Classification of data with optimized fuzzy rough set feature selection on Nash equilibrium based kernel support vector machine

J. Uma Maheshwari¹ and R. Mithra²*

Abstract
In data mining, the reduction of dimension is a highly complicated process, and it should be processed in an efficient way to get an optimal solution. So in this paper, the fuzzy rough set theory is implemented for optimal feature selection in the normalized data of KDD cup 99 datasets. The Nash equilibrium game theory is fed as an input to update the kernel on SVM. The Nash equilibrium gives the rapid update and provides an accurate rate of classification in data reduction. The performance values computed for this method is measured in terms of accuracy, precision, detection rate, F-score, FPR and AUC.

Keywords
Fuzzy rough set theory, Data mining, Nash equilibrium, Kernel-based SVM, KDD cup 99 dataset.

AMS Subject Classification
03E72.

Contents
1 Introduction .................................................... 299
2 Review on Related Works .................................. 300
3 Proposed Methodology ..................................... 302
  3.1 Dataset description ................................. 302
  3.2 Data Normalization ................................. 302
  3.3 Discretion of attributes ............................ 302
  3.4 Fuzzy rough set theory for feature selection .......... 302
  3.5 FRS based optimum subset selection .................. 302
  3.6 Parameter optimization .............................. 303
  3.7 SVM Kernel Updation based on Nash Equilibrium Game Theory .................................. 303
4 Computation of Nash Equilibrium Game theory .......... 303
5 Results and Discussion .................................... 304
6 Summary ...................................................... 305
References ..................................................... 305

1. Introduction
Data mining is referred to retrieve the hidden insights of knowledge base to make a reliable decision by using few ideas like rough set theory based on logic implementations of Boolean function which can be defined as 0 or 1 [1]. It can manage uncertain and vague information. The trending approach used in mathematics is rough set theory is to interpret the undefined knowledge that is tackled by logicians, mathematicians and philosophers. So the proposed fuzzy rough set theory is used for feature selection to select only the particular attributes from a higher-dimensional data. [2] Nash equilibrium is described as an action summary such that no player can attain better outcomes by performing a varied action, providing that other players should follow the action summary. It is defined as the interactions of the best stimulus of all the players’ intricate in the team.

Reduction of data is a significant role in machine learning technologies and sequence recognition with maximum dimensional data. In real-world applications, it is in a hybrid matter, and unique data reducing method should be designed to hold the actual data without any loss. An information metrics is implemented to compute the prominence value of the fuzzy method. We refine the data with the hybrid attribute to ensure the quality outcomes to feed into an application. [3].
2. Review on Related Works

The selection of features is a significant part of getting the best attributes for the input dataset [4]. It explores the feature selection issues by using support vector machines by studying the selection feature subset by clustering methods based on Expectation maximization methods which utilize separability to compute the feature subset. Some ways examined the feature selection as text to get learning efficient with the view of fuzzy-rough relevant features and holds the value of real datasets. It manages the statistical approach to analyze the voluminous discovery of data. The rough set is dependent upon the shunned dataset because the critical data is mislaid as the output of the dissertation. So proper examination on the designed model should be made before deploying it into the live environment.

The learning ability of kernel function in SVM is highly astonishing, which the key to utilizing feature space of higher dimensional data effectively [5]. The kernel approach has gradually improvised the performance of the support vector machine which developed many methodologies in machine learning such as classification of patterns and regression estimation. The SVM is based on single kernel methods and multi-kernel methods which is applied to solve many complex problems which are not rectified for years. Till date, there is no robust application to select and structure the kernel functions. The relationship between several kernels should be established, and the issues in the parameter’s rationality are explored in multi-kernel techniques. The structure of the kernel is still a complicated process to analyze effectively. There are many problems encountered at the time of practical applications.

