

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE OCCIDENTE
Departamento de Economía, Administración y Mercadología

Economía solidaria y trabajo digno

PROYECTO DE APLICACIÓN PROFESIONAL (PAP)
PROGRAMA DE RECONOCIMIENTO DE LAS FUNCIONES SUSTANTIVAS DE
LAS EMPRESAS II



ITESO, Universidad
Jesuita de Guadalajara

3G02 Integración profesional en instituciones financieras

Optimización de procesos y procesamiento de datos para el crecimiento de
la empresa

PRESENTAN

Programas educativos y Estudiantes

Ing. Financiero. Julio Olaf González Guzmán

Profesor PAP: Luis Guillermo Peña Lamas

Tlaquepaque, Jalisco, noviembre de 2022

ÍNDICE

Contenido

REPORTE PAP	2
Presentación Institucional de los Proyectos de Aplicación Profesional	2
Resumen	0
1. Ciclo participativo del Proyecto de Aplicación Profesional.....	0
1.1 Entendimiento del ámbito y del contexto	1
1.2 Caracterización de la organización.....	2
1.3 Identificación de la(s) problemática(s).....	3
1.4. Planeación de alternativa(s).....	3
1.5. Desarrollo de la propuesta de mejora	6
1.6. Valoración de productos, resultados e impactos	24
1.7. Bibliografía y otros recursos	25
1.8. Anexos generales.....	25
2. Productos	25
3. Reflexión crítica y ética de la experiencia.....	26
3.1 Sensibilización ante las realidades	26
3.2 Aprendizajes logrados	27

REPORTE PAP

Presentación Institucional de los Proyectos de Aplicación Profesional

Los Proyectos de Aplicación Profesional (PAP) son experiencias socio-profesionales de los alumnos que desde el currículo de su formación universitaria- enfrentan retos, resuelven problemas o innovan una necesidad sociotécnica del entorno, en vinculación (colaboración) (co-participación) con grupos, instituciones, organizaciones o comunidades, en escenarios reales donde comparten saberes.

El PAP, como espacio curricular de formación vinculada, ha logrado integrar el Servicio Social (acorde con las Orientaciones Fundamentales del ITESO), los requisitos de dar cuenta de los saberes y del saber aplicar los mismos al culminar la formación profesional (Opción Terminal), mediante la realización de proyectos profesionales de cara a las necesidades y retos del entorno (Aplicación Profesional).

El PAP es un proceso acotado en el tiempo en que los estudiantes, los beneficiarios externos y los profesores se asocian colaborativamente y en red, en un proyecto, e incursionan en un mundo social, como actores que enfrentan verdaderos problemas y desafíos traducibles en demandas pertinentes y socialmente relevantes. Frente a éstas transfieren experiencia de sus saberes profesionales y demuestran que saben hacer, innovar, co-crear o transformar en distintos campos sociales.

El PAP trata de sembrar en los estudiantes una disposición permanente de encargarse de la realidad con una actitud comprometida y ética frente a las disimetrías sociales. En otras palabras, se trata del reto de “saber y aprender a transformar”.

El Reporte PAP consta de tres componentes:

El primer componente refiere al ciclo participativo del PAP, en donde se documentan las diferentes fases del proyecto y las actividades que tuvieron lugar durante el desarrollo de este y la valoración de las incidencias en el entorno.

El segundo componente presenta los productos elaborados de acuerdo con su tipología.

El tercer componente es la reflexión crítica y ética de la experiencia, el reconocimiento de las competencias y los aprendizajes profesionales que el estudiante desarrolló en el transcurso de su labor.

Resumen

En este proyecto llamado “**Optimización de procesos y procesamiento de datos para el crecimiento de la empresa**” estuvo enfocado en identificar y ayudar en las problemáticas y tareas que se tenían en los proyectos que se están llevando a cabo por parte de la empresa Multiburo, estas actividades requirieron el conocimiento en las disciplinas de financiera, procesamiento de datos y tecnología, ya que se crearon ciertas herramientas entre las más destacadas: manuales de base de datos, un layout para un reporte de datos como un nuevo servicio que otorgará la empresa, funciones en códigos para el procesamiento de datos que ayudan al entrenamiento de modelos de score crediticio, de igual forma se crearon funciones de código que evalúan el desempeño de dichos modelos.

La realización de este trabajo presentado brinda como producto final herramientas aplicadas en programación en Python con los códigos necesarios para el procesamiento de datos y evaluación de modelos, así como manuales que ayudan a comprender mejor y analizar con los datos con los que se van a trabajar.

Me hacen sentir orgulloso de lo realizado y espero que sus resultados puedan servir tanto de guía como de utilidad. Dando inicio hacia un mayor y mejor beneficio para quienes busquen realizar o seguir trabajando en el mismo proyecto a futuro, de esta manera quienes muestren su interés podrán seguir partir de las bases que se realizaron en este proyecto.

1. Ciclo participativo del Proyecto de Aplicación Profesional

El PAP es una experiencia de aprendizaje y de contribución social integrada por estudiantes, profesores, actores sociales y responsables de las organizaciones, que de manera colaborativa construyen sus conocimientos para dar respuestas a problemáticas de un contexto específico y en un tiempo delimitado. Por tanto, la experiencia PAP supone un proceso en lógica de proyecto, así como de un estilo de trabajo participativo y recíproco entre los involucrados.

1.1 Entendimiento del ámbito y del contexto

El crédito o contrato de crédito es un servicio que brindan las entidades financieras, el requerir de este servicio tiene un costo el cual se ve reflejado en las comisiones o interés que pagamos adicionales al monto original del préstamo. Existe diferentes tipos de crédito como las tarjetas de crédito, créditos automotrices, créditos hipotecarios, créditos de nómina, los créditos personales, entre otros.

El proceso de tramitar un crédito consta en realizar una evaluación del perfil del solicitante el cual puede ser una persona física o moral, el cual comienza en la validación de identidad de la persona o de una empresa que este constituida correctamente, comprobar su domicilio actual, su nivel de ingresos, en algunos casos dependiendo el tipo de crédito que se está solicitando se debe tener un bien o aval que garantice el pago del crédito en caso de incumplimiento del solicitante y como último paso una consulta de su historial crediticio, esto dictamina si el solicitante es elegible para obtener el crédito.

Una vez realizado el proceso mencionado anteriormente, las entidades financieras contratan a instituciones privadas como son las SIC (Sociedades de Información Crediticia), las cuales brindan el servicio del manejo, recolección, entrega y envío de información sobre el historial crediticio de personas físicas y morales. Las cuales se encargan de evaluar a los solicitantes del crédito mediante modelos de riesgo dando como producto final un score crediticio que califica al solicitante, analizando la información de su historial crediticio, este score nos muestra las posibilidades que tiene el solicitante en caer en incumplimiento de sus pagos.

El problema de dichos modelos es que están sesgados dejando fuera a muchos solicitantes que pueden cumplir con la obligación de sus pagos, la correcta evaluación del riesgo crediticio es un elemento fundamental para el otorgamiento de préstamos, los modelos *scoring* perfilan estadísticamente a los solicitantes de acuerdo al riesgo de pago, facilitan la reducción de la asimetría del conocimiento entre el solicitante y la entidad de financiera, gracias a diversas herramientas como la tecnología, la *big data* y la ciencia de datos, las entidades pueden tomar decisiones en cuestión de minutos sin solicitar garantías, patrimonio

u otra documentación, estos modelos tienen el potencial de poder ofertar créditos a un mayor porcentaje de los solicitantes ajustando los parámetros de los modelos dependiendo el giro de las empresas que otorgan créditos.

1.2 Caracterización de la organización

Este proyecto estuvo vinculado con la empresa Multiburo que se encarga de la automatización del análisis y calificación de prospectos para el otorgamiento de créditos y renta de inmuebles. Dentro de la empresa se abrió el área de ciencia de datos que se encarga del procesamiento, limpieza y análisis de los datos para crear modelos de *score* crediticio,

El propósito del PAP es brindar el conocimiento adquirido a lo largo de la carrera y poder realizar tareas esenciales para la empresa en el área de ciencia de datos, donde la empresa tiene establecido un lapso de dos semanas para lograr realizar las tareas asignadas en las reuniones llamadas *sprint* las cuales se llevaban a cabo cada dos semanas, en ellas se aborda en lo que trabajo cada uno de los integrantes del equipo de *developers* y de ciencia de datos y donde se asignan nuevas tareas o continuar con las que quedaron pendientes.

Dentro de este PAP se contaba con 2 integrantes de la carrera de Ingeniería Financiera y 2 personas de la carrera de Administración Financiera. El enfoque de 3 de los integrantes del equipo estuvo en el área de administración y finanzas, mientras que su servidor estuvo en el área de ciencia de datos.

En esta área tuve las tareas asignadas de recolección de datos, así como el procesamiento y limpieza de mismo con la herramienta del lenguaje de programación Python, donde también realicé códigos para la evaluación de modelos de *score* crediticio y conexiones a la Api de Google maps. ayudar a crear archivos tipo *json*, crear cuestionarios en *Google Forms*, explorar modelos de clasificación en el lenguaje de programación R y automatizar layouts en Excel.

1.3 Identificación de la(s) problemática(s)

Desde el comienzo se planificaban las actividades realizadas por cada miembro del área de trabajo, esto se realizaba cada dos semanas en las reuniones *sprint*, donde se resolvían las preguntas: ¿En qué se estuvo trabajando? ¿Cuáles fueron los bloqueos? Y ¿en qué se seguirá trabajando? Aquí surgían las problemáticas que iba teniendo la empresa a lo largo del proyecto y que fuimos abordando conforme iban surgiendo.

En estas juntas se logra identificar ciertas problemáticas como la falta de datos para los entrenamientos de los modelos, manuales de base de datos para ayudar a los *developers* a conocer la estructura de ciertas variables y como transformarlas en sus campos de desarrollo, buscar alternativas de códigos que necesitaran menos recursos para reducir los tiempos de respuesta al momento de ejecutar códigos y con menos margen de error, así como la automatización de reporte de base de datos a Buro de crédito

1.4. Planeación de alternativa(s)

Las actividades que se realizaron a lo largo del proyecto fueron divididas en *sprint*:

En mi primer *sprint* fui conociendo al equipo con el cual me iba a desenvolver, en el cual noté que uno de los principales problemas que se presentaban dentro de la empresa era que los desarrolladores tenían problemas al momento de querer saber cómo manejar e interpretar algunas variables que estaban presentes dentro de los archivos tipo *json* que se reciben al momento de hacer consultas a Buro de Crédito con los cuales ellos trabajaban día a día.

