International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Integration of Generalized Discriminant Analysis and Classification Technique for Identification Well Test Interpretation Model

Saifi Redha[1], Zeraibi Nourreddine[2]

Submitted: 07/05/2023 **Revised**: 16/07/2023 **Accepted**: 05/08/2023

Abstract: This paper presents a hybrid method that combines generalized discriminant analysis and machine learning technique for identifying well test model. The proposed method consists of three stages: (1) nonlinear combination of features spaces to maximize the separability among the class models through generalized discriminant analysis. (2) Construction a set of classifier and classify the new data points by a plurality vote of their prediction. This method is described in detail to ensure successful replication of results. The required training and test dataset were generated by using analytical solution models. In our case, there were used 600 samples: 70% for training, 15% for validation, and 15% for testing. We notice that the generalized discriminant analysis is an effective strategy in reducing the number of input features, simplifying the network structure, and lowering the training time of the ANN. The results obtained show that the proposed model provides better performance when predicting new data with a coefficient of correlation approximately equal to 99%.

Keywords: *well test, neural networks, generalized discriminant analysis, classification.*

1. Introduction

Well-test transient analysis is considered one of the most effective reservoir management tool to estimate the reservoir properties such as initial pressure, well permeability, skin effect…etc. by analyzing pressure transient data vs. time data from a producing or shut-in well. These data allow our engineers to plan the optimal exploitation of the reservoir. Finding the proper well-test interpretation model is the most required task to get a successful interpretation analysis even though we have a good quality of measurement. There are many available techniques to select the best reservoir model. The pressure derivative type curve is one of these important techniques [1], which can be used to identify all flow regimes present in pressure transient data. Selecting the appropriate reservoir model from the pressure derivative type curve is a hard step where different models may apparently satisfy the available information about the reservoir. On other hand, artificial neural networks have found many applications in the petroleum industry because of his ability to learn complex nonlinear relationships and many research works have been done in identifying the reservoir model since the work done by Lee and Al-Kaabi in 1988 [2,3]. They have attempted to improve the efficiency of the classifier by changing classification algorithms (Neural networks, Fuzzy Logic, self-organizing… etc.), training data, and the number of models considered in the classifier [4-6]. Despite considerable time and effort have been devoted to establish a procedure for automatic classification, these procedures have always difficulties to select the list of appropriate models [7, 8]. The focus of this paper is to develop a new classification model for identifying the well-test interpretation based on neural networks[9-11] and generalized discriminant analysis[12] to increase the modeling accuracy and compare the results with the basic methods. To achieve our goal, we conducted this study by distinguishing the followings steps:

- Reducing the dimensionality while maximizing the separability among the class models through generalized discriminant analysis.
- Generating ten automatic reservoir models through artificial neural networks. Models are trained on different subsets of the pre-processed training data by resampling methods.
- The final output prediction is the one that receive more than half of the votes.

The developed model is trained on a set of 600 data points generated using analytical solution models: 300 for training, 100 for cross-validation, and 200 for testing.

This paper is organized as follow: section II introduces artificial neural network and generalized discriminant analysis. Section III provides the methodology of proposed model and the discussion of the results. Finally, the paper is concluding in section IV.

[1] *Physics Engineering of Hydrocarbons(LGPH), Boumerdes, Algeria*
ORCID ID : 0000-0002-5434-6096
[2] *Physics Engineering of Hydrocarbons(LGPH),Boumerdes Algeria*
ORCID ID : 0009-0007-7880-1958
** Corresponding Author Email: r.saifi@univ-boumerdes.dz*

2. Theory Background

2.1. Artificial neural network

The term neural network derives its origin from the human brain or the human nervous system, which consists of a massively large parallel interconnection of a large number of neurons. The idea behind neuron networks is that we have nodes that have some connection between them. We trigger a node with some input and that node, in turn, triggers the nodes it is connected to. Each node is given what is called a transfer function to judge inputs. The most commonly used structure is formed in three layers, called the input layer, hidden layer, and output layer. The input layer is responsible for receiving information (data) and each value from the input layer is sent to all of the hidden nodes, the values entering a hidden node are multiplied by weights and then summed. This produces a single value that is passed through a nonlinear function called sigmoid. Each value in the hidden layer is duplicated and applied to the next layer then the output layer combines and modifies the data to produce values of this network. This can be expressed as follows:

$$y_{jk} = f_k\left(\sum_{i=1}^{N_{k-1}} w_{ijk} y_{i(k-1)} + b_{jk}\right) \quad (1)$$

