Feature Selection for Gabor Filter Based on Level Measurement using Non-Interacting Tanks Level Images

Kalaiselvi B[1], Karthik B[2], C. V. Krishna Reddy[3]

[1]Research Scholar, Dept of ECE, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India

[2]Associate Professor, Dept of ECE, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India

[3]Professor and Director, Nalla Narasimha Reddy Group of Institutions, Hyderabad, Telangana, India

Abstract

Level measurement models using image-based classifiers (pixel-based datasets) are used for estimation purposes. Pre-processing is thought-provoking in proceeding out the image filter technique and classifying the level. The level scenario of a two non-interacting tank system plays a vital role in predicting the level. Level monitoring is done using the supervised learning method using instance-based filters (Gabor Filter) and selected base classifiers for level measurements. The main scope of this case study is to improve the level measurements from the two non-interacting tank scenarios using Artificial Intelligent algorithms. The suggested article includes the finest feature selection process to increase the accuracy performance attained by the designated classifiers like IBK Instance base classifier for different neighbourhood values and Tree category algorithm like Random Forest. The performance accuracy in level prediction obtained is 81.356%, the weighted Average of Receiver operator characteristics of (ROC) 0.931 are obtained by Random Forest Tree Category Classifier.

Key words: Level Monitoring, Gabor filter, machine learning, KNN, Random Forest, ROC.

1.Introduction

The process of constructing models from experimental data is called system identification. System identification involves building a mathematical model of a dynamic system based on set of measured stimulus and response samples. It is a process of acquiring, formatting, processing, and identifying mathematical models based on raw data from the real-world system. Once the mathematical model is chosen, it can be characterized in terms of suitable descriptions such as transfer function, impulse response, or power series expansions and that can be used for controller design. Tri Chandra S. Wibowo.et.al [1] have done System Identification of an Interacting Series Process for Real-Time Model Predictive Control. This paper aims at identifying a linear time-invariant (LTI) with lumped parameters state space model of the gaseous pilot plant which has a typical structure of interacting series process, and the model has been developed around an operating point. Edward P. Gatzke.et.al [2] have done work on Model based control of a four-tank system. In this paper, the authors used sub space process modelling and hence applicable to operating point.Nithya.et.al [3] have done work on model-based controller design for a spherical tank process in real time. They have proposed tangent and point of inflection methods for estimating FOPTD model parameters. The major disadvantage of all these methods is the difficulty in locating the point of inflection in practice and may not be accurate. Gatzke et al [2] perform the parametric identification process of a quadruple tank using subspace system identification method. Such a system has series structure with recycles and the input signals used are the pseudo-random binary sequence (PRBS). The identification process is carried out without considering the prior knowledge of process, and no assumption are made about the state relationships or number of process states. Weyer [5] presents the empirical modeling of water level in an irrigation channel using system identification technique with considering the prior physical information of the system. The identified process is a kind of interacting series process; however, the model only has a single output variable. To accurately control a system, it is beneficial to first develop a model of the system. The main objective for the modeling task is to obtain a good and reliable tool for analysis and control system development. A good model can be used in off-line controller design and implementation of new advanced control schemes. In some applications, such as in an industrial sewing machine [6] , it may be time consuming or dangerous to tune controllers directly on the machinery. In such cases, an accurate model must be used off-line for the tuning and verification of the controller [7]. In the present work, the system identification of Interacting and Non-interacting tank systems is found using genetic algorithm which is working out for full region and the results are compared with Process Reaction curve method, ARX model, and Statistical model of Identification [8]. Also, we propose a method to obtain an accurate nonlinear system model based on neural networks (NNs) and Fuzzy logic [9]. Modelling techniques based on NNs, and Fuzzy logic have proven to be quite useful for building good quality models from measured data [10].

2. Literature survey

Changhyun Choi et al (2019) [11], Advancement of Water Level Forecast Models Utilizing AI in Wetlands utilizing different AI models like fake brain organization (ANN), rules-based class like choice tree (DT), trees-based class like irregular backwoods (RF), and capabilities-based class support vector machine (SVM). F. N. M. Ariff, et al (2020) [12] A Person Division for Programmed Vehicle Tag Acknowledgment In view of Quick K-Means Grouping" proposed to uncover and become mindful of the car with the guide of utilizing dissecting the auto enlistment code numbers and find the distinctions of the plate character. The got photograph has sectioned the utilization of Quick K means grouping and Fluffy 'K' signifies calculations with a precision of an extra the 80%. Realities set of 100 photographs are utilized for tutoring and evaluating purposes. The strategy concerned is division with grouping processes executed. Fluffy like arrangement of rules and in examination for FCM show a scarcely better normal percent of precision. V. V. Mainkar, et al (2020) [13] Involved pre-processing, division, capability extraction, and Post processing on OCR to figure out the individual. A simple way that might be applied is through essentially taking pictures the photograph of the written by hand data into an electric sum gentle tweaked by means of which the individual is perceived. The proposed gadget utilizes an android versatile cell phone to hold onto the photograph and optical individual acknowledgment. The hard test is the exhibition of the gentle sign to comprehend stand-out penmanship styles. The level of exactness in generally execution is prepared 90% for the manually written records analyzed data set. Y. Peng et al (2020) [14], acknowledgment on investigating the water meter man or lady automatically, which is moreover an extraordinary methodology for mechanized personality approach the use of Profound Brain Organization. This also proposes computerized man or lady investigating from the got submerged photo and prompted unnecessary exactness [15]. Independent water meter photo dataset is advanced the utilization of area basically totally based convolution networks for objective discoveries with various emphasis to as it ought to be concentrate on the water meter characters. The essential kind calculations are executed for personality [16]. The outcomes suggest that the prevalence charge is unreasonable that could address the issues of far-reaching organizations.Prescient calculations in AI like Credulous Bayes, Occurrence base classifiers KNN, Choice Tree, and Irregular Woods are utilized for the early location of Diabetes. Outfit cross breed model expanded the precision and execution time. Mohammed M Mazid, A. B. et al (2013) [17]. C4.5 Calculation for Rule-Based Grouping has been moved along. Computerized reasoning, Information Designing, and Information Bases: Ongoing Advances.C4.5 is a well-known calculation a standard based classifier with exceptionally high precision and execution time. Nagaparameshwara careful, S. D. et al (2017) [18]. A Study of Choice Tree Calculations in Information Mining and Their Relative Examination. Imaginative Designing and Data Innovation Applications Worldwide Gathering. An imaginative relative concentrate on ML calculations like choice tree calculations like Truck, ID3, C4.5, CHAID and their exhibition and precision assessment are completed. Jian Huang1 et al (2020) [19] proposed an original programmed simple instrument perusing framework utilizing PC vision and an investigation robot has been introduced. Paaranan Sivasothy, et al (2018) [20] proposed AI based filling level assessment for mass strong storehouses and arrange different filling levels

