This journal provides immediate open access to its content under the Creative Commons BY 4.0 license. Authors who publish with this journal retain all copyrights and agree to the terms of the above-mentioned CC BY 4.0 license.

DOI: 10.2478/seeur-2023-0064

ANALYSIS AND VISUALIZATION OF ROAD ACCIDENTS USING HEATMAPS BASED ON WEB DATA

Luan Sinanaj
Faculty of Contemporary Sciences and Technologies,
South East European University, North Macedonia
ls30441@seeu.edu.mk

Lejla Abazi Bexheti
Faculty of Contemporary Sciences and Technologies,
South East European University, North Macedonia
l.abazi@seeu.edu.mk

ABSTRACT

Road accidents have increased rapidly in recent years for a variety of reasons. Analyzing and visualizing road accidents through heatmaps can help improve policies for their prevention by informing about areas with a high-risk of road accidents.

The purpose of this research is to build a model for the analysis and visualization of road accidents through heatmaps. Information about road accidents is extracted from the news of the main online media portals through scripts in the Python language and Web Scraping techniques. From the extraction of about 30,000 articles from news portals for one year, only 829 were selected in the end that provided information about road accidents.

As a result, and contribution of this research, a corpus was built with the geographic coordinates of road accidents and on this data our model was applied for the analysis and visualization of high-risk areas of road accidents using heatmaps. The visualization of heatmaps was done through a Python script, where it was applied to the geographic coordinates of road accidents.

Key words: Heatmaps, Road Accidents Mapping, Road Accidents Analysis, Web Scraping.
JEL classification: C3, C6

INTRODUCTION

Nowadays, the increase in the use of vehicles, heavy traffic and non-compliance with traffic rules are very worrying problems causing road accidents. In the previous 2 articles (Sinanaj & Bexheti, 2023) and (Sinanaj, Bedalli, & Bexheti, 2023), studies have been made about the various causes of road accidents and their degree of danger in Albania, based on the data obtained on the Web from the news posted by the main online media. As a continuation of the two previous works, part of the focus of this work is the detailed study of the regions and locations of road accidents in order to visualize them in the simplest and most understandable forms, such as heatmaps.

Based on the above problems about road accidents, it is necessary to have good knowledge about the regions and locations of road accidents by presenting them in heatmaps, thus, their prevention by the responsible state institutions can be made easier. Analysis of the areas and addresses of road accidents using heatmaps is also necessary for citizens, not only for state institutions.

From a review of the literature, there are no studies that predict road accidents using heatmaps in the state of Albania. The only information that is published about road accidents in the state of Albania is that of the Ministry of the Interior (Ministria e Brendshme, 2022) and (INSTAT, 2022), which consist of simple statistics.

The main purpose of this research is to analyze and visualize road accidents using heatmaps based on data obtained from the news of the main online media portals. So, the data of the regions and locations of road accidents will be taken from the main online media portals. Knowing these heatmaps of road accidents can provide quick information for citizens to be more careful in these areas, but these heatmaps can also help to undertake road accident prevention or awareness campaigns by state authorities or the various associations that they work to help the citizens and state institutions.

For this research, the methodology followed is presented as follows: (1) first, extracting regions and locations of road accidents through a Python script with Web Scraping techniques from news articles on online portals and saving them in a CSV file (Sinanaj & Bexheti, 2023); (2) secondly, the preparation of a corpus of data containing the geographical coordinates of all the regions and locations of road accidents that have been taken at the above point and saving

them in another CSV file; (3) thirdly, based on the file of the locations of road accidents, through a script in the Python programming language, the visualization and analysis of the areas with the highest-risk of road accidents is done by presenting them through heatmaps.

The related research questions raised for this case study are as follows:

- RQ1: Is it possible to extract the regions and locations of traffic accidents from the news of the main online media portals and present them on geographical maps?
- RQ2: Is it possible to analyze and visualize the locations of traffic accidents using heatmaps?
- RQ3: Based on the visualization of road accidents through heatmaps, which are the areas analyzed to have the highest rate of road accidents?
- RQ4: In which periods of the year are there the highest risk and number of road accidents?
- RQ5: In which regions are there the highest risk and number of road accidents?

