

Selección de parámetros de la voz para el diagnóstico y monitoreo de pacientes con la enfermedad del Parkinson

García Santamaría, Valery Eloisa. Narro Silva, Jairo David

Resumen: La enfermedad de Parkinson Idiopática (IPD) es el segundo trastorno neurodegenerativo más común en el mundo. Debido a la inexistencia de una cura, los tratamientos actuales solo se enfocan en mejorar la calidad de vida al tratar de ralentizar el avance de la enfermedad. Por lo tanto, los pacientes requieren monitoreo constante y continuas visitas a las clínicas. Debido a esto, es necesaria la identificación de marcadores biológicos que permitan no solo el diagnóstico temprano de la enfermedad sino también el telemonitoreo de los pacientes. Distintos parámetros de las señales vocales pueden ser utilizados como potenciales marcadores para el seguimiento de la enfermedad del Parkinson. En este artículo presentamos los resultados de la aplicación de un modelo de Machine Learning para la selección de parámetros de la voz relevantes que permitan monitorear a un paciente con IPD. Esto se hizo a partir de una base de datos que contenía 195 registros de voz provenientes de personas sanas y pacientes con IPD. El modelo de clasificación que obtuvo el mejor resultado fue k-nearest neighbors, con una precisión de 97.8%. Así mismo, se identificó a los atributos HNR, RPDE y DFA como potenciales parámetros de la voz para el diagnóstico y monitoreo de pacientes.

Palabras Clave: *Parkinson, Inteligencia Artificial, Machine Learning, voz humana.*

I. INTRODUCCIÓN

La enfermedad de Parkinson idiopática (IPD, por sus siglas en inglés) es el trastorno neurodegenerativo más común en el mundo, justo después de la enfermedad del Alzheimer [1]. Debido a que no existe una cura para la enfermedad del Parkinson, los pacientes deben someterse a tratamientos y terapias de forma continua para mejorar su calidad de vida. En los Estados Unidos, los medicamentos tienen un costo promedio de 2,500 dólares anuales y la cirugía terapéutica puede costar hasta 100,000 dólares por persona [2]. Los costos de las visitas continuas a la clínica pueden ser reducidos por medio del monitoreo remoto de pacientes con la enfermedad del Parkinson. Debido a esto, es necesaria la identificación de marcadores biológicos que permitan monitorear la enfermedad.

Además de temblores, rigidez y bradicinesia, del 70% al 80% de los pacientes con IPD sufren de disfonía u otro tipo de desorden fonatorio [3]. Los pacientes con IPD generalmente muestran rangos de tono y volumen reducidos, presiones de umbral de fonación subglotales elevadas y temblor vocal [4]. La identificación de parámetros relevantes en las señales vocales de pacientes con IPD resulta de gran importancia para evaluar el seguimiento de la progresión de la enfermedad.

Los principales métodos de medición de las señales vocales incluyen parámetros como la F0 (la frecuencia fundamental o el tono de la oscilación vocal), el nivel de presión sonora absoluto (que indica el volumen relativo del habla), el jitter (el grado de variación de la F0 del habla de un ciclo vocal a otro), el shimmer (el grado de variación de la amplitud del habla de un ciclo a otro) y la relación entre ruido y armónicos (la amplitud del ruido en relación con los componentes tonales del habla) [5].

En la presente investigación, se busca identificar y seleccionar los parámetros de las señales vocales más relevantes que permitan monitorear el avance progresivo de la enfermedad del Parkinson para una futura aplicación en telemedicina. Así mismo, se busca identificar el modelo de clasificación que permita discriminar de forma efectiva a los pacientes con IPD de los pacientes sanos. De esta forma, estos parámetros también pueden ser útiles para el diagnóstico temprano de la enfermedad del Parkinson.

II. METODOLOGÍA

I. Balance de la base de datos

La base de datos utilizada corresponde a 195 registros de voz pertenecientes a 31 personas, de las cuales 23 presentan IPD [5]. La data se encuentra disponible en el repositorio de Machine Learning de la Universidad de California [6]. Contiene 22 parámetros de la voz (variables independientes), una variable que indica el nombre del sujeto y el número de la grabación y un atributo clasificatorio (variable dependiente) denominado 'status'. Este último atributo adquiere un valor igual a 0 si el registro de voz pertenece a una persona sana o igual a 1 si pertenece a una persona con IPD.

