



TRABAJO DE INVESTIGACIÓN

Salud: Detección de anomalías

PREDICCIÓN EPIDEMIAS: CASO COVID-19

[Máster en Data Science y Analítica Avanzada.](#)

Verónica Sánchez
Víctor Aranda
22/12/2023

ÍNDICE

1. Introducción	3
Contextualización del estudio sobre la predicción de pandemias	
2. Contexto Metabiota	4
3. Fundamentos de la Predicción de Pandemias	5
3.1 Vigilancia de Enfermedades en Animales	5
3.2 Secuenciación Genética	5
3.3 Investigación Epidemiológica	5
3.4 Salud de la Vida Silvestre	5
3.5 Salud Pública Global	5
3.6 Investigación Científica Continua	5
3.7 Desarrollo de Vacunas y Tratamientos	5
3.8 Educación y Concienciación Pública	5
4. Análisis de Datos y Modelos Predictivos	6
4.1 Datos Históricos y Tendencias	6
4.2 Metodología	6
4.3 Tendencias Temporales	7 - 8
4.4 Distribución Geográfica	8 - 9
4.5 Simulación de Eventos Epidémicos	9 - 11
4.6 Cambio en el Riesgo a lo Largo del Tiempo	11
5. Conclusiones y Recomendaciones	12
5.1 Recapitulación de puntos clave	12
5.2 Sugerencias para futuras investigaciones y acciones para mejorar la capacidad predictiva y de respuesta ante pandemias.	12 - 13
6. Biografía	14
6.1 Listado de fuentes utilizadas y referencias bibliográficas.	14
7. Recolección de Datos y Recursos Relevantes	15
7.1 Recursos de Datos	15
7.2 Enlaces y fuentes de interés	15

Introducción

Elegimos abordar este tema debido al impacto significativo que la pandemia ha tenido en nuestras vidas, lo cual lo convierte en un tema crucial para el futuro. Existe una amplia gama de fuentes de información, principalmente gubernamentales, que podrían ser fundamentales si deseamos realizar predicciones sobre futuras epidemias. Consideramos que la investigación en este campo es fascinante, ya que existen numerosos factores que influyen en la aparición y propagación de las epidemias, aspectos que exploraremos en nuestro trabajo de investigación.

En este trabajo, trazaremos un recorrido comenzando por los fundamentos en la predicción de pandemias. Luego nos adentraremos en una investigación detallada sobre la empresa **Metabiota**, especializada en la predicción de estos eventos.

Finalmente, concluiremos con nuestras observaciones y dejaremos abierta la posibilidad de realizar un estudio de predicción basado en datos reales específicos de España.

La anticipación de pandemias, especialmente aquellas derivadas de coronavirus zoonóticos, ha adquirido una importancia crucial en la salud pública a nivel mundial. Nuestra investigación se sumerge en el modelo predictivo y estrategias utilizadas para prever futuras amenazas pandémicas, teniendo en cuenta el enfoque innovador implementado por la empresa Metabiota en este campo.

La elección de Metabiota para este estudio se basa, en parte, en su papel pionero durante la pandemia al difundir información casi en tiempo real sobre la evolución del COVID-19.

2. Contexto Metabiota

Metabiota es una startup estadounidense fundada en 2008 que compila datos de todo el mundo para predecir epidemias.

La startup tiene estrechos enlaces con el gobierno de Estados Unidos siendo contratada en ocasiones por el Departamento de Defensa, en los primeros meses de las noticias del COVID se unificó a una empresa de Inteligencia Artificial Canadiense llamada BlueDot por tal de aunar sus conocimientos y accesos a información con expertos en Inteligencia Artificial.

Tanto el gobierno de Estados Unidos como Google tienen inversiones de miles de dólares en dicha empresa y suministran información por tal de que puedan ejercer su labor de predicción de pandemias.

La empresa Metabiota emplea un enfoque de modelado complejo que se basa en la integración de múltiples fuentes de datos, análisis epidemiológicos, y herramientas de inteligencia artificial y aprendizaje automático para predecir futuras epidemias.

Su modelo de predicción utiliza:

Recopilación de Datos Globales: Integra datos de múltiples fuentes, incluyendo información epidemiológica, datos genómicos, movilidad humana, entre otros, para crear una base sólida y comprensiva.

