Journal of Advances in Science and Technology
Vol. 20, Issue No. 2, September-2023, ISSN 2230-9659

DOI: https://doi.org/10.29070/de54jn83

Machine Learning Algorithms to Enhance The Performance of PD Prediction

Lakshmikant Bopche [1] * , Dr. Mrs. A. A. Miraje [2]

1. Research Scholar, Shridhar University, Pillani, Rajasthan, India
luckybopche.780@gmail.com ,
2. Professor, Shridhar University, Pillani, Rajasthan, India

Abstract: Deep learning's strong performance and capacity to learn features automatically have propelled it into the AI mainstream in recent years, where it is being used in a wide variety of applications. In certain fields, like medicine, deep learning may outperform human physicians in terms of accuracy. neurodegenerative diseases, especially Alzheimer's disease, affecting neurons. Tremors, stiffness, and decreased balance are classic motor symptoms of Parkinson's disease, which is caused by a depletion of dopamine neurons in the brain. This disruption in smooth coordination and communication with other nerve cells makes the condition more difficult to control. One of the hallmark clinical aspects of Parkinson's disease is a decline in motor coordination, which shows the central role of dopamine in controlling movement. Motor and non-motor symptoms are also affected by a lack of dopamine neurons. Manual symptom checking is still used to identify Parkinson's disease by doctors and other medical professionals. Researchers have proposed a plethora of ways for Parkinson's disease diagnosis. Voice signals, handwriting traces, MRIs, PET/SPECT scans, and a battery of other modalities are all part of the toolbox for these techniques. Patients with the condition now have improved alternatives for managing symptoms, even though there is no proven medicine to change the illness.

Keywords: Machine Learning, Algorithms, Prediction, computers

---------------------------- X ----------------------------

INTRODUCTION

The field of artificial intelligence known as machine learning is concerned with "teaching" machines new skills and adapt to new situations automatically, rather than relying on pre-programmed instructions. Building algorithms with the ability to self-learn via training on a large number of inputs is the core idea behind Machine Learning. Where traditional algorithm development would be too time-consuming or costly, machine learning algorithms step in to fill the void. Some examples of these uses are computer vision and email filtering.

By using machine learning, massive data sets may be analyzed. It may take more time and resources to train it properly, but it generally gives quicker and more accurate results to detect lucrative opportunities or risky hazards. To further improve its ability to analyses massive amounts of data, machine learning may be combined with artificial intelligence and perceptual technologies. There is a tight relationship between machine learning and computational statistics, the field that employs computers to make predictions. One common way to classify machine learning systems is by the kind of "signal" or "feedback" they may receive; supervised, unsupervised, and semi-supervised learning are the three most common types.

The grid for transferring energy transports massive quantities of power from power plants to substations with high voltage, making it a vital component of the country's energy infrastructure. Due to its complexity, the contemporary power grid needs a security system that is quick, accurate, and dependable. Faults in power networks are unavoidable; in most cases, problems with higher overhead transmission lines are associated with other critical parts. They have far-reaching consequences for end users and for the reliability of the network. Power quality has recently been an important focus in power system engineering due to the fact that 85 to 87% of power system failures occur on distribution lines. It is necessary to anticipate the errors in order to circumvent the problems.

LITERATURE REVIEW

Bode, Gerrit et.al. (2020). In this study, we explore the feasibility of training FDAs based on a dataset of heat pump failures in a controlled environment and then applying them to a real-world application system to see whether any expensive alterations are necessary. We are also considering big data as a potential approach, which would include training the FDAs using data that has been accumulated over an extended duration. This is accomplished by using a dataset kindly provided by NIST, which records typical heat pump failures as identified by an individualized air-water heat pump. Using this information, we determine which attributes are most important for the FDAs to consider and rank them accordingly. We apply the Applying FDAs to our proprietary measurement data after algorithm training on the NIST dataset will help identify any discrepancies. The NIST dataset shows good performance from the trained FDAs, their performance is severely lacking on the real-world dataset.

Earlier research by Ashari et al. (2016) This research focuses on implementing optimisation strategies for distribution networks that include Distributed Generators (DG) using modified Particle Swarm Optimisation (PSO) algorithms. With the help of the advancements in distributed network architecture, the effort to enhance execution was successful. The optimisation primarily aimed at raising voltage efficiency while decreasing dynamic power loss inside the spiral structure, all while keeping the distribution organisations intact.

Saputra et al. (2016) A modified PSO calculation is the basis for the design streamlining approach in this study. A 33-bus spiral distribution network test framework and a 60-bus Bantul distribution network in Indonesia were both used to test the method. The simulation results show how important it is to reconfigure the mechanism to enhance the performance of the distribution system from DG's perspective.

That is, according to Mishra and Ray (2016), The procedure included the suggestion of an additional improvement system, JAYA in particular, to enhance the filter criteria coefficients of the photovoltaic fed distributed static compensator (PV-DSTATCOM). Another development tool that would run without algorithm-specific control settings was the JAYA. In terms of achieving global optimums with less computing effort, it outperforms competing restrictions. When confronted with alternatives like GEM and TLBO, PV stands out. According to DSTATCOM, the results of the simulations and tests show that JAYA advanced PVDSTATCOM is easier to use and more effective.

