# ÖZET

VERİ MADENCİLİĞİ YÖNTEMLERİ KULLANILARAK YÜKSEK TANSİYON HASTALIĞI İÇİN İLAÇ DOZU PLANLANMASI

Çağdaş ÇALIŞ

Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi: Doç. Dr. Adem Karahoca

Haziran 2010, 58 sayfa

Yüksek Tansiyon dünyadaki en yaygın hastalıklardan birisidir. Tedavi sürecinde hastaların sabırlı ve dikkatli olmaları gerekmektedir, çünkü çoğu zaman tedavi süreci sancılı geçmektedir ve bu süreç gelecek yaşamları açısından önemlidir. Düzenli olarak tansiyonlarının ölçülmesi ve doktor gözetiminde ilaç kullanmaları gerekmektedir. Doktorlar genelde ilaç dozuna karar verirken hastanın yaşını, boy kilo endeksini, genetik durumunu ve idrar tahlili sonuçlarını göz önünde bulundurmaktadırlar.

Bu çalışmanın amacı, veri madenciliği teknikleri kullanarak yüksek tansiyon hastalarının ilaç dozlarını planlamaktır. Bu çalışmada ANFIS ve Rough Set veri madenciliği yöntemleri kullanılmıştır. Kullanılan giriş parametreleri cinsiyet, boy kilo endeksi, hba, tansiyon, şeker, kolestrol, kandaki keton, kandaki protein, kandaki mikalb, kandaki mikros, genetik ve ilaç dozudur. Dozaj planlaması Coversyl, Monopril, Tenoretic ve Atacand ilaçları üzerinde yapılmıştır.

Sonuç olarak ANFIS'in RSES'den daha iyi sonuçlar verdiği gözlenmiştir. ANFIS, dosaj planlaması yapılırken en güvenilir yöntemdir.

**Anahtar Kelimeler:** ANFIS, rough set, yüksek tansiyon, dosaj planlama, veri madenciliği.

# ABSTRACT

DRUG DOSAGE PLANNING OF HYPERTENSION DISEASE USING DATA MINING TECHNIQUES

Çağdaş ÇALIŞ

M.S. Department of Computer Engineering
Supervisor: Assoc. Prof. Dr. Adem Karahoca

June 2010, 58 pages

Hypertension is one of the most frequently seen disease all over the world. During the treatment patients must be patient and careful, because sometimes treatment process can be very hard and it's important for their future life. Their blood pressure must be measured regularly and use drugs for compensate blood pressure levels. Generally, doctors decide the dosage of used drug with patient's age, body mass index, genetic condition and urine test results.

The aim of this study is, drug dosage planning of hypertension patients using data mining techniques. In this study, ANFIS and Rough Set data mining methods are used. Input parameters are gender, BMI, hba, blood pressure, glucose, cholesterol, ketone, protein, mikalb, mikros, genetic and dosage of drugs. Dosage planning is applied on drugs Coversyl, Monopril, Tenoretic and Atacand.

As a result, ANFIS give better results than RSES. ANFIS is the most reliable data mining technique for dosage planning.

**Key Words:** ANFIS, rough set, hypertension, dosage planning, data mining.

# 1. INTRODUCTION

## 1.1 PROBLEM DEFINITION

Hypertension is one of the most frequently seen disease all over the world and in Turkey. Generally, if the hypertension is not at risky levels or the patients don't know that it's at risky levels, they don't care too much about it, but treatment is very important for patient's future life. Age and weight is strongly affect hypertension, so the patients that is obese or older ages must be more careful and responsible about disease or instantaneous blood pressure rises can result with deaths.

The treatment process is not easy. Especially if the patients are over middle-age, their blood pressure must be measured regularly, and use drugs for compensate blood pressure levels. One of the most important part of the treatment is to determine the correct amount of drug dosage. This thesis aim is to plan drug dosage effectively.

## 1.2 WHAT IS HYPERTENSION?

Hypertension is a common disorder in which blood pressure remains abnormally high (a reading of 140/90 mm Hg or greater) [2].

Arteries are the blood vessels that carry the blood from the heart through the entire body. High blood pressure results either when the output of the blood pumped by the heart increases, or when there is an increased resistance to the flow of blood through the arteries, or both.

A blood pressure reading has two numbers. The first, or upper, number measures the pressure in your arteries when your heart beats (systolic pressure). The second, or lower, number measures the pressure in your arteries between beats (diastolic pressure). In terms of numbers, a resting blood pressure of 140/90 or greater in an adult is usually considered

to be high. Normal blood pressure levels are lower in children and rise with age (MFMER, 1998).

Everyone's blood pressure goes up in moments of excitement or stress, and that is considered to be normal. High blood pressure is considered a medical condition only when it continues over an extended period of time. This condition can then become a serious threat to health; the higher the pressure and the longer it is untreated, the greater the risk.

People who have hypertension are more likely to suffer a stroke, heart attack, or failure of the kidneys or heart. For this reason, and because there usually are no symptoms, hypertension has been termed "the silent killer." (http://www.humanillnesses.com/original/Her-Kid/Hypertension.html)

### 1.2.1 Types of Hypertension

There are four general types of high blood pressure.

- Prehypertension
- Essential hypertension
- Secondary hypertension
- Isolated systolic hypertension

#### 1.2.1.1 Prehypertension

Prehypertension is a condition between normal blood pressure and high blood pressure (hypertension). Because there are often no symptoms, people typically don't realize they have it until their blood pressure readings are too high [28].

Prehypertension is considered to be blood pressure readings with a systolic pressure from 120 to 139 mm Hg or a diastolic pressure from 80 to 89 mm Hg. Classification of blood pressure is based upon two or more readings at two or more separate occasions (HS. Kelly, 2008 ).

A primary risk factor for prehypertension is being overweight. Other risk factors include a family history of hypertension, a sedentary lifestyle, eating high sodium foods, smoking, and excessive alcohol intake. Blood pressure levels appear to be familiar, but there is no clear genetic pattern (Herb Valley, 2010).

Prehypertension is often asymptomatic (without symptoms) at the time of diagnosis. Only extremely elevated blood pressure (malignant hypertension) can, in rare cases, cause headaches, visual changes, fatigue, or dizziness, but these are nonspecific symptoms which can occur with many other conditions. Thus, blood pressures above normal can go undiagnosed for a long period of time (Besser Kyle).

#### 1.2.1.2 Essential Hypertension

Essential hypertension is the form of hypertension that by definition has no identifiable cause. It is the more common type and affects 90-95% of hypertension patients, it tends to be familiar and is likely to be the consequence of an interaction between environmental and genetic factors (Josh, 2010).

Even though there are no direct causes, there are many risk factors such as sedentary lifestyle, obesity (more than 85% of cases occur in those with a body mass index greater than 25), salt (sodium) sensitivity, alcohol intake and vitamin D deficiency. It is also related to aging and to some inherited genetic mutations. Family history increases the risk of developing hypertension. Renin elevation is another risk factor, Renin is an enzyme secreted by the juxtaglomerular apparatus of the kidney and linked with aldosterone in a negative feedback loop. Also sympathetic over activity is implicated. Insulin resistance which is a component of syndrome X, or the metabolic syndrome is also thought to cause

hypertension. Recently low birth weight has been questioned as a risk factor for adult essential hypertension (Riska Damayanti & Suwitto, 2009).

#### 1.2.1.3 Secondary Hypertension

Secondary hypertension (or, less commonly, inessential hypertension) is a type of hypertension which by definition is caused by an identifiable underlying secondary cause. This type accounts for approximately 5-10% of all cases of hypertension, with the remaining being primary hypertension (Lord Michael, 2010).

There are many known conditions that can cause secondary hypertension. Regardless of the cause, arterial pressure becomes elevated either due to an increase in cardiac output, an increase in systemic vascular resistance, or both. When cardiac output is elevated, it is generally due to either increased neurohumoral activation of the heart or increased blood volume (Nazareth Edward, 2008).

Some other causes for secondary hypertension are listed below:

- Disorders of the adrenal gland (small organs, located above the kidneys, that create hormones), including Cushing's syndrome (a condition caused by an overproduction of cortisol); hyperaldosteronism (too much aldosterone); and pheochromocytoma (a rare tumor that causes over secretion of hormones like adrenaline)
- Kidney disease which may include polycystic kidney disease, kidney tumor, kidney failure, or a narrow or blocked main artery supplying the kidney
- Drugs such as corticosteroids (anti-inflammatory drugs like prednisone), nonsteroidal anti-inflammatory drugs (Motrin, Aleve, Naprosyn, Celebrex), weight loss drugs (such as Meridia), cold medications that include decongestants, like pseudoephedrine, birth control pills (the estrogen component), and migraine medications (such as Imitrex).
- Sleep apnea, a condition in which a person has brief spells in which he stops

breathing during sleep. About half of patients with this condition have high blood pressure.

- Coarctation of the aorta, a birth defect in which the aorta is narrowed
- Preeclampsia, a condition related to pregnancy
- Thyroid and parathyroid problems (Robert J Bryg, 2009).

#### 1.2.1.4 Isolated Systolic Hypertension

In medicine, systolic hypertension is defined as an elevated systolic blood pressure. If systolic blood pressure is elevated (>140) with a normal diastolic blood pressure (<90), it is called "Isolated Systolic Hypertension" [18].

Systolic hypertension may be due to reduced compliance of the aorta with increasing age (Kennedy Ron). This increases the load on the ventricle and jeopardizes coronary blood flow, which can eventually result in left ventricular hypertrophy, coronary ischemia, and heart failure. Contemporary physics shows us an Immersed Boundary Method of computational illustration of a single heartbeat Template:Courant Intitute. Applied to physiologic models, immersed boundary theory sees the heart as a great folded semisolid Sail fielding and retrieving a viscous blood mass. The sail, likened to Windkessel physiology gives and receives a load under time ordered phases. Decreasing compliance of the sail heralds the onset of systolic hypertension [26].

Some signs and symptoms for Isolated Systolic Hypertension are listed below:

- Usually asymptomatic
- Headache
- Vision changes
- Heart palpitations
- Increased nighttime urination
- High systolic blood pressure reading

- Normal diastolic blood pressure reading

(http://www.wrongdiagnosis.com/i/isolated_systolic_hypertension/signs.htm)

### 1.2.2 Risk Factors

Hypertension is one of the most common complex disorders. The etiology of hypertension differs widely amongst individuals within a large population (ME Dickson, 2006). And by definition, essential hypertension has no identifiable cause. However, several risk factors have been identified.

Hypertension may be secondary to other diseases but over 95% of patients have essential hypertension which is of unknown origin. It is observed though that:

- Consuming foods that contain High Fructose Corn Syrup may indefinitely increase one's risk of developing hypertension (Science Daily, 2009).
- Having a personal family history of hypertension increases the likelihood that an individual develops HPT (J. Loscalzo & S. Anthony & E. Braunwald & L. Dennis & S. Hauster & D. Longo, 2008).
- Essential hypertension is four times more common in black than white people, accelerates more rapidly and is often more severe with higher mortality in black patients (Science Daily, 2009).

More than 50 genes have been examined in association studies with hypertension, and the number is constantly growing. One of these genes is the angiotensinogen (AGT) gene, studied extensively by Kim et al. They showed that increasing the number of AGT increases the blood pressure and hence this may cause hypertension (ME Dickson, 2006). Twins have been included in studies measuring ambulatory blood pressure; from these studies it has been suggested that essential hypertension contains a large genetic influence (ME Dickson, 2006). Supporting data has emerged from animal studies as well as clinical studies in human populations. The majority of these studies support the concept that the inheritance is probably multifactorial or that a number of different genetic defects each

has an elevated blood pressure as one of its phenotypic expressions. However, the genetic influence upon hypertension is not fully understood at the moment. It is believed that linking hypertension-related phenotypes with specific variations of the genome may yield definitive evidence of heritability. Another view is that hypertension can be caused by mutations in single genes, inherited on a Mendelian basis (M. Sutters, 2006).

Hypertension can also be age related, and if this is the case, it is likely to be multifactorial. One possible mechanism involves a reduction in vascular compliance due to the stiffening of the arteries. This can build up due to isolated systolic hypertension with a widened pulse pressure. A decrease in glomerular filtration rate is related to aging and this results in decreasing efficiency of sodium excretion. The developing of certain diseases such as renal micro vascular disease and capillary rarefaction may relate to this decrease in efficiency of sodium excretion. There is experimental evidence that suggests that renal micro vascular disease is an important mechanism for inducing salt-sensitive hypertension (T. Kosugi & T. Nakagawa & D. Kamath & RJ. Johnson, 2009).

Obesity can increase the risk of hypertension to fivefold as compared with normal weight, and up to two-thirds of hypertension cases can be attributed to excess weight. More than 85% of cases occur in those with a Body mass index greater than 25 (DW. Haslam & WP. James, 2005). A definitive link between obesity and hypertension has been found using animal and clinical studies; from these it has been realized that many mechanisms are potential causes of obesity-induced hypertension. These mechanisms include the activation of the sympathetic nervous system as well as the activation of the renin–angiotensin-aldosterone system(K. Rahmouni & ML. Correia & WG. Haynes & Al. Mark, 2005).

