Preprints of the 4th IEEE Conference on Energy Internet and Energy System Integration
October 30- November 1, 2020. Wuhan, China

Explainable AI in Deep Reinforcement Learning Models: A SHAP Method Applied in Power System Emergency Control

Ke Zhang
Electrical engineering and automation
Wuhan university
Wuhan, China
zhangke@whu.edu.cn

Peidong Xu
Electrical engineering and automation
Wuhan university
Wuhan, China
xupd@whu.edu.cn

Jun Zhang
Electrical engineering and automation
Wuhan university
Wuhan, China
jun.zhang.ee@whu.edu.cn

2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2) | 978-1-7281-9606-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/EI250167.2020.9347147

***Abstract*—The application of artificial intelligence (AI) system is more and more extensive, using the explainable AI (XAI) technology to explain why machine learning (ML) models make certain predictions as important as the accuracy of the predictions, because it ensures the trust and transparency in the model decision-making process. For deep reinforcement learning (DRL) model, although some outstanding progress based on DRL has been made in many fields, it is difficult to explain and cannot be used in safety related occasions. Especially in power system, for the power system emergency control based on DRL, how to provide an intuitive and reliable XAI technology is urgent and necessary. The Shapley additive explanations (SHAP) method has been adopted to provide a reasonable interpretable model for an open-source platform named Reinforcement Learning for Grid Control (RLGC). Through a series of summary plots, force plots and probability of SHAP value, the under-voltage load shedding of power system based on DRL can be interpreted much easier and clearer. More importantly, this work is unique in the power system field, presenting the first use of the SHAP method and the probability of SHAP value to give explanations for emergency control based on DRL in power system.**

***Index Terms*—Deep reinforcement learning, explainable artificial intelligence, power system, under-voltage load shedding, Shapley additive explanations**

I. Introduction

Due to the potential of AI in improving the efficiency, consistency and accuracy of decision-making, it is being widely deployed in many application fields. Although AI algorithms seem to be very powerful in prediction and become more and more popular, most of these models are regarded as black boxes for users, which are opaque and difficult to understand their internal working mechanisms.

In the field of power system, during the past few years, DRL has been used in different applications of reliable and secure power system operation.In [1], DRL was applied to determine generation unit tripping in case of emergency circumstances. In [2], DRL was used to develop a dynamic load shedding scheme for short-term voltage control.In [3] DRL method was adopted to utilize the operation of storage devices in micro-grids.It is worth noting that although DRL models has have prediction and adaptation ability, it is at the cost of its low interpretability.

AI systems are operationally opaque, however, in many applications human operators need to finalize decisions, and require that AI's recommendations can be self-explained. Nowadays, even experts are difficult to interpret complex AI models, such as integrated models or deep learning models. There is a great gap between accuracy and interpretability. In the presence of this challenge, XAI is proposed and utilized to make AI system more transparent.

Nowadays, there are many interpretable models, as shown in Fig. 1, most of which fall into local and agnostic models. The commonly used agnostic models are mainly visualization methods, knowledge extraction, influence methods and example-based explanations. In order to understand ML models, visualization methods [4], [5] and [6] are often adopted, a natural idea is to visualize and explore the contents hidden in neural units. The method of knowledge extraction in agnostic model is popular at present [7]. Model distillation [8] and [9] in knowledge extraction is a hot topic in the field of AI for complex models. Meanwhile, more and more scholars propose to use knowledge graph [10], as a means of knowledge extraction, to achieve interpretability in ML. Influence method is the most widely used in agnostic models method, which estimates the importance or relevance of features by changing inputs or internal components and documenting the impact of changes on model performance [11]. Additionally, example-based and mixed explanations [12] are also applied in ML.

The Shapley value method applied in this paper belongs to the feature importance method of agnostic models. The Shapley value method calculate the importance value of each feature [13]. Feature importance method quantifies the contribution of each input variable, or feature, to the functionality of a complex ML model. In order to measure the importance of features, the increase of model error is calculated after feature replacement. The replacement of important features increases model error.

