

**CONSIDERACIONES GENERALES PARA LA APLICACIÓN DEL BIG DATA
EN NEGOCIOS DE REDES DE MERCADO**

Monografía para optar por el título de Ingeniero de Sistemas.

CARLOS ALEXANDER SALGUEDO JAIME

Código 66990028

UNIVERSIDAD LIBRE

Facultad de Ingeniería

Departamento de Ingeniería De Sistemas

Diplomado Internacional Big Data

Bogotá D.C., 2017

RESUMEN.

El propósito de este estudio monográfico es el plantear un enfoque personal de la aplicación de técnicas y tecnologías del Big Data en negocios de Redes de Mercadeo, que permitan agilizar el crecimiento de este tipo de compañías.

Se Plantea la hipótesis de que es posible generar una mayor rentabilidad y crecimiento de la industria de las redes de Mercado, a partir de el uso del Big Data y de la Minería de Datos para clasificar a las personas con base en los patrones de conducta en las redes sociales y de está manera generar estrategias de crecimiento para cada grupo.

El estudio inicia con los aspectos que se consideraron para determinar como un caso de aplicación del Big Data, las técnicas sugeridas para el manejo de la información y plantea un primer objetivo clave para medir la viabilidad del proyecto.

Palabras Claves: Big Data, Analítica de Datos, Minería de Datos, Análisis de sentimientos, Redes Sociales, Patrones de Conducta, Redes de Mercadeo, Personalidades.

INTRODUCCIÓN.

El éxito en la distribución de productos de consumo masivo a través de Redes de Mercado se encuentra al involucrar la mayor cantidad de personas posibles a un sistema de distribución, y que cada uno de sus integrantes consuma y venda la mayor cantidad posible de los productos generados por la empresa.

Con este fin, las más importantes empresas que fomentan el Mercadeo en Red han conformado un sistema de educación no formal que consta de videos, audios, libros y eventos que buscan dar a conocer a la comunidad en general las ventajas de esta actividad económica, la manera en como las personas con mayor éxito han desarrollado sus negocios e inspirar a más personas a pertenecer a una red de mercadeo.

Durante años los lideres de las redes se han dedicado a identificar los diferentes motivos que tienen las personas para ingresar a una red de mercado y con base en esa información desarrollan los audios, videos y eventos que buscan expandir sus redes de distribución.

Este material genera diversas reacciones en cada una de las personas y por ejemplo un mismo audio por si solo no genera necesariamente el mismo resultado en dos personas diferentes. Por este motivo decidí que se pueden usar métodos de Big Data para clasificar los miembros de una comunidad de acuerdo a sus personalidades e intereses con el propósito de hacer llegar la información adecuada que le haga tomar la decisión de ingresar a una red de mercadeo.

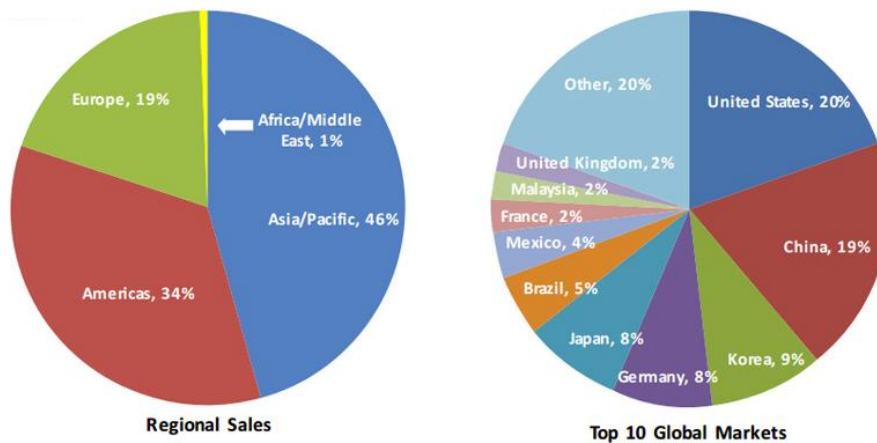
Uno retos en este tipo de sistemas va a hacer la clasificación de las personas por personalidades, motivo por el cual para el desarrollo de esta metodología se sugiere tomar como referencia los estudios de la Dra. Florence Littauer¹, quien clasifica a las personas en 4 diferentes personalidades de acuerdo a sus comportamientos, sentimientos y prioridades en la vida. En el momento de la

¹ LITTAUER FLORENCE, Enriquezca su Personalidad, Editorial Unilit Primera Edición. 2013.

implementación de esta metodología se pueden modificar los criterios de clasificación de las personas.

