International Journal of Engineering and Techniques - Volume 10 Issue 3, June 2024

CLUSTERING FOR CRISIS : AMBULANCE OPTIMIZATION IN ROAD ACCIDENTS

[1]SK.Asiff, [2]Y.V.Ramesh

[1]Associate Professor, [2]Assistant Professor, [1,2]Department of Computer Science & Engineering, Geethanjali Institute of Science and Technology, Gangavaram, Andhra Pradesh, India

Abstract

The number of casualties and fatalities resulting from road accidents rises a significant global concern. To reduce this, the project advocates for a proactive approach by pre-positioning ambulances, thereby reducing response times and ensuring timely medical assistance. Leveraging deep learning techniques, particularly deep-embedded clustering, the project proposes a novel method to predict optimal ambulance locations. Recognizing the relationships between various geographical factors and accident occurrences. The project underscores the importance of preserving such patterns during model construction. This is achieved by integrating Cat2Vec.It is a deep-learning-based model. Comparative analysis with traditional clustering algorithms like K-means, GMM, and Agglomerative clustering highlights the more efficiency of the proposed approach. Additionally, a novel scoring function is introduced to assess real-time response time and distance calculation, further enhancing algorithm performance evaluation. The proposed ambulance positioning system demonstrates remarkable performance, resulting in a 95% accuracy rate with k-fold cross-validation, surpassing results obtained from traditional algorithms. Overall, the project underscores the effectiveness of the deep-embedded clustering-based approach in optimizing ambulance positioning and improving emergency response systems.

Introduction

The project focuses on revolutionizing emergency response systems by optimizing ambulance positioning for road accidents. Recognizing the critical importance of swift medical assistance in reducing casualties and fatalities, the project employs advanced data-driven methodologies to address this pressing challenge. This project predicts optimal ambulance locations by analyzing historical accident data and geographical factors.By preserving geographical patterns through innovative approaches like Cat2Vec integration, the project ensures accurate model construction and efficient ambulance deployment strategies. Moreover, a novel scoring function enhances real-time performance evaluation, further validating the effectiveness of the proposed approach.Through comparative analysis with traditional clustering algorithms, the project demonstrates superior efficiency and accuracy, highlighting the transformative potential of data-driven solutions in emergency response optimization. Ultimately, the project aims to significantly reduce response times, minimize casualties, and save lives in road accidents, making a profound impact on public safety and welfare globally.

Machine Learning

Machine learning (ML) is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks.

Machine learning involves computers discovering how they can perform tasks without being explicitly programmed to do so. It involves computers learning from data provided so that they carry out certain tasks. For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed.

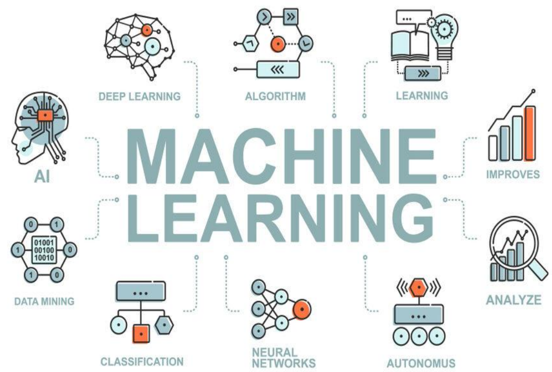


Fig.1. Machine Learning

Machine Learning Approaches

Early classifications for machine learning approaches sometimes divided them into three broad categories, depending on the nature of the "signal" or "feedback" available to the learning system. The Fig 1.2 represents Types of Machine Learning.

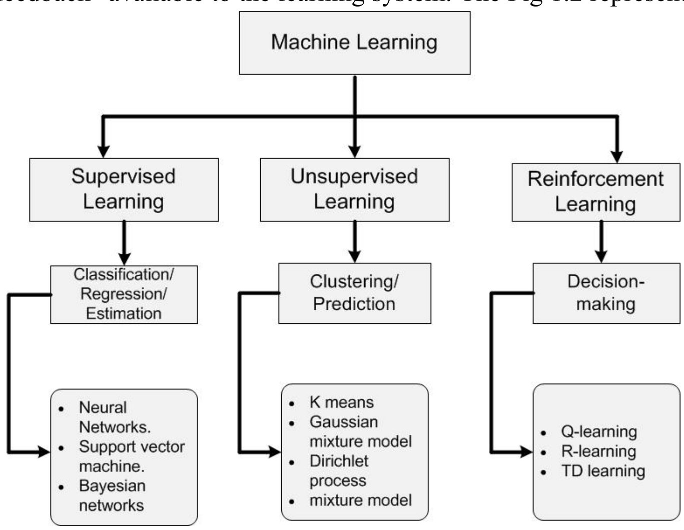


Fig.2. Types of Machine Learning

Supervised Learning: The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs. Some popular examples of supervised machine learning algorithms are: Linear regression for regression problems.

Unsupervised Learning: No labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end (feature learning). For instance, it might group data by temperature or similar weather patterns.

Reinforcement Learning: A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle or playing a game against an opponent) as it navigates its problem space, the program is provided feedback that's analogous to rewards, which it tries to maximize.

