Improving the Correctness of Medical Diagnostics Based on Machine Learning With Coloured Petri Nets

MUHAMMAD NAUMAN\textsuperscript{1}, NADEEM AKHTAR\textsuperscript{1}, OMAR H. ALHAZMI\textsuperscript{2}, MUSTAFA HAMEED\textsuperscript{1}, (Member, IEEE), HABIB ULLAH\textsuperscript{3}, AND NADIA KHAN\textsuperscript{1}

\textsuperscript{1}Faculty of Computing, The Islamia University of Bahawalpur, Bahawalpur, Punjab 63100, Pakistan
\textsuperscript{2}Department of Computer Science, Taibah University, Medina 30001, Saudi Arabia
\textsuperscript{3}Faculty of Management Sciences and Commerce, The Islamia University of Bahawalpur, Bahawalpur, Punjab 63100, Pakistan

Corresponding author: Muhammad Nauman (nauman@iub.edu.pk)

ABSTRACT
Advanced software and storage technologies have enabled medical facilities to record and store vast amounts of data about cancer patients. There is a strong demand for an accurate and interpretable method to perform cancer prognostic for effective treatment. Machine learning algorithms, undoubtedly, demonstrate a remarkable ability to recognize models and extract patterns from data to improve medical prognosis decision-making. But machine learning outcomes are prone to bias and inaccurate labelling. Therefore, to negate the impact of such errors in the prognostic decision-making process, the mechanism to correct such errors is in high demand. This article addresses this problem by proposing the use of Coloured Petri nets formalism to ensure the correctness of the machine learning based prognostic process. Use of formalism makes it possible to ensure that prognostic decisions are correct and understandable. Empirical results show that we have increased the accuracy of prognostic decisions by up to 90%. This research supports improved prognostic decision-making for the effective treatment and identification of cancer patients.

INDEX TERMS
Medical diagnosis, breast cancer prognosis, decision making, supervised machine learning, decision trees, formal verification, formal methods, coloured petri nets.

I. INTRODUCTION
Breast cancer has become a leading cause of death for women all over the world. An accurate and interpretive method is required to diagnose breast cancer patients for a well-executed treatment. Cancer incidence and mortality have increased rapidly in recent decades around the world, with 4,800 new cases are expected every day [1]. In women, Breast cancer is recognized as the most frequently diagnosed cancer and major contributor to cancer death across the globe. In 2018, approximately 2.1 million new cases of breast cancer were diagnosed worldwide, representing close to one in four cases of cancer in women [2]. However, the causes of breast cancer are not yet well understood by medical practitioners. Early detection of breast cancer can help treat life-threatening diseases more effectively [3]–[6].

Conventional methods of diagnosis of most existing diseases depend on human experience in determining cases that match confirmed data profiles. However, this earlier method of diagnosis is exposed to human error and inaccurate diagnosis. It is also long and laborious at the same time and causes unnecessary pain throughout the process. As an alternative, machine learning-based diagnosis continuously improves and is used to assist specialists in determining diagnostic decisions [7]–[9]. Today, Artificial intelligence (AI), Machine learning (ML) is the most dynamic field in the health industry [10]–[14]. AI and ML are in the field of research which processes and enhances technology systems to solve complex tasks by reducing the need for human intelligence [15]–[17]. Machine learning (ML) algorithms allow the extraction of decision rules from vast amounts of available data. In particular, ML algorithms show exceptional ability to recognize models and predict results from the data. Their greatest strength is the capacity to produce models that can facilitate decision-making [18].
Machine learning outcomes are vulnerable to labelling biases and errors [19], [20]. As a result, to mitigate these errors, mechanisms must be put in place to prove that the decisions rendered are correct. Formal verification refers to the process of verifying the correctness and accuracy of complex processes. Therefore, it is considered a valuable tool that guarantees the accuracy of machine learning outcomes [21], [22].

Coloured Petri Nets (CP-Nets) [23] are high-level Petri nets [24] used for modeling systems with discrete, simultaneous and dispersed events. They are used extensively to analyze and ensure the correctness properties of complex systems [25]. A systematic and comprehensive exploration of mathematical system models is used to demonstrate confidence in correct behavioral properties. These features make Petri nets a valuable tool for modeling and analyzing machine learning results.

In this paper, we introduce an approach to analyze breast cancer data using supervised machine learning for improved prognostic decision making. The proposed approach used state space analysis to ensure the correctness of analysis results. This research is different from existing literature as it proposed the use of state space analysis for medical prognostic, which has never before been explored.

This paper focused on the early detection of cancer to assist healthcare professionals. The cancer diagnosis process is assisted with prognostic rules that are derived through supervised machine learning. The correctness of these rules is verified using the State space analysis. The performance measures, including specificity, sensitivity, and accuracy are used to illustrate the effectiveness of the proposed approach. These research findings will help society improve the quality of life by quickly detecting the deadly disease.

The main contributions of this research work are:

1) We propose the use of supervised machine learning to improve the prognosis decision-making process.
2) We propose the use of state space analysis to ensure the correctness of the medical prognostic process.
3) We demonstrate the effectiveness of the proposed approach using breast cancer data sets.

The paper is arranged as follows: in the Section II we present the background. The Section III focuses on the literature review. Subsequently, materials and methods are presented in the Section IV, and finally the Section V presents the conclusion.

II. BACKGROUND
A. COLOURED PETRI NETS

Petri net is a bipartite directed graph \( Pn = (Pl, Tr, In, Ot, M_0) \) where \( Pl \) and \( Tr \) are non-empty finite disjoint sets holding the places and the transitions. An input function \( In \) denotes the directed arcs from places to transitions and defined as \( In: Pl \times Tr \rightarrow Pn \). An output function \( Ot \) denotes directed arcs from transition to places and is defined as \( Ot: Tr \times Pl \rightarrow Pn \). The initial marking is denoted by \( M_0 \) and defined as \( M_0: Pl \rightarrow Pn \).

