

# Calculation of CLV in Banking from a Global Perspective

**Abstract.**In this article a new CLV model is proposed, one that considers the scenario of new relations between bank-client brought about by the entry of new competitors such as *fintech* products and therefore takes into account either monetary and non-monetary aspects. For this purpose it is used the RFM model, which has been proved especially useful for integrating digital and physical interactions of clients in other areas, such as retail. It is also intended that the final model contains the criteria of the main areas of banks that more relation have with clients. In order to solve this, we propose the use of the so-called Multicriteria Decision-Making, in such a manner that the permit to obtain not only the preference or collective criteria of all these members in conjunction, but also it assures that such value has enough agreement between the areas.

**Keywords.**CLV; RFM; Multicriteria Decision-Making; Fintech.

## Calculo de CLV en Banca Bajo una Perspectiva Global

**Resumen.**En este artículo se propone un modelo de CLV que contempla el nuevo escenario de relaciones marco banco-cliente propiciado por la entrada de nuevos competidores, como los de los productos *fintech*, teniendo en cuenta tanto aspectos monetarios como no monetarios a la hora de valorar el cliente. Para ello se hace uso del modelo RFM que se ha mostrado especialmente útil para integrar tanto interacciones digitales como físicas de los clientes en otros ámbitos como el de retail. Se pretende además que el modelo final refleje los criterios de las grandes áreas de los bancos que más relación tienen con los clientes, para lo cual, se propone el uso de los llamados Modelos de Decisión Multi-Criterio que no solo permiten obtener la definición del modelo con el criterio colectivo de todos estos entes conjuntamente sino que además aseguran que el modelo tenga el suficiente grado de consenso entre las áreas.

**Palabras clave.** CLV; RFM; Modelos de Decisión Multi-Criterio; Fintech.

# 1 Introducción

El Valor del Tiempo de Vida del Cliente o CLV (siglas en inglés de *CustomerLifetimeValue*) es una orientación de negocio que evalúa a los clientes como activos financieros, para los cuales la empresa debería medir, gestionar y maximizar dicha puntuación, al igual que otros activos estratégicos de la organización (Valenzuela Fernández et al., 2007). Actualmente se considera una métrica fundamental para la asignación de los recursos del marketing (Venkatesan&Kumar, 2004; Zeithaml et al., 2001). Ciertos autores relacionan el cálculo del CLV con aspectos como la frecuencia y cantidad gastada en las compras (Rust et al., 2004).

En el sector financiero, la relación de los clientes con sus correspondientes bancos está cambiando a un ritmo vertiginoso debido, principalmente, a la transformación digital. Dicha transformación a su vez ha permitido la entrada de seria competencia a la banca comercial tradicional, especialmente, de las compañías TIC que ofrecen nuevos y mejorados servicios con un alto componente de tecnología, conocidos como *fintech* (financiatechnology). El objetivo de este artículo es definir un modelo de CLV que contemple este nuevo escenario de relaciones marco banco-cliente y por tanto tenga en cuenta tanto aspectos monetarios como no monetarios. Con este fin se hace uso del modelo RFM (Recencia, Frecuencia y valor Monetario) que se ha mostrado especialmente útil para integrar tanto interacciones digitales como físicas de los clientes en otros ámbitos como el de retail. En el modelo propuesto: la Recencia se refiere al momento en el que el cliente efectuó su última interacción proactiva con la entidad bancaria independientemente del canal usado; la Frecuencia es el número de veces que un cliente adquiere un producto o hace uso de algún servicio en el período; y el valor Monetario está relacionado con aspectos más clásicos de los modelos CLV en banca: el volumen de negocio (en productos de inversión, financiación...) y, por supuesto, la rentabilidad que genera un cliente para la entidad. Tras una revisión de la bibliografía y consulta a expertos del sector se eligen una serie de subvariables para construir cada una de las dimensiones RFM indicadas.

En un modelo RFM clásico existe la problemática de elegir el peso o importancia de cada una de estas tres dimensiones. En el modelo propuesto, además también hay que elegir la importancia de cada una de las subvariables con las que se calculan cada una de las dimensiones del modelo. Se pretende además que el modelo final refleje los criterios de las grandes áreas de los bancos que más relación tienen con los clientes, a saber: marketing y/o experiencia del cliente; fuerza de ventas; soporte del cliente; y riesgos. Con el fin de resolver este problema, se propone el uso de los llamados modelos de Toma de Decisiones Multi-Criterio (TDMC) (Cid et al., 2016; Carrasco et al., 2015b; Carrasco et al., 2014; Carrasco et al., 2011) de tal manera que permitan obtener no solo la preferencia o criterio de todos estos entes conjuntamente, sino que también aseguren que dicho valor tenga el suficiente grado de consenso entre ellos.

Para desarrollar el modelo propuesto se ha realizando un cuestionario para recoger el criterio de diversos profesionales del sector. Esta información ha sido procesada mediante diversos procesos de TDMC. Una vez obtenido el modelo RFM final se han especificado las distintas fases necesarias para aplicado a los clientes de un banco a partir de la información contenida en su Data Warehouse.

El resto del artículo está estructurado como sigue: en la Sección 2 se introducen los conceptos preliminares, tanto de RFM como de TDMC; en la Sección 3 se propone el modelo RFM para el cálculo del CLV como un modelo secuencial de tres fases; y finalmente se hace una serie de comentarios a modo de conclusiones y se apunta el trabajo futuro.

## 2 Preliminares

En esta sección se va a hacer un repaso de conceptos necesarios para entender la nueva propuesta: el modelo RFM y el modelo TDMC.

