Voltage Stability Prediction Using Active Machine Learning

Vuk Malbasa, *Member, IEEE*, Ce Zheng, *Member, IEEE*, Po-Chen Chen, *Student Member, IEEE*, Tomo Popovic, *Senior Member, IEEE*, and Mladen Kezunovic, *Fellow, IEEE*

***Abstract*—An active machine learning technique for monitoring the voltage stability in transmission systems is presented. It has been shown that machine learning algorithms may be used to supplement the traditional simulation approach, but they suffer from the difficulties of online machine learning model update and offline training data preparation. We propose an active learning solution to enhance existing machine learning applications by actively interacting with the online prediction and offline training process. The technique identifies operating points where machine learning predictions based on power system measurements contradict with actual system conditions. By creating the training set around the identified operating points, it is possible to improve the capability of machine learning tools to predict future power system states. The technique also accelerates the offline training process by reducing the amount of simulations on a detailed power system model around operating points where correct predictions are made. Experiments show a significant advantage in relation to the training time, prediction time, and number of measurements that need to be queried to achieve high prediction accuracy.**

***Index Terms*—Machine learning, active learning, data mining, synchrophasors, prediction methods, power system analysis, power system planning, power system stability, power transmission, smart grids.**

NOMENCLATURE

ANN	Artificial Neural Network
DT	Decision Tree
GPS	Global Positioning System
OP	Operating Point
PMU	Phasor Measurement Unit
PSS®E	Power System Simulator for Engineering
RBF	Radial Basis Function
RF	Random Forest
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machine
VCP	Voltage Collapse Point
VSM	Voltage Stability Margin
WECC	Western Electric Coordinating Council.

Manuscript received October 31, 2016; revised February 27, 2017; accepted April 8, 2017. Date of publication April 12, 2017; date of current version October 19, 2017. This work was supported in part by the PSerc Project S-44 "Data Mining to Characterize Signatures of Impeding System Events or Performance Using PMU Measurements," and in part by the Texas A&M Engineering Experiment Station. Paper no. TSG-01498-2016. *(Corresponding author: Vuk Malbasa.)*

V. Malbasa is with the University of Novi Sad, Novi Sad 21000, Serbia (e-mail: vmalbasa@uns.ac.rs).

C. Zheng is with Electrocon International, Inc., Ann Arbor, MI 48103 USA (e-mail: ce@electrocon.com).

P.-C. Chen and M. Kezunovic are with Texas A&M University, College Station, TX 77840 USA (e-mail: pchen01@tamu.edu; kezunov@ece.tamu.edu).

T. Popovic is with the University of Donja Gorica, 81000 Podgorica, Montenegro (e-mail: tomo.popovic@udg.edu.me).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2017.2693394

I. INTRODUCTION

POWER grid operation needs new monitoring systems that increasingly provide more accurate measurements about the grid behavior to the operator [1]. The utilization of data from new substation equipment such as PMUs contributes to improved decision making and operation [2]. Because the measurements are typically gathered at both high velocity and volume, it becomes imperative to exploit new online data analytics with fast data processing capability.

Traditional time-domain simulation approach, based on the first principles modeling of a power system is computationally intensive and may not meet the requirements of real time applications. Researchers have therefore turned to machine learning techniques to allow for on-line decision-making. The machine learning approach has the generalization ability where the data-based model, if properly trained, can make accurate predictions from measurements that it has not been exposed to previously.

Examples of machine learning applications in power grids include: a) predicting wind farm power output [3]; b) detecting the substations most affected during major disturbances [4]; c) analyzing the faulted transformers [5]; d) obtaining customer load profiles [6]; e) performing dynamic security assessment [7]; f) enhancing the data debugging in power grid operations [8]; g) tackling voltage security concerns [9]; h) building classifiers for stability assessment [10]; i) analyzing power quality data for load estimation [11]; and j) estimating the stability margins from synchrophasor measurements [12]. A wide range of machine learning applications for power grids may be found in [13] and [14].

While the machine learning algorithms excel in their fast decision making capability, two technical difficulties have not been fully resolved yet: a) how to update the machine learning models when predictions contradict the actual system conditions; and b) how to efficiently prepare the training data to eliminate redundant offline power system simulations.

Voltage stability in a transmission system is among major challenges to the operation of an electric grid. For efficiency and economic benefits, grid operators tend to operate the network close to its physical limits [15]. The ability to provide decision-making support in real time is desired and can greatly improve the stability, security and reliability of a power system.

The main contribution of this work is the adaptation of pool-based active learning methodology to power system measurements such as synchrophasor data used to assess voltage stability. It accounts for the uncertainty of machine learning models and the computational burden of training and voltage instability prediction, which have not been taken into consideration by the previously reported work.

The rest of the paper is organized as follows. After the introduction, a background of machine learning applications in power systems is given. The proposed methodology, the results from case studies, and the discussion of results are presented next. The paper ends with the conclusions and references.

II. Background

A. New Voltage Stability Monitoring Techniques Motivated by New Data Acquisition Capabilities

For a transmission system, one critical aspect of operation is maintaining its voltage stability. Considering the large scale of transmission systems, data collection based on traditional SCADA system from geographically dispersed locations may cause fast decision making to become challenging due to large amount of measurements to be processed at each scan period.

New PMU data acquisition technology offers accurately measured GPS-synchronized phasor magnitude and phase data, or synchrophasors, at rates between 30 and 240 samples per second [16]–[18]. The judicious placement and high data-sampling rate of the new data acquisition systems have greatly motivated the application of the measurement-based approaches for monitoring voltage stability problems using the machine learning technique described here.

