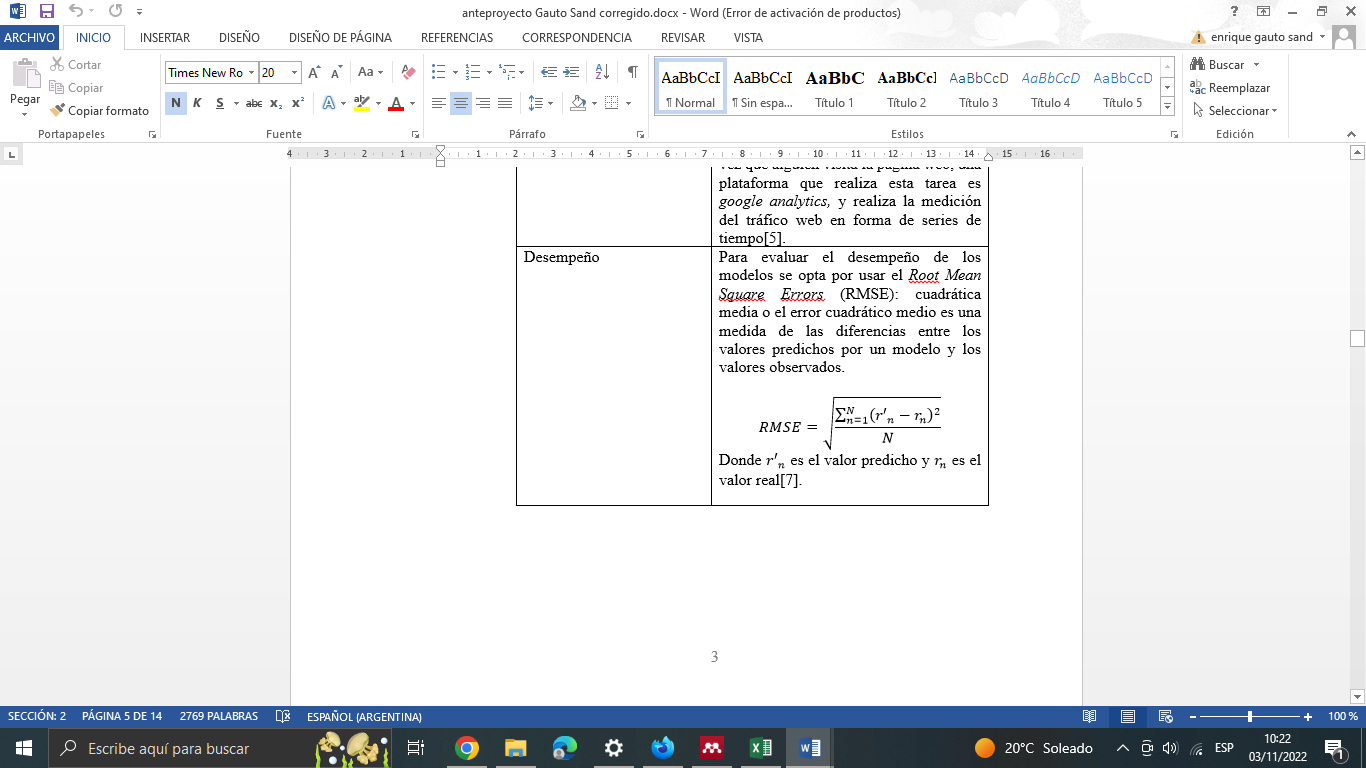
