

Ciudad de México, 16 de mayo de 2017.

Versión Estenográfica de la Sesión Especializada “Casos prácticos de modelaje predictivo para la identificación temprana de fraudes en gastos médicos”, dictada por Marcela Abraham Ortiz y Melissa Ornelas Curiel, en el marco de la 27 Convención de Aseguradores AMIS realizada en el salón Montejo 2, del Centro Banamex, en esta ciudad.

Presentadora: Muy buenas tardes. Les pido, tomen asiento para dar inicio a las pláticas de la tarde con relación a los temas de salud.

Nuestra primera plática va enfocada a mostrar cómo podemos iniciar en el tema de Big Data, Analytics, Minería de Datos, Modelaje Predictivo.

Entonces, hemos venido trabajando en el Comité de Fraudes junto con Willis Towers Watson en tratar de mostrarles algo práctico con la información que ya tenemos y con la operación que ahorita tienen las compañías se seguro en cómo empezar a implementar hoy, empezar a trabajar en estos nuevos temas.

Entonces, tenemos el honor de que nos acompañe Marce Abraham, que es líder de negocio de consultoría de Willis Towers Watson y Melissa Ornelas, experta en este tema y consultora de Willis Towers Watson.

El tema a abordar son casos prácticos de modelaje predictivo para la identificación temprana de fraudes en gastos médicos. Obviamente es gente del sector muy conocida por ustedes, entonces, voy a hacer un súper breve resumen para presentarles y darles la entrada.

Marce es Actuaría titulada de la Universidad Anáhuac del Sur, tiene experiencia en temas de auditoría actuarial, modelaje financiero, reservas, administración de capital, fusiones y adquisiciones.

También en análisis de expansión en mercados y productos en los sectores de seguros generales, vida, pensiones y en portafolios multinacionales. Es responsable de clientes y proyectos estratégicos y de la relación de clientes multinacionales en Latinoamérica, así como

de proyectos internos, tanto de desarrollo, de capital intelectual como de optimización de resultados.

Mel es Actuarial egresado de la Universidad Anáhuac de México, cuenta con un diplomado en operaciones de divisas y estudios de inversión. Es consultora de vida en Willis Towers Watson en la práctica de consultoría actuarial, financiera, de compañías de seguros dentro del segmento de inversiones, riesgos y reaseguro.

Ha trabajado con clientes de México, Panamá, Brasil, Colombia, Estados Unidos, Alemania y Hong Kong y recientemente trabajó en Londres.

Entonces, les pido un aplauso para darles la bienvenida a nuestras expositoras.

Act. Marcela Abraham Ortiz: Bueno, buenas tardes. Muchas gracias, Lolita por la presentación, ya me siento como en casa y después de tantos meses de trabajar con Lolita y con su equipo más. También le quiero dar las gracias a Leslie, a Denise.

Esto que les vamos a presentar es resultado de trabajo de varios meses, hace ya casi un año nos reunimos con Lolita, su equipo, trabajando con el Comité de Fraudes y la petición original era ¿qué hacemos con la información que tenemos de siniestros del sector para detectar fraudes?

De entrada, utilizando un modelo predictivo y de entrada la primera observación es cuando se paga un siniestro que eventualmente por la razón que sea o antes de pagar nos enteramos de que es un fraude, nunca, no se guarda en la base de datos. Ese fue el primer gran problema.

Entonces, tenemos una base de datos de siniestros que se pagaron y los que fueron fraude nunca se guardaron; entonces, supongo que en el público hay muchos no actuarios, muchos no expertos en el tema, perdón si de repente resultamos muy obvias, pero lo que queremos es que quede claro el mensaje.

A pesar de todo, o sea, si lo que queremos es identificar siniestros fraudulentos, desde el inicio tendríamos que haberlos marcado para ver qué características, si es la edad de la persona, el proveedor, el padecimiento, qué propicia las reclamaciones fraudulentas.

De todas formas, había tanta información que lo que platicamos con Lolita es que algo se debe poder hacer.

Entonces, lo que les vamos a presentar el día de hoy es lo que se pudo hacer después de un esfuerzo enorme por parte de Lolita y su equipo tratando de homologar y limpiar la información del sector, más lo que pudimos hacer en las últimas semanas, tratando de buscar la verdad, es decir, estábamos tratando de buscar números morbosos, escandalosos porque nuestro objetivo es que al final de esta plática todas las compañías que están aquí sentadas se lleven como mensaje que vale la pena dedicarle tiempo y ni siquiera es tanto el tiempo que habría que dedicarle para hacer algo muchísimo más sofisticado.

Entonces, bueno, la estructura de la plática; primero vamos a hablar rápidamente de lo que entendemos por fraude, el tema conceptual, después vamos a platicarles cuál fue el criterio que seguimos para analizar la base de datos, el ejemplo y los números que encontramos, es muy importante señalar que este ejemplo fue hecho con números reales de nueve compañías del sector, luego les vamos a dar los detalles.

Después qué podría implicar hacer un análisis a priori, encontrar el fraude antes de pagarlo y finalmente las conclusiones.

Bueno, yo creo que esto no es física cuántica, cómo afecta la siniestralidad o cómo afectan los fraudes o los siniestros que no debieron haber pagado, que no debieron haberse pagado en la operación de Gastos Médicos Mayores.

Bueno, de entrada, para nadie es sorpresa que prácticamente todas las compañías o la gran mayoría de las compañías del sector no están obteniendo resultados aceptables en el ramo de Gastos Médicos Mayores, tanto individual como colectivo.

Además, es un seguro que el día de hoy no deja satisfecho ni al asegurado ni a la aseguradora, ¿qué está pasando? La siniestralidad en los últimos años y ahorita vamos a ver rápidamente una gráfica, ha estado muy por encima de niveles que deberían considerarse como aceptables.

