© Universiti Tun Hussein Onn Malaysia Publisher's Office

JSCDM

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/jscdm

Journal of Soft Computing and Data Mining

A Review of Principal Component Analysis Algorithm for Dimensionality Reduction

Basna Mohammed Salih Hasan[1*], Adnan Mohsin Abdulazeez[2]

[1]IT Department. Technical College of Informatics Akre,
Duhok Polytechnic University, Duhok, Kurdistan Region, IRAQ

[2]Computer Engineering, Duhok Polytechnic University,
Duhok, Kurdistan Region, IRAQ

*Corresponding Author

DOI: https://doi.org/10.30880/jscdm.2021.02.01.003
Received 26 January 2021; Accepted 03 March 2021; Available online 15 April 2021

Abstract: Big databases are increasingly widespread and are therefore hard to understand, in exploratory biomedicine science, big data in health research is highly exciting because data-based analyses can travel quicker than hypothesis-based research. Principal Component Analysis (PCA) is a method to reduce the dimensionality of certain datasets. Improves interpretability but without losing much information. It achieves this by creating new covariates that are not related to each other. Finding those new variables, or what we call the main components, will reduce the eigenvalue /eigenvectors solution problem. (PCA) can be said to be an adaptive data analysis technology because technology variables are developed to adapt to different data types and structures. This review will start by introducing the basic ideas of (PCA), describe some concepts related to (PCA), and discussing. What it can do, and reviewed fifteen articles of (PCA) that have been introduced and published in the last three years.

Keywords: Machine learning, principal component analysis, dimensionality reduction

1. Introduction

Machine Learning (ML) is the automatic training of a computer for specific tasks through algorithms. Automatically learning to know user preferences, application sets of algorithms are used to mine data that discover and filter general rules in large data sets [1]. Big data is an environment in which approaches are extract, information is regularly collected or otherwise stored in data sets that are too much or complicated to be managed by standard data processing application [2]. Current use of the term big data uses predictive analytical systems, user behavioral analytics or some other advanced methods of data analysis that extract value from big data and rarely a particular data set size. In many ways, big data deposits have been constructed by corporations with special needs data processing like operating systems has gradually expanded in sizes and number of available data sets, will need massively parallel software running on tens, hundreds, or even thousands of servers to run parallel software [3].

Big data technology has continued to enhance medical treatment through the provision of customized healthcare and prescriptive analysis, clinical risk response as well as predictive analysis, decrease of duplication and care variabilities, automatic external and internal patient reporting, structured medical terminology and patient registers [4][5]. Dimensionality reduction methods reduce data from higher dimensions to lower dimensions according to some specific criteria [6]. Principal Component Analysis (PCA) is a technique of dimensionality-reduction (DR) that is mostly used to reduce a large set of variables into a smaller one that still contains much of the details in the large set [7]. This paper reviewed will start by introducing the basic ideas of PCA, describe some concepts related to PCA, and discussing articles of PCA that have been introduced and published in the last three years.

*Corresponding author: basna.mhmed@dpu.edu.krd

The primary drive behind PCA is to use the fewest components possible to reduce the dimensions of the data we work with. Hence, we often care about PCA trimming - keep only the top ingredients and discard the rest. There are at least two reasons for this: First, trimming gives us a sense of the complexity of the data set. If the two major components capture a large majority of the contrast, then the dataset is more or less two-dimensional. Second, truncation weakens the data. The conceptual correlation of PCA by regression is useful again here - PCA resembles a smooth curve fit through noisy data as the principal component approximation gives a smooth, less noisy representation of the data. The goals of PCA are to reduce the dimensions of the dataset first, and then to identify meaningful new variables secondly.

In this paper, the review is structured as follows. Part 1 is an introduction to general concepts related to PCA. Part 2 provides the necessary background for Dimensionality Deduction (DR), Machine Learning (ML), and the PCA method. Part 3 related work to the PCA technique the past few years. Part 4 comparison and discussion. Lastly, Part 5 concludes the paper review.

2. Machine Learning

Together with big data technology and high-performance computing, Machine learning (ML) has evolved to build new opportunities in various operating settings to explain, measure, and get data-intensive processes. Machine learning (ML) is regarded as a science field that helps computers to understand, among other things, without being completely programmed. [8]. Machine learning (ML) encompasses more and more research areas per year, including, for example, bioinformatics [9], [10], biochemistry [11], [12], meteorology of medicine [13]–[16], economics [17]–[19], aquaculture [20], chemo-ecology[21], robotics [22]–[25], and climatology [26], [27].

There are tens of thousands of algorithms for machine learning, and every year, hundreds of new algorithms are developed. The three components of each machine learning algorithm are representation, evaluation, and optimization. Machine learning (ML) has 4 types.

- Supervised learning is training data that includes wished-for outputs
- Unsupervised learning is training data that has no real outputs.
- Semi-supervised learning is training data that has some desirable output.
- Reinforcement learning, the most ambitious method of learning [28] in (see Fig.1).

