International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 1 Jan-Feb 2021, pp: 672-678 www.ijaem.net ISSN: 2395-5252

Machine Algorithm Learning and Artificial Intelligence Diagnosis for Liver Cancer Treatment

Yakup Yasin Ozsoy

[1,2]*Master Tesis Student, ISTANBUL AYDIN University Department ofComputer Engineering, Florya, Istanbul, Turkey*

Submitted: 20-01-2021 Revised: 31-01-2021 Accepted: 03-02-2021

ABSTRACT:This research paper is about the information about Liver Cancer is obtained with definitions and some applications in machine language and some algorithms, and with the application I developed, it measures the similarity of the element I specified with the other elements in that array.The ultimate goal is to increase productivity and diversity to be more productive, to be able to look more at liver disease diagnoses and to develop perspectives with calculations.

KEYWORDS: Nash, Liver, Class algorithms , Artificial intelligence.

I. INTRODUCTION

In this paper The liver is known to have hundreds of functions in the body. The main functions are:

1: Production

The liver plays a central role in controlling blood sugar levels. It is very important for the production of bile, protein, fat and some hormones.

2: Detoxification

The liver filters all molecules in the body (drugs, alcohol, food ...)

3: The function of oils

The liver organ processes the fats in our body, and if we talk about them, it is understood that these are production, transformation and transportation. Although the liver does not act as fat storage, it may contain fat "droplets" up to 5% of the cells in a healthy liver.Although the liver is an exceptional organ, it is an incredible regeneration when it is damaged.has the capacity.

The liver is an important organ of ours that is at the center of basic metabolic processes. It balances the good cholesterol (HDL) and bad cholesterol (LDL). This means that the liver is of great importance in the proper functioning of the cardiovascular system.

In antiquity, the liver was seen as the center of our emotions, a role now attributed to the heart. People thought that the starting point for anger and sadness was their livers.The liver is the organ where cancers most frequently spread (metastasize). Metastases are the most common tumors among liver diseases. If we refer to the definition of metastasis, it is the spread of cancer in another organ or tissue of the body to the liver. Cancers in any part of our body can spread to the liver.

Primary liver tumors are malignant tumors that develop from the liver's own tissue. Those developing from hepatocytes forming the liver parenchyma are called "hepatocellular carcinoma, HCC" and those developing from the bile ducts are called "cholangiocellular carcinoma, CSC".

Non-alcoholic fatty liver (NAFLD) and its subgroup, non-alcoholic steatohepatitis (NASH), is another risk factor for HSC. NASH is a liver disease in which fatty liver is accompanied by inflammation (inflammation) that can progress to cirrhosis. Nash is a rapidly growing cause of liver disease driven by the obesity epidemic.

Fatty liver, also known as hepatic steatosis, is a condition where too much fat is stored and accumulated in liver cells. In general, there is a small amount of fat in the liver, but excessive amounts of it cause various health problems.

The liver helps digest the nutrients we get from food and beverages and cleans the blood by filtering harmful substances in the blood. However, too much fat in the liver can cause liver inflammation, which damages the liver and causes scar tissue formation. In more severe cases, these scar tissues can lead to liver failure.

Fatty liver is divided into two types according to its cause; These are divided into two as alcoholic and non-alcoholic fatty liver. When fatty liver occurs in an individual who consumes more alcohol, it is defined as fatty liver disease (AFLD) due to alcohol. Fatty liver developing in individuals who do not consume alcohol or use very little alcohol is defined as non-alcoholic fatty liver (NAFLD).

In some individuals with more severe liver fat, advanced scar tissue may occur due to liver inflammation, ie liver scarring.

Liver scarring is also called liver fibrosis. If severe liver fibrosis develops, this can lead to life-threatening diseases such as cirrhosis. As a result of cirrhosis, an unfavorable situation that progresses to liver failure occurs.