Faris et al. (2018), A large number of researchers have recently focused on the artificial neural network learning process since it is one of the most challenging problems in machine learning. The nonlinear character and the obscurity of the optimal firm of key governing parameters are the primary challenges of neural system preparation. Conventional training methods have the problems of limited mixing speed and stalling of neighbourhood optima.

METHODOLOGY MODALITIES FOR PARKINSON'S DETECTION

More and more research suggests that prion-like processes may be the starting point for proteins with abnormal folding patterns, which in turn might accelerate neuronal degeneration. Research has shown that fibrils derived from recombination alpha-syncline may trigger neuronal death by recruiting soluble endogenous alpha-syncline into inclusion bodies similar to Levy bodies (2014). Recent studies have shown that alpha-syncline might cause harmful side effects. A large number of macaque monkeys and wild-type mice were injected with levy body enriched fractions containing alpha-syncline after they were isolated from the postmortem brains of Parkinson's disease patients (2013). Human alpha-syncline produced from the Levy body did not exhibit the same properties as alpha-syncline derived from a non-Levy body, including internalization into host neurons, collection into integral neurons, and dispersion to linked sites, when administered to control rats.

Neurons progressively degenerated and endogenous alpha-syncline converted abnormally as a result of this. Levy bodies likely act as a defensive mechanism, and it seems that the development of PD is greatly affected by the increase of aberrant alpha-syncline. Animal studies of multiple system atrophy (MSA) provide further credence to the idea by showing that a comparable mechanism is at work. (2013).

Imaging tests including magnetic resonance imaging (MRI), positron emission tomography (PET), and spectral imaging (SPECT) may help diagnose Parkinson's disease since the condition produces brain structural abnormalities. In order to conduct a PET or SPECT scan, a radioactive material must first enter the patient's body, making the scan invasive. The damaged organ collects the radioactive substance. The scan's ability to reveal the radioactive tracer as vivid red dots allows one to assess the degree of organ damage. The functional state of the brain's regions involved in movement is examined by the PET scan. However, dopamine transporter activity is the most common area that spectroscopy is used for. Because MRI reveals the brain's architecture, it could be possible to detect PD symptoms. Image 1.

It is possible to distinguish PD patients from DIP using magnetic resonance imaging (MRI) since PD creates structural abnormalities that may be seen on MRI. It would seem that no structural brain abnormalities are linked to pseudo-Parkinsonism. In order to classify Parkinson's disease (PD), researchers have using MRI sequences that include both structural and functional signals. Magnetic resonance imaging (MRI) comes in two primary varieties: T1 weighted and T2 weighted.A Both kinds vary in the amount of time that the echo lasts and how often it repeats itself during the conduct. Radiologists and physicians rely on this difference in cerebrospinal fluid (CSF) appearance in T1 and T2 weighted MRI scans to interpret brain pictures correctly and diagnose a range of neurological disorders, such as Parkinson's disease.

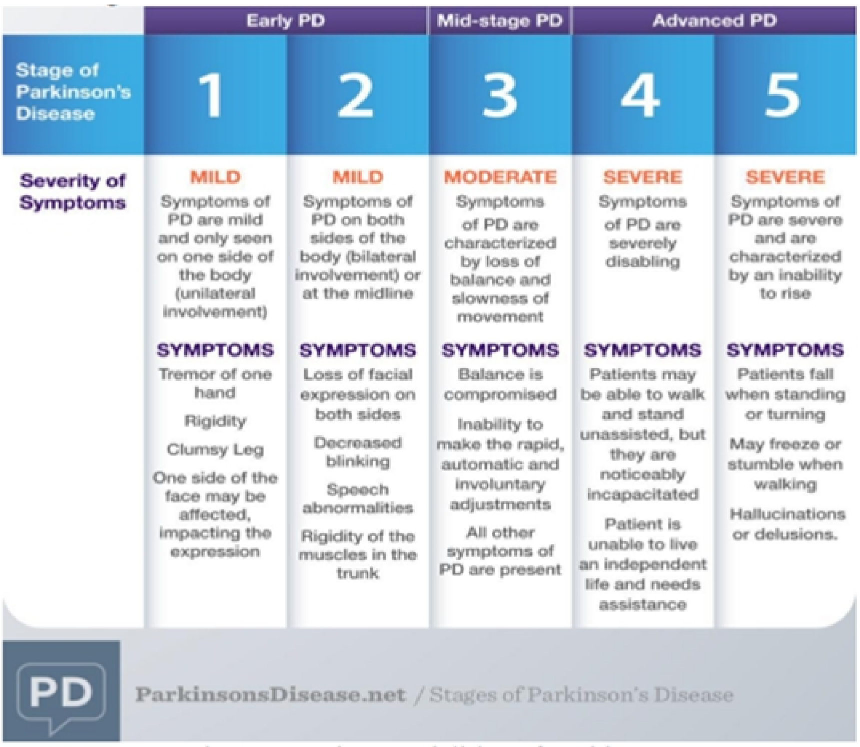

Figure 1 Detection Modalities of Parkinson

Clinical Trials and Research for PD-MCI

The difficulties in treating non-dementia cognitive impairment in PD were brought to light in the 2019 Movement Disorders Society (MDS) update on evidence-based treatment for non-motor symptoms of Parkinson's disease (PD). Regarding cognitive impairment in PD, the study found that several pharmacologic and nonpharmacologic therapies lacked sufficient evidence to be used. The results are shown here.