Coloured Petri net (CP-Net) [23] is an enhanced version of Petri net introduced by C.A. Petri in 1962 [24]. In CP-Nets, colors are introduced to manage data types. It is defined as \( N = (Pl, Tr, F, \Sigma, C, g, f, M_0) \) where \( F \) is a finite set of arcs and \( \Sigma \) is a finite set of non-empty colour sets. Accordingly, \( C \) and \( g \) are colour and guard functions respectively. Here, \( f \) is an arc expression and \( M_0 \) is the initial marking respectively.

The dynamic behavior of systems can be described according to their state and how they are changed [26]. Therefore, for the purpose of simulating the dynamic behaviour of a system, a state or a marking in a Petri net is modified in accordance with the following transitional firing rules:
1) A transition \( t \in Tr \) is said to be enabled if each input place \( p \in Pl \) of \( t \) is marked with at least \( k \) tokens, where \( k \) is the weight of the arc from \( p \) to \( t \).
2) An enabled transition may or may not fire depending on whether or not the event actually takes place.
3) A firing of an enabled transition \( t \) removes \( k \) tokens from each input place \( p \) of \( t \), and adds \( k \) tokens to each output place \( p \) of \( t \), where \( k \) is the weight of the arc from \( t \) to \( p \).

The CP-Net offers an easy-to-use graphical interface and has a large community. It also has extensive documentation and support [27].

B. MODEL CHECKING

Model checking [28] is an automatic static verification method which is done by a tool called model checker. The model checker accepts two inputs. The first input is a system model which is specified based on a rigorous mathematical theory. The second input is a system property which is intended to be verified in the system model expressed by a formalism called temporal logic. However, in some model checkers, it is possible to use different formats instead of LTL or CTL formalism. Afterwards, the model checker constructs a state-transition system called system state space which comprises all reachable states of the system and verifies the given property in each reachable state of the state space.

C. STATE SPACE ANALYSIS

State space analysis is regarded as an excellent method of analysing Petri nets [29]. However, this method is confronted with the famous problem called the state space explosion. In the literature, several alternatives are suggested to address this issue [30]–[33]. A status space represents all possible execution paths in the system to formally check its correctness and prove the error absence in run time traces. Therefore, a state space is a directed graph representing each reachable marking and all arcs for each occurring binding element.

The CPN Tools [34] are used to create Petri nets and state space report to represent the dynamic behaviors of the system. The state-space report includes reachability property, boundedness property, home property, liveness property, and the CP-Net model’s fairness properties.
D. REACHABILITY ANALYSIS

Reachability is a fundamental basis for studying the dynamic properties of any system. The firing of an enabled transition will change the token distribution in a net according to the transition rule. A marking $M_0$, is said to be reachable from a marking $M$ if there exists a sequence of firings that transforms $M_0$ to $M$ [26].

E. MACHINE LEARNING ALGORITHMS

1) DECISION TREES

Decision trees are used for data classification. Their construction is simple and easily understood. Composed of nodes, branches and terminal nodes, these components are interconnected as a directed acyclic graph (DAG). One of the most common methods of classifying data is the decision tree classifiers. The community of researchers from various fields and backgrounds has considered the training decision tree for data analysis and decision making, including machine learning, pattern recognition and statistics [35].

In the decision tree construction process, attribute selection, such as splitting nodes and order, is essential. These essential factors maximize information gain and minimized uncertainty [36].

Entropy can be thought of as a measure of uncertainty, or randomness – the higher the entropy, the higher the uncertainty. Assume that a data set $D$, consists of a number of attributes $A_1, A_2, \ldots, A_n$ and a categorical classifying attribute $C$ possessing $k$ distinct classes. Further, assume that there are $T$ tuples in the data set, each is partitioned into one of the $k$ distinct classes of $C$. The measure of the maximum information contained within $T$ and can be calculated by Eq. 1

$$E(T) = - \sum_{i=1}^{k} \frac{|C_i|}{|T|} \log_2 \left( \frac{|C_i|}{|T|} \right)$$

where $k$ is the cardinality of $C$.

Accordingly; entropy is computed and compared for each attribute in the tuple. We can achieve this by dividing attributes $A_1, A_2, \ldots, A_n$ into sets of tuples $T_1, T_2, \ldots, T_m$. The weighted average of the information required to correctly identify a class of a tuple of $T$ attribute $A_j$ can be determined by Eq. 2

$$E(A_j, T) = - \sum_{i=1}^{m} \frac{|T_i|}{|T|} E(T_i)$$

where $T_i$ is a stochastic data portion from $T$ partitioned along attribute $A_j$’s distinct categories. Hence, the information gain of each non-classifying attribute in $D$ can then be calculated by Eq. 3

$$IG(A_j, T) = E(T) - E(A_j, T)$$

The information gain is calculated to select the attribute with the highest information gain as the initial splitting attribute. These calculations are recursively computed to construct a decision tree. This approach to decision tree induction works effectively. However, it results in biasing towards the attributes that have many categories [37]. In [38] Salzberg developed a method to normalize this bias with a normalization factor known as the gain ratio criterion.

The gain ratio criterion relies on an analysis of the information involved with each split attribute defined by Eq. 4

$$SI(A_j, T) = - \sum_{k=1}^{v} \frac{|T_k|}{|T|} \log_2 \left( \frac{|T_k|}{|T|} \right)$$

where $v$ is the cardinality of attribute $A_j$. The gain ratio is calculated by Eq. 5

$$GR(A_j) = \frac{IG(A_j, T)}{SI(A_j, T)}$$

The goal is to determine which classes differentiate themselves from other classes and simultaneously minimize the depth of the decision tree. This was facilitated by normalizing the information in attribute $A_j$ via Eq. (4). Consequently, less detailed splits are not unfairly penalized. However, if a given attribute has a predominance of one category, Split Information results in a high value, and Gain Ratio results in a low value. Consequently, anomalous findings could result.

2) ANN

An artificial neural network (ANN) [39] is the type of supervised machine learning algorithms, which is inspired by the complex, interconnected neural structure of the brain. Millions of transistors in a computer are interconnected in a similar way to neurons in neural network [40]. ANN offers high noise robustness, adaptive learning, and fault tolerance with high accuracy. For excellent performance, a large data set is necessary to form a neural network. However, training demands a model with a high capacity for learning with deep processing. Research has shown that deep processing also takes place in the human brain [41].

