International Journal of Management, IT & Engineering

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

CRIMINAL JUSTICE SYSTEM USING AI TECHNIQUE BY VERIFYING CREDENTIALS.

Srishti Roy[1], Kanhaiya Lal[2]

[1]Student
BIT MESRA (PATNA CAMPUS)
srishtiroy2706@gmail.com

[2]Assistant professor
Computer science and engineering
BIT MESRA (PATNA CAMPUS)klal@bitmesra.ac.in

Abstract

The criminal justice system is in disarray due to the rising number of cyberattacks. Understanding the motivation for crime and combating it is a key responsibility for law enforcement. In order to categorize the current criminal record while using metadata and forecast crime, The results of this study will show how AI, ML, QA, and soft evidence may be used in practise. More than that, having this database at their disposal would greatly aid law enforcement and intelligence organizations in their quest to solve crimes more quickly and at lower cost to the taxpayers, so contributing to a reduction in crime across the country. The analyst would be better able to track the connections and movements of various criminal elements if they had access to the detailed information included in the papers and records. This study can assist us in understanding how to predict crimes. The evaluation accurately represents the amount of threat posed by 28 Indian states. When this model is loaded with the right data, it is clear from study on the subject that the probability of a prediction is higher and more precise. The study also sought to understand the psychological views on crime and the reasoning behind individuals who engage in it.

KEYWORDS: Machine learning, law enforcement, Artificial Intelligence criminal justice, accuracy, cyber-attacks, information technology laws, motives, prediction algorithm etc.

Introduction

John Mc-Carthy, the son, first presented artificial intelligence (AI) in Dartmouth in 1956 [McCarthy J., 2006: 12–14]. Since technology is the initial layer of thedigital transformation,

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

there are hazards involved [Dmitrik N., 2020: 54–78]. "During the past few years, artificial intelligence (AI) and machine learning (ML)-based technologies have rapidly increased in capability and accessibility, and this trend is not likely to slow down". [Caldwell M., 2020: 1–13]. As such, knowing the pros and cons of the future AI rule will enable humanity determine whether or not to implement it. [Cui Y., 2020: 187–191]. The purpose of AI study and oversight is to find a happy medium between the upsides and downsides of innovation. ["King T., Aggarwal N., Taddeo M"., "Floridi L.,2020: 89–120"]. The Indian government places a high value on artificial intelligence research, development, and deployment because of the positive impact it can have on people's daily lives. [Marda V., 2018: 1–19]. The facts of the depicted crimes and the locations where they took place helped provide context for crimes of a similar nature that have occurred in other nations. ["Furtado V., 2010: 4–17"]. "The Superb and much-needed Roultedge Handbook of technology", crime and "Justice is further illuminated by McGuire and Holt's work" ["McGuire M., Holt T., eds., 2017:1–722"] This demonstrates how criminology has shifted its attention to technological tools. ["Hayward K., Maas M., 2020: 1–25"]. The most crucial application of this knowledge would be to upgrade judges to be experts in the realm of computers; regulations should be put in place requiring that all judges receive enough training in using this technology. 1. Using John McCarthy's invention, artificial intelligence, which is the major rising technology and advantageous since it interprets all data honestly. In contrast, a human psyche must pick or make a decision from among the several pieces of information before thinking, which might result in errors.

Literture Review:

(Thomas J. Holt, 2010) Nowadays, sophisticated statistical inquiries into numerous quantitative data sources have substituted qualitative research methods in criminology. Yet, a growing body of qualitative research is being conducted with data collected from a wide range of online sources. This essay explores how traditional qualitative criminology research can benefit "from the Internet, websites, and various CMCs". All these varieties, in addition to the special methodological and ethical considerations of online investigations, are discussed at length.

(Yazan Mualla, 2019) As AI evolves, it influences many different sectors. Multi-Agent Systems (MAS) are making great progress in their implementation in vital areas such as healthcare, autonomous vehicles, law enforcement, and the economy. This progress adds to the growing interdependence of AI and human cultures. That calls into question the public's faith in AI

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

helpers, which is a problem on multiple levels. The most pressing issues stem from an inability to trust one other, which can be difficult to articulate. Enhancing performance has taken precedence over simplifying systems' interpretation in recent decades. The field of image processing, natural language processing, and automated decision making all owe a debt to this breakthrough. Anxiety about the reliability of AI-based decisions is raising pertinent questions that may lead to the development of innovative methods and solutions that put the user first.

Preparing the Model

Since it allows us to keep tabs on how various states act and the crimes they commit, understanding recidivism is the most significant concept for researching this subject. In Fig. 1, we can see how our data goes through a series of steps before yielding the desired result.The development of ML research and applications is accomplished by many programming environments and languages. Over the ensuing ten years, Python has experienced extensive development in the tensor flow communities, Therefore, in this situation, Python-based ML and deep learning packages are most commonly used [Raschka S. et al, 2020: 193]. When a particular set of data grows to be vast or we need to comprehend some intricate interactions between the variables, Python is utilized to construct the predictive analysis model. We may execute the sculpting of such data for enhanced study purpose that use this paper.

