International Journal of Nanotechnology

ISSN online: 1741-8151 - ISSN print: 1475-7435
https://www.inderscience.com/ijnt

Analysis of high dimensional data using feature selection models

Shubham Mahajan, Amit Kant Pandit

DOI: 10.1504/IJNT.2023.10056472

Article History:

Received:	09 February 2021
Last revised:	16 May 2021
Accepted:	07 June 2021
Published online:	31 May 2023

Analysis of high dimensional data using feature selection models

Shubham Mahajan* and Amit Kant Pandit

School of Electronics & Communication,
Shri Mata Vaishno Devi University,
Katra – 182320, India
Email: mahajanshubham2232579@gmail.com
Email: amitkantpandit@gmail.com
*Corresponding author

Abstract: The determination of features assumes a significant part in enhancing the output of AI models, limiting the computational time taken to make a learning model and improving the exactness of the learning cycle. Hence, analysts give more consideration to the determination of features to expand the exhibition of AI calculations. The choice of the proper technique for the determination of features is significant for a specific AI task through high-dimensional information. It is subsequently important to complete an examination on various strategies for character determination for the exploration network, specifically to improve effective techniques for choice. Method for choosing features to improve the effectiveness of AI undertakings for high-dimensional information. This paper gives the whole writing survey of the different techniques for choosing features for high-dimensional information to accomplish this target.

Keywords: feature selection; signal processing; artificial intelligence; high dimensional data; classification; signal processing; learning models; classifiers.

Reference to this paper should be made as follows: Mahajan, S. and Pandit, A.K. (2023) 'Analysis of high dimensional data using feature selection models', *Int. J. Nanotechnol.*, Vol. 20, Nos. 1/2/3/4, pp.116–128.

Biographical notes: Shubham Mahajan pursuing his PhD from Shri Mata Vaishno Devi University (SMVDU), Katra (India). He received his BTech and MTech degree from the Department of Electronics and Communication Engineering at Baba Ghulam Shah Badshah University and Chandigarh University, Rajouri and Ghauran (India) in 2017 and 2019, respectively. He received the 'Best Research Paper Award' for Paper titled "Design and Investigation of Multiple TX/RX FSO System under Different Weather Conditions" during International Conference (Springer, LNEE) for his research work. He has published more than nine research papers in national and international journal/conferences. He has been Campus Ambassador for IEEE, IIT Bombay, Kanpur, Varanasi and Delhi and various MNC's. He is Google Ads Fundamentals, Campaign Manager, Fundamentals of Digital Marketing, Analytics Individual Qualification, Creative Certification, Display & Video 360 Certification and Mobile Sites Certification Exam Certified and graduate student member of IEEE. He received the Best Student Award from Dr. Kalam Education Trust and IEEE Region-10 Travel Grant Award. His main research interest includes study of image and video processing, wireless communication, optical fibre communication & losses, optical amplifier (EDFA and Raman amplifier), SS-WDM, FSO and Radio over fibre etc.

Amit Kant Pandit is working as an Associate Professor and is Ex-Hod, DECE, SMVDU. He is senior member of IEEE, Member MIR labs and has 18 years of academic experience.

1 Introduction

In the modern world, the administration of tremendous information is a troublesome undertaking for analysts since information is amassed through an assortment of information assortment strategies and gadgets. This accumulated wide crude information diminishes the presentation of AI calculation based on over-fitting, investing a great deal of energy planning AI modes and corrupting their precision as crude information is noisy and has additional features called high-dimensional information. By and large high-dimensional information contains unimportant and repetitive capacities. The insignificant features can be ignored for the learning cycle. A selection of features can likewise settle these problems. The selection of features is a technique to take out repetitive and extra features from dataset to improve productivity of AI calculations. Choosing a capacity is otherwise called choosing a variable or choosing a characteristic. Qualities are otherwise called factors or characteristics. AI calculations can be inexactly isolated into two classes: one is an administered learning calculation, and the other is an unaided learning calculation. Managed learning calculations, learn characterised information and make learning models called classifiers. The classifier is utilised for order or forecast to characterise or anticipate class-mark of unlabeled information. Calculations incline toward plain information and construct learning models called bunching models. Bunch models are utilised for collection or classifications the information to foresee or characterise their classification or bunch. Generally, include determinations are utilised for supervised learning methods since they have experienced high-dimensional space. Along these lines, this work presents a full writing survey of various techniques for choice of features for high-dimensional information.

