Development of a Decision Tree-Based Online Preventive Control Tool for Small-Signal Stability Enhancement of Power Systems

Rizhong Kang, *Student Member, IEEE*, Jin Ma, Yan Xu and Meng Ke, *Member, IEEE,* Z.Y.Dong, *Fellow, IEEE*

Abstract—This paper has developed an improved decision tree (DT)-based Intelligent System (IS) to achieve preventive control to enhance system damping ratio (DR) in large-scale interconnected power systems. System pre-processing using participation factor tuning algorithm (PFTA) for learning algorithms (LAs) is designed for modelling stage in DIgSILENT. Implemented by DIgSILENT programming language (DPL), modal analysis (MA) are performed to analytically evaluate system oscillatory modes. RELIEF-F algorithm and DR sensitivity analysis are then applied for feature reduction to determine the critical generator ranking in terms of generation rescheduling (GR) importance for DR enhancement. The proposed Classification Tree Effectiveness Index (CTEI) enables the analytical comparison of DT performances integrated with market concern to shift an original operation point (OP) to an optimised OP via GR. Two case studies of both medium and large scale interconnected networks are given to validate the proposed DT preventive control scheme for cost-effective optimisation in small-signal security assessment (SSA) tool. Task automation (TA) in this paper is enabled for SSA tool by communicating between DIgSILENT and MATLAB which can facilitate further research of SSA.

Index Terms—**Decision tree (DT), online preventive control, DIgSILENT programming language (DPL), task automation (TA), Classification Tree Effectiveness Index (CTEI), small-signal security assessment (SSA).**

I. Introduction

CURRENT power systems are experiencing profound challenges of increased electricity demand and market deregulation, pushing power systems towards the stability boundary. Insufficient monitoring of violation of system security limit and failure to apply dynamic security assessment (DSA) to address control strategies on corresponding system instability can cause cascading system failure like US 2003 major blackout [1]. So the development of a reliable small-signal stability assessment (SSA) tool is aimed as one of the most important tasks in DSA tool design.

Regarded as the prerequisite stability requirement of the power system, small-signal stability can be mathematically described as large sets of differential algebraic equations (DAEs). Conventional approaches to assess and control power system dynamics to improve small-signal stability are disadvantaged due to complexity in real-time stability analysis heavily based on resolving DAE problems. However, with the advancements in computing hardware technologies, Intelligent Systems (ISs) that depend on knowledge discovery and pattern recognition rather than relying on the system dynamics are recently employed as sound alternatives to address DSA and its control problems. Recent works using IS in power system security assessment majorly dealt with transient and voltage stability, with examples in extreme learning machine (ELM) as fast assessment technique for DSA, artificial neural network (ANN) to screen and determine dynamic security contingencies, core vector machine (CVM) for voltage stability assessment, and decision tree (DT) for multi-fault transient stability analysis. Among different ISs, DT is advantaged for determining power system boundary for its simplicity, interpretability, learning speed, and bi-directionality in SSA tool design [2]-[5].

For small-signal stability controller in this paper, power system stabilizer (PSS) is acknowledged as an effective conventional tool broadly adopted in industry practices for damping ratio (DR) control for decades [6]. However, PSS was found to be insufficient and ineffective when dealing with inter-area oscillations due to time-varying operation conditions (OCs) in recent works [7]-[8]. Alternatively, a modal analysis (MA) for grid operation (MANGO) was proposed to mitigate real-time inter-area oscillations using modal sensitivity ranking [8]. But an inevitable challenge of MANGO stays with the difficulty to compute the system state matrix **A** in large scale interconnected power systems, thus it is feasible to resort to the accurate MA enabled by industry grade power system simulation software of DIgSILENT *PowerFactory* [9].

Evaluation process of SSA can be categorised into pre-disturbance and post-disturbance assessment, thus the main objective of achieving the optimised operation point (OP) with enhanced system DR from the original OP via generation rescheduling (GR) comprises the philosophy of preventive control schemes for SSA without altering control configurations. Corrective control strategies like load-shedding and islanding are taken to remove further violation of security constraints on the other hand.

From the background above, this paper focuses on developing an improved DT-based online preventive control

This research is supported by the School of Electrical and Information Engineering, the University of Sydney, Australia, under the Faculty Research Cluster Program and the Early Career Researcher Development Scheme.

tool for small-signal stability assessment (SSA) in large interconnected power systems. After presenting theoretic foundation in the first two sections, the system pre-processing stage with a proposed tuning algorithm named PFTA to ensure database quality from learning algorithms (LAs) is described in section III. Subsequently in section IV, database is generated via task automation (TA) enabled by DIgSILENT programming language (DPL) and MATLAB. Before the being formulated into DT construction, a distance-based feature reduction algorithm of RELIEF-F and a complementary DR sensitivity analysis are performed to determine the critical generators of system DR enhancement in section V. In section VI, the optimised CT rules are obtained and a comprehensive SSA tool development process is given. Finally in section VII, two case studies of both the validated New England 39-bus system and NEM network with incomplete control settings are presented to validate the DT-based online preventive control tool, where a sensitivity ranking is introduced for the latter network as a fast alternative for RELIEF-F when complete and validated power plant controllers are still unavailable in DIgSILENT modelling stage for complex systems.

II. Mathematical and Theoretic Foundation

Small-signal stability refers to the ability of the power system to maintain synchronism when being subjected to small protuberances, where the system response can be linearized for eigenvalue analysis [6]. According to the classification of power system stability defined by IEEE/CIGRE, small-signal stability in under the category of rotor angle stability [10]. In the context of this paper, small-signal instability is attributed to insufficient damping torque provided by synchronous generator.

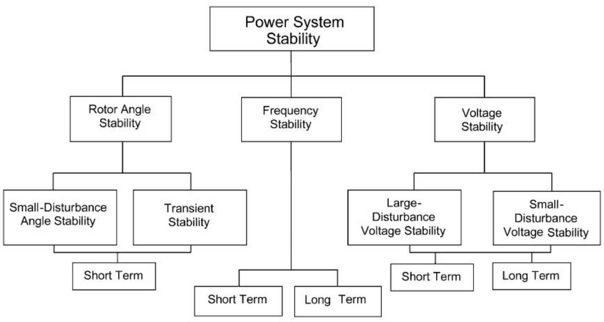

Fig. 1. Power System Stability Classification

A. System Linearization and Modal Analysis

The dynamic behaviour of the power system can be described by a set of DAEs in the following compact form:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, \mathbf{u}) \tag{1}$$

$$\mathbf{y} = f(\mathbf{x}, \mathbf{u}) \tag{2}$$

where x and u denote the state matrix and input matrix, and y denotes the output matrix.