Fue necesario buscar una solución práctica que todos lo tuvieran al alcance para cuando la necesitaran. Por lo que se optó por un manual de la base de datos en un documento Excel interactivo que ayudaría a localizar mejor una variable, con sus descripción y notas echas por los mismos desarrolladores. El cual se dividiría en dos en uno para la base de datos de Persona Física y otra para Persona Moral

Segundo *sprint*:

En la segunda reunión se necesitaba recolectar datos del gasto promedio de los hogares para un modelo llamado “Cuestionario de confiabilidad” en el cual nos indica que tan verdadero es lo que está declarando una persona que gasta dependiendo sus ingresos, esto con la finalidad de evaluar a las personas que no pueden comprobar sus ingresos. La base de datos ayudaría a entrenar el modelo. Se planteo combinar dos bases de datos que se encuentran en la página oficial del INEGI para obtener uno solo más completo y con más información.

Tercer *sprint*:

Se cuenta con un modelo de *score* crediticio llamado “mb-score-model” que se encarga de calcular un *score* que se aproxime al que entrega a Buro de Crédito, pero a diferencia de él, este modelo lo clasifica en estrellas lo cual lo vuelve más amigable a los usuarios evaluar al solicitante de crédito. Actualmente se cuenta con la versión 2 de este modelo, pero no se había evaluado el desempeño de esta nueva versión, por lo que se propuso una función que serviría para evaluar dos modelos con respecto al *score* de Buro de Crédito, y no solo funciones para este caso si no para los modelos futuros que se vayan actualizando.

Dentro de este mismo *sprint* se buscaba obtener la longitud y latitud de una dirección, lo cual se hacia con un robot, pero se busco una forma más facial mediante una conexión a la API de Google maps, la cual reducirá los errores al momento de escribir una dirección manualmente para el modelo “socioeconómico”.

Cuarto y quinto *sprint*:

En esta etapa que duro un mes nos enfrentamos con el desafío de encontrar una base de datos de la cual pudiéramos obtener la información básica del hogar como se muestra en el cuestionario básico por manzana en el portal del INEGI, pero no se encontraba con la amplitud que se buscaba por lo que se recurrió a contactar directamente a las oficinas del INEGI para solicitar un procesamiento de los datos como los necesitábamos, este servicio tenia un costo el cual lo cubrió la empresa.

Estos datos ayudarían a entrenar el modelo de “socioeconómico” en el cual nos indica en una escala del 1 al 7 el nivel económico de la persona que quiere solicitar un préstamo inmobiliario.

También necesitábamos datos concretos mediante una encuesta para evaluar los resultados que arrojaría el modelo “Cuestionario de confiabilidad”.

Se planeo un reto de exploración de un algoritmo de árbol de decisiones, el cual tenía la finalidad de moverte entre los nodos y encontrar el camino más corto a la rama deseada.

Sexto *sprint*:

Las actividades asignadas en este *sprint* fueron relativamente sencillas ya que se necesitaba crear archivos *json* en los cuales se presentarán las recomendaciones que se dan a los clientes, en este caso las empresas, las recomendaciones que se les dan son advertirlos sobre algunas de las complicaciones que tiene sus clientes al momento de solicitar un crédito y a los usuarios que tramitan el crédito las razones por las cuales se les la calificación del score.

Otra de las tareas asignadas fue el estandarizar las respuestas echa en la encuesta anterior para generar respuestas aleatorias que sean falsas para evaluar el modelo de “Cuestionario de confiabilidad”.

Séptimo *sprint*:

Para finalizar este proyecto se presentó un nuevo servicio que se va a ofertar a las entidades financieras, es el reporte del comportamiento de pago de sus clientes de los créditos que les otorgaron a Buro de Crédito. Este fue uno de los grandes retos porque se tenía que optimizar un layout con las restricciones sobre los valores que ingresarían, ya que Buro de Crédito es muy estricto sobre como debe ser el formato y la calidad de los datos que se le entrega.

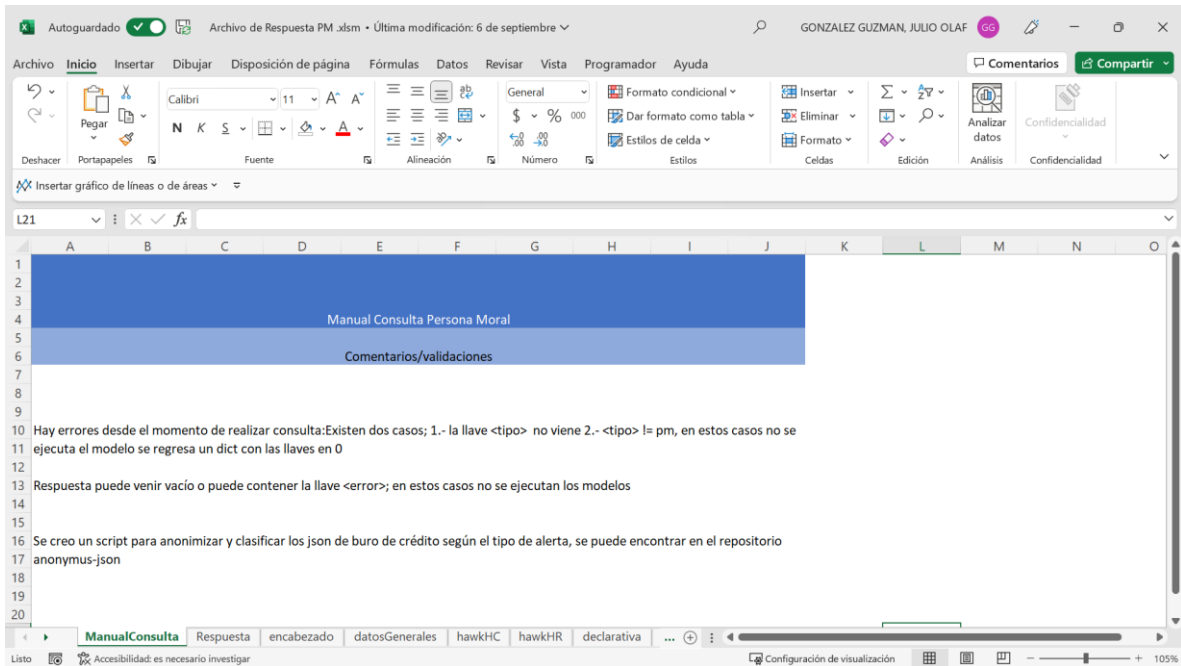
Este layout debía de ser intuitivo para las empresas a la hora de llenarlo, con pequeños mensajes de entrada de cómo es el formato de cada campo a llenar y mostrar mensajes de error sobre los errores que se comenten al momento de hacerlo.

1.5. Desarrollo de la propuesta de mejora

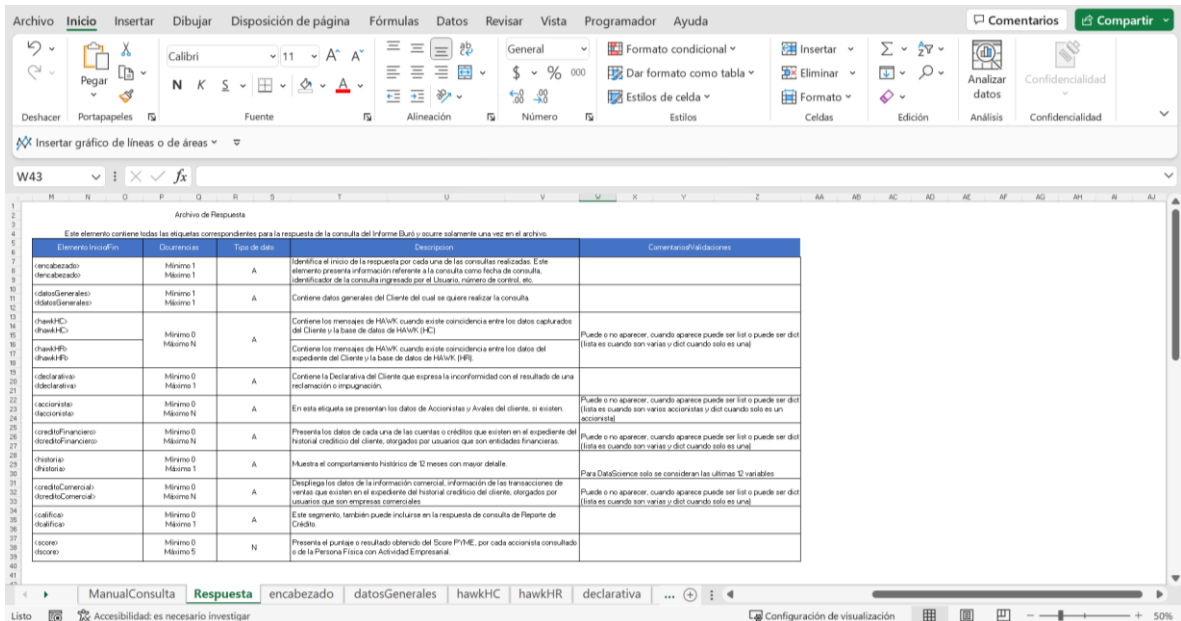
Como se había comentado anteriormente se crearon dos manuales para la base de datos de la empresa de Buro de Crédito, un manual para persona física y otro manual para persona moral.

Se buscó crear un manual que facilitara a los desarrolladores a buscar la solución de como interpretar o trabajar con variables que venían en los archivos con los que ellos trabajan, por lo que estos manuales en Excel cuentan con explicaciones en cada área sobre los errores que se les pudiera presentar desde un inicio hasta el final de un archivo, el cual cuenta con hipervínculos que los guían más rápido a encontrar las variables con las que tenían dificultades para trabajar, de igual forma pueden hacer anotaciones dentro de estos archivos explicando errores o dificultades futuras que se vayan presentando.

