



UNIVERSIDAD
SAN IGNACIO
DE LOYOLA

FACULTAD DE CIENCIAS EMPRESARIALES

Carrera de Marketing

**CUSTOMER ENGAGEMENT EN INSTAGRAM,
CASO DE LAS IMÁGENES PUBLICADAS DESDE
LAS CUENTAS DE CENTROS ESTÉTICOS EN
LIMA**

**Tesis para optar por el Título Profesional de Licenciado en
Marketing**

URSULA GIOVANNA CHIPOCO TORRES

(0000-0002-1321-7765)

Asesor:

MG. MARÍA JOSÉ ROZAS PERALTA

(0000-0001-6581-5196)

Lima- Perú

2020

Dedicado a mi querida abuela Lula
Siempre estarás en mi corazón

Contenido

Resumen	10
Abstract.....	12
Introducción.....	14
1. Problema de Investigación	15
1.1. Planteamiento del Problema	15
1.2. Formulación del Problema.....	19
1.3. Justificación	19
2. Marco Referencial	22
2.1. Antecedentes	22
2.2. Marco Teórico.....	25
2.2.1. Redes Sociales y el marketing.....	25
2.2.2. <i>Customer Engagement</i> y el <i>Business Intelligence</i>	28
2.2.4. Instagram y Marketing 4.0.....	30
2.2.5. Instagram en el Perú	33
2.2.6. Centros Estéticos e Internet en el Perú	36
2.2.7. Servicios de Centros Estéticos.....	40
2.2.8. Analítica de imágenes.....	41
2.2.9. Modelo de Referencia.....	41
2.3. Objetivos.....	42

2.4.	Hipótesis	43
3.	Método	46
3.1.	Tipo y Diseño de Investigación	46
3.1.1.	Tipo de Investigación	46
3.1.2.	Diseño de Investigación.....	46
3.2.	Variables	48
3.3.	Instrumentos.....	49
3.3.1.	<i>Social Blade Analytics</i>	49
3.3.2.	<i>Vision AI Cloud</i> de Google.....	49
3.3.3.	<i>Every Pixel</i>	49
3.3.4.	<i>Eyeem Vision</i>	50
3.3.5.	<i>Microsoft Excel</i>	50
3.3.6.	<i>SPSS Modeler</i>	50
3.3.7.	<i>Downloader for Instagram</i>	51
3.4.	Universo.....	51
3.5.	Muestra	51
3.6.	Recolección de datos	53
3.7.	Plan de análisis.....	53
3.7.1.	Regresión Lineal Múltiple	53
3.7.2.	Red Neuronal.....	55
3.7.3.	Comparación entre modelos	59
4.	Calendario de Actividades y Recursos Disponibles.....	61

4.1. Calendario de Actividades	61
4.2. Recursos Disponibles.....	62
4.3. Presupuesto	62
5. Resultados	63
5.1. Resultados.....	63
5.2. Discusión	75
6. Conclusiones y Recomendaciones	77
6.1. Conclusiones.....	77
6.2. Recomendaciones	80
7. Limitaciones e Investigaciones futuras	85
7.1. Limitaciones.....	85
7.2. Investigaciones Futuras.....	86
Bibliografía.....	88

Índice de Figuras

Figura 1: Evolución del interés hacia Instagram en tópicos de Belleza y Salud según Google (junio 2013-junio2018).....	17
Figura 2: Interés hacia la red social Instagram en el Perú – 2018 de acuerdo a Google	34
Figura 3: Evolución del interés hacia Instagram en territorio peruano (setiembre 2014-setiembre 2019)	35
Figura 4: Interés hacia los Consultorios Estéticos en el Perú – 2018 de acuerdo a Google	37
Figura 5: Crecimiento en el interés hacia centros estéticas en Perú de acuerdo a Google	38
Figura 6: Distribución mensual hacia el interés generado por centros estéticos en Perú de acuerdo a Google en periodo 2013-2018.....	39
Figura 7: Horarios populares y tiempo de permanencia promedio en centros estéticos Lima en.....	39
Figura 8: Tratamientos estéticos que mayores consultas han generado en Lima durante el 2018 de acuerdo a Google	40
Figura 9: Esquema de procesamiento de datos a partir de imágenes	53
Figura 10: Red Neuronal Modelo Perceptrón Multicapa	58
Figura 11: Importancia de las variables predictoras dentro del Modelo de Regresión Lineal (en porcentaje).....	67
Figura 12: Arquitectura de la red neuronal.....	70
Figura 13: Importancia de las variables predictoras dentro del Modelo basado en una red neuronal (en porcentaje).....	72

Figura 14: Comparativo de la importancia dentro de las variables predictoras dentro de los modelos de regresión línea (izquierda) versus red neuronal (derecha) (en porcentaje)

..... 74

Índice de Tablas

Tabla 1: Niveles del Customer Engagement	30
Tabla 2: Variable de Estudio	48
Tabla 3: Características de la muestra	52
Tabla 4: Calendario de Actividades.....	61
Tabla 5: Resumen del modelo Step Forward	63
Tabla 6: ANOVA para los modelos Step Forward.....	64
Tabla 7: Análisis de coeficientes para modelos de regresión.....	65
Tabla 8: Análisis de colinealidad de la regresión lineal	67
Tabla 9: Análisis de autocorrelación entre residuos.....	68
Tabla 10: Test de Normalidad	68
Tabla 11: Resultados de campo para el modelo de regresión lineal.....	69
Tabla 12: Resumen del modelo basado en redes neuronales.....	69
Tabla 13: Estimaciones de parámetros en la red neuronal	71
Tabla 14: Resultados de campo para el modelo basado en redes neuronales.....	72
Tabla 15: Comparativo de resultados entre la regresión lineal y la red neuronal	73

Índice de Anexos

Anexo 1: Interface de Social Blade	98
Anexo 2: Uso de Downloader for Instagram.....	99
Anexo 3: Interface de Eyeem Vision.....	100
Anexo 4: Interface de Every Pixel.....	101
Anexo 5: Interface de Google Cloud Visual AI para medir imágenes propias de la medicina	102
Anexo 6: Interface de Google Cloud Visual AI para medir la cantidad de colores	103
Anexo 7: Interface del SPSS Modeler.....	104
Anexo 8: Matriz de Consistencia.....	105

Resumen

El presente documento tuvo por objetivo relacionar al *Customer Engagement* (suma de *likes* y comentarios) en Instagram para el caso de imágenes publicadas desde las cuentas de centros estéticos en Lima con un conjunto de elementos presentes en la mencionada red social y ciertas características presentes en el contenido fotográfico.

Entre las variables predictoras fueron consideradas siete variables: el número de seguidores con el que contaba la cuenta al momento de subir la imagen, el hecho de que la misma pueda ser clasificada como parte de un procedimiento vinculado a la medicina, la cantidad de días que ha permanecido en Instagram, la cantidad de palabras que se incluye en la foto, el puntaje asociado a la belleza de la imagen, el score vinculado al hecho de que sea asombrosa y la cantidad de colores.

Fueron empleados dos métodos para probar el modelo: la regresión lineal múltiple y la red neuronal. Con el primer algoritmo, fueron cinco las variables con aporte significativo; mientras que en la red neuronal fueron consideradas todos los elementos predictores. Las variables que no aportaban significativamente a la regresión fueron la cantidad de colores y el puntaje correspondiente al atributo asombroso. En cuanto a la precisión de ambos modelos, la regresión registro una mayor precisión que la red neuronal (85.5% vs 85.1%). No obstante, el modelo regresivo no cumple todas las condiciones de validez: los residuos no superaron la prueba de Kolmogórov-Smirnov. La muestra analizada fueron 491 casos.

Las recomendaciones por el lado del marketing están orientadas a la gestión de indicadores, decisiones por parte de los *community manager* y decisiones que pueden tomarse con el modelo propiamente dicho. Cabe mencionar que la analítica de imágenes, uno de los campos de desarrollo del *Data Mining*, fue uno de los pilares que permitió convertir las fotos en indicadores numéricos.

Palabras Clave: *Customer Engagement*, Instagram, Red Neuronal, Regresión Lineal, Centros Estéticos, Lima, Minería de Datos.

Abstract

The purpose of this document was to relate the Customer Engagement (sum of likes and comments) to Instagram for the case of images published from the accounts of aesthetic centers in Lima with a set of elements present in the aforementioned social network and certain features present in the photographic content.

Among the predictor variables were considered seven variables: the number of followers that the account had at the time of uploading the image, the fact that it can be classified as part of a procedure linked to the medicine, the number of days it has remained on Instagram, the amount of words included in the photo, the score associated with the beauty of the image, the score linked to the fact that it is amazing and the amount of colors.

Two methods were used to test the model: multiple linear regression and the neural network. With the first algorithm, there were five variables with significant input; whereas in the neural network, all the predictor elements were considered. The variables that did not significantly contribute to regression were the number of colors and the score corresponding to the staggering attribute. As for the accuracy of both models, the regression registers a greater accuracy than the neural network (85.5% vs 85.1%). However, the regressive model does not fulfil all the conditions of validity: the residues did not pass the Kolmogórov-Smirnov test. The sample analyzed was 491 cases.

The recommendations on the marketing side are oriented to the management of indicators, decisions by the community managers and decisions that can be taken with the model itself. It is worth mentioning that image analytics, one of the Data Mining's development fields, was one of the pillars that allowed the conversion of photos into numeric indicators.

Keywords: *Customer Engagement, Instagram, Neuronal Network, Linear Regression, Aesthetics Centers, Lima, Data Mining.*

Introducción

La red social Instagram es una de las de mayor crecimiento a nivel global. Actualmente, la tendencia de compartir contenido visual en línea crece. Las empresas son conscientes de ello y por ello están empezando a compartir cada vez más contenido visual. Es racional creer que una empresa que tiene una cuenta en Instagram desea tener más interacciones, una cantidad considerable de likes y comentarios es seguramente lo que todo *community manager* en Instagram desea al momento de compartir una imagen en la mencionada red social. En la literatura académica, la adición de likes y comentarios es lo que conoce como *Customer Engagement*.

¿De qué depende el *Engagement*? La cantidad de seguidores es un elemento predictor a considerar, pero deben existir más. Existen datos que no resultan fáciles de medir con métodos como la observación directa: el número de colores de una imagen o el nivel de estética de la misma, por ejemplo, resultan indicadores que pueden explicar el número total de comentarios y *likes* que tendrá una publicación fotográfica en Instagram. Actualmente, analizar el contenido de una imagen es una tarea que puede ser llevada con software especializado.

Cada vez existe un mayor nivel de interés (traducido en búsquedas o consultas) a los temas de belleza y salud en territorio peruano. Por ello, algunos consultorios estéticos participan en Instagram, buscan reacciones a las imágenes que suben, eventualmente alguna de esos me gusta o comentario puede traducirse en una venta. Precisamente el presente estudio va orientado a modelar el *Customer Engagement* en función a las variables abarcadas por otros investigadores en el caso de los centros estéticos en Lima.

El presente documento de investigación se encuentra dividido en siete capítulos. En el primero se plantea el problema de investigación. El segundo incluye aspectos teóricos (antecedentes y marco conceptual) que delimitan la tesis con énfasis en el *Customer Engagement*. El tercer capítulo es el método, sección donde se describe a las variables de estudio, así como la operacionalización de las mismas; la sección termina describiendo los dos algoritmos empleados: la regresión lineal múltiple y la red neuronal. En el cuarto capítulo se muestran los recursos empleados y el cronograma de desarrollo de la investigación. El quinto capítulo abarca los resultados del modelo propuesto. El capítulo seis incluye las conclusiones y recomendaciones. Finalmente, el capítulo siete se limita a mostrar las limitaciones de la investigación, así como investigaciones futuras que pueden ser consideradas por otros investigadores.

1. Problema de Investigación

1.1. Planteamiento del Problema

Quienes han hecho uso de redes sociales como Instagram, Facebook u otras, deben al menos intuir que parte del contenido que se muestra en las páginas de inicio es parte de un algoritmo que obedece tanto a la propia actividad del usuario como a quienes tiene agregado en su red social. En base a ello funcionan las redes sociales. Si por, ejemplo, dos personas empiezan a conversar acerca de salir a comer una pizza, el algoritmo de la red social empezará a mostrar anuncios de pizzerías. Esto no debe sorprender, es simplemente ciencias de la computación aplicada al marketing. En espacios digitales, una empresa quiere incrementar su cantidad de interacciones en redes sociales. Quienes comenten, den un “me gusta”, reproduzcan un contenido o compartan el mismo, tienen al menos conciencia que la marca o empresa que subió el contenido

existe. No necesariamente todos son consumidores finales, pero al menos sus interacciones son útiles para que el contenido se difunda.

Actualmente, los individuos que acceden a la internet comparten mucho contenido en las redes sociales en forma de imágenes, ya sean personales, escenas del día a día, o sus opiniones representadas en forma de caricaturas (Reece y Danforth, 2016). Inspeccionar contenido como este de las redes sociales en sitios web como Flickr, Twitter, Tumblr, etc., puede proporcionar una idea de cómo se siente una persona frente a un contexto específico (Caldeira, 2016) .

Instagram es una aplicación móvil (con versión compatible en PC) utilizada para compartir fotos que permite a los usuarios tomar fotografías, aplicar filtros y compartirlas en la plataforma misma, así como en otras plataformas como Facebook y Twitter (Eagar y Dann, 2015). Según el sitio web de la compañía, Instagram tiene más de 800 millones de usuarios mensuales activos que dieron un total de 14,000 millones de *likes* durante el 2018 (M4rketing E Commerce, 2019)

Las empresas se han dado cuenta de la importancia de Instagram, en especial aquellas que tienen una alta dependencia del componente visual (Dugué, 2017). De ahí que empresas del rubro moda, maquillaje, viaje y similares estén muy interesadas en posicionarse adecuadamente (Vignisdóttir, 2016). El sector de la cosmética también reconoce la importancia de las redes sociales, y dentro de ella está desde luego los centros orientados a la consultoría estética con supervisión médica. En el Perú, Instagram ha incrementado su popularidad, y los tópicos relacionados a belleza y salud también. De eso queda evidencia en las estadísticas de internet.

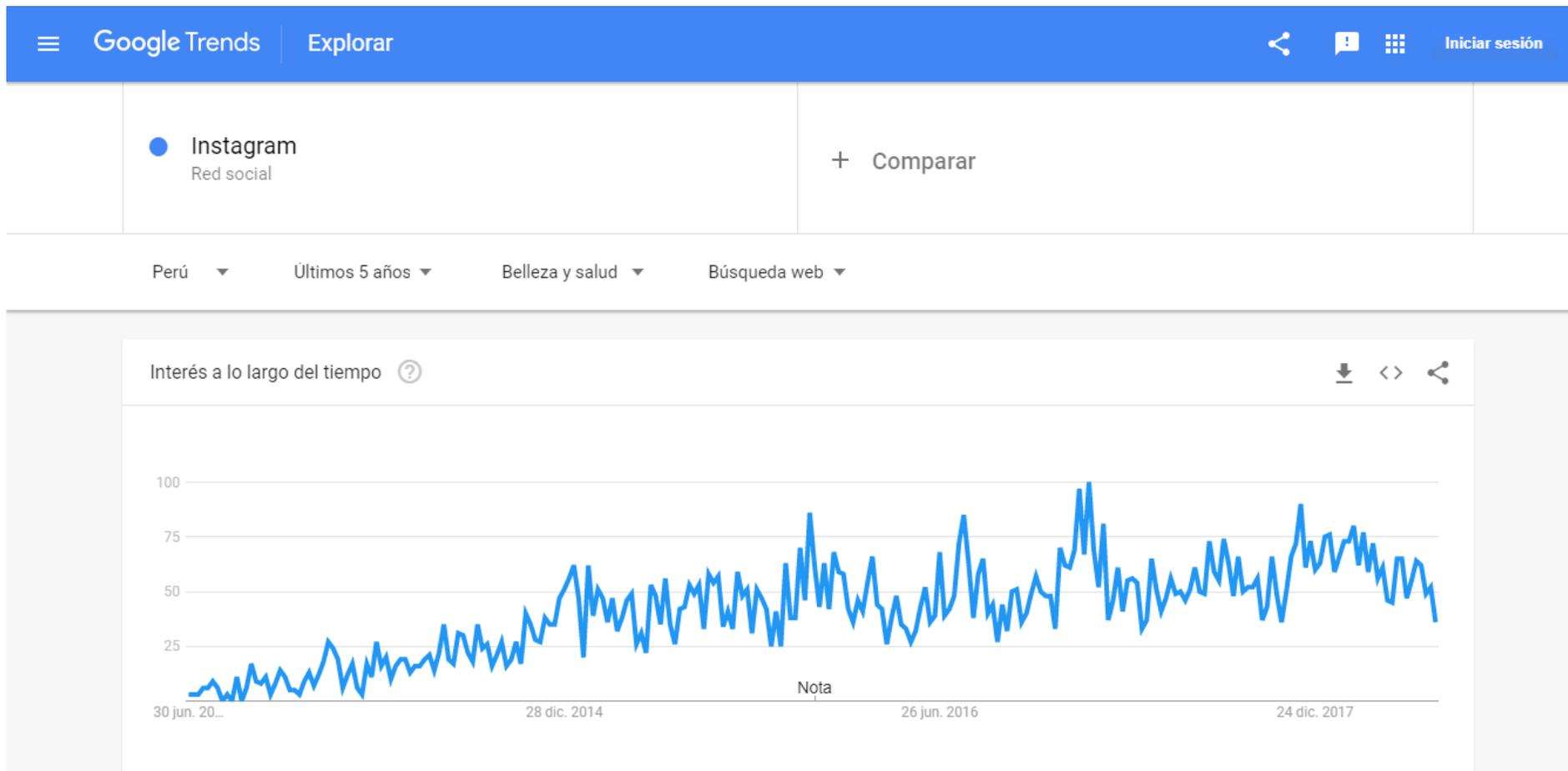


Figura 1: Evolución del interés hacia Instagram en tópicos de Belleza y Salud según Google (junio 2013-junio2018)

Fuente: Google Trends (2018)

El anterior gráfico no hace sino evidenciar que en el periodo 2013-2018 se ha dado un incremento en cuanto a las búsquedas en Instagram de contenido relacionado a la belleza y salud. No es una afirmación gaseosa: son hechos fácilmente verificables en la serie de tiempo presentada en la anterior figura.

Los consultorios estéticos son centros dedicados al cuidado de la apariencia personal con soporte médico. Ofrecen tratamientos de distintos tipos: capilar, cuidado de piel, etc. En un ambiente natural de plena competencia, han optado por ingresar a las redes sociales como parte de sus tácticas de marketing a fin de mostrar sus productos. Naturalmente, al participar en redes sociales, lo que buscan es generar más reacciones de los clientes como likes o comentarios.

Y es que Instagram también es usado en el campo de la salud y medicina (Kamel, Giustini y Wheeler, 2016). De hecho, las imágenes en Instagram orientadas a la salud buscan mostrar a las personas una situación en la que se sienten mejor con ellas mismas (por ejemplo, los centros de nutrición muestran imágenes de personas que han logrado bajar de peso). En el caso de los consultorios estéticos se procuran mostrar imágenes de los tratamientos, las promociones, casos de éxito médico entre otras (Dorfman, Vaca, Mahmood, Fine y Schierle, 2018).

El *Customer Engagement* es una de los indicadores más importantes de las redes sociales. Al referirse al *Engagement* en Instagram, se hace referencia al grado de fidelización por parte de tus seguidores en esta red social (Eagar y Dann, 2015). No se trata únicamente del número de seguidores que tienes una cuenta, sino del grado de implicación, interés e interacción que los seguidores tienen con tus fotografías, vídeos e historias. La interacción puede ser medida de diferentes formas, una de ellas es a través del número de *likes* y cantidad

de comentarios de un contenido (Vignisdóttir, 2016). El *Customer Engagement* de una publicación puede depender de una serie de variables.

En Instagram existen una serie de métricas que son relativamente fácil de recopilar. Se encuentran, por ejemplo, la cantidad de seguidores que tiene una determinada cuenta y la fecha exacta en la que fue publicada. No obstante, existen también características propias de las imágenes que no son resulta fáciles de medir. Bien podría cuestionarse un investigador si la imagen resulta visualmente atractiva o asombrosa, o tal vez si se trata de una imagen vinculado a procesos médicos o si incluye una cantidad de colores adecuada. Precisamente la investigación pretende aproximarse a un modelo que relacione al *Customer Engagement* y el grupo de métricas ya mencionadas.

1.2. Formulación del Problema

¿Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima?

1.3. Justificación

En cuanto a un potencial aporte práctico la tesis, se pretende construir un modelo predictivo que se aproxime al número de interacciones que una publicación debe alcanzar en la red social Instagram. Esto a fin de determinar si un contenido ha registrado un adecuado *Customer Engagement* dentro de Instagram dado un conjunto de condiciones. Si, por ejemplo, una fotografía, de acuerdo al modelo, debe registrar 400 *likes* y solo registró 100, significa que existen condiciones adicionales a tomar en cuenta. La idea básica es brindar a las empresas orientadas a la consultoría estética una herramienta para el marketing online: determinar qué

elementos hacen más probable que una fotografía colgada en Instagram tenga un más alto *Customer Engagement*.

Debe tomarse en consideración que los modelos predictivos de interacciones web no han sido objeto de estudios en la producción académica peruana, al menos no al cierre de la edición de la presente investigación. Para ello basta revisar los repositorios digitales de las universidades peruanas o portales como *Web of Science* o *Scopus*.

Esta investigación incorpora el análisis de imágenes con software de muy fácil uso al ámbito del marketing peruano. El análisis automatizado de fotos es incipiente incluso a nivel de *journals* orientados a negocios (Szeliski, 2010). Los pocos antecedentes que existen empleando el análisis de fotografías por computadora se encuentran dentro del campo de la ingeniería, ciencias políticas y psicología (Forsyth y Ponce, 2010). Por tanto, la investigación pretende servir como arquetipo para quienes deseen emplear este tipo de recursos dentro de sus trabajos de investigación, más aún en un contexto donde la tasa de crecimiento de contenido en formato de imágenes ya linda con lo exponencial (Szeliski, 2010). El extraer data a partir de imágenes es un proceso que únicamente podía ser llevado a cabo por profesionales vinculados a las ciencias de la computación e ingeniería, pero actualmente existe la imperiosa necesidad de aprovechar toda esa cantidad de datos en formato de imagen, los cuales requieren ser transformado para que sean aprovechables. Esto es particularmente relevante en un contexto donde la *Big Data* y la Minería de Datos tienen cada vez mayor relevancia como soporte para los procesos de *Business Intelligence*.

Finalmente, debe tomarse en consideración que la tesis presenta una breve e introductoria comparativa entre dos modelos predictivos: la regresión lineal múltiple y la red neuronal. Ambos modelos son de uso bastante recurrente en publicaciones indexadas en revistas científicas. Comparar ambos modelos, entonces, representa un aporte a la discusión

sobre las ventajas y desventajas de usar cada uno. Por otro lado, debe tomarse en cuenta que en el ámbito peruano de las investigaciones relacionadas a las ciencias administrativas no es de muy frecuente uso el empleo de redes neuronales, método más próximo a las ciencias de la computación que a la inferencia estadística.

El trabajo de investigación resulta útil para gestores de contenido y *community managers* que tomen decisiones acerca de la naturaleza de las imágenes que son compartidas en Instagram. El contenido de la investigación también resulta útil para quienes estén interesados en el estudio de métricas online de redes sociales. Finalmente, la investigación resulta de utilidad para quienes estén interesados en el aprovechamiento de las imágenes como *input* para modelos estadísticos.

