I.J. Modern Education and Computer Science, **2022, 1, 44-55**
Published Online February 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2022.01.04

Artificial Intelligence in Collaborative Information System

Monika Arora
Apeejay School of Management, Dwarka, New Delhi, India
Email: marora.asm@gmail.com

Indira Bhardwaj
Vivekananda School of Business Studies, VIPS, New Delhi, India
Email: indira@dsb.edu.in

Received: 07 February 2021; Revised: 09 March 2021; Accepted: 15 April 2021; Published: 08 February 2022

Abstract: All organizations have a collaborative information system, which is a shared system between employees and teams in the organisation. All such information systems in organizations need to be flawlessly secure. Securing information systems through the latest technologies like Artificial Intelligence, Deep Learning and Blockchain is one of the latest trends in information sciences. This paper tries to explore them in detail through data on user's login time and time spent on the websites along with user actions. The objective is to develop a model that will be used for authentication of the user. This will allow early detection of frauds so that preventive and remedial actions like blocking access to the user can be initiated well in advance. The dataset used to develop this model is the user log data and technique of logistic regression is used to create the regression model for authentication of the user. Logistic regression-based classification is used on the attributes taken to record and analyze entries recorded on the system leading to identification of a cluster based on normal and suspicious users. The accuracy of logistic regression has been analyzed and implemented to secure the collaborative system. This study will help the researcher to implement the AI in the system. It also discusses its future prospects and the disruptive changes in implementation of Information Systems. Finally, the research considers combining blockchain (BC) and deep learning (DL) with Artificial Intelligence (AI) and discusses the revolutionary changes that would result by rapidly advancing the AI field.

Index Terms: AI, Deep Learning, Blockchain, Information System, Security

1. Introduction

Data security has gained momentum in the last few years given the volume of data that gets generated every second. Collaborative security is an abstract concept, where collaboration between different technologies and their compatibility with each other is a significant determinant of their success. Security is often centrally managed and emerging trends such as Artificial Intelligence (AI), Blockchain (BC) and Deep Learning (DL) are used in collaboration to provide high-end security through technology [1]. Implementation of security systems poses challenges related to complexity and compatibility of technologies, sometimes leading to artificial stupidity. Ethical and Legal aspects related to Privacy/Security/Safety of data while implementing Artificial Intelligence (AI), Blockchain (BC) and Deep Learning (DL) are also an increasing cause of concern.

Artificial Neural Networks (ANN) started in the 1950s with the basic programming of machines and it was by the 1990s that the machines had started using explicit programs to evolve their processes and become agile and smart [37]. By the end of the first decade of the 21st Century the concept of Deep Learning had already gained momentum. A collaborative security system recommends the best use of technology such as AI, BC and DL together. Many secure systems offer the security services such as user authentication, use of anti-virus, anti-malware/spyware, intrusion detection etc. The databases used in the organisation and business information can be extracted from these data stores for decision making concerning customer transaction behavior patterns. The organisation is facing increased competition for different reasons, including the user entrance in various applications such as online shopping, banking systems etc. In various online transactions is based on a wide range of offered products and services to the public. As a consequence, the online industry strives to succeed by putting the topic of rapid and changing customer needs on their agenda.

The objectives of corporate security include increased security that leads to gain trust and have better customer response that will improve customer loyalty. The potential areas of application of data-mining techniques are wide. In services, a wide variety of data analysis solutions are provided by the organistaion using computer service personnel and their Software systems implemented such as ERP and MIS depending on the type of data analysis problems encountered. Examples are observed in comprehensive solutions to address the needs of customer/user ever-evolving business scenarios with different frameworks to enable services to critically analyze and evaluate various factors affecting their business transactions, thereby optimizing and enhancing the overall reporting capabilities of their composite application suites.

There are many applications where the authenticity of the user has to be examined thoroughly and it also helps to identify their best customers, implement and measure strategies to retain them, cross-sell and up-sell to them, and make the most effective use of all available assets and channels. Once the correct user login the system, then there are many applications that will help the industries to trap the customer and offer them various services based on its interest and likings such as customer segmentation, profitability, prospecting and acquisition, affinity and cross-sell, retention and attrition, channel utilization, and risk analyses. Customers are distinguished in terms of their profitability. For any business the organisation can build predictive churn models to select the right customer and retain their best customers by identifying symptoms of dissatisfaction and churning. AI can be bitterly used for the same and BC and Deep learning can be looked forward to for the implementation in organization.

The impact of Artificial Intelligence (AI), Blockchain (BC) and Deep Learning (DL) on physical security as shown in Fig.1. Artificial intelligence is the engineering of making astute machines and programs. Artificial intelligence is a superset of many new techniques such as Machine Learning and Deep Learning [2]. Machine learning is defined as the ability of the machine to grasp processes and thus learn without being explicitly programmed for learning. Deep Learning is based on Artificial Neural Networks.

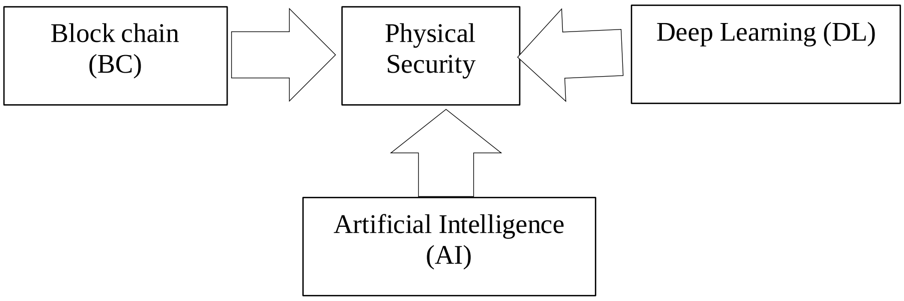

Fig.1. Security Influencer

A new dimension to data security commenced in 2004 in the form of Blockchain. This technology allowed the machines to maintain user audit trails which are one of the most crucial elements of security. The scope of this paper is thus limited to Artificial Intelligence implementation as an element of digital security systems. This disruptive technology together will help in providing secure information systems.

