

Análisis predictivo enfocado a la precisión de detección de enfermos de Parkinson

Sánchez Grigioni, Tomás; Robledo Ferri, Candela; Pérez, Silvia N.

Universidad Nacional de La Matanza, Departamento de Ingeniería e Investigaciones Tecnológicas

Abstract

La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más frecuente, luego del Alzheimer. En la actualidad, se considera que en Argentina el 2% de personas mayores de 65 años cursan la enfermedad. Hasta hace unas décadas la detección de casos se realizaba a partir de la consulta médica. Esto subestimaba la estimación del número de casos y no permitía una detección precoz de la enfermedad, necesaria para la realización de tratamientos y también para la investigación de potenciales poblaciones de riesgo. Investigaciones más recientes han propuesto la detección de la enfermedad a partir de grabaciones de la voz al identificar anomalías en las señales acústicas en presencia de EP. Como parte de un trabajo de beca se propuso, como actividad motivadora de la investigación, seleccionar 100 personas que presenten signos de enfermedad de Parkinson para incluirlas en una investigación de nuevas técnicas de tratamiento. Dicha selección debía realizarse a partir de datos de grabaciones de audios para lo cual en este trabajo se aplicaron modelos estadísticos para la predicción de quienes padecen esta enfermedad sobre el conjunto de datos disponible en la base mPower. Se utilizaron Python y R para ajustar y seleccionar el modelo que obtenga mayor precisión en la detección de EP. Los resultados muestran que la mayor precisión se logró con K-Nearest Neighbors, obteniendo además altos valores predictivos con este modelo.

Palabras Clave

Enfermedad de Parkinson, modelos de clasificación, grabaciones de voz.

Introducción

La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más común, siguiendo al Alzheimer. (Langston, 2002). Sus síntomas incluyen rigidez, alteración del equilibrio, temblor en reposo y lentitud de movimientos. Además de estos síntomas motores, la voz también se ve

afectada. El deterioro de la voz y el habla es un síntoma típico de la EP que se da en la mayoría de los pacientes.

La detección de la EP suele requerir una evaluación profesional por parte de un experto, y la investigación de la voz como biomarcador de la enfermedad podría ser eficaz para evitar un retraso considerable en su diagnóstico. En el caso de la EP y otras enfermedades neurodegenerativas, diversos estudios (Tsanas y otros, 2010; Gómez-Vilda y otros, 2015) han identificado anomalías en grabaciones de voz de los pacientes, poniendo de manifiesto la posibilidad de usar técnicas de análisis de señales acústicas para el diagnóstico precoz de la enfermedad.

Se considera que el diagnóstico a través del análisis de trastornos del habla no solo evitaría las complicaciones de los métodos invasivos, sino que además puede implementarse con facilidad utilizando tecnología. Por lo tanto, se necesita una herramienta clínica económica y objetiva para seguir la progresión de síntomas de la EP. Los avances en el análisis de señales y el procesamiento del habla han introducido soluciones novedosas en el proceso de diagnóstico. Los conocimientos adquiridos gracias a las técnicas más avanzadas de reconocimiento del habla han permitido mejorar la detección de patologías de la voz. El habla del parkinsoniano se caracteriza por tener una sonoridad e intensidad monótona, de bajo tono y pobremente prosódica, que tiende a desvanecerse al final de la fonación. El habla se produce en lentos ataques y significativas pausas para respirar entre

palabras y sílabas, reduciéndose la fluidez verbal y el ritmo (Skodda y Schlegel, 2008). La articulación de los sonidos, tanto linguales, como labiales, se empobrecen reduciendo significativamente su inteligibilidad (Miller et al. 2007) y dificultando la identificación de su estado emocional e intenciones.

La rigidez asociada afecta al sistema respiratorio encargado de facilitar el flujo y la presión de aire necesarios para la emisión vocal lo que se traduce en una disminución de presión subglótica dando como resultado menor intensidad de la voz. El deterioro gradual de la capacidad de comunicación en los pacientes con EP se considera una causa importante de discapacidad.

El análisis acústico de la voz podría ser un medio para una descripción más objetiva de los trastornos de la voz en pacientes con EP y podría permitir una estimación más precisa de los cambios relacionados con la EP en el rendimiento vocal.