Another risk factor is salt (sodium) sensitivity which is an environmental factor that has received the greatest attention. Approximately one third of the essential hypertensive population is responsive to sodium intake (M. Katori & M. Majima, 2006). When sodium intake exceeds the capacity of the body to excrete it through the kidneys, vascular volume expands secondary to movement of fluids into the intra-vascular compartment. This

causes the arterial pressure to rise as the cardiac output increases. Local auto regulatory mechanisms counteract this by increasing vascular resistance to maintain normotension in local vascular beds. As arterial pressure increases in response to high sodium chloride intake, urinary sodium excretion increases and the excretion of salt is maintained at expense of increased vascular pressures (J. Loscalzo & S. Anthony & E. Braunwald & L. Dennis & S. Hauster & D. Longo, 2008). The increased sodium ion concentration stimulates ADH and thirst mechanisms, leading to increased reabsorption of water in the kidneys, concentrated urine, and thirst with higher intake of water. Also, the water movement between cells and the interstitium plays a minor role compared to this. The relationship between sodium intake and blood pressure is controversial. Reducing sodium intake does reduce blood pressure, but the magnitude of the effect is insufficient to recommend a general reduction in salt intake (G. Jürgens, 2004).

Renin elevation is another risk factor. Renin is an enzyme secreted by the juxtaglomerular apparatus of the kidney and linked with aldosterone in a negative feedback loop. In consequence, some hypertensive patients have been defined as having low-renin and others as having essential hypertension. Low-renin hypertension is more common in African Americans than white Americans, and may explain why African Americans tend to respond better to diuretic therapy than drugs that interfere with the Renin-angiotensin system. High renin levels predispose to hypertension by causing sodium retention through the following mechanism: Increased renin → Increased angiotensin II → Increased vasoconstriction, thirst/ADH and aldosterone → Increased sodium resorption in the kidneys (DCT and CD) → Increased blood pressure.

Hypertension can also be caused by Insulin resistance and/or hyperinsulinemia, which are components of syndrome X, or the metabolic syndrome. Insulin is a polypeptide hormone secreted by cells in the islets of Langerhans, which are contained throughout the pancreas. Its main purpose is to regulate the levels of glucose in the body antagonistically with glucagon through negative feedback loops. Insulin also exhibits vasodilatory properties. In normotensive individuals, insulin may stimulate sympathetic activity without elevating mean arterial pressure. However, in more extreme conditions such as that of the metabolic syndrome, the increased sympathetic neural activity may over-ride

the vasodilatory effects of insulin.

It has been suggested that vitamin D deficiency is associated with cardiovascular risk factors (JH. Lee & JH. O'Keefe & D. Bell & DD. Hensrud & MF. Holick, 2008). It has been observed that individuals with a vitamin D deficiency have higher systolic and diastolic blood pressures than average. Vitamin D inhibits renin secretion and its activity, it therefore acts as a "negative endocrine regulator of the renin-angiotensin system". Hence a deficiency in vitamin D leads to an increase in renin secretion. This is one possible mechanism of explaining the observed link between hypertension and vitamin D levels in the blood plasma (JP. Forman & E. Giovannucci & MD. Holmes, 2007).

Also, some authorities claim that potassium might both prevent and treat hypertension (EM. Hamilton & EN. Whitney & FS. Sizer, 1991).

Recent studies claims that obesity is a risk factor for hypertension because of activation of the renin-angiotensin system (RAS) in adipose tissue, and also linked renin-angiotensin system with insulin resistance, and claims that any one can cause the other (S. Saitoh, 2009).

Cigarette smoking, a known risk factor for other cardiovascular disease, may also be a risk factor for the development of hypertension (RO. Halperin, 2008).

### 1.2.3 How is Hypertension Diagnosed?

Diagnosis of hypertension is generally on the basis of a persistently high blood pressure. Usually this requires three separate measurements at least one week apart. Exceptionally, if the elevation is extreme, or end-organ damage is present then the diagnosis may be applied and treatment commenced immediately (R. Felix, 2008).

Obtaining reliable blood pressure measurements relies on following several rules and

understanding the many factors that influence blood pressure reading (R. Felix, 2008).

For instance, measurements in control of hypertension should be at least 1 hour after caffeine, 30 minutes after smoking or strenuous exercise and without any stress. Cuff size is also important. The bladder should encircle and cover two-thirds of the length of the (upper) arm. The patient should be sitting upright in a chair with both feet flat on the floor for a minimum of five minutes prior to taking a reading. The patient should not be on any adrenergic stimulants, such as those found in many cold medications.

When taking manual measurements, the person taking the measurement should be careful to inflate the cuff suitably above anticipated systolic pressure. The person should inflate the cuff to 200 mmHg and then slowly release the air while palpating the radial pulse. After one minute, the cuff should be reinflated to 30 mmHg higher than the pressure at which the radial pulse was no longer palpable. A stethoscope should be placed lightly over the brachial artery. The cuff should be at the level of the heart and the cuff should be deflated at a rate of 2 to 3 mmHg/s. Systolic pressure is the pressure reading at the onset of the sounds described by Korotkoff (Phase one). Diastolic pressure is then recorded as the pressure at which the sounds disappear (K5) or sometimes the K4 point, where the sound is abruptly muffled. Two measurements should be made at least 5 minutes apart, and, if there is a discrepancy of more than 5 mmHg, a third reading should be done. The readings should then be averaged. An initial measurement should include both arms. In elderly patients who particularly when treated may show orthostatic hypotension, measuring lying sitting and standing BP may be useful. The BP should at some time have been measured in each arm, and the higher pressure arm preferred for subsequent measurements (http://en.wikipedia.org/wiki/Talk:Hypertension).

BP varies with time of day, as may the effectiveness of treatment, and archetypes used to record the data should include the time taken. Analysis of this is rare at present (Selasa, 2010).

Automated machines are commonly used and reduce the variability in manually collected readings. Routine measurements done in medical offices of patients with known

hypertension may incorrectly diagnose 20% of patients with uncontrolled hypertension (Selasa, 2010).

Home blood pressure monitoring can provide a measurement of a person's blood pressure at different times throughout the day and in different environments, such as at home and at work. Home monitoring may assist in the diagnosis of high or low blood pressure. It may also be used to monitor the effects of medication or lifestyle changes taken to lower or regulate blood pressure levels. Home monitoring of blood pressure can also assist in the diagnosis of white coat hypertension.

Some home blood pressure monitoring devices also make use of blood pressure charting software. These charting methods provide printouts for the patient's physician and reminders to take a blood pressure reading. However, a simple and cheap way is simply to manually record values with pen and paper, which can then be inspected by a doctor.
Systolic hypertension is defined as an elevated systolic blood pressure. If systolic blood pressure is elevated with a normal diastolic blood pressure, it is called isolated systolic hypertension. Systolic hypertension may be due to reduced compliance of the aorta with increasing age (G. Dwivedi & S. Dwivedi, 2007).

Once the diagnosis of hypertension has been made it is important to attempt to exclude or identify reversible (secondary) causes. Secondary hypertension is more common in preadolescent children, with most cases caused by renal disease. Primary or essential hypertension is more common in adolescents and has multiple risk factors, including obesity and a family history of hypertension. Tests are undertaken to identify possible causes of secondary hypertension, and seek evidence for end-organ damage to the heart itself or the eyes (retina) and kidneys. Diabetes and raised cholesterol levels being additional risk factors for the development of cardiovascular disease are also tested for as they will also require management (GB. Luma & RT. Spiotta, 2006).

## 1.3 RELATED WORKS

Everybody could find articles about hypertension or high blood pressure with Data mining techniques, when they searches about Data mining on Internet or scientific libraries. Some related works about hypertension and data mining are found in the literature search, and mentioned about them in the following section.

When we investigated them, we saw they must be benchmarked and cannot meet our problem defined in the section 1.1.

Some unimportant data in sample data set spoil the classification and increase unwanted calculations leading to decrease in real-time capacity of the medical prediction. An improved hypertension prediction model based on rough set and support vector machine helps to get rid of above-cited problem. Rough set, as an anterior preprocessor of SVM, finds out relevant factors affecting the hypertension disease, and use them as the input vectors of SVM (Support Vector Machine). That is how a hypertension prediction model is conducted in this approach. The data set obtained from the survey covering the people ages from 35 to 92 in the Three Gorges Area refers to many aspects such as demography, medical history, smoke, economic status, education background, medical insurance, family inheritance history and many factors totaling to 133 factors.

The result of the experimental study showed that the hypertension prediction model based on RS and SVM in the Three Gorges Area is better and more efficient than the traditional methods. By minimizing the redundant features involved inside the data set, it is possible to reduce the cost of calculation of the hypertension prediction model.

In contrary to prediction-making based on ambiguous and indirect symptoms, the diagnosis of disease is more often being performed through computer-based probabilistic approaches. This study focuses on the diagnostic accuracy of different neural network models for diagnosing latent arterial hypertension (AH) with the use of results of daily blood pressure monitoring (DBPM). The study was conducted on 34 apparently healthy subjects (24 men and 10 women aged 17-64 years) and 72 AH patients (63 men and 9 women aged 15-62 years). The DBPM was performed for 24 h, every 7.5-15 and 15-30

minutes in the daytime and at night, respectively. The receiver operating characteristic (ROC) analysis was used to compare the accuracy of different diagnostic approaches.

The study shows that the neural network technologies facilitate to develop a model based on DBPM data paving a way for diagnosis of early AH stages in patients not exhibiting an obvious increase in BP. Secondly, through this model the percentage of predictions exceeds 80% when the model is tested using an independent sample. Thirdly, indices recorded during nighttime BP monitoring are very important for diagnosing latent AH. Fourth, indices characterizing the BP variation and the response of the patient's body to hospitalization are also of great importance. Fifth, dependences between the presence of latent AH and DBPM indices are complex and mostly non-linear.

Hypertension has considerable consequences and complications for the societies, as a main factor for coronary heart disease. There are main difficulties associated with hypertension as well as its diagnosis and therapy such as lack of clear clinical effects (pain) in early stages of the disease and treatment non-compliance. There must be effective screening tools to predict the risk of hypertension and to make the early-stage treatment possible. Internet seems to be an appropriate medium in having access to such a tool. This study tries to verify previously developed ANN-based and published using Internet and Java technology hypertension prediction tool, and to identify the system extension with use of additional available data, and test the possibility of use fuzzy nets as the engine of decision system in hypertension prediction area.

The study demonstrated that the ANNs could be successfully used in hypertension risk prediction and in the same time capable to be an engine for Internet-based screening tools. A combination of widespread access via Internet and predictive abilities of ANNs allows for a construction of powerful users.

## 2. MATERIALS & METHODS

In this study, hypertension patients' data is collected from hospitals in Turkey. After the collecting data, information extracting is realized by using data segmentation process. The collected data are arranged for use in data mining techniques.

## 2.1 DATA SET PREPARATION FOR HYPERTENSION

The data set that was used in the thesis consist of 11 variables of 203 subjects who were interviewed in a İstanbul Diabet Hospital and also 4 different treatment type. All of subjects are under hypertension treatment. In our work it's tried to find any relation between hypertension risk and BMI, blood pressure, gender, hba%, glucose, cholesterol, keton, protein, micros, micalb and genetic.

In this research, drug dosage planning made. These drugs are Coversyl, Monopril, Tenoretic, Atacand. ANFIS algorithm is help to planning to degree of opponent medicine for both types of hypertension patients. Some basis population statistics for the training and checking sets of used medicines are shown in Table 2.1.

**Table 2.1: Distribution of data use in the checking and training data sets**

| Drug | Checking Data (n) | Training Data (n) |
|---|---|---|
| Coversyl | 23 | 46 |
| Monopril | 19 | 36 |
| Tenoretic | 15 | 30 |
| Atacand | 11 | 23 |

The mean, ranges of the parameters and standard deviation of parameters are grouped with drugs basis.

As the data in the table Coversyl parameters, the maximum and minimum range and standard deviation and the average are given in Table 2.2.

**Table 2.2: Coversyl dosage planning parameters used for fuzzy modeling**

| Parameters | Mean | Range (minimum-maximum) | Standard Deviation |
|---|---|---|---|
| BMI | 30,81 | 21,39-47,94 | 5,14 |

| | | | |
|---|---|---|---|
| Hba (%) | 8,61 | 5,5-14,2 | 2,12 |
| Cholesterol (mg/dl) | 185,80 | 95-354 | 34,87 |
| Upper Blood Pressure (mmHg) | 145,44 | 100-220 | 20,57 |
| Lower Blood Pressure (mmHg) | 87,05 | 65-120 | 13,48 |
| Glucose (mg/dl) | 175,39 | 14-399 | 79,33 |
| Dosage (mg) | 7,5 | 2,5-10 | 3,26 |

As the data in the table Monopril parameters, the maximum and minimum range and standard deviation and the average are given in Table 2.3.

**Table 2.3: Monopril dosage planning parameters used for fuzzy modeling**

| Parameters | Mean | Range (minimum-maximum) | Standard Deviation |
|---|---|---|---|
| BMI | 28,95 | 13,86-40,23 | 5,29 |
| Hba (%) | 8,87 | 5,5-14 | 2,10 |
| Cholesterol (mg/dl) | 182,91 | 123-333 | 32,44 |
| Upper Blood Pressure (mmHg) | 143,64 | 105-180 | 20,45 |
| Lower Blood Pressure (mmHg) | 87,36 | 65-110 | 13,41 |
| Glucose (mg/dl) | 173,93 | 80-423 | 67,59 |
| Dosage (mg) | 27,82 | 10-40 | 12,88 |

As the data in the table Tenoretic parameters, the maximum and minimum range and standard deviation and the average are given in Table 2.4.

**Table 2.4: Tenoretic dosage planning parameters used for fuzzy modeling**

| Parameters | Mean | Range (minimum-maximum) | Standard Deviation |
|---|---|---|---|
| BMI | 30,56 | 16,17-44,47 | 5,83 |
| Hba (%) | 9,11 | 5,4-15 | 2,19 |
| Cholesterol (mg/dl) | 183,51 | 136-250 | 22,67 |
| Upper Blood | 150,78 | 130-180 | 11,64 |

| | | | |
|---|---|---|---|
| Pressure (mmHg) | | | |
| Lower Blood Pressure (mmHg) | 79,44 | 60-90 | 11,84 |
| Glucose (mg/dl) | 203,76 | 88-394 | 78,99 |
| Dosage (mg) | 36,67 | 25-100 | 15,45 |

As the data in the table Atacand parameters, the maximum and minimum range and standard deviation and the average are given in Table 2.5.

