

UNIVERSIDAD NACIONAL DE JAÉN

CARRERA PROFESIONAL DE INGENIERÍA CIVIL



**“ESTIMACIÓN DE LA VULNERABILIDAD SÍSMICA EN
VIVIENDAS DE ALBAÑILERÍA CONFINADA, MEDIANTE
TÉCNICAS DE MINERÍA DE DATOS, EN EL SECTOR
PUEBLO LIBRE, JAÉN – 2020”**

**TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
CIVIL**

Autores: Bach. Jose Manuel Palomino Ojeda

Bach. Stefano Rosario Bocanegra

Asesores: M.Sc. Ing. Christiaan Zayed Apaza Panca

Mg. Lenin Quiñones Huatangari

JAÉN – PERÚ, MAYO, 2021



ACTA DE SUSTENTACIÓN

En la ciudad de Jaén, el día 25 de mayo del año 2021, siendo las 12:00 horas, se reunieron de manera virtual los integrantes del Jurado:

Presidente: Mg. Billy Alexis Cayatopa Calderón

Secretario: Mg. José Luis Piedra Tineo

Vocal: Mg. Lenin Franchescoleth Núñez Pintado, para evaluar la Sustentación de:

() Trabajo de Investigación

(**X**) Tesis

() Trabajo de Suficiencia Profesional

Titulado:

ESTIMACIÓN DE LA VULNERABILIDAD SÍSMICA EN VIVIENDAS DE ALBAÑILERÍA CONFINADA, MEDIANTE TÉCNICAS DE MINERÍA DE DATOS, EN EL SECTOR PUEBLO LIBRE, JAÉN – 2020,

presentado por los estudiantes: *Jose Manuel Palomino Ojeda y Stefano Rosario Bocanegra,*

de la Carrera Profesional de Ingeniería Civil de la Universidad Nacional de Jaén.

Después de la sustentación y defensa, el Jurado acuerda:

(**X**) Aprobar () Desaprobar (**X**) Unanimidad () Mayoría

Con la siguiente mención:

- | | | |
|----------------|------------|--------|
| a) Excelente | 18, 19, 20 | () |
| b) Muy bueno | 16, 17 | (16) |
| c) Bueno | 14, 15 | () |
| d) Regular | 13 | () |
| e) Desaprobado | 12 ò menos | () |

Siendo las 13:17 horas del mismo día, el Jurado concluye el acto de sustentación confirmando su participación con la suscripción de la presente.

Mg. BILLY ALEXIS CAYATOPA CALDERÓN
Presidente

Mg. JOSÉ LUIS PIEDRA TINEO
Secretario

Mg. LENIN FRANCHESCOLETH NÚÑEZ PINTADO
Vocal

ÍNDICE

ÍNDICE DE TABLAS	VI
ÍNDICE DE FIGURAS	VII
ÍNDICE DE ANEXOS	IX
RESUMEN	X
ABSTRACT	XI
I. INTRODUCCIÓN	1
1.1. Formulación del problema.....	2
1.2. Justificación de la investigación	2
1.3. Formulación de la hipótesis	3
1.4. Antecedentes de la investigación.....	3
1.4.1. Antecedentes internacionales	3
1.4.2. Antecedentes nacionales.....	4
1.4.3. Antecedentes Regionales	4
1.4.4. Antecedentes locales.....	5
1.5. Bases teóricas	6
1.5.1. Sismicidad en el Perú	6
1.5.2. Evaluación de la sismicidad	6
1.5.2.1. Medida de los sismos.....	7
1.5.2.2. Intensidad sísmica.....	7
1.5.2.3. Energía de un sismo.....	7
1.5.3. Tipos de vulnerabilidad sísmica en viviendas de albañilería confinada....	7
1.5.3.1. Vulnerabilidad estructural	7
1.5.3.2. Vulnerabilidad no estructural	8
1.5.4. Albañilería confinada	8
1.5.5. Minería de Datos	8

1.5.5.1.	Clasificación de algoritmos de minería de datos	9
1.5.5.1.1.	Árbol de decisión.....	10
1.5.5.1.2.	Redes bayesianas	10
1.5.5.1.3.	Regresión Lineal.....	10
1.5.5.1.4.	Optimización de Secuencia Mínima (SMOreg)	10
II.	OBJETIVOS.....	11
2.1.	Objetivo general	11
2.2.	Objetivos específicos:.....	11
III.	MATERIAL Y MÉTODOS	12
3.1.	Ubicación del proyecto.....	12
3.2.	Materiales	13
3.3.	Métodos	13
3.4.	Variables.....	13
3.4.1.	Variable Independiente.....	13
3.4.2.	Variable Dependiente	13
3.5.	Tipo de investigación.....	13
3.5.1.	Según su finalidad	13
3.5.2.	Según su diseño	13
3.5.3.	Según su enfoque.....	14
3.5.4.	Línea de investigación	14
3.5.5.	Diseño no experimental	14
3.6.	Instrumentos de recolección de datos	14
3.7.	Desarrollo de la investigación	15
3.7.1.	Matriz de datos a partir de variables que influyen en la vulnerabilidad sísmica en viviendas de albañilería confinada.....	15
3.7.1.1.	Descripción de Variables.....	17

3.7.1.1.1.	Tipo y organización del sistema resistente.	17
3.7.1.1.2.	Calidad del sistema resistente.	17
3.7.1.1.3.	Resistencia convencional.	18
3.7.1.1.4.	Posición del edificio y cimentación.	22
3.7.1.1.5.	Diafragmas horizontales.	22
3.7.1.1.6.	Configuración en planta.	23
3.7.1.1.7.	Configuración en elevación.	24
3.7.1.1.8.	Distancia máxima entre muros o columnas.	25
3.7.1.1.9.	Tipo de cubierta.	26
3.7.1.1.10.	Elementos no estructurales.	26
3.7.1.1.11.	Estado de conservación.	27
3.7.1.1.12.	Vulnerabilidad sísmica.	27
3.7.2.	Modelación de la vulnerabilidad sísmica a través de técnicas de minería de datos empleando la metodología KDD.	28
3.7.2.1.	Etapas de selección.	28
3.7.2.2.	Etapas de pre-procesamiento/limpieza.	29
3.7.2.3.	Etapas de transformación/reducción.	31
3.7.2.3.1.	Reducción del número de variables.	32
3.7.2.3.2.	Formato de los datos.	33
3.7.2.4.	Etapas de minería de datos.	34
3.7.2.4.1.	Software Weka.	34
3.7.2.4.2.	Métodos de evaluación.	36
3.7.2.4.3.	Algoritmos de clasificación.	37
3.7.2.4.4.	Resumen de modelos obtenidos.	52
3.7.2.5.	Etapas de intervención/evaluación de datos.	52
3.7.2.5.1.	Predicción de los modelos.	52

3.7.2.5.2. Evaluación de modelos matriz de confusión	53
3.7.3. Mapa de zonificación, según los niveles de vulnerabilidad sísmica del sector de estudio.	55
3.7.3.1. Elaboración de la cartografía y análisis espacial	55
3.7.3.1.1. Construcción de información temática	56
3.7.3.1.2. Construcción de mapa temático.....	58
IV. RESULTADOS.....	60
4.1. Matriz de datos con sus respectivas variables que influyen en la vulnerabilidad sísmica	60
4.2. Modelación de la vulnerabilidad sísmica con técnicas de minería de datos.....	69
4.3. Mapa de zonificación sísmica según los niveles de vulnerabilidad sísmica del sector Pueblo Libre	82
V. DISCUSIÓN.....	83
VI. CONCLUSIONES Y RECOMENDACIONES.....	85
6.1. Conclusiones.....	85
6.2. Recomendaciones	86
VII. REFERENCIAS BIBLIOGRÁFICAS	87
AGRADECIMIENTO.....	89
DEDICATORIA.....	90
ANEXOS.....	91

ÍNDICE DE TABLAS

Tabla 1. Descripción, tipo y rango de variables recolectadas.	15
Tabla 2. Valores recomendados de esfuerzo cortante máximo para mampostería de edificios históricos.	19
Tabla 3. Tipos de diafragmas, utilizados para el cálculo de coeficiente.	21
Tabla 4. Rankin de atributos según el algoritmo InfoGainAttributeEval.....	28
Tabla 5. Variables seleccionadas que mayor información aportan.	29
Tabla 6. Grupos conformados por 3 variables.	32
Tabla 7. Matriz de confusión.....	53
Tabla 8. Resultados del modelo MDL_1, con sus respectivos algoritmos.....	69
Tabla 9. Resultados del modelo MDL_2, con sus respectivos algoritmos.....	70
Tabla 10. Resultados del modelo MDL_3, con sus respectivos algoritmos	70
Tabla 11. Resultados del modelo MDL_4, con sus respectivos algoritmos	70
Tabla 12. Resultados del modelo MDL_5, con sus respectivos algoritmos	72
Tabla 13. Predicción del modelo MDL_01	72
Tabla 14. Precisión del modelo MDL_01	75
Tabla 15. Predicción del modelo MDL_04.....	76
Tabla 16. Precisión del modelo MDL_04	79
Tabla 17. Predicción del modelo MDL_05	79
Tabla 18. Precisión del modelo MDL_05	82

ÍNDICE DE FIGURAS

Figura 1. Clasificación de técnicas DM	9
Figura 2. Ubicación del proyecto	12
Figura 3. Configuración en planta de la estructura	24
Figura 4. Formas consideradas para la evaluación en elevación de una vivienda	25
Figura 5. Importación de las librerías (Pandas y Numpy) y la base de datos.	30
Figura 6. Eliminación y limpieza de los datos duplicados con el comando duplicates().	30
Figura 7. Visualización de los datos limpiados de las variables seleccionadas.	31
Figura 8. Fichero al que se le aplico minería de datos en formato ARFF.....	34
Figura 9. Interfaz de WEKA 3.9.4.	35
Figura 10. Interfaz de la opción Classify.	36
Figura 11. MDL_1-Clasificador DecisiónStump y validación cruzada	37
Figura 12. MDL_2-Clasificador DecisiónStump y validación cruzada	38
Figura 13. MDL_3-Clasificador DecisiónStump y validación cruzada	38
Figura 14. MDL_4-Clasificador DecisiónStump y validación cruzada	39
Figura 15. MDL_5-Clasificador DecisiónStump y validación cruzada	39
Figura 16. MDL_1-Clasificador HoeffdingTree y validación cruzada	40
Figura 17. MDL_2-Clasificador HoeffdingTree y validación cruzada	41
Figura 18. MDL_3-Clasificador HoeffdingTree y validación cruzada	41
Figura 19. MDL_4-Clasificador HoeffdingTree y validación cruzada	42
Figura 20. MDL_5-Clasificador HoeffdingTree y validación cruzada	42
Figura 21. MDL_1-Clasificador J48 validación cruzada	43
Figura 22. MDL_2-Clasificador J48 validación cruzada	44
Figura 23. MDL_3-Clasificador J48 validación cruzada	44
Figura 24. MDL_4-Clasificador J48 validación cruzada	45

Figura 25. MDL_5-Clasificador J48 validación cruzada	45
Figura 26. MDL_1-Clasificador LMT y validación cruzada	46
Figura 27. MDL_2-Clasificador LMT y validación cruzada	47
Figura 28. MDL_3-Clasificador LMT y validación cruzada	47
Figura 29. MDL_4-Clasificador LMT y validación cruzada	48
Figura 30. MDL_5-Clasificador LMT y validación cruzada	48
Figura 31. MDL_1-Clasificador RandomTreey validación cruzada.....	49
Figura 32. MDL_2-Clasificador RandomTreey validación cruzada.....	50
Figura 33. MDL_3-Clasificador RandomTreey validación cruzada.....	50
Figura 34. MDL_4-Clasificador RandomTreey validación cruzada.....	51
Figura 35. MDL_5-Clasificador RandomTreey validación cruzada.....	51
Figura 36. Visualización de la información espacial del sector Pueblo Libre en ArcMap	55
Figura 37. Ingreso de la muestra en estudio de viviendas de albañilería confinada a ArcMap.	56
Figura 38. Conexión de información a la tabla de atributos de los polígonos.	57
Figura 39. Tabla de atributos, donde se asoció la información espacial con la información temática	58
Figura 40. Conversión de los polígonos a ráster para elaborar el mapa de zonificación de acuerdo a la vulnerabilidad sísmica estimada para cada vivienda	59
Figura 41. Análisis en base de la vulnerabilidad sísmica.....	59
Figura 42. Calidad del sistema resistente	60
Figura 43. Configuración en elevación	61
Figura 44. Configuración en planta.....	61
Figura 45. Diagramas horizontales.....	62
Figura 46. Distancia máxima entre muros.....	62
Figura 47. Elementos no estructurales	63

Figura 48. Estado de conservación.....	63
Figura 49. Posición del edificio y de la cimentación	64
Figura 50. Resistencia convencional	64
Figura 51. Tipo de cubierta	65
Figura 52. Tipo y organización del sistema resistente	65
Figura 53. Vulnerabilidad sísmica	66
Figura 54. Grafico de barras de las variables VR_2, VR_6, VR_7	66
Figura 55. Gráfico de barras de las variables VR_9, VR_3, VR_10, VR_1	67
Figura 56. Gráfico de barras de las variables VR_4, VR_5, VR_11, VR_8	67
Figura 57. Gráfico de barras de las 11 variables de las cuales depende la vulnerabilidad sísmica y que conforman la base de datos.....	68

ÍNDICE DE ANEXOS

ANEXO I.....	92
ANEXO II.	102
ANEXO III.....	111
ANEXO IV.....	114

RESUMEN

El objetivo de la investigación fue pronosticar la vulnerabilidad sísmica de las viviendas de albañilería confinada del sector Pueblo Libre-Jaén, empleando técnica de minería de datos. La matriz de datos fue obtenida a través de una ficha de recolección de información la cual contiene características de viviendas de albañilería confinada. La metodología empleada, fue Knowledge Discovery in Databases (KDD); la cual se desarrolló en diversas etapas, entre ellas la identificación de las variables que influyen en la vulnerabilidad sísmica, la eliminación de datos anómalos y duplicados utilizando las diferentes librerías de Phyton, la reducción de variables que nos permitió crear cinco modelos, y en la etapa de minería de datos se utilizó el software Weka 3.9.4 con cinco algoritmos de clasificación Decision Stump, Hoeffding Tree, J48, Logistic Model Trees (LMT) y Random Tree los cuales obtuvieron una confiabilidad superior al 82% en el pronóstico de la vulnerabilidad sísmica. Además, se generó un mapa de zonificación del sector, con los niveles de vulnerabilidad pronosticados, en el software ArcGIS utilizando la librería de ArcMap, representando las viviendas con vulnerabilidad baja y media.

Palabras clave: Aplicación informática, prevención antisísmica, vivienda.

ABSTRACT

The objective of the research was to predict the seismic vulnerability of confined masonry dwellings in the Pueblo Libre – Jaen sector, using data mining techniques. The data matrix was obtained through a data collection form containing characteristics of confined masonry houses. The methodology used was Knowledge Discovery in Databases (KDD), which was developed in several stages, including the identification of the variables that influence seismic vulnerability, the elimination of anomalous and duplicate data using the different Python libraries, the reduction of variables that allowed us to create five models, and in the data mining stage, the Weka 3 software was used. 9.4 with five classification algorithms Decision Stump, Hoeffding Tree, J48, Logistic Model Trees (LMT) and Random Tree which obtained a reliability higher than 82% in the seismic vulnerability prediction. In addition, a zoning map of the sector was generated, with the predicted vulnerability levels, in the ArcGIS software using the ArcMap library, representing the houses with low and medium vulnerability.

Key words: Computer application, seismic prevention, housing.

I. INTRODUCCIÓN

Las viviendas de albañilería confinada en el mundo son las más usadas por la población ya que son de bajo costo, comparado con otro tipo de sistema estructural. Estas siempre están sometidas a movimientos sísmicos de baja, media y gran intensidad por el movimiento de las placas tectónicas, ocasionado pérdidas de vidas, económicas y materiales.

En el Perú, durante los últimos años se ha evidenciado que una gran cantidad de edificaciones a causa de sismos intensos, han sufrido daños y deterioro de su estructura, tales como, el sismo acontecido en Nazca en 1996 y Arequipa en el 2001, que provocaron daños estructurales irremediables causando el colapso en estas edificaciones (Gómez & Loayza, 2014). Estos eventos sísmicos se han producido frecuentemente, generando diversos daños a las viviendas de albañilería como grietas en paredes, columnas y vigas, dejándolas vulnerables a la ocurrencia de otro evento sísmico.

En el Perú, la mayor parte de las construcciones de albañilería confinada son autoconstruidas y no consideran la norma técnica de albañilería (E.070), y la norma de diseño sismorresistente (E.030) por el contrario, se basan en el asesoramiento de un albañil o maestro de obra con conocimientos empíricos, lográndose obtener estructuras vulnerables ante un posible evento telúrico, lo cual provoca que las viviendas no puedan asegurar la seguridad de los habitantes y el correcto comportamiento estructural.

La presente investigación se realizó con el propósito de estimar la vulnerabilidad sísmica en viviendas de albañilería confinada del sector Pueblo Libre empleando técnicas de minería de datos, para determinar en qué nivel de vulnerabilidad se encuentra y los posibles escenarios de daño, que se puedan suscitar ante la ocurrencia de un evento sísmico, valorando en todas las viviendas de albañilería confinada su nivel de vulnerabilidad en una escala cualitativa de baja, media y alta plasmada en un mapa de zonificación.

1.1. Formulación del problema

El crecimiento poblacional y migración hacia la ciudad de Jaén, ha acrecentado la necesidad de vivienda de sus pobladores. De acuerdo con el censo del Instituto Nacional de Estadística e Informática (INEI), (2017); la población de Jaén tiene un crecimiento promedio anual de 0.91%. Produciendo que la construcción de viviendas, tenga un aumento en su demanda, siendo la albañilería confinada, el sistema constructivo más utilizado.

Las viviendas de albañilería confinada en Jaén son construidas rústicamente con materiales artesanales, y sin la asesoría de un profesional en el proceso constructivo, siendo esto muy frecuentes en nuestro país, por lo que se dispone estudiar estas viviendas en el Sector Pueblo Libre, considerando el peligro sísmico del distrito de Jaén, el cual se ubica en la zona sísmica dos según la norma E.030.

Por lo anterior, se ha decidido abordar la vulnerabilidad sísmica de viviendas de albañilería confinada, así como la implementación de algunas de las técnicas de minería de datos como coadyuvante en la toma de decisiones; ya que, por el desarrollo de distintos programas de software, algoritmos y maneras de visualizar los resultados, resulta una manera eficiente de tratar esta problemática.

En este contexto se planteó el siguiente problema: ¿Cuál será la vulnerabilidad sísmica de viviendas de albañilería confinada aplicando técnicas de minería de datos en el Sector Pueblo Libre, Jaén – 2020?

1.2. Justificación de la investigación

La ciudad de Jaén se encuentra en la zona sísmica dos según el mapa de zonificación sísmica nacional establecida en la norma E.030, la cual presenta aceleraciones de 0.25g, categorizándola como una zona de sismicidad media. Sin embargo, recientemente se han producido sismos de intensidad moderada, que han sido registrados, por la estación acelerográfica de la Universidad Nacional de Jaén (UNJ), lo cual nos permite investigar el estado actual de las viviendas de albañilería confinada del sector Pueblo Libre en la ciudad de Jaén, con la finalidad de reconocer las características de las viviendas y todos los puntos débiles que fallarían al acontecer un evento sísmico.

Para disminuir el nivel de vulnerabilidad sísmica que presentan las diferentes edificaciones en estudio, se planteó realizar un mapa de zonificación sísmica, para que en base a este se realice las recomendaciones pertinentes para la prevención y mitigación de los diferentes daños frente a un evento sísmico.

Debido a la dificultad de pronosticar estos eventos, esta investigación se realizó con la finalidad de contribuir al conocimiento sobre la vulnerabilidad sísmica en viviendas de albañilería confinada, aplicando nuevos métodos como técnicas de minería de datos. La cual considera varios aspectos, entre ellos el estructural y las condiciones de entorno; ya que se debe estar preparados para afrontarlos.

Se desarrolló el estudio para poder categorizar las estructuras, según el nivel de vulnerabilidad en que se encuentren; alta, media o baja para plantear las mitigaciones o acciones pertinentes.

1.3. Formulación de la hipótesis

La vulnerabilidad sísmica en viviendas de albañilería confinada mediante técnicas de minería de datos, se encuentra en el nivel baja a media con un error menor al 25 %, en el sector Pueblo Libre, Jaén – 2020.

1.4. Antecedentes de la investigación

1.4.1. Antecedentes internacionales

López (2019), en la investigación “Análisis cualitativo y comparativo del método Benedetti-Petrini y la NRS 2010, desarrollado en edificaciones de uno y dos pisos en el barrio bijao, municipio del Bagre Antioquia”, donde tuvo como finalidad evaluar la vulnerabilidad sísmica de viviendas en el municipio del Bagre Antioquia, por el método Benedetti-Petrini y la Norma de construcciones Sismo Resistentes (NRS 2010), estas metodologías toman en cuenta la organización del sistema resistente, calidad del sistemas resistente, influencia del terreno y la cimentación, losas y diagramas, configuración en planta, configuración en elevación, conexiones de elementos críticos, tipo de cubierta, elementos no estructurales, estado de conservación; para la interpretación de los resultados se definieron los siguientes rangos: vulnerabilidad < 30% baja, $30\% \geq$ vulnerabilidad $\geq 60\%$ media, vulnerabilidad $\geq 60\%$ alta;

finalmente el estado de vulnerabilidad global de las viviendas del barrio Bijao en el municipio de Bagre, por el método de Benedetti-Petrini fue 72.6% y por la NRS 2010 de 73.8%, donde evidencia la condición de la vulnerabilidad de las viviendas. El aporte de esta investigación, radico en mostrar la cercanía de los resultados entre dos métodos en la determinación de la vulnerabilidad sísmica.

1.4.2. Antecedentes nacionales

Nervi (2017), en la investigación “Análisis de la vulnerabilidad sísmica en viviendas de albañilería confinada según la norma E - 070 del RNE en la Ciudad de Juliaca – Puno”: Elaborado en la Universidad Peruana Unión, que tuvo por finalidad analizar la vulnerabilidad sísmica de las viviendas de albañilería confinada autoconstruidas, sin un previo diseño arquitectónico ni estructural, y que adicionalmente durante la construcción emplearon materiales artesanales que no cumplían con requerimientos normativos. La investigación se fundamentó en la norma E.070 del Reglamento Nacional de Edificaciones (RNE), de igual forma con la metodología planteada por el ingeniero Kuroiwa para la evaluación del riesgo sísmico. Se desarrolló por medio de 40 encuestas a viviendas en la ciudad de Juliaca, la elección estuvo sujeta en su mayoría a factores geográficamente morfológicos. La información del estudio realizado a las viviendas fue compilada en fichas in situ, en el cual se registra la información del proceso constructivo, el tipo de material utilizado y la mano de obra, valorándose en las viviendas el peligro, vulnerabilidad, y riesgo de sismo. Los resultados del estudio documentaron la presencia de riesgo sísmico alto en el caso de originarse un evento telúrico de intensidad media a alta, determinándose el riesgo sísmico del 65% de las viviendas situadas a las afueras de Cusco y el 95% para el sector de salida Huancané.

1.4.3. Antecedentes Regionales

Teobaldo & Santisteban (2015), en la investigación “Vulnerabilidad sísmica en el distrito de Ciudad Eten aplicando índices de Vulnerabilidad (Benedetti - Petrini)”; la metodología aplicada permitió evaluar las edificaciones, obteniendo zonas de vulnerabilidad sísmica para edificaciones de adobe, albañilería y concreto armado en el distrito de Ciudad Eten. Se elaboro una base de datos para las diferentes tipologías de edificaciones del distrito de Ciudad Eten, que fueron utilizados en estudios referentes a vulnerabilidad sísmica y riesgo sísmico. Concluyendo que casi la totalidad de las edificaciones de adobe del distrito presentaron

una vulnerabilidad alta, con un porcentaje de 97.19% y una vulnerabilidad para las edificaciones de albañilería entre baja a media, teniendo un mayor porcentaje para vulnerabilidad media 51.22%. Esto debido a que las edificaciones de albañilería en la zona de estudio han sido afectadas por un regular porcentaje de deficiencias en calidad de muros portante, es decir, se aprecia en muchas edificaciones gran cantidad de muros trabajados con unidades de albañilería artesanal, la cual no provee una buena resistencia ante acciones sísmicas por lo mismo que dicho material no contiene un control de calidad y no puede asegurar una resistencia confiable.

Quiñones & Carrasco (2019), en el artículo científico “Rendimiento académico empleando minería de datos”; crearon una base de datos compuesta de información académica y personal de los estudiantes de la Carrera de Ingeniería de Industrias Alimentarias de la UNJ. A partir de ello, identificaron siete variables que más incidieron en el rendimiento académico, tales como el nombre del colegio, cantidad de cursos desaprobados, ciclo de ingreso a la universidad, lugar del Colegio, número de repeticiones en Matemáticas, el total de cursos y nivel educativo del padre. Finalmente determinaron tres algoritmos de minería de datos que permitieron detectar tres reglas para determinar el rendimiento académico de los estudiantes; estos han sido J48 graft, J48 y PART. Los mismos que fueron determinados con el software Weka con un porcentaje de clasificación correcta mayor de ochenta y tres (83%).

1.4.4. Antecedentes locales

Ramos (2020), en la investigación “Vulnerabilidad sísmica de las viviendas de albañilería confinada ubicadas en el Sector Pueblo Libre en la ciudad de Jaén, Cajamarca – 2020”; determinó el nivel de vulnerabilidad sísmica de las viviendas de albañilería confinada mediante la evaluación de parámetros de medición de vulnerabilidad y las condiciones en las que se encuentran las viviendas y así generar un mapa de zonificación. La evaluación de la vulnerabilidad sísmica por el método de índice de vulnerabilidad desarrollado por Benedetti-Petrini, fue adaptada a la realidad de la zona, donde se determinó el nivel de vulnerabilidad (bajo, medio, alto), en el cual se trabajó con la normativa vigente. Concluyendo que el índice de vulnerabilidad sísmica de las viviendas de albañilería confinada del sector pueblo libre, el 67.36 % poseen vulnerabilidad baja, el 31.34% vulnerabilidad media y además el 1.49% vulnerabilidad alta. Esta investigación permitió la validación de los modelos de minería de datos propuesto en la investigación.

Mogollón (2017), en la investigación “Diagnóstico preliminar de la vulnerabilidad sísmica de las autoconstrucciones de albañilería confinada de ladrillo de arcilla cocida del sector nuevo horizonte de la ciudad de Jaén- Cajamarca” determinó el nivel de vulnerabilidad sísmica y peligro sísmico evaluando las principales características estructurales, calidad de la mano de obra y tipo de suelos de las viviendas autoconstruidas. Aplicando la metodología ya existente, propuesta por Mosqueira, M. Y Tarque, para la determinación del nivel de vulnerabilidad sísmica. Concluyendo que la vulnerabilidad sísmica del sector Nuevo Horizonte se encuentra en un nivel muy alta 53% y 47% alta y un peligro sísmico medio de 57% y 43% bajo no existiendo peligro alto ni muy alto. Esta investigación permitió considerar determinadas variables para la base de datos.

1.5. Bases teóricas

1.5.1. Sismicidad en el Perú

Las zonas más activas del filo occidental de Sudamérica se distinguen del mundo por ser altamente sísmica. El Perú se encuentra entre dos placas: la de Nazca que se inicia en Panamá hasta la integrante sur de guindilla y la sudamericana que nace en la mirilla flota del Perú y recorre la costa sudamericana hasta el macizo del Atlántico Sur. Donde la actividad sísmica está relacionada por la interacción entre la placa de Nazca (oceánica), bajo la placa sudamericana (continental), lo cual genera frecuentes terremotos. Una parte de la variedad de actividad sísmica se produce por las deformaciones que existen a lo largo de la sierra Andina con sismos de menor frecuencia y magnitud. La disposición y acceso de los terremotos en nuestro país ha permitido realizar estándares, utilizando los datos con el fin de enseñanza y de elucubrar las características de la placa de nazca, bajo la influencia de la sudamericana (Tavera & Buforn, 1998, pág. 188).

1.5.2. Evaluación de la sismicidad

El adelanto científico ha permitido aumentar la noción acerca del origen, evaluación del tamaño y manera de extensión, entre otras características, de los terremotos dentro de la corteza terráquea. Los terremotos acontecen cuando el esfuerzo en el suelo alcanza un nivel superior a la resistencia de la roca, causando que los lados contrarios de la misma fallen súbitamente o se deslicen fuertemente pasando de un lado a otro. Estos esfuerzos pueden obrar

perpendicularmente a la falla empujando las rocas entre ellas, o simultáneamente a la falla moviendo las rocas unas versus otras. La tolerancia de la falla está relacionada con la magnitud de estos esfuerzos y la relación de fricción de la materia que la integra, cuando se acumula un esfuerzo suficientemente grande para exceder la resistencia de la rotura, puede acaecer un terremoto produciéndose un estallido en las rocas perdiendo la armonía y liberando la energía almacenada en manera de ondas sísmicas, las cuales mueven las rocas a su entorno (Mena, 2002, pág. 10).

1.5.2.1. Medida de los sismos

Los terremotos o sismos pueden ser expresados cuantitativamente en relación de la cantidad de energía liberada y cualitativamente según el grado de devastación que ellos causan.

1.5.2.2. Intensidad sísmica

La intensidad sísmica, es la fuerza con que se percibe un sismo en la zona damnificada. La medición de la intensidad es cualitativa y se desarrolla con la observación las consecuencias o daños producidos por el sismo en los terrenos, objetos, construcciones, y el efecto que provoca en las personas. Para medir la intensidad se emplea la Escala de Intensidades Modificada de Mercalli (MMI), inicialmente desarrollada por el sismólogo italiano Mercalli y reformado en 1931 que es una escala descriptiva de 12 grados. Existe también la escala de Medvédev-Sponheuer-Kárník (MSK), elaborada por los sismólogos europeos Medvédev, Sponheuer y Kárník (usada en Europa, 1964). Esta escala precisa mejor los efectos de un sismo en las construcciones (Bolaños & Monroy, 2004, pág. 19).

1.5.2.3. Energía de un sismo

“La Energía liberada por un terremoto es la adicción de energía trasferida en forma de ondas sísmicas y la disipada por medio de otros fenómenos en el foco, primeramente, en forma de calor por la alteración en la zona de ruptura” (Yepez, Barbat, & Canas, 1995).

1.5.3. Tipos de vulnerabilidad sísmica en viviendas de albañilería confinada

1.5.3.1. Vulnerabilidad estructural

Se denomina a los efectos dañinos de gran grado que llegan a sufrir todos los componentes estructurales de una edificación a consecuencia de uno o más eventos sísmicos (Vizconde, 2004). Los componentes estructurales están conformados por las partes de la edificación que se encargan de sostener y transmitir el peso de la edificación, y de aquellas fuerzas temporales a la cimentación, la cual las transmite al suelo. Estos elementos principales son: los elementos verticales como las columnas y horizontales como la cimentación, la losa y las vigas.

1.5.3.2. Vulnerabilidad no estructural

Representa a todos aquellos componentes que no forman parte de la estructura. Al producirse un evento sísmico la edificación puede quedar inhabilitada producto de colapsos de algunos equipos o fallas en los diferentes elementos estructurales y arquitectónicos (Vizconde, 2004).

1.5.4. Albañilería confinada

Es el sistema constructivo más utilizado en la construcción de viviendas en el Perú es la albañilería confinada debido al bajo costo que demanda su construcción a diferencia de los demás sistemas constructivos. Este sistema emplea el uso de unidades de albañilería, vigas soleras y columnas de confinamiento para su construcción donde todos los elementos están arriostrados de tal manera puedan soportar eventos sísmicos a lo largo de su vida útil.

Para su procedimiento constructivo se debe construir en primer lugar todos los muros de ladrillos para posteriormente construir las columnas y vigas las que servirán de soporte vertical y horizontal y la losa que estará arriostrada entre sí formando con todos los componentes una sola estructura la cual se despeñará sísmicamente (Aceros Arequipa, 2014).

1.5.5. Minería de Datos

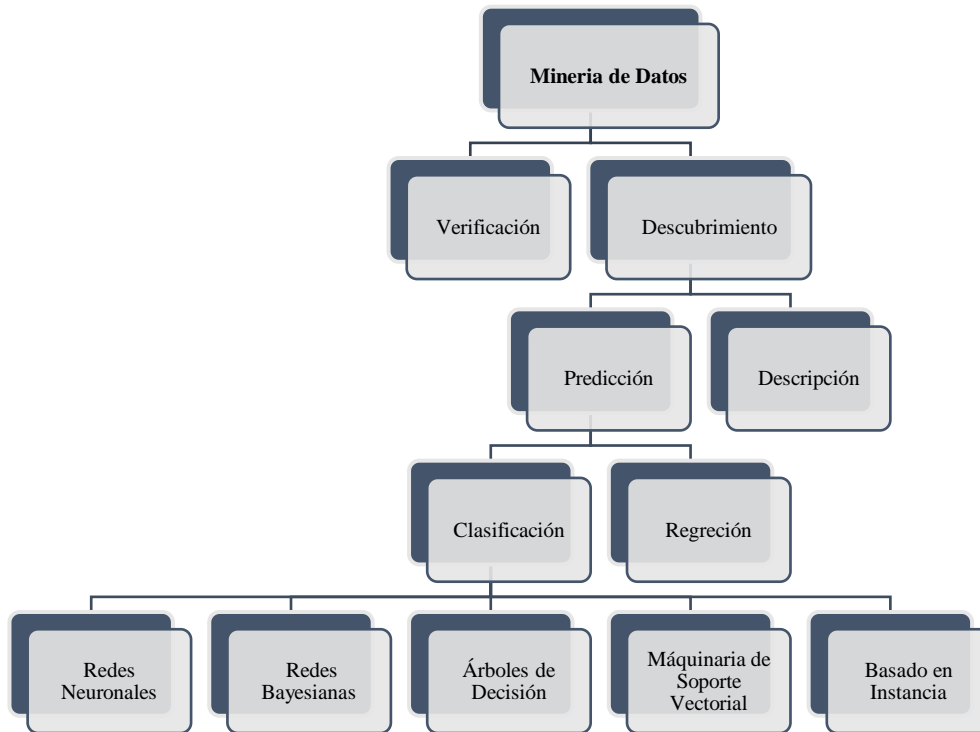
Es aquel proceso que consiste en extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. En ese sentido, la tarea fundamental de la minería de datos es encontrar modelos inteligibles a partir de los datos recogidos (Hernández, Ferrari & Ramírez, 2004). Según Han y Kamber (2006) la minería de datos se refiere a la extracción de conocimiento o minería de grandes cantidades de datos. Ahora bien, al vincular las ideas de datos y extracción, la cual es una opción más óptima para elaborar modelos. Para González (2005), es el proceso por el cual se genera un modelo útil

para la predicción. Dicho modelo se construye teniendo como fundamento los datos que se encuentran en un dataset, a los cuales se le ha aplicado algún algoritmo justamente con el fin de plantear un modelo.

De acuerdo a Maimon y Rokach (2010), está clasificada como la Figura 1.

Figura 1

Clasificación de técnicas DM



Nota, el grafico representa el esquema de la clasificación de la minería de datos. Tomado de Data Mining and Knowledge Discovery Handbook, (2010).

Los sistemas de minería de datos se desarrollan bajo los lenguajes basados en Inteligencia Artificial y utilizando métodos matemáticos tales como:

- Redes Neuronales.
- Sistema Expertos.
- Programación Genética.
- Árboles de Decisión.

1.5.6. Clasificación de algoritmos de minería de datos

1.5.6.1. 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). 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 aplicación, existen diferentes implementaciones comerciales de estas técnicas, así como otras de software libre (Calle y Sánchez, 2007). Los tres tipos de árboles más utilizados son: Chaid, Cart y Quest (Pérez Marqués, 2014, pág. 289).

1.5.6.1.1. Redes bayesianas

Según Aluja (2001), consiste en representar todos los posibles sucesos en los que se está interesado mediante un grafo de probabilidades condicionales de transición entre sucesos. Puede codificarse a partir del conocimiento de un experto o puede ser inferido a partir de los datos. Permite establecer relaciones causales y efectuar predicciones.

1.5.6.1.2. Regresión Lineal

Técnica estadística para determinar la relación entre variables. Permite predecir a partir de un muestreo de datos aleatorio. Se adapta a una amplia variedad de situaciones. Por lo tanto, la regresión ajustada con el error cuadrático medio más bajo, se elige como el modelo final (González & García, 2010). Al aplicar el análisis de funciones automáticamente se genera un modelo de regresión lineal de predicción. La precisión del modelo generado depende en gran manera de la cantidad de datos que se manejen, así, la exactitud de la predicción es directamente proporcional al número de datos disponibles (Lewandowski, 2015).

1.5.6.1.3. Optimización de Secuencia Mínima (SMOreg)

Es el algoritmo de optimización de secuencia mínima desarrollado por John C. Esta implementación sustituye todos los valores que faltan y transforma atributos nominales en otros binarios. También normaliza todos los atributos por defecto (los coeficientes a la salida están basados en los datos normalizados, no en los originales) (González & García, 2010).

II. OBJETIVOS

2.1. Objetivo general

- Estimar la vulnerabilidad sísmica en viviendas de albañilería confinada, mediante técnicas de minería de datos, en el sector pueblo libre, Jaén – 2020

2.2. Objetivos específicos:

- Crear una matriz de datos a partir de variables que influyen en la vulnerabilidad sísmica en viviendas de albañilería confinada, en el sector pueblo libre, Jaén.
- Modelar la vulnerabilidad sísmica a través de técnicas de minería de datos.
- Elaborar un mapa de zonificación, según los niveles de vulnerabilidad sísmica del sector de estudio.

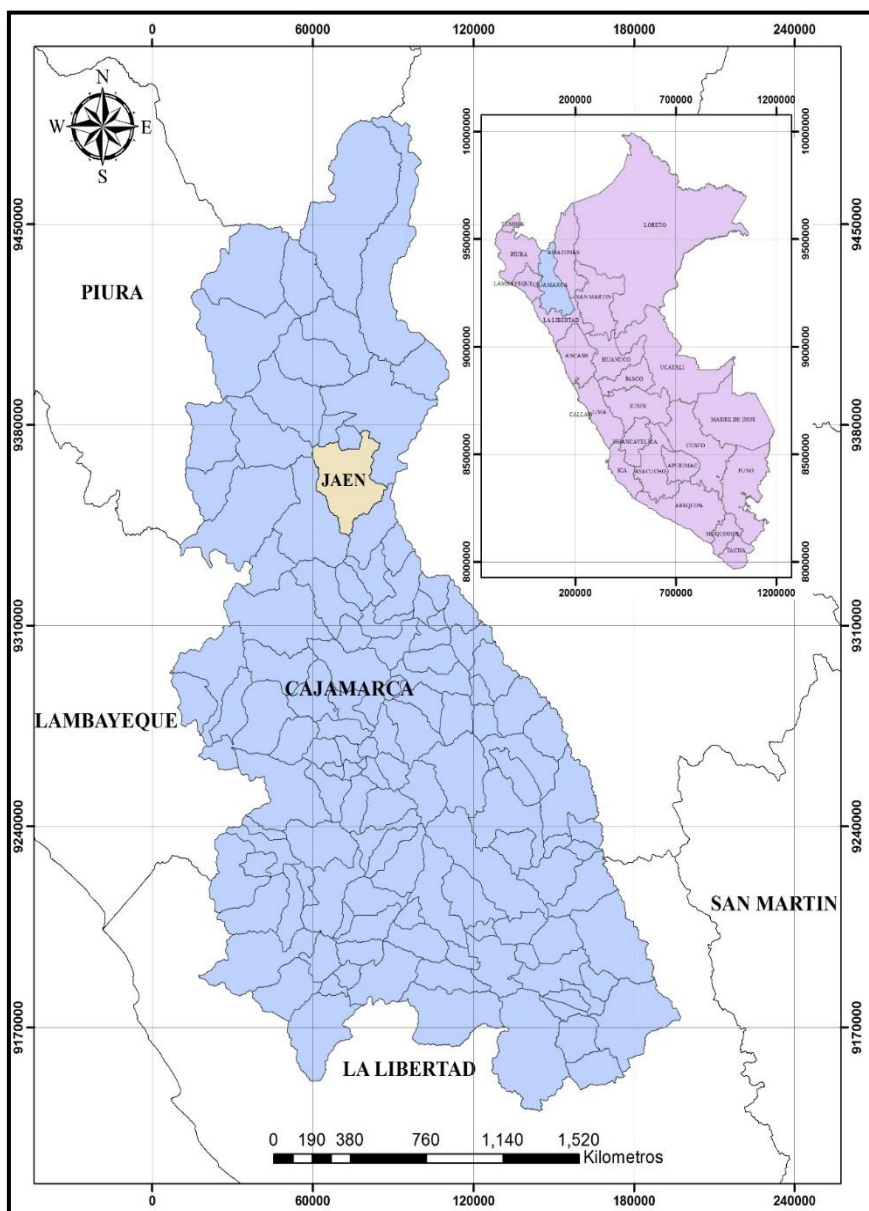
III. MATERIAL Y MÉTODOS

3.1. Ubicación del proyecto

La investigación se desarrolló en el sector pueblo libre, comprendido entre La Calle Iquitos –Avenida Mesones Muro-Calle Coricancha-Calle 28 de julio - la avenida Ejercito y calle pardo miguel localizada en el distrito de Jaén, provincia de Jaén, región Cajamarca, ver Figura 2.