En este trabajo se analiza la clasificación de personas con y sin EP, a partir de parámetros obtenidos a partir de la señal de voz y aplicando distintas alternativas de modelos de predicción. El objetivo, propuesto a partir de una problemática motivadora de investigación de modelos de minería de datos, fue seleccionar 100 personas con EP a quienes se pudiera proponer incluirlos en un nuevo tratamiento. A partir del ajuste de modelos predictivos se buscó seleccionar el que permite captar la mayor cantidad de pacientes que efectivamente presentan EP entre los 100 seleccionados, logrando mejor precisión.

Trabajos Relacionados

El conjunto de datos mPower fue objeto de diversos análisis encontrados en la bibliografía. En Bot, et al (2016) se presenta un análisis demográfico de las características relevadas sobre el conjunto de grabaciones de voz. Por citar algunos trabajos actuales con esta base, en Almaloglou et al, 2020, se propone un modelo para diferenciar las voces patológicas de las sanas basado en redes neuronales aplicadas a los

espectrogramas del habla, logrando altos valores de clasificación. Así también, Tougui et al (2020) comparan distintos clasificadores sobre la base mPower, realizando previamente una selección de características.

En el trabajo de Giuliano et al (2019), se realizó sobre la base mPower una selección de características a partir de análisis de componentes principales y ANOVA. En este artículo se aplicaron modelos de regresión logística y redes neuronales con resultados de alta capacidad predictiva.

Base de datos

El presente trabajo analiza un conjunto de datos de acceso abierto denominado mPower disponible en la plataforma Synapse (www.synapse.org). La base mPower (*Mobil Parkinson Disease Study*, Bot et al 2016), corresponde a un estudio observacional basado en grabaciones realizadas por las personas en sus teléfonos móviles (smartphones) que muestra nuevos enfoques para monitorear los indicadores clave de la progresión y el diagnóstico de la EP.

De la base de datos disponible en la plataforma mPower-Synapse, en este trabajo se consideraron como variables predictivas a las obtenidas mediante procesamiento automático de las fonaciones utilizando el software libre OpenSmile, el cual resulta en 62 características o parámetros. se consideró también la selección realizada por Giuliano et al, (2019) de 5 parámetros de voz para grabaciones de la fonación la vocal /a/. Tomando en cuenta la multicolinealidad presente en la base de 62 características, se realizó un análisis de componentes principales a partir del cual se seleccionaron 19 parámetros de voz, a los que se unieron otros como edad y género. La variable respuesta se toma como la indicadora de que la persona tiene diagnóstico de EP (verdadero/falso).

Si bien la base tiene registros de 2253 personas de 35 años o más, luego de revisar inconsistencias se redujo a 2218 casos de los

cuales 940 corresponden a personas con EP y 1278 a personas sanas.

Metodología

Se buscó identificar el diagnóstico de la enfermedad de Parkinson (EP) y la enfermedad de no Parkinson (no EP) a partir del análisis acústico de la fonación.

Para esto, se ajustaron diferentes modelos predictivos utilizando los softwares R y Python utilizando paquetes específicos, contrastando las posibilidades de cada software para entregar los resultados.

En la literatura revisada observamos que se utilizan diferentes modelos de predicción, de los cuales se consideraron aquí algunos teniendo en cuenta la relación entre simplicidad de aplicación y resultados de predicción. Los modelos considerados para la predicción son los siguientes:

- **Regresión logística:** similar a la regresión lineal, es un método de clasificación supervisado que modela la relación de una variable respuesta categórica con variables numéricas y/o categóricas.
- **Naive Bayes:** es un algoritmo supervisado destinado a la clasificación. Se fundamenta en el teorema de Bayes, realizando suposiciones sobre las variables predictoras.
- **K-Nearest Neighbors (KNN):** es un algoritmo supervisado destinado a la clasificación. Para la clasificación de una nueva observación se basa en las K observaciones del entrenamiento más cercanas. Para calcular la distancia entre dos observaciones existen distintas medidas, como por ejemplo la distancia euclidiana, y siempre se basan en las variables predictoras.

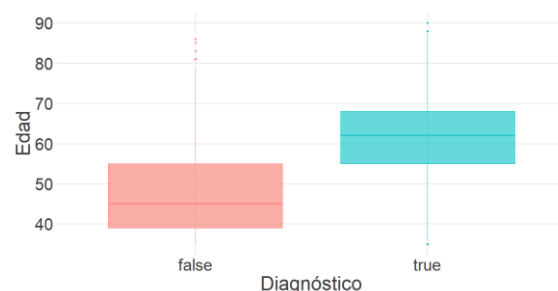
Descripción de los datos analizados

La base de datos analizada está compuesta por 2218 registros de personas de las cuales 940 son enfermos con EP (Diagnóstico=true) y 1278 son personas sanas (Diagnóstico=false).