B. Drawbacks of Existing Machine Learning Applications

When using machine learning in power system studies, many approaches are based on a black box methodology [19], and very few papers try to interpret the trained models [20]. Recently, the most popular choices in literature have been the SVMs [21], ANNs [22], and DTs [23]. Because of the difficulties in interpreting these models, most past work either employs models that are easy to interpret, or applies the SVMs, ANNs and DTs in a potentially limiting way.

The performance of a machine learning tool depends on the capacity of known system behavior represented by the training data set. Therefore, existing machine learning approaches usually employ a brute force method to generate the knowledge base using physical system model-based simulations. Due to the computational complexity from the large amount of required simulations, a single, unchanged knowledge base has been used in many applications. The machine learning tools trained in this way are “passive” and suffer from three main drawbacks when embedded in online applications:

- The initial training set is formulated at a certain physical system operating condition. When subsequent updates are not accounted for, the machine learning tool may fail as system operating conditions evolve;
- When machine learning models are updated, without reducing the amount of required simulations, the training based on the same brute force simulation method greatly delays the resumption of the online application.
- The knowledge base is created solely based on the power system model. When fidelity of the system model is questionable, the prediction accuracy of the machine learning tools is adversely impacted.

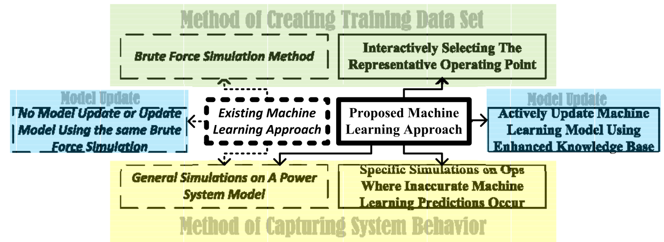

Fig. 1. Proposed approach versus existing approaches.

C. Benefits of the Proposed Active Machine Learning Approach

Active learning is employed in this work to tackle the mentioned drawbacks. Fig. 1 shows the difference between the proposed approach and the existing machine learning algorithms. The active learning technique is iteratively building the knowledge bases, which is optimized for size and accuracy. Instead of using the exhaustive simulation method, it interactively selects the most representative OPs when building the training data set. Instead of relying on all-inclusive simulations based on a power system model, the proposed approach actively searches for the OPs where inaccurate machine learning predictions occur. It then performs special simulations to create new mappings around the identified OPs and adds them to the existing pool of training data set. This helps capture the hidden system behavior not represented in the training process previously.

In most cases, simulations based on power system modeling are used to verify the prediction accuracy and identify any contradictions between machine learning prediction and actual system behavior. In some rarer cases, such contradiction may be observed without simulation verification. When the monitored power system has lost its voltage stability but the voltage stability predictor still reports stable condition is a good example of an obvious contradiction. The proposed approach also includes these directly observed contradictions in the training pool. The objective is to minimize the negative influence from inaccurate models of power system components.

Our approach is general enough to accommodate most common machine learning models, which can function as a probabilistic classifier. The uncertainty of a probabilistic classifier is examined to guide the knowledge base creation. The active learning approach has been successfully applied in the past to tasks such as land cover classification [24], medical image classification [25], and text categorization [26]. More recently the focus has shifted towards stream mining [27], novel applications in medicine [28], collaborative filtering [29], as well as more thorough theoretical analysis [30].

In our work, the proposed approach applied to the voltage stability problem illustrates the potential for improvement by quantifying the performance of several common machine learning models in terms of training time, prediction time, and accuracy. The voltage stability margin is computed to label each OP as stable or unstable given a transmission system state.

With pool-based active learning, a labeled data set is created from a large pool of unlabeled data using an “oracle” which at great cost produces an accurate label. The term “oracle” refers to the steady-state or dynamic simulation based on a detailed power system model. The data set refers to the OPs generated by the physical system model-based simulation at various system operating conditions. In the presented work, the labeled OP can also be the known mapping between streamed measurements and actual system behavior. The machine learning model is then trained to approximate the oracle in a manner which minimizes the number of queries to the oracle. Pool-based active learning has been investigated frequently in similar cases where human experts provide labels for data [28]. These approaches are useful because of the large difference between the speed with which the system operator can provide labels and the duration of the machine learning model’s training and prediction phase.

III. Methodology

A. Voltage Stability in Transmission Systems

Voltage stability refers to the ability of a power system to maintain steady voltages at all buses in the system after being subjected to a disturbance from a given initial operating condition [31]. In this work, the long-term voltage stability of a power transmission system in response to slowly varying load conditions was studied. As mentioned earlier, grid operators tend to operate the network close to its physical limits. When these limits are breached, the system may experience several forms of system-wide failures including voltage collapse [32].

Voltage stability may be approximated by calculating the distance of the current OP from the VCP [33]. As indicated in Fig. 2, voltage instability conditions arise when load demand attempts to move beyond maximum deliverable power. Normal system OPs are above the VCP, along the line defined by the relationship between demand and corresponding voltage magnitude at the load bus. When the OP moves to a point below the VCP, catastrophic consequences may occur.

An indicator of the voltage stability, VSM at a load bus may be computed as:

$$VSM = \frac{P_{max} - P_{current}}{P_{max}} \tag{1}$$

where P_{max} is the maximum deliverable power, $P_{current}$ is the load active power demand at the current OP.

