

Dataton Bancolombia 2018

Análítica para manejo de finanzas personales

Equipo: The Plumbers

Introducción

El aprendizaje de máquinas (ML) y la Inteligencia Artificial (AI) se han utilizado en varios contextos de las finanzas como la estructuración de portafolios, modelos de otorgamiento de crédito, trading algorítmico, entre otros, lo cual ha sido fundamental en la transformación de esta industria. Las entidades financieras han estado tomando ventaja del poder de la tecnología innovadora para potenciar sus relaciones con los clientes y generar mayor satisfacción y rentabilidad, al mismo tiempo que mejorando los costos de sus procesos.

No obstante, aunque la inteligencia artificial ya se está utilizando en las aplicaciones financieras dentro del contexto de las Personal Financial Managers (PFM), todavía está en una etapa prematura. Algunos expertos afirman que los potenciales para la intervención de las metodologías de ML y AI en las finanzas personales son enormes pudiendo transformar vidas y ayudar a las personas a tener una base financiera sólida. Según un estudio de la firma Kantar Worldpan acerca de las tendencias de ahorro en Colombia, menciona lo siguiente:

“Ocho de cada 10 colombianos consultados indicaron que entre sus metas está iniciar un plan de ahorro para cubrir emergencias y otras necesidades. (...) La idea es mantener un ‘colchón monetario’ para enfrentar emergencias y, además, ser más mesurados en sus gastos cotidianos.”[1]

Esto indica un deseo de los colombianos en ahorrar más y gastar menos generando, de esta manera, un nicho de mercado amplio para que aplicaciones PFM que utilicen las metodologías de aprendizaje de máquinas e inteligencia artificial para mejorar su experiencia de usuario, tengan una gran acogida y se conviertan en una idea de negocio exitosa al tiempo que ayuda a las personas a controlar y mejorar sus finanzas.

Dentro del contexto de Bancolombia, el banco más grande de Colombia y entre los más grandes de América Latina, existe una oportunidad de negocio clara con una necesidad marcada de los consumidores: el empoderamiento de las aplicaciones PFM dentro del mercado colombiano. Esta idea bien implementada, podría generar lazos duraderos y rentables entre el banco y sus beneficiarios que permitan crecer a la par de generar un impacto en el crecimiento de la economía colombiana.

Este proyecto presenta ciertas barreras de entrada, tales como la adopción de los teléfonos smartphone, la cobertura del acceso al internet móvil en celulares y hogares, la confianza de los consumidores a tener su información personal financiera en la nube y la veracidad de la información suministrada por el usuario, así como la facilidad y versatilidad del uso de la app como su costumbre. No obstante, estas barreras poco a poco se van estrechando al punto de prácticamente desaparecer en un futuro cercano.

En el marco de la Dataton Bancolombia 2018, se propone una serie de ideas innovadoras utilizando aplicaciones de aprendizaje de máquinas que generan un mayor valor agregado al usuario e incrementa la aceptación de las aplicaciones de administración de finanzas personales. Esta herramienta estará integrada al aplicativo del Banco, lo cual permitirá tener una vista integrada de los usuarios desde los distintos sistemas de información, con variables como sus ingresos, gastos, créditos, comportamientos de pagos, liquidez, etc.

Datos del problema

Información de usuarios

Estos datos corresponden a la información de los usuarios que el Banco posee, esta base de datos cuenta con 166.266 registros. Entre los campos disponibles para la información del usuario están el Id. único de cliente, su segmento estructural, ocupación, tipo de vivienda, nivel académico, estado civil, género, edad, rango de ingreso estimado. Muchas de estas variables vinieron previamente codificadas, con miras a tener más claridad a la hora de hacer análisis y de cara a la construcción final la cual será directamente visualizada por el usuario, se realizó un proceso inverso para tener las variables con su descripción explícita para tener una versión más descriptiva de los datos.

