

# Modelos de machine learning para predecir la estancia en UCI utilizando patrones perioperatorios

\*Note: Sub-titles are not captured in Xplore and should not be used

Alexander Paredes<sup>1,2</sup> Eduardo Cuti<sup>1,2</sup> Angel Dianderas<sup>1,2</sup>.

**Abstract**—La medicina perioperatoria se enfoca en el cuidado de pacientes quirúrgicos en las fases preoperatoria, intraoperatoria y postoperatoria, enfrentando complicaciones impredecibles que pueden ser fatales y a menudo requieren monitoreo en UCI. Reducir riesgos es esencial, con 234 millones de cirugías mayores realizadas anualmente y 1.6 millones de muertes globales. En Perú, las cirugías abdominales tienen una alta mortalidad a corto plazo (15%-20%). La inteligencia artificial (IA) puede ayudar a predecir complicaciones postoperatorias, mejorando los resultados y reduciendo costos. Este proyecto utiliza la base de datos INSPIRE para desarrollar modelos de IA que identifiquen factores de riesgo en cirugías abdominales. La metodología incluye la selección, preprocesamiento y análisis de datos de pacientes mediante machine learning.

**Index Terms**—component, formatting, style, styling, insert

## I. INTRODUCCIÓN

La medicina perioperatoria es una disciplina esencial que incluye el cuidado antes, durante y después de una cirugía. Las complicaciones que ocurren en el periodo postoperatorio representan un desafío significativo para el personal de salud debido a su naturaleza impredecible y pueden aparecer de manera repentina produciendo que el paciente ingrese a la unidad de cuidados intensivos (UCI) y tenga que ser monitoreado constantemente; sin embargo, también se puede dar el caso de que el paciente fallezca sin las complicaciones son muy graves. [1]

Para asegurar una atención segura y minimizar los daños a los pacientes, es crucial identificar y reducir los factores de riesgo de posibles eventos adversos. Por lo tanto, es fundamental realizar una evaluación previa de las necesidades del paciente para evitar tales complicaciones. [2]. Con un análisis adecuado de las necesidades los pacientes pueden tener menos riesgos de sufrir complicaciones durante la operación. Asimismo, es importante tener en cuenta el cuidado post operatorio ya que este varía dependiendo de la gravedad del paciente y del conocimiento de los factores de riesgo durante la operación. Por ejemplo, un paciente con hipertensión debe ser monitoreado constantemente, lo cual es clave para tomar decisiones clínicas en cuanto al suministro de medicamentos prescritos. [3]

Identify applicable funding agency here. If none, delete this.

<sup>1</sup> Estudiante Pregrado de Especialidad de Ingeniería Biomédica Universidad Peruana Cayetano Heredia, Lima, Perú.

<sup>2</sup> Estudiante Pregrado de Especialidad de Ingeniería Biomédica Pontificia Universidad Católica del Perú, Lima, Perú.

Actualmente, se estima que en el mundo se realizan 234 millones de intervenciones quirúrgicas complejas. Estas intervenciones producen ciertas complicaciones que pueden conllevar a la muerte. Por ejemplo, en los países desarrollados se menciona que existe una tasa de mortalidad que va del 0,4% al 0,8%. [4] En el caso de Perú, también se cuenta con esta problemática y las cirugías que cuenta con mayor riesgo postoperatorio son las cirugías abdominales con una tasa de mortalidad a corto plazo que oscila entre el 15% y el 20%, siendo solo superada por la cirugía aórtica de emergencia [5][6]. Actualmente la Inteligencia Artificial (IA), se ha aplicado de diversas formas en el ámbito médico para ayudar en el diagnóstico y prevención de enfermedades [7]. Por ello, la IA también puede ser utilizada para la predicción de complicaciones postoperatorias y así disminuir los casos de enfermedad, la mortalidad y los costos de atención médica en hospitales y el lapso de internación hospitalaria, mejorando los resultados de los pacientes sometidos a cirugía. [8][9]

Para poder generar un modelo de IA, es necesario contar con una gran base de datos que esté verificada y sea de confianza. Se requiere una base de datos que proporcione información sobre los datos perioperatorios para poder generar modelos basados en machine learning que encuentren patrones de clasificación o de regresión. Actualmente, se cuenta con una base de datos de investigación colaborativa llamado INSPIRE, la cual cumple con estos requisitos y será de gran ayuda para desarrollar los modelos propuestos. [10] Finalmente, el objetivo de este proyecto es clasificar el tiempo en el que los pacientes luego de una operación estarán en UCI en base a parámetros perioperatorios utilizando modelos de machine learning con la finalidad de analizar las características que tienen mayor influencia.