### 2.1 El Modelo RFM

El modelo RFM es una técnica usada habitualmente en marketing que identifica patrones comunes de comportamiento de compras permitiendo tipificar a los clientes en base a tres variables (Tsipsis&Chorianopoulos, 2011):

- Recencia (R): Tiempo transcurrido desde la última compra o visita.
- Frecuencia (F): Número total de compras o visitas en el periodo analizado.
- Valor Monetario (M): El montante total de las compras en el periodo analizado.

Las fases habituales para su cálculo podrían ser las siguientes (Carrasco et al., 2015a; Tsipsis&Chorianopoulos, 2011):

1. **Adquisición y preparación de datos.** Una vez elegido el periodo a analizar,  $[Fecha_i, Fecha_f]$ , se seleccionan los clientes que han tenido al menos una compra en dicho periodo. En esta etapa, se obtienen los datos transaccionales (compras) de dichos clientes, se audita su calidad y se efectúa un proceso de limpieza de los mismos preparándolos para las siguientes etapas. Notaremos con *Transacciones* (*ClienteID*, *Fecha*, *Importe*) la tabla con dichos datos transaccionales de los clientes, tal que  $Fecha \in [Fecha_i, Fecha_f]$ .
2. **Agregación de los valores RFM.** A partir de los datos de la tabla *Transacciones*, se obtiene la tabla *ClienteTransacciones* (*ClienteID*, *Recencia*, *Frecuencia*, *Monetario*) aplicando las correspondientes funciones de agregación a nivel de cliente (esto es, sobre el atributo *ClienteID*). Sea  $v_i, i \in \{1, 2, 3\}$  una columna de la tabla *ClienteTransacciones*, tal que  $v_1 = Recencia, v_2 = Frecuencia, v_3 = Monetario$ . Su cálculo sería el siguiente:

$$v_i = \begin{cases} last(Fecha) - Fecha_r, \text{ con } Fecha_r > Fecha_i & \text{si } i = 1, \\ count(ClienteID) & \text{si } i = 2, \\ sum(Importe) & \text{si } i = 3. \end{cases}$$

Esto es, para cada cliente se calcula: el atributo *Recencia*, usando la función *last* que obtiene la fecha de la última transacción realizada por el mismo, a dicho valor se le resta una fecha fija de referencia  $Fecha_r$ ; el atributo *Frecuencia*, contando las distintas transacciones realizadas mediante la función *count*; y el atributo *Monetario*, sumando el importe de todas las compras realizadas mediante la función *sum*.

3. **Cálculo de los scores RFM.** Para cada una de las variables  $v_i$ , se ordenan los clientes respecto al valor de dicha variable y se agrupan en  $n$  clases aproximadamente de igual tamaño. Estos resultados se incluyen en la tabla *ClienteRFM* (*ClienteID*, *RecenciaScore*, *FrecuenciaScore*, *MonetarioScore*, *RFMScore*). Sea  $x_i, i \in \{1, 2, 3\}$  una columna de la tabla *ClienteRFM*, tal que  $x_1 = Recencia, x_2 = Frecuencia, x_3 = Monetario$ , por tanto se tiene que  $x_i \in \{1, \dots, n\}$  y se calcula con:

$$x_i = ntile(n, v_i),$$

donde la función *ntile* obtiene un valor ordinal entre 1 y  $n$ , de tal manera que el valor 1 contiene el  $100/n$  porcentaje de los clientes con los peores valores de la correspondiente variable  $v_i$  y el valor  $n$  el  $100/n$  porcentaje de los mejores clientes según dicha variable. Frecuentemente se usan quintiles, esto es,  $n = 5$ .

4. **Cálculo del score RFM global.** Es de especial utilidad para el cálculo del CLV obtener un único valor, que simbolizamos como *RFMScore*, que caracterice globalmente los tres valores RFM de cada cliente. Para su cálculo, normalmente, se proporcionan los pesos de cada variable,  $w_i, i \in \{1, 2, 3\}$  y se aplica una media ponderada con los correspondientes scores:

$$RFMscore = \frac{\sum_{i=1}^3 w_i \times x_i}{\sum_{i=1}^3 w_i} \quad (1)$$

## 2.2 El Modelo de Toma de Decisiones Multi-Criterio

La toma de decisiones es un proceso complejo que se enfrenta el ser humano cuando se encuentra situaciones en las que se tienen varias alternativas y es necesario escoger una (o varias) siendo necesario, en muchas ocasiones, valorar una serie de criterios. En un modelo de TDMC (Cid et al., 2016; Carrasco et al., 2015b; Carrasco et al., 2014; Carrasco et al., 2011), el objetivo consiste en buscar la/s mejor/es alternativa/s de entre las incluidas en un conjunto,  $alternativas = \{alternativa_1, \dots, alternativa_m\}$ , según las valoraciones proporcionadas por un grupo de expertos, en base a unos criterios preestablecidos y recogidos en el conjunto  $criterios = \{criterio_1, \dots, criterio_n\}$ . Frecuentemente también se asume que la importancia de las valoraciones está determinada por una serie de pesos,  $pesos = \{peso_1, \dots, peso_n\}$ . Las valoraciones se almacenarían en el conjunto  $valoraciones = \{valoración_{ij} \mid \forall i \in \{1, \dots, m\}, \forall j \in \{1, \dots, n\}\}$ .

