COST FORECASTING FOR MEDICAL RESOURCE OF PATIENTS WITH ACUTE HEPATITIS IN EMERGENCY DEPARTMENT

R. J. Kuo
Department of Industrial Management National Taiwan University of Science and Technology, Taipei, Taiwan
rjkuo@mail.ntust.edu.tw

W. C. Cheng
Department of Industrial Management National Taiwan University of Science and Technology, Taipei, Taiwan
jwz.cheng@gmail.com

W. C. Lien
Department of Emergency Medicine, College of Medicine, National Taiwan University and National Taiwan University Hospital, Taipei, Taiwan
dtemer17@yahoo.com.tw

T. J. Yang
Department of Emergency Medicine, Sijih Cathay General Hospital, Taipei, Taiwan
yangteckjin@hotmail.com

ABSTRACT

Taiwan is a hyper endemic area of Hepatitis B virus (HBV). The estimated total number of HBsAg carriers in the general population for more than 20 years old is 3,067,307. Thus, a case record review was conducted from January 2003 to June 2007 of all patients admitted to the Emergency Department (ED) for a well-known teaching-oriented hospital with diagnosis of acute hepatitis. The cost of medical resource utilization is defined as the total medical fee. In this study, support vectors regression (SVR), and artificial neural network (ANN) are employed. A total of 117 patients met the inclusion criteria. 61% patients involved in this study are hepatitis B related. The computational result indicates that the SVR model has better performance. In conclusion, Child-Pugh score and echogram could be used simultaneously to predict the medical resource cost in patients of acute hepatitis without liver cirrhosis in the ED.

INTRODUCTION

Acute hepatitis lasts less than 6 months while chronic hepatitis lasts longer than 6 months. Acute hepatitis has several possible causes, such as Infectious viral hepatitis (hepatitis A, B, C, D, and E), other viral diseases (glandular fever and cytomegalovirus), severe bacterial infections, amoebic infections, medicines (acetaminophen and halothane), and toxins (alcohol and fungal toxins). The severity of illness in acute hepatitis ranges from asymptomatic to fulminant and fatal. Some patients are asymptomatic with abnormalities noted only by laboratory studies, while other patients might have symptoms and signs, such as nausea, vomiting, fatigue, weight loss, abdominal pain, jaundice, fever, splenomegaly, or ascites.

Chronic hepatitis also has several different causes, like Infectious viral hepatitis (hepatitis B, C, D), drug reactions, alcohol, autoimmune hepatitis, Wilson's disease, and hemochromatosis.

According to the World Health Organization (WHO), of the 2 billion people infected with the hepatitis B virus (HBV), more than 350 million have chronic (lifelong) infections (Lee, 1997). Hepatitis B results in 600,000 deaths each year from cirrhosis and hepatocellular carcinoma (Tujios & Lee, 2012).

Taiwan is a hyperendemic area of Hepatitis B virus (HBV). The estimated total number of HBsAg carriers in the general population > 20 years old is 3,067,307, and 61% patients involved in this study are hepatitis B related (Chen et al., 2007).

The typical presentation of severe acute exacerbation is a short onset of jaundice and very high ALT level, sometimes preceded by prodromal constitutional symptoms, in a patient with chronic hepatitis B. It is often misdiagnosed as acute hepatitis especially in those with negative history of hepatitis in the past or had never been examined for hepatitis marker in the past (Chan et al., 2002). The symptoms of severe acute exacerbation of chronic hepatitis B can be very similar to those of acute hepatitis B. Hence, severe acute exacerbation of chronic hepatitis B might be misdiagnosed as acute hepatitis B in some cases (Kumar et al., 2006).

In countries with intermediate or high endemicity for HBV, exacerbations of chronic hepatitis B may be the first presentation of HBV infection (Orenbuch-Harroch et al., 2008). Although these exacerbations are usually transient and asymptomatic, 1%–2.4% of patients later develop hepatic decompensation (Davis & Hoofnagle, 1985; Fattovich et al., 1990; Sheen et al., 1985). In this study, we believe that great majority of patients with acute hepatitis B suffered from acute exacerbation of chronic hepatitis B.

The Child-Pugh classification has been used for decades to measure the severity of chronic liver disease. Recent studies have shown that the model for end-stage liver disease (MELD) more accurately predicts the short and mid-term survival for patients
with cirrhosis compared to the CTP system. MELD, which has 3 parameters (serum bilirubin, creatinine, and prothrombin time) that need logarithmic transformation, has the advantage of a wide-range continuous scale and is more objective and less variable (Kamath et al., 2001).