3. Methodology

The existing resources help to make the suggested framework more understandable. This section also discusses the motivation behind classifiers as well as their representation models.

3.1 Gabor filter

Gabor filter is used for feature extraction. This filter is specially designed for statistical information of character structures. Gabor filter gives outputs to achieve better performance on low pixel quality images. An adaptive sigmoid function is applied to the outputs of Gabor filters to achieve better performance on low-quality images. To enhance the discriminability of the extracted features, the positive and the negative real parts of the outputs from the Gabor filters are used separately to construct histogram features. Experiments show that the proposed method has excellent performance on low-quality machine-printed character recognition and cursive handwritten character recognition.

The transfer function G(k) of a Gabor filter (Fourier transform of the impulse response) is given by

$$G_{mn(k)} = e^{-1/2(k-k_{0mn})(A_{mn}^{-1})(k-k_{0mn})} \tag{1}$$

where $k = k_1 k_2$ is the spatial frequency. To establish a multi-resolution strategy, the image can be filtered with a set of N Gabor filters with different bandwidths and modulation frequencies. If the modulation frequencies are given by

$$k_{0n=\frac{\pi}{2^{n+1}}:n\in[0\ldots\ldots N-1]} \tag{2}$$

and the relative bandwidth is chosen to be constant for all filters the image is decomposed into octaves.

3.2 Gabor Feature filter in the combination of selected Attributes

The collected data is first converted into numeric utilizing the image filter like Gabor, as shown in figure 1 is the procedure in the Weka tool (Using this image filter, the image data set is converted from pixels to statistical features). This ARFF (Attribute Relation File Format) file is used for further classifying process, some of the selected classifiers are used to predict the level and the accuracy performance is evaluated and compared to converge into getting an optimum model.

3.2.1 Parameter Training

Parameter training is for refining the predictions done by the selected classifiers using machine learning algorithms. Here IBk classifier with changing number of nearest neighbours (k) values forecasts the accuracy level and produces the Receiver Operator Characteristics values

3.2.2 Attribute selection in Weka

In Weka, there are 3 options for performing attribute selection

- The built-in approach, using the attribute selection tab.
- Openly using a meta category type classifier.
- Starting with filter approaches.

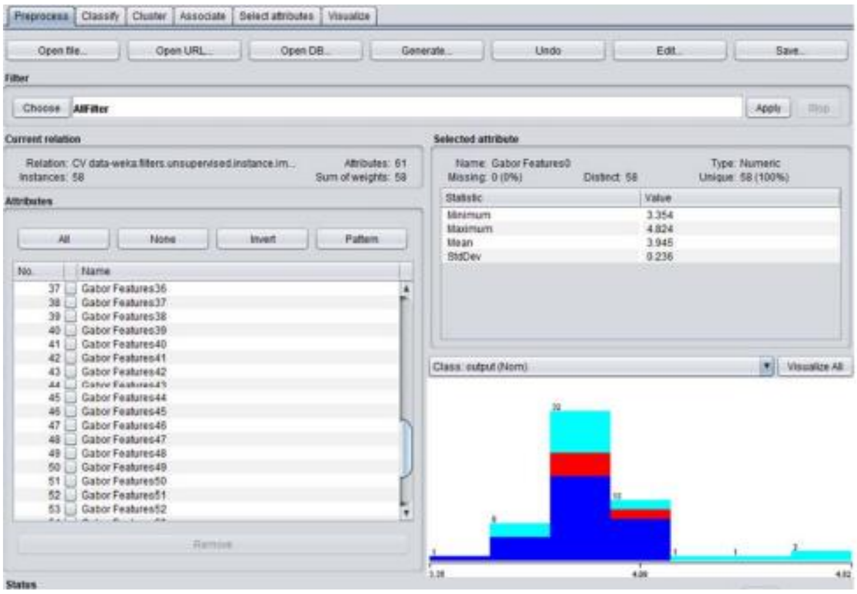

Fig. 1: Weka Gabor Filter based data feature extraction.