People suffering from non-alcoholic fatty liver may also develop non-alcoholic steatohepatitis (NASH), an aggressive form of fatty liver disease. The damage seen on this type of liver tissue is similar to the damage caused by heavy alcohol use and again causes liver failure. In addition, the root cause of NASH is still unknown.

In many cases, fatty liver usually does not cause any obvious symptoms. Among the symptoms of fatty liver, there may be cases of severe pain in the upper right side of the abdomen or when the individual feels extreme tiredness. In rare cases of liver disease caused by fatty liver, various symptoms can be seen on the skin. Redness due to itching and itching may occur. Again, scaling is one of the symptoms of fatty liver and cirrhosis on the skin. Apart from these, symptoms of cirrhosis include:

Loss of appetite, weight loss, weakness, fatigue, nosebleeds, yellow skin,Vascular groups visible under the skin, Abdominal pain, Leg swelling,Breast enlargement in men, Confusion.NAFLD / NASH is blamed for the increase in HSC frequency. The frequency of NAFLD in the world is high (25-30%). These patients can develop HSC without cirrhosis. Strategies are being developed for HSC screening in NAFLD / NASH patients..

Nash (Non AlcoholicSteatoHepatitis) is a more severe form of non-alcoholic fatty liver disease called NAFLD. Nash (Non AlcoholicSteatoHepatitis) is a form of liver disease that belongs to non-alcoholic fatty liver disease. Generally, NASH occurs as a result of a high sugar, high fat diet and inadequate physical exercise. NASH is closely linked to overweight, obesity, and type 2 diabetes. Thus, it is defined as a metabolic disease. Although NASH is an asymptomatic type of disease, unfortunately patients do not show any symptoms until the advanced stage. Thus, NASH is a difficult disease to diagnose.

By the way; The name Hepatitis in the definition of NASH can bring the definition of Hepatitis virus to mind. Hepatitis is not only used to refer to viral diseases i.e. Hepatitis B, C. The word hepatitis is derived from the Greek hepatos, which means liver. Liver disease is defined as hepatitis as it provokes theinflammationof the liver.

If nothing is done to prevent the progression of the disease, the liver can initiate the healing process. Scar tissue forms on the liver: This tissue is called "fibrosis". If this scar tissue does not work properly, the function of the liver is deteriorating day by day. Fibrosis can be classified into four stages: from one to three (mild, moderate, severe), fibrosis evolution can be halted or even reversed.

However, if fibrosis spreads to most of the liver, then stage 4 fibrosis called "cirrhosis" has been reached. Unfortunately, this universe does not come back.Nonalcoholic fatty liver disease is the spectrum of chronic liver disease ranging from simple steatosis to non-alcoholic steatohepatitis and is strongly associated with metabolic syndrome.

Nonalcoholic fatty liver disease significantly increases the prevalence of the development of hepatocellular carcinoma (HCC); however, increased risk of HCC is often misdiagnosed in patients with nonalcoholic fatty liver disease.

To associate non-alcoholic fatty liver disease progression with life-threatening complications, the degree of fibrosis is considered to be the strongest predictive factor.

Various factors contribute to the development of non-alcoholic fatty liver disease or Nash and later HCC development; these factors include genetic and environmental modifiers such as diet or lifestyle.

The pathogenesis of HCC associated with nonalcoholic fatty liver disease is a complex landscape of immune and inflammatory responses, DNA damage, oxidative stress, and the mechanism involved in autophagy. Today, the diagnosis of HCC-associated non-alcoholic fatty liver disease depends on imaging, whereas appropriate HCC staging to assess prognosis is necessary.

II. HOW ARTIFICIAL INTELLIGENCE AFFECTS THE DIAGNOSIS AND TREATMENT OF LIVER DISEASE AND MEDICAL ARTIFICIAL INTELLIGENCE

In recent years, there has been a significant increase in computing capacities and machine language software and the data volumes that enable the progress with these software and their storage. With this increase, the tendency to use artificial intelligence in health services is increasing. These advances make it easier to diagnose a variety of conditions from atrial fibrillation to stroke and to treat others such as depression and anxiety.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 1 Jan-Feb 2021, pp: 672-678 www.ijaem.net ISSN: 2395-5252

Fields such as hepatology are seeing significant changes thanks to such developments, but it is one of the important points to be considered for successful integration and implementation.The term artificial intelligence encompasses many techniques, from advanced statistical modeling to black box applications to deep learning algorithms. The field that provides the most important new applications in healthcare is machine learning, where the machine can be processed with its complex and nonlinear relationships between interests and variables.