At present, Shapley value method has been used in a number of cases to calculate the importance of features in ML model, so as to get an intuitive explanation. In [14] a sampling method for estimating the Shapley values is proposed, stratified sampling method has been used to

This work is supported by the National Key R&D Program of China under Grant 2018AAA0101504 and Science and technology project of SGCC (State Grid Corporation of China): fundamental theory of human-in-the-loop hybrid-augmented intelligence for power grid dispatch and control.

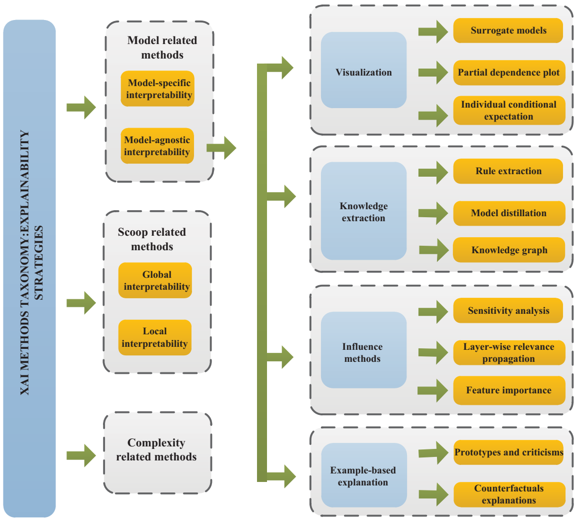


Fig. 1. Categories of XAI

reduce the number of samples of Shapley value method. In [15] a dynamic pricing mechanism of cross-node on-demand bandwidth based on Shapley value has been proposed. The interpretation ability of Shapley value are used to assign a specific predicted important value to each feature.

In recent years, a SHAP method based on Shapley value method has been proposed by [16]. SHAP method can be used for both global explanations and local explanations, which can be adopted to explain individual prediction. At present, SHAP has been utilized in various applications. The discovery of important biomarkers in [17] is helpful for accurate diagnosis and prediction of certain cancer types, which uses gradient enhancement trees and SHAP method. In [16], a tree explainer based on SHAP is proposed, which can obtain better computational performance than Shapley value model, and is a tool for explaining the global model structure based on local interpretations.

Interpretability is always a critical weakness of DRL models, and it is believed that the interpretability technology can help break through this bottleneck, however, few articles investigate the DRL interpretability issue, and how to implement the interpretability technology in DRL models is also necessary and urgent.

A SHAP method has been proposed in this paper to realize the interpretability of DRL model. In order to verify the interpretable performance of DRL models, an adaptive emergency control scheme of power system based on DRL has been put forward as the research background [18]. [18] designed an open-source RLGC platform. The platform is jointly built by Google team and the Pacific Northwest National Laboratory (PNNL) to assist in the development and benchmarking of DRL algorithm for power system control. This paper aims to utilize the SHAP method on the under-voltage load shedding of power system based on DRL.

The rest of the paper is organized as follows. The introduction of XAI and SHAP method are presented in Section II. Section III details the open platform for developing and bench-marking DRL algorithms for grid control. The test cases and results are shown in Section IV. The interpretable conclusions and future work are provided in Section V.

II. SHAP METHOD IN DRL

SHAP is an additive feature attribution methods which is represented by the explanations of Shapley value. In SHAP, Shapley value method is associated with the local interpretable model-agnostic explanations (LIME). At present, many explanation methods of ML models for local prediction belong to additive feature attribution methods, such as LIME, deep learning important features (DeepLIFT), layer-wise relevance propagation, Shapley regression values, Shapley sampling values and quantitative Input Influence. All of the additive feature attribution methods can be unified by SHAP.

Original DRL models cannot be used as their own explanation. Instead, we must use a simple and intuitive explanation model, as explanation approximation. The outcome of the model is interpreted by SHAP as the sum of the attributed values of each input feature.

