Semiconductor Optoelectronics
ISSN: 1001-5868

# IMPLEMENTATION AND PERFORMANCE ENHANCEMENTS OF OPTIMISED KERNEL MSVM MODEL FOR EARLY CHURN PREDICTION IN TELECOM SECTOR

**Pankaj Hooda[1] & Pooja Mittal[2]**
[1]Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, India
[2]Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, India
* Corresponding Author: e-mail: hooda.pankaj@gmail.com

**Abstract:** To reduce the challenges of customer churn analysis, the proposed OKMSVM model has been applied in the Tele-communication sector on "WA_Fn-User-Tele-Customer-Churn" database with over 7042 customers to demonstrate the proposed model's generalization ability. In this model, KPCA algorithm has been used to extract the reliable feature vectors. After this extraction process, an instance selection using ALO optimization is performed. The proposed OKMSVM model when implemented has predicted the test set categories an accuracy value as 91.05%, AUC score of 85.76% and a RMSE score of 2.838 . It is then validated by comparing it with the existing Hybrid two-level SVM (HTL-SVM) model
**Key words:** Customer Churn Analysis, Telecommunication, Optimized Kernel Multiclass Support Vector Machine, Feature extraction using KPCA, Hybrid Two Level SVM model

## 1. INTRODUCTION

Customer churn analysis can be described as the practical tasks carried out over the probability of abandoning products or services. At its most basic, it indicates that customers are left with no choice but to choose another company due to strong competition. Its purpose is to discover customer churn analysis, monitor the circumstances behind customers' abandonment of products or services, and carry out certain preventive activities [1]. Currently, marketplaces are well-defined by demand surpluses, allowing the client to be labelled the true dictator of the marketplace due to technical advancements, changes in marketing tactics, and the formation of a competitive marketplace. As a result, commercial companies need to move their emphasis from consumer goods and regulate their behaviours to prevent customers from switching and maximise earnings with their businesses [2]. CRM (Customer Relationship Management) strictly manages customer details and appropriately manages all consumers to increase their satisfaction. Its primary purpose is to increase consumer pleasure and satisfaction to avoid consumers from churn, which is the most serious issue facing all businesses because even modest differences in consumer attrition can result in major improvements in the company's stock price and profitability [3].

One of the most fundamental concerns for commercial enterprises, primarily shopping malls, is client retention. But on the other side, recruiting new clients comes at a high price, perhaps five times the price of sustaining current ones [4].On the other hand, many businesses experience a loss of valuable clients to the opposition. Customers churning are the term for this. Customer churn is a commercial term that refers to when a client gets engaged in another service or brand, resulting in a decrease in sales and a rise in the prices of acquiring new consumers. Different research models have been conducted to forecast customer turnover and its effects, demonstrating the importance of this topic. According to business experts' surveys, many businesses lose roughly 25% of their clients yearly [5].

In some cases, customer attrition figures have been as high as 36%. A 5% drop in unsatisfied customers, on either hand, will result in a 25% to 85% gain in sales [6]. These figures demonstrate the importance of "customer churn control" in maintaining the company's existence. As a result, client turnover and handling are seen as serious problems in various businesses. Customer churn is a common way to track how many clients are lost. Telecommunications firms frequently lose valued consumers and, as a result, revenue to competitors [7].

The telecommunications business has seen significant transformations in recent years, including the growth of different facilities, scientific breakthroughs, and more competitors due to deregulation [8]. Therefore, customer churn forecasting in the telecommunication sector has become critical to market players who need to shield their customer loyalty, expand their businesses, and enhance customer relationships.

One of the current challenging difficulties in the telecommunications sector is maintaining consumers with a high churn risk [9]. Customers usually choose churn alternatives due to many telecom operators and competitive pressures [10]. As a result, telecommunications companies realized the value of customer retention rather than obtaining new ones.

A variety of reasons influence customer churn. Prepaid clients, unlike post-paid subscribers, are also not constrained by contractual arrangements; therefore, they frequently churn for the most insignificant factors. As a result, predicting their customer churn is challenging [11] [12]. The additional aspect is client devotion, influenced by the phone companies' customer product and service excellence.

Users may be influenced to switch to a competitor with a better range and higher transmission reliability due to concerns such as broadband service and transmission reliability [13]. Slow or insufficient response to criticism and invoicing problems are other variables that increase users' likelihood of defecting to the opposition [14]. Users may defect to the opposition due to shipping expenses, insufficient functionality, and inadequate infrastructure. Regular customers evaluate suppliers and switch to whatever they believe offers substantially better values [15].

Telecommunications business can do OK if it can understand the importance of existing consumers. The typical churn rate for mobile subscribers throughout the telecom business is around 2% worldwide [16], resulting in a total yearly loss of around $100 billion. According to Kotler [17], encouraging a permanent client not to churn toward a rival costs 16 times less than looking for it and comes into contact with such a potential user. In contrast, new customer acquisition costs 5 to 6 times more than existing ones.

According to Sasser and Reichheld [18], an internet provider can boost profitability by 25 to 85% by lowering client attrition by 5%. It demonstrates the significant impact of customer attrition on a company's success. An examination of the churn rate in several businesses reveals that it is especially problematic in the telecommunications sector, in which it varies from 20 to 40% yearly [19].

Advanced technologies have allowed firms to comprehend that their marketing advantages must ensure strong client-customer retention even to compete. It is certainly relevant in the telecommunications sector. As a result, many studies are now being utilized to recognize clients who are likely to walk away [20].

There are several existing models for churn analysis in the telecom industry. For instance, [21] designed a classification model for customer categorization in the telecommunication sector. The model was based on a machine learning model. For the training and testing of the proposed model, the IBM Telecom Customer Churn set of data, which included over 7000 clients, was implemented in the methodology. The essential attributes were selected with the help of multiple clustering models. In the classification phase, SVM (Support vector machine) algorithm was used to categorize churn features. SVR (Support vector regression) algorithm was implemented to predict monthly service charges.