Es un hecho que el incremento de la inflación de los servicios médicos, aunado a una mayor utilización del producto por parte de los asegurados y además a que la población asegurada está envejeciendo, nos está llevando a un incremento en el nivel de siniestralidad que no necesariamente se puede una compañía a la tarifa, o sea, si de por sí las tarifas crecen el crecimiento no podría reflejar todos estos factores y además dentro de esta siniestralidad está el hecho de que varios de los siniestros que se pagan eran siniestros que no procedían, que se pagaron indebidamente para no usar siempre la palabra fraude.

¿Qué han hecho las compañías? Han tomado cualquier cantidad de medidas, pero desgraciadamente una medida muy común es ajustar las coberturas a través de más limitaciones, más exclusiones y estamos acabando en un producto parchado que está tratando de controlar la siniestralidad o las fugas que vemos, pero que al final no lo está haciendo de manera total.

Entonces, cualquier compañía si quiere entrar a atacar este problema, es un problema gigantesco, ahorita vamos a ver, puede haber cualquier cantidad de pagos que no proceden por una cantidad enorme de motivos, entonces, por dónde empiezo, o sea, cómo sé dónde está la mayor parte de los siniestros no procedentes, dónde hay mayor incidencia, hay algún patrón o algún tipo, algún grupo específico de asegurados que me ayude a enfocar mis esfuerzos.

Entonces, lo que se ha estado haciendo en AMIS es tratar de consolidar, de pedir información de siniestros a las compañías del sector, consolidarla, homologarla, que es importantísimo, para que a través de este intercambio de información pueda llegarse a identificar cuáles son las zonas donde tenemos que enfocar nuestros esfuerzos.

Lo que les decía hace un momento, la siniestralidad y vamos, ahora sí que, como un acto de fe, créanme que la siniestralidad debería estar

alrededor del 65 por ciento para decir que estamos operando un ramo de manera rentable y esto es asumiendo un nivel de alrededor del 15, 18 por ciento de gastos de adquisición y otro 8 por ciento o 10 por ciento, todo depende la compañía de gastos de operación, pero vamos a decir que el 65 sería un número razonable.

La siniestralidad de las operaciones de Gastos Médicos Mayores individual ha estado, tal vez cerca, pero siempre por arriba, sólo en el 2016, ya tenemos las cifras, cerramos en el 69 por ciento, pero la siniestralidad de Gastos Médicos colectivo ha estado siempre por arriba de este 65, incluso por arriba del 75, cerramos con el 78 y en ambas gráficas la línea de arriba la representan las compañías más grandes, o sea, aquí no es un tema de volumen o de escala.

Entonces, es un hecho que no estamos dando la rentabilidad que se requiere. Esta gráfica es nada para decir: en México no tenemos la exclusiva, o sea, los fraudes son un problema en el sector asegurador a nivel mundial, es difícilísimo medirlos, pero encontramos aquí de manera anecdótica tres ejemplos. En Malasia el 20 por ciento de las reclamaciones se considera que tienen algún elemento fraudulento; en Estados Unidos se cree que entre el 3 y 10 por ciento de los costos sanitarios totales están relacionados con algún tipo de fraude y por aquí nos encontramos unos 12 millones en Corea del Sur.

El punto es, los fraudes son un elemento muy importante, en términos de pesos y centavos, de hecho, en AMIS se hizo una declaración hace unos dos, tres años, relacionado con el tema de fraudes donde se decía que se estima que, durante el 2014, de enero a septiembre del 2014 el 10 por ciento de la siniestralidad que se pagó correspondía a siniestros fraudulentos donde la mayor parte estaba concentrada en gastos médicos y en autos.

O sea, resumen, el tema de fraude es todo un tema y es algo que podría ayudarnos a hacer la diferencia o acercarnos a hacer la diferencia entre operar un ramo de la manera rentable o no.

Entonces, ahora lo que vamos a platicar es qué hicimos nosotros con la base de datos para ver, aun cuando no habíamos marcado aquellos siniestros que eventualmente pudieron haber sido fraude, se podía hacer algo, se puede hacer algo con esta información y ¿por qué

creemos que se puede hacer algo? Porque si barremos la base de datos y hacemos unos chequeos y aplicamos unas reglas básicas en las que había estado trabajando AMIS durante varios meses, ya podemos encontrar situaciones raras, vamos a decir.

Entonces, en muchos casos a la mejor estamos hablando de siniestros atípicos, que esos existen, tal vez en otros casos estamos hablando de errores de captura, si yo veo una estereotomía género masculino probablemente sea un error de captura, pero de repente la incidencia errores de captura o un siniestro atípico tan atípico o el mismo proveedor siempre pagando siniestros atípicos, nos haría pensar que son siniestros que vale la pena checar dos veces, revisar.

Entonces, aquí vamos a arrancar con la definición de bueno, cuál es la diferencia entre un siniestro que procede la zona gris y un siniestro fraudulento.

Act. Melissa Ornelas Curiel: Entonces, buenas tardes, vamos a empezar con la definición de un siniestro con certeza de pago, ¿qué es un siniestro que definimos como certeza de pago? Es aquel que se apega al contrato de buena fe que se celebró con la aseguradora.

Es decir, tenemos la confianza de pagarlo, no hay nada raro en ese siniestro y lo podemos pagar sin ningún problema. Después de eso vienen todos los atípicos, los atípicos son los raros, los que se salen de la norma y hay dos formas. Es el atípico que se sale de la norma, pero por su naturaleza se sale de la norma, quiere decir que sigue apegado al contrato de buena fe, tiene variables raras fuera de la norma, pero al seguir apegado al contrato se debe de pagar, es atípico, raro, pero se debe de pagar.

Después están todos los que tienen algún elemento extraño, un elemento extraño es un elemento de sobreuso, algún elemento de abuso o algún elemento de lo que estamos llamando fraude. En esta definición, aunque el sobreuso y el abuso ya es un fraude, le estamos llamando fraude a todo aquello que ya es un delito.

