MACHINE LEARNING OPTIMAL DECISION PATH FINDER ALGORITHM FOR CUSTOMIZING TREATMENT IN THE HEALTHCARE INDUSTRY

Srinivas Kolli 1, Parvathala Balakesavareddy2, M Madhuri3

1Assistant Professor, Department of Information Technology, Vnr Vignana Jyothi Institute Of Engineering & Technology, Vignana Jyothi Nagar, Pragathi Nagar, Nizampet (S.O), Hyderabad, Telangana-500090, India.
E-Mail:kollisreenivas@gmail.com

2Assistant Professor, Department of Information Technology, Vnr Vignana Jyothi Institute Of Engineering & Technology, Vignana Jyothi Nagar, Pragathi Nagar, Nizampet (S.O), Hyderabad, Telangana-500090, India.
E-Mail:balakesavareddyp@gmail.com

3Assistant Professor, Department of CSE, Maturi Venkata Subba Rao Engineering College, Nadergul, Hyderabad
Email:mmadhuri_cse@mvsrec.edu.in

ABSTRACT

Present days, machine learning acting a major role in healthcare. Machine learning helps to streamline hospital administrative procedures, map and manage infectious diseases, and customise medical treatments. It can have an effect on improving performance in hospitals and health systems, while reducing treatment costs. In this paper concentrated on path finder for optimal decision and diagnosis for triage, and optimal decision path finder Algorithm used to achieve the requirement. The purpose of this algorithm is to reduce the number of tests to cross the process of treatment. According to this report, based on the pre-test likelihood is most likely to validate a diagnosis calculated from patient knowledge, including symptoms and past tests. Diagnostic decision is taken according to the probabilities in the test.

Keywords: Healthcare, Machine Learning, OPDF Algorithm, Diagnostic Decision Making;

I. INTRODUCTION

Machine learning is a field utilising algorithm in improving output over time. This study uses modern algorithms to discover helpful data for three medical challenges. There are supervised and unsupervised models of machine learning. The decision function transfers a variable's value into an expected score or mark. It's incredibly easy to analyse the prediction output by referring to actual labels. One needs to split a dataset into two subsets in order to show the idea's success.

N-fold cross-validation, which uses the same data once, is a better validation process, with low bias, low variance and simple computation properties. Unsupervised learning strategies draw data from unlabelled data[7]. The decision functions of data mining are the results of supervised learning models. We may apply these functions to several problems. Many supervised learning approaches have a concern for financial fraud [4]. Concern for 'false positive' is not the only instance of mistake. The cost of error may be very low, but the cost of adding an attribute to improve classification may be high.

Due to technological advancements, healthcare can use Internet-connected devices [1]. The technology has today developed instruments for creating alternative employee models, IP capitalisation, intelligent healthcare, and cost reduction through a central role in the treatment, billing and medical records of patients [11]. Machine learning for health care is one such area. Google's profound learning approach is used to recognise mammogram cancer, while Stanford researchers use profound learning methods for classifying skin cancer. In different areas of healthcare, Machine Learning (ML) is already in operation.

www.turkjphysiotherrehabil.org
In this article, we are going to explore some of the influential applications of machine learning in healthcare and how they are going to transform healthcare industry in 2018 and beyond. This article analyses the use of emerging technologies for healthcare services and offers solutions to three different kinds of health problems. They are Deep Structure SVM (ML). They highlight the potential of machine learning for improving healthcare outcomes, reducing healthcare services, and personalised care[13].

II. BACKGROUND

Machine learning is helping us to create fully accurate diagnosis with vast precision. Eventually, there will be a strong use of ML in healthcare as it will enable doctors to figure out new treatment approaches, which were not previously available. The top applications of machine learning of health are described below, and shown in a figure1.

Figure 1: Machine learning application in health care domain

- Disease identification and diagnosis
- Discovery of drug and manufacturing
- Diagnosis of medical image
- Medicine personalization
- Behavioural modification through machine learning
- Machine learning based clinical research and recodes
- Smart health care
- Through machine learning
- Prediction management
- Machine learning based radio therapy
- Data collection and so on…

One of the essential application of ML in healthcare is recognition of diseases and illnesses. Cancer is a tough illness to catch in its initial stages. This is a perfect example of how integrating biological computing with cognitive computing will assist in diagnosing diseases very quickly. In the field of oncology, Berg is using A.I. to improve clinical treatments. P1vital is about creating a one-stop-shop for regular prescription service for the chronically ill. One main area of clinical machine learning is early-stage drug creation [2]. This includes research www.turkjphysiotherrehabil.org
and development of next generation sequencing and precise medicine that will find new avenues for multifactorial disease care. Unsupervised learning can still identify patterns without giving any predictions [13]. Microsoft is using machine learning technology in a variety of projects, including the development of AI-based treatment for cancer and the personalization of AML drug combinations.