CONSIDERACIONES PARA CLASIFICAR ESTE ESTUDIO COMO UN TEMA DE BIG DATA.

La federación mundial de asociaciones de venta directa - WFDSA por sus siglas en ingles (World Federation of Direct Selling Associations), la cual agrupa a las principales asociaciones de venta directa en el mundo, estimó que para el año 2015 cerca de 103.000.000 de personas en el mundo, y 2'181.174 personas en Colombia se encontraban afiliadas con empresa de Redes de Mercadeo.



Global Industry: \$183,729 (USD millions), Up 7.7% in Constant 2015 USD

Region/Country	2015 Retail Sales (1)		% Sales Change (YOY) in Constant 2015 USD (2)	3-Year CAGR in Constant 2015 USD (2012-15)	Direct Sellers (3)
	Local Currency (millions)	USD (millions)			
Global	na	183,729	7.7% ▲	7.2%	103,277,493
Asia/Pacific	na	84,084	11.1% ▲	10.6%	51,922,926

Fuente Estadística de Ventas por Redes de Mercadeo para el Año 2015. ©Copyright WFDSA 2016

Procesar datos del comportamiento y actividad en redes sociales, información financiera de ventas y consumos, asistencia a eventos y/o paginas visitadas de las cantidades de personas vinculadas a las redes de mercadeo en el mundo o en Colombia requiere de un alto nivel de procesamiento y mecanismos de analítica de datos automaticos.

BENEFICIOS DEL USO DE BIG DATA EN LA CONSTRUCCIÓN DE UN NEGOCIO DE MULTINIVEL

Como se menciona en la introducción de este trabajo, el fin de un negocio de multinivel es vincular la mayor cantidad de gente posible a su red de distribución y hacer que el consumo y venta de cada integrante de la red aumente, los ingresos de una persona que se dedique al negocio de venta directa están directamente ligados a la cantidad de productos que logre vender de manera individual o a través de mas personas.

Una de las estrategias que se usan para lograr vincular a más personas a la red es mediante el uso de audios, videos, eventos, libros y otros medios audiovisuales que inspiren a las personas a pertenecer a determinado grupo o red de distribución.

Es por eso que una de las tareas que se realizan al momento de iniciar un negocio de redes de mercadeo, es indagar y detectar que temas o motivos son los que más llaman la atención de la población en general, que hagan que las personas quieran entrar a un sistema de distribución. Adicionalmente también se trata de generar estrategias para que las personas que ya pertenecen a la red de distribución aumenten el nivel de ventas.

En mi concepto esta solución de Big Data tiene sentido en la medida en que dos personas no necesariamente tienen la misma reacción a un hecho en común, es decir, es posible que dos personas que vivan en la misma ciudad y que pertenezcan a al mismo nivel socioeconómico, no necesariamente van a tener la misma reacción al ver un video que los induzca a ingresar a un negocio de multinivel. En la medida en que se logre predecir que material audiovisual va

a tener un mayor impacto en una persona, se puede inferir que el rendimiento de una persona en cuanto a volúmenes de venta va a mejorar.

Lo que busca este trabajo es generar un sistema de información que permita a las empresas de multinivel o a cualquiera de sus afiliados es:

- Análisis Descriptivos
 - Conocer los lugares geográficos en donde se encuentran ubicados cada uno de los miembros de una red de distribución.
 - Conocer el consumo y venta de productos por zonas, y su relación de acuerdo a la cantidad de afiliados.
 - Medir el nivel de influencia de cada uno de los miembros de una red, es decir, que esta solución permitiría determinar la distribución geográfica y la cantidad de personas que conforman una red de un afiliado específico.
 - Medir la cantidad de reproducciones de videos y audios que incentiven a las personas a ingresar a una red, así como la estadística de la cantidad de personas que están visitando determinadas paginas o accediendo a las redes sociales de las empresas de multinivel.
 - Medir la frecuencia y el nivel de asistencia a los eventos de personas que ya hacen parte del multinivel y de personas que aun no han ingresado.