In this work, SVM is enhanced to increase the efficiency of the churn prediction method. The proposed solution depends on a kernel MSVM model that has been optimized. To extract the features in a matrix, the KPCA algorithm is utilized. The dimensionality of data is reduced without significant data loss. The Ant Lion optimizer is used to process the extracted features. ALO is an optimization module in the proposed architecture that helps lower the feature set's error probability. The test set loads the test subset and training module before applying the MSVM prediction model. MSVMs apply a risk-minimization strategy that contains the error, making this model a potential solution for predicting online datasets. This work is classified into different sections: In section 2, prior works are reviewed, and comparison tables are depicted for analysis. The proposed model are explained in section 4 with pseudocode and flowchart. Section 5 presented the experimental results analysis and comparison analysis of different churn analysis models.

## 2. PRIOR WORK

In this section, existing churn prediction models are analyzed, and comparative analyses are depicted for better performance validation. Sarac, F., Şeker et al., (2021) [21] proposed a machine learning-based framework for churn analysis. The framework was divided into different phases. Data was gathered from IBM TELECOM (data about 7000 users) in the initial phase. The essential attributes were selected with the help of multiple clustering models. In the classification phase, SVM (Support vector machine) method was used to categorize churn features. SVR (Support vector regression) method was implemented to predict monthly service charges. The proposed methodology had attained 81.5% of Accuracy and 85.6% of AUC. The monthly service charges with the SVR (Support vector regression) method had achieved 1.27 of root mean squared error (RMSE). The proposed model had attained better outcomes for churn analysis and efficiently forecasted the monthly charges of the service.

Deng, Y., Li, et al., (2021) [22] examined semi-annual financial institution customer data for developing customer churn forecasting models employing ensemble ml (machine learning) methods like Random Forest, Lightgbm, and Catboost to increase forecast efficiency as well as assist banking institutions with lowering costs. According to the empirical observations, the model's accuracy rate was reported at 90%, and the AUC value was greater than 80%. The methodology could also be used to determine if a customer would be abandoned later, plan aside a certain amount of time for user retention efforts, as well as provide a great deal of data to assist advertising people in emerging a practical customer retention strategy with a wide variety of manufacturing applications.

Wu, S., Yau, et al., (2021) [23] developed a churn managing model that integrates customer analytics. Statistics pre-processing, data exploration, churn forecasting, multivariate statistical, client retention, and user behaviour monitoring were six dimensions platforms were implemented in the proposed work. The methodology combined churn forecasting and targeted marketing to give telecommunication providers a comprehensive churn assessment for improved customer management. The tests employed three datasets and six computational models. First, computational intelligence classifiers were used to forecast the consumers' churn state. The Synthetic minority oversampling model was implemented in the training sample to address the issues with large datasets. The results were compared using 10-fold cross-validation. Models analysis was based on Accuracy and F1-score. Since the idea of churn planning was to be possible to perceive users who could churn, the F1-score was considered an essential statistic to evaluate models for outlier detection. According to the results, AdaBoost had achieved the highest in database 1, including accuracy of 77.19% and an f1-score of 63.11 %. In database 2, Random Forest showed the best performance, with an accuracy of 93.6 % and an F1-score of 77.20%. Random Forest was the most accurate in database 3, with a score of 63.09%, whereas Multi-layer Perceptron was the most accurate concerning f1-score, with a value of 42.84 %.

A machine learning-based framework was created by Hu, X., Yang, et al. in 2020 [24] based on an integrated model. The integration model made use of a decision tree and a neural network based on machine learning. This paper creates a statistical model for composite churn prediction and uses statistical results to test its performance. The findings show that, compared to the corresponding churn prediction statistical model, the integrated predictive model had greater results, precision forecasting impact, and the ability to depict the essential features of turnover consumers more transparently.

A regular churn forecasting-based model, based on the user's regularly dynamic features rather than his quarterly behaviour, was proposed by Alboukaey, N., et al. in 2020 [25]. In order to forecast daily turnover, the authors specifically characterised consumer behaviour as multidimensional data and offered four forecasts based on the description. The RFM-based models and the statistical data model relied on features taken from multidimensional data. Other methods, such as the long short term memory (LSTM) method and the CNN (Convolutional neural network) method, use deep learning methods for feature extraction and classification. The optional forecasting model's forecast ability was tested by measuring them that employed a 150-day format compatible with MTN operators across the region.

A methodology was created by Yu, R., An, et al. (2018) [26] based on a machine learning model. It was suggested to use back propagation (BP) networking for particle classification performance adjustment in order to estimate the amount of telephony customers who will churn time to time with the help of PFC (Particle Fitness Computation) and PCO (Particle Classification Optimization).The particle Classification Optimization method divides the particle into three groups based on their fitness scores and uses various calculations to modify the velocities of each group. In every progressive training program of a BP neural network, Particle Fitness Computation determines a particle's optimal solution. The proposed methodology enhanced the Accuracy of churn prediction by optimizing the weight values and the BP (Backpropagation) neural network threshold.

Machine learning was used to build the technique by Lalwani et al. in 2021 [27]. The suggested strategy is composed of six steps. the test dataset outcomes are examined by the confusion matrix and AUC curve. 81.7% accuracy was reported by Adaboost classification model. The XGboost classification method was reported 80.8%. Both XGBoost and Adaboost classification methods got the greatest AUC score of 84%, outperforming other existing methods.

Zheng, et al., (2021) [28] presented two domain-agnostics and easy-to-understand advancements: ClusPred. This broad and interpretable clustering-based churn prediction pipeline integrates all demography and transaction information. An Inhomogeneous Poisson process was implemented to analyze customers' behaviour and identify churners.

Machine learning was utilised by Butgereit et al. (2020) [29] to predict when customers were about to leave and while churn was being looked at. These estimates were then used to look for justifications for why and how the user would be cycling through unstructured or semi-structured user input log files.

A churn estimation model was created by Bayrak, A. T., et al., in 2020 [30] with the aid of a sophisticated learning technique. A framework for estimating customer attrition was developed by Jain et al. in 2020 [31]. The proposed methodology made use of the machine-learning techniques i.e. logit boost and logistic regression.

Ullah, et al., (2019) [32] introduced a churn prediction system, which employed clustering and classification approaches to detect churn consumers and the processes that influence customer attrition within the telecommunication sector. A telecom churn prediction approach that incorporates ensembles layering and uplifting-based techniques was proposed by Ahmed et al. in 2019 [33]. Ruihui et al. 2019 [37] Herrero-Lopez, S. "Multiclass support vector machine", In GPU Computing Gems Emerald Edition, pp. 293-311, 2011 described the uses of DL (Deep learning), several models of DL, and optimizations. It described the various framework and research areas of DL and the new acceleration technology of DL.