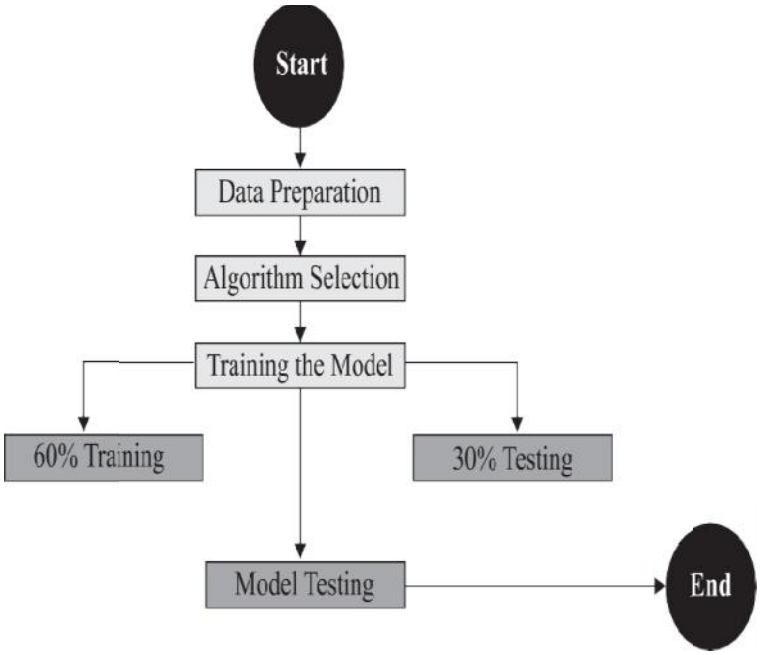


Fig.1.Model Process Flowchart"

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

```
In |2|:  import pandas as                    # pandas is a
         pdimport NumPy                     dataframe
         as np                              library# numpy
         import matplotlib.pyplot as plt    provides
         import seborn as                   N-dim object
         seabornInstancefrom                support#
         sklearn.linear_model import        matplotlib.pyplot
         LogisticRegression                 plots data
         from sklearn.model_selection import train_
         test_split
         from sklearn import metrics
         import os
         % matplotlib inline
```

"Fig.2.Importing Libraries"

Second, it's crucial to regulate and visualize data to make sure it adheres to the models' presumptions. According to the Universal Rule of Law, gains in the justice system made by employers and governments will have an impact on other areas such as human rights, democracy, and development. The primary and most important objectives of the criminal justice system are crime control and prevention, maintaining law and order, protecting the rights of victims and those who come into conflict with the law, punishing and rehabilitating those found guilty of crimes, and safeguarding life and property from the effects of crime and criminality. According to the Indian Constitution, it is seen as the state's main duty [Dhillon K., 2011: 27]. Thus, this article will present an insight of how each police station may update its data and forecast criminal behavior using any accessible data. Since the data in this study work may be freely altered, importing different libraries and functions is a benefit of implementing Python, as can be shown in Figs. 3 and 4.Since it's harder to accurately forecast uncommon occurrences, there is little possibility that they will arise in data or that the algorithm will be built on them.An individual or nation's propensity to learn criminal behavior and purpose can be assessed using just a tiny sample of the event itself. It is easy to determine the frequency of occurrence after pandas has been imported, as we need only search the columns by name. Finally, in the third and last column, you'll find the threat columns, cleanly divided into binary 1s and 0s to indicate if attacks are rapidly increasing or not. Skepticism arises when a murder is predicted using an algorithm or code. Accurate data would boost efficiency. This research would shed much-needed

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

light on an area where numbers alone may be decisive. Predictive analytics and data mining are vital to our day-to-day operations. Using the data at hand, we can pinpoint which states are suffering from a high unemployment rate for this specific reason. Those living in states with the greatest incidences of cybercrime are also more prone to fall prey to online frauds intended to steal their money.This is utilized in order to remain anonymous while destroying someone for revenge or other reasons. When high-speed internet is accessible at a reduced cost, cybercriminals frequently take use of it to conduct various crimes without being identified unless the state has adequately equipped and kept in touch cybersecurity institutes to prohibit such crimes.The CMIE analysis also discloses that all of those between of ages of 40 and 59 have been successful in keeping their careers, while those smaller age 40 have been fired, which has caused to societal turmoil, a desire for retribution, rage, and other incentives for such cyberattacks6.

The information in Fig. (3) illustrates the top cybercrimes that occur in various Indian states up till 2019. This means that people of a certain age range have committed such attacks in order to tarnish the victim's reputation7.To date, this contains crimes like cyberbullying and does not represent full-fledged

4. How code can be trustworthy. Available at: https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337(accessed:05.03.2018)

5What is Data Mining? Definition of Data Mining, Data Mining Meaning — The EconomicTimes(indiatimes.com).Availableat:https://economictimes.indiatimes.com/definition/data-mining(accessed:07.12.2020)

6The recent unemployment data. Available at: https://www.cmie.com/kommon/bin/sr.php?kall=warticle&dt=2020-01-21%2009:51:47&msec=203(accessed:21.01.2020)

7National Crime Records Bureau Empowering Indian Police with Information TechnologyAvailableat:https://ncrb.gov.in/en(accessed:22.10.2020

