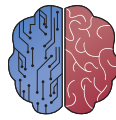




UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería de la Salud



INGENIERÍA
DE LA SALUD

**TFG del Grado en Ingeniería de la
Salud**

**Minería de datos y
aprendizaje automático
aplicado a la predicción de
incidencia de párkinson
basado en la biometereología.**

Presentado por Lorena Calvo Pérez
en Universidad de Burgos

29 de mayo de 2025

Tutores: Antonio Canepa Oneto – Tutor 2



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería de la Salud



D. Tutor 1, profesor del departamento de departamento, área de área.

Expone:

Que el alumno D. Pepe Pérez, con DNI 123456A, ha realizado el Trabajo final de Grado en Ingeniería de la Salud titulado título del trabajo.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 29 de mayo de 2025

Vº. Bº. del Tutor:

Vº. Bº. del Tutor:

D. Tutor 1

D. Tutor 2

Resumen

Como herramienta para el análisis y la creación de modelos predictivos para estimar la prevalencia de la enfermedad de Parkinson a partir de ciertas variables ambientales, se han utilizado técnicas de aprendizaje automático y minería de datos. Este enfoque permite identificar los factores ambientales que más influyen en la aparición de la enfermedad y desarrollar modelos capaces de predecir su prevalencia a nivel mundial. Los resultados contribuyen a una mejor comprensión de la relación entre el entorno y la salud, facilitando el diseño de estrategias preventivas basadas en evidencia.

Descriptores

Palabras separadas por comas que identifiquen el contenido del proyecto Ej: servidor web, buscador de vuelos, android . . .

Abstract

As a tool for the analysis and creation of predictive models to estimate the prevalence of Parkinson's disease based on certain environmental variables, machine learning and data mining techniques have been used. This approach allows for the identification of the environmental factors that most influence the onset of the disease and the development of models capable of predicting its prevalence worldwide. The results contribute to a better understanding of the relationship between the environment and health, facilitating the design of evidence-based preventive strategies.

Keywords

keywords separated by commas.

Índice general

| | |
|--|-----|
| Índice general | iii |
| Índice de figuras | v |
| Índice de tablas | vi |
| Introducción | 3 |
| Objetivos | 5 |
| Conceptos teóricos | 7 |
| 3.1. Enfermedad del Parkinson | 7 |
| 3.2. La Biometeorología y su Relación con la Salud | 8 |
| 3.3. Minería de datos | 10 |
| 3.4. Estado del arte y trabajos relacionados. | 10 |
| 3.5. Estado del arte y trabajos relacionados. | 13 |
| Metodología | 15 |
| 4.1. Descripción de los datos. | 15 |
| 4.2. Procesamiento de datos. | 16 |
| 4.3. Técnicas y herramientas. | 19 |
| Resultados | 21 |
| 5.1. Resumen de resultados. | 21 |
| 5.2. Discusión. | 21 |
| Conclusiones | 23 |
| 6.1. Aspectos relevantes. | 23 |

| | |
|----------------------------------|-----------|
| Lineas de trabajo futuras | 25 |
| Bibliografía | 27 |

Índice de figuras

| | |
|--|----|
| 3.1. Pie de la figura de la figura bla bla bla | 12 |
|--|----|

Índice de tablas

| | |
|----------------|----|
| 3.1. | 12 |
|----------------|----|

Descripción del contenido del trabajo y de la estructura de la memoria y del resto de materiales entregados.

Introducción

El Parkinson es una enfermedad neurodegenerativa que afecta a alrededor de 10 millones de personas en el mundo, siendo esta la segunda enfermedad neurodegenerativa más común tras el Alzheimer. En Europa, su prevalencia oscila entre las 1000 y 2000 casos por cada 100.000 habitantes[[Mayeux et al., 1997]], afectando con más frecuencia a personas mayores de 60 años. La prevalencia de la enfermedad ha aumentado exponencialmente en los últimos años debido al envejecimiento progresivo de la población, por lo que se estima que en 2050 se duplicarán el número de casos.

Aunque la causa exacta de la enfermedad de Parkinson es desconocida, es considerada una patología multifactorial, es decir, influenciada por factores genéticos, ambientales, etc...

