BINNING AND IMPROVED DEEP LEARNING FOR CRIME TRENDS PREDICTION

J.Jeyaboopathiraja[1], Dr.G.Maria Priscilla[2]
[1]Assistant Professor, Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore. jeyaboopathi@gmail.com
[2]Professor and Head , Department of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore. mariajerrin76@gmail.com

Abstract: Future crime can be minimized as well as identified using crime prediction. Past data is used in crime prediction and future crime is predicted by analysing data with time and location. Serial criminal cases occur rapidly in present days. With better performance, accurate future crime prediction is a challenging task. For solving Crime detection problem, data mining methods are highly useful. For crime trends prediction, a new framework is proposed in recent work. In this technique, highly influenced features are selected using improved bat optimization (IBAT). At last, for getting highly accurate crime trends and for avoiding further crime, convolution neural network algorithm is used for predicting crime trends. However data set taken for this work may have noisy data values and it may affect the crime trend prediction performance and it does not focused in existing work. For rectifying those issues, for crime trends prediction, this work introduces an improved frame work. In which first input crime data will be pre-processed using missing value imputation, binning and min - max normalization. And then significant features are selected using improved cuckoo search optimization to improve the prediction. Finally, sparse regularization for convolutional neural network (SRCNN) with rectified linear units (ReLU) in hidden layers is introduced for crime trend prediction. For ReLU's inputs, sparseness is introduced. In learning process, ReLU's inputs are pushed into zero. This prevents unnecessary ReLU's output increase. Experimental results demonstrate that proposed model produces better outcome with respect to accuracy, f-measure, recall and precision for Philadelphia, Chicago, and Francisco dataset.

***Keywords*: Crime prediction, criminal cases, Influenced features, Noisy data values, missing value, Cuckoo search optimization and convolutional neural network**

1. Introduction

The world is urbanized rapidly and undergoes largest urban growth wave in history. Urban population is expected to grow to 4.98 billion in 2030 from 2.86 billion in 2000 as per United Nations report. Roughly around 605 of global population will live in cities by 2030.

Huge environmental, economic and social transformations are already brought by this urbanization and in city management issues, challenges like public safety services, policy, water and air quality, traffic, electricity and water resource planning issues [1,2].

There will be an increase in public safety as cities grow. Higher crime rates are there in larger cities when compared with smaller communities. In urban areas, spiking crime rates is a major concern in various countries with urban population's projected growth.

Policymakers and law enforcement authorities will eventually face immense obstacles deploying perennially inadequate services even more quickly to arrest offenders, infiltrate criminal networks and successfully discourage violence by engaging in crime prevention and mitigation programs in metropolitan communities [3,4].

The reliable and effective analysis of the growing amount of crime data, as well as the variety of data sources that could have predictive value linked to crime, is a significant challenge for all law enforcement agencies. This is especially true in big cities, where law enforcement authorities lack the resources and technology necessary to detect valuable trends in massive amounts of time series data.

From this data, useful information are extracted using data mining techniques as presented by this opportunity and it enhances urban policing efficacy. Limited resources better utilization by police department is enabled using this [5,6].

Moreover, available planning tools can be integrated with advanced analytical tools, which enables quick as well as efficient exploration of huge databases by criminal investigators without training as data scientists. Development of effective techniques are concentrated mainly and they are used for preventing crime and makes highly effective investigation.

For crime trends prediction, a new framework is proposed in recent work. In this technique, highly influenced features are selected using improved bat optimization (IBAT). At last, for

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

getting highly accurate crime trends and for avoiding further crime, convolution neural network algorithm is used for predicting crime trends.

However data set taken for this work may have noisy data values and it may affect the crime trend prediction performance and it does not focused in existing work[7].

For rectifying those issues, for crime trends prediction, this work introduces an improved frame work. In which first input crime data will be pre-processed using missing value imputation, binning and min - max normalization. And then significant features are selected using improved cuckoo search optimization to improve the prediction. At last, in hidden layers, for convolutional neural network (CNN) with rectified linear units (ReLU), sparse regularization is introduced. For ReLU's inputs, sparseness is introduced. In learning process, ReLU's inputs are pushed into zero. This prevents unnecessary ReLU's output increase.

2. Litrature Review

Various crime prediction techniques are reviewed in this section.

