

Análisis comparativo de técnicas de aprendizaje automático para la detección de la enfermedad de Parkinson mediante el reconocimiento de voz

Justo Francisco León Pastor Baak, Giner Alor Hernández,
Nancy Aracely Cruz-Ramos, José Luis Sánchez Cervantes,
Lisbeth Rodríguez Mazahua

Tecnológico Nacional de México, Instituto Tecnológico de Orizaba,
División de Estudios de Posgrado e Investigación,
México

{M22011021, giner.ah, jose.sc.dci,
lisbeth.rm}@orizaba.tecnm.mx, ncruz@ito-depi.edu.mx

Resumen. La enfermedad de Parkinson es una de las enfermedades neurodegenerativas más prevalentes a nivel mundial siendo aproximadamente el 1% en mayores de 60 años; su progresión es gradual y sin cura conocida. En este trabajo se presenta un análisis comparativo de los algoritmos de inteligencia artificial usados para la detección temprana de la enfermedad de Parkinson mediante el reconocimiento de voz. Para este análisis se consideraron 5 conjuntos de datos y 11 algoritmos de aprendizaje automático. Como resultado de este análisis, se obtuvo que los algoritmos Bagged SVM, Random Forest y Gradient Boosting presentaron los mejores resultados en sensibilidad, siendo aplicados en el conjunto de datos Parkinson's Disease Classification.

Palabras claves: Bagged SVM, enfermedad de Parkinson, procesamiento de lenguaje natural, random forest, reconocimiento de voz.

Comparative Analysis of Machine Learning Techniques for Detecting Parkinson's Disease Using Speech Recognition

Abstract. Parkinson's disease is one of the most prevalent neurodegenerative diseases worldwide being approximately 1% in people over 60 years old; its progression is gradual and without known cure. This paper presents a comparative analysis of artificial intelligence algorithms used for early detection of Parkinson's disease using speech recognition. For this analysis, 5 datasets and 11 machine learning algorithms were considered. As a result, Bagged SVM, Random Forest and Gradient Boosting algorithms presented the best results in recall, being applied to the Parkinson's Disease Classification dataset.

Keywords: Bagged SVM, natural language processing, Parkinson's disease, random forest, speech recognition.

1. Introducción

La enfermedad de Parkinson es un proceso neurodegenerativo complejo de aparición en la edad adulta, constituye la segunda enfermedad neurodegenerativa más frecuente por detrás de la enfermedad de Alzheimer [1], los síntomas abarcan, desórdenes motores como; temblores, bradicinesia, rigidez, inestabilidad postural, deterioro cognitivo, bradifenia, hiponimia, disartria, y otros [2].

También se presentan desórdenes del discurso y lenguaje como: mudez, afasia, velocidad rápida, volumen reducido, monótono, mono sonoridad, disartria hipocinética, tartamudeo adquirido, y otros desordenes [3].

Algunas estadísticas como la de la Organización Mundial de la Salud, informan que la enfermedad de Parkinson, está en constante aumento, las estimaciones mundiales indican que en el 2019 más de 8,5 millones de personas fueron afectadas por la enfermedad de Parkinson, lo que provocó el fallecimiento de 329 000 de personas en el año 2020, esto destaca la urgente necesidad de estrategias más efectivas para su diagnóstico temprano [4].

Martínez *et al.* [5] señalan que, en México, la enfermedad de Parkinson aumentó la incidencia de 6.7 en 2014 a 14.9 por 100 000 habitantes en 2023.

Hoy en día esta enfermedad se puede diagnosticar implementando algoritmos de inteligencia artificial, mediante el análisis del movimiento de un paciente [6], análisis de magneto encefalografías [7], análisis de electroencefalogramas [8] e irregularidades en la señal de voz analizando fonemas del paciente [9-19].

Por lo que la importancia de las técnicas de inteligencia artificial en la detección de enfermedades es ampliamente aceptada en la comunidad científica y médica, ya que también, estas herramientas posibilitan la automatización y optimización de los procesos diagnósticos y terapéuticos lo que conduce a una presentación de servicios de salud más eficiente y efectiva, dentro de estas técnicas se encuentran los algoritmos de aprendizaje automático, que es la ciencia (y arte) de programar computadoras para que puedan aprender de datos [10].

Actualmente existen diversos estudios sobre la aplicación del aprendizaje automático en el diagnóstico de la enfermedad de enfermedades con altos porcentajes de éxito en la precisión de su clasificación.

En este trabajo, se presenta un análisis de los algoritmos de aprendizaje automático para el diagnóstico de la enfermedad de Parkinson a través del reconocimiento de voz.

La estructura de este artículo es la siguiente: en la sección 2 se presenta el estado del arte de los trabajos relacionados referente a los algoritmos y técnicas de Procesamiento de Lenguaje Natural para la extracción de elementos vocales para la detección de la temprana de la enfermedad de Parkinson. La sección 3 especifica la descripción de los (*Data Set*, Conjunto de datos) utilizados y sus características para este análisis. En la sección 4 se profundiza en los resultados del análisis y evaluación realizada a los conjuntos de datos con técnicas de inteligencia artificial previamente seleccionadas. En la sección 5 se encuentra la discusión de los resultados obtenidos y finalmente en la sección 6 se presentan las conclusiones y el trabajo futuro.

2. Estado del arte

En esta sección se presentan los estudios previos que utilizaron técnicas de aprendizaje automático en el reconocimiento de voz para la detección de la enfermedad de Parkinson, destacan desarrollos como NeuroSpeech de Orozco-Arroyave *et al.* [11], y modelos LSTM Bidireccionales propuestos por Quan *et al.* [12], también se identificaron herramientas prometedoras con hasta un 92% de precisión.

Orozco-Arroyave *et al.* [11] desarrollaron "NeuroSpeech", un software para evaluar deficiencias del habla en pacientes con Parkinson, con el uso del algoritmo SVM (*Support Vector Machine*, Máquinas de Vector Soporte), se identificaron alteraciones en fonación, articulación, prosodia e inteligibilidad y se demostró una reducción en la inteligibilidad, con un 83% de exactitud en la pronunciación y 50% de similitud, sugiriendo ser una herramienta prometedora en la detección de la enfermedad de Parkinson en el reconocimiento de voz.

