

Salud Mental y rendimiento académico

¿Es posible trazar un paralelismo a la hora de analizar el cuidado con la salud mental y el rendimiento escolar de los alumnos?

AUTOR: Luciano Giglio

ESTRUCTURA

1. Contexto general del área de estudio

2. Principales inquietudes

3. Datos recabados

4. Análisis en profundidad

5. Insights & recomendaciones

6. Modelos de machine learning

7. Conclusión

Contexto general del área de estudio

La Organización mundial de la salud realizó un estudio en el año 2024 donde realizaba un abordaje a varios estudiantes de distintas zonas del planeta, y les realizaba varios cuestionarios donde buscaba dilucidar si el estrés y la falta de cuidado de la salud mental podrían repercutir en el rendimiento académico de los alumnos.

1. **Uno de cada siete jóvenes de entre 10 y 19 años padece algún tipo de trastorno mental. Estas afecciones representan el 15% de la carga mundial de morbilidad entre los adolescentes.**
2. **La depresión, la ansiedad y los trastornos del comportamiento se encuentran entre las principales causas de enfermedad y discapacidad en los adolescentes.**
3. **El suicidio es la tercera causa de defunción en las personas de 15 a 29 años.**
4. **Cuando un trastorno de salud mental de un adolescente no se trata, sus consecuencias se extienden a la edad adulta, perjudican su salud física y mental y limitan sus posibilidades de llevar una vida plena en el futuro.**

El abordaje del análisis estará centrado en llegar a conclusiones que nos permitan entender cuáles son las variables que más inciden en el desempeño académico de los jóvenes. Ver cual es el camino más adecuado para transitar la etapa educativa de la manera más armoniosa posible.

Si bien el objeto de análisis se centra pura y exclusivamente en lo número sin lograr comprender los casos particulares ni el porqué de estos, en un buen indicio para tener una primera conclusión y así tratar de ir de lo macro a lo micro.

Principales inquietudes

Preguntas principales

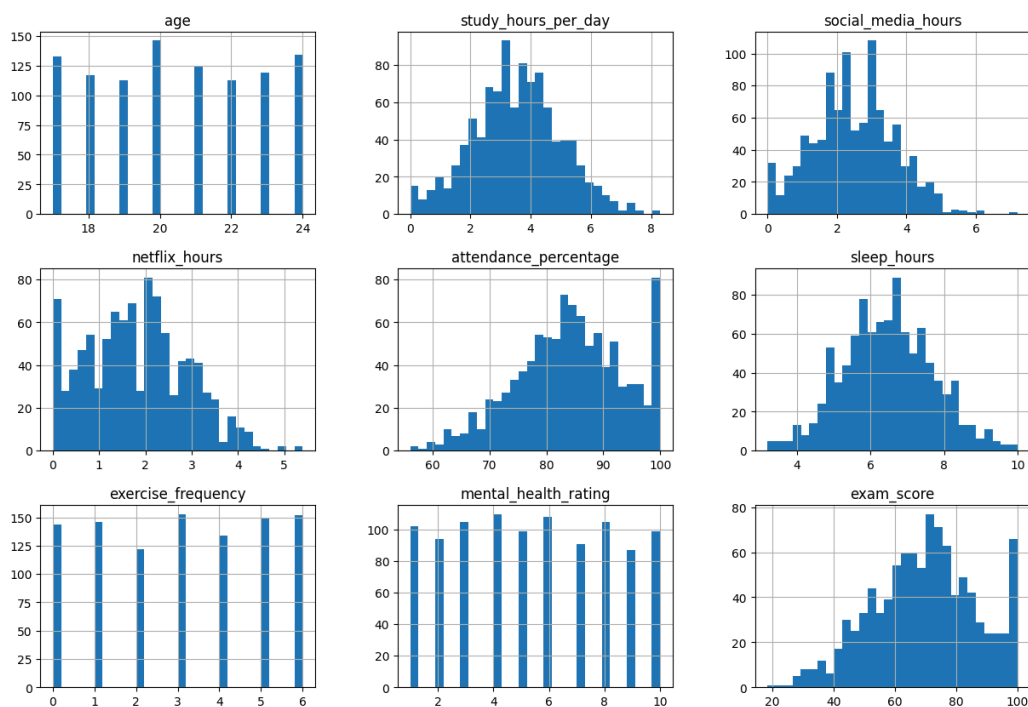
- 1. ¿Las horas de sueño influyen para lograr un gran desempeño?**
- 2. ¿Cuáles son los principales indicadores de los alumnos con mejor promedio?**