The BDA is applied to criminal data in the technique proposed by Feng et al [2019]. For trends prediction and visualization, conducted exploratory data analysis. Various deep learning and data mining methods are used. In Philadelphia, Chicago and San Francisco dataset, criminal data are used for discovering some important patterns and facts using statistical visualization and analysis.

When compared with neural network models, Keras stateful LSTM and Prophet model exhibits a better performance as shown in predictive results. Training data's optimum size is three years. For law enforcement organizations and police departments benefits are provided by these promising outcomes for better understanding of crime issues and for providing insight which can be used for enabling track activities, incident likelihood prediction, effective resource deployment and decision making process optimization.

Over the Baltimore (USA) city, for predicting crime events presence, a Convolutional Neural Network (CNN) together with a Long-Short Term Memory (LSTM) network termed as CLSTM-NN is proposed by Esquivel et al [2020]. On two events types called "larceny" and "street robbery", implemented this model. In past data, available temporal and spatial correlation are considered in proposed procedure for enhancing future prediction.

Using some standard evaluation metrics called AUC-PR, AUC-ROC and accuracy, under controlled plausible scenarios, proposed neural network's performance is predicted.

Across the city, for estimating missing dynamic features, matrix factorization based technique is developed by Rumi et al [2018]. Across various types, correlation is maintained by computed dynamic features with crime occurrence. In various time intervals, proposed technique is evaluated. Dynamic feature inclusion across various crime events types enhances crime prediction performance significantly as verified using experimentation results.

Two techniques namely, modified artificial neural network model and modified autoregressive integrated moving average model is proposed by Jha, et al [2020]. For crime forecasting, this work mainly focuses to make a comparison between univariate time series model's applicability against variate time series model. Testing and training are done using more than two million datasets.

Then generated a rigorous experimental analysis and results. Better forecasting results are yielded using a variate time series model when compared with available techniques predicted values as concluded in this paper. Available techniques are outperformed by proposed techniques as shown results.

On emotions, images having spatiotemporal labels are converted as quantifiable data in the system proposed by Li, et al [2020] and these data are applied for predicting crime. At first, three classes of human emotions are divided namely, positive, neural and negative. Then, for quantifying portrait data, employed facial expression recognition (FER).

Acquired emotion features are imported to crime prediction model which enhances explanatory power of model. At last, on six typical crimes, comparison is made with this technique with kernel density estimation (KDE). Between crimes and emotions, interactions are revealed by introducing emotion data as shown in results and crime prediction performance are enhanced.

Using Naive Bayes classifier, a solution for crime prediction problem is introduced by Kumar and Nagpal [2019]. In this technique, similar crime incidents history is provided with incident-level crime data for predicting specific crime incident's most likely criminal.

A crime dataset is formed using incident-level crime data, which includes criminal ID, crime type, location and incident date. Acquaintances are crime parameters or attributes. Using data learning, for crime prediction problem, tested the proposed system and better results are provided by proposed system as indicated in experimental results and crime patterns and potential solutions are also computed using this.

A Modified k-means clustering method is proposed by Vidyavathi and Neha [2018]. For identifying different crime trends or patterns, on fictitious crime data, this method is applied. From various crime patterns analysis, various predictions are made.

3. Proposed methodology

This section stages proposed model for handwritten text recognition. In which firstly Missing values are imputated

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

using Predictive Mean Matching and binning, Normalization using min max normalization will be performed. And then secondly significant Features are Selected using Improved Cuckoo Search Optimization. Thirdly crime trends are predicted using sparse regularization based convolutional neural network. Proposed work's overall architecture is shown in figure 1.