Quan *et al.* [12] propusieron un modelo LSTM *Bidirectional* (*Long Short-Term Memory*, Red de Gran Memoria de Corto Plazo Bidireccional) para detectar Parkinson mediante análisis de señales de voz, que reveló cambios específicos como aumento de (*Jitter*, Fluctuación) y (*Shimmer*, Brillo), con el uso de técnicas como MFCC (*Mel Frequency Cepstral Coefficients*, Coeficientes Cepstrales en las Frecuencias de Mel) y PCA (*Principal Component Analysis*, Análisis de Componentes Principales) para clasificación. Klempf *et al.* [13] se enfocaron en la clasificación del diagnóstico de la enfermedad de Parkinson basados en sus señales del discurso, usando los métodos de reconocimiento de patrones AdaBoost, y (*Bagged Trees*, Árboles Embolsados), Quadratic SVM y KNN (*Nearest Neighbors*, k-Vecinos más Cercanos) obteniendo un promedio de 82.3% de exactitud y un promedio de 0.88 en el AUC (*Area Under Curve*, Área Bajo la Curva).

Rahman *et al.* [14] propusieron un marco de trabajo para diagnosticar la enfermedad de Parkinson utilizando datos de 726 pacientes de EE.UU. y otros lugares, utilizando los algoritmos SVM, XGBoost (*eXtreme Gradient Boosting*, Potenciación de Gradientes Extremo), LightGBM (*Light - Gradient Boosting Machine*, Luz - Máquina de Potenciación de Gradiente) y RF (*Random Forest*, Bosque Aleatorio) obteniendo 0.753 de AUC con XGBoost. Almeida *et al.* [15] adoptaron un enfoque, centrado en características como el lenguaje nativo y la recolección de datos mediante pruebas grabadas en lituano. Se obtuvo una clasificación de la enfermedad de Parkinson por voz con 92% de exactitud. Younis Thanoun *et al.* [16] investigaron un enfoque para identificar enfermos de Parkinson usando el clasificador SMOTE en una distribución de clases desequilibradas, obteniendo un 96%.52 de exactitud. Zhang *et al.* [17] desarrollaron una aplicación móvil con enfoque en voz, donde SVM y SVR (*Support Vector Regressor*, Regresión de Vectores de Soporte) mostraron el mejor rendimiento para diagnosticar y evaluar la severidad del Parkinson, lograron un 97.03% de sensibilidad en la clasificación. Karan *et al.* [18] propusieron un nuevo enfoque para detectar la enfermedad de Parkinson en la señal de voz basado en descomposición en modo intrínseco tomando en cuenta los bancos de filtros diádicos que capturan eficazmente la dinámica en la señal de la voz, obtuvieron un 100 % de exactitud utilizando el algoritmo de clasificación RF. Quan *et al.* [19] propusieron un modelo de aprendizaje profundo para detectar la enfermedad de Parkinson a partir de señales de voz. Se utilizaron *2Dimensional CNN* (*Convolutional Neural Networks*, Redes

Neuronales Convolucionales) para extraer características dinámicas de series de tiempo y 1Dimensional CNN para capturar dependencias, logrando el 92% de precisión. Hoq *et al.* [20] propusieron 2 modelos híbridos basados en SVM con un PCA y 1 con codificador automático escaso para detectar pacientes con la enfermedad de Parkinson basándose en sus atributos vocales, donde se obtuvo el 93% en exactitud. Khan *et al.* [21] desarrollaron un modelo que clasifica la enfermedad de Parkinson a través de la escala UPDRS (*Unified Parkinson's Disease Rating Scale*, Escala Unificada para la Evaluación de la Enfermedad de Parkinson) aplicando el algoritmo de SVM utilizando la escala Guttman para realizar sistemáticamente la correlación entre variables cualitativas y su orden en el índice UPDRS, obteniendo un 92% de exactitud.

Wroge *et al.* [22] aplicaron diferentes modelos de aprendizaje automático para clasificar la enfermedad de Parkinson usando el conjunto de datos llamado mPower y la biblioteca de Python, PyAudio, de esta manera obtuvieron como resultado 89% de AUC con RF. Soumaya *et al.* [23] decodificaron la señal del habla de la vocal /a/ para clasificar pacientes con la enfermedad de Parkinson, el método propuesto fue un sistema de clasificación que consistió en las 5 fases: conversión de señal, preprocesamiento, extracción de atributos, selección de atributos, y clasificación, en donde el mejor resultado fue obtenido por SVM con un 91.8% de exactitud.

Hyder *et al.* [7] investigaron la conectividad funcional en las redes corticales durante el procesamiento de lenguaje hablado, utilizando MNE-Python 0.19 para analizar magnetoencefalogramas, y LR (*Logistic Regression*, Regresión logística) para la clasificación, logrando un 78% de precisión en la detección de Parkinson mediante el análisis de verbos de acción. Mostafa *et al.* [24] evaluaron el desempeño de 3 métodos para diagnosticar la enfermedad de Parkinson: DT (*Decision Tree*, *Árbol de Decisión*), NV (*Naive Bayes*, Bayes Ingenuo) y NN (*Neuronal Network*, Red Neuronal), el estudio encontró que el algoritmo DT tuvo el mejor desempeño, con un 91,63% de precisión, los resultados sugieren que DT es el método más preciso para diagnosticar la enfermedad de Parkinson. Nissar *et al.* [9] evaluaron el rendimiento del modelo de aprendizaje automático XGBoost para el diagnóstico de la enfermedad de Parkinson mediante la voz, el estudio utilizó un conjunto de datos de voz de 756 pacientes y 64 sujetos sanos, los resultados mostraron que XGBoost obtuvo la exactitud más alta de 88.15%, seguido de RF y DT con 84.86% de exactitud, el modelo de clasificación NB obtuvo la exactitud más baja del 74.34%, en general.