A continuación, se presenta la portada del manual donde se puede apreciar comentarios de los mismos desarrolladores explicación situaciones con las que ellos ya se han presentado cuando reciben un archivo con el cual van a trabajar y desde que lo reciben ya contiene ciertos errores lo cual se documento para que en escenarios futuros sepan que hacer o como tratar estos errores.



En la siguiente hoja del manual se puede observar como hay una tabla que funciona como índice indicando los elementos en los que se dividen los archivos que se reciben, haciendo una descripción de cada elemento, mostrando las veces que aparecen en los archivos, así como los tipos de dato que son y con las notas en los elementos donde se presentan los errores mas comunes.



Donde cada elemento de la tabla función como un hipervínculo que los lleva a la hoja donde se desglosa todas las variables dentro de cada elemento con el mismo formato mostrado, esto se ejemplifica mejor en la siguiente imagen:

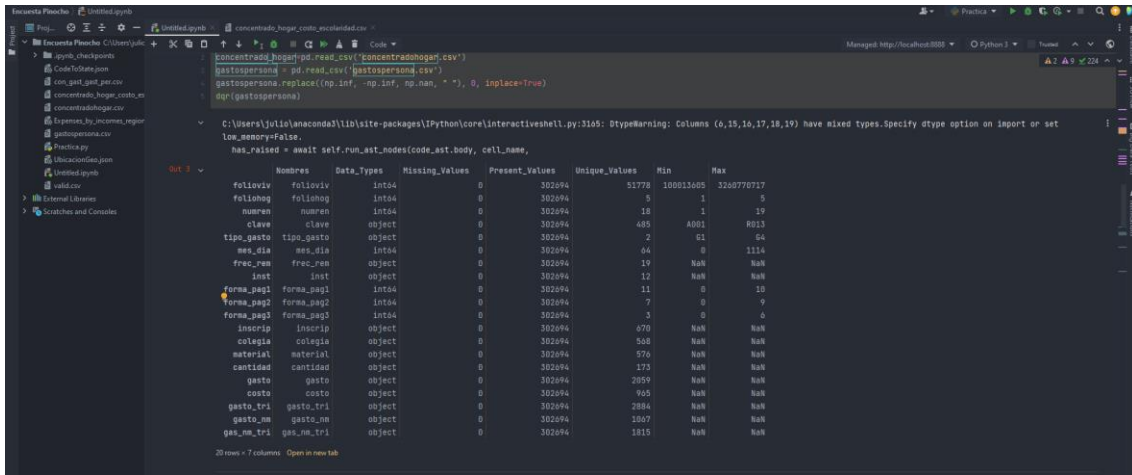
Evento	Ocurrencias	Tipo de dato	Descripción	Comentarios/Adicciones
encabezado encabezado	Mínimo 1 Máximo 1	A	Identifica el inicio de la respuesta por cada una de las consultas realizadas. Este elemento presenta información referente a la consulta como fecha de consulta, identificador de la consulta ingresado por el Usuario, número de control, etc.	
datosGenerales datosGenerales	Mínimo 0 Máximo N	A	Contiene datos generales del Cliente del cual se quiere realizar la consulta.	
hawkHC hawkHC	Mínimo 0 Máximo N	A	Contiene los mensajes de HAWK cuando existe coincidencia entre los datos capturados del Cliente y la base de datos de HAWK (HIC)	Puede o no aparecer, cuando aparece puede ser list o puede ser dot (lista es cuando son varias y dot cuando solo es una)
hawkHR hawkHR	Mínimo 0 Máximo N	A	Contiene los mensajes de HAWK cuando existe coincidencia entre los datos del expediente del Cliente y la base de datos de HAWK (HIC)	Puede o no aparecer, cuando aparece puede ser list o puede ser dot (lista es cuando son varias y dot cuando solo es una)
declarativa declarativa	Mínimo 0 Máximo N	A	Contiene la Declarativa del Cliente que expresa la inconformidad con el resultado de una reclamación o impugnación.	
accionista accionista	Mínimo 0 Máximo N	A	En esta etiqueta se presentan los datos de Accionistas y Avales del cliente, si existen.	Puede o no aparecer, cuando aparece puede ser list o puede ser dot (lista es cuando son varias accionistas y dot cuando solo es un accionista)
creditoFinanciero creditoFinanciero	Mínimo 0 Máximo N	A	Presenta los datos de cada una de las cuentas o créditos que existen en el expediente del titular crédito del cliente, otorgados por usuarios que son entidades financieras.	Puede o no aparecer, cuando aparece puede ser list o puede ser dot (lista es cuando son varias y dot cuando solo es una)
historia historia	Mínimo 0 Máximo 1	A	Muestra el comportamiento histórico de 12 meses con mayor detalle.	Para DataScience solo se consideran las últimas 12 variables
creditoComercial creditoComercial	Mínimo 0 Máximo N	A	Despliega los datos de la información comercial, información de las transacciones de ventas que existen en el expediente del titular crédito del cliente, otorgados por usuarios que son empresas comerciales.	Puede o no aparecer, cuando aparece puede ser list o puede ser dot (lista es cuando son varias y dot cuando solo es una)
credificac credificac	Mínimo 0 Máximo 1	A	Este segmento, también puede incluirse en la respuesta de consulta de Reporte de Crédito.	
score score	Mínimo 0 Máximo 5	N	Presenta el puntaje o resultado obtenido del Score Pyme, por cada accionista consultado de la Persona Física con Actividad Empresarial.	

Con estos hipervínculos presentes en todo el archivo hace más práctico y eficiente el encontrar una variable que presenta un error dentro de un archivo para saber cómo trabajar con él, estas características mencionadas están presente en los dos manuales para los dos tipos de archivos que recibe la empresa, los cuales ya están disponibles para los desarrolladores en la carpeta de Google drive de la empresa.

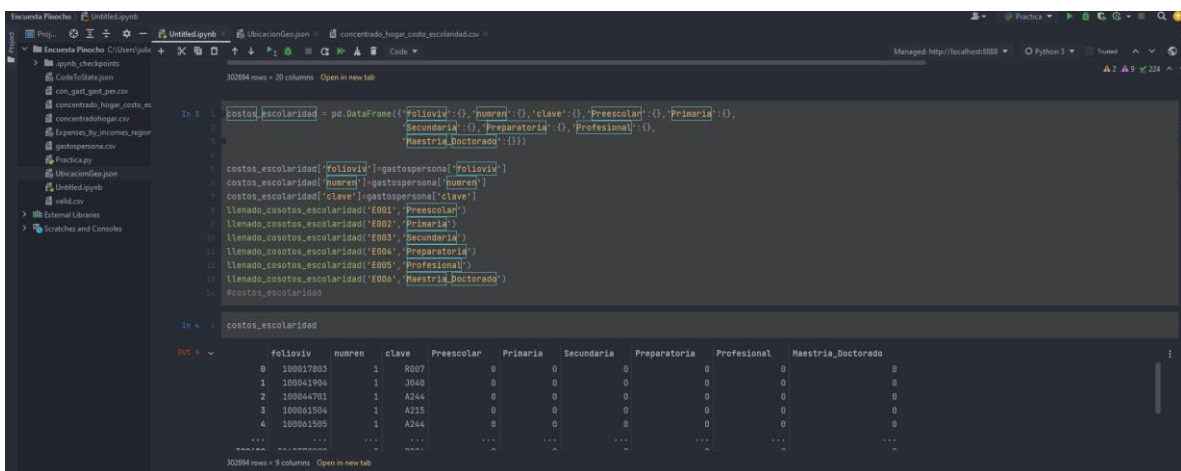
Sprint 2:

Una vez que extrajimos los archivos de la base de datos de INEGI, se hizo un análisis de los datos ver como se presentaba la información con la que íbamos a trabajar, para lo cual se creó una función que nos haría un resumen de los datos el cual nos muestra la lista de las variables que se presentan, el tipo de datos, si hacen falta valores, el total de datos presentes por variables, los valores únicos, mínimos y máximos de cada variable, con esta información

podemos anticiparnos a los posibles problemas con los que nos enfrentaremos al momento de trabajar con la base de datos como se muestra a continuación.



Una vez explorado las dos bases de datos con la que se cuentan, se hace la limpieza de los datos con los que vamos a trabajar. En el archivo de gastospersonas nos interesa la información de los gastos promedio que se tiene a nivel educativo por estado y municipio, por lo que se tuvo que hacer una transformación de los datos, ya que el nivel educativo se presentaba por claves y para que sea más fácil su identificación se tuvo que hacer una transformación de los datos.



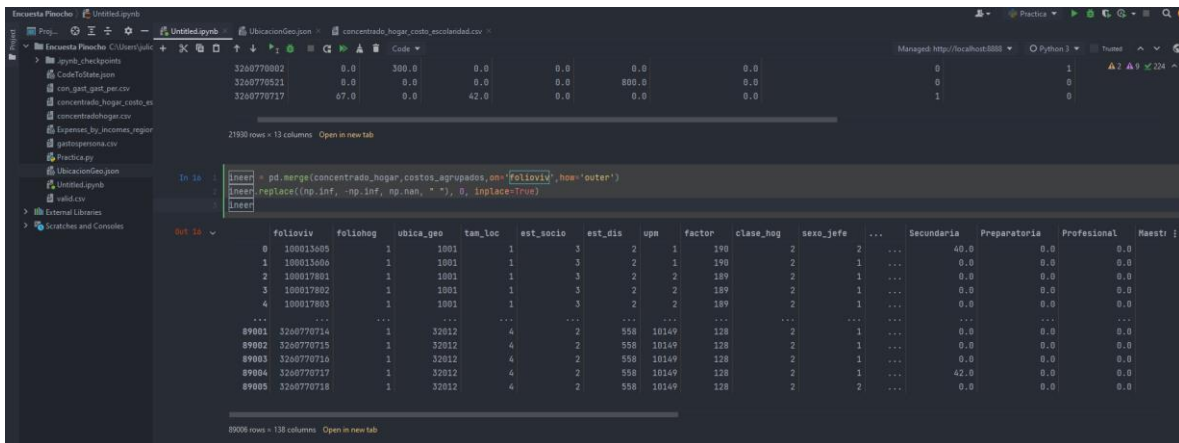
Una vez presentados de esta manera era más fácil identificar cuanto se gasta en promedio dependiendo la escolaridad de la zona.