Reports of predictors for acute hepatitis include MELD Scoring System, Discriminant Function (DF), and Multivariate analysis. However, there are limited reports using abdominal echogram and the Child-Pugh classification as a predictor for acute hepatitis.

Sheth et al. (2002) used MELD score and the Discriminant Function (DF) as a predictor of mortality in 34 patients hospitalized with alcoholic hepatitis. The MELD score performs as well as the DF in predicting mortality at 30 days. A MELD score of greater than 11, or the presence of both ascites and an elevated bilirubin greater than 8 mg/dL should prompt consideration of specific therapeutic interventions to reduce mortality.

Li et al. (2008) tried to find out the prognostic factors for chronic severe hepatitis and constructed a prognostic model. The clinical and laboratory indices of 213 patients with chronic severe hepatitis within 24 hours after diagnosis were analyzed retrospectively. Death or survival was limited to within 3 months after diagnosis. The mortality of all patients was 47.42%. Compared with the survival group, the age, basis of hepatocirrhosis, infection, degree of hepatic encephalopathy (HE) and the levels of total bilirubin (TBil), total cholesterol (CHO), cholinesterase (CHE), blood urea nitrogen (BUN), blood creatinine (Cr), blood sodium ion (Na), peripheral blood leukocytes (WBC), alpha-fetoprotein (AFP), international normalized ratio (INR) of blood coagulation and prothrombin time (PT) were significantly different in the group who died. They concluded that Multivariate analysis excelled univariate anlysis in the prognosis of chronic severe hepatitis, and the regression model was of significant value in the prognosis of this disease.

Basically, the prognosis of liver cirrhosis is assessed by using Child-Pugh score. Since abdominal ultrasound is also a commonly used tool for the evaluation and rapid diagnosis of acute hepatitis in the ED, the objective of this study is to determine whether Child-Pugh score and abdominal ultrasound could be used simultaneously to predict the medical resource cost in patients of acute hepatitis without liver cirrhosis in the ED. In 2008, the consensus recommendations of the Asian Pacific Association for the study of the liver (APASL) indicated several prognostic scores of acute-on-chronic liver failure (ACLF) (S. K. Sarin et al., 2009). Liver-specific scoring systems (Mayo Risk Score, Combined Clinical and Laboratory Index) are adequate, but the Acute Physiology and Chronic Health Evaluation (APACHE) II and III proved to be more powerful, because they include additional physiologic parameters and
therefore also take into account additional complications associated with this liver disorder (S. Sarin et al., 2008; Zauner et al., 1996). An accurate prognosis of the patient helps appropriate selection of a treatment program. This study retrospectively analyzed the results of abdominal echogram and the Child-Pugh classification as predictors in acute hepatitis patients. The attempt is to develop a cost forecasting model.

MATERIALS AND METHODOLOGY

Materials
An electronic search was performed on the medical record database of all patients admitted to the emergency department of a Hospital in Taiwan with an admission diagnosis of hepatitis (ICD-9 code 070 or 570) from January 2003 to June 2007. These patients were then further selected for the presence of liver enzyme (AST/ALT) above 400 and exclude for the presence of liver cirrhosis. 117 patients were included in this study. 76 male and 41 female ranging from aged 21 to aged 77. The ratio of male and female is nearly 2:1.

Artificial neural network
Artificial neural network (ANN) is a system derived through models of neurophysiology. In general, it consists of a collection of simple nonlinear computing elements whose inputs and outputs are tied together to form a network. The learning algorithms of ANNs can generally be divided into three different types: supervised, unsupervised, and hybrid learning. ANNs are nonparametric data driven approaches which can capture nonlinear data structures without prior assumption about the underlying relationship in a particular problem. ANNs are more general and flexible modeling and analysis tools for forecasting applications in that not only can they find nonlinear structures, they also can model linear processes (Zhang et al., 2001).

Figure 1 The structure of ANN.
Many studies have attempted to apply ANN to time-series forecasting. However, their conclusions are often contradictory. Some studies conclude that ANNs are better than conventional methods (Weigend & Rumelhart, 1992), while others reach an opposite conclusion. The ANN approach is a leading contender among statistical modeling approaches. ANN has been found to be useful techniques for modeling any industries (Fildes et al., 2008; Lu & Wang, 2010).