Del total de registros, 48 son de personas sanas y 147 de personas con IPD. En casos donde se tienen muchas más instancias de una clase con respecto a otra, los modelos de clasificación estándar tienden a verse abrumados por las clases grandes e ignoran a las pequeñas [7]. Esto se debe a que la mayoría de los algoritmos de machine learning asumen que el número de instancias en cada clase es similar [8]. Para evitar estos errores de sesgo y obtener un resultado más preciso, se redujo el número de instancias de las personas con IPD a 48. De esta forma, los modelos de clasificación fueron entrenados con una base de datos balanceada.

II. Selección de atributos

La selección de atributos es el proceso de reducir la cantidad de variables de entrada que se utilizarán para entrenar un modelo de machine learning. Este proceso permite reducir el costo computacional del modelamiento y en algunos casos, mejorar el rendimiento del modelo [9]. La tabla I describe los parámetros de las señales vocales que fueron medidos.

Tabla I. Parámetros de las señales vocales

Parámetro	Descripción
MDVP: F0 (Hz)	Frecuencia fundamental
MDVP: Fhi(Hz)	Frecuencia máxima
MDVP: Flo (Hz)	Frecuencia mínima
MDVP: Jitter(%)	Porcentaje de fluctuación de la MDVP
MDVP:Jitter(Abs)	Fluctuación absoluta de la MDVP (ms)
MDVP:RAP	Perturbación de la Amplitud Relativa
MDVP:PPQ	Cociente de Perturbación de 5 puntos
Jitter:DDP	Diferencia absoluta media de las diferencias entre ciclos, dividida por el periodo medio
MDVP:Shimmer	Shimmer local
MDVP:Shimmer(dB)	Shimmer local en decibeles
Shimmer:APQ3	Cociente de Perturbación de 3 puntos
Shimmer:APQ5	Cociente de Perturbación de 5 puntos
MDVP:APQ	Cociente de Perturbación de 11 puntos
Shimmer:DDA	Diferencia absoluta media entre las diferencias consecutivas entre las amplitudes de los períodos consecutivos
NHR	Relación de ruido a armónicos
HNR	Relación de armónicos a ruidos
RPDE	Entropía de densidad del período de recurrencia.

D2	Dimensión de correlación
DFA	Análisis de fluctuación sin tendencia
Spread1, Spread2	Medidas no lineales de la variación de la frecuencia fundamental
PPE	Entropía del período del tono

Debido a que se cuenta con varias variables independientes, la selección de atributos es una parte importante para determinar qué variables serán de utilidad al aplicar los modelos de clasificación. Existen distintos métodos y técnicas para la reducción de dimensionalidad.

Por un lado, la selección de atributos basado en el valor de varianza de cada atributo es uno de los métodos más sencillos. Consiste en retirar todas aquellas variables que presentan un valor de varianza pequeño o cercano a cero. Se considera que estas variables no son adecuadas para el entrenamiento del modelo debido a que presentan un valor similar en cada muestra y, por lo tanto, no brindan información útil. La tabla II muestra el valor de la varianza de cada atributo redondeado a 4 decimales.

Tabla II. Valor de varianza de cada atributo

Parámetro	Varianza
F0 (Hz)	2270.9775
Fhi(Hz)	9786.1821
Flo (Hz)	2558.6014
MDVP: Jitter(%)	0.0000
MDVP:Jitter(Abs)	0.0000
MDVP:RAP	0.0000
MDVP:PPQ	0.0000
Jitter:DDP	0.0000
MDVP:Shimmer	0.0001
MDVP:Shimmer(dB)	0.0109
Shimmer:APQ3	0.0000
Shimmer:APQ5	0.0000
MDVP:APQ	0.0001
Shimmer:DDA	0.0004
NHR	0.0004
HNR	13.1929
RPDE	0.0121
D2	0.1193
DFA	0.0021

Spread1	0.9793
Spread2	0.0066
PPE	0.0042

Sin embargo, este método no toma en cuenta la relación que pueda existir entre las variables independientes o la de una variable independiente con la variable objetivo (variable dependiente). Por lo tanto, existe el riesgo de perder información útil.

