Applying Data Mining Techniques in Predicting Index and non-Index Crimes

Allemar Jhone P. Delima

Abstract—**An increasing incidence of crime has led to the development and use of computer-aided diagnosis system, tools and methods in analyzing, classifying and predicting crimes. This paper clusters municipalities in Surigao del Norte using K-Means algorithm. This is instrumental in finding identical traits, patterns and values in categorizing municipalities with much, more, and most number of recorded index and non-index crimes from 2013-2017. Prediction of its occurrence for the year 2018-2022 was also provided using ARIMA(1,0,7) model. Results showed that Surigao City has the most number of recorded index and non-index crimes. Followed by the municipality of Placer, Claver, and Dapa of cluster 2 and Del Carmen of cluster 3. Further, physical injury, homicide, violation of special laws, car napping, reckless imprudence resulting to physical injury, and other non-index crimes has 26%, 25%, 25%, 24%, 24%, and 23% forecasted increase for the year 2018-2022 with the highest occurrence in years 2018, 2018, 2020, 2018, 2020, and 2020 respectively. Future researchers may utilize other data mining techniques supported by a better accuracy result.**

Index Terms—**Crime forecasting, prediction, arima, data mining, crime mining.**

I. INTRODUCTION

The Police agencies around the world are anxious and eager to reduce crime [1]. Crime is a result of actions which do not conform with the society's norm and moral values [2], condemned by the society for violation to any law. There is a continuous increase in the analysis of crime in the past decades hence, allowing many possibilities of studying and extracting crime data from different disciplines through different perspectives [3].

Data Mining, as the method of analyzing different type of data to extract interesting patterns and knowledge [4], is also used to discover critical information which can help local authorities detect crime [5] as well as predicting areas which have high probability for occurrence of crime and indicate crime prone areas [6] including the type of crimes.

The goal of this paper is to create clusters to identify as to what areas and municipalities within the province of Surigao del Norte established the highest recorded index and non-index crimes from 2013-2017. This is done, by using K-Means clustering algorithm. Data clustering is an unsupervised classification method used in creating groups of objects, or clusters, in such a way that objects in the same cluster are very similar and objects in different clusters are

quite distinct [7], [8]. Another aim of this paper is to predict time, volume and or likelihood of future occurrence of index and non-index crimes for 2018-2022 using ARIMA (1, 0, 7) model. Predicting volume of crimes and the years of its high occurrences is an essential strategy in alleviating future problems before the adversaries can have the chance to execute them. The police authority can identify and come up with a strategic actions and preventive maintenance if the forecasted occurrence of crime and its possible volume of occurrences is identified.

TABLE I: ACTUAL AND FORECASTED DATA OF INDEX CRIMES FROM 2013-2017 AND 2018-2022 RESPECTIVELY IN THE PROVINCE OF SURIGAO DEL NORTE

Year	MURDER	HOMICIDE	PHYSICAL INJURY	RAPE	ROBBERY	THEFT	CAR NAPPING	CATTLE RUSTLING
2013	69	57	911	49	338	821	74	2
2014	53	15	329	62	355	871	107	4
2015	49	13	272	72	278	490	140	5
2016	62	10	198	56	89	170	29	0
2017	57	10	145	52	62	109	23	0
2018	54	40	608	65	146	350	157	5
2019	57	26	478	64	285	575	121	4
2020	55	32	584	57	313	624	33	0
2021	56	22	464	59	257	555	76	2
2022	57	20	374	61	202	461	105	3

TABLE II: ACTUAL AND FORECASTED DATA OF non-INDEX CRIMES FROM 2014-2017 AND 2018-2022 RESPECTIVELY IN THE PROVINCE OF SURIGAO DEL NORTE

Year	RECKLESS IMPRUDENCE RESULTING TO			VIOLATION OF SPECIAL LAWS	OTHER NON INDEX CRIMES
	HOMICIDE	PHYSICAL INJURY	DAMAGE TO PROPERTY		
2014	37	365	297	1024	739
2015	42	283	145	772	570
2016	48	262	165	672	385
2017	39	213	153	506	256
2018	40	281	210	709	344
2019	45	297	146	785	484
2020	41	325	223	893	589
2021	41	293	165	801	574
2022	42	286	214	771	539

II. RELATED LITERATURE

Data mining, as coined with the term Knowledge Discovery in databases (KDD), used diverse approaches of DM analysis such as: decision tree, bayes classifiers, association rules, clustering, neural networks, genetic algorithms, support vector machines, predicting or forecasting and many more. These are instrumental in extracting important information from the database [9]. Along with DM analytics, big data and predictive analysis are relatively new concepts in criminology, while they have become standard practice in disciplines such as business intelligence, biomedical sciences, finance, and marketing. The use of such techniques will also inadvertently impact

social sciences and humanities in general [10] and criminology in particular [11] and is proved to be very useful as it is employed to perform in crime forecasting [12].