**Table 2.5: Atacand dosage planning parameters used for fuzzy modeling**

| Parameters | Mean | Range (minimum-maximum) | Standard Deviation |
|---|---|---|---|
| BMI | 30,60 | 19,34-42,05 | 6,06 |
| Hba (%) | 8,99 | 4,9-14,9 | 2,70 |
| Cholesterol (mg/dl) | 188,38 | 125-290 | 35,12 |
| Upper Blood Pressure (mmHg) | 143,09 | 115-150 | 8,04 |
| Lower Blood Pressure (mmHg) | 89,56 | 75-95 | 3,29 |
| Glucose (mg/dl) | 178,44 | 64-441 | 78,85 |
| Dosage (mg) | 7,18 | 4-16 | 4,09 |

To obtain the data, some algorithms can be used. ANFIS algorithm is one of them. To get the results Anfis model inputs must be entered. You can use the following table for generating ANFIS model is designed according to the data (Table 2.6). Blood Pressure divided in to five types, if LBP is between 60-79 and UBP is between 90-119 then Type 1 (Normal), if LBP is between 80-89 and UBP is between 120-139 then Type 2 (Prehypertension), if LBP is between 90-99 and UBP is between 140-159 then Type 3 (Stage 1), if LBP is greater than 100 and UBP is greater than 160 then Type 4 (Stage 2), if LBP is smaller than 90 and UBP is greater than 140 then Type 5 (Isolated Systolic Hypertension).Gender can take two values male or female. For male patients gender takes 0 value, for female patients it takes 1 value. Body mass index (BMI) is divided 5 parts, between 10-18.5 under weight, 18.51-25 normal, 25.01-30 over weight, 30.01-40 obese

and 40.01-70 over obese. All type of hypertension patients' body mass indexes are usually over weight and obese. In all types of hypertension approximately %33 of the patients affect from heredity. If the patient's parents has hypertension genetic is 1, otherwise genetic is 0. Hba(1c) measures average blood glucose over the past four to six weeks. Hba(%) is divided into 5 parts, between 0-4 then 1, 4.01-8 then 2, 8.01-12 then 3, 12.01-16 then 4 and 16.01-20 then 5. In all types of hypertension Hba is generally 2 and 3. Cholesterol and hypertension are two different things but stroke, hearth attack and artery diseases can be caused by high blood pressure resulting from high cholesterol levels. So it can be very dangerous when blood pressure and cholesterol levels are high together. Cholesterol divided into 5 parts, between 75-135 then 1, 135.01-195 then 2, 195.01-255 then 3, 255.01-315 then 4 and 315.01-375 then 5. In all types of hypertension %80 of the patients' cholesterol level is 2. Glucose is also divided into 5 parts, between 0-60 then 1, 60.01-100 then 2, 100.01-140 then 3, 140.01-200 then 4 and greater than 200 is 5. Glucose of level 1 is seen only on one patient, and in all types of hypertension level 2, level 3 and level 4 Glucose is seen at the same rates.

**Table 2.6: Parameter Values**

| Parameters | Values |
|---|---|
| Gender | Male=1, Female=0 |
| BMI | 10-18.5=1, 18.51-25=2, 25.01-30=3, 30.01-40=4, 40.01-70=5 |
| Genetic | Yes=1, No=0 |
| Blood Pressure (mmHg) | LBP(60-79) && UBP(90-119)=1, LBP(80-89) && UBP(120-139)=2, LBP(90-99) && UBP(140-159)=3, LBP>=100 && UBP>=160=4, LBP>90 && UBP>=140 = 5 |
| Hba 1c (%) | 0-4=1, 4.01-8=2, 8.01-12=3, 12.01-16=4, 16.01-20=5 |
| Cholesterol (mg/dl) | 75-135=1, 135.01-195=2, 195.01-255=3, 255.01-315=4, 315.01-375=5 |
| Glucose (mg/dl) | 0-60=1, 60.01-100=2, 100.01-140=3, 140.01-200=4, >200=5 |
| Ketone (in urine) (mg/dl) | Yes=1, No=0 |
| Protein (in urine) (mg/dl) | Yes=1, No=0 |
| Mikros (in urine) (lokosit) | Yes=1, No=0 |
| Mikalb (in urine) | Yes=1, No=0 |

| | |
|---|---|
| (mg/dl) | |

Attribute Ranking applied to all data sets at WEKA. Ranked values equal to zero, the columns have been eliminated. These eliminated data sets are used with ANFIS.

The ranking results for Coversyl data set is shown in Table 2.7. For these results Mikalb column is discarded.

**Table 2.7: Coversyl Data set's Ranked Attributes**

| Ranked Value | Parameter |
|---|---|
| 0.5464 | Blood Pressure |
| 0.2052 | Cholesterol |
| 0.1483 | Glucose |
| 0.148 | Genetic |
| 0.141 | BMI |
| 0.1227 | Hba |
| 0.0809 | Gender |
| 0.0533 | Protein |
| 0.044 | Ketone |
| 0.044 | Mikros |
| 0 | Mikalb |

The ranking results for Monopril data set is shown in Table 2.8. For these results Gender column is discarded.

**Table 2.8: Monopril Data set's Ranked Attributes**

| Ranked Value | Parameter |
|---|---|
| 0.5804 | Blood Pressure |
| 0.57997 | BMI |
| 0.20328 | Mikros |
| 0.19456 | Cholesterol |
| 0.14962 | Glucose |
| 0.12219 | Hba |
| 0.09935 | Ketone |
| 0.09306 | Mikalb |
| 0.04132 | Genetic |
| 0.01829 | Protein |

| | |
|---|---|
| 0 | Gender |

The ranking results for Tenoretic data set is shown in Table 2.9. For these results no column is discarded.

**Table 2.9: Tenoretic Data set's Ranked Attributes**

| Ranked Value | Parameter |
|---|---|
| 0.2467 | BMI |
| 0.1372 | Glucose |
| 0.0834 | Cholesterol |
| 0.0821 | Protein |
| 0.0607 | Hba |
| 0.0504 | Genetic |
| 0.0497 | Mikalb |
| 0.0381 | Mikros |
| 0.0325 | Ketone |
| 0.0299 | Gender |
| 0.0194 | Blood Pressure |

The ranking results for Atacand data set is shown in Table 2.10. For these results no column is discarded.

**Table 2.10: Atacand Data set's Ranked Attributes**

| Ranked Value | Parameter |
|---|---|
| 0.4731 | Glucose |
| 0.2992 | BMI |
| 0.2516 | Blood Pressure |
| 0.1548 | Mikros |
| 0.1535 | Protein |
| 0.1022 | Cholestorol |
| 0.0892 | Gender |
| 0.0779 | Mikalb |
| 0.0492 | HBA |
| 0.0216 | Ketone |
| 0.0128 | Genetic |

## 2.2 ADAPTIVE NEURO FUZZY INFERENCE SYSTEM (ANFIS)

In this study, Fuzzy Inference System is used to evaluate the hypertension data set for planning the dosages of the medicines.

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. The process of fuzzy inference involves all of the pieces that are described in the previous sections: Membership Functions, Logical Operations, and If-Then Rules.

Fuzzy inference systems are used in fields such as automatic control, data classification, decision analysis, expert systems, and computer vision (Guopeng & Levin 2006). Because of its multidisciplinary nature, fuzzy inference systems are associated with a number of names, such as fuzzy-rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and simply (and ambiguously) fuzzy systems.

You can implement two types of fuzzy inference systems in the toolbox: Mamdani-type and Sugeno-type. These two types of inference systems vary somewhat in the way outputs are determined.

In this research, Takagi and Sugeno type fuzzy if-then rules are used such that the output of each rule is a linear combination of input variables plus a constant term. The final output is the weighted average of each rule's output. ANFIS is a fuzzy rule based classifier in which the rules are learnt from examples that use a standard back propagation algorithm. Anfis uses Sugeno type fuzzy system which is a linear equation (first order Sugeno inference system) or constant coefficients (zero-order Sugeno inference system) (Shafiq, Farooq & Khayam, 2008). The first order Sugeno inference system has two rules expressed as below;

*Rule1: IF x is $A_1$ and y is $B_1$ THEN $f_1 = p_1x + q_1y + r_1$*

*Rule2: IF x is $A_2$ and y is $B_2$ THEN $f_2 = p_2x + q_2y + r_2$*

The inputs are x and y to the node i, $A_i$ and $B_i$ are characterized by convenient membership functions and $p_i$, $q_i$ and $r_i$ are the consequence parameters (i = 1,2,...). The architecture of ANFIS is shown in Figure 2.1. The nodes of the same layer have the same function.

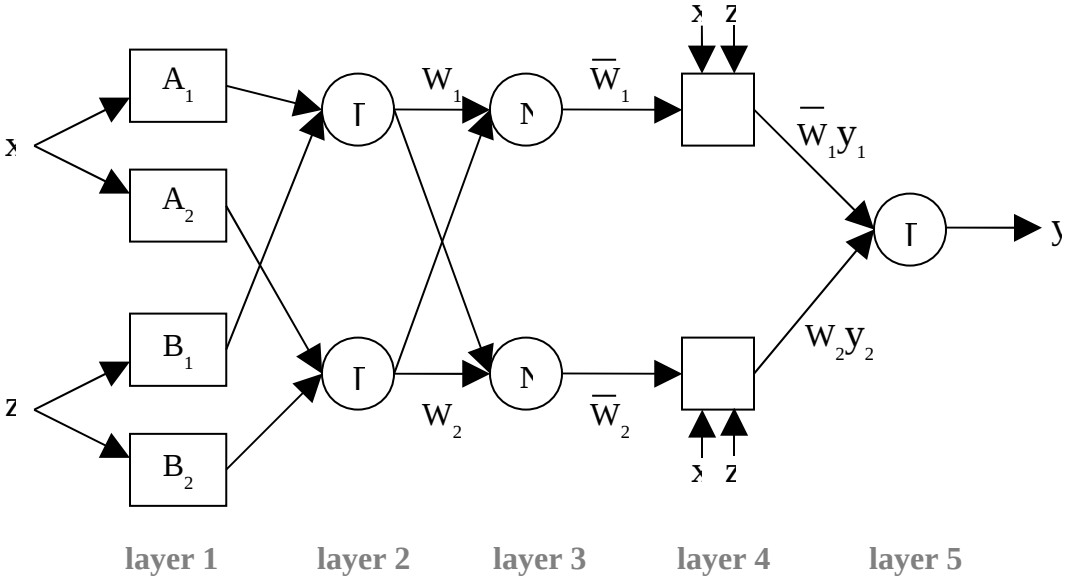


**Figure 2.1 ANFIS architecture**

As you can see in Figure 2.1, the ANFIS consists of five layers. The explanations of these layers are below:

**Layer 0:** It consists of plain input variable set.

**Layer 1:** Every node in this layer is a square node with a node function as given in Formula 2.1;

$$\mu_{A1}(x) = \frac{1}{1+\left[\left(\frac{x-c_1}{a_1}\right)^2\right]^{b_1}}$$

**Formula 2.1**

where A is a generalized bell fuzzy set defined by the parameters {a,b,c} , where c is the middle point, b is the slope and a is the deviation.

**Layer 2:** The function is a T-norm operator that performs the firing strength of the rule, e.g., fuzzy AND and OR. The simplest implementation just calculates the product of all incoming signals.

**Layer 3:** Every node in this layer is fixed and determines a normalized firing strength. It calculates the ratio of the $j^{th}$ rule's firing strength to the sum of all rules firing strength.

**Layer 4:** The nodes in this layer are adaptive and are connected with the input nodes (of layer 0) and the preceding node of layer 3. The result is the weighted output of the rule j.

**Layer 5:** This layer consists of one single node which computes the overall output as the summation of all incoming signals.

## 2.3 ROUGH SET THEORY

The rough set theory is developed by Pawlak. It is interested in classificatory analysis of data sets. Rough Set analysis aim is to synthesize approach of concepts from the acquired data.

In rough set theory, knowledge is interpreted as an ability of classify some objects. These objects form a set called often a *universe of discourse* and their nature may vary case to case: they may be e.g. medical patients, processes participants in a conflict, etc.

Classification of objects consists in this theory in finding a set of subsets of the universe of objects such that each object answers to at least one subset in the sense that it is an element of this subset. Subsets of the universe of objects are called often categories and we may say that classification of objects consists in finding a covering of the universe of objects by a set of categories.

The most clear, unambiguous, classification case happens when the covering of the universe of objects consists of pair-wise disjoint categories as in this case each object answers to a unique category thus classification acquires a functional character. We say that in this case categories included in classification from partition of the universe of objects.

# 3. FINDINGS

The results of this study are mentioned in this part. I applied ANFIS and Rough Set methods to the data sets. The results of ANFIS and Rough set methods are compared. There are 4 data sets which are Coversyl, Monopril, Tenoretic and Atacand data sets. Matlab 7.5.0 Fuzzy Toolbox is used for ANFIS method. ROSETTA software is used for Rough set algorithm. The sensitivity results of ANFIS and RSES are shown in table 3.1.

**Table 3.1: The sensivity result of ANFIS and RSES.**

| Data sets | ANFIS | RSES |
|---|---|---|
| Coversyl | 84% | 82% |
| Monopril | 79% | 74% |
| Tenoretic | 75% | 70% |
| Atacand | 75% | 68% |

The comparison of ANFIS and RSES RMSE rates are shown in table 3.2.