Objective

The primary objective of the above project is to optimize ambulance positioning for road accidents, aiming to enhance emergency response systems and ultimately reduce casualties and fatalities resulting from accidents. This optimization involves predicting the most suitable locations to position ambulances based on historical accident data and various geographical factors. By deploying ambulances strategically, the project aims to minimize response times, ensuring timely medical assistance to accident victims. Ultimately, the overarching goal is to leverage data analytics and machine learning to revolutionize emergency response systems, thereby saving lives and improving public safety on the roads.

Literature Review

From Clustering to Cluster explanations via Neural Networks.

" J.Kauffmann, M.Esders, L.Ruff, G.Montavon ,W.Samek, K.Muller"

The study aimed at proposing a new framework that can explain cluster assignments in terms of input features in an efficient and reliable manner. It is based on the novel insight that clustering models can be rewritten as neural networks or "neuralized." Cluster predictions of the obtained networks can then be quickly and accurately attributed to the input features. Several showcases demonstrate the ability of our method to assess the quality of learned clusters and to extract novel insights from the analyzed data and representations.The quantitative evaluation showed that the explanation method is capable of identifying cluster-relevant input features in a precise and systematic manner, from the simplest k-means model to some of the most recent proposals, such as the SCAN deep clustering model.

Predicting crash injury severity with machine learning algorithm synergized with clustering technique: A promising protocol.

"K. Assi, S. M. Rahman, U. Mansoor, and N. Ratrout"

This study developed machine learning (ML) models to predict crash injury severity using 15 crash-related parameters. Separate ML models for each cluster were obtained using fuzzy c-means, which enhanced the predicting capability.Features that were easily identified with little investigation on crash sites were used as an input so that the trauma center can predict the crash severity level based on the initial information provided from the crash site and prepare accordingly for the treatment of the victims.The SVM-FCM model outperformed the other developed models in terms of accuracy and F1 score in predicting the severity level of severe and non-severe crashes. The FNN had the least accuracy and F1 score values.

Analyzing factors associated with fatal road crashes: A machine learning approach.

"A. J. Ghandour, H. Hammoud, and S. Al-Hajj"

In this study, the authors proposed a model that adopts a hybrid ensemble machine learning classifier structured from sequential minimal optimization and decision trees to identify risk factors contributing to fatal road injuries. The model was constructed, trained, tested, and validated using the Lebanese Road Accidents Platform (LRAP) database of 8482 road crash incidents, with fatality occurrence as the outcome variable. A sensitivity analysis was conducted to examine the influence of multiple factors on fatality occurrence.Evidence gained from the model data analysis will be adopted by policymakers and key stakeholders to gain insights into major contributing factors associated with fatal road crashes and to translate knowledge into safety programs and enhanced road policies.Further research warrants the analysis of additional road injury crashes to enhance the model's prediction performance and the accuracy of its estimated results, as well as to reveal additional factors that contribute to the increased risk of sustaining fatal road injuries.

Explainable K-means and K-medians clustering

"M. Moshkovitz, S. Dasgupta, C. Rashtchian, and N. Frost"

Many clustering algorithms lead to cluster assignments that are hard to explain, partially because they depend on all the features of the data in a complicated way. To improve interpretability, the authors consider using a small decision tree to partition a data set into clusters, so that clusters can be characterized in a straightforward manner. They studied this problem from a theoretical viewpoint, measuring cluster quality by the k-means and k-medians objectives. In terms of negative results, they showed that popular top-down decision tree algorithms may lead to clusterings with arbitrarily large costand any clustering based on a tree with k leaves must incur an log k approximation factor compared to the optimal clustering. On the positive side, for two means/medians, they showed that a single threshold cut can achieve a constant factor approximation, and they gave nearly-matching lower bounds; for general they designed an efficient algorithm that leads to an O(k) approximation to the optimal k-medians and an O(k^2) approximation to the optimal k-means.

Traffic accident's severity prediction: A deep-learning approach-based CNN network.

"M. Zheng, T. Li, R. Zhu, J. Chen, Z. Ma, M. Tang, Z. Cui and Z. Wang"

In traffic accident, an accurate and timely severity prediction method is necessary for the successful deployment of an intelligent transportation system to provide corresponding levels of medical aid and transportation in a timely manner. The existing traffic accidents severity prediction methods mainly use shallow severity prediction models and statistical models. To promote the prediction accuracy, a novel traffic accidents severity prediction-convolutional neural network (TASP-CNN) model for traffic accidents severity prediction is proposed that considers combination relationships among traffic accidents features. Based on the weights of traffic accidents features, the feature matrix to gray image (FM2GI)algorithm is proposed to convert a single feature relationship of traffic accidents data into gray images containing combination relationships in parallel as the input variables for the model. Moreover, experiments demonstrated that the proposed model for traffic accidents severity prediction has a better performance.

Traffic Risk Mining From Heterogeneous Road Statistics

"K. Moriya, S. Matsushima, and K. Yamanishi"

In this paper, the authors proposed a novel framework for mining traffic risk from such heterogeneous data. Traffic risk refers to the possibility of occurrence of traffic accidents. Specifically, they focused on two issues: 1) predicting the number of accidents on any road or at intersection and 2) clustering roads to identify risk factors for risky road clusters. They presented a unified approach for addressing these issues by means of feature-based non-negative matrix factorization (FNMF). In particular, they developed a new multiplicative update algorithm for the FNMF to handle big traffic data. Using real-traffic data in Tokyo, they demonstrated that the proposed algorithm can be used to predict traffic risk at any location more accurately and efficiently than existing methods, and that a number of clusters of risky roads can be identified and characterized by two risk factors. In summary, their work can be regarded as the first step to a new research area of traffic risk mining.