As a contribution to this research, we have:

- the construction of a corpus that contains all the geographical coordinates of the regions and locations of road accidents in the state of Albania;
- Also, the construction of a script in the Python programming language using appropriate libraries for the analysis and visualization of road accidents through heatmap. The corpus prepared with the coordinates of the road accident locations mentioned above was used to train the script to achieve the final result of creating a road accident heatmap.

The organization and structure of this paper are as follows: In Section 2, the review of the literature related to the analysis and visualization of road accident areas, In Section 3, the methodology followed for this research is described, and in Section 4, the results are explained and discussed. In the last two sections, you will find the conclusions and references of this research.

LITERATURE REVIEW

According to a thorough review of scientific papers for the state of Albania, there is no research on the analysis or prediction of the reasons that lead to more accidents or studies on the combination of essential factors that lead to more accidents. The only information available regarding traffic accidents in Albania comes from the Ministry of the Interior (Ministria e Brendshme, 2022) and INSTAT (ISTAT, 2022), although it is insufficient and incomplete.

For the prediction of road accidents, scientific research has been done even in the earliest decades. The author (Kononov, 2002) employed the pattern recognition algorithm and direct diagnostic methods for the modeling of the prediction of road accidents and the diagnosis of the causes of accidents. For example, in the author's (Larsen, 2004) research, he adopted a multidisciplinary approach to provide accurate knowledge about the factors that influence road accidents.

One of the points of reference for the idea of our research has been the study of the authors (Lnenicka, Hovad, Komarkova, & Pasler, 2016), who created an application with Web scraping techniques for extracting information about crimes that have occurred from online portals and presenting them on geographic maps according to high crime areas.

In the last decade, many studies have been done in different countries about road accidents. Many data mining models and algorithms have been used and tested in many studies for the classification and prediction of road accidents. The authors (Yan & Shen, 2022) in their study have achieved high results using a hybrid model that integrates Random Forest and Bayesian Optimization compared to conventional algorithms. However, according to the authors' study (Chen & Chen, 2020), Random Forest proved to be the most effective model for predicting the severity of road accidents among the three compared models, such as Logistic Regression and Classification and Regression Tree. The same conclusion was reached in our previous study (Sinanaj, Bedalli, & Bexheti, 2023), where the Random Forest algorithm turned out to be more performant between the K-Nearest Neighbor and Naïve Bayes Classifier.

In the next references, there are studies done for the country of Bangladesh and that of India. The authors (Siddik, 2021), for the prediction of road accidents, have tested on 1237 data four classification models such as Decision Tree, K-Nearest Neighbors, Naïve Bayes and Logistic Regression. In another study to predict the number of road accidents and their casualties in Bangladesh by the authors (Biswas, Mia, & Majumder, 2019), the results showed that random forest regression is good for predicting and applying in that context. Also, in the state of India from the study of the authors (Mahendra & Roopashree, 2023), for the prediction of road accidents, where they were based on four factors such as the type of collision, the type of road, the location and the weather, it resulted that the Random Forest regressor works better than other compared models.

Many scientific studies have also been done for the prediction of road accidents using heatmaps. The study by the authors (Silva & Saraee, 2019) shows that integrating classification

approaches such as decision trees and time-series calendar heatmaps might be a helpful tool for properly identifying roadside traffic incidents based on injury severity.

The authors' (Szénási, Felde, Kertész, & Nádai, 2018) study proposes a revolutionary image processing technique for quantifying heatmap raw data. Using the morphological function known as erosion, it is possible to provide a more precise separation between black spots and safe places.

The system created by the authors (Feng, Zheng, Ren, & Liu, 2020) consists of three parts: first, they cluster the accident incidents in an interactive Google map to highlight some hotspots; then, they narratively visualize accident attributes to uncover potentially related factors; and finally, they explore several state-of-the-art machine learning, deep learning methods, as well as time series forecasting methods, to predict the number of road accidents in future years.

After geocoding the crash locations on a map, the authors (Xiuqin Chen, 2018) assessed the spatial clustering of accidents and hotspots' spatial densities using Moran's I method, kernel density estimation, Ripley's K-function, and network-based kernel density estimation.

The authors (Sreedhar, P, & Mathew, 2021) used heatmaps, boxplots, and histograms to predict road accidents. Aside from that, they used Geojson to show accident-prone locations, extending the capabilities of Google Maps.

In the authors' article (Saldana-Perez, Torres-Ruiz, & Moreno-Ibarra, 2019), a model is proposed to predict traffic congestion by applying a support vector machine method and through the visualization of heatmaps for Mexico City.

