A Machine Learning Initializer for Newton-Raphson AC Power Flow Convergence

Samuel N. Okhuegbe
University of Tennessee, Knoxville, USA
sokhuegbe@vols.utk.edu

Adedasola A. Ademola
University of Tennessee, Knoxville, and Dominion Energy Virginia Glen Allen, USA
aademol1@vols.utk.edu

Yilu Liu
University of Tennessee, Knoxville, and Oak Ridge National Laboratory, Knoxville, USA
liu@utk.edu

Keywords: Machine Learning, Newton Raphson, Power flow convergence, Power flow initialization, Random Forest

I. INTRODUCTION

Power flow, sometimes called load flow, involves the computation of the bus voltages magnitudes and angles and it gives insight into the steady-state condition of the power system at a particular operating condition. Power flow solutions serve as a base in performing other power system studies such as transient and voltage stability studies. Ideally, a power flow case must be solvable under steady state before any dynamic study can be performed. Grid operators and planners regularly perform several power flow studies under various loading and generation dispatch to ensure reliable grid operation. From the power flow results, transmission planners can have a better understanding of line overloads, losses, and voltage violations in their network [1].

Several iterative and non-iterative methods currently exist for solving the non-linear power flow equations [2]. In practice however, the Newton-Raphson method and its variants stand out as the widely used solution method in industry because of its quadratic convergence property [1], [3]. Although the NewtonRaphson method is quite robust for solving power flow, it takes up significant computation time especially for large power grids with thousands of buses. Linear approximation methods like the DC power flow (DCPF) have been used to achieve faster solutions. DCPF requires several assumptions and may not converge to the true solutions as it does not consider critical parameters such as bus voltage magnitude, reactive power, and losses. Based on the limitations of DCPF, achieving a solved AC power flow (ACPF) case remains a high priority for transmission planners. Getting a converged ACPF can be particularly difficult, especially in large power grids. In some cases, the ACPF problem is unsolvable especially at high loading conditions [4]. Three major challenges exist in obtaining a converged/solved ACPF case for large power grids.

Firstly, the loading condition and generation dispatch affect the solvability of a power flow case. Loads are generally modelled as PQ buses with constant active and reactive power. There exists a maximum loading for each bus above which there would be no solution [4]. This can be further analyzed using the PV curves which determine the maximum power that can be transferred [5]. Loading a bus above this limit (which is the nose of the PV curve) would cause the voltage to drop drastically, which would further affect the convergence of the power flow case. The active power dispatch of all generators is specified in a power flow case except that of the slack generator which typically handles any mismatch between the load (including losses) and generation. Convergence problems would exist in cases where the extra power supplied or absorbed by the slack bus is excessively greater than the slack generator limits [4]. So, the non-slack generators need to be adequately dispatched to prevent unnecessary stress on the slack generator. The next major challenge is due to reactive power support. Reactive power does not travel far over long distances due to the inherently high X/R ratio in long transmission lines, thus reactive power support, if needed, must be provided locally. A power flow case may fail to converge if several buses have insufficient reactive power support. Several algorithms have been developed in [2], [4], [6], [7] to achieve power flow convergence through reactive power planning.

The third major challenge is the choice of initial voltage magnitude and angle guess when solving ACPF using Newton based methods. Previous research has shown that NewtonRaphson is very sensitive to initial conditions [8], [9]. Ref. [10] have shown that Newton-Raphson method would not converge to any solution if the initial guesses were outside the region of attraction to the solution. This shows that even a well dispatched ACPF case may not converge due to the choice of the initial voltage magnitude and angles. The ACPF convergence region or region of attraction are rather complex and fractal in nature. The convergence region also varies with loading conditions and large number of iterations are needed when the initial guesses are far from the convergence region [1]. Ref. [9] developed an analytical iterative method to determine initial guesses to achieve power flow convergence. The iterative method starts with an initial guess and the Jacobian matrix after the first iteration is used to calculate a convergence operator ρ. This convergence operator is related to the non-singularity of the Jacobian matrix. If the convergence operator is greater than 0.5 then an affine matrix is calculated to update the initial guess. Else, an attempt is made to solve the power flow case. This iterative method was applied to small power flow cases up to 118 buses. The technique has not been applied to large scale power grids under varying operating conditions. In practice, power system planners tend to initialize the voltage magnitude and angles using flat start, DCPF solutions, a reference case, or other solved cases; power flow tools developed in [2], [6], [7] used this approach. Machine learning offers a good approach to initialize Newton-Raphson power flow as it can map a dispatch case to the voltage solutions and once sufficiently trained, it requires little computational time to predict its solution.

