



**ESTIMACIÓN DEL PARÁMETRO SOPLOSIDAD DENTRO DEL MÉTODO DE
ANÁLISIS PERCEPTUAL DE LA VOZ CAPE-V.**

MARÍA ALEJANDRA RIVERA MARÍN

**UNIVERSIDAD ANTONIO NARIÑO
FACULTAD DE INGENIERÍA MECATRÓNICA, ELECTRÓNICA Y BIOMÉDICA
PROGRAMA DE INGENIERÍA BIOMÉDICA
POPAYÁN
2020**

**ESTIMACIÓN DEL PARÁMETRO SOPLOSIDAD DENTRO DEL MÉTODO DE
ANÁLISIS PERCEPTUAL DE LA VOZ CAPE-V.**

MARÍA ALEJANDRA RIVERA MARÍN

**PROYECTO DE GRADO PRESENTADO COMO REQUISITO PARCIAL PARA
OPTAR AL TÍTULO DE:
INGENIERO BIOMÉDICO**

DIRECTOR:

MSc. JESÚS DAVID RAMÍREZ MEDINA

LÍNEA DE INVESTIGACIÓN:

PROCESAMIENTO DIGITAL DE SEÑALES ACÚSTICAS.

GRUPO DE INVESTIGACIÓN:

BIOINGENIERÍA

UNIVERSIDAD ANTONIO NARIÑO

FACULTAD DE INGENIERÍA MECATRÓNICA, ELECTRÓNICA Y BIOMÉDICA

PROGRAMA DE INGENIERÍA BIOMÉDICA

POPAYÁN

2020

NOTA DE ACEPTACIÓN

FIRMA DEL DIRECTOR

FIRMA DEL JURADO

FIRMA DEL JURADO

Popayán, mayo de 2020.

(Dedicatoria)

Quiero dedicar este trabajo primeramente a Dios por derramar sus bendiciones sobre mí y llenarme de su fuerza para vencer todos los obstáculos desde el principio de mi vida.

A mis padres, por el amor y apoyo incondicional que me han brindado en todo momento, me han dado todo lo que soy como persona.

A mis hermanos, el mayor regalo que me dieron mis padres, siempre estaré agradecida por las largas horas de risa y buenos momentos.

A mi sobrino que por medio de su alegría me motivó a seguir adelante.

AGRADECIMIENTOS

Agradezco a mi familia por ser la fuerza que me impulsa para seguir adelante, por el amor incondicional, en especial a mis padres, por los sacrificios para darme la posibilidad de una buena educación. Mi madre quien, con sus palabras de aliento, me dio la fuerza para no desfallecer, mi padre que siempre ha tratado de darme el mejor ejemplo, por su trabajo no siempre pudo estar presente en muchos momentos de mi vida, pero nunca me dejó sola, Y a pesar de que no fue económicamente fácil para él, me apoyó hasta el final.

Agradezco a mis amigos, en especial Estefany y Manuela, por acompañarme en todos estos años de formación y llenarlos de buenos momentos, llenos de alegría, gracias por estar siempre que las necesite.

Adrián por apoyarme en estos años, por tu agradable compañía en las noches de traspasado mientras escribía este libro, por cada charla y café que tomamos juntos.

A mi director de tesis el Ingeniero David Ramírez, por compartir su conocimiento conmigo y guiarme con tanta paciencia y dedicación a lo largo del desarrollo de mi trabajo de grado, gracias por las horas dedicadas y la gran colaboración que me brindaste.

También quiero agradecer a todos los docentes que se han visto involucrados en mi formación, por orientarme y motivarme a lo largo de la carrera con su experiencia y profesionalismo.

CONTENIDO

CONTENIDO

1.	Introducción	16
1.1	Planteamiento del problema	16
1.2	Justificación	18
1.3	Objetivos.....	19
1.3.1	Objetivo general	19
1.3.2	Objetivos específicos	19
1.4	Alcance.....	19
2.	Marco referencial	20
2.1	Voz humana.	20
2.1.1	Producción y su fisiología	20
2.1.2	Patologías vocales	21
2.2	Métodos de evaluación vocal.....	23
2.2.1	Métodos invasivos.....	23
2.2.2	Métodos no invasivos.....	24
2.2.3	Método de evaluación perceptual CAPE-V.....	25
2.2.4	Análisis de la señal acústica de voz	26
2.2.5	Principales parámetros del análisis acústico	27
2.3	Machine Learning	32
2.3.1	Tipos de aprendizaje automático.....	32
3.	Estado del arte	35
3.1	Extracción de características en audio con redes neuronales convolucionales.....	35
3.2	Reconocimiento de normas en medidas acústicas para identificar patologías en laringe	36
3.3	Clasificación de trastornos vocales utilizando redes neuronales artificiales	37
3.4	Red neuronal profunda para análisis acústico.	38
3.5	Identificación de parámetros subjetivos de voz utilizando redes neuronales artificiales.	39
3.6	Análisis perceptual de la disfonía: adaptación y validación del método CAPE-V al español.....	40
4	Construcción del dataset.....	41

4.1	Construcción de un banco de voces	41
4.1.1	Base de datos	41
4.1.2	Creación de audios sintéticos.....	42
4.1.3	Edición de los audios	44
4.2	Algoritmo de extracción automático de características	45
5.	Selección de atributos	50
5.1	Software Weka	50
5.1.1	Selección de atributos mediante Weka.....	52
5.1.2	Prueba con la herramienta de selección de atributos	55
6.	Modelo de regresión.....	66
6.1	Modelo de regresión con herramienta Weka.....	66
6.1.1	Modelos de regresión que proporciona Weka	67
6.1.2	Prueba con modelos de regresión y modos de prueba de Weka.....	68
6.2	Modelo de regresión mediante MATLAB	97
6.2.1	Prueba con modelos de regresión de MATLAB.....	99
7.	Resultados	117
8.	Conclusiones.....	122
8.1	Trabajos futuros.....	122
	Bibliografía.....	124

LISTA DE FIGURAS

Figura 1. Esquema aparato fonador..	21
Figura 2. Corte esquemático del aparato fonatorio humano.....	21
Figura 3. pólipo vocal ejemplo trastorno vocal.	22
Figura 4. Equipo necesario para realizar examen de videolaringoestroboscopia. .	24
Figura 5. Representación de las medidas de perturbación en la señal de voz.....	30
Figura 8. clasificación vs regresión fuente	34
Figura 9. base de datos creada para almacenar.	42
Figura 10. Pantalla principal de la interfase gráfica del usuario (IGU).....	43
Figura 11. Pantalla principal de la interfase gráfica del usuario Audacity.	45
Figura 12. Dataset obtenido en MATLAB a partir de los 500 audios creados.	49
Figura 13. Interfaz gráfica de usuario, Weka.....	51
Figura 14. panel de pre procesado, Weka.	51
Figura 15. panel de selección de atributos, Weka.....	52
Figura 16. modelo de atributo evaluador de correlación	56
Figura 17. modelo de atributo evaluador cfssubseteval	58
Figura 18. prueba con evaluador cfssubseteval.....	59
Figura 19. prueba con evaluador Reliefattributeeval	60
Figura 20. prueba con evaluador de atributo de correlación	61
Figura 21. evaluador Reliefattributeeval y método de búsqueda ranker.....	63
Figura 22. bloque de clasificación, interfaz gráfica Weka.....	66
Figura 23. modelo de regresión lineal regression-modo training set.....	69
Figura 24. modelo de regresión lineal regression-modo cross validation.....	70
Figura 25. modelo de regresión lineal regression-modo percentage Split.....	71
Figura 26. modelo de regresión simple linear regression-modo use training set ...	71
Figura 27. modelo de regresión simple linear regression-modo cross validation...	72
Figura 28. modelo de regresión simple linear regression-modo percentage split ..	73
Figura 29. modelo de regresión Additive regression - modo training set.....	74
Figura 30. modelo de regresión Additive regression - modo cross validation.....	74
Figura 31. modelo de regresión Additive regression - modo percentage split	75

Figura 32. modelo de regresión regressionbydiscretization - modo training set	76
Figura 33. modelo de regresión regressionbydiscretization-cross-validation.....	77
Figura 34. modelo de regresión regressionbydiscretization-percentage split	77
Figura 35. modelo de regresión Gaussian processes - modo training set	78
Figura 36. modelo de regresión Gaussian processes - modo cross-validation.....	79
Figura 37. modelo de regresión Gaussian processes - percentage split	80
Figura 38. modelo de regresión M5P - modo training set.....	80
Figura 39. modelo de regresión M5P - modo cross validation.....	81
Figura 40. modelo de regresión M5P - modo percentage split.....	82
Figura 41. modelo linear regression - modo training set	83
Figura 42. modelo linear regression - modo cross-validation	83
Figura 43. modelo linear regression - modo percentage split.....	84
Figura 44. modelo simple linear regression- modo training set	85
Figura 45. modelo simple linear regression- modo cross validation	85
Figura 46. modelo simple linear regression- modo percentage split.....	86
Figura 47. modelo Additive regression- modo training set	87
Figura 48. modelo Additive regression- modo cross validation	87
Figura 49. modelo Additive regression- modo percentage split.....	88
Figura 50. resultados obtenidos modelo regressionbydiscretization	89
Figura 51. resultados obtenidos modelo Regressionbydiscretization	90
Figura 52. resultados modelo Regressionbydiscretization-percentage split	91
Figura 53. resultados obtenidos modelo Gaussian processes- modo training set .	91
Figura 54. resultados obtenidos modelo gaussian processes- cross-validation	92
Figura 55. resultados obtenidos modelo gaussian processes - percentage split ...	93
Figura 56. resultados obtenidos modelo M5P- modo training set.....	93
Figura 57. resultados obtenidos modelo M5P- modo cross-validation	94
Figura 58. resultados obtenidos modelo M5P- modo percentage split	95
Figura 59. Modelos de regresión que proporciona Matlab	99
Figura 60. Gráfico de resultado modelo de regresión lineal en Matlab.....	100
Figura 61. Gráfico de resultado modelo de interacciones lineal en Matlab	100

Figura 62. Gráfico de resultado modelo de robust linear en Matlab	101
Figura 63. Gráfico de resultado modelo fine tree en Matlab	101
Figura 64. Gráfico de resultado o modelo de medium tree en Matlab	102
Figura 65. Gráfico de resultado obtenido con el modelo coarse tree en Matlab ..	102
Figura 66. Gráfico de resultado obtenido con el modelo de linear svm en Matlab	103
Figura 67. Gráfico de resultado modelo de quadratic svm en Matlab.....	103
Figura 68. Grafico de resultado modelo de cubic svm en Matlab	104
Figura 69. Gráfico de resultado modelo de fine gaussian en Matlab	104
Figura 70. Gráfico de resultado modelo de Medium gaussian	105
Figura 71. Gráfico de resultado modelo coarse gaussian en Matlab.....	106
Figura 72. Gráfico de resultado modelo rational quadratic en Matlab	106
Figura 73. Gráfico de resultado modelo de squared exponential	107
Figura 74. Gráfico de resultado modelo de Matern 5/2 en Matlab.....	107
Figura 75. Gráfico de resultado modelo exponential en Matlab	108
Figura 76. Gráfico de resultado modelo de Boosted tres en Matlab.....	109
Figura 77. Gráfico de resultado modelo de bagged trees en Matlab	109
Figura 78. respuesta pronosticada frente a la respuesta verdadera.....	111
Figura 79. gráfica modelo gaussiano optimizado con kernel de 3/2	111
Figura 80. gráfica Modelo gaussiano Matern	112
Figura 81. Opción PCA para reducción de dimensionalidad	113
Figura 82. gráfica entrenamiento con herramienta PCA	113
Figura 83. Reducción de dimensionalidad	115
Figura 84. reducción de dimensionalidad.....	116
Figura 85. interfaz gráfica de inicio Apps Designer Matlab.....	118
Figura 86. interfaz gráfica de usuario de la App.....	118
Figura 87. prueba de funcionamiento de la App.....	119

LISTA DE TABLAS

Tabla 1. Extracción de características en audio	35
Tabla 2. Reconocimiento de normas en medidas acústicas (Santos, 2017).	36
Tabla 3. clasificación de trastornos vocales fuente: (fermino, 2017).	37
Tabla 4. Red neuronal profunda para análisis acústico.....	38
Tabla 5. Identificación de parámetros subjetivos utilizando redes neuronales	39
Tabla 6. Adaptación y validación del método CAPE-V al español (Galan, 2014). .	40
Tabla 7. Atributos seleccionados modo de correlación y modo ranker.....	56
Tabla 8. atributos seleccionados evaluador Cfssubsetval	57
Tabla 9. atributos seleccionados evaluador cfssubetval	58
Tabla 10. atributos seleccionados evaluador Reliefatributeeval	60
Tabla 11. atributos seleccionados evaluador de correlación	61
Tabla 12. Atributos seleccionados evaluador Relieffatributeeval ranker.....	62
Tabla 13. selección final de parámetros	64
Tabla 14. valores obtenidos con el modo de regresión, linear regression	69
Tabla 15. valores obtenidos con el modo de regresión, linear regression	69
Tabla 16. valores obtenidos con el modo de regresión, linear regression	70
Tabla 17. valores obtenidos con el modo de regresión, simple linear regression..	71
Tabla 18. valores obtenidos con el modo de regresión	72
Tabla 19. valores obtenidos con el modo de regresión, simple linear regression..	72
Tabla 20. valores obtenidos con el modo de regresión, additive regression	73
Tabla 21. valores obtenidos con el modo de regresión, additive regression	74
Tabla 22. valores obtenidos con el modo de regresión, additive regression	74
Tabla 23. valores obtenidos con el modo regressionbydiscretization	75
Tabla 24. valores obtenidos con el modo regressionbydiscretization	76
Tabla 25. valores obtenidos con el modo regressionbydiscretization	77
Tabla 26. valores obtenidos con el modo de regresión, Gaussian processes	78
Tabla 27. valores obtenidos con el modo de regresión, Gaussian processes	78
Tabla 28. valores obtenidos con el modo de regresión, Gaussian processes	79
Tabla 29. valores obtenidos con el modo de regresión, M5P - modo training set..	80

Tabla 30. valores obtenidos con el modo de regresión, M5P	81
Tabla 31. valores obtenidos con el modo de regresión M5P	81
Tabla 32. valores modo de regresión, clasificador linear regression	82
Tabla 33. valores obtenidos con el modo de regresión, linear regression	83
Tabla 34. valores obtenidos con el modo de regresión, linear regression	83
Tabla 35. valores obtenidos con el modo de regresión	84
Tabla 36. valores obtenidos con el modo de regresión, clasificador simple linear.	85
Tabla 37. valores obtenidos con el modo de regresión, clasificador simple linear.	86
Tabla 38. valores obtenidos con el modo de regresión, Additive regression	86
Tabla 39. valores obtenidos con el modo de regresión, Additive regression	87
Tabla 40. valores obtenidos con el modo de regresión, Additive regression	88
Tabla 41. valores obtenidos con el modo regressionbydiscretization	88
Tabla 42. valores obtenidos con el modo Regressionbydiscretization	89
Tabla 43. valores obtenidos con el modo Regressionbydiscretization	90
Tabla 44. valores obtenidos con el modo de regresión, Gaussian processes	91
Tabla 45. valores obtenidos con el modo de regresión, gaussian processes	92
Tabla 46. valores obtenidos con el modo de regresión, gaussian processes	92
Tabla 47. valores obtenidos con el modo de regresión, M5P	93
Tabla 48. valores obtenidos con el modo de regresión, M5P	94
Tabla 49. valores obtenidos con el modo de regresión, M5P	94
Tabla 50. Valores obtenidos modelos de regresión	95
Tabla 51. resultados de los diferentes modelos de regresión de Matlab.	110
Tabla 52. valores de error entrenamiento	113

TABLA DE ECUACIONES

(1) Frecuencia fundamental	28
(2) Jitter	29
(3) Jitter relativo	29
(4) Shimmer	30

RESUMEN

La voz es el elemento fundamental de la comunicación humana, a través de ella las personas transmiten sus sentimientos, emociones e ideas, en este contexto, diferentes ramas de la medicina y la ingeniería se han abierto al estudio de este fenómeno para explicar de un modo u otro cómo se origina y las posibles alteraciones (nódulos, quistes, pólipos, entre otras lesiones patológicas) que se interponen a su producción. Para evaluar la voz y analizar las diferentes patologías vocales, se realiza una evaluación en la que se tiene en cuenta todos los aspectos de la voz, en todo el proceso de producción, este examen es realizado por el fonoaudiólogo y se hace a través de exámenes y pruebas que en su mayoría son perceptuales. Se le pide al paciente realizar diferentes tareas vocales para así poder evaluar todos los aspectos de la voz cuando se habla o se canta, el diagnóstico o resultado se da por el fonoaudiólogo en base a su experiencia y conocimientos previos. Existen diferentes escalas que buscan estandarizar la valoración perceptual de la voz, como lo son GRBAS y CAPE-V, pero estas siguen teniendo alto nivel de subjetividad. Diversos estudios se han desarrollado para reducir la subjetividad por parte del profesional en esta prueba, pero la mayoría son desarrollados en torno a la escala GRABS, la cual presenta ciertas desventajas.

Con el propósito de reducir el tiempo de examen y la subjetividad de la prueba se plantea el desarrollo, de una app que permite estimar el parámetro soplosidad dentro del método de evaluación perceptual CAPE-V, mediante procesamiento digital de señales acústicas y aprendizaje de máquina.

Para llegar al desarrollo final de la App, se creó un banco de voces sintético, el cual se utilizó para crear un dataset, que se implementó en modelos de selección de atributos y de regresión, en los softwares Matlab y Weka, obteniendo con esto el modelo de regresión más óptimo para la predicción de la soplosidad. En el transcurso del texto se irá proporcionando información detallada sobre el proceso de diseño, implementación y puesta en marcha de la app, observando cómo este aporte puede llegar a ser favorable para el mundo de la salud y la ingeniería.

Palabras claves: Evaluación perceptual de la voz, dataset, regresión, soplosidad

ABSTRACT

The voice is fundamental element of human communication, through it people transmit their feelings, emotions and ideas. In this context, different branches of medicine and engineering have opened up to the study of this phenomenon to explain in a way or other as it originates and the possible alterations, to detect these disorders of the voice the functional evaluation is used, To carry out this evaluation, the speech pathologist must apply different tests and procedures in order to assess each of the elements that intervene in the production of the spoken and sung voice. These elements are evaluated largely through auditory perceptual evaluation, palpation and visual observation, which makes this examination subjective, and with this that the result depends on the experience of the professional I evaluated. There are different scales that seek to standardize the perceptual assessment of the voice, such as GRABS and CAPE-V, but these still have a high level of subjectivity. Several studies have been developed to reduce professional subjectivity in this test, but most are developed around the GRABS scale, which has certain disadvantages.

With the purpose of reducing the exam time and the subjectivity of the test, the development of an app that allows predicting the breathiness parameter within the CAPE-V perceptual evaluation method is proposed, through digital processing of acoustic signals and machine learning.

To arrive at the final development of the App, a synthetic voice bank was created, which was used to create a dataset, which was implemented in attribute selection and regression models, in the Matlab and Weka softwares, obtaining with this the most optimal regression model for the prediction of soplocity.

Keywords: Perceptual voice note, data set, regression, soplocity.

1. INTRODUCCIÓN

La voz es un instrumento fundamental para la comunicación de los seres humanos especialmente los profesionales de la voz que la utilizan como instrumento de trabajo como: cantantes, profesores, locutores.

Un gran porcentaje de personas de países avanzados desarrollan alguna patología vocal que afecta su voz (mateos, 2017). El porcentaje de patologías de la voz en los docentes durante su vida es de 57.7% en comparación con los no docentes, donde el porcentaje de por vida es del 28.8%. (Trinite, 2017).

Para evaluar la voz y analizar las diferentes patologías vocales, se realiza una evaluación en la que se tiene en cuenta todos los aspectos de la voz, en todo el proceso de producción, este examen es realizado por el fonoaudiólogo y se hace a través de exámenes y pruebas que en su mayoría son perceptuales. Se le pide al paciente realizar diferentes tareas vocales para así poder evaluar todos los aspectos de la voz cuando se habla o se canta, el diagnóstico o resultado se da por el fonoaudiólogo en base a su experiencia y conocimientos previos (Guzmán, 2016).

1.1 PLANTEAMIENTO DEL PROBLEMA

El nivel de percepción depende del nivel de experticia del fonoaudiólogo y puede cambiar de un fonoaudiólogo a otro, ya que el resultado que nos arroje un profesional de la voz puede que no sea el mismo que nos proporcione otro profesional, ya que los resultados obtenidos son subjetivos y dependen de cada uno (Fang-Ling Lu, 2014). Teniendo en cuenta lo anterior, diversos estudios se han adelantado para estandarizar una escala que permita medir el nivel de la calidad vocal como lo son GRBAS (grado, roughness, asthenicity, breathinnes y strain), o CAPE-V (Consensus Auditory-Perceptual Evaluation of Voice).

GRBAS es una escala de evaluación subjetiva o perceptual de la voz, se creó con el fin de estandarizar la calificación dada por los fonoaudiólogos en la evaluación perceptual. Esta escala evalúa 5 aspectos de la voz y que son el acrónimo de la palabra GRBAS. Aunque se estandarizó esta escala sigue existiendo alto nivel de subjetividad en la calificación por parte de fonoaudiólogos (Kreiman, 2010). Algunos estudios se han desarrollado buscando disminuir la subjetividad por medio de una estimación automática de parámetros de la voz basándose en la escala GRBAS, esto mediante características que son extraídas en el estudio del análisis acústico de señales de voz y redes neuronales convolucionales (García Mario Alejandro R. A.), No obstante, estos trabajos presentan inconvenientes ya que su desarrollo es en torno a la escala GRBAS, la cual solo evalúa la calidad en tres niveles, de manera que existe una gran diferencia entre un nivel y otro. Una mala clasificación puede afectar drásticamente el diagnóstico del paciente y así no permitir hacer un buen seguimiento del progreso de este, por esto es difícil evidenciar la evolución de un tratamiento en el paciente, además de esto la evaluación se hace de forma manual por el fonoaudiólogo. Teniendo en cuenta las problemáticas planteadas en la escala Surge CAPE-V, el cual propone una escala de calificación lineal de 0 a 100, permitiendo puntos intermedios en la puntuación y también agrega otros aspectos a evaluar en la voz (severidad global, aspereza, soplosidad, tensión, tono e intensidad). A pesar de que esta escala lineal permite un mejor seguimiento de la calidad vocal del paciente, actualmente los valores de calificación de la prueba son dados de manera subjetiva por el fonoaudiólogo y por lo tanto se sigue teniendo alta subjetividad en el diagnóstico (Zraick, 2011).

Por otro lado, teniendo en cuenta (Zraick, 2011), en el cual se realiza un estado de arte sobre las técnicas de evaluación de la voz y se enfatiza que el CAPE-V es la prueba más adecuada para medir calidad de la voz.

1.2 JUSTIFICACIÓN

Un gran porcentaje de personas de países avanzados desarrollan alguna patología vocal que afecta su voz (mateos, 2017). Estos trastornos de la voz se observan con mayor frecuencia a la edad de los 25 a 45 años. La disfonía en niños puede ser un indicativo de que se puede desarrollar en un futuro un trastorno vocal crónico. Se debe tener en cuenta la disfonía y prebifonia ya que estas nos pueden ayudar a detectar enfermedades a tiempo, como trastornos neurológicos, degenerativos Y oncológicos, esto en pacientes de alto cuidado (mateos, 2017).

El método CAPE-V es una opción positiva para llevar a cabo el análisis perceptual de la voz, y al ser no invasivo es mucho más cómodo para el paciente. Teniendo en cuenta que el método CAPE-V es un proceso tedioso para el profesional, ya que debe realizar todo el proceso evaluativo de forma manual y esto conlleva un cierto tiempo (Galan, 2014). En este proyecto se busca desarrollar un software capaz de estimar el parámetro soplosidad, y a futuro cada uno de los parámetros cape-v de la calidad vocal, esto mediante características que son extraídas en el estudio del análisis acústico de señales de voz y métodos de machine learning, lo cual elimina la subjetividad de la prueba aportando valores objetivos que se ven reflejados en diagnósticos mucho más precisos y permite realizar un seguimiento del paciente en su tratamiento y cuantificar de manera más exacta su evolución.

El proyecto sigue la línea de objetivos de desarrollo sostenible número 3(salud y bienestar) y 4(educación de calidad), de la OMS (organización mundial de la salud), ya que el software puede ser utilizado para educar estudiantes de fonoaudiología, en su oído clínico, propuestos por el programa de las naciones unidas para el desarrollo.

1.3 OBJETIVOS

1.3.1 OBJETIVO GENERAL

Estimar el parámetro SOPLOSIDAD dentro del método de evaluación perceptual CAPE-V, mediante procesamiento digital de señales acústicas y aprendizaje de máquina, para reducir el tiempo de examen y subjetividad por parte del profesional.

1.3.2 OBJETIVOS ESPECÍFICOS

- Determinar parámetros acústicos de la voz relevantes para estimar soplosidad.
- Construir un banco de voces sintético con su respectiva evaluación de soplosidad.
- Implementar un algoritmo para extracción automática de los parámetros acústicos definidos.
- Implementar y validar algoritmo para estimación del parámetro soplosidad en señales de voz.

1.4 ALCANCE

En este proyecto se desarrolló un algoritmo para estimar el parámetro soplosidad dentro del método de evaluación perceptual CAPE-V. Se creó un banco de voces con 500 audios sintéticos mediante el programa EVAPER. El algoritmo de extracción de características automático, se desarrolló mediante el software MATLAB 2019b. El dataset obtenido tiene 3000 muestras con sus características establecidas.

Se hizo uso del programa Weka versión 3.9.4, para análisis de datos. El algoritmo de regresión se implementó con el software MATLAB, haciendo uso de toolbox regression learner. El proyecto se validó con el banco de voces sintéticas creado

2. MARCO REFERENCIAL

2.1 VOZ HUMANA.

La voz se crea por la vibración de los pliegues vocales, esto sucede por el paso del aire proveniente del aparato respiratorio. Es un fenómeno fisiológico con un resultado de naturaleza acústica. Es un vehículo de comunicación entre las personas y uno de los medios esenciales para expresar y comunicar los conocimientos, los pensamientos y los sentimientos propios (Jaume de Montserrat).

2.1.1 PRODUCCIÓN Y SU FISIOLÓGÍA

La producción de la voz comienza en los pulmones donde se almacena el aire, este es dirigido hacia los pliegues vocales haciéndolos vibrar y produciendo así el sonido, este sonido adquiere sus cualidades de timbre propias de cada voz, en el aparato resonador donde es articulado (Preparadores de Oposiciones para la Enseñanza, s.f.), en la Figura 1 y 2, se esquematiza las fases para la producción de voz.

El sistema vocal humano podemos dividirlo en tres partes:

- Aparato respiratorio, que es donde se almacena el aire requerido para hacer vibrar las cuerdas vocales produciendo el sonido. Intervienen diferentes sistemas y partes del organismo como los pulmones, la tráquea y la nariz.
- Aparato fonatorio, Se encarga de producir el sonido transformando el aire que pasa a través de las cuerdas vocales haciéndolas vibrar, intervienen diferentes partes del organismo dentro de estos la laringe y pliegues vocales.
- Aparato resonador, Se encarga de aportarle al sonido obtenido mediante la vibración de las cuerdas vocales, las características que diferencian una voz de otra, en cuanto a timbre. Este sonido es articulado mediante las cavidades bucales, la faringe, el paladar óseo, y los labios.

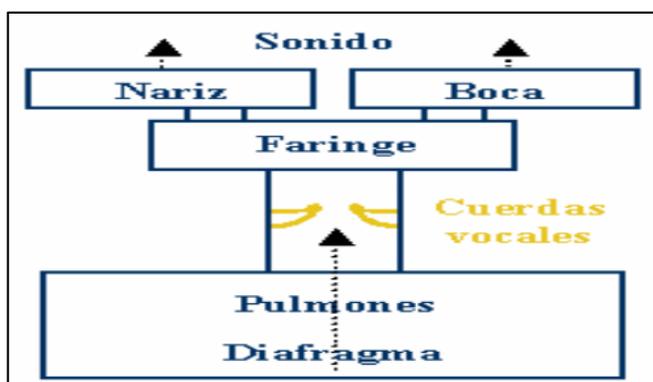


Figura 1. Esquema instrumento vocal. Fuente: (Preparadores de Oposiciones para la Enseñanza, s.f.).