The objectives of the study are

(a) The relationship between blockchain (BC) and deep learning (DL) with Artificial Intelligence (AI)
(b) Mapping of various technologies in collaborative secure information system
(c) To determine the fitness and validation of the model for authentication

This research study is organized in the following way. Section 2 produces the background of the work, which contains the related study, introduction to AI and ANN, and its relation to collaborative security systems. The research design and proposed Model are given in Section 3 and 4. The conclusion drawn is given in Section 5.

2. Background

In 2008, the Blockchain technology was introduced by the name of Bitcoin, by Satoshi Nakamoto.[35] Blockchain is a specific type of database which was initially commenced for Bitcoin. This technology manages data which is received in blocks and assists in tracking transactions in public ledgers between different users. BlockChain technology is used for storing digital information commercially for all applications which use digital assets and virtual records [3, 39]. This information is in encrypted and distributed form which helps to create an extremely secure database.

Technologies like Artificial Neural Networks, work with humans and develop human intelligence. Deep Learning or deep structured learning or hierarchical learning use Machine Learning based methods and are dependent on representations of learning data [4].

Artificial Intelligence can be used in marketing areas where the data and AI create a machine learning system which helps the business in anticipating the customer's reaction and thus helps the business predict and generate profit [9]. The progression of AI to the next level will enable it to provide big data solutions using analytical tools and techniques. Learning can be of three types supervised, semi-supervised or unsupervised. Machine learning is driven by the learning model of human beings. Blockchain is a trusted technology used for the setting up of digital infrastructure which uses the past data to ascertain patterns of behavior that assists in decision making [7]. Human reactions help the system to create machine patterns which then help in the implementation of Artificial Intelligence in systems.

Humans help in creating neural networks to form an artificial network called Artificial Neural Network. This is based on experiences of the human and also helps them build a machine-learning algorithm [3]. The developers of the algorithm are not comfortable in forecasting but they are able to describe the thinking process that has been followed for decision making. The AI-systems operate on the basis of using complex patterns and complex decision trees which provide the results through a black box and match it with human decision. AI uses historical data to give the input to the Blockchain and then AI and Blockchain together work for better decision making. Computers analyze huge amounts of data with the help of technology but the memory used in the machine has no comparison with the brains of the most intelligent people in the world [2]. The data increases exponentially everyday due to increased modes of collection of data from increased number of sources and also in various structures. The algorithms are then developed based on the data and the instructions provided.

Businesses have started using these technologies directly or indirectly to support the ecosystem of business. Artificial intelligence is considered to be a discipline of computer science which uses and imitates the intelligence of humans. The application of AI can be in perception of visuals and recognition of speech that helps in faster and better decision-making. It also helps in translation and creation of machines that use repetitive tasks called machine intelligence. Machines thus are capable of eventually taking care of all repeated human behaviors that have been used to build a system.

Human beings create computer-based systems using past information and enable computers to study the learning pattern from the data stored to use them. The data stores and structures evolve procedures ensuring the correct and accurate use of the data and information.

A blockchain is an immutable, open, append-only transaction log replicated among a network of nodes [15]. One node is connected to the other and assuring through proof- of- work as it gets acknowledged by other participants. Blockchain technology records every transaction in the database. The blockchain is a chain of blocks where any change in any two entities get recorded in a database. Thus, it contains a confirmable recorded history which offers improved transparency and reliability. All the transactions are available for public view over the network. [13]. If there is any communication or transaction between two parties and either of them does not update their records, the corresponding records will not match and it could be considered as fraud. The data thus needs to be updated from moment to moment in time. The blockchain also maintains privacy and transparency.

With evolution of Blockchain, Enterprise Resource planning (ERP) may be outdated and the new technology will define the new market needs and requirements. Many developed countries have already started using the concept of blockchain in their businesses and organization. Blockchain uses distributed databases and the ledger entries are updated in local databases, given the lack of centralized database servers. The real-time update helps the system to minimize the frauds in the businesses. There will be a great shift for e-commerce companies as they shift to Blockchain technology in future. The blockchain can be easily fitted to any business scenario and filtered to any industry. If businesses do not evolve themselves, they will not be able to sustain the challenges posed by other businesses using advanced technologies.

The renowned concept of Deep Learning is defined as the slicing and dicing of data which is studied to create the patterns for machine learning. ML is associated with Deep Learning which creates neural networks which help to create the AI patterns finally translating to machine learning algorithms used for decision making. These networks of Deep Learning study in depth and detail, anything that is extending in layers [14]. The cloud-based implementation uses blockchain and Deep Learning in great extent.

Cloud technology and ontology-based system for security offers a great benefit to the user i.e cost reduction of infrastructure and maximization of profits for an enterprise [10,11]. These compliances help the enterprises to identify and use the best practice for their processes [10]. Model-based ontology attempts to identify the intrusion detection and countermeasures which are immensely useful for administrator, system management, managers and Information Systems [11]. The implementation of ontology and cloud application will give a new dimension to the study considered for security.

3. Literature Review

The relationships between technologies play an important role in improving understanding and use of these technologies. Based on the research paper studied using this technology it was found that AI has a relationship between Deep Learning, Blockchain. The study based on objective-(a) the relationship is as under:

The information layer AI and Blockchain is used in collaboration of various tools. Blockchain makes information available to all the blocks of various businesses. Deep Learning is machine learning applied with AI, which build neural networks where they make their own decisions. Deep works in the concept of layers which emphasize neural networks.

It is based on gaining knowledge based on the experience and experiment. It is also known as hierarchical learning or deep structural learning as shown in the network diagram as shown in Fig.2.