And finally, the authors' article (Netek, Pour, & Slezakova, 2018) describes an exploratory study that examines traffic accident data. Their experimental goal was to provide basic and important recommendations for predicting road accidents by applying heat maps.

METHODOLOGY

In this section, the methodology followed for the realization of this research is explained, along with all the stages that have been passed by referring to the two previous related works, the method of data collection and preparation, as well as the model applied to these data to achieve the objectives of this research. The methodology followed in this research was divided into two important parts:

1- data pre-processing, and
2- the application of a model for the analysis and visualization of road accidents using heatmaps.

Data pre-processing

One of the most important parts of this research has been the preprocessing of the data, which was then applied to the model implemented by us for the analysis and visualization of road accidents through heatmaps, which is also the purpose of our study. The difficulties during this study have been numerous: the first one related to the preparation of a corpus containing all the geographic coordinates of the locations of road accidents extracted from the news of online portals required a long time to prepare and to check manually, and the second one is that the scientific papers and studies related to the processing of the text of the Albanian language or studies dealing with road accidents using geographic coordinates or heatmaps are very limited or non-existent.

During the data pre-processing, the following steps were followed, where the first two steps are based on my two previous studies (Sinanaj & Bexheti, 2023) and (Sinanaj, Bedalli, & Bexheti, 2023) to provide the necessary dataset for this study:

1. During this phase, several online news portals are selected to extract the articles posted daily with Web Scraping techniques by building a Python script with the BeautifulSoup library. During this process, about 30,000 news articles were extracted and saved in a CSV file. The data of the news articles stored in the CSV file are as follows: Number, Date, Title of the article and body of the article.
2. After the completion of phase 1, the second important phase was the selection of all the articles that talked about the accidents that happened in the state of Albania, and this process was carried out using a Python script. The construction of the script had its own difficulties and challenges since there are no libraries for the processing of the text in the Albanian language. It is important to note that this script has been trained by building two data corpora, one that contains key words of the Albanian language related to road accidents, and another that contains all the regions of the state of Albania. By running the script on about 30,000 news articles, about 1,200 articles that talk only about road accidents or other accidents have been filtered and saved in a CSV file. The traffic accident file contains the following information: Number, Date, Region, Location, Type of Road, Injured-Dead, Weight-Accident, Cause and Gender. In order to have the most qualitative data in our study, a very important process during the data preprocessing phase was the manual control of the data and from this selection process we managed to have about 829 articles in the final data set. During the manual process, redundant and repetitive articles about the same accident

and articles unrelated to road accidents were eliminated. This process was long and tiring, but it was very important to have the final data of the highest quality.

3. After the completion of phase 2, the next very important step was the construction of a data corpus containing all the geographical coordinates of the locations of road accidents. The locations of road accidents were taken from the final CSV file of step 2. The construction of this corpus with geographic coordinates was a challenge in itself as the manual work required relatively much time. This was one of the most important phases of this research because the dataset of geographic coordinates of road accidents will be used by our model for the analysis and visualization of the areas with the highest risk of road accidents by presenting them in heatmaps.

Figure 1 below shows schematically all the processes of the three steps described above to have as final data all the geographic coordinates of road accidents based on the news of online portals in the state of Albania.

