BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **23**, No 4
Sofia • 2023 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2023-0039

Convex Least Angle Regression Based LASSO Feature Selection and Swish Activation Function Model for Startup Survival Rate

Ramakrishna Allu, Venkata Nageswara Rao Padmanabhuni

Department of Computer Science and Engineering, GITAM (Deemed to be university), Visakhapatnam, Andhra Pradesh, 530045, India
E-mails: Ramakrishna.allu@gmail.com / Ramakrishna.allu@yandex.com nagesh.uma@hotmail.com

Abstract**:** *A startup is a recently established business venture led by entrepreneurs, to create and offer new products or services. The discovery of promising startups is a challenging task for creditors, policymakers, and investors. Therefore, the startup survival rate prediction is required to be developed for the success/ failure of startup companies. In this paper, the feature selection using the Convex Least Angle Regression Least Absolute Shrinkage and Selection Operator* (*CLAR-LASSO*) *is proposed to improve the classification of startup survival rate prediction. The Swish Activation Function based Long Short-Term Memory* (*SAFLSTM*) *is developed for classifying the survival rate of startups. Further, the Local Interpretable Model-agnostic Explanations* (*LIME*) *model interprets the predicted classification to the user. Existing research such as Hyper Parameter Tuning* (*HPT*)*-Logistic regression, HPT-Support Vector Machine* (*SVM*)*, HPT-XGBoost, and SAFLSTM are used to compare the CLAR-LASSO. The accuracy of the CLAR-LASSO is 95.67% which is high when compared to the HPT-Logistic regression, HPT-SVM, HPT-XGBoost, and SAFLSTM.*

Keywords**:** *Convex Least Angle Regression based Least Absolute Shrinkage and Selection Operator* (*CLAR-LASSO*)*, Crunch base dataset, Startup, Survival rate, Swish activation function based Long Short Term Memory* (*LSTM*)*.*

1. Introduction

In the evolving economies it is required to observe the growth of entrepreneurship, particularly startups, that are pioneering all industries, and these startups are highly correlated with the capacity of economic growth [1]. The report from Global Startup Ecosystem states that the economy of global startup is continuously developed which offers around three trillion USD between 2017 and the first half of 2019. Moreover, the Crunchbase funding report states that investment from the private market has significantly enhanced and forecasts state that 1.5 trillion USD has been globally invested in the recent decade [2]. Small and Medium Enterprises (SMEs) are crucial contributors to the economy, but they struggle to survive in the modern business environment. Worldwide, the SME’s success rate is only around 40% [3, 4]. There are many startups developing effective services and products such as Zomato, Facebook, WhatsApp, Instagram, Snowflake, Paytm, and Uber [5]. The discovery of hopeful startups is a challenging task for creditors, policymakers, and investors. Even though each stakeholder has rich available data, hopeful startup information is required while deciding on probable involvement in a particular startup. This kind of information is rapidly processed which indicates that a simple heuristic is greatly valuable [6, 7].

The startup receives support using knowledge exchange, findings, financial benefits, residence licenses to entrepreneurs, and so on [8]. In exterior cooperation schemes associated with the ability of the company's technology, the startups are referred to as a better candidate for cooperation, due to its advanced technologies and it is considered for forging a modern borderline in high-end industries. Additionally, the startups enhance successful cooperation based on their inexpensiveness and flexible decision-making process [9]. Hence, the success and failure evaluation of startups become crucial for both forthcoming and well-established entrepreneurs, and stakeholders [10]. The startup evaluation is highly subjective and preliminary-stage startups are correspondingly unpredictable as there is less amount of historical data [11]. The survival prediction of the new venture is difficult because the output is mainly based on the environmental improvements and certain complexity of each venture [12]. In this research, an effective feature selection is developed for improving the classifications. Feature selection is used to obtain less amount of highly useful and meaningful features from existing feature sets without losing the data [13]. The high classification accuracy, less overfitting, and enhanced generalization ability for classifiers are achieved by removing the redundant/ irrelevant features which are also used to achieve the classifier with improved interpretability [14, 15].

The research contributions are concise as follows:

- The CLAR-LASSO-based feature selection is accomplished for selecting the optimal features. The CLAR-LASSO is an improved version of LASSO where the group of weights is evenly disseminated while performing the feature selection.
- After selecting the features, the SAFLSTM is used to predict the startup survival rate. Here, the SAF is specifically taken because of its non–non-monotonicity, unsaturation, and smooth features.

The remaining paper is presented as follows: Section 2 provides information about the existing works related to startup survival rate prediction. Detailed information on CLAR-LASSO with SAFLSTM is given in Section 3. The outcomes of the CLAR-LASSO are given in Section 4 whereas the conclusion is presented in Section 5.

2. Related work

The information about the existing works related to startup survival rate prediction is given in this section.

Veganzones [16] has presented the threshold to perform the failure prediction of corporations. The threshold model has been developed among the firm size and failure tendency for computing the size of the firm to divide the sample. The different classifiers considered in this prediction were k-nearest neighbors, neural network, logistic regression, extreme learning machine, and support vector machine. Additionally, the classification has been enhanced by using the hyperparameter optimization for each model while evaluating various inputs.

Ross et al. [17] have developed the machine learning model namely CapitalVX for Capital Venture eXchange for predicting the startups' outcomes. This CapitalVX has been used to identify whether it exited successfully over IPO, failed, or remained private. Here, the integration of k-nearest neighbors, XGBoost, Deep Learning, and Random Forests has been developed to analyze the prediction performances. To confirm the models are up to date, the feeds have been frequently refreshed in the CapitalVX system. The exit prediction has been provided by machine learning models with feature analysis which have been used to discover the company aspects as to whether it is a good investment or not. CapitalVX was required to concentrate on the survival rate prediction for startups.

Żbikowski and Antosiuk [18] have presented supervised machine learning such as SVM, XGBoost, and Logistic regression for estimating the success of business ventures. The feature of scaling approaches such as minimax normalization and standardization have been used to preprocess the data while investigating with logistic regression and SVM. This input scaling is mandatory with logistic regression due to the usage of lasso and ridge regularization. Moreover, the standardization changes each feature's distribution into zero-mean and unit-variance. The exhaustive grid search-based HPT has been used in logistic regression and SVM while a randomized search has been used in XGBoost. The incorporation of feature scaling and HPT are used to enhance the prediction.

Allu and Padmanabhuni [19] developed the prediction of success or failure rate for start-up companies using LSTM with SAF. The conventional ReLU are replaced by Swish units of Feed Forward Neural (FFN) in LSTM for enhancing the deep network training. Each layer has varied in the network for training the fully connected networks. Accordingly, the prediction acquired from LSTM with SAF has been used to eliminate the company that has a failure rate. The classification using LSTM with SAF is processed with all features from the preprocessing which leads to creating misclassification in some situations.