3) SVM

A Support Vector Machine (SVM) [42] is a supervised machine learning technique to identify a set of objects belonging to one of two classes by constructing a hyper-plane in a high dimensional space. The SVM algorithm is capable of creating a complex decision boundary between two classes with excellent accuracy. The SVM was originally introduced for dealing with binary classification issues. However, it was subsequently expanded to deal with multiple classification problems.

4) K-NN

K-nearest neighbour (k-NN) [43] is supervise machine learning algorithm. It applies to training datasets to explore the closest relatives to $k$ in future examples. The closest neighbour algorithm theory is used to define multiple drive samples adjacent to the new point and use them to predict the label.
5) **RF**
Random forests (RF) [44] are a collection of tree predictors. In RF, each tree depends on the values of a randomly sampled independent vector. The generalisation error of a tree classifier forest depends on the strength of each tree in the forest and the correlation between them. Each decision tree in forest gives a classification for each object, called “vote” for that class. The random forest assigns each object the class having a higher number of votes over all trees in the forest.

6) **NB**
Naive Bayes [45] is part of a family of simple probabilistic classifiers based on applying Bayes’ theorem with strong (naive) independence assumptions between the explanatory variables. Bayes’ theorem describes the probability of an event based on conditions relevant to the event. Therefore, given an object classified, characterized by several explanatory variables, Naive Bayes assigns to this object probabilities for each of the possible classes.

### III. LITERATURE REVIEW

#### A. MACHINE LEARNING IN MEDICAL DIAGNOSIS

In machine learning, a number of supervised machine learning methods, such as Linear Regression, Naive Bayes, K-Nearest Neighbors, Support-Vector Machines, Artificial Neural Network and Decision Tree have been widely used for breast cancer diagnosis due to their technical maturity, stability and cost saving. Wang et al. [46] proposed a rule extraction method to derive precise and interpretable classification decision rules from decision tree sets for breast cancer diagnostics. First, Random Forest was used to train a number of decision tree models to generate decision rules. And then a rule retrieval approach was designed to detach the decision rules from the driven tree. Finally, an enhanced multi-objective evolutionary algorithm (MOEA) was used to seek an optimal rule by the compromise between accuracy and intelligibility.

Ornela and Begonya [47] conducted a research study to identify variables that are more relevant to patient survival. The AURIA database was used, containing the electronic health records (EHR) of 20,006 patients with 178 features for breast cancer and 143 for prostate cancer in a particular region of Finland. Punitha et al. [48] proposed an algorithm for breast cancer diagnosis is based on the artificial immune system and bee colony. The focus was to facilitate efficient feature selection and to ensure the optimization process of the parameters of the artificial neural network.

Ronak et al. [49] proposed to use the decision tree for early detection of breast cancer using the Wisconsin breast cancer datasets. In [50] Somayeh proposed three methodologies with the use of deep belief networks (DBNs) to detect breast cancer. The empirical findings suggest that the proposed methods exhibit very high diagnostic performance in the classification of breast cancer. Dancey et al. [51] presented a new rule extraction method to extract a logistics model tree (LMT), based on decision trees, from a neuronal network formed for better explainable decision rules.

#### B. MEDICAL DIAGNOSIS RULE EXTRACTION

In the literature, rule extraction-based decision-making for medical prognosis has attracted much attention in the recent past [52]–[55]. Marian and Filip [56] presented an approach based on medical data using multi-objective evolutionary optimization algorithms (MOEOAs) to generate a set of solutions characterized by various levels of accuracy-interpretability trade-off. Gao et al. [57] proposed a methodology to assist physicians in making weaning decisions using the category-weighted naive Bayes network-based association rules. It incorporated category weighting based on association rules into a naive Bayes classifier and further research was taken into account the classifiers decision tree. The research described above focused on extracting rules directly from the data. While most recent trends have concentrated on extracting interpretable rules from black-box machine learning models.

On the other hand, when a rule-based system makes mistakes, a key issue is determining and correcting the rules that cause those mistakes. An error can be defined as the difference between the belief value generated by the system and the value reported by a presumed correct source of knowledge [58]. Our proposed approach uses supervised machine learning with CP-Nets state space analysis to ensure the correctness of the medical prognostic process.

#### C. FORMAL VERIFICATION OF ML ALGORITHMS

In the literature, correctness guarantees for ML outcomes are discussed in [22] and [59]. In [22] a process mining based approach for medical treatment process is formalised using CP-Nets. The use of CP-Nets-based simulations was explored to analyse the medical treatment process and suggest several alternative solutions. In [59] the authors proposed re-weighting the data points to mitigate the erroneous labeling. Our work used state-space analysis to ensure the correctness of the medical prognostic process. It also mitigates the effect of biasing and mislabelling errors penetrated in decision-making process.

In our previous work, we illustrated the use CPN simulations to formally verify the correctness of ML based decision making in educational domain [20]. This research demonstrates the use of state space analysis to ensure the correctness of the medical prognostic process.

### IV. MATERIALS AND METHODS

In this research work, a hybrid approach is proposed for the analysis of the breast cancer data set using supervised ML and formal verification as shown in Fig. 1. The first stage is the formation of decision tree models. In the second stage, the decision rules are extracted from the trees models, and, finally, CP-Nets state space analysis is applied to prove the correctness of the analysis. The extraction of the decision rule
is performed by following the path of the decision tree from the root node to the leaf nodes.

**A. DATA COLLECTION**

To illustrate our research methodology for breast cancer diagnosis, we conducted experiments on two data sets from the UCI Machine Learning Repository (http://archive.ics.uci.edu/ml). The first data set Breast Cancer Wisconsin Diagnostic (WDBC) is composed of 569 instances with 32 attributes from cell nuclei observed in breast images. The features are extracted from digital mammary mass images by Fine Needle Aspiration (FNA). The data set consists of 212 malignant cases and 357 benign cases. The malignant instances are labeled as "M" and benign as "B".