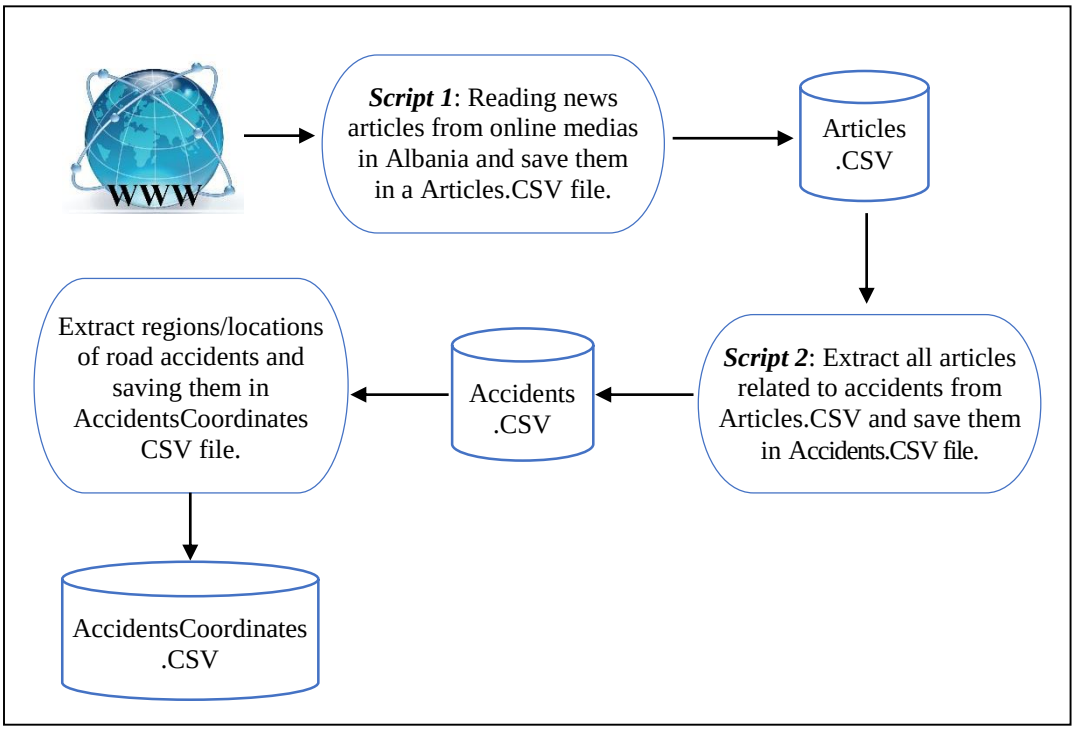


Figure 1: Data Pre-processing Model.

By running script 1, the online portal articles are read and saved to an Articles CSV file. Meanwhile, script 2 reads all the articles from the Articles file and filters those that talk about road accidents, saving them in another CSV file called Accidents. The final process in the figure is the product of the corpus that contains the geographic coordinates of road accidents, saving them in another CSV file called AccidentsCoordinates.

The data variables generated during pre-processing are listed in the table below. These variables are divided into seven categories, namely: Date, Region, Location, Longitude, Latitude, Road Type and Gravity. The Gravity variable is taken from the previous study (Sinanaj, Bedalli, & Bexheti, 2023) and indicates the degree of danger of the accident starting from 1 (low) to 8 (high). The Gravity variable according to the road accident that occurred takes the following values: it takes the value 1 when the road accident was light without involving other cars; it takes the value 2 when the road accident involved at least two cars; it takes the value 3 when the road accident involved also one injured; it takes the value 4 when there were also two injured in the road accident; it takes the value 5 when there were more than two injured in the road accident; it takes the value 6 when there was also one person dead in the road accident; it gets the value 7 when there were two people dead in the road accident; and finally it takes the value 8 when there were more than two dead people in the road accident. The detailed descriptions of these variables are presented in Table 1. The data from these variables have served as the input for the model applied for the analysis and visualization of road accidents through heatmaps.

Table 1: Variables generated during the data pre-processing stage.

VARIABLES	DESCRIPTION
DATE	The date of the publication of the road accident
REGION	The region of the road accident
LOCATION	Possible location or address of the accident
LONGITUDE	The geographic coordinate that indicates the east-west position point on the Earth's surface.
LATITUDE	The geographic coordinate that indicates the east-west position point on the Earth's surface.
TYPE OF ROAD	Interurban, Urban, Rural, Highway
GRAVITY	value = 1 (light risk) to value = 8 (high risk 8)

Application of a model for the analysis and visualization of road accidents using heatmaps

After the successful completion of the data pre-processing phase, where the final output was the AccidentsCoordinates file that contains all the geographic coordinates of the locations of road accidents in the state of Albania, we moved to the next important phase, which was the creation of a Python script for the analysis and visualization of road accidents using heatmaps. Figure 2 below shows the implementation of this model. In the implementation of the model in Figure 2, script 3 reads the geographic coordinates of road accidents from the file AccidentsCoordinates, processes them and as output produces heatmaps of road accidents that become clearer and more understandable for everyone. In order to achieve the results of the presentation of road accidents on geographical maps, the Python language library and package

(Folium, 2023) were used. In a few words and simply, script 3 reads the geographic coordinates of road accidents from the file AccidentsCoordinates and generates as output an HTML file with the analysis and visualization of road accidents through heatmaps.