**Table 3.2: The comparision of ANFIS and RSES RMSE rates.**

| Data sets | ANFIS | RSES |
|---|---|---|
| Coversyl | 20% | 47% |
| Monopril | 17% | 24% |

| | | |
|---|---|---|
| Tenoretic | 12% | 36% |
| Atacand | 17% | 57% |

### 3.1 ANFIS RESULT

First I will start to introduce with Coversyl data set's ANFIS results. 46 data used for training and 23 data used for checking from Coversyl data set. Training and checking data distribution are shown in figure 3.1 and 3.2. The performance of ANFIS for Coversyl data set has seen in figure 3.2. Blue points are actual outputs and red points are predicted outputs.

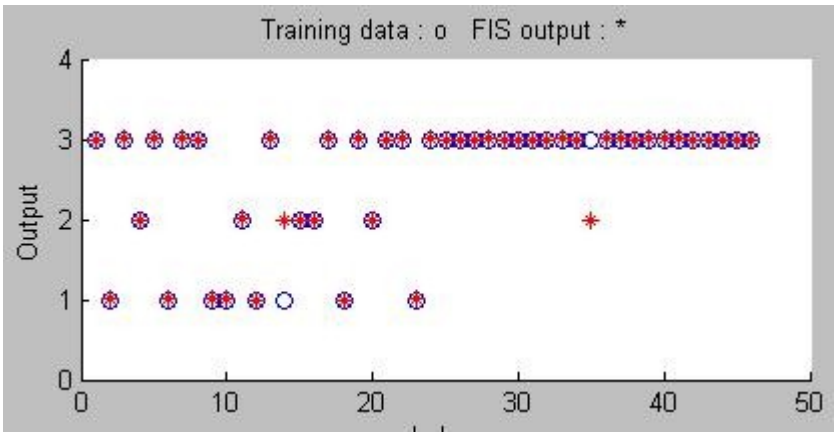


**Figure 3.1: ANFIS Training data plot**

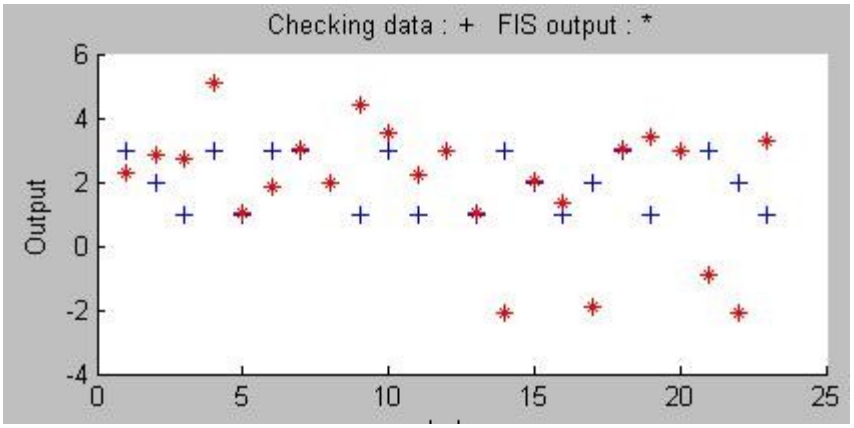


**Figure 3.2: ANFIS Checking data plot**

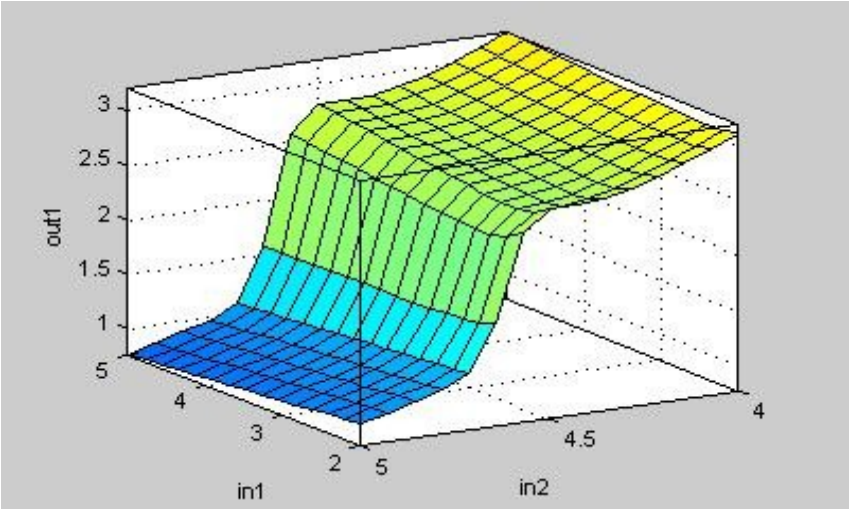


**Figure 3.3: Surface plot BMI and Blood Pressure versus Output**

For Coversyl data set ANFIS generated 7 rules. The checking data error of Coversyl is 0.2083 and the RMSE of Coversyl data set is 20 percent. The sensitivity rate of Coversyl is 84 percent. The correctness of these rules is 78 percent. These rules are expressed below:

Rule 1: [4 5 0 4 3 2 1 1 1 1 1] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 4 and Glucose = 3 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 2: [4 4 0 2 1 2 1 1 1 1 0] [3]

If BMI = 4 and Blood Pressure = 4 and Gender = 0 and Hba = 2 and Glucose = 1 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

Rule 3: [4 4 1 2 1 2 1 1 1 1 1] [3]

If BMI = 4 and Blood Pressure = 4 and Gender = 1 and Hba = 2 and Glucose = 1 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 1.

Rule 4: [4 5 0 2 1 2 0 0 0 0 1] [1]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 2 and Glucose = 1 and Cholesterol = 2 and Ketone = 0 and Protein = 0 and Mikros = 0 and Mikalb = 0 and Genetic = 1 then Output = 3.

Rule 5: [4 5 0 3 1 2 0 0 1 1 0] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 4 and Glucose = 1 and Cholesterol = 2 and Ketone = 0 and Protein = 0 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

Rule 6: [4 5 0 3 2 2 1 0 1 1 0] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 3 and Glucose = 2 and Cholesterol = 2 and Ketone = 1 and Protein = 0 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

Rule 7: [5 5 1 3 3 3 1 1 1 1 0] [2]

If BMI = 5 and Blood Pressure = 5 and Gender = 1 and Hba = 3 and Glucose = 3 and Cholesterol = 3 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

The above rules are in different situations for trained ANFIS model. Each rule is represented by a vector. It consists of input values for the system. Descriptions of the rules which are generated by ANFIS for Coversyl data set are below.

In rule 1, the patient is obese. Her genetic is positive and she has ketone, protein, mikalb and micros in urine and these values affect the output directly. Hba(%) level is high. Cholesterol level is between 135 and 195 mg/dl. Glucose level is between 100 and 140 mg/dl. She is type 5 hypertension patient that means LBP<90 and UPB>=140. The output is 2 that mean Coversyl dosage is 5mg.

In rule 2, the patients body mass index, gender and cholesterol remain in same level with

in rule 1. Her blood pressure and glucose level is higher and Hba level is between 4 and 8 mg/dl. Her genetic is negative. So blood pressure, glucose and Hba is direct proportion to output. According to these parameters system generates output 3 that means Coversyl dosage is 10mg.

Rule 3 is similar to rule 2. Difference from second rule, this patient is male and his parents don't have hypertension. The other parameters are in the same class with rule 2. This variation is not change the output. Output is 3 so Coversyl dosage is 10mg.

In rule 4, the patient is obese again. She is type 5 hypertension patient. Hba, glucose and cholesterol levels are the same with rule 3, but she doesn't have ketone, protein, mikalb and micros in urine also genetic is negative, which means when these parameters' values are lower the dosage is lower, because with this parameters system generate output 1 differently with rule 3. Coversyl dosage is 2.5mg.

In rule 5, the patient is obese female and blood pressure level is 5 same with in rule 4. Also glucose and cholesterol levels are same. Difference is her hba is between 12-16 mg/ dl and she has micros and mikalb in her urine, with these differences output raised to 2. So we can say Hba mikalb and micros is affect on output. That patient starts to use 5mg Coversyl.

Rule 6 is similar to rule 5. Difference from fifth rule, cholesterol level is between 135-195 mg/dl and no ketone find in urine. In spite of these changes output remains same, that means cholesterol and ketone in urine are not directly affect to output. Output is 2 again, so Coversyl dosage is 5mg.

In rule 7, patient is over obese male. Hba level is same with in rule 6 like ketone, micros, mikalb and genetic. Glucose level is between 100-140 mg/dl. He has cholesterol between 195-255 mg/dl, cholesterol level change but output same so it shows cholesterol is not directly affect to output one more time. Also glucose level is not very important for output. Output is 2, dosage is 5 mg again.

Roc curve of Coversyl data set is shown in figure 3.4:

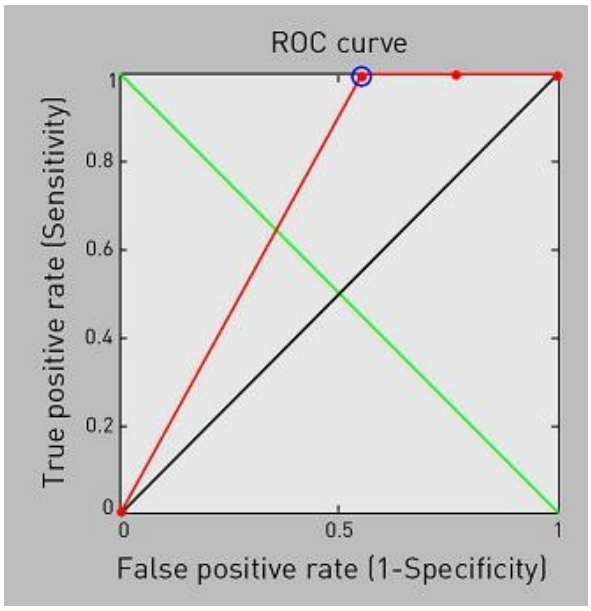


**Figure 3.4: ROC curve of Coversyl data set**

ANFIS model structure of Coversyl data set is shown in figure 3.5. There are 11 inputs and one output.

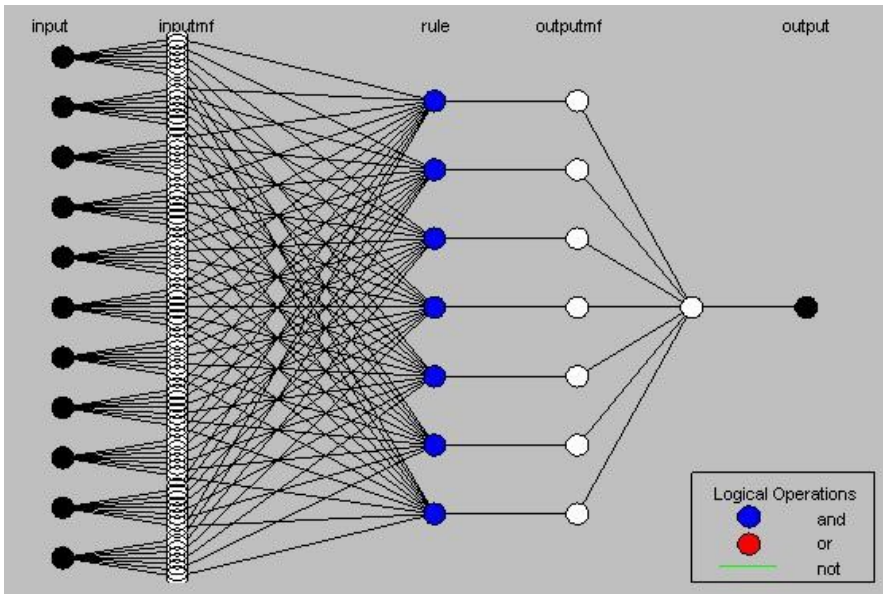


**Figure 3.5: ANFIS model structure of Coversyl data set**

The other data set is Monopril which results I will introduce. 36 data used for training and 19 data used for checking from Monopril data set. Training and checking data distribution are shown in figure 3.6 and 3.7. The performance of ANFIS for Monopril data set has seen in figure 3.7. Blue points are actual outputs and red points are predicted outputs.