The theory outlined in [6] contains a reasonable approach to the processing of valid data rules. This paper examines rough data sets and fuzzy rough data sets with applications in data mining that can handle uncertain and ambiguous data to reach useful conclusions. Data mining is a Information finding that refers to the retrieval of valuable data or secret structures from the knowledge base to make decisions. Rough set approach is a trending mathematical way to unsure facts. Philosophers, mathematicians and philosophers have tackled the issue of incomplete or unclear information. Rough set model is a unique mathematical method to tentative learning. Scholars, mathematicians and logicians have addressed the issue of incomplete or ambiguous data. Analysis of a cluster is the job of combination a collection of objects in such a manner those artefacts in the same cluster are more identical than those in other groups. Clusters analysis is an essential purpose of data mining. A combination scheme that incorporates the benefits of fuzzy sets and hard sets is used to define the objects in the particular classes. The main objective of the fuzzy-rough games is to describe the minor and greater calculation of the set when the fuzzy continuum sets are raw due to equivalence or to transform the equivalence relationship into a related fuzzy relationship. Comparative discernibility relationships are used to classify negligible features in the perceptibility matrix. An algorithm to calculate the smallest factors is then developed. Fuzzy rough sets have different applications in the data mining field that are applied to manage the uncertainty and imprecision of the data. They are not capable of handling indefinite relationships that occur in the data.

The study [7] has a fuzzy rough set theory and a variety of applications, but sometimes it offers ambiguous solutions, and thus the fuzzy approach is introduced to overcome this. The essential principles of graphics engineering are used in many real-time implementations such as routing, information security, map painting and the global mobile service (GSM) [3], mobile phone networks, and so on. The graphic colouring is based on the adaptive Welsh Powell colouring algorithm. In this article, image optimization is done for Indian sign languages (ISL) images, and here the model is considered to be a fuzzy graph defined on a set of pixels where blurred edges embody the distance between the pixels. The purpose of colouring algorithm suggested in this article was to give the decision-maker many alternative representations of the scene. Each of which was obtained using an automated colouring process for each pixel centred on a distance. Such a colouring technique takes into account the actions of each pixel about its neighbouring pixels and recommends a suitable class for each colour. The colouring algorithm known as the Welsh-Powell adaptive method was built on the basis of a simple algorithm. It is converted into a fuzzy sense, which reflects confusion about the dissimilarity between pixels using fuzzy numbers. Such developments make it possible to take a more flexible approach to a complex problem. By using this new algorithm in fuzzy rough set theory, cumulative time can be reduced. Thus, reliable strategy develops a rate of success of about 0.9150, which is said to be much better than the previous strategy and similarity it is evident that the complexity of the process is less and the time consumed is 78s, which is beautifully better than the existing methods. By using this technology, ISL benefits from the fact that the true and transparent hand-cut icon is acquired very effectively. This algorithm also gives a more effective outcome than the usual Welsh-Powell method.

3. Proposed Methodology

The proposed method based on the Nash equilibrium game

![Figure 1. Overview of data mining](image-url)
### Table 1. Comparison of existing Methods [8]

