Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 11
November 2024

Data Reduction Using Feature Reduction Technique and Particle Swarm Optimization

[1]Shubhahshree Pattnayak , [2]Sumant Sekhar Mohanty , [3]Rajanikant Sahu
[1]Computer Science and Engineering Gandhi Institite of Excellent Technocrats
[2]Computer Science and Engineering Gandhi Institite of Excellent Technocrats
[3]Computer Science and Engineering Gandhi Institite of Excellent Technocrats
rajanikantsahu84@gmail.com

Abstract— Feature reduction, commonly referred to as dimensionality reduction, is the process of minimizing the number of features involved in a computation-heavy task without sacrificing crucial information. By decreasing the number of features, the computational load is reduced, enabling more efficient and faster processing. Dimensionality reduction techniques are categorized into two primary types: feature selection and feature extraction. Feature selection focuses on identifying the most relevant features, while feature extraction transforms existing features into a new, reduced set that still retains the essential information. Feature selection - it is where naturally or physically chose which contribute a large portion of the reduction variable or output. In other word, it is a process in which a superior set of features as the best subset is selected. There are three benefits such as reduces over- fitting, improving accuracy, reduces training time. Feature selection is implemented using three techniques such as Wrapper, Filter and Embedded. Wrappers evaluate a specific model sequentially using different potential subsets of features to get the subset that best works in the end. They are profoundly expensive and have a high possibility of over-fitting. Channels techniques are quicker elective that don't test a specific calculation, however rank the first highlights as per their relationship with the issue and simply select the highest point of them. It is a statistical test used to assess the independence of variables, determining whether or not there is a significant dependency between them. This method helps identify relationships or associations between variables by evaluating if the observed distribution of data deviates from what would be expected under the assumption of independence. Few techniques in this category includes Correlation coefficients: removes duplicate features, Information gain or mutual information. A detail discussion on advantages and disadvantages of different filters and wrapper approach for feature reduction is going to be highlighted. Feature selection is a significant information pre-handling procedure, yet it's anything but a troublesome issue due basically to the large search space. Particle swarm optimization (PSO) is a highly effective evolutionary computation technique. Be that as it may, the conventional individual best and global best refreshing component in PSO limits its presentation for highlight choice and the capability of PSO for feature selection has not been completely explored. Evolutionary computation (EC) strategies are notable for their global accessibility. Particle Swarm Optimization (PSO) is a relatively recent evolutionary computation (EC) method, known for being more computationally affordable compared to many other EC algorithms. In this manner, PSO has been utilized as a powerful method in highlight choice.

Keywords—Feature reduction, wrapper, Filter, Embedded, PSO.

1. INTRODUCTION

Feature reduction, otherwise called dimensionality reduction, is the way toward diminishing the number of features in a resource-heavy calculation without losing significant data. It might prompt some measure of information misfortune. PCA will, in general, find direct connections between's factors, which is here and there unwanted. PCA fails in cases where mean covariant are not enough to define the data set. Filter methods determine feature relevance by examining their correlation with the target variable. In contrast, wrapper methods involve training a model to assess the effectiveness of different feature subsets. Feature selection is a significant information pre-handling procedure, yet it's anything but a troublesome issue due basically to the large search space. Particle Swarm

Optimization (PSO) is an effective evolutionary computation (EC) technique. However, the traditional individual best and global best updating mechanism in PSO constrains its performance in feature selection, and its potential in this area has not been fully explored. EC strategies are renowned for their global optimization capabilities. PSO, a relatively recent EC method, is computationally more efficient compared to many other EC algorithms. In this manner, PSO has been utilized as a powerful method in highlight choice. Technically, particle swarm optimization (PSO) is a computational classification that improves a query by duplication working to promote a competitor clarification about concerning a given measure of quality. The PSO algorithm, suggested by Kennedy and Eberhart, implies that meta-heuristic algorithm is based on the idea of swarm intelligence or knowledge or data capable of solving complicated mathematical problems existing in engineering. The fundamental concept of PSO is caused by the social function of birds gathering, fish schooling, and swarm theory.PSO has remained applied to describe a spectrum of optimization barriers, including neural chain practice, and function minimization.PSO holds a computational system that optimizes each query on iteratively working to promote a bidder resolution concerning a given model regarding the group. Particle Swarm Optimization (PSO) is still a developing field, first introduced by Kennedy and Eberhart in 1995. Originating from the concept of simulating social behavior, PSO has grown into a widely applicable optimization method. Some features that attracted scientists were the underlying rules that permitted huge numbers of birds to gather synchronously, frequently replacing way quickly, scattering including regrouping, etc. For certain primary considerations, bird gathering and fish schooling did some regarding the behavioral models which were claimed to be copied. The group distribution about learning amongst the members suggests an evolutionary benefit. Interacting also distributing the data among their community particles, identified since particular best or local best for comparing their neighbor best. At every step of the process, the global best resolution reached during this entire swarm remains refreshed.