There are several explainers in SHAP. tree explainer is a fast and exact method to estimate SHAP values for tree models and ensembles of trees, under several different possible assumptions about feature dependence. The gradient explainer can explain a model using expected gradients. The kernel explainer is a method that uses a special weighted linear regression to compute the importance of each feature. The deep explainer is an enhanced version of the DeepLIFT algorithm which approximates the conditional expectations of SHAP values using a selection of background samples. And the sampling explainer is an extension of the Shapley sampling values explanation method. In this paper, the deep explainer is chosen for DRL interpretation.

The exact Shapley value should be estimated by all possible feature alliances with and without feature xi. However, when the number of features is large, the number of possible alliances will increase exponentially. Therefore, for DRL, the gradient descent method is used in SHAP as the approximate method.

The influence of state features on output actions can be expressed as

$$\nabla y = [\frac{\partial y}{\partial x_1} \; \frac{\partial y}{\partial x_2} \cdots \frac{\partial y}{\partial x_i} \cdots \frac{\partial y}{\partial x_m}]^T \tag{1}$$

where y is the output of DRL model, m is the number of features in data sets. Calculate the difference between the ith feature vector and the data sets. The change of features is expressed as

$$\Delta x = [\Delta x_1 \; \Delta x_2 \cdots \Delta x_i \cdots \Delta x_m]^T \tag{2}$$

The average marginal contribution value of each feature to the output action in the data sets can be expressed as

$$\overline{\Delta \mathrm{y}} = \frac{1}{n}\sum_{k=0}^{n}[\frac{\partial y}{\partial x_1}\Delta x_1 \; \frac{\partial y}{\partial x_2}\Delta x_2 \cdots \frac{\partial y}{\partial x_i}\Delta x_i \cdots \frac{\partial y}{\partial x_m}\Delta x_m]^T \tag{3}$$

where n is the number of feature vectors in data sets.

The XAI algorithm framework based on SHAP method is showed in Fig. 2 with four parts, including the input of SHAP, core algorithm of SHAP, output of SHAP and tasks of SHAP. The inputs of SHAP have three parts, the first part is

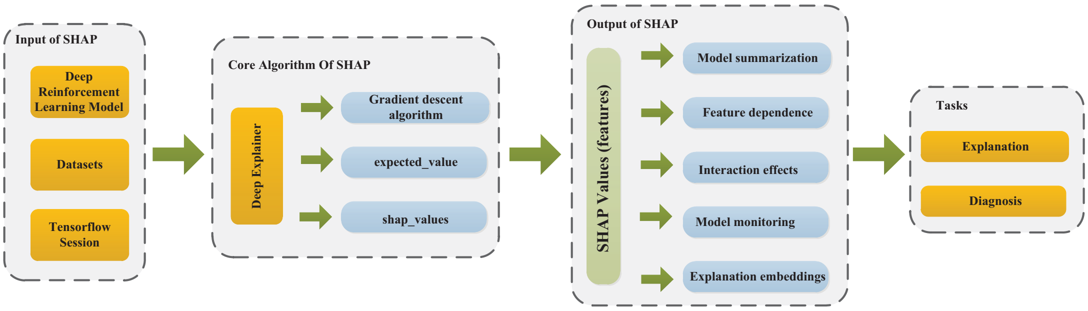


Fig. 2. XAI algorithm framework based on SHAP method

the DRL model, including the observations and actions stored in each step of the model during its training process. Second one is the data sets. The last one is Tensor Flow session, which is used to provide the network architecture of model. The core algorithm of SHAP, deep explainer, is adopted for DRL model. In the deep explainer, gradient descent algorithm is used to calculate the influence of features. Moreover, expected value is the mean value of an action under all the data sets, shap value is the actual contribution value of the action under the current data samples. SHAP values of features are the output of XAI, which can be applied in many aspects, such as model summarization, feature dependence, interaction effects, model monitoring and explanation embedding. Generally speaking, the tasks of XAI is explanation and diagnosis.

III. Power System Emergency Control

An open-source platform named Reinforcement Learning for Grid Control (RLGC) is adopted to verity the performance of SHAP method in DRL models. RLGC is jointly built by Google team and the Pacific Northwest National Laboratory (PNNL).