This is a novel method of evaluation feature subset. This means that it tries every possible combination of the variables and returns the best-performing subset. Perform well in improve accuracy and less misleading data, reduce training time, and reduce overfitting by removing the irrelevant data and redundant data. To create an optimal Machine learning model, it is critical to remove unnecessary attributes from the database. Some automatic feature selection tools can be found in Weka. We choose its attribute selection methods like InfoGain Attribute Evaluator /Ranker Method available in this tool to analyse attribute ranking order since we regard Waikato Environment for Knowledge Analysis (WEKA) to be an open-source tool. To improve the performance of the accuracy percentage, we constantly delete attributes in lower order. To keep the number of iterations to a minimum, we repeat the list of sorted lists of attributes in relation to the selected threshold. The attribute evaluator applied for the selected database is examined in such a way it correctly classifies the output class

3.3 Designated Base Machine Learning Algorithms

Selected two classifiers are that play a vital role in building a model to predict the level parameter using the supervised learning algorithms to fit our modelling procedure and their knowledge and modest implementation foremost to easier clarifications.

3.3.1 IBk Instance base classifiers

One of the traditional algorithms in the machine learning concept that is the lazy based categorized machine learning algorithm like IBk the nearest neighbourhood where the data are distributed in a hyper dimensionally and the manipulation is done depending on the “k” values nearest neighbours by just imitating the tags of the weight without any specific effects on prediction. Although it is very complex calculations, it is very unpretentious for execution.

Lazy category classifier, which Imitates 'k' nearest neighbour, an instance-based classifier IBk with k neighbours is implemented without taking care of any specific model, though it is computationally expensive as it is based on Euclidean distance between any two instances in the data set. Hence the laziness of this classifier for fitting the model turns out to be an advantage for escaping from any representational specific classifier. KNN classifier is a nonparametric algorithm which means it does not make any assumptions on underlying data. Because it just learns from the training set, it stores the data set, at the time of classification it performs an action on the data set. During the training phase, it stores the data set and when it gets new data then it classifies that data into a category that is much like the new data. This algorithm is explained based on the below-mentioned steps:

Step 1: Select the number of K, the neighbors

Step 2: Euclidean distance of K number of neighbor

Step 3: Take just neighbors as per the calculated Euclidean distance

Step 4: For undefined among these k neighbors count the number of data points in each category

Step 5: Assume the new data points to the category for which the number of neighbors maximum

Step 6: The model is built using this algorithm.

Figure 3 specifies KNN Classifier in the parameter tuning option panel provides the option for changing the neighborhood values, the batch size can be varied, and a search algorithm can be varied. It has the capability to handle Binary class, Date class, Missing class values, Nominal class, Numeric class.

It uses nearest neighborhood search algorithms like ball tree, filter neighborhood search, KD tree, and by default linear NN Search applied. The output number in the module is two by default. The maximum number of instances allowed in the training pool is zero. The batch size is 100 by default.

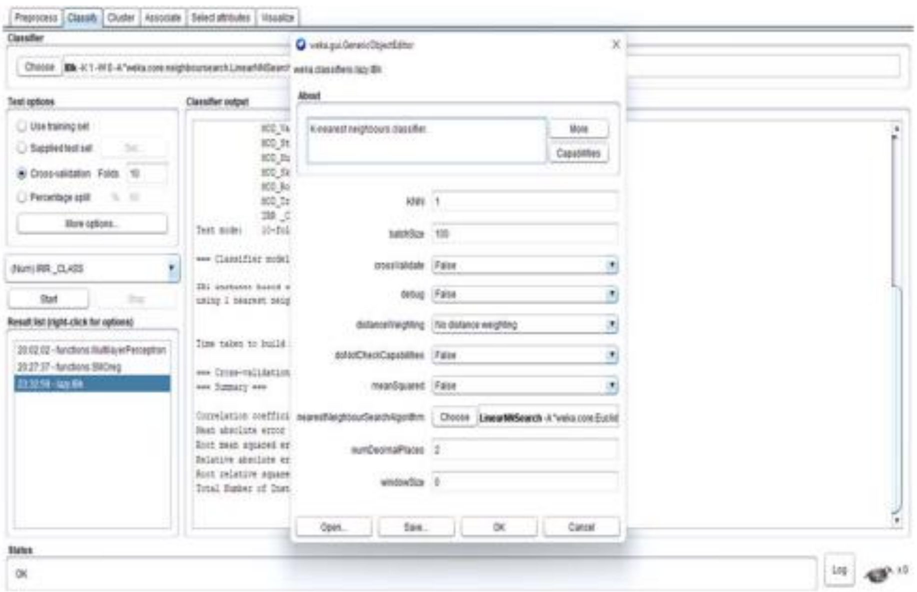

Fig. 2: KNN Classifier in parameter tuning option panel.

3.3.2 Random Forest

This is also a renowned type of tree category machine learning algorithm. The ensemble machine learning algorithm in weka. create multiple decision trees and merge them to get a more accurate and stable prediction. These decision tree-based predictors are best known for their computational power and scalability. However, in the case of highly unbalanced training data, as is often seen in data from medical studies with large control groups, the training algorithm or sampling method should be changed to improve the prediction quality for the classes. minority. In this work, a balanced random forest approach is proposed for WEKA.