Análisis Epidemiológico Avanzado: Realiza un análisis profundo de patrones epidemiológicos pasados para identificar tendencias, factores desencadenantes y correlaciones entre variables que podrían indicar la aparición de nuevas enfermedades.

Inteligencia Artificial y Aprendizaje Automático: Emplea algoritmos avanzados de inteligencia artificial y aprendizaje automático para procesar grandes cantidades de datos, identificar patrones ocultos y realizar predicciones sobre posibles brotes y su potencial gravedad.

Modelado de Escenarios: A partir de la información recopilada y el análisis de datos, Metabiota crea modelos predictivos que simulan diferentes escenarios epidémicos. Estos modelos ayudan a evaluar la probabilidad de que ciertos eventos pandémicos ocurran, así como su posible impacto en términos de propagación, gravedad y afectación.

A continuación vamos a detallar una investigación realizada por **Metabiota** que nos servirá para entender qué ocurrió en las diferentes epidemias y como intentaron predecir con la información que les fue suministrada una futura pandemia.

3. Fundamentos de la Predicción de Pandemias

3.1 Vigilancia de enfermedades en animales.

El seguimiento detallado de enfermedades en poblaciones animales clave, como murciélagos y ciertos mamíferos, se muestra como un pilar esencial para la predicción de pandemias zoonóticas.

3.2 Secuenciación genética.

El análisis exhaustivo de la secuencia genética de virus en animales y humanos permite identificar nuevos coronavirus y evaluar su potencial de infección y propagación.

3.3 Investigación epidemiológica.

El estudio de patrones de enfermedades en poblaciones humanas y animales en áreas de transmisión previa brinda valiosas perspectivas sobre posibles brotes futuros.

3.4 Salud de la vida silvestre.

El monitoreo de la salud de la vida silvestre, especialmente en zonas de interacción humana intensa, es fundamental para detectar puntos de entrada de virus a los humanos.

3.5 Salud pública global.

La colaboración internacional y la notificación rápida de brotes son esenciales, destacando el papel crucial de organizaciones como la OMS en la detección y respuesta a pandemias.

3.6 Investigación científica continúa.

La necesidad de expandir la investigación sobre coronavirus y enfermedades zoonóticas para comprender mejor su biología y sus patrones de transmisión.

3.7 Desarrollo de vacunas y tratamientos.

La preparación para pandemias incluye el desarrollo y despliegue rápido de vacunas y tratamientos antivirales efectivos.

3.8 Educación y concienciación pública.

La importancia de la educación pública sobre medidas preventivas para contener la propagación de virus pandémicos.

4. Análisis de Datos y Modelos Predictivos

En general hay la creencia que un tipo de pandemia como el COVID es algo que ocurre una vez en la vida o cada 100 años, pero los datos informan sobre un crecimiento en las circunstancias de posibilidad de contagios entre animales y humanos, patógenos desconocidos, etc.

Modelos lógicos muestran que eventos como el COVID o de similar magnitud son probables que ocurran en el futuro.

La idea es construir modelos de detección que nos permitan ser ágiles a la hora de identificar una tendencia, y poner medidas de contención de forma más eficiente y con tiempo suficiente.

4.1 Datos Históricos y Tendencias

El análisis detallado de datos históricos del Ébola, Marburgo, Nipah, Machupo y por supuesto el COVID reveló patrones epidemiológicos significativos.

Se identificaron tendencias en la propagación de enfermedades zoonóticas, permitiendo la identificación de eventos críticos y la evaluación de la frecuencia y gravedad de las pandemias.

Estos datos históricos proporcionaron una base crucial para proyectar posibles escenarios futuros, considerando la evolución de las enfermedades a lo largo del tiempo.

4.2 Metodología

La metodología empleada en esta investigación se basó en una recopilación exhaustiva de datos de múltiples fuentes confiables, incluyendo bases de datos epidemiológicas, informes de salud pública global y estudios científicos.

