to prevent these preterm births. The importance of early prediction by health care practitioners can never be overemphasized just as the saying goes prevention is better than cure. Paying serious attention and care to a babies growth during a pregnancy both by the pregnant woman and health care practitioners is very important to be able to quickly notice any anomalies. The scope of this thesis is to research on, identify and predict high risk pre-term births based on the risk factors using artificial neural networks.

1.1 Background of study

1.1.1 What is pre-term birth

Preterm/premature birth is an early birth that takes place any time before the start of the pregnancies 37th week gestational period. Pre-term birth is a global issue because it occurs in underdeveloped, developing and developed countries. Preterm birth can be classified according to the gestational age and or to the birthweight when there are large differences (Gyamfi-Bannerman et al, 2011). Preterm birth serves as a major challenge associated with perinatal healthcare. It serves as a major risk factor for disabilities and various neurological impairments (lee et al, 2008).

1.1.2 Stages of pre-term birth

There are various stages at which pre-term birth occurs each with their various gestational ages.

- **Extremely preterm birth:** This are preterm births that take place before 28 weeks gestation. This period of preterm birth has the highest effects of mortality and morbidity on infants. Infections are mostly responsible for this stage of preterm birth.
- **Very moderate preterm birth:** These are preterm births that take place within 28-31 weeks of pregnancy. It is caused by both issues related with infections and stress and lifestyle.
- **Moderate/late preterm birth:** These are preterm births that take place between 32-36 weeks of gestation basically stress and lifestyle are accountable for these preterm birth, it is associated with higher chances of survival but infants may be faced with issues of disabilities and neurological impairments.

1.1.3 Categories of preterm birth

There are 3 conditions that explain the causes or origin of preterm birth, the risk factors associated with pre-term birth are classified under these categories.

- Medically indicated (iatrogenic) preterm birth: This is pre-term birth that occurs due to medical reason or causes that have a more medical explanation it occurs in about 25% of preterm births. This relates to complications that pregnant women are faced with such as abrupto placentae, intrauterine growth retardation, fetal distress, maternal hypertension.
- Preterm premature rupture of membranes: This also occurs in about 25% of all preterm births, whereby the membranes rupture before time (Goldenberg et al, 2012).
- Spontaneous (idiopathic) preterm birth: This category of preterm birth occurs in about 50% of total preterm births. This refers to the onset of labor before 37 weeks with the fetal membranes being intact. (Lee and Ahn, 2019).

Preterm infants may have health problems because there wasn't enough time for their organs to develop in the womb. Pre-term birth depending on the gestational age, mostly extremely preterm birth and very preterm birth has a lot of long-term effects on the infants after birth, this makes preterm birth a long lasting societal problem, it not only causes cost problems to the healthcare system because preemies have to stay in the incubator till the completion of their gestational age meaning longer hospital time, it also causes long term effects on the development of the preemies, such developmental problems exists such as;

- Intellectual/cognitive impairments that can develop disabilities and learning and therefore the need for special education.
- Motor problems and neurodevelopmental problems.
- Difficulties in feeding, late development of feeding skills, refusal of food, difficulties eating or chewing, and poor appetite (Goldenberg, 1996).
- Behavioral concerns such as hyperactivity and attention problems.
- Emotional problems like anxiety and depression.

Infants that are born preterm may suffer a number of health conditions such as low or very low birth weight, caloric needs that are at an increased rate, Blood problems like anemia, hypoglycemia and jaundice, difficulties in feeding due to a lack of the reflexes needed for sucking and swallowing, problems with breathing such as apnea, bronchopulmonary dysplasia and chronic lung diseases, immaturity in digestion and an impaired absorption of carbohydrates and lipids, cerebral palsy, an impairment of the part of the brain that controls things like movement and muscle tone(Maner and Garfield, 2007), a delay in development and also a low performance in cognitive functions, problems with vision like retinopathy of prematurity (ROP) which may cause blindness, behavioral problems and psychiatric disorders, difficulties hearing, increased risk for necrotizing enterocolitis (NEC) due to underdeveloped gastrointestinal systems, Increased risk of Sudden Infant Death Syndrome (SIDS), problems faced with temperature control, low blood pressure (hypotension) and Heart problems such as patent ductus arteriosus (Blondel et al, 2006).

1.2 Aims and objectives

The aim of this study is to develop an ANN predictive model that has the ability to predict birth age at various stages of full term, moderate/late preterm and very preterm. The area of interest is in this models ability to predict between full-term, late/moderate preterm and high risk very preterm birth as an output given the risk factors associated with very preterm birth as variables.

1.3 Risk factors associated with pre-term birth

Medical and pregnancy conditions:

- Prior pre term birth history.
- Infections such as urinary tract infections, sexually transmitted infections, vaginal infections (Goldenberg, 1996).
- Multiple gestations, carrying more than one baby like twins or triplets, more than 50% of twin births are preterm (M, 2018).
- Use of assisted medical reproductive technology, in-vitro fertilization (Hansen, 2011).
- High blood pressure in pregnancy, diabetes and gestational diabetes.
- Abnormalities of the reproductive organs like women with short cervix are at a high risk of preterm birth or women whose cervix shortens in the second trimester of the pregnancy instead of the third trimester as it should be (Morken et al, 2013).
- Obesity or being underweight right before the pregnancy.
- Developmental abnormalities in the fetus.
- Risk of rupture of the uterine membrane, this happens mostly if there was a subsequent caesarean section.

Social, economic or personal reasons:

- Low or high maternal age, women below the age of 18 are more likely to give birth preterm and women above the age of 35 are also more likely to have pre-term delivery because of the health complications they might have like high blood pressure.
- Black race, preterm birth occurs more prominently in certain races and ethnicities for some unexplainable reasons races like black and American/Indian races are more likely to give birth preterm than women from the white race.

Behavioral characteristics

- Substance abuse like smoking and drugs.
- Late or no healthcare in the course of the pregnancy.
- Physical, emotional or sexual abuse in the course of the pregnancy..
- Exposure to a number of environmental pollutants and contaminants.
- Late prenatal care.
- Alcohol and tobacco use.
- Stress.

Most preterm babies are being faced with the risk of weighing less than 5 pounds 8 ounces that is 2,500g as opposed to 6 or 7 pounds which is the ideal birth weight for fully mature babies.

1.4 Symptoms of being preterm

The symptoms of prematurity may slightly differ in each baby:

- Shiny, pink or red and very thin skin that may even have veins that show (Goldenberg, 1996).
- Very little body fat.
- Low muscle build up and tone.
- Weak cry.
- Not fully developed genital parts that are little in size.
- Small size and low body weight.

1.5 Treatment for preterm babies

Treatment and care for preemies differ based on a number of issues like the severity of the condition.

- Proper care and watch of the babies temperature, blood pressure, oxygen levels, heart and breathing rate and bilirubin levels this care can be administered by using a temperature controlled bed, the use of oxygen masks and a ventilator which is a breathing machine (Allen & Newnham, 2018).
- Administering needed x-ray tests and other imaging tests.
- Practicing of kangaroo care which is regular skin to skin contact with the parents.
- The use of IV (intravenous) feeding, fluids or administering of medication.

Preemies are being taken care of by the NICU (neonatal intensive care unit) until the babies due date or till the baby is fully developed and healed. The signs of full development include being able to stay warm in an open crib, being able to take all kinds of feeding, maintaining the rate at which they are normally expected to grow if they have no signs of apnea. Premature babies are at a higher risk for (SIDS) sudden infant death syndrome.