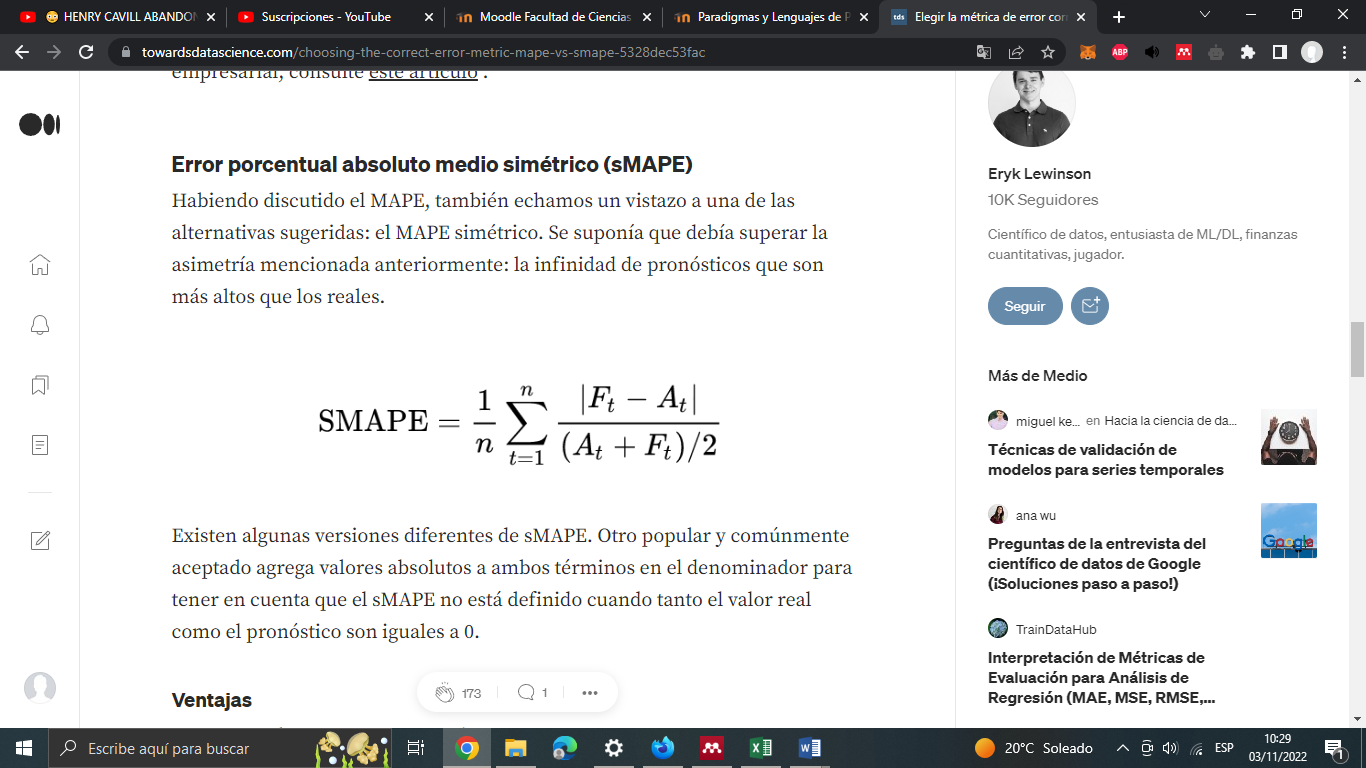
# Columnas Dataframe