Fuertes-Callén, Cuellar-Fernández and Serrano-Cinca [20] have used the first-year financial statement for startup survival discovery. The survival discovery has been utilized for testing the hypotheses and evaluating whether variations in the second-year startup financial statements can denote the survival or bankruptcy of the company up to 8 years after its establishment. The developed work was required to perform pre-processing and to remove the redundant features to enhance the prediction.

Elhoseny et al. [21] have developed an Adaptive Whale Optimization Algorithm with Deep Learning (AWOA-DL) for developing modern financial distress prediction. The AWOA-DL has been used to identify whether the respective company is facing distress or not. The financial distress is predicted by using the Deep Neural Network (DNN) model. The AWOA-based hyperparameter tuning has been done for DNN which has been used to enhance the prediction. However, feature selection is required in AWOA-DL to select the optimum features from the overall feature set.

Kou et al. [22] have presented the bankruptcy prediction for Small and Medium-sized Enterprises (SMEs). The SME considered in this work utilizes transactional information and payment network-based variables where no accounting information is required during the analysis. A two-stage multiobjective feature selection has been developed to overcome the high-dimensional issue. The feature importance is measured, however, the causal relationship between transactional data-based variables and bankruptcy was unknown during the prediction.

3. Proposed method

This research performs an effective startup survival rate prediction using CLAR-LASSO and SAFLSTM.

The main process of this CLAR-LASSO is Dataset acquisition, pre-processing, feature selection, and classification. The CLAR-LASSO-based feature selection is used to remove the redundant features from the feature vector. Therefore, precise classification is achieved using SAFLSTM. The block diagram of the CLAR-LASSO is shown in Fig. 1.

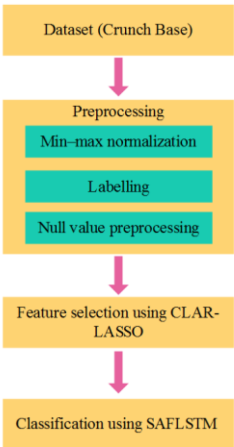

Fig. 1. Block-diagram of the proposed method

3.1. Data acquisition

560,000 community contributors exist in the crunch base [23] where it gains the data in four different ways In-house Data Team, Crunch Base Community, Venture Program, and Machine Learning. In Crunch base, the company is observed from startup to the parent company by utilizing TechCrunch. The data is acquired between 2007 and 2015 where the Tech Crunch is handled by the Crunch base database. The public submits the data to the companies in this dataset which offers a broad list of companies with huge amounts of data. In this proposed method, external and internal environment data is taken to evaluate the success/ failure of start-ups. This Crunch base database totally considers the data from the 48,124 companies. The visualization of the crunch base dataset is shown in Fig. 2. Further, the details about the top 20 companies are shown in Fig. 3 where it total has 27,168 companies. Moreover, the remaining 20,956 companies of the crunch base dataset are categorized as others.

Fig. 2. Visualization of crunch base dataset

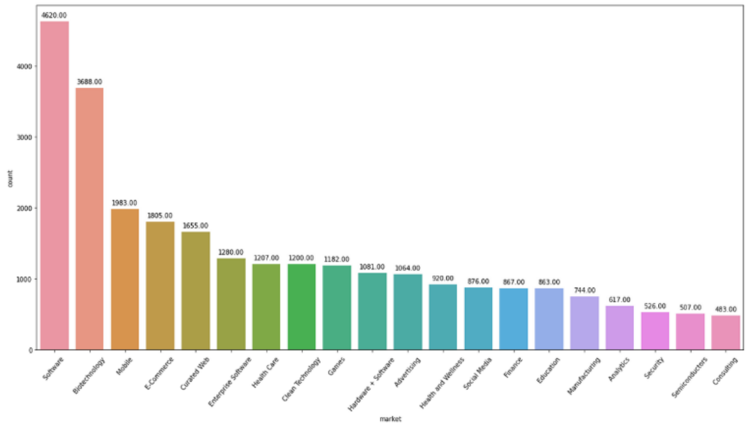

Fig. 3. Top 20 companies from the crunch base dataset

3.2. Preprocessing

The labels are created for Crunch Database and the inferred model is used to accomplish the labeling based on logarithmic rate. However, the labeling itself doesn't provide an improved prediction. The normalization is included by labeling functions for discovering the Trend score. Three different pre-processing approaches are considered in this research such as min-max normalization, labeling, and null value preprocessing processes which are detailed as follows:

- The min-max normalization promises that overall features are converted on the same scale. This min-max normalization is used to overcome the issues that occur by missing data happened to malfunctions, human error, or database failure. The normalization process is expressed in the next equation, which is used to convert the decimal to the whole number in the range of $[0, 1]$:

(1) $$X_{\text{norm}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}},$$

where normalized data is denoted as X_{norm}, data point is X_i, and minimum and maximum values of data points are $X_{\min}$ and $X_{\max}$, respectively. This normalization process fills the structured data and is modified or deleted with uncertain or incomplete information.

- In the labeling process, models or heuristics are applied to obtain the discovery of each row. The trend score saved in decimal form is found using regular expression functions. The function results in a suitable output format when the decimal form is discovered; otherwise, the function returns the value of –1. A unique class label, i.e., 0 or 1 received by each firm which decides as successful or not. The value 1 defines that it is successful whereas a label is obtained as 0 when it is not successful.
- Further, the average value of nearby data instances is included where there is a missing value in any of the input features while performing the null value preprocessing.

3.3. Feature selection using CLAR-LASSO

The preprocessed data is given as input to the CLAR-LASSO for choosing the relevant features to enhance the prediction. The developed CLAR-LASSO adds the sum of the squared value of weights. Therefore, the weights not only have smaller absolute values but also tend to penalize the weight's extremes resulting in the group of weights that are evenly disseminated during the feature selection. This CLAR-LASSO helps to eliminate the problem of multiple regression occurring in high dimensional data because these multiple regressions tend to create misclassification. CLAR-LASSO is utilized to eliminate the irrelevant/extra features or the forcefully connected features in the data. The developed LASSO regression performs the parameter evaluation and model selection while performing the regression analysis. The Objective Variable is represented as OV_i for i-th observation and preprocessed data is denoted as $\rho_i = (\rho_{i1}, \rho_{i2}, \dots, \rho_{im})$ which is the given input features. Next equation expresses the linear regression model,

$$\widehat{\mathrm{OV}}_i = \alpha + \delta\rho_i = \alpha + \sum_{j=1}^{m} \delta_j \rho_{ij}, \tag{2}$$

where: the inner product of ρ_i and the vector $\delta = (\delta_1, \delta_2, \dots, \delta_m)$, denoted as $\delta\rho_i$; δ_j is the feature j's coefficient; α is intercept, and m represents the number of features. Next equation is the L2 regularization penalty which is utilized in the estimator of LASSO ridge regression:

$$\delta_{\text{Ridge}} \operatorname*{argmin}_{\delta} \left\{ \sum_{i=1}^{N} \left(\mathrm{OV}_i - \widehat{\mathrm{OV}}_i\right)^2 + \lambda \sum_{j=1}^{m} \delta_j^2 \right\}, \tag{3}$$

where the $\lambda \sum_{j=1}^{m} \delta_j^2$ denotes the L2 regularization penalty of coefficient δ_j, and $\lambda \geq 0$ denotes the tuning parameter. Therefore, the selected features from CLAR-LASSO is given as input to the SAFLSTM for classifying the success/ failure rate of startups.