The second breast cancer data set (DSBC) used in this study was prepared by Zwitter and Soklic at the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia. The data set contains 2 decision classifications with 9 attributes for 286 patients [60].

The prognostic rules are exploited to monitor and evaluate the cancer diagnosis process in specific patients. These rules leverage the medical practitioners to analyze the patient diagnosis process effectively. It leverages improved decision-making with physicians to maximize a patient’s correct diagnosis to avoid challenging circumstances and save lives. The decision tree is induced to extract decision rules. The data set was split into two groups for training and testing. The training set has 400 records and the test set has 169 records for WDBC data set. The DSBC was divided into 200 records for training and 86 records for validation. The decision tree in Fig. 2 exhibits decision points for the WDBC data set in the tree structure. The prognostic rules extracted from the induced tree are illustrated in Table 3. The derivation of the rules was made via the path from the root to the leaf nodes. A decision rule’s usefulness generally comes down to two factors: support and accuracy. The number of instances that comply with rule conditions is called rule support. The ratio of correctly classifying instances in the data set for a particular rule is known as accuracy for rule. In order to achieve comprehensive prognostic rules, the threshold value for a rule to be taken into account in the decision-making process has been set at five.

In order to transform prognostics decision rule into a CP-Net model, decision choices in a rule needed to be identified. As a result, these decision choices break down the process into branches. In order to guarantee the correctness of prognostic rules for a patient to be classified as “M” or “B”, the CP-Nets for prognostic decision rules are constructed. In CP-Nets the prognostic rules are modeled as transition associated guards and patients are modeled as places. All CPN models are accessible on the link.

![Figure 1: The proposed research methodology to formally verify medical diagnosis decision rules using Coloured Petri-Nets.](https://github.com/mhmmd-nauman/DataSets/raw/master/bc.rar)
CP-Nets transformations for prognostic rules are executed using the algorithm described in our previous work [20]. The proposed hierarchical CP-Nets model is outlined in Fig. 3. It consists of three sub-models, namely `Diagnosis_M_Labeller`, `Diagnosis_B_Label`, and `Merge_Actual_Predicted_Labels`. The sub-modules interact...
TABLE 3. Decision table showing the prognosis rules extracted from the decision tree for the WDBC data set.

| No | Label | Extracted rules                                      | Rule nodes | Rule support | Rule accuracy |
|----|-------|-------------------------------------------------------|------------|--------------|---------------|
| R1 | B     | Worst Perimeter(WP) ≤ 105 AND Worst Concave Points(WCP) ≤ 0.1329 | 2          | 209          | 98.08         |
| R2 | B     | Worst Perimeter(WP) ≤ 105 AND Worst Concave Points(WCP) >0.1329 | 3          | 5            | 100           |
| R3 | M     | Worst Perimeter(WP) < 105 AND Worst Concave Points(WCP) >0.1329 | 3          | 11           | 90.90         |
|    |       | AND Worst Texture(WT) < 23.41                          |            |              |               |
| R4 | M     | Worst Perimeter(WP) ≤ 114.3 AND Mean Texture(MT) ≤ 19.67 AND Mean Radius(MR) ≤ 14.11 | 3          | 173          | 5.20          |
| R5 | B     | Worst Perimeter(WP) ≤ 114.3 AND Mean Texture(MT) ≤ 19.67 AND Mean Radius(MR) >14.11 | 3          | 19           | 100           |
| R6 | M     | Worst Perimeter(WP) ≥ 114.3 AND Mean Texture(MT) >19.67 | 2          | 68           | 38.23         |
| R7 | M     | Worst Perimeter(WP) >114.3                            | 1          | 140          | 98.57         |

FIGURE 3. Top level of the Coloured Petri nets abstract model for medical diagnostic verification process.

TABLE 4. Decision table showing the prognosis rules for the DSBC data set.

| No | Label | Extracted rules                                      | Rule nodes | Rule support | Rule accuracy |
|----|-------|-------------------------------------------------------|------------|--------------|---------------|
| R1 | B     | IN = 0-2                                             | 1          | 213          | 78.40         |
| R2 | M     | IN = 3-5                                             | 1          | 36           | 47.22         |
| R3 | M     | IN = 15-17                                           | 1          | 6            | 50            |
| R4 | M     | IN = 6-8 AND BQ = LEFT_LOW                            | 2          | 7            | 71.42         |
|    |       | AND BQ = LEFT_LOW                                    |            |              |               |
| R5 | M     | IN = 9-11                                            | 1          | 10           | 60            |

The correctness of sub-modules are proven by analysing the state space graph using the algorithm. The positions of the colour tokens are analyzed in the status space generated by the CPN tools [27]. Intuitively, the state space analysis algorithm can change, add, or delete decision rules.

Colour tokens created to represent patient details. Table 5 shows the definitions of the color sets used to model the CP-Net. Fig. 4 shows "M" decision rules CP-Net sub-module after simulating colour tokens.

C. STATE SPACE ANALYSIS ALGORITHM

The state space analysis algorithm 1 first defines the input and output parameters. The input parameter is a state space log generated by the CPN tools. The output parameter is true positive and false positive lists of the tokens in the model. This external loop performs an iteration across all tokens present in the CP-Net model states. In the body of the loop, the first IF Else statement checks whether a token reached on "M" state has "M" or "B" label. The statement guarantees that the actual and predicted labels are comparatively the same, and then adds the token to TP else in the FP lists.
FIGURE 4. The class label “M” CP-Nets after formal verification rules.

