



Predicción de cobertura de vacunación contra H1N1 y gripe estacionaria – Entrega Final

Inteligencia Artificial para las Ciencias e Ingeniería

Eliana Salas Villa, Marisol Correa Gutiérrez, Manuela Ospina Giraldo

Bioingeniería, Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia

eliana.salas@udea.edu.co, marisol.correag@udea.edu.co, manuela.ospinag@udea.edu.co

Noviembre 25, 2023

1. Introducción

La salud pública es un componente esencial en la vida de cualquier sociedad, y la prevención de enfermedades a través de la vacunación desempeña un papel fundamental en este aspecto. En este contexto, el presente trabajo tiene como objetivo principal predecir si una persona tiene acceso o no a las vacunas de H1N1 y de gripe estacional. Se emplearon tres modelos de inteligencia artificial con el fin de determinar cuál tiene un mejor desempeño para la predicción: una Red Neuronal, una Máquina de Vectores de Soporte (SVM) y el algoritmo de agrupamiento K-Means.

La Red Neuronal es un modelo inspirado en la estructura del cerebro humano. Está compuesto por nodos llamados neuronas, organizados en capas, y cada conexión entre nodos tiene un peso que se ajusta durante el entrenamiento. Este modelo es capaz de aprender patrones complejos y no lineales en los datos, lo que lo hace adecuado para problemas de clasificación como el propuesto en este estudio.

Las SVM son modelos de aprendizaje supervisado utilizados para clasificación y regresión. En el contexto de clasificación, SVM busca encontrar el hiperplano óptimo que maximiza la separación entre clases. Este modelo es efectivo incluso en espacios de alta dimensionalidad y puede manejar datos no lineales mediante el uso de funciones de kernel.

K-Means es un algoritmo de agrupamiento que divide un conjunto de datos en k grupos basándose en similitudes. Este método es útil para identificar patrones y segmentar la población en grupos homogéneos. Aunque K-Means es más comúnmente utilizado para tareas de agrupamiento, se explorará su capacidad para clasificar en este contexto.

La base de datos utilizada consta de 35 características sociales recolectadas de 6437 individuos. Estas características pueden abarcar una amplia gama de variables, como edad, ingresos, educación, entre otras. La riqueza de datos sociales proporciona una oportunidad única para identificar factores determinantes en el acceso a las vacunas.

Este estudio es crucial en el contexto de la salud pública, ya que identificar los factores que influyen en el acceso a las vacunas puede ayudar a diseñar estrategias más efectivas de distribución y concienciación. La implementación de modelos de inteligencia artificial permite analizar grandes conjuntos de datos de manera eficiente y descubrir patrones que podrían pasar desapercibidos mediante métodos tradicionales.

La implementación de inteligencia artificial en este estudio permite realizar predicciones precisas y automatizadas basadas en patrones complejos presentes en los datos sociales. Estos modelos pueden ayudar a identificar poblaciones en riesgo, optimizar la asignación de recursos y desarrollar estrategias personalizadas para aumentar la cobertura de vacunación. La inteligencia artificial ofrece un enfoque innovador y poderoso para abordar problemas de salud pública y mejorar la toma de decisiones en la prevención de enfermedades.

2. Base de Datos

El objetivo es predecir la probabilidad de que los individuos reciban las vacunas contra la gripe H1N1 y la gripe estacional. En concreto, predecirá dos probabilidades: una para *h1n1_vaccine* y otra para *seasonal_vaccine*, así como cualquier análisis exploratorio de datos (AED) sofisticado. Cada fila del conjunto de datos representa las respuestas de una persona a la Encuesta nacional sobre la gripe H1N1 realizada en 2009 en Estados Unidos y hay dos variables objetivo:

- *h1n1_vaccine*: Si el encuestado recibió la vacuna contra la gripe H1N1.
- *seasonal_vaccine*: Si el encuestado recibió la vacuna contra la gripe estacional.

Ambas son variables binarias: 0 = No; 1 = Sí. Algunos encuestados no recibieron ninguna de las dos vacunas, otros recibieron sólo una y algunos recibieron ambas. Se trata de un problema multi-etiqueta (y no multiclase).

El conjunto de datos tomados consta de 36 columnas. La primera columna *respondent_id* es un identificador único y aleatorio. Las 35 columnas siguientes corresponden a características como edad del encuestado, nivel de preocupación por el H1N1 (de 0 a 3), si el doctor le ha recomendado o no que se ponga la vacuna (0 o 1), ocupación, región de residencia, entre otras.

La base de datos se encuentra desbalanceada, lo que puede llevar a un sesgo hacia las clases mayoritarias, dificultades en la generalización, y desafíos en la evaluación del desempeño. Para abordar este problema se plantea a futuro balancear los datos de manera manual, escogiendo aleatoriamente los datos usados para entrenamiento, el resto de los datos serán usados para la etapa de validación del modelo.

3. Exploración descriptiva del dataset

Para empezar, se realizó el gráfico de barras de la figura 1, en el que se puede observar cuál es la proporción inicial entre personas vacunadas y no vacunadas para cada una de las vacunas (H1N1 y gripe estacional) respectivamente en el dataset crudo, esto para tener una idea global del mismo previo al preprocesado.