2. Marco Referencial

2.1. Antecedentes

Salem (2012) estudió a los grupos de Facebook que se hicieron específicamente para criticar a las fotografías de los usuarios de Hi5. La investigadora, mediante el análisis de contenido, determinó que los miembros de los grupos en Facebook mencionaban que las fotos en Hi5 eran pobres en estética, abusaban de los colores e incorporaban contenido textual que era visto como “muy cholo”. Si bien es cierto este estudio se hizo dentro del marco de la antropología visual, la autora deja abierta la posibilidad de que se hable ya de una segmentación visual en cuanto a las fotografías en redes sociales. Por otro lado, los hallazgos de la investigadora sugieren que una foto que tenga mucho texto, un rasgo muy frecuente en Hi5, tiende a ser menos atractiva visualmente para los usuarios de Facebook.

Yu (2014) basó su investigación en el análisis de 238 imágenes creadas por los clientes y 62 imágenes creadas por las empresas Nike y Starbucks, todo dentro del entorno de Instagram. La investigadora utilizó un método híbrido de análisis de contenido cualitativo y análisis temático como una forma empírica de explorar el panorama general de este ámbito. La información se recopiló a través de *Keyhole* y *Statigram*, dos herramientas analíticas en línea de redes sociales. El estudio demostró que la práctica de co-creación de valor de marca de los clientes en Instagram se centra en gran medida en la autoexpresión de los individuos. El estudio también descubrió temas valiosos sobre el uso de *selfies* y el sarcasmo autodirigido entre los *instagrammers* (usuarios de Instagram) que comparten imágenes vinculadas a una marca. La investigadora destaca, por último, la importancia de la estética de una fotografía como elemento necesario para el Engagement.

Tatar, Dias de Amorin, Fdida y Antoniadis (2014) orientaron su estudio a analizar distintos modelos que pronosticaban la popularidad de una publicación online. Los autores evidenciaron que en redes sociales como YouTube existían dos variables predictoras especialmente importantes que tenían que ser consideradas para estimar la cantidad de *likes* y comentarios (indicadores asociados a la popularidad de un contenido en la web): el número de suscriptores y la cantidad de días que el contenido ha permanecido online. Entre las opciones de modelado, los investigadores recomendaron el uso de las redes neuronales como uno de los métodos de mayor uso en la predicción de indicadores vinculados a redes sociales.

Qyan, Tang, Penza, y Ferri (2015) utilizaron modelos de regresión para predecir el número de *likes* y comentarios que recibiría una fotografía en Instagram. Los autores entre otros aspectos señalaron que el número de *followers* estaba relacionado directamente con los *likes* y comentarios que obtenía una foto. Por otro lado, indicaron que la cantidad de colores que empleaba una fotografía no guardaba una relación significativa con el número de interacciones generadas. Entre otras sugerencias, sugiere que se introduzca el nivel de atractivo de una fotografía como variable independiente del modelado.

Vignisdóttir (2016) realizó un estudio experimental para determinar el *Customer Engagement* (suma de la cantidad de comentarios y *likes*) de una publicación en Instagram. Luego de un análisis bibliográfico, determinó que el *Customer Engagement* puede determinarse como la cantidad de *likes* más comentarios que alcanza una publicación. Construyó un modelo predictivo en base a 750 fotografías de marcas cosméticas. Entre otros hallazgos determinó que existía una correlación positiva entre el *Customer Engagement* y el número de *followers*. La autora sugiere, entre otras cosas, realizar un análisis del contenido de las fotografías.

Hausman y Toivonen (2017) analizaron un total de 9,059 fotografías de Instagram y 4,558 fotografías de Flickr del *Kruger National Park* (Sudáfrica). Ellos buscaban determinar qué preferían más los viajeros. Los resultados principales fueron que los mamíferos de gran tamaño eran el grupo preferido en ambas plataformas. Sin embargo, se descubrió que Flickr coincidía mejor con la preferencia de los turistas por la biodiversidad menos carismática. Los hallazgos sugieren que el contenido de las redes sociales se puede utilizar como una forma rentable de explorar, y para un monitoreo más continuo de las preferencias por la biodiversidad y las actividades humanas en áreas protegidas. Los autores también indicaron que los escenarios que eran fotografiados con mayor frecuencia eran aquellos que podrían ser calificados como los que ocasionaban mayor nivel de asombro.

Gauthier, Bratberg, Loi y V DiVall (2018) analizaron el rol que cumplía Instagram entre los estudiantes de ciencias de la salud. Los investigadores analizaron a una *fanpage* de contenido médico en Instagram en donde 2550 seguidores de todo el mundo se involucraron con el contenido. La evidencia indica que Instagram ofrece la oportunidad de atraer a una audiencia multidisciplinaria (estudiante de medicina, enfermería, obstetricia, y similares), enriquecer el aprendizaje en línea de los estudiantes y permitir compartir conocimientos e ideas interdisciplinarios. Conclusión: usar Instagram para publicar notas ilustrativas, genera *Engagement*. Los investigadores sugirieron investigar a otros grupos interesados en tópicos médicos debido a que era probable que el hecho de que una imagen pueda ser calificada como propia de un procedimiento médico podría interesar a otros segmentos

Lay y Fewerda (2018) tuvieron por objetivo construir un *framework* que permitiera modelar las personalidades de los usuarios de Instagram en función a los likes que estos daban. Luego de una revisión de casos llegan a determinar que existía una relación entre el hecho de poner like a una foto y el nivel de estética de la misma. Dicho de otra forma, dar un like tenía relación con el atractivo visual del contenido. Entre las implicaciones teórico-prácticas,

mencionan que el atractivo de una fotografía puede servir para predecir la cantidad de likes que obtendría una fotografía.

2.2. Marco Teórico

2.2.1. Redes Sociales y el marketing

Las aplicaciones de los medios sociales, como los blogs, se pueden remontar a la década de 1990, pero han ganado popularidad después de la web 2.0 (Fuchs, 2014). Los medios sociales pueden definirse como “herramientas que aumentan nuestra capacidad de compartir, cooperar entre nosotros y tomar medidas colectivas, todo fuera del marco de las instituciones y organizaciones institucionales tradicionales” (Fuchs, 2014). Sudha y Sheena (2017) mencionaron que el marketing digital se puede identificar como una descripción de tipo paraguas, que retrata la noción de publicitar ciertos servicios o artículos a través de una esfera de tecnologías que son digitales. La mayoría de las veces, esto se lleva a cabo a través de Internet, lo cual implica hacerlo mediante, por ejemplo, la publicidad, así como a través de teléfonos móviles (Sudha y Sheena, 2017).

Las transformaciones que se han producido en relación con el marketing digital se dieron a partir de la década de 1990, aunque fueron más notorias en la década de los 2000, propiciando un cambio en donde las empresas y marcas deciden crear su marketing, mediante el uso de tecnología y opciones digitales relacionadas (Sudha y Sheena, 2017). Además, las redes sociales pueden vincularse con contenido generado por el usuario (UGC por sus siglas en inglés). Una nueva descripción en el mundo

digital comenzó en 2005, cuando Tim O’Riley recalcó los roles de la web 2.0: como usuarios, como creadores y participantes (Fuchs, 2014).

Para las empresas, lo anterior, significó interacciones y compromiso entre los consumidores y las marcas, que se reconfiguraron para una comunicación bidireccional en lugar de la comunicación unidireccional anterior que caracterizó a la web 1.0 (Cosh, Burns y Daniel, 2008). Además, la principal diferencia entre la web 1.0 y la web 2.0 es que le brinda al usuario una experiencia más rica con nuevas aplicaciones como Blogs, Google y Wikipedia (Fuchs, 2014). Además, Hinton y Horta (2013) enfatizaron en la producción creativa dado el contexto de la web 2.0, lo que hace que sea sencillo para cualquier persona que quiera publicar cualquier cosa en Internet. Con respecto al aspecto de los consumidores, la *American Marketing Association* afirma que un consumidor es una persona que consume productos y servicios, así como también ideas (Association American Marketing, 2018). Además, el consumidor no solo tiene que ser el que consume los bienes, sino, también puede dirigirse al comprador, que se conoce como el consumidor final (Association American Marketing, 2018)

Dessart, Veloutsou y Morgan-Thomas (2016) define el compromiso del consumidor en el contexto del marketing, como un estado cognitivo y emocional que tiene lugar a través de experiencias mutuas, interactivas y producidas conjuntamente. Además, las experiencias del consumidor, a menudo se crean con una marca para establecer relaciones. Adicionalmente, al conectar esto con las redes sociales, Zolkepli, Hasno y Mukhiar (2015) indicaron que el compromiso es un concepto formado por varios aspectos, como crear y publicar contenido en línea, que debe ser apropiado y brindar un significado a los consumidores (Zolkepli, Hasno y Mukhiar, 2015). Además, los consumidores pueden compartir y modificar el contenido para despertar la sensación

de estar en una comunidad que está en línea (Zolkepli, et.al. 2015). Por otro lado, la participación de los consumidores en un contexto *online*, también se trata de tener un diálogo en lugar de una comunicación de una sola vía (Zolkepli, et.al. 2015).

Hassan (2014) menciona que el marketing en redes sociales está estrechamente relacionado con el marketing boca a boca y el marketing viral, con el objetivo de influir en los consumidores. El procedimiento del marketing en redes sociales se relaciona con una forma de estrategia para permitir un mayor nivel de tráfico con respecto a un determinado sitio web o para permitir un mayor nivel de atención de los compradores (Sudha y Sheena, 2017). Adicionalmente, Sudha y Sheena (2017) explican que la construcción a través de varias opciones de canales es lo que permite el marketing digital.

Además, según Evans, Phua, Lim y Jun (2017), las empresas utilizan diferentes enfoques en las aplicaciones de redes sociales para interactuar con los consumidores, como el boca a boca electrónico (eWOM), los anuncios pagados y las publicaciones patrocinadas. La expresión eWOM puede definirse como "cualquier declaración positiva o negativa hecha por clientes potenciales, reales o anteriores sobre un producto o compañía, que se pone a disposición de una multitud de personas e instituciones a través de Internet" (Henning-Thurau, Gwinner, Walsh y Gremler, 2004). Por último, según Byrne, Kearney y MacEvilly (2017), el aspecto del concepto de marketing que tiene lugar a través del boca a boca se puede vincular con el concepto de marketing de influencia. Esto ocurre en el sentido de que el marketing influyente puede considerarse como una versión de boca en boca, pero en forma digital (Byrne, Kearney y MacEvilly, 2017).

Sudha y Sheena (2017) afirman que el marketing influyente se ha posicionado como un elemento vital en relación con la orientación que es digital. De acuerdo con esto, De Veirman, Cauberghe y Hudders (2017) mencionan como las marcas se proponen ganar *influencers* que estén dispuestos a respaldar los diversos elementos de la marca, permiten a la marca, la posibilidad de construir la imagen que desean comunicar. Según lo establecido por De Veirman et.al. (2017), alguien que tiene la capacidad de crear contenido y al mismo tiempo tiene una base sólida de seguidores, puede considerarse un *influencer*. Además, esto ocurre a través de plataformas como Blogs o contenido que tiene un lapso más corto, como por ejemplo Instagram o Snapchat (De Veirman, et.al. 2017).

Por último, Chen (2018) menciona la singularidad de Instagram debido a sus características diferenciadas de uso por parte de los consumidores en comparación con Facebook y Twitter, otras plataformas de medios sociales, lo que también lleva a diferencias en la comercialización de Instagram y las percepciones de los consumidores acerca de la comunicación y el marketing. Por lo tanto, uno puede entender que cada plataforma es diferente.

2.2.2. *Customer Engagement* y el *Business Intelligence*

Los indicadores de participación del cliente en el marketing de redes sociales son una de las medidas más importantes para el marketing electrónico actualmente (Barger, 2016). La existencia y difusión de las redes sociales en los últimos años ha modificado el entorno de comercialización de las empresas tremendamente: nuevas oportunidades han sido creadas para que las marcas promocionen sus productos y desarrollen

interacciones con los clientes (Zhang, Guo y Hu, 2017). El impacto de las redes sociales sobre las marcas, ha conllevado al cambio de la forma de interactuar del consumidor, convirtiéndose en el principal medio de comunicación entre una la empresa y los clientes (Barger, 2016).

Las empresas redirigen su enfoque en crear una interacción directa del cliente con una marca y se centran en crear una comunidad. El termino *Customer Engagement* es usado como un concepto estratégico del marketing para definir el compromiso o vinculo de un cliente con una marca, un nivel alto de *Customer Engagement* implica un cliente feliz, con lo cual se obtiene un embajador o promotor de la marca (Barger, 2016). Una de las mejores prácticas de interacción es mediante el uso de las redes sociales, donde el cliente puede asumir el rol de defensor o embajador de la marca como muestra de su lealtad (Barger, 2016). Es innegable el efecto positivo de un cliente satisfecho. La concurrencia entre estos dos términos (*Customer Engagement* y redes sociales) se traduce en la transformación de la *Big Data* dentro de los procesos de *Business Intelligence*, información valiosa, obtenida gracias al Engagement que sienten los clientes hacia un producto, marca o servicio que se pueden medir a través de las reacciones: likes, comentarios positivos y/o negativos, entre otros (Kunz, Askoy, Bart y Theoloudilis, 2017). Esto, mediante el proceso de *Business Intelligence*, permite obtener información de los clientes, productos/ servicios y mercado (Stone y WoodCook, 2013). La recolección de la *Big Data* permitirá que el uso de redes sociales pueda aprovecharse para generar un análisis detallado de la empresa, así como un análisis predictivo para el apoyo en la toma de decisiones, la innovación y la mejora en la competitividad del negocio (Kunz, et.al. 2017).

2.2.3. Beneficios del *Customer Engagement* para personas y empresas, el caso del CRM

Cada interacción que tiene un cliente con la marca contribuye a la imagen de la marca frente a otros y apunta hacia el compromiso o *Engagement* que sienten los clientes frente a la marca (Barger, 2016). Los beneficios de la participación del cliente proporcionan un retorno real y prolongado para la organización; en razón, de ello la interacción que existe en las redes sociales es un pilar básico para la obtención de esta (Giannakis y Boutsouki, 2014).

Desde luego existen diferentes niveles de *Engagement* (Barger, 2016). Dicho de otra forma. Ello genera diferentes beneficios tanto para empresas como para personas, tal como se indica en la siguiente tabla:

Tabla 1: Niveles del Customer Engagement

Nivel de Engagement	Beneficios a las personas	Beneficios a las empresas
Consumidores: visitan y observan	El usuario recibe información que satisface su interés. Observa el contenido ofrecido por la marca.	Cumplimiento de su aspecto más importante al cumplir con la satisfacción y deseos del cliente.
Creadores: dan inicio, aportan y preguntan	Debido a la experiencia de los clientes, estos pueden responder consultas respecto de la marca, generando un <i>feedback</i> positivo entre clientes.	La empresa reconoce lo que le interesa al usuario y le permite crear una estrategia que permita el aumento de ventas.
Contenido Generado por el Usuario: hacen valoraciones sobre el contenido que se hace en redes sociales	Los usuarios sienten que la marca los valora al saber que su contenido influencia sobre la decisión de otros usuarios.	Conocimiento de la empresa sobre la experiencia positiva de otros usuarios, y como esta influye en la decisión de otros a través de sus recomendaciones
Defensores: son creadores y recomiendan	Sabe que su lealtad es reconocida por la empresa y la influencia que tiene frente a los otros usuarios.	La empresa cuenta con firmes defensores y cuenta con que su propuesta comercial se expanda. Nuevas propuestas señaladas por los clientes que le permite innovar.

Fuente: Nebi (2017)

2.2.4. Instagram y Marketing 4.0

Instagram es una aplicación basada en contenido visual que permite a sus usuarios (llamados *instagrammers* de forma informal) compartir solo contenido basado en

vídeos o fotos (Hausman y Toivonen, 2017). El contenido es editable y los usuarios pueden agregar filtros y complementos (animaciones, subtítulos, etiquetas) hasta llegar a un estilo propio (Terttunen, 2017). Los usuarios pueden interactuar a través de comentarios y “Me gusta”, y también tiene incorporado un servicio de mensajería instantánea (Eagar y Dann, 2015).

Instagram tiene ventajas: está plenamente unido a otras plataformas: los usuarios pueden compartir sus fotos y vídeos directamente a otras redes sociales plataformas, al mismo tiempo que se publican en Instagram, el caso de Facebook probablemente sea el máximo referente (Reece y Danforth, 2016). Por otro lado, existen algunos estudios sobre qué tipo de contenido es popular en Instagram, popularidad que debe entenderse como un alto nivel en cuanto a interacciones; no obstante, no hay estudios acerca de qué tipo de contenido es popular en Instagram en la industria de la belleza y salud.

Las empresas recurren cada vez usan más a Instagram como canal de marketing (Felix, Rauschnabel y Hinsch, 2017). Por este motivo, Instagram diseñó perfiles para negocios específicos, similares a los perfiles de usuarios regulares (Yu, 2014). Naturalmente existen diferencias, la información que las organizaciones pueden obtener y la posibilidad de publicitar sus publicaciones. Instagram permite a las organizaciones conocer mejor a sus seguidores y aprende de ellos (Forbes, 2016). Las marcas pueden obtener información sobre sus seguidores a nivel demográfico y una descripción general de los mensajes más populares (Yu, 2014).

Instagram, entonces, es una plataforma que permite visualizar cómo las empresas están construyendo personalidades de la marca y cómo los clientes se identifican con estas (Lin, Lu y Wu, 2012). Con las redes sociales, los consumidores pueden no solo identificarse con las *Brand Personalities* creadas por las empresas:

también terminan asociando nuevos valores e ideas a marcas, lo cual contribuye a enriquecer las diferentes personalidades (Potts, et al. 2008). Las empresas necesitan saber cómo los clientes están asociando sus personalidades a sus marcas (Potts, et al. 2008).

Chen (2018) evidenció que Instagram es el lugar perfecto para la autoidentificación, el reconocimiento, el ocio y las interacciones sociales con amigos y celebridades o personas influyentes. Chen (2018) continúa sugiriendo que las investigaciones futuras en el mismo campo; es decir, las percepciones de los consumidores sobre marketing a través de Instagram, deberían descubrir dimensiones secundarias de los factores con respecto a cómo las personas perciben Instagram, o descubrir la apertura y aceptación de los consumidores hacia el marketing en Instagram.

Además, Ting, Ming, De Run, y Choo (2015) declararon que independientemente del uso de la aplicación, es necesario subrayar que la falta de conocimiento prevalece específicamente en los mercados que pueden percibirse como emergentes. Esto también está acorde con Djarfarova y Rushworth (2017) quienes explican que el interés por estudiar Instagram se debe a una investigación académica limitada, pero también porque los consumidores pasan más tiempo en Instagram que otras aplicaciones de medios sociales similares. Ello evidencia que existe un déficit dentro del comportamiento del consumidor en Instagram.

Según Chen (2018), las marcas y las empresas han implementado ampliamente las redes sociales e Instagram como herramienta de marketing, sin embargo, en el campo académico se están quedando atrás. Hasta hoy, las investigaciones académicas han consistido en estudios orientados a explorar y examinar el uso de Instagram y sus

aplicaciones al marketing (Chen, 2018). Además, Chen (2018) menciona dos vacíos de investigación dentro de la literatura académica. La primera brecha en las redes sociales e Instagram es el hecho de que la mayoría de los estudios realizados examinaron el uso estratégico de Instagram desde la perspectiva de una marca en una industria específica, en lugar de centrarse en la perspectiva del consumidor y su percepción e interés del marketing de Instagram. La segunda brecha enfatizada por Chen (2018) es que la mayoría de las investigaciones realizadas dentro del área, ha utilizado un método que muestra los hallazgos como más descriptivos. En cambio, es necesaria una comprensión más profunda, así como estudios sistemáticos sobre las percepciones de los consumidores acerca la comercialización en Instagram(Chen, 2018). Dado lo explicado anteriormente, es necesario defender la relevancia de estudiar el tema del *Engagement* en Instagram desde una perspectiva del consumidor.

2.2.5. Instagram en el Perú

De acuerdo a una cita brindada al Diario La República, se estima que, a inicios del 2018, en el Perú existían 4.2 millones de usuarios activos en Instagram (Vizcarra, 2018). Ello equivale a un 13% de la población (Vizcarra, 2018). Con respecto al género, el 53% de usuarios de Instagram son mujeres y el resto varones (Vizcarra, 2018).

Por otro lado, el *target* de la red social Instagram en el Perú son adolescentes y adultos jóvenes. Instagram es una de las redes sociales de mayor crecimiento en el Perú. Su atractivo radica en una tendencia global: compartir experiencias a través de imágenes. Otro factor a tomar en cuenta es el uso cada más difundido de smartphones (Vizcarra, 2018). Cabe mencionar que el uso de dicha red social ha sido descentralizado, tal como evidencia la data de Google Trends:

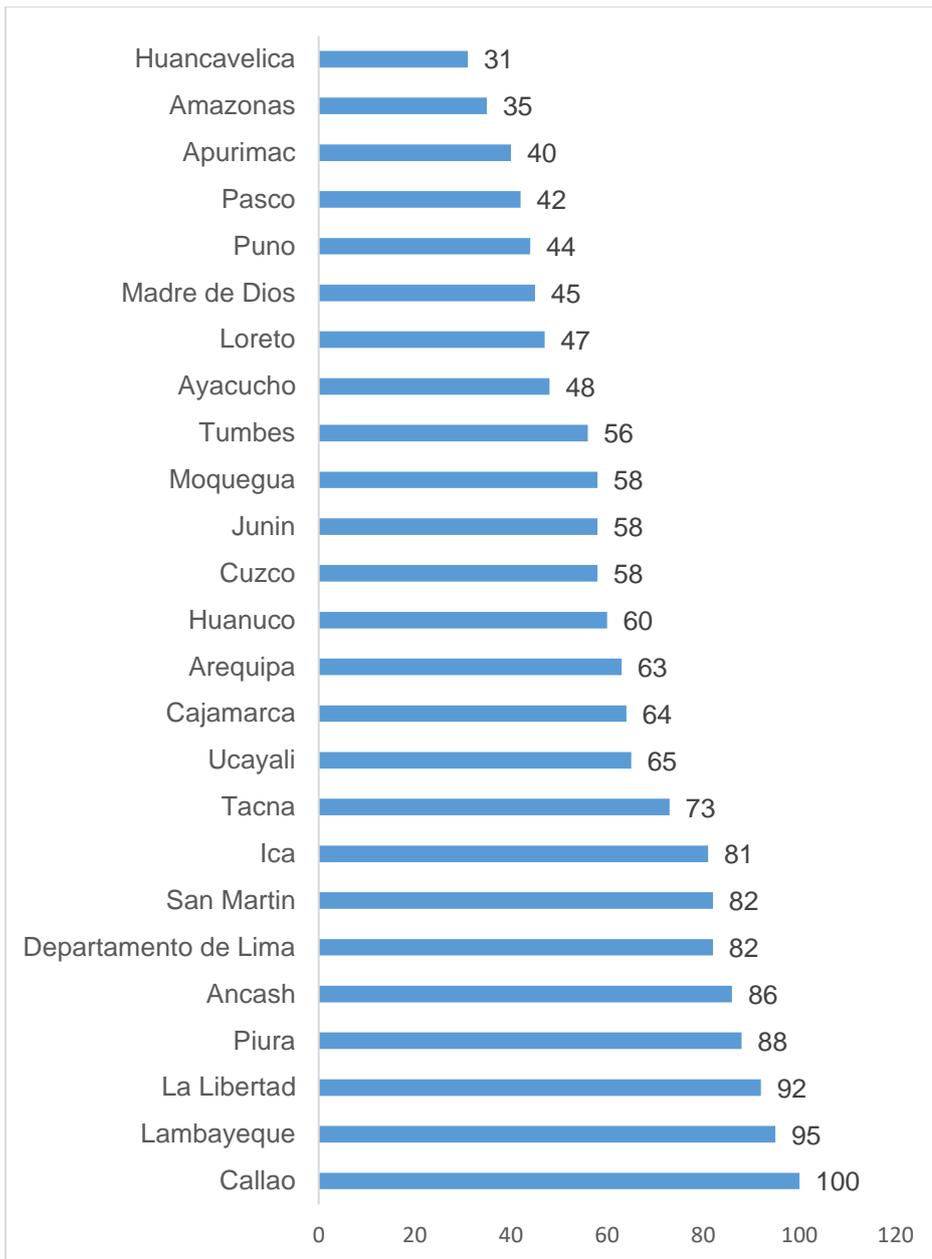


Figura 2: Interés hacia la red social Instagram en el Perú – 2018 de acuerdo a Google

Fuente: Google Trends (2019).