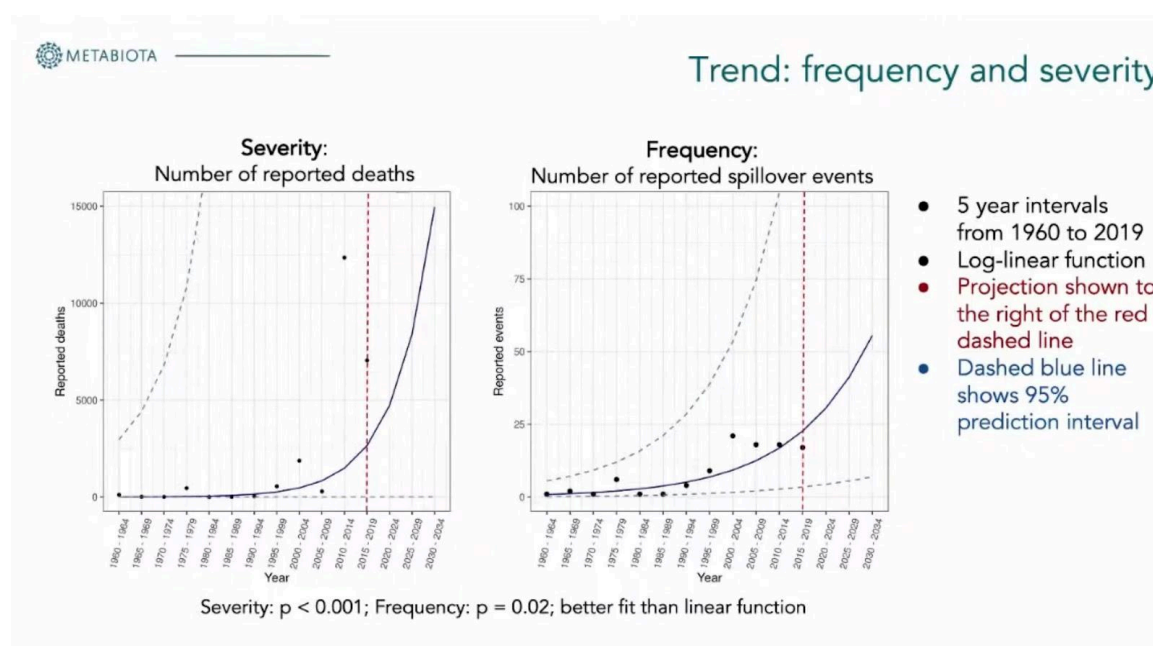
Se utilizaron herramientas estadísticas y modelos predictivos para analizar y correlacionar estos datos, identificando relaciones entre variables clave y estableciendo patrones epidemiológicos relevantes.

- Examinación de bases de datos de +2600 epidemias.
- Enfermedades que se han transmitido de animales hacia humanos
- Si el patógeno ha causado más de 100 muertes.

4.3 Tendencias Temporales

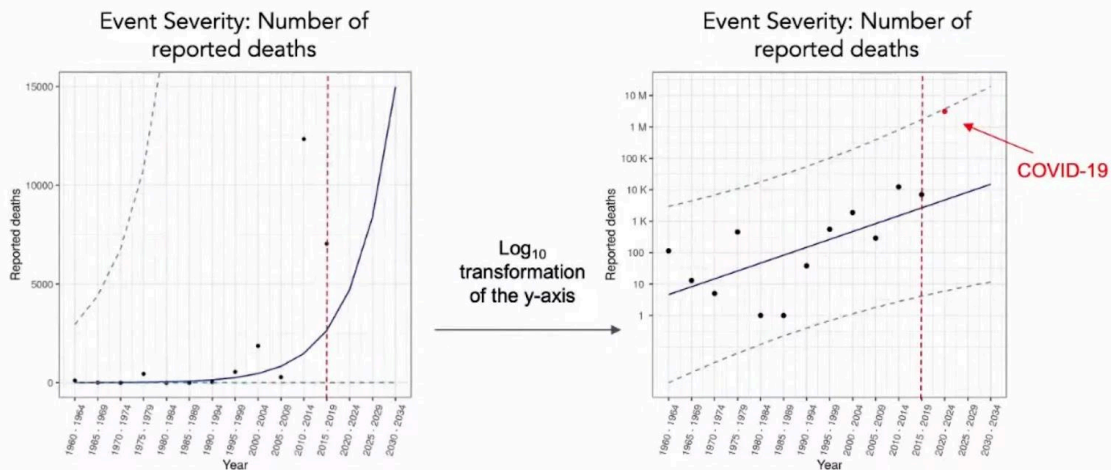
El análisis de tendencias temporales permitió comprender la evolución de enfermedades zoonóticas a lo largo del tiempo.