Entonces, ¿qué es un sobreuso y un abuso? Es todos los procedimientos que no están justificados, todos los lugares en donde la calidad de los estudios era más alta de la que se necesitaba, en

donde se hicieron cirugías más de las que se necesitaban, si existe una preexistencia que no fue declarada o que el médico cubrió y lo que ya es fraude es una suplantación de identidad, es préstamo de facturas entre proveedores, es cuando las facturas se tratan de meter a las aseguradas duplicadas o triplicadas.

Entonces, vamos a partir de esta definición, todo lo que no tiene certeza de pago se divide en el atípico que vamos a pagar y en el que tiene un elemento de fraude.

Act. Marcela Abraham Ortiz: Al final, pues lo que encontramos es que muchas veces los fraudes podrían estarse pagando por esta falta de control y por contar con ciertos elementos básicos de detección.

Ahora, AMIS ya llevaba un rato trabajando con el Comité de Fraudes y ya tenía cierta información y ciertas reglas bastante útiles, bueno, de hecho, ya las tiene; entonces, parte de lo que estamos buscando en esta plática es darles a conocer lo que ya existe para que también exista un mayor incentivo para el sector en seguir trabajando, invirtiéndole tiempo a estos esfuerzos.

Act. Melissa Ornelas Curiel: Entonces, lo que ya existía es claramente la información del sector que se le proporciona a AMIS, también AMIS y el Comité de Fraudes tienen todo el conocimiento técnico de los productos y las coberturas.

También existe un médico con la que trabaja con AMIS, el cual tiene toda la información clínica, recordemos que estamos hablando de enfermedades y padecimientos y todo esto junto es la sospecha del mercado, ¿qué sospecha el mercado? Por dónde es por donde se están metiendo los fraudes.

Se están aplicando las reglas, como dice Marce, esas reglas se están aplicando a posteriori, más adelante vamos a definir qué es a posteriori, pero básicamente se está mirando a la historia, se está mirando con las reglas y las sospechas que yo tengo, estoy mirando la historia y viendo en dónde están los focos.

Los procesos que hay hasta ahora son básicamente creación de catálogos y homologación de información.

Act. Marcela Abraham Ortiz: Perdón, Mel y el objetivo era o con lo que tengo hoy, lo que puedo hacer es ver que pague algo que no debió haber pagado, lo que quisiéramos es empezar a recopilar información que nos permita evitar hacer ese pago.

Act. Melissa Ornelas Curiel: Entonces, voy a platicarles la definición entre a posteriori y a priori, ¿qué es a posteriori? A posteriori es todo lo que voy a revisar de mi información que ya se pagó, ya no puedo hacer nada, son siniestros históricos que ya se pagaron, yo ya no puedo contener el pago de esos siniestros.

¿Cómo lo hago? Tengo que aplicar las reglas con base en el conocimiento que ya tengo del mercado y tengo que aplicar la minería de datos, la minería de datos es estar analizando todas las bases de datos, todas sus variables y todas sus correlaciones.

¿Qué voy a obtener? Voy a obtener los lugares en donde yo tengo que enfocar mis esfuerzos, voy a obtener los lugares en donde me están saltando más situaciones de sobreuso, de abuso, más situaciones de fraude para poder empezar a poner mis esfuerzos en adelante en esos sectores de la información que me están brincando.

¿Qué es a priori? ¿A dónde queremos llegar? A priori es analizar una reclamación que está abierta o una reclamación que viene entrando, de hecho, se hace ya en tiempo real, hay la posibilidad de hacerlo en tiempo real en cuanto se captura la reclamación.

¿Cómo se hace? Tomo los resultados del análisis que hice histórico, veo en dónde me tengo que enfocar, pero también existe el modelaje predictivo. Vamos a ver más adelante la definición de modelaje predictivo. Lo más importante es que el objetivo del análisis a priori es contener el pago de los siniestros, ahí sí ya puedo empezar a reducir la siniestralidad que no debió de haber sido pagada.

Act. Marcela Abraham Ortiz: Entonces, con la información que actualmente disponemos sólo podemos aplicar, de acuerdo a las definiciones que nos explicó Mel, medidas a posteriori.

De entrada, para empezar a hablar todos con los mismos términos, el mismo lenguaje a la hora de revisar el ejemplo que llevamos a cabo, el ejercicio que llevamos a cabo, tenemos que reconocer que nosotros no podemos detectar o no pudimos detectar fraudes, el siniestro ya se pagó y tendría que regresar la compañía a revisar el siniestro que yo le diga: este siniestro no me suena y confirmar que, efectivamente, fue un siniestro que no debió haberse pagado, que se metió de manera dolosa y por lo tanto es fraude.

Entonces, decidimos dejemos de usar la palabra fraude y vamos a usar el término alertas de fraude; entonces, lo que nosotros pudimos hacer en este ejercicio es identificar alertas de fraude, ¿cómo? Definimos una serie de reglas, que son las que nos va a explicar Mel en un momento para que dependiendo de cuántas de estas reglas se cumplieran o no pudiéramos calificar un siniestro como un siniestro que levanta una alerta de fraude o no.

Estas reglas ¿cómo las definimos?

Act. Melissa Ornelas Curiel: Primero estas reglas tienen que estar basadas en, obviamente el conocimiento del mercado, entonces, el conocimiento del mercado, el que sepamos los tipos de productos, las coberturas, nos va a dar la primera visión de qué tipo de reglas tenemos que hacer. Un ejemplo y que se conoce en el mercado es, cuando la aseguradora paga vía reembolso, hay mucha más posibilidad de estar pagando algo que no debió de haber sido pagado, que cuando se hace el pago directo.

También nos tenemos que fijar en los valores o en las combinaciones de valores de las variables que tenemos en la base de datos; es decir, las variables que tenemos son género, edad, tenemos que empezar a revisar todas esas combinaciones con lo que ya sabemos, por ejemplo, tenemos el procedimiento que se aplicó a algún padecimiento, ese monto se sale de mi monto promedio, se sale de mi monto promedio más, menos su varianza.

