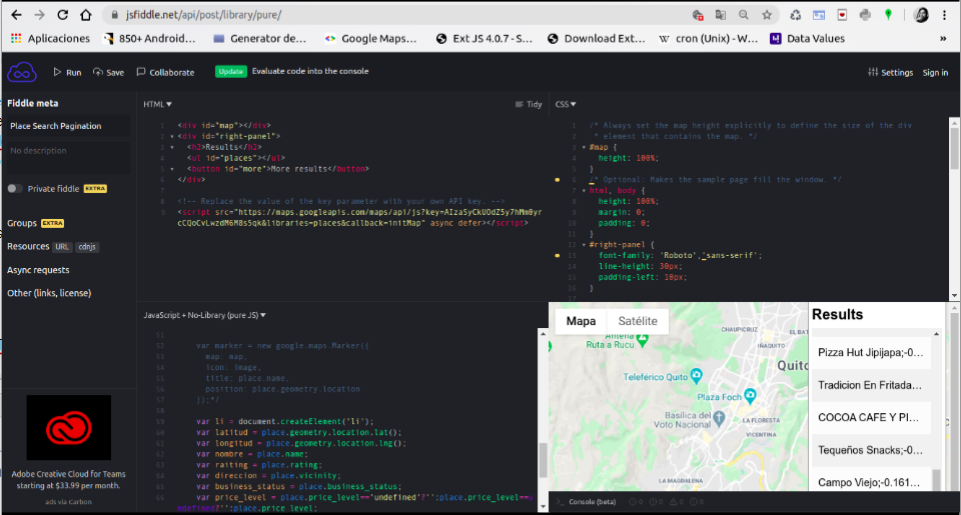
**GG-EAT**

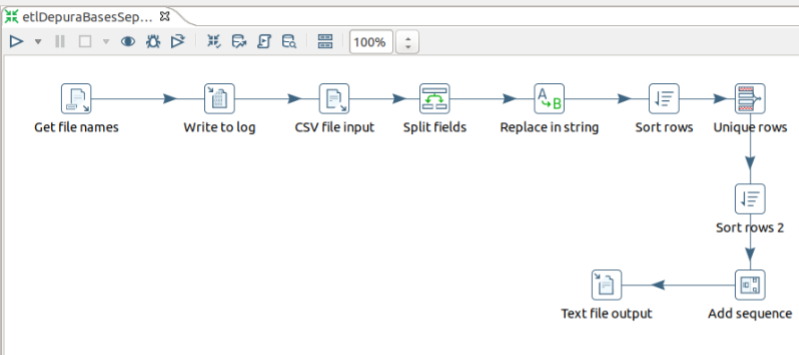
Al momento de realizar un viaje turístico o de negocio, es muy probable que no sepas de un buen lugar para comer y recibir una buena experiencia. Actualmente los restaurantes son calificados por todos sus clientes, sin embargo, no toma en cuenta los gustos de cada cliente. Debido a esto, es posible que no te agrade un restaurante por el precio, cantidad, calidad o ingredientes. Es por ello que se desarrolló el proyecto GG-EAT, cuyo objetivo es recomendar un top 5 de restaurantes de acuerdo a los gustos del usuario.

**Recopilación de Datos**

Después de revisar varios Datasets, incluso extranjeros, se observó que se requería crear un Dataset adecuado. Se utilizó la API de Google para obtener datos de los restaurantes cercanos a un área de Quito.