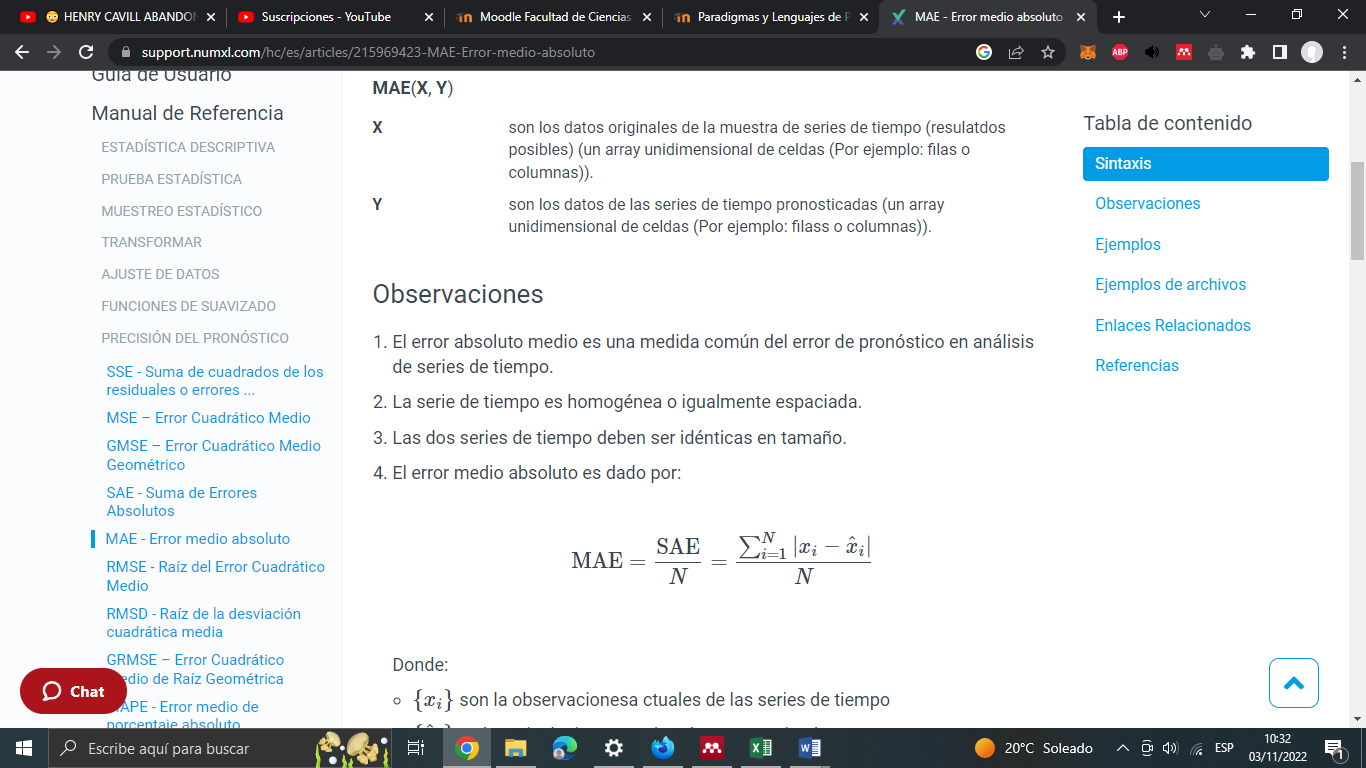
|  |  |  |
| --- | --- | --- |
| Dimensión | Columna | Significado |
|  | Predictor Flag | Predictor Flag |
|  | mes | Agrupación de vistas de 1 mes |
|  | correlacion anual |  |
|  | correlacion Trimestral |  |
|  | Fecha | Fecha |
| Totales | Número de vistas de página |  |
| Usuarios nuevos |  |
| Usuarios |  |
| Sesiones |  |
| Argentina | Vistas de páginas de argentina |
| Otros Paises | Vistas de páginas otros países |
| Categoría de dispositivo | mobile | Vistas de páginas mobile |
| desktop | Vistas de páginas desktop |
| tablet | Vistas de páginas tablet |
|  | Categoría de dispositivo promedio | Promedio de vistas mobile,desktop y tablet |
|  | Categoría de dispositivo std | Desvio estándar de vistas mobile,desktop y tablet |
| Sistema operativo | Windows | Vistas de páginas del sistema operativo Windows |
| Android | Vistas de páginas del sistema operativo Android |
| Linux | Vistas de páginas del sistema operativo Linux |
| Macintosh | Vistas de páginas del sistema operativo Macintosh |
| Sistema operativo (not set) | Vistas de páginas del sistema operativo not set (no encontrado) |
| iOS | Vistas de páginas del sistema operativo iOS |
| Windows Phone | Vistas de páginas del sistema operativo Windows Phone |
| Chrome OS | Vistas de páginas del sistema operativo Chrome OS |
| Tizen | Vistas de páginas del sistema operativo Tizen |
| BlackBerry | Vistas de páginas del sistema operativo BlackBerry |
|  | Sistema operativo promedio | Promedio de vistas de entre los distinos sistemas operativos |
|  | Sistema operativo std | Desvió estándar de vistas de entre los distintos sistemas operativos |
| Agrupación de canales predeterminada | Organic Search | Vistas de páginas agrupadas a través del canal Organic Search, es decir búsqueda orgánica |
| Social | Vistas de páginas agrupadas a través del canal Social |
| Direct | Vistas de páginas agrupadas a través del canal Direct es decir directo a través de un link |
| Referral | Vistas de páginas agrupadas a través del canal Referral, es decir referido |
|  | Agrupación de canales predeterminada promedio | Promedio de Vistas de páginas agrupadas a través de canales |
|  | Agrupación de canales predeterminada std | Desvió estándar de Vistas de páginas agrupadas a través de canales |