2. LITERATURE REVIEW

Different works have just been done in the past on expulsion of filter and wrapper approach. Feature selection seeks to pinpoint an optimal subset of variables that effectively represents the underlying patterns of the training data, minimizing the influence of noise and irrelevant attributes. This process enhances model performance by reducing dimensionality, thereby improving computational efficiency and mitigating the risk of over-fitting. By carefully selecting the most significant features, we ensure that the model focuses on the most informative aspects, leading to more accurate and interpretable results. Filter-based methods offer a more versatile and computationally efficient solution, uninfluenced by the learning algorithm, making them ideal for high-dimensional data. In contrast, wrapper-based methods employ a systematic search to identify suitable feature subsets, evaluating their performance using supervised learning algorithms, which can be a more laborious and exhaustive process, for example, classification performances on a cross-approval of the training set which gave preferred outcomes over filter methods. In any case, wrapper approaches increment the computational expenses. The channel based methodologies are liberated from the coordinated learning estimation and offer all the more comprehensive articulation and they are computationally more affordable portrayed by H Rozas. For taking care of the great dimensional data, channel strategies are appropriate proposed by L Yu, H Liu. Covering based methodology utilizes any of the looking through strategies and assesses utilizing the managed learning calculation regarding arrangement blunder or precision proposed by SB Kotsiantis, I Zaharakis. The benefits and inconveniences of channel, covering, and installed techniques have been inspected in 2011 by two authors Deepa and Ladha, in 2007 by Saeys et al., Canedo et al. (2013), and Canedo et al. (2014).The inserted techniques incorporate more tasteful, for example, choice trees, weighted credulous The study by Saeys, et al. (2007) explores various machine learning methods for feature selection, including Bayesian techniques and Support Vector Machines (SVM). In the context of SVM, a weight vector is pivotal. It defines the direction of the optimal hyper plane that separates data points from different classes in the feature space. The weight vector influences the decision function used for classification, and its magnitude reflects the contribution of each feature in the classification task. In relation to SVM-RFE (Recursive Feature Elimination), as proposed by Guyon et al., Vapnik et al. (2002), the method is a feature selection algorithm that iteratively removes the least important features based on the SVM weight vector. The weight vector in SVM- RFE serves as a ranking mechanism: features with the smallest weight coefficients are considered less important and are eliminated at each iteration. This recursive process aims to retain only the most discriminative features, improving model performance and reducing over fitting. Both studies highlight the critical role of the SVM weight vector in feature selection and classification, with SVM-RFE being a key method for feature ranking through the evaluation of this. The paper by Maldonado, Weber, and Basak (2011) focuses on enhancing feature selection methods, particularly in the context of Support Vector Machines (SVMs). Their research focuses on enhancing the computational efficiency and overall effectiveness of SVM-based methods when dealing with high-dimensional datasets, which often contain noise and irrelevant features. By optimizing the SVM algorithm's ability to select the most informative features and discard irrelevant ones, their work aims to reduce the computational burden while maintaining or even improving predictive performance. This approach is crucial in real-world applications where data dimensionality is high, and preprocessing to eliminate noise is challenging, ensuring that SVMs remain a robust tool for classification and regression in complex and noisy environments. Then Huang et al. (2011), and U˘guz et al., (2012) suggested half and half element determination by consolidating the channel and covering strategies. The channel techniques assess the meaning of highlights by applying a positioning strategy that thus isolates low-scoring highlights. The channel techniques are resolved to be quick, adaptable, computationally straightforward, and sovereign of the classifier. The techniques are arranged into two methods: the univariate channel strategy and the multivariate channel technique. The work by Yongjun et al. (2012) and the earlier study by Yusta (2009) contribute to the field of feature selection and optimization in machine learning, with a specific emphasis on improving the performance of classification models, including Support Vector Machines (SVMs). In their 2007 paper, Saeys et al., Larranaga et al., provide an extensive review of feature selection methods used in machine learning, particularly highlighting their application to high-dimensional data. Their study underscores the critical role of feature selection in reducing dataset dimensionality, which is vital for enhancing both the performance and interpretability of classification models, such as Support Vector Machines (SVMs). By identifying and retaining only the most relevant features, these methods help mitigate the curse of dimensionality, reduce computational complexity, and improve the model's ability to generalize to new data. This is especially important for SVMs, where high-dimensional data can lead to longer training times and increased risk of over fitting. Effective feature selection not only boosts the accuracy and efficiency of SVMs but also makes the models more understandable and easier to interpret by focusing on the most significant variables. The two foundational systems to be discussed subsequently are classified under the multivariate category, whereas the latter two strategies are categorized within the univariate grouping. The channel based approaches are liberated from the organized picking up figuring and offer even more comprehensive articulation and they are computationally more affordable depicted by H Rozas. For taking care of the great dimensional data, channel strategies are reasonable proposed by L Yu, H Liu. Covering based methodology utilizes any of the looking through strategies and assesses utilizing the administered learning calculation as far as characterization blunder or precision proposed by SB Kotsiantis, I Zaharakis. The studies in 2011 by prominent researcher in this field Ladha and Deepa and Saeys et al. (2007) both focus on the importance of feature selection in machine learning, particularly for improving model performance when dealing with high-dimensional data. Each paper, however, approaches the topic from a slightly different perspective. Alonso

Betanzos et al.(2013), and Bol´on-Canedo et al. (2014). The methodology incorporates advanced algorithms, including decision trees, weighted naive Bayes, and SVM weight vectors, as well as feature selection techniques like SVM-RFE and L1-regularized SVM. These approaches, inspired by the work of Saeys et al. (2007), Guyon et al. (2002), and Maldonado et al. (2011), enhance the accuracy and efficiency of the analysis. Concurrently, a series of studies by Hui-Huang et al. (2011), U˘guz (2012), and Naqvi (2012) introduced a blended approach to feature selection, synergizing filter and wrapper methods to enhance performance. The channel techniques assess the meaning of highlights by applying a positioning strategy that thus isolates low-scoring highlights. The channel techniques are resolved to be quick, adaptable, simple computating, and sovereign of the classifier. The procedures are arranged into two divisions: the univariate channel strategy and the multivariate channel technique. The univariate techniques gauge the highlights separately, in this manner dismissing highlight regions and taking care of helpless component subsets (Yongjun et al. 2012; Yusta, 2009). Contrary to univariate techniques that disregard interaction effects and algorithmic correlations, multivariate strategies adopt a holistic perspective, accounting for these crucial factors in a unified framework (Saeys et al. 2007). The initial two strategies that will be clarified in the resulting area fall into the multivariate class, while the last remaining techniques fall into the univariate classification. The PSO algorithm, suggested by Kennedy and Eberhart, implies that metaheuristic algorithm is based on the idea of swarm intelligence or knowledge or data capable of solving complicated mathematical problems existing in engineering.

They noted that Particle Swarm Optimization (PSO) relies heavily on stochastic processes, akin to those in evolutionary programming. It also closely resembles the crossover operation used in genetic algorithms . In the year 2016, Vijay, Kavitha, and Rebecca highlighted that automatic tumor segmentation remains a significant challenge due to issues like low contrast, poorly defined boundaries, and accuracy concerns. To address this, they implemented the Enhanced Discrete Particle Swarm Optimization (EDPSO) algorithm, which demonstrated a higher quality rate in processing all input images compared to the standard Particle Swarm Optimization (PSO) algorithm. This improvement suggests that EDPSO offers better performance in tackling the complexities associated with tumor segmentation. however, they are using gaussian filtering technique in their process which is not good in handling salt pepper noise (very fluent of occurrences in ADC conversion in the medicalequipment). In 2014, Al-Tamimi Sulong mentioned that for the accurate detection of the type of abnormality brain the treatment of brain is highly required for minimization of diagnosis error in it. This detection accuracy can be improved by computer aided diagnosis (CAD) which is used to give the computer output image of the detected area and reduced the image reading time, it also improves the efficiency, consistency, and accuracy of radiological diagnosis [3].

A. Dimensionality reduction refers to the process of decreasing the number of random variables under consideration by deriving a smaller set of significant variables. This process can generally be categorized into two main approaches: feature selection, which involves identifying and selecting a subset of relevant features from the original data, and feature extraction, which transforms the data into a new set of dimensions while retaining most of the original information. Avoiding overfitting is a major motivation for performing di mensionality reduction.As the quantity of features expands, the model turns logically erratic. The more the amount of highlights, the more the shots at overfitting. An AI model that is ready on countless highlights, gets dynamically subject to the data it was ready on and along these lines overfitted, bringing about horrible appearance on authentic information, beating the explanation.