La biometeorología es la ciencia que estudia cómo el clima y las condiciones atmosféricas afectan a los seres vivos incluyendo a la salud humana. Según diversos estudios científicos, se ha demostrado que ciertas variables ambientales pueden influir en el desarrollo y la progresión de la enfermedad de Parkinson.

Por todo ello, en este trabajo se ha utilizado la minería de datos y el aprendizaje automático para analizar qué variables ambientales pueden influir en el desarrollo de la enfermedad a través de la elaboración de modelos predictivos.

Objetivos

Objetivo General

Desarrollar un sistema basado en minería de datos y aprendizaje automático para analizar y predecir la prevalencia de la enfermedad de Parkinson a partir de variables biometeorológicas, con el fin de identificar factores ambientales que puedan influir en su desarrollo.

Objetivos Específicos

■ Objetivos funcionales y técnicos:

- Obtener y extraer datos biometeorológicos de una o varias fuentes para su análisis.
- Procesar y preparar los datos en formatos adecuados para el entrenamiento y validación de modelos de aprendizaje automático.
- Desarrollar y entrenar distintos modelos de *machine learning* para la predicción de la prevalencia de Parkinson.
- Implementar un modelo final basado en el promedio ponderado de los modelos individuales para mejorar la precisión predictiva.
- Desarrollar una aplicación interactiva en *Shiny* con *python* para la visualización de resultados.

■ Objetivos de calidad y fiabilidad:

- Evaluar la calidad, precisión y fiabilidad de los modelos mediante análisis estadísticos, incluyendo desviación estándar y detección de anomalías.

- Optimizar la eficiencia y velocidad de ejecución de los modelos para su aplicación práctica.
- **Objetivos de aprendizaje:**
- Adquirir habilidades en minería de datos y aprendizaje automático aplicados a datos biometeorológicos.
 - Aprendizaje de la utilización del *framework Shiny* con *python* para el desarrollo de la aplicación.

Conceptos teóricos

3.1. Enfermedad del Parkinson

La enfermedad de Parkinson fue descrita por primera vez por James Parkinson, un científico británico, en 1817 en su obra *An Eassy on the shaking palasy* (Un ensayo sobre la parálisis temblorosa) [\[\[Parkinson, 2002\]\]](#), en el que explicó las caraterísitcas clínicas de la patología. Aunque el no fue quien dio el nombre a la enfermedad, posteriormente fue reconocida y nombrada en su honor como .enfermedad de Parkisnon".

La enfermedad del Parkinson es un trastorno neurodegenerativo del SNC (sistema nervioso central), el desorden de movimiento más común. La prevalencia de la enfermedad es significativa, afectando a un 1-2% de la población mayor de 65 años. La edad en la suele manifestarse la enfermedad es entre los 65 y 70 años, pero pueden aparecer en mayores de 50 e incluso en los adolescentes. Cabe destacar que hay una mayor incidencia en hombres que en mujeres [\[\[Armstrong and Okun, 2020\]\]](#).

La enfermedad de Parkinson se origina cuando ciertas neuronas en el cerebro dejan de funcionar adecuadamente. Estas células son responsables de producir dopamina, una sustancia que transmite señales a la parte del cerebro encargada de controlar el movimiento y la coordinación del cuerpo. Las neuronas afectadas se encuentran en una región llamada sustancia negra [\[\[González and Pérez, 2021\]\]](#). El Parkinson se desarrolla cuando estas células empiezan a morir o deteriorarse, y esto es debido a alteraciones en su metabolismo. Un aspecto clave de la enfermedad de Parkinson es la acumulación de una proteína llamada α -sinucleína. Esta proteína, en condiciones normales, tiene una función en el cerebro, pero en los pacientes con Parkinson, se pliega de manera anormal, formando agregados que se

acumulan en las neuronas, lo que contribuye al daño celular. Estos depósitos de α -sinucleína forman estructuras conocidas como cuerpos de Lewy, que son característicos de la enfermedad.[Zhang et al., 2018] La producción insuficiente de dopamina desencadena los principales síntomas de la enfermedad, como temblores, lentitud de movimiento, rigidez y problemas de equilibrio.[Poewe et al., 2017]