Medidas de desempeño del modelo





MAE - Error medio absoluto



# Métricas

|  |  |  |
| --- | --- | --- |
| Métrica | Significado | En paper |
| MAE – Mean Absolute Error  Error medio absoluto | Todos los errores se ponderan en la misma escala ya que se toman valores absolutos. Es útil si los datos de entrenamiento tienen valores atípicos, ya que MAE no penaliza los errores elevados causados por los valores atípicos | si |
| SMAPE - error porcentual absoluto medio simétrico | The symmetric mean absolute percentage error (SMAPE) is an accuracy measure based on percentage (or relative) errors.  Relative error is the absolute error divided by the magnitude of the exact value.  In contrast to the mean absolute percentage error, SMAPE has both a lower bound and an upper bound. Since it's percentage-based, it's scale-independent, which means that it can be used to compare forecast performances between datasets.  A limitation to SMAPE is that if the actual value or forecast value is 0, the value of error will approach 100%. The lower the SMAPE value of a forecast, the higher its accuracy  We have A = 100 and F = 120. The sMAPE is 18.2%. Now a very similar case, in which we have A = 100 and F = 80. Here we come out with the sMAPE of 22.2%. | si |
| RMSE - The root-mean-square error | RMSE uses the squared value of the forecast errors. This helps you identify the impact of outliers. RMSE is a good metric to use for use cases where a few incorrect predictions can be very costly if you implement them.  A forecast model with a lower RMSE value indicates more accurate predictions. |  |
| MAPE - Symmetric mean absolute percentage error | It's an accuracy measure based on the relative percentage of errors. The closer the MAPE value is to zero, the better the predictions.  If the error is greater than the actual value, then the percentage of error can be more than 100%.  For example, if the actual value is 1 and we predicted 3, then that makes the forecast error a 2. The forecast error is greater than the actual value, so the MAPE result is greater than 100%.  As MAPE specifies the size of the error as a percentage rather than actual values, it can compare the forecast error between different data set sizes and different time scales.. |  |
| MSE - Error Cuadrático Medio  Mean Square Error | Representa la distancia al cuadrado entre los valores reales y predichos.  Para errores pequeños, converge a los mínimos de manera eficiente. No hay mínimos locales.  MSE penaliza el modelo por tener grandes errores al elevarlos al cuadrado.  Es particularmente útil para eliminar valores atípicos con grandes errores del modelo al ponerles más peso. |  |

Nose como interpretar las métricas solo sé que mientras más cercanas a cero mejor rendimiento, ahora si es x cantidad de porcentaje o x cantidad de vistas o que es lo que quiere decir la métrica eso no lo sé, aun no lo confirmo.

SMAPE en particular hay distintas fórmulas.

