

Ciudad de México, 16 de mayo de 2017.

Versión Estenográfica de la Sesión Especializada “Desarrollo de Analytics: aplicación práctica”, dictada por Marcos Fernández, en el marco de la 27 Convención de Aseguradores AMIS realizada en el salón Diezmo 3, del Centro Banamex, en esta ciudad.

Presentadora: Le damos la bienvenida a Marcos Fernández, licenciado en Ciencias Económicas y Actuariales por la Universidad Complutense de Madrid, es socio responsable de la práctica global de seguros con más de 22 años de experiencia; skilled person de la Prudential Regulation Authority y la Financial Conduct Authority del Banco of England y profesor en el Master de Inspectores del Banco de España.

Tiene experiencia en proyectos, bueno, en varios proyectos, como planes directores de Solvencia II en 15 entidades aseguradoras, como en España, Portugal, Irlanda, Latinoamérica, así como P&U de proyectos en 10 entidades.

También es responsable del seguimiento normativo de Solvencia II a través de sus distintos grupos de trabajo nacionales e internacionales.

Implantación de end to end finalizada de un sistema que da soporte a los requerimientos de Solvencia II en todos sus pilares, gestión multidisciplinar de servicios para el desarrollo de modelos internos en EY, desarrollo de áreas back office de riesgos de Solvencia II para dar soporte a los requerimientos de TI y proyectos de entidades líderes con servicios y seguros de asistencia, también como rediseño de varios procesos.

Adelante.

Act. Marcos Fernández: Muchas gracias. Sé que es la última, no sé si va a llover o no va a llover aquí en México, por lo cual todos tendrían que salir corriendo, pero prometo que la charla que vamos a tener ahora, no sé si va a ser la mejor pero desde luego de las mejores, ¿por qué? Porque es una aplicación práctica, son cosas que estamos viviendo en Europa y tienen impacto directo en el P&L, con lo cual, me atrevo a decir antes de empezar que son iniciativas reales no teóricas y que todos los que estés aquí, en breve, o estés implantando o estaréis implementando.

Yo deciros agradecer, por supuesto, a la AMIS, la oportunidad de estar aquí hoy, desde abril del 2013 vengo aquí prácticamente todos los meses, creo que sonará esa fecha, por lo cual lo primero es dar la enhorabuena en general al mercado asegurador mexicano, en particular a los asistentes, porque la implementación que habéis hecho es todo un record.

Yo vengo de implementar modelos de Solvencia II en Europa, en países anglosajones y en Latinoamérica, como sabéis, unos empezaron antes que vosotros, pero están todos bastante lejos de tener el tema de conversación que estamos o que estáis teniendo hoy, con lo cual yo creo que desde luego hay una sacrificio por detrás por parte de todos vosotros, de las instituciones que yo creo que es digno de mención, gracias a lo cual estamos haciendo muchos proyectos en otros países, que también soportan.

¿De qué vamos a hablar hoy? Antes de empezar, ¿cuántos actuarios hay en la Sala? Perfecto.

¿Matemáticos?, ¿estadísticos?, ¿físicos? Bien, perfecto, ahora les comentaré por qué, porque yo creo que la charla muy pegada, el panel es de riesgos, vamos a hablar de temas de analytics que obviamente pueden ser utilizados en primera línea o en segunda línea, es decir, pueden ser utilizados por las áreas técnicas o por la función actuarial de segunda línea en riesgos. Lo digo un poco por lo que es el ámbito organizador.

Bien, como he dicho vamos a ir un poquito rápido para cumplir la agenda y la presentación tiene tres bloques, un primer bloque que es el contexto de por qué nace lo que es el analytics o el data science que como sabéis está ahora mismo tan en boga a nivel internacional.

Segundo, y haciendo también mención pues al nombre del evento, de la AMIS, lo quería llamar Customer analytics, porque vamos a hablar de analytics pero con foco cliente; y luego por último, que es donde me voy a parar, si da tiempo, pasaremos también al punto dos, pero al primer punto que es renovación optimizada, prácticamente todo el sector, desde luego en no vida europeo hace este tipo de modelos de renovación optimizada mediante modelos, vamos a decir, modelos GLM o a mí me gusta llamar stating learning mientras que la parte de mejora del proceso de cobranza, sí que se aplican metodologías más de machine learning.

Esto segundo tiene menos sentido en el ámbito asegurador europeo donde prácticamente todas las primas están domiciliadas y se suelen pagar en tiempo y forma, sí que es cierto que se aplica machine learning en utilities, aunque bueno, si te cortan la luz o te cortan el agua, al final, acabas pagando aunque sea un poquito más tarde, pero sí que es cierto que es algo que seguramente vas a estar implementando, estoy convencido, el año que viene por estas fechas, vas a aplicar machine learning en cobranza para mejorar esos ratios y vais aplicar renovación optimizada para ciertos productos.

Vamos allá, primero, contexto. A nadie se le escapa que cada vez hay más datos, más del 90 por ciento de los datos que hoy existen han sido creados en los dos últimos años, no sólo hay más datos, sino que es más barato almacenarlos, su costo se ha reducido en más de un 90 por ciento, pero encima hay más datos, es más fácil almacenarlos y encima es más fácil procesarlo, se han multiplicado desde el año 200 en más de por 300.

Y por último, lo acabo de comentar, se aumenta exponencialmente la aparición del aprendizaje automático, como os habéis compañías, como Facebook, como Amazone, etcétera, ya tiene su machine learning aprendizaje automático implementado y de manera automática capturan información de los clientes y lo aplican dentro de su estrategia.

Abro un paréntesis, imagináros todos los que estáis aquí que representarías compañías de seguros, que en vuestras organizaciones fuereis capaz de recoger toda la información de vuestros clientes en las redes sociales o no son las redes sociales sino también cuando están en el tráfico de vuestra web y fuereis capaz de capturar toda esa información y trasladarle a vuestra red o a vuestro call center quién se va a comprar una casa, quién va a comprarse un coche, quién va a hacer un viaje, quién va a tener un hijo, es decir, que fueres capaz de traducir toda esa información en una campaña concreta para vuestra red, para la venta de un producto. Si eso fuereis capaz de automatizarlo vuestros índices de productividad se aumentarían significativamente.

Esta información también se las voy a dar, pero fijaros en 2012, estábamos por aquí, es decir, cada día, cada día se generaban 2,5 exabyte de datos, en 2012, estamos en 2017, estamos hablando de cerca de 20 mil, con lo cual, bueno, cifras muy altas.

La famosa ley de Moore también, aquí si os dais cuenta en el año 2025 se dice que las máquinas serán capaces de procesar lo que es el human

brain, a veces he oído mucho a los japoneses que ya están pensando en las salas de operaciones en sustituir ciertas tareas por máquinas.

No estamos hablando hoy de eso, pero sí de lo que es la capacidad de procesamiento, vale, y también del precio, es decir, el costo se ha reducido un 93 por ciento.

Esta foto me gusta, es una foto real, acordaros, no hace tanto tiempo, esto era un disco de 5 megas, que ahora lo tenéis en un USB, tampoco ha pasado tanto tiempo.