B. Problem Statement

In order to train a machine learning model to provide the desired predictions in a power grid, a knowledge base that captures sufficient power system behavior needs to be created. In an example where synchrophasor measurements are available, the knowledge base consists of N OPs represented by D data channels, or attributes, recorded by PMUs in the system. Each OP is represented by synchronized magnitude and phase angle measurements of voltages and currents. Synchrophasors summarize the i-th OP with the vector $\mathbf{x}_i = [x_{i1}, x_{i2} \ldots x_{iD}]^T$ where x_{ij} is the feature, or the value of a certain electric quantity recorded from channel j when the system was at operating point i. The set $\mathbf{D}_\mathrm{U}$ is an unlabeled data set, $\mathbf{D}_\mathrm{U} = \mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N]$. For each OP, $\mathbf{x}_i$ is associated with its appropriate quantity of interest y_i, the vector of labels is $\mathbf{Y} = [y_1, y_2, \ldots y_N]$.

For the voltage stability application in transmission systems a three-class problem has been considered, where $y_i = 1$ represents OPs with stability margins that are larger than the mean stability margin value of all OPs, $y_i = 2$ for OPs with a stability margin in the second quartile, and $y_i = 3$ if the stability margin is in the smallest quartile. The range of *VSM* was from 0.1% to 6%, and further details can be found in [12].

Using the labeled data set makes it possible to train a machine learning model *f*, which can then for an unseen OP $\mathbf{x}_j$ provide an approximation of the voltage stability.

The task is to build a knowledge base which will lead to optimal learning per labeled OP. Given $\mathbf{D}_\mathrm{U}$, the objective is to find subsets $\mathbf{D}_\mathrm{L}$ incrementally increasing in size, which optimally increase the prediction accuracy of a machine learning model trained on the current $\mathbf{D}_\mathrm{L}$.

This problem mimics the common situation in a power system where generating OPs in $\mathbf{D}_\mathrm{L}$ are computationally expensive while obtaining OPs for $\mathbf{D}_\mathrm{U}$ is cheap.

C. Classification Problem

The classification problem consists of constructing a mapping f between point $\mathbf{x}_i$ from the input space and the corresponding label $y_i, f(\mathbf{x}_i) \approx y_i$.

In order to solve the multi-class problem where $y_i \in \{1, 2, 3\}$ a set of 3 classifiers approximate $sign(f_c(\mathbf{x}_i)) \approx id(y_i, c)$ for c taking the values 1, 2, and 3, and *id* being the identity function.

D. Probabilistic Classifiers

Probabilistic classifiers predict a probability distribution function across all possible labels. An advantage of such

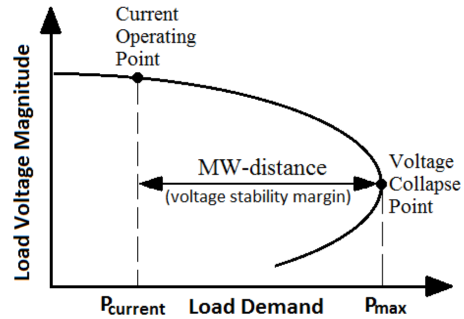

Fig. 2. VCP and VSM on the load demand – voltage magnitude curve.

classifiers is that they can provide confidence along with predictions.

For the 3-class problem of transmission system voltage stability classification, the probabilistic prediction for the i-th OP consists of

$$\left[\sigma(f_1(\mathbf{x}_i))/z, \sigma(f_2(\mathbf{x}_i))/z, \sigma(f_3(\mathbf{x}_i))/z\right] \tag{2}$$

where $z = \Sigma_c[\sigma(f_c(\mathbf{x}_i))]$ is a normalization constant and σ is the logistic sigmoid function.

E. Artificial Neural Networks

The ANN was trained to predict one label as a continuous output. One specific property of ANNs used for binary classification is that they generalize the entire input space even if only very few OPs are used for training. A good heuristic approach to solving this problem involves including at least one OP from each class into the training set. Experiments were performed using the MATLAB Neural Network Toolbox [34].

For the binary case the class probability $\mathrm{p}(y_i = 1|\mathbf{x}_i)$ also indicates $\mathrm{p}(y_i = 0|\mathbf{x}_i) = 1 - \mathrm{p}(y_i = 1|\mathbf{x}_i)$. We obtained the relative uncertainty from the proximity of $\mathrm{p}(\mathbf{x}_i)$ to 0.5, or the proximity of the i-th OP to the decision boundary.

F. Support Vector Machines

During the deployment of applications which are based on the SVM for classification, new OPs are classified by this model according to their similarity to a set of OPs chosen to represent a decision boundary.

The locality of the RBF kernel implies that unlike ANN, confident predictions about OPs dissimilar to data in the training set are not usually made. The LIBSVM library was used for experiments [35].

G. Decision Trees and Random Forests

In an ensemble like the RF, the number of trees used dictates the potential number of values that are obtainable as the probability estimate $\mathrm{p}_f(y_i = 1|\mathbf{x}_i)$ and is therefore more suitable for active learning than pure decision trees. A port of the *randomForest* R-package [36] was used for experiments.

H. Active Learning

Instead of assuming that for all OPs x_i have the appropriate y_i, the data set $\mathbf{D}_\mathrm{L}$ is initially assumed to be empty, or only initialized with an OP from each class. Domain knowledge may typically be used to provide several OPs of each stability state to initialize $\mathbf{D}_\mathrm{L}$, but these OPs may not be relied upon for accurate predictions.

Adding labels to OPs in $\mathbf{D}_\mathrm{U}$ using the oracle increases the size of $\mathbf{D}_\mathrm{L}$. This set can be used to quantify the increase in accuracy, after each labeled OP, for both random sampling and active learning.