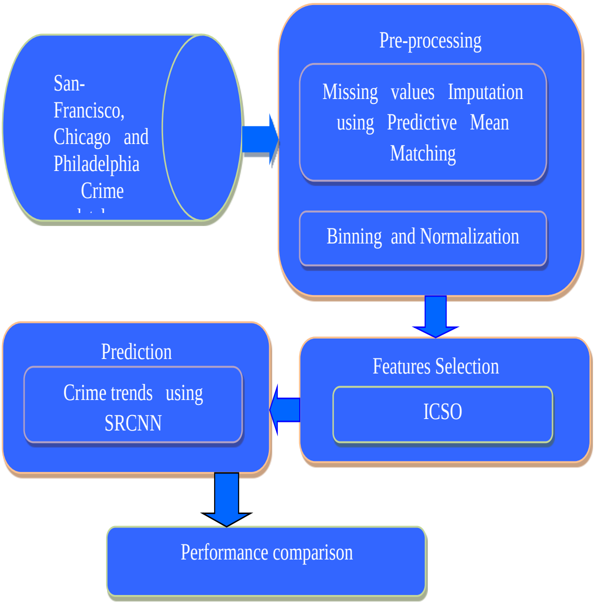

Figure: 1.Overall architecture of the proposed work

3.1. Input

For analysis, three publicly available datasets are used in this work. 3 US cities are covered using this datasets. They are, Philadelphia, Chicago and San-Francisco. There are 2,142,685 crime incidents data are available in San-Francisco crime dataset and this data is collected between 01.01.2003 to 11.08.2017. There exist 5,541,398 records in Chicago dataset and are collected during 2003 to 2017. There are 2,371,416 crime incidents in Philadelphia dataset which are captured between 01.01.2006 to 31.12.2017. In these three datasets, word cloud is used for representing crime text.

3.2. Missing value imputation using Predictive Mean Matching method

Crime data which is used in this work having missing values and it will affect the prediction result. To avoid this problem in this work used Predictive Mean Matching method based missing value imputation.

A state-of-the-art imputation method is Predictive mean matching (*PMM*), where according to instance $Y_{obs,I}$'s observed part, incomplete instance (recipient) Y_i's missing values Y_{miss}, I is computed, which are used for computing nearest donor or instance via distance function. This is find as missing variable's expected value s conditioned on observed covariates rather than finding directly on covariates values. Working of PMM is as follows,

1) Over attribute values, multivariate Gaussian distribution's parameters θ is estimated using *EM* algorithm via all existing data.

2) According to estimated parameters θ, for instance *Yi*'s missing part $Y_{miss,i}$ conditioned on the observed part $Y_{obs,I}$'s conditional expected value is computed.

$$\hat{\mu}_i = E\left(Y_{miss,j} / Y_{obs,i}, \theta\right) \quad (1)$$

3) Every recipient *Yi* is matched with another instance (possible donor) Yj=argminj d(i,j) which has a nearest predictive mean with respect to Mahalanobis distance, defined using residual covariance matrix from missing items regression on observed ones.

$$d(i,j) = \sqrt{\left(\hat{\mu}_i - \hat{\mu}_j\right)^T S^{-1}_{Y_{miss,j}/Y_{obs,i}} \left(\hat{\mu}_i - \hat{\mu}_j\right)} \quad (2)$$

where, empirical covariance matrix is given by *S*.

4) In recipient, every missing value is imputed with its respective values from its nearest donor.

3.3. **Data Binning using bin median points**

After missing value imputation need to perform data smoothing to avoid noisy data. This work uses bin median points. Data values are smoothened using binning techniques where its neighbour value or value around it are referred. Equal buckets or bins count are formed by splitting data in this technique and using bin median points, they are smoothened.

➢

et us have sorted data for crime incident

➢

ow splitting the data into equal groups.

➢

moothing by bin

➢

moothing by bin boundaries

3.3. Data normalization using Min -max Normalization

Input crime data might have variations which lead to provide inaccurate results to avoid this issues it is required to normalize the data. This work uses Min -max Normalization model. In normalization process, mathematical function is used to convert numerical values as new range. In this proposed research work, min-max normalization is utilized to normalize crime dataset. Min-max normalization is one of the most common ways to normalize data. The values in the dataset are normalized within the given range. From dataset, maximum and minimum values and every value are replaced using following expression [15, 16, 17].

$$v' = \frac{v - \min_A}{\max_A - \min_A}\left(new\_max_A - new\_min_A\right) + new\_min_A$$

(3)

Where,

A represents Attribute data,

A's minimum absolute value is represented as Min (A) and maximum absolute value is represented as Max (A).

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

In data, every entry's new value is represented as v'
In data, every entry's old value is represented as v
Boundary range's maximum value is represented as New_ max (A) and its minimum value is represented as new_ min (A).

3.4. Feature selection using improved cuckoo search optimization

After normalize the data, it required to select important features from that database because it might have more number of features and it consume time for computation to get crime trend prediction. To avoid this issues this work uses improved cuckoo search optimization for feature selection.

A novel meta-heuristic algorithm called Cuckoo Search (CS) algorithm is proposed. Some cuckoo species obligate brood parasitism is inspired in this algorithm, where in other host bird's nest, eggs are laid by some cuckoo species. Direct conflict is engaged by some host nest. A host bird throws alien ages away or its simple abandon its nest and new nest is build elsewhere, if it discovers an egg which is not its own.