Figura 2

Ubicación del proyecto



Nota, el mapa muestra la ubicación del proyecto mediante un mapa regional.

3.2. Materiales

- Laptop Toshiba Core i7
- Software ArGIS Pro versión 2.7.0
- Software Weka versión 3.9.4
- Software Python versión 3.8.0

3.3. Métodos

- Metodología KDD
- Árbol de decisión.

3.4. Variables

3.4.1. *Variable Independiente*

- Viviendas de albañilería confinada

3.4.2. *Variable Dependiente*

- Vulnerabilidad Sísmica

3.5. Tipo de investigación

3.5.1. *Según su finalidad*

Es básica, porque tuvo el propósito de evaluar la vulnerabilidad sísmica de las viviendas de albañilería confinada mediante técnicas de minería de datos del sector Pueblo Libre, de la ciudad de Jaén, dichos resultados obtenidos se difundirán para que los interesados en esta actividad se informen y puedan tomar acciones en base a los resultados, para la aplicación de acciones preventivas de mitigación ante un evento sísmico y así reducir esta amenaza.

3.5.2. *Según su diseño*

Es no experimental, porque no se hicieron la modificación de las variables, y los resultados de los parámetros medidos sirvieron para estimar el nivel de vulnerabilidad sísmica en el que se encontraron las viviendas de albañilería confinada, cuyo procedimiento fue descriptivo.

3.5.3. Según su enfoque

Es cualitativa, porque las variables empleadas representan las características de las viviendas de albañilería confinada y son no numéricas, obtenidas de diferentes fuentes de investigación, como tesis y artículos científicos, para estimar la vulnerabilidad sísmica en viviendas de albañilería confinada.

3.5.4. Línea de investigación

Estructuras

3.5.5. Diseño no experimental

La investigación presenta tres etapas diferenciadas, una donde se ha recopilado y analizado la información, luego se ha utilizado minería de datos, y finalmente se realizó un mapa de zonificación, según los niveles de vulnerabilidad sísmica.

En la etapa de recolección y análisis de información se generó una matriz de datos, a partir de investigaciones previas donde se determina el índice de vulnerabilidad sísmica en viviendas de albañilería confinada.

En la etapa de minería de datos se crearon cinco modelos con diferentes variables y se utilizó la metodología KDD, con los algoritmos de árboles de decisiones en el software Weka.

En la etapa de la zonificación sísmica se utilizó el sistema de información geográfica para representar los diferentes niveles de vulnerabilidad obtenido.

3.6. Instrumentos de recolección de datos

Se utilizó una ficha de registro de datos y formatos elaborados de acorde al método de índice de vulnerabilidad sísmica y la normativa vigente con el propósito de organizar y registrar la información para su posterior procesamiento, análisis y discusión.

3.7. Desarrollo de la investigación

3.7.1. Matriz de datos a partir de variables que influyen en la vulnerabilidad sísmica en viviendas de albañilería confinada.

Tomando en cuenta la información obtenida de diferentes tesis relacionadas con la vulnerabilidad sísmica de viviendas de albañilería confinada de varios repositorios del Perú y ficha de recolección a 3 827 viviendas desde el 2017 hasta el 2020. Se realizó una base de datos conformada por 3 827 instancias y 12 variables. Su fin fue generar y validar modelos basados en minería de datos que nos permitirán determinar la vulnerabilidad sísmica de las viviendas de albañilería confinada del sector Pueblo libre.

Las variables contempladas fueron: tipo y organización del sistema resistente, calidad del sistema resistente, resistencia convencional, posición del edificio y de la cimentación, diafragmas horizontales, configuración en planta, configuración en elevación, distancia máxima entre muros, tipo de cubierta, elementos no estructurales, estado de conservación, y vulnerabilidad sísmica.

La descripción, tipo y rango de las variables se muestra en la Tabla 1.

Tabla 1

Descripción, tipo y rango de variables recolectadas.

Variable	Descripción	Tipo	Rango
Tipo y organización del sistema resistente	Califica el grado de organización de los elementos verticales	Ordinal	A B C D
Calidad del sistema resistente	Califica el tipo de mampostería más frecuentemente utilizada, dependiendo del tipo y homogeneidad del material	Ordinal	A B C D
Resistencia convencional	Califica la confiabilidad de la resistencia que puede soportar el edificio frente a cargas horizontales	Ordinal	A B C D

Posición del edificio y de la cimentación	Califica la influencia del terreno y de la cimentación frente al comportamiento sísmico	Ordinal	A B C D
Diafragmas horizontales	Califica la conexión del sistema resistente vertical en la transición de las cargas verticales	Ordinal	A B C D
Configuración en planta	Califica la forma en planta de la edificación	Ordinal	A B C D
Configuración en elevación	Califica la forma en elevación de la edificación	Ordinal	A B C D
Distancia máxima entre muros	Califica el espaciamiento excesivo entre los muros ubicados transversalmente	Ordinal	A B C D
Tipo de cubierta	Califica la tipología y peso determinado de la cubierta de una edificación	Ordinal	A B C D
Elementos no estructurales	Califica los elementos no estructurales presente en una edificación que pueden causar daño	Ordinal	A B C D
Estado de conservación	Califica la presencia de desperfectos internos de la estructura, producidos por fallos el proceso constructivo	Ordinal	A B C D
Vulnerabilidad sísmica	Determina el índice de vulnerabilidad presente en las edificaciones	Ordinal	Media Alta Baja

Nota, La presente tabla describe las variables recolectadas en la base de datos, mencionando su tipo y rango.

3.7.1.1. Descripción de Variables

Las variables obtenidas están basadas en el método Benedetti – Petrini, uno de los métodos más utilizados en Italia con excelentes resultados, que fue adoptado por Gruppo Nazionale Per La Difesa Dei Terremoti (GNDT), y utilizado en el país.

Se detalla cómo se califica cada variable en A, B, C y D, de acuerdo a las normas E.070 (albañilería confinada) y E.030 (diseño sismorresistente) entre otras normas internacionales.

3.7.1.1.1. Tipo y organización del sistema resistente.

Para las viviendas de albañilería confinada, esta variable evalúa el grado de organización del sistema resistente vertical de la estructura sin considerar el material utilizado. De tal manera que se tiene en cuenta la presencia y eficacia de las conexiones entre elementos resistentes verticales ortogonales, procurando de esta manera considerar la posible semejanza del comportamiento de la estructura analizada con una ortogonal cerrada, tipo cajón. Además, se hace énfasis en el uso de la norma E.030 e intervención de un profesional para la construcción de viviendas. Para la evaluación de esta variable se tomó en cuenta la norma E.070 y E.030. por lo tanto, la calificación de la variable, se asignará de acuerdo a los puntos descritos:

A: Edificaciones de albañilería que cumplan lo establecido en la norma E.070.

B: Edificaciones que no cumplen con al menos un requisito de lo establecido en la norma E.070.

C: Edificaciones que presentan vigas y columnas que confinan solo parcialmente los muros portantes debido a deficiencias en el proceso constructivo.

D: Edificaciones sin vigas ni columnas de confinamiento o autoconstrucción sin ningún tipo de orientación técnica. Paredes ortogonales deficientemente conectadas.

3.7.1.1.2. Calidad del sistema resistente.

Para las edificaciones de albañilería esta variable evalúa el tipo de mampostería utilizada, diferenciándola cualitativamente por su resistencia. También, se tiene en cuenta la tipología estructural:

- a. La clase de material utilizado y la configuración del elemento de mampostería.
- b. La uniformidad de dicho material y de los elementos a lo largo de la estructura.

Esta variable toma un valor en base de la información recopilada en la base de datos, respecto al material empleado en la construcción de las diferentes partes de las viviendas de albañilería confinada y el año de construcción.

La calificación de la variable se asigna de acuerdo a:

A: El sistema resistente del edificio presenta tres características:

- ✓ Ladrillo de buena calidad y resistencia con piezas homogéneas y de dimensiones constantes por toda la extensión del muro.
- ✓ Presencia de buen amarre entre unidades de albañilería.
- ✓ Mortero de buena calidad con espesor de 10 mm a 12 mm.

B: El sistema resistente del edificio no presenta una de las características de la clase A.

C: El sistema resistente del edificio no presenta dos de las características de la clase A.

D: El sistema resistente del edificio no presenta ninguna de las características de la clase de acuerdo a lo siguiente A.

3.7.1.1.3. Resistencia convencional.

Para las edificaciones de albañilería confinada esta variable evalúa la confiabilidad de la resistencia que puede soportar las viviendas frente a cargas horizontales. Donde se tiene en cuenta la resistencia de los muros en las direcciones principales.

Esta variable se evalúa utilizando la metodología de Hurtado y Cardona, en donde se propone en seleccionar la clase, no empleando el factor “a” como sucede con el método del índice de Vulnerabilidad, sino utilizando la Demanda de Ductilidad (DD). La metodología se inició en 1990 y se desarrolla de la siguiente manera:

- a) Determinar A_x y A_y las cuales son las áreas totales resistentes de los muros (m^2) en la dirección X y Y .

- b) Determinar la menor resistencia cortante, considerando la menor área de muros en un plano en el primer piso de la edificación. La resistencia cortante se calcula como:

$$VR = \min (Ax, Ay) \quad (1)$$

Donde:

VR = Resistencia a la cortante en muros.

Ax, Ay = Áreas resistentes totales.

Cuando este valor no se obtiene experimentalmente, se puede asumir los valores de la Tabla 2.

Tabla 2

Valores recomendados de esfuerzo cortante máximo para mampostería de edificios históricos.

Propiedades mecánicas de algunos tipos de mampostería de edificios históricos				
Material	Peso volumétrico (ton/m³)	Resistencia a compresión (kgf/cm²)	Resistencia a la cortante (kgf/cm²)	Módulo de elasticidad (kgf/cm²)
Adobe	1.80	2-5	0.5	30 000
Bloques de lepetate con mortero de lodo	1.80	5-10	0.5	50 000
Ladrillo con mortero de lodo	1.60	5-10	1.0	50 000
Ladrillo con mortero de cal	1.60	15-20	2.0	100 000
Mampostería de piedra irregular con mortero de cal	2.0	10-15	0.5	50 000
Mampostería de piedra de buena calidad	2.0	30	2.0	200 000

Nota, la tabla describe las propiedades mecánicas de algunos materiales de mampostería. Fuente: Ingeniería Estructural de los edificios históricos.

- c) Calcular el peso total de la edificación que resiste la estructura (W), como peso de los muros, de los pisos y cubiertas.

$$W = N(Ax + Ay)(h)(Pm) + (M)(Ps)(At) + Ac(Pc) \quad (2)$$

Donde:

A_t : Área total construida en planta (m^2).

A_x, A_y : Áreas totales resistentes de muros (m^2) en las direcciones X y Y respectivamente.

h : Altura Promedio de entrepiso (m).

N : Número de pisos de la edificación.

P_m : Peso específico de la mampostería (tn/m^3)

P_s : Peso por unidad de área del diafragma horizontal (tn/m^2).

M : Es el número de diafragmas horizontales.

A_c : Área total de cubierta (m^2).

P_c : Peso por unidad de cubierta (tn/m^2).

Los valores para P_m , donde la mampostería es de ladrillo solido se utilizarán $1.8 \text{ tn}/m^3$ y para P_s , donde son aligerados $0.3 \text{ tn}/m^2$ y para diafragmas abovedados de 0.40 m de promedio, $0.70 \text{ tn}/m^2$.

Para otros valores de diafragmas horizontales se utilizará la Tabla 3.

Tabla 3*Tipos de diafragmas, utilizados para el cálculo de coeficiente.*

Descripción del forjado	Peso (Kg/m ²)	
	Rango	Promedio
Viguetas de madera y entarimado	40-70	55
Viguetas de madera y bovedillas de yeso	100-160	130
Viguetas de madera y tablero de ladrillo	60-140	100
Viguetas metálicas y bovedillas de ladrillo	130-280	205
Viguetas metálicas y mortero ligero	160-390	275
Viguetas de concreto y bovedillas de ladrillo	180-290	235
Viguetas de concreto y bloques huecos	100-180	440
Losa de concreto armado	190-480	335
Losa aligerada de concreto armado	200-320	260
Losa de cerámica armada	150-240	195

Nota, esta tabla describe los diferentes tipos de forjado con su respectivo peso por metro cuadrado. Fuente: Ingeniería Estructural de los edificios históricos.

Los valores para P_c de coberturas de cemento se utilizarán 0.05 tn/m^2 .

- d) Calcular el coeficiente sismo resistente (CSR)

$$CSR = VR/W \quad (3)$$

- e) Calcular el coeficiente sísmico exigido (CSE), tal como como lo menciona el reglamento sismorresistente. Para albañilería según la norma E-0.30.

$$CSE = ZUSC/R \quad (4)$$

- f) Calcular la demanda de ductilidad (DD) como la relación:

$$DD = CSEW/VR \quad (5)$$

- g) Se podrá clasificar con la siguiente adaptación:

$$A: DD < 0.50$$

B: $0.50 \leq DD < 1.00$

C: $1.00 \leq DD < 1.50$

D: $1.50 \leq DD$

3.7.1.1.4. Posición del edificio y cimentación

Esta variable evalúa hasta donde es posible por medio de una simple inspección visual, la influencia del terreno y de la cimentación en el comportamiento sísmico del edificio. Con ese fin se tiene en cuenta los siguientes aspectos:

- La consistencia y la pendiente del terreno.
- La eventual ubicación de la cimentación a diferente cota.
- La presencia de empuje no equilibrado debido a un terraplén.
- Presencia de humedad, sales, etc.

La calificación de esta variable se asigna de acuerdo a lo siguiente:

A: Edificación cimentada sobre suelo rígido y según la norma E.070, sin presencia de humedad ni sales.

B: Edificación cimentada sobre suelo intermedio y flexible, según la norma E.070, sin presencia de humedad ni sales.

C: Edificación cimentada sobre suelo intermedio y flexible, según la norma E.070, con presencia de humedad y sales.

D: Edificación cimentada sin proyecto aprobado ni asesoría técnica, con presencia de sales y humedad. Estado de conservación deteriorado.

3.7.1.1.5. Diafragmas horizontales.

Esta variable califica la calidad del sistema resistente del piso ya que influye en el funcionamiento de los elementos resistentes verticales. Según con los terremotos pasados registrados, se establece que es una alta significancia que el sistema de diafragmas horizontales esté correctamente conectado al sistema resistente vertical para la transmisión de cargas verticales que soporta el edificio y horizontales producidas por efectos de sismo a los paneles

y afectando a la cimentación. Si las conexiones no evidencian estar en una condición aceptable, la vibración lateral de los paneles resistentes se propensa a separarlos de los diafragmas horizontales, induciendo a la caída de estos últimos y el progreso de un mecanismo de falla con peculiaridad de reacción en cadena afectando a todo el sistema; esta reacción se originará como resultado del actuar de altas cargas de impacto. La calificación de esta variable se asigna de acuerdo:

A: Edificaciones con diafragma que satisfacen las siguientes condiciones.

- ✓ Ausencia de planos a desnivel.
- ✓ la deformabilidad del diafragma es despreciable.
- ✓ la conexión entre el diafragma y el muro es eficaz.

B: Edificación que no cumple con una de las condiciones de la clase A.

C: Edificación que no cumple con dos de las condiciones de la clase A.

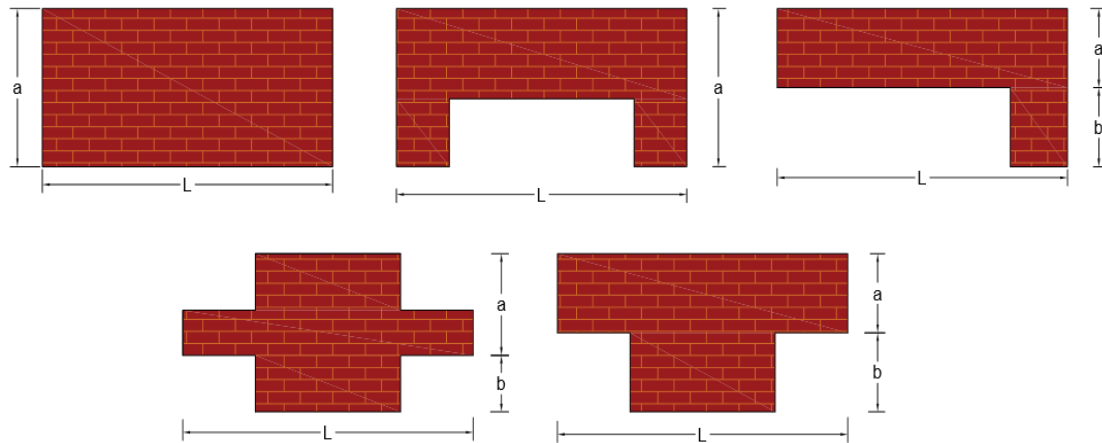
D: Edificación que no cumple con ninguna de las condiciones de la clase A.

3.7.1.1.6. Configuración en planta

Esta variable califica en base a la simetría en planta evaluando la situación de simetría de las viviendas de albañilería, se plantea asignar un valor alto del parámetro cuando las dimensiones en planta sean similares completa o mayoritariamente parcial a secciones cuadradas, sin protuberancias adicionales y disminuyendo el valor asignado si las secciones son excesivamente alargadas o con prominencias demasiado grandes, que pueden producir perturbación en la torsión en planta y la concentración de esfuerzos en las esquinas y en los elementos más distantes de los centros de gravedad y de rigidez, lo cual desencadena en efectos de torsión en la estructura.

Figura 3

Configuración en planta de la estructura



Nota, las figuras muestran las diferentes configuraciones en planta que puede tener una estructura.

Donde la irregularidad en planta (IR) es el máximo valor de lo calculado en el eje X y Y. La calificación de esta variable se asigna de acuerdo:

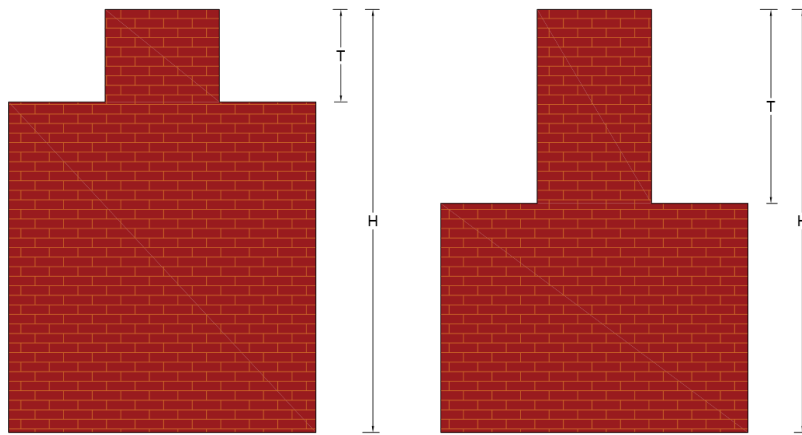
- A:** Edificio con $IR \leq 0.10$
- B:** Edificio con $0.10 < IR < 0.50$
- C:** Edificio con $0.50 < IR < 1.00$
- D:** Edificio con $IR > 1.00$

3.7.1.1.7. Configuración en elevación

Esta variable califica en A, B, C y D, fundamentado en la configuración en elevación que cuando se determina que es inadecuada se tiene problemas como el piso blando, en el cual se incurre cuando hay un cambio brusco de rigidez entre pisos consecutivos. Se considera también a los elementos que integran el sistema resistente y que no tienen continuidad en toda la altura de la edificación.

Figura 4

Formas consideradas para la evaluación en elevación de una vivienda



Nota, las figuras muestran las diferentes configuraciones en elevación que puede tener una estructura.

La calificación de esta variable se asigna de acuerdo a lo siguiente:

A: Edificio con: $\pm \Delta A/A \leq 10\%$

B: Edificio con: $10\% < \pm \Delta A/A \leq 20\%$

C: Edificio con: $20\% < \pm \Delta A/A \leq 50\%$; Presenta discontinuidad en los sistemas resistentes

D: Edificio con: $\pm \Delta A/A \geq 50\%$; Presenta irregularidades de piso blando.

Donde A es el área en elevación y ΔA la diferencia de áreas en elevación.

3.7.1.1.8. Distancia máxima entre muros o columnas

Esta variable evalúa, para edificaciones de albañilería, la distancia máxima entre los muros situados transversalmente a los muros maestros. La clasificación se basa en el resultado de L/S , donde S se refiera al espesor del muro maestro y L al espaciamiento máximo entre los muros transversales.

La calificación de esta variable se asigna de acuerdo a lo siguiente:

A: Edificio con $L/S < 15$

B: Edificio con $15 \leq L/S < 18$

C: Edificio con $18 \leq L/S < 25$

D: Edificio con $L/S \geq 25$

3.7.1.1.9. Tipo de cubierta.

Esta variable toma en cuenta el tipo de cubierta de la edificación, teniendo en cuenta su comportamiento respecto de las solicitaciones de cargas horizontales, y la inestabilidad ante la falta de apoyo en algún extremo de la cubierta sobre un elemento estructuralmente resistente, el cual puede ser una viga o un panel de soporte, no se considera estable si la viga de soporte en la cual se apoya la cubierta es muy esbelta y de poco peralte. La calificación de esta variable se asigna de acuerdo a:

A: El edificio presenta tres características:

- Cubierta estable debidamente amarrada a los muros con conexiones adecuadas, que garanticen el comportamiento de diagrama rígido.
- Provisto de arriostramiento en las vigas y una distancia adecuada entre las mismas.
- Cubierta plana debidamente amarrada y apoyada a la estructura de cubierta de losa aligerada.

B: El edificio no presenta una de las características de la clase A.

C: El edificio no presenta dos de las características de la clase A.

D: El edificio no presenta ninguna de las características de la clase A.

3.7.1.1.10. Elementos no estructurales

Esta variable considera la presencia de elementos no estructurales que puedan dañar la edificación. La calificación de esta variable se asigna de acuerdo a:

A: Edificio que no incluye elementos no estructurales mal conectados al sistema resistente.

B: Edificio con balcones y parapetos bien conectados al sistema resistente.

C: Edificio con balcones y parapetos mal conectadas al sistema resistente. Elementos deteriorados debido a su antigüedad.

D: Edificio que presenta tanques de agua o cualquier otro tipo de elemento en el techo, mal conectado a la estructura. Parapetos u otros elementos de peso significativo, mal construido. Edificios con balcones construidos posteriormente a la estructura principal y conectada a esta de modo deficiente y en mal estado.

3.7.1.1.11. Estado de conservación.

Esta variable es de carácter subjetivo, se califica de manera visual la condición de la estructura, evidenciando algún deterioro o daño, potenciales irregularidades a causa de errores durante el proceso constructivo. La edad de la edificación tiene gran significancia, puesto que cuando se determina el estado de conservación influye la antigüedad de los edificios, dado que el material de construcción se deteriora con el paso del tiempo lo que los que tiene como consecuencia la pérdida de la capacidad resistente. La calificación de esta variable se asigna de acuerdo a:

A: Edificio con en buen estado, sin fisuras visibles.

B: Edificio sin fisuras, pero con cuyos componentes estructurales están levemente deteriorados.

C: Edificio con fisuras y con componentes estructurales deteriorados.

D: Muros con evidente deterioro en sus componentes (presencia de grietas a causa de fallas por flexión, por momento y corte).

3.7.1.1.12. Vulnerabilidad sísmica

Esta variable identifica los parámetros más influyentes en el daño de edificios causados por un evento sísmico, para evaluar los parámetros los calificamos singularmente en una escala nominal donde se resalta su importancia relativa en la obtención del resultado final, la vulnerabilidad sísmica de las viviendas de albañilería confinada en alta, media y baja.

3.7.2. Modelación de la vulnerabilidad sísmica a través de técnicas de minería de datos empleando la metodología KDD

Según Quiroz et al (2012), el término “descubrimiento de conocimiento en bases de datos”, Knowledge Discovery Databases (KDD), fue designado el proceso que pretende obtener conocimiento a partir de datos almacenados en bodegas (data warehouse). En dicho proceso se incluyen la preparación de los datos, el análisis estadístico, el algoritmo usado para la minería de datos y la evaluación e interpretación de los mismos, obteniendo como resultado el descubrimiento de conocimiento que vendría a ser la vulnerabilidad sísmica de las viviendas de albañilería confinada.

3.7.2.1. Etapa de selección.

Identificada la necesidad de estimar la vulnerabilidad sísmica en viviendas de albañilería confinada, en el sector Pueblo Libre, Jaén a través de minería de datos. Se creó un conjunto de datos objetivo, seleccionando la matriz de datos, sobre el cual se realiza el proceso de descubrimiento. Para esta etapa utilizaremos el software Weka, para la selección de los atributos de acuerdo al algoritmo “InfoGainAttributeEval”, el cual asigna pesos a cada variable según la relevancia con la variable de salida (vulnerabilidad sísmica).

Tabla 4

Rankin de atributos según el algoritmo InfoGainAttributeEval.

Variables	Símbolo	Pesos
Estado de conservación	VR_1	0.25853
Tipo y organización del sistema resistente	VR_2	0.24774
Tipo de cubierta	VR_3	0.22988
Posición del edificio y de la cimentación	VR_4	0.19547
Diafragmas horizontales	VR_5	0.19017
Resistencia convencional	VR_6	0.17475
Calidad del sistema resistente	VR_7	0.12815
Configuración en elevación	VR_8	0.0699
Distancia máxima entre muros	VR_9	0.06078
Elementos no estructurales	VR_10	0.03863
Configuración en planta	VR_11	0.00687

Nota, esta tabla muestra la información que aporta cada variable con respecto de la variable de salida (vulnerabilidad sísmica).

Luego de realizar el ranking de atributos se seleccionaron a las variables con un peso mayor a 0.1, por lo cual se trabajó con siete variables (ver Tabla 5) de entrada que son las que mayor información aportan a la variable de salida (vulnerabilidad sísmica).

Tabla 5

Variables seleccionadas que mayor información aportan.

Variables	Símbolo	Pesos
Estado de conservación	VR_1	0.25853
Tipo y organización del sistema resistente	VR_2	0.24774
Tipo de cubierta	VR_3	0.22988
Posición del edificio y de la cimentación	VR_4	0.19547
Diafragmas horizontales	VR_5	0.19017
Resistencia convencional	VR_6	0.17475
Calidad del sistema resistente	VR_7	0.12815

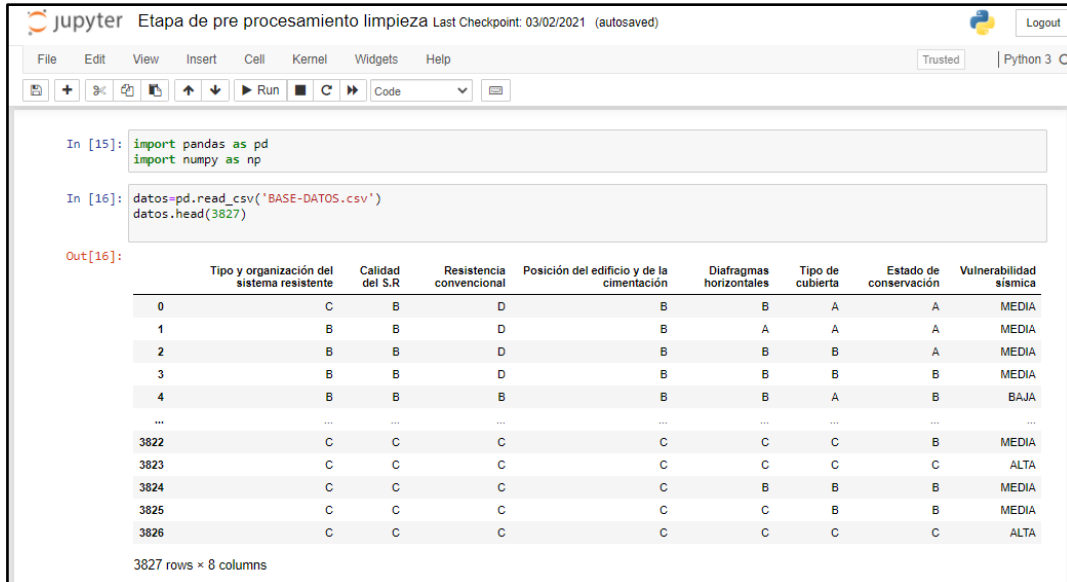
Nota, esta tabla muestra las variables seleccionadas que mayor información aportan a la variable de salida.

3.7.2.2. Etapa de pre-procesamiento/limpieza

Se aplicó la limpieza de los datos con el objetivo de aumentar su calidad, eliminando los datos anómalos. Las tareas de limpieza van encaminadas a resolver dos problemas habituales: la ausencia de valores en algunos registros y la existencia de datos duplicados. Para realizar esta tarea utilizamos python con la interfaz de jupyter y las librerías de análisis de datos: pandas y numpy.

Figura 6

Importación de las librerías (Pandas y Numpy) y la base de datos.



```
In [15]: import pandas as pd
import numpy as np

In [16]: datos=pd.read_csv('BASE-DATOS.csv')
datos.head(3827)
```

Out[16]:

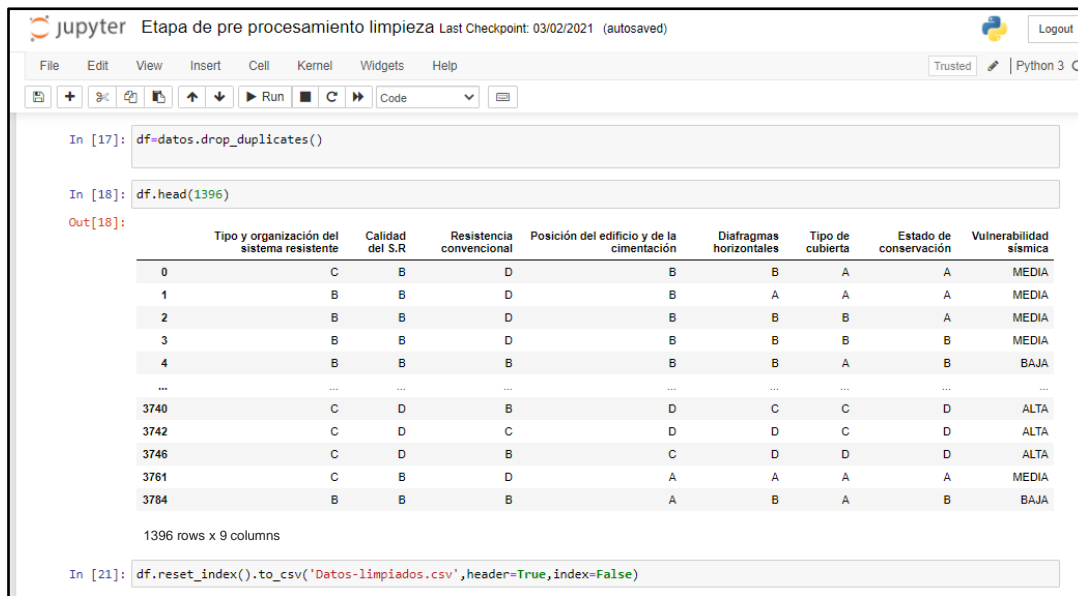
	Tipo y organización del sistema resistente	Calidad del S.R	Resistencia convencional	Posición del edificio y de la cimentación	Diafragmas horizontales	Tipo de cubierta	Estado de conservación	Vulnerabilidad sísmica
0	C	B	D	B	B	A	A	MEDIA
1	B	B	D	B	A	A	A	MEDIA
2	B	B	D	B	B	B	A	MEDIA
3	B	B	D	B	B	B	B	MEDIA
4	B	B	B	B	B	A	B	BAJA
...
3822	C	C	C	C	C	C	B	MEDIA
3823	C	C	C	C	C	C	C	ALTA
3824	C	C	C	C	B	B	B	MEDIA
3825	C	C	C	C	C	B	B	MEDIA
3826	C	C	C	C	C	C	C	ALTA

3827 rows x 8 columns

Nota, la figura muestra el proceso de importación de las librerías de Python y la base de datos.

Figura 5

Eliminación y limpieza de los datos duplicados con el comando duplicates().



```
In [17]: df=datos.drop_duplicates()

In [18]: df.head(1396)
```

Out[18]:

	Tipo y organización del sistema resistente	Calidad del S.R	Resistencia convencional	Posición del edificio y de la cimentación	Diafragmas horizontales	Tipo de cubierta	Estado de conservación	Vulnerabilidad sísmica
0	C	B	D	B	B	A	A	MEDIA
1	B	B	D	B	A	A	A	MEDIA
2	B	B	D	B	B	B	A	MEDIA
3	B	B	D	B	B	B	B	MEDIA
4	B	B	B	B	B	A	B	BAJA
...
3740	C	D	B	D	C	C	D	ALTA
3742	C	D	C	D	D	C	D	ALTA
3746	C	D	B	C	D	D	D	ALTA
3761	C	B	D	A	A	A	A	MEDIA
3784	B	B	B	A	B	A	B	BAJA

1396 rows x 9 columns

```
In [21]: df.reset_index().to_csv('Datos-Limpiados.csv',header=True,index=False)
```

Nota, la figura muestra el proceso de eliminación y limpieza de los datos duplicados.

Finalmente después de haber aplicado el comando “*datos.drop_duplicates()*”, contenido en la librería Pandas, obtenemos 1 396 de los 3 827 instancias de la data recolectada, la cual fue almacenada en un archivo con el nombre ‘Datos-limpiados’ utilizando el comando *df.reset_index().to_csv* (‘Datos-limpiados.CSV’,header = True, index = False). El archivo de salida ‘Datos-limpiados’ se visualiza en la Figura 7.

Figura 7

Visualización de los datos limpiados de las variables seleccionadas.

	A	B	C	D	E	F	G	H
	Tipo y organización del sistema resistente	Calidad del S.R	Resistencia convencional	Posición del edificio y de la cimentación	Diafragmas horizontales	Tipo de cubierta	Estado de conservación	Vulnerabilidad sísmica
1	C	B	D	B	B	A	A	MEDIA
2	B	B	D	B	A	A	A	MEDIA
3	B	B	D	B	B	B	A	MEDIA
4	B	B	D	B	B	B	B	MEDIA
5	B	B	B	B	B	A	B	BAJA
6	B	B	B	C	A	B	B	MEDIA
7	B	B	C	C	B	B	C	MEDIA
8	B	B	D	C	B	B	A	MEDIA
9	B	B	C	B	B	B	B	MEDIA
10	B	C	D	B	C	B	B	MEDIA
11	B	C	C	B	B	B	C	MEDIA
12	B	B	D	B	A	B	C	ALTA
13	B	C	B	B	B	B	C	MEDIA
14	B	C	B	B	B	B	C	MEDIA

Nota, la figura muestra los datos limpiados de la base de datos.

3.7.2.3. Etapa de transformación/reducción

Se buscó características útiles para representar los datos dependiendo de la vulnerabilidad sísmica en viviendas de albañilería confinada, en el sector Pueblo Libre, Jaén. En este sentido se utilizaron métodos de reducción de dimensiones o de transformación para disminuir el número efectivo de variables bajo consideración o para encontrar representaciones invariantes de los datos.

3.7.2.3.1. Reducción del número de variables

Se realizó la reducción de las siete variables obtenidas en la sección 3.7.2.1, a tres variables. Para ello se empleó el análisis combinatorio, ver ecuación (6).

$$C_m^n = \frac{n!}{(n-m)!m!} \quad (6)$$

$$C_3^7 = \frac{7!}{4!3!}$$

$$C_3^7 = 35$$

Obteniendo que las siete variables en grupos de tres, se pueden combinarse de treinta cinco formas diferentes. Se presentan las cinco agrupaciones representativas en la Tabla 6.

Tabla 6

Grupos conformados por tres variables.

Grupo	Variables
Modelo I (MDL_1)	<ul style="list-style-type: none"> ▪ Posición del edificio y de la cimentación ▪ Diafragmas horizontales ▪ Estado de conservación
Modelo II (MDL_2)	<ul style="list-style-type: none"> ▪ Calidad del sistema resistente ▪ Posición del edificio y de la cimentación ▪ Estado de conservación
Modelo III (MDL_3)	<ul style="list-style-type: none"> ▪ Resistencia convencional ▪ Posición del edificio y de la cimentación ▪ Diafragmas horizontales
Modelo IV (MDL_4)	<ul style="list-style-type: none"> ▪ Calidad del sistema resistente ▪ Posición del edificio y de la cimentación ▪ Diafragmas horizontales
Modelo V MDL_5	<ul style="list-style-type: none"> ▪ Calidad del sistema resistente ▪ Resistencia convencional ▪ Posición del edificio y de la cimentación

Nota, esta tabla muestra la agrupación de las variables en grupos de tres conformando cada modelo.

3.7.2.3.2. *Formato de los datos*

El software que se utilizó para realizar la minería de datos fue Weka. Tras la transformación de datos, se obtuvo un fichero que tuvo extensión csv, se decidió convertirlo a formato Attribute-Relation File Format (ARFF), ya que en este formato trabaja Weka. Según Saavedra (2016), la estructura de un fichero con formato ARFF es muy sencilla y se dividen en tres partes: @relation, @attribute y @data.

- ❑ @relation <nombre>: Todo fichero ARFF debe comenzar con esta declaración en su primera línea (no se pueden dejar líneas en blanco al principio). <nombre> será una cadena de caracteres y si contiene espacios se puso entre comillas.
- ❑ @attribute <nombre> <tipo_de_datos>: En esta sección se incluye una línea por cada atributo (o columna) que se vaya a incluir en el conjunto de datos, indicando su nombre y el tipo de dato. Con <nombre> se indicará el nombre del atributo, que debe comenzar por una letra y si contiene espacios se puso entrecomillado. Con <tipo_de_datos> se indicará el tipo de dato para este atributo (o columna) en nuestro caso es <nominal-specification>. Donde todos los datos deben ir separados por coma.
- ❑ @data: En este apartado se incluyen los datos propiamente dichos como las clasificaciones en A, B, C, D y la vulnerabilidad sísmica. Cada columna es separada por comas y todas las filas deben tener el mismo número de columnas, número que coincide con el de declaraciones @attribute, añadidas en la sección anterior.

Figura 8

Fichero al que se le aplico minería de datos en formato ARFF.

```
1 @relation 'DATOS-weka.filters.unsupervised.attribute.'
2
3 @attribute 'Calidad del S.R ' {B,C,D,A}
4 @attribute 'Resistencia convencional' {D,B,C,A}
5 @attribute 'Posición del edificio y de la cimentación ' {B,C,A,D}
6 @attribute 'Vulnerabilidad sísmica' {MEDIA,BAJA,ALTA}
7
8 @data
9 B,D,B,MEDIA
10 B,D,B,MEDIA
11 B,D,B,MEDIA
12 B,D,B,MEDIA
13 B,B,B,BAJA
14 B,B,C,MEDIA
15 B,C,C,MEDIA
16 B,B,C,MEDIA
17 B,D,C,MEDIA
18 B,C,B,MEDIA
19 C,D,B,MEDIA
20 C,C,B,MEDIA
21 B,D,B,ALTA
22 C,B,B,MEDIA
23 C,C,B,MEDIA
24 B,D,C,MEDIA
25 C,D,B,MEDIA
26 B,D,B,MEDIA
27 B,C,B,BAJA
28 C,C,C,MEDIA
29 C,C,B,MEDIA
30 B,D,B,MEDIA
31 B,D,B,MEDIA
```

Nota, la figura muestra el fichero de la base de datos en formato ARFF.