## II. METODOLOGÍA

En primer lugar, se tuvo que seleccionar los datos de entrada. Luego es necesario pre-procesar esta información de tal forma que no cuente con información no deseada. Por último, se desarrollaron modelos de machine learning para evaluar el impacto de los factores durante las operaciones.

### A. Selección y descripción de los datos

La base de datos que se usó fue INSPIRE, a publicly available research dataset for perioperative medicine 1.2 de Physionet. De esta base de datos de extrajo a los pacientes, cuyas operaciones tenían como código 0WJG0 que implican

intervenciones en la cavidad peritoneal. De estos pacientes, solo se tomarán aquellos que cuenten con solo una operación realizada.

### B. Preprocesamiento

- La información de los pacientes por cada operación se saco del archivo vitals.csv. En este archivo se seleccionó por cada operación (op\_id), los valores medidos como ritmo cardiaco (hr), frecuencia respiratoria (rr), perdida de sangre estimada (ebl), presión arterial media (art\_mbp), etc. Los promedios por cada valor se almacenaron en un archivo aparte.
- El target a clasificar fue obtenido del archivo operations.csv. De este archivo, se filtro los datos en base al procedimiento a analizar mencionado y se extrajo el tiempo que el paciente estuvo en UCI .
- El archivo que contenía los datos de las operaciones requirió una limpieza de los datos. En este caso, se retiró a los pacientes que tenían más de dos operaciones. Además, se eliminó las columnas tenían más de 400 datos faltantes. De igual forma, se eliminó las filas que presentaban valores atípicos.
- Analizamos la data y segmentamos el tiempo en UCI, que se encuentra en la columna UCI\_TIME. Lo segmentamos en un tiempo menor a 5 días (4.8) (<7000 ) **[clase 0]** en una estancia corta, 5 a 10 (7000 - 14000) días **[clase 1]** en una estancia media y más de 10 días (>14000) **[clase 2]** en una estancia larga en UCI.
- Visualizamos un desbalance entre clases, por lo que decidimos usar una técnica para reducirlo. Utilizamos SMOTETomek y RandomUnderSampler. La técnica SMOTE (synthetic minority over sampling Technique) permite generar nuevas muestras sintéticas de las clases minoritarias, en nuestro caso, de las clases 1 y 2 . Luego de realizar esto, usamos Tomek (Tomek Links ), para eliminar pares de ejemplos cercanos pero que pertenecen a clases diferentes .Esto limpia/depura ejemplos ambiguos y mejora la separabilidad de clases. Por último, utilizamos Random Under Sampler reduce/elimina aleatoriamente el número de muestras de la clase mayoritaria, a fin de compensar el desbalance de las clases.

### C. Modelos de aprendizaje

Para desarrollar nuestros modelos, se dividió aleatoriamente el conjunto de datos en un 75% para el entrenamiento y un 25% para la prueba. Utilizamos el conjunto de entrenamiento para alimentar seis modelos de aprendizaje automático importados de la biblioteca scikit-learn. Estos modelos fueron seleccionados según su frecuencia de uso, y para cada uno se utilizaron los parámetros predeterminados. Los clasificadores aplicados fueron: K-Vecinos Más Cercanos, Máquina de Vectores de Soporte, Árbol de Decisión, Bosque Aleatorio, Gradient Boosting y Extreme Gradient Boosting. Asimismo, se realizó un modelo Redes Neuronales Multicapa (MLP).

## III. RESULTADOS

En la Fig. 1, se indica las métricas por cada modelo.

	Model	Accuracy (%)	Sensitivity or recall (%)	Specificity (%)	F1-Score (%)	AUC (%)
KNeighbors	Class 0	96.77	89.47	100.00	94.44	98.03
	Class 1	98.39	100.00	97.52	97.75	99.69
	Class 2	98.39	100.00	97.55	97.70	99.39
	Average	97.85	96.49	98.35	96.63	99.03
SVC	Class 0	99.60	98.68	100.00	99.34	100.00
	Class 1	100.00	100.00	100.00	100.00	100.00
	Class 2	99.60	100.00	99.39	99.42	100.00
	Average	99.73	99.56	99.80	99.58	100.00
DecisionTree	Class 0	95.16	94.74	95.35	92.31	95.04
	Class 1	97.18	93.10	99.38	95.86	96.24
	Class 2	97.98	97.65	98.16	97.08	97.90
	Average	96.77	95.16	97.63	95.08	96.40
RandomForest	Class 0	99.60	98.68	100.00	99.34	100.00
	Class 1	100.00	100.00	100.00	100.00	100.00
	Class 2	99.60	100.00	99.39	99.42	100.00
	Average	99.73	99.56	99.80	99.58	100.00
GradientBoosting	Class 0	98.79	96.05	100.00	97.99	99.98
	Class 1	99.60	100.00	99.38	99.43	100.00
	Class 2	99.19	100.00	98.77	98.84	99.97
	Average	99.19	98.68	99.38	98.75	99.98
XGBoost	Class 0	98.79	96.05	100.00	97.99	99.66
	Class 1	99.60	100.00	99.38	99.43	100.00
	Class 2	99.19	100.00	98.77	98.84	99.99
	Average	99.19	98.68	99.38	98.75	99.89