A continuación, se describen los conjuntos de datos de reconocimiento de voz para la detección de la enfermedad de Parkinson reportados en la literatura para su análisis.

3. Conjuntos de datos para la identificación del Parkinson

En esta sección se describen los distintos conjuntos de datos utilizados para el análisis en la identificación de la enfermedad de Parkinson mediante el reconocimiento de voz. Se utilizaron 5 conjuntos de datos, cada uno con distintos enfoques y características:

- **Parkinson's Disease Data Set** [25]: Es un conjunto de datos, que se diseñó en [26] para determinar si un paciente se encuentra sano o tiene enfermedad de Parkinson tomando en cuenta los atributos relacionados con la detección de la disfonía, estos

atributos se conforman, 195 registros de fonaciones vocales de 31 hombres y mujeres, de los cuales 23 fueron diagnosticados con la enfermedad de Parkinson.

- **Parkinson Speech Dataset with Multiple Types of Sound Recordings** [27]: Se diseñó en [28] para clasificar el registro de voz de un paciente como sano o con la enfermedad de Parkinson utilizando 2 algoritmos de clasificación: SVM y KNN, siendo el algoritmo de SVM el que obtuvo un mejor rendimiento general en la clasificación, contiene atributos relacionados con la disfonía, los datos de entrenamiento pertenecen a 20 pacientes con la enfermedad de Parkinson y 20 sujetos sanos, de todos los sujetos, se tomaron 26 tipos de grabaciones de sonidos, de las cuales se extrajo un grupo de 26 características lineales y basadas en frecuencia de tiempo, puntuación UPDRS, en total el conjunto de datos contiene 1040 instancias de enteros reales.
- **Parkinson's Disease Classification** [29]: Se diseñó en [30] con el objetivo de usar la TQWT (*Tunable Q-factor Wavelet Transform*, Transformada Ondícula de Factor Q Sintonizable) a las señales de voz de pacientes con la enfermedad de Parkinson para la extracción de características, que tiene una resolución de frecuencia más alta que la clásica transformada ondícula discreta. La extracción de características del conjunto de datos se basó en los métodos usados en el diagnóstico de los desórdenes vocales en la enfermedad de Parkinson. Los datos se recopilaron de 188 pacientes con enfermedad de Parkinson (107 hombres y 81 mujeres) con edades comprendidas entre 33 y 87 años ($65,1 \pm 10,9$) en el Departamento de Neurología de la Facultad de Medicina de CerrahpaSa, Universidad de Estambul, el grupo control estuvo formado por 64 individuos sanos (23 hombres y 41 mujeres) con edades que oscilan entre los 41 y los 82 años ($61,1 \pm 8,9$), durante el proceso de recolección de datos, el micrófono se configuró a 44,1 KHz y luego del examen del médico, se recolectó la fonación sostenida de la vocal /a/ de cada sujeto con 3 repeticiones, el conjunto de datos contiene 758 instancias de enteros reales, organizados por base (26), Intensidad (3), Ancho de banda y Formantes (8), MFCC (84), WT (*Wavelet transform, Transformada Wavelet*) aplicada a F0 (182), Pliegue vocal (22), TQWT (432) con un total de 756 tributos.
- **Parkinson Replicated Acoustic Features Dataset** [31]: Se utilizó en [32] con el objetivo de ser usado por un sistema capaz de discriminar automáticamente personas sanas de personas con la enfermedad de Parkinson partiendo de múltiples grabaciones de la voz de los mismos sujetos. Los atributos se utilizaron en un enfoque de BLR (*Bayesian Linear Regression*, Regresión Lineal Bayesiana) para manejar mediciones replicadas y tiempo. Estos atributos se identificaron como los mejores para alimentar el modelo de predicción por el método LASSO (*Least Absolute Shrinkage and Selection Operator*, Operador de Selección y Contracción Mínima Absoluta), que minimiza el error cuadrado sujeto a la restricción no diferenciable expresada en términos de la norma L1 de los coeficientes. El conjunto de datos contiene características extraídas de 3 grabaciones de voz replicadas, por lo que no se puede utilizar cada fila de forma independiente, ya que es 1 de las 3 réplicas de 1 individuo, existen 240 instancias, pero solo para 80 sujetos, 40 de ellos con enfermedad de Parkinson, por lo que no son independientes.

Tabla 1. Comparación de los conjuntos de datos para el diagnóstico de la enfermedad de Parkinson.

| Conjunto de datos | # de atributos | # de registros | Descripción de los datos |
|--|----------------|----------------|--|
| Parkinson's Data Set | 24 | 195 | Datos de amplitud, frecuencia y ruido en los componentes tonales de la voz |
| Parkinson Multiple Sound Recording | 29 | 1040 | Datos de irregularidad en la frecuencia, amplitud, y número de frenos en la voz |
| Parkinson's Disease Classification | 756 | 758 | Datos de irregularidad en la frecuencia, amplitud, y de la transformada ondícula de factor Q sintonizable. |
| Parkinson Replicated Acoustic Features Dataset | 240 | 48 | Datos de irregularidad en la frecuencia, amplitud, y de la relación de excitación glótica / ruido |
| Parkinson's telemonitoring | 22 | 5875 | Datos de irregularidad en la frecuencia, amplitud, e índice UPDRS |

- **Parkinson's telemonitoring** [33]: Contiene atributos relacionados con el diagnóstico de la disfonía en sujetos con la enfermedad de Parkinson. Este diagnóstico se realizó a través del uso de fonaciones de vocales sostenidas durante el desarrollo del artículo [34], donde se logró determinar si un registro contiene rasgos de síntomas vocales pertenecientes a la enfermedad de Parkinson, a través de analizar la señal, extraer rasgos representando las características de la señal, y mapear estos rasgos a la escala UPDRS usando métodos de regresión, el estudio se inició asumiendo que el deterioro en el rendimiento del habla es debido a la enfermedad de Parkinson y no a otra patología, los datos se obtuvieron de 5923 señales, usando un micrófono montado en la cabeza a 5 cm de los labios del paciente, el conjunto de datos se compone de una serie de mediciones de la voz de 42 personas con enfermedad de Parkinson en fase temprana, reclutadas para un ensayo de 6 meses, de un dispositivo de telemonitorización para la monitorización remota de la progresión de los síntomas, en donde las grabaciones se capturaron automáticamente en los hogares de los pacientes, las columnas de la tabla contienen el número de sujeto, la edad del sujeto, el sexo del sujeto, el intervalo de tiempo desde la fecha de reclutamiento inicia, UPDRS motor, UPDRS total y 16 medidas de voz biomédicas en donde cada fila corresponde a una de las 5,875 grabaciones de voz de estas personas siendo el objetivo principal de los datos predecir las puntuaciones motoras y totales de UPDRS a partir de las 16 medidas de voz, los datos están en formato ASCII CSV y las filas del archivo CSV contienen una instancia correspondiente a una grabación de voz.