The paper of [5] analyzed the crimes such as theft, homicide and various drug offences along with suspicious activities, noise complaints and burglar alarm by using qualitative and quantitative approach. Using K-means clustering data mining approach on a crime dataset from New South Wales region of Australia, crime rates of each type of crimes and cities with high crime rates have been found.

Moreover, a method to execute a problem-oriented fitness function to predicted spatiotemporal patterns of criminal activity, using memetic differential fuzzy approach which is a unique time series approach from fuzzy clustering, in the city of San Francisco, USA was presented in the paper of [3]. Results show that series approach of fuzzy clustering for criminal patterns is a feasible method and is effective in producing a forecast of criminal patterns.

Furthermore, [13] implemented data mining techniques to understand certain trends and pattern of terrorist attacks in India. K-means clustering was used to determine the year wherein the terrorist groups were most active and also which terrorist group has affected the most. The experimental result is implemented in Rapidminer tool to determine the active group and the affected year.

As stated by [14], on the basis of data source, predicting crimes can be categorized in to two strategies. One approach is to use usual statistic techniques and models such as STL [15], ARIMA, kernel density distribution, and more in identifying hotspots of crime. Crime hotspots refers to geospatial locations where the probability for a crime to occur is high.

Meanwhile, precise crime prediction for small areas like police precincts is proposed in [16]. In [17], autoregressive integrated moving average (ARIMA) is employed for near future prediction of property crime. Based on 50 weeks property crime data, an ARIMA model was built to predict crime number of 1 week ahead. It is found that ARIMA model has higher fitting and prediction precision than exponential smoothing.

An approach based on auto-regressive models to reliably forecast crime trends in areas in Chicago was also performed. In particular, ARIMA as a model to forecast the number of crimes that is likely to occur in rolling time horizons was used to predict the number of crimes with an accuracy of 84% on one year-ahead forecasts and of 80% on two-year-ahead forecasts [18].

In the Philippines, a recent study of [19] clustered the index crime data in the province of Misamis Occidental using K-Means algorithm and forecasted its occurrence for the next five years using ARIMA(1,0,1) model. Based on the result, 3 of the indexed crime data were categorized in the first cluster and the other five index crimes are found to be in the cluster 2. Moreover, rape, cattle rustling, physical injury, robbery, and theft showed an increasing pattern based on the forecasted data from 2015 to 2020.

The study of [20] visualized forecasting algorithm performance using time series instance space. They proved that the ARIMA method does the best overall forecasting performance.

III. METHODOLOGY

The study was anchored on the concept of [19] but differs in many ways. Although K-Means algorithm was also used, it was utilized to find clusters within the city and municipalities in the province of Surigao del Norte to identify areas with much, more, and most number of crimes recorded. The ARIMA (1,0,7) model was utilized to predict future occurrence of index and non-index crimes in the province from 2018-2022.
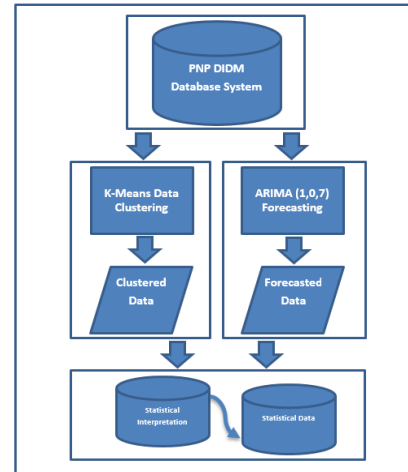

Fig. 1. Conceptual framework of the study.

A. Clustering

Clustering is an unsupervised data analyzing technique used to divide the same data into the same group and the different data into the other group [21]. Index and non-index crime datasets of the city and municipalities from the province of Surigao del Norte for the year 2013-2017 were used as presented in the Table III.

