

Laboratorio 4 – Inteligencia de Negocios

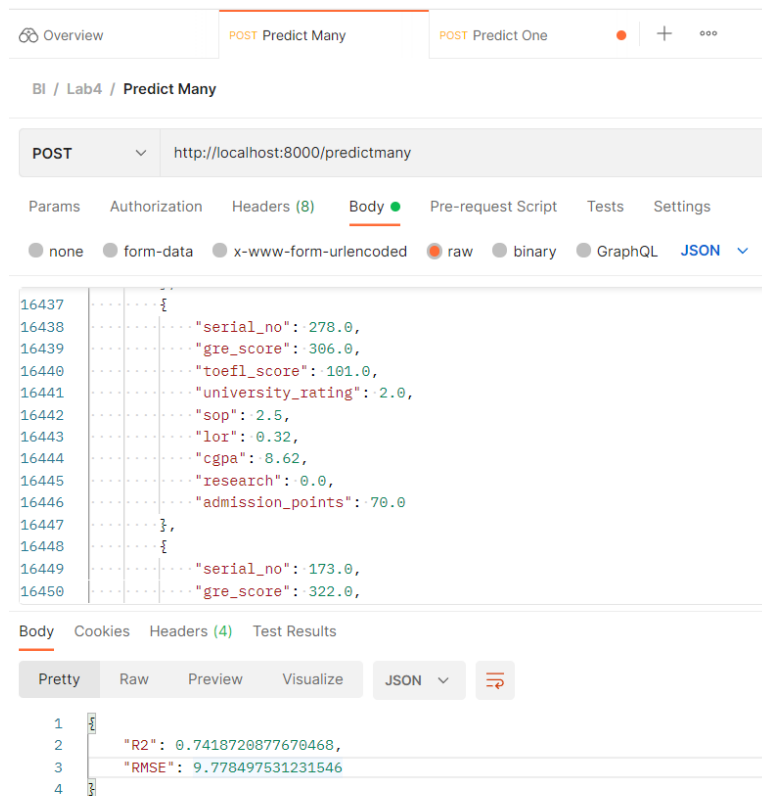
Grupo 7

- Kevin Cohen Solano – 202011864
- Juan Carlos Marín Morales – 202013973
- Juan Felipe Castro Vanegas – 201818130

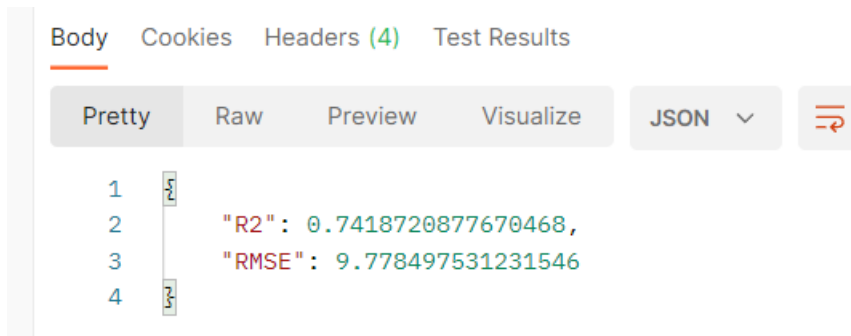
Escenarios de prueba:

Para la realización de este laboratorio implementamos una API utilizando FastAPI, que nos permitió realizar peticiones HTTP por medio de Postman y así consumir nuestro Pipeline construido en el laboratorio 3.

Realizamos el entrenamiento del modelo a través del pipeline con la totalidad de los datos, aunque se puede realizar con diferentes volúmenes de datos (que se encuentran en la carpeta requestBody/retrain) mediante el archivo “Predict Many” como se muestra a continuación.

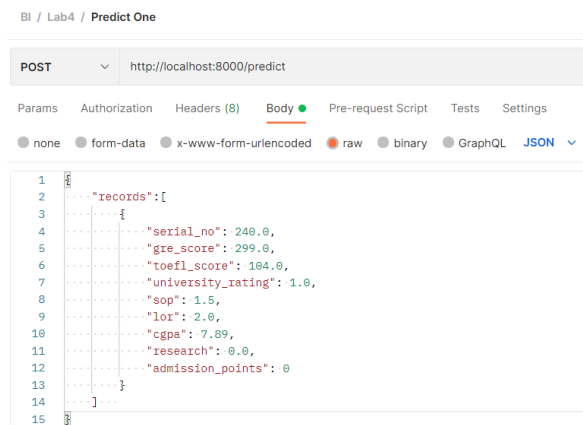


Las métricas obtenidas fueron las siguientes:



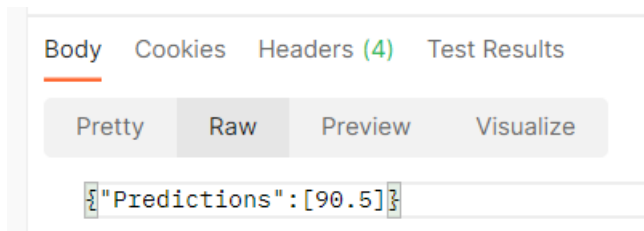
Recordamos que R2 nos indica que tan eficaz es el modelo a la hora de predecir los resultados, mientras que RMSE nos indica el error promedio que se espera de cada una de las predicciones, siendo una predicción buena cuando su error es menor que el promedio.