1.6 Problems preemies may face

- Preterm babies may have problems breathing.
- Preterm babies may have problems with blood related issues like low red blood cells (RBC) count anemia, jaundice which is the skin having a yellow color from breaking down red blood cells, hypoglycemia which is low blood sugar (Rogers and Velten, 2011).
- Infections and kidney problems.
- Problems digesting food properly due to dead sections of the intestines, (necrotizing enterocolitis) and inflammations (Stower, 2018).
- Problems with the nervous system like seizures and brain bleeds.

CHAPTER 2

LITERATURE REVIEW

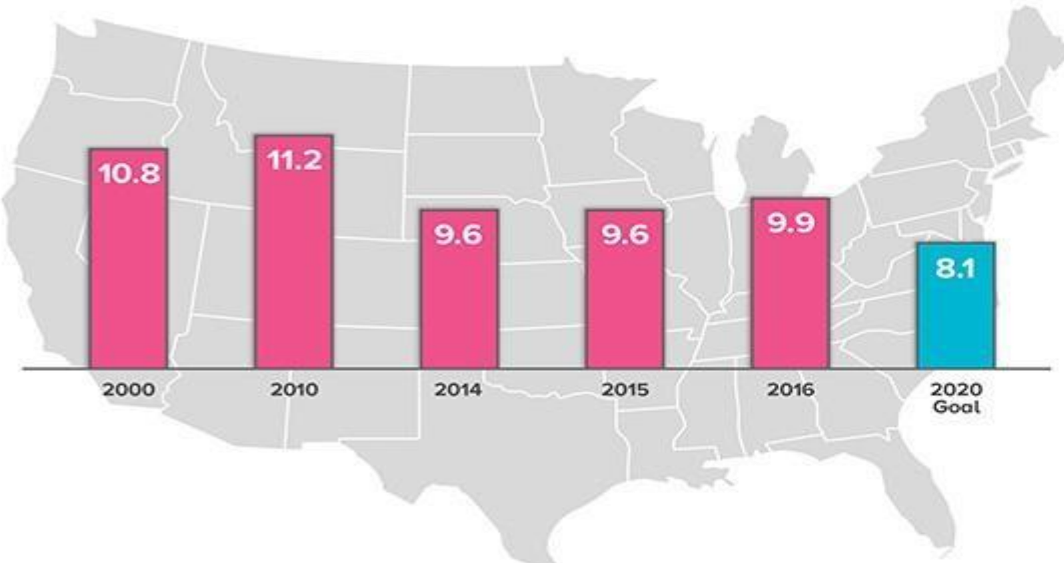
In children under 5 years and the first few months of life preterm birth has proven to be the second leading cause of death (Liu et al, 2012) and for the preemies that do not die a large percentage of those cases suffer developmental conditions. A global team consisting of the leading international organizations like World Health Organization, United Nations International Children's Fund, National Institute For Health And Care Excellence and many other academic institution to find scientifically proven solutions for preterm birth (Who, 2012). Public health, child and maternal health has gone through a series of transition in the course of time. In the case of the decline in maternal and child mortality, drugs and technologies that were responsible in industrialized countries in the 20th century (De Brouwere et al, 1998). Preterm birth posed a great issue not only because of the health risks it has on both the mother and child but also on other aspects like financial aspects, the expenses spent on hospital and healthcare cost were significantly higher for premature and low-birthweight newborns, in comparison to newborns that had uncomplicated births (March of Dimes, 2012). In 2005, there were approximately 3 million preterm births, this caused a global outcry for the issue of preterm birth to be looked into, the WHO Department of reproductive health research (RHR) released global and regional estimates of preterm birth in 2008 (Beck et al., 2010). This shift in attention on preterm birth and a trend in series of events led to a number of establishments being formed in the upcoming years like the International Preterm Birth Collaborative (IPBC) and the launch of the Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) in 2009 (lawn et al, 2010). Preterm birth is one the largest single conditions in the Global Burden of Disease analysis given the high mortality and the considerable risk of lifelong impairment (Who, 2008). Late preterm birth are at a lower risk of mortality and other conditions affecting preterm births but, very preterm birth and extremely preterm births still face higher mortality rates and other health conditions especially in countries with low income. Significant improvements in medical and health care have led to an improved rate of survival and long-term outcomes in high income countries. Among very and extremely preterm babies (Saigal and Doyle, 2008). The rate of preterm birth is gradually showing a significant increase in countries that have available data. This is because more records are being kept and because more preterm births are being registered. maternal age, multiple pregnancies, access to infertility treatment and underlying health problems in the mother include the possible reasons for this, especially with the changes that are now occurring in obstetrics and increasing age of pregnancy (Joseph et al., 2002). Still there are problems being faced with diagnosing preterm birth by early recognition. Prediction in medical sciences is really a way forward and with the help of ANNs this will really change the approach to preterm birth by curbing the risks and dangers associated. Although this is a study and field that needs a whole lot of improvement.

2.1 Statistics for preterm birth

- 11.4 percent of total pregnancies end in early birth.
- Complications from Preterm birth which are responsible for approximately 1 million deaths as of the year 2015. Are the leading cause of death among children under the age of five years (Bick, 2012).
- The rate of preterm birth is within the range of 5% to 18% of the total number of babies born across 184 countries (Bick, 2012).
- In the United States, about 450,000 babies are from preterm birth. And around the world a whopping 15 million preterm babies that is about 1 in 10 births are preterm.
- There has been a drop in the premature birth rate in the United States for the seventh consecutive year, to 11.4 percent of all births in 2013. This is the lowest rate in 17 years (Reedy, 2007)
- Three-quarters of these deaths could be prevented with current, cost-effective interventions.
- Asian and sub-Saharan African countries accounted for 78·9% of livebirths and 81·1% of preterm births globally in 2014. New estimates show that In this same year, the pre term birth ranged from 8.7% in Europe to 13.4% in north Africa (Howson et al, 2013).
- About 80 percent of preterm births are unanticipated.
- Approximately 45-50% of preterm births are idiopathic that is they are unknown causes (Bick, 2012).
- 30% of pre term birth are related to preterm rupture of membranes (PROM) (Bick, 2012). 15-20% of preterm birth are attributed to medically indicated causes.
- The medical expenses for a baby born prematurely average about $54,000, compared with $4,000 for a healthy full-term newborn.

- Preterm birth rates remain higher among certain racial and ethnic groups in the U.S., including black, Native American and Hispanic women as opposed to some other ethnic groups.

Figure 2.1: Bar chart showing the global preterm rates.



2.1.1 Statistics according to gestational age

- More than 70% of premature babies are born between 34 and 36 weeks gestation. That is most pre term babies are born late preterm. (Blencowe et al, 2013)
- 12% of premature babies are born between 32 and 33 weeks gestation.
- 10% of premature babies are born between 28 and 32 weeks gestation.
- 6% of premature babies are born before 28 weeks gestation that is they are extremely preterm (Hansen, 2011).
- Preemies born at 23 weeks have a 17 percent chance of survival.
- Preemies born at 24 weeks have a 39 percent chance of survival.
- Preemies born at 25 weeks have a 50 percent chance of survival.
- Preemies born at 26 weeks have an 80 percent chance of survival.
- Preemies born at 27 weeks have a 90 percent chance of survival.
- Preemies born between 28-31 weeks gestation have at 90-95 percent chance of survival (Blencowe et al., 2013).

- Preemies born between 32-33 weeks have a 95 percent chance of survival.
- Preemies born 34 weeks or greater have the same likelihood of survival as a full term infant.
- This shows that the gestational age has a direct proportionality to the chances of survival among preterm birth babies the higher the gestational age, the higher their chances of survival.
- With singleton pregnancy there is a 7% risk of preterm birth.
- With multiple pregnancy there is a 57% risk of preterm birth.