In order to be able to update itself, the algorithms are fed with large amounts of data containing variables mapped to the results of interest. In doing so, they can reveal previously unidentified relationships, saying that traditional statistical methods cannot. ML also analyzes data types that were previously unsuitable for advanced computer-based analysis, such as imaging and text data.

These techniques have many potential uses in hepatology, from finding new blood marker variation patterns to treat or diagnose liver disease, automating image analysis, predicting or diagnosing liver disease, to automating image analysis to identify new blood marker variation patterns to diagnose; From identifying areas of the liver at risk of irradiation toxicity to using the drug structure to predict the risk of liver injury. All of them can increase diagnostic accuracy, improve decision making by improving predictive capabilities, and increase efficiency through automation.

III. CLASSIFICATION ALGORITHMS IN MACHINE LEARNING: HOW THEY WORK AND TOP 5 CLASSIFICATION ALGORITHMS IN MACHINE LEARNING

Classification is one of the most important and concepts in data science. Classification algorithms are predictive calculations used to analyze sets of training data and assign data to predetermined categories.

What is Classification?

Classification is the process of recognizing, understanding, and grouping ideas and objects into predetermined categories or "sub-populations". Using pre-modernized training datasets, machine learning programs are used in various algorithms to categorize future datasets.

Classification algorithms in machine learning use the training data that subsequent data reaches to predict the probability of falling into one of the predetermined categories. One of the most used in classification is that they filter emails as "spam" or "non-spam."

To evaluate, the classification is to use similar patterns like words or similar thoughts, number sequences, etc. in future datasets. We can define it as a form of "pattern recognition" with classification algorithms applied to training data to find it.

By using the classification algorithms we will describe in more detail, text analysis software can be used to perform studies such as sensitivity analysis to categorize unstructured text according to the poles of view (positive, negative, neutral and beyond).

Previously trained sensitive classifications should be run to understand how the classification algorithms work in the application.

Top 5 Classification Algorithms for Machine Learning

Classification work in statistics is very extensive, and there are various classification algorithms you can use depending on the dataset you are working on. Below are five of the most common algorithms in machine learning.

Various classification algorithms:

Logistic regression

Naive Bayes Classifier

K-Nearest Neighbors

Decision tree

Random Forest

Supporting vector machines

Logistic regression

Logistic regression is a calculation used to predict a binary result: something happens or not. This is Yes / No, Pass / Fail, Alive / Dead etc. can be specified as.

Independent variables are analyzed to determine the binary result with results falling in one of the two categories. Independent variables can be categories or numeric, but the dependent variable is always categorical. It is stated as follows:

$P(Y = 1 | X)$ or $P(Y = 0 | X)$

When X is defined as the independent variable, the probability of the dependent variable Y is calculated.

It can be used to calculate the probability of a word with a positive or negative connotation (on a scale of 0, 1, or between). Or the object (tree, flower, grass, etc.) contained in a photograph, they are used to indicate probability data between 0 and 1 for each object.Applications of Classification Algorithms

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 1 Jan-Feb 2021, pp: 672-678 www.ijaem.net ISSN: 2395-5252

We all understand the mathematical operations associated with classifications, but what can these machine learning algorithms do with real-world data?
Emotion Analysis
Email Spam Classification
Document Classification
Image Classification
Emotion Analysis