It is an ensemble type of classifier. Mainly the bagging with base tree classifier. The main advantage is being an ensemble faster and more accurate than the base tree classifiers. One of the most powerful algorithms is Random Forest whose accuracy is high and training time is low. Moreover, multiclass object detection can be accommodated.

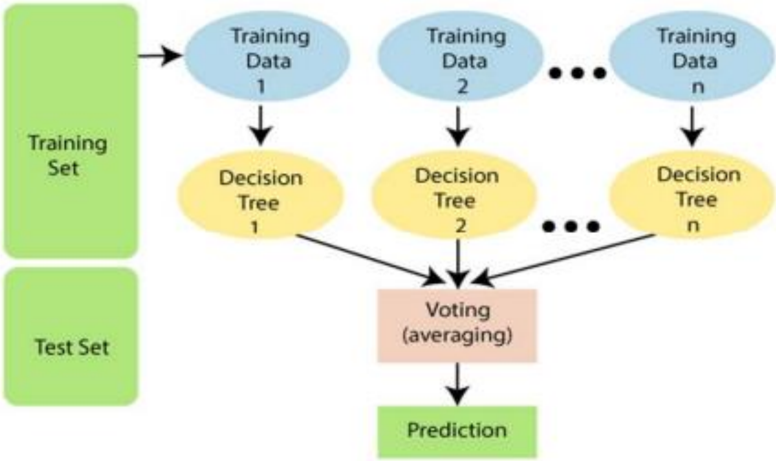

Fig. 3: Splitting of the dataset in Random Forest classifier.

Below are some points that explain why you should use the Random Forest algorithm as shown in figure 3.

- Training time is reduced compared to other algorithms.
- Predict output with high accuracy, even for large datasets that run efficiently.
- Maintain accuracy even if a large amount of data is missing.

The work process can be described in the following steps and diagrams.

Step 1: Select a random K data point from the training set.

Step 2: Build a decision tree associated with the selected data points (subset).

Step 3: Select the number N of decision trees to build.

Step 4: Repeat steps 1 and 2.

3.4 Evaluating the Performance Criteria

The performance measures are derived from the prediction patterns provided by the classifiers in use, which are represented in the confusion matrices' available inputs. High Level (above the setpoint), Low Level (below the setpoint), and Medium Level (in between the setpoints) are the three levels (around 5 percent plus or minus the setpoint). The Receiver Operating Characteristic (ROC) area, a performance metric, is calculated at the end.

3.5 Data Demonstration and Report Generation

The main materials come from the existing laboratories, which contain a system of 2 tanks that do not interact, and the photos of the multi-level images are taken with a conventional smartphone and the level is measured with the lever indicator attached to the tanks. Table 1: Experimental dataset for level prediction S. No Level type Instances count Class description of Level 1 High level 148 Above the set point 2 Medium level 112 Around +-5% with set point 3 Low level 94 Below the set point This is the first processing step. The final level to achieve the optimal model is the main objective to measure the level. Table 1 demonstrates how the three-class values are distributed with their Instances count. The classes High Level (above the setpoint), Low level (below the setpoint), and Medium (around 5% plus or minus the setpoint). Flow chart for Level Monitoring describes the flow chart for the proposed system is shown and the experimental setup to predict the level is done by collecting the data from non-interacting tank scenarios from a normal camera image in JPEG format and the level is predicted using the conventional method. The collected data is pre-processed in WEKA software for the implementation of the machine learning algorithm with the Gabor filter (feature extraction). Data is classified using selected classifiers such as IBk and Random Forest. The next step in the progression is to check for better precision when the precision level is lower. We go to the pre-processing step called Info gain attribute evaluator and ranker for increasing the accuracy performance percentage and the classification steps are performed with the same base classifiers and the level of precision is checked. If the threshold is reached, the model is the optimal model.

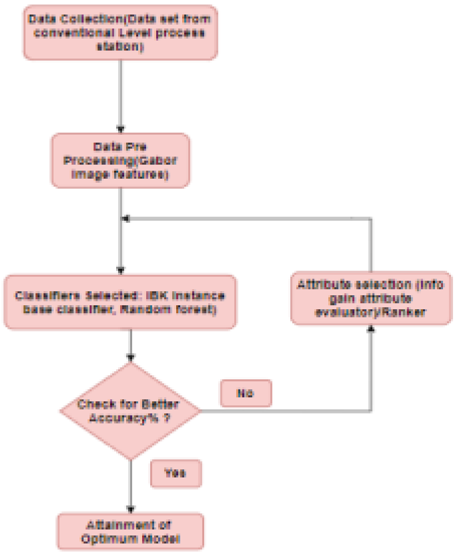

Fig. 4: Flow diagram for level Monitoring.

Figure 4 shows the experimental set up for predicting the Control valve stem position in the following steps:

- Data collection as raw images that must be converted into an accessible format like Atriff or comma delimited value files CSV
- Data pre-processing using Gabor filter and the image to numeric values is processed for further classification and building the model using selected classifiers
- Classification and building the model for training and obtaining the examination performance results
- Observations of the results for better accuracy percentage and area under the curve called weighted receiver operator characteristics (ROC) is better or not, for repeating the above

process by varying the 'K' neighborhood values from 1 to 3 values to improve the overall performance of the built model.

- Attainment of the optimum model to predict the control valve stem position using machine-learning algorithms.