Primer punto importante. Tenemos más datos, más facilidad para almacenarlos, más facilidad para trabajar con ellos y más facilidad para modelizar, pero al mismo tiempo, nuestros clientes también cambian, nuestros clientes consultan el movi, las personas, más de 100 veces al día, poner vosotros cuántas veces consultáis, a lo mejor eso queda hasta corto, cada 60 segundos se enlazan 2,4 millones de búsquedas. La mayor parte de los consumidores consultan antes en las redes sociales, antes de ir a comprar nuestro producto.

Hay una correlación prácticamente directa entre la calidad del servicio percibida y la siguiente compra, ¿qué vas a hacer? Y el 50 por ciento de los clientes utilizan la oficina selectivamente, es el famoso this is ROPO (Research Online Purchase Offline) la gente mira por la web pero todavía muchos se resisten a dar su cuenta corriente o su tarjetita y todavía van a la oficina, pero van por temas muy concretos y, por supuesto, entradas de nuevos players.

Sin entrar en detalles y luego también, por supuesto, muy importante la segmentación por edades de la utilización —esto es España— la segmentación por edades de la utilización de cada una de las redes sociales.

Por lo cual, todo lo que se ha dicho ahora, perdonen que he ido un poquito rapidito, estamos todos de acuerdo que hay más datos, más fácil de almacenarlos, de procesarlos y de modelizarlos y como habéis levantado la mano la mayor parte que sois actuarios ya vais viendo que no sólo tenéis que modelizar sino que tenéis que ser capaces de ver cómo trabajo con todos estos datos, cómo manipularlos y cómo orientarlos a la gestión.

Por eso, todos los que estéis aquí presentes, más allá de que estéis en primera línea o en segunda línea de defensa, me refiero, tenéis que

tener unas capacidades de análisis del dato, toda esta información, habéis oído de, esto se trata con Python, habéis oído hablar de Hajdu y de cada vez de nombres de lenguajes de programación que antes hablábamos del SQL, del SPS, ahora cada vez hay más lenguajes de programación que son capaces de ayudarte a tratar toda esta información.

Segundo, poder analítico, es decir, tenéis que ser mucho más riguroso en estos modelos predictivos para ponerlos en valor. Tecnología, se acabó y no me enrollar aquí, lo que eran los famosos Data Warehouse con los Data Mart, que traían información del mundo operacional, ahora la información viene no sólo del mundo operacional sino de muchísimos más sitios y la base de datos, ya no es un Data Warehouse, sino son los famosos Data Lake's que hay que atacarlos con otro tipo de herramienta.

Luego, eso sí, todos los que estéis aquí representáis área de riesgos, no se les olvide, dos riesgos que no sé si están o no en la LISF, están en el pilar dos, obviamente, pero no resaltados como tal, que es el gobierno y calidad de datos y el MRM que es el Model Risk Management.

En países como Estados Unidos y, por supuesto, en Europa, estos dos riesgos ya están regulados, es decir, en vuestras áreas de riesgos son dos ámbitos, protección del gobierno de dato y el Model Risk Management que tenéis que monitorear.

Aquí veis que los modelos o las técnicas cada vez van creciendo, si ésta es la light, el año que viene la volvemos a ver, cada vez hay más modelos de clasificación, algunos como vamos a ver ahora mismo, los famosos GLM con los que estaréis seguramente muy familiarizados, son modelos multivariantes, otros que a lo mejor estáis menos familiarizados que son los random forest, que también son árboles de clasificación pero en vez de tomar toda la muestra poblacional, toma muestras más pequeñas y luego hace una media de todas estas muestras. Bien.

Regresiones, motores de optimización, muy importante, cada vez hay más motores de optimización que te ayudan fijando una variable, restringirlo por otra; modelos de clustering, que seguramente estaréis utilizando, simulación para la parte de riesgos, se ha comentado en las pláticas anteriores, la necesidad de utilizar motores de simulación, geometría computacional, etcétera, etcétera. Lo que quiero deciros con este mensaje es que cada vez más hay más modelos, vale, y estos

modelos no son gratuitos, cada uno está pensado para una acción concreta.

Muy importante, yo creo que es algo que nos tenemos un poco que enorgullecer, yo soy actuario, también llevo 23 años en la profesión, no empecé a los cinco años como Ángeles, empecé un poquito más tarde, pero deciros que ahora estamos enhorabuena y ahora mismo en casi todos los países la figura del data science es una de las figuras más demandadas, más remuneradas, seguro que alguno de vosotros estaría frotándose las manos si yo soy un data science.

A ver, un data science tiene que tener ciertas características y es capacidades, me consta que aquí las tenéis, capacidades de ser capaz de manejar muchos datos, con lo cual, quizás un reto o un challenge para vuestras organizaciones en general y para vosotros en particular, y tercero, cero conocimiento del negocio.

Perdonarme porque si no lo suelto reviento, o sea, una competidora nuestra está haciendo un proyecto, una compañía de decesos en España para un modelo de renovación, que lo vamos a ver a continuación cómo se hace y una de las variables, seguro de decesos o de funeraria.

Una de las variables que ha metido en el modelo es la siniestralidad es histórica, imagináros una póliza de decesos, la siniestralidad histórica, sólo te puedes morir una vez; entonces, quiere decir que no sólo sirve hacer buenos modelos, manejar datos sino el conocimiento del negocio.

Si yo pregunto aquí la contactibilidad, ¿pensáis que es bueno o malo a la hora de renovar una prima? Si yo le quiero renovar una prima, por ejemplo, a Carlos, el mes que viene de hogar, si él ha contactado muchas veces con la compañía nada o poco, ¿es bueno o es malo para renovarle, para subirle la prima? Es una pregunta que si no conoces el negocio.

Yo por ejemplo soy miembro de una sociedad deportiva y me dijo uno de los directores que teníamos que contactar con la mitad de nuestros asociados que llevan más de cinco años sin pisar las instalaciones pero pagan su cuota, ¿les contactarías o no? No, lo mismo él, pero eso es porque tú conoces, ¿a quién vamos a contactar? A los que se han metido hace dos años o menos de un año y llevan cuatro o cinco meses, quiere decir, es muy importante el conocimiento del negocio.

Aquí entramos ya en materia, esto ya sí que es parte de la presentación que he venido ya a contaros.

He pintado aquí un ciclo de vida del cliente, el cliente nace, lo primero que hay que hacer es captarlo, ¿no? Una vez que lo captamos, lo que podemos hacer es venderle alguna póliza más o alguna cobertura más.

Tercero, hay que fidelizarlo, una vez que lo tenemos fidelizado hay que retenerlo y una vez que se nos ha ido, tenemos que intentar recuperarlo, bien.

Entonces, os preguntareis cómo podemos ayudar dentro del ciclo de vida del cliente desde el mundo analytics del cual vosotros os tenéis que formar parte de esta explotación, ¿en qué tipo de modelos se están haciendo? En Europa, todo lo que veis aquí, al final dejaré unos trípticos, son cosas que se están haciendo, no es un marco teórico, son cosas que están utilizando para mejorar la cuenta del P&L.

En la captación, me faltaría aquí un punto, que es el que he comentado antes, imaginaros si vosotros pudierais tener acceso a toda la información de vuestros clientes, en redes sociales, ser capaces de bajar la información y enteraros, por ejemplo, que Marcos Fernández, en este caso yo, soy miembro de una asociación de dos mil personas de deportistas, por ponernos un ejemplo, y una gente vuestro es el que me está vendiendo la póliza de autos o de hogar. ¿Se entiende? Es decir, se puede traducir, si sois capaces de traducir cada una de esas informaciones en una acción comercial concreta, eso vale su peso en oro.