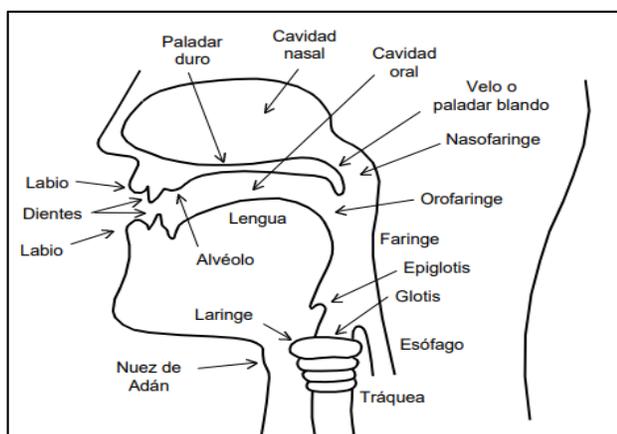


Figura 2. Corte esquemático del aparato fonatorio humano. Fuente: (Brockmann, 2011).

2.1.2 PATOLOGÍAS VOCALES

Un gran porcentaje de personas de países avanzados desarrollan alguna patología vocal que afecta su voz (mateos, 2017). Estos trastornos de la voz se observan con mayor frecuencia a la edad de los 25 a 45 años, sin embargo, pueden ocurrir en cualquier grupo de edad. Se dice que existe un trastorno de la voz cuando sus características tímbricas y de intensidad, son diferentes a comparación de las características de las voces de otras personas relacionadas con su edad, género y entorno cultural, una voz con alteraciones puede ser indicativo de la existencia de alguna enfermedad o un trastorno de la comunicación. Los trastornos de la voz

pueden clasificarse de diferente manera, una de estas es clasificarlos por sus mecanismos etiopatogénicos que intervienen en la fonación (Ignacio Cobeta, 2013):

- Trastornos orgánicos: esta se da cuando se ve afectado la anatomía o estructuras que intervienen en la producción de la voz, en la Figura 3. Se puede observar los trastornos de la voz, resultado de cambios en la estructura como los nódulos o pólipos del pliegue vocal.
- Trastornos funcionales: no se evidencia ningún cambio en cuanto la anatomía en las estructuras que producen la voz.
- Trastornos orgánico-funcionales: Si se presenta un mal uso de la voz, esta puede conllevar a producir lesiones orgánicas, que no son y no causaran patologías vocales

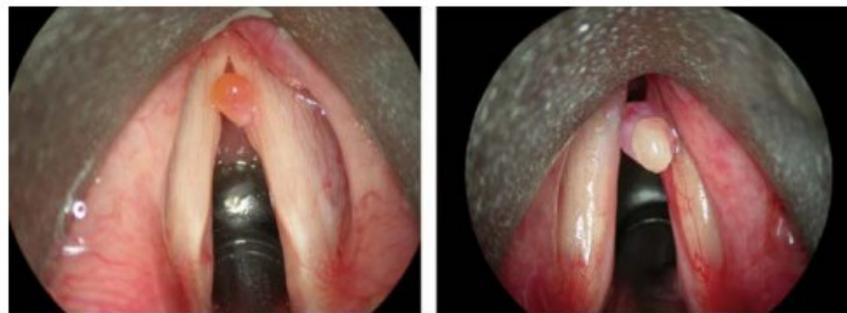


Figura 3. Pólipo vocal ejemplo trastorno vocal. Fuente: (Ignacio Cobeta, 2013).

Los trastornos de la voz pueden tener efectos enormes en la calidad de vida personal y profesional de las personas, por tanto, el diagnóstico y tratamiento eficaces son fundamentales para evitar consecuencias negativas para el paciente y minimizar los costos de atención en salud (Brockmann, 2011).

2.2 MÉTODOS DE EVALUACIÓN VOCAL

Los métodos de evaluación vocal pueden clasificarse en dos métodos, invasivos y no invasivos.

2.2.1 MÉTODOS INVASIVOS

En términos de evaluación vocal, los métodos invasivos se basan en la obtención de imágenes internas de la laringe del paciente, lo que permite la evaluación de las condiciones de las cuerdas vocales. Al observar el uso de la videolaringoestroboscopia, por ejemplo, se percibe la mayor ventaja de los métodos invasivos: la capacidad de visualizar, a través de una cámara, el comportamiento real de las cuerdas vocales y las posibles patologías que pueden estar afectando la producción vocal del individuo. El movimiento vibratorio de las cuerdas vocales es muy rápido para la percepción del ojo humano, por lo que es necesario insertar una micro cámara capaz de capturar a alta velocidad junto con un aparato de luz estroboscópica. Los pulsos de luz se ajustan a una frecuencia similar a la frecuencia fundamental de la voz analizada, lo que permite observar las características relevantes en la evaluación vocal. Sin embargo, a pesar de sus ventajas, es un examen de alto costo donde es necesario insertar el equipo en la orofaringe del paciente, causando molestias e incluso dolor (Fermino, 2017). Se puede ver en la Figura 4 El equipo necesario para realizar la videolaringoestroboscopia.



Figura 4. Equipo necesario para realizar examen de videolaringoestroboscopia
Fuente: (Fermino, 2017).

2.2.2 MÉTODOS NO INVASIVOS

Los métodos no invasivos se pueden clasificar en dos categorías básicas: los métodos de evaluación perceptual auditivo y la evaluación acústica. Es posible definirlos de la siguiente manera (Fermino, 2017).

- **Evaluación audioperceptiva:**

Esto implica describir la voz únicamente escuchándola, es decir, utilizando la percepción auditiva. Los terapeutas del habla que se especializan en trabajar con personas con trastornos de la voz han recibido capacitación para describir las características relevantes de una voz desordenada. Las escalas de evaluación perceptual más utilizados son la escala GRBAS y CAPE-V (Williamson, 2014). Se pueden presentar diferencia de inter juicios entre fonoaudiólogos, ya que los profesionales basan su diagnóstico en diferentes aspectos como los es su experiencia y el entorno en el que se desarrolla su trabajo, estos juzgan basándose en escalas calificativas diferentes. (Galan, 2014).

- **Evaluación acústica:**

Permite cuantificar, medir y describir las características de la voz. Obtener datos extraídos mediante un procesamiento computacional apropiado. También permite una integración entre el análisis auditivo-perceptual con el plan fisiológico, componiendo una visión más detallada sobre todo el proceso de generación de la señal audible, sus modificaciones y disfonía. Para ello, utiliza parámetros acústicos

que forman la señal (periodicidad, amplitud, duración, composición espectral) que caracterizan los atributos físicos de la voz en el dominio del tiempo, la frecuencia y también su intensidad (Fermino, 2017).

2.2.3 MÉTODO DE EVALUACIÓN PERCEPTUAL CAPE-V

CAPE-V es un método de evaluación perceptual de la voz, llamado así por sus siglas (Consensus Auditory-Perceptual Evaluation of Voice) es una escala que se obtuvo mediante la participación de varios fonoaudiólogos expertos en este tema reunidos en un consenso donde se creó un protocolo clínico que se utiliza para realizar evaluaciones subjetivas de la voz de forma más estandarizada, aunque no se ha creado un método totalmente seguro para obtener la calificación evaluativa de la voz, este método resulta ser el más actualizado en cuanto conocimiento respecto a la evaluación perceptual realizada por los fonoaudiólogos en clínica vocal y todos los factores que se deben tener en cuenta para realizar este examen.

En este método se evalúan importantes parámetros de la calidad vocal de los pacientes. Dichos atributos son: a) severidad global; b) aspereza; c) soplosidad; d) tensión; e) tono e f) intensidad. Y en este se pondera cada atributo evaluado en una escala lineal de 0 a 100mm. El fonoaudiólogo debe establecer desde su juicio perceptual la desviación que presenta la voz evaluada con respecto a la escala normal y posteriormente marcar sobre la escala calificativa en la línea. Se asigna la calificación por la observación de las características de la voz de forma subjetiva por parte del profesional. Los atributos nombrados son los mínimos que se deben evaluar en este método (Galan, 2014).

▪ **Parámetro Soplosidad**

El comportamiento fisiopatológico que subyace en este parámetro es el cierre glótico incompleto o ineficaz. Común en las enfermedades que impidan el correcto cierre de las cuerdas en la fonación como en las lesiones tipo masa, Figura 3. Las cuales se presentan como una masa fibrosa en los pliegues vocales o en las parálisis laríngeas (neurológica) (Polo, 2015). La causa más frecuente de un cierre glótico incompleto es la existencia de una masa en la superficie de una o ambas cuerdas vocales que impide su completo contacto en la línea media. Las lesiones más comunes son los nódulos vocales, los pólipos y los carcinomas (Batalla, 2015).

2.2.4 ANÁLISIS DE LA SEÑAL ACÚSTICA DE VOZ

En el análisis acústico de la voz se extraen parámetros que son estudiados mediante procesamiento de señales, estos no sirven para evaluar el estado de una voz, pero se debe tener conocimiento de todos los parámetros que se obtienen para así poder interpretar bien los resultados que se obtengan y dar un diagnóstico preciso. En general, el análisis de la voz acústica instrumental se refiere a una familia de técnicas computarizadas que miden las propiedades de la señal, cuando el paciente realiza una tarea vocal o pronuncia una frase (Ignacio Cobeta, 2013).

▪ **Ventajas de la evaluación acústica**

- ✓ Identificar parámetros de la voz que puedan causar disfonía, para obtener un diagnóstico más preciso.
- ✓ Establecer el grado de disfonía presente.
- ✓ Establecer el avance en la disminución de la disfonía frente a un tratamiento aplicado y determinar si se debe continuar con este o por otro lado cambiarlo.

- ✓ Los equipos utilizados para el análisis de la señal son fáciles de usar. Permitiendo que puedan ser manejados por un amplio grupo de personas.
- ✓ Facilita enormemente el cálculo de los parámetros los cuales anteriormente, se obtenían de forma perceptual y manual.
- ✓ Esta es una técnica que, al no ser invasiva, no es incómoda ni genera dolor para el paciente en comparación con la evaluación invasiva como la laringoestroboscopia.

- **Inconvenientes del análisis acústico**

- ✓ Si no se sabe interpretar los resultados no se obtendrá su mayor potencial ya que se hará mal uso de la herramienta.
- ✓ No se tienen estándares de normalidad, ya que al obtenerse tantos valores se debe determinar si estos se asocian algún valor que indique patología.

2.2.5 PRINCIPALES PARÁMETROS DEL ANÁLISIS ACÚSTICO

De la señal de voz es posible extraer un conjunto de parámetros de la voz. Los parámetros obtenidos por el análisis acústico tienen la ventaja de describir la voz objetivamente. Con la existencia de bases de datos normativas que caracterizan la calidad de la voz o el uso de herramientas inteligentes que combinan los diversos parámetros, es posible distinguir entre la voz normal y patológica o incluso identificar o sugerir la patología. Estas herramientas permiten el monitoreo del punto de vista clínico y reducen el grado de subjetividad del análisis perceptivo (Carla Oliveira, 2013)

A continuación, se describen los parámetros más importantes utilizados para el desarrollo de este proyecto.

- **Frecuencia fundamental (F_0)**

La frecuencia fundamental, también llamada Pitch, está representada por F_0 , está relacionada con la masa, la tensión (elasticidad) y la longitud de las cuerdas vocales, además de la presión subglótica y la configuración del tracto vocal (Santos, 2017). Es el número de vibraciones de las cuerdas vocales por segundo. La frecuencia fundamental varía dependiendo del sexo y edad de cada persona. El tiempo de duración en cada ciclo se le conoce como periodo (T) sus unidades son segundos. Podemos conocer el valor del periodo teniendo el valor de la frecuencia y de igual manera de forma inversa (Ignacio Cobeta, 2013):

$$T = 1/F_0 \text{ y } F_0 = 1/T \quad (1)$$

- **Determinación de Jitter**

Variaciones de forma involuntaria de la F_0 (frecuencia fundamental) que se observan en la señal de la voz, estas representan una medida de estabilidad en la producción de la voz (Cobeta, Núñez, & Fernández, 2013). Fisiológicamente las causas de estas perturbaciones son diversas y se explican a continuación (Ford & Bless, 1986):

- ✓ De tipo Neurológico: se debe al control ineficaz del sistema nervioso ocasionando el no control de los músculos de la boca.
- ✓ De tipo aerodinámico: cuando se tiene un cierre glótico incompleto o ineficaz permitiendo que escape el aire.
- ✓ De tipo mecánica: cuando no se tiene una simetría en los pliegues vocales o algunos cambios en las estructuras que intervienen en la fonación.

$$jitta = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}| \quad (2)$$

La ecuación (2). Representa la diferencia absoluta promedio entre dos períodos consecutivos y se conoce como Jitta. El valor umbral para detectar patologías en adultos es 83,2. (Carla Oliveira, 2013)

- **Determinación de Jitter relativo (Jitt)**

Mide la variación inter-ciclos en porcentaje representando la diferencia absoluta promedio entre dos períodos consecutivos, dividido por el período promedio Figura 5. Se conoce como jitt y se ve representada en la ecuación 3. (Carla Oliveira, 2013).

$$jitt = \frac{jitta}{\frac{1}{N} \sum_{i=1}^N T_i} \times 100 \quad (3)$$

- **Determinación de shimmer**

Este es un parámetro relacionado con la intensidad, esta es la amplitud de la variación de la presión del sonido que se produce al propagarse la voz a través de un medio y sus unidades son en decibeles (Teixeira, Oliveira, & Lopes, 2013)

Su valor depende principalmente de la amplitud en la que vibran los pliegues vocales y la presión subglótica, al aumentar estos valores, incrementa la intensidad de la voz (Brockmann, 2011). Este parámetro se puede ver disminuido por: un ejercicio respiratorio incorrecto, un ciclo glótico incompleto y unos pliegues vocales demasiado rígidos con poca flexibilidad (Cobeta, Núñez, & Fernández, 2013).

$$shimmer = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \times 100 \quad (4)$$

La ecuación (4) Representa la diferencia absoluta promedio entre las amplitudes de dos períodos consecutivos, dividido por la amplitud promedio (Carla Oliveira, 2013).

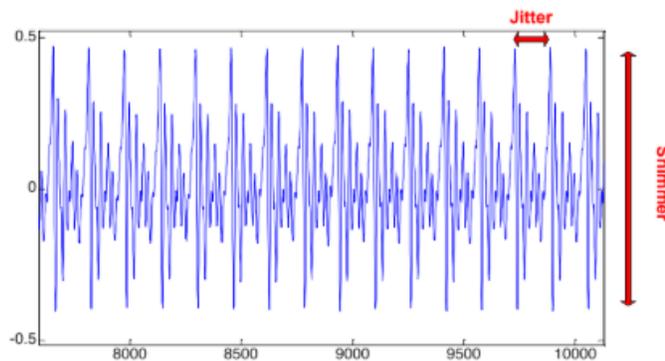


Figura 5. Representación de las medidas de perturbación Jitter y Shimmer en la señal de voz Fuente: (Carla Oliveira, 2013).

- **HNR (relación armónica - ruido)**

Establece la relación entre el armónico y el nivel de ruido que compone la señal, separa la señal en dos canales distintos y luego realiza una comparación entre la intensidad de los dos (Ignacio Cobeta, 2013).

- **SNR (Relación Señal a Ruido)**

Proporción existente entre el nivel de potencia de la señal y el nivel de potencia de ruido que la corrompe. Usualmente se expresa en decibeles (dB) (Nicolas Viancha, 2018).

- **THD (Distorsión total armónica)**

La distorsión armónica total (Total Harmonic Distortion -THD), de una señal es una medida de la distorsión armónica presente y se define como el cociente de la suma de las potencias de todos los componentes armónicos, o de un armónico en particular respecto a la potencia de la frecuencia fundamental (Nicolas Viancha, 2018).

- **Ancho de banda:**

Valor en Hz correspondiente a la diferencia de frecuencias que se obtiene al descender 3 decibeles (dB) del valor máximo del pico del espectro (Botero Tobón, 2008).

- **Cepstrum**

El cepstrum de la señal de voz se define como la transformada inversa de Fourier del logaritmo de su espectro localizado. El término cepstrum es indicativo de haber realizado una transformación inversa del espectro. La variable independiente del cepstrum se denomina cufrecuencia, término formado a partir de la palabra frecuencia, y tiene carácter temporal. La principal característica del cepstrum es que permite separar del espectro de la señal de voz la estructura fina y los formantes (Pericas, 1993).

- **Periodograma**

El método del periodograma es el método no paramétrico más común para calcular la densidad espectral de potencia (PSD) de una serie de tiempo. Este método calcula el PSD con la transformada rápida de Fourier (FFT) (Martínez, 2011)

2.3 MACHINE LEARNING

Con la cantidad cada vez mayor de datos en forma electrónica, la necesidad de métodos automatizados para el análisis de datos continúa creciendo. El objetivo del aprendizaje automático es desarrollar métodos que puedan detectar automáticamente patrones en los datos, y luego usar los patrones descubiertos para predecir datos futuros u otros resultados de interés. Por lo tanto, el aprendizaje automático está estrechamente relacionado con los campos de la estadística y la minería de datos, pero difiere ligeramente en términos de su énfasis y terminología (Murphy, 2012). El rápido crecimiento de la información, ha hecho imposible que un humano sea capaz de procesar y analizar tal cantidad de datos, por lo tanto, técnicas como el aprendizaje de máquina se han popularizado por la variedad de modelos con los que se puede analizar grandes volúmenes de datos de forma rápida y precisa. El resultado de usar aprendizaje de máquina son predicciones de alto valor que pueden guiar y tomar las mejores decisiones sin necesidad de la intervención humana (Gerardo Gutiérrez, 2015).

2.3.1 Tipos de aprendizaje automático

El aprendizaje automático generalmente se divide en dos tipos principales.

- **Aprendizaje no supervisado**

En el aprendizaje supervisado, los algoritmos trabajan con datos “etiquetados” (labeled data), intentando encontrar una función que, dadas las variables de entrada (input data), les asigne la etiqueta de salida adecuada. El algoritmo se entrena con un “histórico” de datos y así “aprende” a asignar la etiqueta de salida adecuada a un nuevo valor, es decir, predice el valor de salida (Simeone, 2018).

- **Aprendizaje supervisado**

En los métodos de aprendizaje supervisado se le proporciona al sistema los datos de entrada y los de salida, posteriormente se entrena el modelo buscando que el sistema aprenda o se entrene con estos datos y que a futuro pueda estimar un nuevo valor con datos nuevos que se le ingresen, una salida nueva del sistema (Serna, 2018). Estos métodos pueden dividirse en, clasificación y regresión.

- **Clasificación**

La clasificación hace parte del aprendizaje supervisado, donde se intenta estimar valores de salida del sistema basándose en el conocimiento adquirido durante la fase de entrenamiento del modelo, en este modelo se trabaja con valores discretos que pertenecen a grupos ya establecidos (Serna, 2018).

- **Regresión**

La regresión hace parte del aprendizaje supervisado, en la que se hace estimación de valores continuos, esto mediante el aprendizaje que se obtiene al entrenar el modelo con datos que se proporcionan para la entrada y salida del sistema, cuando el modelo encuentra relación entre estos datos puede estimar un valor de salida con datos nuevos que se le proporcionen, las variables que se utilizan en este proceso son salidas de valores continuos (Serna, 2018), esta diferencia entre clasificación y regresión se ilustra en la Figura 6.

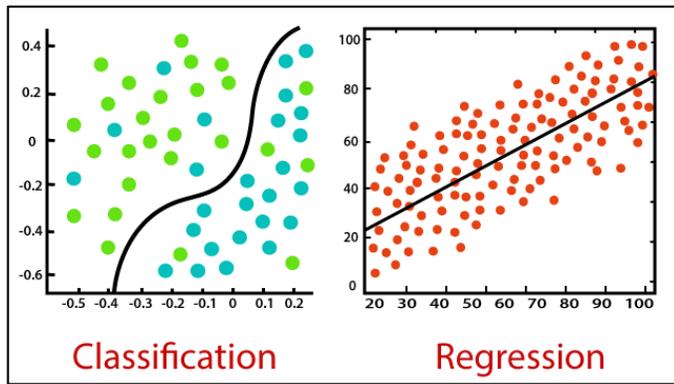


Figura 6. Clasificación vs regresión fuente: (Gerardo Gutiérrez, 2015).

3. ESTADO DEL ARTE

A continuación, se presenta una revisión bibliográfica de las investigaciones y trabajos realizados con el procesamiento digital de señales acústicas, las cuales son base primordial para el desarrollo de este trabajo.

3.1 EXTRACCIÓN DE CARACTERÍSTICAS EN AUDIO CON REDES NEURONALES CONVOLUCIONALES

Tabla 1. Extracción de características en audio con redes neuronales convolucionales fuente: (García Mario Alejandro R. A.)

Año	2019
País/Ciudad	Argentina
Resumen	Este proyecto tiene como objetivo la realización de una estimación automática de parámetros de la voz basándose en la escala GRBAS, esto mediante características que son extraídas en el estudio del análisis acústico de señales de voz. En este trabajo se exponen resultados relacionados con un diseño de extracción de características de una red neuronal que se enfoca en la clasificación de la calidad de la voz
Aportes	Se realizó una evaluación de la voz enfocándose en una escala perceptual usada por fonoaudiólogos, mediante técnicas de machine learning.

Brechas	Este trabajo se desarrolló en torno a la escala GARBS, la cual solo evalúa la calidad en cuatro niveles, de manera que existe una gran diferencia entre un nivel y otro y por esto tener una mala clasificación
---------	---

3.2 RECONOCIMIENTO DE NORMAS EN MEDIDAS ACÚSTICAS PARA IDENTIFICAR PATOLOGÍAS EN LARINGE

Tabla 2. Reconocimiento de normas en medidas acústicas (Santos, 2017).

Año	2017
País/Ciudad	Curitiba
Resumen	A través de técnicas de procesamiento de señal digital (aplicadas a la vocal sostenida), medidas acústicas tales como jitter (o alteración de la frecuencia fundamental de la voz), shimmer (alteración de la amplitud de la señal de voz durante el tiempo de vocalización) y relación entre energía armónica y no armónica se puede obtener para cuantificar el funcionamiento anormal de la laringe con el fin de obtener resultados más precisos. A partir de la evaluación estadística de estas medidas, utilizando el uso de redes neuronales artificiales, es posible definir un modelo para la identificación de

	patologías que afectan las estructuras de la laringe.
Aportes	Con este proyecto se minimiza la incomodidad generada por los procedimientos de diagnóstico invasivos y Se reduce la subjetividad en el diagnóstico de enfermedades de la laringe utilizando técnicas de procesamiento de señal y redes neuronales.
Brechas	Las medidas utilizadas demostraron ser insuficientes para una buena clasificación. No se tuvo en cuenta una escala de clasificación vocal para estimar el estado de la voz.

3.3 CLASIFICACIÓN DE TRASTORNOS VOCALES UTILIZANDO REDES NEURONALES ARTIFICIALES

Tabla 3. Clasificación de trastornos vocales fuente: (Fermino, 2017).

Año	2017
País/Ciudad	Brasil/Cornélio procópio
Resumen	En este trabajo se aplican redes neuronales artificiales para clasificar patrones relacionados con trastornos vocales en las señales del habla. Se propone utilizar el paquete wavelet transformada para extraer características de la señal en análisis

	de frecuencia de tiempo, redes neuronales artificiales para el reconocimiento y clasificación de los patrones presentados por trastornos vocales.
Aportes	El clasificador modelado demostró ser consistente en las pruebas desarrolladas, demostrando ser prometedor para una aplicación futura como un método no invasivo.
Brechas	No se tuvo en cuenta una escala de clasificación vocal para estimar el estado de la voz.

3.4 RED NEURONAL PROFUNDA PARA ANÁLISIS ACÚSTICO.

Tabla 4. Red neuronal profunda para análisis acústico (García Mario Alejandro R. A., 2016).

Año	2016
País/Ciudad	Córdoba
Resumen	Por medio de análisis acústico de la voz, se intentan reconocer parámetros importantes que permitan clasificar de forma automática audios, y determinar si son voces patológicas, esto teniendo como estándar de calificación la escala GRBAS, con ayuda de técnicas de machine learning y Deep learning.

Aportes	Se diseña desde cero una red neuronal profunda para la clasificación de la calidad de la voz. Se reduce la subjetividad en el diagnóstico del estado de la voz.
Brechas	Este trabajo se desarrolló en torno a la escala GRABS.

3.5 IDENTIFICACIÓN DE PARÁMETROS SUBJETIVOS DE VOZ UTILIZANDO REDES NEURONALES ARTIFICIALES.

Tabla 5. Identificación de parámetros subjetivos de voz utilizando redes neuronales artificiales (Higor E. Pavoni, 2016).

Año	2016
País/Ciudad	São Carlos
Resumen	El método consiste en la clasificación de muestras de audio: rugosidad, soplosidad y tensión mediante el análisis de parámetros acústicos extraídos de señales previamente grabadas. Utilizando la energía extraída de cada uno de los coeficientes del tercer nivel de Wavelet Packet Transform como valores de entrada, una red neuronal artificial para clasificar las muestras de la clase deseada.
Aportes	La metodología propuesta pudo generalizar adecuadamente las muestras analizadas, resultando en un promedio de precisión del 95%.

Brechas	Este trabajo se desarrolló en una escala de evaluación perceptual de poca confiabilidad.
---------	--

3.6 ANÁLISIS PERCEPTUAL DE LA DISFONÍA: ADAPTACIÓN Y VALIDACIÓN DEL MÉTODO CAPE-V AL ESPAÑOL.

Tabla 6. Adaptación y validación del método CAPE-V al español (Galan, 2014).

Año	2014
País/Ciudad	Oviedo
Resumen	Con este trabajo se buscó adaptar la escala de evaluación perceptual CAPE-V al español, ya que dolo se encontraba disponible en idioma inglés.
Aportes	Adaptaron esta escala que es lineal, permitiendo puntos intermedios en la puntuación y contribuye en un mejor seguimiento de la calidad vocal del paciente.
Brechas	Los valores de calificación de la prueba son dados por el fonoaudiólogo y por lo tanto se tiene alta subjetividad en el diagnóstico. Al no estar automatizado se convierte en un proceso tedioso para el profesional, ya que debe realizar todo el proceso evaluativo de forma manual y esto conlleva un cierto tiempo.

4 CONSTRUCCIÓN DEL DATASET

Este capítulo presenta cada uno de los componentes teóricos y prácticos necesarios en la implementación y desarrollo del dataset, durante el desarrollo de este trabajo se emplea la metodología CDIO (concebir, diseñar, implementar y operar), por medio del cual se puede realizar un proceso de aprendizaje, permitiendo la creación y fijación del objetivo teniendo en cuenta los requisitos necesarios para desarrollarlo, al culminar esta fase se puede realizar el diseño y comenzar con su ejecución, para finalmente emplear y utilizar el dataset.

4.1 CONSTRUCCIÓN DE UN BANCO DE VOCES

4.1.1 Base de datos

Se crea una base de datos utilizando el programa Excel, esta se utilizará para clasificar el nivel de cada uno de los parámetros correspondientes a cada audio sintético que será generado mediante el programa EVAPER, este es un sistema de evaluación perceptual basado en el método psico-acústico de producción de la magnitud y “matching”, creado por el grupo de investigación CONICET de Argentina, este programa se adquirió en la universidad mediante un proyecto de investigación en la facultad.

En esta base de datos se agregarán los datos que se irán obteniendo conforme al desarrollo del algoritmo.

NOMBRE AUDIO	TIMBRE VOCAL	RONQUERA ASPEREZA	SOPLOCIDAD	ASTENIA	TENSION
2	MASCULINO	0	0	-3	0
2	MASCULINO	0	0	-3	0
2	MASCULINO	0	0	-3	0
2	MASCULINO	0	0	-3	0
2	MASCULINO	0	0	-3	0
2	MASCULINO	0	0	-3	0
3	MASCULINO	0	0	-2	0
3	MASCULINO	0	0	-2	0
3	MASCULINO	0	0	-2	0
3	MASCULINO	0	0	-2	0
3	MASCULINO	0	0	-2	0
3	MASCULINO	0	0	-2	0
4	MASCULINO	0	0	0	1
4	MASCULINO	0	0	0	1
4	MASCULINO	0	0	0	1
4	MASCULINO	0	0	0	1
4	MASCULINO	0	0	0	1
4	MASCULINO	0	0	0	1
5	MASCULINO	0	0	0	2
5	MASCULINO	0	0	0	2

Figura 7. Base de datos creada para almacenar datos importantes del desarrollo del algoritmo.