Se identificaron períodos de aumento significativo en la incidencia de enfermedades, proporcionando información valiosa sobre la frecuencia y la aparición de pandemias. Esta evaluación temporal se centró en identificar patrones estacionales, ciclos de brotes y cambios en la transmisibilidad de patógenos entre especies.



En los gráficos previos, se evidencia un incremento significativo de múltiples epidemias, excluyendo el impacto del COVID. Por un lado, se observa un aumento progresivo en términos de la mortalidad generada por diversas epidemias. Por otro lado, en el gráfico de frecuencia, se aprecia un aumento en el número de pandemias registradas a lo largo del tiempo.

Trend: frequency and severity



En los gráficos anteriores se puede observar una tendencia al alza de que existan diferentes virus que se vayan transmitiendo entre animales y humanos, el COVID en particular es un evento extremo que se sale de la predicción hasta ese momento.

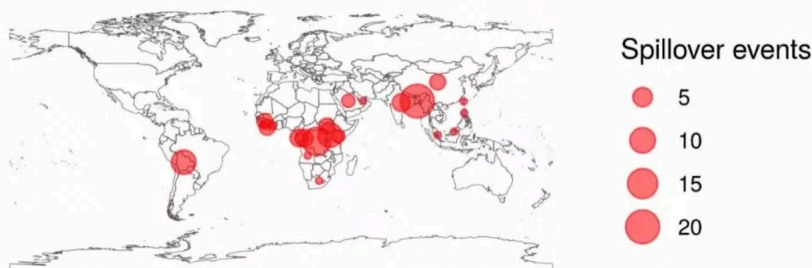
4.4 Distribución Geográfica

El examen exhaustivo de la distribución geográfica de las pandemias reveló patrones regionales en la propagación de enfermedades zoonóticas.

Se detectaron áreas de mayor incidencia y se analizó la relación entre factores geográficos, socioeconómicos y ambientales con la aparición y propagación de pandemias.

Esta comprensión geoespacial fue fundamental para identificar puntos críticos y para la implementación de estrategias preventivas más efectivas.

Geographic distribution of spillover events

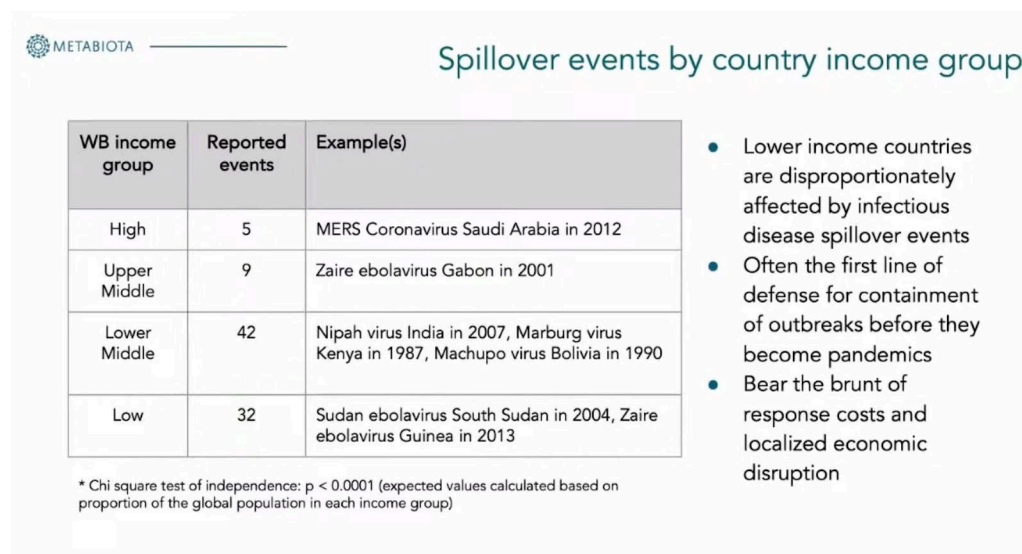


Notes:

- World map displays the distribution of 88 reported spillover events (points) since 1960 for selected pathogens (epidemic coronaviruses, filoviruses, Nipah, Machupo).
- the size of the point is the proportional to the number of spillover events.