I. Pool-Based Active Learning

The details of pool-based active learning are illustrated in Fig. 3. A larger pool of unlabeled data representing power system operating conditions is iteratively labeled with an oracle (simulation based on detailed system modeling) in order to create a labeled data set which a machine learning model may be trained on. The machine learning model's uncertainty is used to select data points for labeling by the oracle. At each iteration, a partially trained classifier chooses an example $\mathbf{x}^*$ from the unlabeled data pool about which the classifier is most uncertain about.

When embedded in online power system applications, the unlabeled data pool refers to the time-series measurements

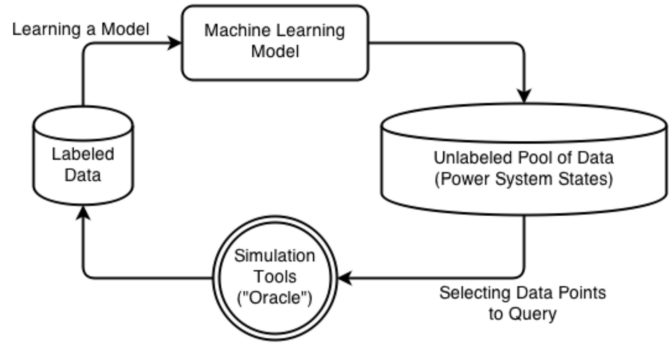

Fig. 3. The pool based active learning scheme.

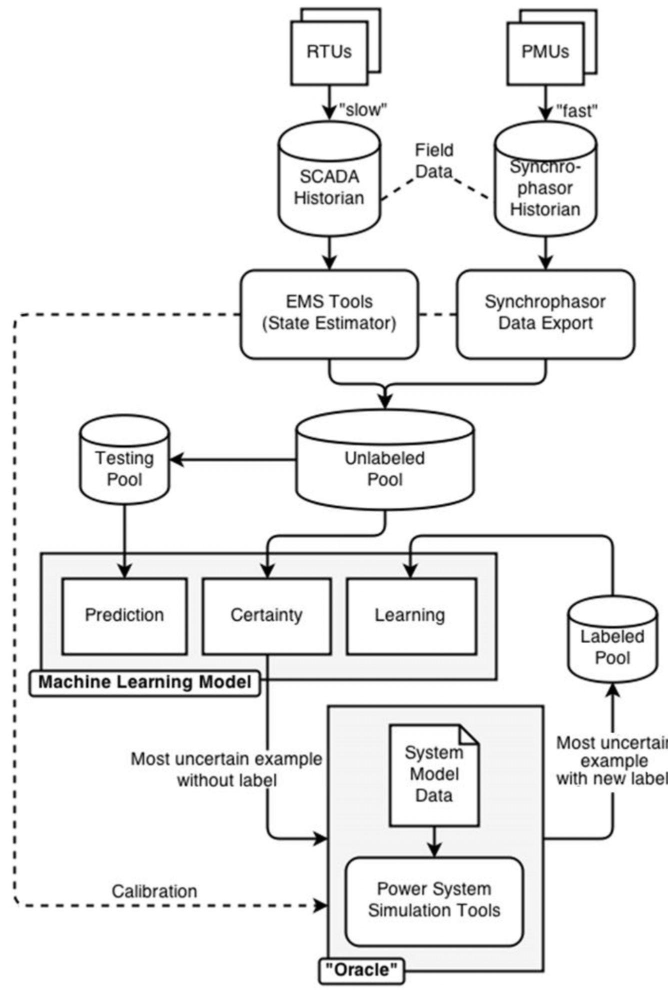

Fig. 4. Integration of the proposed active learning approach in power systems data acquisition infrastructure, representing data flow.