1.1 Motivation

- Multimedia technology (Image and Video) is increasing interest, and everyone involved in technology must understand the possibilities it presents and the current limitations of technology.
- The problem of efficient multimedia technology (Image and Video) management is an important issue.
- The semantic gap in machine learning methods, such as classification, can be enhanced.
- Efficiency: reducing the computational time and processing.
- Segmentation and Feature selection on trivial images are difficult tasks (computational time and processing) in image processing [1].

On a fundamental level, this importance of features is decided by comparing its relationship with class variables to other features. These features can and must be removed before training a Machine Learning model to compensate for the loss from overfitting. This being said, the curse mentioned above will also be cured as not much information will be lost [2] upon removal of these non-important features. These techniques are most useful in cases where there are too many features but relatively few samples.

Feature selection is an optimisation problem. The main steps are shown below:

a search space of feature subsets

b pick's subset that is ideal or approximately ideal w.r.t. some objective function.

One of such algorithms would be the exhaustive search, which is simply testing iteratively every possible feature sub-space and evaluating each of them to find one with minimum error. The three types are filter-based, wrapper and embedded methods [3,4], as shown in Figure 1.

1 Wrapper method [3]

- It measures the 'utility' of features based on performance.
- It has a low speed.
- It has high accuracy.

Further divided into three categories:

- Step forward.
- Step backward.
- Exhaustive.

2 Filter method

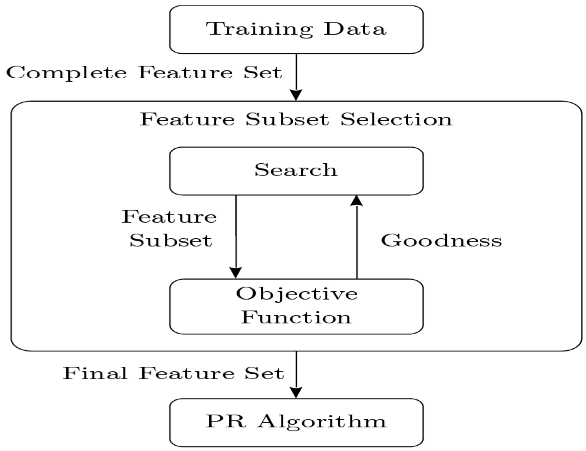
- It picks up essential properties of features.
- It has high speed as they do not involve training the models.
- It has low accuracy.

3 Embedded methods

- They perform feature selection during model training, i.e., why they are called embedded approaches.
- In embedded approaches, feature selection methods are combined as part of learning algorithm.

Apart from this, Filter-based methods utilise a pseudo-measure, unlike that in Wrapper methods. This proxy measure is computationally efficient and faster than the formerly discussed set of methods and is more valuable in very high dimensional data.

Figure 1 Main steps of feature selection



The paper is planned as Section 2 outlines technique of features selection and summarises the survey in the feature section. Section 3 shows the conclusion of the paper.

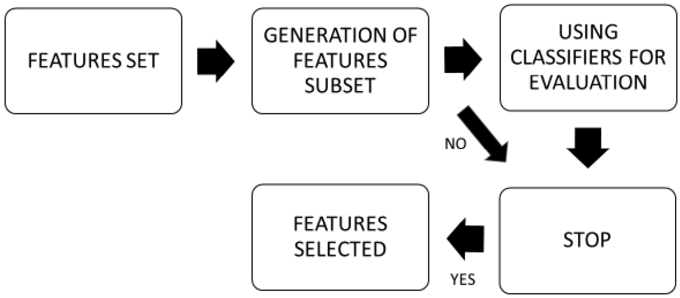
2 Feature selection

Feature selection is a method to extract obsolete and repetitive features from the information set to help the effectiveness of AI calculations and precision and time to develop a model. The feature selection method is separated into two gatherings, particularly component subset determination and highlight positioning techniques, contingent upon how the features are joined for assessment. The component subset choice strategy creates a potential number of mixes of highlight subsets utilising any of pursuit strategies, for example, covetous forward choice, eagerness in reverse disposal, and so forth, to test the individual element subset by a selection measurement, for example, consistency, co-relation and so on This methodology requires more space and computational intricacy because of the age and assessment of subsets [5].

In the feature selection process, every component is picked by a decision estimation; for instance, such as vulnerability, proportion, data and highest-level features are selected as important features by predefined limit. This strategy is computationally less expensive, and the unpredictability of space is not as much as that of the subset technique. It doesn't, nonetheless, manage excess capacities. Also, the feature selection process is separated into four gatherings, specifically covering, inserted filter, and half breed draws near, contingent upon how administered learning calculation is utilised in the feature selection process.