Por otro lado, con respecto al interés a lo largo del tiempo que ha generado la red social Instagram se tiene:

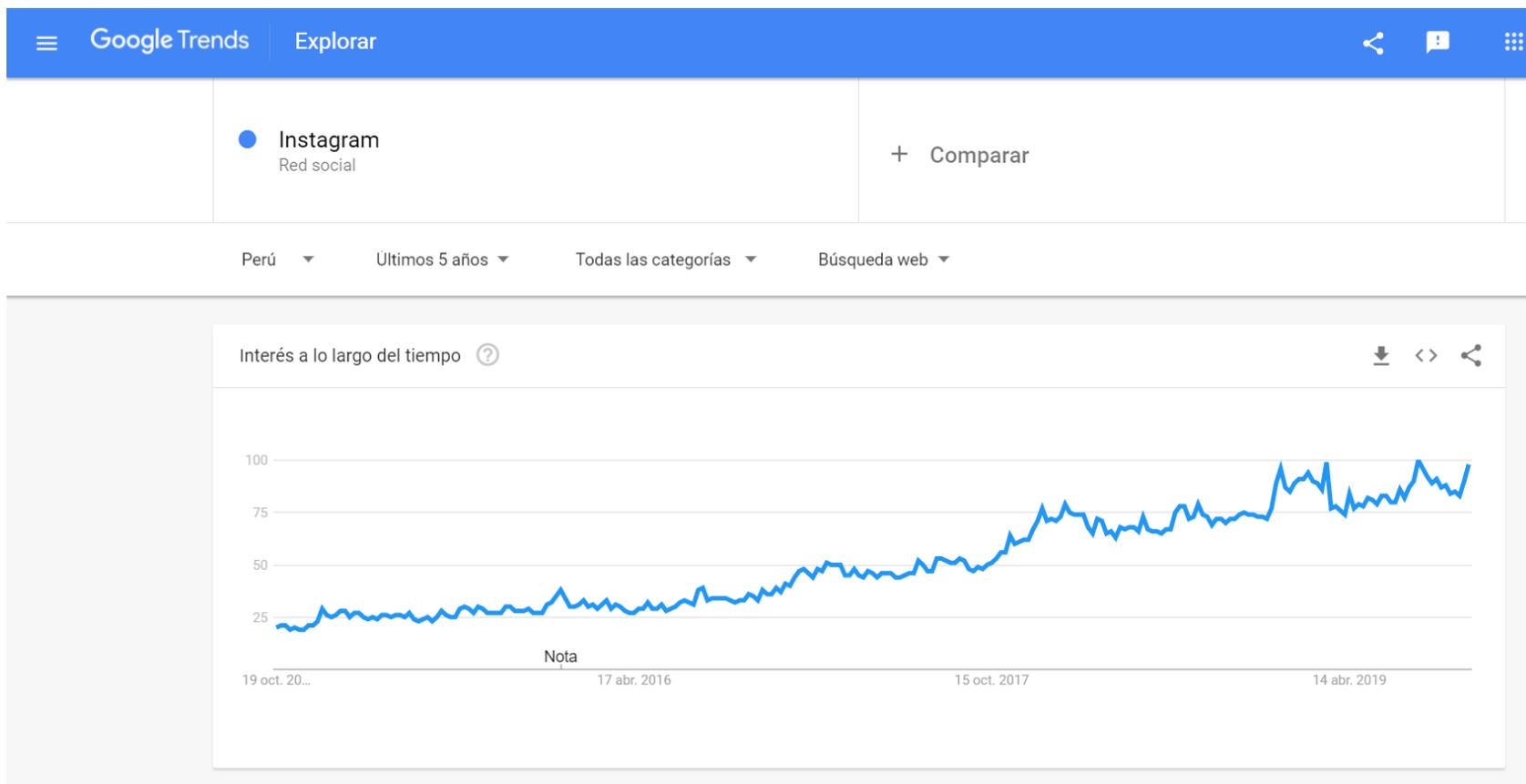


Figura 3: Evolución del interés hacia Instagram en territorio peruano (setiembre 2014-setiembre 2019)

Fuente: Google Trends (2019)

2.2.6. Centros Estéticos e Internet en el Perú

Al investigar la bibliografía en un plano más local, no se han encontrado estudios sobre el negocio de la medicina estética realizados en Lima, o si se han realizado, estos no se han hecho públicos. Por otra parte, según Peru Retail (2016), el consumidor peruano gasta entre el 10% y 15% de su sueldo mensual en estética y belleza, siendo los sectores socioeconómicos A, B y C+ quienes generan una mayor demanda de este tipo de productos y servicios. Asimismo, de acuerdo a lo citado por Abregu, De la Cruz, Fabian, y Gordillo (2018), el sector de belleza y estética crece, aproximadamente, entre el 9% y el 13% anualmente y la rentabilidad de este negocio está entre el 30% y 45%.

Formalmente, poco se sabe sobre qué factores contribuyen a la participación de los clientes en las publicaciones de marcas de Instagram dentro del sector de cuidado estético (Forbes, 2016). En el Perú, nos obstante, sí existe un interés descentralizado a los consultorios estéticos, tal como lo señalan las estadísticas de *Google Trends*.

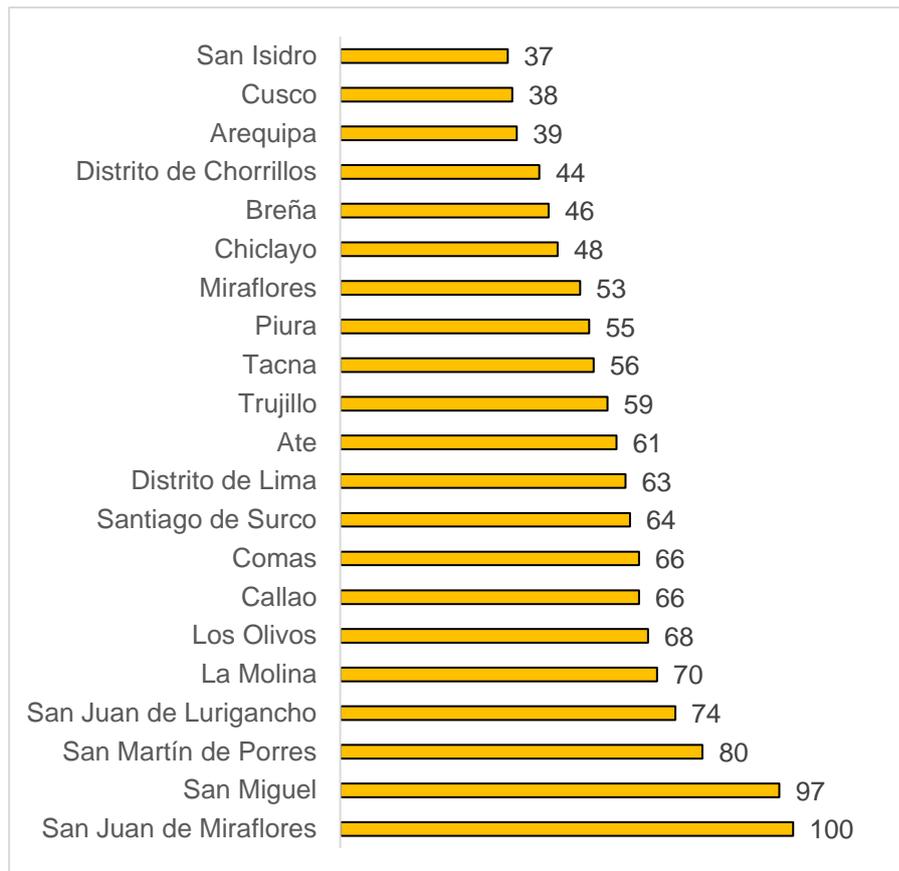


Figura 4: Interés hacia los Consultorios Estéticos en el Perú – 2018 de acuerdo a Google

Fuente: Google Trends (2018)

La figura anterior evidencia que el interés hacia centros estéticos es descentralizado, aunque se concentra en distritos de Lima. No solo es descentralizado geográficamente, sino que además llega a un sector amplio de niveles socioeconómicos. Por tanto, es posible afirmar que la actividad en internet vinculada a las consultas de Centros Estéticos tiende a ser representativa no solo en espacio geográfico, sino también en NSE. Este detalle es importante, ya que coincide con el fenómeno de penetración en el acceso a internet, donde la telefonía móvil juega un rol fundamental.

Por otro lado, cabe mencionar que el interés hacia los centros estéticos en territorio peruano ha crecido sostenidamente en los últimos cinco años, tal como evidencia la siguiente figura:

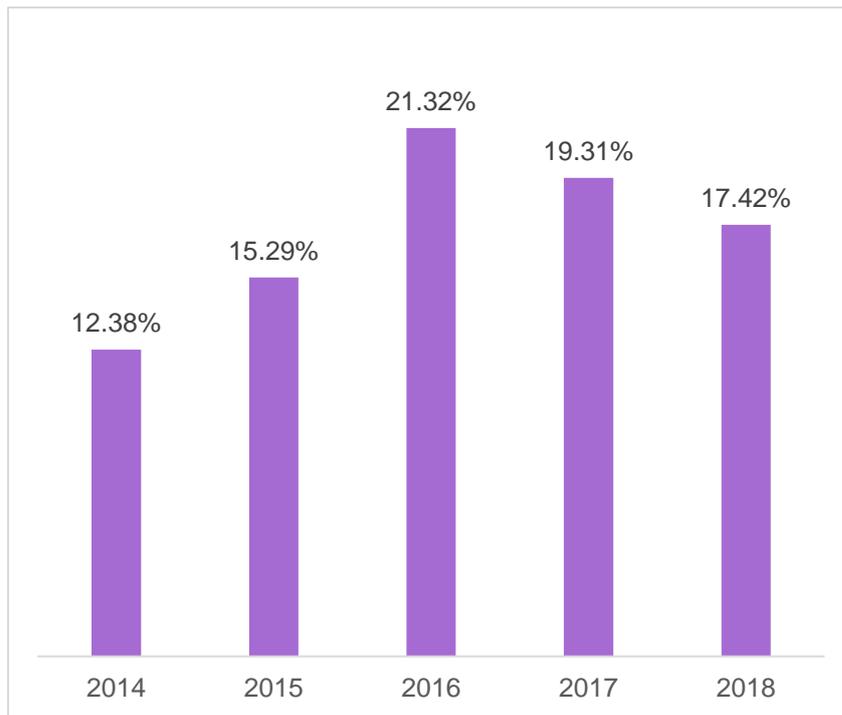


Figura 5: Crecimiento en el interés hacia centros estéticas en Perú de acuerdo a Google

Fuente: Google Trends (2019) – Elaboración: Propia

Nota: Calculado en base a los datos de índices relativos de búsqueda interanuales.

Por otro lado, si se trata de estimar los meses en que se produce un mayor interés por los servicios de centros estéticos en territorio peruano, se cuenta con la siguiente data:

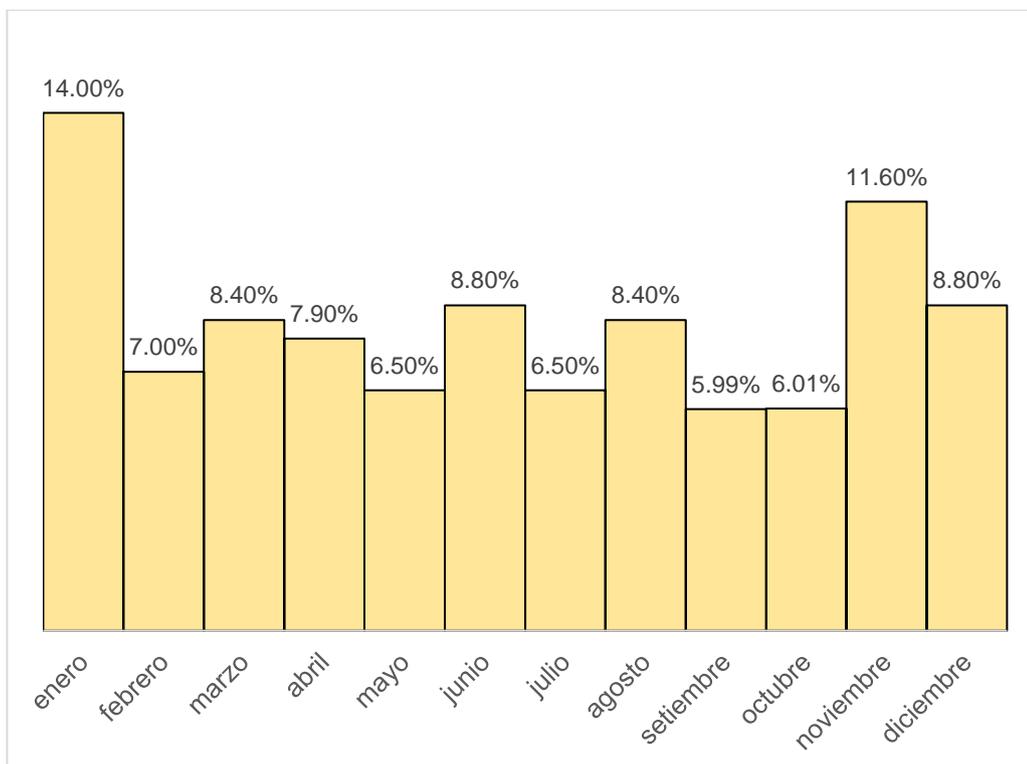
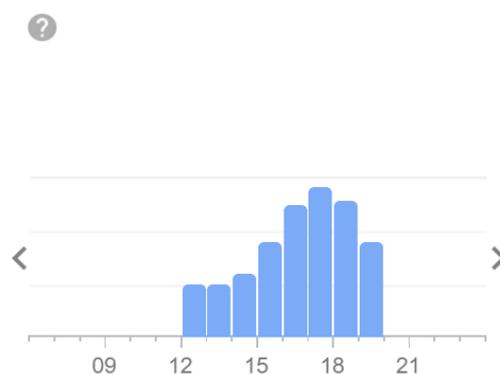


Figura 6: Distribución mensual hacia el interés generado por centros estéticos en Perú de acuerdo a Google en periodo 2013-2018

Fuente: Google Trends (2019)

En lo relacionado a los tiempos en los que las personas acuden a este tipo de negocios se obtuvo:

Horarios populares



Planifica tu visita

Promedio de permanencia: **1-2.5 horas**

Figura 7: Horarios populares y tiempo de permanencia promedio en centros estéticos Lima en

Fuente: Google Maps (2019)

La máxima concurrencia a centros estéticos se produce alrededor de las 6:00 pm. Por otro lado, el tiempo promedio de permanencia de un cliente oscila entre 1 y 2.5 horas. Todo ello de acuerdo a *Google Maps*.

Luego, con respecto a los tratamientos que mayor número cantidad de búsquedas han registrado en Lima Metropolitana de acuerdo a Google, se obtiene:

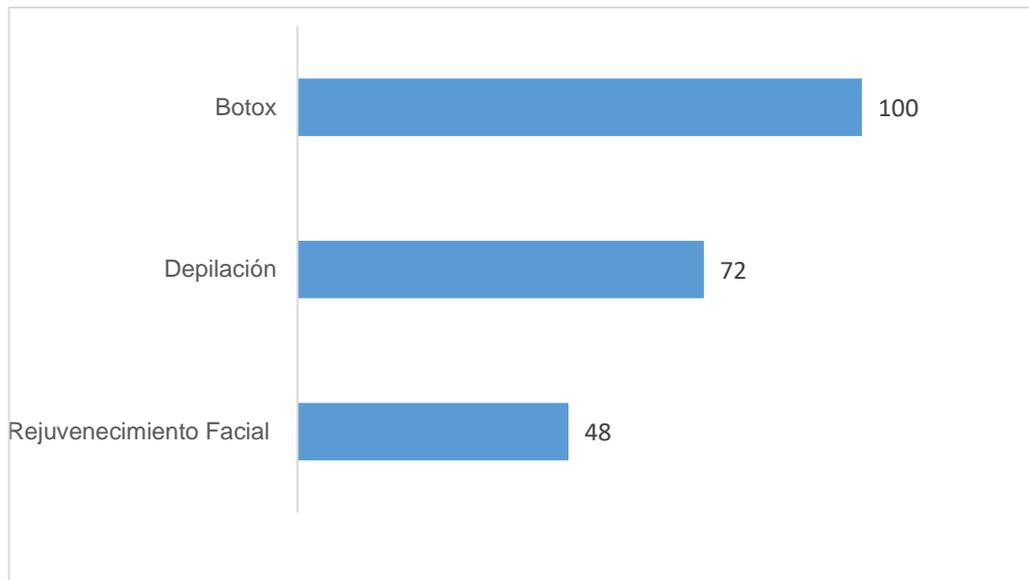


Figura 8: Tratamientos estéticos que mayores consultas han generado en Lima durante el 2018 de acuerdo a Google

Fuente: Google Trends (2019)

2.2.7. Servicios de Centros Estéticos

Los centros estéticos se dedican principalmente al mejoramiento de la apariencia del paciente tanto facial como corporal (Martínez y Padilla, 2020). Se realizan tratamientos tópicos y mínimamente invasivos (Guerra y Salazar, 2016) Entre los más requeridos se pueden mencionar:

- Botox
- Depilación Láser
- Rejuvenecimiento facial,
- Tratamiento con Ácido hialurónico
- Tratamiento con Plasma rico en plaquetas

Estos tratamientos son constantemente publicados en redes sociales como Instagram, plataforma que sirve para evidenciar los tratamientos y casos de éxito en los mismos (Guerra y Salazar, 2016).

2.2.8. Analítica de imágenes

Las personas emplean el sentido de la vista y su capacidad cerebral para ver y percibir visualmente el mundo que los rodea (Szeliski, 2010). La analítica de imágenes es una rama de la visión por computadora (o *Computer Vision*), área del conocimiento que pretende dotar de una capacidad similar, si no mejor, a un computador o máquina en lo que a analizar imágenes de refiere (Ordenes y Zhang, 2019). La visión por computadora está relacionada con la extracción automática, el análisis y la comprensión de información útil de una única imagen o un conjunto de imágenes (Forsyth y Ponce, 2010).

Precisamente, la analítica de imágenes representa un esfuerzo por aprovechar uno de los dos grupos de mayor tamaño disponible en la web (el otro son los textos) (Ordenes y Zhang, 2019). Esto es particularmente importante en contextos donde aplicaciones como Instagram gana popularidad entre los usuarios (Ordenes y Zhang, 2019). La analítica de imágenes forma parte de los procesos de *Business Intelligence* (Baier y Brusch, 2012).

2.2.9. Modelo de Referencia

El modelo de referencia para la investigación es una recopilación de los antecedentes:

$$\text{Customer Engagement} = f(\text{seguidores}, \text{procedimiento_médico}, \text{cantidad_días}, \text{cantidad_palabras}, \text{score_belleza}, \text{score_asombroso}, \text{cantidad_colores})$$

Donde el *Customer Engagement* es la cantidad de likes y comentarios que tiene una determinada publicación de acuerdo a Vignisdóttir (2016), mientras que el hecho de que la imagen pueda ser calificada como propia de un procedimiento vinculado a la medicina fue introducido por Gauthier et.al. (2018). Luego, tanto

la cantidad_palabras y la cantidad_colores fue abordada por Salem (2012) en la comparativa de Facebook vs. Hi5. Precisamente, este investigador también menciona la importancia de que una foto sea considerada como bella al igual que Yu (2014) y Lay y Fewerda (2018). En cuanto al score_asombroso, es decir, el nivel en que una foto pueda ser calificada como asombrosa, fue mencionada por Hausman y Toivonen (2017). Finalmente, la cantidad de días transcurridos desde que una imagen fue subida a Instagram fue una variable propuesta por Tatar, et.al. (2014).

2.3. Objetivos

Dado que la investigación gira en torno a un modelo matemático que debe ser sometido a validez, bien puede emplearse un modelo de regresión. De acuerdo a Harrell (2015), resulta matemáticamente adecuado primero preguntar si el modelo resulta significativo para luego concentrarse en la relación de cada variable predictora con aquello que se quiere predecir.

Objetivo General: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Específico 1: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de seguidores que se tenía en dicho momento en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Específico 2: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el nivel en que la foto podría ser clasificada como

propia de un procedimiento médico en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Específico 3: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de días que ha permanecido en dicha red social en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Específico 4: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de palabras visibles en la misma para el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Específico 5: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente a la belleza de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Especifico 6: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente al atributo sorprendente de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Objetivo Especifico 7: Determinar si existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de colores de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

2.4. Hipótesis

Las hipótesis planteadas corresponden a cada uno de los objetivos planteados en la sección 2.3.

Hipótesis General: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Específica 1: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de seguidores que se tenía en dicho momento en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Específica 2: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el nivel en que la foto podría ser clasificada como propia de un procedimiento médico en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Específica 3: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de días que ha permanecido en dicha red social en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Específica 4: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de palabras visibles en la misma para el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Específica 5: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente a la belleza de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Especifica 6: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente al atributo sorprendente de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

Hipótesis Especifica 7: Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de colores de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.

3. Método

3.1. Tipo y Diseño de Investigación

3.1.1. Tipo de Investigación

La investigación busca modelar el *Customer Engagement* en función a una serie de variables que pueden obtenerse en Instagram o a partir de la propia imagen. Que este modelo sea consistente depende de una comprobación estadístico – matemática: la investigación es del tipo cuantitativa (Bairagi y Munot, 2019). Toda gira en torno a una serie de indicadores estadísticos que están asociados a la validez de un modelo (Kothari, 2004). Evaluar dicho modelo permitirá alcanzar los objetivos de la investigación y determinar si las hipótesis (general y específicas) resultan ser verdaderas o falsas (Jason y Glenwick, 2016).

3.1.2. Diseño de Investigación

La investigación en cuanto a su diseño posee tres características:

- Correlacional: Pretende determinar si el cambio de alguna de las variables independientes está relacionado con la variación del *Customer Engagement*. Este tipo de nexo es propio de las investigaciones correlaciones (Kothari, 2004). Así, por ejemplo, es probable que un aumento en la cantidad de *followers* esté relacionado con la cantidad de interacciones, pero este hecho debe comprobarse mediante el análisis correlacional entre la variable dependiente (*Customer Engagement*) y la variable independiente (número de seguidores).
- No experimental: no se ejerce control sobre alguna de las variables obtenidas de Instagram (Jason y Glenwick, 2016). Todas ellas tienen su origen en la red

social Instagram. Los estudios que toman como base *data* secundaria son principalmente del tipo no experimental (Kothari, 2004). Un trabajo experimental implicaría controlar alguna de las variables (Vignisdóttir, 2016) (subir fotos con una serie de determinadas características, por ejemplo).

- Transversal: El modelo incluye a la cantidad de días que han transcurrido desde el día en que se compartió el contenido hasta la fecha de recolección. Pero no se pretende estudiar a un mismo grupo a lo largo del tiempo. Los elementos de la muestra son tomados en un único lapso. Por otro lado, debe recordarse que Instagram está asociada al concepto de instantaneidad en cuanto a las interacciones que genera (Eagar y Dann, 2015). En este sentido es diferente a YouTube y Facebook (Eagar y Dann, 2015).

