

**Evaluación de vulnerabilidad sísmica actual de albergues temporales en
Bucaramanga aplicando algoritmos de clasificación supervisada.**

Angie Yulima Cifuentes Urrego

Jessica Liliana Fuentes Jerez

Trabajo de Grado para Optar al Título de Ingeniero Industrial

Director:

Daniel Orlando Martínez Quezada

M.Sc en Ingeniería Industrial

Codirector:

Henry Lamos Díaz

Ph.D en Física, Matemáticas

Universidad Industrial de Santander

Facultad de Ingeniería Físico-mecánicas

Escuela de Estudios Industriales y Empresariales

Bucaramanga

2018

AGRADECIMIENTOS

Gracias a Dios por permitirnos tener la dicha de vivir y disfrutar de este gran logro, por cruzarnos en el camino al profesor Daniel Orlando Martínez Quezada, Director de Grado y guía durante todo este tiempo, agradecemos enormemente por compartirnos sus conocimientos, orientarnos y sobre todo por su paciencia y motivación cuando las cosas se complicaban. ¡Muchas gracias Dani!

Gracias al profesor Henry Lamos porque con su experiencia y sabiduría fue un apoyo especial en el desarrollo de este trabajo de investigación. De igual manera agradecemos al grupo de investigación OPALO, por acogernos en su familia de investigadores y encaminarnos en nuevos retos que finalmente dieron frutos.

No menos importante, agradecemos a la prestigiosa Universidad Industrial de Santander por abrirnos sus puertas cuando éramos unas niñas y que hoy, seis años después nos entrega al mundo laboral con valiosas herramientas para enfrentar los retos que nos ponga la vida.

Gracias a todos los profesores que nos compartieron sus conocimientos a lo largo de estos años.

Finalmente agradecemos a cada una de las personas que de una u otra forma contribuyeron tanto en nuestro camino por la Universidad, como en el desarrollo de nuestro Proyecto de Grado. ¡A todos ellos, muchas gracias!

DEDICATORIA

Este trabajo lo dedico a Dios, a mi Mamá, Ortencia Urrego y a mi equipo de trabajo. A Dios porque me ha permitido llegar a escribir estas palabras. ¡Sin la voluntad de Dios no se mueve ni la hoja de un árbol! A mi Mamá, quién ha sido la persona que me ha acompañado, apoyado y ha velado por mi bienestar y educación todos estos años. Siempre confiando en mí y enseñándome a ser una mujer independiente, que lucha por conseguir lo que quiere. A mi gran equipo de trabajo, Daniel Martínez director de Proyecto de grado y Jessica Fuentes mi compañera de Proyecto. Unas personas maravillosas con quien desarrolle un trabajo arduo, un trabajo que siempre fue de tres, y que hoy cierra su ciclo llenando todas nuestras expectativas. Por último, vuelvo nuevamente a mi Mamá, a ella le debo la persona que soy hoy en día. ¡Soy Ingeniera!

Angie Cifuentes Urrego

Primeramente, dedico este trabajo a Dios por permitirme ser una mujer llena de principios y valores, por ser la luz en mi camino y no permitir que perdiera mi rumbo. A mi madre, Luz Jerez Sepúlveda por brindarme su cariño, amor y apoyo durante todos estos años, a ella le debo todo lo que soy hoy. A ella dedico este trabajo, el esfuerzo en toda mi carrera profesional y mi crecimiento personal. Y a mi equipo de trabajo Daniel Martínez y Angie Cifuentes, por su dedicación, compromiso y entrega total. Finalmente agradezco a todas las personas que de una u otra manera hicieron parte de mi camino y dejaron grandes enseñanzas en mi vida.

Jessica Fuentes Jerez

Contenido

	Pág.
Introducción.....	15
1. Generalidades de la Investigación	21
1.1 Planteamiento del problema	21
1.2 Objetivos	23
1.2.1 Objetivo General.....	23
1.2.2 Objetivos específicos:.....	23
2. Revisión de la literatura	24
2.1 Gestión de desastres.....	26
2.2 Aplicaciones del aprendizaje automático	28
2.2.1 Aplicaciones de máquinas de soporte vectorial.....	30
2.2.2 Aplicaciones de Arboles de decisión:.....	33
3. Marco de antecedentes	36
4. Marco Teórico.....	39
4.1 Aprendizaje automático.	39
4.1.1 Aprendizaje supervisado.....	41
4.1.2 Aprendizaje no supervisado:.....	41
4.2 Sistema de expertos	42
4.3 Validación cruzada	43
4.3.1 K-folds:	44
4.3.2 Hold-out:.....	45
4.4 Máquinas de soporte vectorial.	46
4.4.1 Función Kernel	47

4.5	Árbol de decisión.....	49
4.5.1	Cart:	50
4.5.1.1	Criterios de división.....	51
4.5.1.1.1	Entropía:	51
4.5.1.1.2	Índice de Gini	51
4.5.1.1.3	Error de mala clasificación:	51
4.5.2	Chaid.....	52
4.5.3	Quest	52
4.5.4	Crecimiento del árbol:	52
5.	Caracterización de albergues temporales.....	53
5.1	Año de construcción	54
5.2	Tipología de Estructuras	55
5.3	Niveles	60
5.4	Distancia a fallas.....	61
5.5	Zonas en riesgo de remoción de masa	62
5.6	Llenos antrópicos y mecánicos	63
5.7	Vulnerabilidad	64
5.8	Albergue.....	67
5.8.1	Albergue temporal:	68
5.8.2	Albergues comunitarios	69
5.8.3	Metodología de selección de albergues temporales.....	69
6.	Caso de Estudio.....	76
6.1	Análisis de Datos	77
6.2	Ejecución de los algoritmos.....	81
6.3	Definición de los parámetros	84

6.3.1 Máquinas de soporte vectorial	84
6.3.2 Árbol de decisión.....	88
7. Conclusiones	93
8. Recomendaciones	95
Referencias Bibliográficas.....	96

Lista de Tablas

	Pág
Tabla 1. Cumplimiento de objetivos de la investigación	20
Tabla 2. Desarrollo histórico de la investigación artificial según Jackson (1989).....	30
Tabla 3. Parámetros de las máquinas de soporte vectorial.....	48
Tabla 4. Parámetros del árbol de decisión.....	53
Tabla 5. Caracterización de las fallas de Bucaramanga.	61
Tabla 6. Niveles de vulnerabilidad.....	67
Tabla 7. Metodología de selección de albergues temporales por autor.....	70
Tabla 8. Factores asociados a la vulnerabilidad estructural	77
Tabla 9. Proceso de ejecución de los modelos de aprendizaje automático.	83
Tabla 10. Rango de parámetros para las máquinas de soporte vectorial.....	85
Tabla 11. Porcentaje de Mala clasificación para los modelos y tiempo de ejecución	87
Tabla 12. Valores asignados al rango de parámetros para el árbol de decisión.	88

Lista de figuras

	Pág
Figura 1. Clasificación del Aprendizaje Automático.....	40
Figura 2. Algoritmo de evaluación VC.....	44
Figura 3. Máquinas de Soporte Vectorial.....	47
Figura 4. Truco del Kernel para el caso no linealmente separable.....	48
Figura 5. Árbol de decisión.....	50
Figura 6. Esquema estructural de muros de adobe o tapia pisada.....	56
Figura 7. Esquema estructural de la mampostería en ladrillo confinado.....	57
Figura 8. Esquema estructural de la mampostería no confinada.....	58
Figura 9. Esquema estructural de la mampostería reforzada.....	58
Figura 10. Esquema estructural en pórtico.....	59
Figura 11. Esquema estructural de muros en concreto.....	60
Figura 12. Tipología de albergues temporales.....	69
Figura 13. Representación de los albergues temporales.....	72
Figura 14. Localización de Instituciones Educativas.....	73
Figura 15. Localización de Zonas verdes.....	74
Figura 16. Localización de Zonas deportivas.....	75
Figura 17. Análisis de los niveles de vulnerabilidad en cada Tipología.....	79
Figura 18. Análisis de datos por año de construcción.....	80
Figura 19. Análisis de datos por clases de vulnerabilidad.....	81
Figura 20. Comportamiento del Error contra la variación del Cost.....	87
Figura 21. Comportamiento del Error contra la variación del Cp.....	89
Figura 22. Comportamiento del Error contra la variación del Maxdepth.....	90
Figura 23. Comportamiento del Error contra la variación del Minisplit.....	90

Resumen

TÍTULO: EVALUACIÓN DE VULNERABILIDAD SÍSMICA ACTUAL DE ALBERGUES TEMPORALES EN BUCARAMANGA APLICANDO ALGORITMOS DE CLASIFICACIÓN SUPERVISADA. *

AUTOR: CIFUENTES URREGO, Angie Yulima.

FUENTES JEREZ, Jessica Liliana. **

PALABRAS CLAVE: Sismo, Albergues temporales, Aprendizaje automático, Árbol de decisión, Máquinas de soporte vectorial, Vulnerabilidad.

DESCRIPCIÓN: En la siguiente investigación se planteó, el diseño de modelos de clasificación supervisada para evaluar la vulnerabilidad sísmica de albergues temporales en la ciudad de Bucaramanga, con base al criterio dado por un experto, con el fin de validar los refugios para seguridad en la reubicación de las comunidades afectadas ante un desastre natural, así como, anticipar medidas de reducción de impactos, como acciones correctivas sobre las infraestructuras esenciales y reforzar las zonas que se cataloguen altamente vulnerables identificadas a través de los mapas de riesgo. Para tal objetivo, se emplearon dos algoritmos de aprendizaje supervisado: Máquinas de soporte vectorial y Árbol de decisión. Estos modelos permitieron predecir el nivel de vulnerabilidad para 14 albergues de prueba. Se concluyó que, el modelo que mejor clasificó los datos con un rendimiento del 64,29%, fue el de máquinas de soporte vectorial con función de kernel lineal, medido a partir del porcentaje de buena clasificación de un conjunto de prueba. Esta investigación permite a las entidades gubernamentales y/o interesados obtener una aproximación a la vulnerabilidad sísmica de albergues en Bucaramanga y asimismo se tomen las decisiones necesarias para reforzar aquellas estructuras que se asocian con un nivel de vulnerabilidad alto. De esta forma se minimizarán los impactos en la vida y se brindará mayor seguridad a la población.

** Facultad de Ingeniería Físico-mecánicas. Escuela de Estudios Industriales y Empresariales. Programa de Ingeniería Industrial. Director: M.Sc. Daniel Orlando Martínez Quezada. Codirector: Ph. D. Henry Lamos Díaz.

Abstract

TITLE: EVALUATION OF CURRENT SEISMIC VULNERABILITY OF TEMPORARY SHELTERS IN BUCARAMANGA APPLYING MONITORED CLASSIFICATION ALGORITHMS. *

AUTHOR: CIFUENTES URREGO, Angie Yulima.

FUENTES JEREZ, Jessica Liliana **

KEY WORDS: Earthquake, Temporary shelters, Machine learning, Decision tree, Vector support machines, Vulnerability

DESCRIPTION: In the following investigation, the design of supervised classification models to evaluate the seismic vulnerability of temporary shelters in the city of Bucaramanga was established, based on the criteria given by an expert, in order to validate the shelters for security in the relocation of the communities affected by a natural disaster, as well as anticipate measures to reduce impacts, such as corrective actions on essential infrastructures and reinforce the areas that are classified as highly vulnerable, identified through risk maps. For this purpose, two supervised learning algorithms were used: Vector support machines and Decision tree. These models allowed to predict the level of vulnerability for 14 test shelters, in a matter of seconds. The conclusion was that the model that best classified the data, with a performance of 64, 29%, where the machines of vertical support with linear kernel function, measured from the classification of a test set. This investigation allows government entities and/or interested parties to obtain an approximation to seismic vulnerability of shelters in Bucaramanga and therefore make the necessary decisions to reinforce structures with high levels of vulnerability. This way the impact on life will be minimized and greater safety will be achieved for the population.

* Graduation Project

** Facultad de Ingeniería Físico-mecánicas. Escuela de Estudios Industriales y Empresariales. Programa de Ingeniería Industrial. Director: M.Sc. Daniel Orlando Martínez Quezada. Codirector: Ph. D. Henry Lamos Díaz.

Introducción

Colombia es un país con alta actividad sísmica debido a su ubicación en el cinturón de fuego del pacífico. Actualmente el 86% de la población colombiana se localiza en zonas de amenaza sísmica alta y media; considerándose las regiones Pacífica y Andina las poblaciones más expuestas. A nivel municipal, Cali representa la mayor población expuesta a amenaza sísmica alta, seguido por otras capitales como Cúcuta, Bucaramanga, Pereira, Villavicencio, Pasto y Manizales (Mundial, 2012).

Fenómenos como los sismos fuertes y las erupciones volcánicas severas, por lo general ocasionan grandes pérdidas concentradas en un territorio y en un lapso de tiempo relativamente corto, por esto han sido denominados riesgos intensivos (EIRD/ONU, 2011). Estos eventos pueden ser poco frecuentes, pero con grandes impactos, como el terremoto de Popayán (31 de marzo de 1983), el desastre provocado por la erupción del volcán Nevado del Ruiz y la avalancha que destruyó a Armero (13 de noviembre de 1985), los sismos del Atrato Medio (17 y 18 de octubre de 1992), el terremoto de Páez (6 de junio de 1994) y el sismo del Eje Cafetero (25 de enero de 1999), que han generado importantes pérdidas de vidas y un sin número de viviendas destruidas (Mundial, 2012, pág. 21).

En vista de los acontecimientos sucedidos en los últimos años las entidades gubernamentales gestoras de desastres han tomado medidas para la gestión de riesgos con el objetivo de prevenir y mitigar impactos sociales, económicos y ambientales en el país. En

Colombia se cuentan con avances estructurales en viviendas y construcciones esenciales, que cuentan con normas de sismo resistencia y de ordenamiento territorial, las cuales han contribuido parcialmente, a la reducción del riesgo en edificaciones nuevas. Sin embargo, enfrenta la seria situación de que cerca del 50% de las viviendas fueron construidas antes de que existieran normas de sismo resistencia y es producto de la fuerte actividad de consolidación de asentamientos informales que han generado un incremento del riesgo, pues no se cumplen los estándares constructivos y muchos de ellos se localizan en zonas expuestas a algún tipo de amenaza (Mundial, 2012, pág. 219).

Bucaramanga es considerada una zona propensa a amenazas sísmicas debido a su ubicación geográfica como lo menciona el Ministerio de Ambiente, Vivienda y Desarrollo Territorial en la comisión asesora permanente para el régimen de construcciones sismo resistentes (Creada por la Ley 400 de 1997); según García (2015), esta ciudad se encuentra a 50 kilómetros del nido sísmico considerado el segundo más activo en el mundo, con epicentro localizado a 150 km de profundidad en el que se registran alrededor de 85 movimientos telúricos diarios de baja magnitud en la escala de Richter. Es por esto que la gestión integral de riesgos debe ser un pilar fundamental en la unidad municipal de desastres de la ciudad de Bucaramanga con énfasis en las tres fases de desastres: antes, durante y después del desastre. Las actividades que abarcan cada una de estas etapas en la unidad gestora de desastres de la alcaldía de Bucaramanga se explican a continuación:

- Antes: la fase pre-desastres actualmente es el eslabón más débil de la gestión integral de desastres en Bucaramanga puesto que las actividades empleadas en la mitigación del desastre se centran exclusivamente en simulacros de evacuación y planes de contingencia de desastres.

Esta fase debe considerarse la más importante, sin embargo, el desarrollo de actividades de mitigación como la identificación y reforzamiento de las infraestructuras vulnerables no se lleva a cabo; este es un tema crucial considerando que la mayoría de edificaciones de la ciudad son construcciones antiguas por lo que es de esperarse que estas no cuentan con la normatividad vigente de construcciones sismo-resistentes.

- Durante: la fase de respuesta abarca el lapso de ocurrencia del desastre natural en el que se realizan actividades de evacuación, búsqueda y rescate.
- Después: la fase post-desastres incluye la distribución de recursos, reubicación de la población afectada, adecuación de las infraestructuras esenciales deterioradas y otras actividades logísticas.

Es muy importante la fase pre-desastres para corroborar que los albergues temporales de la ciudad sean seguros y reubicar las comunidades afectadas y/o en caso contrario realizar los reforzamientos necesarios para mitigar los impactos en la vida humana. En general, en “los municipios no se utilizan los mecanismos de protección y aseguramiento de edificaciones y patrimonio particular, lo que aumenta la presión y el riesgo público” (Mundial, 2012, pág. 177). De acuerdo con lo anterior se resalta la importancia de la mitigación del riesgo ante un acontecimiento sísmico a través una evaluación de vulnerabilidad actual de albergues temporales en Bucaramanga y así validar los refugios seguros para la reubicación de las comunidades afectadas ante un desastre natural. Además, es una herramienta de apoyo para la toma de decisiones ante las entidades gubernamentales y unidades interesadas. También permite anticipar medidas de reducción de impactos, como acciones correctivas sobre las

infraestructuras esenciales y asimismo reforzar las zonas que se cataloguen altamente vulnerables identificadas a través de los mapas de riesgo.

No obstante, en términos de percepción, es importante aclarar que la sismo resistencia no implica que la edificación sea “antisísmica”, pues no existen tales estructuras. Lo que realmente se pretende es una reducción considerable de los efectos de los fenómenos en la vida humana. (Mundial, 2012, pág. 291)

En consultas previamente realizadas en artículos científicos y tesis a nivel nacional e internacional en evaluaciones de vulnerabilidad se utilizan diversos métodos y herramientas, entre los que se encontraron: Métodos Delphi, proceso jerárquico analítico (AHP), sistemas de información geográfica (SIG), índice de vulnerabilidad (IV), lógica difusa, curvas de fragilidad, simulación, redes neuronales artificiales y árboles de eventos.

Desde la perspectiva de la ingeniería industrial se plantea la solución a problemas de la vida real aplicando modelos estadísticos que logren representar una situación de estudio; además, la interpretación y análisis de datos como apoyo en la toma de decisiones. Por consiguiente, para esta investigación se propone el desarrollo de un modelo de aprendizaje automático que represente el conocimiento de un experto en evaluación de vulnerabilidad estructural ante la ocurrencia de un sismo de albergues temporales en la ciudad de Bucaramanga. Para este estudio se evaluaron dos algoritmos de clasificación supervisada: máquinas de soporte vectorial (SVM) y árboles de decisión (DT) con los cuales se pretende replicar el conocimiento de un experto. Este enfoque permite reducir capital económico dado

que el desarrollo de evaluaciones basadas en expertos es muy costoso, sin dejar a un lado las limitaciones de movilidad a diferentes lugares donde se hace necesario los estudios de vulnerabilidad; facilita la toma de decisiones basadas en la etapa pre-desastres como lo es la mitigación del desastre y minimización del tiempo de respuesta.

Este documento está organizado de la siguiente manera. En la sección 1 se presenta las generalidades de la investigación. En la sección 2 y 3 se presenta la revisión de la literatura y el marco de antecedentes. En la sección 4 se explica el marco teórico. En la sección 5 se describe la caracterización de los albergues temporales. En la sección 6 se presenta el caso de estudio para la ciudad de Bucaramanga junto con los resultados obtenidos. En la Sección 7 se presentan las conclusiones y recomendaciones de la investigación realizada. Finalmente, en la Sección 8 se encuentran las referencias bibliográficas consultadas.

Tabla de cumplimiento de objetivos

Tabla 1.

Cumplimiento de objetivos de la investigación

Objetivos específicos	Apartado Relacionado
1. Realizar una revisión de literatura sobre las aplicaciones de los modelos máquinas de soporte vectorial y árboles de decisión.	2.0
2. Caracterizar los albergues temporales propuestos en la ciudad de Bucaramanga.	5.0
3. Desarrollar modelos de clasificación supervisada utilizando máquinas de soporte vectorial y árboles de decisión para la evaluación de la vulnerabilidad estructural de albergues temporales a través del criterio de un experto.	6.0
4. Determinar el algoritmo de clasificación supervisada que mejor describa la evaluación de vulnerabilidad sísmica actual de albergues en Bucaramanga.	6.0-7.0
5. Realizar un artículo académico de carácter publicable basado en los resultados de la investigación.	Apéndice A

1. Generalidades de la Investigación

1.1 Planteamiento del problema

La ley N° 46 de 1988 es la ley con la cual se crea y organiza el Sistema Nacional para la Prevención y Atención de Desastres; define un desastre como “El daño grave o la alteración grave de las condiciones normales de vida en un área geográfica determinada, causada por fenómenos naturales y por efectos catastróficos de la acción del hombre en forma accidental, que requiera por ello de la especial atención de los organismos del Estado y de otras Entidades de carácter humanitario o de servicio social”.

Colombia está conformada por 3 placas tectónicas: la del Caribe, la Nazca y la Suramericana, éstas placas están en constante movimiento y al colisionar entre sí generan perturbaciones en la tierra, catalogándolo como un país altamente sísmico por su ubicación en el Cinturón de Fuego del Pacífico.

Por otro lado, Bucaramanga es considerada una ciudad vulnerable a sucesos naturales no controlables por el hombre; se encuentra a 50 kilómetros del nido sísmico de Bucaramanga en el que se registran movimientos telúricos diarios de baja magnitud en la escala de Richter y alta profundidad.