Machine learning and deep learning are also made possible by computer vision technologies. This is an introduction app introduced by Microsoft for image recognition. You can see more data from medical imaging increase in their capacity as AI-driven diagnostics grows. With the aid of predictive analytics, personalised approaches are more efficient and more validated by empirical studies. Due to this restriction, doctors can only include a set of diagnoses or predict the patient's risk based on their symptomatic history and genetic information. Thanks to medical systems such as IBM Watson, there are all sorts of breakthroughs in medicine with the use of patient medical records and biometrics. More sophisticated health measurement systems and biosensors will become possible to use artificial intelligence in healthcare.

Behavioural therapy is an important method that can play a major role in preventive medicine. Somatix is a data analytics company which can detect and track our trends in order to make better decisions as regards our actions. It is an exhaustive process to monitor and comply with the health records. Machine learning is to make the process quicker and simpler. Methods using image classification and various machine learning technology-based OCR recognition programmes are becoming popular. Today, MIT is at the forefront of applying predictive modelling methods to health records in order to aid decisions in evaluating treatment options for patients suffering from chronic conditions.

There are many applications of machine learning in healthcare field. Clinical trials are lengthy and expensive, and they can often take years or even decades to complete. Predictive analytics focused on Machine Learning can assist in the systematic classification of potential candidates for clinical trials. Machine learning is also applicable in keeping trial data's records and individuals in the testing process. Crowdsourcing have become the buzz in the healthcare world as it is being facilitated by individuals. People can view medicine with less credibility if live health data is extracted. Google's testing kit allows iOS users to try to cure Asperger's and Parkinson's disease with a on facial recognition programme. Based on crowd sourcing knowledge, IBM has collaboration with Medtronic to decode, gather, and exchange real-time diabetes and insulin data[5]. There are also major applications in using these body sensors in the healthcare sector.

The application of machine learning in medical imaging is one of the most requested areas. This is the field where simulation model cannot be expressed easily. Since ML algorithms learn from the many examples available, diagnosis and selection becomes very easy. Machine learning is commonly used in both bioinformatics and medical image processing as one example of image classification. By partnering with DeepMind Medicine, London's UCLH National Health System hospital is using artificial intelligence to speed up cancer therapies.

AI technology and machine learning are currently used in global disease tracking and prediction. The scientists today have a large amount of expertise and descriptions of data collected through the space and the social media, which can expose aspects that may be ignored in the past. Artificial neural networks use math and science to keep an eye on anything from the Zika virus to cancer [3]. In third-world countries, several outbreaks may be identified due to lack of scientific equipment and schools. Especially ProMED-mail, an Internet-based reporting network [6] that monitors emerging diseases and provides real-time outbreak updates.

III. METHODS

In our work, we proposed an efficient decision-making algorithm. This would assist in reducing the expense of general research. Experience and deductive reasoning is the basis for diagnostic thinking. Medical workers either gather more details or treat in the face of proof of confusion. Clinicians can alter the likelihood of the disease depending on each test, according to Bayesian theory. The therapeutic threshold is essentially defined by how sure we are that the illness and disorder has been identified correctly [9]. The risk of illness has a different type depending on whether the treatment (or non-treatment) threshold has been reached. Specificities and sensitivities of the test are critical when determining test parameters. The steps of the algorithm are shown in Figure 2. It's basically like mathematical models.
It would be ideal if a person or group of people identified the test sequences as optimal ones to minimise cost while optimising decision-making confidence. One needs to know at what stage the outcome is most likely for a given patient with a given clinical scenario.

The OPDF algorithm is shown in Figure 2. Each next step is evaluated to maximise the efficacy of procedure.

![Figure 2: Steps for OPDF Algorithm](image)

Initially, we might have only the symptoms or some test results which are instantly available. A machine learning classifier is set up to predict whether or not a patient has a disease. LUCC then runs the forecast without attempting the procedure. The computer gives assessment if the prediction is capable (see Figure 4) and ends the process. Otherwise, the computer is looking for a test that will help it make its next prediction (in Step 2). The health professional may make a no treatment decision or an option to refer to another health provider. To tackle this problem, lazy support vector machine was used. It was a common way to get over hardships.