To the knowledge of the authors in this paper, no attempt has been successfully made to use machine learning to achieve ACPF convergence in large scale power grids by accurately predicting the initial conditions. Although machine/deep learning has been regularly used to predict power flow solutions[11]–[17], no attempt has been made to solve previously unsolved power flow cases by predicting adequate initial bus voltage magnitude and angles. Perhaps the closest attempt was by researchers in [13] where previously solved ACPF test cases were initialized to achieve faster solution time. In [13], it was assumed that all dispatch cases that form the training and testing dataset converged, cases that failed to converge were not used.

This paper aims to address the third major power flow convergence issue highlighted earlier. Here we introduce a new way to achieve ACPF convergence in large power grids by estimating the initial voltage magnitude and angles using Random Forest machine learning algorithm. Previously nonconverging power flow cases were solved by providing better initial voltage magnitude and angle estimates within the convergence region. The machine learning initializer was applied to a real ERCOT 6102 bus power flow case under various operating conditions. The simulation was performed using Python and PSS/E.

II. OVERVIEW OF AC POWER FLOW FORMULATION

The power flow problem is formulated as non-linear algebraic equations that maps the bus active and reactive power injection to the voltage phasor as shown in (1) and (2) [7].

$$P_i^{inj} = V_i * \sum_{k=1}^{N} V_k (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik}) \quad (1)$$

$$Q = V_i * \sum_{k=1}^{N} V_k (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik}) \quad (2)$$

Where P_i^{inj} and Q_i^{inj} are the net active and reactive power injection. G_{ik} and B_{ik} are real and imaginary parts of the admittance matrix element $Y_{ik} \cdot \theta_{ik}$ represent the bus voltage angle difference between bus *i* and *k*. V_i and V_k are the bus voltage magnitude at bus *i* and *k* while *N* is the total number of buses. For the ERCOT case $N = 6102$. Although there is no guarantee that the machine learning algorithm would learn the actual power flow equations in (1) and (2), the machine learning performs satisfactorily in providing key mappings between the active and reactive powers and the bus voltage solutions.

III. DATASET GENERATION

In this paper, the dataset was generated using the generator and load dispatch obtained from the U.S Energy Information Administration (EIA) [18]. The data contains the hourly generation and loadings in ERCOT for the year 2022 and the first hour of 2023, so a total of 8761 hourly data. To generate the actual power flow case, a solved reference ERCOT case is required. The active power of the generators P_{gen} and loads P_{load} in the reference PSS/E case are then varied as specified in the EIA data. The conventional (hydro, gas, thermal), nuclear and renewable (solar, wind) generators in the PSSE reference case are scaled uniformly based on the EIA data for each generator resource category. The active power of the loads in the PSS/E case are scaled uniformly with the reactive power maintained at a constant P/Q ratio. At this stage 8761 power flow cases were generated corresponding to various loading conditions. The minimum load in the dispatch is 31.9 GW while the peak load is 79.8 GW. After modifying the reference case to create the dispatch, full Newton-Raphson power flow initialized with the reference case was applied to solve all cases as shown in Fig. 1. 4,862 power flow cases successfully converged while 3,899 power flow cases did not converge. The 4,862 solved cases would form the training and validation case. The 3,899 unsolved (non-converged) cases formed the testing case.

IV. PROPOSED FRAMEWORK AND MODEL SETUP

Random Forest (RF) is a widely used machine learning algorithm in power system applications. One of its key advantages is that it avoids overfitting and takes advantage of an ensemble of trees for improved accuracy. A RF regressor was trained to learn the mapping between the power injections and power flow solutions (voltage magnitude and angle). The input data are $P_{gen,i \ldots Ngen}$, $P_{load,i \ldots Nload}$ and $Q_{load,i \ldots Nload}$, where $P_{gen,i \ldots Ngen}$ is the active power of the generator at the bus *i* to the last generator at N_{gen} likewise, $P_{load,i \ldots Nload}$ and $Q_{load,i \ldots Nload}$ are the active and reactive power load at bus i to the last load at bus N_{load}.