TABLE III: INDEX AND NON-INDEX CRIME DATASET PER CITY AND MUNICIPALITY IN THE PROVINCE OF SURIGAO DEL NORTE

MUNICIPALITY	INDEX CRIME	NON-INDEX CRIME
ALEGRIA	151	313
BACUAG	148	297
BURGOS	27	61
CLAVER	366	617
DAPA	222	437
DEL CARMEN	151	216
GENERAL LUNA	161	294
GIGAQUIT	183	264
MAINIT	194	303
MALIMONO	76	149
PILAR	98	138
PLACER	316	801
SAN BENITO	65	121
SAN FRANCISCO	103	298
SAN ISIDRO	88	111
SISON	120	308
STA. MONICA	85	185
SOCORRO	88	124
SURIGAO CITY	3562	3705
TAGANAAN	165	367
TUBOD	139	329

In K-means algorithm, the user specifies the k centroids. This K centroids refers to the number of the wanted clusters. Each cluster must have a centroid that is a mean of a cluster. Then a nearest centroid is assigned to each data record. When all input data records have been assigned, the centroid changes in each cluster and is updated by calculating the mean

cluster. These processes will be repeated until the latest centroids do not change [22].

The experimental result for clustering was implemented using KNIME (Konstanz Information Miner) [23] analytics platform. Fig. 2 shows the node structure of K-means clustering executed in KNIME. The node for the K-Means is connected and then positioned after the node of the imported csv file of the dataset. The node color manager comes after as it put distinctions to the results to be generated later. The node scatter plot shows the scatter plot of the clusters while the interactive table is used to view the result in a table manner
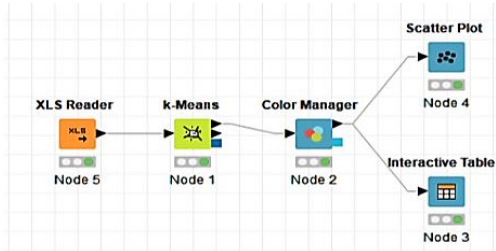
Fig. 2. Node structure of K-Means clustering in KNIME.

Surigao City, as the only place that belongs to Cluster 1, denotes that it does not possess any similar traits with the other places. It is evident in the Table IV that having the highest value makes the algorithm categorized it differently from the others. Municipalities under Cluster 2 is then interpreted as next to Surigao City to have the highest value of recorded crimes. Meanwhile, municipalities that are under Cluster 3 has the least among the groups.

TABLE IV: CLUSTERING RESULT IN KNIME

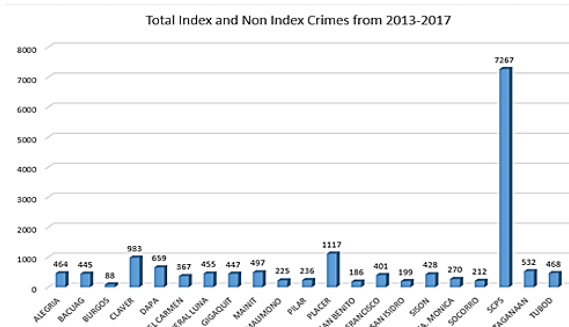| MUNICIPALITY	INDEX CRIME	NON-INDEX CRIME	CLUSTER
ALEGRIA	151	313	2
BACUAG	148	297	2
BURGOS	27	61	3
CLAVER	366	617	2
DAPA	222	437	2
DEL CARMEN	151	216	3
GENERAL LUNA	161	294	2
GIGAQUIT	183	264	2
MAINIT	194	303	2
MALIMONO	76	149	3
PILAR	98	138	3
PLACER	316	801	2
SAN BENITO	65	121	3
SAN FRANCISCO	103	298	2
SAN ISIDRO	88	111	3
SISON	120	308	2
STA. MONICA	85	185	3
SOCORRO	88	124	3
SURIGAO CITY	3562	3705	1
TAGANAAN	165	367	2
TUBOD	139	329	2

Fig. 3. Crime rates in Surigao del norte.

Fig. 3 shows that among the city and all municipalities in Surigao del Norte, Surigao City has the highest crime rate with a total of 7,267 recorded index and non-index crimes from the year 2013-2017.