There evolves some other species where female parasitic cuckoos are very specialized often in mimicry in colour and in few selected host species egg pattern.

Eggs abandoned probability is minimized using this and its re-productivity is enhanced using this. On the other side, Levy flights typical characteristics are shown by various insects and animals flight behaviour as illustrated in various studies. The CS algorithm is proposed here by considering those flight and breeding behaviour [18-21].

Next three idealized rules are followed in CS algorithm:

1) At a time, one egg is laid by every cuckoo and in randomly selected nest, its egg are dumped.

2) To next generations, best nests having high eggs quality are carried over.

3) There will be fixed available host count and with probability $p_a \in [0, 1]$, host bird discovers egg laid by cuckoo. In this condition, host bird can either throw the egg away or abandon nest and completely new nest is build.

The n nest's fraction p_a are replaced using new nests having new random solution and it can be used for approximating last assumption as suggested by authors.

Solution's fitness is assumed as objective function called classification accuracy for feature selection problem. In a nest, solution is represented using every egg in this algorithm and ne solution is represented using cuckoo egg. In nest, not-so good solution are replaced using new as well as potentially better cuckoos called solutions. According to these three rules, Cuckoo Search (CS)'s basic steps are summarized as following pseudo code,

In above mentioned three idealized rules, new bird's nest location path search is expressed as,

$$h_i^{(t+1)} = h + \alpha \oplus Levy\ (\lambda);\ i = 1,2, \ldots \ldots \ldots, n \quad (4)$$

In t generation, i^{th} bird's nest position is represented as $h_i^{(t+1)}$, step size control is represented as α, $\alpha > 0$, in general, $\alpha = 1$ Levy(λ) is Levi's random search path, its expression is given as,

$$Levy\ (\lambda) = t^{-\lambda}; 1 < \lambda < 3 \quad (5)$$

Step size is a limitation of traditional cuckoo search and at initialization, fixed value is set to CS algorithm's discovery probability and in subsequent iterations, this value will not change. With large step size, convergence is made easy and search accuracy is minimized and with small step length, search speed is minimized and it may be trapped with local optimum easily. An improved CS algorithm is introduced in this work for overcoming those issues.

Iterations count are combined with step size in ICO algorithm and at iteration's beginning, longer step size is set and step size is minimized as iteration progress. In iteration, algorithm has a huge step. Iterative speed and global optimization is achieved using and in algorithm iteration's latter part, local optimization is achieved using with small step size and it has enhances search accuracy. Enhanced formula is expressed as,

$$\propto_i = a_{max} * \frac{1}{\left(\frac{a_{max}}{a_{amin}}\right)^{\frac{t}{T}}} * ran_i \ * 0.01 \quad (6)$$

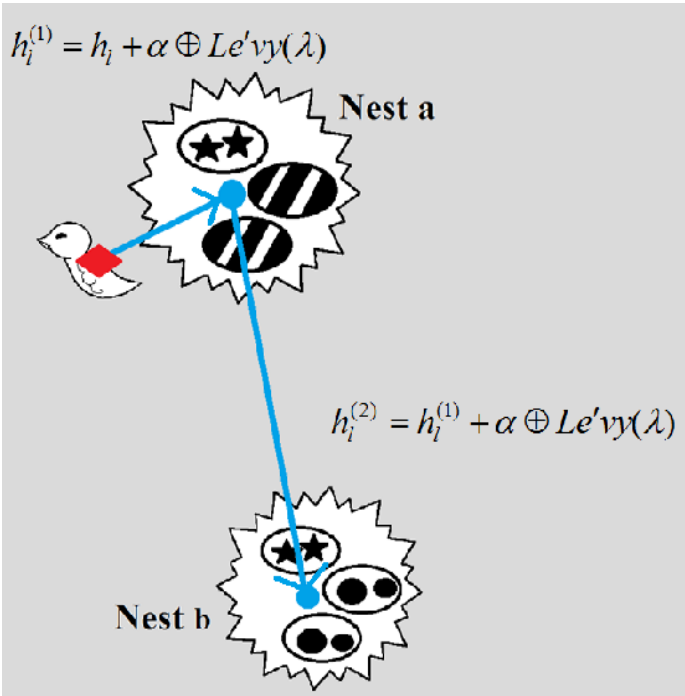

Figure: 2.Cuckoo search working procedure

Where, maximum step size is represented as a_{max}, minimum step size is represented as a_{amin}, total iterations count is given by T. Current iteration number is represented as T, in dataset, i^{th} dimension's scope is given by ran_i.