PROPOSED SYSTEM:

Deep Embedded Clustering:

Deep Embedded Clustering (DEC) is a machine learning technique that combines deep learning and clustering algorithms to perform unsupervised learning tasks. It was introduced as a way to learn useful representations from data while simultaneously clustering similar data points together.

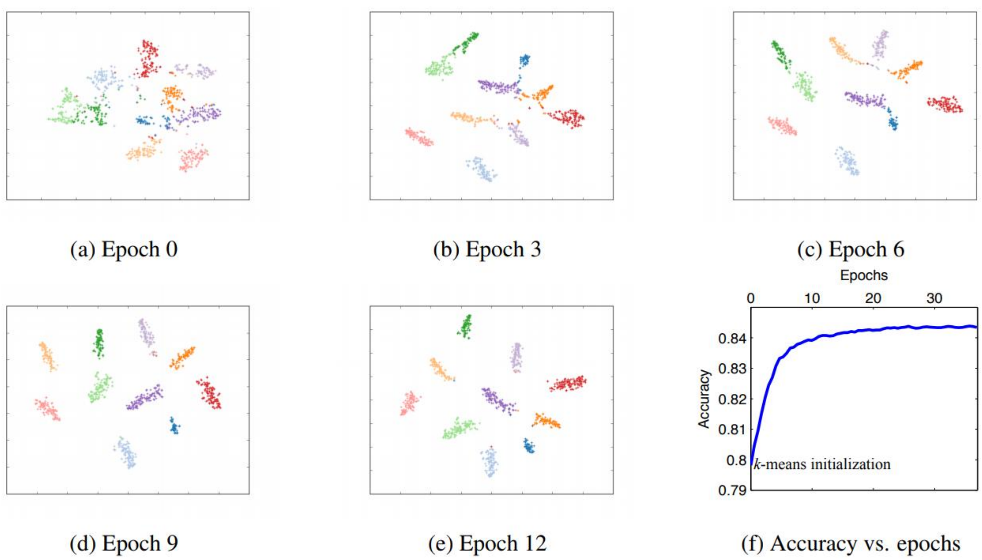


Fig.3. Deep Embedded Clustering

In this we use, Deep Embedded Clustering using Cat2Vec and Novel Distance Scoring.

Cat2Vec:

Cat2Vec is a deep learning-based embedding technique.It transforms categorical variables into deep embeddings. Preserves relationships between categories and patterns in data.Cat2Vec enhances the performance of the ambulance positioning framework. It improves the efficiency of the clustering model.Using Cat2Vec a low-dimensional continuous vector is automatically learned for each category in each field.

Novel Distance Scoring:

Novel distance scoring technique used to score or measure the dissimilarity or similarity between data points in a novel or innovative way.In various fields like machine learning, statistics, and data analysis, distance metrics play a crucial role in quantifying the similarity or dissimilarity between observations or features.It could involve designing custom distance functions tailored to specific datasets or problem domains, incorporating domain knowledge or heuristics, or leveraging advanced techniques such as deep learning for feature representation learning.