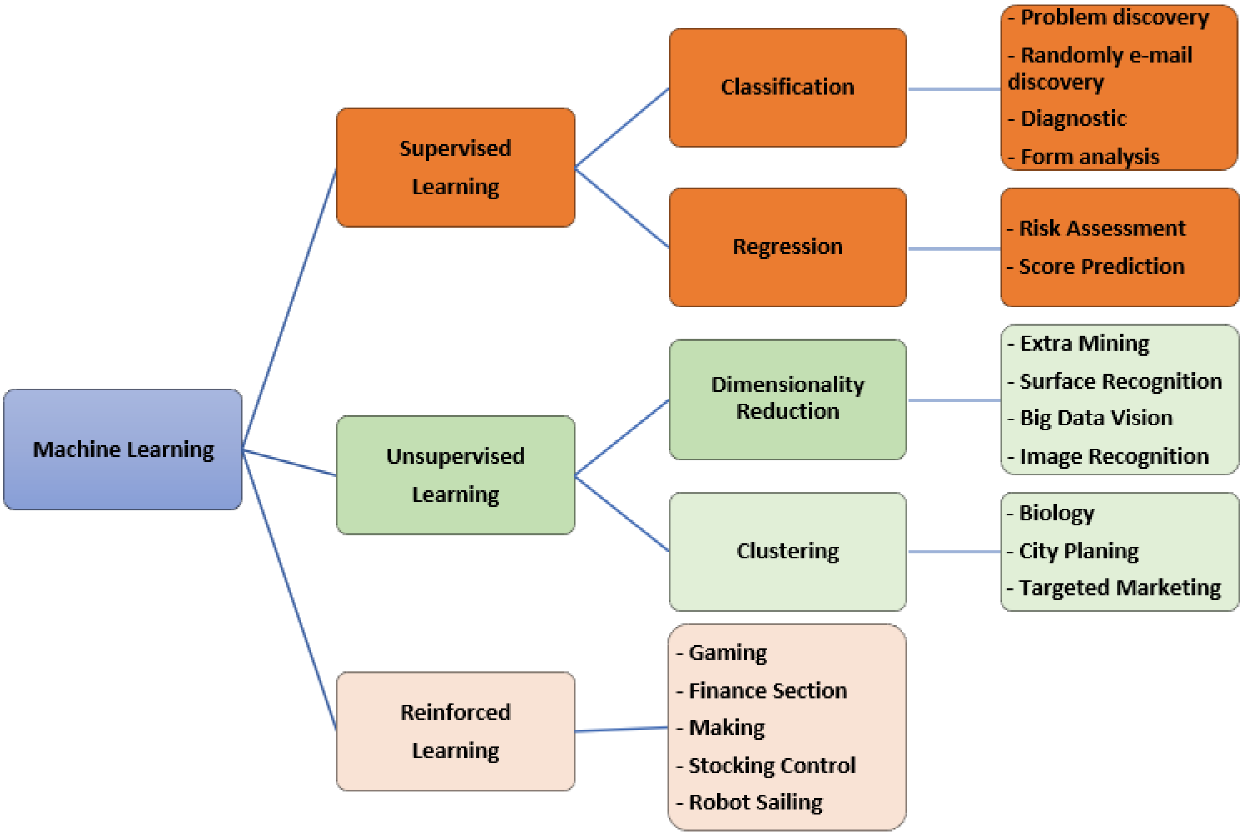

Fig. 1 - Type of machine learning [29].

2.1 Dimensionality Reduction Techniques (DR)

Dimensionality Reduction (DR) algorithm, which aims to reduce the distance in a latent space between distributions of different data sets to allow efficient transfer learning [30]. The results point that for each device individually, the findings with Dimensionality Reduction (DR) are much preferable to those without decreased dimensionality [31].

The low dimensional data representation of the initial data tends to overcome the issue of the dimensionality curse, and can be easily analyzed, processed, and visualized [32][33]. Advantage of dimensionality reduction techniques applied to a dataset. (i) decrease the number of dimensions, and data storage space. (ii) It requires less time to compute. (iii) Irrelevant, noisy, and redundant data can be deleted. (iv) Data quality may well be optimized. (v) helps an algorithm to work efficiently and improves accuracy. (vi) allow to visualize data (vii) It simplifies the classification and increases performance as well [34], [35].

Generally, the DR techniques are classified into two main different techniques: Feature Selection (FS) and Feature Extraction (FE). Feature Selection (FS) is considered an important method since data is constantly produced at an ever-increasing rate; with this method, some significant dimensionality concerns can be minimized, such as effectively decreasing redundancy, eliminating unnecessary data, and enhancing comprehensibility of results. Moreover, Feature Extraction (FE) addresses the issue of finding the most distinctive, informative, and reduced set of features to improve the efficiency of both data processing and storage [32], [36]–[40] (see Fig.2).

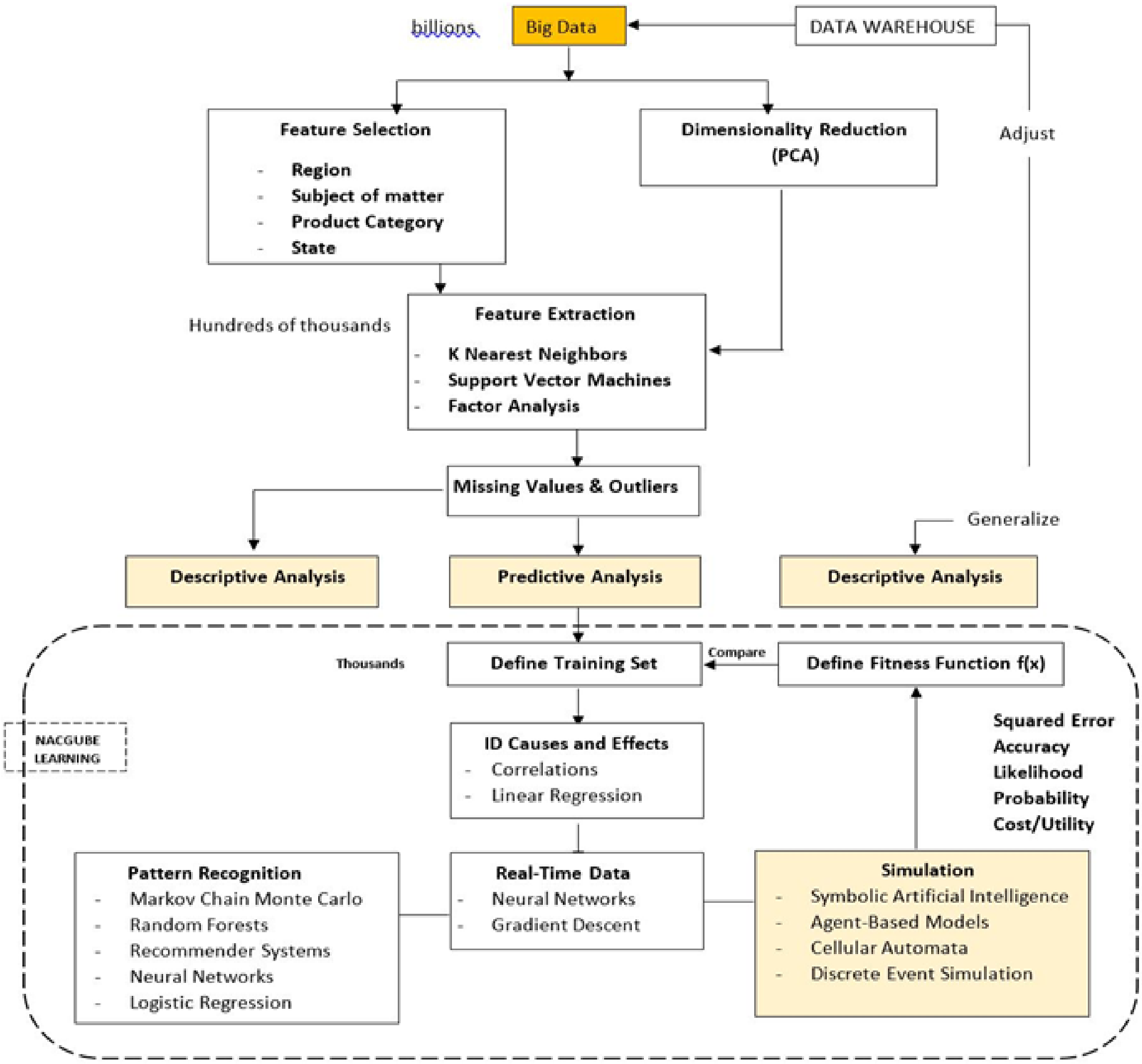

Fig. 2 - Type of Dimensionality Reduction (DR) [41].