As Figure 4 shows, we use a simple weight algorithm (C_i), which indicates a question and three training cases. For upgrading example weights, the multiplier is $\beta$ (Figure 1). The initial weight for all training cases is 1, because we have no prior training case experience. In step 1, we don't pass the trust level and choose Test A, but only training case 3 has the same value that the question after the test result has been measured. Thus, his example weight (C_3) is multiplied with a $\mu$ (in this example, we are using $\mu = 2$). In step 2, training cases 2 and 3 have the same value as the query and are multiplied by 2, (C_2 and C_3). In step 3 only C_2 and C_3 are repeated by 2.
If ODPF works, all patient values are expected to be "missing", and we have no problem with a query with missing values. When I validate ODPF query, but I need to slightly adjust the case to validate reasonably. We avoid performing any tests if test values of the query is lacking. Some forms of examination will not yield adequate evidence for a particular diagnosis. It must be decided in this case who is afflicted.

IV. EVALUATION AND RESULTS

In figure 5, we will summarise the verification processes. We take care of each patient like our own family member (A). If research does not pass either threshold, further investigation is needed (from B to C). This can be done by having a limited ambiguity in C programming. A test will be conducted each time a sequence runs through A, B, C, A, if the prediction of the next number is above threshold. If prediction is accurate then the prediction is proven to be correct.

Table 1 demonstrates the accuracy and cost saving of four diagnosis techniques. The notation can show how many extra tests a patient is subjected to. It's 1 test completed, although it's 4 performances done. This is a situation where one make a prediction with no more medical advice. We cannot sum the results for all patients because of test variability. The score is a predictive outcome in a selection of a normal probability distribution.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ProbCont</td>
<td>Correct</td>
<td>239</td>
<td>145</td>
<td>78</td>
<td>108</td>
<td>126</td>
<td>72</td>
<td>46</td>
<td>51</td>
<td>45</td>
<td>218</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>266</td>
<td>181</td>
<td>82</td>
<td>121</td>
<td>153</td>
<td>81</td>
<td>51</td>
<td>63</td>
<td>56</td>
<td>306</td>
</tr>
</tbody>
</table>
The number of tests needed to determine an injured individual also differs. Some patients require few tests, while others need many. Therefore patients are more likely to choose the cheaper treatment, and the positive finding is summarised into ten strata (ite0 to ite9). Here, the prediction at stopping point is applied to the entire lines, and it shows that this is the right prediction on this situation. If it is right, then be consistent with it. The percentage of correct predictions in the map is 91%. In this statistical process, the overall accuracy is calculated from the percent accuracy of each stratum. In this way, we can concentrate on the total amount of tests used, and the average price saving of those used without costing to all patients. According to the tables he is a very high performing medical practitioner. Some patients do not need any examination, but some require almost all tests (ite9). The problem can be grouped into various categories for further study. It refers to the traditional notion that in the third case the ultimate risk turns to be less than zero. The patients would remain in a separate studio while being treated in hospital. Detection of early stage prostate cancer is more successful (ite0 to ite8)[8].

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th>0.8</th>
<th>0.812</th>
<th>0.929</th>
<th>0.909</th>
<th>0.837</th>
<th>0.867</th>
<th>0.8</th>
<th>0.807</th>
<th>0.864</th>
<th>0.717</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>minUC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>239</td>
<td>374</td>
<td>147</td>
<td>128</td>
<td>37</td>
<td>35</td>
<td>17</td>
<td>16</td>
<td>14</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>266</td>
<td>416</td>
<td>158</td>
<td>154</td>
<td>48</td>
<td>37</td>
<td>25</td>
<td>25</td>
<td>14</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td><strong>expUC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>236</td>
<td>186</td>
<td>216</td>
<td>97</td>
<td>77</td>
<td>57</td>
<td>37</td>
<td>26</td>
<td>23</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>266</td>
<td>221</td>
<td>233</td>
<td>112</td>
<td>96</td>
<td>63</td>
<td>51</td>
<td>33</td>
<td>24</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td><strong>IWexpUC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>237</td>
<td>187</td>
<td>211</td>
<td>101</td>
<td>171</td>
<td>53</td>
<td>38</td>
<td>31</td>
<td>18</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>264</td>
<td>221</td>
<td>224</td>
<td>115</td>
<td>101</td>
<td>63</td>
<td>52</td>
<td>36</td>
<td>22</td>
<td>256</td>
<td></td>
</tr>
</tbody>
</table>

TA=Total Accuracy; AT= Average Accuracy; CS= Cost Saving;

This table shows our algorithm’s ability to classify illness correctly. Some patients are not supposed to undergo any exams but some others are required to undergo all the tests (ite9). People with serious diseases are kept in the unit wards from 04 to 23. This sign that p value is a storey by itself. The boundary wall is to stay quiet for
patients all the time. There is a high effectiveness in the early detection of patients [8], (ite0 to ite8). For any patient who remains in this category, it must be made a final decision regarding his condition. Patients in the ite9 category are difficult to diagnose since they show no visible symptoms. Patients are categorised based on disorders based on the level of their diseases.