TABLE 5. The CP-Nets colour sets definitions.

| Color Set Definition | Description |
|----------------------|-------------|
| colset NO = INT;     | Specifies the serial number for patient record. |
| colset WP = REAL;    | Specifies the Worst Perimeter. |
| colset WCP = REAL;   | Specifies the Worst Concave Points. |
| colset WT = REAL;    | Specifies the Worst Texture. |
| colset MT = REAL;    | Specifies the Mean Texture. |
| colset MR = REAL;    | Specifies the Mean Radius. |
| colset L = string;   | Specifies the patient record classification label. |
| colset P = product NO*WP*WCP*WT*MT*MR*L; | Cartesian product specifies patient record attributes present in feature list. |
| colset P_LAB = product NO*WP*WCP*WT*MT*MR*L; | Cartesian product specifies patient record with class label attribute. |
| colset P_ACL_PRL = product NO*WP*WCP*WT*MT*MR*L; | Cartesian product specifies patient record with actual and predicted class label attributes. |
| colset ACTUAL_LAB = product NO*L; | Cartesian product specifies actual label for a patient record. |
| var sr : NO;         | Variable declarations for serial number. |
| var wcp : WCP;       | Variable declarations for Worst Concave Points values. |
| var wp : WP;         | Variable declarations for Worst Perimeter values. |
| var wt : WT;         | Variable declarations for Worst Texture values. |
| var mt : MT;         | Variable declarations for Mean Texture values. |
| var mr : MR;         | Variable declarations for Mean Radius values. |
| var label : L;       | Variable declarations for actual and predicted labels. |

**Algorithm 1 State Space Analysis Algorithm**

**Input:** State Space log of Model  
**Output:** TP, FP  

1. for Tokens in State Space log for a State in Model do  
   2. if State.M ← Token then  
      3. if Token.label == M then  
          4. TP ← Token  
      5. else  
          6. FP ← Token  
      7. end  
   8. else  
      9. if Token.label == B then  
          10. TP ← Token  
      11. else  
          12. FP ← Token  
      13. end  
   14. end  

D. CORRECTNESS VERIFICATION
CPN tools provide a state space tool to walk through Petri net models’ reachability graphs. The status of the patient in the CP-Net model is modelled as a transition place. Unique transitions within the model are a single decision rule. The formal verification is performed by simulating colour tokens and analysing state space for place Compare_Actual_Predicted_Labels which represent predicted and actual labels. Table 6 shows revised comprehensible decision rules after implementing Algorithm 1. In Fig. 4 & Fig. 5, the transition place represents the revised rules obtained after investigating state space analysis. Fig. 7 (b) illustrate the submodule Merge_Actual_Predicted_Labels.

E. REACHABILITY ANALYSIS
The state space report for WDBC data set is shown in Fig. 7(a). The CPN provides a state space tool to generate the analysis report automatically. The absence of cycles is proved by the equal number of arcs in the SCC graph and marking in the analysis report. The fairness property shows that there are no infinite occurrence sequences in model.
The properties described above have desired values. We discuss the properties for state space analysis as analysed by Hasiba et al. in [63]. The reachability analysis for formal verification of selected properties is presented in the section below.

1) MALIGNANT DECISION RULES VERIFICATION

The CP-Nets template for malignant cancer diagnosis decision rules is depicted in the Fig. 4. Once the model has its initial marking, we are interested in the verification of the following properties.

**Property 1:** Predictions From CP-Nets Are Always Correct for Cases Labelled As Malignant: This property ensures that the tokens satisfying malignant case scenarios are always reach the place Patient_Predicted_Label and label as “M”.

**Property 2:** The CP-Nets Model Put Predicted Labels on the Colour Token “M”: The marking of the place Label_M on transition Label_Patient_M ensures that colour tokens are always labelled as “M” on the place Patient_Predicted_Label.

2) BENIGN DECISION RULES VERIFICATION

**Property 1:** Predictions From CP-Nets Are Always Correct for Cases Labelled As Benign: This property ensures that
Property 2 The CP-Nets Model Put Predicted Labels on the Colour Token “B”: The marking of the place Label_B on transition Label_Patient_B proves that colour tokens are always labeled as “B” with the place Patient_Predicted_Label.

3) PREDICTIVE LABEL VERIFICATION

Property 1 The Predictive Label Colour Token Reaching Compare_Actual_Predicted_Labels Place Have Same Label: The property ensures that colour tokens after passing from transitions should reach Compare_Actual_Predicted_Labels place for label comparison. The correctness of the diagnostic and decision rules is ensured by analyzing the colour tokens at this stage.

The Compare_Actual_Predicted_Labels place marking proves that a token with actual and predicted labels always reaches the specified place.

F. PERFORMANCE METRICS

A decision tree classification model was used to label the cancer status of the patient either as Malignant (“M”) or Benign (“B”). To classify cases, the Confusion Matrix was calculated as shown in Table 8. In order to illustrate the effectiveness of proposed approach, we have chosen Accuracy (ACC), Sensitivity (TPR), Specificity (TNR) as evaluation parameters.

Here the total number of instances is represented as N and computed such as:

\[ N = TP + TN + FP + FN \]  \hspace{1cm} (6)

1) ACCURACY (ACC)

ACC specifies the accuracy of the classification model such that:

\[ ACC = \frac{(TP + TN)}{N} \]  \hspace{1cm} (7)

2) SENSITIVITY (TPR)

TPR specifies the correct classification rate of positive instances such as:

\[ TPR = TP/(TP + FN) \]  \hspace{1cm} (8)

3) SPECIFICITY (TNR)

TNR specifies the correct classification rate of negative instances such as:

\[ TNR = TN/(TN + FP) \]  \hspace{1cm} (9)

G. RESULTS

The effectiveness of the proposed approach is illustrated using two breast cancer data sets from University of Wisconsin, Clinical Sciences Center and University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia respectively. The WEKA tool is used to induce decision trees using J48 algorithm shown in Fig. 2. Table 3 shows the induced prognostic rules from decision trees by traversing a path from root to leaf nodes. The transformation of prognostic rules to CP-Nets is implemented using the transformation algorithm described in [20]. Fig. 3 illustrates the Hierarchical CP-Nets for the proposed approach. The colour tokens are constructed using the data from WDBC and DSBC. The comparative CP-Nets models after applying the state space analysis algorithm 1 are shown in Fig. 4 and Fig. 5.
TABLE 9. Comparisons with other supervised machine learning methods.