2.2 Principal Component Analysis

There are many Dimensionality Reduction (DR) technologies, but the most used one is Principal Component Analysis (PCA), like Fourier analysis, or wavelet decomposition this has the benefit of comparing findings across data sets and resolving the importance of the components, then why do PCA transform instead of factor analysis? First, in some states the purpose of doing a factor analysis is different technique. We could number the features and get cosines of the characteristic numbers, etc., but it just seems crazy no such difficulties attend PCA.

Second, when applying a fixed attitude of segments, there is no warranty that a tiny number of features will give a good regeneration of the main data, PCA confirms that the first q features will make a better (mean-square) function of reconstructing the main data than any other linear method using alone q features.

Third, it is great at maintaining measures between the points the features scores give the optimal linear multidimensional calibration, the PCA gives us uncorrelated features, which are popularly not confident features; for that you need independent features analysis, the PCA views for linear features of the main features; one could completely make more agreeable by finding nonlinear features, the PCA is basically a descriptive method it does not what forecast the future data .that'll be curves or surfaces rather than directions in feature space. The concept is to reduce the dimensionality of data while retaining the maximum 'variance'. While it is used and has often been reinvented, is at its core, a statistical technique in several various disciplines and therefore most of its development has been by statisticians [42][43].

The aim of PCA is to find an optimal position for the best Information variance and vector dimensional features reduction. The PCA is an unsupervised learning technique that reduces the dimensionality of data [44]. PCA is a dimensionality reduction technique invented by Karl Pearson in 1901[45]. It is a dimensional reduction technique that is often used to reduce the features of large data sets into smaller features that contain the most information in a big data set. PCA is an orthogonal statistical technique that converts a set of observations of related variables into a set of non-linearly related values [38], [46]–[48].

PCA has been found used in a wide range of fields ranging from spike-triggered covariance analysis in neuroscience [49], [50], to quantitative finance [51]–[54] with the most common application being facial recognition [54]–[56] and other applications like medical data correlation [57]–[60].

This indicates that "to preserve as far deviation as possible" involves the discovery of new variables which are linear features of those in the original data set, and the successive optimization of variance which is irrelevant. The key components (PCs) discovering those new variables reduces the issue of own eigenvalue/eigenvector solving.

Pearson [61] and Hotelling's [62] of early PCA literature, however it was only decades after that modern computers became widely available that it was used on small databases computationally. Its use has since expanded and a range of variants have been provided in various disciplines. Many books have been written on the subject and whole books on differences in PCA of unique data types have been released [63]. PCA is often used for analyzing data in the most various filed. This paper review reviewed the PCA approach to several theoretical and practical aspects of PCA. it starts by providing in basic principles and accessible manner, the basic principles underlying PCA and its applications.

Advantages of PCA include: (i) is used to overcome duplication in features in the data set. (ii) Valuable information is obtained explaining the high contrast providing the best resolution. (iii)It makes a better visualization of the data. (iv) It reduces complexity and increases computational efficiency. (v)It is a very flexible tool and allows analysis of datasets that may contain, (vi) multi-collinearity, (vii) missing values, (viii) categorical data, (ix) and imprecise measurements. The goal is to extract the important information from the data and to express this information as a set of summary indices called principal components. (x) One of the most important applications of PCA is for speeding up machine learning algorithms. (xi) PCA as Dimensionality Reduction Techniques (DR) [64]–[68].

2.3 Principal Component Analysis Techniques

PCA is an unsupervised, non-parametric statistical technique mainly used for dimensionality reduction in machine learning [69]. PCAs may have different units of measure so there will be some unwanted features, and to solve these unwanted features it is usual to start calibrate the variants. The standard deviation from the vector measurements centers and segments each data value. The PCA (variance) standard that is based on the units of measure indicates computers based on a matrix covariance can change as units change in one or more variables [70].

Since the standard dataset covariance matrix is essentially the correlation matrix of the original data set, a PCA is often referred to as the PCA correlation matrix for standard data. The eigenvectors of this matrix describe linear combinations of the standard variables of the uncorrelated maximum-variance [42]. These PCAs are nor are they mainly correlated to the previously defined PCAs of the covariance matrix. Each PCA will also differ by percentage variance, and more PCAs are expected to cover the same percentage as the substantial percentage than the covariance matrix PCAs [71]. The matrix of correlation is directly proportional to the number of variables used in the analysis, so that any correlation matrix PCA is the proportion of total variance which is divided into the number of variables of the PCA. PCAs are the appropriate choice for datasets where various scale changes are conceivable for every variable and are invariant for linear changes in measurement units. [72].

Some mathematical program suggests that a PCA is a matrix correlation PCA by nature and, often, the standardization for loading vectors is not the normal In the PCA matrix the correspondence coefficient between the column is the vector variable. The PCA matrix covariance is meaningful. A PCA correlation matrix results in the same way because variances in the original component are not so different. The first two matrix PCs reflect a percentage of the overall variance. Differences can be more important for other datasets [73].

3. Review for PCA Algorithm

This section will review the latest literature regarding of the Principal Component Analysis (PCA) and dimensional reduction techniques (DR). Table 1 shows a summary of recent literature. Reddy et al in [74] investigated the two pioneer DR techniques in machine learning (ML), PCA and Linear Discriminant Analysis (LDA). They applied on a dataset in the UCI machine learning repository. Features have been reduced to 26 from 36 dependent attributes by retaining 95% of the dataset by PCA, while Linear Discriminant Analysis (LDA) reduced the features to 1, using Decision Tree (DT) classifier, Naive Bayes classifier, Random Forest classifier, and Support Vector Machine (SVM). It is observed that the performance of classifiers with PCA is better than that of Linear Discriminant Analysis (LDA).