The convex least angle regression added in LASSO is expressed in the equation

$$\nabla_j f(\delta) - w_j^{-1} \operatorname{sgn}\left(\nabla_j f(\delta)\right) s(t) = 0, j \in \mathcal{A}, \tag{4}$$

where: the partial derivative of $f(\delta)$ is denoted as $\nabla_j f(\delta)$; the active index is denoted as $\mathcal{A}$; the sign function is denoted as sgn; the predictor-specific weight is denoted as w_j. This CLAR-LASSO is used for accomplishing even distribution of a group of weights that helps to choose the features (x). The chosen features and their importance are shown in the Table 1 and Fig. 4.

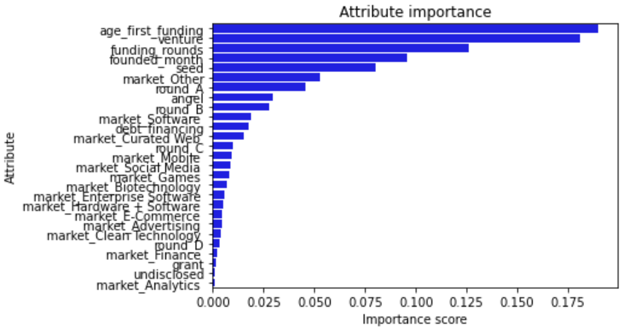

Fig. 4. Importance of feature chosen by CLAR-LASSO

Table 1. Chosen features with the Importance value

Attribute	Importance
age_first_funding	0.189865
venture	0.181118
funding_rounds	0.125819
founded_month	0.095607
seed	0.080231
market_Other	0.052832
round_A	0.045549
angel	0.029827
round_B	0.027900
market_Software	0.018805
debt_financing	0.017495
market_Curated Web	0.015530
round_C	0.009634
market_Mobile	0.009595
market_Social Media	0.008555
market_Games	0.008054
market_Biotechnology	0.006705
market_Enterprise Software	0.005703
market_Hardware + Software	0.005010
market_E-Commerce	0.004740
market_Advertising	0.004740
market_Clean Technology	0.004239
round_D	0.003391
market_Finance	0.002235
grant	0.001541
undisclosed	0.001272
market_Analytics	0.000963

The contribution of the selected feature using CLAR-LASSO is given as follows:

- **age_first_funding.** It denotes the startup age when it received its 1st round of funding. It is a pointer of how rapidly the startup is able to gain investment, that is a positive signal if it's comparatively low.
- **Venture.** It is a binary variable that denotes whether the startup gathered venture capital funding or not. Venture capital is frequently linked with high-growth potential, so it is a robust interpreter of success.

- **funding_rounds.** An amount of funding rounds a startup has gone through indicates its capacity to gain continuous support and investment. Startups that have effectively secured multiple rounds of funding are considered highly promising.
- **founded_month.** The month in which the startup originated has seasonal or market-specific consequences. Startups in specific industries fare better when it is created in specific months.
- **Seed, round_A, round_B, round_C.** These factors denote various levels of creating rounds (Seed, Series A, Series B, Series C, etc.). The amount of funding at every level and the capacity to develop via these levels are robust indicators of success.
- **Angel.** Angel investment is a primary indication of investor assurance. This binary variable designates whether the startup got angel funding or not.
- **market_Other, market_Software, market_Curated Web.** These binary variables probably denotes the market sector or industry in that the startup functions. A dissimilar markets has different stages of competition and growth, so the market sector is an important forecaster of success.
- **debt_financing.** Debt financing specify a startup's capability for securing the non-equity financing that is considered as an indication of financial stability and creditworthiness.

The aforementioned features are given as input to the SAFLSTM to perform the survival rate prediction of startup.

3.4. Classification using SAFLSTM

After extracting the features from the CLAR-LASSO, the chosen features are given as input to SAFLSTM for classification. In this phase, the conventional LSTM [24] is combined with SAF [25] to enhance classification performance. The SAFLSTM cell contains input, forgetting, and output gates, and data flow denoting a long-term memory is incorporated to develop a black box of input x and output state s. These aforementioned features support the SAFLSTM to train effectively, therefore historical sequence data is completely used in the classification. The process of SAFLSTM is formulated in the following equations:

(5) $$f_t = \text{Swish}\ (W_\text{f}.[h_{t-1}, x_t] + b_\text{f}),$$

(6) $$i_t = \text{Swish}\ (W_\text{i}.[h_{t-1}, x_t] + b_\text{i}),$$

(7) $$\tilde{c}_t = \tanh(W_\text{c}.[h_{t-1}, x_t] + b_\text{c}),$$

(8) $$c_t = f_t.c_{t-1} + i_t.\tilde{c}_t,$$

(9) $$o_t = \text{Swish}\ (W_\text{o}.[h_{t-1}, x_t] + b_\text{o}),$$

(10) $$h_t = o_t.\tanh(c_t).$$

Here: the output of forgetting, input, and output gates are denoted as f_t, i_t, and o_t, respectively; current input is denoted as x_t; the previous and current state from the hidden layer is denoted as h_{t-1} and h_t, respectively; the previous and current state's memory information is denoted as c_{t-1} and c_t, respectively; $\tilde{c}_t$ is candidate memory cell; the weight matrices forgetting input connect, and the hidden layer output are denoted as $W_\text{f}, W_\text{i}, W_\text{c}$, and W_o; offset vectors are denoted as b_f, b_i, b_c and b_o, and the swish activation function is denoted as Swish.

In this SAFLSTM, the SAF is used instead of the sigmoid activation function for enhancing the classification. Swish is motivated by the utilization of the sigmoid function for gate control in LSTM and highway networks where this SAF is expressed in the equation

$$\text{Swish}\ (x) = x.Sigmoid\ (x). \tag{11}$$

Further, the derivative expression of the swish is expressed by

$$\text{Swish}'(x) = \frac{e^x(1+e^x+x)}{(1+e^x)^2}. \tag{12}$$

The impressive features of SAF are non–non-monotonicity, unsaturation, and smoothness, which help to achieve effective performances with local response normalization. Moreover, the weights of SAFLSTM are updated using an input training set which helps to reduce the error during prediction. Generally, the LSTM is a deep neural network and data broadcasted in the data flow includes the key information, and historical memory is retained based on weight adjustment.