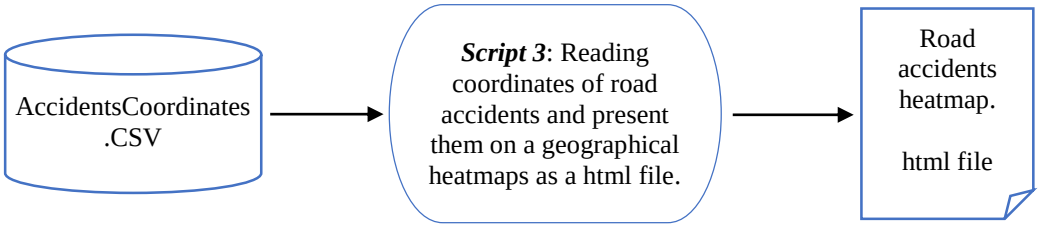


Figure 2: Heatmap Generation Model.

RESULTS & DISCUSSIONS

By implementing the steps of the methodology followed for this research, which is described in the section above, the data preprocessing phase was first performed. The final result of this phase is Table 3, which contains all the geographic coordinates of the regions and locations of road accidents. The geographic coordinates of road accidents are expressed through two variables: Longitude and Latitude. For the generation of geographic coordinates, it has been a real challenge since a manual check and work has been done to have the highest quality data. Also, as a result of this phase, data was generated for each accident, including date, region, location, type of road and gravity (refer to Table 2). In order to generate the data in Table 3, the data of road accidents in Table 2 were taken as a reference. Although in Table 2 the data are presented as an example, in reality, there are about 829 records of road accidents processed.

Table 2: Road accident data results during the data pre-processing stage.

Nr	Date	Region	Location	Type of Road	Gravity
1	31 12 2021 16:15	durres	aksi rrugor Maminas Durrës	interurban	4
2	30 12 2021 19:34	puke	pukë	urban	5
3	30 12 2021 18:47	tirane	Rrugën e Arbrit, Bastar i Mesëm	interurban	4
⋮	⋮	⋮	⋮	⋮	⋮
827	03 01 2021 08:08	lezhe	Lezhë Milot, pranë kryqëzimit të Ishull Shëngjin	interurban	6
828	02 01 2021 14:10	durres	lagja nr 7	urban	3
829	01 01 2021 17:04	librazhd	rrugë kombëtare Librazhd Përrenjas, “Lapidari i Xhyrë”	interurban	5

Table 3: The output of the geographic coordinates of road accidents during the data pre-processing.

Nr	Longitude	Latitude
1	41,37340500	19,60640200

2	42,04414000	19,90018000
3	41,43262500	19,98763300
⋮	⋮	⋮
827	41,77044100	19,63310600
828	41,31887000	19,44513000
829	41,11502200	20,41963900

After the first phase of data pre-processing, we moved to the second phase, where script 3 (see figure 2) was applied to the data in table 3 and as a result generated a heatmap of road accidents (refer to figure 3). The result of the analysis and visualization of road accidents through the heatmap also answers the research questions raised at the beginning.