Saraswathi and Gupta [75] proposed a multi-class tumor classification approach using a Random Forest (RF) classifier and also presented a comparative analysis of Random Forest (RF), RF-PCA, and RF-PCA approaches with random selection techniques. Through the experimental results, it has been shown that the random selection of RF-PCA provides better accuracy than other methods. Additionally, Jamal et al in [76] focused on the number of features for the classification of breast cancer from the original White Blood Cell (WBC) data set can be reduced by the feature extracting. The metric measurement results that the dimensionality reduction using the K-means cluster is almost as good as PCA.

On the other hand, Salo et al [77], proposed A novel hybrid technique combining of Information Gain (IG) and PCA to discard irrelevant features and retain the optimum attribute subset. The robustness of the proposed approach yielded promising results in both NSL-KDD and Kyoto 2006+ datasets.

evaluated PCA performance against Isomap a deep autoencoder, and a variational autoencoder. Three widely utilized picture datasets were experimented with: MNIST, Fashion-MNIST, and CIFAR-10. The processing period for PCA was two orders of magnitude faster than its equivalents in the neural network. The two auto-encoders had a sufficiently broad dimension [78].

Amir et al in [79] used a fine-tuned deep Neural network for the Hoda dataset, which is the largest dataset for Persian handwritten digit classification to extract useful discriminative characteristics. Then, for the classification portion, these features are fed to a linear Support Vector Machine (SVM). In the next experiment, he used the PCA to decrease the measurements of the extracted characteristics to increase accuracy and computational load, then we fed it to Support Vector Machine (SVM). To the best of our experience, in terms of precision calculation, the suggested approach was stronger than other approaches.

An innovative technique focused on the acquisition of X-ray fluorescence spectroscopy XRF spectra using a PCA-related portable energy dispersive XRF instrument (pED-XRF) instrument and machine learning algorithms has been successfully applied and tested on a set of 18 meteorite and meteor-wrong samples of various kinds and origins. This technique allowed meteorite classification and discrimination of macro-groups, e.g., iron, stony-iron, and meteor-wrong, to be rapid and trustworthy. The findings revealed that the Cubic Support Vector Machine (CSVM), Fine Kernel Nearest Neighbor (FKNN), Subspace Discriminant-Ensemble Classifiers SD-EC, and SKNN algorithms on uniform spectra obtained in the key instrumental energy spectrum achieved 100 percent precision in classifying meteorites. These first findings confirm that a mixture of X-ray fluorescence spectroscopy XRF spectra and machine learning algorithms is a very powerful and promising solution to identifying and classifying every actual or supposed meteorite [80].

Work in [81] focused on dynamic Cone Beam X-ray luminescence computed tomography (CB-XLCT), a 4D temporal-locative remodeling procedure is proposed based on main PCA in the projection space. Firstly, any 3D image compresses angle measurements to decrease the noise initially. In order to decorate the 4D issue with various 3D problems in the PCA domain, the projection data are then used to do a temporal PCA. Qualitative and quantitative review of the experiment results shows the preferred solution to overcome numerous targets and recover complex concentrates with great success, strong computational efficiency.

The suggested use as a tool for dimension reduction PCA and the support vector machine SVM in order to improve the accuracy of cancer detection as a classification for microarray data classification optimized by kernel functions. The recommended schema was extended to 7 data sets with 5-fold cross-validation and then both in terms of precision and time was measured and analyzed. The outcome showed that when Linear and Cubic kernel functions are used, the scheme can achieve 100 percent accuracy for Ovarian and Lung Cancer data. The PCA substantially decreased the running period for all data sets in terms of running time [82].

Table 1 - Review for PCA algorithm.

Ref.	Year	Dataset	Techni	Contribution	Accuracy
[74]	2020	DR & IDS	LDA - PCA With some classification algorithms	Ignoring redundant features reduces the burden on machine learning algorithms	95%
[75]	2019	brain tumor dataset having 3064	GLCM – LBP - PCA	Identifying and classifying tumors helps the radiologist make an accurate diagnosis	88.72% 85.56%.
[76]	2018	breast cancer -WBC	PCA and SVM PCA and XGBoost	measurement for the sake of early detection of breast cancer.	97.8%
[79]	2020	Hoda dataset	optical character recognition (OCR) SVM PCA	Persian handwritten digits recognition	SNN=95% PCA,kNN=97.11 % ANFIS.=97.3%
[80]	2020	meteorite samples of different nature and origin	XRF and PCA CSVM and SD-EC KNN	To identifying and classifying every actual or supposed meteorite	100%
[85]	2020	the Nansi Lake Basin, China	PCA and FCE	discovery of the most important pollutions and heavy metals into water	60%
[82]	2018	morphological and clinical	PCA and SVM	to improve the accuracy of cancer detection	100%
[83]	2017	Numerous high-hroughput omics	meta-analytic framework of PCA	effective dimensional reduction for exploratory visualization.	>80%
[84]	2018	MIT-BIH arrhythmia	PCA-Net and SVM	Echocardiographic noise influences ECG classification models and negatively affects skewed data	97.77%
[87]	2017	12 standard word similarity benchmarks	PCA and post-processing	construct word embeddings of higher dimensions	50%

Also, a research by [83], combined multiple omics datasets from similar biological hypotheses and presented two different types of PCA meta-analysis framework, namely, Meta PCA. To facilitate selection of scattered features in MetaPCA. combined multiple omics datasets from similar biological hypotheses and presented two different types of PCA meta-analysis framework, namely, MetaPCA. To facilitate selection of scattered features in Meta PCA. They used simulations, three meta-analysis transcriptional studies in the yeast cycle, prostate cancer, rat metabolism, and a The Cancer Genome Atlas (TCGA) study for pan-cancer methylation. The result enhanced detailed visualisation, robustness and discovery of the proposed structure.

The work of [84] presented a novel heartbeat recognition method. The goal of this study is to apply principal component analysis network (PCANet) focused on a noisy EigenECG Network (ECG) signal for feature extraction. A linear support vector machine (SVM) was applied to increase the classification speed. To check the effectiveness of their algorithm, they defined five types of imbalanced initial and noise-free EigenECG Networks (ECGs) in the Massachusetts Institute of Technology (MIT)-Boston's Beth Israel Hospital (BIH) Arrhythmia Database (MIT-BIH) and achieved 97.77 % and 97.08 % accuracy, respectively [84].