| Method | Breast Cancer Wisconsin Diagnostics (WDBC) | Breast Cancer data set (DSBC) |
|--------|-------------------------------------------|-------------------------------|
|        | Sensitivity | Specificity | Accuracy | Sensitivity | Specificity | Accuracy |
| ANN    | 97.25       | 97.31       | 97.14    | 25          | 82.25       | 66.27    |
| SVM    | 95.98       | 94.44       | 96.82    | 37.5        | 83.87       | 72.5     |
| K-NN   | 95.21       | 95.86       | 94.02    | 18.51       | 95.58       | 73.68    |
| RF     | 95.73       | 97.84       | 92.64    | 33.68       | 88          | 71.16    |
| NB     | 92.39       | 93.51       | 90.47    | 50          | 77.41       | 69.76    |
| Proposed approach | 91.71 | 85.20 | 88.46 | 43.13 | 86.57 | 75.5 |

TABLE 10. Comparisons with other rule extraction methods.

| Method | Breast Cancer Wisconsin Diagnostics (WDBC) | Breast Cancer data set (DSBC) |
|--------|-------------------------------------------|-------------------------------|
|        | Sensitivity | Specificity | Accuracy | Sensitivity | Specificity | Accuracy |
| Decision Tables | 27 | 93.75 | 89.59 | 96.91 | 11 | 73.68 | 77.01 | 75.35 |
| OneR   | 1 | 88.5 | 87.28 | 89.42 | 1 | 40.47 | 72.13 | 56.30 |
| PART   | 6 | 91.75 | 89.59 | 93.39 | 18 | 58.33 | 76.05 | 67.19 |
| RIPPER | 3 | 92.5 | 91.32 | 93.39 | 2 | 56 | 75.84 | 65.92 |
| Proposed approach | 6 | 91.71 | 85.20 | 88.46 | 5 | 44.11 | 78.84 | 65.11 |

FIGURE 8. Confusion Matrix before formal verification in (a) Confusion Matrix after formal verification in (b) Bar chart comparison for specificity, sensitivity and accuracy in (c) for WDBC data set.

TABLE 11. Prognosis rule correctness analysis for WDBC data set.

| Rule No | Before CP-Nets formalism | After CP-Nets formalism |
|---------|--------------------------|--------------------------|
|         | nodes | Rule accuracy | nodes | Rule accuracy |
| 1       | 2     | 98.08         | 2     | 100           |
| 2       | 3     | 100           | 3     | 83.72         |
| 3       | 3     | 90.90         | 3     | 84.84         |
| 4       | 3     | 5.20          | 3     | 92.30         |
| 5       | 3     | 100           | 2     | 90            |
| 6       | 2     | 38.23         | 1     | 91.66         |
| 7       | 1     | 98.57         | -     | -             |
| Average | 2.42 | 75.85         | 2.33  | 90.42         |

The critical malignant stage for a particular patient can be inferred from prognostic rules No.3, No.4, No.6 and No.7 from the Table 3. It should be noted that the accuracy of rule No.4 and No.6 is 5.20% and 38.23% respectively, which is evidence of data errors in the data set. In order to mitigate the effects of these errors and to facilitate the rules of prediction without error, this work proposed the use of the CP-Net formalism.

TABLE 12. Prognosis rule correctness analysis for DSBC data set.

| Rule No | Before CP-Nets formalism | After CP-Nets formalism |
|---------|--------------------------|--------------------------|
|         | nodes | Rule accuracy | nodes | Rule accuracy |
| 1       | 1     | 78.40         | 1     | 78.40         |
| 2       | 2     | 47.22         | 2     | 80            |
| 3       | 3     | 50            | 2     | 23.29         |
| 4       | 4     | 71.42         | 2     | 66.66         |
| 5       | 5     | 60            | 2     | 71.42         |
| Average | 1.2   | 61.41         | 1.71  | 63.95         |

Algorithm 1 is applied to obtain comprehensible prognostic rule analyses and verification. The formally verified prognostic rules, including No.3, No.4, No.6 and No.7 for critical malignant stage prediction from WDBC data set are shown in Table 6. It is evident from empirical results that prognostic rules accuracy has improved after applying Algorithm 1. It should be noted that accuracy for “M” classification is nearly 87%. The predictive accuracy for the “B” classification is 97%.
It is observed from Table 9 that the proposed approach has almost the accuracy as it was obtained using baseline classifiers including Artificial Neural Networks, Support vector Machine, K-Nearest Neighbours, Random Forest, and Naïve Bayes in terms of accuracy, sensitivity and specificity. The proposed approach has the primary advantage of applying the CP-Net formalism to prove the correctness of prognostics rules with almost the same accuracy from base line classifiers. From Table 10, we can see that the proposed approach shows almost same performance in terms of sensitivity, specificity, and accuracy against the other rule extraction methods.

The empirical results in Table 11 and Table 12 illustrate that we obtained correctness guaranteed prognostics rules from proposed approach. The prognostic prediction confusion matrix before and after applying the proposed approach is shown in Fig. 8 (a) & (b) respectively. It shows that correct classified instances has increased after apply CP-Nets formalism. Fig. 8 (c) & (d) shows the bar chart to visualise the comparison for specificity, sensitivity and accuracy WDBC and DSBC cancer data sets respectively. The chart analysis clearly indicates that specificity, sensitivity and accuracy have increased with the implementation of the proposed approach.

H. DISCUSSION
The use of machine learning to efficiently improve medical prognosis is still an emerging and active field of research in the recent past. This can improve the ability of the physician to make prognosis decisions. The proposed methodology has guaranteed the correctness of medical prognostic process with CP-Nets formalism. The proposed accuracy is very close to baseline machine learning algorithms with correct prognostic rules for improved decision-making. This research indicates that supervised machine learning is appropriate and effective for analyzing medical data.

