

**POLARIDAD EN REDES SOCIALES Y PRINCIPALES
INDICADORES FINANCIEROS DEL SECTOR ASEGURADOR**

Andrés Camilo Ramírez Gaita

**Escuela Colombiana de Ingeniería Julio Garavito
Decanatura de ingeniería de sistemas
Maestría Gestión de Información
BOGOTÁ D.C., Colombia
2017**

**POLARIDAD EN REDES SOCIALES Y PRINCIPALES
INDICADORES FINANCIEROS DEL SECTOR ASEGURADOR**

Andrés Camilo Ramírez Gaita

Trabajo de investigación para optar al título de
Magíster en Gestión de Información

Director
Mauricio Alberto Martínez Franki
Ingeniero de Sistemas

**Escuela Colombiana de Ingeniería Julio Garavito
Decanatura de ingeniería de sistemas
Maestría Gestión de Información
BOGOTÁ D.C., Colombia
2017**

DEDICATORIA

“Este trabajo es dedicado a mi madre que hoy día ve los frutos de tantos años de esfuerzo, apoyo y dedicación”

AGRADECIMIENTOS

Agradezco especialmente:

A Ingrid Piragauta mi novia, por su aliento, aporte y apoyo incondicional brindado a lo largo de todo el proyecto de grado.

A mi tutor, el profesor Mauricio Martinez Franki quien siempre estuvo dispuestos a apoyarme y a orientarme en la consecución del trabajo a través de sus conocimientos y experiencia.

A mi hermano, Jeison Ramírez y a Nury Piragauta por el tiempo dedicado a la construcción del modelo.

A la escuela colombiana de ingeniería Julio Garavito por darme la oportunidad y facilidad de estudiar mi maestría.

A todos mis profesores, porque de cada uno aplique algo en el proyecto y por supuesto en mi vida profesional.

Resumen

El presente documento tiene como objetivo dotar a los líderes de compañías de seguros con una herramienta de análisis de gestión comparativa, la cual les permita mediante un tablero de control analizar el comportamiento de indicadores de gestión financiera y la opinión del mercado sobre sus compañías. El documento se basa en la información de 25 aseguradoras que participan activamente en Facebook o en Twitter y se encuentran afiliadas a Fasesolda (Federación de Aseguradores Colombianos). También se muestra la utilidad de usar herramientas libres para convertir datos en información y está en conocimiento.

Abstract

The purpose of this document is to equip leaders of insurance companies with a comparative management analysis tool, which allows it through a dashboard to analyze the behavior of financial management indicators and market opinion about companies. The document is based on the information of 25 companies that participate actively in Facebook and Twitter which are affiliated to Fasesolda (Federation of Colombian Insurers). It also shows the utility of using free tools to convert data into information and is in the know.

CONTENIDO

| | | |
|--------|---|----|
| 1 | PLANTEAMIENTO DEL PROBLEMA | 13 |
| 1.1. | DESCRIPCIÓN DEL PROBLEMA..... | 13 |
| 1.2. | ALCANCES Y DELIMITACIONES..... | 14 |
| 2 | JUSTIFICACIÓN | 16 |
| 3 | OBJETIVOS GENERAL Y ESPECÍFICOS | 18 |
| 3.1. | OBJETIVO GENERAL..... | 18 |
| 3.2. | OBJETIVOS ESPECÍFICOS | 18 |
| 4 | MARCO TEÓRICO..... | 19 |
| 4.1. | SECTOR SEGUROS EN COLOMBIA. | 19 |
| 4.1.1. | Seguros por asegurador | 19 |
| 4.1.2. | Seguros por ramo..... | 21 |
| 4.1.3. | Percepción del cliente en la industria aseguradora..... | 22 |
| 4.1.4. | Marco Legal del sector asegurador..... | 23 |
| 4.2. | REDES SOCIALES..... | 24 |
| 4.2.1. | Twitter | 27 |
| 4.2.2. | Facebook..... | 28 |
| 4.3. | BALANCED SCORECARD..... | 29 |
| 4.3.1. | Beneficios del Balance ScoreCard | 29 |
| 4.3.2. | Perspectivas | 30 |
| 4.3.3. | Indicadores claves de rendimientos (KPI)..... | 32 |
| 4.3.4. | Casos de éxito del Balance ScoreCard | 35 |
| 4.4. | TEORIA DE GRAFOS..... | 36 |
| 4.4.1. | Definición | 36 |
| 4.4.2. | Propiedades de los grafos | 38 |
| 4.5. | ANALÍTICA..... | 40 |
| 4.5.1. | Introducción a la Inteligencia de negocios. | 40 |
| 4.5.2. | Introducción a la Minería De Datos | 45 |
| 4.6. | POLARIDAD | 47 |

| | | |
|--------|--|----|
| 4.6.1. | Analizando opiniones de las redes sociales en sector turístico | 48 |
| 4.6.2. | Modelo léxico-sintáctico para la polaridad en redes sociales..... | 50 |
| 5. | DISEÑO METODOLÓGICO | 52 |
| 5.1. | ENTENDIMIENTO DEL NEGOCIO..... | 53 |
| 5.1.1. | Métricas e indicadores propuestos..... | 53 |
| 5.1.2. | Tableros de control propuestos..... | 60 |
| 5.2. | ENTENDIMIENTO DE LOS DATOS | 62 |
| 5.2.1. | Aseguradoras postuladas y sus redes sociales. | 62 |
| 5.2.2. | Fuentes de información y campos. | 64 |
| 4.1.2. | Estado de la información y restricciones | 66 |
| 4.2. | PREPARACIÓN DE LOS DATOS | 69 |
| 4.2.1. | Normalización e Integración de datos | 69 |
| 4.2.2. | Herramientas De Analítica | 70 |
| 4.2.3. | Arquitectura | 73 |
| 4.3. | MODELO | 74 |
| 4.4. | EVALUACIÓN | 75 |
| 4.5. | DESPLIEGUE | 76 |
| 5. | VALIDACIÓN CON EXPERTOS | 80 |
| 6. | CONCLUSIONES | 83 |
| 7. | OPORTUNIDADES | 84 |
| 8. | BENEFICIOS..... | 85 |
| 9. | BIBLIOGRAFÍA..... | 86 |

LISTA DE TABLAS

| | |
|--|----|
| Tabla 1: Actividad en redes sociales | 16 |
| Tabla 2: Rentabilidad del sector asegurador..... | 19 |
| Tabla 3: Rentabilidad por tipo de seguro..... | 20 |
| Tabla 4: Resultado técnico por ramo..... | 21 |
| Tabla 5: Marco legal sector asegurador..... | 24 |
| Tabla 6: Propiedades de grafos relacionados con las redes sociales | 54 |
| Tabla 7: Indicadores y métricas a desarrollar..... | 58 |
| Tabla 8: Metas de cada indicador financiero..... | 59 |
| Tabla 9: Metas de cada indicador financiero..... | 59 |
| Tabla 10: Rangos para cada indicador financiero | 59 |
| Tabla 11: Rangos del indicador de clientes | 60 |
| Tabla 12: Aseguradoras y sus redes sociales..... | 63 |
| Tabla 13: Problemas de calidad en la localización del usuario | 68 |
| Tabla 14: Matriz de confusión..... | 75 |
| Tabla 15: Indicadores financieros propuestos para el sector asegurador | 90 |
| Tabla 16: Indicadores de clientes propuestos para el sector asegurador | 91 |

LISTA DE GRAFICOS

| | |
|--|----|
| Gráfico 1: Porcentaje de rentabilidad del sector asegurador | 20 |
| Gráfico 2: Evaluación de las redes sociales..... | 25 |
| Gráfico 3: Causa – Efecto entre las perspectivas del BSC | 32 |
| Gráfico 4: Ejemplo Básico de un BSC y sus indicadores | 35 |
| Gráfico 5: Ejemplo básico de un grafo..... | 37 |
| Gráfico 6: Ejemplo básico de un grafo dirigido | 37 |
| Gráfico 7: Agrupación de nodos..... | 40 |
| Gráfico 8: Actores principales en Inteligencia de Negocios. | 42 |
| Gráfico 9: Arquitectura de BI..... | 42 |
| Gráfico 10: Evolución de la minería de datos | 46 |
| Gráfico 11: Arquitectura para analizar redes sociales..... | 48 |
| Gráfico 12: Arquitectura modelo léxico-sintético | 50 |
| Gráfico 13: Comparación de metodologías para proyectos de minería de datos. | 52 |
| Gráfico 14: <i>Red social como un grafo</i> | 53 |
| Gráfico 15: BSC del sector asegurador | 55 |
| Gráfico 16: Perspectiva financiera del sector asegurador. | 56 |
| Gráfico 17: Perspectiva cliente del sector asegurador..... | 57 |
| Gráfico 18: Dashboard Financiero | 61 |
| Gráfico 19: Dashboard Cliente | 62 |
| Gráfico 20: herramientas más usadas por los científicos de datos | 71 |
| Gráfico 21: Cuadrante de gartner BI | 72 |
| Gráfico 22: Cuadrante de gartner BI. | 73 |

| | |
|---|----|
| Gráfico 23: Proceso de modelado..... | 74 |
| Gráfico 24: Tablero de control financiero | 76 |
| Gráfico 25: Tablero de control financiero – Métricas | 77 |
| Gráfico 26: <i>Tablero de control clientes</i> | 78 |
| Gráfico 27: <i>Tablero de control clientes - Métricas</i> | 78 |

INTRODUCCIÓN

“El 92% de los CEOs colombianos creen que la confianza en las empresas puede verse amenazada por las redes sociales”¹. Esto se debe de gran manera a la forma en que ahora se buscan referencias de un producto o un servicio. Antes, las personas solo recibían referencias de unos cuantos expertos, hoy en día y gracias al avance de la tecnología las personas prefieren buscar referencias del producto o el servicio en personas que lo hayan adquirido y tengan alguna experiencia, ya sea positiva o negativa. Es por esto que el cliente prefiere buscar comentarios en redes sociales o blogs que le den una percepción sobre las empresas que proveen el producto o el servicio. Esto hace que las empresas comiencen a preocuparse más sobre como los perciben las personas en redes sociales pues saben que malos comentarios en exceso hará que los clientes pierdan su confianza en ellos.

La tecnología también ha ayudado a mejorar la atención del cliente con los llamados contact center, mecanismos usados para atender quejas, reclamos o solicitudes de información de los clientes, no solo por teléfonos sino también chats y redes sociales. principalmente Facebook y Twitter², esto ha permitido que los clientes escriban en estas redes su percepción frente al servicio haciendo comentarios buenos o malos. Es por esto que muchas compañías han empezado a determinar indicadores de satisfacción de sus clientes en redes sociales, para de esta forma controlar y mejorar la calidad de los servicios. Estos indicadores son construidos de la mano con modelos de minería de texto quienes se apoyan en análisis de sentimiento y el procesamiento de lenguaje natural (PLN)³. El presente documento pretende realizar un

¹PWC. (20 de Mayo de 2017). *CEO Survey*. Obtenido de Para crecer en un mundo cambiante: Innovación con talento y tecnología: <http://www.pwc.com/co/es/assets/document/ceo-survey-colombia-7a-edicion.pdf>

²Asociación Española para la Calidad . (Diciembre de 2012). Obtenido de La calidad en la atención al cliente: https://www.aec.es/c/document_library/get_file?uuid=a17b6ac7-e816-4145-aeb9-a0c12e479d1b&groupId=10128

³Yuvila M, S., & Darnes Vilariño, M. J. (2015). Modelos para detectar la polaridad de los mensajes en. *Research in Computing Science*, 29-42.

modelo que le permita a las aseguradoras monitorearse a sí misma y al mercado mediante tableros de control financieros y de cliente a través de indicadores financieros del sector asegurador y de un modelo de polaridad de redes sociales usando herramientas libres.

El documento se divide en 5 capítulos principales, el planteamiento del problema donde se muestra el problema que se desea resolver, los alcances y delimitaciones del documento, luego viene la justificación del ejercicio y los objetivos propuestos. Luego entra el marco teórico donde se revisa toda la bibliografía, conceptos y trabajos realizados sobre teoría de grafos, redes sociales, polaridad, inteligencia de negocios, entre otros. Posteriormente se documenta la aplicación de la metodología Crisp-DM en el proyecto y por último se muestran las conclusiones, oportunidades y beneficios del proyecto.

1 PLANTEAMIENTO DEL PROBLEMA

1.1.DESCRIPCIÓN DEL PROBLEMA

A menudo, las compañías realizan anual o semestralmente un ejercicio de satisfacción de clientes con el fin de mejorar la calidad de sus servicios, ya que con el paso de los años han aprendido que la efectividad de la dirección de la calidad (TQM) puede reforzar su competitividad y proporcionar ventajas de este tipo en el mercado.⁴ Medir la calidad de los servicios por medio de la satisfacción del cliente le permite a las empresas determinar en qué punto de la cadena de valor deben intervenir para mejorar la calidad del servicio y así mejorar la percepción del cliente.

Las aseguradoras poseen diversas formas de calificar la percepción del cliente, SURA, por ejemplo, usa la metodología QSA de invamer group, Mapfre usa la metodología NPS (Net Promoter Score) y Bolívar posee su propia metodología. El problema es que estas metodologías son extensas de implementar y aunque permiten medir la satisfacción de los clientes efectivamente, estas son semestrales o anuales.

Ahora, como manifestó *Jorge Humberto Botero Presidente de Fasecolda*“... El entorno social está cambiando, los consumidores en todo tipo de productos y más en los consumos intangibles que son los consumos financieros, entre ellos los seguros, demandan productos más a su medida, requieren respuestas más prontas y el sector asegurador debe adaptarse a esa realidad, como también le corresponde desplegar nuevas plataformas tecnológicas para llegar de manera más eficiente a los consumidores de seguros...”⁵. Esto sugiere que medir la satisfacción de los clientes mensualmente o trimestralmente no es la mejor opción hoy en día ya que con el paso del tiempo, el avance de las tecnologías y un mercado cada vez más competitivo los directivos necesitan conocer sus indicadores en el menor tiempo posible para así determinar dónde y en qué momento actuar antes de que los problemas se materialicen.

⁴Tarí Guilló, J. J. (2000). *Calidad Total: fuente de ventaja competitiva*. Publicaciones Universidad de Alicante.

⁵HUMBERTO BOTERO, J. (06 de 05 de 2016). Entr. (B. Radio, Entrevistador)

1.2.ALCANCES Y DELIMITACIONES

Debido a que los análisis en redes sociales pueden abarcar temas de bastante complejidad se hace necesario acotar el proyecto para finalizarlo en los meses que dura el trabajo de grado. A continuación, se listan los alcances del proyecto.

- Se utilizará el informe de indicadores de gestión que Fasecolda emite mensualmente por medio un archivo de Excel, para compilar los resultados financieros desde el año 2010 hasta el 2016 y realizar los análisis pertinentes.
- Las redes sociales que se usarán para los análisis serán Facebook y Twitter.
- Los indicadores se realizarán únicamente para las compañías de seguros que posean páginas en Facebook o en Twitter.
- Solo se analizarán la información sobre los comentarios realizados, más no sobre variables demográficas como el género de la persona que publicó o el tipo de publicador (Asegurado, no asegurado, ex asegurado, corredor, etc.).
- Los análisis e indicadores estarán enfocados en compañías de seguros, no se discriminarán comentarios por ramo o por tipo de seguros (Vida, SOAT, ARL etc.).
- El cuadro de mando contendrá información de redes sociales a partir del primero de octubre de 2016 y se actualizará día a día permitiendo ver el historial por días, meses o años.
- El cuadro de mando contendrá información de los indicadores de gestión a partir del año 2010 y dependerá de la periodicidad con la que Fasecolda publique el informe.

- El análisis de sentimiento se realizará únicamente con palabras en español, de esta forma se tendrán en cuenta únicamente los comentarios en este mismo idioma.

2 JUSTIFICACIÓN

Hoy en día muchas organizaciones hacen uso de las redes sociales con el fin de llegar al público de una manera más directa, rápida y gratuita⁶ En estas redes se incluye mucha información de sus productos, mensajes de salud en el trabajo o información que le pueda interesar a sus clientes y mejorar la rentabilidad de su negocio; sin embargo, los usuarios están utilizando las redes sociales para incluir día a día comentarios buenos y malos sobre el producto. Comentarios que deberían ser incluidos en las mediciones de satisfacción del cliente ya que es un indicador muy útil para saber si los usuarios están conformes o no con los productos que ofrece cada compañía. A continuación, se muestra una tabla donde se puede evidenciar que varias compañías del sector asegurador están activas en las redes sociales de Facebook y Twitter.

| ACTIVIDAD EN REDES SOCIALES | | |
|-----------------------------|-----------------------------------|-----------------|
| Compañía | Facebook | Twitter |
| Suramericana | Seguros Sura | @SegurosSURA |
| Bolivar | Seguros Bolivar | @SegurosBolivar |
| Axa Colpatria | AXA COLPATRIA Oficial | @AXACOLPATRIA |
| Generali | GENERALI | @GENERALI_es |
| Metlife | MetLife | @MetLife |
| Old Mutual | Old Mutual | @OldMutualSA |
| Allianz | AllianzColombia | @Allianz_es |
| BBVA | BBVA | @BBVA_Colombia |
| Colmena | Colmena Seguros | @ColmenaSeguros |
| La equidad | Club Deportivo La Equidad Seguros | @SegurosEquidad |
| Liberty | Liberty Seguros Colombia | @LibertySegCol |
| Mapfre | MAPFRE | @MAPFRE_CO |
| Positiva | Positiva Compañía de Seguros | @PositivaCol |

Tabla 1: Actividad en redes sociales

En este trabajo se propone usar indicadores de satisfacción de cliente soportados en información de las redes sociales; lo anterior busca obtener decisiones en tiempos más cortos

⁶AROYO ALMARAZ, I., BALADRÓN PAZOS, A., & MARTÍN NIETO, R. (2013). La comunicación en redes sociales: Percepciones y usos de las Ong españolas. *Universidad Rey Juan Carlos*.

y detectar cambios en el servicio que estén alterando dichos indicadores. También, al finalizar el ejercicio, se tendrá un cuadro de mando donde los directivos de cada compañía podrán ver, además de la situación financiera de cada compañía, la evolución de la satisfacción de los clientes a través de un indicador de polaridad y comparar sus indicadores con los del mercado para de esta forma ser más competitivos en el sector asegurador.

3 OBJETIVOS GENERAL Y ESPECÍFICOS

3.1. OBJETIVO GENERAL

Diseñar e implementar un tablero de control, que les permita a los líderes de compañías de seguros, analizar:

- El estado y comportamiento de indicadores de gestión financiera.
- La opinión del mercado sobre sus compañías, en Facebook y Twitter.

3.2. OBJETIVOS ESPECÍFICOS

- Implementar técnicas de análisis de información, para desarrollar un análisis de polaridad sobre las opiniones acerca de las compañías aseguradoras del país, registradas en Facebook y Twitter.
- Diseñar e implementar una bodega de datos, que permita integrar la información de redes sociales y Fasescolda.
- Aplicar y automatizar técnicas de extracción de información de interés, en las redes sociales Facebook y Twitter.
- Diseñar e implementar un tablero de control integrado con la información de las compañías del sector asegurador en Colombia, que permita analizar indicadores de industria sobre su gestión financiera e indicadores de polaridad sobre las opiniones de sus clientes en redes sociales.