The outputs are the bus voltage magnitude, bus angles, star bus voltages and star bus angles. Large scale grids regularly have three-winding transformers which are modelled in commercial software's like PSS/E and PowerWorld as three separate two-winding transformers connected to a fictitious star point [19], [20]. This star point or star bus has a calculated voltage magnitude and angle and is crucial for the effectiveness of the machine learning initializer.

The 4,862 solved cases were split into training and validation based on a 90/10 split. The main preprocessing done is the minmax normalization of the input data. Also, the bus/star bus angles were converted to radian to avoid large negative angles during training. The ERCOT power flow model has 6102 buses and 134-star buses (corresponding to 134 three-winding transformers). Fig. 3 shows the setup for training the RF model. Four RF models were trained for each output categories namely bus voltage magnitude, bus angle, star bus voltage magnitude and star bus angle. The same hyperparameters were used for all 4 RF models. Fig. 2 shows the entire machine learning initializer framework. After the models have been sufficiently trained to learn the mappings between the bus power injections and the voltage/angle solutions, the trained model was then applied to the previously 3,899 non-converging cases. The RF model then predicts the solutions of the bus/star bus voltage magnitude and angle. This solution was then used as the initial values of the 3,899 unsolved cases. This is based on the concept that although there is no guarantee that the machine learning model would predict the exact voltage magnitude and angle solution, the values would be close to the actual solution and therefore within the convergence region. Full Newton-Raphson power flow is then applied to this newly initialized power flow case.

V. RESULTS

A. Random Forest Initializer

This section looks at the performance of the RF models. Table I shows the accuracy of the RF model based on the Root Mean Square Error (RMSE). The RMSE is a suitable accuracy metric when dealing with a regression problem. RMSE is computed using (3), where n is the total data points, Y_i is the true value while $\hat{Y}_i$ is the predicted value.

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \qquad (3)$$

From Table I, RF provides a very small root mean square error although with significant training time. But once sufficiently trained the model takes only a few seconds to predict.

Regardless of the accuracy of the algorithm, the main task is to investigate the ability of the algorithm to assist in NewtonRaphson convergence. The RF models are applied to the 3,899 non-converging cases for which we do not know the actual solutions, and the predicted voltage magnitudes and angles were then used to initialize their respective power flow cases. 2,106 power flow dispatch cases successfully converged due to the better initialization provided by the RF models. This shows that the RF initializer was able to assist the Newton-Raphson solution method in converging 54% of the previous 3,899 nonconverging cases. The remaining 1,793 cases would need to be further investigated, because even though the problem of initialization was addressed in this work, non-convergence could be due to several other problems. Although the power flow cases successfully converged there exist certain voltage violations. From Table II a typical converged testing case has about 315 buses with voltage magnitude below 0.9pu or above 1.1pu. These voltage violations were reduced by adding switched shunts using the automated process described in [6]. After adding the switched shunts, the average number of voltage violations (1.1pu) dropped from 315 to about 9 buses per converged testing power flow case.

B. Comparison with Other Initialization Methods

The results obtained from RF were compared with other analytical and machine learning methods. Table III compares the performance of RF with other machine learning methods in terms of RMSE. The RF models provided better accuracy with the least RMSE, while linear regression had the worst RMSE.

Although the RF model had the best accuracy in terms of RMSE it required significant training time as it is an ensemble of many regression trees as shown in Table IV. Linear Regression (LR) on the other hand takes very little training time but offers poor accuracy. The Decision Tree (DT) had a relatively shorter training time when compared with RF, but its accuracy was not as good. Table V compares the convergences performance between the RF initializer and other initialization techniques. In practice, power system planners regularly initialize power flow cases using DCPF or flat start on rare occasions. In this work, none of the 3,899 unsolved power flow cases converged with flat start. It is difficult to achieve convergence from flat start for power flow cases with thousands of buses, although for small systems with a few hundred of buses, flat starting may be useful. However, initializing with DCPF which provides voltage angle estimates helped solve about 758 cases which represent about 19.44% of the 3,899 test cases. The process for initializing with DCPF is similar to that of the other machine learning methods. First the 3,899 cases were solved with DCPF with the voltage magnitude assumed as 1 p.u since DCPF only solves for the voltage angles. The voltage magnitude (1pu) and calculated angles are then used to initialize all 3,899 unsolved cases and a full Newton-Raphson ACPF was then applied.