The work of [85] provided the built-in the Fuzzy Comprehensive Evaluation-Principal Component Analysis (FCE-PCA) Model useful water quality evaluation details. The model was applied to the waters of the Lake Nancy Basin in China. The efficiency of extraction of main contaminants was improved by generating organic functions through the semisinusoidal distribution system, measuring weight via several additional standard methods, then solving self-equation using Jacobi and removing the main components based on inherent values, the percentage of the contribution accumulated as well as packing. The result has been four components removed which are the 14 pollutant indices for water quality evaluation. And then, Lazcano in [86] has presented a review of the algorithm parallelism PCA and its adaptation to (MPPA) a Massively Parallel Processor Array architecture of 256 centers spanning 16 clusters. The emphasis was on bringing hyper-spectral image processing on multiple platforms by changes to real-time limitations fixed by the image capture rate of the hyperspectral sensor. Real-time is an awkward objective for hyperspectral image processing because hyperspectral images are made of exceedingly large volumes of data and this problem is effectively solved by increasing the size of the image before the processing is begun [86].

Another work presented a new algorithm that effectively combines the reduction of PCA -based dimensionality with a recently suggested post-processing algorithm to create lower-dimensional word embeddings [87]. Empirical assessments on 12 standard word similarity benchmarks showed that our algorithm decreases the dimensionality of the embedding by 50 percent, thus obtaining equivalent or (more often) greater efficiency than the embedding of the higher dimension. And lastly, [88] proposed a simple and effective super pixels (SuperPCA) approach to learning low-dimensional essential features of hyperspectral images (HSIs). the three sets of data proved that the (SuperPCA) approach significantly superior to traditional baselines to reduce the dimensions based on PCA for classification HIS.

From the literature that has been described above, it is clear that the PCA algorithm has proven effective in reducing the dimensions of big data with high accuracy and improving the work of classification. Although the PCA algorithm is commonly used in many areas, it has many of the modifications that make it usable for a range of scenarios and various types of data [89]. The literature about PCA is extensive and includes many disciplines, and research is still ongoing to suggest the algorithm with other techniques to reduce the dimensions of big data and save space and time for processing. In addition to its higher efficiency compared to its counterparts in work, for example, we noted in research [74], it is observed that the performance of classifiers with PCA is better than that of Linear Discriminant Analysis (LDA). Although it is in low-dimensional data it is less efficient.

4. Conclusion

The most frequently used approach for reduction of dimensionality is PCA. In essence, it reduces the high dimensions in a large data set to fewer dimensions to speed up the storage and processing process, and the data is more interpretable and faster when processing. It is a statistical technique that preserves the largest amount of information and eliminates redundant noise and data. There is a lot of literature on PCA. In this work, previous papers related to PCA technology were reviewed by presenting their basic ideas, describing some related concepts, and discussing what they could do.

Acknowledgement

The authors would like to thank the IT Department. Technical College of Informatics Akre, Duhok Polytechnic University and the Computer Engineering, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq.

References

[1] Peng, D., Jin, L., Wu, Y., Wang, Z., and Cai, M. (2019) A fast and accurate fully convolutional network for end-to-end handwritten chinese text segmentation and recognition, International Conference on Document Analysis and Recognition

[2] Anam, M., Ponnusamy, V., Hussain, M., Nadeem, M. W., Javed, M., Goh, H.G., Qadeer, S. (2021) Osteoporosis prediction for trabecular bone using machine learning: a review, Computers, Material Continua, vol. 67, no. 1, pp. 89–105, doi: 10.32604/cmc.2021.013159

[3] Awan, M. J., Mohd Rahim, M. S., Salim, N., Mohammed, M. A., Garcia-Zapirain, B., and Abdulkareem, K. H. (2021). Efficient Detection of Knee Anterior Cruciate Ligament from Magnetic Resonance Imaging Using Deep Learning Approach, Diagnostics, vol. 11, no. 1, p. 105, doi: 10.3390/diagnostics11010105

[4] Gupta, M., Jain, R., Arora, S., Gupta, A., Javed Awan, M., Chaudhary, G., Nobanee., H. (2021). AI-enabled COVID-19 outbreak analysis and prediction: Indian states vs. union territories, Computers, Material Continua., vol. 67, no. 1, pp. 933–950, doi: 10.32604/cmc.2021.014221

[5] Awan, M. J., Mohd Rahim, M. S., Nobanee, H., Munawar, A., Yasin, A., Mohd Zain, A. (2021). Social Media and Stock Market Prediction: A Big Data Approach, doi: 10.32604/cmc.2021.014253

[6] Van Der Maaten, L. J. P., Postma, E. O., and Van Den Herik, H. J. (2009). Dimensionality Reduction: A Comparative Review, J. Mach. Learn. Res., vol. 10, pp. 1–41, doi: 10.1080/13506280444000102

[7] Wold, S., Esbensen, K. and Geladi, P. Principal Component Analysis. (1987). Chemom. Intell. Lab. Syst., vol. 2, no. 1–3, pp. 37–52, [Online]. Available: http://files.isec.pt/documentos/servicos/biblio/documentos de acesso remoto/principal components analysis.pdf.

[8] Samuel, L. (1959). Eight-move opening utilizing generalization learning. (See Appendix B, Game G-43.1 Some Studies in Machine Learning Using the Game of Checkers, IBM J., pp. 210–229

[9] Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 2007 Jul;35(Web Server issue): W345-9. doi: 10.1093/nar/gkm391. PMID: 17631615; PMCID: PMC1933232