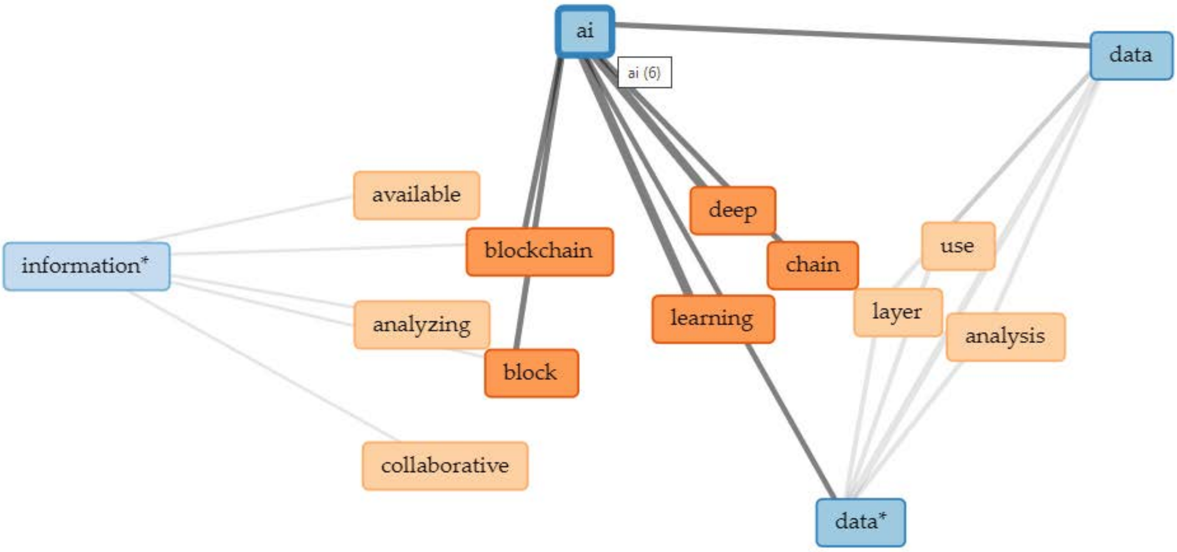

Fig.2. AI relationships with Deep Learning

The relationship between all three has emerged stronger in increasing data security and information security. The collaborations AI, Deep Learning and Blockchain use Content-based Filtering, which matches knowledge of a single user with the knowledge of available items. The above discussed relationships between AI, Deep Learning and Blockchain are thus useful for the enhanced security of the system.

The data is input into the AI system, which uses the transparent layer where non genuine users are not allowed to enter the system. The filtered data input thus flows for Deep Learning, which works on hierarchical study, which further confirms the genuinity of the user. The data transactions verify and acknowledges the same. The blockchain starts using the chains for validating connection and authorization [1]

Learning in depth means the acquisition of knowledge and skills through study and experience that are used in making and building the machine learning patterns. Learning can be in different forms; it can be unsupervised, partially supervised and supervised. The supervised learning can be a subset of Machine Learning (ML) in Artificial Intelligence that is based on defined patterns with labels. These labels create a hierarchy which are useful for creating a hierarchy for the machine pattern which assists in taking a decision.

Collaborative security systems

The use of AI helps in efficiency of data/information of the blockchain technology. All the transactions are verified/authenticated by the miners and AI trains the machine algorithm thereby helping in efficient ways for taking decisions. The decision making is efficient because while using blockchain technology, every piece of information is secure and transparent and recorded over the network. As the data storage and the number of blocks increases every moment the chain becomes heavier. The Blockchain with the use of machine learning algorithms thus optimizes the data storing methods. The integration of various technologies where AI is decentralized and allows parallel computing is Blockchain. Machine learning uses all the computation power to the highest degree and it also analyzes huge amount of data quickly [7].

The different uses of technique in securing business systems have mapped and over the objective- (b) are shown in Table 1.

Table 1. Summary of Related work on used for Collaborative security systems

Technology Used	Description	Reference
Blockchain and Artificial Intelligence	AI is used as intelligence tools and decision-making capabilities for machines and emerging blockchain applications, platforms, and protocols specifically targeting the AI area. It can be used as a potential capacity in the fields of international payment, secure data sharing and marketing, and supply chain management. Artificial Intelligence is used to develop the creation of machines capable of performing tasks that need intelligence. Blockchain, IoT and Artificial Intelligence can influence future cloud computing systems and provide security as well.	[3, 5, 36]

Artificial Intelligence, Deep Learning, Internet of Things, cloud computing, and blockchain, in the new generation of big data and Industry 4.0.	Used for multifaceted systematic analysis of AI in practical applications. Artificial neural networks (ANN) predict new customer behavior from previously observed customer behavior after executing the process of learning from existing data. An evaluation process determined that it performed.	[6,37]
Big data, Artificial Intelligence and IOT	Block IoT Intelligence architecture on device, fog, edge and cloud intelligence use parameters such as accuracy, latency, security and privacy, computational complexity and energy cost in assessing IoT applications.	[4]
Artificial Intelligence, Deep Learning IOT and Blockchain	Blockchain and AI technologies help form a sustainable smart society. Blockchain can be used as security enhancement solutions, blockchain-AI based intelligent, transportation systems are an example. An integrated solution using the edge computing frameworks to build intelligent edge for dynamic, adaptive edge maintenance and management. These techniques are used in applications to solve security and privacy issues.	[8,2,7]

Logistics model and its interpretation

The past papers on logistics and its interpretation in various models [26] show that there are many industries where logistics regression is used as a predictive model, such as real estate, banking, insurance industry, retail industry etc. Please refer to Table 2.

Table 2. Literature Review

Model Discussed	Dataset/Description	Reference
Regression model was used to find out the relationships	Financial ratios and corporate governance	[27]
Empirical paper on financial ratios	Financial ratios of three companies in UK	[28]
Implementation of Linear Regression Model and was used for the predictive model. Explanatory variables (EV) were analyzed. Regression coefficient was observed	Residential energy and Forest dataset	[29, 30]
Regression and multiple regression models were applied. The spatial predictive modeling was made and Multivariate analysis.	Spices dataset, Sugar and acid content from Citrus fruits	[31, 32]
k-Nearest Neighbor performance in Simulated using Annealing, a metaheuristic search algorithm	Realtime dataset used	[37]

4. Result Analysis and Discussion

The open source Voyant tool has been used for visual analytics for the purpose of research. A network diagram was created for the analysis and it was found that the data and information bifurcate the use of technology. Both AI and Deep Learning use historical data and store it for future use in analysis and reporting purposes.

