Survival Analysis of People with Cardiac Problems in a Simulated Earthquake Environment

W. Velásquez, *Member, IEEE*[*], A. Munoz-Arcentales[†], M. Chalén T., *Member, IEEE*[‡], Joaquín Salvachúa[§]
[*§] *Departamento de Ingeniería de Sistemas Telemáticos, Escuela Técnica Superior de Ingenieros de Telecomunicación Universidad Politécnica de Madrid, Spain, Email:[*]wa.velasquez@alumnos.upm.es, [‡]joaquin.salvachua@upm.es*
[*†‡]*Escuela Superior Politécnica del Litoral, ESPOL, FIEC P.O. Box 09-01-5863, Guayaquil, Ecuador*
Email: ([*]wavelasq, [†]joanmuno, [‡]mchalen)@espol.edu.ec

***Abstract*—This paper states the development of a machine learning model capable of classifying whether a person who has cardiac problems is able to escape from a building if it is affected by an earthquake. The data used to train the model was obtained from a simulation which considered some parameters like earthquake magnitude, BMI, cardiovascular disease, person location within the building, and current evacuation routes. The result of this implementation presents a model able to determine whether a person inside the building could or could not go out of it depending on the parameters previously described, for achieving this goal the use of supervised learning algorithms like KNN and Decision Tree was needed. Also, the case study presents two scenarios, the first scenario presents an evaluation metric at normal walking speed and the second at fast walking speed for a person to escape from an affected building.**

***Keywords*-KNN, Decision Tree, Machine Learning, Earthquake, Simulation**

I. Introduction

Nowadays, the term 'big data' has become important in the professional world, it considers not only huge amounts of data, also, very high rates of process, and its variety; managing certain relationships to generate some predictions or inferences to improve business models. [1]. In engineering areas and similar fields where data analysis has been implemented, it has contributed to have a better agriculture [2], education [3], health [4], etc. Which leads us to talk about an interesting topic, whether the data can help prevent behaviours or improve different areas of the industry, why not use them to save human lives in emergencies?.

Some Latin American countries, especially those that sit around the Ring of Fire, like Ecuador, Perú, Chile, Colombia, present a seismic situation on their coastal regions, because of the inflection point between the Nazca Plate and the South American Plate [5], being a natural hazard that threats communities physical infrastructure and human lives. On April 16, 2016, in Ecuador, an earthquake with 7.8 magnitude on the Richter Scale in the coastal zone happened, where communes from Esmeraldas and Manabí provinces were the most affected with 670 of human losses, and uncountable economic loss in physical infrastructure [6].

This work focuses on a data set generated from an emergency evacuation simulation with several sensors in the offices of a three-story building and different metrics that intervene in the route of a person to go to the nearest exit. In order to predict the number of people that could be hemmed inside of a building in case of an earthquake it considers some metrics based on parameters such as the weight of the person, speed of the person, floor area, and especially a person with a heart disease condition.

The following parts of this paper are divided into: Section II, "A description of the state of the art". Section III, "A presentation of the scheme and configuration of the simulation". Section IV, "Analysis of the different datasets and their behaviors", and Section V, "A conclusion of the results and data analyzed".

II. State of the Art

This study is the continuation of the article entitled: "Relation between BMI and Exit Time of a Building in an Emergency Situation: Earthquake" [7], which describes the behavior of people evacuating of three-story building simulation environment affected by several metrics, such as: velocity of the person, adrenaline, type surface of the floor, Body Mass Index (BMI) and clothing, allowing to know the relation that exists between the BMI and the time that takes a person in leaving a building while an earthquake happens. The important aspects of this development are described below:

A. Earthquake

An earthquake is a sudden and violent shaking of the surface of the Earth caused by unexpected relaxation of energy, accumulated by deformation of the Lithosphere; it propagates itself in the form of seismic waves. It can be presented instantly causing serious damage depending on their intensity, making it one of the most damaging and destructive natural disasters on the planet. It could be the cause of several negative consequences to people around the world like deaths, injuries, people leaving infrastructure; also, it unleashes other natural phenomenon like floods, landslides[8].

B. Body Mass Index

The World Health Organization, WHO, has recommended to use the BMI as a measure for indicating nutritional status in adults, children and adolescents. However, during childhood and adolescence, BMI should be obtained considering the age and sex of the subject, using the equation 1 [9]; during puberty, the correlation between BMI and fat mass decreases notably [10],

$$\mathbf{BMI} = weight/height^2(Kg/m^2) \tag{1}$$

BMI is one of the most important factors in this study, because there is a relationship between the speed and the way on individual moves. As indicated by Zoltan Pataky et all. [11], in the study of the effects of obesity on functional capacity.

