

# Trabajo de Grado

**CAOBA – PONTIFICIA UNIVERSIDAD JAVERIANA**



**Sergio Armando Losada Bello  
Libardo Andrés Leal Verdugo**

**Asesores:**

**Juan Pablo Pájaro Hernández  
Juan Pablo Mora López**

**Trabajo de grado para optar al título de Magister en Analítica para la Inteligencia de  
Negocios**

**Noviembre de 2021**

**Maestría en Analítica para la Inteligencia de Negocios  
Facultad de Ingeniería  
Pontificia Universidad Javeriana  
Bogotá D.C**

## INTRODUCCIÓN

Las redes sociales se comportan como comunidades virtuales donde las personas por medio de un usuario intercambian información y datos en diversas formas, por videos, textos, imágenes o una combinación de estos, creando redes de conexión entre los usuarios con pensamientos en común o diversos independientemente del tema de interés. En este comportamiento social convergen grandes volúmenes de datos por diferentes medios de interacción, páginas web, redes sociales o blogs de opinión. Todo este flujo de opiniones puede convertirse en información estratégica de cara a encontrar solución a un problema de negocio específico. La idea de este proyecto es desarrollar una solución estratégica de detección de los alimentos publicados en redes sociales por medio de fotografías e imágenes compartidas por los usuarios.

Este trabajo tiene como base la metodología CRISP DM con el fin de garantizar un correcto desarrollo prueba y experimentación de los modelos considerados para el objetivo del trabajo, adicionalmente, para la construcción de los modelos se realizó un proceso de transfer learning con modelos del reconocimiento de imágenes, los cuales enfocados en el negocio específico de Nutresa como usuario final del modelo, permitirán una implementación rápida a la necesidad de la empresa.

## CONTENIDO

ENTENDIMIENTO DE NEGOCIO .....	5
1.1 Contexto de la industria de alimentos y las redes sociales .....	5
1.1.1 Movimiento Foodie .....	5
1.1.2 Recomendaciones gastronómicas.....	6
1.1.3 Cifras de imágenes compartidas en las redes sociales .....	7
1.1.4 Ontología alimentaria.....	7
1.1.5 Estado del arte del entrenamiento de maquina en reconocimiento de imágenes ...	9
1.1.6 Uso de la Red Neuronal Convolutiva en la detección de imágenes .....	10
1.1.7 Reconocimiento de imágenes en Python.....	11
1.1.8 Análisis comparativo entre diferentes arquitecturas de redes neuronales profundas.....	13
1.1.9 Evolución de las redes neuronales profundas .....	13
1.1.10 Comparativo de diferentes arquitecturas.....	13
1.2 Pregunta de Negocio.....	15
1.3 Objetivo de negocio .....	16
1.4 Objetivos específicos del negocio.....	16
1.5 Métricas de Negocio .....	16
1.6 Hardware & Software .....	16
1.7 Requerimientos del proyecto .....	17
1.8 Supuestos del proyecto .....	17
1.9 Restricciones del proyecto .....	17
1.10 Objetivos del Data Mining.....	18
1.11 Métricas de Data Mining .....	18
1.11.1 Pipeline del proyecto.....	18
1.11.2 Pipeline analítico .....	19
1.11.3 Pipeline del Modelamiento y Evaluación.....	21
EXPLORACIÓN DE LOS DATOS.....	22
2.1 Ontología .....	22
2.2 Construcción del set de datos.....	22
2.3 Entendimiento del set de datos .....	22

PREPARACIÓN DE DATOS.....	23
3.1 Limpieza de la Data .....	23
3.2 Selección de la data de entrenamiento:.....	25
3.3 Partición de la Data:.....	26
MODELAMIENTO Y EXPERIMENTACIÓN.....	28
4.1 Modelamiento .....	28
4.2 Arquitectura del proyecto .....	30
4.3 Protocolo de experimentación .....	30
4.4 Tratamiento de las imagenes.....	31
4.4.1 Transformación del set de datos de imágenes.....	31
4.4.2 Dummificación de la variable objetivo:.....	33
4.4.3 Normalización de las imágenes:.....	33
4.4.4 Tratamiento del set de imágenes:.....	33
4.5 Implementación de los modelos seleccionados .....	34
4.6 Resultados de los modelos .....	37
4.7 Selección del modelo Final.....	38
4.8 Evaluación .....	38
4.8.1 Validación de Resultados: .....	39
4.8.2 ETL de alimentos en redes sociales.....	39
4.8.3 Gridsearch Manual.....	40
CONCLUSIONES .....	43
RECOMEDACIONES.....	43
BIBLIOGRAFÍA .....	45
Anexo 1 .....	46
Implementación del Modelo .....	46

## ENTENDIMIENTO DE NEGOCIO

### 1.1 CONTEXTO DE LA INDUSTRIA DE ALIMENTOS Y LAS REDES SOCIALES

Cuando hablamos de redes sociales se asocia inmediatamente con opiniones masivas de las personas alrededor del mundo en temas personales o de interés común, estas opiniones en muchos casos son consideradas como inductivas para otros usuarios dependiendo su línea crítica, positiva o negativa, lo cual en términos de negocio puede ayudar al aumento o disminución en el consumo de un producto o servicio. Por esto cada vez son las industrias que se preocupan más por la opinión de sus clientes, dado que el “voz a voz” se ha convertido en un frecuente medio de mercadeo alrededor del mundo. Ante esta realidad, una de las tendencias de mayor auge en las redes sociales es el “Socialfoodie”.

La tendencia de publicar fotos de comida esta por todos lados tanto en Twitter, Facebook, Instagram o cual red social esté disponible en el mundo digital, se encuentran llenas de fotos de recetas, fotos de desayunos saludables, aperitivos apetitosos y un amplio menú de alimentos inundan las redes sociales, tanto que se ha convertido en el oficio diario de algunos.

#### 1.1.1 Movimiento Foodie

*Foodie* define a todas las personas apasionadas por la comida y las bebidas, esta tendencia popular prácticamente en todos los países de Europa, Reino Unido y América ya existía desde los años ochenta explica Beatriz Garcia periodista y especialista de comunicación audiovisual (García).

El movimiento foodie es más que una tendencia gastronómica, “*Se ha convertido en un estilo de vida*” explica García. Según la autora, existe confusión entre lo que es una persona *foodie* y una *gourmet*, estas últimas son personas muy selectas de paladar entrenado, en cambio un *foodie* está más preocupado por el consumo donde toda la información alrededor de la comida sea una experiencia (García).

Alejados de la comida tradicional, los *foodie* se dedican la mayor parte de su tiempo a conocer nuevas tendencias culinarias que cumplan con los requisitos de ser atractivas para el mercado tanto como por su sabor, su imagen y su experiencia, que pueda ser considerados de moda en la industria de las comidas. Esta tendencia ha incrementado la cantidad de festivales locales e internacionales donde los restaurantes promocionan sus productos impulsados por la proliferación de los *foodies* en todo el mundo.

De acuerdo con el diario La Vanguardia una de las necesidades vitales de la forma de ser un *foodie* es volcar sus opiniones y fotografías lo antes posible a las redes sociales (TOMÀS, 2013). Sin embargo, esto no sería posible sin seguidores, y hoy por hoy abundan los usuarios que buscan recomendaciones gastronómicas en las redes sociales, un *foodie* sin seguidores es como un pez fuera del agua.

No cabe duda de que las redes sociales y las aplicaciones para teléfonos inteligentes contribuyen cada vez más al crecimiento de esta tendencia de amantes de todo lo relacionado

con la comida. Natalia San Juan, experta en community management y redes sociales, además de aficionada a la fotografía, nos cuenta: “Existen varios fenómenos dentro de la tendencia *foodie* de los cuales las redes sociales se han hecho eco. Uno es, por ejemplo, la creciente tendencia en volcarse en la cocina, retomando el interés por guisar e intercambiar recetas, al mismo tiempo que aumenta la preocupación por una alimentación saludable. Todo ello lo vemos claramente reflejado en el creciente número de aplicaciones destinadas a compartir este tipo de contenidos y al intercambio de consejos y experiencias. En este contexto, si existe una red que ha puesto en el candelero a los foodies es, sin duda, Instagram” (TOMÀS, 2013)

### 1.1.2 Recomendaciones gastronómicas

Anteriormente encontrar un restaurante adecuado para la cita perfecta era una tarea de larga duración, la llegada de las redes sociales y los amantes a la comida “*foodies*” cambió el paradigma. En Colombia estos conocedores culinarios han venido aganado terreno y son tan populares como periodistas o deportistas.



**Figura 1:** Movimiento Foodie en el mundo y Colombia  
Fuente Diario La República

De acuerdo con Juan Andrés Mogollón, funcionario de Gold Fish, empresa encargada de manejar el marketing de la nueva generación de influenciadores en redes sociales, las personas están empezando a consumir el contenido partiendo de alternativas más visuales. “Instagram es una plataforma con mayor alcance donde los foodies conectan marcas con audiencias potenciales” (Murillo).

La tendencia en las redes sociales cada día tiene más impacto en las decisiones de los consumidores, es por esto que se debe estar constantemente vigilando los cambios del mercado en las plataformas digitales y más aún las opiniones de los usuarios.

### 1.1.3 Cifras de imágenes compartidas en las redes sociales

Parte del interés en el mundo de las comidas a través de redes sociales, llevo al siguiente análisis de Food and Social Media publicado por la periodista y bloguera Yuriana Cross a finales de 2018, este estudio fue publicado con cifras de las tendencias tanto en España como América Latina alrededor de la industria de la comida. (Coss, s.f.)

- ✓ El 29% de las personas de entre 18 y 35 años comparten fotos de platos (50% de 18 a 24 años).
- ✓ El 31% de encuestados dice que cocina platos después de descubrirlos en Facebook.
- ✓ El 69% de los Millennials toman una foto o un vídeo de su comida antes de comer.
- ✓ El 75% de los "foodies" están inspirados en las publicaciones de Facebook o Instagram en su consumo.
- ✓ #Food contabiliza más de 1 billón de interacciones al mes en Facebook, y más de 250 millones de publicaciones en Instagram.
- ✓ Los food lovers consumen 4 veces más de contenido en Instagram que un usuario promedio.
- ✓ El 40% de los usuarios de Facebook ven contenido de "Food" en la plataforma y un 38% en Instagram.
- ✓ El 28% de los usuarios de Facebook comparten contenido "Food" en la plataforma, y el 27% en Instagram.
- ✓ El universo de Food es muy consistente en Instagram en el mundo: es una de las categorías más grandes, una de las más populares. Los adictos a los alimentos se conectan a Instagram desde la mañana hasta la noche, con un promedio de 18 veces al día.
- ✓ La gastronomía ha generado 22 millones de tuits en todo el mundo en 2017, un post cada segundo y medio.
- ✓ Entre los usuarios que tuitean sobre gastronomía, el 62% tiene entre 18 y 34 años y el 28% son mujeres entre 18 y 24 años.
- ✓ Los 2 emojis principales tuiteados son Pizza (382k) y Pasteles (341k) 🍕 🍰
- ✓ El veganismo está a la vanguardia de estas discusiones con casi 680,000 tweets.
- ✓ Entre los usuarios que tuitean sobre gastronomía, el 62% tiene entre 18 y 34 años y el 28% son mujeres entre 18 y 24 años.
- ✓ La comida es la 5ta categoría más vista en YouTube frente a los juegos y la belleza
- ✓ Los mejores canales: en España "Cocina para todos"(1.7M), en México "Marisolpink" (4.1M), "FastGoodCuisine" en francés (2.4M), "Laura in the kitchen" en los EE. UU. (2.8M) o "Kinoshita Yuka" en Japón (+ 2.9M)