3.7.2.4. Etapa de minería de datos.

3.7.2.4.1. Software Weka

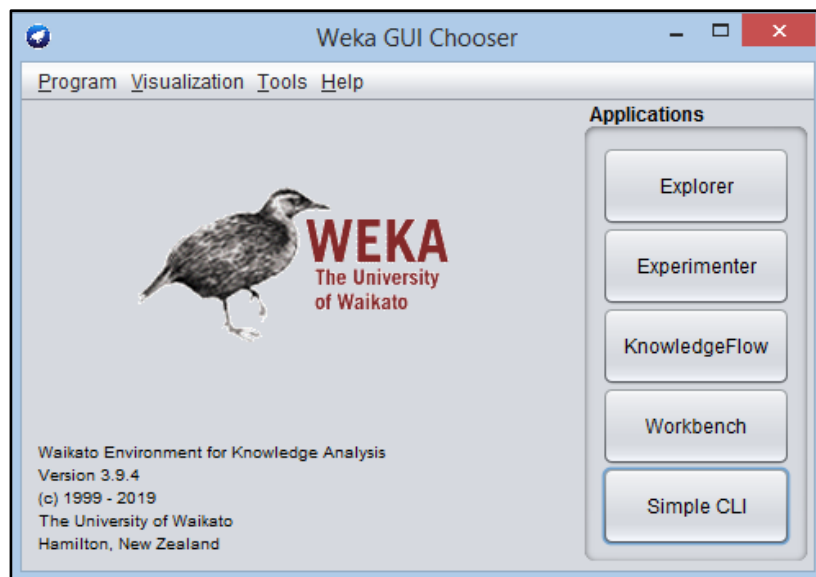
Es un software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato, donde se pudo realizar mediante una interfaz gráfica. La investigación utilizó la interfaz gráfica de WEKA 3.9.4, donde cada una de las tareas de minería de datos viene representada por una pestaña en la parte superior, como se detalla:

- ✓ Preprocess: Visualización y pre procesado de los datos. En esta pestaña tenemos una serie de filtros que aplicamos al conjunto de datos.
- ✓ Classify: Algoritmos de clasificación y regresión.
- ✓ Cluster: Algoritmos de agrupamiento.

- ✓ Associate: Algoritmos para obtener reglas de asociación.
- ✓ Select Attributes: Selección de atributos.
- ✓ Visualize: Nos permite visualizar pares de atributos
- ✓ Experimenter: Sirvió para aplicar diferentes algoritmos de minería de datos sobre distintos conjuntos de datos.
- ✓ Knowledge Flow: Nos permitió realizar lo mismo que la herramienta “explorer” pero de manera gráfica.

Figura 9

Interfaz de Weka 3.9.4.

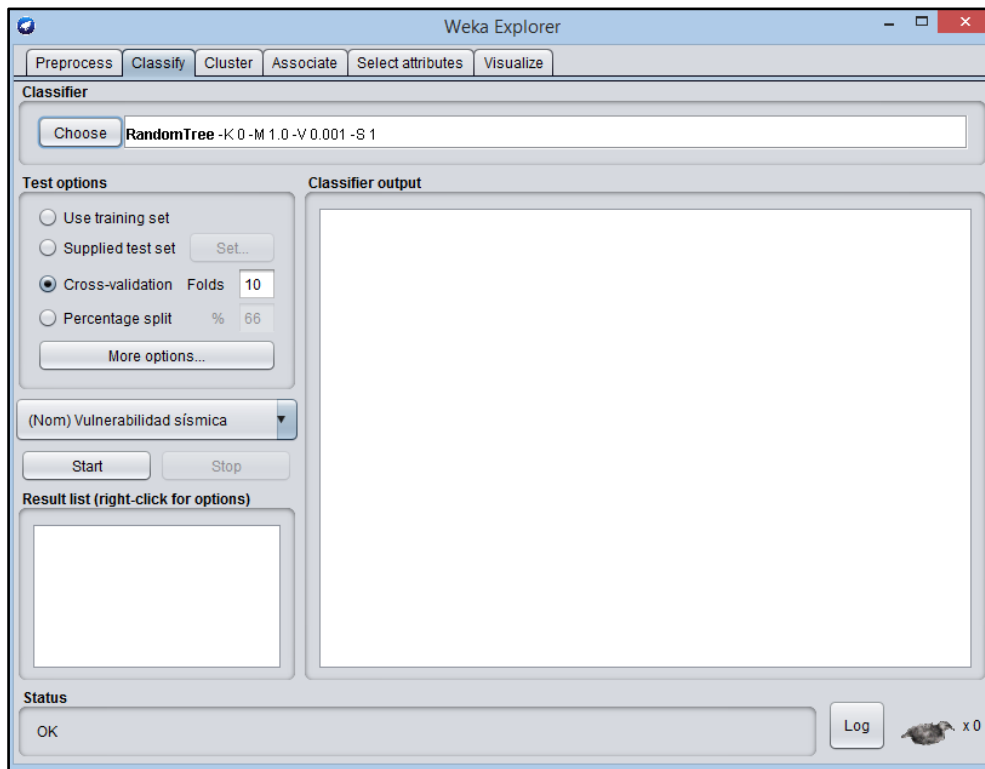


Nota, la figura muestra la interfaz gráfica de Weka 3.9.4.

De todas las funcionalidades mencionadas hicimos uso de la función “Explorer”, utilizando la herramienta “Classify” (Ver Figura 10).

Figura 10

Interfaz de la opción Classify.



Nota, la figura muestra la interfaz de Classify de la herramienta de Explorer.

3.7.2.4.2. Métodos de evaluación

Existen cuatros formas de evaluar los diferentes algoritmos de minería de datos utilizados:

- Use training set: Nos permitió entrenar el modelo con todos los datos disponibles y a posteriori realizar la evaluación sobre los mismos datos.
- Supplied test set: Nos permitió cargar un conjunto de datos, normalmente diferentes a los de aprendizaje, con los cuales se realizó la estimación de la vulnerabilidad sísmica.
- Cross-validation: Evaluación con validación cruzada, donde se dividen las instancias en diferentes carpetas como indica el parámetro Folds, en cada evaluación se tomaron las instancias de cada carpeta como datos de test y el resto como datos de entrenamiento para construir el modelo. Los errores calculados serán el promedio de todas las ejecuciones.

- ❑ Percentage split: Se definió un porcentaje de entrenamiento para el modelo. La evaluación se realiza con los datos restantes.

En esta etapa se ha utilizado la opción de *Cross-Validation* y *Supplied test set* para el entrenamiento del modelo y la predicción de la vulnerabilidad sísmica.

3.7.2.4.3. Algoritmos de clasificación

Los algoritmos fueron aplicados a los diferentes grupos de la Tabla 6.

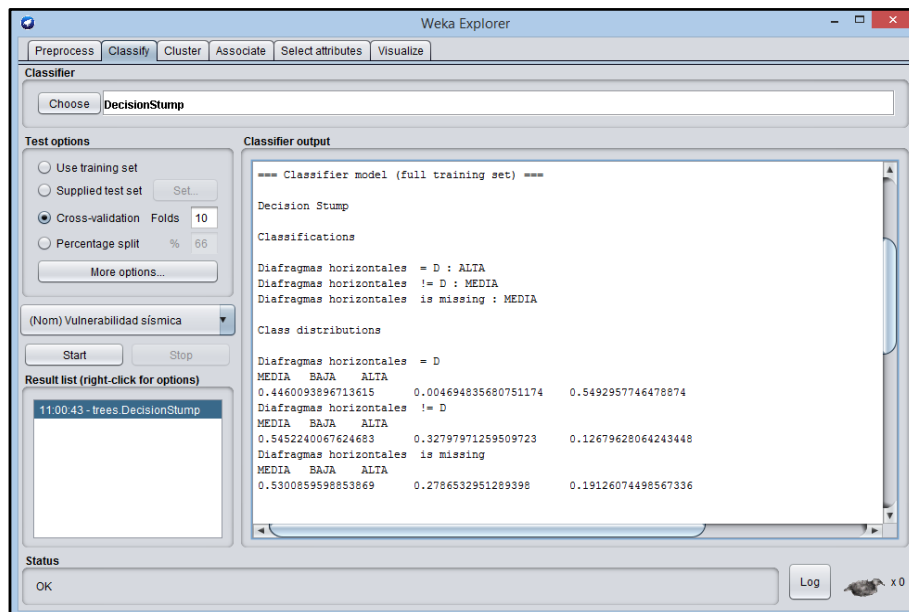
▪ Decisión Stump

Este algoritmo construyó un simple árbol de decisión binaria de único nivel, dicho algoritmo se aplicó indistintamente a conjuntos de datos nominales o numéricos. El manejo que realizó de los valores inclasificables consistió en la adición de una tercera rama unida al nodo raíz (Ferrero & Martínez, 2017).

✓ MDL_1

Figura 11

MDL_1-Clasificador Decision Stump y validación cruzada

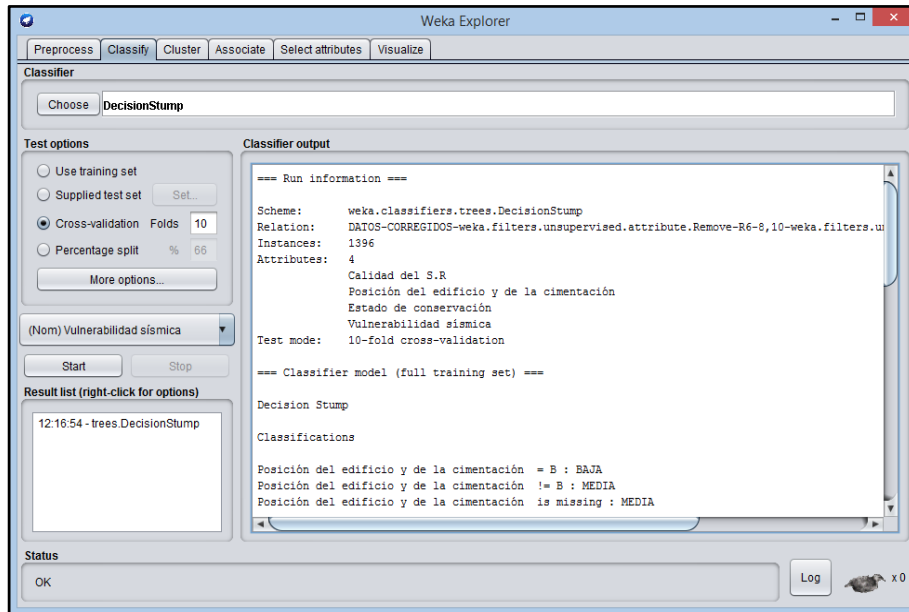


Nota, la figura muestra la clasificación del modelo I con el algoritmo Decision Stump y validación cruzada.

✓ MDL_2

Figura 12

MDL_2-Clasificador Decision Stump y validación cruzada

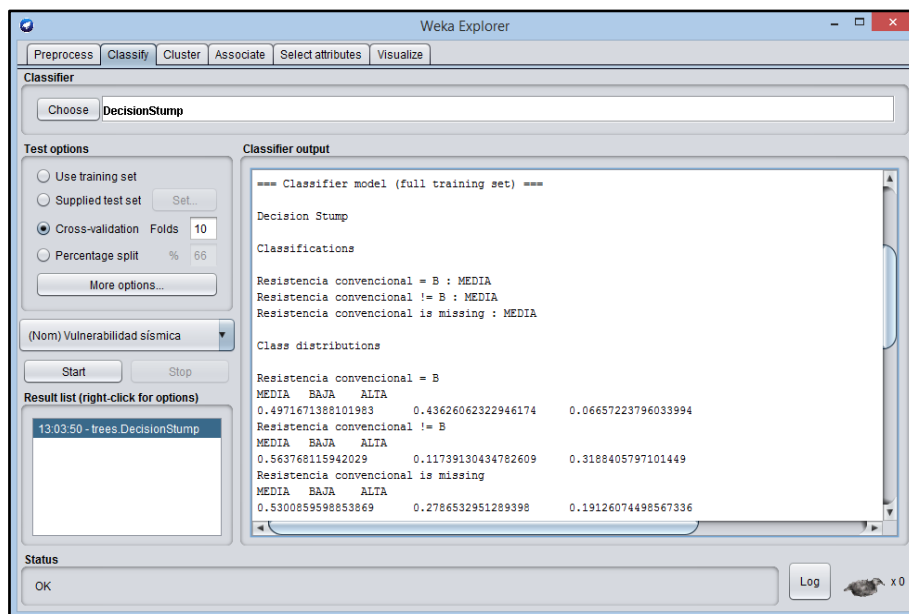


Nota, la figura muestra la clasificación del modelo II con el algoritmo Decision Stump y validación cruzada.

✓ MDL_3

Figura 13

MDL_3-Clasificador Decision Stump y validación cruzada

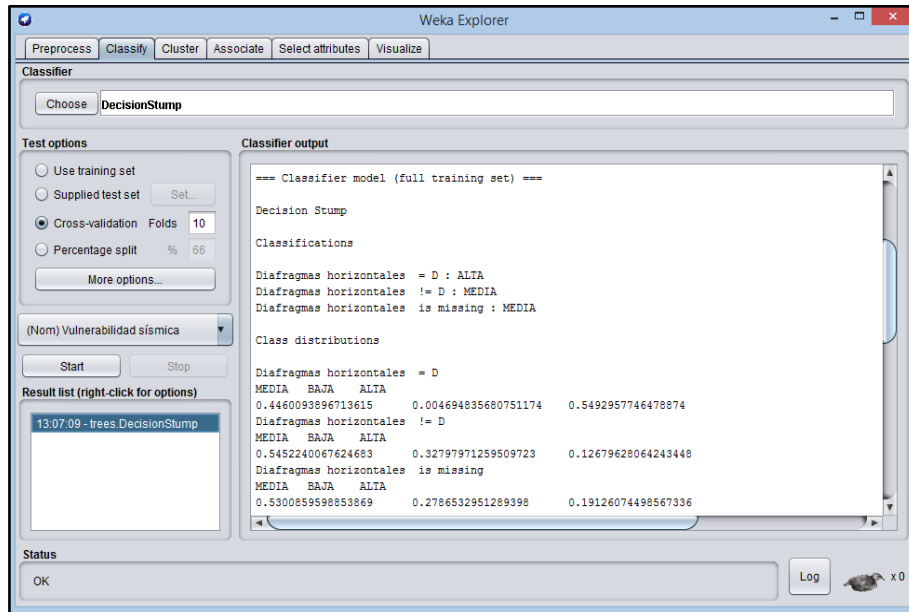


Nota, la figura muestra la clasificación del modelo III con el algoritmo Decision Stump y validación cruzada.

✓ MDL_4

Figura 14

MDL_4-Clasificador Decision Stump y validación cruzada

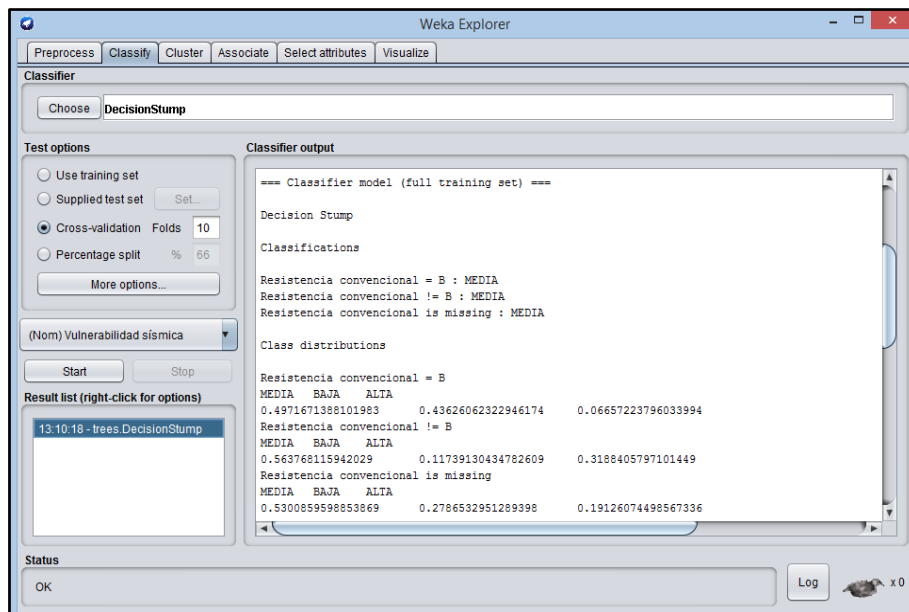


Nota, la figura muestra la clasificación del modelo IV con el algoritmo Decision Stump y validación cruzada.

✓ MDL_5

Figura 15

MDL_5-Clasificador Decision Stump y validación cruzada



Nota, la figura muestra la clasificación del modelo V con el algoritmo Decision Stump y validación cruzada.

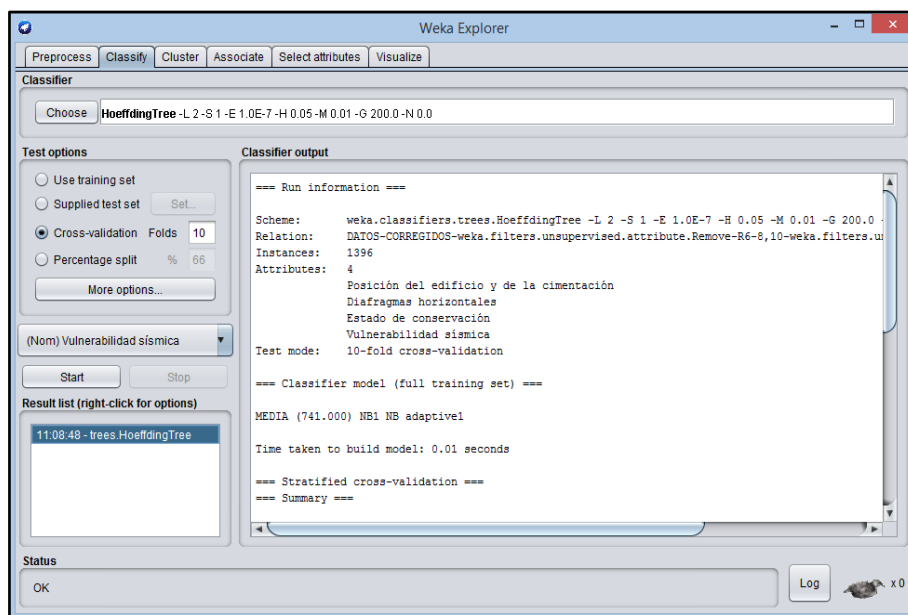
▪ Hoeffding Tree

Es un algoritmo de inducción de árbol de decisión incremental, en cualquier momento, que es capaz de aprender de flujos de datos masivos, asumiendo que los ejemplos de generación de distribución no cambian con el tiempo. Los árboles hoeffding aprovechan el hecho de que una pequeña muestra a menudo puede ser suficiente para elegir un atributo de división óptimo.

✓ MDL_1

Figura 16

MDL_1-Clasificador HoeffdingTree y validación cruzada

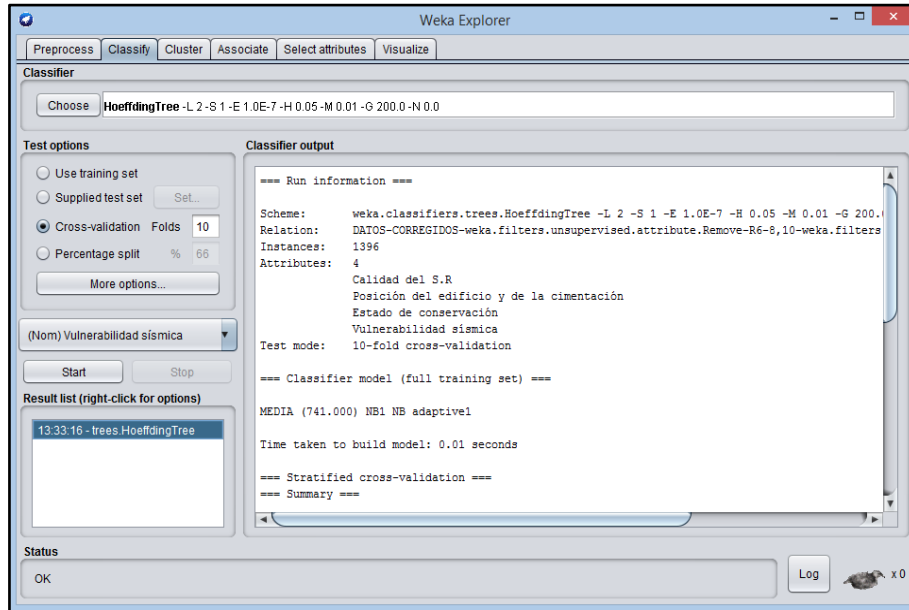


Nota, la figura muestra la clasificación del modelo I con el algoritmo HoeffdingTree y validación cruzada.

✓ MDL_2

Figura 17

MDL_2-Clasificador HoeffdingTree y validación cruzada

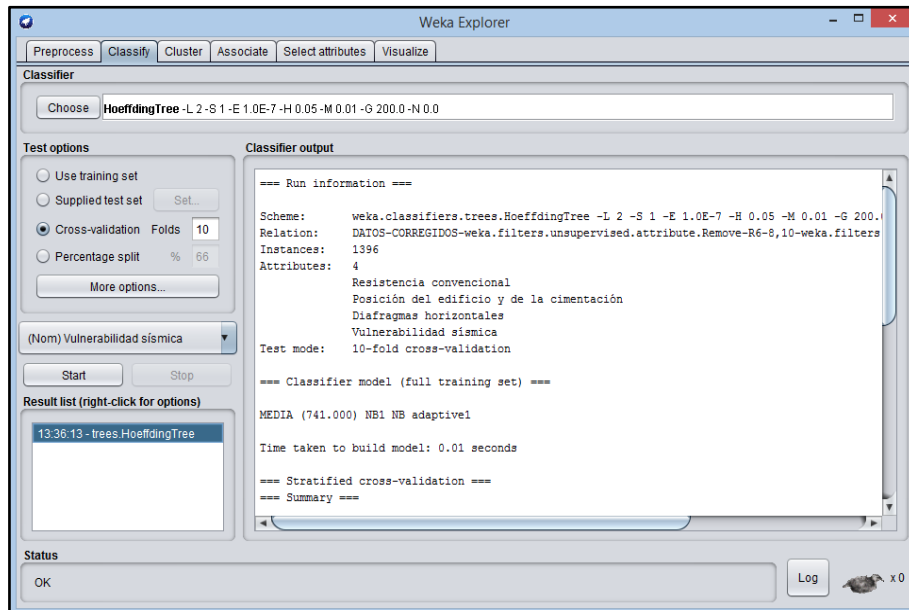


Nota, la figura muestra la clasificación del modelo II con el algoritmo HoeffdingTree y validación cruzada.

✓ MDL_3

Figura 18

MDL_3-Clasificador HoeffdingTree y validación cruzada

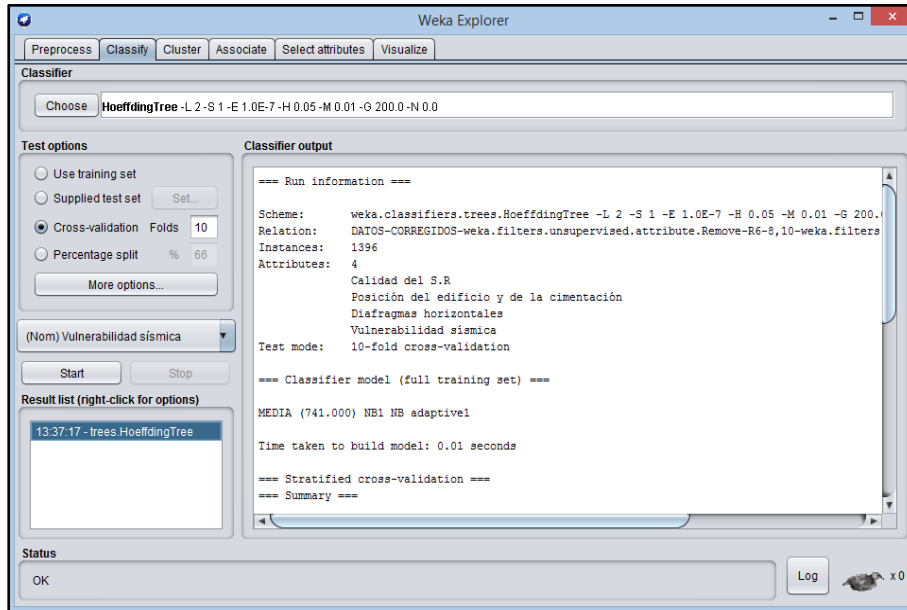


Nota, la figura muestra la clasificación del modelo III con el algoritmo HoeffdingTree y validación cruzada.

✓ MDL_4

Figura 19

MDL_4-Clasificador HoeffdingTree y validación cruzada

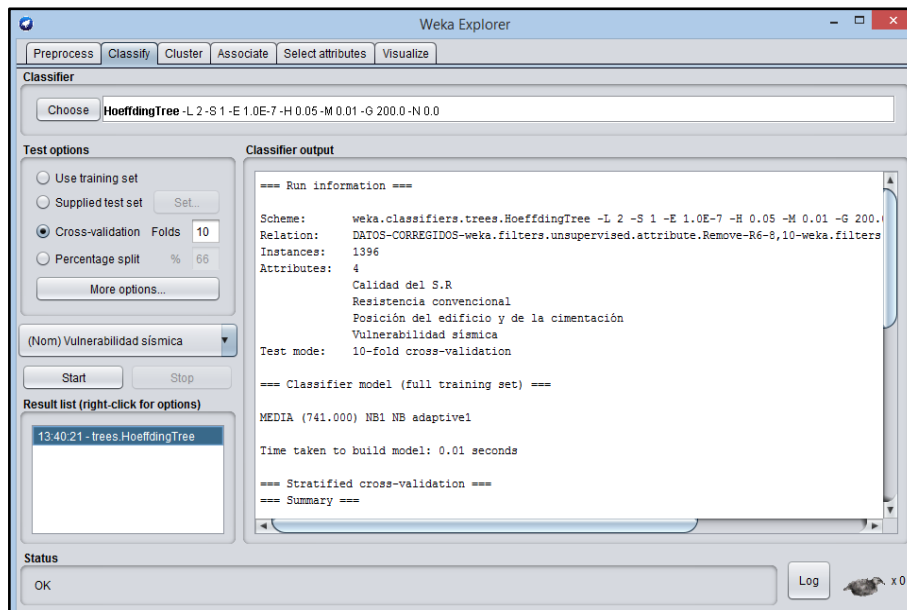


Nota, la figura muestra la clasificación del modelo IV con el algoritmo HoeffdingTree y validación cruzada.

✓ MDL_5

Figura 20

MDL_5-Clasificador HoeffdingTree y validación cruzada



Nota, la figura muestra la clasificación del modelo V con el algoritmo HoeffdingTree y validación cruzada.

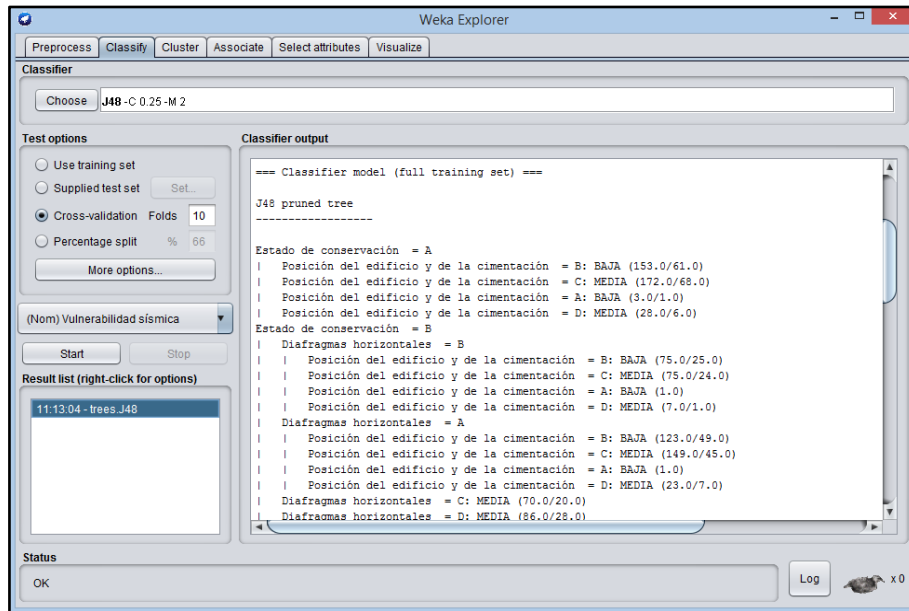
- **J48**

Es un algoritmo que ofrece la posibilidad de poder parar, antes de alcanzar las hojas en cada subárbol; esto dio lugar a árboles menos refinados, y ayudo a evitar el "overfitting". Se pueden utilizar dos parámetros para detener el algoritmo. Por un lado, tenemos el “minNumObj” que establece un límite inferior de elementos, para ser divididos en un conjunto de datos. Cuando un nodo contiene menos de “m” instancias no se le divide o reparte más. Por otro lado, está el “confidenceFactor” el algoritmo que considera la terminación del proceso para un subconjunto “E”, cuando la tasa de desclasificados es menor al umbral especificado.

- ✓ MDL_1

Figura 21

MDL_1-Clasificador J48 validación cruzada

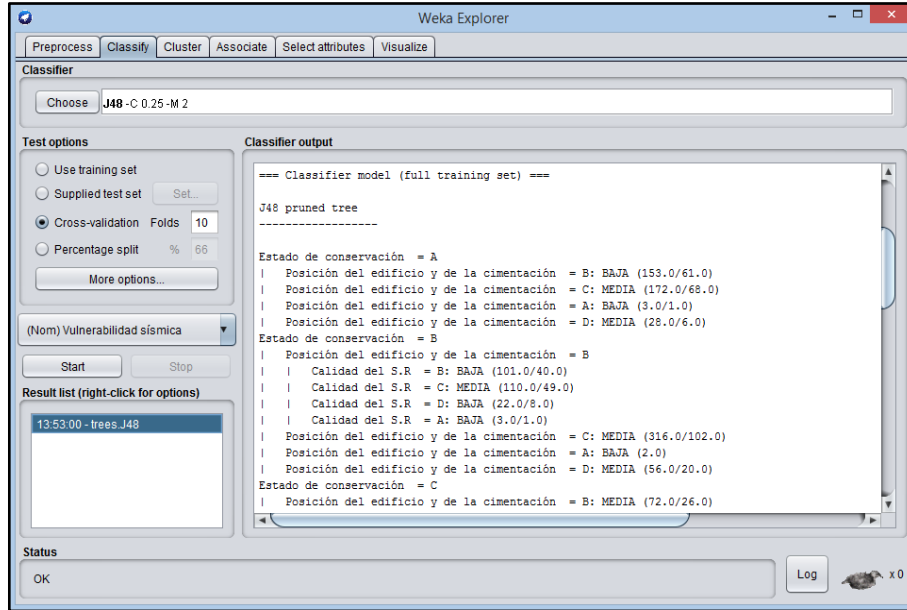


Nota, la figura muestra la clasificación del modelo I con el algoritmo J48 y validación cruzada.

✓ MDL_2

Figura 22

MDL_2-Clasificador J48 validación cruzada

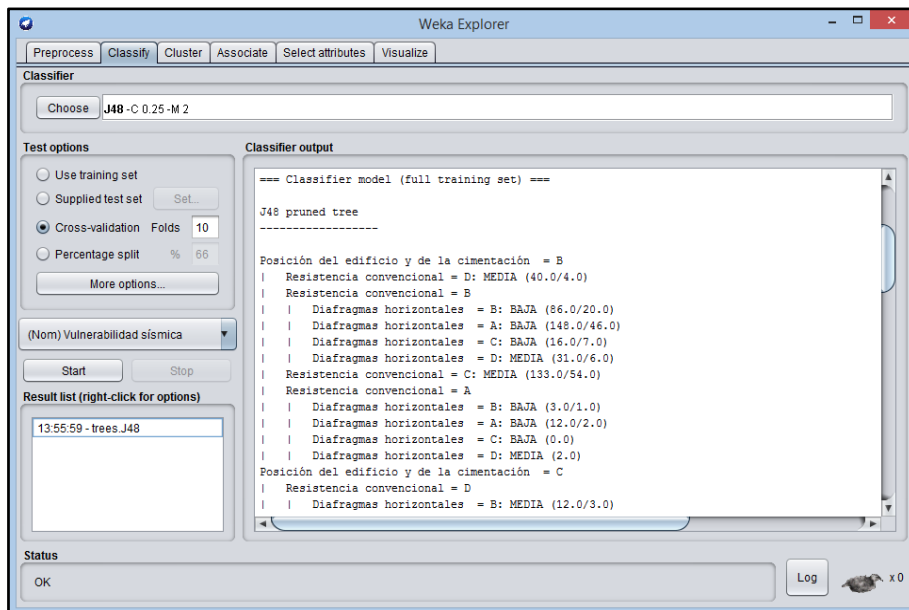


Nota, la figura muestra la clasificación del modelo II con el algoritmo J48 y validación cruzada.

✓ MDL_3

Figura 23

MDL_3-Clasificador J48 validación cruzada

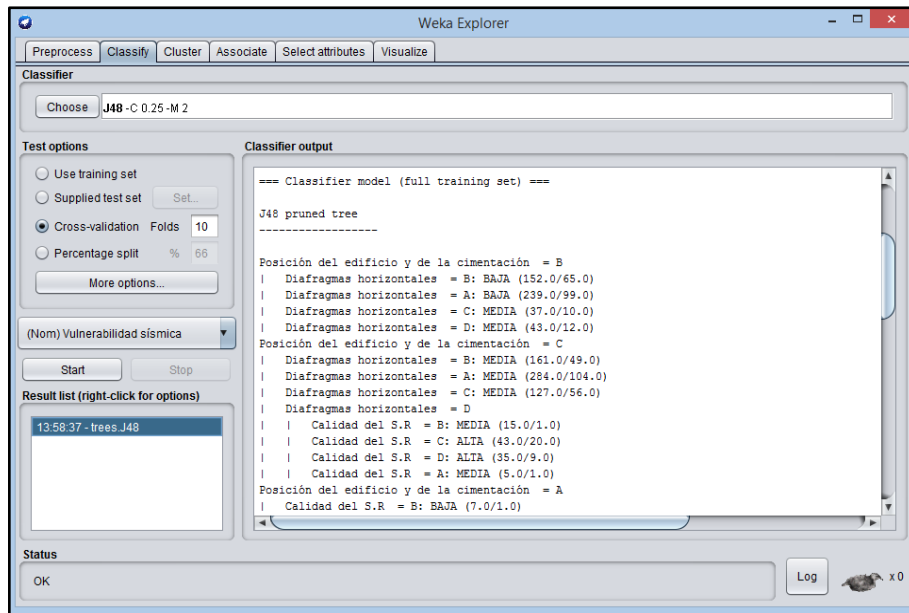


Nota, la figura muestra la clasificación del modelo III con el algoritmo J48 y validación cruzada.

✓ MDL_4

Figura 24

MDL_4-Clasificador J48 validación cruzada

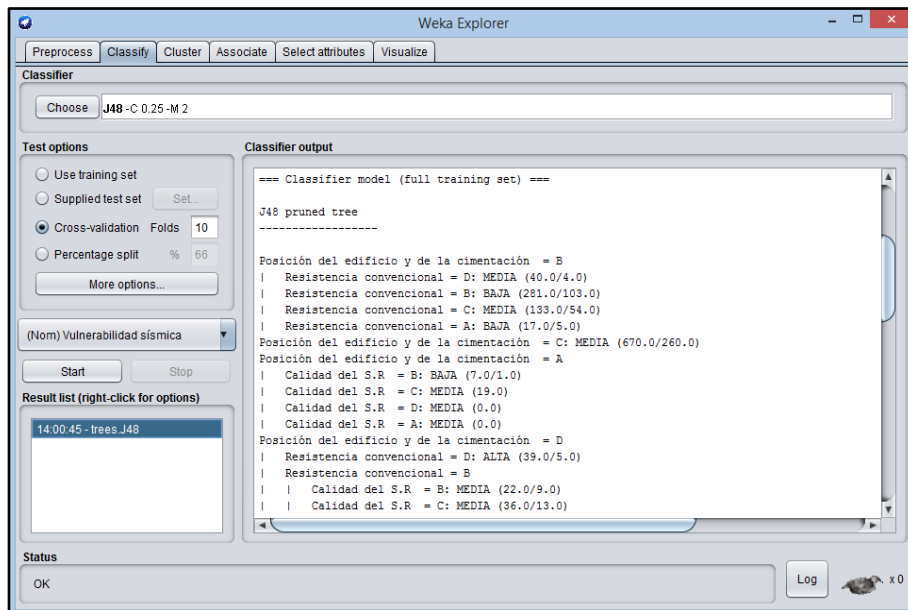


Nota, la figura muestra la clasificación del modelo IV con el algoritmo J48 y validación cruzada.

✓ MDL_5

Figura 25

MDL_5-Clasificador J48 validación cruzada



Nota, la figura muestra la clasificación del modelo V con el algoritmo J48 y validación cruzada.

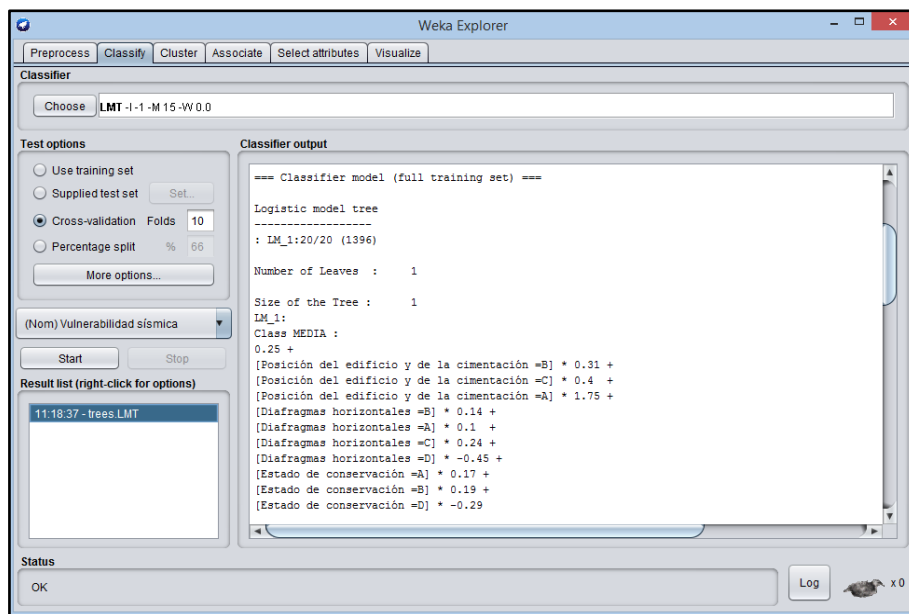
- **LMT**

Algoritmo clasificador para construir '*árboles modelo logísticos*', que son árboles de clasificación con funciones de regresión logística en las hojas. El algoritmo puede tratar con variables objetivo binarias y de clases múltiples, atributos numéricos y nominales y valores perdidos.