2.1.2 Statistics of preterm birth based on ethnicity

- Bangladeshi: 8%
- Indian: 7%
- Pakistani:7%
- Black African: 8%
- Black Caribbean: 10%
- White British: 7%
- White Other: 6%

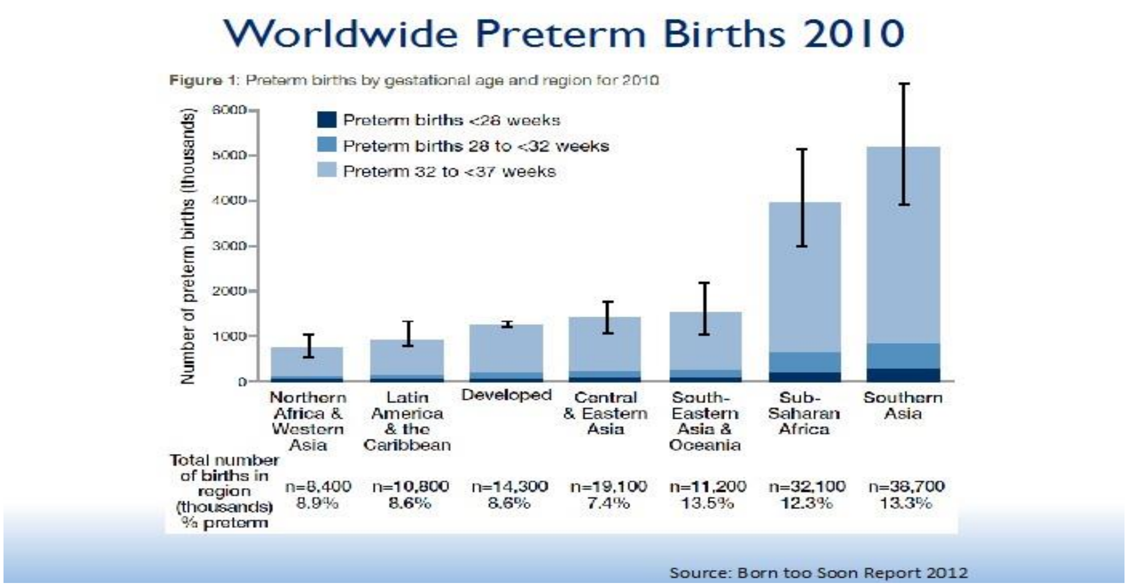
2.1.3 The 10 countries with the greatest number of preterm births.

- India: 3 519 100
- China: 1 172 300
- Nigeria: 773 600
- Pakistan: 748 100
- Indonesia: 675 700
- United States of America: 517 400
- Bangladesh: 424 100
- Philippines: 348 900
- Democratic Republic of the Congo: 341 400
- Brazil: 279 300

2.1.4 The 10 Countries With The Highest Rates Of Preterm Birth Per 100 Live Births.

- Malawi: 18.1 preterm births per 100 births
- Comoros: 16.7
- Congo: 16.7
- Zimbabwe: 16.6
- Equatorial Guinea: 16.5
- Mozambique: 16.4
- Gabon: 16.3
- Pakistan: 15.8
- Indonesia: 15.5
- Mauritania 15.4 (Howson et al, 2013).

Figure 2.2: Worldwide preterm rates in 2010.



2.2 Artificial Neural Networks

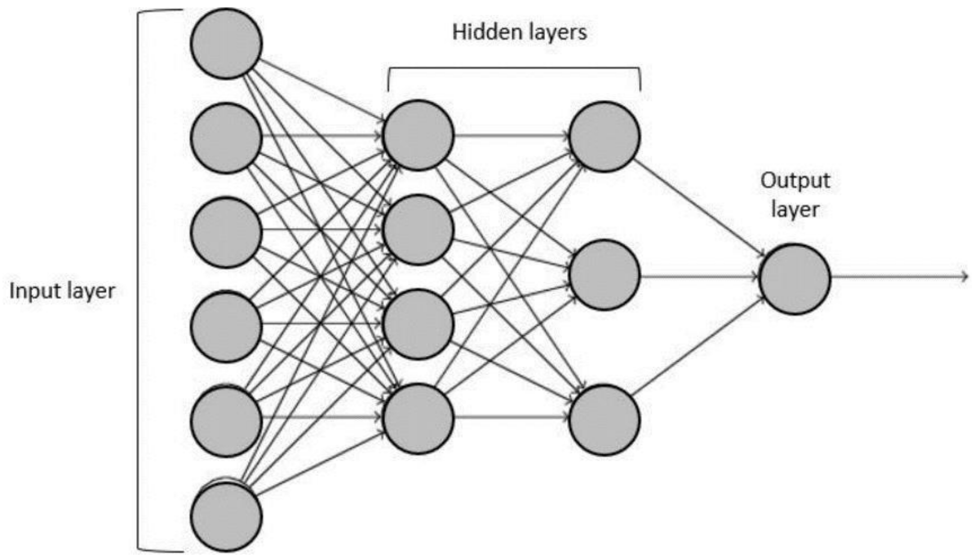
Neural network is a vast field that is very much in use in our world today, there are so many practical applications that use neural networks, the world is growing at a fast pace and neural networks, deep learning, machine learning. Neural networks look at data and tries to figure out the function or set of calculations that will turn the input variable into the output variable. Neural networks forms the basis of deep learning which a subfield of machine learning (Ripley, 1996). Neural networks are inspired by the structure of the human brain basically the idea of this neural networks is to create a system that has the ability to decipher, recognize, read and make decisions(Michel et al, 1998). A neural network will take in data and train itself to recognize the patterns in this data and will further predict the output or form a new set of similar data.

Neural networks are being made up of layers of neurons. And a neuron is a code processing unit of the network that holds a number. In a neural network we have the input layer, hidden layer (which can be two or more layers), output layer (Wolpert, 1992). The hidden layers perform most of the computations required by the network. In the working system of the neural networks, neurons of one layer are connected to neurons of the next layer through channels and each of these channels are assigned numerical values known as weights. The inputs are multiplied by the corresponding weights and the sum of the multiplication is being sent as input to the neurons that are in the hidden layer. Each of this neurons is associated with a numerical value called the bias. The bias is then added to the input sum. The resulting value of the hidden layer is passed through an activation function which determines the potency of each of the neurons being activated or not. The activation neurons will further transmit data into the neurons of the next layer over the channels. Therefore, the data is being propagated to the next channel in what is called a forward propagation. Then the neurons with the highest value determines the course of the output (Karaca, 2016)

The neural network has to be trained and usually in the process, the input and output are both being fed to the neural networks (Bello et al, 2010). In this case, the predicted output is being compared to the actual output. In order for the error being made in prediction to be realized. When the error is realized, the magnitude of the error which is how far the predicted value differs from the actual value is used to indicate how wrong the network is positive or negative, it suggest if the error is higher or lower which is the direction usually indicated by arrows (Karaca, 2016).

This information shows how to reduce the error. The information is then transferred backward through our network in what is known as the back propagation. The weights are further adjusted in order to suit the model. Forward and backward propagation is usually being performed with multiple inputs and the process is being repeated until the weights are being corrected in such a way that the network can predict the output correctly. Neural networks can take a long time to train depending on its complexity, the time differs.

Figure 2.3: An ANN model.