Power system emergency control is the final defense line to ensure the safety and recovery of power grid. Fault-induced delayed voltage recovery refers to the phenomenon that the system voltage remains at a significantly reduced level for several seconds after the fault is cleared. In this way, under-voltage load shedding was introduced to solve the problem of fault-induced delayed voltage recovery.

An XAI framework based on RLGC is shown in Fig. 3 The test is on IEEE 39-bus system, and buses of 4, 7, and 18 are heavy load area. The causes of fault is considered as stalling and tripping of an air conditioning motor. After fault clearance, the standard requires that voltages should return to at least 0.8, 0.9 and 0.95 p.u. within 0.33 s, 0.5 s and 1.5 s, respectively [18].

The structure of neural network in Fig. 3 is fully connected network. Ni is the input layer of neural network, including 83 observations. The observations include the voltage magnitudes of buses 4, 7, 8 and 18 and low-voltage sides of the step-down transformers connected with them in 10 observation times. What is more, the percentage of load of buses 4, 7, and 18 at current time are also belong to observations.

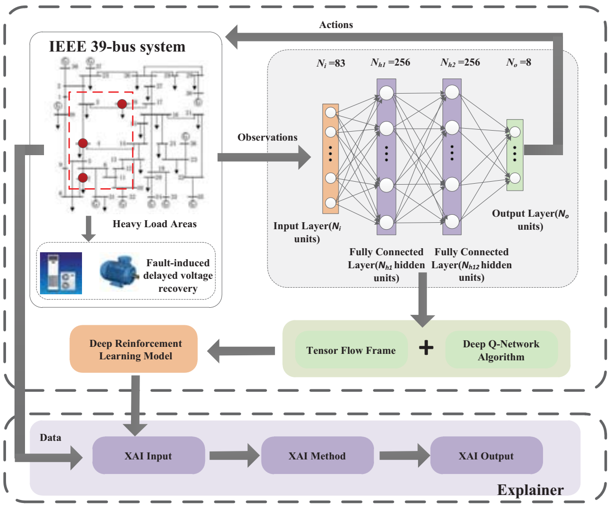


Fig. 3. XAI framework based on RLGC

N_o is the output layer of neural network, including 8 actions. The control actions for buses 4, 7, and 18 at each action time step include no load shedding and 20% load shedding. The above actions in total will be used to control the way of under voltage load shedding in IEEE 39-bus system.

There are two hidden layers N_{h1} and N_{h2}, each with 256 units. The algorithm for DRL implementation is the deep Q-network (DQN) algorithm. Finally, both of the trained DRL model and data from IEEE 39-bus system will be utilized as the input of XAI.

IV. Validation and Analysis

The results in this section validate that the under-voltage load shedding model based on DRL in power system emergency control can be well explained with the SHAP method. The explanation of the model can be presented in both global interpretability and local interpretability.

A. Global interpretability

The SHAP method can realize the global explanation for the whole model, the influence of all input features and the results of all output actions in the model can be computed and presented. To obtain an overview of which features are most important for the model, one can plot the SHAP values

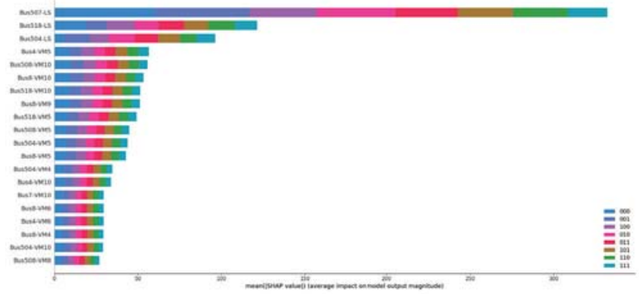


Fig. 4. Rich summaries of entire model based on SHAP

of every feature for every sample. Fig. 4 is the summary plot, which sorts features by the summation values of SHAP value magnitudes over all samples, and uses SHAP values to show the distribution of the impacts each feature has on the model output. In Fig. 4, horizontal axis is the average impact on output magnitude. Vertical axis is the features, it shows the top 20 of 83 features. The eight colors in Fig. 4 represent eight output actions respectively. The meaning of symbols in the Fig. 4 is explained in table I in detail.