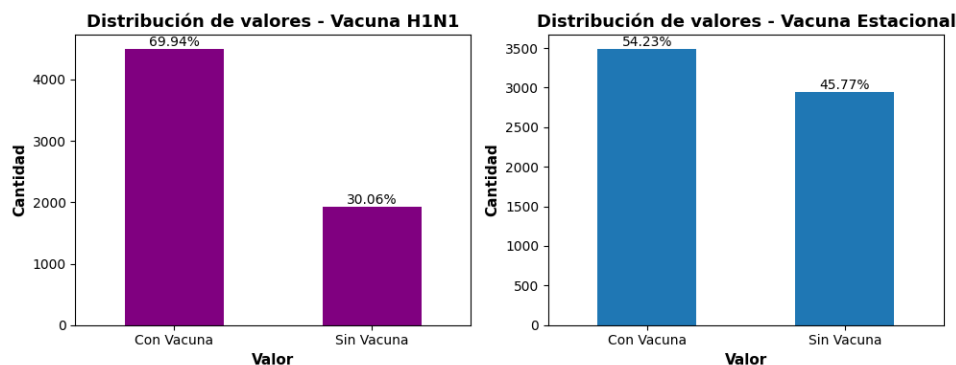


Figura 1. Distribución de resultados esperados para cada vacuna.

4. Iteraciones de desarrollo.

Preprocesado de datos

Para poder lograr que el *dataframe* sea apto para el entrenamiento de un algoritmo de *machine learning* se siguieron ciertos pasos:

1. Se eliminaron las últimas dos columnas que correspondían a las etiquetas de los datos para cada vacuna.
2. Se eliminó la primera columna que correspondía al ID de cada participante
3. Se eliminaron las filas que contenían caracteres nulos.
4. Se convirtieron todas las columnas del *dataframe* en variables categóricas.
5. Se realizó la codificación one-hot de las variables categóricas para que todas las entradas al algoritmo sean numéricas.

División de datos en entrenamiento y prueba utilizando sobre muestreo

- Datos balanceados

Se tomó el 80% de los datos para entrenamiento y el 20% para prueba. Se estableció una semilla fija de 42 para asegurarse de que la división de datos sea reproducible. Luego se aplicó SMOTE al conjunto de entrenamiento y al conjunto de prueba para abordar el desequilibrio de clases generando instancias sintéticas de la clase minoritaria. En las figuras 2 y 3 se observa la proporción de datos para el set de entrenamiento y prueba respectivamente para H1N1 y Gripe estacional.

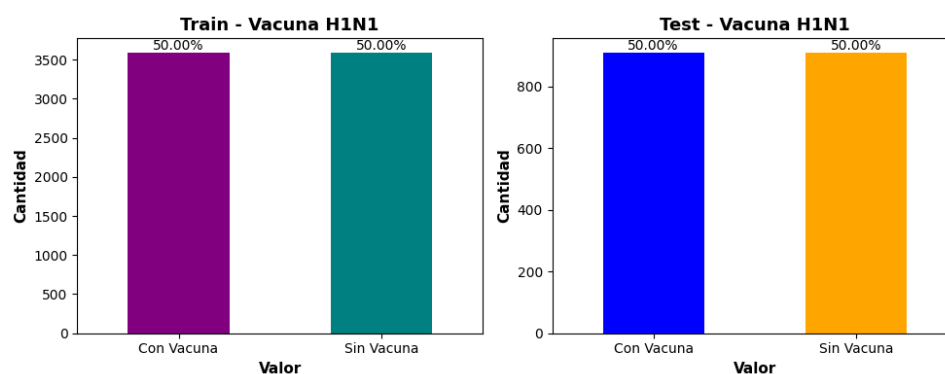


Figura 2. Proporción de sujetos vacunados y no vacunados contra H1N1 para sets de entrenamiento y prueba.

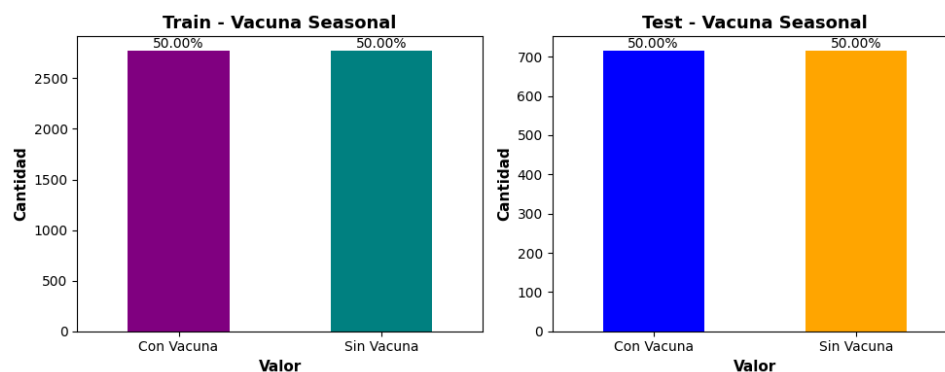


Figura 3. Proporción de sujetos vacunados y no vacunados contra Gripe estacional para sets de entrenamiento y prueba.

Modelos supervisados

Red Neuronal

La Red Neuronal es un modelo inspirado en la estructura del cerebro humano. Está compuesto por nodos llamados neuronas, organizados en capas, y cada conexión entre nodos tiene un peso que se ajusta durante el entrenamiento. Este modelo es capaz de aprender patrones complejos y no lineales en los datos, lo que lo hace adecuado para problemas de clasificación como el propuesto en este estudio.