En la tabla 1 se muestran características como: nombre del conjunto de datos, número de atributos, cantidad de registros y una breve descripción de los datos.

A continuación, se muestran los algoritmos de aprendizaje automático, de tipo (*Boosting*, Impulso) y (*Bagging*, Ensamble) utilizados en el análisis de los conjuntos de datos, así como su configuración y el rendimiento de las métricas seleccionadas.

Tabla 2. Comparación de las métricas de los algoritmos implementados en el conjunto de datos Parkinson's Data Set.

| Algoritmo | Exactitud | Precisión | Sensibilidad | Especificidad | F1-Score | AUC |
|-------------------|-----------|-----------|--------------|---------------|----------|--------|
| SVM | 87.18 | 90.63 | 93.55 | 62.50 | 92.06 | 0.8185 |
| KNN | 76.92 | 92.31 | 77.42 | 75.00 | 84.21 | 0.8972 |
| NB | 61.54 | 100.00 | 51.61 | 100.00 | 68.09 | 0.8790 |
| LR | 82.05 | 90.00 | 87.10 | 62.50 | 88.52 | 0.8347 |
| DT | 76.92 | 95.83 | 74.19 | 87.50 | 83.64 | 0.8085 |
| AdaBoost | 87.18 | 93.33 | 90.32 | 75.00 | 91.80 | 0.9637 |
| Gradient Boosting | 82.05 | 92.86 | 83.87 | 75.00 | 88.14 | 0.9194 |
| XGBoost | 87.18 | 96.43 | 87.10 | 87.50 | 91.53 | 0.9395 |
| RF | 79.49 | 89.66 | 83.87 | 62.50 | 86.67 | 0.8931 |
| Bagged DT | 76.92 | 89.29 | 80.65 | 62.50 | 84.75 | 0.8992 |
| Bagged SVM | 84.62 | 87.88 | 93.55 | 50.00 | 90.63 | 0.8306 |

Tabla 1. Comparación de las métricas de los algoritmos implementados en el conjunto de datos Parkinson Multiple Sound Recording.

| Algoritmo | Exactitud | Precisión | Sensibilidad | Especificidad | F1-Score | AUC |
|-------------------|-----------|-----------|--------------|---------------|----------|--------|
| SVM | 63.46 | 60.17 | 71.00 | 56.48 | 65.14 | 0.6925 |
| KNN | 59.13 | 56.64 | 64.00 | 54.63 | 60.09 | NA |
| NB | 53.85 | 51.47 | 70.00 | 38.89 | 59.32 | NA |
| LR | 58.17 | 55.75 | 63.00 | 53.70 | 59.15 | 0.6519 |
| DT | 61.06 | 59.79 | 58.00 | 63.89 | 58.88 | 0.6094 |
| AdaBoost | 62.02 | 60.82 | 59.00 | 64.81 | 59.90 | 0.6769 |
| Gradient Boosting | 63.94 | 62.89 | 61.00 | 66.67 | 61.93 | 0.6996 |
| XGBoost | 66.35 | 65.31 | 64.00 | 68.52 | 64.65 | 0.7339 |
| RF | 65.38 | 64.89 | 61.00 | 69.44 | 62.89 | 0.7133 |
| Bagged DT | 60.10 | 58.95 | 56.00 | 63.89 | 57.44 | 0.5994 |
| Bagged SVM | 63.46 | 60.17 | 69.00 | 59.26 | 64.79 | 0.6897 |

4. Análisis comparativo de técnicas de aprendizaje automático

En esta sección se detalla el análisis realizado a los conjuntos de datos revisados con distintas técnicas de inteligencia artificial, las cuales aplican técnicas de aprendizaje automático. Dentro de los algoritmos de aprendizaje automático se utilizaron SVM, KNN, NB, LR, DT, Dentro de los algoritmos de ensamble de tipo Boost: AdaBoost, (*Boostraping*, Autoensamble), (*Gradient Boosting*, Potenciación de Gradiente), XGBoost, y dentro de los algoritmos de ensamble de tipo *Bagging* se utilizaron RF, Bagged DT, Bagged SVM. Para llevar a cabo el análisis, se utilizó el lenguaje de programación Python en su versión 3.12 para realizar los scripts.

Como primer paso se importó la biblioteca Pandas, para manipular los datos de forma organizada y eficiente, después, se determinó la forma del df (Dataframe,

Tabla 4. Comparación de resultados obtenidos en el conjunto de datos Parkinson’s Disease Classification.

| Algoritmo | Exactitud | Precisión | Sensibilidad | Especificidad | F1-Score | AUC |
|-------------------|-----------|-----------|--------------|---------------|----------|--------|
| SVM | 80.92 | 88.99 | 85.09 | 68.42 | 87.00 | 0.8802 |
| KNN | 86.18 | 86.61 | 96.49 | 55.26 | 91.29 | 0.9535 |
| NB | 73.03 | 90.11 | 71.93 | 76.32 | 80.00 | 0.7946 |
| LR | 85.53 | 89.66 | 91.23 | 68.42 | 90.43 | 0.8996 |
| DT | 76.92 | 95.83 | 74.19 | 87.50 | 83.64 | 0.8085 |
| AdaBoost | 90.13 | 91.60 | 95.61 | 73.68 | 93.56 | 0.9148 |
| Gradient Boosting | 87.50 | 87.40 | 97.37 | 57.89 | 92.12 | 0.9578 |
| XGBoost | 90.13 | 90.24 | 97.37 | 68.42 | 93.67 | 0.9663 |
| RF | 90.13 | 89.60 | 98.25 | 65.79 | 93.72 | 0.9675 |
| Bagged DT | 88.82 | 89.43 | 96.49 | 65.79 | 92.83 | 0.9481 |
| Bagged SVM | 88.16 | 86.36 | 100.00 | 52.63 | 92.68 | 0.9224 |