3.2. Variables

Tabla 2: Variable de Estudio

Variable	Descripción	Operacionalización
<i>Customer Engagement</i>	Esta es la variable dependiente de la investigación y es del tipo cuantitativa. De acuerdo a la recopilación bibliográfica realizada por Vignisdóttir (2016)	Puede ser determinada como la suma de <i>likes</i> y comentarios. Basta el registro basado en observación directa para su obtención.
Seguidores	Es la cantidad de seguidores que tiene la cuenta de referencia el día en el que fue compartida la imagen. Es una variable cuantitativa. También se obtiene directamente de Instagram La correlación entre está y el <i>Customer Engagement</i> fue planteada Vignisdóttir, (2016) y Qyan, et.al. (2015).	Es obtenida mediante observación directa. Basta anotar la cantidad de seguidores de la cuenta al momento de recolecta la <i>data</i> .
procedimiento_médico	Variable ordinal que mide si una imagen puede ser clasificada como propia de un procedimiento médico (desde very unlikely = 1 hasta very likely =5). Es una variable independiente sugerida por los estudios de Gauthier et.al. (2018).	Es determinada con el instrumento <i>Vision AI Cloud</i> de Google
cantidad_dias	Variable cuantitativa e independiente que mide la cantidad de días que ha permanecido el contenido en la red social Instagram. Esta variable fue considerada tomando en cuenta el aporte de Tatar, et.al. (2014)	Puede obtenerse mediante una operación matemática simple. Para ello se registraron dos fechas (subida y registro en la muestra)
cantidad_palabras	Es el número de palabras que aparecen en la oferta. (Por ejemplo, “aproveche esta oportunidad, deslice hacia arriba” tiene 6 palabras). Es una variable cuantitativa e independiente. Fue una variable indicada por Salem (2012)	Calculada mediante observación directa y comprobada con <i>Vision AI Cloud</i> de Google
score_belleza	Una fotografía puede obtener un puntaje estético entre 0 y 100%. Es una variable cuantitativa e independiente. Fue una variable indicada por Salem (2012), Yu (2014) y Lay y Fewerda (2018).	Calculada con el instrumento <i>Eyeem Vision</i>
score_awesome	Una fotografía puede obtener un puntaje en cuanto al atributo sorprendente de entre 0 y 100%. Es una variable cuantitativa e independiente introducida por Hausman y Toivonen (2017).	Calculada con el instrumento <i>EveryPixel</i>
cantidad_colores	Representa a la cantidad de colores dominantes (principales) presente en la fotografía. Por tanto, es una variable cuantitativa. También es una variable independiente. Fue explorada por Salem (2012)	Calculada con el instrumento <i>Vision AI Cloud</i> de Google

Elaboración: Propia

3.3. Instrumentos

3.3.1. *Social Blade Analytics*

Es un instrumento que se emplea para extraer información de Instagram (Goanta y Ranchordás, 2020). Para usarlo basta instalar el complemento para Chrome desde e ingresar la cuenta que se desea analizar. El uso de ese instrumento radica en la extracción de información de Instagram. *Social Blade Analytics* es gratuito y fue desarrollada bajo el sistema de programación Python (Goanta y Ranchordás, 2020). Con esta herramienta es posible medir la cantidad de *followers* (seguidores). La versión de *premium* brinda información con un mayor nivel de detalle (Social Blade, 2019).

3.3.2. *Vision AI Cloud de Google*

Se trata de una inteligencia artificial (IA) diseñada en los laboratorios de *Google* y especializada en la analítica de imágenes (Hosseini, Xiao, y Poovendran, 2016). Esta IA ha sido parcialmente entrenada por las búsquedas realizadas por internautas que emplean el mencionado navegador (Stokman, 2014). Cuenta con una versión gratuita y otra de pago. Forma parte del *suite de Google Cloud* (Hosseini, et.al. 2016).

3.3.3. *Every Pixel*

De acuerdo a Stokman (2014), *Every Pixel* es una red neuronal entrenada para establecer en qué nivel una imagen resulta asombrosa (o *awesome* en inglés). El instrumento permite subir un archivo de imagen como *input* y como salida se obtiene un número en la escala de 0 a 100% (Stokman, 2014). Dicho número representa el nivel de asombro que puede causar la fotografía (Stokman, 2014). El instrumento fue aplicado a las fotografías que compartieron los centros estéticos mencionados. Debe mencionarse que el hecho de que una fotografía reciba el

calificativo de asombrosa tiene relación con la capacidad de la misma para generar admiración o sorpresa entre quienes la visualizan (Stokman, 2014)

3.3.4. Eyeem Vision

Es una IA basada una red neuronal del tipo *Deep Learning*, entrenada para medir el nivel de belleza (estética o *aesthetics*) presente en una foto (Stokman, 2014). Para ello, la interface del software requiere adjuntar un archivo en formato de imagen, luego se obtiene una calificación que va desde 0 a 100% dependiendo de un algoritmo computacional orientado a estimar la belleza (Stokman, 2014). El instrumento fue aplicado a las fotos de Instagram.

3.3.5. Microsoft Excel

Es una hoja de cálculo diseñada para la serie Office de Microsoft (Ose, 2016). En la investigación este instrumento se usa para almacenar toda la información obtenida a partir de los instrumentos *Social Blade Analytics* y los instrumentos que transforman imágenes en patrones numéricos.

3.3.6. SPSS Modeler

El *SPSS Modeler* es un producto diseñado por IBM orientado a la analítica y desarrollo de modelos (Xiao, Xu, y Xu, 2015). El *SPSS Modeler* contiene un grupo de procesos que facilitan la construcción de flujos orientados al *Data Mining* que permite desarrollar rápidamente modelos de clasificación y predicción mediante técnicas diversas y desplegarlos en operaciones empresariales para brindar un mejor soporte a la toma de decisiones (Xiao, et.al. 2015). Este instrumento será empleado para construir los modelos que permitan estimar el *Customer Engagement*.

3.3.7. Downloader for Instagram

Es un complemento diseñado para el navegador Chrome y es de libre instalación. Este aplicativo permite descargar el contenido (fotografías y vídeos) de la red social Instagram en formatos jpeg (Chrome Web Store, 2017). Su funcionamiento es sencillo: solo se necesita ingresar desde un computador a la página (o *fanpage*) de Instagram y la aplicación aparecerá como un ícono pequeño en la parte superior de cada fotografía o vídeo, luego solo es necesario hacer clic en el dicho ícono para descargar el contenido deseado. También puede integrarse a Chrome para extraer datos en forma masiva.

3.4. Universo

La Universo estuvo formada por todas las publicaciones en formato de fotografía que compartieron entre los meses de setiembre del 2018 y febrero del 2019 las siguientes empresas: Medicina Estética Montero, Belleza Peruana y *360 Mei*. Estas empresas han sido elegidas debido a que pueden ser examinadas de acuerdo a las estadísticas del *Social Blade*. Seguramente existen más, pero el instrumento requiere de una cantidad mínima de interacciones para poder realizar mediciones. Esto es lo que se conoce como masa crítica de información, concepto mencionado por Qyan, et.al. (2015). Por otro lado, son las que también registran un tráfico mínimo de acuerdo a las métricas de Google Trends (2018). El criterio de seleccionar cuentas con un mínimo de registro en cuanto a búsquedas también fue señalado y sugerido por Eagar y Dann (2015). Finalmente hay otras empresas que pueden ser tomadas en consideración, pero dadas ciertas condiciones de privacidad no es posible que se analicen son *Social Blade*.

3.5. Muestra

El muestreo fue abierto y por conveniencia. Ello debido a que Instagram no muestra automáticamente todas las fotos que pueden ser objeto de análisis. Dicho de otra forma,

no se conoce el tamaño exacto de la población y, por tanto, no es posible realizar un muestreo aleatorio (Malhotra, Nunan y Birks, 2017). Debe mencionarse además que existen ciertos criterios de exclusión. Entre ellos cabe mencionar:

- Fotografías que sean parte de un mosaico (secuencia de imágenes que formen parte de una mayor)
- Historias en formato de fotografía o en texto puro.
- Contenido que no fuera propiamente del negocio (apoyo a la selección peruana, por ejemplo)
- Contenido en donde se promocioe otro servicio (o producto) no vinculado con el negocio.

Los criterios de exclusión previamente mencionados son una adaptación de las investigaciones de Ordenes y Zhang (2019). En general, se acepta que el investigador puede eliminar de la muestra aquellos elementos que no considera apropiados para analizar (Vignisdóttir, 2016). La muestra tiene una característica principal : evidenciar los servicios que brindan los negocios. En resumen, no fue incorporada a la muestra cualquier contenido que originara alguna duda sobre el mensaje principal: fotos que busquen promocionar los servicios de clínica estética. En total fue recolectada la data de 491 casos (249 de Belleza Peruana, 183 de Medicina Estética Montero y 59 de 360 Mei). Esta composición de la muestra obedece a la disponibilidad de datos que pudieron ser analizados con *Social Blade*, además de la cantidad de publicaciones de cada empresa. Es necesario recordar que para poblaciones de tamaño desconocido por lo general se incluyen 384 casos. Las características de la muestra son:

Tabla 3: Datos descriptivos de la muestra

Variable	Promedio
<i>Customer Engagement</i> Seguidores	146.12 24,312

procedimiento_médico	2.83
cantidad_días	43.21
cantidad_palabras	12.3
score_belleza	64.50%
score_awesome	4.20%
cantidad_colores	9.15

Elaboración: Propia – Basado en los datos de las publicaciones que incluyeron foto.

3.6. Recolección de datos

Los datos relacionados a seguidores fueron registrados con Social Blade, la cantidad de días fue cotejada con el mismo instrumento. Por otro lado, para analizar las imágenes propiamente dichas es necesario primero descargarlas con *Downloader for Instagram*. La cantidad de palabras se realizó con observación directa, pero fue comprobada con *Vision Cloud*. Lo complicado y trabajoso es trabajar con las variables vinculadas a características de las fotos.

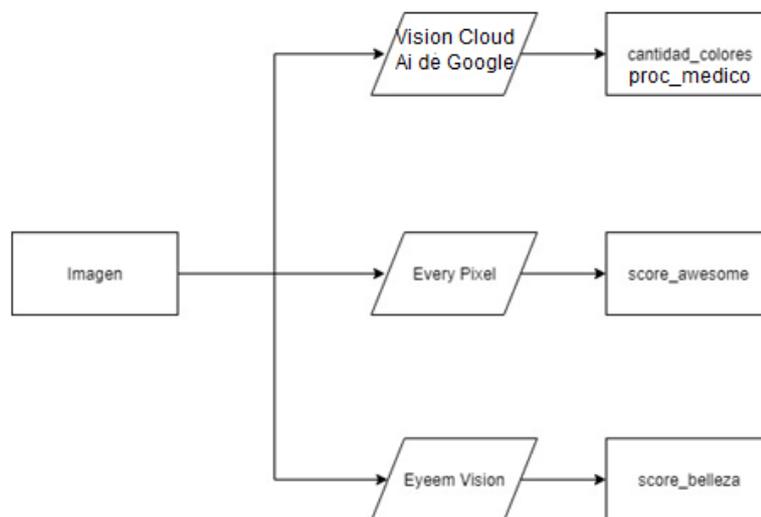


Figura 9: Esquema de procesamiento de datos a partir de imágenes

Elaboración Propia

Todos los datos fueron trasladados a una plantilla en *Excel* para su posterior procesamiento con el *SPSS Modeler*.

3.7. Plan de análisis

3.7.1. Regresión Lineal Múltiple

Para la etapa numérica, una vez completada la base de datos en *Excel*, procedió a emplearse el *SPSS Modeler*. El primer modelo a emplear fue regresión lineal, de forma más precisa fue empleada la regresión con paso hacia adelante, el cual incorporada cada una de las variables explicativas del *Customer Engagement* en forma secuencial y progresiva (Khademi, Jamal, Deshpande y Londhe, 2015). Este algoritmo del SPSS consiste en la selección de variables hacia adelante (*step forward*) y para ello verifica la consistencia del mismo empleando el R-cuadrado o el R-cuadrado ajustado (Bagirov y Ugon, 2018). Este es una medida de la calidad relativa para un modelo estadístico dado un conjunto de datos.

La regresión múltiple debe tener en cuenta algunos supuestos de apoyo para decidir si es un método idóneo, aunque en muchas investigaciones en las que se ha usado este tipo de regresión, no se ha cumplido con todos los criterios necesarios o siquiera han sido mencionados por el investigador (Khademi, Jamal, Deshpande y Londhe, 2015). El escenario más probable es que no se cumplan todos los criterios de validez; sin embargo, es un modelo que se usa con bastante frecuencia dada su facilidad de desarrollo y comprensión (Besley, 1991, pág. 124). Las condiciones para aplicar la regresión lineal múltiple son:

- Aditividad y multicolinealidad: Esto significa que el modelo regresivo es aditivo: cada variable explicativa por sí sola suma contribuye a explicar la variable dependiente, *Customer Engagement* en este caso (Vatcheva, Lee, McCormick, y Rahbar, 2016). No hay relación de dependencia entre las variables explicativas. Si ello se produjera dentro del modelo, se tendría un problema llamado multicolinealidad (Vatcheva, et.al. 2016). La multicolinealidad se da cuando dos variables explicativas están relacionadas (Besley, 1991, pág. 125). Para medir la multicolinealidad es muy frecuente recurrir el análisis del Factor Inflación de la

Varianza (VIF), un valor de este que supera a 5 indica que al menos dos variables explicativas están relacionadas (Vatcheva, et.al. 2016). Valores por debajo de dicho umbral son tolerables para el modelo de regresión (Kurtner, et.al. 2004).

- Autocorrelación entre los residuos: Es una causal de sesgo bastante frecuente en las regresiones con *outputs* continuos. Al emplear la regresión lineal como opción de modelado, se debe cumplir que los residuos deben ser independientes entre sí (Osborne y Waters, 2002). Para verificar si esto se cumple fue necesario calcular el estadístico de Durbin-Watson (Kurtner, et.al. 2004). Conviene recordar que los residuos son las diferencias entre los valores predichos y observados de la regresión.

- Normalidad: Los residuos deben tener una distribución normal, aproximarse a la distribución en forma de campana invertida (Osborne y Waters, 2002). Para ese requerimiento se recurre a la prueba de Kolmogorov – Smirnov aplicado a los residuos no estandarizados (Harrell, 2015)

El nivel de significancia empleado para la prueba ANOVA y los pruebas t de los coeficientes asociados a la regresión lineal múltiple es de 0.05; Este último fue fijado con el propósito de incluir el máximo número de variables pertinentes en el modelo. El valor de 0.05 obedece a parámetros internacionales que no pretenden ser discutidos en el documento (Amrhein y Greenland, 2018).

3.7.2. Red Neuronal

Es posible que no se cumplan todos los requisitos de la regresión múltiple (Khademi, Jamal, Deshpande y Londhe, 2015). Por lo que fue necesario considerar el uso de una red neuronal para poder contrastar los resultados de la

regresión múltiple. Una red neuronal puede ser conceptualizada como un sistema computarizado compuesto por un gran número de elementos simples, elementos de procesos muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas (Matich, 2001).

IBM (2011) y Van Gerven y Bohte (2017) indicaron que las redes neuronales son una de las herramienta preferidas por quienes se especializan en minería de datos debido a que son potentes, flexibles y fáciles de usar. Las redes neuronales predictivas son especialmente útiles en aplicaciones cuyo proceso subyacente sea complejo, tal como ocurre en el *Customer Engagement* de Instagram, donde se pretende modelar una realidad simplificada.

Una red neuronal puede combinar una diversa gama de modelos estadísticos sin que sea un imperativo el plantear hipótesis en forma previa que las vinculen (Matich, 2001). Las relaciones se establecen durante el proceso de aprendizaje *machine learning* (es decir, donde la máquina es entrenada). Si una relación de grado 1 no es la más adecuada, la red neuronal tiende a aproximarse de forma automática a la estructura correcta dada una serie de algoritmos (Kavuncuoglu, et al., 2018). Dado el fin anterior, la data analizada se particiona en dos (Kavuncuoglu, et al., 2018). Una proporción de la data es empleada para el entrenamiento del algoritmo que pretende modelar el *Customer Engagement*, mientras que otra es empleada para la comprobar los resultados del entrenamiento. En este caso la proporción de la data de entrenamiento fue del alrededor de 70% (Xiao, Xu, y Xu, 2015). La proporción de datos restantes fue empleada para comprobar la eficacia del algoritmo. En la red neuronal no es necesario cumplir con todos los supuestos de la Regresión Lineal Múltiple (Kavuncuoglu, et al., 2018).

En forma más precisa, ha sido empleada el modelo perceptrón multicapa. Los perceptrones son las unidades básicas de la red neuronal (equivalentes a las neuronas del cerebro) (Berzal, 2018). La estructura más sencilla presenta tres capas:

- Capa de entrada: Capa donde están las variables predictoras (Berzal, 2018). En este caso son las variables predictoras del *Customer Engagement*.
- Capa oculta: En una red neuronal simple, tanto la capa de entrada como la capa de salida son visibles desde el exterior de la red (Berzal, 2018). Si se añade nuevas capas intermedias, estas capas ya no serán visibles desde el exterior, lo que obliga a utilizar algoritmos como *backpropagation* para ajustar sus parámetros internos (Berzal, 2018).
- Capa de salida: Capa asociada a la variable respuesta u *output* del algoritmo (Berzal, 2018).

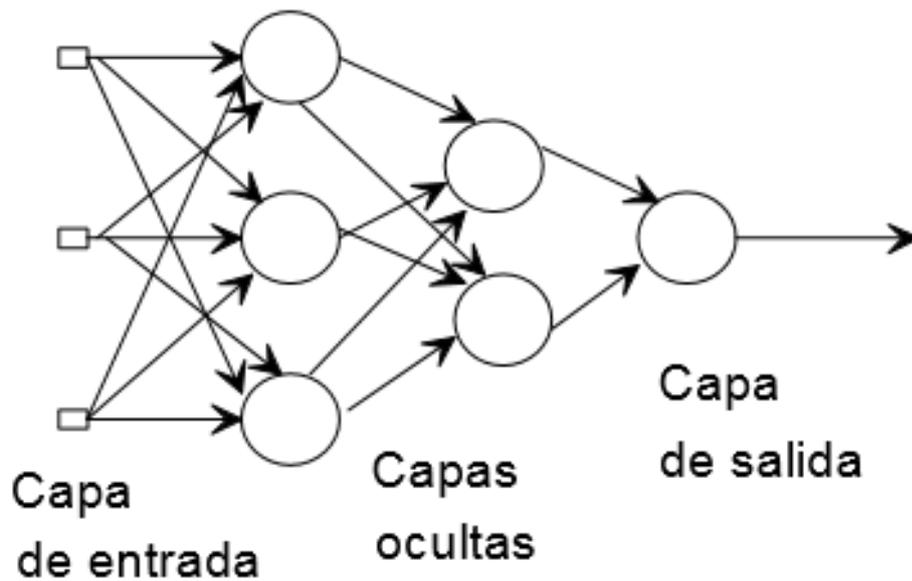


Figura 10: Red Neuronal Modelo Perceptrón Multicapa

Fuente: Jimenez, Paz-Arias, y Larco (2015)

Las redes neuronales son parte del *machine learning*. Para 'aprender' con cada experiencia de entrada (conjunto de datos), el algoritmo alterará las conexiones internas hasta que se descifra cómo lograr una salida deseada dentro de un nivel de precisión especificado (Brenner, 2018). Una vez que el algoritmo ha aprendido, se pueden introducir más entradas y el ANN (*Artificial Neural Network*) proporciona una predicción factible (Brenner, 2018).

Las redes neuronales se utilizan en todas las industrias: en medicina, ingeniería, finanzas, química, biología y otras (Berzal, 2018). También están transformando el conjunto disponible de recursos de tecnología aplicable al marketing, dando a los comercializadores herramientas nuevas, más eficientes y dinámicas en la predicción del comportamiento de los consumidores, comprensión de segmentos de compradores más sofisticados, y la automatización en la creación y entrega de contenidos de marketing (Brenner, 2018).

La aplicación más recurrente de las redes neuronales se encuentra en el campo de la analítica predictiva (DeTienne y DeTienne, 2017). En este caso, las redes neuronales pueden ayudar a los profesionales del marketing a realizar predicciones sobre el resultado de una campaña al reconocer las tendencias de las campañas anteriores (Brenner, 2018). Mientras que las redes neuronales han existido durante décadas, es la aparición más reciente de la *Big Data* la que ha propiciado que esta tecnología sea increíblemente útil para la comercialización (Brenner, 2018).

Las redes neuronales, alimentadas con datos suficientes, son capaces de proporcionar información más precisa y predicciones, ayudando a los tomadores de decisiones de marketing a medir mejor las expectativas (Brenner, 2018). Esta tecnología también está permitiendo un nivel más dinámico de automatización, que no sólo está transformando el flujo de trabajo de marketing, sino que está creando una experiencia con menos fricciones para el consumidor (DeTienne y DeTienne, 2017).

3.7.3. Comparación entre modelos

Finalmente se compara ambos modelos. La regresión lineal múltiple sirve para comprobar el valor de verdad de las hipótesis, mientras que la red neuronal es utilizada para contrastar los resultados del primer modelo (Berzal, 2018).

De ambas variables se ha contrastado la precisión (*accuracy*), los niveles de error y la correlación lineal (entre valor predicho y observado). Todas estas mediciones son proporcionadas por el *SPSS Modeler* (Wendler y Gröttrup, 2016). La labor del investigador es interpretar los indicadores.

Adicionalmente se presentan los resultados que involucran a la importancia de cada variable predictora dentro del mismo modelo. Ello con el fin de alcanzar un mejor nivel de comprensión de la problemática estudiada desde una perspectiva estadístico-matemática (Wendler y Gröttrup, 2016). El *SPSS Modeler* permite calcular la importancia de cada predictor dentro de su respectivo modelo (Wendler y Gröttrup, 2016).

4. Calendario de Actividades y Recursos Disponibles

4.1. Calendario de Actividades

El plan preestablecido para las actividades de investigación consiste en una serie de etapas secuenciales que abarcan el lapso de 15 meses indicado dentro del reglamento aplicable a la Carrera de Marketing de la Universidad San Ignacio de Loyola. A continuación, se muestra el detalle del cronograma.

Tabla 4: Calendario de Actividades

Actividad	Jul-18	Ago-18	Set-18	Oct-18	Nov-18	Dic-18	Ene-19	Feb-19	Mar-19	Abr-19	May-19	Jun-19	Jul-19	Ago-19	Sep-19	Oct-19
Redacción del Plan de Tesis	X															
Primera modificación del Plan de Tesis		X	X													
Recolección de data de Instagram			X	X	X	X	X	X								
Procesamiento de Data con Instrumentos (fotos)						X	X	X	X							
Segunda modificación del Plan de Tesis								X	X	X						
Realización de Pruebas Estadísticas – SPSS Modeler										X	X					
Tercera modificación del Plan de Tesis											X	X				
Entrega Final												X				
Levantamiento de Observaciones													X	X	X	X
Sustentación																X

Elaboración: Propia

4.2. Recursos Disponibles

Los principales recursos para una tesis de esta naturaleza es la data online de *Instagram* y los instrumentos que permiten extraer y procesar las imágenes. La data es abierta (lo que en investigación se conoce como *open data*). Por otro lado, la herramienta que permitieron extraer datos de Instagram y los que viabilizan su posterior procesamiento tienen versiones gratuitas y están disponibles para cualquier persona que quisiera usarlas (al menos antes del cierre de edición del documento). Finalmente, el *SPSS Modeler* cuenta con una versión gratuita de 15 días, plazo suficiente para realizar las pruebas matemáticas que permitieron cumplir con los objetivos de la investigación.

4.3. Presupuesto

El monto en soles destinado a la investigación propiamente dicha fue de casi cero. Ello incrementa la viabilidad de cualquier investigación. Tal como indica Locharoenrat (2017), las investigaciones deben procurar ser viables en cuanto a aspectos financieros.

5. Resultados

5.1. Resultados

El primer resultado consiste en el modelo step-forward (regresión múltiple), el cual irá agregando variables al modelo predictivo del *Customer Engagement*. Ello a medida que se incremente el R-cuadrado y disminuya el error (Rajabi Vandechali, Abbaspour-Fard y Rohani, 2017).