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

State UT	"Personal"_I	"Anger"	"Fraud"	"Extortion"	"Causing"_I	"Prank"	"Sexual"	Political	Terrorist	Inciting Hate Country	Disrupt Service	Sale purchase	Developing	Spreading	Psycho	Steal	Abetment	Others	Risk
"Andhara Pradesh"	34	26	733	45	7	0	92	12	1	1	1	0	2	14	2	0	1	236	0
"Arunachal Pradesh"	"0"	"0"	"2"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"5"	"0"
"Assam"	239	46	389	153	234	0	113	9	4	3	0	0	0	0	0	0	0	832	1
"Bihar"	"5"	"8"	351	"2"	"0"	"0"	"8"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
"Chhattisga"	"0"	"1"	"23"	"2"	"25"	"0"	"2"	"4"	"0"	"1"	"0"	"0"	"1"	"0"	"0"	"0"	"0"	"61"	"0"
"Goa"	"0"	"0"	"11"	"0"	"12"	"0"	"4"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"2"	"0"
"Gujarat"	"17"	"32"	"401"	"24"	154	16	23	0	0	17	0	0	1	0	0	3	0	14	1
"Haryana"	"6"	"9"	"137"	"21"	"11"	"2"	"75"	"0"	"12"	2	"0"	0	2	"0"	"0"	"0"	0	141	1
"Himachal"	"4"	"1"	"18"	"1"	"3"	"1"	"15"	"4"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"22"	"0"
J&K	2	0	20	7	7	3	10	3	1	2	3	0	0	1	0	0	0	14	0
"Jharkhand"	16	6	783	44	16	0	16	1	16	0	0	0	32	0	0	0	0	0	0
"Karnataka"	27	10	5441	97	49	1	85	22	1	3	3	0	5	5	1	1	0	88	1
"Kerala"	"69"	"18"	"93"	"8"	"48"	"3"	"50"	18	3	"0"	"0"	0	6	"0"	"0"	"0"	0	24	0
"MadhyaP"	"93"	"10"	"230"	"19"	109	2	49	1	3	29	1	0	4	20	0	0	1	169	1
"Maha	99	129	1998	31	64	18	724	20	0	33	2	2	13	6	0	3	0	369	1

rash"																			
"Manipur"	"0"	"0"	14	3	"0"	"0"	9	"0"	1	1	"0"	"0"	"0"	"0"	"0"	"0"	"0"	1	"0"
"Meghalaya"	"0"	"0"	35	"0"	3	"0"	3	3	"0"	11	"0"	2	"0"	6	"0"	"0"	"0"	11	"0"
"Mizoram"	"0"	4	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	1	"0"	"0"	"0"	1	"0"	"0"	"0"	"0"
"Nagaland"	1	"0"	"0"	"0"	1	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
"Odisha"	7	"0"	506	224	"0"	"0"	37	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	69	1
"Punjab"	"14"	"7"	"48"	"15"	"19"	"4"	"85"	"2"	"0"	"3"	"0"	"2"	"2"	"0"	"0"	"0"	"0"	"38"	"0"
"Rajasthan"	"9"	"11"	"499"	"31"	"66"	"14"	"60"	"3"	"0"	"9"	"0"	"0"	"17"	"2"	"0"	"0"	"0"	"383"	"1"
"Sikkim"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"
"TamilNad"	"39"	"13"	"55"	"7"	"17"	"6"	"36"	"52"	2	39	1	0	3	1	0	4	0	20	0
"Telangana"	19	3	732	51	3	2	77	14	0	3	0	0	0	0	0	0	0	301	1
"Tripura"	"5"	"0"	"8"	"0"	"0"	"0"	"3"	"4"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
"Uttar Pradesh"	"47"	"73"	"2351"	"199"	"343"	"191"	"343"	"45"	"0"	"59"	"9"	"0"	"75"	"614"	"0"	"0"	"0"	"1931"	"1"
"Uttarakhand"	"3"	"41"	"46"	"16"	"12"	"22"	"13"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"5"	"0"	"13"	"0"

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

"West Bengal	"28"	"9"	"68"	"25"	"2"	"3"	"39"	"1"	"0"	"2"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"158"	"1"
"A&NIsland"	"0"	"0"	"3"	"0"	"3"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"
"Chandigar"	"0"	"0"	"19"	"3"	"0"	"7"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"
"D&N Have"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
"Daman &"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"
"Delhi UT"	"11"	"4"	"36"	"11"	"4"	"1"	"35"	"0"	"0"	"0"	"0"	"0"	"35"	"2"	"0"	"0"	"0"	"50"	"0"
"Lakshadow"	"0"	"0"	"1"	"0"	"0"	"0"	"2"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"0"	"1"	"0"

"Fig.3.Raw Data(7)"

The term "feature" refers to any component whose function is to decide an outcome, column, on the other hand, might mean anything physical that stores the value of some characteristic or outcome (as shown in Fig. 4).A 36 by 13 grid is used to display the data in Fig. 12 using the shape function. In this study, we look at the missing and null values in the datasets shown in Fig. The crossing plot function in the Matplot module helps us see if two features are connected. The data is then analyzed to find out which parts can be removed. The same technique can be used to get rid of duplicate values. As visual inspection has its limitations and fails to solve the critical

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

problem of related columns, we resort to this strategy to sort our data. In order to determine if a value in a data frame is null, the pandas Is Null method performs the checks shown in Figure 6. This allows us to spot blank spots in our data where critical details could be missing. Using the Matplot library, we can build a function that shows characteristics to identify data relationship; in this case, as can be seen in Fig. 11, the colour yellow indicates a very strong link while other colours suggest lower correlation. The titles along Fig. 11's horizontal and vertical axes serve as a matrix, designating which column in the figure carries the information corresponding to the given values.

```
In [3]:  os.getcwd()
         os.chdir ('C:/Users/Cybercare/CRIME
         RECORD')os.getcwd()

Out[3]:  'C:\\Users\\Cybercare\\CRIME RECORD'

In [20]: Cyber_data = pd.read_csv('Cyber new.csv') // read dataset
         Cyber_data.head()

Out[20]:
```