### 1.1.4 Ontología alimentaria

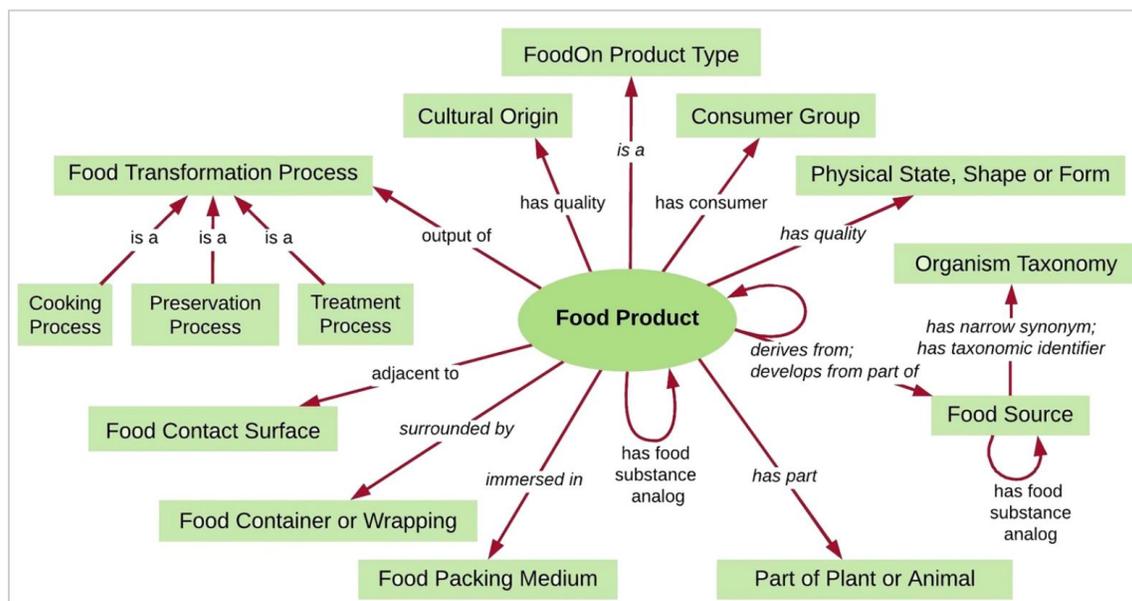
La ontología constituye un vocabulario sencillo que describe recetas, ingredientes, menús y dietas, basado en todos los alimentos disponibles para el consumo humano. Esta ontología permite acomodar las jerarquías pertinentes desde la producción, procesamiento y producto final del alimento, a menudo en forma de taxonomías. Las jerarquías permiten actuar como facetas desde una generalidad y poco a poco hacia el detalle y especificidad a medida que se avanza en la jerarquía dentro de la ontología, un ejemplo de esto es que en su mayoría los

productos alimenticios pueden vincular categorías de diferentes de menús nacionales e internacionales bajo la jerarquía de “tipo de producto”, así puede este contenido en otra categoría como “Fuente de alimento” bien sea animal o vegetal. (Dooley, y otros, 2018)

Una ontología facilita en términos de vocabulario los axiomas lógicos permitiendo a una computadora leer e interpretar la información asociada al alimento definido. La computadora estará en capacidad de razonar automáticamente si existen interferencias entre los supuestos establecidos, contradicciones no deseadas, incluso identificar ingredientes sospechosos capaces de transmitir enfermedades por alimentos.

La ambigüedad de las referencias alimentarias básicas es un problema común tanto para la industria del sector público y privado que se decidió resolver en consorcio entre varias organizaciones, entre estas, la Universidad de Columbia Británica y el Laboratorio de Salud Pública del Centro para el Control de Enfermedades de la Columbia Británica, desarrollaron el proyecto **FoodOn**, el cuál inicialmente se enfocó en la falta de estandarización básica de los alimentos asociados a investigaciones sobre brotes de enfermedades transmitidas por alimentos, este proyecto fue apoyado por OBO Foundry, (Foundry, s.f.) quienes dieron un paso adelante llenando los vacíos existentes en la ontología. Este proyecto promueve un sistema práctico y sencillo de la descripción de los alimentos, para que los investigadores, consumidores y la industria lo puedan reutilizar. Esta estandarización se enfocó en reducir esta ambigüedad mejora el uso a través de sistemas transaccionales o conectado por IoT y permite reducir los costos de intercambio de los datos.

El proyecto FoodOn decidió usar el idioma inglés con el fin de garantizar el uso del mismo a nivel mundial, el cual pueda integrar diccionarios terminológicos y sistemas de datos que traspasen fronteras, el cual garantice un uso en la descripción de los alimentos y su producción de forma coherente. (Dooley, y otros, 2018)



**Figura 2:** Diagrama de productos alimenticios y sus cualidades descriptivas, componentes y procesos relacionados.

**Fuente:** FoodOn - Ontología alimentaria.

FoodOn es un recurso de ontología integral de código abierto compuesto por facetas jerárquicas como se muestra en figura 2 de términos que cubren los ingredientes básicos de la fuente de alimentos crudos, los términos del proceso para el envasado, la cocción y la conservación, y una variedad de nivel superior de esquemas de tipos de productos en los que se pueden clasificar los productos alimenticios.

FoodOn tiene como objetivo cubrir los productos alimenticios y los pasos amplios del procesamiento de alimentos, actuando como un centro más generalista que interactúa con ontologías de dominio más especializadas que involucran el modelado técnico de la ciencia de los alimentos. (Dooley, y otros, 2018)

La siguiente figura muestra un árbol de 15 ramas tópicas dentro de FoodOn de nivel superior dentro de la cual para el reconocimiento de imágenes nos centraremos en tipo de producto alimenticio.

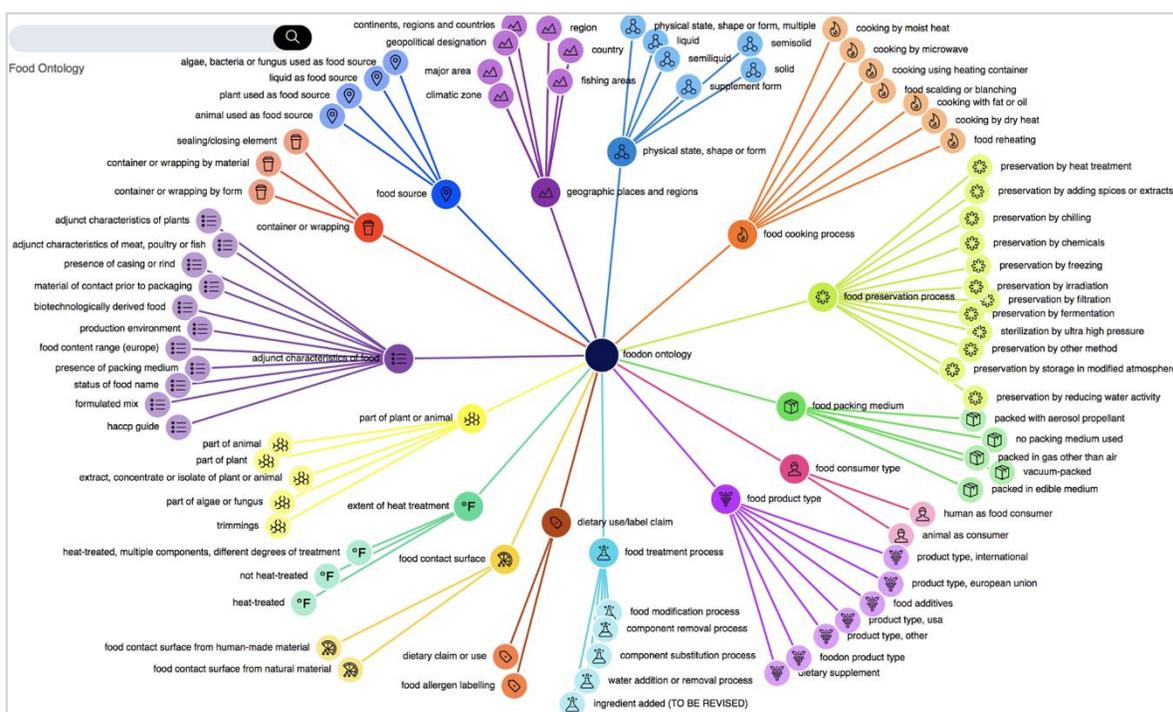


Figura 3: Diagrama de tópicos relacionados con los productos alimenticios

Fuente: FoodOn- Ontología alimentaria.

### 1.1.5 Estado del arte del entrenamiento de maquina en reconocimiento de imágenes

La importancia del reconocimiento de imágenes ha crecido en los últimos años gracias a la evolución de los algoritmos que año tras año hacen más precisa la identificación de estas. Este reconocimiento de imágenes es fundamental en la Inteligencia Artificial la cual ha visto en los últimos años, a medida que la capacidad y acceso a tecnologías de procesamiento y computo se incrementa, la aplicación de los algoritmos de entrenamiento de máquina también crecen exponencialmente, principalmente por la importancia y necesidad del reconocimiento

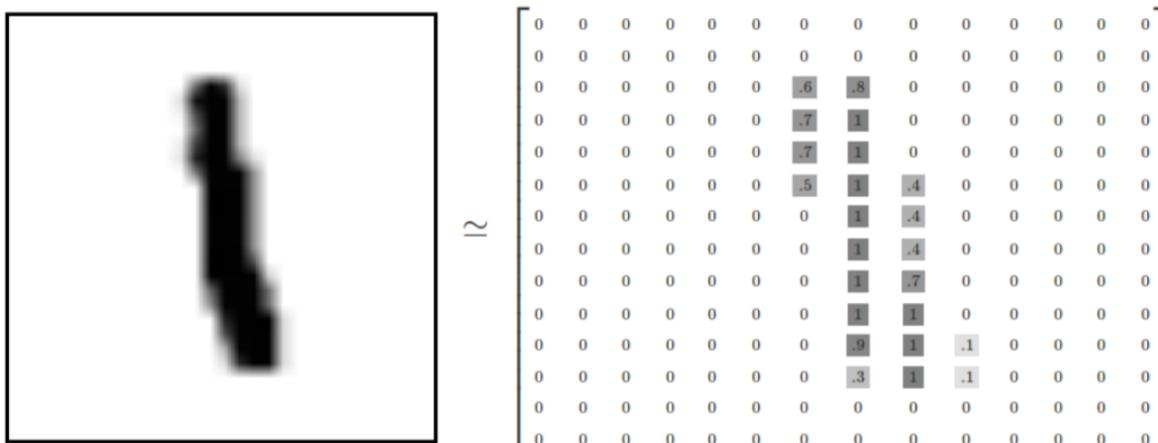
de patrones en tiempo real, los cuales suministran gran cantidad de información estableciendo vínculos y relaciones entre propiedades de los datos. Promotores de esto son los gigantes tecnológicos como lo son Google, Amazon, Facebook y Microsoft los cuales ofrecen soluciones basadas en Inteligencia Artificial, desde videojuegos hasta asistentes virtuales.

De las principales áreas que más han experimentado mayor evolución es el reconocimiento de imágenes, principalmente por la utilidad social de las aplicaciones desarrolladas. Tal es la eficiencia del reconocimiento de imágenes, que usos como el reconocimiento facial implica restricciones éticas bueno uso de la información detectada.

### 1.1.6 Uso de la Red Neuronal Convolutiva en la detección de imágenes

Un enfoque utilizado para la extracción de características usa aprendizaje profundo centrado en el entrenamiento de Redes Neuronales Convolucionales (CNN por sus siglas en inglés). La importancia de la utilización de las CNN está enfocada en la extracción de características robustas, posibilidad de extraer patrones complejos por la cantidad de atributos y el uso de valores tipo punto flotante. Así mismo este entendimiento de patrones ha hecho posible el desarrollo de modelos y técnicas en algoritmos de aprendizaje automático y aprendizaje profundo para mejorar la estructura de los datos y a su vez aumente la eficiencia de los algoritmos.

