

Documentación del parcial

Fecha: 23/05/2022

Autor: Pedro Luis Cárdenas Díaz

En el parcial de tercer corte de la materia HPC para IA del grupo 01, se busca la realización de 2 proyectos diferentes, uno realizado en Python por medio de Colab, y el otro realizado a través de QtCreator a través de C++, ambos proyectos tienen como punto centrar la ejecución de un modelo de regresión lineal por medio de un dataset dado a nosotros por el docente.

Observaciones:

A la hora de realizar el parcial se utilizó trabajos realizados en lo que respecta a el primer semestre académico del 2022 de la materia HPC para IA.

Cambios realizados en el proyecto:

Python:

- En Python debido a un problema a la hora de realizar la parte de C++ se imposibilitaba la impresión de las gráficas de comparación entre Python y C++, por ende, se tuvo que realizar un cambio a la hora de buscar por los datos a graficar, limitando los datos de 0 a 20 para así poder imprimir el grafico.

C++:

- En C++ debido a que la primera fila esta escrita en letras y no en números, esta fila se encontraba llena de 0, dificultando así la correcta medición e datos, para dar solución a este, se procedió a dirigirnos a la carpeta donde se encontraba el dataset y modificar este de forma que la primera fila sea la segunda.

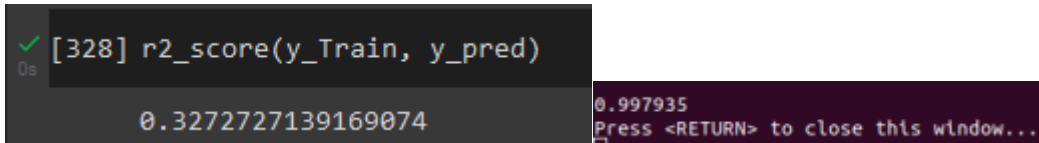
Métrica de rendimiento utilizada:

R2 Score:

- R2 Score es una métrica de rendimiento la cual revela el porcentaje de correlación existente entre 2 variables, siendo 1/100% que estas variables están perfectamente relacionadas y 0/0% siendo que estas no tienen relación alguna, estos valores también pueden ser negativos.

Conclusiones:

- Los valores a la hora de comparar la métrica de rendimiento no son los mismos, mientras la métrica realizada por C++ produce una correlación del 99.7%, la métrica realizada por Python produce una correlación del 32% significando que hay un problema con lo realizado en C++ o en Python.



```
[328] r2_score(y_Train, y_pred)
0.3272727139169074
0.997935
Press <RETURN> to close this window...
```

Referencia:

[https://es.wikipedia.org/wiki/Normalizaci%C3%B3n_\(estad%C3%ADstica\)](https://es.wikipedia.org/wiki/Normalizaci%C3%B3n_(estad%C3%ADstica))

https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

<https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.normalize.html>

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

<https://www.bmc.com/blogs/mean-squared-error-r2-and-variance-in-regression-analysis/>

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.r2_score.html