Vemos que en los países de renta más baja es donde en su mayoría se originan estos eventos y por lo tanto se propagan con mayor intensidad entre su población.

La siguiente tabla es una representación numérica del mapa arriba mostrado, hay una repartición desproporcionada de donde una epidemia surge y se empieza a transmitir sin un control adecuado, en países en precariedad o con muy poco acceso a un buen sistema sanitario se detectan muchos más eventos de alta transmisión que en países de renta alta y con un sistema sanitario adecuado.



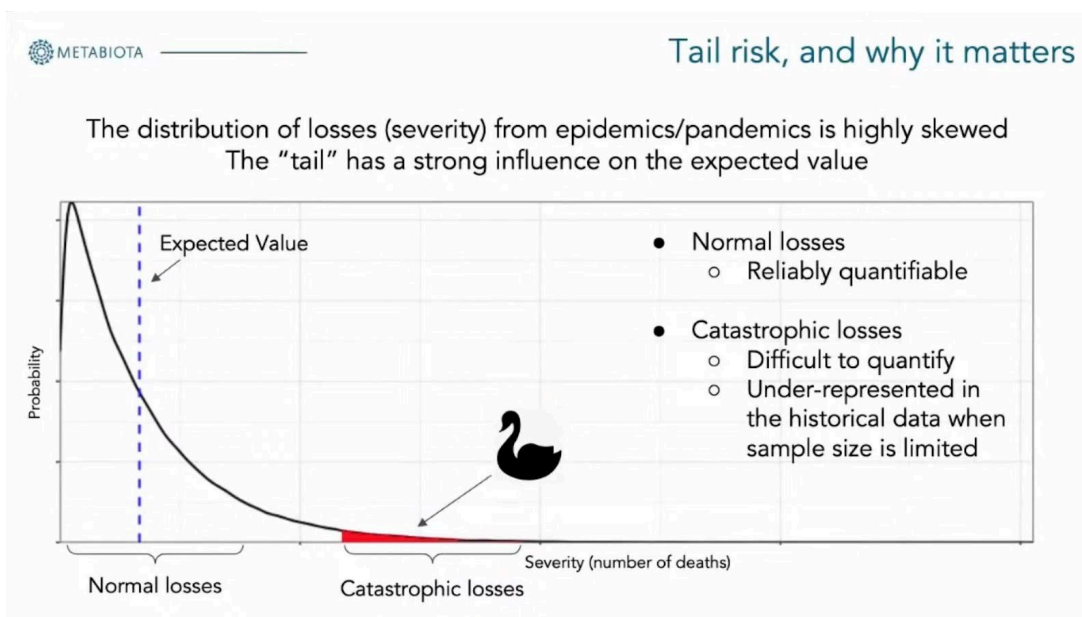
4.5 Simulación de Eventos Epidémicos

Los modelos predictivos y las simulaciones de eventos epidémicos proporcionaron una evaluación cuantitativa de la probabilidad y gravedad de futuras pandemias.

Se utilizaron técnicas de modelado avanzado para proyectar escenarios potenciales, considerando diferentes variables como la transmisibilidad del patógeno, la efectividad de las medidas de control y las condiciones ambientales.

Estos modelos permitieron evaluar estrategias de mitigación y preparación ante posibles pandemias.

Las pandemias que causan pérdidas catastróficas son difíciles de prever ya que hay muy poca información histórica sobre ellas.



En la gráfica anterior podemos observar como eventos que pueden llegar a causar una mortalidad inusual (eventos llamados cisnes negros debida a su alta severidad), como pasó con el COVID, no son eventos habituales y son difíciles de prever pero no porque no vayan a existir en un futuro si no porque hay muy poca información sobre los mismos.