From Table III and Table V, it can be observed that the lower the RMSE of the machine learning method, the higher the number of cases to converge. Amongst the machine learning methods, the RF algorithm had the least RMSE and best convergence rate, followed by DT and finally LR. The LR model performed poorly as it could not capture the relevant non-linear mappings. But even at that it solved about 246 cases. Fig. 4, shows overlapping and unique cases solved by different initialization methods. RF and DT had the most overlap due to the similarity in their algorithms. From Fig. 4a and Fig.4c, it can be observed that 1413 dispatch cases were solved by both RF and DT initialization methods.

Although both Linear Regression and DC power flow are linear initialization techniques, there is very little overlap in the cases solved by both methods as seen in Fig. 4b. Just 58 cases are solved by both LR and DCPF, even though LR solved a total of 246 cases and DCPF initialization solved 758 cases. The key inference from all this is that the initialization method used has an impact on what particular power flow case would be solved.

VI. CONCLUSION AND FUTURE WORK

In this paper, machine learning was used to predict the initial voltage/angle guesses to initialize Newton-Raphson power flow. The developed Random Forest initializer successfully converged 2,106 power flow cases which did not converge originally due to bad initialization. The RF initializer performed better when compared with popular analytical methods like DCPF initialization which is used in industry. After achieving a converged/solved power flow case, power system planners regularly analyze these cases for (voltage and thermal) violations and then propose adequate measures to solve these violations. In future, the capabilities of physics based deeplearning initializers need to be further investigated and compared with already established machine learning methods. In addition to this, retraining the model with more data and varying topology configurations could provide further insights and improve the success rate of the model.

REFERENCES

[1] J. J. Deng and H. D. Chiang, "Convergence region of Newton iterative power flow method: Numerical studies," J Appl Math, vol. 2013, 2013, doi: 10.1155/2013/509496.

[2] B. Wang and J. Tan, "DC-AC Tool: Fully Automating the Acquisition of the AC Power Flow Solution: National Renewable Energy Laboratory," 2022. [Online]. Available: www.nrel.gov/publications.

[3] S. Dutto, G. Masetti, S. Chiaradonna, and F. Di Giandomenico, "On Extending and Comparing Newton-Raphson Variants for Solving Power-Flow Equations," IEEE Transactions on Power Systems, vol. 34, no. 4, pp. 2577–2587, Jul. 2019, doi: 10.1109/TPWRS.2019.2897640.

[4] A. B. Birchfield, T. Xu, and T. J. Overbye, "Power flow convergence and reactive power planning in the creation of large synthetic grids," IEEE Transactions on Power Systems, vol. 33, no. 6, pp. 6667–6674, Nov. 2018, doi: 10.1109/TPWRS.2018.2813525.

[5] Michael Starke, Fangxing Li, Leon M. Tolbert, and Burak Ozpineci, "AC vs DC Distribution Maximum Transfer Capability," in 2008 IEEE Power and Energy Society General Meeting, 2008.

[6] B. Vyakaranam, Q. H. Nguyen, T. B. Nguyen, N. A. Samaan, and R. Huang, "Automated Tool to Create Chronological AC Power Flow Cases for Large Interconnected Systems," IEEE Open Access Journal of Power and Energy, vol. 8, pp. 166–174, 2021, doi: 10.1109/OAJPE.2021.3075659.

[7] A. A. Ademola, S. You, and Y. Liu, "Automated Two-stage Conversion of Hourly Optimal DC Flow Solution to AC Power Flow," CSEE Journal of Power and Energy Systems, vol. 9, no. 1, pp. 28–36, Jan. 2023, doi: 10.17775/CSEEJPES.2021.08870.

[8] F. Casella and B. Bachmann, "On the choice of initial guesses for the Newton-Raphson algorithm," Appl Math Comput, vol. 398, Jun. 2021, doi: 10.1016/j.amc.2021.125991.

[9] Q. Sun, L. Liu, D. Ma, and H. Zhang, "The initial guess estimation Newton method for power flow in distribution systems," IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 2, pp. 231–242, Apr. 2017, doi: 10.1109/JAS.2017.7510514.