C. Cardiovascular Disease

CVD remains the leading cause of death among Europeans and around the world. The Global Burden of Disease study estimated that 29.6% of all deaths worldwide in 2010 (15 616.1 million deaths) were caused by CVD, more than caused by communicable diseases, maternal, neonatal and nutritional disorders combined, and double the number of deaths caused by cancers [12]. But, what does BMI and CVD have in common?, this is described in depth in the article entitled: ”Relation between BMI and Walking Speed in Men with Cardiovascular Disease in an Emergency Situation: Earthquake” [13], which sets patient's walking speed with heart problems to the simulation, so that those who present some pathology take a little more time to leave a building than those who do not.

D. Classification Algorithms

When using machine learning in big data processes, it should be kept in mind what type of algorithm to use to predict or infer data, especially if you want to work with supervised, non-supervised and reinforcement learning algorithms, because specific rules for each category do not exist. In this study, we have considered working with the following algorithms, having as a precedent that we have a set of data (Table III), with an output value (Survive or No to an earthquake):

- *K nearest neighbors* is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition since in the beginning of 1970's as a non-parametric technique [14].
- *Decision tree* builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The result is a tree with decision nodes and leaf nodes [15].

III. Experimental Setup

This study is based on the simulation of a three-story building, which has different floor surfaces types (wood, ceramics, concrete and marble); where users must walk to reach the nearest exit [16], [17], each office has a set of sensors (temperature, carbon dioxide, carbon monoxide, natural gas and vertical liquid); these are used to alert or prevent an accident in the building, such as fires, floods and earthquake [18]. The interesting thing about the simulation is that people are walking on a surface with a different speed according to the floor type, using average speed on a surface ΔVS (wood, ceramic, marble and concrete), average speed taking into account the BMI $\Delta VBMI$ (UW, NO, OM, AO, SO, VSO) and average speed of people with heart problems[19] ΔVHP (60, 60-70, 70 years old). It generates the values shown in table I:

$$\Delta V = (\Delta VS + \Delta VBMI + \Delta VHP)/3 \tag{2}$$

Additionally, each scenario used two types of velocities (normal and fast) to consider that different friction force corresponding to the type of surface and the persons foot.

A. Escape Routes

This section describes the evacuation routes that users should follow in an emergency situation, depending on where they are inside the building. In the Table II, the meaning of the nodes in Figures 1 is explained in detail:

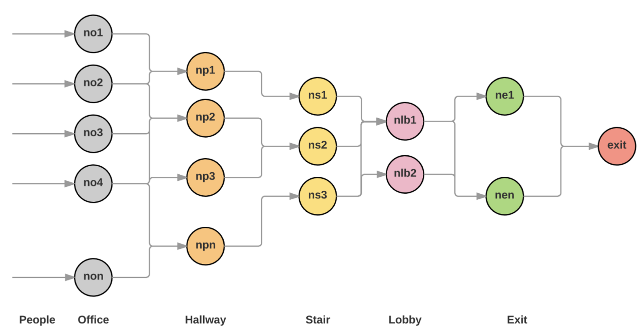

Figure 1. Evacuation routes that people should follow if they are on the first floor.

B. Experimental Data

The simulation assumes three magnitudes of an earthquake (1: 4.01, 2: 5.75, 3: 6.47) and a sample of 2079 men that have some cardiac pathology. Three datasets were obtained (a different magnitude per each one), each one has the fields described in Table III. Aditionally, it was added a column to the dataset (survive), that contains the result of each person result (0: Does not go out, 1: Goes out), depending on exit time, where: $TimeNormal < 40s \rightarrow 1$ else 0, the same situation is when it is using $TimeFast$.