**Impact of spillover reduction -
Estimates based on simulated event catalogs**

Pathogen group	Expected Deaths over next decade	Deaths averted over 10 years from 1% reduction in spillover events
Pandemic Flu (excluding seasonal flu)	2,200,000	22,000
Coronaviruses	1,500,000	15,000
Viral Hemorrhagic Fevers (e.g. ebola virus, nipah, etc.)	300,000	3,000
Total	4,000,000	40,000

Expected deaths over 10 years due to epidemics and the number of deaths that could be averted by a 1% reduction in spillover events (i.e. 1% of spillover events are prevented)

Source: Oppenheim B, Stephenson N., Bertozzi S., Gadsden P., Guerrero J., Madhav N, Jamison DT, The Value of Viral Surveillance. (2021) Working paper.

El modelo predictivo creado por Metabiota basado en las diferentes epidemias del pasado arrojó los resultados de la tabla anterior se puede observar como en los próximos 10 años pueden existir diferentes pandemias que pueden causar cientos de miles de muertes siendo la gripe pandémica la que mas predomina entre todas las posibilidades de epidemias. La última columna indica como con una correcta mitigación y si los distintos gobiernos son capaces de reaccionar a tiempo frente a estos eventos se puede reducir la mortalidad.

Hay infinidad de posibles escenarios.

Estos se ven alterados por muchos factores, ambientales, decisiones políticas, patrones de conducta.

Teniendo en cuenta distintas simulaciones se estima un % que vuelva a ocurrir un evento como el COVID19.

2,5 - 3,3% de que haya un evento que cause 3.5 millones de muertes en un año.

22 - 28% de que ocurra en los próximos 10 años.

47 - 57% de que ocurra en los próximos 25 años.

4.6 Cambio en el Riesgo a lo Largo del Tiempo

El análisis del cambio en el riesgo de pandemias a lo largo del tiempo consideró múltiples factores, incluyendo aspectos ambientales, políticos y de comportamiento humano.

Se examinaron cambios significativos en la dinámica de transmisión de enfermedades zoonóticas, identificando las influencias fluctuantes que podrían aumentar o disminuir el riesgo de brotes pandémicos.

Esta evaluación fue crucial para adaptar estrategias de respuesta y mitigación a medida que evolucionan las condiciones epidemiológicas y sociales.

La última década se han visto una acumulación de enfermedades infecciosas:

- Múltiples epidemias de ébola en el oeste y centro de África
- Transmisión del MERS-CoV en Surcorea.
- Pandemias como el Zika y el COVID.

No solo es importante crear un modelo predictivo y entenderlo, si no la vigilancia, el monitoreo de datos en tiempo lo más real posible para detectar el inicio de una posible epidemia o transmisión y ser capaces de reaccionar a tiempo.

Para ello es importante incluir todas las fuentes de información posibles a la hora de monitorear en tiempo real, esto incluye redes sociales, no solo información gubernamental y ser capaces de aislar el ruido.

5. Conclusiones y recomendaciones

5.1 Recapitulación de Puntos Clave

Durante la investigación y análisis exhaustivo de datos sobre la predicción de pandemias, se identificaron varios puntos fundamentales:

Vigilancia Integral de Enfermedades: La vigilancia activa de enfermedades en poblaciones animales, la secuenciación genética y la investigación epidemiológica son componentes cruciales para predecir y prevenir la propagación de pandemias zoonóticas.

Interacción entre Animales y Humanos: El análisis reveló una tendencia al alza en la transmisión de virus entre animales y humanos, con eventos como el COVID-19 evidenciando la urgencia de monitorear estrechamente esta interacción.

Necesidad de Modelos Predictivos Mejorados: Los modelos predictivos actuales han demostrado ser útiles, pero aún enfrentan desafíos para prevenir pandemias catastróficas debido a la falta de datos históricos y a la complejidad de factores variables.

Distribución Geográfica y Factores Socioeconómicos: Los países de renta más baja experimentan una mayor incidencia de eventos pandémicos, resaltando la importancia de abordar las disparidades socioeconómicas en la respuesta a las pandemias.

5.2 Sugerencias para Futuras Investigaciones

Basándonos en las conclusiones anteriores, se proponen las siguientes recomendaciones para mejorar la capacidad predictiva y de respuesta ante pandemias:

Mayor Colaboración Internacional: Fomentar una colaboración más estrecha entre países y organizaciones para compartir datos de manera más amplia y rápida, mejorando así la capacidad de predicción y respuesta.