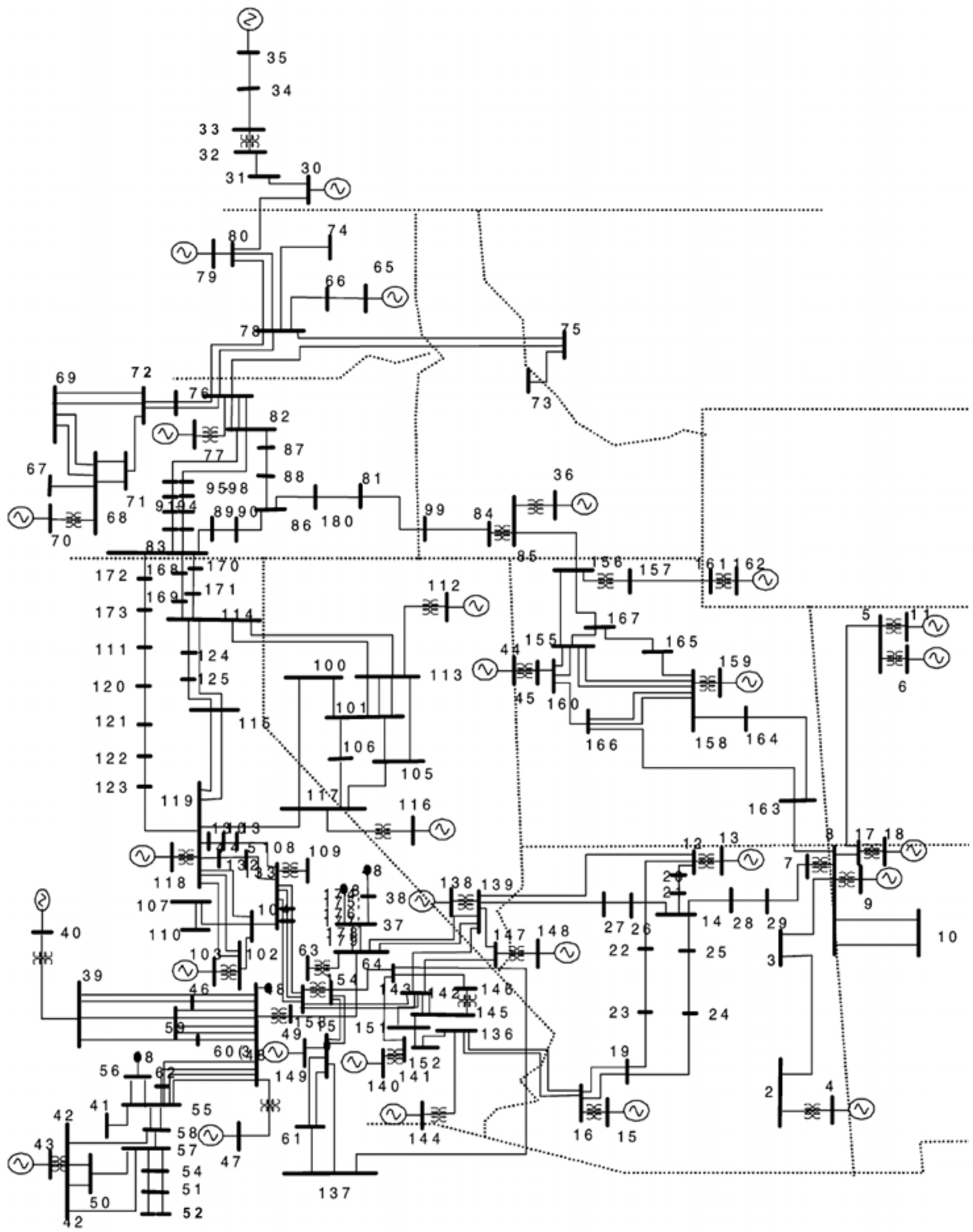

Fig. 5. One-line diagram of the WECC 179-bus equivalent system.

streamed to a central facility where the machine learning model resides. The proposed technique continually compares the machine learning model's prediction with actual system behavior. Once a contradiction is identified, the corresponding OP is recorded. The "oracle" is used to generate a precise label for the OP through model-based simulation. In the presented work, voltage stability status is determined as a label y^*, and assigned to the OP. The newly labeled OP may now be included into the pool of $\mathbf{D}_{\mathrm{L}}$, so that it can be used in the next iteration of learning.

J. Multi-Class Sampling Strategy

For multi-class problems, several approaches can be applied within the described framework. In order to take advantage of additional information obtained from a prediction in the multi-class problem, the margin sampling approach chooses the OP with the smallest difference in probability between the first and second most probable label, as computed from the prediction.

K. Active Learning in Power Systems

As previously discussed, in the study of transmission system voltage stability the VSM is used as the indicator, or label. For a large power network, it may take hours to create labeled OPs by using iterative continuation power flow calculation on the detailed system model built in the commercial stability program PSS®E [12], [37].

The integration of the proposed active learning approach in power system applications is illustrated in Fig. 4. The synchrophasor measurements are streamed from PMUs into the Unlabeled Pool. In this example, activity does not occur simultaneously. The "oracle" (model-based simulation) is calibrated

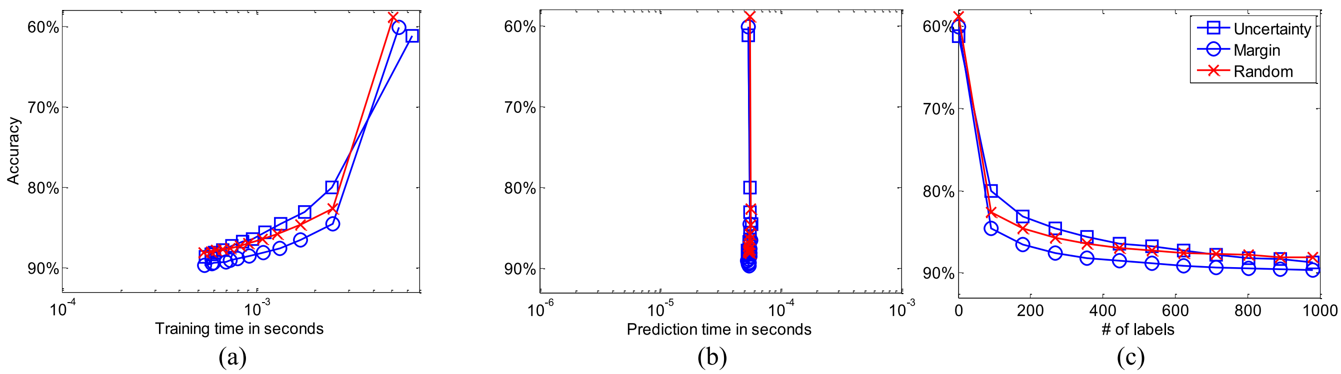

Fig. 6. Voltage stability prediction using ANNs: a) training time per OP, (b) testing time per OP, and (c) number of labeled OPs (c) on the x-axis; accuracy on the y-axis.

off-line, data in the Unlabeled Pool may be historical DT, RF, or SVM. The "oracle" is used for labeling OPs, which are then included in the Labeled Pool for learning.

By making predictions on the unlabeled pool, the most uncertain OPs are identified. The OPs are then assigned precise labels by the "oracle", and stored in the labeled pool so they can be used for later learning. Field measurements can be used to verify and calibrate simulation tools during the initial setup of the system.

The machine learning model has three relevant properties in this work: 1) it is trained using the labeled pool, 2) uncertainty is estimated through probabilistic classification, and 3) an explicit class may be assigned to the i-th OP by finding $p_f(y_i|\mathbf{x}_i)$, the maximal value in (2).

