

COMPARACIÓN CON SIMILARES

DESCRIPCIÓN DEL PROYECTO:

Estudios y Comparación de Gastos entre Clientes Similares para conseguir ahorros

Ya sea en el extracto mensual del banco o en la página web de cliente al conectarse al banco el cliente puede ver con un sistema PFM (Personal Finance Management) tanto un reporte total de sus gastos mensualizados como un desglose de estos agrupados por categorías (Hogar, Deporte, Educación, Impuestos, Supermercado, Inversiones y Ahorro, Seguros, ...).

Adicionalmente podemos realizar una segmentación de la base de clientes del PFM a partir de los datos que ya tiene el banco o a través de los seguros que ya tiene contratado el cliente (hogar, automóvil, salud, ...) o a través de cuestionarios ad hoc que se podrían hacer a los clientes solicitándoles datos al conectarse a su página web de cliente de su banco o bien mediante un email con un enlace al cuestionario. Al cliente se le explicaría que sus datos serían agrupados y comparados con los clientes de su segmento (es decir, con los clientes que sean similares a él) para poder ofrecerles estudios comparativos de gasto entre él y sus similares para que puede comprobar, en cada categoría de gasto, si está gastando igual, más o menos que los clientes similares a él de su segmento.

Los datos a solicitar al cliente o con los que trabajar para segmentarlo, entre otros serían: sexo, edad, nivel de estudios (sin estudios, bachillerato/FP, grado, máster/licenciatura, posgrado, doctorado), profesión (a elegir de entre una lista de opciones), sector laboral (a elegir de entre una lista de opciones), nivel laboral (administrativo, técnico, jefe de equipo, mando intermedio, directivo), si trabaja por cuenta propia (autónomo) o ajena, estado civil, número de hijos y edades, número de convivientes en la unidad familiar, ingresos anuales, código postal del domicilio, metros cuadrados y número de habitaciones del domicilio, número de coches de la unidad familiar, modelos y matrículas (número de años de cada vehículo), seguros contratados (hogar, coche, salud, vida, ...) y compañías con los que se contrataron (también modalidad de seguro contratado cuando sean comparables; en el caso de coche si es a terceros, con robo y lunas, con franquicia, a todo riesgo, etc.); cuantía, compañías y modalidad de servicios contratados de los principales recibos periódicos (internet/móvil/televisión, electricidad/gas, agua, calefacción), otros recibos o gastos periódicos (impuestos (IBI, IVTM), comunidad, colegios profesionales, agrupaciones deportivas o de otra clase, asociaciones de antiguos alumnos, etc.), otros servicios contratados por suscripción (Netflix, Spotify, Disney+, Viber, Microsoft Office; LinkedIn, Evernote, LastPass, ...), cuota de hipoteca/alquiler.

Con estos datos correctamente anonimizados y agregados a cada cliente se le puede ofrecer una información muy valiosa para la gestión de sus finanzas. Y no sólo mediante estudios comparativos con similares; sino que mediante creación de modelos con algoritmos de Machine Learning de aprendizaje supervisado se podría predecir cuanto debería estar pagando el cliente en cada categoría de gasto incluso en cada recibo y, además de presentarle dichos datos en su extracto mensual o área de cliente, también se le podrían enviar alertas o resaltarle en su extracto mensual o área de cliente, por ejemplo, aquellas oportunidades de ahorro superiores a un umbral (10% por ejemplo) en una categoría o incluso, en un recibo concreto, lo cual sería mucho más útil para el usuario.

Por ejemplo, al cliente se le podrían emitir alertas como las siguientes:

- Si está pagando por el seguro de su coche de más de 8 años más de un 30% en su seguro de automóvil de lo que debería estar pagando de acuerdo con sus características personales (su segmentación de clientes) o comparado con clientes similares a él y con un vehículo similar al suyo e indicarle, por ejemplo, que la mayoría de los clientes de su segmento tienen contratado un seguro a terceros en lugar de un seguro a todo riesgo.
- Por ejemplo, si está pagando de luz en media más de un 10% de lo que predice el modelo o de la media que están pagando los clientes de su segmento se le podría enviar una alerta o bien resaltárselo en su extracto mensual o en la web o aplicación móvil en su área de cliente. Si el modelo de Machine Learning indica que una de las características que más influye en la cuantía de la factura es la tarifa contratada (sin discriminación horaria, con discriminación horaria de 2 o 3 tramos, tarifa de último recurso, etc.) y se detecta que el cliente que está pagando en media más de un 10% en la factura mensual tiene una tarifa sin discriminación horaria y

la mayoría de los clientes de su segmento tiene contratada una tarifa con discriminación horaria de 2 tramos se le podría indicar dicha situación o incluso sugerirle el cambio a dicha tarifa.

Una de las ventajas de predecir el gasto total o el gasto en cada categoría o en cada recibo mediante modelos creados con algoritmos de Machine Learning es que muchos de los algoritmos que los crean permiten además conocer de forma cuantificada cuales son las variables de entrada (input o dependientes) que más influyen en el valor final de la variable target (objetivo o independiente).

Esto no solamente permite obtener modelos que hagan predicciones más precisas ya que sólo tomaran en cuenta en el cálculo de la predicción aquellas variables de entrada que realmente influyan en el valor final de la variable target sino que podrá dar información más precisa al PFM de las variables que más influyan en el valor de la predicción. Esto se conoce como 'Ingeniería de Características' o 'Feature Engineering'.