Tabla 5: Resumen del modelo Step Forward

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	,914 ^a	.836	.836	29.2805
2	,921 ^b	.849	.848	28.1583
3	,923 ^c	.852	.851	27.8795
4	,924 ^d	.854	.853	27.7379
5	,925 ^e	.855	.854	27.6494

a. Predictores: (Constante), Seguidores

b. Predictores: (Constante), Seguidores, procedimiento médico

c. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras

d. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras, cantidad_días

e. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras, cantidad_días, score_belleza

De la tabla 5 es posible inferir que alguna de las variables independientes añadidas una a una tiende a explicar mejor el *Customer Engagement*. En este caso las variables independientes que aportan a explicar la variable dependiente son: la cantidad de seguidores (*followers*), el nivel en que la foto pueda ser calificada como propia de un procedimiento médico, la cantidad de palabras presente en la imagen, la cantidad de días que ha permanecido la publicación en Instagram y el score asignado a la belleza. El valor

de R-cuadrado del modelo final es de 85.5%. Esta puede ser considerada como la precisión del modelo.

Existen variable que no deben formar parte del modelo. Estas son el score asignado al atributo *awesome* y la cantidad de colores dominantes presentes en la fotografía. El resumen de los modelos sirve para comprobar si la hipótesis general y las hipótesis específicas son verdaderas. Algunas variables del modelo explican el *Customer Engagement* (cantidad de *likes* más mensajes que tiene una publicación). El resumen de los modelos de regresión puede ser verificados con los resultados de la prueba ANOVA. Estos fueron:

Tabla 6: ANOVA para los modelos Step Forward

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2139152.249	1	2139152.249	2495.089	,000 ^a
	Residuo	419241.727	489	857.345		
	Total	2558393.976	490			
2	Regresión	2171463.922	2	1085731.961	1369.336	,000 ^b
	Residuo	386930.053	488	792.889		
	Total	2558393.976	490			
3	Regresión	2179864.572	3	726621.524	934.841	,000 ^c
	Residuo	378529.403	487	777.268		
	Total	2558393.976	490			
4	Regresión	2184469.951	4	546117.488	709.805	,000 ^d
	Residuo	373924.024	486	769.391		
	Total	2558393.976	490			
5	Regresión	2187617.367	5	437523.473	572.309	,000 ^e
	Residuo	370776.609	485	764.488		
	Total	2558393.976	490			

Elaboración propia

a. Predictores: (Constante), Seguidores

b. Predictores: (Constante), Seguidores, procedimiento médico

c. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras

d. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras, cantidad_días

e. Predictores: (Constante), Seguidores, procedimiento_médico, cantidad_palabras, cantidad_días, score_belleza

De la tabla 6 es posible inferir que para las cinco regresiones empleadas para explicar *Customer Engagement*, el nivel de significancia de la prueba ANOVA permanece en .000. Esto significa que, para cada uno de los modelos propuestos, alguna de las variables independientes (Seguidores, procedimiento_médico, cantidad_palabras, cantidad_días, score_belleza) no deben ser retiradas o deben de permanecer en el modelo.

La prueba ANOVA es una de las condiciones básicas de la regresión lineal múltiple: si el nivel de significancia superaba en algún momento el límite de .05, ello implicaría que ninguna variable debería permanecer en el modelo. Dicho de otra forma, la combinación lineal de las variables independientes no resultaría útil para pronosticar el *Customer Engagement*: todas las variables independientes deben salir del modelo.

El anterior hallazgo, luego, debe ser complementado con el análisis de coeficientes para cada uno de los cinco modelos del *Customer Engagement*. Esto fue realizado con el fin de saber si existe relación directa o inversa entre cada una de las variables independientes en relación con la variable de respuesta.

Tabla 7: Análisis de coeficientes para modelos de regresión

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	
	B	Error estándar	Beta			
1	(Constante)	76,950	2,577	29,861	,000	
	Seguidores	,003	,000	,914	49,951	,000
2	(Constante)	60,006	3,631	16,524	,000	
	Seguidores	,003	,000	,912	51,811	,000
	procedimiento_médico	5,746	,900	,112	6,384	,000
3	(Constante)	71,141	4,940	14,402	,000	
	Seguidores	,003	,000	,914	52,406	,000
	procedimiento_médico	5,627	,892	,110	6,309	,000
	palabras_mensaje	-,450	,137	-,057	-3,288	,001

	(Constante)	64,682	5,579		11,594	,000
	Seguidores	,003	,000	,904	50,552	,000
4	procedimiento_médico	5,686	,888	,111	6,405	,000
	cantidad_palabras	-,435	,136	-,055	-3,189	,002
	cantidad_días	,205	,084	,044	2,447	,015
<hr/>						
	(Constante)	50,116	9,081		5,519	,000
	Seguidores	,003	,000	,904	50,741	,000
	procedimiento_médico	5,779	,886	,113	6,522	,000
5	cantidad_palabras	-,412	,136	-,053	-3,021	,003
	cantidad_días	,202	,084	,043	2,421	,016
	score_belleza	19,556	9,638	,035	2,029	,043

Elaboración propia

El modelo quedaría de la siguiente forma:

$$Customer\ Engagement = 50.116 + 0.003 (Seguidores) + 5.779(procedimiento_médico) - 0.412(cantidad_palabras) + 0.202(cantidad_días) + 19.556(score_belleza)$$

Bajo la ecuación indicada anteriormente, existe una relación directa entre el *Customer Engagement*, la cantidad de seguidores (*followers*), el nivel en que la foto pueda ser calificada como propia de un procedimiento médico, la cantidad de días que tiene la publicación y el score asignado a la belleza. En contraste, la variable dependiente es inversamente proporcional a la cantidad de palabras presentes en la foto.

Naturalmente, no todas las variables independientes tienen la misma importancia: algunas variables deben tener un mayor peso dentro del modelo propuesto. De ello queda evidencia en la siguiente figura:

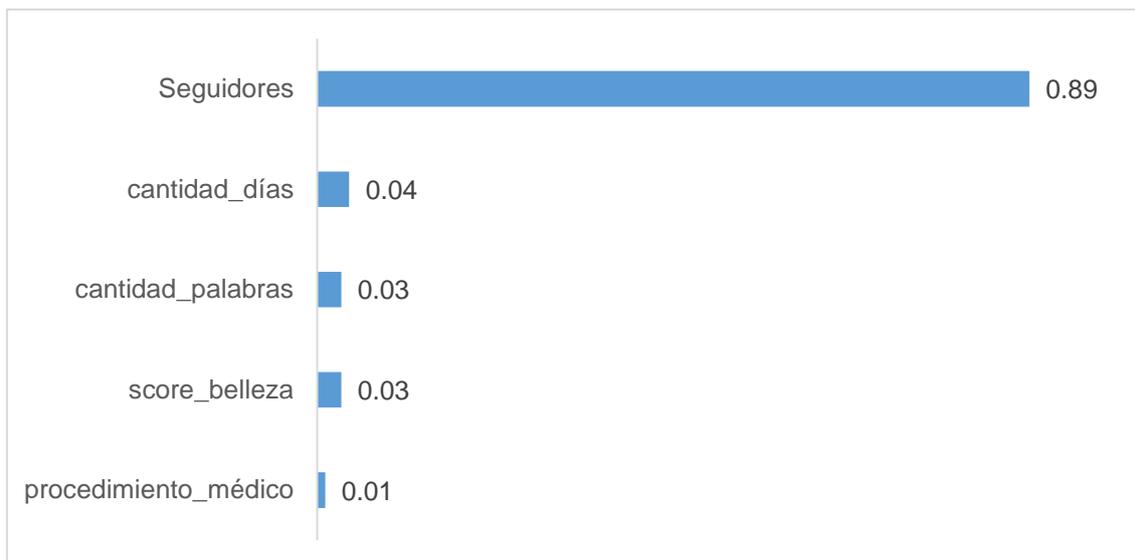


Figura 11: Importancia de las variables predictoras dentro del Modelo de Regresión Lineal (en porcentaje)- *Elaboración propia*

Nota: Los niveles de importancia deben sumar uno

Entonces, dentro del modelo de regresión lineal múltiple consolidado, la variable predictora de mayor peso es la cantidad de seguidores (0.89). A ella le siguen la cantidad de días que tiene la publicación en Instagram (0.04), la cantidad de palabras que incluye la foto (0.03), el score correspondiente a la belleza (0.03), y, por último, el nivel en que la imagen pueda ser calificada como propia de un procedimiento médico (0.01).

Es necesario que la regresión lineal cumpla con una serie de condiciones para que sea un modelo adecuado en el sentido estrictamente estadístico. En primer lugar, fue necesario estudiar si entre las variables independientes existía una relación de dependencia lineal (por ejemplo, que el *score* correspondiente la belleza tuviera relación con la cantidad de colores). Este es un sesgo común que se da en los modelos de regresión lineales.

Tabla 8: Análisis de colinealidad de la regresión lineal

Variable	Estadísticas de colinealidad	
	Tolerancia	VIF
Followers	0.941	1.063

procedimiento_médico	0.995	1.005
cantidad_palabras	0.989	1.012
cantidad_días	0.941	1.063
score_belleza	0.990	1.010

Elaboración propia

A partir de la tabla 8 es posible inferir que los valores VIF se encuentra próximos a uno. Este resultado favorece a la regresión lineal múltiple: no hay relación de dependencia lineal entre las variables independientes del modelo consolidado.

Otra condición a analizar para determinar la idoneidad del modelo de regresión es la no autocorrelación entre residuos. Para ello es necesario recurrir a la prueba de Durbin-Watson.

Tabla 9: Análisis de autocorrelación entre residuos

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar	Durbin-Watson
5	,925	,855	,854	276,494	1,951

Elaboración propia

Lo que favorece a una regresión es que el valor del estadístico de Durbin-Watson se aproxime a 2. Ello evidencia que no se produce autocorrelación entre los residuos (diferencia entre los valores esperados y predichos). Efectivamente, dicha condición se ha cumplido: el estadístico de Durbin-Watson es de 1.951.

Por último, es necesario comprobar otra de las condiciones que idealmente debe darse en una regresión lineal múltiple para ser considerada un modelo adecuado: los residuos deben seguir una distribución normal:

Tabla 10: Test de Normalidad

	Kolmogórov-Smirnov		
	Estadístico	gl	Sig.
Unstandardized Residual	.067	491	.000

Elaboración propia

Los residuos no estandarizados deben cumplir con la condición de distribución normal. Si el nivel de significancia del test de Kolmogórov-Smirnov es inferior al delimitado en la investigación (0.05) entonces los residuos no siguen una distribución normal. Ello no es adecuado para la regresión lineal múltiple como opción de modelado (Bagirov y Ugon, 2018). Esta situación, no obstante, se da con bastante frecuencia, ya que la normalidad de las variables se da casi siempre en aquellas magnitudes vinculadas a la naturaleza (temperatura, distancia entre los astros, humedad, velocidad del viento, etc.) (Rajabi Vandechali, et.al. 2017). Sin embargo, en diversas publicaciones de revistas internacionales indexadas, muchas veces esto no se menciona o se omite. Ello debido a que el hecho de obtener un R-cuadrado muy próximo a 1 ocasiona que muchos investigadores no pretendan profundizar en el estricto sentido estadístico-matemático del modelo.

Finalmente, algunos indicadores de calidad del modelo consolidado de regresión lineal múltiple.

Tabla 11: Resultados de campo para el modelo de regresión lineal

Indicador	Resultado_ regresión
Error mínimo	-47,599
Error máximo	51,061
Error promedio	-0.001
Error absoluto promedio	23,908
Desviación estándar	27,508
Correlación lineal	0,925
Ocurrencias	491

Elaboración propia

Dado que en un sentido estricto no se han cumplido todas las condiciones de la regresión lineal fue necesario emplear explorar otra opción de modelado. En este caso fue seleccionado el modelo perceptrón multicapa.

Tabla 12: Resumen del modelo basado en redes neuronales

Objetivo	<i>Customer Engagement</i>
Modelo	Perceptrón Multicapa
Reglas de Parada	El error no se puede disminuir más
Neuronas en la capa oculta	5
Precisión del modelo	85.10%

Elaboración propia

La precisión del modelo, es decir 85.1%, es inferior al de 85.5% registrado por la regresión lineal (ver R-cuadrado de la tabla 5). Ello significa que la precisión de la red neuronal es inferior a la regresión. Pero para comprender mejor esta situación conviene seguir profundizando en la red neuronal. Con respecto a la arquitectura de la misma se obtuvo:

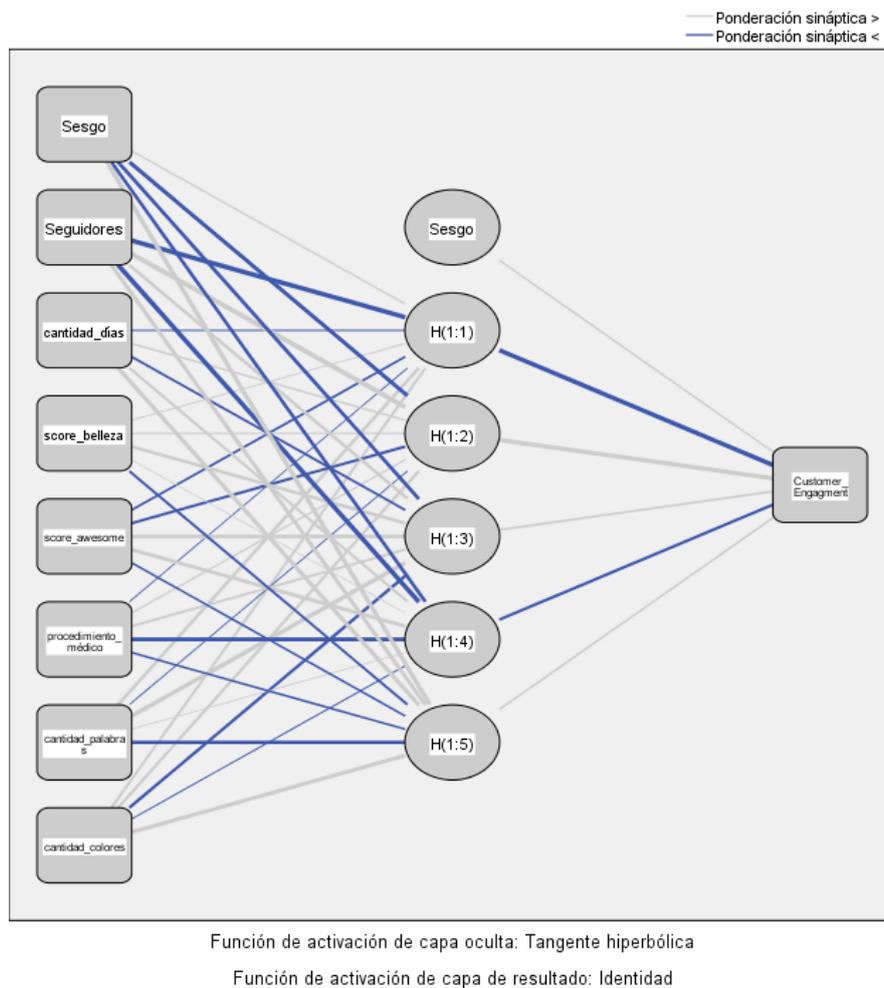


Figura 12: Arquitectura de la red neuronal - Elaboración propia

El modelado anterior, que incluyó todas las variables independientes, generó un total de cinco elementos ocultos, excluyendo el sesgo, el cual tiene una función similar al término independiente en una regresión lineal. Hasta este punto, todas las variables forman parte del modelo, es decir, ninguna ha sido excluida. Las funciones de activación han sido: identidad (análoga a los modelos lineales) (Berzal, 2018) y tangente hiperbólica (en el segundo tramo y primero, respectivamente).

De acuerdo a la figura anterior es posible apreciar que se han formado 5 unidades ocultas de la forma $H(1; n)$, las cuales son también conocidas como perceptrones. Estas representan las combinaciones de las variables predictoras (o independientes) con distintos pesos. Luego, esas 5 unidades en combinación con el sesgo dan como resultado el *Customer Engagement*, la variable respuesta. En cuanto a lo relacionado con los parámetros que permitieron obtener la variable respuesta *Customer Engagement* a partir de los inputs de entrada se obtuvo:

Tabla 13: Estimaciones de parámetros en la red neuronal

Predictor		Pronosticado					<i>Customer_Engagement</i>	
		Capa oculta 1						Capa de salida
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)		
Capa de entrada	(Sesgo)	.189	-.229	.070	-.145	-.366		
	Seguidores	-.439	-.597	-.931	.332	-.179		
	cantidad_días	-.078	.093	-.184	-.305	.213		
	score_belleza	-.382	-.178	.054	-.311	.145		
	score_awesome	-.225	.221	.014	.156	-.368		
	procedimiento_médico	-.428	-.345	-.044	-.372	.457		
	cantidad_palabras	.191	-.368	.022	-.456	-.086		
	cantidad_colores	-.050	-.089	.247	.225	.072		
Capa oculta 1	(Sesgo)						.102	
	H(1:1)						-.259	
	H(1:2)						-.386	
	H(1:3)						-.847	
	H(1:4)						.311	
	H(1:5)						.123	

Elaboración propia

Por otro lado, es natural que los *inputs* del modelo tengan distintos niveles de importancia como predictores. Los resultados en cuanto a la importancia de los *inputs* en la predicción del *customer engagement* de las publicaciones en Instagram para los centros estéticos en Lima dentro de la red neuronal son los siguientes:

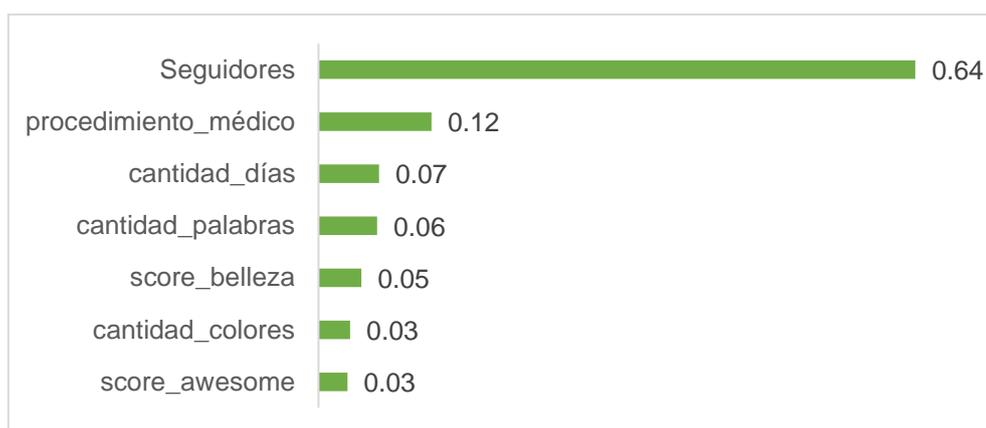


Figura 13: Importancia de las variables predictoras dentro del Modelo basado en una red neuronal (en porcentaje) - Elaboración propia

Hasta este punto existen una diferencia notable con lo determinado en la regresión lineal múltiple. La principal radica en la importancia del predictor vinculado a calificar a la imagen como propia de un procedimiento asociado a la medicina. En la regresión lineal múltiple, dicha variable es la quinta en importancia: mientras que la red neuronal es la segunda. Por otro lado, cabe mencionar que la importancia de la variable predictora cantidad de seguidores ha caído considerablemente de 0.89 en la regresión lineal a 0.64 en la red neuronal. Por otro lado, en cuanto a los resultados de campo para la data de entrenamiento y la de comprobación se obtuvo:

Tabla 14: Resultados de campo para el modelo basado en redes neuronales

Indicador	1_Entrenamiento	2_Comprobación
Error mínimo	-63,771	-56,761
Error máximo	56,777	64,227
Error promedio	0,793	-2,905
Error absoluto promedio	23,24	24,585
Desviación estándar	27,276	28,703
Correlación lineal	0,923	0,924

Elaboración propia

Para que los resultados se comprendan de una mejor manera puede fusionarse los resultados con los resultados alcanzados por la regresión múltiple en la tabla 11. Así se obtuvo:

Tabla 15: Comparativo de resultados entre la regresión lineal y la red neuronal

Indicador	Regresión	1_Entrenamiento	2_Comprobación
Error mínimo	-47,599	-63,771	-56,761
Error máximo	51,061	56,777	64,227
Error promedio	-0.201	0,793	-2,905
Error absoluto promedio	23,908	23,24	24,585
Desviación estándar	27,508	27,276	28,703
Correlación lineal	0,925	0,923	0,924
<i>Accuracy</i> (Precisión)	85.5%	85.1%	
Ocurrencias	491	348	143

Elaboración propia

En cuanto a indicadores vinculados a la comparativa de ambos modelos, la regresión lineal registra menores niveles en de error y un mayor nivel de correlación entre los resultados observados y predichos. El algoritmo estadístico es un modelo de mayor precisión, pero no cumple con las condiciones de validez necesarias para ser considerado como un modelo adecuado desde una perspectiva matemática.

Por otro lado, con respecto a la importancia de las variables predictoras entre los modelos de regresión lineal y red neuronal se obtuvo:

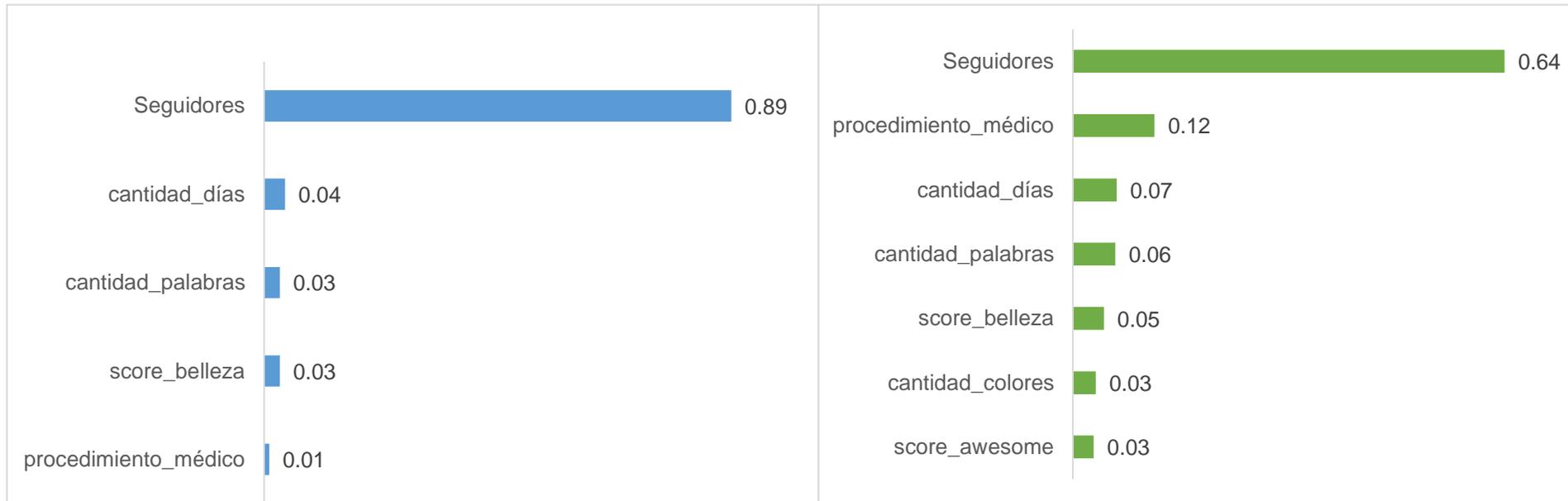


Figura 14: Comparativo de la importancia dentro de las variables predictoras dentro de los modelos de regresión línea (izquierda) versus red neuronal (derecha) (en porcentaje)-
Elaboración propia

La red neuronal incluye a todas las variables predictoras: mientras que la regresión múltiple, únicamente a 5 que resultan significativas dentro del modelado. En a la red neuronal, la variable vinculada a la cantidad de seguidores es menor que en la regresión, mientras que el indicador relacionado a un procedimiento propio de la medicina tiene mayor relevancia en la red neuronal, desplazando a la cantidad de palabras, la cantidad de días y el score de la belleza. Las dos variables de menor grado de importancia son la cantidad de colores y el score vinculado al atributo *awesome*.