Deep Learning is also called deep structured learning or hierarchical learning used for creation of machine learning patterns. The objective of this paper is to validate the user based on the entries already available in the table and find out the correct entries through a prediction model creation with its validation. The objective estimate of the probability of user validation from user log enables data to validate the correct user with greater precision. The best measure of the predictive power of an equation is destination of packet, size of packet and the source of the packets. A simple equation is formed to identify the relationship with the valid user. The training dataset was thus created and the logistic regression model was applied.

Logistics regression is used and can be applied to a big dataset to create a model for authentication of users from a security point of view. This can be used to identify the valid users. As a new user enters the system, it identifies the parameters whether the class of the user is valid or not. This then helps the system to be alert and if there is a suspicious entry done by the user in the system, it immediately blocks the entry and does not allow the user to do further transactions. This will be considered as an application of AI as the prediction model based on the history data and thus helps as an application to enhance the security of the system.

Data Model Used

AI implements the use of attack models and security protocols for the implementation of security. It tries to automate the system and uses the directed graphs for the implementation of the model.[21].

Captcha has been introduced as an automatic test to bypass human and solution of AI problems. Captcha may be used in many applications and it can enhance the security. Captcha based voice not picture recognition helps to identify the correct user [19] It is also considered to be the best fit model currently used [20].

A past research article has discussed various types of logit options available for modelling [21]. This has helped to explore the different options which have been examined and used for the category wise predictive modelling. There are many researchers that have used these types of regression for their study. The generalized linear model (GLM) is a more flexible generalization of ordinary linear regression that allows for response variables used for training dataset.

They have used commands such as GLM, it has been used with class, Analysis Of Variance(anova) etc. They further describe the details about the model and also the model parameters to be studied for the interpretation [22, 23]. Poisson regression data has also been applied to the real estate data and the array of applications from the social to the physical sciences has been examined and analyzed [24].

The research examines the risk factors of victims of cybercrime and the analysis is done at the dataset that shows that a logistic GLM, where the conventional methods using the Wald test have been applied, the deviance of change can be examined [25]

The model is proposed using the data and information layer. The Information layer uses summarization, analysis and collaboration of data using blockchain and AI. The data layer enables Deep Learning and AI, which makes a very secure use of data and information in different layers. This mitigates the risk and increases the trust and security in the system. The privacy/ personal data threats are taken care of.

The model discussed in this paper analyzes a major concern of protection and the storage of data. The collaborative system designed is based on the three futuristic techniques i.e AI, Blockchain and Deep Learning. We have divided our model in five stages- Stage 1: AI; Stage 2: Deep Learning; Stage 3: Blockchain; Stage 4: Clustering; and Stage 5: Mainframe. Stages 1 to 3 work as the Data layer and stage 4 and stage 5 work as the information layer.

Stage 1, 2 and 3 use and analyze the data and work confidently in AI and Deep Learning and Block chain. The input data is from different sources and is in the form of structured and unstructured data used for the item of interest. The data can be used with AI for creating patterns, which flows into the data layer. AI approach is used to check for any robotic interference and the filters are analyzed for making of a pattern.

The stage AI gathers the data and passes on to Deep Learning, which analyzes the data to identify filters for the consideration of different categories. Genuine data is thus passed on for Deep Learning. This stage filters the pattern as per the requirement and based on the filters the pattern is designed. The AI further connects to the block chain where the information is required for analysis, summarizing and collaboration of the information.

The remaining data is passed to stage 3 for block chain formation where blocks of users are created which help to accommodate all data without revealing an individual's identity as shown in Fig.3. In this whole model the security and privacy of data is ensured.

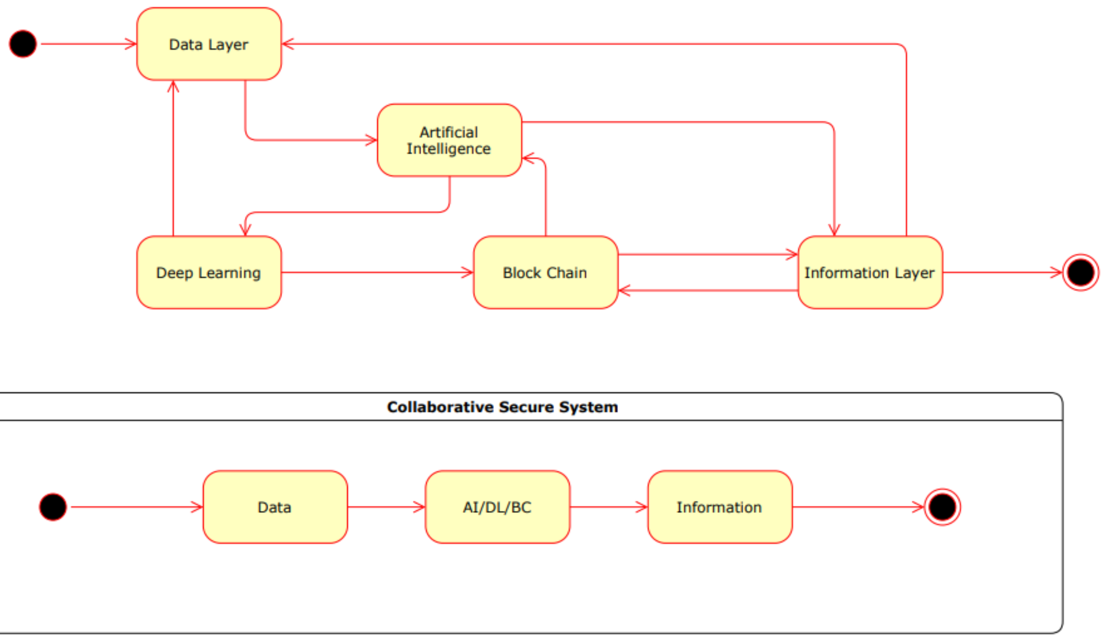

Fig.3. Proposed Security Model

4.1 Logistic Regression Model

In the collaborative secure system this paper discusses the AI, DL and BC. In the present study, this paper considers only AI implementation for a collaborative secure system. The DL and BC for collaborative secure systems will be considered in future. This research paper only discusses the implementation of AI to the system to make it secure. In the present logistic model, study, dependent or response variable is categorical. Hence, a logistic panel is suitable to predict the authenticity of the user. In a logistic model, the response variable is identified as integer’s values. It means a logistic model contains qualitative dependent variables. When the response variable (y) is quantitative, its objective is to estimate expected or mean value but if it qualitative in nature, its objective is to find the probability of something happening. Therefore, it is known as probability models.