3.5. Local Interpretable Model-agnostic Explanations (LIME) method

After completing the classification, the LIME method is used to interpret the predicted outcomes to the user. This LIME method is used to represent which feature favors what class with a probability ratio. Generally, the LIME [26] observes the locally interpretable model around predicting and it is an approach that details what it observes based on the classifier's prediction. LIME is integrated with the SAFLSTM for mixing the input data and noticing the black box model's output for understanding how classification varies with various observations in the succeeding process. The sample report for the LIME explanation is shown in Fig. 5. Fig. 5 describes the data sample that is predicted as a successful startup performance. The parameters of many founders' exits, investments, and news articles create a positive impact on success whereas the number of founded organizations creates a negative impact on success.

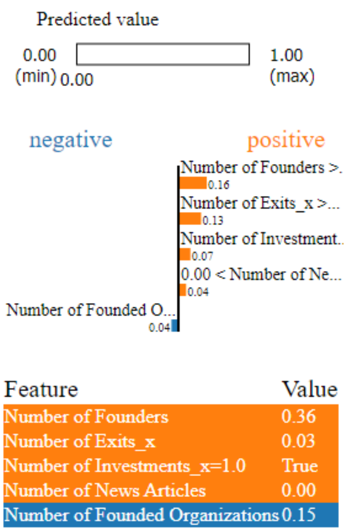

Fig. 5. Sample report for LIME explanation

4. Results and discussion

The proposed method has been implemented and simulated using Anaconda Navigator and Python 3.6 software. The system has been configured with Windows 10 operating system, 128 GB RAM, 22 GB RAM for RTX 2080 Ti GPU, 1 TB memory, and i9 processor. The developed proposed method is used to find the success and failure prediction of startups using the Crunch base dataset. The proposed method is evaluated using accuracy, precision, recall, F-measure, AUC, MAE, MSE, and RMSE, which are expressed in the next equations:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TN+TP+FN+FP}} \times 100\%, \tag{13}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \times 100\%, \tag{14}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \times 100\%, \tag{15}$$

$$\text{F-measure} = \frac{2\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \times 100\%, \tag{16}$$

$$\text{AUC} = \frac{1}{2} \times \left(\frac{\text{TP}}{\text{TP+FN}} + \frac{\text{TN}}{\text{TN+FP}}\right) \times 100\%, \tag{17}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}|, \tag{18}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2, \tag{19}$$

$$\text{RMSE} = \sqrt{\text{MSE}}. \tag{20}$$

Here: TP is the True Positive; TN is the True Negative; FP is a False Positive and FN is a False Negative; y_i and $\hat{y}$ represents original and predicted values; N is the data amount. Specifically, the prediction of startup success as success is denoted as TP; the prediction of startup failure as failure is denoted as FN; the prediction of startup success as failure is denoted as TN, and the prediction of failure as success is denoted as FP.

4.1. Performance analysis

The proposed method is analyzed with two different cases such as analysis with different feature selection approaches and analysis with different classifiers. A detailed analysis of the proposed method is given in the following sections.

4.1.1. Evaluation of the proposed method for different feature selection approaches

The CLAR-LASSO-based feature selection used in this research is compared with Pearson correlation, logistic regression, LASSO, and LAR LASSO. The comparison of CLAR-LASSO with Pearson correlation, logistic regression, LASSO, and LAR-LASSO is shown in Table 2. Further, the graph for CLAR-LASSO comparison in terms of classification measures is displayed in Fig. 6.

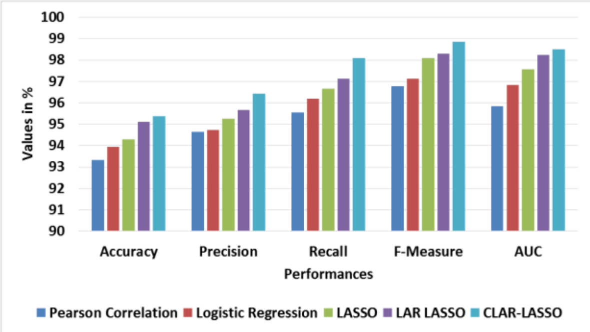

Fig. 6. Classification measure graph for different feature selection approaches

This analysis shows that the CLAR-LASSO achieves better performance than the Pearson correlation, logistic regression, LASSO, and LAR-LASSO. The accuracy of CLAR-LASSO is 95.36% where the Pearson correlation obtains 93.334%, logistic regression obtains 93.928%, LASSO obtains 94.299% and LAR LASSO obtains 95.094%. The CLAR incorporated in the LASSO is used to accomplish an even distribution of a group of weights that is used to select optimal features that predict the precise class. Therefore, the CLAR-LASSO provides better performances than the Pearson correlation, logistic regression, LASSO, and LAR LASSO.

Table 2. Analysis of the proposed method for different feature selection approaches

Feature selection methods	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Pearson correlation	93.334	94.653	95.553	96.773	95.847	0.588	0.587	0.795
Logistic regression	93.928	94.725	96.19	97.124	96.828	0.587	0.587	0.79
LASSO	94.299	95.262	96.644	98.077	97.572	0.582	0.582	0.781
LAR LASSO	95.094	95.661	97.133	98.291	98.222	0.573	0.581	0.777
CLAR-LASSO	95.36	96.43	98.09	98.84	98.51	0.565	0.574	0.775

4.1.2. Evaluation of a proposed method for different classifiers

The SAFLSTM used in this research has been analyzed with different classifiers such as logistic regression, decision tree, random forest, EXtreme Gradient Boost (XGB), Support Vector Machine (SVM), Artificial Neural Network (ANN), Recurrent Neural Network (RNN) sigmoid, RNN TanH, RNN Relu, RNN Softmax, RNN Swish, LSTM Sigmoid, LSTM TanH, LSTN Relu, and LSTM Softmax. In this evaluation, the classifiers have been examined with and without CLAR-LASSO as well and the classifiers have been evaluated for different k-fold validations such as 5, 10, 15, and 20. Tables 3, 4, 5, and 6 show the proposed method evaluation of different classifiers without CLAR-LASSO for k-fold values of 5, 10, 15, and 20 respectively. This analysis shows that SAFLSTM provides a better classification than the Logistic Regression, Decision Tree, Random Forest, XGB, SVM, ANN, RNN sigmoid, RNN TanH, RNN Relu, RNN Softmax, RNN Swish, LSTM Sigmoid, LSTM TanH, LSTN Relu and LSTM Softmax. Specifically, the SAFLSTM with $k = 20$ provides better classification than the k-fold values of 5, 10, and 15. The graph of classification parameters for LSTM with different activation functions using $k = 20$ without CLAR-LASSO is shown in Fig. 7. The accuracy of SAFLSTM for $k = 20$ is 72.419% which is high when compared to conventional classifiers. The multiplier that exists in SAFLSTM is used to make the classification highly near to the exact class of startup survival rate prediction which increases the classification and error metrics.