Avoiding overfitting is a crucial reason for conducting dimensionality reduction. With fewer features in our training data, our model makes fewer assumptions and becomes less complex, thus reducing the risk of overfitting. Nevertheless, that isn't all and dimensionality decrease has substantially more focal points to bring to the table, as Figure 1 is showing the graph of feature reduction performance with the number of features. [3]. Figure 2.1 Feature reduction performance vs number of features

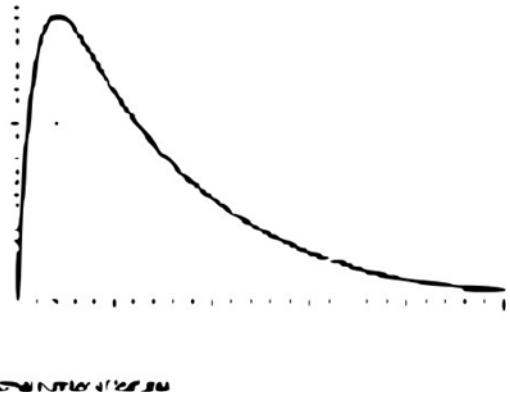

Figure 2.1: Feature reduction performance vs number of features

Reducing misleading information leads to improved model accuracy. Fewer dimensions result in lower computational demands, allowing algorithms to train faster. This reduction also requires less storage space. Moreover, with fewer dimensions, algorithms that are unsuitable for high-dimensional data become viable options. Additionally, this process eliminates redundant features and noise. There are various beneficiary uses of feature reduction.

Some of them are as follows: Feature selection aids in data compression, reducing the amount of storage required by eliminating unnecessary variables. By streamlining the modeling process, it lowers computation time, removes redundant or irrelevant features, and improves data quality. These enhancements lead to more accurate and reliable models, which are better at identifying significant patterns and relationships in the data, ultimately increasing their effectiveness and performance.

Feature selection, also referred to as variable selection, attribute selection, or variable subset selection, is the process of identifying a subset of relevant features (variables or predictors) to be used in building a model. This process consists of two key components: identifying the most significant features that contribute to the predictive power of a model and eliminating redundant or irrelevant features that do not enhance model performance. By focusing on the most relevant data, feature selection helps improve model accuracy, reduce complexity, and prevent over-fitting a search mechanism that proposes novel feature subsets, and a evaluation criterion that scores their effectiveness. Algorithms for feature selection are typically categorized into three main types: wrappers, filters, and embedded methods. The application of feature selection techniques aims to achieve several objectives, including simplifying models to enhance interpretability for researchers and users, reducing training times, and mitigating the effects of the "curse of dimensionality." By narrowing down the feature set to only the most relevant variables, feature selection makes models more comprehensible, less computationally intensive, and more robust, especially when dealing with high-dimensional data. Wrapper methods function by evaluating subsets of features through the use of machine learning algorithms, employing a search procedure to explore the possible combinations of features. Each subset is assessed based on how well the algorithm performs using that specific set of features. These methods are often referred to as greedy algorithms because their goal is to find the optimal combination of features that produces the highest-performing model. However, this approach is computationally intensive due to the exhaustive search for the best subset. [5].

Figure 2.2 Model of a wrapper method.

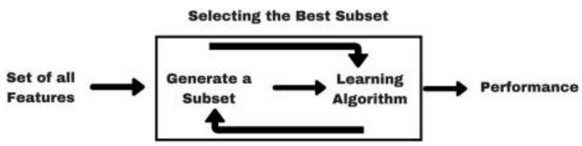

Figure 2.2: Model of a wrapper method

Wrapper Method Processes and Analysis: Using a search method (depicted next), we select a subset of features from the accessible ones. In this progression, a picked ML algorithm is prepared on the previously chosen subset of features and at long last, we assess the newly-trained ML model with a chosen metric. The entire procedure begins again with another subset of features, another ML model prepared, etc. Filter methods for feature selection identify the most relevant features from a dataset independent of any machine learning algorithm. These methods evaluate the inherent characteristics of each variable to determine their importance, filtering out irrelevant or redundant features before the training phase begins. Various filter techniques exist, each with distinct principles, such as assessing correlation, mutual information, or statistical significance, to systematically refine the dataset and enhance model performance.

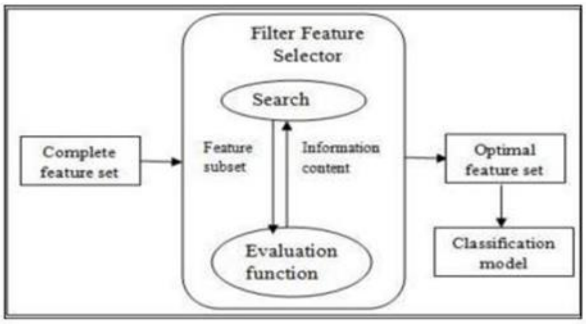

Figure 2.3: Model of a filter method

Filter methods independently select features from a dataset without relying on a specific machine learning algorithm. These techniques solely consider the inherent properties of the features themselves, filtering out irrelevant or redundant ones before the model training begins. A variety of filter methods exist, each employing different principles, such as correlation, statistical measures, or mutual information, to determine feature relevance and enhance the efficiency and accuracy of the learning process. By selecting the most relevant features, these methods streamline the dataset, removing redundant and irrelevant information that could otherwise hinder model performance. Benefits of filter methods are that they have a low calculation time and will not overfit the information. However, one drawback is that they are incognizant in regards to any associations or relationships between's highlights. Figure 3 is showing the overall usefulness of filter approach. Wrapper methods employ a machine learning algorithm to evaluate a subset of features, using a search process to navigate the possible feature combinations. This approach assesses each subset based on its performance, aiming to identify the optimal feature set that yields the best model. Due to their iterative nature, wrapper methods can be computationally intensive, as they strive to find the most effective feature combination. The wrapper method evaluates subsets of features by iterating through the entire training set and testing various combinations of features until the optimal subset is found. While effective in improving model performance, this method has two key disadvantages. First, it is computationally expensive, especially when dealing with datasets containing a large number of features. Second, it can lead to overfitting, particularly when there are limited data points, as the model may fit too closely to the training data. Figure 2 illustrates the general workflow of the wrapper approach.

Search methods are based on finding the best feature subset among features. There are four wrapper search methods are follows:

Forward Feature Selection: This iterative approach builds a model by progressively adding the most significant features. It starts with a blank slate, then identifies the feature with the strongest correlation (smallest p-value) and incorporates it into the model. Next, it assesses the remaining features in combination with the selected one, adding the most impactful feature at each step. This process continues until all substantial features are included, while irrelevant ones are omitted.

Backward Feature Elimination: This method takes a holistic approach, starting with a comprehensive model that includes all features. It then systematically removes the least significant features, beginning with the one having the highest p-value (most insignificant). Through repeated iterations of model testing and p-value calculation, the weakest features are eliminated until only the most influential ones remain.

Exhaustive Feature Selection: This strategy attempts all conceivable feature combinations. This strategy attempts all conceivable feature combinations. Exhaustive feature selection based algorithm is defined wrapper approach for animal power assessment of feature subsets; the best subset is chosen by upgrading a predetermined exhibition metric given a

subjective regress or classifier. Bidirectional Search: This last one does both forward and backward feature selection all the while so as to get one unique solution.

Filter methods can be categorized into two types: univariate and multivariate.

