**Aplicación de ML + Informe final individual**

**Tarea a automatizar:**

Clasificación de rendimiento y apropiación en la utilización de la plataforma E-ducativa de los estudiantes ingresantes a las carreras de Formación Docente en el periodo 2020-2021 de Instituto de Educación Superior de Charata.

**Objetivos**

* Comprender los procesos de aplicación de Machine Learning.
* Generar un marco de recursos introductorios para la utilización de ML.
* Generar, depurar y utilizar datos de contexto educativos para la utilización en ML.
* Probar algoritmos de clasificación con los datos obtenidos.

**Metodología de desarrollo:**

Enfoque Supervisado

Algoritmo de Clasificación – Naives Bayes

Métrica de desempeño a utilizar *matriz de confusión* (se encuentra circunscripto a posibles cambios debido a la existencia de un proceso de aprendizaje)

**Herramientas (incluyendo los datos)**

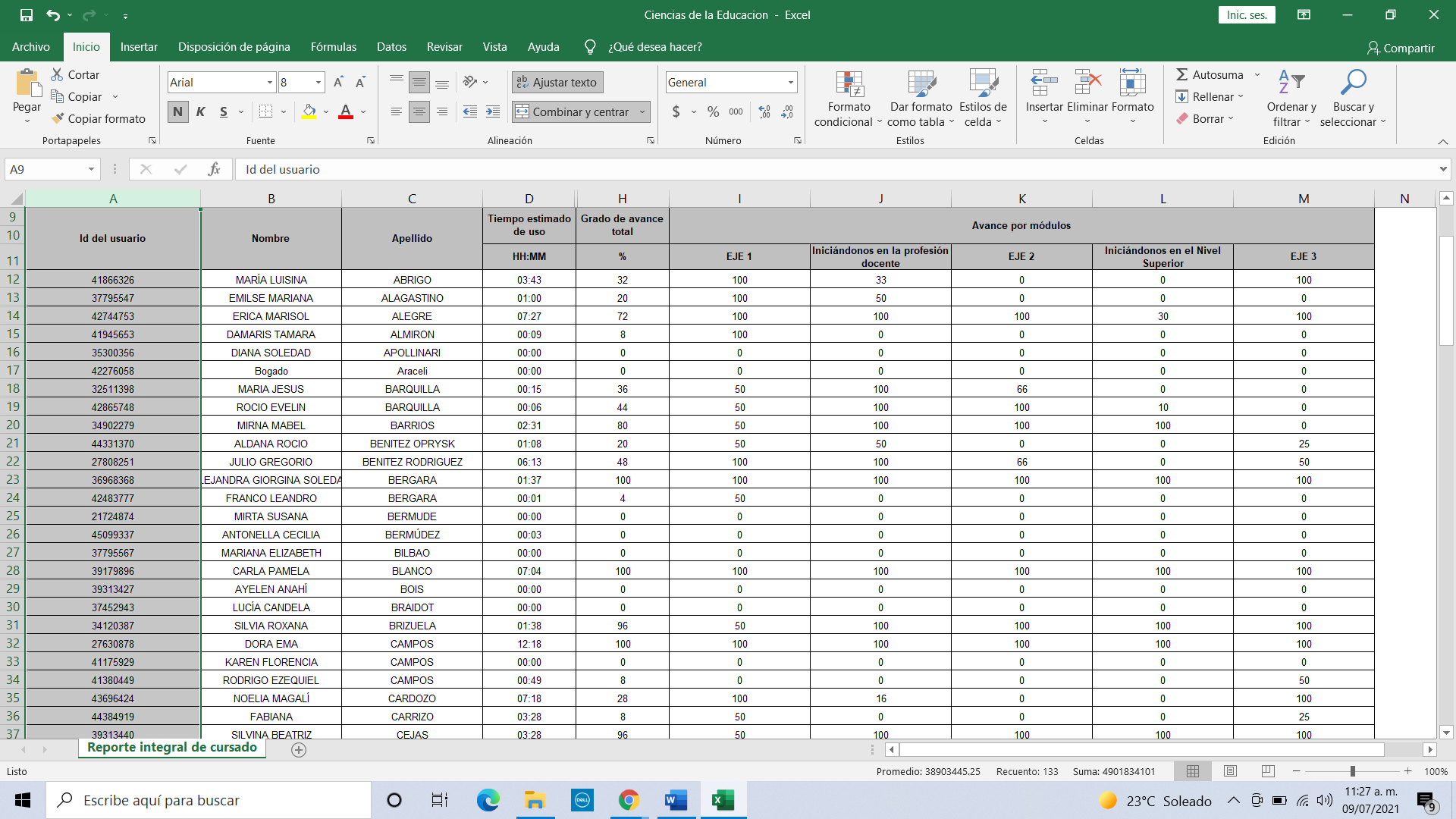
Google Colab, Jupyter, Python, y SciKit-Learn.

