



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

CENTRO DE CIENCIAS BÁSICAS

TESIS

CLASIFICACIÓN DE LA DISFUNCIÓN TEMPOROMANDIBULAR
CON ALGORITMOS DE APRENDIZAJE SUPERVISADO

PRESENTA

Carlos Alberto López Hernández

PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS DE LA
COMPUTACIÓN

COTUTORES

Dr. Rogelio Salinas Gutiérrez

Dr. Omar Gutiérrez Navarro

INTEGRANTE DEL COMITÉ TUTORAL

Dra. Ivonne Bazán Trujillo

Aguascalientes, Ags, 3 de septiembre de 2024



M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS
P R E S E N T E

Por medio del presente como **Miembros del Comité Tutoral** designado del estudiante **CARLOS ALBERTO LÓPEZ HERNÁNDEZ** con ID 223348 quien realizó la tesis titulada: **CLASIFICACIÓN DE LA DISFUNCIÓN TEMPOROMANDIBULAR CON ALGORITMOS DE APRENDIZAJE SUPERVISADO**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia damos nuestro consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que nos permitimos emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Ponemos lo anterior a su digna consideración y sin otro particular por el momento, le enviamos un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"
Aguascalientes, Ags., a 26 de agosto de 2024.



Dr. Rogelio Salinas Gutiérrez
Co-Tutor de tesis



Dr. Omar Gutiérrez Navarro
Co-Tutor de tesis



Dra. Ivonne Bazán Trujillo
Asesora de tesis

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad
Aprobado por: Depto. Control Escolar/Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-16
Actualización: 00
Emisión: 17/05/19



DICTAMEN DE LIBERACIÓN ACADÉMICA PARA INICIAR LOS TRÁMITES DEL EXAMEN DE GRADO



Fecha de dictaminación dd/mm/aaaa: 26/08/2024

NOMBRE: Carlos Alberto López Hernández ID 223348

PROGRAMA: Maestría en Ciencias de la Computación, Matemáticas Aplicadas LGAC (del posgrado): Computación-Inteligencia Artificial

TIPO DE TRABAJO: (X) Tesis () Trabajo Práctico

TÍTULO: Clasificación de la Disfunción Temporomandibular con Algoritmos de Aprendizaje Supervisado

IMPACTO SOCIAL (señalar el impacto logrado): Se generaron modelos de clasificación que permiten la diferenciación entre individuos sanos y los que tienen trastorno temporomandibular artralgia. La comparación del desempeño de estos modelos nos permitió identificar el mejor para esta tarea de clasificación.

INDICAR	SI	NO	N.A. (NO APLICA)	SEGÚN CORRESPONDA:
Elementos para la revisión académica del trabajo de tesis o trabajo práctico:				
SI				El trabajo es congruente con las LGAC del programa de posgrado
SI				La problemática fue abordada desde un enfoque multidisciplinario
SI				Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI				Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI				Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI				El trabajo demuestra más de una aportación original al conocimiento de su área
NO				Las aportaciones responden a los problemas prioritarios del país
SI				Generó transferencia del conocimiento o tecnológica
SI				Cumple con la ética para la investigación (reporte de la herramienta antiplagio)
El egresado cumple con lo siguiente:				
SI				Cumple con lo señalado por el Reglamento General de Docencia
SI				Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc)
SI				Cuenta con los votos aprobatorios del comité tutorial, en caso de los posgrados profesionales si tiene solo tutor podrá liberar solo el tutor
N.A.				Cuenta con la carta de satisfacción del Usuario
SI				Coincide con el título y objetivo registrado
SI				Tiene congruencia con cuerpos académicos
SI				Tiene el CVU del Conacyt actualizado
N.A.				Tiene el artículo aceptado o publicado y cumple con los requisitos institucionales (en caso que proceda)
En caso de Tesis por artículos científicos publicados				
N.A.				Aceptación o Publicación de los artículos según el nivel del programa
N.A.				El estudiante es el primer autor
N.A.				El autor de correspondencia es el Tutor del Núcleo Académico Básico
N.A.				En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación.
N.A.				Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
N.A.				La aceptación o publicación de los artículos en revistas indexadas de alto impacto

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado: SI No

Elaboró: FIRMAS

* NOMBRE Y FIRMA DEL CONSEJERO SEGÚN LA LGAC DE ADSCRIPCIÓN: Dr. Rogelio Salinas Gutiérrez

NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO: Dra. Mariana Álvarez Gómez

* En caso de conflicto de intereses, firmará un revisor miembro del NAB de la LGAC correspondiente distinto al tutor o miembro del comité tutorial, asignado por el Decano

Revisó: Dr. Alejandro Padilla Díaz

NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO:

Autorizó: M. en C. Jorge Martín Álvarez Chávez

NOMBRE Y FIRMA DEL DECANO:

Nota: procede el trámite para el Depto. de Apoyo al Posgrado
 En cumplimiento con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: ... Cuidar la eficiencia terminal del programa de posgrado y el Art. 105F las funciones del Secretario Técnico, llevar el seguimiento de los alumnos.

Elaborado por: D. Apoyo al Posg.
 Revisado por: D. Control Escolar/D. Gestión de Calidad.
 Aprobado por: D. Control Escolar/ D. Apoyo al Posg.

Código: DO-SEE-FD-15
 Actualización: 01
 Emisión: 28/04/20

Agradecimientos

Al CONAHCYT por el apoyo financiero brindado a través del programa de becas nacionales, para la realización de mi posgrado.

A la Universidad Autónoma de Aguascalientes que me proporcionó herramientas, espacios y oportunidades para desarrollar mi proyecto de investigación.

Al Dr. Rogelio Salinas Gutiérrez por su constante guía y supervisión, que a lo largo de estos últimos dos años han contribuido enormemente a mi formación profesional.

Al Dr. Omar Gutiérrez Navarro por la confianza y apertura para colaborar en sus proyectos de investigación, así como por su consejo y ayuda brindada siempre que fue necesario.

A la Dra. Ivonne Bazán Trujillo, por su tiempo invertido en la asesoría y mejora constante de mi proyecto de investigación.

A mi familia y amigos, que me acompañaron siempre en los momentos más arduos de estos últimos dos años.

Índice general

Índice de figuras	3
Índice de tablas	6
Acrónimos	8
1. Introducción	11
1.1. Planteamiento del problema	11
1.2. Justificación	12
1.3. Objetivos	12
1.4. Articulación Temporomandibular	13
1.5. Trastornos Temporomandibulares	14
1.6. Datos Disponibles	16
2. Marco Teórico	19
2.1. Visión Computacional y Seguimiento de Movimientos	19
2.2. Aprendizaje Supervisado	20
2.2.1. Máquina de Soporte Vectorial	22
2.2.2. K-Vecinos más Cercanos	24
2.2.3. Perceptrón Multicapa (Red Neuronal)	26
2.3. Clasificación Probabilística	29
2.3.1. Clasificador Ingenuo de Bayes	29
2.3.2. Funciones Cópula	29
2.3.3. Clasificador Probabilístico con Cópulas	33
2.4. Estado del Arte	33
3. Metodología	37
3.1. Preprocesamiento de las Señales	38
3.1.1. Filtrado y Segmentación de Señales	38

<i>ÍNDICE GENERAL</i>	2
3.1.2. Descarte por Frecuencia de Muestreo	40
3.2. Extracción de Atributos	40
3.2.1. Atributos Biomecánicos	40
3.2.2. Atributos Derivados del Desplazamiento	41
3.2.3. Atributos en el dominio de la frecuencia	42
3.2.4. Atributos en el Tiempo	42
3.2.5. Estandarización de Atributos	42
3.3. Selección de Atributos	43
3.3.1. Algoritmo UMDA	43
3.4. Entrenamiento de Clasificadores	45
3.4.1. Modelos a Entrenar	45
3.4.2. Validación Cruzada	46
3.5. Clasificadores con Funciones Cópula	46
3.5.1. Clasificador Probabilístico	46
3.5.2. Modelo Gráfico	47
3.6. Análisis de Resultados	49
3.6.1. Métricas de desempeño	49
4. Resultados	52
4.1. Resultados por Método de Selección	52
4.1.1. Resultados con todos los Atributos (sin selección)	53
4.1.2. Resultados con Atributos Seleccionados con Filtro de Correlación de Pearson	55
4.1.3. Resultados con Atributos Seleccionados con Filtro EDA	57
4.1.4. Resultados con Atributos Seleccionados con Filtro de Correlación de Pearson y EDA	59
4.2. Comparación de Resultados	61
5. Discusiones	70
5.1. Limitaciones del Trabajo	72
5.2. Trabajo a Futuro	72
6. Conclusiones	74
Bibliografía	76
Anexos	83

Índice de figuras

1.1. Localización y componentes la articulación temporomandibular. Por University of Dundee [1]	14
1.2. Clasificación de los TTM	15
1.3. Localización de los Marcadores en el Rostro del Paciente. Fuente: Elaboración propia.	16
1.4. Posición Tridimensional del Marcador 1	17
1.5. Posición Tridimensional del Marcador 2	17
2.1. Categorías del Aprendizaje Máquina	20
2.2. Diferencias entre Clasificación y Regresión. Elaboración propia.	21
2.3. Hiperplanos separando dos clases, recuperado de: [2]	22
2.4. Vectores de soporte y el margen de un hiperplano. Recuperado de [3]	22
2.5. Margen de un hiperplano. Recuperado de [2]	23
2.6. Ejemplo de transformación usando funciones kernel, recuperado de [2]	24
2.7. Ejemplo del funcionamiento de un clasificador K-vecinos más cercanos, recuperado de [4]	25
2.8. Modelo de McCullogh y Pitts [5]	26
2.9. Modelo de perceptrón [2]	27
2.10. Convergencia del error según el valor de η [5]	28
2.11. Comparación de redes neuronales [5]	28
2.12. Funciones asociadas a una cópula Gaussiana bivariada. Caso Cópula Frank. Obtenida con [6]	31
2.13. Estructura de dependencia para diferentes cópulas bivariadas	32
3.1. Diagrama de etapas en la metodología	37
3.2. Segmentación de las señales corregidas y filtradas.	39
3.3. Funcionamiento del Algoritmo UMDA	44
3.4. Ejemplo de un árbol de dependencia para un conjunto de d atributos	48

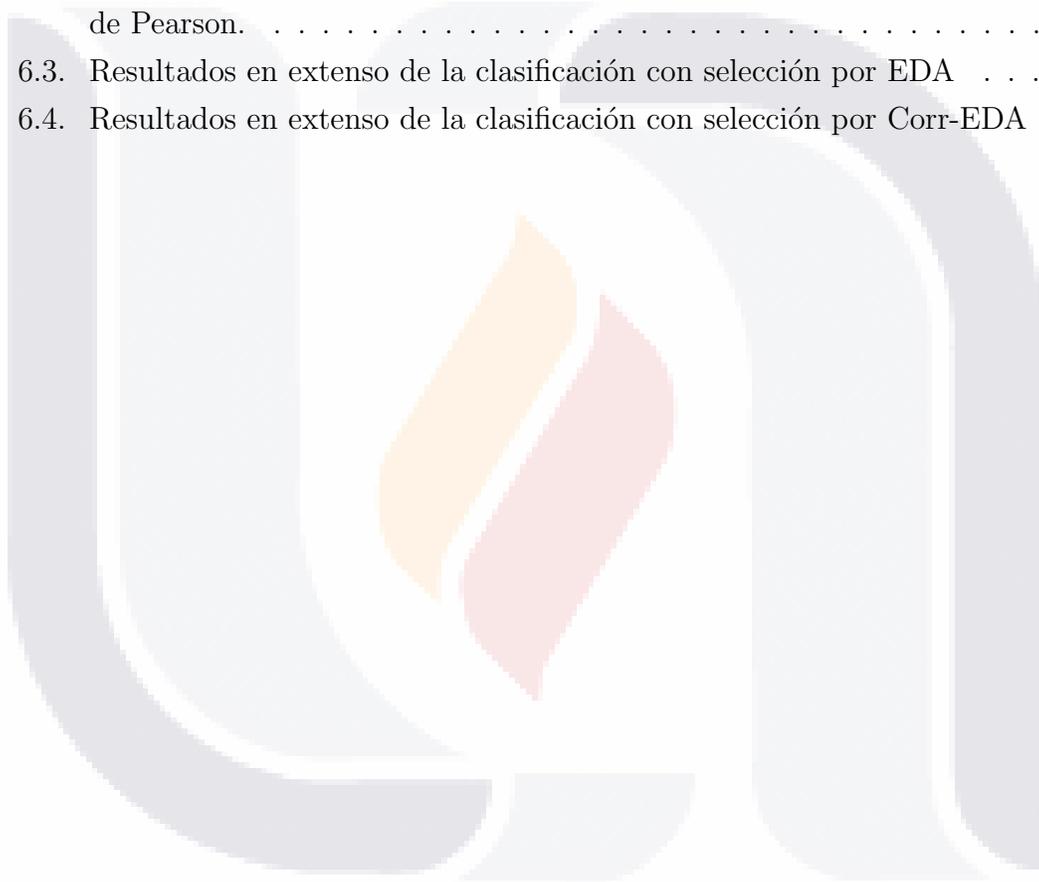
<i>ÍNDICE DE FIGURAS</i>	4
3.5. Ejemplo de una curva ROC	50
4.1. Curvas ROC para cada uno de los clasificadores analizados sin selección de atributos.	54
4.2. Curvas ROC para cada uno de los clasificadores analizados. Selección por correlación.	56
4.3. Curvas ROC para cada uno de los clasificadores analizados. Selección con EDA	58
4.4. Curvas ROC para cada uno de los clasificadores analizados. Selección con correlación y EDA.	60
4.5. Desempeño en exactitud por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.	62
4.6. Desempeño en AUC por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.	64
4.7. Desempeño en sensibilidad por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.	66
4.8. Desempeño en especificidad por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.	68
6.1. curva ROC para el clasificador con cópula Clayton con todos los atributos	83
6.2. curva ROC para el clasificador con cópula Frank con todos los atributos	83
6.3. curva ROC para el clasificador con cópula Gaussiana con todos los atributos	84
6.4. curva ROC para el clasificador KNN con todos los atributos	84
6.5. curva ROC para el clasificador NN con todos los atributos	84
6.6. curva ROC para el clasificador con todos los atributos	85
6.7. curva ROC para el clasificador SVM con todos los atributos	85
6.8. curva ROC para el clasificador con cópula Clayton con selección por correlación	85
6.9. curva ROC para el clasificador con cópula Frank con selección por correlación	86
6.10. curva ROC para el clasificador con cópula Gaussiana con selección por correlación	86

6.11. curva ROC para el clasificador KNN con selección por correlación . . .	86
6.12. curva ROC para el clasificador NN con selección por correlación	87
6.13. curva ROC para el clasificador con ingenuo de Bayes con selección por correlación	87
6.14. curva ROC para el clasificador SVM con selección por correlación . . .	87
6.15. curva ROC para el clasificador con cópula Clayton con selección por EDA	88
6.16. curva ROC para el clasificador con cópula Frank con selección por EDA	88
6.17. curva ROC para el clasificador con cópula Gaussiana con selección por EDA	88
6.18. curva ROC para el clasificador KNN con selección por EDA	89
6.19. curva ROC para el clasificador NN con selección por EDA	89
6.20. curva ROC para el clasificador ingenuo de bayes con selección por EDA	89
6.21. curva ROC para el clasificador SVM con selección por EDA	90
6.22. curva ROC para el clasificador con cópula Clayton con selección por correlación y EDA	90
6.23. curva ROC para el clasificador con cópula Frank con selección por correlación y EDA	90
6.24. curva ROC para el clasificador con cópula Gaussiana con selección por correlación y EDA	91
6.25. curva ROC para el clasificador KNN con selección por correlación y EDA	91
6.26. curva ROC para el clasificador NN con selección por correlación y EDA	91
6.27. curva ROC para el clasificador con ingenuo de Bayes con selección por correlación y EDA	92
6.28. curva ROC para el clasificador SVM con selección por correlación y EDA	92

Índice de tablas

2.1. Kernels comúnmente usados en máquinas de soporte vectorial	24
2.2. Funciones de distribución de cópulas bivariadas	32
2.3. Funciones de densidad de cópula bivariadas	33
3.1. Atributos biomecánicos extraídos	41
3.2. Atributos derivados del desplazamiento	41
3.3. Atributos del dominio de la frecuencia	42
3.4. Atributos en el Tiempo	42
3.5. Hiperparámetros de los Modelos de Clasificación	46
3.6. Matriz de Confusión	49
4.1. Media y desviación estándar de las métricas de los clasificadores. Sin selección de atributos.	53
4.2. Nombre de los atributos seleccionados por correlación de Pearson	55
4.3. Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con correlación de Pearson	56
4.4. Nombre de los atributos seleccionados con EDA	57
4.5. Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con EDA	57
4.6. Nombre de los atributos seleccionados por correlación de Pearson y EDA	59
4.7. Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con correlación de Pearson y EDA	59
4.8. Valores p de prueba Kruskal-Wallis para cada métrica de desempeño y método de selección.	61
4.9. Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica exactitud. $\alpha = 0.05$	63
4.10. Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica AUC. $\alpha = 0.05$	65

4.11. Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica Sensibilidad. $\alpha = 0.05$	67
4.12. Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica Especificidad. $\alpha = 0.05$	69
5.1. Atributos seleccionados por los 3 métodos de selección	71
6.1. Resultados en extenso de la clasificación sin selección de atributos.	92
6.2. Resultados en extenso de la clasificación con selección por Correlación de Pearson.	99
6.3. Resultados en extenso de la clasificación con selección por EDA	106
6.4. Resultados en extenso de la clasificación con selección por Corr-EDA	113



Acrónimos

Abreviatura	Descripción
ATM	Articulación Temporomandibular
AUC	Área bajo la curva ROC
Corr	Método de selección de atributos basado en la correlación de Pearson
Corr-EDA	Método de selección de atributos basado en correlación de Pearson y EDA
DC/TMD	Criterio Diagnóstico de Trastornos Temporomandibulares
EDA	Algoritmo evolutivo de estimación de distribución
FPR	Tasa de falsos positivos
KNN	Clasificador K vecinos más cercanos
MLP	Perceptrón multicapa. similar a red neuronal
NN	Clasificador Red Neuronal
ROC	Curva de Característica Operativa del Receptor
SVM	Clasificador Máquina de Soporte Vectorial
TPR	Tasa de verdaderos positivos
TTM	Trastorno Temporomandibular

Resumen

Los trastornos temporomandibulares (TTM) son un conjunto de afecciones de la articulación temporomandibular y los músculos asociados. Estas afecciones se presentan comunmente en la población en general y repercuten en la calidad de vida de quien las padece.

Este trabajo aborda la evaluación de clasificación de individuos sanos y con diagnóstico de TTM utilizando modelos de aprendizaje supervisado. Se extrajeron un total de 121 características de 59 registros de movimientos mandibulares obtenidos mediante un sistema de seguimiento de marcadores con una cámara de profundidad.

Se utilizan 3 métodos de selección de atributos para entrenar los modelos de clasificación. Siete clasificadores fueron entrenados con las selecciones de atributos, incluidos tres basados en funciones de cópula que modelan estructuras de dependencia. Se comparan las actuaciones de estos clasificadores. Los resultados muestran un buen desempeño para los clasificadores SVM y NN en las 4 métricas de desempeño obtenidas. Las funciones cópula muestran buenos resultados en la métrica sensibilidad, aunque no muestran una diferencia estadísticamente significativa con los de modelos como NN, SVM o Naive Bayes.

Abstract

Temporomandibular disorders (TMD) are a group of conditions that affect the temporomandibular joint and associated muscles. These conditions occur commonly in the general population and impact the quality of life of those who suffer from them.

This work addresses the classification of healthy individuals and those diagnosed with TMD using supervised learning models. A total of 121 features were extracted from 59 mandibular movement records obtained using a marker tracking system with a depth camera.

Three feature selection methods are used to train the classification models. Seven classifiers were trained with the attribute selections, including three based on copula functions that model dependency structures. The performances of these classifiers are compared. The results show good performance for the SVM and NN classifiers in the 4 performance metrics obtained. The copula functions show good results in metric sensitivity, although they do not show a statistically significant difference with those of models such as NN, SVM or Naive Bayes.

Capítulo 1

Introducción

1.1. Planteamiento del problema

Actualmente, el proceso de diagnóstico de los trastornos temporomandibulares (TTM) se basa en gran medida en la examinación física del paciente, el historial médico, el uso de criterios como el DC/TMD (Criterio Diagnóstico de Trastornos Temporomandibulares, por sus siglas en inglés) [7] y el empleo de recursos de imagenología. Esto implica que el diagnóstico de los TTM es realizado principalmente por médicos utilizando variables cualitativas. En este sentido, las propuestas de metodologías que se enfoquen en un análisis cuantitativo de datos podrían ampliar los recursos de diagnóstico disponibles para pacientes con estos trastornos.

Recientemente, la visión computacional ha sido propuesta como auxiliar en el diagnóstico de TTM [8], ya que mediante el análisis de material videográfico se pueden obtener datos cuantitativos sobre los movimientos mandibulares de un paciente. A partir de estos datos, es posible extraer información para caracterizar dichos movimientos y posteriormente emitir un diagnóstico. Analizar este tipo de datos presenta un reto para proveer información certera que pueda ser usada con fines diagnósticos.

En este proyecto, se propone el uso de algoritmos de aprendizaje supervisado para la clasificación de trastornos temporomandibulares en pacientes y analizar su desempeño en la tarea de diagnóstico de TTM. Por lo tanto, es relevante responder a la pregunta: ¿El uso de algoritmos de aprendizaje supervisado permite clasificar adecuadamente los TTM en pacientes?

1.2. Justificación

Los trastornos temporomandibulares (TTM) tienen una incidencia considerable tanto en la población de México como en el resto del mundo. Sus causas son multifactoriales y pueden ser de índole biológica, social, psicológica, entre otras, siendo el estrés y la ansiedad algunos de los desencadenantes más importantes [9]. Actualmente, el diagnóstico de los TTM se realiza principalmente mediante un examen físico del paciente realizado por un médico, y el uso de métodos cuantitativos basados en nuevas tecnologías es poco común.

El uso de la visión computacional para el diagnóstico de TTM es una propuesta no invasiva reciente y su aplicación es aún limitada. En este contexto, la propuesta y el análisis de algoritmos de clasificación para los TTM proporcionan información que puede guiar en la búsqueda del método óptimo para el análisis de este tipo de datos.

El uso de algoritmos de aprendizaje supervisado ofrece beneficios al proporcionar una metodología fiable para el análisis de datos de movimientos mandibulares, lo cual facilita la adopción generalizada de la visión computacional como herramienta diagnóstica. Esto aporta seguridad tanto a los pacientes como al personal médico, asegurando que los resultados ofrecidos son precisos.

1.3. Objetivos

- Objetivo General: Evaluar el desempeño de algoritmos de aprendizaje supervisado en la clasificación de pacientes en las clases sanos y con trastornos temporomandibulares.
- Objetivos Específicos:
 1. Caracterizar los datos de movimientos mandibulares para obtener información necesaria para el uso en los algoritmos de aprendizaje supervisado.
 2. Implementar un método de selección de características.
 3. Implementar al menos un algoritmo probabilístico para el problema de clasificación de TTM.
 4. Implementar al menos un algoritmo no probabilístico para el problema de clasificación de TTM.
 5. Evaluar los algoritmos implementados mediante métricas de desempeño.

1.4. Articulación Temporomandibular

La articulación temporomandibular (ATM) es la articulación responsable del movimiento de la mandíbula. Esta articulación es esencial en actividades como masticar y hablar. La disfunción de la ATM puede provocar problemas como dolor, dificultad para masticar y problemas del habla [10, 9]. La ATM es una articulación sinovial bilateral que conecta el hueso de la mandíbula con el hueso temporal del cráneo.

La ATM está formada por varios componentes, entre ellos dos superficies articulares: el cóndilo mandibular y la fosa glenoidea del hueso temporal. El cóndilo mandibular es una proyección redondeada en la parte superior de la mandíbula, mientras que la fosa glenoidea es una estructura cóncava en forma de cavidad ubicada en el hueso temporal [11, 12].

En el espacio entre las superficies articulares encontramos un disco fibrocartilaginoso, también llamado menisco. Este disco divide la articulación en dos compartimentos (superior e inferior) y proporciona amortiguación, estabilidad y asegura deslizamientos suaves durante los movimientos mandibulares [13].

La ATM recibe su suministro de sangre de las ramas de la arteria maxilar y de la arteria temporal superficial. La inervación sensorial de la articulación la proporcionan el nervio auriculotemporal y el nervio masetero, que transmiten información sensorial y percepción del dolor desde la articulación al sistema nervioso central [12].

Varios músculos participan en el movimiento de la ATM, incluidos los músculos de la masticación (por ejemplo: el temporal, masetero, pterigoideo medial y pterigoideo lateral), que son responsables de masticar y morder, así como otros músculos faciales que ayudan en el movimiento de la mandíbula y en la expresión facial [14, 15]. La Figura 1.1 muestra la articulación, su localización, así como sus componentes y algunos de sus músculos asociados.

Los músculos de la masticación desempeñan un papel central en la fisiología de la ATM al generar las fuerzas necesarias para el movimiento de la mandíbula. Los músculos temporal y masetero elevan la mandíbula durante la mordida y la masticación, mientras que los músculos pterigoideos lateral y medial contribuyen a los movimientos laterales y protrusivos de la mandíbula [12, 16].

La ATM funciona mediante una combinación de movimientos de deslizamiento y bisagra, lo que permite los movimientos complejos necesarios para actividades como masticar, hablar y tragar. Durante la apertura y el cierre de la mandíbula, el cóndilo mandibular experimenta un movimiento de rotación dentro de la fosa glenoidea. Esta acción similar a una bisagra permite el movimiento hacia arriba y hacia abajo de la

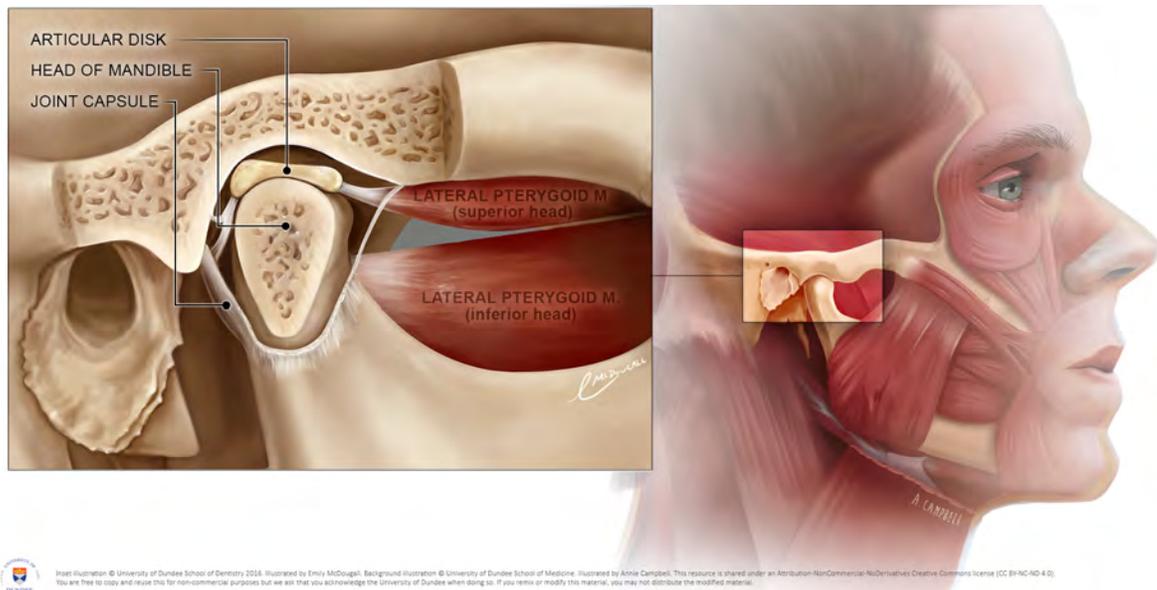


Figura 1.1: Localización y componentes la articulación temporomandibular. Por University of Dundee [1]

mandíbula inferior, facilitando la masticación y la mordida. Además del movimiento de bisagra, la ATM también presenta un deslizamiento o traslación del disco articular a lo largo de las superficies articulares. A medida que la mandíbula se mueve hacia adelante o hacia atrás, el disco articular se desliza a lo largo de la fosa glenoidea, lo que permite un movimiento mandibular suave y coordinado [16].

1.5. Trastornos Temporomandibulares

Los trastornos temporomandibulares (TTM) son un grupo de condiciones musculoesqueléticas asociadas al mal funcionamiento de la articulación temporomandibular. Los TTM pueden involucrar la articulación, los músculos masticatorios y todos los tejidos asociados. Estos trastornos son una de las causas más comunes de dolor en la boca y la cara, y tienen el potencial de causar dolores crónicos [9, 14]. Los trastornos de interés en este proyecto son aquellos que producen dolor en la articulación. Estos se pueden dividir en tres categorías: a) aquellos que producen dolor en la articulación, b) los que producen dolor en los músculos masticatorios y c) los dolores de cabeza asociados a estos dos últimos (vea Figura 1.2). Al trastorno doloroso de la articulación se le conoce como artralgia y al trastorno doloroso muscular se le conoce como mialgia [10, 17].

Las causas detrás de los trastornos temporomandibulares (TTM) son aún objeto de debate; sin embargo, se sugiere que son multifactoriales, incluyendo factores biológicos,

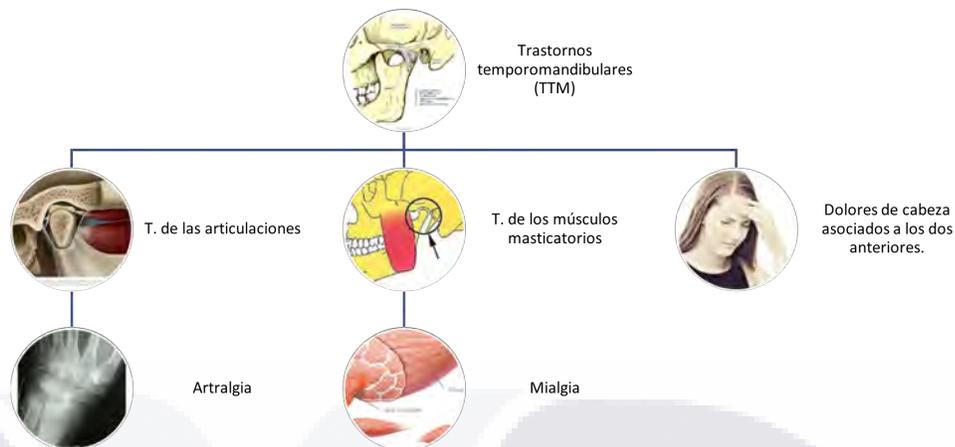


Figura 1.2: Clasificación de los TTM

psicológicos y sociales, entre otros. Estos factores pueden ser iniciadores, predisponentes y perpetuadores de un TTM [9, 18]. Por ejemplo, estudios han demostrado que la incidencia de trastornos temporomandibulares se duplica en personas con depresión y que la incidencia de dolores miofaciales aumenta en un 180 % en personas con ansiedad [19].

El diagnóstico de los TTM se basa en gran parte en historia clínica y resultados de evaluaciones físicas. Existen criterios de diagnóstico de los TTM que son de alta fiabilidad como el DC/TMD de [7] que usan un cuestionario de seis partes que puede responder el paciente. Entre los resultados de una evaluación física que sugieren la existencia de un TTM se encuentran los movimientos mandibulares anormales, un rango reducido de movimiento, sensibilidad en músculos masticatorios, signos de bruxismo, etc. [19]. Otros métodos de diagnóstico de trastornos temporomandibulares incluyen la obtención de imágenes por medio de mecanismos como resonancia magnética [17]; el análisis de biomarcadores asociados a los trastornos también puede ser usado para el diagnóstico [20].

Solo entre el 5 y el 10 % de los pacientes requieren tratamiento de TTM, y el 40 % de los pacientes tienen síntomas que se resuelven por sí solos. En un estudio de seguimiento a largo plazo, entre el 50 % y el 90 % de los pacientes experimentaron alivio del dolor después del tratamiento conservador. El enfoque multidisciplinario para el manejo de TTM ha tenido éxito. Los objetivos iniciales del tratamiento deben centrarse en la resolución del dolor y la disfunción. Más de 1500 personas en un registro de TTM en línea informaron que usaban antiinflamatorios (73 %), analgésicos de venta libre (56 %), antidepresivos (50 %), opioides (48 %), ansiolíticos (41 %) y relajantes musculares (40 %)

[19].

En México, los trastornos temporomandibulares se caracterizan por tener una prevalencia de entre el 30 % y el 70 % de la población. Generalmente se presenta en comorbilidad con enfermedades como artritis degenerativa y desarreglos oclusales. También se observa que una de las causales de los TTM es el estrés y que la incidencia de estos padecimientos es mayor en mujeres mayores a 15 años [21].