In Fig. 4, violation of special laws is the top recorded crime in Surigao City. It is followed by other non-index crimes, theft, physical injury, and robbery. The least recorded crimes in the city are the cattle rustling, homicide, and reckless imprudence resulting to homicide.
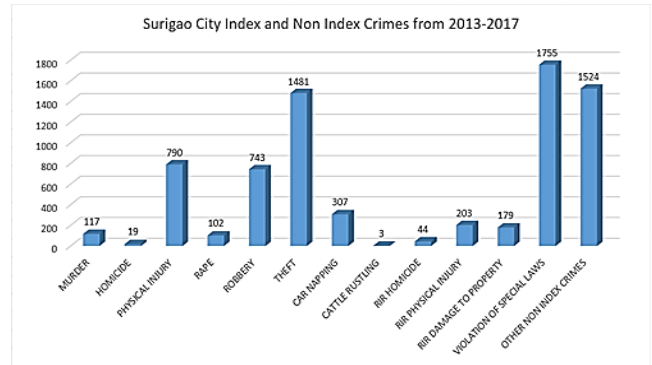

Fig. 4. Index and non-index crime rates in Surigao city.

Municipalities that belongs to Cluster 2 is shown in Fig. 5. The municipality of Placer has the highest recorded incident of index and non-index crimes next to Surigao City of the Cluster 1. It is followed by the municipality of Claver and Dapa making them as the second, third and fourth most number of recorded crime incidents from the whole province of Surigao del Norte from the year 2013-2017.

In Cluster 3, the municipality of Del Carmen has the most number of recorded index and non-index crimes from the year 2013-2017 followed by Sta. Monica and Pilar. The municipality of Burgos has the lowest number of recorded index and non-index crimes for the past five years as shown in Fig. 6.
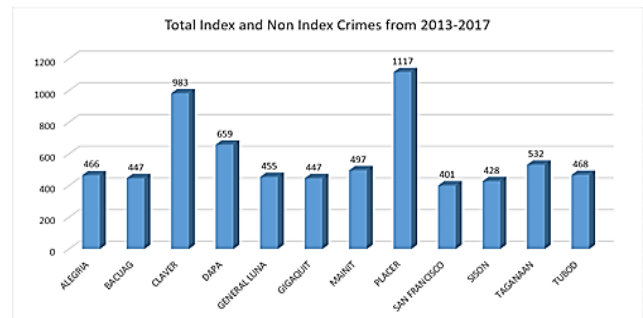

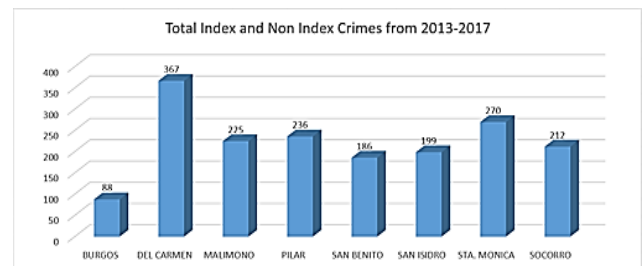Fig. 5. Crime rates of each municipality from cluster 2.

Fig. 6. Crime rates of each municipality from cluster 3.

B. Forecasting Index Crimes

The study [19] used ARIMA(1,0,1) model in forecasting

index crimes, while the ARIMA(1,0,7) model was used in this paper in determining the occurrence of index crimes for the next five years. Fig. 7 to Fig. 14 showed the graph of the predicted index crimes from 2018 to 2020 having 80% and 95% interval.

An autoregressive integrated moving average (ARIMA) model makes prediction of time series values based upon prior values (AR terms) as well as the errors made by previous predictions (MA terms). This allows the model to adjust itself to sudden changes in the time series. Therefore, the ARIMA forecasting equation for a stationary time series is a linear regression equation in which the predictors are the lags of the dependent variable and/or lags of the prediction errors. This model is explained in more detail in [18], [24]. In this paper, this method was implemented in R Studio using R language.

It is shown in Fig. 7 that there is a decrease of recorded murder cases in the province of Surigao del Norte since 2016 to 2017. The highest predicted rate of murder is 57 in the years 2019 and 2022. The total predicted murder case for the year 2018-2022 is 279 which is 4% lesser from the 290 actual murder cases from the past five years.

It is evident in Fig. 9 that there is a rapid increase rate of physical injury from 2017 to 2018 with crime data of 608 which is also considered as the highest. A decrease is forecasted in year 2018 to 2019 and an increase a year after. The total forecasted crime data of physical injury from 2018 to 2022 in the province of Surigao del Norte is 2,508 which is 26% higher from the past five years.
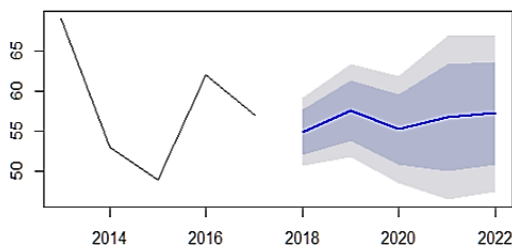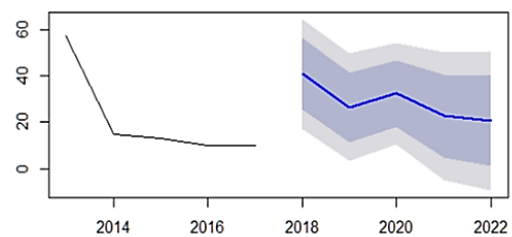
Fig. 7. Forecasted murder from 2018-2022.