Tabla 5. Comparación de resultados obtenidos en el conjunto de datos Parkinson Replicated Acoustic Features Dataset.

| Algoritmo | Exactitud | Precisión | Sensibilidad | Especificidad | F1-Score | AUC |
|-------------------|-----------|-----------|--------------|---------------|----------|--------|
| SVM | 68.75 | 70.83 | 68.00 | 69.57 | 69.39 | 0.8035 |
| KNN | 85.42 | 87.50 | 84.00 | 86.96 | 85.71 | 0.8904 |
| NB | 87.50 | 88.00 | 88.00 | 86.96 | 88.00 | 0.8765 |
| LR | 75.00 | 76.00 | 76.00 | 73.91 | 76.00 | 0.8383 |
| DT | 77.08 | 85.00 | 68.00 | 86.96 | 75.56 | 0.7748 |
| AdaBoost | 77.08 | 76.92 | 80.00 | 73.91 | 78.43 | 0.7878 |
| Gradient Boosting | 83.33 | 83.33 | 83.33 | 82.61 | 83.33 | 0.8591 |
| XGBoost | 85.42 | 87.50 | 84.00 | 86.96 | 85.71 | 0.8852 |
| RF | 87.50 | 88.00 | 88.00 | 86.96 | 88.00 | 0.8983 |
| Bagged DT | 81.25 | 81.27 | 81.25 | 78.26 | 81.23 | 0.8583 |
| Bagged SVM | 75.00 | 75.09 | 75.00 | 69.57 | 74.91 | 0.7965 |

Estructura de Datos Bidimensional), se obtuvo la información sobre el total de columnas y nombres de las columnas, se verifico si existían valores nulos y cuáles eran los tipos de datos dentro del df.

Como segundo paso, se filtraron las características importantes de uno de cada conjunto de datos, excluyendo aquellas características cuya presencia no aporta valor al resultado, después se asignaron los atributos relacionados con la predicción y el atributo target.

Continuando, como tercer paso, los conjuntos de datos se dividieron en un 80% de entrenamiento y un 20% de pruebas, para después estandarizar los valores con la

Tabla 2. Comparación de las métricas de los algoritmos implementados en el conjunto de datos Parkinson's telemonitoring.

| Algoritmo | Exactitud | Precisión | Sensibilidad | Especificidad | F1-Score | AUC |
|-------------------|-----------|-----------|--------------|---------------|----------|--------|
| SVM | 57.28 | 60.71 | 27.52 | 84.01 | 37.87 | 0.6488 |
| KNN | 70.72 | 70.46 | 65.65 | 75.28 | 67.97 | 0.7727 |
| NB | 54.38 | 56.94 | 14.75 | 89.98 | 23.43 | 0.5943 |
| LR | 57.96 | 58.24 | 39.39 | 74.64 | 47.00 | 0.6470 |
| DT | 65.28 | 64.12 | 60.43 | 69.63 | 62.22 | 0.6503 |
| AdaBoost | 53.19 | 58.33 | 3.78 | 97.58 | 07.09 | 0.5635 |
| Gradient Boosting | 67.15 | 68.89 | 55.76 | 77.38 | 61.63 | 0.7575 |
| XGBoost | 70.64 | 70.81 | 64.57 | 76.09 | 67.54 | 0.7943 |
| RF | 71.66 | 72.34 | 64.93 | 77.71 | 68.44 | 0.7933 |
| Bagged DT | 68.51 | 71.04 | 56.47 | 79.32 | 62.93 | 0.7604 |
| Bagged SVM | 57.87 | 60.00 | 32.91 | 80.29 | 42.51 | 0.6504 |

función `StandardScaler()` de la biblioteca `Scikit-Learn`, la cual se utiliza para preprocesar datos antes de ajustar un modelo de aprendizaje automático, estandarizando características, eliminando la media y escalando a la varianza unitaria. Hasta este paso la configuración es la misma para todos los algoritmos de inteligencia artificial, en el cuarto paso, se crearon los modelos de aprendizaje automático, de *Boosting* y *Bagging*. Para todos los algoritmos, se obtuvieron las siguientes métricas: (Exactitud, *Accuracy*), (*Precision* Precisión), (*Recall*, Sensibilidad), (*Specificity*, Especificidad), F1 score y el AUC.

Para evaluar el rendimiento del modelo de clasificación al momento de predecir a que clase pertenece la instancia de datos, se utilizó la matriz de confusión y otras métricas por ser ampliamente aceptadas en la literatura.

A continuación, las tablas 2-6 presentan una comparación de las métricas obtenidas por los algoritmos de inteligencia artificial implementados a los conjuntos de mayor uso en los trabajos relacionados: Parkinson's Data Set, Parkinson Multiple Sound Recording, Parkinson's Disease Classification, Parkinson Replicated Acoustic Features Data Set y Parkinson telemonitoring, en donde los resultados de las métricas, se representan como valores de porcentaje, con excepción del AUC ya que es una medida entre 0 y 1.

En la tabla 2 se presenta una evaluación comparativa de 11 algoritmos de Inteligencia Artificial aplicados al conjunto de datos Parkinson's Disease Data Set, en donde se analizan métricas clave como exactitud, precisión, sensibilidad, especificidad, F1-Score y AUC para determinar el algoritmo con mejor desempeño en la detección de la enfermedad de Parkinson.