Previo a la creación de cada modelo, se creó una instancia de escalador estándar para que las características del set de entrenamiento tuviesen una media de 0 y desviación estándar de 1, ayudando a estabilizar y normalizar las características. Luego, se realizaron distintas variaciones de los hiperparámetros de la red neuronal (capas ocultas y número de neuronas) y empleando para todas, la función de activación sigmoide. Finalmente, para todos los casos se compiló el modelo escogido especificando el optimizador Adam, la función de pérdida '*Binary_crossentropy*' para clasificación binaria y las métricas a seguir (en este caso, la precisión). Fueron probados 7 modelos con combinaciones de hiperparámetros diferentes y los que proporcionaron la mejor precisión se presentan en la tabla 1 :

Tabla 1. Variación de Estructura – Vacuna H1N1

| Modelo | Capas Ocultas | Neuronas | Accuracy Entrenamiento | Accuracy Prueba |
|-----------------|---------------|-------------|------------------------|-----------------|
| Modelo 1 | 1 | 32 | 83.54% | 84% |
| Modelo 2 | 1 | 64 | 83.62% | 84% |
| Modelo 3 | 1 | 128 | 83.39% | 83% |
| Modelo 4 | 2 | 32, 32 | 82.69% | 83% |
| Modelo 5 | 2 | 64, 32 | 83.23% | 83% |
| Modelo 6 | 2 | 64, 64 | 83.23% | 83% |
| Modelo 7 | 3 | 64, 64, 128 | 83.15% | 83% |

Tabla 2. Hiperparámetros de la red neuronal que mejor predijo H1N1

| Capas ocultas | Neuronas | Función de activación | Optimizador | Función de pérdida | Número de épocas | Tamaño del lote | Accuracy set prueba |
|---------------|----------|-----------------------|-------------|---------------------|------------------|-----------------|---------------------|
| 1 | 64 | Sigmoide | Adam | Binary crossentropy | 15 | 64 | 79.09% |

En las figuras 4 y 5 se muestran las matrices de confusión para cada set de datos y métricas de desempeño adicionales.

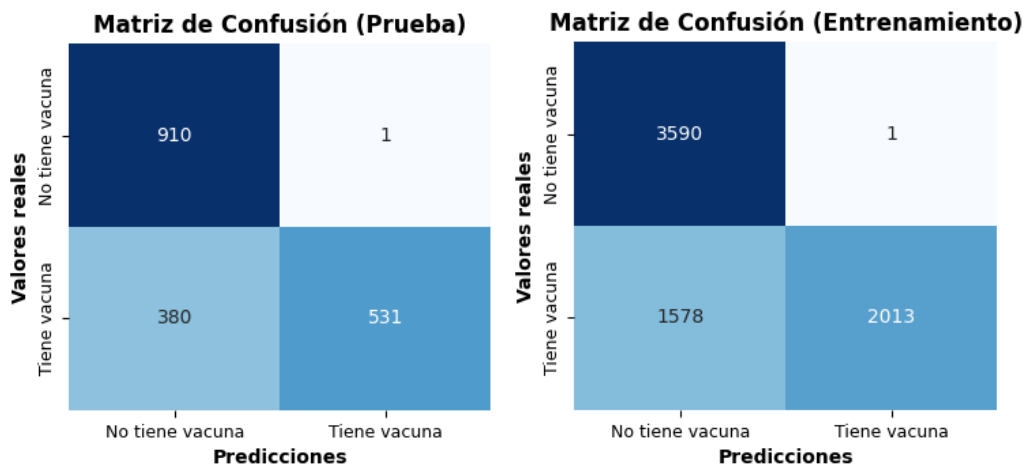


Figura 4. Matrices de confusión

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.74 | 0.88 | 0.80 | 716 |
| 1 | 0.85 | 0.69 | 0.76 | 716 |
| accuracy | | | 0.78 | 1432 |
| macro avg | 0.79 | 0.78 | 0.78 | 1432 |
| weighted avg | 0.79 | 0.78 | 0.78 | 1432 |

Figura 5. Métricas adicionales

Support Vector Machine

SVM busca encontrar el hiperplano óptimo que maximiza la separación entre clases. Este modelo es efectivo incluso en espacios de alta dimensionalidad y puede manejar datos no lineales mediante el uso de funciones de kernel.

Tabla 3. Hiperparámetros del modelo de SVM mejor predijo H1N1 y Seasonal flu

| Kernel | Penalización | Accuracy set de prueba H1N1 | Accuracy set de prueba Seasonal |
|--------|--------------|-----------------------------|---------------------------------|
| Lineal | 10 | 85% | 74% |

En las figuras 6 y 7 se muestran las matrices de confusión para cada set de datos y métricas de desempeño adicionales.