Table I
AVERAGE OF USER SPEEDS FOR EACH TYPE OF FLOOR IN THE BUILDING, CLASSIFIED BY BMI

		Normal (m/s)				*Fast (m/s)*			
Age	**BMI**	**Wood**	**Ceramic**	**Marble**	**Concrete**	**Wood**	**Ceramic**	**Marble**	**Concrete**
<60	*UW*	1,40	1,43	1,42	1,44	1,59	1,62	1,61	1,63
	NO	1,39	1,42	1,41	1,43	1,57	1,59	1,59	1,60
	OM	1,37	1,39	1,39	1,40	1,53	1,56	1,55	1,57
	AO	1,34	1,36	1,36	1,37	1,50	1,52	1,52	1,53
	SO	1,28	1,31	1,30	1,32	1,47	1,49	1,49	1,50
	VSO	1,27	1,30	1,29	1,31	1,43	1,46	1,45	1,47
60 -70	*UW*	1,37	1,40	1,39	1,41	1,57	1,59	1,59	1,60
	NO	1,36	1,39	1,38	1,40	1,54	1,57	1,56	1,58
	OM	1,34	1,37	1,36	1,38	1,51	1,53	1,53	1,54
	AO	1,31	1,34	1,33	1,35	1,47	1,50	1,49	1,51
	SO	1,26	1,28	1,28	1,29	1,44	1,47	1,46	1,48
	VSO	1,25	1,27	1,27	1,28	1,41	1,43	1,43	1,44
>70	*UW*	1,33	1,35	1,35	1,36	1,52	1,55	1,54	1,56
	NO	1,32	1,34	1,34	1,35	1,49	1,52	1,51	1,53
	OM	1,29	1,32	1,31	1,33	1,46	1,49	1,48	1,50
	AO	1,26	1,29	1,28	1,30	1,42	1,45	1,44	1,46
	SO	1,21	1,24	1,23	1,25	1,39	1,42	1,41	1,43
	VSO	1,20	1,23	1,22	1,24	1,36	1,39	1,38	1,40

UW: Under Weight, NO: Normal, OM: Obesity Mild, AO: Average Obesity, SO: Severe Obesity, VSO: Very Severe Obesity – Each speed depending on the person presents a: ±0,02

Table II
DESCRIPTION OF EVACUATION ROUTE NODES

Node	**Type**	**Description**
no	Office	Internal administrative offices that do not have a direct access to the corridor
np	Hallway	Administrative offices that have direct access to the hallway or exits of the building.
ns	Stair	These indicate that users are going through a ladder.
ne	Exit	They represent the exits of the building
nob	Lobby	Represents the lobby of the building

’ Second Floor, ” Ground Floor

Table III
FIELDS CONTAINING THE DATASET

Field	**Description**
BMI	User BMI
TimeNormal	Time it takes a user to leave the building with normal speed
TimeFast	Time it takes a user to leave the building with fast speed
Office	Office where the user is located when the event begins
Floor	Floor of the building where the office is in the event
Path	Optimum path that the user walks out of the building
Length	Distance in meters that the user walks

IV. RESULTS

Depending on the magnitude of the earthquake, serious damage can be caused to a buildings infrastructure, for this reason the present study is based on the premise that the building collapses in 60 seconds. Taking this consideration into account creates a data-set where the magnitude of the earthquake is 4.018, since other datasets with higher earthquake magnitude generate exit times less than 40 seconds. If higher exit times are taken into account, the described model adapts doing a classification with 94% confidence without tuning the algorithm.

In the figure 2, it is important to mention that the number of persons that get out of the building on the established time is 183 out of 664 that were on the ground floor, other floorswere consider. BMI presents a correlation of 0.11 with TimeNormal and 0.10 with TimeFast.

A. Normal Speed Time

In table IV and figure 3, results of the algorithm assuming a normal speed are presented. It can be appreciated the classification of the person who goes out the building or not in the established time is 92% of precision using the knn and 100% using decision tree. Aditionally, $K = 3$ was used to improve the knn algorithm obtaining a better classification by 3%.

Table IV
COMPARISON OF ALGORITHMS WITH NORMAL TIME SPEED

Field	**Knn**	**Decision tree**
Precision	0,92	1
Tuning Algorithm	0,95	1
False Negative	11	0
False Positive	2	0
Avg. Classification 0	0,98	1
Avg. Classification 1	0,84	1
Avg. Precision Classification	0,93	1