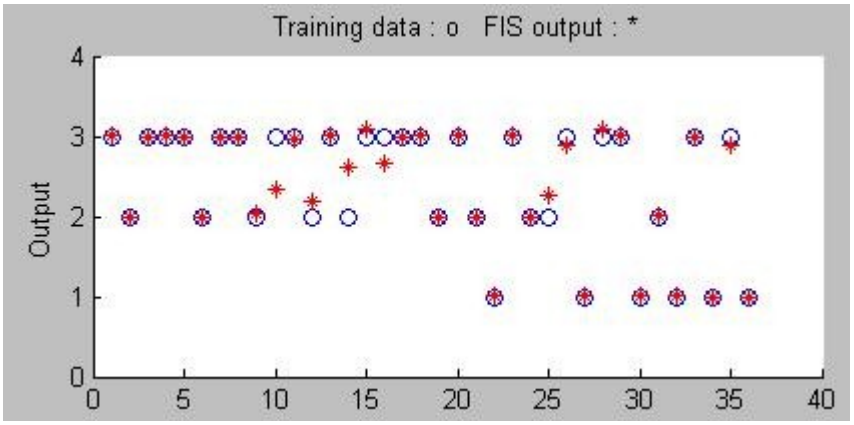


**Figure 3.6: ANFIS Training data plot**

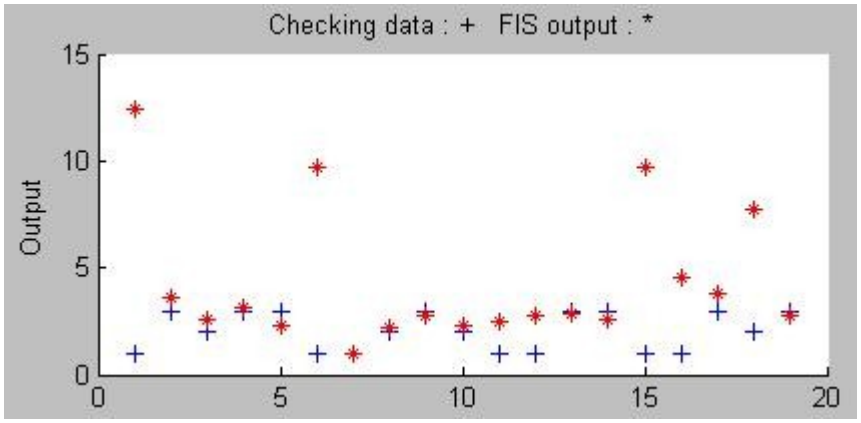


**Figure 3.7: ANFIS Cheching data plot**

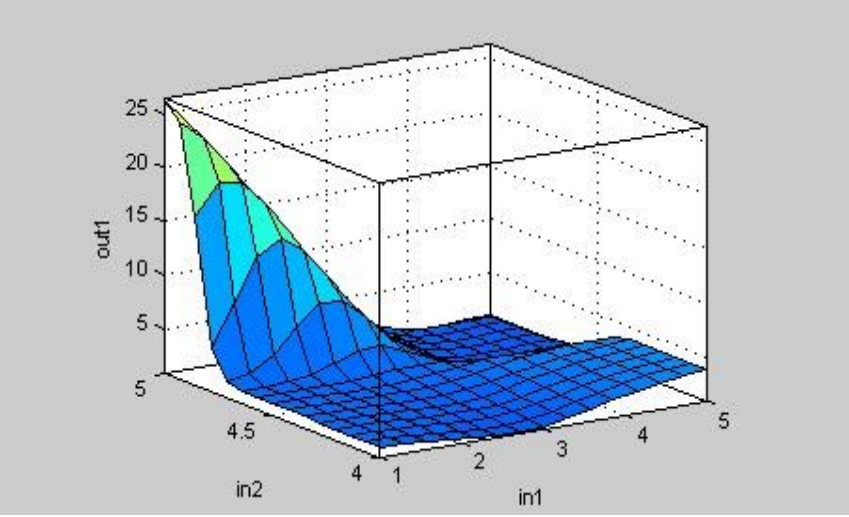


**Figure 3.8: Surface plot BMI and Blood Pressure versus Output**

For Monopril data set ANFIS generated 7 rules. The checking data error of Monopril is 0.17246 and the RMSE of Monopril data set is 17 percent. The sensitivity rate of Monopril is 78 percent. The correctness of these rules is 70 percent. These rules are expressed below:

Rule 1: [2 5 1 3 3 4 1 1 1 1 0] [1]

If BMI = 2 and Blood Pressure = 5 and Gender = 1 and Hba = 3 and Glucose = 3 and Cholesterol = 4 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 1.

Rule 2: [3 5 0 2 4 4 1 1 1 1 1] [2]

If BMI = 3 and Blood Pressure = 5 and Gender = 0 and Hba = 2 and Glucose = 4 and Cholesterol = 4 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 3: [4 5 0 2 3 4 0 0 0 0 1] [3]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 2 and Glucose = 3 and Cholesterol = 4 and Ketone = 0 and Protein = 0 and Mikros = 0 and Mikalb = 0 and Genetic = 1 then Output = 3.

Rule 4: [2 5 1 3 3 3 1 0 0 0 01] [1]

If BMI = 2 and Blood Pressure = 5 and Gender = 1 and Hba = 3 and Glucose = 3 and Cholesterol = 3 and Ketone = 1 and Protein = 0 and Mikros = 0 and Mikalb = 0 and Genetic = 0 then Output = 1.

Rule 5: [3 4 1 4 4 4 1 1 1 1 1] [2]

If BMI = 2 and Blood Pressure = 4 and Gender = 1 and Hba = 4 and Glucose = 4 and Cholesterol = 4 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 6: [4 5 0 1 2 4 0 1 0 0 1] [3]

If BMI = 5 and Blood Pressure = 5 and Gender = 0 and Hba = 1 and Glucose = 2 and Cholesterol = 4 and Ketone = 0 and Protein = 1 and Mikros = 0 and Mikalb = 0 and Genetic = 1 then Output = 3.

Rule 7: [2 5 0 1 2 1 1 0 1 1 0] [2]

If BMI = 2 and Blood Pressure = 5 and Gender = 0 and Hba = 1 and Glucose = 2 and Cholesterol = 1 and Ketone = 1 and Protein = 0, and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

The above rules are in different situations for trained ANFIS model. Each rule is represented by a vector. It consists of input values for the system. Descriptions of the rules which are generated by ANFIS for Monopril data set are below.

In rule 1, the patient is normal weight. Hba(%) level is between 8-12mg/dl. Genetic is negative and she has ketone, protein, mikalb and micros in urine. Cholesterol level is high. Glucose level is between 100 and 140 mg/dl. She is type 5 hypertension patient that means LBP<90 and UPB>=140. With these values output is 1 that means Monopril dosage is 5mg.

In rule 2, the patient’s blood pressure, cholesterol remain in same level with in rule 1.

Also ketone, protein, mikalb and micros remain same too. But body mass index, glucose and genetic values are different. Body mass index increase from level 2 to level 3, glucose level increase from level 3 to level 4 that means 140-200mg/dl, hba decreases from 8-12mg/dl to 4-8mg/dl and genetic is positive. So these parameters are strongly affect output. According to these parameters system generates output 2 that means Monopril dosage is 10mg.

In rule 3 none of ketone, protein, mikalb and micros found on patient's urine. Gender, cholesterol, hba and genetic values are same with rule 2. This patient is obese, as we can see in rule 2 and rule 3 body mass indexes are increased, at the same time the dosage level increased too, then we can say body mass index is an important parameter on output. With this variation output increased to 3 so Monopril dosage is 20mg.

Rule 4 is similar to rule 1. Difference is cholesterol decreases 1 level, and ketone, protein, micros and mikalb are negative. These changes doesn't change the output, that means these parameters are not affect the output seriously, so the output is 1 like in rule 1. Monopril dosage is 5 mg.

In rule 5, the patient is male. Ketone and genetic values are same with rule 4. Hba, body mass index, glucose and cholesterol values are all increased from level 3 to level 4 when compared to rule 4. His blood pressure level is 4. By these parameters we can see body mass indexes affect on output it increases again and output increase too. At the same time blood pressure affects the output like body mass index.. With these affects output raised to 2 when compared to rule 4. That patient starts to use 10mg Monopril.

Rule 6 is similar to rule 3. Difference from third rule, glucose level is between 60-100 mg/dl, hba level is between 0-4 mg/dl and no protein is positive. In spite of these changes output remains same, that means glucose and protein in urine are not directly affect to output. Output is 3 again, so Monopril dosage is 10mg.

In rule 7, patient is normal weight male. Hba level is between 0-4 mg/dl. He is type 5

hypertension patient. Glucose level is between 60-100 mg/dl. He has cholesterol between 75-135 mg/dl. Ketone, micros and mikalb parameters are positive. With these variations the output is 2, so Monopril dosage is 10mg.

Roc curve of Monopril data set is shown in figure 3.9:

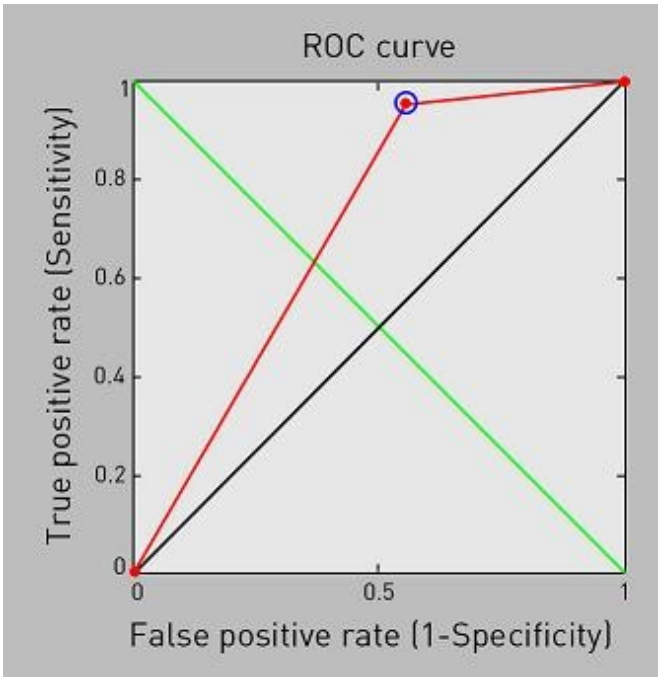


**Figure 3.9: ROC curve of Monopril data set**

ANFIS model structure of Monopril data set is shown in figure 3.10. There are 11 inputs and one output.

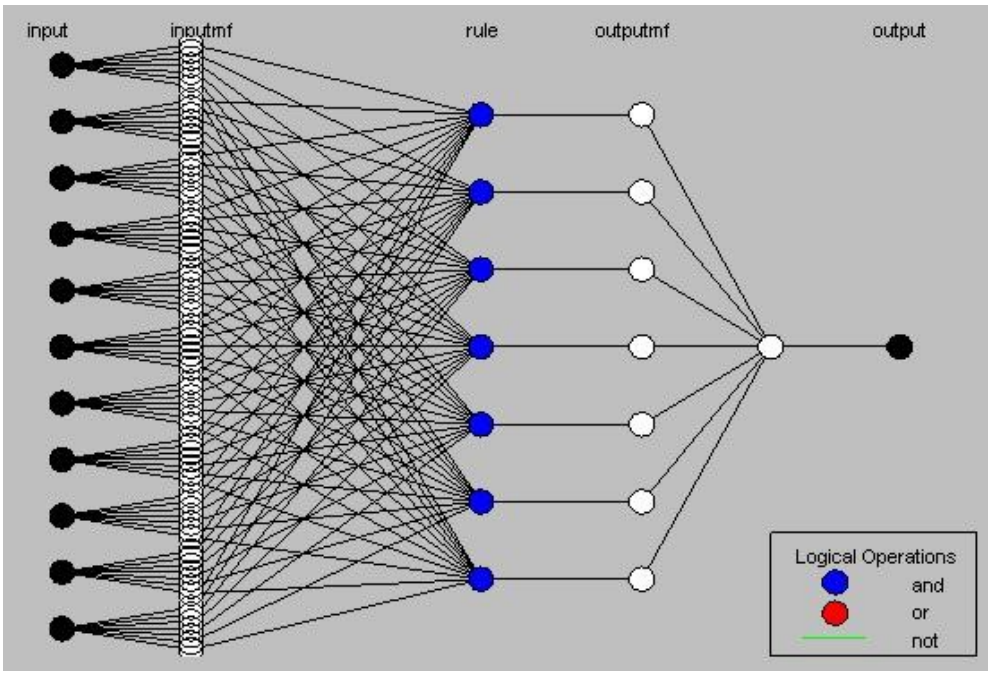


**Figure 3.10: ANFIS model structure of Monopril data set**

The other data set is Tenoretic which results I will introduce. 30 data used for training and 15 data used for checking from Tenoretic data set. Training and checking data distribution are shown in figure 3.11 and 3.12. The performance of ANFIS for Tenoretic data set has seen in figure 3.12. Blue points are actual outputs and red points are predicted outputs.


**Figure 3.11: ANFIS Training data plot**

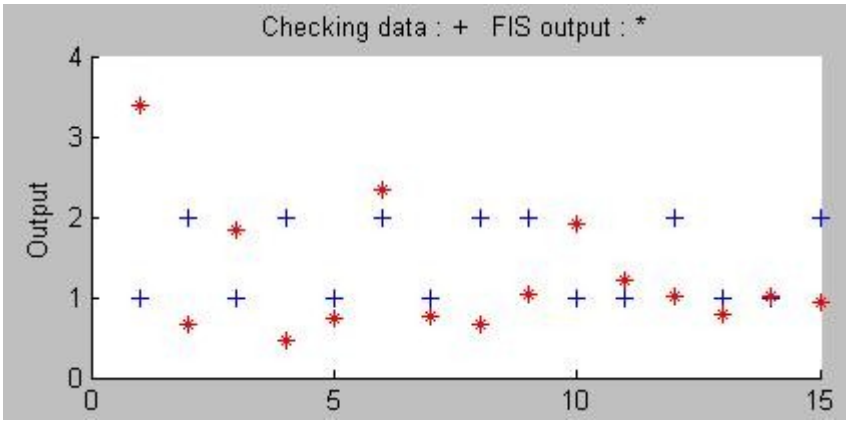


**Figure 3.12: ANFIS Cheching data plot**

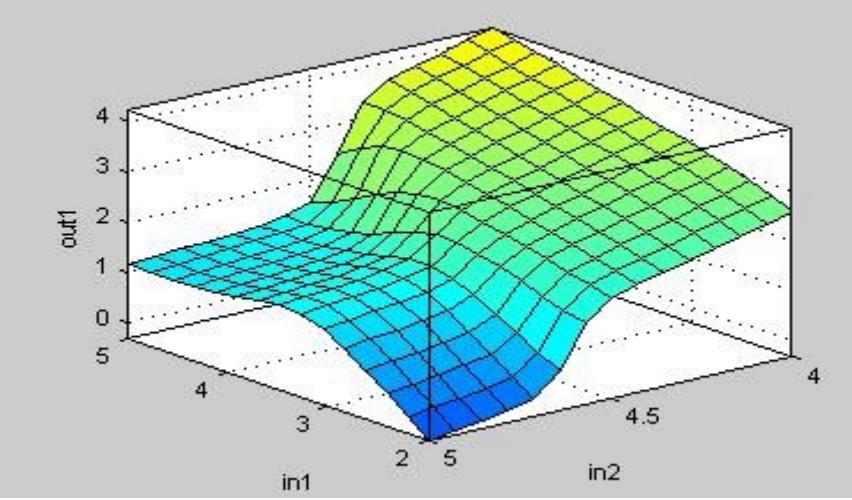


**Figure 3.13: Surface plot BMI and Blood Pressure versus Output**

For Tenoretic data set ANFIS generated 7 rules. The checking data error of Tenoretic is 0.074536 and the RMSE of Tenoretic data set is 7 percent. The sensitivity rate of Tenoretic is 81 percent. The correctness of these rules is 74 percent. These rules are expressed below:

Rule 1: [4 5 0 3 4 2 1 1 1 1 1] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 3 and Glucose = 4 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 2: [4 4 0 2 2 2 1 1 1 1 0] [3]

If BMI = 4 and Blood Pressure = 4 and Gender = 0 and Hba = 2 and Glucose = 2 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 3.