4.1.2 Creación de audios sintéticos

Se obtuvieron 500 audios sintéticos mediante el uso del programa EVAPER, el cual cuenta con una interfaz que nos permite observar la forma de onda correspondiente a cada audio que se va a generar, y cuenta con botones de parámetros variables como se observa en la Figura (8), estos pueden ser variados de 0 (no introduce ningún grado de cambio) a 6 (mayor grado de cambio), los parámetros ajustables son:

- Aspereza/ronquera (roughness)
- Ruido o escape de aire o soplosidad (breathiness)
- Astenia o tensión (asthenia-strain)

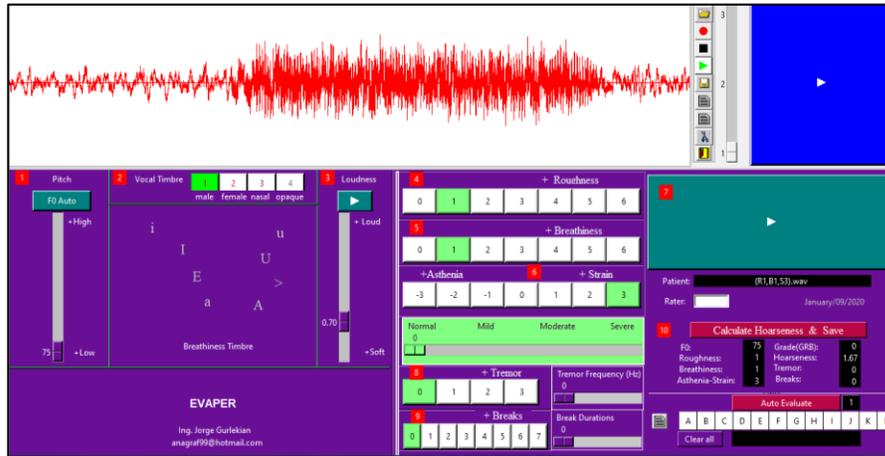


Figura 8. Pantalla principal de la interface gráfica del usuario (IGU) del sistema EVAPER.

El programa nos permite seleccionar el nivel de intensidad en cada audio (normal, leve, moderado y severo).

Para obtener estos audios, se variaron todos los niveles de soplosidad que proporciona el programa (0 a 6), y para los otros parámetros (ronquera, astenia y tensión), también se variaron los valores en diferentes combinaciones, en los cuales esto se hizo en la grabación de cada audio, de tal forma que se pudieran generar para un valor de soplosidad diferentes voces que pueden estar clasificadas como, voces normales, sanas o con patologías.

Las combinaciones implementadas para la creación de audio se pueden observar en la Figura 9. En estas se variaron todos los parámetros ajustables, con el fin de que cada audio creado contara con algún grado o nivel de soplosidad, aspereza, astenia o tensión.

Figura 9. Combinaciones utilizadas para la creación de audios sintético con Evaper, Se usaron abreviaciones que representan los parámetros, donde R, es Aspereza, B es soplosidad, A es astenia y S tensión.

(R0,B0,A-2)	(R1,B0,A-2)	(R2,B0,A-2)	(R3,B0,A-2)	(R4,B0,A-2)	(R5,B0,A-2)	(R6,B0,A-2)
(R0,B0,A-1)	(R1,B0,A-1)	(R2,B0,A-1)	(R3,B0,A-1)	(R4,B0,A-1)	(R5,B0,A-1)	(R6,B0,A-1)
(R0,B0,S1)	(R1,B0,S1)	(R2,B0,S1)	(R3,B0,S1)	(R4,B0,S1)	(R5,B0,S1)	(R6,B0,S1)
(R0,B0,S2)	(R1,B0,S2)	(R2,B0,S2)	(R3,B0,S2)	(R4,B0,S2)	(R5,B0,S2)	(R6,B0,S2)
(R0,B0,S3)	(R1,B0,S3)	(R2,B0,S3)	(R3,B0,S3)	(R4,B0,S3)	(R5,B0,S3)	(R6,B0,S3)
(R0,B1,A-3)	(R1,B1,A-3)	(R2,B1,A-3)	(R3,B1,A-3)	(R4,B1,A-3)	(R5,B1,A-3)	(R6,B1,A-3)
(R0,B1,A-2)	(R1,B1,A-2)	(R2,B1,A-2)	(R3,B1,A-2)	(R4,B1,A-2)	(R5,B1,A-2)	(R6,B1,A-2)
(R0,B1,A-1)	(R1,B1,A-1)	(R2,B1,A-1)	(R3,B1,A-1)	(R4,B1,A-1)	(R5,B1,A-1)	(R6,B1,A-1)
(R0,B1,S1)	(R1,B1,S1)	(R2,B1,S1)	(R3,B1,S1)	(R4,B1,S1)	(R5,B1,S1)	(R6,B1,S1)
(R0,B1,S2)	(R1,B1,S2)	(R2,B1,S2)	(R3,B1,S2)	(R4,B1,S2)	(R5,B1,S2)	(R6,B1,S2)
(R0,B1,S3)	(R1,B1,S3)	(R2,B1,S3)	(R3,B1,S3)	(R4,B1,S3)	(R5,B1,S3)	(R6,B1,S3)
(R0,B2,A-3)	(R1,B2,A-3)	(R2,B2,A-3)	(R3,B2,A-3)	(R4,B2,A-3)	(R5,B2,A-3)	(R6,B2,A-3)
(R0,B2,A-2)	(R1,B2,A-2)	(R2,B2,A-2)	(R3,B2,A-2)	(R4,B2,A-2)	(R5,B2,A-2)	(R6,B2,A-2)
(R0,B2,A-1)	(R1,B2,A-1)	(R2,B2,A-1)	(R3,B2,A-1)	(R4,B2,A-1)	(R5,B2,A-1)	(R6,B2,A-1)
(R0,B2,S1)	(R1,B2,S1)	(R2,B2,S1)	(R3,B2,S1)	(R4,B2,S1)	(R5,B2,S1)	(R6,B2,S1)
(R0,B2,S2)	(R1,B2,S2)	(R2,B2,S2)	(R3,B2,S2)	(R4,B2,S2)	(R5,B2,S2)	(R6,B2,S2)
(R0,B2,S3)	(R1,B2,S3)	(R2,B2,S3)	(R3,B2,S3)	(R4,B2,S3)	(R5,B2,S3)	(R6,B2,S3)
(R0,B3,A-3)	(R1,B3,A-3)	(R2,B3,A-3)	(R3,B3,A-3)	(R4,B3,A-3)	(R5,B3,A-3)	(R6,B3,A-3)
(R0,B3,A-2)	(R1,B3,A-2)	(R2,B3,A-2)	(R3,B3,A-2)	(R4,B3,A-2)	(R5,B3,A-2)	(R6,B3,A-2)
(R0,B3,A-1)	(R1,B3,A-1)	(R2,B3,A-1)	(R3,B3,A-1)	(R4,B3,A-1)	(R5,B3,A-1)	(R6,B3,A-1)
(R0,B3,S1)	(R1,B3,S1)	(R2,B3,S1)	(R3,B3,S1)	(R4,B3,S1)	(R5,B3,S1)	(R6,B3,S1)
(R0,B3,S2)	(R1,B3,S2)	(R2,B3,S2)	(R3,B3,S2)	(R4,B3,S2)	(R5,B3,S2)	(R6,B3,S2)
(R0,B3,S3)	(R1,B3,S3)	(R2,B3,S3)	(R3,B3,S3)	(R4,B3,S3)	(R5,B3,S3)	(R6,B3,S3)
(R0,B4,A-3)	(R1,B4,A-3)	(R2,B4,A-3)	(R3,B4,A-3)	(R4,B4,A-3)	(R5,B4,A-3)	(R6,B4,A-3)
(R0,B4,A-2)	(R1,B4,A-2)	(R2,B4,A-2)	(R3,B4,A-2)	(R4,B4,A-2)	(R5,B4,A-2)	(R6,B4,A-2)
(R0,B4,A-1)	(R1,B4,A-1)	(R2,B4,A-1)	(R3,B4,A-1)	(R4,B4,A-1)	(R5,B4,A-1)	(R6,B4,A-1)
(R0,B4,S1)	(R1,B4,S1)	(R2,B4,S1)	(R3,B4,S1)	(R4,B4,S1)	(R5,B4,S1)	(R6,B4,S1)
(R0,B4,S2)	(R1,B4,S2)	(R2,B4,S2)	(R3,B4,S2)	(R4,B4,S2)	(R5,B4,S2)	(R6,B4,S2)
(R0,B4,S3)	(R1,B4,S3)	(R2,B4,S3)	(R3,B4,S3)	(R4,B4,S3)	(R5,B4,S3)	(R6,B4,S3)
(R0,B5,A-3)	(R1,B5,A-3)	(R2,B5,A-3)	(R3,B5,A-3)	(R4,B5,A-3)	(R5,B5,A-3)	(R6,B5,A-3)
(R0,B5,A-2)	(R1,B5,A-2)	(R2,B5,A-2)	(R3,B5,A-2)	(R4,B5,A-2)	(R5,B5,A-2)	(R6,B5,A-2)
(R0,B5,A-1)	(R1,B5,A-1)	(R2,B5,A-1)	(R3,B5,A-1)	(R4,B5,A-1)	(R5,B5,A-1)	(R6,B5,A-1)
(R0,B5,S1)	(R1,B5,S1)	(R2,B5,S1)	(R3,B5,S1)	(R4,B5,S1)	(R5,B5,S1)	(R6,B5,S1)
(R0,B5,S2)	(R1,B5,S2)	(R2,B5,S2)	(R3,B5,S2)	(R4,B5,S2)	(R5,B5,S2)	(R6,B5,S2)
(R0,B5,S3)	(R1,B5,S3)	(R2,B5,S3)	(R3,B5,S3)	(R4,B5,S3)	(R5,B5,S3)	(R6,B5,S3)
(R0,B6,A-3)	(R1,B6,A-3)	(R2,B6,A-3)	(R3,B6,A-3)	(R4,B6,A-3)	(R5,B6,A-3)	(R6,B6,A-3)
(R0,B6,A-2)	(R1,B6,A-2)	(R2,B6,A-2)	(R3,B6,A-2)	(R4,B6,A-2)	(R5,B6,A-2)	(R6,B6,A-2)

4.1.3 Edición de los audios

Los audios creados con el programa EVAPER son recortados haciendo uso de un editor de audio llamado Audacity (Figura 10), el cual nos permite eliminar la parte del inicio y final de los audios que no nos aporta información en el análisis de estos. Con esta herramienta se logra que todos los audios queden del mismo tamaño (1,30s).

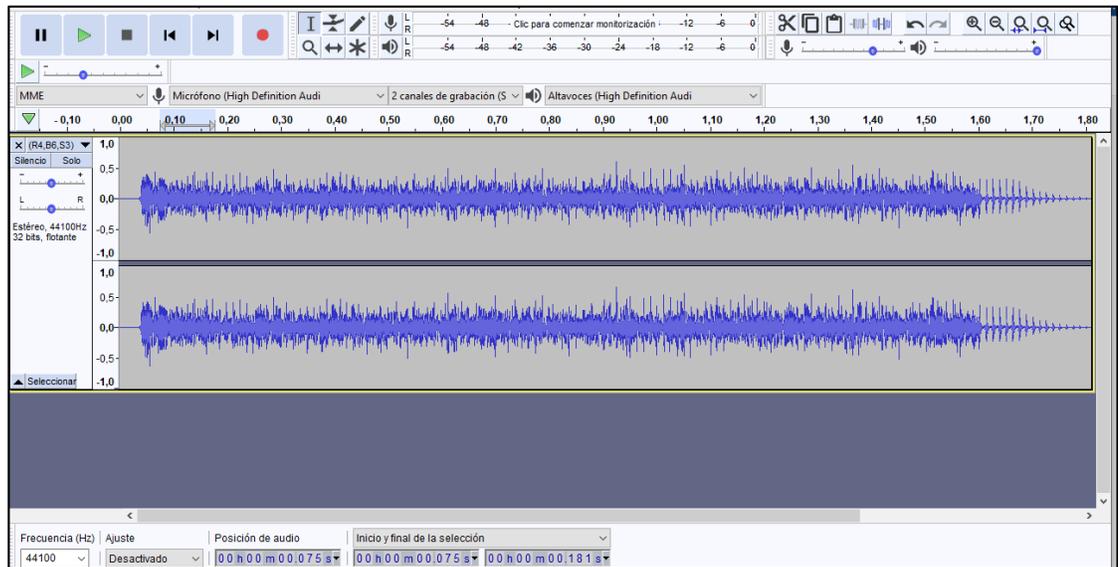


Figura 10. Pantalla principal de la interface grfica del usuario del programa Audacity, para edici3n de audios.

4.2 ALGORITMO DE EXTRACCI3N AUTOMTICO DE CARACTERSTICAS

Se gener3 un dataset aplicando el algoritmo de extracci3n automtico al banco de voces creado, Figura 11. En este algoritmo se busc3 analizar y extraer 42 diferentes parmetros acsticos de los audios grabados, estos se pueden clasificar como parmetros temporales, espectrales, cepstrales y de potencia.

Se crea una funci3n en MATLAB la cual calcula cada parmetro, los parmetros que se calcula son:

- a) Loudness (volumen): este parmetro se calcul3 mediante una funci3n de MATLAB(integratedLoudness), la cual mide el volumen integrado y el rango de volumen de la seal.
- b) Presi3n del sonido: este parmetro se calcul3 haciendo uso de la funci3n de MATLAB(splMeter), esta permite medir el nivel de presi3n sonora de la seal de audio.

- c) Jitter, jitt, shimmer: para el cálculo de estos parámetros, se normaliza la señal y se invierte, con el fin de que también se tenga en cuenta los picos negativos de esta. Seguidamente se identificaron los picos máximos y mínimos, esto para la señal normal y también la señal invertida, esto se hace mediante la función de MATLAB (findpeaks) estableciendo el valor de prominencia que se va a tener en cuenta para la selección de los picos, el valor de prominencia establecido fue de (1.1) y (1.0) respectivamente para cada señal. Posteriormente se calcula el vector de periodos y la amplitud pico a pico. Por último, se calcula estos parámetros teniendo en cuenta las ecuaciones de jitter (2), jitt (3) y shimmer (4).
- d) Pitch con periodos: se obtuvo implementando la ecuación ya mencionada anteriormente (1), utilizando el promedio de los periodos ya obtenidos.
- e) Ensanchamiento espectro: para obtener este parámetro, se hizo por medio de la función de MATLAB(spectralSpread), que nos permite encontrar la difusión espectral en señales de audio, que representa el "ancho de banda instantáneo" del espectro. Se utiliza como una indicación del dominio de un tono.
- f) Punto de caída espectral: se implementó la función de MATLAB (spectral Rolloff Point), esta mide el ancho de banda de la señal de audio determinando el intervalo de frecuencia bajo el cual existe un porcentaje dado de la energía total.
- g) Centroide espectral: representa el "centro de gravedad" del espectro. Se usa como una indicación de brillo, para obtenerlo se usó la función de MATLAB (Spectral Centroid), esta calcula la suma ponderada en frecuencia normalizada por la suma no ponderada.

- h) Asimetría espectral: este parámetro se calculó mediante una función de MATLAB (Spectral Skewness), la asimetría espectral mide la simetría alrededor del centroide.
- i) Curtosis espectral: mide la planitud, o no Gaussianidad, del espectro alrededor de su centroide. Por el contrario, se usa para indicar el pico de un espectro y se obtuvo por medio de la función de MATLAB (Spectral Kurtosis).
- j) Flujo espectral: es una medida de la variabilidad del espectro a lo largo del tiempo, se usó la función de MATLAB (Spectral Flux).
- k) Planitud espectral: La planitud espectral es una indicación del pico del espectro. Una planitud espectral más alta indica ruido, mientras que una planitud espectral más baja indica tonalidad, se utilizó la función de MATLAB (Spectral Flatness), esta mide la relación entre la media geométrica del espectro y la media aritmética del espectro.
- l) Entropía espectral: la entropía es una medida del desorden, las regiones del habla sonora tienen una entropía menor en comparación con las regiones del habla no sonora, para calcularla se hace uso de la función de MATLAB (Spectral Entropy), la cual mide el pico del espectro.
- m) Pitch con función: se estimó la frecuencia fundamental de la señal de audio, mediante la función de MATLAB(pitch).
- n) Análisis pico del cepstrum: Este parámetro se calculó mediante la función de MATLAB (cpeps), la cual devuelve el cepstrum complejo de la secuencia de datos real utilizando la transformada de fourier dos veces.

- o) Coeficientes cepstrales de frecuencia de mel: coeficientes que son calculados mediante diferentes bandas de frecuencias que se aplican a modo de filtro, este parámetro se obtuvo haciendo uso de la función de MATLAB (mfcc), la cual devuelve los coeficientes cepstrales de frecuencia de mel (MFCC) para la entrada de audio, muestreados a una frecuencia de Hz.

- p) Periodograma: es una estimación de la densidad espectral de una señal, la función de MATLAB (periodogram) nos permite obtenerlo, calculando la estimación de densidad espectral de potencia de periodograma, con este también se obtiene el formante uno, que viene siendo la frecuencia fundamental basada en el periodograma.

- q) Cálculo del ancho de banda: este parámetro se obtuvo haciendo uso de la función de MATLAB (powerbw), esta devuelve el ancho de banda de 3 dB. Y con este parámetro podemos llegar a obtener también una frecuencia fundamental basada en el ancho de banda.

- r) Relación señal a ruido: es la proporción existente entre el nivel de potencia de la señal y el nivel de potencia de ruido que la corrompe. Usualmente se expresa en decibeles (dB), para obtener este parámetro se utilizó la función de MATLAB (snr), esta devuelve la relación señal-ruido (SNR) en decibelios de una señal, calculando la relación de su magnitud cuadrada sumada a la del ruido.

- s) Distorsión total armónica: es una medida de la distorsión armónica presente y se define como el cociente de la suma de las potencias de todos los componentes armónicos, o de un armónico en particular respecto a la potencia de la Frecuencia fundamental, este parámetro se obtuvo mediante la función de MATLAB (thd), esta es determina a partir de la frecuencia fundamental y los primeros cinco armónicos utilizando un periodo grama modificado de la misma longitud que la señal de entrada.

t) Relación armónico ruido: Establece la relación entre el armónico y el nivel de ruido que compone la señal, separa la señal en dos canales distintos y luego realiza una comparación entre la intensidad de los dos. Esta se calculó despejando a partir de la fórmula de la relación señal a ruido y teniendo así que la potencia de los armónicos divididos de la potencia del ruido nos da como resultado la relación armónico ruido.

Después de esto se tomó el banco de voces y se submuestreo 6 veces sacando 300 ms con solapamiento de 100 ms en cada audio.

audiodataset															
id_audio	genero	rugosidad	soplosidad	astenia	tension	presion_son...	ensancham...	asimetria_e...	rolloffPoint	kurtosis_es...	flujo_espec...	planitud_es...	shimmer	jitta	
Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
2	0	0	0	-3	0	65.014	341.13	14.65	674.07	649.91	0.014353	0.0017785	0.078029	10.798	
3	0	0	0	-3	0	65.023	340.15	14.678	674.07	651.26	0.014597	0.0017708	0.042931	10.798	
4	0	0	0	-3	0	65.012	339.37	14.68	669.14	650	0.014724	0.0017668	0.03603	10.798	
5	0	0	0	-3	0	64.965	338.55	14.72	667.9	652.55	0.0148	0.0017623	0.040515	10.798	
6	0	0	0	-3	0	64.903	337.68	14.754	667.9	654.51	0.014894	0.0017563	0.044246	10.798	
7	0	0	0	-3	0	64.837	336.88	14.79	667.9	656.86	0.014995	0.0017513	0.046037	10.798	
8	0	0	0	-2	0	65.322	353.62	16.774	725.93	687.74	0.0091478	0.0014313	0.30566	11.878	
9	0	0	0	-2	0	65.133	352.44	16.783	725.93	690.22	0.0103	0.0014142	0.043541	11.878	
10	0	0	0	-2	0	65.119	351.4	16.762	725.93	690.53	0.010379	0.0014028	0.037064	11.878	
11	0	0	0	-2	0	65.067	350.55	16.75	725.93	691.38	0.010399	0.0013915	0.039265	11.878	
12	0	0	0	-2	0	64.997	349.69	16.743	725.93	692.11	0.010435	0.0013804	0.043725	11.878	
13	0	0	0	-2	0	64.926	348.83	16.734	725.93	693.33	0.010477	0.0013692	0.061203	11.878	
14	0	0	0	0	1	63.267	501.86	6.5275	1200	134.77	0.0071355	0.0012858	0.14055	11.878	
15	0	0	0	0	1	63.321	501.1	6.5599	1200	135.88	0.0072088	0.0012958	0.11205	11.878	
16	0	0	0	0	1	63.349	502.89	6.5353	1200	134.66	0.0072266	0.0013051	0.11631	11.878	
17	0	0	0	0	1	63.343	504.5	6.5208	1200	133.79	0.007222	0.0013156	0.095232	11.878	
18	0	0	0	0	1	63.319	506.52	6.4885	1204.9	132.3	0.0072003	0.0013245	0.090611	11.878	
19	0	0	0	0	1	63.283	508.49	6.4712	1206.2	131.33	0.0071847	0.0013361	0.089645	11.878	
20	0	0	0	0	2	62.837	676.45	5.1297	2021	75.55	0.0052467	0.0023427	0.16607	10.798	
21	0	0	0	0	2	62.917	673.91	5.1553	1980.2	76.289	0.0053047	0.0023281	0.13731	10.798	

Figura 11. Dataset obtenido en MATLAB a partir de los 500 audios creados.

El dataset que se obtuvo, está compuesto por 3000 muestras en su totalidad Figura 11.

5. SELECCIÓN DE ATRIBUTOS

En este capítulo se da a conocer en detalle cada uno de los modelos y programas utilizados para la selección de atributos que se implementó con el dataset creado en el capítulo 4.

5.1 SOFTWARE WEKA

La selección de atributos y la clasificación se hizo por medio del software Weka, el cual nos permite hacer análisis de datos mediante técnicas de machine learning, este fue creado por la universidad de Waikato. Cuenta con diferentes herramientas que nos permite realizar diferentes procesos, dentro de estos, modelos de regresión y selección de atributos, proporcionando diferentes algoritmos que nos permite analizar datos. En su interface de usuario nos muestra las diferentes opciones que se pueden seleccionar. Figura 12. Este software proporciona varias herramientas de análisis de datos dentro de estos, procesos de selección de atributos, clasificación y regresión.

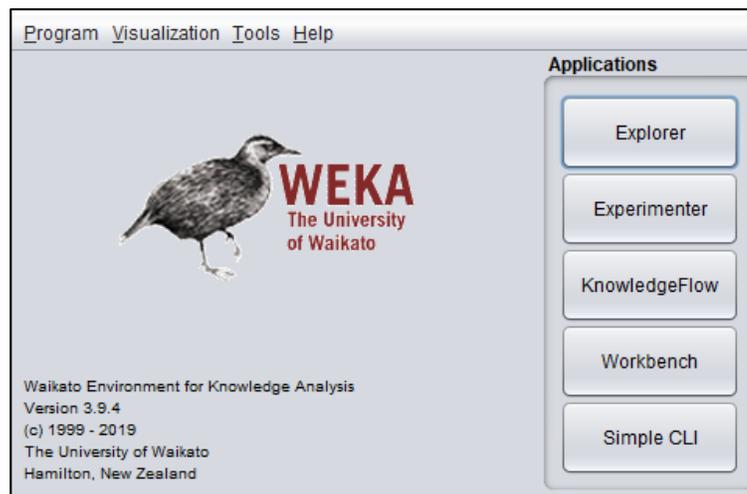


Figura 12. Interface gráfica de usuario, Weka.

Este software cuenta con una interface del usuario que me permiten acceder al entorno de trabajo:

El entorno de trabajo me proporciona la opción de preprocesamiento de datos, en el que puedo importar una base de datos y hacer un tratamiento de los mismos haciendo un algoritmo en el que me permite filtrar ciertos atributos, este entorno se puede visualizar en la Figura 13.

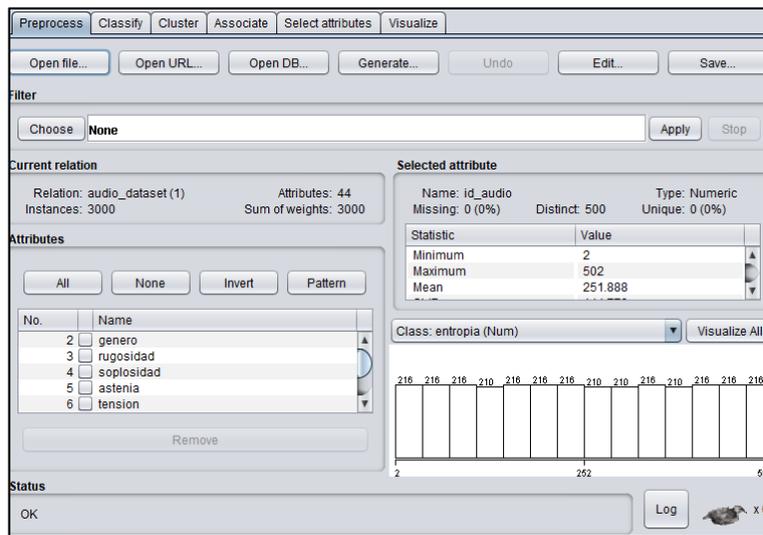


Figura 13. Panel de pre procesado, Weka.

El panel de trabajo proporciona modelos para seleccionar los parámetros más predictivos en un conjunto de datos Figura 14.

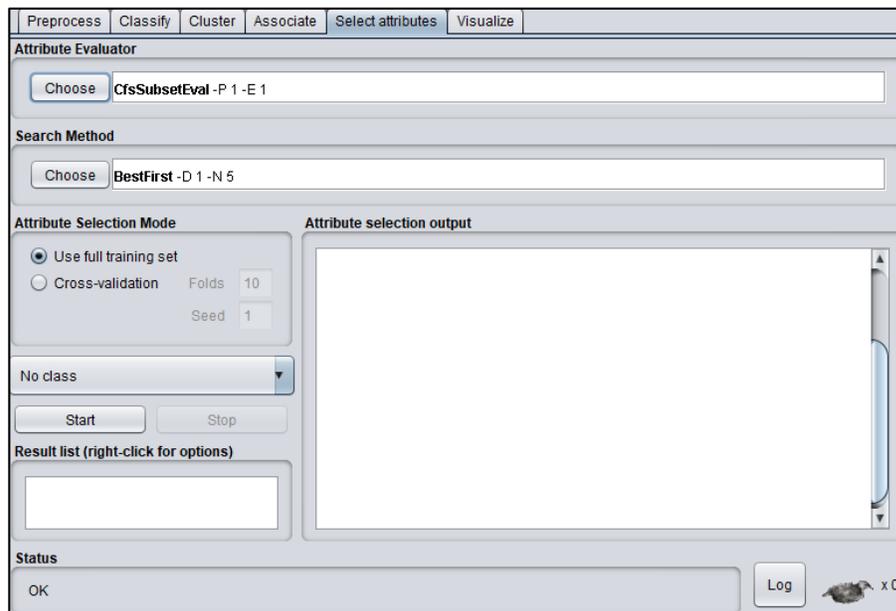


Figura 14. Panel de selección de atributos, Weka.

En la interface se presenta un entorno de trabajo que permite realizar procesos de clasificación y regresión a los subconjuntos de datos, permitiendo establecer el error obtenido con cada modelo.

5.1.1 Selección de atributos mediante Weka

Se selecciona mediante este método los parámetros más importantes en el momento de realizar la regresión, es decir, los parámetros más predictivos de un conjunto de datos ingresados.

Se debe seleccionar el método que evaluara cada uno de los atributos, este le asignara un peso a cada parámetro basándose en la evaluación realizada. Una vez seleccionado el método de evaluación y el de generación del espacio se elegí el modo de selección de atributo, el atributo que representa la clasificación conocida, en nuestro caso soplosidad. En la ventana attribute selection output nos mostrará los resultados habiendo tomado los mejores atributos.