<table>
<thead>
<tr>
<th>Application model</th>
<th>Problem identification</th>
<th>Solution</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPRADM Using Game Theory</td>
<td>Prevention of Collusion in Privacy Preservation</td>
<td>Nash Equilibrium</td>
<td>Developing the game-theoretic solutions with some specific PPDP approaches</td>
</tr>
<tr>
<td>Sequential Game Model</td>
<td>Trade-off resolution between privacy preservation and data utility</td>
<td>Subgame Perfect Nash Equilibrium</td>
<td>Developing the game-theoretic solutions for some specific PPDP approaches</td>
</tr>
<tr>
<td>Sequential Game &amp; Static Game</td>
<td>Agreement between parties concerning the trade of incentives and private information.</td>
<td>Nash Equilibrium</td>
<td>Due to this reason the most of practical solutions are obtained using game theory</td>
</tr>
<tr>
<td>Cooperative Game</td>
<td>Data Publishing of Privacy Preservation in using Cooperative Privacy concept. Game Model on Honeypot DDoS</td>
<td>Nash Equilibrium</td>
<td>Lack of theoretical models for more than three players and the stochastic game model requires more realistic.</td>
</tr>
<tr>
<td>Bayesian Honeypot Game Model</td>
<td></td>
<td>Bayesian-Nash Equilibrium</td>
<td></td>
</tr>
<tr>
<td>Stackelberg Model</td>
<td>Network Hardening Problem.</td>
<td>Stackelberg equilibrium</td>
<td></td>
</tr>
<tr>
<td>Static games of data</td>
<td>Risk assessment of a network, Information warfare</td>
<td>Nash Equilibrium, Mixed Strategy Nash Equilibrium</td>
<td></td>
</tr>
<tr>
<td>Dynamic Games of perfect and entire data</td>
<td>Stackelberg network intrusion detection game</td>
<td>Stackelberg Equilibrium</td>
<td></td>
</tr>
<tr>
<td>Dynamic Games of perfect and complete information</td>
<td>Stackelberg network intrusion detection game</td>
<td>Stackelberg equilibrium</td>
<td></td>
</tr>
<tr>
<td>Dynamic Games of imperfect information and incomplete information. The zero-sum static game, N + M player non-zero sum stochastic game, Zero-sum stochastic game, Nonzero-sum game with dynamic data Incomplete information nonzero-sum stochastic game, Repeated game.</td>
<td>Two-player Multi-stage Bayesian game IDS configuration Optimization</td>
<td>Perfect Bayesian equilibrium</td>
<td>Improved techniques are required for game-theoretic integration models into real-world scenarios.</td>
</tr>
<tr>
<td>Zero-sum static game, Zero-sum stochastic game, Nonzero-sum game, N + M player non-zero sum stochastic game,</td>
<td>Resource allocation optimization</td>
<td>Nash equilibrium, Bayesian nash equilibrium</td>
<td></td>
</tr>
<tr>
<td>Stochastic game, MDP, Incomplete data zero-sum sequential game, Nonzero-sum Incomplete information nonzero-sum sequential game,</td>
<td>Issues of IDS optimization and countermeasure optimization</td>
<td>Nash equilibrium</td>
<td></td>
</tr>
</tbody>
</table>
theory to update the SVM kernel function to get maximum accuracy in the reduction of dimensional data. After the process, of data normalization and discretion of attributes, the fuzzy rough set theory is used to select only the specific qualities which are more suitable for the data reduction process. The overview of the proposed method is shown in Fig. 2.

3.1 Dataset description
The input dataset taken in the proposed way is KDD Cup 99 dataset can be applied in all sort of domains for multiple purposes.

3.2 Data Normalization
This method is used to eliminate the repeated data from the database and save the non-repeated and constant information into it. It is used to merge the multiple data tables into the whole, which can be queried rapidly. It is a primary process to structure the relational data concerning to the nominal data to minimize the data redundancy and maximize the data integrity. The entities are organized as relations and attributes based on the dependencies and constraints. It is attained by providing a few formal rules by synthesis.

3.3 Discretion of attributes
Discretion of attributes is a significant method for compressing the information and simplification, which is indeterminable in recognition of pattern and a view on a rough set domain. The tool of discretization depends on diving the cut point. Initially, there are five different views of the discretization algorithm. Such as local or global, dynamic vs static, unsupervised vs supervised, incremental or direct. The continuous attributes which are to be discretized in many algorithms, such as tag sort or it can be rule statement particularly rough set theory in the field of data mining. The decision table which should be followed is not discriminate at time of discretion.

3.4 Fuzzy rough set theory for feature selection
The feature selection method is to choose the optimal features from higher dimensional dataset to limit the computational time [9]. The vast number of attributes existing in real-life data sets and only a section of code is effectively applied to mark the accurate dataset. The proposed algorithm reduces the count of unwanted attributes by selecting only the essential qualities. The significant is based on dependent value. Higher the significance value and maximum dependency are auxiliaries to the crucial qualities. The steps are repeated until the characteristics are minimized to significant values.