Where

- y_{jk} are j th neuron outputs from k th layer a b_{jk} is the bias weight for neuron j in layer k
- w_{ijk} are the connection weight
- f_k are the transfer function

Given a set of training examples of the form (x, m) where $x \in X$ represent model signature (pressure derivative). A classifier is considered a mathematical function that maps input data to a category class. We denote the probability of signal whose target is i th class that classified as j th class with symbol

$$p(m_i^* / m_j) = \frac{S_j^i}{\sum_{i=1}^{N} S_j^i} \quad (2)$$

Where

- m_i^* is the label of i th output class
- m_j is the label of j th target class
- S_j^i are the number of samples whose target are the i th class that were classified as j
- N the number of class

An optimal scenario is the statement when $p(m_i^* / m_j) = 0$ in case $i \neq j$ and

$p(m_i^* / m_j) = 1$ in case $i = j$

In other words, this value represents the classification accuracy of the classifier network to class the signal from the i ith class model into the j th class model.

2.2. Generalized discriminant analysis

GDA, also known as kernel fisher discriminant analysis, is a nonlinear combination of features spaces that maximizes the separability among two or more classes.

Let $X = \{x^{(1)}, x^{(2)}, ..., x^{(m)}\}$ be the dataset given d -dimensional vectors. Each data point belongs to one c model classes $\{X_1, X_2, ..., X_c\}$. If we map our dataset into high dimension space via some function ϕ, it is easy to find a hyper-plane w that can separate different class. In this new features space F, the hyper-plane w is created according to two criteria (considered simultaneously):

- maximize The generalization of the between-class scatter S_b^ϕ
- minimize The generalization of the within-class scatter S_w^ϕ

The function that needs to be maximized, with respect to w, is:

$$w_{opt}^\phi = \arg J(w) = \frac{w^T S_B^\phi w}{w^T S_w^\phi w} = [w_1^\phi, ..., w_1^N] \quad (3)$$

The between class scatter matrix and within class scatter matrix are defined as:

$$S_B^\phi = \sum_{c=1}^{C} N_c m_c^\phi \left(m_c^\phi\right)^T \quad (4)$$

$$S_w^\phi = \sum_{c=1}^{C} \sum_{x \in X_c} \phi(x) \phi(x)^T \quad (5)$$

Where m_c^ϕ is the mean of class X_c in F and N_c is the number of samples belongs to X_c. The vector w_{opt} can be found as the solution of the generalized eigenvalue problem:

$$S_B^\phi w_{opt}^\phi = \lambda_i S_w^\phi w_{opt}^\phi \quad (6)$$

The kernel trick is used to avoid the explicit the mapping function ϕ in which the dot product in the new feature space F is replaced by a kernel function

$$k(x, y) = \phi(x)^T \phi(y).$$
(7)

3. Case Study

The required dataset analyzed during the current study were generated by using analytical model (software). We have used 600 samples as follows: 400 for training, 15% for validation, and 15% for testing. The input vector $x'_1 y'_1\ x'_2 y'_2 \ldots x'_{30} y'_{30}$ are sampled from the normalized curve of derivative pressure (figure 2) where:

$$x_i' = x_i \Big/ \sqrt{\sum x_j^2} \quad (8)$$

$$y_i' = y_i \Big/ \sqrt{\sum y_j^2} \quad (9)$$

This data preprocessing is a necessary step before building any model with machine learning in order to make the training less sensitive to the scale of features. Any normalized scheme can be used. the output vector has 6 element where one element is 1 and the others are 0. Each value in the element is interpreted as the probability of membership for each class model. The list of the class model label in this study are:

- node #1: Homogenous+ infinite action radial flow
- node #2: Homogenous+ constant pressure boundary
- node #3: Homogenous+ no flow boundary
- node# 4: 2-porosity + infinite action radial flow
- node #5: 2-porosity + constant pressure boundary
- node #6: 2-porosity + no flow boundary

The proposed method consists of two steps:

1. Nonlinear discriminant analysis to extract low dimensional and discriminating axes that maximizing the separation between classes and use the features vectors were used for classification.
2. Create ten classifiers using artificial neural network. each model is trained using different subset of the training data