El diagnóstico de la enfermedad de Parkinson (EP) se basa principalmente en la evaluación clínica. Los criterios más reconocidos fueron establecidos por la UK Parkinson Disease Society - Brain Bank e incluyen cuatro signos clave: bradicinesia o acinesia, temblor en reposo, rigidez e inestabilidad postural.[Marín et al., 2018]

El tratamiento de los síntomas motores de la enfermedad de Parkinson se basa en la terapia de reemplazo de la dopamina o la utilización de agonistas dopaminérgicos. No obstante, la dopamina no puede atravesar la barrera hematoencefálica, por lo que el tratamiento de referencia es la administración de levodopa (L-Dopa), su precursor, que se convierte en dopamina en el cerebro por la acción de la enzima di-hidroxi-fenilalanina descarboxilasa[Hurtado et al., 2016].

En resumen, el tratamiento de la enfermedad combina opciones farmacológicas (principalmente levodopa en monoterapia o en combinación de otros), así como no farmacológicas como el ejercicio físico, fisioterapia, terapia de lenguaje entre otras.

3.2. La Biometeorología y su Relación con la Salud

La Biometeorología es la rama de la ciencia que trata las relaciones entre los procesos atmosféricos y los seres vivos.[Ramos, 2014] Por otro lado, la biometeorología médica, estudia cómo los fenómenos meteorológicos repercuten en el cuerpo humano y cómo los cambios del clima a lo largo de un año provocan variaciones importantes en la salud.[RTVE, 2017] Además, diferentes investigaciones han demostrado que las condiciones meteorológicas y climáticas, como la temperatura, humedad entre otras, puede influir en la aparición de diversas enfermedades por lo que estas variables actúan como factores de riesgo para la salud humana[Rodríguez et al., 2015].

Entre los fenómenos meteorológicos de mayor impacto en la salud están las olas de calor. La frecuencia y la intensidad de estas ha aumentado exponencialmente en los últimos años como consecuencia del cam-

bio climático[[Organización Mundial de la Salud, 2021](#)]. Esto ha provocado un incremento en el número de enfermedad así como de la mortalidad, destacando los periodos de tiempo en los que se prolonga una elevada temperatura[[Fortoul van der Goes, 2022](#)]. De hecho se ha demostrado a través de investigaciones que las olas de calor tienen un impacto directo en la salud pública, especialmente en poblaciones vulnerables como personas mayores. Un ejemplo significativo es el verano de 2003, cuando España experimentó tres olas de calor que causaron un exceso de mortalidad del 8 %, concentrado en mayores de 75 años, con aumentos del 15 % y 29 % en los grupos de edad de 75-84 y mayores de 85 años, respectivamente[[Simón et al., 2005](#)]. A su vez existen otros componentes que influyen en la salud de las personas:

- Calidad del aire: Entre los contaminantes que representan un grave riesgo para la salud pública se encuentran las partículas en suspensión, el monóxido de carbono, el ozono, el dióxido de nitrógeno y el dióxido de azufre. La contaminación del aire, tanto en interiores como en exteriores, causa enfermedades respiratorias y de otro tipo, y es una fuente importante de morbilidad y mortalidad[[World Health Organization, 2025](#)]
- Agua y saneamiento: El acceso a agua limpia y un saneamiento adecuado son esenciales para prevenir enfermedades transmitidas por el agua. La contaminación de fuentes de agua, por desechos industriales y urbanos, puede generar graves problemas de salud.
- Alimentos y seguridad alimentaria: Los contaminantes ambientales como pesticidas y metales pesados afectan la cadena alimentaria y pueden causar enfermedades tanto agudas como crónicas.
- Suelos y contaminación: La contaminación del suelo por productos químicos y desechos peligrosos afecta la salud humana a través del contacto directo o por la ingestión de alimentos contaminados.
- Cambio climático: Las alteraciones en el clima aumentan los riesgos para la salud, como la propagación de enfermedades, el estrés térmico y los desastres naturales, además de causar desplazamientos poblacionales.[[Díaz Cordero, 2012](#)] [[Ambientum, 2025](#)]