The Logistic Regression defines a regression model where the response variable has a categorical value such as True/False or 0/1 i.e., dependent variable. The measure of the probability of a response which is binary i.e. normal, suspicious and unknown under the variable “class”. The probability of a binary response actually measures the value of the response variable defined in a mathematical equation involving the predictor variables. The general mathematical equation for logistic regression is –

$$y = 1/(1+e\text{^}-(a+b1x1+b2x2+b3x3+...)) \tag{1}$$

Where the use of the parameters is as under. The response

$$\text{variable } y = 1/(1+e\text{^}-(a+b1x1+b2x2+b3x3+...)) \tag{2}$$

Where the use of the parameters is as under. The response variable i.e “y”, predictor variable i.e x and the coefficients which are numeric constants i.e. a and b. The function used to create the regression model in using the R statistical software using the glm() function. In the logistics panel the p value plays an vital role in estimation and creation of the model [33,34].i.e “y”, predictor variable i.e x and the coefficients which are numeric constants i.e. a and b. The function used to create the regression model in using the R statistical software using the glm() function. In the logistics panel the p value plays a vital role in estimation and creation of the model [33,34].

4.2 Data collection

The basic objective of this research paper is to create the corrected model for the determination of a valid user. This will help in establishment of the security systems in the implementation of AI, as it will be creating a model for predicting the validity of users. The dataset for the experiment was downloaded from the kaggle data warehouse website. The data columns are Date.first.seen, Duration, Proto, Src. IP.Addr, Src.Pt , Dst.IP. Addr,Dst.Pt, Packets, Bytes, Flows, Flags, Tos, class. The columns which contain a numeric value want to be a part of creating the model and it was found that the Duration, Dst.Pt, Packets, Src.Pt and class are used. Also, the columns which are not useful can be removed from the dataset. The csv file is imported in a dataset and the dataset named user_log was created.

```
#Import Data
user_log <- read.csv("user_log.csv",stringsAsFactors = TRUE)
```

The column details are the sample data of all the columns are as under:

```
'data.frame': 172838 obs. of  16 variables:
 $ Date.first.seen  : Factor w/ 29659 levels "00:00.0","00:00.1",..: 21572
 $ Duration         : num  81413 81413 81505 81505 82101 ...
 $ Proto            : Factor w/ 4 levels "GRE  ","ICMP ",..: 3 3 3 3 3 3 3 3 3
 $ Src.IP.Addr      : Factor w/ 10539 levels "10000_214","10001_101",..:
 $ Src.Pt           : int  8082 56978 8082 56979 8082 51649 8082 37039
 $ Dst.IP.Addr      : Factor w/ 10478 levels "10000_214","10001_101",..:
 $ Dst.Pt           : num  56978 8082 56979 8082 51649 ...
 $ Packets          : int  3057 4748 8639 12024 11012 14186 9974 16476
 $ Bytes            : Factor w/ 4695 levels "   1.2 M","   1.4 M",..: 4 6 25 29
 $ Flows            : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Flags            : Factor w/ 25 levels "  0x53","  0xc2",..: 21 21 21 21 22
 $ Tos              : int  0 0 0 0 0 0 0 0 0 0 ...
 $ class            : Factor w/ 3 levels "normal","suspicious",..: 1 1 1 1 1 1 1
 $ attackType       : Factor w/ 1 level "---": 1 1 1 1 1 1 1 1 1 1 ...
 $ attackID         : Factor w/ 1 level "---": 1 1 1 1 1 1 1 1 1 1 ...
 $ attackDescription: Factor w/ 1 level "---": 1 1 1 1 1 1 1 1 1 1 ...
```

Source: Authors own calculation

The columns which are of not use can be removed from the dataset the command run to remove the column are

User_type$Date.first.seen -> null

Similarly the columns such as proto, Src.IP.Addr, Dst.IP.Addr, Bytes, Flows , Flags , attackType , attackID, attackDescription and Tos are removed from the dataset.

4.3 Results of Descriptive Statistics

The descriptive /summary statistics are as under refer to Table 3.

Table 3. descriptive /summary statistics

	Duration	Dst.Pt	Packets	Src.Pt	class
Min	0.0	0	1	0	normal: 49606
1 Qrt	0.1	23	5	23	supcious:107344
Median	5.9	8000	7	8000	Unknown:15888
Mean	136.3	23310	13.81	22445	
3 Qrt	16.8	49002	17	49145	
Max	519611.2	65535	34136	65535	

Source: Authors own calculation

Data slicing done for creating the training data and it helps in dividing the records in the ratio of 60% and 40%, where the 60% records are used for the training the model and 40% is used for applying the test and compare with the actual values used for calculating the success rate of model defined.

The command used for creating the logistics regression model is as under:

```
> myModel <- glm(class~ Dst.Pt+Packets+Src.Pt,data=train,family = 'binomial')
> summary(myModel)
```

Results of the logistics Model refer to Table 4. is as under

Table 4. Results of Logit Model

Coefficients	Estimate	Std.Error	z value	Pr(>|z|)	Sig
(Intercept)	1.595e+01	1.217e-01	0.7430	131.088	***
Dst.Pt	2.028e-06	-130.927	-2.656e-04	< 2e-16	***
Packets	-2.749e-04	8.831e-05	-3.113	0.00185	**
Src.Pt	-2.654e-04	2.027e-06	-130.936	< 2e-16	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 (Dispersion parameter for binomial family taken to be 1) Null deviance: 124559 on 103763 degrees of freedom Residual deviance: 53510 on 103760 degrees of freedom AIC: 53518 Number of Fisher Scoring iterations: 8					

Source: Authors own calculation

The results suggest that the greater number of stars defines the significance. It means that Dst_Pt, Src_Pt give the more significant contribution in measure. Also, the z value is less than .05 as depicts that they are more significant in the contribution of model reliability. The balance model is more reliable.