The word "investigation" appeared on all of the evaluated methods. Physical exercise and medication may assist people with PD who do not have dementia reach their cognitive objectives (2019). Intense physical therapy increased global cognition, whereas aerobic exercise (such riding a recumbent bike or treadmill) had little effect on executive function. Additional research is needed to identify pharmacologic and non-pharmacologic methods, as well as combination approaches, that may be used to target PD-MCI. Image 2. At the same time, there is mounting evidence that exercise may help those with Parkinson's disease.

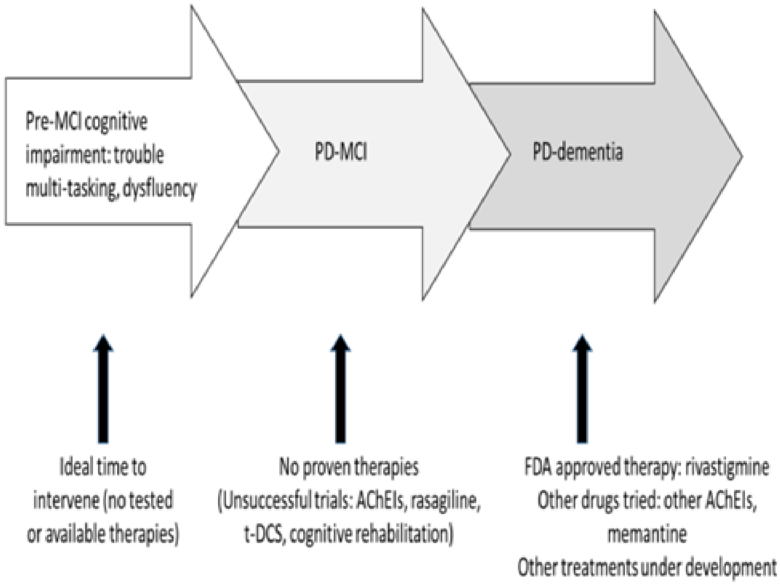

Figure 2. Clinical Approach Parkinson

Discussion and Future

Work Finding the most important features for PD categorization and studying how imbalances in medical data affect accuracy are the goals of your research. Improving the reliability and accuracy of PD diagnostic tools is one of the main goals of this study. With these considerations in mind, three approaches will be used for PD classification, all of which rely on training on features discovered by PCA and obtained after data set balancing. The algorithms used by each method are detailed below:

Classification Model Development

Find the scaled data's covariance matrix. To get the eigenvectors and primary components, you may use eigenvalue or singular value decomposition. Rank the principal components by their explained variance ratios and choose the top five. Sorting out the data by column variance and using Principal Component Analysis (PCA) using models like SVM, logistic regression, random forest, and KNN. Following this, the accuracy, confusion curve, and classification results are evaluated.

STRUCTURED DATA ALGORITHMS

Artificial neural network

Theory

Artificial neural networks (ANNs) are multi-layered, with one or more NNs per layer. A neuron gets inputs from several sources. To begin, a randomly initialized network parameter called a weight is applied to each input. Put the total of each neuron's weighted inputs and deviation values into the activation function (nonlinear variation function), after which you'll get the answer. At its heart, NN is an activation function. Because it adds nonlinearity to the network, it may learn more complicated functions. The output of the neurons in the preceding layer serves as the activation function, which is used as an input by each subsequent layer of neurons. The optimal weight distribution will be found by the network during repeated training, and its optimality will be determined by the loss function. An ANN with three layers is shown schematically in Figure 3. Every network needs an input layer, a hidden layer (or layers), and an output layer to function properly. When put into practice, the network's layer count may easily exceed hundreds.