Del conjunto total de registros en la base mPower, el dataset considerado corresponde a personas de al menos 35 años de edad. En la Figura 1 observamos los diagramas de

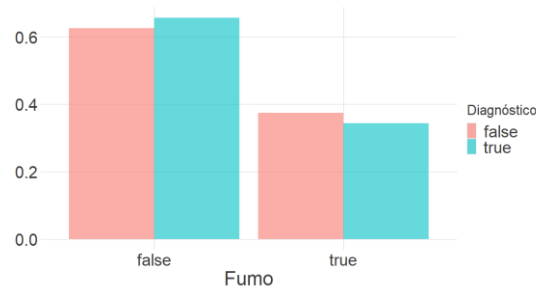
caja o box-plots para la edad según la variable respuesta donde se distingue una diferencia pronunciada entre los que cursan la enfermedad y los individuos sanos. Esto es coherente con el hecho de que la enfermedad tiene mayor incidencia en edades mayores.

Figura 1: Box-plot de edad según el diagnóstico



Por otro lado, se planteó de interés analizar la variable indicadora de si la persona fumó tabaco en algún momento. Al observar las proporciones cuando se la divide según el diagnóstico (Figura 2), se observa que son similares en ambos casos por lo que no se encuentran indicios de que sea útil para la predicción de EP.

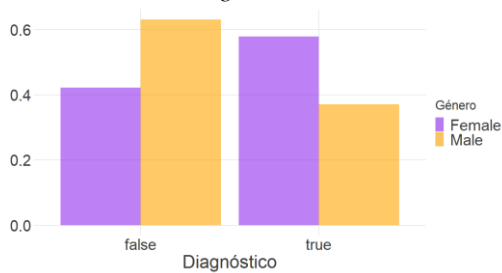
Figura 2: Proporción de fumadores según el diagnóstico



También se cuenta con la información del género de los registros que indica que casi el 80% son hombres y 20% mujeres.

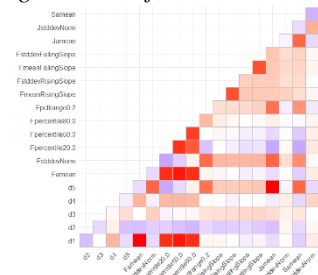
Si se observa la proporción de hombres y mujeres dentro de las personas con EP existe mayor cantidad de mujeres que hombres, y para las personas sin EP se invierte la relación (Figura 3).

Figura 3: Gráfico de barras género según el diagnóstico



Sobre el conjunto de 62 características de la señal de audio registrada para cada caso, y procesadas con OpenSmile, se consideró una preselección de 19 variables atendiendo a la multicolinealidad presente en el conjunto total. En la Figura 4 se muestran las correlaciones de las características a considerar como predictoras en los modelos.

Figura 4: Gráfico de correlaciones



Si bien puede verse que existe alta correlación entre algunas variables, representadas con colores fuertes en el gráfico, puede notarse que entre algunos subgrupos no lo son, por ejemplo, entre d1, d2...d5.

Resultados de predicción

Para el entrenamiento y validación se partitionaron los datos destinando un 70% para el entrenamiento y un 30% para la validación.

Se elije usar la curva ROC y el área bajo esta (AUC) como métricas para la evaluación del poder predictivo de los modelos. La medida de AUC se considera un buen indicador de la capacidad predictiva de un modelo (Hastie, Tibshirani & Friedman, 2009).

Para cada modelo se realizaron varias alternativas, considerando distintos conjuntos de predictores y/o transformaciones en los datos. Aquí solo se

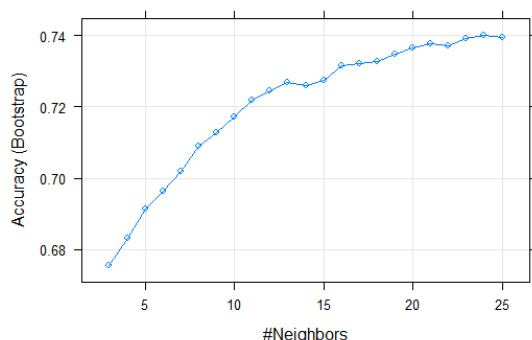
mostrarán las versiones que obtuvieron la mejor métrica de cada modelo.