Entonces, tengo que empezar a revisar, por ejemplo, cuántas son las facturas promedio que hay para un padecimiento dado, si hay 20 facturas, 50 facturas, tengo que empezar a revisar todos esos valores de las variables que ya conozco.

Por último, el conocimiento clínico, ese nos lo da un médico y hay ciertos padecimientos que son específicos de cierto género, tenemos que tener cuidado, otros que aplican solamente para algunos rangos de edad.

Entonces, con eso vamos a definir las reglas que son parte de una alerta.

Act. Marcela Abraham Ortiz: Ahora, les decía que trabajamos, o sea, cuando nos dimos cuenta que no estaban los siniestros que eventualmente fueron fraudulentos, pero teníamos tal cantidad de información que algo se tendría que poder hacer, decidimos que, bueno, vamos a tratar de hacer algo, pero también estamos conscientes de que yo no puedo regresar con una compañía de seguros y decirle: mira, encontré mil 500 siniestros en tu base de datos que no me suenan, porque va a haber errores de captura, porque puede haber diferencias que realmente no, suenan a desviaciones, pero no son relevantes.

¿Qué hay que hacer? Hay que darle prioridad; entonces, yo puedo regresar a una compañía con un número limitado de siniestros para tratar de comprobar si mis alertas de fraude, si mis reglas están funcionando o no.

Entonces, lo que hay que hacer es darle valor a esas alertas, ¿cuáles son las alertas que me preocupan y cuáles son las alertas con la que puedo vivir y puedo seguir operando? En este caso para este ejercicio nosotros usamos un criterio tipo semáforo que nos va a explicar en un momento Mel, pero también podemos haber utilizado otro criterio que se usa en otros países y en otros mercados, que es un criterio tipo scoring o puntuación.

Entonces, yo puedo definir una serie de reglas y decir: okay, si se cumple esta regla vale 10 puntos, esta regla vale 30, la sumo cualquier siniestro que llegue a los 500 puntos es un siniestro que ha incumplido o cumplido, si lo queremos ver positivo o negativo, tantas reglas que merece la pena ser revisados dos veces. O sea, en nuestro caso nosotros usamos un criterio tipo semáforo.

Act. Melissa Ornelas Curiel: Y lo usamos tipo semáforo porque es muy fácil reconocer los colores, entonces, el verde es aquel siniestro que ya habíamos definido como con certeza de pago, son los que no me preocupan, después definimos el color ámbar, el color ámbar es aquéllos que ya tienen un elemento de fraude, pero o todavía no es una alta probabilidad o no me preocupan tanto.

Entonces, en los colores ámbar lo que estamos proponiendo es, hay que darle seguimiento, hay que tener cuidado y darle seguimiento, que no se vayan a convertir en rojo, ¿por qué? Porque en rojo están todos aquéllos que me preocupan, que levantaron más de equis número de alertas o de reglas, que levantaron una regla que para mí es muy importante, como el monto promedio.

Entonces, todo lo que está en rojo, para la parte a posteriori, que es la parte histórica, es en donde voy a enfocar, es en donde a poner el ojo y donde voy a empezar a revisar y para la parte a priori es en donde puedo empezar a contener los pagos en lo que investigo si fue un fraude o no, tenemos que tener mucho cuidado con los errores de captura, los errores de captura pueden generarnos falsos positivos o falsos negativos en nuestros semáforos, entonces, es muy importante que tengamos mucho cuidado, tanto con la calidad como con la cantidad de información que estamos guardando.

Algo que es muy importante decir es que esta clasificación es nuestra propuesta, pero también es personalizable, es decir, depende del apetito de riesgo de la compañía, depende del tamaño de la compañía, depende del tipo de productos que esté vendiendo la compañía, a quién pone, una misma regla la puede poner en ámbar y hay una compañía que la puede poner en rojo porque le preocupa muchísimo más por su tamaño.

Entonces, siempre podemos personalizar la escala, por ahí vamos a ver colores naranjas también, la escala en donde vamos a poner lo que nos preocupa.

Act. Marcela Abraham Ortiz: Entonces, lo que hicimos fue cuando recibimos la base de datos definimos una serie de reglas y decidimos qué criterios iban a hacer que se levantara una alerta verde, por ejemplo, ámbar o roja; entonces, en este caso...

Act. Melissa Ornelas Curiel: Primero vamos a definir qué es un registro, la siniestralidad viene, obviamente, por siniestro, pero recordemos que un siniestro puede tener varias reclamaciones.

Entonces, el registro puede estar a nivel reclamación, pero a su vez también puede tener cierto número de facturas, en este caso cuando yo diga registro estoy hablando de registro a nivel factura.

Entonces, aplicamos, recibimos un registro de la base de datos, es una factura y yo conozco el CPT, el CPT es el Código Internacional para los Procedimientos Aplicados; entonces, yo conozco el monto para ese CPT y también tengo información interna del costo promedio que tiene ese CPT en la historia de mi información.

Mi pregunta es, ¿el costo que se le aplica a ese CPT está fuera de mi monto promedio? En este caso es falso, quiere decir que está adentro, no me preocupa, pasa como verde.

La siguiente regla, voy a revisar el CIE, el CIE es el Código Internacional para Padecimientos, pues voy a revisar si el CIE no concuerda con el procedimiento que se aplicó, es decir, el padecimiento y el procedimiento no concuerdan. En este caso es falso, quiere decir que no hay nada raro, es verde.

También voy a revisar el género contra el padecimiento, entonces si el género está fuera de los géneros especiales para ciertos padecimientos, en este caso es falso también, lo propongo verde.

Por último, si la edad está fuera del rango de edades de cierto padecimiento.

Es verde, es falso y entonces decidimos en nuestra propuesta que si nuestras cuatro reglas son falsas entonces mi alerta es verde, no tengo nada qué preocuparme y voy a pagar el siniestro.