2.2.1 Types of ANN

- Feed forward neural network: this is a simple type of neural network with no back propagation, here the data moves in one direction through different nodes until it reaches the output layer the sum of the products of inputs together with their weights are being fed to the output layer.
- Radial basis function: this neural network has two layers it considers the distance of any point relative to the center point. The features are combined with the radial basis function in the inner layer. The output of this is taken into use when calculating the same output the next time.
- Multilayer perceptron: here every single node in a layer is connected to each node in the following layer, this is a type of ANN that can be said to be fully connected. Usually used in classification of data that can not be linearly separated.
- Convolution neural network (CNN): this uses a variation of the multilayer perceptron's. A CNN contains one or more than one convolutional layers that can either pooled or be completely interconnected. The convolution network uses the convolution operation on the input before passing it to the next layer.

- Recurrent neural network: this is a type of ANN in which the output of a particular layer is saved and fed back to the input. In order to help predict the outcome of the output layer. The first layer is being formed like that of a feed forward ANN.
- Modular neural network: this consist of a number of different networks that function independently and work in performing sub-tasks. There is usually no interaction with the different networks they don't even signal each other during the computation process. Therefore they work independently towards achieving the output.
- Sequence to sequence models: this model is especially applicable in case whereby the input and output data do not have the same length. It consists of two recurrent neural networks. Where an encoder that processes the input and a decoder that processes the output. The encoder and decoder can either use the same or different parameters.

2.3 Types of Activation Functions

- Identity
- Softmax
- Sigmoid
- Hyperbolic tangent

2.4 Applications Of Neural Networks

- Facial recognition
- Weather forecasting like detecting weather changes and chances of rainfall happening or detecting the rises and falls in stock prices.
- Undersea mine detection
- Music composition
- Speech recognition
- 3d objects recognition and texture analysis
- Hand written words recognition
- Prediction and diagnosis of medical problems such as hepatitis
- Data validation
- Industrial process control and risk management.

2.5 Artificial Neural Networks for Prediction of High Risk Preterm Birth

The previous commonly used procedure of finding out if a patient was at a high risk of giving birth pre term was through the procedure of the use of clinical testing's that were very costly and invasive on maternal populations that were high risks (Catley et al, 2010). The use of artificial neural networks (ANN) is a reengineered way of approach to the early prediction of preterm birth. It is therefore taken in as a complimentary technique. ANNs are employed as a tool for screening for preterm birth on a population of mothers that is heterogeneous. In order to estimate the risks, obstetrical variables are used that are being made available to the physicians before 23 weeks of gestation. The goal of these is to make an assessment if ANNs have a potential use in the estimation of obstetrical outcome in both high and low risks maternal populations (Ertuğrul, 2018). The backpropagation feedforward ANN will be trained and tested on some cases with a certain number of input variables that explain the obstetrical history of the patients. The output variables will be full term, very pre-term birth and late/moderate preterm birth.

Thousands of variables are usually examined when trying to determine the risk factors following preterm birth as we know there are a lot of risk factors associated with preterm birth. The risk factors all range within the scope of the basic characteristics, environment, medical history, occupational factors of parents, and infant related variables. However there are some factors that can be considered the most important risk factors related to pre-term birth on a general global scale as they are some of the most recurrent factors. This factors include, multiple birth, hemorrhage during pregnancy, age, previous preterm history, disease, body weight before pregnancy and height of pregnant women, and paternal life style risk factors related to drinking and smoking. These are the factors that could be used as the input variables for prediction in our ANN models. Therefore findings from this study will be useful for medical workers, public health workers and parents in detecting the high risk of pre-term birth in a maternal population, and this information will be used to provide early prevention or to take preventive measure that will reduce the occurrence of pre-term birth.

Knowing that ANNs are computerized systems that reveal previously weak relationships or unrecognized relationships between both the output and input variables through the use of a statistical analysis that is nonlinear. These ANNs are adaptive models that connect and have more than one layer as was in initially discussed. This means that when presented with various datasets they have the capability of dynamically finding the connections between these datasets (Ertuğrul, 2018). When a certain phenomenon is given and an ANN has been trained with use of appropriate data to select or determine the rules that will better expatiate the phenomenon it also can make some correct generalizations responding correctly to data it has not yet processed (Barron, 1991). This means that ANNs have the capability to make generalizations. In other words, after a certain period of training and learning, the ANN will have the ability to predict what the output will be on inputs of future cases that are unknown. ANNs have a pretty outstanding superiority as opposed to the use of classic logistic regression (Ripley, 1996). They have proved thus most especially in cases with weak linear correlation indexes among dependent and independent variables which can impair the logistics modelling and in cases of excessive collinearity among independent variables (Geman et al, 1992).

The performance indexes of an ANN method has to also be evaluated after carrying out the ANN procedure and having the assurance that the model has a great degree of accuracy in being able to predict PTB(pre term birth) some of the performance indexes we will be analyzing are,

- Sensitivity: this is a statistical measure that measures the performance of a classification test that is binary. Sensitivity is also known as the recall, true positive rate (TPR) or probability of detection. It works to measure the proportions of positives that can actually be correctively identified as being positives. For example the percentage of preemies that are correctly identified as being pre-term births.
- Specificity: this is also a statistical measure that measures the performance of a binary classification test. It is known as the true negative rate (TNR) . It measures the proportion of negatives that are correctly identifies as negatives. For example the percentage of full term babies that are correctly defined as being non preterm.
- Accuracy : this refers to how close a measured value is to the actual value or the known value if there is a big difference between this two it can be said that the measured value isn’t accurate. This can be used to measure the accuracy of the ANN model.
- P value: this is the probability value it is used to determine the significance of the results, it is used in hypothetical testing.

- Area under the ROC curve: this is a measure of the rate at which a model can distinguish between two groups correctly like normal/ disease in a case of a diagnostic group. Usually sensitivity is plotted against specificity.

The use of the ANNs are very important move for predictive and diagnostic medicine, technology and medicine keep enhancing to make quality of life better for humans. A clinical decision support system (CDSS) is a designed system of health information technology that provides physicians and other health professionals with a system of assistance with their tasks of clinical decision making which is clinical decision support (CDS). CDSS is a major topic or a major sector where artificial intelligence is constantly in use in medicine. "CDSS link observations made in medicine and healthcare with health knowledge that help to make an influence choices in health care. To improve medicine and health care.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This shows the various methods, and techniques that were used in carrying out this research. From the data collection, data sampling, processing and analysis.

3.2 Study design

This is a primary research study

3.3 Sample selection

This sample was selected using simple random sampling.

3.4 Methods

The data used in this study was obtained from the maternal and obstetrics records of the labor ward of Elite Ivf and Research Hospital, Lefkosia, North Cyprus. The data was gotten from the duration of 2015 to 2019. This information was obtained from 231 women who had deliveries in this duration, the information used was on 218 women which had delivered live births. 153 (70.2%) of the births were full term, 27 (12.4%) were moderate/late preterm and 38 (17.4%) were very preterm. Here labor is defines as regular uterine contractions with changes in cervix. 13 cases where excluded in total of which cases of indicated preterm birth were excluded and vaginal induced delivery was excluded so in other words all induced deliveries were excluded and caesarean deliveries whereby the woman went through labor was included. 7 independent variables which serve as our risk factors were used as the input, namely; in vitro fertilization (IVF), multiple birth (multbirth), gestational diabetes (gestdiabetes), hypertension, premature rupture of membranes (PROM), body mass index (BMI) classified age (ageclassified).

3.5 Data collection

The data was collected from the hospitals stored record files. This data was collected from the open reserves maternal, fetal/neonatal and obstetrics records.