Table 3. Analysis of the proposed method for different classifiers without CLAR-LASSO and $k = 5$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	46.760	46.870	47.210	45.940	44.380	1.010	0.863	1.289
Decision Tree	48.280	48.730	49.050	47.420	45.560	0.970	0.792	1.213
Random Forest	50.170	49.990	50.110	49.080	47.140	0.878	0.709	1.135
XGB	51.870	51.500	51.850	50.250	48.700	0.832	0.631	1.132
SVM	53.800	53.210	53.390	51.920	50.440	0.764	0.627	1.032
ANN	54.890	54.480	55.340	53.120	52.230	0.750	0.577	0.953
RNN sigmoid	56.070	56.050	56.750	54.740	53.320	0.728	0.538	0.880
RNN TanH	57.330	57.820	57.930	56.370	54.390	0.707	0.470	0.829
RNN Relu	58.430	59.500	59.090	58.290	56.030	0.683	0.468	0.762
RNN Softmax	59.570	61.240	60.810	59.390	57.610	0.647	0.434	0.760
RNN Swish	61.230	62.430	62.160	61.290	59.420	0.647	0.391	0.746
LSTM Sigmoid	62.420	63.630	63.340	62.580	61.220	0.614	0.334	0.660
LSTM TanH	64.580	65.660	65.340	65.390	63.760	0.602	0.321	0.609
LSTN Relu	66.980	68.080	67.610	68.120	66.020	0.547	0.265	0.518
LSTM Softmax	69.020	71.060	69.880	70.980	68.690	0.454	0.229	0.462
SAFLSTM	71.640	73.450	72.810	73.270	70.920	0.383	0.206	0.453

Table 4. Analysis of proposed method for different classifiers without CLAR-LASSO and $k = 10$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	44.385	48.658	47.439	47.847	44.766	1.238	0.773	1.375
Decision Tree	46.345	50.148	48.509	49.577	46.206	1.230	0.743	1.335
Random Forest	47.625	52.058	50.109	51.537	47.976	1.153	0.701	1.252
XGB	49.485	53.938	52.069	52.797	49.266	1.111	0.630	1.210
SVM	50.915	55.408	53.709	53.827	50.796	1.066	0.620	1.125
ANN	52.635	56.748	54.789	54.957	52.456	0.987	0.603	1.035
RNN sigmoid	54.455	58.348	56.059	56.257	54.126	0.914	0.568	0.969
RNN TanH	55.475	59.718	57.409	58.107	55.566	0.883	0.474	0.870
RNN Relu	56.815	61.138	59.399	59.947	57.176	0.842	0.472	0.858
RNN Softmax	58.415	62.468	61.069	61.187	58.986	0.837	0.390	0.765
RNN Swish	60.285	63.568	62.499	63.177	60.206	0.767	0.316	0.763
LSTM Sigmoid	61.765	64.588	63.709	64.377	61.236	0.736	0.291	0.703
LSTM TanH	64.125	66.628	65.759	67.017	64.106	0.649	0.274	0.697
LSTN Relu	66.925	69.278	67.909	69.317	66.706	0.552	0.222	0.604
LSTM Softmax	69.525	72.028	70.729	71.347	69.026	0.462	0.218	0.531
SAFLSTM	71.715	74.288	73.519	73.387	71.746	0.383	0.205	0.453

Table 5. Analysis of proposed method for different classifiers without CLAR-LASSO and $k = 15$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	42.948	50.947	48.187	45.175	43.732	0.995	0.841	1.235
Decision Tree	44.818	52.077	49.607	47.175	45.072	0.983	0.757	1.162
Random Forest	46.108	53.517	51.447	49.095	46.672	0.924	0.687	1.162
XGB	47.218	54.627	53.107	50.895	47.992	0.878	0.669	1.126
SVM	48.948	55.817	54.667	52.465	49.292	0.856	0.584	1.068
ANN	50.318	57.237	56.077	54.265	50.922	0.816	0.490	0.993
RNN sigmoid	52.298	58.307	57.357	55.545	52.592	0.740	0.475	0.950
RNN TanH	53.808	59.667	58.517	56.725	54.082	0.721	0.462	0.858
RNN Relu	55.418	61.007	59.687	58.475	55.312	0.711	0.363	0.780
RNN Softmax	57.288	62.647	60.797	60.335	57.232	0.628	0.346	0.735
RNN Swish	59.248	63.967	62.147	62.115	59.112	0.565	0.322	0.648
LSTM Sigmoid	61.218	65.377	63.207	63.585	61.062	0.505	0.297	0.588
LSTM TanH	63.958	68.357	65.527	66.075	63.622	0.491	0.285	0.512
LSTN Relu	66.738	70.777	68.037	68.445	66.312	0.437	0.256	0.470
LSTM Softmax	69.338	72.877	70.757	71.325	68.932	0.406	0.215	0.456
SAFLSTM	71.738	75.167	73.597	74.105	71.912	0.383	0.205	0.453

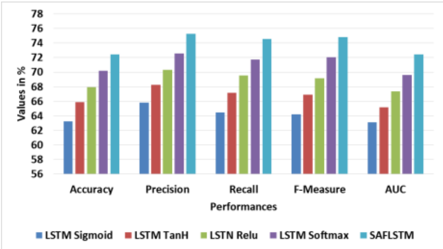

Fig. 7. Classification Measure Graph of LSTM's different activation functions for $k = 20$ without CLAR-LASSO

Table 6. Analysis of the proposed method for different classifiers without CLAR-LASSO and $k = 20$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	46.769	49.496	48.862	46.836	45.245	1.184	0.876	1.039
Decision Tree	48.439	51.406	50.262	48.196	46.885	1.125	0.843	0.941
Random Forest	49.599	52.516	51.482	49.876	48.785	1.113	0.808	0.915
XGB	51.529	53.556	53.082	51.796	50.275	1.042	0.804	0.868
SVM	52.569	54.996	54.302	53.106	51.805	0.980	0.732	0.840
ANN	53.979	56.766	55.682	54.756	53.415	0.954	0.675	0.825
RNN sigmoid	55.889	58.306	56.712	55.906	55.105	0.874	0.640	0.801
RNN TanH	56.899	60.156	58.402	57.186	56.665	0.822	0.588	0.773
RNN Relu	58.489	61.716	59.842	59.126	57.925	0.742	0.508	0.716
RNN Softmax	59.509	63.326	61.542	61.026	59.775	0.730	0.495	0.687
RNN Swish	61.359	64.326	62.732	62.556	61.125	0.662	0.408	0.652
LSTM Sigmoid	63.259	65.806	64.462	64.236	63.125	0.632	0.377	0.630
LSTM TanH	65.899	68.246	67.152	66.906	65.185	0.533	0.296	0.578
LSTN Relu	67.959	70.326	69.552	69.166	67.345	0.506	0.255	0.524
LSTM Softmax	70.229	72.606	71.712	72.086	69.605	0.444	0.249	0.485
SAFLSTM	72.419	75.256	74.572	74.806	72.465	0.383	0.205	0.453

Table 7, 8, 9, and 10 shows the proposed method evaluation of different classifiers with CLAR-LASSO for k-fold values of 5, 10, 15, and 20, respectively.