Por tanto, el modelo TDMC que se va a usar en este trabajo podría formalizarse mediante la función  $\Delta$  siguiente:

$$\begin{aligned} \Delta: & \begin{array}{c} \text{criterio}_1 \dots \text{criterio}_n \\ \text{peso}_1 \dots \text{peso}_n \\ \text{alternativa}_1 \left[ \begin{array}{ccc} \text{valoración}_{11} & \dots & \text{valoración}_{n1} \\ \vdots & \ddots & \vdots \\ \text{valoración}_{1m} & \dots & \text{valoración}_{nm} \end{array} \right] \\ \vdots & \\ \text{alternativa}_m & \end{array} \\ \rightarrow & \begin{array}{c} \text{alternativa}_1 \left[ \begin{array}{cc} \text{preferencia\_colectiva}_1 & , \text{valor\_cosenso}_1 \\ \vdots & \vdots \end{array} \right] \\ \vdots & \\ \text{alternativa}_m \left[ \begin{array}{cc} \text{preferencia\_colectiva}_m & , \text{valor\_cosenso}_m \end{array} \right] \end{array} \end{aligned} \quad (2)$$

Las fases que va a tener este proceso de decisión van a ser las clásicas de este tipo de problemas:

- **Fase de Agregación:** Se combinan las preferencias de los expertos, esto es, los elementos del conjunto  $valoraciones$ , según los pesos expresados en el conjunto  $pesos$ , mediante el operador de agregación correspondiente, obteniendo el conjunto  $preferencias\_colectivas = \{preferencia\_colectiva_1, \dots, preferencia\_colectiva_m\}$  con el valor de preferencia colectivo para cada una de las alternativas incluidas en el conjunto  $alternativas$ . Se hace necesario, en este tipo de problemas, que dicha agregación se haga con el suficiente grado de consenso entre expertos (criterios emitidos por ellos), expresándose dichos valores en el conjunto  $valores\_consenso = \{valor\_consenso_1, \dots, valor\_consenso_m\}$ . Dichos grados de consenso son la coincidencia entre relaciones de preferencia expresadas por diferentes actores (Herrera-Viedma et al., 2007), es decir, serían medidas para calcular el grado de acuerdo grupal a partir de preferencias individuales de los expertos. Las medidas de consenso suelen estar basadas en el empleo de métricas de similitud o distancia para calcular la cercanía entre las preferencias de los expertos. Las medidas de consenso y de preferencias colectivas a la que nos referiremos en este trabajo está basada en los planteamientos de Cid et al. (2016), definiéndose como sigue:

$$\begin{aligned}
\text{valor\_consenso}_i &= 1 - \frac{\sum_{j=1}^n | \text{preferencia\_colectiva}_i - \text{valoración}_{ij} | \times \text{peso}_j}{\sum_{j=1}^n \text{peso}_j} \\
\text{preferencia\_colectiva}_i &= 1 - \frac{\sum_{j=1}^n \text{valoración}_{ij} \times \text{peso}_j}{\sum_{j=1}^n \text{peso}_j} \\
\forall i \in \{1, \dots, m\}.
\end{aligned}$$

- **Fase de Explotación:** Si el problema exige seleccionar solo las mejores alternativas, se seleccionan aquellas del conjunto *alternativas* con mayor valor de preferencia colectiva, esto es, con mayores valores en el conjunto *preferencias\_colectivas*. En otras ocasiones se exige que estén altamente consensuadas por los expertos, esto es, se seleccionarían solo aquellas cuyos valores correspondientes en el conjunto *valores\_consenso* igualan o superan un umbral mínimo de consenso  $\gamma$ .

### 3 Modelo RFM para cálculo CLV en banca

El uso del modelo RFM en el sector financiero ha sido ya propuesto en varias ocasiones. Como se ha visto en el Epígrafe 2.1 este modelo parte de los datos transaccionales de los clientes, así diversos autores (Wu et al., 2005; Hsieh, 2004; Sohrabi&Khanlari, 2007) lo proponen para tipificar a los clientes según sus compras con tarjetas de crédito. Mohammadi et al. (2014) lo usan instrumentalmente para establecer rankings de crédito. Por otro lado, Aggelis&Christodoulakis (2005) proponen un modelo RFM que parte de las transacciones de banca electrónica que realizan los clientes.

Sin embargo, la banca comercial suele tener sus estrategias comerciales centradas en el cliente y soportadas por CRM (*Customer Relationship Management*) que contienen mucha información no solo transaccionales realizadas por los clientes, sino también otro tipo de interacciones con los mismos tanto personales, como electrónicas. Además el catálogo de productos de los bancos suele ser bastantes extensos y complejos. De esta forma, Nikumanesh&Albadvi (2014) aunque siguen restringiendo el cálculo de la frecuencia a las operaciones hechas con tarjeta, amplían los indicadores usados para cálculo del valor monetario a los depósitos, préstamos y número medio de cuentas que tiene el cliente.

En este epígrafe se va a proponer un modelo global para el cálculo del CLV que abarque todas las posibles relaciones tanto contractuales como no contractuales de los clientes con la entidad, que además esté ampliamente aceptado por todas las áreas de la entidad que más relación tienen con los clientes y que sea fácilmente trazable. Para ello se van a seguir tres pasos: 1) definición inicial del modelo RFM según los indicadores de clientes; 2) cálculo de la importancia consensuada de los pesos del modelo RFM; y 3) cálculo del modelo RFM para cada cliente. A continuación se explican con más detalle:

#### 3.1 Definición inicial del modelo RFM según los indicadores de cliente

El primer paso, es definir el concepto de RFM aplicado a entidades financieras pero cubriendo todas las posibles relaciones e interacciones entre el cliente y la entidad:

- **R:** Momento en el que el cliente efectuó su última interacción proactiva con la entidad bancaria.

- F: Número de veces que un cliente adquiere un producto o usa un servicio en un período determinado.
- M: Valor monetario en dicho periodo que el cliente mantiene contractualmente con la entidad.

