

**MAESTRIA EN DATA SCIENCE**



***CURSO***: MODELOS GRAFICOS PROBABILISTICOS

***TEMA****:*

“MODELO PROBABILÍSTICO DE RECLAMOS REGISTRADOS EN LA CADENA DE RESTAURANTES FRIDAYS”

***PROFESOR***: Patricia Daniela Reyes Silva

***EQUIPO***: 3

***INTEGRANTES:***

Carlos Enrique Belloda Saavedra e202210911

Christian Enrique Mercado Justo e202210783

Rossmery Mercedes Salvatierra Barrantes e202210810

2023

Contenido

[Introducción 6](#_Toc152711802)

[1. Definición del problema 6](#_Toc152711803)

[2. Definición del conjunto de datos 6](#_Toc152711804)

[2.1. Origen de los datos 6](#_Toc152711805)

[2.2. Características de los datos 6](#_Toc152711806)

[3. Propuesta del modelo 7](#_Toc152711807)

[3.1. Objetivo 7](#_Toc152711808)

[3.2. Técnicas y metodologías 7](#_Toc152711809)

[3.3. Representación del modelo 8](#_Toc152711810)

[3.4. Tabla de distribución de probabilidades 9](#_Toc152711811)

[3.5. Preprocesamiento de datos 1](#_Toc152711812)1

[3.6. Generación de CPDs a partir de la data 12](#_Toc152711813)

[3.7. Inferencias 13](#_Toc152711814)

[4. Conclusiones 15](#_Toc152711815)

[5. Bibliografía 16](#_Toc152711816)

Índice de ilustraciones

[Ilustración 1 Ejemplo Archivo Json 4](#_heading=h.1t3h5sf)

índice de tablas

[Tabla 1- Diccionario de datos 6](#_heading=h.4d34og8)

# Introducción

El presente proyecto se enfoca en el análisis del sistema de libro de reclamaciones de la cadena americana de restaurante Fridays, la cual es franquiciada en Perú a través de la empresa peruana Franquicias Alimentarias S.A.

El numero locales en el 2023 asciende a 19 tiendas, distribuidas 17 en Lima, 1 en Arequipa y 1 en Trujillo.

El proyecto busca proponer un modelo probabilístico que ayude analizar los puntos de dolor o mayores factores que hacen que un cliente formalice un reclamo, a través del libro de reclamaciones.

# Definición del problema

A mediados de octubre del 2021 la compañía implemento el sistema de libro de reclamación a través de su pagina web, donde se viene registrando y procesando los reclamos de los clientes lo cual afecta directamente el nivel de satisfacción de los clientes (NPS), por lo cual, se necesita entender el efecto las diversas variables del negocio hacia un posible reclamo, lo que nos lleva a la búsqueda de un modelo que nos ayude a conocer la probabilidad de un posible reclamo conjugando las variables encontradas.

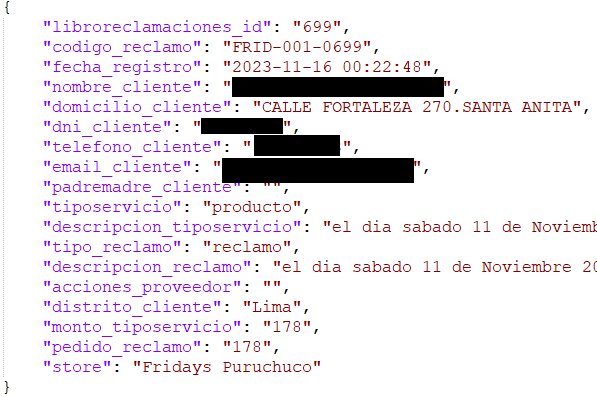
# Definición del conjunto de datos

## Origen de los datos

Los datos se obtienen del sistema de libro de reclamaciones que puede ser accedido por clientes a través de la pagina web de la compañía en [www.fridaysperu.com](http://www.fridaysperu.com) y en el enlace libro de reclamaciones que se encuentren el pie de página acompañado del logo normado por INDECOPI

## Características de los datos

Los datos son obtenidos a través de un servicio web que devuelve un archivo Json con la información detallada del reclamo.