Algorithm for ICSO

INPUT: San-Francisco, Chicago, and Philadelphia crime database

OUTPUT: Optimal features

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

1: Initial population of N host nest x_i ∀i, i = 1, ..., n is generated
2: **while** t < Max Generation or (stop criterion) **do**
3: Get a cuckoo randomly using Levy flights and e its fitness Fi is computed
4: Among N, nest j is randomly selected
5: **if** $F_i > F_j$ **then**
6: Replace j using new solution.
7: **end if**
8: Abandoned a fraction (p_a) of worse nest and built new ones.
9: Best solutions are maintained (or nest having quality solutions).
10: Solutions are ranked and current best is computed
11: **end while**

Chicago Selected data	Philadeplia Selected data	San francisco Selected data
1	1	1
2	4	2
3	7	3
4	8	4
5	10	5
8	13	7
13		8
14		9
20		10
		17
		30
		33

Table 1.Feature selection results

3.5. Crime trends prediction using sparse regularization based convolutional neural network

After feature selection it needs to predict the crime trends. This work uses sparse regularization based convolutional neural network for crime trend prediction.

Distinct layers stack is used for forming deep CNN architecture. Class scores are formed by transforming input image using this. In general, four types of layers are used, namely, classification, fully connected, pooling and convolution layer. Various pooling and convolution layers are repeated at first and there will be fully connected layer followed by classification layer.

CNN's building block is convolution layer [22,23,24]. In convolution layer, every neuron has small receptive field in input image and output is computed using receptive field's convolution with linear filter. If k-th feature map's weight value is represented as wk and bias value is represented as b_k, then filter output is h_k is expressed as,

$$h_k = w_k^T x + b_k \quad (7)$$

Where, neuron input is represented as x. Then unit $f(h_k)$'s output is computed as,

$$f(h_k) = \sigma(h_k) \quad (8)$$

Where, neuron's activation function is represented as $\boldsymbol{\sigma}$. In input layer, if there exist various channels, filtered outputs sum is used as neuron's activity. Feature maps down-sampling is performed in pooling layer. For instance, non-overlapping rectangle sub-regions set is formed by partitioning feature map and output corresponds to maximum value is every sub-region. This is termed as max pooling. Invariance against input image's small shift is made using this pooling layer. Representation size is minimized using this pooling layer. In network, computation and parameters count are minimized because of this. In general, in a deep CNN architecture, in-between successive convolution layers, pooling layer is inserted periodically.

Various pooling techniques are proposed in addition with max pooling. Such pooling technique includes L2-norm pooling and average pooling. Historically, average pooling is applied very often, but in recent days, due to better practical working, max pooling is applied in various applications. MLP is formed using classification and fully connected layers.

Using fully connected layers, integrated the low-level image features after various pooling and convolution layers pair. In previous layer, in fully connected layer, neurons have full connections to all output. This connection is similar to standard multi-layer Perceptron.

In network, last layer used in general is classification layer and final decisions are made using this layer. Between true and predicted labels, deviation penalization by network training is specified using this layer. In general, in classification layer, soft-max function is used.

The ReLU is given with unnecessary negative inputs by traditional CNN, which is a major disadvantage of it. Sparse regularized CNN is introduced in this work for rectifying the same.

In model, for preventing unnecessary parameters increase, sparseness is introduced generally. For instance, in weight decay, weight sparseness is introduced. Training samples over fitting can be prevented using this and trained model's generalization ability is also enhanced using this.

Likewise, ReLU's inputs is induced with sparseness called linear filter output in this work. Unnecessary increase in ReLU's output is prevented while minimizing ReLU's unnecessary negative inputs by introducing sparseness in ReLU's input. Batch normalization effect can be realized using this proposed technique while enhancing generalization ability.

For the mini-batch samples in a layer, assume neuron input as $\{x_i | i = 1, \ldots, m\}$, where, in mini-batch, samples count is represented as m. For mini-batch samples in layer, neuron input is defined as,

$$\mu_x = \frac{1}{m}\sum_{i=1}^{m} x_i \quad (9)$$

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

For the mini-batch samples in layer, neuron input's variance is defined as follows,

$$\sigma_x^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \mu_\beta\right)^2 \qquad (10)$$

Then can normalize the inputs by

$$\hat{x}_i = \frac{x_i - \mu_x}{\sqrt{\sigma_x^2 + \epsilon}} \qquad (11)$$

In statistics, this process is termed as standardization and normalized values $\hat{x}_i$'s mean value becomes 0 and variation becomes 1. Then this normalized input will be given to the ReLU.