El siguiente paso que se va a seguir es definir los indicadores necesarios para el modelo. Siguiendo la premisa marcada, esto es, la de simplicidad y fácil trazabilidad del mismo, en la Tabla 1 se proponen dichos indicadores base tras consulta a diversos expertos del sector.

Tabla 1. Indicadores a usar en el modelo RFM

Indicador	Descripción
<i>Medios Pago</i>	Operaciones correspondientes a pagos efectuados por parte del cliente: tarjetas (en sus distintas modalidades), cheques, pago por móvil, recibos domiciliados, etc.
<i>Medios Cobro</i>	Operaciones correspondientes a ingresos efectuados al cliente: nóminas, prestaciones, pensiones u otros ingresos como pagos de alquileres, etc.
<i>Interacciones Físicas</i>	Entrevistas, reuniones... entre el gestor comercial (u otro personal del banco) y el cliente ya sea en la oficina bancaria u otro lugar pero siempre con presencia física de ambas partes.
<i>Resto Interacciones</i>	Llamadas, emails, video-conferencias, SMS, mensajes de whatsapp o cualquier otra comunicación no física entre cliente y entidad.
<i>Financiación Hipotecaria</i>	Producto que permite al cliente, recibir una determinada cantidad de dinero (capital del préstamo) de la entidad de financiera (prestamista), a cambio del compromiso de devolver dicha cantidad, junto con los intereses correspondientes, mediante pagos periódicos (cuotas). La entidad cuenta con una garantía especial para el recobro de la cantidad prestada: una hipoteca sobre un inmueble (una vivienda, por lo general) que suele ser propiedad del cliente.
<i>Resto Financiaciones</i>	Otro tipo de préstamos no incluidos en el apartado anterior, como por ejemplo préstamos personales o créditos de tarjeta. El préstamo personal responde a las mismas características que la financiación hipotecaria, pero en este caso, se llaman personales porque en este tipo de préstamos la entidad no suele contar con una garantía especial para el recobro de la cantidad prestada.
<i>Inversión Vista</i>	Productos como cuentas corrientes o depósitos a la vista. En la cuenta corriente el cliente puede ingresar dinero en la entidad, quedando ésta obligada a su devolución en cualquier momento en que lo solicite. Por lo general, la entidad ofrece al cliente una remuneración o tipo de interés por el dinero depositado. Otro rasgo fundamental de la cuenta corriente es que la entidad presta un “servicio de caja” muy activo, de manera que se compromete a realizar los pagos y cobros que el cliente le encargue: ingresos de efectivo, abonos en general (nóminas, etc), ordenar y recibir transferencias bancarias, domiciliar recibos, retiradas de dinero en efectivo o pagar cheques.
<i>Inversión Ahorro</i>	Comprende los depósitos a plazos fijos o fondos de inversión. Los depósitos son productos en los que el cliente entrega una cantidad de dinero a la entidad durante un tiempo determinado. Transcurrido ese plazo, la entidad se lo devuelve, junto a la remuneración pactada (intereses), salvo que se acuerde con la entidad el cobro periódico de intereses mientras dure la operación. La diferencia fundamental entre este tipo de depósitos y los depósitos “a la vista” consiste en que en éstos el cliente puede disponer del dinero de forma inmediata sin ningún tipo de penalización. A cambio, la remuneración de los depósitos a plazo suele ser superior. Hay otros tipos de depósitos en los que la rentabilidad está vinculada a la evolución de un índice, al valor de una cesta de acciones o incluso a que tenga lugar o no un acontecimiento futuro, depósitos que no permiten la cancelación anticipada o si la permiten es en unas condiciones muy distintas a las de los depósitos tradicionales.
<i>Productos Parabancarios</i>	Productos que no computan en el balance de la entidad, la cual actúa como comisionista, por ejemplo: valores, fondos de pensiones, seguros...
<i>Rentabilidad</i>	Margen de explotación u ordinario generada por el cliente. Es el resultado de la diferencia entre los ingresos derivados de la prestación de servicios por la entidad a sus clientes, es decir, los ingresos correspondientes a las comisiones cobradas por los distintos productos que cada entidad suministra a sus clientes, y los costes de explotación derivados de la prestación de esos mismos servicios.

En base a estos indicadores (Tabla 1), a continuación se define cada una de las variables del modelo RFM,  $v_i$ ,  $i \in \{1, 2, 3\}$ , con  $v_1 = Recencia$ ,  $v_2 = Frecuencia$ ,  $v_3 = Monetario$ , en base a una serie de subvariables,  $v^i_j$ , con  $j \in \{1, \dots, \#v_i\}$ , siendo  $\#v_i$  el número de subvariables de las que están compuesto la variable  $v_i$ . Dicha definición se muestra en la Tabla 2.

Tabla 2. Indicadores que componen el modelo RFM

Variable RFM ( $v_i$ )	Indicador	Subvariable RFM ( $v^i_j$ )
R	Medios de pago	$v^1_1$
	Medios de cobro	$v^1_2$
	Interacciones directas	$v^1_3$
	Resto de interacciones	$v^1_4$
F	Medios de pago	$v^2_1$
	Medios de cobro	$v^2_2$
	Interacciones directas	$v^2_3$
	Resto de interacciones	$v^2_4$

M	Financiación hipotecaria	$v^3_1$
	Otros tipos financiación	$v^3_2$
	Productos parabancarios	$v^3_3$
	Inversión productos a la vista	$v^3_4$
	Inversión productos ahorro	$v^3_5$
	Rentabilidad	$v^3_6$