3.5 FRS based optimum subset selection

\[ FIS = \{ (V, B, U, e) \mid V \text{ represent dataset which includes instances for entire network connection} \} \]

\[ B \text{ represents attribute set consist of a finite number of attributes} \]

\[ \text{The domain set is } U = \bigcup_{b \in B} U_b \text{ where } U_b \text{ represents a set of attributes } b, \text{ also known as the domain of } b, e: V \times B \rightarrow U \text{ represents information function which assigns particular values to objects from attribute domain.} \]

\[ H \text{ represents a conditional attribute} \]

\[ T \text{ represent decision attribute} \]

Four tuple information system is expressed as

\[ FIS = \{ V, H \cup Q, U, e \} \quad (3.1) \]

Fuzzy equivalence matrix is expressed as \( N \in 4^m \times m \).

Fuzzy similarity matrix \( S \in 4^m \times m \).

\[ s_{ji} = \max \left( \min \left( \frac{y_j - (y_j - \sigma_b)}{y_j - (y_j - \sigma_b) + (y_j + \sigma_b - y_i)}, 0 \right), \right) \quad (3.2) \]

or

\[ s_{ji} = \begin{cases} 
1 - 4 \times \frac{|y_j - y_i|}{|b_{\text{max}} - b_{\text{min}}|}, & |y_j - y_i| \leq 0.25 \\
0, & \text{otherwise} 
\end{cases} \quad (3.3) \]

where \( y_j, y_i \) are feature values of two objects on feature \( b \) with maximal \( b_{\text{max}} \) and minimal \( b_{\text{min}} \) of all objects, \( \sigma_b \) standard deviation value of an instance of attribute \( b \).

Fuzzy relation matrix \( S \) is \( t(S) = V_{m=1}^m S^m \) which is incalculable. In real application \( N = t(S) \) could be streamlined following theorem.

**Theorem 3.1.** Fuzzy similarity matrix \( R \) has a transitive closure which is a fuzzy equivalent matrix as \( N = t(S) \) where \( t(S) - S^m \).

**Proof.** We need to satisfy that fuzzy similarity matrix \( S \) with transitive closure, \( t(S) = S^m \) and satisfies that \( t(S) \) has the property of symmetry and reflexivity.
Reflexivity: As fuzzy similarity matrix $S$ is reflexive, namely, $J_{m \times m} \subseteq S$, and $S \subseteq S^2 \subseteq \cdots \subseteq S^m$ then we attain $t(S) = V_{m=1}^m S^m = S^m \supseteq J_{m \times m}$ thus $t(S) = S^m$ and $S^m$ is reflexive.

Symmetry: Since fuzzy similarity matrix $S$ is symmetric, namely, $S^T = S$, then $(S^T)^m = S^m = (S^m)^T$.

To summarize, since $t(S) = S^m$ is reflexive, symmetric and transitive $t(S)$ is a fuzzy equivalent matrix.

Description 1 (Fuzzy partition): Input Fuzzy information system, is $FIS = \{ V, H \cup Q, U, e \}$, the fuzzy partition of a universe set $V$ based on FER $S$ is defined as

$$V/S = \{ [y_j]_S \}_{y_j=1}^m \text{ r.t. } [y_j]_S = \frac{s_{1j}}{y_1} + \frac{s_{2j}}{y_2} + \frac{s_{mj}}{y_m} \tag{3.4}$$

$[y_j]_S$ fuzzy equivalence class generated by $y_j$ and fuzzy equivalence relation $S$, due to FER $S$, $V/S$ is fuzzy partition and $[y_j]_S$ is a fuzzy set.

Description 2: Information quantity based on FER $R$ is define as:

$$C(S) = -\frac{1}{m} \sum_{j=1}^m \log \frac{|[y_j]_S|}{m} \tag{3.5}$$

where $|[y_j]_S| = \sum_{j=1}^m s_{ij}$ is the cardinality of $[y_j]_S$.