4 MARCO TEÓRICO

4.1.SECTOR SEGUROS EN COLOMBIA.

4.1.1. Seguros por asegurador

El sector asegurador en Colombia está conformado por 35 empresas aseguradoras reguladas por medio de la Superintendencia Financiera de Colombia y que a diciembre de 2016 dejaron una utilidad de 693.403 Millones. A continuación, se muestran las mejores empresas aseguradoras en términos de rentabilidad positiva.

| ASEGURADORA | RENTABILIDAD POSITIVA | |
|--------------|-----------------------|-------------------------|
| | (En Millones) | PORCENTAJE RENTABILIDAD |
| SURAMERICANA | 417.695 | 33% |
| ALFA | 280.337 | 22% |
| BOLIVAR | 262.183 | 21% |
| BBVA SEGUROS | 133.633 | 11% |
| COLMENA | 117.716 | 9% |
| OTROS | 47.026 | 4% |
| TOTAL | 1.258.589 | 100% |

Tabla 2: Rentabilidad del sector asegurador⁷

Como lo muestra la tabla 2, las aseguradoras principales de Colombia son Suramericana, Alfa y Bolívar ya que entre las tres obtienen el 76% de la rentabilidad del sector asegurador (Ver gráfico 1). La rentabilidad de estas empresas asciende a 960.215 Millones de pesos.

⁷FASECOLDA. (Marzo de 2016). *Indicadores de gestión.[En Línea]*. Recuperado el 14 de 02 de 2016, de <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>

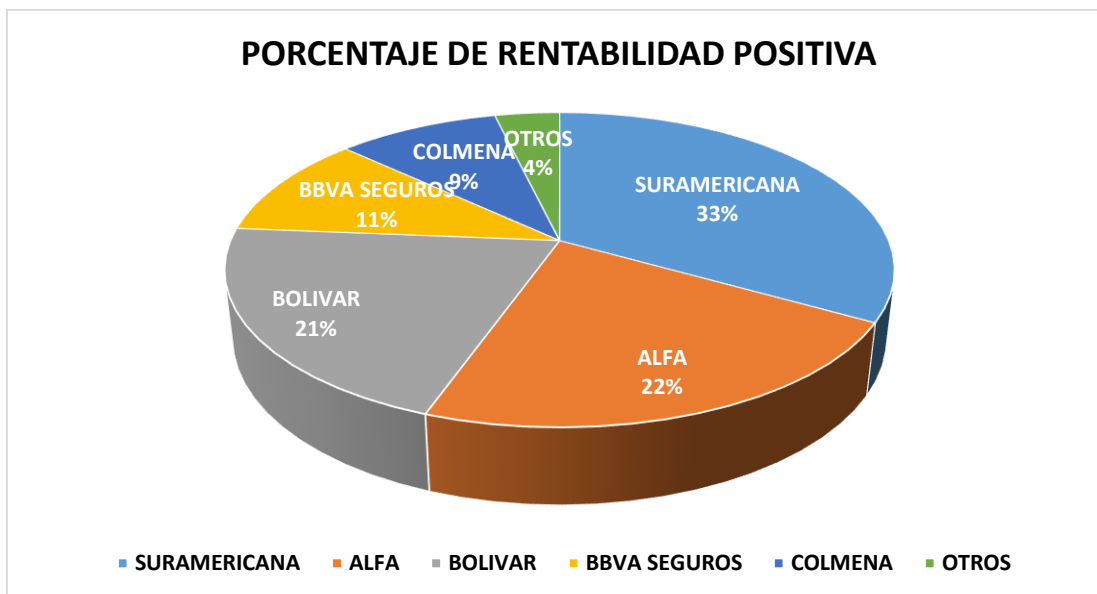


Gráfico 1: Porcentaje de rentabilidad del sector asegurador⁸

Los seguros se dividen en dos partes: los seguros de vida, enfocados en proteger la vida del asegurado o la calidad de vida; y los seguros generales, enfocados en proteger el patrimonio de los asegurados. La siguiente tabla muestra la distribución de la rentabilidad:

| TIPO DE SEGURO | RENTABILIDAD (En Millones) |
|----------------|----------------------------|
| GENERALES | -652.138 |
| VIDA | 1.345.542 |
| TOTAL | 1.258.589 |

Tabla 3: Rentabilidad por tipo de seguro⁹

⁸FASECOLDA. (Marzo de 2016). *Indicadores de gestión.* [En Línea]. Recuperado el 02 de febrero de 2016, de <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>

⁹FASECOLDA. (Marzo de 2016). *Indicadores de gestión.* [En Línea]. Recuperado el 02 de febrero de 2016, de <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>

Como se puede observar, los seguros de vida componen la mayor parte de la rentabilidad con un valor superior al billón de pesos.

4.1.2. Seguros por ramo

En Colombia existen 33 ramos que componen el sector asegurador, estos se miden por la gestión sobre el ramo, es decir que no se tiene en cuenta las inversiones sobre las reservas que realice la aseguradora. Como consecuencia, la variable por la cual se miden los ramos es el resultado técnico, esto es, la gestión pura del ramo. A continuación, se muestran los mejores ramos del sector asegurador.

| RAMO | RESULTADO TÉCNICO |
|------------------------|-------------------|
| VIDA GRUPO Y COLECTIVO | \$ 254.172 |
| ACCIDENTES PERSONALES | \$ 70.465 |
| DESEMPLEO | \$ 68.735 |
| SUSTRACIÓN | \$ 44.204 |
| CUMPLIMIENTO | \$ 22.267 |
| TRANSPORTE | \$ 21.768 |
| MANEJO | \$ 14.356 |
| HOGAR | \$ 7.986 |
| SEGUROS DE CRÉDITO(2) | \$ 2.510 |
| NAVEGACION Y CASCO | \$ 1.547 |
| OTROS | -\$ 2.727.988 |
| TOTAL | -\$ 2.219.980 |

Tabla 4: Resultado técnico por ramo¹⁰

Como se puede observar los mejores ramos son: vida grupo y colectivo; accidentes personales; desempleo y sustracción.

¹⁰FASECOLDA. (Marzo de 2016). *Indicadores de gestión.[En Línea]*. Recuperado el 23 de Mayo de 2016, de <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>

4.1.3. Percepción del cliente en la industria aseguradora

La percepción del cliente está relacionada directamente con la calidad de los servicios prestados por la empresa o, en este caso, las aseguradoras. Según Zeithaml y Bitner la calidad se define como “*La función de la discrepancia existente entre las expectativas de los usuarios sobre el servicio que van a recibir y sus percepciones sobre el servicio efectivamente prestado por la empresa.*”¹¹ Además, “la excelencia en la calidad le permitirá a una empresa ser más competitiva en cualquier sector de la industria.”¹² Es por esto que muchas de ellas se han enfocado en darle al cliente un servicio de calidad que les permita competir en el mercado.

Las aseguradoras poseen diferentes formas de determinar la satisfacción de los clientes, a continuación, se muestra la metodología a través de la que algunas de ellas miden la percepción del cliente:

- **SURAMERICANA:** Usa cada año la metodología QSA de Invamer Gallup, con ella determina la percepción de los clientes en términos de satisfacción, lealtad y recomendación. Los resultados son insumo para los planes de mejora en los procesos. (Informe de sostenibilidad 2014)¹³.
- **BOLIVAR:** Posee su propia metodología para conocer la opinión de sus clientes y la determina desde tres puntos de vista: experiencia, satisfacción y recomendación. Para

¹¹PRIETO, A. T., & HERNANDEZ, C. (20 de Junio de 2014). *calidad de servicio como factor de competitividad en los postgrados de las instituciones universitarias públicas*. Recuperado el 03 de Julio de 2016.

¹²REQUENA PONCE, M. V., & SERRANO LÓPEZ, G. C. (02 de Novimebre de 2007). *Calidad del servicio desde la perspectiva de los clientes, usuarios y auto-percepción de empresas de captación de talento*. Recuperado el 04 de Julio de 2016.

¹³SURAMERICANA;Filial de Grupo SURA. (2014). *Informe anual*.

el año 2015 la compañía realizó alrededor de 30 estudios de experiencia en los cuales incluyeron cerca de 11374 clientes. (Informe de gestión y estados financieros 2015)¹⁴.

- **MAPFRE:** Crea el Observatorio de Calidad Mapfre en el que pretende concentrar todos los análisis que se realicen frente a la calidad de los productos ofrecidos a sus clientes, estos análisis se realizan semestralmente por medio de encuestas hacia los clientes y se evalúan por medio de la metodología NPS (Net Promoter Score)¹⁵.

4.1.4. Marco Legal del sector asegurador

La normativa del sector asegurador es bastante densa, sin embargo se intentó condensar la información y extraer las leyes, decretos y artículos más sobresalientes del sector.

| | | |
|--|---|--|
| LEY 100 DE 1993 A LEY DE SEGURIDAD SOCIAL, APROBADA | Abre al sector asegurador un gran potencial en la suscripción de riesgos profesionales, pensiones y planes de salud obligatorios para la población general. | CONGRESO DE LA REPUBLICA |
| DECRETO 2150 DE 198 | Permite a las aseguradoras diseñar y comercializar planes alternativos y complementarios para pensiones, seguros de pensiones, planes complementarios, y administrar pasivos pensionales. | CONGRESO DE LA REPUBLICA |
| RESOLUCIÓN 5148 DEL 31 DE DICIEMBRE DE 1991 | Autoriza la operación a la Compañía en el ramo de vida individual | SUPERINTENDENCIA FINANCIERA DE COLOMBIA |
| DECRETO 410 DE 1971 | Establece los principios comunes para los seguros de vida y de personas | CÓDIGO DE COMERCIO COLOMBIANO |
| LEY 33 FEB. 1986 | Se crea el seguro por daños a personas en accidentes de tránsito, es obligatorio y el | CONGRESO DE LA REPÚBLICA |

¹⁴SEGUROS BOLIVAR. (2015). *Informe de Gestión*.

¹⁵MAPFRE. (2015). *Responsabilidad Social - Informe anual*.

| | | |
|--|---|-------------------------------------|
| | perjudicado tiene acción directa contra el asegurador. | |
| LEY 675 DE 2001 ARTICULO 15 PARAGRAFO 1 | Establece la obligatoriedad de la constitución del seguro de incendio para propiedad horizontal | CONGRESO DE LA REPUBLICA |
| ART 188 NRO 1 | Fija seguros para transporte terrestre, aéreo, fluvial en territorio Colombiano. | ESTATUTO ORGÁNICO FINANCIERO |
| ARTÍCULO 5 LEY 1581 DE 2012 | Determina la definición de datos sensibles. | EL CONGRESO DE COLOMBIA |
| LEY 1266 DE 2008 | Dicta las disposiciones generales del habeas data y regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios así como la proveniente de países terceros, dicta otras disposiciones | EL CONGRESO DE LA REPUBLICA |
| LEY 1581 DE 2012 | Desarrolla el derecho constitucional que tienen todas las personas a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bases de datos o archivos. | EL CONGRESO DE COLOMBIA |

Tabla 5: Marco legal sector asegurador

4.2.REDES SOCIALES

Una red social se define como un servicio que permite a los individuos (1) construir un perfil público o semipúblico dentro de un sistema delimitado, (2) articular una lista de otros usuarios con los que comparten una conexión, y (3) ver y recorrer su lista de las conexiones

y la lista de conexiones realizadas por otros dentro del sistema¹⁶. Su inicio viene desde los años 90 con Classmate y SixDegrees pasando por LinkedIn en 2003, Facebook 2004, Youtube en 2005, Twitter 2006 hasta llegar a WhatsApp, Instagram y Google Plus, como se puede ver a continuación en el gráfico 3.

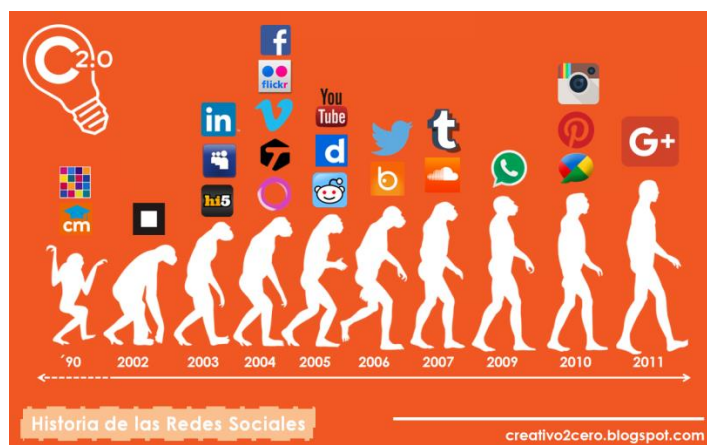


Gráfico 2: Evaluación de las redes sociales

Según Pablo Burgeño¹⁷ las redes se clasifican de la siguiente forma:

Redes sociales Horizontales: Son aquellas dirigidas a todo tipo de usuario y sin una temática definida. Se basan en una estructura de celdillas que permiten la entrada y participación libre y genérica sin un fin definido distinto del de generar masa. A continuación, podemos encontrar algunos ejemplos:

- **Redes de contactos:** Son redes sociales dedicadas a conectar amigos, familiares y personas conocidas, los usuarios pueden compartir fotos y videos. Algunos ejemplos

¹⁶BOYD, D. M., & ELLISON, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*.

¹⁷BURGEÑO, P. F. (02 de Marzo de 2009). *Clasificación de Redes Sociales*. Recuperado el 04 de Julio de 2016, de <http://www.pabloburgueno.com/2009/03/clasificacion-de-redes-sociales/>

son: Facebook (www.facebook.com), Bebo(www.bebo.com), Myspace (www.myspace.com), Tuenti (www.tuenti.com), y Orkut(www.orkut.com).

- **Redes profesionales:** Redes sociales para crear relaciones profesionales, las más conocidas son: LinkedIn (www.linkedin.com), Xing(www.xing.com), y Plaxo (www.plaxo.com).
- **Miroblogging:** Son redes sociales dedicadas al envío de mensajes cortos vistos por cualquier usuario de la red. Las más usadas son Twitter (www.twitter.com), PingGadget (www.pinggadget.com), Jaiku (www.jaiku.com), Plurk(www.plurk.com), Beeing(www.beeing.com), y Pownce (pownce.com).
- **Redes sociales para compartir información basada en la geolocalización del usuario:** Foursquare (foursquare.com).

Redes sociales Verticales: Están concebidas sobre la base de un eje temático agregado. Su objetivo es el de congregar un colectivo concreto en torno a una temática definida. A su vez, pueden clasificarse en función de su especialidad:

- **Redes sociales Verticales Profesionales:** Están dirigidas a generar relaciones profesionales entre los usuarios. Los ejemplos más representativos son Viadeo, Xing y Linked In.
- **Redes sociales Verticales De Ocio:** Su objetivo es congregar a colectivos que desarrollan actividades de ocio, deporte, usuarios de videojuegos, fans, etc. Como ejemplo cabe mencionar a Wipley, Minube, Dogster, Last.FM y Moterus.

Redes sociales Verticales Mixtas: Ofrecen a usuarios y empresas un entorno específico para desarrollar actividades tanto profesionales como personales en torno a sus perfiles, algunos ejemplos son: Yuglo, Unience, PideCita, 11870. Los ejemplos de estas redes sociales son:

- **Bytepr**(www.bytepr.com): Red social de los profesionales de la comunicación y el marketing

- **Esanum**(www.esanum.es): Red social para los profesionales del mundo de la salud y la medicina.
- **Nosolotop**(www.nosolotop.com): Red social para fotógrafos, modelos, maquilladores, estilistas y peluqueros.
- **Hospitalityhub** (www.hospitalityhub.com): Red social para facilitar el contacto entre los profesionales del sector turismo y alojamiento.
- **CinemaVIP** (www.cinemavip.com): Red profesional del sector audiovisual
- **Domestika** (www.domestika.org): Red social diseñada para profesionales especializados en diseño y creatividad¹⁸.

Para el desarrollo de este proyecto se hará uso de las dos redes sociales horizontales más importantes en la actualidad Twitter y Facebook.

4.2.1. Twitter

Fundada por Jack Dorsey en 2006, esta red social se define como una plataforma de microblogging que permite compartir mensajes cortos (hasta 140 caracteres) con cualquier usuario de internet, y cada usuario tiene una lista de contactos que sigue estos mensajes¹⁹

El funcionamiento de Twitter es relativamente sencillo: Los usuarios pueden publicar mensajes cortos denominados tweets para comentar noticias, eventos, su estado de ánimo u opinión sobre un asunto, etc.

Los mensajes que se expresan allí son públicos y cualquier usuario de la red social puede consultarlo. Sin embargo, si así lo prefiere, el dueño de la cuenta puede restringir las personas que acceden a los tweets publicados.

¹⁸Gómez Vieites, A., & Otero Barros, C. (2013). *Redes Sociales en la empresa*. Bogotá, Colombia: Ediciones de la U.

¹⁹Perez Soler, S. (19 de 02 de 2017). *Tesis en red*. Obtenido de http://www.tesisenred.net/bitstream/handle/10803/362644/Tesi_Susana%20Perez.pdf

Al acceder a la página web de Twitter se observa que hay cuatro parámetros importantes para el usuario: “Tweets”, “Siguiendo”, “Seguidores” y “Listas”. La información compartida puede contener opiniones personales con enlaces a páginas web, vídeos, fotos, etc. Los tweets permiten interactuar en línea facilitando que las personas respondan de forma fácil al usuario que ha publicado algún tweet mediante la opción “reply”. Además, cuando un usuario comparte el tweet que ha recibido origina un retweet, produciéndose una retroalimentación de la información en la red; de esta forma, los usuarios utilizan los tweets para comunicarse entre sí y difundir la información de forma viral. El parámetro “siguiendo” (following) hace referencia a las cuentas que sigue el usuario en su perfil de Twitter. Mientras que el parámetro “seguidores” (followers) muestra quienes siguen la cuenta del usuario. En las «Listas» se encuentran los diferentes temas fijados, que generan las conversaciones²⁰.

4.2.2. Facebook

Facebook es una red social horizontal creada por Mark Zuckerberg en el año 2006 que cuenta con ciertas características que la hacen única y líder en las redes sociales, más de 1500 millones de usuarios en el mundo, la convierten en un medio masivo al alcance tanto de personas con conocimientos informáticos, como de personas sin conocimientos específicos de tecnología informática. A pesar de no ser la primera red social, ha sido la que más ha impactado; dejando de ser la red social de las universidades elite de los Estados Unidos, para convertirse en una plataforma de fácil acceso para todas las personas alrededor del mundo. Su interfaz amigable e interactiva es una de sus grandes cualidades²¹.

Básicamente, Facebook ofrece:

- Posibilidad de subir fotos y videos, organizados por caretas o álbumes.

²⁰Guzmán Duquea, A. P., Moral Pérez, M. E., & Ladrón de Guevara, F. G. (2016). Usos de Twitter en las universidades iberoamericanas. *RELATEC*, 28.