In [46] a rule extraction method was proposed to derive accurate and interpretable classification rules from a decision tree ensemble for breast cancer diagnosis. Our proposed method not only extracts the design rules, but also demonstrates the correctness of those rules with CP-Nets formalism. We have applied the decision tree algorithm with CP-Net formalism to obtain decision rules the accuracy over 90%. In [53] Obregon et al. demonstrated a novel approach to combine and simplify the output of tree ensembles. The research was aimed at generating a set of rules for producing interpretable decisions. We focused on using formalism to generate the correct decision rules by providing formal assurances. Ahmad et al. [64] implemented ANN on Wisconsin Breast Cancer Dataset for early detection of breast cancer. This approach has less interpretable rules, as ANNs are generally seen as a black box, making them less useful to medical practitioners. However, our approach uses the decision tree with the CP-Nets formalism which creates understandable rules. This makes the approach very helpful to the physicians.

The results of machine learning algorithms are mostly verified with the help of the theorem proving formalism. We introduced an approach for verifying machine learning outcomes with CP-Nets state space analysis. To the best of our knowledge, this is the first attempt to apply the state space analysis of CP-Nets to the verification of prognostic rules based on machine learning. The empirical results show that the application of state space analysis not only guaranteed the accuracy of the medical prognosis process, but also maintained the accuracy of the results. As a result, it can be concluded that our proposed approach has improved the quality of the medical decision-making process.

V. CONCLUSION
Machine Learning has shown an exceptional ability to enhance the cancer prognosis process. At the same time, the results of machine learning are subject to bias and labelling errors in the data. In order to lessen the effects of these errors, we proposed the use of the CP-Nets formalism with the state space analysis. Formal verification is a technique used to ensure that complex software systems and process are correct. In the proposed approach, the decision tree algorithm is used to elicit prognostic rules based on breast cancer data sets. Prognostic rules are transformed into CP-Nets to prove correctness and reduce labeling errors. The CP-Nets state space is analysed to ensure the correctness of prognostic rules.

The proposed approach can be concluded that the accuracy of the derived prognostic rule can be ensured by the formalism of CP-Nets. Research results can serve to improve the decision-making process of physicians. The resulting prognosis rules can provide information on the patient’s diagnosis to physicians to quickly detect lethal disease and save lives.

In this work, only supervised machine learning was used with the CP-Nets formalism for the prognostic process. The proposed approach has limitations and needs further investigation:

1) Research has been carried out on two breast cancer data sets, for generalized results, The proposed methodology must be validated using multiple sets of medical diagnostic data.

2) At present, our proposed approach uses only prognosis rules based on the decision tree. Additional machine learning algorithms are required to improve the induction of prognostic rules.

In the future, we are considering further study of machine learning techniques for the induction of prognostic rules. In this work, only the CP-Nets formalism is implemented, other formal verification techniques such as the theorem proving needed to be studied to guarantee the correctness of the decisions based on the prognostic rules.
REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2019,” CA, Cancer J. Clinicians, vol. 69, no. 1, pp. 7–34, 2019.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA, Cancer J. Clin., vol. 68, no. 6, pp. 394–424, 2018.

[3] S. A. Alansari, M. M. Kamruzzaman, N. I. L. Sarker, M. Abruwa, Y. Alhwawi, N. Alshammarri, and M. H. Siddiqui, “Boosting breast cancer detection using convolutional neural network,” J. Healthcare Eng., vol. 2021, pp. 1–11, Apr. 2021.

[4] L. A. Torre, B. Trabert, C. E. DeSantis, K. D. Miller, G. Samimi, M. D. A. C. Jayatilake and G. U. Ganegoda, “Involvement of learning models to assess two important quality factors: Maintainability and reusability,” in Advances in Swarm Intelligence (Lecture Notes in Computer Science), vol. 6146, Y. Tan, Y. Shi, and K. C. Tan, Eds. Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-13498-2_46.

[5] R. L. Siegel, K. D. Miller, A. G. Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, “Colonorectal cancer statistics, 2020,” CA A Cancer J. Clinicians, vol. 70, no. 3, pp. 145–160, 2020.

[6] R. Lourdusamy and J. C. Anderson, “Clinical decision support systems and predictive analytics,” in Machine Learning With Health Care Perspective: Predictive Analytics, vol. 2, no. 3, pp. 284–296, Jul. 2018.

[7] L. Peng, W. Chen, W. Zhou, F. Li, J. Yang, and J. Zhang, “An immune-inspired semi-supervised algorithm for breast cancer diagnosis,” Comput. Methods Programs Biomed., vol. 134, pp. 259–265, Oct. 2016.

[8] S. M. D. A. C. Jayatilake and G. U. Ganegoda, “Involvement of machine learning tools in healthcare decision making,” J. Healthcare Eng., vol. 2021, pp. 1–20, Jan. 2021.

[9] R. R. Janghel, A. Shukla, R. Tiwari, and R. Kala, “Intelligent decision support system for breast cancer,” in Advances in Swarm Intelligence (Lecture Notes in Computer Science), vol. 6146, Y. Tan, Y. Shi, and K. C. Tan, Eds. Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-13498-2_46.

[10] R. L. Siegel, K. D. Miller, A. G. Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, “Colonorectal cancer statistics, 2020,” CA A Cancer J. Clinicians, vol. 70, no. 3, pp. 145–160, 2020.

[11] R. Lourdusamy and J. C. Anderson, “Clinical decision support systems and predictive analytics,” in Machine Learning With Health Care Perspective: Predictive Analytics, vol. 2, no. 3, pp. 284–296, Jul. 2018.

[12] L. A. Torre, B. Trabert, C. E. DeSantis, K. D. Miller, G. Samimi, M. D. A. C. Jayatilake and G. U. Ganegoda, “Involvement of learning models to assess two important quality factors: Maintainability and reusability,” in Advances in Swarm Intelligence (Lecture Notes in Computer Science), vol. 6146, Y. Tan, Y. Shi, and K. C. Tan, Eds. Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-13498-2_46.

[13] L. Peng, W. Chen, W. Zhou, F. Li, J. Yang, and J. Zhang, “An immune-inspired semi-supervised algorithm for breast cancer diagnosis,” Comput. Methods Programs Biomed., vol. 134, pp. 259–265, Oct. 2016.