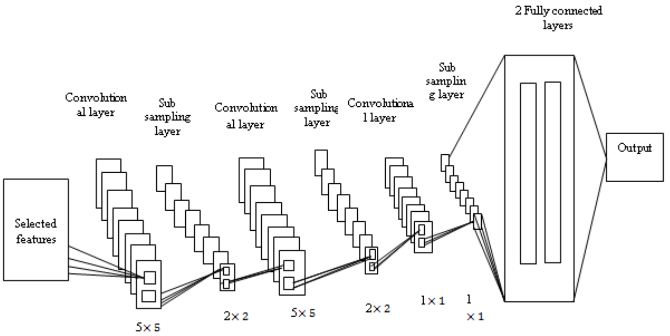

Figure: 3. Convolutional neural network

Figure: 4. Classification result

Table: 2.Performance comparison results

Different methods	Datasets	Metrics			
		Precision	**Accuracy**	**Recall**	**F - Measure**
NN	Chicago	99.12	84	98.90	99.44
	San-Francisco	25	88.29	60	35.29
	Philadelphia	50	83.33	33.33	40
LSTM	Chicago	99.66	84.28	98.89	99.28
	San-Francisco	21.42	88.94	60	31.57
	Philadelphia	25	80.30	38.88	40
CNN	Chicago	99.54	87	98.73	99.36
	San-Francisco	95.66	89.19	99.82	97.65
	Philadelphia	99.14	85.50	98.97	99.48
SRCNN	Chicago	87.11	99.74	99.16	99.71
	San-Francisco	89.46	96.02	99.95	97.99
	Philadelphia	99.44	85.56	99.15	99.52

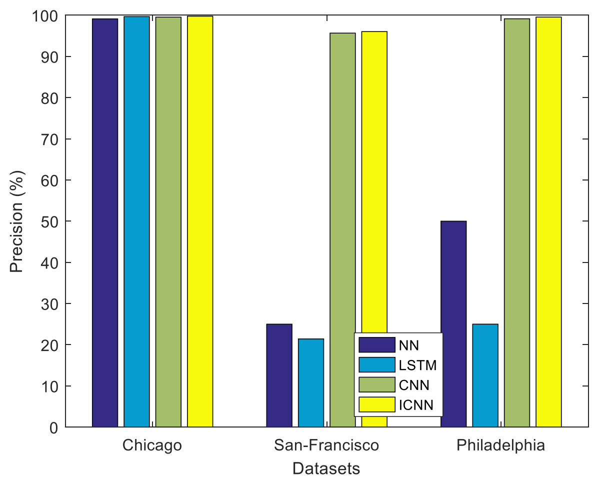

Figure .5: Precision results Comparison of Various Classifiers

4. Results And Discussion

Detailed demonstration about proposed model's experimentation results are presented in this section. Using MATLAB 2013b, implemented this proposed model. With respect to performance evaluation metrics like f-measure, accuracy, recall and precision, available CNN, LSTM and NN models are compared with proposed SRCNN model. Three publicly available crime datasets which covers 3 US cities are used in experimentation. In these three dataset, word cloud is used for representing crime text.1

Precision performance metric comparison among proposed SRCNN and existing CNN, LSTM and NN methods are illustrated in figure 5. For standardizing input values, min max normalization is used in this proposed technique. Various methods are represented in X-axis of above graph and precision values are plotted in Y axis. A precision value of 96% is produced by newly introduce SRCNN model for Chicago dataset as observed in experimentation results, whereas, 99.91% is produced by CNN, 99.66% is produced by LSTM and 99.16% is produced by NN techniques.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