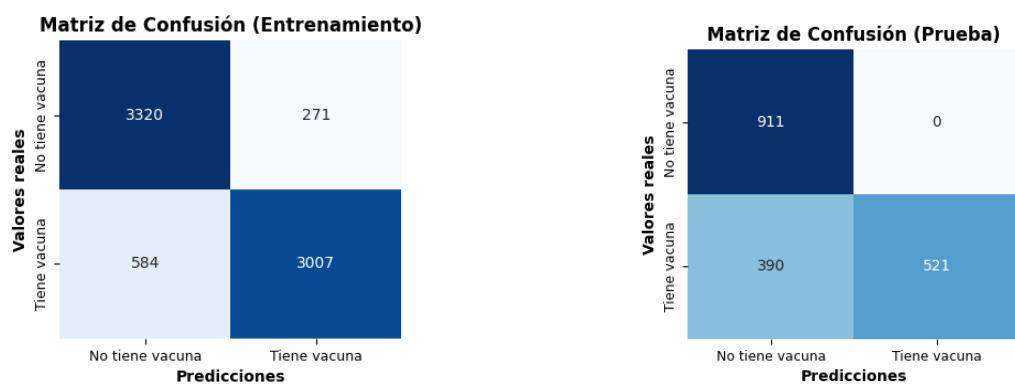


Figura 6. Matrices de confusión

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.70 | 1.00 | 0.82 | 911 |
| 1 | 1.00 | 0.57 | 0.73 | 911 |
| accuracy | | | 0.79 | 1822 |
| macro avg | 0.85 | 0.79 | 0.78 | 1822 |
| weighted avg | 0.85 | 0.79 | 0.78 | 1822 |

Figura 7. Métricas adicionales

Como este es el modelo que mejor predijo ambos grupos, se obtuvo adicionalmente una curva de aprendizaje.

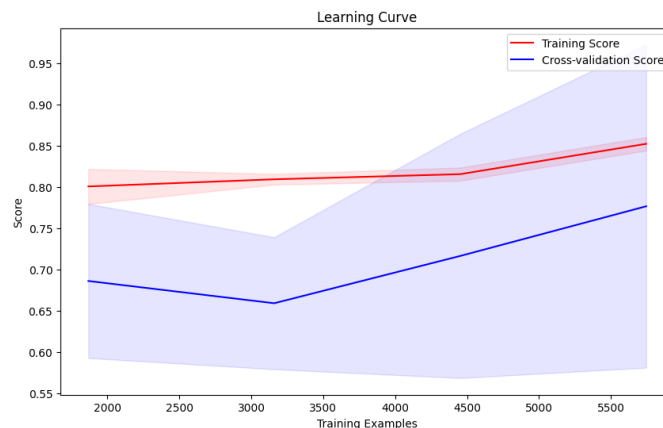


Figura 8. Curva aprendizaje SVM

Validación cruzada

Se realizó validación cruzada k-fold con el objetivo de generalizar el modelo en datos no vistos. En esta técnica, el conjunto de datos se divide en k pliegues (folds), y el modelo se entrena y evalúa k veces. En cada iteración, un pliegue se utiliza como conjunto de prueba, mientras que los otros se utilizan como conjunto de entrenamiento. Las métricas de rendimiento se promedian a lo largo de las k iteraciones para obtener una estimación más precisa del rendimiento del modelo.

El código implementa la validación cruzada estratificada utilizando las bibliotecas scikit-learn y Keras en Python. Primero, normaliza los datos de entrenamiento con StandardScaler de scikit-learn, lo cual es una práctica común para garantizar que todas las características tengan la misma escala y mejorar así la convergencia del modelo. Luego, se configura un modelo de Support Vector Machine (SVM) con un kernel lineal y un parámetro de regularización C igual a 0.1.

La configuración de la validación cruzada se realiza utilizando la clase KFold de scikit-learn, dividiendo el conjunto de datos en un número específico de pliegues (folds), en este caso, el valor de num_folds que varía a lo largo de las iteraciones del bucle. Durante cada iteración, se evalúa el modelo utilizando la función cross_val_score de scikit-learn, que realiza la validación cruzada y devuelve los puntajes de rendimiento para cada pliegue.

Los resultados se almacenan en un diccionario llamado results, donde la clave es el número de pliegues y el valor es el accuracy promedio obtenido durante la validación cruzada. Estos resultados se imprimen en la consola para su análisis.

Finalmente, el modelo se entrena utilizando todos los datos de entrenamiento. Se realizan predicciones tanto en el conjunto de entrenamiento como en el conjunto de prueba, lo que permite evaluar el rendimiento del modelo entrenado en datos no vistos. Este experimento se centra exclusivamente en la evaluación del rendimiento de un modelo de Support Vector Machine (SVM)

Modelo No Supervisado

K-Means

Este algoritmo agrupa los datos disponibles de un dataset de acuerdo con sus similitudes en las características introducidas al algoritmo. Con este algoritmo los resultados no son buenos puesto que se presenta un alto porcentaje de error el cual alcanza un valor de 45.71% y se utilizaron 2 clusters puesto que según los resultados de la prueba Silhouette, este número de cluster es el más adecuado para los datos que se tienen a disposición.

- Datos No Balanceados

Se tomó el 80% de los datos para entrenamiento y el 20% para prueba. Se estableció una semilla fija de 42 para asegurarse de que la división de datos sea reproducible. En las figuras 9 y 10 se observa la proporción de datos para el set de entrenamiento y prueba respectivamente para H1N1 y Gripe estacional.