Linearise (1) and (2) using Taylor's series expansion:

$$\frac{d\Delta\mathbf{x}}{dt} = f(\Delta\mathbf{x}, \Delta\mathbf{u}) = \mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{u} \tag{3}$$

$$\Delta\mathbf{y} = g(\Delta\mathbf{x}, \Delta\mathbf{u}) = \mathbf{C}\Delta\mathbf{x} + \mathbf{D}\Delta\mathbf{u} \tag{4}$$

where **A** is the state matrix, **B** is the control matrix, **C** is the output matrix and **D** is the feedthrough matrix, whose mathematical expressions are as follows:

$$\mathbf{A} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_r} \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} \frac{\partial g_1}{\partial u_1} & \cdots & \frac{\partial g_1}{\partial u_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial u_1} & \cdots & \frac{\partial g_m}{\partial u_r} \end{pmatrix} \tag{5}$$

Thus the eigenvalue λ of the state matrix **A** is determined by:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{6}$$

Here λ is in the form of:

$$\lambda = \sigma \pm j\omega \tag{7}$$

where the real part σ is the damping and ω is the oscillation frequency.

According to Lyapunov's First Method, if $\sigma > 0$, then the system is deemed unstable, otherwise the system is asymptotically stable if $\sigma < 0$.

Another important stability indicator of damping ratio (DR) of system mode can be derived as:

$$\zeta = \frac{-\sigma}{\sqrt{\sigma^2 + \omega^2}} \tag{8}$$

The participation factor identifying the relationship between state variables and modes can be written as:

$$\mathbf{P} = [\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}, \dots \mathbf{p_n}] \tag{9}$$

$$\mathbf{p}_i = \begin{bmatrix} \mathbf{p}_{1i} \\ \mathbf{p}_{3i} \\ \mathbf{p}_{3i} \\ \vdots \\ \mathbf{p}_{ni} \end{bmatrix} = \begin{bmatrix} \Psi_{i1}\Phi_{1i} \\ \Psi_{i2}\Phi_{2i} \\ \Psi_{i3}\Phi_{3i} \\ \vdots \\ \Psi_{in}\Phi_{ni} \end{bmatrix} \tag{10}$$

where Φ_{ki} the k^{th} entry of the right eigenvector Φ_i, and Ψ_{ik} the kth entry of the left eigenvector Ψ_i.

B. Decision Tree (DT) as an Intelligent System (IS)

1) Decision Tree (DT)) as a Supervised Learning Tool

DT was initially proposed by Breiman in 1980s. DT is notable for its supervised tree-structure predictive model comprised of both classification and regression modes based on their training targets. For analysis in SSA, an objective magnitude of the system DR is labelled into security categories, and thus a classification model is used to match the nature of security classification.

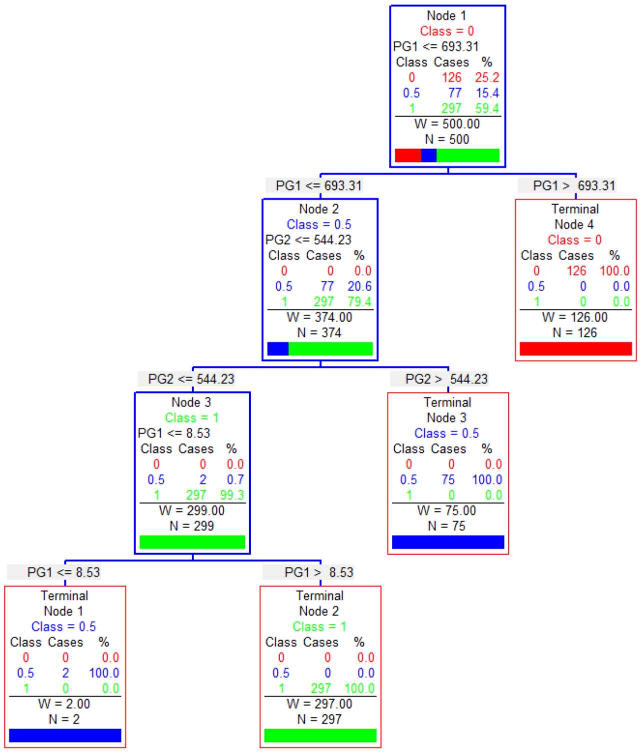

Fig.2. A DT Example of Dynamic Stability Assessment (DSA)

For DTs with optimal rules for small-signal stability enhancement, stability status in DT growing is categorical of either ‘stable’ (denoted by 1) or ‘unstable’ (denoted by 0) for two-class CT. A DT is developed from both a training set and a test set of data. The tree growing process is initiated by constructing a large enough tree and then by splitting the parent node into two new child nodes recursively. Each splitting rule calculated by GINI rule related to predictors are scored and ranked by the performance of separation among different classes of cases in parent node [11].

For the testing set, a common evaluation index is the misclassification cost expressed as below:

$$R^{TS} = \frac{1}{N^{TS}} \sum_{i,j} C(i|j) \times N_{ij}^{TS} \qquad (11)$$

where R^{TS} is the misclassification cost, N^{TS} the number of test cases, $C(i|j)$ the cost of misclassifying a class *j* case as a class *i* case, N_{ij}^{TS} the number of class *j* cases predicted as *i* class.

Another correctness index for classification can be expressed as:

$$CR_i^{TS} = \frac{N_{ii}^{TS}}{N_i^{TS}} \times 100\% \qquad (12)$$

where CR_i^{TS} is the correctness rate of classifying class *i* cases, N_i^{TS} the number of class *i* as test cases, N_{ii}^{TS} the number of correct classification of class *i*.

The statistical standard error for misclassification cost can be computed as:

$$\Delta R^{TS} = \sqrt{\left[\frac{R^{TS}(1-R^{TS})}{N^{TS}}\right]} \qquad (13)$$

where ΔR^{TS} denotes the error of R^{TS}.

2) Classification Tree Reliability and Classification Tree Effectiveness Index (CTEI)

Classification tree (CT) can have either two or multi classifiers depending on the labelling requirement. Hereby System Operator (SO) may define DR into various upper and lower borders rather than simply categorising DR status into secure or insecure.

In order to assess an analytical performance comparison between two-class CT and multi-class CT regarding its effectiveness index for small-signal stability enhancement, a classification tree effectiveness index (CTEI) is proposed in this paper to fulfil the requirement as (14):

$$CTEI_i = \begin{cases} [EW_i \cdot \frac{TC_{o,i}}{TC_{n,i}} + SW_i \cdot (\frac{DR_i - DR_0}{DR_{r,i} - DR_i})] \times 100\%, (\frac{DR_i}{DR_{r,i}} < 1) \\ [EW_i \cdot \frac{TC_{o,i}}{TC_{n,i}} + SW_i \cdot (\frac{\frac{DR_i}{DR_{r,i}} - \min(\frac{DR_i}{DR_{r,i}})}{\max(\frac{DR_i}{DR_{r,i}}) - \min(\frac{DR_i}{DR_{r,i}})})] \times 100\%, (\frac{DR_i}{DR_{r,i}} > 1) \end{cases} \qquad (14)$$

where i (i=1,2,…,n) denotes the identification number of CT, EW_i the economic weight, $TC_{o,i}$ the total optimised costs of the original base case, i.e. that of the original OP, TC_{ni} the total optimised costs of the new OP with enhanced DR, SW_i the security weight, DR_0 the damping ratio of initial OP, $DR_{r,i}$ the required damping ratio, and DR_i the actual DR of the ith new OP.