IV.DATA SET INFORMATION

This data set includes 416 liver patient records and 167 non-hepatic patient records. The data set was collected from the north east of Andhra Pradesh, India. Selector is a class tag used to categorize into groups (with or without liver disease). This data set includes 441 male patient records and 142 female patient records. Every patient over the age of 89 is listed as "90".
Attribute Information:
1. Age of Patient
2. Gender of the Patient's Sex
3.TB Total Bilirubin
4. DB Direct Bilirubin
5. Alkphos Alkaline Phosphatase
6.Sgpt Alamin Aminotransferase
7.Sgot Aspartate Aminotransferase
8. TP Total Protiens
9.ALB Albumin
10.A / G Ratio Albumin and Globulin Ratio
11. Selector field used to divide the data into two groups (labeled by experts)

From the statistics, it was determined that the minimum age was 4 and the maximum age was 90. Based on the information in this data set, everyone over the age of 85 was evaluated as 90. So we can change that. A new data framework is being created. Also, we see that there are missing values in the "Albumin_and_Globulin_Ratio" column, which we can briefly discuss. Finally, the healthy patient ranges will be calculated and found to understandwhereeach patient lies.

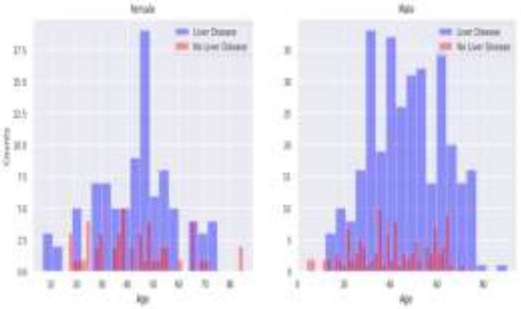
In the Albumin and Globulin Ratio property column, there are only 4 missing values equal to 1% of all data.

Healthy ranges of feature results
Healthy Ranges for 10 property columns
Total Bilirubin = 0.1 to 1.2 mg / dL = 1.71 to 20.5 umol / L
Direct_Bilirubin = <0.3 mg / dL = <5.1 umol / L
Alkaline_phosphatase = 44 to 147 IU / L (High ALP levels are seen in children and pregnant women undergoing growth)
Alamine_Aminotransferase = 29 to 33 IU / L (Age and gender can affect value)
Aspartate_Aminotransferase = 1-45 U / L (Values are slightly lower in women) Total_proteins = 6.0 - 8.3 g / dL
Albumin = 3.4 - 5.4 g / dL
Albumin_and_Globulin_Ratio = Adult: 3.7 to 5.2 g / dL; Older Adult: 3.2 to 4.6 g / dL; > 90 years old: 2.9 - 4.5 g / dL
Note: These values may differ according to different guidelines or hospitals. The above values are taken fromgoogle

Coping with missing values

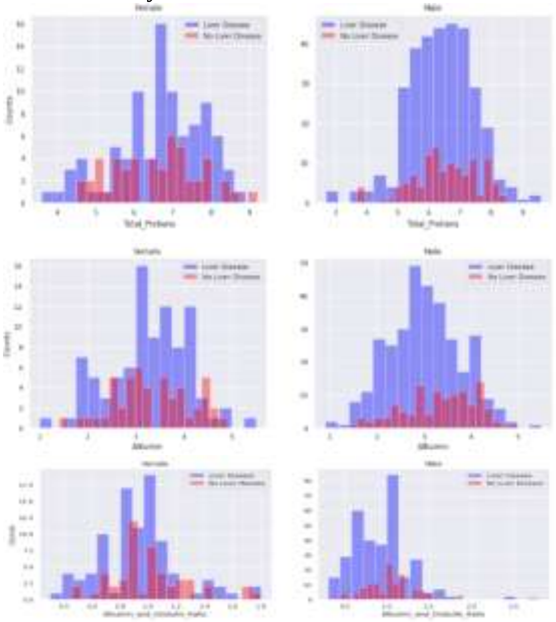
Total	%	
Albumin_and_Globulin_Ratio	4	1.0
Dataset	0	0.0
Albumin	0	0.0
Total_Protiens	0	0.0
Aspartate_Aminotransferase	0	0.0
Alamine_Aminotransferase	0	0.0
Alkaline_Phosphotase	0	0.0
Direct_Bilirubin	0	0.0
Total_Bilirubin	0	0.0
Gender	0	0.0
Age	0	0.0