Los computadores reciben las imágenes de forma diferente a como lo recibimos los humanos, la imagen se transforma en un vector bidimensional con valores relacionados a pixeles. En la figura 4 se muestra un ejemplo del set de datos MNIST y como es recibida la imagen de un número interpretada por un computador. (Martínez Llamas & De Mingo)

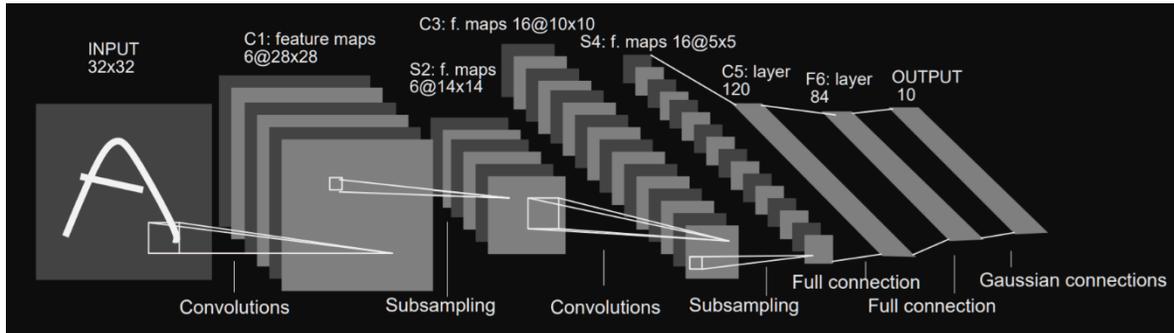


**Figura 4:** Vector de reconocimiento de imagen de MNIST.

**Fuente:** Universidad Politécnica de Madrid – Reconocimiento de Imágenes Mediante Redes Neuronales Convolucionales.

La red convolutiva sigue una secuencia basada en tres capas: Convolutional Layer, Pooling Layer y Fully-Connected Layer. Se efectuar las convoluciones y reducciones mediante una

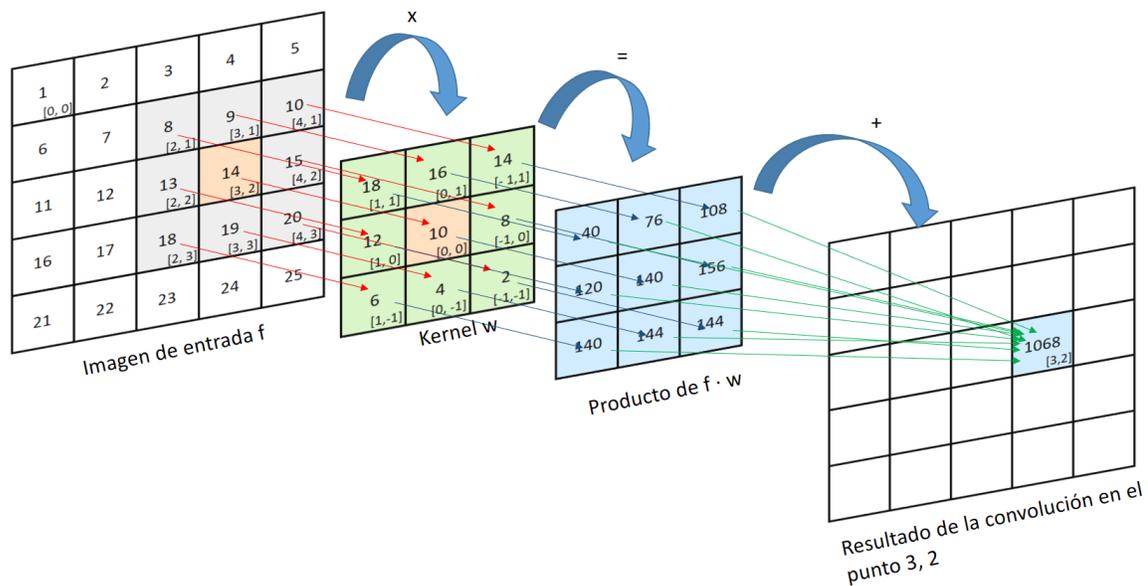
serie de capas conectadas (perceptrón multicapa) se obtiene la salida deseada. (LeCun, Bottou, Haffner, & Bengio, 2020)



**Figura 5:** Arquitectura LeNet-5.

*Fuente:* Entendimiento e Implementación de la Arquitectura CNN LeNet-5.

La convolución desarrolla un producto y suma entre dos matrices, una matriz de partida y una matriz de filtro o Kernel de tamaño menor, esto reduce la dimensión de la matriz de entrada lo cual permite generalizar las características de salida. (Medina, 2019)



**Figura 6:** Proceso de Convolución.

*Fuente:* Convolución en 2D (Filtrado espacial)

### 1.1.7 Reconocimiento de imágenes en Python

Unas de las herramientas más comunes para el reconocimiento y clasificación de las imágenes es TensorFlow y Keras.

El reconocimiento de imágenes se refiere a la tarea de introducir una imagen en una red neuronal y hacer que ésta produzca algún tipo de etiqueta para esa imagen. La etiqueta que

la red produzca corresponderá a una clase predefinida. Puede haber varias clases a las que la imagen pueda ser etiquetada, o sólo una. Si hay una sola clase, se suele aplicar el término "reconocimiento", mientras que una tarea de reconocimiento de varias clases se suele llamar "clasificación". (Nelson, s.f.)

Un subconjunto de la clasificación de imágenes es la detección de objetos, en la que se identifican instancias específicas de objetos como pertenecientes a una determinada clase, como animales, coches o personas.

#### - **TensorFlow**

TensorFlow es una librería abierta que compila diferentes algoritmos y modelos que permite tanto procesamiento de lenguaje natural, como para este caso, el reconocimiento de imágenes. Desarrollada por Google con el objetivo de satisfacer las necesidades del entrenamiento de máquina, fue publicada bajo la licencia de código abierto en el año 2015 y ha sido considerada un referente en el aprendizaje profundo y las redes neuronales.

#### - **Keras**

Keras consiste en una biblioteca de redes neuronales de alto nivel desarrollada para Python y se puede usar en la biblioteca de TensorFlow, fue integrado en 2017 como interfaz del código.

#### - **CUDA**

El alto número de cálculos que se efectúan en el desarrollo de redes convolucionales eleva la complejidad del entrenamiento, por tanto hacen que sea necesario un hardware de altas prestaciones, esto ha facilitado el uso de GPU's (Unidad de Procesamiento Gráfico) en el desarrollo de algoritmos de Deep Learning. La GPU usada principalmente en la ejecución de juegos, al usarse en el entrenamiento de redes neuronales permite reducir considerablemente los tiempos de procesamiento en comparación con el uso de la CPU. El anterior proceso se ve favorecido por el alto número de núcleos que permite el procesamiento en paralelo.

Debido a lo anterior, CUDA, la arquitectura de cálculo paralelo de NVIDIA y la librería para redes neuronales profundas cuDNN, permite el uso de GPUs en TensorFlow.

### **Extracción de características**

Para llevar a cabo el reconocimiento/clasificación de imágenes, la red neuronal debe realizar la extracción de características. Las características son los elementos de los datos que le interesan y que serán introducidos en la red. En el caso concreto del reconocimiento de imágenes, las características son los grupos de píxeles, como bordes y puntos, de un objeto que la red analizará en busca de patrones. (Nelson, s.f.)

El reconocimiento de rasgos (o la extracción de rasgos) es el proceso de extraer los rasgos relevantes de una imagen de entrada para poder analizarlos. Muchas imágenes contienen anotaciones o metadatos sobre la imagen que ayudan a la red a encontrar las características relevantes. (Nelson, s.f.)

*Técnicas de reducción de dimensionalidad:*

1. Análisis de componentes independientes
2. Isomap
3. Kernel PCA
4. Análisis semántico latente
5. Mínimos cuadrados parciales
6. Análisis de componentes principales
7. Reducción de dimensionalidad multifactorial
8. Reducción de dimensionalidad no lineal
9. Análisis multilineal de componentes principales
10. Aprendizaje subespacial multilineal
11. Incrustación semidefinida
12. Autoencoder

### **1.1.8 Análisis comparativo entre diferentes arquitecturas de redes neuronales profundas**

Un estudio realizado en octubre de 2018 por los autores Simone Bianco, Luigi Celona y Paolo Napoletano, puso en comparación diferentes modelos de redes neuronales profundas (DNNs) para el reconocimiento de imágenes. Este estudio evaluó diferentes prestaciones de cada DNN, tal como la precisión del reconocimiento, la complejidad del modelo, el costo computacional, la memoria usada y por último el tiempo empleado en cada uno. Para garantizar una correcta comparación de la eficiencia de los modelos, se corrieron en máquinas con diferentes capacidades computacionales. Esto permitirá tomar la decisión de cuales modelos de reconocimiento de imágenes por las variables de requerimiento de maquina y tiempo de ejecución podrían ser útiles en el marco del proyecto.

### **1.1.9 Evolución de las redes neuronales profundas**

Desde los primeros intentos de mejorar los porcentajes de precisión en modelos de reconocimiento de imágenes, las DNNs han tenido logros sobresalientes. AlexNet fue una de las primeras en mejorar la precisión del reconocimiento de imágenes en más de 10 puntos porcentuales por el año 2012, en comparación con los métodos tradicionales de ImageNet, a partir de este punto se han diseñado redes más precisas y computacionalmente más eficientes.

### **1.1.10 Comparativo de diferentes arquitecturas**

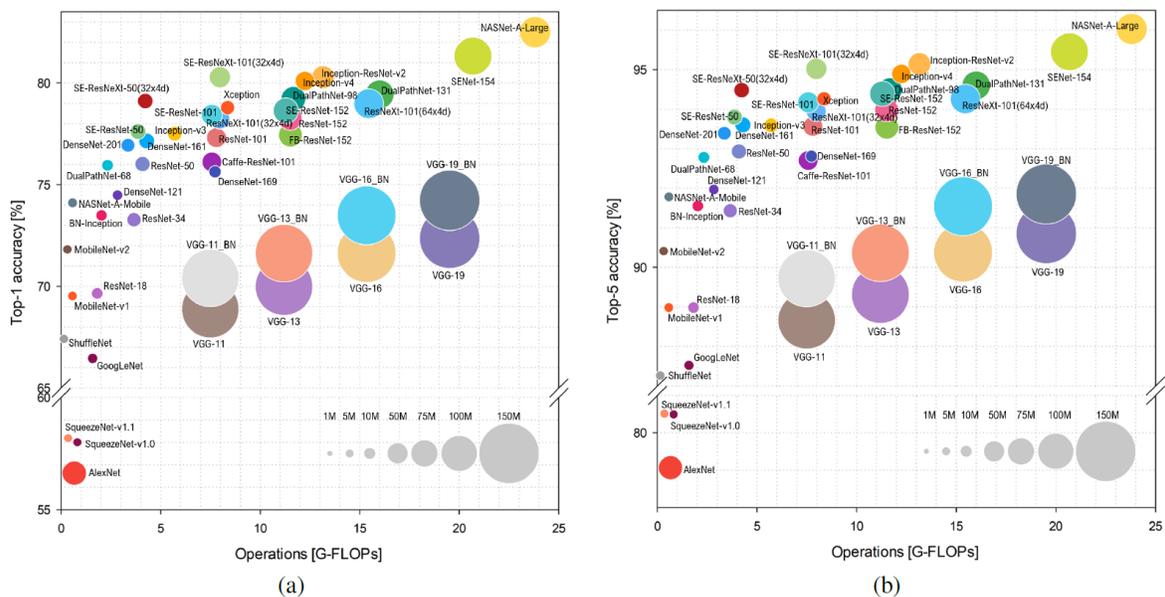
El estudio tuvo en cuenta diferentes arquitecturas y familias de arquitecturas:

- AlexNet

- VGG
- Inception
- GoogLeNet
- SqueezeNet
- ResNet
- DenseNet
- ResNeXt
- Xception
- DualPathNet
- SE-ResNet
- SENet
- SE-ResNet
- SE-ResNeXt
- NASNet-A-Large
- NASNet-A-Mobile

Las métricas de comparación para este estudio correspondieron a índice de precisión, complejidad del modelo, uso de la memoria, complejidad computacional y tiempo empleado en cada modelo.