*Ilustración 1 Ejemplo Archivo Json*

Los datos poseen el siguiente diccionario

|  |  |  |
| --- | --- | --- |
| Campo | Tipo | Descripción |
| libroreclamaciones\_id | Entero | Identificador interno del reclamo |
| codigo\_reclamo | Alfanumérico | Código del reclamo |
| fecha\_registro | Fecha | Fecha del reclamo |
| nombre\_cliente | Cadena | Nombre del cliente |
| domicilio\_cliente | Cadena | Dirección del cliente |
| dni\_cliente | Cadena | Documento de identidad del cliente |
| telefono\_cliente | Numérico | Número de teléfono del cliente |
| email\_cliente | Cadena | Correo |
| padremadre\_cliente | Cadena | Nombre del padre o madre si fuese menor de edad |
| tiposervicio | Cadena | Tipo de servicio brindado |
| descripcion\_tiposervicio | Cadena | Descripción del servicio |
| tipo\_reclamo | Cadena | Tipo de reclamo |
| descripcion\_reclamo | Cadena | Descripción del reclamo |
| acciones\_proveedor | Cadena | Acciones solicitadas hacia el proveedor |
| distrito\_cliente | Cadena | Distrito del cliente |
| monto\_tiposervicio | Decimal | Monto reclamado |
| pedido\_reclamo | Decimal | Monto de pago efectuado |
| store | Cadena | Tienda implicada en el reclamo |

*Tabla 1- Diccionario de datos*

# Propuesta del modelo

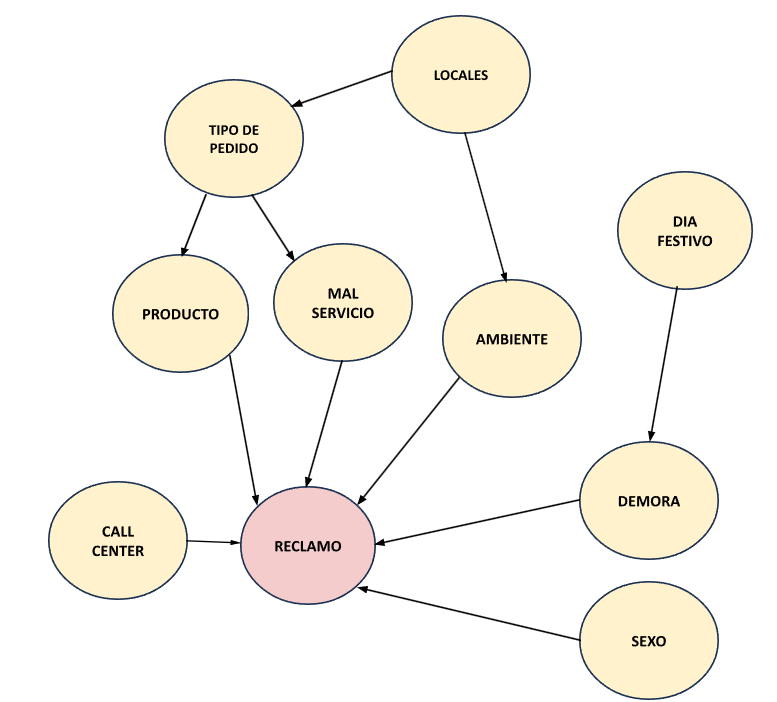
## Objetivo

Se quiere evaluar, y encontrar los factores principales que hacen que un cliente Fridays presente un reclamo.

## Técnicas y metodologías

Un modelo bayesiano es un tipo de modelo estadístico que utiliza el teorema de Bayes para actualizar la probabilidad de una hipótesis a medida que se obtienen nuevos datos o evidencias. Este enfoque se basa en la inferencia bayesiana, que es un marco probabilístico para la actualización de creencias sobre un evento o una hipótesis a medida que se adquiere nueva información.