Scatter charts for all liver function tests



This distribution graph is what it means to us,

1) On average, women tend not to have liver disease than men.
2) The largest number of women without liver disease was around 38 years old.
3) Girls around the age of 10 had liver disease, this may be a genetic link.
4) Men are more prone to liver disease (may be caused by alcoholism).
5) The largest number of men without liver disease was about 38 years old.



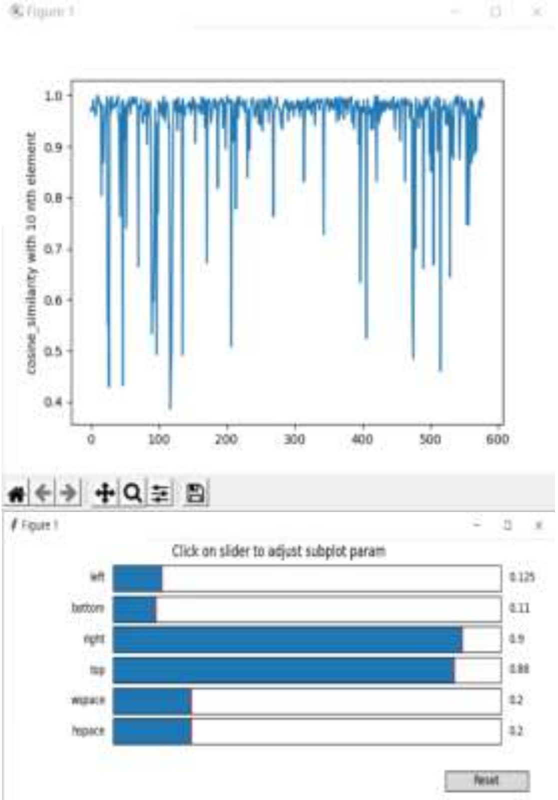
Calculating with cosine scikit-learn
We call the cosine method from within sklearn, it draws the graph and prints the similarities to us on the screen.
Exactly how you use it to the terminal
python cosine.py 10
You write the 10th element in csv finds how similar it is to others
or python cosine.py 27
You write, the 27th element in the csv finds how similar it is to the others and measures the neighborhood distance. Measures 27th element similarity to other elements in that array
27 elements in an exepction those that look like 27 elements 48, 120, 474 ...
Now these elements look alike if the cancer risk is high because I do not know the details of the data. Or you can add a new line to our csv and look at its status. Its data is close to whatever group it looks like. cosine comparison method The 27th element has 33% similar and 89% similar.

```
import numpy as np
import pandas as pd
import os
```