Note that of the above parameters, when DR_i / $DR_{r,i}$ <1, which is common due to LA training practicalities, the weight of an optimised DR is evaluated according to the distance from original DR in percentage. When DR_i / $DR_{r,i}$ >1, normalisation is performed on DR_i / $DR_{r,i}$ in order to scale it within [0, 100%]. SW_i and EW_i are defined by SO according to specific security and economic requirement of the operation condition. $CTEI_i$ is a percentage within [0,100%] for the ith assessed OP. The maximum DR_i / $DR_{r,i}$ will have the weight of 100% while the minimum DR_i / $DR_{r,i}$ is weighted 0.

Correctness index in (12) can be a viable assessing index, but with more obvious limitation since it only considers errors in tree splitting process. Nonetheless, it is still regarded as an important evaluation index in complex DT growing.

C. *Interface and Modelling in DIgSILENT*

DIgSILENT is a commercialised power system analysis software widely employed in European power industry. Modal analysis (MA) function in DIgSILENT computes the accurate eigenvalues of the integrated power system with power plant models and controllers. The main working principle behind is that DIgSILENT firstly calculates the natural oscillatory modes of the system, and then the system state matrix **A** is solved by deploying numerically iterative algorithms. Then the computation is followed by automatic triggering of oscillatory perturbations into the system and corresponding responses of eigenvalues are calculated [9].

Fig. 3. demonstrates the modelling approach using composite and common models for controller design in DIgSILENT. Such approach is applied in system pre-processing in section III.

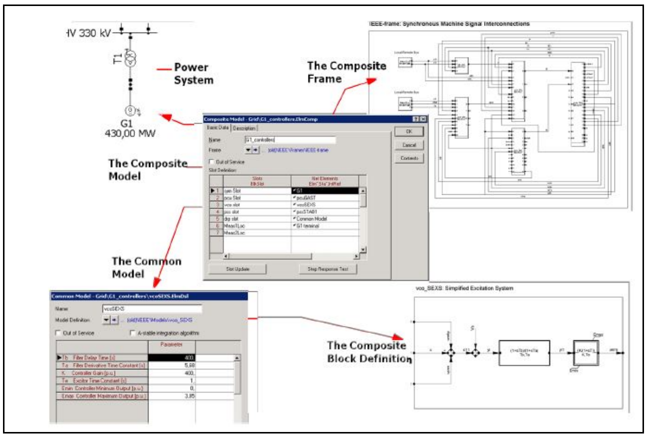

Fig. 3. Controller Modelling Approach in DIgSILENT

D. OPF based Generation Rescheduling

Optimal Power Flow (OPF) determines the minimising objective function of total cost when the power system is subject to security constraints. According to [12] and [13], OPF with a fuel cost minimisation quadratic object function under security constraints can be expressed as:

$$\text{Objective function} = \min(f(x,u)) = \sum_{i\in\Omega} a_i P_{gi}^2 + b_i P_{gi} + c_i \tag{15}$$

$$\text{Subject to } g(x,u) = 0 \text{ and } h(x,u) \le 0 \tag{16}$$

where $f(x,u)$ denotes the fuel cost function of each generator, $g(x,u)=0$ is the power flow equality constraints and $h(x,u)\le 0$ is the inequality constraints represented by operational constraints including security constraints. x and u are respectively the dependant variables and control variables. Ω denotes the set of OPs.

In case studies, MATPOWER package is used to the compute the minimised total fuel cost of the system given the splitting rules from the DT growing to be inputted as the generator inequality matrix [14]. Hence, an original OP can be shifted towards a new optimised OP with minimum fuel cost subject to security constraints.

III. SYSTEM PRE-PROCESSING

It is common in DIgSILENT modelling stage that a complex network may have various problems and challenges to be yet adopted as a fully validated test system. Due to the lack of accompanying load compensation and undefined power plant controller in some cases, the prerequisite control settings are usually unsatisfactory to realise the optimal preventive control rules by GR. Thus system pre-processing is crucial for training by DT learning algorithms.

A. Applying Load Compensation and Controller Model

Load compensation is crucial for efficient and reliable power system operations. Since the voltage levels of different equipment at different terminals rely on the magnitude of reactive power and will affect the system stability, thus it is imperative to introduce a cost-effective way to mitigate reactive power imbalance. Among the most popular compensation methods, shunt reactors, shunt capacitors, and Static Var System (SVCs) are generally applied in order to stabilise the system.

B. Participation Factor Tuning Algorithm (PFTA)

PFTA proposed hereby aims at fast tuning method to improve system DR when the complex system is subject to incomplete control settings yielding highly unstable DRs that affect feature quality during the SSA development process in DIgSILENT. PFTA is based on the philosophy of a quick adjustment of active power output (APO) of each generation unit in order to obtain a small-signal stable OP according to its participation factor ranking of rotation speed as state variables, since rotation speed is positively related to the generator APO. PFTA generally deals with the closest mode to the origin in right half plane (RHP) visualised on the eigenvalue plot in DIgSILENT. According to the definition of participation factor (PF), PF with negative magnitude implies an under-generated APO while a positive PF represents excessive APO.

TABLE I
PFTA Algorithm Procedures

Step 1. Perform Modal Analysis (MA) on an initially unstable OP_i *(i=1, 2, 3….N)* and acquire all system modes with their damping ratios (DRs)
Step 2. Search and identify the most sensitive modes (closest mode to the origin from RHP). **Step 3.** Obtain participation factors (PFs) of rotor speed (active power) of each contributing generators and rank their PFs in descending order. **Step 4.** Adjust and modify the impacting active power outputs of the unstable modes according to their contributing PF magnitudes to move those RHP modes towards the origin, i.e. $PG_j = PG_j + \varepsilon$, *(j=1,2,3….M)* where *M* is the number of contributing generators and ε is a user-defined step size according to the scale of power system and PF value of generators of interests. **Step 5.** Update the new set of OP_{i+1}, where *i=1, 2, 3…, N*. If system minimum of this OP satisfies *DR>0*, stop and exit this loop. Otherwise, go back the step one **Step 6.** Return the active power output PG_j, active load power $Pload_k$ (*k=1, 2, 3…, L)* and DR of the stable OP_i as the base case for further database generation in MATLAB and DIgSILENT

IV. DATABASE GENERATION

Database comprised of input features, i.e. active power output (APO) of each generator, and output features, i.e. system minimum DR can be generated via varying load consumption and APO of PV buses to a certain intervals. It is notable that there is trade-off between database size and training quality regarding the tolerance of error-free load flow calculation and MA in DIgSILENT.

DIgSILENT programming language (DPL) interface is designed for task automation (TA) in both *PowerFactory* and MATLAB. It is notable that DPL syntax can be categorized into the following parts [9]:

- Variables definitions
- Parameter Assignments and Mathematical Expressions
- Control loops
- Method and Object (External/Internal) Calls

Hereby, DPL code aims to fulfil the functionalities for data generation comprised of different code blocks as skeleton below:

A. Define Variable, Object and Set Types, File and Their Names

- Define generator active power output as double (PG_i, i=1,2,3,…n).
- Define load active power consumption as double (PL_j, j=1,2,3,…m).
- Define iterator identifier as integer i, j, …etc. to represent database size, number of network elements etc..
- Define objects of synchronous generator and load.
- Define sets of synchronous generator and load.
- Define file types, modes and input/output directories.