## Representación del modelo



*Ilustración 2- DAD Reclamos Fridays*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Índice | Nodo | Campo | Valores | Cardinalidades |
| 1 | REG | Locales | 0: Norte  1: Sur | 2 |
| 2 | TDP | Tipo de pedido | 0: Local  1: Delivery | 2 |
| 3 | PROD | Producto | 0: Malo  1: Bueno | 2 |
| 4 | SERV | Servicio | 0: Malo  1: Bueno | 2 |
| 5 | AMB | Ambiente | 0: Agradable  1: Desagradable | 2 |
| 6 | DFEST | Dia Festivo | 0: No  1: Si | 2 |
| 7 | DEM | Demora | 0: No  1: Si | 2 |
| 8 | CALL | Call center | 0: No  1: Si | 2 |
| 9 | SEX | Sexo | 0: Masculino  1: Femenino | 2 |
| 10 | REC | Reclamo | 0: No  1: Si | 2 |

*Tabla 2- Diccionario de variables*

## Tabla de distribución de probabilidades

|  |  |
| --- | --- |
| P(REG) |  |
| REG\_0 | REG\_1 |
| 0.61 | 0.39 |

|  |  |  |
| --- | --- | --- |
| P(TPD|REG) | TPD\_0 | TPD\_1 |
| REG\_0 | 0.68 | 0.32 |
| REG\_1 | 0.27 | 0.73 |

|  |  |  |
| --- | --- | --- |
| P(AMB|REG) | AMB\_0 | AMB\_1 |
| REG\_0 | 0.28 | 0.72 |
| REG\_1 | 0.35 | 0.65 |

|  |  |  |
| --- | --- | --- |
| P(PROD|TPD) | PROD\_0 | PROD\_1 |
| TPD\_0 | 0.05 | 0.95 |
| TPD\_1 | 0.15 | 0.85 |

|  |  |  |
| --- | --- | --- |
| P(SERV|TPD) | SERV\_0 | SERV\_1 |
| TPD\_0 | 0.18 | 0.82 |
| TPD\_1 | 0.05 | 0.95 |

|  |  |
| --- | --- |
| P(DFEST) |  |
| DFEST\_0 | DFEST\_1 |
| 0.16 | 0.84 |

|  |  |  |
| --- | --- | --- |
| P(DEM|DFEST) | DEM\_0 | DEM\_1 |
| DFEST\_0 | 0.92 | 0.08 |
| DFEST\_1 | 0.33 | 0.67 |

|  |  |
| --- | --- |
| P(SEX) |  |
| SEX\_0 | SEX\_1 |
| 0.55 | 0.45 |

|  |  |
| --- | --- |
| P(CALL) |  |
| CALL\_0 | CALL\_1 |
| 0.72 | 0.28 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P(REC|PROD,SERV,AMB,DEM,SEX,CALL) | | |  |  |  |  |  |
| PROD | SERV | AMB | DEM | SEX | CALL | RECL\_0 | RECL\_1 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_0 | SEX\_0 | CALL\_0 | 0.10 | 0.90 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_0 | SEX\_0 | CALL\_1 | 0.75 | 0.25 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_0 | SEX\_1 | CALL\_0 | 0.00 | 1.00 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_0 | SEX\_1 | CALL\_1 | 0.70 | 0.30 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_1 | SEX\_0 | CALL\_0 | 0.32 | 0.68 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_1 | SEX\_0 | CALL\_1 | 0.23 | 0.77 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_1 | SEX\_1 | CALL\_0 | 0.55 | 0.45 |
| PROD\_0 | SERV\_0 | AMB\_0 | DEM\_1 | SEX\_1 | CALL\_1 | 0.65 | 0.35 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_0 | SEX\_0 | CALL\_0 | 0.37 | 0.63 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_0 | SEX\_0 | CALL\_1 | 0.49 | 0.51 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_0 | SEX\_1 | CALL\_0 | 0.19 | 0.81 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_0 | SEX\_1 | CALL\_1 | 0.77 | 0.23 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_1 | SEX\_0 | CALL\_0 | 0.23 | 0.77 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_1 | SEX\_0 | CALL\_1 | 0.72 | 0.28 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_1 | SEX\_1 | CALL\_0 | 0.38 | 0.62 |
| PROD\_0 | SERV\_0 | AMB\_1 | DEM\_1 | SEX\_1 | CALL\_1 | 0.14 | 0.86 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_0 | SEX\_0 | CALL\_0 | 0.66 | 0.34 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_0 | SEX\_0 | CALL\_1 | 0.50 | 0.50 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_0 | SEX\_1 | CALL\_0 | 0.69 | 0.31 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_0 | SEX\_1 | CALL\_1 | 0.19 | 0.81 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_1 | SEX\_0 | CALL\_0 | 0.62 | 0.38 |
| PROD\_0 | SERV\_1 | AMB\_0 | DEM\_1 | SEX\_0 | CALL\_1 | 0.35 | 0.65 |