Rule 3: [3 4 1 4 4 2 1 1 1 1 1] [3]

If BMI = 3 and Blood Pressure = 4 and Gender = 1 and Hba = 4 and Glucose = 4 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 3.

Rule 4: [5 5 0 2 2 2 0 1 0 0 0] [1]

If BMI = 5 and Blood Pressure = 5 and Gender = 0 and Hba = 2 and Glucose = 2 and Cholesterol = 2 and Ketone = 0 and Protein = 1 and Mikros = 0 and Mikalb = 0 and Genetic = 0 then Output = 1.

Rule 5: [3 4 1 1 2 2 1 0 1 0 0] [3]

If BMI = 3 and Blood Pressure = 4 and Gender = 1 and Hba = 1 and Glucose = 2 and Cholesterol = 2 and Ketone = 1 and Protein = 0 and Mikros = 1 and Mikalb = 0 and Genetic = 0 then Output = 3.

Rule 6: [4 4 1 1 2 2 1 1 1 1 1] [3]

If BMI = 4 and Blood Pressure = 4 and Gender = 1 and Hba = 1 and Glucose = 2 and Cholesterol = 2 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 3.

Rule 7: [3 5 0 3 4 1 1 0 1 1 0] [3]

If BMI = 3 and Blood Pressure = 4 and Gender = 0 and Hba = 3 and Glucose = 4 and Cholesterol = 1 and Ketone = 1 and Protein = 0, and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 3.

The above rules are in different situations for trained ANFIS model. Each rule is represented by a vector. It consists of input values for the system. Descriptions of the

rules which are generated by ANFIS for Tenoretic data set are below.

In rule 1, the patient is obese. Her genetic is positive and she has ketone, protein, mikalb and micros in urine. Hba(%) level is normal. Cholesterol level is between 135 and 195 mg/dl. Glucose level is between 140 and 200 mg/dl. She is type 5 hypertension patient that means LBP<90 and UPB>=140. The output is 2 that mean Tenoretic dosage is 50mg.

In rule 2, the patients body mass index, gender and cholesterol remain in same level with in rule 1. Her blood pressure is higher. Hba level is between 4 and 8 mg/dl. Glucose level is between 60-100 mg/dl. Her genetic is negative. According to these parameters system generates output 3 that means Tenoretic dosage is 100mg.

In rule 3, this patient is normal weight female and her parents have hypertension too. Her blood pressure level is 4 which is very high. Hba and glucose levels are same, level 4. Cholesterol level is between 135 and 195 mg/dl. Ketone, protein, micros and mikalb parameters are same with rule 2. This variation is not change the output. Output is 3 so Tenoretic dosage is 100mg.

In rule 4, the patient is over obese and she is a type 5 hypertension patient. Cholesterol and protein levels are the same with rule 3, but she doesn't have ketone, mikalb and micros in urine also hba and glucose values are level 2 and genetic is negative, which means when these parameters' values are lower the dosage is lower, because with this parameters system generate output 1 different form rule 3. Tenoretic dosage is 25mg.

In rule 5, the patient is normal weight female and blood pressure level is 4. Glucose and cholesterol levels are same with rule 4. Difference is her hba is between 0-4 mg/dl and she has ketone and mikros in her urine, with these differences output raised to 3. That patient starts to use 100mg Tenoretic.

Rule 6 is similar to rule 5. Difference from fifth rule, body mass index level is between 30-40, protein and mikalb found in urine and genetic is positive. In spite of these changes

output remains same, that means cholesterol, genetic, protein and mikalb in urine are not directly affecting output. Output is 3 again, so Tenoretic dosage is 100mg.

In rule 7, patient is over weight male. Hba level is between 8-12 mg/dl. Blood pressure, ketone, micros and mikalb remains same level with rule 6. Genetic is positive. Glucose level is between 140-200 mg/dl. He has cholesterol between 75-135 mg/dl, cholesterol and genetic levels are change when compared to rule 6 but output same so it shows these are not directly affect to output. Output is 3, dosage is 100mg again.

Roc curve of Tenoretic data set is shown in figure 3.14:

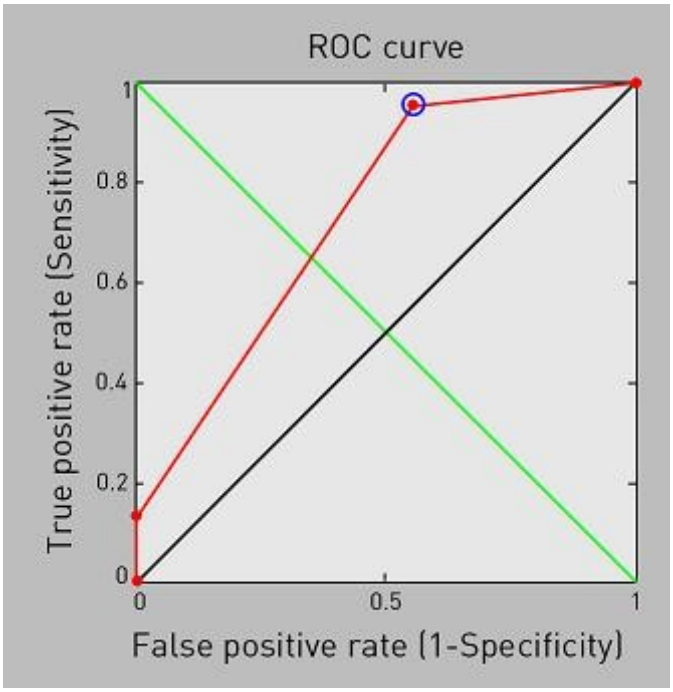


**Figure 3.14: ROC curve of Tenoretic data set**

ANFIS model structure of Tenoretic data set is shown in figure 3.15. There are 11 inputs and one output.

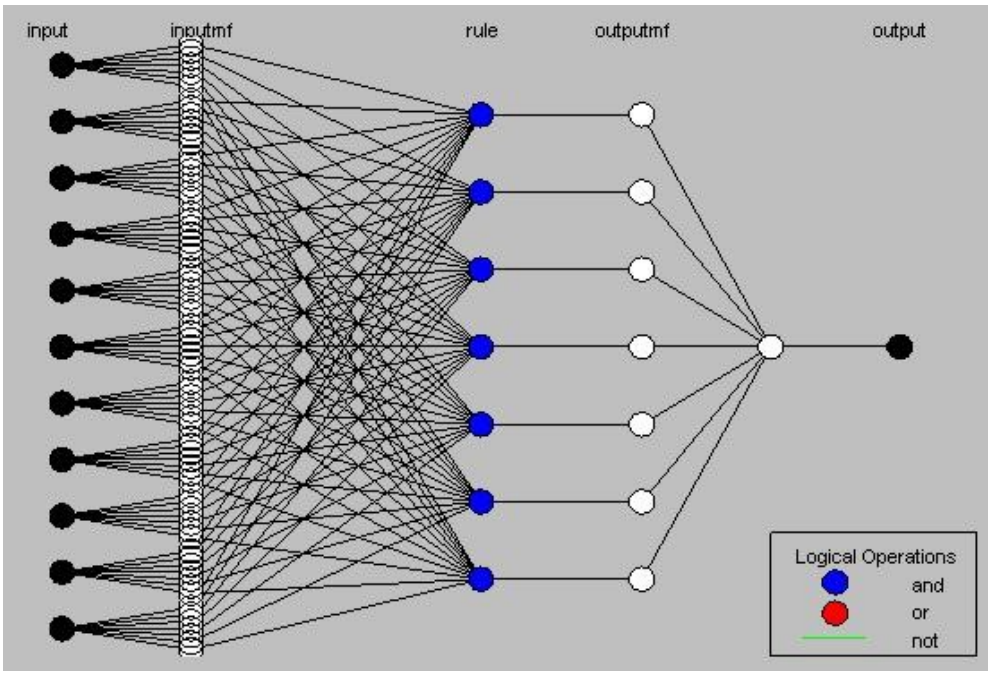


**Figure 3.15: ANFIS model structure of Tenoretic data set**

The other data set is Atacand which results I will introduce. 23 data used for training and 11 data used for checking from Atacand data set. Training and checking data distribution are shown in figure 3.16 and 3.17. The performance of ANFIS for Atacand data set has seen in figure 3.17. Blue points are actual outputs and red points are predicted outputs.


**Figure 3.16: ANFIS Training data plot**

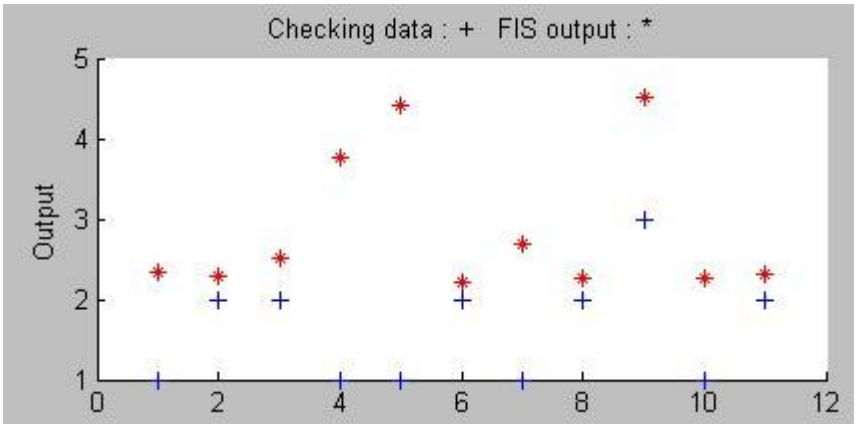


**Figure 3.17: ANFIS Cheching data plot**

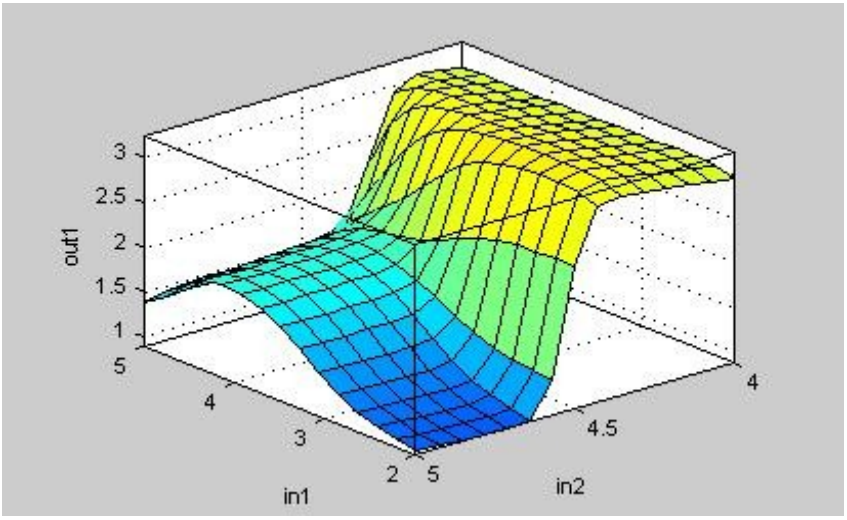


**Figure 3.18: Surface plot BMI and Blood Pressure versus Output**

For Atacand data set ANFIS generated 7 rules. The checking data error of Atacand is 0.068317 and the RMSE of Atacand data set is 6 percent. The sensitivity rate of Atacand is 78 percent. The correctness of these rules is 71 percent. These rules are expressed below:

Rule 1: [4 5 0 3 3 3 1 1 1 1 1] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 3 and Glucose = 3 and Cholesterol = 3 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 2: [2 4 1 2 2 3 1 0 0 0 0] [3]

If BMI = 2 and Blood Pressure = 4 and Gender = 1 and Hba = 2 and Glucose = 2 and Cholesterol = 3 and Ketone = 1 and Protein = 0 and Mikros = 0 and Mikalb = 0 and Genetic = 0 then Output = 3.

Rule 3: [4 4 0 3 4 3 1 1 1 1 1] [3]

If BMI = 4 and Blood Pressure = 4 and Gender = 0 and Hba = 3 and Glucose = 4 and Cholesterol = 3 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 3.

Rule 4: [2 5 1 3 3 4 1 1 1 1 0] [2]

If BMI = 2 and Blood Pressure = 5 and Gender = 1 and Hba = 3 and Glucose = 3 and Cholesterol = 4 and Ketone = 1 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 2.