Univariate filter methods evaluate and rank individual features based on specific criteria, treating each feature independently of others in the dataset. These methods operate by assessing each feature in isolation, assigning a score according to predefined metrics, and then selecting the highest-ranked features. While this approach can be effective, it has a notable limitation: it may select redundant features because it does not account for relationships or interactions between different features. Consequently, highly correlated or overlapping variables may still be chosen, reducing the overall model efficiency.

The univariate filter methods are the kind of techniques where individual features are positioned by explicit rules. The top N highlights are then chosen. One of the significant disadvantage of univariate channel strategies is that they may choose repetitive feature on the grounds that the connection between singular feature isn't considered while deciding. Multivariate filter methods, then again, evaluate the entire feature space. They consider includes different ones in the informational index. These strategies can deal with copied, repetitive, and related features. Multivariate filter methods are equipped for expelling repetitive features from the information since they consider the common connection between the features. Multivariate filter methods can be utilized to expel copy and corresponded features from the data.

There are some kind of benefits we are getting from filter approach in feature reduction processes are simple interacts with the classifier, small over lifting risk, less computational, prone to local optima, consider the dependence among features. Similarly, wrapper approach give us some benefits as well. These are for giving better performance, less prone to local optima, interacts with classifier, models measure dependencies, higher performance accuracy.

The filter method is a broad category of feature selection techniques that operate independently of any specific machine learning algorithm. These methods are generally faster and less computationally intensive compared to wrapper approaches because they evaluate the relevance of features based on intrinsic statistical properties, such as correlation or mutual information, without involving model training. This simplicity makes filter methods less susceptible to over-fitting since they do not rely on the iterative training of models.

On the other hand, the wrapper method employs a specific machine learning algorithm to evaluate and select the optimal subset of features. By assessing each combination of features through model performance, wrappers can provide highly tailored feature sets for a given algorithm. However, this approach is computationally expensive, especially for datasets with a large number of features, as it involves repeatedly training the model on different subsets of data. The repetitive model training increases the risk of over-fitting, as the method can become overly tuned to the training data, capturing noise rather than generalizable patterns.

PSO is a commutating system that enhances each query on repetitively working to promote a bidder resolution concerning a given model regarding the group. PSO, first proposed in 1995, is still a developing optimization technique. It originated from simulating social behavior, but has expanded into a versatile method for solving complex problems. Some features that attracted scientists included the underlying rules that allowed large flocks of birds to gather and move synchronously, frequently replacing way quickly, scattering including regrouping, etc. For certain primary considerations, bird gathering and fish schooling did some regarding the behavioral models which were claimed to be copied. The group distribution about learning amongst the members suggests an evolutionary benefit. Interacting also distributing the data among their community particles, identified since particular best or local best for comparing their neighbor best. At every step of the process, the global best resolution reached during this entire swarm remains refreshed.

"Particle Swarm Optimization (PSO) encompasses various techniques, PSO techniques such as Canonical PSO, Hierarchical PSO (HPSO), Time varying acceleration coefficient (TVAC) PSO, Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients (HPSO-TVAC), Stochastic inertia weight (Sto-IW) PSO, and Time-varying inertia weight (TVIW) PSO are as follows: Canonical PSO excels in solving Expectation-Maximization (EM) optimization problems. In K-means clustering, Canonical PSO-based algorithms have been evaluated on diverse datasets, including air contamination, customer data, wine quality, and transport datasets. Evaluations compared Canonical PSO-based K-means to standard K-means, simple PSO-based K-means, DBSCAN, and hierarchical clustering methods, using inter-class and intra- class indices as performance metrics." The outcomes were then compared to assess the effectiveness and accuracy of each method in different contexts. A hierarchical variant regarding this particle swarm optimization technique is termed as H-PSO. In H-PSO the particles are made into a powerful position that is applied to decide a local area association. In H-PSO all particles are coordinated in a progression that portrays the local area development. Every molecule is adjoined to itself and the parent in the pecking order. In this paper we study ordinary tree-like chains of command, i.e., the fundamental geography is an (nearly) standard tree. The chain of command is characterized by the stature h, the degreed, i.e., the most extreme (0ut) level of the inward nodes, and the all out number of nodes m of the comparing tree. In H-PSO every one of the particles are coordinated in a tree in such a way that shapes the order so every hub of the tree contains precisely one molecule. To give the best particles in the multitude high/raised impact particles move up and descending the progression. In the event that a molecule at a youngster hub has discovered an answer that is superior to the best so far arrangement of the molecule at the beginning hub the two particles are exchanged. One inspiration for the overview of H-PSO is that it offers a reformist neighborhood yet because of the fixed tree structure in H-PSO the new speed and area of every molecule can be assessed quick. We think about alterations of H-PSO where the greatest speed increase of a particle depends on its location in the hierarch. The time-varying acceleration coefficient is the effect of a cognitive element on new velocity decreases, concurrently the impact of a social component on new speed upgrades by changing the speedup coefficient “b1 and b2” with time. It presents an alternate heuristic calculation for deciding economic dispatch (ED) issues, by working emphasis/repetition molecule swarm streamlining with the the time-varying acceleration coefficients (IPSO-TVAC) technique is an effective approach for solving complex optimization problems, particularly in scenarios involving economic dispatch (ED). Due to the presence of valve-point effects and prohibited operating zones (POZs) in the cost functions of generating units, the ED problem is inherently nonlinear and non-convex. The complexity further increases when transmission losses are considered. The proposed IPSO-TVAC technique has been rigorously validated through comprehensive experiments conducted on three different test systems, accounting for factors such as valve-point effects, POZs, ramp rate constraints, and transmission losses.

The numerical results demonstrate that the IPSO-TVAC method exhibits excellent convergence properties. Moreover, the generation costs obtained using this technique are significantly more economical compared to other optimization algorithms reported in the recent literature. This indicates that the IPSO-TVAC approach not only provides a robust solution to the ED problem but also offers cost-effective performance advantages. The Inertia Weight parameter controls the degree to which a particle's momentum is preserved from one time step to the next. By introducing stochasticity into the Inertia Weight, the PSO algorithm gains enhanced convergence properties and improved robustness. This modified algorithm is then applied to clustering tasks, where its performance is evaluated using various UCI data sets. The results of this experimentation reveal that the stochastic PSO-based clustering algorithm surpasses traditional K-means clustering in terms of quantization error, while also demonstrating insensitivity to particle swarm size, making it a reliable and efficient clustering solution.