Description 3: It is based on Condition Entropy. In Fuzzy decision system, $FIS = \{ V, H \cup Q, U, e \}$. Assume $A$ is a subset of $H$, the conditional entropy of decision attribute $Q$ conditioned on $A$ is defined as:

$$C(Q | A) = -\frac{1}{m} \sum_{j=1}^m \log \frac{|[y_j]_A \cap [y_j]_Q|}{|[y_j]_A|} \tag{3.6}$$

Description 4: It is Mutual Information. Given Fuzzy information system $FIS = \{ V, H \cup Q, U, e \}$.

And assume that $A$ is subset of $H$, the mutual information of $A$ and $Q$ is

$$J(A; Q) = C(Q) - C(Q | A) \tag{3.7}$$

Description 5: It is based on Information Gain. Input, Fuzzy decision system $FIS = \{ V, H \cup Q, U, e \}$. Let $A$ is current chosen subset of $H$ and $b$ is new independent attribute excluded in $A$, namely, $b \in H - A$, $JG$ of attribute $b$, $Gain(b, A, Q)$, tentatively incorporating it into chosen attribute subset, can defined as

$$Gain(b, A, Q) = J(A \cup \{b\}; Q) - J(A; Q) \tag{3.8}$$

Description 6: It is a Information gain ratio the IGR of attribute $b$, $Gain\_ratio(b, A, Q)$ can defined as

$$Gain\_ratio(b, A, Q) = \frac{Gain(b, A, Q)}{C(b)} = \frac{J(A \cup \{b\}; Q) - J(A; Q)}{C(b)} \tag{3.9}$$

### 3.6 Parameter optimization

The output of the feature selection is further optimized by selecting only its significant values. That selected values are fed as input to a classifier, which is SVM based on the Nash equilibrium game theory. Now the attained values are without noises, mislaid values, fake values, repeated values, dummy values and only the optimized values are fed as input to the classifier.

### 3.7 SVM Kernel Updation based on Nash Equilibrium Game Theory

The Nash equilibrium game theory is introduced to update the kernel function of SVM. The SVM is defined by a kernel function and entirely on the training of dataset. It is implemented with kernel function, which is an inner component and complete symmetric tasks which should have the rapid computation capacity in the reduction of dimensions of optimized data. The new kernel functions are selected to maximize the ability of classification and dimensional reduction of feature space. Evaluation of the kernel function in input tactically reacts to the calibration in the high dimensional region. Hence to determine the nature of feature region kernel function should be updated periodically for effective performance in the classifier.

### 4. Computation of Nash Equilibrium Game theory

In game $E = \{ m, R^*, V_m \}$, $R^*$ is set of strategies available to user $j$ ($j = 1, 2, \ldots, m$). Player $j$ randomly chose strategy with probability distribution $Q_j = \{ q_{j1}, q_{j2}, \ldots, q_{jl} \}$ in $l$ optimal strategy. This is referred to mixed strategy, $0 \leq 1 \ (i = 1, 2, \ldots, l)$ and $q_{j1} + q_{j2} + \cdots + q_{jl} = 1$.

$$Q_{sae} = \left\{ \begin{array}{ll} q, & \text{change} \\ 1 - q, & \text{no change} \end{array} \right. \tag{4.1}$$

and

$$Q_{slow} = \left\{ \begin{array}{ll} \rho, & \text{change} \\ 1 - \rho, & \text{no change} \end{array} \right. \tag{4.2}$$

The possibility of kernel function altering for a sample is $q$ by using $L_{sae}$ and possibility will be represented as $p$ for $L_{slow}$.

For $L_{sae}$ player, the benefit gained by choosing the “change” strategy is $U_{sae}(1, p) = \Delta \rho p + (1 + \Delta \rho) \times (1, q)$; and the benefit received by choosing the “no change” strategy is $U_{sae}(0, p) = 0 + 2(1\rho)$. Then we use the principle of solving the Nash equilibrium of the mixed strategy to calculate the $p$-value and set $U_{sae}(1, p) = U_{sae}(0, p)$. Then, it can be got $p = 1 - \Delta \rho$. The situation is symmetric when we consider issues from $L_{slow}$ player’s point of view, and evaluate the payoffs from a play of probability $q$ by $L_{sae}$.