There are several procedures that must be followed when applies this method. 

Step 1. Set up the parameters including learning rate ($\eta$), momentum ($\alpha$), number of training iterations.

Step 2. Set up the connecting weight $W_{xh}$ (input layer to hidden layer), $W_{hy}$ (hidden layer to output layer), and the bias weight $\theta_h$ (in hidden layer), $\theta_y$ (in output layer) randomly.

Step 3. Input the training data.

Step 4. Compute the output $Y$ by Eq. (1) to Eq. (4).

\[
net_h = \sum_i W_{xh_i} \cdot X_i - \theta_h \quad (1)
\]

\[
H_h = f(\text{net}_h) = 1/1+\exp(-\text{net}_h) \quad (2)
\]

\[
net_j = \sum_h W_{hy_h} \cdot H_h - \theta_y \quad (3)
\]

\[
Y_j = f(\text{net}_j) = 1/1+\exp(-\text{net}_j) \quad (4)
\]

where $i = 1\ldots N$, $N$ is number of inputs, $h = 1\ldots M$, $M$ is number of hidden layer nodes, $j = 1\ldots K$, $K$ is number of outputs.

Step 5. Compute the amount of difference $\delta$ by Eq. (5) and Eq. (6).

\[
\delta_j = Y_j(1-Y_j)(T_j-Y_j) \quad (5)
\]

\[
\delta_h = H_h(1-H_h)\sum_j W_{hy_h} \delta_j \quad (6)
\]

where $M$ is number of hidden layer nodes, $j = 1\ldots K$, $K$ is number of outputs.

Step 6. Compute the updating amount of connecting weight $\Delta W_{hy}$ (hidden layer to output layer), $\Delta W_{xh}$ (input layer to hidden layer) and the amount of bias weight $\Delta \theta_y$ (in output layer), $\Delta \theta_h$ (in hidden layer) by Eq. (7) to Eq. (10).

\[
\Delta W_{hy_h} = \eta \delta_j H_h + \alpha \cdot \Delta W_{hy_h} \quad (7)
\]

\[
\Delta \theta_y = -\eta \delta_j + \alpha \cdot \Delta \theta_y \quad (8)
\]

\[
\Delta W_{xh_i} = \eta \delta_h X_i + \alpha \cdot \Delta W_{xh_i} \quad (9)
\]

\[
\Delta \theta_h = -\eta \delta_h + \alpha \cdot \Delta \theta_h \quad (10)
\]

where $i = 1\ldots N$, $N$ is number of inputs, $h = 1\ldots M$, $M$ is number of hidden layer nodes, $j = 1\ldots K$, $K$ is number of outputs.

Step 7. Update $W_{xh}$, $W_{hy}$, $\theta_y$, and $\theta_h$ by Eq. (11) to Eq. (14).

\[
W_{hy_h} = W_{hy_h} + \Delta W_{hy_h} \quad (11)
\]

\[
\theta_y = \theta_y + \Delta \theta_y \quad (12)
\]

\[
W_{xh_i} = W_{xh_i} + \Delta W_{xh_i} \quad (13)
\]

\[
\theta_h = \theta_h + \Delta \theta_h \quad (14)
\]
where \( i = 1 \ldots N \), \( N \) is number of inputs, \( h = 1 \ldots M \), \( M \) is number of hidden layer nodes, \( j = 1 \ldots K \), \( K \) is number of outputs.

Step 8. Repeat steps 3 to 7 until the termination criteria is satisfied.

**Support vectors regression**

Support vector machines (SVM) method, which was proposed by Vapnik (Vapnik, 1995), is used to solve the pattern recognition problems at the beginning. Later Vapnik (Vapnik, 1999) promoted the SVM method to deal with the function fitting problems in 1999, which forms the support vector regression (SVR) method. It is assumed that a set of data \( G = \{(x_i, d_i)\}, i = 1 \ldots N, x_i \in \mathbb{R}^n \) is a n-dimensional input vector, The basic concept of SVR is as follows. A nonlinear mapping \( \varphi(.) : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is defined to map the input data (training data set) \( \{(x_i, y_i)\}_{i=1}^N \) into a so-called high dimensional space (which may have infinite dimensions), \( \mathbb{R}^m \). Then, in the high dimensional feature space, there theoretically exists a linear function, \( f \), to formulate the nonlinear relationship between input data and output data. Such a linear function, namely SVR function, is as Eq. (15),