¿Qué pasa si hay dos que me dan positivo? Es decir, el CIE y el CPT no concuerdan, entonces, mi regla es positiva y qué pasa si la edad está fuera del rango de edades definidas para cierto padecimiento. Yo

decido que con dos mi alerta ya va a ser color ámbar, ya me empieza a preocupar, ya le tengo que dar seguimiento.

Act. Marcela Abraham Ortiz: Ahora aquí nada más enfatizando, la intención de esta plática es hacer una muestra de cómo aplicaría o cómo funcionaría un ejemplo aplicando reglas y alertas, pero esto no quiere decir que esas sean las reglas importantes, entonces, por eso lo que decía Melissa de, se puede personalizar, cada quien decide cuáles son las que cambian de verde a ámbar y de ámbar a rojo.

Act. Melissa Ornelas Curiel: Y en rojo entonces proponemos que, con más de dos positivas, es decir, el monto de la factura para ese procedimiento se sale de los parámetros que yo considero como mi costo promedio y entonces pongo mi alerta roja.

Entonces, éstas son las que ya me preocupan mucho, las que ya tengo que investigar, las que ya en el futuro voy a poner mucha atención.

Act. Marcela Abraham Ortiz: Ahora sí vámonos al resultado del ejercicio.

Entonces, recibimos una base de datos de siniestros del sector, aplicamos estos criterios y las reglas que vamos a describir a continuación y les vamos a mostrar lo que encontramos.

Primero, cuál fue la información con la que trabajamos.

Act. Melissa Ornelas Curiel: Recibimos una base de datos, es una muestra que corresponde a nueve compañías del sector, no conocemos las compañías, no conocemos más detalles, pero sabemos que hay nueve compañías, por lo que es una buena base de datos, la siniestralidad que estamos usando es la reportada a 2016, ¿por qué a 2016? Porque es la base de datos que tiene mejor calidad y mejor calidad me refiero a que AMIS y el Comité tienen un proceso de calidad, entonces, esta base de datos ya pasó por el proceso de calidad, de tal forma que podemos confiar en la base de datos.

Son 1.3 millones de registros a nivel factura y son 9 mil 371 millones de pesos pagados en siniestralidad en sólo un año. La base de datos

tiene 24 campos a lo ancho, éstos son los campos que me van a definir el identificador del siniestro, el género, toda la parte que me da la estadística descriptiva de cada registro.

Es muy importante decir que la información es limitada porque es solamente una muestra, es solamente un año y solamente estamos aplicando un subconjunto de reglas y de alertas, como bien decía Marce, no son las únicas, de hecho, son unas que son simples de explicar y que tienen resultados bastante fuertes, pero esto nos va a permitir estimar el impacto económico o financiero de todo lo que no debimos de haber pagado en el 2016, pero también nos va a permitir establecer ciertos controles para toda la siniestralidad que viene adelante y también nos va a permitir implementar procesos para administrar mejor mis siniestros.

Act. Marcela Abraham Ortiz: Exacto, o sea, escogimos la muestra que podíamos manipular en unas semanas y cuyos resultados les podríamos presentar en 45 minutos, si no hubiéramos necesitado horas.

¿Cuáles fueron las reglas que definimos? Estas son las seis reglas que nos van a levantar alertas verdes, ámbar o roja.

Act. Melissa Ornelas Curiel: La primera regla es identificamos todos aquellos registros que tenían relación con una lista de proveedores que se conocen al sector que tienen ciertas actividades fraudulentas, entonces esas actividades las dividimos, los amarillos son las actividades que tenemos sospecha, pero es empírica, todavía no tenemos ningún sustento.

Los naranjas son aquellos que han sido reincidentes, proveedores que ya se sabe que tienen cierta reincidencia, aunque todavía no se ha podido comprobar un fraude y los rojos son aquéllos a los que ya se les ha comprobado que están teniendo actividades fraudulentas.

Entonces, vamos a cruzar la base de datos y ver cómo se separa, en este caso. La regla número 2 es precisamente la que explicábamos hace rato, identificar los padecimientos reportados que no concuerdan con los procedimientos.

La número 3 es identificar todos los siniestros que tengan por reclamación más de ene facturas. Más adelante vamos a ver qué ene, no sabíamos qué nos íbamos a encontrar, vamos a ver cuál es ese parámetro.

La regla 4 es todos los gastos reportados que estén fuera del rango del costo promedio, más menos dos veces la desviación ¿por qué más menos? Porque no me preocupan sólo los grandes, también me preocupan los montos chiquitos.

La tercera que la edad del paciente, voy a identificar todos aquellos donde la edad del paciente esté fuera del rango de edades definido para el padecimiento.

Por último, el género, que el género del paciente sí está fuera de aquellos padecimientos que son propios de algún género.

La regla 4, 5 y 6, como podrán ver, pueden por sí mismas fácilmente errores de captura, puede ser, como decía Marce, una estereotomía que se le puso género masculino, bueno, podría ser un error de captura; entonces, las vamos a revisar juntas porque ya dos o tres errores en un mismo registro nos van a preocupar.

Vuelvo a recalcar, son reglas genéricas, son reglas que se le aplican a cualquier registro, pero no son las únicas, de hecho, podríamos hacer reglas específicas para ciertos padecimientos, reglas específicas para ciertos procedimientos, para ciertos tipos de coberturas, para ciertos tipos de pago y nuestro objetivo principal es cuál es el impacto económico de estas seis reglas.

A su vez con la herramienta de minería de datos lo que hicimos fue poder responder a estas preguntas por cada segmento que tenga la base de datos; es decir, yo puedo responder las preguntas de cuál es el impacto económico por tipo de proveedor, por entidad, por cualquier variable que tenga mi base de datos o por cualquier conjunto de variables que tenga mi base de datos.