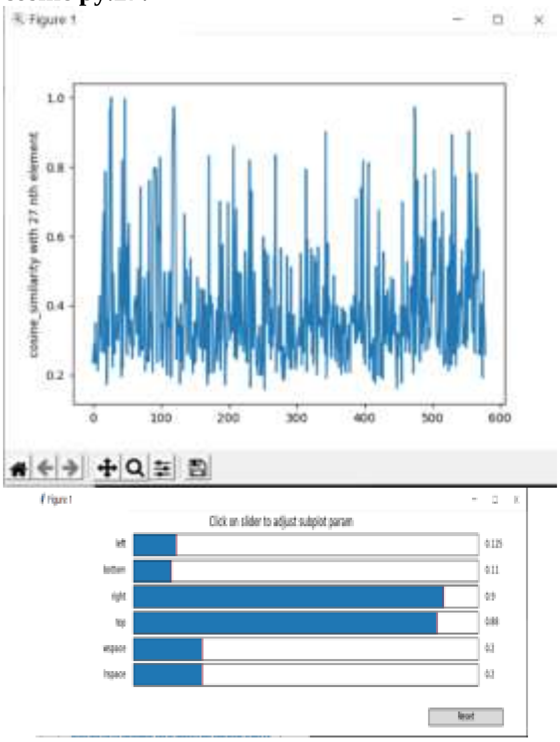
```
from sklearn.metrics.pairwise import cosine_simila
rity
import sys
import matplotlib.pyplot as plt

arg = 0
if len(sys.argv) > 1:
    arg = int(sys.argv[1])

for eachArg in sys.argv:
    print(eachArg)

path = os.getcwd()
print(arg)
my_data=pd.read_csv(path +'/indian_liver_patient.
csv', sep=',',header=0 )
original_headers = list(my_data.columns.values)
my_data.dropna(inplace=True)
numeric_headers = list(my_data.columns.values)
numpy_array = my_data.values

temp =numpy_array[arg].reshape(1,-1)
print(temp)
sim = cosine_similarity(numpy_array,temp)
print(arg,' element similar with : ', sim)

plt.plot(sim.flatten() )
plt.ylabel('cosine_similarity with ' +str(arg)+ ' nth e
lement')
plt.show()

from math import sqrt

def get_cosine_similarity(A, B):


    dot_prod = get_dot_prod(A,B)
    eucl_magn = get_eucl_magn(A,B)
    return dot_prod / eucl_magn if eucl_magn else N
one

def get_dot_prod(A,B):
    dot_prod = 0
    for a, b in zip(A, B):
        dot_prod += a * b
    return dot_prod

def get_eucl_magn(A,B):

    A_len = 0
    B_len = 0
    for a, b in zip(A,B):
        A_len += a * a
        B_len += b * b
    return sqrt(A_len * B_len)
```

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 1 Jan-Feb 2021, pp: 672-678 www.ijaem.net ISSN: 2395-5252

cosine py.10:



cosine py.27:



V. CONCLUSION

Looking at this project, the overall result of the results will be observed. This can be judged by how well our goals are met. One of our goals was to do what happens with the general algorithms and you will see the result with them. Applications of machine language with various approaches and their results. With various approaches I found thatmy best model is Logistic regression.

Score	Model	Accuracy
71.329	Logistic Regression	72.000
70.860	GB Classifier	64.571
70.110	MLP Classifier	72.000
69.134	Random Forest Classifier	68.571
67.152	KNeighborsClassifer	65.143
64.957	Decision Tree Classifier --- gereksizzz	65.143
55.854	GaussianNB	54.857

REFERENCES

[1]. Yu SJ. A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010-2016. Clin Mol Hepatol. 2016Mar;22(1):7-17.doi: 10.3350/cmh.2016.22.1.7.

[2]. Heimbach J, Kulik LM, Finn R, Sirlin CB, Abecassis M, Roberts LR, Zhu A, Murad MH, Marrero J. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology. 2017 Jan 28. doi: 10.1002/hep.29086.

[3]. Manzini G, Henne-Bruns D, Porzsolt F, Kremer M. Is There a Standard for Surgical Therapy of Hepatocellular Carcinoma in Healthy and Cirrhotic Liver? BMJ Open Gastro. 2017;4(1):e000129

[4]. Khan SA, Davidson BR, Goldin RD, Heaton N, Karani J, Pereira Sp, Rosenberg WMC, Tait P, Taylor-Robinson SD, Thillainayagam AV, Thomas HC, Wasan H. Guidelines for the diagnosis and treatment of cholangiocarcinoma: an update. Gut 2012;61:1657–1669. doi:10.1136/gutjnl-2011-301748

[5]. Alvaro D. The Challenge of Cholangiocarcinoma Diagnosis: The Turning Point Is in Extracellular Vesicles? J Hepatol (2016), http://dx.doi.org/ 10.1016/j.jhep.2016.04.001

[6]. European Association for the Study of the Liver (EASL). EASL Clinical Practice Guidelines on the management of benign liver tumours. J Hepatol. 2016 Aug;65(2):386-98. doi: 10.1016/j.jhep.2016.04.001. Epub 2016 Apr 13.