Vacuna H1N1

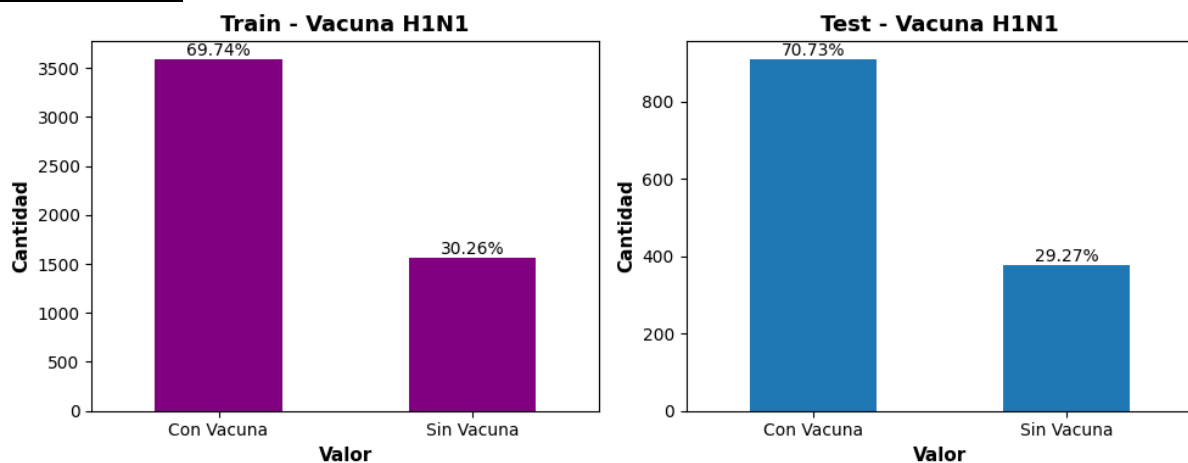


Figura 9. Proporción de sujetos vacunados y no vacunados contra H1N1 para sets de entrenamiento y prueba.

Gripe estacional

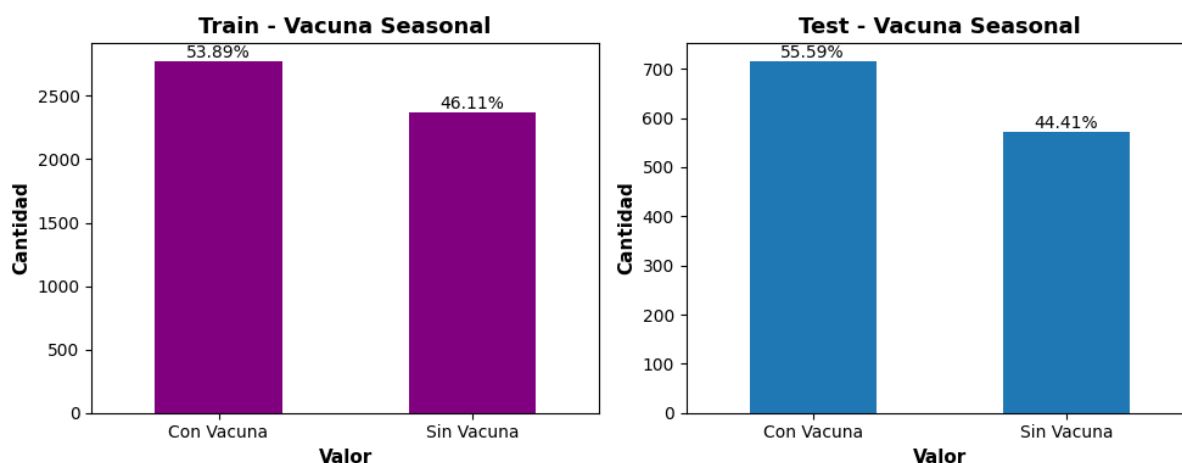


Figura 10. Proporción de sujetos vacunados y no vacunados contra Gripe estacional para sets de entrenamiento y prueba.

Modelos Supervisados

Red Neuronal

Previo a la creación de cada modelo, se creó una instancia de escalador estándar para que las características del set de entrenamiento tuviesen una media de 0 y desviación estándar de 1, ayudando a estabilizar y normalizar las características. Luego, se realizaron distintas variaciones de los hiperparámetros de la red neuronal (capas ocultas y número de neuronas) y empleando para todas, la función de activación sigmoide. Finalmente, para todos los casos se compiló el modelo escogido especificando el optimizador Adam, la función de pérdida *'Binary_crossentropy'* para clasificación binaria y las métricas a seguir (en este caso, la precisión). En la tabla a continuación, se observa el valor de precisión (accuracy) para el set de entrenamiento y prueba para cada uno de los modelos escogidos, así como los hiperparámetros que corresponden a cada uno.