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

	State_UT	Personal_	Anger	Fraud	Extortion	Causin	Prank	Sexual	Politic	Terrori	Inciting Hate	Disrupt	Sale	Developi
0	Andhra Pradesh	34	26	733	45	7	0	92	12	1	1	1	0	
1	"Arunachal Pradesh"	0	0	2	0	0	0	0	0	0	0	0	0	
2	"Assam"	"239	46	389	153	234	0	113	9	4	3	0	0	
3	"Bihar"	5	8	351	2	0	0	8	0	0	0	0	0	
4	"Chhattisga"	0	1	23	2	25	0	21	4	0	1	0	0	

"Fig.4.Selecting the data"

Figures (4) and (5) show how data is retrieved from the file location and applied to metadata "cleaning and correlating".

```
In [3]:  os.getcwd()
         os.chdir ('C:/Users/Cybercare/CRIME
         RECORD')os.getcwd()

Out[3]:  'C:\\Users\\Cybercare\\CRIME RECORD'

In [20]: Cyber_data = pd.read_csv('Cyber new.csv') # read dataset
         Cyber_data.head()

Out[20]:
```

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Causing_Disrepute	Prank	Sexual Exploitation	Political Motives	Terrorist Activities	Inciting Hate against Country	Disrupt Public Service	Sale purchase illegal drugs	Developing own business	Spreading	PsychoorPervert	StealInformation	Abetment toSuicide	Others	Risk
7	0	92	12	1	1	1	0	2	14	2	0	1	236	0
0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
234	0	113	9	4	3	0	0	0	0	0	0	0	832	1
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0
25	0	21	4	0	1	0	0	1	0	0	0	0	61	0

Fig.5.Showing the data"

Molding the Data

To check for problems, we first test the data after it has been cleaned by removing any extraneous columns or null values. Data molding, as shown in Fig. 1, allows us to use this data to train and run the algorithm despite the fact that algorithms are mathematical models that work best with numerical data.In order to precisely shape the data, calculations are carried out for 6 counts, mean, std, etc. As a result, a lot of data transformation is done in machine learning in order to experiment, learn from failures and make the most accurate forecasts.It is quite simple to modify the meaning of data when it has been manipulated, which also facilitates it determining determine determine where data manipulation has gone wrong. Keeping track of all the updates and modifications has been done automatically because the numerical model was generated in Jupyter Notebook [Perkel J. et al., 2018: 145–147]. As seen in Figs. 6 and 7, we also have the reactive Python interpreter, which we could employ to clarify our data for the prediction.

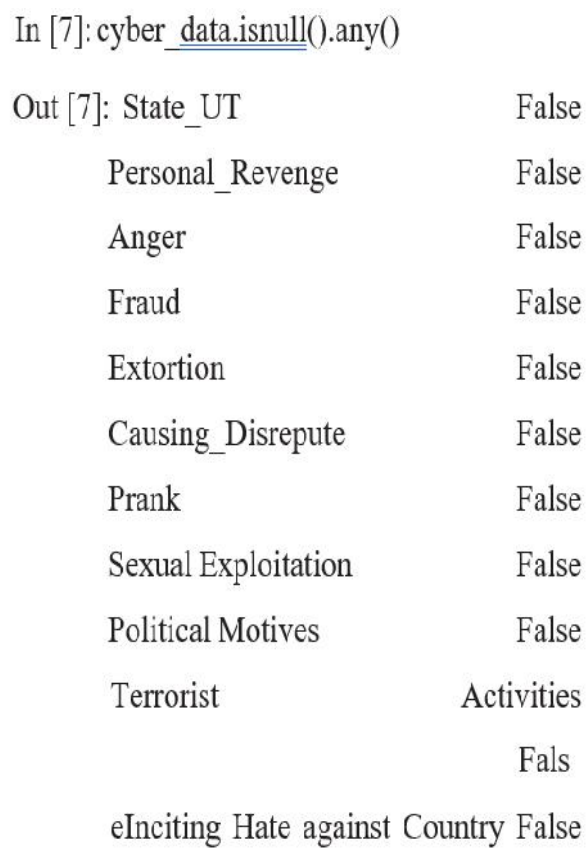
n[6]:Cyber_data.describe()

Out[6]:

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

	Personal_Revenge	Anger	Fraud	Extortion	Causing_Disrepute	Prank	Sexual Exploita-tion	Political Motives	Terrorist Activities	"Inciting Hate against Country"	"Disrupt Public Service"	"Sale purchase illegal drugs"
count	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000
mean	22.055556	12.805556	418.083333	29.166667	33.666667	8.222222	56.388889	6.055556	1.222222	6.055556	0.583333	0.166667
std	44.940560	25.484807	1007.891615	54.345193	72.200119	31.825516	129.880886	12.094732	3.330475	13.259917	1.645340	0.560612
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	6.750000	0.000000	0.000000	0.000000	2.750000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	5.000000	3.500000	41.000000	7.500000	3.500000	0.000000	15.500000	0.500000	0.000000	0.000000	0.000000	0.000000
75%	21.000000	10.250000	392.000000	26.500000	20.500000	3.000000	52.500000	4.000000	1.000000	3.000000	0.000000	0.000000
max	239.000000	129.000000	5441.000000	224.000000	343.000000	191.000000	724.000000	52.000000	16.000000	59.000000	9.000000	2.000000