Voy a ir pasando, si os parece, aquí vamos a seguir en captación y vamos a ir viendo.

Segmentación. Imaginaros si vuestra compañía segmenta mejor que la compañía de al lado a la hora de estrategia comercial por segmento, campañas comerciales, acciones de fidelización, aquí normalmente la técnica que más se suele utilizar, la más sencillita son los clúster, los clustering que seguramente los venís utilizando.

Esto es muy importante, cuando hablaba al principio de utilizar stating learning o machine learning, la machine learning normalmente utiliza muchas más variables, muchas más variables y son cosas que están cambiando mucho más en el tiempo y una persona hoy está en un segmento pero mañana puede estar en otro segmento, se casa o se

divorcio o si tiene un hijo o si no lo tiene o si se va a vivir al extranjero, es decir, todo ese tipo de cosas le pueden cambiar dentro de lo que es su clúster.

Segundo —perdonar que voy rápido para centrarme en el caso práctico— pricing de no vida y pricing de vida. En el pricing cada vez más se utilizan, yo aquí he hecho bastantes mystery shopping, entonces, muy poquitas variables, los que, algunos de los aquí que hemos tenido la suerte o la desgracia de trabajar en países anglosajones, yo, la primera vez que llegué, un seguro del hogar, te preguntaban en la casa cuántas habitaciones tenías, no sólo cuántos metros cuadrados, te preguntaban cosas para un seguro del hogar, te preguntan cosas de ti no de la casa, cuántos metros tiene, si está en altura, etcétera, sino cosas de ti, es decir, que tu comportamiento afectaba a lo que es el pricing. Quiere decir que yo reconozco que en el proceso de comercialización no podemos pedir muchos datos a los clientes porque eso inflaciona el tiempo comercial y puede ser que el cliente se canse, pero sí podemos obtener muchos datos de esa información que pedimos relacionados.

Como os comento, esto ha sido un cambio, esta información la dejaré bastante, bastante estructurado, esto es un ejemplo real, esto es Londres, desde un barrio a otro barrio, una misma persona viviendo en un barrio del centro o el otro, en 50 kilómetros tiene diferencia de 20 libras para un mismo seguro de vida riesgo, eso es un poco el nivel del fine tune de zonificación, en este caso, de zonificación que tiene ahora de tarificado.

Propensión a la compra, aquí veo algunos colegas de España que lo conocen muy bien, en países, como es el caso de España, hay competidores puramente digitales, como puede ser línea directa, compañías que para ese mismo producto son puramente de redes agencial entre medio, banco y aseguradoras y todas compiten juntas.

Claro, fijaros aquí, ¿qué hacen los analytics? Tienes que hacer un estudio de propensión a la compra de tus clientes antes de acercarte a él, hay clientes que lo que quieren es la prima mínima del mercado. Otros, la prima media y otros la prima máxima.

Una compañía con un portfolio muy grande está más cerca de por aquí, mientras que compañías que acaban de irrumpir o de directo están más por aquí.

Entonces, ¿nuestra misión cuál es como analytics? Es identificar cuál es la probabilidad de compra de cada uno de los clientes. Si me habláis de autos, en España todo el mundo, como lleva multitarificador, siempre piensan en el más barato, en la verde, pero si por ejemplo es una póliza de hogar, como no han llegado los multitarificadores suele buscar tres precios. Por lo cual se va más a la prima media, a la hora un poco de ir pensando cuál puede ser la propensión a la compra. Bien, se entiende.

Luego, profit testing, esto está muy vinculado a la charla anterior, es decir, los profit testing, como sabéis son, para lanzamiento de un nuevo producto no para el pricing de un anual renovable, tenemos que proyectar variables económicas, variables no económicas y también tenemos que proyectar el capital, esto sobre todo tiene muchísimo impacto en los productos de vida ahorro y productos más con componente financiero.

Esto son precios de competencia, yo sé que muchos de los que estáis aquí presentes ya lo estáis haciendo en vuestras organizaciones, pero si algunos participáis en pricing no se puede hacer un pricing sin conocer los precios de competencia y ver cómo está posicionada vuestra cartera.

Yo, me atrevo un poco a preguntaros, me podéis contestar o no, si vosotros conocéis o es elegible el ramo que queráis, salud, autos, hogar, si vuestra cartera, imaginaros que es esto, ¿qué porcentaje está a la izquierda? Es decir, esto es más barato que la competencia y qué porcentaje está a la derecha, es decir, más caro.

Si encima os pregunto, para cada una de estas columnitas por edades, es decir, en este caso se ve que los que son de edades más jóvenes, esta compañía, son más baratos, mientras que los que son más mayores son más caros, justifica cartera de viejecitos, que esos no le baja la prima y los deja ahí quietecitos, pero los que están entrando o más jóvenes les tienes que poner una prima más barata.

Esta proporción es fundamental, disponer de modelos que te permitan ver cómo están tus precios respecto a la competencia y por zona geográfica. España es igual que México, igual que la mayor parte de los países, la zonificación es muy importante.

Después tenemos lo que son la parte de geolocalización de oficinas, lo tengo puesto en captación, tiene normalmente dos usos, un primer uso es ayudar a vuestros agentes y a vuestras oficinas a mejorar su

performance de nueva producción, de retención, ver cuánto tiene que hacer más descuentos o menos descuentos.

También se pueden utilizar la geolocalización para si se van a abrir oficinas ver en cuáles tienen más potencial o menos potencial.

Up selling y cross selling, no sé si estáis utilizando o no actualmente en vuestras organizaciones la información que tenéis para darle al área comercial un numerito de cuál es la probabilidad de que compre otro producto o de los que tiene los eleva. Esto que puede ser un tema que a lo mejor lo veis desde lejos, esto es si tenéis los datos para en vuestro portfolio en base a antigüedad, en base a siniestralidad, etcétera, podéis acertar mucho más en la probabilidad de compra de un nuevo producto y por tanto ahorrar muchos costos comerciales, de campañas, etcétera.

Infraseguro, esto pasa mucho, sobre todo, en seguros del hogar, no se sabéis, cuando tú te compras una casa, pones a los 30 años, ¿qué tiene la casa? Un frigorífico, una televisión muy grande y una cama, pero cuando vas cumpliendo años y llegas a los 55, a los 60, la casa parece un museo, sin embargo, porque tiene jarrones, tiene cuadros, tiene de todo y sin embargo, habéis podido pensar en vuestra compañía si renováis los capitales, las coberturas, etcétera. Esto pasa muchas veces que tiene la misma prima y con los mismos capitales.

Esto es un tema que también, insisto, cada uno de estos puntos requiere una plática aislada que es el Nest Best Opción y Nest Best Action.

El Nest Best Opción es en vez de crearle un producto por cliente es cuál es la cesta óptima del mismo, puede haber dos personas de 45 años con la misma edad, con los mismos estudios, mismo trabajo pero uno casado, otro no, o los dos casados, pero uno con hijos, otro sin hijos o los dos casados con hijos, pero uno viaja y otro no viaja, al final, la cesta óptima la construye también una serie de patrones o de uso de la información de vuestros clientes; y lo otro es fundamental que es cuándo haces esta acción para completar la cesta óptima, que esto es el camino o el pasillo del cliente en la contratación, en la misión, en el siniestro.