Tabla 4. Variación de Estructura – Vacuna Gripe estacional

| Modelo | Capas Ocultas | Neuronas | Accuracy Entrenamiento | Accuracy Prueba |
|-----------------|---------------|---------------|------------------------|-----------------|
| Modelo 1 | 1 | 32 | 83.15% | 85% |
| Modelo 2 | 1 | 64 | 83.31% | 85% |
| Modelo 3 | 1 | 128 | 83.39% | 85% |
| Modelo 4 | 2 | 32, 32 | 83.31% | 85% |
| Modelo 5 | 2 | 64, 32 | 83.46% | 85% |
| Modelo 6 | 2 | 64, 64 | 83.15% | 86% |
| Modelo 7 | 3 | 64, 64, 128 | 80.36% | 82% |

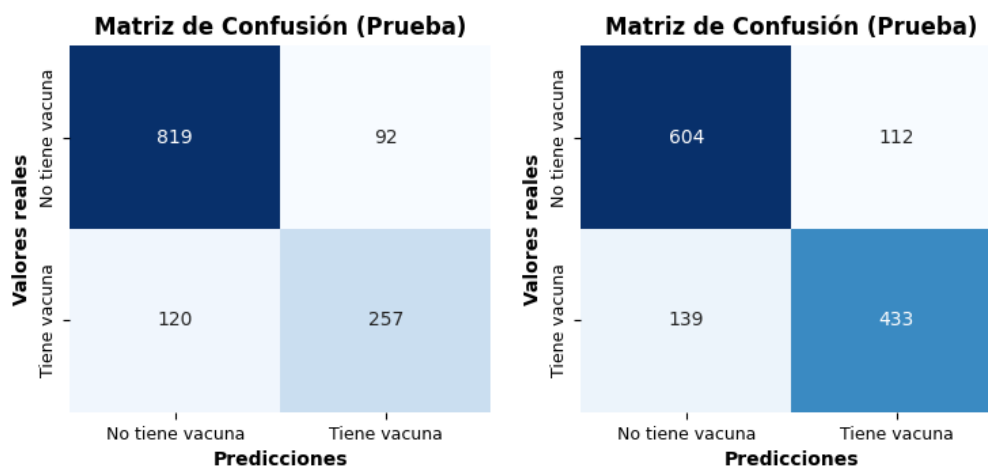


Figura 11. Matrices de confusión de H1N1 y Seasonal respectivamente

Support Vector Machine

Se obtienen mejores resultados pues el modelo es entrenado con una mayor variabilidad para la clase de “No tiene acceso a vacuna” dándole un mejor entrenamiento para esta clase al modelo y permitiendo alcanzar mejores valores de accuracy como se muestra en la tabla 5 y en la figura 12.

Tabla 5. Hiperparámetros del modelo de SVM mejor predijo H1N1 y Seasonal flu

| Kernel | Penalización | Accuracy set de prueba H1N1 | Accuracy set de prueba Seasonal |
|--------|--------------|-----------------------------|---------------------------------|
| Lineal | 10 | 82.74% | 78.79% |

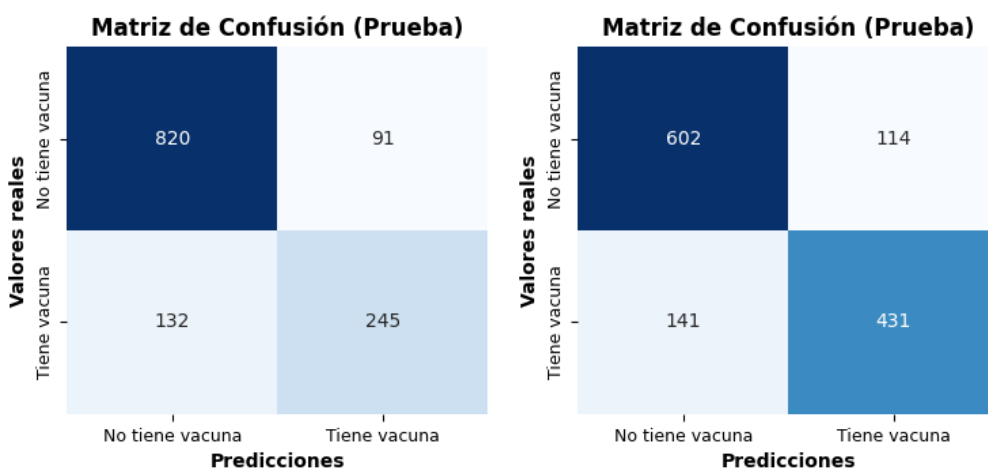


Figura 11. Matrices de confusión de H1N1 y Seasonal respectivamente

Modelo No Supervisado

K-Means

En este caso se obtuvo un error menor al alcanzado con un set de datos balanceado y el error obtenido fue de 39.99% en el caso de H1N1 y de 34.63% en el caso de gripe estacional.

Como se observa anteriormente, en general se tienen mejores resultados al tener un dataset desbalanceado pues se da una mayor variabilidad a la clase de “No tienen acceso a vacuna” que es la clase que, en prueba, tendrá una mayor variación por lo que se disminuyen los

casos de falsos positivos y falsos negativos. Por esto los modelos compuestos expuestos a continuación seguirán siendo entrenados con un set de datos desbalanceado.

- Modelos compuestos

K-Means + Redes Neuronales

K-Means es un algoritmo de agrupamiento que divide un conjunto de datos en k grupos basándose en similitudes. Este método es útil para identificar patrones y segmentar la población en grupos homogéneos. Aunque K-Means es más comúnmente utilizado para tareas de agrupamiento, se explorará su capacidad para clasificar en este contexto.

Se aplicó una normalización de datos y luego una red neuronal que ayudaron a identificar estructuras en los datos que no eran evidentes en su forma original. Posteriormente, K-Means se utilizó para agrupar las instancias de entrenamiento en clústeres y analizar la frecuencia y frecuencia relativa de los valores de la variable objetivo dentro de cada clúster, proporcionando una visión adicional sobre la distribución de las etiquetas en los grupos identificados por K-Means.