- Análisis Diagnostico:
 - Determinar las causas de el crecimiento o decrecimiento de las redes de distribución tanto en número de personas como en ventas en las diferentes zonas, teniendo en cuenta los datos obtenidos en los análisis descriptivos.
 - Conocer la opinión, sentimientos y el impacto causado en las personas de cada uno de los audios, videos, libros o eventos.

- Agrupar o clasificar a los usuarios de las redes sociales e independiente de que pertenezcan o no a una red de distribución de acuerdo a sus patrones de conducta. De cada uno de los grupos medir el impacto en cuanto al crecimiento de la red de acuerdo a los audios, videos y eventos.
 - Determinar los mejores sitios para los eventos de acuerdo a la cantidad de personas y distribución geográfica de las personas pertenecientes a la red de mercadeo.
 - Medir los resultados obtenidos al aplicar determinadas estrategias a grupos pequeños antes de replicarlas a los demás grupos de similares características.
- **Análisis Predictivo y Prescriptivo**
 - Establecer la estrategia para cada uno de los grupos, referente a los temas a tratar en los audios, publicaciones en redes sociales, eventos que permitan obtener la mayor productividad en cuanto a ventas y cantidad de personas en la red.
 - Contar con un sistema de información automático que permita predecir los temas y materiales audiovisuales recomendados a una persona nueva con base a los patrones de comportamiento identificado en las redes sociales y que personas de ese mismo comportamiento tienen reacciones positivas y han tenido resultados aceptables en la distribución de productos

FUENTES DE INFORMACIÓN DISPONIBLES EN LAS REDES DE MERCADEO

En mi concepto de acuerdo a lo visto en el desarrollo del diplomado, el éxito de una solución de Big Data depende en gran medida a la calidad de los datos con los que se trabaja. Es por eso que dedico este capítulo del trabajo para

describir las fuentes de información con las que se cuenta al hacer una solución de Big Data enfocada a la industria de las redes de mercadeo.

Datos de la Empresa

La mayor parte de las empresas que distribuyen sus productos a través de venta directa, brindan a sus afiliados o vendedores herramientas de software que les permiten consultar información de su red como por ejemplo:

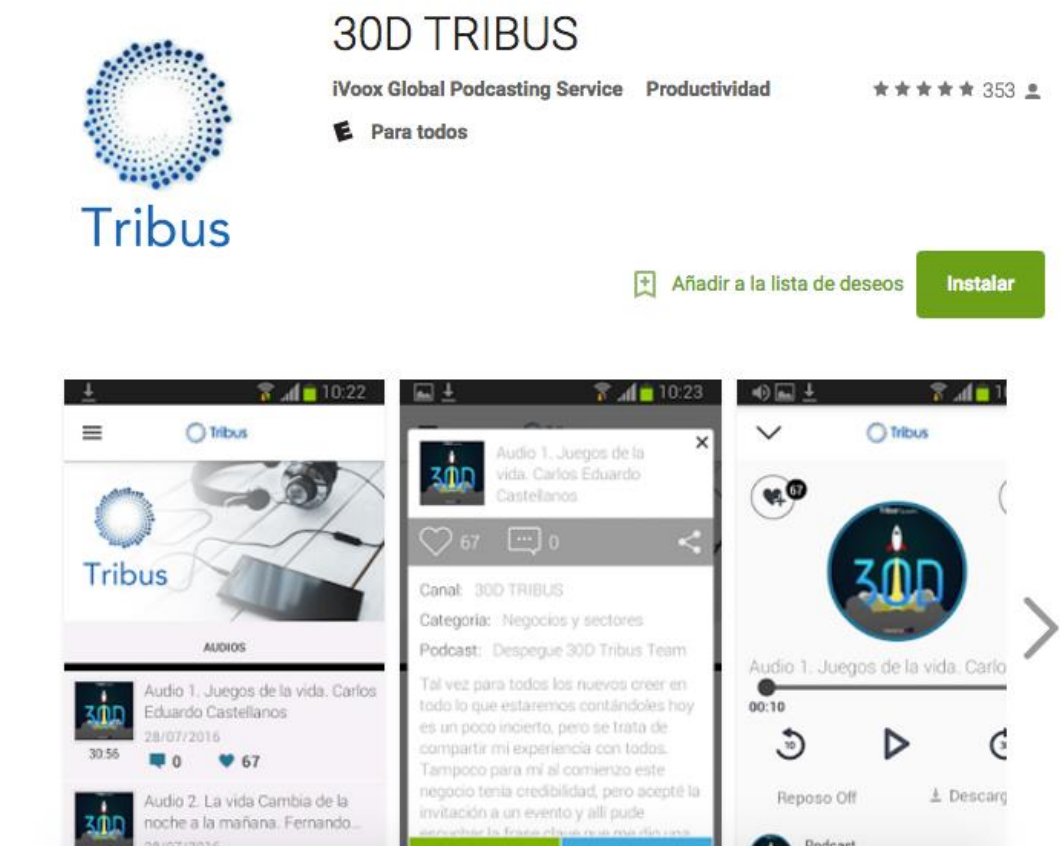
- Numero de Personas relacionadas con su código.
- Datos básicos de las personas asociadas a su código (Nombre, fecha de nacimiento, números de cedula, correo electrónico, ciudad de residencia).
- Datos de los consumos relacionados a cada código. En el caso de que la empresa tenga varias líneas de producto muestra el consumo por cada una de las líneas de cada uno de los miembros de la red.
- Nivel. Por lo general las empresas de Redes de Mercado clasifican a los afiliados de acuerdo al nivel de consumo que generan de forma personal o grupal. Esta información también puede ser consultada con la empresa.