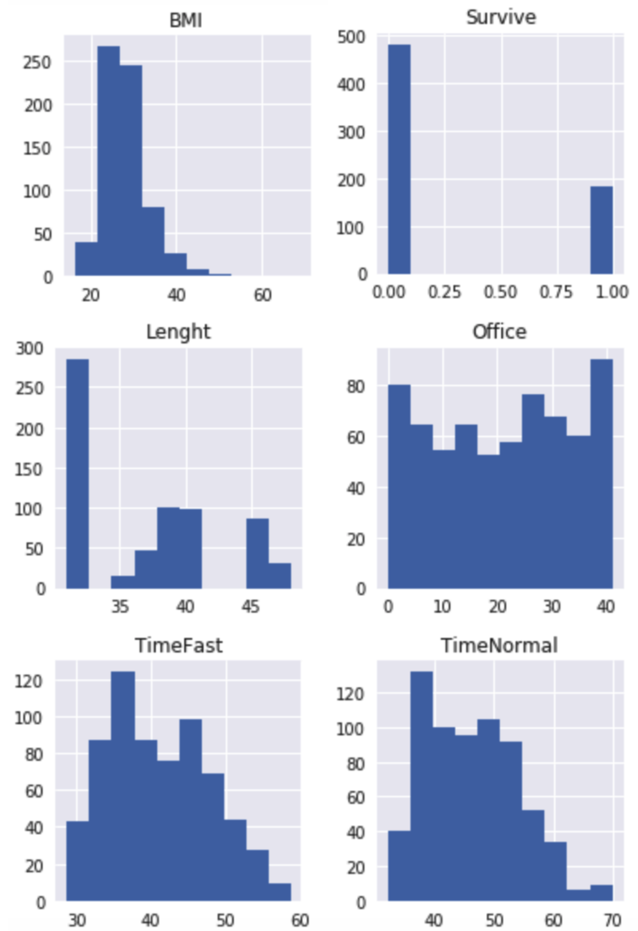

Figure 2. Result from a dataset of 4.01 Magnitude Earthquake - BMI (Person/BMI Category), Survive (Person/0: Don't Survive, 1: Survive), Length (Meters/Person), Office (Office/Person), TimeFast and TimeNormal (Seconds/Person)

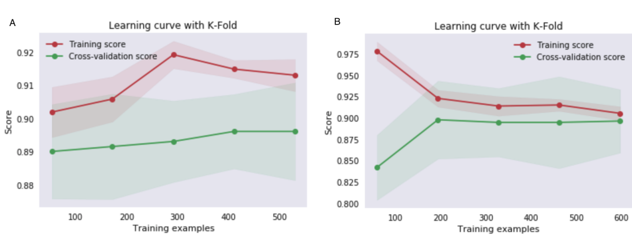

Figure 3. Graphic of k-fold (A: Original, B: Tunning) Knn Algorithm

B. Fast Speed Time

If fast speed time is considered, there are variations in both algorithms, because of the existence of more data with the value of $1 \rightarrow Survive$, considering that persons have more possibilities to get out the building.

It can be appreciated in table V that the precision value is 87% in both algorithms. But, when the tuning is done, there is a slight change, 88%→ KNN (K=5) y 89%→ Decision Tree. But, Decision Tree presents less False Positives and False Negatives than KNN.

Table V
COMPARISON OF ALGORITHMS WITH FAST TIME SPEED

Field	KNN	Decision tree
Precision	0,87	0,87
Tuning Algorithm	0,88	0,89
False Negative	17	15
False Positive	3	1
Avg. Classification 0	0,97	0,99
Avg. Classification 1	0,77	0,74
Avg. Precision Classification	0,90	0,91

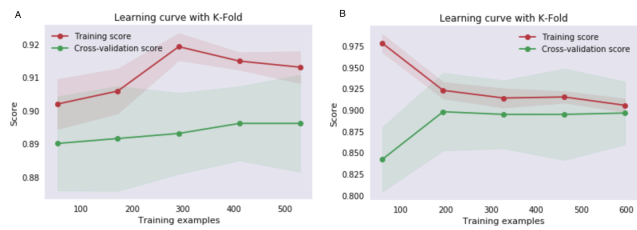

Figure 4. Graphic of k-fold (A: Original, B: Tunning) KNN Algorithm

C. Survival Analysis

It was taken into account the times (Normal Speed, Fast Speed) collected during the simulation for each magnitude of the earthquake and as a premise a time $t = 40s$ before a catastrophe occurs (building collapse, people trapped by debris and even death). The Table VI presents an analysis of the people who could evacuate the building at the suggested time, with this information it can be concluded that people who are on lower floors are more likely to leave a building than those on upper floors. It is advisable that people who can not evacuate in these times follow the rules of survival in an earthquake.

Table VI
PERCENTAGE OF EVACUATION OF THE BUILDING

		Earthquakes		
Floor	**Q**	**M1**	**M2**	**M3**
Ground	933	86%	3%	1%
First	805	20%	0%	0%
Second	341	5%	0%	0%

Q: Quantity, M1: 4.01, M2: 5.75, M3: 6.47

V. Discussion and Future Works

The Decision Tree algorithm presents a better classification of the persons that could get out of a building in a time of 40 seconds or less than the KNN; being assumed two specific parameters: the time of the building collapse and the time a person on ground floor takes to go to the nearest exit.