4. Experiment and Results

In this section, we'll go over the recommended experimental setup. Figure 5 depicts the suggested framework methodology. Table 2 shows the results of applying the Java version of the Random Forest decision tree classifier to selected characteristics. With selected properties and training/parameter fitting to neighborhood size, the IBk (KNN) classifier was developed. "k" is shown in Table 3. The experimental statistical analysis of aspect level versus precision and attribute level prediction to the area under the curve called weighted average receiver operator characteristics are shown in Figures 2 and 3, correspondingly. 5.1 Proposed Methodology The data enhancing techniques like data cropping and assigning the ground truth by a skilled person who already did the traditional level predictions. The categorized images are not evenly disseminated among the nominated classes and images with labels as LOW, MEDIUM, and HIGH. Because this implies a highly skewed distribution of images, training this image data set will skew the results and impair classification accuracy. To avoid this bias, data replication techniques were used to boost the count of photos in the remaining classes. [25] This pre-processed data is subjected to a feature extraction method with the help of image filters (Gabor) and then the features have been extracted, the characteristics. Data training activities such as Data pre-processing, Parameter Training, and Attribute Selection were performed on both Random and K-Nearest Neighbor classifiers in the process of obtaining the best model. For various setups, the performance of both classifiers is compared, and a better model based on accuracy is obtained.

4.1 Accuracy Performance Examinations

The Random and IBk classifiers were used to test performance outcomes on feature selection and parameter training/tuning. Reiteration is a typical strategy for enhancing the performance of machine learning models and reducing training time. The information gain Attribute Eval/Ranker in WEKA is used to select the attributes. The IBk classifier modifies the nearby neighborhood to tune the parameters (k). A quantity of various characteristic choice strategies as benefit the information from the associated works choice of method of the statistics is taken into consideration. Additionally, the movement of various forms of a subset of statistics set mode the use of processing and classifying the use of in which is set of rules silly and after modern. The reason is to pick out an excellent tool mastering method and to try and construct a version for every of the statistics. The approach of assessment of consequences is performed earlier than making use of that it's far a top of the line order the statistics set to the satisfactory overall performance furthermore this greater used on maximum issues for mastering algorithms. The characteristic statistics characteristic choice is partitioned into parts: Attribute Evaluator and seek technique having more than one strategy in every direction from which possible pick out. The characteristic evaluator is the technique with the aid of using which the characteristic of a characteristic with inside the statistics set is evaluated like every, the significance of the classes, etc. Techniques and navigation methodologies are combos of capabilities with inside the statistics set with the intention to attain the most accuracy in overall performance or the most fulfilling length of a listing of decided on capabilities. Attribute evaluator and seek technique strategies may be configured via the talk packing containers as special with inside the characteristic evaluator. This technique with the aid of using which every characteristic or statistics characteristic with inside the statistics set is evaluated like the magnificence characteristic of characteristic instance low, high, and medium. The simple looking method and of a distinctive aggregate of capabilities with inside the statistics set are completed for accomplishing the most fulfilling version with the chosen capabilities. The assessment method does now no longer paintings for all varieties of statistics sets, it varies with the request from a particular seek methodology. Implementation of correlation characteristic can most effective be carried out with a particular kind or seek technique this is primarily based totally at the valuation of every characteristic. Mind choosing distinctive characteristic evaluators it's far the requirement to alternate the compatibility of seek strategies like grasping set of rules for the breadth-first set of rules with selected method. Data characteristic choice primarily based totally on correlation is a well-known technique for choosing the maximum applicable attributes with inside the statistics set is figuring out the correlation in statistics. It is a greater widespread connection with a green correlation Coefficient in which we are able to calculate the dependencies with the aid of using the notation of correlation among every characteristic and the magnificence of variables. Based in this calculation, the choice technique follows most effective the ones attributes which have a pleasing common to superb or poor correlation and coffee correlation value. Correlation characteristic assessment method that calls for using ranker seek technique in Weka as shown in figure 5.