It will have $q = 1 \Delta U$. According to (3.4), we get the final form of $q$ and $p$, shown as

$$Q = \frac{1}{c}, \quad p = \sigma \tag{4.2}$$
The intersection point of red line and blue line is the Nash equilibrium, which is more intuitive description of our game model.

Thus we determine the value of the two players, and the ensemble kernel function based on the game theory is shown as

\[
L(y,z) = \frac{1}{e} L_{sa} + \sigma L_{qloy}, \quad \frac{1}{e} + \sigma = 1
\]

(4.3)

We name (4.3) as the game kernel function, symbolically \( K_{GT} \).

In this section, we use the Nash equilibrium kernel function in feature space transformation, and we introduce stratified random sampling based on grid search for cross validation to improve the quality of learning samples. The algorithm NK-SVM is proposed.

According to the Mercer condition, assume that \( L_1 \) and \( L_2 \) are kernels of \( Y \times Y, Y \subseteq S^n \). Constant \( \alpha > 0 \), then the following functions \([7]\) are kernels:

\[
L(y,z) = L_1(y,z) + L_2(y,z)
\]

(4.4)

\[
L(y,z) = \alpha L_1(y,z)
\]

(4.5)

Therefore, our game kernel function (4.3) satisfies the Mercer condition. At the same time, it does not change the original mapping space.

### 5. Results and Discussion

The performance metrics of the proposed method FST-NKSVM is measured in terms of accuracy, detection rate, precision, F-score, FPR and AUC. It is compared with different existing approaches and proved that the proposed method outperforms all the current methods.

- Accuracy is an essential attribute which measures the efficacy of the proposed method.
- Detection rate defines the correct identification of the significant qualities evaluating from true positive.
- Precision is a measure of positive predicted values, and it’s a part of the relevant instance from the retrieved cases.
- FPR is the false positive rate which should be minimum in the detection of valid values as false values.

#### Table 2. Comparison of performance metrics for a different method and proposed method [10]

<table>
<thead>
<tr>
<th>Methods</th>
<th>Accuracy</th>
<th>Detection Rate</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA-MLP</td>
<td>92.52</td>
<td>89.23</td>
<td>88.23</td>
</tr>
<tr>
<td>PSO-MLP</td>
<td>90.52</td>
<td>89.15</td>
<td>90.29</td>
</tr>
<tr>
<td>TLBO-MLP</td>
<td>85.34</td>
<td>85.34</td>
<td>85.34</td>
</tr>
<tr>
<td>GOA-MLP</td>
<td>93.53</td>
<td>93.62</td>
<td>94.52</td>
</tr>
<tr>
<td>GA-NB</td>
<td>88.63</td>
<td>89.23</td>
<td>88.23</td>
</tr>
<tr>
<td>PSO-NB</td>
<td>88.63</td>
<td>87.13</td>
<td>86.61</td>
</tr>
<tr>
<td>TLBO-NB</td>
<td>86.34</td>
<td>87.21</td>
<td>83.29</td>
</tr>
<tr>
<td>GOA-NB</td>
<td>94.52</td>
<td>94.52</td>
<td>93.62</td>
</tr>
<tr>
<td>GA-SVM</td>
<td>96.31</td>
<td>97.26</td>
<td>96.54</td>
</tr>
<tr>
<td>PSO-SVM</td>
<td>96.15</td>
<td>96.32</td>
<td>97.32</td>
</tr>
<tr>
<td>TLBO-SVM</td>
<td>95.32</td>
<td>94.12</td>
<td>94.93</td>
</tr>
<tr>
<td>GOA-SVM</td>
<td>98.51</td>
<td>97.52</td>
<td>95.993</td>
</tr>
<tr>
<td>FST-NKSVM</td>
<td>99.21</td>
<td>98.63</td>
<td>96.89</td>
</tr>
</tbody>
</table>