B. Assign and Extract Parameters

- Extract text file generated by MATLAB containing active power of synchronous generators and PL_j, (j=1,2,3,…m) loads from different OPs generated to be substituted.
- Load OP data and assignment them to PG_i, (i=1,2,3,…n) and PL_j, (j=1,2,3,…m).

C. Construct Control Loop to Achieve TA

- Define of iterator identifier to determine the database size.
- Search and pinpoint synchronous generators and loads in the system.
- Perform initial condition calculation.
- Perform Modal Analysis command.
- Return error information and reset calculation in case of failure in power flow/initial condition convergence.

D. Export and Import Result Files

- Export and write out system eigenvalues computed from Modal Analysis as text files named after iterator identifier.
- Import and read the next line of a new OP data and repeat the loop again.

E. Control Loop to Achieve Inter-software Communication

- Export initial OP data containing PG_i *(i=1,2,3,…n)* and PL_j *(j=1,2,3,…m)* from DIgSILENT to MATLAB.
- Generate new OPs that vary both PG_i *(i=1,2,3,…n)* and PL_j *(j=1,2,3,…m)* without interrupting the power flow balance, e.g. concurrently increase PG_i *(i=1,2,3,…n)* and PL_j *(j=1,2,3,…m)* in the manner that the total power generation and total load consumption remains balanced, respectively in separate files based on initial OP data and export them to DIgSILENT.
- Perform DPL scripts described in part C commenced with extraction of data in new OPs.
- Import text files containing eigenvalues of each OP into MATLAB
- Determine the damping ratios (DRs) of all system modes of each OP in MATLAB and find their minimum DR to be output features of corresponding inputs of each OP.

V. CRITICAL GENERATOR DETERMINATION

A. RELIEF-F Algorithm

The working principle of RELIEF algorithm is based on iteratively updating the relevant weight of each feature separately. In this paper, the generators' features of active power are considered and their higher RELIEF yielded weights are to be inputted for DT construction. The algorithm can be expressed as [15]:

$$W[A] = W[A] - diff(A, R_i, H)/m + diff(A, R_i, M)/m, i \in m \tag{19}$$

where A above denotes a feature, R_i is the instance randomly sampled in this iteration. H is the nearest instance from R_i in the same class as the nearest hit, and *M* represents the nearest instance from a different class of R_i, i.e. a nearest miss. *m* is the amount of sampled instances ensuring that all of them fall within the interval of [-1, 1]. Function *diff (A,X,Y)* computes the difference between the *A* values of *X* and *Y*.

However, RELIEF is insufficient when dealing with multi-class problem which will be examined later. As a feasible solution, RELIEF-F is devised with extended class component for the weight updating algorithm [16].

$$\begin{aligned} W[A] = W[A] - \sum\nolimits_{j=1}^{k} diff(A, R_i, H_j)/m\cdot k + \\ \sum_{C\neq class(R_i)}^{k} \frac{P(C)}{1-P(class(R_i))}\cdot diff(A, R_i, M_j(C))/m\cdot k \end{aligned} \tag{20}$$

where *k* is an user-defined parameter with *C* as the class label and *P(X)* as the prior probability of a class.

B. Damping Ratio (DR) Sensitivity Analysis

While RELIEF-F depends heavily on distance among instances, it may fail to distinguish system features of large-scale power systems with incomplete control settings as shown in NEM network in case studies. Thus it is essential to introduce an alternatively fast feature reduction scheme regarding with system DR enhancement during system modelling stage without considering the rigid and time-consuming design of controllers. According to the definition of sensitivity, the DR sensitivity is calculated as:

$$S_i\Big|_{\Delta P_i, \Delta\zeta_i \to 0} = \frac{P}{\zeta}\cdot\frac{\Delta\zeta_i}{\Delta P_i} \tag{21}$$

where ΔP_i denotes the difference of APO and $\Delta\zeta_i$ the difference of DR, P and ζ are respectively the base value of active power output and DR. Hereby $P = 1MW$ and $\zeta = 0.01\%$.

From the DR sensitivity, an APO feature ranking of each generation unit is enabled through magnitude comparison.

VI. DERIVING OPTIMAL RULES BY DECISION TREES

After DT growing from specification of input and output features, DT-based preventive control rules obtained can be summarised as follows:

$$R_{pc} = \{N \in C_S : n_i < \tau_i \cup n_i > \vartheta_i, i \in \Theta\} \tag{22}$$

where R_{pc} denotes the splitting rules for preventive control, N the terminal nodes with class labels, C_S the class of 'secure', n_i the node i, τ_i the upper boundary and lower boundary ϑ_i, and Θ the critical generator set.

When incorporated into online SSA, (17) can be revised as below.

$$R_{pc} = \{N \in C_S : n_i < k_{ub,i}\tau_i \cup n_i > k_{lb,i}\vartheta_i, i \in \Theta\} \tag{23}$$

$$k_{ub,i} = k_{ub,i} + \varepsilon \tag{24}$$

$$k_{lb,i} = k_{lb,i} - \eta \tag{25}$$

where $k_{lb,i}$ and $k_{ub,i}$ denote respectively the gains of lower bound and upper bound, and ε and η are tuning parameters defined by the SO.

The computation process can be performed off-line by DT growing before being synchronised with the online application for SSA, whose process can be shown as Fig.4. as below:

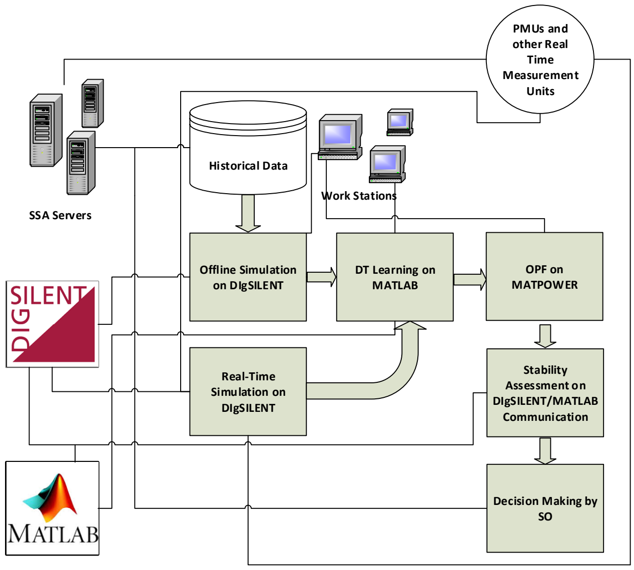

Fig. 4. SSA Process Flow Chart

It is notable that offline training shall be accomplished before online synchronisation of SSA.