Act. Marcela Abraham Ortiz: Entonces, ya que teníamos esta información, las reglas definidas, aplicamos las reglas al 1.3 millones de registros, vimos cuáles se volvían positivas, hicimos todos los

cortes y agrupaciones que nos permitieran ver si había algún grupo, área, patrón identificable o algo que pudiera traducirse en una medida inmediata para que una compañía de seguros sin haber hecho un análisis muy largo, muy sofisticado, muy científico, pudiera implementar una medida de control mañana.

Entonces, nada más antes de entrar a los resultados, algunas láminas que describen el perfil de la base de siniestros que estábamos analizando.

Act. Melissa Ornelas Curiel: Entonces, era importante tener todos los tipos de seguros, entonces, bueno, del ramo, entonces, tenemos el casi 24 por ciento para gastos médicos colectivo, el casi 35 para individual y tenemos una buena parte en AP individual.

También revisamos cuál era el tipo de pago, el perfil de tipo de pago de la siniestralidad, el 75 por ciento es pago directo y el 25 es reembolso, les recuerdo que ya tenemos la sospecha, tengamos en mente, ya tenemos la sospecha de que en reembolso hay más probabilidades de que haya un fraude.

Por tipo de proveedor es un poco obvio, el hospital se lleva el mayor porcentaje, después los honorarios médicos y después está muy repartido entre farmacia, terapias, enfermera ambulatorios.

Por entidad también claramente el Distrito Federal se lleva un buen pedazo, lo que aquí me gustaría que mantengan en mente es por qué hay 14 por ciento en desconocido. Es un porcentaje bastante algo y vamos a ver más adelante que está ligado al reembolso, cuando el pago del siniestro es por reembolso la práctica del sector es no poner la entidad, ya no ponemos la entidad en el registro.

Entonces, tenemos que tener mucho cuidado con eso porque es un monto también bastante alto.

Act. Marcela Abraham Ortiz: Entonces qué pasa cuando revisamos cuántas de estas facturas tienen alguna relación con uno de esos proveedores que de una u otra forma ya se han encontrado relacionados con alguna actividad fraudulenta, se haya o no se haya comprobado.

Act. Melissa Ornelas Curiel: Y es sorprendente que es casi el 16 por ciento de la siniestralidad; entonces, está separado entre los amarillos que dijimos que era una sospecha empírica, por ahora no nos van a preocupar.

Los naranjas ya tienen cierto sustento y es un porcentaje bastante alto, casi el 5 por ciento, pero lo que realmente preocupa es que el 2.5 por ciento de la siniestralidad cayó con proveedores que ya se sabe que tienen actividades fraudulentas.

Entonces, el impacto que vamos a ver, el impacto potencial que hubo en esa siniestralidad es de 230 millones de pesos y me estoy enfocando solamente en los rojos.

Lo que estamos proponiendo es cuál es el control, qué control podemos tener. Primero tenemos que mantener actualizada esa base de datos de proveedores entre el sector y AMIS.

Segundo, también tenemos que hacer obligatoria la captura de todos los detalles que nos permitan mantener y utilizar esas listas. Si yo en mi registro que va a la base de datos no capturo el RFC completo del proveedor, no voy a poder revisar si está en alguna de estas listas.

El proceso que proponemos para a priori es que cuando llega una reclamación yo lo cruzo con esta base de datos, veo si corresponde a alguno de estos proveedores y contengo el pago en lo que investigo, si resulta que sí fue fraude entonces voy a detener el pago de ese siniestro.

Act. Marcela Abraham Ortiz: Esta es una medida que tampoco suena imposible implementar mañana.

Otra cosa interesante que nos encontramos es, cómo compara la distribución de los siniestros que caían con proveedores fraudulentos contra la distribución total de los siniestros.

Act. Melissa Ornelas Curiel: Y lo importante aquí es mirar gastos médicos colectivo, se va del 24 por ciento al 50 por ciento, quiere decir que en gastos médicos colectivos estamos usando muchos o los

asegurados están usando muchos de estos proveedores, se hace muy grande la proporción.

En individual también hay un incremento, hay un incremento bastante grande; entonces, tendríamos que enfocar nuestros esfuerzos en revisar este tipo de seguro y es claro que AP no tiene este problema.

Act. Marcela Abraham Ortiz: Después aquí sí fue oh, desilusión y el problema aquí es que no se está dedicando suficiente tiempo para mantener bien cargado y actualizados los catálogos que cruzan enfermedades con tratamientos.

Aquí ya AMIS hizo un esfuerzo considerable en las últimas semanas, cuando cruzamos la base y nos encontramos que el 41 por ciento de las facturas no podíamos ligar la enfermedad con el procedimiento nos dimos cuenta que éste era un error de captura, entonces ya era imposible.

De hecho, ya después hubo un trabajo adicional por parte de AMIS y esa cifra se mejoró.

Act. Melissa Ornelas Curiel: Entonces, tratamos esta regla como no concluyente, pero sabemos que ya con esta misma información se reduce a menos del 7 por ciento, si nosotros ya sabemos que sólo el 7 por ciento no cae dentro de esta lista, ésta debería de ser una regla muy poderosa porque es una regla clínica en donde si el procedimiento no cuadra con el padecimiento no tengo por qué pagarlo. Entonces, ha sido AMIR la que ha homologado todos estos catálogos.

Act. Marcela Abraham Ortiz: Y en la siguiente regla, la tercera, que era donde menos decía, no sabíamos bien a bien qué íbamos a encontrar, aunque existía la sospecha del robo hormiga lo que hicimos es ver cuántas facturas había por reclamación y después agruparlas por el número de facturas.

Act. Melissa Ornelas Curiel: Y vemos que, de 10 a 30 facturas por reclamación, recordemos que en la base de datos puede haber un siniestro con 10, 15 o 20 reclamaciones, esto es sólo de una reclamación, está al 3 por ciento.