Los hiperparámetros que proporcionaron el mejor desempeño del modelo fueron:

Red neuronal:

Tabla 6. Red neuronal escogida

| Capas ocultas | Neuronas | Función de activación | Optimizador | Función de pérdida | Número de épocas | Tamaño del lote |
|---------------|----------|-----------------------|-------------|---------------------|------------------|-----------------|
| 1 | 64 | Sigmoide | Adam | Binary crossentropy | 15 | 64 |

K-means:

Tabla 7. K-means

| Número de clústeres | Número de inicializaciones | Número máximo de iteraciones |
|---------------------|----------------------------|------------------------------|
| 2 | 1000 | 100000 |

El accuracy proporcionado por este algoritmo compuesto fue de 82% para H1N1 y de 62% para la gripe estacional.

En las figuras 12 y 13 se muestra la matriz de confusión y las métricas adicionales.

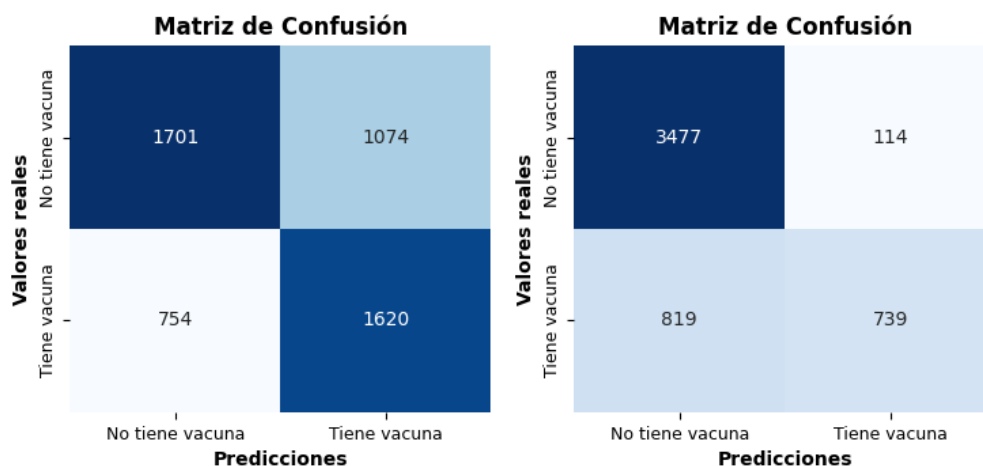


Figura 12. Matrices de confusión H1N1 y Seasonal

| | precision | recall | f1-score | support | | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|--------------|-----------|--------|----------|---------|
| 0 | 0.81 | 0.97 | 0.88 | 3591 | 0 | 0.69 | 0.61 | 0.65 | 2775 |
| 1 | 0.87 | 0.47 | 0.61 | 1558 | 1 | 0.60 | 0.68 | 0.64 | 2374 |
| accuracy | | | 0.82 | 5149 | accuracy | | | 0.64 | 5149 |
| macro avg | 0.84 | 0.72 | 0.75 | 5149 | macro avg | 0.65 | 0.65 | 0.64 | 5149 |
| weighted avg | 0.83 | 0.82 | 0.80 | 5149 | weighted avg | 0.65 | 0.64 | 0.65 | 5149 |

Figura 13. Métricas adicionales H1N1 y Seasonal

K-means + SVM

El objetivo de esta combinación de algoritmos era entender cómo se agrupaban las instancias en clústeres y cómo se distribuían las etiquetas de clase en esos clústeres, buscando patrones en la estructura subyacente de los datos que pudieran ser capturados tanto por SVM como por KMeans.

Se entrenó un modelo SVM con kernel lineal mediante validación cruzada para evaluar la precisión promedio del modelo. Después, aplicó KMeans al conjunto de entrenamiento para agrupar las instancias en dos clústeres. Examinó la distribución de las etiquetas de clase en cada clúster, calculó la frecuencia y frecuencia relativa de cada valor de la variable objetivo dentro de los clústeres y, finalmente, realizó una predicción basada en las etiquetas del clúster utilizando el modelo SVM entrenado.

El accuracy proporcionado por este algoritmo compuesto fue de 64% para H1N1 y para Seasonal fu.

En las figuras 14 y 15 se muestra la matriz de confusión y las métricas adicionales.