Según el centro de conocimiento en salud pública y desastres el colapso de edificios es el factor de riesgo que más muertes puede causar. Existen muchas formas de gestionar la

prevención de riesgos en las comunidades, y la evaluación al estado de las infraestructuras es una de ellas; el hecho de saber que tan vulnerable es un edificio permite a las entidades gubernamentales y demás interesadas en la gestión de desastres tomar acciones pertinentes que ayuden a reducir los impactos ante la ocurrencia de un sismo, para este caso el objetivo principal es conocer una medida de vulnerabilidad de albergues temporales ubicados en Bucaramanga que permita analizar y decidir si es necesario el reforzamiento de dicha estructura para disminuir las probabilidades de un posible colapso.

Es evidente que los nuevos logros que se alcanzan en el campo del diseño de estructuras pueden aplicarse únicamente a las nuevas construcciones, cuyo número es insignificante respecto al número total de estructuras existentes. Por otra parte, la ocurrencia de los fenómenos sísmicos es una tarea difícil de predecir, por esto se hace necesario mejorar la resistencia de las estructuras existentes, para así mitigar las pérdidas que los terremotos están produciendo en el mundo (Yépez, Barbat y Torres, 1995, p. 1).

La evaluación de vulnerabilidad sísmica de estructuras es una actividad que requiere de grandes recursos económicos, considerando aspectos asociados a herramientas, personal involucrado, expertos y tiempos de respuesta. El uso de un sistema de expertos concierne una serie de beneficios en torno a la solución de problemas en las diferentes disciplinas (Castillo, Gutiérrez, & Hadi, 2011, p. 7):

- Personal con poca experiencia puede resolver problemas que requieren un conocimiento de experto. Esto es también importante en casos en los que hay pocos expertos humanos.

- Los sistemas expertos pueden responder a preguntas y resolver problemas más rápido que un experto humano. Por ello, los sistemas son muy valiosos en casos en los que el tiempo de respuesta es crítico.

- Debido a la capacidad de los ordenadores de procesar un elevado número de operaciones complejas de forma rápida y aproximada los sistemas de expertos suministran respuestas rápidas y fiables en situaciones en las que los expertos humanos no pueden.

La integración de un experto a cualquier proceso representa un costo muy elevado. Por esto el presente proyecto plantea el desarrollo de un modelo de aprendizaje automático con el fin de replicar el conocimiento de un experto a partir del desarrollo de un modelo de clasificación múltiple supervisado teniendo como insumo el criterio de expertos y factores asociados a la vulnerabilidad de la estructura, siendo un instrumento de valor que pueda ser utilizado en gran variedad de infraestructuras. Además, sirve de apoyo para la toma de decisiones preventivas -Físicas, Sociales y Económicas- de las entidades gestoras de desastres.

1.2 Objetivos

1.2.1 Objetivo General: Diseñar modelos de clasificación supervisada para evaluar la vulnerabilidad sísmica de albergues temporales en la ciudad de Bucaramanga con base al criterio dado por un experto.

1.2.2 Objetivos específicos:

- Realizar una revisión de literatura sobre las aplicaciones de los modelos máquinas de soporte vectorial y árboles de decisión.

- Caracterizar los albergues temporales propuestos en la ciudad de Bucaramanga.
- Desarrollar modelos de clasificación supervisada utilizando máquinas de soporte vectorial y árboles de decisión para la evaluación de la vulnerabilidad estructural de albergues temporales a través del criterio de un experto.
- Determinar el algoritmo de clasificación supervisada que mejor describa la evaluación de vulnerabilidad sísmica actual de albergues en Bucaramanga.
- Realizar un artículo académico de carácter publicable basado en los resultados de la investigación.

2. Revisión de la literatura

La literatura en logística de emergencia clasifica las etapas de un desastre de acuerdo a si consideran las operaciones realizadas antes o después del impacto del mismo. Las operaciones realizadas antes de la ocurrencia de un evento devastador se denominan operaciones pre-desastre y desempeñan un papel importante en la planificación estratégica o mitigación de la catástrofe. Las operaciones realizadas después de la ocurrencia de un desastre se denominan operaciones post-desastre que se llevan a cabo inmediatamente después de la ocurrencia del mismo y comprende actividades como distribución de socorro, evacuación de personas desplazadas, transporte y tratamiento de víctimas de desastres (Caunhye, Nie y Pokharel, 2012, p. 5).

Se identifican diferentes enfoques para llevar a cabo una evaluación de vulnerabilidad. Dependiendo si el objetivo es económico, social u ambiental se deben seleccionar los factores

más relevantes y un modelo que permita representar la realidad de la situación a evaluar. Aunque no se pueda predecir con exactitud cuándo y cómo se presentará una eventualidad si se puede conocer con cierto grado de incertidumbre el nivel del riesgo al que están expuestas las personas y, así mismo, tomar las medidas necesarias para estar preparados al momento en que ocurra un evento.

El aumento en el número de desastres ocurridos y los fuertes impactos que estos han provocado en los últimos años se ha despertado el interés de los autores sobre investigaciones en temas de estudios de desastres.

Algunos terremotos ocurridos presentan daños en las infraestructuras de edificios (Safina Melone, 2003, pp. 6-8), como se mencionan a continuación:

- Terremoto de Alaska (1964): alcanzó una magnitud de 7,9 en la escala de Richter; colapso del Hospital Elmendorf, instituciones educativas afectadas y daño en los servicios públicos.
- Terremoto de Caracas (1967) – Venezuela: se registra un movimiento telúrico de magnitud 7,9 en la escala de Richter; se genera colapso en diferentes edificaciones.
- Terremoto de San Francisco (1971) – Estados Unidos: registro una magnitud de 6,8 en la escala de Richter; diferentes hospitales salieron de operación debido a los daños generados: colapso del Hospital Olive View que fue construido semanas antes de ocurrido el evento sísmico y el Hospital Holy Cros tuvo que ser demolido después del sismo.
- Terremoto de Tachira (1981) – Venezuela: alcanzó una magnitud de 5,4 en la escala de Richter; se produjo agrietamiento en la mampostería del Hospital Central de San Cristóbal, edificio de la gobernación y policía técnica judicial de San Cristóbal.

- El Terremoto de Popayán (1983) – Colombia: alcanzó una magnitud de 5,5 en la escala de Richter; se afectaron más de 10 edificaciones esenciales; la capacidad de funcionamiento del Hospital Universitario San José se reduce una vez ocurrido el sismo.

Los artículos consultados y analizados a continuación fueron desarrollados para la etapa pre-desastre, en donde mediante la aplicación de un determinado modelo se pueda predecir el grado de vulnerabilidad y sus respectivas consecuencias ante la presencia de un desastre.

2.1 Gestión de desastres

Según Maldonado Rondón, Chio Cho y Gómez Araujo (2007), para el desarrollo de un modelo de evaluación de vulnerabilidad generalmente es necesario contar con información sobre eventos que hayan ocurrido en el pasado. Es por eso que en esta investigación se construyó un modelo aplicable a zonas que no cuentan con información histórica de sismos. El objetivo de este trabajo fue obtener índices de vulnerabilidad sísmica para edificaciones construidas en mampostería teniendo en cuenta la opinión de expertos nacionales e internacionales y las características de mayor impacto en edificaciones de mampostería ante un evento sísmico. Para reducir la subjetividad de la opinión de los expertos se aplicaron técnicas difusas para que por medio de un promedio ponderado difuso se obtuviera el índice de vulnerabilidad de la edificación de mampostería. Por otro lado, se utilizó el programa IVEM (índice de vulnerabilidad sísmica en edificaciones de mampostería basado en la opinión de expertos) para analizar la vulnerabilidad siendo ésta una herramienta sistematizada y aplicable a cualquier ciudad.

De acuerdo con Benjamin Hernandez y Lockhart Castro (2011), la evaluación de vulnerabilidad sísmica de las edificaciones puede hacerse con métodos cualitativos o cuantitativos. En República Dominicana se llevó a cabo una evaluación de vulnerabilidad sísmica de edificaciones de hormigón armado con el método FEMA-154; es un método cualitativo utilizado por la Federal Emergency Management Agency (FEMA) en los Estados Unidos. Emplean un formulario en el que se recogen todos los datos necesarios de la infraestructura para calcular el índice de vulnerabilidad. Dependiendo del resultado se toma la decisión de reforzar la estructura, no reforzarla o emplear un método de evaluación cuantitativo que mejor represente la situación real de la edificación. El formulario registra datos como la localización, número de pisos, año de construcción, área de construcción, foto de la edificación, tipo de suelo y número de personas que la ocupan.

Khamespanah, Delavar, Alinia, y Zare (2013), integraron el modelo de computación granular con la teoría de Dempster-Shafer para extraer las reglas de clasificación que permitan identificar las zonas y su respectiva intensidad de vulnerabilidad sísmica en Teheran, capital de Irán. La computación granular se centra en grupos de datos que han sido clasificados previamente en diferentes clases o granularidades. Los factores que se tuvieron en cuenta para ejecutar el modelo fueron el número de pisos, edad y calidad de los edificios y la pendiente e intensidad del terremoto. La teoría de Dempster-Shafer se utiliza para combinar diferentes puntos de vista y conseguir una decisión unificada de un grupo de expertos. Debido a su posición en el sistema montañoso, Irán es catalogada como una zona de alto riesgo sísmico y al ser eventos impredecibles surge la necesidad de desarrollar herramientas como mapas de vulnerabilidad sísmica que ayuden a las organizaciones de manejo de desastres a planificar e

implementar medidas de seguridad pertinentes para reducir el impacto de un terremoto y sus consecuencias.

Fallah Aliabadi, Sarsangi, y Modiri (2015), propusieron un nuevo enfoque utilizando el método AHP con un sistema de información geográfica (GIS) y la opinión de expertos obteniendo dos mapas de vulnerabilidad. Se identifican varios enfoques en las evaluaciones de vulnerabilidad sísmica; para este caso se tuvieron en cuenta factores sociales y físicos para evaluar la vulnerabilidad global del distrito de Fahadan en la ciudad de Yazd. Si bien el 90% de las muertes directas se debieron a la caída de edificios, el factor físico no es el único que influye en la vulnerabilidad de una zona, factores económicos y sociales también tiene gran relación con la vulnerabilidad: las personas de bajos recursos suelen vivir en casas con materiales débiles como adobe y barro o en zonas de asentamiento informal y de alto riesgo sísmico. En 1970 se introdujo la valoración de un factor social en las evaluaciones de vulnerabilidad dejando a un lado la idea de tomar decisiones basándose únicamente en los factores físicos. La vulnerabilidad social es una medida tanto de la sensibilidad de una población a los peligros naturales como de su capacidad para responder y recuperarse de un desastre.

2.2 Aplicaciones del aprendizaje automático

A mediados de siglo XX surge la idea de la inteligencia artificial o inteligencia computacional y no más adelante en el año 1958 se formaliza la IA durante la conferencia de Dartmouth. La principal motivación desde tiempos remotos de esta ciencia computacional se fundamenta en intentar construir máquinas que puedan pensar como el ser humano o al menos emular cierta capacidad de su conocimiento. Diversas disciplinas científicas se han esmerado en lograr reproducir el conocimiento humano a través de sistemas computacionales que

permitan realizar de forma automática tareas hechas por el hombre. Marvin Minsky define la inteligencia artificial como “la ciencia orientada al diseño y construcción de máquinas que implementen tareas propias de humanos dotados de inteligencia” (Pajares Martinsanz & Santos Peñas, 2006).

Las metodologías de la inteligencia artificial incluyen, la lógica difusa, sistemas de expertos, aprendizaje, sistemas de planificación, algoritmos de búsqueda, sistemas híbridos, diagnóstico de fallos y reconfiguración, autómatas, redes de Petri, redes neuronales, computación evolutiva, etc (Pajares Martinsanz & Santos Peñas, 2006, pág. 6). Los hitos en la evolución de la inteligencia artificial son:

- En el siglo XVII, el filósofo Gottfried Leibniz propuso construir una máquina capaz de llevar a cabo cualquier proceso de razonamiento por el que se le preguntara.
- En el siglo XIX, Charles Babbage, cuyo “ingenio analítico” era capaz de realizar operaciones algebraicas, pensó que se podría implementar un sistema formal en el cual toda la pluralidad de los fenómenos físicos pudiese ser adecuadamente descrita.
- En el siglo XX, en 1936, Alan Turing diseñó una máquina que pensara y realizara tareas en un dominio específico de forma equiparable a los hombres (Pajares Martinsanz & Santos Peñas, 2006, pág. 10)

Tabla 2.

Desarrollo histórico de la investigación artificial según Jackson (1989)

Periodo	Investigaciones desarrolladas
Periodo clásico (1960 – 1965)	<ul style="list-style-type: none"> • Gestación [McCulloch y Pitts, Shannon, Turing] • Inicio – Darmouth Collage, 1956 [Minsky, McCarthy] • Redes neuronales, robótica (Shakey) • Búsqueda en un espacio de estados, heurísticas, LISP • Resolvedor general de problemas (GPS) [Newell, Simon] • Juegos, pruebas de teoremas • Limitaciones como consecuencia de la dedicación a la búsqueda pura, explosión combinatoria.
Periodo romántico (1965 – 1975)	<ul style="list-style-type: none"> • Representación “general” del conocimiento • Redes semánticas [Quillian] • Prototipos (frames) [Minsky] • Perceptrón [Minsky y Papert] • Lógica [Kowalski] • Mundo de bloques (SHDRLU) [Winograd] • Compresión de lenguaje, visión, robótica • Dificultades de representación general, problemas “juguete”.
Periodo moderno (1975 - Hoy)	<ul style="list-style-type: none"> • Inteligencia “Específica” vs “general” • Representación explícita del conocimiento específico del dominio • Sistemas expertos o basados en conocimiento • Regreso de redes neuronales [Hopfield, Rumelhart, Hinton], algoritmos genéticos [Holland, Goldberg], reconocimiento de voz (HMM), incertidumbre (RB, Lógica difusa), planificación, aprendizaje. Aplicaciones “reales” (medicina, finanzas, ingeniería, exploración, etc.) • Limitaciones: conocimiento “superficial”.

Nota: Adaptado de Pajares Martinsanz & Santos Peñas. (2006). Inteligencia artificial e ingeniería del conocimiento. (Páginas 13- 14). Capítulo I. Mexico D.F. Alfa omega grupo editor.

2.2.1 Aplicaciones de máquinas de soporte vectorial: Según Méndez, Fdez Riverola, Díaz y Corchado (2007), la detección de correos Spam es una de las aplicaciones del

aprendizaje automático. Según Spamhaus, organización internacional sin fines de lucro que controla el spam y amenazas cibernéticas relacionadas, el término spam aplicado a mensajería electrónica hace referencia a aquellos correos electrónicos enviados de forma masiva y que no han sido solicitados por sus destinatarios. Los correos spam se pueden detectar y filtrar mediante modelos basados en la colaboración de usuarios o basados en el análisis de contenido. Los modelos basados en la colaboración de usuarios pierden precisión en la medida que se agregan nuevos datos y en ocasiones se hace necesaria la reconstrucción completa del modelo. De modo que, para obtener predicciones precisas en la clasificación de textos se utilizan los modelos basados en el análisis del contenido, a fin de usar técnicas de aprendizaje automático capaces de realizar las adaptaciones en cada situación y poder obtener resultados fiables sin afectar la estructura del modelo. Algunos modelos de clasificación utilizados para la identificación de correos spam son: SVM, Flexible Bayes, Random Forests, Casos de Similitud, Spam Hunting, entre otros.

Venegas (2007) tiene como objetivo en su investigación la clasificación de textos académicos en función de su contenido léxico-semántico, utilizando y comparando dos métodos de categorización automática (Naives Bayes y máquinas de soporte vectorial). La muestra utilizada en este problema está conformada por 216 textos (30.886.081 palabras) divididos en cuatro áreas del conocimiento e ingeniería: 26 usados en Ingeniería en Construcción, 31 en Química, 64 en Trabajo Social y 95 en Psicología. En la metodología de esta investigación se realizó una definición de características que se encuentran asociadas a la clasificación de textos por áreas de conocimiento, por lo que se plantea una etapa que permita construir un archivo de texto agrupado para cada área disciplinaria que incluya todos los textos correspondientes a la

temática. Durante este proceso: se eliminan lexemas que no aportan contenido semántico e información no lingüística; se identifican todos los sustantivos, verbos y adjetivos presentes en cada uno de los cuatro archivos; se determinan cuáles de estos eran compartidos entre los cuatro documentos de texto y se construye una matriz en la que se registra la frecuencia de cada uno de los lexemas compartidos por cada uno de los textos que componen los archivos. Usando la determinación de valor umbral para la frecuencia del documento (DF) y ganancia de la información (IG) se obtuvieron 74 variables útiles para la clasificación de textos por áreas disciplinarias.

Finalmente se realiza la etapa de entrenamiento, prueba y clasificación general para cada uno de los modelos. Para el análisis de los resultados se realiza una matriz de confusión que permita observar el comportamiento de los modelos en cada etapa. Para determinar el modelo más eficiente se utiliza el porcentaje de mala clasificación. En los resultados obtenidos se plantea que las máquinas de soporte vectorial clasifican más eficientemente los textos académicos, con altos niveles de precisión y exhaustividad.

Velásquez, Olaya y Franco (2009) en su investigación proponen una metodología para seleccionar las constantes (σ , C , ϵ) que controlan el desempeño de una máquina de soporte vectorial (SVM) con kernel RBF o kernel gaussiano; esto con la finalidad de desarrollar problemas de predicción de series de tiempo. La búsqueda de estas constantes se realiza mediante una técnica de búsqueda discreta donde se evalúan secuencialmente los puntos en la vecindad, este proceso termina cuando se logra disminuir el error de ajuste. El segundo objetivo es presentar una metodología de SVM para pronosticar cinco series de tiempo (AIRLINE,

POLLUTION, INTERNET, LYNX y PAPER) del mundo real que presentan diferentes características estructurales comparándolas con otros modelos (ARIMA y MLP). Por último, se difunde el uso de las SVM como una técnica competitiva para la predicción de series de tiempo que puedan producir mejores pronósticos que otros métodos tradicionales.

Rodríguez y Martínez (2015) proponen en su investigación el uso de las máquinas de soporte vectorial (SVM) para clasificar las transformaciones en las células de la región del cuello uterino en condiciones normal y anómala. Para la ejecución del modelamiento se hace uso del lenguaje Matlab y las funciones propias del mismo, tomando como referencia la base de datos Herlev que consta de 917 imágenes de células previamente clasificadas donde se consideran un conjunto primario de rasgos geométricos: área, perímetro, diámetro corto, diámetro más largo y la redondez. La metodología de la investigación propone seleccionar el 80% de los datos para el entrenamiento del clasificador y el 20% restante para la prueba del modelo. Los parámetros considerados para la definición de la máquina de soporte vectorial son el tipo de kernel (Lineal y RFB) y para evaluar el rendimiento del clasificador se utilizan los siguientes indicadores: la sensibilidad, la especificidad, la predictividad positiva y negativa, y la tasa de clasificación correcta para evaluar el rendimiento del clasificador.

2.2.2 Aplicaciones de Árboles de decisión: Calle y Sánchez (2007) presentan la integración de los árboles de clasificación y regresión en la investigación clínica. Con un total de 146 datos de pacientes con cáncer de próstata buscaron identificar los factores predictores de la posible progresión de un tumor. La muestra fue segmentada en dos grupos homogéneos a través de los cuales las particiones que sucesivamente se realizaban en los nodos generaban un

aumento en el poder predictivo del árbol. Mediante el índice Gini se calculó el porcentaje de impureza presente en cada nodo eligiendo así las particiones que tenían mayor reducción de impureza. Por último, para controlar el sobreajuste de los datos, se utilizó el parámetro de complejidad que permitió llegar al árbol óptimo en relación a la complejidad del modelo y el error de clasificación.

Şengezer, Ansal y Bilen (2008) aplicaron las técnicas de árboles de decisión a una base de datos que contiene información relevante de los daños de terremotos recientemente ocurridos en 9.400 edificios de las ciudades de Erzincan, Dinar y Kocaeli con el objetivo de identificar los factores que controlan el daño de los edificios durante un terremoto para evaluar sus efectos y poder así desarrollar modelos de clasificación para daños sísmicos en las edificaciones de Turquía. Los factores que se consideraron influyentes en la vulnerabilidad de las estructuras fueron la magnitud del terremoto (M), la intensidad (I) en la ciudad, la distancia al epicentro (ED), los tipos de edificios (BT), el número de pisos (NS), condiciones del suelo (SS), posición de construcción (BP) en el sitio y, las condiciones del sitio (SC). Una vez probados los 4 algoritmos de árboles de decisión, se determinó que el algoritmo Chaid presenta los mejores resultados, en donde se clasificó la vulnerabilidad en 6 categorías (sin daños, daños leves, daños moderados, daños graves, parcialmente en colapso y totalmente colapso) y 3 códigos de colores (rojo, amarillo y verde). El modelo entrenó varias veces combinando los parámetros, es decir, el nivel de daños en función de las combinaciones de los parámetros y se concluye que el tipo de construcción es el factor que más afecta las estructuras de los edificios en Turquía y, en segundo lugar están las condiciones del sitio y el número de pisos.