The wrapper technique incorporates regulated learning calculation to approve created feature subsets utilising both of the pursuit techniques, as appeared in Figure 2. High order precision is accomplished distinctly for a similar learning calculation received. As a result, it doesn't have a high consensus, and the computationally complex nature is greater than inserted and filter techniques.

Figure 2 Shows selection of features by wrapper method



The embedded strategy utilises supervised learning calculation for the feature selection and delivers improved exactness for learning calculation utilised in the selection process. As a result, it doesn't have a solid consensus and is practical than the filter, which is less than wrapper form.

Filter method chooses features that are not influenced by supervised learning calculation, as shown in Figure 3. It consequently works for all orders of calculation and accomplishes additional consensus with reduced computational multifaceted nature than wrapper and embedded approaches. It is in this way ideal for high-dimensional space. The combination of these two methods is called the hybrid [6] method.

Figure 3 Shows selection of features by filter method

Since the determination of features is utilised in various AI applications, the researchers have extraordinary records. Feature Selection is a pre-processing procedure used to pick features from the dataset by removing unnecessary and excess features to upgrade the proficiency of AI calculations. The feature selection process is separated into various methodologies relying upon how features are joined for assessment in feature selection method and how supervised learning calculation is utilised to assess features. This paper discusses different techniques for the determination of features and talks about their benefits and faults.

2.1 Selection by combination of features for evaluation

This section discusses different methodologies for collecting features depending on how features are merged for assessment by which the important features can be selected from the dataset. They are categorised into subset-based features and rank-based approaches.

2.1.1 Subset based method in feature

The features are joined as potential mixes of highlight subsets utilising both inquiry systems in the component-based subset strategy. The element subsets are then assessed utilising both measurable techniques or regulated learning calculations to notice every subset and main subset's significance as a huge component subset for the dataset. If the subset is assessed utilising a regulated learning calculation, this technique is called the covering strategy. The best model for a subset-based element strategy is the correlation-based subset feature selection (CRFS) [7]. In this strategy, two relationship tests are thought of; one is a component class connection, and the other is an element that includes connection. At first, N number of features are joined as possible mixes of highlight subsets utilising heuristic-based best-first hunt. Afterward, every subset is assessed with two connection measures alluded to above. A subset with a lower work highlights connection, and a higher capacity class relationship contrasted with other component subsets is viewed as chose significant element subset for order task. Liu and Setiono [7] projected a component subset-based element determination measure, to be specific consistency-based subset selection (COFS). This methodology utilises class consistency as an assessment measure to pick a critical component subset from dataset. These techniques are filter dependent strategies, as they don't utilise an administered learning calculation to approve subsets and utilise a factual measure to assess subsets of features.

Specifically, a thorough or full inquiry must create $2N$ subsets to produce the greatest number of conceivable subset mixes of highlights from N number of highlights for assessment. This thorough inquiry methodology is like this computationally outrageous, because of heuristic pursuit systems, for example, particle swarm optimisation (PSO), colony optimisation (ACO), tabu searching (TS), simulated annealing (SA), genetic algorithm (GA) etc. [8] a few analysts are utilising it to get the ideal arrangement by creating fewer highlights of the subsets. In the inquiry measure, the heuristic capacity acquires earlier information to guide the hunting cycle to deliver subsets, and these subsets are assessed utilising a regulated AI calculation. These variables make the capacity subset-based strategies computationally excessive, and they will be the covering approach.

A few analysts utilised a virtual research to create a subset of highlights for assessments. For instance, Lin [9] utilised a reenacted toughening search to produce subsets of highlights and assessed them by a regulated learning calculation, to be specific a back-spreading network (BPN) to pick a superior subset of highlights. Wu et al. [10] utilised recreated tempering-based scope of promoting applications. In many component choice techniques, a tab search is utilised for the age of subsets, for example, Shi et al. [11], which has built up a tab search-based element determination. In this cycle, the subsets created by the tab search are assessed utilising the grouping blunder measures to locate a superior subset of highlights. Tahir et al. [12] created include subsets utilising forbidden hunt, and afterward, these subsets are assessed utilising KNN with an ordering blunder as assessment rules to acquire a critical element subset.

