

A doctor in a white coat with a stethoscope around their neck is shown from the chest up. They are holding a small blue shield icon with a white cross inside. The background is a blurred clinical setting.

Cobertura Oncológica 2022: Un Análisis de Impacto y Tendencias en Perú

Harold Eustaquio

Objetivo del Proyecto

Evaluar las tendencias y patrones en la atención oncológica de 2022 para identificar brechas de cobertura y mejorar la calidad de los servicios ofrecidos a pacientes con cáncer en Perú.

Cobertura

Este proyecto cubre la atención oncológica en Perú, enfocándose en los departamentos y centros de salud con mayor demanda para los principales tipos de cáncer



Caso de Negocio

Permitir a FISSAL e IPRESS identificar áreas con alta demanda de atención oncológica, optimizar recursos y mejorar la calidad de vida de los pacientes en Perú.



Problemática

- Distribución Desigual de Servicios
- Optimización de Recursos
- Enfoque Personalizado en la Atención

Beneficios

- Mejora de Accesibilidad
- Eficiencia en la Asignación de Recursos
- Atención Adaptada al Paciente



CRISP DM

Comprensión del Negocio

Objetivos

Identificar cómo el análisis de las atenciones oncológicas puede mejorar la cobertura y distribución de recursos en Perú.

Preguntas clave

¿Cuáles son los tipos de cáncer más atendidos?

¿En qué regiones hay mayor demanda?

¿Cuál es el perfil de los pacientes en términos de edad, género y tipo de seguro?

Comprensión de los Datos

Objetivos

Obtener datos detallados sobre consultas, diagnósticos, edades, géneros, tipos de seguro y ubicación geográfica de las atenciones oncológicas.

Exploración Inicial

Examinar la distribución de consultas por departamento, tipo de cáncer, y perfil del paciente.

Preparación de Datos

Objetivos

Limpieza de datos
Transformación de variables
Selección de datos relevantes

Modelado

Objetivos

Selección de algoritmos:

- Series de Tiempo
- Regresión Logística
 - Naive Bayes
 - Clustering

Entrenamiento de modelos

Ajuste de parámetros

Evaluación

Objetivos

Validación del modelo

Interpretación práctica

Selección del mejor modelo

Despliegue

Objetivos

Implementación de resultados

Visualización y reportes



<https://github.com/haroldeustaquio/Data-Mining-UNAM/tree/main/final-project>



haroldeustaquio

Situación 1: Clasificación de Tipo de Servicio

Objetivo

- Predecir el tipo de servicio requerido en función de las características del paciente:
 - Departamento
 - Diagnóstico
 - Sexo
 - Edad
 - tipo de seguro

Conclusiones

CategoricalNB sería la mejor elección aquí, ya que muestra un balance sólido entre precisión y recall, con el **mayor F1 Score y accuracy** entre los modelos.

Esto sugiere que logra capturar correctamente tanto positivos como negativos con menos trade-off, siendo más confiable para clasificaciones equilibradas.

Modelos

- RandomForestClassifier
- Logistic Regression
- ComplementNB
- CategoricalNB
- DecisionTreeClassifier

	Model	Accuracy	Precision	Recall	F1 Score
0	RandomForestClassifier	0.442730	0.608599	0.442730	0.470853
1	LogisticRegression	0.393007	0.504617	0.393007	0.423366
2	ComplementNB	0.286261	0.512249	0.286261	0.257599
3	CategoricalNB	0.581283	0.569050	0.581283	0.573507
4	DecisionTreeClassifier	0.504679	0.604745	0.504679	0.522288

Situación 2: Análisis de Segmentación de Pacientes

Objetivo

- Agrupar a los pacientes en segmentos similares en función de sus características
 - Departamento
 - Diagnóstico
 - Sexo

Conclusiones

En esta situación, en los clúster destacan dos perfiles principales:

- mujeres de edades variadas con cáncer de mama
- jóvenes con leucemia

Todos beneficiarios del SIS Gratuito en Lima.

