

AVANCE DEL COVID-19 EN LATINOAMERICA Y EL CARIBE

RESUMEN

En este proyecto se analizará datos registrados por los diferentes países de latinoamérica y el caribe. Mediante la visualización de los datos en serie de tiempo y modelos de regresión se intentará extraer y resumir información relevante, y realizar predicciones acerca del progreso del Covid-19.

INTRODUCCIÓN Y JUSTIFICACIÓN

El Covid-19 ha causado estragos en el sistema sanitario, económico y social a nivel mundial. Ha cobrado una gran cantidad de vidas en el año 2020 y ha afectado la vida de todas las personas, debido a eso es necesario un monitoreo continuo y un análisis de datos acerca del progreso de la enfermedad, para tomar decisiones que puedan salvaguardar tanto la vida de las personas así como su calidad de vida.

OBJETIVOS

- + Evaluar la situación del avance del covid-19 hasta el 10 de junio del 2020
- + Identificar los países más afectados
- + Predecir el avance del covid-19 desde el 10 de junio del 2020 hasta la actualidad
- + Comparar las predicciones del modelo con los datos reales
- + Evaluar la metodología utilizada

METODOLOGÍA

Se utilizaron herramientas de análisis exploratorio y machine learning para extraer las características de los datos y hacer predicciones del avance del covid-19, tales como Python, Jupyter, Pandas, Geopandas, Matplotlib y Sklearn. Se evaluó el modelo mediante distintas funciones de costos y se contrastó con los datos reales mediante la visualización, para identificar fallas y sesgos en el modelo. El modelo que se utilizó es una regresión ajustada a una función logística.



Conclusión

El modelo utilizado para el ajuste de la trayectoria de la pandemia parece ser adecuado debido a la alta precisión arrojada en la etapa de evaluación del modelo sobre el conjunto de datos de prueba, se pudo crear una proyección. Sin embargo, no es suficiente para predecir el progreso real debido a las decisiones, cambios de política y circunstancias por las que atraviesan los distintos países de Latinoamérica y el Caribe, se puede decir que el modelo es adecuado solamente cuando la pandemia sigue su curso natural.

RESULTADOS

En la Gráfico 1 se puede observar el avance de la pandemia desde el periodo de 31/12/2019 hasta el 10/06/2020 y al resultado de la aplicación del modelo de regresión sobre la curva logística, se puede observar que la trayectoria del avance se ajusta con mucha exactitud a la trayectoria de la descripción.