The ant colony optimisation (ACO) was used as the basis for application text classification [13]. Kanan and Faez [14] recommended a technique for choosing highlights utilising an insect state improvement for the facial recognition framework. In this strategy, the closest neighbour classifier is utilised to assess created subset utilising subterranean insect settlement-based learning enhancement. Sivagaminathan and Ramakrishnan [15] have built up an insect state streamlining based choice of artificial neural networks (ANNs) for a clinical analytic technique. The made component subsets are approved utilising ANN in this cycle. Sreeja and Sankar [16] proposed a case-based example PMC [17] ACO-based feature choice.

In some feature examination works, a hereditary calculation is utilised to produce highlight subsets for assessment, and a directed AI calculation is utilised to assess the created subsets. Welikala et al. [17] assorted highlights are utilising a support vector machine (SVM) to mining a clinical dataset. A hereditary calculation and a fake neural organisation were used to order an electroencephalogram (EEG) signal [18]. Oreski and Oreski [19] recommended a framework for determining highlights dependent on a hereditary calculation with neural organisations to appraise credit hazard. A hereditary calculation with a help vector machine for hyper-otherworldly picture arrangement [20]. Das et al. [21] has detailed a hereditary calculation with machine-based vector uphold for the application for manually written digit acknowledgement. Wang et al. [22] utilised a hereditary calculation for subset age with a help vector machine in the information order application determination measure.

In literature, a few investigations utilised PSO [8] to produce include subsets and approve them by a managed AI calculation to group a critical component subset. Xue et al. [23] built up a PSO-based characterisation determination work. In this progression, PSO-created includes subsets are assessed utilising a managed learning calculation [23]. Chen et al. [24] proposed a strategy for choosing highlights utilising particle swarm optimisation for rest problem determination. A built-up of PSO-based component determination for land cover characterisation was used by Yang et al. [25]. The comprehensive or complete query items in high computational unpredictability create $2N$ subsets from N assessment highlights from subset-based element determination writing. This inquiry methodology can't be a superior alternative for high-dimensional space. Heuristic inquiry techniques can regularly increase computational unpredictability because of the requirement for earlier information and the requirement for each produced subset to make a characterisation model test them to acquire an iterative ideal subset of highlights reason these pursuit methodologies are not appropriate for high-dimensional space. Be that as it may, these heuristic inquiry techniques depend on a covering-based methodology. These strategies are consequently computationally exorbitant and can yield higher-order precision for specific arrangement calculations utilised to approve subset, so they can't accomplish high over-simplification.

2.1.2 Ranking based method in feature

In a feature-based methodology, each dataset is weighted depending on whether factual or data hypothetical measures and highlights are appraised depending on their weight. The higher positioned highlights are then chosen as critical highlights utilising predefined limit that regulates number of highlights to be chosen from the dataset. Chi-square-based element choice (CQFS) is the best model for a feature-based element framework. In this framework, Liu and Setiono [26] utilised a chi-square measurement test to quantify the heaviness of the highlights to rank them for the assortment of huge highlights. Additionally, hypothetical data measurements, such as data increase, symmetric vulnerability, gain proportion, and so forth, are utilised to weigh and rank the individual capacity.

Also, it is noticed that feature depend positioning methods use hypothetical data trials to weigh a single element by observing the importance between the single element and objective class. Like this, these strategies set aside less effort to run, yet don't dispose of excess highlights [7]. Feature-based methodologies receive a feature-based methodology since these methodologies don't need a directed learning calculation to decide the significance of the highlights. As an outcome, these strategies are autonomous of the managed learning calculation and, in this way, acquire over-simplification and less computational intricacy. In this way, a feature-based strategy for positioning might be a sensible decision for choosing critical highlights from a high-dimensional space with a viable repetition investigation instrument.

2.2 Utilising supervised machine learning in feature selection

This section discusses different methodologies for selecting features based on the machine learning method is utilised. They are classified as embedded, hybrid, wrapper and filter.

2.2.1 Wrapper method

Wrapper-based technique accomplishes highlight subsets utilising both hunt methods and tests these subsets utilising the administered learning calculation regarding arrangement blunder or precision [4]. The wrapper approach will, in general, be the 'beast power' approach. This strategy appears in Figure 3. Kohavi and John [27] have fabricated a wrapper approach technique for choosing huge highlights from the dataset. This methodology comprises a subset age web index and a subset assessment calculation. Besides, they look at the productivity of this methodology to order exactness with slope climbing and best-first inquiry methodologies utilising choice tree and guileless Bayes classifiers. In any case, they found that the wrapper approach had issues, for example, overhead looking, overfitting, and expanded runtime.