Por otro lado, la selección de atributos basado en la correlación tiene en cuenta tanto la relación de las variables independientes entre sí como la relación entre la variable objetivo con las variables independientes. Si existe una alta correlación entre dos variables, se considera que llevan información similar y por lo tanto sólo una de ellas sería útil para el entrenamiento del modelo. Además, aquellas variables que presenten una correlación baja con la variable objetivo pueden considerarse irrelevantes para el entrenamiento del modelo. La figura I muestra los resultados de la correlación entre los 22 atributos y la variable objetivo 'status'. La figura II muestra los resultados de la correlación sin tener en cuenta los atributos que presentan un valor bajo de varianza.

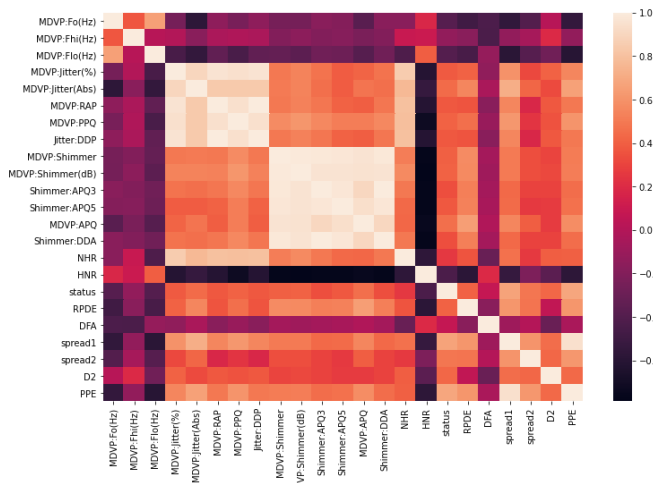


Figura I. Resultados de la correlación para todos los atributos

Por ejemplo, se puede observar que las variables MDVP:Shimmer, MDVP:Shimmer(dB), Shimmer(APQ3), Shimmer(APQ5), MDVP(APQ) y Shimmer DDA presentan una gran correlación entre ellas. Por lo tanto, se puede considerar que sólo una de ellas tendrá información relevante para el entrenamiento del modelo y las demás pueden considerarse como información redundante.

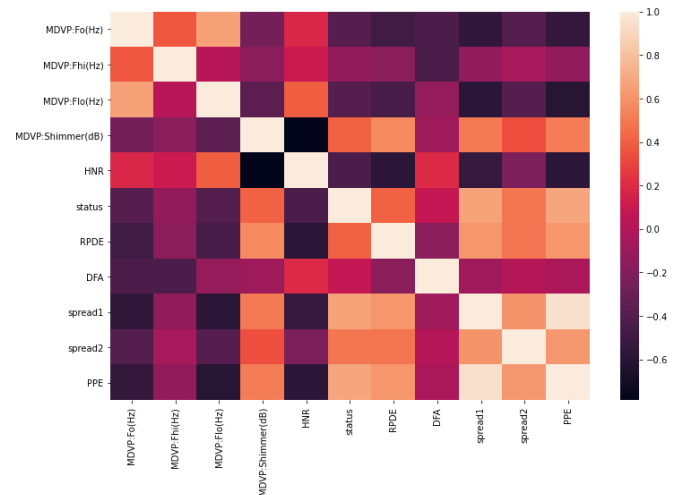


Figura II. Resultados de la correlación sin tener en cuenta los atributos con poca varianza

Otros métodos para la selección de atributos son forward selection, backward selection y stepwise selection. Por un lado, forward selection inicia entrenando el modelo con un solo atributo y añade más atributos de uno a uno en cada paso, seleccionando las combinaciones que otorgan los mejores resultados. Por otro lado, backward elimination inicia entrenando el modelo con todos los atributos, eliminando en cada paso el atributo menos relevante. Finalmente, el proceso en stepwise selection es similar al de forward selection, pero también puede eliminar en cada paso los atributos que no considera relevantes.

Para la base de datos, se aplicaron los métodos de forward y backward selection. Para cada método se obtuvo una combinación de 8 atributos. Además, teniendo en cuenta los valores de la varianza y los resultados de la correlación, se realizaron otras combinaciones de parámetros con los que se entrenaron los modelos.