Este mismo proceso se ejecutó para los gastos que se tienen para seguros, como varios claves de gastos se repetían, se agruparon con suma de las demás variables para tener datos más concretos.

Como siguiente paso era calcular el número de personas en cada grupo de escolaridad, lo cual se tuvo que tomar varias validaciones con otras variables para poder encontrar el número de personas.

En este punto nos dimos cuenta de que la variable de gastos en seguro no nos estaba aportando la información suficiente para tomarla en consideración, por lo que se optó por eliminarla de la base de datos que estamos creando. Como paso final se agruparon el foliovienda para solo tener un registro único y combinarlo con la base de datos de concentrado_hogar, de esta forma aumentamos la información que nos brindaba una sola base de datos.

Este resultado se puede ver en la siguiente imagen.

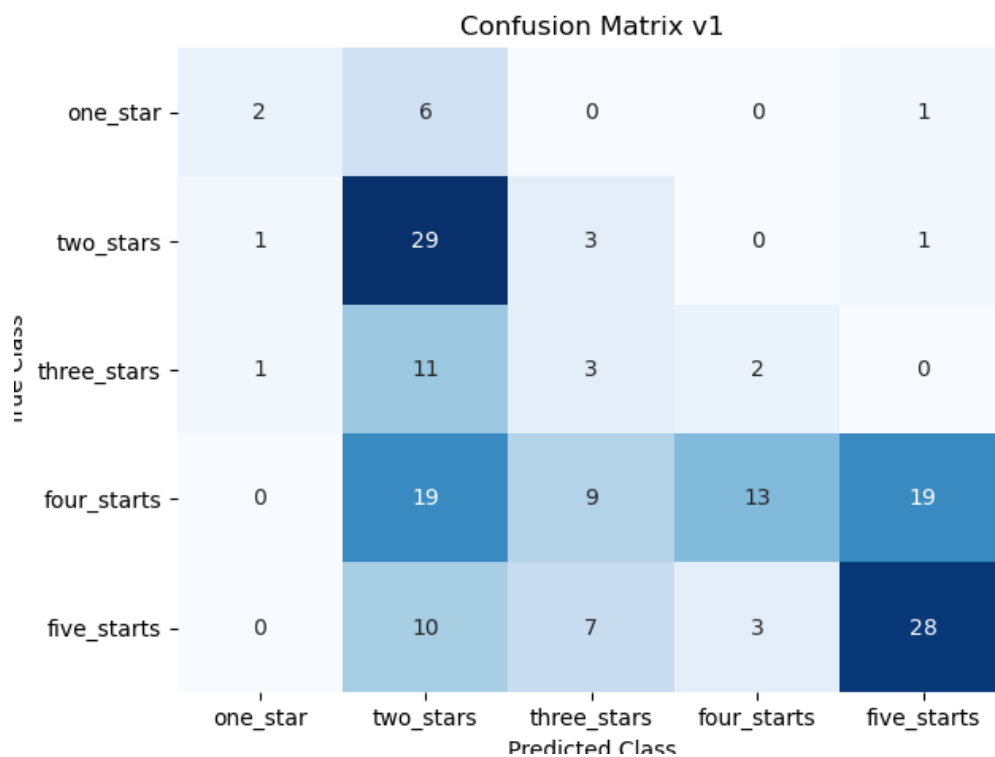


Esta base de datos nos ayuda a entrenar el modelo de score “Cuestionario de confiabilidad” que nos permitirá dictaminar que tan confiable es una persona al momento de declarar sus gastos.

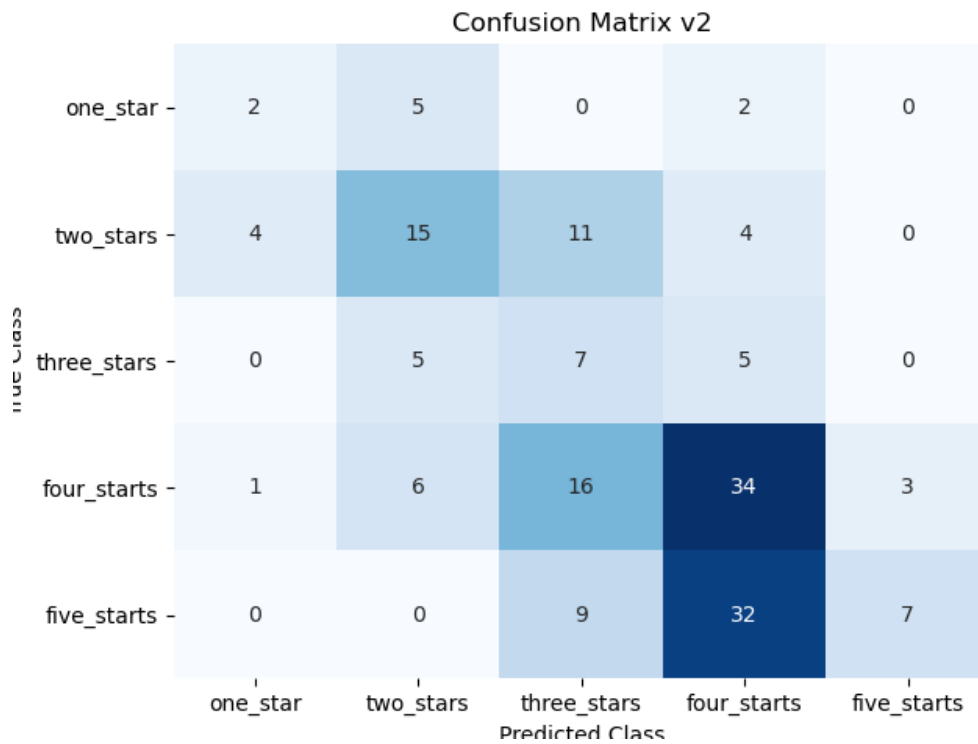
Sprint 3:

En esta etapa del proyecto se desarrolló una función en lenguaje Python que evalúa el desempeño de las dos versiones que hay del modelo “mb-score-model”, como *output* nos arroja dos imágenes que nos muestra la matriz de confusión, esta herramienta nos permite analizar los resultados de cómo trabaja un algoritmo de aprendizaje supervisado, la cual se presenta en forma de tabla, de manera que en cada columna aparecen el número de predicciones de cada clase, mientras que cada fila muestra el número real de instancias de cada clase. Para ejemplificarlo mejor se puede observar el ejemplo obtenido del algoritmo creado.

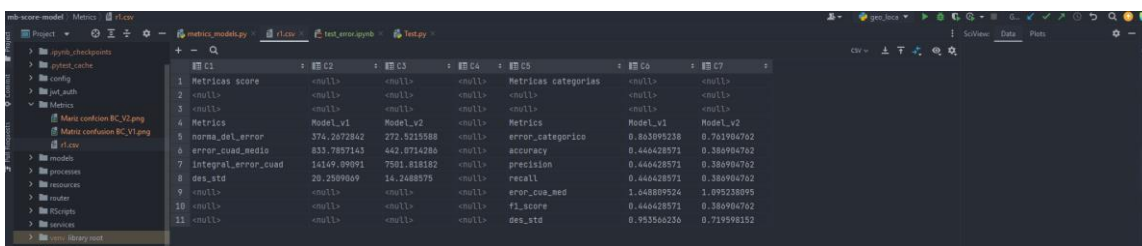
Esta imagen representa la matriz de confusión de la versión uno del modelo en el cual podemos notar que este modelo es bueno al momento de clasificar las puntuaciones más alta y que la dispersión de los valores predichos no es tan significativa.



A diferencia de la versión uno del modelo podemos notar que la versión dos tiene un número mayor de valores predichos correctamente, pero los valores sesgados están más dispersos por lo que podemos decir que la versión 2 al momento de clasificar es muy bueno pero los errores presentes son muchos mayores que en el anterior.



Como ultimo *output* genera un *DataSet* crea un archivo csv donde vemos las métricas de desempeño de cada modelo:



Estos resultados nos mostraron que algo estaba ocurriendo en la versión dos del modelo que estaba dando incongruencias por lo que llevo a su ajuste para aumentar su desempeño y reducir el error del modelo.

En la segunda tarea asignada dentro de este *sprint* fue conectarme a la API de Google maps la cual agilizó el tiempo de respuesta al momento de obtener la latitud y longitud de una

dirección para el modelo socioeconómico que se tiene lo cual arroja el resultado en mucho menor tiempo y con un margen de error al momento de escribir la dirección mayor, sin necesidad de ser tan precisa.



```
import webbrowser as wb
import pyinputgui as pt
import pandas as pd
import time
import pyperclip
import googlemaps as gm
pip install googlemaps
import re
import os
import math
import json

def addressToLocation(address:str):
    API_key = 'AIzaSyCfr8v51z6Vg3Nv_an_Ryyd00hza1k'
    map_client = gm.Client(API_key)
    response = map_client.geocode(address)
    locat = (response[0]['geometry']['location']['lat'],response[0]['geometry']['location']['log'])
    locat = str(locat)
    locat = locat.replace(",")
    locat = locat.replace(" ")
    return locat
```

Sprint 4 y 5:

Nos enfocamos en recolectar la información que necesitábamos para generar un modelo de *score* “sociodemográfico” el cual se base en el modelo AMAI, el cual con una serie de preguntas básicas podemos clasificar el nivel socioeconómico de una persona, las variables que considera dicho modelo son:

- El grado de estudios del jefe del hogar
- Número de baños completo en la vivienda
- Número de vehículos en el hogar
- Cuenta con internet la vivienda
- Número de personas que trabajan del hogar
- Cuantos dormitorios hay en la vivienda