Jun-ki Hong 2021[38] described the prediction of product sales as they related to alters in temp. was applied. LSTM model, which defined better performance for TS (time-series) prediction. For identification proves has introduced the sales of short pants, flip flops, etc. were predicted depends on sales in temp. and TS sales information of clothing products gathered from 2015 to 2019.

In Table 1, various existing models of churn prediction along with performance metrics and research gap are depicted for a better understanding.

***Table 1.*** *Research gaps in existing models of Churn Analysis*

| Ref. | Proposed Model(s) | Performance metrics | Research gaps/Problems |
|---|---|---|---|
| [21] | The hybrid two-level SVM model | AUC , Accuracy, RMSE | Need to enhance the ability of generalization for selecting different features |
| [22] | Catboost, Lightgbm, RF (random forest) | Accuracy , AUC , F1-Score, Recall | Need to select more essential features for effective results |
| [23] | LR , DT, RF, NB, AbaBoost, Multi-layer Perceptron, K-means | AUC , Recall, Precision, F1-score, Accuracy, Computation Time | Imbalanced datasets generated under sampling and oversampling issues |
| [24] | DT and NN | Accuracy | Low stability and accuracy outcomes |
| [25] | LSTM with CNN model, RF | AUC ,F1-Score, Log Loss, Lift, EMPC | Insufficient results Interpretability issues |
| [26] | PCO, PDCCP | Prediction Accuracy : TP,TN,FP,FN | Poor results of classification |

The description of dataset and simulation tool along with conclusion are showed in Table 2.

***Table 2.*** *Existing models of churn analysis: dataset, simulation tool, conclusion*

| Ref. | Dataset | Simulation Tool | Conclusion |
|---|---|---|---|
| [21] | IBM Telecom dataset | e-commerce tools | Evaluated and performed at the similar time with a maximum accuracy rate. |
| [22] | Bank user dataset | - | The model helps determine the relationship among a user's and provider's characteristic, client sales information, and revenue, along with particular mathematical formulas or guidelines for calculating customer churn risk. |
| [23] | IBM Telecom Dataset 1 Kaggle Telecom Dataset 2 | Jupyter Notebook with Python 3 | Hybrid optimized model will be implemented in further enhancement |

| | | | |
|---|---|---|---|
| | Cell2Cell Dataset 3 | | |
| [24] | Supermarket customer information Dataset | - | More data will be collected for reliability of the system |
| [25] | MTN dataset | Python 3 using Scikit-learn 0.22.0 Keras 2.2.4 with TensorFlow 2.1.0 | The suggested model is more logistically economical than the monthly modelesand also anticipates churners faster and much more correctly. |
| [26] | CMCC | - | Proposed model provided highly efficient results with maximum Accuracy |

## 3. EXISTING MODEL

In this section, various methods i.e. SVM, SVR, and MCFS algorithm for classification are briefly discussed for existing HTL-SVM Model [21].

### 3.1. Hybrid Two-Level SVM (HTL-SVM) Model

This section aims to provide a hybrid two-level SVM-SVR model that combines the essential qualities (features) of both SVR and SVM. The fundamental concept is to exploit the stochastic nature of SVR to identify difficult-to-classify cases, which will then be used to develop a Performance Oriented Support vector machine. This existing article has surveyed the detailed description of SVM, SVR, MCFS (multi-cluster feature selection) algorithms utilized for the mathematical validation of all methods analyzed.

#### 3.1.1. SVM and SVR Methods

SVM can be thought of as an arithmetic ML technique that has been applied for both classification and regression (R) purposes. Support vector regression (SV) is the name given to the "R" version of SVM because a specific number of the training feature set is selected as SVs to map the link between inputs and outputs, and the weight sum of these SVs is increased to form the regression model. Both the SVM and SVR techniques are created using the proposed study's LibSvm library.

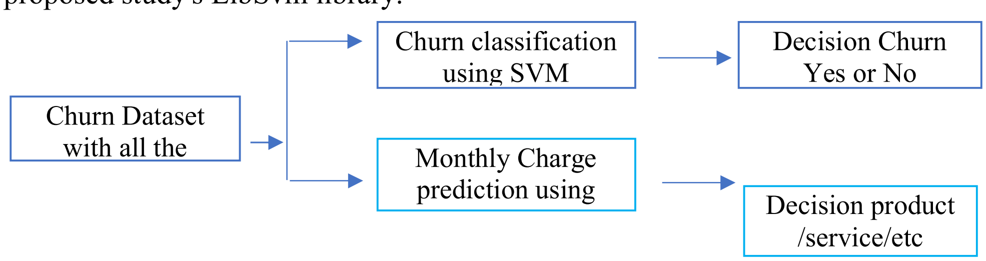

*Fig. 1. CA with the entire future set for the classification model of churn and the detection of product[21]*

### 3.1.2. MCFS Algorithm

An un-supervised FS (feature selection) algorithm chooses the best sub-set of the feature vectors for the detection tasks. It is one of the efficient models that achieve spectral “R” with l1 norm regularization to verify required feature sets, and It is autonomous of target outcomes.

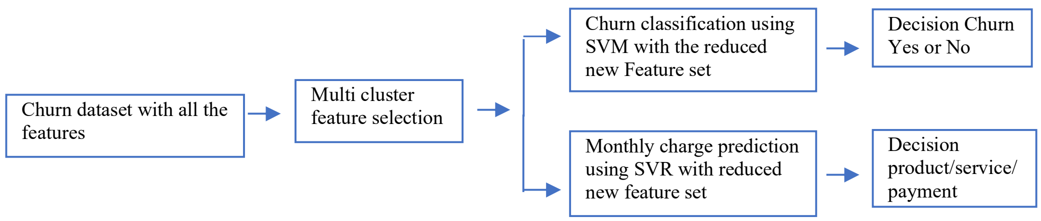


*Fig. 2. CA uses a multi-cluster feature selection based novel feature-set for the CC and the detection of service, etc. [21]*

It may be utilized for both un-supervised and supervised FS motives and classification and detection tasks.