In this method, looking is overhead because the hunting technique doesn't have area data. To conquer the overhead pursuit time, Inza et al. [28] utilised Bayesian organisation calculation assessment to highlight subset choice utilising naive Bayes and ID3 (Iterative dichotomised 3). Specifically, inquiry strategy can prompt an expansion in computational intricacy as the preparation information is separated for assessment. To conquer this issue, Grimaldi et al. [29] utilised a successive inquiry conglomeration rule. Dy and Brodley [30] built up a covering-based way to deal with unaided picking up utilising a request acknowledgement (perceiving the number of groups in the information) with an expectation-maximisation (EM) bunching calculation dependent on the most machine learning (ML) measure. Aha and Bankert [31] gave a bar search and IB1 classifier covering based cycle. They likewise contrasted its proficiency and notable quest calculations for highlight determination, like sequential forward selection (FSS) and backward sequential selection (BSS).

Maldonado and Weber [32] have executed a wrapper model-based element determination by incorporating a support vector machine (SVM) with bit capacities. This methodology utilises a consecutive in reverse determination for the age of subset highlights, and these subsets are approved regarding a characterisation blunder to decide the best subset [32]. Gutlein et al. [33] utilised the hunt calculation, particularly ORDERED-FS, to lessen the overhead inquiry, which arranges the highlights as far as the re-replacement mistake to characterise their insignificance. Kabir et al. [34] has built up a useful way to deal with a constructive approach to collecting features (CAFS) using a neural network (NN). In this methodology, a connection measure is utilised to diminish excess in the pursuit procedure to support NN yield [34]. Stein et al. [35] recommended a streamlining-based subterranean insect settlement highlight choice with a wrapper technique.

Here, ant colony optimisation is utilised as an inquiry apparatus to limit search overheads, for example, daze search or forward choice or in reverse end search [35].

Selection method by using SVM and genetic algorithms were used for classifying image of hyperspectral characterises. Overfitting is illuminated by post-cutting, jitter, and early halting techniques in the covering approach. Post-pruning is completed while the choice tree is being created [36]. In the jitter approach, the loud information that makes the learning cycle more muddled is disposed of to coordinate the preparation information to kill the over-fitting [37]. The early halting methodology forestalls over-fitting utilising the neural organisation by halting the preparation stage when the yield on an approval set starts to disintegrate [38]. Analysts have attempted to limit over-fitting by early halting utilising the GAWES algorithm [39,40].

Moreover, it is seen that the wrapper-based techniques are liable to overhead looking, overfitting [41], and have large computational unpredictability with fewer consensus since they utilise supervised learning calculation to assess created subsets utilising inquiry strategy. Accordingly, these approaches are not the right decisions for high-dimensional space.

2.2.2 Embedded method

Embedded methodologies utilise some portion of the learning cycle of supervised learning calculation for the assortment of features. Embedded strategies limit computational costs comparative with the covering technique [42]. This embedded strategy is inexactly characterised by a pruning technique, an integrated mechanism, and a regularisation model. In pruning-based methodology, all highlights are first brought into the arrangement model preparing cycle and highlights with a lower relationship coefficient esteem are recursively eliminated utilising the SVM [18,38]. A piece of preparation cycle of C4.5 and ID3 is important for the incorporated component-based feature selection system [5]. Supervised learning calculations are utilised to pick a capacity. In the regularisation cycle, fitting mistakes are decreased utilising target capacities, and highlights with almost zero relapse coefficients are disposed off [43].

An embedded feature selection system for separating critical features from both manufactured and genuine world datasets. In their strategy, direct and non-straight SVMs are utilised in the selection process utilising DCA [44]. Xiao proposed an embedded system for choosing critical features from sound signs for the characterisation of feelings. This methodology was presented based on the idea of the mass-work proof hypothesis, and the main features discovered are gradually added for order [45]. Maldonado et al. [46] has built up an implanted strategy for choosing critical features from the imbalanced grouping information with a few target capacities.

Besides, it is noticed that embedded strategies are computationally more effective than wrapper techniques and computationally more expensive than filter strategies; thus, they can't be worthy decision for high-dimensional space and have low consensus because inserted strategies utilise regulated learning calculation [47].

2.2.3 Filter method

Filter-based techniques are self-governing of supervised learning calculation and subsequently, give more over-simplification and are computationally less expensive than coverings and implanted methodologies. Filter methods are adequate for preparing high-dimensional information instead of covering and implanted techniques.

Ordinarily, the selection process of features is pointed toward choosing proper features. The best model is Relief [48], which has been worked with a distance-subordinate measurement work that loads each element dependent on their significance (correlation) to the objective class. In any case, Relief is wasteful as it can just arrange with two-class issues and doesn't manage excess capacities. The appropriate example is Relief [49], which is equipped to tending to multi-class issues and working with inadequate and loud datasets. It doesn't, in any case, erase repetitive capacities.