Una experiencia concreta de esta semana, me dieron un golpe por detrás con el coche, me llama, llamé para dar el parte, esto es verídico, mientras estaba dando el parte, me ofrecieron un seguro del hogar, o sea, en un despiste, pero yo preocupado del golpe, os lo prometo, no es, o sea, me estaban ofreciendo un seguro del hogar y además sabían que me vencía a los dos meses y me están diciendo que me hacían un

descuento del 20 por ciento, yo bastante encabronado al call center que además era como un ordenador, o sea, era pa, pa, pa, mira, solucióname lo otro primero; es que cuando acabó me ofreció el de salud, entonces, esto no es una broma, o sea, es, me ofreció dos seguros durante la declaración de un siniestro.

Entonces, si no lo llevamos tanto al extremo, que a mí me parece eso que es un poco osado, pero intentar a verlo en positivo, si el momento hubiera sido después de arreglar del coche, que te lo dejan limpio, ¿no? Que te lo dan, que todo lo dan, pues a lo mejor ahí con un estudio de propensión hubiera sido mucho más fácil.

Entonces, es muy importante qué producto y en qué momento, y eso se hace también, también con clustering, algoritmo sobre grafos y también modelos GLM multivariantes.

Por supuesto, valor del cliente, supongo que todos hacéis cálculos de valor del cliente, el valor actual es más sencillo con las pólizas que hay, el valor futuro también es predecir los ingresos y egresos probabilizados y éste que es más complicado te lo suele dar también las redes sociales y la información por fuera que es, relaciones de valor, grupos de afinidad, si eres miembro de tal, en qué compañía trabajas, etcétera.

Y bueno, relación optimizada, me lo dejo porque lo vamos a ver ahora, bonus-malus, modelos de cobranza, esto realmente lo que persigue es segmentar todos los que no te están pagando por intención de pago y capacidad de pago, clusterizarlos y poner medidas concretas para cada uno de ellos.

Es impresionante cómo se mejora la cobranza, el coste y la velocidad de gestión cuando aplicas metodologías de machine learning que lo que haces es coger toda esta información e irte montando una serie de patrones. Si nos da tiempo lo veremos.

Y luego por supuesto fuga anticipada, que es un tema que yo estuve viendo en México, como he dicho yo, cuatro años viniendo a México de manera muy continuada y veo que es algo que también nos ocurrió en España, los colegas que estáis aquí que lo conocen, les sonará más cuando se van comoditizando ciertos productos y la gente empieza a cambiar mucho más de compañía, si no los tiene segmentados y no predices quién se te va a ir, pues no puedes hacer una serie de campañas de fidelización. Todo esto se hace también con dicha información.

Como lo prometido es deuda vamos allá. Lo mismo me sorprendéis y algunos ya lo vienen haciendo y quedo mal. Primero, los modelos de renovación, primero, ya son modelos multivariantes, es decir, ¿qué quiere decir multivariantes? No se quedan sólo en árboles con hojas, edad, por zona geográfica o tipo de casa o coche, etcétera, sino que estos arbolitos con hojas se van traduciendo ya modelos multivariantes, estos modelos multivariantes incluyen también datos internos del cliente, es decir, no sólo del bien asegurado sino también, oye, este cliente qué valor tiene, cuántos productos, tiene más de uno o tiene dos, qué antigüedad tiene, tenemos algún tipo de indicador de satisfacción o transaction a NPS que pueda utilizar, etcétera, etcétera.

Después de incorporar datos internos, se empiezan a incorporar datos externos, esto, hay muchas compañías, me consta México porque trabajamos también con ellas, que lo que hacen es ir al mercado a traerte información de zonificación, de tus clientes, de capacidades económicas, etcétera, etcétera.

Esto que veis aquí es lo que hemos comentado antes, que son los precios de competencia y éste, digamos, es el clave, cuando vas a subir o bajar la prima en el modelo de renovación tienes que hacer un price elasticity o elasticidad al precio de ese cliente concreto porque no somos todos iguales. No nos podemos fijar solo en los modelos de prima pura que os sonarán, que eso también, quien más siniestralidad tiene que lo pague, pero también, oye, tenemos que ver cuál es el price elasticity de ese cliente.

Luego vienen modelos de optimización, modelos matemáticos muy sencillos de optimización de variables, vamos allá, básicamente, la relación optimizada lo que consiste es, imaginemos que tenemos dos pólizas y que el objetivo de renovación de la dirección es un 4 por ciento, que quiere crecer.

Estas dos pólizas si las veis, básicamente, una es de 261 euros, es muy parecida, su siniestralidad, sus gastos, su beneficio, lo que tratamos es, en vez de subirle el 4 por ciento a Miguel y a mí, oye, a lo mejor a Miguel se le puede subir un 5 por ciento porque tiene una sensibilidad al precio menor, pero a mí, subírmelo un 3 por ciento, al final, lo que conseguimos es que la tasa de renovación conjunta sea mejor, es decir, sobre un clúster, no subas a todo el mundo un 4 por ciento, inténtalo ver cuál es el price elasticity de cada uno de ellos.

Para ello, ¿qué tenéis que hacer? Tenéis que hacer un modelo de prima pura, que lo vamos a ver, un modelo de fuga y un modelo de optimización.

¿Cómo se hace el modelo de prima pura? Lo conocéis, cuál es lo único diferencial y os lanzo también la pregunta, en los modelos de renovación miráis o no miráis la siniestralidad, predecís o no predecís la siniestralidad de esa póliza a futuro, porque a lo mejor decís, no Marcos, subimos un 4 por ciento o un 3, lo que dice el Jefe, o a unos es un 3, otros son un 4 o 5, pero no vais póliza a póliza, tratando de prever porque la siniestralidad pasada no quiere decir que vaya a ser la siniestralidad futura. Entonces, tenéis que intentar predecir cuánto va a ser, esto es, digamos, lo más importante.

Los modelos sabéis hacerlo, son modelos GLM que como sabéis son sencillos, ¿lo importante qué son? Las variables, es ir dando en la tecla con qué variables os van a predecir más la siniestralidad futura de esas pólizas que tenéis en vuestro portfolio. ¿De acuerdo?

Lo segundo es algo nuevo que muchas veces organizativamente está el área de marketing trabajando con ello y lo que tenemos que hacer es integrarlo en los pricing's de renovación que son los modelos de fuga, es obvio, cuanto más precio, cuanto más precio por la probabilidad de caída, es mayor, cuanto menos precio la probabilidad de caída es mucho menor.

Aquí lo importante, ¿qué es? Y seguro que también les lanzo la pregunta, necesitamos contar con los incrementos de prima pasados, cuántas compañías guardáis en vuestro operacional cuando un cliente al renovar, no ha renovado y se ha ido, habéis guardado el perfil de ese cliente, cuánto se le ha subido la prima y queda anulado, porque si a ti te han subido la prima un 10 por ciento y has cancelado, al año siguiente si hay 10 como tú dentro de tu portfolio tienes que aprender de la elasticidad al precio de ese perfil, que han subido un 10 por ciento se han ido.

Entonces, lo que hay que hacer es reconstruirlo, se puede hacer muchas veces por la carta, supongo que aquí en México envías una carta para lo que es la renovación, es decir, se puede reconstruir.