²¹Ricardo Barreto, C., & Chavarro Jimenez, A. (2010). El uso de Facebook y Twitter en educación. *Instituto de estudios en educación - IESE*, 5

- Publicación de mensajes de texto con información del estado del usuario, notas, fotos, videos y enlaces en el muro asociado al perfil del usuario.
- Gestión de listas de amigos.
- Definición y gestión de eventos.
- Creación de preguntas y análisis de los resultados.
- Definición de grupos abiertos y cerrados.

4.3.BALANCED SCORECARD

El Balanced ScoreCard es un modelo de gestión creado por Kaplan y Norton en la universidad de Harvard, que se traduce en la estrategia de una organización en objetos relacionados, medidos a través de indicadores y ligados a unos planes de acción que permitan alinear el comportamiento de los miembros de la organización²². Por lo tanto el BSC se concibe como un proceso descendente que consiste en traducir la misión y la estrategia global de la empresa en objetivos y medidas más concretos que puedan inducir a la acción empresarial oportuna y relevante²³.

4.3.1. Beneficios del Balance ScoreCard

EL balance Scorecard es una de las herramientas más utilizadas para la planificación y ejecución de la estrategia dado que posee los siguientes beneficios²⁴:

²²Fernández, A. (2001). El balance ScoreCard: Ayudando a implantar la estrategia. *Universidad de Navarra - IESE*, 32.

²³Blanco Dopico, M. I., Guzmán, B., & Agra, S. (1999). El enfoque conductual contable y su reflejo en un cuadro de mando integral. *Revista española de financiación y contabilidad*, 77-104

²⁴Alvarez Medina , M. T., & Moreno Velarde, S. A. (s.f.). El balance ScoreCard, Una herramienta para la planeación estratégica. *Itson*.

1. **Alineación de los empleados hacia la visión de la empresa:** Permite que toda la organización conozca hacia donde se dirige la empresa y como contribuye para lograr la visión.
2. **Mejora de la comunicación hacia todo el personal de los objetivos y su cumplimiento:** Permite que cada uno de los colaboradores de la compañía conozca los objetivos que ha plasmado la organización para lograr la visión y de esta forma saber cómo puede ayudar con su labor diaria.
3. **Redefinición de la estrategia con base en resultados:** El Balance ScoreCard permite evaluar periódicamente el cumplimiento de la estrategia, de esta forma y en cualquier momento se puede determinar si es necesario ajustar o redefinir la estrategia de la compañía.
4. **Traducción de la visión y la estrategia en acción:** El Balance ScoreCard le permite a la organización traducir la estrategia y la visión en acciones concretas que garantizaran su cumplimiento.
5. **Orientación de la compañía hacia la creación de valor:** El Balance ScoreCard fomenta el trabajo en equipo y garantiza que los colaboradores trabajen todos en función de un objetivo común, crear valor para el cliente a través de la cadena de valor.
6. **Mejora de la capacidad de análisis y de la toma de decisiones:** Los directivos mejoran los métodos para la toma de decisiones debido a que las decisiones que se toman alrededor del Balance ScoreCard son decisiones estratégicas y tendrán un gran impacto en la compañía.

4.3.2. Perspectivas

El Cuadro de Mando debe contar la historia de la estrategia, empezando por los objetivos financieros a largo plazo para luego vincularlos a la secuencia de acciones que deben realizarse con los procesos financieros, los clientes, los procesos internos y después con los empleados y los sistemas, para entregar la deseada actuación económica.²⁵De aquí Kaplan y Norton establecen cuatro perspectivas que permitirán convertir la estrategia de una organización en acciones y resultados:

Perspectiva Financiera: Tiene como finalidad u objetivo dar respuesta a las expectativas de los socios. Su orientación está centrada en la creación de valor para los accionistas, con altos índices rentables y garantías de desarrollo y mantenimiento del negocio²⁶.

Perspectiva del cliente: Tiene como objetivo verificar que la compañía le entregue el valor agregado al cliente. Esto se realiza mediante la medición de su satisfacción, percepción y necesidades. Con ello se garantiza la fidelización del cliente, posibilidades de nuevos productos y el apoyo en el cumplimiento de la perspectiva financiera.

Perspectiva de procesos: Esta perspectiva monitorea la correcta ejecución de los procesos claves que garantizan la generación de valor al cliente y así cumplir las expectativas de los accionistas.

Perspectiva de aprendizaje y crecimiento: se enfoca en las capacidades del personal de la organización que ejecutan los procesos clave, la infraestructura necesaria para que los procesos puedan lograr sus objetivos y el clima organizacional que garantiza un buen trabajo al interior de la organización.

Las perspectivas son una cadena de causa y efecto. Para cumplir con los objetivos plasmados en la perspectiva financiera se deben cumplir los objetivos de la perspectiva del cliente, ya

²⁵Kaplan, R., & Norton, D. (2011). Cuadro de mando integral: the Balanced Scorecard. *Gestión 2000*

²⁶Alveiro Montoya, C. (2011). El banaced scorecard como herramienta de evaluación en la gestión administrativa. *Visión de futuro*.

que los clientes son el centro fundamental del sostenimiento de la empresa; estos se fidelizan siempre y cuando se les garantice la propuesta de valor, la cual se alcanza con la efectividad en los procesos clave de la compañía, que a su vez son efectivos cuando se tiene un buen clima laboral, personal calificado e infraestructura tecnológica, ver gráfico 4.

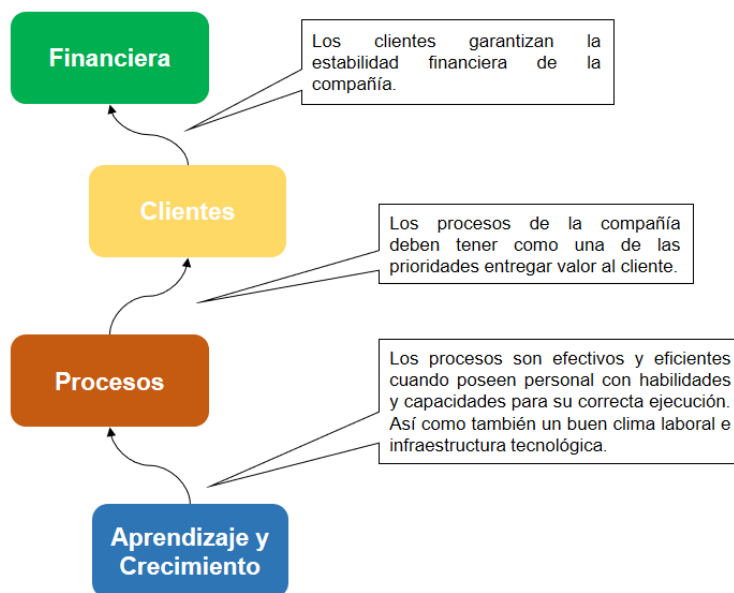


Gráfico 3: Causa – Efecto entre las perspectivas del BSC

4.3.3. Indicadores claves de rendimientos (KPI)

Las perspectivas del Balanced ScoreCard están compuestas de indicadores clave de rendimiento que permiten monitorear la consecución de la estrategia, para ello revisaremos que significa cada uno de estos conceptos:

La métrica es una valoración cuantitativa de estadísticas que describe tanto los eventos como las tendencias de un determinado evento²⁷. Algunos ejemplos son: Número de alumnos,

²⁷Kaushik, A. (2011). *analitica web 2.0*. Barcelona: Gestión 2000.

Número de profesores, horas de clase, horas de descanso etc. Un indicador clave de rendimiento (KPI) es una métrica que ayuda de un modo más racional a conseguir los objetivos previstos, dos ejemplos de KPI son: porcentaje de estudiantes aprobados y tasa de ingreso en la universidad. Como se puede observar todos los KPI son métricas, pero no todas las métricas son KPI o no son lo suficientemente importantes para ser consideradas KPI²⁸.

Cada perspectiva tiene sus KPI agrupados en objetivos estratégicos los cuales garantizan el cumplimiento de la estrategia. A continuación, se muestran algunos indicadores básicos de un Balance SocreCard²⁹:

KPI Financieros

- Valor económico agregado –EVA
- Retorno sobre capital empleado – ROCE
- Margen de operación
- Ingresos
- Rotación de activos
- Retorno de la inversión –ROI
- Relación deuda/patrimonio
- Inversión como porcentaje de la venta

KPI Clientes

- Satisfacción de clientes
- Desviaciones de acuerdos de servicios
- Reclamos resueltos del total de reclamos
- Incorporación y retención de clientes
- Mercado

²⁸Aguilar, L. J. (2015). *Sistemas de Información en la empresa*. Mexico: Alfaomega.

²⁹RAMÍREZ RAMÍREZ, C. P. (2010). *El Balanced Scorecard como herramienta metodológica para evaluar la gestión en la unidad de compras en el municipio de Medellín*. Tesis de Especialización en Alta Gerencia, Universidad de Medellín

KPI Procesos Clave

- Tiempo del ciclo del proceso
- Costo unitario por actividad
- Niveles de producción
- Costos de falla
- Costos reprocesos, desperdicio (costos de calidad)
- Beneficios derivados del mejoramiento continuo
- Eficiencia de uso de los activos

KPI Aprendizaje y Crecimiento.

- Brecha de competencias clave (personal)
- Desarrollo de competencias clave
- Retención de personal clave
- Captura y aplicación de tecnologías y valor agregado
- Ciclo de toma de decisiones clave
- Disponibilidad y uso de información estratégica
- Progreso en sistemas de información estratégica
- Satisfacción del personal
- Clima organizacional

En el gráfico 5 se muestra un ejemplo de un BSC con sus perspectivas, indicadores claves de rendimiento y la relación de causa-efecto

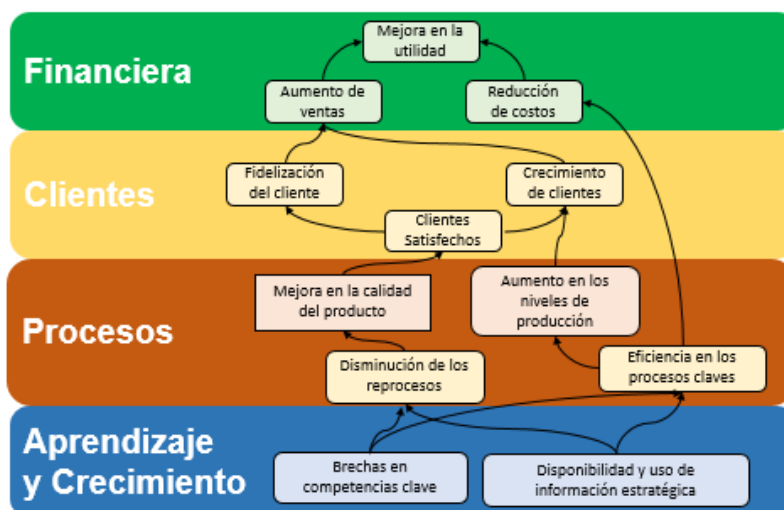


Gráfico 4: Ejemplo Básico de un BSC y sus indicadores

4.3.4. Casos de éxito del Balance ScoreCard

Kaplan y Norton en su libro “como utilizar el cuadro de mando integral” exponen dos casos de éxito los cuales se mencionan a continuación³⁰:

- **Mobil North America Marketing and Refining**

Esta empresa con 1500 Millones de dólares al año, ocupaba el último lugar en rentabilidad entre sus iguales del sector y necesitaba de una inyección de capital cercana a los 500 millones de dólares de casa matriz solo para mantener y mejorar las instalaciones. El equipo directivo desarrolló una nueva estrategia orientada al cliente y se está introdujo en el cuadro de mando para comunicar la estrategia. Mobil ocupó el primer lugar en la clasificación luego de dos años y logró obtener beneficios con un 56% superior a la media del sector. Brian Baker vicepresidente ejecutivo manifestó lo siguiente: “En 1997 ocupamos el primer lugar de la clasificación por tercer año consecutivo, algo que no tiene precedentes en la gran industria del petróleo... El

³⁰Kaplan, R., & Norton, D. (2009). *Cómo utilizar el Cuadro de Mando Integral*. Barcelona: Gestión 2000.

cuadro de mando se lleva la parte del león en este mérito. Con el cuadro de mando integral creamos una mentalidad en la que los resultados tienen una máxima importancia”.

- **CIGNA Property & casualty insurance**

En el año 1993, la Property & Casualty Division de CIGNA perdió cerca de 257 Millones de dólares, siendo así sus resultados los peores del sector. La división estaba casi en bancarrota. El equipo directivo reformó la estrategia y decidieron transformarse en especialistas, centrándose en nichos donde tenían una ventaja competitiva gracias a la información que poseían. El equipo directivo desplegó la estrategia a las 24 unidades de negocio en 1994 usando el Balance ScoreCard como proceso integral de gestión. Los resultados fueron rápidos y drásticos, en dos años la compañía obtuvo utilidad manteniendo y mejorando su resultado por los siguientes cuatro años. Según Gerald Isom, Presidente de CIGNA Property & Casualty el cuadro de mando integral tuvo un papel importante destacado en el éxito de esta historia: “CIGNA utilizó el cuadro de mando integral para lograr su transformación”.

4.4.TEORIA DE GRAFOS

4.4.1. Definición

En matemáticas y ciencias de la computación, un grafo, del griego *grafos* (dibujo, imagen), es un conjunto de objetos llamados vértices (nodos, agentes o actores) unidos por enlaces llamados aristas o arcos, que permiten representar relaciones binarias entre los elementos de un conjunto³¹

³¹Trudeau, R. (1993). Introduction to Graph Theory

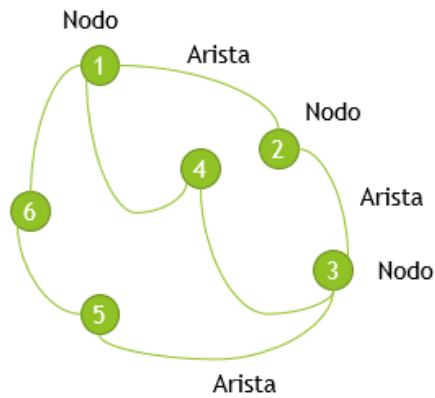


Gráfico 5: Ejemplo básico de un grafo

Existen dos tipos de grafos, los grafos no dirigidos y los dirigidos: los grafos no dirigidos se caracterizan por ser de doble vía, es decir, el nodo 1 puede generar entradas al nodo 2 y viceversa, el nodo 2 puede generar entrada al nodo 1. El grafico 6 muestra un grafo no dirigido; los grafos dirigidos se caracterizan por ser de una sola vía, es decir, el nodo 1 puede generar entradas al nodo 2 pero el nodo dos no puede generar entradas al nodo 1. Este tipo de grafos también se conocen como dígrafos. El grafico 7 muestra ejemplo de un grafo dirigido.

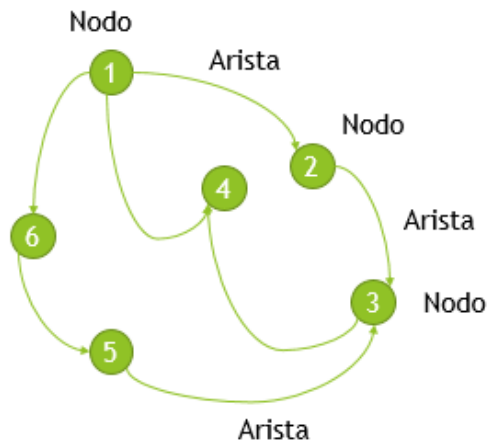


Gráfico 6: Ejemplo básico de un grafo dirigido

4.4.2. Propiedades de los grafos

Una vez entendido el concepto de grafos, se pueden revisar las propiedades de estos³²:

- **Tamaño de la red:** Se determina en términos del número de aristas o del número de nodos, siendo este último crítico para la estructura de la red a causa de los recursos limitados que cada actor pueda tener, destinándolos a la construcción y el sostenimiento de sus vínculos.
- **Densidad de la red:** Es el número de aristas de la red expresado como una proporción del número total de aristas posibles. En una red dirigida de tamaño n , el número de aristas posibles es $n \times (n-1)$, mientras que en una red no dirigida del mismo tamaño, el número de aristas posibles es $n \times (n-1)/2$.
- **Grado de los nodos:** En una red no dirigida, esta medida es el número de aristas que tiene determinado nodo, mientras que en una red dirigida hay que distinguir entre las aristas que llegan (grado de entrada) y las que salen (grado de salida).
- **Distancia entre nodos:** Número mínimo de aristas que hay al ir de un nodo al otro.
- **Recorrido:** Secuencia de nodos y aristas que comienza y termina con nodos. Puede ser cerrado (si inicia y termina en el mismo nodo) y no son restringidos. Esta secuencia también puede incluir al mismo nodo o a la misma arista en múltiples ocasiones, razón que los convierte en la forma más general de conexión entre dos nodos.

³²UIAF. (2014). *Unidad de Información y Análisis Financiero*. Obtenido de UIAF:
[http://www.urosario.edu.co/observatorio-de-lavado-de-activos/Archivos_Lavados/Aplicabilidad-de-la-mineria-de-datos-\(1\).pdf](http://www.urosario.edu.co/observatorio-de-lavado-de-activos/Archivos_Lavados/Aplicabilidad-de-la-mineria-de-datos-(1).pdf)

- **Ciclo:** es un camino cerrado de 3 o más nodos diferentes excepto por el nodo origen/destino.
- **Camino:** Cualquier recorrido en el que una arista es incluida una sola vez.
- **Ruta:** Recorrido en el que cada nodo es incluido una sola vez.
- **Excentricidad de los nodos:** Es la mayor distribución de las distancias mínimas del nodo al resto de actores, esta mide qué tan lejos está un actor de su contraparte más lejana.
- **Diámetro y radio de la red:** El diámetro es la máxima excentricidad entre todos los nodos de la red y da el número de pasos suficientes para ir desde cualquier nodo a cualquier otro; el radio es la mínima excentricidad entre todos los nodos de la red.
- **Accesibilidad:** Cualidad de un nodo y otro de desarrollar una serie de vértices que logren la conexión entre ellos.
- **Agrupación (Clustering):** Agrupaciones de nodos que se relacionan ampliamente de manera constante por variables como la distancia o la cantidad de relaciones entre ellos. En la gráfica 8 se pueden observar 5 Clusters de un grafo.

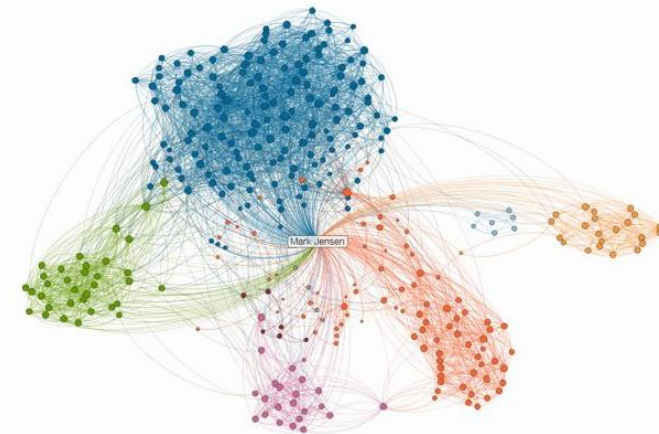


Gráfico 7: Agrupación de nodos

- **Centralidad de intermediación:** Son nodos que están en medio de dos grupos de nodos que sin él no se lograrían comunicar.