System Architecture

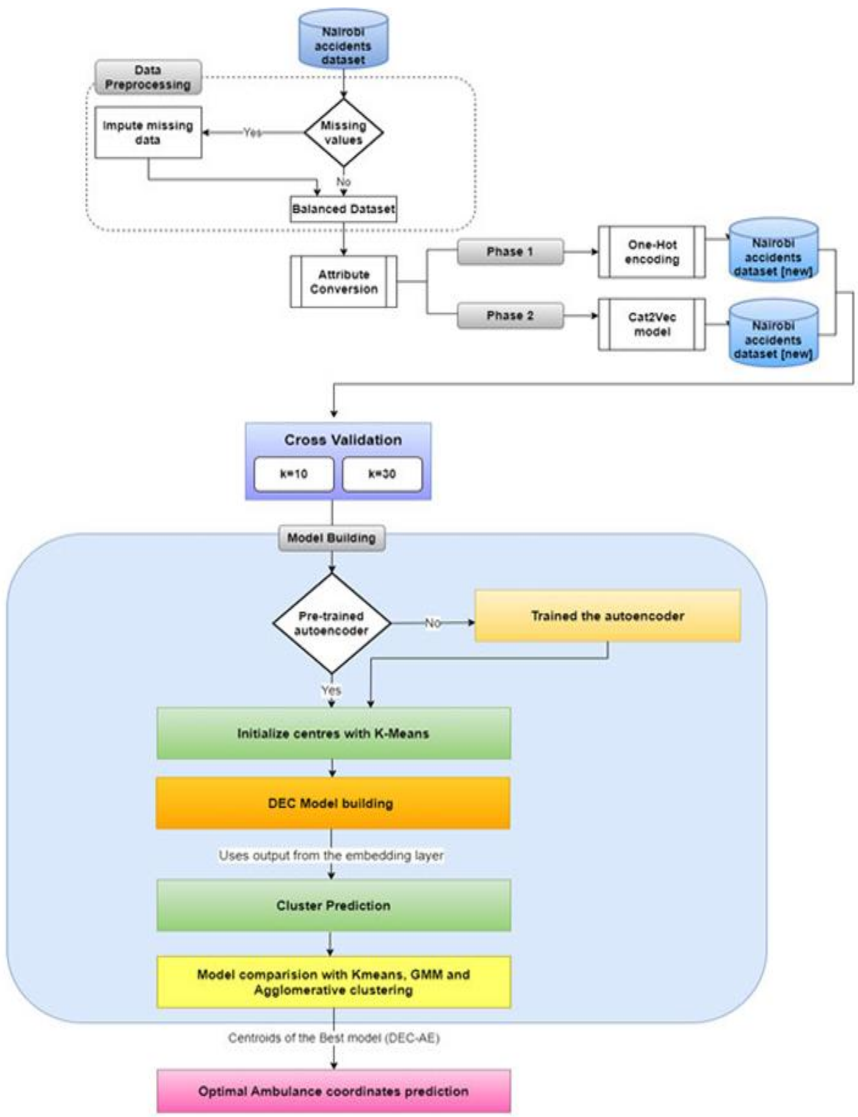


Fig.4. System architecture

The architecture for the optimal ambulance positioning is shown. Firstly, Dataset named Nairobi is taken.

Data Collection and Preprocessing:

Collection of road accident data: This involves gathering data related to road accidents, including information such as accident locations (latitude and longitude), timestamps, severity, and other relevant features.

Data preprocessing: Cleaning the data, handling missing values, feature engineering, and transforming the data into a suitable format for modeling.

Feature Extraction and Selection:

Extracting relevant features: Identifying and extracting features from the raw data that are deemed important for ambulance positioning, such as accident locations, traffic patterns, weather conditions, etc.

Feature selection: Choosing the most relevant features to be used in the modeling process. This step may involve techniques such as mutual information, correlation analysis, or domain knowledge.

Deep Embedded Clustering

Deep learning model architecture: Designing a deep neural network architecture suitable for embedding the accident data.

Training the model: Using the accident data to train the deep embedded clustering model. This involves optimizing the model parameters to minimize a clustering loss function while simultaneously learning meaningful representations of the data.

Optimal Ambulance Positioning

Determining ambulance positions: Using the clustered accident data and other relevant factors, such as population density, traffic patterns, and ambulance availability, to determine optimal locations for positioning ambulances.

Evaluation: Assessing the effectiveness of the proposed ambulance positioning strategy using metrics such as response time, distance traveled, and accuracy in predicting accident locations.

Comparison with Traditional Methods

Comparison with traditional clustering algorithms: Evaluating the performance of the deep embedded clustering approach against traditional clustering algorithms like K-means, GMM, and Agglomerative clustering.

Performance evaluation: Comparing the proposed method with traditional approaches in terms of accuracy, efficiency, and scalability.

System integration and Deployment

Integration with real-time systems: Developing a system that integrates the optimized ambulance positioning strategy into real-time emergency response systems.

Deployment considerations: Addressing practical considerations such as computational efficiency, scalability, and adaptability for deployment in real-world scenarios.

DATASET

We have datasets named:

Nairobi Accidents Dataset(2018-19), Train dataset, Segment_info, Sample Submission, k-Means.

Nairobi Accident Dataset:

Date	precipitable_water_entire_atmosphere	relative_humidity_2m_above_ground	specific_humidity_	temperature_	u_component_	v_component_of_wind_10m_above_ground
01-01-2018	24.10000038	72.70000458	0.00956	15.161493	-0.01218628	-0.9339
02-01-2018	27.36228752	74.90000153	0.010462524	16.047998	0.314543456	-0.4855
03-01-2018	30.24661827	86.59999847	0.01193	15.838525	-0.28273192	-0.654
04-01-2018	33.01068878	90.90000153	0.01278	16.169244	0.21157226	0.90812
05-01-2018	27.25037193	82.41335297	0.011391523	15.923456	1.047275424	-0.0377
06-01-2018	24.30000114	81.30000305	0.01066	15.124994	-0.36769041	-1.7204
07-01-2018	23.80000114	74.73505402	0.009574572	14.816248	0.544003904	-0.8133
08-01-2018	20.60062981	64.90000153	0.00818	14.496667	-0.00381592	-1.8618
09-01-2018	21.7869873	72.59728241	0.009780956	15.587091	-0.60039306	-0.6937
10-01-2018	22	68.07369232	0.008876207	15.062982	-0.88420653	-1.481
11-01-2018	27.70000076	83.37598419	0.011370239	15.711389	-1.52468622	-0.7495
12-01-2018	24.60000038	72.30000305	0.008959999	14.288538	-0.07512451	-0.9005
13-01-2018	23.41394615	76.90000153	0.00937	13.981531	-2.30096674	-1.7738
14-01-2018	23.39999962	71.40000153	0.00862	13.847192	-1.32664299	-1.212
15-01-2018	19.89999962	49.90000153	0.00739	17.023248	-1.1594702	-0.633
16-01-2018	18.10000038	50.5	0.00657	14.979639	0.038762204	-1.5346
17-01-2018	18.11301231	78.09999847	0.00942	13.808466	-0.92187256	-0.7401
18-01-2018	20.93750572	81.90000153	0.01048	14.711938	-1.53145754	-1.8585
19-01-2018	20.68883705	69.5	0.00895	14.814081	-0.96940428	-1.5953
20-01-2018	23.39999962	91.40000153	0.012120004	15.234064	-2.39337873	-1.3531
21-01-2018	18.81626892	69.90000153	0.00893	14.713586	-1.175354	-2.1108
22-01-2018	14.57213116	62.90000153	0.00792	14.4414	-1.82892334	-2.5337
23-01-2018	18.40672112	56.70000076	0.00769	15.599298	-1.7039355	-1.794
24-01-2018	19.75696373	77.5	0.009610163	14.217249	-2.01549792	-2.303
25-01-2018	23.51301384	86.5	0.011351202	15.077936	-3.32868886	-2.5414