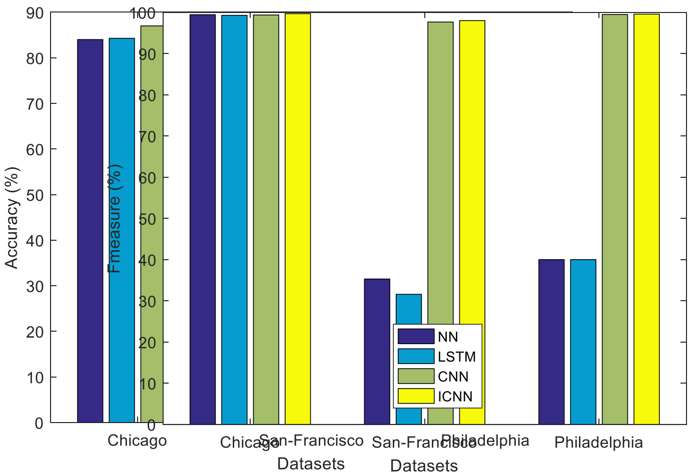

Figure: 6. Accuracy results vs. classification methods

Accuracy performance metric comparison among proposed SRCNN and existing CNN, LSTM and NN methods are illustrated in figure 6. For significant features, probability function is used in proposed SRCNN technique, which enhances accuracy value. Various methods are represented in X-axis of above graph and accuracy values are plotted in Y axis. An accuracy value of 88% is produced by newly introduce SRCNN model for Chicago dataset as observed in experimentation results, whereas, 87% is produced by CNN, 84.28% is produced by LSTM and 84% is produced by NN techniques.

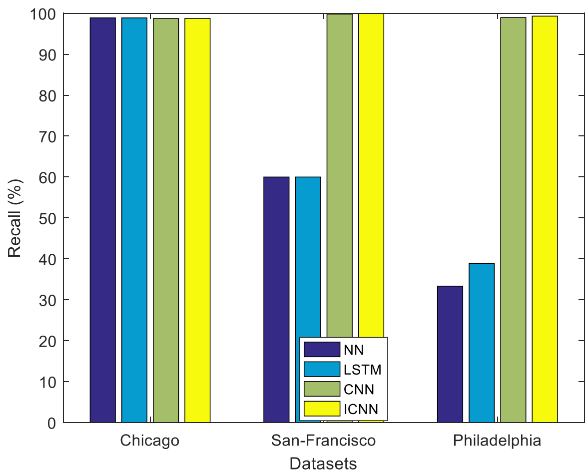

Figure: 7. Recall results vs. classification methods

Recall performance metric comparison among proposed SRCNN and existing CNN, LSTM and NN methods are illustrated in figure 7. A recall value of 98.73% is produced by newly introduce SRCNN model for Philadelphia dataset as observed in experimentation results, whereas, 38.88% is produced by CNN and 33.33% is produced by LSTM techniques.

Figure: 8. F -Measure results vs. classification methods

F-measure performance metric comparison among proposed SRCNN and existing CNN, LSTM and NN methods are illustrated in figure 8. For learning features, more convolution function is used in CNN, which enhances f-measure value. A F-measure value of 99.36% is produced by newly introduce SRCNN model for Chicago dataset as observed in experimentation results, whereas, 99.28% is produced by CNN and 99.44% is produced by LSTM techniques.

5. Conclusion And Future Work

In our society, important role is played by cities due to population growth and continuous urbanization. However, accidents and violent crimes are also getting developed with this growth of population and urbanization. For tackling these problems, much effort is devoted by safety institutions, analysts and sociologists for mining potential factors and patterns. However, processing large volume of available data is a challenging task. For crime trends prediction, an improved framework is provided in this work.

In which first pre-processing of crime data is performed using predictive mean matching based missing value imputation, binning and min - max normalization. Then feature selection computed using improved cuckoo search optimization.

Finally, sparse regularization for convolutional neural network (SRCNN) with rectified linear units (ReLU) in hidden layers is introduced for crime trend prediction. Experimental results demonstrate that proposed model produces better outcome with respect to accuracy, f-measure, recall and precision. However deep learning has more computational complexities which leads to other methods usage for forecasting in future.

References

[1] Feng, M., Zheng, J., Han, Y., Ren, J. and Liu, Q., 2018, July. Big data analytics and mining for crime data analysis, visualization and prediction. In International Conference on Brain Inspired Cognitive Systems (pp. 605-614). Springer, Cham.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

[2] Shamsuddin, N.H.M., Ali, N.A. and Alwee, R., 2017, May. An overview on crime prediction methods. In 2017 6th ICT International Student Project Conference (ICT-ISPC) (pp. 1-5). IEEE.

[3] Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y. and Liu, Q., 2019. Big data analytics and mining for effective visualization and trends forecasting of crime data. IEEE Access, 7, pp.106111-106123.