IV. Case Study

The proposed approach is evaluated in experiments using synthetic data obtained from simulations on detailed power system model. Its performance is quantified in terms of prediction and training time, and prediction accuracy.

The experiment focuses on predicting the voltage stability margins in a transmission network. The test network is the simplified WECC system which consists of 29 generators, 179 buses, 263 transmission lines, 42 shunts, and 104 loads. The one-line diagram is shown in Fig. 5. The knowledge base prepared by the oracle includes 5078 "stable", 2540 "alert", and 2529 "critical" labeled OPs. A total of 256 channels of simulated phasor data were gathered, covering 10147 selected OPs.

Because this is a 3-class problem, uncertainty and margin sampling were both compared to random sampling by labeling 1000 OPs and quantifying the relevant metrics.

In the case of three classes, the following accuracy metric was used,

$$Accuracy = \sum_{i \leq N} id(y_i, f(x_i))/N, \tag{3}$$

where id is the identity function, $y_i \in \{1, 2, 3\}$.

In this experiment 1000 random OPs from each class for a total of 3000 OPs were used for testing. Table I shows the final accuracy for the transmission system voltage stability prediction, while Table II summarizes the speedups achieved over the baseline random sampling approach. The results of this experiment are illustrated in Figs. 6, 7 and 8.

TABLE I
Accuracy in Transmission System Experiment

	RF	ANN	SVM
Uncertainty	89.91%±0.5%	88.72%±0.9%	86.59%±0.8%
Margin	**90.01**%±0.4%	**89.73**%±0.9%	**86.7**%±0.8%
Random	84.31%±1%	88.16%±1%	81.79%±-0.9%

TABLE II
Speedups in Transmission System Experiment

		RF	ANN	SVM
Margin	A vs #L	91%±20%	68%±82%	121%±70%
	A vs TR	313%±128%	11%±19%	977%±603%
	A vs TE	128%±32%	6%±11%	307%±73%
Uncertainty	A vs #L	140%±55%	217%±150%	125%±61%
	A vs TR	319%±128%	17%±19%	973%±572%
	A vs TE	124%±33%	5%±11%	310%±74%

V. Discussion

A feed forward, back-propagation, artificial neural network's prediction time is theoretically independent of the size of the knowledge base, and this fact is reflected in all experiments involving ANN in Fig. 6. The specific architecture and parameterization used are well suited to the transmission problem, where data is not as noisy.

In Fig. 6 (a) the training times decrease as the number of OPs is increased. This may be interpreted as the training process necessitating less iterations where more OPs are available, where each individual iteration is more informative.

ANN prediction time is hardly affected by the training set size in Fig. 6 (b), the RF does increase significantly in Fig. 7 (b) as well as the SVM in Fig. 8 (b). Margin sampling outperforms uncertainty sampling in the transmission experiment in Fig. 7 (c), indicating that the additional information used in this approach is useful for ANN training set creation. The accuracy of uncertainty sampling improves over random sampling only after approximately 700 OPs are labeled in Fig. 7 (c).

In Fig. 7 (c), margin sampling outperforms random sampling only after 300 points are labeled, while uncertainty sampling outperforms it from the beginning. The prediction time remains mostly the same even after several hundred OPs

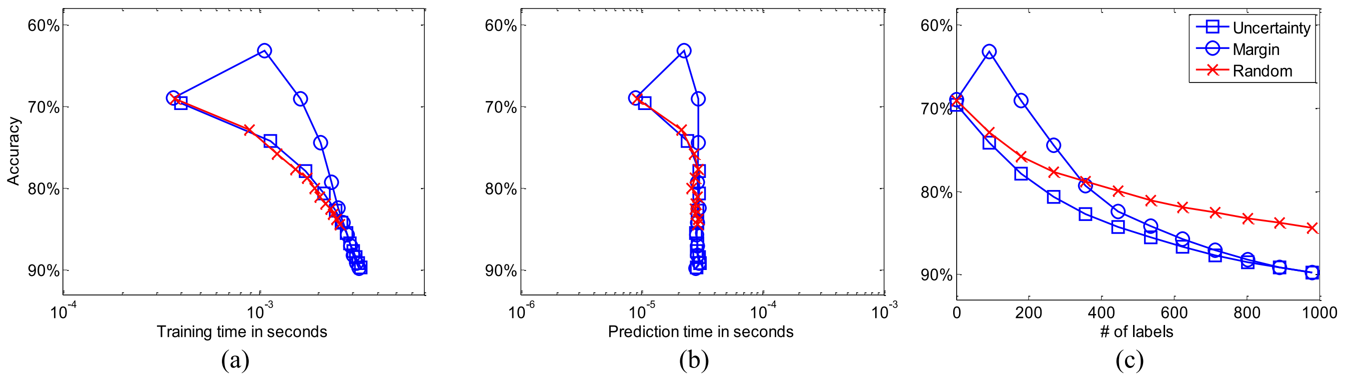

Fig. 7. Voltage stability prediction using RFs: a) training time per OP, (b) testing time per OP, and (c) number of labeled OPs (c) on the x-axis; accuracy on the y-axis.