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A



"Fig.6.Null values are checked"

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

In[6]:Cyber_data.describe()

Out[6]:

"Prank"	Sexual Exploitation	Political Motives	Terrorist Activities	Inciting Hate against Country	Disrupt Public Service	Sale purchase illegal drugs	Developing own business	Spreading	Psycho or Pervert	Steal Information	Abetment to Suicide	Others	Risk
36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000
8.222222	56.388889	6.055556	1.222222	6.055556	0.583333	0.166667	5.500000	18.638889	0.111111	0.444444	0.055556	137.666667	0.305556
31.825516	129.880886	12.094732	3.330475	13.259917	1.645340	0.560612	14.484474	102.147275	0.398410	1.229273	0.232311	349.023454	0.467177
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	2.750000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000
0.000000	15.500000	0.500000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	14.000000	0.000000
3.000000	52.500000	4.000000	1.000000	3.000000	0.000000	0.000000	2.250000	1.000000	0.000000	0.000000	0.000000	101.250000	1.000000
191.000000	724.000000	52.000000	16.000000	59.000000	9.000000	2.000000	75.000000	614.000000	2.000000	5.000000	1.000000	1931.000000	1.000000

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

"Fig.7.Null Values shown in risk column"

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

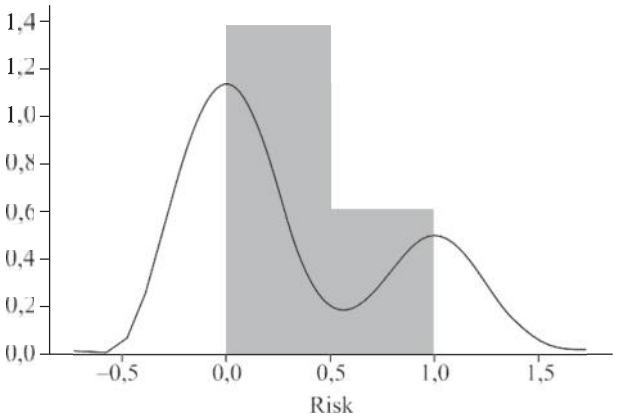
Testing Model's Accuracy

The characteristics of the machine learning algorithm itself will be discussed here. Algorithms may be at the heart of the process, which could explain the observed behavior. We will utilize scikit-learn and the algorithm's logic to create an educated judgement about the future based on our analysis of the data that acts as examples for our forecast. The algorithm uses this evaluation to adjust some of its internal parameters, improving its ability to estimate the parameters it has been taught to utilize and so yielding more precise results in light of the input qualities. The math and logic underlying an algorithm are examined here. It is recommended to utilize a function tailored to the algorithm being tested. The model can be trained after the fit parameters have been saved. With this approach, we can extrapolate future outcomes from the current set of information. Our team makes advantage of the Sci-kit Learn extension for Python to create a predictive model. The Python code and the parameters of the trained model are used to calculate the state's susceptibility to cyberattack. The hardest element of our study for this paper was picking an efficient algorithm from Scikit-Learn.

I wanted to focus on excluding all other techniques, despite the fact that prediction is supervised learning and may be further subdivided into classification and regression (where regression refers to a continuous collection of variables). Then, we disregarded any approaches that didn't use classification, and especially binary classification, to establish whether or not the danger was real. Three algorithms—the Naive Bayes classifier, logistic regression, and the decision tree—greatly enhance and simplify the understanding of more complex machine learning techniques.

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

```
In [8]:   plt. Figure(fig size=(15,10))
          plt.tight_layout() seabornInstance.distplot(cyber_data['Risk'])
Out [8]: <matplotlib.axes._subplots.AxesSubplot at 0x1c303d03808>
```



"Fig.8.Graph Denoting the Risk"

```
In [13]: df.corr()
Out [13]:
```

	Online Banking Frauds	**"Cyber Blackmailing/Threatening Sec506,503, 384"**	**"Fake News on Social Media Sec505"**	**"Cyber Terrorism sec66F"**	**Tampering Computer Source**	**Identity"Theft sec66C"**	**Computer "related offence sec66"**	**Ransom ware**	**"Offences other than Ransoware"**	**"Cyber Stalking/Bullying of Women/ChildrenSec354DIPC"**
"Online Banking Frauds"	1.000000	0.297261	0.210440	– 0.081656	0.297991	– 0.029883	0.278153	0.276576	0.322202	0.799312
Cyber Blackmailing/ThreateningSec506, 503,384	0.297261	1.000000	0.472718	0.368497	0.154649	– 0.049758	0.156595	0.128745	0.154715	0.300828
Fake News	0.210440	0.472718	1.00000	0.589458	0.252916	–	0.231064	0.23535	0.27013	0.205406