Con base en los resultados de la tabla 2, se concluye que el algoritmo XGboost presenta el mejor desempeño general para la detección de la enfermedad de Parkinson. La tabla 3 presenta una evaluación comparativa al conjunto de datos Parkinson Speech Dataset with Multiple Types of Sound Recordings.

Tabla 7. Comparación de los resultados obtenidos por los algoritmos de aprendizaje automático en cada conjunto de datos.

| Dataset | Algoritmo destacado | Exactitud | Precisión | Sensibilidad | Especificidad | Índice F1 | AUC |
|--|---------------------|-----------|-----------|--------------|---------------|-----------|--------|
| Parkinson's Data Set | SVM | 87.18 | 90.63 | 93.55 | 62.50 | 92.06 | 0.8185 |
| | Bagged SVM | 84.62 | 87.88 | 93.55 | 50.00 | 90.63 | 0.8306 |
| | AdaBoost | 87.18 | 93.33 | 90.32 | 75.00 | 91.80 | 0.9637 |
| Parkinson Multiple Sound Recording | SVM | 63.46 | 60.17 | 71.00 | 56.48 | 65.14 | 0.6925 |
| | NB | 53.85 | 51.47 | 70.00 | 38.89 | 59.32 | NA |
| | Bagged SVM | 63.46 | 60.17 | 69.00 | 59.26 | 64.79 | 0.6897 |
| Parkinson's Disease Classification | Bagged SVM | 88.16 | 86.36 | 100.00 | 52.63 | 92.68 | 0.9224 |
| | RF | 90.13 | 89.60 | 98.25 | 65.79 | 93.72 | 0.9675 |
| | Gradient Boosting | 87.50 | 87.40 | 97.37 | 57.89 | 92.12 | 0.9578 |
| Parkinson Replicated Acoustic Features | RF | 87.50 | 88.00 | 88.00 | 86.96 | 88.00 | 0.8983 |
| | NB | 87.50 | 88.00 | 88.00 | 86.96 | 88.00 | 0.8765 |
| | XGBoost | 85.42 | 87.50 | 84.00 | 86.96 | 85.71 | 0.8852 |
| Parkinson's telemonitoring | KNN | 70.72 | 70.46 | 65.65 | 75.28 | 67.97 | 0.7727 |
| | RF | 71.66 | 72.34 | 64.93 | 77.71 | 68.44 | 0.7933 |
| | XGBoost | 70.64 | 70.81 | 64.57 | 76.09 | 67.54 | 0.7943 |

En la tabla 3, se observa que el algoritmo XGBoost presenta el mejor desempeño general en términos de exactitud (66.35%), precisión (65.31%), F1-Score (64.65%) y AUC (73.39%), estos resultados sugieren que XGBoost es el algoritmo más efectivo para la detección de la enfermedad de Parkinson a partir de grabaciones de voz. En la tabla 4 continúa la evaluación de los algoritmos y el análisis de los resultados del conjunto de datos Parkinson's Disease Classification.

En la tabla 4 se observa un rendimiento notable de varios algoritmos, XGBoost, RF y Bagged SVM destacan por su alta precisión (90.24%, 89.60% y 86.36% respectivamente), F1-Score (93.67%, 93.72% y 92.24%) y AUC (0.9663, 0.9675 y 0.9224). En la tabla 5 se presenta la evaluación del conjunto de datos Parkinson Replicated Acoustic Features Dataset.

En la tabla 5 RF, NB y KNN destacan con un 88% y 88-86% F1-Score, NB y *Gradient Boosting* sobresalen con 88%-83% de sensibilidad, mientras que NB, KNN y XGBoost presentan un buen equilibrio entre precisión 88-87% y sensibilidad 88-84% según el AUC 89-88%. En la tabla 6 se muestran los resultados de la evaluación del conjunto de datos Parkinson's telemonitoring.

En la Tabla 6, se observa un rendimiento notable del algoritmo RF con 72.34% de precisión y 68.44% de F1-Score, NB presenta la sensibilidad más alta del 89.98%, pero con una precisión baja del 56.94%. La tabla 7 muestra una comparación de las métricas obtenidas por los algoritmos de aprendizaje automático destacados en la métrica de sensibilidad, implementados en todos los conjuntos de datos.

En el conjunto de datos Parkinson's Disease Data Set, AdaBoost sobresale con un 90.32% de sensibilidad, seguido de SVM con un 93.55%. Para Parkinson Speech Dataset with Multiple Types of Sound Recordings, SVM (71.00%) y NB (70.00%) son los más sensibles, para Parkinson's Disease Classification, Bagged SVM (100.00%) y RF (98.25%) lideran en sensibilidad. RF (88.00%) y NB (88.00%) nuevamente se destacan en Replicated Acoustic Features Dataset, mientras que en Parkinson's telemonitoring, KNN (65.65%) y RF (64.93%) son los más sensibles.

5. Discusión

Los resultados obtenidos durante el análisis realizado arrojaron un rango de sensibilidad de 50% al 100% de sensibilidad, la sensibilidad de los algoritmos de inteligencia artificial, es crucial para diagnosticar la enfermedad de Parkinson porque determina la capacidad del modelo para detectar correctamente los casos positivos de la enfermedad, independientemente, el rendimiento de cada modelo varía significativamente según el conjunto de datos utilizado. Este resultado podría deberse a la calidad o naturaleza de los datos, la cantidad de datos de entrenamiento y la complejidad del modelo utilizado. Ahora, profundizando más en los resultados obtenidos, se tiene que los 3 algoritmos con mejor rendimiento en la métrica de sensibilidad fueron, Bagged SVM 100%, RF 98%, y *Gradient Boosting* 97%, el rendimiento de exactitud en la clasificación de paciente con enfermedad de Parkinson obtuvo un 87%, 88% y 90% respectivamente, en el conjunto de datos Parkinson's Disease Classification, mientras que para el conjunto de datos Parkinson's Data Set tuvo valores arriba del 84% de exactitud con los algoritmos SVM, Bagged SVM, y AdaBoost; para el resto de los conjuntos de datos (Parkinson Multiple Sound Recording, Parkinson Replicated Acoustic Features, Parkinson's telemonitoring) fueron los menos sensibles, arrojando resultados entre 90-93%. Por último, considerando los resultados de este análisis, se determina que los atributos más óptimos para la detección de la enfermedad de Parkinson son la combinación de atributos con, 1) datos de irregularidad en la frecuencia ya que se estima que con un porcentaje de irregularidad superior a 1.040 % ya existiría patología y 2) Datos de irregularidad en la

amplitud (Shimmer) ya que se estima que una voz es patológica a partir de un porcentaje de irregularidad del 3.810.