- ✓ MDL_1

Figura 26

MDL_1-Clasificador LMT y validación cruzada

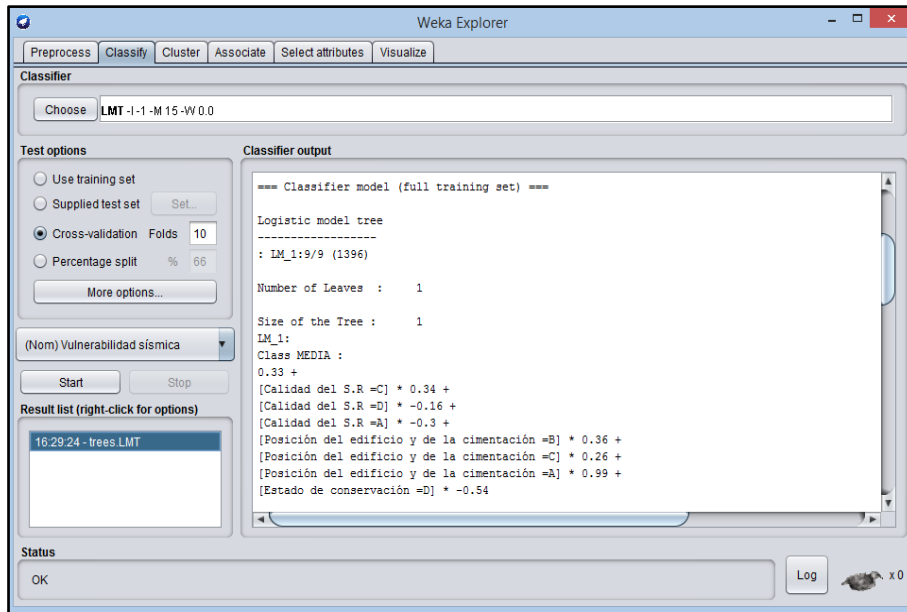


Nota, la figura muestra la clasificación del modelo I con el algoritmo LMT y validación cruzada.

- ✓ MDL_2

Figura 27

MDL_2-Clasificador LMT y validación cruzada

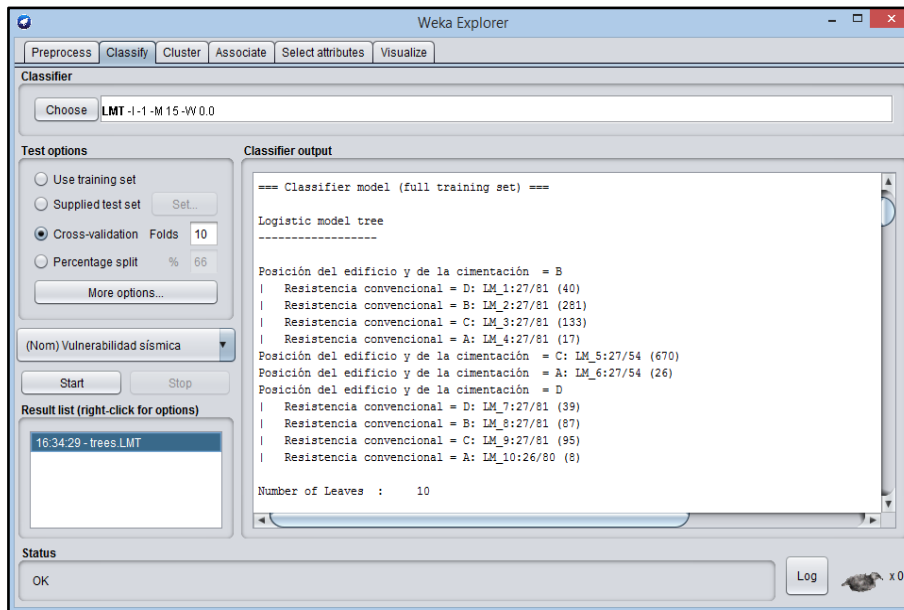


Nota, la figura muestra la clasificación del modelo II con el algoritmo LMT y validación cruzada.

✓ MDL_3

Figura 28

MDL_3-Clasificador LMT y validación cruzada

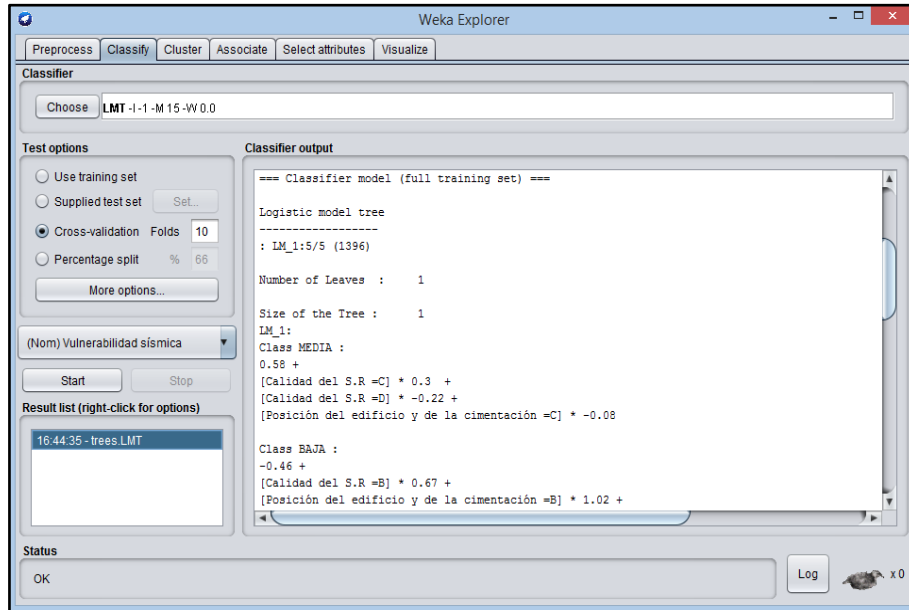


Nota, la figura muestra la clasificación del modelo III con el algoritmo LMT y validación cruzada.

✓ MDL_4

Figura 29

MDL_4-Clasificador LMT y validación cruzada

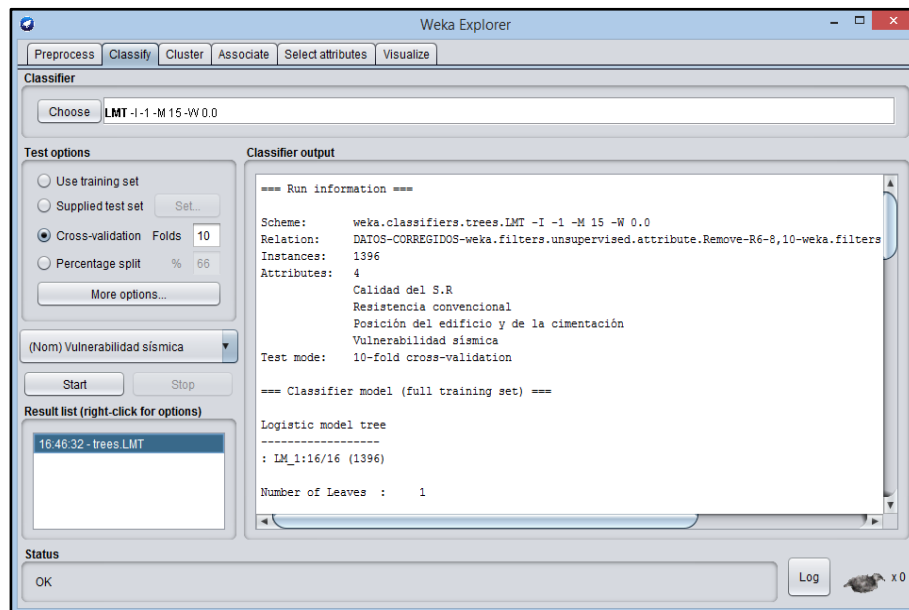


Nota, la figura muestra la clasificación del modelo IV con el algoritmo LMT y validación cruzada.

✓ MDL_5

Figura 30

MDL_5-Clasificador LMT y validación cruzada



Nota, la figura muestra la clasificación del modelo V con el algoritmo LMT y validación cruzada.

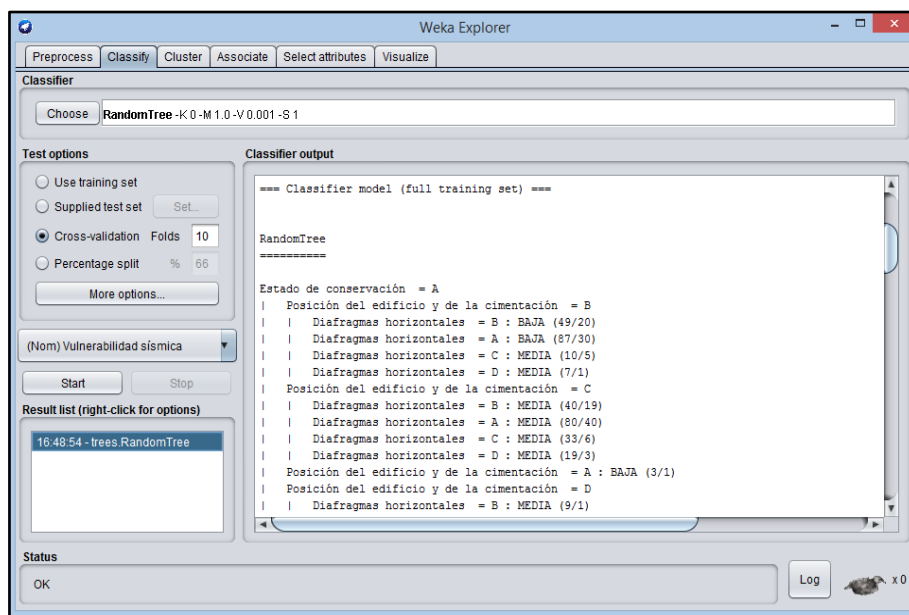
- **Random Tree**

Este algoritmo construye un árbol que considere K atributos elegidos al azar en cada nodo, no realiza podas. También tiene una opción para permitir la estimación de probabilidades de clase (o media objetivo en el caso de regresión) basada en un conjunto de retención (backfitting).

✓ MDL_1

Figura 31

MDL_1-Clasificador Random Tree y validación cruzada

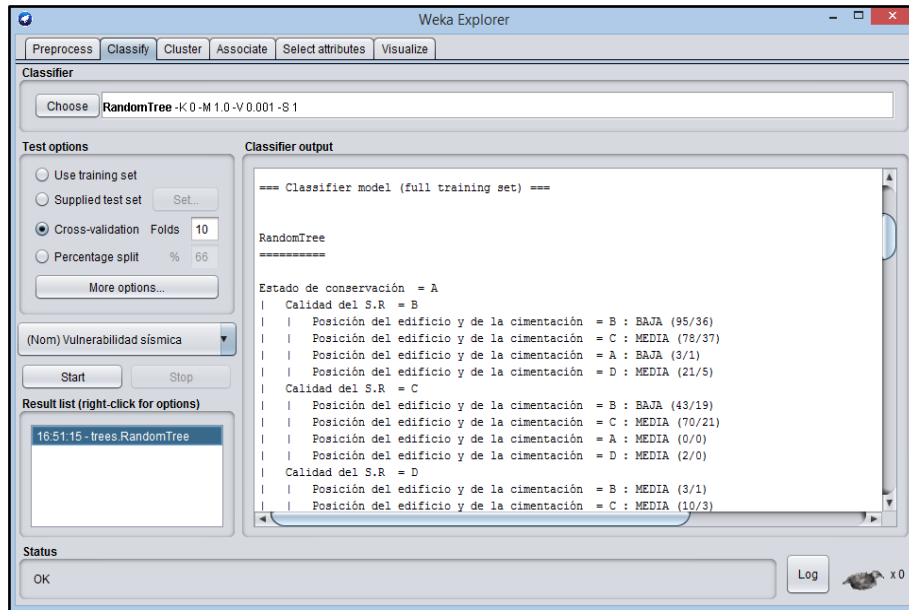


Nota, la figura muestra la clasificación del modelo I con el algoritmo Random Tree y validación cruzada.

✓ MDL_2

Figura 32

MDL_2-Clasificador Random Tree y validación cruzada

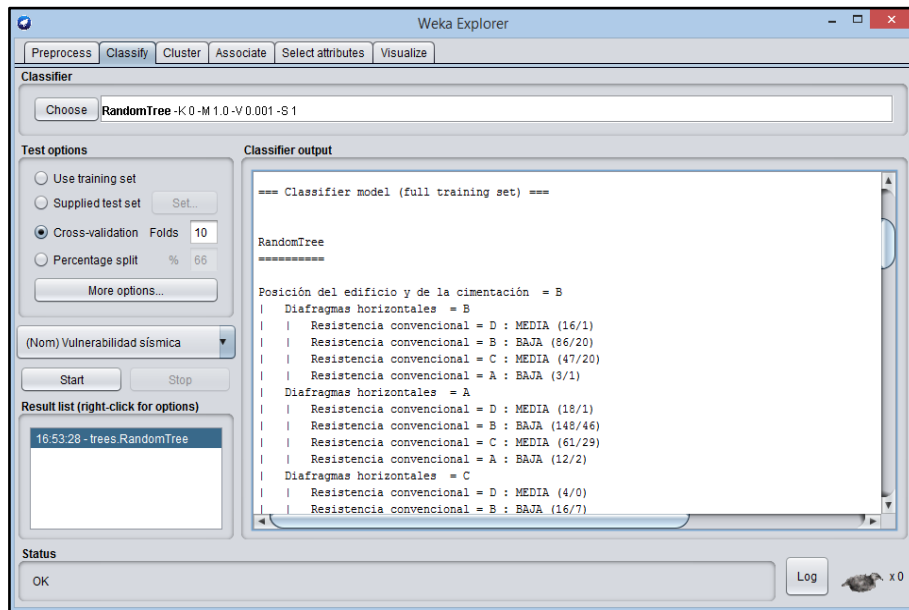


Nota, la figura muestra la clasificación del modelo II con el algoritmo Random Tree y validación cruzada.

✓ MDL_3

Figura 33

MDL_3-Clasificador Random Tree y validación cruzada

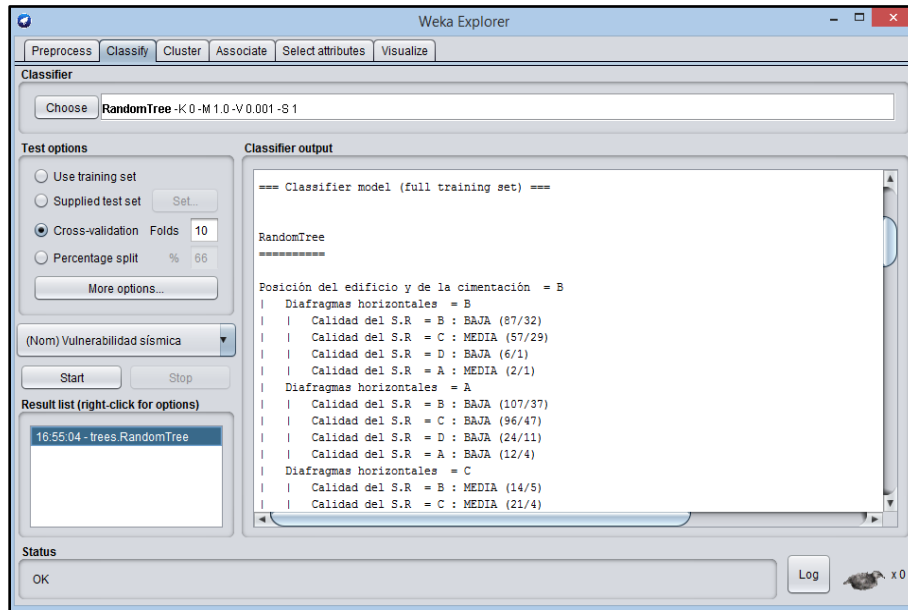


Nota, la figura muestra la clasificación del modelo III con el algoritmo Random Tree y validación cruzada.

✓ MDL_4

Figura 34

MDL_4-Clasificador Random Tree y validación cruzada

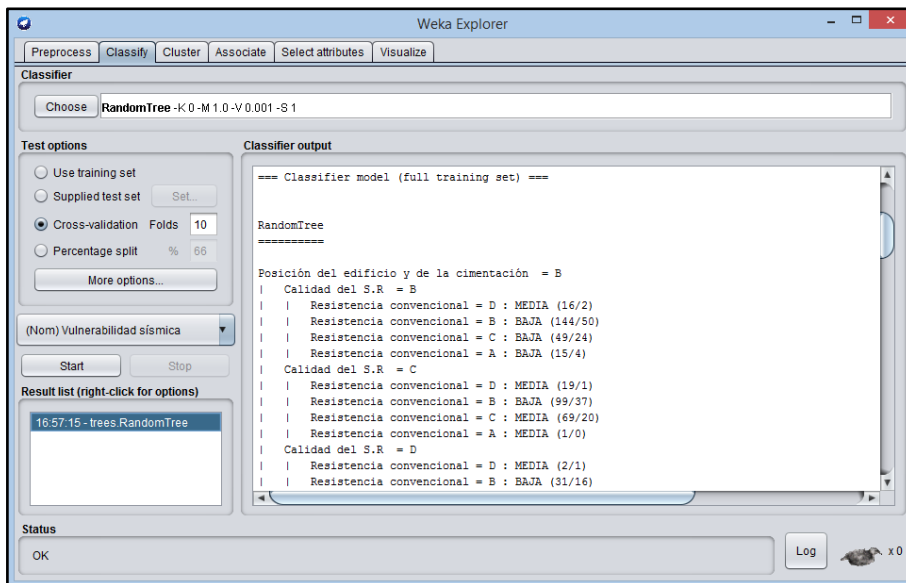


Nota, la figura muestra la clasificación del modelo IV con el algoritmo Random Tree y validación cruzada.

✓ MDL_5

Figura 35

MDL_5-Clasificador Random Tree y validación cruzada



Nota, la figura muestra la clasificación del modelo V con el algoritmo Random Tree y validación cruzada.

3.7.2.4.4. Resumen de modelos obtenidos

Tras haber construido los modelos haciendo uso de los algoritmos que contiene Weka en la sección 3.7.2.4.3, obtenemos un resumen de dicho modelo.

Los diferentes modelos obtenidos tuvieron la estructura:

- Número de instancias clasificadas correctamente.
- Numero de instancias clasificadas incorrectamente.
- El estadístico Kappa es un índice que compara el nivel de coincidencia entre varios expertos con el nivel de coincidencia que se podría dar por casualidad.
- Error cuadrático medio: es un estimador que mide el promedio de los errores al cuadrado, es decir, la diferencia entre el estimador y lo que se estima al cuadrado.
- Error absoluto: La diferencia entre el valor obtenido y su valor real.
- Error absoluto relativo: La fórmula $[(xi - xt)/xt] \times 100$, donde $xi - xt$ es el error absoluto y xt es el valor real.
- Número total de instancias.

En el Anexo I se detalló el resumen de los modelos de minería de datos.

3.7.2.5. Etapa de intervención/evaluación de datos.

3.7.2.5.1. Predicción de los modelos

Después de haber entrenado los cinco modelos, con la base de datos optima, se escogió tres modelos para la validación, ya que estos presentaron mayor porcentaje de predicción durante el entrenamiento, siendo el modelo MDL_1, MDL_4 y MDL_5.

Para esta validación se utilizaron los datos recopilados por Ramos (2020), en su tesis “Vulnerabilidad Sísmica de las viviendas de Albañilería confinada ubicadas en el sector Pueblo Libre en la ciudad de Jaén, Cajamarca-2020”, donde determinó la vulnerabilidad sísmica del sector Pueblo Libre. Se basó en sus resultados, para realizar la comparación de lo obtenido

según el método Benedetti Petrini y los modelos de minería de datos propuestos. En el Anexo II se detallan la predicción de los modelos de minería de datos.

3.7.2.5.2. Evaluación de modelos matriz de confusión

Luego de haber estimado la vulnerabilidad sísmica, los diferentes modelos fueron evaluados a través de la matriz de confusión, el cual permitió determinar que algoritmo tuvieron menor error durante la predicción. Se presentan los principales criterios de predicción:

Tabla 7

Matriz de confusión

		Detectado	
		Positivo	Negativo
Actual	Positivo	Positivo Verdadero (<i>PV</i>)	Negativo Falso (<i>NF</i>)
	Negativo	Positivo Falso (<i>PF</i>)	Negativo Verdadero (<i>NV</i>)

Nota, matriz de confusión para la evaluación de los modelos. Obtenido de Quiñones & Carrasco (2020).

Instancias Correctamente Clasificadas (*ICC*): Es la proporción de instancias correctamente clasificadas, dividida por el total de instancias.

$$ICC = \frac{PV + NV}{PV + NV + PF + NF} \quad (7)$$

Instancias Incorrectamente Clasificadas (*IIC*): Es la proporción de instancias incorrectamente clasificadas dividida por el total de instancias.

$$IIC = \frac{PF + NF}{PV + NV + PF + NF} \quad (8)$$

Tasa TP: Es la tasa de casos positivos que están correctamente identificados o proporción de casos que la prueba declara positivos y que son positivos.

$$Tasa TP = \frac{PV}{PV + NF} \quad (9)$$

Tasa FP: Es la tasa de casos negativos que fueron erróneamente clasificados como positivos o proporción de casos que la prueba declara positivos y que en realidad son negativos.

$$Tasa FP = \frac{PF}{PF + NV} \quad (10)$$

Precisión: Es la proporción del total de predicciones positivas correctas sobre el total de instancias clasificadas como esa clase.

$$Precisión = \frac{PV}{PV + PF} \quad (11)$$

F-Mesure: Es la media armónica de la Precisión con el Recall.

$$FMesure = 2 * \frac{Precisión * Recall}{Precisión + Recall} \quad (12)$$

Accuracy: Es la proporción de las predicciones correctas sobre el total de instancias.

$$Accuracy = \frac{PV + NV}{PV + NV + PF + NF} \quad (13)$$

Sensibilidad: Es la proporción de las predicciones positivas sobre el total de instancias que pertenecen a esa clase.

$$Sensibilidad = \frac{PV}{PV + NF} \quad (14)$$

Especificidad: Es la proporción de las predicciones negativas sobre el total de instancias que pertenecen a esa clase.

$$Especificidad = \frac{NV}{NV + PF} \quad (15)$$

3.7.3. Mapa de zonificación, según los niveles de vulnerabilidad sísmica del sector de estudio.

Se hizo uso del Sistema de Información Geográfica (SIG), por su capacidad de almacenar, manejar y analizar datos geográficos georreferenciados, los cuales se pueden representar de manera gráfica en mapas temáticos (Ludeña Ocampo, 2005).

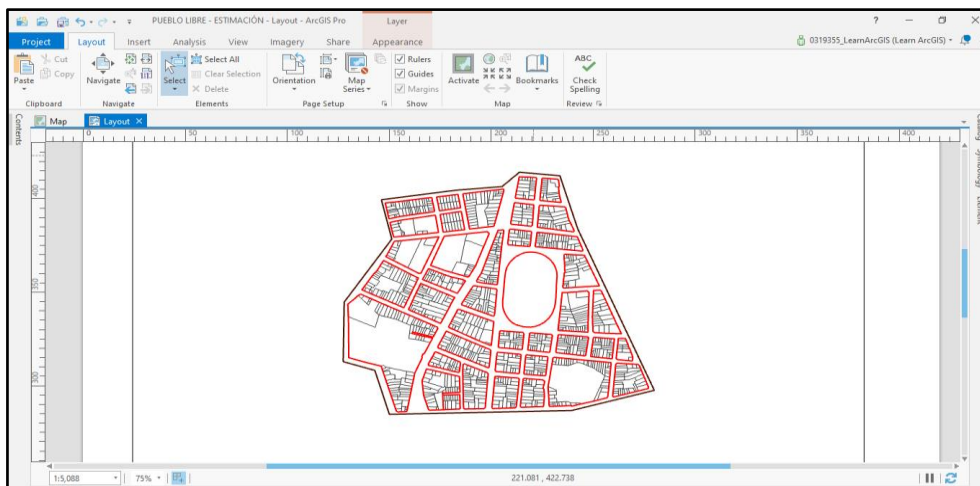
3.7.3.1. Elaboración de la cartografía y análisis espacial

Los datos de carácter espacial y no espacial se recolectaron de la Sub Gerencia de Desarrollo Urbano y Catastro – Municipalidad Provincial de Jaén y de investigaciones en el sector Pueblo Libre – Jaén, la información obtenida fue consolidada y sistematizada acorde a los requerimientos que permiten analizar la vulnerabilidad sísmica del sector en estudio. Se utilizó el sistema de referencia espacial durante todo el proceso al Datum WGS – 84, proyección UTM en la Zona 17.

Utilizando “ArcCatalog” se conectó al directorio la carpeta que contenía la información del sector Pueblo Libre – Jaén y mediante “ArcMap” visualizamos los polígonos del contorno, manzanas y lotes del sector en estudio.

Figura 36

Visualización de la información espacial del sector Pueblo Libre en ArcMap

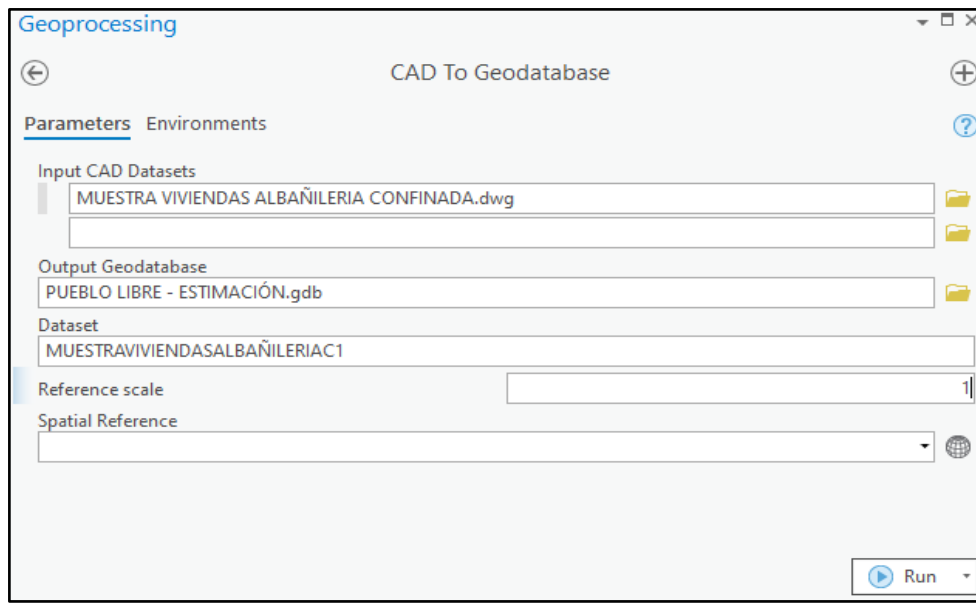


Nota, la figura muestra el plano en planta del sector Pueblo Libre en Arc Map.

Las viviendas de albañilería confinada en estudio fueron sesenta y siete, representadas a través de polígonos georreferenciados en Computer Aided Design (CAD), y procesados en ArcMap con la herramienta “Conversion” en la opción “CAD to Geodatabase”.

Figura 37

Ingreso de la muestra en estudio de viviendas de albañilería confinada a ArcMap



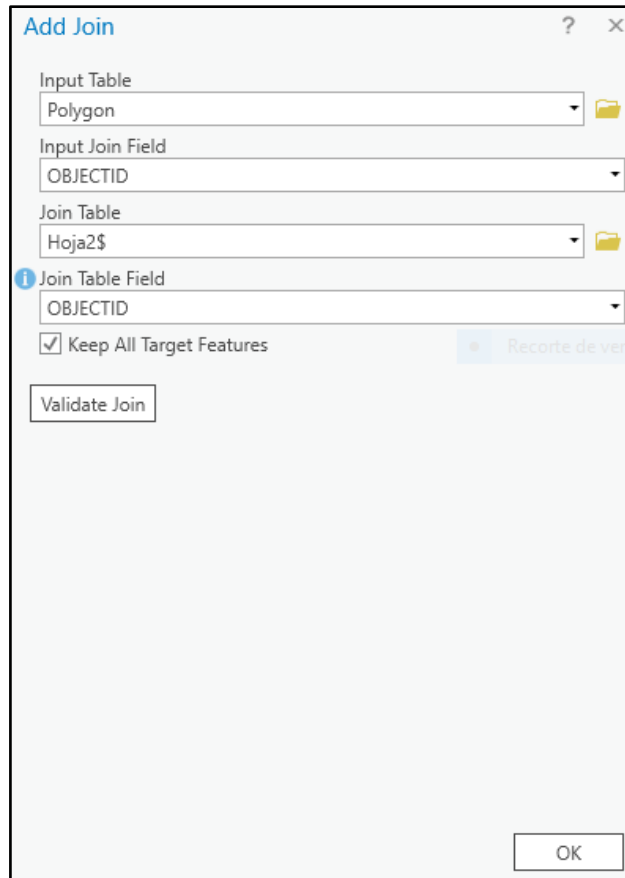
Nota, la figura muestra el proceso de ingreso de las viviendas en estudio a Arc Map.

3.7.3.1.1. Construcción de información temática

Identificado la tabla de atributos de los polígonos que se generaron en la conversión a “Geodatabase”, en la columna “OBJECTID” nos basamos en esta para generar un archivo de Excel con extensión “.xls” en el que se agregó información del perímetro, área, código que hace referencia al número de manzana y su respectivo lote, y vulnerabilidad sísmica asociada a un valor (de acuerdo a lo estimado por el algoritmo RandomTree en el MDL_5). El archivo con extensión “.xls” en específico la “Hoja2\$” se conectó a la tabla de atributos, utilizando la opción “Join Data” y se verifico que la base de unión, sea la columna “OBJECTID”.

Figura 38

Conexión de información a la tabla de atributos de los polígonos

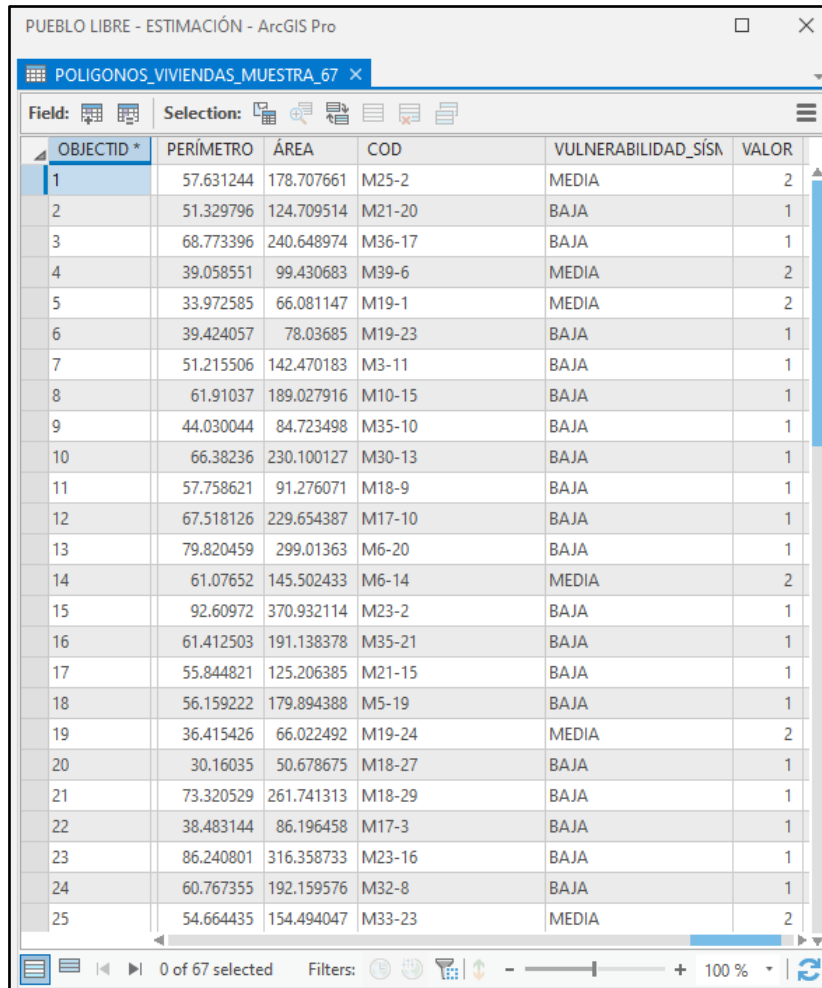


Nota, la figura muestra el proceso de conexión de información a la tabla de atributos.

Asociada la información espacial y la información temática, se relacionó cada uno de los sesenta y siete polígonos que representan una vivienda, con sus características entre ellas la vulnerabilidad sísmica.

Figura 39

Tabla de atributos, donde se asoció la información espacial con la información temática



OBJECTID *	PERÍMETRO	ÁREA	COD	VULNERABILIDAD_SÍSN	VALOR
1	57.631244	178.707661	M25-2	MEDIA	2
2	51.329796	124.709514	M21-20	BAJA	1
3	68.773396	240.648974	M36-17	BAJA	1
4	39.058551	99.430683	M39-6	MEDIA	2
5	33.972585	66.081147	M19-1	MEDIA	2
6	39.424057	78.03685	M19-23	BAJA	1
7	51.215506	142.470183	M3-11	BAJA	1
8	61.91037	189.027916	M10-15	BAJA	1
9	44.030044	84.723498	M35-10	BAJA	1
10	66.38236	230.100127	M30-13	BAJA	1
11	57.758621	91.276071	M18-9	BAJA	1
12	67.518126	229.654387	M17-10	BAJA	1
13	79.820459	299.01363	M6-20	BAJA	1
14	61.07652	145.502433	M6-14	MEDIA	2
15	92.60972	370.932114	M23-2	BAJA	1
16	61.412503	191.138378	M35-21	BAJA	1
17	55.844821	125.206385	M21-15	BAJA	1
18	56.159222	179.894388	M5-19	BAJA	1
19	36.415426	66.022492	M19-24	MEDIA	2
20	30.16035	50.678675	M18-27	BAJA	1
21	73.320529	261.741313	M18-29	BAJA	1
22	38.483144	86.196458	M17-3	BAJA	1
23	86.240801	316.358733	M23-16	BAJA	1
24	60.767355	192.159576	M32-8	BAJA	1
25	54.664435	154.494047	M33-23	MEDIA	2

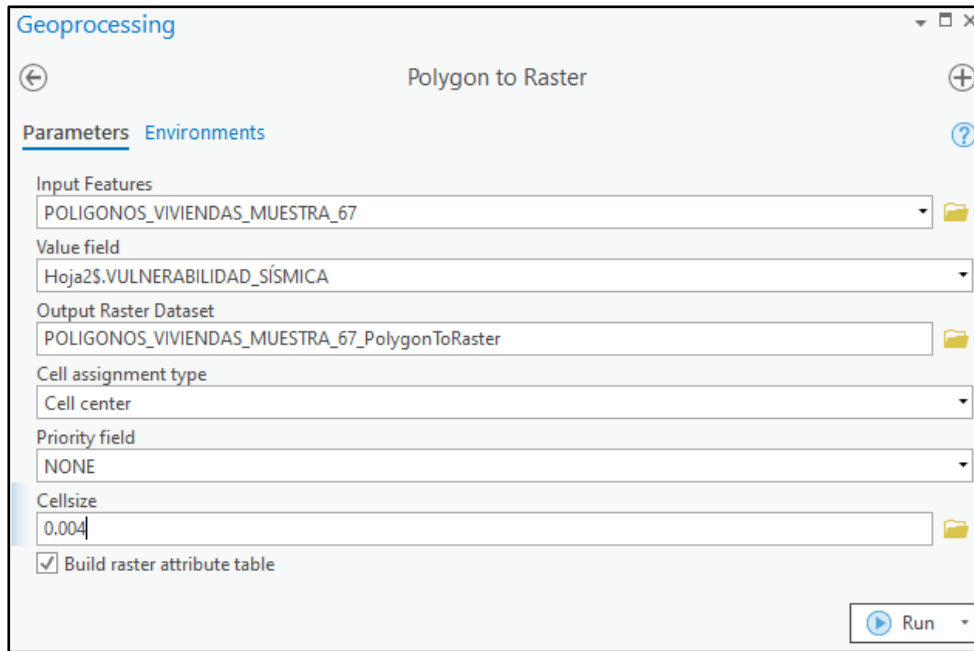
Nota, la figura muestra la asociación de la información espacial con la temática.

3.7.3.1.2. Construcción de mapa temático

Para la elaboración del mapa, se utilizó la herramienta “Conversion” en la opción “To raster” y seleccionamos “Polygon to Raster”, se configuró la entrada “POLIGONOS_VIVIENDAS_MUESTRA_67” y se evaluó de acuerdo a la columna “VALOR” la vulnerabilidad sísmica para cada polígono.

Figura 40

Conversión de los polígonos a ráster para elaborar el mapa de zonificación de acuerdo a la vulnerabilidad sísmica estimada para cada vivienda

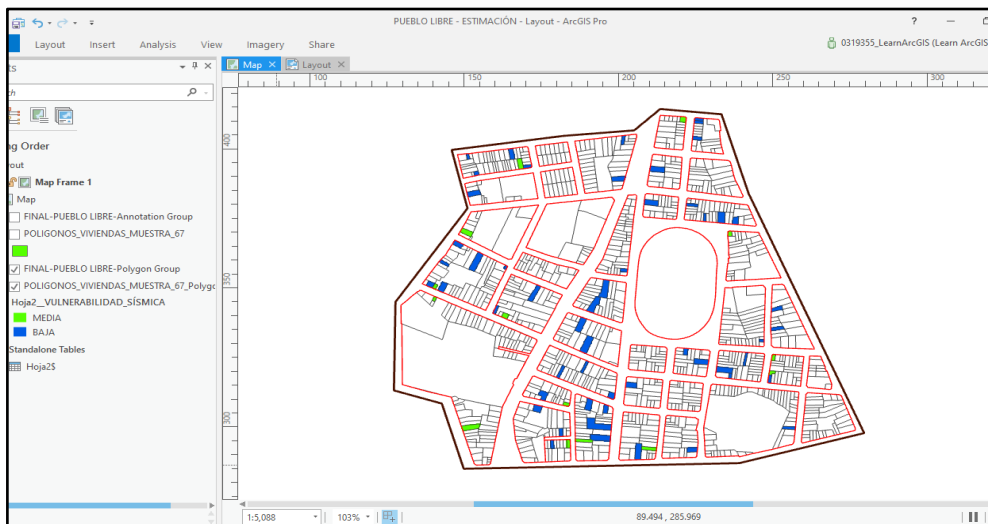


Nota, la figura muestra el proceso de conversión de los polígonos a ráster.

Habiendo utilizado la información estimada por el algoritmo Random Tree del MDL_5, la cual clasificó la vulnerabilidad sísmica de las sesenta y siete viviendas de albañilería confinada, como baja y media. Se representó de color azul y verde los niveles de vulnerabilidad ver Figura 41.

Figura 41

Análisis en base de la vulnerabilidad sísmica



Nota, la figura muestra la asignación de la vulnerabilidad a cada vivienda.

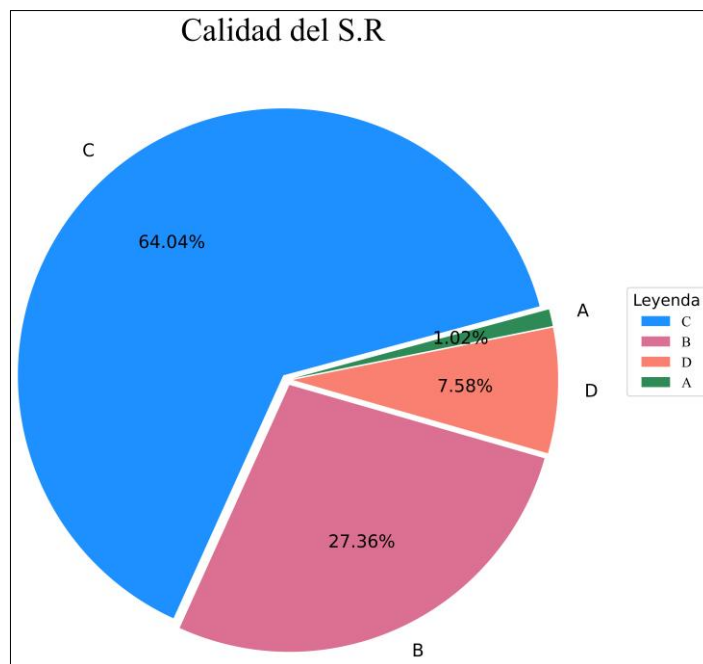
IV. RESULTADOS

4.1. Matriz de datos con sus respectivas variables que influyen en la vulnerabilidad sísmica

La matriz de datos estuvo conformada por 3 827 registros de viviendas de albañilería confinada y doce variables recopiladas de tesis de varios repositorios del Perú. Las variables se basan en el método de índice de vulnerabilidad, calificando a cada atributo en A, B, C y D, con un nivel de vulnerabilidad baja, media y alta. En las figuras 42-57, se representan todas las variables mediante gráficas circulares y barras.