Debido a los problemas de eficiencia y escalabilidad que presentan los métodos estadísticos tradicionales de análisis de datos, surge la necesidad de usar nuevos métodos que permitan el análisis de información masiva generada a diario. Es por esto que Barrientos Martínez, et al. (2009) en su investigación, evalúan el desempeño de clasificación de datos médicos mediante arboles de decisión como herramienta de apoyo en la toma de decisiones del proceso de tratamiento y diagnóstico médico. Para el desarrollo del modelo se utilizaron dos bases de datos con características que los patólogos consideraron sintomatología de cáncer de seno, en donde la primera base de datos cuenta con datos emitidos por un solo especialista mientras que la segunda contiene datos emitidos por 19 especialistas. Edad, dimensión celular, lumina intracitoplasmática, agrupaciones de células epiteliales tridimensionales, núcleos bipolares, macrófago espumoso, nucléolos, pleomorfismo nuclear, tamaño nuclear, células epiteliales necróticas y cambios apócrinos fueron las características que se tuvieron en cuenta para emitir un diagnóstico. En el primer experimento se dividió aleatoriamente la primera base de datos, dos terceras partes del total para entrenar el modelo y una tercera parte del total para probarlo; en el segundo experimento se usaron dos terceras partes del total de la primera base de datos para el entrenamiento y el total de la segunda base de datos para prueba. A partir del porcentaje de casos clasificados correctamente se analizó cual base de datos representa mejor la situación real con el modelo propuesto, concluyéndose que la combinación de diferentes opiniones de expertos incrementa la subjetividad en el resultado así que la mejor clasificación se logró con la base de datos emitidos por un solo especialista.

La gestión de desastres a gran escala es una tarea compleja y ha sido atendida en base a la experiencia de los gestores de las emergencias, pero las limitaciones de recursos y el flujo de

información en tiempo real que se maneja, generan la necesidad de un sistema robusto de apoyo a las decisiones. Romanowski et al. (2015), en su investigación proponen una metodología de apoyo basada en datos históricos de emergencias que proporciona recomendaciones específicas para las necesidades de recursos en respuesta a un evento de gran escala. Esto permite a los administradores de emergencias ubicar y organizar los recursos en una comunidad y determinar la combinación adecuada de voluntarios, profesionales y representantes de la comunidad; aunque se aplicó en una región de tamaño medio esta metodología está diseñada para una región de cualquier tamaño. Se consideró una base de datos que contenía 7 años (2005-2012) de llamadas al 911, es decir, más de 8 millones de registros de casos y 100 millones de transcripciones de llamadas realizadas durante los eventos en el condado de Monroe, Nueva York. El caso de estudio se enfocó en la ocurrencia de incendios y 2 etapas se llevaron a cabo en el desarrollo del modelo: primero se identificaron 4 clúster con datos de grandes incendios estructurales empleando la técnica k-means para definir las características de los diferentes tipos de incendios, y así mismo poder identificar los recursos necesarios; después emplearon los árboles de decisión para realizar la predicción de los recursos. Finalmente se observó que para las estaciones grandes con más de 10 unidades los incendios consumieron entre el 10% y el 18% de los recursos disponibles y para estaciones pequeñas el clima, la estación del año y la hora son factores que también influyen en la respuesta.

3. Marco de antecedentes

Mosquete Rosario (2012) en su trabajo “Evaluación del riesgo sísmico en edificios especiales: escuelas. Aplicación a Barcelona” emplea el método del índice de vulnerabilidad

(MIV) igualmente conocido como Risk-UE nivel I, una técnica avanzada de análisis de riesgo sísmico apoyada en el sistema de información Geográfica. En la metodología de esta investigación se realiza una matriz de tipológica de edificios educativos en la que consolidan diferentes características de las construcciones a fin de realizar una clasificación de vulnerabilidad de los colegios y posteriormente efectuar los análisis cuantitativos de la información. A partir de la elaboración de esta matriz, se desarrollan los modelos de predicción de daños y pérdidas para cada tipología, las probabilidades de los estados de daño se obtienen suponiendo una distribución de probabilidad binomial o beta equivalente. Las consolidaciones de los resultados obtenidos en este trabajo se presentan en un mapa de riesgos con seis escalas de vulnerabilidad definidas por el autor: Nulo, leve, moderado, severo, extensivo y colapso.

Cuevas (2014) en su Tesis doctoral “Evaluación de la vulnerabilidad sísmica urbana basada en tipologías constructivas y disposición urbana de la edificación. Aplicación en la ciudad de Lorca, región de Murcia”. Busca analizar procedimientos y metodologías basados en tipologías constructivas para poder caracterizar y graduar los parámetros urbanísticos que hacen que las ciudades sean más o menos vulnerables sísmicamente. En dicho estudio se emplea una metodología empírica basada en la experiencia sobre el comportamiento de diferentes tipos de edificaciones durante sismos con base en un sistema de información geográfica (SIG). En el desarrollo de este trabajo se definen los parámetros urbanísticos fundamentales que determinan una respuesta sísmica irregular en las edificaciones, algunos de ellos son: la irregularidad geométrica en planta, irregularidad geométrica en alzado, piso blando y débil, efecto de la columna corta y cautiva, tipología edificatoria, entre otros. Una vez identificados y

caracterizados estos parámetros urbanísticos, se estudia la mayor o menor relación que pueden tener con el daño producido en las edificaciones tras un terremoto.

Los investigadores Barrera y Nieves (2015) realizan su trabajo de grado para optar el título de ingeniero civil de la universidad de Cartagena en el año 2015 titulado “Determinación de la vulnerabilidad en las casas coloniales ubicadas en el barrio de San Diego de la ciudad de Cartagena”. Esta investigación tiene como objetivo determinar el índice de vulnerabilidad en el barrio San Diego considerado patrimonio histórico de la ciudad de Cartagena.

El desarrollo de esta investigación se dividió en dos etapas. La primera fase es la exploratoria; en ella los autores realizan observaciones de las casas coloniales del barrio San Diego y efectúan una selección de 5 casas para la muestra en esta investigación. La segunda fase consistió en un análisis de la información. Esta investigación es de carácter cualitativo en la que no se incluye modelamiento matemático. El procedimiento empleado durante el desarrollo de este trabajo se basó en el índice de vulnerabilidad considerando un método cualitativo. Los investigadores definen los rangos de vulnerabilidad para el estudio.

$Vulnerabilidad < 15\%$ *Baja*

$15\% \leq Vulnerabilidad < 35\%$ *Media*

$Vulnerabilidad \geq 35\%$ *Alta*

Los autores realizan una inspección visual detallada de la muestra, consideran once parámetros para determinar la vulnerabilidad de cada casa coloquial, estos fueron determinados bajo consultas con expertos en el tema y concepto propio de los autores. Para realizar el cálculo

de cada uno de los parámetros, los autores requieren información como: Año de construcción, normatividad bajo la cual se construyó la estructura, materiales empleados, forma de la mampostería, número de pisos, pendiente del terreno, distancia entre muros, tipo de cubierta e identificación de grietas. Con esta información se calcula un peso W_i para cada parámetro y se determina el índice de vulnerabilidad para cada casa colonial en la muestra y se clasifica entre los rangos ya definidos. Posteriormente se realiza una tabulación consolidada con las 5 casas de la muestra y sus índices de vulnerabilidad correspondientes para obtener un índice de vulnerabilidad promedio para la zona de estudio.

Se concluye y realiza recomendaciones para las infraestructuras considerando juicios de expertos en las que se incluyen: corregir humedad, fisuras, corrosión, goteras, ligar elementos no estructurales a las edificaciones como cornisas y parapetos e incluso sustituir o remplazar maderas deterioradas de la cubierta.

4. Marco Teórico

4.1 Aprendizaje automático.

“Se dice que un programa de computadora aprende de la experiencia E con respecto a alguna clase de tareas T y medida de rendimiento P , si su desempeño en tareas en T , medida por P , mejora con la experiencia E ” (Mitchell, 1997, p. 2).

Macías, Gómez, Tous y Torres (2015) definen aprendizaje automático (en inglés, Machine learning) al hecho de que los ordenadores pueden “aprender” mediante programas capaces de

generalizar y automatizar comportamientos a partir de los datos de los que dispone. El aprendizaje automático devuelve como resultado un modelo para predecir datos y comportamientos futuros a partir de los datos usados para el aprendizaje (p. 137).

Esta rama de la inteligencia artificial es útil en la medida en que se logra conservar a lo largo del tiempo la experiencia de los expertos y así mismo nutrir continuamente el modelo desarrollado haciendo que este sea más acertado en sus predicciones. El aprendizaje automático facilita la toma de decisiones a priori ante circunstancias imprevistas en los diferentes campos del conocimiento, minimizando recursos económicos y tiempos. En la figura 1, se ilustra la división del aprendizaje automático en aprendizaje supervisado y aprendizaje no supervisado, junto con sus algoritmos de aplicación.

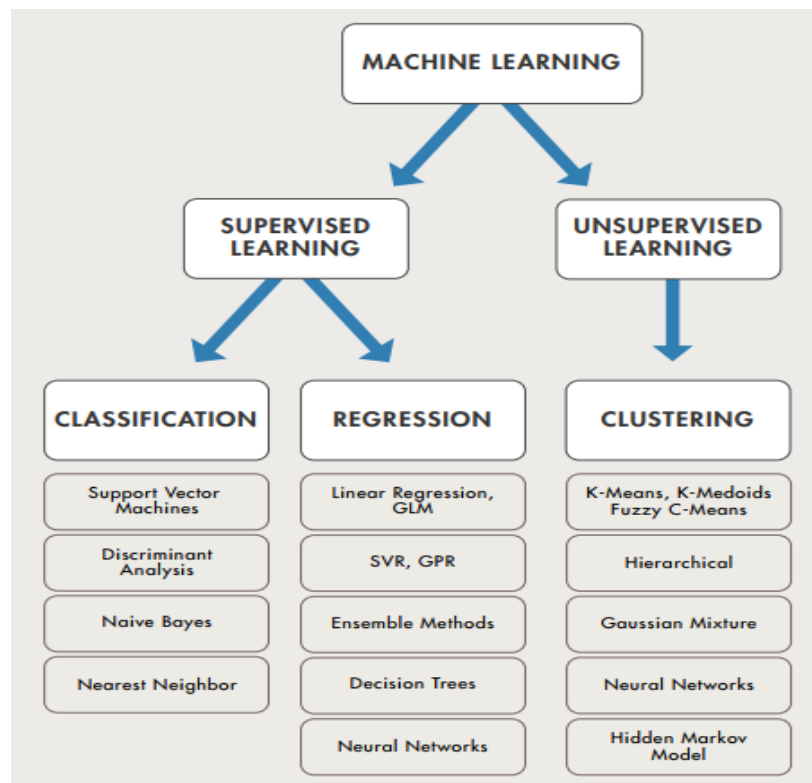


Figura 1. Clasificación del Aprendizaje Automático. Adaptado de Math Works. (2016).

https://www.mathworks.com/tagteam/89703_92991v00_machine_learning_section1_ebook_v12.pdf.

4.1.1 Aprendizaje supervisado: El aprendizaje supervisado se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar el sistema a partir de una entrada determinada. El supervisor controla la salida del sistema y en caso de que ésta no coincida con la deseada, se procederá a modificar los parámetros, con el fin de conseguir que la salida obtenida se aproxime a la deseada. (Thapelo, 2014, p. 20).

El aprendizaje supervisado requiere que el supervisor defina a priori la clasificación de los datos o etiquetas de salida. El objetivo del aprendizaje supervisado es lograr que un modelo clasifique acertadamente los datos de acuerdo a los parámetros de entrada suministrados por los expertos sin intervención humana.

4.1.2 Aprendizaje no supervisado: En la categoría sin supervisión (llamada también aprendizaje no supervisado), el problema es descubrir la estructura del conjunto de datos si es que hay alguna. Esto generalmente significa que el usuario quiere saber si hay grupos en los datos, y qué características hacen que los objetos sean similares dentro del grupo y diferentes a través de los grupos. (Kuncheva, 2004, p. 2).

A diferencia del aprendizaje supervisado, este no requiere definir previamente las clasificaciones (Etiquetas) ya que las salidas determinan las agrupaciones de un conjunto de datos. En la literatura es común encontrar algoritmos de Clustering para el desarrollo del aprendizaje no supervisado.

4.2 Sistema de expertos

Barr, Feigenbaum y Roads (1982) definen la Inteligencia Artificial (AI) como la parte de la informática que se ocupa del diseño de sistemas de computación inteligentes, es decir, sistemas que muestran las características que asociamos a la inteligencia en el comportamiento humano que se refiere a comprensión del lenguaje, aprendizaje, razonamiento, resolución de problemas, etc (p. 3).

García y Britos (como se citó en Sierra et al., 2005) definen los sistemas expertos como una clase de programas que son capaces de manejar problemas que normalmente requieren para su resolución de intervención humana especializada. Los sistemas de expertos son un gran avance de la Inteligencia Artificial y se encargan de crear un modelo que permita generalizar y replicar el conocimiento de los expertos con el objetivo de resolver problemas con los datos de entrada suministrados.

Todo sistema experto está constituido por los siguientes componentes (Quesada, 2007, p. 52):

- La adquisición del conocimiento: por lo general, un ingeniero del conocimiento se encarga de recopilar toda la información disponible sobre el campo de acción del sistema y, además, realiza las entrevistas o rastreos pertinentes a expertos en el área de conocimiento.
- La base de conocimiento: que está formada por hechos (la información documentada y codificada) y por reglas que usualmente son heurísticas (a manera de predicados) que le dan al sistema los “volados” para resolver las consultas de los usuarios.

- El motor de inferencias: se considera el cerebro del sistema, define una metodología de búsqueda para utilizar la base de conocimiento y las consultas de los usuarios con la finalidad de brindar una respuesta.
- Interfase: posibilita la comunicación hombre-máquina.
- El subsistema de explicación o justificador: una particularidad de los sistemas expertos es contar con un módulo de explicaciones, el cual se encarga de explicarle al usuario por qué se ha llegado a una determinada conclusión o recomendación; inclusive algunos sistemas pueden explicar por qué un determinado estado del problema no se eligió como meta.

Los estudios de vulnerabilidad sísmica de las edificaciones se realizan en su gran mayoría con el criterio de un experto, que debido a su experiencia es quien da un juicio sobre la probabilidad de una eventualidad no deseada.

4.3 Validación cruzada

Para comparar dos modelos existen numerosas técnicas estadísticas dentro de las cuales se encuentra la validación cruzada. Esta técnica se recomienda como una de las mejores formas de probar un modelo, pues se introduce un sesgo al probar la validez de un modelo con los mismos datos con los que fueron estimados sus parámetros (Cooil et al., 1987). La validación cruzada hace uso de distintos subconjuntos de los datos disponibles para realizar el entrenamiento del modelo y su posterior validación (Arlot y Celisse, 2010).

4.3.1 K-folds: El K-Folds se utiliza en el campo de máquinas de aprendizaje para determinar la precisión con la que un algoritmo será capaz de predecir datos que no fueron entrenados en él (Kale, Mehrotra y Manza, 2010, p. 540). En el campo de aprendizaje automático, el rendimiento de un clasificador se mide generalmente en términos de error de predicción. (Rodríguez, Pérez, & Lozano, 2010, p. 569).

En este método el conjunto de datos inicial se divide en K subconjuntos; de los K subconjuntos un solo subconjunto se guarda como datos de validación para probar el modelo, y los otros K-1 subconjuntos son usados como datos de entrenamiento. El proceso *crossvalidation* se repite K veces (particiones), donde cada uno de los K subconjuntos se usa exactamente una vez como datos de validación. Los resultados de las particiones se promedian (o se combinan de otra manera) para producir un solo estimador de la exactitud del clasificador (ver figura 2) (Rodríguez et al., 2010, p. 9).

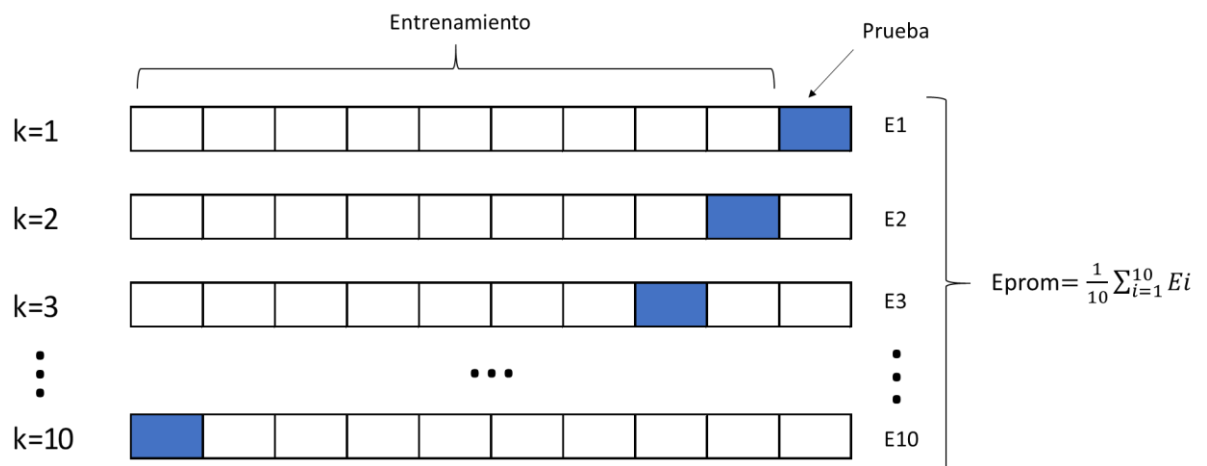


Figura 2. Algoritmo de evaluación VC. Adaptado de P. Mayorga Ortiz, C. Druzgalski, M. A. Criollo Arellano, O. H. González Arriaga (2013). <http://www.scielo.org.mx/pdf/rmib/v34n2/v34n2a2.pdf>.

4.3.2 Hold-out: Cuando no se proporciona una muestra de prueba independiente a la muestra de entrenamiento, se selecciona aleatoriamente una porción de la muestra de entrenamiento para realizar la prueba y se construye un clasificador con la muestra restante. La tasa de error real del clasificador construido es estimada con la muestra de prueba seleccionada; todo este proceso se repite muchas veces, y el promedio de las estimaciones de tasa de error obtenidas repetidamente se denomina repeated Hold-out estimate (estimación de retención repetida). (Rodríguez et al., 2010, p. 3739).

Las metodologías de validación cruzada K-Fold y Hold-out tienen gran similitud en el procedimiento para evaluar un algoritmo. Los datos disponibles se dividen aleatoriamente en un conjunto de entrenamiento y un conjunto de prueba. El conjunto de entrenamiento se subdivide, a su vez, en dos conjuntos Disjuntos (Herrera, Hervás, Otero, & Sánchez, 2004, p. 2):

- El conjunto de estimación, usando para seleccionar el algoritmo.
- El conjunto de validación, usando para probar o validar el algoritmo.

Una ventaja que tiene la metodología K-Fold es que hace uso de todos los datos, es decir, todos los k subgrupos se integran tanto en las pruebas de entrenamiento como en las de pruebas de validación generando mayor fiabilidad en el resultado. Por otro lado, el método Hold-out, al escoger al azar el subgrupo de validación no asegura la contribución de todos los datos a las dos pruebas lo que podría sesgar la información, en ocasiones se realizan varias iteraciones con el objetivo de ponderar los resultados y poder así mitigar el riesgo de trabajar con muestras que no sean representativas.

4.4 Máquinas de soporte vectorial.

Una Máquina de Soporte Vectorial (SVM) aprende de la superficie decisión de dos clases distintas de los puntos de entrada. La descripción dada por los datos de los vectores de soporte es capaz de formar una frontera de decisión alrededor del dominio de los datos de aprendizaje con muy poco o ningún conocimiento de los datos fuera de esta frontera. Los datos son mapeados por medio de un kernel Gaussiano u otro tipo de kernel a un espacio de características en un espacio dimensional más alto, donde se busca la máxima separación entre clases. Esta función de frontera, cuando es traída de regreso al espacio de entrada, puede separar los datos en todas las clases distintas, cada una formando un agrupamiento (Betancourt, 2005a, p. 67).

Esta estrategia de aprendizaje fue introducida por Valpnik y sus colaboradores, es un principio y método muy poderoso que en los pocos años desde su introducción ya ha superado a la mayoría de los otros sistemas en una amplia variedad de aplicaciones. (Cristianini & J. Shawe-Taylor, 2000, p. 7).

Mientras la mayoría de los métodos de aprendizaje se centran en minimizar los errores cometidos por el modelo generado a partir de los ejemplos de entrenamiento (error empírico), el sesgo inductivo asociado a las SVMs radica en la minimización del denominado riesgo estructural. (Betancourt, 2005b, p. 2)

En la figura 3 se ilustra el hiperplano que separa dos clases de datos, los límites mostrados con una línea a trazos definen el margen de distancia existente entre las clases, en una máquina

de soporte vectorial se pretende maximizar el margen de distancia y de este modo obtener el mínimo riesgo estructural.

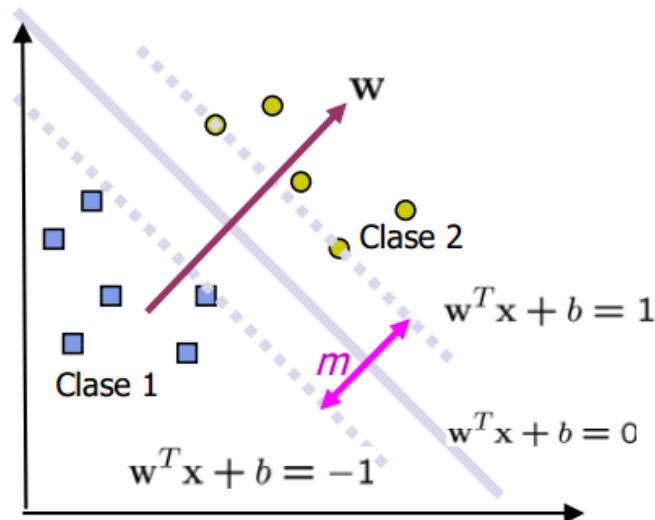


Figura 3. Máquinas de Soporte Vectorial. Adaptado de Gustavo A. Betancourt (2005).