### 3.2 Cálculo de la importancia consensuada de los pesos del modelo RFM

Como se ha visto en la etapa anterior, cada una de las variables del modelo RFM  $v_i$  está compuesta a su vez de otras subvariables  $v^j_i$ . Ahora en esta fase, se trata de establecer la importancia relativa de cada una de estas subvariables, que se notará como  $w^j_i$ , según la opinión colectiva de los expertos bancarios. Además también se quiere obtener, según la misma opinión, los pesos  $w_i$  de cada una de las variables  $v_i$  del RFM de Eq. (1). Para ambos propósitos se va a usar el modelo TDMC explicado en la Sección 2.2. Es importante destacar que se exigirá un grado de consenso mínimo en dichos TDMC, ya que sin él no se podría dar validez al modelo, especialmente para las subvariables que lo componen. Los tres pasos que se siguen en esta fase son los siguientes:

#### 1. Obtención de los datos de los expertos

La solución planteada para este punto ha consistido en la creación del cuestionario, notado como  $Q$ , que está incluido en el Anexo I y explicado con más detalle en (Hernández, 2014). Este cuestionario consta de tres partes: la primera referente a la importancia que tienen cada una las tres variables RFM de cara a valorar al cliente; la segunda sobre las distintas subvariables del modelo (Tabla 2); y la tercera sobre los años de experiencia del encuestado. Para las dos primeras partes se ha usado una escala 7-Likert. En este artículo, para facilitar el entendimiento de la variable a la que corresponde cada cuestión se ha incluido en dicho Anexo la columna  $V$ .

Con la encuesta preparada, el método utilizado para difundirla entre la población objetivo (profesionales del sector bancario, seleccionados en función de sus competencias relacionadas con el cliente, pertenecientes a: marketing y/o experiencia del cliente; fuerza de ventas; soporte del cliente; y riesgos), ha sido el correo electrónico. La población que ha respondido a esta encuesta asciende a un total de 29 expertos, que pertenecen a distintas entidades bancarias situadas en España y con los años de experiencia que se muestran en la Tabla 3. En lo que sigue, se nota las respuestas efectuadas a dicho cuestionario como  $Q_k(v)$ , siendo  $v$  una variable de la columna  $V$  y  $k$  el correspondiente encuestado, con  $k = \{1, \dots, s\}$ ,  $s = 22$ .

Tabla 3. Detalle de los encuestados por años de experiencia

Experiencia	Encuestados
Menos de un año	1
Entre uno y tres años	2
Entre tres y cinco años	5
Entre cinco y diez años	7
Más de diez años	14

#### 2. Modelos TDMC para obtener los pesos de cada subvariable del RFM

Se realiza tres procesos TDMC distintos para cada una de las variables  $v_i$ ,  $i \in \{1, 2, 3\}$ .

En cada proceso  $i$ , las entradas a la función  $\Delta$  (ver Eq. (2)) son las siguientes:

- *alternativas* =  $\{v^j_i\}$ , con  $j \in \{1, \dots, \#v^j_i\}$ , esto es, las distintas subvariables;
- *criterios* =  $\{k\}$ , con  $k \in \{1, \dots, s\}$ , que sería los encuestados;
- *valoraciones* =  $\{Q_k(v^j_i)\}$ , respuestas de los encuestados;

- $pesos = \{f(Q_k(e))\}$ , esto es, se usa como peso de cada encuestado  $k$  el resultado de una función  $f$  aplicada sobre la respuesta a la pregunta correspondiente a la variable  $e$  (en columna  $V$  del Anexo I) referente a los años que lleva el experto en el sector bancario. En la Fig. 1 se puede observar la función  $f$  elegida, de tal manera que se asume el supuesto de que la cantidad aprendida en relación con el tiempo va incrementándose linealmente en el horizonte temporal entre los 0 y 10 años de experiencia, momento en el que el aprendizaje llega al máximo.

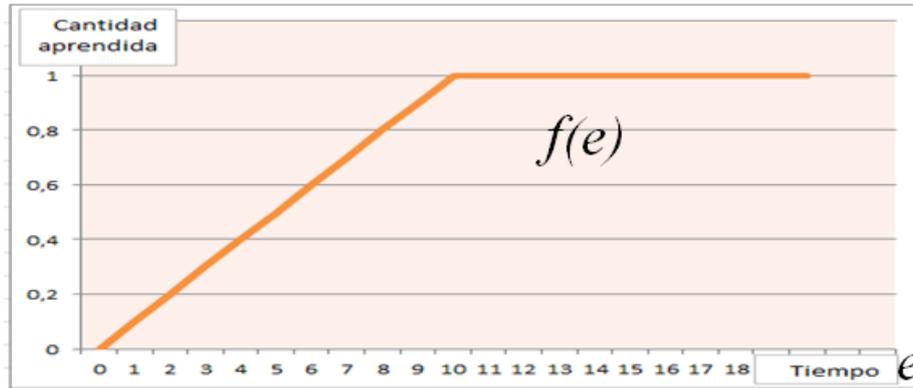


Figura 1. Definición de la función  $f$  sobre la experiencia  $e$

Las salidas de dicho proceso son:

- $preferencias\_colectivas = \{w_j^i\}$ , con  $j \in \{1, \dots, \#v_i\}$ , esto es, corresponde al grado de importancia de las distintas subvariables  $v_j^i$  para cada variable  $v_i$ ;
- $valores\_consenso = \{valor\_consenso_j\}$ , en este caso usamos como umbral  $\gamma = 0.65$ , ya que el consenso entre los distintos criterios (expertos) debe ser suficiente, lo cual se interpreta como que están de acuerdo en el uso de dicha subvariable para el modelo. Si alguna subvariable no lo supera implicaría sustituirla en el proceso TDMC e iniciar una nueva ronda (Herrera-Viedma et al., 2007). Si son varias las que no lo superan, el modelo entero debería de replantearse.