Figura 42

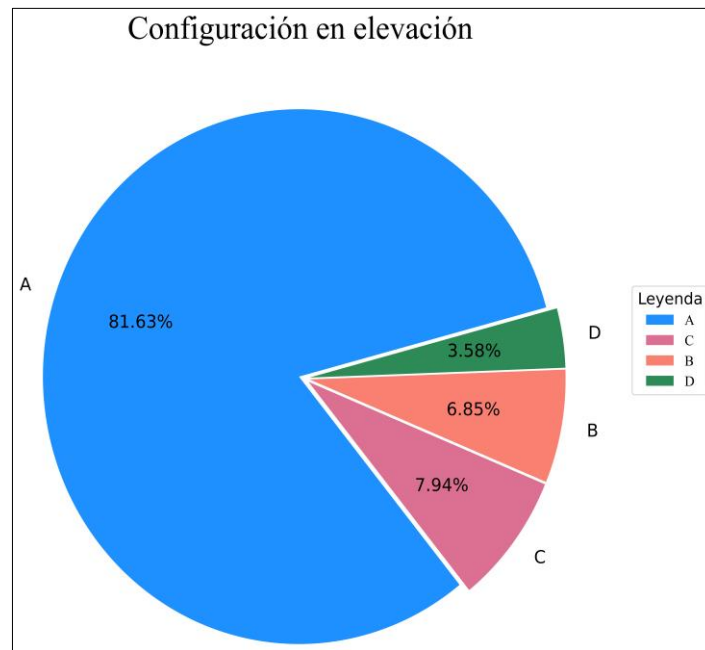
Calidad del sistema resistente



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.2, donde la mayor parte de las viviendas no presenta ladrillos de buena calidad y mortero de entre 10 mm a 12 mm de espesor.

Figura 43

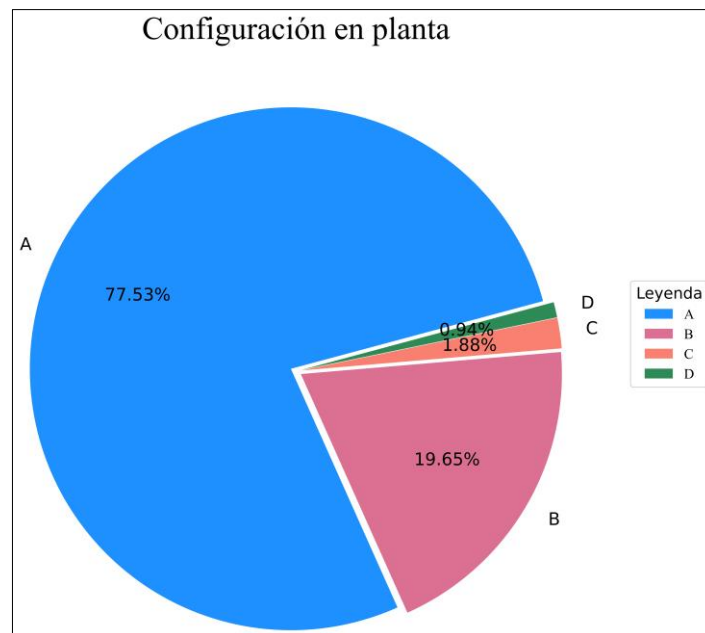
Configuración en elevación



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.7, donde la mayor parte de las viviendas cumplen con la variación de áreas respectivas menores al 10%.

Figura 44

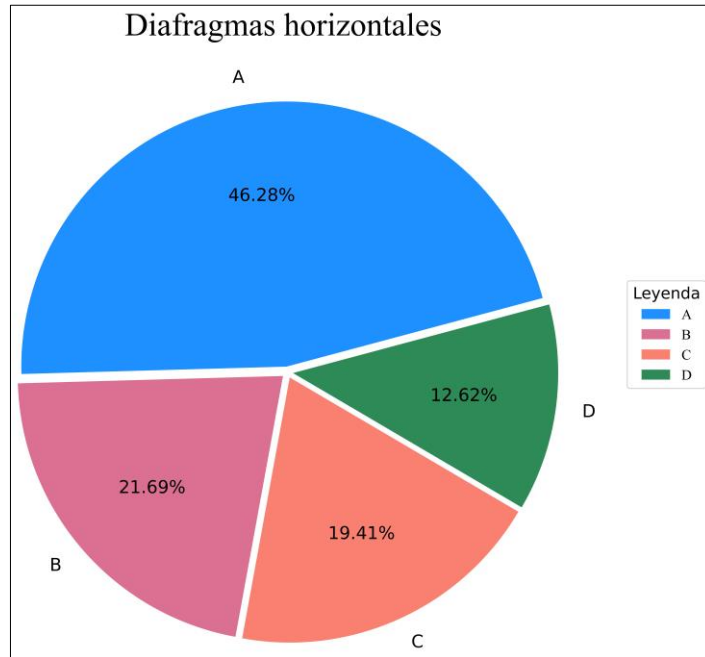
Configuración en planta



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.6, donde la mayor parte de las viviendas cumplen con la irregularidad en planta siendo esta menor al 0.10.

Figura 45

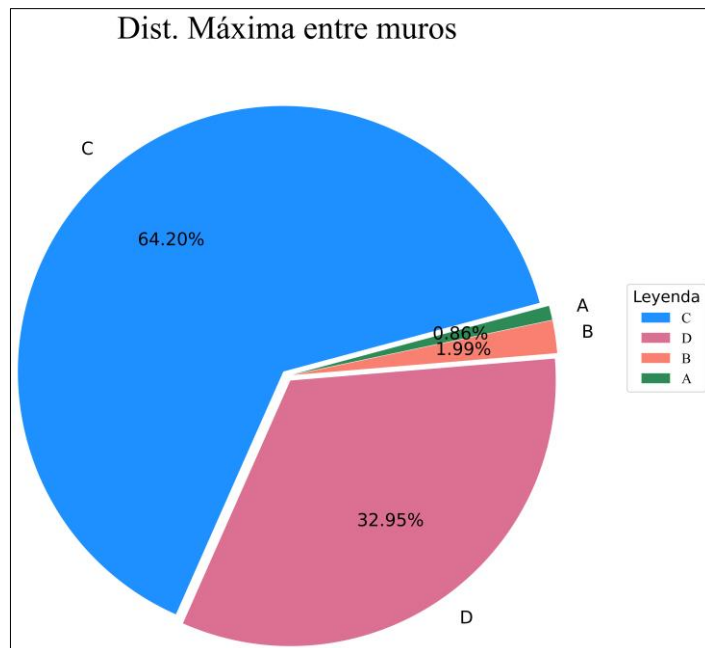
Diagramas horizontales



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.5, donde la mayor parte de las viviendas tienen ausencia de desnivel y la deformabilidad del diagrama es despreciable, entre otros.

Figura 46

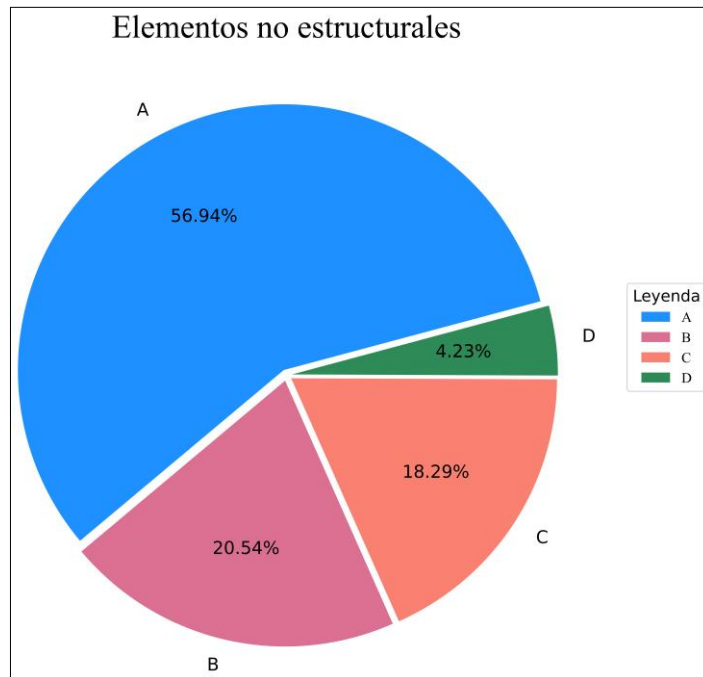
Distancia máxima entre muros



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.8, donde la mayor parte de las viviendas tienen una relación entre el espaciamiento y la sección transversal de los muros de 18 a 25.

Figura 47

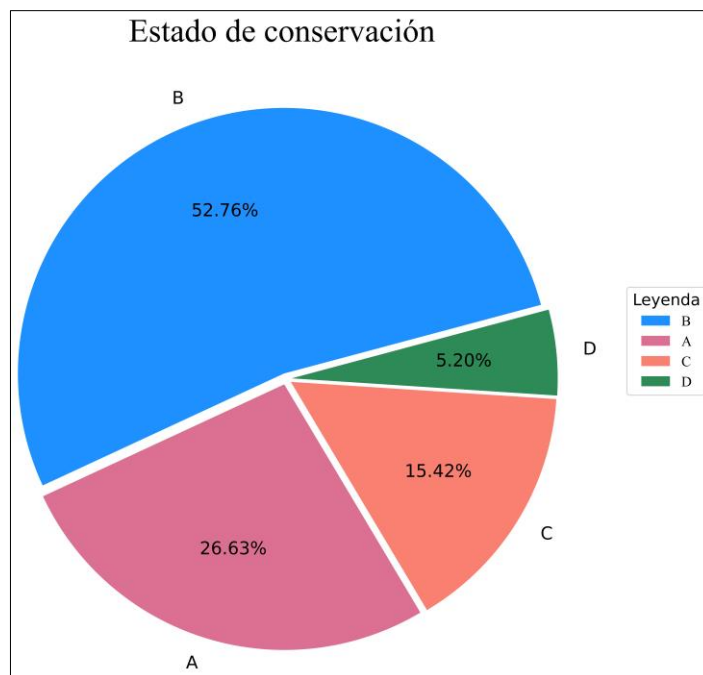
Elementos no estructurales



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.10, donde la mayor parte de las viviendas no incluyen elementos no estructurales, mal conectados al sistema resistente.

Figura 48

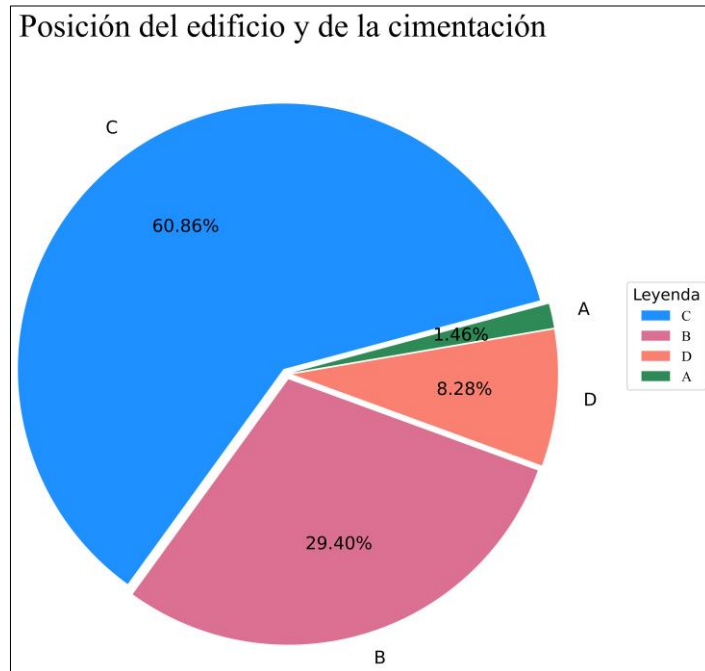
Estado de conservación



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.11, donde la mayor parte de las viviendas no presentan fisuras, pero cuyos componentes estructurales están levemente deteriorados.

Figura 49

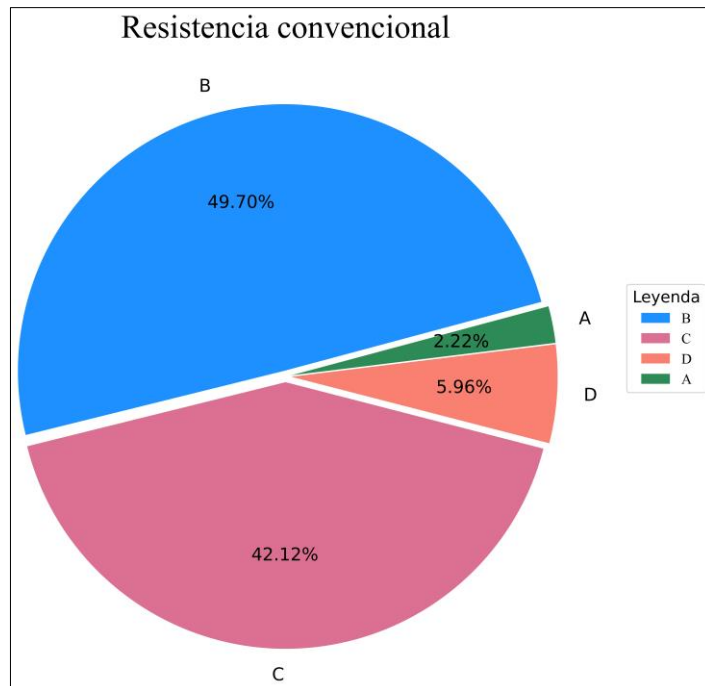
Posición del edificio y de la cimentación



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.4., donde la mayor parte de las viviendas están cimentadas sobre suelos intermedios y flexibles según la norma E.070.

Figura 50

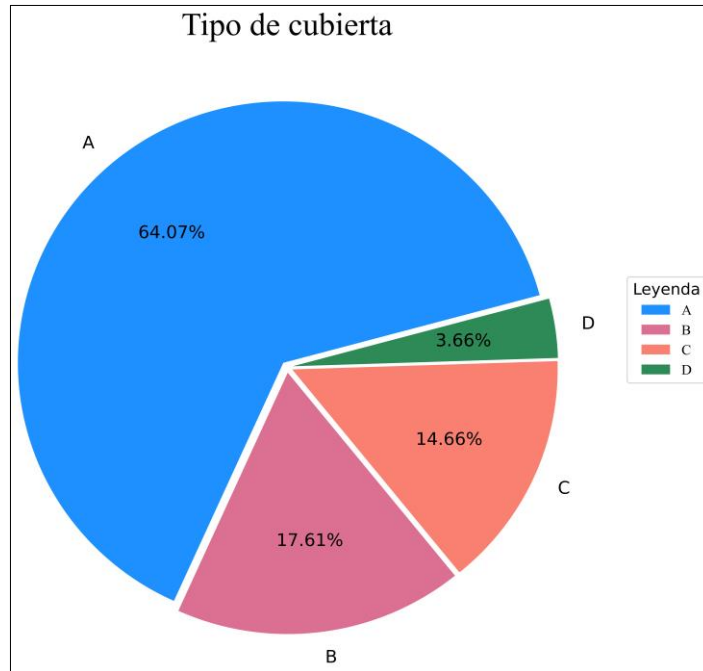
Resistencia convencional



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.3, donde la mayor parte de las viviendas tiene una demanda de ductilidad mayor e igual a 1.00 y menor a 1.50.

Figura 51

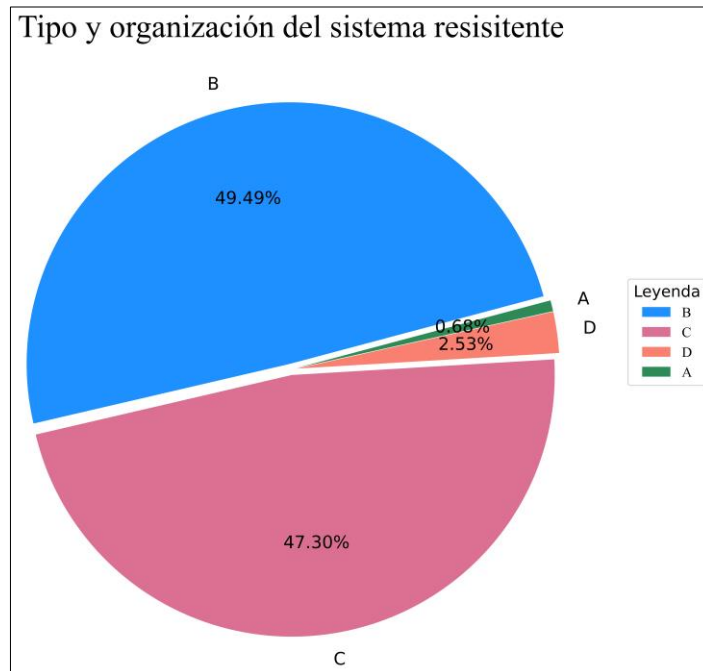
Tipo de cubierta



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.9, donde la mayor parte de las viviendas tiene cubierta estable, plana y arriostramiento en las vigas.

Figura 52

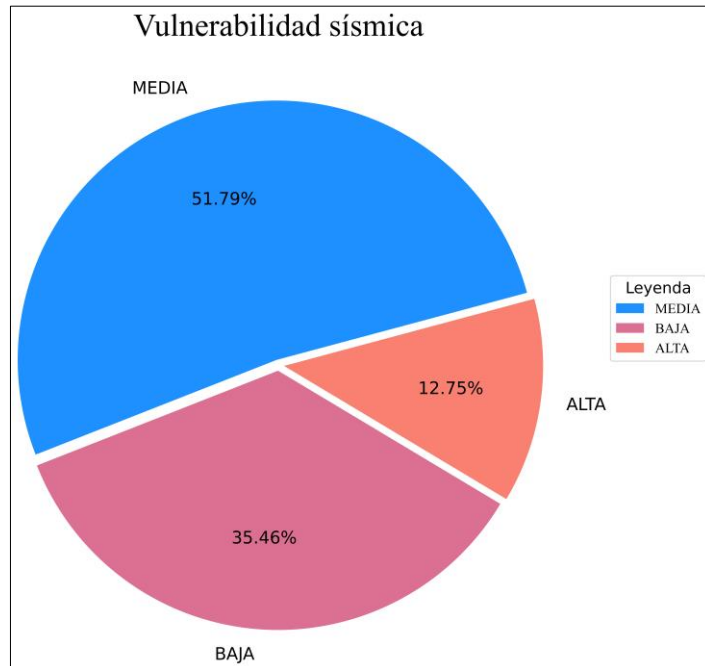
Tipo y organización del sistema resistente



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.1, donde la mayor parte de las viviendas presentan vigas y columnas que confinan parcialmente con los muros.

Figura 53

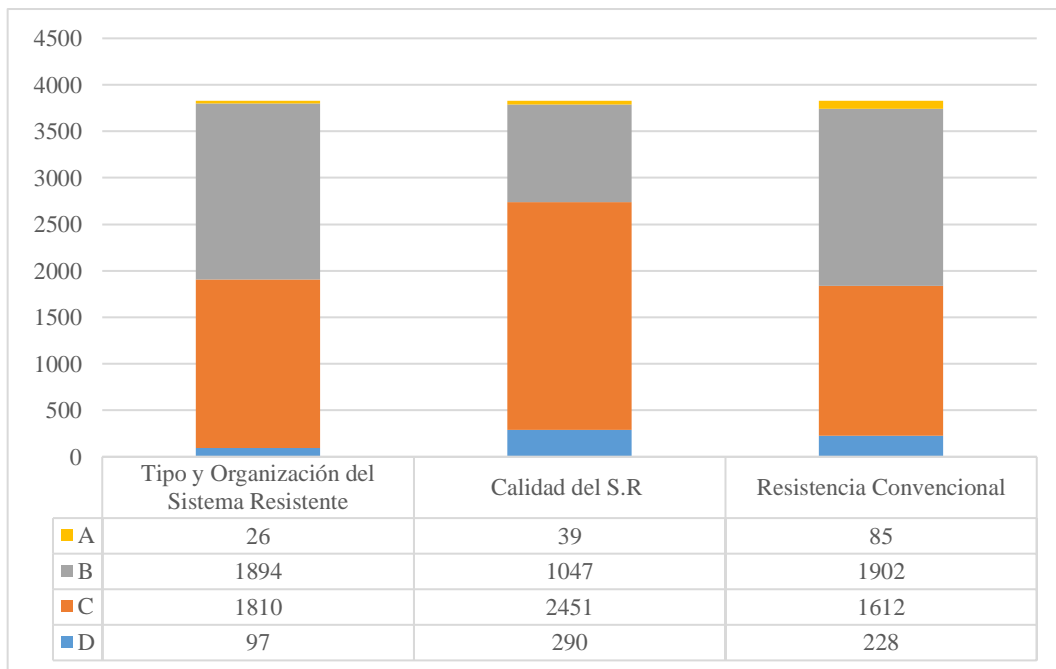
Vulnerabilidad sísmica



Nota, la figura muestra los porcentajes de la calificación de la variable según el ítem 3.7.1.1.12, donde la mayor parte de las viviendas presentan vulnerabilidad media.

Figura 54

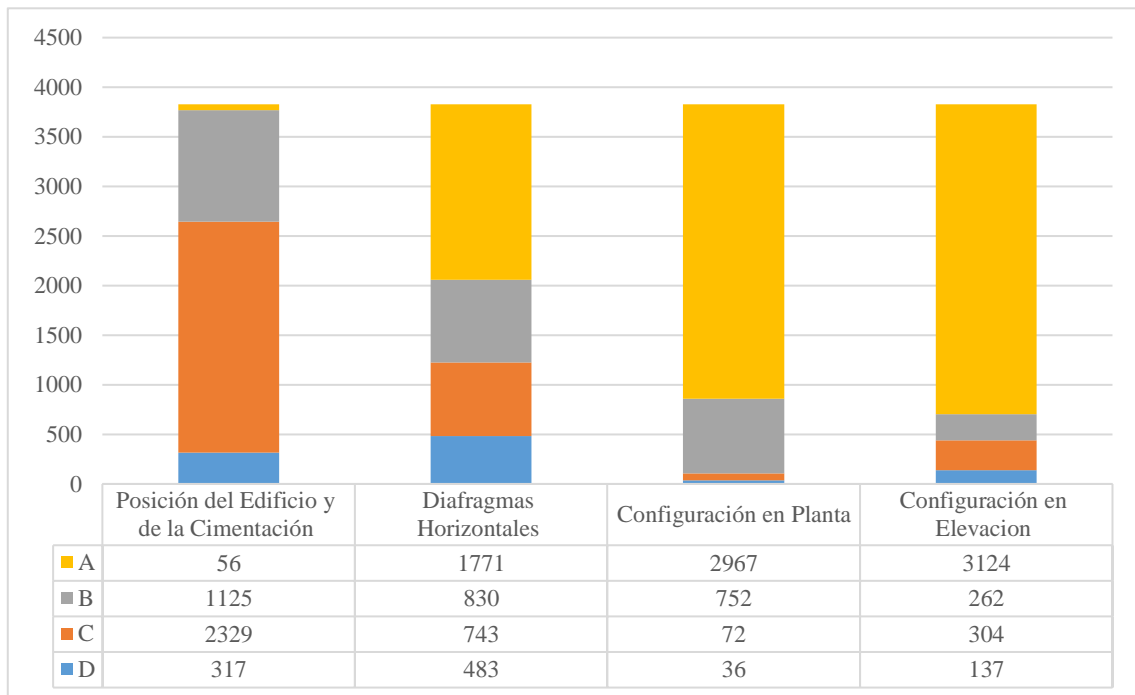
Gráfico de barras de las variables VR_2, VR_6, VR_7



Nota, la figura muestra las cifras de la calificación de la variable 2, 6 y 7, donde la mayor parte de las viviendas tiene una calificación predominante de B.

Figura 56

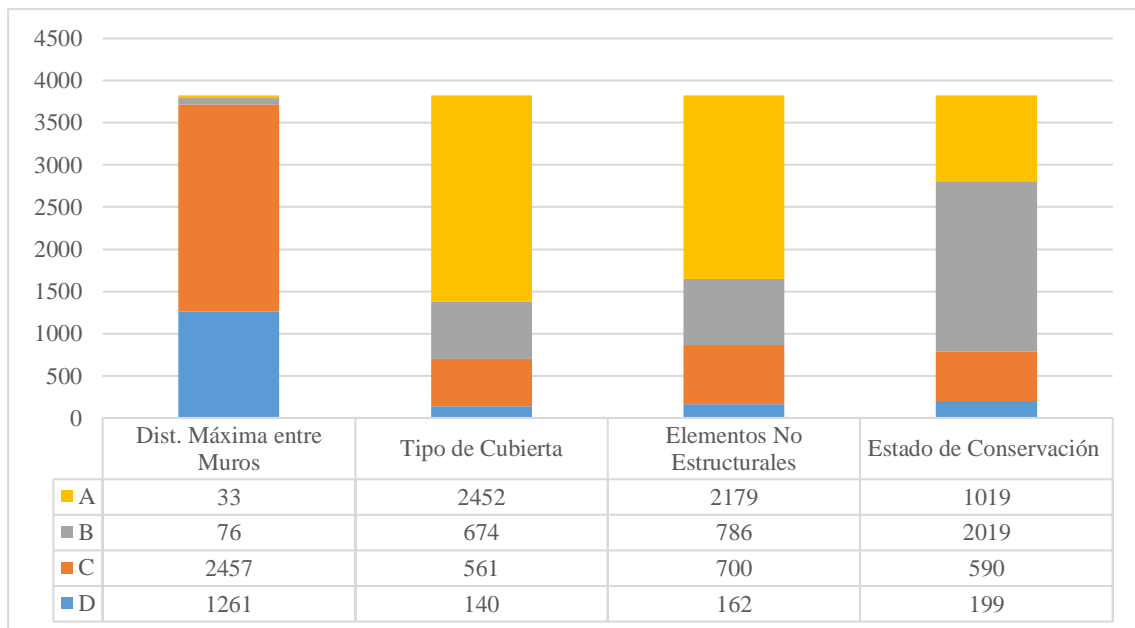
Gráfico de barras de las variables VR_4, VR_5, VR_11, VR_8



Nota, la figura muestra las cifras de la calificación de la variable 4, 5, 8 y 11, donde la mayor parte de las viviendas tiene una calificación predominante de A.

Figura 55

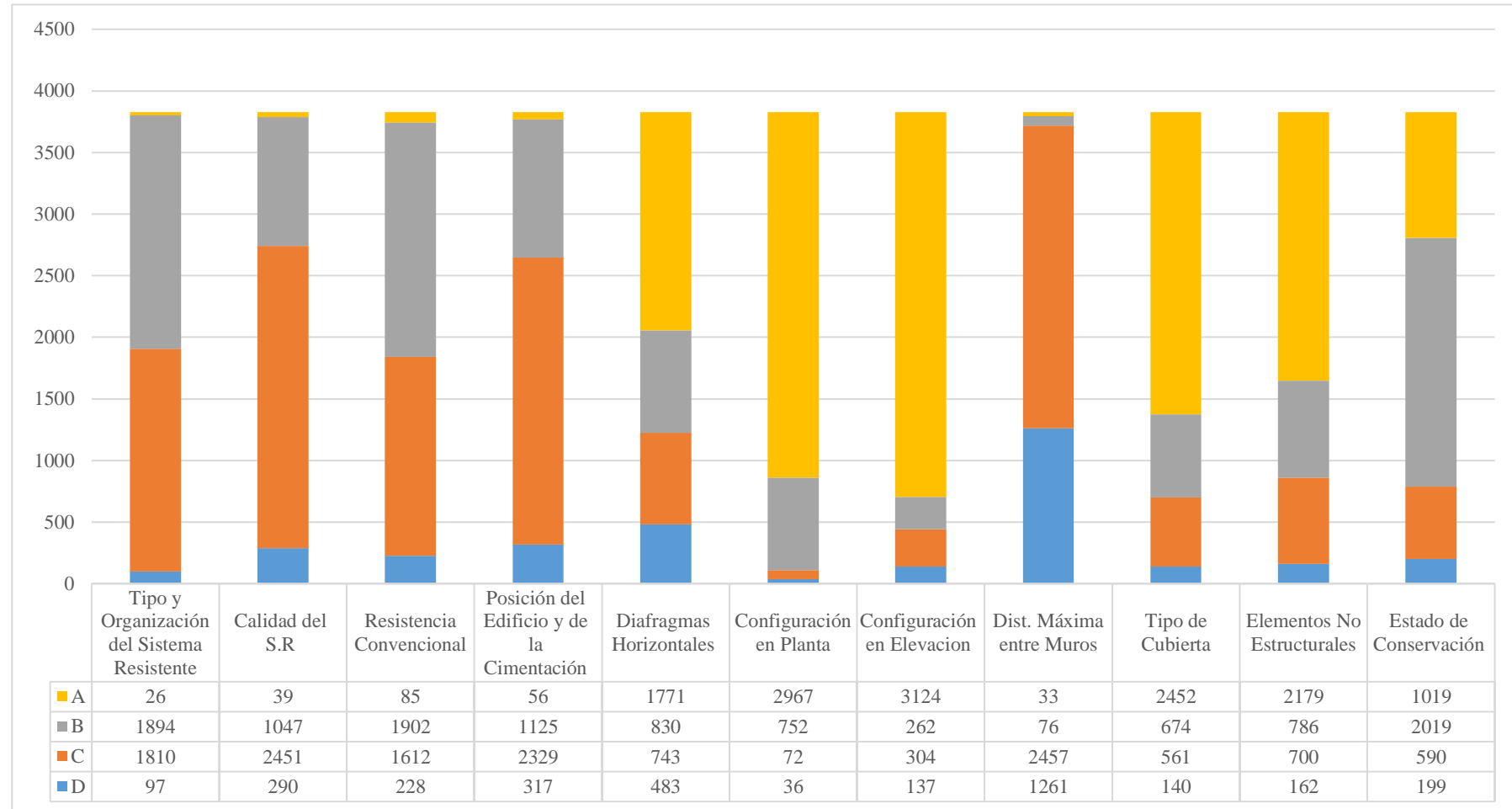
Gráfico de barras de las variables VR_9, VR_3, VR_10, VR_1



Nota, la figura muestra las cifras de la calificación de la variable 1, 3, 9 y 10, donde la mayor parte de las viviendas tiene una calificación predominante de A y B.

Figura 57

Gráfico de barras de las 11 variables de las cuales depende la vulnerabilidad sísmica y que conforman la base de datos



Nota, la figura muestra todas las variables de la base de datos y las cifras de la calificación según el ítem 3.7.1.1, donde la mayor parte de las viviendas tiene una calificación predominante de A, B y C.

4.2. Modelación de la vulnerabilidad sísmica con técnicas de minería de datos.

Los resultados obtenidos de la simulación con el test de validación cruzada, del MDL_1 con los algoritmos de minería de datos como Decision Stump, Hoeffding Tree, J48, LMT y Random Tree, se representa en la Tabla 8, donde se obtuvo el número de hojas, el tamaño del árbol y la precisión de aciertos en porcentaje.

Tabla 8

Resultados del modelo MDL_1, con sus respectivos algoritmos

Clasificador	Tipo de test	Num. hojas	Tam. árbol	Precisión (% aciertos)
Decision Stump	Validación cruzada	1	1	53.30
Hoeffding Tree	Validación cruzada	1	1	67.77
J48	Validación cruzada	28	37	67.34
LMT	Validación cruzada	1	1	69.05
Random Tree	Validación cruzada	1	73	67.84

Nota, esta tabla muestra los algoritmos, el tipo de test, el número de hojas, el tamaño del árbol y la precisión del Modelo I.

Los resultados obtenidos de la simulación con el test de validación cruzada, del MDL_2 con los algoritmos de minería de datos como Decision Stump, Hoeffding Tree, J48, LMT y Random Tree se representa en la Tabla 9, donde se obtuvo el número de hojas, el tamaño del árbol, y la precisión de aciertos en porcentaje.

Tabla 9*Resultados del modelo MDL_2, con sus respectivos algoritmos*

Clasificador	Tipo de test	Num. hojas	Tam. árbol	Precisión (% aciertos)
Decision Stump	Validación cruzada	1	1	53.94
Hoeffding Tree	Validación cruzada	1	1	62.75
J48	Validación cruzada	22	29	64.47
LMT	Validación cruzada	1	1	64.33
Random Tree	Validación cruzada	1	81	64.04

Nota, esta tabla muestra los algoritmos, el tipo de test, el número de hojas, el tamaño del árbol y la precisión del Modelo II.

Los resultados obtenidos de la simulación con el test de validación cruzada, del MDL_3 con los algoritmos de minería de datos como Decision Stump, Hoeffding Tree, J48, LMT y Random Tree, se representa en la Tabla 10, donde se obtuvo el número de hojas, el tamaño del árbol y la precisión de aciertos en porcentaje.

Tabla 10*Resultados del modelo MDL_3, con sus respectivos algoritmos*

Clasificador	Tipo de test	Num. hojas	Tam. árbol	Precisión (% aciertos)
Decision Stump	Validación cruzada	1	1	53.01
Hoeffding Tree	Validación cruzada	1	1	68.19
J48	Validación cruzada	34	45	67.48
LMT	Validación cruzada	1	1	68.05
Random Tree	Validación cruzada	1	73	68.19

Nota, esta tabla muestra los algoritmos, el tipo de test, el número de hojas, el tamaño del árbol y la precisión del Modelo III.

Los resultados obtenidos de la simulación con el test de validación cruzada, del MDL_4 con los algoritmos de minería de datos como Decision Stump, Hoeffding Tree, J48, LMT y Random Tree, se representa en la Tabla 11, donde se obtuvo el número de hojas, el tamaño del árbol y la precisión de aciertos en porcentaje.

Tabla 11

Resultados del modelo MDL_4, con sus respectivos algoritmos

Clasificador	Tipo de test	Num. hojas	Tam. árbol	Precisión (% aciertos)
Decision Stump	Validación cruzada	1	1	53.72
Hoeffding Tree	Validación cruzada	1	1	84.67
J48	Validación cruzada	19	25	82.02
LMT	Validación cruzada	1	1	87.32
Random Tree	Validación cruzada	1	73	82.09

Nota, esta tabla muestra los algoritmos, el tipo de test, el número de hojas, el tamaño del árbol y la precisión del Modelo IV.

Los resultados obtenidos de la simulación con el test de validación cruzada, del MDL_5 con los algoritmos de minería de datos como Decision Stump, Hoeffding Tree, J48, LMT y Random Tree, se representa en la Tabla 12, donde se obtuvo el número de hojas, el tamaño del árbol y la precisión de aciertos en porcentaje.

Tabla 12*Resultados del modelo MDL_5, con sus respectivos algoritmos*

Clasificador	Tipo de test	Num. hojas	Tam. árbol	Precisión (% aciertos)
Decision Stump	Validación cruzada	1	1	53.72
Hoeffding Tree	Validación cruzada	1	1	80.52
J48	Validación cruzada	19	25	79.73
LMT	Validación cruzada	1	1	79.15
Random Tree	Validación cruzada	1	73	82.45

Nota, esta tabla muestra los algoritmos, el tipo de test, el número de hojas, el tamaño del árbol y la precisión del Modelo V.

Los resultados obtenidos de la predicción de la vulnerabilidad sísmica con los diferentes algoritmos de minería de datos, entrenados y validados, se presentan en la Tabla 13, donde los modelos MDL_01, MDL_04 y MDL_05, presentan el menor error en la estimación de la vulnerabilidad sísmica.

Tabla 13*Predicción del modelo MDL_01*

		Predicción				
		Algoritmo				
Vivienda	Real	Decision Stump	Hoeffding Tree	J48	LMT	Random Tree
M2 – 31	Baja	Media	Baja	Baja	Baja	Baja
M3 – 11	Baja	Media	Baja	Baja	Baja	Baja
M5 – 19	Baja	Media	Media	Media	Media	Media
M6 – 14	Media	Media	Media	Media	Media	Media

M6 – 20	Media	Media	Media	Media	Media	Media
M7 – 3	Baja	Media	Baja	Baja	Baja	Baja
M7 – 6	Media	Media	Baja	Baja	Baja	Baja
M7 – 12	Baja	Media	Baja	Baja	Baja	Baja
M7 – 14	Media	Media	Baja	Baja	Baja	Baja
M7 – 23	Media	Media	Baja	Baja	Baja	Baja
M8 – 3	Baja	Media	Baja	Baja	Baja	Baja
M8 – 8	Media	Media	Alta	Alta	Alta	Alta
M9 – 16	Baja	Media	Baja	Baja	Baja	Baja
M9 – 26	Baja	Media	Baja	Baja	Baja	Baja
M10 – 9	Media	Media	Media	Media	Media	Media
M10 – 15	Baja	Media	Baja	Baja	Baja	Baja
M11 – 8	Media	Media	Baja	Baja	Baja	Baja
M11 – 17	Baja	Media	Baja	Baja	Baja	Baja
M11 – 30	Baja	Media	Baja	Baja	Baja	Baja
M12 – 3	Baja	Media	Media	Media	Media	Media
M12 – 62	Baja	Media	Media	Media	Media	Media
M13 – 1	Baja	Media	Baja	Baja	Baja	Baja
M13 – 6	Baja	Media	Baja	Baja	Baja	Baja
M13 – 12	Baja	Media	Baja	Baja	Baja	Baja
M13 – 31	Baja	Media	Baja	Baja	Baja	Baja
M13 – 33	Media	Media	Media	Media	Media	Baja
M13 – 34	Media	Media	Baja	Baja	Baja	Baja
M13 – 40	Baja	Media	Media	Media	Media	Baja
M14 – 1	Baja	Media	Baja	Baja	Baja	Baja
M14 – 15	Media	Media	Baja	Baja	Baja	Baja

M14 – 32	Media	Media	Baja	Baja	Baja	Baja
M15 – 4	Alta	Media	Media	Alta	Media	Alta
M15 – 16	Baja	Media	Baja	Baja	Baja	Baja
M17 – 3	Baja	Media	Baja	Baja	Baja	Baja
M17 – 10	Baja	Media	Baja	Baja	Baja	Baja
M18 – 9	Baja	Media	Baja	Baja	Baja	Baja
M18 – 27	Baja	Media	Baja	Baja	Baja	Baja
M18 – 29	Baja	Media	Baja	Baja	Baja	Baja
M19 – 1	Media	Media	Alta	Alta	Alta	Alta
M19 – 23	Baja	Media	Baja	Baja	Baja	Baja
M19 – 24	Media	Media	Media	Media	Media	Media
M21 – 15	Baja	Media	Baja	Baja	Baja	Baja
M21 – 20	Baja	Media	Baja	Baja	Baja	Baja
M22 – 13	Baja	Media	Baja	Baja	Baja	Baja
M23 – 2	Baja	Media	Baja	Baja	Baja	Baja
M23 – 16	Baja	Media	Baja	Baja	Baja	Baja
M25 – 2	Media	Media	Media	Media	Media	Media
M27 – 18	Baja	Media	Baja	Baja	Baja	Baja
M28 – 1	Media	Media	Media	Media	Media	Media
M29 – 20	Baja	Media	Baja	Baja	Baja	Baja
M29 – 23	Baja	Media	Baja	Baja	Baja	Baja
M29 – 28	Baja	Media	Media	Media	Media	Media
M29 – 37	Media	Media	Baja	Baja	Baja	Baja
M30 – 13	Baja	Media	Media	Media	Media	Media
M32 – 8	Baja	Media	Media	Media	Media	Baja
M33 – 2	Baja	Media	Media	Baja	Baja	Baja

M33 – 13	Media	Media	Baja	Baja	Baja	Baja
M33 – 22	Baja	Media	Media	Media	Media	Media
M33 – 23	Media	Media	Media	Media	Media	Media
M33 – 35	Baja	Media	Media	Baja	Baja	Baja
M35 – 10	Baja	Media	Baja	Baja	Baja	Baja
M35 – 21	Baja	Media	Baja	Baja	Baja	Baja
M36 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 11	Baja	Media	Baja	Baja	Baja	Baja
M38 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 18	Media	Media	Baja	Baja	Baja	Baja
M39 – 6	Media	Media	Media	Media	Media	Media

Nota, en la presente tabla se detallan los resultados de la estimación, por los cinco algoritmos de minería de datos, del MDL_1 categorizándose a la vulnerabilidad sísmica de baja a media.