III. Selección del modelo de clasificación

Se entrenaron 10 modelos de clasificación con las distintas combinaciones realizadas. Los diez modelos son los siguientes:

1. K-Nearest Neighbors
2. Linear Support Vector Machine
3. Radial Basis Function Kernel
4. Gaussian Process
5. Decision Trees
6. Naive Bayes
7. Random Forest
8. Red neuronal
9. Quantitative Descriptive Analysis)
10. AdaBoost

Para todos los modelos se aplicó el método k-fold cross validation (validación cruzada) con un k igual a 10. En este método, los datos son divididos en k segmentos de igual tamaño. Posteriormente, se realizan k iteraciones de entrenamiento y validación de manera que dentro de cada iteración se mantiene un segmento diferente de los datos para su validación, mientras que los k-1 pliegues restantes se utilizan para el aprendizaje [10].

IV. RESULTADOS

Teniendo en cuenta los valores de varianza y los resultados de la correlación entre los atributos, se realizaron diferentes combinaciones para la obtención del mejor resultado. Para todas las combinaciones, el mejor modelo de clasificación fue el k-nearest neighbors. La tabla III muestra el puntaje de precisión (accuracy score) de todos los modelos de clasificación para la mejor combinación.

Teniendo en cuenta lo discutido anteriormente, aunque se tenga un buen resultado de correlación.

Tabla III. Resultados de los métodos de clasificación para la mejor combinación

Método	Accuracy Score
k-nearest neighbors	0.98
Gaussian Process	0.93
Radial Basis Function Kernel	0.83
Random Forest	0.83
Neural Networks	0.81
Support Vector Machine	0.81
Adaptive Boosting	0.78
Decision Trees	0.77
Naive Bayes Classifier	0.73
Quadratic Discriminant Analysis	0.72

Las combinaciones obtenidas con los métodos de Forward y Backward Selection contaban con diferentes parámetros pero obtuvieron el mismo puntaje de precisión. La tabla IV muestra los atributos que se seleccionaron con estos métodos.

Tabla IV. Combinaciones obtenidas con los métodos Forward y Backward Selection

Método	Atributos seleccionados
Forward Selection	MDVP:F0(Hz), MDVP:Fhi(Hz), MDVP:Jitter(Abs), NHR, HNR, RPDE, DFA y spread1

Backward Selection	MDVP:F0(Hz), MDVP:RAP, MDVP:Shimmer(dB), HNR, RPDE, DFA, spread1 y D2.
--------------------	--

La tabla V muestra todas las combinaciones que se realizaron y sus respectivos puntajes de precisión utilizando el modelo k-nearest neighbors.

Tabla V. Resultados de las distintas combinaciones

Combinaciones	Accuracy Score KNN
MDVP:F0(Hz), MDVP:Fhi(Hz), MDVP:Jitter(Abs), NHR, HNR, RPDE, DFA, spread1 (8)	0.9789
MDVP:F0(Hz), MDVP:RAP, MDVP:Shimmer(dB), HNR, RPDE, DFA, spread1, D2 (8)	0.9789
MDVP:F0(Hz), MDVP:Fhi(Hz), Shimmer:DDA, HNR, RPDE, DFA, PPE (7)	0.9667
MDVP:F0(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:Shimmer(dB), HNR, RPDE, DFA, spread1, spread2, PPE (10)	0.9344
MDVP:F0(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:Shimmer, MDVP:Shimmer(dB), MDVP:APQ, Shimmer:DDA, NHR, HNR, RPDE, DFA, spread1, spread2, D2, PPE (15)	0.9133
Todos (22)	0.8911

V. DISCUSIONES

Se redujo el número de variables independientes de 22 a 8 atributos. Sin embargo, existen dos combinaciones de atributos diferentes obtenidas que otorgan el mismo resultado de precisión. Estas combinaciones se pueden observar en la tabla IV. En ambas combinaciones, se repiten los atributos MDVP:F0(Hz), HNR, RPDE, DFA y spread1.

De acuerdo a los resultados del modelo de clasificación, se puede considerar que todos los atributos seleccionados por los métodos de forward y backward selection son parámetros relevantes para discriminar a una persona con IPD de una persona sana. Por lo tanto, pueden ser utilizados para el telemonitoreo de pacientes con IPD. Sin embargo, se presentan algunas variables confusoras que no fueron consideradas en el modelo de clasificación que pueden afectar la viabilidad del uso de un parámetro.