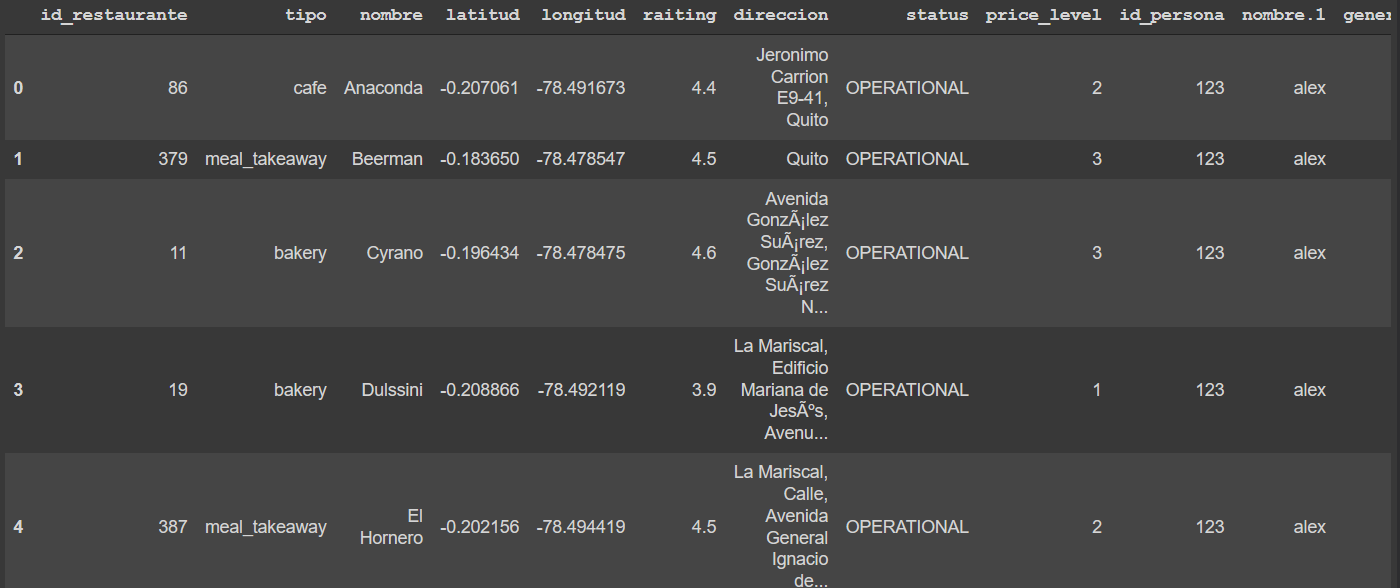
Luego se realizó una limpieza de los datos utilizando la herramienta ETL



Finalmente se realizó una encuesta en línea para adquirir las preferencias de las personas y si les gustaban los restaurantes de la zona.

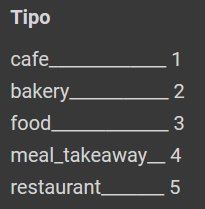
**Adecuación del Dataset**

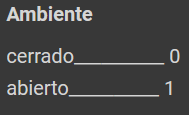
Los datos adquiridos con ayuda de la encuesta fueron tanto cualitativos como cuantitativos.



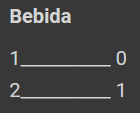
Para realizar un modelo de aprendizaje se debe contar con datos numéricos que representen a los datos originales. Por lo que se realizó lo siguiente

1. Se categorizo cada restaurante (del 1 al 5) a partir del rating obtenido (-1 al 4)
2. Se cuantificaron las variables “tipo” y “ambiente”

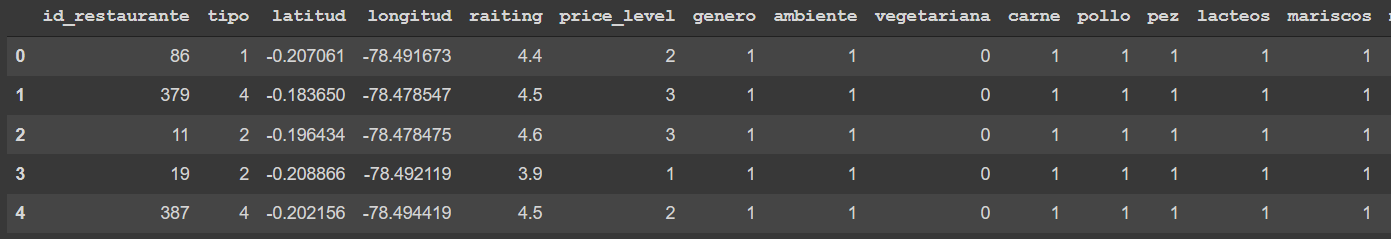




1. Se normalizan los valores de la variable “bebida” para su posible uso en una regresión logística



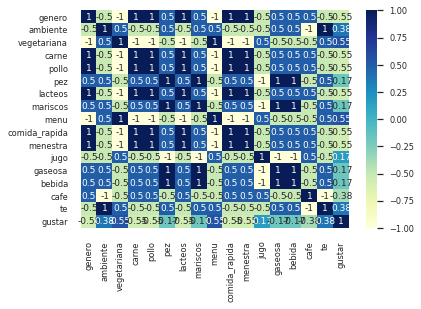
1. Se eliminaron variables innecesarias: “nombre”, “dirección”, “status”, etc.
2. Finalmente se obtienen los datos finales con los que se trabaja.