Información de transacciones PSE

Estos datos corresponden a la información, guardada en un archivo separado por comas con codificación UTF-8, de las transacciones PSE que el Banco posee, desde el 09/2016 hasta 10/2018 con un total de 11.853.782 registros. En primer lugar se retiraron espacios extras que contenía la base de datos, este procedimiento se realizó con código bash:

```
sed -z 's/\n\s\s/ /g' dt_trxpse_personas_2016_2018_muestra_adjt.csv  
> dt_trxpse_personas_2016_2018_muestra_adjt_mod_1.csv  
sed -z 's/\n,/,/g'  
dt_trxpse_personas_2016_2018_muestra_adjt_mod_1.csv >  
dt_trxpse_personas_2016_2018_muestra_adjt_mod_2.csv
```

```
sed -z 's/\n[^0-9]/ /g'  
dt_trxpse_personas_2016_2018_muestra_adj_t_mod_2.csv >  
dt_trxpse_personas_2016_2018_muestra_adj_t_mod_final.csv  
rm dt_trxpse_personas_2016_2018_muestra_adj_t_mod_1.csv  
dt_trxpse_personas_2016_2018_muestra_adj_t_mod_2.csv
```

Posteriormente, se encontró que en la columna ‘descripción’ de la transacción había casos de separación de comas, lo cual era interpretado como una o más columnas extras por parte del compilador incurriendo así a errores en tiempo de ejecución. Para evitar esto, se realizó un script adicional en R en donde se removía dicha coma adicional, este código primero partió la base de datos transaccional en 11 bloques o “chunks” de manera que se pudiera procesar más eficiente dadas las capacidades de cómputo y tener conjuntos de datos que se pudieran probar sin necesidad de cargar todos los datos a memoria, donde el script miraba línea a línea la presencia o no de una coma extra y se removía. Adicionalmente se configuraron los formatos de fechas para la presentación de gráficos representados a través del tiempo

Objetivo y alcance del proyecto

El objetivo de este proyecto es construir un aplicativo tipo PFM que permita, mediante la interacción del usuario con la plataforma, generar un valor agregado de sus finanzas y mayor y mejor relacionamiento con los servicios y productos que ofrece el Banco. Esta aplicación integrará diversas herramientas y técnicas analíticas pensadas y construidas para brindar una experiencia memorable al usuario, en términos generales:

1. Control, seguimiento de los gastos y segmentación del usuario
2. Determinación de capacidad de ahorro dada sus transacciones en PSE y estimación de ingresos.
3. Asignación de crédito ágil
4. Sistemas de recomendación mediante ChatBots

Implementación detallada de la solución

Como punto de partida se construyó una segmentación de los usuarios de la aplicación por medio de técnicas analíticas, de manera que se caracterizaron los grupos de usuarios en torno a sus patrones de consumo obtenidos de la base de datos de PSE.

Segmentación transaccional

Como complemento a la segmentación estructural del Banco, se construyó una segmentación transaccional con base a la información de PSE del tipo Recencia, Frecuencia, Monto (RFM) donde el histórico usado fue de los últimos tres meses de transacciones de la base de datos, la intención de esto es no penalizar a los clientes más nuevos (y que por lo tanto podrían tener un