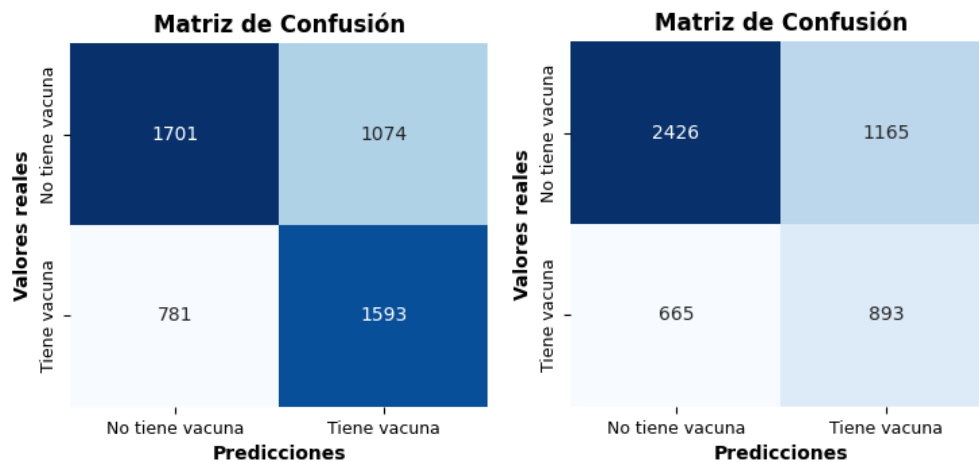


Figura 14. Matrices de confusión H1N1 y Seasonal

| | precision | recall | f1-score | support | | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|--------------|-----------|--------|----------|---------|
| 0 | 0.78 | 0.68 | 0.73 | 3591 | 0 | 0.69 | 0.61 | 0.65 | 2775 |
| 1 | 0.43 | 0.57 | 0.49 | 1558 | 1 | 0.60 | 0.67 | 0.63 | 2374 |
| accuracy | | | 0.64 | 5149 | accuracy | | | 0.64 | 5149 |
| macro avg | 0.61 | 0.62 | 0.61 | 5149 | macro avg | 0.64 | 0.64 | 0.64 | 5149 |
| weighted avg | 0.68 | 0.64 | 0.66 | 5149 | weighted avg | 0.64 | 0.64 | 0.64 | 5149 |

Figura 15. Métricas adicionales H1N1 y Seasonal

Retos y consideraciones de despliegue

En los algoritmos presentados como aquellos con mejor desempeño de cada modelo, no se dio overfitting ya que, aunque el desempeño en el set de entrenamiento fue siempre superior, a la hora de predecir en el set de prueba, su desempeño no decayó por completo. Se observó que en los algoritmos no supervisados la predicción para no vacunados y vacunados en cada caso fue similar, sin embargo, los algoritmos compuestos presentaron una tendencia a predecir mejor los casos negativos (no tienen vacuna) que los positivos para H1N1.

Para mejorar el desempeño obtenido en cada modelo, lo que haría sería:

Modelo de red neuronal:

1. Introducir técnicas de regularización como capas de abandono o regularización L1/L2 para evitar el sobreajuste.
2. Experimentar con diferentes funciones de activación, arquitecturas de red y tasas de aprendizaje para mejorar el rendimiento.

Modelo SVM:

1. Optimizar los hiperparámetros de la SVM utilizando técnicas como la búsqueda en cuadrícula (grid search) o la búsqueda aleatoria (random search).
2. Explorar diferentes funciones de núcleo y parámetros de regularización para mejorar el rendimiento del modelo.

Modelo de agrupación K-means:

1. Experimentar con diferentes números de clusters y evaluar las puntuaciones de silueta para encontrar el número óptimo de clusters.

Consideramos que el desempeño mínimo para desplegar en producción nuestro modelo es que no se presente overfitting durante el entrenamiento y que tenga un accuracy en el set de entrenamiento superior a 82%.

Los modelos pueden desplegarse utilizando herramientas de contenedorización como Docker y desplegarse en plataformas en la nube o en servidores locales. Los procesos de supervisión del rendimiento pueden incluir la detección de desviaciones, la activación del reentrenamiento del modelo y la evaluación periódica del rendimiento del modelo en función de umbrales predefinidos.

Conclusiones

- **Evaluación del Mejor Modelo:** La elección del modelo SVM basada en su destacado desempeño, especialmente con una precisión del 85% para H1N1 y del 74% para la vacuna estacional, sugiere que este modelo puede ser una elección sólida para predicciones futuras.
- **Consideraciones para Puesta en Producción:** Se propone la implementación del modelo SVM específicamente para el caso de H1N1 en un entorno de producción, donde superó las condiciones establecidas previamente. Esto podría contribuir a la toma de decisiones en el ámbito de la salud pública.
- **Recolectar Nuevos Datos:** Dada la naturaleza cambiante de las condiciones de salud pública, se sugiere realizar nuevamente la encuesta en el mismo grupo focal, recogiendo las mismas características del dataset original. Esto permitirá evaluar la capacidad de generalización del modelo en contextos actualizados y específicos de la situación pandémica.
- **Comparación con la Situación de la Pandemia de COVID-19:** La intención de traspolar el modelo a la situación de la pandemia de COVID-19 demuestra una visión estratégica. Realizar encuestas similares en el contexto actual permitirá evaluar la robustez del modelo en escenarios epidemiológicos diferentes.
- **Potencial Referente en Salud Pública:** La propuesta de utilizar el modelo como referente en estudios de salud pública abre posibilidades para la aplicación de la inteligencia artificial en la predicción de la probabilidad de vacunación. Esto puede contribuir a la planificación y ejecución de programas de vacunación más efectivos.
- **Necesidad de Monitoreo Continuo:** Dada la dinámica de las condiciones de salud y la evolución de la pandemia, se recomienda un monitoreo continuo del modelo, ajustándolo según sea necesario para mantener su precisión y relevancia en la toma de decisiones.
- **Perspectivas Futuras:** La experiencia y lecciones aprendidas de este proyecto sientan las bases para futuros estudios y aplicaciones de modelos predictivos en el campo de la salud pública. La adaptación constante a nuevas circunstancias permitirá que el modelo sea una herramienta valiosa para la toma de decisiones informada.