Fig.5. Nairobi Accident Dataset

This dataset contains 730 samples. This has the weather in Nairobi when the accident occurred. It contains details such as Date, Precipitable water entire atmposhere, Relative humidity, Specific Humidity,Temperature.

Train Dataset

uid	datetime	latitude	longitude
1	01-01-2018 00:25	-1.18884981	36.93138244
2	01-01-2018 02:02	-0.66293876	37.20873
3	01-01-2018 02:31	-0.66293876	37.20873
4	01-01-2018 03:04	-1.28808711	36.8265834
5	01-01-2018 03:58	-1.18884981	36.93138244
6	01-01-2018 04:04	-0.66293876	37.20873
7	01-01-2018 05:31	-1.16522821	36.9600077
8	01-01-2018 07:57	-1.30130261	36.8209033

Fig.6. Train Dataset

This train dataset contains the accident occurred datetime, its location namely latitude and longitude. This dataset contains 6300 samples.

K-Means Dataset

	date	A0_Latituc	A0_Longit	A1_Latituc	A1_Longit	A2_Latituc	A2_Longit	A3_Latituc	A3_Longit	A4_Latituc	A4_Longit	A5_Latituc	A5_Longitude
0	01-07-2019 00:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
1	01-07-2019 03:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
2	01-07-2019 06:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
3	01-07-2019 09:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
4	01-07-2019 12:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
5	01-07-2019 15:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
6	01-07-2019 18:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
7	01-07-2019 21:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
8	02-07-2019 00:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
9	02-07-2019 03:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
10	02-07-2019 06:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
11	02-07-2019 09:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
12	02-07-2019 12:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
13	02-07-2019 15:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
14	02-07-2019 18:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
15	02-07-2019 21:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498
16	03-07-2019 00:00	-1.27938	36.86373	-2.68201	37.43737	-1.06138	37.03223	-1.47827	37.06678	-1.25588	36.73792	-2.11581	37.46498

Fig.7. K-Means Dataset

This dataset contains the locations of the 6 ambulances that need to be placed at certain location to have optimal results when accident occurs on that date and time. This dataset contains 1400 samples.

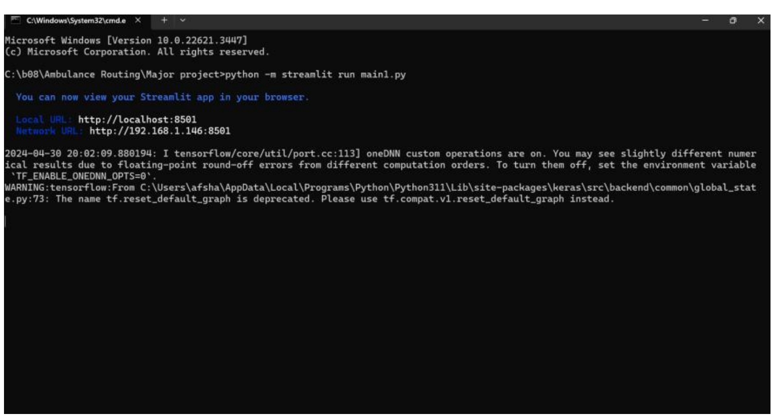
Baseline Static Dataset

Date	precipitabl	relative_h	specific_h	temperatu	u_compon	v_compon	precipitabl	relative_h	specific_h	temperatu	u_compon	v_compon	3h	year	weekend/	dayofweel	month
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	0	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	3	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	6	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	9	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	12	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	15	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	18	2019	0	1	1
01-01-2019	20.1864	85.1	0.01032	13.91317	-2.6281	-3.06291	mid	high	mid	low	mid	low	21	2019	0	1	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	0	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	3	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	6	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	9	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	12	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	15	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	18	2019	0	2	1
02-01-2019	19.26901	92.96189	0.01147	14.1799	-1.9874	-2.12117	mid	high	high	low	mid	low	21	2019	0	2	1
03-01-2019	18.9	86.20886	0.011421	15.28896	-2.55027	-2.69084	mid	high	high	mid	mid	low	0	2019	0	3	1
03-01-2019	18.9	86.20886	0.011421	15.28896	-2.55027	-2.69084	mid	high	high	mid	mid	low	3	2019	0	3	1
03-01-2019	18.9	86.20886	0.011421	15.28896	-2.55027	-2.69084	mid	high	high	mid	mid	low	6	2019	0	3	1
03-01-2019	18.9	86.20886	0.011421	15.28896	-2.55027	-2.69084	mid	high	high	mid	mid	low	9	2019	0	3	1
03-01-2019	18.9	86.20886	0.011421	15.28896	-2.55027	-2.69084	mid	high	high	mid	mid	low	12	2019	0	3	1

Fig.8. Baseline Static Dataset

This datset contains the locations of the 6 ambulances including the Datetime, Weather Conditions, Week, 3 hourTimesplit, Month, Year. This dataset contains 3000 samples.