4.5.ANALÍTICA

4.5.1. Introducción a la Inteligencia de negocios.

La inteligencia de negocios es un conjunto de procesos y tecnologías que le permiten a las diferentes áreas de la organización tomar decisiones basadas en datos coherentes con el negocio y únicos para la organización. El activo más importante en inteligencia de negocios son las personas y su éxito dependerá de ellas, sin personas para extraer conclusiones sobre la información y formular planes de acción ella partir de esta, la inteligencia de negocios se volverá un costoso reportador de información. Por esta razón, la inteligencia de negocios es más que una herramienta tecnológica; se trata de la cultura, el análisis estratégico, y el valor de la información. La tecnología posibilita la inteligencia de negocios, sin embargo, demasiada importancia puesta en la tecnología puede sabotear las iniciativas de inteligencia de negocios. Vale la pena señalar que son las personas quienes harán de los esfuerzos de BI un enorme éxito o un profundo fracaso.

El proceso de inteligencia de negocios se basa en transformar los datos en información para luego convertirla en conocimiento, sin embargo, estos conceptos aún son similares o iguales para muchas personas, a continuación, se muestra lo que significa cada termino:

Dato: Un hecho bruto, simple y discreto³³, es decir una cifra sin procesar, sin ninguna interpretación ni ningún análisis añadido.; un buen ejemplo de dato seria precio, petróleo, \$80 dólares, barril.

Información: Son datos procesados, organizados y con significado, necesarios para la creación de conocimiento³⁴.; un ejemplo de información seria “El precio del crudo de petróleo ha subido de \$70 a \$80 dólares por barril”, esto aporta significado a los datos, por eso se dice que aporta información a la persona que estudia los precios del crudo.

Conocimiento: Es información con sentido y valiosa³⁵ para la organización; básicamente el sentido y lo valiosa que sea la información se definen por las personas y su experiencia. El siguiente es un ejemplo de conocimiento: Cuando los precios del crudo suben \$10 dólares por barril, es probable que los precios de la gasolina suban 2 centavos por litro.

Como se puede observar en la siguiente gráfica, la tecnología transforma los datos en información y las personas convierten esa información en conocimiento. Es decir, el éxito de BI depende de las personas y de tecnología.

³³Bhatt, G. (2001). Knowledge management in organisations: Examining the interaction between technologies, techniques, and people. *Journal of Knowledge*, 8-18

³⁴Spek, R., & Spijkervet, A. (1997). Knowledge management: Dealing Intelligently with knowledge. *Utrecht: Kenniscentrum CIBIT*, 25.

³⁵Davenport, T., & Prusak, L. (1998). Working Knowledge: How organisations manage what they know. *Harvard Business School Press*, 224.

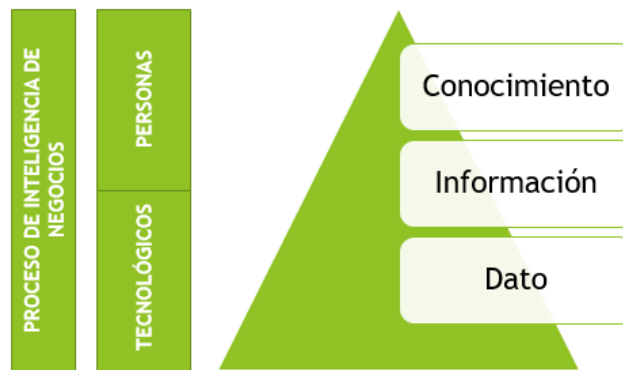


Gráfico 8: Actores principales en Inteligencia de Negocios.

Para convertir los datos en información se deben tener procesos tecnológicos bien definidos que permitan transformar y unificar los datos, estos procesos se describen a continuación:

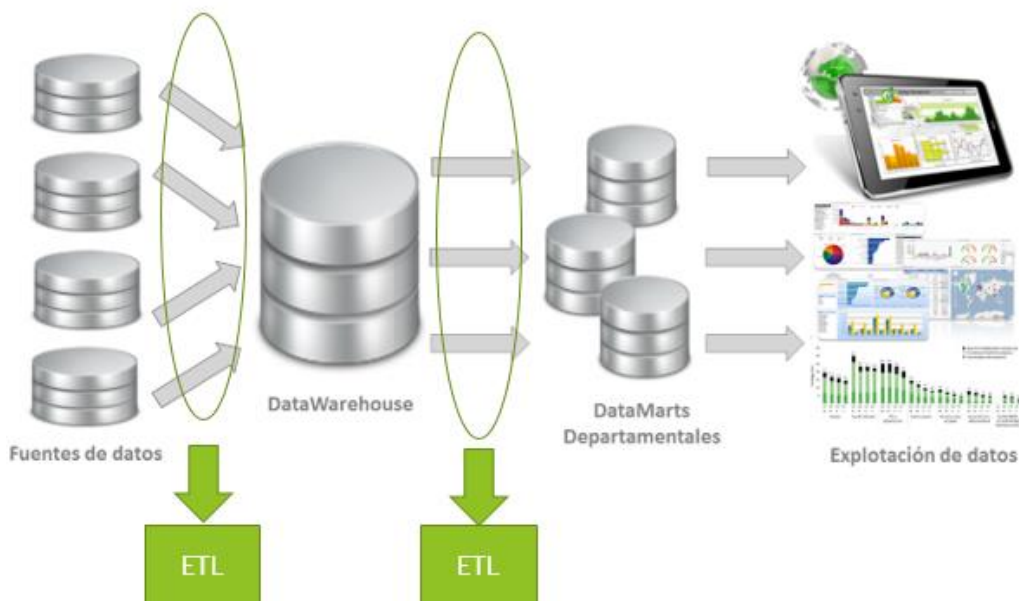


Gráfico 9: Arquitectura de BI

Fuentes de datos: Son bases de datos relacionales que soportan la operación del día a día del negocio y están diseñadas para tener respuestas rápidas y con una tolerancia muy baja a fallos.

ETL: Es un término usado para referirse al movimiento y transformación de los datos. Se basa en la extracción de diferentes fuentes de datos (principalmente transaccionales) que luego son transformados, integrados según la necesidad y cargados en otra base de datos, ya sea un DataWarehouse o Datamarts

DataWarehouse: Es una base de datos diseñada para analizar la información, tiene una estructura diferente de una base de datos relacional, soporta gran cantidad de estos, garantiza la unicidad en la información a nivel de compañía y la trazabilidad de los datos.

DataMarts: Es un subconjunto del DataWarehouse que se enfoca en proveerle información netamente para análisis a departamentos.

Explotación de datos: Son herramientas enfocadas en la visualización de datos, con esta se busca que los usuarios exploten la información y generen conocimiento. Esto se realiza a través de cuadros de mando, indicadores y métricas de información.

Para que el proceso de BI tenga éxito es necesario aplicar un concepto denominado gestión visual, el cual se encarga de asegurar que la información sea interiorizada por todos los integrantes de la compañía (directivos y operarios). Al interiorizar la información y reflejar su trabajo en el resultado de los indicadores, los integrantes de la organización impulsarán la mejora de los resultados de los tableros de control sabiendo que dependen de su rendimiento personal.

Las siguientes son características de una buena gestión visual³⁶:

- **Simplicidad:** No sacrificar la facilidad de uso por la funcionalidad del sistema.
- **Apoyo:** Proporcionar el control sobre el sistema y suministrarle asistencia al usuario para facilitar la realización de las tareas.

³⁶Suárez Vicente, R. (2015). Gestión Visual - Aplicación de herramientas LEAN en el área de mantenimiento de una empresa minera.

- **Familiaridad:** Construir el producto según el conocimiento previo del usuario, lo que le permitirá progresar rápidamente.
- **Evidencia:** Hacer los objetos y sus controles visibles e intuitivos. Emplear siempre que se pueda representaciones del mundo real en la interfaz.
- **Estímulo:** Hacer las acciones previsibles y reversibles. Las acciones de los usuarios deberían producir los resultados que ellos esperan.
- **Satisfacción:** Evocar una sensación de progreso y logro en el usuario.
- **Disponibilidad:** Hacer todos los objetos disponibles de forma que el usuario pueda usarlos en cualquier secuencia y en cualquier momento.
- **Seguridad:** Evitarle errores al usuario proporcionándole diferentes tipos de ayuda bien sea de forma automática o a petición del propio usuario.
- **Versatilidad:** Soportar diversas técnicas de interacción, de forma que el usuario pueda seleccionar el método de interacción más apropiado para su situación.
- **Personalización:** Permitir a los usuarios adaptar la interfaz según sus necesidades.
- **Afinidad:** Permitir que los objetos sean afines a otros de la realidad cotidiana mediante un buen diseño visual.

La Inteligencia de negocios tiene mucha relación con la filosofía del BSC, ya que logra detallar muchos de sus indicadores y detectar dónde están los actores importantes que hacen que el indicador se cumpla o no. De hecho, la inteligencia de negocios es un complemento del BSC. A continuación, se muestran las relaciones que existen entre el BSC y los tableros de Inteligencia de Negocios:

1. Los tableros en BSC son tableros de Gestión, es decir, están más enfocados en reuniones ejecutivas y determinan si los indicadores se cumplieron; si no se cumplieron intentan exponer las causas y las acciones a seguir. Los tableros de inteligencia de negocios están más enfocados en el análisis de información y la recopilación de conocimiento.
2. El BSC define y alinea los indicadores clave de rendimiento (KPI) a partir de los objetivos estratégicos de la compañía. Por su parte, los tableros de inteligencia de

negocios se componen de indicadores y métricas que tienen como función apoyar el cumplimiento de los KPI. En pocas palabras, en inteligencia de negocios se desglosan los KPI del BSC.

3. El BSC vigila el desarrollo de la estrategia mensualmente, reportando el cumplimiento de los KPI, la causa y el efecto correspondiente. Además de definir planes de acción para el cumplimiento de los objetivos estratégicos. Mientras que los tableros de BI se enfocan en el detalle y la actualización puede llegar a ser en tiempo real.
4. Las aplicaciones en las cuales se soporta un BSC son transaccionales y de bajo costo, mientras que el tablero de BI se apoya en servidores de alto rendimiento y bases de datos especiales para realizar análisis sobre grandes volúmenes de información.

4.5.2. Introducción a la Minería De Datos

Las empresas han almacenado por años información que en su momento fue valiosa para su negocio pero que, con el paso del tiempo, se acumulará hasta perderse. Es importante para las empresas entender cuáles son sus datos más relevantes dentro de todo el conjunto de datos y que estos le sirvan para mejorar la forma en que toma decisiones. La minería de datos es el proceso que tiene como propósito descubrir, extraer y almacenar conocimiento de amplias bases de datos, a través de programas de búsqueda e identificación de patrones y relaciones globales, tendencias, desviaciones y otros indicadores aparentemente caóticos que tengan una explicación que pueda descubrirse mediante diversas técnicas de esta herramienta³⁷.

La minería de datos tiende a confundirse con la inteligencia de negocios. Aunque sean procesos distintos, los dos son complementarios; la gran diferencia es que la minería de datos es un análisis a futuro y resuelve preguntas como ¿Cuáles serán las sucursales más malas el próximo año y qué hacer para revertir evitar el suceso? ¿Cuáles serán las ventas del próximo

³⁷ANGELES LARRIETA, M. I., & SANTILLAN GOMÉZ, A. M. (s.f.). Minería de datos: Concepto, características, estructura y aplicaciones.

año? El proceso de Inteligencia de negocios se encarga de realizar un análisis del pasado y resuelve preguntas como ¿Cuáles fueron las ventas del día de ayer? ¿Cuáles son las mejores sucursales? Son distintos, pero como todos sabemos el futuro se predice a partir de los hechos que hayan sucedido en el pasado. La siguiente imagen muestra la evolución del análisis de datos, desde los reportes y cubos OLAP (Pasado - Inteligencia de Negocios) hasta la Minería de datos y la disposición en tiempo real.

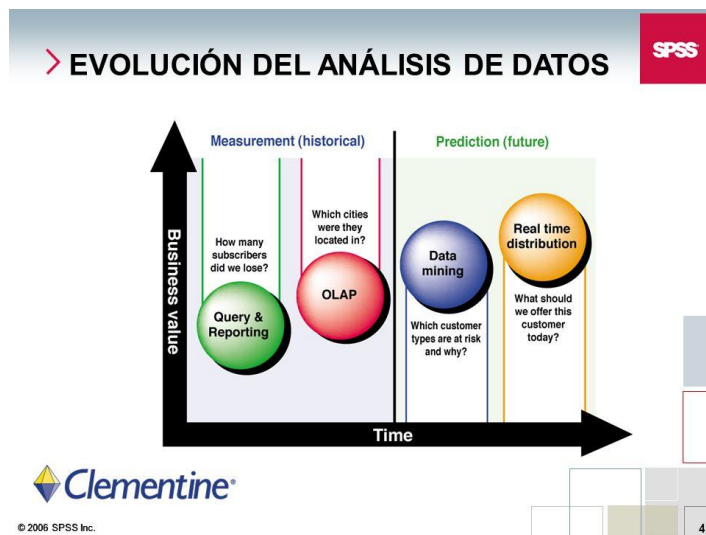


Gráfico 10: Evolución de la minería de datos

La minería de datos ha tenido bastante aplicación en el sector empresarial a continuación se listan algunas aplicaciones³⁸:

1. El FBI analiza bases de datos comerciales de cada individuo para detectar terroristas, basándose en si fuma, la talla y tipo de ropa que usa, la cantidad de arrestos, el salario, las revistas a las que está suscrito/a, la altura, el peso, las contribuciones que haya realizado a una iglesia o culto, los grupos políticos u organizaciones no gubernamentales a las que pertenece, las enfermedades crónicas que padece (diabetes, asma u otros), los libros que lee, los productos de supermercado que compra, si tomó clases de vuelo o si tiene cuentas de banco abiertas, entre otros.

³⁸VIRSEDA BENITO, F., & ROMÁN CARRILLO, J. (s.f.). Minería de datos y aplicaciones.

2. El sistema Falcon Fraud Manager es un sistema basado en redes neuronales que le permite a las instituciones financieras detectar fraudes teniendo en cuenta variables como: transacciones, propietarios de tarjetas y datos financieros entre otras. Falcón le ha ahorrado a las instituciones financieras más de seiscientos millones de dólares.
3. Una operadora móvil española creó un proyecto en el cual buscaba determinar el perfil de los clientes que se habían retirado de la compañía y con ellos determinar los clientes que aún permanecían y que eran propensos a retirarse de la compañía. Para ello se tuvieron en cuenta variables como: Estado Civil, edad, género, nacionalidad, morosidad, frecuencia y horario de servicio, entre otras. Como resultado del proyecto se recomendó a la operadora hacer un estudio de sus ofertas y analizar profundamente las incidencias recibidas por los clientes.

4.6. POLARIDAD

La polaridad es un valor que se le asigna a un término que expresa opinión³⁹, dependiendo del significado lingüístico de la palabra y se basa en metodologías de análisis de sentimiento o minería de opiniones. Estas metodologías permiten extraer comentarios de redes sociales como Twitter, Facebook, Youtube, entre otros y determinar su polaridad. El valor de la polaridad puede variar entre diferentes rangos, por ejemplo, Negativo (NG), Neutro (NE), Positivo (PO), estos rangos se pueden ampliar con dos rangos de más Muy Positivos (MP) o Muy Negativos (MN)

A través del tiempo se han realizado bastantes métodos para determinar la polaridad de los comentarios en redes sociales. Esto se debe a que actualmente, el valor de las opiniones de grandes expertos frente a un servicio o un producto se pierde y las opiniones de cientos de

³⁹Torres Samboni, L. A. (2015). Análisis de sentimientos sobre el posconflicto colombiano utilizando herramientas de minería de texto.

usuarios quienes tienen menos experiencia y conocimientos, los cuales juzgan los temas de acuerdo a su experiencia con el proveedor comienzan a tener mayor relevancia. A continuación, se muestran algunos ejemplos de los trabajos realizados alrededor de la polaridad en redes sociales.

4.6.1. Analizando opiniones de las redes sociales en sector turístico⁴⁰

Es un trabajo realizado por integrantes de la universidad de Alicante de España, el cual basó en determinar la polaridad del sector turístico de las localidades de la provincia de Alicante, para ello se utilizó la siguiente arquitectura (Gráfico 12):



Gráfico 11: Arquitectura para analizar redes sociales.

Básicamente la arquitectura se compone de:

- **Redes sociales:** Son fuentes de información donde las personas ponen sus comentarios, de estas se determinarán los atañen al ejercicio y se almacenarán en una base de datos.

⁴⁰Fernández, J., Llopis, F., Martínez, P., Gutiérrez, Y., & Díez, Á. (2016). Analizando opiniones en las redes sociales. *Procesamiento del Lenguaje Natural*, 141-148

- **Base de datos:** Son repositorios de información, se inicia con las **colecciones** de términos de interés para el usuario (hashtags, palabras o menciones de usuarios). Luego sigue un **repositorio** de los mensajes y comentarios extraídos de las diferentes fuentes y por última se dan los **índices** creados a partir de los comentarios y mensajes, Podríamos decir que en esta base de datos es donde esta almacenada la polaridad.
- **Módulo de escucha:** Este módulo se encarga de extraer la información de las redes sociales con base en la recolección de términos de interés. Por lo general se usa la API de las redes sociales, las cuales tienen algún tipo de restricción en cuanto al acceso de la información. Posteriormente, este módulo almacena dicha información en un repositorio.
- **Módulo de procesamiento:** Este módulo realiza la extracción de datos, detección de la localización (a través de una base de datos de lugares) y análisis de opiniones (polaridad y reputación) de los mensajes descargados mediante el módulo de escucha. Para determinar la polaridad se usan algoritmos basados en técnicas de aprendizaje automático apoyado en un listado de palabras positivas y negativas (también conocido como corpus); para determinar la reputación tiene en cuenta la distancia que haya tenido un comentario, es decir que, si se dispersó ampliamente por la red social, la polaridad de este comentario tendrá un peso mayor dentro de la reputación de la organización.
- **Módulo de presentación:** Este modelo se refiere a la interfaz del usuario; los usuarios podrán generar conocimiento de la información y ver los indicadores de esta.