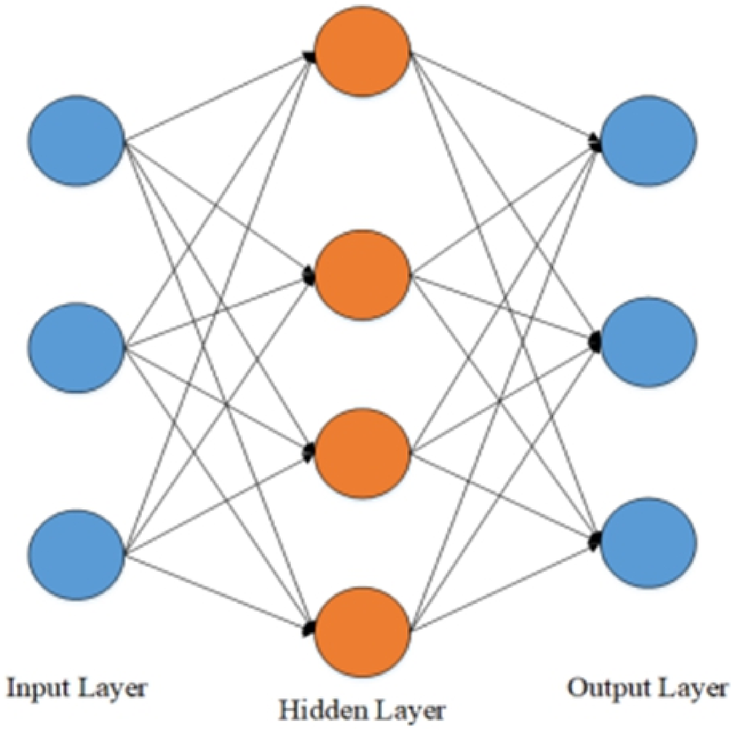

Figure 3 Artificial neural network diagram

Disease application

There is a dearth of study on ANN because to its relatively basic structure, which prevents it from exhibiting the exceptional qualities seen in CNN and RNN. Utilizing NN models with 1, 2, and 3 hidden layers, respectively, with epochs of 200, 400, and 800, Khanam and Foo (2021) used the model to predict diabetes. With 400 epochs and an accuracy of 88.6%, hidden layer 2 outperforms several machine learning models, including Decision Tree, K-Nearest Neighbor (KNN), Random Forest, Logistic Regression, Support Vector Machines (SVM), and many more. After comparing ANN with other machine learning models for AD detection, Soundarya et al. (2021) concluded that, given enough data, ANN attained the best accuracy. In order to enhance the accuracy of cardiovascular illness predictions, Pasha et al. (2020) used ANN. Conventional machine learning models struggle with massive datasets, but ANN may be a lifesaver. What this means is that ANN is going to be a big deal in the future, and that deep learning, which it embodies, will be the go-to technique for illness prediction.

FM-deep learning

Theory

When using linear regression, it is common practice to use a second-order cross term to account for features' second-order interactions:

$$y_{FM} = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{ij} x_i x_j.$$

There are nðn 1Þ=2 parameters in the second-order intersection component. On the other hand, characteristics xi and xj can't both be zero for wij to be found. With the sparse data, especially after one-hot coding, we can be confident that xi and xj will not be zero at the same time. Since there aren't enough instances, the training set doesn't have enough examples of how features interact with each other, leading to over-fitting and erroneous learnt weights. In order to solve this problem, FM separates the input vector wij into its hidden components, vi and vj, which are represented as wij ½ hvi; vji. In this case, vi= (vi1, vi2;...; vik). W, the matrix consisting of wij, may be expressed in the following way:

$$W = \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_k \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_k \end{pmatrix}.$$

The initial number of wij has been significantly reduced to n * k binomial parameters.

The claim that hidden vectors can address data sparsity begs the question: why? Since vh may be learned from any sample that has non-zero feature combinations of xh. Consider the parameters hvh; vii for xhxi and hvh; vji for xhxj. Since they share an object vh, we may make an educated guess as to its worth. Because of this, data sparsity may be significantly mitigated.

The model's generalizability may be enhanced by the use of the implicit vector technique as well. whether FM is able to solve sparsity, then it doesn't care whether a certain combination of features has happened while learning the embedded hidden vector weight of a single feature. In the case of the unique feature combination xixj

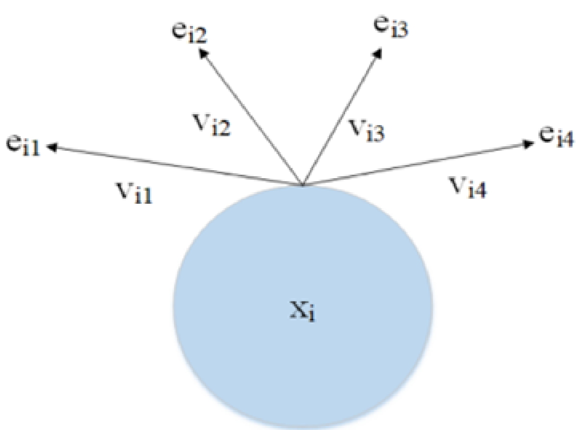

Figure 4 Embedding of feature x_i

Previously, FM had great generalizability because, after learning the hidden vectors for xi and xj, It is possible to get the weight of this feature combination by using the inner product. Here is the FM formula:

$$y_{FM} = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j.$$

The following procedures may lower the complexity of FM from O (n2kÐ) to O (n * k), as is seen from the complexity of FM:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle v_i, v_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^{n} \langle v_i, v_i \rangle x_i^2$$

$$= \frac{1}{2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{f=1}^{k} v_{if} v_{jf} x_i x_j - \sum_{i=1}^{n} \sum_{f=1}^{k} v_{if}^2 x_i^2 \right)$$

$$= \frac{1}{2} \sum_{f=1}^{k} \left(\left(\sum_{i=1}^{n} v_{if} x_i \right) \left(\sum_{j=1}^{n} v_{jf} x_j \right) - \sum_{i=1}^{n} v_{if}^2 x_i^2 \right)$$

$$= \frac{1}{2} \sum_{f=1}^{k} \left(\left(\sum_{i=1}^{n} v_{if} x_i \right)^2 - \sum_{i=1}^{n} v_{if}^2 x_i^2 \right).$$