Fig. 1. Metricas de modelos de clasificación

Asimismo, se graficó las matrices de confusión por cada modelo de machine learning (Fig. 2). En esta matriz se observa que el modelo con los mejores resultados es Random Forest.

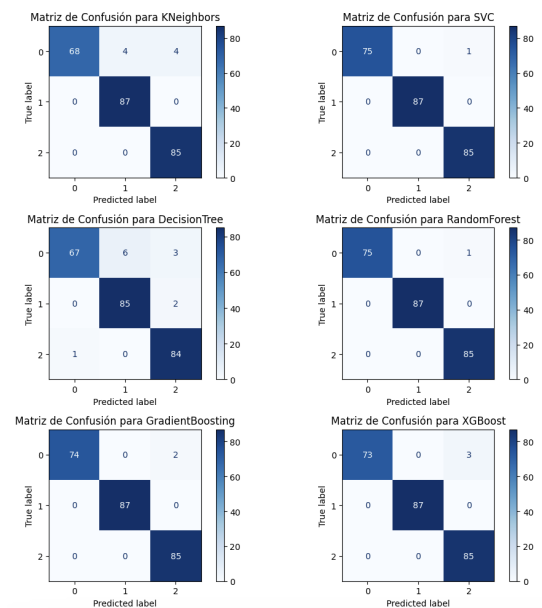


Fig. 2. Matriz de confusión de modelos de clasificación

Además, se realizó el análisis de las características que tenían mayor influencia en el modelo de SVC. Para ello se utilizó RFE y el resultado se muestra en la Fig 3.

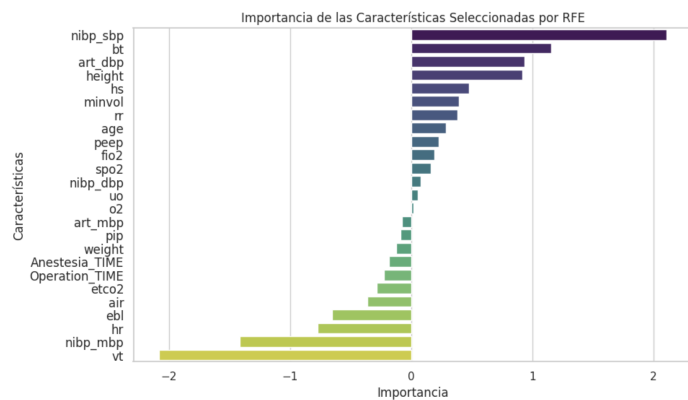


Fig. 3. Importancia de las características usando RFE

#### IV. DISCUSIÓN

Todas las métricas tienen valoraciones muy altas (llegando a 1.00): MLP, Random Forest, X Gradient Boosting y XG Boost ; sin embargo , el modelo decision tree tiene métricas más reales ; no obstante, decae en algunos valores cuando se realiza el cross validation score.

Las métricas del MLP tienen a la perfección (posible indicador de overfitting); sin embargo, cuando graficamos los plots de train - validation loss y accuracy, no se visualiza un overfitting a pesar de usar un epoch de 100.

#### V. CONCLUSIÓN

Todas las métricas tienen valoraciones muy altas (llegando a 1.00): MLP, Random Forest, Gradient Boosting y XG Boost ; sin embargo , el modelo decision tree tiene métricas más reales ; no obstante, decae en algunos valores cuando se realiza el cross validation score.

Las métricas del MLP tienen a la perfección (posible indicador de overfitting); sin embargo, cuando graficamos los plots de train - validation loss y accuracy, no se visualiza un overfitting a pesar de usar un epoch de 100.