3.6 Analysis of data

The collected data was further being analyzed through the use of artificial neural networks which is a computational model that has the ability to carry out various functions of which in this case is prediction. The ANN works as a predictive model that will be able to predict very preterm or full term birth from a number of risk factors which will be the input variables. The data collected will be analyzed by using statistical packages. We made the use of "IBM SPSS 25". The type of ANN used for this analysis is the multilayer perceptron neural network, the activation function that was used in the hidden layer is the hyperbolic tangent, this activation function is a nonlinear activation function that is an alternative to the sigmoid activation function, it is also known as the tan h activation function and the derivative of this activation function is used for calculation of error effects on weights. The derivative of the hyperbolic tangent function comprises of a simple form and that is why it is usually common in ANNs. This derivative can be gotten by:

tanh = sinh/cosh   tanh=y and;

$\sinh(x) = (e^{x} – e^{-x}) / 2$

$\cosh(x) = (e^{x} + e^{-x}) / 2$

$$y = \frac{\frac{e^{x}-e^{-x}}{2}}{\frac{e^{x}+e^{-x}}{2}} = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$ we find the derivatives;

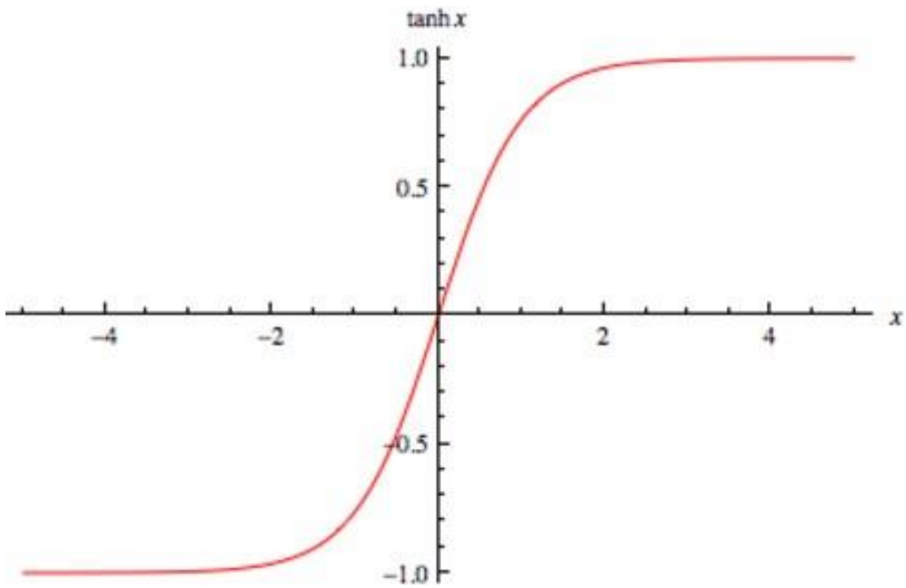
$$= \frac{d\,\sinh}{dx} = \frac{e^{x}+e^{-x}}{2} = \cosh$$

$$\frac{d\,\cosh}{dx} = \frac{e^{x}-e^{x}}{2} = \sinh$$ therefore going back to the calculation of tanh

$$\frac{dy}{dx} = \frac{\frac{d\,\sinh}{dx}.\cosh - \frac{d\,\cosh}{dx}.\sinh}{cosh^{2}} = \frac{cosh^{2} - sinh^{2}}{\cosh^{2}} \;\; ;$$

$\frac{dy}{dx} = 1\text{-}y^2 = 1\text{- } \tanh^2$

Fig: 3.1: Graph showing a hyperbolic tangent function.



CHAPTER 4

RESULTS

This discusses the ANN analysis and the results gotten from the analysis on the data of maternal and obstetrics records of the ELITE IVF and research hospital from August 2018 to November 2019 on the use of an artificial neural network to predicting preterm birth based on their high risk factors which are the variables. In this data we had 218 women which had delivered uninduced live births. 153 (70.2%) of the births were full term, 27 (12.4%) were moderate/late preterm and 38 (17.4%) were very preterm. Below is a table showing the descriptive statistics of the variables.

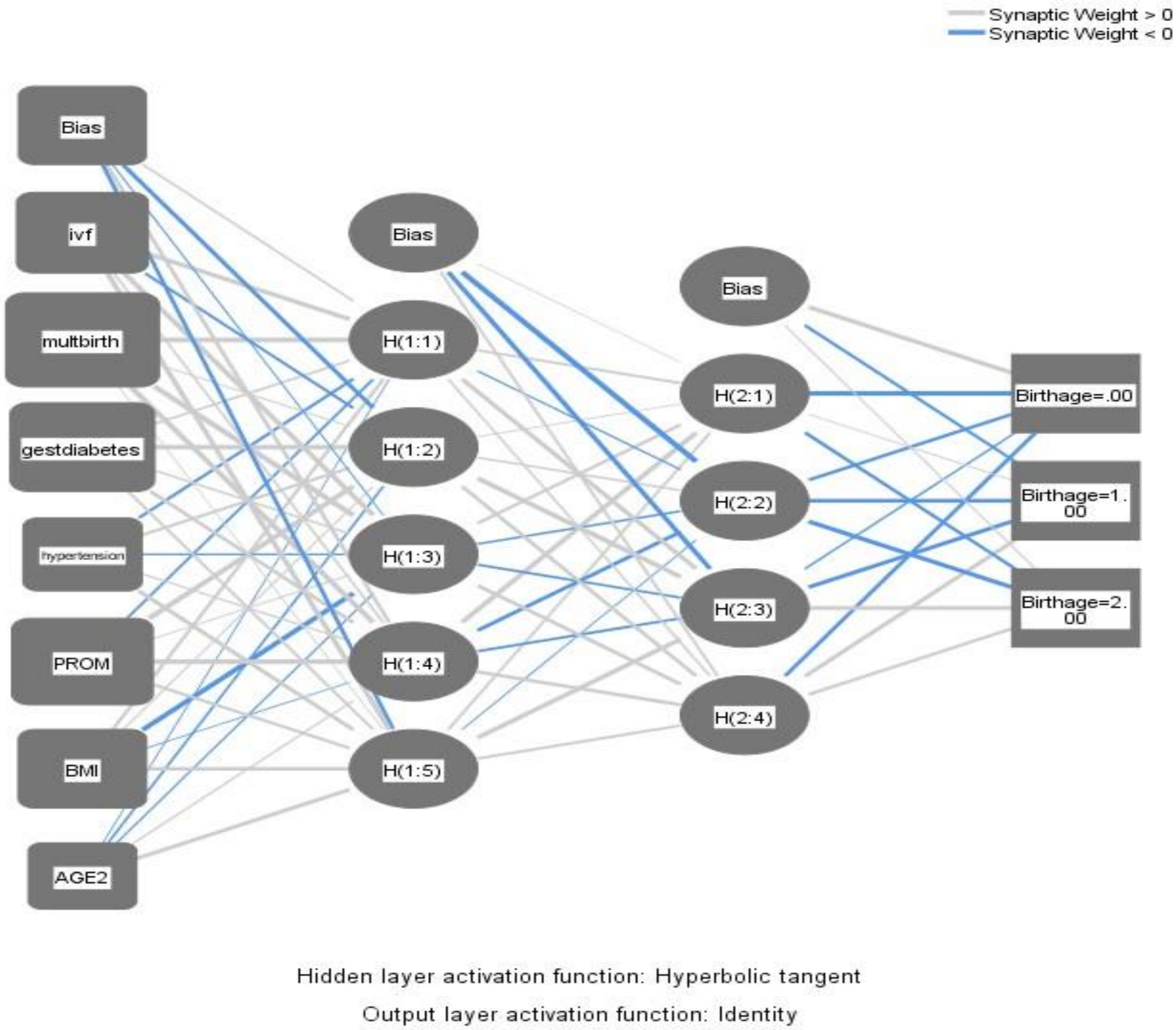
Table: 4.1: A table showing the descriptive statistics.