- **Opciones de selección que proporciona el evaluador de atributos**

Valor de atributo de correlación: evalúa el valor de un atributo midiendo la correlación (Pearson) entre él y la clase. Los atributos nominales se consideran valor por valor al tratar cada valor como un indicador. Se llega a una correlación general para un atributo nominal a través de un promedio ponderado, este nos entrega un valor de correlación entre -100 y +100, para los valores de interpretación de la escala se utiliza la siguiente referencia (Cabrera, 2009), en la que se interpreta la correlación de la siguiente manera:

0 – 0,25: Escasa o nula

0,26-0,50: Débil

0,51- 0,75: Entre moderada y fuerte

0,76- 1,00: Entre fuerte y perfecta

Cfssubseteval: Estudia un subconjunto de datos conformado con diferentes atributos, seleccionando aquellos parámetros o atributos con un mayor peso para una buena estimación. Se tiene preferencia sobre los grupos de atributos que están altamente correlacionados con la salida, mientras que tienen una baja intercorrelación entre estos.

Atributo del clasificador eval: Estudia el valor de un parámetro, del conjunto de datos haciendo uso de un clasificador que se seleccione.

Evaluador de subconjunto clasificador: evalúa los subconjuntos de atributos en los datos de entrenamiento o un conjunto de pruebas de retención independiente. Utiliza un clasificador para estimar el "mérito" de un conjunto de atributos.

Componente principales-transformador de atributos: Esta opción analiza los componentes o atributos principales escogiendo los más importantes para la estimación. La reducción de la dimensionalidad se logra eligiendo suficientes vectores propios para dar cuenta de algún porcentaje de la varianza en los datos originales, por defecto 0,95 (95%).

Relieffatributeeval: El principio de este método está en asignar pesos a los atributos con mayor influencia en la estimación, los parámetros más predictivos, basándose en los vecinos más cercanos, teniendo la capacidad de distinguir las características que están más cerca la una de la otras.

Wrappersubseteval: evalúa conjuntos de atributos utilizando un esquema de aprendizaje. La validación cruzada se usa para estimar la precisión del esquema de aprendizaje para un conjunto de atributos.

- **Métodos de búsqueda que proporciona la selección de atributos**

BestFirst: busca en el espacio de subconjuntos de atributos mediante escalada codiciosa aumentada con una función de retroceso. Es un caso en el cual se selecciona un nodo para la expansión basada en una función de evaluación, esta función devuelve un número que sirve para representar lo deseable o indeseable que sería la expansión de un nodo. Se expande primero aquel nodo que tiene mejor evaluación.

GreedyStepwise: Este método de búsqueda, realiza la inspección sin retroceso por cada subconjunto, y selecciona el subconjunto más óptimo para una buena estimación, este método sigue buscando y va seleccionando el mejor de todos hasta que la solución y la elección ya no puede mejorar.

Ranker: clasifica los atributos por sus evaluaciones individuales

- **Modos de selección de atributos**

Training set: Se hace uso de todos los datos para realizar el entrenamiento del método y luego se emplea el método con los mismos datos usados para el entrenamiento.

Validación cruzada: Con este método se entrena el modelo haciendo una validación cruzada del número de folds dados. En este método primero se entrena el modelo con una de las particiones dada y se hace test con el resto de datos dado, posteriormente toma la siguiente partición y se entrena con todos los datos disponibles de nuevo, repite este proceso hasta terminar con el número de folds.

5.1.2 Prueba con la herramienta de selección de atributos

Se realizará la selección de atributos probando con las diferentes opciones que proporciona Weka. En primer lugar, se realizó la selección de atributos en el modo de búsqueda use training set, y posteriormente se procedió hacer pruebas de selección de atributos en el modo de Cross validation.

- **Prueba con modo de selección Use training set**

Valor de atributo de correlación- Ranker

Este evaluador y método de búsqueda, ordena los atributos por importancia, basándose en la medida de correlación de cada atributo respecto a la clase. De los resultados obtenidos a partir de este evaluador y método de búsqueda se seleccionaron 4 atributos basándonos en la escala de correlación ya nombrada anteriormente, en el rango entre 0,51-0,75, la cual establece una correlación entre moderada y fuerte. Los resultados obtenidos se pueden observar en la Figura 15. Los atributos seleccionados se muestran en la Tabla 7.

Tabla 7. Atributos seleccionados modo de correlación y método de búsqueda Ranker

Atributos seleccionados	Medida de correlación del atributo con la clase
Jitt	0.76738
Shimmer	0.73452
Jitta	0.64501
Mffcc_10	0.6289

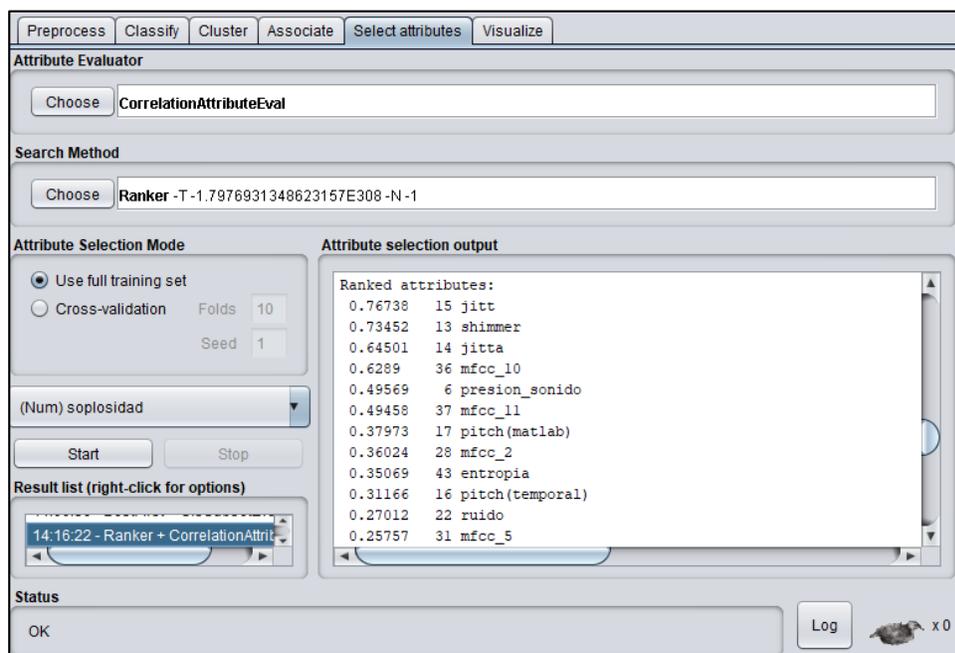


Figura 15. Selección de atributos modelo de atributo evaluador de correlación y método de búsqueda Ranker

Cfssubseteval- bestfirst

Los resultados obtenidos a partir de este evaluador y método de búsqueda se pueden observar en la Figura 16. En este se considera el valor predictivo (correlación) de cada atributo y de su redundancia. El subconjunto de atributos seleccionado se muestra en la Tabla 8.

Tabla 8. Atributos seleccionados evaluador Cfssubsetval y método de búsqueda bestfirst

Subconjunto seleccionado
presion_sonido
asimetria_espectro
Shimmer
Jitt
SNR
mfcc_4
mfcc_5
mfcc_7
mfcc_9
mfcc_10
mfcc_11
mfcc_14

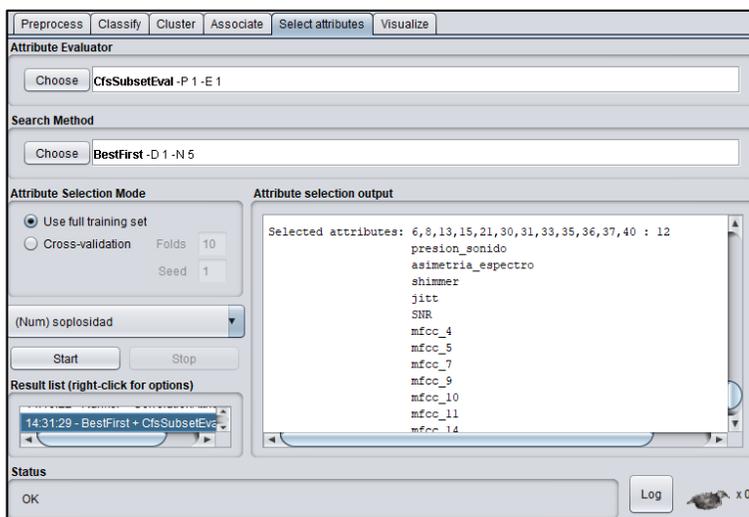


Figura 16. Selección de atributos modelo de atributo evaluador cfssubseteval y método de búsqueda bestfirst

Cfssubseteval –greedystepwise

Los resultados obtenidos a partir de este evaluador y método de búsqueda se pueden observar en la Figura 17, en este método evaluador se considera el valor de correlación de cada atributo con respecto a la clase. El método busca seleccionar un solo subconjunto de atributos, este subconjunto se muestran en la Tabla 9.

Tabla 9. Atributos seleccionados evaluador cfssubseteval y método de búsqueda greedystepwise

Subconjunto seleccionado
presion_sonido
asimetria_espectro
Shimmer
Jitt
SNR
mfcc_4
mfcc_5

mfcc_7
mfcc_9
mfcc_10
mfcc_11
mfcc_14

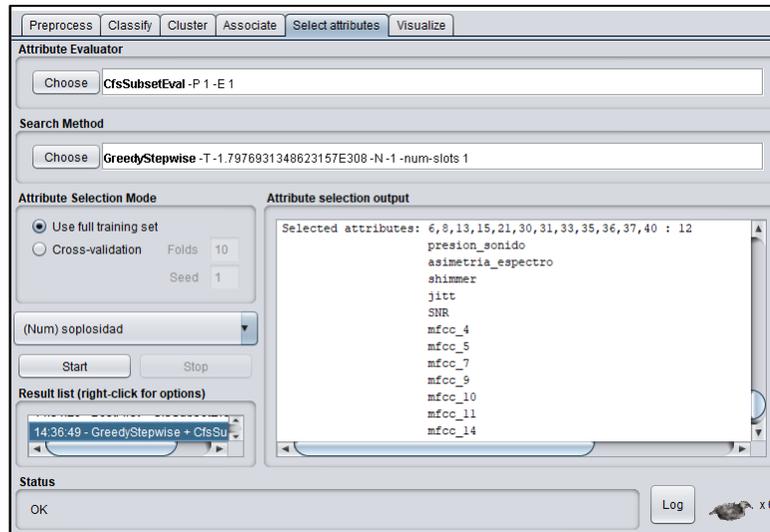


Figura 17. Resultado obtenido prueba con evaluador cfssubetval y método de búsqueda greedystepwise

Relieffattributeeval- ranker

Como criterio de selección se tuvieron en cuenta las referencias (villacampa, 2015) (Portocarrero, 2007), sin embargo no se encontró una escala de interpretación de los resultados estándar y por lo tanto se tuvieron en cuenta los primeros 10 atributos según el ranker que se obtuvo a partir de este evaluador y método de búsqueda, el cual estudia y establece un peso para cada atributo, seleccionando varias veces una instancia y buscando el valor del vecino más cercano de la misma salida, los resultados obtenidos se pueden observar en la Figura 18, los atributos escogidos se muestran en la Tabla 10.

Tabla 10. Atributos seleccionados evaluador Reliefattributeeval y método de búsqueda Ranker

Atributos seleccionados	Clasificación de atributo por peso
presion_sonido	0.0190038
Entropía	0.0186947
Jitt	0.0177283
Centroide	0.016591
Mfcc_14	0.0156204
Mfcc_10	0.0154299
Loudness	0.0146215
SNR	0.013504
Shimmer	0.0132657
mfcc_11	0.0118696

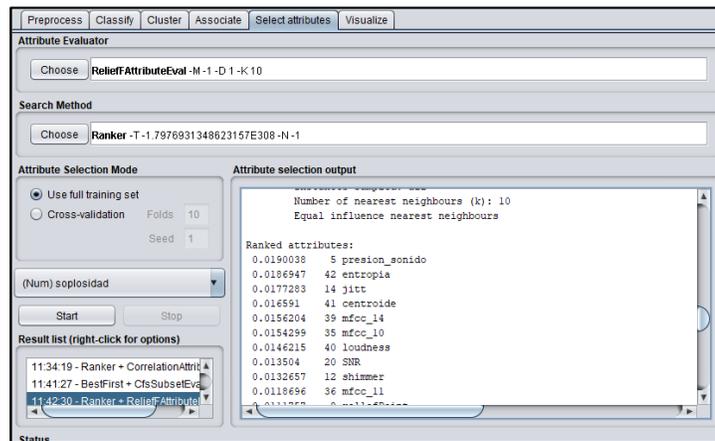


Figura 18. Resultado obtenido prueba con evaluador Reliefattributeeval y método de búsqueda ranker

- **Prueba con modo de selección cross-validation**

Mediante Weka se realizó una validación cruzada del número de particiones dadas en nuestro caso 10 folds. El método de validación cruzada utiliza muy bien los datos al calcular el porcentaje de aciertos esperado.

Valor de atributo de correlación- Ranker

Este evaluador y método de búsqueda, ordena los atributos por importancia, basándose en la medida de correlación de cada atributo respecto a la clase. De los resultados obtenidos a partir de este evaluador y método de búsqueda se seleccionaron 6 atributos basándonos en la escala de correlación ya nombrada anteriormente, en el rango entre 0,51-0,75, la cual establece una correlación entre moderada y fuerte. Los resultados obtenidos se pueden observar en la Figura 19, los atributos escogidos se muestran en la Tabla 11.

Tabla 11. Atributos seleccionados evaluador de correlación y método de búsqueda Ranker

Atributos seleccionados	Correlación atributo respecto a la clase
Jitt	0.767
Shimmer	0.735
Jitta	0.645
mfcc_10	0.629

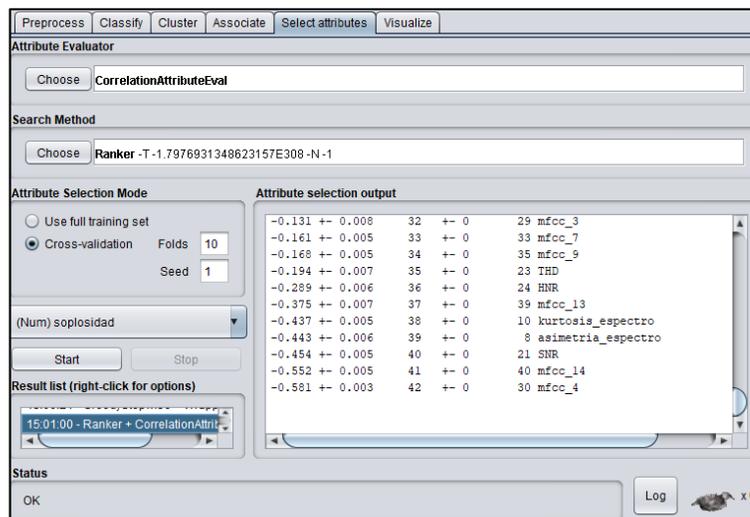


Figura 19. Resultado obtenido prueba con evaluador de atributo de correlación y método de búsqueda Ranker

Relieffatributeeval- Ranker

Como criterio de selección se tuvieron en cuenta las referencias (villacampa, 2015) (Portocarrero, 2007), sin embargo no se encontró una escala de interpretación de los resultados estándar y por lo tanto se tuvieron en cuenta los primeros 10 atributos según el ranker que se obtuvo a partir de este evaluador y método de búsqueda, el cual estudia y establece un peso para cada atributo, seleccionando varias veces una instancia y buscando el valor del vecino más cercano de la misma salida, los resultados obtenidos se pueden observar en la Figura 20, los atributos escogidos se muestran en la Tabla 12.

Tabla 12. Atributos seleccionados evaluador Relieffatributeeval y método de búsqueda Ranker

Atributos seleccionados	Clasificación de atributo por peso
Entropia	0.019
presion_sonido	0.018
Jitt	0.018
Centroide	0.016
Mfcc_14	0.016
Mfcc_10	0.015
Loudness	0.014
SNR	0.013
Shimmer	0.013
Mfcc_11	0.012

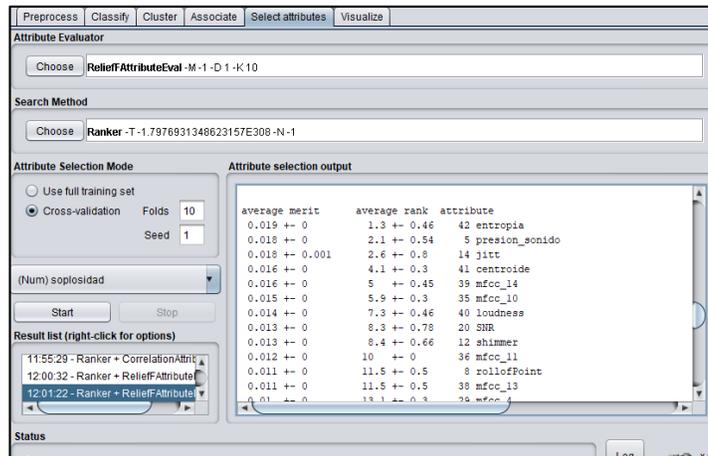


Figura 20. Resultados obtenidos con evaluador Relieffattributeeval y método de búsqueda Ranker

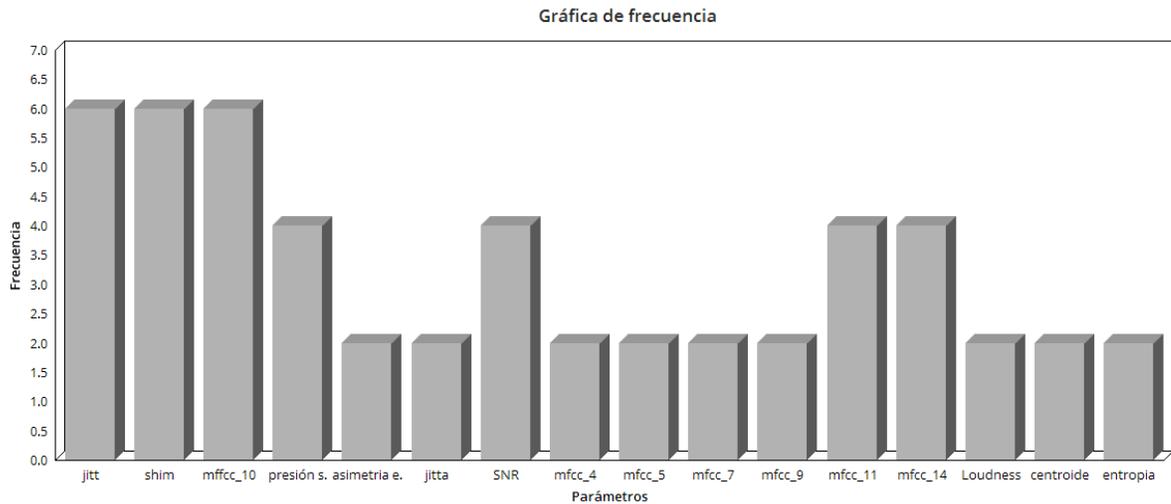
A partir de los resultados obtenidos con los diferentes evaluadores y métodos de búsqueda, se creó una tabla y grafico de frecuencia Tabla 13, Figura 21, para identificar los atributos que tenían una mayor recurrencia y así seleccionarlos para ser utilizados en el proceso de regresión.

Tabla 13. Frecuencia de datos obtenidos con evaluadores y métodos de búsqueda implementados en selección de atributos

Atributos	Frecuencia
Jitt	6
Shimmer	6
mfcc_10	6
presión_sonido	4
asimetria_espectro	2
Jitta	2
SNR	4
mfcc_4	2
mfcc_5	2
mfcc_7	2
mfcc_9	2
mfcc_11	4
mfcc_14	4
Loudness	2

centroide	2
entropia	2

Figura 21. Gráfica basada en tabla de distribución de frecuencias Tabla 13.



Basándonos en los resultados obtenidos a partir de la prueba realizada con las diferentes posibilidades que nos ofrece la herramienta de selección de atributos de Weka y teniendo en cuenta el análisis realizado con Tabla 13 y Figura 21. Se seleccionaron los siguientes 7 atributos que pueden observarse en la Tabla 14. Estos serán utilizados más adelante en el proceso de regresión.

Tabla 14. Selección final de parámetros a través de la herramienta selección de atributos de Weka

Parámetros seleccionados a partir de las pruebas realizadas
Jitt
Shimmer
Mfcc_10
Presión de sonido

SNR
mfcc_11
mfcc_14

6. MODELO DE REGRESIÓN

6.1 MODELO DE REGRESIÓN CON HERRAMIENTA WEKA

Weka nos proporciona la herramienta para realizar regresión en su interface Figura 22. Se puede utilizar la opción de regresión por diferentes algoritmos, con los datos ya cargados, en nuestro caso los datos que utilizamos son los 42 atributos del dataset generado en el capítulo anterior.

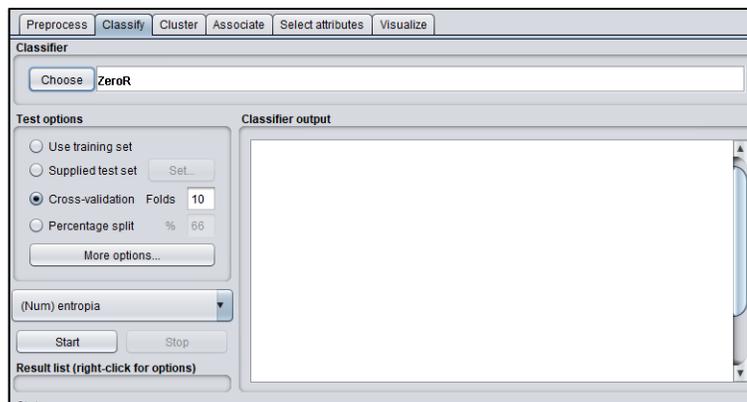


Figura 22. Bloque de clasificación, interfaz gráfica Weka

Después de elegir el modelo que será utilizado para el proceso de regresión, se establece el modo de entrenamiento (Test Options). Podemos elegir entre cuatro opciones de modo de prueba

training set: Al seleccionar esta opción el programa entrena el modelo con todos los atributos disponibles y luego, aplica el modelo ya entrenado sobre los mismos atributos

Supplied test set: Con esta elección el software nos proporciona la opción de escoger un conjunto de datos con el que el modelo será entrenado.

Validación cruzada: Con este método se entrena el modelo haciendo una validación cruzada del número de folds dados. En este método primero se entrena el modelo con una de las particiones dada y se hace test con el resto de datos dado, posteriormente toma la siguiente partición y se entrena con todos los datos disponibles de nuevo, repite este proceso hasta terminar con el número de folds

Percentage dividido: En este caso se divide el conjunto de datos o atributos en dos, la primera parte se utiliza para entrenar el modelo y la segunda para realizar el proceso de test.

6.1.1 Modelos de regresión que proporciona Weka

Los modelos de regresión que ofrece Weka y que serán utilizados, son los siguientes:

Linear regression: La regresión lineal es una técnica estadística que intenta construir un modelo para los datos analizados, y a través de éste predecir los datos futuros. Este modelo cuantifica la relación entre dos variables continuas: “la variable dependiente o la variable que intentamos predecir y la variable independiente o la variable predecible”.

Simple linear regression: modelo de regresión lineal simple. Selecciona el atributo que da como resultado el error cuadrático más bajo. Solo puede tratar con atributos numéricos.

Additive regression: meta clasificador que mejora el rendimiento de un clasificador. Cada interacción ajusta un modelo a los residuos que dejó el clasificador en la iteración anterior. La predicción se logra agregando las predicciones de cada clasificador. La reducción del parámetro de contracción (tasa

de aprendizaje) ayuda a evitar el sobreajuste y tiene un efecto de suavizado, pero aumenta el tiempo de aprendizaje.

Reregressionbydiscretization: esquema de regresión que emplea cualquier clasificador en una copia de los datos que tiene el atributo de clase discretizado. El valor predicho es el valor esperado del valor medio de la clase para cada intervalo discretizado (basado en las probabilidades pronosticadas para cada intervalo). Esta clase ahora también admite la estimación de densidad condicional al construir un estimador de densidad univariante a partir de los valores objetivo en los datos de entrenamiento, ponderado por las probabilidades de clase.

Gaussian processes: Implementa procesos gaussianos para la regresión sin ajuste de hiperparámetros. Los modelos de regresión de procesos gaussianos (GPR) son modelos probabilísticos no paramétricos basados en kernel.

M5P: Implementa rutinas básicas para generar árboles y reglas del Modelo M5, un árbol de decisión es un conjunto de condiciones o reglas organizadas en una estructura jerárquica, de tal manera que la decisión final se puede determinar siguiendo las condiciones que se cumplen desde la raíz hasta alguna de sus hojas, Cada hoja tiene asociada una clase que permite calcular el valor estimado de la instancia mediante una regresión lineal.

6.1.2 Prueba con modelos de regresión y modos de prueba de Weka

Se realizaron pruebas primero con 42 atributos del dataset, después de esto, se hicieron pruebas con los 7 atributos seleccionados anteriormente con la herramienta de selección de atributos de Weka, estos atributos seleccionados se indican en la Tabla 14.

- **Prueba con la totalidad de atributos del dataset**

Clasificador linear regression - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 23. Y registrados en Tabla 15.

Tabla 15. Valores obtenidos con el modo de regresión, linear regression - modo training set

Coefficiente de correlación	0.9732
Error absoluto medio	0.3243

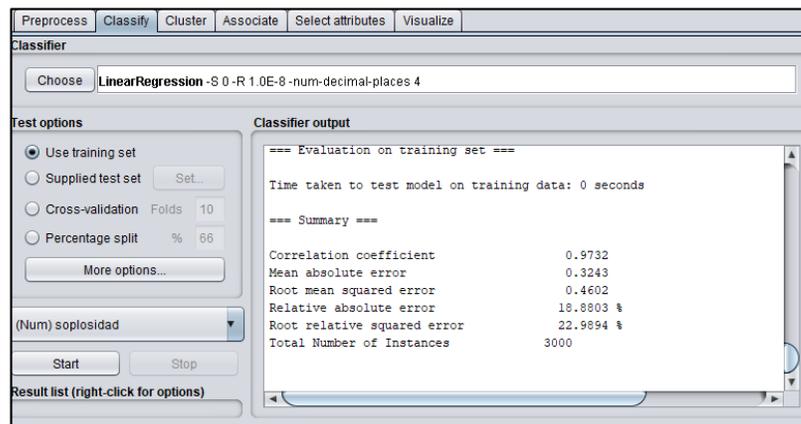


Figura 23. Resultados obtenidos modelo de regresión linear regression-modo training set

Clasificador linear regression – cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 24. Y registrados en Tabla 16.

Tabla 16. Valores obtenidos con el modo de regresión, linear regression – modo cross validation

Coefficiente de correlación	0.9722
Error absoluto medio	0.3294

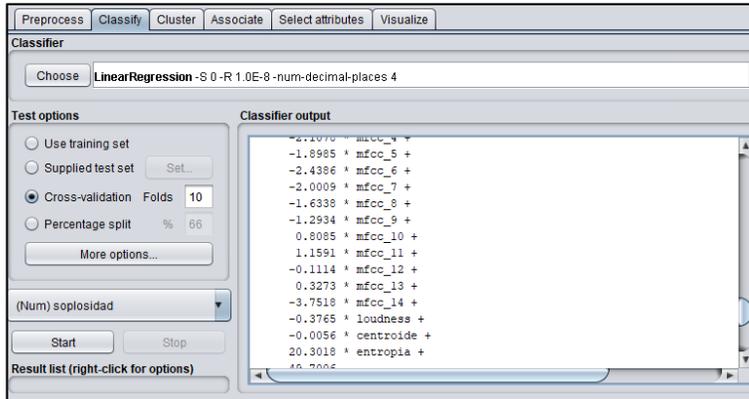


Figura 24. Resultados obtenidos modelo de regresión linear regression-modo cross validation

Clasificador linear regression – percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 25. Y registrados en Tabla 17.

Tabla 17. Valores obtenidos con el modo de regresión, linear regression – modo percentage Split

Coefficiente de correlación	0.9759
Error absoluto medio	0.321

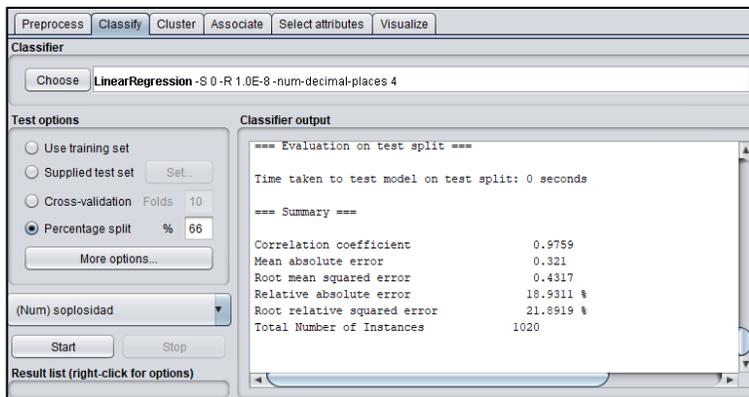


Figura 25. Resultados obtenidos modelo de regresión lineal regression-modo percentage Split

Clasificador simple linear regression - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 26. Y registrados en Tabla 18.