[7]. European Association For The Study Of The Liver; European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma.J Hepatol. 2012 Apr;56(4):908-43. doi: 10.1016/j.jhep.2011.12.001.

[8]. Ma, H., Xu, C.-F., Shen, Z., Yu, C.-H., Li, Y.-M. Application of machine learning techniques for clinical predictive modeling: A cross-sectional study on nonalcoholic fatty liver disease in China. Biomed Res Int. 2018; 2018: 4304376

[9]. Kuppili, V., Biswas, M., Sreekumar, A., Suri, HS, Saba, L., Edla, DR et al, Overlearning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization. J Med Syst. 2017; 41: 152

[10]. Byra, M., Styczynski, G., Szmigielski, C., Kalinowski, P., Michałowski, Ł, Paluszkiewicz, R. et al. Learning transmission with deep convolral neural network for liver steatosis assessment on ultrasound images. International J Comput Assist Radiol Surg. 2018; 13: 1895--1903

[11]. Biswas, M., Kuppili, V., Edla, DR, Suri, HS, Saba, L., Marinhoe, RT, et al. Symtosis: liver ultrasound tissue characterization and risk stratification in an optimized deep learning paradigm. Computer Methods Programs Biyomed. 2018; 155: 165--177

[12]. Cristoferi, L., Nardi, A., Ronca, V., Invernizzi, P., Mells, G., Carbone, M. Prognostic models in primary biliary cholangitis. J Autoimmun. 2018; 95: 171--178

[13]. Bertsimas, D., Kung, J., Trichakis, N., Wang, Y., Hirose, R., Vagefi, PA Development and validation of optimized mortality prediction for liver transplant candidates. J Transplant. 2018;

[14]. Briceño, J., Cruz-Ramírez, M., Prieto, M., Navasa, M., Ortiz de Urbina, J., Orti, R.et al., Artificial donor-recipient matching model for the liver The use of intelligence transplant: results of a multicenter Spanish study. J Hepatol. 2014; 61: 1020--1028

[15]. Dou, Q., Yu, L., Chen, H., Jin, Y., Yang, X., Qin, J. et al. 3D deeply controlled network for automatic segmentation of volumetric medical images. Med Image Anal. 2017; 41: 40--54

[16]. Ibragimov, B., Toesca, D., Chang, D., Yuan, Y., Koong, A., Xing, L. Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT. Med Phys. 2018; 45: 4763--4774

[17]. Saini, N., Bakshi, S., Sharma, S. In-silico approach for drug-induced liver injury prediction: recent developments. Toksikol Lett. 2018; 295: 288-295

[18]. MARMARA UNIVERSITY SULTANAHMET CAMPUS MEETING ROOM June 12, 2018: https://dosyaism.saglik.gov.tr/Eklenti/32577,1-uluslararasi-nas-nedir-bilgilendirmepdf.pdf?0

[19]. MARMARA UNIVERSITY SULTANAHMET CAMPUS MEETING ROOM June 12, 2018: https://dosyaism.saglik.gov.tr/Eklenti/32575,1-uluslararasi-nas-aciklayici-rehberpdf.pdf?0 Bendi

[20]. Venkata Ramana, Prof. MS Prasad Babu and Prof. NB Venkateswarlu, “A Critical Comparative Review of Liver Patients from the USA and India: An Exploratory Analysis”, International Journal of Computer Science Problems, ISSN: 1694-0784, May 2012.Bendi

[21]. Venkata Ramana, Prof. MS Prasad Babu and Prof. NB Venkateswarlu, “A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis”, International Journal of Database Management Systems (IJDMS), Vol. 3, No. 2, ISSN: 0975-5705, PP 101-114, May 2011.

International Journal of Advances in Engineering and Management
ISSN: 2395-5252

IJAEM

Volume: 03 Issue: 01 DOI: 10.35629/5252

www.ijaem.net

Email id: ijaem.paper@gmail.com