VII. CASE STUDIES

A. New England 39-Bus System

The proposed preventive control scheme is firstly tested and validated on the IEEE 39-Bus New England System, as shown in Fig.5 below. Regarded as a medium scale power system brought up firstly in 1970s, it supplies six different states in regional New England area of Massachusetts, New Hampshire, Main, Connecticut, Vermont. This test system is a simplified equivalent model of more than 350 generators and over 8000 miles of transmission line comprising 10 generators ranging from nuclear, hydro to coal power plant and 21 loads, with altogether 39 buses in total [17]. It is acknowledged as a widely employed benchmark test system for power system stability analysis and serves as a good simulated approximation of realistic systems for research purpose. All system configuration data are available from [18], and the fuel cost parameters are obtained from [19].

Here for the convenience of symbol representation, the generator identification is changed from G30 to G39 to G01 to G10 as tabulated in the following table.

So it is imperative that G01 and G10 should exchange data with each other in further OPF calculations when inputting generator data matrix for consistency with [20]. Note that controller models of New England System are constructed on DIgSILENT composite and common model units in [20] as predefined control settings to be remained unchanged. In [20], PSS and AVR are constructed, while governors (GOVs) employ IEEE Type G1 as steam turbine on G02-G09 and IEEE Type G3 as hydro turbine is installed on G10.

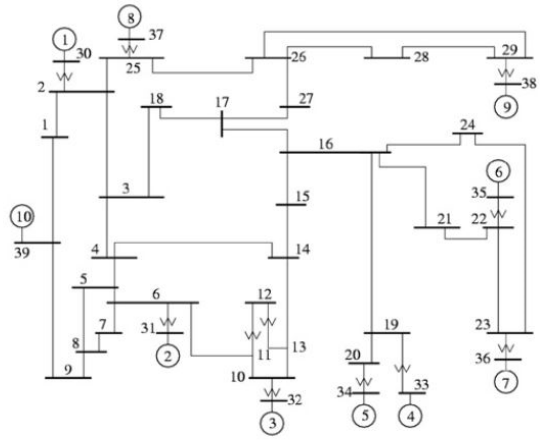

Fig. 5. New England 39-Bus System

TABLE II
Generator Conversion Table of New England-39 Bus System in DIgSILENT

Generator ID. in IEEE 39 Bus System	*Generator ID. in DIgSILENT*
G30	G10
G31	G02
G32	G03
G33	G04
G34	G05
G35	G06
G36	G07
G37	G08
G38	G09
G39	G01

1) System Pre-processing and Database Generation

As the first step, it is necessary to tune the constructed 39-bus system in DIgSILENT to closely small-signal stable by applying pre-processing stage described in Part IV. This procedure is achieved by applying the same power plant model in Composite Model Frames on G02 to G09 in the DIgSILENT-based system [20]. Herein, IEEE Type 1 Excitation System with acronym of AVR_IEEET1, IEEE Type 1 Speed-Governing Model with acronym of GOV_IEEEG1 and Speed Sensitive Stabilizing Model with acronym of PSS_CONV are adopted as the backbone elements of the power plant model for maintaining oscillatory stability as the prerequisite of investigating the effects of generation rescheduling on small-signal stability.

TABLE III
Initial OP of IEEE 39-Bus System

Generator (DIgSILENT ID.)	G30 (G10)	G31 (G02)	G32 (G03)	G33 (G04)	G34 (G05)	G35 (G06)	G36 (G07)	G37 (G08)	G38 (G09)	G39 (G01)
Active Power Output (MW)	250.00	894.60	650.00	632.00	580.00	650.00	560.00	540.00	830.00	1000.00

Otherwise, most of the generated cases would be small-signal unstable and would result in a uniform damping ratio of -1.00, thus severely affect the quality of DT training as input features. Here, altogether 500 valid OPs are generated.

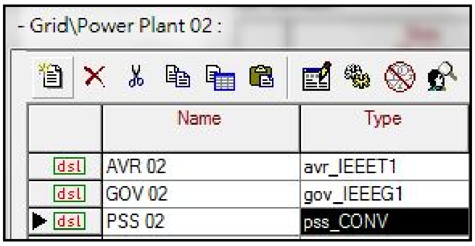

Fig. 6. Power Plant Controller Model in New England 39-Bus System

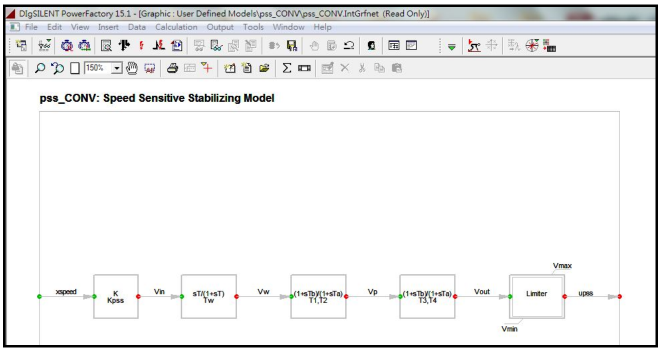

Fig. 7. PSS_CONV Block Diagram in DIgSILENT

2) *Identification of Critical Generators*

Followed by this step, RELIEF-F algorithm is used to assess the APO importance weight ranking of the operation 10 generators in the system ranging from G01 to G10. The resulting weight ranking is demonstrated below in Fig. 8. It is notable that the *k* constant of RELIEF in this case is chosen as 100 to achieve the best discrimination among instances via trail-and-error approach.

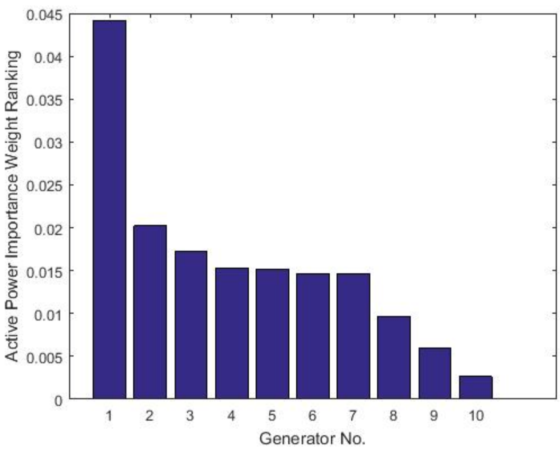

Fig. 8. RELIEF-F Weight Ranking of Generator Active Output in IEEE 39-Bus System

3) *Optimised Rule from Two-Class CT*

If the DR threshold is defined as 6.00%, i.e. only DRs greater than 6.00% will be regarded as the class of 'Secure (1)' and the remaining DRs will be labelled as 'Insecure (0)' in a two-class CT.

From the critical generator features obtained above, an optimised CT is grown by employing CART software with Gini as the splitting method and 10-folder cross-validation is chosen for CT growing.

Hence, the CT can be obtained as Fig. 9.