1.6. Datos Disponibles

Los datos disponibles para este proyecto de investigación provienen de una colaboración con la Escuela Nacional de Estudios Superiores Unidad León, UNAM. Se cuenta con 72 registros de pacientes realizando movimientos mandibulares de apertura/cierre, laterales y de protrusión/retrusión. De estos registros, 36 provienen de personas diagnosticadas con trastorno temporomandibular artralgia y 36 de personas sanas.

Los registros de los pacientes fueron obtenidos mediante la grabación de los movimientos mandibulares, realizando un seguimiento de marcadores en el rostro del paciente. Los movimientos fueron grabados con una cámara Intel Realsense Depth D435. En total, se colocaron dos marcadores en cada paciente: el marcador 1 (marcador de referencia) se colocó en el punto craneométrico nasión y el marcador 2 en el punto craneométrico gnación, tal como se muestra en la Figura 1.3.

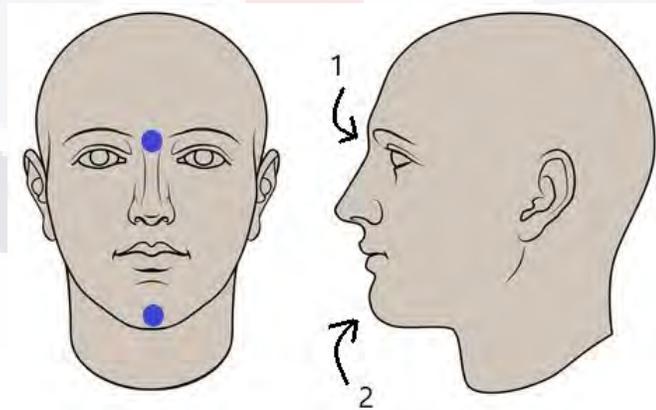


Figura 1.3: Localización de los Marcadores en el Rostro del Paciente. Fuente: Elaboración propia.

El seguimiento de los marcadores mencionados anteriormente se realizó utilizando la librería *OpenCV* en python. De este seguimiento se obtienen señales de la posición tridimensional de los marcadores durante los movimientos mandibulares. Es decir, cada

una de las señales (tres en total, por marcador) nos muestra como cambia la posición del marcador respecto al eje vertical, horizontal o de profundidad en el tiempo.

En las Figuras 1.4 y 1.5 se puede observar las señales como fueron obtenidas por el seguimiento de marcadores para un paciente diagnosticado con trastorno temporomandibular artralgia. La Figura 1.4 corresponde al marcador 1 y la Figura 1.5 corresponde al marcador 2. En ambos casos se muestra en el tercio superior la posición respecto al eje horizontal, en el tercio medio la posición respecto al eje vertical y en el tercio inferior la posición respecto al eje de profundidad.

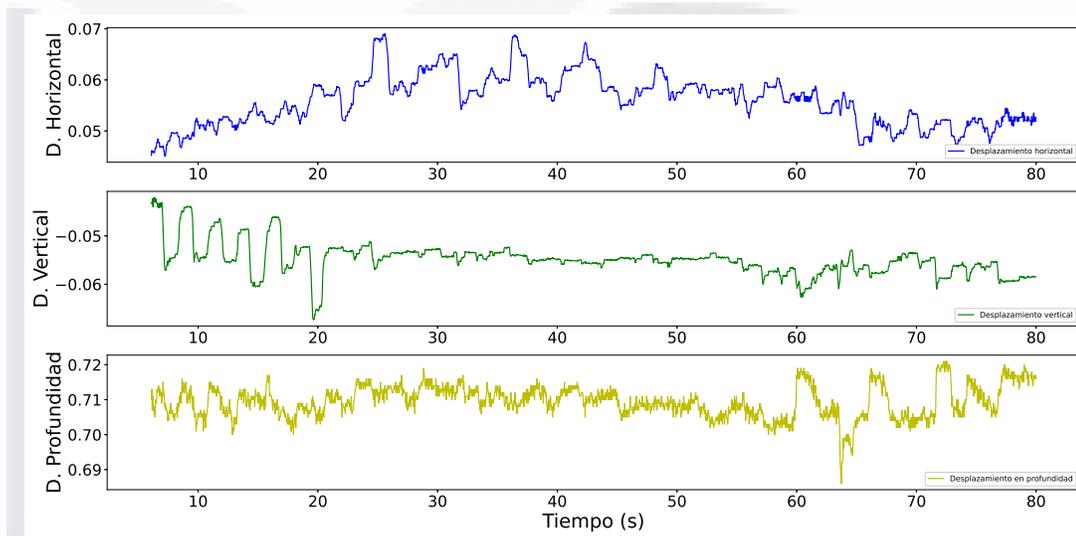


Figura 1.4: Posición Tridimensional del Marcador 1

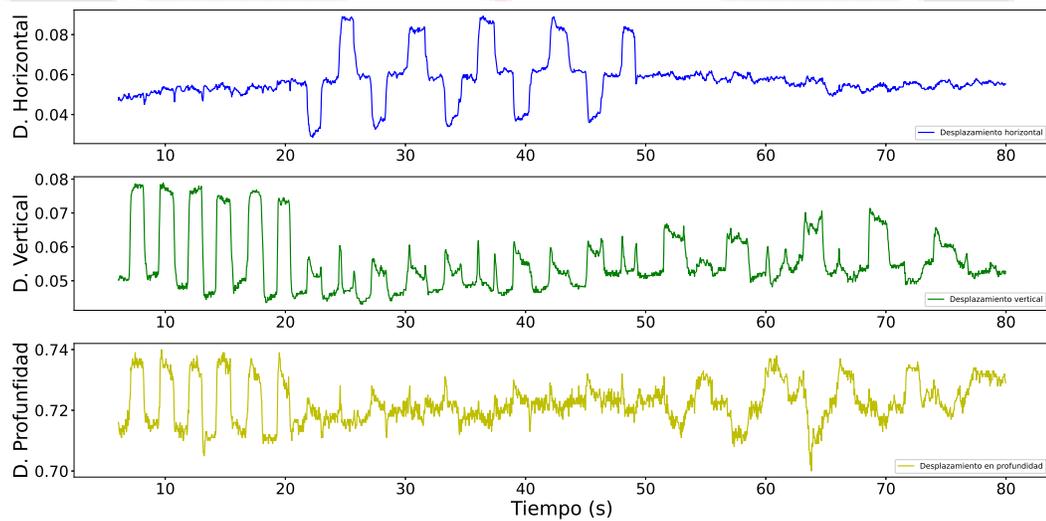


Figura 1.5: Posición Tridimensional del Marcador 2

Por cada uno de los 72 registros se cuenta con un archivo en formato *csv* (valores separados por comas, por sus siglas en inglés) que contiene la información de la posición tridimensional de ambos marcadores en los instantes de tiempo grabados. La duración de los registros es variable, ya que cada uno de los individuos realiza los movimientos con distintas velocidades. En promedio la duración de los registros es de 68 segundos, con una desviación estándar de 10.2 segundos. Los registros fueron adquiridos durante el segundo semestre de 2023.



TESIS TESIS TESIS TESIS TESIS

Capítulo 2

Marco Teórico

2.1. Visión Computacional y Seguimiento de Movimientos

Un sistema de visión computacional permite a una computadora analizar e interpretar el mundo visual usando algoritmos y técnicas para extraer información significativa de imágenes o transmisiones de video [22]. En primera instancia, un sistema hace una adquisición de imágenes utilizando cámaras u otros sensores. Luego se aplican técnicas de pre-procesamiento como reducción de ruido para mejorar los datos de entrada [23].

Además de capturar los datos, el sistema analiza las imágenes adquiridas para identificar patrones, objetos y relaciones. Este paso requiere de la implementación de una variedad de tareas como segmentación, extracción de características y detección de objetos. Esto con el objetivo de localizar y clasificar objetos dentro de la imagen. Finalmente, el sistema asigna significado a los datos visuales analizados haciendo uso de algoritmos de reconocimiento de objetos o de clasificación de imágenes. Para ello generalmente se utilizan técnicas de aprendizaje computacional [24, 23].

Los sistemas de visión computacional tienen un campo extenso de aplicaciones, abarcando muchas áreas del conocimiento como la manufactura, meteorología, robótica o las ciencias de la salud. Esta última es de particular importancia en este trabajo. En la literatura existen casos que describen el uso de sistemas de visión con el objetivo de capturar información del movimiento de la mandíbula [8, 25]. Asimismo, estudios han mostrado resultados positivos en la confiabilidad de las mediciones de movimientos mandibulares realizadas con sistemas de visión [26, 27].

Los datos utilizados en este trabajo de investigación tienen su origen en un sistema

de visión computacional similar al utilizado en [8]. Es decir, fueron adquiridos con un sistema de fotometría capaz de capturar información tridimensional de los objetos. Utilizando librerías de OpenCV, información de la posición de marcadores en el rostro de pacientes fue obtenida de las imágenes originales. Para obtener más detalles revise la Sección 1.6 de la Introducción.

2.2. Aprendizaje Supervisado

Se le llama aprendizaje computacional a las técnicas que permiten a una computadora mejorar el desempeño de un sistema o tarea aprendiendo de la experiencia (datos). Estas técnicas nos permiten construir, usando métodos computacionales, modelos que identifiquen patrones en los datos y permitan hacer predicciones [2, 4]. El aprendizaje computacional tiene varias categorías, como se denota en el esquema de la Figura 2.1.

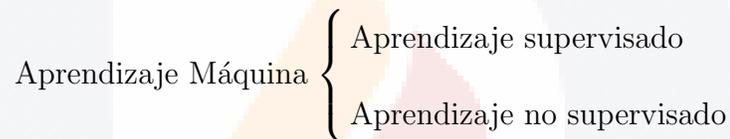


Figura 2.1: Categorías del Aprendizaje Máquina

Las características del aprendizaje supervisado son las siguientes:

1. Los datos disponibles se encuentran etiquetados.
2. Los algoritmos de aprendizaje buscan generar un modelo para predecir o catalogar instancias correctamente.

Existen dos subcategorías de aprendizaje supervisado: regresión, donde la variable a predecir es numérica continua; y clasificación, donde la variable a predecir es categórica. En otras palabras, un modelo de regresión busca ajustar un modelo que siga el patrón de los datos de entrenamiento, mientras que un modelo de clasificación busca patrones en los datos que permitan encontrar un límite de separación entre las categorías en los datos. [2, 28] En la Figura 2.2 se ejemplifican las diferencias entre un modelo de clasificación y regresión.

Como se mencionó anteriormente, el objetivo del aprendizaje supervisado es generar modelos que puedan hacer predicciones. Estos modelos tienen una gran variedad de aplicaciones en campos como:

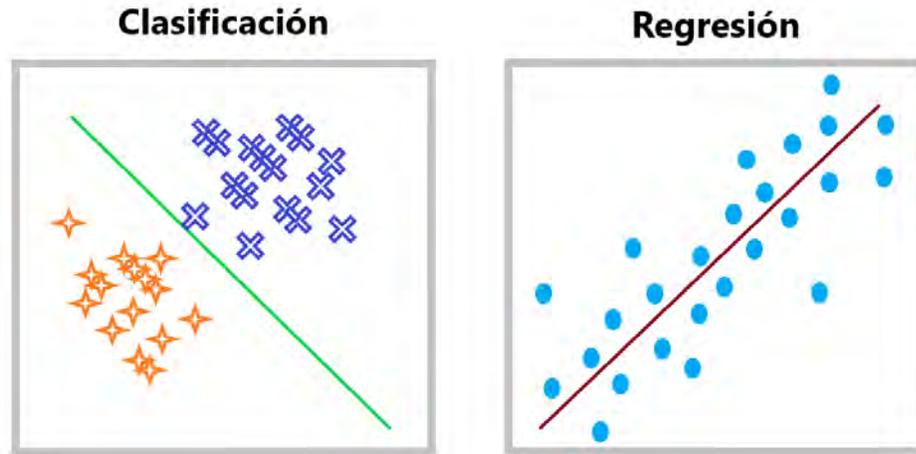


Figura 2.2: Diferencias entre Clasificación y Regresión. Elaboración propia.

1. Visión computacional
2. Procesamiento del lenguaje natural
3. Minería de Datos.

Al día de hoy el aprendizaje supervisado ha permeado a muchos ámbitos de la vida diaria, donde se usa en predicciones climáticas, monitoreo ambiental, análisis de datos de venta y clientes, optimizaciones de inventario, etc. [2].

Existen gran cantidad de métodos para generar modelos de aprendizaje supervisado, tanto de regresión como de clasificación. Entre los más conocidos y usados se encuentran el Clasificador ingenuo de Bayes [29], las Máquinas de Soporte vectorial[3], Regresión Logística [30], Redes Neuronales [5], Árboles de decisión[31], entre otros.

Dentro del aprendizaje supervisado existe una rama que recientemente ha tomado relevancia en el área de aprendizaje máquina. El deep learning engloba técnicas de clasificación y regresión que se enfocan en el uso de redes neuronales de gran tamaño para entrenar modelos. Este tipo de técnicas muestran buenos resultados en tareas como la clasificación de imágenes [32, 33].

A continuación se describirán los métodos de aprendizaje supervisado aplicados en la solución de el problema de investigación planteado. Los métodos encontrados en esta sección son tres: Máquina de Soporte Vectorial, K-Vecinos más Cercanos y Perceptrón Multicapa (Red Neuronal). Otros métodos de aprendizaje considerados en este trabajo se describirán en secciones posteriores.

2.2.1. Máquina de Soporte Vectorial

Una máquina de soporte vectorial es un algoritmo de aprendizaje supervisado usado para tareas de clasificación y regresión. El objetivo del algoritmo es encontrar un hiperplano que separe de la mejor manera las instancias de dos clases distintas [2, 3]. En la Figura 2.3 se muestran ejemplos de diferentes hiperplanos (En este caso, rectas) posibles para un conjunto de entrenamiento de 2 atributos con dos clases.

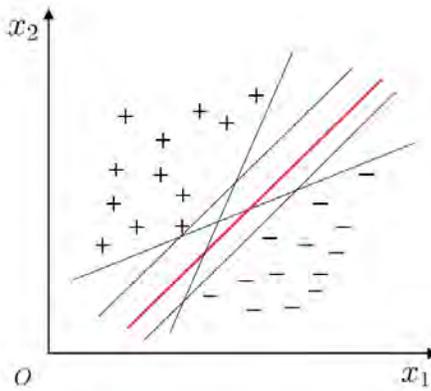


Figura 2.3: Hiperplanos separando dos clases, recuperado de: [2]

El hiperplano que mejor separa las clases es aquel que maximice el margen entre el hiperplano y los vectores de soporte. Los vectores de soporte son instancias de entrenamiento más cercanas al hiperplano [2]. En la Figura 2.4 se muestra un ejemplo de vectores de soporte.

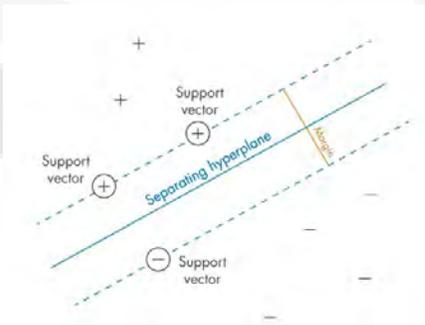


Figura 2.4: Vectores de soporte y el margen de un hiperplano. Recuperado de [3]

Entonces, el margen de un hiperplano se define como la región entre los vectores de soporte donde no hay más puntos de datos. Según [2], sea $w^T x_i + b = 0$ un hiperplano que clasifica correctamente las instancias de entrenamiento tal como muestra la Ecuación

(2.1), donde w es el vector normal w_1, \dots, w_d que determina la dirección de hiperplano, y b es el sesgo que controla la distancian ente el hiperplano y el origen.

$$\begin{aligned} w^T x_i + b &\geq 1, y_i = 1 \\ w^T x_i + b &\leq -1, y_i = -1 \end{aligned} \tag{2.1}$$

Donde y_i es la etiqueta de la i -ésima instancia de entrenamiento. Entonces el margen del hiper plano es como se muestra en la Ecuación (2.2):

$$\gamma = \frac{2}{\|w\|} \tag{2.2}$$

En la Figura 2.5 se muestra de manera gráfica los conceptos de hiperplano, vector de soporte y margen del hiper plano.

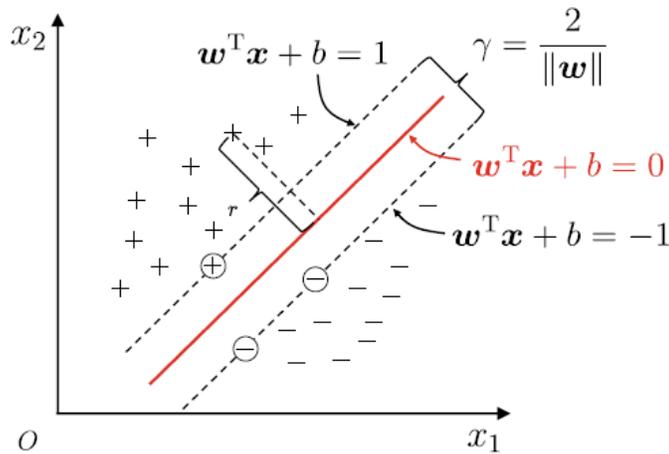


Figura 2.5: Margen de un hipeplano. Recuperado de [2]

Entonces, de acuerdo con [2], el hiperplano con el margen máximo que tenga w y b tal que maximicen γ , como se muestra a continuación en la Ecuación (2.3):

$$\text{máx}_{w,b} \frac{2}{\|w\|} \tag{2.3}$$

Visto de otra manera si el margen óptimo se obtiene maximizando $\|w\|^{-1}$, es equivalente a minimizar $\|w\|^2$, tal como nos muestra la Ecuación (2.4):

$$\text{mín}_{w,b} \frac{1}{2} \|w\|^2 \tag{2.4}$$

Truco del Kernel

Una desventaja de las máquinas de soporte vectorial es que sólo pueden encontrar un hiper plano para clases que son linealmente separables. Cuando se tenga un conjunto de datos que no es separable linealmente se puede aplicar una función Kernel que transforme los datos a un espacio dimensional superior al original, en el que las clases sean linealmente separables. A estos procedimientos se les conoce como el truco del Kernel [5, 3]. En la Figura 2.6 se muestra un ejemplo de problema de clasificación que no se puede solucionar linealmente hasta aplicar una función Kernel.

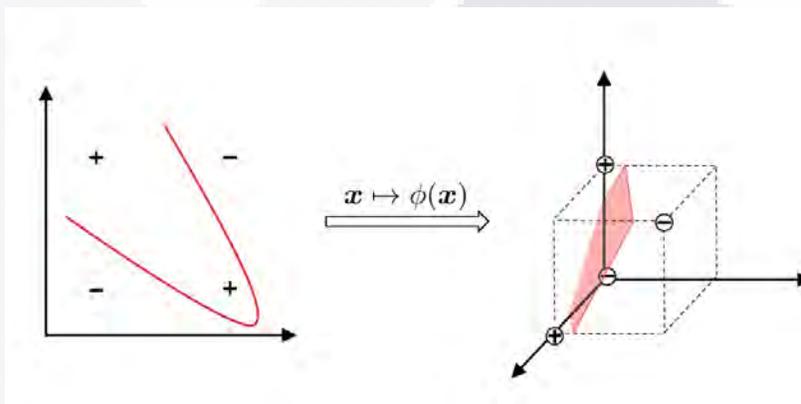


Figura 2.6: Ejemplo de transformación usando funciones kernel, recuperado de [2]

En la Tabla 2.1 se muestran algunos de los kernel más comunes usados en SVM

Tabla 2.1: Kernels comúnmente usados en máquinas de soporte vectorial

Nombre	Expresión	Parámetros
kernel Lineal	$k(x_i, x_j) = x_i^T x_j$	N/A
kernel Gaussiano	$k(x_i, x_j) = \exp -\frac{ x_i - x_j ^2}{2\sigma^2}$	$\sigma > 0$ anchura del kernel
kernel polinomial	$k(x_i, x_j) = (x_i^T x_j)^d$	$d > 1$ el grado del polinomio

2.2.2. K-Vecinos más Cercanos

El K-vecinos más cercanos (KNN) es un algoritmo de aprendizaje supervisado sencillo e intuitivo que se utiliza para tareas de clasificación. Su funcionamiento se basa en las similitudes que hay entre las instancias de una clase. A diferencia de otros clasificadores no es necesario entrenar propiamente el modelo, para comenzar a hacer predicciones solo se requiere de un conjunto de datos etiquetados descritos por un vector de atributos.

Para clasificar un punto de datos nuevo, se obtiene una medida de similitud de esta instancia con los datos etiquetados y se toman como referencia las k instancias etiquetadas con mayor similitud a la instancia a clasificar. La cantidad k de instancias es un parámetro del modelo, debe ser un número entero. Entonces, la instancia nueva se clasifica de acuerdo a un voto mayoritario de las clases de las k instancias más cercanas [4, 5].

La Figura 2.7 muestra un ejemplo de lo descrito anteriormente. Dado un conjunto de entrenamiento descrito por dos variables x y y con dos clases distintas, un punto nuevo a clasificar con 2 de sus 3 vecinos más cercanos miembros de la clase positiva y el restante de la negativa debe ser clasificado como positivo por voto mayoritario.

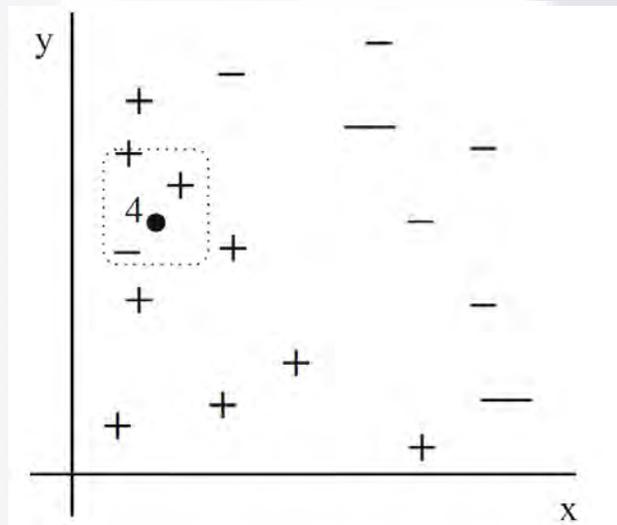


Figura 2.7: Ejemplo del funcionamiento de un clasificador K-vecinos más cercanos, recuperado de [4]

Existen diversas métricas de similitud que pueden ser usadas para obtener los k vecinos más cercanos. Las más comunes incluyen la distancia euclidiana, la distancia de Manhattan o la distancia de Hamming. La selección de la métrica correcta dependerá de los atributos del conjunto de entrenamiento. Por ejemplo, para atributos continuos, la distancia euclidiana es una buena opción, mientras que la distancia de Hamming es más adecuada para atributos discretos.

Por ejemplo, la distancia de euclidiana para dos instancias de datos $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ con las características de los datos de la Figura 2.7 sería $d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Generalizando para n atributos, el proceso queda como se muestra en la Ecuación (2.5).

$$d(p_1, p_2) = \sqrt{\sum_{i=1}^n (p_1^i - p_2^i)^2} \quad (2.5)$$

Donde $d(p_1, p_2)$ es la distancia euclidiana entre dos puntos p_1 y p_2 , n el número de atributos y p_1^i es el i -ésimo atributo en p_1 .

El valor del parámetro k en KNN se selecciona generalmente basado en conocimiento del dominio. Un valor menor de k conduce a límites de decisión más complejos, mientras que un valor mayor de k conduce a límites de decisión más suaves. En casos de clasificación binaria k debería ser un número impar para evitar empates [5].

2.2.3. Perceptrón Multicapa (Red Neuronal)

Un perceptrón es un modelo matemático de una neurona, que busca imitar el funcionamiento del cerebro, en particular el procesamiento paralelo, y puede ser usado para clasificación [2]. Este modelo fue propuesto por McCulloch y Pitts en 1943 [5], y es similar al mostrado en la Figura 2.8.

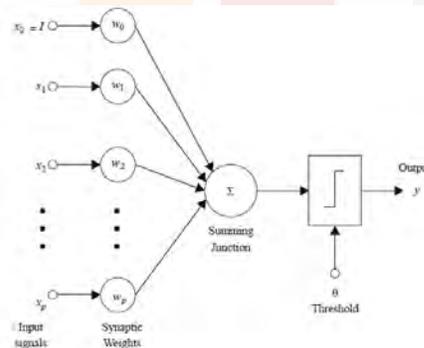


Figura 2.8: Modelo de McCulloch y Pitts [5]

Un perceptrón tiene 2 capas de neuronas (nodos), la capa de entrada y una de salida con las siguientes características [2]:

1. La capa de entrada recibe información externa y la transmite.
2. La capa de salida recibe información y genera una salida a partir de una función de activación.

Cada neurona recibe información de otras neuronas y su salida es una suma ponderada de las entradas [5] de manera que para un conjunto de entrenamiento x con n atributos la suma y es como se muestra en la Ecuación (2.6):

$$y = \sum_{i=1}^n \omega_i x_i + \omega_0 \quad (2.6)$$

Donde ω_i es el peso asignado a cada entrada, x_i es el valor de la entrada y ω_0 es un término de sesgo que hace el modelo más general.

En la Figura 2.9 se observa un diagrama del modelo del perceptrón descrito hasta ahora.

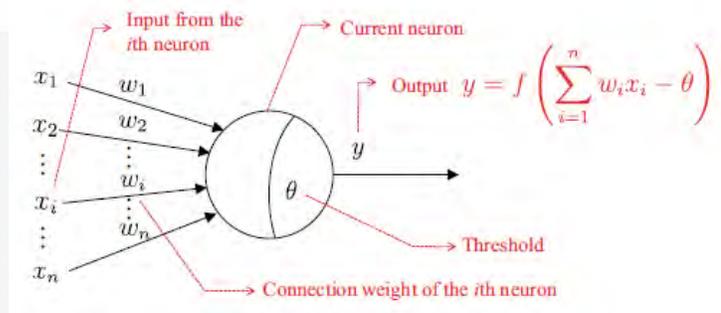


Figura 2.9: Modelo de perceptrón [2]

La información recibida por una neurona puede activarla o no, para tomar la decisión se usa una función de activación f . En general estas funciones son funciones sigmoides como la función logística. Cuando la suma ponderada de las entrada sobrepasa un umbral [5] θ la neurona se activa y genera una salida, como se muestra en la Ecuación (2.7):

$$salida = f\left(\sum_{i=1}^n \omega_i x_i + \omega_0\right) \quad (2.7)$$

Este tipo de algoritmos se entrenan siguiendo un algoritmo llamado "Backpropagation" [5]. En este algoritmo consta de las siguientes partes:

1. Los pesos ω se asignan aleatoriamente.
2. Se clasifican las instancias disponibles.
3. El error se calcula utilizando las etiquetas
4. Los pesos se recalculan
5. se repiten los pasos 2 al 4 hasta cumplir un criterio de paro.

Los pesos recalculados una vez terminada la clasificación se obtienen de acuerdo con la Ecuación (2.8):

$$\omega_j^{(k+1)} = \omega_j^{(k)} + \eta(y_i - \hat{y}_i^{(k)})x_{ij} \tag{2.8}$$

Donde $\omega_j^{(k+1)}$ es la siguiente generación de pesos, $\hat{y}_i^{(k)}$ es la estimación del clasificador en la generación actual, y_i es la etiqueta, x_{ij} es el valor del atributo j en la instancia i y η es el learning rate.

El learning rate es un valor entre 0 y 1 usado para controlar el ajuste de los pesos en cada iteración. Entre más alto sea η el peso será mas sensible de ajustes en la iteración actual y resultará en una convergencia más rápida, pero también habrá más riesgos de divergencias. [5]. La Figura 2.10 muestra lo anterior de manera gráfica.

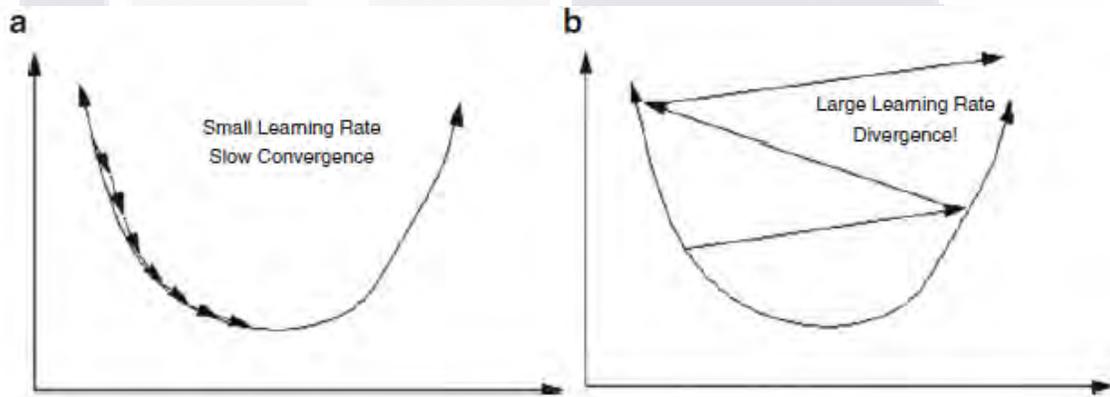


Figura 2.10: Convergencia del error según el valor de η [5]

Un perceptrón es una red neuronal muy sencilla, que consta de solo 2 capas, a diferencia de otras redes neuronales que cuentan con un número mayor de capas entre la de entrada y salida llamadas capas escondidas [2]. Tal como se muestra en la Figura 2.11.

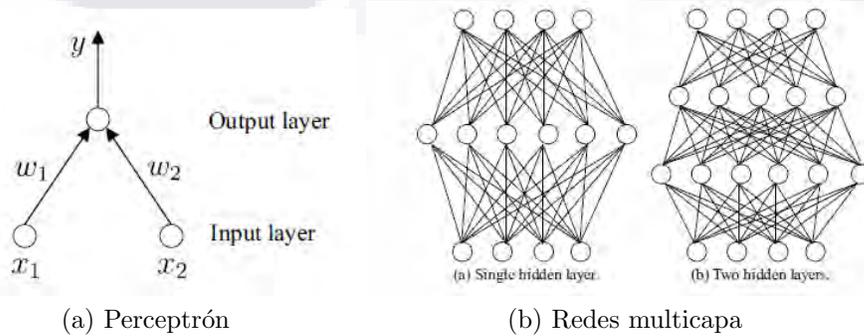


Figura 2.11: Comparación de redes neuronales [5]

2.3. Clasificación Probabilística

2.3.1. Clasificador Ingenuo de Bayes

Un clasificador probabilístico comúnmente usado es el clasificador ingenuo de Bayes. Este método se basa en la probabilidad condicional y el teorema de Bayes. En términos simples, el clasificador calcula y compara la probabilidad de una instancia de pertenecer a alguna de las clases, asignando la clase con la probabilidad mayor [2, 4]. La probabilidad se calcula de acuerdo con la Ecuación (2.9):

$$p(c|x) = \frac{p(c)p(x|c)}{p(x)} \quad (2.9)$$

Donde $p(c|x)$ es la probabilidad de una instancia de pertenecer a la clase c dados los valores x de sus atributos, $p(c)$ es la probabilidad a priori de la clase c , $p(x|c)$ es la probabilidad de tener los valores x en los atributos entre las instancias de la clase c . Finalmente $p(x)$ es la probabilidad de obtener los valores x en los atributos.

El clasificador Ingenuo de Bayes toma su nombre de la presunción "ingenua" de independencia condicional en los atributos. Dicho de otra manera, este método asume que entre las instancias de una clase, los atributos no están relacionados entre sí [4]. Este es uno de los métodos de clasificación más antiguos, y aunque en diversas aplicaciones provee de buenos resultados, excluir del modelo información de asociación entre variables puede limitar el potencial del clasificador.

Una forma de incluir información de asociación en un clasificador probabilístico como el ingenuo de Bayes, es usando funciones cópula. En las siguientes secciones se describirá a detalle lo que es una función cópula y cómo se puede aplicar en clasificadores probabilísticos.

2.3.2. Funciones Cópula

Las funciones de cópula fueron introducidas por primera vez por Sklar [34] como una forma de diferenciar el efecto de las dependencias del de las distribuciones marginales en una función de distribución conjunta. Al utilizar funciones de cópula, es posible modelar dependencias lineales y no lineales entre atributos. Estas asociaciones proporcionan nueva información, útil a la hora de abordar tareas de clasificación. La Definición 1 y el Teorema 1 describen el concepto y características de una cópula.