La Organización Panamericana de la Salud (OPS) resalta que el cambio climático representa un riesgo significativo para la salud y el bienestar.[([PAHO](#)), 2025]

3.3. Minería de datos

La minería de datos es una disciplina que se centra en el desarrollo de métodos y algoritmos diseñados para extraer automáticamente información relevante, lo que facilita la identificación de patrones ocultos en grandes volúmenes de datos. Además, uno de los objetivos de la minería de datos es garantizar que la información obtenida tenga capacidad predictiva, optimizando así, el proceso de análisis de datos.[Martinez, 2001]

Tanto la minería de datos como el aprendizaje automático se han vuelto fundamentales en el campo de la salud, ya que permiten procesar y analizar grandes cantidades de datos clínicos y biométricos, lo que facilita la detección de patrones, la predicción de enfermedades y el apoyo en la toma de decisiones médicas.[Raul et al., 2016]

Conclusión

En conclusión, la combinación de la minería de datos y el aprendizaje automático en la predicción de la incidencia del Parkinson, teniendo en cuenta la biometeorología, permite identificar relaciones entre los variables ambientales y la salud de los pacientes. Contribuyendo a una mejor comprensión de la interacción del entorno y la salud.

3.4. Estado del arte y trabajos relacionados.

En esta sección se nombrarán algunos proyectos o investigaciones que se han llevado a cabo relacionado con el trabajo, esta revisión bibliográfica puede ser de ayuda a la elaboración de este trabajo.

1. **Predicción de la enfermedad de Parkinson mediante análisis acústico** [Requena Sánchez, 2024] En este trabajo se aborda la predicción de la enfermedad de parkinson a partir de un análisis obtenido de la voz de los pacientes. Para ello, utilizaron grabaciones de voz recogidas de un aplicación móvil, sin supervisión profesional. El objetivo principal era observar si , analizando diferentes aspectos de la voz (variaciones en el tono o la intensidad), se podían encontrar patrones comunes en perosnas con Parkinson. Tras procesar y reducir el numero de variables ,mediante técncias estadísticas, se entrenó un molde SVM (*Support Vector Machine*) para predecir la presencia de la enferemdad solamente a partir de la voz. En cuanto a los resultado obtenidos,el

modelo mostró limitaciones a la hora de identificar de forma correcta a los pacientes enfermos (baja sensibilidad), pero a pesar de ello este tipo de estudios sugiere que este tipo de herramienta es útil como sistema complementario, accesible y no invasivo para poder apoyar el diagnóstico temprano de Parkinson.

2. **Estudio longitudinal del declive cognitivo en pacientes con Parkinson de novo mediante modelos predictivos y modelos de progresión de la enfermedad** [Dick et al., 2007] Este trabajo se centra en estudiar cómo evoluciona el deterioro cognitivo en personas con Parkinson de reciente diagnóstico, utilizando datos recogidos a lo largo del tiempo (estudio longitudinal). Para ello, se emplearon tanto modelos predictivos como modelos de progresión de la enfermedad.

Se utilizaron múltiples fuentes de datos del estudio PPMI, incluyendo imágenes de resonancia magnética (MRI), biomarcadores del líquido cefalorraquídeo (CSF), estudios del transportador de dopamina (DAT), y pruebas clínicas y neuropsicológicas.

En cuanto al análisis, se aplicaron modelos de machine learning con validación cruzada para predecir qué pacientes podían desarrollar deterioro cognitivo leve (MCI), seleccionando las variables más relevantes mediante el método mRMR. Además, se emplearon modelos de efectos mixtos lineales (LME) para analizar cambios estructurales en el cerebro y modelos de progresión tipo GRACE para estimar trayectorias cognitivas individuales. También se usaron modelos de supervivencia de Cox para estudiar los tiempos de conversión a MCI.

Los resultados mostraron que ciertos marcadores, como el grosor cortical y el volumen del hipocampo, junto con algunos datos clínicos, son útiles para anticipar el deterioro cognitivo. En general, el trabajo demuestra que los modelos predictivos pueden ser una herramienta valiosa para mejorar el diagnóstico temprano y personalizar el seguimiento de pacientes con Parkinson.