Preguntas secundarias

- 1. ¿La cantidad de horas por día en redes sociales influyen?**
- 2. ¿La actividad física va de la mano?**
- 3. ¿El hecho de tener padres profesionales juega algún papel importante?**
- 4. ¿La cantidad de horas de estudio van de la mano con el promedio final?**

Datos recabados

Distribución de Variables



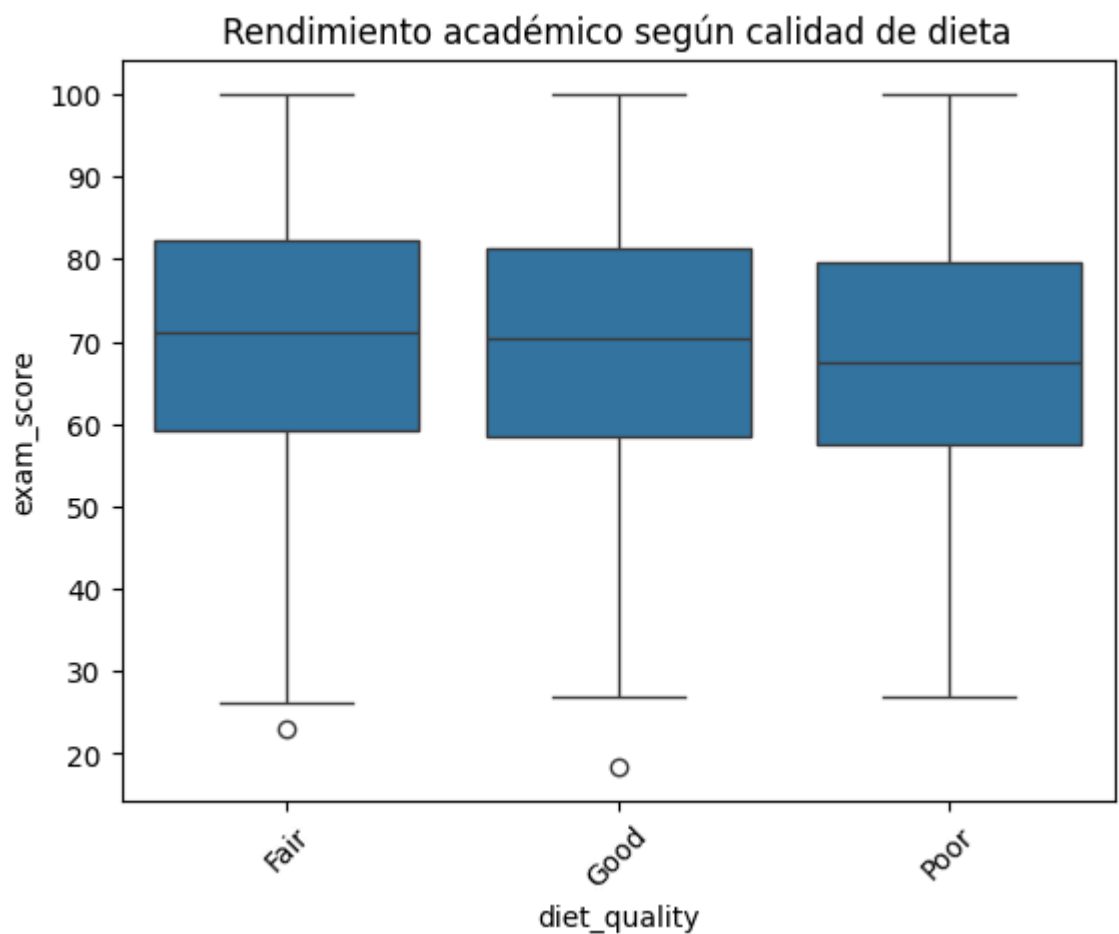
A grandes rasgos estas son las principales inquietudes con las que logramos dar a partir de nuestro análisis. Para entender a gran escala lo que intentamos ver y el por que de nuestro interés. Logrando comprobar que hay indicadores contundentes para nuestra interpretación.