TABLE I
SYMBOL INTERPRETATION IN FIG. 4

Symbol	Meaning
Bus$i$-VMj	Represents the observed value of voltage magnitude in high-voltage side for the ith bus at jth moment
Bus5i-VMj	Represents the observed value of voltage magnitude in low-voltage side for the ith bus at jth moment
Bus5i-LS	Represents the observed value of the percentage load at the ith bus
Action “000”	Means no load reduction
Action “001”	Means 20% load reduction of bus 4
Action “010”	Means 20% load reduction of bus 7
Action “011”	Means 20% load reduction of both buses 4 and 7
Action “100”	Means 20% load reduction of bus 18
Action “101”	Means 20% load reduction of both buses 4 and 18
Action “111”	Means 20% load reduction of both buses 4, 7 and 18

The total number of interaction steps in training is 1200000, 1000 as the random input of deep explainer. The result is 945 times of output action “000”, 6 times of “001”, 49 times of “010”. That means, in most cases, the voltage magnitudes of the emergence control system can be recovered within the specified time. When the voltage magnitudes cannot be recovered by itself, the most cost-effective operation will be choose by the trained model.

For the output actions of the system, the loads on buses 7, 18 and 4 have the greatest influence, especially the loads on bus 7. Next is the impact of voltage magnitudes. After observing the entire samples, different output actions can be analyzed respectively.

B. Local interpretability

The data sets of different output actions are shown in Fig. 5 The mean absolute values of the SHAP for each feature are in Fig. 5(a), which is a bar chart of the average SHAP value magnitude. The loads of bus 7 has the greatest influence for the case of no load reduction and then are the loads of buses 18 and 4.

Fig. 5(b) is a set of bee swarm plots, where each dot corresponds to an individual data in the study. The color represents the feature value, red color means high and the blue color means low. The data sets of 20% load reduction of bus 4 are shown in Fig. 5(c) and Fig. 5(d). The load of bus 7 has the greatest influence for the case of 20% load reduction of bus 4, followed by the loads of buses 18 and 4. And the voltage magnitudes at low voltage side of buses of 7, 8 and 18 are concentrated in the area with lower value, and has a negative impact on the current output. And the voltage magnitudes at high voltage side of buses of 4, 7 and 8 are concentrated in the area with lower value, and has a positive impact on the current output. Similarly, the data sets of 20% load reduction for bus 7 are shown in Fig. 5(e) and Fig. 5(f)

Fig. 6(a) is an example in the case of no load reduction, it shows the contribution of each feature for the output. Output value in the figure is cumulative return value of certain action. Features promoting the result are shown in red, features in blue are hindering the result.The observed values of voltage magnitude in high-voltage side for the bus 18 at the $5th$ and $10th$ moment are 1.01101 and 1.01103 respectivly, which increase output value and promote current outcome. The value of load at bus 7 is 1.0, which reduces current prediction. That means, under the case of no load reduction, whether voltage magnitudes can be restored to the standard value is the main factor.