Revisión bibliográfica de que se está haciendo en la industria o la academia relativo al problema que se está tratando.

Enumeración y resumen de todos los trabajos relacionados de interés.

Sub Subsección

En esta sección y el resto de secciones de la memoria puede ser necesario incluir listas de items.

| a | b | c |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |

Tabla 3.1:

- item1
- item2
- item3
- item4

Listas enumeradas.

1. item1
2. item2
3. item3

Figuras, como la figura 3.1 que aparece en la página 12.

Puedes aprender más de las figuras en la dirección https://es.overleaf.com/learn/latex/Inserting_Images



Figura 3.1: Pie de la figura de la figura bla bla bla

También se pueden insertar tablas como 3.1, que ha sido generada con <https://www.tablesgenerator.com/>.

Es necesario que todas las figuras y tablas aparezca referenciadas en el texto, como estos ejemplos.

Todos los conceptos teóricos deben de estar correctamente referenciados en la bibliografía. Por ejemplo, aquí estoy citando la página de L^AT_EX de Wikipedia [[Wikipedia](#), 2015].

También puede ser necesario utilizar notas al pie ¹, para aclarar algunos conceptos.

3.5. Estado del arte y trabajos relacionados.

Revisión bibliografica de que se está haciendo en la industria o la academia relativo al problema que se está tratando.

Enumeración y resumen de todos los trabajos relacionados de interés.

¹como por ejemplo esta

Metodología

4.1. Descripción de los datos.

Los datos utilizados en este trabajo provienen de la plataforma ***Our World in Data (OWID)***, una fuente de datos abiertos sobre diversas temáticas globales. Para obtener los datos necesarios, primero se exploraron los conjuntos de datos disponibles en su web, seleccionando aquellos que eran más relevantes para el trabajo.

4.1.1. Obtención de datos.

Una vez identificados los datasets de interés, se procedió a localizar los endpoints de la API correspondientes, lo que facilitó la automatización de la obtención de los datos. Estos endpoints permiten acceder a los datos en tiempo real, lo que asegura que la información utilizada se mantenga actualizada.

El proceso consistió en obtener las URLs de los endpoints específicos, los cuales proporcionan los datos en formato JSON, un formato adecuado para su posterior manipulación y análisis. En algunos casos, al acceder a los metadatos, los datos proporcionados inicialmente incluían solo información descriptiva (como años y países), por lo que fue necesario cambiar la extensión del endpoint de "metadata.json" a "data.json" para obtener los valores reales de los indicadores, organizados por país y año. Este proceso permitió integrar los datos automáticamente en el análisis.

4.1.2. Limitaciones de Acceso y Obtención Manual de Datos

. Algunos conjuntos de datos presentan restricciones de redistribución que impiden su acceso directo a través de la API. En estos casos, al intentar acceder a los endpoints correspondientes, se muestra el mensaje: "Los datos de este gráfico no están disponibles para descargar debido a restricciones de licencia". Esta limitación impide la obtención automática de los datos en ciertos casos.

Ante esta restricción, se optó por un enfoque manual utilizando herramientas del navegador, como la función Control + Shift + I, para extraer la información directamente desde la página web de Our World in Data (OWID). A través de este proceso, se pudieron obtener los valores requeridos y generar los datasets necesarios para el análisis.datos.

4.2. Procesamiento de datos.

Una vez obtenidos los datos desde la plataforma Our World in Data (OWID) fue necesario estructurarlos y procesarlos para su análisis.

Sin embargo, la información necesaria no se encontraba en un único archivo JSON, sino que estaba distribuida en dos fuentes de datos distintas dentro de la misma plataforma:

- Fuente de datos principal: Contiene los valores de los indicadores analizados, organizados en una estructura donde los datos están asociados a identificadores numéricos de países y años.
- Fuente de metadata: Proporciona información adicional necesaria para interpretar correctamente los datos, como la correspondencia entre identificadores numéricos y nombres de países, así como la lista de años disponibles en el conjunto de datos.