Comenzando con la regresión logística, el mejor modelo resultó con 5 variables predictoras asociadas a características de la señal de audio (d1,... d5) y además la edad. Una de las ventajas de este modelo reside en la interpretabilidad de los coeficientes, lo que en este caso resulta de interés para la predictora “edad”. Puede mencionarse que el coeficiente asociado a la edad dio como valor 0,11. Esto indica que, según este modelo, la chance de presentar la enfermedad EP aumenta un 12% con año de edad (Hastie, Tibshirani & Friedman, 2009).

Para el modelo Naive Bayes se utilizó un número mayor de variables, en total 22, dentro de las cuales también se encuentra incluida la edad. Si bien este modelo puede considerarse excesivamente sencillo y requiere condiciones de independencia que pueden considerarse como un supuesto demasiado fuerte, da resultados competitivos con otros modelos de clasificación. Y este fue el caso, si observamos la capacidad predictiva obtenida (Tabla 1).

Por último, un modelo KNN se ajustó aplicando normalización de variables numéricas y creación de variables dummies para las categóricas. En este caso, luego de realizar selección de variables, se utilizaron 8 variables predictoras, incluyendo variables demográficas y parámetros de la voz. Se buscó el valor óptimo para el hiperparámetro K mediante la técnica de 10 folds cross-validation en los datos de entrenamiento. Los resultados obtenidos se muestran en la Figura 5, donde el valor elegido para K fue el 24. Este parámetro K corresponde a la cantidad de observaciones “vecinas” a tener en cuenta para realizar la predicción.

Figura 5: Accuracy de KNN según valores de K considerados



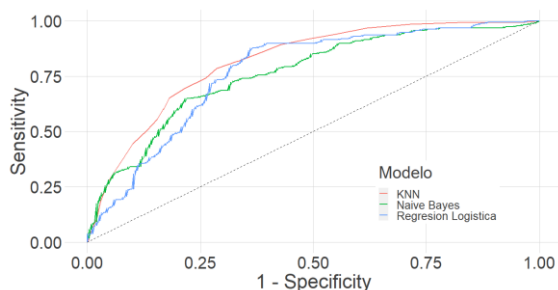
Finalmente, se comparan en Tabla 1 los modelos ajustados según la métrica AUC calculada sobre el conjunto de validación. El modelo que alcanza la mayor capacidad predictiva según la métrica AUC resultó el de KNN, con AUC=0.82. A este siguen la regresión logística y Naive Bayes, con valores muy cercanos.

Tabla 1: AUC para los modelos ajustados

Modelo	AUC
KNN	0.82
Regresión Logística	0.77
Naive Bayes	0.76

En la Figura 6 se puede observar las curvas ROC obtenidas para los tres modelos considerados, donde se aprecia también que la correspondiente a KNN es la que mejora la métrica considerada (área bajo la curva).

Figura 6: Comparación curva ROC de todos los modelos



El objetivo propuesto en este trabajo considera no sólo obtener un modelo de alta capacidad predictiva sino también aquel que logre una mejor precisión al detectar la mayor cantidad de personas con EP. Dicha precisión se define como el porcentaje de

verdaderamente enfermos relativo al total de personas que el modelo clasifica como enfermos con EP.

Para lograr esto, se seleccionó con cada modelo una muestra de 100 personas las cuales se tienen alta probabilidad de ser EP a los cuales se considera el modelo predice como enfermos. Entre estos, se contó el número de personas que efectivamente son enfermos con EP. Este valor corresponde a la precisión relativa a estos 100 casos.

En la tabla 2 se compara la precisión para cada modelo ajustado, esto es, el porcentaje de casos correctamente detectados.

Tabla 2: Porcentaje de casos detectados por modelo

Modelo	Porcentaje de casos con EP detectados correctamente
KNN	81%
Regresión Logística	67%
Naive Bayes	70%

Según esta medida, resulta nuevamente KNN como modelo ganador. Esto último permite responder a la pregunta motivadora de esta investigación:

¿Cómo seleccionar, a partir de datos obtenidos de grabaciones de voces, un conjunto de personas de la población que sean enfermos con EP?