Aquí también es súper importante las variables, esto es lo que aporta fundamentalmente estos proyectos, cuando lo hacéis una vez vais aprendiendo y os vais dando cuenta mediante grupos de control cuáles

son las variables que realmente son más discriminantes a la hora de tener elasticidades al precio.

Nosotros hemos aprendido cosas, como por ejemplo la variable que llamamos de (...), lo es lo mismo quedar en un siniestro, los últimos tres meses que quien lo ha tenido en el último año, es brutalmente distinto la memoria de que lo ha tenido hace poco tiene mucha más probabilidad, la zonificación, en el caso de hogar, es decir, hay ciertas variables, el piso, la altura de la vivienda, son variables que al final tú no las tienes en cuenta en tu prima pura pero cuando las metes dentro de tu elasticidades al precio te das cuenta la diferencia que tienes.

Bien, una vez que tenemos, como he dicho, las variables para modelo de fuga, tenemos que hacer un ejercicio, vale, aquí obviamente quitar los nombres, un ejercicio de cada uno de los perfiles de vuestra cartera, supongamos, por ejemplo, automóviles, cuánto está cobrando la competencia para ese mismo perfil, lo ideal, por zona geográfica y por segmento, como actuarios que sois, digo, esto sí tengo una cartera imaginarios de un millón de pólizas, esto es complicadísimo, no, lo vamos a ver después.

Nosotros con 300 pólizas puedes perfectamente replicar carteras enormes si lo haces segmentado, obviamente, esto es la prueba de algodón, que coges una póliza te vas a la competencia a ese perfil y si no te sale una varianza más o menos del 5 por ciento, quiere decir que tienes que volver a seguir haciendo más ejercicios.

Fundamental saber lo que está cobrando la competencia, como veíamos, para ver que el 44 por ciento de tus pólizas son más baratas, perdón, las más caras, en este caso, y el 56 son más baratas.

Por último, un concepto que os empezará a sonar cada vez más que se llama el concepto La frontera eficiente, ¿qué es la frontera eficiente? Y aquí os lanzo la pregunta, ¿cómo se hace el pricing en vuestras compañías? Es decir, tienes un Comité de pricing, me imagino, y ahí cómo se fijan los precios de renovación. Un 4 por ciento, se fija, es que quiero como mínimo tener una retención del 80 por ciento, no quiero que se me caiga más del 20 por ciento o por el contrario, se dice, quiero sacar un margen o una rentabilidad para cada póliza de un equis por ciento.

La frontera eficiente lo que hace es, busca todos los puntos que van optimizando, que van optimizando la tasa de renovación con el beneficio

o el margen, entendiendo por margen, por margen primas, obviamente, multiplicado por la caída, menos siniestros y menos comisión, no nos metemos en gastos por el tema de la elasticidad de tener 10 pólizas más o menos.

Esta frontera que lo vamos a ver ahora, se utilizan, insisto, por las compañías líderes de países que desde luego en Europa y en España, sobre todo, para productos de autos, hogar y salud y cada vez se van incorporando nuevos productos, como decesos o vida-riesgo.

Daros cuenta que la subida de la prima es lo único que está en nuestro tejado, el resto son variables de nuestro cliente, el cliente es como es, el bien asegurado es como es, pero la subida de prima es la palanca que tenemos, es la palanca que tenemos y es vuestra responsabilidad, nuestra responsabilidad porque en función de ello nos tenemos que acercar a esta frontera, esta compañía podemos estar aquí, ¿por qué estamos aquí? Porque estamos subiendo un 4 por ciento a todos los que estamos aquí en vez de subir, a unos son un 6, a otros es un 5 y a otros son 4 y a algunos son 3 para mantener la misma tasa.

Vamos a hacer el ejemplo práctico. Vamos a ver una renovación tradicional y vamos a ver una renovación optimizada.

Renovación tradicional, si no lo hacéis así, por favor, me corregís. Dos clientes, perdonarme que he elegido el ramo del hogar que es el que estará mismo en Europa más en boga, autos está ya más que superado, un cliente que pagó 229 euros, este 284, producto base, una zona, una antigüedad en la casa de cero y 25 años, metros cuadrados de la casa, uso habitacional o podría ser de alquiler, la edad de la persona, pago semestral o anual y no tuvo siniestros, tuvo un siniestro de 150 euros.

Bien, ¿cómo se tarifica ahora mismo o cómo se tarificaba la mayor parte de las aseguradoras? Se cogía o la prima de media de nueva producción o la prima de cartera con el IPC, simplemente, entonces, saldrían 270 y luego empezamos con árboles de decisión, con hojas, estos árboles, lo que vamos haciendo es, pues en este caso por zona geográfica, que aparece en un seguro del hogar por la siniestralidad, efectos meteorológicos, zonificación importante, después antigüedad de la persona y después metros de la casa.

Estas dos personas que hemos ubicado se encuentran aquí y aquí, luego lo que hacen las compañías es, poner unas restricciones de si han tenido siniestro o no han tenido siniestro, una restricción es, oye, no

vamos al que ha tenido siniestro bajarle menos de un 7 por ciento ni subirle más de un 20 por ciento, se ponen ciertas horquillas también reputacionales, imaginar vuestra compañía que luego sale en el mercado que sube a precios de más del 20 por ciento.

Entonces, por cerrarlo ¿cuál sería la tarificación de esta compañía en la renovación tradicional? Sería 229, 284 que es lo venían pagando en el año ene menos uno, en el año ene me sale 245 y 290 y lo que hago es, le aplico estas restricciones. En concreto, cómo está un 7 por ciento que está dentro de esta horquilla pues se queda en los 245 iniciales, como éste, el 2 por ciento se ha quedado por debajo y tuvo siniestro, le subo al 7 por ciento. Se entiende, ¿no? Perdonar que voy un poquito de prisa.

Y esto es, seguro que lo hacéis un poco parecido, pero más o menos, corregirme, es así como tarificáis, ¿no? Más o menos.

Bien, ¿qué es lo que proponemos hacer? Primero, modelo de prima pura, lo que tienes que hacer es buscaros a vuestro asegurado alfa, el alfa es el que más se repite y vuestro asegurado alfa, que es el que más se repite, luego va a tener a otros que se diferencian por una serie de betas y estas betas son distintas variables, cuanto más ricas sean las variables betas más granular será la diferencia con el alfa.

¿Este alfa qué tiene? ¿Qué hemos dicho que haríamos lo primero? Vamos a predecir su siniestralidad futura porque como mínimo, como ya es póliza, póliza le tengo que cobrar eso, estamos de acuerdo.

Bien, pues la frecuencia del alfa es 0,4, es un 40 por ciento de siniestralidad que en España para hogar sería razonable.

Lo que hace es, el que veis en oscurito y en clarito son nuestros dos asegurados y lo que veis es, para cada una de estas variables o estas variables que nos hacen a la beta, he puesto seis ejemplos, metros, siniestros, antigüedad, producto o forma de pago, se van desviando del alfa en función de ese trabajo obviamente que ha hecho una modelización.

En resumen, nuestros dos clientes que les podía haber dejado, si fueran alfa, serían 0,4, voy multiplicando todos estos betas y al final este me queda joven en 0.30 y el mayor en 0.44, es decir la frecuencia, en vez de un 0.40 del cliente uno es 0.30 y del segundo 0.44.