La siguiente imagen compara la mejor precisión en el lado (a) y la quinta precisión del mismo modelo en el lado (b), con lo cual se obtuvieron los siguientes resultados:



**Figura 7:** Comparativo entre Top-1 y Top-5 de precisión vs. Complejidad computacional de los modelos.

**Fuente:** Benchmark Análisis of Representative Deep Neural Network Architectures.

Los resultados marcan que tanto en la prueba (a) como en la prueba (b) NASNet-A-Large es la arquitectura de reconocimiento de imágenes que más precisión alcanza, sin embargo, la cantidad de G-FLOPs hace de este un modelo computacionalmente complejo, seguido este

de SNet-154, estos a su vez son los modelos que más tiempo de procesamiento emplean como se observa en la siguiente tabla.

DNN	1	2	4	8	16	32	64	DNN	1	2	4	8	16	32	64
AlexNet	1.28	0.70	0.48	0.27	0.18	0.14	0.15	AlexNet	28.88	13.00	8.58	6.56	5.39	4.77	
BN-Inception	5.79	3.00	1.64	1.10	0.87	0.77	0.71	BN-Inception	35.52	26.48	25.10	23.89	21.21	20.47	
CaffeResNet-101	8.20	4.82	3.32	2.54	2.27	2.16	2.08	CaffeResNet-101	84.47	91.37	70.33	63.53	56.38	53.73	
DenseNet-121 (k=32)	8.93	4.41	2.64	1.96	1.64	1.44	1.39	DenseNet-121 (k=32)	66.43	50.87	50.20	43.89	40.41	38.22	
DenseNet-169 (k=32)	13.03	6.72	3.97	2.73	2.14	1.87	1.75	DenseNet-169 (k=32)	137.96	130.27	110.82	100.56	92.97	88.94	
DenseNet-201 (k=32)	17.15	9.25	5.36	3.66	2.84	2.41	2.27	DenseNet-201 (k=32)	84.57	61.71	62.62	53.73	49.28	46.26	
DenseNet-161 (k=48)	15.50	9.10	5.89	4.45	3.66	3.43	3.24	DenseNet-161 (k=48)	103.20	76.11	77.10	68.32	62.73	59.14	
DPN-68	10.68	5.36	3.24	2.47	1.80	1.59	1.52	DPN-68	113.08	52.73	42.85	43.32	38.18	36.40	36.22
DPN-98	22.31	13.84	8.97	6.77	5.59	4.96	4.72	DPN-98	243.51	148.51	135.30	125.92	123.34	118.68	117.27
DPN-131	29.70	18.29	11.96	9.12	7.57	6.72	6.37	DPN-131	330.15	204.69	184.89	172.25	165.59	162.67	160.66
FBResNet-152	14.55	7.79	5.15	4.31	3.96	3.76	3.65	FBResNet-152	133.68	147.75	113.48	105.78	94.26	97.47	
GoogLeNet	4.54	2.44	1.65	1.06	0.86	0.76	0.72	GoogLeNet	32.11	27.19	23.29	21.66	19.77	19.96	
Inception-ResNet-v2	25.94	14.36	8.82	6.43	5.19	4.88	4.59	Inception-ResNet-v2	198.95	141.29	127.97	130.25	117.99	116.47	
Inception-v3	10.10	5.70	3.65	2.54	2.05	1.89	1.80	Inception-v3	79.39	59.04	56.46	51.79	47.60	46.85	
Inception-v4	18.96	10.61	6.53	4.85	4.10	3.77	3.61	Inception-v4	158.00	120.23	106.77	102.01	95.51	95.40	
MobileNet-v1	2.45	0.89	0.68	0.60	0.55	0.53	0.53	MobileNet-v1	15.06	11.94	11.34	11.03	10.82	10.58	10.55
MobileNet-v2	3.34	1.63	0.95	0.78	0.72	0.63	0.61	MobileNet-v2	20.51	14.58	13.67	13.56	13.18	13.10	12.72
NASNet-A-Large	32.30	23.00	19.75	18.49	18.11	17.73	17.77	NASNet-A-Large	437.20	399.99	385.75	383.55	389.67		
NASNet-A-Mobile	22.36	11.44	5.60	2.81	1.61	1.75	1.51	NASNet-A-Mobile	133.87	62.91	33.72	30.62	29.72	28.92	28.55
ResNet-101	8.90	5.16	3.32	2.69	2.42	2.29	2.21	ResNet-101	84.52	77.90	71.23	67.14	58.11		
ResNet-152	14.31	7.36	4.68	3.83	3.50	3.30	3.17	ResNet-152	124.67	113.65	101.41	96.76	82.35		
ResNet-18	1.79	1.01	0.70	0.56	0.51	0.41	0.38	ResNet-18	21.16	15.30	14.35	13.82	11.99	10.73	12.45
ResNet-34	3.11	1.80	1.20	0.96	0.82	0.71	0.67	ResNet-34	39.88	28.82	27.51	24.97	20.41	18.48	17.97
ResNet-50	5.10	2.87	1.99	1.65	1.49	1.37	1.34	ResNet-50	53.09	44.84	41.20	38.79	35.72		
ResNeXt-101 (32x4d)	17.05	9.02	6.27	4.62	3.71	3.25	3.11	ResNeXt-101 (32x4d)	115.37	90.93	84.64	79.66	77.63		
ResNeXt-101 (64x4d)	21.05	15.54	10.39	7.80	6.39	5.62	5.29	ResNeXt-101 (64x4d)	177.40	155.77	144.82	137.43	134.07		
SE-ResNet-101	15.10	9.26	6.17	4.72	4.03	3.62	3.42	SE-ResNet-101	118.13	105.11	96.71	91.68	80.99		
SE-ResNet-152	23.43	13.08	8.74	6.55	5.51	5.06	4.85	SE-ResNet-152	169.73	155.08	139.72	133.59	116.97		
SE-ResNet-50	8.32	5.16	3.36	2.62	2.22	2.01	2.06	SE-ResNet-50	69.65	61.37	55.33	51.87	47.80		
SE-ResNeXt-101 (32x4d)	24.96	13.86	9.16	6.55	5.29	4.53	4.29	SE-ResNeXt-101 (32x4d)	139.62	122.01	112.05	105.34	102.39		
SE-ResNeXt-50 (32x4d)	12.06	7.41	5.12	3.64	2.97	3.01	2.56	SE-ResNeXt-50 (32x4d)	80.08	69.86	67.20	62.66	61.19		
SNet-154	53.80	30.30	19.32	13.27	10.45	9.41	8.91	SNet-154	309.48	240.80	221.84	211.00	207.06	201.49	201.66
ShuffleNet	5.40	2.67	1.37	0.82	0.66	0.59	0.56	ShuffleNet	36.58	22.61	13.80	13.36	12.91	12.66	12.50
SqueezeNet-v1.0	1.53	0.84	0.66	0.59	0.54	0.52	0.53	SqueezeNet-v1.0	17.00	16.47	15.03	13.97	13.25	12.89	12.70
SqueezeNet-v1.1	1.60	0.77	0.44	0.37	0.32	0.31	0.30	SqueezeNet-v1.1	11.05	9.88	8.80	7.90	7.38	7.20	7.04
VGG-11	3.57	4.40	2.89	1.56	1.19	1.10	1.13	VGG-11	106.44	125.84	85.84	60.10	32.56	30.51	32.27
VGG-11_BN	3.49	4.60	2.99	1.71	1.33	1.24	1.27	VGG-11_BN	101.58	122.82	86.26	54.50	47.81	47.31	41.26
VGG-13	3.88	5.03	3.44	2.25	1.83	1.75	1.79	VGG-13	122.59	148.80	108.28	75.99	70.57	64.88	62.79
VGG-13_BN	4.40	5.37	3.71	2.42	2.05	1.97	2.00	VGG-13_BN	129.69	153.68	113.90	81.19	76.39	70.59	67.38
VGG-16	5.17	5.91	4.01	2.84	2.20	2.12	2.15	VGG-16	151.52	169.92	129.89	96.81	91.72		
VGG-16_BN	5.04	5.95	4.27	3.06	2.45	2.36	2.41	VGG-16_BN	163.37	176.35	136.85	103.45	98.11		
VGG-19	5.50	6.26	4.71	3.29	2.59	2.52	2.50	VGG-19	178.04	192.86	152.28	117.92	112.39		
VGG-19_BN	6.17	6.67	4.86	3.56	2.88	2.74	2.76	VGG-19_BN	185.18	198.66	159.20	124.88	119.15		
Xception	6.44	5.35	4.90	4.47	4.41	4.41	4.36	Xception	98.96	93.40	90.49	87.65	86.89		

Tabla 1: Relación entre tiempo y procesamiento de las arquitecturas de reconocimiento de imágenes, al lado izquierdo los resultados con una tarjeta gráfica Titan Xp y al lado derecho una tarjeta gráfica Jetson TX1.

Fuente: Benchmark Analysis of Representative Deep Neural Network Architectures.

La decisión de selección de los modelos a entrenar entonces debe considerar las métricas de rendimiento evaluadas anteriormente, si bien existen modelos con altos niveles de precisión, para la implementación dentro de este proyecto, se deben contemplar los requerimientos del sistema y capacidad computacional disponible dentro del ejercicio. Por tanto un balance entre precisión y rendimiento computacional se convierten en los decisores del modelo a utilizar, garantizando el cumplimiento de los KPI's determinados dentro del proyecto y el tiempo disponible para la realización del mismo.

Adicional a lo anterior es fundamental el entendimiento de los modelos seleccionados, lo cual garantice un correcto uso de las arquitecturas disponibles, con el fin de aplicar esta transferencia de conocimiento previo a la necesidad del negocio.

### 1.2 PREGUNTA DE NEGOCIO

Para la industria de comidas es relevante la opinión de sus clientes para la toma de decisiones informadas en cuando a aspectos de mejora, cambio o robustecimiento de su menú gastronómico. En este contexto tener información actualizada y lo más recientemente posible

es de gran valor tanto para las e, aquí es donde las redes sociales son de vital importancia dada la variedad de información que allí puede circular, esta información es casi que instantánea y mantiene en constante actualización las tendencias de comida en cualquier parte del mundo. Dado lo anterior, contar con un instrumento capaz de reconocer la tendencia de comidas y el impacto de cada producto en los usuarios se convierte en una estrategia de negocio fundamental para crecer o simplemente no sucumbir en la industria.

¿Cómo generar valor a Nutresa con la identificación de los alimentos compartidos en redes sociales a través de imágenes y fotografías?

### 1.3 OBJETIVO DE NEGOCIO

Realizar identificación de alimentos en las imágenes compartidas en redes sociales con el fin de generar valor, conocimiento y entendimiento de las tendencias del mercado para Nutresa.

### 1.4 OBJETIVOS ESPECÍFICOS DEL NEGOCIO

- Identificar y describir las tendencias del contenido digital de alimentos en redes sociales.
- Dar a conocer el análisis de las redes sociales y su uso como activo estratégico de Nutresa.
- Implementar una taxonomía de alimentos, en español, basado en una ontología existente.