Conociendo estas respuestas el modelo puede clasificar en 7 niveles diferentes el nivel económico de esa vivienda para ello era necesario recolectar información de todo el país para poder entrenar el modelo, para ellos se contactó al INEGI para solicitar dicha información a través de un servicio de laboratorio de datos, el cual tenía un costo y nos entregaban los resultados hasta nivel manzana de toda la república mexicana, este proceso duro alrededor de dos semanas, pero de la extracción de datos pudimos conseguir las siguientes variables:

- Clave de entidad federativa
- Entidad federativa

- Clave de municipio o demarcación territorial
- Municipio o demarcación territorial
- Clave de localidad
- Localidad
- Clave del AGEB
- Clave de manzana
- Población de 14 años y más ocupada
- Personas de referencia sin escolaridad
- Personas de referencia con preescolar
- Personas de referencia con primaria
- Personas de referencia con secundaria
- Personas de referencia con preparatoria o bachillerato general
- Personas de referencia con bachillerato tecnológico
- Personas de referencia con estudios técnicos o comerciales con primaria terminada
- Personas de referencia con estudios técnicos o comerciales con secundaria terminada
- Personas de referencia con estudios técnicos o comerciales con preparatoria terminada
- Personas de referencia con normal con primaria o secundaria terminada
- Personas de referencia con normal de licenciatura
- Personas de referencia con licenciatura
- Personas de referencia con especialidad
- Personas de referencia con maestría
- Personas de referencia con doctorado
- Grado promedio de escolaridad
- Viviendas particulares habitadas que disponen de excusado con sanitario
- Viviendas particulares habitadas que no disponen de automóvil o camioneta
- Viviendas particulares habitadas que disponen de automóvil o camioneta
- Viviendas particulares habitadas con un dormitorio
- Viviendas particulares habitadas con dos dormitorios
- Viviendas particulares habitadas con tres dormitorios

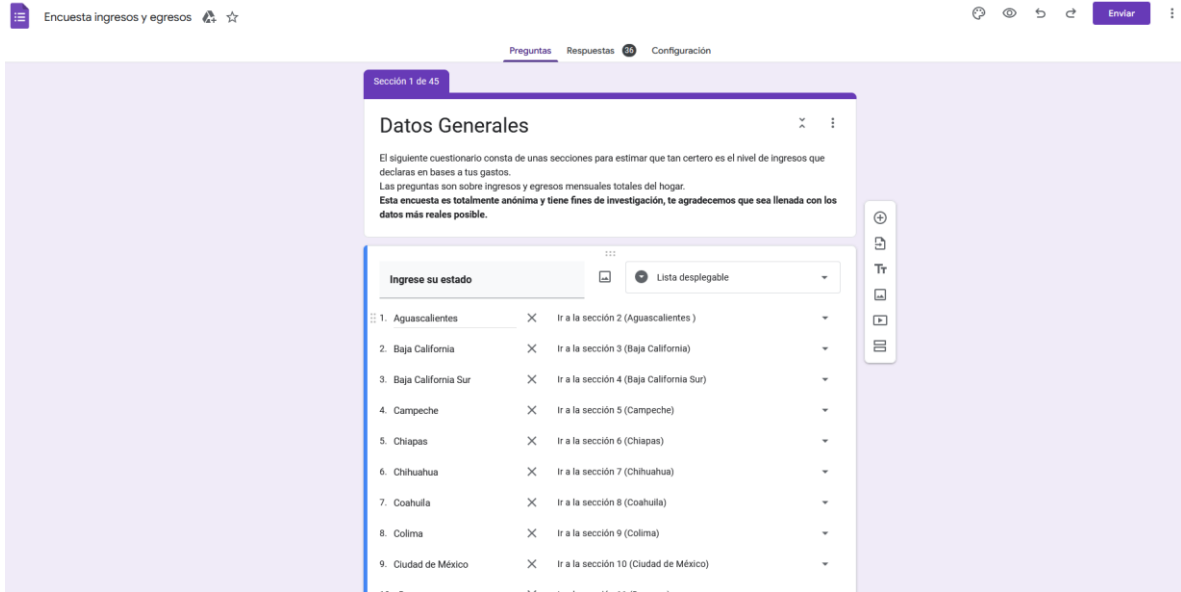
- Viviendas particulares habitadas con cuatro dormitorios y más

Con esta información fue de gran utilidad para entrenar dicho modelo y que su precisión fuera muy elevada.

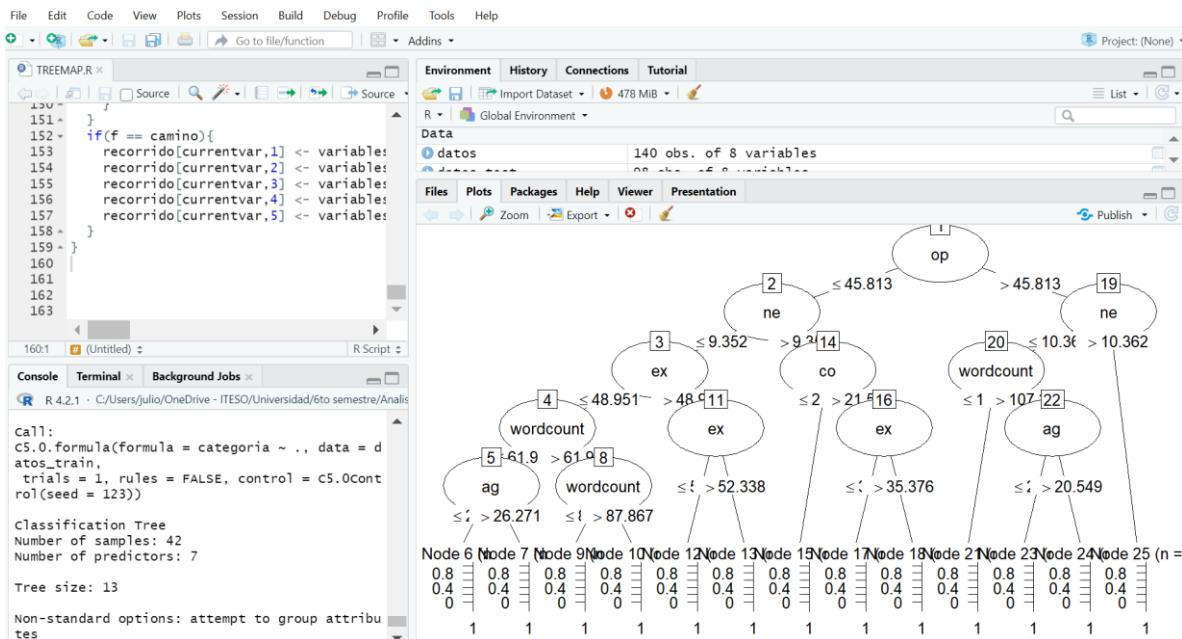
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ENT	NOM_ENT	MUN	NOM_MUN	LOC	NOM_LOC	AGEB	POCUPADA	PJEFE_SE	PJEFE_PRES	PJEFE_PRIM	PJEFE_SEC	PJEFE_PREP	PJEFE_BTEC	PJEFE_TPRIM/PJEF
2	1	Aguascalient	1	Aguascalient	1	Aguascalient	17	1077	9 *		97	346	116	16 *	*
3	1	Aguascalient	1	Aguascalient	1	Aguascalient 006A		655	0 *		13	31	51	7 *	
4	1	Aguascalient	1	Aguascalient	1	Aguascalient	106	1661	0 *		26	67	166	19	5
5	1	Aguascalient	1	Aguascalient	1	Aguascalient	163	1362	8	0	30	213	265	22	0
6	1	Aguascalient	1	Aguascalient	1	Aguascalient	182	1072	0	0	15	39	67	9	0
7	1	Aguascalient	1	Aguascalient	1	Aguascalient	229	204 *			0	22	9	15	0 *
8	1	Aguascalient	1	Aguascalient	1	Aguascalient	233	774	15	0	79	65	45	15	4
9	1	Aguascalient	1	Aguascalient	1	Aguascalient	286	1609	6	0	154	106	122	9	16
10	1	Aguascalient	1	Aguascalient	1	Aguascalient	290	851	12	3	86	61	51	7	19
11	1	Aguascalient	1	Aguascalient	1	Aguascalient	303	1101	20	0	154	122	106	5	18
12	1	Aguascalient	1	Aguascalient	1	Aguascalient	322	1113	15	3	177	111	86	17	16
13	1	Aguascalient	1	Aguascalient	1	Aguascalient	337	625	22 *		115	85	44	8	5
14	1	Aguascalient	1	Aguascalient	1	Aguascalient	341	882	10 *		160	87	71	9	13
15	1	Aguascalient	1	Aguascalient	1	Aguascalient	356	624	9	0	92	78	46	12	5
16	1	Aguascalient	1	Aguascalient	1	Aguascalient	360	854	13	0	156	123	71	7	9
17	1	Aguascalient	1	Aguascalient	1	Aguascalient	375	458	3	0	70	53	35 *		6
18	1	Aguascalient	1	Aguascalient	1	Aguascalient 038A		1468	38 *		317	183	145	30	9
19	1	Aguascalient	1	Aguascalient	1	Aguascalient	407	282	0	0	11	21	27 *		*
20	1	Aguascalient	1	Aguascalient	1	Aguascalient	411	364	10 *		61	54	23 *		*
21	1	Aguascalient	1	Aguascalient	1	Aguascalient	426	1037	33 *		215	170	55	12	9

En la espera de la recolección de datos que se esperaba por parte del INEGI, ayude en el desarrollo de un formulario en Google forms para el modelo de “Cuestionario de confiabilidad” el cual nos ayudaría a evaluar el desempeño de dicho modelo y calibrar los parámetros que considera si una persona miente en sus gastos.

El desarrollo de esta encuesta tuvo varios módulos de validación que dependiendo las respuestas de las personas era al siguiente modulo que avanzarían, este formulario cuenta con 89 preguntas totales pero el usuario al final solo respondería 20 como máximo dependiendo de cuales hayan sido sus respuestas.