Definición 1 Una cópula C es una función de distribución conjunta de variables uniformes U_1, U_2, \dots, U_d .

$$C(u_1, \dots, u_d) = P(U_1 \leq u_1, \dots, U_d \leq u_d)$$

donde cada variable sigue una distribución uniforme estándar $U_i \sim \text{Uniforme}(0, 1)$ para $i = 1, 2, \dots, d$

Teorema 1 (Teorema de Sklar) Sea F una función de distribución d -dimensional con distribuciones marginales F_1, F_2, \dots, F_d . entonces hay una cópula C tal que para todo x en \mathbb{R}^d ,

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d))$$

Donde \mathbb{R} denota la línea real extendida $[-\infty, \infty]$. Si $F_1(x_1), F_2(x_2), \dots, F_d(x_d)$ son continuas entonces C es única. De lo contrario C está unívocamente determinada en $\text{Ran}(F_1) \times \text{Ran}(F_2) \times \dots \times \text{Ran}(F_d)$

Como consecuencia del teorema de Sklar, como se muestra en la Ecuación (2.10) cualquier densidad conjunta d -dimensional f y sus densidades marginales f_1, f_2, \dots, f_d también están relacionados y se pueden representar como:

$$f(x_1, x_2, \dots, x_d) = c(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \cdot \prod_{i=1}^d f_i(x_i) \quad (2.10)$$

donde c es la densidad de la cópula C , X_i es una variable aleatoria continua, $F_i(x_i)$ es la función de distribución marginal de X_i y $f_i(x_i)$ es la densidad marginal de X_i . La Ecuación (2.10) muestra que la estructura de dependencia está modelada mediante una función de cópula y que las densidades marginales pueden tener diferentes distribuciones.

Entonces, una función cópula es una función de distribución conjunta cuyas marginales son distribuciones uniformes. Su sustento teórico está dado por el teorema de Sklar que garantiza la existencia de una cópula para variables continuas y sus propiedades se extienden también a las funciones de densidad conjunta.

Las funciones cópula nos proveen de flexibilidad al modelar dependencias entre atributos al hacer posible separar los efectos propios de la asociación entre variables de los de las marginales de las variables [35, 36]. Este tipo de modelado no es posible para otras funciones de distribución conjunta que no son cópulas, como la normal multivariada, la cual forzosamente modela marginales que también son normales [37].

Funciones Cópula Bivariadas

Las funciones cópula bivariadas modelan dependencias entre dos variables. Partiendo de la Definición 1, el Teorema 1 y la Ecuación (2.10) la función cópula bivariada C se define como en la Ecuación (2.11), U_1, U_2 son variables con distribución uniforme:

$$C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2) \tag{2.11}$$

Asimismo, la función de densidad c asociada a una función cópula bivariada está definida en la Ecuación 2.12. Donde X_i es una variable aleatoria continua, $F_i(x_i)$ es la función de distribución marginal de X_i y $f_i(x_i)$ es la densidad marginal de X_i :

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \cdot \prod_{i=1}^2 f_i(x_i) \tag{2.12}$$

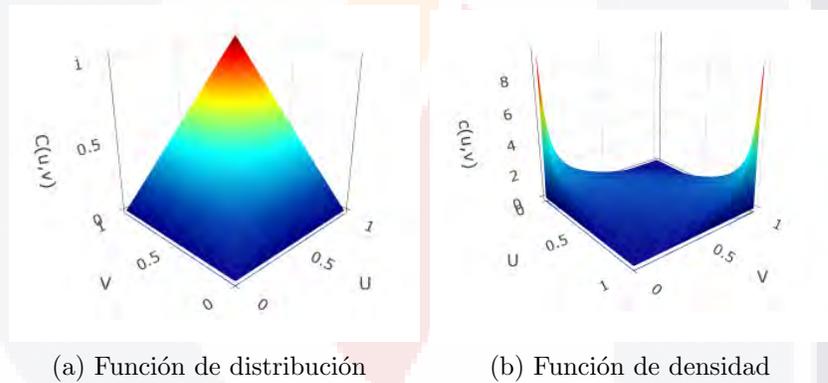


Figura 2.12: Funciones asociadas a una cópula Gaussiana bivariada. Caso Cópula Frank. Obtenida con [6]

Existe un gran número de funciones cópula bivariadas. Cada una modela distintos tipos de asociación entre variables y distintos rangos de fuerza de asociación. La fuerza de asociación en funciones cópula bivariadas se mide por el parámetro θ . Este parámetro para el caso de las funciones cópula bivariadas está relacionado con la Tau de Kendall tal como se muestra en la Ecuación (2.13) [38]. θ puede ser estimado usando el método de máxima verosimilitud y la estimación de la Tau de Kendall [39].

$$\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \tag{2.13}$$

Las Tablas 2.2 y 2.3 muestran algunos ejemplos de las funciones de distribución y densidad, respectivamente, de las funciones cópula bivariadas AMH (Ali-Mikhail-

Haq), Clayton, FGM (Farlie-Gumbel-Morgenstern), Frank, Gaussiana y Gumbel, que son comunmente usadas. La Tabla 2.2 muestra también el rango de los parámetros de dependencia de cada cópula.

Copula	Función de distribución
AMH	$C(u_1, u_2) = \frac{u_1 u_2}{1 - \theta(1-u_1)(1-u_2)}$; $\theta \in [-1, 1)$
Clayton	$C(u_1, u_2) = \max(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, 0$; $\theta \in [-1, \infty)/0$
FGM	$C(u_1, u_2) = u_1 u_2 (1 + \theta(1-u_1)(1-u_2))$; $\theta \in [-1, 1]$
Frank	$C(u_1, u_2) = -\frac{1}{\theta} \ln(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1})$; $\theta \in (-\infty, \infty)/0$
Gaussian	$C(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{e^{(-1/2)t'\Sigma^{-1}t}}{2\pi \Sigma ^{1/2}} dt_1 dt_2$; $\theta \in (-1, 1)$ Donde Σ es una matriz de correlación con $\Sigma_{12} = \theta$
Gumbel	$C(u_1, u_2) = \exp(-(m_1^\theta + m_2^\theta)^{1/\theta})$; $\theta \in [1, \infty]$ Donde $m_1 = -\ln(u_1)$ y $m_2 = -\ln(u_2)$

Tabla 2.2: Funciones de distribución de cópulas bivariadas

La Figura 2.13 muestra la estructura de dependencia para cada una de las cópulas mencionadas en las Tablas 2.2 y 2.3, con un parámetro θ dentro del intervalo especificado para cada cópula.

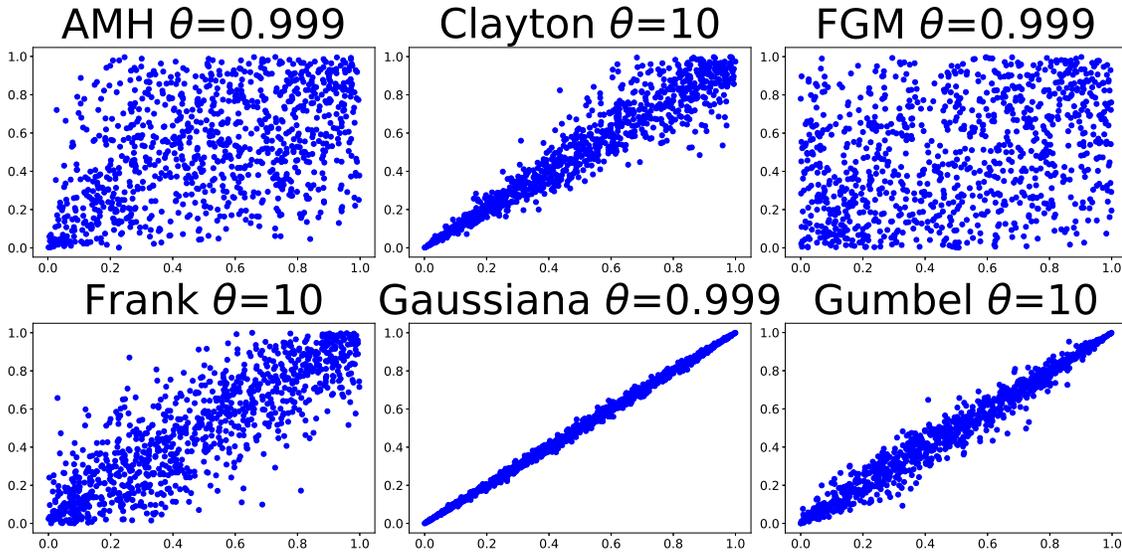


Figura 2.13: Estructura de dependencia para diferentes cópulas bivariadas

Copula	Función de densidad
AMH	$c(u_1, u_2) = \frac{1+\theta(u_1+u_2+u_1u_2-2)-\theta^2(u_1+u_2+u_1u_2-1)}{(1-\theta(1-u_1)(1-u_2))^3}$
Clayton	$c(u_1, u_2) = (1 + \theta)(u_1u_2)^{-(\theta+1)}(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{2\theta+1}{\theta}}$
FGM	$c(u_1, u_2) = 1 + \theta(1 - 2u_1)(1 - 2u_2)$
Frank	$c(u_1, u_2) = \frac{-\theta(e^{-\theta}-1)e^{-\theta(u_1+u_2)}}{((e^{-\theta u_1}-1)(e^{-\theta u_2}-1)+(e^{-\theta}-1))^2}$
Gaussian	$c(u_1, u_2) = (1 - \theta^2)^{-\frac{1}{2}} \exp\left(\frac{n_1^2+n_2^2-2\theta n_1 n_2}{2(1-\theta^2)} - \frac{n_1^2+n_2^2}{2}\right)$ donde $n_1 = \Phi^{-1}(u_1)$ y $n_2 = \Phi^{-1}(u_2)$
Gumbel	$c(u_1, u_2) = \frac{\exp(-(m_1^\theta+m_2^\theta)^{-1/\theta})(m_1m_2)^{\theta-1}}{u_1u_2(m_1^\theta+m_2^\theta)^{2-1/\theta}}((m_1^\theta + m_2^\theta)^{1/\theta} + \theta - 1)$ Donde $m_1 = -\ln(u_1)$ y $m_2 = -\ln(u_2)$

Tabla 2.3: Funciones de densidad de cópula bivariadas

2.3.3. Clasificador Probabilístico con Cópulas

El clasificador ingenuo de Bayes, como el de la Ecuación (2.14) asume independencia condicional entre las características [40]. Por ejemplo, la probabilidad de que una instancia descrita por d características pertenezca a una clase A es:

$$P(A|x_1, \dots, x_d) = \frac{P(A) \times \prod_{i=1}^d f(x_i|A)}{f(x_1, \dots, x_d)} \quad (2.14)$$

Si se considera la asociación entre atributos usando cópulas, el clasificador probabilístico quedaría como se muestra en la Ecuación (2.15)

$$P(A|x_1, \dots, x_d) = \frac{c(F_1(x_1|A), \dots, F_d(x_d|A)) \times P(A) \times \prod_{i=1}^d f_i(x_i|A)}{f(x_1, \dots, x_d)} \quad (2.15)$$

2.4. Estado del Arte

El análisis de la cinemática mandibular utilizando sistemas de captura de movimientos tiene una larga trayectoria. Sistemas de visión computacional utilizando cámaras infrarrojas, como los realizados en [27] y [41] son utilizados para grabar movimientos mandibulares y analizar los datos con objetivos de generar modelos que puedan ser usados como apoyo en el diagnóstico de enfermedades relacionadas con el mal funcio-

namiento de la mandíbula.

Otros estudios como el de [42], el de [43] y el de [26] muestran que el uso de sistemas de captura de movimientos mandibulares presentan un considerable grado de confiabilidad en sus mediciones, respecto al movimiento real de la mandíbula. Por lo que podemos decir que el uso de este tipo de sistemas provee de datos veraces para procesamientos y predicciones posteriores.

Existen trabajos relacionados con el procesamiento de datos obtenidos del seguimiento de movimientos mandibulares, datos de imágenes mandibulares, o señales acústicas relacionadas al funcionamiento de la mandíbula. Estos trabajos tienen el objetivo ya sea de hacer una clasificación o plantear una metodología para análisis y extracción de atributos de los datos. Ejemplos de ello se describen en los siguientes párrafos.

Fuentes et al. [44] hace la determinación de protocolos para obtención de imágenes de los movimientos mandibulares. Definieron y desarrollaron cuatro protocolos para analizar los movimientos de apertura y cierre, así como movimientos de borde en los planos frontal, sagital y horizontal. Para obtener la información utilizaron un articulograma electromagnético o EMA. Para la recolección de información fueron necesarios 16 sensores optoeléctricos para poder medir el movimiento funcional y el límite de estos. La ejecución de los movimientos fue de apertura y cierre, lateralidades, protrusión y retrusión. Los autores no mencionan cuál protocolo da mejores resultados determinación del movimiento funcional y el límite.

Calil et al. [25] hacen una propuesta de características para el análisis de movimientos mandibulares obtenidos por medio de cámaras infrarrojas de una muestra de 40 personas que divide en grupos según su condición (sano, TTM mialgia y TTM artroalgia). Obtienen 7 características para los movimientos de apertura y cierre, 4 para los movimientos laterales y 5 para los movimientos de protrusión y retrusión. Con la información obtenida de sus datos hacen una clasificación de pacientes sanos y con TTM (subclasificación en mialgia y artroalgia) usando los métodos de K-nearest neighbor, máquina de soporte vectorial, random forest y Bayes ingenuo. De los resultados, el clasificador que mejor se desempeñó fue el K-nearest neighbor con una precisión por encima de 90

Por otro lado, en 2021, Perpetuini et al. [45], utilizan información obtenida por electromiograma e imágenes térmicas de infrarrojo buscando encontrar la relación entre la circulación superficial y el aumento de temperatura con la actividad muscular de los músculos temporal y masseter de la mandíbula. Basando en que el músculo temporal presenta más actividad en casos de TTM. Con esta información utilizan el algoritmo

de modelado general lineal para hacer una regresión y estimar la relación entre sus variables. Posteriormente con su modelo hacen una clasificación de pacientes sanos y con TTM. Finalmente utilizan análisis ROC para medir el desempeño de su clasificador.

Kakkeri & Bormane [46] hacen un análisis de datos de electromiograma (EMG) de 81 personas, de los cuales 41 son grupo de control (Sin TTM). Midieron señales de los músculo masseter y temporalis cuando los pacientes hacían movimientos de apretar la mandíbula. Con estos datos se entrenaron algoritmos de clasificación dividiendo sus datos en 70 % para entrenamiento y 30 % datos de prueba. Los algoritmos utilizados fueron: máquina de soporte vectorial, regresión logística, K-nearest neighbor, Adaboost, y Bayes ingenuo. Siendo estos dos últimos los que tuvieron un mejor desempeño con 93 % y 92 % de exactitud respectivamente.

Otros estudios que analizan datos de electrocardiograma para tareas de aprendizaje son los de Sonmezocak [47] y el de Troka [48]. El primero busca clasificar pacientes con bruxismo a partir de información de músculos asociados al funcionamiento mandibular como el masetero. Algoritmos como Redes neuronales, Máquina de Soporte Vectorial y K-vecinos más cercanos fueron implementados con buenos resultados. El trabajo de Troka utiliza datos similares pero la tarea de aprendizaje en este caso es de agrupamiento de acuerdo a las respuestas musculares, es decir, aprendizaje no supervisado. Utilizando su enfoque pudieron encontrar similitudes en los datos de diferentes pacientes que sugieren, podrían usarse para proporcionar información de relevancia clínica.

Existen también trabajos que utilizan imágenes médicas como datos de entrenamiento para clasificadores, como el de [49]. En este trabajo se usan imágenes de tomografía de articulaciones temporomandibulares para entrenar diversos algoritmos de aprendizaje supervisado. Utilizando técnicas de segmentación y reducción de dimensionalidad en sus imágenes obtienen datos con los que entrenan clasificadores como el ingenuo de Bayes, K-vecinos más cercanos y árboles aleatorios.

Los autores Taşkıran & Çunkaş [50], proponen un sistema de clasificación de trastornos temporomandibulares utilizando procesamiento de sonido. Micrófonos colocados en el paciente graban los sonidos producidos por la articulación temporomandibular mientras el médico realiza una examinación del paciente. Posteriormente los datos son digitalizados y procesados utilizando técnicas de aprendizaje profundo y redes neuronales. Los resultados de su estudio mostraron que el método propuesto de clasificación más exitoso de los propuestos fue el de aprendizaje profundo con un 94.5 % de clasificación correcta en los casos analizados.

Otro ejemplo de clasificación de la disfunción temporomandibular lo encontramos

en [8], que hace un estudio de pacientes sanos y con TTM donde por medio de cámaras graba movimientos mandibulares de los cuales posteriormente obtiene información de marcadores que coloca en el rostro de los sujetos que fueron parte del estudio. Las etiquetas para la clasificación las obtiene aplicando los criterios DC/TMD a cada uno de los sujetos de estudio. Para la clasificación utiliza una máquina de soporte vectorial (SVM, por sus siglas en inglés), que arrojó resultados del 93 % para la sensibilidad y especificidad.

Aunque existe un amplio número de trabajos que abordan los trastornos temporo-mandibulares, particularmente tareas de clasificación en ninguno de los consultados se utilizan modelos de aprendizaje que hagan uso de funciones cópula.

Las funciones de cópula son funciones de distribución conjunta que proporcionan una forma más flexible de modelar la asociación entre características permitiendo separar la estructura de dependencia de las distribuciones marginales de las características[51]. Al ser capaces de modelar dependencias tanto lineales como no lineales, las funciones de cópula se han utilizado en campos como economía y finanzas [52], hidrología [53], clasificación supervisada [54], clasificación de señales cerebrales [55], entre otros.

Por ejemplo, [56] Propone un modelo multivariado utilizando funciones cópula para usarse en la detección de nodos defectuosos en redes de sensores inalámbricos. Por otro lado, en [57] utilizan modelos con funciones cópula como propuesta de índices de capacidad de proceso en el área de manufactura con el objetivo de medir el grado de variación que experimenta un proceso en relación con sus límites de especificación.

En el área de algoritmos evolutivos, trabajos como [39] y [58] utilizan las funciones cópula como modelos de distribución en algoritmos de estimación de la distribución, conocidos como EDAs. Estos algoritmos son puestos a prueba con la optimización de funciones como la Ackley y Griewangk.

Clasificación con datos sintéticos con el objetivo de comparar el desempeño de clasificadores tradicionales con los clasificadores que incluyen funciones cópula ha sido abordado en trabajos como [59] y [60]. Estos autores utilizaron modelos clasificadores, con varias funciones de cópula, con y sin modelos gráficos.

En el área de clasificación de imágenes, existen trabajos que utilizan las funciones cópula para entrenar modelos en tareas de clasificación como identificar pixeles de fondo y frente de una imagen [61], o clasificar áreas de tierra de acuerdo al uso que se les den [62] [63]. Otro trabajo similar hace clasificación de acuerdo al tipo de roca del que se compone la tierra, no al uso que se le da [64]. Para ello utilizan cópulas elípticas.

Capítulo 3

Metodología

En este capítulo se describe el proceso metodológico seguido en este proyecto de tesis para la clasificación de pacientes con trastornos temporomandibulares. El proceso se llevó a cabo en 5 etapas: preprocesamiento de las señales, Extracción de atributos, selección de atributos, implementación de los modelos de clasificación y por último la evaluación del desempeño de los clasificadores. En la Figura 3.1 se puede observar un diagrama que menciona las etapas seguidas. A lo largo del capítulo se describe a detalle cada una de ellas.

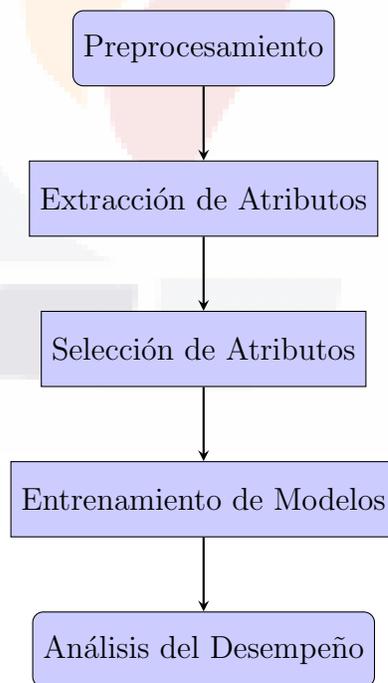


Figura 3.1: Diagrama de etapas en la metodología

3.1. Preprocesamiento de las Señales

Previo al entrenamiento de los modelos de clasificación, las señales han sido preprocesadas, siendo filtradas y divididas en segmentos correspondientes a tipos de movimientos mandibulares realizados. Posteriormente de cada segmento de señal se han obtenido atributos para caracterizar los movimientos del paciente.

3.1.1. Filtrado y Segmentación de Señales

El filtrado de las señales se realizó en dos fases: la fase de corrección de movimientos de cabeza, y la fase de filtrado pasa bajas. En los datos originales disponibles se tiene información de dos marcadores, el primero de ellos (M_{mov}) contiene la información tridimensional de los movimientos manibulares combinada con cualquier otro movimiento realizado por el paciente; el segundo, llamado marcador de control (M_c) contiene información de los movimientos que no corresponden a movimientos mandibulares y no son de interés.

En la primera fase del filtrado se recuperó la información de los movimientos mandibulares restando la información del marcador de control al marcador de los movimientos mandibulares, tal como se muestra en la ecuación (3.1)

$$M_{corregido} = M_{mov} - M_c \quad (3.1)$$

Donde $M_{corregido}$ es el marcador con información correspondiente solo a movimientos mandibulares, M_i es el marcador con la información combinada y M_c es el marcador de control.

Una vez corregidas las señales, fue aplicado un filtro gaussiano, que elimina componentes de frecuencia altas correspondientes a ruido de las señales. Este ruido es de tipo aleatorio, causado por variaciones en los valores de los píxeles en fotogramas consecutivos de los videos de los que se obtiene las señales. Es decir, variaciones en los píxeles de las imágenes ocasionan que la detección del centro de los marcadores fluctúe. Esto puede deberse a variaciones en las condiciones de iluminación, inestabilidad del sensor, entre otros.

Para la implementación de este filtro se utilizó la librería *scipy* de python, específicamente la función *gaussian_filter* con un parámetro $\sigma = 2$ correspondiente a la desviación estándar del kernel usado en la convolución con la señal.

A continuación, se segmentaron las señales filtradas, donde cada segmento corresponde a un tipo de movimiento mandibular realizado. En los datos disponibles se gra-

baron 3 tipos de movimientos y fueron realizados por los individuos en orden el : 1) Apertura/Cierre, 2) Laterales y 3) Protrusión/Retrusión. por lo que se tienen 3 segmentos en total. Para obtener los segmentos de las señales se hizo una detección de eventos de movimientos laterales en el eje x de las señales. Dado que el orden en que los participantes del estudio realizan los movimientos, al detectar el inicio y fin de los movimientos mandibulares se obtienen los dos cortes necesarios para seccionar las señales.

Para la detección de eventos se utilizaron las funciones *findpeaks* y *peakwidths* de *scipy* en python. Con la primera se obtienen los eventos de movimientos mandibulares, con la segunda el inicio y fin de cada movimiento. El primer corte se hace en el momento de tiempo en que ocurre el inicio del primer evento, el segundo corte se hace en el momento de tiempo en que ocurre el final del último evento.

Cada señal tiene un componente en x , y , y z , correspondiente al los desplazamientos en el eje horizontal, vertical y de profundidad, respectivamente, por lo que en total se obtuvieron 9 segmentos de señales. En la Figura 3.2 se muestra un ejemplo de las señales segmentadas.

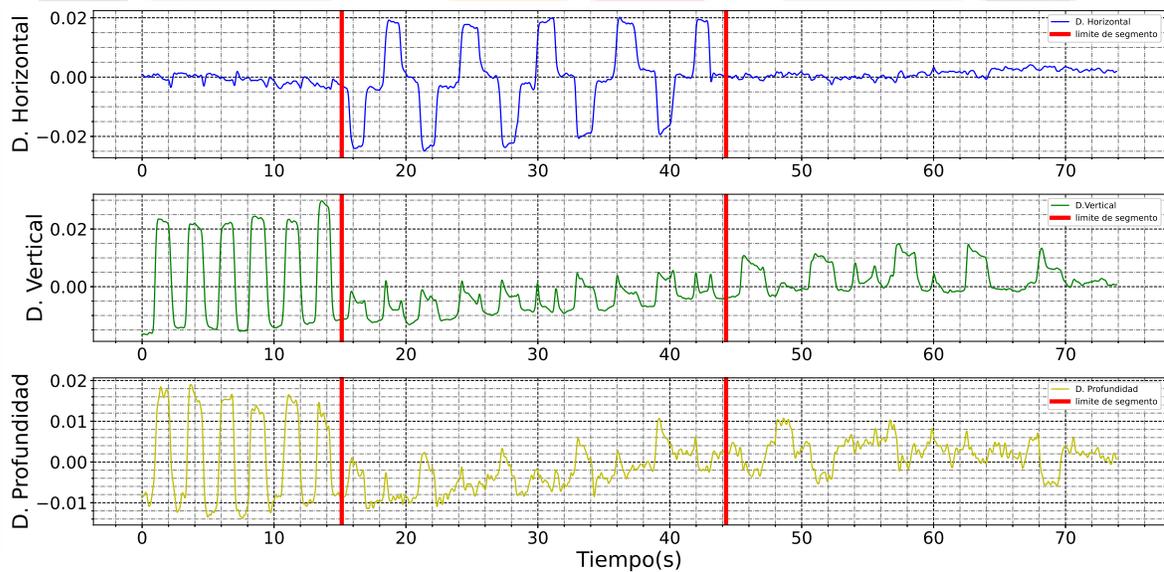


Figura 3.2: Segmentación de las señales corregidas y filtradas.

Finalmente, a cada uno de los segmentos obtenidos se les estima una línea base utilizando la librería *peakutils* de python y se les substraen dicha línea base. El algoritmo de *peakutils* realiza iterativamente un ajuste polinomial en los datos para detectar su línea de base. En cada iteración, los pesos del modelo se ajustan en las regiones con picos para identificar únicamente la línea base. Para este problema se optó por ajustar

un polinomio de grado 2.

3.1.2. Descarte por Frecuencia de Muestreo

Con la finalidad de eliminar datos atípicos, se descartan aquellas instancias de pacientes cuyos datos hayan sido muestreados a frecuencias distintas al resto. La variación en la frecuencia de muestreo en estos casos se debe principalmente a saltos temporales en los que el algoritmo de seguimiento de los marcadores no pudo localizar los marcadores en el rostro del individuo.

Para calcular la frecuencia de muestreo, en primer lugar se obtiene el periodo de muestreo promedio de la señal. Calculando el inverso del periodo promedio se obtiene la frecuencia promedio. Con este dato calculamos cuartiles y descartamos aquellas instancias por encima de $Q_3 + 1.5 \times IQR$ y por debajo de $Q_1 - 1.5 \times IQR$. IQR es el rango intercuartílico calculado $Q_3 - Q_1$.

3.2. Extracción de Atributos

Retomando de la Sección 3.1, se obtuvieron nueve segmentos en total de las señales originales, tres por cada movimiento mandibular, correspondientes a los desplazamientos en los ejes x , y , y z . De cada uno de estos segmentos se extrajeron una serie de atributos, catalogados en cuatro clases: biomecánicos; derivadas del desplazamiento; del dominio de la frecuencia; y de tiempo. Cada una de estas clases se describe a continuación.

3.2.1. Atributos Biomecánicos

Los atributos biomecánicos son mediciones de las capacidades del individuo para realizar los movimientos mandibulares, en este caso, fueron tomados de [25] donde se utilizan en una tarea de clasificación similar a la presentada en este trabajo. La Tabla 3.1 describe cada uno de estos atributos. De las cuatro clases de atributos obtenidos, este caso, es el único en el que se definen y extraen de manera distinta dependiendo del movimiento que describen.

Tabla 3.1: Atributos biomecánicos extraídos

Nombre	Descripción
OCY	Desplaz. máximo en el eje y durante los movimientos de apertura
OCX	Desplaz. máximo en el eje x durante los movimientos de apertura
OCZ	Desplaz. máximo en el eje z durante los movimientos de apertura
LLX	Desplaz. máximo a la izquierda (eje x) en los movimientos laterales.
LRX	Desplaz. máximo a la derecha (eje x) en los movimientos laterales.
LY	Desplaz. máximo en el eje y durante los movimientos laterales.
LZ	Desplaz. máximo en el eje z durante los movimientos laterales.
PY	Desplaz. máximo en el eje y durante los movimientos de protrusión
PX	Desplaz. máximo en el eje x durante los movimientos de protrusión
PZ	Desplaz. máximo en el eje z durante los movimientos de protrusión

3.2.2. Atributos Derivados del Desplazamiento

Los atributos derivados del desplazamiento se obtuvieron con derivadas numéricas. Las tres primeras derivadas del desplazamiento, correspondientes a la velocidad, aceleración y jerk, de los segmentos de señal fueron extraídas. De estas señales, se obtuvieron los atributos descritos en la Tabla 3.2, donde s representa el segmento del que se realice el análisis y t el tiempo.

Tabla 3.2: Atributos derivados del desplazamiento

Nombre	Descripción	Fórmula
\dot{s}	Velocidad	$\frac{\Delta s}{\Delta t}$
\dot{s}_{prom}	Velocidad promedio	$\frac{\sum \dot{s}}{ \dot{s} }$
$\sigma(\dot{s})$	Desviación estándar de la velocidad	$\sqrt{\frac{\sum (\dot{s} - \dot{s}_{prom})^2}{ \dot{s} }}$
\ddot{s}	Aceleración	$\frac{\Delta \dot{s}}{\Delta t}$
\ddot{s}_{prom}	Aceleración promedio	$\frac{\sum \ddot{s}}{ \ddot{s} }$
$\sigma(\ddot{s})$	Desviación estándar de la aceleración	$\sqrt{\frac{\sum (\ddot{s} - \ddot{s}_{prom})^2}{ \ddot{s} }}$
$\dot{\ddot{s}}$	Jerk	$\frac{\Delta \ddot{s}}{\Delta t}$
$\dot{\ddot{s}}_{prom}$	Jerk promedio	$\frac{\sum \dot{\ddot{s}}}{ \dot{\ddot{s}} }$
$\sigma(\dot{\ddot{s}})$	Desviación estándar del jerk	$\sqrt{\frac{\sum (\dot{\ddot{s}} - \dot{\ddot{s}}_{prom})^2}{ \dot{\ddot{s}} }}$

3.2.3. Atributos en el dominio de la frecuencia

Los atributos en el dominio de la frecuencia son aquellos que se obtienen de la transformada de Fourier de un segmento de señal en lugar del segmento original. La transformada de Fourier nos entrega la potencia de la señal en cada uno de sus componentes de frecuencia. En esta clase de atributos se consideraron 3 que se describen en la Tabla 3.3. En este contexto es importante resaltar que después del descarte por frecuencia de muestreo, todas las señales de la base de datos están muestreadas a 30 Hz, por lo que la frecuencia de Nyquist es de 15 Hz en todos los casos. P_i en la Tabla 3.3 hace referencia a la potencia del componente en frecuencia i en las señales.

Tabla 3.3: Atributos del dominio de la frecuencia

Nombre	Descripción	Fórmula
P_{tot}	Potencia total	$\sum P_i$
P_{prom}	Potencia promedio	$\frac{\sum P_i}{m}$
F_{ratio}	Relación de frecuencias	$\frac{\sum_{7.5}^{15} P_i}{\sum_0^{7.5} P_i}$

3.2.4. Atributos en el Tiempo

Cada uno de los movimientos realizados por los pacientes, ya sean de apertura, laterales, o de protrusión, tienen una duración determinada. Utilizando las funciones *find_peaks* y *peak_widths* del módulo *signal* de la librería *scipy* [65] se detectaron los eventos de cada movimiento. Con esto fue posible obtener los atributos en la Tabla 3.4.

Tabla 3.4: Atributos en el Tiempo

Nombre	Descripción
t_{prom}	Tiempo promedio en la realización de los movimientos mandibulares

3.2.5. Estandarización de Atributos

Una vez extraídos los atributos de las señales y formada una matriz de observaciones de n filas y m columnas, donde n es el número de instancias o pacientes y m es el número de atributos extraídos, se deben estandarizar los datos, de manera que los datos tengan el mismo tipo de contenido y de formato. Siguiendo la Ecuación (3.2) se estandarizan cada uno de los atributos, de manera que tengan una media $\bar{X} = 0$ y una desviación estándar $s = 1$.

$$a_j^{(i)} = \frac{x_j^{(i)} - \bar{X}_j}{s_j} \quad (3.2)$$

Donde A_j es el atributo estandarizado j , a_j^i es el valor en la posición i de A_j , \bar{X}_j es la media muestral del atributo j y s_j es la desviación estándar del atributo j . Esto para $i = 1, 2, 3, \dots, n$ y $j = 1, 2, 3, \dots, m$.

3.3. Selección de Atributos

De la Sección 3.2 obtuvimos una matriz de observaciones con n filas y m columnas, que representan el número de instancias y atributos respectivamente. El número de atributos hasta ahora es muy grande, y entre ellos podría haber algunos que no sean relevantes para la tarea de clasificación o sean redundantes con otros, por lo que es necesario un método de selección.

En [66] se menciona que es común el uso de estrategias como algoritmos evolutivos o métodos de envoltura (*Wrapper* en inglés) como métodos de selección de atributos en tareas de clasificación. En este trabajo se aplicó un método mixto usando un algoritmo evolutivo cuyo parámetro de aptitud es el desempeño de un modelo lineal de clasificación. El algoritmo evolutivo propuesto es uno de estimación de distribución (EDA, por sus siglas en inglés), en específico un algoritmo de distribuciones marginales univariado (UMDA por sus siglas en inglés). Por otro lado el modelo lineal propuesto es un Perceptrón. A continuación se describe el funcionamiento del método de selección propuesto.