### 1.5 MÉTRICAS DE NEGOCIO

Se identificaron las siguientes métricas de negocio asociadas a este proyecto, para las cuales recomendamos identificar las variables que junto con los resultados de este proyecto, afectan la medición en estos dos KPI's.

- ✓ **Riesgo de Reputación (NPS):** la medición de la reputación permite identificar la satisfacción de los clientes de productos de comidas específicos y la posible recomendación o no de los productos asociados a Nutresa.
- ✓ **Eliminación de marcas:** de acuerdo con el contexto y percepción de los productos, y de acuerdo con los registros de ventas e ingresos, se puede determinar si productos específicos se deban eliminar del portafolio de Nutresa.

### 1.6 HARDWARE & SOFTWARE

- ✓ Procesador: Intel(R) Core™ i7 – 5500U CPU@ 2.40GHz.
- ✓ Tarjeta gráfica: 8 GPU de NVIDIA GTX 1080 TI – 11GB de memoria.

- ✓ Memoria RAM: 256 GB, se utilizarán los servidores de la Universidad para el modelamiento.
- ✓ Sistema Operativo: UBUNTO 16.04.7.
- ✓ Software: Tensorflow.2.0, Keras 2.3.1, Python 3.7.11.

## 1.7 REQUERIMIENTOS DEL PROYECTO

- ✓ **Participación efectiva de CAOBA**  
Acompañamiento durante el desarrollo de cada una de las etapas CRISP-DM, especialmente en las etapas de entendimiento del negocio y entendimiento de los datos.
- ✓ **Base de datos**  
Se deberá contar con un conjunto de datos apropiada para el proceso de entrenamiento del modelo de reconocimiento de imágenes aprobado por CAOBA
- ✓ **Entregables**  
Entrega de la caracterización de las variables resultado del proyecto
- ✓ **Alcance del proyecto**  
Enmarcado en las fases establecidas por la metodología CRISP-DM, dentro de las cuales se desarrollarán todas las establecidas a excepción de la última fase que corresponde a la implementación del proyecto.

## 1.8 SUPUESTOS DEL PROYECTO

- ✓ La información se encuentra en formato estructurado.
- ✓ El equipo de CAOBA está en capacidad de transferir el conocimiento requerido al equipo de consultores.
- ✓ El equipo de CAOBA asignará espacios suficientes para el seguimiento y desarrollo del proyecto.
- ✓ El equipo de la Pontificia Universidad Javeriana cuenta con las habilidades requeridas para identificar la problemática de negocio y abordarla con técnicas analíticas.

## 1.9 RESTRICCIONES DEL PROYECTO

- ✓ Tiempo: El proyecto se ejecutará entre el viernes 6 de agosto y el 26 de noviembre de 2021.
- ✓ Información: Para las etapas de entendimiento de datos, preparación de datos y modelado, se tendrá en cuenta información de data sets existentes de imágenes que puedan ser entrenadas.

- ✓ Dada la complejidad que supone la construcción de un modelo de detección o reconocimiento de imágenes y la restricción de tiempo para la ejecución del proyecto, se contempla la ejecución de una primera etapa enfocada al reconocimiento de imágenes con la segmentación realizada de los data set propuestos.
- ✓ Metodología: De acuerdo con el alcance propuesto, no se contempla la aplicación de modelos de reconocimiento de imágenes.

## 1.10 OBJETIVOS DEL DATA MINING

- Realizar el reconocimiento de imágenes y clasificar los diferentes tipos de alimentos en una única fotografía.
- Encontrar un modelo que cumpla la métrica de medición del Data Mining.
- Implementar el entrenamiento del modelo de clasificación de imágenes con la unión de dos o más set de datos de imágenes “tagueadas”, que permita obtener un modelo confiable de reconocimiento.
- Implementar técnicas de reducción de dimensionalidad para la detección de objetos.

## 1.11 MÉTRICAS DE DATA MINING

- ✓ **Precisión:** obtener un modelo con una precisión de clasificación de imágenes superior al 70%.

### 1.11.1 Pipeline del proyecto

El conjunto de eventos establecidos responde al propósito de modificar los retos del aprendizaje del modelamiento de clasificación de imágenes. La figura 8 muestra la representación del pipeline y un subconjunto de eventos dentro del mismo correspondiente a la manipulación de los datos, manipulación del procesamiento y manipulación del algoritmo. Estas tres categorías y las subcategorías dentro de cada una se presentan en la figura 9.



*Figura 8: Pipeline del proyecto.  
Fuente: Elaboración propia.*

- ❖ **Entendimiento del negocio:** Comprender la situación actual de las redes sociales en la industria de alimentos tanto en el contexto mundial y local identificando la importancia de las publicaciones de fotos relacionadas con alimentos.
- ❖ **Construcción del set de datos:** Se utilizarán bases de datos existentes relacionadas con alimentos, las cuales se relacionarán con la ontología FoodOn.
- ❖ **Modelamiento:** Se seleccionarán tres modelos de reconocimiento de imágenes que sean considerados buenos predictores y presenten rendimientos computacionales óptimos dados los requerimientos del sistema.
- ❖ **Evaluación:** De acuerdo con el modelo de mejor rendimiento se evaluarán los resultados con nuevas imágenes y se contrastará con la medida de desempeño seleccionada para la minería de datos.
- ❖ **Modelamiento:** Aplicación de técnicas de aprendizaje automático que permita identificar métricas de evaluación.
- ❖ **Evaluación de resultados:** Contraste entre los resultados obtenidos por parte de los modelos y los esperados en los objetivos de minería de datos.

### 1.11.2 Pipeline analítico

A nivel analítico se ejecutarán las siguientes actividades de acuerdo con el objetivo propuesto:



*Figura 9: Pipeline analítico.  
Fuente: Elaboración propia*

### Preprocesamiento de la data – Construcción del set de datos



*Figura 9: Pipeline de la construcción del set de datos.  
Fuente: Elaboración propia*

- ❖ **Recolección de datos:** para la construcción del set de datos se seleccionaron las bases existentes de imágenes de alimentos *UECFood-256*, *Food-101* y *Food X251*.
- ❖ **Validación y exploración de datos:** Luego de construido el set de datos, se iniciará una validación de las imágenes, con el objetivo de comprender las imágenes que comprenden cada una de las categorías seleccionadas en relación con la ontología empleada.
- ❖ **Limpieza de datos:** Para lograr un entrenamiento adecuado de los modelos, eliminando el sesgo posible que pueda existir en el set de datos, se debe realizar una limpieza manual de las imágenes por categoría, eliminando las imágenes que dificulten el entrenamiento por la existencia de varios objetos no relacionados con la categoría o que simplemente sean ajenas al contexto al que pertenecen.
- ❖ **Selección de la Categorías:** La selección definitiva de las categorías para el entrenamiento se hacen con dos objetivos, el primero es acotar la cantidad de datos usados para el entrenamiento del modelo por las restricciones del sistema. Segundo, se definieron imágenes que fueran más relacionadas a los productos de la compañía Nutresa.
- ❖ **Validación de la cantidad de imágenes:** con el fin de garantizar un correcto modelamiento es necesario validar si las categorías contienen la cantidad de datos suficiente para el ejercicio.

### Preprocesamiento de la data – Preprocesamiento de las imágenes

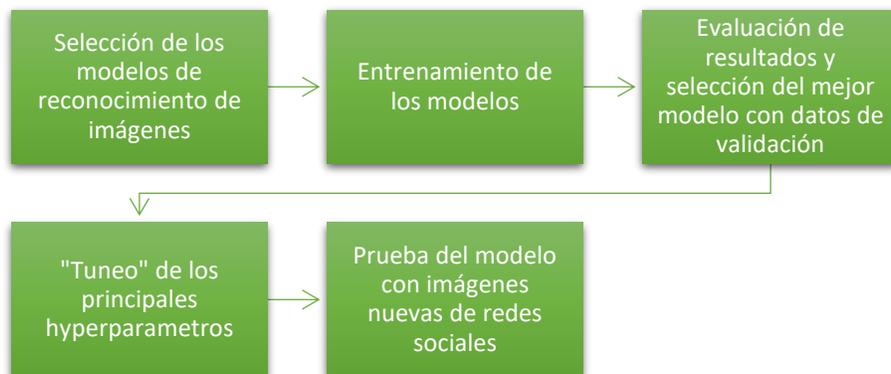


*Figura 10: Pipeline del preprocesamiento de imágenes.*

*Fuente: Elaboración propia*

- ❖ **Creación de conjuntos de entrenamiento y validación:** Se garantizará que las imágenes seleccionadas para el entrenamiento tengan similares porcentajes en la cantidad de imágenes de entrenamiento y validación por categoría a entrenar.
- ❖ **Preprocesamiento de las imágenes de entrenamiento:** Definir la técnica de preprocesamiento a utilizar para el entrenamiento del modelo.
- ❖ **Selección de la técnica de codificación de las imágenes:** Con el objetivo que el modelamiento se logre por la capacidad y requerimientos de sistema, la técnica de codificación permitirá que el código corra y logre el entrenamiento de los modelos para la identificación de las imágenes.

### 1.11.3 Pipeline del Modelamiento y Evaluación



*Figura 11: Pipeline del modelamiento y Evaluación.*

*Fuente: Elaboración propia*

- ❖ **Selección de los modelos de reconocimiento de imágenes:** Basados en la teoría de entrenamiento de imágenes se seleccionan tres modelos con el fin de comparar los resultados y poner a prueba el modelo con los mejores rendimientos.
- ❖ **Entrenamiento de los modelos:** Los modelos seleccionados se entrenarán con las 30 categorías de imágenes seleccionadas.
- ❖ **Evaluación de resultados:** Los modelos entrenados se comparan con los datos de validación de las 30 categorías de alimentos y dada la métrica de selección se definirá el modelo que se pondrá a prueba con imágenes nuevas.
- ❖ **“Tuneo” de los principales hiperparámetros:** Se seleccionarán los hiperparámetros para realizar “tuning” sobre el modelo seleccionado que tuvo mejores resultados en el proceso de validación.
- ❖ **Prueba del modelo con imágenes nuevas:** Con el insumo del equipo de ETL’s encargado de suministrar nuevas imágenes de cuentas de redes sociales, se pondrá a prueba el modelo bajo los parámetros escogidos de mejores rendimientos en la validación, con el fin de garantizar un correcto ejercicio se deben construir las mismas categorías por imágenes del proceso de entrenamiento.

## EXPLORACIÓN DE LOS DATOS

### 2.1 ONTOLOGÍA

La ontología sugerida para el desarrollo del modelo de reconocimiento de imágenes corresponde a la ontología FoodOn, esta ontología contempla un complejo ecosistema de desde la siembra producción y preparación de alimentos. En esta se describen más de 9.600 categorías de alimentos con un vocabulario detallado sin ambigüedades y fáciles de referenciar. (FoodOn, s.f.)

### 2.2 CONSTRUCCIÓN DEL SET DE DATOS

Se cuenta con tres data sets los cuales se combinan para proceder con el proceso de entrenamiento del modelo.

1. Data set de imágenes 1: **UECFood-256 Dataset Ver.1.0**, este data set cuenta con 256 categorías de diferentes clases de alimentos, tiene un porcentaje alto de información de comida japonesa pero también cuenta con varias categorías comunes a nivel mundial (FoodCam, s.f.).
2. Data set de imágenes 2: The **Food-101** Data Set, en este data set se compone de 101 categorías de alimentos con más de 101 mil imágenes para cada clase. (Bossard, s.f.)
3. Data set de imágenes 3: **Food X251** es un set de datos compuesto por 251 categorías con más de 158 mil imágenes recolectadas. (Parneet Kaur, Karan Sikka, Weijun Wang, Serge Belongie, & Ajay Divakaran)

### 2.3 ENTENDIMIENTO DEL SET DE DATOS

Después de realizar el cruce de información con la ontología propuesta FoodOn para el desarrollo del modelo de clasificación de imágenes se encontraron **113** categorías diferentes de alimentos.