#### REFERENCES

- [1] "Association Between Postoperative Troponin Levels and 30-Day Mortality Among Patients Undergoing Noncardiac Surgery," *JAMA*, vol. 307, no. 21, pp. 2295–2295, Jun. 2012, doi: <https://doi.org/10.1001/jama.2012.5502>. Available: <https://jamanetwork.com/journals/jama/fullarticle/1172044>. [Accessed: Apr. 14, 2024]
- [2] C. Mayor, A. De Coordinadora, A. Ma, and L. Alonso, "Guía Cuidados en." Available: [https://www.asecma.org/Documentos/Blog/LIB\\_I\\_6\\_C\\_1.PDF](https://www.asecma.org/Documentos/Blog/LIB_I_6_C_1.PDF). [Accessed: Apr. 14, 2024]
- [3] Fernández P, Espuñes E. Manual de Enfermería Quirúrgica. 1era ed. Esp. Servei de Publicacions. 2012; vol 1: 17-19. Disponible en: [https://publicacions.uab.cat/pdf\\_lilibrs/TRII008.pdf](https://publicacions.uab.cat/pdf_lilibrs/TRII008.pdf)
- [4] "En el mundo se realizan 234 millones de operaciones de cirugía mayor al año," *Www.imq.es*, 2019. Available: [https://www.imq.es/corporativo/es\\_ES/particular/imq-corporativo/sala-prensa/Noticias/en-el-mundo-se-realizan-234-millones-operaciones-cirugia-mayor-al-ano](https://www.imq.es/corporativo/es_ES/particular/imq-corporativo/sala-prensa/Noticias/en-el-mundo-se-realizan-234-millones-operaciones-cirugia-mayor-al-ano). [Accessed: May 08, 2024]

- [5] W. R. Calmet Rocca, "Mortalidad de pacientes ingresados a la unidad de cuidados intensivos después de una cirugía abdominal de emergencia en un hospital de tercer nivel de Lima, Perú", *Mortality of patients admitted to the intensive care unit after emergency abdominal surgery in a tertiary level hospital at Lima, Peru*, 2023, Accedido: 10 de junio de 2024. [En línea]. Disponible en: <https://repositorio.upch.edu.pe/handle/20.500.12866/14366>
- [6] A. Clarke, H. Murdoch, M. J. Thomas, T. M. Cook, and C. J. Peden, "Mortality and postoperative care after emergency laparotomy," *European journal of anaesthesiology*, vol. 28, no. 1, pp. 16–19, Jan. 2011, doi: <https://doi.org/10.1097/eja.0b013e32833f5389>. Available: <https://pubmed.ncbi.nlm.nih.gov/20829702/>. [Accessed: May 08, 2024]
- [7] X.-Y. Zhou, Y. Guo, M. Shen, and G.-Z. Yang, "Application of artificial intelligence in surgery," *Frontiers of Medicine*, vol. 14, no. 4, pp. 417–430, Jul. 2020, doi: <https://doi.org/10.1007/s11684-020-0770-0>. Available: <https://link.springer.com/article/10.1007/s11684-020-0770-0>. [Accessed: Apr. 14, 2024]
- [8] S. Johnston, M. Louis, Leonid Churilov, R. Ma, C. Christophi, and L. Weinberg, "Health costs of post-operative complications following rectal resection: a systematic review," *ANZ journal of surgery/ANZ journal of surgery*, vol. 90, no. 7–8, pp. 1270–1276, Feb. 2020, doi: <https://doi.org/10.1111/ans.15708>. Available: <https://pubmed.ncbi.nlm.nih.gov/32053858/>. [Accessed: Apr. 14, 2024]
- [9] K. Estefanía and G. Correa, "Aplicaciones de la inteligencia artificial en cirugía," *Salud ConCiencia*, vol. 2, no. 2, pp. e31–e31, Oct. 2023, doi: <https://doi.org/10.55204/scc.v2i2.e31>. Available: <https://saludconciencia.com.ar/index.php/scc/article/view/31>. [Accessed: Apr. 14, 2024]
- [10] Gonçalves DM, Henriques R, Costa RS. Predicting Postoperative Complications in Cancer Patients: A Survey Bridging Classical and Machine Learning Contributions to Postsurgical Risk Analysis. *Cancers (Basel)*. 2021 Jun 28;13(13):3217. doi: 10.3390/cancers13133217. PMID: 34203189; PMCID: PMC8269422.
- [11] Stone K, Zwiggelaar R, Jones P, Mac Parthaláin N. A systematic review of the prediction of hospital length of stay: Towards a unified framework. *PLOS Digit Health*. 2022 Apr 14;1(4):e0000017. doi: 10.1371/journal.pdig.0000017. PMID: 36812502; PMCID: PMC9931263.
- [12] INSPIRE, a publicly available research dataset for perioperative medicine v1.2. (2023, 28 diciembre). <https://physionet.org/content/inspire/1.2/>
- [13] P. Haya, "La metodología CRISP-DM en ciencia de datos - IIC," *Instituto de Ingeniería del Conocimiento*, Nov. 29, 2021. <https://www.iic.uam.es/innovacion/metodologia-crisp-dm-ciencia-de-datos/> (accessed Jun. 03, 2024).