6. Conclusiones

La falta de un diagnóstico temprano y accesible para la detección de la enfermedad de Parkinson evita que las personas con esta enfermedad reciban el tratamiento adecuado a tiempo, generando problemas en su entorno social, laboral y personal. En este trabajo se analizaron diferentes conjuntos de datos con distintas técnicas de aprendizaje automático, de *Boosting* y *Bagging*, para identificar que algoritmos ofrecen los mejores resultados, así como también los atributos de la voz más relacionados para la identificación de la enfermedad de Parkinson, durante el desarrollo de este análisis se detectó que la exactitud en la detección puede variar significativamente según el conjunto de datos utilizado y las técnicas de inteligencia artificial utilizadas en este trabajo. Los algoritmos con mayor puntaje obtenido para esta métrica fueron Bagged SVM, RF, *Gradient Boosting*, aplicados sobre el conjunto de datos Parkinson's Disease Classification, donde los resultados arrojaron un rango entre un 87% y 90% de exactitud. El análisis de la voz con aprendizaje automático y colaboración médica permite la detección temprana del Parkinson de forma precisa y eficiente. En resumen, este análisis presenta resultados prometedores, demostrando que las técnicas de aprendizaje automático, *Boosting* y *Bagging* son útiles en la detección de la enfermedad de Parkinson, pero se necesita un conjunto de datos con suficiente información y características específicas para realizar una correcta identificación. Como trabajo a futuro se considera el uso de más conjuntos de datos similares o diferentes para complementar la información, también la evaluación de métricas en otros lenguajes de programación como R.

Agradecimientos. Este trabajo de investigación fue patrocinado por el Consejo Nacional de Humanidades, Ciencia y Tecnología de México (CONAHCYT) y la Secretaría de Educación Pública (SEP) de México a través del programa PRODEP. Adicionalmente, se agradece al Tecnológico Nacional de México (TecNM) por apoyar este proyecto.

Referencias

1. Martínez-Fernández, R., Gasca-Salas, C., Sánchez-Ferro, Á., Obeso, J.Á.: Actualización en la enfermedad de Parkinson. *Revista Médica Clínica Las Condes*, vol. 27, pp. 363–379 (2016). DOI: 10.1016/j.rmcl.2016.06.010.
2. Jankovic, J.: Parkinson's disease: Clinical Features and Diagnosis. *Journal of neurology, Neurosurgery & Psychiatry*, vol. 79, no. 4, pp. 368–376 (2008)
3. Archive: Bradley's neurology in clinical practice. <https://archive.org/details/bradleysneurolog00unse> (2024)
4. WHO: Enfermedad de Parkinson. <https://www.who.int/es/news-room/fact-sheets/detail/parkinson-disease> (2024)
5. Martínez-Ramírez, D., Rodríguez-Violante, M., Velázquez-Ávila, E.S., Cervantes-Arriaga, A., González-Cantú, A., Corona, T., Velásquez-Pérez, L.: Incidencia y distribución

- geográfica de la enfermedad de Parkinson en México. *Salud Pública de México*, vol. 62, no. 6, pp. 873–875 (2020). DOI: 10.21149/11750.
6. Jiang, H., Lim, W.Y.B., Ng, J.S., Wang, Y., Chi, Y., Miao, C.: Towards Parkinson's Disease Prognosis using Self-supervised Learning and Anomaly Detection. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings*, pp. 3960–3964 (2021). DOI: 10.1109/ICASSP39728.2021.9414840.
 7. Hyder, R., Jensen, M., Højlund, A., Kimppa, L., Bailey, C.J., Schaldemose, J.L., Shtyrov, Y.: Functional Connectivity of Spoken Language Processing in Early-stage Parkinson's disease: An MEG Study. *NeuroImage: Clinical*, vol. 32, pp. 102718 (2021). DOI: 10.1016/j.nicl.2021.102718.
 8. Nogales, A., García-Tejedor, Á.J., Maitín, A.M., Pérez-Morales, A., Del Castillo, M.D., Romero, J.P.: BERT Learns from Electroencephalograms about Parkinson's Disease: Transformer-based Models for Aid Diagnosis. *IEEE Access*, vol. 10, pp. 101672–101682 (2022). DOI: 10.1109/ACCESS.2022.3201843.
 9. Nissar, I., Rizvi, D.R., Masood, S., Mir, A.N.: Voice-based Detection of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study. *EAI endorsed Transactions on Pervasive Health and Technology*, vol. 5, no. 19, pp. e2–e2 (2019). DOI: 10.4108/eai.13-7-2018.162806.
 10. Géron, A.: *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc. (2022)
 11. Orozco-Arroyave, J.R., Vásquez-Correa, J.C., Vargas-Bonilla, J.F., Arora, R., Dehak, N., Nidadavolu, P.S., Noeth, E.: NeuroSpeech: An Open-source Software for Parkinson's Speech Analysis. *Digital Signal Processing*, vol. 77, pp. 207–221 (2018). DOI: 10.1016/j.dsp.2017.07.004.
 12. Quan, C., Ren, K., Luo, Z.: A Deep Learning based Method for Parkinson's Disease Detection using Dynamic Features of Speech. *IEEE* vol. 9, pp. 10239–10252 (2021). DOI: 10.1109/ACCESS.2021.3051432.
 13. Klempíř, O., Krupička, R.: Machine Learning using Speech Utterances for Parkinson Disease Detection. *Lékař a technika-Clinician and Technology*, vol. 48, no. 2, pp. 66–71 (2018)
 14. Rahman, W., Lee, S., Islam, M.S., Antony, V.N., Ratnu, H., Ali, M.R., Hoque, E.: Detecting Parkinson Disease using a Web-based Speech Task: Observational Study. *Journal of Medical Internet Research*, vol. 23, no. 10, pp. e26305 (2021). DOI:10.2196/26305.
 15. Almeida, J.S., Rebouças-Filho, P.P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., de Albuquerque, V.H.C.: Detecting Parkinson's Disease with Sustained Phonation and Speech Signals using Machine Learning Techniques. *Pattern Recognition Letters*, vol. 125, pp. 55–62 (2019). DOI: 10.1016/j.patrec.2019.04.005.
 16. Thanoun, M.Y., Yaseen, M.T., Aleesa, A.M.: Development of Intelligent Parkinson Disease Detection System based on Machine Learning Techniques using Speech Signal. *International Journal on Advanced Science Engineering and Information Technology*, vol. 11, no. 1, pp. 388–392 (2021). DOI: 10.18517/ijaseit.11.1.12202.
 17. Zhang, L., Qu, Y., Jin, B., Jing, L., Gao, Z., Liang, Z.: An Intelligent Mobile-enabled System for Diagnosing Parkinson Disease: Development and Validation of a Speech Impairment Detection System. *JMIR Medical Informatics*, vol. 8, no. 9, pp. e18689 (2020). DOI: 10.2196/18689.
 18. Karan, B., Sahu, S.S., Mahto, K.: Parkinson Disease Prediction using Intrinsic Mode Function based Features from Speech Signal. *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 249–264 (2020). DOI: 10.1016/j.bbe.2019.05.005.