Hiperparámetros:

Se menciona en alguna parte el uso de “SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization”

Lista de hiperparametros en hparams.py

*# Manually selected params*

|  |  |  |
| --- | --- | --- |
| Nombre | Valor por defecto | Descripcion |
| batch\_size | 256 | Tamaño del lote |
| train\_window | 283 |  |
| train\_skip\_first | 0 |  |
| rnn\_depth | 267 | Unidades de la red neuronal GRU por layer |
| use\_attn | False |  |
| attention\_depth | 64 |  |
| attention\_heads | 1 |  |
| encoder\_readout\_dropout | 0.4768781146510798 |  |
| encoder\_rnn\_layers | 1 | Numero Capa de encoder de red neuronal |
| decoder\_rnn\_layers | 1 | Numero Capa de decoder de red neuronal |
| decoder\_input\_dropout | [1.0, 1.0, 1.0] |  |
| decoder\_output\_dropout | [0.975, 1.0, 1.0] |  |
| decoder\_state\_dropout | [0.99, 0.995, 0.995] |  |
| decoder\_variational\_dropout | [**False**, **False**, **False**] |  |
| decoder\_candidate\_l2 | 0.0 |  |
| decoder\_gates\_l2 | 0.0 |  |
| fingerprint\_fc\_dropout | 0.8232342370695286 |  |
| gate\_dropout | 0.9967589439360334 |  |
| gate\_activation | **'none'** |  |
| encoder\_dropout | 0.030490422531402273 |  |
| encoder\_stability\_loss | 0.0 | No aplicar esta funcion |
| encoder\_activation\_loss | 1e-06 |  |
| decoder\_stability\_loss | 0.0 |  |
| decoder\_activation\_loss | 5e-06 |  |
|  |  |  |

# Hiperparams tuner keras

<https://keras.io/guides/keras_tuner/getting_started/>

RandomSearch <https://keras.io/api/keras_tuner/tuners/random/>

Hyperband <https://keras.io/api/keras_tuner/tuners/hyperband/>

Este si hay video en youtube de cómo hacer.

BayesianOptimization <https://keras.io/api/keras_tuner/tuners/bayesian/>

# Que se puede tunear?

hp\_units = hp.Int(‘units\_hp’, min\_value = 32, max\_value = 128, step = 32)

is used to set the range of a hyperparameter whose values are integers, like for ‘number of filters’ in Convolutional Neural Networks and ‘number of units’ in Dense layer/Fully Connected layer.

hp.Float(‘dropout’, 0, 0.5, step=0.1)

Para variaciones flotantes como una layer dropout

hp.Choice() offers us way to provide specific values (integers or decimal) and functions to be used for hyperparameter tuning.

Example: hp.Choice(‘learning\_rate’, values = [1e-2, 1e-3, 1e-4])

Example: hp.Choice(‘pooling’ , [‘avg’, ‘max’])

# Hyperparametros que se optimizaron

|  |  |  |
| --- | --- | --- |
| Nombre | Valor por defecto | Description |
| units | - | Número de unidades de GRU |
| Dropout | - | Float between 0 and 1. Fraction of the units to drop for the linear transformation of the inputs. Default: 0. |
| Recurrent dropout | - | Float between 0 and 1. Fraction of the units to drop for the linear transformation of the recurrent state. Default: 0. |
| Batch size | - | The batch size defines the number of samples that will be propagated through the network.  For instance, let's say you have 1050 training samples and you want to set up a batch\_size equal to 100. The algorithm takes the first 100 samples (from 1st to 100th) from the training dataset and trains the network. Next, it takes the second 100 samples (from 101st to 200th) and trains the network again. |
| Learning rate | - | Ratio de aprendizaje, La tasa de aprendizaje afecta a la velocidad a la que el algoritmo alcanza (se converge en) las ponderaciones óptimas. |
| N\_days | 7-17-27-37 | Optimizado a mano, cuantos días atrás hay que tomar de, es un parámetro de entrada. |
|  |  |  |
|  |  |  |