[10] Mackowiak, S.D., Zauber, H., Bielow, C. et al. (2015). Extensive identification and analysis of conserved small ORFs in animals. Genome Biol 16, 179
[11] Richardson, A. Signor, B.M., Brett A. Lidbury, Badrick, T. (2016). Clinical chemistry in higher dimensions: Machine-learning and enhanced prediction from routine clinical chemistry data, Clin. Biochem., vol. 49, no. 16–17, pp. 1213–1220, doi: 10.1016/j.clinbiochem.2016.07.013
[12] Wildenhain, J., Spitzer, M., Dolma, S., Jarvik, N., White, R., Roy, M., Griffiths, E., Bellows, D.S., Wright, G.D., Tyers, M. (2015). Prediction of Synergism from Chemical-Genetic Interactions by Machine Learning, Cell Syst., vol. 1, no. 6, pp. 383–395, doi: 10.1016/j.cels.2015.12.003
[13] Cramer,S., Kampouridis, M., Freitas, A. A., Alexandridis, A. K. (2017). An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives, Expert Syst. Appl., vol. 85, pp. 169–181, doi: 10.1016/j.eswa.2017.05.029
[14] Zhang B, He X, Ouyang F, Gu D, Dong Y, Zhang L, Mo X, Huang W, Tian J, Zhang S. (2017). Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. Cancer Lett. 2017 Sep 10;403:21-27. doi: 10.1016/j.canlet.2017.06.004. Epub 2017 Jun 10. PMID: 28610955
[15] Asadi H, Dowling R, Yan B, Mitchell P. (2014). Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. PLoS One. 2014 Feb 10;9(2):e88225. doi: 10.1371/journal.pone.0088225. PMID: 24520356; PMCID: PMC3919736
[16] Kang J, Schwartz R, Flickinger J, Beriwal S. (2015) Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. Int J Radiat Oncol Biol Phys. 2015 Dec 1;93(5):1127-35. doi: 10.1016/j.ijrobp.2015.07.2286. Epub 2015 Nov 11. PMID: 26581149
[17] Zounemat-Kermani, M., Ramezani-Charmahineh, A., Razavi, R. et al. (2020). Machine Learning and Water Economy: a New Approach to Predicting Dams Water Sales Revenue. Water Resour Manage 34, 1893–1911. https://doi.org/10.1007/s11269-020-02529-0
[18] Paz-Marín, M., Gutiérrez, P.A., Hervás-Martínez,C. (2015). Classification of countries' progress toward a knowledge economy based on machine learning classification techniques, Expert Systems with Applications, Volume 42, Issue 1, pp. 562-572
[19] Barboza, F. Kimura, H., Altman, E. (2017). Machine learning models and bankruptcy prediction, Expert Syst. Appl., vol. 83, pp. 405–417, doi: 10.1016/j.eswa.2017.04.006
[20] López-Cortés, X. A., Nachtigall, F. M., Olate, V. R., Araya, M., Oyanedel, S., Diaz, V., Jakob, E., Ríos-Momberg, M., Santos, L. S. (2017). Fast detection of pathogens in salmon farming industry, Aquaculture, vol. 470, pp. 17–24, 2017, doi: 10.1016/j.aquaculture.2016.12.008
[21] Mitchell JB. Machine learning methods in chemoinformatics. (2014). Wiley Interdiscip Rev Comput Mol Sci. 2014 Sep 1;4(5):468-481. doi: 10.1002/wcms.1183. PMID: 25285160; PMCID: PMC4180928
[22] Takahashi, K., Ogata, T., Nakanishi, J., Cheng, G. and Sugano, S. (2017). Dynamic motion learning for multi-DOF flexible-joint robots using active–passive motor babbling through deep learning, Adv. Robot., vol. 31, no. 18, pp. 1002–1015, doi: 10.1080/01691864.2017.1383939
[23] Thomaz, A., Hoffman, G., & Breazeal, C. (2005). Real-Time Interactive Reinforcement Learning for Robots. AAAI Work. - Tech. Rep., vol. WS-05-04, pp. 9–13
[24] Saito, N., Kim, K., Murata, S., Ogata, T. and Sugano, S. (2019). Tool-Use Model Considering Tool Selection by a Robot Using Deep Learning, IEEE-RAS Int. Conf. Humanoid Robot., vol. 2018-Novem, pp. 814–819, doi: 10.1109/HUMANOIDS.2018.8625048
[25] Mahadevan, S. (1996). Machine Learning for Robots: A Comparison of Different Paradigms, Workshop on Towards Real Autonomy, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-96)
[26] Foresti, L., Sideris, I. V., Nerini, D., Beusch, L., and Germann, U. (2019). Using a 10-year radar archive for nowcasting precipitation growth and decay: A probabilistic machine learning approach, Weather Forecast., vol. 34, no. 5, pp. 1547–1569, doi: 10.1175/WAF-D-18-0206.1
[27] Ashley, W. S., Haberlie, A. M., & Strohm, J. (2019). A Climatology of Quasi-Linear Convective Systems and Their Hazards in the United States, Weather and Forecasting, 34(6), 1605-1631
[28] Shalev-Shwartz, S. and Ben-David, S. (2013). Understanding machine learning: From theory to algorithms, vol. 9781107057
[29] Team, D. (2019). Learn Types of Machine Learning Algorithms with Ultimate Use Cases, Data Flair. 2019, [Online]. Available: https://data-flair.training/blogs/types-of-machine-learning-algorithms/
[30] Pan, S. J., Kwok, J. T., and Yang, Q. (2008). Transfer learning via dimensionality reduction. In Proceedings of the 23rd national conference on Artificial intelligence - Volume 2 (AAAI'08). AAAI Press, 677–682
[31] Lin, Y., Zhu, X., Zheng, Z., Dou, Z., & Zhou, R. (2017). The individual identification method of wireless device based on dimensionality reduction and machine learning. The Journal of Supercomputing, 75, 3010-3027
[32] Verleysen M., François D. (2005) The Curse of Dimensionality in Data Mining and Time Series Prediction. In: Cabestany J., Prieto A., Sandoval F. (eds) Computational Intelligence and Bioinspired Systems. IWANN 2005. Lecture Notes in Computer Science, vol 3512. Springer, Berlin, Heidelberg