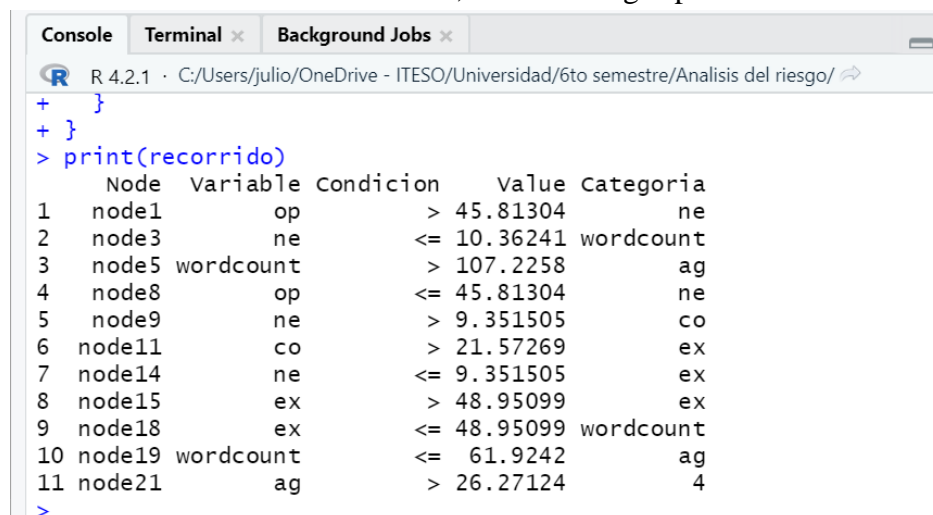
Durante el proceso de estos dos *sprints* también se estuvo explorando como se comporta un algoritmo de clasificación, C5.0 es un algoritmo que crea árboles de clasificación simple, basado en reglas, esto con el objetivo de ver como se desplazan los usuarios en el modelo de “mb-score-model” y poder crear recomendaciones sobre como mejorar su calificación crediticia, buscado en el árbol el camino más corto por el cual el usuario tenia que pasar para obtener una calificación mayor a la que obtuvo.



En este punto ya podíamos ver el árbol en desarrollo, solo faltaba ver como poder movernos entre las ramas para encontrar el camino más corto a la mejor calificación, para ello fue necesario encontrar las validaciones que debería cumplir para movernos entre las ramas

```
> print(dtreestring)
[1] "op > 45.81304:"
[2] "ne > 10.36241: 6 (3/1)"
[3] "ne <= 10.36241:"
[4] "wordcount <= 107.2258: 2 (3)"
[5] "wordcount > 107.2258:"
[6] "ag <= 20.54886: 1 (9/1)"
[7] "ag > 20.54886: 2 (3/1)"
[8] "op <= 45.81304:"
[9] "ne > 9.351505:"
[10] "co <= 21.57269: 1 (4/1)"
[11] "co > 21.57269:"
[12] "ex <= 35.37618: 8 (3)"
[13] "ex > 35.37618: 1 (2/1)"
[14] "ne <= 9.351505:"
[15] "ex > 48.95099:"
[16] "ex <= 52.33754: 2 (3)"
[17] "ex > 52.33754: 4 (2)"
[18] "ex <= 48.95099:"
[19] "wordcount <= 61.9242:"
[20] "ag <= 26.27124: 6 (3)"
[21] "ag > 26.27124: 4 (2)"
[22] "wordcount > 61.9242:"
[23] "wordcount <= 87.8672: 9 (3/1)"
[24] "wordcount > 87.8672: 4 (2/1)"
> |
```

El siguiente paso era encontrar el camino que deberíamos seguir si queremos encontrar el camino más corto al valor buscado, el cual se logro plasmar en un *DataFrame*.



```
Console Terminal x Background Jobs x
R 4.2.1 · C:/Users/julio/OneDrive - ITESO/Universidad/6to semestre/Analisis del riesgo/
+ }
+ }
> print(recorrido)
  Node Variable Condicion Value Categoria
1 node1      op          > 45.81304      ne
2 node3      ne          <= 10.36241 wordcount
3 node5 wordcount          > 107.2258      ag
4 node8      op          <= 45.81304      ne
5 node9      ne          > 9.351505      co
6 node11     co          > 21.57269      ex
7 node14     ne          <= 9.351505      ex
8 node15     ex          > 48.95099      ex
9 node18     ex          <= 48.95099 wordcount
10 node19 wordcount          <= 61.9242      ag
11 node21     ag          > 26.27124      4
>
```

Aquí podemos observar el nodo en el cual empezamos y vemos la condición y el valor de debe cumplir para pasar al siguiente nodo, y así sucesivamente hasta llegar al camino 11 que era el esperado. No pudimos continuar más adelante con esta tarea porque surgieron unas actividades con mayor importancia.

Sprint 6:

Se realizaron 4 archivos tipo *json* en los cuales podemos encontrar recomendaciones para arrendadoras sobre acciones que pudieran tomar o avisos, que les ayude a obtener mejor garantía al momento de firmar un contrato de arrendamiento.

```
1 {
2
3   "historial_crediticio": "Se recomienda solicitar algunos meses de renta por adelantado debido a su historial
4   "legal_no_id_vigente": "La identificación del inquilino no se encuentra vigente. Se recomienda firmar hasta
5   "legal_fiador_no_id_vigente": "La identificación del fiador no se encuentra vigente. Se recomienda firmar r
6   "legal_extanjero_no_forma": "Se recomienda firmar una vez que se el solicitante cuente con su forma migratori
7   "legal_pld": "Se recomienda extremar precauciones por coincidencias en listas negras",
8   "legal_pld_normales": "Se recomienda tomar precauciones por coincidencia en listas negras",
9   "legal_penales": "Se recomienda extremar precauciones debido a que presenta denuncia penal",
10  "legal_representante": "Se recomienda que sea el representante legal con facultades quien firme",
11  "capacidad_80_menos": "Se recomienda reforzar garantía debido a la capacidad de pago",
12  "capacidad_ingresos_efectivo": "El solicitante dice recibir sus ingresos en efectivo, por lo que no pudieron
13  "capacidad_ingresos_2omas": "Los ingresos demostrados son del solicitante y de otra persona, por lo que se re
14  "capacidad_ingresos_otrapersona": "Los ingresos demostrados son de otra persona por lo que se recomienda ql
15  "ingresos_sin_comprobar": "No se pudieron comprobar los ingresos, por lo que se recomienda reforzar garanti
16  "garantia_inmueble_compartido": "Se considera el 100% en garantía siempre y cuando firmen todos los dueños de
17  "garantia_gravamen": "Se recomienda sustituir la garantía o firmar un convenio de justicia alternativa par
18  "garantia_otrapersona": "Se recomienda que sea el dueño del inmueble quien firme como fiador. Firmar un conve
19  "garantia_solo_obligado": "Se recomienda complementar la garantía o firmar un convenio de justicia alternat
20  "garantia_no": "Se recomienda firmar un convenio de justicia alternativa para aumentar la certeza o solícite
21 }
```

El archivo de inquilino nacional funciona como las razones por las cuales se obtuvo el *score* sobre su historial crediticio, el cual cuenta con las razones del *score* para persona física y persona moral, junto con las recomendaciones que deberían seguir para mejorar su puntuación.


```

[{"PF": {
  "Razones_Score": {
    "credito_max_ab": "Se ha estado muy cerca del límite de crédito",
    "credito_max_ce": "Se ha estado muy cerca del límite de crédito",
    "credito_max_revolvente": "Se ha estado muy cerca del límite en créditos revolventes",
    "credito_max_limite_norevolvente": "Se ha(n) tenido atraso(s) en algún pago de las cuentas no revolventes",
    "credito_max_rev_sinlimite": "Se ha utilizado mucho las cuentas revolventes que no poseen límite de crédito",
    "saldo_actual_ab": "El saldo actual de créditos activos es alto",
    "saldo_actual_ce": "Existen saldos en cuentas cerradas",
    "saldo_actual_revolvente": "El saldo actual en créditos revolventes es alto",
    "saldo_actual_limite_norevolvente": "Aún existe un saldo alto por pagar en cuentas no revolventes",
    "saldo_actual_rev_sinlimite": "El saldo actual en créditos revolventes sin límite preestablecido",
    "saldo_actual_max": "El mayor saldo actual es alto",
    "saldo_vencido_ab": "Existe deuda vencida en cuentas actuales",
    "saldo_vencido_ce": "Existe deuda vencida en cuentas cerradas",
    "saldo_vencido_revolvente": "Existe deuda vencida en cuentas revolventes",
    "saldo_vencido_limite_norevolvente": "Existe deuda vencida en cuentas no revolventes",
    "saldo_vencido_rev_sinlimite": "Existe deuda vencida en cuentas sin límite preestablecido",
    "saldo_vencido_max": "Existe deuda vencida",
    "limite_credito_revolvente": "El límite de crédito revolvente obtenido es bajo aún",
    "limite_credito_limite_norevolvente": "No hay control",
    "limite_credito_max": "El mayor límite de crédito es bajo aún",
    "tiempo_ultimo_atraso": "Existen atrasos recientes en los pagos",
    "creditos_aprobados_6m": "Hay cuentas abiertas recientemente",
    "num_aperturas_3meses": "Hay cuentas abiertas recientemente",