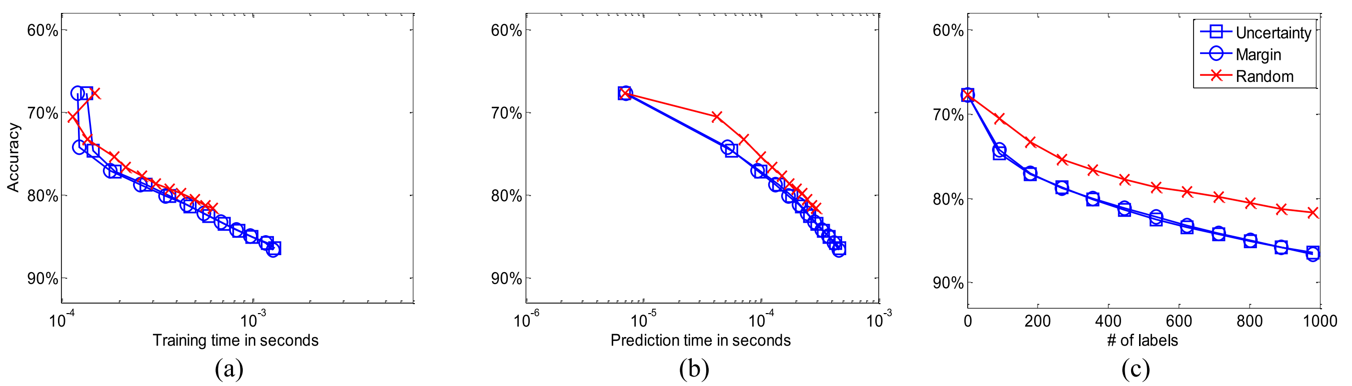

Fig. 8. Voltage stability prediction using SVMs: a) training time per OP, (b) testing time per OP, and (c) number of labeled OPs (c) on the x-axis; accuracy on the y-axis.

are labeled in Fig. 7 (b), while training time per OP increases slightly with an increase in knowledge base size in Fig. 7 (a).

Note the significant increase in training and prediction times per OP as the number of labeled OPs increases in Figs. 8 (a) and 8 (b). In Fig. 8 (c) the accuracy of uncertainty sampling closely matches that of margin sampling.

The SVM shows increases in prediction time as more OPs are added in Fig. 8 (b). However, even being an order of magnitude slower in testing, it produces competitive results, and is trained significantly more quickly than the ANN. In the transmission system experiment the ANN starts to train faster than the SVM after more than 500 OPs are included in $\mathbf{D}_{\mathrm{L}}$. The proposed sampling strategies outperform the baseline random sampling in each case.

Experiments revealed that it is not necessary to initialize the $\mathbf{D}_{\mathrm{L}}$ for SVM training with more than a single OP, because of the locality of the kernel used. The ANN however did not perform adequately with a single training OP, and therefore for comparison sake each initial $\mathbf{D}_{\mathrm{L}}$ contained one OP from each class. This ensured a realistic decision boundary for ANN models and results comparable across machine learning models.

The RF were not typically impacted beyond the training stage by training set size in Fig. 7 (b). Even though not well suited for the active learning approach because of the flat decision surfaces, the procedure as it is described here performs well. However, at 100 labeled OPs the margin sampling strategy performs significantly worse than random sampling, indicating a potential interplay of factors discussed in Section III-H.

While the active machine learning shows promising performance, it is worth noting that there are two possible drawbacks of the proposed approach. One drawback is that as a heuristic method it does not offer theoretical guarantees on improving performance in terms of accuracy or the number of OPs that need to be simulated in the time domain. Another drawback is that data which is intentionally misleading about the underlying uncertainty of the data generating process may adversely affect the estimates of uncertainty by the machine learning model resulting in poor predictive accuracy.

VI. Conclusion

The following conclusions can be reached:

- A weakness that the training data sets are always efficiently and sufficiently built, which is often overlooked in applying machine learning problems to power systems has been identified.
- The proposed pool-based active learning approach can build data sets for a machine learning model to train on more efficiently.
- The described approach enhances the existing machine learning models by identifying the operating points where model predictions contradict with reality, and adding labeled data sets around those points to the knowledge base.

- The approach also accelerates the offline training process by reducing the amount of model-based simulations around other operating points where correct predictions were made.
- The approach was employed to tackle voltage stability in transmission systems. Promising performance has been achieved.
- The improvements above the baseline random sampling are quantified in all relevant metrics. Results from the experiments analyzed each machine learning model within a comparable framework. Significant improvements have been observed and were discussed in Section VI.
- On average the most efficient sampling strategy across all experiments conducted was margin sampling, and the most accurate predictor the RF.

Future directions of this work may include the significant regression based applications in power systems which appear to be overlooked thus far in literature.

References

[1] T. Hong *et al.*, "Guest editorial big data analytics for grid modernization," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2395–2396, Sep. 2016.

[2] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2561–2570, Sep. 2016.

[3] A. Kusiak, H. Zheng, and Z. Song, "Short-term prediction of wind farm power: A data mining approach," *IEEE Trans. Energy Convers.*, vol. 24, no. 1, pp. 125–136, Mar. 2009.

[4] S. K. Tso *et al.*, "Data mining for detection of sensitive buses and influential buses in a power system subjected to disturbances," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 563–568, Feb. 2004.

[5] J. A. Steel, J. R. McDonald, and C. D'Arcy, "Knowledge discovery in databases: Applications in the electrical power engineering domain," in *Proc. Inst. Elect. Eng. Colloquium IT Strategies Inform. Overload*, Dec. 1997, pp. 1–4.

[6] B. D. Pitt and D. S. Kirschen, "Application of data mining techniques to load profiling," in *Proc. 21st IEEE Int. Conf. Power Ind. Comput. Appl.*, Santa Clara, CA, USA, 1999, pp. 131–136.

[7] L. A. Wehenkel, *Automatic Learning Techniques in Power System*. Norwell, MA, USA: Kluwer, 1998.