Table 7. Analysis of proposed method for different classifiers with CLAR-LASSO and $k = 5$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	86.66	87.56	88.7	89.66	90.7	0.6631	0.6749	0.8785
Decision Tree	87.25	88.49	89.3	90.55	91.44	0.6569	0.6698	0.8718
Random Forest	88.79	89.79	90.79	91.79	92.79	0.6501	0.6636	0.8659
XGB	89.69	90.75	91.65	92.72	93.4	0.6431	0.6561	0.8581
SVM	88.99	90	90.69	92.05	92.62	0.6373	0.65	0.8522
ANN	88.19	89.05	89.83	91.07	91.65	0.6299	0.6449	0.8465
RNN sigmoid	89.02	89.92	90.46	92.03	92.15	0.6219	0.6372	0.8401
RNN TanH	89.52	90.69	91.37	92.89	93.11	0.6157	0.6313	0.8331
RNN Relu	90.3	91.31	92.26	93.77	94.11	0.6085	0.6257	0.8272
RNN Softmax	91.15	92.2	92.97	94.45	94.65	0.6012	0.618	0.8219
RNN Swish	92.11	92.98	93.9	95.19	95.53	0.595	0.6123	0.8139
LSTM Sigmoid	92.67	93.64	94.8	95.7	96.26	0.5899	0.6054	0.806
LSTM TanH	93.29	94.39	95.45	96.43	96.79	0.5839	0.5981	0.7982
LSTN Relu	93.92	95.11	96.19	97.4	97.65	0.5779	0.5919	0.7915
LSTM Softmax	94.78	95.63	96.71	97.9	98.38	0.5712	0.5842	0.7838
SAFLSTM	95.67	96.32	97.26	98.77	99.17	0.5635	0.5783	0.7774

This analysis shows that SAFLSTM with CLAR-LASSO provides a better classification than the other classifiers. Specifically, the SAFLSTM with $k = 5$ provides better classification than the k-fold values of 10, 15, and 20.

Table 8. Analysis of proposed method for different classifiers with CLAR-LASSO and $k = 10$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	86	87	88	89	90	0.67	0.68	0.886
Decision Tree	86.93	87.8	88.92	89.97	90.68	0.6647	0.6746	0.8781
Random Forest	88.79	89.79	90.79	91.79	92.79	0.6584	0.6669	0.8713
XGB	89.39	90.66	91.48	92.34	93.68	0.6523	0.6619	0.8663
SVM	88.58	90	90.92	91.37	92.77	0.6471	0.6546	0.8602
ANN	87.72	89.14	90.06	90.51	91.85	0.6391	0.6469	0.8522
RNN sigmoid	88.62	90.11	90.59	91.13	92.84	0.6314	0.6395	0.8453
RNN TanH	89.19	90.79	91.54	91.66	93.48	0.6245	0.6316	0.8381
RNN Relu	89.7	91.59	92.12	92.32	94.43	0.6165	0.6264	0.8311
RNN Softmax	90.33	92.1	92.88	92.82	94.98	0.6113	0.6186	0.8247
RNN Swish	91.14	92.68	93.43	93.57	95.56	0.6033	0.612	0.8193
LSTM Sigmoid	91.83	93.22	94.09	94.41	96.49	0.5978	0.6049	0.8113
LSTM TanH	92.73	93.76	94.99	95.08	97.24	0.5922	0.5992	0.8054
LSTN Relu	93.68	94.64	95.93	96.05	98.09	0.5855	0.5924	0.7996
LSTM Softmax	94.48	95.25	96.55	96.68	98.71	0.5778	0.5852	0.7943
SAFLSTM	95.47	96.08	97.05	97.5	99.48	0.5715	0.578	0.7878

Table 9. Analysis of proposed method for different classifiers with CLAR-LASSO and $k = 15$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	86.71	87.73	88.84	89.79	90.97	0.6632	0.6729	0.8785
Decision Tree	87.23	88.35	89.54	90.68	91.81	0.6557	0.6668	0.8731
Random Forest	88.79	89.79	90.79	91.79	92.79	0.6491	0.6593	0.8653
XGB	89.46	90.45	91.43	92.67	93.71	0.6435	0.6523	0.8599
SVM	88.72	89.55	90.63	92.14	92.75	0.6382	0.6461	0.8548
ANN	87.75	88.73	89.75	91.24	91.94	0.6318	0.6388	0.8488
RNN sigmoid	88.43	89.62	90.58	92.02	92.88	0.6248	0.6324	0.8409
RNN TanH	89.1	90.49	91.23	92.88	93.65	0.6186	0.6261	0.8344
RNN Relu	89.63	91.21	91.76	93.74	94.42	0.6125	0.6205	0.8286
RNN Softmax	90.59	92.06	92.74	94.32	95.32	0.6054	0.6134	0.821
RNN Swish	91.48	92.92	93.37	94.82	95.95	0.5984	0.6073	0.8154
LSTM Sigmoid	92.18	93.55	94.25	95.47	96.82	0.5923	0.6016	0.8094
LSTM TanH	92.78	94.48	95.02	96.19	97.71	0.5863	0.5951	0.8036
LSTN Relu	93.69	95.41	95.83	96.99	98.26	0.5783	0.5879	0.7981
LSTM Softmax	94.49	96.39	96.72	97.96	98.8	0.5729	0.5815	0.7907
SAFLSTM	95.16	97.22	97.34	98.46	99.33	0.5656	0.5738	0.7842

The graph of classification parameters for LSTM's different activation functions using $k = 5$ with CLAR-LASSO is shown in Fig. 8. The accuracy of SAFLSTM for $k = 5$ is 95.67% which is high when compared to conventional classifiers. The following strategies are used to achieve better performances such as, 1) The CLAR-LASSO accomplishes the even distribution of a group of weights that helps to select optimal features, and 2) the non-monotonicity, unsaturation, and smooth features of SAF used in LSTM are used to enhance the classification.