variables	levels	frequency	percentage
Birth age	Full term	153	70.2
	Moderate/late preterm	27	12.4
	Very preterm	38	17.4
IVF	No	170	78.00
	Yes	48	22.00
Mult birth	Single birth	187	85.8
	Multiple birth	31	14.2
Gest diabetes	Negative	197	90.4
	Positive	21	9.6
Hypertension	Negative	199	91.3
	Positive	19	8.7
PROM	Negative	189	86.7
	Positive	29	13.3
BMI	Underweight	5	2.3
	Normal weight	139	64.1
	Overweight	59	27.2

	Obese	14	6.5
Age classified	Younger	71	32.6
	midage	101	46.3
	Older	46	21.1

The dependent variable at the output layer which is the results of our predictions was labelled birthage with fullterm birth =0.00, moderate/late preterm birth = 1.00 and very preterm birth =2.00. The risk factors associated with preterm birth which served as the input independent variables, in vitro fertilization (IVF), multiple birth (multbirth), gestational diabetes (gestdiabetes), hypertension, premature rupture of membranes (PROM), body mass index (BMI) range 18-25, classified age (ageclassified). Is shown in the above table with their individual groupings. The age had a mean of 32.17±4.57 and the body mass index had a mean of 24.55±3.26 this shows that the data had a majority of midage women with a lot of normal weighted women which were in delivery. The system was divided with 70% being in the training set and 30% being in the test set.

Figure 4.1: ANN model of preterm birth.



The figure above shows the ANN model for this analysis, it had 7 input independent variables, with two hidden layers and 1 output layers. It has 3 biases for the input layer, hidden layer 1 and hidden layer 2 respectively, hidden layer 1 has 5 units and hidden layer 2 has 4 units. It used a hyperbolic tangent activation function for the hidden layers which is a non linear activation and as such best suited for the ANN, and an identity activation function which is a linear activation function was used for the output layer. The connecting lines signify the weights and the grey lines mean the weights are >0 and the blue lines signify that the weights are <0 which is the amplitude or strength of connection between the two neurons. the lines show the relations that were estimated, the darker the blue line and the bigger the line is, the stronger the relations, the bias serves as some kind of error term and we can see that the bias of the input layer where the line is strongest lies with the unit H(1:5) and H(1:2) and these goes for the bias in hidden layer 1 which lies with H(2:2) and H(2:3). The bias that lies with the output layer has the strongest effect on birthage 1.00 which is late/moderate preterm birth..

Table 4.2: A table showing training and test set.

	Number of values (N)	Percentage	Sum of squares error	Percent incorrect predictions (%)
Training set	152	70	18.912	16.4
Test set	65	30	7.925	18.5

The table above shows characteristics of the training and test set, this model took the time interval of 0:00:00:09 to be trained. It stopped when there was 1 step with no decrease in the error. Of the 70% that was the training set, 16.4% had incorrect predictions and the test set had 18.5% incorrect predictions as it didn't train with the testing cases. This table shows a summary of the quality of our ANN

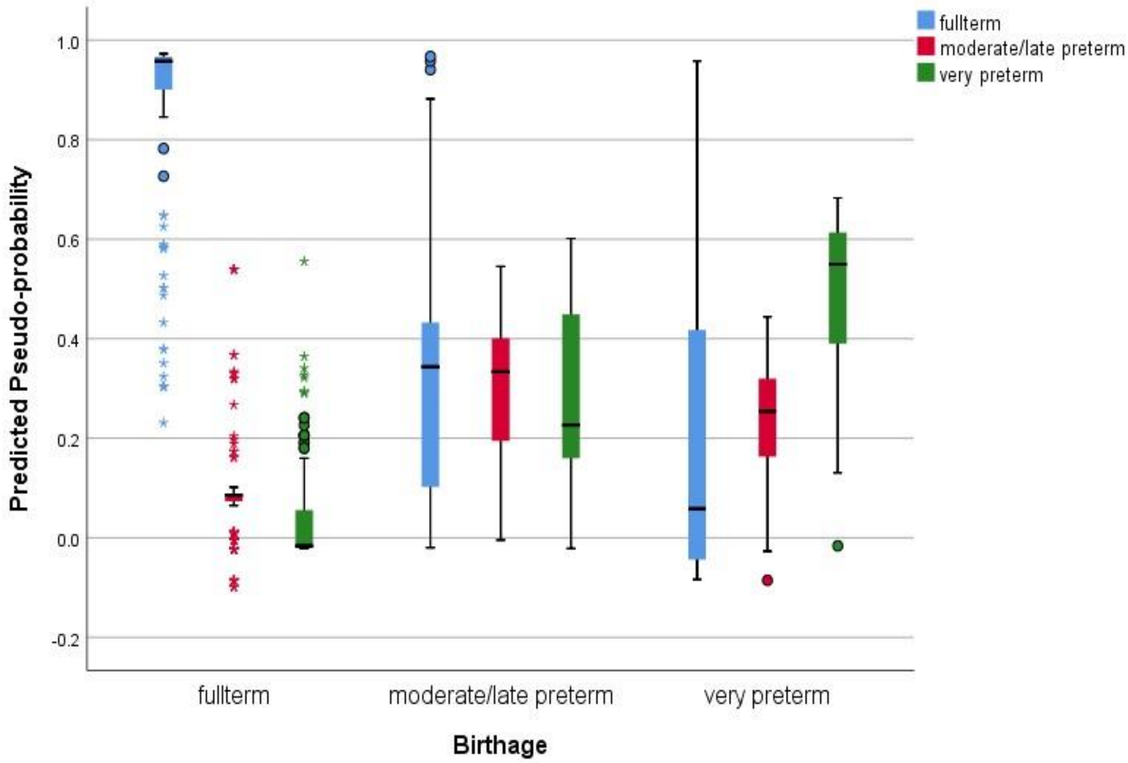
Table 4.3: Classification table.

		Predicted			
Sample	Observed	Full term	Moderate/late preterm	Very preterm	Percent correct
Training	Full term	103	1	1	98.1%
	Moderate/late preterm	10	6	7	26.1%
	Very peterm	5	1	18	75.0%

	Overall percent	77.6%	5.3%	17.1%	83.6%
Testing	Full term	43	3	1	91.5%
	Moderate/late preterm	1	1	2	25.0%
	Very preterm	4	1	9	64.3%
	Overall percent	73.8%	7.7%	18.5%	81.5%

The table above shows the classification scheme that tells us where the different wrong assignments occurred. In the training phase for full term birth they were 2 cases that the model wrongly assigned, and for the moderate/late preterm birth, there were 17 cases that were wrongly assigned this had the most error as shown earlier. For the very preterm birth, it had 6 wrongly assigned cases with a percentage correct of 75%. For the testing phase; note that the data used for the training set is different from the set used for testing. The observed set has a lower percentage, for the full term birth set 4 cases were wrongly predicted and 43 cases. For the moderate/late preterm 3 cases were wrongly predicted and 1 case were predicted correctly and for the very preterm cases 5 cases were wrongly predicted and 9 cases were correctly predicted.

Figure 4. 2: A box plot showing the ANN pseudo probability.



The figure above shows a box plot of the representation of the output. The blue colored boxes were full term birth, the red colored boxes are moderate/late preterm and the green colored boxes are for very preterm birth. This shows the predicted pseudo probability.