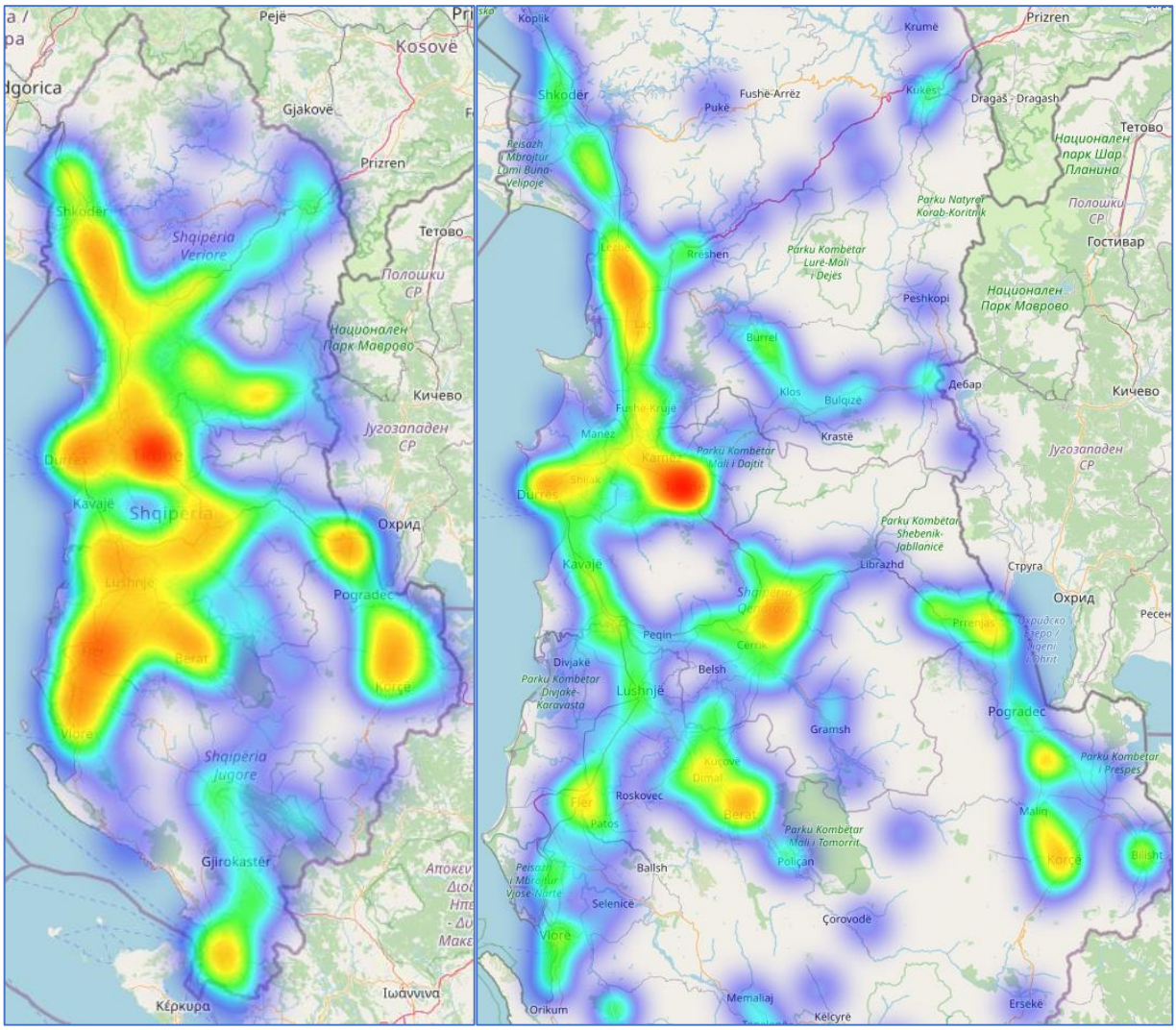


Figure 3: Heatmaps of road accident areas in the state of Albania.

Regarding research questions 1 and 2, with the help of the results of our previous research in the reference (Sinanaj, Bexheti, 2023), it was possible to find the regions, the locations of road accidents and to present them on geographical maps. The results are presented

in tables 2, 3 and in Figure 3 which presents geographically the analysis and visualization of the heatmap of road accidents. Table 2 presents the model of the results of road accidents for 829 records, where for each accident, the Number, Date, Region, Location, Type of Road and Gravity are presented. The data of this table were generated from our previous study in reference to (Sinanaj, Bexheti, 2023). Whereas, from the Location variable of Table 2, were generated the geographic coordinates of the roads where the accidents happened, and those data are presented in Table 3. Whereas, based on the geographic coordinates data in Table 3, it was generated the heatmap of Figure 3, was generated, which is also explained in detail in the following paragraph. Since this study was new in its kind for the state of Albania, we cannot compare it with other scientific works because no other similar research results from the review of the literature. These results were difficult to compare with the data published by the state institutions since they only publish basic statistics for the number of accidents that occurred without including other necessary details.

The danger of road accidents of areas in the state of Albania is depicted in the heatmap using colors ranging from green to dark red, indicating those with the lowest risk to those with the highest risk. According to the interpretation of the heatmap in Figure 3 related to RQ3, as you can see on the map, the areas with the highest risk of road accidents are those in Tirana, Elbasan, those in the direction from Tirana to Durres, those in the direction from Fushe-Kruje to Shkodra and those in the direction from Fier to Vlore and this is shown by the red color. In order to generate the result of the heatmap in Figure 3, script 3 was executed on the data of the geographical coordinates as presented in the example in Table 3. The information presented in the heatmap can be useful for citizens to be informed and more careful in these areas.