Tabla 14

Precisión del modelo MDL_01

Modelo	Árbol	Precisión (%)	Instancias correctamente clasificadas	Instancias incorrectamente clasificadas
MDL_01	Decision Stump	31.34	21	46
	Hoeffding Tree	65.67	44	23
	J48	70.14	47	20
	LMT	68.66	46	21
	Random Tree	71.64	48	19

Nota, la presente tabla detalla la precisión de los algoritmos, en base a las instancias correctamente e incorrectamente clasificadas, donde el algoritmo Random Tree presenta mayor precisión.

Tabla 15*Predicción del modelo MDL_04*

Vivienda	Real	Predicción				
		Algoritmo				
		Decision Stump	Hoeffding Tree	J48	LMT	Random Tree
M2 – 31	Baja	Media	Baja	Baja	Baja	Baja
M3 – 11	Baja	Media	Baja	Baja	Baja	Baja
M5 – 19	Baja	Media	Baja	Baja	Baja	Baja
M6 – 14	Media	Media	Media	Media	Media	Media
M6 – 20	Media	Media	Baja	Baja	Baja	Baja
M7 – 3	Baja	Media	Media	Baja	Media	Baja
M7 – 6	Media	Media	Baja	Baja	Baja	Baja
M7 – 12	Baja	Media	Baja	Baja	Baja	Baja
M7 – 14	Media	Media	Media	Baja	Media	Baja
M7 – 23	Media	Media	Media	Baja	Media	Baja
M8 – 3	Baja	Media	Media	Baja	Media	Baja
M8 – 8	Media	Media	Media	Media	Media	Media
M9 – 16	Baja	Media	Baja	Baja	Baja	Baja
M9 – 26	Baja	Media	Baja	Baja	Baja	Baja
M10 – 9	Media	Media	Media	Media	Media	Media
M10 – 15	Baja	Media	Baja	Baja	Baja	Baja
M11 – 8	Media	Media	Media	Baja	Media	Baja
M11 – 17	Baja	Media	Baja	Baja	Baja	Baja
M11 – 30	Baja	Media	Baja	Baja	Baja	Baja

M12 – 3	Baja	Media	Media	Media	Media	Media
M12 – 62	Baja	Media	Media	Media	Media	Media
M13 – 1	Baja	Media	Baja	Baja	Baja	Baja
M13 – 6	Baja	Media	Baja	Baja	Baja	Baja
M13 – 12	Baja	Media	Media	Baja	Media	Baja
M13 – 31	Baja	Media	Baja	Baja	Baja	Baja
M13 – 33	Media	Media	Media	Media	Media	Media
M13 – 34	Media	Media	Baja	Baja	Baja	Baja
M13 – 40	Baja	Media	Media	Baja	Media	Baja
M14 – 1	Baja	Media	Baja	Baja	Baja	Baja
M14 – 15	Media	Media	Media	Baja	Media	Baja
M14 – 32	Media	Media	Media	Baja	Media	Baja
M15 – 4	Alta	Media	Media	Baja	Media	Media
M15 – 16	Baja	Media	Baja	Baja	Baja	Baja
M17 – 3	Baja	Media	Media	Baja	Media	Media
M17 – 10	Baja	Media	Baja	Baja	Baja	Baja
M18 – 9	Baja	Media	Baja	Baja	Media	Baja
M18 – 27	Baja	Media	Baja	Baja	Baja	Baja
M18 – 29	Baja	Media	Baja	Baja	Baja	Baja
M19 – 1	Media	Media	Media	Media	Media	Media
M19 – 23	Baja	Media	Baja	Baja	Baja	Baja
M19 – 24	Media	Media	Media	Media	Media	Media
M21 – 15	Baja	Media	Baja	Baja	Baja	Baja
M21 – 20	Baja	Media	Baja	Baja	Baja	Baja
M22 – 13	Baja	Media	Media	Baja	Media	Media

M23 – 2	Baja	Media	Media	Baja	Media	Baja
M23 – 16	Baja	Media	Media	Baja	Media	Media
M25 – 2	Media	Media	Media	Media	Media	Media
M27 – 18	Baja	Media	Baja	Baja	Baja	Baja
M28 – 1	Media	Media	Media	Media	Media	Media
M29 – 20	Baja	Media	Baja	Baja	Media	Baja
M29 – 23	Baja	Media	Media	Baja	Media	Baja
M29 – 28	Baja	Media	Media	Baja	Media	Media
M29 – 37	Media	Media	Baja	Baja	Baja	Baja
M30 – 13	Baja	Media	Baja	Baja	Baja	Baja
M32 – 8	Baja	Media	Media	Baja	Media	Baja
M33 – 2	Baja	Media	Media	Baja	Media	Baja
M33 – 13	Media	Media	Baja	Baja	Baja	Baja
M33 – 22	Baja	Media	Baja	Baja	Baja	Baja
M33 – 23	Media	Media	Media	Media	Media	Media
M33 – 35	Baja	Media	Media	Baja	Media	Baja
M35 – 10	Baja	Media	Baja	Baja	Baja	Baja
M35 – 21	Baja	Media	Baja	Baja	Baja	Baja
M36 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 11	Baja	Media	Baja	Baja	Baja	Baja
M38 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 18	Media	Media	Baja	Baja	Baja	Baja
M39 – 6	Media	Media	Media	Media	Media	Media

Nota, en la presente tabla se detallan los resultados de la estimación, por los cinco algoritmos de minería de datos, del MDL_04, categorizándose a la vulnerabilidad sísmica de baja a media.

Tabla 16*Precisión del modelo MDL_04*

Modelo	Árbol	Precisión (%)	Instancias correctamente clasificadas	Instancias incorrectamente clasificadas
	Decision Stump	31.34	21	46
	Hoeffding Tree	67.16	45	22
MDL_04	J48	79.10	53	14
	LMT	64.18	43	24
	Random Tree	73.13	49	18

Nota, la presente tabla detalla la precisión de los algoritmos, en base a las instancias correctamente e incorrectamente clasificadas, donde el algoritmo J48 presenta mayor precisión.

Tabla 17.*Predicción del modelo MDL_05*

Vivienda	Real	Predicción				
		Algoritmo				
		Decision Stump	Hoeffding Tree	J48	LMT	Random Tree
M2 – 31	Baja	Media	Baja	Baja	Baja	Baja
M3 – 11	Baja	Media	Baja	Baja	Baja	Baja
M5 – 19	Baja	Media	Baja	Baja	Baja	Baja
M6 – 14	Media	Media	Media	Media	Media	Media
M6 – 20	Media	Media	Baja	Baja	Baja	Baja
M7 – 3	Baja	Media	Baja	Baja	Baja	Baja
M7 – 6	Media	Media	Baja	Baja	Baja	Baja
M7 – 12	Baja	Media	Baja	Baja	Baja	Baja
M7 – 14	Media	Media	Baja	Baja	Baja	Baja
M7 – 23	Media	Media	Media	Media	Media	Media

M8 – 3	Baja	Media	Baja	Baja	Baja	Baja
M8 – 8	Media	Media	Media	Media	Media	Media
M9 – 16	Baja	Media	Baja	Baja	Baja	Baja
M9 – 26	Baja	Media	Baja	Baja	Baja	Baja
M10 – 9	Media	Media	Media	Media	Media	Media
M10 – 15	Baja	Media	Baja	Baja	Baja	Baja
M11 – 8	Media	Media	Baja	Baja	Baja	Baja
M11 – 17	Baja	Media	Baja	Baja	Baja	Baja
M11 – 30	Baja	Media	Media	Media	Media	Baja
M12 – 3	Baja	Media	Media	Media	Media	Media
M12 – 62	Baja	Media	Baja	Media	Media	Media
M13 – 1	Baja	Media	Baja	Baja	Baja	Baja
M13 – 6	Baja	Media	Baja	Baja	Baja	Baja
M13 – 12	Baja	Media	Baja	Baja	Baja	Baja
M13 – 31	Baja	Media	Baja	Baja	Baja	Baja
M13 – 33	Media	Media	Media	Media	Media	Media
M13 – 34	Media	Media	Baja	Baja	Baja	Baja
M13 – 40	Baja	Media	Media	Baja	Media	Baja
M14 – 1	Baja	Media	Baja	Baja	Baja	Baja
M14 – 15	Media	Media	Baja	Baja	Baja	Baja
M14 – 32	Media	Media	Media	Baja	Media	Media
M15 – 4	Alta	Media	Baja	Baja	Baja	Baja
M15 – 16	Baja	Media	Baja	Baja	Baja	Baja
M17 – 3	Baja	Media	Baja	Baja	Baja	Baja
M17 – 10	Baja	Media	Baja	Baja	Baja	Baja

M18 – 9	Baja	Media	Baja	Baja	Baja	Baja
M18 – 27	Baja	Media	Baja	Baja	Baja	Baja
M18 – 29	Baja	Media	Baja	Baja	Baja	Baja
M19 – 1	Media	Media	Media	Media	Media	Media
M19 – 23	Baja	Media	Baja	Baja	Baja	Baja
M19 – 24	Media	Media	Baja	Media	Media	Media
M21 – 15	Baja	Media	Baja	Baja	Baja	Baja
M21 – 20	Baja	Media	Baja	Baja	Baja	Baja
M22 – 13	Baja	Media	Baja	Baja	Baja	Baja
M23 – 2	Baja	Media	Baja	Baja	Baja	Baja
M23 – 16	Baja	Media	Baja	Baja	Baja	Baja
M25 – 2	Media	Media	Media	Media	Media	Media
M27 – 18	Baja	Media	Baja	Baja	Baja	Baja
M28 – 1	Media	Media	Baja	Media	Media	Media
M29 – 20	Baja	Media	Baja	Baja	Baja	Baja
M29 – 23	Baja	Media	Baja	Baja	Baja	Baja
M29 – 28	Baja	Media	Baja	Baja	Baja	Baja
M29 – 37	Media	Media	Media	Media	Media	Baja
M30 – 13	Baja	Media	Baja	Baja	Baja	Baja
M32 – 8	Baja	Media	Baja	Baja	Baja	Baja
M33 – 2	Baja	Media	Media	Baja	Media	Baja
M33 – 13	Media	Media	Baja	Baja	Baja	Baja
M33 – 22	Baja	Media	Baja	Baja	Baja	Baja
M33 – 23	Media	Media	Media	Media	Media	Media
M33 – 35	Baja	Media	Baja	Baja	Baja	Baja

M35 – 10	Baja	Media	Baja	Baja	Baja	Baja
M35 – 21	Baja	Media	Media	Media	Media	Baja
M36 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 11	Baja	Media	Media	Media	Media	Baja
M38 – 17	Baja	Media	Baja	Baja	Baja	Baja
M38 – 18	Media	Media	Media	Media	Media	Baja
M39 – 6	Media	Media	Media	Media	Media	Media

Nota, en la presente tabla se detallan los resultados de la estimación, por los cinco algoritmos de minería de datos, del MDL_05, categorizándose a la vulnerabilidad sísmica de baja a media.

Tabla 18

Precisión del modelo MDL_05

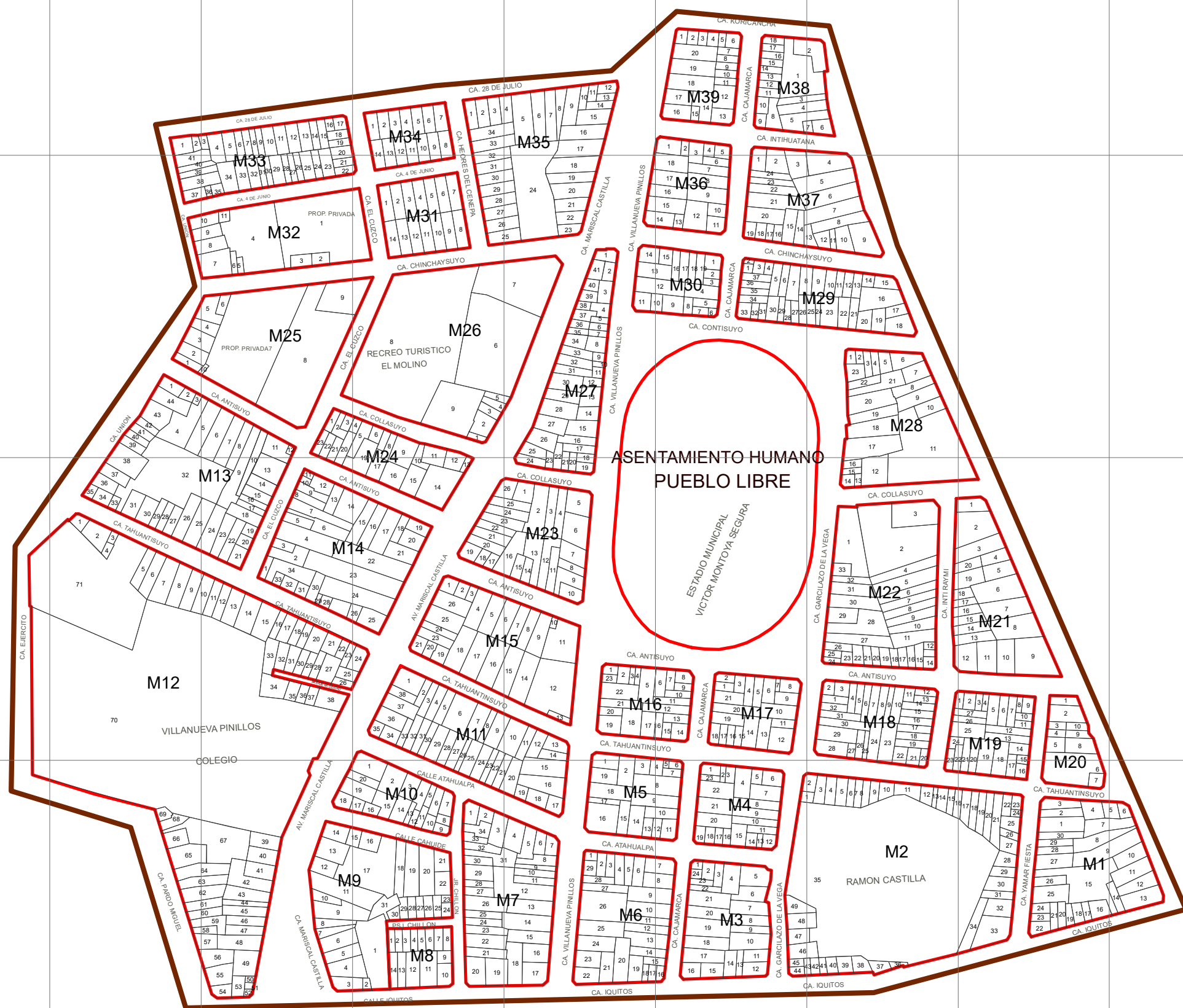
Modelo	Árbol	Precisión (%)	Instancias correctamente clasificadas	Instancias incorrectamente clasificadas
	Decision Stump	31.34	21	46
	Hoeffding Tree	76.12	51	16
MDL_05	J48	79.10	53	14
	LMT	77.61	52	15
	Random Tree	82.09	55	12

Nota, la presente tabla detalla la precisión de los algoritmos, en base a las instancias correctamente e incorrectamente clasificadas, donde el algoritmo Random Tree presenta mayor precisión.



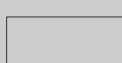
4.3. Mapa de zonificación sísmica según los niveles de vulnerabilidad sísmica del sector Pueblo Libre

El primer mapa presenta la ubicación, las viviendas y los lotes del sector Pueblo Libre con sus respectivas manzanas. El segundo mapa presenta todas las viviendas de albañilería confinada del sector representadas de color celeste claro por cada mazana. El tercer mapa presenta los niveles de vulnerabilidad sísmica representando de color azul las viviendas con vulnerabilidad baja, y de color verde claro las viviendas con vulnerabilidad media.

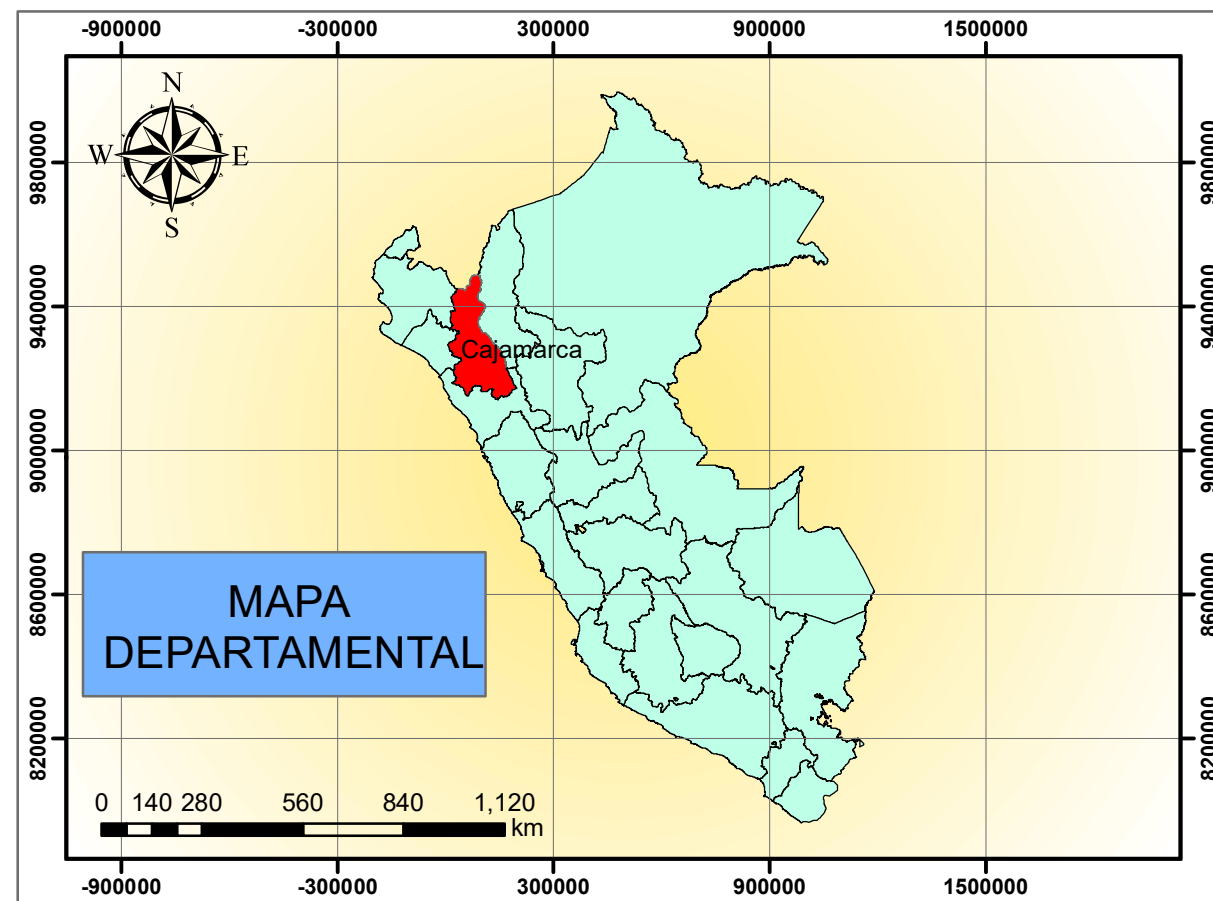
SECTOR - PUEBLO LIBRE



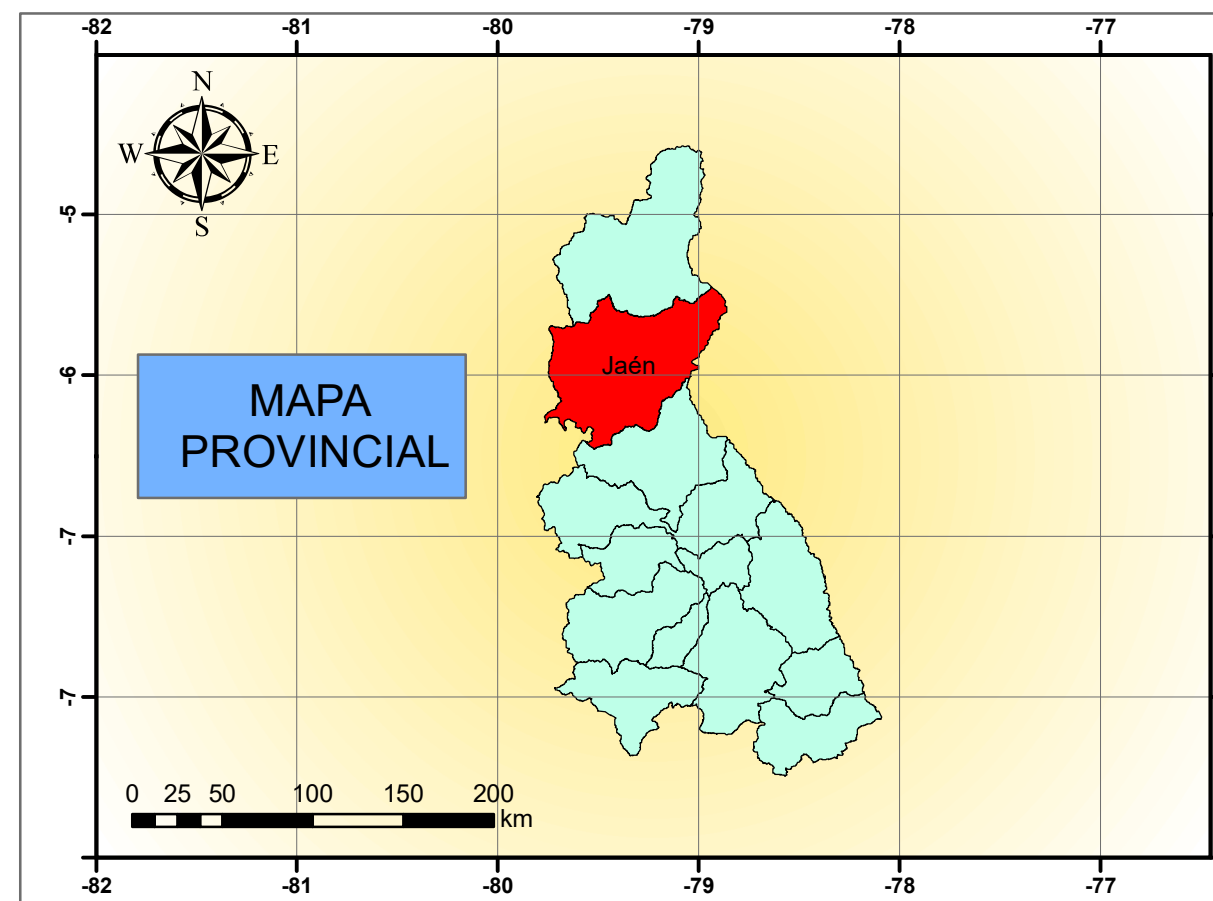
LEYENDA

-  Contorno Pueblo Libre
-  Manzanas-
-  Viviendas

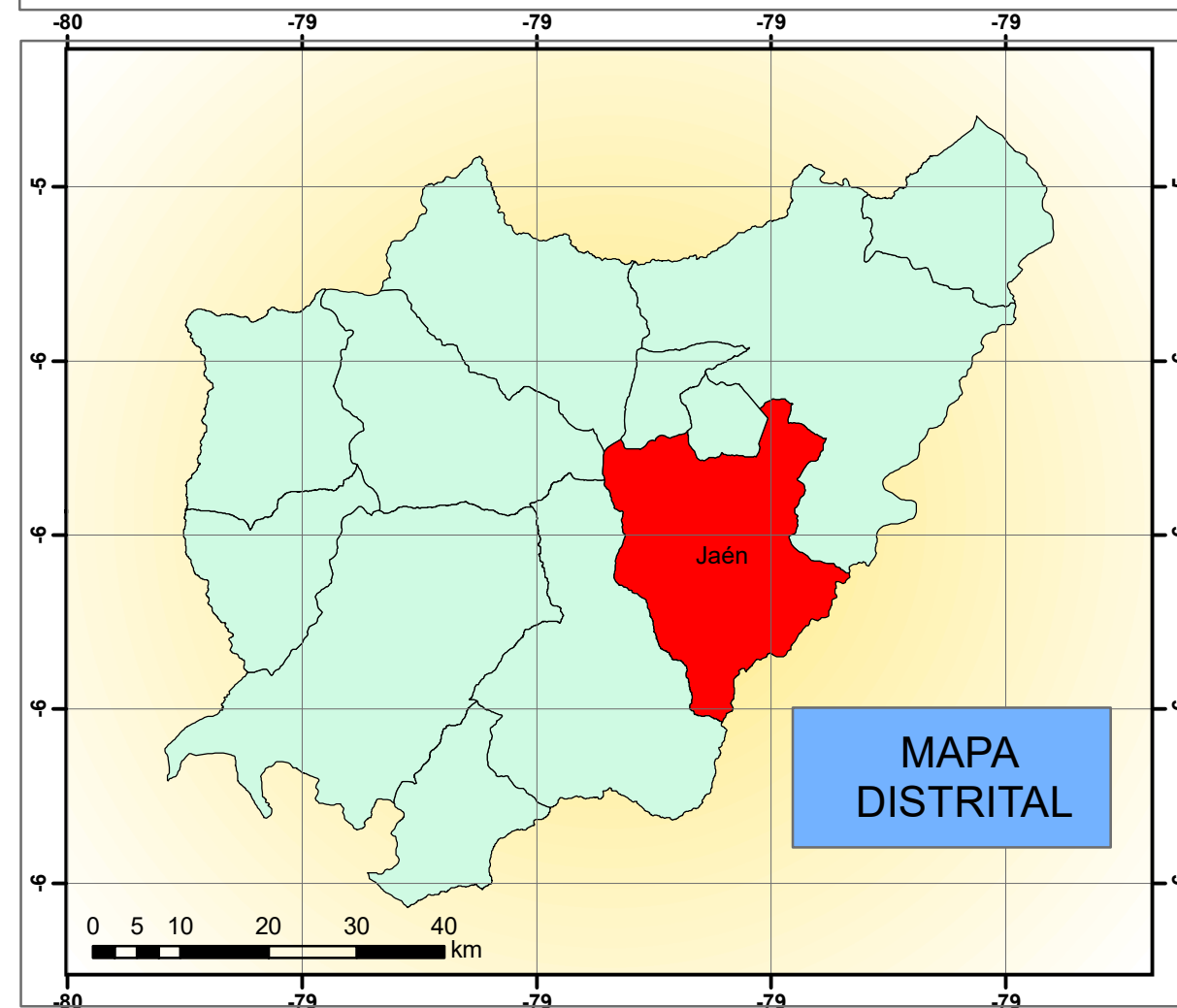
MAPA DEPARTAMENTAL



MAPA PROVINCIAL



MAPA DISTRITAL

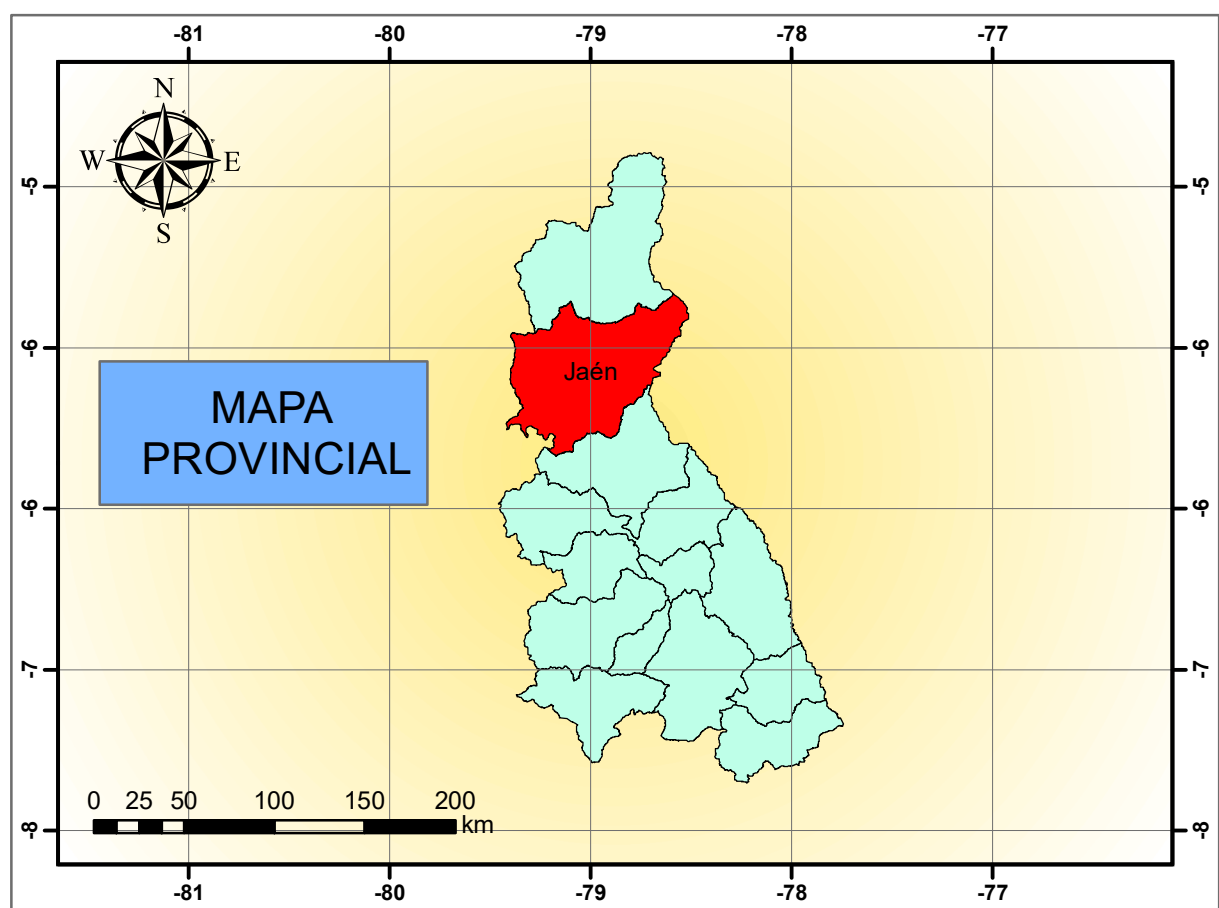
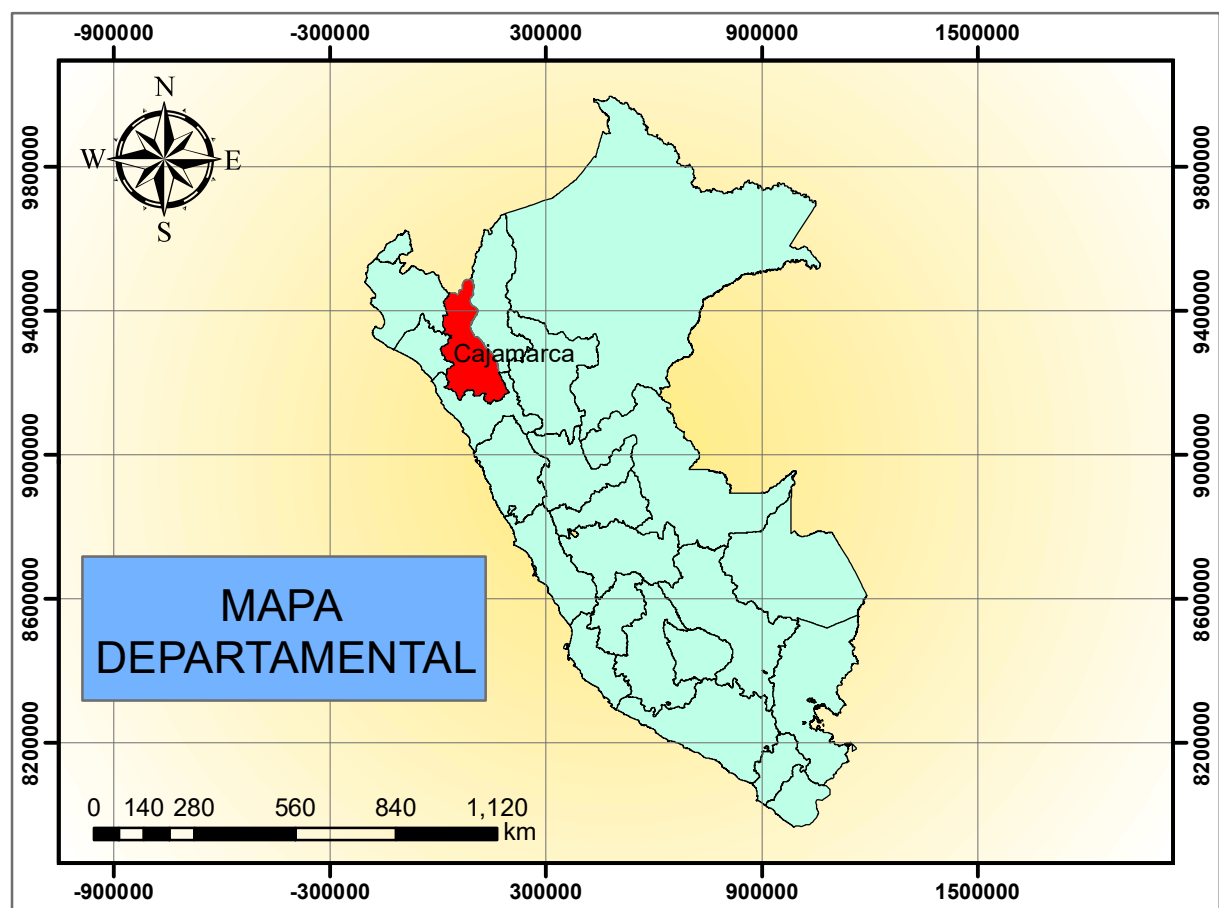
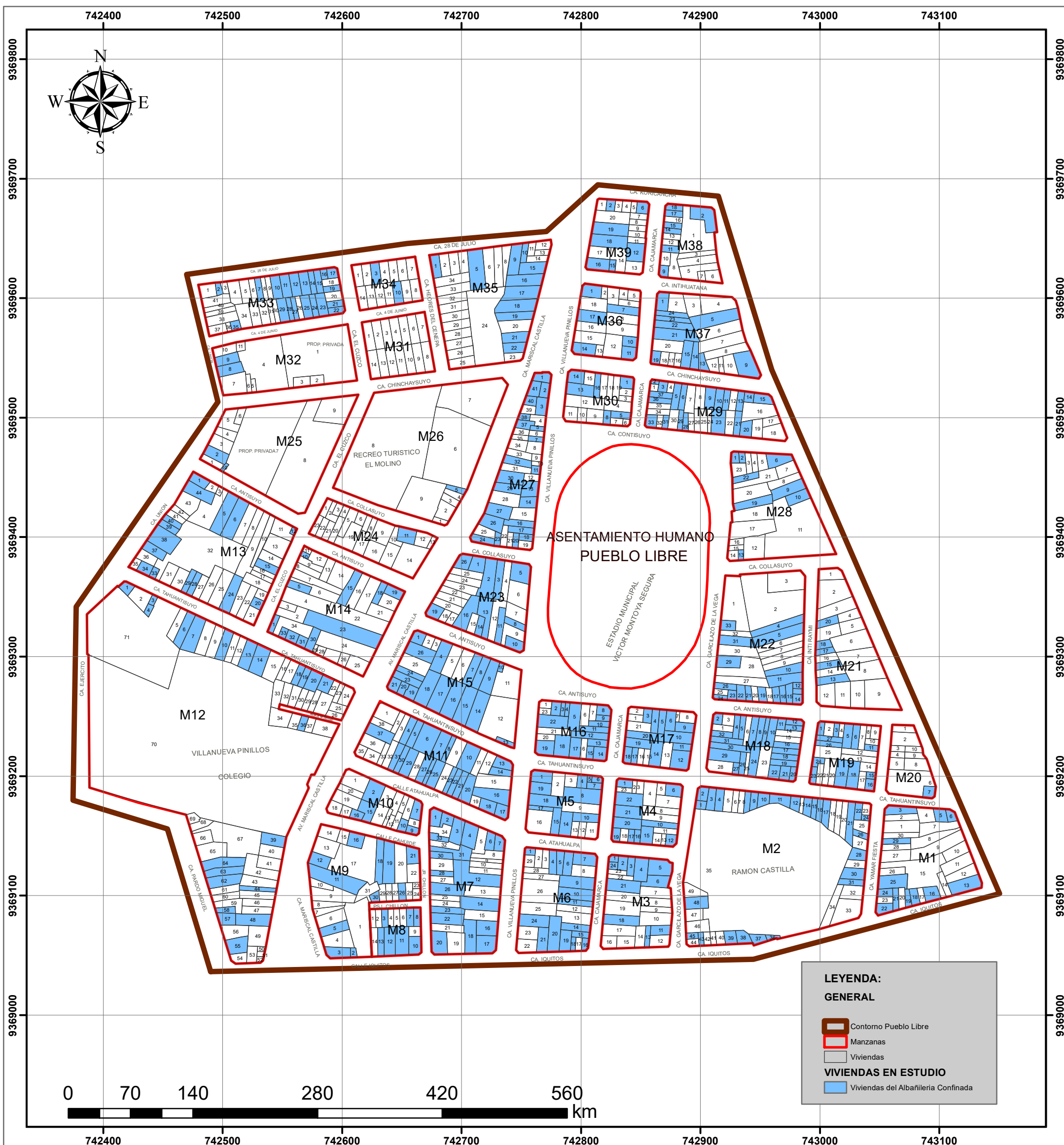


UNIVERSIDAD NACIONAL DE JAÉN
CARRERA PROFESIONAL DE INGENIERÍA CIVIL



Tesis: "ESTIMACIÓN DE LA VULNERABILIDAD SÍSMICA EN VIVIENDAS DE ALBAÑILERÍA CONFINADA, MEDIANTE TÉCNICAS DE MINERÍA DE DATOS, EN EL SECTOR PUEBLO LIBRE, JAÉN - 2020"

Lámina:	PLANO DE UBICACIÓN DEL SECTOR PUEBLO LIBRE DE LA CIUDAD DE JAÉN (ÁREA DE ESTUDIO)	REGIÓN:	CAJAMARCA
Responsables:	PALOMINO OJEDA JOSE MANUEL ROSARIO BOCANEGRA STEFANO	PROVINCIA:	JAÉN
Asesores:	M.Sc. Ing. Christiaan Zayed Apaza Panca Mg. Lenin Quiñones Huatangari	DISTRITO:	JAÉN
DATUM: WGS - 84	SIST. DE PROYEC.: UTM	FECHA:	FEBRERO - 2021
	ZONA 17	ESCALA:	INDICADA
		PLANO:	U - 01