Dado que cada fuente presenta una estructura diferente, se tuvo que diseñar un proceso de integración para combinar ambas y construir un conjunto de datos estructurado en forma de DataFrame, donde cada fila representa un país en un año determinado con su valor correspondiente del indicador analizado.

Para llevar a cabo esta integración, primero se cargaron los datos desde las API correspondientes, asegurando que la información estuviera correctamente obtenida antes de su procesamiento.

4.2.1. Carga de Datos

Para obtener los datos se implementó una función que se encarga de realizar peticiones a las direcciones URL de los datos y la metadata, obteniendo así la información en formato JSON. Esto permitió estructurar los datos de manera uniforme para su posterior manipulación y análisis.

4.2.2. Estructuración de Datos

El procesamiento de los datos es un paso esencial para estructurar la información de manera homogénea. Primero, se extrajeron los valores numéricos de los indicadores desde los datos obtenidos. Luego, se identificaron los años disponibles en el archivo metadata, ya que estos definen la dimensión temporal del análisis. Finalmente, se recuperó la lista de países o entidades presentes en el dataset, lo que permitió organizar la información de manera estructurada.

Si alguno de estos elementos no está presente en los datos obtenidos, se han establecido condiciones para detectar y gestionar estos errores, garantizando la coherencia de la información utilizada.

4.2.3. Conversión a DataFrame

Para facilitar el análisis, los datos procesados fueron organizados en un formato tabular utilizando un DataFrame. Se estructuraron en filas donde cada una representa una combinación de país y año, asignándole el valor correspondiente del indicador. En los casos en los que no existió un valor disponible para una combinación específica, se dejó un espacio vacío para mantener la coherencia del conjunto de datos.

Gracias a este procesamiento, se logró una organización clara y accesible de los datos, lo que permitió su integración en el análisis posterior de manera eficiente.

4.2.4. Gestión y adaptación de los datos para los modelos

Una vez realizado este procesamiento, se obtuvieron seis dataframes, uno de ellos correspondiente a la variable dependientes y el resto a las diferentes variables independientes. Como bien se ha mencionado anteriormente, no todas las variables presentan el mismo número de filas, debido a que la disponibilidad de los años y los países no era la misma para los diferentes

conjuntos de datos. Debido a esto se procedió a la unión de los distintos dataframes, tomando como referencia la variable objetivo, es decir, la prevalencia de parkinson. De esta manera, el resto de variables se alinean solamente con los países y los años disponibles en el conjunto de datos de la variable dependiente. En los casos en los que para una variables no hubiera datos en un año y país determinado, se dejaron esos valores como nulos.

Antes de realizar una imputación de datos, se calculó el porcentaje de valores nulos que presentaba el dataframe unificado, cuyo resultado fue de un 5.4 %. Estos valores ausentes se encontraban distribuidos en diferentes columnas del dataframe por lo que al realizar una eliminación de lo nulos pasamos de tener 7264 filas a 5404, conservando un 75 % de los datos aproximadamente. Por todo ello, no se llevó a cabo la imputación de datos ya que se mantenían mas del 50 % de los datos.

A partir del dataframe final, se obtuvieron dos nuevos conjuntos de datos, uno para el entrenmientto de modleos y otro para la predicción. El dataframe de entrenamiento contenía tanto las varibales independientes como la dependiente, así como los países, Este conjunto abarcaba todos los años disponibles a excepción del último, ya que, es el año que se reservó para la predicción. En cuanto al dataframe de predicción contenía unicamente las variables independientes y los países, utilizando solamente los datos correpondientes al último año disponible en el dataframe final.

4.2.5. Modelos predictivos

Antes de la etapa de modelado, se realizó un análisis exploratorio preliminar con el objetivo de evaluar la multicolinealidad entre las variables independientes y en estudiar el tipo de relación que cada una mantiene con la variable objetivo. Estos resultados y análisis preliminares se encuentran detallados en el Anexo G. Para la predicción de la prevalencia de Parkinson en el año 2021, se utilizaron seis modelos de aprendizaje supervisado con el objetivo de abordar el trabajo desde diferentes metodologías y perspectivas. Los modelos seleccionados para la predicción fueron Modelo GLM (binomial negativo), modelos basado en arboles (Random Forest Regressor, XGBoosting Regressor), SVR Regressor, KNN Regressor y MLP Regressor. El motivo de la selección de estos, junto a su ajuste y evaluación se encuentran en el Anexo G.