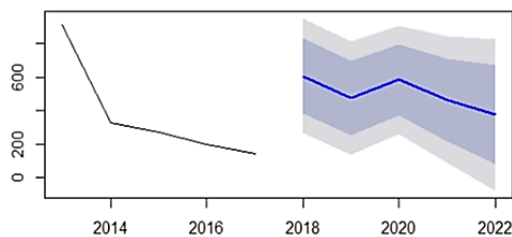Fig. 8. Forecasted homicide from 2018-2022.

Fig. 9. Forecasted physical injury from 2018-2022.

The total predicted homicide from 2018-2022 in the province of Surigao del Norte is 140 which is 25% higher

from the actual homicide data from the past five years. It can be seen in Fig. 8 that there is a predicted increase of homicide from year 2017 to 2018 but a decreasing trend from 2020 to 2022.

The total forecasted rape from 2018-2022 in the province of Surigao del Norte is 306 which is 5% higher from the actual rape data of year 2013-2017. It is forecasted that there is a considerable increase of rape cases from 2017 to 2018 and a decrease years after. A slightly increase pattern from year 2020 to 2022 is also evident in Fig. 10.
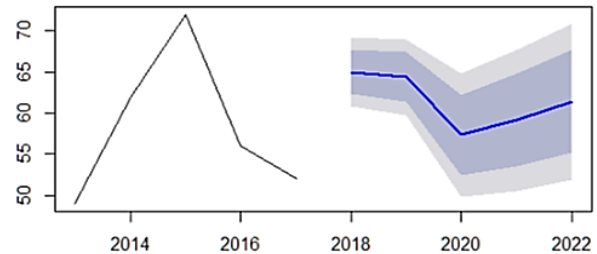
Fig. 10. Forecasted rape from 2018-2022.

It is presented in Fig. 11 that there is a decreasing trend of robbery from year 2013 to year 2017. Meanwhile, the forecasted crime data of robbery shows a rapid increase from 62 actual data for the year 2017 to 146 forecasted data in year 2018. An increasing and decreasing pattern is shown from 2017 to 2020 and 2020 to 2022, respectively. The total forecasted crime data of robbery from 2018-2022 in the province of Surigao del Norte is 1,203 which is 7% higher from the past.
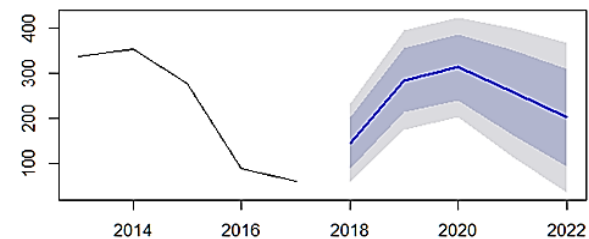
Fig. 11. Forecasted robbery from 2018-2022.

The total forecasted theft from 2018-2022 in the province of Surigao del Norte is 2,565 which is 4% higher from the actual theft data of 2,461 from 2013-2017. An increasing trend from 2017 to 2020 is evident in the Fig. 12.
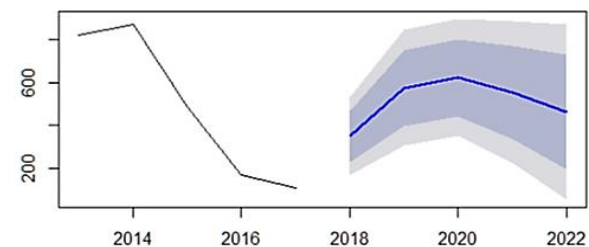
Fig. 12. Forecasted theft from 2018-2022.