Tabla 18. Valores obtenidos con el modo de regresión, simple linear regression – modo training set

Coeficiente de correlación	0.7674
Error absoluto medio	1.0646

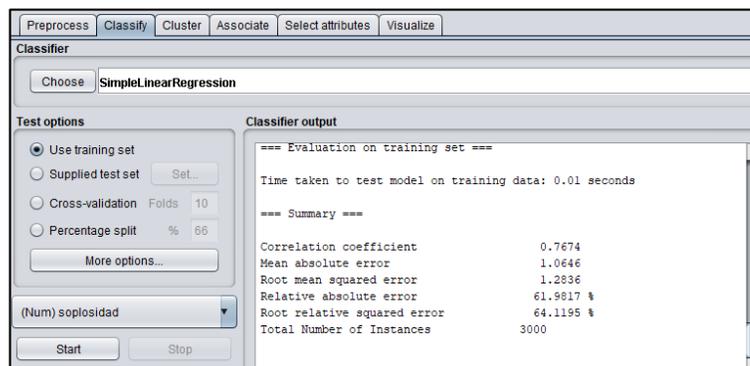


Figura 26. Resultados obtenidos modelo de regresión simple linear regression-modo use training set

Clasificador simple linear regression - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 27. Y registrados en Tabla 19.

Tabla 19. Valores obtenidos con el modo de regresión, simple linear regression crossvalidation

Coefficiente de correlación	0.7668
Error absoluto medio	1.0659

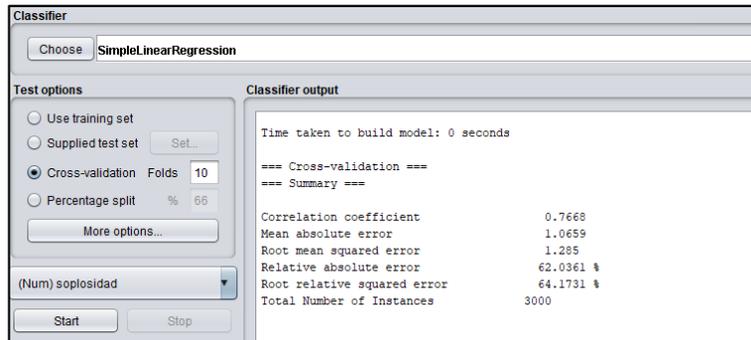


Figura 27. Resultados obtenidos modelo de regresión simple linear regression- modo cross validation

Clasificador simple linear regression - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 28. Y registrados en Tabla 20.

Tabla 20. Valores obtenidos con el modo de regresión, simple linear regression - percentage split

Coefficiente de correlación	0.7625
Error absoluto medio	1.0519

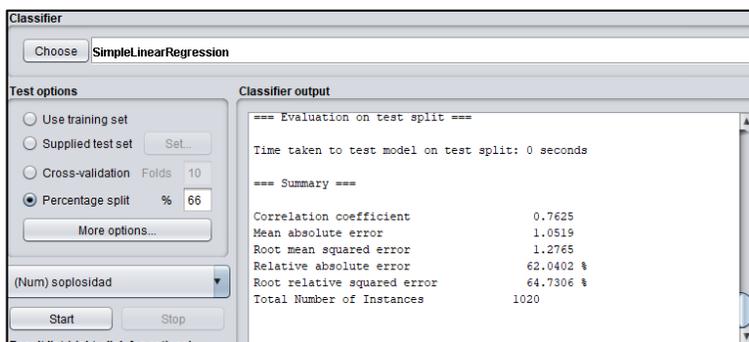


Figura 28. Resultados obtenidos modelo de regresión simple linear regression- modo percentage split

Additive regression – modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 29. Y registrados en Tabla 21.

Tabla 21. Valores obtenidos con el modo de regresión, Additive regression - modo training set

Coefficiente de correlación	0.9215
Error absoluto medio	0.6053

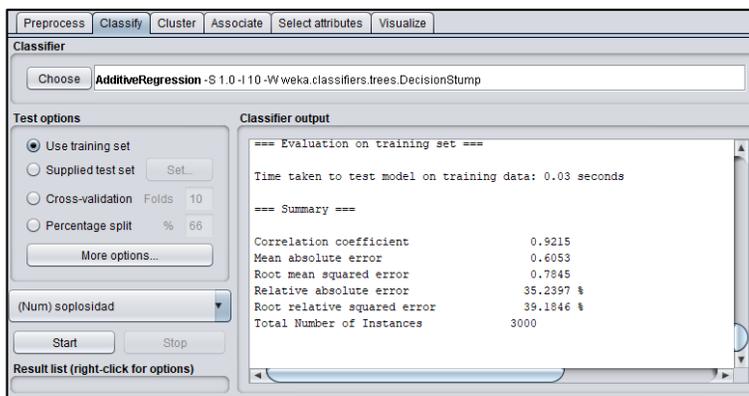


Figura 29. Resultados obtenidos modelo de regresión Additive regression - modo training set

Additive regression - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 30. Y registrados en Tabla 22.

Tabla 22. Valores obtenidos con el modo de regresión, additive regression - modo cross validation

Coefficiente de correlación	0.9153
Error absoluto medio	0.6292

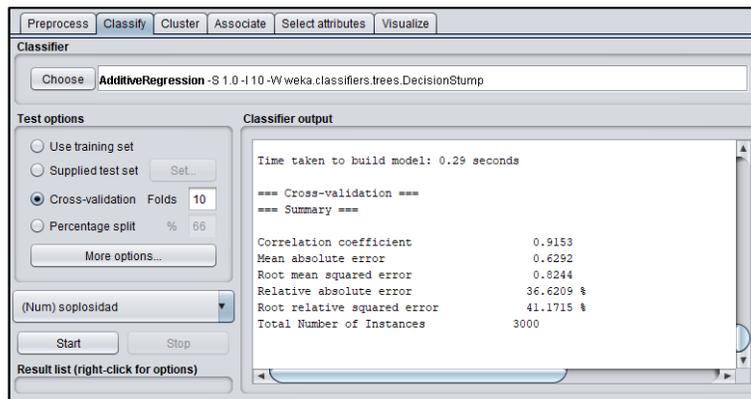


Figura 30. Resultados obtenidos modelo de regresión Additive regression - modo cross validation

Additive regression - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 31. Y registrados en Tabla 23.

Tabla 23. Valores obtenidos con el modo de regresión, additive regression - percentage Split

Coefficiente de correlación	0.9154
Error absoluto medio	0.6377

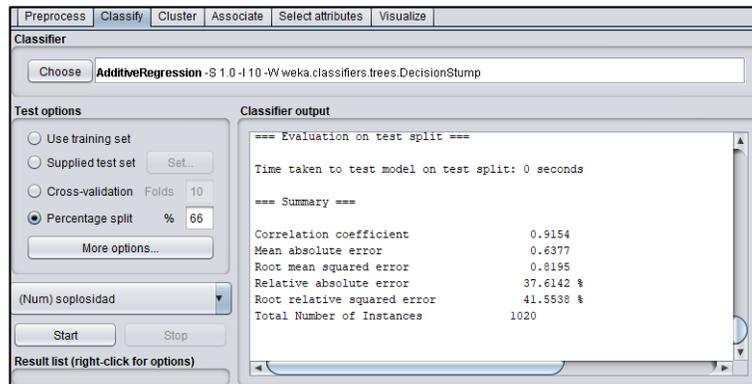


Figura 31. Resultados obtenidos modelo de regresión Additive regression - modo percentage split

Reregressionbydiscretization - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 32. Y registrados en Tabla 24.

Tabla 24. Valores obtenidos con el modo de regresión, regressionbydiscretization - modo training set

Coefficiente de correlación	0.9967
Error absoluto medio	0.0307

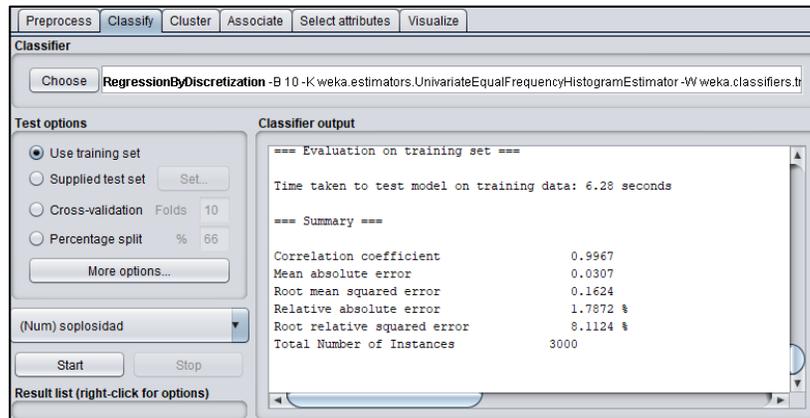


Figura 32. Resultados obtenidos modelo de regresión regressionbydiscretization - modo training set

Reregressionbydiscretization - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la figura 33. Y registrados en tabla 25.

Tabla 25. Valores obtenidos con el modo de regresión, regressionbydiscretization - modo cross-validation

Coefficiente de correlación	0.9779
Error absoluto medio	0.1445

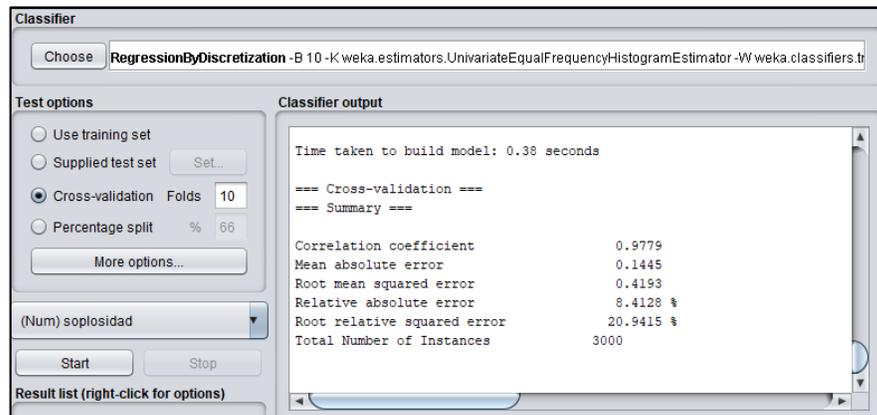


Figura 33. Resultados obtenidos modelo de regresión regressionbydiscretization-cross-validation

Reregressionbydiscretization - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 34. Y registrados en Tabla 26.

Tabla 26. Valores obtenidos con el modo de regresión, regressionbydiscretization - modo percentage split

Coefficiente de correlación	0.9733
Error absoluto medio	0.1743

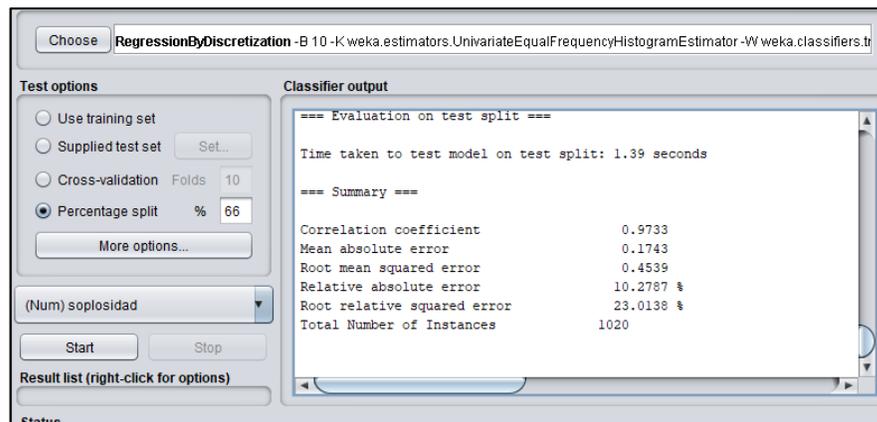


Figura 34. Resultados obtenidos modelo de regresión regressionbydiscretization-percentage split

Gaussian processes- modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 35. Y registrados en Tabla 27.

Tabla 27. Valores obtenidos con el modo de regresión, Gaussian processes - modo training set

Coeficiente de correlación	0.9702
Error absoluto medio	0.3465

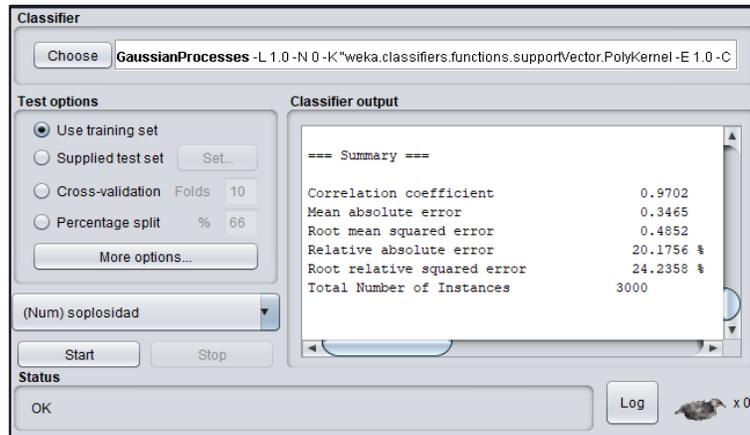


Figura 35. Resultados obtenidos modelo de regresión Gaussian processes - modo training set

Gaussian processes- cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 36. Y registrados en Tabla 28.

Tabla 28. Valores obtenidos con el modo de regresión, Gaussian processes - modo cross-validation

Coeficiente de correlación	0.9692
Error absoluto medio	0.3527

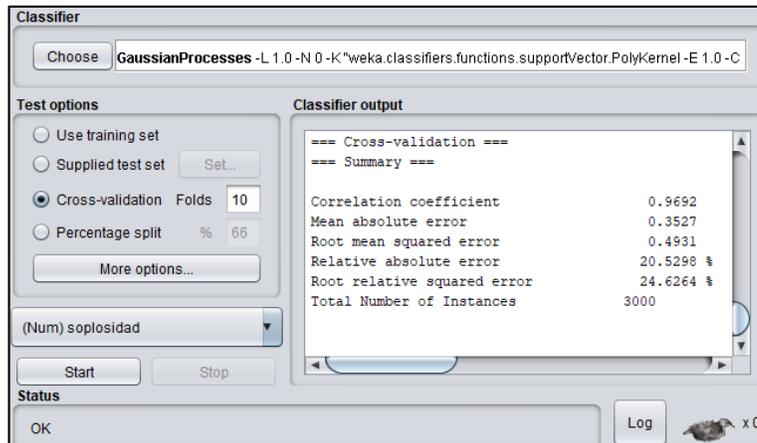


Figura 36. Resultados obtenidos modelo de regresión Gaussian processes - modo cross-validation

Gaussian processes- percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 37. Y registrados en Tabla 29.

Tabla 29. Valores obtenidos con el modo de regresión, Gaussian processes - modo percentage split

Coefficiente de correlación	0.972
Error absoluto medio	0.345

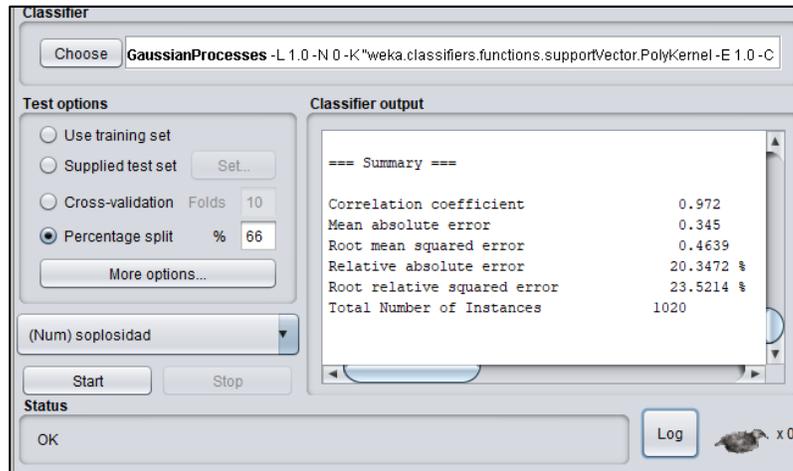


Figura 37. Resultados obtenidos modelo de regresión Gaussian processes - percentage split

M5P- modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 38. Y registrados en Tabla 30.

Tabla 30. Valores obtenidos con el modo de regresión, M5P - modo training set

Coefficiente de correlación	0.9906
Error absoluto medio	0.1429

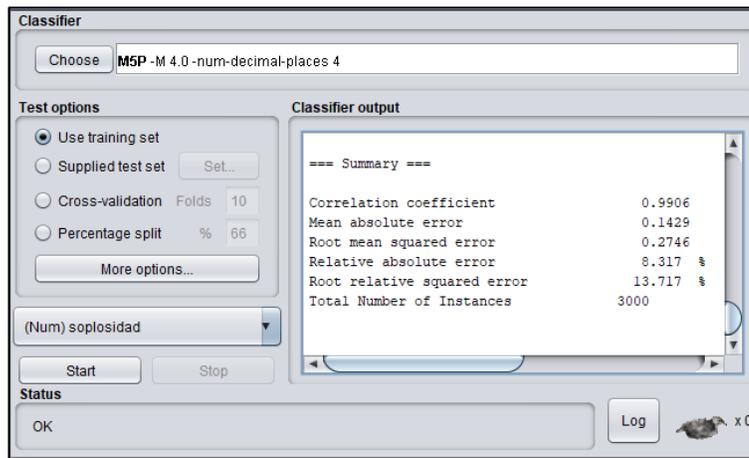


Figura 38. Resultados obtenidos modelo de regresión M5P - modo training set

M5P- cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 39. Y registrados en Tabla 31.

Tabla 31. Valores obtenidos con el modo de regresión, M5P - modo cross-validation

Coefficiente de correlación	0.9847
Error absoluto medio	0.1958

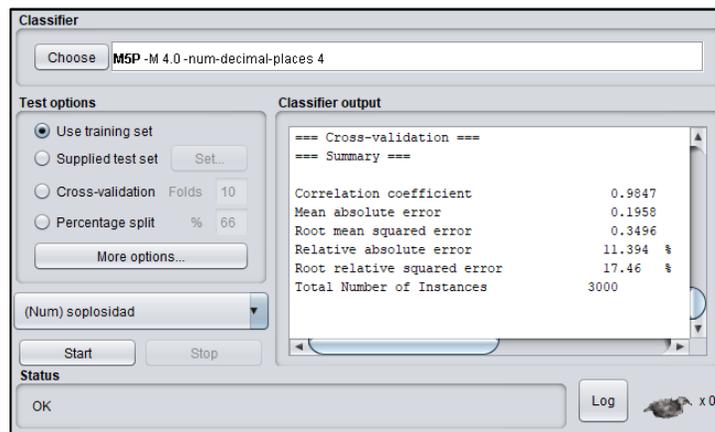


Figura 39. Resultados obtenidos modelo de regresión M5P - modo cross validation

M5P- percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 40. Y registrados en Tabla 32.

Tabla 32. Valores obtenidos con el modo de regresión M5P - modo percentage split

Coefficiente de correlación	0.9868
Error absoluto medio	0.2043

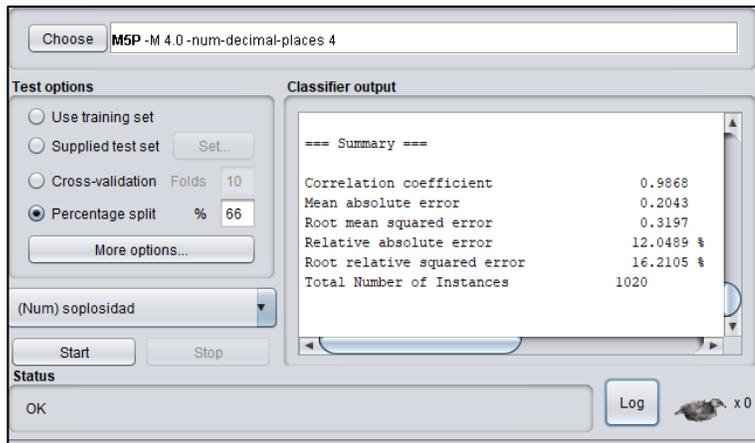


Figura 40. Resultados obtenidos modelo de regresión M5P - modo percentage split

- **Prueba con los atributos seleccionados anteriormente.**

Clasificador linear regression - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 41. Y registrados en Tabla 33.

Tabla 33. Valores obtenidos con el modo de regresión, clasificador linear regression - modo training set

Coefficiente de correlación	0.8969
Error absoluto medio	0.6787

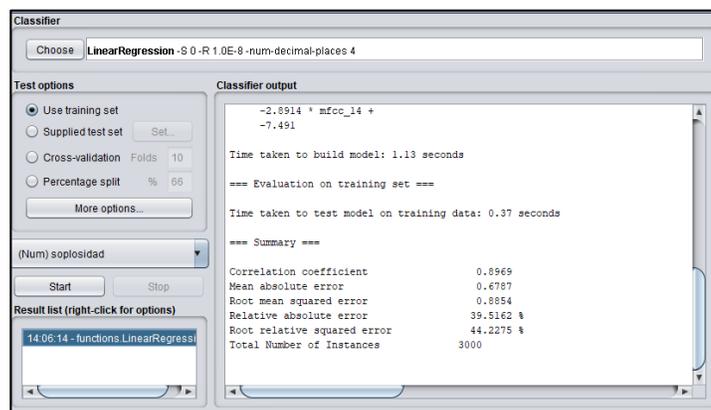


Figura 41. Resultados obtenidos modelo linear regression - modo training set

Clasificador linear regression – cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 42. Y registrados en Tabla 34.

Tabla 34. Valores obtenidos con el modo de regresión, clasificador linear regression - modo cross-validation

Coefficiente de correlación	0.8962
Error absoluto medio	0.6805

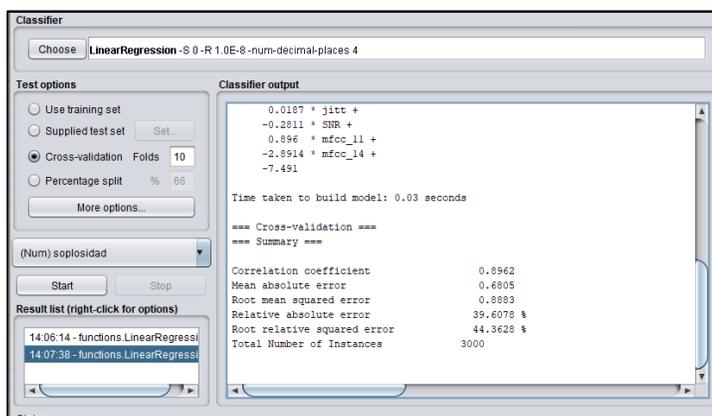


Figura 42. Resultados obtenidos modelo linear regression - modo cross-validation

Clasificador linear regression – percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 43. Y registrados en Tabla 35.

Tabla 35. Valores obtenidos con el modo de regresión, clasificador linear regression- percentage split

Coefficiente de correlación	0.89
Error absoluto medio	0.6896

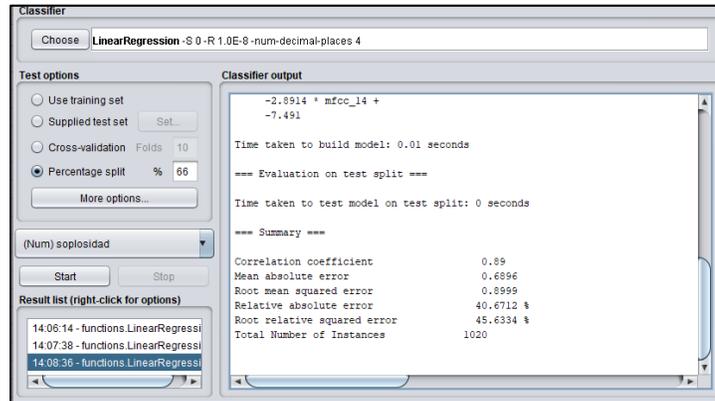


Figura 43. Resultados obtenidos modelo linear regression - modo percentage split

Clasificador simple linear regression - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 44. Y registrados en Tabla 36.

Tabla 36. Valores obtenidos con el modo de regresión, clasificador simple linear regression- modo training set

Coefficiente de correlación	0.7674
Error absoluto medio	1.0646

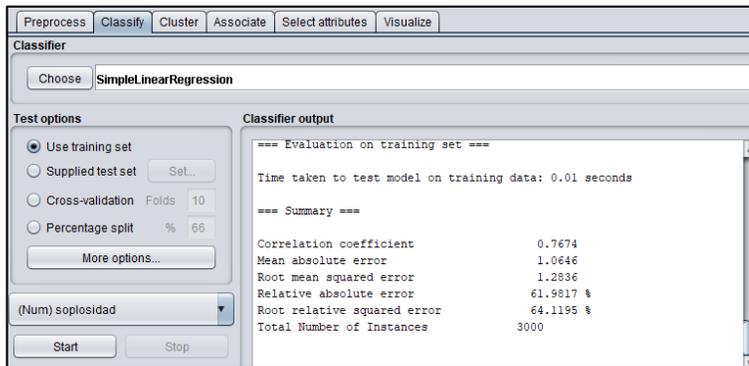


Figura 44. Resultados obtenidos modelo simple linear regression- modo training set

Clasificador simple linear regression - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 45. Y registrados en Tabla 37.

Tabla 37. Valores obtenidos con el modo de regresión, clasificador simple linear regression- modo cross validation

Coeficiente de correlación	0.7674
Error absoluto medio	1.0646

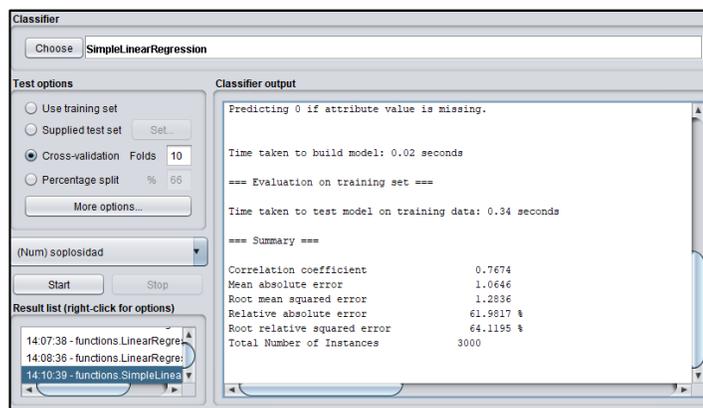


Figura 45. Resultados obtenidos modelo simple linear regression- modo cross validation

Clasificador simple linear regression - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 46. Y registrados en Tabla 38.

Tabla 38. Valores obtenidos con el modo de regresión, clasificador simple linear regression- modo percentage split

Coeficiente de correlación	0.7625
Error absoluto medio	1.0519

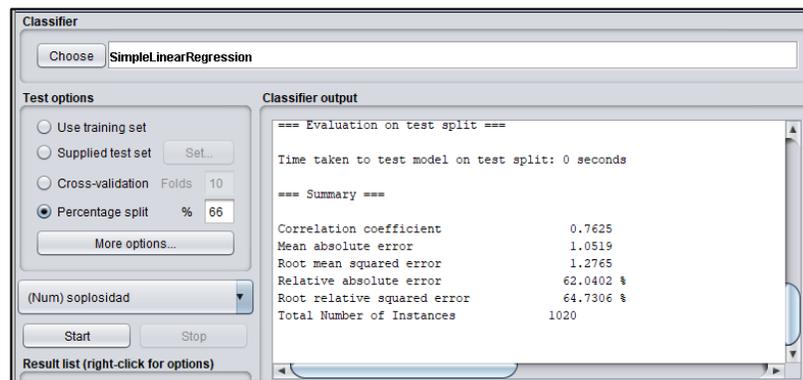


Figura 46. Resultados obtenidos modelo simple linear regression- modo percentage split

Additive regression - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 47. Y registrados en Tabla 39.