Modelos

- Kmeans (Elbow Method)
- DBSCAN

	departamento	grupo_diagnostics	sexo	edad	tipo_seguro	servicio
cluster						
0	Lima	CANCER DE MAMA	FEMENINO	70.0	SIS GRATUITO	Consulta externa
1	Lima	CANCER DE MAMA	FEMENINO	39.0	SIS GRATUITO	Consulta externa
2	Lima	CANCER DE MAMA	FEMENINO	54.0	SIS GRATUITO	Consulta externa
3	Lima	LEUCEMIA	MASCULINO	11.0	SIS GRATUITO	Consulta externa

Situación 3: Modelo de Serie de Tiempo para la Evaluación de la Cantidad de Pacientes por Mes

Objetivo

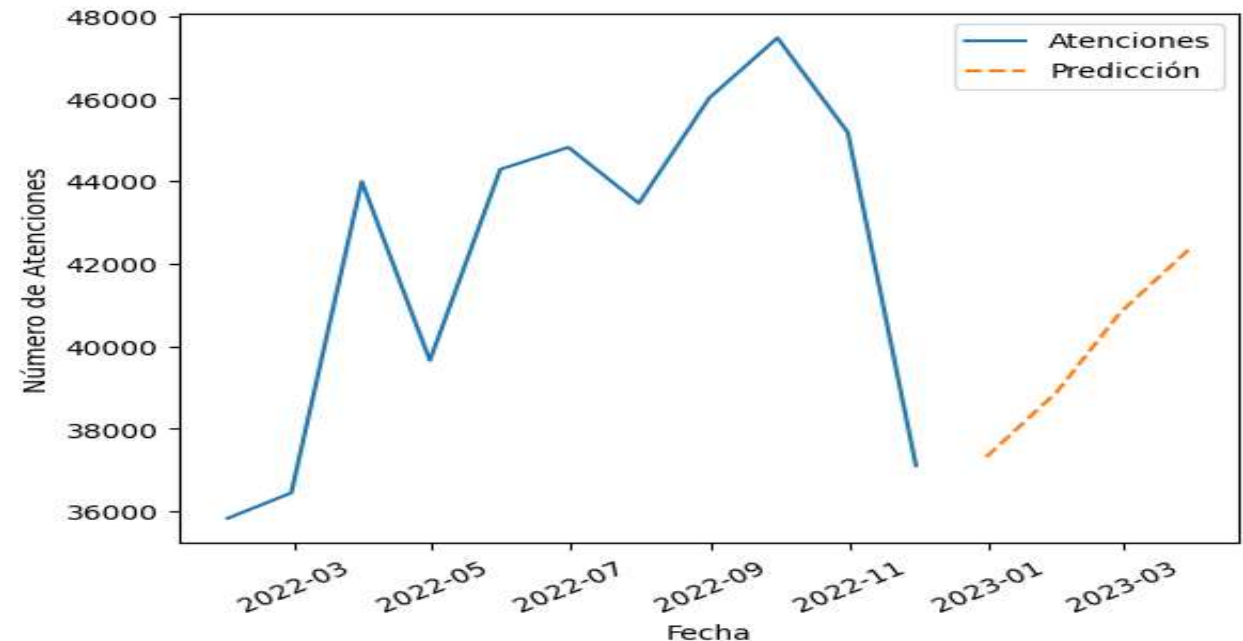
- El objetivo de este modelo es analizar la tendencia en la cantidad de pacientes atendidos mensualmente en función de los diagnósticos de cáncer más comunes.
 - Fecha
 - # de atenciones

Conclusiones

La gráfica muestra un aumento en las atenciones oncológicas hasta octubre de 2022, seguido de una caída al final del año. La proyección para 2023 indica una recuperación en la demanda, sugiriendo la necesidad de planificar recursos adicionales para los meses de alta demanda..

Modelos

- ARIMA



Dashboard Overview