The inertia weight is typically set to decrease linearly over time, which can lead to a limited exploration of the search space. This poses a dilemma: if the search space is too narrow in the early stages, the algorithm may miss the optimal solution; if it's too broad, the search process slows down. To address this, researchers have proposed using a PSO algorithm with a stochastic inertia weight, where the weight is randomly generated from a probability distribution. This allows for more flexibility and adaptability in the search process. Unlike the basic PSO algorithm, which uses a fixed inertia weight, the stochastic version uses different random number distributions to execute modified PSO calculations, enabling more effective exploration of the search space. Normal irregular number examples incorporate uniform request, ordinary appropriation, and Poisson distribution. Experiments expose that the PSO algorithm with stochastic inertia weight is definitely more useful in combining rate than the essential PSO calculation. At the point when the ordinary example irregular latency weight is chosen, the calculation's worldwide examination information is more valuable than the calculations with other distributed arbitrary inertia weight. Time-varying inertia weight is defined as the inertia weight linearly decreases concerning the time (iteration number). In the initial phases of the examination cycle, a large inertia weight is generally recommended to enhance global exploration, allowing the algorithm to search broadly across new regions of the solution space. Conversely, in the later stages, the inertia weight is reduced to facilitate local exploitation, which helps in fine-tuning around the best-performing regions identified earlier. This adaptive adjustment of inertia weight helps balance exploration and exploitation, improving the overall effectiveness of the optimization process. The mathematical representation for the same is given as follows: Where is the initial value of the inertia weight is the final value of the inertia weight. Iter is the current iteration. Max iter is the maximum number of allowable iterations. inertia weight = $w^1 - w^2$maxiter − iter/iter + w^2

The main advantages of the PSO algorithm are reviewed as an easy concept, simple implementation, robustness to measure parameters, and computational ability when confronted with numerical algorithms and additional mheuristic optimization procedures. It is amazingly simple for implementing and has few parameters to adjust. It is robust and takes short computation times with an ability of parallel computation of different particles at the same time. These are less possibility of getting the wrong position regardless of the damaging tissue pattern. It has a higher probability and efficiency of finding the location of abnormalities in the brain (global optima).Particles in the PSO can coverage fast without any overlapping and mutation and give the location of abnormality efficiently.

The issue of undesirable increment in dimension is firmly identified with others. That was to obsession of assessing/recording data at a far granular level then it was done in past. There has been a huge addition in the manner sensors are being used in the business. These sensors continually record data and store it for assessment at a later point. Hence, highlight decrease began acquiring significance recently because of a flood in information. The reduction of dimensionality plays a crucial role in enhancing AI productivity, optimizing information sampling, improving pattern recognition accuracy, and ensuring the effectiveness of data mining. By decreasing the number of irrelevant or redundant features, models can focus on the most significant variables, thereby improving their ability to generalize from data and recognize patterns accurately. This process also contributes to better computational efficiency, as it simplifies the data structures involved, reducing the processing time and resource consumption. Moreover, it refines the quality of the analysis by concentrating on the most relevant data points, thereby enhancing the reliability of the results to a considerable extent. Some of the typical applications are discussed. The mining of quality articulation profile information can recognize cancer types. Bearings are important assets for most industrial applications. The non-damaging finding of these components needs a precise and solid obtaining of its dynamic vibration signals influenced by noise and the other part of system, such as, gears, shafts, and so on. For the early detection of bearing degradation we often use feature reduction. Effective data mining solutions have for since quite a while ago been expected in Customer Relationship Management (CRM) to precisely predict customer behaviour. It categorizes or classifies the documents (e.g. Politics, Sport, etc.) based on the common characteristics among the documents. In large multimedia databases, content based image and video retrieval use good data structures for similarity search and indexing.

It groups together microarray data and to reduce the number of features. SVD(Singular Value Decomposition) and other form of feature reduction methods are used to detect the face of a certain person from large database of facial data. Human view of dimensions is typically constrained to a few degrees. Any further increment in the quantity of dimensions as a rule prompts the difficulty in visual creative mind for any individual. Henceforth, machine learning specialists normally need to defeat the scourge of dimensionality in high dimensional feature sets with dimensionality reduction techniques. Intrusion detection is the way toward observing and investigating the occasions happening in a computer system so as to identify indications of security issues. E-mail classification is a significant way to deal with distinguish those spam emails. In light of various AI calculations, a novel semantic-based methodology for email is commonly used. The methodology investigations the substance of the email and assigns out a load to each term that can help in classifying it into spam or ham email. It applies this system both to the classification of malware and the recognizable proof of malware from a set joined with clean-ware.

Kennedy and Eberhart introduced the initial application of Particle Swarm Optimization (PSO) in 1995, focusing on neural network training and providing the original computational framework. Since then, PSO has been effectively applied across a broad spectrum of fields, including telecommunications, system control, data mining, power systems, design, combinatorial optimization, signal processing, network training, and numerous other areas. Nowadays, PSO algorithms have evolved to address constrained problems, multi-objective optimization issues, efficiently changing landscapes, and to identify multiple solutions. Originally, the primary PSO algorithm was predominantly used for solving unconstrained, single-objective optimization problems. Some notable applications of these advancements include image classification, image fusion, image noise cancellation, photo timestamp recognition, defect detection, character recognition, image registration, microwave imaging, pixel classification, object detection, and scene matching, 3D recovery with structured beam matrix.

The fields of robotic control encompass a wide array of applications, including the manipulation and control of robotic arms, motion planning, and execution, as well as robot operation. Key areas include swarm robotics, robot vision, and collective robotic search. Other significant domains involve transport robots, unsupervised robotic learning, voice-controlled robots, obstacle avoidance, unmanned vehicle navigation, and environment mapping. Additional significant applications encompass radar networks, Bluetooth networks, routing systems, auto-tuning for Universal Mobile Telecommunication Systems (UMTS), optimal equipment placement in mobile communications, TCP network control, and wireless networks. Further applications extend to economic dispatch problems, assembled and deferred broadcasting, bandwidth reservation, transmission network planning, and voltage regulation, all represent various domains where efficient network design, control, and optimization techniques are crucial for performance enhancement. Each of these areas requires precise management of resources, such as transfer speed, energy consumption, and signal distribution, to achieve optimal communication, connectivity, and system stability. Human tremor analysis for the diagnosis of Parkinson's disease, human movement biomechanics optimization, DNA motif detection, biomarker selection, drug design, biometrics and so on. Optimization of nuclear electric propulsion systems, optimization of internal combustion engines, floor planning, packing and knapsack, satisfiability, path optimization, layout optimization, Water quality prediction and classification, ecological models, time series prediction, meteorological predictions, electric load forecasting, predictions of elephant migrations, Design of neurofuzzy networks, fuzzy classification, clustering, dynamic clustering, data mining, feature selection. Key areas of focus include the optimal control and design of phased arrays, broadband antenna design and modeling, and reflector antennas. This also extends to the synthesis of antenna arrays. Additionally, important fields are VLSI design, RF circuit synthesis, worst- case electronic design, CMOS wideband amplifier design, power plants and systems, and customer satisfaction models. Moreover, advancements have been made in the optimization and application of on-chip inductors, circuit synthesis, AC transmission system control, and electromagnetic design. These developments also extend to microwave filter design and various other optimization applications.

Neuroimaging or mind imaging is the utilization of different methods in directly or by implication picture the construction, capacity of the sensory system. It is a similarly current order inside medicine, neuroscience, and treatment. Doctors who practice, execution and comprehension of neuroimaging in the clinical background are neuro-radiologists. Accurate acquisition technique of neuro-image with a better accuracy would be beneficial factor for neuro-radiologist to diagnose disorder. In the modern era, the understandings of human brain's self-adoption of its neurons according to its behavioral and environmental changes, commonly known as neuroplasticity, is essential for proper treatment. To address the structural and functional neuroplasticity basically neuroimaging is of two types:

Structural Magnetic Resonance Imaging (MRI) is a non-invasive neuroimaging technique that provides detailed information on the morphology of the nervous system, enabling the detection and characterization of macroscopic intracranial lesions, including tumors, cysts, and structural damage resulting from traumatic brain injury or other conditions.