\[
 f(x) = w^T \varphi(x) + b ,
\]

where \( f(x) \) denotes the forecasting values and the coefficient \( w (w \in \mathbb{R}^m) \) and \( b (b \in \mathbb{R}) \) are adjustable. SVR method aims at minimizing the empirical risk as Eq. (16):

\[
 R_{emp}(f) = \frac{1}{N} \sum_{i=1}^{N} \Theta_{\varepsilon}(y_i, w^T \varphi(x_i) + b) ,
\]

where \( \Theta_{\varepsilon}(y_i, w^T \varphi(x_i) + b) \) is the \( \varepsilon \)-insensitive loss function and defined as Eq. (17):

\[
 \Theta_{\varepsilon}(y_i, w^T \varphi(x_i) + b) = \begin{cases} 
 \left| w^T \varphi(x_i) + b - y_i \right| - \varepsilon, & \text{if } \left| w^T \varphi(x_i) + b - y_i \right| \geq \varepsilon \\
 0, & \text{otherwise.}
\end{cases}
\]

In addition, \( \Theta_{\varepsilon}(y_i, w^T \varphi(x_i) + b) \) is employed to find out an optimum hyper plane on the high dimensional feature space to maximize the distance separating the training data into two subsets. Thus, the SVR focuses on finding the optimum hyper plane and minimizing the training error between the training data and the \( \varepsilon \)-insensitive loss function. Then, the SVR minimizes the overall errors as follows:

\[
 \text{Min}_{w,b,\varepsilon^+, \varepsilon^-} R_\varepsilon(w, \varepsilon^+, \varepsilon^-) = \frac{1}{2} w^T w + C \sum_{i=1}^{N} (\varepsilon_i^+ + \varepsilon_i^-) \]

Subject to:

\[
 y_i - w^T \varphi(x_i) - b \leq \varepsilon + \varepsilon_i^+, \ i = 1,2,\ldots,N \\
 -y_i + w^T \varphi(x_i) + b \leq \varepsilon + \varepsilon_i^-, \ i = 1,2,\ldots,N \\
 \varepsilon_i^+ \geq 0, \ i = 1,2,\ldots,N \\
 \varepsilon_i^- \geq 0, \ i = 1,2,\ldots,N
\]

The first term of Eq. (18), employing the concept of maximizing the distance of
two separated training data, is used to regularize weight sizes, to penalize large weights, and to maintain regression function flatness. The second term penalizes training errors of \( f(x) \) and \( y \) by using the \( \varepsilon \)-insensitive loss function. \( C \) is a parameter to trade off these two terms. Training errors above \( \varepsilon \) are denoted as \( \xi_i^* \), whereas training errors below \( \varepsilon \) are denoted as \( \xi_i \).

After the quadratic optimization problem with inequality constraints is solved, the parameter vector \( w \) in Eq. (19) is obtained,

\[
w = \sum_{i=1}^{N} (\beta_i^* - \beta_i) \varphi(x_i),
\]

where \( \beta_i^*, \beta_i \) are obtained by solving a quadratic program and are the Lagrangian multipliers. Finally, the SVR function is obtained as Eq. (20) in the dual space:

\[
f(x) = \sum_{i=1}^{N} (\beta_i^* - \beta_i) k(x_i, x_j) + b,
\]

where \( k(x_i, x_j) \) is called the kernel function, and the value of the kernel equals the inner product of two vectors, \( x_i \) and \( x_j \) in the feature space \( \varphi(x_i) \) and \( \varphi(x_j) \), respectively; that is, \( K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \). Any function that meets Mercer’s condition can be used as the kernel function. There are several types of kernel function. The most used kernel functions are the Gaussian radial basis functions (RBF) with a width of \( \sigma \): \( K(x_i, x_j) = \exp(-0.5 \|x_i - x_j\|^2 / \sigma^2) \) and the polynomial kernel with an order of \( d \) and constants \( a_1 \) and \( a_2 \): \( K(x_i, x_j) = (a_1 x_i x_j + a_2)^d \). If the value of \( \sigma \) is very large, the RBF kernel approximates the use of a linear kernel (polynomial with an order of 1). However, the Gaussian RBF kernel is not only easier to implement, but also capable to nonlinearly map the training data into an infinite dimensional space, thus, it is suitable to deal with nonlinear relationship problems. Therefore, the Gaussian RBF kernel function is specified in this study. It is well known that the forecasting accuracy of an SVR model depends on a good setting of hyper parameters \( C, \varepsilon \), and the kernel parameters (\( \sigma \)). Thus, the determination of all three parameter selection is further an important issue.