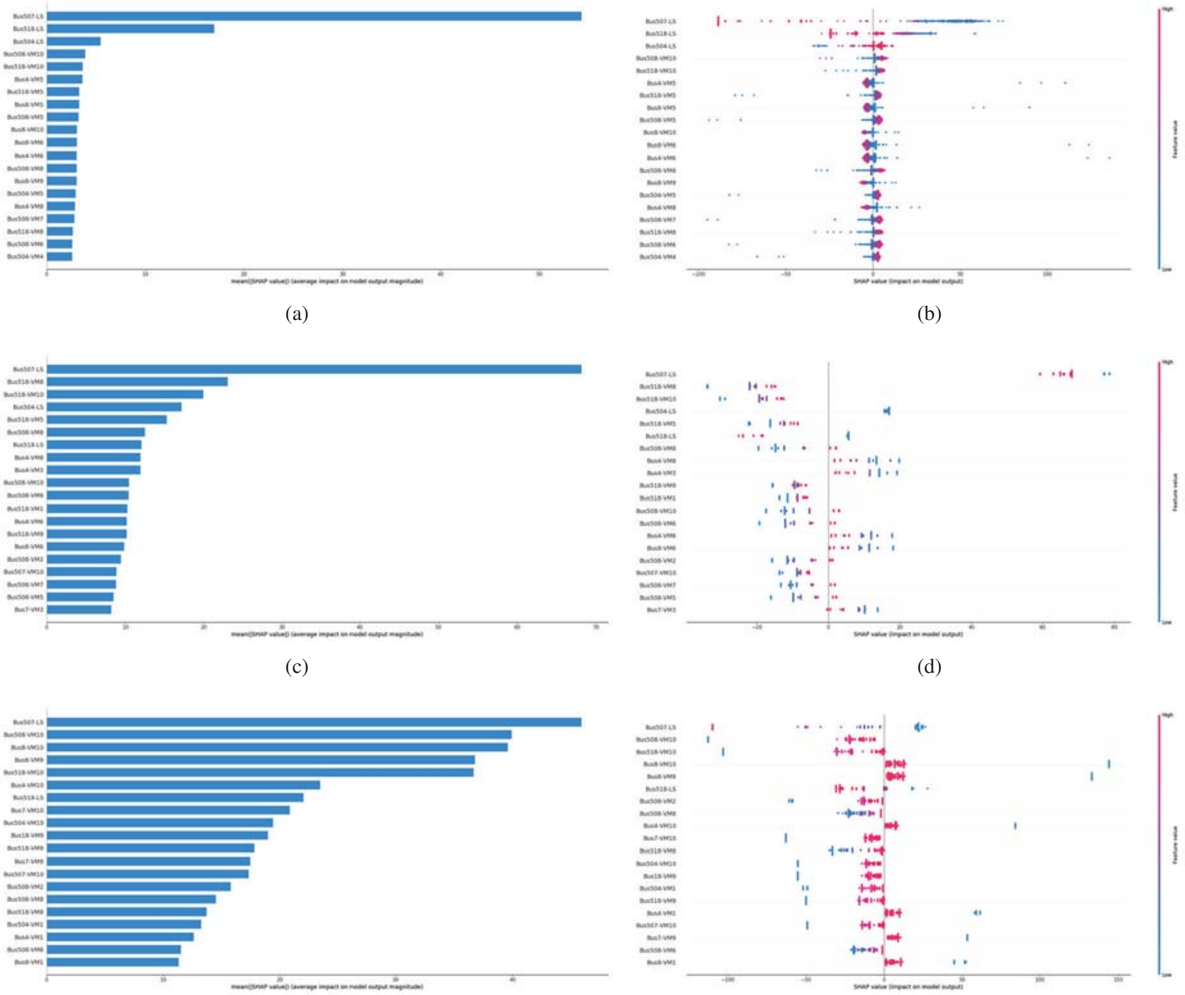
Fig. 6(b) is an example in the case of 20% load reduction for bus 4. The load of bus 7 is 0.2 and the load of bus 4 is 1.0, which increase output value and good for current prediction. And the observed value of voltage magnitude in high-voltage side for the bus 18 at the $8th$ moment is 0.90254, these data hinder current outcome. It shows that the voltage magnitude at high voltage side of bus 18 does not reach 0.95 p.u, but it may recover in a certain period of time.

Similarly, Fig. 6(c) is an example in the case of 20% load reduction for bus 7, the values of load at buses 7 and 4 are 0.4 and 1.0 respectively, which increase output value and promote current outcome. And the figure shows the voltage magnitude at high voltage side of bus 18 does not reach 0.95 p.u, but it may recover in a certain period of time.