4.6.2. Modelo léxico-sintáctico para la polaridad en redes sociales⁴¹

Este modelo está basado en algoritmos de calificación como Bayes Naive y máquina de soporte vectorial. Para el desarrollo del modelo se utilizó el corpus de la competencia SemEval 2014, el cual contiene un conjunto de tweets etiquetados con su correspondiente polaridad (Positiva, Negativa o neutra). La arquitectura del modelo se describe a continuación:

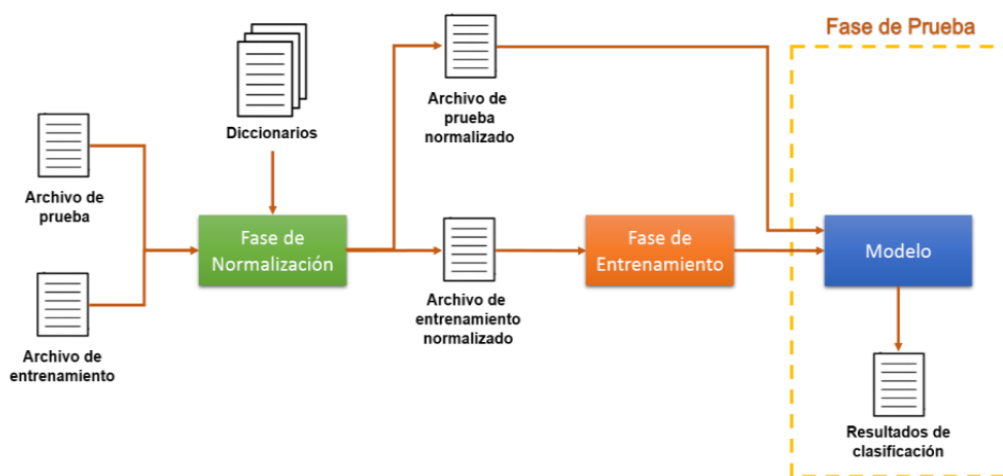


Gráfico 12: Arquitectura modelo léxico-sintético

- **Bases de datos:** Como insumo se tiene el conjunto de tweets con polaridad de la fuente SemEval 2014. Este conjunto de tweets se dividió en dos partes: un **archivo de entrenamiento** que se utilizó para que el modelo extraiga los patrones que determinan la polaridad de un tweet; el otro archivo, **archivo de prueba**, se usó para validar el modelo y determinar su precisión, por lo general el corpus usa el 70% para la construcción del modelo y el 30% para validar y calibrar el modelo. Por otro lado, se tiene un **diccionario** con el significado de cada emoción y de cada abreviatura usada en las redes sociales.

⁴¹Yuvila M, S., & Darnes Vilariño, M. (2015). Modelos para detectar la polaridad de los mensajes en. *Research in Computing Science*, 29-42

- **Modelo:** El modelo se divide en tres fases: La **fase de normalización** que se encarga de transformar todos los tweets a un mismo formato (minúsculas, reemplazar los emoticones y abreviaturas por su significado, y eliminar URL, hashtags y nombres de usuario). En este punto los archivos de prueba y entrenamiento estarán normalizados y listos para iniciar el modelo. La **fase de entrenamiento**, que extrae los patrones de clasificación del archivo de entrenamiento mediante los modelos de bayes naive y máquina de soporte vectorial. Por último, se realiza la **fase de prueba** del modelo, que busca determinar la polaridad de cada tweet del archivo de pruebas a través de los patrones descubiertos con el archivo de entrenamiento. Con esto el modelo determina una probabilidad de precisión. Con este modelo se pudo determinar una precisión del 56.58 %. Es decir que de cada 100 tweets el modelo clasificará correctamente la polaridad de 56 o 57.

5. DISEÑO METODOLÓGICO

Kdnuggets es una reconocida página web dedicada a la minería de datos, análisis de información, Big Data, y Data Science. Esta página realizó la siguiente pregunta a 200 Data Science: ¿Cuál es la principal metodología que usa en sus proyectos de Minería de datos? El 43% de las personas contestó que usaba la metodología CRISP-DM⁴². Dicho resultado permitió que el presente trabajo se realizara a través de esta metodología.

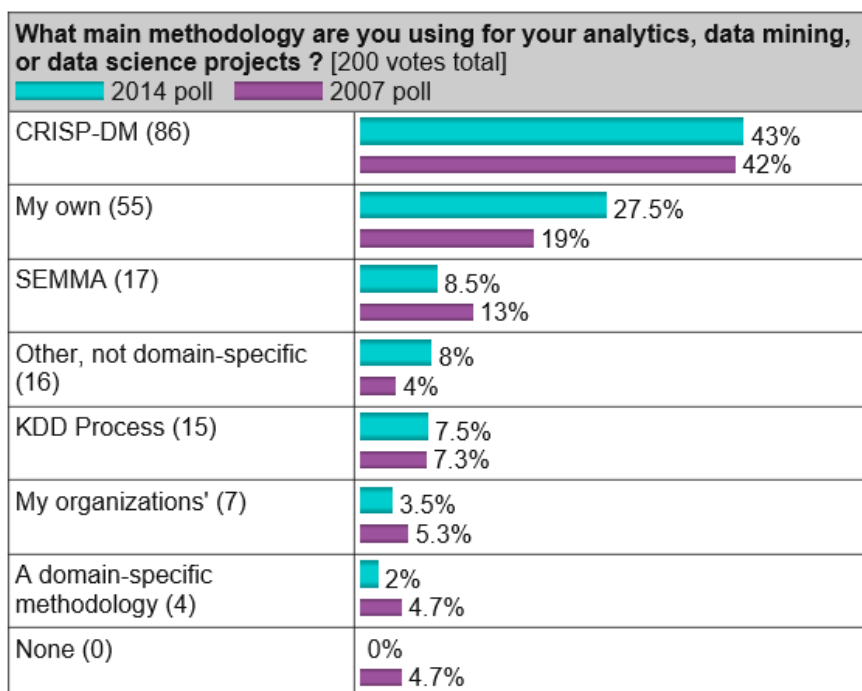


Gráfico 13: Comparación de metodologías para proyectos de minería de datos.

CRISP-DM (Cross Industry Standard Process for Data Mining) es una metodología de Minería de datos de libre distribución, fue creada en el año 1999 por un consorcio de empresas europeas como NCR (Dinamarca), AG(Alemania), SPSS (Inglaterra), OHRA (Holanda), Teradata, y Daimler-Chrysler. Esta metodología se compone de seis fases⁴³ las cuales se llevan a cabo en esta sección

⁴²KDNUGGETS. (2014). Recuperado el 20 de 07 de 2016, de <http://www.kdnuggets.com/polls/2014/analytics-data-mining-data-science-methodology.html>

⁴³SPSS. (1999). *CRISP-DM 1.0 Step by step Data Mining Guide*.

5.1. ENTENDIMIENTO DEL NEGOCIO

Lo primero que se debe realizar en cualquier modelo de minería de datos es entender el negocio, se recomienda crear un bosquejo del resultado final ya que este tipo de modelos es consumido por medio de un tablero de control. Este paso es muy importante porque define el alcance del modelo y los entregables a realizar. Para construir un tablero útil de indicadores relevantes para las aseguradoras se revisó la bibliografía relacionada con las redes sociales y su comportamiento, para así determinar los indicadores que logran complementar la polaridad de los comentarios en redes sociales. Lo más relevante que se ha encontrado es la relación que estas tienen con la teoría de grafos, como lo dice Carlos Lozares en su artículo denominado “la teoría de redes sociales”. “... Al tratar los orígenes de las redes sociales la teoría de grafos ha sido de uso habitual. Esta teoría proporciona una representación apropiada y un conjunto de conceptos para estudiar.”⁴⁴ Básicamente, un grafo es una serie de nodos que se conectan entre sí mediante aristas o líneas. En este documento los nodos en redes sociales son los usuarios y las aristas es la relación que se crea entre estos mediante los comentarios, respuestas o publicaciones, ver gráfico 13.



Gráfico 14: Red social como un grafo.

5.1.1. Métricas e indicadores propuestos

Ahora, teniendo clara la relación que existe entre la teoría de grafos y las redes sociales se lograron extrapolar algunas propiedades del grafo a métricas de redes sociales, las cuales

⁴⁴Lozares, C. (1996). La teoría de redes sociales. *Universidad autónoma de barcelona*, 103-126.

apoyarán posteriormente la construcción de los indicadores. La relación entre las propiedades de los grafos y las métricas se muestran a continuación.

| GRAFOS | | | MÉTRICAS RED SOCIAL | |
|--------|-------------------------------|---|-------------------------------------|---|
| # | Propiedad del grafo | Descripción de la propiedad | Nombre de la métrica | Descripción |
| 1 | Tamaño del grafo | Número de Nodos que contiene la red | Tamaño de la red social | Cantidad de usuarios que realizan por lo menos un comentario sobre alguna aseguradora o que sean amigos o seguidores de una alguna de ellas |
| 2 | Grado de los nodos | Grado del nodo. Número de aristas que salen de nodo. | Publicaciones | Cantidad de post y tweets que realice una aseguradora |
| 3 | Grado de los nodos | Grado del nodo. Número de aristas que salen de nodo. | Atención | Cantidad de respuestas a quejas que realicen los usuarios por medio de las redes sociales |
| 4 | Peso de cada arista | El peso que tiene una arista en un grafo. La suma del peso de las aristas de un nodo dará como resultado el peso del nodo | Interacción | Cantidad de me gusta + cantidad de comentarios + cantidad de veces que se ha compartido un post + cantidad de retweets |
| 5 | Centralidad de intermediación | Son nodos que están en medio de dos grupos de nodos que sin él no se lograrían comunicar | Posibles Intermediarios Potenciales | Cantidad de amigos o seguidores de cada uno de los seguidores o amigos de cada compañía de seguros. |

Tabla 6: Propiedades de grafos relacionados con las redes sociales

En el anexo 1 se puede ver la ficha de cada uno de los indicadores que nacen a partir de las métricas extrapoladas de la teoría de grafos.

Ahora, para entender cómo se relacionan estos indicadores para lograr mejorar la utilidad de la compañía, se utiliza la filosofía del Balance ScoreCard; se propone construir dos perspectivas: la financiera y la de clientes. En la perspectiva de clientes se proponen 3 objetivos principales: clientes satisfechos, fidelización del cliente y crecimiento de cliente. En la perspectiva financiera se proponen cuatro objetivos: obtener un resultado técnico positivo o un índice combinado menor al 100%; aumentar de las primas o ingresos; optimizar los costos; y optimizar los gastos. Todos estos objetivos se relacionan de la siguiente forma: Si los clientes están satisfechos, se mejora su fidelización, los cuales ayudarán a transmitir las buenas experiencias a otros clientes potenciales, y esto hará que la cantidad de clientes de la aseguradora crezca constantemente. Si se logra crecer en clientes objetivo se logra un

aumento en las primas, ayudando a mejorar la utilidad de la compañía, esto siempre y cuando se haga una optimización adecuada de los costos y de los gastos. Ver gráfico 14.

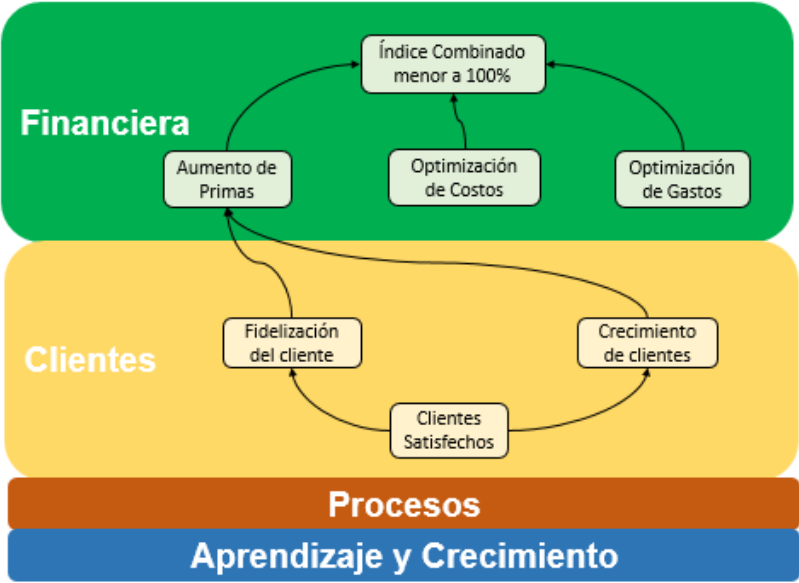


Gráfico 15: BSC del sector asegurador

Para que los objetivos se cumplan se asignan indicadores los cuales se presentan a continuación: los costos están divididos en el dinero que se destina para la atención de accidentes y enfermedades laborales; y la comisión a los intermediarios. Es por esto que el índice de siniestralidad y el índice de comisión son medidos. Para el caso de los gastos, estos se dividen en gastos de personal (Salarios y bienestar de los empleados) y gastos administrativos (Arriendos, licencias, servicios públicos etc), por esta razón, el índice de gastos de personal e índice de gastos administrativos son medidos. Los indicadores se muestran en el anexo 1 y la relación con los objetivos financieros en el grafico 15.

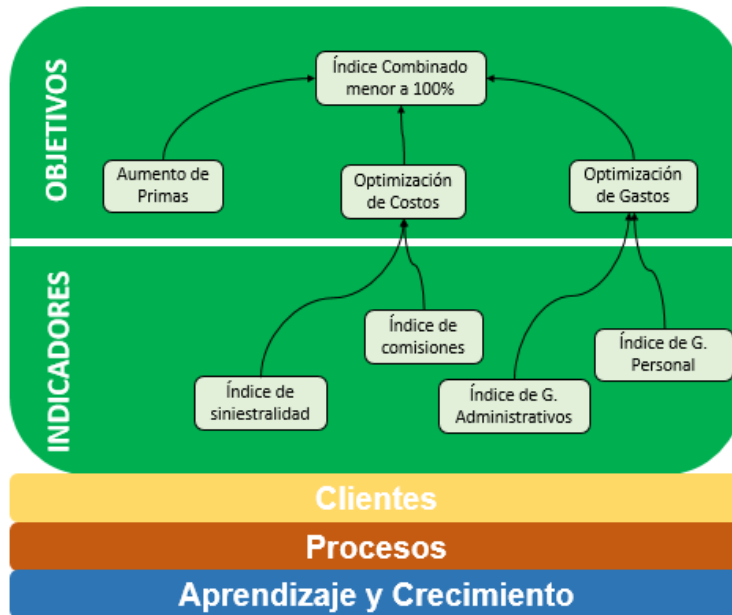


Gráfico 16: Perspectiva financiera del sector asegurador.

El aumento de las primas tiene una relación directa con el crecimiento de los clientes objetivo y la fidelización o mantenimiento de los clientes de la aseguradora; para esto se proponen los indicadores relacionados en la figura 16. Si el tamaño de la red social crece, las atenciones por este medio crecerán, si estas son oportunas y de buena calidad se obtendrán reacciones positivas de los clientes y esto implicará mejoras en la satisfacción del cliente. Por otro lado, si tenemos intermediarios que compartan las publicaciones o los comentarios, se generará más interacción entre los clientes; si el indicador de reacciones positivas está cumpliendo la meta, se podrán obtener nuevos clientes objetivos.

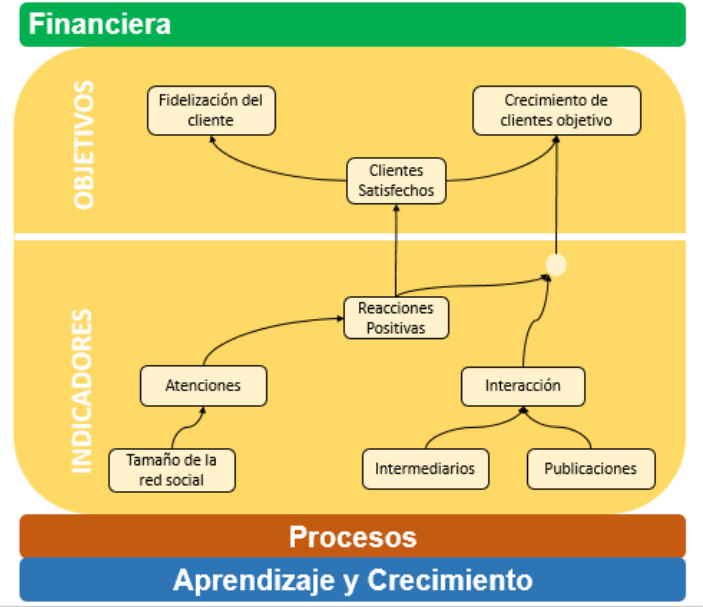


Gráfico 17: Perspectiva cliente del sector asegurador.

Este trabajo implementa algunos de los indicadores propuestos y algunas métricas que se pueden extraer de las redes sociales en un tablero de control. Los siguientes datos son la descripción de los indicadores, sus metas, y las métricas a desarrollar.

| Tip o | Perspectiv a | # | Nombre | Descripción | Meta |
|--------------------|-------------------|-----------------|----------------------------------|--|---|
| INDICADORES | FINANCIERA | 1 | Índice Combinado | Porcentaje de los ingresos que se usan para atender el core del negocio | 106% |
| | | 2 | Índice Siniestralidad | Porcentaje de los ingresos que se destinan para atender los costos de los accidentes | 63% |
| | | 3 | Índice de Comisiones | Porcentaje de los ingresos que se entregan a los intermediarios | 10% |
| | | 4 | Índice de Personal | Porcentaje de los ingresos que se destinan a los salarios y bienestar de los empleados | 9% |
| | | 5 | Índice de Gastos Administrativos | Porcentaje de los ingresos que usan para los gastos de la compañía (Servicios públicos, licencias tecnológicas etc.) | 18% |
| | | Clientes | 6 | Reacciones | Crecimiento de reacciones positivas en redes sociales |
| Métricas | Clientes | 7 | Publicaciones | Cantidad de post y tweets realizados por cada aseguradora | NA |
| | | 8 | Interacciones | Cantidad de Me gusta + Cantidad de Comentarios + Cantidad de veces compartido + Cantidad de retweets | NA |

Tabla 7: Indicadores y métricas a desarrollar

La meta de cada indicador financiero se calculó con el promedio de los resultados de los últimos 6 años menos una desviación estándar ($\mu - \sigma$). En la tabla 7 se muestra el análisis realizado.

| Año | Ind. Siniestralidad | Ind. Intermediarios | Ind. Personal | Gasto | Ind. Administrativos | Gastos | Ind. Combinado |
|----------------------------|---------------------|---------------------|---------------|-------|----------------------|--------|----------------|
| 2011 | 66% | 13% | 11% | | 20% | | 115% |
| 2012 | 65% | 14% | 10% | | 18% | | 113% |
| 2013 | 63% | 14% | 10% | | 18% | | 108% |
| 2014 | 64% | 13% | 9% | | 18% | | 109% |
| 2015 | 62% | 9% | 9% | | 19% | | 104% |
| 2016 | 67% | 11% | 10% | | 21% | | 116% |
| Promedio | 65% | 12% | 10% | | 19% | | 111% |
| Desviación estándar | 2% | 2% | 1% | | 1% | | 5% |
| Meta | 63% | 10% | 9% | | 18% | | 106% |

Tabla 8: Metas de cada indicador financiero

La escala de colores que se plantea para cada uno de los indicadores es extraída estadísticamente con base en el promedio (μ) de los seis años y la desviación estándar (σ), en la tabla 8 se puede observar la escala para cada color. Como se puede observar en la tabla 9, esta escala se aplica directamente a cada indicador y se obtienen los rangos para cada uno.

| Límite entre | Estadísticamente | Descripción |
|--------------------|------------------|---|
| Rojo y Naranja | $\mu + \sigma$ | El promedio más una desviación estándar |
| Naranja y Amarillo | μ | El promedio |
| Amarillo y Verde | $\mu - \sigma$ | El promedio menos una desviación estándar |

Tabla 9: Metas de cada indicador financiero

| Color | Índice de siniestralidad | de Índice intermediarios | de Índice de gastos de personal | Índice de Gastos Administrativos | Índice Combinado |
|----------|--------------------------|--------------------------|---------------------------------|----------------------------------|------------------------|
| Rojo | $x \geq 67\%$ | $x \geq 14\%$ | $x \geq 11\%$ | $x \geq 20\%$ | $x \geq 116\%$ |
| Naranja | $65\% \leq x < 67\%$ | $12\% \leq x < 14\%$ | $10\% \leq x < 11\%$ | $19\% \leq x < 20\%$ | $111\% \leq x < 116\%$ |
| Amarillo | $63\% \leq x < 65\%$ | $10\% \leq x < 12\%$ | $9\% \leq x < 10\%$ | $18\% \leq x < 19\%$ | $106\% \leq x < 111\%$ |
| Verde | $x < 63\%$ | $x < 10\%$ | $x < 9\%$ | $x < 18\%$ | $x < 106\%$ |

Tabla 10: Rangos para cada indicador financiero

La meta del indicador de reacciones de clientes se basó en la escala Likert que cuenta con un rango de 5 categorías, las cuales se adecuaron para calcular/identificar/encontrar la polaridad en las redes sociales. En la tabla 10 se puede observar la escala Likert con la correspondiente categoría de polaridad.

| Escala Likert | Polaridad | Rango | Color |
|-----------------------------|--------------|----------------------|----------|
| Totalmente de acuerdo | Muy Positivo | $x \geq 80\%$ | Verde |
| De acuerdo | Positivo | $60\% \leq x < 80\%$ | Amarillo |
| Ni de acuerdo ni desacuerdo | Neutral | $40\% \leq x < 60\%$ | Gris |
| En desacuerdo | Negativo | $20\% \leq x < 40\%$ | Naranja |
| Totalmente en desacuerdo | Muy Negativo | $x < 20\%$ | Rojo |

Tabla 11: Rangos del indicador de clientes

5.1.2. Tableros de control propuestos

Finalmente, luego de entender el negocio y tener claros los indicadores, se realiza un bosquejo del tablero de control que se construirá a partir de los datos.