LEYENDA:
GENERAL

- Contorno Pueblo Libre
- Manzanas
- Viviendas

VIVIENDAS EN ESTUDIO

- Viviendas del Albañilería Confinada

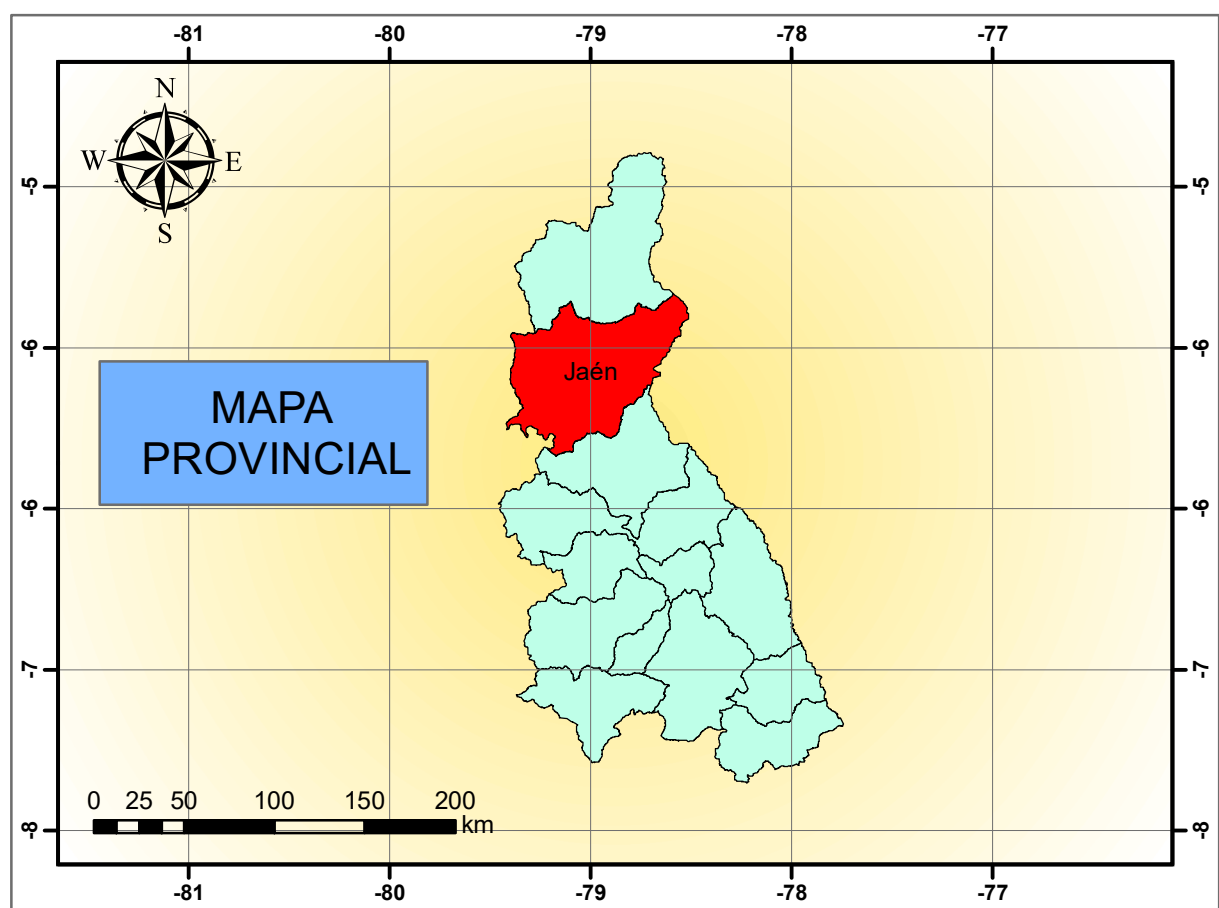
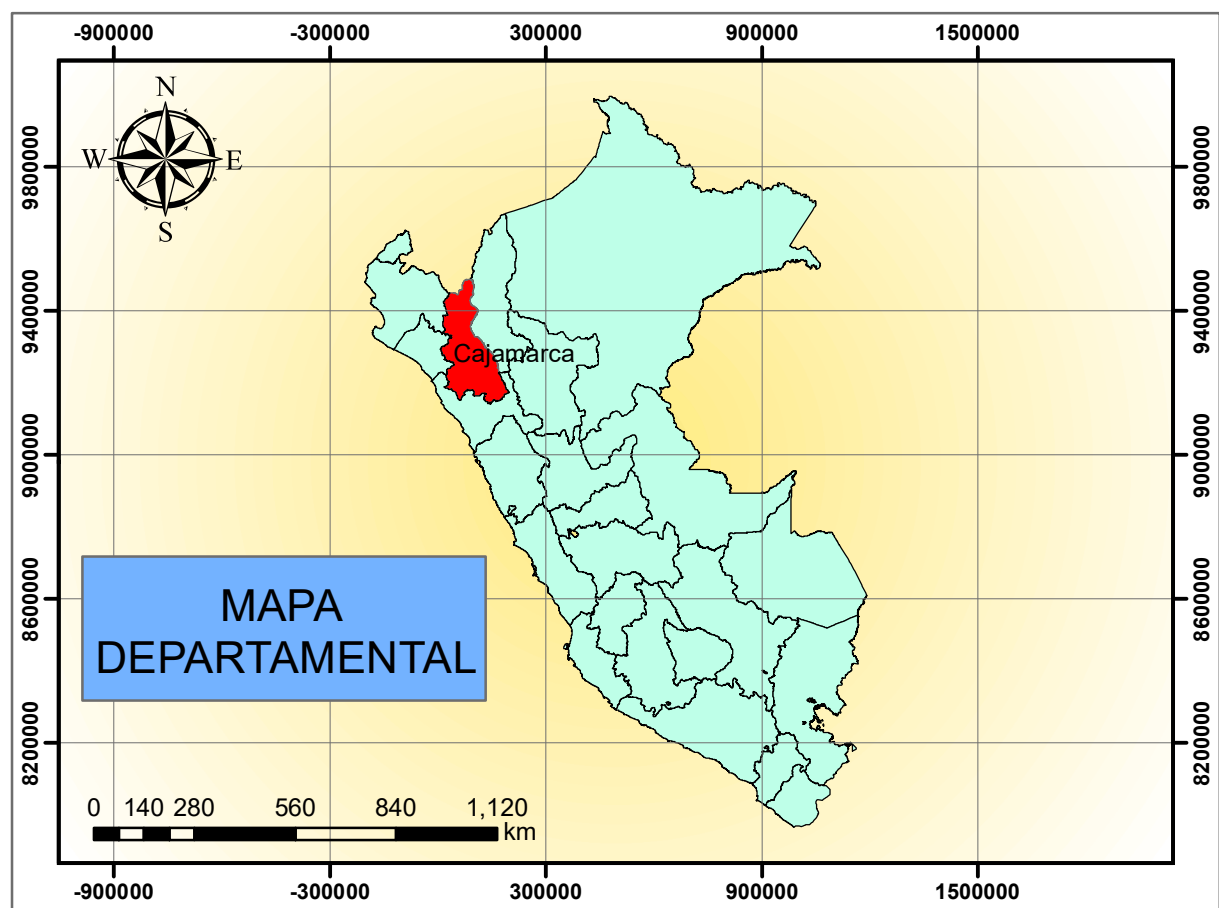
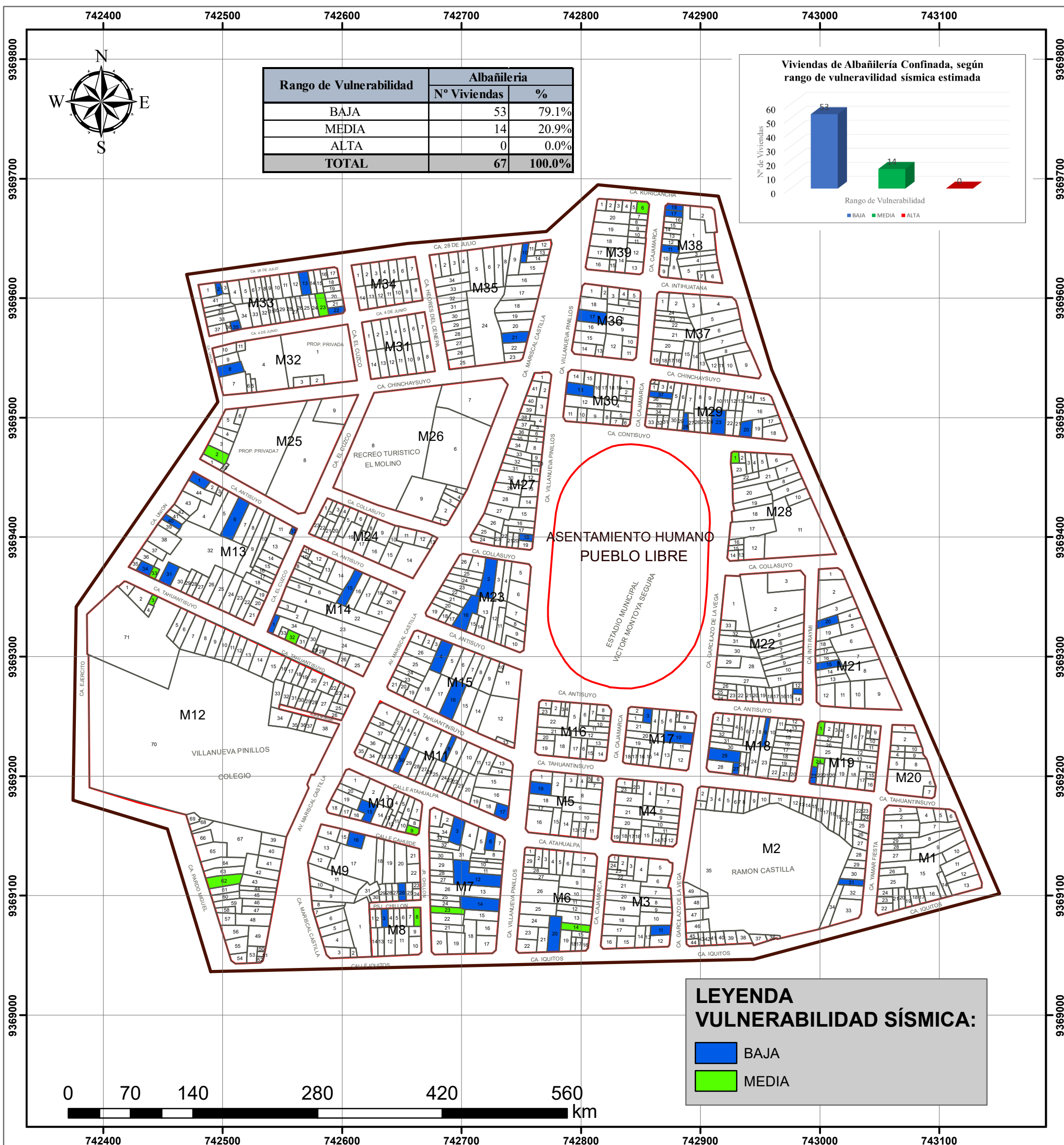
UNIVERSIDAD NACIONAL DE JAÉN


CARRERA PROFESIONAL DE INGENIERÍA CIVIL

Tesis: "ESTIMACIÓN DE LA VULNERABILIDAD SÍSMICA EN VIVIENDAS DE ALBAÑILERÍA CONFINADA, MEDIANTE TÉCNICAS DE MINERÍA DE DATOS, EN EL SECTOR PUEBLO LIBRE, JAÉN - 2020"

<i>Lámina:</i> PLANO DE LOTES IDENTIFICADOS DE ALBAÑILERÍA CONFINADA EN EL SECTOR - PUEBLO LIBRE (ÁREA DE ESTUDIO)	REGIÓN: CAJAMARCA
<i>Responsables:</i> PALOMINO OJEDA JOSE MANUEL ROSARIO BOCANEGRA STEFANO	PROVINCIA: JAÉN
<i>Asesores:</i> M.Sc. Ing. Christian Zayed Apaza Panca Mg. Lenn Quñones Huatangari	DISTRITO: JAÉN
	FECHA: FEBRERO - 2021
	ESCALA: INDICADA
	PLANO: PL - 01


DATUM: WGS - 84 SIST. DE PROYEC.: UTM ZONA 17





UNIVERSIDAD NACIONAL DE JAÉN

CARRERA PROFESIONAL DE INGENIERÍA CIVIL



Tesis: "ESTIMACIÓN DE LA VULNERABILIDAD SÍSMICA EN VIVIENDAS DE ALBAÑILERÍA CONFINADA, MEDIANTE TÉCNICAS DE MINERÍA DE DATOS, EN EL SECTOR PUEBLO LIBRE, JAÉN - 2020"

<i>Lámina:</i>	ZONIFICACIÓN - VULNERABILIDAD SÍSMICA ESTIMADA DE LAS VIVIENDAS EN ESTUDIO	REGIÓN: CAJAMARCA
		PROVINCIA: JAÉN
<i>Responsables:</i>	PALOMINO OJEDA JOSE MANUEL	DISTRITO: JAÉN
	ROSARIO BOCANEGRA STEFANO	FECHA: FEBRERO - 2021
<i>Asesores:</i>	M.Sc. Ing. Christian Zayed Apaza Panca	ESCALA: INDICADA
	Mg. Lenin Quiñones Huatangari	PLANO: VS - 01

DATUM: WGS - 84 SIST. DE PROYEC.: UTM ZONA 17

V. DISCUSIÓN

A partir de los resultados obtenidos aceptamos la hipótesis general que establece que la vulnerabilidad sísmica en viviendas de albañilería confinada mediante técnicas de minería de datos, se encuentra en el nivel baja a media con un error menor al 25%, en el sector Pueblo Libre-Jaén, obteniéndose un error de 17.91% con una estimación de vulnerabilidad sísmica de baja a media. Estos valores guardan relación con lo que sostiene Ramos (2020), donde determina que la vulnerabilidad sísmica se encuentra en el nivel media a baja, al igual que Quiñones & Carrasco (2020), donde obtienen un 83% de clasificación correcta, con un error de 17% con la metodología de minería de datos.

La matriz de datos estuvo conformada por 3 827 registros y doce variables recopiladas de tesis de varios repositorios del Perú, la cual presentan variables de tipo ordinal semejante a la base de datos obtenida por Quiñones & Carrasco (2020), en el artículo científico "*Rendimiento académico empleando minería de datos*" donde las variables son de tipo ordinal, nominal y discreto. Además, las variables de la matriz de datos están basadas en las Normas E-070 (albañilería confinada), y E-030 (Diseño sismorresistente).

La base de datos presenta variables cualitativas las cuales se clasifican en A, B, C y D al igual que Teobaldo & Santisteban (2015), en la tesis "*Vulnerabilidad sísmica en el distrito de ciudad Eten aplicando índices de vulnerabilidad (Benedetti Petrini)*" donde las variables son cualitativas y cumplen las normas internacionales y nacionales asignado A las variables que cumplen con todos los requisitos de la norma y D a las que no cumplen.

La modelación de la vulnerabilidad sísmica con técnica de minería de datos se realizó mediante cinco modelos MDL_1, MDL_2, MDL_3, MDL_4, MDL_5 los cuales están conformados con tres variables diferentes, entrenados con cinco algoritmos de árboles de decisiones en Weka obteniéndose un porcentaje de clasificación correctamente de 69.05%, 64.67%, 68.19%, 87.32% y 82.45% respectivamente, los cuales son semejantes a los resultados obtenidos por Saavedra (2016), Espino & García (2018) y Timarán & et al (2017) de 83.35%, 80.00% y 81.70% instancias correctamente clasificadas.

Los algoritmos seleccionados para la comparación son Decision Stump, Hoeffding Tree, J48, LMT y Random Tree siendo los que mejor representan los patrones de un conjunto de datos,

según Beltrán (2018) el cual en su investigación utiliza algoritmos de árboles de decisión como Decision Stump, Hoeffding Tree, J48, LMT, Random Forest, Random Tree y Rep Tree para elegir al mejor algoritmo aplicado al mismo conjunto de datos.

En la elaboración del mapa de zonificación de vulnerabilidad sísmica del sector Pueblo Libre, se sistematizó la información recopilada y obtenida de la estimación de la vulnerabilidad sísmica de las viviendas de albañilería confinada, relacionando la información espacial y no espacial, el mismo procedimiento seguido por Flores (2016), que le permitió realizar el geoprocésamiento para determinar los grados de vulnerabilidad de distintos aspectos analizados, obteniendo resultados satisfactorios.

El mapa de zonificación muestra los niveles de vulnerabilidad sísmica estimada para las viviendas de albañilería confinada del sector en estudio, donde gráficamente se puede indicar que el 79.1%, 20.9% y 0% representa el rango de vulnerabilidad baja, media y alta, con un error menor al 18%, los cuales son semejantes a la investigación de Ramos (2020) donde obtuvo valores de 67.16%, 31.34% y 1.49%.

VI. CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

- Se estimó la vulnerabilidad sísmica en viviendas de albañilería confinada, mediante técnicas de minería de datos. Esta se encuentra en el nivel de baja a media con una confiabilidad del 82.09%, en el sector Pueblo Libre, Jaén.
- Se creó una matriz de datos compuesta por 3 827 registros de viviendas de albañilería confinada y doce variables de tipo ordinal, clasificadas en A, B, C y D recopiladas de tesis de varios repositorios del Perú, para el entrenamiento y validación de los diferentes modelos de minería de datos.
- Se modeló la vulnerabilidad sísmica a través de técnica de minería de datos planteándose cinco modelos MDL_1, MDL_2, MDL_3, MDL_4, MDL_5, que utilizaron la metodología KDD con los algoritmos Decision Stump, Hoeffding Tree, J48, LMT y Random Tree, obteniéndose el menor error en el MDL_5 con el algoritmo Random Tree del 17.91%.
- Se elaboró un mapa de zonificación modelado espacial y georreferenciado por medio de una plataforma de Sistemas de Información Geográfica (SIG) permitiendo una representación e identificación gráfica de la vulnerabilidad sísmica estimada de las viviendas de albañilería confinada en el sector Pueblo Libre, donde se determinó que el 79.1% (53 viviendas) presentan vulnerabilidad baja, el 20.9% (14 viviendas) vulnerabilidad media y el 0.00% (ninguna vivienda) vulnerabilidad alta; se establece de esta manera que el índice de vulnerabilidad estimado que predomina es de baja a media.

6.2. Recomendaciones

- A los investigadores se les sugiere que, para estimar la vulnerabilidad sísmica en viviendas de albañilería confinada, deben considerar como mínimo un 85% de confiabilidad para determinar si esta se encuentra en el nivel baja, media o alta, de manera eficiente.
- A los estudiantes e investigadores que diseñen modelos con minería de datos que, para estimar la vulnerabilidad sísmica en viviendas de albañilería confinada, se requiere la recopilación de mayor información de viviendas y sus características, ya que esto reduciría el margen de error estimado.
- A los interesados en aplicar técnicas de minería de datos para estimar la vulnerabilidad sísmica en viviendas de albañilería se sugiere la aplicación de nuevas variables y parámetros que puedan influir en la vulnerabilidad sísmica y también se sugiere utilizar la metodología KDD con los diferentes algoritmos de árboles de decisión ya que son eficaces y reducen el margen de error.
- A los interesados y autoridades municipales se recomienda en base del mapa de zonificación producto de la investigación, diseñar un plan de mitigación y evacuación para la gestión del riesgo sísmico en el sector Pueblo Libre.
- A los estudiantes e investigadores se recomienda teniendo en cuenta la información espacial y no espacial de la presente investigación, actualizar los estudios y la información necesaria para determinar el índice de vulnerabilidad sísmica de las viviendas, y así permitir la actualización de los planes de gestión de riesgo basados en evaluación de resultados.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Beltrán , C. (2018). *Desarrollo de un Modelo Algorítmico Basado En Árboles de Decisión para la Predicción de la Permanencia de un Paciente en un Proceso Psicoterapéutico*. Lima.
- Calle, M. L., & Sánchez-Espigares, J.A. (2007). Árboles de clasificación y regresión en la investigación biomédica. *Medicina clínica*, 129(18),702-706.
- Daza, A. (2016). *Data Mining, Minería de datos*. Perú: Editorial Macro.
- Espino, L., & García, M. (2018). *Aplicación de Minería de Datos basado en Árboles de Decisión para Predecir el Riesgo de Morosidad de los Clientes en la Empresa de Seguros ONCOSALUD S.A.C. 2018*. Lima.
- Ferrero, F., & Martínez, J. (2017). *Minería de Datos Weka* . España.
- Flores, R. (2016). *Modelamiento Geoespacial para la Determinación del Grado de Vulnerabilidad Sísmica del Distrito de Leimebamba-Amazonas*. Tesis para optar el Título Profesional, Universidad Nacional Agraria de la Molina, Facultad de Ingeniería Agrícola, Leimebamba. Recuperado el Marzo de 2021
- GIL, Q., VALENCIA, N. L., & Andrés, C. (2012). Aplicación del proceso de KDD en el contexto de bibliomining. *Revista Interamericana de Bibliotecología*, 97-108.
- Gómez, W., & Loayza, A. (2014). *Evaluación de la vulnerabilidad sísmica de centros de salud del distrito de Ayacucho*. Huancavelica.
- González, C., & García, F. (2010). *Inteligencia en Comunicaciones Minería de Datos*. En Predicción de las condiciones meteorológicas.
- González, L. (2005). *Una arquitectura para el análisis de información que integra procesamiento analítico en línea con minería de datos*. Universidad de las Américas Puebla.
- Hernández, Ferrari, & Ramírez. (2004). *Introducción a la minería de*. Madrid: Pearson Education.

- Ludeña Ocampo, M. M. (2005). *Diseño de puesta en marcha de un Sistema de Información Geográfica de la Comunidad de Madrid para el estudio de la distribución de metales pesados*. Tesis de grado, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Minas. Recuperado el Febrero de 2021
- Maimon, O., & Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook*. Obtenido de In Data Mining and Knowledge Discovery Handbook: http://doi.org/10.1007/0-387-25465-x_2
- MOGOLLÓN ADANAQUÉ, E. J. (2017). *Diagnóstico preliminar de la vulnerabilidad sísmica de las autoconstrucciones de albañilería confinada de ladrillo de arcilla cocida del sector nuevo horizonte de la ciudad de jaén- cajamarca*. Jaén.
- Nanfuñay, H., & Santisteban , E. (2015). *Vulnerabilidad Sísmica en el distrito de ciudad de Eten aplicando índices de vulnerabilidad (Benedetti-Petrini)*. Lambayeque.
- Nervi Laura, M. (2017). *Análisis de la vulnerabilidad sísmica en viviendas de albañilería confinada según la norma E - 070 del RNE en la Ciudad de Juliaca Puno*. Tesis, Universidad Peruana Unión, Puno, Juliaca.
- Quiñones, L., & Carrasco, Y. (2020). *Rendimiento académico empleando minería de datos. Espacios*.
- Ramos, R. (2020). *Vulnerabilidad sísmica de las viviendas de albañilería confinada ubicada en el sector pueblo libre en la ciudad de Jaén, Cajamarca-2020*. Jaén.
- Saavedra, A. (2016). *Estimación del estado del flujo de tráfico mediante preprocesado y minería de datos*. Las Palmas de Gran Canaria.
- Teobaldo, H., & Santisteban, E. (2015). *Vulnerabilidad sísmica en el distrito de ciudad Eten aplicando índices de vulnerabilidad (Benedetti - Petrini)*. Chiclayo.
- Timarán, R., Calderón, A., & Hidalgo, A. (2017). *Aplicación de los árboles de decisión en la identificación de patrones de lesiones*. 11-12.

AGRADECIMIENTO

Agradecer a todas las personas que formaron parte de nuestra formación profesional y personal, así como aquellos que hicieron posible la realización del presente trabajo de investigación.

Agradecer a Dios por permitirnos llegar hasta este punto de nuestra carrera profesional con salud y en compañía de nuestros seres queridos.

A nuestros padres, por su infinito amor, su trabajo y sacrificio, gracias por sus consejos y su apoyo incondicional. La culminación de este proyecto de investigación es el fruto de todo lo que nos han brindado.

Al **Mg. Lenin Quiñones Huatangari** por su asesoría y tiempo brindado durante la ejecución del proyecto de investigación.

Al **M.Sc. Ing. Christiaan Zayed Apaza Panca** por su asesoría y tiempo brindado durante el curso de seminario de investigación científica.

A nuestra casa superior de estudios, la **Universidad Nacional de Jaén** por brindarnos una enseñanza de calidad, por habernos formado profesionalmente y por el apoyo brindado para lograr nuestro proyecto de investigación.

A todos nuestros compañeros que nos brindaron su apoyo de manera directa e indirecta en esta labor y nos permitieron realizarla.

DEDICATORIA

A nuestros **PADRES:**

Quienes nos brindan su apoyo incondicional, nos guían desde pequeños, son nuestros primeros maestros de quienes aprendemos a superar las adversidades y nos enseñan lecciones valiosas día a día y son el pilar fundamental de nuestras vidas por ello siempre estaremos orgullosos, porque sin ellos nada de esto sería realidad; les estaremos infinitamente agradecidos.

A Dios, quien supo guiarnos en todo momento por el buen camino, por darnos fuerzas para seguir adelante y no desmayar en los problemas que se nos presentan, enseñándonos a superar las adversidades sin desfallecer en el intento.

A nuestros **TÍOS Y HERMANOS:**

Por el gran cariño y apoyo hacia nosotros, por sus consejos, sus palabras de aliento, que fueron de gran ayuda para poder realizar este proyecto; se les estima mucho.

ANEXOS

ANEXO I.
RESUMEN DE MODELOS
OBTENIDOS

➤ **Modelo MDL_1**

✓ **Decision Stump**

Anexo 1.

MDL_1-Resumen algoritmo Decision Stump

=== Summary ===		
Correctly Classified Instances	744	53.2951 %
Incorrectly Classified Instances	652	46.7049 %
Kappa statistic	0.1298	
Mean absolute error	0.3802	
Root mean squared error	0.438	
Relative absolute error	94.2768 %	
Root relative squared error	97.5615 %	
Total Number of Instances	1396	

✓ **Hoeffding Tree**

Anexo 2.

MDL_1-Resumen algoritmo Hoeffding Tree

=== Summary ===		
Correctly Classified Instances	946	67.765 %
Incorrectly Classified Instances	450	32.235 %
Kappa statistic	0.4521	
Mean absolute error	0.2851	
Root mean squared error	0.3775	
Relative absolute error	70.6904 %	
Root relative squared error	84.0805 %	
Total Number of Instances	1396	

✓ **J48**

Anexo 3.

MDL_1- Resumen algoritmo J48

=== Summary ===		
Correctly Classified Instances	940	67.3352 %
Incorrectly Classified Instances	456	32.6648 %
Kappa statistic	0.4469	
Mean absolute error	0.2859	
Root mean squared error	0.383	
Relative absolute error	70.8979 %	
Root relative squared error	85.3008 %	
Total Number of Instances	1396	

✓ LMT

Anexo 4.

MDL_1-Resumen algoritmo LMT

=== Summary ===		
Correctly Classified Instances	964	69.0544 %
Incorrectly Classified Instances	432	30.9456 %
Kappa statistic	0.4728	
Mean absolute error	0.2819	
Root mean squared error	0.3733	
Relative absolute error	69.8903 %	
Root relative squared error	83.1346 %	
Total Number of Instances	1396	

✓ Random Tree

Anexo 5.

MDL_1-Resumen algoritmo Random Tree

=== Summary ===		
Correctly Classified Instances	947	67.8367 %
Incorrectly Classified Instances	449	32.1633 %
Kappa statistic	0.4532	
Mean absolute error	0.2719	
Root mean squared error	0.3759	
Relative absolute error	67.4111 %	
Root relative squared error	83.7102 %	
Total Number of Instances	1396	

➤ **Modelo MDL_2**

✓ Decision Stump

Anexo 6.

MDL_2-Resumen algoritmo Decision Stump

=== Summary ===		
Correctly Classified Instances	753	53.9398 %
Incorrectly Classified Instances	643	46.0602 %
Kappa statistic	0.1666	
Mean absolute error	0.3775	
Root mean squared error	0.4354	
Relative absolute error	93.6143 %	
Root relative squared error	96.9687 %	
Total Number of Instances	1396	

✓ Hoeffding Tree

Anexo 7.

MDL_2-Resumen algoritmo Hoeffding Tree

=== Summary ===		
Correctly Classified Instances	876	62.7507 %
Incorrectly Classified Instances	520	37.2493 %
Kappa statistic	0.3606	
Mean absolute error	0.2993	
Root mean squared error	0.3956	
Relative absolute error	74.2259 %	
Root relative squared error	88.1027 %	
Total Number of Instances	1396	

✓ J48

Anexo 8.

MDL_2- Resumen algoritmo J48

=== Summary ===		
Correctly Classified Instances	900	64.4699 %
Incorrectly Classified Instances	496	35.5301 %
Kappa statistic	0.3842	
Mean absolute error	0.3073	
Root mean squared error	0.3954	
Relative absolute error	76.188 %	
Root relative squared error	88.056 %	
Total Number of Instances	1396	

✓ LMT

Anexo 9.

MDL_2-Resumen algoritmo LMT

=== Summary ===		
Correctly Classified Instances	898	64.3266 %
Incorrectly Classified Instances	498	35.6734 %
Kappa statistic	0.3719	
Mean absolute error	0.3098	
Root mean squared error	0.3927	
Relative absolute error	76.8065 %	
Root relative squared error	87.4679 %	
Total Number of Instances	1396	

✓ Random Tree

Anexo 10.

MDL_2-Resumen algoritmo Random Tree

=== Summary ===		
Correctly Classified Instances	894	64.0401 %
Incorrectly Classified Instances	502	35.9599 %
Kappa statistic	0.3828	
Mean absolute error	0.3028	
Root mean squared error	0.3973	
Relative absolute error	75.0823 %	
Root relative squared error	88.4977 %	
Total Number of Instances	1396	

➤ **Modelo MDL_3**

✓ Decision Stump

Anexo 11.

MDL_3-Resumen algoritmo Decision Stump

=== Summary ===		
Correctly Classified Instances	740	53.0086 %
Incorrectly Classified Instances	656	46.9914 %
Kappa statistic	0	
Mean absolute error	0.3751	
Root mean squared error	0.4332	
Relative absolute error	93.0033 %	
Root relative squared error	96.4721 %	
Total Number of Instances	1396	

✓ Hoeffding Tree

Anexo 12.

MDL_3-Resumen algoritmo Hoeffding Tree

=== Summary ===		
Correctly Classified Instances	952	68.1948 %
Incorrectly Classified Instances	444	31.8052 %
Kappa statistic	0.4324	
Mean absolute error	0.2914	
Root mean squared error	0.3776	
Relative absolute error	72.2606 %	
Root relative squared error	84.1062 %	
Total Number of Instances	1396	

✓ J48

Anexo13.

MDL_3- Resumen algoritmo J48

```
=== Summary ===
```

Correctly Classified Instances	942	67.4785 %
Incorrectly Classified Instances	454	32.5215 %
Kappa statistic	0.4262	
Mean absolute error	0.2888	
Root mean squared error	0.3867	
Relative absolute error	71.6005 %	
Root relative squared error	86.117 %	
Total Number of Instances	1396	

✓ LMT

Anexo 14.

MDL_3-Resumen algoritmo LMT

```
=== Summary ===
```

Correctly Classified Instances	950	68.0516 %
Incorrectly Classified Instances	446	31.9484 %
Kappa statistic	0.4354	
Mean absolute error	0.2811	
Root mean squared error	0.3768	
Relative absolute error	69.7037 %	
Root relative squared error	83.9138 %	
Total Number of Instances	1396	

✓ Random Tree

Anexo 15.

MDL_3-Resumen algoritmo Random Tree

```
=== Summary ===
```

Correctly Classified Instances	952	68.1948 %
Incorrectly Classified Instances	444	31.8052 %
Kappa statistic	0.4363	
Mean absolute error	0.2739	
Root mean squared error	0.3764	
Relative absolute error	67.9143 %	
Root relative squared error	83.8387 %	
Total Number of Instances	1396	

➤ **Modelo MDL_4**

✓ **Decision Stump**

Anexo 16.

MDL_4-Resumen algoritmo Decision Stump

Correctly Classified Instances	750	53.7249 %
Incorrectly Classified Instances	646	46.2751 %
Kappa statistic	0.1797	
Mean absolute error	0.3677	
Root mean squared error	0.4293	
Relative absolute error	91.1786 %	
Root relative squared error	95.624 %	
Total Number of Instances	1396	

✓ **Hoeffding Tree**

Anexo 17.

MDL_4-Resumen algoritmo Hoeffding Tree

=== Summary ===		
Correctly Classified Instances	1182	84.6705 %
Incorrectly Classified Instances	214	15.3295 %
Kappa statistic	0.7424	
Mean absolute error	0.1727	
Root mean squared error	0.2794	
Relative absolute error	42.8205 %	
Root relative squared error	62.2337 %	
Total Number of Instances	1396	

✓ **J48**

Anexo 18.

MDL_4- Resumen algoritmo J48

Correctly Classified Instances	1145	82.0201 %
Incorrectly Classified Instances	251	17.9799 %
Kappa statistic	0.6991	
Mean absolute error	0.1697	
Root mean squared error	0.3145	
Relative absolute error	42.0902 %	
Root relative squared error	70.0482 %	
Total Number of Instances	1396	

✓ LMT

Anexo 19.

MDL_4-Resumen algoritmo LMT

=== Summary ===		
Correctly Classified Instances	1219	87.3209 %
Incorrectly Classified Instances	177	12.6791 %
Kappa statistic	0.788	
Mean absolute error	0.1447	
Root mean squared error	0.2523	
Relative absolute error	35.8912 %	
Root relative squared error	56.186 %	
Total Number of Instances	1396	

✓ Random Tree

Anexo 20.

MDL_4-Resumen algoritmo Random Tree

=== Summary ===		
Correctly Classified Instances	1146	82.0917 %
Incorrectly Classified Instances	250	17.9083 %
Kappa statistic	0.7021	
Mean absolute error	0.1354	
Root mean squared error	0.3177	
Relative absolute error	33.5744 %	
Root relative squared error	70.7489 %	
Total Number of Instances	1396	

➤ **Modelo MDL_5**

✓ Decision Stump

Anexo 21.

MDL_5-Resumen algoritmo Decision Stump

=== Summary ===		
Correctly Classified Instances	750	53.7249 %
Incorrectly Classified Instances	646	46.2751 %
Kappa statistic	0.1797	
Mean absolute error	0.3677	
Root mean squared error	0.4293	
Relative absolute error	91.1786 %	
Root relative squared error	95.624 %	
Total Number of Instances	1396	

✓ Hoeffding Tree

Anexo 22.

MDL_5-Resumen algoritmo Hoeffding Tree

=== Summary ===		
Correctly Classified Instances	1124	80.5158 %
Incorrectly Classified Instances	272	19.4842 %
Kappa statistic	0.6707	
Mean absolute error	0.2008	
Root mean squared error	0.3076	
Relative absolute error	49.7972 %	
Root relative squared error	68.5007 %	
Total Number of Instances	1396	

✓ J48

Anexo 23.

MDL_5- Resumen algoritmo J48

=== Summary ===		
Correctly Classified Instances	1113	79.7278 %
Incorrectly Classified Instances	283	20.2722 %
Kappa statistic	0.6629	
Mean absolute error	0.1973	
Root mean squared error	0.3281	
Relative absolute error	48.9126 %	
Root relative squared error	73.071 %	
Total Number of Instances	1396	

✓ LMT

Anexo 24.

MDL_4-Resumen algoritmo LMT

=== Summary ===		
Correctly Classified Instances	1105	79.1547 %
Incorrectly Classified Instances	291	20.8453 %
Kappa statistic	0.6532	
Mean absolute error	0.1652	
Root mean squared error	0.3311	
Relative absolute error	40.9531 %	
Root relative squared error	73.7517 %	
Total Number of Instances	1396	

✓ Random Tree

Anexo 25.

MDL_5-Resumen algoritmo Random Tree

=== Summary ===		
Correctly Classified Instances	1151	82.4499 %
Incorrectly Classified Instances	245	17.5501 %
Kappa statistic	0.7085	
Mean absolute error	0.1707	
Root mean squared error	0.2876	
Relative absolute error	42.339 %	
Root relative squared error	64.0486 %	
Total Number of Instances	1396	

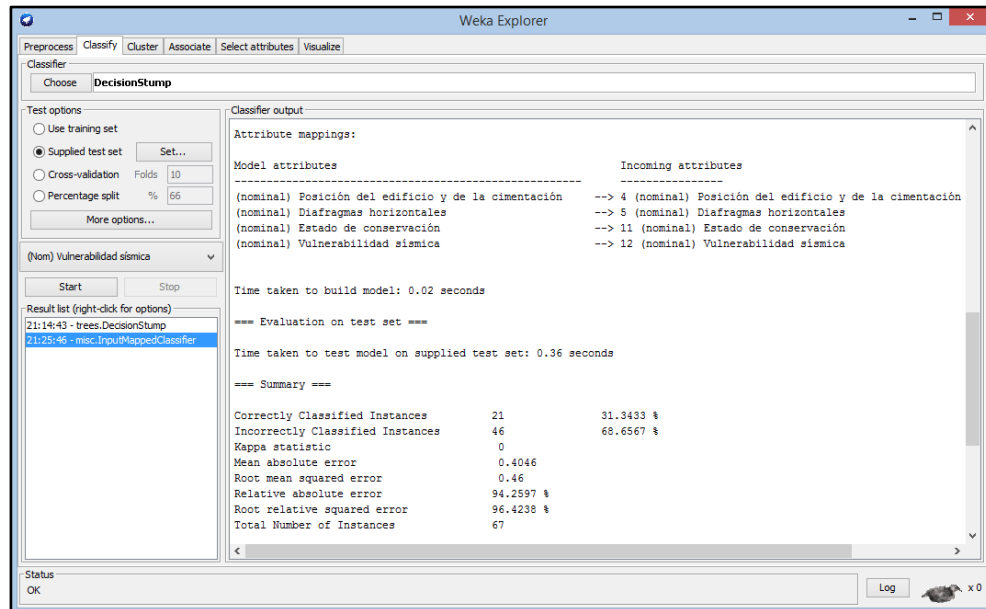
ANEXO II.
PREDICCIÓN DE LOS
MODELOS DE MINERÍA
DE DATOS

➤ MDL_1

✓ Decision Stump

Anexo 26.

MDL_1-Predicción del algoritmo Decision Stump



The screenshot shows the Weka Explorer interface with the Decision Stump classifier selected. The classifier output window displays the following information:

Attribute mappings:

Model attributes	Incoming attributes
(nominal) Posición del edificio y de la cimentación	--> 4 (nominal) Posición del edificio y de la cimentación
(nominal) Diafragmas horizontales	--> 5 (nominal) Diafragmas horizontales
(nominal) Estado de conservación	--> 11 (nominal) Estado de conservación
(nominal) Vulnerabilidad sísmica	--> 12 (nominal) Vulnerabilidad sísmica

Time taken to build model: 0.02 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.36 seconds

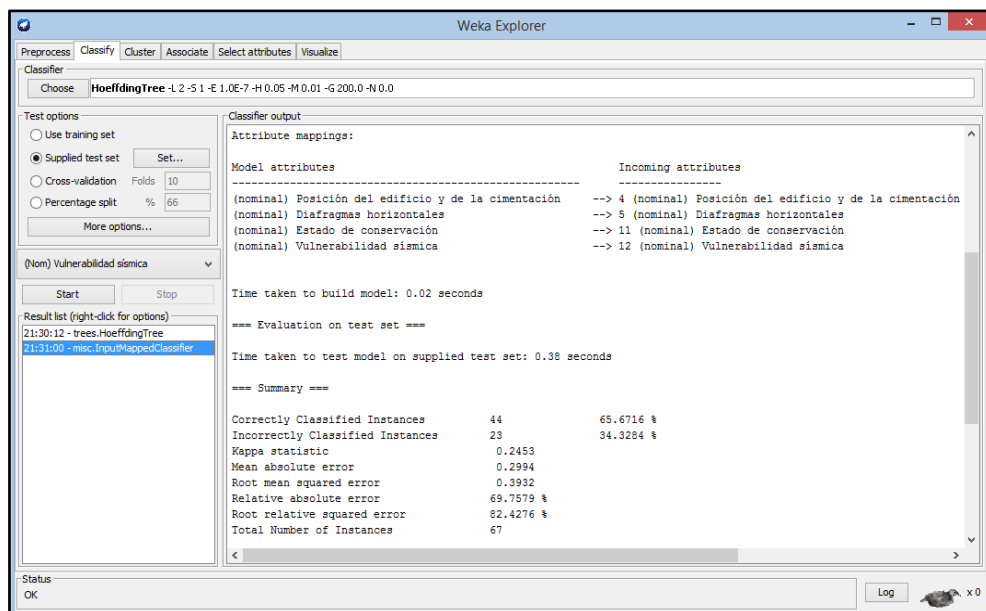
=== Summary ===

Correctly Classified Instances	21	31.3433 %
Incorrectly Classified Instances	46	68.6567 %
Kappa statistic	0	
Mean absolute error	0.4046	
Root mean squared error	0.46	
Relative absolute error	94.2597 %	
Root relative squared error	96.4238 %	
Total Number of Instances	67	

✓ Hoeffding Tree

Anexo 27.