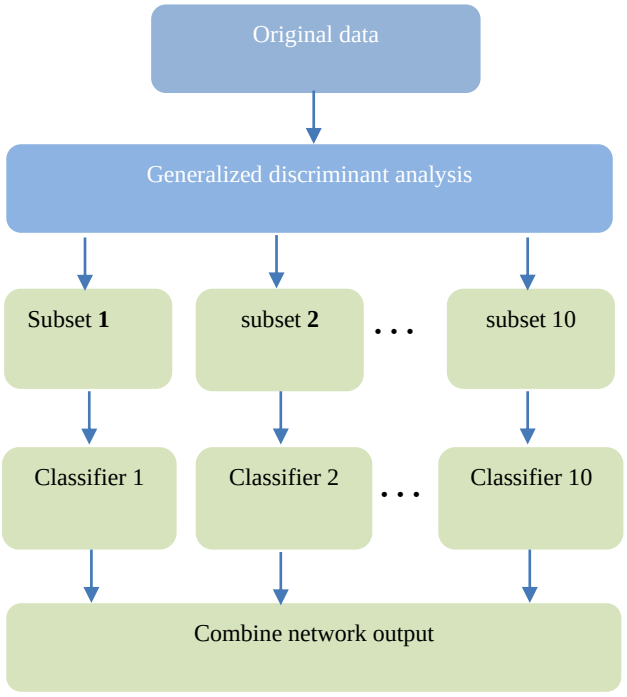

Fig 1. Structure of the proposed well-test identification model

The new reduced training dataset are obtain by projection the raw training data into the eigenvector correspond to d-largest eigenvalue of the generalized characteristic equation. Each eigenvalue indicate how well the eigenvector differentiate the classes.

figure 3 showed that 5 discriminant components are sufficient enough to explain 99% of the total variance in the original dataset while the remaining can be safely be dropped without losing information. Therefore, we have reduced 60-dimensional data into 5-dimensions, which are used as input data to train the ten different models. The set of neural network were created by Matlab neural network tool. We have used the most common structure neural network architecture which consists of three layers as input, hidden, and output layers. Input layer with five nodes $GD_1, GD_2, \ldots, GD_5$ representing the critical features considered from the output of GDA. Different hidden layer architectures were tested in this study to arrive at optimum network design. Two hidden layers with eight neurons were determined according to the test performance to get the maximum value. The output layer has six nodes, where each node corresponds to a specific model class in the data. The value of each output node essentially represents the probability of that class being the correct class.