Después tenemos registros que tienen, perdón, reclamaciones que tienen entre 30 y 50 registros para esa reclamación y es el 1.6 por ciento del total, pero lo que realmente nos preocupó es que hay registros que tienen más de 50 facturas por reclamación en un año, de hecho, encontramos registros que tienen más de 140 facturas por reclamación.

Entonces, ¿cuál es el impacto potencial de la siniestralidad? Sólo de los rojos, sólo de los que tienen más de 50 facturas son 175 millones de pesos, ¿cuál es el control? Tengo que tener todos los detalles de mis facturas, folios, cuándo se emitió la factura, de cuándo es el procedimiento, esa factura cuándo fue pagada, para poder empezar a revisar por qué hay tantas facturas por reclamación.

Y el proceso es si yo empiezo a capturar todo eso voy a empezar ver las reclamaciones que están abiertas, que empiezan a tener cierto volumen de facturas, lo que tengo es que contener el pago hasta revisar si realmente es un robo hormiga o por ejemplo, se trata de una enfermedad crónica o se trata de terapias que pueden tener un gran número de facturas.

Act. Marcela Abraham Ortiz: Y aquí el perfil de esta base de datos donde tuvo un cambio muy relevante fue en el tipo de pago.

Act. Melissa Ornelas Curiel: En el tipo de pago vemos que el 100 por ciento cayó en reembolso, entonces, es muy importante ver que de todas éstas que tienen una gran cantidad de facturas todo fue pagado vía reembolso, entonces, confirmo un poquito nuestra sospecha del inicio, ¿qué está pasando en reembolso? No estamos revisando bien nuestra información y no estamos teniendo proceso ahí que nos cuiden del robo hormiga.

Act. Marcela Abraham Ortiz: Después cuando revisamos de manera conjunta, porque como habíamos comentado, esto se presta, estas tres últimas reglas se prestan a errores de captura, ¿cuáles eran los registros que salían del costo promedio para el procedimiento? Que además tenían, que salían del rango de edades para el que ese procedimiento o esa enfermedad es normal y que salían del género normal para ese tipo de padecimiento.

Entonces, si analizáramos esas tres reglas de manera conjunta encontramos lo siguiente.

Act. Melissa Ornelas Curiel: Entonces, es alrededor del 18 por ciento y el 7 por ciento son las que me levantan una regla, entonces, fácilmente yo podría decir: es un error de captura.

Bueno, si es un error de captura realmente no debería de desecharlo luego luego, tengo que tener cuidado con mi base de datos porque a la mejor por ahí, ya saben que como es un error de captura estoy dejando pasar ese siniestro y realmente no es un error de captura es una fuga.

Los que tienen dos reglas es el 6.3 por ciento, ya me empiezo a preocupar, pero realmente el 4.6 por ciento de la siniestralidad levantó tres de estas reglas, realmente tenemos tres errores de captura en cada registro.

Entonces, el impacto potencial es bastante grande, 434 millones de pesos, ¿qué tenemos que hacer en el control? Mantener la calidad de nuestra información, tener cuidado con los errores de captura, que no haya falsos positivos o falsos negativos en nuestra información y ¿cuál es el proceso? Es que voy a contener el pago de una reclamación que me levante más de dos o las tres alertas hasta que no identifique yo si realmente es un error de captura y si es fraude entonces puedo ahí detener y bajar el costo de la siniestralidad.

Act. Marcela Abraham Ortiz: Y viendo cuál era el tipo o la clasificación que cambiaba más para este tipo de alertas respecto de la población o de la base total de siniestros encontramos que aquí era la entidad federativa.

Act. Melissa Ornelas Curiel: Y bueno, la gráfica, los colores están extraños, pero en realidad veo que el Distrito Federal se va del 32 por ciento al 50 por ciento, es algo un poco obvio, es la concentración de la población, pero me preocupa el desconocido, les decía yo que había 14 por ciento, se basta el 34 por ciento.

Entonces, estamos viendo que todo lo que viene desconocido resulta que también levantó tres alertas, tenemos que tener mucho cuidado con eso porque hay mucha fuga ahí y tenemos que cuidar la captura de la entidad también.

Act. Marcela Abraham Ortiz: Finalmente, hicimos el ejercicio enfocándonos dentro de estas reglas en aquellas que eran las más preocupantes, o sea, qué pasa si sumamos todos los registros o facturas relacionadas con proveedores que tuvieron ya comprobado algún, que cometieron algún fraude de manera comprobada, más todas aquellas reclamaciones que tienen más de 50 facturas, más todas aquellas facturas con más de tres errores de captura.

Act. Melissa Ornelas Curiel: Vemos que si sumamos todos son 879 millones de pesos, es el 9.2 o 4 por ciento de la siniestralidad total que potencialmente no debimos de haber pagado en el 2016, todo eso sería dinero que jugaría a nuestro favor en utilidad.

Act. Marcela Abraham Ortiz: Y todo esto es sin haber aplicado ningún tipo de modelaje predictivo, simplemente aplicando filtros y reglas a la información que estamos procesando.

Act. Melissa Ornelas Curiel: Y bueno ya lo repetimos varias veces, no son las únicas reglas, hay reglas mucho más robustas, por ejemplo, alguna es guardar la historia del asegurado, el padecimiento que ya tuvo, de tal forma que, si ya existe una estereotomía o extirpación del útero en una mujer, pues en el futuro no podría tener ninguna relación relacionada con el útero, entonces, podríamos contener por ahí también ciertos fraudes.

La que nos preocupa también es la reclamación hormiga, a la mejor estamos dejando pasar montos pequeños, pero no nos estamos fijando en la frecuencia de las reclamaciones.

Act. Marcela Abraham Ortiz: Ahora sí, ¿qué implicaría o cuál sería la diferencia entre haber hecho el ejercicio que nosotros llevamos a cabo y hacer un análisis a priori.