## 4. PROPOSED MODEL

In this section, the three methods used in proposed OKMSVM model are briefly discussed i.e., the KPCA algorithm for feature extraction, ALO used for optimization and MSVM for classification.

### 4.1. Proposed Optimized Kernel MSVM (OK-MSVM) Model

This section explains the proposed optimized kernel MSVM model. For the optimization purpose, the ALO technique is used, and KPCA is used to reduce the dimensions of the features. The feature patterns combine M*N matrix numbers that make up an exclusive combination against a certain case. The KPCA modelling suggested architecture handles the feature extraction procedure. The features of a matrix are extracted using the KPCA technique. Without considerable data loss, the dimensionality of the data is reduced. In contrast to previous methods, it is applied to the dataset that can be linearly separated. To project the database into a high dimensional, linearly separable feature space, KPCA uses a kernel function (KF).The Ant Lion optimizer processes the extracted characteristics. The suggested architecture's ALO optimization module aids in lowering the likelihood of an error in the feature set. The training set with smaller error probabilities increases the trained model's accuracy. Finding the optimal cost solutions for input feature patterns is an iterative process. The data element is processed with the initialization module of the prediction model once all the cases have been processed.

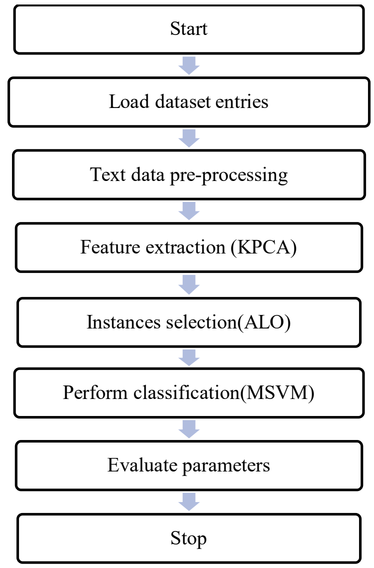


*Fig.5. Basic Proposed Model Using OKMSVM Model*

The optimized data elements are processed by this module, which also creates various subsets of the original datasets. The prediction model is trained, validated, and tested using these subsets. During this phase, the training subset and labels for those patterns are processed with the training module. To load and carry out test phases, it introduces the OKMSVM model [39] and stores it in the secondary storage device. Testing the Churn analysis method makes predictions based on the available input dataset. The test set loads the training module and test subset before starting the MSVM prediction algorithm. Because MSVMs employ a risk-minimization algorithm that contains the mistake, this model holds promise for forecasting online datasets. Following the prediction module, the performance of the suggested architecture is calculated using a variety of performance metrics. To create the comparison sets and verify the contribution of this improvement, the computed performance metrics are employed. The flowchart of the optimized kernel MSVM model is presented in fig.5.

#### 4.1.1. Feature Extraction Using KPCA Algorithm

Kernel Principal Component Analysis is an algorithm for simplifying data by transforming linear data into unique coordinates with many variations. It's a multi-variant numerical instrument for estimating data with several dimensions. It is utilized in every field of research to manipulate a huge number of factors. It assisted in lowering the data set's dimension, which included a huge number of connected variables, and it recalls the maximal variation in real statistics. It is accomplished by reducing the number of variables in the real variable set. As a result, a unique variable is linked to the categorization of actual values. The principal components are organized where small components are available in the current variables. As a result, the PCA's main goal is to recognize previously undiscovered patterns in data sets to reduce the dimensionality of information by removing noise and duplication [34]. It is determined by the covariance matrix's eigenvalue and eigenvector. As a result, executing the eigenvalue and eigenvector requires a time-consuming matrix change that may not be viable to perform manually. The PCA model's basic phases are listed below:

- Centre the Records

- Compute the covariance matrix
- Compute the eigenvector and eigenvalue
- Compute the result

| ***Feature Extraction Pseudo Code:*** |
|---|
| *Step 1: Construction of covariance matrix from the given dataset*<br>*Step 2: Matrix Eigenvector evaluation*<br>*Step 3: The original data set is reconstructed using high Eigenvalues of Eigenvectors.*<br>*Step 4: Principal components are considered from variables with significant variance.* |

#### 4.1.2. Optimization Using ALO Algorithm

This algorithm is a new meta-heuristic that numerically represents the dynamic relationship between ant lions and ants. Ant lion optimization is a framework that theoretically represents how ant lions and ants interact in nature. Doodlebugs is another name for ant lions. The ANTLION's life is separated into two phases, the larvae phase and the second of which is the mature phase. The mature phase has a lifespan of three to five weeks. Their hunting action is fascinating when they are in phase one (Larvae). Ant lions make the small traps in the shape of a cone, which can be observed in nature to capture ants. Ant lions lie beneath that trap and hold back for prey to be trapped. After chowing down on the prey's flesh, the leftovers are tossed outside the pit by ant lions, following which that pit is amended for the next hunt [35]. It has been noticed that bigger pits are dug up by ant lions when they are hungry, which is exactly the main insight for this model. In fig. 3 flow chart of ALO is depicted. The five most important steps of ALO are described below as:

- Ants random walk
- Trap building
- Entrapment of ants in traps
- Sliding ants towards ant lions (catching prey)
- Rebuilding the pit and Elitism.