<http://revistas.utp.edu.co/index.php/revistaciencia/article/view/6895/4139>

La SVM encuentra el hiperplano óptimo utilizando el producto punto con funciones en el espacio de características que son llamadas kernels (Betancourt, 2005a, p. 67). Este tipo de método se aplica para clasificaciones no lineales.

A la hora de definir el hiperplano, sólo se consideran los ejemplos de entrenamiento de cada clase que caen justo en la frontera de dichos márgenes. Estos ejemplos reciben el nombre de vectores soporte (Carmona Suarez, 2013, p. 2).

4.4.1 Función Kernel: La metodología para generalizar SVM se basa en un procedimiento denominado Kernel Trick (Aizerman *et al.*, 1964), que consiste en que cuando los datos de entrada no son linealmente separables, existe la posibilidad de transformar los datos a un espacio euclidiano de mayor dimensión (el espacio de características) en el que los puntos sí

pueden ser separados por un hiperplano (ver figura 4), y en el que aproximadamente se tiene una estructura lineal (Alpaydin, 2011; Fernando y Gutiérrez, 2011).

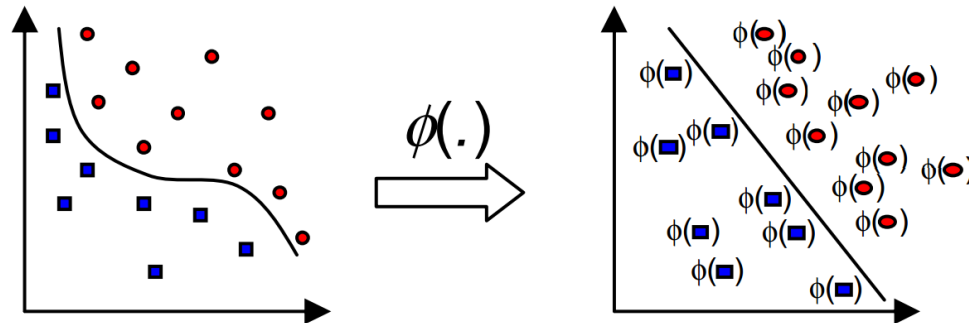


Figura 4. Truco del Kernel para el caso no linealmente separable. Adaptado de Gustavo A. Betancourt (2005). <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/6895/4139>

En esta investigación se emplea la función svm de la librería e1071, la cual requiere de un conjunto de parámetros necesarios para evitar el “Sobreajuste” del modelo, estos se muestran en la tabla 3.

Tabla 3.
Parámetros de las máquinas de soporte vectorial

Parámetro	Descripción
Cost	Este parámetro determina la penalización permitida sobre el margen de separación.
Gamma	El gamma define hasta donde llega la influencia de un conjunto de datos de entrenamiento
Degree	Es el grado del polinomio necesario para la función de kernel de tipo polinomial
Coef0	Es el coeficiente que acompaña la variable en cada función de kernel.

4.5 Árbol de decisión

El método de los árboles de clasificación y regresión se consolidó en 1984 con la edición del libro de Breiman et al, que sigue siendo una referencia básica. En cuanto a su aplicabilidad, existen diferentes implementaciones comerciales de estas técnicas, así como otras de software libre. Entre estas últimas, la más utilizada son las librerías tree y rpart en lenguaje estadístico R. La librería tree incluye las funciones básicas para trabajar con árboles, mientras que la librería rpart proporciona una implementación más rápida y con mayor número de funciones de soporte (Calle y Sanchez, 2007).

Un árbol de decisión es un modelo de predicción cuyo objetivo principal es el aprendizaje inductivo a partir de observaciones y construcciones lógicas. Son muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que suceden para dar solución de un problema. (Barrientos et al., 2009, p. 20). Si la variable dependiente es continua, se genera un árbol de regresión y si la variable dependiente es categórica, se genera un árbol de clasificación. El modelo de árbol de decisión permite predecir o asignar un elemento a una categoría previamente establecida, inicialmente se escoge la variable que mejor explique el conjunto de los datos dando como resultado el nodo raíz, éste se divide en 2 nodos hijos que contienen los posibles valores de los atributos. Luego se analizan los nodos individualmente y se repite el proceso hasta llegar a los nodos terminales que contienen el valor de clasificación correspondiente como se muestra en la figura 5. La partición de los nodos se hace en base a unos criterios estadísticos que se explicaran más adelante.

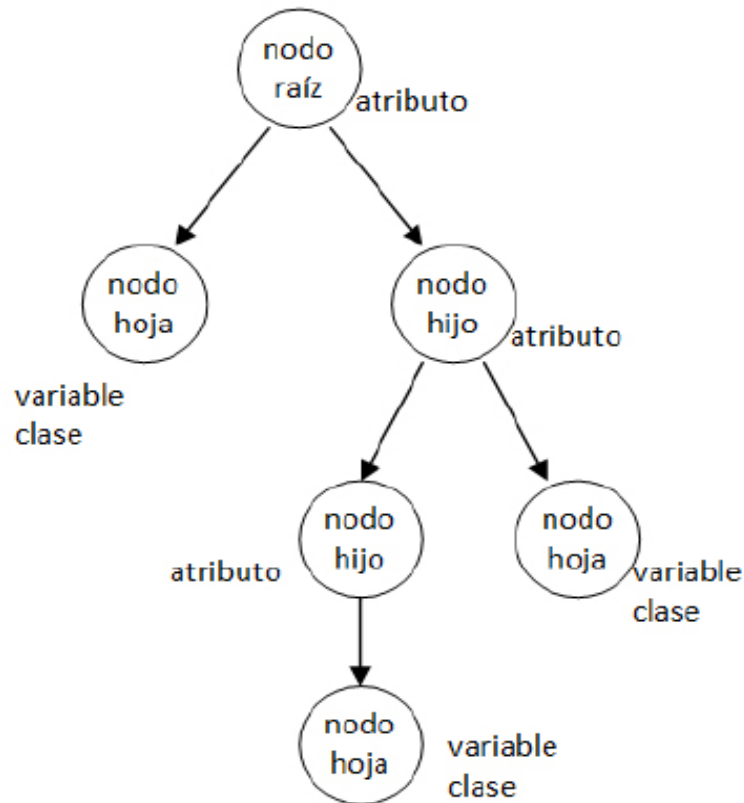


Figura 5. Árbol de decisión. Adaptado de Barrientos Rocio, Cruz Nicardo, Acosta Héctor & Rabatte Ivonne (2009). Árboles de decisión como herramienta en el diagnóstico médico. Facultad de física e inteligencia artificial. México. <http://www.medigraphic.com/pdfs/veracruzana/muv-200>

Los tres tipos de árboles más utilizados son: Chaid, Cart y Quest (Pérez Marqués, 2014, pág. 289), los cuales se explican en los siguientes apartados.

4.5.1 Cart: Genera un árbol de clasificación y regresión que permite predecir o clasificar observaciones futuras. El método utiliza la partición reiterada para dividir los registros de entrenamiento en segmentos minimizando las impurezas en cada paso, donde un nodo se considera “puro” si el 100% de los casos del nodo corresponden a una categoría específica del campo objetivo. Los campos de entrada y objetivo pueden ser continuos (rango numérico) o categóricos (nominal, ordinal o marca). Todas las divisiones son binarias (sólo se crean dos subgrupos) IBM (s.f).

4.5.1.1 Criterios de división: La impureza en los nodos de un árbol de decisión tipo CART se puede medir con la entropía, el índice de Gini o el error de mala clasificación.

4.5.1.1.1 Entropía: La entropía es una medida que permite determinar el grado de incertidumbre en la clasificación de una muestra. Puede presentarse que la muestra sea completamente homogénea esto quiere decir que se tiene un nivel de incertidumbre mínimo $h=0$ (asúmase h como entropía), como también puede presentarse que el nivel de incertidumbre es máximo $h=1$ esto se da cuando la muestra ha sido distribuida igual entre las clasificaciones, por lo que no se tiene absoluta certeza de que el árbol de decisión sea el más indicado ya que se presentarían conflictos a la hora de clasificar un nuevo conjunto de datos. (Wang & Suen, 1984, p. 407)-

$$Entropía(t) = - \sum_j p(j|t) \log_2 p(j|t)$$

4.5.1.1.2 Índice de Gini: Al igual que la entropía, el índice de Gini es una medida de impureza de la información. Su valor está comprendido entre 0 y 1.

$$GINI(t) = 1 - \sum_j [p(j|t)]^2$$

4.5.1.1.3. Error de mala clasificación: Se entiende por mala clasificación el evento de clasificar un dato en una categoría al que no pertenece, y se define como el cociente entre las observaciones mal clasificadas sobre el número total de observaciones.

$$Error(t) = 1 - \max[p(j|t)]$$

Donde $p(j|t)$ es la probabilidad de pertenecer a la clase “j” estando en el nodo t.

Los criterios de división son medidas que toman valores entre 0 y 1, donde 0 significa que el nodo es totalmente homogéneo, es decir que no se pueden realizar más particiones, y 1 para un nodo totalmente heterogéneo.

4.5.2. *Chaid*: Genera árboles de decisión utilizando estadísticos de chi-cuadrado para identificar las divisiones óptimas. A diferencia de los nodos CART y QUEST, CHAID puede generar árboles no binarios, lo que significa que algunas divisiones generarán más de dos ramas. Los campos de entrada y objetivo pueden ser continuos (rango numérico) o categóricos. IBM (s.f).

4.5.3. *Quest*: Proporciona un método de clasificación binario para generar árboles de decisión; está diseñado para reducir el tiempo de procesamiento necesario para realizar los análisis de CART y reducir la tendencia de los métodos de clasificación de árboles para favorecer a las entradas que permitan realizar más divisiones. Los campos de entrada pueden ser continuos (rango numérico), sin embargo, el campo objetivo debe ser categórico y todas las divisiones son binarias. IBM (s.f).

4.5.4. *Crecimiento del árbol*: Entre mayor sea la profundidad del árbol, mayor va a ser la probabilidad de que se genere un sobreajuste en los datos y la clasificación no sea tan eficiente. El objetivo de los modelos de árbol de decisión siempre será el de elegir las variables que mejor

expliquen los datos para obtener un árbol simple. Para el desarrollo de este trabajo se utilizaron 3 parámetros de la librería rpart que controlan el crecimiento del árbol, estos se muestran en la tabla 4.

Tabla 4.
Parámetros del árbol de decisión.

Parámetro	Descripción
Maxdepth	Hace referencia a la máxima profundidad del árbol, es decir los niveles que el árbol tendrá. Profundidad cero hace referencia al nodo raíz, y por debajo de este nodo se cuentan los demás niveles.
Minsplit	Determina el número de datos mínimo para que se ejecute la división de un nodo.
Cp	El parámetro de complejidad indica que, si el criterio de impureza no se reduce en más de $Cp \cdot 100$, entonces no se ejecuta la división de un nodo. Su valor oscila entre 0 y 1.

5. Caracterización de albergues temporales

Cada año más de 200 millones de personas resultan afectadas por las sequías, inundaciones, ciclones tropicales, terremotos, incendios forestales y otros peligros. Con este alarmante panorama es importante resaltar que el conocimiento y el acceso a información sobre gestión de desastres es una herramienta de vital importancia para gestionar y reducir los riesgos y consecuencias que ocasiona un sismo, salvar vidas, reducir el sufrimiento humano y las pérdidas materiales (Ulloa, 2011, pag.8).

La prevención consiste en la implementación de procesos y medidas con el objeto de evitar que se produzcan desastres o minimizar su impacto. Esto implica que la toma de decisiones debe darse antes de que se presente un fenómeno natural adverso (Lavell, 2001). Por tal razón, surge interés por parte de los investigadores en el desarrollo de trabajos enfocados en este tema, buscando proveer herramientas de apoyo que permitan identificar, planear y ejecutar programas de contingencia ante eventos catastróficos.

En este apartado se realiza una descripción de los factores asociados a la vulnerabilidad estructural de albergues temporales en Bucaramanga propuestos en este trabajo de investigación; y, posteriormente, se detalla la localización de los albergues temporales propuestos en otros trabajos de investigación a través de la herramienta de georreferenciación: Google Maps.

Para determinar los factores que tienen mayor influencia en la vulnerabilidad sísmica estructural de los albergues, se utilizó información secundaria, donde se identificaron los factores más relevantes propuestos en la revisión de literatura. Finalmente se consideraron 8 variables de entrada para los modelos, descritas a continuación:

5.1 Año de construcción

El año de construcción permite tener una idea acerca de la calidad de los materiales y las técnicas de construcción empleadas en una estructura. Después de la ocurrencia del terremoto de Popayán en el año 1983, se generó interés por parte de la Asociación Colombiana de Ingeniería Sísmica en crear el reglamento colombiano de construcción sismo resistente (Pinzón, 2014). La publicación de dicho reglamento en el año 1984 representa un hito importante en la

construcción de las edificaciones colombianas. Antes de la implementación de la norma las construcciones carecían de técnica y seguridad. A partir de 1984 las nuevas edificaciones tienen un diseño sísmo resistente que disminuye no solo los impactos y daños generados en las edificaciones sino también el número de personas afectadas.

5.2 Tipología de Estructuras

Colombia es un país que despertó su interés por mitigar los riesgos estructurales en consecuencia a las cifras de mortalidad registradas después de la ocurrencia de un sismo. Por esa razón se crea el reglamento colombiano de construcción sísmo resistente en 1984. Los daños ocasionados después de la acción de un sismo en una edificación dependen en gran medida de las características de construcción; las estructuras tienen como función principal soportar y resistir un conjunto de cargas, por consiguiente, se hace importante el desarrollo de un buen diseño que cumpla la normativa sísmo resistente. El movimiento que genera un sismo somete las edificaciones a fuerzas que van en sentido horizontal y vertical, lo adecuado sería que las estructuras tengan la capacidad de resistir un sismo sin sufrir daños graves, pero las personas lo evitan debido al incremento en los costos que esto representa.

La edificación en mampostería es considerada una de las tipologías más frecuentes en las ciudades colombianas (Maldonado Rondón, Chío Cho y Gómez Araujo, 2007). Además de los materiales empleados en dichas construcciones, los elementos estructurales toman vital importancia debido a la resistencia que proporcionan a los muros. Es por esto que cuando se presentan los movimientos de un sismo las estructuras que sufren daños más graves son las que no cuentan con elementos estructurales adecuados que disipen la energía transmitida por las ondas sísmicas dirigiendo el impacto directamente a los muros de la edificación. La tipología

de mampostería en adobe es considerada la más vulnerable; mampostería confinada, no confinada y reforzada presentan vulnerabilidad si no cuentan con elementos estructurales y muros ubicados uniformemente. Para este trabajo se consideraron las tipologías descritas en el libro visión estructural de las edificaciones en Bucaramanga y en los albergues propuestos se encuentran las tipologías: mampostería confinada y no confinada y pórtico dúctil y no dúctil. Según Maldonado y Chío (2007); las tipologías de las estructuras en Colombia son:

- **Mampostería en adobe o tapia pisada:** Es un sistema conformado por muros construidos en adobe o en tapia pisada, los principales componentes de dichas construcciones son arena y arcilla, en el cual las cargas verticales y horizontales son resistidas por los muros. Este sistema se presenta en las edificaciones más antiguas de la ciudad (ver figura 6).

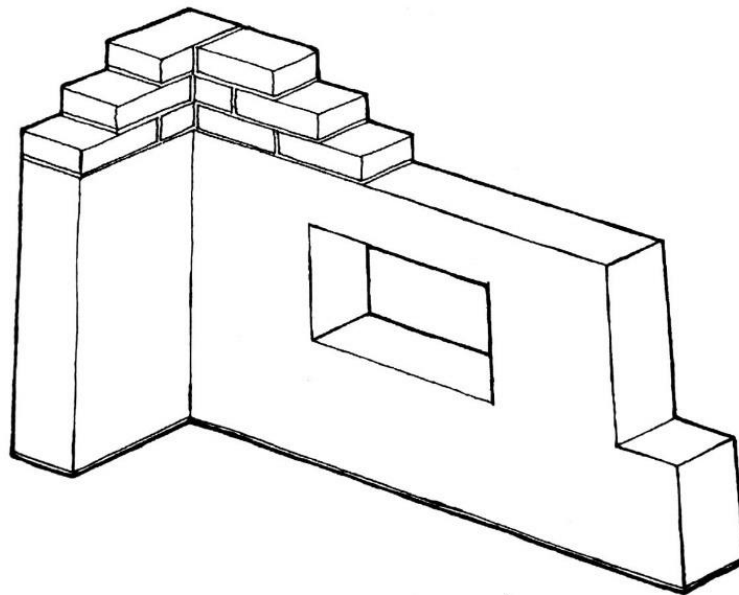


Figura 6. Esquema estructural de muros de adobe o tapia pisada. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

• **Mampostería Confinada:** Este sistema está conformado por muros construidos con ladrillos, pegados con cemento y adecuadamente confinados por vigas y columnas, de tal manera que el comportamiento estructural mejora al contar con elementos que le ayudan principalmente al soporte de las fuerzas horizontales ocasionadas por un sismo, en la figura 7 se ilustra esta tipología.

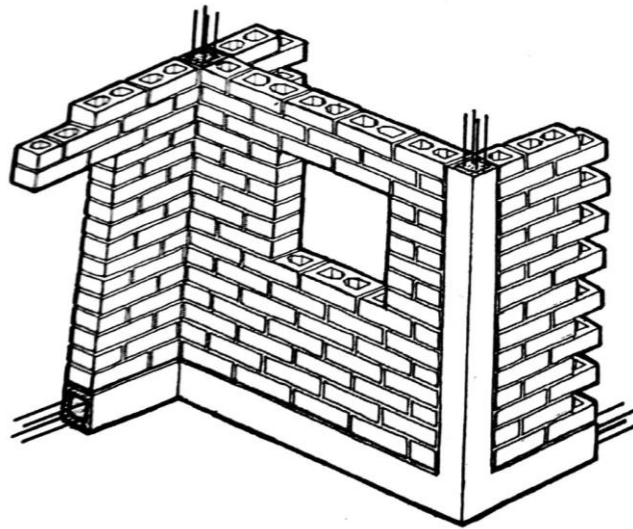


Figura 7. Esquema estructural de la mampostería en ladrillo confinado. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

• **Mampostería no confinada:** En esta clasificación se encuentran todas las construcciones elaboradas con mampostería en los diferentes tipos de ladrillo que se encuentran en la ciudad. Dichos muros no cuentan con un adecuado confinamiento o, en algunos casos, no poseen elementos de soporte y apoyo lo cual hace que las cargas verticales y horizontales sean soportadas directamente por los muros de mampostería (ver figura 8).

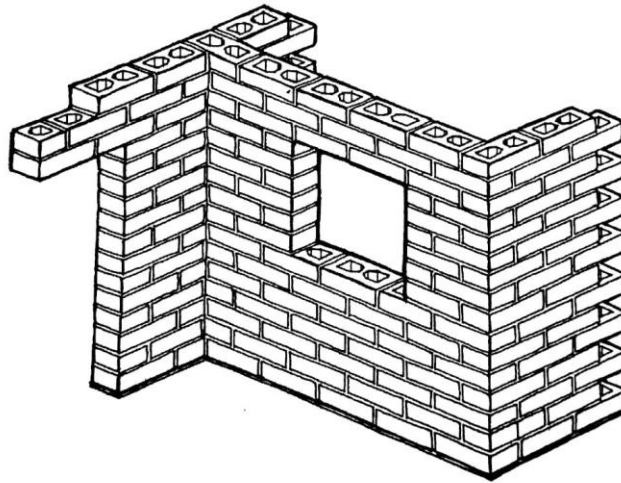


Figura 8. Esquema estructural de la mampostería no confinada. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

• **Mampostería reforzada:** En este sistema las edificaciones están reforzadas mediante la colocación de barras de acero en el interior de las celdas en las líneas de unión de las unidades. En algunos casos puede estar acompañado de elementos que confinan los muros de mampostería, los cuales mejoran la respuesta estructural (ver figura 9).

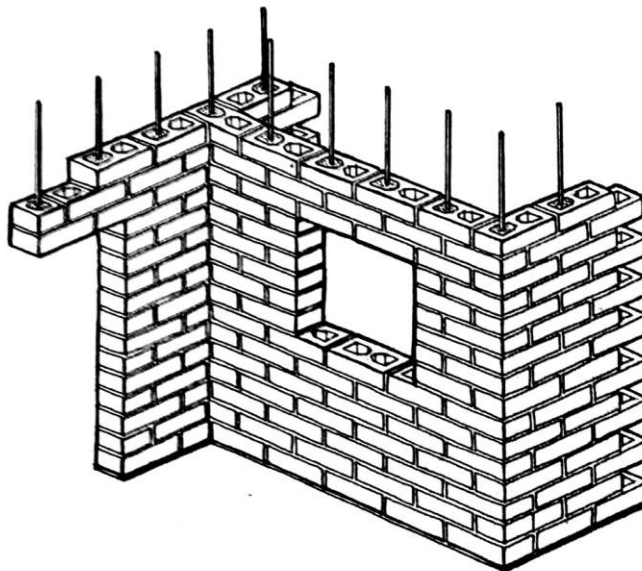


Figura 9. Esquema estructural de la mampostería reforzada. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

• **Pórtico:** Es una tipología compuesta por un arreglo espacial entre columnas y vigas, resistente a momentos, el cual soporta todas las fuerzas verticales y horizontales. Este sistema puede tener un comportamiento dúctil o no dúctil en función del reforzamiento, confinamiento y configuración estructural de sus elementos (ver figura 10).