Los datos de la empresa son los que me van a permitir la medición del impacto de las estrategias en el crecimiento del negocio de redes de mercado en el que se aplique este modelo de Big Data en la medida que las ventas de un grupo sean superiores al promedio de los demás.

Aplicaciones Propias

En la actualidad las personas que ya pertenecen a negocios de redes de mercadeo están impulsando el desarrollo de aplicaciones web de manera particular, como herramienta que permita a los usuarios de su red compartir información.

Un ejemplo de este tipo de aplicaciones es **30D Tribus**, la cual permite a los usuarios acceder a audios e información sobre redes de mercado de manera gratuita. Con el uso de este tipo de herramientas, en las cuales se accede usando una cuenta de usuario de correo electrónico o de redes sociales es posible medir que audios esta escuchando cada usuario y que audios son los mas populares.



Fuente: <https://play.google.com>

Redes Sociales

Hoy por hoy las redes sociales se han convertido en una fuente muy popular de compartir información a nivel general, miles de empresas de publicidad monitorean a diario el uso de las redes de sociales con el fin de generar sus

propias estrategias y han generados datos muy interesantes del uso actual de internet en el mundo.

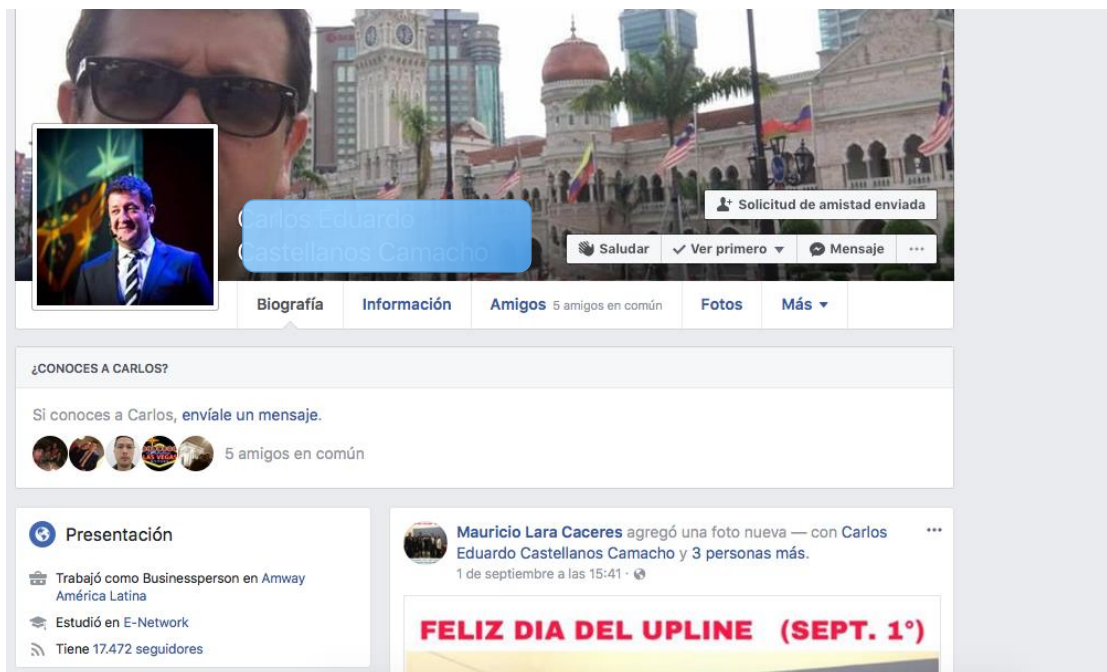


Fuente: Tomado de <https://es.statista.com/>

Actualmente también las personas que hacen parte de empresas de multinivel hacen uso frecuente de estas herramientas para transmitir sus mensajes. Las siguientes imágenes son las paginas de las redes sociales de algunas de las personas mas conocidas en Colombia en el negocio de las redes de mercadeo (Se ocultan sus nombres debido a que no contamos con permiso explicito).



Fuente: <https://www.youtube.com/>



Fuente: <https://www.facebook.com/>



Fuente: <https://www.instagram.com/>

Estos son solo algunos ejemplos que muestran el comportamiento y número de seguidores que tienen en este momento personas dedicadas a la distribución de productos a través de Redes de Mercadeo.

CAPTURA DE DATOS REGISTRADOS EN LAS REDES SOCIALES

