

---

---

# Test Bankaya

Por: Angélica León González

---

# Modelo de Originación

## Regresión Logística

Realizamos un modelo que nos ayuda a saber si un cliente es riesgoso o no, la forma en la que se puede implementar el modelo es utilizando una Lambda de AWS, y crear el ETL que genere las variables que utiliza el modelo y así con cada nueva aplicación se realizaría la evaluación del cliente y dependiendo la probabilidad de evento (que no pague), se podría tomar la decisión de rechazarlo, o tal vez darle el préstamo a una tasa más alta.

---

---

# Regresión Logística

Algunas observaciones y mejoras para el modelo son:

- Podemos obtener mejores resultados utilizando más información de la solicitud del crédito, personal, geoespacial, incluso de Buró de Crédito.
  - Podemos generar clusters para segmentar a nuestra población y complementar los resultados de nuestro modelo.
  - El modelo se debe de estar supervisando para en cuanto se tengan más aplicaciones o información, poder re-calibrar o reentrenar el modelo.
  - También podemos utilizar otro tipo de algoritmo más potente, en este caso se utilizó la regresión logística porque parte de la solución es realizar Credit Scoring.
  - La ROC del modelo no es muy alta, así que hay que evaluar sus resultados ya en la operación diaria para saber si se requiere alguna modificación.
-

# Credit Scoring

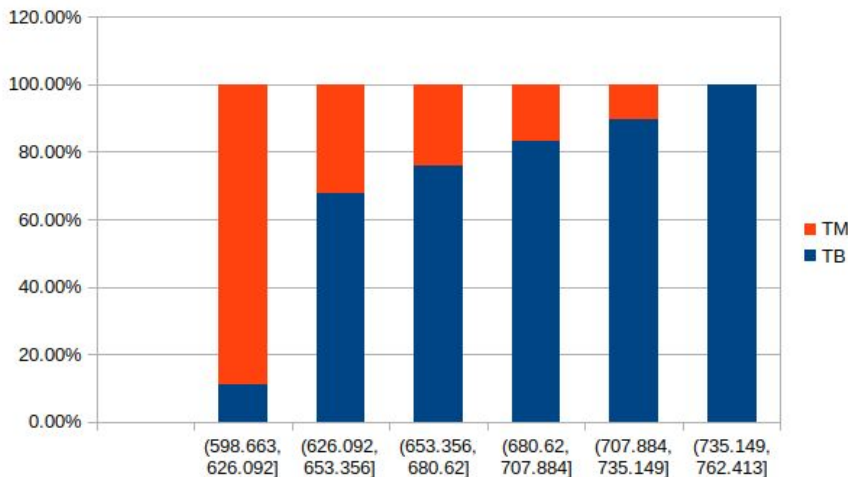
Podemos utilizar sólo el modelo anterior o realizar la evaluación de las solicitudes utilizando nuestro sistema de puntuación.

---

---

# Resultados Credit Scoring

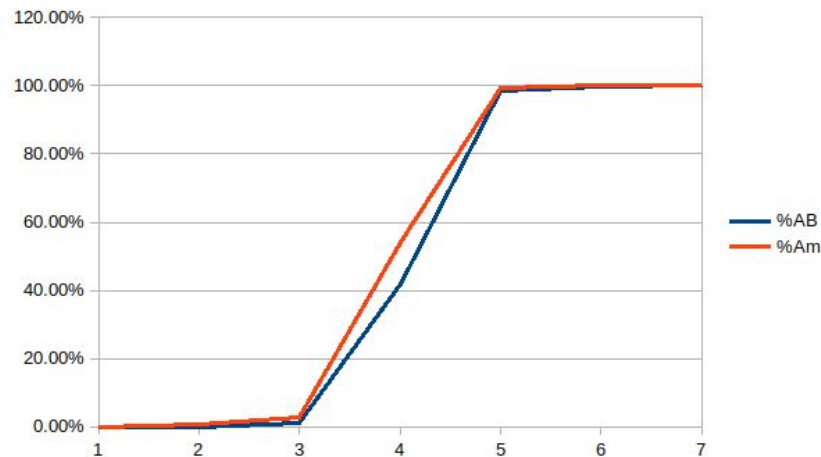
En esta gráfica vemos la distribución de los clientes riesgosos (TM) y los no riesgosos (TB) según la cantidad de puntos que obtengan. Vemos que conforme más alta es la puntuación va desapareciendo la población riesgosa, por lo que nosotros buscaríamos evaluar de esa manera a nuestros nuevos clientes.



---

# Resultados Credit Scoring

Esta gráfica nos muestra el acumulado de créditos buenos y malos en cada rango de score, en este caso tenemos 6 rangos, y en cada uno podemos observar como nuestra población no se mezcla (para ambas poblaciones se inicia en 0% y se termina en 100%). Esto indica que nuestro modelo está separando bien los casos riesgosos de los no riesgosos.



## Ejemplo de uso de la ScoreCard

En la scorecard que generamos tenemos todas nuestras variables con su rango, valor de la variable y puntos correspondientes. Se debe realizar un proceso que evalúe a nuestros clientes nuevos según este sistema de puntos.

Por ejemplo, si llega un cliente que realizó la aplicación en Enero, se le asignaría 59.76 puntos y así sucesivamente con cada variable, al final se hace la suma de todos los puntos y se toma la decisión según el rango en el que haya caído.

Nosotros buscaríamos tener la mayor cantidad de nuestros clientes con una puntuación dentro del rango **[735.149, 762.413]**, pero en caso de estar en alguno de los otros rangos podemos asignar tasas más altas, montos más bajos o a modificar el plazo, según sea el caso.

v_APPLICATION_MONTH	W_v_APPLICATION_MONTH	P_v_APPLICATION_MONTH
1	-0.0662155058133157	59.7602288482832
2	-0.0193828436228043	61.2797853877306
3	0.0437134816155259	63.3270410667047
4	0.148796566282531	66.7366206426316
5	0.380014206359121	74.2388265942111
6	0.459877590083463	76.8301148388778
7	0.182928547226041	67.844084374589
8	0.199894081384337	68.3945567835796
9	-0.153821943814257	56.917705499504
10	-0.352272882538095	50.4786647191494
11	-0.144017943717636	57.2358111070597
12	-0.105344426200443	58.4906318603114