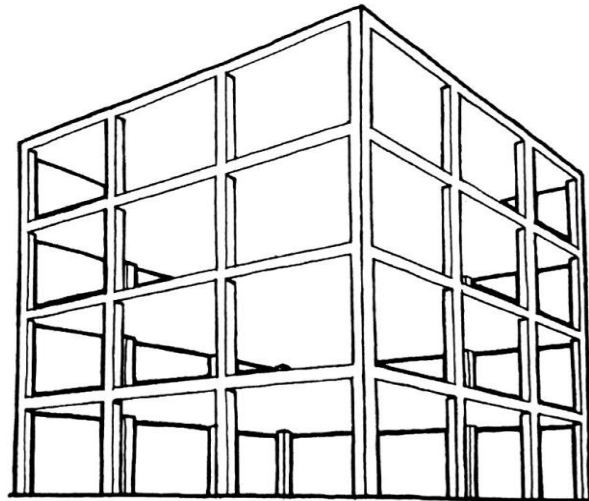


Figura 10. Esquema estructural en pórtico. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

• **Pórtico dúctil:** Se consideran dúctiles los pórticos cuya construcción es posterior a 1984, por haber sido diseñados y construidos bajo los requerimientos sismo-resistentes plasmados en el Código Colombiano de Construcciones Sismo Resistentes publicado en 1984 y la Norma Sismo Resistentes en 1998.

• **Pórtico no dúctil:** Se consideran no dúctiles la mayoría de pórticos construidos antes de 1983, época en que las estructuras no contaban con requerimientos para resistir las fuerzas sísmicas.

• **Muro Estructural:** Es un sistema estructural que no dispone de un pórtico esencialmente completo y, en el cual, las cargas verticales son resistidas por los muros de carga y las fuerzas horizontales son soportadas por muros estructurales o pórticos con diagonales (ver figura 11).

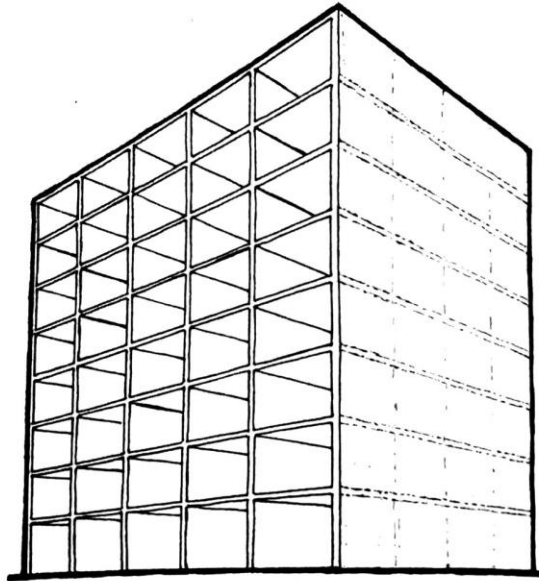


Figura 11. Esquema estructural de muros en concreto. Adaptado de Maldonado Rondón y Chío Cho (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga.

5.3 Niveles

Según datos históricos, el número de pisos de una estructura también es considerado como un parámetro que controla el daño estructural. Éste es un problema importante en los países en desarrollo, las principales razones del daño de los edificios de mediana y gran altura son la mala calidad de la construcción, la deficiencia en la aplicación de las normas sismo resistentes, la existencia de piso suave, los cambios en la resistencia del edificio y los fuertes efectos de edificios adyacentes (Şengezer, Ansal y Bilen, 2008).

Se considera el número mayor de pisos del albergue, es decir, si un albergue cuenta con 2 torres dentro de sus instalaciones, una con 3 pisos y la otra de 2 pisos, al albergue se le asignan 3 pisos.

5.4 Distancia a fallas

Una falla geológica es una grieta en la corteza terrestre en donde existe un desplazamiento de los bloques. La distancia a falla es la Mínima distancia euclidiana entre el albergue y cada una de las fallas que afectan a Bucaramanga y toda su área metropolitana: Chimita, Rio Suratá y Bucaramanga (Martínez 2007). La alta actividad sísmica en el departamento de Santander en términos generales, es superficial a intermedia. La región nororiental se ve especialmente influenciada por la presencia de los sistemas de fallas geológicas de Bucaramanga, Chimita y Suratá, y más al occidente por el sistema de fallas de la Salina, límite oriental del valle del Magdalena (Hermelín, 2007, pág 94). La distancia a falla se considerada un factor no estructural que afecta la vulnerabilidad de las edificaciones, siendo más vulnerable aquellas estructuras de mayor aproximación a la falla.

Tabla 5.

Caracterización de las fallas de Bucaramanga.

Falla	Tasa de movimientos	Longitud	Trayecto	Zonas afectadas
Bucaramanga – Morrorrico	10 mm/año	220 Km en dirección Norte-Sur	Recorre el costado del municipio, aproximadamente desde el caserío Chitota hasta Morrorrico1 tiene un rumbo de N15° W y se extiende desde la región de Curos a Umpalá, donde se producen algunas atenuaciones de energía sísmica	Pan de Azúcar Lagos del Cacique Cerro de Morrorrico Zona dos o corregimiento dos de Bolarqui.

				originada en epicentros alrededor de la región de Umpalá. Continúa su recorrido por Río Negro y San Alberto para unirse con la falla de Santa Marta.	
Suárez-Chimitá	0.01 a 0.1 mm/año	170 Km en dirección Sur de NNW y al Norte N-NE con inclinación SW y NW respectivamente en ángulos entre 45 y 70 grados.	Recorre el límite occidental del municipio, siguiendo el trazo del Río de Oro, para luego chocar contra la falla de Bucaramanga. A través de ella se producen algunas atenuaciones de energía sísmica originada en epicentros alrededor de la región de Vélez.	Meseta de los Santos de Ruitoque Terraza de Bucaramanga	
Chitota – Río Suratá	Dirección Noreste	Se extiende hacia el Noreste en el costado Norte de la ciudad, aproximadamente paralela al Río Suratá	Se extiende hacia el Noreste en el costado Norte de la ciudad, aproximadamente paralela al Río Suratá	Escarpa Norte de la Meseta de Bucaramanga	

Nota: Adaptado de unidad nacional para la gestión de riesgo de desastres – Colombia (2013). Plan municipal de gestión del riesgo de desastres de Bucaramanga. (Página 10)

5.5 Zonas en riesgo de remoción de masa

La Cruz Roja Colombiana define la remoción de masas como un fenómeno de la naturaleza que consiste en el movimiento pendiente abajo, lento o súbito de una ladera, formado por

materiales naturales como roca, suelo, vegetación o bien de rellenos artificiales. Los deslizamientos o derrumbes se presentan sobre todo en la época lluviosa o durante períodos de actividad sísmica (Cruz Roja, 2003).

Los procesos de erosión hídrica de remoción en masa y de erosión fluvial intensa se convirtieron en la principal amenaza exógena de la ciudad desde la década de 1950 debido al crecimiento poblacional. La remoción en masa en la tipología de deslizamientos, ha ocurrido al nororiente de la ciudad (siguiendo la vía hacia Cúcuta) en los asentamientos y barrios de Buenos Aires, Miraflores, el Diviso, Vegas de Morrorrico, la Malaña, la parte alta del barrio Albania y la vereda de Santa Bárbara. En la escarpa del Norte se ha presentado en los últimos años un deslizamiento rotacional a semitraslacional, de gran magnitud y velocidad extremadamente lenta, que afecta los barrios de La Esperanza, Lizcano, El Mirador, José María Córdoba, Villa Rosa y Villa Helena. Al Oriente urbano en los barrios La Trinidad, Los Guanábanos, Santa Lucía y Miramanga, en el sector de la quebrada Suratoque y entre los barrios La Cumbre y la Urbanización Bucarica. Otros sectores están sobre la vía que comunica a la urbanización San Bernardo con el barrio El Reposo, en la microcuenca de la quebrada Zapamanga, taludes y laderas que bordean a lagos del Cacique, antiguos Polvorines, y en los afluentes de la quebrada El Dínamo, sur del lago Miramanga. (Hermelín, 2007, pág. 106).

5.6 Llenos antrópicos y mecánicos

Se entiende por llenos antrópicos, todos aquellos materiales que han sido depositados por el hombre, también llamados rellenos o llenos, están integrados por: basuras y/o escombros y/o suelos removidos o por la mezcla de éstos. Dichos materiales son depositados en laderas y

depressiones naturales por las que transitan corrientes intermitentes o permanentes de agua. De acuerdo al método utilizado para depositar los materiales que conforman el lleno, estos se clasifican en: Llenos hidráulicos, llenos mecánicos, botaderos y rellenos sanitarios.

Los llenos mecánicos corresponden a materiales de cortes de ladera y de vías que han sido utilizados para llenar depressiones y cauces. Están compuestos primordialmente por terrones de ceniza volcánica embebidos en una matriz de ceniza volcánica remoldeada, mezclada con suelos orgánicos y materiales de construcción. Se distinguen por su homogeneidad, textura y color (Narváez, Pérez y Lavell, 2009).

5.7 Vulnerabilidad

La terminología de la Estrategia Internacional para la Reducción de Desastres de las Naciones Unidas (UNISDR, 2009) define los siguientes conceptos:

- **Amenaza:** Un fenómeno, sustancia, actividad humana o condición peligrosa que pueden ocasionar la muerte, lesiones u otros impactos a la salud, al igual que daños a la propiedad, la pérdida de medios de sustento y de servicios, trastornos sociales y económicos, o daños ambientales.

- **Desastre:** Una seria interrupción en el funcionamiento de una comunidad o sociedad que ocasiona una gran cantidad de muertes al igual que pérdidas e impactos materiales, económicos y ambientales que exceden la capacidad de la comunidad o la sociedad afectada para hacer frente a la situación mediante el uso de sus propios recursos.

- **Riesgo de desastres:** Las posibles pérdidas que ocasionaría un desastre en términos de vidas, las condiciones de salud, los medios de sustento, los bienes y los servicios, y que podrían ocurrir en una comunidad o sociedad particular en un período específico de tiempo en el futuro.

- **Susceptibilidad:** es el grado de fragilidad interna de un sujeto, objeto o sistema para enfrentar una amenaza y recibir un posible impacto debido a la ocurrencia de un evento adverso.

La oficina de las naciones unidas para la reducción del riesgo de desastres UNISDR, define vulnerabilidad como la incapacidad de resistencia cuando se presenta un fenómeno amenazante, o la incapacidad para reponerse después de que ha ocurrido un desastre. La vulnerabilidad sísmica de una estructura se define como la predisposición intrínseca a sufrir daño ante la ocurrencia de un movimiento sísmico y está asociada directamente con sus características físicas y estructurales de diseño. Díaz y León (2013) afirman que en muchos estudios realizados se ha confirmado que uno de los factores de incidencia en el número de afectados es la resistencia de los edificios, especialmente en los países en desarrollo se presenta alta construcción de viviendas en zonas riesgosas y muy vulnerables, además de que están vacías de técnicas ingenieriles y materiales de mala calidad. Un diseño adecuado de las edificaciones disminuye las probabilidades de que estas sufran grandes daños en caso de presentarse un sismo, pero no sólo la intervención de ingenieros es suficiente para que los daños sean insignificantes, la ubicación del edificio y las condiciones del entorno son variables que también deben ponerse en estudio.

Según cifras proporcionadas por el DANE, en el año 1995 la población colombiana estaba expuesta al riesgo sísmico de la siguiente manera: el 35% en riesgo alto, el 51% en riesgo

intermedio y 14% restante en riesgo bajo. Es decir que gran parte de los colombianos se encuentran expuestos a riesgos sísmicos; en especial, ciudades consideradas muy vulnerables, como lo es Bucaramanga, se hace relevante el hecho de anticiparse y controlar los daños que pueda generar un evento catastrófico. Los elementos expuestos y el volumen de dichos elementos son las únicas variables que pueden ser modificadas por el hombre, los programas preventivos modifican el grado de vulnerabilidad una vez desarrollado un modelo de pérdidas que califique su nivel de exposición, y una vez modificadas se puede afectar o reducir el nivel de riesgo existente

La percepción del riesgo es diferente en todas las personas y estandarizar una escala se hace necesario para una comprensión clara de los resultados. En la Tabla 6 se muestra una descripción de los niveles de vulnerabilidad asociados a los albergues; descritos a partir del juicio del experto, quien realiza la etiqueta de los datos y estima los daños esperados para un sismo de diseño, se clasifica la vulnerabilidad de los albergues en 5 clases correspondientes a las variables de Salida de los modelos. Considérese 5 el nivel más alto de vulnerabilidad y 1 el nivel más bajo de daño esperado.

Tabla 6.
Niveles de vulnerabilidad.

Clase	Nivel de vulnerabilidad	Descripción
5	Muy Alto	La edificación presenta caída de sus elementos estructurales, muros y vidrios rotos, aunque no en un 100% pero su funcionalidad es nula.
4	Alto	La edificación presenta daños graves en los muros y/o caídas de algunos de ellos. Se requiere de un mediano plazo para su recuperación.
3	Medio	La edificación presenta daños en algunos de los muros, fisuras en las vigas y aun puede seguir en uso.
2	Bajo	La edificación presenta daños marginales como pequeñas grietas en los muros.
1	Muy Bajo	La edificación no presenta daños.

5.8 Albergue

Según el Proyecto Esfera (2011) la reducción del riesgo de desastres: incluye el concepto y la práctica de reducir el riesgo de desastres mediante un esfuerzo sistemático de análisis y de gestión de los factores que causan los desastres, especialmente limitando la exposición a las amenazas y paliando la vulnerabilidad de las personas, dando muestras de sensatez en el manejo de la tierra y el entorno natural, y mejorando la preparación para enfrentar fenómenos peligrosos; entre ellos desastres naturales como tormentas, inundaciones, sequías y aumento del nivel del mar. Dichos fenómenos parecen volverse cada vez más variables y graves y se le atribuye cada vez más al cambio climático mundial.

Por consiguiente, la fase Post-Desastres es trascendental si lo que se quiere es reducir el impacto del desastre. Este trabajo de investigación tiene por objeto la identificación de potenciales albergues en Bucaramanga, mediante una evaluación de vulnerabilidad que permita determinar las infraestructuras aptas para albergar las poblaciones más afectadas y asimismo servir como una herramienta de apoyo para las entidades gubernamentales en el área de gestión de desastres de Bucaramanga.

La Federación Internacional de la Cruz Roja y la Media Luna Roja define Albergue como: “Lugar físico creado e identificado como un lugar seguro, que cuenta con todos los medios necesarios para hospedar por un periodo corto, mediano y largo plazo a un grupo de personas afectadas por los resultados del impacto de una amenaza, con las garantías esenciales para garantizar la dignidad humana, conservando la unidad familiar y la cultura de las personas afectadas así como su estabilidad física y psicológica. Promoviendo la organización comunitaria” (Esfera, 2011, pag.17).

5.8.1 Albergue temporal: Es el lugar donde se proporciona temporalmente techo, alimentación, vestido y salud a personas vulnerables, antes, durante y/o después de la ocurrencia de un fenómeno destructivo (Esfera, 2011, pag.17). La tipología de albergues temporales se encuentra afectado por la complejidad de la situación de la emergencia o el desastre, como también por la capacidad local, regional o nacional, por tal razón es importante considerar las instalaciones o los espacios en los cuales se pueda ubicar a la población afectada, según las condiciones y el número de familias o personas que requieran del albergue. En la Figura 12 se discrimina la tipología de los albergues temporales.

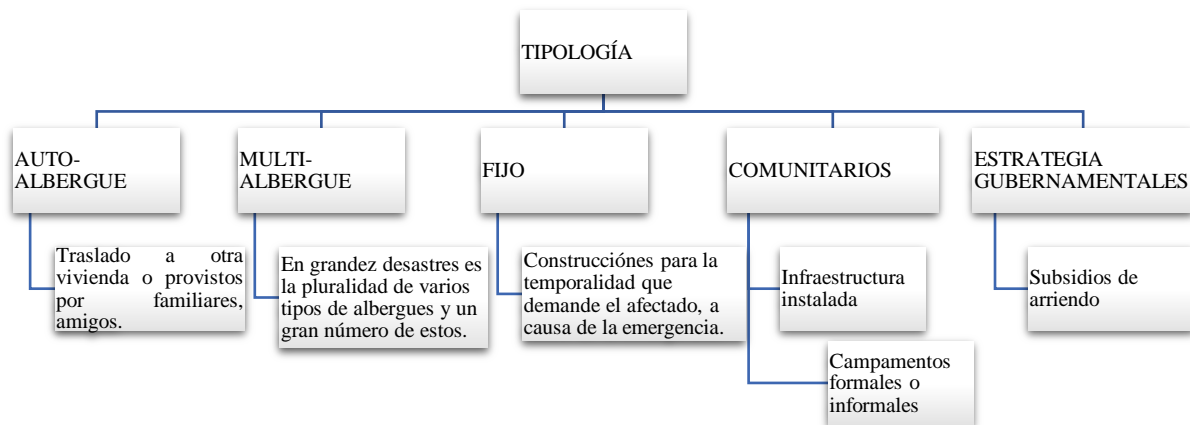


Figura 12. Tipología de albergues temporales. Adaptado del Proyecto Esfera (2011).

<http://www.acnur.org/t3/fileadmin/Documentos/Publicaciones/2011/8206.pdf?view=1>

5.8.2 Albergues comunitarios: Como se vio en la Figura 12, el tipo de albergue comunitario hace referencia a una infraestructura instalada. Este tipo de albergue se ajusta al prototipo de estructuras empleadas en el desarrollo de esta investigación como son aquellos escenarios deportivos, escuelas, colegios, salones comunales, entre otros. Estos lugares que se utilizan de manera transitoria, para el alojamiento de personas o comunidades que se ven afectadas por los efectos de un desastre. Sin embargo, en muchas regiones su uso se ve restringido, en especial escuelas y colegios, ya que al usar estos espacios como albergues se limita la posibilidad de retomar el objeto para el que fueron construidos. Generalmente este tipo de espacios cuentan con los servicios básicos, áreas comunes amplias, y se facilita el acceso de salud y la administración general del albergue (Esfera, 2011, pag.19).

5.8.3 Metodología de selección de albergues temporales: La identificación de posibles albergues temporales para el caso de estudio se efectúa a partir de una revisión de trabajos de investigación realizados en la Escuela de Estudios Industriales y Empresariales de la

Universidad Industrial de Santander. Estas investigaciones entre sus objetivos abarcan la localización de instalaciones potenciales a albergues temporales en la ciudad de Bucaramanga. En la Tabla 7 se muestran los trabajos de investigación tomados como referencia y sus respectivas metodologías aplicadas para la selección de albergues.

Tabla 7.

Metodología de selección de albergues temporales por autor.

Autores	Titulo	Método de elección de albergues	Tipos de albergues
Barrera y Hernández (2016)	“Un algoritmo evolutivo para el problema de distribución de recursos Post-Desastres sísmicos en la ciudad de Bucaramanga”	Se realiza una búsqueda exploratoria por medio de la herramienta de georreferenciación (Googlemaps).	<ul style="list-style-type: none"> • Canchas • Parques • Colegios • Centro de distribución
Contreras y Martínez (2016)	“Un problema de localización multi-objetivo en la fase de preparación para desastres por sismos en la ciudad de Bucaramanga”	Se basan en “El manual de albergues temporales de la Cruz Roja colombiana” en el cual se definen las condiciones necesarias para designarse un albergue: fácil acceso, debe contar con servicios básicos, servicios sanitarios y espacio suficiente (3.5m ² por persona) para albergar un alto número de personas.	<ul style="list-style-type: none"> • Colegios • Parques • Zonas deportivas
Aguilar (2016)	Un modelo multi-objetivo de localización-Ruteo para la planeación logística en la fase de	La localización de albergues se realiza por medio de una geo-localización de los albergues propuestos sobre el mapa de	<ul style="list-style-type: none"> • Instituciones de educación de carácter publico • Zonas verdes

	preparación a sismos en Bucaramanga”	Bucaramanga y se excluyen las instalaciones ubicadas en las zonas más vulnerables según Maldonado y Chío (2008) en su libro “Vulnerabilidad sísmica en centros urbanos.”	<ul style="list-style-type: none"> • Zonas deportivas
Barreto y Niño (2016)	“Un algoritmo memético para el problema de localización ruteo con ventanas de tiempo para la atención de desastres sísmicos en la ciudad de Bucaramanga”	La solución del algoritmo memético con ventanas de tiempo hace referencia a los posibles albergues que deben ser considerados en la ciudad de Bucaramanga	<ul style="list-style-type: none"> • Colegios • Parques • Polideportivos
Morales y Palacios (2017)	“Localización de instalaciones y asignación de recursos mediante un algoritmo heurístico para la atención Post-Desastres, en caso de inundación o remoción de masa”	Se hace uso de la metodología de ponderación de factores para comparar y elegir las mejores alternativas a albergues temporales. Los criterios considerados fueron: Accesibilidad, seguridad, criterios no estructurales, topografía, criterios geológicos, hidrológicos y climáticos.	<ul style="list-style-type: none"> • Colegios • Parques • Canchas

Partiendo de la información recolectada en las investigaciones mostradas en la Tabla 7, se realiza una consolidación de la información obteniendo en total 132 albergues para Bucaramanga (Ver en Apéndice B) en los que se incluye: Instituciones educativas, Zonas verdes y Zonas deportivas. Sin embargo, para el desarrollo y ejecución de este proyecto de

investigación se hace necesario el levantamiento de la información correspondiente a diferentes factores asociados a la vulnerabilidad estructural de una edificación, logrando la obtención de este tipo de información para un total de 68 albergues.