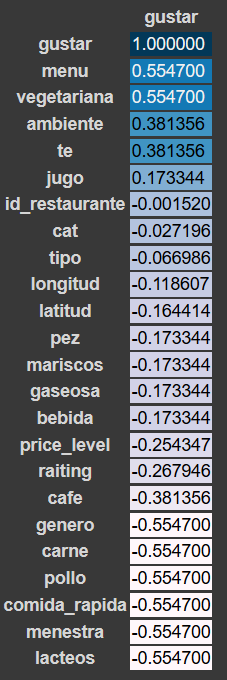
**Análisis de Correlación**

Con ayuda de las herramientas de Python, se obtuvo la correlación entre todas las variables. Debido a que muchas tenían una correlación igual a cero (0), se prefirió graficar 2 mapas de correlación de las variables. En ambos mapas se puede verificar que la variable “gustar”, la cual refleja si las personas les gustó dicho restaurante, se relaciona con el resto de variables.





Debido a esta característica, se decidió que la variable de salida del modelo sea “gustar”, y a partir de esa salida se puede lograr identificar los restaurantes más adecuados de acuerdo a los gustos de la persona que lo utilizara. Para definir que variables definitivamente serán la entrada del modelo, se grafica la correlación entre la salida (gustar) y el resto de variables. Finalmente se excluyen las variables cuyo valor sea menor a 0.38.

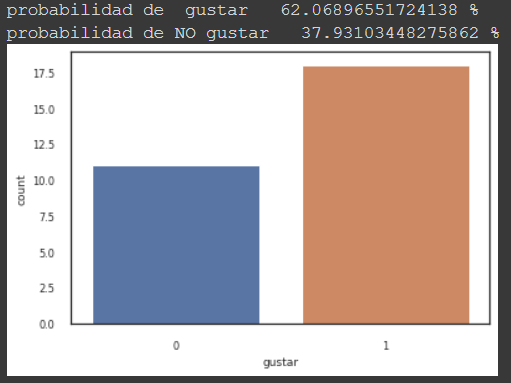


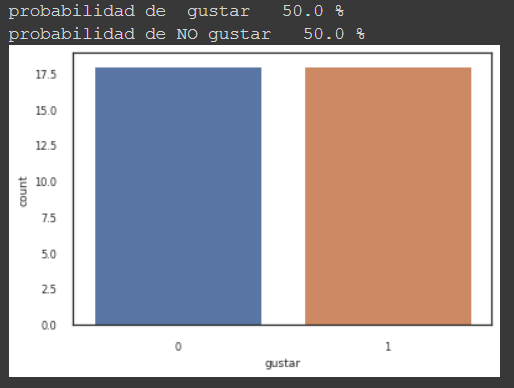
**Modelos de Aprendizaje y exportación de modelo**

Luego de definir las entradas y las salidas, se intentó utilizar distintos modelos de Machine Learning (ML). Debido a que la variable de salida era 1 o 0 (Le gusta o no un restaurante), se eligieron modelos de clasificación.

**Regresión Logística**

El primer intento fue la regresión logística, para lo cual se sobre muestreó los datos para que la probabilidad de que a una persona le guste un restaurante sea igual a que no le guste



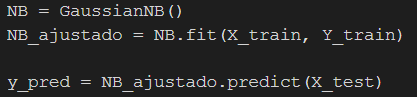


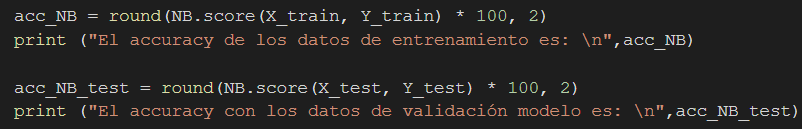
Sin embargo, la Regresión logística fracaso debido a la (relativamente) poca información. Por lo que se probó otros modelos.