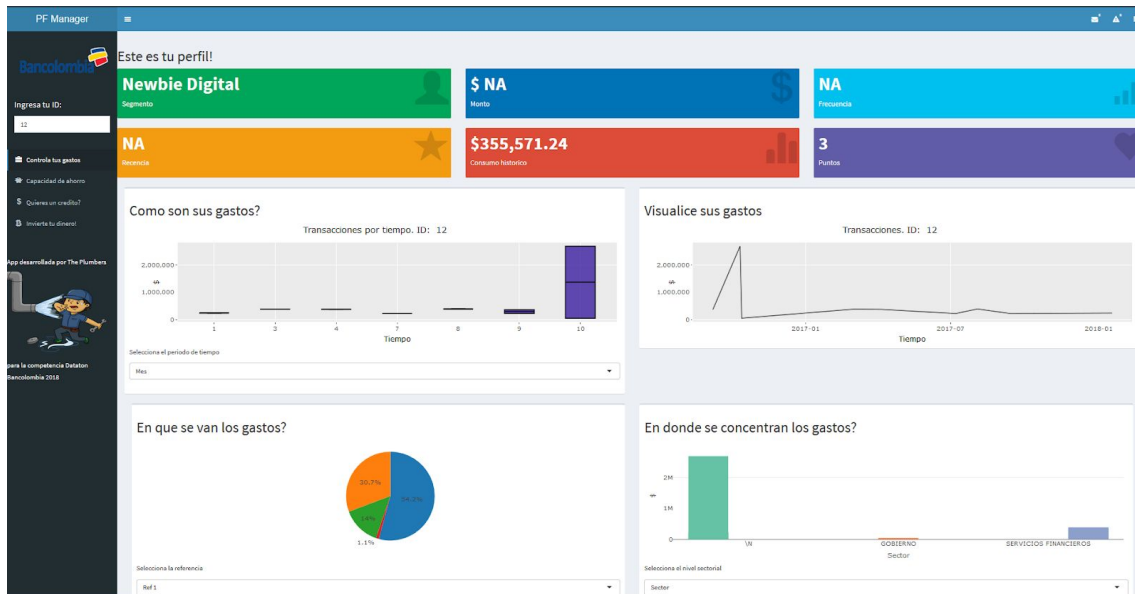
histórico menor con respecto a otros usuarios), además, de premiar los usos más recientes del sistema PSE. Para esto, se definió un sistema de puntaje por percentiles por variables de Recencia, Frecuencia y Monto, por ejemplo, para el caso de la frecuencia, se asignó hasta el percentil 40 un puntaje de 1, mientras que para los valores comprendidos entre el percentil 40 y 70 se asignó un puntaje de 2, para los valores de frecuencia entre 70 y 90 se les asignó un puntaje de 3 y para los valores entre el percentil 90 y 100 se les asignó un puntaje de 4, este procedimiento se realizó de manera análoga para las variables de recencia y monto. Adicional, para incluir también información histórica de todos los registros se incluyó la mediana de los valores de las transacciones, como un estimador robusto del valor medio de la transacción por usuario. Con esto, se definió un esquema piramidal de segmentación de tal forma que el primer 40% con puntaje más bajo definiera la base de la pirámide y aumentando de segmento con los restantes 30%, 20% y 10%. Los segmentos definidos, en orden de los patrones de consumos más bajos a los más altos, son 'Newbie Digital', 'Aprendiz Digital', 'Asociado Digital' y 'Experto Digital'. El uso de este tipo de segmentación tiene la ventaja de describir muy bien el comportamiento transaccional por cliente, además de ser fácilmente interpretable de cara al usuario final, permitiendo que se le puedan poner de manera sencilla objetivos determinados para aumentar de segmento.

App integrada de soluciones analíticas enfocadas al usuario

Como se mencionó anteriormente la aplicación se encuentra dividida en 4 partes principales las cuales se describen con detalle a continuación:

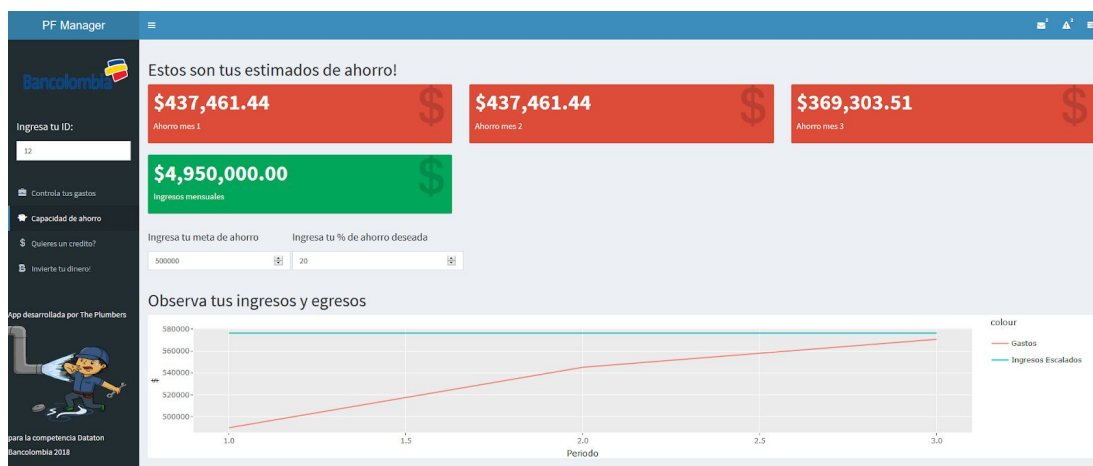
1. Controla tus gastos.