4) *Optimised Rule from Multi-Class CT*

If a multi-class CT is selected to develop the optimal tree, the multi-level of security status can be labelled as three states of Very Secure (1), Medium Secure (0.5), Marginal Secure (0), Medium Insecure (-0.5), and Extremely Insecure (-1). The two DR thresholds can now be presumed to be (26) as follows:

$$Class = \begin{cases} 1(DR \geq 6.00\%) \\ 0.5(3.00\% \leq DR < 6.00\%) \\ 0(DR < 3.00\%) \end{cases} \tag{26}$$

Therefore, the multi-class CT with preventive control rules is grown as Fig. 10.

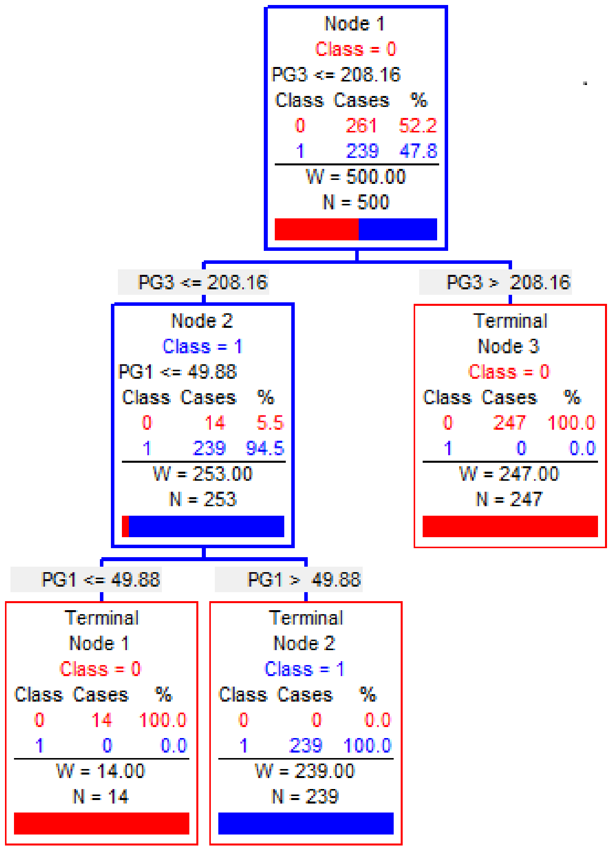

Fig. 9. Two-Class CT Preventive Control with DR threshold of 6.00% in IEEE 39-Bus System

5) *Performance Test Using CTEI*

Both two-class and multi-class CT are feasible when subject to different customised system requirements. However, it is essential to conduct a reliability test on both methods in terms of how well both CT types meet the system desired DR requirement.

a) *Security Concern*

DR of original OP under OPF is 4.38%.

DR of the new OP under OPF using two-class CT is 4.56%.

DR of the new OP under OPF using multi-class CT to is 5.34%.

DR threshold required by SO's requirement is presumed to be 6.00%.

b) *Economic Concern*

Total cost of original OP under OPF is $100,950.64/hr

Total cost of the new OP under OPF using two-class CT is $103,455.79/hr.

Total cost of the new OP under OPF using multi-class CT to is $105,480.50/hr.

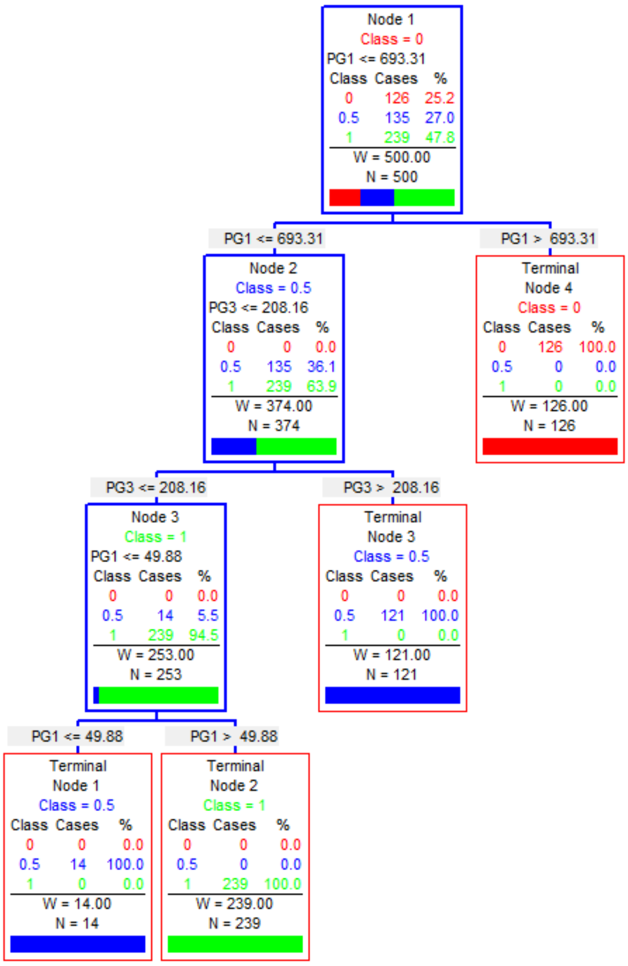

Fig. 10. Multi-Class CT Preventive Control with DR threshold of 6.00% in IEEE 39-Bus System

Refer to the predefined Classification Tree Reliability Index (CTEI) in (14), as $DR_i/DR_{r,i}<1$, relevant index calculations can be performed as (27). Hereby, both the economic weight (EW) and security weight (SW) is assumed as equally 50% for convenience.

$$CTEI_i = [EW_i \cdot \frac{TC_{o,i}}{TC_{n,i}} + SW_i \cdot (\frac{DR_i - DR_0}{DR_{r,i} - DR_i})] \times 100\% \qquad (27)$$

where i=1,2.

Substitute the above data, the following CTEI of both two-class CT and multi-class CT can be obtained as:

CTEI (two-class) = 54.34% (28)

CTEI (multi-class) = 77.48% (29)

Hence, the results turn out that optimal rules using multi-class CT is much more reliable and effective with regards to both system security and economic considerations given the same set of database. This is reaffirmed by the applicable practice of SO's defining multi levels of oscillatory security requirements for different regions throughout the network. CTEI can be improved by better quality and large size of database generated.

B. *NEM 14-Generator Network (Incomplete)*

National Energy Market (NEM) 14-Generator network was firstly proposed by Gibbard with the University of Adelaide in 2010. Fig.11. shown is the schematic diagram of NEM 14-Generator network, which contains Area 1 to Area 5 loosely representing respectively the Snowy Hydro (SH), New South Wales (NSW), Victoria (VIC), Queensland (QLD) and South Australia (SA) [21]. This network of an incomplete DIgSILENT version was developed by Gregor Verbic with the University of Sydney and is used as the testing system roughly simulating the South Eastern Australian coastal areas for research on stability, renewable energy penetration and power market studies.

The area-generator configuration of NEM 14-Generator network can be categorised according to the responsible power supply area of corresponding generator groups. For example, generator XPS_Y_N (X is an alphabet, Y is the area code, N is the number identification in the generator group) is the Nth generator responsible for supplying area Y, which a participating member of the generator group of XPS.