3.3.1. Algoritmo UMDA

Un algoritmo de estimación de la distribución es una estrategia evolutiva que busca optimizar una función. A diferencia de otros algoritmos evolutivos los EDAs no requieren de estrategias como la mutación o la cruce. El algoritmo UMDA estima la distribución de probabilidad de una sola variable. En el caso del problema de selección de atributos, el UMDA estima una distribución Bernuolli para cada atributo disponible, que nos muestra la probabilidad de ser seleccionado[67].

En el UMDA se le llama individuo a una solución posible, que está codificada y en este caso representa si seleccionamos o no a un atributo. La Figura 3.3 muestra un ejemplo del funcionamiento del algoritmo de selección de atributos. El Algoritmo 1 muestra de manera más detallada el funcionamiento del mismo.

Cuando se cumple el criterio de paro, que en este caso es un número específico de poblaciones (100), se obtiene el mejor individuo de la última población y se convierte en la selección de atributos para el problema de clasificación.

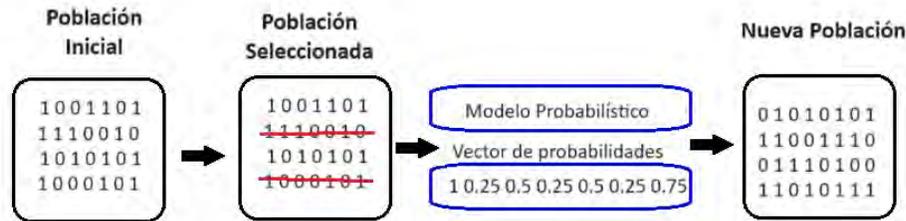


Figura 3.3: Funcionamiento del Algoritmo UMDA

Algorithm 1 Pseudocódigo para el UMDA

- 1: Asigna $t \leftarrow 0$
 - 2: Genera aleatoriamente una población inicial P_0 con n individuos.
 - 3: **while** No se cumpla criterio de paro **do**
 - 4: Genera un modelo y estima su aptitud
 - 5: Selecciona los mejores individuos en P_t , envía el mejor directamente a la siguiente población
 - 6: Calcula las probabilidades de seleccionar los atributos en los mejores individuos de P_t
 - 7: Estima distribuciones binomiales para cada atributo
 - 8: Genera $n - 1$ nuevos individuos usando las distribuciones estimadas.
 - 9: Asigna $t \leftarrow t + 1$
 - 10: **end while**
-

Otro criterio de selección de atributos en este trabajo utilizado de manera independiente al algoritmo EDA es la correlación de Pearson, que mide la fuerza de asociación lineal entre dos variables aleatorias, tomando valores de -1 a 1. La correlación de Pearson se utiliza para encontrar aquellos pares de variables que sean redundantes, es decir, aporten la misma información. Entonces, una correlación alta en términos absolutos, se toma como un indicador de redundancia entre variables aleatorias. En problemas de clasificación, se considera que utilizar dos variables que sean redundantes no ayuda en la diferenciación entre clases, por lo que utilizar variables redundantes no mejora el desempeño de clasificadores que desde su concepción no tomen en cuenta asociación entre variables. El Algoritmo 2 describe el proceso seguido para eliminar variables redundantes de la matriz de observaciones inicial obtenida de la extracción de características.

Algorithm 2 Pseudocódigo para selección con correlación de Pearson

```

1: while  $\max(cnt\_reds) > 1$  do
2:   Calcula matriz de correlaciones de Pearson  $M$  de  $n \times n$ , donde  $n$  representa el
   número de atributos.
3:   Calcula  $|M|$ 
4:   Obtener matriz de ceros de dimensiones  $n \times n$ 
5:   Calcula  $M_{aux}[M > 0.75] = 1$ 
6:   suma filas o columnas para obtener vector  $cnt\_reds$  tamaño  $n$ .
7:   Calcula indice  $k$  donde  $\max(cnt\_reds)$ 
8:   elimina el  $k$ -ésimo atributo de la matriz de observaciones.
9: end while

```

3.4. Entrenamiento de Clasificadores

Los modelos de clasificación tienen el objetivo de asignar a una instancia una de las clases para las que el modelo fue entrenado a distinguir. En este trabajo el problema de clasificación es binario, teniendo dos clases posibles: sano y con artralgia. A continuación se mencionan los elementos a considerar en el entrenamiento de los clasificadores:

- Tarea: Clasificación de pacientes en las clases sano o con Trastorno temporomandibular artralgia.
- Experiencia: Atributos extraídos de señales de movimientos mandibulares provenientes de pacientes previamente diagnosticados.
- Medida de desempeño: Error en la clasificación de los pacientes:

$$\sum_i \hat{y}_i - y_i$$

Donde \hat{y}_i es la predicción del modelo de la i -ésima instancia de validación y y_i es la etiqueta real de esa misma instancia.

3.4.1. Modelos a Entrenar

Para este trabajo se abordó el entrenamiento de siete modelos distintos de clasificación. Los modelos se pueden dividir en dos clases: modelos tradicionales y modelos con funciones cópula. Los modelos tradicionales considerados son cuatro: Máquina de Soporte Vectorial, Perceptrón Multicapa, K-vecinos e Ingenuo de Bayes. Los modelos con funciones cópula considerados son tres: cópula Gaussiana, cópula Clayton y cópula

Frank. En la Sección 2.2 del marco teórico se describen a detalle cada uno de los modelos. En la Tabla 3.5 se mencionan los hiperparámetros considerados en el entrenamiento de los modelos tradicionales. La Sección 3.5 describe a detalle el entrenamiento de los modelos con funciones cópula.

Tabla 3.5: Hiperparámetros de los Modelos de Clasificación

Modelo	Hiperparámetros
Máquina de Soporte Vectorial	$C = 1.0$ $Kernel = lineal$
Perceptrón Multicapa	$optim = adam$ $f_{activacion} = relu$ $lr = 0.01$ $capasmedias = 1 : 10000$
K-vecinos	$k = 5$ $pesos = \frac{1}{distancia}$
Ingenuo de Bayes	Densidades Gaussianas

3.4.2. Validación Cruzada

En el entrenamiento de los modelos de clasificación se utiliza un esquema de validación cruzada de 30 divisiones con instancias seleccionadas aleatoriamente y una proporción para el conjunto de entrenamiento del 80% del total de las instancias y 20% del total para el conjunto de validación. Para este fin se utilizó la función *ShuffleSplit* de la librería *Scikit learn* [68] de Python.

3.5. Clasificadores con Funciones Cópula

En esta sección se describe el proceso seguido en el modelado de los clasificadores con funciones cópula. Para comenzar, se narra la estructura de un clasificador con cópulas, posteriormente se describe el proceso para la obtención del modelo gráfico del clasificador y la estimación de sus parámetros.

3.5.1. Clasificador Probabilístico

El clasificador ingenuo de Bayes, vea Ecuación (3.3), asume independencia condicional entre atributos. Por ejemplo, la probabilidad de que una instancia descrita por

los atributos $x \in \{x_1, x_2, \dots, x_d\}$ donde d es el número total de atributos, pertenezca a una clase A es:

$$P(A|x_1, \dots, x_d) = \frac{P(A) \times \prod_{i=1}^d P(x_i|A)}{P(x_1, \dots, x_d)} \quad (3.3)$$

Pero si se considera la asociación entre atributos, la probabilidad queda de la siguiente manera:

$$P(A|x_1, \dots, x_d) = \frac{c(F_1(x_1|A), \dots, F_d(x_d|A))P(A) \times \prod_{i=1}^d P_i(x_i|A)}{P(x_1, \dots, x_d)} \quad (3.4)$$

Modelar todas las asociaciones entre atributos con una sola cópula, aunque matemáticamente posible, no es práctico porque requiere la estimación simultánea de muchos más parámetros de dependencia, lo que aumenta los costos computacionales. Además, una sola cópula tendría un modelo gráfico completamente conectado que no excluiría dependencias débiles. Modelar cópulas bivariadas es la solución propuesta a este problema.

3.5.2. Modelo Gráfico

Las funciones de cópula pueden modelar dependencias entre variables, pero no todas las variables muestran fuertes dependencias por pares, entonces es importante identificar cuáles son las más relevantes. Esta es la finalidad del modelo gráfico, en este trabajo se propuso el uso de un modelo gráfico de árbol basado en el algoritmo de árbol de expansión mínima (*MST* en inglés)[69] de Kruskal.

En primer lugar, guardamos en una matriz $M_{d \times d}$, donde d representa el número total de variables, la máxima log-verosimilitud para la combinación de las i -ésima y j -ésima variables, de acuerdo con la Ecuación (3.5), que nos servirá para generar el modelo gráfico.

$$M_{i,j} = \ell(\theta|u_i, u_j). \quad (3.5)$$

Donde θ es el parámetro que mide la fuerza de dependencia en una cópula, u_i y u_j son la i -ésima y j -ésima variables transformadas vía $u_i = F_i(x_i)$ y $u_j = F_j(x_j)$.

Antes de que podamos aplicar el algoritmo MST, necesitamos transformar los valores de máxima verosimilitud en M en costos, lo que se puede hacer fácilmente multiplicándolos por -1 . Cuando se aplica, el algoritmo MST generará un modelo gráfico en forma de árbol que muestra las dependencias más fuertes que se modelarán con un

conjunto de datos determinado. El Algoritmo 3 detalla el procedimiento para obtener este modelo gráfico.

Algorithm 3 Pseudocódigo para la obtención del modelo gráfico

- 1: Separa los datos de entrenamiento por clase y por cada una de ellas:
 - 2: Estima las funciones de distribución F_1, F_2, \dots, F_d de los atributos, por ejemplo, usando kernels gaussianos.
 - 3: Transforma los atributos x_1, x_2, \dots, x_d a u_1, u_2, \dots, u_d usando $u_i = F_i(x_i)$
 - 4: **for** i in $\{1, 2, \dots, d-1\}$ **do**
 - 5: **for** j in $\{i+1, i+2, \dots, d\}$ **do**
 - 6: Estima el parámetro $\theta_{i,j}$ usando máxima verosimilitud y guarda $M_{ij} = \ell(\theta|u_i, u_j)$
 - 7: **end for**
 - 8: **end for**
 - 9: Multiplica la matriz M por -1
 - 10: Aplica a M el algoritmo del árbol de expansión mínima
-

Ahora utilizando un ejemplo de modelo gráfico generado con el Algoritmo 3 como el de la Figura 3.4, un clasificador probabilístico con cópulas sería como en la Ecuación (3.6).

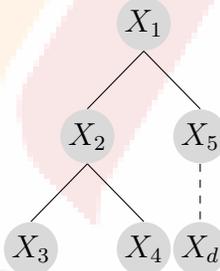


Figura 3.4: Ejemplo de un árbol de dependencia para un conjunto de d atributos

$$P(A|x_1, \dots, x_d) = \frac{c(u_1, u_2)c(u_1, u_5)\dots c(u_{d-1}, u_d)P(A) \times \prod_{i=1}^d f_i(x_i|A)}{f(x_1, \dots, x_d)} \quad (3.6)$$

Donde $u_i = F_i(x_i)$ para $i = 1, 2, \dots, d$. Luego, volviendo a la Figura 3.4, cada una de las líneas que conectan un par de características representa una cópula que modela la asociación entre ellas.

La implementación de los modelos de clasificación con funciones cópula se hizo importando a python el paquete de R 'MLCOPULA' [70], que provee de herramientas

para entrenar este tipo de clasificadores. Este paquete está disponible en el repositorio CRAN.

3.6. Análisis de Resultados

Para el análisis de resultados se obtuvieron 4 métricas de desempeño de los modelos de clasificación entrenados. Estas métricas son: Exactitud (*Accuracy* en inglés), Sensibilidad (*Sensitivity* en inglés), Especificidad (*Specificity* en inglés) y la curva ROC con el área bajo la curva. A continuación se describen estas métricas.

3.6.1. Métricas de desempeño

En un problema de clasificación binaria tenemos 4 posibles combinaciones de la clase verdadera con la clase predicha por un modelo, las cuales se pueden representar en un arreglo bidimensional como el la Tabla 3.6:

Tabla 3.6: Matriz de Confusión

Clase Real	Clase Predicha	
	Positivo	Negativo
Positivo	<i>TP</i>	<i>FN</i>
Negativo	<i>FP</i>	<i>TN</i>

Exactitud

La exactitud es la proporción de instancias que fueron clasificadas correctamente, y su fórmula está descrita en la Ecuación (3.7).

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.7)$$

Sensibilidad

La sensibilidad o recall (*R*) es la razón de las instancias que son realmente positivas y que fueron clasificadas como tal, se calcula de acuerdo con la Ecuación (3.8).

$$R = \frac{TP}{TP + FN} \quad (3.8)$$

Especificidad

La especificidad (E) es la razón de las instancias negativas que son clasificadas correctamente, se calcula de acuerdo con la Ecuación (3.9).

$$E = \frac{TN}{TN + FP} \quad (3.9)$$

Curva ROC

Las predicciones de un modelo frecuentemente son valores reales o probabilidades a los cuales se les aplica un umbral de clasificación para asignarles una clase. La curva ROC mide la capacidad de generalización de un modelo de aprendizaje, evaluando la capacidad discriminativa del modelo. Ordenando las predicciones del modelos para las instancias de validación y gradualmente moviendo el umbral de clasificación obtiene 2 métricas (sensibilidad y 1-especificidad) que son el eje y y x del gráfico respectivamente. La Figura 3.5 muestra un ejemplo de una curva ROC.

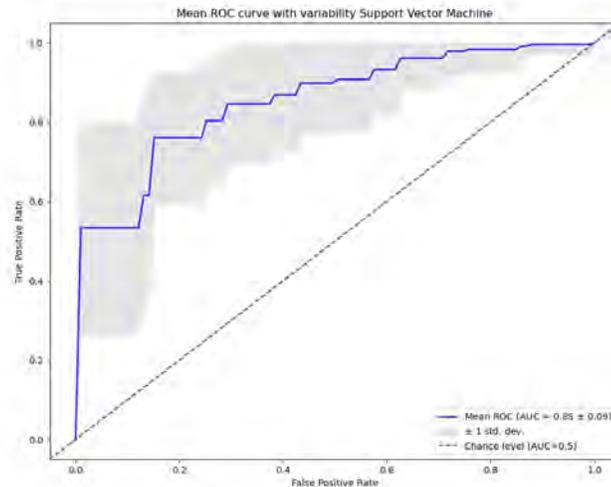


Figura 3.5: Ejemplo de una curva ROC

En otras palabras, la curva ROC muestra la compensación entre sensibilidad y especificidad. Los clasificadores que dan curvas más cercanas a la esquina superior izquierda (ángulo de 90 grados) indican un mejor rendimiento. Cuanto más se acerque la curva a la diagonal de 45 grados del espacio, el desempeño del clasificador es peor, siendo esta curva el desempeño de un clasificador aleatorio donde $sensibilidad = 1 - especificidad$.

Para medir los resultados de una curva ROC se utiliza la métrica AUC, que es el área bajo la curva ROC. Esta métrica puede tomar valores entre 0 y 1. Siendo 1 el desempeño de un clasificador ideal (con una curva ROC en un ángulo de 90 grados). La AUC de un clasificador aleatorio (curva ROC diagonal) es de 0.5.



TESIS TESIS TESIS TESIS TESIS

Capítulo 4

Resultados

En este capítulo se presentan los resultados del desempeño de 7 clasificadores entrenados con 4 distintas selecciones de atributos. Las métricas obtenidas para medir este desempeño son: exactitud, sensibilidad, especificidad y área bajo a curva ROC (AUC). Para este caso, para el cálculo de la sensibilidad y especificidad, se consideró un caso positivo aquel clasificado con TTM y un caso negativo el clasificado como sano. Posterior a la presentación de los resultados, se hace una comparación de los resultados en cada uno de los subconjuntos de atributos escogidos.

4.1. Resultados por Método de Selección

Se extrajeron un total de 121 atributos para la clasificación de TTM. Los atributos de velocidad, aceleración, jerk y frecuencia (vea Tablas 3.2 y 3.3) se obtuvieron para cada uno de los segmentos de señal correspondientes a cada uno de los movimientos mandibulares. Es decir, se extrajeron las 12 características propuestas de velocidad, aceleración, jerk y frecuencia para cada una de las tres señales (x , y y z) y para cada uno de los movimientos (OC, LAT y PROT). Eso suma 108 atributos, a las que sumamos los diez atributos biomecánicos (vea Tabla 3.1) dando una suma de 118 atributos. Finalmente, los atributos de tiempo (vea Tabla 3.4) se extrajeron solo por cada movimiento (OC, LAT y PROT), de manera que se obtuvieron tres. Entonces, en total tenemos 121 atributos.

4.1.1. Resultados con todos los Atributos (sin selección)

En la Tabla 4.1 se muestra el desempeño de los clasificadores cuando se entrenan con todos los atributos disponibles, es decir, sin utilizar un método de selección. Los resultados muestran que respecto a la exactitud, el mejor clasificador es el Ingenuo de Bayes con un resultado de 0.669. Para sensibilidad, el mejor clasificador fue el de cópula Clayton con 0.822. En cuanto especificidad, nuevamente el ingenuo de Bayes tiene el mejor desempeño con 0.639. Finalmente, en términos de AUC, es también el ingenuo de Bayes el mejor.

Tabla 4.1: Media y desviación estándar de las métricas de los clasificadores. Sin selección de atributos.

Clasificador	Exactitud		Sensibilidad		Especificidad		AUC	
	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)
Clayton Copula	0.542	(0.14)	0.822	(0.16)	0.261	(0.22)	0.550	(0.13)
Frank Copula	0.644	(0.15)	0.750	(0.19)	0.539	(0.23)	0.634	(0.17)
Gaussian Copula	0.611	(0.14)	0.706	(0.22)	0.517	(0.22)	0.618	(0.17)
KNN	0.508	(0.13)	0.383	(0.21)	0.630	(0.19)	0.558	(0.16)
NN	0.561	(0.12)	0.544	(0.20)	0.578	(0.16)	0.575	(0.15)
Ingenuo de Bayes	0.669	(0.10)	0.700	(0.18)	0.639	(0.21)	0.722	(0.15)
SVM	0.600	(0.16)	0.578	(0.21)	0.622	(0.19)	0.496	(0.18)

La Figura 4.1 muestra las curvas ROC para los clasificadores entrenados sin selección previa de atributos. En ella podemos observar que todos los clasificadores tuvieron resultados cercanos al nivel de azar ($AUC=0.5$), siendo el resultado más destacable el del clasificador ingenuo de Bayes, seguido por los clasificadores con cópulas Frank y Gaussiana.

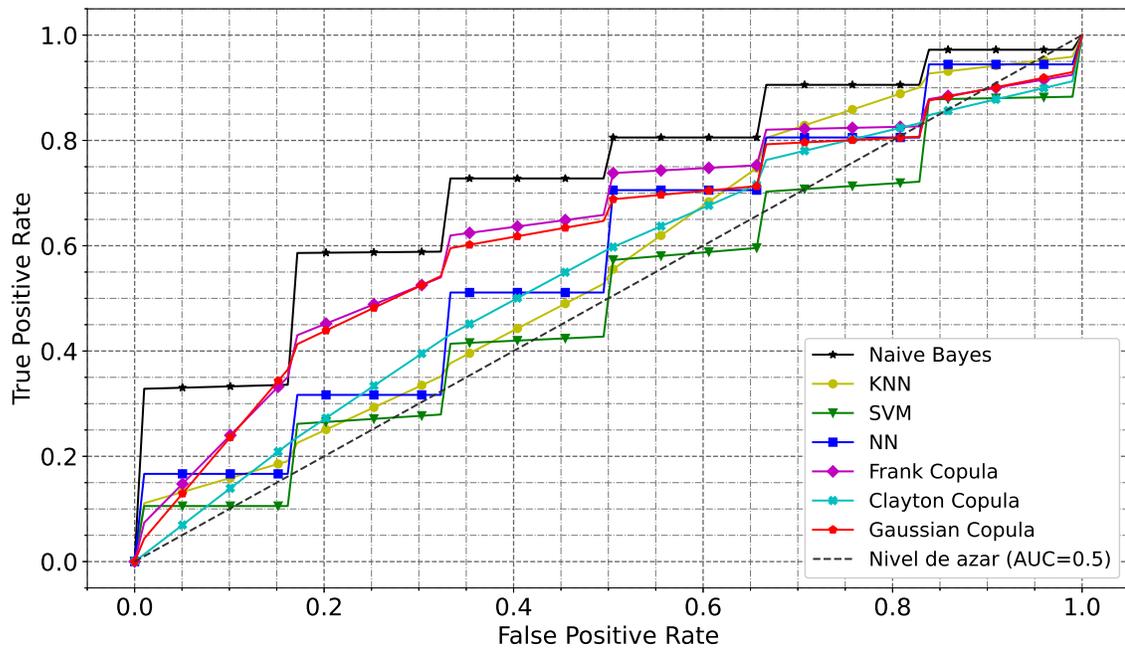


Figura 4.1: Curvas ROC para cada uno de los clasificadores analizados sin selección de atributos.

4.1.2. Resultados con Atributos Seleccionados con Filtro de Correlación de Pearson

En la Tabla 4.2 se muestran los atributos seleccionados con el método de correlación de Pearson descrito en la metodología. Los atributos, LZ, PX y PZ son como en la Tabla 3.1. Para el resto de los atributos, el código OC, LAT y PROT representan el movimiento mandibular del que se extrajo el atributo: apertura/cierre, lateral y protrusión, respectivamente. Las letras X, Y y Z representan el eje del movimiento del que se extrajo el atributo: horizontal, vertical y de profundidad, respectivamente. Se seleccionaron 63 atributos de los 121 originales.

Tabla 4.2: Nombre de los atributos seleccionados por correlación de Pearson

LAT_tiempro_prom	LAT_acprom_x	LAT_acprom_y	LAT_acprom_z
LAT_jerk_maxx	LAT_jerk_maxy	LAT_jerk_maxz	LAT_jerk_prom_x
LAT_jerk_prom_y	LAT_jerk_prom_z	LAT_velmax_z	LAT_velprom_x
LAT_velprom_y	LAT_velprom_z	LAT_x_pot_ratio	LAT_y_pot_ratio
LAT_y_pot_total	LAT_z_pot_ratio	LAT_z_pot_total	LRX
LZ	OC_acprom_x	OC_acprom_z	OC_jerk_maxx
OC_jerk_maxy	OC_jerk_maxz	OC_jerk_prom_y	OC_jerk_prom_z
OC_velmax_x	OC_velmax_z	OC_velprom_x	OC_velprom_y
OC_velprom_z	OC_x_pot_ratio	OC_x_pot_total	OC_y_pot_ratio
OC_y_pot_total	OC_z_pot_ratio	OC_z_pot_total	PROT_tiempro_prom
PROT_ac_maxy	PROT_ac_maxz	PROT_acprom_x	PROT_acprom_y
PROT_acprom_z	PROT_jerk_maxx	PROT_jerk_maxy	PROT_jerk_prom_x
PROT_jerk_prom_y	PROT_jerk_prom_z	PROT_jerk_std_x	PROT_jerk_std_z
PROT_velmax_z	PROT_velprom_x	PROT_velprom_y	PROT_velprom_z
PROT_x_pot_ratio	PROT_y_pot_ratio	PROT_y_pot_total	PROT_z_pot_ratio
PROT_z_pot_total	PX	PZ	

La Tabla 4.3 se muestra la media (desviación estándar entre paréntesis) de las métricas obtenidas para cada uno de los clasificadores probados. El mejor resultado en términos de exactitud fue el Ingenuo de Bayes. Respecto a sensibilidad, el clasificador con cópula Gaussiana fue el mejor. Respecto a especificidad y AUC, nuevamente el ingenuo de Bayes, presenta los mejores resultados.

Analizando las curvas ROC de los clasificadores con selección por correlación (vea Figura 4.2), se puede apreciar una mejora con respecto a los clasificadores de sin selección

Tabla 4.3: Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con correlación de Pearson

Clasificador	Exactitud		Sensibilidad		Especificidad		AUC	
	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)
Clayton Copula	0.478	(0.12)	0.678	(0.16)	0.278	(0.16)	0.482	(0.12)
Frank Copula	0.631	(0.12)	0.761	(0.17)	0.500	(0.22)	0.662	(0.11)
Gaussian Copula	0.614	(0.13)	0.778	(0.18)	0.450	(0.21)	0.664	(0.13)
KNN	0.589	(0.12)	0.500	(0.15)	0.678	(0.18)	0.599	(0.16)
NN	0.692	(0.12)	0.744	(0.18)	0.639	(0.18)	0.740	(0.10)
Ingenuo de Bayes	0.714	(0.12)	0.722	(0.16)	0.706	(0.18)	0.771	(0.13)
SVM	0.672	(0.11)	0.711	(0.17)	0.633	(0.20)	0.758	(0.11)

de atributos. Particularmente los clasificadores KNN, NN, y SVM tuvieron mejoras considerables de hasta 26 puntos porcentuales en AUC.

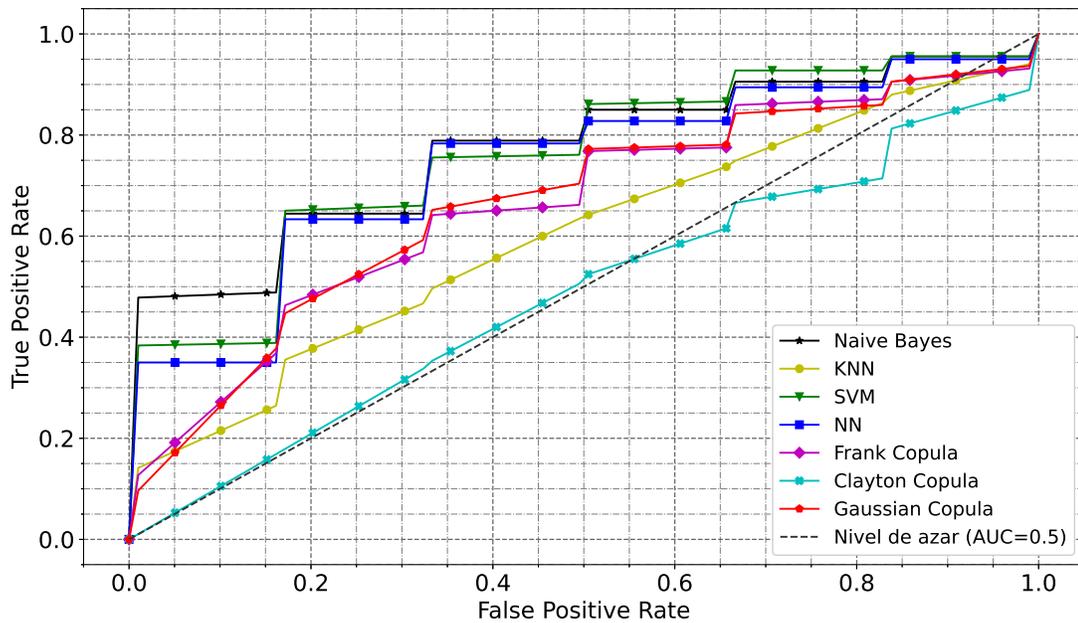


Figura 4.2: Curvas ROC para cada uno de los clasificadores analizados. Selección por correlación.

4.1.3. Resultados con Atributos Seleccionados con Filtro EDA

En la Tabla 4.4 se muestran los atributos seleccionados por el EDA. Los códigos OC, LAT y PROT, así como X, Y y Z tienen el mismo significado mencionado anteriormente en la Sección 4.1.2. Se seleccionaron 39 atributos de los 121 extraídos originalmente.

Tabla 4.4: Nombre de los atributos seleccionados con EDA

LAT_ac_std_y	LAT_acprom_x	LAT_acprom_z	LAT_jerk_maxy
LAT_jerk_prom_x	LAT_velmax_x	LAT_velprom_y	LAT_x_frec_prom
LAT_y_frec_prom	OCX	OCY	OC_tiempo_prom
OC_ac_maxz	OC_acprom_z	OC_jerk_maxx	OC_jerk_maxz
OC_jerk_prom_x	OC_jerk_prom_y	OC_jerk_prom_z	OC_vel_std_x
OC_vel_std_y	OC_vel_std_z	OC_velmax_x	OC_y_frec_prom
OC_y_pot_ratio	PROT_ac_std_x	PROT_ac_std_y	PROT_acprom_y
PROT_acprom_z	PROT_jerk_maxx	PROT_jerk_prom_z	PROT_vel_std_y
PROT_velmax_x	PROT_velprom_y	PROT_velprom_z	PROT_x_frec_prom
PROT_x_pot_ratio	PROT_y_frec_prom	PROT_y_pot_ratio	

La Tabla 4.5 muestra el desempeño de los clasificadores entrenados con la selección de atributos con el algoritmo EDA. En términos de exactitud, los clasificadores SVM y NN tuvieron el mejor resultado con 0.767. Respecto a sensibilidad, el clasificador con cópula Clayton obtuvo los mejores resultados con 0.822. En cuanto a especificidad el KNN obtuvo el mejor resultado con 0.872. Finalmente, el clasificador con el mejor resultado para AUC es nuevamente SVM.

Tabla 4.5: Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con EDA

Clasificador	Exactitud		Sensibilidad		Especificidad		AUC	
	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)
Clayton Copula	0.589	(0.10)	0.822	(0.15)	0.356	(0.15)	0.609	(0.10)
Frank Copula	0.672	(0.12)	0.800	(0.16)	0.544	(0.19)	0.734	(0.15)
Gaussian Copula	0.681	(0.13)	0.778	(0.15)	0.583	(0.21)	0.716	(0.15)
KNN	0.594	(0.09)	0.317	(0.18)	0.872	(0.12)	0.640	(0.13)
NN	0.767	(0.12)	0.744	(0.16)	0.789	(0.19)	0.808	(0.12)
Ingenuo de Bayes	0.678	(0.11)	0.722	(0.18)	0.633	(0.19)	0.750	(0.12)
SVM	0.767	(0.11)	0.728	(0.19)	0.806	(0.19)	0.836	(0.11)

En la Figura 4.3 se muestran las curvas ROC de los clasificadores para este método de selección. Podemos apreciar que los clasificadores NN y SVM obtienen resultados muy similares entre sí y los mejores del grupo. Les siguen con un desempeño medio, los clasificadores probabilísticos, excluyendo al de cópula Clayton. Por último, el clasificador KNN y el de cópula Clayton tuvieron los peores resultados.

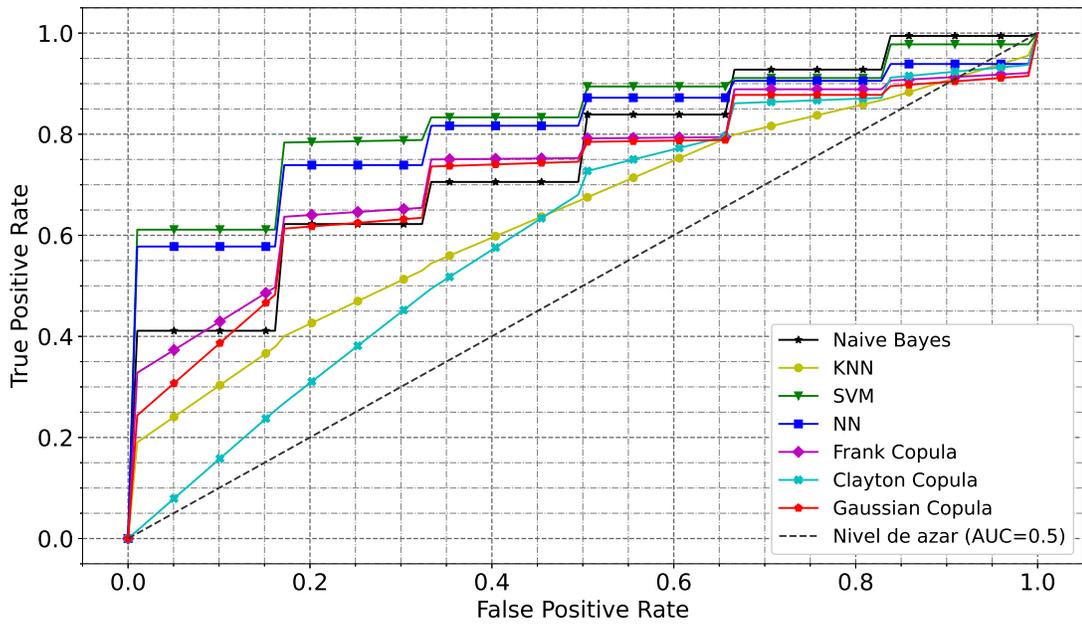


Figura 4.3: Curvas ROC para cada uno de los clasificadores analizados. Selección con EDA

4.1.4. Resultados con Atributos Seleccionados con Filtro de Correlación de Pearson y EDA

El último caso de selección es una combinación de los dos métodos de selección vistos: correlación y EDA. Es decir, primero se filtran los atributos con correlación y de esta primera selección se hace una segunda selección con el algoritmo EDA. La Tabla 4.6 muestra los atributos seleccionados con este método. Los códigos OC, LAT y PROT, así como X, Y y Z tienen la misma interpretación, previamente mencionada en la Sección 4.1.2. Se seleccionados 20, de los 121 atributos originales.