Tabla 39. Valores obtenidos con el modo de regresión, Additive regression- modo training set

Coeficiente de correlación	0.9176
Error absoluto medio	0.6058

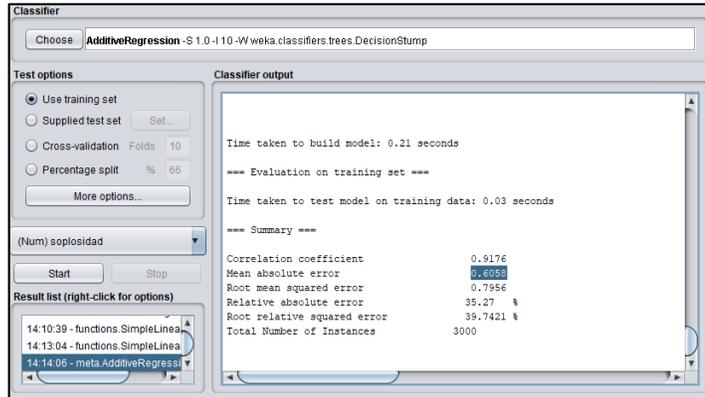


Figura 47. Resultados obtenidos modelo Additive regression- modo training set

Additive regression - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 48. Y registrados en Tabla 40.

Tabla 40. Valores obtenidos con el modo de regresión, Additive regression- modo cross validation

Coeficiente de correlación	0.9103
Error absoluto medio	0.6471

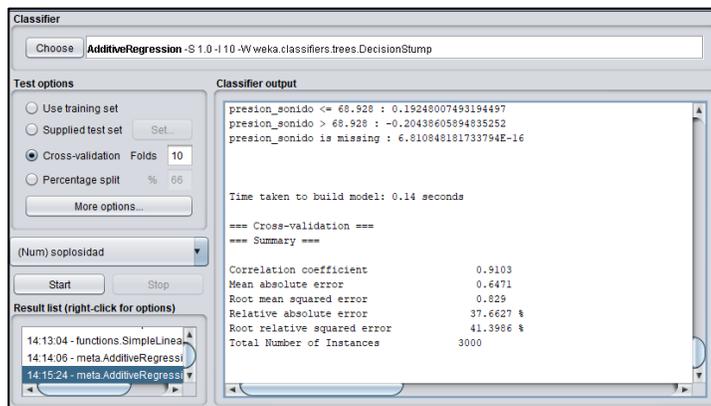


Figura 48. Resultados obtenidos modelo Additive regression- modo cross validation

Additive regression - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 49. Y registrados en Tabla 41.

Tabla 41. Valores obtenidos con el modo de regresión, Additive regression- modo percentage split

Coeficiente de correlación	0.9097
Error absoluto medio	0.6382

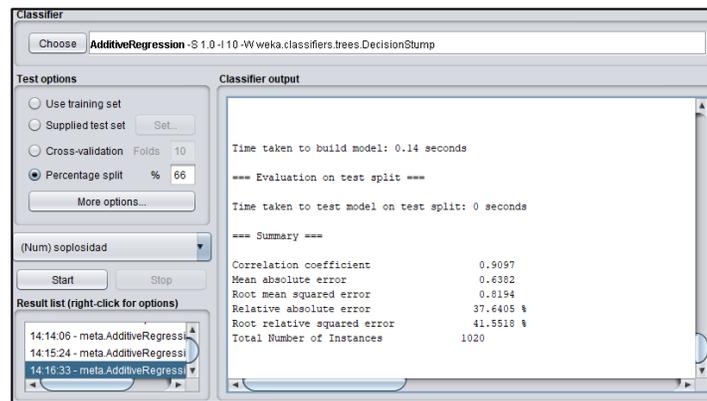


Figura 49. Resultados obtenidos modelo Additive regression- modo percentage split

Regressionbydiscretization - modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 50. Y registrados en Tabla 42.

Tabla 42. Valores obtenidos con el modo de regresión, regressionbydiscretization - modo training set.

Coeficiente de correlación	0.9876
Error absoluto medio	0.1085

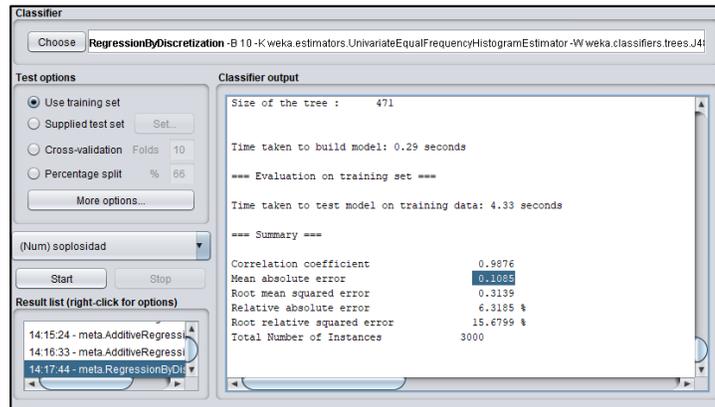


Figura 50. Resultados obtenidos modelo regressionbydiscretization - modo training set

Regressionbydiscretization - cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 51. Y registrados en Tabla 43.

Tabla 43. Valores obtenidos con el modo de regresión, Regressionbydiscretization - cross-validation

Coefficiente de correlación	0.9607
Error absoluto medio	0.2873

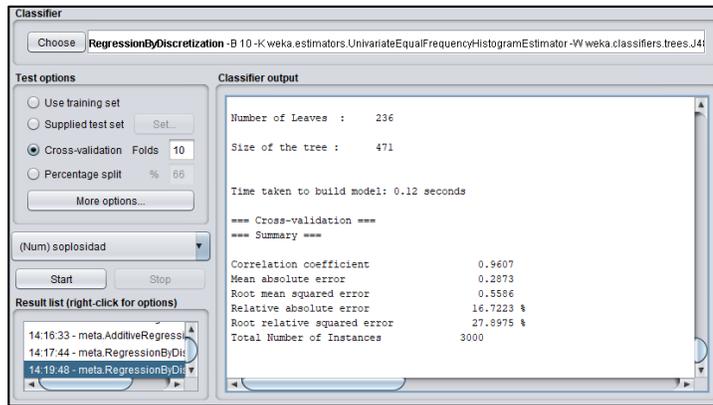


Figura 51. Resultados obtenidos modelo Regressionbydiscretization - cross-validation

Regressionbydiscretization - percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 52. Y registrados en Tabla 44.

Tabla 44. Valores obtenidos con el modo de regresión, Regressionbydiscretization-percentage split

Coeficiente de correlación	0.9635
Error absoluto medio	0.287

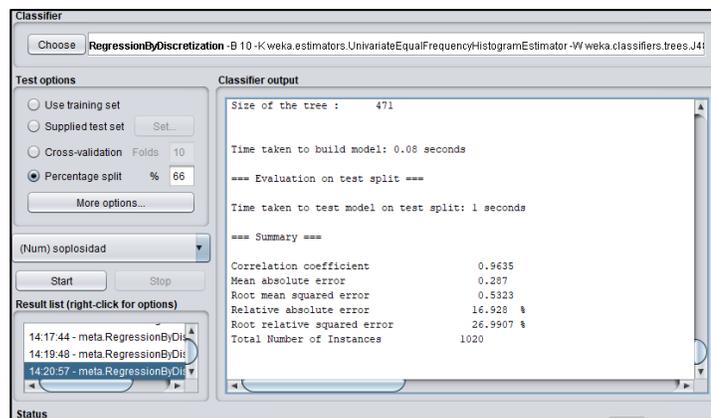


Figura 52. Resultados obtenidos modelo Regressionbydiscretization-percentage split

Gaussian processes- modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 53. Y registrados en Tabla 45.

Tabla 45. Valores obtenidos con el modo de regresión, Gaussian processes- modo training set

Coeficiente de correlación	0.8871
Error absoluto medio	0.7158

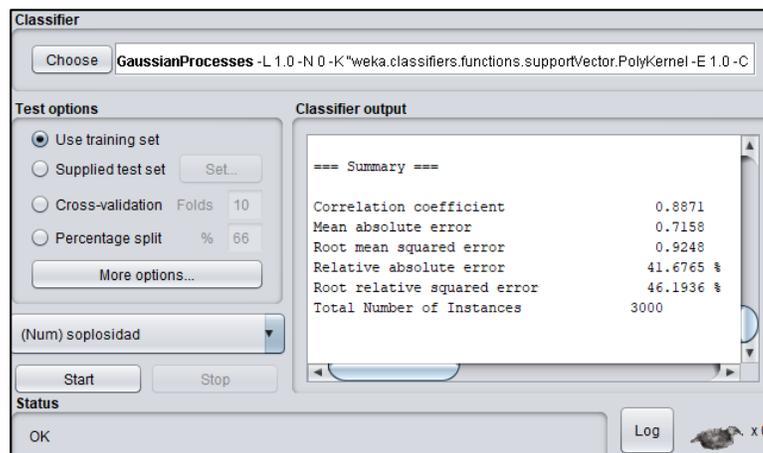


Figura 53. Resultados obtenidos modelo Gaussian processes- modo training set

Gaussian processes- cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 54. Y registrados en Tabla 46.

Tabla 46. Valores obtenidos con el modo de regresión, gaussian processes- cross-validation

Coeficiente de correlación	0.8858
Error absoluto medio	0.7195

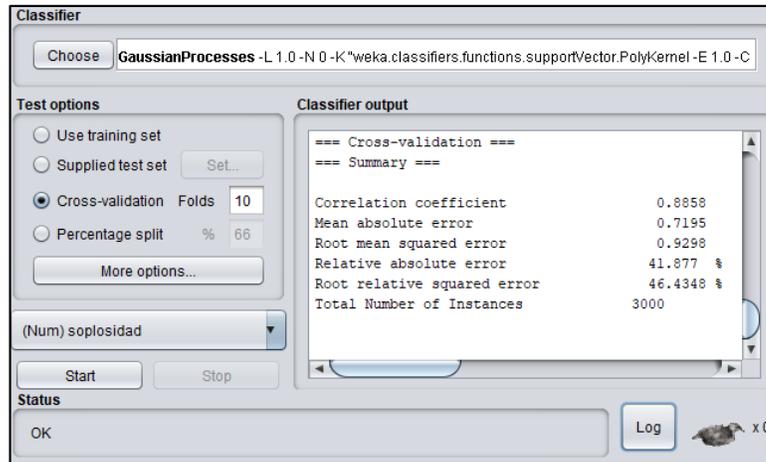


Figura 54. Resultados obtenidos modelo gaussian processes- cross-validation

Gaussian processes- percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 55. Y registrados en Tabla 47.

Tabla 47. Valores obtenidos con el modo de regresión, gaussian processes - percentage split

Coeficiente de correlación	0.8761
Error absoluto medio	0.7394

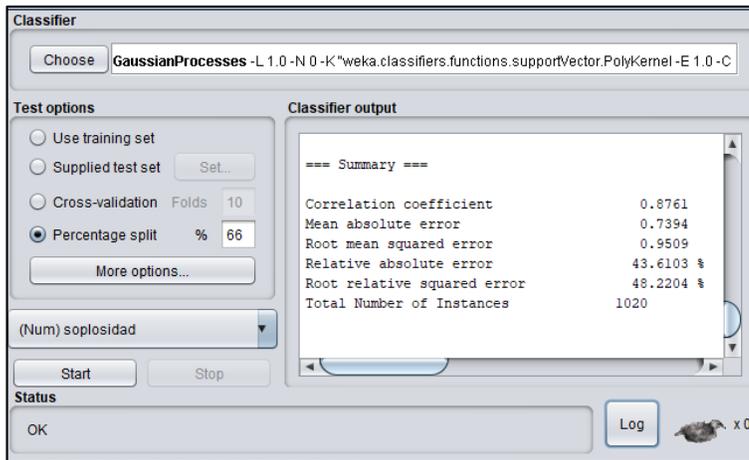


Figura 55. Resultados obtenidos modelo gaussian processes - percentage Split

M5P- modo training set

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 56. Y registrados en Tabla 48.

Tabla 48. Valores obtenidos con el modo de regresión, M5P- modo training set

Coeficiente de correlación	0.6221
Error absoluto medio	1.7352

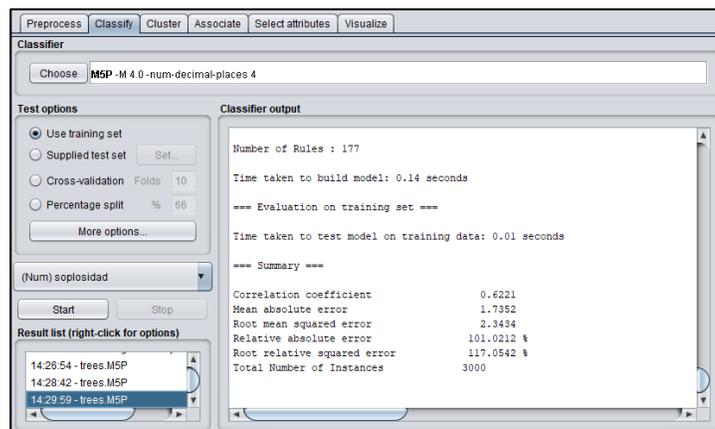


Figura 56. Resultados obtenidos modelo M5P- modo training set

M5P- cross-validation

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 57. Y registrados en Tabla 49.

Tabla 49. Valores obtenidos con el modo de regresión, M5P- modo cross-validation

Coefficiente de correlación	0.4106
Error absoluto medio	2.193

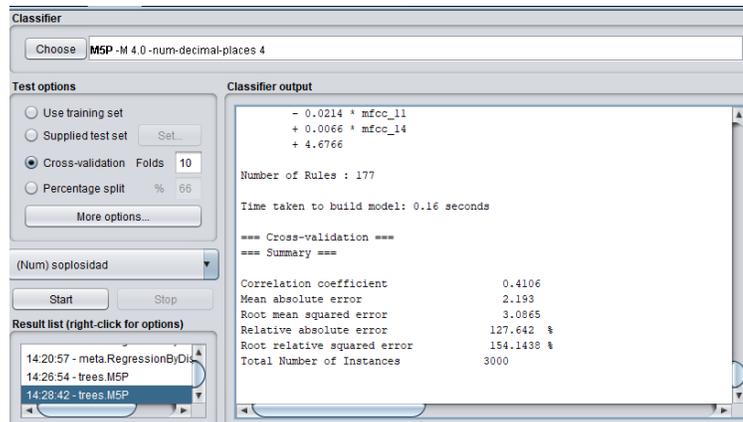


Figura 57. Resultados obtenidos modelo M5P- modo cross-validation

M5P- percentage Split

Con este clasificador y modo de prueba, se obtuvieron los resultados visualizados en la Figura 58. Y registrados en Tabla 50.

Tabla 50. Valores obtenidos con el modo de regresión, M5P- modo percentage split

Coefficiente de correlación	0.7232
Error absoluto medio	2.2455

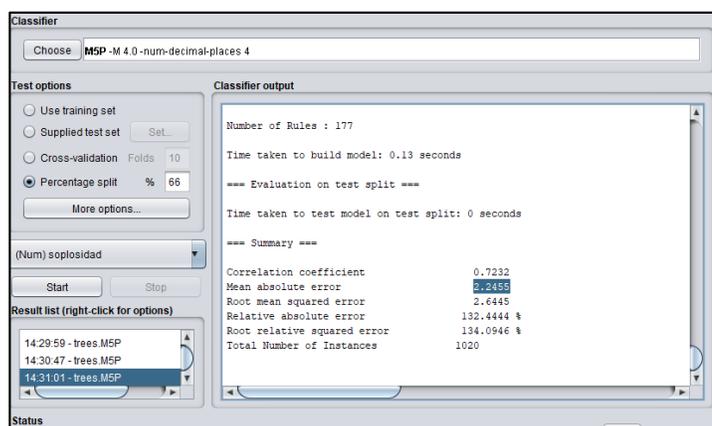


Figura 58. Resultados obtenidos modelo M5P- modo percentage split

Teniendo en cuenta los resultados obtenidos con las pruebas realizadas los cuales se pueden observar en la Tabla 51, el clasificador con la mejor métrica (mayor coeficiente de correlación y menor error), junto con el modo de prueba, es regression by discretization-modo training set, sin embargo el modo training set puede causar sobre entrenamiento (overfit), por lo tanto el modo seleccionado es cross validation, también se pudo concluir en que se obtiene un mejor resultado en la regresión utilizando todos los 42 parámetros.

Tabla 51. Valores obtenidos en pruebas realizadas con diferentes modelos de regresión

Clasificador	Modo de prueba	Coeficiente de correlación	Error absoluto medio
Classificador linear regression	Training set	0.9732	0.3243
Classificador linear regression	Cross validation	0.9722	0.3294
Classificador linear regression	Percentage Split	0.9759	0.321
Simple linear regression	Training set	0.7674	1.0646

Simple linear regression	Cross validation	0.7668	1.0659
Simple linear regression	Percentage Split	0.7625	1.0519
Additive regression	Training set	0.9215	0.6053
Additive regression	Cross validation	0.9153	0.6292
Additive regression	Percentage Split	0.9154	0.6377
Regressionbydiscretization	Training set	0.9967	0.0307
Regressionbydiscretization	Cross validation	0.9779	0.1445
Regressionbydiscretization	Percentage Split	0.9733	0.1743
Gaussian processes	Training set	0.9702	0.3465
Gaussian processes	Cross validation	0.9692	0.3527
Gaussian processes	Percentage Split	0.972	0.345
Mp5	Training set	0.9906	0.1429
Mp5	Cross validation	0.9847	0.1958
Mp5	Percentage Split	0.9868	0.2043
Prueba de clasificación con los parámetros seleccionados			
Classificador linear regression	Training set	0.8969	0.6787
Classificador linear regression	Cross validation	0.8962	0.6805
Classificador linear regression	Percentage Split	0.89	0.6896
Simple linear regression	Training set	0.7674	1.0646
Simple linear regression	Cross validation	0.7674	1.0646

Simple linear regression	Percentage Split	0.7625	1.0519
Additive regression	Training set	0.9176	0.6058
Additive regression	Cross validation	0.9103	0.6471
Additive regression	Percentage Split	0.9097	0.6382
Regressionbydiscretization	Training set	0.9876	0.1085
Regressionbydiscretization	Cross validation	0.9607	0.2873
Regressionbydiscretization	Percentage Split	0.9635	0.287
Gaussian processes	Training set	0.8871	0.7158
Gaussian processes	Cross validation	0.8858	0.7195
Gaussian processes	Percentage Split	0.8761	0.7394
Mp5	Training set	0.6221	1.7352
Mp5	Cross validation	0.4106	2.193
Mp5	Percentage Split	0.7232	2.2455

6.2 MODELO DE REGRESIÓN MEDIANTE MATLAB

Matlab proporciona una serie de opciones y herramientas diferentes de machine learning y Deep learning, dentro de estas, Regression Learner, de la cual haremos uso, esta hace parte del toolbox Statistics and Machine Learning de MATLAB.

Machine learning de la caja de herramientas de Matlab: Mediante esta herramienta se pueden utilizar diferentes modelos para analisis de datos, proporcionando diferentes graficas que permiten el analisis exploratorio de estos, se pueden obtener modelos predictivos de bajo error. Matlab selecciona el mejor

modelo basándose en el error más bajo destacándolo de los demás modelos entrenados.

Matlab proporciona herramientas para seleccionar las características más importantes dentro de un conjunto de datos, es decir los parámetros que influyan en una mejor estimación, y poder aplicar una reducción de dimensionalidad, dentro de estas herramientas está el PCA en el que se hace reducción de dimensionalidad por análisis de componentes principales. Este toolbox me permite hacer uso de diferentes modelos de regresión muy utilizados en machine learning.

Regression learner: Entrena modelos de regresión para predecir datos. Con esta aplicación, puede explorar sus datos, seleccionar características, especificar esquemas de validación, entrenar modelos y evaluar resultados. Puede realizar una capacitación automatizada para buscar el mejor tipo de modelo de regresión, incluidos los modelos de regresión lineal, los árboles de regresión, los modelos de regresión de procesos gaussianos, las máquinas de vectores de soporte y los conjuntos de árboles de regresión.

Se probaron todos los modelos que ofrece la herramienta de regresión, con el fin de establecer cuál es el más óptimo al presentar el menor error, los modelos son los que se observan en la Figura 59.

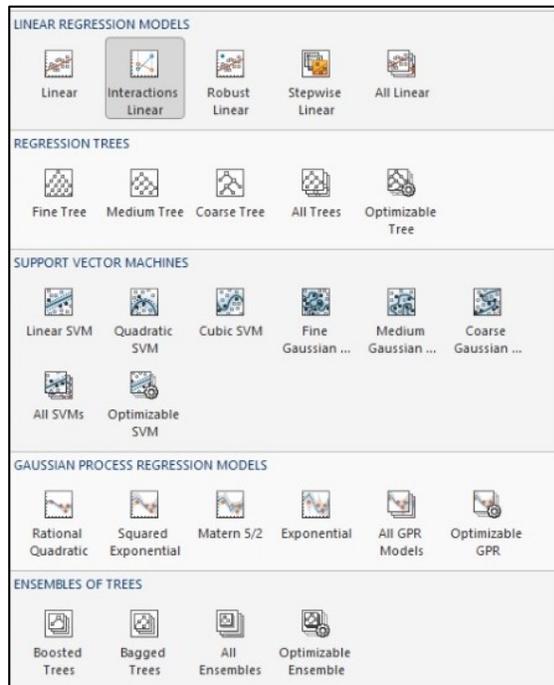


Figura 59. Modelos de regresión que proporciona Matlab

6.2.1 Prueba con modelos de regresión de MATLAB

- **Modelos de regresión lineal**

Linear: Un modelo de regresión lineal con solo intersección y términos lineales.

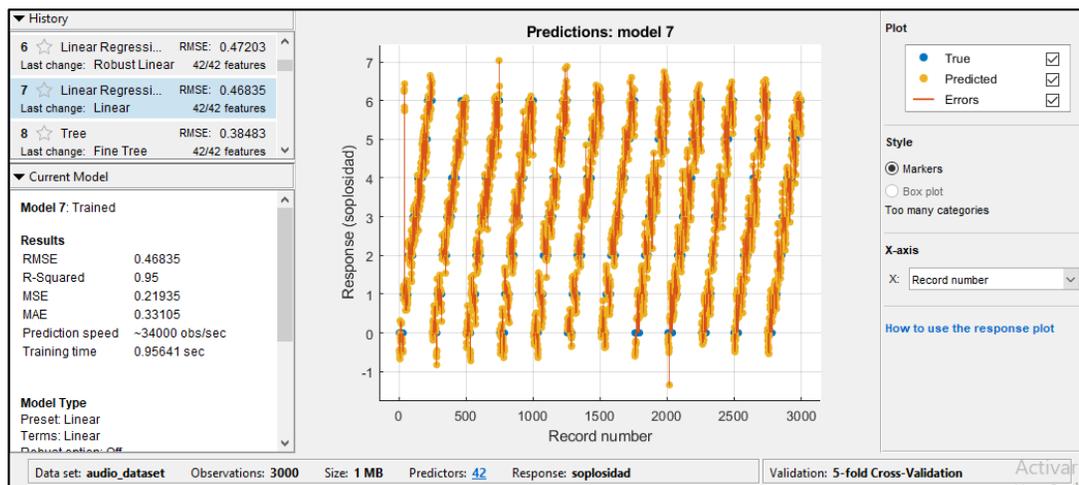


Figura 60. Gráfico de resultado obtenido con el modelo de regresión lineal en Matlab

El error cuadrático medio (RMSE) que se obtiene con este modelo, es de 0.46835.

En el gráfico ubicado en la Figura 60, se ubica en el eje x, los 3000 audios, y en el eje y, los valores de soplosidad, el cuadro superior derecho ubicado en la imagen, permite ver el valor verdadero representado con el color azul, el valor de predicción que arroja el software, representado en color amarillo y el valor de error que presenta la regresión, representado en color rojo.

Interactions linear: un modelo de regresión lineal con términos de intercepción, lineal e interacción.

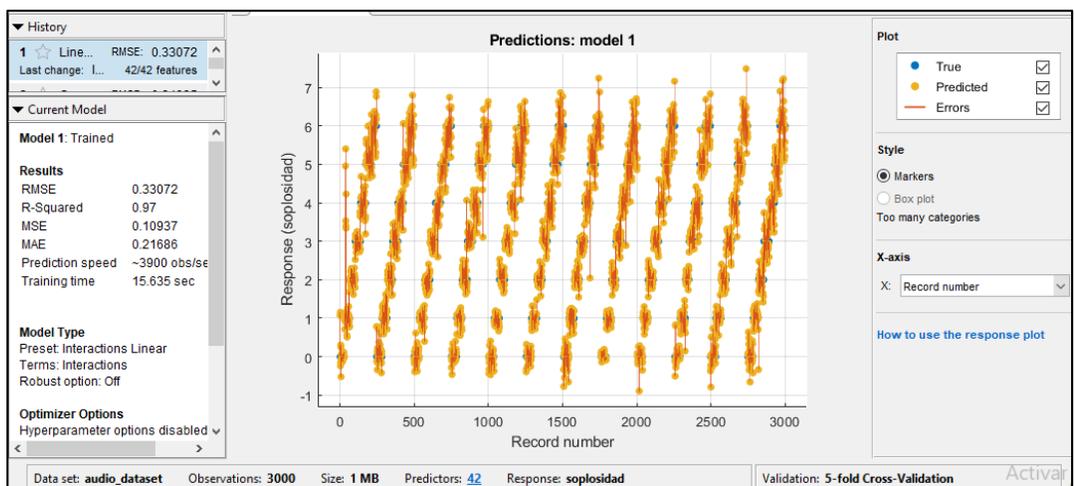


Figura 61. Gráfico de resultado obtenido con el modelo de interactions linear en Matlab

El resultado obtenido con este modelo es un error cuadrático medio (RMSE) de 0.33072, este se puede visualizar en la Figura 61.

Robust linear: Un modelo de regresión lineal robusto con solo intersección y términos lineales.

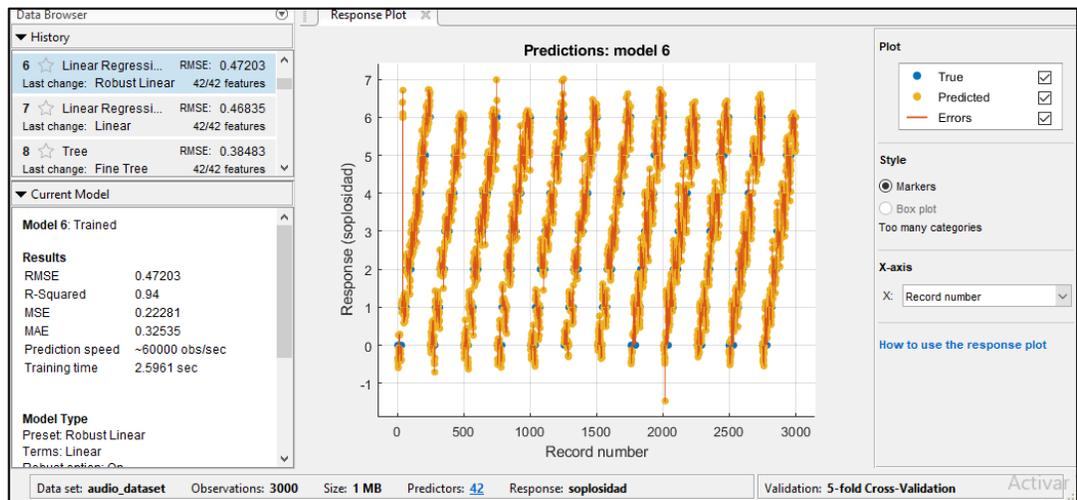


Figura 62. Gráfico de resultado obtenido con el modelo de robust linear en Matlab

El resultado obtenido con este modelo es un error cuadrático medio (RMSE) de 0.47203, este se puede visualizar en la Figura 62.

- **Modelos de regresión de árboles**

Fine tree: árbol de regresión, su tamaño mínimo de hoja es 4.