Ahora nos vamos a la severidad, la severidad como sabéis es el coste medio de los siniestros, pues es el coste medio de mi siniestro alfa es 650 euros y busco aquí también cuáles son sus variables más discriminantes que nos llevan a buscar una dispersión de ese coste medio.

En este caso me vuelven a salir también unas variables, esto es un ejemplo, como sabéis en un modelo GLM son puntos, lo que se hace agruparlos para el ejemplo y estos 650 se convierten en 787, en un caso y en el otro caso, 638.

En resumen, multiplicando sus frecuencias y sus severidades ajustando al alfa con sus distintas desviaciones, nos salen 490 euros, ¿qué es 490 euros? La siniestralidad esperada de mi portfolio.

Ahora vamos a seguir, ¿qué tenemos que hacer ahora? El modelo de fuga, modelo de fuga, igual, ¿qué tenemos? El alfa es 0.2 que es la media de mi caída, es decir, mi cartera se me está cayendo para este producto un 20 por ciento, intentar ajustarlo a cada uno de los ramos, lo que buscamos, estas son las variables que os comentaba, he puesto un ejemplo de seis, qué variables van haciendo que sobre este asegurado base sea mayor o menor la probabilidad de fuga, hasta aquí algo bastante normal.

Bien, esto que veis aquí es lo más importante porque eso es la palanca que vuestras compañías tienen para optimizar la renovación, que es lo que estamos hablando, lo he juntado, pero aquí hay dos modelos.

Un primer modelo es la reconstrucción de vuestra historia, como he comentado, los que he ido renovando o no me han renovado, o me han renovado, con esto lo que voy a construyendo es ante incrementos de prima o bajadas de prima, qué porcentaje, qué porcentaje de asegurados con los distintos perfiles me han anulado o no me ha anulado.

Viejecita de más de 65 años que lleva más de 10 etcétera, etcétera, le he subido durante los últimos años un 5 por ciento y no he dicho nada, pues entonces claramente más de un 10 por ciento, pues tendrá un porcentaje muy bajo.

Pero por otro lado lo que tengo también son mis precios de competencia que lo que miden es la diferencia del precio actual de esa póliza de vuestra cartera con lo que cobraría en la competencia en su zona

geográfica y cuánto lejos está obviamente del más barato, es imposible replicar las tarifas de una compañía nada más, obviamente, del más barato o del más caro.

Se entiende, entonces con esos dos modelos construimos la probabilidad ante incrementos o bajada de prima de caída, de cómo se incrementa la caída de este asegurado, con lo cual está un poco exagerado, como podéis ver, pero el que llevaba aquí 20 años con nosotros, pues fíjate, aunque le suba un 20 por ciento, la probabilidad de caída es un 17 por ciento, que está por debajo del 1,2, mientras que el que lleva poquito tiempo con nosotros, menos de un año, si le subo, aun bajándole a un 10 por ciento la prima su porcentaje de caída es un 13 por ciento, ¿por qué estoy pensando en auto? Porque es en autos, acaba de llegar y a lo mejor está muy comoditizado y su precio lo puedo obtener mucho más barato en la competencia.

Con esto, aquí se ve mucho más fácil, esta curva suena, esto es, retención por incremento de primas, son nuestros dos asegurados, este que tenéis aquí es más inelástico, más inelástico a subidas de primas, mientras que este que veis aquí es mucho más elástico a la subida de primas, está más comoditizado. Esto obviamente es beneficio, un euro más, un euro más a P&L, con lo cual esto es totalmente plano.

Aquí el reto está en buscar en la frontera de eficiencia los distintos puntos, los distintos puntos donde manteniendo un porcentaje de retención deseado, obviamente, voy obteniendo un mayor o un menor beneficio.

Antes de que penséis, oye, Marcos, pero esto es un poco arriesgado, me juego la renovación de toda mi cartera, me juego la renovación de toda mi cartera a estos modelos que estoy haciendo, están las restricciones de negocio, las restricciones valen dinero. Si vosotros hacéis la tarificación y os dice el Jefe, no, pero es que no le podemos subir el precio más de un 10 por ciento a los que llevan más de 10 años con nosotros. Perfecto, pero ya tiene un precio y ese precio es que esta frontera, este puntito va a estar más debajo de esa frontera por no hacerlo bien.

Aquí os he puesto unos ejemplos de restricciones que se suelen aplicar, oye, no quiero subir más la prima a un 2 por ciento, una retención mínima del 85. Oye, para que nuestros agentes no se molesten y piensen que hay subida del 15 por ciento, sólo son cuando hay una siniestralidad por encima de cierto importe, es decir, las restricciones de

negocio nos sirven para a la hora de aplicar las subidas de prima en la renovación a cada una de ellas, ponerle, digamos, ciertos stop loss de distinta naturaleza de negocio.

Esto que veis aquí es un ejemplo real, hay unos que lo pueden ver con unos ojos y otro con otros, y os animo a que lo veáis con ojos de analytics.

Imaginaos la situación, Europa, tasa de tipo de interés negativas, precios comoditizados, ya ha habido trabajos de eficiencia, con despido de personas, etcétera, etcétera, los actuarios ahora mismo son los que están salvando la cuenta de resultados, ¿cómo? Mejorando el pricing, el pricing que era algo que estaba ahí apartado y que lo que había que hacer era vender, vender y vender y vender, como no crecen las primas también por la época de la crisis que venimos, digamos, con crecimientos planos de volumen de primas, la única manera y la única palanca que queda es optimizar el pricing, eso no sólo es lo que tardaba pero acabará llegando, acabará llegando a México.

Esto es un ejemplo real, he cambiado algunos números obviamente, esto que veis aquí es la frontera que sale de los modelos sin ningún tipo de restricción, es decir, si os dejaran como analytics que sois reprocesarías toda la cartera con estos precios y saldría esta curva y esta que veáis aquí son con restricciones de negocio que puso la compañía, esto es un ejemplo donde se ve que manteniendo un mismo nivel de retención prácticamente este con este, retención innovación, el beneficio en este caso subía casi un 7 o un 8 por ciento. Y con esto acabo el ejemplo práctico.

Proyectos que hemos hecho en Europa han demostrado beneficios en primas entre el 2 y el 6 por ciento, o sea, las primas, después de aplicar un modelo de renovación entre un 2 y un 6, os decís, es muy grande, esto es como jamón, el primer corte hay mucho, pero cuando vas cortando vas tocando hueso, no es lo mismo una compañía que no haya hecho ningún modelo de estos en el pasado que una compañía que ya lo venía haciendo, no es lo mismo una compañía que venía de tener primas muy altas a una compañía que venía a hacer primas muy bajas, en función del portfolio y de la política de pricing de todos estos años de vuestras compañías, vuestros business key sería, por lo menos el quehacer en España, entre un 2, 6 por ciento en primas y entre un 7 y 10 por ciento en P&L en bottom line, es decir, hay compañías que han ganado un 10 por ciento con sus mismas pólizas simplemente ajustando el precio a lo que es un precio de mercado o un precio elasticity.

Y para acabar pido dos minutos, esto es una aplicación que en México, nosotros desde MS, desde Management Solutions ya estamos aplicando en alguna compañía mexicana, que son modelos de machine learning para mejorar el proceso de cobranza.

Lo voy a hacer bastante rápida, tiene una necesidad de información, una aplicación de unos modelos, una definición de una estrategia de medidas y por último, un gobierno del mismo.