2.2.4 Hybrid method

Hybrid methods typically are mixture of filter and wrapper-based methodologies [50]. By and large, the preparing of high-dimensional information is a troublesome undertaking with the wrapper strategy. In this way, the creators Bermejo et al. [51] have constructed a mixture feature choice technique called the filter wrapper method. In this methodology, a factual measure was utilised to rank features based on significance and higher positioned features. The wrapper strategy with the end goal that the quantity of appraisals required for the wrapper technique is direct. Complexity for medical data is decreased by the method of hybrid [51]. Ruiz et al. [52] has built up a quality (feature) determination calculation to choose huge qualities for the clinical demonstrative technique. They utilised a measurable positioning way to deal with sifting includes high-dimensional space and the separated feature taken care of into wrapper method. This mix of filter and wrapper method has recognised significant qualities that cause disease in the analysis cycle of data [52,53].

3 Conclusion

This paper examines some systems for the selection of features proposed by various researchers. Past exploration has indicated that include-based positioning strategies are superior to subset-based techniques as far as memory space and computational intricacy, and positioning-based strategies don't limit excess. Besides, the covering, implanted, and crossover approaches are computationally wasteful than the channel framework and have a powerless consensus. The determination of features for high-dimensional data can consequently be set up utilising a filter method with a ranking technique for the important selection of features in high-dimensional space. Moreover, a clustering procedure might be actualised to address the disadvantages.

Future scope: In the future, we will be proposing an efficient image segmentation method with different feature selection methods, which will help in reducing the computational time that is need of hour in real-time applications.

References

1 Mahajan, S., Mittal, N. and Pandit, A.K. (2021) 'Image segmentation using multilevel thresholding based on type II fuzzy entropy and marine predators algorithm', *Multimed. Tools Appl.*, Vol. 80, pp.19335–19359, https://doi.org/10.1007/s11042-021-10641-5

2 Bermingham, M.L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campbell, H., Wright, A.F. *et al.* (2015) 'Application of high-dimensional feature selection: evaluation for genomic prediction in man', *Sci. Rep.*, Vol. 5, No. 1, pp.1–12.

3 Guyon, I. and Elisseeff, A. (2003) 'An introduction to variable and feature selection', *J. Mach. Learn. Res.*, Vol. 3, Mar, pp.1157–1182.

4 Chandrashekar, G. and Sahin, F. (2014) 'A survey on feature selection methods', *Comput. Electr. Eng.*, Vol. 40, No. 1, pp.16–28.

5 Asir Antony Gnana Singh, D., Appavu Alias Balamurugan, S. and Jebamalar Leavline, E. (2016) 'Literature review on feature selection methods for high-dimensional data', *Int. J. Comput. Appl.*, Vol. 975, p.8887.

6 Remeseiro, B. and Bolon-Canedo, V. (2019) 'A review of feature selection methods in medical applications', *Comput. Biol. Med.*, Vol. 112, p.103375.

7 Liu, H. and Setiono, R. (1996) 'A probabilistic approach to feature selection-a filter solution', *ICML*, Vol. 96, pp.319–327.

8 Lisnianski, A., Frenkel, I. and Ding, Y. (2010) 'Aging multi-state systems', *Multi-state System Reliability Analysis and Optimization for Engineers and Industrial Managers*, Springer, London, pp.273–319.

9 Lin, S.W., Chen, S.C., Wu, W.J. *et al.* (2009) 'Parameter determination and feature selection for back-propagation network by particle swarm optimization', *Knowl. Inf. Syst.*, Vol. 21, pp.249–266.

10 Wu, S., Hu, Y., Wang, W., Feng, X. and Shu, W. (2013) 'Application of global optimization methods for feature selection and machine learning', *Math. Probl. Eng.*

11 Shi, L., Wan, Y., Gao, X. and Wang, M. (2018) 'Feature selection for object-based classification of high-resolution remote sensing images based on the combination of a genetic algorithm and tabu search', *Comput. Intell. Neurosci.*

12 Tahir, M.A., Bouridane, A. and Kurugollu, F. (2007) 'Simultaneous feature selection and feature weighting using Hybrid Tabu Search/K-nearest neighbor classifier', *Pattern Recognit. Lett.*, Vol. 28, No. 4, pp.438–446.