The final FM equation is:

$$y_{FM} = w_0 + \sum_{i=1}^{n} w_i x_i + \frac{1}{2} \sum_{f=1}^{k} \left(\left(\sum_{i=1}^{n} v_{if} x_i \right)^2 - \sum_{i=1}^{n} v_{if}^2 x_i^2 \right)$$

Essentially, In the end, FM is all about embedding and engagement. Implicit vectors vi ¼ ðvi1; vi2; vi3, vi4) are allocated for every feature xi, which consists of discrete features that have previously been one-hot encoded. For k = 4, the embedding layer converts the high-dimensional data to a dense vector e, which is low-dimensional. To get ei, the corresponding hidden vector vi is multiplied by xi, as illustrated in Figure 4.

Figure 5 shows the whole Embedding layer:

Figure 5 shows the general outline of FM's structure, where $y_{Linear} = w_0 + \sum_{i=1}^{n} w_i x_i,$. $y_{FM2} = \frac{1}{2} \sum_{f=1}^{k} \left(\left(\sum_{i=1}^{n} v_{if} x_i \right)^2 - \sum_{i=1}^{n} v_{if}^2 x_i^2 \right)$

Development history

A FM Supported NN (FNN) was suggested by Zhang et al. (2016) in 2016. The model uses a DNN with embedded layers to complete the CTR prediction. By pre-training the FM model, this network obtains the dense vector of every feature. After that, the sample's embedded vectors are combined and input to DNN for training.