1) *System Pre-processing*

The originally provided NEM network in DIgSILENT authored by Dr. Gregor Verbic has all 6 operation scenarios under small-signal instability before applying any amendments. It can be observed in Fig.13. that the two closest system modes to the origin in the RHP have real parts of 0.4102, which is unsatisfactory for data generation.

Since the given NEM system is of high complexity and oscillatory instability yielding uniformly system minimum DR of -1.00 by DIgSILENT when performing a generation rescheduling even within the convergence of power flow calculations, it is essential to tune the current system OP by reducing active load power and raising active power generation. Area 2 of NSW is responsible for the heaviest stress of load so that focused considerations are devoted to both the load and generation statuses in this region. Such method is a module comprising parameter tuning in [6].

After applying parameter tuning method along with load compensation, the closest RHP mode is now moved towards the origin with a reduced real part to 0.056. However, DR still cannot be tuned positive until PFTA is employed. Fig.12. below illustrates the participation factor bar plot with active power output (APO) as state variable generated by DIgSILENT. Positive participation factor implies contribution of excessive APO to oscillatory instability of the outlined mode thus its corresponding APO needs to be reduced.

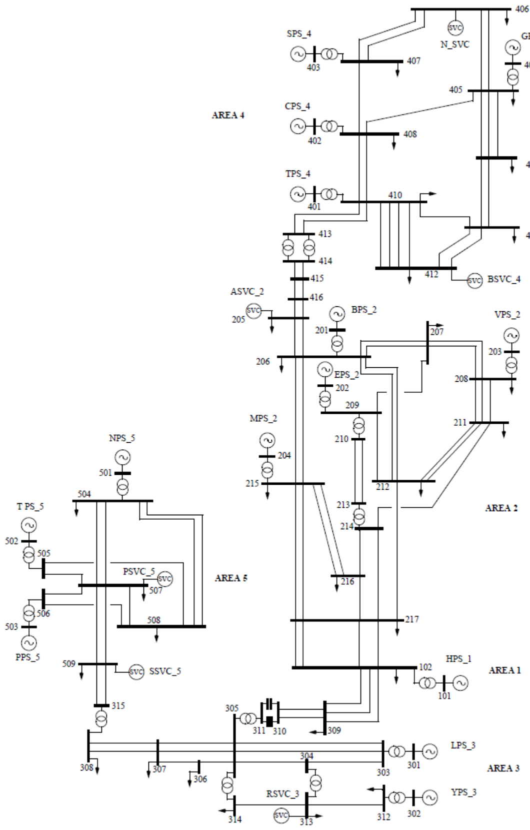

Fig. 11. NEM Network

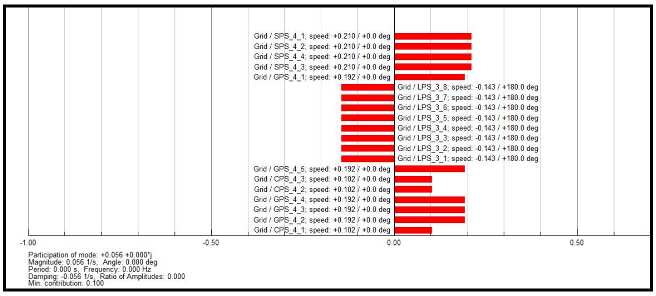

Fig. 12. Participation Factor Bar Plot of the Closest RHP Mode in NEM

Fig.13. shows the enhancement process of system DR and movement of the closest RHP mode to the original benefited from different stages of system pre-processing applied.

2) *Data Generation and Critical Generators*

Due to the unavailability of control configurations in this NEM 14-Generator network on DIgSILENT, RELIEF-F does not yield such distinguishable importance weight ranking of G2 to G14 as that of IEEE 39-Bus System. RELIEF-F heavily relies on distance among instances and a negligible distance caused by incomplete control settings in complex system like NEM during DIgSILENT modelling stage will disable its usage. However, it is still feasible to resort to the sensitivity analysis in order to rank the criticalness of those generators for DR enhancement. In (21), ΔP_i is the APO difference approaching zero and can be set to incremental 0.1MW to each generator per time.

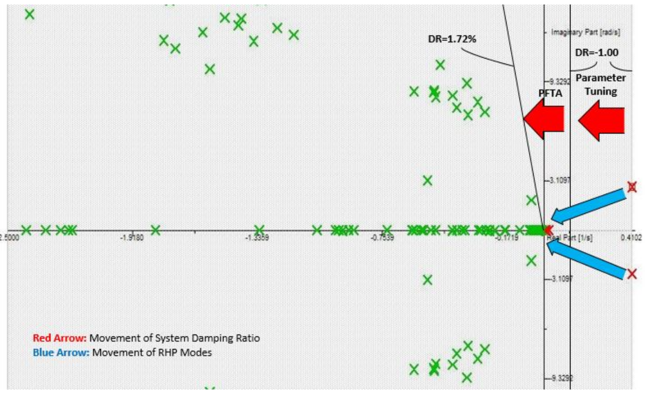

Fig. 13. Fast PFTA to Achieve Fast System Tuning in NEM

TABLE IV
Generator Damping Ratio (DR) Sensitivity in NEM

Generator ID.	*Damping Ratio Sensitivity*
G1 (HPS_1)(Slack)	N/A
G2 (BPS_2)	2.37824×10^{1}
G3 (EPS_2)	2.04749×10^{0}
G4 (MPS_2)	-8.07428×10^{0}
G5 (VPS_2)	1.211141×10^{2}
G6 (LPS_3)	-3.69589×10^{1}
G7 (YPS_3)	3.14988×10^{1}
G8 (CPS_4)	-9.10849×10^{1}
G9 (GPS_4)	-3.96895×10^{1}
G10 (SPS_4)	3.96895×10^{1}
G11 (TPS_4)	-3.41242×10^{0}
G12 (NPS_5)	8.63672×10^{2}
G13 (PPS_5)	-6.00532×10^{2}
G14 (TPS_5)	2.78344×10^{2}

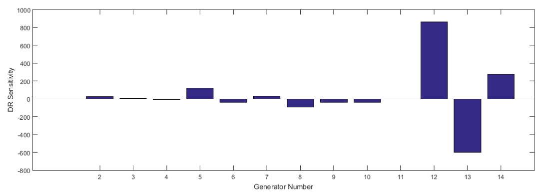

Fig. 14. Generator DR Sensitivity Ranking in NEM

After generating the power active output database with 0.1MW incremental power injected to one generator per time in ascending order (i.e. from G2 to G14) and maintaining the rest of generator active power output constant per time, DIgSILENT reads in those 14 × 13 generator feature matrix keeping the original OP in the first line, and perform automatic MA through DPL. It is then followed by MATLAB's communication of reading through the 13 output text-files containing 432 complex eigenvalues each to compute the system minimum DR of each case. Finally, substituting parameters into (21) will obtain the relevant DR sensitivity of each generator. In order to assure the numerical accuracy of the DR sensitivity, ΔP_i can be assigned as 10^{-n} MW, where *n=1, 2, 3,* ... etc. and the resulting DR sensitivity can be averaged by *n* to acquire relatively accurate outcomes. Table IV above is tabulated to illustrate the sensitivities of each generator (G2-G14), with Fig.14. showing the bar plot of generator DR sensitivity of G2 to G14.