## 

## Preprocesamiento de datos

Hacemos el preprocesamiento de datos, el campo de donde extraemos es un campo descriptivo, es no estructurado y con base a esta información la transformamos a data estructurada

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

La data obtenida:

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

## Generación de CPDs a partir de la data

Hemos pasado por tres procesos, para la generación de CPDs, escogimos un estimador Bayesiano.

* Haciendo un conteo de valores, observamos datos faltantes, debido a que solo contamos con data de reclamos y existen combinaciones de variables que no se encuentran en la cuenta, pues no hay registro de reclamos por esas condiciones.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

* Usamos máxima verosimilitud, para calcular una mejor distribución de probabilidad, nos enfocamos en el CDP de presenta reclamo:

Tabla

Descripción generada automáticamente

* Para los casos en que no hay registro de reclamo el algoritmo de Máxima verosimilitud llena las probabilidades con el 50 %para Sí y 50% para No Y si hay un registro presente para esa combinación de variables le pone probabilidad 100%. Por tal motivo, optamos por un algoritmo un poco menos concluyente respecto a la probabilidad, usamos un estimador bayesiano

Tabla

Descripción generada automáticamente

## Inferencias

* + 1. **¿Cuál es la probabilidad de que se genere un reclamo de la zona Sur por un mal producto vendido por el canal Delivery?**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* + 1. **Interfaz de usuario gráfica, Texto

       Descripción generada automáticamente¿Cuál es la probabilidad de que se genere un reclamo por demora en un día festivo cuando el pedido fue consumido en el local?**
    2. **¿Cuál de los géneros tiene probabilidad de presentar un reclamo solo por demora?**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* + 1. **¿Qué probabilidad hay que se formalice un reclamo si un cliente llama o es llamado por call center antes de efectuarlo?**

De los clientes que tuvieron contacto con el call center antes de formalizar un reclamo hay una probabilidad de aproximadamente 30% de que no se presente un reclamo:

Diagrama

Descripción generada automáticamente

* + 1. **¿Cuánto impacta un ambiente desagradable en los reclamos?**

Si el ambiente es desagradable, o el cliente percibe el ambiente como desagradable, hay una probabilidad del 97% que se presente un reclamo.

Diagrama

Descripción generada automáticamente con confianza media

## Conclusiones

* Hemos estudiado la interacción entre variables relacionadas a los reclamos de la cadena de restaurante Fridays, los resultados de las CPDs generadas por esta data, se aproximaron a lo formulado inicialmente por el experto.
* Hemos podido, gracias a las redes bayesianas descubrir, la importancia del centro de atención telefónica para amortiguar el impacto de los reclamos.
* Se debe mejorar el modelo, tomando en cuenta un poco más de detalle en las razones por las cuales los clientes presentan un reclamo.

## Bibliografía

* "Probabilistic Graphical Models: Principles and Techniques" de Daphne Koller y Nir Friedman - Este libro cubre una amplia gama de temas relacionados con modelos gráficos probabilísticos, incluidas las redes bayesianas.
* "Bayesian Networks and Decision Graphs" de Thomas Dyhre Nielsen y Finn Jensen - Un libro que proporciona una introducción completa a las redes bayesianas y su aplicación en la modelización probabilística.