Results



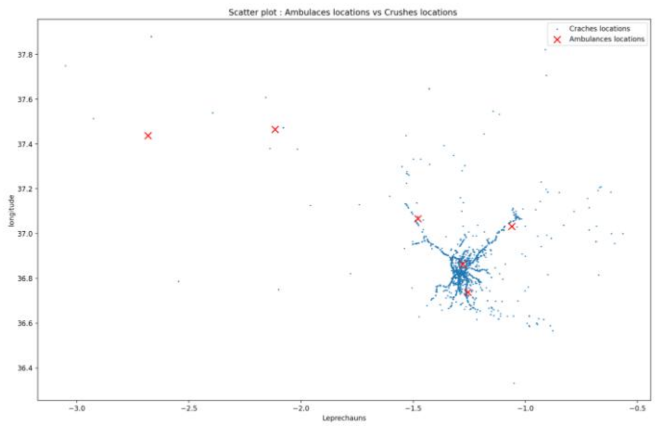
STEP 1: *Launching Application Streamlit*

STEP 2: *A Overview on Abstract*

STEP 3:*Training the Model*

The distance from latitude and longitude is calculated for 6 ambulances. The closest ambulance is added to the overall distance.

We find the 6 virtual ambulance locations by grouping over the crashes locations into 6 groups. We use k-means and ignore the time feature and predict a fixed place. Plot the points in the map.



STEP 4: *Plotting the Ambulances for static positions*

	date	A0_Latitude	A0_Longitude	A1_Latitude	A1_Longitude	A2_Latitude	A2_Longi
0	2019-07-01 00:00:00	0	0	0	0	0	
1	2019-07-01 03:00:00	0	0	0	0	0	
2	2019-07-01 06:00:00	0	0	0	0	0	
3	2019-07-01 09:00:00	0	0	0	0	0	
4	2019-07-01 12:00:00	0	0	0	0	0	

Details of Whether Data

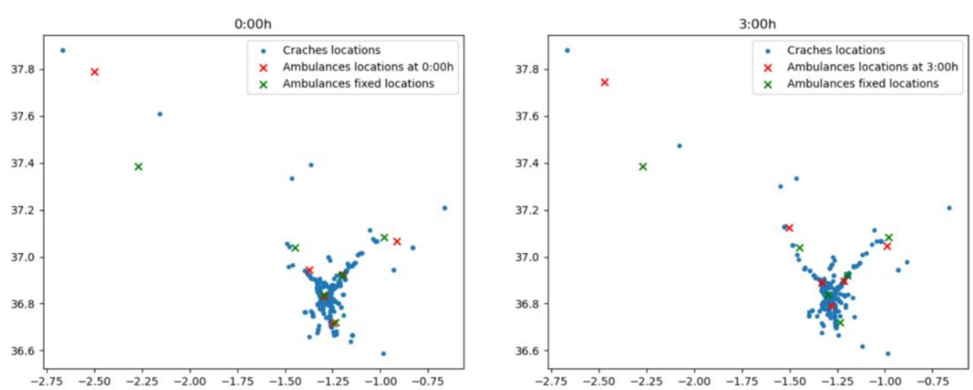
	0
0	Date
1	precipitable_water_entire_atmosphere
2	relative_humidity_2m_above_ground
3	specific_humidity_2m_above_ground
4	temperature_2m_above_ground
5	u_component_of_wind_10m_above_ground
6	v_component_of_wind_10m_above_ground

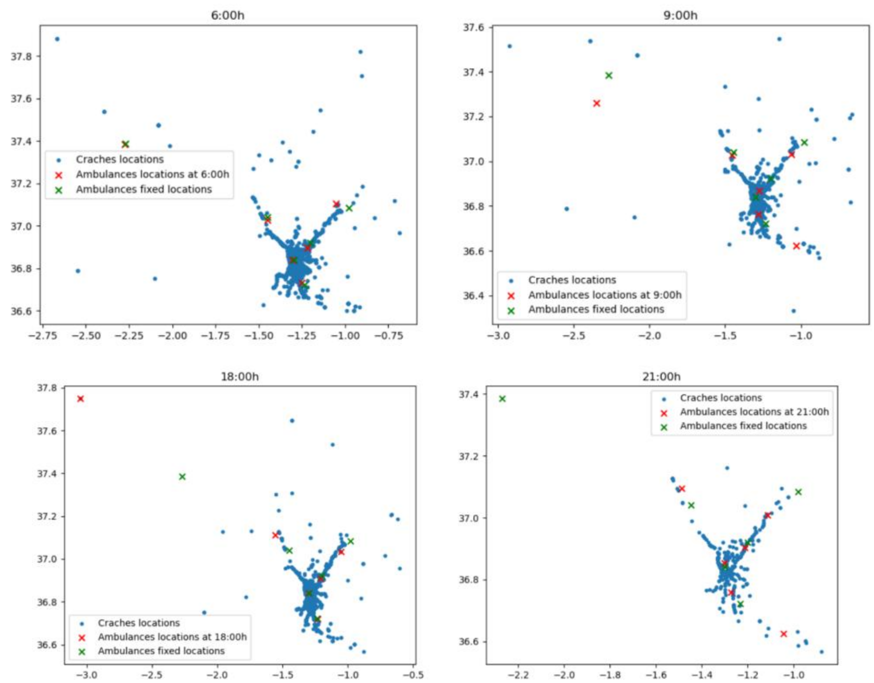
	Date	precipitable_water_entire_atmosphere	relative_humidity_2m_above_ground	spec
0	2018-01-01 00:00:00	24.1	72.7	
1	2018-01-02 00:00:00	27.3623	74.9	
2	2018-01-03 00:00:00	30.2466	86.6	
3	2018-01-04 00:00:00	33.0107	90.9	
4	2018-01-05 00:00:00	27.2504	82.4134	