Selection of Variables The best measure of the predictive power of an equation is the Destination of the Packet, size of packet and the source of the packets. A simple equation is formed to identify the relationship with the valid user. The training a dataset was thus created and the logistic regression model was applied with these variables. The dataset was chosen specifically for the purpose of defining a model to determine the authenticity of the user. The column that specifies the same is class. This dataset defines whether the class is normal, suspicious or unknown. There are 16 variables and found that as per the user, suspicious activity can be in terms of following: source of packet- their origin, packet- size of data used and destination of packet-their destination. These variables are considered for the model for testing the activity of the user. This will analyse whether the user is normal or not.

The identification of the variables for the model are the dependent variable whether the class will be normal or suspicious or unknown. Also, all the numerical variable with wide ranges of values is considered as the independent variables for the model. They are Src.Pt, Dst.Pt and the Packet size. That is the sources and destination of packets and also the size of packets.

G2 = −2 logL is used to calculate the Residual deviance, which is the difference in a maximal model. These are the separate parameters for each cell in the model and the built model. The Changes in the deviance (the difference in the quantity −2 logL) for two models will be calculated by chisquare. χ 2. The distribution with degree of freedom value is equal to the change in the number of estimated parameters. Thus the difference in deviances can be tested against the χ 2 distribution for significance.

The predictor used for testing the model is the use of test and train data which is useful for the testing the model using training dataset

```
# testing of model with the test dataset
res <- predict(myModel,test,type="response")
res
# testing of model with the train dataset
res <- predict(myModel,train,type="response")
res
```

The use of actual and predicted value the dataset is created with response more than 50%

```
>confmatrix <- table(Actual_value=train$class, Predicted_value = res > 0.5)
> confmatrix
```

Predicted_value	**Actual_value**	
	FALSE	TRUE
normal	23639	6228
suspicious	5458	58895
unknown	719	8825

Source: Authors own calculation

The accuracy has been calculated as 79%. This model is used to check the user authentication.

```
> (confmatrix[[1,1]]+confmatrix[[2,2]])/sum(confmatrix)
[1] 0.7954011
```

The predictive accuracy is 79.5%, which is considered for the fitness of the model and further validates the model for its further use. As the value is 79.5%, there are more hidden correlations in this model needed to make it 98% accurate. But still the normal user can be judged based on the selected parameters used in the model. There may be hidden correlations which have not been considered to be present in the model. The validity of the model can be determined by the use of predictive accuracy. As the accuracy is 79.5% it means the model is accurate and also validate.