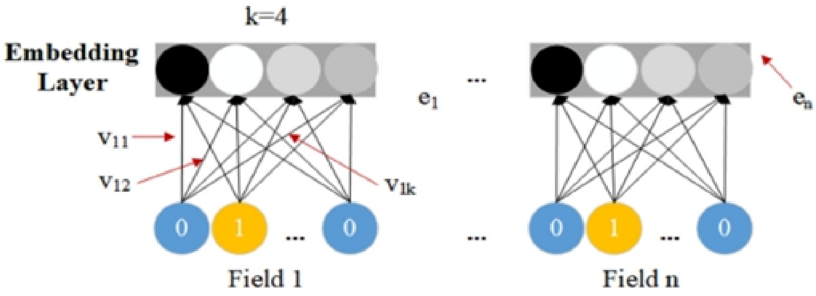

Figure 5 Embedding layer of FM

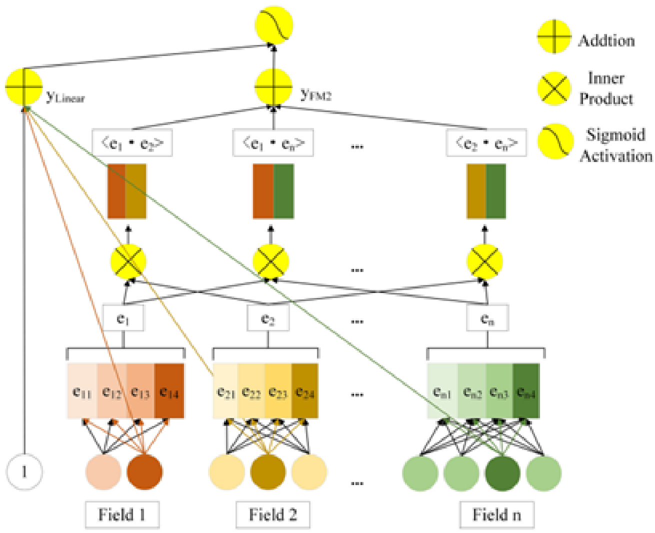

Figure 6 Overall structure diagram of FM

One characteristic of FNN is that the FM model trains the embedding vector of each feature in advance. As a result, the training overhead of a DNN model is decreased, allowing for quicker convergence. The overall network performance, however, is limited by FM's capabilities. At the same time, PNNs were suggested by Qu et al. (2016), who combined completely linked and embedding layers with a product layer. PNN discovers feature relationships by taking the inner or outer product of the features, but it doesn't account for low-order feature interaction, which means it can miss out on useful information in the initial vector. After researching the issue of recommendations with limited input data, he and colleagues introduced Neural FM (NFM) (2017). NFM takes a page out of Wide&Deep (2016) playbook by processing second-order cross information using a Bi-Interaction Layer (Bilinear interaction) structure; DNN structures are able to learn cross-feature information more easily as a result of this, which facilitates the learning of higher-order cross-feature information as well. By integrating Deep and FM, Guo et al. (2017) created DeepFM, a technique for learning low-level feature interaction. In order to achieve its goal, DeepFM ran two algorithms simultaneously: DNN for high-level feature interaction and FM for low-level feature interaction. Plus, there is no difference between the two components' inputs. No pre-training or feature engineering is required since the output layer takes as inputs not only the final first-order features but also the interactions between second-order and higher-order features. Attention FM (AFM) was proposed by him and colleagues (2017) as an expansion of NFM. The interpretability and representation ability of NFM were substantially enhanced once the attention mechanism was included into the Bi-Linear interactive pooling process. It is not possible for AFM to use DNN since the deeper network does not include the quadratic term; all it does is add an attention mechanism based on FM. A novel NN structure for learning called Deep AFM (DeepAFM) was suggested by Zhang et al. (2021) that integrated DeepFM with AFM. By include the feature domain structure, our technique is able to learn the weighted interaction between features more effectively than current deep learning models—all without feature engineering. Investigations into the mechanisms of attention are numerous as well. In order to improve upon DeepFFM, A novel model, FAT-DeepFFM, was presented by Zhang et al. (2019). Prior to the explicit feature interaction procedure, this model dynamically captures the significance of each feature using CENet domain attention. To find out how important co-occurrence features are at different dimensional granularities, Tao et al. developed a cross-interaction layer, updated feature representations by aggregating the representations of other co-occurrence features, and employed a bit-by-bit attention mechanism in their 2020 publication. To further account for the interplay of high-order sparse features, they also suggested Higher-order AFM (HoAFM). Based on these two competing concerns, Yu et al. (2021) proposed Gated AFM (GAFM), which makes advantage of the gate structure to control accuracy and speed. Wen et al. (2021) improved the FM by developing the Neural Attention Model (NAM) with completely connected layers. In order for NAM to understand the varied significance of interactions involving low-order features, the attention mechanism is used. To enable non-linear modeling of higher-order feature interactions, NAM builds fully linked layers atop the attention component. A new neural network architecture called EMD2FNN was introduced by Yang and colleagues in 2019. The non-stationarity of the data may be overcome with the aid of empirical mode decomposition, and the nonlinear interaction between inputs can be mastered with the help of FM. According to Zhang et al. (2019), HCFM stands for High-order Cross-Factor FM. Achieving high-order display interactions was the goal when they developed Cross-Weight Network (CWN). In order to achieve a balance between the weights of high-order and low-order feature interactions, the weight pooling layer learns the weights of different interaction orders. On the other hand, the cross and compression layers of CWN are designed to successfully learn important feature combinations. A Dual-Input FM (DIFM) was proposed by Lu et al. (2021) that can learn multiple representations of a feature depending on the input instances. As an added bonus, it may re-weight the original feature representation by learning input aware factors at both the bit-wise and vector levels concurrently. The DIFM is an end-to-end model that purposefully incorporates several components, such as DNN, residual networks, and multi-head self-attention. The novel Deep Field-weighted FM (DeepFwFM) introduced by Deng et al. (2020) combines FwFM and regular DNN components; it demonstrates distinct benefits in structure pruning and, when used, significantly decreases inference time. Yu et al. (2020) created the Neural Pairwise Ranking Factorization Machine (NPRFM), a model that uses a multilayer perceptual NN in Pairwise Ranking Factorization Machine. Notably, a double-interaction layer records higher-order and nonlinear interactions between features, while a multi-layer perceptual neural network encodes second-order interactions between features. Both Deep FEFM (DeepFEFM) and Field Embedding FM (FEFM) were suggested by Pande (2020). For every field pair and feature, FEFM learns the symmetric matrix embedding and the single vector embedding. By combining

DNN with the FEFM interaction vector learned by the FEFM components, DeepFEFM is able to learn high-order feature interactions. To address the "short expression" issue and enhance the capturing of interactions including several density features, Qi and Li (2021) suggested Deep Field-Aware Interaction Machine (DeepFIM). A novel "hierarchy expression" based on field identifiers was suggested by them for feature interaction expressions. They built a cross-interaction layer to identify field and field interaction based on this and used an attention mechanism to distinguish between aspects of various significance. Use of a dynamic bi-pool layer enhances feature acquisition at the high-order level.

Furthermore, a CNN/FM hybrid does exist. The DGFFM model, proposed by Zhang et al. (2019), uses a wide-deep architecture to train both DenseNet and GFFM at the same time. Aiming to combine the advantages of traditional machine learning methods—like their faster learning speed for low-rank features and their ability to extract high-dimensional features GFFM drastically cuts computation time by utilizing the corresponding positional relationship between field indices and feature indices. Chanaa and El Faddouli (2020) introduced Latent Graph Predictor FM (LGPFM), which used convolutional neural networks (CNNs) to derive interaction weights for every feature pair. Because convolutional neural networks (CNNs) work so well with grid topologies, LGPFM combines the greatest aspects of FM and CNN to provide much more accurate outcomes.