It is shown in Fig. 13 that there is a rapid increase from 23 recorded data of car napping in 2017 to forecasted data of 157 in the year 2018. A slightly decrease pattern is shown from year 2018 all the way to year 2020. Meanwhile, an increasing pattern from 2020 to 2022 is also visible. The total forecasted

crime data of car napping from 2018-2022 in the province of Surigao del Norte is 492 which is 24% higher from the past five years.
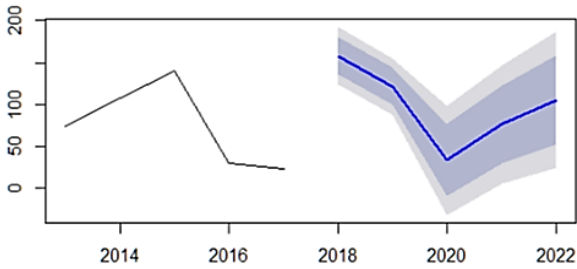

Fig. 13. Forecasted car napping from 2018-2022.

Among all the index crimes in the province of Surigao del Norte, the cattle rustling has the lowest number of crime reported from the year 2013-2017. The forecasted crime data of cattle rustling from 2018-2022 is 14 which has an estimated increase of 3 from the actual data of 11 from year 2013-2017. An increase from the year 2017 to 2018 and a decrease in the next two years is shown in the Fig. 14. Further, a slightly increase trend is evident in 2020-2022.
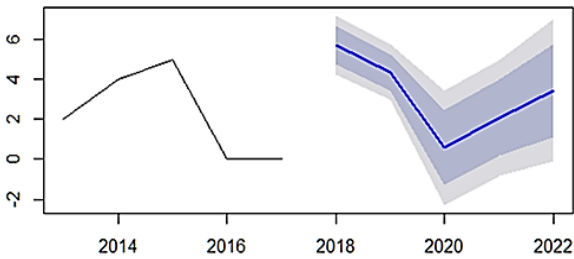

Fig. 14. Forecasted cattle rustling from 2018-2022.

C. Forecasting Non-Index Crimes

ARIMA(1,0,6) model was used to forecast the occurrences of non-index crimes for the next five years. Fig. 15 to Figure 19 showed the graph of the predicted non-index crimes from 2018 to 2020 having 80% and 95% interval.

It is shown in Fig. 15 that there is an increase of recorded reckless imprudence resulting to homicide cases in the province of Surigao del Norte since 2014 to 2016 and a decreasing pattern thereafter. The highest predicted rate of reckless imprudence resulting to homicide is 45 in the year 2019. Further, a decreasing pattern is shown from year 2019 to 2020 and a steady trend right after. The total predicted case for the year 2018-2022 is 209 which is 20% higher from the 166 actual cases from the past five years.
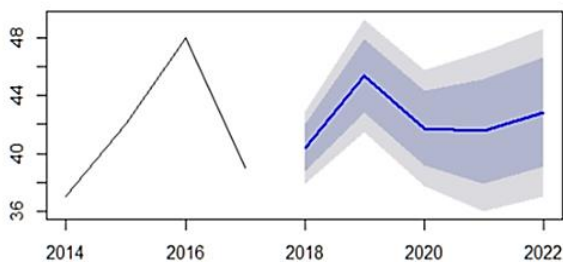

Fig. 15. Forecasted reckless imprudence resulting to homicide from 2018-2022.

The total predicted data for reckless imprudence resulting to physical injury from 2018-2022 in the province of Surigao del Norte is 1,482 which is 24% higher from the 1,123 actual data from the past five years. It can be seen in Fig. 16 that there is a predicted increase from year 2017 to 2020 but a decreasing trend from 2020 to 2022.
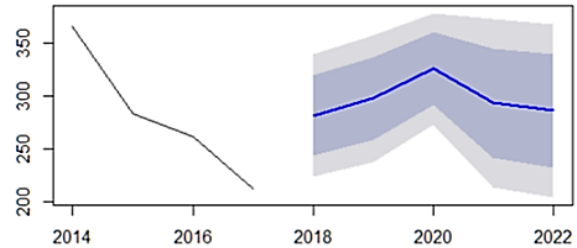

Fig. 16. Forecasted reckless imprudence resulting to physical injury from 2018-2022.

It is shown in Fig. 17 that there is almost a steady trend from 153 recorded data of reckless imprudence resulting to damage to property in 2017 to forecasted data of 210 in the year 2018. A decrease and increase pattern is shown from year 2018 to 2019 and 2019-2020 respectively and then repeat all the way to year 2022. The total forecasted crime data of reckless imprudence resulting to damage to property from 2018-2022 in the province of Surigao del Norte is 958 which is 20% higher from the 760 actual data from 2014-2017.
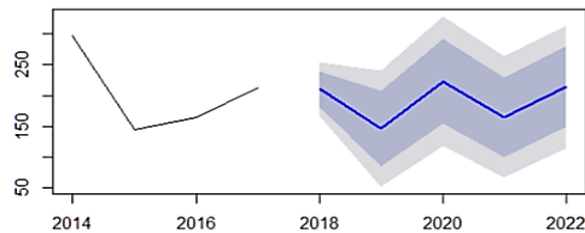

Fig. 17. Forecasted reckless imprudence resulting to damage to property from 2018-2022.