STEP 5:*Using Nairobi weather as our test set, Creating the Submission File.*

Statistical details of the dataset

	uid	datetime	latitude	longitude
count	6,318	6318	6,318	6,318
mean	3,159.5	2018-09-14 01:40:53.082146560	-1.2703	36.8555
min	1	2018-01-01 00:25:46	-3.05	36.3322
25%	1,580.25	2018-05-03 17:57:09	-1.3166	36.802
50%	3,159.5	2018-08-29 16:51:39	-1.2717	36.8446
75%	4,738.75	2019-02-07 08:01:31.500000	-1.2337	36.8956
max	6,318	2019-06-30 20:06:14	-0.5654	37.8795
std	1,823.9938	None	0.1252	0.1129

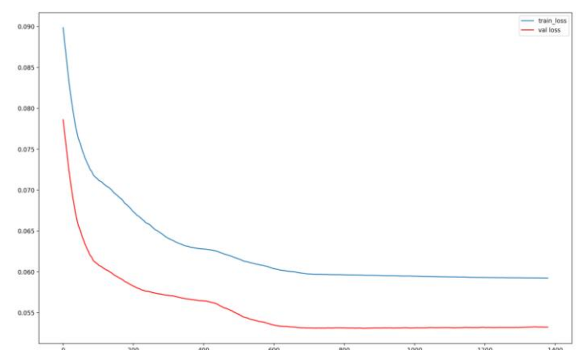




STEP 6:*Locating the optimal positions for 6 ambulances with 3 hours time-split and save the data of the weather conditions and the datetime of it.*

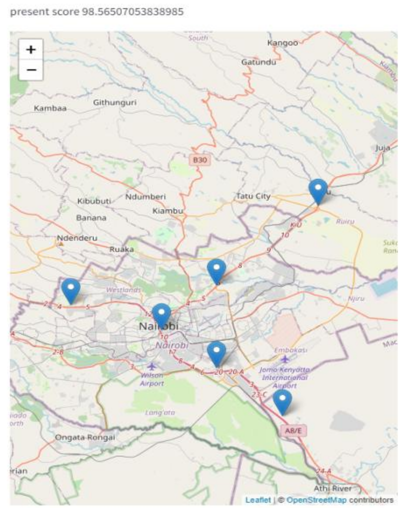
Split the Training dataset into 8 chunks (3 hours each) and create a separate dataset.

Calculate the loss functions.




```
tensor([[-1.2782, 36.9286], [-1.2621, 36.7392], [-1.3341, 36.8862], [-1.2429, 36.8674], [-1.1610,
36.9571], [-1.2893, 36.8200]], requires_grad=True)
```

STEP 7: *Calculates the losses (Train vs Val)*



STEP 8:*Create a Folium Map centered at the mean of coordinates and add markers to the map.*

CONCLUSION

Optimal ambulance positioning plays a key role in saving the victim's life. Here, it offers a promising solution to address the critical challenge of optimizing ambulance positioning for road accidents. Through the integration of techniques, the system demonstrates its ability to predict optimal ambulance locations, thereby reducing response times and ensuring timely medical assistance.This underscores the importance of leveraging geographical factors and accident occurrences to enhance emergency response systems. By preserving patterns through advanced methodologies like Cat2Vec and deep-embedded clustering, the system achieves remarkable accuracy rates, surpassing traditional clustering algorithms.Furthermore, the introduction of a novel scoring function for real-time response time and distance calculation enhances algorithm performance evaluation. The system's effectiveness is evidenced by its ability to achieve a 95% accuracy rate with k-fold cross-validation, showcasing its superiority over conventional approaches.In addition to its technical capabilities, the system's Streamlit web application interface enhances usability, allowing users to interact intuitively with the predictive model and visualize ambulance positioning recommendations. This user-friendly interface, coupled with the system's robust performance, makes it a valuable tool for emergency response organizations and policymakers.Overall, the Ambulance Positioning System represents a significant advancement in optimizing ambulance positioning for road accidents. Its combination of deep learning techniques, novel scoring functions, and user-friendly interface positions it as a practical and effective solution for improving emergency response systems and ultimately saving lives on the road.