As previous sections explain, the persons who are in higher floors should not go to the exits (building structure), they must follow precaution rules from the assistance organism when an earthquake occurs (magnitude does not matter).

Depending on the BMI and the speeds of the individuals both algorithms and exit times present a classification of persons that do not get out of the building is higher than the persons who do. That's why it is recommended to do physical activities to get more opportunities to get out of a building in this kind of situation.

More studies where BMI and certain user speeds on specific floor surfaces are analyzed in other emergency cases like fire, should be done.

ACKNOWLEDGMENT

Thanks to the Escuela Superior Politécnica del Litoral (ESPOL) and to the Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) for financing the continuity of evacuation systems study in emergencies.

REFERENCES

[1] F. Provost and T. Fawcett, "Data science and its relationship to big data and data-driven decision making," *Big Data*, vol. 1, no. 1, pp. 51–59, 2013.

[2] M. Carolan, "Publicising food: Big data, precision agriculture, and co-experimental techniques of addition," *Sociologia Ruralis*, 2016.

[3] J. Saltz and R. Heckman, "Big data science education: A case study of a project-focused introductory course," *Themes in Science and Technology Education*, vol. 8, no. 2, pp. 85–94, 2016.

[4] G. Roesems-Kerremans, "Big data in healthcare," *Journal of Healthcare Communications*, 2016.

[5] J. Audefroy, "La problemática de los desastres en el hábitat urbano en américa latina," *Revista INVI*, vol. 18, no. 47, 2003.

[6] " A un ano de la Tragedia que sacudio a Ecuador, qu eha pasado desde el terremoto, Abril 2016," http://cnnespanol.cnn.com/2017/04/13/, [Online].

[7] W. Velásquez, A. Munoz-Arcentales, and J. Salvachúa, "Relation between bmi and exit time of a building in an emergency situation: Earthquake," *Engineering Letters, vol. 25. no. 4*, 2017.

[8] F. V. Sánchez, "Los terremotos y sus causas," in *El estudio de los terremotos en Almería*. Instituto de Estudios Almerienses, 1994, pp. 17–38.

[9] R. Burrows, N. Díaz, and S. Muzzo, "Variaciones del índice de masa corporal (imc) de acuerdo al grado de desarrollo puberal alcanzado," *Revista médica de Chile*, vol. 132, no. 11, pp. 1363–1368, 2004.

[10] R. Chamorro and M. G. Lorenzo, "Indice de masa corporal y composición corporal: Un estudio antropométrico de 2500 deportistas de alto nivel," *Lecturas: Educación física y deportes*, vol. 76, 2004.

[11] Z. Pataky, S. Armand, S. Müller-Pinget, A. Golay, and L. Allet, "Effects of obesity on functional capacity," *Obesity*, vol. 22, no. 1, pp. 56–62, 2014.

[12] M. Nichols, N. Townsend, P. Scarborough, and M. Rayner, "Cardiovascular disease in europe 2014: epidemiological update," *European heart journal*, p. ehu299, 2014.

[13] W. Velásquez, A. Munoz-Arcentales, W. Yanez, and J. Salvachúa, "Relation between bmi and walking speed in men with cardiovascular disease in an emergency situation: Earthquake," in *The International Conference on Advanced Management and Information Technology Services*, 2017.

[14] "K Nearest Neighbors - Classification," http://www.saedsayad.com/k_nearest_neighbors.htm, [Online].

[15] "Decision Tree - Classification," http://www.saedsayad.com/decision_tree.htm, [Online].

[16] W. V. Vargas, A. Munoz-Arcentales, and J. S. Rodríguez, "A distributed system model for managing data ingestion in a wireless sensor network," in *Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual*. IEEE, 2017, pp. 1–5.

[17] W. A. V. Vargas, "Algoritmo de recomendación sensible a contexto de elementos educativos reutilizables con almacenamiento orientado a grafos," *Revista Tecnológica-ESPOL*, vol. 30, no. 1, 2017.

[18] A. Munoz-Arcentales, W. Yánez-Pazmino, and W. V. Vargas, "Proposal of a communication structure model for activating reactive signaling in an emergency evacuation systems," in *Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual*. IEEE, 2017, pp. 1–5.

[19] G. Chiaranda, E. Bernardi, L. Codecà, F. Conconi, J. Myers, F. Terranova, S. Volpato, G. Mazzoni, and G. Grazzi, "Treadmill walking speed and survival prediction in men with cardiovascular disease: a 10-year follow-up study," *BMJ open*, vol. 3, no. 10, p. e003446, 2013.