En la Tabla 4 pueden observarse los resultados de estos procesos. El menor consenso es para el valor monetario *Inversión en productos ahorro* pero por encima del umbral establecido. La preferencia colectiva ha determinado que las más influyentes en el modelo son: la frecuencia en el uso de *Medios de pago* y el valor monetario de *Rentabilidad*.

Tabla 4. Resultados de los procesos TDMC para obtener los pesos de las subvariables

Variable RFM( $v_i$ )	Indicador	Subvariable RFM( $v_j^i$ )	$preferencia\_colectivas$	$valores\_consenso$
R	Medios de pago	$v_1^1$	0.464	0.791
	Medios de cobro	$v_2^1$	0.502	0.753
	Interacciones directas	$v_3^1$	0.455	0.791
	Resto de interacciones	$v_4^1$	0.475	0.844
F	Medios de pago	$v_1^2$	0.682	0.714
	Medios de cobro	$v_2^2$	0.515	0.800
	Interacciones directas	$v_3^2$	0.475	0.804
	Resto de interacciones	$v_4^2$	0.483	0.860
M	Financiación hipotecaria	$v_1^3$	0.348	0.719
	Otros tipos financiación	$v_2^3$	0.305	0.721
	Productos parabancarios	$v_3^3$	0.535	0.776
	Rentabilidad	$v_4^3$	0.631	0.728
	Inversión productos a la vista	$v_5^3$	0.555	0.786
	Inversión productos ahorro	$v_6^3$	0.543	0.684

### 3. Modelos TDMC para obtener los pesos de cada variable del RFM

Al igual que otros autores han propuesto (Rezaeinia et al., 2012; Zaheri et al., 2002), se realiza un proceso TDMC para obtener los pesos  $w_i$  de las variables  $v_i$ ,  $i \in \{1, 2, 3\}$  (Eq. (1)).

Las entradas a la función  $\Delta$  (Eq. (2)) son las siguientes:

- *alternativas* =  $\{v^i\}$ , esto es, las distintas variables RFM;
- *criterios* =  $\{k\}$ , con  $k \in \{1, \dots, s\}$ , que sería los encuestados;
- *valoraciones* =  $\{Q_k(v_i)\}$ , respuestas de los encuestados;
- *pesos* =  $\{f(Q_k(e))\}$ , con la misma función  $f$  definida en la Fig. 1.

Las salidas de dicho proceso son:

- *preferencias\_colectivas* =  $\{w_i\}$ , corresponde al grado de importancia las distintas variables  $v_i$ ;
- *valores\_consenso* =  $\{valor\_consenso_i\}$ , estableciendo otra vez el umbral mínimo  $\gamma = 0.65$  para dar validez al modelo.

En la Tabla 5 se muestran los resultados obtenidos. En este caso, los valores de consensos son bastante altos, pudiéndose interpretar que los expertos están bastante de acuerdo con el uso del modelo RFM de cara al cálculo del valor del cliente. Según la preferencia colectiva la variable más influyente es claramente el valor monetario muy por encima de las demás, por otro lado, la recencia está algo por encima de la frecuencia.

Tabla 5. Resultados del proceso TDMC para obtener los pesos de las variables

Variable RFM( $v_i$ )	<i>preferencias_colectivas</i>	<i>valores_consenso</i>
R	0.454	0.828
F	0.410	0.777
M	0.700	0.808

### 3.2 Cálculo del modelo RFM para cada cliente

Una vez definido el modelo RFM, se aplica a los clientes de la entidad siguiendo los pasos análogos al proceso convencional (explicados en la Sección 2.1) de tal manera que se redefinen los pasos 1 y 2, según se explica a continuación, permaneciendo inalterables los pasos 3 y 4:

1. **Adquisición y preparación de datos.** Sea *Transacciones* (*ClienteID*, *Fecha*, *Indicador*, *Importe*) la tabla con los datos de los clientes, tal que  $Fecha \in [Fecha_i, Fecha_f]$ , esto es, el periodo analizar, e *Indicador* con los valores definidos en la Tabla 1. Por tanto esta tabla tendría, el detalle de uso de los servicios de medios de pago, medios de cobro, interacciones directas e indirectas con el cliente, y posiciones diarias en productos de financiación hipotecaria, resto de financiaciones, inversión en vista y ahorro, productos parabancarios y la rentabilidad imputada al cliente mediante el proceso de contabilidad analítica de la entidad.
2. **Agregación de los valores RFM.** A partir de los datos de la tabla *Transacciones*, se obtiene la tabla *ClienteTransacciones* (*ClienteID*, *Recencia*, *Frecuencia*, *Monetario*). Con dicho fin, en primer lugar se obtienen las subvariables del modelo aplicando las correspondientes funciones de agregación a nivel de cliente y teniendo en cuenta el valor del atributo *Indicador* para determinar el correspondiente valor  $j$  (ver Tabla 2):

$$v_j^i = \begin{cases} last(Fecha) - Fecha_r, & \text{con } Fecha_r > Fecha_f \text{ si } i = 1, \\ count(ClienteID) & \text{si } i = 2, \\ avg(Importe) & \text{si } i = 3. \end{cases}$$

Se hace uso de las funciones *last* (ver Sección 2.1) para el cálculo de las subvariables de recencia, *count* (también vista en Sección 2.1) para las de frecuencia y *avg* para las de los valores monetarios, de tal forma que obtiene el valor medio de dichas posiciones monetarias en el periodo analizado, para cada cliente.