Información, por supuesto, del recibo, de la póliza y de capacidad financiera, si eso es una banca aseguradora más fácil, si no es una banca aseguradora tenéis que inferir la capacidad económica de esa persona de otra manera, dónde vive, qué coche tiene, es decir, con otro tipo de variables.

Los modelos, que es donde me quería parar, son: de segmentación, es fundamental, veis que volvemos otra vez a lo mismo, cómo segmento mi cartera y como analizó o cómo calculo la probabilidad de impago de esta persona. En este caso, en energía suelo utilizar seis clúster, en seguros, la verdad que nosotros estamos utilizando seis también. Un ejemplo didáctico donde lo que tienes es capacidad de pago a voluntad de pago y cuantas más personitas tengas, dependiendo de estos grupos obviamente, la estrategia va a ser distinta, si los tienes aquí, pues estarías conmigo, que si tienen capacidad de pago y voluntad de pago no hay que hacer nada, ellos sólo se corrigen.

Si tenéis muchos con buena capacidad de pago pero con mala voluntad de pago, eso lo que hay que hacer es perseguirlos, estos que tenéis aquí, lo que hay que hacer es ayudarles a lo mejor a venderles otro seguro con menos coberturas, con menos capitales porque a lo mejor es que no lo pueden pagar.

Obviamente, los modelos de segmentación también se sofistican y cada vez pues puedes ir perfilando a tu cartera desde el punto de vista de capacidad, insisto, económica, de voluntad y capacidad de pago, imprudentes que a lo mejor viven al día, se retrasan pero acaba pagando, a lo mejor lo que hay que hacer es adelantarle qué va a tener que pagar, otros son ingenuos, desconocen las consecuencias de la deuda. En el caso de la luz y el agua es que te cortan la luz, en el caso del seguro que puedes perder la protección a tus familiares, en el caso de un seguro de salud o que tu auto se puede quedar sin estar

asegurado. Desafortunado, bueno, inmoral o arruinado, bueno, hay distintas tipologías como veis.

Estos son los modelos que venimos utilizando, venimos utilizando, como veis los modelos de machine learning son modelos que lo que haces es combina la sabiduría de todo el colectivo y son modelos, en este caso, el random forest, si lo veis es como un arbolito de los que venís haciendo vosotros en el pasado y seguro que habéis trabajado, pero en vez de coger toda la muestra va cogiendo porcentajes, 5 por ciento de la cartera, 5 por ciento de la cartera y todos estos luego lo que hace es una media entre todos estos arbolitos.

Y lo mismo para el grading bustin que lo que va haciendo es, va aprendiendo de uno y los que se van saliendo de la curva, los va utilizando en el siguiente modelo para ir aprendiendo.

Son técnicas, insisto, que aunque veis aquí que es un marco teórico, que funcionan, o sea, que funcionan y compañías como Gas Natural, como Fenosa, se utiliza mucho en la banca para, también para procesos de recobros y de refinanciaciones y en el mundo de seguros es verdad, no está todavía tan implementado, es una realidad, es casi más un challenge que tenemos.

También os digo, no sirve de nada segmentarlos bien, hacer un modelito, oye que funcione si luego al final las medidas no son buenas, oye, sé que hay 20 tíos que es que pueden pagar pero hay que ir a su casa a tocarles la puerta, si vas, tocas la puerta y no te la abren, al final no es el que el modelo esté mal, es que la contactabilidad está mal.

Por último, muy importante, medirlo y esto también, os lo aconsejo, cualquier proyecto de analytics que hagáis, ya sea el que he dicho anteriormente, que es un modelo renovación o un modelo de cobranza, esto es como el algodón, o sea, si tú pones que tienes 10 pólizas en tu compañía y las ibas a renovar y te iban a dar equis, las has renovado en base a este modelo y te han dado equis más y, eso se llama que has ganado 10, equis siendo y, equis más 10, has ganado 10 como consecuencia de aplicar bien este modelo.

Y aquí lo mismo, si tú venías cobrando 80 por ciento, 85 por ciento de tu cartera y has pasado a aumentar 5 puntos porcentuales como consecuencia de hacer campaña segmentada y has gastado menos coste y has disminuido el número de días, eso lo traduces a dinero y eso es la mejor financiación para las áreas de analytics. Bien.

Eso es un poco lo que venía a contaros, en la puerta dejaremos cierta información porque como he visto aquí que hay muchos analistas con componente técnico, nosotros pues tenemos formación de todos y cada uno de estos modelos y el enganche con la aplicación, con la tipología de modelos concretos que hemos hablado.

¿No sé si hay alguna pregunta?

Dra. María de los Ángeles Yáñez Acosta: Muchísimas gracias por el tema, Marcos, yo creo que es muy fresco y está muy en el contexto de esta convención que se llama “El cliente del futuro en seguros” y que en el documento sectorial que desarrollamos este año en mi área y que va a estar público en la página de AMIS, por cierto, hablamos mucho de lo que son los modelos de customer centricity, son modelos centrados en el cliente, todo lo que es general realmente un valor y experiencias memorables para los asegurados, diferenciarlos de una manera adecuada, entonces, todo este tema de la segmentación.

Entonces, yo quisiera preguntarte, bueno, dos cosas que tienen que ver directamente con el tema de analytics y una que es cosa curiosa que mencionaste en tu presentación, entonces, si quisieres empezamos por la, bueno.

En la mañana en una de nuestras sesiones plenarias nos dijeron que todas estas nuevas empresas que hacen insurtech y que la mayoría de ellas están ayudando como a distribución o a enfocar, no son compañías de seguros como tal, sino que su herramienta en tal se asocia con alguna aseguradora para ayudarle a hacer esto, la mayoría usan data analytics, entonces es pregunta, ¿alguno de tus clientes o has visto tú que en Europa se den este tipo de análisis en apoyo a aseguradoras? Porque estos modelos de renovación, pero hasta de segmentación como tal creo que son clave en tener la atención al nuevo cliente.

Act. Marcos Fernández: Nosotros ahora mismo como caso real, caso práctico acabamos de llevar una aseguradora que tiene la fábrica, como comentas, asegura, las tiene, como digo yo, tuberías, las personas, no los operaciones a vender productos de ahorro y productos financieros a través de fintech, es decir, van a chiringuitos que daban chiringuitos financieros o distribuidores o agencias como son Magallanes, como son, una serie de compañías que lo que hacen es, ponen la red, la red pero de fintech a través de web, clientes muy jóvenes con alto poder adquisitivo y han tenido, se puede decir, es público, aquí hay algún

español, es la compañía Cacer, ha empezado a vender en menos de siete meses, ha incrementado en 100 millones de euros la cartera de productos de ahorro gracias a capturar a través de la fintech y posicionarse como fábrica.

También dentro de tu pregunta sí que conocemos compañías que se están aliando con las Googles, sabes, con las Alibaba, etcétera, precisamente para pagarles lo que es la parte de la suscripción como comisionistas, es decir, simplemente, tú me capturas los clientes, me das la información.

Una que salió, empresa fue la Idealista, por ejemplo, la idealista es una compañía que se dedica a temas inmobiliarios, es una página web de compra, venta de inmuebles, que es la más famosa pues vende la información transaccional de la gente que está buscando casa o que va a cambiar de casa para vendérselo a aseguradoras y a bancos para posesionarse.