Figure 4.3: An ROC curve of the ANN model

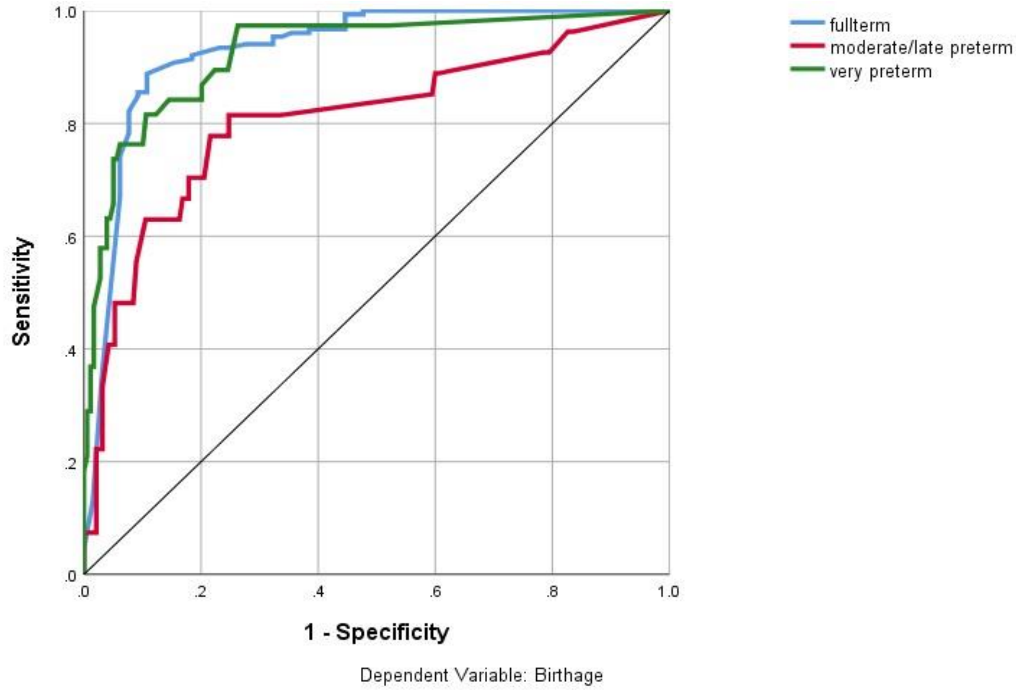


Table 4.4 Area Under the Curve

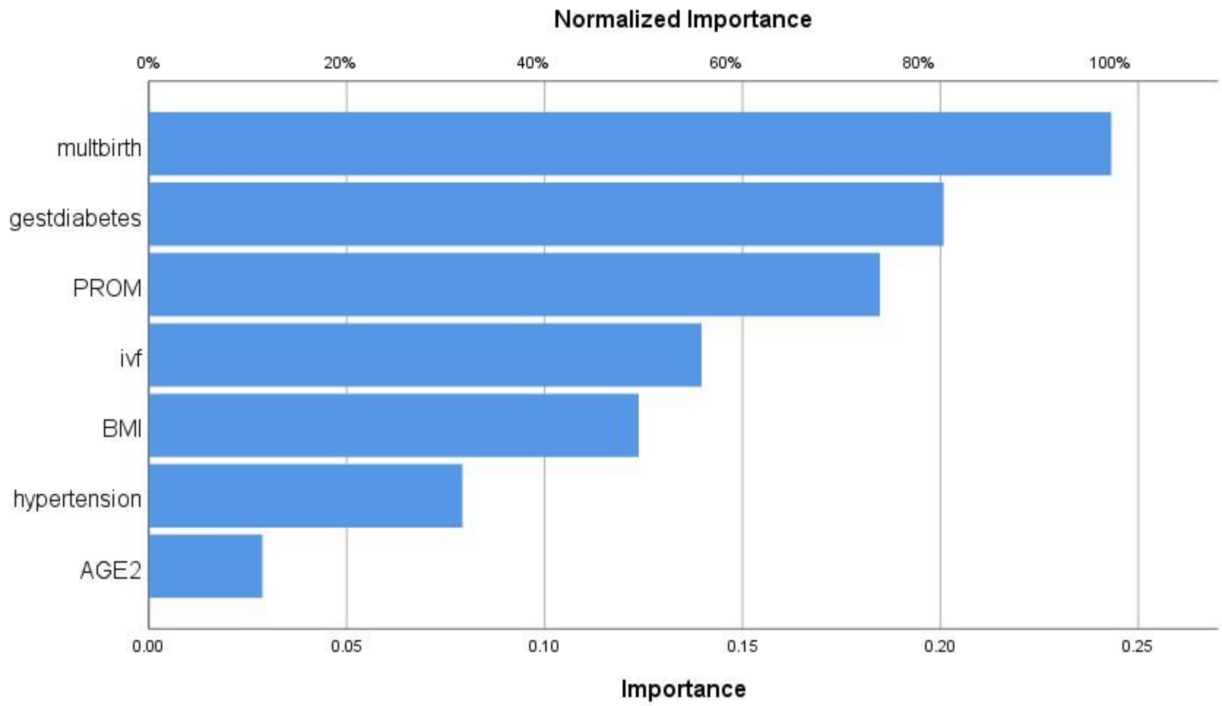
		Area
Birthage	Fullterm	.930
	Moderate/late preterm	.806
	Very preterm	.925

The figure above shows the ROC curve , from the graph it gives quite a clear representation of the accuracy of the test, the ROC curve shows the tradeoff between the sensitivity and the false positive rate (FPR) we can see from the curve that the full term birth is closer to the top left corner. This also shows the difference in their sensitivity respectively with full term births having greater sensitivity and as the sensitivity increases 1 point the FPR decreases. Form the AUC we can see the accuracy of the test.

Table 4.5: Independent Variable Importance.

	Importance	Normalized importance
IVF	.140	57.4%
Multiple birth	.243	100.0%
Gestational diabetes	.201	82.6%
Hypertension	.079	32.6%
Premature rupture of membranes	.185	76.0%
BMI	.124	50.9%
Age 2	.029	11.8%

Figure 4.4: A bar graph of the normalized importance



Above shows the figure and the table showing the importance of each of the independent variables as evaluated by the ANN , from the above we can see that multiple births has a greater importance 100% in the classification of birthage based on the data. This means that it plays a major role in the birthage being fullterm, moderate/late preterm or very preterm. Followed by gestational diabetes and which is of 8.2.6% importance, then premature rupture of membranes (PROM) which is 76% and then IVF which is 57.4% then BMI, followed by hypertension and age2 which plays the least important role.