[10] Mariesa L Crow, Computational Methods for Electric Power Systems, Second Edition. Taylor & Francis, 2010. [Online]. Available: www.EngineeringEbooksPdf.com

[11] B. Huang and J. Wang, "Applications of Physics-Informed Neural Networks in Power Systems - A Review," IEEE Transactions on Power Systems, vol. 38, no. 1, pp. 572–588, Jan. 2023, doi: 10.1109/TPWRS.2022.3162473.

[12] A. B. Jeddi and A. Shafieezadeh, "A Physics-Informed Graph Attention-based Approach for Power Flow Analysis," in Proceedings - 20th IEEE International Conference on Machine Learning and Applications, ICMLA 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1634–1640. doi: 10.1109/ICMLA52953.2021.00261.

[13] L. Chen and J. E. Tate, "Hot-Starting the Ac Power Flow with Convolutional Neural Networks," Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.09342

[14] K. Baker, "Learning Warm-Start Points for AC Optimal Power Flow," in 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), 2019.

[15] X. Hu, H. Hu, S. Verma, and Z. L. Zhang, “Physics-Guided Deep Neural Networks for Power Flow Analysis,” IEEE Transactions on Power Systems, vol. 36, no. 3, pp. 2082–2092, May 2021, doi: 10.1109/TPWRS.2020.3029557.

[16] K. Chen and Y. Zhang, “Physics-guided Residual Learning for Probabilistic Power Flow Analysis,” Jan. 2023, [Online]. Available: http://arxiv.org/abs/2301.12062

[17] B. Vyakaranam et al., “Novel Data-Driven Distributed Learning Framework for Solving AC Power Flow for Large Interconnected Systems,” IEEE Open Access Journal of Power and Energy, vol. 8, pp. 281–292, 2021, doi: 10.1109/OAJPE.2021.3092264.

[18] EIA, “Hourly Electric Grid Monitor,” EIA. Accessed: Jun. 05, 2023. [Online]. Available: https://www.eia.gov/electricity/gridmonitor/dashboard/electric_over view/regional/REG-TEX

[19] PowerWorld, “Three Winding Transformer Information.” Accessed: Sep. 09, 2023. [Online]. Available: https://www.powerworld.com/WebHelp/Content/MainDocumentati on_HTML/Three_Winding_Transformer_Information.htm

[20] Siemens, Program Operation Manual - PSS®E 34.8.1. Siemens, 2020.