Por ejemplo, los pacientes con IPD presentan un valor alto en el parámetro MDVP:F0(Hz) (frecuencia máxima) [11]. Esto puede observarse en la función de densidad de probabilidad de la figura III. Sin embargo, este parámetro también puede verse afectado por el género o las preferencias individuales del paciente [12]. Por otro lado, debido a que muchas de las señales de voz son ruidosas, el valor de D2 se aumenta falsamente, por lo que resulta ser un parámetro poco confiable [5].

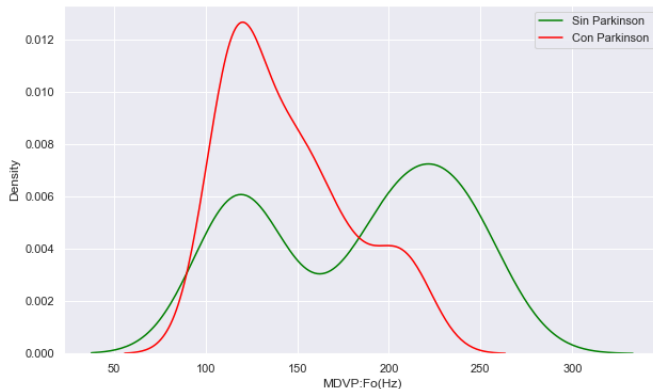


Figura III. Función de densidad de probabilidad para el parámetro MDVP:F0(Hz)

El parámetro HNR se presenta como una perturbación en la razón de ruidos/armónicos y evalúa el grado de ronquera de un paciente con IPD [11]. Este parámetro presenta un valor más bajo en los pacientes con IPD por lo que resulta útil en el entrenamiento del modelo. Esto se puede observar en la función de densidad de probabilidad de la figura IV.

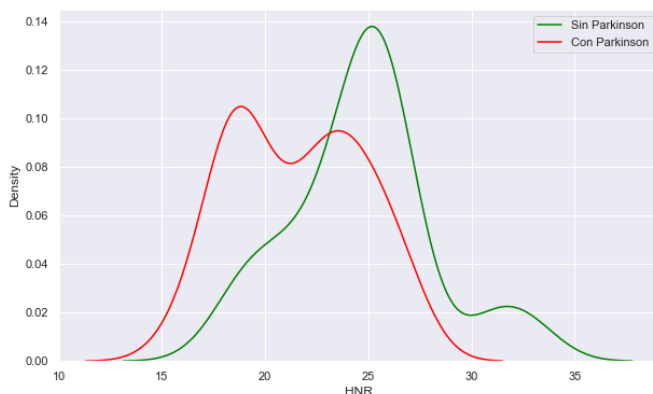


Figura IV. Función de densidad de probabilidad para el parámetro HNR

La mayoría de las medidas tradicionales del Shimmer y Jitter producen valores cercanos a cero, mientras que las medidas novedosas y no estándar y las relaciones de armónicos a ruido se distribuyen de manera más uniforme en un rango más amplio de valores [5].

La combinación de los parámetros RPDE (Entropía de Densidad del Período de Recurrencia) y DFA (Análisis de Fluctuación sin tendencia) permiten una mejor precisión en el proceso de clasificación [13]. Esto debido a que, a diferencia de los parámetros estándar, estos parámetros fueron medidos teniendo en cuenta la aleatoriedad y el ruido inherentes a las señales vocales [5].

VI. CONCLUSIONES

La identificación de parámetros de la voz resulta de gran ayuda para el diagnóstico temprano del Parkinson, ya que los trastornos del habla pueden aparecer mucho antes de los temblores musculares. Además, los parámetros seleccionados resultan útiles para monitorear a los pacientes con IPD. El seguimiento de la variación de estos parámetros ayudaría a identificar qué tratamientos o medicamentos resultan más efectivos para ralentizar la progresión de la enfermedad en el paciente. Tanto el monitoreo como el diagnóstico se pueden realizar por medio de un smartphone y conectividad a internet, reduciendo así el número de visitas a las clínicas [14]. Otros indicadores relevantes que pueden ser utilizados para el monitoreo de pacientes con IDP son los cambios en la marcha, la memoria y el tapping [15]. Finalmente, es importante tener en cuenta las características del paciente. Esto para evitar los errores que puedan ocurrir en el diagnóstico o monitoreo debido a la presencia de variables confusoras que puedan aumentar falsamente el puntaje de precisión del modelo de clasificación.

