

3 Auditoría

Equipo auditor: ERP BI

Equipo auditado: Big Eye BA

Asignatura: Proyectos IV

Juan Carlos Ávila*

Javier Coque†

Alejandro Gallego‡

Chantal López§

April 20, 2024

Introducción

El equipo Big Eye BA presentó sus avances en el proyecto en ese último Sprint antes de la entrega al cliente final en un par de semanas. Aunque tengan tiempo de poder añadir o modificar pequeñas cosas, se esperaba que se nos presentase un producto casi terminado —y así fue.

Resumen

El equipo auditado estuvo comentando que no todo había ido como esperaban y algunos errores —junto con sus soluciones— fueron comentadas:

- En primer lugar, mencionar que, inesperadamente, lo que más tiempo les llevó fue el refinamiento

*juan.avila@live.u-tad.com

†javier.coque@live.u-tad.com

‡alejandro.gallego@live.u-tad.com

§chantal.lopez@live.u-tad.com

de los modelos en general, pero, especialmente el de Catastro. Concretamente, fue el preproceso, con el objetivo de simplificar los datos obtenidos de Catastro, lo que más problema dio. Al parecer, las clasificación entre *edificio* y *no edificio* seguía dando problemas (se clasificaba como edificio zonas que no eran un edificio). Como medida respecto que, cabe mencionar que es muy buena, es tener en cuenta aquellos "edificios" con **al menos** un planta. Aunque todavía los resultados no son los esperados, nos parece un buen punto de partida del que partir para seguir mejorando el modelo y obtener unos resultados algo más satisfactorios.

- Respecto al modelo predictivo, TCN, en la última auditoría se comentó que se había elegido un sitio particular de Almería en donde no llovía casi. Esto porque las imágenes para estudiar el NDVI¹ son satelitales y los sitios muy nublados ocasionaban mucho ruido. Pese a que con esto se consigue mejorar el problema, no se consigue paliarlo del todo. La solución que aportan es no tener en cuenta estos datos, tratarlos como NaN y hacer una interpolación cúbica. Aquí, nuestro equipo sugiere, para futuras ocasiones, que algo común y aparentemente más fácil es utilizar métodos como ffill ó bfill junto con fillna para solventar el mismo problema. De todas maneras, su solución también parece correcta.
- Ligado con este punto anterior, también se comentó las dificultades que surgieron a la hora de desplegar los modelos de Pytorch en MLflow. Como respuesta a esto, se han desplegado en local. Algo que no les ha dado tiempo es la investigación del despliegue con la otra librería de *machine learning*: Tensorflow. El cliente Pablo les sugiere hacerlo, pues es más fácil este proceso con la librería recientemente mencionada que con Pytorch.

Recomendaciones

Como recomendaciones, el equipo ERP BI sugiere al equipo auditado que se reúnan para establecer que tareas pendientes son las más importantes para finalizar. Por ejemplo, el refinamiento de modelos parece ser más importante que mejorar el diseño y UX de la interfaz gráfica. Una vez ordenas las tareas por prioridad, que asignen personas a cada una de ellas según el tiempo y esfuerzo que en un primer momento se espere de ellas.

Esto con el fin de poder completar un MVP lo más completo posible para la presentación del cliente final.

¹Índice de vegetación de diferencia normalizada.

Conclusiones

Se ha notado que el equipo, especialmente en este último Sprint ha estado más unido o mejor comunicado que en otros Sprints. Por ejemplo, el miembro del equipo Álvaro, no tuvo problema en comentar la parte de Javier. Esto es algo que queremos remarcar como un punto muy **positivo**.

Por último —y más importante— queremos felicitar al equipo por su gran trabajo. La dificultad del proyecto es muy alta y aunque los resultados son mejorables, se nota el esfuerzo, dedicación y compromiso por parte de todo el equipo. Nuestro equipo no tiene ninguna duda de que conseguir entregar un MVP que logre satisfacer las expectativas del cliente final.