19. Quan, C., Ren, K., Luo, Z., Chen, Z., Ling, Y.: End-to-end deep Learning Approach for Parkinson's disease Detection from Speech Signals. *Biocybernetics and Biomedical Engineering*, vol. 42, no. 2, pp. 556–574 (2022). DOI: 10.1016/j.bbe.2022.04.002.
20. Hoq, M., Uddin, M.N., Park, S.B.: Vocal Feature Extraction-based Artificial Intelligent Model for Parkinson's Disease Detection. *Diagnostics*, vol. 11, no. 6, pp. 1076 (2021). DOI: 10.3390/diagnostics11061076.
21. Khan, T., Westin, J., Dougherty, M.: Classification of Speech Intelligibility in Parkinson's disease. *Biocybernetics and Biomedical Engineering*, vol. 34, no. 1, pp. 35–45 (2014). DOI: 10.1016/j.bbe.2013.10.003.
22. Wroge, T.J., Özkanca, Y., Demiroglu, C., Si, D., Atkins, D.C., Ghomi, R.H.: Parkinson's disease Diagnosis using Machine Learning and Voice. In *IEEE Signal Processing in Medicine and Biology Symposium*, pp. 1–7 (2018). DOI: 10.1109/SPMB.2018.8615607.
23. Soumaya, Z., Taoufiq, B.D., Benayad, N., Yunus, K., Abdelkrim, A.: The Detection of Parkinson disease using the Genetic Algorithm and SVM classifier. *Applied Acoustics*, vol. 171, pp. 107528 (2021). DOI: 10.1016/j.apacoust.2020.10752.
24. Mostafa, S.A., Mustapha, A., Khaleefah, S.H., Ahmad, M.S., Mohammed, M.A.: Evaluating the Performance of three Classification Methods in Diagnosis of Parkinson's disease. In *Recent Advances on Soft Computing and Data Mining: Proceedings of the Third International Conference on Soft Computing and Data Mining*, Springer International Publishing, vol. 700, pp. 43–52 (2018). DOI: 10.1007/978-3-319-72550-5_5.
25. Kaggle: Parkinson's Disease Data Set, <https://www.kaggle.com/datasets/vikasukani/parkinsons-disease-data-set> (2024)
26. Little, M., McSharry, P., Hunter, E., Spielman, J., Ramig, L.: Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's disease. *Nature Proceedings*, pp. 1015–1022 (2009)
27. ICS: Parkinson Speech Dataset with Multiple Types of Sound Recordings - UCI Machine Learning Repository. <https://archive.ics.uci.edu/dataset/301/parkinson+speech+dataset+with+multiple+types+of+sound+recordings> (2024)
28. Sakar, B.E., Isenkul, M.E., Sakar, C.O., Sertbas, A., Gurgun, F., Delil, S., Kursun, O.: Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound recordings. *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828–834 (2013). DOI: 10.1109/JBHI.2013.2245674.
29. ICS: Parkinson's Disease Classification - UCI Machine Learning Repository. <https://archive.ics.uci.edu/dataset/470/parkinson+s+disease+classification> (2024)
30. Sakar, C.O., Serbes, G., Gunduz, A., Tunc, H.C., Nizam, H., Sakar, B.E., Apaydin, H.: A Comparative Analysis of Speech Signal Processing Algorithms for Parkinson's Disease Classification and the Use of the Tunable Q-factor Wavelet Transform. *Applied Soft Computing*, vol. 74, pp. 255–263 (2019). DOI: 10.1016/j.asoc.2018.10.022.
31. ICS: Parkinson Dataset with replicated acoustic features - UCI Machine Learning Repository. <https://archive.ics.uci.edu/dataset/489/parkinson+dataset+with+replicated+acoustic+features> (2024)
32. Naranjo, L., Pérez, C.J., Martín, J.: Addressing Voice Recording Replications for Tracking Parkinson's disease Progression. *Medical & Biological Engineering & Computing*, vol. 55, pp. 365–373 (2017). DOI: 10.1007/s11517-016-1512-y.
33. ICS: Parkinsons Telemonitoring - UCI Machine Learning Repository. <https://archive.ics.uci.edu/dataset/189/parkinsons+telemonitoring> (2024)
34. Tsanas, A., Little, M., McSharry, P., Ramig, L.: Accurate Telemonitoring of Parkinson's Disease Progression by Non-invasive Speech Tests. *Nature preceding* (2010). DOI: 10.1038/npre.2009.3920.1.