**COMPUTATIONAL RESULTS**

**Descriptive statistic of samples**

From January 2003 to June 2007, a total of 117 patients were enrolled. There are 76 males and 41 females. **TABLE 1** shows the number of samples in gender and age.
TABLE 1  DESCRIPTIVE STATISTIC OF SAMPLES

<table>
<thead>
<tr>
<th>Features</th>
<th>N</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>41</td>
<td>35%</td>
</tr>
<tr>
<td>Male</td>
<td>76</td>
<td>65%</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>23</td>
<td>20%</td>
</tr>
<tr>
<td>30-39</td>
<td>28</td>
<td>24%</td>
</tr>
<tr>
<td>40-49</td>
<td>25</td>
<td>21%</td>
</tr>
<tr>
<td>50-59</td>
<td>27</td>
<td>23%</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>14</td>
<td>12%</td>
</tr>
</tbody>
</table>

**Construct forecasting model**

In this study, to construct the forecasting model by ANN and SVR. In building the SVR model, the LIBSVM package developed by Chang and Lin (Chang & Lin, 2011) is employed in this study. The ANN is coded in C++ programming language. The forecasting performance is evaluated using the following performance measures, the mean square error (MSE) and mean absolute difference (MAD). The definitions of the measures are as Eq. (21) and Eq. (22):

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (T_i - Y_i)^2
\]

and

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} |T_i - Y_i|
\]

where \(T_i\) is the actual value, \(Y_i\) is the forecasting value and \(n\) is the total number of data.

The parameters setting in both algorithms are important. Thus, there are the experiments results in TABLE 2 and TABLE 3. In the experiments, the criterion is to minimize MSE. The parameters of ANN are learning rate (\(\eta\)), and momentum (\(\alpha\)). The parameters of SVR are (\(C\)) and (\(\varepsilon\)). TABLE 2 shows the testing result of ANN, and the result of SVR are shown in TABLE 3. According to the results, the parameters of ANN (\(\alpha, \eta\)) are (0.99, 0.01), and (\(C, \varepsilon\)) are (4, -3) in SVR.

For comparing the forecasting models performance, ANN and SVR are tested in K-fold cross validation. In the experiment, the data are separated in 10 partitions. Then, there are 90% data for training, and 10% for testing. The results are shown in TABLE 4 and TABLE 5. Performance evaluation using MSE, training data MSE in SVR is 0.0116, and in testing data is 0.009. It indicates that SVR model has better performance than ANN model.
## TABLE 2  RESULT OF ANN PARAMETERS TESTING

<table>
<thead>
<tr>
<th>α</th>
<th>η</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>0.10</td>
<td>0.01289</td>
</tr>
<tr>
<td>0.95</td>
<td>0.10</td>
<td>0.01346</td>
</tr>
<tr>
<td>0.90</td>
<td>0.10</td>
<td>0.01352</td>
</tr>
<tr>
<td>0.99</td>
<td>0.05</td>
<td>0.01254</td>
</tr>
<tr>
<td>0.95</td>
<td>0.05</td>
<td>0.01346</td>
</tr>
<tr>
<td>0.90</td>
<td>0.05</td>
<td>0.01352</td>
</tr>
<tr>
<td>0.99</td>
<td>0.01</td>
<td>0.01212</td>
</tr>
<tr>
<td>0.95</td>
<td>0.01</td>
<td>0.01347</td>
</tr>
<tr>
<td>0.90</td>
<td>0.01</td>
<td>0.01352</td>
</tr>
</tbody>
</table>

## TABLE 3  RESULT OF SVR PARAMETERS TESTING

<table>
<thead>
<tr>
<th>C</th>
<th>ε</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-3</td>
<td><strong>0.01312</strong></td>
</tr>
<tr>
<td>6</td>
<td>-3</td>
<td>0.01352</td>
</tr>
<tr>
<td>9</td>
<td>-3</td>
<td>0.01365</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>0.01384</td>
</tr>
<tr>
<td>7</td>
<td>-2</td>
<td>0.01389</td>
</tr>
</tbody>
</table>