Tabla 4.6: Nombre de los atributos seleccionados por correlación de Pearson y EDA

LAT_jerk_maxy	LAT_jerk_prom_x	LAT_velprom_y	OC_acprom_z
OC_jerk_maxx	OC_jerk_prom_z	OC_velmax_z	OC_velprom_y
OC_y_pot_ratio	OC_y_pot_total	OC_z_pot_ratio	PROT_tiempo_prom
PROT_acprom_z	PROT_jerk_maxx	PROT_jerk_prom_y	PROT_jerk_prom_z
PROT_jerk_std_x	PROT_velprom_y	PROT_velprom_z	PROT_z_pot_ratio

La Tabla 4.7 muestra el desempeño de los clasificadores entrenados con la selección combinada de correlación y EDA. En las cuatro métricas obtenidas, el clasificador SVM presenta los mejores resultados con 0.858, 0.844, 0.872 y 0.921 para exactitud, sensibilidad, especificidad, y AUC, respectivamente. La red neuronal, en este caso queda un segundo lugar, pero con métricas muy cercanas a las del SVM.

Tabla 4.7: Media y desviación estándar de las métricas de los clasificadores. Atributos seleccionados con correlación de Pearson y EDA

Clasificador	Exactitud		Sensibilidad		Especificidad		AUC	
	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)	Media	(desviación estándar)
Clayton Copula	0.625	(0.10)	0.594	(0.16)	0.656	(0.175)	0.645	(0.13)
Frank Copula	0.711	(0.11)	0.733	(0.16)	0.689	(0.19)	0.769	(0.12)
Gaussian Copula	0.717	(0.13)	0.778	(0.17)	0.656	(0.22)	0.747	(0.14)
KNN	0.597	(0.11)	0.356	(0.16)	0.839	(0.16)	0.640	(0.11)
NN	0.842	(0.09)	0.811	(0.13)	0.872	(0.13)	0.906	(0.06)
Ingenuo de Bayes	0.658	(0.11)	0.667	(0.14)	0.650	(0.19)	0.757	(0.11)
SVM	0.858	(0.07)	0.844	(0.09)	0.872	(0.113)	0.921	(0.06)

En la Figura 4.4 donde se muestran las curvas ROC de los clasificadores se pueden distinguir tres grupos distintivos de acuerdo a su desempeño: alto, medio y bajo. Los clasificadores NN y SVM tienen el desempeño alto. EL desempeño medio se observa

en los clasificadores probabilísticos excluyendo el de cópula Clayton. Finalmente los clasificadores con cópula Clayton y el KNN tienen un desempeño pobre.

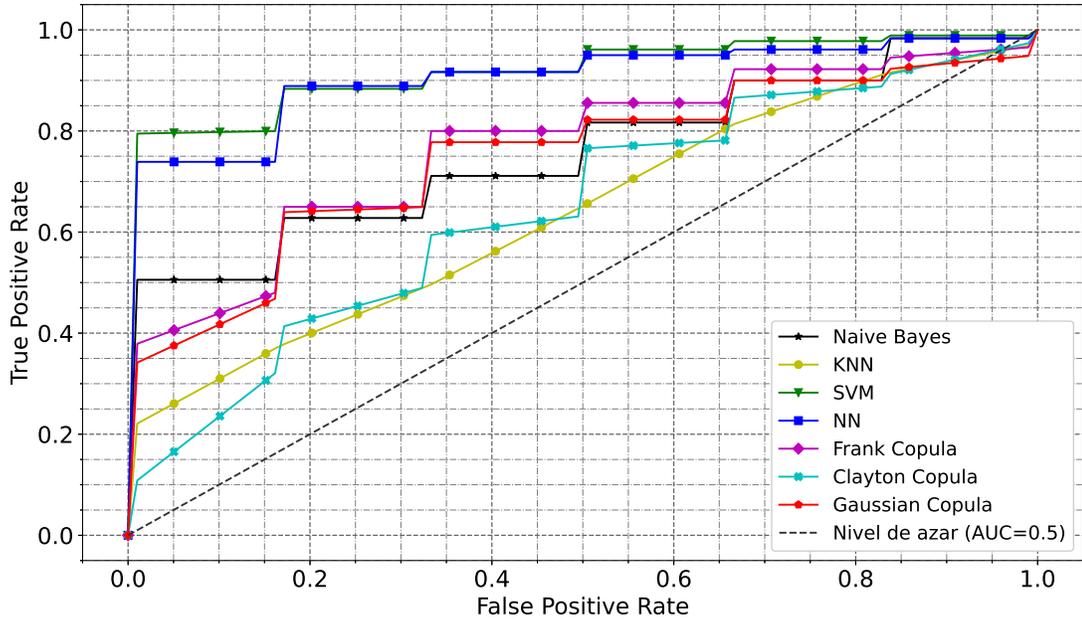


Figura 4.4: Curvas ROC para cada uno de los clasificadores analizados. Selección con correlación y EDA.

4.2. Comparación de Resultados

Para comparar el desempeño de los clasificadores en cada una de las métricas y métodos de selección se aplicó una prueba de hipótesis no paramétrica de Kruskal-Wallis (Vea Tabla 4.8). Con un nivel de significancia $\alpha = 0.05$ podemos concluir que se rechaza la hipótesis nula de igualdad de medias en todos los casos. Es decir, hay diferencias en el desempeño de los clasificadores para las 4 métricas obtenidas y los 4 métodos de selección.

Tabla 4.8: Valores p de prueba Kruskal-Wallis para cada métrica de desempeño y método de selección.

	Sin selección	Correlación	EDA	Corr-EDA
Exactitud	8.348e-05	1.293e-09	3.872e-09	1.764e-18
Sensibilidad	1.363e-12	3.007e-07	1.098e-13	2.364e-19
Especificidad	2.064e-08	5.039e-14	3.623e-18	4.004e-10
AUC	2.591e-05	3.064e-14	2.007e-10	1.144e-20

Partiendo de que tenemos diferencias en el desempeño de los clasificadores para todas las métricas y todos los métodos de selección, analizaremos cada uno de los casos para identificar cuáles clasificadores se desempeñaron mejor y cuáles peor.

Comenzando con la métrica exactitud, la Figura 4.5 muestra una comparación del desempeño de los 7 clasificadores en los 4 métodos de selección de atributos. La figura muestra una tendencia creciente entre los métodos de selección. La única excepción es el clasificador ingenuo de Bayes que empeora su desempeño al aumentar la complejidad del método de selección.

Por otro lado, la Tabla 4.9 muestra los resultados de pruebas Wilcoxon por pares de clasificadores en cada uno de los métodos de selección. La corrección del valor p de estas pruebas se hizo con el método Benjamini-Hochberg y el nivel de significancia $\alpha = 0.05$.

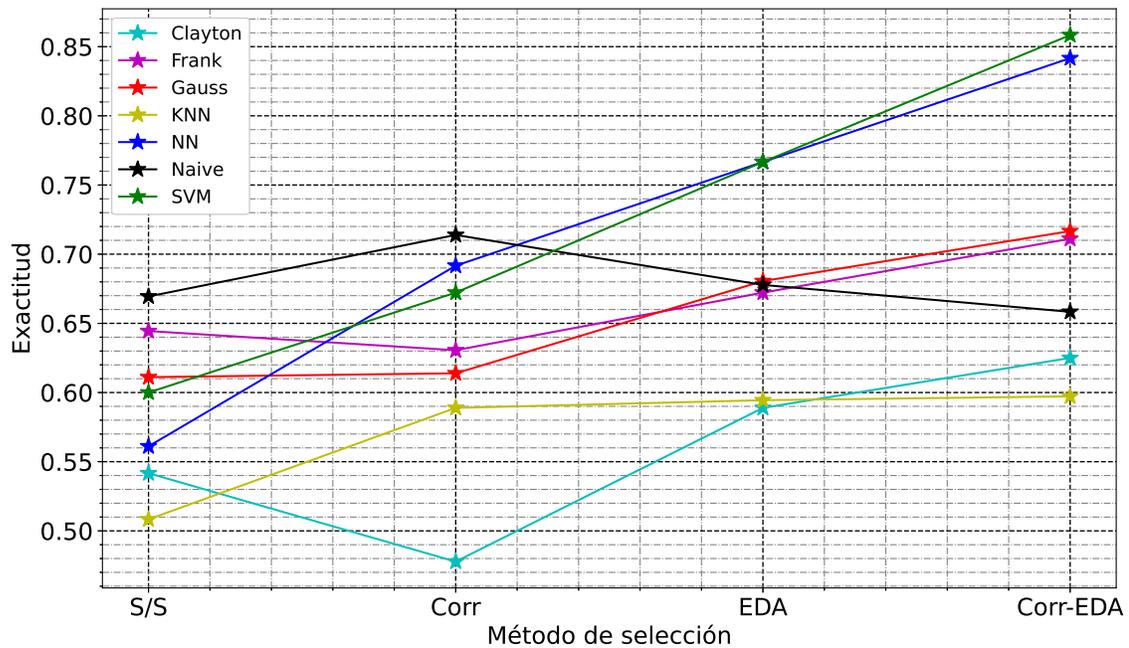


Figura 4.5: Desempeño en exactitud por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.

Analizando la Figura 4.5 y la Tabla 4.9 podemos llegar a hacer varias aseveraciones. Primero en el caso de los clasificadores entrenados sin selección de atributos (S/S en la Figura) los tres mejores resultados, el Ingenuo de Bayes, cópula Frank y copula Gaussiana no tienen diferencias significativas entre sí, pero sí con el resto de clasificadores. Para los clasificadores con selección de atributos por correlación, los mejores clasificadores son el ingenuo de Bayes, la red neuronal (NN) y la máquina de soporte vectorial (SVM), que no tienen diferencia en desempeño entre sí, pero sí con el resto de clasificadores. Para la selección con EDA y Corr-EDA, la máquina de soporte vectorial (SVM) y la red neuronal (NN) se mantienen como los mejores clasificadores, con una diferencia significativa al resto.

Tabla 4.9: Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica exactitud. $\alpha = 0.05$.

Par de clasificadores		Valor p corregido			
Clf A	Clf B	S/S	Corr	EDA	Corr-EDA
Clayton Copula	Frank Copula	0.039	0.0	0.018	0.011
Clayton Copula	Gaussian Copula	0.129	0.001	0.018	0.011
Clayton Copula	KNN	0.385	0.008	0.909	0.264
Clayton Copula	NN	0.722	0.0	0.0	0.0
Clayton Copula	Naive Bayes	0.003	0.0	0.016	0.223
Clayton Copula	SVM	0.33	0.0	0.0	0.0
Frank Copula	Gaussian Copula	0.569	0.65	0.909	0.78
Frank Copula	KNN	0.008	0.229	0.018	0.001
Frank Copula	NN	0.082	0.099	0.018	0.0
Frank Copula	Naive Bayes	0.569	0.029	0.909	0.186
Frank Copula	SVM	0.33	0.229	0.018	0.0
Gaussian Copula	KNN	0.028	0.45	0.018	0.001
Gaussian Copula	NN	0.164	0.049	0.032	0.001
Gaussian Copula	Naive Bayes	0.19	0.01	0.909	0.103
Gaussian Copula	SVM	0.698	0.099	0.038	0.0
KNN	NN	0.251	0.007	0.0	0.0
KNN	Naive Bayes	0.0	0.001	0.016	0.034
KNN	SVM	0.091	0.012	0.0	0.0
NN	Naive Bayes	0.003	0.538	0.018	0.0
NN	SVM	0.478	0.653	0.909	0.502
Naive Bayes	SVM	0.13	0.229	0.018	0.0

La Figura 4.6 muestra la comparación del desempeño de los clasificadores en términos de la métrica AUC. Al igual que el caso de la exactitud, la AUC tiene una tendencia creciente: conforme aumenta la complejidad del método de selección de atributos también aumenta la AUC promedio de los clasificadores. Con excepción del clasificador ingenuo de Bayes, que se mantiene su desempeño relativamente estable sin importar el método de selección de atributos.

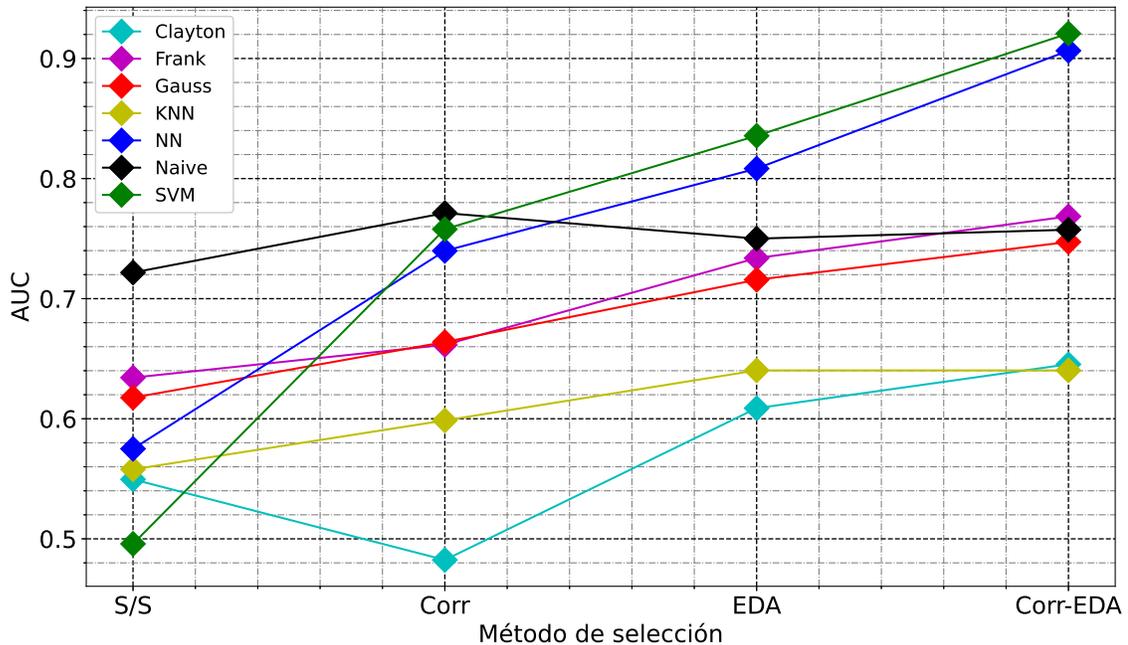


Figura 4.6: Desempeño en AUC por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.

La Tabla 4.10 confirma las tendencias observadas en la Figura 4.6. Para el caso de clasificadores sin selección de atributos el Ingenuo es el mejor, con una diferencia significativa comparado con los clasificadores KNN, NN y SVM. Los clasificadores con cópula Frank y Gaussiana le siguen en segundo lugar, pero con un desempeño estadísticamente igual al del ingenuo de Bayes. El clasificador con el peor desempeño en este caso es la máquina de soporte vectorial.

Con método de selección por correlación los clasificadores SVM y NN aumentan considerablemente su desempeño y se empatan con los clasificadores ingenuo de Bayes. El único clasificador que empeora su desempeño es el de cópula Clayton. Por último, con la selección con EDA y Corr-EDA los clasificadores forman tres grupos distintos con diferencias significativas en desempeño. El primero en orden de mayor a menor

desempeño, compuesto por SVM y NN que empatan entre sí; el segundo compuesto por el ingenuo de Bayes, y cópulas Frank y Gaussiana, que también empatan entre sí; el tercero, compuesto por el KNN y cópula Clayton.

Tabla 4.10: Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica AUC. $\alpha = 0.05$.

Par de clasificadores		Valor p corregido			
clf A	clf B	S/S	Corr	EDA	Corr-EDA
Clayton Copula	Frank Copula	0.129	0.0	0.003	0.001
Clayton Copula	Gaussian Copula	0.129	0.0	0.011	0.003
Clayton Copula	KNN	1.0	0.009	0.288	0.877
Clayton Copula	NN	1.0	0.0	0.0	0.0
Clayton Copula	Naive Bayes	0.001	0.0	0.0	0.001
Clayton Copula	SVM	0.167	0.0	0.0	0.0
Frank Copula	Gaussian Copula	0.924	0.953	0.646	0.791
Frank Copula	KNN	0.129	0.085	0.033	0.0
Frank Copula	NN	0.162	0.016	0.079	0.0
Frank Copula	Naive Bayes	0.129	0.001	0.679	0.877
Frank Copula	SVM	0.016	0.007	0.019	0.0
Gaussian Copula	KNN	0.195	0.085	0.095	0.002
Gaussian Copula	NN	0.296	0.015	0.024	0.0
Gaussian Copula	Naive Bayes	0.105	0.001	0.432	0.877
Gaussian Copula	SVM	0.032	0.007	0.006	0.0
KNN	NN	0.869	0.001	0.0	0.0
KNN	Naive Bayes	0.003	0.0	0.006	0.0
KNN	SVM	0.296	0.0	0.0	0.0
NN	Naive Bayes	0.003	0.159	0.086	0.0
NN	SVM	0.162	0.555	0.601	0.401
Naive Bayes	SVM	0.0	0.431	0.024	0.0

La Figura 4.7 muestra el desempeño de los clasificadores en términos de sensibilidad. En ella observamos un comportamiento particular que no habíamos observado en otras métricas: La sensibilidad se mantiene relativamente estable y con valores altos sin importar el tipo de selección de atributos usado, siendo el clasificador KNN la única excepción a este comportamiento.

El clasificador con Cópula Clayton en un caso particularmente interesante, pues pasa de ser uno de los clasificadores con peor desempeño a uno de los mejores. En los casos en los que no se usa selección y cuando se usa el EDA como método de selección, este clasificador tiene el mejor desempeño de todos los clasificadores.

Aunado a esto, los clasificadores con cópulas presentan un comportamiento particular en el que cuando se utiliza correlación (ya sea en solitario o con EDA) como método de selección de atributos, el desempeño es menor comparado con los otros dos métodos de selección.

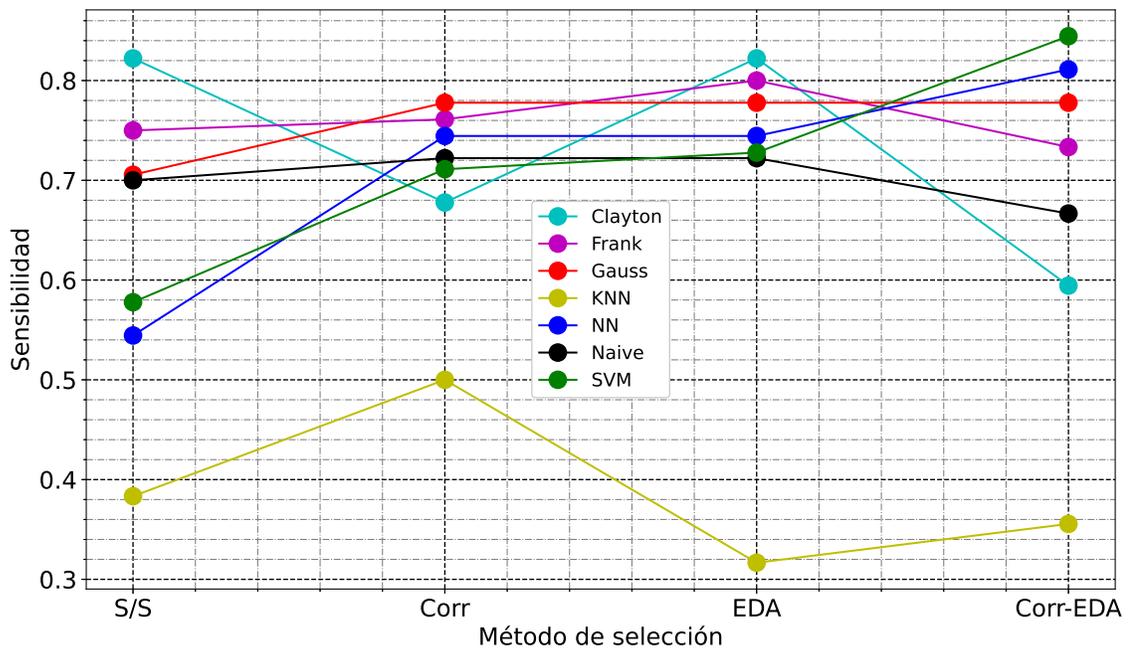


Figura 4.7: Desempeño en sensibilidad por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.

La Tabla 4.11 confirma lo descrito anteriormente. La mayoría de las diferencias significativas en sensibilidad entre los clasificadores se dan con el KNN, que es el dato atípico en este caso. Entre el resto de los clasificadores no es posible identificar uno que destaque sobre el resto. Por ejemplo, para los clasificadores entrenados con selección por

correlación no hay diferencias significativas entre los 6 clasificadores mejor posicionados. Este resultado se repite para los clasificadores entrenados con atributos seleccionados por EDA.

Tabla 4.11: Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica Sensibilidad. $\alpha = 0.05$.

Par de clasificadores		Valor p corregido			
clf A	clf B	S/S	Corr	EDA	Corr-EDA
Clayton Copula	Frank Copula	0.154	0.238	0.756	0.004
Clayton Copula	Gaussian Copula	0.042	0.086	0.4	0.0
Clayton Copula	KNN	0.0	0.001	0.0	0.0
Clayton Copula	NN	0.0	0.378	0.172	0.0
Clayton Copula	Naive Bayes	0.012	0.585	0.141	0.14
Clayton Copula	SVM	0.0	0.69	0.163	0.0
Frank Copula	Gaussian Copula	0.498	0.73	0.645	0.277
Frank Copula	KNN	0.0	0.0	0.0	0.0
Frank Copula	NN	0.001	0.757	0.316	0.105
Frank Copula	Naive Bayes	0.351	0.585	0.247	0.105
Frank Copula	SVM	0.006	0.474	0.287	0.01
Gaussian Copula	KNN	0.0	0.0	0.0	0.0
Gaussian Copula	NN	0.009	0.628	0.683	0.541
Gaussian Copula	Naive Bayes	0.902	0.331	0.568	0.004
Gaussian Copula	SVM	0.042	0.283	0.568	0.14
KNN	NN	0.009	0.0	0.0	0.0
KNN	Naive Bayes	0.0	0.0	0.0	0.0
KNN	SVM	0.005	0.0	0.0	0.0
NN	Naive Bayes	0.004	0.73	0.858	0.001
NN	SVM	0.618	0.643	0.858	0.405
Naive Bayes	SVM	0.034	0.757	0.963	0.0

Por último, la Figura 4.8 muestra el desempeño en términos de especificidad. En esta métrica recuperamos el comportamiento creciente observado en las métricas exactitud y AUC. En este caso destaca el clasificador KNN, que en las otras métricas presenta malos resultados, en especificidad tiene uno de los mejores. Los clasificadores con cópula tienen valores de especificidad muy bajos, aunque mejoran con métodos de selección como el Corr-EDA.

Los clasificadores KNN, SVM y NN observan una mejora en el desempeño conforme incrementa la complejidad del método de selección de atributos, superando por más de 10 puntos porcentuales el desempeño del resto de clasificadores cuando se utiliza selección de atributos por EDA y Corr-EDA. Finalmente el clasificador ingenuo de Bayes se mantiene relativamente estable sin importar el método de selección de atributos.

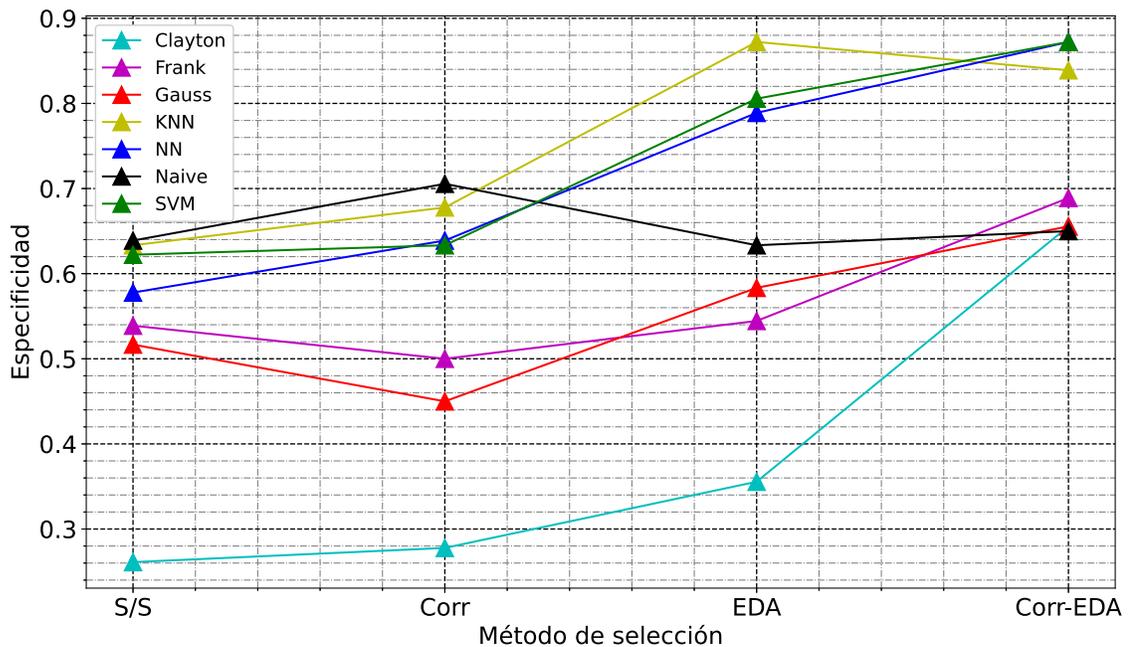


Figura 4.8: Desempeño en especificidad por clasificador y método de selección. S/S: Sin selección, Corr: selección por Correlación, EDA: Selección por EDA, Corr-EDA: Selección por Correlación y EDA.

La Tabla 4.12 muestra que en los clasificadores entrenados sin selección de atributos las diferencias no son significativas entre 6 de los 7 clasificadores. El único que presenta diferencias significativas es el de cópula Clayton, que para este caso tiene un resultado muy pobre.

En el resto de métodos de selección de atributos, los clasificadores forman dos grupos: el primero compuesto por el KNN, SVM, y NN; el segundo compuesto por todos

los clasificadores probabilísticos. Intragrupo los resultados son muy similares, es decir, no presentan diferencias significativas. Sin embargo, intergrupo sí existen diferencias significativas, siendo el grupo del KNN, SVM y NN el que se desempeña mejor.

Tabla 4.12: Valores p de la prueba Wilcoxon por pares de clasificadores para el análisis de la métrica Especificidad. $\alpha = 0.05$.

Par de clasificadores		Valor p corregido			
clf A	clf B	S/S	Corr	EDA	Corr-EDA
Clayton Copula	Frank Copula	0.0	0.0	0.001	0.612
Clayton Copula	Gaussian Copula	0.0	0.003	0.0	0.904
Clayton Copula	KNN	0.0	0.0	0.0	0.0
Clayton Copula	NN	0.0	0.0	0.0	0.0
Clayton Copula	Naive Bayes	0.0	0.0	0.0	0.92
Clayton Copula	SVM	0.0	0.0	0.0	0.0
Frank Copula	Gaussian Copula	0.776	0.539	0.51	0.753
Frank Copula	KNN	0.278	0.003	0.0	0.004
Frank Copula	NN	0.71	0.011	0.0	0.0
Frank Copula	Naive Bayes	0.211	0.0	0.132	0.612
Frank Copula	SVM	0.278	0.011	0.0	0.0
Gaussian Copula	KNN	0.134	0.0	0.0	0.002
Gaussian Copula	NN	0.388	0.002	0.001	0.0
Gaussian Copula	Naive Bayes	0.119	0.0	0.399	0.904
Gaussian Copula	SVM	0.134	0.002	0.0	0.0
KNN	NN	0.484	0.539	0.128	0.686
KNN	Naive Bayes	0.817	0.609	0.0	0.0
KNN	SVM	0.976	0.539	0.312	0.753
NN	Naive Bayes	0.317	0.248	0.008	0.0
NN	SVM	0.401	0.956	0.712	0.904
Naive Bayes	SVM	0.796	0.268	0.002	0.0

TESIS TESIS TESIS TESIS TESIS

Capítulo 5

Discusiones

El objetivo de esta investigación fue el de evaluar el desempeño de distintos modelos de aprendizaje supervisado en la tarea de clasificar individuos como sanos o con trastorno temporomandibular artralgia. Para ello se realizó una serie de experimentos, que bajo diversas condiciones pusieron a prueba el rendimiento de 7 clasificadores distintos.

De los resultados de los experimentos podemos deducir varios puntos: en primer lugar, podemos deducir que los clasificadores SVM y NN tuvieron el mejor desempeño de todos los clasificadores probados y en todas las métricas obtenidas; en segundo lugar, podemos decir que el desempeño de los clasificadores depende ampliamente de la selección de atributos con los que se entrenan.

La primera afirmación del párrafo anterior se justifica en las pruebas de hipótesis realizadas que muestran diferencias significativas en el desempeño medio de los clasificadores SVM y NN con el resto de los clasificadores, esto se muestra en la Tabla 4.9. Lo anterior tomando como base el método de selección Corr-EDA, que constantemente provee de los mejores resultados, vea Figura 4.4.

La afirmación de que el desempeño de los clasificadores depende de la selección de atributos con los que se entrene tiene también una justificación. Para los distintos métodos de selección probados, los clasificadores muestran también resultados distintos, es decir, al modificar la selección de atributos, se afecta o beneficia el desempeño del clasificador.

Cada uno de los clasificadores probados presenta patrones de cambio distintos. Por ejemplo, el SVM y NN consistentemente mejoran su desempeño cuando aumenta la complejidad del método de selección y a su vez, disminuye el número de atributos usado en el entrenamiento. Por el contrario, el clasificador Ingenuo de Bayes mantiene su desempeño estable sin importar la selección y atributos, y en algunos casos disminuye

conforme aumenta la complejidad del método de selección.

Los clasificadores con funciones cópula son un caso particularmente interesante, pues a diferencia de modelos como SVM o Ingenuo de Bayes (creciente en el caso de SVM y estable en el caso del ingenuo de Bayes, vea Figuras 4.5, 4.6, 4.7 y 4.8), que presentan el mismo patrón de desempeño en todas las métricas, los clasificadores con cópulas muestran comportamientos distintos para cada una de sus métricas. El caso más evidente de esto es el de la métrica sensibilidad, que para los clasificadores con cópula tiende a reducirse considerablemente cuando se utiliza algún tipo de selección por correlación.

Los clasificadores probabilísticos basados en funciones cópula funcionan mejor entre más fuertes sean las asociaciones entre los atributos. Es ese precisamente el objetivo de las funciones cópula: modelar dependencias. Al utilizar métodos de selección de atributos que filtran aquellos altamente correlacionados, se limita la capacidad de estos clasificadores. Es decir, eliminar atributos altamente correlacionados puede propiciar que el desempeño de los clasificadores con cópula no sea el óptimo.

Desde esta perspectiva es importante mencionar que, si se busca maximizar el desempeño de varios clasificadores, utilizar un método de selección de atributos para todos los clasificadores sea inadecuado cuando los principios de funcionamiento de los clasificadores sean muy distintos entre sí. En estos casos, utilizar métodos de selección distintos diseñados específicamente para potenciar el desempeño del clasificador a entrenar.

Aún con lo mencionado anteriormente, existe la probabilidad de que haya atributos que funcionen para todos los clasificadores. En la Tabla 5.1 se presentan los nombres de los atributos que fueron seleccionados de manera independiente por los 3 métodos de selección probados: Corr, EDA y Corr-EDA. Estos atributos podemos decir que cumplen los criterios establecidos por los 3 métodos de selección.

Tabla 5.1: Atributos seleccionados por los 3 métodos de selección

OC_jerk_maxx	OC_acprom_z	OC_jerk_prom_z
OC_y_pot_ratio	LAT_jerk_prom_x	LAT_velprom_y
LAT_jerk_maxy	PROT_jerk_maxx	PROT_velprom_y
PROT_velprom_z	PROT_acprom_z	PROT_jerk_prom_z

Comparando los resultados de esta investigación con trabajos que también analizan movimientos mandibulares como [25, 8], encontramos que son similares para clasificadores como el ingenuo de Bayes y SVM, sin embargo, para el clasificador KNN, los resultados obtenidos en este trabajo son muy inferiores a las presentadas en otros trabajos. Respecto a otros trabajos enfocados a la misma tarea de clasificación pero con

distinto tipo de datos encontramos que los modelos entrenados en este trabajo tienen un desempeño más bajo que aquellos propuestos por [50] y [49] que utilizan señales de audio e imágenes respectivamente. En este trabajo el mejor resultado lo brinda un modelo SVM con 0.86 en exactitud. En [50] logran un 0.94 de exactitud en clasificación. En este caso es importante mencionar que la tarea de clasificación en [50] es distinta a la propuesta en este trabajo, ya que clasifican la presencia de ruidos articulares al hacer movimientos de la mandíbula. La presencia de estos ruidos está altamente relacionada con TTM.