Mejora de Modelos Predictivos: Investigar y desarrollar modelos más sofisticados que integren datos multidisciplinarios, considerando factores ambientales, comportamentales y socioeconómicos para mejorar la precisión de las predicciones.

Énfasis en la Prevención: Priorizar la inversión en programas de prevención y control de enfermedades, incluyendo el desarrollo de vacunas más rápidas y eficaces ante patógenos emergentes.

Enfoque en la Educación y la Concienciación: Impulsar campañas de educación pública para aumentar la conciencia sobre medidas preventivas, fomentando la notificación temprana de enfermedades inusuales.

Monitoreo Continuo y Mejora de Recursos de Datos: Mejorar la infraestructura para el monitoreo continuo, incluyendo el uso de redes sociales y datos en tiempo real, así como ampliar y enriquecer los recursos de datos disponibles.

Estas recomendaciones buscan fortalecer la capacidad de anticipación y respuesta ante futuras pandemias, abordando tanto las lagunas identificadas en la investigación como las áreas clave para el avance en la predicción y prevención de pandemias.

6. Bibliografía

6.1 Listado de fuentes utilizadas y referencias bibliográficas.

Datasets Datos Metabiota COVID:

https://gh.bmj.com/content/bmjgh/suppl/2021/05/06/bmjgh-2021-005542.DC1/bmjgh-2021-005542supp001_data_supplement.pdf

Datos Metabiota vendidos a Ginkgo:

<https://www.epidemictracker.com/>

CGDEV: What's Next? Predicting Frequency and Scale Future Pandemics:

<https://www.cgdev.org/event/whats-next-predicting-frequency-and-scale-future-pandemics>

National Geographic: ¿Puede una nueva aplicación predecir la próxima pandemia?:

<https://www.nationalgeographic.es/ciencia/2021/04/puede-una-nueva-aplicacion-predecir-la-proxima-pandemia>

Organización Mundial de la Salud: Preparación para las Emergencias - COVID-19:

<https://www.woah.org/es/que-ofrecemos/preparacion-para-las-emergencias/covid-19/>

Harvard: Preventing and Predicting Infectious Diseases:

<https://ccdd.hsph.harvard.edu/research/preventing-and-predicting-infectious-diseases/>

ScienceDirect: Amenazas, Desafíos y Preparación para Futuras Pandemias:

<https://www.sciencedirect.com/science/article/pii/S1567134822000144>

Organización Mundial de la Salud: Imaginando el Futuro de las Pandemias:

<https://iris.who.int/bitstream/handle/10665/361252/9789240052093-eng.pdf?sequence=1>

Royal Society Publishing: Características de los 100 Brotos de Enfermedades Zoonóticas Modernas más Grandes:

<https://royalsocietypublishing.org/doi/10.1098/rstb.2020.0535>

1885 El científico Louis Pasteur aplicó en Francia la primera vacuna antirrábica a un niño de 9 años, que había sido mordido por un perro. (Zoonosis)

<https://www.paho.org/es/noticias/3-7-2020-dia-mundial-zoonosis-pasteur-al-abordaje-salud>

7. Recolección de Datos y Recursos Relevantes

7.1 Recursos de Datos

En el caso que quisiéramos crear un modelo predictivo propio centrado exclusivamente en España, utilizamos estos enlaces de referencia.

Open data Salud Gobierno de España:

https://datos.gob.es/es/catalogo?theme_id=salud

Open data hospitales de España:

<https://opendata.esri.es/datasets/68745a7fb7a348b6b0d722c8517790af/explore?location=39.069736%2C-0.307884%2C5.92>

Opendata España portal Esri:

<https://opendata.esri.es/>

7.2 Enlaces y fuentes de interés

Lista completa de enlaces y referencias a fuentes relevantes utilizadas para la investigación.

Organización Mundial de la Sanidad Animal | Wahis:

<https://wahis.woah.org/#/dashboards/qd-dashboard>

Comisión Europea / Animal disease information- Overview reports:

https://food.ec.europa.eu/animals/animal-diseases/animal-disease-information-system-ads_en