13 Aghdam, M.H., Ghasem-Aghaee, N. and Basiri, M.E. (2009) 'Text feature selection using ant colony optimization', *Expert Syst. Appl.*, Vol. 36, No. 3, pp.6843–6853.

14 Kanan, H.R. and Faez, K. (2008) 'An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system', *Appl. Math. Comput.*, Vol. 205, No. 2, pp.716–725.

15 Sivagaminathan, R.K. and Ramakrishnan, S. (2007) 'A hybrid approach for feature subset selection using neural networks and ant colony optimization', *Expert Syst. Appl.*, Vol. 33, No. 1, pp.49–60.

16 Sreeja, N.K. and Sankar, A. (2015) 'Pattern matching based classification using ant colony optimization based feature selection', *Appl. Soft Comput.*, Vol. 31, pp.91–102.

17 Welikala, R.A., Fraz, M.M., Dehmeshki, J., Hoppe, A., Tah, V., Mann, S., Williamson, T.H. and Barman, S.A. (2015) 'Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy', *Comput. Med. Imaging Graph.*, Vol. 43, pp.64–77.

18 Erguzel, T.T., Ozekes, S., Tan, O. and Gultekin, S. (2015) 'Feature selection and classification of electroencephalographic signals: an artificial neural network and genetic algorithm based approach', *Clin. EEG Neurosci.*, Vol. 46, No. 4, pp.321–326.

19 Oreski, S. and Oreski, G. (2014) 'Genetic algorithm-based heuristic for feature selection in credit risk assessment', *Expert Syst. Appl.*, Vol. 41, No. 4, pp.2052–2064.

20 Li, S., Wu, H., Wan, D. and Zhu, J. (2011) 'An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine', *Knowl. Based Syst.*, Vol. 24, No. 1, pp.40–48.

21 Das, N., Sarkar, R., Basu, S., Kundu, M., Nasipuri, M. and Basu, D.K. (2012) 'A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application', *Appl. Soft Comput.*, Vol. 12, No. 5, pp.1592–1606.

22 Wang, Y., Chen, X., Jiang, W., Li, L., Li, W., Yang, L., Liao, M. *et al.* (2011) 'Predicting human microRNA precursors based on an optimized feature subset generated by GA–SVM', *Genomics*, Vol. 98, No. 2, pp.73–78.

23 Xue, B., Zhang, M. and Browne, W.N. (2012) 'Particle swarm optimization for feature selection in classification: a multi-objective approach', *IEEE Trans. Cybern.*, Vol. 43, No. 6, pp.1656–1671.

24 Chen, Li-Fei, Chao-Ton Su, Kun-Huang Chen, and Pa-Chun Wang (2012) 'Particle swarm optimization for feature selection with application in obstructive sleep apnea diagnosis', *Neural Comput. Appl.*, Vol. 21, No. 8, pp.2087–2096.

25 Yang, H., Du, Q. and Chen, G. (2012) 'Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification', *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, Vol.5, No. 2, pp.544–554.

26 Liu, H. and Setiono, R. (1995) 'Chi2: Feature selection and discretization of numeric attributes', *Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence*, IEEE, pp.388–391.

27 Kohavi, R. and John, G.H. (1997) 'Wrappers for feature subset selection', *Artif. Intell.*, Vol. 97, Nos. 1–2, pp.273–324.

28 Inza, I., Larrañaga, P., Etxeberria, R. and Sierra, B. (2000) 'Feature subset selection by Bayesian network-based optimization', *Artif. Intell.*, Vol. 123, Nos. 1–2, pp.157–184.

29 Grimaldi, M., Cunningham, P. and Kokaram, A. (2003) 'An evaluation of alternative feature selection strategies and ensemble techniques for classifying music', *Workshop on Multimedia Discovery and Mining*, Vol. 2.

30 Dy, J.G. and Brodley, C.E. (2000) 'Feature subset selection and order identification for unsupervised learning', *ICML*, pp.247–254.

31 Aha, D.W. and Bankert, R.L. (1996) 'A comparative evaluation of sequential feature selection algorithms', *Learning from data*, Springer, New York, NY, pp.199–206.

32 Maldonado, S. and Weber, R. (2009) 'A wrapper method for feature selection using support vector machines', *Inf. Sci.*, Vol. 179, No. 13, pp.2208–2217.

33 Gutlein, M., Frank, E., Hall, M. and Karwath, A. (2009) 'Large-scale attribute selection using wrappers', *2009 IEEE Symposium on Computational Intelligence and Data Mining*, IEEE, pp.332–339.