La popularidad de las redes sociales en la actualidad, han volcado el interés de muchas personas y empresas hacia el desarrollo de aplicaciones que permitan la captura y análisis de los datos almacenados allí². La siguiente es un resumen de diferentes aplicación con las que se puede capturar la información:

Herramienta	SO	Descargable desde	Plataformas
Mozdeh	Windows (recomendable PC)	http://mozdeh.wlv.ac.uk/installation.html	Twitter
Webometric Analyst	Windows	http://lexiurl.wlv.ac.uk/	Twitter (con capacidades de extracción de imagen) YouTube Flickr Mendeley Otros recursos web
NodeXL	Windows	http://nodexl.codeplex.com/	Twitter YouTube Flicker
Netlytic	Web based	https://netlytic.org/	Twitter Facebook YouTube Instagram

² Lista de aplicaciones en <http://socialmediadata.wikidot.com/>

Twitter Archiving Google Spreadsheet (TAGS)	Web based	https://tags.hawksey.info/	Twitter
Chorus	Windows (Desktop advisable)	http://chorusanalytics.co.uk/chorus/request_download.php	Twitter
DiscoverText (Gratis prueba de 30 días)**	Web based	http://discovertext.com/	Twitter Facebook Blogs Forums Plataformas de Noticias online
COSMOS Project	Windows MAC OS X	http://www.cosmosproject.net/	Twitter

De las redes sociales se puede determinar los siguientes datos que quizás permitan determinar la personalidad de un usuario:

- Los gustos musicales.
- Los libros leídos.
- Sentimientos de los comentarios realizados.
- La clasificación de las fotos publicadas y comentadas para detectar si se trata de viajes, eventos, paisajes, animales, carros o personas.
- Cantidad de amigos en las redes sociales.
- Comentarios y datos acerca de deportes.
- Usuarios o personalidades que siguen en redes sociales.
- Sentimientos acerca de publicaciones y hashtag específicos.
- Los videos vistos en Youtube.

MINERÍA DE DATOS

La Minería de Datos es una etapa del descubrimiento de conocimiento en Bases de Datos (KDD - por sus siglas en inglés: Knowledge Discovery in Databases) y consiste en el uso de técnicas de tecnología, base de datos, estadística y aprendizaje de maquina con el propósito de³:

- Predecir tendencias y comportamientos de manera automática.
- Descubrimiento de patrones en grandes volúmenes de datos.