Referring to research questions RQ4 and RQ5, the results and the answer are presented below in tables 4 and 5. The time periods with the highest number and risk of road accidents are those in the months of December, November, August and October (table 4). Meanwhile, the regions with the highest number of road accidents listed are those in Tirana, Durres, Elbasan, Vlore, Fier, Lezhe, Berat, Korce, Fushe Kruje, Lushnje, Librazhd, Gjirokaster, Pogradec, Shkoder, etc. (Table 5).

Table 4: The results of the period times with the highest number of road accidents.

The number of road accidents by period times			
Season	**Months**	**Number of Accidents**	**Total**
Winter	December	95	217
	January	62	
	February	60	

Spring	March	54	173
	April	47	
	May	72	
Summer	June	71	215
	July	62	
	August	82	
Autumn	September	57	224
	October	80	
	November	87	
			829

Table 5: The results of the regions with the highest number of road accidents.

The number of road accidents by region				
Region	**Accident Numbers**		**Region**	**Accident Numbers**
Tirane	164		Cerrik	4
Durres	69		Lac	4
Elbasan	68		Maliq	4
Vlore	50		Mamurras	4
Fier	42		Mirdite	4
Lezhe	42		Perrenjas	4
Berat	36		Thumane	4
Korce	33		Burrel	3
Fushe Kruje	26		Klos	3
Lushnje	20		Ksamil	3
Librazhd	19		Kurbin	3
Gjirokaster	18		Tepelene	3
Pogradec	18		Vore	3
Shkoder	15		Belsh	2
Kruje	13		Finiq	2
Kamez	11		Kelcyre	2
Milot	11		Kucove	2
Sarande	11		Levan	2
Gramsh	8		Puke	2
Rreshen	8		Shengjin	2
Shijak	8		Sukth	2
Bulqize	7		Bilisht	1
Kukes	7		Dajt	1
Kavaje	6		Delvine	1
Peqin	6		Frakull	1
Divjake	5		Kashar	1
Maminas	5		Kolonje	1
Rrogozhine	5		Konispol	1

CONCLUSION

Nowadays, a series of factors, such as the increase in the use of vehicles, heavy traffic or non-compliance with traffic rules have increased the number of road accidents. Based on the above-mentioned problems, the need to analyze or predict and prevent road accidents becomes essential. The analysis and visualization of the areas with the highest risk of road accidents using heatmaps based on the data obtained from the news articles of the main online media was an innovation for the state of Albania (Figure 3). Knowing the regions and locations

through heatmaps can help citizens to be more careful, as well as state institutions to take preventive measures in the most problematic areas.

The purpose of this research was to apply a model for visualizing and predicting road accidents using heatmaps based on data obtained from the news of the main online media portals. Reading and saving news articles from online media portals were done using Web Scraping techniques. The objectives of this study were successfully achieved and in tables 2, 3, 4 and 5, as well as in figure 3, there are answers to the research questions raised at the beginning. Figure 3 shows the analysis of road accidents using heatmaps. Meanwhile, tables 4 and 5 present the regions and time periods with the highest number and risk of road accidents.