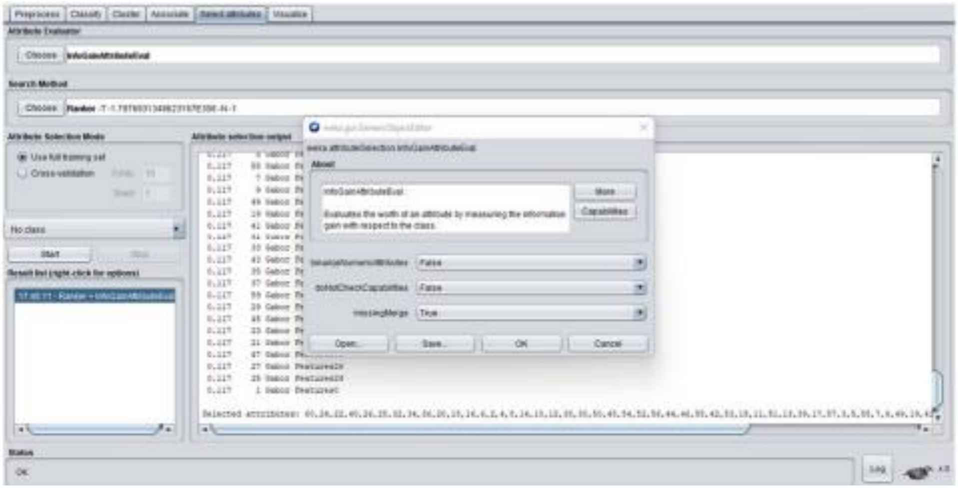

Fig. 5: Option panel for select attribute /Ranker in Weka

In Weka attribute selection info gain attribute evaluator can evaluate the worth of an attribute by measuring the information gain with respect to the class. It has the capability like handling binary class, missing class values, Nominal attributes, Numeric attributes, and Unary attributes. Attribute can be of the form of binary attributes, date attributes, empty nominal attributes and missing values nominal attributes numeric attributes and unary attributes. The selected attribute starts from the range of 60, 24, 22, 40, 26, 28, 32, 34, 36, 20, 18, 16, 6, 2, 4, 8, 14, 10, 12, 38, 30, 50, 48, 54, 52, 56, 44, 46, 58, 42, 53, 15, 11, 51, 13, 39, 17, 57, 3, 5, 55, 7, 9, 49, 19, 41, 31, 33, 43, 35, 37, 59, 29, 45, 23, 21, 47, 27, 25, 1 : 60}. As shown in figure 5. It uses full dataset for training and 10-fold cross validation method without any class specification is done.

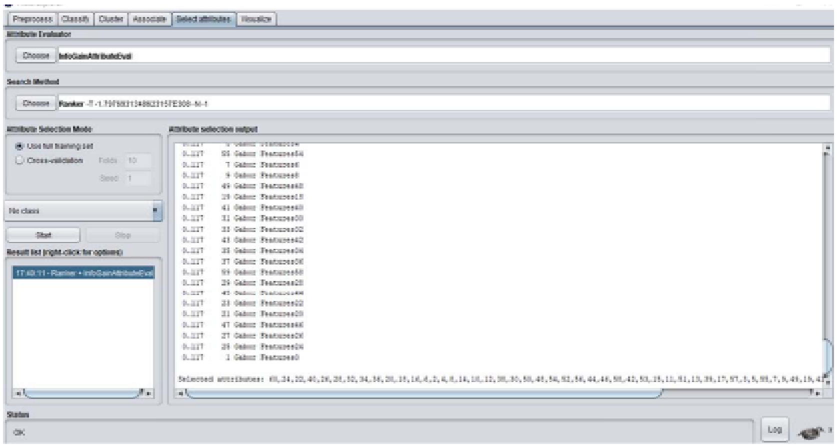

Fig. 6: Option panel for select infogain /Ranker in Weka.

The run information of info gain attribute value it is as follows: The number of instances is 140 and 61 attributes. The ranks of all the features according to the information present in the features after evaluating the entire dataset. Attribute selection on all input data and the search method adopted is an attribute ranking and it is of information gained ranking filter that gives an output rank for all the 60 features. The performance characteristics of the data set is then applied to the classifier tab where two different classifiers random forest and instance based classifiers are been applied to the data set with the retained attributes, number of attributes for various category the accuracy and the weighted-average receiver operator characteristics are tabulated in table 1 and 2. The performance of random forest with feature selection module is done for both retaining all the attributes and removing the attributes in step by step with an account of 10 at a time. The result of the random forest classifier is such a way that accuracy is increased from 80.791% to 81.3559%. The average area under the curve or the weighted-average receiver operator Characteristics is 0.931 and 0.932.

Table 1: Performance of Random Forest with feature selection module.

S.No	**Uninvolved attribute list**	**Attribute selection session (Attribute assessor/Search method)**	**Tree based RF - Random Forest Classifier accuracy (%)**	**Tree based RF - Random Forest Classifier Area under the Curve [ROC]**
1	Retaining all attributes (61 attributes are retained)	**InfoGainAttributeEval / Ranker**	80.791	0.931
2	Gobor Features removed - 16,02,18,20,24,40,14,46,12,48 and retained 51 attributes	**InfoGainAttributeEval / Ranker**	81.3559	0.932
3	Gobor Features removed - 58,38,30,52,32,26,44,50,4,6,22 and retained 40 attributes	**InfoGainAttributeEval / Ranker**	80.791	0.930
4	Gobor Features removed – 10,36,54,8,42,56,28,34,3and retained 31 attributes	**InfoGainAttributeEval / Ranker**	74.011	0.875
5	Gobor Features removed – 51,53,17,13,59,55,57,11,23,31,49 and retained 20 attributes	**InfoGainAttributeEval / Ranker**	75.988	0.895
6	Gobor Features removed – 37,45,27,49,9,7,1 and retained 13 attributes	**InfoGainAttributeEval / Ranker**	75.706	0.907

The performance of instant based classifier IBK @l,2,3 selection module shows the performance characteristics of this classifier whose output accuracy percentage for retaining all the attributes is 75.141% with a weighted average ROC of point 0.848 as shown in table 2.

Performance characteristics of Random Forest, IBK Classifiers versus accuracy percentage for different attribute levels. On observing the results that the Random Forest Depending on the approach used to choose attributes, the accuracy ranges from 74.011 percent to 81.356 percent. When comparing 31 and 51 features, the Random Forest classification algorithm has the greatest accuracy rate of 81.356 percent.

Table. 2: Presentation of IBK @ (K = 1,2,3) with data feature selection segment.