Posteriormente se georreferencia esta información a través de Google Maps. En el siguiente link se puede visualizar los albergues temporales determinados para este trabajo de investigación:

<https://www.google.com.co/maps/@7.1226753,73.1724992,12z/data=!3m1!4b1!4m2!6m1!1s1fHgVodXexoyeaE2KIQTD3Nqiwmg?hl=es-419>

El conjunto de datos es clasificado en 3 tipos de albergues temporales. En la Figura 13 se muestra el símbolo empleado para cada tipo de albergue.




<i>Tipo de Albergue</i>	<i>Representación</i>
<i>Instituciones Educativas</i>	
<i>Zonas verdes</i>	
<i>Zonas deportivas</i>	

Figura 13. Representación de los albergues temporales.

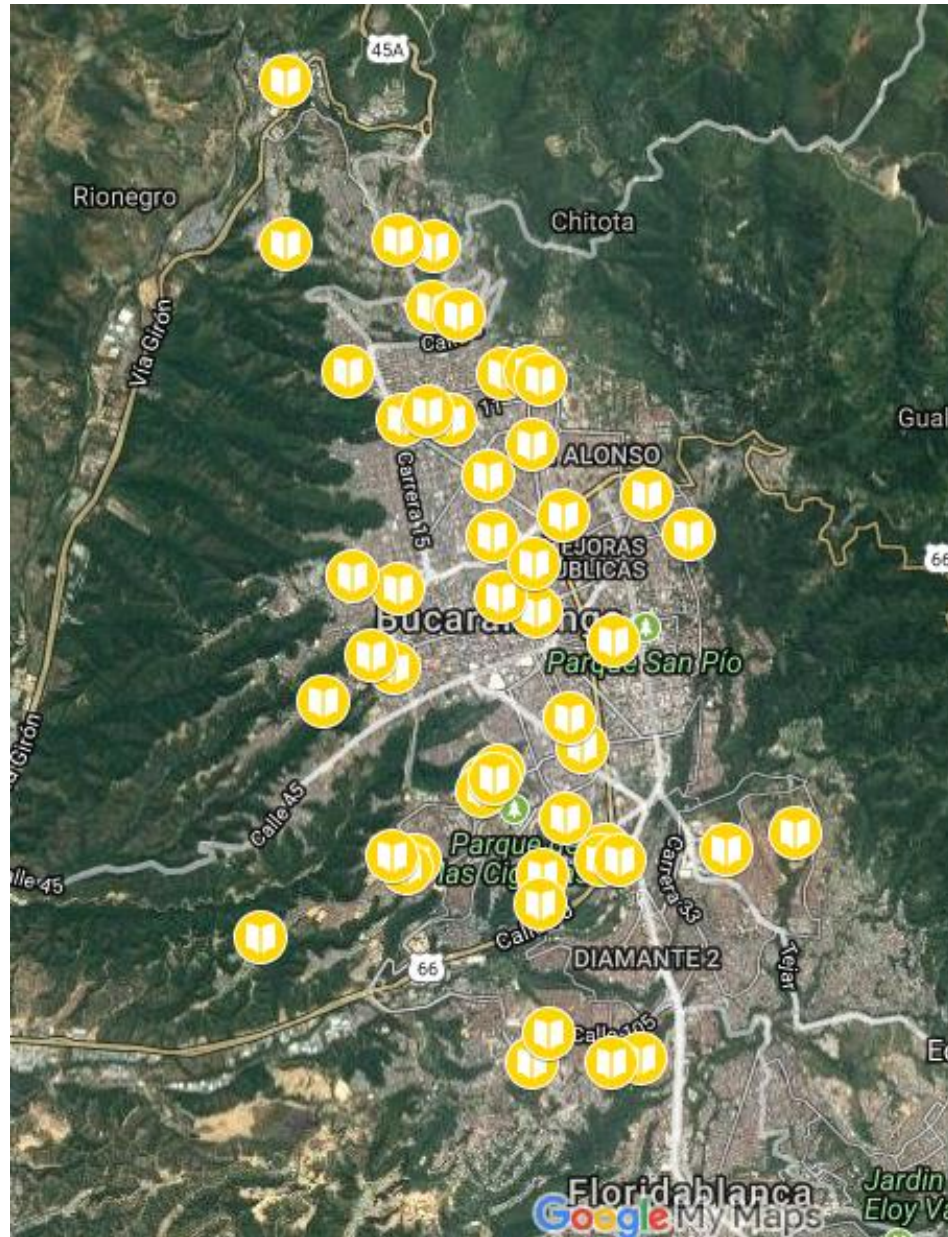


Figura 14. Localización de Instituciones Educativas

En la Figura 14 se ilustra la ubicación geo-referencial de los albergues temporales de tipo educativo identificados con el símbolo 🏫, en los que se incluyen: universidades, colegios públicos y privados de Bucaramanga. En el Apéndice C se amplía esta información.

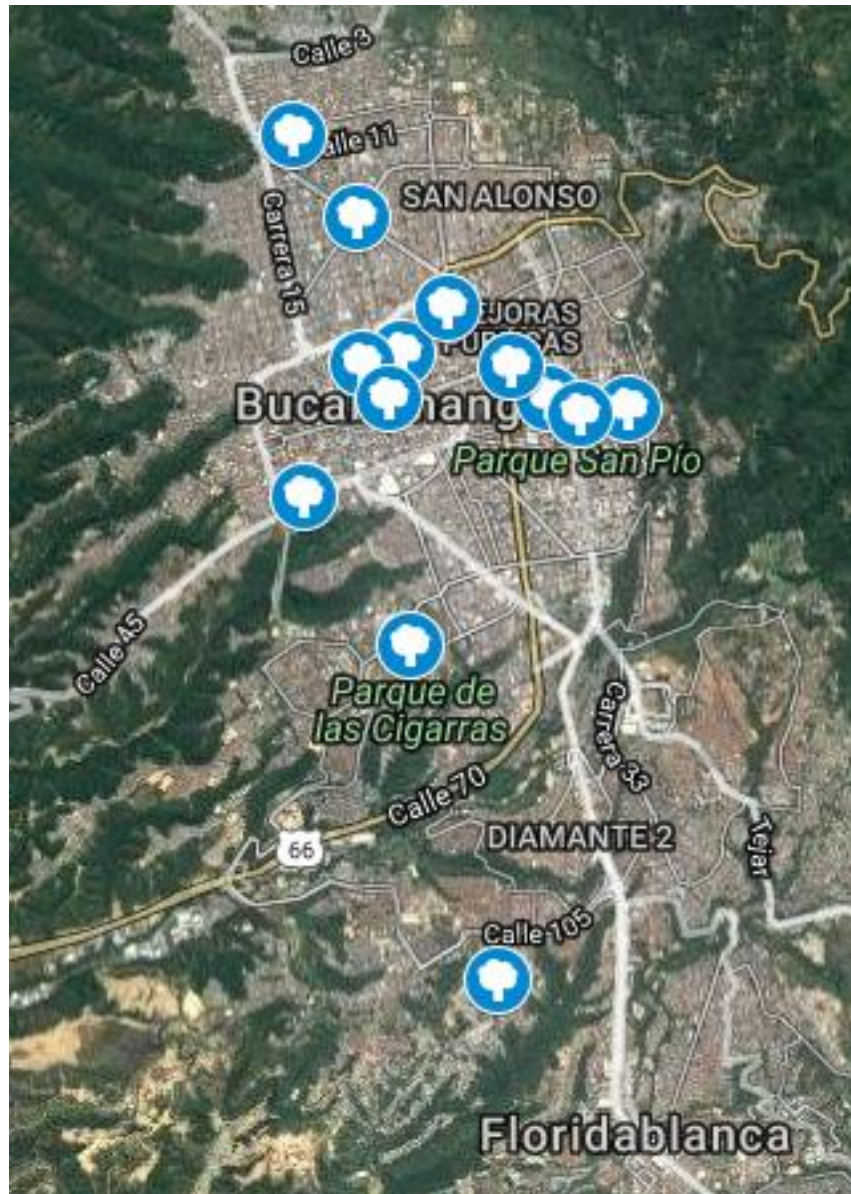



Figura 15. Localización de Zonas verdes

La Figura 15 ilustra los albergues temporales ubicados en zonas verdes, estos son identificados con el símbolo . Como se menciona en el proyecto Esfera (2011), estas localidades son consideradas los más seguros dados su carencia de tipología estructural. En el Apéndice D se amplía esta información.

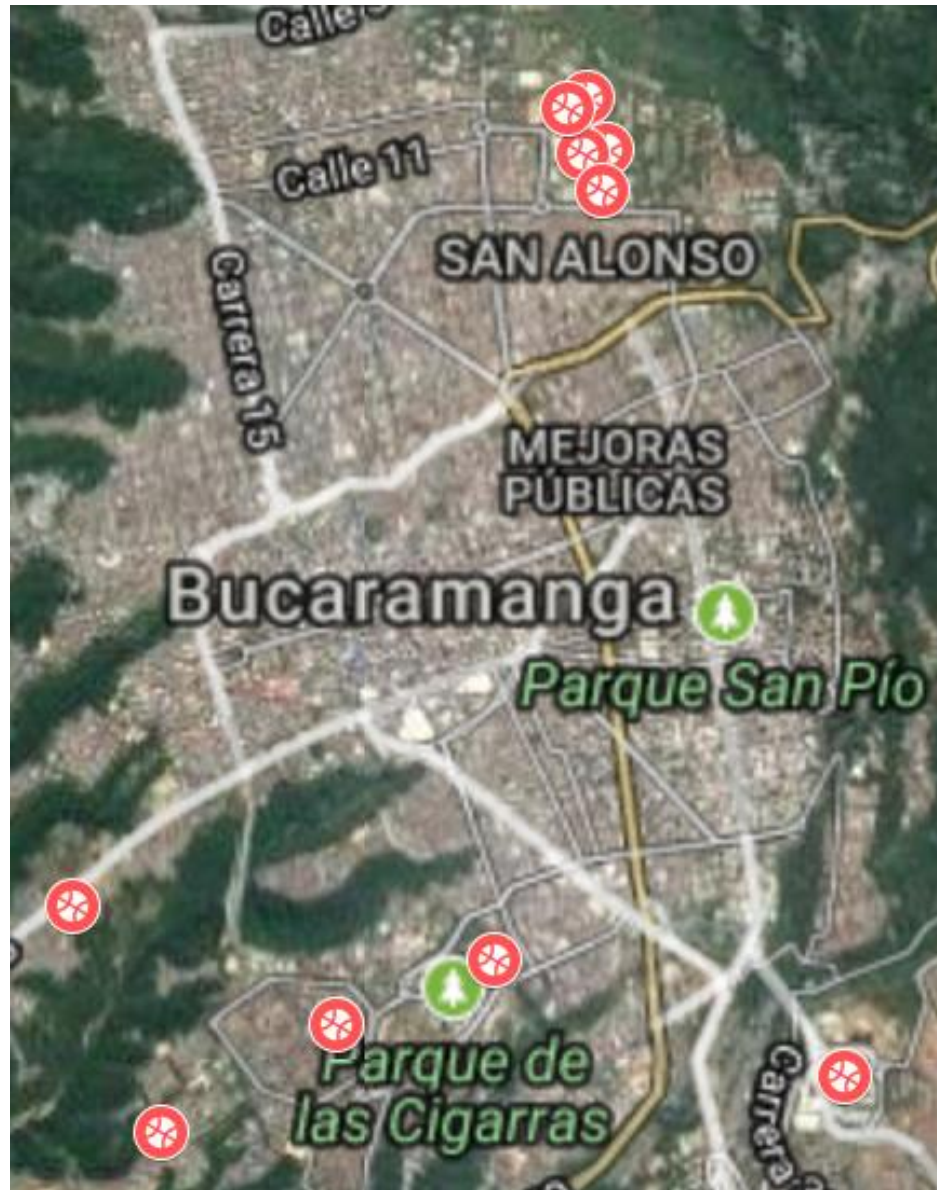



Figura 16. Localización de Zonas deportivas

La Figura 16 se ilustra los albergues de tipo deportivo en los cuales se incluye polideportivos, canchas y estadios, estos son identificados con el símbolo . En el Apéndice E se amplía esta información.

6. Caso de Estudio

Bucaramanga es una ciudad altamente afectada por movimientos telúricos, por esto se considera que debería contar con fuertes estrategias y planes de acción en la prevención de desastres. Por el contrario, se evidenció que carece de planificación, organización y falta de atención por parte del personal encargado. Los albergues identificados en la ciudad se asocian a un nivel de vulnerabilidad “alto” según el criterio del experto consultado; es por eso que este proyecto presenta una solución basada en técnicas de aprendizaje automático, que permita intervenir de forma correctiva las estructuras de los albergues a través de acciones orientadas a disminuir o reducir los niveles de vulnerabilidad sísmica y las condiciones de amenaza cuando sea posible.

El caso de estudio se realiza para los albergues temporales ubicados en la ciudad de Bucaramanga, pero contribuye a una iniciativa para otras ciudades. Empleando las técnicas de máquinas de soporte vectorial y árboles de decisión, se realizó un entrenamiento del modelo utilizando el software RStudio, para posteriormente probar el modelo con el conjunto de datos de prueba verificando la eficiencia en los resultados de las clasificaciones. Bajo el criterio del porcentaje de mala clasificación, se determinó cuál de los dos algoritmos tiene mayor precisión al momento de replicar el conocimiento del experto.

6.1 Análisis de Datos

El nivel de vulnerabilidad es afectado por 8 variables divididas en factores estructurales y factores no estructurales como se muestra en la Tabla 8. Los factores estructurales son aquellas variables que influyen directamente en las condiciones físicas de la edificación, los factores no estructurales están relacionados con el entorno en el que se encuentra ubicada la estructura.

Tabla 8.
Factores asociados a la vulnerabilidad estructural

Tipo	Factores
Factores Estructurales	<ul style="list-style-type: none"> • Año de Construcción • Tipología • Niveles
Factores no Estructurales	<ul style="list-style-type: none"> • Llenos antrópicos y mecánicos • Remoción de masa • Falla Bucaramanga • Falla Chimitá • Falla Suratá

Mampostería, mampostería confinada, pórtico dúctil y pórtico no dúctil son las tipologías con que fueron construidos los albergues, predominando aquellos construidos en pórtico no dúctil. En la Figura 17 se observa que la mampostería y mampostería confinada representan la proporción de datos más pequeña. Estos albergues se caracterizan por ser de tipo educativo y tener altos niveles de vulnerabilidad debido a la carencia de elementos estructurales adecuados, y una distribución uniforme de los muros que puedan resistir los movimientos generados.

Como se mencionó en la sección 5, el pórtico no dúctil es característico de las instituciones educativas, que son calificadas como el tipo de albergue más vulnerable a efectos de un desastre.

El 90.47% de los albergues construidos en pórtico no dúctil son etiquetados en niveles de vulnerabilidad “alto” y “muy alto”, esto se debe a que ésta tipología es caracterizada por su año de construcción, es decir, aquellas estructuras que fueron construidas antes de 1983 y no se ajustaron a la normatividad de construcciones sismo resistentes del 98. Contrario al pórtico no dúctil, el pórtico dúctil es característico de albergues educativos construidos después de 1983, bajo los lineamientos de la normatividad, por lo que el 81.81% de los albergues construidos con esta tipología se etiquetan en niveles de vulnerabilidad “bajo” y “muy bajo”, es decir que están en capacidad de responder favorablemente a la ocurrencia de un sismo.

Finalmente, se observa que aquellos albergues que carecen de tipología (Ninguna), el 66.67% fueron etiquetados con un nivel de vulnerabilidad “alto”. Según el proyecto Esfera, los parques y polideportivos se consideran instalaciones seguras para alojar la población más vulnerable, debido a que son espacios abiertos que no poseen ningún tipo de estructura, la diferencia de los resultados obtenidos en la Figura 17 con lo mencionado por el proyecto esfera, se debe a que la evaluación de vulnerabilidad sísmica se hizo en función de ocho variables (Tabla 8), si por el contrario se hiciera en función únicamente de la variable tipología, muy seguramente el nivel de vulnerabilidad asociado sería “muy bajo”, pues no se presenta ningún tipo de riesgo.

Niveles de Vulnerabilidad por Tipología

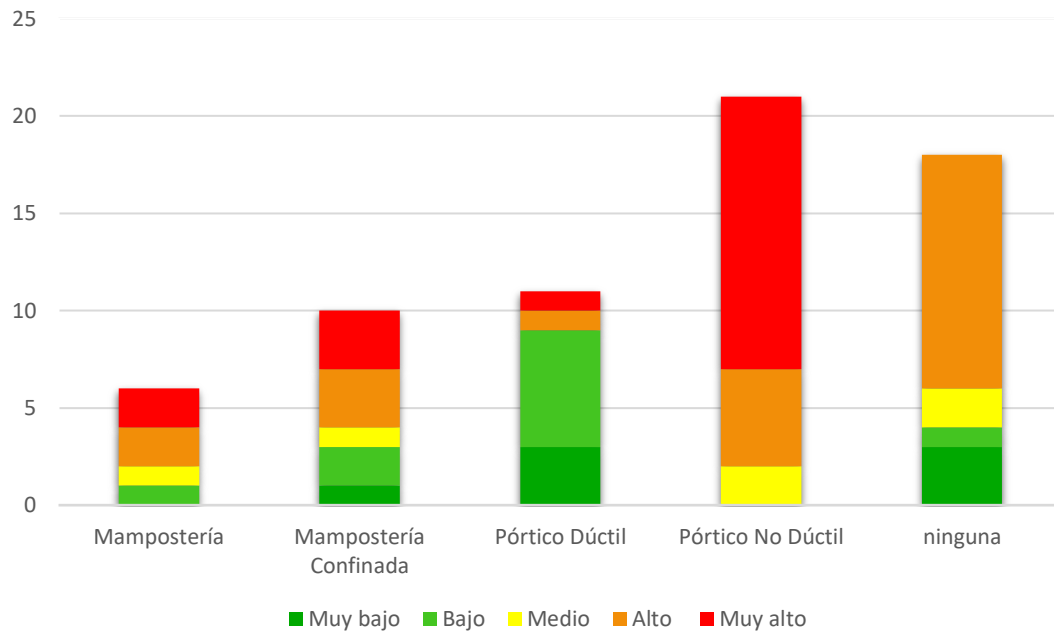


Figura 17. Análisis de los niveles de vulnerabilidad en cada Tipología.

El aumento de zonas vulnerables y personas expuestas a fenómenos peligrosos en Colombia, se debe principalmente al crecimiento de la vivienda informal, incumplimiento de estándares de diseño y construcción, a su localización en zonas de peligro, a la posible presencia de fenómenos naturales, y a la falta de estrategias para intervenir las construcciones anteriores a la primera norma de sismo resistencia (1984) (Banco Mundial, 2012, pag.327).

Bajo este concepto el año de construcción de los albergues temporales influye directamente en si este cuenta o no con los requerimientos mínimos exigidos por la NRS-98, en la cual se establece parámetros de diseño para 500 años de período de retorno, en las que se considera que un evento puede tener el 10% de probabilidad de ser excedido durante los 50 años de vida útil

de una edificación, de manera que no colapse y se asegure la protección de la vida, y que pueden existir ciertos niveles de daños aceptables (Banco Mundial, 2012, pag.291).

Como se ha venido mencionando, el año de construcción es un factor crítico en la determinación de la vulnerabilidad estructural ya que se toma como un punto de partida para evaluar la resistencia de la estructura. El 72% de los albergues fueron construidos antes de 1984 como se muestra en la figura 18, es decir que gran parte de los albergues presentan falencias en sus diseños de construcción y son más propensos a sufrir daños y/o colapsar dado que no se construyeron bajo los lineamientos de la normatividad sismo resistente.

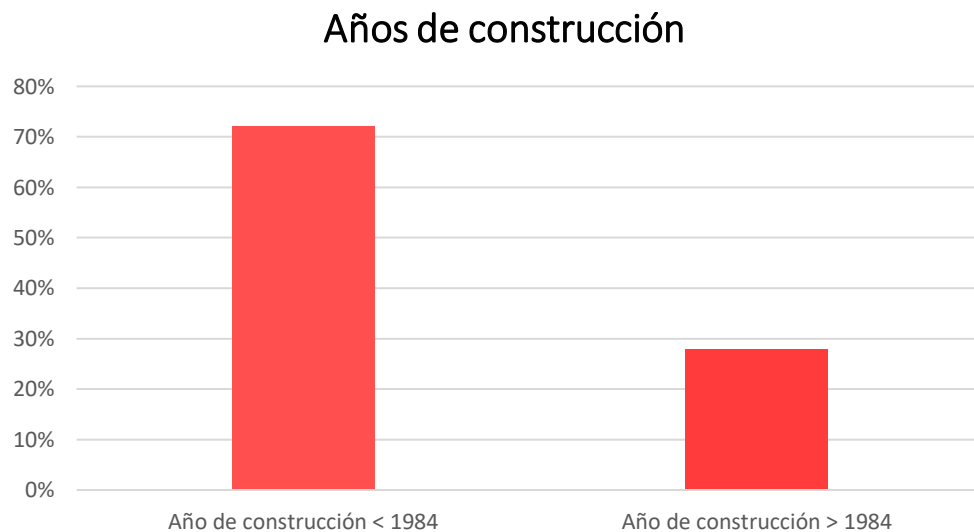


Figura 18. Análisis de datos por año de construcción.