TABLES AND FIGURES

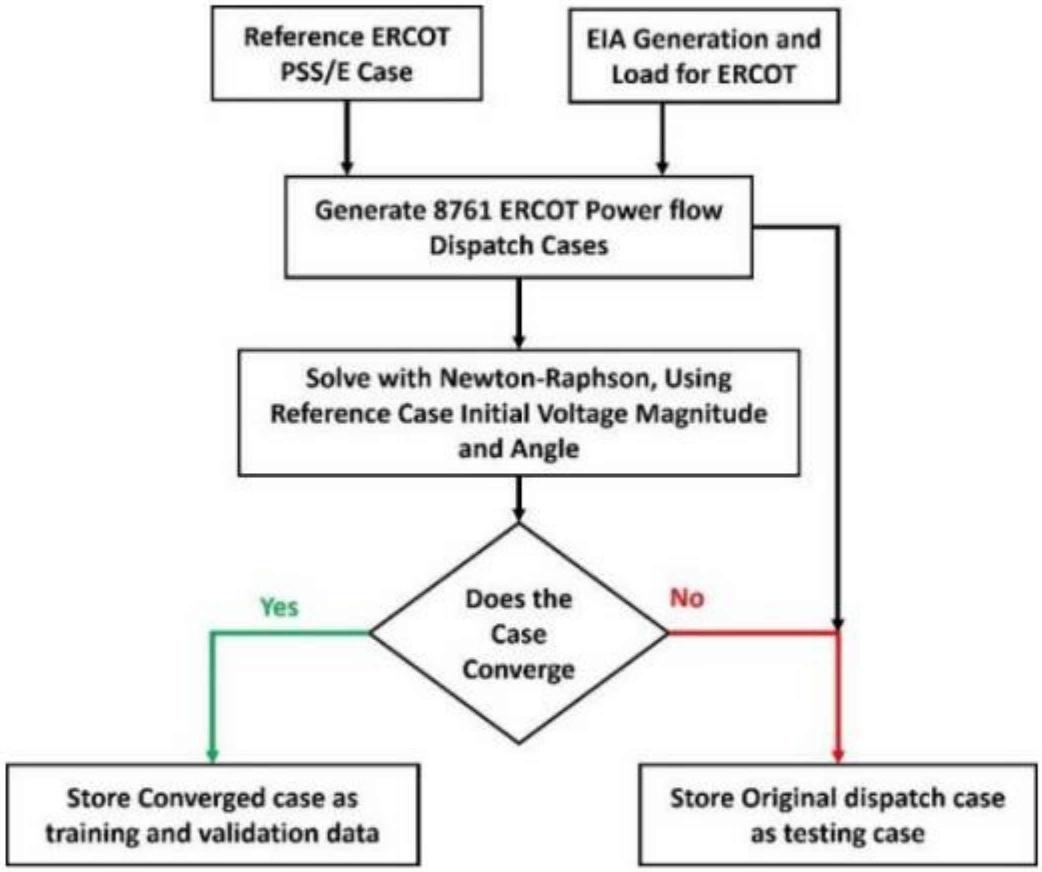

Fig. 1 Generation of 8761 Hourly Dispatch Power Flow Cases

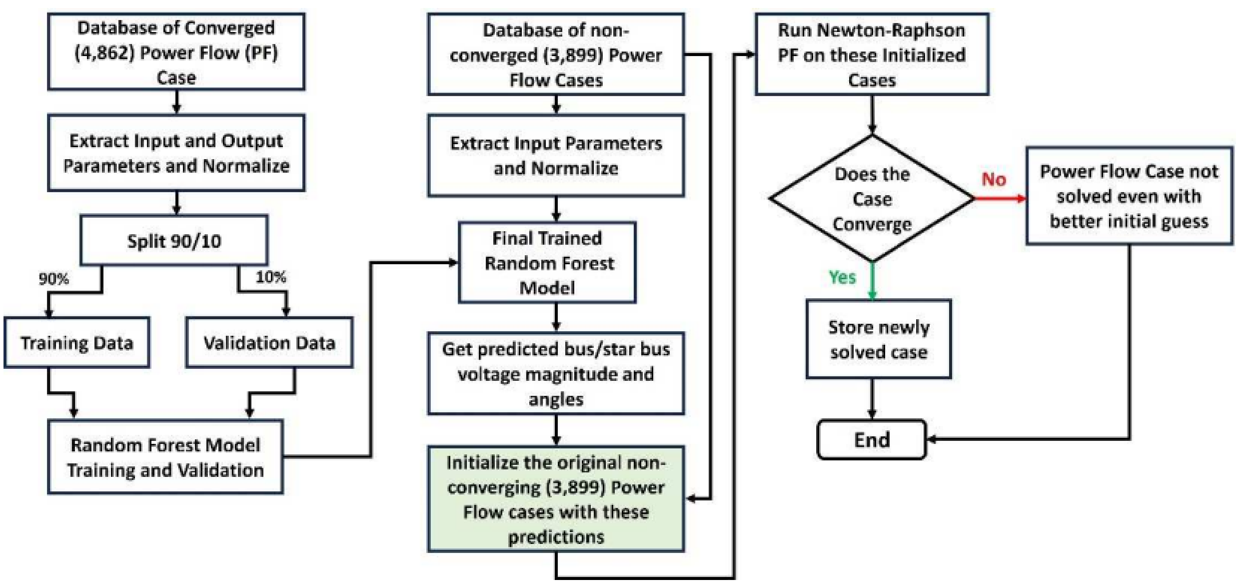

Fig. 2 Proposed Framework for Machine Learning Initializer

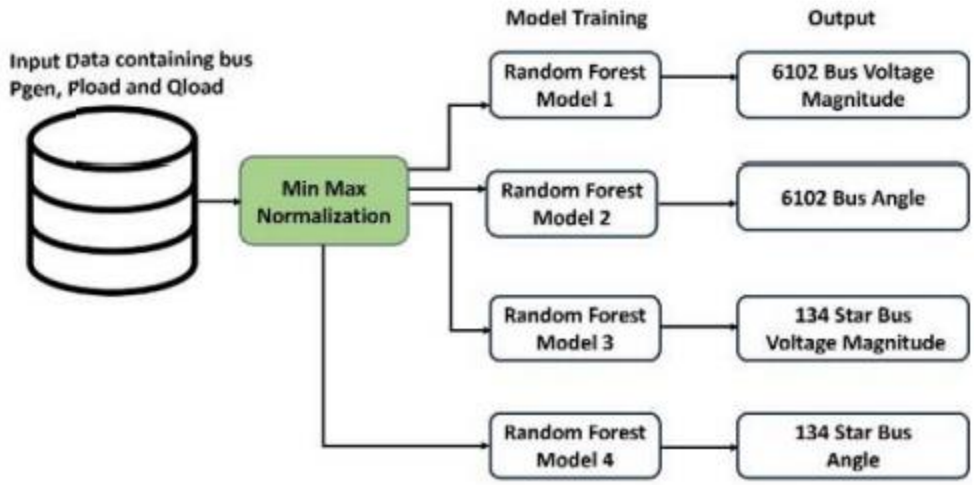

Fig. 3 Model Training Setup

Random Forest Model	***RMSE on Validation Data***	***Training Time***
Bus Voltage Magnitude	0.01629	15.16 Hours
Bus Voltage Angle	0.0138	14.15 Hours
Star Bus Voltage Magnitude	0.0086	23.89 Minutes
Star Bus Voltage Angle	0.01024	23.06 Minutes

TABLE I. PERFOMANCE OF RANDOM FOREST

Voltage Magnitude Ranges (p.u)	***Average No. of Violating Buses per Converged Testing Case (Without Extra Voltage Support)***	***Average No. of Violating Buses per Converged Testing Case (With Automated Voltage Support)***
<0.90 or > 1.1	315	9
<0.85 or >1.12	192	1
<0.80 or > 1.2	144	0

TABLE II. SOLVED CASES WITH VOLTAGE VIOLATING LIMITS

Model	*Training Time Random Forest*	*Training Time Decision Trees*	*Training Time Linear Regression*