Future Work

Predicting optimal ambulance positions in Nairobi using Deep Embedded Clustering (DEC) highlights several avenues for future. Firstly, enhancing the dataset by including additional variables like road type, construction, speed limit, accident severity, driving behavior, and road conditions would offer a more comprehensive understanding of factors influencing accidents, thereby improving ambulance positioning accuracy. Extending the analysis period beyond 2018-2019 could reveal temporal trends, seasonal variations, and the impact of evolving road infrastructure or safety measures, capturing changes in traffic patterns over time.Furthermore, including cities with varying urbanization levels and traffic conditions would provide insights into model performance across diverse settings, offering valuable information on safety measures and policy interventions' effectiveness. Comparative studies using datasets from different socio-economic contexts would elucidate the proposed approach's generalizability and robustness, aiding in understanding accident factors' variations. Time series analysis could be employed to examine accident trends over time, while integrating real-time data feeds such as traffic flow, weather updates, and accident reports would enhance ambulance positioning models' responsiveness and adaptability. By dynamically adjusting ambulance positions based on live data streams, these models could better respond to changing traffic conditions and real-time demands, ultimately improving emergency response system’s effectiveness and saving lives on the road.

References

1. Global Status Report on Road Safety, World Health Organization, Geneva, Switzerland, 2015.
2. T. Sivakumar and R. Krishnaraj, ‘‘Road traffic accidents due to drunken driving in India–challenges in prevention,’’ in Proc. Int. J. Res. Manage. Technol., vol. 2, no. 4, p. 1, 2012.
3. A.F.G.G.Ferreira,D.M.A.Fernandes,A.P.Catarino,andJ. L. Monteiro, ‘‘Localization and positioning systems for emergency responders: A survey,’’ IEEE Commun. Surveys Tuts., vol. 19, no. 4, pp. 2836–2870, 4th Quart., 2017
4. Y. Wen, J. Wang, T. Chen, and W. Zhang, ‘‘Cat2Vec: Learning distributed representation of multi-field categorical data,’’ Tech. Rep., 2016.
5. J. Xie, R. Girshick, and A. Farhadi, ‘‘Unsupervised deep embedding for clustering analysis,’’ in Proc. Int. Conf. Mach. Learn., vol. 48, 2015, pp. 478–487.
6. T. A. Alhaj, M. M. Siraj, A. Zainal, H. T. Elshoush, and F. Elhaj, ‘‘Feature selection using information gain for improved structural-based alert correlation,’’ PLoS ONE, vol. 11, no. 11, Nov. 2016, Art. no. e0166017.
7. J. T. VanEssen, J. L. Hurink, S. Nickel, and M. Reuter, ‘‘Models for ambulance planning on the strategic and the tactical level,’’ Univ. Eindhoven, Beta Res. School Oper. Manage. Logistics, Eindhoven The, Netherlands, Beta Work. Paper WP-434, 2013.
8. M. Zheng, T. Li, R. Zhu, J. Chen, Z. Ma, M. Tang, Z. Cui, and Z. Wang, ‘‘Traffic accident’s severity prediction: A deep-learning approach-based CNNnetwork,’’ IEEE Access, vol. 7, pp. 39897–39910, 2019.
9. X. Xiong, L. Chen, and J. Liang, ‘‘A new framework of vehicle collision prediction by combining SVM and HMM,’’ IEEE Trans. Intell. Transp. Syst., vol. 19, no. 3, pp. 699–710, Mar. 2018.
10. J. Olusina and W. A. Ajanaku, ‘‘Spatial analysis of accident spots using weighted severity index (WSI) and density-based clustering algorithm,’’ J. Appl. Sci. Environ. Manage., vol. 21, no. 2, pp. 397–403, Apr. 2017.
11. A. Alqahtani, X. Xie, J. Deng, and M. W. Jones, ‘‘A deep convolutional auto-encoder with embedded clustering,’’ in Proc. 25th IEEE Int. Conf. Image Process. (ICIP), Oct. 2018, pp. 4058–4062.
12. F. Tian, B. Gao, Q. Cui, E. Chen, and T.-Y. Liu, ‘‘Learning deep representations for graph clustering,’’ in Proc. 28th AAAI Conf. Artif. Intell., Jun. 2014, pp. 1–6.
13. K. G. Dizaji, A. Herandi, C. Deng, W. Cai, and H. Huang, ‘‘Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 5747–5756
14. M. Taamneh, S. Taamneh, and S. Alkheder, ‘‘Clustering-based classification of road traffic accidents using hierarchical clustering and artificial neural networks,’’ Int. J. Injury Control Saf. Promotion, vol. 24, no. 3, pp. 388–395, Jul. 2017.
15. S. Alkheder, M. Taamneh, and S. Taamneh, ‘‘Severity prediction of traffic accident using an artificial neural network,’’ J. Forecasting, vol. 36, no. 1, pp. 100–108, Jan. 2017.
16. S. H.-A. Hashmienejad and S. M. H. Hasheminejad, ‘‘Traffic accident severity prediction using a novel multi-objective genetic algorithm,’’ Int. J. Crashworthiness, vol. 22, no. 4, pp. 425–440, Jul. 2017.
17. B.Ghosh,M.T.Asif,andJ.Dauwels,‘‘Bayesianpredictionoftheduration of non-recurring road incidents,’’ in Proc. IEEE Region 10 Conf. (TEN CON), Nov. 2016, pp. 87–90.
18. S. Sasaki, A. J. Comber, H. Suzuki, and C. Brunsdon, ‘‘Using genetic algorithms to optimise current and future health planning–the example of ambulance locations,’’ Int. J. Health Geographics, vol. 9, no. 1, pp. 1–10, Dec. 2010