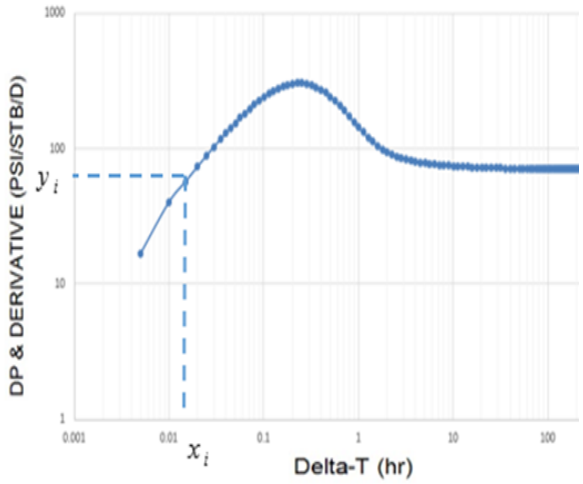

Fig 3. .Derivative pressure plot

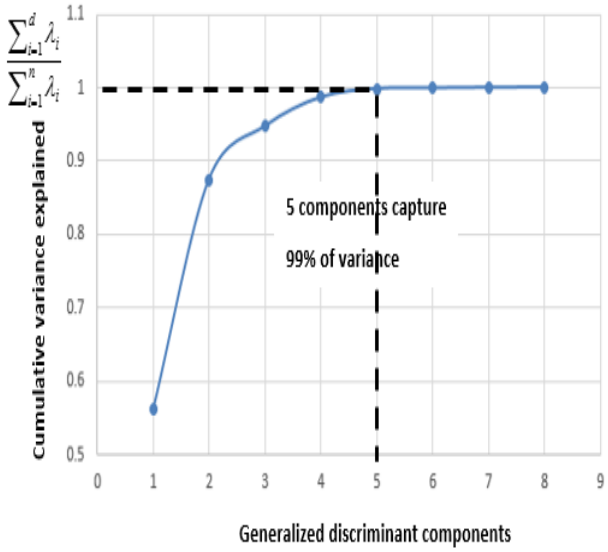

Fig 2. Cumulative variance explained for 60 discriminat components

Figure 4 and Figure 5 show the confusion matrix results for the first ANN model where we get a classification accuracy of 98.7% for transformed data through GDA and about 95.8% when using raw data. Dimensionality reduction using GDA is one of the data preprocessing concepts in machine learning, which hugely affects the performance of the **model.** A high dimensional data, network architecture with a huge input layer, is an obstacle for solving optimization problems in machine learning and renders it less accurate because the number of samples needed to estimate connection weights grows exponentially with the number of input variables.

There are two risks in a bunch of networks classification:

- Good class is not detected
- Appearance of bad class

For the first case the probability of getting bad class output is P(none detected by the first ANN model) * P(none detected by the second ANN model)*…*P(none detected by the tenth ANN model)= 0.1^10 , so we can neglect this case.

For the second case, even if the risk factor value is equal to 0.1, the probability to get a least one misclassified is P(at least one)= 1- p(none)^(number of the model) =1- 0.9^10= 0.65, so the most embarrassing case is the appearance of bad classes. In this case, we will use the most simple and intuitive approach to find the final output prediction is based on majority voting rule. (table I) show that we can get by this method, ensemble learning, a classification accuracy of 100%

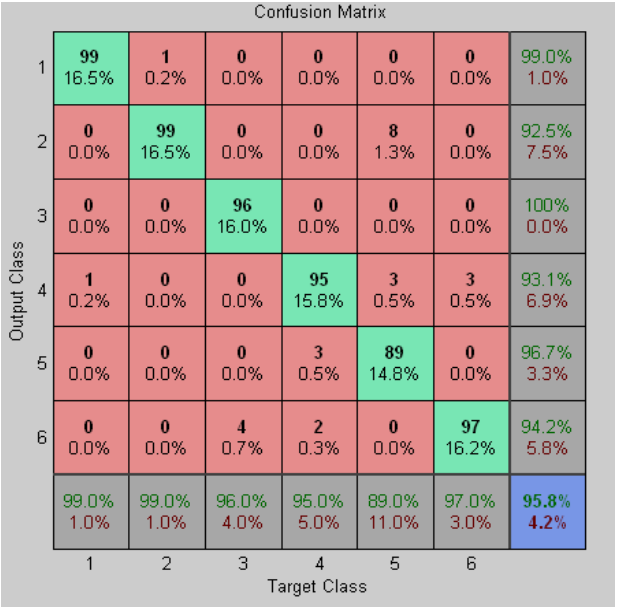

Fig 4. Confusion matrix for all classes and all attributes of the first neural network without GDA analyis

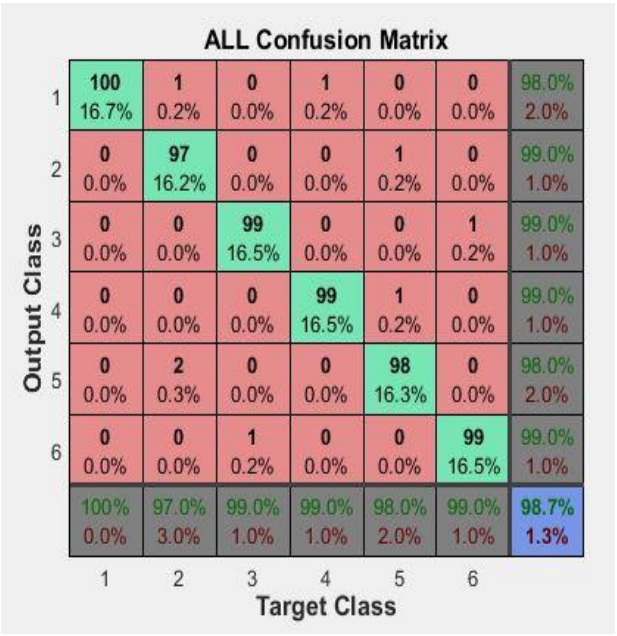

Fig 5. Confusion matrix for all classes and all attributes of the first neural network with GDA analysis

Table 1. Confusion matrix of the ensemble ANN models

		Actual class					
		m_1	m_2	m_3	m_4	m_5	m_6
Predicted Class	m_1^*	100	0	0	0	0	0
	m_2^*	0	100	0	0	0	0
	m_3^*	0	0	100	0	0	0
	m_4^*	0	0	0	100	0	0
	m_5^*	0	0	0	0	100	0
	m_6^*	0	0	0	0	0	100