5.2. Discusión

Salem (2012), Yu (2014) y Lay y Fewerda (2018) habían mencionado la importancia de la estética en las fotografías para generar *Engagement* entre las personas. Esta variable es relevante en el contexto de Instagram y los centros estéticos. Es significativa para el modelo de regresión lineal múltiple, pero tiene un mayor peso en la red neuronal.

Salem (2012) también había indicado que el hecho de que una fotografía tuviera una determinada cantidad de palabras la convertía en menos atractiva. Para el segmento de centros estéticos estudiado, sucede que la cantidad de palabras incide en el *Engagement*, pero de forma negativa. Ello es apreciable en el modelo de regresión lineal.

Vignisdóttir, (2016) y Qyan, et.al. (2015). Indicaron que el *Customer Engagement* tiene relación con el número de seguidores de una cuenta. En el caso de las publicaciones realizadas desde cuentas de centros estéticos localizados en Lima ocurre lo mismo. De hecho, es la variable con mayor nivel de importancia.

Gauthier et.al. (2018) indicaron que las imágenes que pueden ser clasificados como propias de un procedimiento médico generan *Engagement* en profesionales de la salud. También se ha verificado que lo mismo ocurre en las publicaciones de Instagram para el segmento ya estudiado. Existe una relación directa entre *Customer Engagement* y la naturaleza de la fotografía como propia del ámbito de la medicina.

Hausman y Toivonen (2017) indicaron que existe una relación entre el hecho de que una foto sea calificada de asombrosa y el hecho de que la misma sea considerada como popular dentro del ámbito del turismo. Si bien es cierto que dichos investigadores no exploraron el *Customer Engagement*, la inclusión de la variable mencionada (*score_awesome*) resulta útil para modelar el *Engagement* en el caso de las publicaciones de centros estéticos localizados en Lima, pero ello solo se da con las redes neuronales, pero no en la regresión lineal múltiple.

Salem (2012) había mencionado que los usuarios de *Facebook* consideraban que las fotografías de los usuarios de la red social Hi5 no eran visualmente atractivas debido a que tenían demasiados colores. En el caso de las fotografías publicadas por centros estéticos limeños, la cantidad de colores es independiente del *Engagement*. Ello ocurre en el modelo de regresión lineal. Por otro lado, con la red neuronal, la mencionada variable independiente es de las de menor importancia.

Tatar, et.al. (2014) habían indicado que la cantidad de días que permanece un video en YouTube tiene relación con los comentarios y *likes* del mismo. En el caso de Instagram y los centros estéticos ocurre lo mismo, aunque la incidencia de la cantidad de días es relativamente pequeña. Ello ocurre tanto en la regresión lineal como en la red neuronal. Debe tomarse en cuenta que Instagram funciona de forma diferente a YouTube en cuanto a la difusión de elementos. Los contenidos de Instagram son de menor duración.

6. Conclusiones y Recomendaciones

6.1. Conclusiones

La conclusión principal de la investigación es que si se puede modelar el *Customer Engagement* de una imagen compartida en Instagram en función de un grupo de métricas mencionadas.

Hipótesis general: se corrobora como verdadera. Dicho de otra forma, el modelo propuesto resulto significativo para algunas de las variables independientes propuestas, dicho de otra forma: existe relación entre el *Customer Engagement* y métricas estudiadas.

Hipótesis específica 1: Se corrobora que es verdadera. Existe relación entre la cantidad de *followers* (seguidores) y el *Customer Engagement* de las fotos en Instagram subidas por los centros estéticos de Lima.

Hipótesis específica 2: Se corrobora que es verdadera. El *Customer Engagement* tiene relación con el nivel en que la foto podría ser clasificada como propia de un procedimiento médico. Luego, las fotos de este tipo no solo son apreciadas por estudiantes de medicina o ramas afines, sino también por quienes están interesados en los servicios de centros estéticos.

Hipótesis Específica 3: Se corrobora su veracidad. El *Customer Engagement* tiene relación con la cantidad de días que ha permanecido en dicha red social. Dicha relación es del tipo directa.

Hipótesis Específica 4: Se corrobora su veracidad. Existe relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de palabras visibles en la misma.

Hipótesis Específica 5: Se corrobora su veracidad. Existe una relación significativa entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente a la belleza de la misma.

Hipótesis Especifica 6: Se corrobora que es falsa. No hay relación entre el *Customer Engagement* de una imagen compartida en Instagram y el score correspondiente al atributo sorprendente (*awesome*) de la misma, al menos bajo el modelo de la regresión lineal múltiple.

Hipótesis Especifica 7: Se corrobora que es falsa. No hay relación entre el *Customer Engagement* de una imagen compartida en Instagram y la cantidad de colores de la misma.

Bajo los dos modelos propuestos, la variable de mayor incidencia en el *Customer Engagement* es la cantidad de seguidores. La sumatoria de la cantidad de likes y el número de comentarios que recibe una fotografía tiene una particular relación con la cantidad de seguidores que tenía la cuenta para cualquiera de los dos modelos estudiado. Ello implica que para generar mayores niveles de *Engagement* es necesario establecer políticas orientadas a incrementar la cantidad de seguidores.

El hecho de que fotografía pueda ser clasificada como del tipo médico tiene relación con el *Customer Engagement*. Esta es del tipo directa. Esto es un aporte de la investigación a la literatura académica, y es que es una variable de reciente incorporación medible con un instrumento basado en una inteligencia artificial. Esto tiene implicancias en el marketing como campo teórico. Los procedimientos médicos llaman la atención no solo de los profesionales vinculados a la medicina, sino también de personas interesadas en procedimientos estéticos.

La cantidad de días que ha transcurrido la publicación tiene una relación significativa con el *Customer Engagement*. Esto era esperable, pero la incorporación de esta variable era necesaria para dimensionar mejor la variable dependiente.

La cantidad de palabras presentes en una imagen guarda relación inversa con la cantidad de likes y comentarios que obtiene la misma. La idea, entonces, no es que los profesionales en marketing eliminen las palabras de las imágenes, lo adecuado es minimizar la cantidad de palabras que son empleadas en la imagen. No es sugerible abusar del contenido en formato de texto, al menos bajo los resultados de la muestra analizada. Ello debe ser tomado en cuenta por quienes tienen a su cargo la gestión del contenido en publicidad online.

Por otro lado, la belleza de la imagen juega un rol significativo en el *Customer Engagement*. Cabe mencionar que esta variable es medida empleando un software especializado respaldado por un algoritmo entrenado. Esto último no se da con el *score_awesome*. Dicha variable no aporta significativamente al *Engagement* bajo el modelo de regresión lineal múltiple y es la última en importancia en el modelo basado en redes neuronales. Ello implica que, en el ámbito del marketing, los modelos predictivos son simplificaciones del mundo real. Se busca la predictibilidad a fin de reducir la incertidumbre.

La cantidad de colores no aporta significativamente al modelo de regresión lineal propuesto, al menos para la data empleada. Cabe mencionar que en la red neuronal dicha variable tiene uno de los aportes de menor relevancia. No debe confundirse la cantidad de colores con la armonía que presentan los mismos dentro de una imagen.

La regresión lineal múltiple no es el modelo más propicio. Una red neuronal, en cambio, si es más adecuada en el sentido estrictamente matemático. El modelo de red

neuronal incluye todas las variables independientes dentro del modelo. Sin embargo, la red neuronal no permite verificar el valor de verdad de las hipótesis. Esta es una de las limitantes del uso de este método. Las redes neuronales siguen siendo válidas en cuanto a su empleo matemático. Su contraste con la regresión múltiple obedece a que los principios de esta última no siempre se cumplen. De hecho, lo habitual es que no se cumplan todos los requisitos que la regresión implica. Los departamentos de marketing pueden integrar ambos modelos, dado que ambos tienen un nivel de ajuste relativamente alto. Probablemente el modelo de regresión sea el más fácil de implementar.

De alguna manera el *Customer Engagement* puede ser comparado con un ratio de respuesta. Así es posible comparar los resultados esperados versus lo que finalmente se obtuvo. Esto con el fin de llevar un control más preciso de una de las métricas vinculadas al uso de las redes sociales más allá de la simple recolección de *likes*. El *Customer Engagement* posee una estructura matemática. El modelo propuesto es solo una de las rutas que pretenden comprender el fenómeno de las interacciones de las personas con las empresas en un contexto de marketing 4.0.

Otra aplicación al marketing consiste en el hecho de aproximarse a seleccionar una fotografía con características que resulten propicias para general *Engagement* entre quienes siguen a dicha cuentas. Ello puede ser respaldado con software gratuito.

6.2. Recomendaciones

El modelo puede incorporar otras variables a medida que un investigador decida explorar otras variables. Así, por ejemplo, es posible incorporar el análisis de entidades. Los instrumentos pueden medir, por ejemplo, en qué proporción una foto tiene actividad humana. Pero es necesario contar con cierto respaldo teórico. En los antecedentes académicos mencionados existen casos en otros contextos (turismo, otras redes sociales, etc.), pero que han servido de respaldo para la investigación.

Los instrumentos mencionados pueden servir para no solo para construir modelos predictivos, sino también para segmentar. Instagram no es la única red social de la que se pueden extraer fotografías. El estudio puede ser replicable en Facebook. Dado que las empresas ejercen estrategias multicanal, entonces es probable que el Engagement en Facebook tenga sus propios parámetros.

Cuando las personas comentan o brindan un “me gusta” es posible acceder a su información, de forma más específica, es posible aproximarse a la idea de cómo ellos pretenden mostrarse en Instagram. Por otro lado, cabe destacar que la analítica de imágenes puede emplearse para analizar perfiles personales. De hecho, hay software que puede estimar la edad de una persona e incluso determinar su grupo racial. Naturalmente todos los datos deberían ser anonimizados para así no incurrir en falta éticas.

Si bien es cierto que los modelos obtenidos son útiles para los centros estéticos. Es probable que sea aplicable en otros contextos como los spas o salones de belleza. Probablemente aquí deba excluirse la variable vinculada a los procedimientos médicos. Pero los demás indicadores deberían ser relevantes.

En cuanto a los profesionales vinculados al marketing y *Business Intelligence* que deseen incorporar el modelo a los métodos de estimación, es poco probable que puedan replicar la red neuronal debido a que se requiere software especializado o dominio sobre ciertos lenguajes de programación, además de la data de entrenamiento. En este sentido queda como alternativa la regresión lineal múltiple, ya que al tener una ecuación planteada ya es fácil de replicar mediante operaciones aritméticas sencillas. Por ello, resulta de mayor facilidad aproximarse al *Customer Engagement* ingresando los valores indicados en una ecuación de regresión.

Desde un punto de vista de la gestión, es probable establecer metas. Obtener una cantidad determinada de reacciones, por ejemplo, ya debería ser accesible dadas ciertas variables que puede controlar un *community manager*. Un modelo predictivo contribuye al establecimiento de metas dada su respaldo matemático. Así, pueden establecerse metas reales. Precisamente la analítica de imágenes contribuye a ello, ya que permite obtener los indicadores numéricos que posibilita el procesamiento estadístico (o computacional) pertinente.

Desde un enfoque estratégico y de gestión, el método empleado puede contribuir a la construcción de *Key Performance Indicators*. Así, por ejemplo, el *community manager* puede trabajar dos indicadores:

- *Customer Engagement* esperado: Cantidad de *likes* y comentarios que se esperan de una publicación.
- Diferencial entre *Customer Engagement* esperado y el real, el cual surge de la siguiente fórmula:

$$\left(\frac{C.Engagement Esperado - C.Engagement Real}{C.Engagement Real} \right) * 100\%$$

Quienes tengan a su cargo la elección de contenido fotográfico, pueden apoyarse en el software de análisis de imágenes. Resulta posible estimar el nivel en el que una foto puede ser calificada como bella o asombrosa. El uso de los instrumentos empleados en la investigación representa un ejemplo de ello.

El punto anterior tiene que ver con el entrenamiento de quienes tengan a su cargo el proceso que consiste en tomar y editar las fotografías. Por ejemplo, quienes capturan y procesan imágenes pueden determinar si su capacidad para generar contenido visualmente atractivo va mejorando a lo largo del tiempo.

Si es que el *community manager* desea seleccionar alguna característica en especial de una foto (que pertenezca a un procedimiento médico de nivel 3 o superior, por ejemplo), ahora puede hacerlo. Lo mismo ocurre con todos los atributos visuales. Es posible agrupar las fotos y realizar un proceso de selección de acuerdo a aquellos que desee reforzarse (belleza, capacidad de asombrar, cantidad de colores, etc.).

La ecuación estimada sirve incluso para monitorear a un grupo de empresas del sector. Por ejemplo, si el valor esperado del *Customer Engagement* para una publicación determinada y en una fecha específica es de 500 y al final termina siendo 1,200; ello significaría que la publicación en particular podría haber sido objeto de algún elemento externo (publicidad de una *landing page*, difusión de pago en Instagram, compra de *likes* y otros). Lo relevante sigue siendo el hecho de poder predecir el *Engagement*. Sin un modelo predictivo resultaría casi imposible poder identificar este tipo de casos. Esta situación representa una aplicación de lo que se conoce como identificación de casos atípicos.

Si los creadores de contenido carecen provisionalmente de ideas en cuanto a contenido fotográfico que pueda generar *Engagement*, pueden recurrir a analizar otras realidades. De acuerdo a Nurbayani y Annuraini (2017), Instagram ha jugado un rol importante como difusor de lo que conoce como *Korean Wave* (Onda Coreana). Ello se debe a que en Corea del Sur se ha alcanzado niveles de calidad bastante altos en lo que contenido fotográfico se refiere. Nurbayani y Annuraini (2017) mencionan que el contenido de Instagram incluso ha logrado incidir en el turismo médico que proviene desde Emiratos Árabes hacia Corea del Sur. Los investigadores mencionan que el contexto cultural de ambas naciones es muy diferente, pero a pesar de ello el *Engagement* se da. Por tanto, es posible analizar el contenido en Instagram de otros contextos culturales y adaptarlo al escenario peruano. Si se requiere saber si dicha adecuación

resulta visualmente atractiva, puede recurrirse a los instrumentos ya mencionados en la tesis. Todo ello con el fin de estructurar la información que conllevará a una mejor toma de decisiones, sostenida en evidencia cuantitativa.

7. Limitaciones e Investigaciones futuras

7.1. Limitaciones

Una limitación a considerar es la propia técnica de analítica de imágenes. Los *softwares* empleados están basados en algoritmos que obedecen a patrones actuales. Los estándares de lo que hoy es considerado bello o asombro no necesariamente deberían permanecer estáticos. Lo esperable es que dichos conceptos cambien con el tiempo. Los instrumentos requieren, entonces, ser actualizados cada cierto tiempo.

Existen externalidades que la investigación no puede medir. La compra de *likes*, por ejemplo, es un fenómeno que puede presentarse. En este sentido existen investigaciones en los que se han desarrollado algoritmos que permiten detectar cuentas falsas. Esta limitación no es propia de Instagram, sino también puede darse en YouTube o Facebook, tal como lo indica Ross (2019). Si los *likes* son comprados es una situación que solo pueden ser quienes administran las cuentas oficiales. Intentar aproximarse a cuantificar los *likes* que provienen de cuentas falsas involucraría una investigación propia de las ciencias de la computación. En las investigaciones llevadas a cabo por Dugué (2017) para predecir *likes* en Instagram no se hace referencia a mecanismo alguno para filtrar los mismos. Por ello es un límite para la investigación.

Ciertamente, los hallazgos de la tesis pueden ser reforzados si es que se realiza un muestreo aleatorio en lugar de uno por conveniencia. Pero ello implica superar uno de las principales limitaciones de la investigación: la ausencia de un comando automático que permita filtrar todas las fotos que puedan ser objeto de análisis.

Al cierre de edición del documento, han sido integradas nuevas políticas en la red social Instagram. La principal está referida a la visualización de los likes. La red social propiedad de Facebook está haciendo pruebas para ocultar la información sobre el total

de “me gusta” que acumula una foto, y que se muestran en forma de corazón junto a cada publicación (Vives, 2019). Los likes no van a desaparecer, pero la etapa de prueba restringe el hecho de poder visualizarlos con rapidez (Vives, 2019). En otras palabras, si Pedro, usuario de Instagram, compartía un contenido y recibe cinco me gusta de Juan, María, Rodolfo, Miguel y Claudia, únicamente él (Pedro) podría ver ese consolidado total. Si otra persona quisiera ver el número de likes del contenido publicado por Pedro, entonces tendría que contabilizar cada uno de los “me gusta” de forma manual. Estas modificaciones están en etapa piloto (Vives, 2019).

7.2. Investigaciones Futuras

Las investigaciones futuras pueden ser del tipo experimental. Dicho de otra forma, es posible que un investigador asuma el rol de *community manager* de una empresa para el sector mencionado y así realizar publicaciones. De esa forma el investigador puede controlar algunas variables.

Otro camino de investigación está vinculado a las imágenes médicas en Instagram. Es un campo poco explorado por revistas especializadas. No obstante, los hallazgos del trabajo evidencian que juegan un rol importante con los *likes* y comentarios que obtiene una foto en Instagram.

Puede emplearse el análisis textual del contenido que aparece en las imágenes. Para ello no sería recomendable el análisis sentimental, dado que se espera que todos los mensajes incorporados a las fotografías sean positivos. En cambio, si es recomendable emplear el análisis de tópicos. Por ejemplo, es probable que una frase anunciando un 50% de descuento genere un mayor nivel de *Engagement* que otro en donde se anuncia la apertura durante un día feriado. En estos casos sería necesario clasificar los tipos de mensaje.

Existen otras variables que pueden ser analizadas: el enfoque empleado para la foto, las emociones (si es que aparecen rostros), plano desde que fue capturada, contextos, balance de colores, etc. (Muñoz y Towner, 2017). Pero la inclusión de dichas variables requiere de un marco teórico de mayor amplitud.

Instagram cuenta también con datos en formato de video. Otros investigadores pueden transformar esta data en indicadores numéricos. Esto ya forma parte de otro campo de aplicación conocida como *Video Analytics* (Kholod y Balkan, 2014). Posteriormente sería posible comparar los resultados del *Engagement* generado por los formatos de imagen y video. Otra opción interesante es la construcción de clústeres a partir de fotos (Donaire, Camprubí y Galí, 2014).

Bibliografía

- Abregu, J., De la Cruz, D., Fabian, C., & Gordillo, V. (2018). *Plan de negocio para la creación de un servicio de medicina estética en la Clínica Santa María Magdalena en la ciudad de Ayacucho - Perú*. Tesis de maestría, ESAN, Lima.
- Amrhein, V., & Greenland, S. (2018). Remove, rather than redefine, statistical significance. *Nature Human Behaviour*, 2(4). doi:10.1038/s41562-017-0224-0
- Association American Marketing. (15 de Mayo de 2018). *Dictionary*. Available Online.: Obtenido de <https://www.ama.org/resources/Pages/Dictionary.aspx?dLetter=C>
- Bagirov, A., & Ugon, J. (2018). Nonsmooth DC programming approach to clusterwise linear regression: optimality conditions and algorithms. *Optimization Methods and Software*, 33(1), 194-219. doi:10.1080/10556788.2017.1371717
- Baier, D., & Bruschi, I. (2012). Image data analysis and classification in marketing. *Advances in Data Analysis and Classification*, 6(4). doi:10.1007/s11634-012-0116-0
- Bairagi, V., & Munot, M. (2019). *Research Methodology: A Practical And Scientific Approach*. Chapman and Hall/CRC.
- Barger, V. (2016). Social media and consumer engagement: a review and research agenda. *Journal of Research in Interactive Marketing*, 10(4), 268-287. doi:10.1108/JRIM-06-2016-0065
- Berzal, F. (2018). *Redes Neuronales & Deep Learning*. Granada: Edición independiente.
- Besley, D. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. John Wiley & Sons.
- Brenner, M. (23 de Junio de 2018). Artificial Neural Networks: What Every Marketer Should Know. *Marketing Insider Group*. Recuperado el 30 de Octubre de 2019, de <https://marketinginsidergroup.com/content-marketing/artificial-neural-networks-every-marketer-know/>
- Buss, O., & Begorgis, G. (2015). *The Impact of Social Media as a Customer Relationship Management Tool*. Tesis de maestría, Karlsstads Universitet, Business School. Recuperado el 20 de Junio de 2018, de <http://www.diva-portal.org/smash/get/diva2:821090/FULLTEXT01.pdf>
- Byrne, E., Kearney, J., & MacEvelly, C. (2017). The Role of Influencer Marketing and Social Influencers in Public Health. *Proceedings of the Nutrition Society*, 76(3), 114-128.

- Caldeira, S. (2016). Identities in Flux: An Analysis to Photographic Self-Representation on Instagram. *Observatorio*, 10(3). Recuperado el 17 de Julio de 2019, de http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1646-59542016000400009
- Chen, H. (20 de Marzo de 2018). *College-Aged Young Consumers' Perceptions of Social Media Marketing: The story of Instagram*. . Obtenido de Journal of Current Issues & Research in Advertising, [e-journal] vol. 39, no. 1, Available through: LUSEM Library website: e <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Chrome Web Store. (2017). Downloader for Instagram - Descripción. Recuperado el 14 de Febrero de 2018, de <https://chrome.google.com/webstore/detail/downloader-for-instagram/olkpikmlhoajbbmmpejnimiglejmboc?hl=en>
- Chu, D., Shen, Z., Zhang, Y., & Yang, S. (2017). Real-Time Popularity Prediction on Instagram. *28th Australasian Database Conference*. Brisbane. doi:10.1007/978-3-319-68155-9_21
- CMOSurvey.org. (21 de Marzo de 2018). *CMO Survey Report: Highlights and Insights*. Obtenido de https://cmosurvey.org/wp-content/uploads/sites/15/2018/02/The_CMO_SurveyHighlights_and_Insights_Report-Feb-2018.pdf
- Cosh, K., Burns, R., & Daniel, T. (22 de Marzo de 2008). *Content Clouds: Classifying content in Web 2.0*. Obtenido de Library Review, [e-journal] vol. 57, no. 9, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing Through Instagram Influencers: The impact of number of followers and product divergence on brand attitude. 36(5). Obtenido de International Journal of Advertising, [e-journal]: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Dehghan, A., Ortiz, E., Shu, G., & Zain, S. (2017). DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Networks. *Computer Vision and Pattern Recognition*.
- Dessart, L., Veloutsou, C., & Morgan-Thomas, A. (2016). Capturing Consumer Engagement: Duality, dimensionality and measurement. *Journal of Marketing Management*, vol. 32, no. 5-6, pp.399-426 .
- DeTienne, K., & DeTienne, D. (2017). Neural networks in strategic marketing: exploring the possibilities. *Journal of Strategic Marketing*, 25(4), 289-300. doi:10.1080/0965254X.2015.1076881

- Djafarova, E., & Rushworth, C. (2017). *Exploring the Credibility of Online Celebrities' Instagram Profiles in Influencing the Purchase Decisions of Young Females*. . Obtenido de Computers in Human Behavior, [e-journal] vol 68, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Donaire, J., Camprubí, R., & Galí, N. (2014). Tourist clusters from Flickr travel photography. *Tourism Management Perspectives*, 11, 26-33.
- Dorfman, R., Vaca, E., Mahmood, E., Fine, E., & Schierle, C. (2018). Plastic Surgery-Related Hashtag Utilization on Instagram: Implications for Education and Marketing. *Aesthetics Surgery Journal*, 38(3), 332–338. doi:10.1093/asj/sjx120
- Dugué, C. (2017). *Predicting the number of likes on Instagram*. Recuperado el 30 de Enero de 2018, de <https://towardsdatascience.com/predict-the-number-of-likes-on-instagram-a7ec5c020203>
- Eagar, T., & Dann, S. (2015). Capturing and Analyzing Social Media Composite Content: The Instagram Selfie. *Research in Consumer Behavior*, 18(2).
- Elo, S., & Kyngas, H. (2008). The Qualitative Content Analysis Process. *Journal of Advanced Nursing*, 62(1), 107-115.
- Evans, N., Phua, J., Lim, J., & Jun, H. (2017). *Disclosing Instagram Influencer Advertising: The effect of disclosure language on advertising recognition, attitudes, and behavioral intent*. Obtenido de Journal of Interactive Advertising, [e-journal] vol. 12, no. 2, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- EveryPixel. (2019). EveryPixel Demo. Recuperado el 12 de Junio de 2019, de <https://labs.everypixel.com/api/demo>
- Eyeem Vision. (2019). Online demo of Eyeem Vision. Recuperado el 2 de Junio de 2019, de <https://www.eyem.com/blog/introducing-eyevision>
- Felix, R., Rauschnabel, P., & Hinsch, C. (2017). Elements of strategic social media marketing: A holistic framework. *Journal of Business Research*, 70, 117-126.
- Forbes, C. (2016). Examining the Beauty Industry's Use of Social Influencers. *Elon Journal of Undergraduate Research in Communications*, 7(2), 78-87.
- Forsyth, A., & Ponce, M. (2010). *Computer Vision: A Modern Approach*. Cambridge University Press.
- Fuchs, C. (2014). *Social Media: A critical introduction*. London: SAGE Publications Ltd.