Act. Melissa Ornelas Curiel: Nos vamos a enfocar un poquito en platicar muy rápido qué es el modelaje predictivo, para poder hacerme

del modelaje predictivo lo primero que tenemos que tener, la primera situación que tiene que tener las bases de datos es tener una banderita, un indicador de todo aquello que fue fraude o todo aquello que no fue fraude.

Entonces, es básico, sin ese identificador no podemos hacer modelaje predictivo. Ventajas puedo tener el número de variables que yo quiera en mi base de datos no sólo 24, como trabajamos ahora, puedo tener 100, 150, las que yo quiera, ¿por qué? Porque van a ir a un modelo, ese modelo lo que va a hacer con técnicas matemáticas y estadísticas bastante robustas va a revisar en palabras simples la correlación entre todas las variables, todas con todas, va a revisar en qué se parecen todas esas variables, tanto de las que fueron marcadas como fraude como de las que no.

De tal forma que cuando yo tengo mi modelo llega mi reclamación actual en tiempo real, si así lo deseo, pasa por el modelo, el modelo revisa todas las características de eso junto con las características de la historia que ya revisó y me va a decir en términos fáciles fraude o no fraude.

¿Cuáles son las ventajas? Uno, son modelos matemáticos, son modelos estadísticos. Y dos, mientras más información tengamos vamos a poder calibrar este modelo y vamos a poder hacer que su capacidad de predicción sea mucho más exacta, ¿por qué? Porque yo ya conozco la historia ya sé lo que fue fraude o no, si yo aplico el modelo a la historia, pues debería de replicarme lo que ya fue fraude o no, con eso puedo empezar a calibrar mi modelo hasta que su capacidad de predicción sea de un nivel de confianza que nosotros queramos.

Las grandes ventajas, bueno, todo este análisis es a priori, lo puedo revisar en cuanto es recibido, los modelos pueden analizar y revisar grandes cantidades de variables, pero también de registros y dependiendo de las correlaciones, el modelo él solito va a decidir sus reglas y va a decidir todo lo que una reclamación me va a decir si es fraude o no, no las tengo que mencionar yo, como hicimos en el ejercicio.

Act. Marcela Abraham Ortiz: Entonces, voy a irme rápidamente a las conclusiones para evitar alargarnos y por si alguien quisiera hacer alguna pregunta, y es simplemente reforzar los mensajes que hemos estado dando a lo largo de esta presentación, yo creo que los esfuerzos que está haciendo AMIS y el Comité de Medidas Preventivas y Correctivas contra Abusos y Fraudes de Gastos Médicos Mayores son muy acertados.

El trabajo analítico que hemos hecho hasta el momento indica que sí hay forma de mejorar la siniestralidad sin hacer grandes esfuerzos y aplicando reglas bastante básicas que un elemento clave para que esto funcione es compartir información y conocimiento del mercado y que después cada compañía decida cómo medir, perdón, cómo aplicar estas reglas en función a su perfil de cartera, a su tamaño, a su capacidad.

Existe lo que se llaman *quick wins* en inglés, o sea, sí es posible aplicando ciertas medidas muy concretas que requieren poco esfuerzo, existe manera de ver resultados, existe la posibilidad de ver resultados de manera inmediata.

Finalmente, creo que el mensaje es que toda inversión en análisis, minería, todo lo que se refiere a información, modelaje predictivo, es una inversión que nos sirve no nada más para detectar fraudes en gastos médicos, para conocer mejor al cliente, no por nada es el tema de modo, entonces, creo que las compañías, el sector en general sigan trabajando de manera conjunta con AMIS en este esfuerzo.

Muchas gracias por su atención. No sé si hay algún comentario o pregunta.

Presentadora: ¿Tienen alguna pregunta?

Pregunta: Hola, buenas tardes. Bien, nada más quería preguntarte cuál fue la herramienta de minería que usaron para hacer todo esto.

Act. Marcela Abraham Ortiz: Nosotros usamos un software nuestro, no quisimos hacer el comercial descargado, pero se llama Emblem, pero es una y la ventaja que teníamos y que fue el valor que le pudimos agregar a AMIS que no teníamos que hacer filtros usando

Excel y otro tipo de software, entonces, pudimos hacer muchos filtros de manera muy rápida, pero hay varias.

Presentadora: ¿Alguna otra pregunta?

Pregunta: Cuando hablamos de fraude esto implica que es un acto criminal, hay una discusión con AMIS que vamos a hacer cuando encontramos fraude, acto criminal, vamos a pasar las autoridades o qué porque esto en otros países lo hacen y hay un buen trabajo junto con las autoridades y los deudores, ¿estamos pensando esto?

Presentadora: Bueno, la compañía cuando detecta un fraude, que como bien dicen, es que vaya en contra de la ley, la gran mayoría de las compañías al tener todos los elementos y los argumentos, pues presentan la demanda y hay casos que sí se han llevado a los tribunales, etcétera.

El gran tema aquí a veces es que no se tienen los suficientes elementos como para que llevarlos a fraude. Entonces, lo que estamos viendo es que gran parte de los casos que tenemos son de abusos, entonces, también esa parte en donde vemos que ciertos proveedores, médicos, etcétera, cometen abusos, también a veces los llegamos a calificar para que en un momento dado ese tipo de proveedores los podamos revisar en el momento que haya un número o un número determinado de alertas con la finalidad que la aseguradora puede tener mayor elementos para ver si es fraude o no es fraude o simplemente de alguna manera pone restricciones en su operación o en las pólizas.

¿Alguna otra pregunta?

Quiero agradecer a Marce y a Mel porque tratamos de preparar esta plática con la intención de darles un ejemplo práctico y sencillo de todo esto y cómo lo empezamos a ver, espero que haya sido de utilidad y espero que sea también de alguna manera que se empiece a implementar y que tomemos conciencia de la importancia de la calidad de la información y que de alguna manera el que vayamos mejorando esa calidad nos puede repercutir en beneficios desde el producto, la siniestralidad y los resultados.

Muchas gracias.

---ooo0ooo---