S.No	Uninvolved attribute list	Attribute selection session (Attribute assessor/Search method)	IBk (k@1) Accuracy%	weighted Avg. ROC (k@1)	IBk (k@2) Accuracy%	weighted Avg. ROC (k@2)	IBk (k@3) Accuracy%	weighted Avg. ROC (k@3)
1	Retaining all attributes (61 attributes are retained)	InfoGainAttributeEval / Ranker	69.492	0.777	68.079	0.836	75.141	0.848
2	Gobor Features removed - 16,02,18,20,24,40,14,46,12,48 and retained 51 attributes	InfoGainAttributeEval / Ranker	68.644	0.772	66.949	0.828	75.424	0.851
3	Gobor Features removed - 58,38,30,52,32,26,44,50,4,6,22 and retained 40 attributes	InfoGainAttributeEval / Ranker	66.384	0.758	66.667	0.827	72.881	0.837
4	Gobor Features removed – 10,36,54,8,42,56,28, 34,3and retained 31 attributes	InfoGainAttributeEval / Ranker	62.429	0.720	59.322	0.770	62.994	0.789
5	Gobor Features removed – 51,53,17,13,59,55,57,11,23,31,49 and retained 20 attributes	InfoGainAttributeEval / Ranker	63.842	0.734	62.7119	0.799	65.538	0.810
6	Gobor Features removed – 37,45,27,49,9,7,1 and retained 13 attributes	InfoGainAttributeEval / Ranker	66.667	0.766	63.277	0.871	64.972	0.811

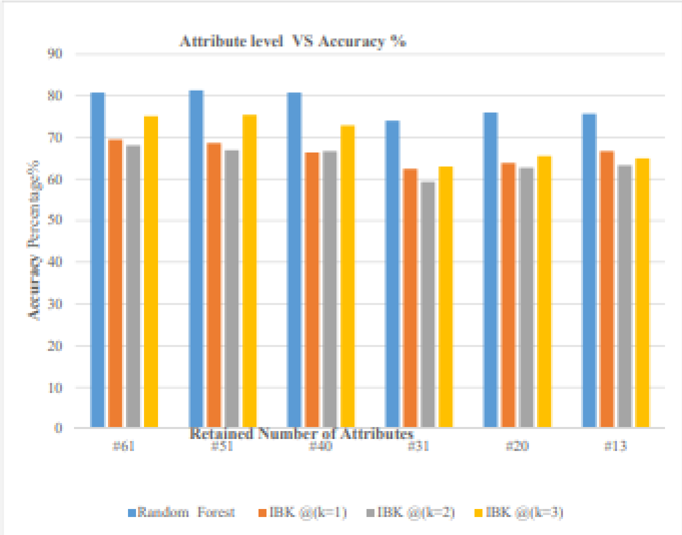

Fig. 7: Presentation of RF - Random Forest & IBK For k=1,2,3 Selected attribute level vs accuracy.

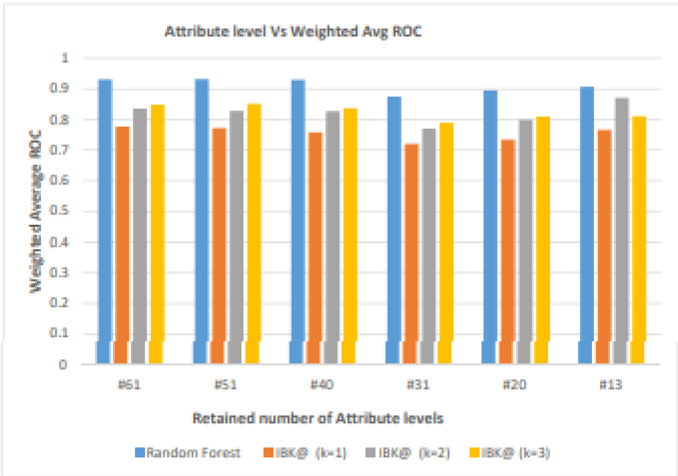

Fig. 8: Performance of Random Forest & IBK (K = 1,2,3) attribute level vs Weighted Average ROC.

The comparison of the Random & IBK algorithms with multi-k values is depicted in the graph above. Accuracy from a neighborhood of size k = 1,2.3 is represented by bars. With a k = 3 value, IBK achieves 75.424 percent accuracy. Based on the feature selection approach, we find that the Weight Avg. ROC achieved ranges from 0.851 to 0.720. Table 2 shows that when trained with 61, 51, and 40 selected features, the Random Forest classification algorithm yielded 0.932 high weighted average receiver operator characteristics (ROC) values. With varying k values, the IBK method is iterated. In the WEKA tool, the yellow line represents the standard k value, the blue line represents the nearest neighbour k = 2, and the red line represents the standard k value.

5. Deduction

The proposed system is made up of two primary modules with filters and classifications: attribute selection among filtered features in the first component, and iterations with instance-based closest neighbor models in the second. In the first component, the ideal accuracy with maximum performance was discovered to be 74.59 percent, and in the second component, it was 81.99 percent. These findings are the first of their type in this framework for level prediction utilizing digital image processing that is based on the Gabor image filter.

Table. 4: Conclusion Summary of research article inferences.

S.No	Interpretation of results	Performance Evaluating Criteria	Alterations in the examinations	Final output result
1	Feature Selection for Gabor filter based on level measurement using Non - interacting Tanks Level images	Accuracy: 81.356%, and Weighted average ROC: 0.931	Selected attributes with information gain and ranker, Increased K value in Instance base Classifier	Diverse approaches and obtained results in a structure arising new emergence of measurement.

5.1 Future Possibilities

[1] The Gabor image filter allows us to identify the image set of level scenarios in a more efficient manner. These filters and expanded classifiers should be investigated further and evaluated for improved performance and accuracy

[2] Trying other image-based parameter measurements and monitoring can be implemented.