REFERENCES

1. Biswas, A. A., Mia, M. J., & Majumder, A. (2019). Forecasting the Number of Road Accidents and Casualties using Random Forest Regression in the Context of Bangladesh. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). Kanpur, India: IEEE. doi:10.1109/ICCCNT45670.2019.8944500
2. Chen, M.-M., & Chen, M.-C. (2020). Modeling Road Accident Severity with Comparisons of Logistic Regression, Decision Tree and Random Forest. *Information*, pp. 11(5), 270. doi: https://doi.org/10.3390/info11050270
3. Feng, M., Zheng, J., Ren, J., & Liu, Y. (2020). Towards Big Data Analytics and Mining for UK Traffic Accident Analysis, Visualization & Prediction. *12th International Conference on Machine Learning and Computing (ICMLC '20)* (pp. 225–229). New York, NY, USA: ACM. doi: https://doi.org/10.1145/3383972.3384034
4. Folium. (2023, August). *Folium — Folium 0.12.1 documentation*. Retrieved May 2022, from https://python-visualization.github.io/folium/
5. INSTAT. (2022). *Transporti, Aksidentet dhe Karakteristikat e Mjeteve Rrugore*. (Instituti i Statistikave - Tiranë) Retrieved May 2022, from http://www.instat.gov.al/al/temat/industria-tregtia-dhe-sh%C3%ABrbimet/transporti-aksidentet-dhe-karakteristikat-e-mjeteve-rrugore
6. Kononov, J. (2002). Road accident prediction modeling and diagnostics of accident causality: A comprehensive methodology. Denver: University of Colorado at Denver.
7. Larsen, L. (2004). Methods of multidisciplinary in-depth analyses of road traffic accidents. *Journal of Hazardous Materials, 111*(1-3), 115-122.
8. Lnenicka, M., Hovad, J., Komarkova, J., & Pasler, M. (2016). A proposal of web data mining application for mapping crime areas in the Czech Republic. *10th International Joint Conference on Software Technologies (ICSOFT)* (pp. 1-6). Colmar, France: IEEE.
9. Mahendra, G., & Roopashree, R. (2023). Prediction of Road Accidents in the Different States of India using Machine Learning Algorithms. *2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1-6). Raichur, India: IEEE. doi:10.1109/ICICACS57338.2023.10099519
10. Ministria e Brendshme. (2022, December). *Raporti Mujor (Monthly Report)*. (Ministry of Interior in Albania) Retrieved May 2022, from https://mb.gov.al/en/raporti-mujor/
11. Netek, R., Pour, T., & Slezakova, R. (2018, August 15). Implementation of Heat Maps in Geographical Information System – Exploratory Study on Traffic Accident Data. *Open Geosciences, 10*(1), pp. 367-384. doi: https://doi.org/10.1515/geo-2018-0029
12. Saldana-Perez, M., Torres-Ruiz, M., & Moreno-Ibarra, M. (2019). Geospatial Modeling of Road Traffic Using a Semi-Supervised Regression Algorithm. *IEEE Access, 7*, pp. 177376-177386. doi:10.1109/ACCESS.2019.2942586
13. Siddik, M. A. (2021). Predicting the Death of Road Accidents in Bangladesh Using Machine Learning Algorithms. *ICACDS 2021: Advances in Computing and Data Sciences* (pp. 160–171). Springer, Cham. doi: https://doi.org/10.1007/978-3-030-88244-0_16
14. Silva, C., & Saraee, M. (2019). Predicting Road Traffic Accident Severity using Decision Trees and Time-Series Calendar Heatmaps. *2019 IEEE Conference on Sustainable Utilization and Development in Engineering and Technologies (CSUDET)* (pp. 99-104). Penang, Malaysia: IEEE. doi:10.1109/CSUDET47057.2019.9214709
15. Sinanaj, L., & Bexheti, L. A. (2023). Predicting Road Accidents with Web Scraping and Machine Learning Techniques. *6th International Scientific Conference on Business and Economics (ISCBE 2023).* Tetovo, North Macedonia.

16. Sinanaj, L., Bedalli, E., & Bexheti, L. A. (2023). A Classification Model for Predicting Road Accidents Using Web Data. *ENTerprise REsearch InNOVAtion Conference - ENTRENOVA.* Dubrovnik, Croatia.

17. Sreedhar, M. T., P, K. K., & Mathew, L. S. (2021, Mar). Road Traffic Accident Analysis and Visualization of Accident Prone Areas. *International Journal for Research in Applied Science & Engineering Technology (IJRASET), 9*(III), pp. 552-561. doi:https://doi.org/10.22214/ijraset.2021.33280

18. Szénási, S., Felde, I., Kertész, G., & Nádai, L. (2018). Road Accident Black Spot Localisation using Morphological Image Processing Methods on Heatmap. (pp. pp. 000251-000256). Budapest, Hungary: IEEE. doi: doi: 10.1109/CINTI.2018.8928248

19. Xiuqin Chen, L. H. (2018, Feb 12). Hotspots of road traffic crashes in a redeveloping area of Shanghai. *International Journal of Injury Control and Safety Promotion, 25*(3), pp. 293-302. doi:10.1080/17457300.2018.1431938

20. Yan, M., & Shen, Y. (2022). Traffic Accident Severity Prediction Based on Random Forest. *Sustainability*, pp. 14(3), 1729. doi: https://doi.org/10.3390/su14031729