[4] Chauhan, C. and Sehgal, S., 2017, May. A review: crime analysis using data mining techniques and algorithms. In 2017 International Conference on Computing, Communication and Automation (ICCCA) (pp. 21-25). IEEE.

[5] Chen, X., Cho, Y. and Jang, S.Y., 2015, April. Crime prediction using Twitter sentiment and weather. In 2015 Systems and Information Engineering Design Symposium (pp. 63-68). IEEE.

[6] Catlett, C., Cesario, E., Talia, D. and Vinci, A., 2019. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. Pervasive and Mobile Computing, 53, pp.62-74.

[7] Dash, S.K., Safro, I. and Srinivasamurthy, R.S., 2018, December. Spatio-temporal prediction of crimes using network analytic approach. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 1912-1917). IEEE.

[8] Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y. and Liu, Q., 2019. Big data analytics and mining for effective visualization and trends forecasting of crime data. IEEE Access, 7, pp.106111-106123.

[9] Esquivel, N., Nicolis, O., Peralta, B. and Mateu, J., 2020. Spatio-temporal prediction of Baltimore crime events using CLSTM neural networks. IEEE Access.

[10] Rumi, S.K., Deng, K. and Salim, F.D., 2018. Crime event prediction with dynamic features. EPJ Data Science, 7(1), p.43.

[11] Jha, S., Yang, E., Almagrabi, A.O., Bashir, A.K. and Joshi, G.P., 2020. Comparative analysis of time series model and machine testing systems for crime forecasting. NEURAL COMPUTING & APPLICATIONS.

[12] Li, Z., Zhang, T., Jing, X. and Wang, Y., 2020. Facial expression-based analysis on emotion correlations, hotspots, and potential occurrence of urban crimes. Alexandria Engineering Journal.

[13] Kumar, R. and Nagpal, B., 2019. Analysis and prediction of crime patterns using big data. International Journal of Information Technology, 11(4), pp.799-805.

[14] Vidyavathi, B.M. and Neha, D., 2018. Prediction of Crime Trends Using Mk-MC Technique. In Data Engineering and Intelligent Computing (pp. 417-426). Springer, Singapore.

[15] Shehab, M., Khader, A.T. and Al-Betar, M.A., 2017. A survey on applications and variants of the cuckoo search algorithm. Applied Soft Computing, 61, pp.1041-1059.

[16] Chitara, D., Niazi, K.R., Swarnkar, A. and Gupta, N., 2018. Cuckoo search optimization algorithm for designing of a multimachine power system stabilizer. IEEE Transactions on Industry Applications, 54(4), pp.3056-3065.

[17] Zhao, J., Liu, S., Zhou, M., Guo, X. and Qi, L., 2018. An improved binary cuckoo search algorithm for solving unit commitment problems: Methodological description. IEEE Access, 6, pp.43535-43545.

[18] Cui, Z., Zhang, M., Wang, H., Cai, X. and Zhang, W., 2019. A hybrid many-objective cuckoo search algorithm. soft computing, 23(21), pp.10681-10697.

[19] Cao, Z., Lin, C., Zhou, M. and Huang, R., 2018. Scheduling semiconductor testing facility by using cuckoo search algorithm with reinforcement learning and surrogate modeling. IEEE Transactions on Automation Science and Engineering, 16(2), pp.825-837.

[20] Biswas, A. and Chandrakasan, A.P., 2018, February. Conv-RAM: An energy-efficient SRAM with embedded convolution computation for low-power CNN-based machine learning applications. In 2018 IEEE International Solid-State Circuits Conference-(ISSCC) (pp. 488-490). IEEE.

[21] Pandey, A. and Wang, D., 2019. A new framework for CNN-based speech enhancement in the time domain. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(7), pp.1179-1188.

[22] Abbas, S.M. and Singh, S.N., 2018, February. Region-based object detection and classification using faster R-CNN. In 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT) (pp. 1-6). IEEE.

[23] Zhang, Q., Wang, X., Cao, R., Wu, Y.N., Shi, F. and Zhu, S.C., 2020. Extracting an explanatory graph to interpret a CNN. IEEE transactions on pattern analysis and machine intelligence.

[24] Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A. and Baik, S.W., 2019. Multi-grade brain tumor classification using deep CNN with extensive

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

d ata augmentation. Journal of computational science, 30, pp.174-182.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731