El grafico 17 muestra el tablero de control financiero, el cual contiene los indicadores de índice combinado, índice de siniestralidad, índice de comisión, índice de gastos de personal e índice de gastos administrativos, los cuales se podrán filtrar por compañía, año, mes y tipo de seguro.

DASHBOARD FINANCIERO

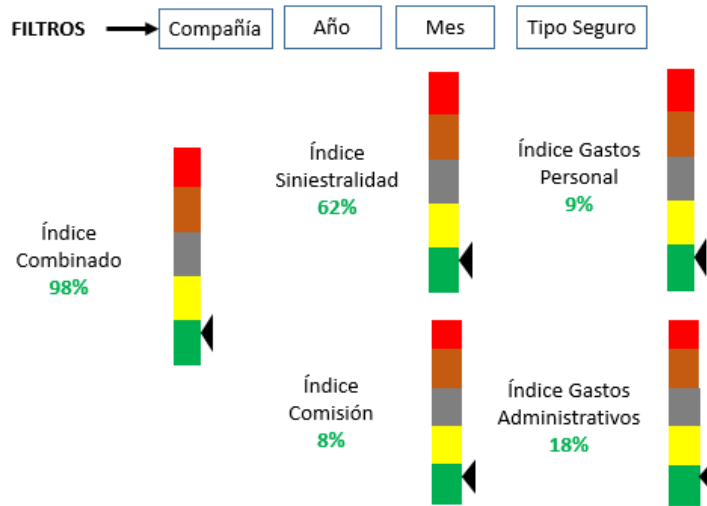


Gráfico 18: Dashboard Financiero

El gráfico 18 muestra el tablero de control del cliente que posee el indicador de reacciones. Este tablero se compone del indicador con los rangos correspondientes, también muestra/expone/señala el mapa de Colombia en el cual se podrá ver el indicador de reacciones por cada uno de los departamentos y las métricas tanto de las publicaciones como las interacciones. Este tablero se podrá filtrar por compañía, año, red social y país.

DASHBOARD CLIENTE

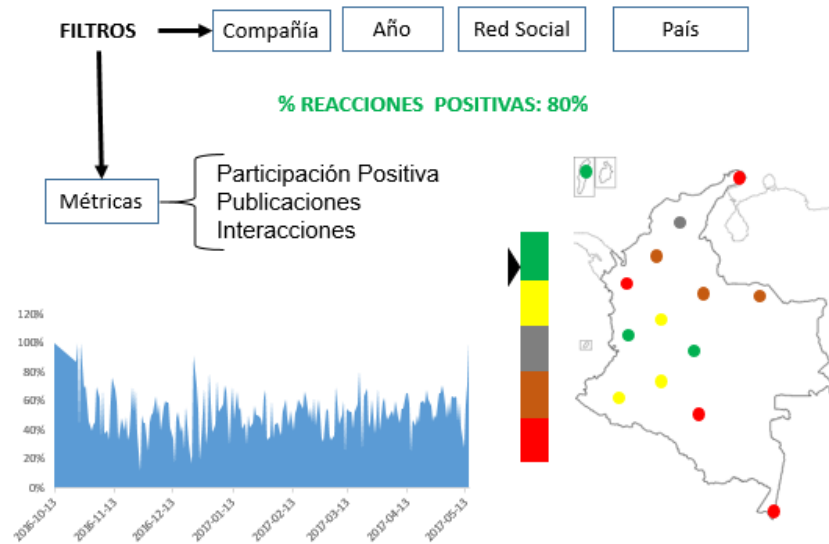


Gráfico 19: Dashboard Cliente

5.2. ENTENDIMIENTO DE LOS DATOS

Luego de tener claro el negocio y lo que se debe entregar al finalizar el ejercicio se deben entender los datos, fuentes de extracción, restricciones y estado de la información.

5.2.1. Aseguradoras postuladas y sus redes sociales.

Lo primero que se realizó fue determinar un conjunto de compañías aseguradoras con las que se realizará el ejercicio, para ello se revisaron las empresas afiliadas a Fasecolda (Federación de aseguradores colombianos, en total 35 empresas), luego se validó que tuvieran Twitter o Facebook, que los comentarios estuvieran en español y que no tuvieran más de dos productos sobre la misma cuenta, por ejemplo BBVA tiene en una sola cuenta de Twitter que incluye la parte financiera y la parte de seguros, haciendo difícil saber qué comentarios corresponden al banco BBVA y qué comentarios corresponden a la aseguradora BBVA. A continuación, se muestran las compañías de seguros postuladas que poseen Twitter o Facebook y su correspondiente cuenta.

| COMPAÑÍA ASEGURADORA | TWITTER | FACEBOOK |
|----------------------|----------------------------------|---------------------------|
| ALLIANZ | @Allianz es | AllianzColombia |
| AXA COLPATRIA | @AxaColpatria | AxaColpatria |
| BERKLEY | @WRBerkleyCorp | NA |
| BOLIVAR | @SegurosBolivar | SegurosBolivar |
| CHUBB | @Chubb Seguros | NA |
| COFACE | @CofaceCol | NA |
| COLMENA | @ColmenaSeguros | ColmenaSeguros |
| EQUIDAD | @SegurosEquidad | LaEquidadSegurosOC |
| GENERALI | @GENERALI es | GENERALIes |
| LIBERTY | @LibertySegCol | LibertySegurosCO |
| MAPFRE | @Mapfre co | Mapfreco |
| METLIFE | @MetLife | MetLifeColombia |
| MUNDIAL | @SegurosMundial | SegurosMundial |
| OLD MUTUAL | @OldMutualSA | OldMutualColombia |
| PANAMERICAN | @PanAmericanLife | NA |
| POSITIVA | @PositivaCol | PositivaCompaniaDeSeguros |
| PREVISORA | @SomosPrevisora | Previsora.Seguros |
| QBE | @QBEColombia | QBEColombia |
| RSA | @rsagroup | NA |
| SEGUREXPO | @Segurexpo | SegurexpodeColombiaS.A |
| SOLIDARIA | @SolidariaSeguro | NA |
| SOLUNION | @SolunionSeguros | NA |
| SURAMERICANA | @SegurosSURA | SegurosSURAColombia |
| ZURICH | @Zurich | NA |

Tabla 12: Aseguradoras y sus redes sociales

5.2.2. Fuentes de información y campos.

Para el ejercicio se usaron tres fuentes de información: los indicadores de gestión⁴⁵ de Fasecolda, en donde reposan las cifras financieras de cada aseguradora en archivos planos de Excel los cuales son de libre acceso; Twitter; y Facebook. De estas dos fuentes se extraerán los comentarios, las publicaciones y las interacciones que los usuarios hacen sobre las aseguradoras.

5.2.2.1. Twitter

Twitter posee una API (Interfaz de Programación de Aplicaciones, por sus siglas en inglés) que permite extraer la información publicada en la red social⁴⁶. Esta API es de uso gratuito, sin embargo, para lograr extraer la información hay que tener conocimientos de programación y bases de datos. Para este caso se dividió la extracción en dos partes: publicaciones de la aseguradora y comentarios de los usuarios. Los campos usados para extraer las publicaciones de una aseguradora son los siguientes:

Id: Identificación única del tweet

Text: Texto de la publicación del tweet

Created_at: Fecha en que se creó cada tweet

Favorite_count: Cantidad de “me gusta” sobre el tweet

Retweet_count: Cantidad de veces que se comparte el tweet

screen_name: Alias de la aseguradora que realizó el tweet

Los campos usados para extraer los comentarios de los usuarios son los siguientes:

Id: Identificación única del comentario

Text: texto del comentario

Created_at: Fecha en que se creó el comentario

⁴⁵Fasecolda. (15 de 05 de 2017). *Fasecolda*. Obtenido de Federación de Aseguradores Colombianos: <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>

⁴⁶Twitter. (2017). *Twitter*. Obtenido de Twitter: <https://dev.Twitter.com/>

screen_name: Alias de la persona que realizó el comentario

user_location: Localización de nacimiento del usuario

geo.coordinates[0]: Longitud de la localización desde donde se emite el comentario

geo.coordinates[1]: Latitud de la localización desde donde se emite el comentario

screen_name: Alias de la persona que realizó el tweet

In_reply_to_screen_name: Alias de la persona que realizó el comentario original en caso de ser un comentario compartido

5.2.2.2. Facebook

Facebook al igual que Twitter posee una API para extraer la información de la red social la cual es usada en este proyecto⁴⁷. Para Facebook los campos utilizados para extraer la información.

Id: Id del Post

Message: Mensaje incluido en el post

created_time: Fecha en la que se creó el post

likes: Likes del post

id_comentario: Id del comentario que se le realiza al post

comments: Texto del comentario que se realizan sobre el post

created_time_comentario: Fecha en que se crea el comentario

4.1.1.1. Indicadores de gestión

Los indicadores de gestión es un informe que publica mensualmente Fasecolda con la información financiera de cada compañía de seguros. De este informe se puede obtener

⁴⁷Facebook. (2017). *Developer Facebook*. Obtenido de Developer Facebook: <https://developers.facebook.com/>

información como la utilidad de cada compañía, los ingresos y los costos. Esta información se usa para realizar el tablero de control financiero. Los campos que se extraen de este archivo son los siguientes:

Año: Año en el cual se publican las cifras

Mes: Mes en el cual se publican las cifras

Tipo de seguro: Generales o Vida

Compañía: Cada compañía de seguros.

Primas devengadas: Ingresos de cada compañía por concepto de primas.

Siniestros incurridos: Dinero reservado para la atención de accidentes de trabajo o enfermedades laborales.

Costos Netos de intermediación: Valor que se le reconoce a los intermediarios por traer el negocio.

Otros Costos de ingresos de seguros: Otros costos no incluidos en los anteriores rubros.

Gastos de personal: El dinero destinado a los salarios y el bienestar de los empleados.

Gastos administrativos: El dinero destinado a los gastos de la compañía (Servicios públicos, licencias tecnológicas, etc)

Inversiones: Ingresos de cada compañía por concepto de inversiones.

4.1.2. Estado de la información y restricciones

Luego de revisar cada uno de los campos y extraer una muestra de la información se encontró lo siguiente:

- La periodicidad con que se tiene la información de Twitter y Facebook es diaria, sin embargo, la de Fasecolda es mensual con un rezago de 2Meses, esto hace que el tablero de control no tenga el mismo periodo de actualización en las dos perspectivas

- La API de Twitter permite una conexión durante 15 minutos⁴⁸ y lo que se podrá extraer de información serán uno o dos meses de antigüedad según el grado de movimiento de la cuenta, es decir, no se puede extraer información tan antigua, la misma se debe ir construyendo de a pedazos.
- Los tweets encontrados son de máximo 140 caracteres, de esta forma se encontrarán tweets muy concretos lo que permite hacer un modelo de polaridad mucho más eficiente.
- Para publicar los tweets en un mapa hay que tener la georreferenciación del tweet, sin embargo, la cantidad de tweets con georreferenciación son menos del 1%. De esta forma es necesario usar otro método para recolectar la ubicación del tweet. En el caso de Facebook la API tiene restricciones para acceder a la localización de los usuarios, básicamente solo deja consultar las personas que sean amigos del dueño de la cuenta autorizada para usar la API.
- Para el problema de Twitter se determina tomar la localización de nacimiento de usuario y no por la localización de donde se está emitiendo el tweet, esto se realiza mediante el campo “user_location” de Twitter el cual posee en un 78% la localización. Sin embargo, es importante comentar que este campo es un campo de texto abierto, es decir los usuarios de Twitter pueden poner el nombre de su ubicación de cualquier forma, aunque Twitter tiene un autocompletar a la hora de crear la localización. En la siguiente tabla se muestra un ejemplo de como vienen los datos de localización desde Twitter.

⁴⁸TWITTER. (2017). *Twitter Developer Documentation*. Obtenido de Twitter Developer Documentation: <https://dev.twitter.com/rest/public/rate-limiting>

| Localización | user_location | Cantidad de tweets |
|--------------|--------------------------------|--------------------|
| Bogotá | Bogotá, D.C., Colombia | 540 |
| Bogotá | Bogotá, Colombia | 352 |
| Bogotá | Bogotá-Colombia | 20 |
| Bogotá | Bogotá, D.C. | 12 |
| Bogotá | Bogotá, DC, Colombia | 11 |
| Bogotá | Bogotá. Colombia | 6 |
| Bogotá | Bogotá, Colombia. | 4 |
| Bogotá | Bogotá, Col | 1 |
| Bogotá | Bogotá, Colombia, Suramérica | 1 |
| Bogotá | Bogotá, La Atenas Suramericana | 1 |
| Bogotá | Bogotá. La felicidad existe... | 1 |

Tabla 13: Problemas de calidad en la localización del usuario

- Para la información de Fasecolda es necesario homologar los nombres de las aseguradoras pues no en todos los años las compañías tenían los mismos nombres, por ejemplo, Colpatria y Axa Colpatria. También existe el caso en que las compañías dividen los productos por ejemplo Sura ARL es la parte de la empresa suramericana destinada netamente a ARL. Pero finalmente pertenece a suramericana.
- Para realizar un modelo de polaridad es importante obtener un diccionario de palabras positivas y negativas, las cuales determinarán la polaridad de un comentario. En este documento se recomienda un listado de palabras determinado por Ismael Díaz Rangel, Grigori Sidorov y Sergio Suárez-Guerra en su artículo “*creación y evaluación de un diccionario marcado con emociones y ponderado para el español*”⁴⁹ el cual es de libre acceso y contienen aproximadamente 2036 diferentes palabras. Es importante leer el artículo completo para comprender la metodología que se usó para crear el diccionario de palabras. Otro método y el que se utilizó fue crear nuestro propio diccionario de

⁴⁹Díaz Rangel, I., Sidorov, G., & Suárez-Guerra, s. (2014). Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. *instituto politécnico nacional*

palabras a partir de una cantidad de tweets. El proceso fue escoger 2.708 y determinar manualmente la polaridad de cada tweet con esto se obtiene dos cosas, la primera construir el diccionario y la segunda obtener una población con la verdadera polaridad de tweets para construir el modelo.

4.2.PREPARACIÓN DE LOS DATOS

Luego de conocer los datos, sus restricciones y dificultades se debe revisar cómo hacer frente a estas dificultades. Para este fin se revisaron métodos para su solución y herramientas de extracción de información, de almacenamiento y de Inteligencia de Negocios

4.2.1. Normalización e Integración de datos

- Los tableros de control se actualizarán de diferentes formas, el tablero financiero se actualizará cada mes mientras que el tablero de clientes se actualizará diariamente. Por otro lado, el tablero financiero tendrá información histórica desde el año 2010, mientras que (debido a la imposibilidad de adquirir información histórica) en las redes sociales tendrán información desde octubre de 2016.
- El API de twitter permite solo conectarse por 15 minutos a la aplicación, esto hace que solo se pueda recolectar información desde octubre de 2016. De esta forma y para que la información tenga consistencia la información en Facebook también se tendrá en cuenta a partir de esta misma fecha.
- Los tweets tienen un límite de 140 caracteres, mientras que los comentarios en Facebook pueden ser de más de 5000 caracteres. De esta forma el modelo necesitará un mayor tiempo de procesamiento para determinar la polaridad de los comentarios en Facebook.
- Para lograr homologar la localización de nacimiento del usuario en Twitter Se usan métodos de limpieza de datos. En el caso del nombre de las compañías de seguros se realiza una homologación manual ya que son 35 aseguradoras por cada uno de los meses de cada año. Este procedimiento se tendrá que hacer cada vez que se actualice el tablero financiero.

- Aunque se recomienda utilizar un diccionario de palabras ya construido, en este proyecto se decidió crear un propio diccionario de palabras a partir de una cantidad de tweets. El proceso fue escoger 2.708 comentarios y determinar manualmente la polaridad de cada uno, con esto se obtuvieron dos cosas, la primera el diccionario como tal y la segunda una población con su polaridad que servirá para entrenar el modelo.

4.2.2. Herramientas De Analítica

Para realizar todo el proceso de inteligencia de negocios y de minería de datos existen diferentes herramientas tecnológicas que ayudarán a transformar los datos en información y estos en conocimiento. A continuación, se muestran estas herramientas y las razones por las cuales se escogieron.

- **PYTHON:** Para la extracción de información de redes sociales se usó Python debido a que ocupa el segundo lugar dentro de las 85 herramientas usadas por los científicos de datos, es libre y además tiene un crecimiento significativo entre los años 2010 y 2014 según la encuesta de Kdnuggets. En la siguiente tabla se muestra las herramienta más usada por los científicos de datos según Kdnuggets⁵⁰.