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

on Social Media Sec 505			0			0.041483		4	6	
Cyber Terrorism sec 66F	-0.081656	0.368497	0.589458	1.000000	0.028105	–0.042616	0.061458	0.034039	0.025907	–0.047158
Tampering Computer Source	0.297991	0.154649	0.252916	0.028105	1.000000	0.065687	0.991130	0.992945	0.937691	0.028343
"Identity Theft sec 66C"	-"0.029883"	-"0.049758"	-"0.041483"	-"0.042616"	"0.065687"	"1.000000"	"0.014784"	"0.00353"	"0.074809"	"0.000353"
"Computer related offence sec66"	"0.278153"	"0.156595"	"0.231064"	"0.061458"	"0.991130"	"0.014784"	"1.000000"	"0.998242"	"0.955151"	"0.010488"
"Ransomware"	"0.276576"	"0.128745"	"0.235354"	"0.034039"	"0.992945"	"0.00353"	"0.998242"	"1.000000"	"0.949917"	"0.006252"
"Offences other than Ransom ware"	0.322202	0.154715	0.270136	0.025907	0.937691	0.074809	0.955151	0.949917	1.000000	0.060779
"Cyber Stalking/Bullying of Women/Chil-drenSec354 DIPC"	0.799312	0.300828	0.205406	–0.047158	0.028343	0.000353	0.010488	0.006252	0.060779	1.000000

Fig.9.CorrelationPerformed

The logistic regression method is sometimes misunderstood since its name suggests continuous values when in fact it only yields binary findings from statistical analyses. The algorithm delves into the interconnections between features, taking into account how they all contribute to the final result. A threat is indicated by a curve that intersects with the outcome and value in Fig. 8

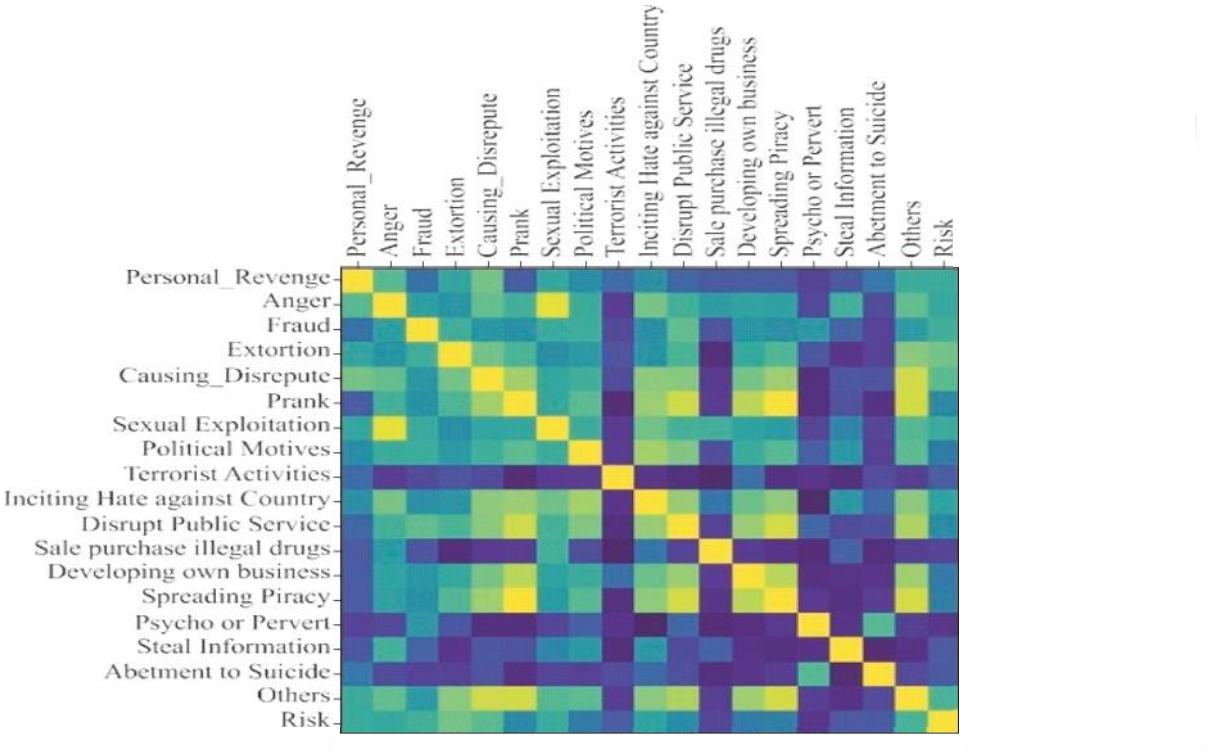
Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

```
def     plot_corr (df, size=11):
        “””
        Function plots a graphical correlation matrix for each pair of columnsin the
        dataframe
        Input:
          df: pandas DataFrame
          size: vertical and horizontal size of the plot
        Displays:
          matrix of correlation between columns. Blue-cyan-yellow-red-darkred
          => less to more correlated
                                          0 ----------------------> 1
                                          Expect a yellow line running from topleft to
        “””                               bottom right
        corr = df. corr ()

        fig, ax = Pl. Subplots (fig size (size, size)) ax. matshow(corr)
        plt. xticks (range (len (corr. columns)), corr.c olumns)plt.
        yticks(range(len(corr. columns)), corr.c olumns)
        plt. setp (ax.get_xticklabels (), rotation=90, horizontalalignment=’right’)
```

“Fig.10.Giving the values for correlation” “In[10]:plot_corr(cyber_data”

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A



***"Fig.11.*Correlation graph"**

Training the Model

We trained the system to employ the training set while keeping the test set available for evaluation after separating the cyber data into these two distinct subsets. "The data was split into a training set (consisting of around 70%) and a testing set (consisting of about 30%)". "This training method produces a training model based on the logic of the algorithm and the quantities of the elements in the training data".Since data drives model training, care has been taken to refrain from utilizing all the data. Scikit Learning is a Python library that maintains machine learning, training, and evaluating procedures. It provides a variety of straightforward and efficient tools that can handle numerous machine-learning tests.