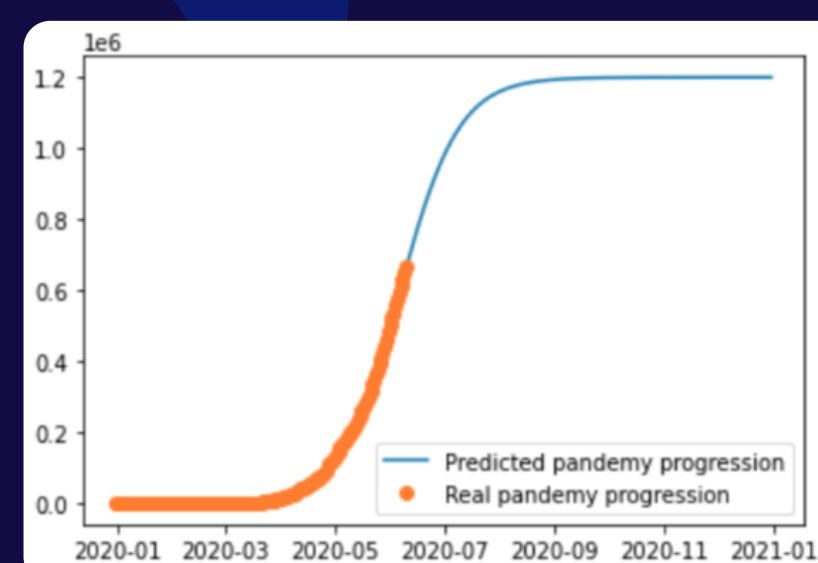


Gráfico 1. Análisis a priori del progreso de la pandemia hasta junio del 2020

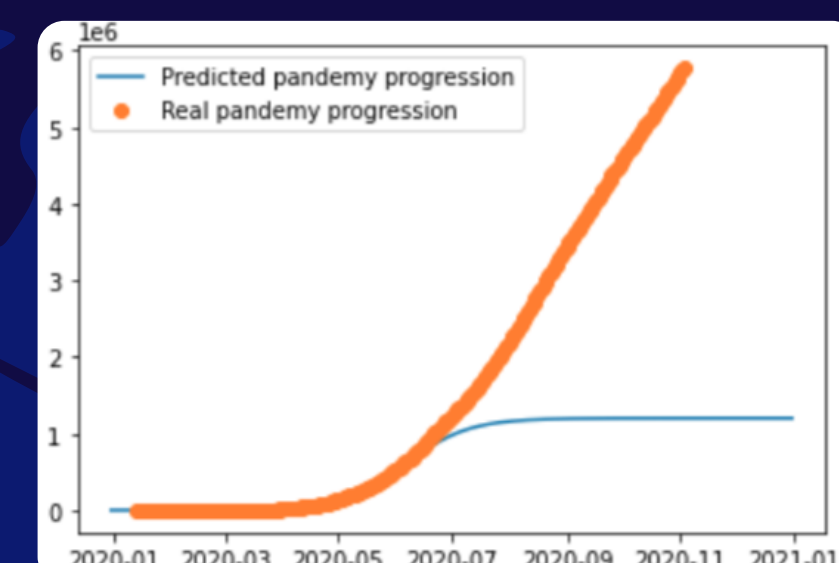


Gráfico 2. Análisis a posteriori del progreso de la pandemia hasta octubre de 2020

En el Gráfico 2 se aprecia como la trayectoria del avance de la pandemia se empieza a dissociar de la curva, siguiendo una trayectoria lineal en un punto en donde la trayectoria debería haber sido logarítmica, lo cual nos indica un aporte exponencial sobre el aporte logarítmico del progreso natural. Esto puede deberse principalmente a las políticas de liberación progresiva de la cuarentena en varios países de Latinoamérica y el Caribe.

Las métricas arrojadas por el modelo son la siguientes:

MSE: 24759360.0

RMSE: 4975.878

RRSE: 0.03159888

Precisión: 96.84 %

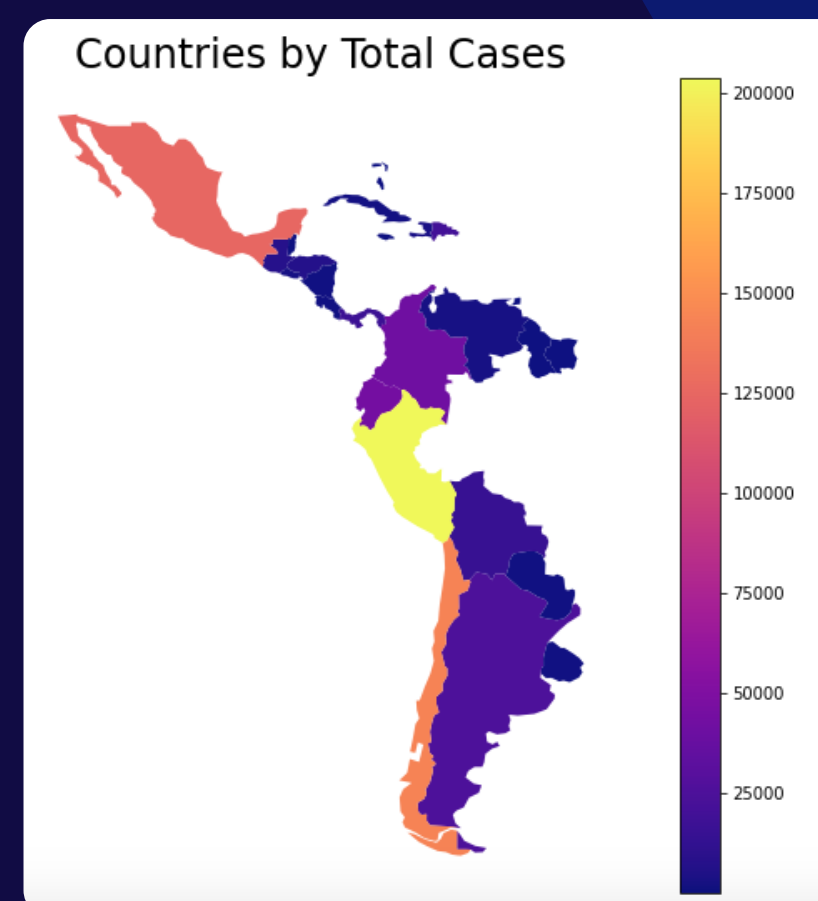


Gráfico 3. Mapa de calor de países por cantidad de infectados

Para el 10 de junio de 2020 los 10 países con mayor cantidad de casos son:

- 1 - Peru: 203736 casos
- 2 - Chile: 142759 casos
- 3 - Mexico: 124301 casos
- 4 - Ecuador: 43917 casos
- 5 - Colombia: 42078 casos
- 6 - Argentina: 24748 casos
- 7 - Dominican Republic: 20415 casos
- 8 - Panama: 17233 casos
- 9 - Bolivia: 14644 casos
- 10 - Guatemala: 7866 casos