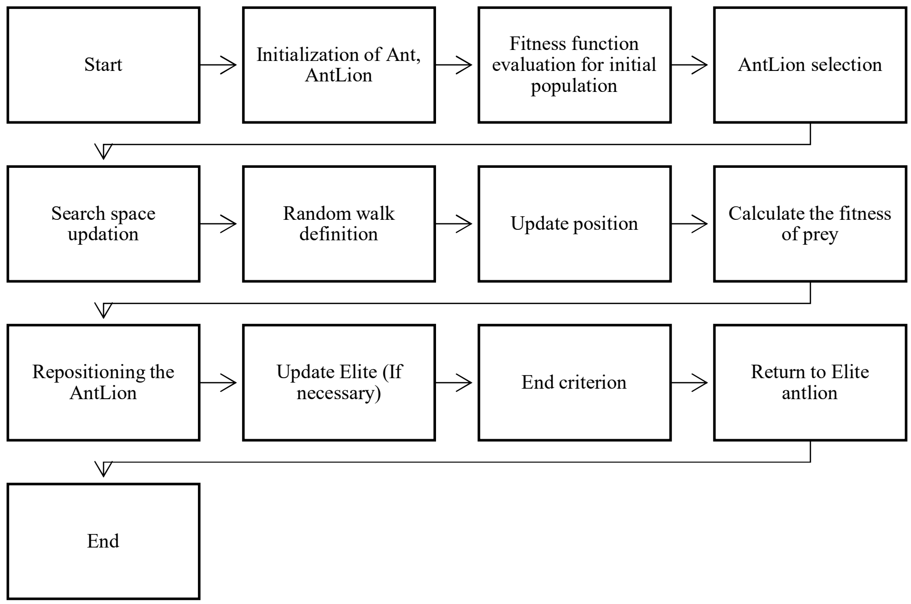


*Fig. 3. Workflow of Ant Lion optimization algorithm [35]*

***ALO-base Optimization Algorithm Pseudo Code***

***INPUT:*** *ANTS and ANTLIONS, ITERATION*

***OUTPUT:*** *Elitist ant lion and their fitness value (FV)*

*Step 1: Initialization of the agents (random locations) inside upper bound (ub) and lower bound (lb).*

*Step 2: Evaluate the population fitness value.*

*Step 3: Elite antlion selection.*

*While (until the condition is not met)*

*do*

*for (every ANT) do*

*Step 4: Selection the antlion A with the roulette wheel*

*Step 5: Run ANT random walk corresponding ANTLION*

*Step 6: ANT location updation*

*Step 7: Fitness value updation*

*Step 8: Generate the fitness metric and select the fittest agent*

*Step 9: Elite ant updation (if new Antlion is better)*

### 4.1.3 Multi-Class Support Vector Machine

Multi- SVM is the easy way to categorize multiple categories of data. MSVM, one of the ML- models based on SL (supervised learning), helps in classification and regression. The main aim of SVM is to provide a hyper-plane and produce the various class clusters. The MSVM is used to produce perception between two or more classes. The several steps of MSVM are shown in fig. 2. The process is started from preprocessing stage; after that, features are extracted from the image. The method is trained with extracted features and an MSVM classifier is used to classify data in multiple classes. Fig. 4 is the multi-class representation of SVM.

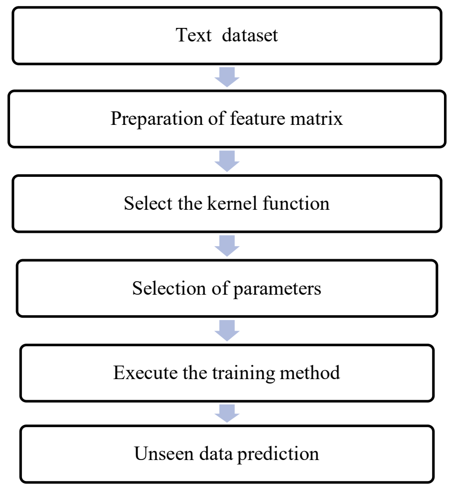


*Fig. 4. Steps of MSVM [36] Algorithm*

## 5. Experiment Results Analysis

This section gives insight to the tools and dataset used, experiment results analysis and covers the CA (churn analysis) experiment findings. The following work is divided into training, testing and validation Phases.

### 5.1. Tools and Dataset

The "***WA_Fn-UseC-Telco-Customer-Churn***" dataset [14] was used in the proposed research. This proposed database has 21 columns(referred to as "properties") and 7043 rows (referred to as "users") of raw data. 21 characteristics are used as the target position for the regression and categorization activities. Our target values are in the churn columns.

Python, a language with a scripting foundation, is used to make the training phase. It will develop the man-to-machine interfaces so that the user may click on them and view the results. Jupyter is the new web-based interactive development framework for notebooks, data and code. A reliable interface gives users to validate and manage workflows in the DS (data science).

### 5.2. Evaluation Parameters

Three parameters are used for the classification work with the optimised kernel MSVM to gauge how well the CCs are evaluated (Churn Classification). AUC, Accuracy Rate (%), and RMSE rate are the three. The area under the curve can be interpreted as the total amount of classification assessment required to carry out all anticipated categorization strategies. The accuracy percentage could be written as;

$$Accuracy = 100 * \frac{no.of\ correctly\ classified\ samples}{no.of\ all\ samples\ in\ the\ database} \text{.......(i)}$$

The calculation of the monthly charges is measured using the RMSE (Root Mean Square Error Rate) for the prediction and classification method with the optimised kernel MSVM algorithm. It could be stated as;

$$RMSE = \sqrt{\frac{\sum_i^n (y1_i - y1'_i)^2}{n1}} \quad \text{................................ (ii)}$$

The accuracy rate for the classification method using the optimised kernel MSVM algorithm is the proportion of correctly identified churn values to the entire dataset. It could be described as;

$$accuracy = \frac{tp+t}{tp+fp+fn+tn} \text{.................................(iii)}$$

Here, tn stands for True Negative, fp for False Positive, tp for True Positive, fn for False Negative, n1 for the total sample count in the tele-comm dataset, and y1i and y'1i for the intended and expected values, respectively.

**5.3. Implementation Results**

The database is split into two sections so that each strategy, such as training and testing modules, may be evaluated separately. The training programme is used to train the models, and the testing part is used to gauge the effectiveness of the models.

To effectively complete the analysis, a 10-fold cross-validation methodology that incorporates both CC (churn categorization) and forecasting activities is appropriate. It means that there are equal amounts of data in each of the subgroups into which the information has been separated.