En segundo lugar, se calcula  $v_i$ , que se corresponde a las columnas de la tabla *ClienteTransacciones*, de tal modo que  $v_1 = \text{Recencia}$ ,  $v_2 = \text{Frecuencia}$  y  $v_3 = \text{Monetario}$ , mediante la siguiente fórmula:

$$v_i = \frac{\sum_{j=1}^{\#v_i} \text{ntil}(n, v_j^i) \times w_j^i}{\sum_{j=1}^{\#v_i} w_j^i}.$$

Por tanto, la variable  $v_i$  se calcula en función de todas sus subvariables, previamente transformadas mediante la función *ntile* (ver Sección 2.1) que obtiene ordinal entre 1 y  $n$ , de tal manera que el valor 1 contiene los peores valores de la correspondiente subvariable  $v_j^i$  y el valor  $n$  los mejores clientes según dicha variable. Dicho valor es ponderado por el correspondiente peso  $w_j^i$  de dicha subvariable.

## 4 Conclusiones y trabajo futuro

En este artículo se ha propuesto un modelo RFM útil para el cálculo del CLV con una visión más amplia que los existentes (Wu et al., 2005; Hsieh, 2004; Sohrabi & Khanlari, 2007; Mohammadi et al. 2014; Aggelis & Christodoulakis, 2005; Nikumanesh & Albadvi, 2014) de lo que la dimensión cliente significa actualmente para las entidades financieras, integrando cualquier posición en productos que el cliente pudiera mantener, uso de servicios y posibles interacciones realizadas por los clientes, tanto presenciales como de otro tipo. Gracias a los modelos TDMC, se ha podido configurar el modelo final con amplio consenso entre las áreas de la entidad que más relación tienen con los clientes incluyendo marketing, la fuerza de ventas y el soporte al cliente. Se ha considerado también la experiencia de los integrantes de dichas áreas en aras de conseguir una mayor fiabilidad de los resultados. Fruto de dichos procesos TDMC, los expertos han terminado dando bastante importancia a los aspectos monetarios de los clientes con predominio de un aspecto tradicional, básico en cualquier modelo CLV, como es el de la rentabilidad del cliente.

Este último aspecto menos disruptivo, con respecto a modelos actuales, se considera crucial y permitirá implantar más fácilmente el modelo para la gestión directa e individual de los clientes. Por otra parte, los aspectos más novedosos del modelo permitirán adaptarse a las entidades financieras a nuevos competidores como los de los productos *fintech* que basan sus estrategias comerciales en ofrecer inicialmente servicios adicionales a los clientes de los bancos sin que éstos tengan que dejar de ser clientes de sus bancos tradicionales, con lo que mantienen, en una primera etapa, tanto posiciones como rentabilidades. La irrupción de estos nuevos competidores podrá ser detectada por el modelo ya que tanto la recencia como la frecuencia de las interacciones con la entidad (incluidas las no contables) se verán claramente afectadas, provocando la correspondiente bajada del valor del cliente que empieza a usar *fintech*. Además se podría identificar que subvariables en concreto han provocado dicha bajada pudiéndose adoptar las medidas oportunas para dichos clientes.

Además de esta aplicación práctica del modelo, se está trabajando en su extrapolación a otros sectores como el de telecomunicaciones.

## Referencias

- Aggelis, V., & Christodoulakis, D. (2005). RFM analysis for decision support in e-banking area. *WSEAS Transactions on Computers*, 4(8), 943-950.
- Carrasco, R. A., Villar, P., Hornos, M. J., & Herrera-Viedma, E. (2011). A linguistic multi-criteria decision making model applied to the integration of education questionnaires. *International Journal of Computational Intelligence Systems*, 4(5), 946-959.
- Carrasco, R. A., Blasco, M. F., & Herrera-Viedma, E. (2014, July). An implementation of a linguistic multi-criteria decision making model: an application to tourism. In *International Conference on Rough Sets and Current Trends in Computing* (pp. 232-239). Springer International Publishing.
- Carrasco, R. A., Blasco, M. F., & Herrera-Viedma, E. (2015a). A 2-tuple Fuzzy Linguistic RFM Model and Its Implementation. *Procedia Computer Science*, 55, 1340-1347.
- Carrasco, R. A., Sánchez-Fernández, J., Muñoz-Leiva, F., Blasco, M. F., & Herrera-Viedma, E. (2015b). Evaluation of the hotels e-services quality under the user's experience. *Soft Computing*, 1-17.
- Cid-López, A., Hornos, M. J., Carrasco, R. A., & Herrera-Viedma, E. (2016). Applying a linguistic multi-criteria decision-making model to the analysis of ICT suppliers' offers. *Expert Systems with Applications*, 57, 127-138.
- Hsieh, N. C. (2004). An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert systems with applications*, 27(4), 623-633.
- Hernández, C. (2014). El Valor del Cliente aplicado a entidades financieras. *E-prints UCM*.
- Herrera-Viedma, E.; Alonso, S.; Chiclana, F. and Herrera, F. (2007) A consensus model for group decision making with incomplete fuzzy preference relations. *IEEE Transactions on Fuzzy Systems*, New York, NY: Institute of Electrical and Electronics Engineers, Inc.15:863-877
- Mohammadi, R., Bidabad, B., Nourasteh, T., & Sherafati, M. (2014). Credit Ranking of Bank Customers (An Integrated Model of RFM, FAHP and K-means). *European Online Journal of Natural and Social Sciences*, 3(3), 564.
- Nikumanesh, E., & Albadvi, A. (2014). Customer's life-time value using the RFM model in the banking industry: a case study. *International Journal of Electronic Customer Relationship Management*, 8(1-3), 15-30.
- Rezaeinia, S. M., Keramati, A., & Albadvi, A. (2012). An integrated AHP-RFM method to banking customer segmentation. *International Journal of Electronic Customer Relationship Management*, 6(2), 153-168.
- Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *Journal of marketing*, 68(1), 109-127.
- Sohrabi, B., & Khanlari, A. (2007). Customer lifetime value (CLV) measurement based on RFM model. *Iranian Accounting & Auditing Review*, 14(47), 7-20.
- Tsiptsis, K. K., & Choriantopoulos, A. (2011). *Data mining techniques in CRM: inside customer segmentation*. John Wiley & Sons.
- Valenzuela Fernández, L., García de Madariaga Miranda, J., & Blasco López, M. F. (2007). Orientación al valor del cliente y las nuevas métricas de marketing. Revisión y análisis. *Panorama Socioeconómico*, (34), 70-75.
- Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of marketing*, 68(4), 106-125.
- Wu, R. C., Chen, R. S., & Chen, C. C. (2005, July). Data mining application in customer relationship management of credit card business. In *29th Annual International Computer Software and Applications Conference (COMPSAC'05)* (Vol. 2, pp. 39-40). IEEE.