Bus Voltage Magnitude	15.16 Hours	1.6 Hours	2.267 Minutes
Bus Voltage Angle	14.15 Hours	1.5 Hours	2.59 Minutes
Star Bus Voltage Magnitude	23.89 Minutes	5.048 Minutes	24.21 Seconds
Star Bus Voltage Angle	23.06 Minutes	4.908 Minutes	23.19 Seconds

TABLE IV. TRAINING COMPARISON WITH DIFFERENT MACHINE LEARNING ALGORITHMS

Parameter	*RF Initializer*	*DT Initializer*	*DCPF Initializer*	*LR Initializer*
Total (Initial Non-Converged Power Flow Cases)	3,899 Cases	3,899 Cases	3,899 Cases	3,899 Cases
Power Flow Cases Converged by Initialization	2,106 Cases	1,783 Cases	758 Cases	246 Cases
Percentage (%) of Cases Solved by Initialization	54.01%	45.73%	19.44%	6.31%
Remaining Non-Converged Power Flow Dispatch Cases	1,793 Cases	2,116 Cases	3,141 Cases	3,653 Cases

TABLE V. CONVERGENCE PERFORMANCE COMPARISON

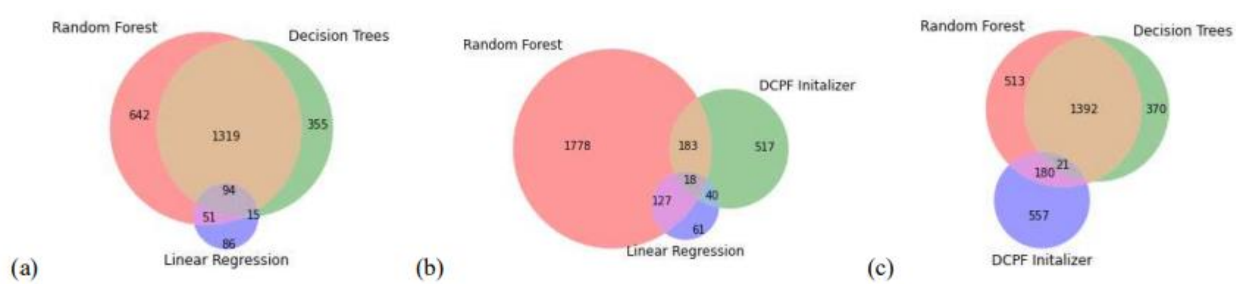

Fig. 4 Overlap in Power Flow Cases Solved by different Initialization Methods