

PREDICCIÓN DE BROTES EPIDÉMICOS BASADA EN DATOS DE SALUD

MAXIMILIANO LOBOS - JOSÉ PALMA - ALEXIS VILLARROEL

INTRODUCCIÓN

LA PREDICCIÓN DE BROTES EPIDÉMICOS ES UNA HERRAMIENTA FUNDAMENTAL EN LA SALUD PÚBLICA MODERNA, YA QUE PERMITE LA DETECCIÓN TEMPRANA DE POSIBLES EMERGENCIAS SANITARIAS. ESTA CAPACIDAD DE ANTICIPACIÓN ES CRUCIAL PARA IMPLEMENTAR MEDIDAS PREVENTIVAS QUE REDUZCAN EL IMPACTO DE LOS BROTES EN LA POBLACIÓN, SALVAGUARDANDO LA SALUD DE LAS COMUNIDADES Y OPTIMIZANDO EL USO DE LOS RECURSOS DISPONIBLES.

OBJETIVO

EL PROYECTO TIENE COMO PRINCIPAL OBJETIVO DESARROLLAR UN MODELO PREDICTIVO EFICIENTE QUE PERMITA IDENTIFICAR LA PROBABILIDAD DE OCURRENCIA DE BROTES INFECCIOSOS A PARTIR DE DATOS DE SALUD. ESTO INCLUYE LA APLICACIÓN DE TÉCNICAS DE PREPROCESAMIENTO DE DATOS, REDUCCIÓN DE DIMENSIONALIDAD MEDIANTE PCA Y EL USO DE UN MODELO DE REGRESIÓN LOGÍSTICA COMPARÁNDOLO CON UN MODELO DE ÁRBOL DE DECISIÓN CON EL FIN DE VISUALIZAR CUÁL ES EL APTO PARA ESTE PROYECTO

METODOLOGÍA

PREPROCESAMIENTO DE DATOS: SE REALIZÓ UNA LIMPIEZA EXHAUSTIVA DE LA BASE DE DATOS, ELIMINANDO COLUMNAS CON MÁS DEL 50% DE VALORES FALTANTES Y RELLENANDO LOS VALORES RESTANTES CON LA MEDIANA PARA CONSERVAR LA ROBUSTEZ.

ANÁLISIS DE COMPONENTES PRINCIPALES (PCA): ESTA TÉCNICA SE UTILIZÓ PARA REDUCIR LA DIMENSIONALIDAD, PRESERVANDO LA MAYOR CANTIDAD DE VARIABILIDAD POSIBLE Y MEJORANDO LA EFICIENCIA DEL MODELO.

MODELO DE REGRESIÓN LOGÍSTICA: ELEGIDO POR SU IDONEIDAD PARA PROBLEMAS DE CLASIFICACIÓN BINARIA. EL MODELO FUE ENTRENADO CON UN 80% DE LOS DATOS Y VALIDADO CON EL 20% RESTANTE, LOGRANDO UNA PRECISIÓN DE 96% Y UN ÁREA BAJO LA CURVA (AUC) DE 0.96.

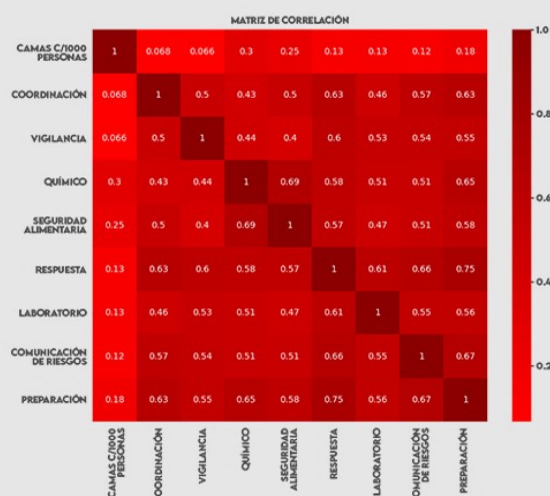
MODELO DE ÁRBOL DE DECISIÓN: IMPLEMENTADO PARA IDENTIFICAR RELACIONES NO LINEALES ENTRE LAS VARIABLES PREDICTORAS. EL MODELO FUE ENTRENADO CON EL 70% DE LOS DATOS Y VALIDADO CON EL 30% RESTANTE, ALCANZANDO UNA PRECISIÓN DEL 97% Y UN AUC-ROC DE 0.91. ADEMÁS, PERMITIÓ INTERPRETAR DE MANERA INTUITIVA LAS INTERACCIONES ENTRE LAS VARIABLES CLAVE.

RESULTADOS

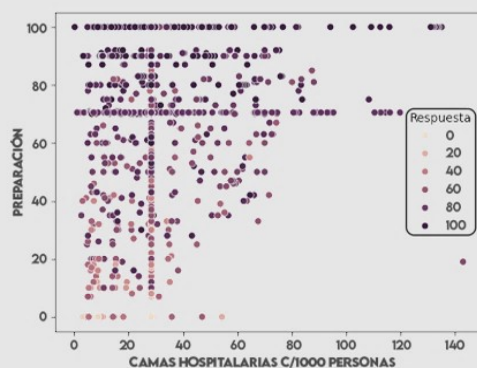
PRECISIÓN DEL MODELO: TANTO EL MODELO DE REGRESIÓN LOGÍSTICA COMO EL MODELO DE ÁRBOL DE DECISIÓN LOGRARON UNA PRECISIÓN SATISFACTORIA AL PREDECIR BROTES INFECCIOSOS EN EL CONJUNTO DE PRUEBA.



CORRELACIONES CLAVE: EL PCA REVELÓ GRUPOS DE VARIABLES CORRELACIONADAS, DESTACÁNDOSE INDICADORES COMO LA PREPARACIÓN Y LA CAPACIDAD DE RESPUESTA QUE INFLUYEN SIGNIFICATIVAMENTE EN LA PREDICCIÓN



VISUALIZACIÓN DE PATRONES: LA REDUCCIÓN DE DIMENSIONALIDAD FACILITÓ LA IDENTIFICACIÓN DE PATRONES CLAVE, AYUDANDO A INTERPRETAR MEJOR LOS DATOS.



CONCLUSIONES

ESTE PROYECTO DEMUESTRA EL POTENCIAL DE LAS TÉCNICAS DE MACHINE LEARNING EN LA SALUD PÚBLICA. LAS PREDICCIONES OBTENIDAS PUEDEN SER USADAS PARA PLANIFICAR RECURSOS Y ESTRATEGIAS DE INTERVENCIÓN, MEJORANDO LA RESPUESTA ANTE BROTES. EL TRABAJO MOSTRÓ LA IMPORTANCIA DEL PREPROCESAMIENTO Y EL USO DE PCA PARA OPTIMIZAR LOS MODELOS PREDICTIVOS. LA EXPERIENCIA REFORZÓ LA NECESIDAD DE UN ENFOQUE DE INVESTIGACIÓN Y ADAPTACIÓN CONSTANTE ANTE LOS RETOS QUE SURGIERON, COMO LA LIMPIEZA DE DATOS INCOMPLETOS Y LA SELECCIÓN DE MÉTRICAS ADECUADAS.