Entonces, es algo Ángeles que está, ahora mismo en boga pero que ya hay experiencias de compañías que lo que tienen es la red de distribución, lo que llaman, no está en la reunión plenaria pero le llaman los abstractas de distribución que no lo que tienen es la parte de captación para traerlo y trasladarlo a la fábrica, entonces, sí que hay experiencias.

Dra. María de los Ángeles Yáñez Acosta: Súper bien, porque todo eso está basado en estos conocimientos finalmente.

Act. Marcos Fernández: Sin duda, perdona que te interrumpa, Ángeles, sobre todo, por la gente un poco que quizás se quiera subir también a nivel formación en el carro que requiere, sobre todo, lo más complicado son los lenguajes de programación y no son complejos para capturar y manejar toda esa data, o sea, no es tanto el modelo en sí como el traerte lo que son, lo que comentado los Python, los Hajduc, también funciona R que seguro que ya aquí todo el mundo lo conoce, pero hay algunos lenguajes que son más complejos.

Dra. María de los Ángeles Yáñez Acosta: En términos de inversión para crear realmente una base de información que te permita hacer estos modelos, ¿qué tan caro es para una aseguradora organizar esto? Porque, mira, en una asegurado puede haber mil, tú lo has visto, repositorios de información, pero darles esa estructura clara y ordenada, ¿qué tan cara puede resultar?, ¿en cuánto a tiempo crees que se puede

lograr un esfuerzo de ese tiempo, de ese tipo de información? y tres, ¿qué te están pidiendo en Solvencia con lo del gobierno del dato?

Act. Marcos Fernández: Empiezo por la de en medio que es la más fácil, tiempo, pues normalmente Ángeles, esto sí que es importante, no menos de un año. Las compañías que han construido el famoso data lake, que es el data lake no deja de ser como un repositorio con información externa, no menos de un año, con lo cual, traducido a coste, pues imagínate, pues 3, 4 o 5 personas, dedicadas mil 800 horas, o sea, sin duda.

Lo que ocurre que también, Ángeles, los data lake's, tú le puedes ir rellenando de información, entonces, lo único que tienen es que cuando tú haces un data warehouse tienes que modelizar y pensar antes, la estructura, dimensiones y los vínculos mientras que estos data lake's están pensados para ir trayendo información desestructurada y estructurarla, entonces, se pueden hacer como con capas de cebolla, siendo más o menos ambicioso, entonces, lo importante es que pueden ser también los Big Data pueden ser, no quiero decir Little Data pero que pueden ser incrementales, con lo cual yo te diría que un año, pero no hablamos tampoco de cantidades, digamos, se pueden refinanciar.

Es decir, imagínate el tema de fraude que seguro que esta mañana también ha habido una, realmente cuando tú buscas un patrón de fraude y encuentras realmente, refinancias también o en autos también que se está dando.

El último tema es crítico, Ángeles, de hecho a mí me extraña, por ejemplo, que gobierno y calidad de datos, digo, de modelos quizás no, pero gobierno y calidad de datos y protección de datos de clientes que es un reglamento, que ya es una directiva europea que entra en vigor en marzo del año que viene, que México, que estamos muy adelantados en todo lo que tiene que ver los requerimientos de capital, el pilar 1, el pilar 2, etcétera, o sea, es una normativa que hasta ahora sólo se viene aplicando, hasta ahora en seguros, aquellos que van a modelos internos, es decir, si tú vas a...

Dra. María de los Ángeles Yáñez Acosta: Pero protección de datos personales sí tenemos una ley y muy fuerte.

Act. Marcos Fernández: Que es el reconocimiento tácito de si puedes hacer uso o no uso de la información.

Dra. María de los Ángeles Yáñez Acosta: Datos sensibles.

Act. Marcos Fernández: Sí, gobierno de calidad de datos.

Dra. María de los Ángeles Yáñez Acosta: Perdón, Marcos, es una regulación independiente, bueno, está todo implícito pero sí es y tenemos al IFAI, el INAI, perdón, de Acceso a la Información y todo. Entonces, sí, realmente sí, incluso, AMIS ha hecho cosas de certificación en ese tipo de temas.

El tema de la calidad de dato mucho está como que el gobierno corporativo es responsable de esto y la función actuarial deberá decir la calidad pero nada más, no es algo así como más estricto.

Act. Marcos Fernández: No, no, te contesto, la normativa más avanzada en gobierno y calidad de dato es la que se llama (...) que es RDA que es de la banca, que es la de RDA que es de banca.

Incluso en Estados Unidos la figura del chief data officer, es obligatoria, es decir, el chief data officer es una figura que tiene que ser creada en las organizaciones, lo que no dictamina es si organizativamente tiene que depender de tecnología, operaciones, área financiera, etcétera, pero sí te dictamina que tiene que haber un responsable del gobierno y dueño del dato, entendiendo por gobierno, estructura de comités, quién es el responsable de cada dato, quién se responsabiliza cuando un dato va a información externa o información interna, tiene que estar definidas las políticas, los procedimientos, el report.

Es decir, eso en cuanto a gobierno como tú has diferenciado muy bien, que es distinto luego de calidad, que calidad son ya pruebas de validación, etcétera, pero el gobierno de dato, yo os diría la normativa de RDA en seguros Solvencia II la tiene pero la tiene sólo como tal, tiene igual que la LISF, a aquellas compañías que tenía modelo interno, hay un consultation paper que desarrolla el artículo de Solvencia II, que ahí sí que ya tienes un documentito que sí te va explicando todos los elementos del gobierno y calidad de dato.

Pero aquellas compañías que han ido a lo que es el SSR que sería al RCS de aquí, o sea, realmente lo que te piden es validación de datos, pero no te piden un gobierno o un propietario al que apuntar como responsable.

Deciros de la normativa anglosajona y americana tiene incluso incentivos, está incentivada la figura del chief data officer y de ir mejorando en lo que son todos los datos, los missing, los que están mal, etcétera.

Esto hasta el año pasado era obligatorio sólo para la banca con riesgo sistémico, los bancos inclusive, pero de 2016 es obligatoria para toda la banca. Desde luego es uno de los retos.

Dra. María de los Ángeles Yáñez Acosta: Es algo que en AMIS hemos dicho mucho, que hace falta este responsable de los datos. El hecho de que ya hay una figura de chief data officer es un...

Act. Marcos Fernández: Está definida y hay bibliografía y literatura de dónde ubicarla, con qué responsabilidades, etcétera.

Dra. María de los Ángeles Yáñez Acosta: Ya está hecha la chamba.

Act. Marcos Fernández: Ya está hecha.

Dra. María de los Ángeles Yáñez Acosta: ¿Alguna otra pregunta?

Marcos, muchísimas gracias por la buena charla. Miguel te va a hacer entrega de un reconocimiento.

Le damos un aplauso, por favor.

Act. Marcos Fernández: Que va al despacho, aunque con esto de que ya no hay despachos ni nada de eso, yo no sé dónde lo vamos a poner, virtual.

Dra. María de los Ángeles Yáñez Acosta: Y nos has dejado unos folletos con más información.

Con esto damos por concluido las sesiones de la Sala de Riesgos, muchísimas gracias por acompañarnos y si quieren seguir viendo algunas de las pláticas, todavía hay varias.

Hasta luego.

---ooo0ooo---