5.1. Limitaciones del Trabajo

La principal limitante en el desarrollo de este trabajo fue la cantidad de individuos de los que se contaba para el entrenamiento de los modelos de clasificación. Inicialmente se contaba con un total de 72 registros de movimientos mandibulares del mismo número de individuos. Posterior al procesamiento de estos registros se descartaron 13 de ellos debido a las fluctuaciones en la frecuencia de muestreo de las señales. Este descarte debido a que cambios en la frecuencia de muestreo a lo largo de la señal podría resultar en una extracción de características inadecuada.

El bajo número de registros fue lo que llevó a tomar la decisión de usar un esquema de validación cruzada de muestreo aleatorio con un número definido de muestras y en proporciones entrenamiento- validación definidas. Esto en contraste con la forma más común de validación cruzada que considera grupos de validación mutuamente excluyentes.

5.2. Trabajo a Futuro

Trabajo futuro en la clasificación de trastornos temporomandibulares con datos de movimiento debería contar con un mayor número de registros de individuos con la finalidad de que los resultados de la clasificación puedan ser generalizables. Asimismo, en trabajos de comparación del desempeño de clasificadores deberán considerar la utilización de métodos de selección de atributos que se ajusten a los principios de funcionamiento del clasificador. Por ejemplo, la utilización de métricas como la rho de Spearman, tau de Kendall, o información mutua, bajo el principio de mantener atributos que mantienen fuertes relaciones entre ellos. Finalmente dados los resultados en sensibilidad y especificidad para los modelos con funciones cópula, experimentos que

modelen los datos de entrenamiento de cada clase con cópulas distintas deberían considerarse. Por ejemplo, cópula Clayton para la clase positiva y cópula Gaussiana para la clase negativa podrían generar un modelo con mejor desempeño que el que se obtiene al modelar ambas clases con las mismas cópulas.



TESIS TESIS TESIS TESIS TESIS

Capítulo 6

Conclusiones

En este trabajo se propone la evaluación del desempeño de siete modelos de aprendizaje supervisado, incluyendo 3 modelos de clasificación con funciones cópula, en la tarea de clasificación de trastornos temporomandibulares.

Un total de 121 atributos fueron extraídos de registros de movimientos mandibulares de 59 individuos participantes en el estudio. Los modelos de clasificación fueron entrenados con tres métodos distintos de selección de atributos. Cuatro métricas fueron utilizadas para medir el desempeño de los clasificadores: Exactitud, sensibilidad, especificidad y AUC.

Los resultados de la evaluación muestran que los mejores clasificadores en todas las métricas evaluadas son la máquina de soporte vectorial (SVM) y la red neuronal (NN). El resto de los clasificadores presentan un desempeño por debajo de estos dos. Los clasificadores basados en funciones cópula presentan su mejor desempeño en la métrica de sensibilidad, donde está casi a la par que el SVM y NN.

Asimismo, los resultados sugieren que el método de selección de atributos influye en gran medida en el desempeño de algunos clasificadores. Por ejemplo, para SVM y NN, entre menor sea el número de atributos con los que se entrenan, presentan mejor desempeño. Por otro lado, los clasificadores con funciones cópula tienen un mejor desempeño cuando la selección de atributos no incluye filtros de correlación entre variables.

Las limitantes de este trabajo fueron el número de registros disponibles, con los que se contó para el entrenamiento de los modelos de clasificación. En trabajos futuros relacionado con la clasificación de trastornos temporomandibulares se sugiere la recopilación de más registros de movimientos mandibulares y el uso de más métodos de selección de atributos.

Finalmente, podemos decir que sí es posible clasificar individuos en las clases TTM

artralgia y sanos caracterizando movimientos mandibulares y utilizando esas características como datos de entrenamiento. Sin embargo, la clasificación de TTM sigue siendo un área de oportunidades para la investigación, especialmente en lo referente a los métodos de extracción y selección de atributos para el entrenamiento de los modelos de clasificación.



Bibliografía

- [1] U. of Dundee, “Temporomandibular joint (tmj) by emily mcdougall and annie campbell,” 2016.
- [2] Z. H. Zhou, *Machine Learning*. Springer Nature, 1 2021.
- [3] Mathworks, “Support vector machine (svm),” 2023. <https://la.mathworks.com/discovery/support-vector-machine.html> accessed: 2023/11/14.
- [4] M. Kubat, *An Introduction to Machine Learning*. Springer International Publishing, 9 2017.
- [5] G. Dougherty, *Pattern Recognition and Classification: An Introduction*. Springer-Verlag New York Incorporated, 1 2012.
- [6] A. Montoya, “Visualización de copulas.” <https://abrahammontoyacalzada.shinyapps.io/Copulas/>, 2024. [Accessed 22-07-2024].
- [7] E. Schiffman, R. Ohrbach, E. Truelove, J. Look, G. Anderson, J.-P. Goulet, P. Svensson, and Y. Gonzalez, “Diagnostic criteria for temporomandibular disorders (dc/tmd) for clinical and research applications: Recommendations of the international rdc/tmd consortium network * and orofacial pain special interest group † hhs public access,” 2014.
- [8] M. Castillo-Juárez, “Análisis cuantitativo de movimientos mandibulares para la clasificación de trastornos temporomandibulares por medio de un sistema de visión computacional,” *Universidad Autonoma de Aguascalientes*, 2022.
- [9] J. Durham, T. R. Newton-John, and J. M. Zakrzewska, “Temporomandibular disorders,” *BMJ (Online)*, vol. 350, 3 2015.

- [10] N. I. of Health, “Los trastornos temporomandibulares,” *National Institutes of Health*, 2022. <https://www.nidcr.nih.gov/espanol/temas-de-salud/los-trastornos-temporomandibulares>.
- [11] N. S. Norton, “Netter’s head and neck anatomy for dentistry,” *Elsevier*, pp. 256–270, 2016.
- [12] D. Bhargava and P. Gurjar, *Anatomy and Basic Biomechanics of the Temporomandibular Joint*, pp. 9–21. Springer Nature, 1 2021.
- [13] T. von Arx and S. Lozanoff, *Temporomandibular Joint*, pp. 525–534. Springer International Publishing, 2017.
- [14] M. N. Shaun, *Etiology, Presentation, and Diagnosis*, pp. 1–9. Springer, 2018.
- [15] A. A. Patel and C. Patel, *Temporomandibular Joint Disorder*, pp. 25–27. Springer International Publishing, 2017.
- [16] J. E. Petsavage-Thomas, “Unlocking the jaw: Advanced imaging of the temporomandibular joint,” *American Journal of Roentgenology*, vol. 203, pp. 1047–1058, 11 2014.
- [17] R. Shaikh, D. Findlay, and R. Reti, *Temporomandibular Joint Dysfunction*, pp. 23–41. Springer International Publishing, 1 2020.
- [18] W.-S. B.-B. B. Sojka AZarowski, MSteinborn BHedzelek, “Temporomandibular disorders in adolescents with headache,” *Advances in Clinical and Experimental Medicine*, vol. 27, pp. 193–199, 2 2018.
- [19] R. L. Gauer and M. J. Semidey, “Diagnosis and treatment of temporomandibular disorders,” *Am Fam Physician*, vol. 91, 2015.
- [20] J. Kostrzewa-Janicka, E. Mierzwinska-Nastalska, P. Jurkowski, P. Okonski, and M. Nedzi-Gora, “Assessment of temporomandibular joint disease,” *Advances in Experimental Medicine and Biology*, vol. 788, pp. 207–211, 2013.
- [21] M. M. Rojas, P. L. Mendieta, and M. O. A.F., “Perfil epidemiológico del trastorno temporomandibular en mexicanos con maloclusión,” *Revista Mexicana de Ortodoncia*, vol. 3, pp. 79–83, 2015.

- [22] R. Cipolla, S. Battiato, and G. Farinella, *Computer Vision Detection, Recognition and Reconstruction*. Springer-Verlag, 2010.
- [23] R. Bajcsy, *Computer Vision*. Springer US, 2014.
- [24] L. E. Sucar and M. G. Gómez, *Visión Computacional*. El Instituto Nacional de Astrofísica, Óptica y Electrónica, INAOE, s.f. <https://ccc.inaoep.mx/esucar/Libros/vision-sucar-gomez.pdf>.
- [25] B. C. Calil, D. V. D. Cunha, M. F. Vieira, A. D. O. Andrade, D. A. Furtado, D. P. B. Junior, and A. A. Pereira, “Identification of arthropathy and myopathy of the temporomandibular syndrome by biomechanical facial features,” *BioMedical Engineering Online*, vol. 19, 4 2020.
- [26] R. M. Andrade, L. R. Guimarães, A. P. Ribeiro, A. P. Marques, O. Crivello, B. K. G. de Carvalho, and S. M. A. João, “Reliability in mandibular movement evaluation using photogrammetry in patients with temporomandibular disorders,” *Journal of Manipulative and Physiological Therapeutics*, vol. 42, pp. 267–275, 5 2019.
- [27] D. A. Furtado, A. A. Pereira, A. D. O. Andrade, D. Peres, B. Junior, and M. R. D. Silva, “A specialized motion capture system for real-time analysis of mandibular movements using infrared cameras,” 2013.
- [28] IBM, “¿qué es el aprendizaje supervisado?,” 2023. <https://www.ibm.com/mx-es/topics/supervised-learning>.
- [29] G. I. Webb, *Naïve Bayes*, pp. 895–896. Springer US, 2017.
- [30] D. G. Kleinbaum and M. Klein, *Introduction to Logistic Regression*, pp. 1–39. Springer New York, 2010.
- [31] V. Podgorelec and M. Zorman, *Decision Trees Decision tree*, pp. 1826–1845. Springer New York, 2009.
- [32] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” 5 2015.
- [33] F. Chollet, *Deep Learning with Python*. Manning Publications Co., 1st ed., 2017.
- [34] A. Sklar, “Fonctions de répartition á n dimensions et leurs marges.,” *Publications de l’Institut de Statistique de l’Université de Paris*, vol. 8, pp. 229–231, 1959.

- [35] N. Balakrishna and C. D. Lai, *Continuous bivariate distributions*. Springer New York, 2009.
- [36] Ivannkojadinovicc and M. Junnyan, *UseR ! Elements of Copula Modeling with R*. Springer, 2018. <http://www.springer.com/series/6991>.
- [37] G. A. F. Seber, *Multinomial Distribution*, pp. 882–884. Springer Berlin Heidelberg, 2011.
- [38] R. Nelsen, *An Introduction to Copulas*. Springer-Verlag, 1 2006.
- [39] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. R. Villa-Diharce, “Dependence trees with copula selection for continuous estimation of distribution algorithm,” in *GECCO '11: Proceedings of the 13th annual conference on Genetic and evolutionary computation*, pp. 585–592, ACM, 2011.
- [40] G. James, D. Witten, T. Hastie, and R. Tibshirani, “Springer texts in statistics an introduction to statistical learning,” *Springer*, 2013. <http://www.springer.com/series/417>.
- [41] M. García, J. A. Cabrera, A. Bataller, S. Postigo, and J. J. Castillo, “3d kinematic mandible model to design mandibular advancement devices for the treatment of obstructive sleep apnea,” *Bio-Design and Manufacturing*, vol. 4, pp. 22–32, 3 2021.
- [42] O. Röhrle, J. N. Waddell, K. D. Foster, H. Saini, and A. J. Pullan, “Using a motion-capture system to record dynamic articulation for application in cad/cam software,” *Journal of Prosthodontics*, vol. 18, pp. 703–710, 12 2009.
- [43] M. F. Laird, “Variation in human gape cycle kinematics and occlusal topography,” *American Journal of Physical Anthropology*, vol. 164, pp. 574–585, 11 2017.
- [44] R. Fuentes, P. Navarro, A. Curiqueo, and N. E. Ottone, “Determination of mandibular border and functional movement protocols using an electromagnetic articulograph (ema),” 2015.
- [45] D. Perpetuini, N. Trippetti, D. Cardone, L. Breda, M. D’Attilio, and A. Merla, “Detection of temporomandibular joint dysfunction in juvenile idiopathic arthritis through infrared thermal imaging and a machine learning procedure,” in *IFMBE Proceedings*, vol. 80, pp. 372–381, Springer Science and Business Media Deutschland GmbH, 2021.

- [46] R. B. Kakkeri and D. S. Bormane, “Analysis and prediction of temporomandibular joint disorder using machine learning classification algorithms,” in *Techno-Societal 2020* (P. M. Pawar, R. Balasubramaniam, B. P. Ronge, S. B. Salunkhe, A. S. Vibhute, and B. Melinamath, eds.), pp. 51–61, Springer International Publishing, 2021.
- [47] T. Sonmezocak and S. Kurt, “Machine learning and regression analysis for diagnosis of bruxism by using emg signals of jaw muscles,” *Biomedical Signal Processing and Control*, vol. 69, p. 102905, 8 2021.
- [48] M. Troka, W. Wojnicz, K. Szepietowska, M. Podlasiński, S. Walerzak, K. Walerzak, and I. Lubowiecka, “Towards classification of patients based on surface emg data of temporomandibular joint muscles using self-organising maps,” *Biomedical Signal Processing and Control*, vol. 72, p. 103322, 2 2022.
- [49] H. A. A and S. Kolahi, “Diagnosis of tempromandibular disorders using local binary patterns,” *J Biomed Phys Eng*, vol. 8, 2018.
- [50] U. Taşkıran and M. Çunkaş, “A deep learning based decision support system for diagnosis of temporomandibular joint disorder,” *Applied Acoustics*, vol. 182, 11 2021.
- [51] R. Salinas-Gutiérrez, A. Hernández-Aguirre, M. J. J. Rivera-Meraz, and E. R. Villa-Diharce, *Using Gaussian Copulas in Supervised Probabilistic Classification*. Springer, 2010.
- [52] K. Dowd, “Copulas in macroeconomics,” *Journal of International and global Economic Studies*, vol. 1, pp. 1–26, 7 2008.
- [53] C. Genest and A.-C. Favre, “Everything you always wanted to know about copula modeling but were afraid to ask,” *Journal of Hydrologic Engineering*, vol. 12, no. 4, pp. 347–368, 2007.
- [54] R. Salinas-Gutiérrez, A. Hernández-Aguirre, M. J. J. Rivera-Meraz, and E. R. Villa-Diharce, “Supervised probabilistic classification based on gaussian copulas,” in *LNAI*, pp. 104–115, Springer, 2010.
- [55] D. Carrera, R. Santana, and J. A. Lozano, “Vine copula classifiers for the mind reading problem,” *Progress in Artificial Intelligence*, vol. 5, pp. 289–305, 2016.

- [56] S. K. Ghalem, B. Kechar, A. Bounceur, and R. Euler, “A probabilistic multivariate copula-based technique for faulty node diagnosis in wireless sensor networks,” *Journal of Network and Computer Applications*, vol. 127, pp. 9–25, 2 2019.
- [57] L. A. Rodríguez-Picón, L. C. Méndez-González, V. H. Flores-Ochoa, and M. A. Rodríguez-Medina, “Capability indices for circular tolerance regions based on a gaussian copula,” *The International Journal of Advanced Manufacturing Technology*, vol. 104, pp. 4143–4153, 2019.
- [58] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. R. Villa-Diharce, “Using copulas in estimation of distribution algorithms,” in *LNAI 5845*, pp. 658–668, Springer, 2009.
- [59] S. Sen, N. Diawara, and K. M. Iftekharruddin, “Statistical pattern recognition using gaussian copula,” *Journal of Statistical Theory and Practice*, vol. 9, pp. 768–777, 10 2015.
- [60] K. Vogelaere, “Using truncated vine copulas in supervised probabilistic classification,” *Université catholique de Louvain*, 2020.
- [61] Ángela Paulina Pérez-Díaz, R. Salinas-Gutiérrez, A. Hernández-Quintero, and O. Dalmau-Cedeño, “Supervised classification based on copula functions,” in *Research in Computing Science 133*, pp. 11–18, Springer, 2017.
- [62] Y. Zhang, X. Wang, D. Liu, C. Li, Q. Liu, Y. Cai, Y. Yi, and Z. Yang, “Joint probability-based classifier based on vine copula method for land use classification of multispectral remote sensing data,” *Earth Science Informatics*, vol. 13, pp. 1079–1092, 2020.
- [63] C. Tamborrino and F. Mazzia, “On the classification of hyperspectral images with different copula family,” in *AIP Conference Proceedings*, vol. 2849, American Institute of Physics Inc., 9 2023.
- [64] X. Huang and Z. Wang, “Probabilistic spatial prediction of categorical data using elliptical copulas,” *Stochastic Environmental Research and Risk Assessment*, vol. 32, pp. 1631–1644, 2018.
- [65] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern,

- E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . Contributors, “Scipy 1.0: Fundamental algorithms for scientific computing in python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [66] I. Guyon and A. M. De, “An introduction to variable and feature selection andré elisseeff,” 2003.
- [67] P. Larranaga and J. Lozano, *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Springer New York, 1 2002.
- [68] F. Pedregosa, G. Varoquaux, A. Gramfort, B. M. V., Thirion, O. Grisel, M. Blondel, R. P. P., Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [69] J. B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proc. Amer. Math. Soc.*, vol. 7, pp. 48–50, 1956.
- [70] R. Salinas Gutiérrez, A. Hernández Quintero, P. A. Montoya Calzada, C. A. López Hernández, and J. M. Marquez Romero, *MLCOPULA: Classification Models with Copula Functions*, 2024. R package version 1.0.0.

Anexos

Anexo I: Gráficas ROC por Clasificador

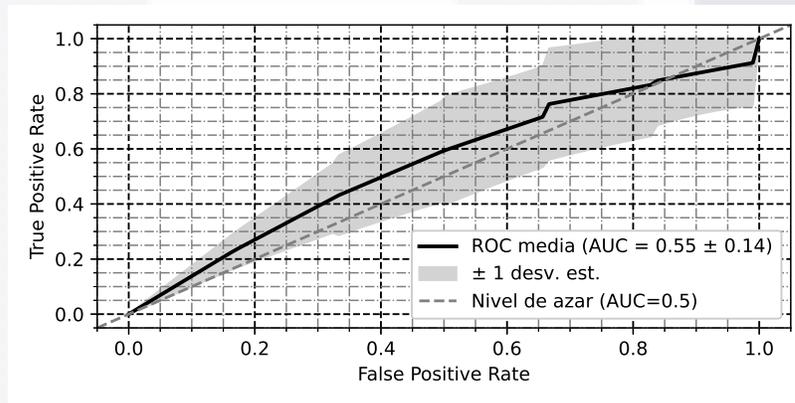


Figura 6.1: curva ROC para el clasificador con cópula Clayton con todos los atributos

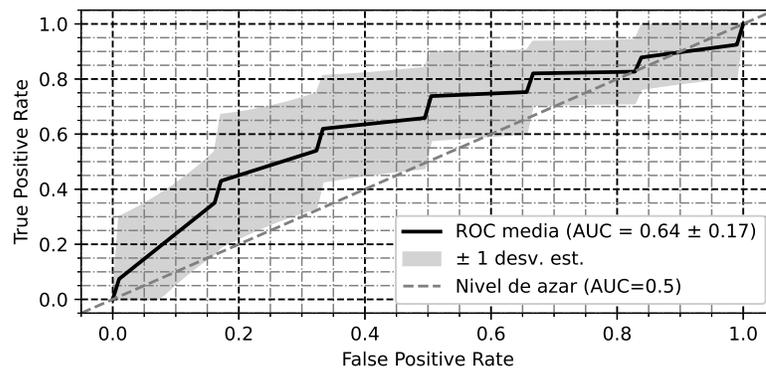


Figura 6.2: curva ROC para el clasificador con cópula Frank con todos los atributos

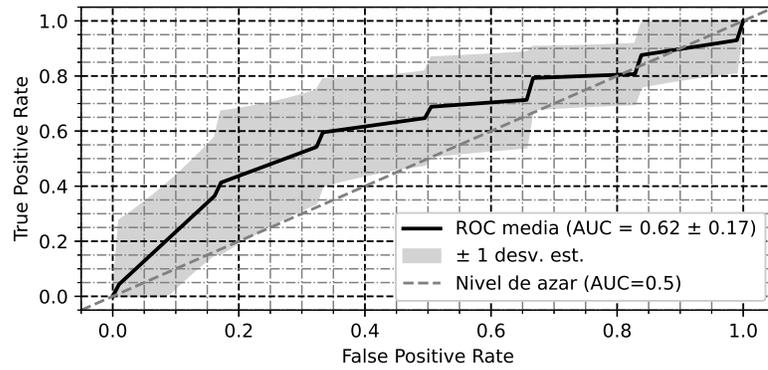


Figura 6.3: curva ROC para el clasificador con cópula Gaussiana con todos los atributos

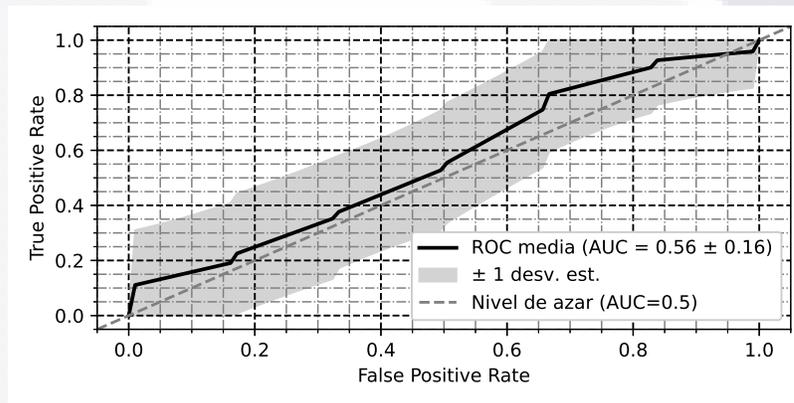


Figura 6.4: curva ROC para el clasificador KNN con todos los atributos

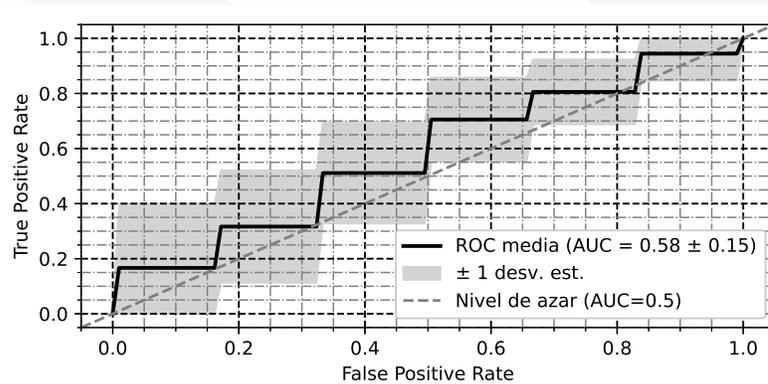


Figura 6.5: curva ROC para el clasificador NN con todos los atributos

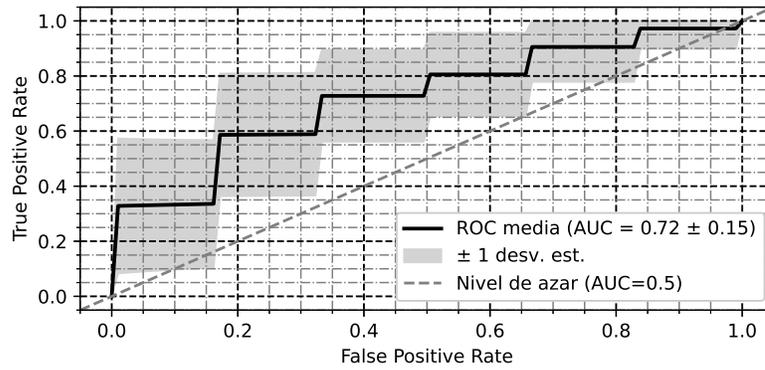


Figura 6.6: curva ROC para el clasificador con todos los atributos

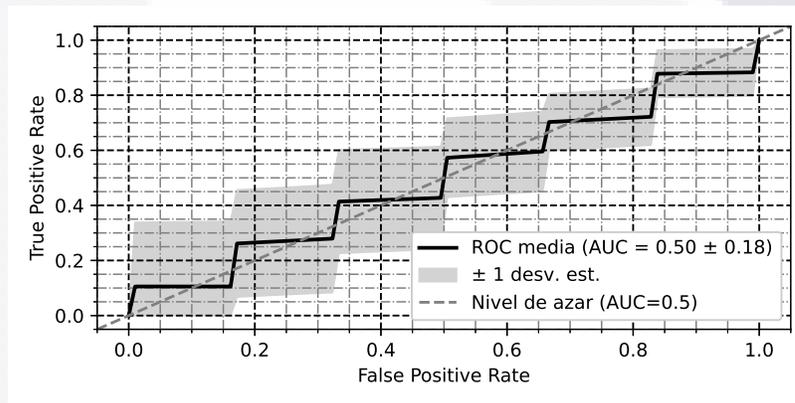


Figura 6.7: curva ROC para el clasificador SVM con todos los atributos

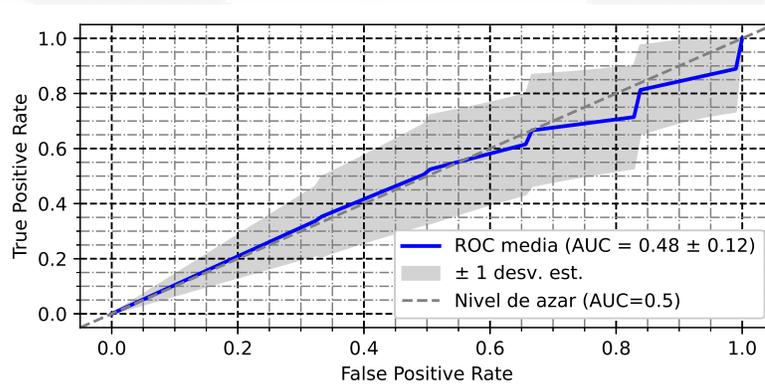


Figura 6.8: curva ROC para el clasificador con cópula Clayton con selección por correlación

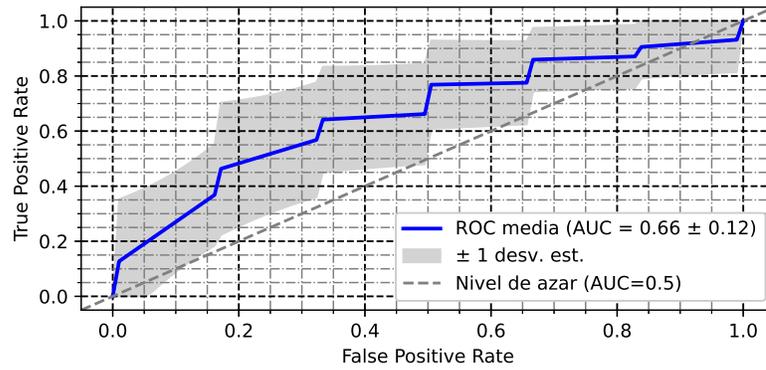


Figura 6.9: curva ROC para el clasificador con cópula Frank con selección por correlación

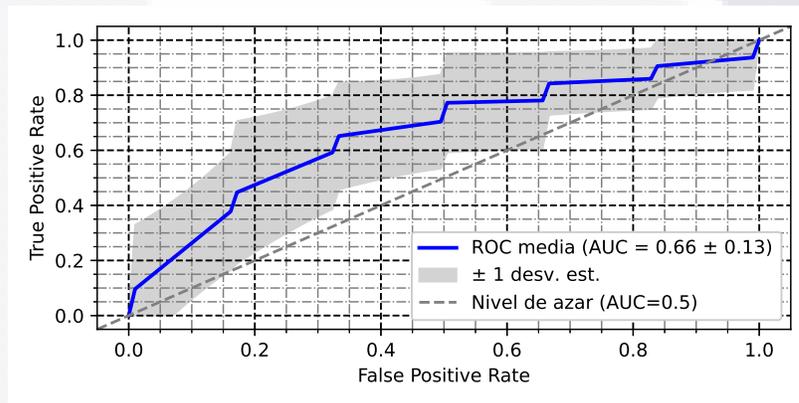


Figura 6.10: curva ROC para el clasificador con cópula Gaussiana con selección por correlación

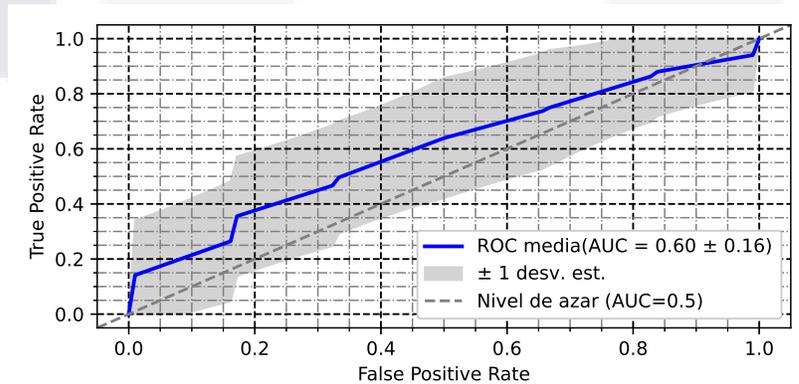


Figura 6.11: curva ROC para el clasificador KNN con selección por correlación

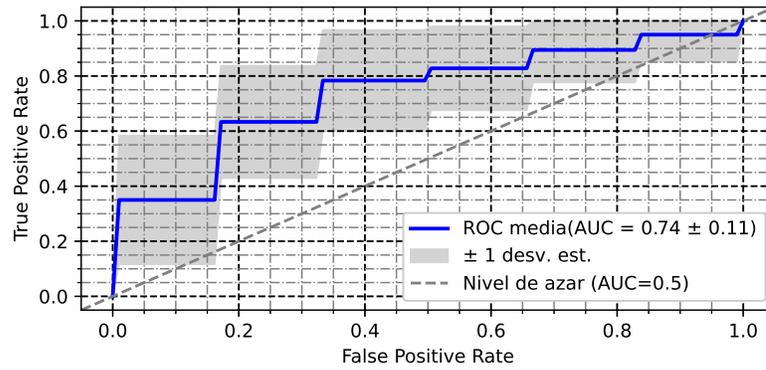


Figura 6.12: curva ROC para el clasificador NN con selección por correlación

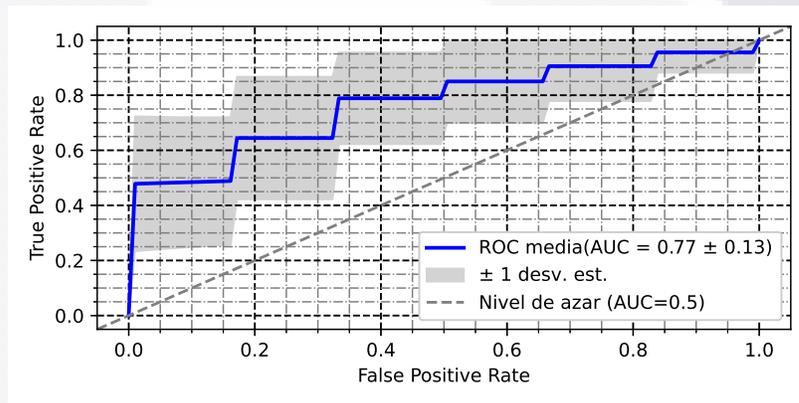


Figura 6.13: curva ROC para el clasificador con ingenuo de Bayes con selección por correlación

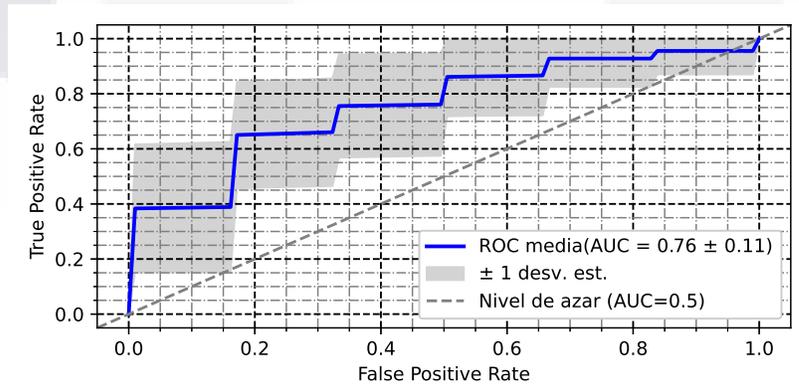


Figura 6.14: curva ROC para el clasificador SVM con selección por correlación

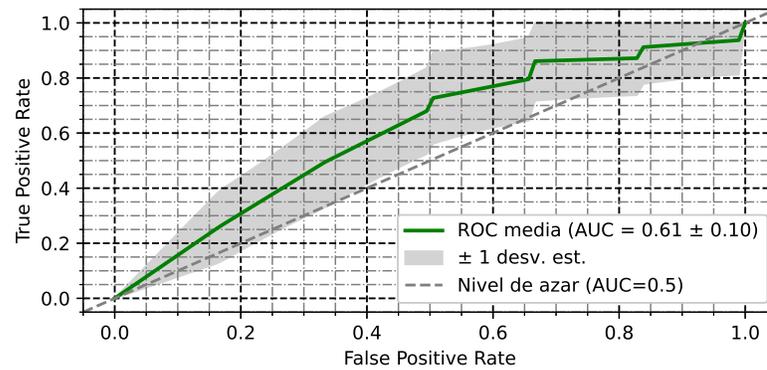


Figura 6.15: curva ROC para el clasificador con cópula Clayton con selección por EDA