- Gauthier, T., Bratberg, J., Loi, K., & V DiVall, M. (2018). Delivery of educational content via Instagram. *Medical Education*, 50(5), 575-576. doi:10.1111/medu.13009
- Giannakis, C., & Boutsouki, C. (2014). Customer Relationship Management in the Era of Social Web and Social Customer: An Investigation of Customer Engagement in the Greek Retail Banking Sector. *Procedia - Social and Behavioral Sciences*, 148, 67-78.
- Goanta, C., & Ranchordás, S. (2020). *The Regulation of Social Media Influencers*. Edward Elgar Pub .
- Google Cloud Vision AI. (2019). Demo Google Cloud Visual AI. Recuperado el 23 de Junio de 2019, de <https://cloud.google.com/vision/>
- Google Trends. (2018). Recuperado el 25 de Junio de 2018, de <https://trends.google.es/trends/explore?cat=44&date=today%205-y&geo=PE&q=%2Fm%2F0glpjll>
- Guerra, G., & Salazar, Y. (2016). *Implementación y desarrollo del plan de marketing para el Centro Estético IPS Bioestetik Express*. Tesis de pregrado, Universidad Cooperativa de Colombia. Obtenido de <http://repository.ucc.edu.co/handle/ucc/10692>
- Guntuku, S., Roy, S., & Weisi, L. (2015). Evaluating visual and textual features for predicting user 'likes'. *IEEE International Conference on Multimedia and Expo (ICME)*. Turín. doi:10.1109/ICME.2015.7177381
- Gürkaynak, G., Kama, Ç., & Ergün, B. (25 de Marzo de 2018). *Navigating the Uncharted Risks of Covert Advertising in Influencer Marketing*. Obtenido de Business Law Review, [e-journal] vol. 39, no. 1, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- Harrell, F. (2015). *Regression Modeling Strategies* (Segunda ed.). Springer.
- Hassan, A. (2014). *Do Brands Targeting Women Use Instamarketing Differently: A content analysis*. Obtenido de Marketing Management Association 2014 Annual Spring Conference Proceedings, : Available Online: <http://www.mmaglobal.org/publications/Proceedings/2014-MMA-SpringConference-Proceedings.pdf#page=78>
- Hausman, A., & Toivonen, T. (2017). Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas. *Conservation Letters*. doi:10.1111/conl.12343

- Henning-Thurau, T., Gwinner, K., Walsh, G., & Gremler, D. (2004). *Electronic Word-of-mouth via Consumer-opinion Platforms: What motivates consumers to articulate themselves on the internet?* Obtenido de Journal of Interactive Marketing, [e-journal] vol. 18, no. 1, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Hidalgo, M., & Segarra, J. (2017). The Youtuber Phenomenon and its Transmedia Expansion. Analysis of Youth Empowerment in Social Media. *Revista de Investigación USAL*, 3(15). doi:10.14.201
- Hinton, S., & Hjorth, L. (2013). *Understanding Social Media*. London: SAGE Publications Ltd.
- Hosseini, H., Xiao, B., & Poovendran, R. (2016). Google's Cloud Vision API is Not Robust to Noise. *16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. Cancun: IEEE. doi:10.1109/ICMLA.2017.0-172
- Hu, M., Milner, J., & Wu, J. (2016). *Liking and Following and the Newsvendor: Operations and marketing policies under social influence*. Obtenido de Management science, [e-journal] vol. 62, no. 3, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch> [
- Huang, Z., Zhou, E., & Cao, Z. (2015). Coarse-to-fine Face Alignment with Multi-Scale Local Patch Regression. *arXiv*. Obtenido de https://www.researchgate.net/publication/284096812_Coarse-to-fine_Face_Alignment_with_Multi-Scale_Local_Patch_Regression
- Jason, L., & Glenwick, D. (2016). *Handbook of methodological approaches to community-based research : qualitative, quantitative, and mixed methods*. Oxford University Press.
- Jimenez, D., Paz-Arias, H., & Larco, A. (2015). Desarrollo de un sistema inteligente para la clasificación de documentos ya digitalizados aplicando redes neuronales supervisadas. *Revista Tecnológica ESPOL – RTE*, 28(1), 8-23.
- Kamel, M., Giustini, D., & Wheeler, S. (2016). Instagram and WhatsApp in Health and Healthcare: An Overview. *Future Internet*, 8(37). doi:10.3390/fi8030037
- Kapitan, S., & Silvera, D. (2016). From Digital Media Influencers to Celebrity Endorsers: Attributions drive endorser effectiveness. 27(3). Obtenido de Marketing Letters, [e-journal]: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Kavuncuoglu, H., Kavuncuoglu, E., Karatas, S., Benli, B., Sagdic, O., & Yalcin, H. (2018). Prediction of the antimicrobial activity of walnut (*Juglans regia* L.) kernel aqueous

- extracts using artificial neural network and multiple linear regression. *Journal of microbiological methods*, 148, 78-86.
- Khademi, F., Jamal, S., Deshpande, N., & Londhe, S. (2015). Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression. *International Journal of Sustainable Built Environment*, 5(2), 355-369.
- Kholod, T., & Balkan, S. (2014). Is Video Analytics a Game Changer for Market Research? *DSS 2.0-Supporting Decision Making with New Technologies*.
- Klette, R. (2014). *Concise Computer Vision: An Introduction into Theory and Algorithms*. Springer.
- Kothari, C. (2004). *Research Methodology, Methods and Techniques* (Segunda ed.). New Age International.
- Kumar, A. (2016). An Incorporation of Artificial Intelligence Capabilities in Cloud Computing. *International Journal Of Engineering And Computer Science*, 5(11), 19070-19073.
- Kunz, W., Askoy, L., Bart, Y., & Theoloudilis, E. (2017). Customer Engagement in a Big Data World. *Journal of Services Marketing*, 31(2). doi:10.1108/JSM-10-2016-0352
- Kurtner, M., Nachtsheim, C., & Neter, J. (2004). *Applied Linear Regression Models* (Cuarta ed.). McGraw-Hill Irwin.
- Lay, A., & Fewerda, B. (2018). Predicting Users' Personality Based on Their 'Liked' Images on Instagram. *Humanize*.
- Lin, M., Lu, K. Y., & Wu, J. J. (2012). The effects of visual information in eWOM communication. *Journal of Research in Interactive Marketing*, 6(1), 7-26.
- Locharoenrat, K. (2017). *Research Methodologies for Beginners*. Pan Stanford.
- Lu, L., Chang, W., & Chang, H. (2014). *Consumer Attitudes Toward Blogger's Sponsored Recommendations and Purchase Intention: The effect of sponsorship type, product type, and brand awareness*. Obtenido de Computers in Human Behavior, [e-journal] vol. 34, no.1, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- M4rketng E Commerce. (2019). 14.000 millones de likes en un año: repasamos las grandes cifras de Instagram en 2018. Recuperado el 22 de Abril de 2019, de <https://marketing4ecommerce.net/14-000-millones-likes-cifras-de-instagram-2018/>

- Malhotra, C., Nunan, D., & Birks, D. (2017). *Marketing Research* (Quinta ed.). Trans-Atlantic Publications.
- Martínez, J., & Padilla, L. (2020). Innovación organizacional y competitividad empresarial. *Revista de ciencias sociales*, 26(2), 120-132.
- Matich, D. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Material de Clase, Universidad Tecnológica Regional - Rosario. Recuperado el 22 de Diciembre de 2018, de https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monografias/matich-redesneuronales.pdf
- Mochon, D., Johnson, K., Schwartz, J., & Ariely, D. (2017). *What Are Likes Worth? A Facebook page field experiment*. . Obtenido de Journal of Marketing Research, [e-journal] vol. LIV, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- Muñoz, C., & Towner, T. (2017). The Image is the Message: Instagram Marketing and the 2016 Presidential Primary Season. *Journal of Political Marketing*, 16(3-4). doi:10.1080/15377857.2017.1334254
- Nebi. (2017). *Customer Engagment y su importancia para las emppresas*. Recuperado el 22 de Junio de 2018, de <http://www.nebi.co/bi/customer-engagement-y-su-importancia-para-las-empresas/>
- Nurbayani, S., & Annuraini, A. (2017). The New Way of Beauty Experience from Korean Wave as Vocational Guidance in Indonesia. *Innovation of Vocational Technology Education*, 13(2).
- Ordenes, F., & Zhang, S. (2019). From words to pixels: text and image mining methods for service research. *Journal of Service Management*. doi:10.1108/JOSM-08-2019-0254
- Osborne, J., & Waters, E. (2002). Four Assumptions Of Multiple Regression That Researchers Should Always Tests. *Practical Assement, Research & Evaluation*, 8(2).
- Ose, S. (2016). Using Excel and Word to Structure Qualitative Data. *Journal of Applied Social Science*, 10(2). doi:10.1177/1936724416664948
- Peru Retail. (2016). Consumidor peruano gasta entre el 10% y 15% de su sueldo mensual en belleza y estética. Recuperado el 19 de Junio de 2019, de <https://www.peru-retail.com/consumidor-peruano-gasto-belleza-estetica/>

- Pornsakulvanich, V. (2017). *Personality, Attitudes, Social Influences, and Social Networking Site Usage Predicting Online Social Support*. . Obtenido de *Computers in Human Behavior*, [e-journal] vol. 76, Available through: LUSEM Library website : <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- Qyan, C., Tang, J., Penza, M., & Ferri, C. (2015). *Instagram Popularity Prediction via Neural Networks and Regression Analysis*. Princeton University, Computer Science. Recuperado el 12 de Mayo de 2018, de http://cjqian.github.io/docs/instagram_paper.pdf
- Rajabi Vandechali, M., Abbaspour-Fard, M., & Rohani, A. (2017). Multiple linear regression analysis of some important performance parameters of a conventional diesel engine in different working conditions. *Modares Mechanical Engineering*, 17(5), 363-373.
- Reece, A., & Danforth, C. (2016). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6, 15-21.
- Roncha, A., & Radclyffe-Thomas, N. (2016). *How TOMS' "One Day Without Shoes" Campaign Brings Stakeholders Together and Co-creates Value for the Brand Using Instagram as a Platform*. Obtenido de *Journal of Fashion Marketing and Management*, [e-journal] vol. 20, no. 3, Available through: LUSEM Library website: [e https://www.lub.lu.se/en/search-systems-and-tools/lubsearch](https://www.lub.lu.se/en/search-systems-and-tools/lubsearch)
- Ross, S. (2019). Being Real on Fake Instagram: Likes, Images, and Media Ideologies of Value. *Journal of Linguistic Anthropology*, 29(3), 359-374.
- Sánchez, H., Reyes, C., & Mejía, K. (2018). *Manual de términos en investigación científica, tecnológica y humanística*. Lima: Universidad Ricardo Palma.
- Shalev, E., & Morwitz, V. (2008). *The Surprising Influencers: How the inferred attributes of observed consumers shape observer consumers' buying intention, working paper*. Obtenido de *Advances in consumer research*, [e-journal] vol. 35, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Social Blade. (2019). Recuperado el 2 de Junio de 2019, de <https://socialblade.com/>
- Stokman, H. (2014). The Future of Smart Photography. *IEEE MultiMedia*, 21(3), 66-71.
- Stone, M., & WoodCook, M. (2013). Social intelligence in customer engagement. *Journal of Strategic Marketing*. doi:10.1080/0965254X.2013.801613
- Sudha, M., & Sheena, K. (2017). Impact of Influencers in Consumer Decision Process: The fashion industry. *SCMS Journal of Indian Management [e-journal]*, 14(3). Obtenido de

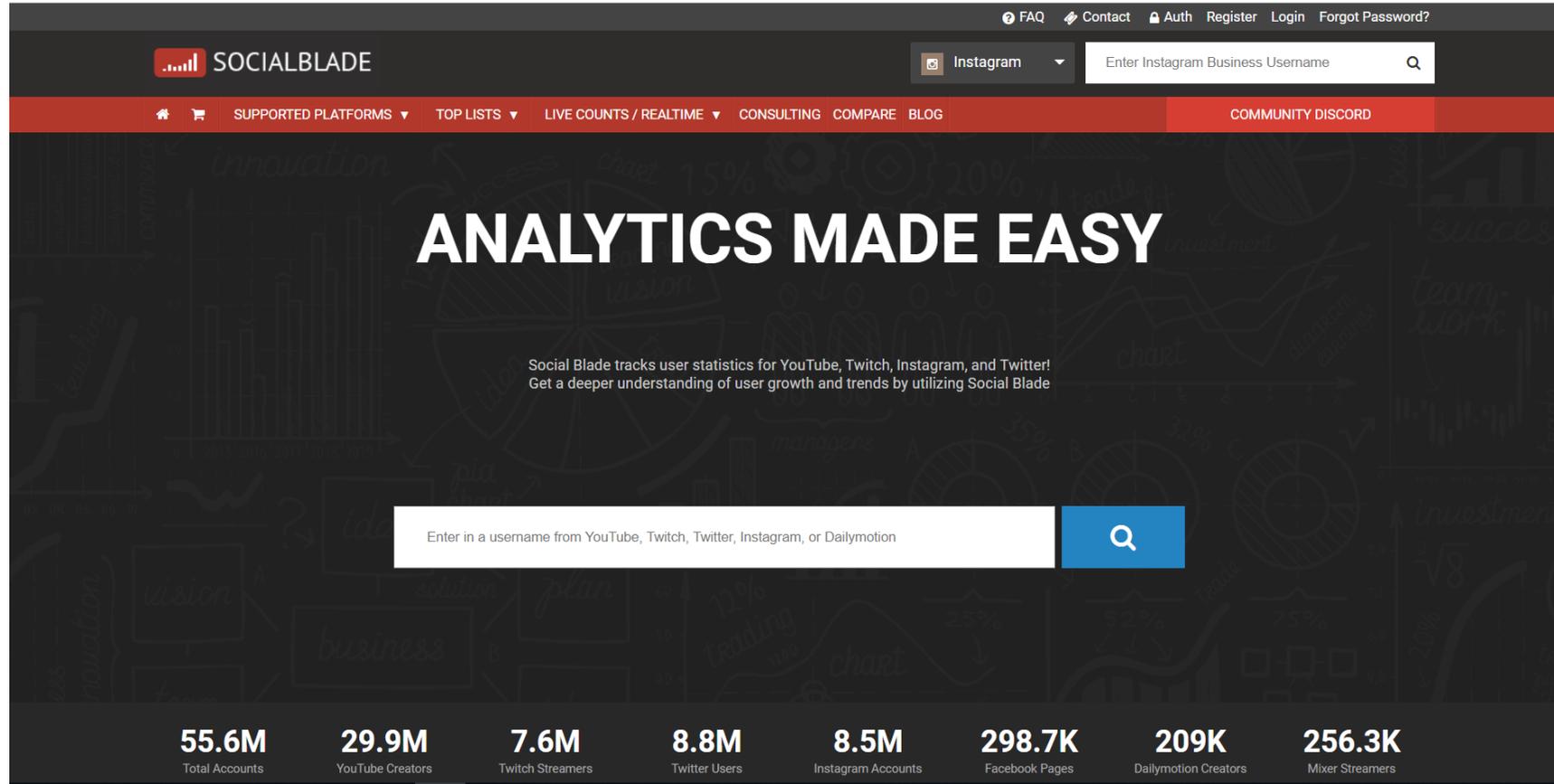
- SCMS Journal of Indian Management, [e-journal] vol. , no. 3, Available: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- Sudha, M., & Sheena, K. (2017). *Impact of Influencers in Consumer Decision Process: The fashion industry*. Obtenido de SCMS Journal of Indian Management, [e-journal] vol. 14, no. 3, Available: <https://www.lub.lu.se/en/search-systems-andtools/lubsearch>
- Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer.
- Tatar, A., Dias de Amorin, M., Fdida, S., & Antoniadis, P. (2014). A survey on predicting the popularity of web content. *Journal of Internet Services and Applications*, 5(8). doi:10.1186/s13174-014-0008-y
- Terttunen, A. (2017). *The influence of Instagram on consumers' travel planning and destination choice*. Tesis de maestría. Obtenido de https://www.theseus.fi/bitstream/handle/10024/129932/Terttunen_Anna.pdf?sequence=1
- Thomas, V., & Vinuales, G. (2017). *Understanding the Role of Social Influence in Piquing Curiosity and Influencing Attitudes and Behaviors in a Social Network Environment*. Obtenido de Psychology and Marketing, [e-journal] vol. 34, no. 9, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>
- Ting, H., Ming, W., De Run, E., & Choo, S. (2015). Beliefs About the Use of Instagram: An exploratory study. *International Journal of Business and Innovation*, vol. 2, no. 2,.
- Van Gerven, M., & Bohte, S. (2017). Editorial: Artificial Neural Networks as Models of Neural Information Processing. *Frontiers in Computational Neuroscience*, 19. doi:10.3389/fncom.2017.00114
- Vatcheva, K., Lee, M., McCormick, J., & Rahbar, M. (2016). Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiol*, 6(227). doi:10.4172/2161-1165.1000227
- Vignisdóttir, A. (2016). *CUSTOMER ENGAGEMENT ON INSTAGRAM BRAND PAGES IN THE MAKE-UP COSMETICS INDUSTRY*. Recuperado el 12 de Junio de 2018, de https://skemman.is/bitstream/1946/26717/1/MAR0117_Thesis_Annika_Vignisdottir_Customer%20Engagement%20on%20Instagram%20Brand%20Pages%20In%20the%20Make-up%20Cosmetics%20Industry.pdf
- Vives, J. (2 de Mayo de 2019). Instagram quiere ocultar el número de likes de las fotos. *La Vanguardia*. Recuperado el 1 de Noviembre de 2019, de

<https://www.lavanguardia.com/tecnologia/20190502/461995110682/instagram-quiere-ocultar-numero-likes-fotos.html>

- Vizcarra, E. (2018). Perú en Redes Sociales 2018 (Parte I). *La República*. Recuperado el 15 de Octubre de 2019, de <https://larepublica.pe/sociedad/1198456-peru-en-redes-sociales-2018-parte-i/>
- Voorveld, H., Van Noort, G., Muntinga, D., & Bronner, F. (5 de Abril de 2018). *Engagement With Social Media and Social Media Advertising: The differentiating role of platform type*. Obtenido de Journal of Advertising, [e-journal] vol. 47, no. 1, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch> [Accessed 5 April 2018]
- Wendler, T., & Gröttrup, S. (2016). *Data Mining with SPSS Modeler, Theory, Exercises and Solutions*. Berlin: Springer.
- Xiao, X., Xu, H., & Xu, S. (2015). Using IBM SPSS modeler to improve undergraduate mathematical modelling competence. *Computer Applications in Engineering Education* , 23(4), 603-609. doi:10.1002/cae.21632
- Yu, C. (2014). *Visualizing brand personality and personal branding : case analysis on Starbucks and Nike brand value co-creation on Instagram*. Tesis de maestría, Iowa. Obtenido de <http://ir.uiowa.edu/etd/1304/>
- Zhang, M., Guo, L., & Hu, M. (2017). Influence of customer engagement with company social networks on stickiness: Mediating effect of customer value creation. *International Journal of Information Management*, 37(3), 229-240. doi:10.1016/j.ijinfomgt.2016.04.010
- Zolkepli, I., Hasno, H., & Mukhiar, S. (2015). *Online Social Network Citizen Engagement on Instagram Crowdsourcing: A conceptual framework*. Obtenido de The Electronic Journal of Knowledge Management, [e-journal] vol. 13, no 4, Available through: LUSEM Library website: <https://www.lub.lu.se/en/search-systems-and-tools/lubsearch>

Anexos

Anexo 1: Interface de Social Blade



The screenshot displays the Social Blade website interface. At the top, there is a navigation bar with links for FAQ, Contact, Auth, Register, Login, and Forgot Password?. The main header features the Social Blade logo and a search bar for Instagram Business Usernames. Below the header is a secondary navigation bar with links for Home, Supported Platforms, Top Lists, Live Counts / Realtime, Consulting, Compare, Blog, and Community Discord. The main content area has a dark background with the text "ANALYTICS MADE EASY" and a sub-header: "Social Blade tracks user statistics for YouTube, Twitch, Instagram, and Twitter! Get a deeper understanding of user growth and trends by utilizing Social Blade". A search bar is present with the placeholder text "Enter in a username from YouTube, Twitch, Twitter, Instagram, or Dailymotion". At the bottom, a row of statistics is displayed:

Platform	Count
Total Accounts	55.6M
YouTube Creators	29.9M
Twitch Streamers	7.6M
Twitter Users	8.8M
Instagram Accounts	8.5M
Facebook Pages	298.7K
Dailymotion Creators	209K
Mixer Streamers	256.3K

Figura A1: Interface de Social Blade

Fuente: Social Blade (2019)