La interpretación de los modelos predictivos estudiados y aplicados aquí, muestra que puede responderse a esta pregunta a partir de diferentes modelos de clasificación, y que la elección del mejor método deberá realizarse no sólo atendiendo a la capacidad predictiva sino también, particularmente para este objetivo, a la precisión.

Conclusión y Trabajos Futuros

En este trabajo se estudiaron distintos modelos que permitan detectar la enfermedad de Parkinson a partir de análisis acústico de voces obtenidas de grabaciones con smartphones. Los resultados muestran que esta tarea puede realizarse considerando las probabilidades predichas por cada modelo para la condición de enfermo con EP, siendo KNN el modelo que mejor

precisión obtuvo para esto. Los resultados obtenidos son comparables con los encontrados por Giuliano et al, 2019, sobre este mismo conjunto de datos, aunque utilizando redes neuronales.

El análisis acústico de la voz da paso a una metodología que permitirá identificar cambios en los parámetros de la voz para predecir el empeoramiento de la enfermedad y orientar la intervención específica.

Este trabajo contribuye a mostrar el potencial de las grabaciones de voz para la detección de la EP e indica que la voz puede servir como un biomarcador que contribuirá a lograr una medicina de precisión al mejorar la velocidad, la accesibilidad y el costo de la gestión de la EP.

Como trabajos a futuro es de interés investigar otras metodologías de análisis de este tipo de datos aplicando *deep learning* o *reinforcement learning*.

Agradecimientos

El presente trabajo se realizó bajo el marco de beca CIN del alumno Tomás S.G y beca UNLaM de la alumna Candela R.F. Ambas becas se vinculan con los proyectos PROINCE-C224 y C239. Universidad Nacional de La Matanza, Argentina.

Referencias

- [1] Langston, J. Parkinson's disease: Current and future challenges, *Neurotoxicology* 23 (4-5) (2002) 443-450.
- [2] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Accurate telemonitoring of Parkinson's disease progression by noninvasive speech test, *IEEE Transactions of Biomedical Engineering* 57 (4) (2010) 884-893.
- [3] Gómez-Vilda, P., Vicente-Torcal, M., Ferrández-Vicente, J., Álvarez Marquina, A., Rodellar-Biarge, V., Nieto-Lluis, V., et al. Parkinson's disease monitoring from phonation biomechanics. In: Ferrández Vicente, J.M., Álvarez Sánchez, J.R., de la

Paz López, F., Toledo-Moreo, F.J., Adeli, H., editors. *Artificial Computation in Biology and Medicine* (2015); vol. 9107 of *Lecture Notes in Computer Science*.

[4] Skodda S, Schlegel U. Speech rate and rhythm in Parkinson's disease. *Mov Disord* 2008; 23: 985-92.

[5] Miller N, Allcock L, Jones D, Noble E, Hildreth AJ, Burn D. Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2007; 78: pp, 1188-1190.

[6] Bot, B., Suver, C., Neto, E. *et al.* The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci Data* 3, 160011 (2016). <https://doi.org/10.1038/sdata.2016.11>

[7] E. I. E. Almaloglou, S. Geronikolou, G. Chroussos and C. Kotropoulos, "Design and validation of a new diagnostic tool for the differentiation of pathological voices in Parkinsonian patients", In Proc. 4th IEEE World Congress Genetics, Geriatrics, and Neurodegenerative Diseases (GeNeDis 2020), October 8-11, 2020.

[8] Tougui I, Jilbab A, Mhamdi JE. Analysis of Smartphone Recordings in Time, Frequency, and Cepstral Domains to Classify Parkinson's Disease. *Healthc Inform Res.* 2020;26(4):274-283. doi:10.4258/hir.2020.26.4.274

[9] Giuliano, M., García-López, A., Pérez, S., Pérez, F. D., Spositto, O., & Bossero, J. (2019). Selection of voice parameters for Parkinson's disease prediction from collected mobile data. In 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA). IEEE.

[10] Hastie, T.; Tibshirani, R. y Friedman, J. (2001). *The Elements of Statistical Learning – Data Mining, Inference, and Prediction*. Springer, Nueva York.

Datos de Contacto:

Tomás Sánchez Grigioni. Universidad Nacional de La Matanza. Florencio Varela 1903, San Justo, Buenos Aires.

E-mail: tomassanchezgrigioni98@gmail.com