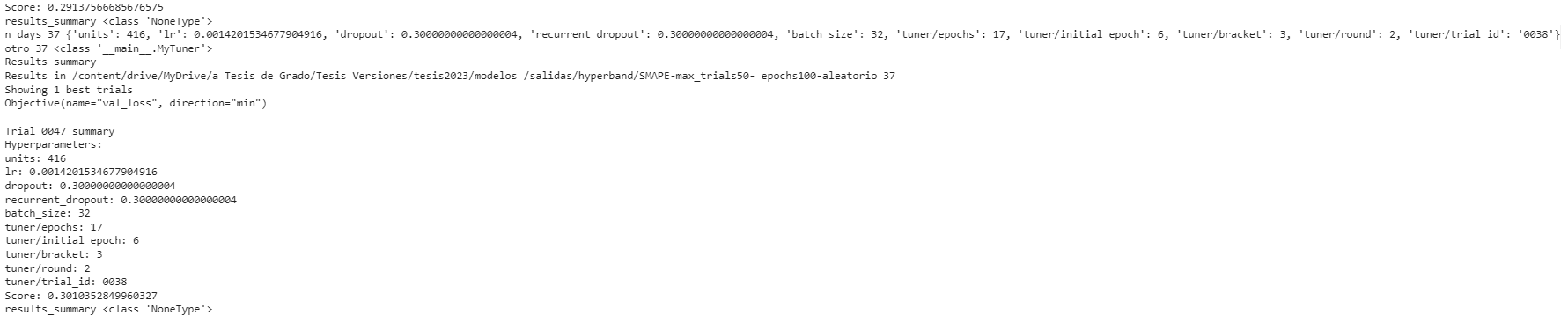
Buscar características de equipos de google colab

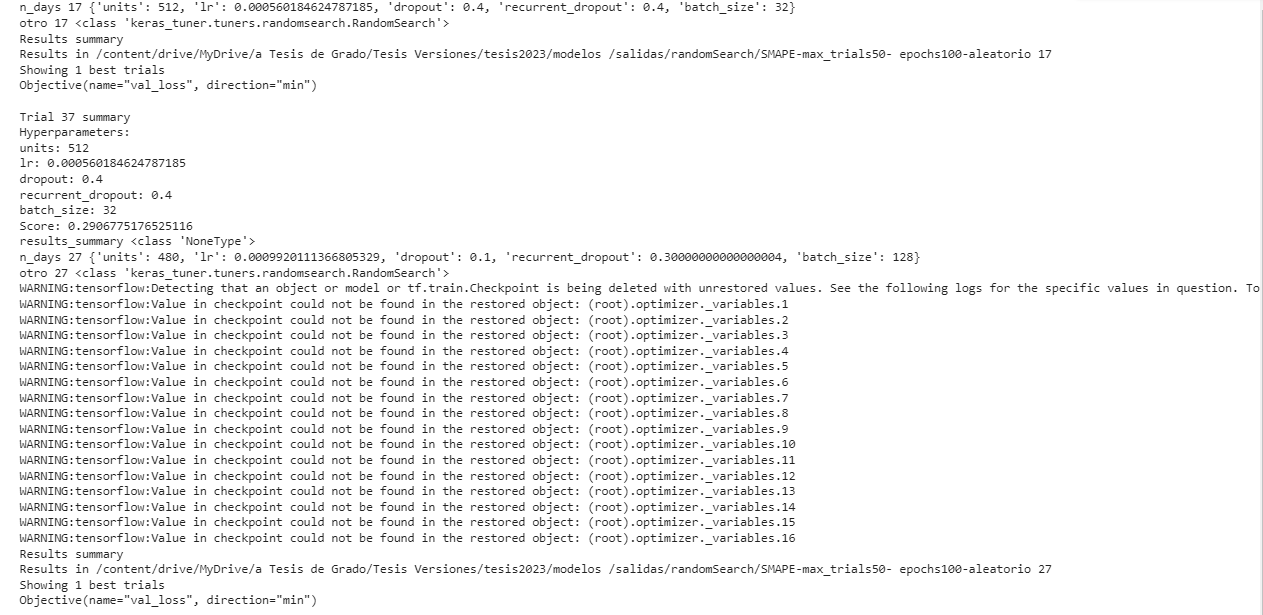
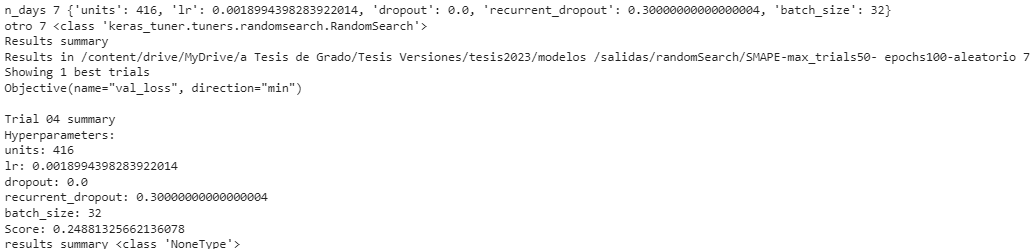
Resultados entrenamiento con tuner hyperparametross con hyperband