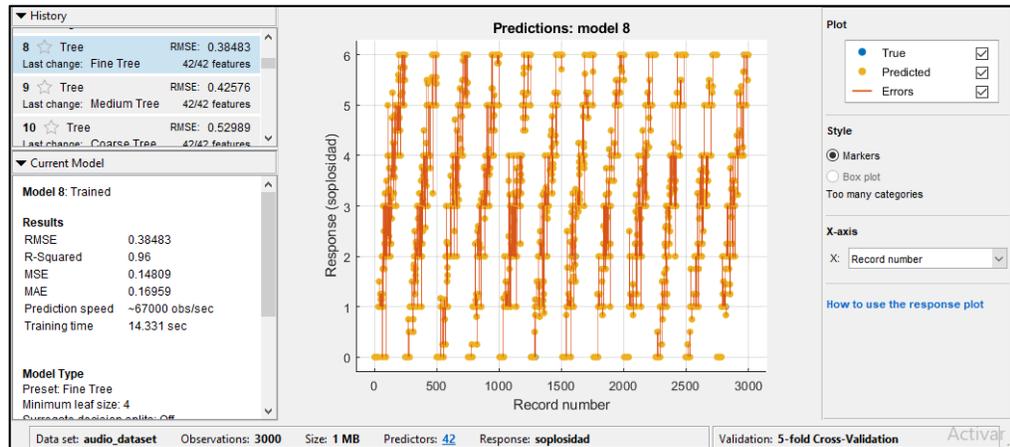


Figura 63. Gráfico de resultado obtenido con el modelo fine tree en Matlab

El Resultado obtenido con este modelo es un error cuadrático medio (RMSE) de 0.38483, este se puede visualizar en la Figura 63.

Medium tree: un árbol de regresión medio, el tamaño mínimo de la hoja es 12

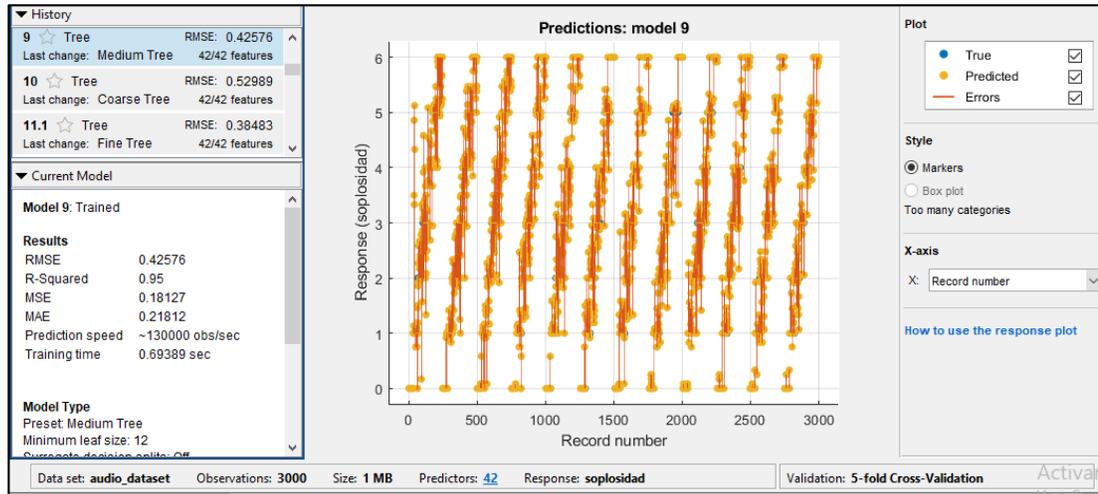


Figura 64. Gráfico de resultado obtenido con el modelo de medium tree en Matlab

El resultado obtenido con este modelo es un error cuadrático medio (RMSE) de 0.42576, este se puede visualizar en la Figura 64.

Coarse tree: un árbol de regresión grueso, el tamaño mínimo de la hoja es 36

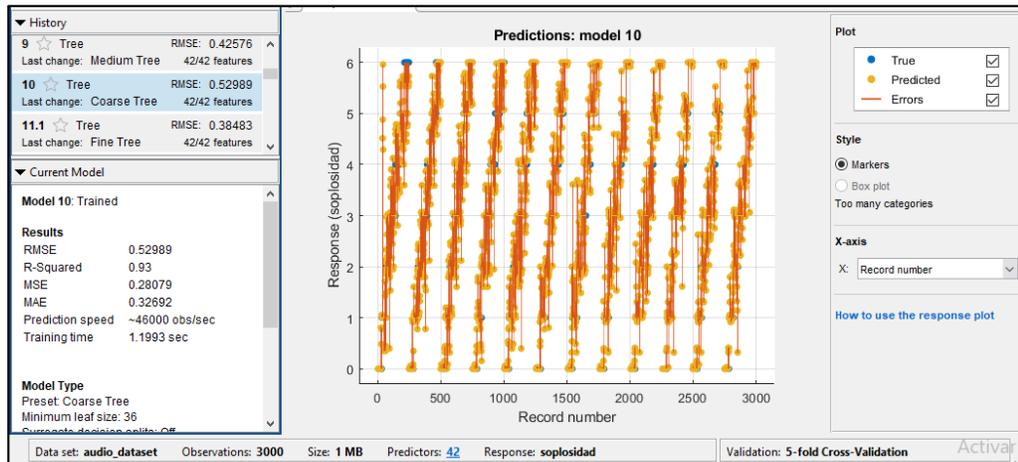


Figura 65 Gráfico de resultado obtenido con el modelo coarse tree en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.52989, este se puede visualizar en la Figura 65.

- **Soporte de máquinas de vectores**

Linear svm: Una máquina de vectores de soporte que sigue una estructura lineal simple en los datos, utilizando el núcleo lineal. El SVM más fácil de interpretar.

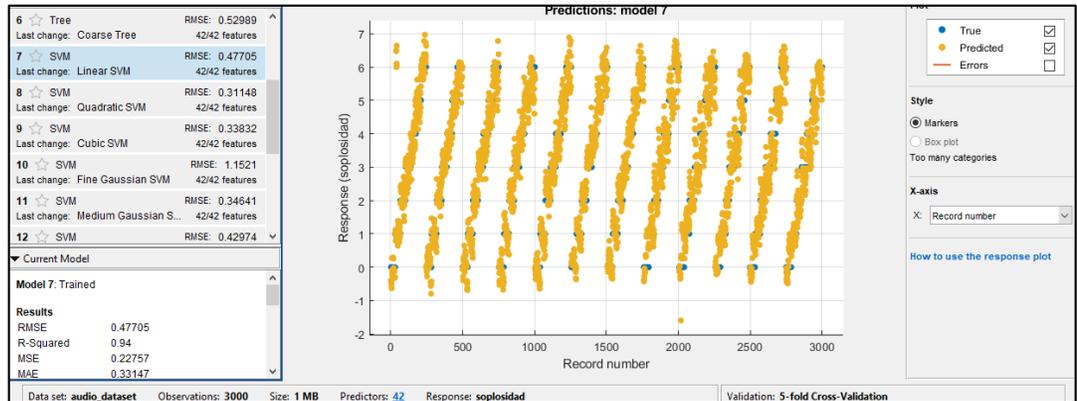


Figura 66. Gráfico de resultado obtenido con el modelo de linear svm en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.47705, este se puede visualizar en la Figura 66.

Quadratic svm: una máquina de vectores de soporte que usa el núcleo cuadrático

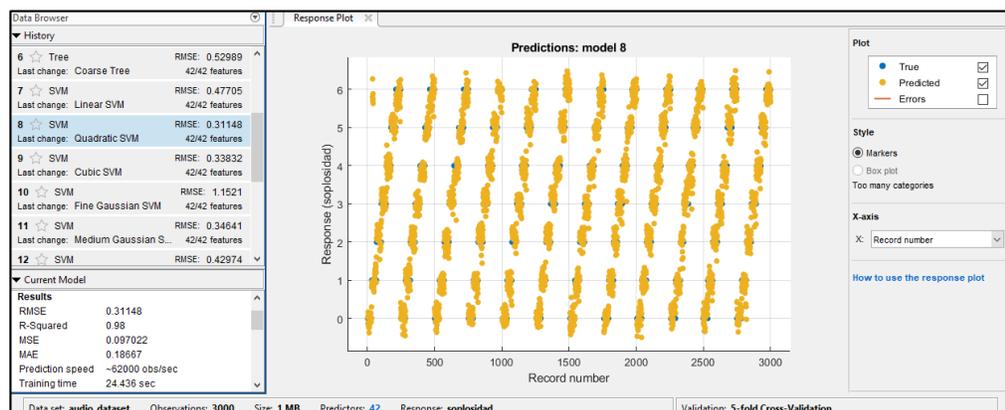


Figura 67. Gráfico de resultado obtenido con el modelo de quadratic svm en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.31148, este se puede visualizar en la Figura 67.

Cubic svm: una máquina de vectores de soporte que usa el núcleo cúbico

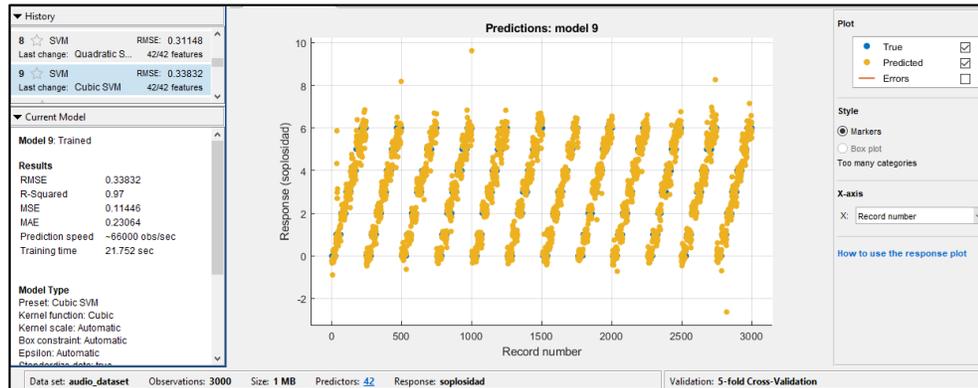


Figura 68. Gráfico de resultado obtenido con el modelo de cubic svm en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.33832, este se puede visualizar en la Figura 68.

Fine gaussian: Una máquina de vectores de soporte que sigue una estructura finamente detallada en los datos. Utiliza el núcleo gaussiano con escala de núcleo $\sqrt{p} / 4$, con p el número de predictores

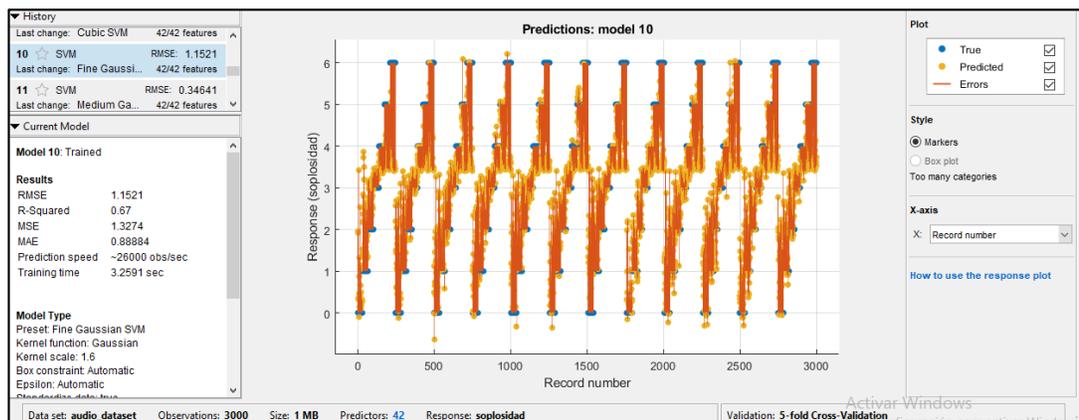


Figura 69. Gráfico de resultado obtenido con el modelo de fine gaussian en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 1.1521, este se puede visualizar en la Figura 69.

Medium gaussian: una máquina de vectores de soporte que encuentra una estructura menos fina en los datos. Utiliza el núcleo gaussiano con escala de núcleo \sqrt{p} , con p el número de predictores.

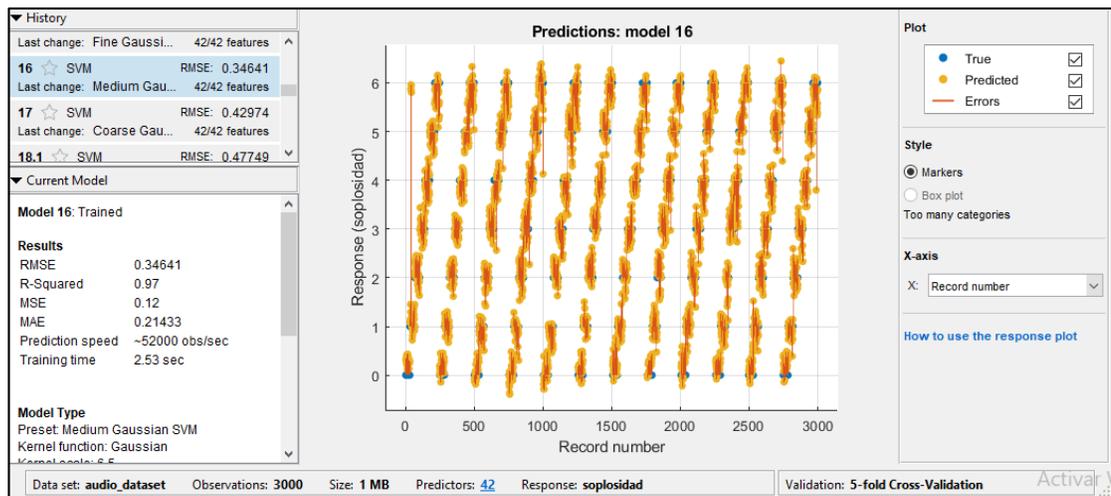


Figura 70. Gráfico de resultado obtenido con el modelo de Medium gaussian en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.34641, este se puede visualizar en la Figura 70.

Coarse gaussian: una máquina de vectores de soporte que sigue una estructura gruesa en los datos. Utiliza el núcleo gaussiano con la escala del núcleo $\sqrt{p} * 4$, con p el número de predictores.

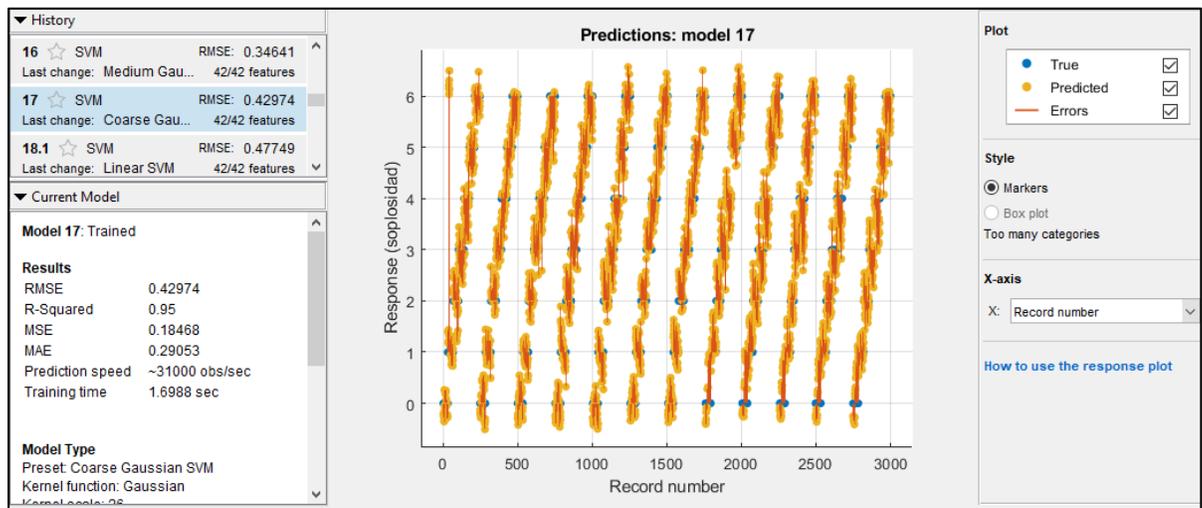


Figura 71. Gráfico de resultado obtenido con el modelo coarse gaussian en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.42974, este se puede visualizar en la Figura 71.

- **Modelos de regresión de procesos gaussianos**

Rational quadratic: Un modelo de proceso gaussiano que usa el núcleo cuadrático racional.

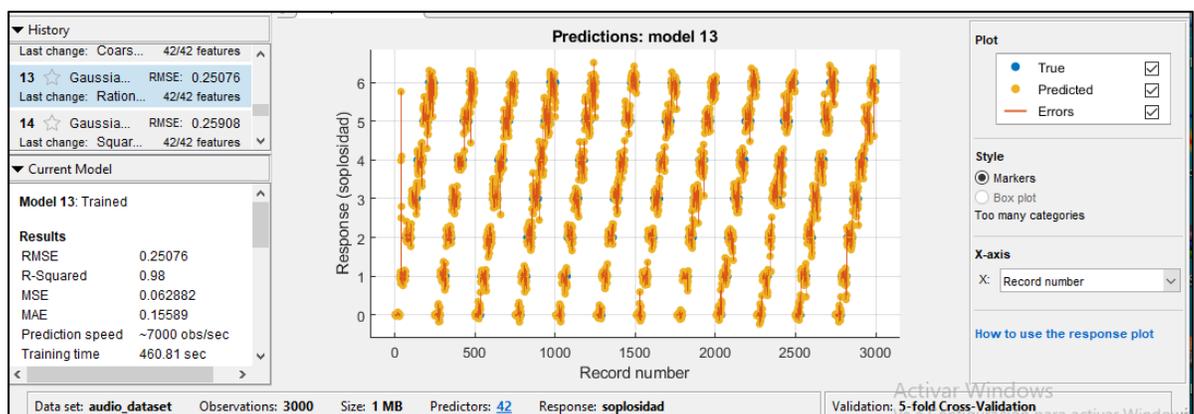


Figura 72. Gráfico de resultado obtenido con el modelo rational quadratic en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.25076, este se puede visualizar en la Figura 72.

Squared exponential: Un modelo de proceso gaussiano que utiliza el núcleo exponencial cuadrado

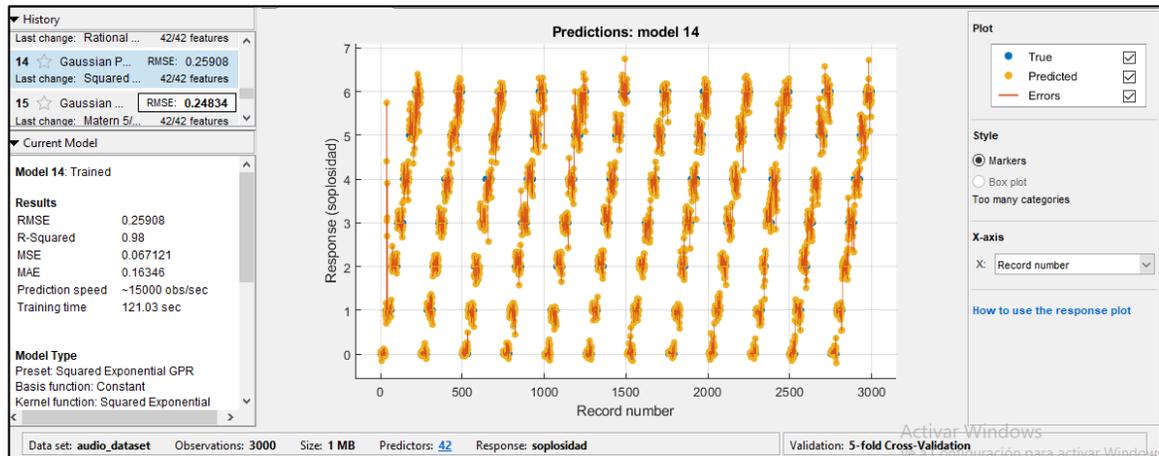


Figura 73 Gráfico de resultado obtenido con el modelo de squared exponential en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.25908, este se puede visualizar en la Figura 73.

Matern 5/2: Un modelo de proceso gaussiano que utiliza el núcleo matern 5/2

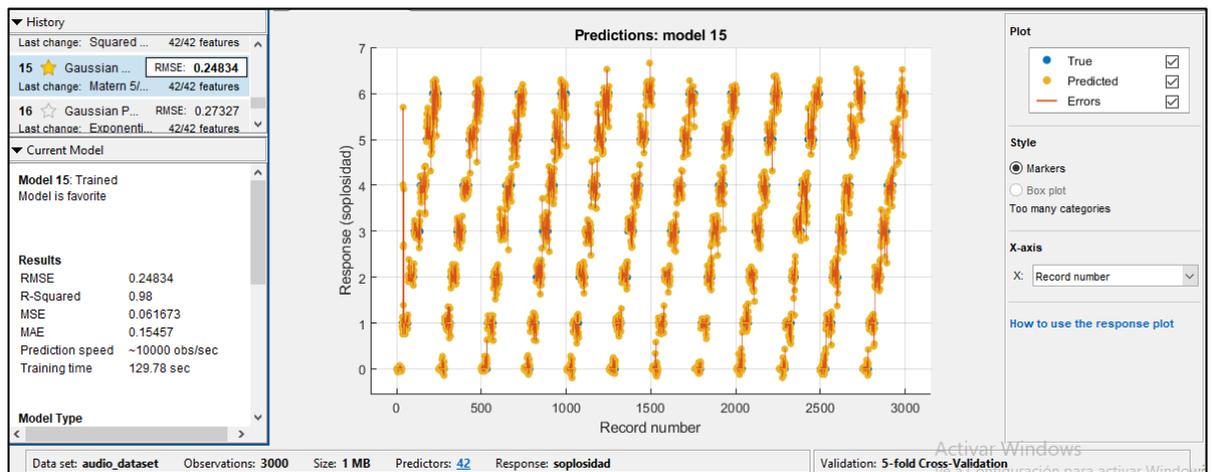


Figura 74. Gráfico de resultado obtenido con el modelo de Matern 5/2 en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.24834, este se puede visualizar en la Figura 74.

Exponential: Un modelo de proceso gaussiano que usa el núcleo exponencial

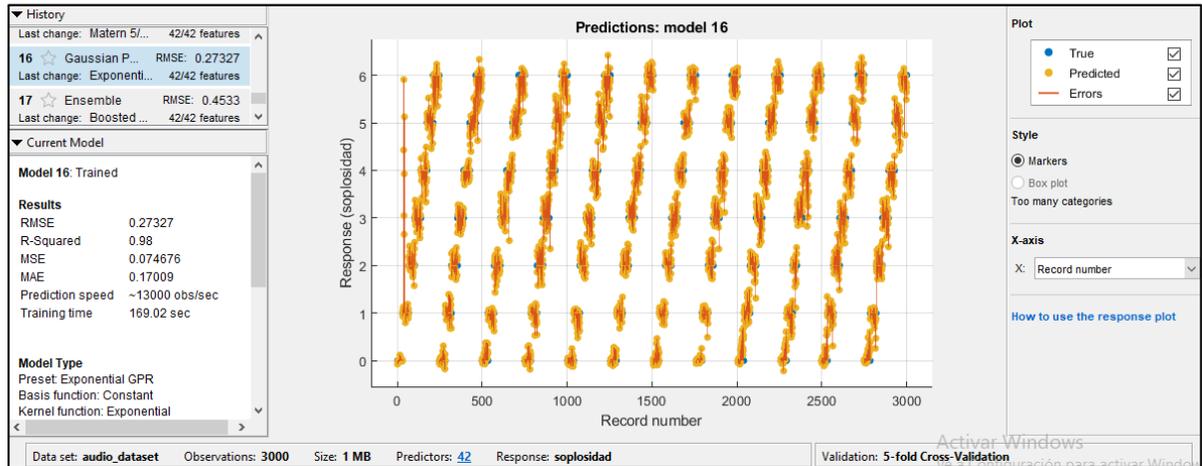


Figura 75. Gráfico de resultado obtenido con el modelo exponencial en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.27327, este se puede visualizar en la Figura 75.

▪ Conjunto de arboles

boosted trees: Un conjunto de árboles de regresión utilizando el algoritmo LSBoost. En comparación con el bagging, los algoritmos de refuerzo usan relativamente poco tiempo o memoria, pero podrían necesitar más miembros del conjunto.

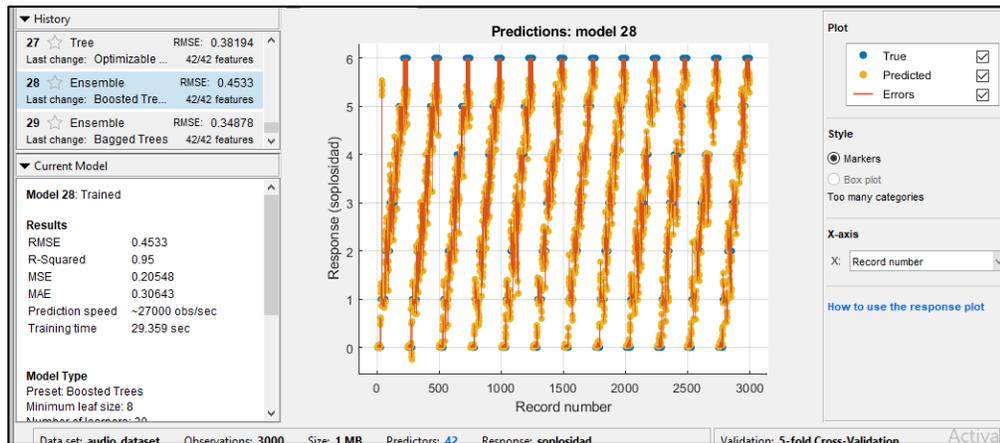


Figura 76. Gráfico de resultado obtenido con el modelo de Boosted tres en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.4533, este se puede visualizar en la Figura 76.

Bagged trees: conjunto agregado de árboles de regresión. A menudo es muy preciso, pero puede ser lento e intensivo en memoria para grandes conjuntos de datos

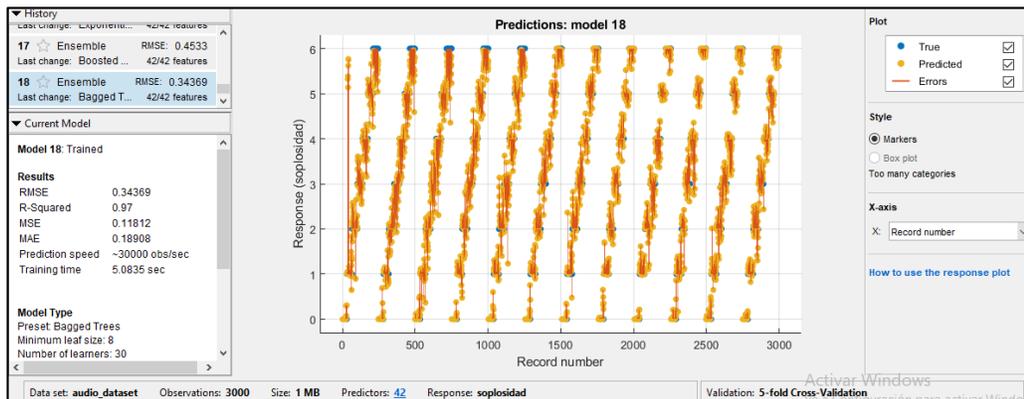


Figura 77. Gráfico de resultado obtenido con el modelo de bagged trees en Matlab

El resultado obtenido con este modelo de error cuadrático medio (RMSE) es de 0.34369, este se puede visualizar en la Figura 77.

Se realizó un cuadro comparativo con los diferentes valores obtenidos en las regresiones implementadas en Matlab, este se puede observar en la Tabla 52.

Tabla 52. Resultados obtenidos de los diferentes modelos de regresión de Matlab.

		RMSE
MODELOS DE REGRESIÓN LINEAL	LINEAR	0.46835
	INTERACTIONS LINEAR	0.33072
	ROBUST LINEAR	0.47203
MODELOS DE ATRIBUTOS DE REGRESIÓN	FINE TREE	0.38483
	MEDIUM TREE	0.42576
	COARSE TREE	0.52989
SOPORTE DE MAQUINA DE VECTORES	LINEAR SVM	0.47705
	QUADRATIC SVM	0.31148
	CUBIC SVM	0.33832
	FINE GAUSSIAN	1.1521
	MEDIUM GAUSSIAN	0.34641
	COARSE GAUSSIAN	0.42974
MODELOS DE REGRESIÓN DE PROCESOS GAUSSIANOS	RATIONAL QUADRATIC	0.25076
	SQUARED EXPONENTIAL	0.25908
	MATERN 5/2	0.24834
	EXPONENTIAL	0.27327
CONJUNTO DE ÁRBOLES	BOOSTED TRESS	0.4533
	BAGGET TREES	0.34369

El modelo de regresión escogido es, Matern 5/2 de procesos gaussianos, ya que presenta el menor error frente a los demás modelos de regresión, y su respuesta pronosticada frente a la respuesta verdadera es satisfactoria como podemos observar en la Figura 78. Se seguirá trabajando con este modelo para optimizarlo aún más y conseguir un error menor.