**Otros Modelos de clasificación**

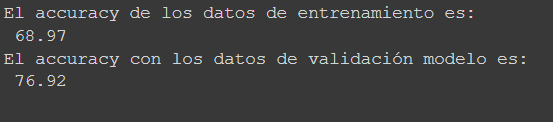
A diferencia de la regresión logística, los modelos de clasificación de Scikit-learn tienen 3 pasos básicos:

1. Definir Modelo
2. Entrenar Modelo
3. Calificar Modelo



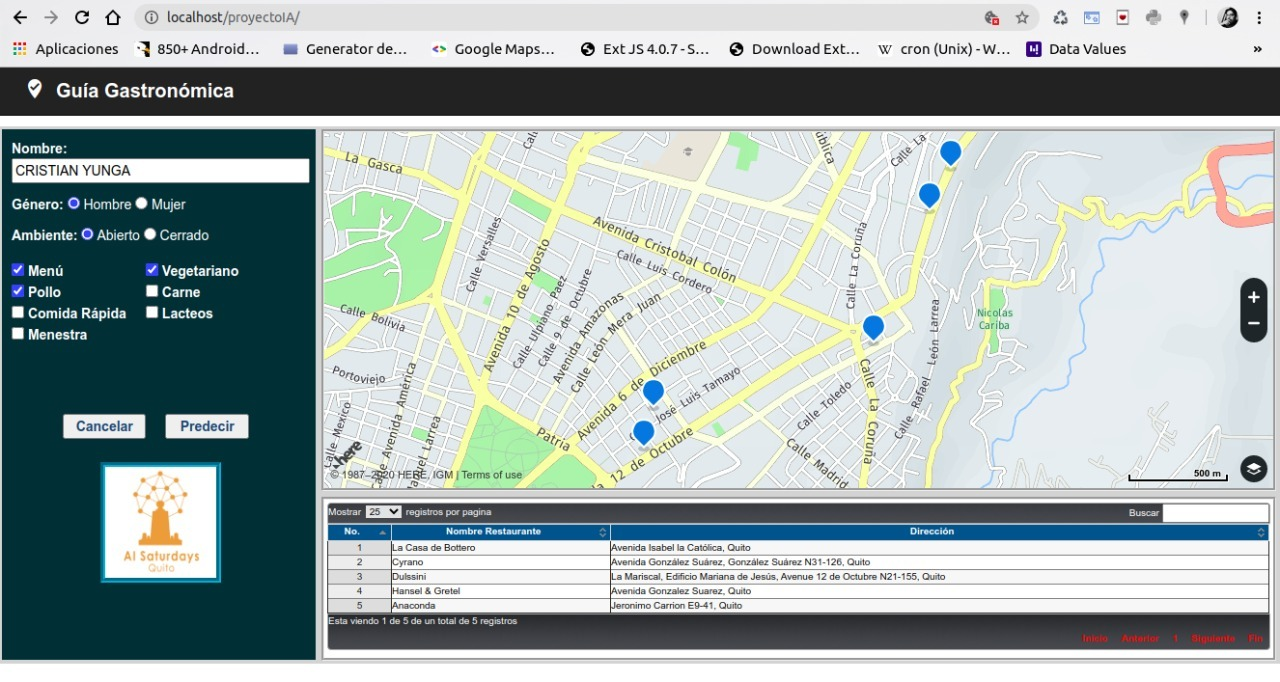


Por lo que se probó los modelos de Árbol de Decisión, Random Forest y Naive Bayes. Este último reflejo una precisión más alta que el resto de modelos por lo que se eligió como modelo de ML y se exportó para utilizarlo en una aplicación Web. Cabe recalcar que esta es solo una fase y el objetivo final es utilizar un modelo de ML no supervisado para que procese la nueva información que recopila la aplicación Web y mejore la precisión de la aplicación



**Aplicación WEB**

El producto final de este proyecto es una aplicación web que permite al usuario ingresar sus datos, preferencias y gustos para obtener una lista de restaurantes con su ubicación en el mapa



**Conclusiones**

Esta aplicación será de beneficio mutuo tanto para los usuarios como para los restaurantes, debido a que nuestro sistema ayuda a seleccionar la mejor opción para alimentarse y también difundir varios locales generando una gran variedad de opciones.

En los datos recopilados se pudo observar que las variables que definen el gusto de las personas son variables: conocidas y comunes, dentro de una región determinada (Sierra).