3) Optimised Rules from CT

Similarly with 641 OPs containing input and output features, CT for preventive control in NEM is grown to achieve SSA. Due to practicalities of incomplete control settings in current DIgSILENT NEM model, system DR of those 641 OPs only ranges from 1.38%-1.72%. However, it still suffices for machine learning to obtain CT optimised preventive control rules. If the DR threshold is set to 1.50%, the CT is grown as Fig.15. below.

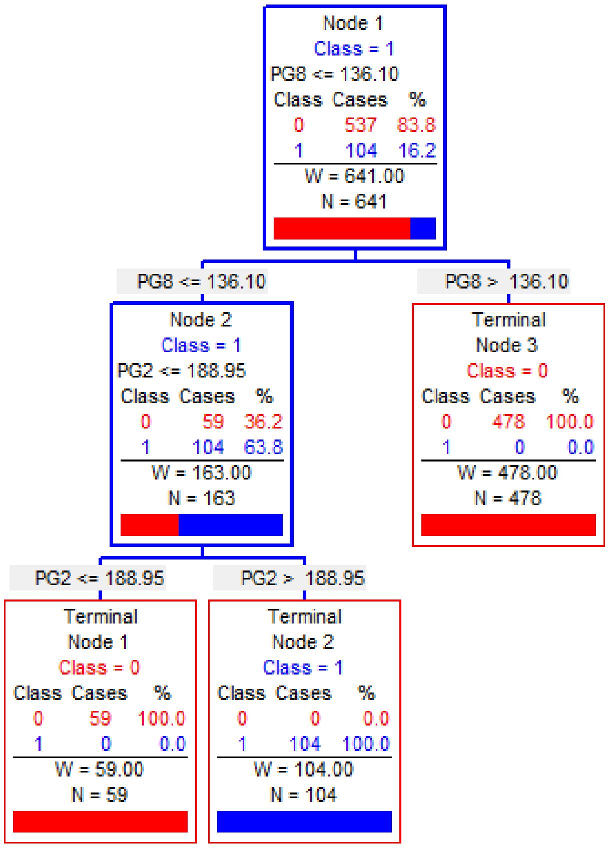

Fig. 15. Two-Class CT Preventive Control in NEM

It is observed that PG2 and PG8 are regarded as stakeholders in system DR adjustment, which reinforces the findings from the former DR sensitivity analysis that generators with the 4th and 6th highest sensitivity participate in controlling a comparatively low system DR around 1.50% via generation rescheduling (GR).

VIII. CONCLUSION AND FUTURE WORK

This paper presents a DT-based online preventive control tool to enhance system DR adopted as the major part of small-signal stability assessment (SSA) tool developed on DIgSILENT/MATLAB. SSA tool enables the further development of reliable, diversified and powerful modelling and computational functionalities in power system planning, modelling, and control under renewable energy penetration. The DPL and MATLAB scripts from this paper can be the backbone of the automated SSA tool interface. Apart from demonstrating a novel scheme for quick system adjustment to find stable OP with positive DR enabled by PFTA, DT performance is also evaluated by the proposed CTEI regarding both the effectiveness to achieve desirable system DR and minimum security cost.

Further work is recommended on validating the tool in larger scale power systems. Additionally, a Python-based interface of SSA communicating between DIgSILENT and MATLAB can be realised to effectively reduce the operational complexity of DIgSILENT and accelerate the computation process for relevant future researches.

IX. REFERENCES

[1] Anderson, G., et al. "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance". *Power Systems, IEEE Transactions,* vol. 20.4, pp. 1922-1928, 2005.

[2] Y.Xu , et al. "A reliable intelligent system for real-time dynamic security assessment of power systems." *Power Systems, IEEE Transactions on* 27.3, 2012.

[3] Mansour, Yakout, Ebrahim Vaahedi, and Mohammed El-Sharkawi. "Dynamic security contingency screening and ranking using neural networks." *Neural Networks, IEEE Transactions on* 8.4, 1997.

[4] Mohammadi, M., and G. B. Gharehpetian. "Application of core vector machines for on-line voltage security assessment using a decision-tree-based feature selection algorithm." *IET generation, transmission & distribution* 3.8, 2009.

[5] Y. Xu, et al. "A decision tree-based on-line preventive control strategy for power system transient instability prevention", *International Journal of Systems Science*, 45:2, 176-186, 2014.

[6] P. Kundur, *Power System Stability and Control*, McGraw-Hill Inc., New York, NY, 1994.

[7] J.F. Hauer, J.R. Hunt, in association with the WSCC System Oscillations Work Groups, V Symposium of Specialists in the Electric Operational and Expansion Planning (SEPOPE), Recife (PE), Brazil, May 19-24, 1996

[8] Huang, Zhenyu, et al. "Improving small signal stability through operating point adjustment." *Power and Energy Society General Meeting, 2010 IEEE*. IEEE, 2010.

[9] DIgSILENT GmbH, *PowerFactory User's Manual*, *Gomaringen, Germnay, May*, 2011.

[10] Kundur, Prabha, et al. "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions." *Power Systems, IEEE Transactions on* 19.3 2004.

[11] L. Breiman, J.Friedman, R.A.Olshen, and C.J. Stone, *Classification and regression trees*. CRC press, 1984.

[12] Genc, Istemihan, et al. "Decision tree-based preventive and corrective control applications for dynamic security enhancement in power systems." *Power Systems, IEEE Transactions on* 25.3, 2010.

[13] Zhu, Jizhong. *Optimization of power system operation*. John Wiley & Sons, 2015.

[14] Zimmerman, Ray D., Carlos E. Murillo-Sánchez, and Deqiang Gan. "MATPOWER: A MATLAB power system simulation package." *Manual, Power Systems Engineering Research Center, Ithaca NY* 1, 1997.

[15] Kira, Kenji, and Larry A. Rendell. "A practical approach to feature selection."*Proceedings of the ninth international workshop on Machine learning*. 1992.

[16] Kononenko, Igor. "Estimating attributes: analysis and extensions of RELIEF."*Machine Learning: ECML-94*. Springer Berlin Heidelberg, 1994.

[17] Meng, Ke, et al. "Speed-up the computing efficiency of power system simulator for engineering-based power system transient stability simulations." *IET generation, transmission & distribution* 4.5, 2010.

[18] Pai, Anantha. *Energy function analysis for power system stability*. Springer Science & Business Media, 2012.

[19] Nguyen, Tony B., and M. A. Pai. "Dynamic security-constrained rescheduling of power systems using trajectory sensitivities." *IEEE Transactions on Power Systems* 18.2, 2003.

[20] DIgSILENT GmbH, *39 Bus New England System*, Gomaringen, Germnay, 2015.

[21] Gibbard, M. and Vowles, D., "Simplified 14-Generator Model of the SE Australian Power System," The University of Adelaide, TechRep., 2010.