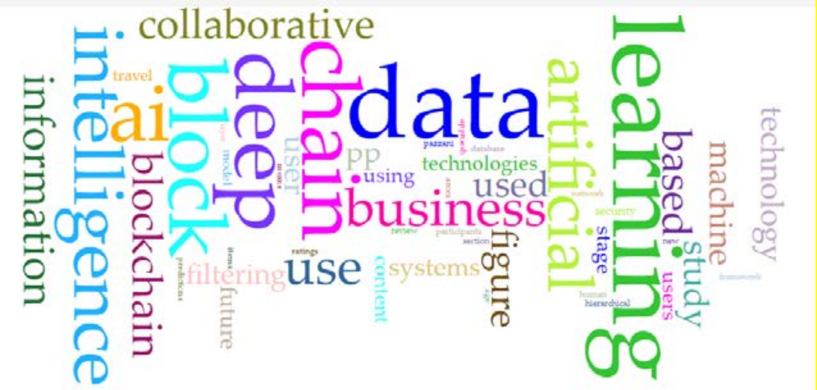

Fig. 4. Word Cloud

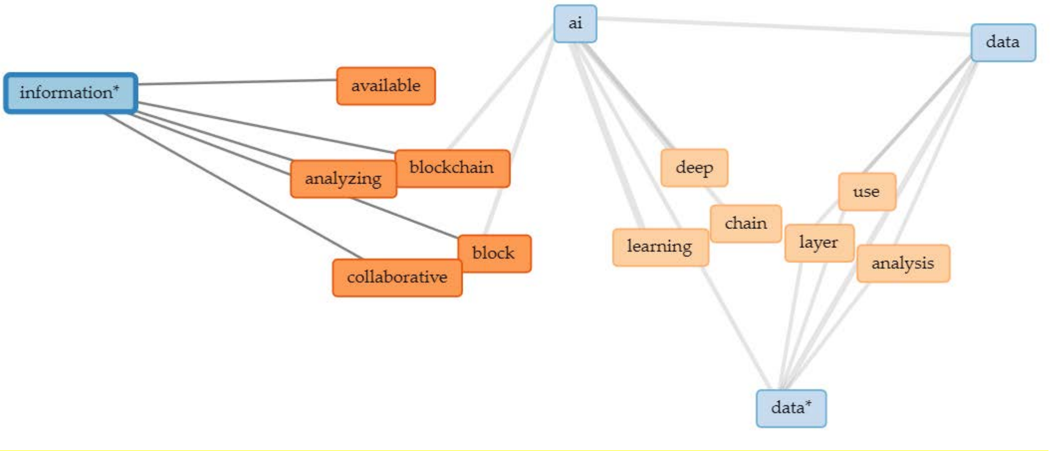

Fig.5. Information based connections

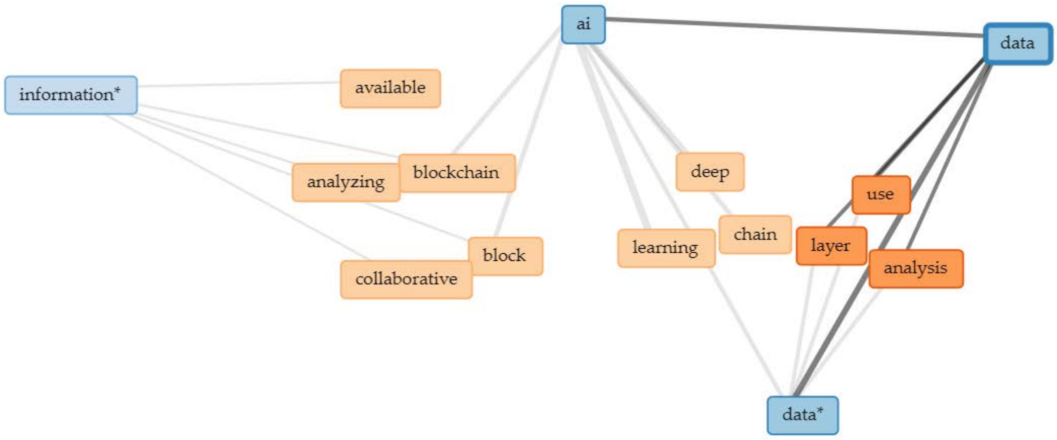

Fig.6. Data based connections

5. Conclusion and Future Recommendation

Logistic models are capable of doing the fit test. They are involved in ROC analysis, Hosmer-Lemes goodness-of-fit test is based on the nature of variables, their link test and residual analysis. These models classify and test the data. The comparison can be based on the expected probabilities at different levels, and their consistency across levels is examined and used for further analysis. This paper also presents the AI implementation with Deep Learning and Block chain to implement collaborative security. Initially, the paper started with establishing the networking between the data and the information. This relationship also defines the relevance of relationships between AI, Deep Learning and Block chain.

AI for security was implemented as the user authentication for security systems. The training data was tested and accuracy was calculated as 79%. It used a training set in the model, to check the feasibility of study. While conducting the experimental study, an AI approach was used as the first level of security wherein only the relevant and genuine data was passed to the next level, i.e., Deep Learning for hierarchical study.

This study was limited to AI used in the transactions of user log and was not applied to Deep Learning & Blockchain. The data collected was collected from open source for testing purposes. We recommend, the dataset to be collected physically real-time transaction for purpose will provide a solution of implementation, setting up model programs; and design of Enterprise into controllable portions. All business units or departments of the organization should be integrated in a small, adjustable method. Thus, for an effective business solution, should consider blockchain technology and Deep Learning.

With both AI and blockchain technology together in one place, data usage could be controlled. With all the emerging technologies used in combination they are able to provide a different edge to the business processes. This research has also redefined the parameters of data security and attempts to provide a solution for the same for the future. The extension of model has been discussed as follows:

The dataset has other parameters in the security system which have not been explored and this will define the future scope for research in this area. Other techniques to enhance security of data and information can be also identified and analyzed further for future studies. The world cloud depicts the use of bigger words as compared to small words. The most occurrences of the words discussed in the Fig.4. are as follows:

Data, chain, Deep Learning, Block chain, AI, information etc

Based on the popularity of different words in the document. The use of the words and the network diagram will help us to create a flow chart. This proposed model will finally help in the implementation of the project as shown in figure defining in Fig.5. and Fig.6. Fig.5. discusses the involvement of information in AI, Block chain and Deep Learning. Fig.6. discusses the involvement of data in AI, Blockchain and Deep Learning.

Fig.5. discusses that the information is dependent on blockchain directly and connected to AI indirectly through blockchain and Further AI is linked to Deep Learning and is directly linked to the data.

Fig.6. discusses the data connection which is directly connected to AI and Deep Learning. Further the AI is connected to blockchain and finally linked to the information.

All the three technologies are important but it has been seen that these originate from AI. Thus, AI can be broadly used for the user authentication for the check whether the user is valid or not. This will be at the very first stage where one can stop the user from entering the system making it a very secure system. The tracking of the data filled by every user can be checked and validated. For the experiment, the sample dataset was considered and logistic regression was applied.

References

[1] Arora M., Chopra A.B., Dixit V.S. (2020) An Approach to Secure Collaborative Recommender System Using Artificial Intelligence, Deep Learning, and Blockchain. In: Choudhury S., Mishra R., Mishra R., Kumar A. (eds) Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, vol 989. Springer, Singapore. https://doi.org/10.1007/978-981-13-8618-3_51

[2] Waheed, N., He, X., Ikram, M., Usman, M., Hashmi, S. S., & Usman, M. (2020). Security and privacy in IoT using machine learning and blockchain: Threats and countermeasures. ACM Computing Surveys (CSUR), 53(6), 1–37.

[3] Salah, K., Rehman, M. H. U., Nizamuddin, N., & Al-Fuqaha, A. (2019). Blockchain for AI: Review and open research challenges. IEEE Access, 7, 10127–10149.

[4] Singh, S., Sharma, P. K., Yoon, B., Shojafar, M., Cho, G. H., & Ra, I.-H. (2020). Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. Sustainable Cities and Society, 63, 102364.

[5] Ekramifard, A., Amintoosi, H., Seno, A. H., Dehghantanha, A., & Parizi, R. M. (2020). A systematic literature review of integration of blockchain and artificial intelligence. In Blockchain Cybersecurity, Trust and Privacy (pp. 147–160). Springer.

[6] Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29.

[7] Wang, X., Han, Y., Leung, V. C., Niyato, D., Yan, X., & Chen, X. (2020). Convergence of edge computing and Deep Learning: A comprehensive survey. IEEE Communications Surveys & Tutorials, 22(2), 869–904.

[8] Singh, S. K., Rathore, S., & Park, J. H. (2020). Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence. Future Generation Computer Systems, 110, 721–743

[9] Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. Metabolism, 69, S36-S40.

[10] Mustapha, A. M., Arogundade, O. T., Misra, S., Damasevicius, R., & Maskeliunas, R. (2020). A systematic literature review on compliance requirements management of business processes. International Journal of System Assurance Engineering and Management, 11(3), 561-576.

[11] Arogundade, O. T., Abayomi-Alli, A., & Misra, S. (2020). An Ontology-Based Security Risk Management Model for Information Systems. Arabian Journal for Science and Engineering, 1-16.

[12] Kamta Nath Mishra, "A Proficient Mechanism for Cloud Security Supervision in Distributive Computing Environment", International Journal of Computer Network and Information Security, Vol.12, No.6, pp.57-77, 2020.

[13] Arora M and Arora A. (2018), "Digital Information Tracking Framework using Blockchain, Journal of Supply Chain management Systems (UGC Approved), Volume 7 Issue 2, pp 1-7, ISSN: 2277-1387.