MDL_1- Predicción algoritmo Hoeffding Tree



The screenshot shows the Weka Explorer interface with the Hoeffding Tree classifier selected. The classifier output window displays the following information:

Attribute mappings:

Model attributes	Incoming attributes
(nominal) Posición del edificio y de la cimentación	--> 4 (nominal) Posición del edificio y de la cimentación
(nominal) Diafragmas horizontales	--> 5 (nominal) Diafragmas horizontales
(nominal) Estado de conservación	--> 11 (nominal) Estado de conservación
(nominal) Vulnerabilidad sísmica	--> 12 (nominal) Vulnerabilidad sísmica

Time taken to build model: 0.02 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.38 seconds

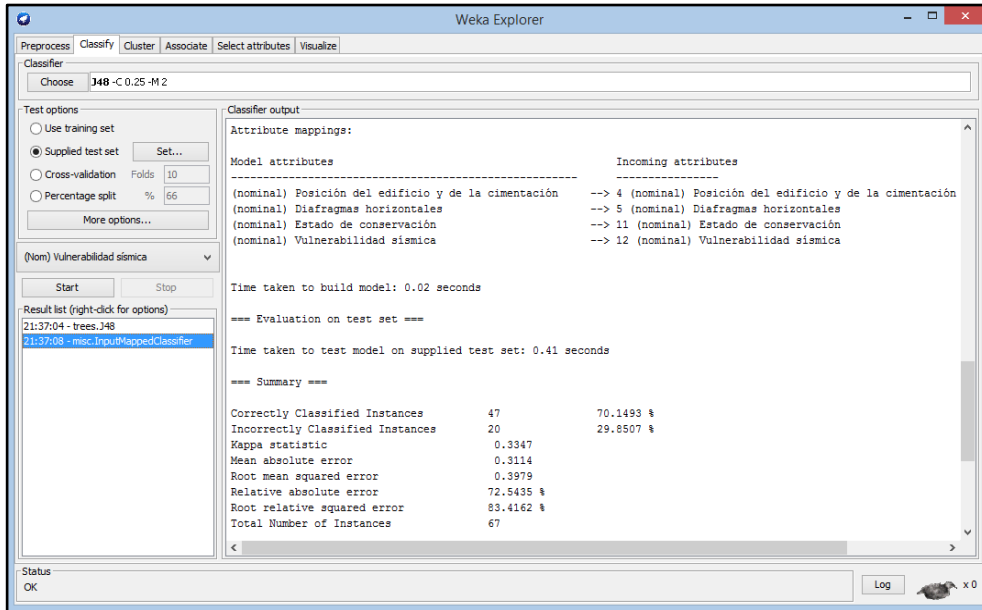
=== Summary ===

Correctly Classified Instances	44	65.6716 %
Incorrectly Classified Instances	23	34.3284 %
Kappa statistic	0.2453	
Mean absolute error	0.2994	
Root mean squared error	0.3932	
Relative absolute error	69.7579 %	
Root relative squared error	82.4276 %	
Total Number of Instances	67	

✓ J48

Anexo 28.

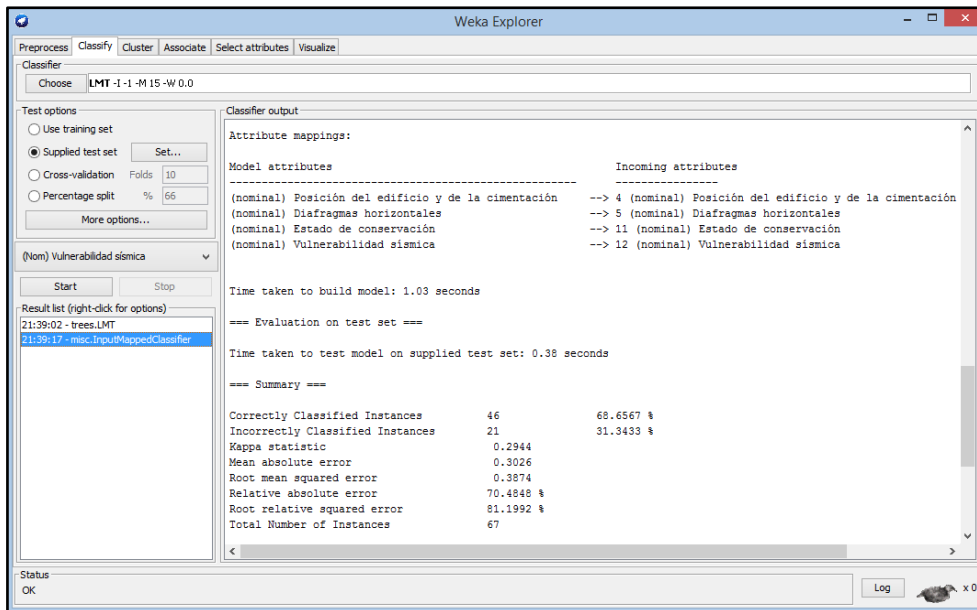
MDL_1- Predicción algoritmo J48



✓ LMT

Anexo 29.

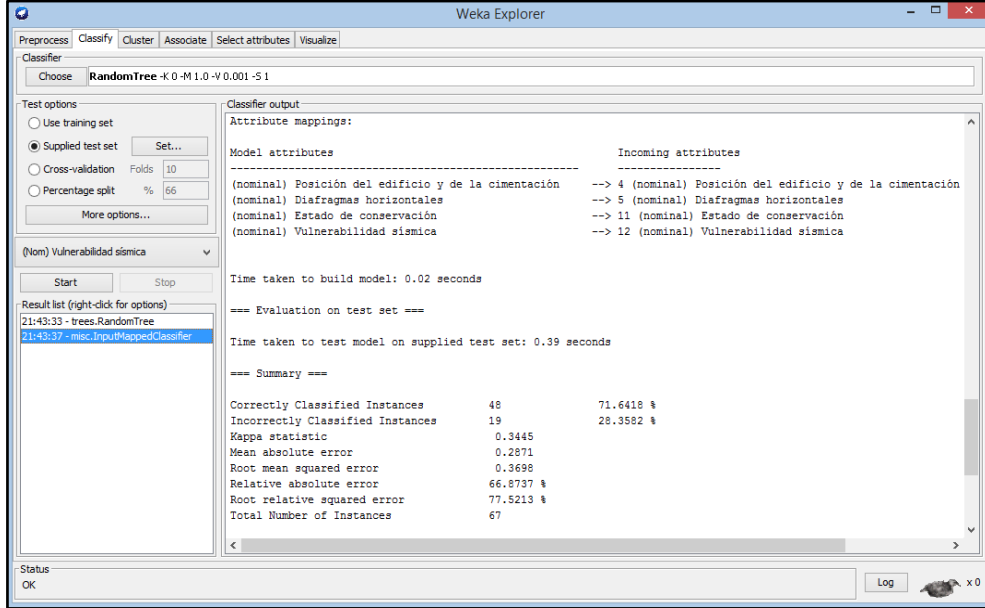
MDL_1- Predicción algoritmo LMT



✓ Random Tree

Anexo 30.

MDL_1- Predicción del algoritmo Random Tree

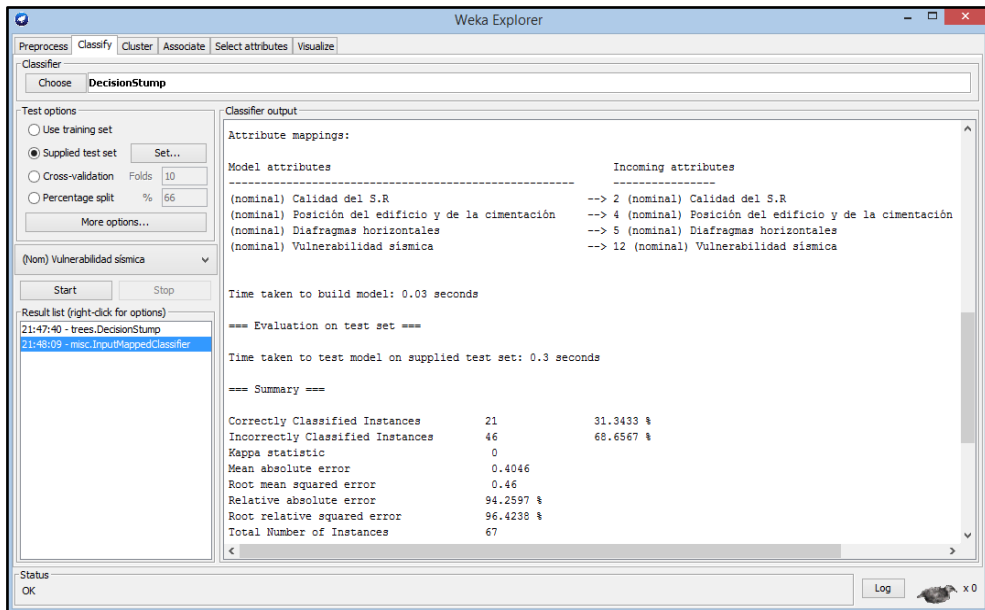


➤ MDL_4

✓ Decision Stump

Anexo 31.

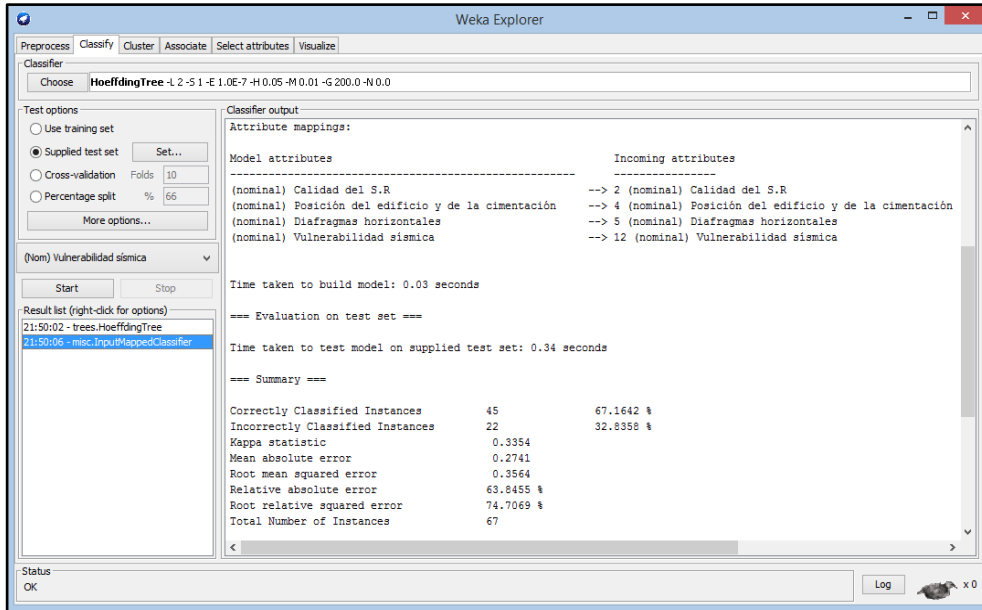
MDL_4-Predicción del algoritmo Decision Stump



✓ Hoeffding Tree

Anexo 32.

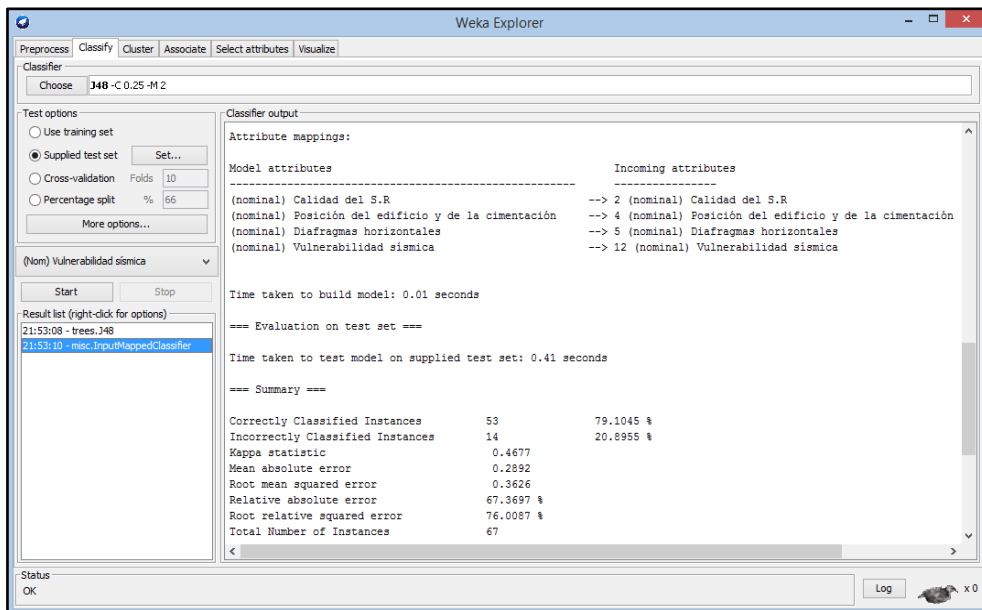
MDL_4-Predicción del algoritmo Hoeffding Tree



✓ J48

Anexo 33.

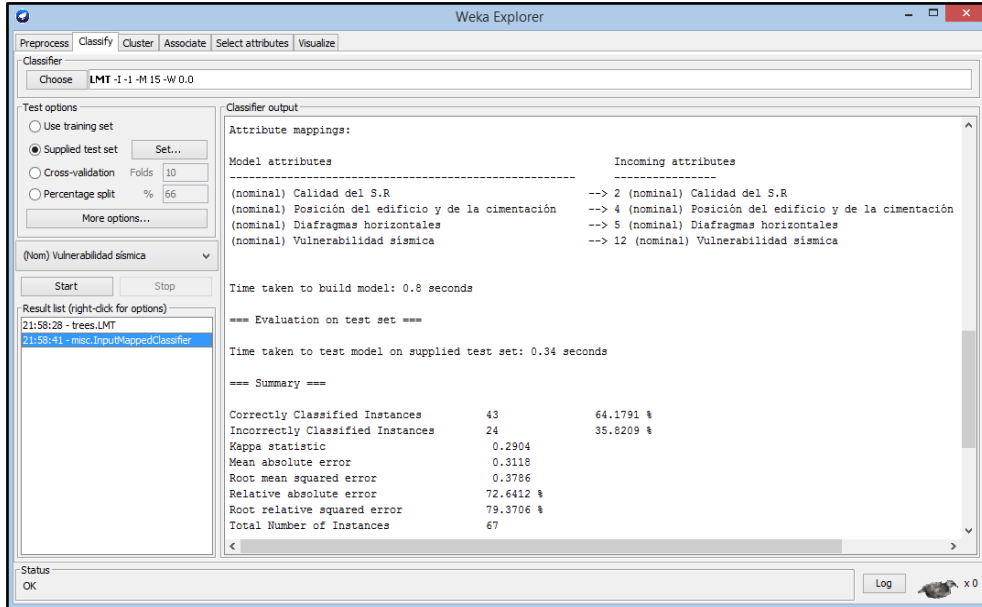
MDL_4- Predicción algoritmo J48



✓ LMT

Anexo 34.

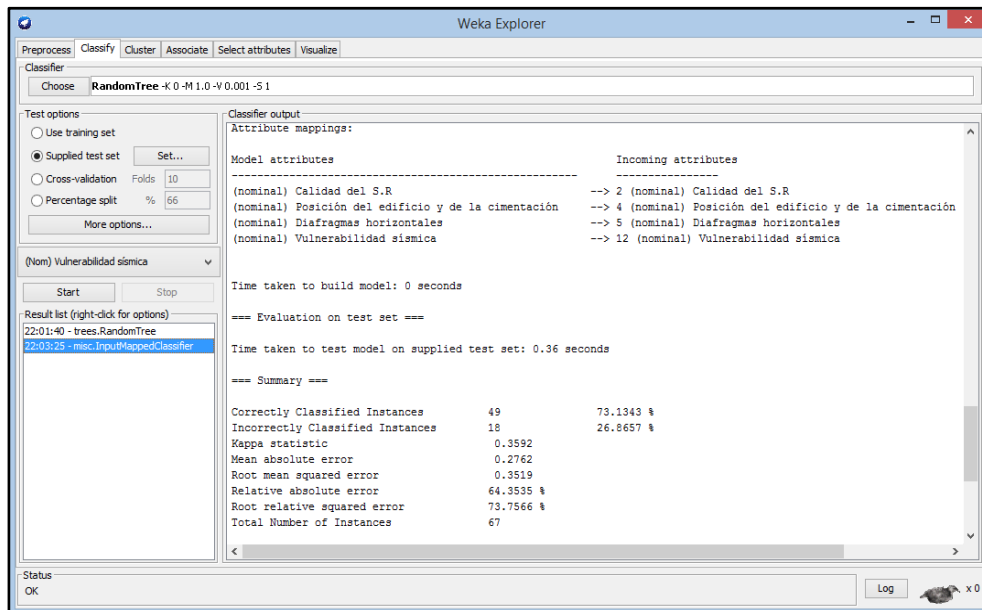
MDL_4- Predicción algoritmo LMT



✓ Random Tree

Anexo 35.

MDL_4- Predicción del algoritmo Random Tree

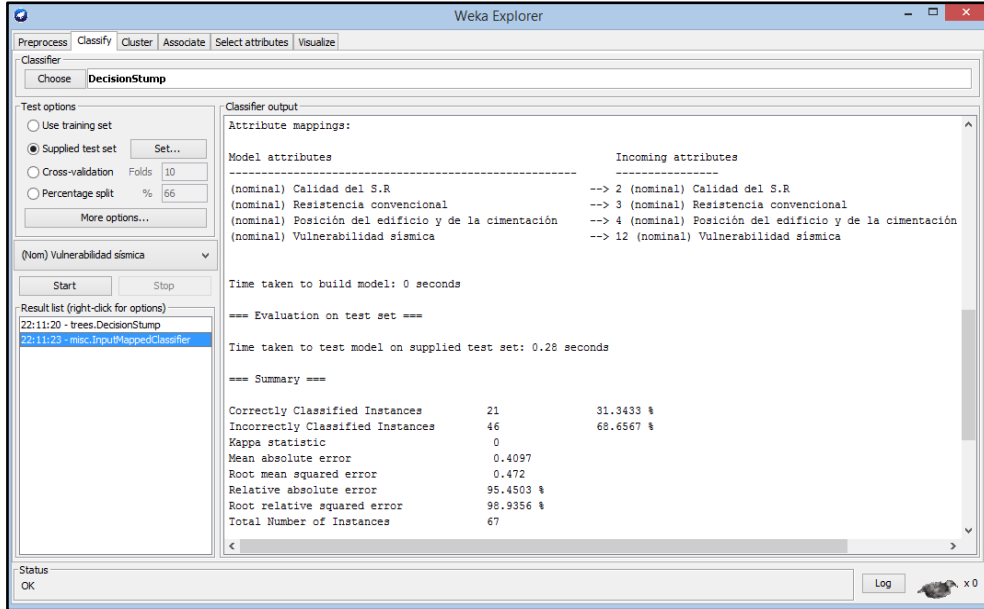


➤ MDL_5

✓ Decision Stump

Anexo 36.

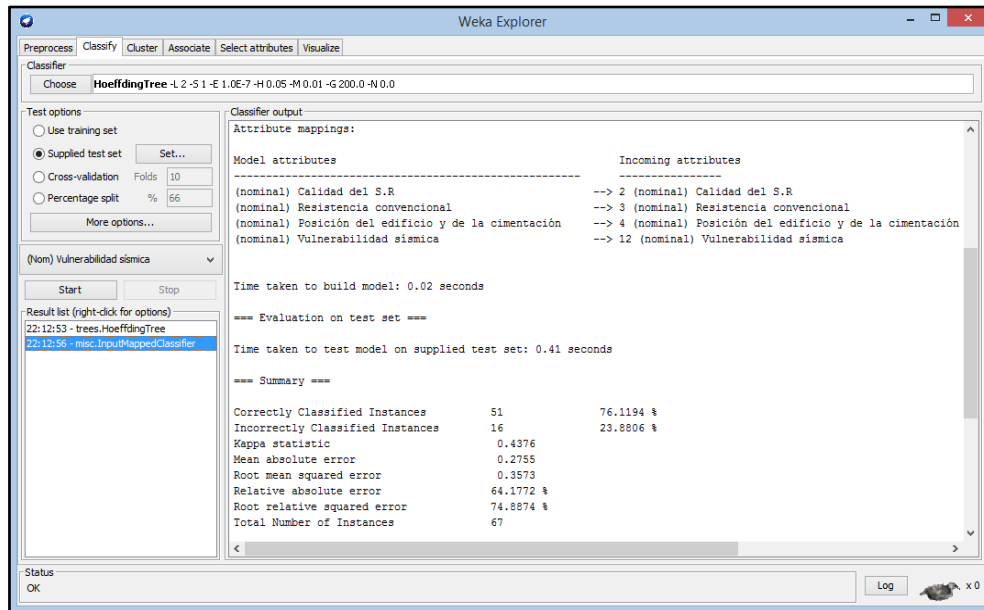
MDL_5- Predicción del algoritmo Decision Stump



✓ Hoeffding Tree

Anexo 37.

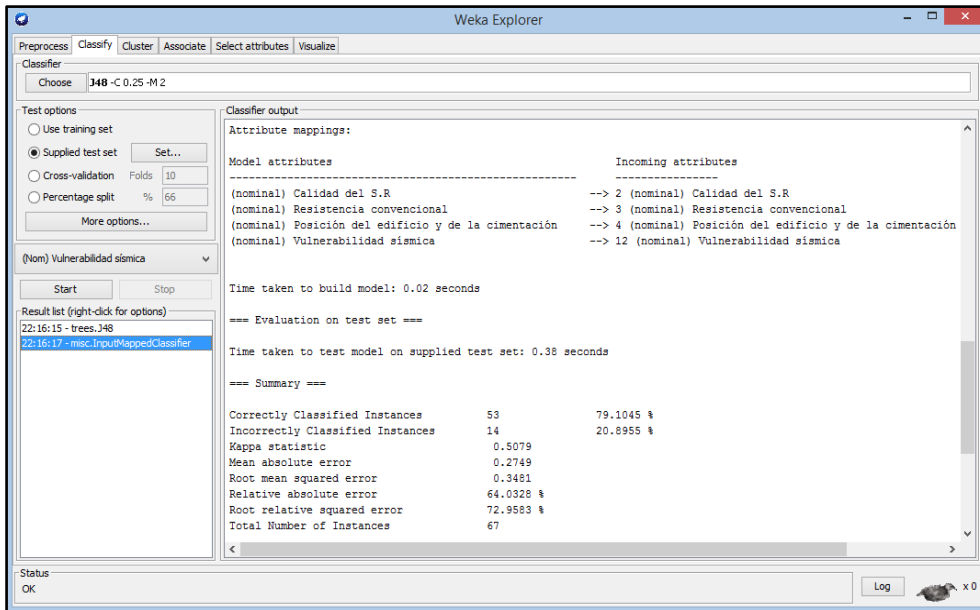
MDL_5-Predicción del algoritmo Hoeffding Tree



✓ J48

Anexo 38.

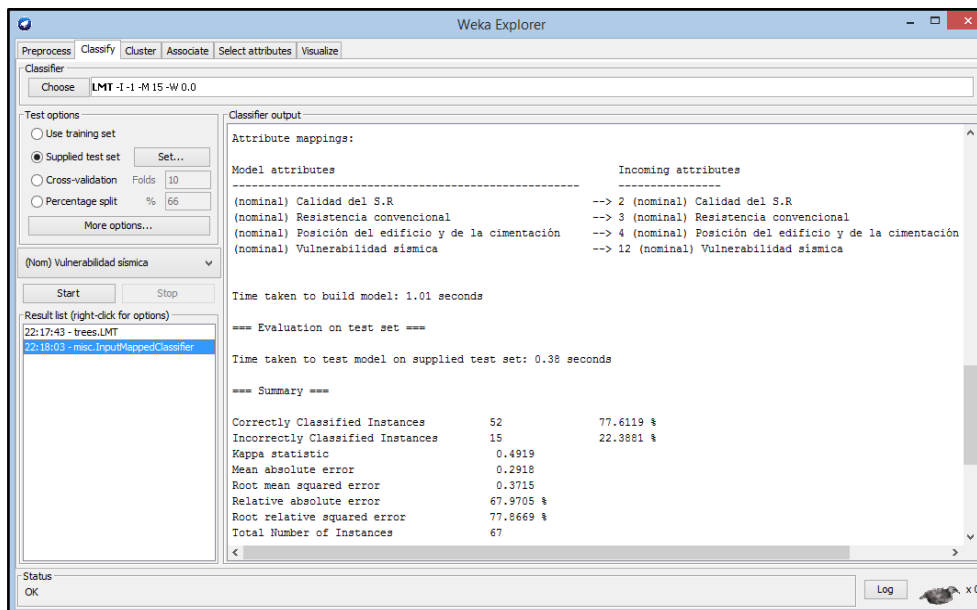
MDL_5- Predicción algoritmo J48



✓ LMT

Anexo 39.

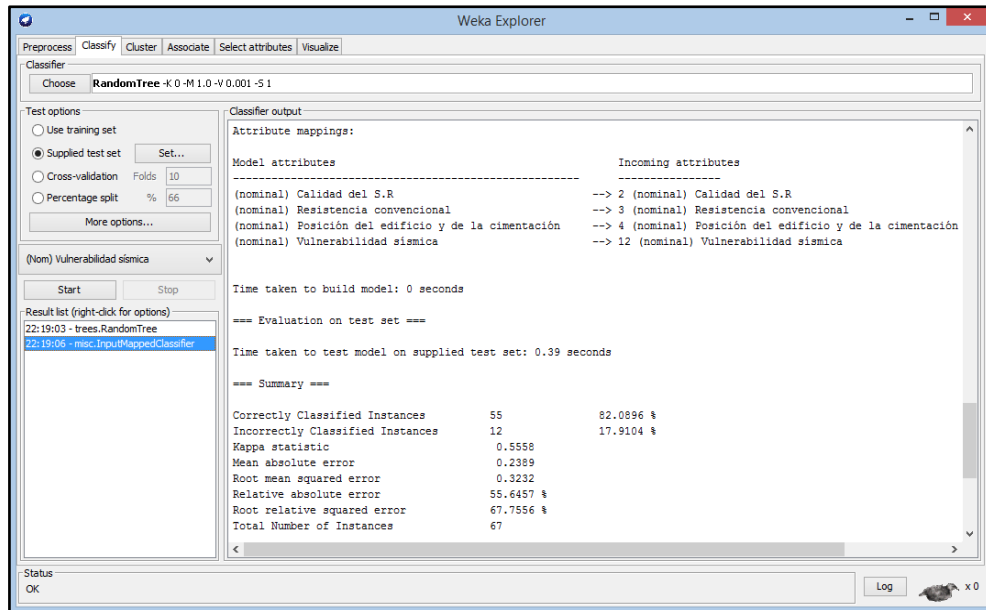
MDL_5- Predicción algoritmo LMT



✓ Random Tree

Anexo 40.

MDL_5- Predicción del algoritmo Random Tree



ANEXO III.
ALGORITMO
RAMDOM-TREE

Anexo 41.

Algoritmo que predice la vulnerabilidad sísmica

```
RandomTree
=====

Posición del edificio y de la cimentación = B
|
|  Calidad del S.R = B
|  |
|  |  Resistencia convencional = D : MEDIA
|  |  Resistencia convencional = B : BAJA
|  |  Resistencia convencional = C : BAJA
|  |  Resistencia convencional = A : BAJA
|  |
|  |  Calidad del S.R = C
|  |  |
|  |  |  Resistencia convencional = D : MEDIA
|  |  |  Resistencia convencional = B : BAJA
|  |  |  Resistencia convencional = C : MEDIA
|  |  |  Resistencia convencional = A : MEDIA
|  |  |
|  |  |  Calidad del S.R = D
|  |  |  |
|  |  |  |  Resistencia convencional = D : MEDIA
|  |  |  |  Resistencia convencional = B : BAJA
|  |  |  |  Resistencia convencional = C : MEDIA
|  |  |  |  Resistencia convencional = A : MEDIA
|  |  |  |
|  |  |  |  Calidad del S.R = A
|  |  |  |  |
|  |  |  |  |  Resistencia convencional = D : MEDIA
|  |  |  |  |  Resistencia convencional = B : BAJA
|  |  |  |  |  Resistencia convencional = C : MEDIA
|  |  |  |  |  Resistencia convencional = A : BAJA
|  |  |  |
|  |  |  |  Posición del edificio y de la cimentación = C
|  |  |  |  |
|  |  |  |  |  Resistencia convencional = D
|  |  |  |  |  |
|  |  |  |  |  |  Calidad del S.R = B : MEDIA
|  |  |  |  |  |  Calidad del S.R = C : ALTA
|  |  |  |  |  |  Calidad del S.R = D : ALTA
|  |  |  |  |  |  Calidad del S.R = A : MEDIA
|  |  |  |  |  |
|  |  |  |  |  |  Resistencia convencional = B
|  |  |  |  |  |  |
|  |  |  |  |  |  |  Calidad del S.R = B : MEDIA
|  |  |  |  |  |  |  Calidad del S.R = C : MEDIA
|  |  |  |  |  |  |  Calidad del S.R = D : MEDIA
|  |  |  |  |  |  |  Calidad del S.R = A : MEDIA
|  |  |  |  |  |  |
|  |  |  |  |  |  |  Resistencia convencional = C
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  Calidad del S.R = B : MEDIA
|  |  |  |  |  |  |  |  Calidad del S.R = C : MEDIA
|  |  |  |  |  |  |  |  Calidad del S.R = D : ALTA
|  |  |  |  |  |  |  |  Calidad del S.R = A : BAJA
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  Resistencia convencional = A
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  Calidad del S.R = B : MEDIA
|  |  |  |  |  |  |  |  |  Calidad del S.R = C : MEDIA
|  |  |  |  |  |  |  |  |  Calidad del S.R = D : MEDIA
|  |  |  |  |  |  |  |  |  Calidad del S.R = A : MEDIA
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  Posición del edificio y de la cimentación = A
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  Calidad del S.R = B
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  Resistencia convencional = D : MEDIA
|  |  |  |  |  |  |  |  |  |  Resistencia convencional = B : BAJA
|  |  |  |  |  |  |  |  |  |  Resistencia convencional = C : MEDIA
|  |  |  |  |  |  |  |  |  |  Resistencia convencional = A : BAJA
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  Calidad del S.R = C : MEDIA
|  |  |  |  |  |  |  |  |  |  Calidad del S.R = D : MEDIA
|  |  |  |  |  |  |  |  |  |  Calidad del S.R = A : MEDIA
|  |  |  |  |  |  |  |  |  |
```

```

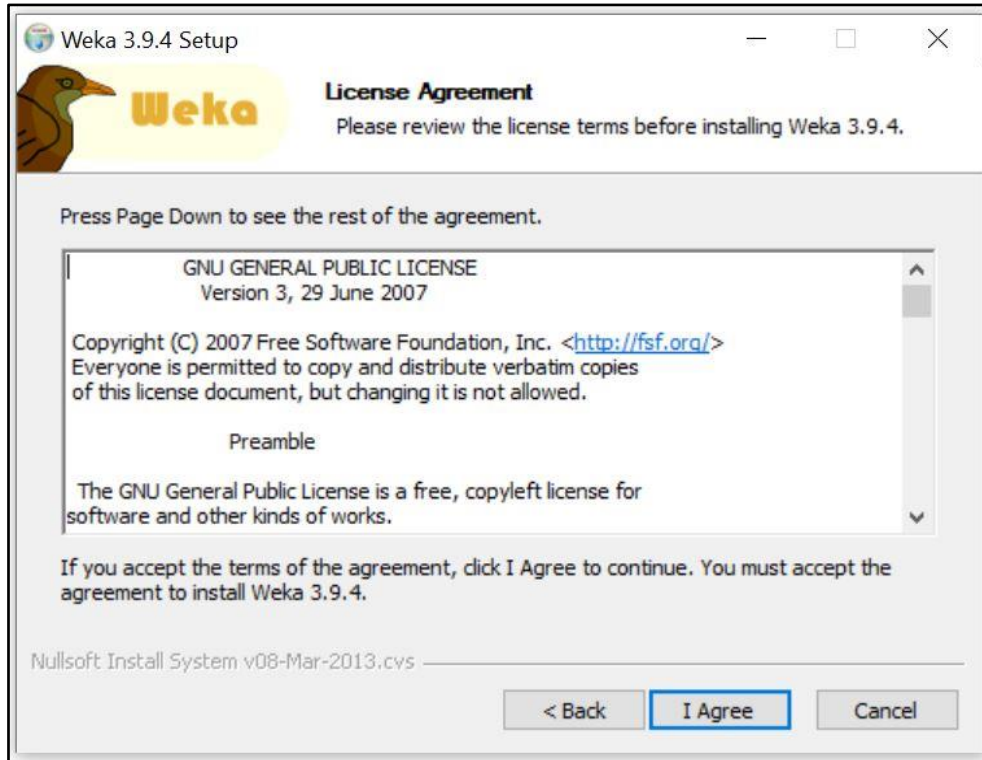
Posición del edificio y de la cimentación = D
| Resistencia convencional = D
| | Calidad del S.R = B : MEDIA
| | Calidad del S.R = C : ALTA
| | Calidad del S.R = D : ALTA
| | Calidad del S.R = A : MEDIA
| Resistencia convencional = B
| | Calidad del S.R = B : MEDIA
| | Calidad del S.R = C : MEDIA
| | Calidad del S.R = D : ALTA
| | Calidad del S.R = A : MEDIA
| Resistencia convencional = C
| | Calidad del S.R = B : MEDIA
| | Calidad del S.R = C : ALTA
| | Calidad del S.R = D : ALTA
| | Calidad del S.R = A : MEDIA
| Resistencia convencional = A
| | Calidad del S.R = B : MEDIA
| | Calidad del S.R = C : MEDIA
| | Calidad del S.R = D : ALTA
| | Calidad del S.R = A : MEDIA

```


ANEXO IV.
LICENCIA DE SOFTWARE

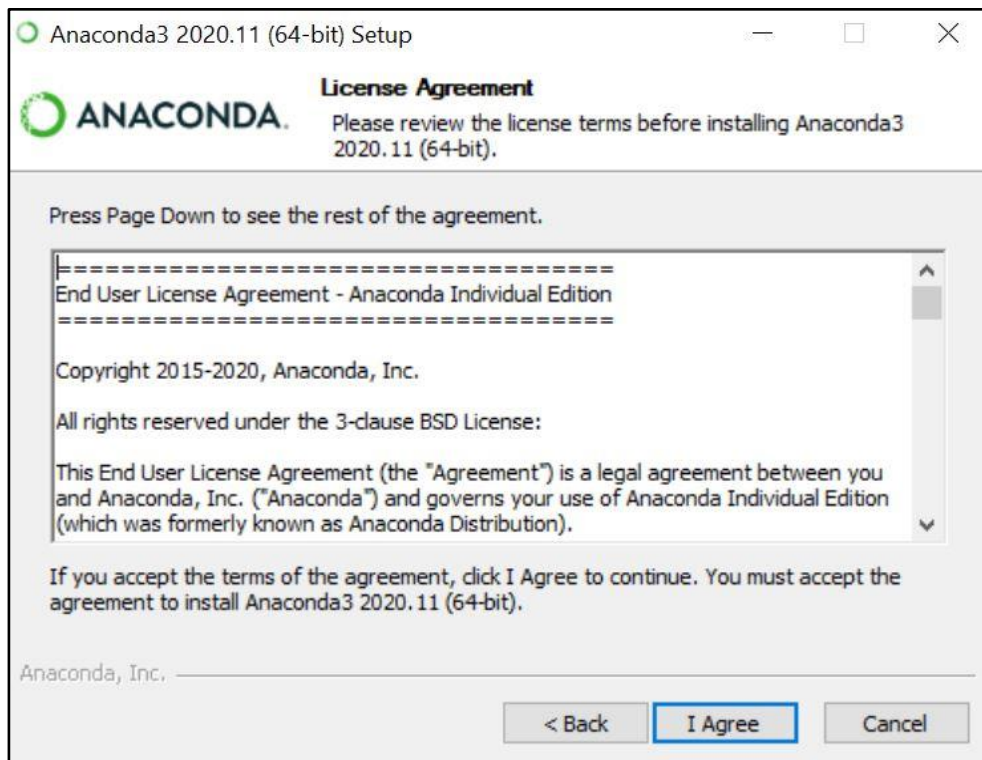
Anexo 42.

Licencia del software Weka 3.9.4.



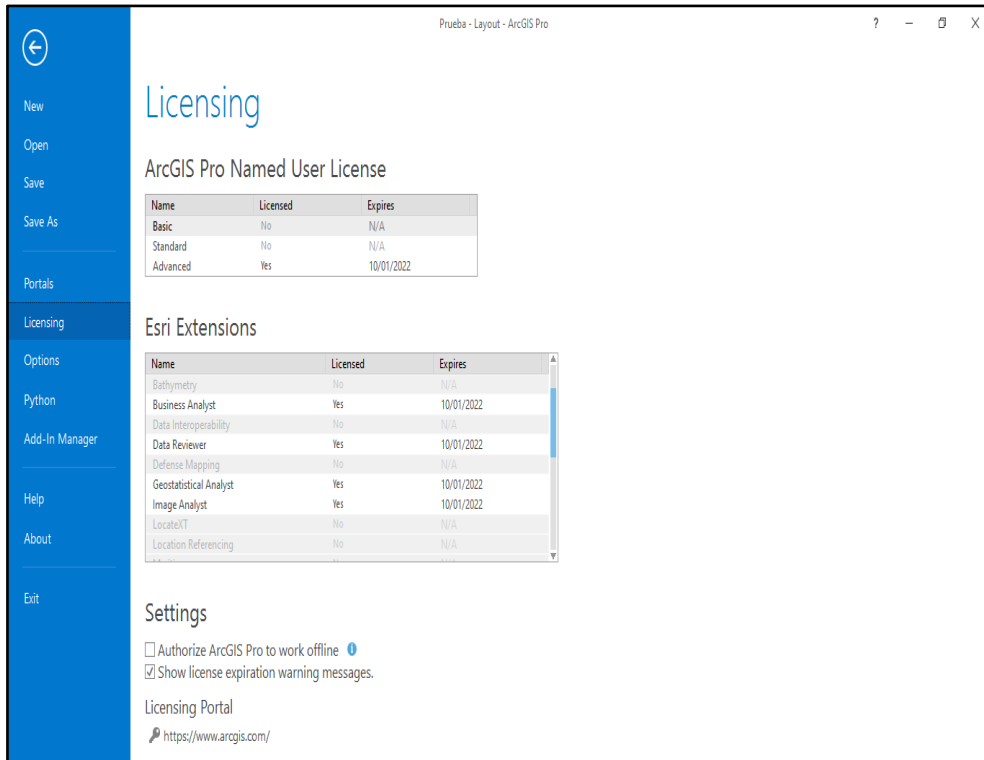
Anexo 43.

Licencia del software Anaconda-Phyton



Anexo 44.

Licencia del software ArcGIS Pro



The screenshot shows the 'Licensing' window in ArcGIS Pro. The window title is 'Prueba - Layout - ArcGIS Pro'. The main heading is 'Licensing'. Below this, there are three sections: 'ArcGIS Pro Named User License', 'Esri Extensions', and 'Settings'.

ArcGIS Pro Named User License

Name	Licensed	Expires
Basic	No	N/A
Standard	No	N/A
Advanced	Yes	10/01/2022

Esri Extensions

Name	Licensed	Expires
Bathymetry	No	N/A
Business Analyst	Yes	10/01/2022
Data Interoperability	No	N/A
Data Reviewer	Yes	10/01/2022
Defense Mapping	No	N/A
Geostatistical Analyst	Yes	10/01/2022
Image Analyst	Yes	10/01/2022
LocateXT	No	N/A
Location Referencing	No	N/A

Settings

- Authorize ArcGIS Pro to work offline ⓘ
- Show license expiration warning messages.

Licensing Portal

<https://www.arcgis.com/>