[33] Zebari , R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. Journal of Applied Science and Technology Trends, 1(2), 56 - 70
[34] Juvonen, A., Sipola, T., Hämäläinen, T. (2015). Online anomaly detection using dimensionality reduction techniques for HTTP log analysis, Comput. Networks, vol. 91, pp. 46–56, doi: 10.1016/j.comnet.2015.07.019.
[35] Liu, L., Zsu, M. T. (2009). Encyclopedia of Database Systems, Springer Publishing Company
[36] Abd-Alsabour, Nadia. (2018). On the Role of Dimensionality Reduction. Journal of Computers. 571-579. 10.17706/jcp.13.5.571-579
[37] Eesa, A., Orman, Z., & Brifcani, A. (2015). A new feature selection model based on ID3 and bees algorithm for intrusion detection system. Turkish Journal of Electrical Engineering and Computer Sciences, 23, 615-622
[38] Eesa, A. S., Orman, Z., Abdulazeez Brifcani, A. M. (2015). A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems, Expert Syst. Appl., vol. 42, no. 5, pp. 2670–2679, doi: 10.1016/j.eswa.2014.11.009
[39] Zeebaree, D. Q., Haron, H., and Abdulazeez, A. M. (2018). Gene Selection and Classification of Microarray Data Using Convolutional Neural Network, ICOASE 2018 - Int. Conf. Adv. Sci. Eng., pp. 145–150, doi: 10.1109/ICOASE.2018.8548836
[40] Abdulqader, D. M., Abdulazeez, A. M., and Zeebaree, D. Q. (2020). Machine learning supervised algorithms of gene selection: A review, Technol. Reports Kansai Univ., vol. 62, no. 3, pp. 233–244
[41] Jaiswal, S. (2018). "Dimensionality-reduction-technique." JavaTpoint, [Online]. Available: https://www.javatpoint.com/dimensionality-reduction-technique
[42] Abdi, H., and Williams, L. J. (2010). Principal component analysis. WIREs Comput. Stat. 2, 4 (July 2010), 433–459
[43] I. Compression, "Applications of PCA Image Compression with PCA," pp. 1–14
[44] Hassan, O. M. S., Abdulazeez,, A. M. and Tiryaki, V. M. (2018). Gait-Based Human Gender Classification Using Lifting 5/3 Wavelet and Principal Component Analysis, ICOASE 2018 - Int. Conf. Adv. Sci. Eng., pp. 173–178, doi: 10.1109/ICOASE.2018.8548909
[45] Bartholomew, D. J. (2010). Principal components analysis, Int. Encycl. Educ., pp. 374–377, doi: 10.1016/B978-0-08-044894-7.01358-0
[46] Najim Adeen, I. M., Abdulazeez, A. M., and Zeebaree, D. Q. (2020). Systematic review of unsupervised genomic clustering algorithms techniques for high dimensional datasets, Technol. Reports Kansai Univ., vol. 62, no. 3, pp. 355–374
[47] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M. and Zebari, D. A. (2019). Machine learning and Region Growing for Breast Cancer Segmentation, 2019 Int. Conf. Adv. Sci. Eng. ICOASE 2019, pp. 88–93, doi: 10.1109/ICOASE.2019.8723832
[48] Mahmood, M. R. and Abdulazeez, A. M. (2019). Different Model for Hand Gesture Recognition with a Novel Line Feature Extraction, 2019 Int. Conf. Adv. Sci. Eng. ICOASE 2019, no. 2018, pp. 52–57, doi: 10.1109/ICOASE.2019.8723731
[49] Zhuang, X., Yang, Z., and Cordes, D. (2020). A technical review of canonical correlation analysis for neuroscience applications, Hum. Brain Mapp., vol. 41, no. 13, pp. 3807–3833, doi: 10.1002/hbm.25090.
[50] Yahata, N., Kasai, K. and Kawato, M. (2017). Computational neuroscience approach to biomarkers and treatments for mental disorders, Psychiatry Clin. Neurosci., vol. 71, no. 4, pp. 215–237, doi: 10.1111/pcn.12502.
[51] Simó, C. (2020). Principal component analysis of financial statements. A compositional approach, Análisis en componentes principales de los estados financieros. Un enfoque composicional, no. 29, pp. 18–37, [Online]. Available: https://www.upo.es/revistas/index.php/RevMetCuant/article/view/3580/3930
[52] Pasini, G. (2017). Principal Component Analysis for Stock Portfolio Management, Int. J. Pure Apllied Math., vol. 115, no. 1, pp. 153–167, doi: 10.12732/ijpam.v115i1.12
[53] Dai, P.F., Xiong, X., Zhou, W. X. (2020). A global economic policy uncertainty index from principal component analysis, Financ. Res. Lett., p. 101686, doi: 10.1016/j.frl.2020.101686
[54] Yi, S., Lai, Z., He, Z., Cheung, Y. M. and Liu, Y. (2017). Joint sparse principal component analysis, Pattern Recognit., vol. 61, pp. 524–536, doi: 10.1016/j.patcog.2016.08.025
[55] Deshpande, N. T. and Ravishankar, S. (2016). Face Detection and Recognition using Viola-Jones algorithm and fusion of LDA and ANN, IOSR J. Comput. Eng., vol. 18, no. 6, pp. 1–6, [Online]. Available: https://pdfs.semanticscholar.org/c5cf/c1f5a430ad9c103b381d016adb4cba20ce4e.pdf.
[56] Wang, M., Song, L., Sun, K. and Jia, Z. (2020). F-2D-QPCA: A Quaternion Principal Component Analysis Method for Color Face Recognition, IEEE Access, vol. 8, pp. 217437–217446, doi: 10.1109/ACCESS.2020.3041847
[57] Saqlain Shah, S. M., Batool, S., Khan, I., Ashraf, M. U., Abbas, S. H., Hussain, S. A. (2017). Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis, Phys. A Stat. Mech. its Appl., vol. 482, pp. 796–807, doi: 10.1016/j.physa.2017.04.113