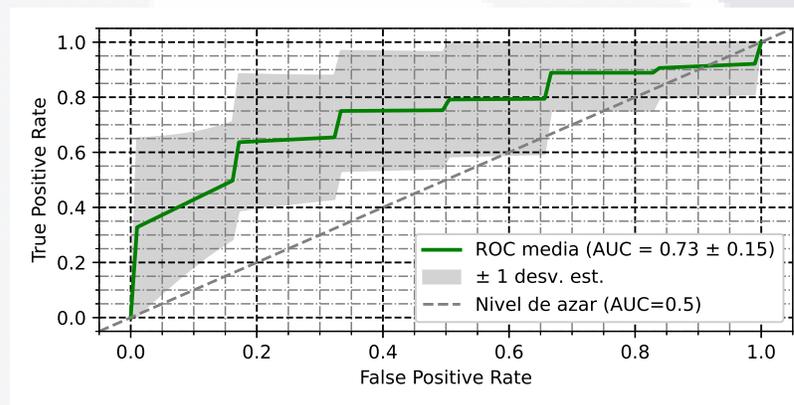


Figura 6.16: curva ROC para el clasificador con cópula Frank con selección por EDA

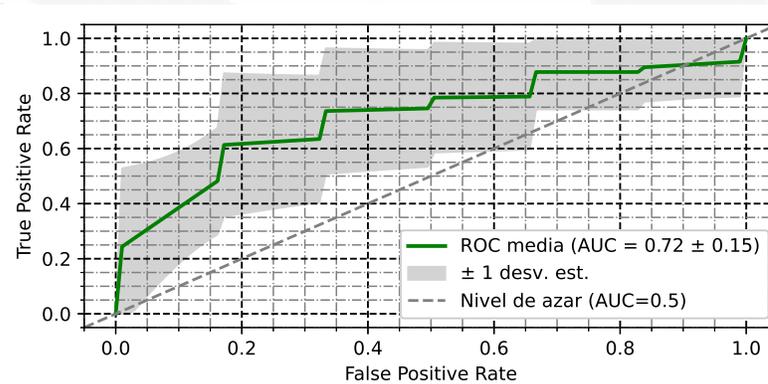


Figura 6.17: curva ROC para el clasificador con cópula Gaussiana con selección por EDA

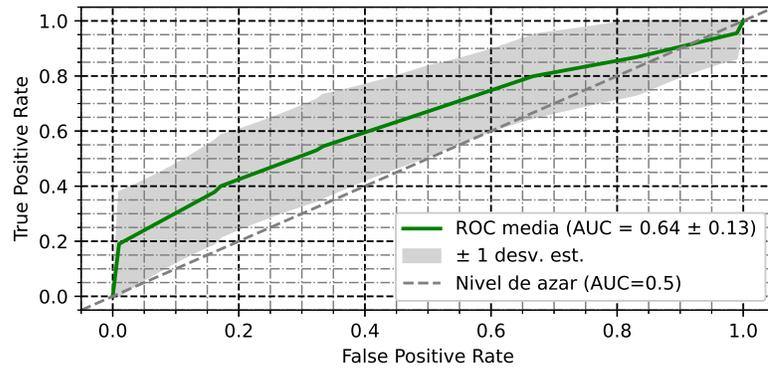


Figura 6.18: curva ROC para el clasificador KNN con selección por EDA

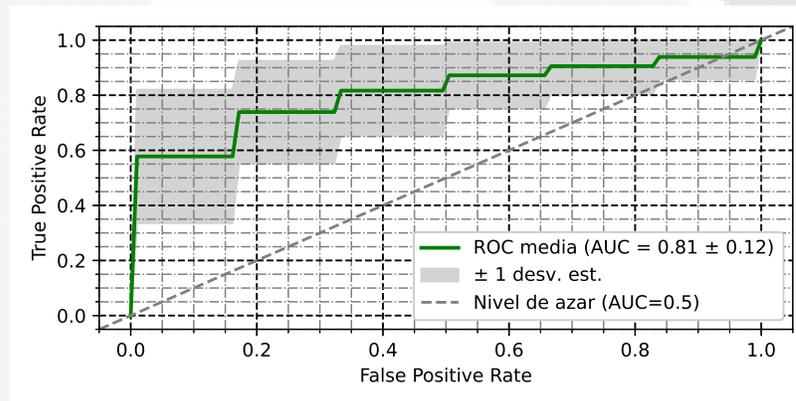


Figura 6.19: curva ROC para el clasificador NN con selección por EDA

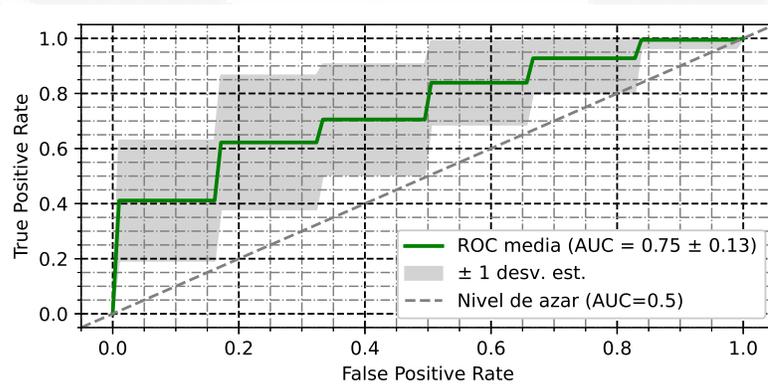


Figura 6.20: curva ROC para el clasificador ingenuo de bayes con selección por EDA

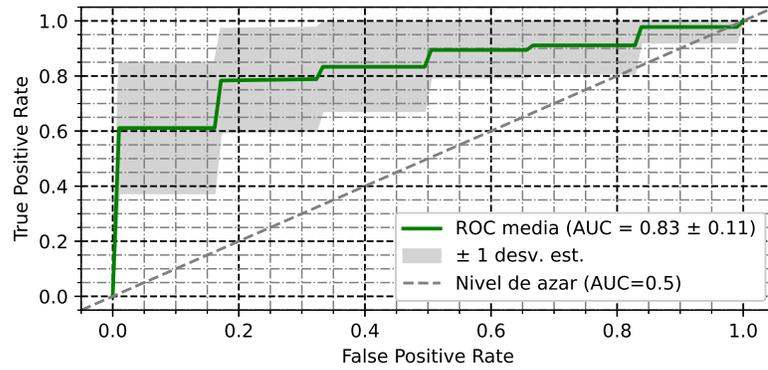


Figura 6.21: curva ROC para el clasificador SVM con selección por EDA

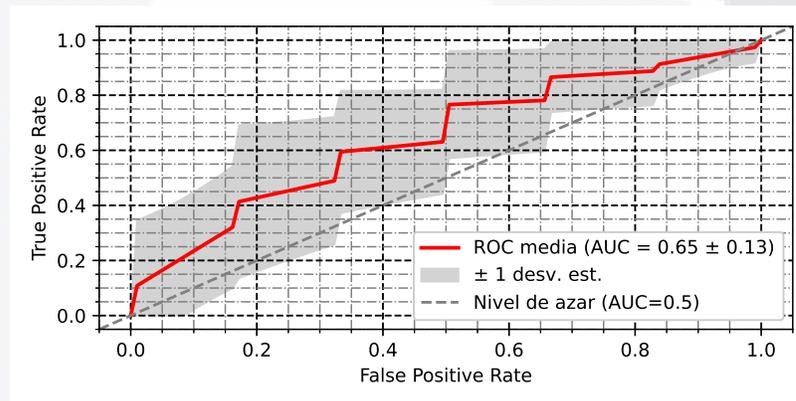


Figura 6.22: curva ROC para el clasificador con cópula Clayton con selección por correlación y EDA

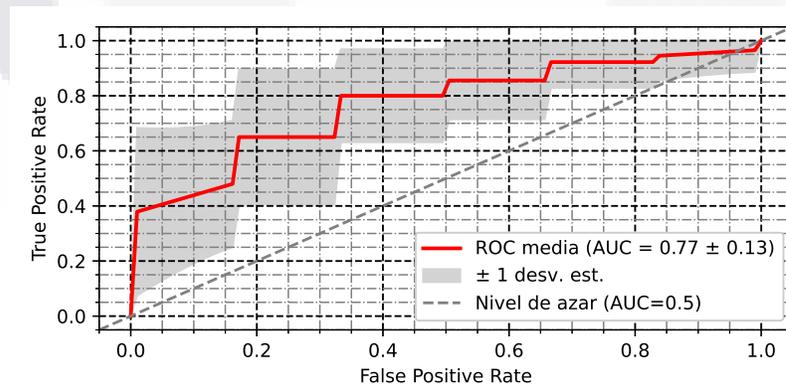


Figura 6.23: curva ROC para el clasificador con cópula Frank con selección por correlación y EDA

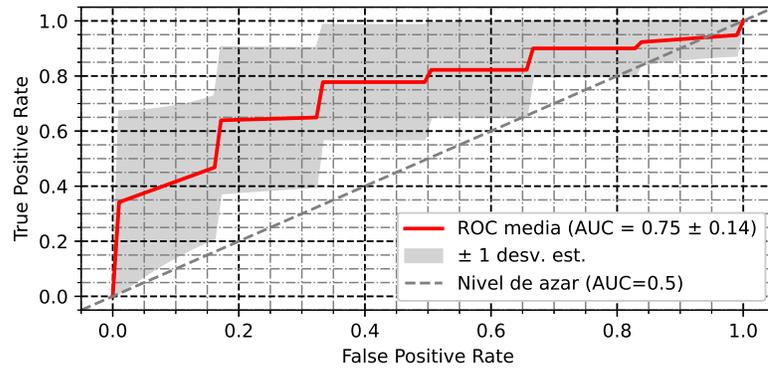


Figura 6.24: curva ROC para el clasificador con cópula Gaussiana con selección por correlación y EDA

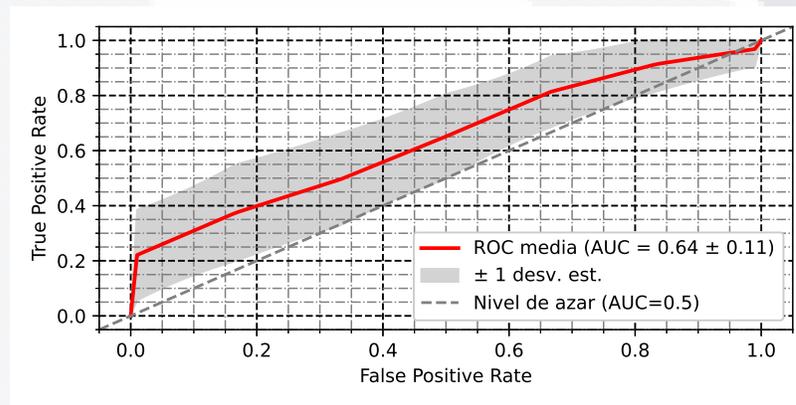


Figura 6.25: curva ROC para el clasificador KNN con selección por correlación y EDA

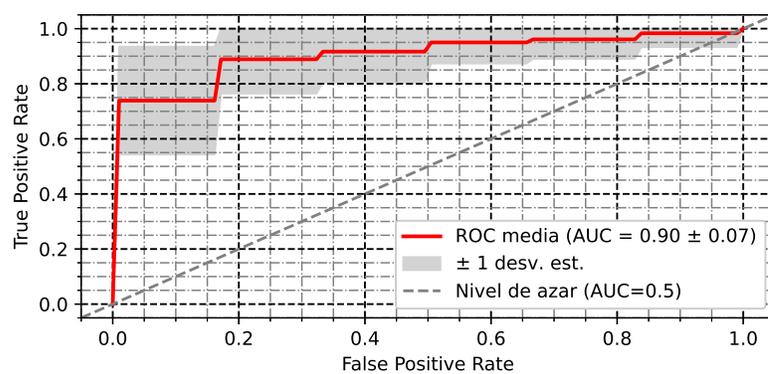


Figura 6.26: curva ROC para el clasificador NN con selección por correlación y EDA

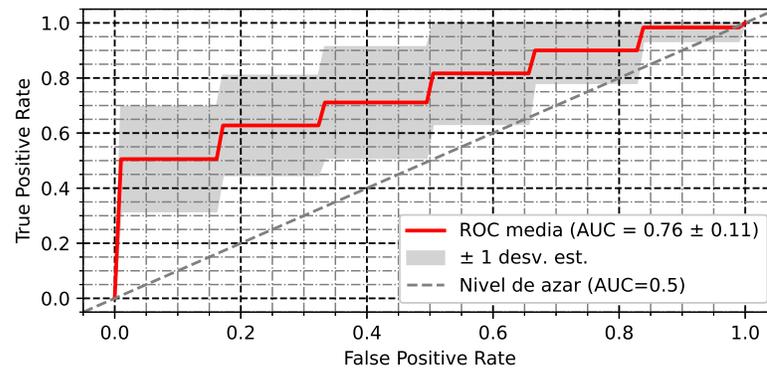


Figura 6.27: curva ROC para el clasificador con ingenuo de Bayes con selección por correlación y EDA

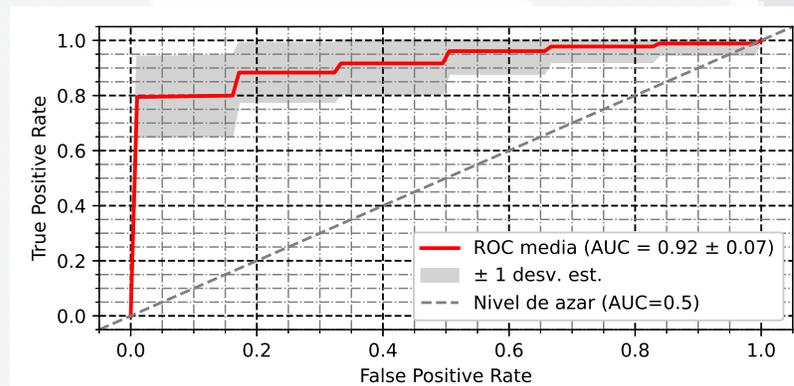


Figura 6.28: curva ROC para el clasificador SVM con selección por correlación y EDA

Anexo II: Resultados en Extenso

Tabla 6.1: Resultados en extenso de la clasificación sin selección de atributos.

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	1.0	0.5	0.8333	Naive Bayes
0.75	0.6667	0.8333	0.8611	Naive Bayes
0.75	0.8333	0.6667	0.9167	Naive Bayes
0.5833	0.8333	0.3333	0.7222	Naive Bayes
0.4167	0.6667	0.1667	0.5	Naive Bayes
0.5	0.5	0.5	0.5139	Naive Bayes

Continua en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	0.8333	0.6667	0.7222	Naive Bayes
0.6667	0.6667	0.6667	0.6944	Naive Bayes
0.5833	0.6667	0.5	0.5278	Naive Bayes
0.5833	0.8333	0.3333	0.6667	Naive Bayes
0.6667	0.6667	0.6667	0.8611	Naive Bayes
0.75	0.6667	0.8333	0.8056	Naive Bayes
0.75	0.6667	0.8333	0.8333	Naive Bayes
0.75	0.8333	0.6667	0.8611	Naive Bayes
0.8333	0.8333	0.8333	0.9167	Naive Bayes
0.5833	0.6667	0.5	0.6111	Naive Bayes
0.75	0.5	1.0	0.9444	Naive Bayes
0.5833	0.8333	0.3333	0.6389	Naive Bayes
0.5833	0.6667	0.5	0.6389	Naive Bayes
0.5833	0.8333	0.3333	0.5	Naive Bayes
0.75	0.8333	0.6667	0.8889	Naive Bayes
0.75	0.6667	0.8333	0.8889	Naive Bayes
0.6667	0.8333	0.5	0.6111	Naive Bayes
0.5833	0.1667	1.0	0.5833	Naive Bayes
0.6667	0.5	0.8333	0.5556	Naive Bayes
0.5	0.3333	0.6667	0.4722	Naive Bayes
0.8333	0.8333	0.8333	0.6944	Naive Bayes
0.75	0.6667	0.8333	0.7778	Naive Bayes
0.6667	0.5	0.8333	0.6667	Naive Bayes
0.75	1.0	0.5	0.9444	Naive Bayes
0.3333	0.3333	0.3333	0.5139	KNN
0.5	0.3333	0.6667	0.5972	KNN
0.6667	0.6667	0.6667	0.5833	KNN
0.6667	0.5	0.8333	0.7778	KNN
0.4167	0.3333	0.5	0.3194	KNN
0.5	0.5	0.5	0.5	KNN
0.4167	0.3333	0.5	0.375	KNN
0.6667	0.5	0.8333	0.7778	KNN

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5	0.1667	0.8333	0.5278	KNN
0.4167	0.5	0.3333	0.4861	KNN
0.6667	0.6667	0.6667	0.8056	KNN
0.8333	0.6667	1.0	0.9444	KNN
0.6667	0.3333	1.0	0.625	KNN
0.5833	0.6667	0.5	0.6389	KNN
0.5	0.5	0.5	0.6111	KNN
0.5833	0.5	0.6667	0.5556	KNN
0.6667	0.3333	1.0	0.6667	KNN
0.4167	0.5	0.3333	0.3194	KNN
0.4167	0.3333	0.5	0.5	KNN
0.4167	0.3333	0.5	0.4306	KNN
0.5	0.5	0.5	0.5278	KNN
0.5	0.3333	0.6667	0.5139	KNN
0.25	0.0	0.5	0.3333	KNN
0.3333	0.0	0.6667	0.4167	KNN
0.3333	0.1667	0.5	0.4306	KNN
0.5	0.3333	0.6667	0.3333	KNN
0.5833	0.3333	0.8333	0.7361	KNN
0.3333	0.0	0.6667	0.3889	KNN
0.4167	0.0	0.8333	0.7083	KNN
0.6667	0.8333	0.5	0.7917	KNN
0.6667	0.8333	0.5	0.2222	SVM
0.8333	0.8333	0.8333	0.8056	SVM
0.5	0.3333	0.6667	0.3889	SVM
0.9167	1.0	0.8333	0.0833	SVM
0.4167	0.5	0.3333	0.3889	SVM
0.3333	0.5	0.1667	0.4167	SVM
0.75	0.8333	0.6667	0.7222	SVM
0.5	0.6667	0.3333	0.5556	SVM
0.5	0.5	0.5	0.4722	SVM
0.5833	0.5	0.6667	0.5833	SVM

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.8333	0.8333	0.8333	0.8056	SVM
0.8333	0.6667	1.0	0.3472	SVM
0.75	0.6667	0.8333	0.5556	SVM
0.5833	0.5	0.6667	0.6389	SVM
0.75	0.8333	0.6667	0.2778	SVM
0.4167	0.3333	0.5	0.5278	SVM
0.5833	0.6667	0.5	0.7778	SVM
0.5	0.6667	0.3333	0.4722	SVM
0.8333	0.8333	0.8333	0.8194	SVM
0.5	0.5	0.5	0.5556	SVM
0.6667	0.6667	0.6667	0.375	SVM
0.5	0.3333	0.6667	0.4444	SVM
0.5	0.5	0.5	0.4722	SVM
0.5833	0.3333	0.8333	0.5	SVM
0.5833	0.5	0.6667	0.3889	SVM
0.3333	0.1667	0.5	0.4722	SVM
0.4167	0.3333	0.5	0.5833	SVM
0.6667	0.5	0.8333	0.5833	SVM
0.4167	0.1667	0.6667	0.4444	SVM
0.75	0.8333	0.6667	0.1944	SVM
0.5833	0.6667	0.5	0.6111	NN
0.6667	0.6667	0.6667	0.7778	NN
0.5	0.5	0.5	0.5278	NN
0.6667	0.6667	0.6667	0.75	NN
0.4167	0.5	0.3333	0.3889	NN
0.4167	0.3333	0.5	0.3889	NN
0.5833	0.6667	0.5	0.5556	NN
0.5833	0.6667	0.5	0.6389	NN
0.5833	0.5	0.6667	0.5278	NN
0.4167	0.6667	0.1667	0.4722	NN
0.8333	1.0	0.6667	0.8333	NN
0.9167	0.8333	1.0	0.9167	NN

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.4167	0.1667	0.6667	0.5556	NN
0.5	0.5	0.5	0.4722	NN
0.6667	0.6667	0.6667	0.7222	NN
0.5833	0.6667	0.5	0.4722	NN
0.6667	0.6667	0.6667	0.8611	NN
0.4167	0.5	0.3333	0.4444	NN
0.4167	0.5	0.3333	0.5	NN
0.5	0.3333	0.6667	0.4444	NN
0.5833	0.6667	0.5	0.5	NN
0.5833	0.5	0.6667	0.5278	NN
0.5	0.5	0.5	0.3889	NN
0.6667	0.5	0.8333	0.5833	NN
0.5833	0.6667	0.5	0.5556	NN
0.4167	0.1667	0.6667	0.3333	NN
0.5	0.3333	0.6667	0.6111	NN
0.5	0.3333	0.6667	0.5	NN
0.5	0.1667	0.8333	0.5278	NN
0.6667	0.8333	0.5	0.8611	NN
0.4167	0.5	0.3333	0.5833	Frank Copula
0.6667	0.6667	0.6667	0.6944	Frank Copula
0.8333	1.0	0.6667	0.8889	Frank Copula
0.4167	0.6667	0.1667	0.3056	Frank Copula
0.3333	0.5	0.1667	0.3056	Frank Copula
0.5833	0.8333	0.3333	0.6111	Frank Copula
0.75	0.8333	0.6667	0.8056	Frank Copula
0.8333	1.0	0.6667	0.875	Frank Copula
0.5833	0.6667	0.5	0.4444	Frank Copula
0.75	0.8333	0.6667	0.75	Frank Copula
0.5833	0.8333	0.3333	0.6389	Frank Copula
0.4167	0.8333	0.0	0.5278	Frank Copula
0.8333	1.0	0.6667	0.875	Frank Copula
0.5	0.6667	0.3333	0.3333	Frank Copula

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5833	0.6667	0.5	0.5556	Frank Copula
0.5833	0.8333	0.3333	0.6667	Frank Copula
0.8333	0.8333	0.8333	0.8056	Frank Copula
0.8333	1.0	0.6667	0.7778	Frank Copula
0.5833	0.8333	0.3333	0.6389	Frank Copula
0.5833	0.6667	0.5	0.5556	Frank Copula
0.8333	0.8333	0.8333	0.8333	Frank Copula
0.9167	1.0	0.8333	0.8889	Frank Copula
0.6667	0.8333	0.5	0.5972	Frank Copula
0.6667	0.3333	1.0	0.5833	Frank Copula
0.5833	0.3333	0.8333	0.4028	Frank Copula
0.5	0.5	0.5	0.4583	Frank Copula
0.5833	0.6667	0.5	0.5417	Frank Copula
0.5833	0.6667	0.5	0.625	Frank Copula
0.6667	0.6667	0.6667	0.6111	Frank Copula
0.8333	1.0	0.6667	0.8472	Frank Copula
0.4167	0.6667	0.1667	0.4722	Clayton Copula
0.6667	0.8333	0.5	0.6528	Clayton Copula
0.5	0.8333	0.1667	0.5556	Clayton Copula
0.5	1.0	0.0	0.4167	Clayton Copula
0.4167	0.8333	0.0	0.4167	Clayton Copula
0.6667	0.6667	0.6667	0.6389	Clayton Copula
0.8333	0.8333	0.8333	0.8194	Clayton Copula
0.6667	1.0	0.3333	0.6667	Clayton Copula
0.5	0.8333	0.1667	0.4861	Clayton Copula
0.5833	1.0	0.1667	0.6667	Clayton Copula
0.5833	1.0	0.1667	0.5833	Clayton Copula
0.5	1.0	0.0	0.5	Clayton Copula
0.5	0.8333	0.1667	0.5694	Clayton Copula
0.5	0.8333	0.1667	0.4861	Clayton Copula
0.6667	1.0	0.3333	0.6111	Clayton Copula
0.5833	1.0	0.1667	0.5833	Clayton Copula

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5	0.6667	0.3333	0.5	Clayton Copula
0.5	0.6667	0.3333	0.5556	Clayton Copula
0.5	1.0	0.0	0.5	Clayton Copula
0.3333	0.6667	0.0	0.3333	Clayton Copula
0.6667	1.0	0.3333	0.6667	Clayton Copula
0.3333	0.5	0.1667	0.4028	Clayton Copula
0.5	0.8333	0.1667	0.5	Clayton Copula
0.75	0.8333	0.6667	0.7361	Clayton Copula
0.5833	0.8333	0.3333	0.5556	Clayton Copula
0.4167	0.6667	0.1667	0.3889	Clayton Copula
0.1667	0.3333	0.0	0.1667	Clayton Copula
0.5833	0.8333	0.3333	0.6389	Clayton Copula
0.5	0.6667	0.3333	0.5833	Clayton Copula
0.8333	1.0	0.6667	0.8333	Clayton Copula
0.5	0.5	0.5	0.5139	Gaussian Copula
0.6667	0.6667	0.6667	0.6111	Gaussian Copula
0.75	0.8333	0.6667	0.7222	Gaussian Copula
0.4167	0.6667	0.1667	0.3194	Gaussian Copula
0.25	0.5	0.0	0.2222	Gaussian Copula
0.6667	1.0	0.3333	0.6667	Gaussian Copula
0.75	0.8333	0.6667	0.7361	Gaussian Copula
0.75	1.0	0.5	0.9028	Gaussian Copula
0.5833	0.5	0.6667	0.4444	Gaussian Copula
0.5833	0.6667	0.5	0.6528	Gaussian Copula
0.6667	1.0	0.3333	0.6806	Gaussian Copula
0.4167	0.6667	0.1667	0.5	Gaussian Copula
0.75	1.0	0.5	0.7083	Gaussian Copula
0.5	0.6667	0.3333	0.3889	Gaussian Copula
0.6667	0.8333	0.5	0.6667	Gaussian Copula
0.5	0.6667	0.3333	0.6667	Gaussian Copula
0.8333	0.8333	0.8333	0.8056	Gaussian Copula
0.75	0.8333	0.6667	0.7222	Gaussian Copula

Continúa en la siguiente página

Tabla 6.1 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.6667	0.8333	0.5	0.7222	Gaussian Copula
0.4167	0.5	0.3333	0.4306	Gaussian Copula
0.8333	0.8333	0.8333	0.8333	Gaussian Copula
0.6667	1.0	0.3333	0.8889	Gaussian Copula
0.5833	0.6667	0.5	0.5972	Gaussian Copula
0.5833	0.3333	0.8333	0.5972	Gaussian Copula
0.6667	0.3333	1.0	0.4583	Gaussian Copula
0.4167	0.1667	0.6667	0.4306	Gaussian Copula
0.5833	0.6667	0.5	0.5972	Gaussian Copula
0.5	0.6667	0.3333	0.5139	Gaussian Copula
0.5833	0.5	0.6667	0.625	Gaussian Copula
0.8333	1.0	0.6667	0.9028	Gaussian Copula

Tabla 6.2: Resultados en extenso de la clasificación con selección por Correlación de Pearson.

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5833	0.8333	0.3333	0.75	Naive Bayes
0.6667	0.5	0.8333	0.75	Naive Bayes
0.6667	0.6667	0.6667	0.8333	Naive Bayes
0.5	0.5	0.5	0.6111	Naive Bayes
0.8333	0.6667	1.0	0.7778	Naive Bayes
0.75	0.8333	0.6667	0.8333	Naive Bayes
0.8333	1.0	0.6667	0.8889	Naive Bayes
0.6667	0.6667	0.6667	0.6389	Naive Bayes
0.8333	0.8333	0.8333	0.9167	Naive Bayes
1.0	1.0	1.0	1.0	Naive Bayes
0.75	0.8333	0.6667	0.8889	Naive Bayes
0.8333	1.0	0.6667	0.8611	Naive Bayes
0.6667	0.6667	0.6667	0.6667	Naive Bayes
0.75	0.8333	0.6667	0.8056	Naive Bayes

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	0.6667	0.8333	0.8889	Naive Bayes
0.8333	0.8333	0.8333	0.8889	Naive Bayes
0.4167	0.6667	0.1667	0.4722	Naive Bayes
0.75	0.6667	0.8333	0.7222	Naive Bayes
0.6667	0.6667	0.6667	0.75	Naive Bayes
0.5833	0.3333	0.8333	0.8333	Naive Bayes
0.8333	0.6667	1.0	0.8056	Naive Bayes
0.6667	0.6667	0.6667	0.8056	Naive Bayes
0.5	0.5	0.5	0.5833	Naive Bayes
0.8333	0.8333	0.8333	0.8611	Naive Bayes
0.5833	0.5	0.6667	0.4444	Naive Bayes
0.75	0.8333	0.6667	0.8333	Naive Bayes
0.6667	0.6667	0.6667	0.6111	Naive Bayes
0.6667	0.8333	0.5	0.75	Naive Bayes
0.8333	0.8333	0.8333	0.8889	Naive Bayes
0.75	0.6667	0.8333	0.7778	Naive Bayes
0.4167	0.5	0.3333	0.4167	KNN
0.6667	0.6667	0.6667	0.6944	KNN
0.8333	0.6667	1.0	0.8611	KNN
0.5	0.5	0.5	0.4861	KNN
0.5	0.3333	0.6667	0.5417	KNN
0.6667	0.5	0.8333	0.6111	KNN
0.5833	0.5	0.6667	0.7083	KNN
0.6667	0.6667	0.6667	0.6111	KNN
0.4167	0.5	0.3333	0.5417	KNN
0.75	0.6667	0.8333	0.7917	KNN
0.5	0.5	0.5	0.5833	KNN
0.75	0.6667	0.8333	0.8333	KNN
0.5833	0.3333	0.8333	0.4722	KNN
0.5833	0.5	0.6667	0.6806	KNN
0.75	0.6667	0.8333	0.75	KNN
0.5833	0.5	0.6667	0.6806	KNN

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5833	0.5	0.6667	0.5417	KNN
0.5833	0.3333	0.8333	0.5278	KNN
0.6667	0.5	0.8333	0.6944	KNN
0.5833	0.5	0.6667	0.5833	KNN
0.5833	0.5	0.6667	0.5278	KNN
0.3333	0.3333	0.3333	0.3889	KNN
0.5	0.6667	0.3333	0.4861	KNN
0.6667	0.5	0.8333	0.7778	KNN
0.5833	0.5	0.6667	0.5972	KNN
0.5	0.3333	0.6667	0.4722	KNN
0.5	0.3333	0.6667	0.5278	KNN
0.5	0.5	0.5	0.4722	KNN
0.9167	0.8333	1.0	0.9583	KNN
0.4167	0.0	0.8333	0.1389	KNN
0.5	0.6667	0.3333	0.6389	SVM
0.4167	0.8333	0.0	0.8056	SVM
0.75	0.6667	0.8333	0.6944	SVM
0.75	0.8333	0.6667	0.7778	SVM
0.6667	0.5	0.8333	0.7361	SVM
0.75	0.8333	0.6667	0.8472	SVM
0.6667	0.6667	0.6667	0.75	SVM
0.75	0.6667	0.8333	0.8333	SVM
0.75	1.0	0.5	0.7222	SVM
0.75	0.8333	0.6667	0.8611	SVM
0.75	0.8333	0.6667	0.7778	SVM
0.9167	1.0	0.8333	1.0	SVM
0.75	0.8333	0.6667	0.6944	SVM
0.6667	0.6667	0.6667	0.7778	SVM
0.6667	0.5	0.8333	0.7222	SVM
0.6667	0.5	0.8333	0.75	SVM
0.4167	0.5	0.3333	0.5833	SVM
0.6667	0.5	0.8333	0.6389	SVM

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.6667	1.0	0.3333	0.8472	SVM
0.6667	0.6667	0.6667	0.8611	SVM
0.5	0.5	0.5	0.5	SVM
0.75	0.6667	0.8333	0.8611	SVM
0.75	0.8333	0.6667	0.8611	SVM
0.8333	1.0	0.6667	0.8611	SVM
0.6667	0.6667	0.6667	0.5833	SVM
0.5833	0.8333	0.3333	0.8611	SVM
0.5833	0.6667	0.5	0.6667	SVM
0.75	0.6667	0.8333	0.9167	SVM
0.6667	0.6667	0.6667	0.75	SVM
0.5	0.3333	0.6667	0.5556	SVM
0.6667	0.8333	0.5	0.6111	NN
0.8333	0.8333	0.8333	0.7778	NN
0.5833	0.5	0.6667	0.7222	NN
0.5	0.6667	0.3333	0.6389	NN
0.6667	0.6667	0.6667	0.6944	NN
0.9167	1.0	0.8333	0.8333	NN
0.6667	0.6667	0.6667	0.7222	NN
0.6667	0.6667	0.6667	0.7222	NN
0.75	1.0	0.5	0.8333	NN
0.8333	1.0	0.6667	0.7778	NN
0.8333	1.0	0.6667	0.7222	NN
0.75	1.0	0.5	0.8611	NN
0.75	0.8333	0.6667	0.7222	NN
0.75	0.8333	0.6667	0.7778	NN
0.6667	0.5	0.8333	0.7778	NN
0.6667	0.5	0.8333	0.7222	NN
0.5833	0.6667	0.5	0.5556	NN
0.75	0.6667	0.8333	0.75	NN
0.4167	0.6667	0.1667	0.7222	NN
0.6667	0.6667	0.6667	0.75	NN