Figure 2.4: Structural MRI

Functional neuroimaging techniques are employed to examine the brain's functional and metabolic activities, allowing researchers to detect subtle changes linked to neurological and psychiatric conditions, like Alzheimer's disease. These techniques facilitate the development of targeted therapies by providing detailed insights into brain function and pathology. Additionally, functional imaging plays a crucial role in the development of brain-computer interfaces and the study of cognitive processes.

Figure 2.5: Functional MRI

Functional neuroimaging has become a crucial tool in various scientific fields, including therapy, medicine, and neurology, with numerous institutions vying for access to cutting-edge technologies. To visualize brain activity, researchers employ a range of techniques, including: Computerized axial tomography (CAT Scan), Magnetic resonance imaging (MRI), Functional magnetic resonance imaging (fMRI), Positron emission tomography (PET), Single-photon emission computed tomography (SPECT), Magneto encephalography (MEG), Magnetic resonance spectroscopy (MRS), Transcranial magnetic stimulation (TMS), Event-related encephalograms (EEG). A set of feature reduction processes in data mining and machine learning involves numerous challenges. Some of the major difficulties in feature reduction are discussed below. Detection of geographic areas: Feature reduction based on local identities for bundle adjustment of images. Bio-medical application: Bio- medical data have the unusual feature comprising a very large numbers of variables. EMG signal detection: It is used to differentiate the useful data which is hidden in the surface EMG signal and to remove the undesired part and involvement. Information loss and PCA issues: It might prompt some measure of information loss.

Vol 45 No. 11
November 2024

PCA will in general find direct connections between's factors, which is once in a while unwanted. PCA flops in situations where mean and covariance are lacking to describe datasets. We may not know the number of head parts to keep by and by, some thumb rules is applied. Likewise, other algorithms there are certain drawbacks in this algorithm as well for particle swarm optimization. It is quite challenging sometimes to define initial parameters. Particles are not following any types defined rule for their initial parameters. They choose random movements. It is very difficult to define initial design parameter. It cannot work out on problems of scattering to avoid the curse of dimensionality. It can converge prematurely. It can be trapped into a local minimum especially with complex problems.

3. METHODOLOGY AND ALGORITHM OF PARTICLE SWARM OPTIMIZATION

Algorithm Steps:

1. For each particle (n = 1 to Q):
- Initialize position (xn) with a random vector within bounds (blo, bup)
- Set initial best-known position (pn) to current position (xn)
- If the fitness function (f) of pn is better than the swarm's best-known position (g), update g to pn
- Initialize velocity (vi) within bounds (-bup-blo, -bup-blo)
2. While termination criterion is not met:
- For each particle (n = 1 to Q):
- For each dimension (d = 1 to n):
- Generate random numbers (rp, rg) between 0 and 1
- Update particle velocity (vi) using rp and rg

Radiology particles are move to the tissue in the brain through the scalp and skull. The magnetic resonance images are named a grayscale value that ranges from 0 de- noted pure black and 255 denoted pure white which represents unwanted background signals and bones. Three parameters are used in particle swarm optimization i.e; current direction, neighbor best (regional best), and overall best. In individual trace of the method, damaged tissue detection task is used. To model the swarm, respective particle maneuver in multidimensional area give to the position x^n and velocity v^t values which are extremely vulnerable on regional best x and overall best g information. The planned PSO

t n t

Algorithm is as follows :

Load swarm (Initialize x^n, v^n, x, g)

Loop: t t t t

For any particle in swarm Calculate the damage tissue detection method of particle in−terms of velocity and location. $v^n + 1 = wv^n + p1r1(x - x^n) + p\ r\ (n - x^n)$//velocity of next particle

t t t t

2 2 t

t

4. RESULT ANALYSIS

The Filter approach is more useful than the wrapper approach. The wrapper method is computationally intensive, and its results are highly dependent on the chosen classifier. In contrast, the filter approach offers a faster alternative, providing a more efficient way to select features without relying on a specific classifier. When dealing with a vast number of features, filters provide a crucial initial step in assessing feature importance. They help eliminate the majority of irrelevant features, making it possible to apply more comprehensive feature selection methods subsequently. PSO method performs the search of the optimal solution through agents, referred to as particles. Here we have experimented on 5000 particles and 1000 iterations. Experimenting on higher number of particles tend to give much accurate result. First data has been updated in a seven columns matrix where each row has been assigned to one of the particles. Initially columns have been updated with its row number. The trajectory of each particle is influenced by both stochastic and deterministic elements. Each particle's movement is dictated by its own 'best' encountered position and the overall group's 'best' position, though it retains a degree of randomness. A particle *'i'* is characterized by its position vector p_n and its velocity vector v_n. With each iteration, the particle updates its state based on the newly calculated velocity. Through feature reduction I decreased the dataset so unwanted data were removed, so we have less data's and for time consumption we used feature reduction in PSO. Input of our thesis work is assigned for each particle with respect to its positional value in the matrix. Fetching of the input data's from a medical image will be taken into consideration in my future work. In figure 2 each of the particle’s initial position has been plotted in the axial boundary of -1000 to 5000 for each axis. As per initial assignment of columns with its row value the graph more tends to have a linear progression of slope 45. [?]. Figure 4.1 Initial distribution of 5000 particles in axial mapping of [- 1000 5000].

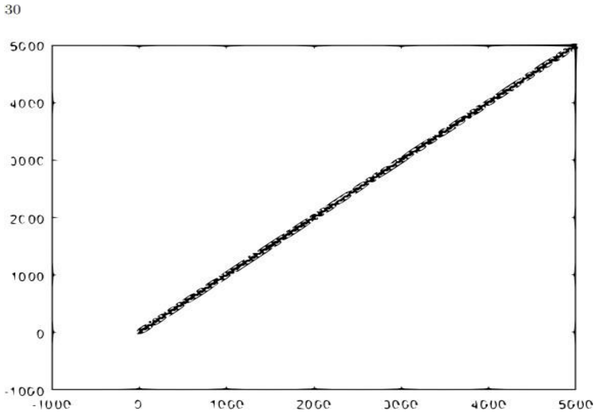

Figure 2.6:

Initial distribution of 5000 particles in axial mapping of [-1000 5000] In figure 4 the final scattering of particle’s position has mapped on the graph of axial boundary of [-1000 5000] in both X and Y axis. It is clearly visible that the density of particles is more on the axial boundary of [0 1500] in both X and Y axis. This position on brain cell is one millionth part of a brain cell which tend to give us very precise result on examination of abnormality. Figure 4 Result of dispersion of 5000 particles after 1000 iterations in axial mapping of [-1000 5000].