[8] S.-J. Huang and J.-M. Lin, "Enhancement of power system data debugging using GSA-based data-mining technique," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1022–1029, Nov. 2002.

[9] T. V. Cutsem, L. Wehenkel, M. Pavella, B. Heilbronn, and M. Goubin, "Decision tree approaches to voltage security assessment," *IEE Proc. C Gener. Transm. Distrib.*, vol. 140, no. 3, pp. 189–198, May 1993.

[10] I. Kamwa, S. R. Samantaray, and G. Joos, "Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 258–270, Feb. 2009.

[11] *Power Quality Monitoring Data Mining for Power System Load Modeling*, EPRI, Palo Alto, CA, USA, 2003.

[12] C. Zheng, V. Malbasa, and M. Kezunovic, "Regression tree for stability margin prediction using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1978–1987, May 2013.

[13] S. Madan, W. K. Son, and K. E. Bollinger, "Applications of data mining for power systems," in *Proc. Can. Conf. Elect. Comput. Eng. Eng. Innov. Voyage Disc.*, vol. 2. St. John's, NL, Canada, 1997, pp. 403–406.

[14] J. Morais *et al.*, "An overview of data mining techniques applied to power systems," in *Data Mining and Knowledge Discovery in Real Life Applications*, J. Ponce and A. Karahoca, Eds. Rijeka, Croatia: I-Tech Educ. Publ., 2009.

[15] A. Bose, "Power system stability: New opportunities for control," in *Stability and Control of Dynamical Systems With Applications*. New York, NY, USA: Springer, 2003, ch. 16.

[16] P.-C. Chen, Y. Dong, V. Malbasa, and M. Kezunovic, "Uncertainty of measurement error in intelligent electronic devices," in *Proc. IEEE/PES Gener. Meeting*, National Harbor, MD, USA, 2014, pp. 1–5.

[17] *IEEE Standard for Synchrophasor Measurements for Power Systems*, IEEE Standard C37.118.1-2011, 2011, pp. 1–61.

[18] A. Radovanovic, "Using the Internet in networking of synchronized phasor measurement units," *Int. J. Elect. Energy Syst.*, vol. 23, no. 3, pp. 245–250, 2001.

[19] N. Al-Messabi, C. Goh, and Y. Li, "Heuristic grey-box modelling for photovoltaic power systems," *Syst. Sci. Control Eng.*, vol. 4, no. 1, pp. 235–246, 2016.

[20] M. G. Doborjeh, G. Y. Wang, N. K. Kasabov, R. Kydd, and B. Russell, "A spiking neural network methodology and system for learning and comparative analysis of EEG data from healthy versus addiction treated versus addiction not treated subjects," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1830–1841, Sep. 2016.

[21] K. Thirumala, S. P. Maganuru, T. Jain, and A. Umarikar, "Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances," *IEEE Trans. Smart Grid*, to be published.

[22] S. Padrón, M. Hernández, and A. Falcón, "Reducing under-frequency load shedding in isolated power systems using neural networks. Gran Canaria: A case study," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 63–71, Jan. 2016.

[23] S. Samantaray, P. Achlerkar, and M. S. Manikandan, "Variational mode decomposition and decision tree based detection and classification of power quality disturbances in grid-connected distributed generation system," *IEEE Trans. Smart Grid*, to be published.

[24] S. Rajan, J. Ghosh, and M. M. Crawford, "An active learning approach to hyperspectral data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 4, pp. 1231–1242, Apr. 2008.

[25] S. C. H. Hoi, R. Jin, J. Zhu, and M. R. Lyu, "Batch mode active learning and its application to medical image classification," in *Proc. 23rd Int. Conf. Mach. Learn.*, Pittsburgh, PA, USA, 2006, pp. 417–424.

[26] S. C. H. Hoi, R. Jin, and M. R. Lyu, "Large-scale text categorization by batch mode active learning," *Proc. 15th Int. Conf. World Wide Web*, Edinburgh, U.K., 2006, pp. 633–642.

[27] E. Lughofer and M. Pratama, "On-line active learning in data stream regression using uncertainty sampling based on evolving generalized fuzzy models," *IEEE Trans. Fuzzy Syst.*, to be published.

[28] A. Holzinger, "Interactive machine learning for health informatics: When do we need the human-in-the-loop?" *Brain Informat.*, vol. 3, no. 2, pp. 119–131, 2016.

[29] M. Elahi, F. Ricci, and N. Rubens, "A survey of active learning in collaborative filtering recommender systems," *Comput. Sci. Rev.*, vol. 20, pp. 29–50, May 2016.

[30] I. Zliobaite, A. Bifet, B. Pfahringer, and G. Holmes, "Active learning with drifting streaming data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 1, pp. 27–39, Jan. 2014.

[31] P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[32] A. Bose, "Power system stability: New opportunities for control," in *Stability and Control of Dynamical Systems With Applications*. New York, NY, USA: Springer, 2003, ch. 16.

[33] T. V. Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Boston, MA, USA: Kluwer, 1998.

[34] MathWorks Documentation Center. (2014). *Neural Network Toolbox*. [Online]. Available: http://www.mathworks.com/help/nnet/

[35] C.-C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, Apr. 2011, Art. no. 27.

[36] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[37] PSS®E, *Power Transmission System Planning Software*, Siemens, Munich, Germany.

Vuk Malbasa, photograph and biography not available at the time of publication.

Ce Zheng, photograph and biography not available at the time of publication.

Po-Chen Chen, photograph and biography not available at the time of publication.

Tomo Popovic, photograph and biography not available at the time of publication.

Mladen Kezunovic, photograph and biography not available at the time of publication.