⁵⁰KDNUGGETS. (2016). *KDNUGGETS*. Obtenido de KDNUGGETS: <http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html/2>

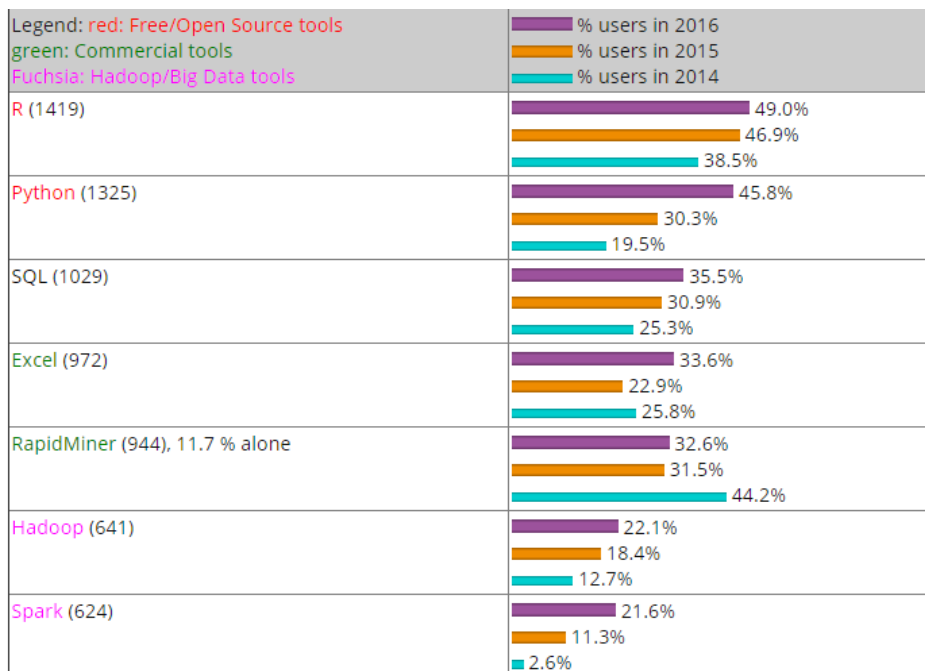


Gráfico 20: herramientas más usadas por los científicos de datos

Para la selección de herramientas de inteligencia de negocios se usó el cuadrante mágico de Gartner, el cual se ha vuelto un referente para quienes desean adquirir herramientas tecnológicas. En el gráfico 19 se puede ver el cuadrante de las empresas que tienen herramientas de inteligencia de negocios, el cual las mide desde dos puntos de vista principales, la visión que tenga a futuro la empresa (Eje x) y la facilidad de uso de la herramienta (Eje y). Así mismo el cuadrante se divide en 4 cuadrantes, en la esquina superior derecha se ubican las empresas que lideran el mercado, para el año 2017 solo existen 3, Microsoft, Tableau y Qlik. En la esquina inferior derecha se encuentran los visionarios, son empresas que poseen mucha visión del mercado pero que tienen dificultades con el uso de las herramientas por parte de los usuarios finales y los vendedores, entre estas empresas están SAP, IBM y Salesforce. En la esquina superior izquierda están las empresas que tienen herramientas fáciles de usar pero que deben mejorar la visión de sus productos si quieren ser líderes, para el informe del 2017 no existe ninguna empresa en este cuadrante. Y en la esquina inferior izquierda se tienen las empresas de nicho, es decir empresas que deben mejorar su visión y su facilidad de uso, entre ellas están Oracle y Pentaho.



Gráfico 21: Cuadrante de gartner BI

- PENTAHO:** Para la transformación e integración de la información y la construcción de los indicadores, se utilizó la herramienta libre de inteligencia de negocios Pentaho en su versión más estable 5.0. Aunque su puntuación en el cuadrante de Gartner no es la mejor debido a su debilidad en la facilidad de uso por parte de los usuarios finales, fue escogida por ser una de las herramientas de BI que enfocan su esfuerzo en la transformación de datos a través de la analítica avanzada mediante su producto estrella, Pentaho Data Integration (PDI)⁵¹. Es decir, su fuerte está en la transformación de los datos en conocimiento.

Con esta herramienta se seleccionaron los comentarios que estuvieran escritos en español, se descartó BBVA debido a que tiene los servicios financieros y de seguros dentro de la misma cuenta. Se escogieron los campos necesarios como la localización, el texto, la

⁵¹Gartner. (Febrero de 2017). *Magic Quadrant for Business Intelligence and Analytics Platforms*. Obtenido de Gartner: <https://www.gartner.com/doc/reprints?id=1-3TYE0CD&ct=170221&st=sb>

fecha de creación y la compañía de seguros. Este conjunto de programas libres también se usó para normalizar los departamentos, ciudades y países.

- **QLIK SENSE:** Pentaho es fuerte en la transformación de los datos, pero débil en el uso de la información por parte de los usuarios finales, básicamente porque es poco intuitiva a la hora de crear un cuadro de mando y en varias ocasiones se necesita de ingenieros que conozcan de sentencias SQL; es por esto que para la construcción y visualización de los cuadros de mando se usará la herramienta QlikSense de la empresa Qlik ubicada en el cuadrante de líderes y usada en su versión escritorio.
- **PostgreSQL:** Para el almacenamiento de la información se usó la base de datos PostgreSQL en la cual se creó el modelo de datos para almacenar los comentarios, el diccionario de palabras, los departamentos normalizados y el modelo como tal.

4.2.3. Arquitectura

Luego de determinar las herramientas con las cuales se realizará el modelo, se construye la arquitectura para el desarrollo, la cual se puede ver en la grafica 20

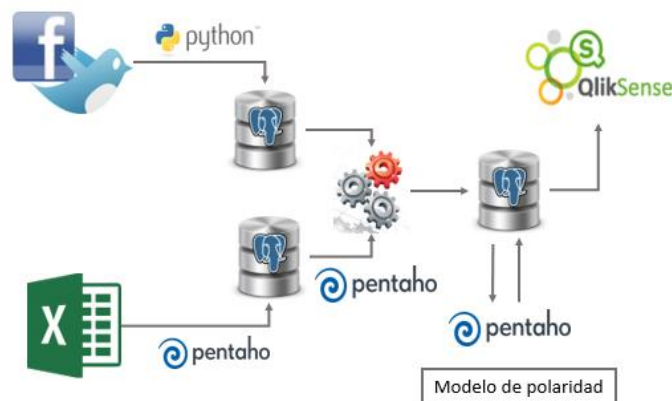


Gráfico 22: Cuadrante de gartner BI.

Las fuentes de datos son Facebook, Twitter y el informe de indicadores de gestión; por medio de python se extrae la informacion de redes sociales y se almacenan en una base de datos postgres, asi mismo se usa pentaho para llevar la informacion del excel de indicadores de gestion a una base de datos postgres. Posteriormente se integra la informacion con pentaho y se unifica en una sola fuente de datos de los cuales se realizará el modelo de polaridad y se crearán los indicadores y metricas. Luego se usara la aplicación de Qlik Sense para crear los tableros de control financiero y de clientes.

4.3. MODELO

Para el modelo de polaridad se utilizó el proceso descrito en el siguiente gráfico:

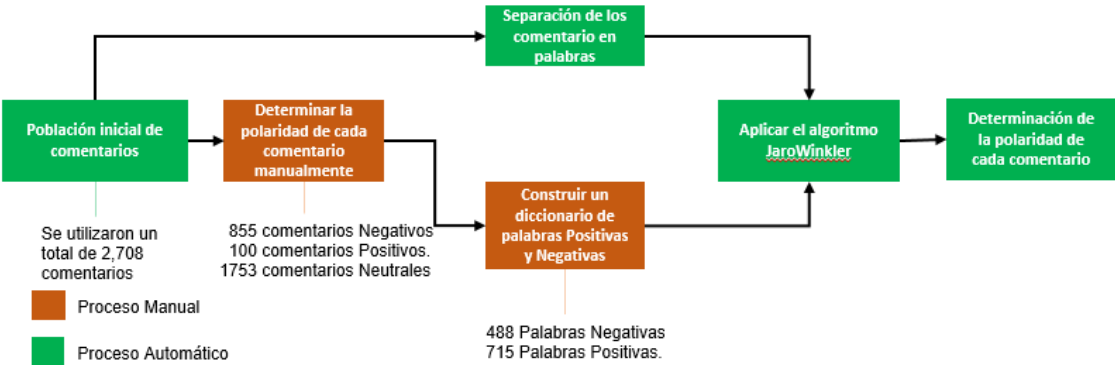


Gráfico 23: Proceso de modelado

Primero se recolectaron 2.708 comentarios de las diferentes aseguradoras, con los cuales se construyó el modelo, de ellos a todos se les determino manualmente la polaridad (Positiva, Neutra o Negativa). En total fueron 2708 comentarios de los cuales 100 fueron positivos, 855 Negativos y 1753 Neutrales. De este conjunto de comentarios se determinó el diccionario de palabras positivas o negativas que determinaban su polaridad. El diccionario se conforma por 488 palabras negativas y 715 palabras positivas. Luego, con ayuda de pentaho y lenguaje SQL, se separan cada uno de los comentarios en palabras, para compáralas posteriormente con el diccionario construido mediante el algoritmo JaroWinkler. Este algoritmo determina

la probabilidad de coincidencia entre dos palabras, es decir un numero entre 0 y 1⁵². Para el caso solo se determinaron los comentarios que tuvieran una coincidencia mayor o igual a 0.9. Posteriormente, se hizo un conteo de la cantidad de palabras negativas y positivas dentro de un mismo comentario, si la cantidad de palabras positivas es mayor a las negativas se considera un comentario positivo; si por el contrario hay mayor cantidad de palabras negativas se considera un comentario negativo.

4.4. EVALUACIÓN

Luego de crear el modelo, se procede a evaluarlo para determinar su consistencia, para ello se utiliza el método “matriz de confusión” el cual es una herramienta de visualización que se emplea en los modelos predictivos. Cada columna de la matriz representa el número de predicciones de cada clase (Positivo, negativo y neutral), mientras que cada fila representa a las instancias en la clase real⁵³, ver tabla 10.

| | | Predicción Polaridad de los tweets | | | |
|-----------------------|-----------|---|------------|------------|-------|
| | | Positivo | Negativos | Neutral | Total |
| Polaridad Real | Positivo | 66 | 5 | 29 | 100 |
| | Negativos | 24 | 569 | 262 | 855 |
| | Neutral | 593 | 211 | 949 | 1753 |

Tabla 14: Matriz de confusión

Como se puede observar el modelo clasifico 66 tweets positivos correctamente de un total de 100. Es decir, la precisión para predecir los tweets positivos es de 66%. Ahora, para determinar la precisión del modelo en general se deben sumar todos los tweets que se clasificaron correctamente y dividirlos sobre la población escogida. El resultado se expresa en la siguiente formula.

⁵²Dreßler, K., & Ngonga Ngomo, A.-C. (s.f.). *On the Efficient Execution of Bounded Jaro-Winkler Distances*. AKSW Research Group

⁵³SANTAMARÍA RUÍZ, W. (2010). *Modelo de detección de fraude basado en el descubrimiento simbólico de reglas de clasificación extraídas de una red neural*. Bogotá D.C: Universidad nacional de colombia

$$\text{Precisión} = \frac{66 + 569 + 949}{2.708} = 0.5849$$

La precisión del modelo es de 58,49% esto quiere decir que de cada 100 comentarios el modelo predecirá correctamente la polaridad de 58 o 59.

4.5. DESPLIEGUE

Para la visualización el modelo y construcción de los tableros de mando se utilizó la herramienta QlikSense de la empresa Qlik, la cual es muy intuitiva de manejar y bastante interactiva con el usuario final. A Continuación, se muestran los cuadros de mando construidos.

- **Tablero de control financiero:** El gráfico 22 muestra el tablero de control con los indicadores financieros filtrados por el tipo de seguro, año, mes y compañía.

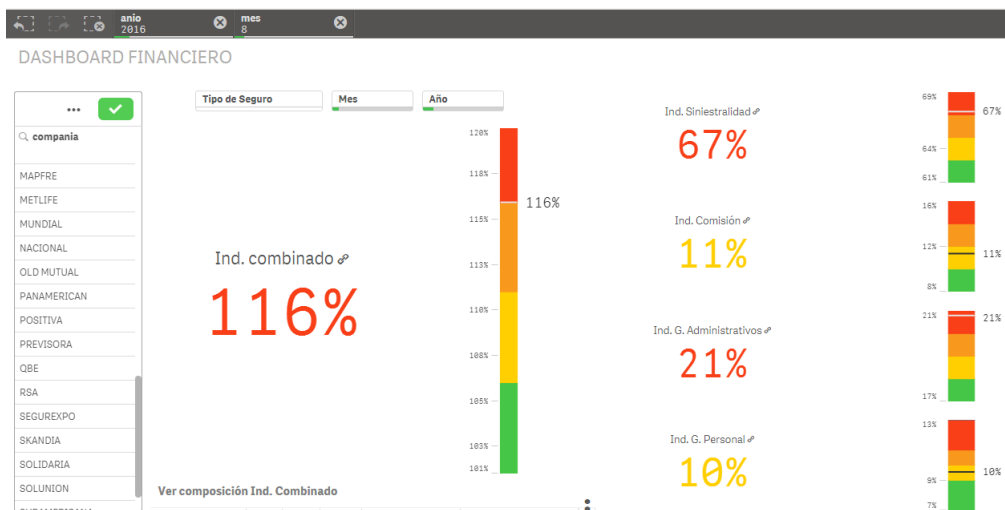


Gráfico 24: Tablero de control financiero

Como se puede observar el sector esta en rojo. Posee un indice combinado de 116% y es debido en gran parte al indice de siniestralidad y los gastos administrativos.

Este tablero se ha complementado con algunas métricas, en el gráfico 23 se puede observar la evolución del índice combinado del sector, pues al parecer el año 2016 es un año especial dado que el sector llevaba 6 años con un índice combinado por debajo del 116%. Las demás métricas son la evolución de las primas, del resultado técnico (Utilidad antes de inversiones) y la utilidad

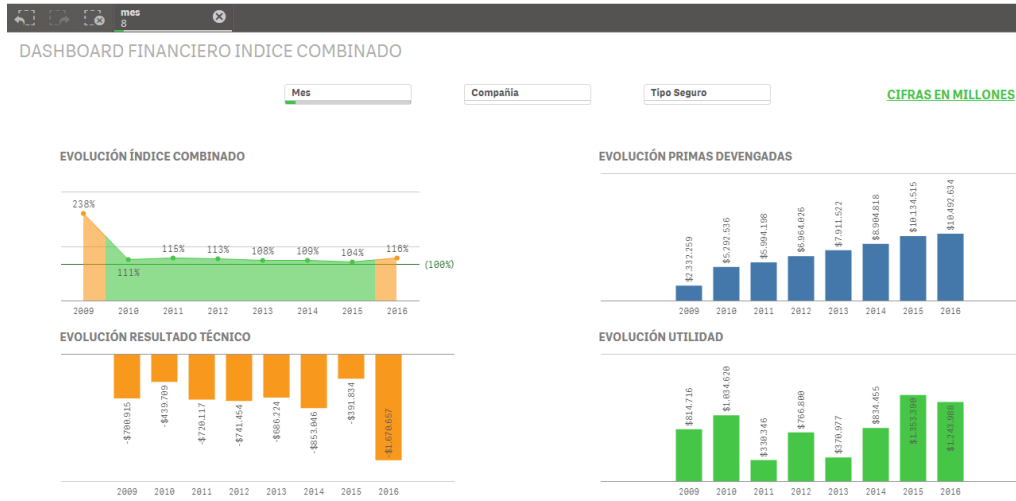


Gráfico 25: Tablero de control financiero – Métricas

- Tablero de control clientes:** Este tablero muestra el indicador de reacciones por cada país y en el caso de Colombia se desagrega a departamento. También se muestra la cantidad de palabras positivas y negativas dentro de los comentarios, la evolución del indicador día a día y cada uno de los comentarios con su correspondiente polaridad. Esta información se puede filtrar por año, red social, polaridad y país.

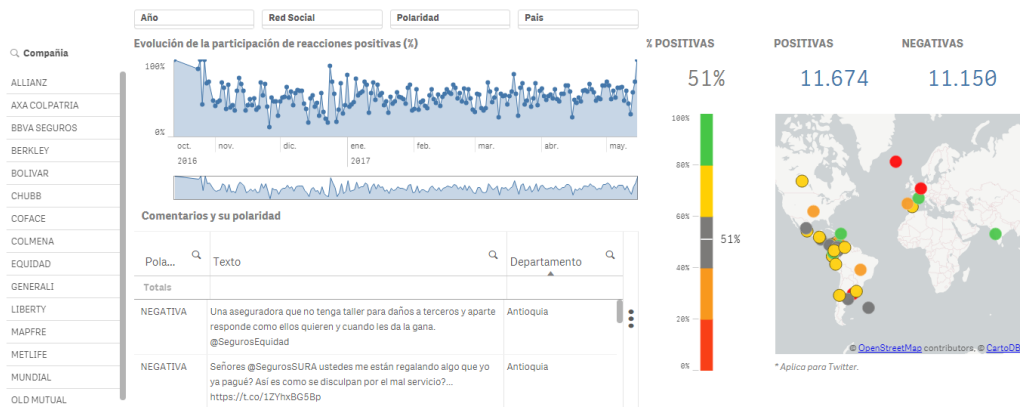


Gráfico 26: Tablero de control clientes

En el gráfico 24 se puede observar que la cantidad de reacciones positivas por parte de los usuarios está en el 51%, es decir, la polaridad es neutra en el sector asegurador.

Al igual que en el tablero de control se muestran algunas métricas que complementan el indicador. El gráfico 25 se muestra el crecimiento de las interacciones y publicaciones en redes sociales.

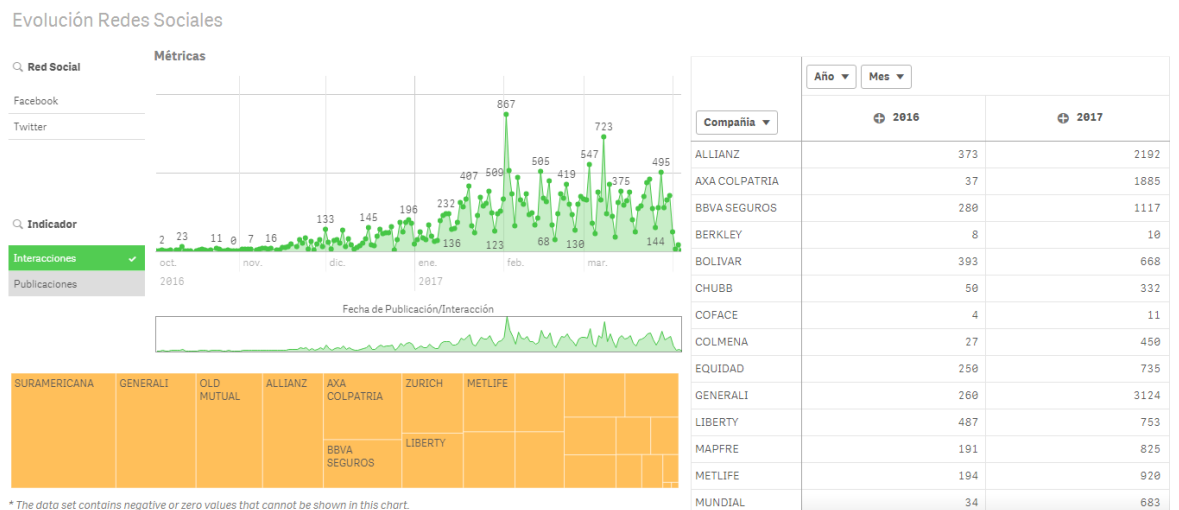


Gráfico 27: Tablero de control clientes - Métricas

También se pueden determinar las empresas que más interacción o publicaciones poseen en redes sociales, Las empresas que encabezan la interacción son suramericana, Generali y Old Mutual

5. VALIDACIÓN CON EXPERTOS

Por último, se realizaron reuniones de validación con las áreas de planeación y mercadeo de una compañía de seguros y con la profesora Luz Angélica Rodríguez. Ingeniera industrial de la universidad Distrital Francisco Jose de Caldas y Ph.D. (candidate) Industrial Ecology and Sustainability, Erasmus University of Rotterdam.