REFERENCIAS

- [1] W. Poewe *et al.*, “Parkinson disease”, *Nat Rev Dis Primers*, vol. 3, núm. 1, pp. 1–21, mar. 2017, doi: 10.1038/nrdp.2017.13.
- [2] “Statistics”, *Parkinson’s Foundation*. <https://www.parkinson.org/Understanding-Parkinsons/Statistics> (consultado jul. 19, 2021).
- [3] Y. Wu *et al.*, “Dysphonic Voice Pattern Analysis of Patients in Parkinson’s Disease Using Minimum Interclass Probability Risk Feature Selection and Bagging Ensemble Learning Methods”, *Computational and Mathematical Methods in Medicine*, vol. 2017, p. e4201984, may 2017, doi: 10.1155/2017/4201984.
- [4] G. K. Sewall, J. Jiang, y C. N. Ford, “Clinical evaluation of Parkinson’s-related dysphonia”, *Laryngoscope*, vol. 116, núm. 10, pp. 1740–1744, oct. 2006, doi: 10.1097/01.mlg.0000232537.58310.22.
- [5] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, y L. O. Ramig, “Suitability of dysphonia measurements for telemonitoring of Parkinson’s disease”, *IEEE Trans Biomed Eng*, vol. 56, núm. 4, p. 1015, abr. 2009, doi: 10.1109/TBME.2008.2005954.

- [6] “UCI Machine Learning Repository: Parkinsons Data Set”. <https://archive.ics.uci.edu/ml/datasets/parkinsons> (consultado may. 23, 2021).
- [7] N. V. Chawla, N. Japkowicz, y A. Kotcz, “Editorial: special issue on learning from imbalanced data sets”, *SIGKDD Explor. Newsl.*, vol. 6, núm. 1, pp. 1–6, jun. 2004, doi: 10.1145/1007730.1007733.
- [8] B. Krawczyk, “Learning from imbalanced data: open challenges and future directions”, *Prog Artif Intell*, vol. 5, núm. 4, pp. 221–232, nov. 2016, doi: 10.1007/s13748-016-0094-0.
- [9] J. Brownlee, “How to Choose a Feature Selection Method For Machine Learning”, *Machine Learning Mastery*, nov. 26, 2019.
<https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/> (consultado jun. 15, 2021).
- [10] P. Refaeilzadeh, L. Tang, y H. Liu, “Cross-Validation”, en *Encyclopedia of Database Systems*, L. LIU y M. T. ÖZSU, Eds. Boston, MA: Springer US, 2009, pp. 532–538. doi: 10.1007/978-0-387-39940-9_565.
- [11] R. Chiamonte y M. Bonfiglio, “Acoustic analysis of voice in Parkinson’s disease: a systematic review of voice disability and meta-analysis of studies”, *Rev Neurol*, vol. 70, núm. 11, pp. 393–405, jun. 2020, doi: 10.33588/rn.7011.2019414.
- [12] L. Cnockaert, J. Schoentgen, P. Auzou, C. Ozsancak, L. Defebvre, y F. Grenez, “Low-frequency vocal modulations in vowels produced by Parkinsonian subjects”, *Speech Communication*, vol. 50, núm. 4, pp. 288–300, abr. 2008, doi: 10.1016/j.specom.2007.10.003.
- [13] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. E. Costello, y I. M. Moroz, “Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection”, *Biomed Eng Online*, vol. 6, p. 23, jun. 2007, doi: 10.1186/1475-925X-6-23.
- [14] S. Singh y W. Xu, “Robust Detection of Parkinson’s Disease Using Harvested Smartphone Voice Data: A Telemedicine Approach”, *Telemedicine and e-Health*, vol. 26, núm. 3, pp. 327–334, mar. 2020, doi: 10.1089/tmj.2018.0271.
- [15] B. M. Bot *et al.*, “The mPower study, Parkinson disease mobile data collected using ResearchKit”, *Sci Data*, vol. 3, núm. 1, p. 160011, mar. 2016, doi: 10.1038/sdata.2016.11.