Una vez se ha construido la segmentación de los usuarios y que se cuenta con los patrones de consumo, se realiza un modelo de análisis exploratorio a partir del histórico de transacciones de PSE para cada usuario donde se presentan gráficamente los comportamientos de ingresos y pronósticos de gastos, como gráficos de bigotes con periodicidad personalizada, gráfica de líneas que permiten ver la tendencia en sus periodos de consumo, así como la posibilidad de visualizar en qué sectores se concentran mayormente sus gastos, todo esto de forma interactiva para el usuario. Así mismo, el usuario puede observar el segmento al que pertenece y las variables que definen su relación con la aplicación del banco.



2. Capacidad de ahorro.

Para la capacidad de ahorro de los usuarios, se calculó una predicción de gastos a partir del histórico de transacciones de PSE utilizando un modelo de SMA, luego se estimó sus ingresos con la información suministrada de la tabla de usuarios, posteriormente el usuario ingresa un porcentaje de ahorro deseado y una meta de ahorro, ambas en la aplicación de manera personalizada. Luego el aplicativo le calcula de manera automática cuales son sus estimaciones de ahorro para los siguientes tres periodos. Esta información se visualizará en el aplicativo PFM donde finalmente por medio de un semáforo le indicará cómo está el usuario frente a sus metas de ahorro y su comportamiento de gastos dadas sus parametrizaciones. Por otra parte, se presentan gráficamente los comportamientos de ingresos y pronósticos de gastos.



3. ¿Quieres un crédito?

Para la asignación de créditos ágil se construyó un modelo de otorgamiento de crédito basado en el perfil del usuario, donde se cuenta con la información de la capacidad de ahorro del apartado anterior y la segmentación estructural que le definió el banco en la información suministrada. De manera que dependiendo el perfil integral del usuario el modelo le asigna o no el crédito, una vez que el usuario sea beneficiado con el crédito, el modelo le informa el monto máximo a prestar, la tasa de intereses asociada y el número de periodos dependiendo de la calificación que se le otorgue por el modelo.

El sistema de asignación de crédito consistió en un modelo de clasificación en el cual, dada la información de los usuarios como su segmento estructural, ocupación, tipo de vivienda, nivel de estudio y estado civil, define las características más relevantes que permitan definir cuándo otorgar o no un crédito a un usuario determinado. Para esto, se construyó un árbol de decisión. Una vez definido, si el usuario se le puede otorgar o no un crédito se procede a definir el monto que se le puede asignar como préstamo. Para esto se define una capacidad de ahorro, que básicamente consiste en la estimación, mediante modelos de Smooth Moving Average de los consumos futuros del usuario que junto a sus ingresos y al segmento estructural al que pertenezcan definan un crédito para un interés dado.

Una vez el crédito se encuentre aprobado, el cliente puede acceder a este de dos formas. A través del link que redirecciona a la página de Bancolombia o descargando el certificado en PDF y mostrarlo en una sucursal, donde tras un papeleo breve, se desembolsa el crédito. Esto permite un acceso fácil y mejor para el usuario.



4. ¡Invierte tu dinero!