In the table, 1 is a plus, while 0 is a minus (alpha=5%). The baseline confidence level is 0.8407. All states have successfully completed the necessary tasks. The accuracy of these models is 0.852, far better than the baseline estimate. The features and capabilities of all the devices have been further enhanced. A variety of experiments saved 75 to 80 percent of costs. The speed-based approach is less complex than the cost-based strategy. Speed-based methods also cost less than the cost-based ones. This will depend on the dataset.

Figure 6 shows the accuracy with rising threshold. Standard refers to the form which was pre-processed with all the training data. LUCC gives greater importance to the findings. However, the recorded values for cumulative accuracies are 0.813, 0.852, and 0.847 for the cases of 0.75, 0.85, and 0.95 respectively. A high threshold can have no substantial effect on accuracy. Early detection can avoid later complications from cancer.

Figure 7 shows how frequently SAT will be needed across thresholds. This condition can be diagnosed early (i.e., before exhausting all tests). If the lower cutoff rate is in place, less patients will receive diagnoses. This incident shows that, as time passes, the threshold becomes higher. Integration of the precision and accuracy can lead to 100 percent accuracy. Diagnostic strata I to 8 is relatively precise, but strata 9 is very accurate. The chance of a better than null outcome from this analysis is not close to 0.05. The study shows that the average accuracies of total diagnoses improve with fewer patients who were granted diagnosis (ite0 to ite8). The threshold should be chosen to improve the accuracy. These irregularities were because of very small sample from strata ite4 to ite8.
In Figure 8, there is a negative relation between saving costs and the cost threshold. Both methods can provide more details than a random sample. The accuracy rates of minUC are much greater than those of random. For most points with comparable accuracies, costs for manual UC are substantially greater than costs for randomised guessing.

It is a trade-off between the decrease in cost and accuracy. It has been shown that LUCC provides greater economic benefit than marginal replacement. If cost saving is 50%, the accuracy for minUC is 81.9% and the accuracy for costExpUC is 78.3%. Speed-based approach result in high cost savings by using fewer and far more effective diagnostic tests. One thing you can do to cut back on costs is to pick cheap tests. These measures serve an excellent function of discovering or diagnosing the syndrome. This finding suggests that expensive tests at the start might be preferable sometimes. Sequence optimization would allow more patients to be diagnosed correctly which will cost less. Confident forecasts can be used to boost one's ability to predict outcomes. Figures 9 and 10 display the strength of application to each stratum. This study showed that sensitivity is appropriate for difficult threshold values. Bad diagnoses and treatment will result in fluctuation on the health graph. This shows the precision of the classifier's predictions. It enhances the response time of individuals who is already suspicious.
V. CONCLUSION

The earlier approach uses the prior findings for recommendation. This role can only be acquired by those countries which have big political population. The second important benefit is to help each other (interactive vs. non-interactive). In the alternative form, information should be revised periodically. Under this process, the doctors are told about the signs and symptoms when there are ample evidences. The doctor will need to carefully consider the option of treatment until he is unsure of which to use. The rationale of accuracy and saving costs are different in this article. In addition, by prescribing treatment, you can also advise the consistency of the action taken. This is due to the accuracy of the entire populace. Eliminating unnecessary assessments means less financial loss and less errors.

Medical professionals must take part in interacting with computers and machines to give assistance. Dynamic test is determined dynamically. For treatment without equipment, the second approach is employed. They are as follows: Second, the study of the data for all the patients in the heart disease database checked the three non-invasive cardiac examinations of their patients. The extent of medical applications of ODPF is still underestimated. A lot of data cannot be used in this analysis. While, the hepatitis dataset are omitted we didn't determine the utility of the test sequences for primary care providers. In order to address this issue, there needs to be fair limits on how and when to assess in clinical practise. Third, there are several false positive results present in the data.

For future work, to reduce data set is required to contain the categorical result. Health experiences are also more subtle. Medical experiments are created for the purpose of identifying diseases. Our main aim is to expand the ODPF so that it is applicable to more clinical settings. Some doctors perform several test processes concurrently rather than one at a time. Here we will compare two most promising studies. In the future, we intend to perform a range of blood and urine testing to pick the best among them.

REFERENCES