The predicted crime data for the violation of special laws from 2018-2022 in the province of Surigao del Norte is considerably high from the year 2017 to 2020 with forecasted 893 crime records. Meanwhile, there is a decreasing pattern from 2020 to 2022 as shown in Fig. 18. The total forecasted crime data of violation of special laws from 2018-2022 in the province of Surigao del Norte is 3,959 which is 25% higher from the 2,974 actual data from 2014-2017.
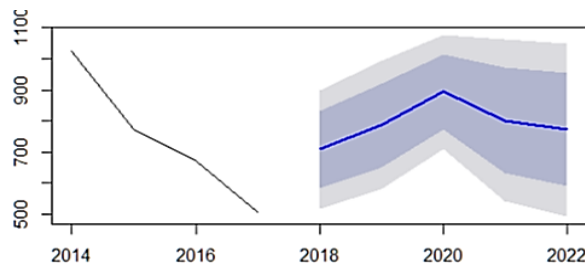

Fig. 18. Forecasted violation of special laws from 2018-2022.

It is presented in Fig. 19 that there is a decreasing trend of non-index crimes from year 2014 to year 2017. Meanwhile,

the forecasted crime data shows a rapid increase from 256 actual data for the year 2017 to 344 forecasted data in year 2018. An increasing and decreasing pattern is shown from 2017 to 2020 and 2020 to 2022, respectively. The total forecasted crime data of other non-index crimes from 2018-2022 in the province of Surigao del Norte is 2,530 which is 23% higher from the 1,950 actual data from 2014-2017.
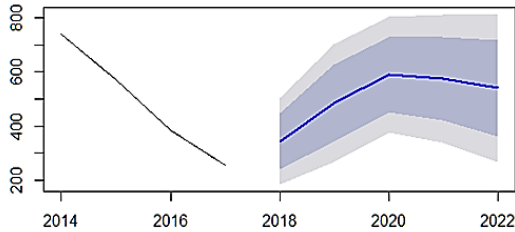

Fig. 19. Forecasted other non-index crimes from 2018-2022.

IV. CONCLUSION

With the use of K-Means clustering algorithm, determining the groupings of municipality with identical traits and values became possible. In cluster 1, the Surigao City topped as the municipality in Surigao del Norte with most number of reported index and non-index crimes. In cluster 2, the municipality of Placer, Claver and Dapa has the highest crime rate. Meanwhile in cluster 3, Del Carmen, Sta. Monica and Pilar was identified.

Among the index crimes in the province of Surigao del Norte, theft was identified as the highest number of recorded crime with a total of 2,565 from 2013-2017 with the highest occurrence in 2014. Furthermore, the highest predicted crime for the year 2018-2022 is the physical injury having the predicted value of 2,508 or 26% increase from 2014-2017. Moreover, the least reported crime in the province is cattle rustling.

For the non-index crimes, violation of special laws was identified as the highest reported incident in the province with the highest occurrence in 2014. Moreover, violation of special laws has the highest predicted value of 3,959 or 25% increase from the data of year 2014-2017 with the highest occurrence in 2020.