[14] S. M. D. A. C. Jayatilake and G. U. Ganegoda, “Involvement of machine learning tools in healthcare decision making,” J. Healthcare Eng., vol. 2021, pp. 1–20, Jan. 2021.
[47] O. Bardhi and B. G. Zapirain, “Machine learning techniques applied to electronic healthcare records to predict cancer patient survivability,” Comput. Mater. Continua, vol. 68, no. 2, pp. 1595–1613, 2021.

[48] P. S., F. Al-Turjman, and T. Stephan, “An automated breast cancer diagnosis using feature selection and parameter optimization in ANN,” Comput. Elect. Eng., vol. 90, Mar. 2021, Art. no. 106958.

[49] R. Sumbaly, N. Vishnusri, and S. Jeyalatha, “Diagnosis of breast cancer using decision tree data mining technique,” Int. J. Comput. Appl., vol. 98, no. 10, pp. 16–24, Jul. 2014.

[50] S. Ronoud and S. Asadi, “An evolutionary deep belief network extreme learning-based for breast cancer diagnosis,” Soft Comput., vol. 23, no. 24, pp. 13139–13159, Dec. 2019.

[51] D. Dancey, Z. A. Bandar, and D. McLean, “Logistic model tree extraction from artificial neural networks,” IEEE Trans. Syst. Man, Cybern. B, Cybern., vol. 37, no. 4, pp. 794–802, Aug. 2007.

[52] M. Shahabi, H. Hassanpour, and H. Mashayekhi, “Rule extraction for fatty liver detection using neural networks,” Neural Comput. Appl., vol. 31, no. 4, pp. 979–989, Apr. 2019.

[53] J. Obregon, A. Kim, and J.-Y. Jung, “RuleCOSIE: Combination and simplification of production rules from boosted decision trees for imbalanced classification,” Expert Syst. Appl., vol. 126, pp. 64–82, Jul. 2019.

[54] J. Sultana, M. U. Rani, and M. A. H. Farqund, “Knowledge discovery from recommender systems using deep learning,” in Proc. Int. Conf. Smart Syst. Inventive Technol. (ICSSIT), Nov. 2019, pp. 1074–1078.

[55] J. R. Quinlan, “Generating production rules from decision trees,” in Proc. IJCAI, vol. 87, 1987, pp. 304–307.

[56] M. B. Gorzałczany and F. Rudziński, “Interpretable and accurate medical data classification—A multi-objective genetic-fuzzy optimization approach,” Expert Syst. Appl., vol. 71, pp. 26–39, Apr. 2017.

[57] Y. Gao, A. Xu, P. J.-H. Hu, and T.-H. Cheng, “Incorporating association rule networks in feature category-weighted naive Bayes model to support weaning decision making,” Decis. Support Syst., vol. 96, pp. 27–38, Apr. 2017.

[58] L.-M. Fu, “Recognition of semantically incorrect rules: A neural-network approach,” in Proc. 3rd Int. Conf. Ind. Eng. Appl. Artif. Intell. Expert Syst. (IEA/AIE), Jun. 1990, pp. 1013–1018.

[59] H. Jiang and O. Nachum, “Identifying and correcting label bias in machine learning,” in Proc. 23rd Int. Conf. Artif. Intell. Statist., vol. 108, S. Chiappa and R. Calandra, Eds. Aug. 2020, pp. 702–712. [Online]. Available: http://proceedings.mlr.press/v108/jiang20a.html

[60] R. S. Michalski, I. Mozetic, J. Hong, and N. Lavrač, “The multi-purpose incremental learning system aq15 and its testing application to three medical domains,” in Proc. AAAI, vol. 1986, 1986, pp. 1–041.

[61] J. R. Quinlan, C4. 5: Programs for Machine Learning. Amsterdam, The Netherlands: Elsevier, 2014.

[62] E. Frank, M. A. Hall, and I. H. Witten, “The WEKA workbench,” in Data Mining: Practical Machine Learning Tools and Techniques, 4th ed. San Mateo, CA, USA: Morgan Kaufmann, 2016.

[63] H. Ben Attia, L. Kahloul, S. Benhazrallah, and S. Bourrekakche, “Using hierarchical timed coloured Petri nets in the formal study of TRBAC security policies,” Int. J. Inf. Secur., vol. 19, no. 2, pp. 163–187, Apr. 2020.

[64] F. Ahmad, N. A. M. Isa, Z. Hassain, M. K. Osman, and S. N. Sulaiman, “A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer,” Pattern Anal. Appl., vol. 18, no. 4, pp. 861–870, 2015.

[65] MUHAMMAD NAUMAN received the Ph.D. degree in computer science from The Islamia University of Bahawalpur (IUB), Pakistan. He is currently working as a Programmer with The Islamia University of Bahawalpur (IUB). His research interests include formal approaches, machine learning, data mining, artificial neural networks, decision trees, predictive models, and big data analytics.

[66] NADIA KHAN received the Master of Science degree in computer science (MScS) from The Islamia University Bahawalpur, in 2015. Her research interests include multi-agent systems for education, educational data science, and formal methods for data science. She became a Student Member of IEEE Education Society, in 2020.

[67] HABIB ULLAH received the bachelor’s and master’s degrees in computer science, the master’s degree in business administration, and the M.Phil. degree in human resource management. His research interests include technology management, machine learning, big data analytics, tourism management, use of technology in sustaining tourism, project management, risk management, and tourism sustainability during pandemic with the use of technological advancement.

[68] OMAR H. ALHAZMI received the Ph.D. degree from Colorado State University, in 2007. He is currently a Full Professor with the Department of Computer Science, Taibah University, Medina, Saudi Arabia. He has also worked as a consultant to several organizations and the e-government national program. His research interests include software vulnerabilities, security, reliability of software, and the Internet of Things.

[69] MUSTAFA HAMEED (Member, IEEE) received the Master of Science degree in computer science (MSCS) from The Islamia University Bahawalpur, Pakistan, in 2015. His research interests include multi-agent systems for education, educational data science, and formal methods for data science. He became a Student Member of IEEE Education Society, in 2020.