```

```

    "atrasos_24m": "Existen atrasos recientes en pagos"
  },
  "Recomendaciones": {
    "credito_max_ab": "No usar más del 70% del límite de crédito, se puede solicitar una ampliación del mismo.",
    "credito_max_ce": "No usar más del 70% del límite de crédito",
    "credito_max_revolvente": "No usar más del 70% del límite en créditos revolventes, solicitar una ampliación",
    "credito_max_limite_norevolvente": "Evitar que ocurran atrasos",
    "credito_max_rev_sinlimite": "No generar un gasto alto en cuentas revolventes antes de pagarlas",
    "saldo_actual_ab": "Reducir el saldo actual en créditos activos",
    "saldo_actual_ce": "Liquidar por completo al cerrar cualquier crédito",
    "saldo_actual_revolvente": "Reducir endeudamiento respecto a el límite en créditos revolventes",
    "saldo_actual_limite_norevolvente": "Realizar pagos por adelantado en cuentas no revolventes, ayudará a liqu",
    "saldo_actual_rev_sinlimite": "Reducir el endeudamiento en créditos revolventes sin límite preestablecido a",
    "saldo_actual_max": "Reducir el endeudamiento respecto al límite de crédito",
    "saldo_vencido_ab": "Pagar el saldo vencido en cuentas actuales",
    "saldo_vencido_ce": "Liquidar por completo al cerrar cualquier crédito",
    "saldo_vencido_revolvente": "Priorizar pagar deudas vencidas en cuentas revolventes y evitar atrasos nuevamen",
    "saldo_vencido_limite_norevolvente": "Priorizar pagar deudas vencidas en cuentas no revolventes y evitae a",
    "saldo_vencido_rev_sinlimite": "Priorizar pagar deudas vencidas en cuentas sin límite y evitar atrasos nueva",
    "saldo_vencido_max": "Priorizar pagar la deuda vencida y evitar atrasos nuevamente",
    "limite_credito_revolvente": "Buscar obtener el máximo de límite de crédito que sea posible, realizando pa",
    "limite_credito_limite_norevolvente": "No recomendación",
    "limite_credito_max": "Entre mayor sea el límite de crédito aprobado que se obtenga, aumentará el SCORE",
    "tiempo_ultimo_atraso": "Esperar un tiempo y evitar que ocurran nuevos atrasos",
    "creditos_aprobados_6m": "Esperar un tiempo sin abrir cuentas",

```

Como se había mencionado anteriormente se realizó una encuesta que ayudaría a evaluar el modelo “Cuestionario de confiabilidad” con datos reales, esto para obtener los límites de los parámetros a los cuales una persona pudiera en ciertas variables. Una vez que se obtuvieron las respuestas, se empezó a trabajar con esa base de datos.

Primero se calculó la diferencia entre los resultados de los gastos que se obtuvieron en la encuesta, los cuales se agruparon por ciudad, a esos valores se les restaría el gasto promedio por ciudad de los gastos del hogar que se había conseguido en la base de datos del INEGI, que nos dio como resultado lo siguiente:

The screenshot shows a Jupyter Notebook with the following code in the cell above the table:

```

data_diff['Diferencia entretenimiento'] = c_den
data_diff['Ahorra dinero'] = c_ad
data_diff['Monto de ahorro'] = c_ma
data_diff

```

The table below shows the output of the code, with columns: Ciudad, Hijos en escuela, Número de hijos, Diferencia educación, Diferencia food, and Di.

Ciudad	Hijos en escuela	Número de hijos	Diferencia educación	Diferencia food	Di
co Zapopan	Si	2	252.152778	-857.518400	-
co Coyoacan	No	0	-2800.000000	-2136.714204	-
co Zapopan	No	0	-500.000000	-1348.947239	-
ma Manzanillo	No	0	-2800.000000	-2617.714303	-
co Gustavo A. Madero	No	0	-500.000000	-974.810565	-
co Tlalpan	No	0	-3000.000000	-703.914426	-
co Zapopan	Si	1	-2200.000000	651.052761	-
po Morelia	No	0	-3000.000000	-2493.595481	-
co Zapopan	Si	1	-234.334936	2590.866667	-
co San Pedro Tlaquepaque	Si	2	-204.166667	-893.741071	-
co Zapopan	Si	2	25787.092391	-257.471113	-
as Victoria	Si	2	-1601.463038	-1741.357476	-
ma Manzanillo	Si	2	-2528.703112	-1640.004000	-

The screenshot shows a Jupyter Notebook with the following code in the cell above the table:

```

data_diff['Diferencia entretenimiento'] = c_den
data_diff['Ahorra dinero'] = c_ad
data_diff['Monto de ahorro'] = c_ma
data_diff

```

The table below shows the output of the code, with columns: Diferencia servicios, Diferencia salud, Diferencia deudas, Diferencia transporte, and Diferencia entreten.

	Diferencia servicios	Diferencia salud	Diferencia deudas	Diferencia transporte	Diferencia entreten
3	806.421176	-2.394790e+03	-1882.205770	-1448.214286	-1
4	-58.844583	-7.407270e+02	16.393333	1777.726471	-1
9	76.511026	-9.337515e+02	-5996.578789	1160.786061	-1
3	-798.543385	-8.542251e+02	-79854.639551	-1051.156288	-4
5	1060.575455	-4.083000e+01	-999.305808	99.731000	-1
6	-765.486190	-2.934950e+03	0.694192	-945.366471	-1
1	-923.488974	-1.433751e+03	-7996.578789	160.786061	-1
7	-4068.736667	-4.617065e+02	0.694192	389.502187	-2
1	620.960000	-4.337196e+02	0.054527	721.730000	-1
1	-948.542162	-3.671175e+02	-1496.578789	2075.809767	-
3	1239.546618	-1.533985e+03	841.285362	2378.524571	-3
6	-884.710000	-1.883086e+03	-3999.305808	-1432.898421	-
3	3625.470000	-9.217400e+02	-9999.512999	-3419.360000	-1

Con esta información ya podíamos observar que los gastos promedio de la base de datos del INEGI están un poco sesgados y por lo que se tenía que ajustar el modelo para que considere

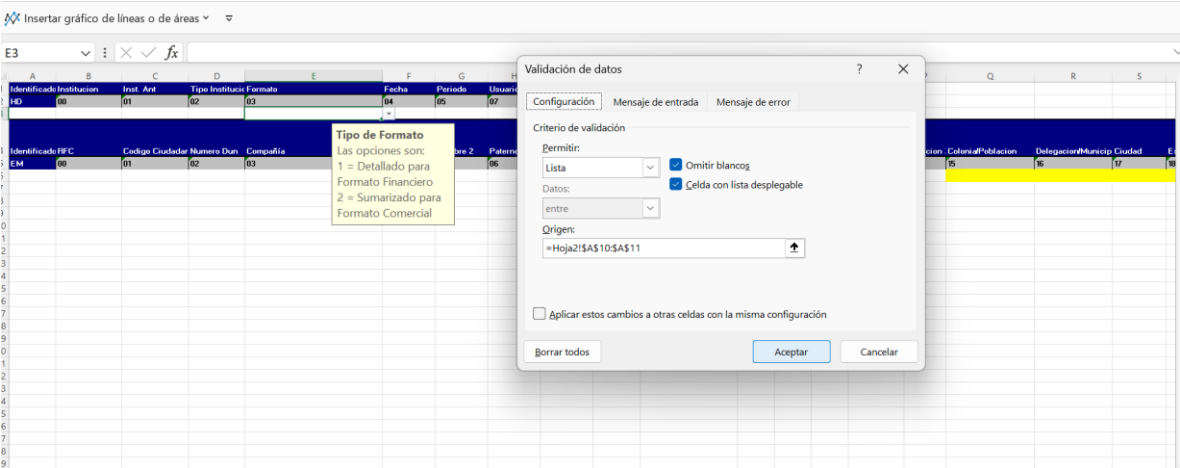
que tanto se pueden mover estas variables y no clasificar erróneamente a los solicitantes de créditos. Por lo que se calculó las diferencias estandarizadas para saber el margen en el cual se podían mover los usuarios al momento de reportar sus gastos y los puedan seguir considerando como verdadero lo que está reportando.

	Promedio diferencias estandarizadas
Diferencia educación estandar	0.888889
Diferencia food estandar	0.472222
Diferencia servicios estandar	0.333333
Diferencia salud estandar	0.333333
Diferencia deudas estandar	0.277778
Diferencia transporte estandar	0.500000
Diferencia entretenimiento estandar	0.388889
Monto de ahorro estandar	0.166667

Sprint 7:

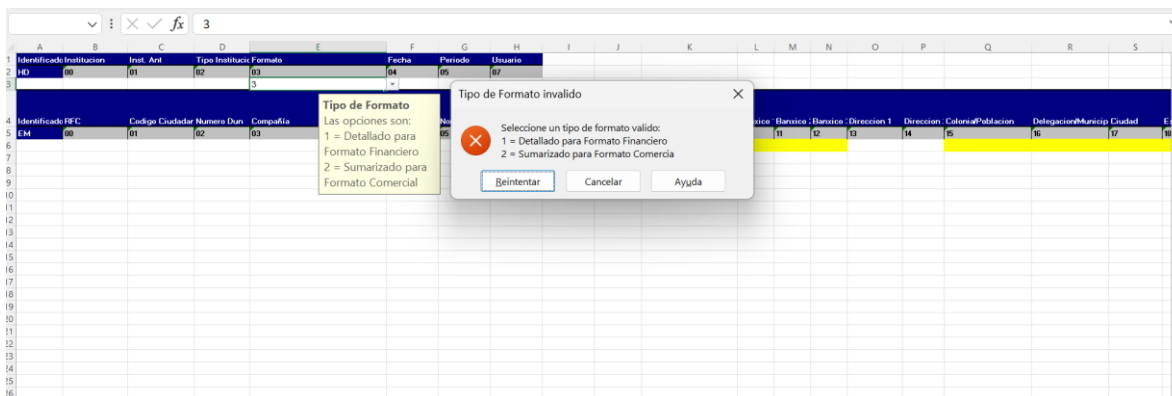
La última actividad realiza en este proyecto fue el desarrollar un layout automatizado con las restricciones que se deben de cumplir al momento de ser realizado, en el cual fue basado con las reglas especificadas en el Manual Técnico de Entrega de Buro de Crédito para personas morales.

Como primer paso fue el realizar todas las validaciones a todos los campos del layout, Excel nos permite crear validaciones para celdas cuyo valor solo será aceptado cuando se cumplan con reglas específicas para cada uno.