- Zaheri, F., Farughi, H., Soltanpanah, H., Alaniazar, S., & Naseri, F. (2002). Using multiple criteria decision making models for ranking customers of bank network based on loyalty properties in weighted RFM model. *Management Science Letters*, 2(2), 697-704.
- Zeithaml, V. A., Lemon, K. N., & Rust, R. T. (2001). Driving customer equity: How customer lifetime value is reshaping corporate strategy. *Simon and Schuster*.

## Anexo I. Cuestionario Q

V	Pregunta	Na- da	Muypo- co	Po- co	Normal	Mu- cho	Muchí- simo	Total- mente
$v_1$	La <u>Recencia</u> se refiere al momento en el que el cliente efectuó su última interacción proactiva con la entidad bancaria. En general, puntúa la importancia que tiene que una operación sea <u>reciente</u> para valorar a nuestro cliente.	0	1	2	3	4	5	6
$v_2$	La <u>Frecuencia</u> es el número de veces que un cliente adquiere un producto o hace un uso de un servicio en un período determinado. En general, puntúa la importancia que tiene que una operación sea frecuente para valorar a nuestro cliente.	0	1	2	3	4	5	6
$v_3$	El <u>Valor monetario</u> hace referencia a la cuantía del producto o servicio. En general, puntúa la importancia que tiene el volumen de negocio global del cliente (inversión, financiación, productos par bancarios y rentabilidad) en su valoración.	0	1	2	3	4	5	6
$v_1^2$	Respecto, al uso de los servicios de <u>medios de pago</u> por parte del cliente (tarjetas, cheques, recibos domiciliados, transferencias...): indique la importancia de la <u>frecuencia</u> del uso de dichos servicios.	0	1	2	3	4	5	6
$v_1^1$	Respecto, al uso de los servicios anteriores: Indique la importancia de que el uso de dichos servicios sea <u>reciente</u> .	0	1	2	3	4	5	6
$v_2^2$	Respecto, al uso de los servicios de <u>medios de cobro</u> (nóminas, prestaciones, pensiones, otros ingresos recurrentes...): Indique la importancia de la <u>frecuencia</u> del uso de dichos servicios:	0	1	2	3	4	5	6
$v_2^1$	Respecto, al uso de los servicios anteriores: indique la importancia de que el uso de dichos servicios sea <u>reciente</u> :	0	1	2	3	4	5	6
$v_3^2$	Respecto, a las <u>interacciones directas</u> que tiene el cliente con la entidad con presencia física en oficina u otro lugar: Indique la importancia de la <u>frecuencia</u> de dichas interacciones.	0	1	2	3	4	5	6
$v_3^1$	Respecto, a las interacciones anteriores: Indique la importancia de que las interacciones sean <u>recientes</u> .	0	1	2	3	4	5	6
$v_4^2$	Respecto, al <u>resto de interacciones</u> que tiene el cliente con la entidad (electrónica, en cajero, telefónica...): Indique la importancia de la <u>frecuencia</u> de dichas interacciones.	0	1	2	3	4	5	6
$v_4^1$	Respecto, a estas últimas interacciones: Indique la importancia de que las interacciones sean <u>recientes</u> .	0	1	2	3	4	5	6
$v_4^3$	El valor monetario <u>invertido</u> en productos a la <u>vista</u> (cuentas corrientes, libretas, etc.) afecta en la valoración del cliente.	0	1	2	3	4	5	6
$v_5^3$	El valor monetario <u>invertido</u> en productos de <u>ahorro</u> (plazos fijos, fondos de inversión...) afecta en la valoración del cliente.	0	1	2	3	4	5	6
$v_1^3$	El valor monetario de la <u>financiación hipotecaria</u> total de un cliente afecta en su valoración.	0	1	2	3	4	5	6
$v_2^3$	El valor monetario de <u>otros tipos de financiación</u> (préstamos personal, crédito tarjeta...) total de un cliente afecta en su valoración.	0	1	2	3	4	5	6

$v^3_3$	El valor monetario de los <u>productos parabancarios</u> (valores, Planes de pensiones, Seguros de inversión, etc.) de mi cliente afectan a su valoración.	0	1	2	3	4	5	6
$v^3_6$	La <u>rentabilidad</u> (margen de explotación u ordinario) generada por mi cliente afecta a su valoración.	0	1	2	3	4	5	6
$e$	Número de años de <u>experiencia</u> en el sector bancario:							