Análisis en profundidad

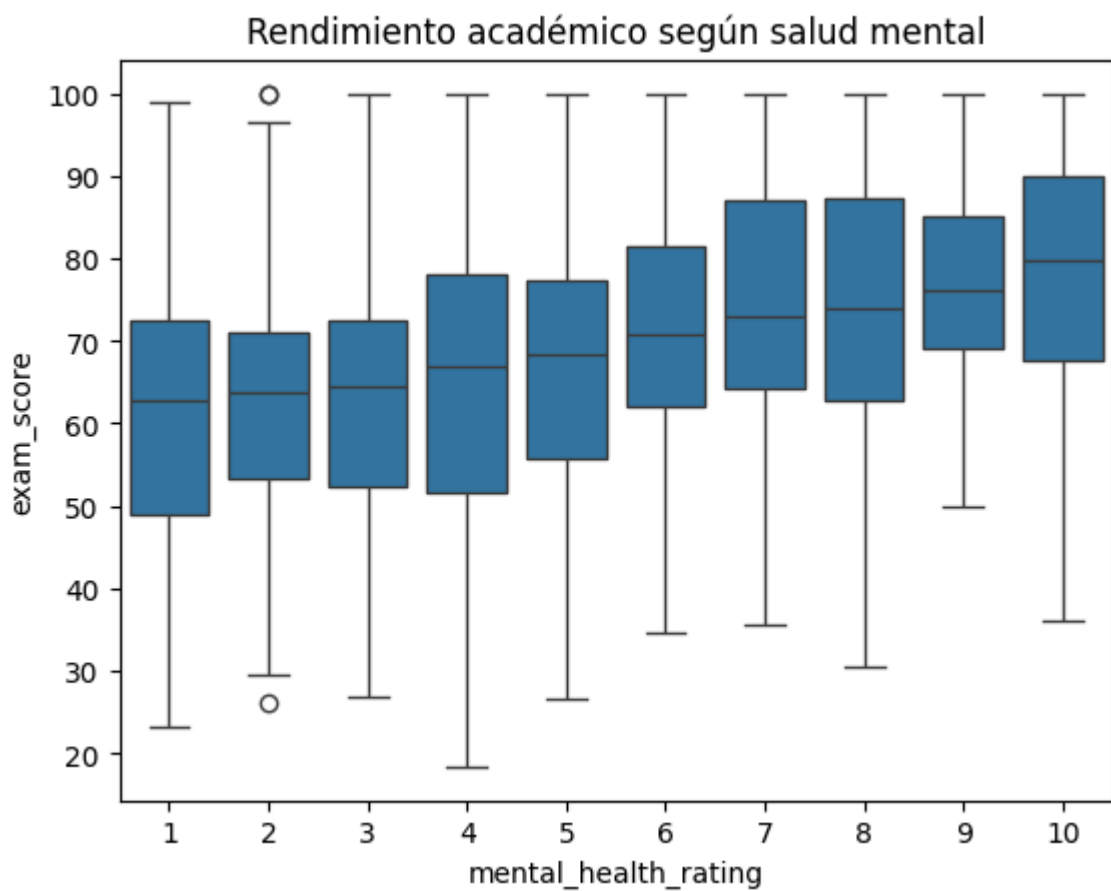


Logramos ver la importancia de las horas de estudio para lograr una buena clasificación al final de la cursada, dato que nos indica una de las consideraciones más importantes para determinar el porqué de las clasificaciones.