Rule 5: [3 4 0 3 3 2 1 0 1 1 0] [3]

If BMI = 3 and Blood Pressure = 4 and Gender = 0 and Hba = 3 and Glucose = 3 and Cholesterol = 2 and Ketone = 1 and Protein = 0 and Mikros = 1 and Mikalb = 1 and Genetic = 0 then Output = 3.

Rule 6: [4 5 0 1 3 3 0 1 1 1 1] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 0 and Hba = 1 and Glucose = 3 and Cholesterol = 3 and Ketone = 0 and Protein = 1 and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

Rule 7: [4 5 1 1 2 3 1 1 1 1 1] [2]

If BMI = 4 and Blood Pressure = 5 and Gender = 1 and Hba = 1 and Glucose = 2 and Cholesterol = 3 and Ketone = 1 and Protein = 1, and Mikros = 1 and Mikalb = 1 and Genetic = 1 then Output = 2.

The above rules are in different situations for trained ANFIS model. Each rule is

represented by a vector. It consists of input values for the system. Descriptions of the rules which are generated by ANFIS for Atacand data set are below.

In rule 1, the patient is obese. Cholesterol level is between 195 and 255 mg/dl. Hba(%) level is between 8-12 mg/dl. Her genetic is positive and she has ketone, protein, mikalb and micros in urine. Glucose level is between 140 and 200 mg/dl. She is type 5 hypertension patient. The output is 2 that mean Atacand dosage is 16mg.

In rule 2, the patient's cholesterol and ketone remain in same level with in rule 1, but all other parameters are different. Her blood pressure is level 4. Hba level is between 4 and 8 mg/dl. Glucose level is between 60-100 mg/dl. Her genetic is negative. By the affect of these changes, system generates output 3 that means Atacand dosage is 32mg.

In rule 3, this patient is obese male and his parents are hypertension patient too. He is type 4 hypertension patient which is very risky. Hba and cholesterol levels are same, level 4. Glucose level is between 140 and 200 mg/dl. Ketone, protein, micros and mikalb parameters are positive differently from rule 2. This variation is not change the output. Output is 3 so Atacand dosage is 32mg.

In rule 4, the patient has a hypertension of type 5. She is normal weight. Hba, ketone, protein, micros and mikalb levels are the same with rule 3, but her genetic is negative also body mass index and glucose values are level 2 that is lower than the values in rule 3, which means when these parameters' values are lower the dosage is lower, because with this parameters system generate output 2 different form rule 3. Atacand dosage is 16mg.

In rule 5, the patient is normal weight male and blood pressure level is 4. Glucose and hba levels are same with rule 4. Difference is his cholesterol is between 135-195 mg/dl, genetic is negative and he doesn't has protein in his urine, with these differences output raised to 3. That patient starts to use 32mg Atacand.

Rule 6 is similar to rule 1. Difference from first rule, hba level is between 0-4 mg/dl and ketone in urine is negative. In spite of these changes output remains same, that means hba and ketone in urine are not directly affect to output. Output is 2 again, so Atacand dosage is 16mg.

In rule 7, patient is obese female. Hba level is between 0-4 mg/dl. Body mass index, blood pressure, cholesterol, protein, micros, mikalb and genetic remains same level with rule 6. Glucose level is between 60-100 mg/dl. Only gender, glucose and ketone levels are change when compared to rule 6 but output is same so it shows these parameters are not directly affect to output. Output is 2, dosage is 16mg again.

Roc curve of Atacand data set is shown in figure 3.19:

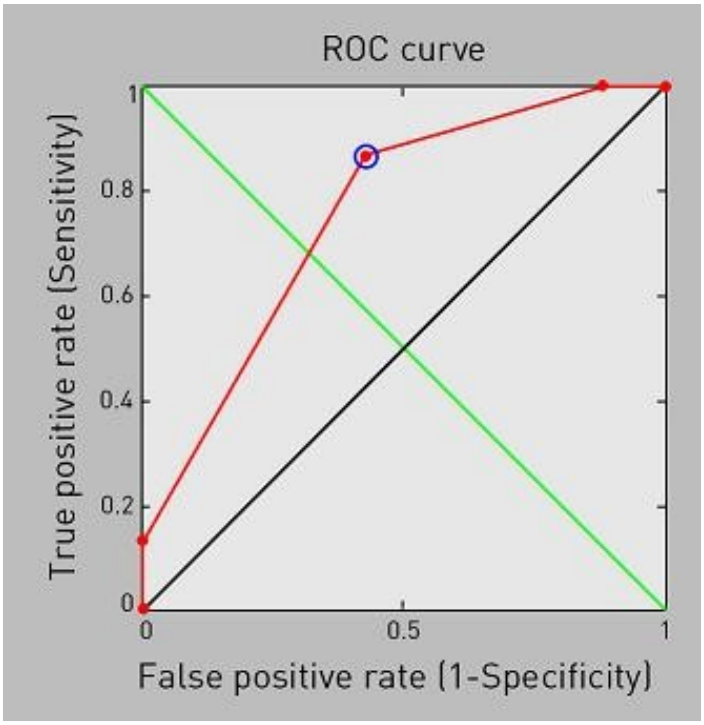


**Figure 3.19: ROC curve of Atacand data set**

ANFIS model structure of Atacand data set is shown in figure 3.20. There are 11 inputs and one output.

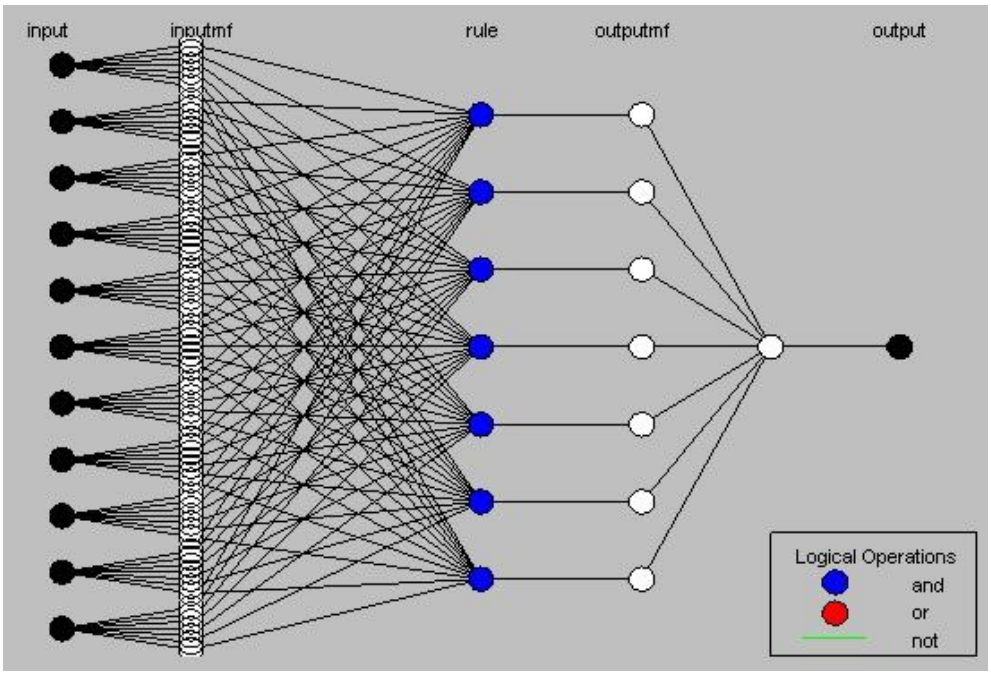


**Figure 3.20: ANFIS model structure of Atacand data set**

## 3.2 ROUGH SET RESULTS

Rosetta software is used for RSES (Rough Set algorithm implementation) algorithm.

I will start to introduce with Coversyl data set's RSES results. RSES algorithm generates 23 rules. Among those, I choose the specific ones and examine them. The selected rules are below:

Rule 1: Blood Pressure(4) AND Gender(1) AND Cholesterol(2) => Coversyl(3)

Rule2: Hba(4) => Coversyl(3) OR Coversyl(1)

Rule3: Blood Pressure(4) AND Gender(1) AND Glucose(5) => Coversyl(1)

Rule4: Blood Pressure(5) AND Glucose(3) AND Cholesterol(2) =>Coversyl(1) OR Coversyl(3)

Rule5: BMI(5) => Coversyl(3)

Rule6: BMI(3) AND Blood Pressure(5) AND Gender(1) AND Genetic(0) => Coversyl(1) OR Coversyl(3)

Rule7: Blood Pressure(5) AND Glucose(1) AND Cholesterol(2) AND Genetic(1) => Coversyl(2)

When rules of above applied, possible outputs seen by the given inputs. To obtain result for a given input set, we use the resultant classification and each point of rules. The root mean square error of Coversyl data set is 0.4781 and the RMSE rate is 47 percent.

Rule 1, patient's blood pressure is 4, gender is female and cholesterol is 2, then the output is 3.

Rule 2, patient's hba is 4, then the output can be 3 or 1.

Rule 3, patient's blood pressure is 4, gender is female and glucose is 5, then the output is 1.

Rule 4, patient's blood pressure is 5, glucose is 3 and cholesterol is 2. The output can be 1 or 3.

Rule 5, patient is over obese, then the output is 3.

Rule 6, patient is over-weight, blood pressure is 5, gender is 1 and genetic is 0.The output can be 1 or 3.

Rule 7, patient's blood pressure is 5, glucose is 1, cholesterol is 2 and genetic is 1, then the output is 2.

Roc curve of Coversyl data set is shown in figure 3.21:

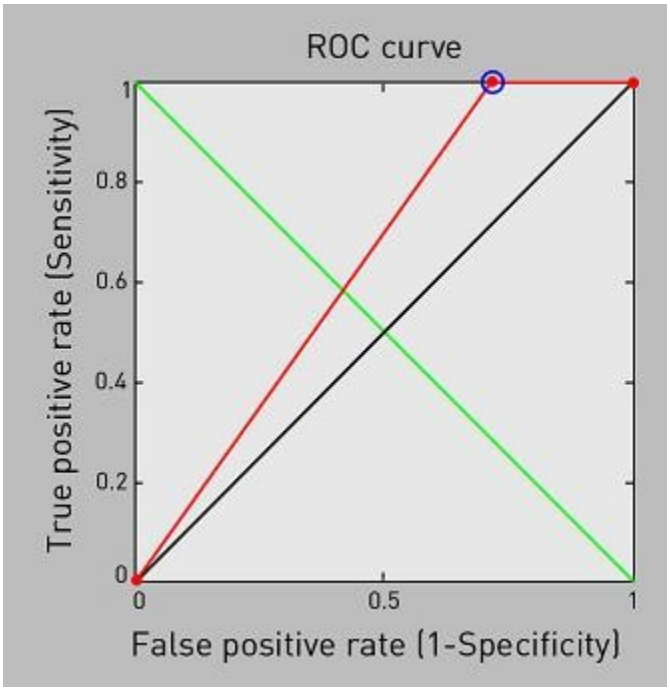


**Figure 3.21: ROC curve of Coversyl data set**

Monopril data set RSES results are introduced below. RSES algorithm generates 31 rules. Among those, I choose the specific ones and examine them. The selected rules are below:

Rule 1: BMI(2) => Monopril(2)

Rule 2: BMI(4) AND Blood Pressure(4) AND Hba(3) => Monopril(3)

Rule 3: BMI(4) AND Blood Pressure(5) AND Hba(4) AND Genetic(1) => Monopril(1)

Rule 4: Cholesterol(3) => Monopril(1) OR Monopril(2)

Rule 5: BMI(2) AND Blood Pressure(4) AND Glucose(5) => Monopril(2)

Rule 6: BMI(4) AND Blood Pressure(4) AND Hba(2) AND Glucose(3) AND Genetic(0) => Monopril(3)

Rule 7: Blood Pressure(4) AND Hba(4) AND Genetic(1) => Monopril(2)

When rules of above applied, possible outputs seen by the given inputs. To obtain result for a given input set, we use the resultant classification and each point of rules. The root mean square error of Monopril data set is 0.247 and the RMSE rate is 24 percent.

Rule 1, patient's weight is normal then the output is 2.

Rule 2, patient is obese, blood pressure is 4 and Hba is 3, then the output is 3.

Rule 3, patient is obese, blood pressure is 5, hba is 4 and genetic is 1, then the output is 1.

Rule 4, patient's cholesterol is 3, then the output can be 1 or 2.

Rule 5, patient is normal weight, blood pressure is 4 and glucose is 5, then the output is 2.

Rule 6, patient is obese, blood pressure is 4, hba is 2, glucose is 3 and genetic is 0, then the output is 3.

Rule 7, patient's blood pressure is 4, hba is 4 and genetic is 1, then output is 2.