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.4167	0.5	0.3333	0.5556	NN
0.75	0.6667	0.8333	0.8056	NN
0.6667	0.8333	0.5	0.6667	NN
0.8333	1.0	0.6667	0.8889	NN
0.5833	0.5	0.6667	0.4722	NN
0.5833	0.8333	0.3333	0.8056	NN
0.6667	0.6667	0.6667	0.7778	NN
0.8333	0.8333	0.8333	0.9167	NN
0.9167	1.0	0.8333	0.9722	NN
0.5833	0.3333	0.8333	0.6389	NN
0.5833	1.0	0.1667	0.75	Frank Copula
0.5833	0.5	0.6667	0.6111	Frank Copula
0.5833	0.6667	0.5	0.625	Frank Copula
0.5833	0.6667	0.5	0.625	Frank Copula
0.75	0.8333	0.6667	0.75	Frank Copula
0.5	1.0	0.0	0.6667	Frank Copula
0.8333	1.0	0.6667	0.8056	Frank Copula
0.4167	0.3333	0.5	0.4167	Frank Copula
0.5833	0.8333	0.3333	0.6389	Frank Copula
0.6667	1.0	0.3333	0.6667	Frank Copula
0.75	0.8333	0.6667	0.8194	Frank Copula
0.75	1.0	0.5	0.6667	Frank Copula
0.6667	0.8333	0.5	0.6111	Frank Copula
0.5833	0.8333	0.3333	0.7222	Frank Copula
0.8333	1.0	0.6667	0.7778	Frank Copula
0.75	0.6667	0.8333	0.7778	Frank Copula
0.5	0.6667	0.3333	0.4583	Frank Copula
0.5	0.6667	0.3333	0.5	Frank Copula
0.5	0.6667	0.3333	0.6111	Frank Copula
0.8333	0.8333	0.8333	0.8611	Frank Copula
0.6667	0.8333	0.5	0.7639	Frank Copula
0.6667	0.8333	0.5	0.7222	Frank Copula

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5	0.5	0.5	0.4444	Frank Copula
0.6667	0.6667	0.6667	0.625	Frank Copula
0.5833	0.5	0.6667	0.5278	Frank Copula
0.5833	0.6667	0.5	0.7361	Frank Copula
0.75	1.0	0.5	0.7083	Frank Copula
0.3333	0.6667	0.0	0.5417	Frank Copula
0.8333	0.6667	1.0	0.8333	Frank Copula
0.5833	0.6667	0.5	0.5833	Frank Copula
0.4167	0.6667	0.1667	0.4444	Clayton Copula
0.6667	0.6667	0.6667	0.6944	Clayton Copula
0.4167	0.5	0.3333	0.4167	Clayton Copula
0.3333	0.6667	0.0	0.3333	Clayton Copula
0.3333	0.5	0.1667	0.3889	Clayton Copula
0.6667	1.0	0.3333	0.6111	Clayton Copula
0.6667	0.8333	0.5	0.625	Clayton Copula
0.4167	0.3333	0.5	0.3472	Clayton Copula
0.25	0.5	0.0	0.25	Clayton Copula
0.4167	0.6667	0.1667	0.4861	Clayton Copula
0.5833	0.8333	0.3333	0.5	Clayton Copula
0.5	0.8333	0.1667	0.5	Clayton Copula
0.5	0.8333	0.1667	0.5556	Clayton Copula
0.5833	0.8333	0.3333	0.6806	Clayton Copula
0.5833	0.8333	0.3333	0.6111	Clayton Copula
0.3333	0.3333	0.3333	0.2778	Clayton Copula
0.5833	0.8333	0.3333	0.5694	Clayton Copula
0.4167	0.5	0.3333	0.3889	Clayton Copula
0.25	0.5	0.0	0.2917	Clayton Copula
0.5833	0.6667	0.5	0.5972	Clayton Copula
0.5	0.6667	0.3333	0.5556	Clayton Copula
0.5833	0.8333	0.3333	0.5278	Clayton Copula
0.3333	0.6667	0.0	0.5	Clayton Copula
0.5	0.6667	0.3333	0.4306	Clayton Copula

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5833	0.6667	0.5	0.5972	Clayton Copula
0.5	0.8333	0.1667	0.4444	Clayton Copula
0.5833	0.8333	0.3333	0.625	Clayton Copula
0.3333	0.5	0.1667	0.3472	Clayton Copula
0.5	0.8333	0.1667	0.5139	Clayton Copula
0.4167	0.5	0.3333	0.3611	Clayton Copula
0.5	1.0	0.0	0.7222	Gaussian Copula
0.5833	0.5	0.6667	0.625	Gaussian Copula
0.5833	0.6667	0.5	0.5139	Gaussian Copula
0.5	0.8333	0.1667	0.6389	Gaussian Copula
0.75	0.8333	0.6667	0.7222	Gaussian Copula
0.5	1.0	0.0	0.5139	Gaussian Copula
0.75	0.8333	0.6667	0.7222	Gaussian Copula
0.4167	0.5	0.3333	0.4444	Gaussian Copula
0.5833	0.8333	0.3333	0.6111	Gaussian Copula
0.6667	1.0	0.3333	0.7222	Gaussian Copula
0.75	0.8333	0.6667	0.7778	Gaussian Copula
0.75	1.0	0.5	0.7083	Gaussian Copula
0.6667	0.8333	0.5	0.6389	Gaussian Copula
0.5	0.8333	0.1667	0.6528	Gaussian Copula
0.6667	1.0	0.3333	0.75	Gaussian Copula
0.6667	0.6667	0.6667	0.8056	Gaussian Copula
0.5833	0.6667	0.5	0.5972	Gaussian Copula
0.5	0.6667	0.3333	0.5278	Gaussian Copula
0.4167	0.6667	0.1667	0.5833	Gaussian Copula
0.75	0.8333	0.6667	0.8611	Gaussian Copula
0.75	0.8333	0.6667	0.7639	Gaussian Copula
0.6667	0.8333	0.5	0.6667	Gaussian Copula
0.3333	0.3333	0.3333	0.3472	Gaussian Copula
0.6667	0.8333	0.5	0.7778	Gaussian Copula
0.5833	0.5	0.6667	0.5833	Gaussian Copula
0.6667	0.6667	0.6667	0.6667	Gaussian Copula

Continúa en la siguiente página

Tabla 6.2 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	1.0	0.5	0.7083	Gaussian Copula
0.5	0.8333	0.1667	0.7361	Gaussian Copula
0.9167	1.0	0.8333	1.0	Gaussian Copula
0.5	0.5	0.5	0.5278	Gaussian Copula

Tabla 6.3: Resultados en extenso de la clasificación con selección por EDA

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.8333	0.8333	0.8333	0.9167	Naive Bayes
0.6667	0.8333	0.5	0.8333	Naive Bayes
0.5833	0.3333	0.8333	0.7222	Naive Bayes
0.75	0.6667	0.8333	0.75	Naive Bayes
0.5833	0.5	0.6667	0.7778	Naive Bayes
0.75	0.8333	0.6667	0.8611	Naive Bayes
0.6667	0.8333	0.5	0.6944	Naive Bayes
0.5833	0.6667	0.5	0.6944	Naive Bayes
0.5	0.5	0.5	0.4444	Naive Bayes
0.6667	0.8333	0.5	0.8056	Naive Bayes
0.5833	0.8333	0.3333	0.7778	Naive Bayes
0.75	0.6667	0.8333	0.8889	Naive Bayes
0.6667	1.0	0.3333	0.75	Naive Bayes
0.8333	0.8333	0.8333	0.9167	Naive Bayes
0.8333	0.8333	0.8333	0.8333	Naive Bayes
0.75	1.0	0.5	0.9167	Naive Bayes
0.75	0.8333	0.6667	0.75	Naive Bayes
0.8333	1.0	0.6667	0.9444	Naive Bayes
0.5833	0.8333	0.3333	0.6389	Naive Bayes
0.5833	0.6667	0.5	0.5833	Naive Bayes
0.5	0.6667	0.3333	0.5556	Naive Bayes
0.75	0.8333	0.6667	0.7778	Naive Bayes
0.5833	0.5	0.6667	0.6389	Naive Bayes

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.6667	0.5	0.8333	0.7222	Naive Bayes
0.8333	0.8333	0.8333	0.7778	Naive Bayes
0.6667	0.8333	0.5	0.7222	Naive Bayes
0.4167	0.5	0.3333	0.5278	Naive Bayes
0.8333	0.8333	0.8333	0.8611	Naive Bayes
0.75	0.5	1.0	0.8611	Naive Bayes
0.5833	0.3333	0.8333	0.5556	Naive Bayes
0.6667	0.5	0.8333	0.75	KNN
0.6667	0.3333	1.0	0.7778	KNN
0.5	0.0	1.0	0.6389	KNN
0.5	0.0	1.0	0.4167	KNN
0.5833	0.1667	1.0	0.6806	KNN
0.5833	0.1667	1.0	0.6806	KNN
0.5833	0.3333	0.8333	0.5417	KNN
0.75	0.5	1.0	0.6667	KNN
0.5	0.3333	0.6667	0.4861	KNN
0.5833	0.1667	1.0	0.6528	KNN
0.5	0.3333	0.6667	0.5972	KNN
0.6667	0.5	0.8333	0.6806	KNN
0.6667	0.6667	0.6667	0.7778	KNN
0.75	0.5	1.0	0.8472	KNN
0.6667	0.5	0.8333	0.5972	KNN
0.5833	0.1667	1.0	0.7917	KNN
0.5833	0.5	0.6667	0.6667	KNN
0.75	0.6667	0.8333	0.8194	KNN
0.5833	0.3333	0.8333	0.6667	KNN
0.4167	0.0	0.8333	0.375	KNN
0.5833	0.3333	0.8333	0.6667	KNN
0.5833	0.1667	1.0	0.7222	KNN
0.5833	0.1667	1.0	0.4444	KNN
0.5833	0.3333	0.8333	0.6667	KNN
0.5833	0.3333	0.8333	0.4722	KNN

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5	0.3333	0.6667	0.4167	KNN
0.5	0.3333	0.6667	0.6111	KNN
0.75	0.5	1.0	0.7778	KNN
0.6667	0.3333	1.0	0.7917	KNN
0.4167	0.0	0.8333	0.5278	KNN
0.9167	1.0	0.8333	1.0	SVM
0.75	0.6667	0.8333	0.8194	SVM
0.75	0.5	1.0	0.7778	SVM
0.6667	0.8333	0.5	0.7222	SVM
0.8333	0.6667	1.0	0.9722	SVM
1.0	1.0	1.0	1.0	SVM
0.5833	0.8333	0.3333	0.7778	SVM
0.6667	0.6667	0.6667	0.7778	SVM
0.75	0.8333	0.6667	0.7778	SVM
0.9167	1.0	0.8333	1.0	SVM
0.5833	0.6667	0.5	0.75	SVM
0.75	0.6667	0.8333	0.7778	SVM
0.9167	1.0	0.8333	0.9444	SVM
0.6667	0.8333	0.5	0.8056	SVM
0.6667	0.8333	0.5	0.8889	SVM
0.8333	0.6667	1.0	0.8889	SVM
0.9167	0.8333	1.0	0.8333	SVM
0.9167	0.8333	1.0	1.0	SVM
0.6667	0.8333	0.5	0.5833	SVM
0.75	0.5	1.0	0.9167	SVM
0.6667	0.3333	1.0	0.7222	SVM
0.75	0.6667	0.8333	0.9167	SVM
0.6667	0.5	0.8333	0.9167	SVM
0.6667	0.5	0.8333	0.6667	SVM
0.6667	0.5	0.8333	0.6389	SVM
0.8333	0.8333	0.8333	0.8333	SVM
0.9167	1.0	0.8333	0.8889	SVM

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	0.6667	0.8333	0.8056	SVM
0.6667	0.3333	1.0	0.7222	SVM
0.9167	0.8333	1.0	0.9444	SVM
0.75	1.0	0.5	0.9167	NN
0.8333	0.6667	1.0	0.8333	NN
0.8333	0.6667	1.0	0.8611	NN
0.5	0.3333	0.6667	0.5278	NN
0.75	0.6667	0.8333	0.8611	NN
0.8333	1.0	0.6667	0.9167	NN
0.75	0.8333	0.6667	0.9444	NN
0.5833	0.6667	0.5	0.7222	NN
0.6667	0.5	0.8333	0.6389	NN
0.8333	0.8333	0.8333	0.9167	NN
0.5833	0.6667	0.5	0.6944	NN
0.9167	0.8333	1.0	0.9167	NN
0.8333	1.0	0.6667	0.8611	NN
0.9167	0.8333	1.0	0.8889	NN
0.75	0.8333	0.6667	0.8056	NN
0.8333	0.6667	1.0	0.8889	NN
0.75	0.8333	0.6667	0.8333	NN
1.0	1.0	1.0	1.0	NN
0.5833	0.6667	0.5	0.5	NN
0.75	0.8333	0.6667	0.7778	NN
0.5833	0.5	0.6667	0.6667	NN
0.9167	0.8333	1.0	0.9444	NN
0.8333	0.6667	1.0	0.7778	NN
0.75	0.6667	0.8333	0.6944	NN
0.75	0.6667	0.8333	0.7778	NN
0.5833	0.6667	0.5	0.6944	NN
0.75	0.8333	0.6667	0.8056	NN
0.9167	0.8333	1.0	0.8611	NN
0.9167	0.8333	1.0	0.8611	NN

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	0.5	1.0	0.8611	NN
0.75	1.0	0.5	0.9722	Frank Copula
0.6667	1.0	0.3333	0.6389	Frank Copula
0.6667	0.6667	0.6667	0.6111	Frank Copula
0.8333	0.8333	0.8333	0.875	Frank Copula
0.6667	0.8333	0.5	0.7639	Frank Copula
0.8333	0.8333	0.8333	0.8889	Frank Copula
0.5	0.6667	0.3333	0.5833	Frank Copula
0.5833	0.8333	0.3333	0.6944	Frank Copula
0.4167	0.6667	0.1667	0.3056	Frank Copula
0.75	0.8333	0.6667	0.7778	Frank Copula
0.5833	0.8333	0.3333	0.6667	Frank Copula
0.75	1.0	0.5	0.9444	Frank Copula
0.5833	0.8333	0.3333	0.8333	Frank Copula
0.8333	1.0	0.6667	0.8889	Frank Copula
0.8333	1.0	0.6667	0.9722	Frank Copula
0.6667	0.8333	0.5	0.8333	Frank Copula
0.5	0.5	0.5	0.6111	Frank Copula
0.8333	1.0	0.6667	0.9444	Frank Copula
0.5	0.8333	0.1667	0.5556	Frank Copula
0.5833	0.5	0.6667	0.6528	Frank Copula
0.6667	0.8333	0.5	0.6528	Frank Copula
0.8333	1.0	0.6667	0.8333	Frank Copula
0.6667	0.6667	0.6667	0.6806	Frank Copula
0.75	0.6667	0.8333	0.7083	Frank Copula
0.75	0.8333	0.6667	0.8056	Frank Copula
0.8333	1.0	0.6667	0.8333	Frank Copula
0.5	0.6667	0.3333	0.5278	Frank Copula
0.5833	0.8333	0.3333	0.7083	Frank Copula
0.6667	0.5	0.8333	0.5278	Frank Copula
0.5833	0.5	0.6667	0.7222	Frank Copula
0.5833	0.8333	0.3333	0.5556	Clayton Copula

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5	0.8333	0.1667	0.4861	Clayton Copula
0.5	1.0	0.0	0.6111	Clayton Copula
0.75	1.0	0.5	0.75	Clayton Copula
0.5833	0.6667	0.5	0.6389	Clayton Copula
0.6667	0.8333	0.5	0.6944	Clayton Copula
0.5833	0.8333	0.3333	0.7222	Clayton Copula
0.4167	0.6667	0.1667	0.5	Clayton Copula
0.5	0.8333	0.1667	0.5139	Clayton Copula
0.6667	0.8333	0.5	0.7222	Clayton Copula
0.5833	0.8333	0.3333	0.6111	Clayton Copula
0.6667	1.0	0.3333	0.6111	Clayton Copula
0.5833	1.0	0.1667	0.6667	Clayton Copula
0.6667	0.8333	0.5	0.6528	Clayton Copula
0.5833	0.8333	0.3333	0.6111	Clayton Copula
0.5833	1.0	0.1667	0.6667	Clayton Copula
0.6667	1.0	0.3333	0.6667	Clayton Copula
0.4167	0.6667	0.1667	0.4306	Clayton Copula
0.6667	1.0	0.3333	0.75	Clayton Copula
0.75	1.0	0.5	0.75	Clayton Copula
0.4167	0.6667	0.1667	0.4444	Clayton Copula
0.6667	1.0	0.3333	0.6111	Clayton Copula
0.5	0.5	0.5	0.4861	Clayton Copula
0.5833	0.6667	0.5	0.5833	Clayton Copula
0.4167	0.5	0.3333	0.375	Clayton Copula
0.75	1.0	0.5	0.7083	Clayton Copula
0.5833	0.6667	0.5	0.5556	Clayton Copula
0.5833	0.6667	0.5	0.5694	Clayton Copula
0.5	0.6667	0.3333	0.5278	Clayton Copula
0.75	0.8333	0.6667	0.7917	Clayton Copula
0.75	0.6667	0.8333	0.8333	Gaussian Copula
0.5833	0.8333	0.3333	0.5556	Gaussian Copula
0.5	0.5	0.5	0.5833	Gaussian Copula

Continúa en la siguiente página

Tabla 6.3 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.8333	0.8333	0.8333	0.875	Gaussian Copula
0.8333	1.0	0.6667	0.7917	Gaussian Copula
0.8333	0.8333	0.8333	0.9167	Gaussian Copula
0.5	0.6667	0.3333	0.5	Gaussian Copula
0.5833	0.6667	0.5	0.6111	Gaussian Copula
0.4167	0.6667	0.1667	0.3889	Gaussian Copula
0.75	0.8333	0.6667	0.7639	Gaussian Copula
0.5	0.6667	0.3333	0.5694	Gaussian Copula
0.6667	1.0	0.3333	0.8472	Gaussian Copula
0.5833	0.8333	0.3333	0.7778	Gaussian Copula
0.8333	1.0	0.6667	0.8889	Gaussian Copula
0.75	1.0	0.5	1.0	Gaussian Copula
0.75	0.6667	0.8333	0.6389	Gaussian Copula
0.6667	0.6667	0.6667	0.6806	Gaussian Copula
0.8333	1.0	0.6667	0.9167	Gaussian Copula
0.5833	0.8333	0.3333	0.6667	Gaussian Copula
0.75	0.6667	0.8333	0.7222	Gaussian Copula
0.6667	0.6667	0.6667	0.6111	Gaussian Copula
0.8333	1.0	0.6667	0.8056	Gaussian Copula
0.5833	0.6667	0.5	0.625	Gaussian Copula
0.8333	0.6667	1.0	0.7361	Gaussian Copula
0.75	0.8333	0.6667	0.7222	Gaussian Copula
0.8333	1.0	0.6667	0.8889	Gaussian Copula
0.5	0.6667	0.3333	0.5139	Gaussian Copula
0.5	0.6667	0.3333	0.625	Gaussian Copula
0.5833	0.5	0.6667	0.5556	Gaussian Copula
0.8333	0.8333	0.8333	0.8611	Gaussian Copula

Tabla 6.4: Resultados en extenso de la clasificación con selección por Corr-EDA

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.6667	0.8333	0.5	0.7778	Naive Bayes
0.6667	0.6667	0.6667	0.8333	Naive Bayes
0.75	0.6667	0.8333	0.8333	Naive Bayes
0.75	0.8333	0.6667	0.8333	Naive Bayes
0.75	0.6667	0.8333	0.8056	Naive Bayes
0.75	0.8333	0.6667	0.8333	Naive Bayes
0.75	0.8333	0.6667	0.8056	Naive Bayes
0.75	0.6667	0.8333	0.7778	Naive Bayes
0.8333	0.8333	0.8333	0.9444	Naive Bayes
0.6667	0.6667	0.6667	0.75	Naive Bayes
0.75	0.5	1.0	0.6389	Naive Bayes
0.5833	0.5	0.6667	0.6667	Naive Bayes
0.5	0.3333	0.6667	0.4722	Naive Bayes
0.4167	0.5	0.3333	0.6389	Naive Bayes
0.5	0.6667	0.3333	0.6111	Naive Bayes
0.5833	0.6667	0.5	0.75	Naive Bayes
0.6667	0.5	0.8333	0.8333	Naive Bayes
0.8333	0.6667	1.0	0.8611	Naive Bayes
0.75	0.8333	0.6667	0.9167	Naive Bayes
0.4167	0.5	0.3333	0.5556	Naive Bayes
0.6667	0.8333	0.5	0.6667	Naive Bayes
0.6667	0.5	0.8333	0.75	Naive Bayes
0.6667	0.6667	0.6667	0.7222	Naive Bayes
0.6667	0.6667	0.6667	0.8333	Naive Bayes
0.75	0.6667	0.8333	0.75	Naive Bayes
0.6667	0.8333	0.5	0.8889	Naive Bayes
0.5833	0.5	0.6667	0.7778	Naive Bayes
0.5	0.8333	0.1667	0.7778	Naive Bayes
0.75	0.8333	0.6667	0.8333	Naive Bayes
0.5	0.5	0.5	0.5833	Naive Bayes
0.5833	0.5	0.6667	0.6389	KNN

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.75	0.5	1.0	0.7222	KNN
0.6667	0.5	0.8333	0.7083	KNN
0.8333	0.8333	0.8333	0.8194	KNN
0.5833	0.3333	0.8333	0.6389	KNN
0.6667	0.5	0.8333	0.6667	KNN
0.75	0.5	1.0	0.75	KNN
0.4167	0.3333	0.5	0.4306	KNN
0.8333	0.6667	1.0	0.8889	KNN
0.5833	0.3333	0.8333	0.6944	KNN
0.5	0.3333	0.6667	0.6528	KNN
0.4167	0.5	0.3333	0.4444	KNN
0.5	0.1667	0.8333	0.4167	KNN
0.5833	0.1667	1.0	0.6806	KNN
0.4167	0.1667	0.6667	0.5	KNN
0.6667	0.3333	1.0	0.75	KNN
0.5833	0.1667	1.0	0.625	KNN
0.5833	0.3333	0.8333	0.5972	KNN
0.5833	0.1667	1.0	0.5833	KNN
0.6667	0.3333	1.0	0.7222	KNN
0.5	0.3333	0.6667	0.5972	KNN
0.5	0.3333	0.6667	0.5139	KNN
0.5833	0.1667	1.0	0.5278	KNN
0.6667	0.3333	1.0	0.75	KNN
0.6667	0.5	0.8333	0.6944	KNN
0.5	0.1667	0.8333	0.5694	KNN
0.5833	0.3333	0.8333	0.7083	KNN
0.5833	0.3333	0.8333	0.6667	KNN
0.5	0.1667	0.8333	0.5278	KNN
0.6667	0.3333	1.0	0.7222	KNN
0.8333	0.6667	1.0	0.8333	SVM
0.6667	0.6667	0.6667	0.8056	SVM
0.8333	0.8333	0.8333	0.8611	SVM

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.9167	0.8333	1.0	0.9444	SVM
0.8333	0.8333	0.8333	0.8889	SVM
0.9167	1.0	0.8333	1.0	SVM
0.9167	0.8333	1.0	0.9722	SVM
0.9167	0.8333	1.0	0.8333	SVM
0.8333	0.8333	0.8333	0.9167	SVM
0.8333	0.8333	0.8333	0.9722	SVM
0.9167	0.8333	1.0	0.9722	SVM
0.8333	1.0	0.6667	1.0	SVM
0.8333	0.8333	0.8333	0.9444	SVM
0.8333	1.0	0.6667	0.9167	SVM
0.8333	0.8333	0.8333	0.9167	SVM
0.9167	0.8333	1.0	0.9722	SVM
0.6667	0.6667	0.6667	0.7778	SVM
0.9167	0.8333	1.0	0.9444	SVM
0.9167	0.8333	1.0	0.9444	SVM
0.8333	0.8333	0.8333	0.9167	SVM
0.8333	0.8333	0.8333	0.9583	SVM
0.9167	1.0	0.8333	0.9444	SVM
0.9167	0.8333	1.0	0.8611	SVM
0.75	0.6667	0.8333	0.75	SVM
0.9167	1.0	0.8333	1.0	SVM
1.0	1.0	1.0	1.0	SVM
0.9167	0.8333	1.0	0.9167	SVM
0.8333	0.8333	0.8333	0.9722	SVM
0.8333	0.8333	0.8333	0.9722	SVM
0.8333	0.8333	0.8333	0.9167	SVM
0.75	0.8333	0.6667	0.8611	NN
0.75	0.8333	0.6667	0.8056	NN
0.8333	0.8333	0.8333	0.8611	NN
0.9167	1.0	0.8333	1.0	NN
0.6667	0.5	0.8333	0.8333	NN

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
1.0	1.0	1.0	1.0	NN
0.9167	0.8333	1.0	0.9167	NN
0.9167	0.8333	1.0	0.8333	NN
0.9167	0.8333	1.0	1.0	NN
0.9167	0.8333	1.0	0.9722	NN
0.9167	0.8333	1.0	0.8611	NN
0.8333	0.8333	0.8333	0.9722	NN
0.8333	0.8333	0.8333	0.8333	NN
0.8333	1.0	0.6667	0.9167	NN
0.5833	0.6667	0.5	0.75	NN
1.0	1.0	1.0	1.0	NN
0.8333	0.8333	0.8333	0.8889	NN
0.9167	0.8333	1.0	0.8889	NN
0.8333	0.6667	1.0	0.9722	NN
0.75	0.6667	0.8333	0.9167	NN
0.75	0.6667	0.8333	0.8889	NN
0.8333	0.8333	0.8333	0.9167	NN
0.75	0.5	1.0	0.8333	NN
0.8333	0.6667	1.0	0.8889	NN
0.9167	0.8333	1.0	0.9722	NN
0.75	0.6667	0.8333	0.9444	NN
1.0	1.0	1.0	1.0	NN
0.75	0.8333	0.6667	0.8889	NN
0.9167	1.0	0.8333	0.9167	NN
0.8333	0.8333	0.8333	0.8611	NN
0.9167	1.0	0.8333	0.9444	Frank Copula
0.75	0.8333	0.6667	0.6944	Frank Copula
0.8333	0.8333	0.8333	0.8333	Frank Copula
0.6667	0.8333	0.5	0.6667	Frank Copula
0.75	0.5	1.0	0.75	Frank Copula
0.8333	0.8333	0.8333	0.8333	Frank Copula
0.6667	0.8333	0.5	0.8194	Frank Copula

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.8333	0.8333	0.8333	0.8056	Frank Copula
0.9167	1.0	0.8333	1.0	Frank Copula
0.75	0.8333	0.6667	0.7083	Frank Copula
0.75	0.5	1.0	0.7083	Frank Copula
0.5	0.5	0.5	0.4306	Frank Copula
0.5833	0.3333	0.8333	0.6528	Frank Copula
0.5833	0.5	0.6667	0.8333	Frank Copula
0.5833	0.8333	0.3333	0.7222	Frank Copula
0.75	0.6667	0.8333	0.75	Frank Copula
0.8333	0.8333	0.8333	0.9444	Frank Copula
0.6667	0.6667	0.6667	0.7083	Frank Copula
0.6667	0.5	0.8333	0.6667	Frank Copula
0.6667	0.8333	0.5	0.8056	Frank Copula
0.75	0.6667	0.8333	0.9167	Frank Copula
0.5833	0.8333	0.3333	0.5139	Frank Copula
0.5	0.6667	0.3333	0.6667	Frank Copula
0.6667	0.8333	0.5	0.7639	Frank Copula
0.6667	0.8333	0.5	0.8056	Frank Copula
0.8333	0.8333	0.8333	0.8889	Frank Copula
0.6667	0.8333	0.5	0.6389	Frank Copula
0.8333	0.8333	0.8333	0.9444	Frank Copula
0.6667	0.5	0.8333	0.9167	Frank Copula
0.6667	0.6667	0.6667	0.7222	Frank Copula
0.6667	0.6667	0.6667	0.7639	Clayton Copula
0.5833	0.6667	0.5	0.5833	Clayton Copula
0.75	0.6667	0.8333	0.7222	Clayton Copula
0.5	0.6667	0.3333	0.4444	Clayton Copula
0.6667	0.5	0.8333	0.7778	Clayton Copula
0.8333	0.8333	0.8333	0.7917	Clayton Copula
0.5833	0.6667	0.5	0.5833	Clayton Copula
0.6667	0.6667	0.6667	0.7222	Clayton Copula
0.75	0.6667	0.8333	0.6667	Clayton Copula

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.4167	0.1667	0.6667	0.4583	Clayton Copula
0.75	0.6667	0.8333	0.75	Clayton Copula
0.5833	0.6667	0.5	0.5417	Clayton Copula
0.4167	0.1667	0.6667	0.4861	Clayton Copula
0.5833	0.6667	0.5	0.4722	Clayton Copula
0.5	0.5	0.5	0.5278	Clayton Copula
0.75	0.5	1.0	1.0	Clayton Copula
0.6667	0.8333	0.5	0.6389	Clayton Copula
0.6667	0.3333	1.0	0.8333	Clayton Copula
0.75	0.6667	0.8333	0.7917	Clayton Copula
0.6667	0.6667	0.6667	0.5556	Clayton Copula
0.5833	0.5	0.6667	0.625	Clayton Copula
0.5833	0.5	0.6667	0.4722	Clayton Copula
0.5	0.5	0.5	0.5972	Clayton Copula
0.5833	0.6667	0.5	0.6111	Clayton Copula
0.6667	0.6667	0.6667	0.6667	Clayton Copula
0.4167	0.5	0.3333	0.4167	Clayton Copula
0.5833	0.5	0.6667	0.6944	Clayton Copula
0.75	0.8333	0.6667	0.7222	Clayton Copula
0.6667	0.5	0.8333	0.7778	Clayton Copula
0.6667	0.8333	0.5	0.6667	Clayton Copula
0.8333	1.0	0.6667	0.8889	Gaussian Copula
0.6667	0.8333	0.5	0.6944	Gaussian Copula
0.8333	0.8333	0.8333	0.8333	Gaussian Copula
0.5833	0.8333	0.3333	0.6111	Gaussian Copula
0.6667	0.5	0.8333	0.75	Gaussian Copula
0.75	0.8333	0.6667	0.8333	Gaussian Copula
0.6667	0.8333	0.5	0.7917	Gaussian Copula
0.75	0.6667	0.8333	0.8056	Gaussian Copula
1.0	1.0	1.0	1.0	Gaussian Copula
0.75	0.8333	0.6667	0.7361	Gaussian Copula
0.75	0.5	1.0	0.7639	Gaussian Copula

Continúa en la siguiente página

Tabla 6.4 – continúa de la página anterior

Exactitud	Sensibilidad	Especificidad	AUC	Clasificador
0.5833	0.6667	0.5	0.5556	Gaussian Copula
0.5833	0.1667	1.0	0.5139	Gaussian Copula
0.75	0.8333	0.6667	0.8056	Gaussian Copula
0.5	0.6667	0.3333	0.625	Gaussian Copula
0.8333	0.8333	0.8333	0.7222	Gaussian Copula
0.8333	0.8333	0.8333	0.9444	Gaussian Copula
0.8333	0.8333	0.8333	0.8472	Gaussian Copula
0.8333	0.8333	0.8333	0.7778	Gaussian Copula
0.5833	0.6667	0.5	0.6111	Gaussian Copula
0.6667	0.8333	0.5	0.8056	Gaussian Copula
0.4167	0.6667	0.1667	0.3333	Gaussian Copula
0.5	0.6667	0.3333	0.6806	Gaussian Copula
0.5833	0.8333	0.3333	0.5972	Gaussian Copula
0.6667	0.8333	0.5	0.7361	Gaussian Copula
0.9167	1.0	0.8333	0.875	Gaussian Copula
0.75	0.8333	0.6667	0.6667	Gaussian Copula
0.8333	0.8333	0.8333	0.9722	Gaussian Copula
0.75	0.8333	0.6667	0.8333	Gaussian Copula
0.8333	1.0	0.6667	0.8056	Gaussian Copula