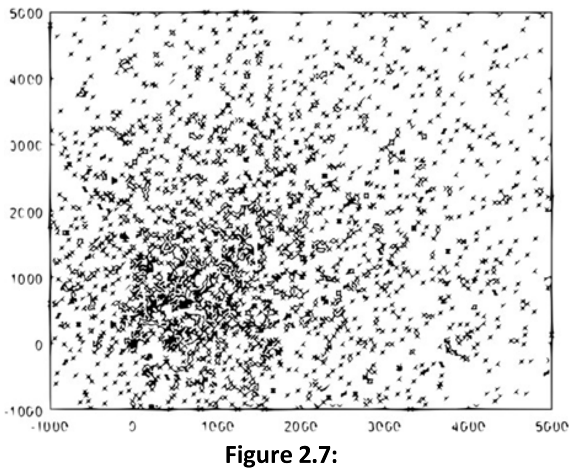

Figure 2.7:

In both figure 5 and 6, the figure 4 has zoomed out where each unit value is equal to the unit value multiplied with 10. Both figures give us a precise and pinpoint accuracy to detect abnormal tissue within the brain cells, and neuron system. [?]. Figure 4 Zoom out of Result of dispersion of 5000 particles after 1000 iterations.].

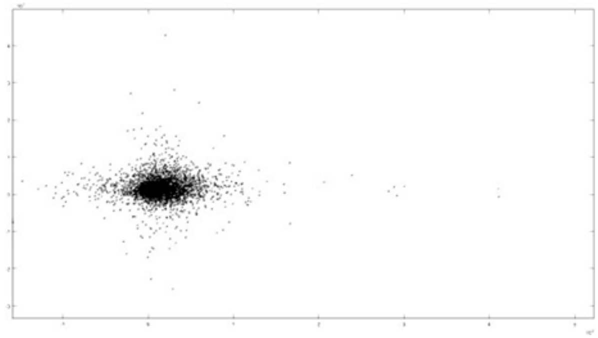

Figure 2.8:
[?]. Figure 4 Zoom out of Result of dispersion of 5000 particles after 1000 iterations.].

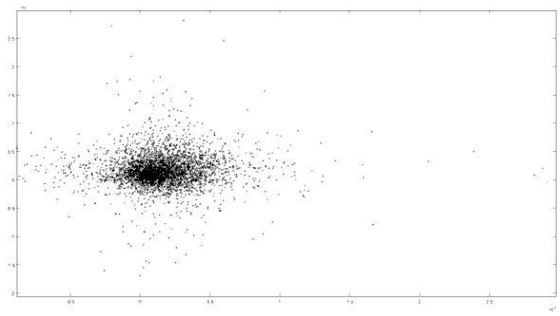

Figure 2.9:

Wrapper technique looks for an ideal element subset tailored to a specific calculation and an area. We study the strengths and shortcomings of the wrapper approach. Filtering performs feature selection as a preliminary step without the need for an inductive process. The general properties of the training material are used to select features (e.g. distance or class statistics). This method is faster than the wrapper approach and more efficient because it does its own inductive algorithm. However, it tends to select subsets with more features (or even all features), so there needs to be a threshold for selecting subsets.

5. CONCLUSION OF OUR WORK

The primary objective of feature selection algorithms is to develop a computationally efficient strategy that effectively addresses the challenge of identifying the most relevant features from a dataset. By focusing on selecting a subset of features that contributes most to the predictive accuracy of a model, these algorithms aim to enhance model performance, reduce computational costs, and mitigate the risk of over fitting. The goal is to strike a balance between retaining meaningful information and discarding irrelevant or redundant data, thereby simplifying the model without compromising its ability to generalize to new data. This research paper provides an in-depth review of the literature, concentrating on two widely used feature selection methodologies: the filter approach, which evaluates features independently, and the wrapper approach, which assesses features based on their impact on model performance. The filter method proves to be more advantageous compared to the wrapper method. The wrapper technique is notably time-intensive, with its outcomes heavily reliant on the chosen classifier. In contrast, the filter method offers a swift alternative, although it operates independently of any specific classifier. The findings of this study confirm that feature selection methods play a crucial role in boosting the accuracy and efficiency of learning algorithms. By effectively selecting and retaining the most informative features and discarding redundant or irrelevant ones, feature selection methods not only improve model performance but also offer a deeper understanding of the methodology driving each algorithm's feature selection decisions. Tumor detection is a great difficulty due to complex brain structures. The granted methodology is based on PSO to explicitly segment the entire tumor region only. While manual brain segmentation offers high reliability, it is hindered by its time-consuming nature and human bias. To address these limitations, computerized methods can be employed to streamline the process. The effectiveness of computerized MRI brain image segmentation for tumor identification can be evaluated based on two key factors: processing time and segmentation accuracy.

.

REFERENCES

[1] Chandrashekar, Girish, and Ferat Sahin.”A survey on feature selection methods.” Computers & Electrical Engineering

[2] 40.1 (2014): 16-28.

[3] Cheng, Zhun, and Zhixiong Lu.”A novel efficient feature dimensionality reduction method and its application in engineering.” Complexity 2018 (2018).

[4] S´anchez-Maro˜no, Noelia, Amparo Alonso-Betanzos, and Mar´ıa Tombilla-Sanrom´an.”Filter methods for feature selection–a comparative study.”International Conference on Intelligent Data Engineering and Automated Learning.Springer, Berlin, Heidelberg, 2007.

[5] Jovi´c, Alan, Karla Brki´c, and Nikola Bogunovi´c.”A review of feature selection methods with applications.”2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO). Ieee, 2015.

[6] Konstantinos E. Parsopoulos and Michael N. Vrahatis, Multi-Objective Particles Swarm Optimization Approaches, 2008.

[7] Lihua Gong and Zhengyuan Jia, ”Multi-criteria Human Resource Allocation for Optimization Problems Using Multi- objective particle Swarm Optimization Algorithm,” International Conference on Computer Science and Software Engineering, vol. 1, pp. 1187-1190, 2008.

[8] Kennedy, James, and Russell Eberhart. ”Particle swarm optimization.” Proceedings of ICNN’95-international conference on neural networks. Vol. 4. IEEE, 1995.

[9] Vijay, Vasupradha, A. R. Kavitha, and S. Roselene Rebecca. ”Automated brain tumor segmentation and detection in MRI using enhanced Darwinian particle swarm optimization (EDPSO).” Procedia Computer Science 92 (2016): 475-480.

[10] Al-Tamimi, Mohammed Sabbih Hamoud, and Ghazali Sulong. ”TUMOR BRAIN DETECTION THROUGH MR IMAGES: A REVIEW OF LITERATURE.” Journal of Theoretical Applied Information Technology 62.2 (2014).

[11] Balthazard, Pierre A., and Robert W. Thatcher. ”Neuroimaging modalities and brain technologies in the context of organizational neuroscience.” Organizational neuroscience. Emerald Group Publishing Limited, 2015.