It is also possible to integrate metric learning with the FM algorithm. Guo et al. (2020) proposed a GML-based FM architecture. By projecting features into a new space using a semi-positive definite matrix, the features must comply to certain linear restrictions in order for the metric approach based on Mahalanobis distance to be used. To take use of the powerful representational capabilities of both the metric learning approach and NN, To capture the nonlinear feature correlation, a distance function was built using DNN. Additionally, each pair of attributes now has a learnable weight that may be used to enhance the distance function's effectiveness.

Disease application

For the purpose of diagnosing sepsis in children, Chen and Qian (2017) suggested using NN and FM. While NN excels at handling patients' test index result values, FM is more suited to handling sparsely structured patients' test index state data. Ronge et al. (2021) developed a deep FM model for the diagnosis of Alzheimer's disease. An embedding layer handles sparse categorical data, a Factorization Machine expedites pairwise interaction learning, and the system is comprised of these three parts and a DNN for implicit modeling of higher order interactions. None of the above-mentioned FM-Deep Learning algorithms—which have superior performance—are employed; instead, they are basic NN and FM combinations. In contrast, Fan et al. (2021) 354 patients at Peking Union Medical College Hospital who had obtained initial postoperative remission were able to have their Cushing's disease returns predicted using Deep FM after transsphenoidal surgery; and they outperformed competing models with respect to both AUC (0.869) and logistic loss (0.256).

CONCLUSION

The use of deep learning algorithms for illness prediction is discussed in this article. There are two main categories of algorithms: those that work with structured data and those that work with unstructured data.

Both ANN and FM-Deep Learning are deep learning algorithms that can handle structured data. Unstructured data networks include a variety of types, including RNNs, CNNs, and others. The literature review makes it quite evident that a reliable PD detection approach is necessary. Doctors still depend on characteristics to help them make a diagnosis. Researchers have created several frameworks for the identification of Parkinson's disease (PD) utilizing a range of modalities, such as voice signals, SPECT scans, PET pictures, MRIs, handwriting drawings, and FTT, however they cannot be used in isolation to diagnose PD. People should avoid invasive and potentially dangerous imaging examinations like SPECT and PET. You can only use FTT and scrawled drawings to look for upper-limb deficiencies, and voice signals aren't particularly reliable. Research employing MRI sequences has recently allowed for PD identification.

References

1. Bode, gerrit & thul, simon & baranski, marc & mueller, dirk. (2020). Real-world application of machine-learning-based fault detection trained with experimental data. Energy. 198. 117323. 10.1016/j.energy.2020.117323.

2. Syahputra, R, Soesanti,I & Ashari, M2016,‗Performance enhancement of distribution network with dg integration using modified PSO algorithm', Journal of Electrical Systems, vol. 12, no. 1, pp. 6-12.

3. Syahputra, R, Soesanti,I & Ashari, M2016,‗Performance enhancement of distribution network with dg integration using modified PSO algorithm', Journal of Electrical Systems, vol. 12, no. 1, pp. 6-12.

4. Kumar, RA, Kumar, GS, Kumar, BK & Mishra, MK 2009, ‗Compensation of voltage sags with phase-jumps through DVR with minimum VA rating using PSO ', in TENCON 2009 IEEE Region 10 Conference IEEE, pp. 1-6.

5. Aljarah, I, Faris, H & Mirjalili, S 2018, ‗Optimizing connection weights in neural networks using the whale optimization algorithm ', Soft Computing, vol. 22, no. 1, pp. 1-15.

6. Aybar-Ruiz, A, Jiménez-Fernández, S, Cornejo-Bueno, L, CasanovaMateo, C, Sanz-Justo, J, Salvador-González, P & Salcedo-Sanz, S, 2016, 'A novel grouping genetic algorithm–extreme learning machine approach for global solar radiation prediction from numerical weather models inputs', Solar Energy, vol. 132, pp.129-142.

7. Azimi, R, Ghayekhloo, M & Ghofrani, M 2016, 'A hybrid method based on a new clustering technique and multilayer perceptron neural networks for hourly solar radiation forecasting', Energy Conversion and Management, vol. 118, pp.331-344.

8. Cornejo-Bueno, L, Casanova-Mateo, C, Sanz-Justo, J & Salcedo-Sanz, S 2019, 'Machine learning regressors for solar radiation estimation from satellite data', Solar Energy, vol. 183, pp.768-775.

9. Cheng, HY, 2016, 'Hybrid solar irradiance now-casting by fusing Kalman filter and regressor', Renewable Energy, vol. 91, pp.434-441.

10. Bouzgou, H & Gueymard, CA 2017, 'Minimum redundancy–maximum relevance with extreme learning

machines for global solar radiation forecasting: Toward an optimized dimensionality reduction for solar time series', Solar Energy, vol. 158, pp.595-609.