- ❖ Media: 872 imágenes por categoría.
- ❖ Max: Chicken\_Wing con 2.238 imágenes.
- ❖ Min: Sausage con 63 imágenes.

Chicken_Wing	2390	Spaghetti_bolognese	1465	boiled_egg	538	fried_shrimp	115
Baby_back_ribs	2059	Ravioli	1462	linguine	533	pork_fillet_cutlet	115
omelet	1902	Ice_cream	1458	meatball	529	pork_loin_cutlet	115
lasagna	1684	hamburger	1456	cannelloni	523	meat_loaf	113

french_toast	1656	nachos	1445	fondue	513	muffin	112
apple_pie	1649	pancake	1443	scrambled_egg	513	roast_chicken	110
Caesar_salad	1626	Falafel	1431	brisket	511	baked_salmon	109
churro	1620	pizza	1425	roast_duck	489	beef_stroganoff	109
grilled_salmon	1607	Fried_calamari	1421	ham_and_eggs	484	macaroni_salad	109
Ceviche	1600	Mussels	1421	chop_suey	475	brownie	108
risotto	1599	Foie_gras	1412	profiterole	467	cream_puff	108
burrito	1597	waffle	1400	souffle	464	roll_bread	107
Dumplings	1577	Cup_cakes	1399	rigatoni	456	bagel	105
Macaroni_and_cheese	1570	Spaghetti_carbonara	1390	penne	450	sweet_and_sour_pork	105
Chocolate_mousse	1569	tacos	1383	flan	439	salmon_meuniere	104
Carrot_cake	1550	hot_dog	1382	poached_egg	434	chicken_nugget	103
paella	1541	shortcake	1379	tortellini	433	scone	103
Grilled_cheese_sandwich	1535	Red_velvet_cake	1378	tagliatelle	419	tortilla	102
Chicken_quesadilla	1529	french_fries	1366	osso_buco	412	french_bread	101
tiramisu	1511	Chocolate_cake	1355	fish_and_chips	405	raisin_bread	101
Cheesecake	1506	steak	1108	pate	362	sausage	63
doughnut	1491	Beignets	1000	chili	251		
Filet_mignon	1487	fried_egg	659	toast	220		
Hummus	1487	rice	620	spaghetti	151		
egg_roll	1481	adobo	612	sirloin_cutlet	140		
Onion_rings	1478	compote	611	potato_salad	128		
Guacamole	1470	couscous	603	fried_fish	121		
Greek_salad	1466	chow_mein	578	croissant	120		
sushi	1466	Eggs_benedict	573	oatmeal	117		
Garlic_bread	1465	enchilada	571	popcorn	116		

**Tabla 2:** Cantidad de imágenes por categoría dentro del set de entrenamiento.

*Fuente:* Elaboración propia

## PREPARACIÓN DE DATOS

### 3.1 LIMPIEZA DE LA DATA

Dentro de la etapa de preparación de la Data de trabajo se realizó una limpieza manual dentro de las 113 categorías, en estas categorías se encontraron imágenes que no correspondían a la categoría específica de ese alimento.

Algunos ejemplos de esas imágenes que podrían hacer ruido al entrenamiento del modelo aparecen a continuación:

Categoría Nachos:

Como se observa en las imágenes, los data sets seleccionados para esta categoría tienen imágenes en las cuales sus denominaciones son relacionadas con la palabra “nacho”, ejemplo de esto son las imágenes encontradas de la película Nacho Libre y Nacho el jugador de fútbol del Real Madrid, estas imágenes no ayudan al modelo en su entrenamiento, por tanto se requiere la inmediata eliminación del set de datos.

Imágenes requeridas para el modelamiento	Imágenes no relacionadas con la categoría
	

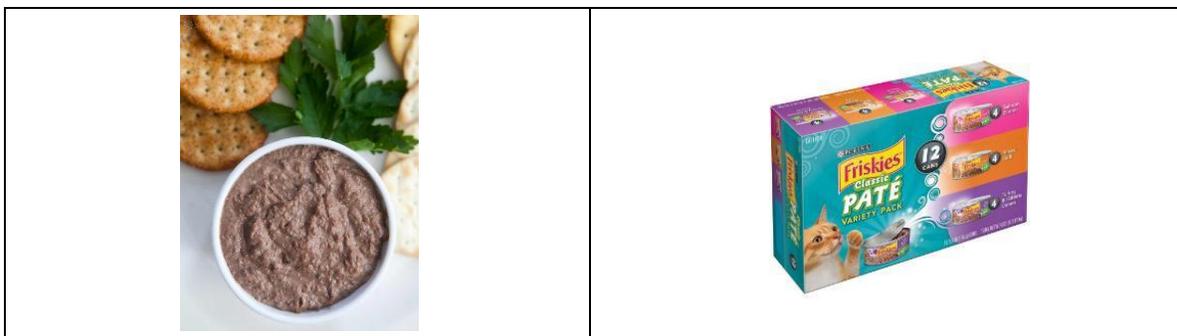
*Tabla 3: Ejemplo de limpieza manual de las categorías de alimentos.*

*Fuente: Elaboración propia*

Categoría Pate:

En esta categoría se encuentran imágenes de comida para gato que nuevamente como en el ejemplo anterior, por la denominación tienen relación con la categoría, no necesariamente corresponden al producto que se desea incluir dentro de las imágenes seleccionadas:

Imágenes requeridas para el modelamiento	Imágenes no relacionadas con la categoría
	



**Tabla 4:** Ejemplo de limpieza manual de las categorías de alimentos.

*Fuente:* Elaboración propia<sup>4</sup>

Esta limpieza se realizó para cada categoría dentro de la integración de los tres (3) data sets seleccionados para el modelo, lo cual implicó un conocimiento del producto y un tiempo para la depuración manual de las imágenes.

### 3.2 SELECCIÓN DE LA DATA DE ENTRENAMIENTO:

Teniendo en cuenta estas 113 categorías iniciales, se depuró para llegar a 30 categorías relacionadas con las marcas más populares de Nutresa, esto con el fin de realizar una segmentación del mercado objetivo inicial. La decisión de seleccionar 30 categorías se debe a que la gran cantidad de imágenes utilizadas en el entrenamiento del modelo ocupaba toda la memoria del servidor y detenía la continuidad del modelamiento.

Las 30 categorías que permitieron al modelo ser entrenado y posteriormente probado, son las siguientes:

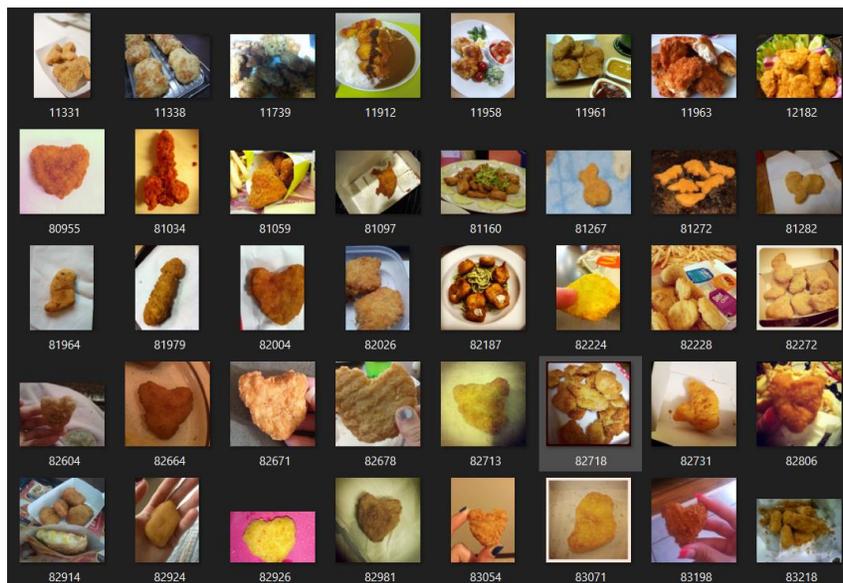
1. Baby_back_ribs	2. brisket	3. burrito	4. Caesar_salad	5. cannelloni
6. chicken_nugget	7. Chicken_Wing	8. fish_and_chips	9. french_fries	10. french_toast
11. Fried_calamari	12. fried_shrimp	13. Greek_salad	14. Grilled_cheese_sandwich	15. hamburger
16. hot_dog	17. Ice_cream	18. lasagna	19. Macaroni_and_cheese	20. meatball
21. penne	22. pizza	23. potato_salad	24. Ravioli	25. roast_chicken
26. sausage	27. spaghetti	28. steak	29. tagliatelle	30. tortellini

**Tabla 5:** Categorías finales para el modelamiento.

*Fuente:* Elaboración propia

El data set para el procesamiento toma las imágenes de cada una de las categorías seleccionadas y las nombra de la siguiente forma “Categoría/Nombre\_de\_la\_imagen”.

Ejemplo de una categoría de alimentos, “chicken\_nugget”:



**Figura 12:** Ejemplo de imágenes dentro del data set de alimentos.  
**Fuente:** Elaboración propia integración de los data set utilizados.

### 3.3 PARTICIÓN DE LA DATA:

Para realizar el entrenamiento y posterior prueba del modelo implementado se definió trabajar con una partición de 70% train y 30% test, para garantizar un correcto entrenamiento del modelo se particiono en estos porcentajes cada una de las categorías seleccionadas, y haciendo un random por categoría se obtuvieron estos porcentajes para cada una de ellas.

N°	Categoría Alimento	Entrenamiento		Prueba	
		Imágenes	Porcentaje	Imágenes	Porcentaje
1	Baby_back_ribs	1434	70%	625	30%
2	brisket	348	68%	163	32%
3	burrito	1129	71%	468	29%
4	Caesar_salad	1125	69%	501	31%
5	cannelloni	388	74%	135	26%
6	chicken_nugget	69	67%	34	33%
7	Chicken_Wing	1748	70%	759	30%
8	fish_and_chips	285	70%	120	30%
9	french_fries	931	68%	435	32%
10	french_toast	1163	70%	493	30%
11	Fried_calamari	987	69%	434	31%
12	fried_shrimp	79	69%	36	31%
13	Greek_salad	1023	70%	443	30%

14	hamburger	1017	70%	439	30%
15	hot_dog	979	71%	403	29%
16	Ice_cream	1028	71%	430	29%
17	lasagna	1171	70%	513	30%
18	linguine	371	70%	162	30%
19	macaroni_salad	84	77%	25	23%
20	meatball	364	69%	165	31%
21	penne	328	73%	122	27%
22	pizza	1006	71%	419	29%
23	potato_salad	81	63%	47	37%
24	Ravioli	1010	69%	452	31%
25	roast_chicken	75	68%	35	32%
26	sausage	84	71%	34	29%
27	spaghetti	99	66%	52	34%
28	steak	757	68%	351	32%
29	tagliatelle	289	69%	130	31%
30	tortellini	286	66%	147	34%
<b>Total general</b>		<b>52437</b>	<b>70%</b>	<b>22323</b>	<b>30%</b>