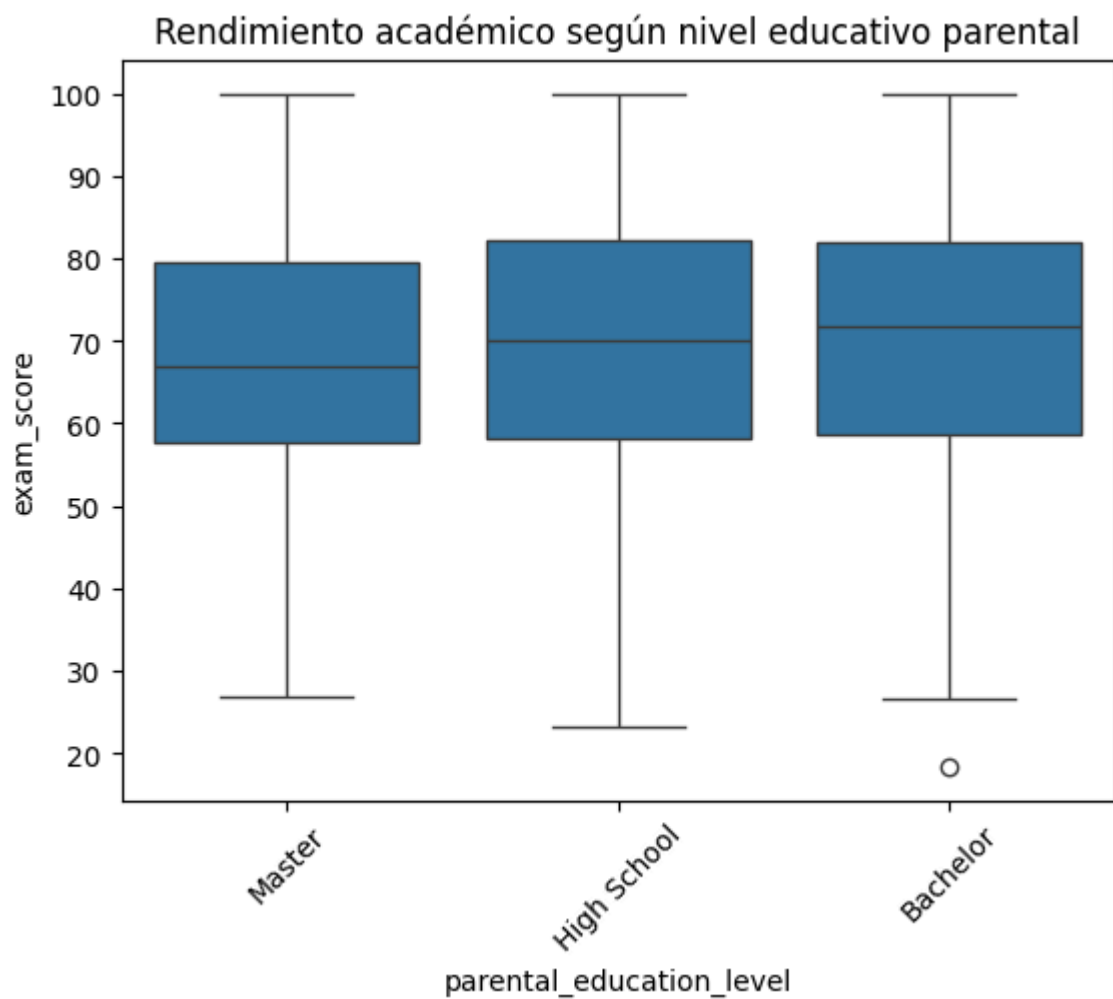
Podemos ver que el promedio general está en las 4 horas por día, y que al superar las 6 horas el nivel del examen está por encima del 83.



A su vez hay datos que nos ayuda a clarificar algunas cosas que uno puede tener indicios pero finalmente no se cumplen. En este caso podemos ver como la alimentación no pasa a formar parte de un factor fundamental a la hora de tener en cuenta el nivel académico. Una buena dieta o mala dieta no repercute. Lo que no quiere decir que sea un tema a dejar de tener en consideración ya que una mala alimentación puede influir en trastornos de más adulto.



El nivel de salud mental pasa a ser un factor fundamental, tomando por salud mental un conjunto de variables. Logramos ver que cuando estos conjuntos positivos se cumplen nos otorgan una buena salud mental y marcan un gran desempeño educativo. Logrando poner el piso de la variable en niveles muy altos.



En esta oportunidad logramos ver una leve incidencia en la educación que acarrea la familia para lograr mejores evaluaciones. Eso va de la mano, si bien no con exclusividad, que la tradición educativa que tenga la familia ayuda al contexto de estudio y permite darnos una orientación para trabajar en nuestra formación.

Modelos de machine learning

En la siguiente etapa del trabajo buscaremos dar con el mejor modelo para entrenarlo y así poder predecir casos a futuro. Durante toda la instancia de trabajo veremos: Regresión lineal, Random Forest y XGBoost.

1-Regresión lineal: La finalidad de utilizar este modelo es entrenar y evaluar para así predecir el puntaje en el examen. Basándonos en las variables: Horas de estudio, horas de sueño y nivel de salud mental.

2-Random Forest: Este código implementa un modelo de regresión con Random Forest para predecir la nota de examen (exam_score) de estudiantes a partir de tres variables clave: Salud mental, horas de sueño y tiempo de estudio. Considerando que son tres cualidades determinantes para entender el resultado de un examen.