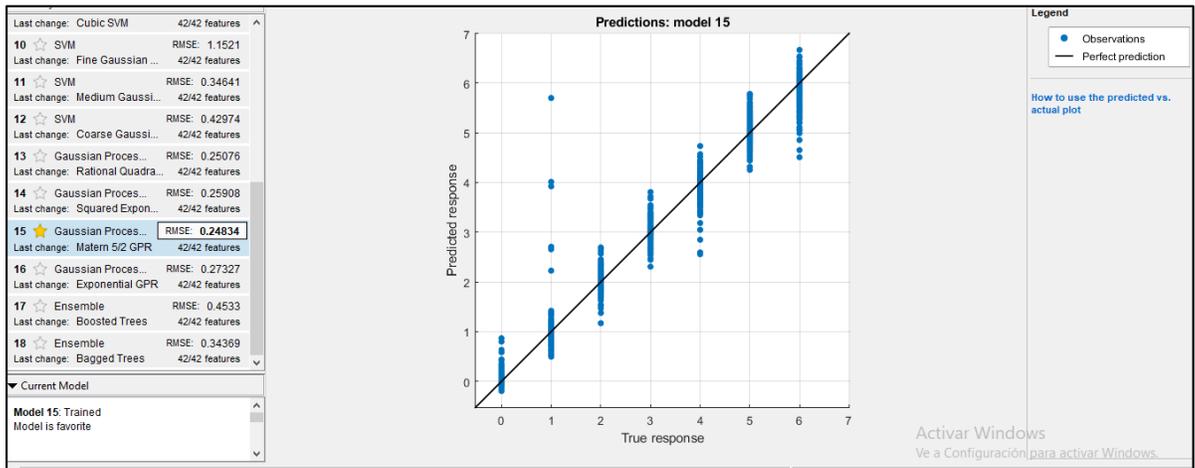


Figura 78. Gráfica respuesta pronosticada frente a la respuesta verdadera del modelo Matern 5/2, la línea trazada en color negro, nos indica la predicción óptima, a la cual se busca llegar, los puntos azules no indican la respuesta del modelo.

Para optimizar el modelo escogido, se seleccionó una función avanzada que proporciona Matlab, esta permite pasar de un kernel de 5/2 a uno de 3/2. Con esto se obtuvo un menor error como se observa en la Figura 79, con un error de 0.24513.

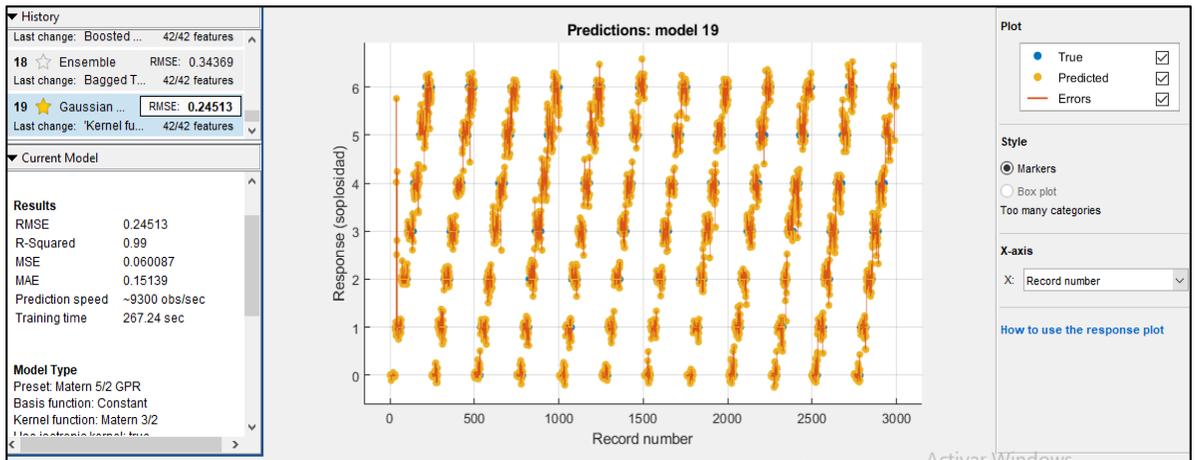


Figura 79. Gráfica modelo gaussiano optimizado con kernel de 3/2

Otra herramienta que se utilizó de MATLAB para intentar optimizar el modelo, es la selección de características, está la implementamos con los 7 parámetros

seleccionados anteriormente con el programa WEKA, estos están registrados en la Tabla 13, esto junto con el modelo escogido en MATLAB, el resultado obtenido no fue satisfactorio con un error de 0.35317 como se observa en la Figura 80.

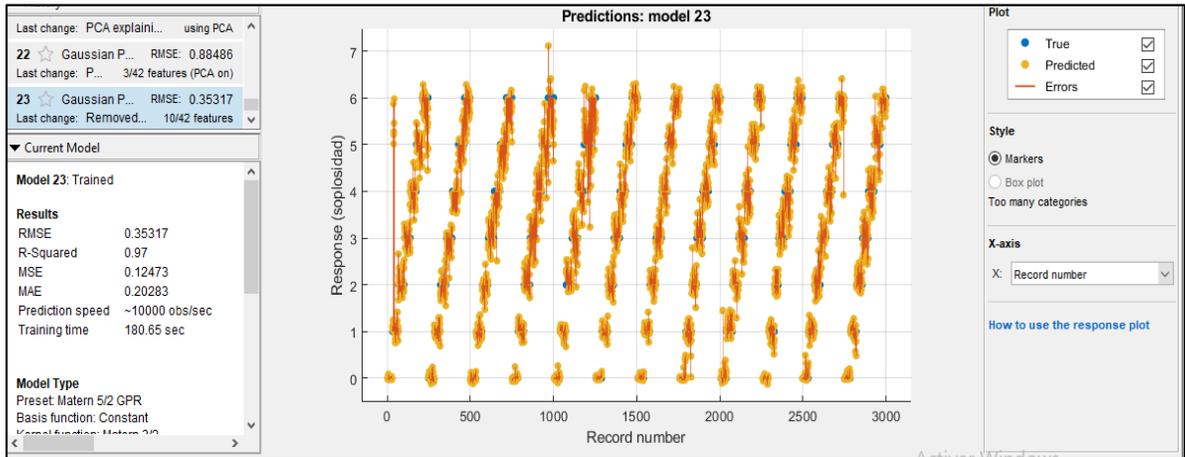


Figura 80. Gráfica Modelo gaussiano Matern, entrenado con características seleccionadas

Intentando optimizar un poco más el modelo, se utilizó la opción de PCA que ofrece Matlab para reducir dimensionalidad Figura 81, esta me proporciona el criterio por varianza, este hace que las variables escogidas tengan una varianza de más del 95%, este valor se sugiere por defecto en Matlab, además según la referencia (Rio, 2019), los valores de confianza sugeridos deben estar entre 90%, 95% y 99%.

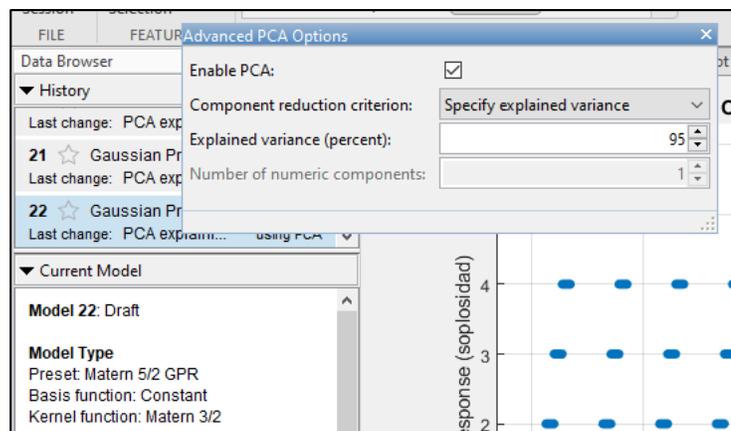


Figura 81. Opción PCA para reducción de dimensionalidad

Al seleccionar la opción del PCA de varianza, el modelo selecciona 3 características de 42, el resultado obtenido no es satisfactorio ya que no contribuye en la disminución del error, por el contrario aumenta en gran medida, por lo tanto esta opción es descartada, el resultado obtenido se puede observar en la Figura 82.

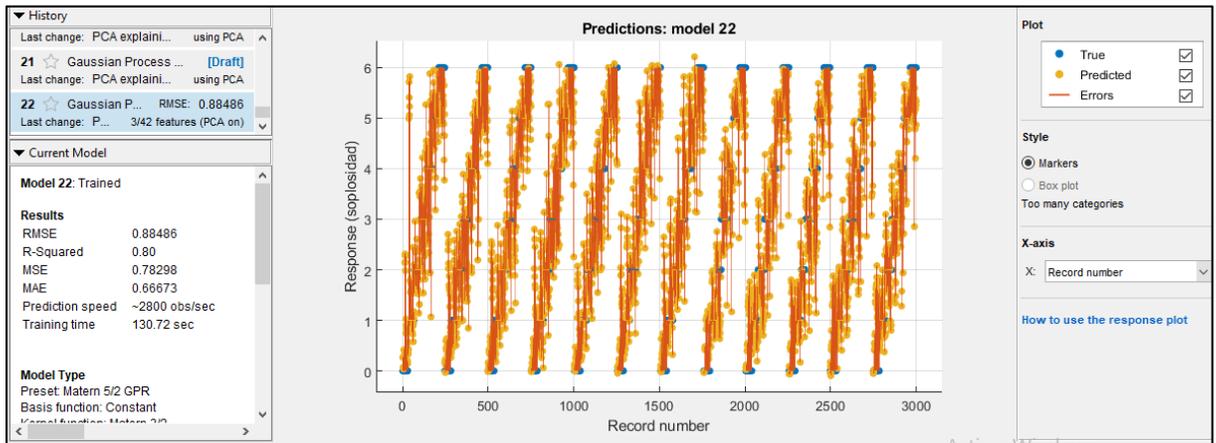


Figura 82. Gráfica entrenamiento con herramienta PCA, con la opción de varianza

Otra opción que ofrece el PCA es la reducción de dimensionalidad de número de componentes principales, con esta opción se fueron seleccionando componentes de tres en tres hasta completar 42, esto junto con la opción del modelo de regresión escogido Matern 3/2 de procesos gaussianos, dado que fue el modelo que obtuvo mejor resultado, los valores de error obtenidos se observan en la Tabla 53.

Tabla 53. Valores de error obtenidos en entrenamiento con diferente número de componentes y modelo de regresión Matern 3/2

Número de componentes	Error
3	0.88492
6	0.87248
9	0.67443

12	0.5994
15	0.46771
18	0.37955
21	0.33589
24	0.29905
27	0.29026
30	0.28305
33	0.2761
36	0.26652
39	0.26552
42	0.26481

Se pudo concluir que la reducción de dimensionalidad con análisis de componentes principales no arroja un resultado positivo para la reducción del error, ya que se obtiene un mejor resultado haciendo uso de todas las características como se puede visualizar en la Figura 83, donde a medida que se aumentan gradualmente las características, el error que se obtiene disminuye.

Reducción de dimensionalidad

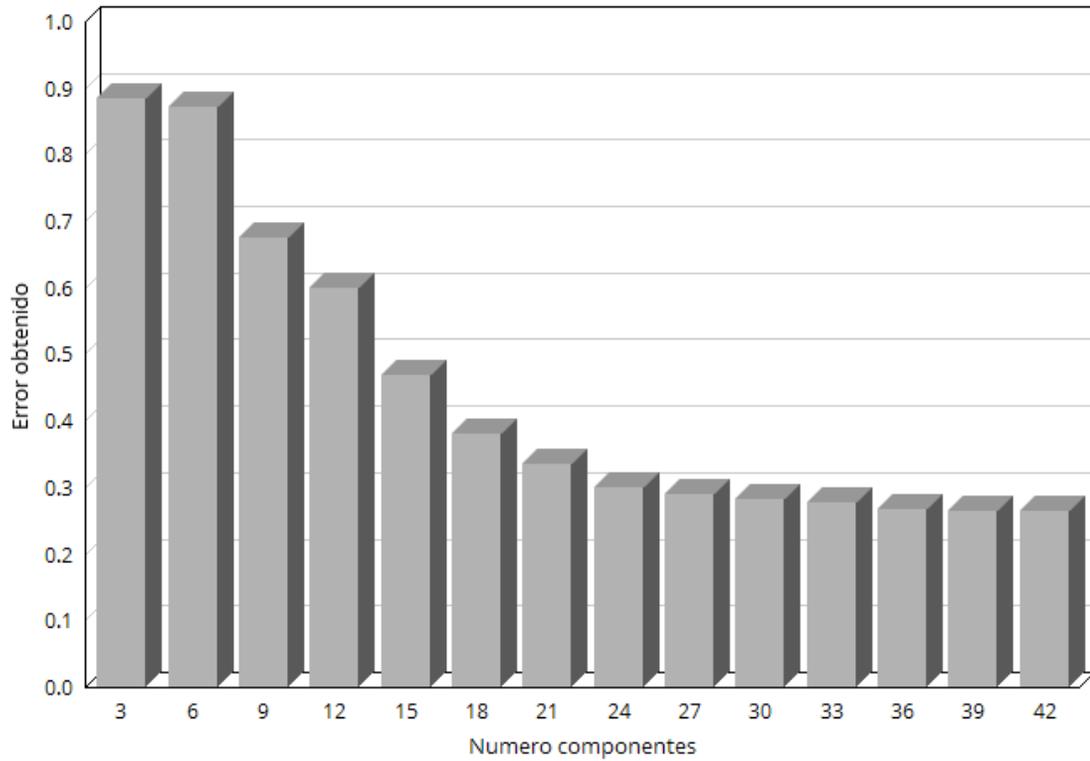


Figura 83. Reducción de dimensionalidad

Otra opción de reducción de dimensionalidad es descartando los parámetros que están duplicados, es decir, el mismo parámetro calculado por diferentes métodos. Se dejó la opción más confiable y se descartaron los otros, con esta opción se obtuvo un modelo confiable y con error de 0.22879 como se puede observar en la Figura 84.

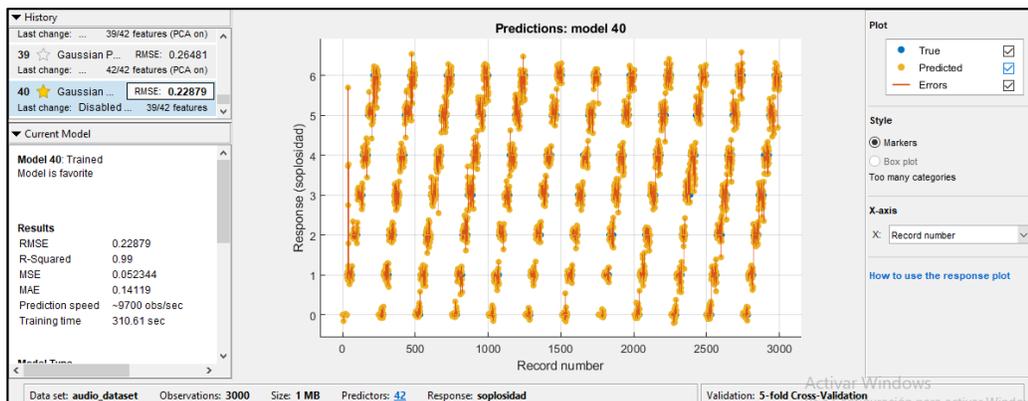


Figura 84. Reducción de dimensionalidad con método de parámetros descartados

El propósito del proyecto en el cual se enmarca este trabajo, es predecir todo el examen de evaluación perceptual CAPE-V, por lo tanto, adicionalmente se buscará predecir la rugosidad, astenia y tensión, es por esto que estas características no se deben tomar para realizar este entrenamiento, al retirarlas y entrenar el modelo que ya se seleccionó anteriormente, el error aumenta 0.29496, pero es viable para continuar el desarrollo del proyecto.

7. RESULTADOS

El modelo seleccionado es Matern 3/2 de procesos gaussianos, que presentó un error de 0,29496.

Después de crear modelos de regresión de forma interactiva en la aplicación regresión Learner, Matlab nos permite exportar el mejor modelo al área de trabajo. Después de exportar un modelo al espacio de trabajo desde el aprendizaje de regresión, o ejecutar el código generado desde la aplicación, se obtiene una estructura que puede usarse para realizar predicciones con datos nuevos.

Se crea una función que entrega valores de soplosidad de 0 a 6, ya que nuestros datos de soplosidad siempre estuvieron establecidos en esos rangos desde la creación de los audios, pero debido a que se plantea la estimación de la soplosidad en escala CAPE-V, la cual como se mencionó en un capítulo anterior, evalúa de 0 a 100, se debe escalar la función para que nos arroje valores dentro de este rango, y se redondean para que solo tome valores enteros.

Después de obtener la soplosidad, hacemos uso de la función num2str, la cual convierte números a matriz de caracteres.

Esta función posteriormente se llama desde la herramienta de Apps Designer de Matlab Figura 85, La cual nos permite crear aplicaciones, y ofrece un amplio entorno de desarrollo que ofrece un gran conjunto de componentes interactivos, un flujo de programación sólido, y un administrador de diseño.



Figura 85. Interface gráfica de inicio Apps Designer Matlab

Con esta herramienta se creó una App que cuenta con un botón, que nos permite seleccionar el archivo de sonido que se quiere analizar, y otro botón el cual, al seleccionarlo, nos arrojará el valor de soplosidad estimado. La interfaz gráfica de la App, en la que se encuentran estas opciones la podemos observar en la Figura 86.



Figura 86. Interfaz gráfica de usuario de la App, calculadora automática de soplosidad

El resultado es satisfactorio, ya que al realizar la prueba de funcionamiento con audios sintéticos del cual ya se conoce su valor real de soplosidad, la App nos arroja el valor correcto de soplosidad correspondiente a cada audio, este proceso lo podemos observar en la Figura 87, donde se selecciona el archivo de audio. Figura 88, donde ya se ha cargado el archivo que se va analizar. Figura 89, cuando se selecciona la opción calcular soplosidad, nos arrojó el valor de soplosidad estimado.

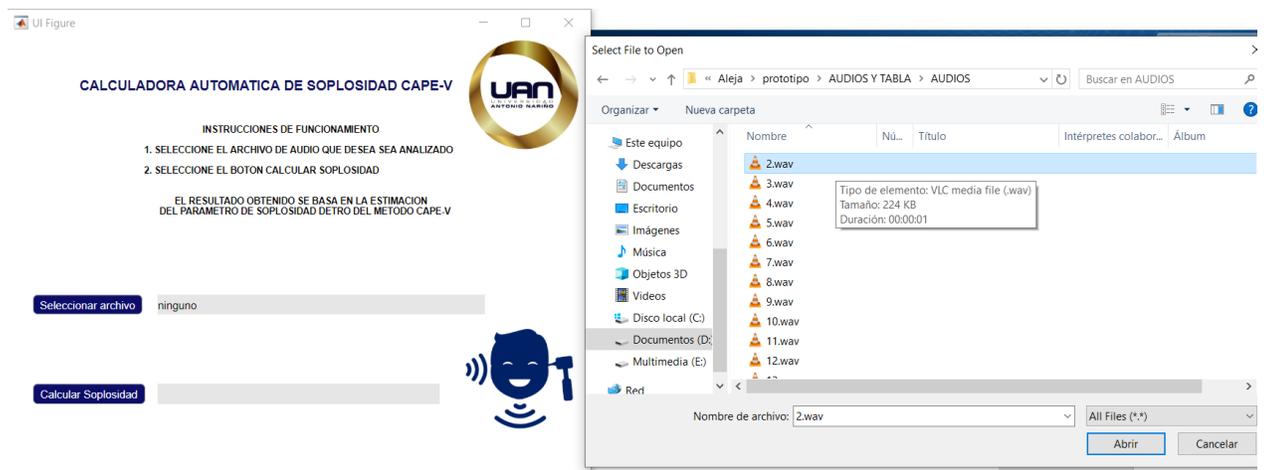


Figura 87. Prueba de funcionamiento de la App calculadora automática de soplosidad, selección de archivo de audio.



Figura 88. Prueba de funcionamiento de la App calculadora automática de soplosidad, archivo ya cargado para ser analizado



Figura 89. Prueba de funcionamiento de la App calculadora automática de soplosidad, resultado de soplosidad estimado.

7.1 PRUEBAS REALIZADAS CON AUDIOS DE VOCES REALES

Se realizó una pequeña prueba con 6 voces reales, en las cuales el fonoaudiólogo superviso el protocolo de toma de estas. Los resultados obtenidos en la prueba se pueden observar en la Tabla 54.

Tabla 54. Resultados obtenidos a partir de pruebas realizadas en voces reales.

Nombre Audio	Regresión original	Valor escalado
ID142_A	1.38	23
ID145_A	-0.55	-9
ID149_A	-0.25	-4
ID150_A	-1.4	-23
ID157_A	-0.8	-14
ID158_A	1.6	26

Los resultados obtenidos no son satisfactorios en algunos casos, pero en otros sí. Se debe mirar los valores que se tienen y analizar los parámetros acústicos, además

de esto, validar las muestras que se tienen con un fonoaudiólogo con experiencia, lo ideal es que sean varios fonoaudiólogos en un consenso, en el que estos evalúen las muestras y les asignen un valor de soplosidad. Y con base en esto obtener más muestras, para ser analizadas y realizar más pruebas, ya que 6 muestras son muy pocas para determinar el buen o mal funcionamiento de la App. El sistema debe ser validado con muchas más muestras para tener un resultado más confiable, además de esto comparar los resultados de soplosidad dados por los fonoaudiólogos con las muestras y determinar si existe algún tipo de correlación.

8. CONCLUSIONES

- En la selección de atributos, se pudo concluir que, todas las características extraídas (temporales, espectrales, cepstrales y de potencia.) son necesarias para lograr una buena regresión con un error bajo.
- La reducción de dimensionalidad utilizando la herramienta Weka y sus diferentes criterios de selección de atributos, no dio el resultado esperado, porque a pesar de utilizar el conjunto de atributos seleccionado, el error en la regresión aumento.
- La reducción de dimensionalidad utilizando el método PCA de Matlab, no fue satisfactorio, dado que, entre menor cantidad de atributos seleccionados, el error aumento.
- Matlab es un software con gran potencial para hacer regresión, presenta diversas ventajas sobre el programa Weka, dentro de estas está que las gráficas son más fáciles de comprender e interpretar que las de Weka, además de esto Matlab permite exportar el modelo de regresión.
- Un inconveniente es que Matlab no proporciona la ecuación del modelo, para que pueda ser implementado en otras plataformas.
- La creación de un banco de voces sintético para realizar procesos de regresión es una opción viable, y resulta ser más objetivo que un banco de voces reales.
- Al utilizar voces sintéticas la variabilidad no es muy grande, lo que permitió realizar procesos de submuestreo, permitiendo ampliar el tamaño del banco de voces.
- La aplicación de Matlab regression learner permite de una manera fácil aplicar varios algoritmos de regresión e identificar el más eficiente.

8.1 TRABAJOS FUTUROS

- En un futuro, se puede aplicar el modelo aquí expuesto, a todo el método de evaluación perceptual CAPE-V, en el cual se buscaría predecir los otros parámetros del examen como son: rugosidad, astenia y tensión.

- El modelo exportado se puede llegar a trabajar en una aplicación web o móvil que sea más portable para que pueda ser usado por fonoaudiólogos en trabajo de campo.
- El modelo debe ser probado con voces humanas ya que en este caso se creó y se validó con audios sintéticos.
- Se podría probar otro método de regresión basado en redes neuronales y deep learning, para analizar si la eficiencia es mejorada.
- Se puede experimentar con otro método de regresión que no se base en la extracción de características si no que la entrada, sea el audio completo.

BIBLIOGRAFÍA

- Batalla, D. F. (2015). *sociedad española de otorrinolaringología y cirugía de cabeza y cuello*. Obtenido de seorl.net: <https://seorl.net/wp-content/uploads/2015/09/CURSO-DE-VOZ.pdf>
- Botero Tobón, L. M. (2008). *Caracterización de los indicadores acústicos*. caldas: El artista.
- Brockmann, M. (2011). *Improving jitter and shimmer measurements in normal voices*. reino unido: Newcastle University.
- Brockmann, M. (2011). *Improving jitter and shimmer measurements in normal voices*. reino unido: Newcastle University.
- Carla Oliveira, C. L. (2013). Vocal Acoustic Analysis - Jitter, Shimmer and HNR Parameters. *ELSEVIER*, 1112-1122.
- Cobeta, I., Núñez, F., & Fernández, S. (2013). Clasificación de las enfermedades neurodegenerativas. En p. d. voz. Barcelona: ICG Marge, SL.
- Fang-Ling Lu, S. M. (2014). Tareas del habla y confiabilidad entre evaluadores en la evaluación perceptiva de la voz. *Journal of Voice. Elsevier.28(6)*, 725-732.
- fermino, m. a. (2017). Clasificación de trastornos vocales utilizando redes neurales artificiales. Cornélio Procópio, brasil.
- Ford, C., & Bless, D. (1986). Clinical experience with injectable collagen for vocal fold augmentation. *Laryngoscope*. En C. Ford, & D. Bless.
- Galan, M. M. (julio de 2014). Análisis perceptual de la disfonía: adaptación y validación del método CAPE-V al español. oviedo.
- García Mario Alejandro, R. A. (2016). Red neuronal profunda para analisis acustico. *Red neuronal profunda para analisis acustico*. Universidad Tecnológica Nacional Facultad Regional Córdoba.
- García Mario Alejandro, R. A. (s.f.). *Extracción de Características en Audio con Redes Neuronales convolucionales*. Cordoba: Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN FRC).
- Gerardo Gutiérrez Gutiérrez, A. R. (2015). *centro de ciencias de la computación*. Obtenido de Instituto Tecnológico Metropolitano: <http://c3.itm.edu.co/~gerardo.gutierrez/machinelearning/Introducci%C3%B3n%20Aprendizaje%20de%20M%C3%A1quina.pdf>
- Higor E. Pavoni, M. A. (2016). Identificación de parámetros subjetivos de voz utilizando redes neuronales artificiales. *Identificación de parámetros subjetivos de voz utilizando redes neuronales artificiales*. Cornélio procopio, Brasil.
- Ignacio Cobeta, f. n. (2013). *patología de la voz*. Marge medica books.
- Jaume de Montserrat, A. O. (s.f.). *El uso profesional*. catalunya: Reversible SCP.
- Kreiman, J. &. (2010). Perceptual assessment of voice quality: past,. *Perspectives on Voice and Voice Disorders*.
- Martinez, J. F. (2011). *Estimación espectral*.
- mateos, c. (18 de abril de 2017). *sociedad española de otorrinolaringología*. Obtenido de https://seorl.net/wp-content/uploads/2016/05/NP_Dia-Mundial-de-la-Voz.pdf

- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Londres, Inglaterra: Massachusetts Institute of Technology.
- N., M. G. (2016). evaluación funcional de la voz. *Artículo de divulgación científica en el área vocal*. Chile.
- Nicolas Viancha, C. M. (2018). *studocu*. Obtenido de <https://www.studocu.com/co/document/universidad-antonio-narino/electronica-analoga/practica/lab-espectro-senalesfinalizadob/4874733/view>
- Pericas, F. J. (1993). TECNICAS DE PROCESADO Y REPRESENTACION DE LA SEÑAL DE VOZ PARA EL RECONOCIMIENTO DEL HABLA EN AMBIENTES RUIDOSOS. Barcelona.
- Polo, N. (2015). *sottovoce*. Obtenido de <https://sottovoce.hypotheses.org/tag/voz-area>
- Preparadores de Oposiciones para la Enseñanza*. (s.f.). Obtenido de La voz humana y su fisiología: <https://www.preparadores.eu/secundaria/Musica/Musica-Tema.pdf>
- R.I. Zraick, G. K. (2011). Establishing validity of the Consensus Auditory-. *Journal*, 14-22.
- Santos, J. f. (2017). Reconocimiento de patrones en medidas acústicas para identificar patologías laríngeas. Curitiba, Brasil.
- Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Analisis acustico vocal- Parametros de Jitter, Shimmer y HNR. *Conferencia sobre sistemas y tecnologias de la informacion en salud y asistencia social*, (págs. 1-2). Braganca.
- Trinite, B. (2017). Epidemiology of Voice Disorders in Latvian School Teachers. *Elsevier*, 508.e1-508.e9.
- universidad de alcala*. (s.f.). Obtenido de departamento de cirugía.
- Williamson, G. (1 de 02 de 2014). *Evaluación perceptiva de la voz*. Obtenido de SLTinfo: <https://www.sltinfo.com/perceptual-assessment-of-voice/>