Como podemos ver en la imagen anterior, la celda señalada tiene como restricción que solo podrá ser llenado con valores de una lista creada y no aceptará otro valor que no se encuentre

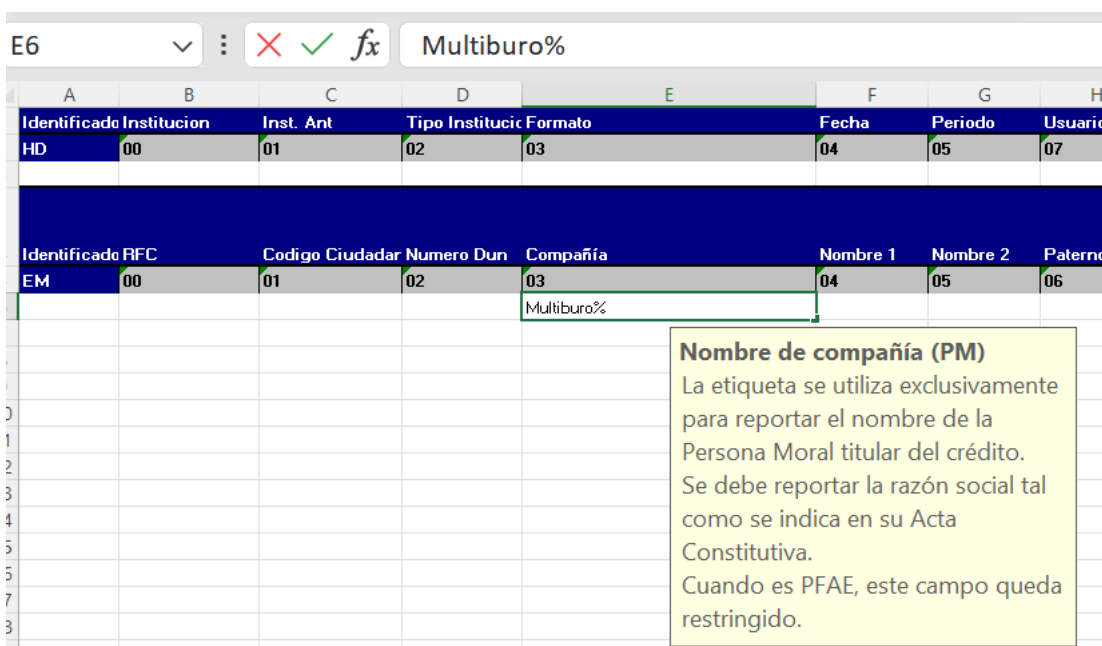
en la lista, también al seleccionar una celda saldrá una pequeña leyenda de los valores que acepta y las opciones que se tiene en el caso que tenga opciones, de lo contrario solo mostrara las reglas o el formato que debe cumplir dicho campo.



Aquí podemos observar que si intento agregar un valor diferente a la de esa lista nos arroja un mensaje de error con una breve descripción de los valores que están permitidos.

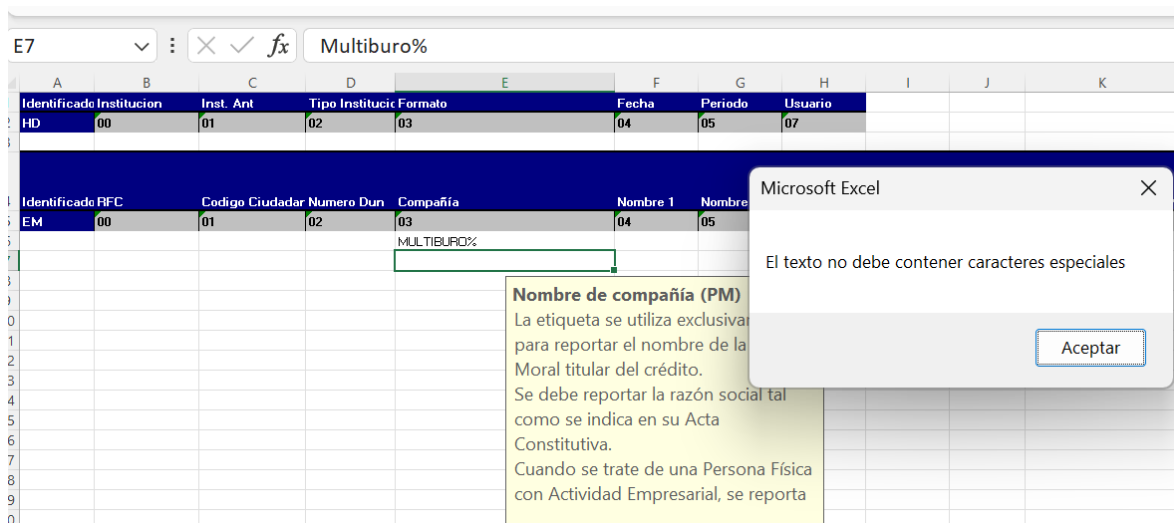
El layout está automatizado para que todos valores automáticamente después de ser llenados, en caso de ser minúsculas la transforme en mayúsculas y detecte algún carácter especial que no está permitido usar. Esto se puede ejemplificar mejor en las dos imágenes siguientes.

Antes:



Como se muestra en el ejemplo al yo escribir el nombre de una empresa con minúsculas y con un carácter especial que no esta permitido, al momento de yo dar enter ocurre lo siguiente:

Después:



El texto me lo convierte automáticamente en mayúsculas y me arroja un mensaje mencionando que el texto no debe contener caracteres especiales.

Esto facilita a los usuarios al momento de hacer su reporte del historial de pagos que los datos cumplan con todos los criterios de evaluación y el formato que acepta Buro de Crédito, para que así sean aceptados en su primer intento.

1.6. Valoración de productos, resultados e impactos

Realizar este proyecto PAP fue una experiencia de alto aprendizaje, el cual dedico de mucho esfuerzo para lograr dar soluciones eficientes a los problemas que se iban presentado. Se adquirieron conocimientos de nuevas herramientas como Docker, conexión a una API, se reforzo los conocimientos en los repositorios en Github, ver como es el desarrollo de una *startup* y cómo funciona su modelo de negocio y la evolución que van teniendo dependiendo del mercado al cual se están enfocando.

Como primero resultados se agilizo la búsqueda de errores dentro de los archivos json con lo que se trabajan, al igual que ayudar a probar, entrenar y evaluar diferentes modelos de *score* lo cual acelero su etapa para que estuvieran listo para que entraran como nuevos productos que la empresa va a lanzar al mercado, así como la recolección de datos y automatización de reportes para minimizar errores y que las bases de datos presente una calidad superior al 95% en la calidad de los datos.

Estas actividades mencionadas dan como resultado para la empresa herramientas que podrán seguir utilizando en un futuro, las cuales ayudarán a minimizar los tiempos de respuesta en errores y procesamiento de datos.

1.7. Bibliografía y otros recursos

- Mexico - *Encuesta Nacional de Ingresos y Gastos de los Hogares 2020, Nueva serie.* (2020). INEGI. https://www.inegi.org.mx/rnm/index.php/catalog/685/data-dictionary/F28?file_name=concentradohogar
- Comité de Niveles Socioeconómicos AMIA. (2021). Nivel Socioeconómico AMAI 2022. NSE. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjVns_N5tf7AhXQk2oFHR-rCMoQFnoECAwQAQ&url=https%3A%2F%2Fwww.amai.org%2Fdescargas%2FNota_Metodologico_NSE_2022_v5.pdf&usg=AOvVaw0IOF4xRSfrXgHBNUenfQo
- Rodrigo, J. A. (s. f.). *Arboles de decision, Random Forest, Gradient Boosting y C5.0.* https://www.cienciadedatos.net/documentos/33_arboles_decision_random_forest_gradient_boosting_c50

1.8. Anexos generales

2. Productos

Formulario en Google forms para el modelo de “Cuestionario de confiabilidad”
https://docs.google.com/forms/d/1FIT4E7k5OM6JeksuzvO8FJd3lGgeR9926SD7s4r48/edit?allow_large_form

Archivo con lenguaje Python con lenguaje de programación con los códigos procesamiento de datos, evaluación de modelos, exploración del modelo de clasificación en R, conexión a la Api de Google maps mediante python, archivos json, manual buro, Layout automatizado.
https://iteso01my.sharepoint.com/:f:/g/personal/if714007_iteso_mx/EnZ_n2RRz0BNsxuoQwfxDGEBUBVcYYMBkJFye1AYHyfRSA?e=nq1six

3. Reflexión crítica y ética de la experiencia

El RPAP tiene también como propósito documentar la reflexión sobre los aprendizajes en sus múltiples dimensiones, las implicaciones éticas y los aportes sociales del proyecto para compartir una comprensión crítica y amplia de las problemáticas en las que se intervino.

3.1 Sensibilización ante las realidades

Unos de obstáculos al tratar de conseguir la información que necesitábamos y conseguirla en forma que queríamos, ya que eso ocurre en la vida real podemos obtener bases de datos que nos están procesadas, lista para trabajar, para ello necesitábamos realizar limpieza de datos y un preprocesamiento de los datos para poder trabajar con ellos. Actualmente en la vida laboral esto es lo que ocurre no es fácil acceder a la información que uno busca de manera segura, confiable y completa, requiere de un trabajo de investigación y recolección.

Una vez conseguido la información el trabajar con ella nos damos cuenta que no somos expertos en lo que hacemos día a día siempre encontramos nuevas formas de realizar las cosas de una manera más rápida y eficiente.

Siempre habrá nuevos retos a los cuales no estamos preparados pero el actualizarnos e investigar nos llevará a la superación de estos retos con éxito.

3.2 Aprendizajes logrados

Aprendí el trabajo en conjunto con programadores sobre las herramientas básicas que debo adquirir como el manejo de Docker, la conexión con procesos ya establecido al área de producción, aprendí las diferentes técnicas de solucionar problemas cuando estamos atorados al momento de programar y que nunca esta mal visto pedir consejos o ayuda a alguien con alguna línea de código en las que estamos atorados.

Puse en practico mis conocimientos de programación para procesamiento de datos capas de obtener una calidad y limpieza de los datos optima para trabajar con modelos de *score* crediticos, para poder así evaluarlos, de igual forma mis conocimiento en programación crecieron así como mis habilidades en Excel.

Gracias a todas las experiencias vividas a lo largo de este proyecto me ayudaron a entender, saber y prepararme a la vida laboral en la cual me quiero encaminar, me quedo con una gran motivación gracias a los excelentes compañeros de trabajo dentro de la empresa los cuales me enseñaron una de las mejores convivencias laborales, aconsejándome y aprendiendo de cada uno de ellos, en un ambiente laboral sano y respetuoso.