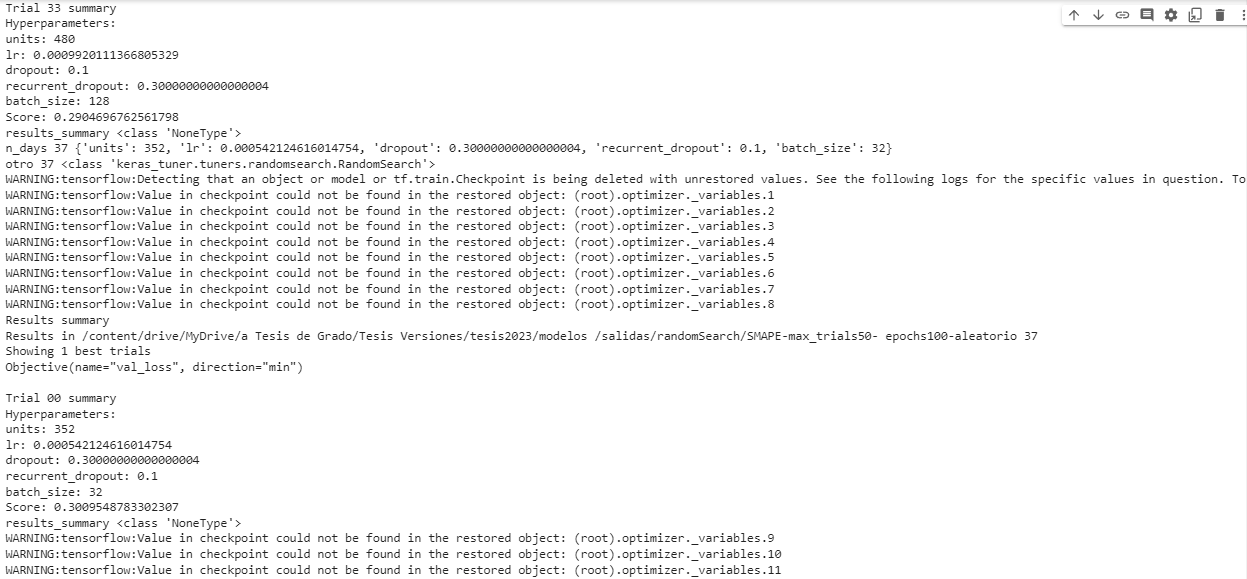
Se intentó entrenar sin tunear el dropout pero no se vio mejoras por lo que procedió a entrenar tuneando el dropout







Resultados entrenamiento con tuner hyperparametross con randomsearch Ndivisiones predictor flag 30



Cambiando el rango de divisiones del predictor flag de 30 divisiones a la mitad 15



0.0834= 8.34% de error

Más del 90% de acierto, todavía no lo probé con los hyperparametros tunning

Falta probar:

El algoritmo de hyperparameter tunning bayesian

Pronosticar más columnas en vez de solo el predictor flag

Con datos en semanal(es en caso de que fallen los otros)

El tema de las estadísticas

Definir que es cada cosa

Arreglar la métrica de medición

Otros temas

Lunes que viene 10-07 cirugía de extracción de muela de juicio

A partir de que día comienza vacaciones en la universidad?

Cuando siguiente reunión?

Jueves 3 o viernes 4 de agosto

Fuentes

<https://stats.stackexchange.com/questions/153531/what-is-batch-size-in-neural-network>

<https://aws.amazon.com/es/what-is/hyperparameter-tuning/>

<https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/training-parameters1.html>