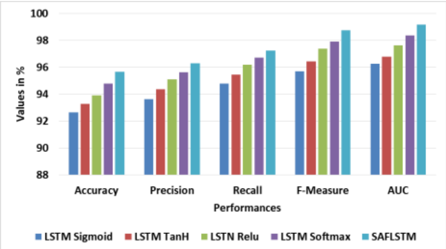

Fig. 8. Classification Measure Graph of LSTM's different activation functions for $k = 5$ with CLAR-LASSO

Table 10. Analysis of proposed method for different classifiers with CLAR-LASSO and $k = 20$

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	AUC (%)	MAE	MSE	RMSE
Logistic Regression	86.51	87.84	88.76	90	90.72	0.6649	0.6722	0.8789
Decision Tree	87.34	88.6	89.53	90.87	91.65	0.6576	0.6651	0.8711
Random Forest	88.79	89.79	90.79	91.79	92.79	0.6526	0.659	0.8633
XGB	89.66	90.74	91.74	92.79	93.34	0.6448	0.6529	0.8576
SVM	88.84	89.98	91.01	91.9	92.44	0.6376	0.6471	0.8513
ANN	88.02	89.02	90.01	91.01	91.5	0.6298	0.6399	0.8437
RNN sigmoid	88.66	89.76	90.88	91.77	92.04	0.6221	0.6319	0.8357
RNN TanH	89.49	90.73	91.61	92.58	92.8	0.6148	0.6252	0.828
RNN Relu	90.18	91.47	92.13	93.11	93.3	0.6097	0.6192	0.8221
RNN Softmax	91.05	92.32	92.86	93.9	94.11	0.6043	0.613	0.8144
RNN Swish	91.57	93.12	93.83	94.88	94.72	0.5981	0.6067	0.8092
LSTM Sigmoid	92.46	93.81	94.83	95.54	95.46	0.5902	0.6011	0.8024
LSTM TanH	93.4	94.43	95.78	96.2	96.35	0.5832	0.5953	0.7956
LSTN Relu	93.98	95.05	96.51	97.19	97.33	0.5758	0.5886	0.7885
LSTM Softmax	94.65	95.8	97.22	98.1	97.95	0.5707	0.5814	0.7826
SAFLSTM	95.36	96.43	98.09	98.84	98.51	0.5649	0.5737	0.7754

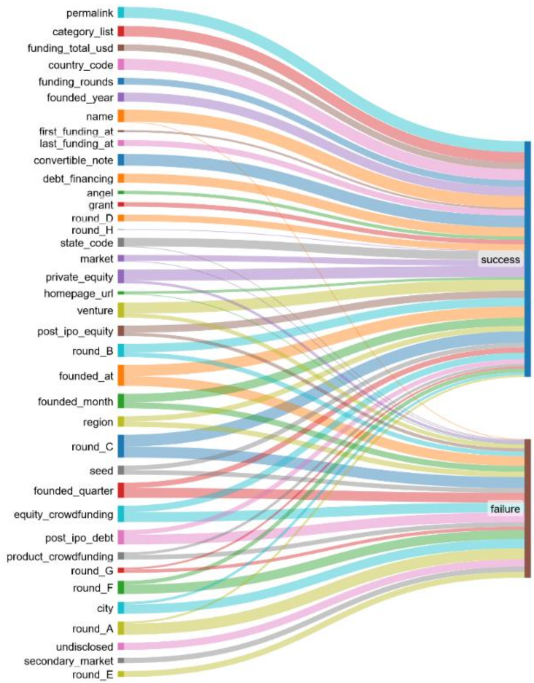

Fig. 9. Analysis of survival rate prediction of startup

The analysis of the survival rate prediction of startups is shown in Fig. 9. This analysis shows 45,521 companies have survived and 2603 companies are failure from 48,124 companies. The survival and failure rates of the companies are found by using the following factors such as age_first_funding, venture, funding_rounds, founded_month, seed, round_A, market_Other, round_B, angel, market_Software, debt_financing, market_Curated Web and round_C. These aforementioned factors are potentially used in a SAFLSTM to assess or predict startup success. Each of these factors provides valuable information to predict the success and its individual contributions. Here, the acquired startups are also considered in the success ratio.

This analysis shows that 94.59% of companies survived and 5.41% of companies were failures which is predicted by using the SAFLSTM with CLAR-LASSO.

The accuracy identified in this work is used to define the prediction of the success and failure of the startup companies. The prediction of startup survival rate using CLAR-LASSO provides the following advantages to the startup companies.

- **Market trends and demand prediction.** The SAFLSTM used in this research is trained according to the external data sources i.e., customer behavior, market trends, and industry data. The aforementioned factors are used to come up with better-informed decisions related to customer acquisition, product development, and marketing strategies.
- **Financial forecasting.** The short-term and long-term financial forecasts are obtained by SAFLSTM based on the training using historical financial data that comprises profits, revenue, and expenses. These factors are used by startups for planning budgets, making financial decisions, and fixing realistic financial goals.
- **Investor relations.** Investors frequently seek data-driven perceptions before determining to invest in a startup. Thus, the performance identification using SAFLSTM provides the possible investors with a clear view of the growth potential of the startup, reduces the uncertainty, and makes it highly attractive for investors.
- **Risk assessment.** The SAFLSTM model can be used by startups for evaluating and avoiding risks.
- **Iterative improvement.** The SAFLSTM is constantly updated with modern information and allows startups to refine their performance predictions along with time. This iterative process helps companies adapt to various market conditions and enhance their decision-making process.

4.2. Comparative analysis

The comparative analysis of CLAR-LASSO with existing research is given in this section. The existing researches considered for comparison are HPT-Logistic regression [18], HPT-SVM [18], HPT-XGBoost [18], and SAFLSTM [19]. Table 11 and Table 12 show the comparison of classification and error measures respectively. This analysis shows that the CLAR-LASSO achieves better performance than the HPT-Logistic regression [18], HPT-SVM [18], HPT-XGBoost [18], and SAFLSTM [19]. An optimal selection of features using CLAR-LASSO is used to improve the classification of startup survival rate.

Table 11. Comparative analysis in terms of classification measures

Methods	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
HPT-Logistic regression [18]	86	67	21	32
HPT-SVM [18]	84	49	31	38
HPT-XGBoost [18]	85	57	34	43
SAFLSTM [19]	71.64	NA	NA	NA
CLAR-LASSO	95.67	96.32	97.26	98.77

Table 12. Comparative analysis in terms of error measures

Methods	MAE	MSE	RMSE
SAFLSTM [19]	0.3829	0.2055	0.4533
CLAR-LASSO	0.5715	0.578	0.7878

5. Conclusion

In this research, the startup survival rate prediction is performed using the CLAR-LASSO and SAFLSTM. Initially, the data acquired from the Crunch base dataset is enhanced by using such as min–max normalization, labeling, and null value preprocessing approaches. Next, the CLAR-LASSO is used to extract the optimal features according to the even distribution of a group of weights. Further, the classification is done by using the SAFLSTM to predict the success rate of startups followed by LIME is used to interpret the predicted classification to the user. The performance evaluation shows that the CLAR-LASSO achieves better performance than the HPT-Logistic regression, HPT-SVM, HPT-XGBoost, and SAFLSTM. The accuracy of the CLAR-LASSO is 95.67% which is high when compared to the HPT-Logistic regression, HPT-SVM, HPT-XGBoost, and SAFLSTM. In the future, hyperparameter tuning will be developed to improve the performances of startup survival rate prediction.