Los análisis realizados anteriormente se hicieron de forma aislada, es decir, identificando la relación directa que existe entre cada variable y la vulnerabilidad que genera en la estructura. La figura 19 permite analizar la vulnerabilidad estructural integrando todas las variables

consideradas proporcionando datos más reales. El 66% de los albergues de Bucaramanga tienen un nivel “alto” y “muy alto” de vulnerabilidad sísmica, y solo un 25% son considerados como espacios en condiciones óptimas para prestar alojamiento temporal seguro a la comunidad (niveles muy bajo o bajo).

Análisis por Clases

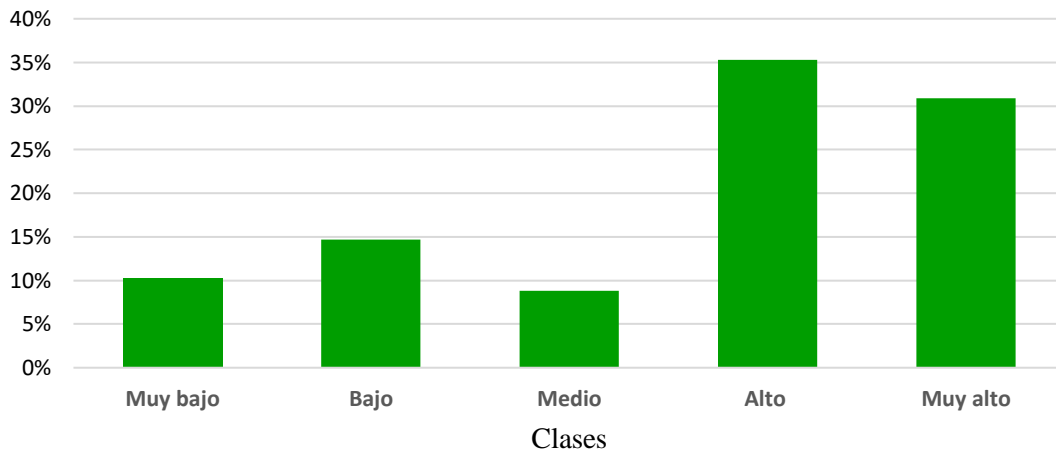


Figura 19. Análisis de datos por clases de vulnerabilidad

6.2 Ejecución de los algoritmos

Para la ejecución y desarrollo de este modelo se utilizó una base de datos que contiene información característica de 68 albergues temporales identificados en la ciudad de Bucaramanga. El conjunto de datos fue etiquetado por el criterio de un solo experto asociando el nivel de vulnerabilidad correspondiente en una escala de 1 a 5, donde 1 representa nivel muy bajo de vulnerabilidad y 5 un nivel muy alto.

Se definen dos tipos de modelos de aprendizaje automático para predecir la clase y el nivel de vulnerabilidad asociado a zonas catalogadas como albergues temporales: Máquinas de

soporte vectorial y árboles de decisión. El costo computacional de los algoritmos propuestos se encuentra afectado por el tipo de computador en el cual se ha ejecutado el código de programación, y el método empleado para determinar el valor de los parámetros, que para este caso fue la búsqueda exhaustiva y consiste en realizar todas las iteraciones posibles entre los rangos definidos para cada uno de los parámetros del modelo i , como resultado de la ejecución de las iteraciones se obtiene el valor óptimo de los parámetros.

Las características de los computadores empleados en la ejecución de los modelos son:

- Para las máquinas de soporte vectorial:

Procesador: Intel Core i-3, CPU 2.40 GHz

RAM: 6, 00 GB

Tipo de sistema operativo: 64 bits

- Para el árbol de decisión:

Procesador: Intel Core i7-5500U, CPU 2.40 GHz

RAM: 8, 00 GB

Tipo de sistema operativo: 64 bits

Se realizan 4 variaciones de las máquinas de soporte vectorial: SVM con kernel Lineal, SVM con kernel Polinomial, SVM con kernel radial y SVM con kernel Sigmoide.

Considere que:

$i \rightarrow 1: 5$

con: {*SvmLineal*, *SvmPolinomial*, *SvmRadial*, *SvmSigmaide*, *Arbol de decisión*}

En la Tabla 9 se describe los pasos que se deben seguir para la ejecución de los modelos:

Tabla 9.

Proceso de ejecución de los modelos de aprendizaje automático.

Paso	Descripción
1	Ingresar a la base de datos.
2	La variable tipología tiene la particularidad de ser una variable categórica, y para integrarla en el lenguaje de programación se transformó a una variable binaria, asignando 1 a la tipología que corresponde al albergue y 0 para las demás tipologías. Igualmente sucede con los llenos antrópicos y la remoción de masas; se asigna 1 si el albergue está construido en una zona de llenos antrópicos o de remoción de masas o de lo contrario se asigna 0.
3	Normalizar el conjunto de datos para evitar que el modelo asigne pesos diferentes a las variables, es decir, que tengan igual importancia al momento de elegir los criterios de selección.
4	Elegir el porcentaje datos para el entrenamiento del modelo i. De forma pseudoaleatoria se selecciona el conjunto de datos de entrenamiento.
5	El porcentaje restante serán los datos empleados para probar la validación del modelo. De forma pseudoaleatoria se selecciona el conjunto de datos de prueba.
6	Determinar los rangos de valores a evaluar en cada parámetro para obtener los valores óptimos. Esta técnica es conocida como búsqueda exhaustiva.
7	Entrenar el modelo i con los valores óptimos de los parámetros obtenidos en el paso anterior.
8	Validar del modelo i con el conjunto de datos de prueba.
9	Determinar la eficiencia del modelo i, bajo el criterio de porcentaje de mala clasificación.

6.3 Definición de los parámetros

Cuando se evalúa la calidad de un modelo, es importante medir el error en el conjunto de entrenamiento y en el conjunto de predicción (prueba). La utilización exclusiva del error del conjunto de entrenamiento puede conducir a resultados engañosos. Estos errores pueden conducir a un fenómeno de sobreajuste (overfitting), en el cual el modelo se ajusta muy bien a los datos existentes, pero tiene un pobre rendimiento para predecir nuevos resultados (Cárdenas, s.f, pag.1). Es decir, el modelo tiende a seguir el ruido de los datos de entrenamiento ajustándose de tal manera que produce un modelo que no es capaz de generalizar.

Para solucionar este problema es necesario reducir el tamaño de la red mientras se mantiene un buen rendimiento; la idea central es obtener una red con un tamaño óptimo, menos propensa a aprender el ruido en los datos de entrenamiento y a incurrir en el sobreajuste, y por ende, que generalice con mayor precisión en un tiempo computacional menor que una red de mayor tamaño. (Villa, Velásquez y Sánchez, 2015). La determinación de los parámetros óptimos mediante búsqueda exhaustiva, es el método o la medida que se usó para controlar el sobreajuste en los modelos.

6.3.1 Máquinas de soporte vectorial: La función Kernel es descrita como una medida de similitud que puede ser vista como el producto punto en un espacio llamado de características. El propósito de esto, es proporcionar un entendimiento intuitivo de dichas funciones en vez de presentar la teoría analítica de cuales clases de kernel admiten la representación de un producto punto en un espacio de características dado (Henao, 2004, pag.87).

Las ecuaciones del 1 al 4 muestran los parámetros que se requieren para cada tipo de kernel:

$$\text{Kernel Lineal} \rightarrow U' * V \quad (1)$$

$$\text{Kernel Polinomial} \rightarrow (\text{Gamma} * U' * V + \text{Coef0})^{\text{degree}} \quad (2)$$

$$\text{Kernel Radial} \rightarrow \text{Exp}(-\text{Gamma} * |U - V|^2) \quad (3)$$

$$\text{Kernel Sigmoide} \rightarrow \tanh(\text{Gamma} * U' * V + \text{Coef0}) \quad (4)$$

El parámetro Cost (V) se evaluó para valores comprendidos entre el rango de 0.1 a 4.6, Gamma de 0.1 a 5, Degree de 1 a 3 y Coef0 en 1 y 2; como se muestra en la tabla 10:

Tabla 10.
Rango de parámetros para las máquinas de soporte vectorial.

Parámetro	Rango
Cost	0.1 0.6 1.1 1.6 2.1 2.6 3.1 3.6 4.1 4.6
Gamma	0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
	1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2
	2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3
	3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 4
Degree	1 2 3
Coef0	1 2

Para controlar el sobreajuste de las máquinas de soporte vectorial se realiza una variación del Cost (v), necesario para los cuatro tipos de Kernel enunciados en las ecuaciones 1-4, este parámetro determina la penalización permitida en el margen de separación. La asignación de valores muy grandes a este parámetro puede llevar a provocar el sobreajuste del modelo y

valores de cost cercanos a 0 ocasionan una menor penalización a las observaciones permitiendo un mayor porcentaje de datos mal clasificados.

El gamma controla el comportamiento del Kernel (Polinomial, Radial, Sigmoide) en cuanto se aumenta su valor la función kernel se hace más flexible.

Emplear valores superiores a 5 para el grado del polinomio puede llevar a problemas de sobreajuste. El coef0 es el parámetro necesario para los kernel de tipo polinomial y radial tal y como se puede ver en las ecuaciones 2 y 4; esta medida generalmente toma valores de 0 y 1

Por lo que en esta investigación se realiza una búsqueda en malla para determinar el valor de los parámetros de tal forma que se ajusten al problema, los rangos definidos en la tabla 10 para cada parámetro se determinan ensayando un conjunto de valores e identificando el punto en el cual se obtiene el menor error de validación en un conjunto de entrenamiento.

En la Tabla 11 se registran los porcentajes de mala clasificación para cada modelo junto con su respectivo costo computacional. En las máquinas de soporte se hicieron interacciones para cada tipo de kernel con el objetivo de disminuir el costo computacional, razón por la cual se observa primero el porcentaje de mala clasificación obtenido por el árbol; 42.86% de datos mal clasificados en comparación con 35.71% que obtuvieron las máquinas con kernel lineal. Dado que las máquinas arrojaron el menor porcentaje de clasificación, se entiende que este algoritmo posee mayor aproximación a replicar el conocimiento del experto.

Tabla 11.

Porcentaje de Mala clasificación para los modelos y tiempo de ejecución

	Modelo	%Mala clasificación	Tiempo ejecución
SVM	Lineal	35,71	6 min : 31 seg
	Polinomial	35,71	6 min: 22 seg
	Radial	42,86	7 min: 17 seg
	Sigmoide	42,86	6 min: 55 seg
	DT	42,86	21 min: 15 seg

Cabe resaltar que para las máquinas de soporte vectorial con kernel lineal solo se analiza la variación del Cost, puesto que es el parámetro que afecta a este tipo de kernel (ver ecuación 1). La variación del error presenta una tendencia decreciente tal como se puede ver en la figura 20, a medida que aumenta el cost, el error disminuye. En 4.6 se alcanza el mínimo error, es decir, el valor óptimo con el cual se validó el modelo.

Cost Vs Error

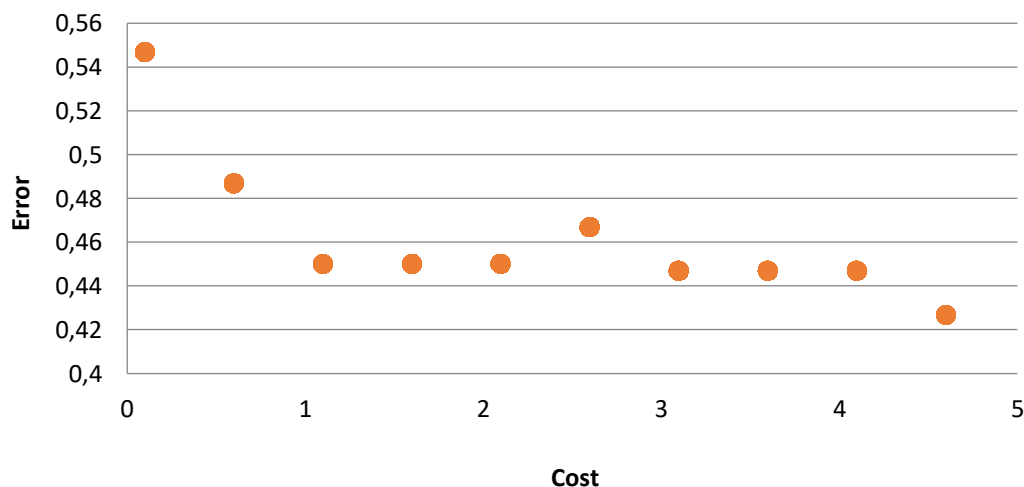


Figura 20. Comportamiento del Error contra la variación del Cost.

6.3.2 Árbol de decisión: El desarrollo del algoritmo para el árbol de decisión se realizó empleando la función rpart de la librería e1071. Esta función requiere de un conjunto de parámetros que permitan controlar el sobreajuste en el modelo: el costo de complejidad “Cp” es una medida que toma valores entre 0 y 1; el maxdepth se evaluó entre un rango de 1 a 10 debido a que después del valor 3 el error se estabiliza (Ver la figura 21); por último, el minsplit que representa el número mínimo de datos para ejecutar una división en un nodo, no puede tomar valores superiores al número total del conjunto de entrenamiento, 54 para este caso. (Ver tabla 12).

Tabla 12.
Valores asignados al rango de parámetros para el árbol de decisión.

Parámetro	Rango									
Cp	0.001	0.011	0.021	0.031	0.041	0.051	0.061	0.071	0.081	0.091
	0.101	0.111	0.121	0.131	0.141	0.151	0.161	0.171	0.181	0.191
	0.201	0.211	0.221	0.231	0.241	0.251	0.261	0.271	0.281	0.291
	0.301	0.311	0.321	0.331	0.341	0.351	0.361	0.371	0.381	0.391
	0.401	0.411	0.421	0.431	0.441	0.451	0.461	0.471	0.481	0.491
	0.501	0.511	0.521	0.531	0.541	0.551	0.561	0.571	0.581	0.591
	0.601	0.611	0.621	0.631	0.641	0.651	0.661	0.671	0.681	0.691
	0.701	0.711	0.721	0.731	0.741	0.751	0.761	0.771	0.781	0.791
	0.801	0.811	0.821	0.831	0.841	0.851	0.861	0.871	0.881	0.891
	0.901	0.911	0.921	0.931	0.941	0.951	0.961	0.971	0.981	0.991
Maxdepth	1	2	3	4	5	6	7	8	9	10
Minsplit	1	2	3	4	5	6	7	8	9	10
	11	12	13	14	15	16	17	18	19	20
	21	22	23	24	25	26	27	28	29	30
	31	32	33	34	35	36	37	38	39	40
	41	42	43	44	45	46	47	48	49	50
	51	52	53	54						

Cp vs Error

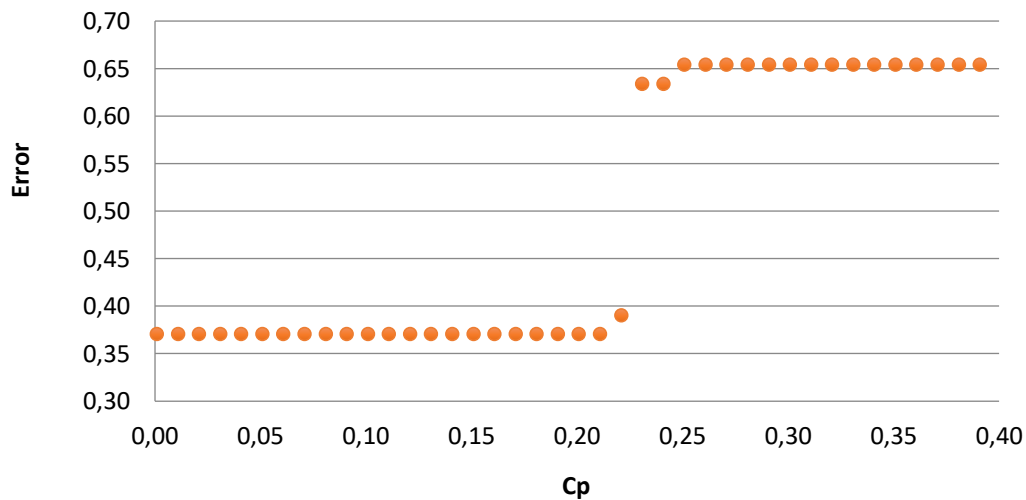


Figura 21. Comportamiento del Error contra la variación del Cp.

La variación del error en función del Cp no es significativa, como se observa en la figura 21, en el intervalo [0,001 , 0,21] el error es 0,37 , y para el intervalo [0,25 , 0,991] el error es 0,65; se mantiene constante en estos dos rangos, lo que significa que el Cp no afecta de manera significativa al sobreajuste del modelo. La validación del modelo se hace con el valor óptimo de los parámetros y esto ocurre donde el error es mínimo, es decir que se puede elegir cualquier valor que contenga el intervalo [0,001 , 0,21], para la ejecución de este algoritmo se usó un Cp igual a 0,001.

El análisis del comportamiento de los parámetros permite analizar que el Cp y el Maxdepth no tienen relación directa con la variación del error. Por otro lado, se observa que el comportamiento del Minsplit no presenta una tendencia en los datos (figura 23). Por esta razón se concluye que la variación del error en el árbol de decisión se ve afectada directamente por el Minsplit. Para validar el modelo se puede escoger cualquier valor entre 23 y 33, puesto que es donde se alcanza el mínimo error.

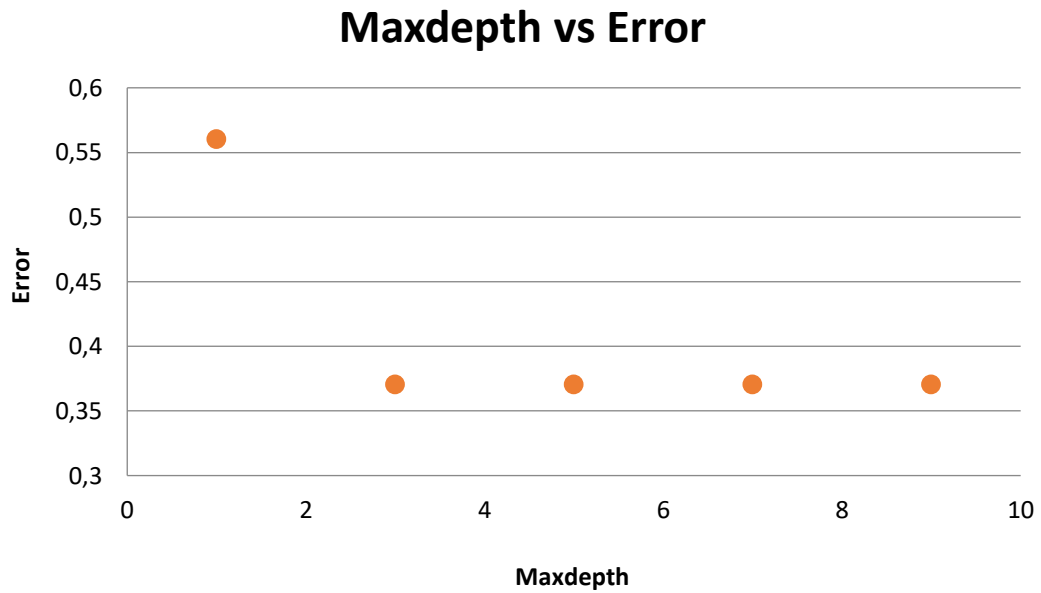


Figura 22. Comportamiento del Error contra la variación del Maxdepth.

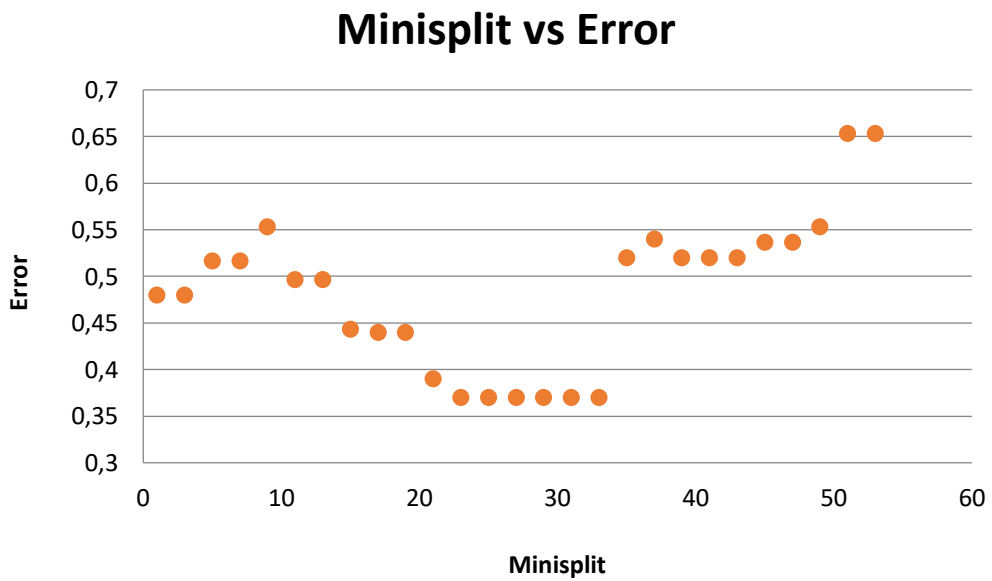


Figura 23. Comportamiento del Error contra la variación del Minisplit.