[12] Sahab, Mohammed Ghasem, Vassili V. Toropov, and Amir Hossein Gandomi. ”A review on traditional and modern structural optimization: problems and techniques.” Metaheuristic

applications in structures and infrastructures (2013): 25-47.

[13] Zeng, Nianyin, et al. ”A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer’s disease.” Neurocomputing 320 (2018): 195-202.

[14] Thirioux, B´erang`ere, et al. ”The cognitive and neural time course of empathy and sympathy: an electrical neuroimaging study on self–other interaction.” Neuroscience 267 (2014): 286-306.

[15] Ahmed, Md Rishad, et al. ”Neuroimaging and machine learning for dementia diagnosis: Recent advancements and future prospects.” IEEE reviews in biomedical engineering 12 (2018): 19-33.

[16] Yang, Shih-Ting, et al. ”Discrimination between Alzheimer’s disease and mild cognitive impairment using SOM and PSO-SVM.” Computational and mathematical methods in medicine 2013 (2013).

[17] Sharif, Muhammad, et al. ”An integrated design of particle swarm optimization (PSO) with fusion of features for detection of brain tumor.” Pattern Recognition Letters 129 (2020): 150-157.

[18] Remamany, Krishna Priya, et al. ”Brain tumor segmentation in MRI images using integrated modified PSO-fuzzy approach.” Int. Arab J. Inf. Technol. 12.6A (2015): 797-805.

[19] Karthik, R., et al. ”Neuroimaging and deep learning for brain stroke detection-A review of recent advancements and future prospects.” Computer Methods and Programs in Biomedicine (2020): 105728.

[20] Sivakumar, V., N. Janakiraman, and S. Naganandhini. ”Machine Learning Approach improves the Quality of the MRI Images in Tumor Detection and Diagnosis: A PSO based Cluster Analysis.” 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW). IEEE, 2018.

[21] Taie, Shereen A., and Wafaa Ghonaim. ”CSO-based algorithm with support vector machine for brain tumor’s disease diagnosis.” 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE, 2017.

[22] Shehab, Mohammad, Ahamad Tajudin Khader, and Mohammed Al-Betar. ”New selection schemes for particle swarm optimization.” IEEJ Transactions on Electronics, Information and Systems 136.12 (2016): 1706-1711.

[23] Mahalakshmi, S., and T. Velmurugan. ”Detection of brain tumor by particle swarm optimization using image segmentation.” Indian Journal of Science and Technology 8.22 (2015): 1.

[24] Lin, Juan, Chenglian Liu, and Yongning Guo. ”Particle swarm optimization and its application in MEG source localization using single time sliced data.” AIP Conference Proceedings. Vol. 1618. No. 1. American Institute of Physics, 2014.

[25] 24] Ozekes, Serhat, and Turker Tekin Erguzel. ”Artificial intelligence ap- ¨ proaches in psychiatric disorders.” (2014).

[26] Farouk, Y., Rady, S. (2018, September). Supervised classification techniques for identifying Alzheimer’s disease. In International Conference on Advanced Intelligent Systems and Informatics (pp. 189-197). Springer, Cham.

[27] Tyagi, Vaishali, and Anil K. Ahlawat. ”To detect normal and abnormal neurological disorder of mri image in human using convolutional neural netwok.” 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). Vol. 1. IEEE, 2019.

[28] Xin, Jiang, et al. ”Brain differences between men and women: Evidence from deep learning.” Frontiers in neuroscience 13 (2019): 185.

[29] Ma, Xinpei, et al. ”Brain response pattern identification of fMRI data using a particle swarm optimization-based approach.” Brain Informatics 3.3 (2016): 181-192.

[30] Ibrahim, Shafaf, Noor Elaiza Abdul Khalid, and Mazani Manaf. ”Evaluation method for MRI brain tissue abnormalities segmentation study.” 15th WSEAS International Conference on Computers. 2011.

[31] Shaikh, Tawseef Ayoub, and Rashid Ali. ”Automated atrophy assessment for Alzheimer’s disease diagnosis from brain MRI images.” Magnetic resonance imaging 62 (2019): 167-173.

[32] Ait-Aoudia, Samy, El-Hachemi Guerrout, and Ramdane Mahiou. ”Medical image segmentation using particle swarm optimization.” 2014 18th International Conference on Information Visualisation. IEEE, 2014.

[33] Sountharrajan, S., P. Thangaraj, and E. Suganya. ”PARTICLE SWARM OPTIMIZED FEATURE

SELECTION FOR ALZHEIMER CLASSIFICATION.” IIOAB JOURNAL 7.9 (2016): 787-794.

[34] Blum, Christian, and Daniel Merkle, eds. Swarm intelligence: introduction and applications. Springer Science Business Media, 2008.

[35] Mwangi, Benson, Tian Siva Tian, and Jair C. Soares. ”A review of feature reduction techniques in neuroimaging.” Neuroinformatics 12.2 (2014): 229- 244.

[36] Phinyomark, Angkoon, Pornchai Phukpattaranont, and Chusak Limsakul. ”Feature reduction and selection for EMG signal classification.” Expert systems with applications 39.8 (2012): 7420-7431.

[37] Brunzell, H˚akan, and Jonny Eriksson. ”Feature reduction for classification of multidimensional data.” Pattern Recognition 33.10 (2000): 1741-1748.

[38] Lam, Savio LY, and Dik Lun Lee. ”Feature reduction for neural network based text categorization.” Proceedings. 6th international conference on advanced systems for advanced applications. IEEE, 1999.

[39] Heisele, Bernd, et al. ”Hierarchical classification and feature reduction for fast face detection with support vector machines.” Pattern recognition 36.9 (2003): 2007-2017.

[40] Kondo, Masanari, et al. ”The impact of feature reduction techniques on defect prediction models.” Empirical Software Engineering 24.4 (2019): 1925- 1963.

[41] Kalivas, John H., ed. Adaption of simulated annealing to chemical optimization problems. Elsevier, 1995.

[42] Datti, Rupali, and Bhupendra Verma. ”B.: Feature reduction for intrusion detection using linear discriminant analysis.” International Journal on Engineering Science and Technology. 2010.

[43] Kalaivani, P., and K. L. Shunmuganathan. ”Feature reduction based on genetic algorithm and hybrid model for opinion mining.” Scientific Programming 2015 (2015).

[44] Aksoy, Selim. ”Feature reduction and selection.” Department of Computer Engineering Bilkent University, CS 551 (2008).

[45] R¨uger, Stefan M., and Susan E. Gauch. Feature reduction for document clustering and classification. Imperial College of Science, Technology and Medicine, Department of Computing, 2000.

[46] Kumar, Gulshan, Krishan Kumar, and Monika Sachdeva. ”An empirical comparative analysis of feature reduction methods for intrusion detection.” International Journal of Information and Telecommunication Technology 1.1 (2010): 44-51.

[47] Mirniaharikandehei, Seyedehnafiseh, et al. ”A novel feature reduction method to improve performance of machine learning model.” Medical Imaging 2021: Computer-Aided Diagnosis. Vol. 11597. International Society for Optics and Photonics, 2021.