Una vez seleccionados, ajustados los modelos, se procede a realizar el entrenamiento para cada uno de los países disponibles. A continuación, se realiza la predicción de la prevalencia de parkinson para cada país y una

posterior representación visual en forma de mapa para su interpretación e integración en la aplicación.

A partir de los resultados de la predicción de los seis modelos, se construyó un nuevo conjunto de datos que contenía el promedio de las predicciones individuales. Esta medida fue la utilizada como predicción final de la prevalencia de Parkinson, con el objetivo de poder combinar las características de cada uno de los modelos y obtener un resultado lo mas representativo posible.

4.3. Técnicas y herramientas.

Esta parte de la memoria tiene como objetivo presentar las técnicas metodológicas y las herramientas de desarrollo que se han utilizado para llevar a cabo el proyecto. Si se han estudiado diferentes alternativas de metodologías, herramientas, bibliotecas se puede hacer un resumen de los aspectos más destacados de cada alternativa, incluyendo comparativas entre las distintas opciones y una justificación de las elecciones realizadas. No se pretende que este apartado se convierta en un capítulo de un libro dedicado a cada una de las alternativas, sino comentar los aspectos más destacados de cada opción, con un repaso somero a los fundamentos esenciales y referencias bibliográficas para que el lector pueda ampliar su conocimiento sobre el tema.

Resultados

5.1. Resumen de resultados.

Breve resumen de los resultados. En caso de ser un trabajo muy experimental, los resultados completos pueden aparecer en su anexo correspondiente.

Debería haber una correspondencia entre los objetivos y los resultados explicados en esta sección

5.2. Discusión.

Discusión y análisis de los resultados obtenidos.

Conclusiones

Todo proyecto debe incluir las conclusiones que se derivan de su desarrollo. Éstas pueden ser de diferente índole, dependiendo de la tipología del proyecto, pero normalmente van a estar presentes un conjunto de conclusiones relacionadas con los resultados del proyecto y un conjunto de conclusiones técnicas.

6.1. Aspectos relevantes.

Este apartado pretende recoger los aspectos más interesantes del **desarrollo del proyecto**, comentados por los autores del mismo.

Debe incluir los detalles más relevantes en cada fase del desarrollo, justificando los caminos tomados, especialmente aquellos que no sean triviales.

Puede ser el lugar más adecuado para documentar los aspectos más interesantes del proyecto y también los resultados negativos obtenidos por soluciones previas a la solución entregada.

Este apartado, debe convertirse en el resumen de la experiencia práctica del proyecto, y por sí mismo justifica que la memoria se convierta en un documento útil, fuente de referencia para los autores, los tutores y futuros alumnos.

Lineas de trabajo futuras

Este capítulo debería ser informe crítico indicando cómo se puede mejorar el proyecto, o cómo se puede continuar trabajando en la línea del proyecto realizado.

Bibliografía

- [Ambientum, 2025] Ambientum (2025). Enfermedades emergentes: El rol crucial de la salud ambiental. https://www.ambientum.com/ambientum/cambio-climatico/enfermedades-emergentes-el-rol-crucial-de-la-salud-ambiental.asp?utm_source=chatgpt.com. Accedido: 2025-03-26.
- [Armstrong and Okun, 2020] Armstrong, M. J. and Okun, M. S. (2020). Diagnosis and treatment of parkinson disease: A review. *JAMA*, 323(6):548–560.
- [Díaz Cordero, 2012] Díaz Cordero, G. (2012). El cambio climático. *Ciencia y sociedad*.
- [Dick et al., 2007] Dick, F. D., De Palma, G., Ahmadi, A., Scott, N. W., Prescott, G. J., Bennett, J., Semple, S., Dick, S., Counsell, C., Mozzoni, P., Haites, N., Wettinger, S. B., Mutti, A., Otelea, M., Seaton, A., Söderkvist, P., Felice, A., and study group, G. (2007). Environmental risk factors for parkinson’s disease and parkinsonism: the geoparkinson study. *Occupational and environmental medicine*, 64(10):666–672.
- [Fortoul van der Goes, 2022] Fortoul van der Goes, T. I. (2022). Cambio climático, la onda de calor y sus efectos en la salud. *Revista de la Facultad de Medicina (México)*, 65(5):3–5.
- [González and Pérez, 2021] González, J. C. and Pérez, M. L. (2021). James parkinson y la parálisis agitante: 200 años después. *Revista de Neurología*, 72(12):501–506.