34 Kabir, M.M., Islam, M.M. and Murase, K. (2010) 'A new wrapper feature selection approach using neural network', *Neurocomputing*, Vol. 73, Nos. 16–18, pp.3273–3283.

35 Stein, G., Chen, B., Wu, A.S. and Hua, K.A. (2005) 'Decision tree classifier for network intrusion detection with GA-based feature selection', *Proceedings of the 43rd Annual Southeast Regional Conference-Volume 2*, pp.136–141.

36 Zhuo, L., Zheng, J., Li, X., Wang, F., Ai, B. and Qian, J. (2008) 'A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine', *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images*, SPIE, Vol. 7147, pp.503–511.

37 金田重郎(1995) 'Quinlan, JR: C4.5 Programs for machine learning, Morgan Kaufmann, San Mateo, California (1992)', 人工知能, Vol. 10, No. 3, pp.475–476.

38 Baluja, S. (1994) *Population-based Incremental Learning. A Method for Integrating Genetic Search based Function Optimization and Competitive Learning*, Carnegie-Mellon Univ Pittsburgh Pa Dept. of Computer Science.

39 Buntine, W. (1991) 'Theory refinement on Bayesian networks', *Uncertainty Proceedings 1991*, Morgan Kaufmann, pp.52–60.

40 Gadekallu, T.R., Rajput, D.S., Reddy, M.P.K., Lakshmanna, K., Bhattacharya, S., Singh, S., Jolfaei, A. and Alazab, M. (2020) 'A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU', *J. Real-Time Image Process*, Vol. 18, pp.1383–1396.

41 Mahajan, S., Raina, A., Gao, X-Z. and Kant Pandit, A. (2021) 'Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost', *Symmetry*, Vol. 13, No. 2, p.356.

42 Loughrey, J. and Cunningham, P. (2004) 'Overfitting in wrapper-based feature subset selection: The harder you try the worse it gets', *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, Springer, London, pp.33–43.

43 Chandrashekar, G. and Sahin, F. (2014) 'A survey on feature selection methods', *Comput. Electr. Eng.*, Vol. 40, No. 1, pp.16–28.

44 Guyon, I., Weston, J., Barnhill, S. and Vapnik, V. (2002) 'Gene selection for cancer classification using support vector machines', *Mach. Learn.*, Vol. 46, No. 1, pp.389–422; Quinlan, J.R. (1986) 'Induction of decision trees', *Mach. Learn.*, Vol. 1, No. 1, pp.81–106.

45 Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005) 'Sparsity and smoothness via the fused lasso', *J. R. Stat. Soc., B*, Vol. 67, No. 1, pp.91–108.

46 Fu, H., Xiao, Z., Dellandréa, E., Dou, W. and Chen, L. (2009) 'Image categorization using ESFS: a new embedded feature selection method based on SFS', *ACIVS*, Vol. 5807, pp.288–299.

47 Xiao, Z., Dellandrea, E., Dou, W. and Chen, L. (2008) *ESFS: A New Embedded Feature Selection Method based on SFS*, PhD diss., Ecole Centrale Lyon; Université de Lyon; LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/École Centrale de Lyon; Tsinghua University, Bejing, China.

48 Maldonado, S., Weber, R. and Famili, F. (2014) 'Feature selection for high-dimensional class-imbalanced data sets using support vector machines', *Inf. Sci.*, Vol. 286, pp.228–246.

49 Kira, K. and Rendell, LA. (1992) 'A practical approach to feature selection', *Machine Learning Proceedings 1992*, Morgan Kaufmann, pp.249–256.

50 Parsa, M.G., Zare, H. and Ghatee, M. (2020) 'Unsupervised feature selection based on adaptive similarity learning and subspace clustering', *Eng. Appl. Artif. Intell.*, Vol. 95, p.103855.

51 Bermejo, P., Gámez, J. and Puerta, J. (2008) 'On incremental wrapper-based attribute selection: experimental analysis of the relevance criteria', *Proceedings of IPMU'08*, pp.638–645.

52 Ruiz, R., Riquelme, J.C. and Aguilar-Ruiz, J.S. (2006) 'Incremental wrapper-based gene selection from microarray data for cancer classification', *Pattern Recognit.*, Vol. 39, No. 12, pp.2383–2392.

53 Jawhar, Q., Thakur, K. and Singh, K.J. (2020) 'Recent advances in handling big data for wireless sensor networks', *IEEE Potentials*, Vol. 39, No. 6, pp.22–27, November–December, doi: 10.1109/MPOT.2019.2959086.