[3] Alterations are done on many more selections of ensembles other than Random Forest adopted in this study and framework.

References

[1] Tri Chandra S.Wibowo, Nordin Saad, and Mohd Noh Karsiti, System Identification of an Interacting Series Process for Real-Time Model Predictive Control, American Control Conference, pp.4384- 4389, June2009

[2] E. P. Gatzke, E. S. Meadows, C. Wang, and F. J.Doyle, “Model based control of a four-tank system", Computer and Chemical Engineering, Vol. 24, pp 1503-1509, 2000.

[3] S. Nithya1, N.Sivakumaran, T .Balasubramanian and N.Anantharaman, “Model based controller design for spherical process in real time”, IJSSST, Vol.9, No.4, pp.25-31, November 2008.

[4] E. Weyer, "System identification of an open water channel," Control Engineering Practice, Vol. 9, pp.1289-1299, 2001.

[5] W. Favoreel, B. D. Moor, and P. V. Overschee, “Subspace state space system identification for industrial processes", Journal of Process Control, Vol. 10, pp. 149-155, 2000.

[6] Il-Hwan Kim, Stanley Fok, Kingsley Fregene, Dong-Hoon Lee,Tae-Seok Oh and David W. L.Wang “Neural Network-Based System Identification and Controller Synthesis for an Industrial Sewing Machine”, International Journal of Control, Automation, and Systems, Vol. 2, No. 1, March 2004.

[7] Pardeep Rohilla, Vineet Kumar. Fuzzy I+PD controller for stiction compensation in pneumatic control valve. International Journal of Applied Engineering Research 2017; Volume 12, no.13, pp. 3566-3575, ISSN 0973-4562.

[8] Srinivasan Arumugam and R.C. Panda. Identification of stiction nonlinearity for pneumatic control valve using and method. International Journal of Engineering & Technology. 2014; vol. 6, no. 2, pp. 570-578.

[9] Fei Li, You Jun Fan, Hus Tian Zhao. The Research of a Piezoelectric Valve Positioner Control Algorithm[J]. Advanced Materials Research, 2014; 3470

[10] Wang J, Wang D J D, Moore P R. Modelling Study, Analysis and Robust Servo Control of Pneumatic Cylinder Actuator System[J]. Control Theory and Applications, 2014; 148(1): 35-41.

[11] Choi, Changhyun, Jungwook Kim, Heechan Han, Daegun Han, and Hung S. Kim. Development of Water Level Prediction Models Using Machine Learning in Wetlands: A Case

Study of Upo Wetland in South Korea. 2020; Water 12, no. 1: 93. https://doi.org/10.3390/w12010093

[12] F. N. M. Ariff, A. S. A. Nasir, H. Jaafar and A. N. Zulkifli, "Character Segmentation for Automatic Vehicle License Plate Recognition Based on Fast K-Means Clustering," 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET), 2020, pp. 228-233, doi: 10.1109/ICSET51301.2020.9265387.

[13] V. V. Mainkar, J. A. Katkar, A. B. Upade and P. R. Pednekar, "Handwritten Character Recognition to Obtain Editable Text," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 599-602, DOI: 10.1109/ICESC48915.2020.9155786.

[14] Y. Peng and Z. Chen, "Application of Deep Residual Neural Network to Water Meter Reading Recognition," 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), 2020, pp. 774-777, doi: 10.1109/ICAICA50127.2020.9182460

[15] N.Saravanan, V Gayathiri. Performance and Classification Evaluation of J48 Algorithm and Kendall's Based J48 Algorithm (KNJ48). International Journal of Computational Intelligence and Informatics. 2018; Vol. 7: No. 4.

[16] P. Wijerathna and L. Ranathunga. Rice Category Identification using Heuristic Feature Guided Machine Vision Approach. 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), 2018; pp. 185-190, doi: 10.1109/ICIINFS.2018.8721396.

[17] O.D. Trier, A.K. Jain. Goal-directed evaluation of binarization methods. IEEE Trans. 1995; PAMI 16

[18] Mohammed M Mazid, A. B. Improved C4.5 Algorithm for Rule-Based Classification. Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases. 2013.

[19] Nagaparameshwara chary, S. D. A Survey on Comparative Analysis of Decision Tree Algorithms in Data Mining. International Conference on Innovative Applications in Engineering and Information Technology (ICIAEIT-2017.

[20] Dr. Anjali B Raut, A. A. Students Performance Prediction Using Decision Tree Technique. International Journal of Computational Intelligence Research, 2017; 1735-1741. (12) 1191–1201.

[21] W. Niblack. An Introduction to Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ, 1986.

[22] Gonzalez, Digital Image Processing, Electronic Industry Press.2005.

[23] J.R. Parker. Image Processing and Computer Vision Algorithms and Applications, Tsinghua University Press.2006.

[24] Paaranan Sivasothy, Jens Wagener and Jörg Seewig. Machine learning based filling level estimation for bulk solid silos Institute for Measurement and Sensor-Technology, Technische Universität Kaiserslautern, Kaiserslautern, Rheinlandpfalz, GERMANY, 2018; Vol. 35, 055002

[25] Vijayan T, Sangeetha M , A. Kumaravel, Karthik B , Gabor Filter and Machine learning Based Diabetic Retinopathy Analysis and Detection, Microprocessors and Microsystem, 2020; doi: https://doi.org/10.1016/j.micpro.2020.103353