- [Hurtado et al., 2016] Hurtado, F., N Cárdenas, M. A., Cardenas, F., and León, L. A. (2016). La enfermedad de parkinson: Etiología, tratamientos y factores preventivos. *Universitas Psychologica*, 15(SPE5):1–26.
- [Marín et al., 2018] Marín, D. S., Carmona, H., Ibarra, M., and Gámez, M. (2018). Enfermedad de parkinson: fisiopatología, diagnóstico y tratamiento. *Revista de La Universidad Industrial de Santander. Salud*, 50(1):79–92.
- [Martinez, 2001] Martinez, B. B. (2001). Minería de datos. *Cómo hallar una aguja en un pajar. Ingenierías*, 14(53):53–66.
- [Mayeux et al., 1997] Mayeux, R., Marder, K., Cote, L. J., and et al. (1997). Prevalence of parkinson’s disease in europe: the europarkinson collaborative study. *Neurology*, 48(1):9–13.
- [Organización Mundial de la Salud, 2021] Organización Mundial de la Salud (2021). Cambio climático y salud. <https://www.who.int/es/news-room/fact-sheets/detail/climate-change-and-health>. Hoja informativa. Consultado el 28 de mayo de 2025.
- [(PAHO), 2025] (PAHO), P. A. H. O. (2025). Cambio climático y salud. https://www.paho.org/es/temas/cambio-climatico-salud?utm_source=chatgpt.com. Accedido: 2025-03-26.
- [Parkinson, 2002] Parkinson, J. (2002). An essay on the shaking palsy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(2):223–236. PMID: 11983801.
- [Poewe et al., 2017] Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkman, J., Schrag, A.-E., and Lang, A. E. (2017). Parkinson disease. *Nature reviews Disease primers*, 3(1):1–21.
- [Ramos, 2014] Ramos, M. B. (2014). Biometeorología humana en la ciudad de punta alta.
- [Raul et al., 2016] Raul, A., Patil, A., Raheja, P., and Sawant, R. (2016). Knowledge discovery, analysis and prediction in healthcare using data mining and analytics. In *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, pages 475–478. IEEE.
- [Requena Sánchez, 2024] Requena Sánchez, C. (2024). *Predicción de la enfermedad de Parkinson mediante análisis acústico*. PhD thesis, UPC, Facultat de Matemàtiques i Estadística.

- [Rodríguez et al., 2015] Rodríguez, Y., García, M. d. l. A., Martínez, M. d. C., and Rodríguez, M. d. C. (2015). Variabilidad y cambio climáticos: su repercusión en la salud. *Revista Cubana de Salud Pública*, 41(7):1–12.
- [RTVE, 2017] RTVE (2017). ¿cómo afecta el clima a nuestra salud?
- [Simón et al., 2005] Simón, F., López-Abente, G., Ballester, E., and Martínez, F. (2005). Mortality in spain during the heat waves of summer 2003. *Euro Surveillance*, 10(7):156–161.
- [Wikipedia, 2015] Wikipedia (2015). Latex — wikipedia, la enciclopedia libre. [Internet; descargado 30-septiembre-2015].
- [World Health Organization, 2025] World Health Organization (2025). Air pollution. https://www.who.int/health-topics/air-pollution#tab=tab_1. Accedido: 2025-03-26.
- [Zhang et al., 2018] Zhang, G., Xia, Y., Wan, F., Ma, K., Guo, X., Kou, L., Yin, S., Han, C., Liu, L., Huang, J., et al. (2018). New perspectives on roles of alpha-synuclein in parkinson’s disease. *Frontiers in aging neuroscience*, 10:370.