Durante la planeación del proyecto se debe pensar con el cumplimiento de los siguientes procedimientos:

- Fase de Exploración: Se refinan las datos y se identifican los problemas de calidad que pueda contener la información.
- Preparación de los datos: Limpiar y dar formato a los datos recolectados.
- Modelamiento: Selección de las técnicas y métodos que se deben usar para el caso.
- Evaluación: Evaluar los resultados previos del modelo.
- Despliegue: La implementación o ejecución del modelo con una base de datos mayor.

Lo primero que se marca de este estudio es que se clasifica como un caso de Aprendizaje Automático no Supervisado, esta clasificación se da debido a que no tenemos un conocimiento histórico de los comportamientos o patrones que podemos encontrar en la base de datos que me marquen la personalidad de un usuario de las redes sociales.

Este planteamiento es similar al trabajo realizado en minería de datos para la detección de fraudes⁴, en las cuales su usan métodos de Clustering, que

³ RANTES, Mónica. Detección de Fraude Usando Tecnicas de Clustering. Peru. P. 12.

⁴ RIOS VILLEGAS, Andrés Felipe. Minería de datos aplicada a la detección de clientes con alta probabilidad de fraudes en sistemas de distribución. Colombia. P. 20.

consiste en realizar una búsqueda de patrones en una base datos, con la característica que no se cuenta con información previa de los patrones.

Las técnicas de Clustering se pueden clasificar en tres tipos:

- Clustering Jerárjico.
- Clustering basado en particiones.
- Clustering basado en grid

Método Sugerido

Basado a la revisión de las técnicas de Clustering, se sugiere aplicar un método basado en técnicas predictivas, el cual permitirá encontrar los grupos y los valores atípicos de la base de datos con el fin de detectar los usuarios de las redes sociales con similares patrones de comportamiento.

La metodología sugerida incluye los siguientes pasos:

1. Emplear el algoritmo de K-Means para la detección de los grupos.
2. Ubicación de los puntos mas lejanos con el propósito de identificar los casos atípicos y excluirlos del análisis.
3. Detectar los grupos o Clusters con pocos elementos, para detectar que no se traten de casos atípicos y del ser el caso poderlos excluir del análisis.
4. Tomar una muestra de cada uno de los grupos, tomando como tamaño mínimo de la muestra de acuerdo a la estimación de una proporción y conociendo el tamaño de la población de estudio.

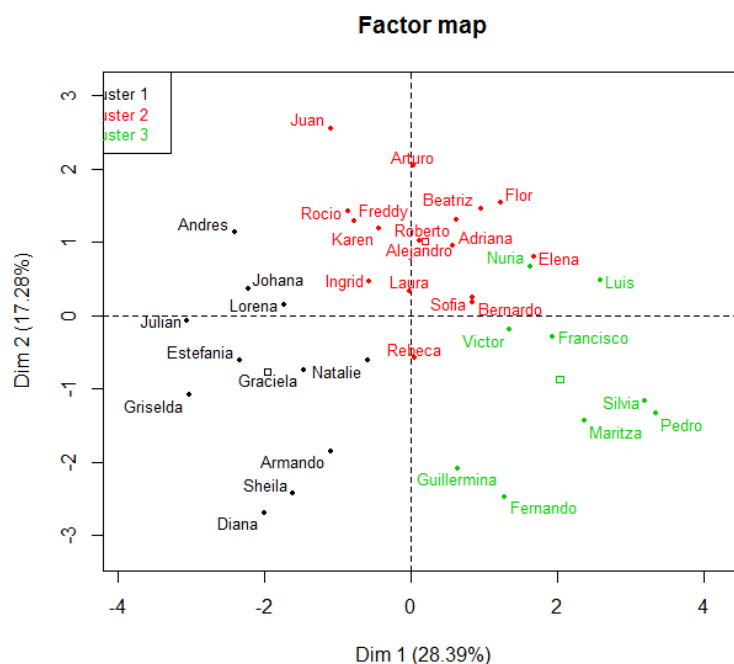
$$n = \frac{N \times Z^2 \times p \times q}{d^2 \times (N - 1) + Z^2 \times p \times q}$$

Donde N = Tamaño de la población del cluster seleccionado ; Z = Coeficiente de confianza; p = Probabilidad de que pertenezca a cualquiera de las 4

personalidades; $q = 1 - p$; $d =$ Precisión (Error máximo admisible en términos de proporción. Ej. 5% $\rightarrow 0.05$)

5. Con esta información tomar a la cantidad de usuarios definida en la estimación de la muestra y establecer un análisis de personalidad a cada uno de los usuarios de redes sociales incluidos en la muestra.
6. Establecer el porcentaje de usuarios de cada una de las personalidades en cada uno de los clusters.