The above analysis can be used as an important basis for decision-making of human beings. XAI has a great contribution to human-computer interaction. The interpretability of DRL is an implicit solution to understand human-computer interaction. In the process of actual scheduling, although ML can help human to give prediction results, the final decision must be made by human beings. The interpretation of XAI provides the additional information and confidence for human beings, so that human beings can act wisely and decisively, and then make better decision and deployment.

V. CONCLUSION

AI has become an important means of decision making, and reliability and transparency of AI systems are with great demand. For the power system emergency control based on DRL, one urgent need is to provide intuitive and reliable XAI technologies is urgent and necessary. In this paper, the SHAP value method has been adopted to provide a reasonable interpretable model for RLGC. The reasons why DRL model make certain actions under three different outputs have been



(a) (b) (c) (d) (e) (f)

Fig. 5. (a) Bar chart of the average SHAP value magnitude for no load reduction. (b) Bee swarm plot of SHAP value magnitude for no load reduction. (c) Bar chart of the average SHAP value magnitude for 20% load reduction of bus 4. (d) Bee swarm plot of SHAP value magnitude for 20% load reduction of bus 4. (e) Bar chart of the average SHAP value magnitude for 20% load reduction of bus 7. (f) Bee swarm plot of SHAP value magnitude for 20% load reduction of bus 7.

analyzed. For the first time, the application of SHAP in DRL model of power system is realized.

The research on XAI technologies in power system is still at its beginning stage. There are some key points for future research, such as how to improve the performance of ML models through XAI, and how to obtain a better interpretation effect through heterogeneous information network (HIN). Additionally, human-computer interaction in power system are also the main directions of future work:

1) Improve ML models with XAI In the process of model learning, DRL models are easy to fall into the problem of local optimization, which can be found by the help of XAI. Through XAI technology, intelligence of human beings is added to ML models, and the verification from human will improve the performance of the model.

2) Heterogeneous information network The heterogeneous information network provides an effective solution to achieve the interpretation in DRL, especially in the field of power system. Building a HIN model for power grid, through the research of entity recognition, entity relationship extraction and representation learning, makes the interpretation of DRL models more intuitive.

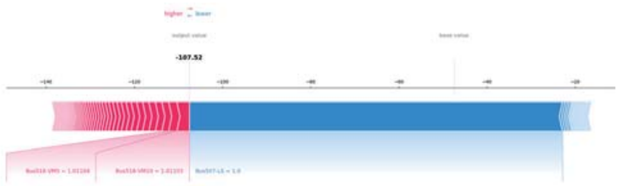
3) Human-computer interaction The interpretability of DRL is an implicit solution to understand human-computer interaction. ML can make decisions through multiple curves, and its cognitive ability is stronger than human beings. The interpretation of XAI provides the additional information and confidence for human beings, so that human beings can act wisely and decisively, and then make better decision and deployment. In the process of actual scheduling, although ML can help human to give prediction results, the final decision must be made by human beings. XAI technology makes human decision-making more secure and reliable.

References

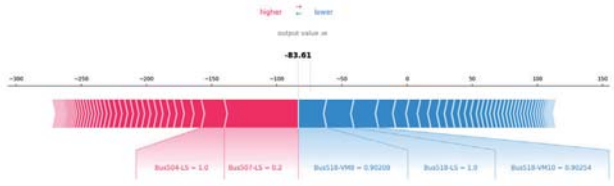
[1] L. Wei, Z. Dongxia, W. Xinying, H. Jinxiu, and L. Liping, "A decision making strategy for generating unit tripping under emergency circumstances based on deep reinforcement learning," *Proc CSEE*, vol. 38(1), p. 109–119, 2018.

[2] J. Zhang, C. Lu, J. Si, J. Song, and Y. Su, "Deep reinforcement leaming for short-term voltage control by dynamic load shedding in china southem power grid," 07 2018, pp. 1–8.

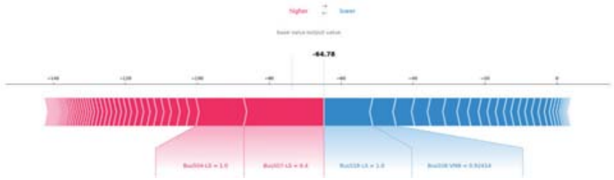
[3] N. Tomin, A. Zhukov, and A. Domyshev, "Deep reinforcement learning for energy microgrids management considering flexible energy sources," *EPJ Web of Conferences*, vol. 217, p. 01016, 01 2019.