[58] Zhou, W., Wu, C., Chen, D., Yi, Y. and Du, w. (2017). Automatic Microaneurysm Detection Using the Sparse Principal Component Analysis-Based Unsupervised Classification Method," IEEE Access, vol. 5, no. c, pp. 2563–2572, doi: 10.1109/ACCESS.2017.2671918
[59] Thanh D.N.H., Hien N.N., Surya Prasath V.B., Erkan U., Khamparia A. (2020) Adaptive Thresholding Skin Lesion Segmentation with Gabor Filters and Principal Component Analysis. In: Solanki V., Hoang M., Lu Z., Pattnaik P. (eds) Intelligent Computing in Engineering. Advances in Intelligent Systems and Computing, vol 1125. Springer, Singapore
[60] Zhang J, Cui W, Guo X, Wang B, Wang Z. (2020). Classification of digital pathological images of non-Hodgkin's lymphoma subtypes based on the fusion of transfer learning and principal component analysis, Med. Phys., vol. 47, no. 9, pp. 4241–4253, 2020, doi: 10.1002/mp.14357
[61] Karl Pearson F.R.S. (1901) LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2:11, 559-572, doi: 10.1080/14786440109462720
[62] Hotelling, H. (1933). Analysis of a complex of statistical variables into Principal Components, Jour. Educ. Psych., 24, 417-441, 498-520, J. Educ. Psychol., vol. 24, pp. 417–441.
[63] Arif, M. T. (2010). Book Review: Book Review, Asia Pacific J. Public Heal., vol. 22, no. 4, pp. 507–508, doi: 10.1177/1010539510380245
[64] Ning, C. and You, F. (2018). Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods, Comput. Chem. Eng., vol. 112, pp. 190–210, doi: 10.1016/j.compchemeng.2018.02.007
[65] Zhang, M., Wang, H., Cui, Z. and Chen, J. (2018). Hybrid multi-objective cuckoo search with dynamical local search, Memetic Comput., vol. 10, no. 2, pp. 199–208, 2018, doi: 10.1007/s12293-017-0237-2
[66] Zhao, H., Zheng, J. Xu, J., and Deng, W. (2019). Fault diagnosis method based on principal component analysis and broad learning system, IEEE Access, vol. 7, pp. 99263–99272, doi: 10.1109/ACCESS.2019.2929094
[67] Fadhel, S., Delpha, C., Diallo, D., Bahri, I., Migan, A., Trabelsi, M., Mimouni, M.F. (2019). PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system," Sol. Energy, vol. 179, no. December 2018, pp. 1–10, doi: 10.1016/j.solener.2018.12.048
[68] Gajjar, S., Kulahci, M. and Palazoglu, A. (2018). Real-time fault detection and diagnosis using sparse principal component analysis," J. Process Control, vol. 67, pp. 112–128, 2018, doi: 10.1016/j.jprocont.2017.03.005
[69] Ghodsi, A. (2006). Dimensionality Reduction A Short Tutorial, Science (80-.)., p. 25, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.88.3592&rep=rep1&type=pdf%5Cnhttp://stats.stackexchange.com/questions/7111/pca-for-images-arrays-with-high-dimensionality
[70] Pořízka, P., Klus, J., Képeš, E., Prochazka, D., Hahn, D. W., and Kaiser, J. (2018). On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, Spectrochim. Acta - Part B At. Spectrosc., vol. 148, pp. 65–82, doi: 10.1016/j.sab.2018.05.030
[71] Liu, J., Wang, J., Xu, C., Jiang, H., Li, C., Zhang, L., Lin, J., & Shen, Z. (2018). Advanced Energy Storage Devices: Basic Principles, Analytical Methods, and Rational Materials Design, Adv. Sci., vol. 5, no. 1, doi: 10.1002/advs.201700322
[72] Fujiwara, T., Kwon, O. H. and Ma, K. L. (2020). Supporting Analysis of Dimensionality Reduction Results with Contrastive Learning, IEEE Trans. Vis. Comput. Graph., vol. 26, no. 1, pp. 45–55, doi: 10.1109/TVCG.2019.2934251
[73] Jackson, J. E. (1991). A User' s Guide to Principal Components. by J. Edward Jackson Review by: Brian D. 1991.
[74] Reddy, G. T. et al., (2020). Analysis of Dimensionality Reduction Techniques on Big Data, IEEE Access, vol. 8, pp. 54776–54788, doi: 10.1109/ACCESS.2020.2980942
[75] Saraswathi, V. and Gupta, D. (2019), Classification of Brain Tumor using PCA-RF in MR Neurological Images, 2019 11th Int. Conf. Commun. Syst. Networks, COMSNETS 2019, vol. 2061, pp. 440–443, doi: 10.1109/COMSNETS.2019.8711010
[76] Jamal, A., Handayani, A., Septiandri, A. A., Ripmiatin, E. and Effendi, Y. (2018). Dimensionality Reduction using PCA and K-Means Clustering for Breast Cancer Prediction, Lontar Komput. J. Ilm. Teknol. Inf., vol. 9, no. 3, p. 192, doi: 10.24843/lkjiti.2018.v09.i03.p08
[77] Salo, F., Nassif, A. B. and Essex, A. (2019). Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection, Comput. Networks, vol. 148, pp. 164–175, doi: 10.1016/j.comnet.2018.11.010
[78] Fournier, Q. and Aloise, D. (2019). Empirical comparison between autoencoders and traditional dimensionality reduction methods, Proc. - IEEE 2nd Int. Conf. Artif. Intell. Knowl. Eng. AIKE 2019, pp. 211–214, doi: 10.1109/AIKE.2019.00044
[79] Amir, M., Mousavi, H., and Bossaghzadeh, A. (2020). Improving Persian Digit Recognition by Combining Deep Neural Networks and SVM and Using PCA, Iran. Conf. Mach. Vis. Image Process. MVIP, vol. 2020-Febru, pp. 0–4, doi: 10.1109/MVIP49855.2020.9116893

[80] Allegretta, I. et al. (2020). Macro-classification of meteorites by portable energy dispersive X-ray fluorescence spectroscopy (pED-XRF), principal component analysis (PCA) and machine learning algorithms, Talanta, vol. 212, no. October 2019, p. 120785, 2020, doi: 10.1016/j.talanta.2020.120785
[81] Pu, H., Gao, P., Liu, Y., Rong, J., Shi, F., and Lu, H. (2019). Principal Component Analysis Based Dynamic Cone Beam X-Ray Luminescence Computed Tomography: A Feasibility Study," IEEE Trans. Med. Imaging, vol. 38, no. 12, pp. 2891–2902, doi: 10.1109/TMI.2019.2917026
[82] Astuti, W. and Adiwijaya, (2018). Support vector machine and principal component analysis for microarray data classification, J. Phys. Conf. Ser., vol. 971, no. 1, doi: 10.1088/1742-6596/971/1/012003
[83] Kim, S., Kang, D., Huo, Z., Park, Y. and Tseng, G. C. (2018). Meta-analytic principal component analysis in integrative omics application, Bioinformatics, vol. 34, no. 8, pp. 1321–1328, doi: 10.1093/bioinformatics/btx765
[84] Yang, W., Si, Y., Wang, D., & Guo, B. (2018). Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine. Computers in biology and medicine, 101, 22–32. https://doi.org/10.1016/j.compbiomed.2018.08.003
[85] Xu, S. et al., (2020). The fuzzy comprehensive evaluation (FCE) and the principal component analysis (PCA) model simulation and its applications in water quality assessment of Nansi Lake Basin, China, Environ. Eng. Res., vol. 26, no. 2, pp. 200022–0, doi: 10.4491/eer.2020.022
[86] Lazcano, R. et al. (2017). "Porting a PCA-based hyperspectral image dimensionality reduction algorithm for brain cancer detection on a manycore architecture," J. Syst. Archit., vol. 77, pp. 101–111, doi: 10.1016/j.sysarc.2017.05.001
[87] Raunak, V. (2017). Simple and Effective Dimensionality Reduction for Word Embeddings, arXiv, no. Nips, pp. 1–6
[88] Jiang, J., Ma, J., Chen, C., Wang, Z., Cai, Z. and Wang, L. (2018). SuperPCA: A Superpixelwise PCA Approach for Unsupervised Feature Extraction of Hyperspectral Imagery, IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4581–4593, doi: 10.1109/TGRS.2018.2828029
[89] Jollife, I. T. and Cadima, J. (2016) Principal component analysis: A review and recent developments, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., vol. 374, no. 2065, doi: 10.1098/rsta.2015.0202