La idea de este análisis es que logremos identificar los patrones que logren dividir la población de usuarios en clusters con una tendencia bien definida en cuanto a la personalidad de sus miembros. Si logramos identificar una tendencia en cada grupo podemos inferir que logramos identificar los patrones que permitan predecir la personalidad de un usuario de las redes de sociales.



Fuente: Autor

A manera de ejemplo, el proposito del agrupamiento sería conseguir clusters en donde se identifique una marcada tendencia de sus miembros hacia una determinada tendencia.

- Cluster 1 – Puntos en color negro en la gráfica:
 - Porcentaje de la muestra que tiene la Personalidad 1: 85%
 - Porcentaje de la muestra que tiene la Personalidad 2: 5%
 - Porcentaje de la muestra que tiene la Personalidad 3: 10%
 - Porcentaje de la muestra que tiene la Personalidad 4: 0%
- Cluster 2 Personas en color rojo en la grafica:
 - Porcentaje de la muestra que tiene la Personalidad 1: 15%
 - Porcentaje de la muestra que tiene la Personalidad 2: 5%
 - Porcentaje de la muestra que tiene la Personalidad 3: 70%
 - Porcentaje de la muestra que tiene la Personalidad 4: 10%
- Cluster 3 Personas en color verde en la grafica:
 - Porcentaje de la muestra que tiene la Personalidad 1: 3%
 - Porcentaje de la muestra que tiene la Personalidad 2: 2%
 - Porcentaje de la muestra que tiene la Personalidad 3: 5%
 - Porcentaje de la muestra que tiene la Personalidad 4: 90%

CONCLUSIONES

En mi concepto la viabilidad de aplicar un proyecto de Big Data que permita predecir la estrategia adecuada para una personal que se dedique a la distribución de productos por medio de Redes de Mercadeo, esta directamente ligada al hecho de lograr agrupar a los miembros de una población con similares personalidades, basados principalmente en los datos registrados en las redes sociales.

El numero actual de usuarios que estan ingresando a las redes sociales y la información que se comparte en las diferentes platarformas son una importante fuente de información, con la que es posible predecir y anticipar estrategias de crecimiento en diversas áreas de la industria, incluidas las redes de mercadeo.

Con los avances de en el aprendizaje de maquina en temas de comprensión automatica de texto, la identificación de imágenes y el analisis de videos y audio permite encontrar nuevos parametros de conducta en las redes sociales con los que es posible identificar y clasificar a los usaurios.

BIBLIOGRAFÍA

ASENSIO BLASCO, Elena – Aplicación de Técnicas de Minería de Datos en Redes Sociales. Valencia – España : UNIVERSIDAD POLITECNICA DE VALENCAI, 2015 [Revisado en 05 Septiembre de 2017] disponible en: <https://riunet.upv.es/bitstream/handle/10251/56102/ASENSIO%20-%20Apliaci%C3%B3n%20de%20t%C3%A9cnicas%20de%20miner%C3%ADa%20de%20datos%20en%20redes%20sociales.pdf>

RANTES GARCÍA, Mónica – Detección de Fraude Usando Tecnicas de Clustering. Lima – Peru : UNIVERSIDAD NACIONAL MAYOR DE SANMARCOS, 2010 [Revisado en 06 Septiembre de 2017] disponible en: http://repositorio.uchile.cl/bitstream/handle/2250/130479/cf-montesinos_lg.pdf

RIOS VILLEGAS, Andrés Felipe – Minería de datos aplicada a la detección de clientes con alta probabilidad de fraudes en sistemas de distribución. Pereira - Colombia: UNIVERSIDAD TECNOLOGICA DE PEREIRA, 2013 [Revisado en 04 Septiembre de 2017] disponible en: <http://repositorio.utp.edu.co/dspace/handle/11059/3856>