Realizamos 5 escenarios de prueba, de los cuales 3 de ellos fueron casos exitosos con métricas coherentes para el negocio y 2 pruebas erróneas con resultados incoherentes. A continuación, observaremos los resultados de cada una, junto con su respectivo JSON con los datos utilizados en cada prueba, así como la imagen de su respectiva prueba en Postman (adjuntamos una imagen de referencia de cómo se deben ver los JASON en Postman [en el archivo “Predict One”], igualmente se pueden colocar varios de estos elementos en un mismo JASON separándolos por comas para obtener diferentes resultados en una misma ejecución):



Escenario 1:

-Imagen prueba de postman:



-Script JSON:

```

"serial_no": 337.0,
"gre_score": 331.0,
"toefl_score": 74.0,
"university_rating": 4.0,
"sop": 4.92,
"lor": 0.45,
"cgpa": 9.44,
"research": 1.0,
"admission_points": 0

```

-Resultado esperado:

```

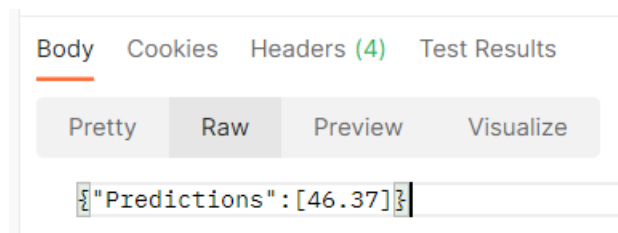
"admission_points": 88.16

```

El resultado se considera coherente ya que si observamos nuestro RSME solo se desvía 2 unidades con respecto a la predicción de 90 dada por el modelo.

Escenario 2:

-Imagen prueba de postman:



-Script JSON:

```

"serial_no": 240.0,
"gre_score": 299.0,
"toefl_score": 104.0,
"university_rating": 1.0,
"sop": 1.5,
"lor": 2.0,
"cgpa": 7.89,
"research": 0.0,
"admission_points": 0

```

-Resultado esperado:

```

"admission_points": 59.0

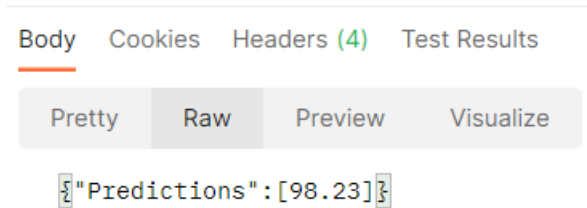
```

El resultado se considera incoherente porque la predicción dada por el modelo sobrepasa por 3 unidades el error esperado. Si observamos los valores de entrada, la mayoría están dentro de los valores esperados así que creemos que este resultado depende de las

variables que no tuvimos en cuenta para nuestro modelo pesaron para la decisión lo cual generó un resultado atípico.

Escenario 3:

-Imagen prueba de postman:



-Script JSON:

```
"serial_no": 424.0,  
"gre_score": 334.0,  
"toefl_score": 75.0,  
"university_rating": 5.0,  
"sop": 4.5,  
"lor": 5.0,  
"cgpa": 9.54,  
"research": 1.0,  
"admission_points": 0
```

-Resultado esperado:

```
"admission_points": 94.0
```

El resultado se considera coherente gracias a que la predicción del modelo está bastante cerca del valor esperado y no supera el valor del error.

Escenario 4:

-Imagen prueba de postman:



-Script JSON:

```
"serial_no": 76.0,  
"gre_score": 329.0,  
"toefl_score": 114.0,
```

```
"university_rating": 2.0,  
"sop": 2.0,  
"lor": 4.0,  
"cgpa": 8.56,  
"research": 1.0,  
"admission_points": 0
```

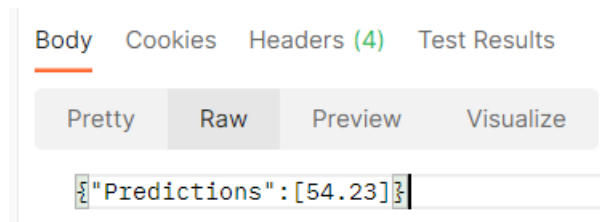
-Resultado esperado:

```
"admission_points": 72.0
```

El resultado se considera bastante coherente, ya que fue la prueba que más se acercó entre los valores de la predicción y valor esperado. Esto posiblemente debido a que los valores se encuentran bastante centrados a su respectivo valor esperado.

Escenario 5:

-Imagen prueba de postman:



-Script JSON:

```
"serial_no": 316.0,  
"gre_score": 308.0,  
"toefl_score": 104.0,  
"university_rating": 2.0,  
"sop": 1.6,  
"lor": 3.0,  
"cgpa": 8.07,  
"research": 0.0,  
"admission_points": 0
```

-Resultado esperado:

```
"admission_points": 65.0
```

El resultado se considera incoherente porque fue la prueba que más se alejaron los valores de la predicción y el valor esperado, esto al igual que en la otra prueba incoherente pudo haber sido por las variables que no tuvimos en cuenta o también por un SOP muy bajo.

Finalmente, para mitigar las incoherencias en las predicciones, podemos evaluar las variables que escogemos revisando más minuciosamente los valores que tenemos en cada columna, para poder arreglarlas o eliminarlas. También podríamos ver que variables que no escogimos pueden ser más influyentes para añadirlas a un nuevo procesamiento. En general los resultados del modelo son aceptables, pero siempre lo mejor para incrementar las métricas de un modelo es conseguir muchos más datos, algo que es más trabajoso pero muy efectivo.