Anexo 2: Uso de Downloader for Instagram

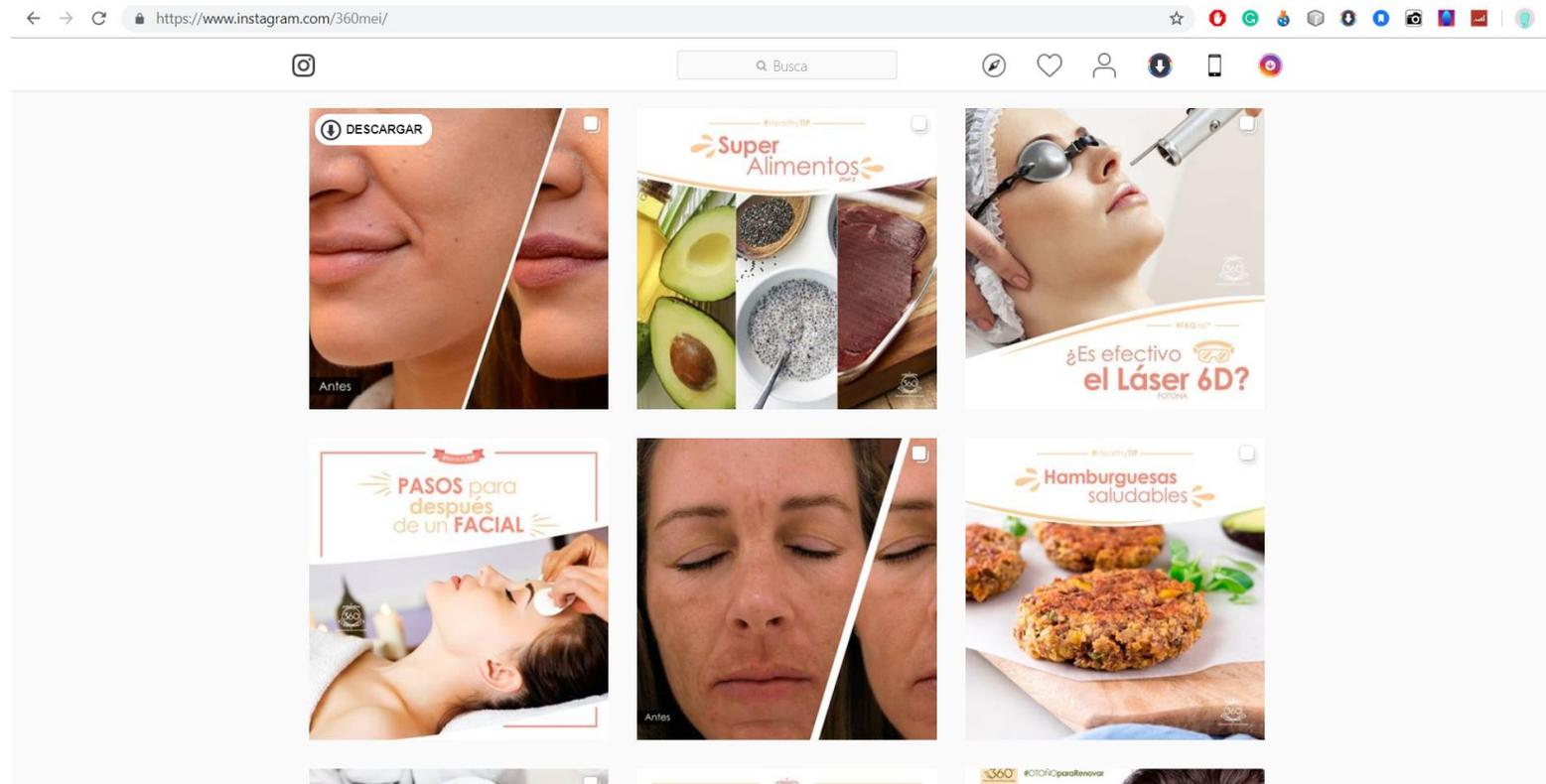


Figura A2: Interface de Social Blade

Fuente: Instagram (2019)

Nota: Observar en la imagen superior izquierda un botón de descarga que no aparece en la versión normal de Instagram

Anexo 3: Interface de Eyeem Vision

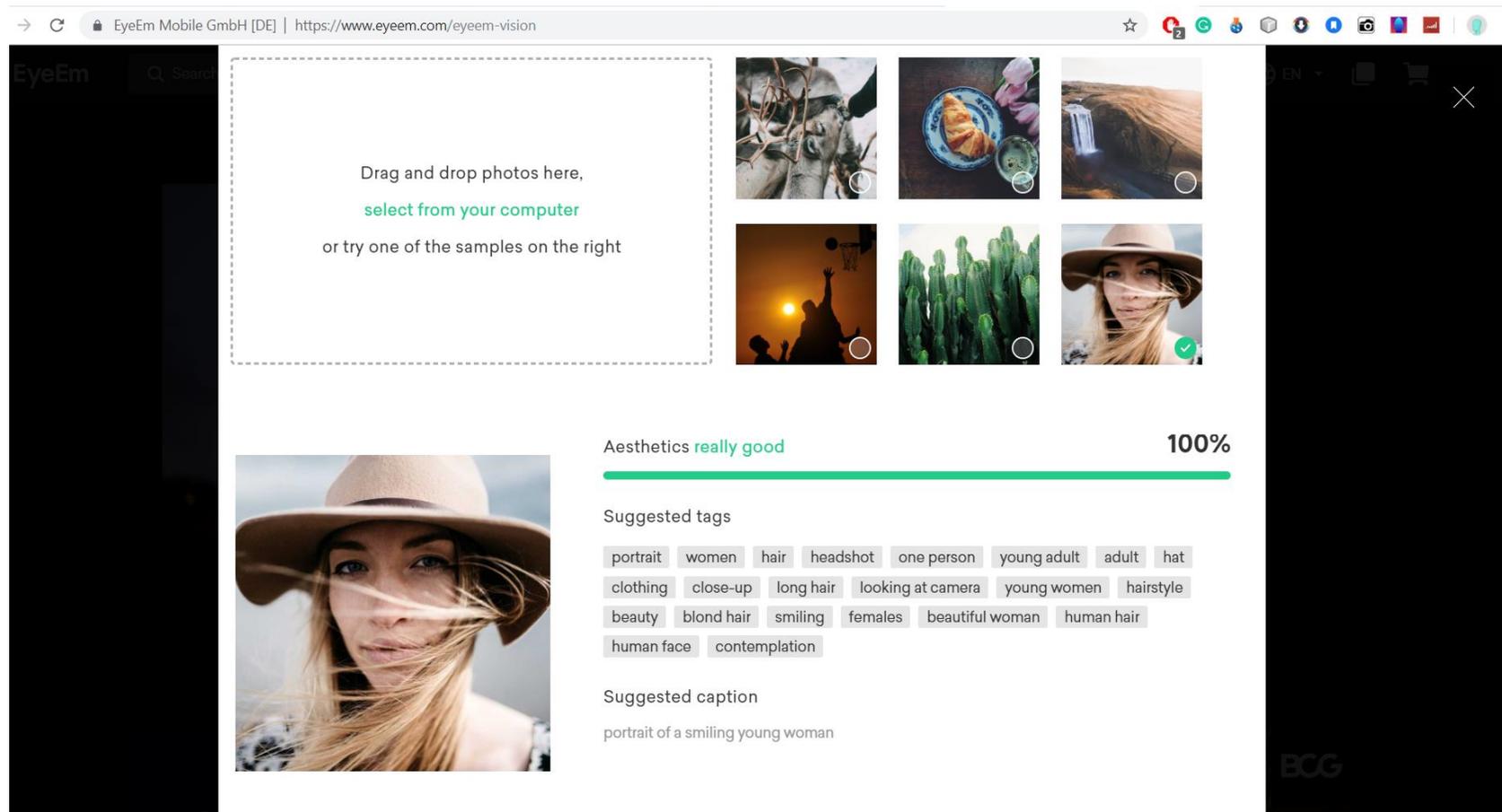


Figura A3: Interface de Eyeem Vision

Fuente: Eyeem Vision (2019)

Anexo 4: Interface de Every Pixel

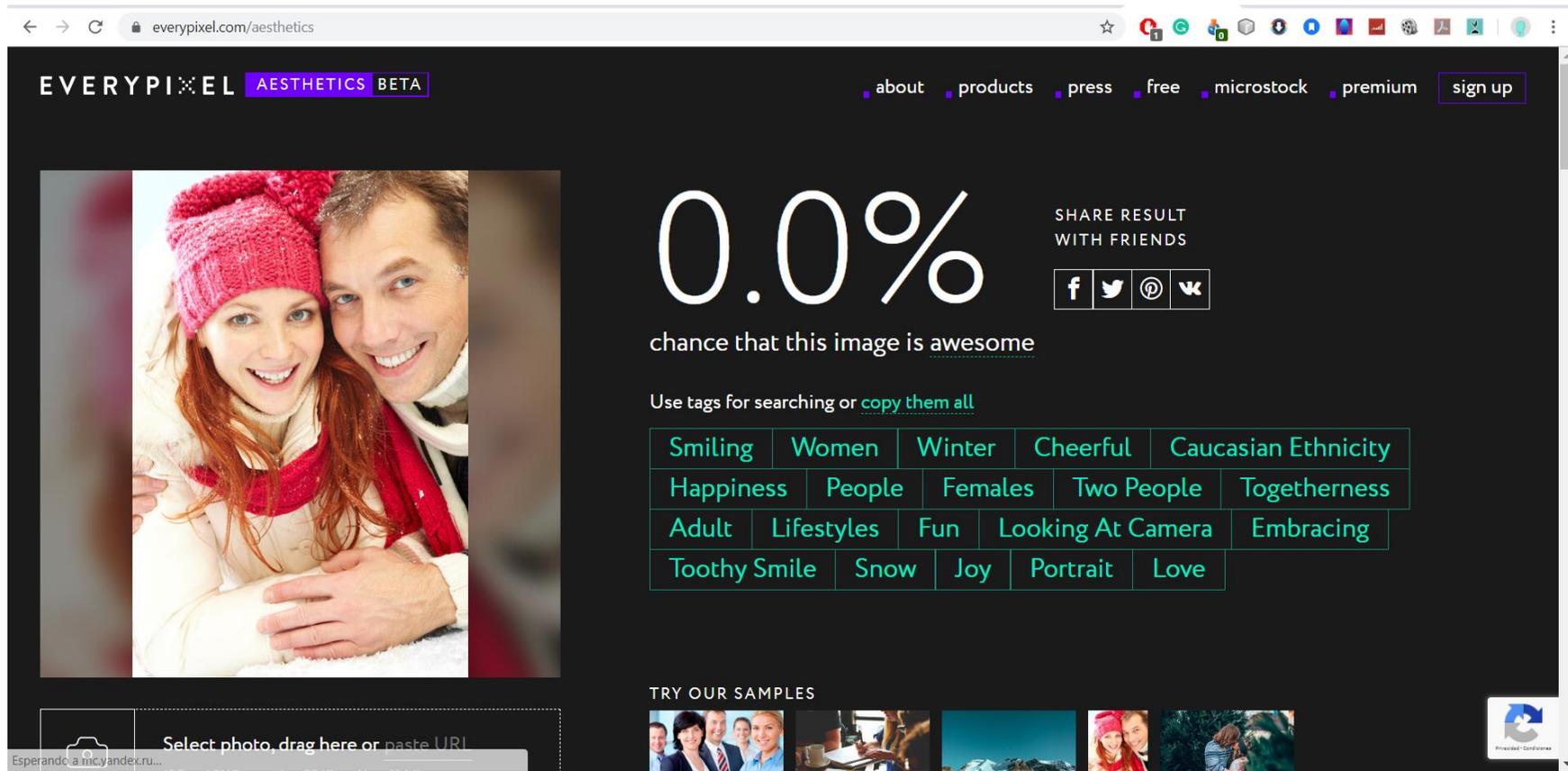


Figura A4: Interface de Eyeem Vision

Fuente: EveryPixel (2019)

Anexo 5: Interface de Google Cloud Vision AI para medir imágenes propias de la medicina

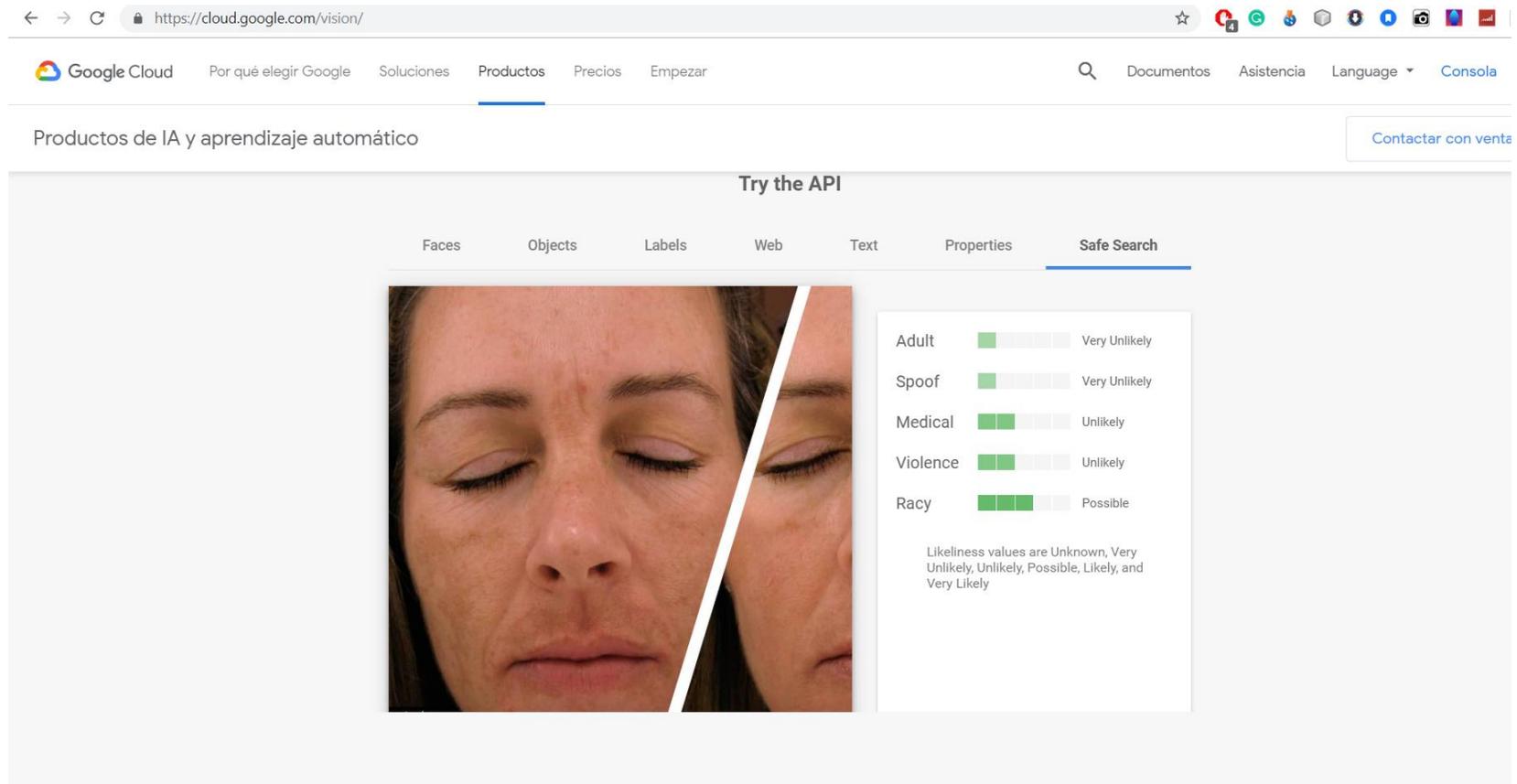


Figura A5: Interface de Google Cloud para la clasificación de imágenes vinculadas a la medicina

Fuente: Vision Cloud AI de Google (2019)

Nota: El puntaje relaciona a las imágenes médicas se encuentra en la derecha (ver categoría Medical). En este caso el puntaje es de nivel 2

Anexo 6: Interface de Vision AI para medir la cantidad de colores

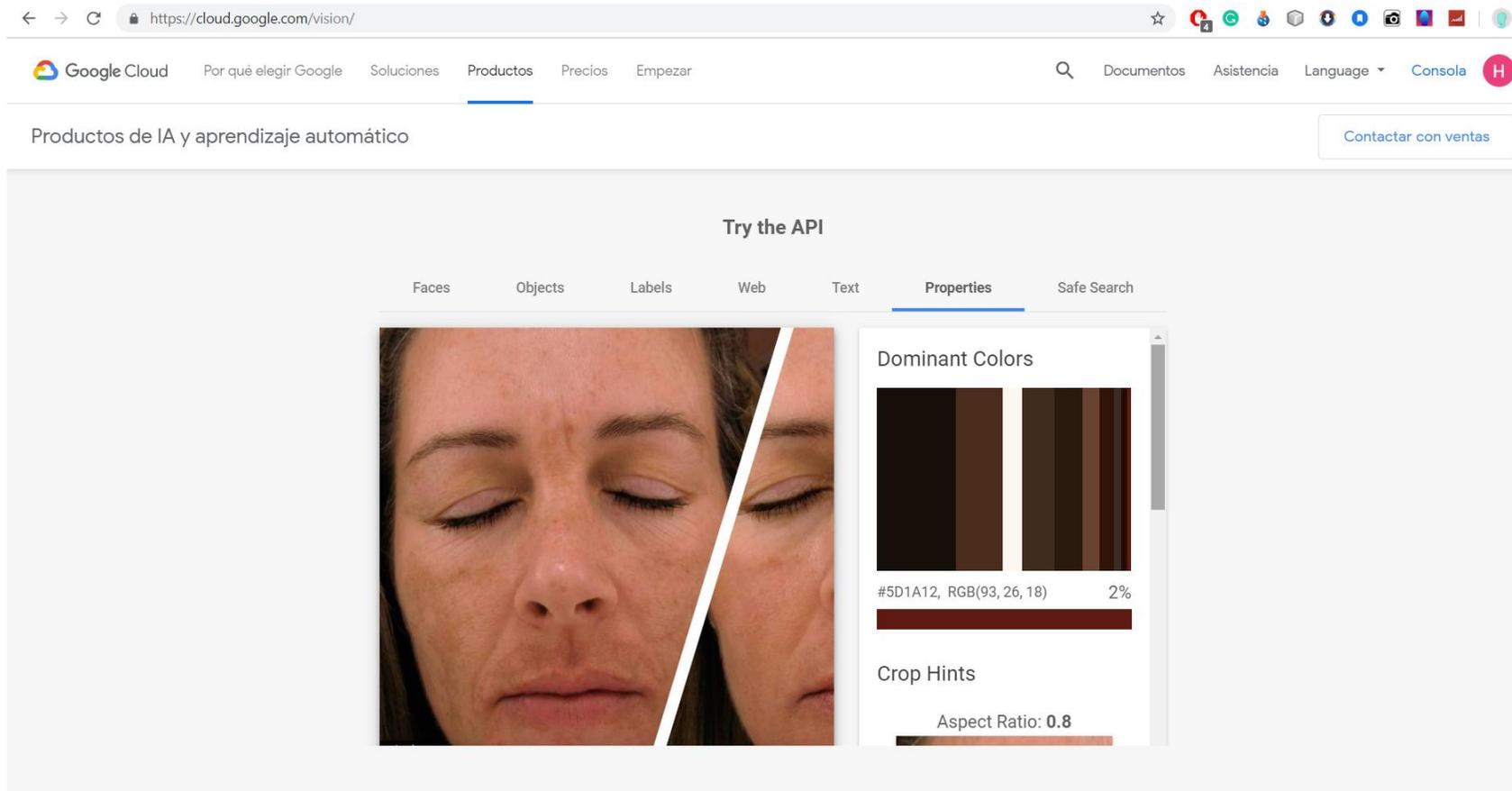


Figura A6: Interface de Google Cloud para en donde se aprecian los colores dominantes

Fuente: Google Cloud Visual AI (2019)

Nota: Para contabilizar la cantidad de colores es necesario pasar el cursor por la margen derecha en *Dominant Colors*

Anexo 7: Interface del SPSS Modeler

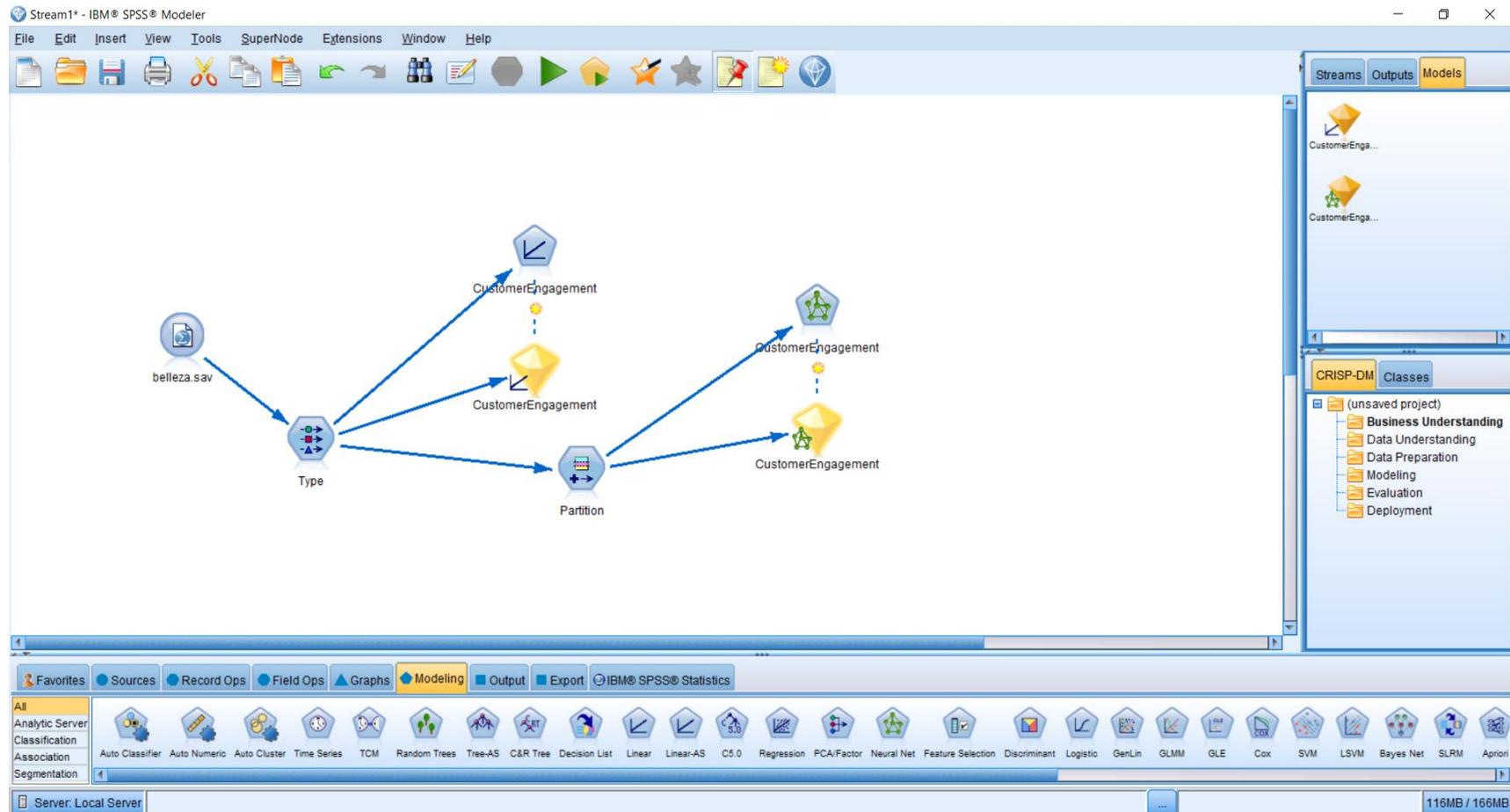


Figura A7: Interface de uso de SPSS Modeler

Fuente: IBM (2019)

Anexo 8: Matriz de Consistencia

Problema	Objetivos	Hipótesis	Variable	Metodología	Muestra	Plan de análisis
¿Existe relación entre el Customer Engagement de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima?	Objetivo General: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis General: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y un grupo de métricas en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	<i>Customer Engagement</i>	Cuantitativa	491 casos	Regresión Lineal
	Objetivo Específico 1: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de seguidores que se tenía en dicho momento en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 1: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de seguidores que se tenía en dicho momento en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Seguidores			
	Objetivo Específico 2: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el nivel en que la foto podría ser clasificada como propia de un procedimiento médico en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 2: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el nivel en que la foto podría ser clasificada como propia de un procedimiento médico en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	procedimiento_médico	Correlacional No experimental		
	Objetivo Específico 3: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de días que ha permanecido en dicha red social en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 3: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de días que ha permanecido en dicha red social en el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	cantidad_dias			
	Objetivo Específico 4: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de palabras visibles en la misma para el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 4: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de palabras visibles en la misma para el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	cantidad_palabras			
	Objetivo Específico 5: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el score correspondiente a la belleza de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 5: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el score correspondiente a la belleza de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	score_belleza			
	Objetivo Específico 6: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el score correspondiente al atributo sorprendente de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 6: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y el score correspondiente al atributo sorprendente de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	score_awesome			
	Objetivo Específico 7: Determinar si existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de colores de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	Hipótesis Específica 7: Existe relación entre el <i>Customer Engagement</i> de una imagen compartida en Instagram y la cantidad de colores de la misma el caso de publicaciones realizadas desde cuentas de centros estéticos localizados en Lima.	cantidad_colores			

Elaboración: Propia