



Get Started with Data Science and Python

FrameWork Report

FrameWork



El marco de trabajo a utilizar será el BADIR ya que, según la necesidad del negocio, es la herramienta que mas se ajusta para cumplir con los objetivos planteados.

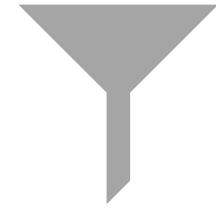


A continuación, se detallan las razones del por que se contempla Badir como la mejor opción y como será utilizado:

Objetivos de negocio



Identificar la causa raíz del por que va en aumento el número de clientes de Credit One que incumple el pago de sus prestamos



Creación de modelos analíticos para generar un scoring y poder así clasificar correctamente los clientes.

Plan de análisis

- Justamente este es uno de los puntos por el cual se definió utilizar el marco de trabajo BADIR, ya que con la información (DataSet) que se cuenta hasta el momento nos permite poder identificar claramente los objetivos del negocio, además de que con base a esto se puede definir una hipótesis y poder definir un plan de trabajo para poder cumplir con los objetivos planteados.

Colección de los datos

- Con respecto a los datos, se cuenta con datos históricos los cuales representan cantidad amplia de registros (30000) además de contener variables importantes que nos ayudarán a poder identificar la causa raíz que nos exige la organización, además de poder generar un Score que permita filtrar los mejores clientes, a partir de acá se procederá a realizar una depuración y validación de los datos, para obtener resultados más certeros.
- Estructura de fuente de datos:
 - Monto del crédito
 - Genero del cliente
 - Nivel de escolaridad
 - Estado civil
 - Edad
 - Pago a tiempo de los últimos 6 meses: Indica si se realizo el pago a tiempo, retrasado por un mes o mas
 - Monto de la factura para los últimos 6 meses
 - Monto de los pagos realizados para los últimos 6 meses
 - Comportamiento del cliente

Gestión de los datos

- Como parte de la administración o gestión de los datos, se debe realizar varias acciones importantes para poder aprovechar de la mejor manera la información contenida en el set de datos, las acciones a ejecutar a los datos serían las siguientes:
- Depuración
- Transformación
- Reducción
- Discretización

Set de datos

- Parte del análisis inicial que se pudo realizar a los datos se puede identificar que existen datos socioeconómicos que podrían ser importante de incorporar y así poder obtener resultados mas valiosos, por ejemplo, datos referentes a su empleabilidad, sus ingresos mensuales, deudas en otras entidades, etc son algunos de datos que podrían ayudar a seleccionar mejor a los clientes.

Descubrimiento

- Una vez hayamos realizado la depuración de la información, procederemos a realizar un análisis inicial de los datos con graficas “simples” para empezar a identificar algunos patrones y poder así realizar hallazgos que posteriormente en conjunto con las hipótesis podremos confirmar con la construcción de los modelos necesarios que justamente nos ayuden a poder dar respuesta a los objetivos planteados por el negocio.

Proceso a seguir



Objetivos de negocio

Definición de objetivos e hipótesis a comprobar



Exploración y Colección de Datos

Conocer los datos

Validación: ¿Tenemos todo lo que necesitamos?

Identificar posibles problemas en el data Set.



Depuración de los Datos

Limpieza de datos.

Aplicar transformación de los datos.



Crear los modelos

Con base al data set depurado y a los objetivos del negocio definidos, procederemos a crear los modelos analíticos necesarios que nos permita poder generar las respuestas y resultados necesarios.



Presentación de resultados

Una vez creados los modelos, con los resultados obtenidos procederemos a presentar dicho insumo al negocio.