De la reunión con el área de planeación se determinaron las siguientes

Fortalezas:

- El tablero permite monitorear los principales indicadores del sector asegurador y permite la comparación entre compañías.
- El rango de los indicadores tiene un fundamento estadístico fuerte, lo cual respalda bastante el ejercicio.
- La herramienta es muy intuitiva para el usuario final, esto permitirá que los directivos la asimilen rápido y le saquen provecho.

Mejoras:

- Es importante para hacer un buen análisis, garantizar que la información que reportan las aseguradoras a Fasecolda sea la correcta pues es difícil creer que Alfa tenga un índice de gastos administrativos del 2% y de personal de 0%, mientras el índice de siniestralidad esté en 115%.
- Sería bueno desglosar el tablero de control en los diferentes ramos (ARL, Vida, Vida Grupo, Exequias, etc.) de cada compañía para llegar a un análisis por ramo.
- Aunque los rangos tienen un análisis estadístico riguroso cada compañía tendrá que ajustar los rangos a su estrategia.

De la reunión con el área de mercadeo se determinaron las siguientes

Fortalezas:

- El modelo permitirá optimizar el desempeño de las redes sociales.
- El indicador de reacciones positivas complementa los indicadores de satisfacción que posee la compañía.
- El modelo permitirá identificar las tendencias de los consumidores y las acciones a realizar para mejorar la calidad del servicio.

Mejoras:

- Separar las redes sociales, pues son muy diferentes, Facebook es una red social más familiar y de ocio mientras que Twitter tiende a usarse para expresar opiniones con un nivel menor de censura. Esto hace que los indicadores de cada una puedan no ser los mismos
- El porcentaje de precisión aun es bajo para poder implementar acciones con base en la información. Se recomienda madurar el modelo para poder tomar decisiones con él.

De la reunión con la profesora Luz Angélica Rodríguez se determinaron los siguientes aspectos:

Fortalezas:

- El modelo permite hacer benchmarking de la competencia y verlo a nivel visual, (visual management). Esto facilita determinar los principales competidores de una empresa aseguradora.

Mejoras:

- El método para determinar la polaridad por sectores económicos estratégicos para que los directivos puedan enfocar sus esfuerzos en estos sectores debería buscarse.

- La relación entre el indicador de reacciones positivas y los ingresos de las compañías de seguros debería determinarse, para establecer la causalidad entre el tablero de clientes y el tablero financiero.

6. CONCLUSIONES

1. Las redes sociales están cobrando cada vez más importancia en el mundo, los clientes las usan para manifestar sus felicitaciones o su inconformidad con algún servicio y esto hace que se genere más información para realizar análisis sobre el comportamiento y las necesidades del cliente.
2. No es necesario invertir en herramientas costosas para obtener datos. Varias compañías han trabajado en el desarrollo de instrumentos de libre acceso que le permitan a los ingenieros y usuarios finales extraer conocimiento de los datos.
3. La fase más importante en la metodología Crisp-DM es la del entendimiento del negocio, porque permite entender a la perfección las necesidades del negocio y acotar los alcances, si esta fase no se tiene clara, el desarrollo tendrá varios reprocesos que terminarán alargando el tiempo de ejecución y su costo.
4. Se puede montar un modelo de polaridad usando herramientas netamente libres como Python para la extracción de información de redes sociales; Pentaho para la transformación de los datos y creación del modelo; y QlikSense para la visualización de los resultados.
5. Los modelos analíticos de cualquier índole deben estar fundamentados en información de calidad. Si no es así, el modelo estaría propenso a llegar a conclusiones erróneas y por ende los directivos tomarán decisiones erróneas que pueden costar un monto considerable de dinero.

7. OPORTUNIDADES

1. El modelo propuesto posee un 58% de precisión, sin embargo, se puede mejorar con un diccionario más específico, obteniendo no solo polaridad de palabras sino también polaridad de frases. De esta forma tweets como “increíble lo que me está pasando. Juegan con el pan de mis hijos” que pueden contener palabras que no reflejan polaridad por sí solas, pero al unir las en una frase lo determinan como un tweet negativo. Por otro lado, también hay que determinar sarcasmos en los tweets ya que estos se califican como positivos, pero el sentimiento es realmente negativo. Además de validar si es necesario ponderar los comentarios pues las personas comentan más acerca de un mal servicio que de un excelente servicio.
2. Una de las hipótesis que se manejan en el documento es la relación que debería existir entre el indicador de las reacciones positivas y la rentabilidad de la empresa, aunque hay más factores que pueden influir sería interesante realizar una correlación de variables para determinar si en efecto esta existe o no.
3. Luego de tener las personas que tienen comentarios negativos sería interesante perfilarlas para determinar en qué parte de la cadena de valor está fallando la compañía y cuál es el público más difícil de complacer.
4. El ejercicio se realiza con una herramienta de visualización propietaria versión escritorio, es decir los cuadros realizados solo se podrán tener en la máquina instalada, por esta razón se recomienda revisar adquirir la licencia o implementar el tablero en otra herramienta de visualización que permita publicar a toda la compañía.
5. A medida que el tiempo pase la información crecerá bastante y se deberá validar la necesidad de aplicar técnicas de Big Data

8. BENEFICIOS

1. Medir la satisfacción del cliente a través de su comportamiento en las redes sociales
2. Optimizar esfuerzos para desarrollar análisis comparativos del sector, sobre la opinión de clientes en redes sociales y la gestión financiera.
3. Identificar nuevos prospectos de cliente para realizar gestión comercial
4. Monitorear y mejorar el seguimiento a las actividades de retención y fidelización de clientes.
5. Optimizar en termino de recursos y tiempo los análisis que se realizan sobre redes sociales
6. Evidenciar un amplio conjunto de posibilidades de bajo costo, para desarrollar ejercicios de BI, Minería de Datos y Big Data, sobre la información disponible en redes sociales, gremios y entes de control.
7. Predecir comportamientos de los clientes con base en la red social y realizar modelos de experiencia al cliente.
8. Medir el desempeño de las redes sociales

9. BIBLIOGRAFÍA

- Aguilar, L. J. (2015). *Sistemas de Información en la empresa*. Mexico: Alfaomega.
- Alvarez Medina , M. T., & Moreno Velarde, S. A. (s.f.). El balance ScoreCard, Una herramienta para la planeación estratégica. *Iton*.
- Alveiro Montoya, C. (2011). El banaced scorecard como herramienta de evaluación en la gestión administrativa. *Visión de futuro*.
- Angeles Larrieta, M. I., & Santillan Gómez, A. M. (s.f.). *Minería de datos: Concepto, características, estructura y aplicaciones*. (s.f.).
- Arooyo Almaraz, I., Baladrón Pazos, A., & Martín Nieto, R. (2013). *La comunicación en redes sociales: Percepciones y usos de las Ong españolas*. Universidad Rey Juan Carlos. (s.f.).
- Asociación Española para la Calidad . (Diciembre de 2012). Obtenido de La calidad en la atención al cliente: https://www.aec.es/c/document_library/get_file?uuid=a17b6ac7-e816-4145-aeb9-a0c12e479d1b&groupId=10128
- Bhatt, G. (2001). Knowledge management in organisations: Examining the interaction between technologies, techniques, and people. *Journal of Knowledge*, 8-18.
- Blanco Dopico, M. I., Guzmán, B., & Agra, S. (1999). El enfoque conductual contable y su reflejo en un cuadro de mando integral. *Revista española de financiación y contabilidad*, 77-104.
- Botero, J. H. (6 de Mayo de 2016). *Entr. (B. Radio, Entrevistador)*.
- Boyd, D. M., & Ellison, N. B. (2007). *Social network sites: Definition, history, and scholarship*. *Journal of Computer-Mediated Communication*. (s.f.).
- Burgeño, P. F. (02 de Marzo de 2009). *Clasificación de Redes Sociales*. Recuperado el 04 de Julio de 2016, de <http://www.pabloburgueno.com/2009/03/clasificacion-de-redes-sociales/>. (s.f.).
- Davenport, T., & Prusak, L. (1998). *Working Knowledge: How organisations manage what they know*. *Harvard Business School Press*, 224.
- Díaz Rangel, I., Sidorov, G., & Suárez-Guerra , s. (2014). *Creación y evaluación de un diccionario marcado con emociones y ponderado para el español*. instituto politécnico nacional. (s.f.).

- Dreßler, K., & Ngonga Ngomo, A.-C. (s.f.). *On the Efficient Execution of Bounded Jaro-Winkler Distances*. AKSW Research Group. (s.f.).
- Facebook. (2017). *Developer Facebook*. Obtenido de Developer Facebook: <https://developers.facebook.com/>
- Fasecolda. (Mayo de 2017). *Fasecolda*. Obtenido de Federación de Aseguradores Colombianos: <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>
- Fasecolda. (Marzo de 2016). *Indicadores de gestión*. [En Línea]. Recuperado el 14 de 02 de 2016, de <http://www.Fasecolda.com/index.php/Fasecolda/estadisticas-del-sector/indicadores-de-gestion/>. (s.f.).
- Fernández, A. (2001). El balance ScoreCard: Ayudando a implantar la estrategia. *Universidad de Navarra - IESE*, 32.
- Fernández, J., Llopis, F., Martínez, P., Gutiérrez, Y., & Díez, Á. (2016). Analizando opiniones en las redes sociales. *Procesamiento del Lenguaje Natural*, 141-148.
- Gartner. (Febrero de 2017). *Magic Quadrant for Business Intelligence and Analytics Platforms*. Obtenido de Gartner: <https://www.gartner.com/doc/reprints?id=1-3TYE0CD&ct=170221&st=sb>
- Gómez Vieites, A., & Otero Barros, C. (2013). *Redes Sociales en la empresa*. Bogotá, Colombia: Ediciones de la U.
- Guzmán Duquea, A. P., Moral Pérez, M. E., & Ladrón de Guevara, F. G. (2016). Usos de Twitter en las universidades iberoamericanas. *RELATEC*, 28.
- Kaplan, R., & Norton, D. (2009). *Cómo utilizar el Cuadro de Mando Integral*. Barcelona: Gestión 2000.
- Kaplan, R., & Norton, D. (2011). Cuadro de manto integral: the Balanced Scorecard. *Gestion 2000*.
- Kaushik, A. (2011). *analitica web 2.0*. Barcelona: Gestión 2000.
- Kdnuggets. (2014). *Kdnuggets*. Obtenido de <http://www.kdnuggets.com/polls/2014/analytics-data-mining-data-science-methodology.html>

- Kdnuggets.* (2016). *Kdnuggets.* Obtenido de *Kdnuggets:* <http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html/2>. (s.f.).
- Lozares, C. (1996). La teoría de redes sociales. *Universidad autonoma de barcelona*, 103-126.
- Mapfre.* (2015). *Responsabilidad Social - Informe anual.* (s.f.).
- Perez Soler, S. (19 de 02 de 2017). *Tesis en red.* Obtenido de http://www.tesisenred.net/bitstream/handle/10803/362644/Tesi_Susana%20Perez.pdf
- Prieto, A. T., & Hernández, C.* (20 de Junio de 2014). *calidad de servicio como factor de competitividad en los postgrados de las instituciones universitarias públicas.* Recuperado el 03 de Julio de 2016. (s.f.).
- PWC. (20 de Mayo de 2017). *CEO Survey.* Obtenido de Para crecer en un mundo cambiante: Innovación con talento y tecnología: <http://www.pwc.com/co/es/assets/document/ceo-survey-colombia-7a-edicion.pdf>
- RAMÍREZ RAMÍREZ, C. P. (2010). *El Balanced Scarecard como herramienta metodológica para evaluar la gestión en la unidad de compras en el municipio de Medellín.* Tesis de Especialización en Alta Gerencia, Universidad de Medellín.
- Requena Ponce, M. V., & Serrano López, G. C.* (02 de Novimebre de 2007). *Calidad del servicio desde la perspectiva de los clientes, usuarios y auto-percepción de empresas de captación de talento.* Recuperado el 04 de Julio de 2016. (s.f.).
- Ricardo Barreto, C., & Chavarro Jimenez, A. (2010). El uso de Facebook y Twitter en educación. *Instituto de estudios en educación - IESE*, 5.
- Santamaria Ruíz, W.* (2010). *Modelo de detección de fraude básado en el descubrimiento simbólico de reglas de clasificación extraídas de una red neural.* Bogotá D.C: *Universidad Nacional de Colombia.* (s.f.).
- Seguros Bolivar.* (2015). *Informe de Gestión.* (s.f.).
- Spek, R., & Spijkervet, A. (1997). Knowledge management: Dealing Intelligently with knowledge. *Utrecht: Kenniscentrum CIBIT*, 25.
- SPSS. (1999). *CRISP-DM 1.0 Step by step Data Mining Guide.*

Suárez Vicente, R. (2015). Gestión Visual - Aplicación de herramientas LEAN en el área de mantenimiento de una empresa minera.

Suramericana;Filial de Grupo SURA. (2014). Informe anual. (s.f.).

Tarí, G. J. (2000). *Calidad Total: fuente de ventaja competitiva. Publicaciones Universidad de Alicante.*

Torres Samboni, L. A. (2015). ANÁLISIS DE SENTIMIENTOS SOBRE EL POSCONFLICTO COLOMBIANO UTILIZANDO HERRAMIENTAS DE MINERÍA DE TEXTO.

Trudeau, R. (1993). Introduction to Graph Theory.

Twitter. (19 de 02 de 2016). *twitter*. Obtenido de <https://support.twitter.com/articles/352810?lang=es>

Twitter. (2017). *Twitter Developer Documentation. Obtenido de Twitter Developer Documentation: https://dev.Twitter.com/rest/public/rate-limiting. (s.f.).*

Twitter. (2017). *Twitter. Obtenido de Twitter: https://dev.Twitter.com/. (s.f.).*

UIAF. (2014). *Unidad de Información y Análisis Financiero. Obtenido de UIAF: http://www.urosario.edu.co/observatorio-de-lavado-de-activos/Archivos_Lavados/Aplicabilidad-de-la-mineria-de-datos-(1).pdf*

Virveda Benito, F., & Román Carrillo, J. (s.f.). Minería de datos y aplicaciones. (s.f.).

Yuvila M, S., & Darnes Vilariño, M. J. (2015). Modelos para detectar la polaridad de los mensajes en. *Research in Computing Science, 29-42.*

ANEXO I

Indicadores propuestos para el sector asegurador

| Perspectiva | Nombre del indicador | Medición | A | B | Periodicidad | Formula | Tendencia |
|-------------|----------------------------------|--|--------------------------------------|----------------------|--------------|-----------|-------------|
| FINANCIERA | Índice Combinado | Porcentaje de los ingresos que se usan para atender el core del negocio | Suma de los costos y gastos | Primas emitidas en t | Mensual | $A/B*100$ | Decreciente |
| | Índice Siniestralidad | Porcentaje de los ingresos que se destinan para atender los costos de los accidentes | Valor de la siniestralidad incurrida | Primas emitidas en t | Mensual | $A/B*100$ | Decreciente |
| | Índice de Comisiones | Porcentaje de los ingresos que se entregan a los intermediarios | Valor de la intermediación Neta | Primas emitidas en t | Mensual | $A/B*100$ | Decreciente |
| | Índice de Personal | Porcentaje de los ingresos que se destinan a los salarios y bienestar de los empleados | Valor del gasto del personal | Primas emitidas en t | Mensual | $A/B*100$ | Decreciente |
| | Índice de Gastos Administrativos | Porcentaje de los ingresos que usan para los gastos de la compañía (Servicios públicos, licencias tecnológicas etc.) | Valor de los gastos administrativos | Primas emitidas en t | Mensual | $A/B*100$ | Decreciente |

Tabla 15: Indicadores financieros propuestos para el sector asegurador

| Perspectiva | Nombre del indicador | Medición | A | B | Periodicidad | Formula | Tendencia |
|-------------|-------------------------------------|--|--|--|--------------|---------------|-----------|
| CLIENTES | Tamaño de la red social | Crecimiento en usuarios que realizan por lo menos un comentario sobre alguna aseguradora o que sean amigos o seguidores de una alguna de ellas | Cantidad de usuarios que realizan por lo menos un comentario sobre alguna aseguradora o que sean amigos o seguidores de una alguna de ellas en t | Cantidad de usuarios que realizan por lo menos un comentario sobre alguna aseguradora o que sean amigos o seguidores de una alguna de ellas en t-1 | Mensual | $(A/B-1)*100$ | Creciente |
| | Publicaciones | Crecimiento de publicaciones de cada aseguradora en redes sociales | Cantidad de post y tweets que realice una aseguradora en t | Cantidad de post y tweets que realice una aseguradora en t-1 | Mensual | $(A/B-1)*100$ | Creciente |
| | Atención | Crecimiento de atenciones realizadas por medio de las redes sociales | Cantidad de respuestas a quejas que realicen los usuarios por medio de las redes sociales en t | Cantidad de respuestas a quejas que realicen los usuarios por medio de las redes sociales en t-1 | Mensual | $(A/B-1)*100$ | Creciente |
| | Interacción | Crecimiento en la interacción de los usuarios en la redes sociales | Cant. Me gusta + Cant. Comentarios + Cant. Veces compartido + Cantidad de retweets en t | Cant. Me gusta + Cant. Comentarios + Cant. Veces compartido + Cantidad de retweets en t-1 | Mensual | $(A/B-1)*100$ | Creciente |
| | Posibles Intermediarios Potenciales | Crecimiento de los intermediarios que permiten acceder a más público objetivo. | Cantidad de amigos o seguidores de cada uno de los seguidores o amigos de cada compañía de seguros en t | Cantidad de amigos o seguidores de cada uno de los seguidores o amigos de cada compañía de seguros en t-1 | Mensual | $(A/B-1)*100$ | Creciente |
| | Reacciones Positivas | Crecimiento de reacciones positivas en redes sociales | Cantidad de reacciones positivas | Total de reacciones | Mensual | $(A/B-1)*100$ | Creciente |

Tabla 16: Indicadores de clientes propuestos para el sector asegurador