While other subgroups are employed for training, the method is also evaluated using the residual set. This process is continued until every subgroup has been individually examined and the training phase has been finished.

```
In [160]: plt.figure(figsize=(8,4))
          g = sns.catplot(x="PaymentMethod", kind="count", data=df, hue="Churn")
          for ax in g.axes.ravel():
              ax.set_xticklabels(ax.get_xticklabels(), rotation=45)
          plt.title("Count of churned customer by payment method")
          plt.show()
```

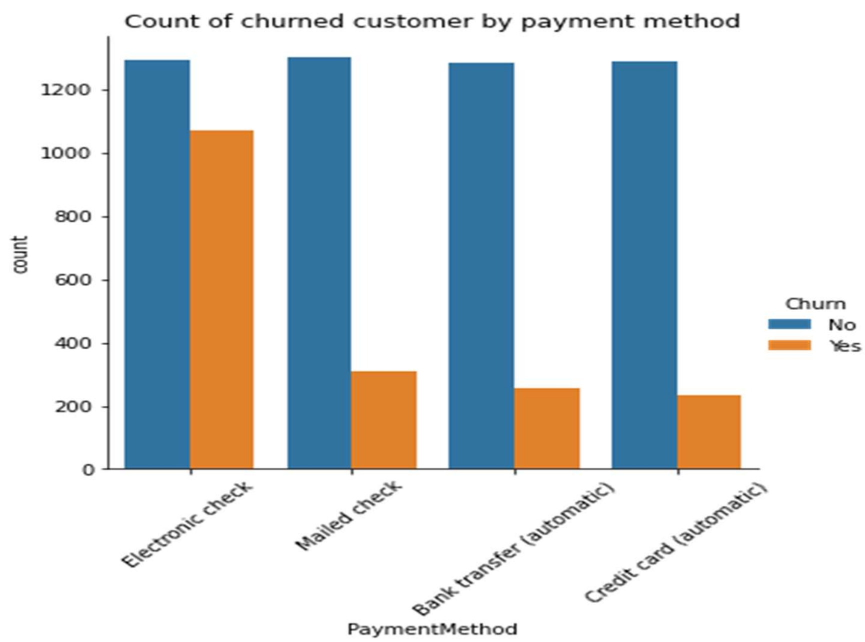


*Fig.7. Count of churn customers on basis of payment methods*

Figure 7 clearly illustrates that customers who pay through electronic cheques are more likely to churn as compared to customers using other modes of payment.

Table 3 discusses the proposed performance metrics such as Accuracy, AUC, and RMSE score. It sows various users and services.

***Table 3.*** *Comparison between proposed and existing classifier Models*

| ***Parameters*** | ***OK-MSVM*** | ***HTL-SVM*** |
|---|---|---|
| *Accuracy (%)* | *91.05* | *81.5* |
| *AUC Score* | *85.76* | *85.6* |
| *RMSE Score* | *2.838* | *3.01* |

```
In [187]: plt.figure(figsize=(8,4))
          g = sns.catplot(x="Partner", kind="count", data=df, hue="Churn")
          for ax in g.axes.ravel():
              ax.set_xticklabels(ax.get_xticklabels(), rotation=45)
          plt.title("count of churn by Partner")
          plt.show()
```

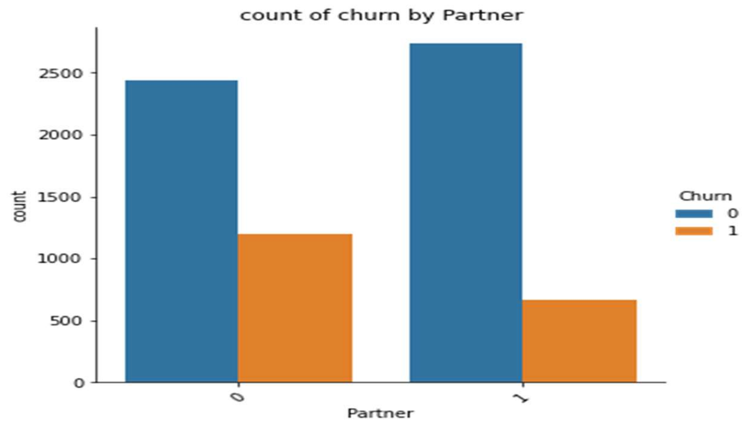


*Fig. 8. Count of churn on basis of partner*

Figure 8 clearly depicts that customers with no partner have higher churn rate.

The accuracy values for the proposed and existing models are defined in Table 3. The outcomes define that the research OK-MSVM based model with different FS (feature selected) using ALO the high classification accuracy rate of 91.05% and AUC 85.76%, which outperforms the outcomes defined in table 3. The proposed model with all the feature sets has a better root means square error rate than the existing HTL-SVM (hybrid two-level SVM) model. As OK-MSVM has optimized feature selection and classification model has been implemented and improved the performance metrics compared with the existing model.