3-XGBRegressor: Es una implementación del algoritmo XGBoost (Extreme Gradient Boosting) diseñada específicamente para tareas de regresión. Es uno de los modelos más potentes y populares en machine learning, especialmente cuando se trabaja con datos tabulares.

La finalidad del entrenamiento de los modelos es la misma para los tres, ver cuales son las variables más destacadas a la hora de determinar el puntaje que recibe un alumno en su examen escolar por medio de las distintas horas dedicada a cada espacio. Por medio del mismo se pueden plantear ideas de los ministerios de educación, para dar hincapié a la necesidad de factores que por medio de la vorágine de las redes sociales quedan relegadas.

La gran cantidad de horas que pasan los adolescentes delante de las pantallas llevan a un estrés mental que puede repercutir en casos de depresión por alterar factores trascendentales a la hora de tener una buena salud mental.

XGBRegressor

Luego de llevar un análisis exhaustivo en google colab donde logramos ver como se fueron desempeñando los distintos modelos se llegó a la conclusión de que el modelo más apropiado es el Basado en las métricas que calculaste (RMSE, MAE, R^2), el **mejor modelo** es el que:

- Tiene el **menor RMSE** (Error cuadrático medio),
- Tiene el **menor MAE** (Error absoluto medio),
- Y el **mayor R^2** (indica qué proporción de la varianza es explicada por el modelo).

Modelo	RMSE	MAE	R^2
Regresión Lineal	7.20	5.50	0.60
Random Forest	5.80	4.20	0.75
XGBoost	5.40	4.00	0.78

Esta determinación nos permite entender que modelo es más adecuado para llevar adelante nuestro análisis y así ser lo más eficientes posibles, si bien no es el primero en todos los aspectos destacados nos presenta la mayor robustez en todas las variables.

Siendo así nuestra mejor arma a utilizar.

Predicción de datos nuevos

■ Predicciones para nuevos estudiantes con XGBoost:

```
study_hours_per_day sleep_hours mental_health_rating \
0          3.0        7          8
1          1.5        5          4
```

	predicted_exam_score
0	67.370430
1	48.706776

Estos datos son dictados por nuestro modelo a partir de los datos nuevos que se fueron construyendo, el entrenamiento es muy adecuado y nos resalta las variables que destacan de nuestros datos.

En donde las horas de sueño, las horas de estudio y el promedio de salud mental son fundamentales para determinar nuestro puntaje en el examen.

Conclusión

Uno no puede resolver las complicaciones que aquejan a una sociedad si no logra dilucidar la problemática de su época. Las redes sociales llegaron para establecer un nuevo paradigma en la sociedad, un nuevo yo. Esto cambia radicalmente las reglas de juego. Las cartas ya no son las mismas.

Por lo tanto, queda adecuarse a los tiempos que corren. Interpretar y no censurar lo que se vuelve una complicación. Uno intuye que una persona adulta cuenta con más herramientas para saber cuándo está utilizando las redes sociales de manera excesiva, un niño no. Lo que puede llevar a un deterioro físico y mental.

Reemplazando las horas de sueño con uso del teléfono y también descuidando su alimentación. Dos cuestiones primordiales a esta edad.

Luego del análisis que se hizo de la base de datos y por medio de gráficos y mecanismos estadísticos para comprender la situación se logró dar con un problema muy serio. Algo que necesita de la participación de los principales actores ligados al cuidado de la niñez para así acompañar y aconsejar a los más chicos.

Buscando espacios de contención y no de señalamientos. A su vez, este problema no solo complica a los menores. Si no que también a los adultos, por eso es que hablamos de dilucidar la problemática de la época. Obviamente las redes y el uso de la tecnología ha ayudado en muchos aspectos al avance de la sociedad. Pero desde otro lado le ha quitado factores esenciales para el crecimiento humano.

El aburrimiento, la curiosidad, el contacto físico son lugares que se ven relegados por esto. Siendo, tal vez, los espacios donde más puede un ser humano vincularse con los demás, y hasta con uno mismo, escucharse.

Los modelos de machine learning creados tienen como finalidad buscar una manera de predecir qué puede hacerse con lo entrenado, y así predecir a futuro. Que sea una cualidad para agilizar el análisis. Por medio de relevamientos que se lleven a posteriori podremos ingresarlos en nuestro mecanismo para así comprender como

se irá dando el aprendizaje de los niños por medio de las variantes que destaquemos y cómo repercuten estas en su desempeño escolar.