En este apartado se construyó un asistente virtual que a partir de la capacidad de ahorro y el segmento que tiene el cliente le recomienda diferentes opciones de inversión y donde el usuario puede interactuar mediante unas preguntas determinadas con el asistencial, de manera que el aplicativo según su perfil transaccional y sus ingresos estimados le dará una recomendación para invertir en fondos de inversión de Bancolombia.

Por otra parte, al desarrollar e integrar chats con Asistentes Interactivos al PFM, se facilita la ayuda a la hora de ofrecer soluciones financieras a la medida del usuario, lo cual, partiendo de la segmentación inicial, se pueden desarrollar modelos de recomendación buscando aumentar el cross-selling del usuario en el portafolio de inversión del banco como por ejemplo ahorros programados o inversiones en fondos cartera colectiva logrando una inclusión mayor en los mercados financieros a través de estos bots.



Conclusiones

Se construyó una solución analítica enfocada a las PFM en la cual se realiza un control y seguimiento de gastos, permitiendo además definir una segmentación del usuario. Adicional, se determinó la capacidad de ahorro del cliente mediante sus ingresos y transacciones en PSE. Además, para dicha solución se construyó un sistema de otorgamiento de créditos ágiles. Finalmente, esta solución también integra sistemas de recomendación mediante chatBots.

Los PFM generan una relación más fuerte entre el usuario y el Banco permitiendo generar un valor agregado a partir del historial de transacciones así como de otras fuentes adicionales, convirtiéndose esto en mayores ingresos para el Banco y una mayor satisfacción por parte del cliente debido a la personalización de servicios y productos.

El uso de herramientas analíticas en el contexto de las PFM permite mejorar la experiencia del usuario, lo cual podría dar una gran acogida por parte de los clientes hacia este tipo de soluciones, dando la oportunidad de convertirse en un negocio exitosa que a su vez ayuda a las personas a manejar de mejor manera sus finanzas.

Recomendaciones para trabajos futuros

Para trabajos futuros, de la mano de expertos del negocio y expertos en análisis de datos, se podrán construir distintos clases modelos buscando siempre una mirada estratégica del negocio. Posteriormente realizar pruebas de validación partiendo los usuarios entre grupos de entrenamiento y control permitiendo verificar la capacidad de predicción y generalización de los

modelos para los diferentes problemas de negocio. Una vez se obtenga un umbral de desempeño aceptable para el negocio, se procederá a la automatización y despliegue del modelo.

Como parte de una etapa posterior a la construcción del modelo se propone realizar mejoras al modelo de segmentación mediante la inclusión de variables demográficas, edad, estado civil, tipos de productos financieros, nivel de estudios, historial crediticio, etc., así como fuentes de información externas como las encontradas en redes sociales de manera que se pueda tener un perfil más integral de los usuarios.

Desde un punto de vista comercial, se pueden brindar incentivos a los usuarios con mayor número de transacción o mayores montos con el otorgamiento de puntos colombia que puedan ser canjeados con los distintos aliados del banco, logrando un mejor relacionamiento e incrementando la fidelización de los usuarios.

Gracias a la integración de la PMF a los sistemas de información transaccionales, así como a la generación de los llamados perfiles 360 del usuario y recomendaciones de los asistentes interactivos, se podrá construir una solución a la medida del cliente, permitiendo un mayor vínculo con el Banco derivado en un aumento del valor agregado para el usuario y una mayor rentabilización.

Finalmente, cuando las recomendaciones del asistente tenga una efectividad con el usuario, se pueden programar modelos analíticos de alertas para reportar seguimientos de cartera para predecir posibles defaults que permitirán anticiparse y generar estrategias de reestructuración de gastos del usuario para lograr cumplir con su obligación; de igual forma, el asistente puede informar alertas de rentabilidad de las inversiones del cliente lo cual actualizará su flujo de caja y tenga un panorama más realista de la liquidez con la que va a contar en el corto, mediano y largo plazo.

Referencias

[1] <https://www.eltiempo.com/economia/finanzas-personales/estudio-sobre-tendencias-de-ahorro-de-los-colombianos-en-2017-134210>