[14] Lotter, W., Kreiman, G., Cox, D (2016)." Deep predictive coding networks for video prediction and unsupervised learning". arXiv preprint arXiv:1605.08104

[15] Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1997). Neural network applications in business: A review and analysis of the literature (1988–1995). Decision Support Systems, 19(4), 301-320.

[16] Gleichauf, R. E., Teal, D. M., & Wiley, K. L. (2002). U.S. Patent No. 6,499,107. Washington, DC: U.S. Patent and Trademark Office.

[17] Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain intelligence: go beyond artificial intelligence. Mobile Networks and Applications, 23(2), 368-375.

[18] Park,H.,D., Kim., K., H., Choi, Y.,II., and Kim, K.,J., (2012). "A literature review and classification of systems research", Expert Systems with Applications (39: 1) pp. 10059-10072

[19] Ahn L., Blum M., Hopper N.J., Langford J. (2003) CAPTCHA: Using Hard AI Problems for Security. In: Biham E. (eds) Advances in Cryptology — EUROCRYPT 2003. EUROCRYPT 2003. Lecture Notes in Computer Science, vol 2656. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-39200-9_18

[20] McLeod, A. I., & Xu, C. (2010). bestglm: Best subset GLM. URL http://CRAN. R-project. org/package= bestglm accessed on 15 December 2020.

[21] Bozic, J., & Wotawa, F. (2017). Planning the attack! or how to use ai in security testing?. In IWAISe: First International Workshop on Artificial Intelligence in Security (Vol. 50) http://iwaise.it.nuigalway.ie/wp-content/uploads/2017/02/IWAISe-17_paper_10-jb.pdf accessed on 20 December 2020.

[22] Hilbe, J. M. (2009). Logistic regression models. CRC press.

[23] Hilbe, J. M. (2011). Logistic Regression. International encyclopedia of statistical science, 1, 15-32. https://encyclopediaofmath.org/images/6/69/Logistic_regression.pdf accessed on 20 December 2020

[24] Sellers, Kimberly F., and Galit Shmueli. "A flexible regression model for count data." The Annals of Applied Statistics (2010): 943-961.

[25] Sfetsos, A., & Coonick, A. H. (2000). Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Solar Energy, 68(2), 169-178.

[26] Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. The annals of applied statistics, 2(4), 1360-1383

[27] Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. Journal of Business Research, 68(5), 1012-1025.

[28] Ciampi, F., & Gordini, N. (2008, January). Using economic-financial ratios for small enterprise default prediction modeling: An empirical analysis. In 2008 Oxford Business & Economics Conference Proceedings, Association for Business and Economics Research (ABER) (pp. 1-21) accessed on 20 December 2020.

[29] Fumo, N., & Biswas, M. R. (2015). Regression analysis for prediction of residential energy consumption. Renewable and Sustainable Energy Reviews, 47, 332–343

[30] Kumar, R., Nandy, S., Agarwal, R., & Kushwaha, S. P. S. (2014). Forest cover dynamics analysis and prediction modeling using logistic regression model. Ecological Indicators, 45, 444–455.

[31] Lehmann, A., Overton, J. M., & Leathwick, J. R. (2002). GRASP: Generalized regression analysis and spatial prediction. Ecological Modelling, 157(2–3), 189–207

[32] Song, S. Y., Lee, Y. K., & Kim, I.-J. (2016). Sugar and acid content of Citrus prediction modeling using FT-IR fingerprinting in combination with multivariate statistical analysis. Food Chemistry, 190, 1027–1032

[33] Steyerberg, E. W., Harrell Jr, F. E., Borsboom, G. J., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology, 54(8), 774–781

[34] Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Computing and Applications, 19(8), 1165-1195.

[35] Nakamoto, S., & Bitcoin, A. (2008). A peer-to-peer electronic cash system. Bitcoin.–URL: https://bitcoin. org/bitcoin.pdf accessed on 8 January 2021

[36] Gill, S. S., Tuli, S., Xu, M., Singh, I., Singh, K. V., Lindsay, D. ... & Garraghan, P. (2019). Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges. Internet of Things, 8, 100118.

[37] N. M. Tahir, Adam N. Ausat, Usman I. Bature, Kamal A. Abubakar, Ibrahim Gambo, "Off-line Handwritten Signature Verification System: Artificial Neural Network Approach", International Journal of Intelligent Systems and Applications, Vol.13, No.1, pp.45-57, 2021.

[38] Anozie Onyezewe, Armand F. Kana, Fatimah B. Abdullahi, Aminu O. Abdulsalami, "An Enhanced Adaptive k-Nearest Neighbor Classifier Using Simulated Annealing", International Journal of Intelligent Systems and Applications, Vol.13, No.1, pp.34-44, 2021.

[39] Hossein Mohammadinejad, Fateme Mohammadhoseini, "Privacy Protection in Smart Cities by a Personal Data Management Protocol in Blockchain", International Journal of Computer Network and Information Security, Vol.12, No.3, pp.44-52, 2020.

Authors' Profiles

Monika Arora is currently working as a Associate Professor at the Apeejay School of Management. She has more than 24 years of blended experience of teaching experience and Industry such as nucleus software exports Limited, Global Information System Technology and Platinum EDU. She has completed her PhD in computer science from MPBHOU, Bhopal She has done her Bachelor of Science in Mathematics BSc (Maths Hons) from Devi Ahiliya University, Indore and Master of Computer Application from Barkatullah University, Bhopal. Her research interest includes Analytics, Artificial Intelligence, Blockchain Technology and Information Security.

Indira Bhardwaj is currently working as a Dean at Vivekananda School of Business Studies. She has more than 20 years of teaching experience across institutions in India. Her PhD is from AMU in Understanding Corporate Value Using Intellectual Capital Assets. Her research interests include Corporate Valuation and Corporate Finance. She has been Research Associate at IIM Indore and has worked as Knowledge Manager with Forbes EduMetry, providing consultancy for AACSB accreditation to business schools in US and SE Asia. Her other areas of interest include sustainable management and sustainability reporting.

How to cite this paper: Monika Arora, Indira Bhardwaj, "Artificial Intelligence in Collaborative Information System", International Journal of Modern Education and Computer Science(IJMECS), Vol.14, No.1, pp. 44-55, 2022.DOI: 10.5815/ijmecs.2022.01.04