Panda's data frames are supported by the Scikit Python library, which is a machine learning framework that relies on NumPy, SciPy, and Matplotlib. Data pre-processing, data selection based on importance, training the model, and model tuning are all part of this process.

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

```
In [11]: x = cyber_data. Drop (['Risk', 'State UT'], axis=1) # Independent
             Variables
         y = cyber_data[['Risk']]  # Dependent Variable

In [12]: Shape

Out [12]: (36, 18)
```

```
In [24]: # Splitting the data into train and test
         X_train, X_test, y_train, y_test = train_test_split (X, y, test_
         size=0.4, random_state=1)

In [14]: # Building Linear Regressionreg
         = LogisticRegression () reg.fit
         (X_train, _train)
         M: \Users\Cybercare\anaconda3\lib\site-packages\sklearn\utils\
         validation.py:760: DataConversionWarning: A column-vector y was
         passe d when a 1d array was expected. Please change the shape of yto
         (n_samples,), for example using ravel ().
         y = column_or_1d (y, warn=True)
         M: \Users\ Cybercare\anaconda3\lib\site-packages\sklearn\linear_
         model\_logistic.py:940: ConvergenceWarning: 1bfgs failed to
         converge (status=1):
         STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.
```

"Fig. 12. Applying regression algorithm"

Checking the Accuracy

Explanation for code":

Since we are aware that employing factors like righteous vengeance, anger, fraud, etc., our research sought to determine if a specific State/UT is more susceptible to cybercrime. We have

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

employed the statistical approach of logistic regression to predict this correlation. Before beginning a modelling project, it's crucial to find out if the independent variables are related (for example Personal revenge, Anger, Fraud, etc.)Since the link between Personal Vengeance and Personal Retribution is not of relevance, we can ignore the diagonal block in the correlation diagram. Correlation levels are shown on a continuum from light green (moderate) to dark green (low) to black (no correlation). When a link is particularly strong, the bar will be shaded yellow. The yellow blocks representing these variables suggest a possible relationship between sexual exploitation and rage, the spread of piracy, and practical jokes, among other things. Spreading pirate content and damaging reputations, executing practical jokes, and inciting patriotism are all negatively linked as if These variables' block, which is yellow in color, is evident in the plot. In a similar vein, we might assert that there is less non-correlation for the variables with darker blocks. The dependent variable is characterized by the letter y, while the independent variable is marked by the letters X. Our dependent variable is a risk, and our independent variables include personal vengeance, anger, fraud, etc.

So, after evaluating the shape of X (purely to make sure it made sense), We were able to train the logistic regression model with "X train" and "Y train", and then test it with "X test" and "Y test", because we had divided the variable into train and test.In this last step, we have applied the function to data from trains X and Y in order to create a logistic regression model. Logistic regression is utilized when the answer variable is a Boolean one (True/False, Present/Absent, etc.). With this model in hand, we can utilize the X test to anticipate y values. To test the efficacy of the model, we now look at the projected values for y. The accuracy of the model can be determined by counting the number of times the State/UT was accurately predicted to be in danger relative to the total number of times it was predicted to be safe. Even though we have often assured the public that the State/UT is not under danger, it was in fact threatened. Specifically, the model's ability to accurately estimate the threat in the states is depicted in Figures 11, 12, and 13.

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

```
Out [14]: LogisticRegression (C=1.0, class_weight=None, dual=False, fit_
intercept=True,
                    intercept scaling=1, 11_ratio=None, max_iter=100,
                    multi_class='auto', n_jobs=Nonc, pcnalty='12',
                    random_state=None, solver='1bfgs', tol=0.0001,
                    verbose=0,warm_start=False)

In [15]: # Predicting  the cases
          y_pred = reg.
          predicts(X_tcst)y_pred

Out |15|: array (|0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0|, dtype=int64)

In [18]:  Metrics.accuracy_score (y_test, y_pred)
Out [18]: 0.6666666666666666
In [19]:  cyber_data. head (5)

Out [19]:
```

Causing Disrepute	Prank	Sexual Exploitation	Political Motives	Terrorist Activities	Inciting Hate against Country	Disrupt PublicService	Sale purchase illegal drugs	Developing own business	Spreading	Psycho orPervert	Steal Information	Abetment to Suicide	Others	Risk
7	0	92	12	1	1	1	0	2	14	2	0	1	236	0
0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
234	0	113	9	4	3	0	0	0	0	0	0	0	832	1
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0
25	0	21	4	0	1	0	0	1	0	0	0	0	61	0

Fig.13.Model predicting the accuracy

Conclusion

We may infer from the study above that by experimenting with different algorithms and learning from our errors, we were able to use the Logistic Regression Algorithm to derive some important observations. This study might be useful for law enforcement officials in determining the root

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

cause of a crime if there was a political upheaval, natural calamity, or other significant dropout rate in the state where the crime was committed. We all know that the government and law enforcement can only impose the law on us, thus it's important to investigate the accused's upbringing, society, and teachings to better understand the motivation for the crime. Because we can't put all our faith in this approach, we end up with unfair sentences for the accused. But, the root of this criminal activity must be isolated and eliminated. The main question is how and to what degree technology will help the justice system. "[Sourdin T., 2018. Judge v. Robot: Artificial Intelligence and Judicial Decision- Making. UNSWLJ, 41, pp: 1114]". Many sorts of businesses stand to gain from this cutting-edge tech, so long as it isn't abused or put to harmful uses. "[Cath C., 2018: 1–8]". For this reason, the legitimacy of this approach can be understood by a court with proper information monitoring procedures in place.