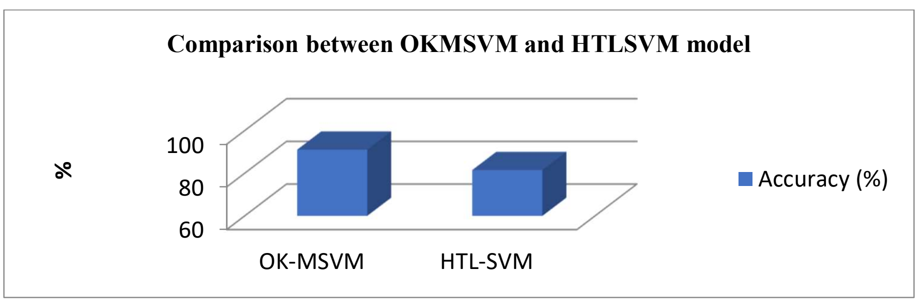


*Fig.9. Comparison between proposed and existing models with Accuracy (%)*

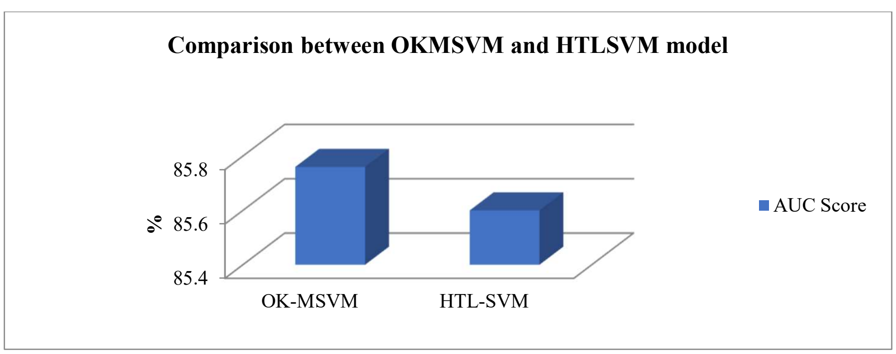


*Fig.10. Comparison between proposed and existing models with AUC*

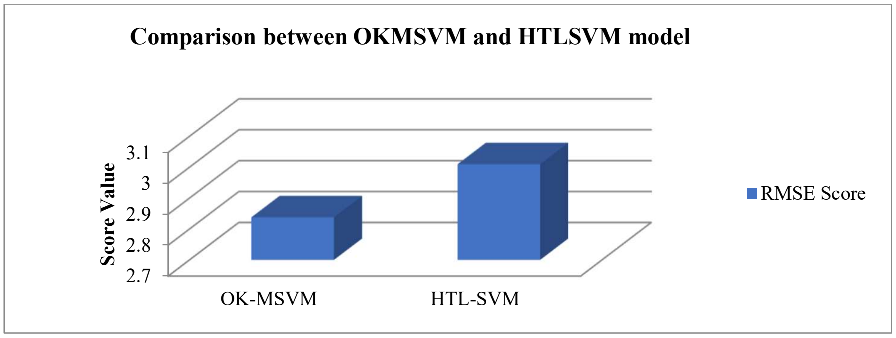


*Fig. 11. Comparison between proposed and existing models with RMSE Score*

Figs 9, 10, and 11 show the comparative analysis with the proposed model. This proposed model has improved the accuracy rate by 91%, the AUC rate by 85.76%, and reduced the error rate, which is RMSE value to 2.838 score. This proposed model has improved the error rate value and high accuracy rate associated with the HTL-SVM model.

## 5. Conclusion and Future Scope

In this work, the proposed OKMSVM model is implemented, validated and then compared with existing HTLSVM model. The proposed framework is divided into different phases. In the initial phase, the dataset is loaded, and then data is pre-processed for further processing. In the suggested design, the KPCA algorithm manages the feature extraction process. The features of a matrix are extracted using the KPCA technique. To project the database into a high dimensional, linearly separable feature space, KPCA uses a kernel function (KF). The ALO optimization is carried out using the extracted features. The optimized data items are processed in the classification stage, which also creates distinct subsets of the original datasets. The prediction model is trained, validated, and tested using these subsets using Python with the help of Jupyter notebook. The proposed framework has achieved efficient outcomes and performs better than the existing hybrid two-level SVM model. The OK-MSVM proposed model has achieved a high accuracy rate of 91%, an AUC value of 85.76%, and reduced the error rate value is 2.838 compared to the existing hybrid two-level SVM model.

## REFERENCES

[1] J. Ahn, J. Hwang, D. Kim, H. Choi and S. Kang, A Survey on Churn Analysis in Various Business Domains, *in IEEE Access*, vol. 8, pp. 220816-220839, 2020, doi: 10.1109/ACCESS.2020.3042657.

[2] Mahajan, V., Misra, R., & Mahajan, R., Review on factors affecting customer churn in telecom sector, *International Journal of Data Analysis Models and Strategies*, Vol. 9, pp.122-144, 2017, doi: 10.1504/IJDATS.2017.10006960

[3] Khodabandehlou, S., & Rahman, M. Z. Comparison of supervised machine learning models for customer churn prediction based on analysis of customer behavior, *Journal of Systems and Information Technology*, 2017, doi: 10.1108/JSIT-10-2016-0061.

[4] Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., & Abbasi, U. Improved churn prediction in telecommunication industry using data mining models, *Applied Soft Computing*, Vol. 24, pp. 994-1012, 2014. DOI: 10.1016/j.asoc.2014.08.041

[5] A. Ahmed and D. M. Linen, A review and analysis of churn prediction methods for

customer retention in telecom industries, *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2017, pp. 1-7, doi: 10.1109/ICACCS.2017.8014605.

[6] Almana, Amal M., Mehmet Sabih Aksoy and Rasheed AlZahrani, A Survey On Data Mining Techniques In Customer Churn Analysis For Telecom Industry, *International Journal of Engineering Research and Applications*, Vol. 4, pp.165-171, 2014.

[7] Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S., Customer churn prediction in telecommunication industry using data certainty, *Journal of Business Research*, Vol. 94, pp. 290-301, 2019, doi: https://doi.org/10.1016/j.jbusres.2018.03.003

[8] Abou el Kassem, Essam, et al., Customer Churn Prediction Model and Identifying Features to Increase Customer Retention based on User Generated Content, *International Journal of Advanced Computer Science and Applications*, Vol. 11, 2020, doi: 10.14569/IJACSA.2020.0110567

[9] Anjum, Adnan, et al., Optimizing coverage of churn prediction in telecommunication industry, *International Journal of Advanced Computer Science and Applications*, Vol. 8, pp. 179-188, 2017. DOI: 10.14569/IJACSA.2017.080523

[10] Sabbeh, Sahar F., Machine-learning techniques for customer retention: A comparative study, *International Journal of Advanced Computer Science and Applications*, Vol. 9, 2018, doi: 10.14569/IJACSA.2018.090238.

[11] Abou el Kassem, Essam, et al, Customer Churn Prediction Model and Identifying Features to Increase Customer Retention based on User Generated Content, *International Journal of Advanced Computer Science and Applications*, Vol. 11, 2020, doi: 10.14569/IJACSA.2020.0110567

[12] Wei, Chih-Ping, and I-Tang Chiu., Turning telecommunications call details to churn prediction: a data mining approach, *Expert systems with applications*, Vol. pp. 103-112, 2002. DOI :10.1016/S0957-4174(02)00030-1

[13] Anjum, Adnan, et al., Optimizing coverage of churn prediction in telecommunication industry, *International Journal of Advanced Computer Science and Applications*, Vol. 8, pp. 179-188, 2017, doi: 10.14569/IJACSA.2017.080523

[14] Khan, Yasser, et al, Customers churn prediction using artificial neural networks (ANN) in telecom industry, *Editorial Preface From the Desk of Managing Editor*, Vol. 10, 2019, doi: 10.14569/IJACSA.2019.0100918

[15] De, Soumi, P. Prabu, and Joy Paulose, Effective ML Techniques to Predict Customer Churn, *2021 Third International Conference on Inventive Research in Computing Applications*, IEEE, 2021.DOI: 10.1109/ICIRCA51532.2021.9544785

[16] Santharam, Akshara, and Siva Bala Krishnan, Survey on customer churn prediction techniques, *International Research Journal of Engineering and Technology*, Vol. 5, 2018, doi: 10.5120/ijca2016912237

[17] Kotler, P., Analysis, planning, implementation and control. Prentice Hall International, 1994.