Y en el caso de que un usuario que esté pagando de más en un recibo o en una categoría y se detecte que tiene un valor distinto (en el caso de una variable categórica; por ejemplo tarifa sin/con discriminación horaria, seguro de coche a terceros o a todo riesgo) o muy distinto (en el caso de una variable continua, por ejemplo potencia contratada o número de partes enviados por año) se le podrá informar el cliente del PFM no sólo que está pagando de más en una categoría o recibo sino en cual de las variables que influyen en el valor final de la predicción el usuario tiene un valor distinto o muy distinto al resto de usuarios que están pagando menos que él. Por ejemplo, se le podría informar al usuario que la mayoría de los clientes como él tienen una tarifa de electricidad con discriminación horaria de 2 tramos y el tiene una tarifa sin discriminación horaria; o que la mayoría de los usuarios que tienen un coche como el suyo tienen un seguro del automóvil a terceros mientras que él lo tiene a todo riesgo, etc.

Este tipo de predicciones realizadas sobre los datos de los clientes mediante algoritmos de Machine Learning tienen multitud de aplicaciones y no sólo para optimizar gastos. Algunos otros ejemplos de aplicación son los siguientes:

- Se podría predecir si un usuario debiera tener o no contratado un seguro del hogar y en el caso de que no lo tenga y el modelo prediga que debería tenerlo, informarle de dicha circunstancia. Igualmente se podría hacer lo mismo con un seguro de salud.
- Se podría predecir la cantidad media que un usuario debería tener en cuentas corrientes y de ahorro de bajo interés y en el caso de que difiera de la cantidad media real difiera en más de un porcentaje determinado (por ejemplo, un 40%) y una de las variables categóricas influyentes en el valor de la predicción sea la contratación de productos financieros como fondos de inversión, y el usuario no los tenga contratados, entonces indicarle dicha circunstancia. Igualmente, si la cantidad que el usuario tiene en cuentas corrientes y de ahorro es inferior de un porcentaje umbral predeterminado frente a lo que prediga el modelo de Machine Learning, también se podría informar al usuario de dicha circunstancia.
- Se podrían hacer modelos Machine Learning de scoring personalizados para cada usuario. Es decir, predecir qué cantidad máxima de dinero se le podría conceder a un usuario en un préstamo hipotecario o en un crédito al consumo dadas unas condiciones (plazo de devolución) en función de sus características personales y de otras hipotecas y créditos concedidos a otros usuarios similares a él. Los resultados de esta predicción se podrían mostrar en la página web de cliente o en su aplicación móvil a petición del cliente o bien se le podría mostrar al cliente en su página de inicio. Esto incrementaría el cross-selling de productos del banco y también la experiencia de cliente UX del usuario al conocer, de forma casi inmediata, la cantidad pre concedida que tiene tanto para una hipoteca como para un prestamos para el consumo personal.
- En cuanto a la estimación de gastos en suscripciones se podrían catalogar como variables categóricas True/False si un usuario tiene una suscripción o no con los servicios más comunes como Netflix, Spotify, Amazon Prime, Disney+, Office365, LinkedIn, Apple iCloud, Google Cloud, Viber y, al predecir el gasto que debería tener un usuario en suscripciones, el modelo debería presentar, de aquellas variables características que más influencia tienen en el valor de la variable final target, aquellas que tengan un valor distinto al de la mayoría de usuarios similares a él que estén pagando de forma apreciable menos dinero en suscripciones. Por ejemplo, si detectamos que la mayoría de los usuarios como él pagan por Netflix, Spotify, pero no están suscritos a Disney+, Amazon Prime y a los Cloud de Google y Apple se podría informar al cliente de dicha circunstancia.

BENEFICIOS ESPERADOS

Aumento de la satisfacción del cliente y de su lealtad hacia el banco pues le estaremos ofreciendo de forma gratuita sugerencias e información para gestionar mejor sus finanzas que no le estarán ofreciendo en otros bancos.

Esta información al usuario para la mejor gestión de sus finanzas será una ventaja competitiva para aquellos bancos que la tengan frente a los que no.

Además, al tener usuarios que sepan gestionar mejor sus finanzas reduciremos el riesgo de tener clientes morosos y de impagos.

IMPACTO ESPERADO

Mejora de la reputación del banco al mejorar la opinión que los clientes tendrán del mismo por los motivos descritos en el apartado anterior. Esa mejora reputacional se puede traducir, con el tiempo, en un incremento de clientes para el banco además de una reducción de la tasa de churning o clientes que abandonan el banco.

ESTIMACION CREACION NETA DE EMPLEO

Para la creación de un producto básico se necesitaría un Ingeniero de Datos y un Científico de Datos (Ingeniero de Machine Learning) durante 4 meses y un Desarrollador Backend y otro Frontend durante 3 meses. Una vez puesto en producción el producto básico, éste necesitaría ser mantenido y evolucionado de forma indefinida por un equipo formado por un Ingeniero de Datos, un Científico de Datos (Ingeniero de Machine Learning), un Desarrollador Frontend y un Desarrollador Backend.

SECTOR

FinTech: Financiero y de Tecnologías de la Información. Con especial foco en las disciplinas de Big Data (Ingeniería de Datos), Aprendizaje Automático (Machine Learning) y Desarrollo Software de Backend y Frontend.

OBJETIVO DE TRANSFORMACION DIGITAL

Crear un servicio adicional de información al usuario sobre el estado y la gestión de sus finanzas. Contribuyendo, además, a su educación financiera, reduciendo las posibilidades de impagos, morosidad e incrementando los beneficios económicos del usuario al presentarle información para realizar una gestión más inteligente de su economía personal.

PRESUPUESTO ESTIMADO

Apartado ya contestado en horas-hombre en la sección 'Estimación Creación Neta de Empleo'.

PORCENTAJE DE FINANCIACION PUBLICA SOLICITADO

A concretar tras la aceptación de una oferta concreta por parte de un cliente.