Declarations

Funding: This research received no external funding.

Conflict of Interest: The authors declare that they have no conflict of interest.

Data Availability Statements: The datasets generated and/or analyzed during the current study are available in the [Crunch base dataset] repository:
Crunch base dataset link = https://www.kaggle.com/datasets/arindam235/startup-investments-crunchbase

References

1. Gosztonyi, M., J. Csákné Filep, S. Zsigmond-Heinczinger. Comparison of Applicability of Startup Life Cycle Theories Based on Natural Language Processing. – Eastern European Economics, Vol. **60**, 2022, No 6, pp. 511-539.
2. Savin, I., K. Chukavina, A. Pushkarev. Topic-Based Classification and Identification of Global Trends for Startup Companies. – Small Business Economics, Vol. **60**, 2023, No 2, pp. 659-689.
3. Bangarigadu, K., R. Nunkoo. A Success Versus Failure Prediction Model for Small Firms. – Journal of African Business, 2022, pp. 1-17.
4. Díaz-Santamaría, C., J. Bulchand-Gidumal. Econometric Estimation of the Factors that Influence Startup Success. – Sustainability, Vol. **13**, 2021, No 4, 2242.
5. Pasayat, A. K., A. Mitra, B. Bhowmick. Determination of Essential Features for Predicting Start-Up Success: An Empirical Approach Using Machine Learning. – Technology Analysis & Strategic Management, 2022, pp. 1-19.
6. Kaiser, U., J. M. Kuhn. The Value of Publicly Available, Textual and Non-Textual Information for Startup Performance Prediction. – Journal of Business Venturing Insights, Vol. **14**, 2020, e00179.
7. Archibald, T. W., E. Possani. Investment and Operational Decisions for Start-Up Companies: A Game Theory and Markov Decision Process Approach. – Annals of Operations Research, Vol. **299**, 2021, No 1-2, pp. 317-330.

8. G u p t a, V., L. R u b a l c a b a. Competency-Industry Relatedness (C-IR) Framework for Sustained Business Growth in Startups during and Beyond Pandemic: Myths and Lessons from Publicly Funded Innovative Startups. – Sustainability, Vol. **13**, 2021, No 9, 4632.
9. K i m, H. J., T. S. K i m, S. Y. S o h n. Recommendation of Startups as Technology Cooperation Candidates from the Perspectives of Similarity and Potential: A Deep Learning Approach. – Decision Support Systems, Vol. **130**, 2020, 113229.
10. G y i m a h, P., K. O. A p p i a h, R. N. L u s s i e r. Success Versus Failure Prediction Model for Small Businesses in Ghana. – Journal of African Business, Vol. **21**, 2020, No 2, pp. 215-234.
11. B a i, S., Y. Z h a o. Startup Investment Decision Support: Application of Venture Capital Scorecards Using Machine Learning Approaches. – Systems, Vol. **9**, 2021, No 3, 55.
12. A n t r e t t e r, T., I. B l o h m, D. G r i c h n i k, J. W i n c e n t. Predicting New Venture Survival: A Twitter-Based Machine Learning Approach to Measuring Online Legitimacy. – Journal of Business Venturing Insights, Vol. **11**, 2019, e00109.
13. Z e r h a r i, B., A. L e h c e n, S. M o u l i n e. A New Horizo-Vertical Distributed Feature Selection Approach. – Cybernetics and Information Technologies, Vol. **18**, 2018, No 4, pp. 15-28.
14. A k h i a t, Y., Y. M a n z a l i, M. C h a h h o u, A. Z i n e d i n e. A New Noisy Random Forest Based Method for Feature Selection. – Cybernetics and Information Technologies, Vol. **21**, 2021, No 2, pp. 10-28.
15. A l h e n a w i, E., H. A l a z z a m, R. A l-S a y y e d, O. A b u A l g h a n a m, O. A d w a n. Hybrid Feature Selection Method for Intrusion Detection Systems Based on an Improved Intelligent Water Drop Algorithm. – Cybernetics and Information Technologies, Vol. **22**, 2022, No 4, pp. 73-90.
16. V e g a n z o n e s, D. Corporate Failure Prediction Using Threshold-Based Models. – Journal of Forecasting, Vol. **41**, 2022, No 5, pp. 956-979.
17. R o s s, G., S. D a s, D. S c i r o, H. R a z a. CapitalVX: A Machine Learning Model for Startup Selection and Exit Prediction. – The Journal of Finance and Data Science, Vol. **7**, 2021, pp. 94-114.
18. Ż b i k o w s k i, K., P. A n t o s i u k. A Machine Learning, Bias-Free Approach for Predicting Business Success Using Crunchbase Data. – Information Processing & Management, Vol. **58**, 2021, No 4, 102555.
19. A l l u, R., V. N. R. P a d m a n a b h u n i. Predicting the Success Rate of a Start-Up Using LSTM with a Swish Activation Function. – Journal of Control and Decision, Vol. **9**, 2022, No 3, pp. 355-363.
20. F u e r t e s-C a l l é n, Y., B. C u e l l a r-F e r n á n d e z, C. S e r r a n o-C i n c a. Predicting Startup Survival Using First Years Financial Statements. – Journal of Small Business Management, Vol. **60**, 2022, No 6, pp. 1314-1350.
21. E l h o s e n y, M., N. M e t a w a, G. S z t a n o, I. M. E l-H a s n o n y. Deep Learning-Based Model for Financial Distress Prediction. – Annals of Operations Research, 2022, pp. 1-23.
22. K o u, G., Y. X u, Y. P e n g, F. S h e n, Y. C h e n, K. C h a n g, S. K o u. Bankruptcy Prediction for SMEs Using Transactional Data and Two-Stage Multiobjective Feature Selection. – Decision Support Systems, Vol. **140**, 2021, 113429.
23. Crunch Base Dataset. **https://www.kaggle.com/datasets/arindam235/startup-investments-crunchbase**
24. Z h o u, S., L. Z h o u, M. M a o, H. -M. T a i, Y. W a n. An Optimized Heterogeneous Structure LSTM Network for Electricity Price Forecasting. – IEEE Access, Vol. **7**, 2019, pp. 108161-108173.
25. Z h u, H., H. Z e n g, J. L i u, X. Z h a n g. Logish: A New Nonlinear Nonmonotonic Activation Function for Convolutional Neural Network. – Neurocomputing, Vol. **458**, 2021, pp. 490-499.
26. T o ğ a ç a r, M., N. M u z o ğ l u, B. E r g e n, B. S. B. Y a r m a n, A. M. H a l e f o ğ l u. Detection of COVID-19 Findings by the Local Interpretable Model-Agnostic Explanations Method of Types-Based Activations Extracted from CNNs. – Biomedical Signal Processing and Control, Vol. **71**, 2022, 103128.

Received: 22.06.2023; Second Version: 30.09.2023; Third Verion: 13.10.2023; Accepted: 20.10.2023