[18] Reichheld, F. F., & Sasser, W. E., Zero Defections: Quality Comes to Services, *Harvard business review*, 68(5), 105-111, 1990.

[19] Shaaban, E., Helmy, Y., Khedr, A., & Nasr, M., A proposed churn prediction model, *International Journal of Engineering Research and Applications*, Vol. 2, pp. 693-697, 2012.

[20] Umayaparvathi, V., and K. Iyakutti, A survey on customer churn prediction in telecom industry: Datasets, methods and metrics, *International Research Journal of Engineering and Technology*, Vol. 3, 2016, doi: 10.1016/j.jbusres.2018.03.003

[21] Sarac, Ferdi and Şeker, Hüseyin and Lisowski, Marcin and Timothy, Alan, A Hybrid Two-Level Support Vector Machine-Based Method for Churn Analysis. *In: 2021 5th International Conference on Cloud and Big Data*, 13-15 August 2021, doi: https://doi.org/10.1145/3481646.3481665

[22] Deng, Y., Li, D., Yang, L., Tang, J., & Zhao, J., Analysis and prediction of bank user churn based on ensemble learning model, *In 2021 IEEE International Conference on Power Electronics, Computer Applications* (ICPECA), pp. 288-291, 2021. DOI: 10.1109/ICPECA51329.2021.9362520

[23] S. Wu, W. -C. Yau, T. -S. Ong and S. -C. Chong, Integrated Churn Prediction and Customer Segmentation Framework for Telco Business, *in IEEE Access*, vol. 9, pp. 62118-62136, 2021, doi: 10.1109/ACCESS.2021.3073776.

[24] Hu, X., Yang, Y., Chen, L., & Zhu, S., Research on a customer churn combination prediction model based on decision tree and neural network, *In 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics,* pp. 129-132, 2020, doi: 10.1109/ICCCBDA49378.2020.9095611

[25] Alboukaey, N., Joukhadar, A., & Ghneim, N., Dynamic behavior based churn prediction in mobile telecom, *Expert Systems with Applications*, Vol. 162, pp. 113779, 2020.DOI: https://doi.org/10.1016/j.eswa.2020.113779

[26] Yu, R., An, X., Jin, B., Shi, J., Move, O. A., & Liu, Y., Particle classification optimization-based BP network for telecommunication customer churn prediction, *Neural Computing and Applications*, Vol. 29, pp. 707-720, 2018, doi: 10.1007/s00521-016-2477-3

[27] Lalwani, P., Mishra, M.K., Chadha, J.S. et al., Customer churn prediction system: a machine learning approach, *Computing* 104, 271–294, 2022, doi: https://doi.org/10.1007/s00607-021-00908-y

[28] Zheng, Hanming, Ling Luo, and GoceRistanoski, A Clustering-Prediction Pipeline for Customer Churn Analysis, *International Conference on Knowledge Science, Engineering and Management,* Springer, Cham, 2021.DOI: 10.1007/978-3-030-82153-1_7

[29] Butgereit, Laurie, Big Data and Machine Learning for Forestalling Customer Churn Using Hybrid Software, *2020 Conference on Information Communications Technology and Society*, IEEE, 2020, doi: 10.1109/ICTAS47918.2020.233972

[30] Bayrak, Ahmet Tuğrul, et al, Churn prediction with sequential data using long short term memory, *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies*, IEEE, 2020, doi: 10.1109/ISMSIT50672.2020.9254679

[31] Jain, Hemlata, Ajay Khunteta, and Sumit Srivastava, Churn prediction in telecommunication using logistic regression and logit boost, *Procedia Computer Science*, Vol. 167, pp. 101-112, 2020, https://doi.org/10.1016/j.procs.2020.03.187

[32] Ullah, Irfan, et al., A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector, *In IEEE access*, Vol.7, pp. 60134-60149, 2017, doi: 10.1109/ACCESS.2019.2914999

[33] Ahmed, Ammar AQ, and D. Maheswari, An enhanced ensemble classifier for telecom churn prediction using cost based uplift modeling, *International Journal of Information*

*Technology*, Vol. 11, pp.381-391, 2019, doi: 10.1007/s41870-018-0248-3

[34] Suguna, R., Devi, S., & Mathew, R., Customer churn predictive analysis by component minimization using machine learning, *International journal of innovative technology and exploring engineering*, Vol. 8, pp. 3229-3233, 2019.

[35] Abualigah, L., Shehab, M., Alshinwan, M., Mirjalili, S., & Elaziz, M. A., Ant lion optimizer: a comprehensive survey of its variants and applications, *Archives of Computational Models in Engineering*, Vol. 28, No.3, pp. 1397-1416, 2021, doi: 10.1007/s11831-020-09420-6

[36] Pankaj Hooda, Pooja Mittal, “An Optimized Kernel MSVM Machine Learning-based Model for Churn Analysis” International Journal of Advanced Computer Science and Applications (IJACSA), 13(5), 2022.

[37] Herrero-Lopez, S., Multiclass support vector machine, *In GPU Computing Gems Emerald Edition*, pp. 293-311, 2011.

[38] Ralph Sherwin, A. Corpuz, An Application Method of Long Short-Term Memory Neural Network in Classifying English and Tagalog-Based Customer Complaints, Feedbacks, and Commendations, *International Journal on Information Technologies and Security*, vol. 13, pp. 89-100, 2021.

[39] Romansky, Radi, Statistical analysis of empirical network traffic data from program monitoring, *International Journal on Information Technologies and Security*, vol. 14, pp. 15-24, 2022.

***Information about the author(s):***

**Pankaj Hooda**–Mr. Pankaj Hooda has completed his Masters in Computer Applications from M.D.University, Rohtak. He is currently pursuing his Ph.D from Department of Computer Science & Applications, M.D.University, Rohtak, India. The main area of research include Data Mining, Artificial Intelligence and Machine Learning.

**Pooja Mittal** – Dr. Pooja Mittal is currently working as Assistant Professor at Department of Computer Science & Applications, M.D.University, Rohtak, India. She obtained her Ph.D from Maharshi Dayanand University. Her area of research and specialization includes Data Mining, Data Warehousing and Computer Science. She has published more than 50 research papers in renowned international and SCI Journals and attended more than 30 conferences.