#### Table 3. Comparison of FPR, F-measure and AUC [10]

<table>
<thead>
<tr>
<th>Methods</th>
<th>FPR</th>
<th>F-measure</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA-MLP</td>
<td>10.03</td>
<td>92.45</td>
<td>91.37</td>
</tr>
<tr>
<td>PSO-MLP</td>
<td>11.41</td>
<td>94.21</td>
<td>93.62</td>
</tr>
<tr>
<td>TLBO-MLP</td>
<td>12.33</td>
<td>93.52</td>
<td>92.73</td>
</tr>
<tr>
<td>GOA-MLP</td>
<td>9.05</td>
<td>91.53</td>
<td>93.6</td>
</tr>
<tr>
<td>GA-NB</td>
<td>13.4</td>
<td>73.52</td>
<td>88.9</td>
</tr>
<tr>
<td>PSO-NB</td>
<td>18.54</td>
<td>64.51</td>
<td>89.79</td>
</tr>
<tr>
<td>TLBO-NB</td>
<td>17.65</td>
<td>92.51</td>
<td>88.92</td>
</tr>
<tr>
<td>GOA-NB</td>
<td>12.11</td>
<td>70.43</td>
<td>89.93</td>
</tr>
<tr>
<td>GA-SVM</td>
<td>12.2</td>
<td>98.5</td>
<td>97.43</td>
</tr>
<tr>
<td>PSO-SVM</td>
<td>18.41</td>
<td>98.72</td>
<td>97.42</td>
</tr>
<tr>
<td>TLBO-SVM</td>
<td>15.54</td>
<td>98.31</td>
<td>97.62</td>
</tr>
<tr>
<td>GOA-SVM</td>
<td>15.01</td>
<td>97.6</td>
<td>97.53</td>
</tr>
<tr>
<td>FST-NKSVM</td>
<td>9.21</td>
<td>98.86</td>
<td>98.45</td>
</tr>
</tbody>
</table>

- F-measure is termed as the sensitivity of the model, which only focus on only to detect the positive benefits.
- AUC is the area under the curve which is placed for research and is the difference between the region of the curve and below the part of the curve.

#### Figure 3. Performance metrics of different methods

From Table 2, it is displayed that the proposed method gets the value of accuracy as 99.21, precision as 96.89, and detection rate as 98.63. It is compared with a different existing...
Classification of data with optimized fuzzy rough set feature selection on Nash equilibrium based kernel support vector machine — 305/305

Figure 4. Comparison of FPR, F-measure and AUC

process such as multi-layer perception, NB and traditional SVM along with feature selection techniques such as GA, TLBO, PSO and GAO. The proposed method is based on fuzzy rough set theory for feature selection and SVM kernel update by Nash equilibrium function.

From Fig. 3, it is shown that the proposed method gives high values in F-measure and AUC in the reduction of dimension when compared to existing methods.

From Table 3, it is shown that the proposed method, FST-NKSVM, gives minimum detection in false-positive ratio as 9.21, F-measure as 98.86 and AUC as 98.45.

From Fig. 4, it is shown that the proposed method outperforms the existing process in terms of FPR, F-measure and AUC.

6. Summary

Game theory is the procedure of modelling the strategically interface between dual or many players in an environment that comprises an array of rules and results. Here the Nash equilibrium game theory which is fed as an input to update the kernel function in the reduction of dimensions of a higher dimensional data. The performance metrics computed for the proposed method outperforms all the existing process in terms of accuracy, precision, detection rate, AUC, FPR, F-score. Mathematically in data mining by selecting fuzzy rough set theory for optimal feature selection by introducing optimal variables. It is fed into Nash equilibrium based kernel SVM which effectively helps in the classification of data and training the dataset in the reduction of dimensions which is implemented in data mining.

References