## TABLE 4  RESULT OF ANN IN 10-FOLD TESTING

<table>
<thead>
<tr>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAD</td>
</tr>
<tr>
<td>Fold 1</td>
<td>0.0856</td>
</tr>
<tr>
<td>Fold 2</td>
<td>0.0849</td>
</tr>
<tr>
<td>Fold 3</td>
<td>0.0812</td>
</tr>
<tr>
<td>Fold 4</td>
<td>0.0849</td>
</tr>
<tr>
<td>Fold 5</td>
<td>0.0781</td>
</tr>
<tr>
<td>Fold 6</td>
<td>0.0836</td>
</tr>
<tr>
<td>Fold 7</td>
<td>0.0887</td>
</tr>
<tr>
<td>Fold 8</td>
<td>0.0823</td>
</tr>
<tr>
<td>Fold 9</td>
<td>0.0849</td>
</tr>
<tr>
<td>Fold 10</td>
<td>0.0894</td>
</tr>
<tr>
<td><strong>avg.</strong></td>
<td>0.0844</td>
</tr>
<tr>
<td><strong>std.</strong></td>
<td>0.0033</td>
</tr>
</tbody>
</table>
TABLE 5  RESULT OF SVR IN 10-FOLD TESTING

<table>
<thead>
<tr>
<th></th>
<th>Training MAD</th>
<th>Training MSE</th>
<th>Testing MAD</th>
<th>Testing MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fold 1</td>
<td>0.0898</td>
<td>0.0135</td>
<td>0.0151</td>
<td>0.0003</td>
</tr>
<tr>
<td>Fold 2</td>
<td>0.0895</td>
<td>0.0135</td>
<td>0.0161</td>
<td>0.0011</td>
</tr>
<tr>
<td>Fold 3</td>
<td>0.0906</td>
<td>0.0136</td>
<td>0.0210</td>
<td>0.0009</td>
</tr>
<tr>
<td>Fold 4</td>
<td>0.0900</td>
<td>0.0136</td>
<td>0.0147</td>
<td>0.0004</td>
</tr>
<tr>
<td>Fold 5</td>
<td>0.0893</td>
<td>0.0135</td>
<td>0.0094</td>
<td>0.0002</td>
</tr>
<tr>
<td>Fold 6</td>
<td>0.0884</td>
<td>0.0128</td>
<td>0.0672</td>
<td>0.0139</td>
</tr>
<tr>
<td>Fold 7</td>
<td>0.0908</td>
<td>0.0136</td>
<td>0.0434</td>
<td>0.0057</td>
</tr>
<tr>
<td>Fold 8</td>
<td>0.0334</td>
<td>0.0023</td>
<td>0.1650</td>
<td>0.0669</td>
</tr>
<tr>
<td>Fold 9</td>
<td>0.0748</td>
<td>0.0071</td>
<td>0.0177</td>
<td>0.0005</td>
</tr>
<tr>
<td>Fold 10</td>
<td>0.0868</td>
<td>0.0126</td>
<td>0.0137</td>
<td>0.0003</td>
</tr>
<tr>
<td>avg.</td>
<td>0.0823</td>
<td>0.0116</td>
<td>0.0383</td>
<td>0.0090</td>
</tr>
<tr>
<td>std.</td>
<td>0.0178</td>
<td>0.0038</td>
<td>0.0479</td>
<td>0.0208</td>
</tr>
</tbody>
</table>

CONCLUSIONS

Echogram is a commonly used tool for the evaluation and rapid diagnosis of acute hepatitis in the ED and the parameters of the Child-Pugh classification are easily accessed in the ED. This study constructs the cost forecasting model by using echogram and Child-Pugh classification data. From the experimental result, the SVR gives better forecasting result than ANN. It is proved by the MSE and MAD. These findings indicate that by the cost forecasting model, the echogram and the Child-Pugh classification could be a practical tool for cost forecasting of acute hepatitis admitted to the ED. In future study, more data can be collected in order to improve the forecasting accuracy. In addition, it is possible to integrate other learning algorithms to develop the forecasting model.

REFERENCES


Chen, Chien-Hung, Yang, Pei-Ming, Huang, Guan-Tarn, Lee, Hsuan-Shu, Sung,