REFERENCES

[1] G. Dudfield, C. Angel, L. W. Sherman, and S. Torrence, "The 'power curve' of victim harm: targeting the distribution of crime harm index values across all victims and repeat victims over 1 year," *Cambridge J. Evidence-Based Polic.*, vol. 1, no. 1, pp. 38–58, 2017.
[2] M. Sevri, H. Karacan, and M. A. Akcayol, "Crime analysis based on association rules using apriori algorithm," *International Journal of Information and Electronics Engineering*, vol. 7, no. 3, 2017.
[3] C. D. R. Rodriguez, D. M. Gomez, and M. A. M. Rey, "Forecasting time series from clustering by a memetic differential fuzzy approach: An application to crime prediction," in *Proc. 2017 IEEE Symp. Ser. Comput. Intell.*, 2018, pp. 1–8.
[4] K. Rajalakshmi, S. S. Dhenakaran, and N. Roobini, "Comparative analysis of K-means algorithm in disease prediction," *Int. J. Sci. Eng. Technol. Res.*, vol. 4, no. 7, pp. 2697–2699, 2015.
[5] J. Agarwal, "Crime analysis using K-means clustering," *Int. J. Comput. Appl.* , vol. 83, no. 4, pp. 975–8887, 2013.
[6] O. Vaidya, S. Mitra, R. Kumbhar, S. Chavan, and R. Patil, *Comprehensive Comparative Analysis of Methods For Crime*, pp. 715–718, 2018.
[7] D. Kaur and K. Jyoti, *Enhancement in the Performance of K-means Algorithm*, vol. 2, no. 1, pp. 29–32, 2013.
[8] A. Bansal, M. Sharma, and S. Goel, "Improved K-mean clustering algorithm for prediction analysis using classification technique in data mining," *Int. J. Comput. Appl.*, vol. 157, no. 6, pp. 975–8887, 2017.
[9] E. Şuşnea, "Using data mining techniques in higher education," *High. Educ.*, vol. 1, no. 1, pp. 68–72, 1996.
[10] R. Kitchin, "Big Data, new epistemologies and paradigm shifts," *Big Data Soc.*, vol. 1, no. 1, p. 205395171452848, 2014.
[11] J. Chan and L. B. Moses, "Is big data challenging criminology?" *Theor. Criminol.*, vol. 20, no. 1, pp. 21–39, 2016.
[12] C. Yu, M. W. Ward, M. Morabito, and W. Ding, *Crime Forecasting Using Data Mining Techniques*, 2011.
[13] P. Gupta, A. S. Sabitha, and T. Choudhury, Terrorist Attacks Analysis Using Clustering Algorithm, © Springer Nat. Singapore Pte Ltd., pp. 317–328, 2018.
[14] J. Azeez and D. J. Aravindhar, "Hybrid approach to crime prediction using deep learning," in *Proc. 2015 Int. Conf. Adv. Comput. Commun. Informatics*, 2015, pp. 1701–1710.
[15] A. Malik, R. Maciejewski, S. Towers, S. Mccullough, and D. S. Ebert, "Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement," *IEEE Trans Vis Comput Graph.*, vol. 20, no. 1, pp. 1863–1872, 2014.
[16] W. Gorr, A. Olligschlaeger, and Y. Thompson, "Short-term forecasting of crime," *Int. J. Forecast.*, vol. 19, no. 4, pp. 579–594, 2003.
[17] P. Chen, H. Yuan, and X. Shu, "Forecasting crime using the ARIMA model," in *Proc. 5th Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD*, 2008, vol. 5, pp. 627–630.
[18] E. Cesario, C. Catlett, and D. Talia, "Forecasting crimes using autoregressive models," in *Proc. 2016 IEEE 14th Int. Conf. Dependable, Auton. Secur. Comput.*, pp. 795–802, 2016.
[19] M. Y. Orong, A. M. Sison, and A. A. Hernandez, "Mitigating vulnerabilities through forecasting and crime trend analysis," in *Proc. 2018 5th Int. Conf. Bus. Ind. Res.*, 2018, pp. 57–62.
[20] Y. Kang, R. J. Hyndman, and K. Smith-Miles, "Visualising forecasting algorithm performance using time series instance spaces," *Int. J. Forecast.*, vol. 33, no. 2, pp. 345–358, 2017.
[21] A. Ben Ayed, M. Ben Halima, and A. M. Alimi, "Survey on clustering methods: Towards fuzzy clustering for big data," in *Proc. 6th Int. Conf. Soft Comput. Pattern Recognition*, 2015, pp. 331–336.
[22] A. Thammano and A. K. Algorithm, *Enhancing K-means Algorithm for Solving Classification Problems*, pp. 1652–1656, 2013.
[23] L. Feltrin, "KNIME an open source solution for predictive analytics in the geosciences [software and data sets]," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 4, 2015.
[24] A. Rege *et al.*, "Predicting adversarial cyber intrusion stages using autoregressive neural networks," *IEEE Intell. Syst.*, 2018.

Allemar Jhone P. Delima was born in Surigao City, Surigao del Norte, Philippines on March 24, 1994. He completed master in information technology degree at Surigao State College of Technology, Surigao City, Philippines in 2016. With his quest in improving his professional career, he took doctoral degree and is now a student of doctor in information technology at Technological Institute of the Philippines, Quezon City, Philippines. He is currently connected in Surigao State College of Technology as a faculty in the College of Engineering and Information Technology. His research interest is in data mining and data analytics.