(a)



(b)



(c)

Fig. 6. (a) Influence of features in individual example for the case of no load reduction; (b) Influence of features in individual example for the case of 20% load reduction of bus 4; (c) Influence of features in individual example for the case of 20% load reduction of bus 7.

[4] M. Ribeiro, S. Singh, and C. Guestrin, “”why should i trust you?” explaining the predictions of any classifier,” *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144, 2016.

[5] T. Xu, J. Zhan, O. G. B. Garrod, P. H. S. Torr, S. Zhu, R. A. A. Ince, and P. G. Schyns, “Deeper interpretability of deep networks,” *CoRR*, vol. abs/1811.07807, 2018. [Online]. Available:

[15] W. Shi, C. Wu, and Z. Li, “A shapley-value mechanism for bandwidth

http://arxiv.org/abs/1811.07807

[6] N. Frosst and G. E. Hinton, “Distilling a neural network into a soft decision tree,” *CoRR*, vol. abs/1711.09784, 2017. [Online]. Available: http://arxiv.org/abs/1711.09784

[7] O. Bastani, C. Kim, and H. Bastani, “Interpreting blackbox models via model extraction,” *CoRR*, vol. abs/1705.08504, 2017. [Online]. Available: http://arxiv.org/abs/1705.08504

[8] S. Luo, X. Wang, G. Fang, Y. Hu, D. Tao, and M. Song, “Knowledge amalgamation from heterogeneous networks by common feature learning,” *CoRR*, vol. abs/1906.10546, 2019. [Online]. Available: http://arxiv.org/abs/1906.10546

[9] J. Song, Y. Chen, X. Wang, C. Shen, and M. Song, “Deep model transferability from attribution maps,” *ArXiv*, vol. abs/1909.11902, 2019.

[10] M. Alshammari, O. Nasraoui, and S. Sanders, “Mining semantic knowledge graphs to add explainability to black box recommender systems,” *IEEE Access*, vol. 7, pp. 110 563–110 579, 2019.

[11] D. Philps, T. Weyde, A. S. d’Avila Garcez, and R. Batchelor, “Continual learning augmented investment decisions,” *CoRR*, vol. abs/1812.02340, 2018. [Online]. Available: http://arxiv.org/abs/1812.02340

[12] G. R. Vásquez-Morales, S. M. Martínez-Monterrubio, P. Moreno-Ger, and J. A. Recio-García, “Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning,” *IEEE Access*, vol. 7, pp. 152 900–152 910, 2019.

[13] S. Lundberg and S. Lee, “A unified approach to interpreting model predictions,” *Advances in Neural Information Processing Systems 30*, 2017.

[14] Q. Yu, J. Xie, X. Chen, K. Yu, L. Gan, and L. Chen, “Loss allocation for radial distribution networks including dgs using shapley value sampling estimation,” *IET Generation, Tandsmission and Distribution*, vol. 13(8), pp. 1382–1390, 2019.

on demand between datacenters,” *Ieee Transactions on Cloud Computing*, vol. 6(1), pp. 19–32, 2018.

[16] S. Lundberg, G. Erion, H. Chen, A. DeGrave, J. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, “From local explanations to global understanding with explainable ai for trees,” *Nature Machine Intelligence*, vol. 2, 01 2020.

[17] G. R. Vásquez-Morales, S. M. Martínez-Monterrubio, P. Moreno-Ger, and J. A. Recio-García, “Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning,” *IEEE Access*, vol. 7, pp. 152 900–152 910, 2019.

[18] Q. Huang, R. Huang, W. Hao, and R. Fan, “Adaptive power system emergency control using deep reinforcement learning,” *IEEE Transactions on Smart Grid*, vol. 11(2), pp. 1171–1182, 2020.