**Tabla 6:** Partición de las categorías de alimentos en entrenamiento y validación.

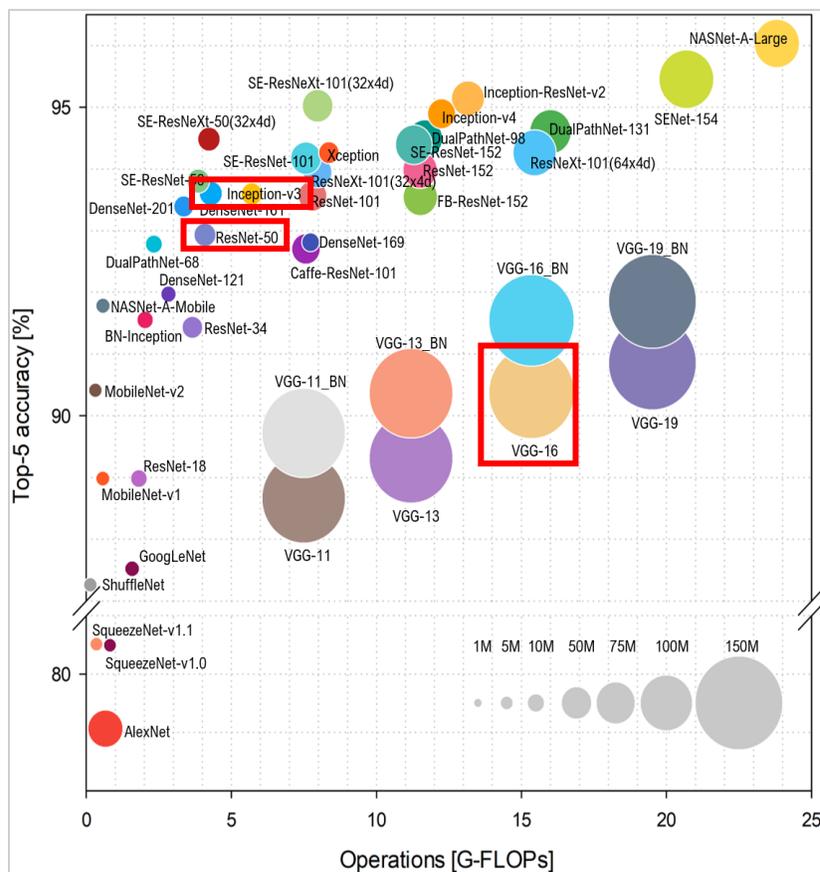
**Fuente:** Elaboración propia

## MODELAMIENTO Y EXPERIMENTACIÓN

### 4.1 MODELAMIENTO

Para encontrar la mejor técnica de modelamiento para el reconocimiento de imágenes, se buscaron diversos modelos existentes con el fin de realizar transfer learning al ejercicio planteado en este proyecto.

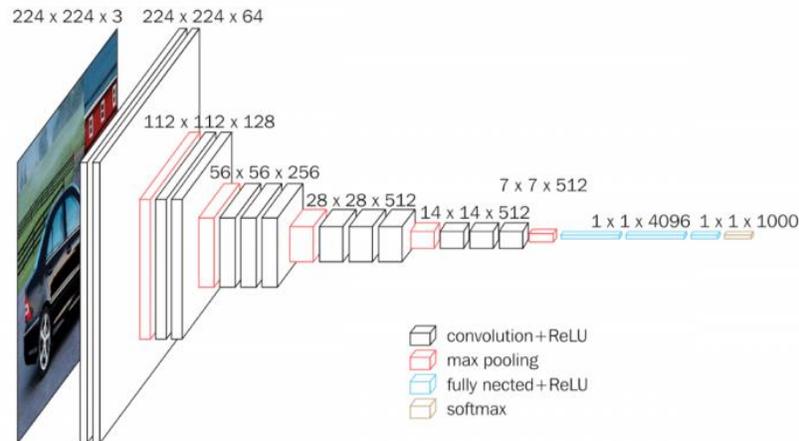
De acuerdo con los resultados de evaluación de modelos de clasificación de imágenes implementados por Bianco, Napoletano y Celona, se seleccionaron 3 modelos que para fines académicos presentan buenos resultados de clasificación y bajos costos computacionales: VGG16, Resnet50 e InceptionV3,



**Figura 13:** Comparativo entre precisión vs. Complejidad computacional de los modelos.  
**Fuente:** Benchmark Analysis of Representative Deep Neural Network Architectures.

- **VGG16** Esta arquitectura de red neuronal ganadora de la competición ILSVR(Imagenet) en 2014 ha sido considerada una de las mejores arquitecturas de reconocimiento de imágenes, sin embargo, es costosa computacionalmente dada la cantidad de parámetros con que cuenta, aproximadamente 138 millones. Sin

embargo, su facilidad de implementación hace de esta arquitectura muy popular y por tanto muy útil en múltiples aplicaciones. (Perumanoor, 2021)



**Figura 14:** Arquitectura VGG13

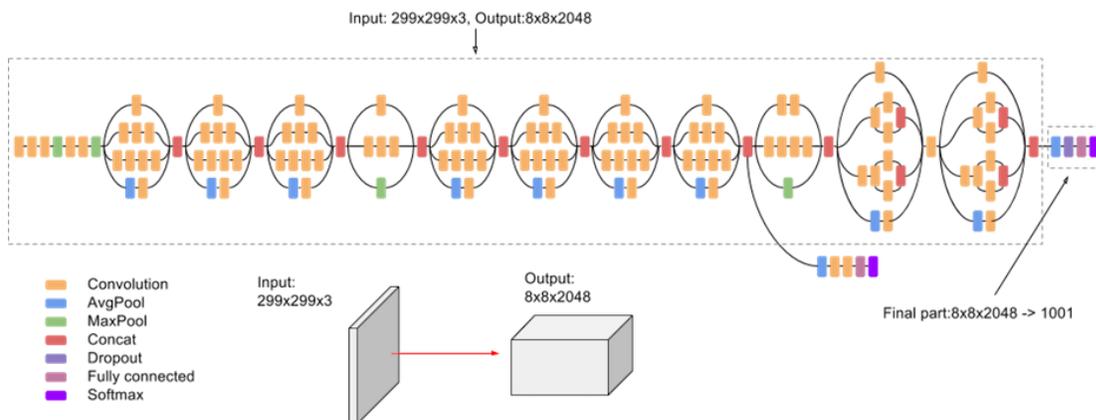
**Fuente:** Informe web- What is VGG16? Introduction to VGG16

- **ResNet50**

Esta arquitectura es una variante del modelo ResNet el cual cuenta con 48 capas de convolución, más una capa de MaxPool y una capa promedio.

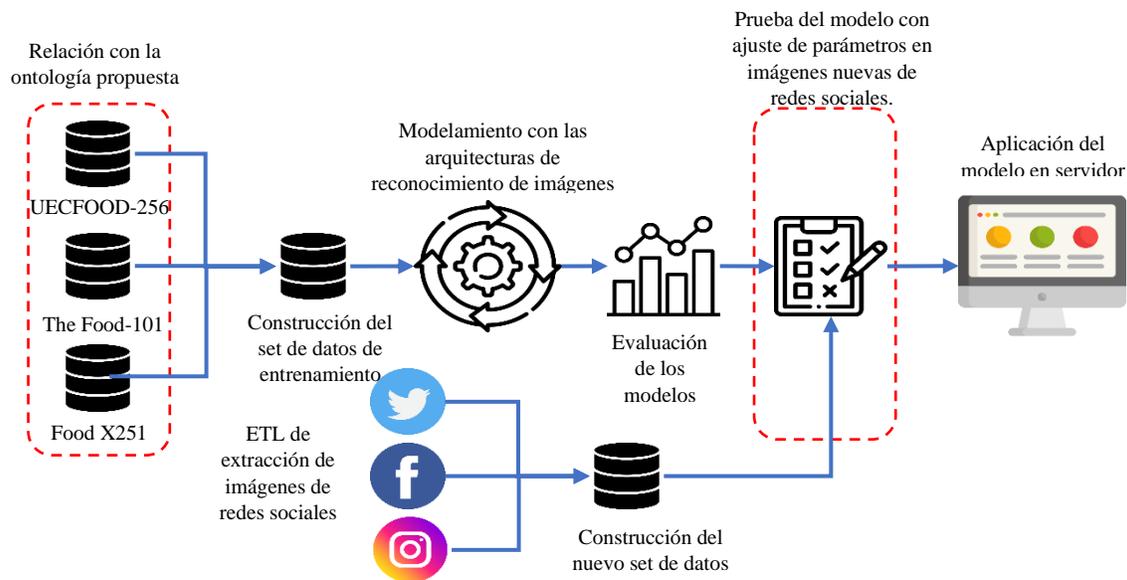
- **InceptionV3**

Red neuronal convolucional factorizadas de que fue pensado originalmente para reducir los costos computacionales mediante la modificación de arquitecturas de Inception anteriores. InceptionV3 ha demostrado ser más eficiente, tanto en términos de parámetros generados por la red, así como también el costo incurrido en memoria y otros recursos computacionales. (Szedy, Vanhoucke, Ioffe, & Shlens, 2015)



**Figura 15:** Arquitectura Inception V3

## 4.2 ARQUITECTURA DEL PROYECTO



**Figura 16:** Arquitectura del proyecto.

**Fuente:** Elaboración propia.

## 4.3 PROTOCOLO DE EXPERIMENTACIÓN

Con el fin de garantizar un correcto protocolo de experimentación, se propone el siguiente modelo de ejecución de pasos que permita llegar a los resultados esperados en los objetivos del proyecto:

1. Se seleccionan los modelos que se usarán para el modelamiento sin la modificación de los hiperparámetros que por defecto están definidos en cada modelo.
2. Ejecución del entrenamiento de los modelos para las categorías de alimentos seleccionadas.
3. Se evaluará los modelos inicialmente con los datos de validación de las 30 categorías seleccionadas.
4. Selección del mejor modelo y se realiza un “tuning” de los hiperparámetros.
5. Se pone a prueba el modelo con imágenes nuevas de redes sociales.

## 4.4 TRATAMIENTO DE LAS IMAGENES

### Librería utilizada:

Para el entrenamiento del modelo se trabajó con la librería de TensorFlow la cual cuenta con la función **ImageDataGenerator** la cual se usa para realizar el procesamiento de las imágenes con las siguientes opciones aplicadas:

#### 4.4.1 Transformación del set de datos de imágenes

- ❖ Rotación de la imagen de 0 a 90 grados:  
rotation\_range = 90, # randomly rotate images in the range (degrees, 0 to 180)



*Figura 17: Ejemplo de rotación de imágenes.  
Fuente: Elaboración propia.*

- ❖ Zoom de la imagen en un 20%:  
zoom\_range = 0.2, # Randomly zoom image



*Figura 18: Ejemplo de zoom de imágenes.  
Fuente: Elaboración propia.*

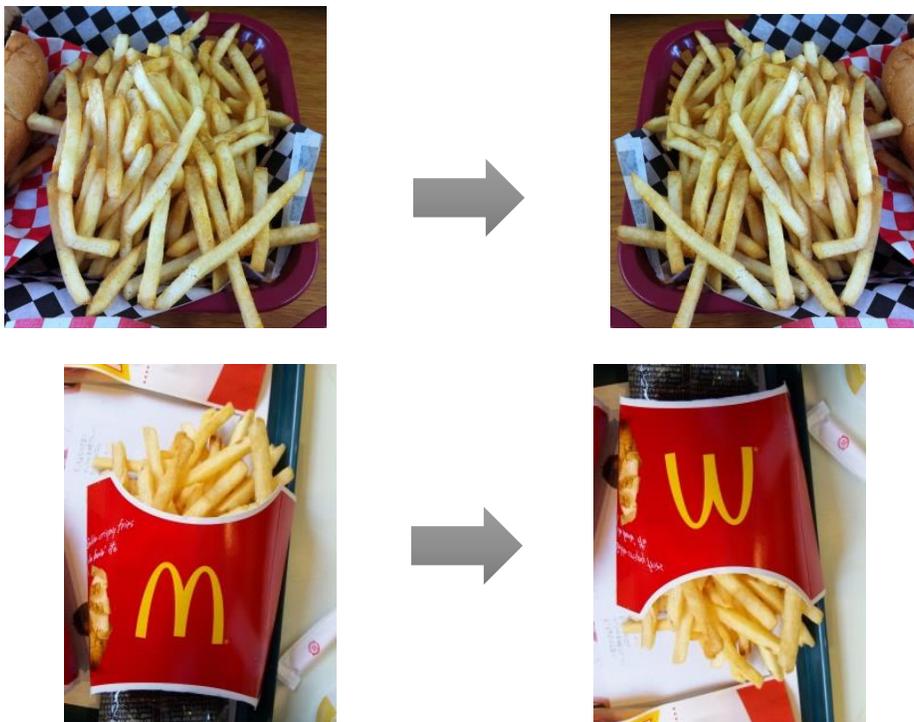
- ❖ Movimiento horizontal y vertical de la imagen en un 10%:  
width\_shift\_range = 0.1, # randomly shift images horizontally (fraction of total width)

height\_shift\_range = 0.1, # randomly shift images vertically (fraction of total height)



**Figura 19:** Ejemplo de desplazamiento vertical de las imágenes.  
**Fuente:** Elaboración propia.

- ❖ Inversión aleatoria de algunas imágenes:  
horizontal\_flip = True, # randomly flip images  
vertical\_flip = True) # randomly flip images



**Figura 20:** Ejemplo de inversión vertical y horizontal de las imágenes.  
**Fuente:** Elaboración propia.

- ❖ Partición del data set de prueba  
validation\_split = 0.2.