Roc curve of Monopril data set is shown in figure 3.22:

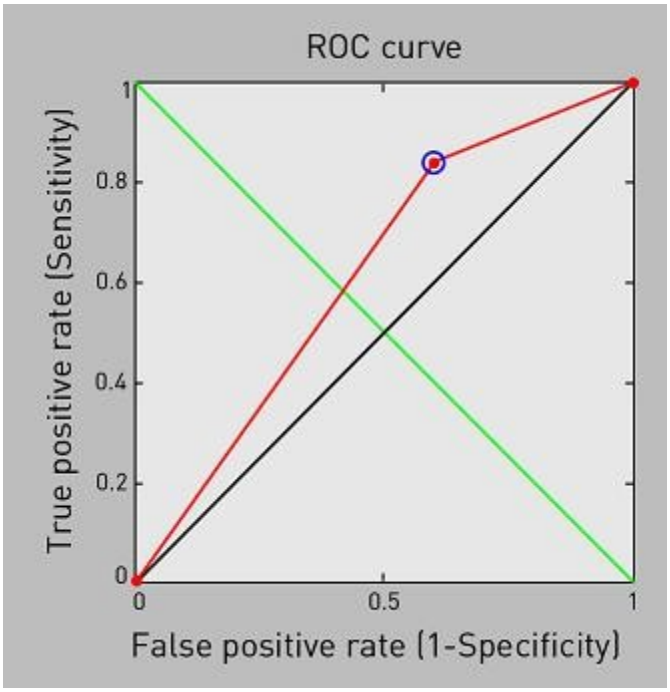


**Figure 3.22: ROC curve of Monopril data set**

Tenoretic data set RSES results are introduced below. RSES algorithm generates 31 rules. Among those, I choose the specific ones and examine them. The selected rules are below:

Rule 1: Glucose(5) => Tenoretic(1) OR Tenoretic(2)

Rule 2: BMI(4) AND Gender(0) AND Hba(3) => Tenoretic (1)

Rule 3: BMI(5) AND Gender(1) AND Hba(2) AND Mikalb(0) => Tenoretic (3)

Rule 4: Cholesterol(1) => Tenoretic (1)

Rule 5: BMI(3) AND Gender(1) AND Glucose(3) AND Protein(0) AND Ketone(0) => Tenoretic (1) OR Tenoretic(2)

Rule 6: BMI(5) AND Cholesterol(2) AND Mikros(0) => Tenoretic (3)

Rule 7: BMI(4) AND Gender(0) AND Hba(3) AND Glucose(4) AND Cholesterol(3) => Tenoretic (2)

When rules of above applied, possible outputs seen by the given inputs. To obtain result for a given input set, we use the resultant classification and each point of rules. The root mean square error of Tenoretic data set is 0.361 and the RMSE rate is 36 percent.

Rule 1, patient's glucose is 5 then the output can be 1 or 2.

Rule 2, patient is obese, gender is male and Hba is 3, then the output is 1.

Rule 3, patient is over obese, gender is female, hba is 2 and mikalb is 0, then the output is 3.

Rule 4, patient's cholesterol is 1, then the output is 1.

Rule 5, patient is normal weight, gender is female, glucose is 3, protein is 1 and ketone is 0, then the output can be 1 or 2.

Rule 6, patient is over obese, cholesterol is 2 and mikros is 0, then the output is 3.

Rule 7, patient is obese, gender is male, hba is 3 and glucose is 4 and cholesterol is 3, then output is 2.

Roc curve of Tenoretic data set is shown in figure 3.23:

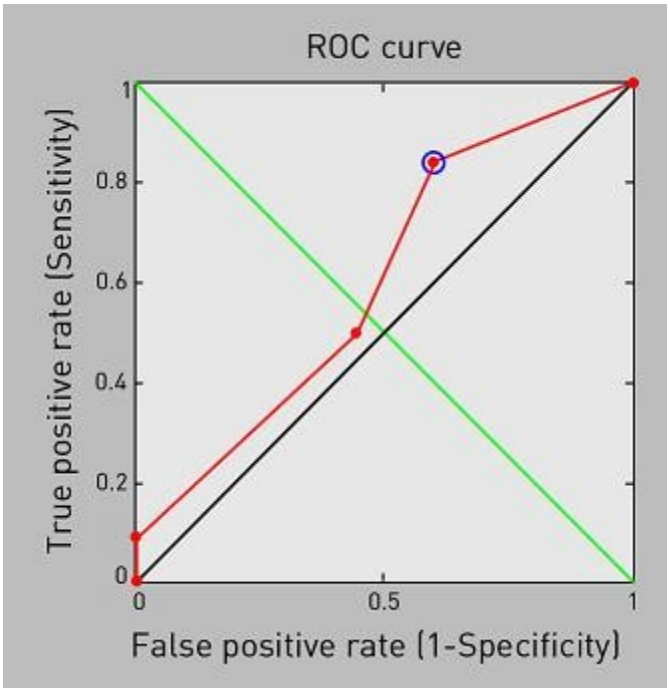


**Figure 3.23: ROC curve of Tenoretic data set**

Atacand data set RSES results are introduced below. RSES algorithm generates 21 rules. Among those, I choose the specific ones and examine them. The selected rules are below:

Rule 1: BMI(5) => Atacand (3)

Rule 2: BMI(2) AND Gender(1) AND Hba(3) => Atacand (2)

Rule 3: BMI(3) AND Gender(1) AND Glucose(5) AND Cholesterol(2) => Atacand (2)

Rule 4: Cholesterol(1,4) => Atacand (1)

Rule 5: BMI(4) AND Gender(1) AND Hba(2) AND Glucose(3) AND Cholesterol(2) => Atacand (1) OR Atacand (2)

Rule 6: Gender(1) AND Hba(3) AND Keton(0) AND Genetic(1) => Atacand (3)

Rule 7: BMI(4) AND Gender(1) AND Hba(2) AND Glucose(5) AND Cholesterol(2) => Atacand (2)

When rules of above applied, possible outputs seen by the given inputs. To obtain result for a given input set, we use the resultant classification and each point of rules. The root mean square error of Atacand data set is 0.57923 and the RMSE rate is 57 percent.

Rule 1, patient is over obese then the output can is 3.

Rule 2, patient BMI is 2, gender is female and Hba is 3, then the output is 2.

Rule 3, patient is over-weight, gender is female, glucose is 5 and cholesterol is 2, then the output is 2.

Rule 4, patient's cholesterol cluster range is between 1 and 4, then the output is 1.

Rule 5, patient is obese, gender is female, hba is 2, glucose is 3 and cholesterol is 2, then the output can be 1 or 2.

Rule 6, patient is female, hba is 3, ketone is 0 and genetic is 1, then the output is 3.

Rule 7, patient is obese, gender is female, hba is 2 and glucose is 5 and cholesterol is 2, then output is 2.

Roc curve of Atacand data set is shown in figure 3.24:

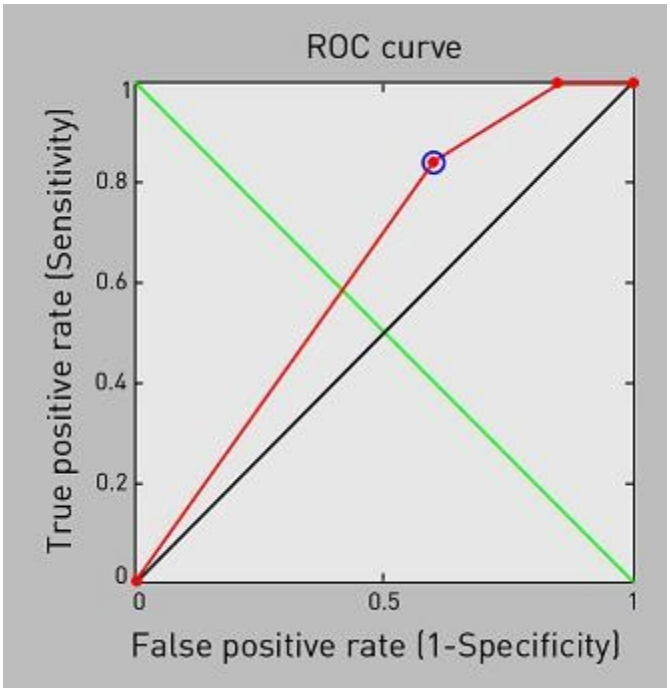


**Figure 3.24: ROC curve of Atacand data set**

# 4. CONCLUSION

Hypertension is very common and one of the dangerous disease in the world. There are many works have been done to treat this disease correctly all over the world. Data mining technique is one of the best technique in these works and there are lots of data mining studies can found about the hypertension. During these studies the most important part is to find the right data and perpetrate them correctly.

The aim of this thesis study is dosage planning of hypertension with using data mining techniques. In this thesis study, I try to determine the amount of drug dosage that will be use for perfect treatment. Total of 203 patient data were used and these data were classified into 4 classes, because there are 4 different drug used for treatment. As a result of this classification, 4 data sets are occupied. These are Coversyl data set, Monopril data set, Tenoretic data set and Atacand data set. Weka is the first data mining technique used

in this study. Weka is used for obtaining the ranking attributes of used parameters. By these ranking values found, the parameters that have a ranking value of zero were eliminated, because the eliminated parameters were not having much more importance as the others.

The other method used in this study is ANFIS. According to the ANFIS results Blood Pressure, BMI, glucose and cholesterol parameters more dominant for most of the data sets. Also doctors used to use these parameters to decide the drug dosage. As an example, generally if a patient BMI level is very high which means obese or over obese and blood pressure level is 4 which means lower blood pressure is greater than 100 and upper blood pressure is greater than 160 (very risky), then doctor decided to begin the highest dosage of given drug. This result can be seen also in ANFIS results, so we can say we achieve successful results.

The final method used in this study to compare with ANFIS is RSES algorithm. When these two methods compared to each other, ANFIS has very acceptable results than RSES. I can say all the compared result values are better in ANFIS, so ANFIS is great method for deciding drug dosage for hypertension patients.

I hope this thesis study will be helpful for correct treatment of hypertension for patients and doctors.

# 5. REFERENCES


1. Bigus, J. P. 1996. Data mining with neural networks. New York: McGraw Hill.

2. Cardiovascular Diseases, "Prevention and Control", WHO, 2001-2002.

3. Carretero, O. A., & Oparil, S. (2000). Essential hypertension. Part I: Definition and etiology. Circulation, 101, 329–335.

4. Chao-Ton Su, Chien-Hsin Yang, "Feature selection for the SVM: An application to hypertension diagnosis", 2006, pp.754-763.

5. Folkow, R. (1993). The pathophysiology of hypertension. Differences between young and elderly patients. Drugs, 46(Suppl), 2 (see also pp. 3–7).

6. G. Mancia, A. Ferrari, L. Gregorini, et all: "Blood pressure and heart rate variabilities in normotensive and hypertensive human beings", Circ. Res. 1983; 53: 96-104.

7. Guidelines Subcommittee (1999). 1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Journal of Hypertension, 17(2), 151–183.

8. Guojun Zhang, "An Improved Hypertension Prediction Model Based on RS and SVM in the Three Gorges Area", College of computer science & technology, Huazhong University of Science and Technology, Hubei,China, 2008, pp.810-814.

9. J.L. Palma, M.D. Isasa, M.A. Recarte, V. Palazuelos, E. Lbpez: "Study and classification of circadian patterns of blood pressure in both normotensive and hypertensive patients", Joumal of Ambulatory Monitoring 1992;5:21-31.

10. M. Burnier, M.P. Schneider, A. Chiolero, C.L. Fallab, H.R. Brunner, Objective monitoring of drug compliance: an important step in the management of hypertension resistant to drug therapy, Am. J. Hypertens. 12 (1999) 129.

11. Mancia, G., Gamba, P.L., and Omboni, S., Ambulatory Blood Pressure Monitoring, J. Hypertens., 1996, vol. 14, p. S61.

12. McInnis BC, Deng LZ: Automatic control of blood pressures with multiple drug inputs. Ann Biomed Eng 13: 217-225, 1985.

13. Mevlut Ture, Imran Kurt, A. Turhan Kurum, Kazim Ozdamar, "Comparing classification techniques for predicting essential hypertension", Expert Systems, 2005, pp.583-588.

14. R. Poli, S. Cagnoni, G. Coppini, R. Livi, G. Val& “A Neural Network Expert System for Diagnosing and Treating Hypertension,” Computer, Vo1.24, No.3, Mar.1991, pp. 64-71.

15. S Cagnoni, G Coppini, R Livi, R Poli, PT Scarpelli, G. Valli, “A Neural Network Expert System for Computer-Assisted Analysis of Blood-Pressure Data”, 1992, pp.473-476.

16. S. Polak, A. Mendyk “Artificial neural networks based Internet hypertension prediction tool development and validation”, Science Direct 2007, pp. 734-739.

17. Slate, J. B., & Sheppard, L. C. (1982). Automatic control of blood pressure by drug infusion. IEE Proceedings, 129A (9), 639-645.

18. Smulyan H, Safar ME, “The diastolic blood pressure in systolic hypertension”, Ann Intern Med., 2000

19. T.R. Rebbeck, S.T. Turner, C.F. Sing, Probability of having hypertension: effects of sex, history of hypertension in parents and other risk factors, J. Clin. Epidemiol. 49 (1996) 727–734.

20. V. G. Vilkov, R. G. Oganov, and S. A. Shal’nova, “Comparative Accuracy of Neural Network Models for Diagnosing Latent Arterial Hypertension on the Basis of Data on Daily Blood Pressure Monitoring”, State Research Center for Preventive Medicine, Ministry of Health of the Russian Federation, Moscow, Russia, Mar. 2005, pp. 657-661.

21. Y. Hayashi, “Neural Expert System using Fuzzy Teaching Input and Its Application to Medical Diagnosis”, Proc. of the 2nd Intl. Conf. on Fuzzy Logic & Neural Networks (IIZUKA’96), pp.989-993( 1992).

22. Y. M. Chae, S.H. Lee, S.H. Ho, M.Y. Bae, H.C. Ohrr, Medical decision support system for the management of hypertension, Informatica 21 (1997) 219–225.

23. Y. Sirinivas, W. D. Timmons, J. Durkin “A comparative study of three expert systems for blood pressure control”, Expert Systems 2001, pp. 267-274.

24. Yamashita Y, Suzuki M, Kambe K: Fuzzy control of blood pressure by drug infusion. J Chem Eng Japan 21: 541-543, 1988.

25. Ying H, Sheppard LC: Real-time expert-system-based fuzzy control of mean arterial pressure in pigs with sodium nitroprusside infusion. Med Prog Technol 16: 69-76, 1990.

26. http://en.wikipedia.org/wiki/Hypertension

27. http://www.humanillnesses.com/original/Her-Kid/Hypertension.html

28. http://hypertension.emedtv.com/prehypertension/prehypertension.html