References

- Alarie B., Niblett A. & Yoon A. (2018) How artificial intelligence will affect thepracticeoflaw.*University of Toronto Law Journal*, vol.68, supplement 1,pp.106–124.
- CaldwellM.etal(2020)AI-enabledfuturecrime.*CrimeScience*,no1,pp.1–13.
- Cath C. (2018) Governing artificial intelligence: ethical, legal and technicalopportunitiesandchallenges.*Phil.Trans.Royal.Society*,issue2133,pp.1–8.
- Chen H. et al (2003) COPLINK Connect: information and knowledge manage-mentforlawenforcement.*Decisionsupportsystems*,no3,pp.271–285.
- Cui Y. (2020) Building AI-assisted rule of law for the future, seeking advantagesand avoiding disadvantages to make AI better benefit mankind. In: *ArtificialIntelligenceandJudicialModernization*.Singapore:Springer,pp.187–191.
- Dhillon K. (2011) The police and the criminal justice system in India. *The Police,State,andSociety:PerspectivesfromIndiaandFrance*.Pearson,pp.27–59.
- DmitrikN.(2020)DigitalState,DigitalCitizen:MakingFairandEffectiveRulesforaDigital World.*LegalIssuesintheDigitalAge*, no1, pp.54–78.
- Furtado V. et al (2010) Collective intelligence in law enforcement–The Wiki-Crimessystem.*InformationSciences*,no1,pp.4–17.

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

- HauckR.etal(2002)UsingCoplinktoanalyzecriminal-justicedata.*Computer*,no3,pp.30–37.
- Hayward K., Maas M. (2020) Artificial intelligence and crime: A primer forcriminologists.*Crime,Media,Culture*,pp.1–25.
- Isaac W. (2017) Hope, hype, and fear: the promise and potential pitfalls ofartificialintelligenceincriminaljustice.*OhioSt.J.Crim.*L.,vol.15,p.543.
- King T., Aggarwal N., Taddeo M., Floridi L. (2020) Artificial intelligence crime:Aninterdisciplinaryanalysisofforeseeablethreatsandsolutions.*Scienceandengineeringethics*,no1,pp.89–120.
- MardaV.(2018)ArtificialintelligencepolicyinIndia:aframeworkforengagingthe limits of data-driven decision-making. *Philosophical Transactions of theRoyal Society: Mathematical, Physical and Engineering Sciences,* vol. 376,pp.1–19.
- McCarthyJ.etal(2006)AproposalfortheDartmouthsummerresearchprojectonartificialintelligence.*AImagazine*,no4,pp.12–14.
- McGuire M., Holt T. (eds.) (2017) The Routledge Handbook of Technology,CrimeandJustice.L.:Taylor&Francis,pp.1–722.
- Nath S. (2006) Crime pattern detection using data mining. In: 2006 IEEE/WIC/ACMInternationalConferenceonWebIntelligenceandIntelligentAgent,pp.41–44.
- Perkel J. (2018) Why Jupyter is data scientists' computational notebook ofchoice.*Nature*,vol.563,pp.145–147.
- Raschka S. et al (2020) Machine Learning in Python: Main developments andtechnologytrendsindatascience,machinelearning,andartificialintelligence.*Information*,no4,p.193.
- Sourdin T. (2018) Judge v. Robot: Artificial Intelligence and Judicial Decision-Making.UNSW*LawJournal*,vol.41,pp.11–14.
- Završnik A. (2020) Criminal justice, artificial intelligence systems, and humanrights.*ERAForumSpringer*,no4,pp.567–583.
- Duan Dai, S. B. et. a. (2021). A Review of Artificial Intelligence to Enhance the Security of Big Data Systems: State-of-Art, Methodologies, Applications, and

Vol. 13 Issue 3, March 2023,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Challenges. Springerlink. https://link.springer.com/article/10.1007/s11831-021-09628-0

- Hind Benbya, S. P. et. a. (2021). Artificial Intelligence in Organizations: Current State and Future Opportunities. SSRN, 19(04). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3741983
- Jamie Grace. (2019). Machine Learning Technologies and Human Rights in Criminal Justice Contexts. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3487454
- Thomas J. Holt. (2010). Exploring Strategies for Qualitative Criminological and Criminal Justice Inquiry Using On-Line Data. Taylor & Francis, 21(04). https://www.tandfonline.com/doi/abs/10.1080/10511253.2010.516565
- Yazan Mualla, A. N. et. a. (2019). Explainable Multi-Agent Systems Through Blockchain Technology. Springer Link. https://link.springer.com/chapter/10.1007/978-3-030-30391-4_3