Para la construcción del árbol se emplea el criterio del índice de Gini explicado en la sección 4. En la figura 24 se visualiza el árbol de decisión obtenido para el caso de estudio,

tomando como variable respuesta la clase a la que pertenece cada dato para posteriormente asociarle un nivel de vulnerabilidad. Las variables que influyen en el desarrollo del árbol son: el año de construcción y el número de pisos. La variable que mejor separa los datos es el año de construcción, por lo que se ubica en el nodo raíz, como era de esperarse el punto crítico está en el año 1984 por la implementación de la normatividad sismo resistente. En promedio, los albergues temporales en Bucaramanga son construidos de 2 pisos, por ende, se hace razonable que el mejor criterio de partición para separar los albergues que fueron construidos antes de 1983 de los que fueron construidos después de este año, sea el número de pisos. Siguiendo con la construcción del árbol se separan los datos en función de la variable año de construcción nuevamente, para finalmente llegar al nodo terminal en el que se clasifica el albergue en las clases 4 o 5, es decir, en albergues con niveles de vulnerabilidad “alto” o “muy alto” según corresponda.

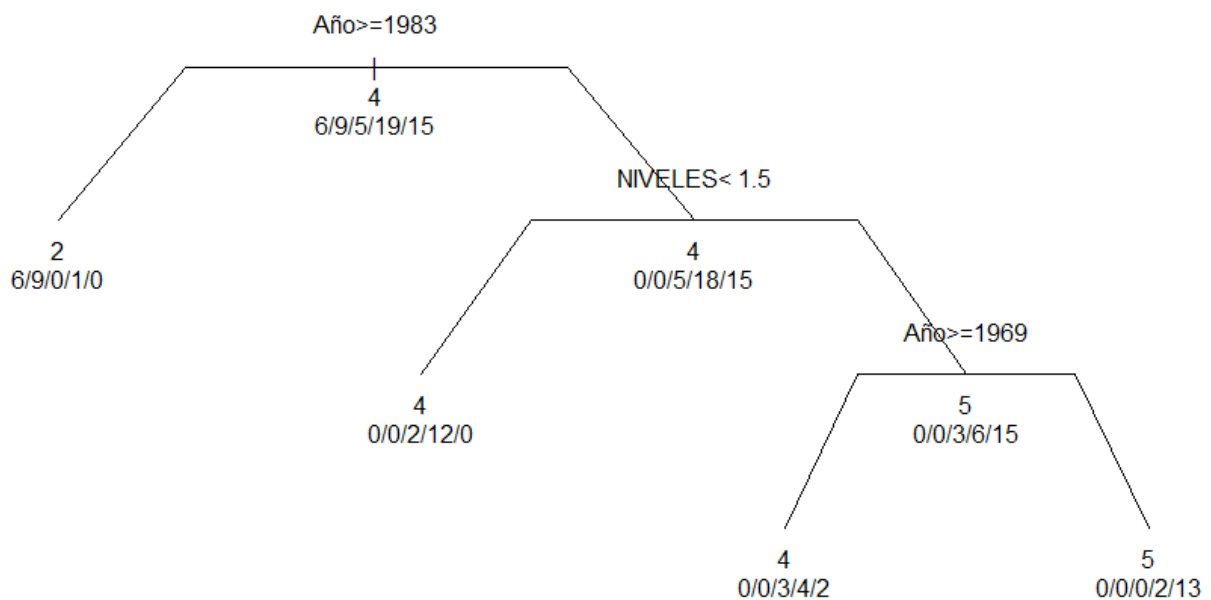


Figura 24. Construcción del árbol de decisión

Una vez se validan los modelos de aprendizaje supervisado propuestos en esta investigación se determina que las máquinas de soporte vectorial son mejores que el árbol de decisión porque cuentan con mayor precisión en el momento de clasificar los datos, es decir, se aproximan en mayor medida a replicar el conocimiento del experto.

7. Conclusiones

- El objetivo de esta investigación fue elegir un modelo de aprendizaje automático supervisado que permitiera clasificar un conjunto de albergues temporales identificados en la ciudad de Bucaramanga en 5 clases de vulnerabilidad. El conocimiento de un experto en evaluaciones de vulnerabilidad es una herramienta clave pero muy costosa, razón por la cual se dio inicio a este trabajo de investigación, con el objetivo de replicar el conocimiento de un experto en evaluaciones de vulnerabilidad sísmica de albergues.

- El aprendizaje automático es una herramienta que aporta grandes contribuciones a problemas de clasificación basados en sistemas de expertos, en este estudio se destacan los siguientes beneficios: Permite que cualquier persona y/o entidad interesada, obtenga una aproximación a la vulnerabilidad sísmica de albergues en Bucaramanga, disminuye el tiempo de respuesta y permite tomar decisiones correctivas en búsqueda de disminuir los riesgos ante un movimiento telúrico de alta magnitud.

- Varios son los estudios realizados en evaluaciones de vulnerabilidad sísmica a estructuras, en la revisión de la literatura se encontró que los problemas de clasificación de desastres se realizan en su gran mayoría por otros métodos diferentes al aprendizaje automático. El proceso de análisis jerárquico “AHP” es un método que goza de una alta participación en la resolución de estos problemas. Sin embargo, desde hace algún tiempo se ha venido desarrollando una

creciente ola de investigación en gestión de desastres implementado técnicas de Machine learning.

- Los árboles de decisión son una herramienta de clasificación muy popular. Se conocen por ser un método efectivo y sencillo para la estimación o predicción, al igual que otros modelos alternativos de aprendizaje automático, pero ofrece una ventaja fundamental al ser un método de fácil entendimiento para personas que no cuentan con conocimientos avanzados en estadística (Hernández, 2004).

- Las máquinas de soporte vectorial se introducen como un modelo estadístico que permite resolver problemas de clasificación de tipo binario. Con ayuda del aprendizaje automático se desarrolla la multi-clasificación de las máquinas de soporte vectorial empleando la función kernel. Esta técnica se hace necesaria para problemas de clasificación en los cuales el conjunto de datos no es linealmente separable. El kernel permite proyectar un conjunto de datos a un espacio de mayor dimensión, permitiendo que las máquinas de soporte determinen un hiperplano óptimo en el que se obtenga el máximo margen de separación entre cada clase.

- Los modelos de aprendizaje supervisado propuestos por los autores, permiten predecir el nivel de vulnerabilidad en cuestión de segundos para un conjunto de datos de prueba que contiene 14 albergues. El Accuracy es una medida de rendimiento que muestra el porcentaje de datos clasificados correctamente, bajo esta métrica se obtiene un rendimiento de 64,29% y 57,14% para las máquinas de soporte y el árbol de decisión respectivamente. En síntesis, se concluye que las máquinas de soporte vectorial con kernel lineal representan la mejor

clasificación de los datos, es decir, es el mejor modelo para dar solución a este problema de investigación.

- La mampostería, como principal sistema de construcción en Colombia y en especial en los albergues temporales evaluados, se considera como una de las tipologías más vulnerable a impactos en sus muros estructurales frente a la ocurrencia de un sismo en comparación con otros sistemas estructurales. La evaluación e identificación de los albergues más vulnerables en Bucaramanga permite la intervención de la alcaldía o entidades interesadas en la implementación de planes de acción. Las técnicas de reforzamiento permitirían disminuir el grado de vulnerabilidad de los albergues y asimismo ofrecer mayor seguridad a la comunidad en caso de un desastre.

8. Recomendaciones

- Para futuras investigaciones se recomienda aumentar el tamaño de la muestra y recopilar proporciones similares de información para cada clase, debido a que la heterogeneidad que se presenta en los datos puede generar fluctuaciones aumentando el error y disminuyendo la eficiencia del modelo. Por otro lado, son muchos los factores que pueden afectar la vulnerabilidad de un albergue, en esta investigación se tuvieron en cuenta aquellos que, según la revisión de la literatura tienen mayor influencia en los niveles de vulnerabilidad. Razón por la cual se recomienda tener en cuenta la integración de diferentes opiniones de expertos y analizar nuevos factores para comprobar si es posible un aumento en la eficiencia de los modelos.

Referencias Bibliográficas

- Barr, A., Feigenbaum, E., & Roads, C. (1982). *The Handbook of Artificial Intelligence, Volume 1*.
- Barrera Ramos, O. E., & Nieves Corredor, O. D. (2015). *Determinación de la vulnerabilidad en las casas coloniales ubicadas en el barrio de san diego de la ciudad de cartagena. Universidad de cartagena*. Universidad de Cartagena.
- Barrientos, R., Cruz, N., Acosta, H., Rabatte, I., Gogeochea, M. del C., Pavón, P., & Blázquez, S. (2009). *Árboles De Decisión Como Herramienta En El Diagnóstico Médico*.
- Benjamin Hernandez, J. D., & Lockhart Castro, S. A. (2011). *Metodología para la evaluación de la vulnerabilidad sísmica de edificaciones de hormigón armado existente. Instituto Tecnológico de Santo Domingo, ciencia y sociedad*. [Instituto Tecnológico de Santo Domingo]. Retrieved from <http://repositoriobiblioteca.intec.edu.do/handle/123456789/1363>
- Betancourt, G. (2005a). Las máquinas de soporte vectorial (svms). *Scientia et Technica*, 1(27). Retrieved from <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/6895>
- Betancourt, G. (2005b). Las máquinas de soporte vectorial (svms). In *Scientia et technica*. Retrieved from <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/6895>
- Bonett Díaz, R. L. (2003). *Vulnerabilidad y riesgo sísmico de edificios. Aplicación a entornos urbanos en zonas de amenaza alta y moderada*. Universitat Politècnica de Catalunya.
- Calle, M. L., & Sánchez-Espigares, J. A. (2007). Árboles de clasificación y regresión en la investigación biomédica. *Medicina clínica*, 129(18), 702-706.
- Camps, D. (2007). *Estudio bibliométrico general de colaboración y consumo de la información en artículos originales de la revista Universitas Médica, período 2002 a 2006. Universitas Médica* (Vol. 48). Facultad de Medicina de la Pontificia Universidad Javeriana. Retrieved from <http://www.redalyc.org/html/2310/231018670002/>
- Cárdenas-Montes, M. PREPROCESADO DE DATOS PARA MINERÍA DE DATOS.
- Carmona Suarez, E. (2013). *Tutorial sobre Máquinas de Vectores Soporte (SVM). Dpto. de Inteligencia Artificial, ETS de Ingeniería*. Madrid.
- Castillo, E., Gutiérrez, J., & Hadi, A. (2011). Sistemas expertos y modelos de redes probabilísticas.
- Caunhye, A. M., Nie, X., & Pokharel, S. (2012). Optimization models in emergency logistics: A

- literature review. *Socio-Economic Planning Sciences*, 46(1), 4–13. <https://doi.org/10.1016/j.seps.2011.04.004>
- Costarricense, C. R. (2003). Sistema de alerta temprana de deslizamientos de tierras en Costa Rica. *San José, CR*.
- Cristianini, N., & J. Shawe-Taylor. (2000). *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*, Cambridge University Press. Retrieved from https://books.google.es/books?hl=es&lr=&id=_PXJn_cxv0AC&oi=fnd&pg=PR9&dq=support+vector+machine+and+other+kernel-based+learning+methods&ots=xSM9I2tXc&sig=y2uRMpyXSqMYxowEpu1oBndLYiQ
- Cuevas, S. M. (2014). *Evaluación De La Vulnerabilidad Sísmica Urbana Basada En Tipologías Constructivas Y Disposición Urbana De La Edificación. Aplicación En La Ciudad De Lorca, Región De Murcia*. Universidad Politécnica de Madrid. <https://doi.org/10.1017/CBO9781107415324.004>
- De La Comisión, S., Colombiano, R., Construcción, D., & Resistente, S. (n.d.). Ministerio de Ambiente, Vivienda y Desarrollo Territorial COMISION ASESORA PERMANENTE PARA EL REGIMEN DE CONSTRUCCIONES SISMO RESISTENTES (Creada por la Ley 400 de 1997) NSR-10 TÍTULO A — REQUISITOS GENERALES DE DISEÑO Y CONSTRUCCIÓN SISMO RESISTENTE. Retrieved from <https://www.idrd.gov.co/sitio/idrd/sites/default/files/imagenes/titulo-a-nsr-100.pdf>
- EIRD/ONU. (2011). *Revelar el riesgo, replantear el Desarrollo*. Green Ink, Devon, Reino Unido.
- Esfera, P. (2011). *Carta Humanitaria y normas mínimas para la respuesta humanitaria*.
- Fallah Aliabadi, S., Sarsangi, A., & Modiri, E. (2015). The social and physical vulnerability assessment of old texture against earthquake (case study: Fahadan district in Yazd City). *Arabian Journal of Geosciences*, 8(12), 10775–10787. <https://doi.org/10.1007/s12517-015-1939-8>
- Franco, Carrasco, S. &, & Trinidad, J. F. M. (2013). Decision Tree based Classifiers for Large Datasets. *Computación Y ...*, 17(1), 95–102. Retrieved from <http://www.scielo.org.mx/pdf/cys/v17n1/v17n1a10.pdf>
- García, L. E. (2015). Desarrollo de la normativa sismo resistente colombiana en los 30 años desde su primera expedición. <https://doi.org/10.16924/riua.v0i41.785>

- Henao, R., Hurtado, J. E., & Castellanos, G. (2005). Selección de hiperparámetros en máquinas de soporte vectorial utilizando adaptación de matriz de covarianza. *Scientia et Technica*, 1(27).
- Hermelín, M. (2007). *Entorno natural de 17 ciudades de Colombia*. Universidad Eafit.
- Hernández, P. A. C. (2004). Aplicación de árboles de decisión en modelos de riesgo crediticio. *Revista colombiana de estadística*, 27(2), 139.
- Herrera, F., Hervás, C., Otero, J., & Sánchez, L. (2004). Un estudio empírico preliminar sobre los tests estadísticos más habituales en el aprendizaje automático. *De La Minería de Datos En ...*. Retrieved from <http://www.lsi.us.es/~riquelme/red/Capitulos/LMD35.pdf>
- Kale, K. V., Mehrotra, S. C., & Manza, R. R. (2010). *Computer vision and information technology: advances and applications*. IK International Pvt Ltd. Retrieved from https://books.google.es/books?hl=es&lr=&id=MRGJN7GI42IC&oi=fnd&pg=PA3&dq=Computer+vision+and+information+technology+advances+and+applications+kale+2010&ots=oIhauYg_9v&sig=ha_O_24GpiiP3ZVlgzXjjO49eSg
- Khamespanah, F., Delavar, M. R., Alinia, H. S., & Zare, M. (2013). Granular Computing and Dempster–Shafer Integration in Seismic Vulnerability Assessment (pp. 147–158). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33218-0_11
- Kuncheva, L. (2004). *Combining pattern classifiers: methods and algorithms*. Canada: WILEY INTERSCIENCE. Retrieved from <https://books.google.es/books?hl=es&lr=&id=B4TZtp7X82oC&oi=fnd&pg=PR5&dq=Combining+Pattern+Classifiers+Methods+and+Algorithms+Kuncheva+2004&ots=Igw4nmY-Np&sig=akJ1CUXE8pVSgbZSuGZcW5TKCUc>
- Lavell, A. (2001). Sobre la gestión del riesgo: apuntes hacia una definición. *Biblioteca Virtual en Salud de Desastres-OPS*. Consultado el, 4.
- Macías, M., Gómez, M., Tous, R., & Torres, J. (2015). *Introducción a Apache Spark*. Retrieved from https://books.google.es/books?hl=es&lr=&id=JqunDAAAQBAJ&oi=fnd&pg=PT7&dq=Introducción+a+Apache+Spark:+para+empezar+a+programar+el+big+data+macias&ots=_d3wy8mL4l&sig=MsWnIQ-gmigTyg_0yeRxp6RK_9Q
- Magidson, J., & Vermunt, J. K. (2005). An extension of the CHAID tree-based segmentation

- algorithm to multiple dependent variables. *Studies in Classification, Data Analysis, and Knowledge Organization*, 176–183. <https://doi.org/10.1007/3-540-28084-7-18>
- Maldonado, E., & Chio, G. (2007). Visión estructural de las edificaciones de la ciudad de Bucaramanga. *Ediciones Universidad Industrial de Santander, Bucaramanga*, 156.
- Maldonado Rondón, E., Chio Cho, G., & Gómez Araujo, I. (2007). *Índice de vulnerabilidad sísmica en edificaciones de mampostería basado en la opinión de expertos*. Pontificia Universidad Javeriana. Facultad de Ingeniería. E SPARC (Organization) (Vol. 11). Pontificia Universidad Javeriana. Retrieved from <http://www.redalyc.org/html/477/47711201/>
- Manzano, J. A. (n.d.). El análisis de segmentación jerárquica mediante CHAID, 1–21.
- Méndez, J. R., Fdez Riverola, F., Díaz, F., & Corchado, J. M. (2007). Sistemas inteligentes para la detección y filtrado de correo spam: una revisión. *Inteligencia artificial, Redalyc*. Obtenido de <http://www.redalyc.org/html/925/92503405/>
- Mitchell, T. (1997). *Machine learning*. WCB.
- Moquete Rosario, F. E. (2012). *Evaluación del Riesgo Sísmico en Edificios Especiales: Escuelas. Aplicación a Barcelona*. Universidad Politécnica de Cataluña.
- Mundial, B. (2012). *Análisis de la gestión del riesgo de desastres en Colombia: un aporte para la construcción de políticas públicas*. Obtenido de <http://gestiondelriesgo.gov.co/sigpad/archivos/GESTIONDELRIESGOWEB.pdf>
- Narváez, L., Pérez Ortega, G., & Lavell, A. (2009). La gestión del riesgo de desastres. Un enfoque basado en procesos.
- Pajares Martinsanz, G., & Santos Peñas, M. (2006). *INTELIGENCIA ARTIFICIAL E INGENIERÍA DEL CONOCIMIENTO*. México, D.F.: Alfaomega Grupo Editor.
- Pérez Marqués, M. (2014). *Minería de datos a través de ejemplos*. Madrid, España: RC Libros.
- Pinzón, A. L. G. (2014). La historia no contada del primer código de construcción en Colombia—30 años. *Revista de Ingeniería*, (40), 82-84.
- Quesada, E. (2007). Sistemas expertos para la enseñanza y el aprendizaje de la matemática en la educación superior. *Cuadernos de Investigación*, 2, 45–67. Retrieved from http://cimm.ucr.ac.cr/cuadernos/cuaderno3/cuaderno3_c2.pdf

- Roche, A. (2009). Árboles de decisión y Series de tiempo, 83.
- Rodríguez, J. D., Pérez, A., & Lozano, J. A. (2010). Sensitivity analysis of kappa-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 569–575. <https://doi.org/10.1109/TPAMI.2009.187>
- Safina Melone, S. (2003). Vulnerabilidad sísmica de edificaciones esenciales. Análisis de su contribución al riesgo sísmico. *TDX (Tesis Doctorals En Xarxa)*. Retrieved from <https://upcommons.upc.edu/handle/2117/93538>
- Sengezer, B., Ansal, A., & Bilen, Ö. (2008). Evaluation of parameters affecting earthquake damage by decision tree techniques. *Natural hazards*, 47(3), 547-568.
- Sepúlveda, J. F. D. (2012). Comparación entre Árboles de Regresión CART y Regresión Lineal, 1–85.
- Sierra, E. A., Hossian, A. A., García Martínez, R., & Marino, P. D. (2005). Sistema experto para control inteligente de las variables ambientales de un edificio energéticamente eficiente. *XI Reunión de Trabajo En Procesamiento de La Información Y Control*, 446–452. Retrieved from <http://www.sistemamid.com/panel/uploads/biblioteca/1/349/1259/6572/6578/77462.pdf>
- Thapelo, S. (2014). Técnicas de aprendizaje automatizado para el pronóstico de temperaturas mínimas en el Centro Meteorológico de Villa Clara, Santa Clara. Retrieved from <http://dspace.uclv.edu.cu/handle/123456789/7341>
- Ulloa, F. (2011). Manual de gestión del riesgo de desastre para comunicadores sociales. *Lima: UNESCO*. Recuperado de <http://unesdoc.unesco.org/images/0021/002191/219184s.pdf>.
- UNISDR, E., & UNISDR. (2009). Terminología sobre Reducción del Riesgo de Desastres. *Estrategia Internacional para la Reducción de Desastres de las Naciones Unidas*.
- Villa, G., Fernán, A., Velásquez, H., Juan, D., Sánchez, S., & Paola, A. (2015). Control del sobreajuste en redes neuronales tipo cascada correlación aplicado a la predicción de precios de contratos de electricidad. *Revista Ingenierías Universidad de Medellín*, 14(26), 161-176.
- Wang, Q. R., & Suen, C. Y. (1984). Analysis and design of a decision tree based on entropy reduction and its application to large character set recognition. *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, 6(4), 406–17.
<https://doi.org/10.1109/TPAMI.1984.4767546>

Yépez, F., Barbat, H. B., & Torres, J. C. (1995). Riesgo, peligrosidad y vulnerabilidad sísmica de edificios de mampostería. Retrieved from <https://upcommons.upc.edu/handle/2117/27297>