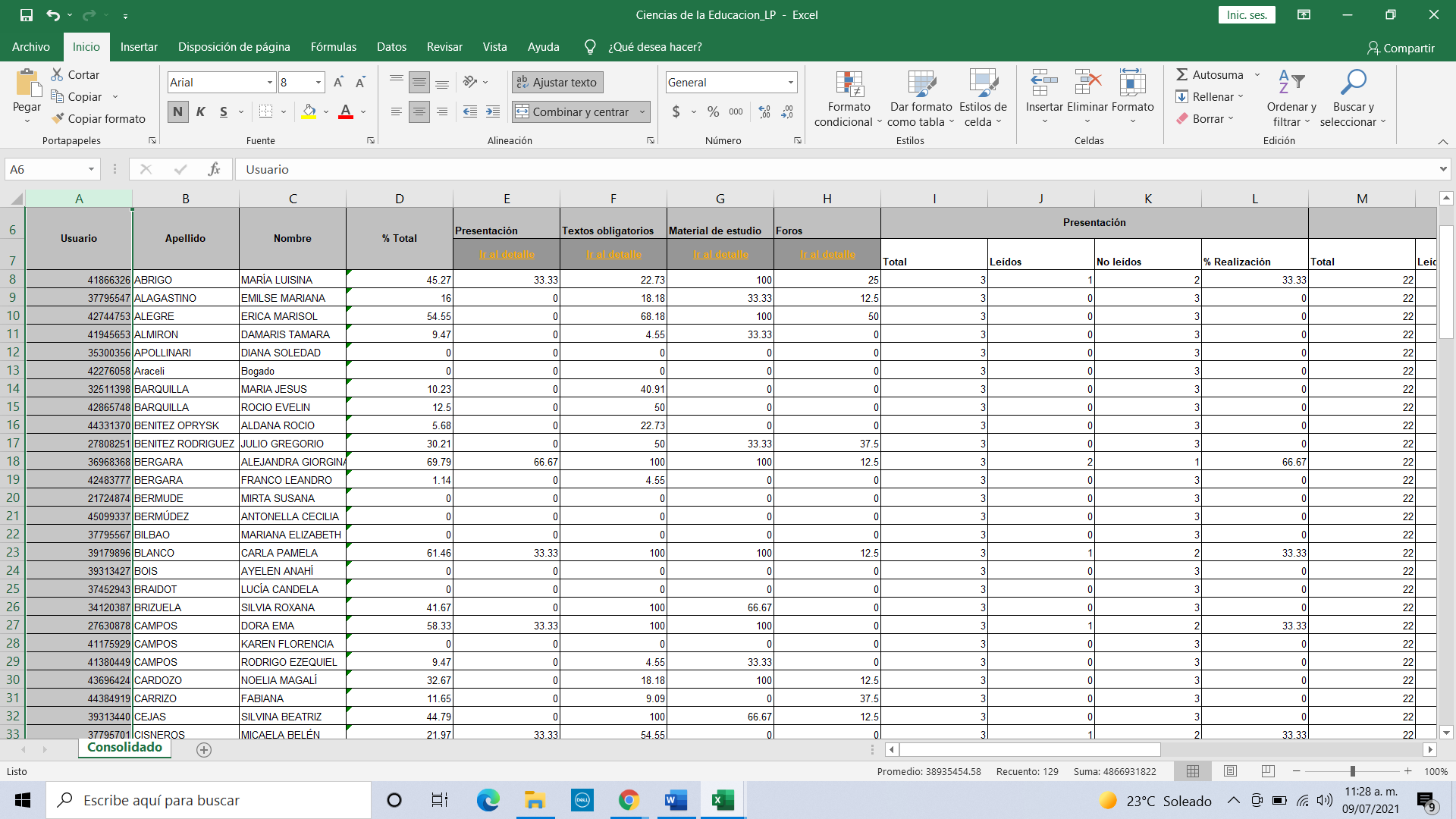
Probar una segunda alternativa de trabajo mediante:

Google Cloud  Vertex AI.

**Composición de los datos**

Datos extraídos de la plataforma virtual E-ducativa del ciclo lectivo 2020-2021 del Instituto de Educación Superior de Charata.





Estos datos van a ser procesados y preparados para poder utilizarlos en el proyecto, para ellos se va a: estandarizar los datos, normalizar los datos y realizar una eliminación de columna.

Actualmente se dispone un archivo con 756 registros.

**Resultados esperados**

Clasificar mediante aprendizaje automático a los estudiantes que tienen un rendimiento inferior en la utilización de la plataforma virtual del IES de Charata y permitir la implementación de estrategias alternativas de uso de plataformas virtuales o ecosistemas educativos multicomponentes y posibilitar la correcta toma de decisiones sobre los datos obtenidos.

**Indicar el repositorio de control de versiones a utilizar**

Google Colab.

<https://colab.research.google.com/drive/1hSKgQqvRkhCuMnk3htpRvvgIxCyCKx-0?usp=sharing>

GitHub

<https://github.com/juanejara/repojuan>

**Indicar una planificación temporal para la realización del trabajo**

|  |  |  |
| --- | --- | --- |
| Actividad \ Mes | Julio | Agosto |
| Entender los datos | 12/07 - 5 días – |  |
| Introducción a la libreria Sci-kit Learn | 12/07 - 5 dias – |  |
| Dividir los datos en grupos de entrenamiento y prueba | 17/07 - 5 dias - |  |
| Procesar nuestro conjunto de datos | 18/07 – 7 dias - |  |
| Implementación de Naive Bayes con Sci-kit Learn | 25/07 – 10 dias | 05/08 – 5 dias |
| Evaluación del modelo | 25/07 – 10 dias | 05/08 – 5 dias |
| Escritura del informe | 5 días | 10 días – 10/08 |

**Bibliografía de referencia**

[1] De-La-Hoz, E. J., De-La-Hoz, E. J., & Fontalvo, T. J. (2019). Metodología de aprendizaje automático para la clasificación y predicción de usuarios en ambientes virtuales de educación. Información tecnológica, 30(1), 247-254.

[2] Esparza, G. G., Fuentes, M. D. L. M., del Real, T. A. R., & Reich, J. C. (2017). Un modelo basado en el Clasificador Naïve Bayes para la evaluación del desempeño docente. RIED. Revista Iberoamericana de Educación a Distancia, 20(2), 293-313.

[3] Murphy, K. P. (2006). Naive bayes classifiers. University of British Columbia, 18(60), 1-8.

[4] Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46).

[5] Nieto Acevedo, Y. V. (2020). UDLearn: modelo de aprendizaje de máquina que facilita la toma de decisiones académicas en las instituciones de Educación Superior.

[6] Feng, Y., Zhou, M., & Tong, X. (2020). Imbalanced classification: an objective-oriented review. arXiv preprint arXiv:2002.04592.

[7] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

[8] Raschka, S., & Mirjalili, V. (2017). Python Machine Learning: Machine Learning and Deep Learning with Python. Scikit-Learn, and TensorFlow. Second edition ed.