4. Conclusions

- the generalized discriminant analysis is an effective strategy in reducing the number of input features, in simplifying the network structure ANN and in improving the performance of the ANN
- Model based on bagging neural networks and principal component analysis is an important technique to overcome the curse of dimensionality and produce accurate prediction in classification model.
- We can get by this method a classification accuracy of 100%. thus, implementation of GDA-ANN can be part of well-testing analysis to identify well-test interpretation model

References

[1] Bourdet D, Whittle T ,Douglas A, Piard.Y . A new set of type curves simplifies well test analysis. World Oil (1983) May: 95.

[2] Al-Kaabi A, Lee W. Using artificial neural networks to identify the well test interpretation model . Paper Society of Petroleum Engineers 20332 presented at the SPE petroleum computer conference(1993), Denver. doi.org/10.2118/20332-PA

[3] Ahmadi, R., Shahrabi, J. Aminshahidy, B.Automatic well-testing model diagnosis and parameter estimation using artificial neural networks and design of experiments. Journal of Petroleum Exploration and Production Technology (2017) 7: 759. doi.org/10.1007/s13202-016-0293-z

[4] Ali S. Al-Bemani, Boyun Guo & Ali Ghalambor (2003) The Challenge of Model Identification in Well Test Interpretation—A Unique Build Up Analysis Case Study, Petroleum Science and Technology, 21:5-6, 879-899, DOI: 10.1081/LFT-120017455

[5] Kumoluyi, A, Daltaban, T, Archer J. Identification of Well-Test Models by Use of Higher-Order Neural Networks. Society of Petroleum Engineers (1995). doi:10.2118/27558-PA

[6] Sinha S, Panda M. Well-Test Model Identification With Self-Organizing Feature Map. Society of Petroleum Engineers(1996). doi:10.2118/30216-PA

[7] Iraj E,Woodbury J. Examples of Pitfalls in Well Test Analysis. Society of Petroleum Engineers(1985). doi:10.2118/12305-PA

[8] Juniardi I,Ershaghi I. Complexities of Using Neural Network in Well Test Analysis of Faulted Reservoirs. Society of Petroleum Engineers(1993). doi:10.2118/26106-MS

[9] H. Alizadeh, B. Minaei Bidgoli, "Introducing A Hybrid Data Mining Model to Evaluate Customer Loyalty," Engineering, Technology & Applied Science Research, vol. 6, no. 6, pp. 1235-1240 December 2016.

[10] L. Belhaj Salah and F. Fourati, "Systems Modeling Using Deep Elman Neural Network", *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 2, pp. 3881–3886, Apr. 2019.

[11] A. Sokhal, Z. Benaissa, S. A. Ouadfeul, and A. Boudella, "Dynamic Rock Type Characterization Using Artificial Neural Networks in Hamra Quartzites Reservoir: A Multidisciplinary Approach", *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 4, pp. 4397–4404, Aug. 2019.

[12] Manukumar N.M and Manjunatha Hiremath, A Hybrid Algorithm for Face Recognition using PCA, LDA and ANN, International Journal of Mechanical Engineering and Technology 9(3), 2018, pp. 85–93.

[13] Mr. R. Senthil Ganesh. (2019). Watermark Decoding Technique using Machine Learning for Intellectual Property Protection . International Journal of New Practices in Management and Engineering, 8(03), 01 - 09. https://doi.org/10.17762/ijnpme.v8i03.77

[14] Sreeramulu, T. ., Devi, L. N. ., Rao, A. R. ., & Laxminarayana, G. . (2023). Ring Laser Gyros: Improving Precision and Accuracy Through Soft-Core Processor-Based Active Current Balance Control Approach- Simulation and Implementation Results. International Journal on Recent and

Innovation Trends in Computing and Communication, 11(4), 141–147. https://doi.org/10.17762/ijritcc.v11i4.6397

[15] Anand, R., Khan, B., Nassa, V. K., Pandey, D., Dhabliya, D., Pandey, B. K., & Dadheech, P. (2023). Hybrid convolutional neural network (CNN) for kennedy space center hyperspectral image. Aerospace Systems, 6(1), 71-78. doi:10.1007/s42401-022-00168-4