CHAPTER 5

CONCLUSIONS

5.1 Discussions

From the ANN model we could see that the data had a majority of midage women with a lot of normal weighted women which were in delivery from observations, there was a stronger connection from the error function of the bias to the moderate/late preterm which shows that the model had more difficulty classifying the moderate/ late preterm and this fact goes further to prove itself in the classification table where we could see that the training and test sets had higher number of incorrectly predicted cases at the point of the moderate/late preterm, the data had a majority of midage women with a lot of normal weighted women which were in delivery from observations, this problem could be from the dataset and also because it is a hard tax predicting preterm birth as preterm birth doesn't have fixed proven risk factors, that is to say that a pregnant woman can deliver preterm and not exhibit any of this risk factors that were shown above as the input variables or not even exhibit any symptoms or risk factors at all. This is because preterm birth is not easily medically diagnosed. The aim of this study is to be able analyze and predict preterm birth from an early stage using ANNs. The bias that lies with the output layer has the strongest effect on birthage 1.00 which is late/moderate preterm birth. Late/ moderate preterm birth has a lot of risk factors and anything could easily trigger such preterm birth which is part of the reasons it was more complex for the ANN to train and predict this group. There are a lot of risk factors that determine preterm birth, all at different levels of importance. And due to the data obtained, the input variables were selected based on the most important risk factors commonly associated with preterm birth. The ANN could predict very preterm and full term birth much more easily because in order to give birth very preterm it is mostly associated with more distinct causes so the ANN could easily identify that, and the data had majority of full term birth, as it is the most commonly occurring births. The quality of the ANN model was on average because the mean squared errors and the percentage wrong predictions we had were on average, it was able to show the predictive ability of the ANN. The ROC and AUC show the models accuracy which is satisfactory, from the table and bar chart showing the variable importance, we could see that multiple births had the greatest importance and age had the least importance the levels of importance shown was quite true because, the effect of age isn't much And multiple birth also has a great effect on preterm birth, hypertension should normally have much of an effect than BMI thus there are also some misgivings about this importance level. It isn't as accurate but it gives a good representation of the relationship between the variables and the output in the ANN. The ANN has a good future in predicting preterm birth, of course there needs to be more study and improvements to this but in the near future the ANN will be able to solve the problems of preterm birth in the world and also many diagnostic c problems in predictive medicine too, in our world today, ANNs are playing a major role in predictive medicine. Early prediction is very important because it aids prevention in the sense the preventive measures will be taking before it is even late.

REFERENCES

Allen, S., Belcher, C., & Newnham, J. (2018). The Preterm Birth Prevention Initiative - safely lowering the rate of Preterm Birth in Western Australia. *Women And Birth, 31,* S53. doi: 10.1016/j.wombi.2018.08.158

Aggarwal, R. and Yonghua Song (1998). Artificial neural networks in power systems. Part 2: Types of artificial neural networks. *Power Engineering Journal*, 12(1), pp.41-47.

Artificial neural systems engineering and analysis. (1988). *Neural Networks*, 1, p.434.

Assessing Pre-term Labor and Birth Risk. (2004). *AWHONN Lifelines*, 8(1), p.45.

Bello, G., Rodríguez, M., Olivieri, M., Rosemberg, E., & Bronfenmajer, S. (2010). Artificial neural networks for the prediction of IVF outcome using clinical and morphological data. *Fertility And Sterility*, *94*(4), S250. doi: 10.1016/j.fertnstert.2010.07.966

Bick, D. (2012). Born too soon: The global issue of preterm birth. *Midwifery*, 28(4), pp.401-402.

Catley, C., Frize, M., Walker, R. and Petriu, D. (2006). Predicting High-Risk Preterm Birth Using Artificial Neural Networks. *IEEE Transactions on Information Technology in Biomedicine*, 10(3), pp.540-549.

Ertuğrul, Ö. (2018). A novel type of activation function in artificial neural networks: Trained activation function. *Neural Networks*, *99*, 148-157. doi: 10.1016/j.neunet.2018.01.007

Goldenberg, R. (1996). Predicting Preterm Birth. *ACOG Clinical Review*, *1*(5), 5. doi: 10.1016/1085-6862(96)85126-0

Goldenberg, R., Iams, J., Mercer, B., Meis, P., Moawad, A., Copper, R., Das, A., Thom, E., Johnson, F., McNellis, D., Miodovnik, M., Van Dorsten, J., Caritis, S., Thurnau, G. and Bottoms, S. (1998). The preterm prediction study: the value of new vs standard risk factors in predicting early and all spontaneous preterm births. NICHD MFMU Network. *American Journal of Public Health*, 88(2), pp.233-238.

Gupta, S. (2019). Artificial neural network modeling and exposure assessments: a new scaling approach. *Human and Ecological Risk Assessment: An International Journal*, pp.1-20. Howson, C., Kinney, M., McDougall, L. and Lawn, J. (2013). Born Too Soon: Preterm birth matters. *Reproductive Health*, 10(S1).

Johnson, S. (2015). Understanding social development following very preterm birth. *Developmental Medicine & Child Neurology*, 57(10), pp.890-890.

Johnson, S., Gilmore, C., Gallimore, I., Jaekel, J. and Wolke, D. (2015). The long-term consequences of preterm birth: what do teachers know?. *Developmental Medicine & Child Neurology*, 57(6), pp.571-577.

Karaca, Y. (2016). Case study on artificial neural networks and applications. *Applied Mathematical Sciences*, *10*, 2225-2237. doi: 10.12988/ams.2016.65174

Kim, Y. (2019). Analysis of Spontaneous Preterm Labor and Birth and Its Major Causes Using Artificial Neural Network. *Journal of Korean Medical Science*, 34(16).

Lamont, R. and Jaggat, A. (2007). Emerging drug therapies for preventing spontaneous preterm labor and preterm birth. *Expert Opinion on Investigational Drugs*, 16(3), pp.337-345.

Lawn, J., Kinney, M., Belizan, J., Mason, E., McDougall, L., Larson, J., Lackritz, E., Friberg, I. and Howson, C. (2013). Born Too Soon: Accelerating actions for prevention and care of 15 million newborns born too soon. *Reproductive Health*, 10(S1).

Lee, K. and Ahn, K. (2019). Artificial Neural Network Analysis of Spontaneous Preterm Labor and Birth and Its Major Determinants. *Journal of Korean Medical Science*, 34(16).

Maner, W. and Garfield, R. (2007). Identification of Human Term and Preterm Labor using Artificial Neural Networks on Uterine Electromyography Data. *Annals of BiomedicalEngineering*, 35(3), pp.465-473.

M, D. (2018). Study of Risk Factors for Preterm Labour. *Journal Of Medical Science And Clinical Research*, *6*(2). doi: 10.18535/jmscr/v6i2.53

Morken, N., Källen, K. and Jacobsson, B. (2013). Predicting Risk of Spontaneous Preterm Delivery in Women with a Singleton Pregnancy. *Paediatric and Perinatal Epidemiology*, 28(1), pp.11-22.

Simhan, H. and Canavan, T. (2005). Preterm premature rupture of membranes: diagnosis, evaluation and management strategies. *BJOG: An International Journal of Obstetrics & Gynaecology*, 112, pp.32-37.

Stout, M., Macones, G. and Tuuli, M. (2016). 329: Preterm birth classification: Is birth certificate data adequate?. *American Journal of Obstetrics and Gynecology*, 214(1), pp.S185-S186.

Thomas, J. and Kulanthaivel, G. (2017). Optimized Rule Generation for Preterm Birth Prediction Using Soft Computing. *International Journal of Data Mining And Emerging Technologies*, 7(2), p.43.

Sen, C. (2017). Preterm labor and preterm birth. *Journal of Perinatal Medicine*, 45(8).

Shennan, A. (2018). Prediction and prevention of preterm birth: a quagmire of evidence. *Ultrasound in Obstetrics & Gynecology*, 51(5), pp.569-570.

Stower, H. (2018). Predicting preterm birth. *Nature Medicine*, *24*(8), 1088-1088. doi: 10.1038/s41591-018-0154-z

Suzuki, S. (1998). Constructive function-approximation by three-layer artificial neural networks. *Neural Networks*, 11(6), pp.1049-1058.

Thomas, J. and Kulanthaivel, G. (2010). Data Mining Approach in Preterm Birth Prediction. *Mapana - Journal of Sciences*, 9(1), pp.18-30.

Zernikow, B., Holtmannspoetter, K., Michel, E., Pielemeier, W., Hornschuh, F., Westermann, A. and Hennecke, K. (1998). Artificial neural network for risk assessment in preterm neonates. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 79(2), pp.F129F134.

APPENDICES

APPENDIX 1

Convolusion neural network