11. Benali, L, Notton, G, Fouilloy, A, Voyant, C & Dizene, R 2019, 'Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components', Renewable energy, vol. 132, pp.871-884.
12. Babar, B, Luppino, LT, Boström, T & Anfinsen, SN, 2020, 'Random Forest regression for improved mapping of solar irradiance at high latitudes', Solar Energy, vol. 198, pp.81-92.
13. Belaid, S & Mellit, A 2016, 'Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate', Energy Conversion and Management, vol. 118, pp.105-118.
14. Bouzerdoum, M, Mellit, A & Pavan, AM 2013, 'A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant', Solar energy, vol. 98, pp.226-235.
15. Mathiesen, P & Kleissl, J 2011, 'Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States', Solar Energy, vol. 85, no. 5, pp.967-977.
16. C. Chauhan, et al., "Multiple Disease Prediction Using Machine Learning Algorithms," 2021.
17. A. Kamboj, et al., "A Machine Learning Model for Early Prediction of Multiple Diseases to Cure Lives," 2020.
18. S. Kolli, et al., "Symptoms Based Multiple Disease Prediction Model using Machine Learning Approach," 2021.
19. P. Krishnaiah, et al., "Predictive Modeling for Multiple Diseases Using Machine Learning with Feature Engineering," 2015.
20. H. Al-Mallah, et al., "Multiple Disease Prediction Using Hybrid Deep Learning Architecture," 2016.
21. Y. Gamo, et al., "Machine Learning Based Clinical Decision Support Systems for Multi-Disease Prediction: A Review," 2020.
22. W. Li, et al., "Towards Multi-Disease Prediction Using Graph Neural Networks," 2020. [8] R. Ribeiro, et al., "A Survey on Explainable AI Techniques for Diagnosis and Prognosis in Healthcare," 2020.
23. E. Char, et al., "Ethical Considerations in AI-Driven Healthcare," 2020. M. Wild, et al., "Streamlit for Machine Learning: Creating Interactive Web Apps in Python," 2022.
24. R. Liang, S. Wu, P. Chi, N. Peng, and Y. Li, ''Optimal placement of UHF sensors for accurate localization of partial discharge source in GIS,'' Energies, vol. 12, no. 6, p. 1173, Mar. 2019.
25. Z. Chen, ''Review of direction of arrival estimation algorithms for partial discharge localisation in transformers,'' IET Sci., Meas. Technol., vol. 13, no. 4, pp. 529–535, Jun. 2019.

26. Q. Xie, T. Li, J. Tao, X. Liu, D. Liu, and Y. Xu, ‘‘Comparison of the acoustic performance and positioning accuracy of three kinds of planar partial discharge ultrasonic array sensors,’’ IET Radar, Sonar Navigat., vol. 10, no. 1, pp. 166–173, Jan. 2016.
27. N. Zhou, L. Luo, H. Song, G. Sheng, and X. Jiang, ‘‘A substation UHF partial discharge directional of arrival estimation method based on maximum likelihood estimation,’’ Trans. China Electr. Soc, vol. 34, no. 15, pp. 3285–3292, 2019.
28. B. Han, L. Luo, G. Sheng, and X. Jiang, ‘‘Partial discharge direction-ofarrival estimation in air-insulated substation by UHF wireless array and RSSI maximum likelihood estimator,’’ IET Gener., Transmiss. Distrib., vol. 14, no. 23, pp. 5491–5497, Dec. 2020.
29. Q. Liu, M. Zhu, Y. Wang, J. Deng, and G. Zhang, ‘‘Partial discharge orientation method in substation based on ultra-high frequency phased array theory,’’ Proc. CSEE, vol. 37, no. 20, pp. 6126–6135, 2017.
30. X. Pang, H. Wu, X. Li, Y. Qi, H. Jing, J. Zhang, and Q. Xie, ‘‘Partial discharge ultrasonic detection based on EULER-MUSIC algorithm and conformal array sensor,’’ IET Gener., Transmiss. Distrib., vol. 12, no. 14, pp. 3596–3605, 2018.
31. N. Zhou, L. Luo, G. Sheng, and X. Jiang, ‘‘Direction of arrival estimation method for multiple UHF partial discharge sources based on virtual array extension,’’ IEEE Trans. Dielectr. Electr. Insul., vol. 25, no. 4, pp. 1526–1534, Aug. 2018.
32. Q. Xie, Y. Wang, T. Li, X. Bian, H. Zhang, and Y. Xu, ‘‘Application of signal sparse decomposition in the detection of partial discharge by ultrasonic array method,’’ IEEE Trans. Dielectr. Electr. Insul., vol. 22, no. 4, pp. 2031–2040, Aug. 2015.
33. S. Gao, Y. Zhang, Q. Xie, Y. Kan, S. Li, D. Liu, and F. Lü, ‘‘Research on partial discharge source localization based on an ultrasonic array and a step-by-step over-complete dictionary,’’ Energies, vol. 10, no. 5, p. 593, Apr. 2017.
34. L. Perfetto and G. D’Antona, ‘‘Experimental results of partial discharge localization in bounded domains,’’ Sensors, vol. 21, no. 3, p. 935, Jan. 2021.
35. W. Yi-Fan, G. Jin-Long, J. Jian-Ling, T. Zhi-Guo, and W. Fei, ‘‘Field measurement and verification of the UHF calibration platform based on the GTEM cell,’’ J. Eng., vol. 2019, no. 16, pp. 1585–1588, Mar. 2019.