Otras posibles transformaciones sobre las imágenes se nombran a continuación, estas transformaciones no fueron realizadas para este entrenamiento:

- ❖ featurewise\_center=False.
- ❖ samplewise\_center=False.
- ❖ featurewise\_std\_normalization=False.
- ❖ samplewise\_std\_normalization=False.
- ❖ zca\_whitening=False.

#### 4.4.2 Dummificación de la variable objetivo:

La variable objetivo corresponde a la categoría de cada alimento, por tanto, para estas 30 categorías de alimentos, la variable irá de 0 a 29.

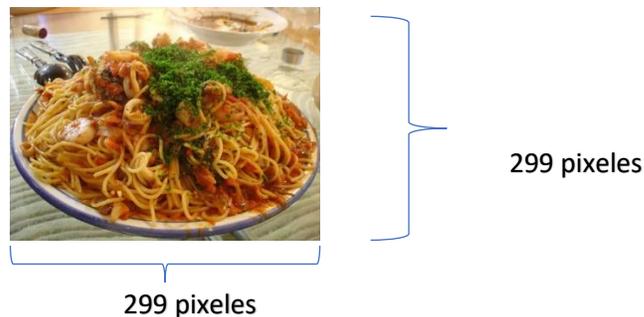
#### 4.4.3 Normalización de las imágenes:

Dentro de la estandarización de las imágenes, se usó el siguiente preprocesamiento de las imágenes, dividir todas las imágenes por 255, restar por 0.5 y multiplicar por 2. Con esto se garantiza que las imágenes cumplan con las mismas condiciones de tamaño.

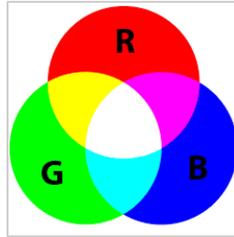
#### 4.4.4 Tratamiento del set de imágenes:

Modelo basado en redes convolucionales CNN para el reconocimiento y detección de imágenes, a continuación, se explican los requerimientos e hiperparámetros utilizados:

- ❖ Tamaño de la imagen: 299 píxeles por 299 píxeles.  
Como parte de un preprocesamiento de las imágenes se realizó un escalamiento a todas las imágenes en un tamaño de 299 píxeles.



- ❖ RGB: 3 capas de color por imagen.

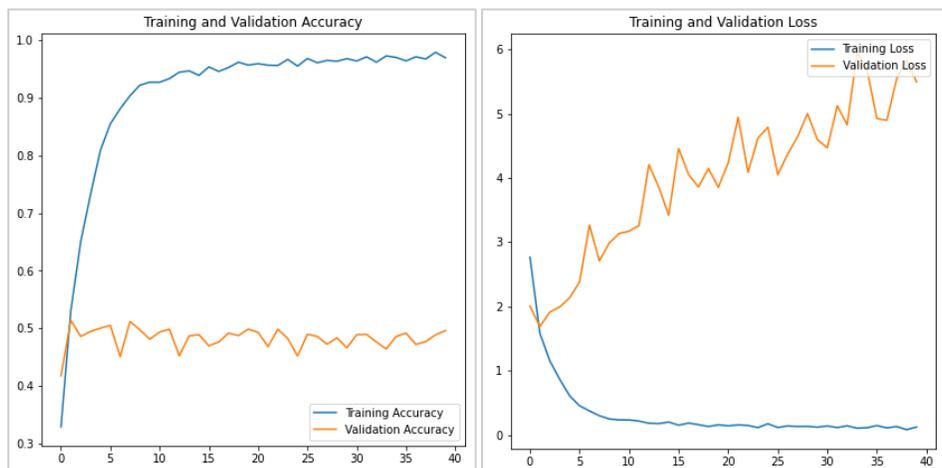


- ❖ Modelo de predicción con n-classes: número de neuronas por cada categoría, en este caso 30 neuronas de salida para todas las categorías seleccionadas.
- ❖ Pérdida del modelo: se define `categorical_crossentropy`.
- ❖ Métrica de medición: Accuracy
- ❖ Número de épocas: 40.

#### 4.5 IMPLEMENTACIÓN DE LOS MODELOS SELECCIONADOS

##### Modelo - VGG16:

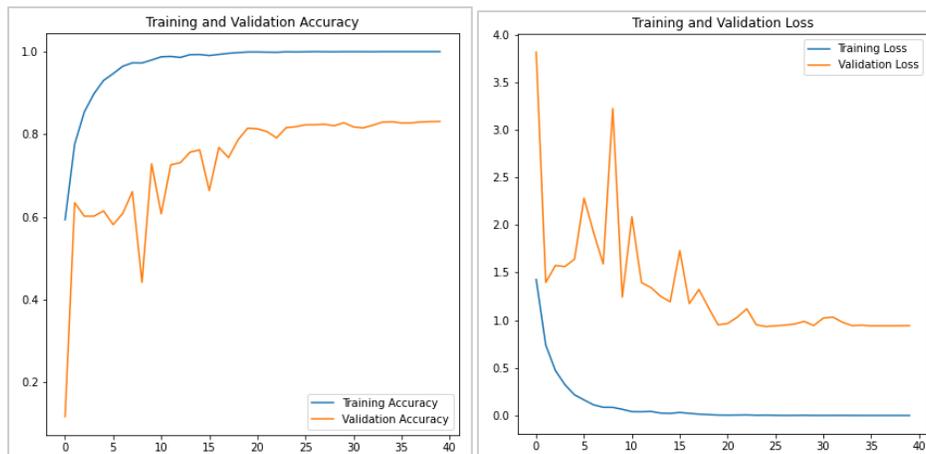
El modelo VGG16 apenas logra resultados promedios de clasificación de todas las imágenes en las categorías seleccionadas, el cual después de la segunda época mantiene una precisión en el reconocimiento de las imágenes por categoría constante sin un aumento en la métrica de medición:



	precision	recall	f1-score	support
Ice_cream	0.49	0.66	0.57	430
brisket	0.46	0.33	0.38	163
Ravioli	0.43	0.35	0.38	452
Chicken_wing	0.59	0.59	0.59	759
lasagna	0.39	0.63	0.48	513
Fried_calamari	0.40	0.68	0.50	434
Caesar_salad	0.70	0.54	0.61	501
meatball	0.31	0.34	0.32	165
hamburger	0.66	0.48	0.56	439
Grilled_cheese_sandwich	0.47	0.52	0.49	447
steak	0.44	0.28	0.34	351
french_fries	0.76	0.67	0.71	435
roast_chicken	0.70	0.20	0.31	35
penne	0.45	0.20	0.28	122
chicken_nugget	0.30	0.18	0.22	34
spaghetti	0.64	0.40	0.49	52
Ribs	0.53	0.43	0.47	625
Macaroni_and_cheese	0.66	0.49	0.56	470
cannelloni	0.33	0.15	0.21	135
hot_dog	0.44	0.62	0.51	403
tagliatelle	0.48	0.36	0.41	130
french_toast	0.39	0.52	0.44	493
pizza	0.76	0.48	0.59	419
sausage	0.19	0.09	0.12	34
burrito	0.38	0.41	0.40	468
potato_salad	0.23	0.11	0.14	47
tortellini	0.37	0.38	0.37	147
Greek_salad	0.52	0.65	0.58	443
fried_shrimp	0.46	0.17	0.24	36
fish_and_chips	0.45	0.14	0.22	120
accuracy			0.50	9302
macro avg	0.48	0.40	0.42	9302
weighted avg	0.52	0.50	0.49	9302

### Modelo - ResNet50:

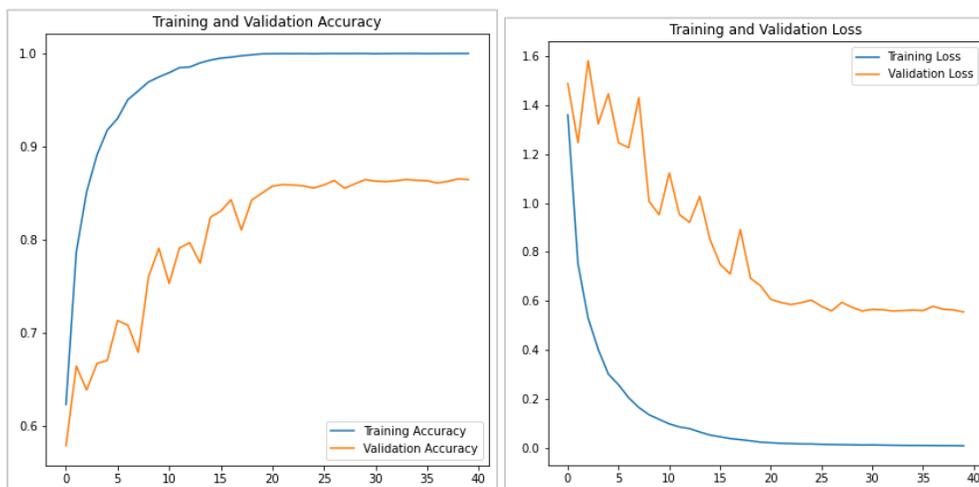
El modelo Resnet50 presenta mejores resultados comparativamente con el modelo VGG16 logrando resultados constantes a partir de la época 20 superando el 80 % de precisión en el reconocimiento de las imágenes:



	precision	recall	f1-score	support
Ice_cream	0.80	0.89	0.84	430
brisket	0.77	0.65	0.70	163
Ravioli	0.79	0.75	0.77	452
Chicken_wing	0.85	0.90	0.88	759
lasagna	0.73	0.79	0.76	513
Fried_calamari	0.89	0.84	0.86	434
Caesar_salad	0.89	0.91	0.90	501
meatball	0.82	0.76	0.79	165
hamburger	0.87	0.83	0.85	439
Grilled_cheese_sandwich	0.87	0.82	0.84	447
steak	0.77	0.74	0.76	351
french_fries	0.85	0.91	0.88	435
roast_chicken	0.96	0.63	0.76	35
penne	0.87	0.70	0.77	122
chicken_nugget	0.81	0.62	0.70	34
spaghetti	0.82	0.77	0.79	52
Ribs	0.81	0.82	0.81	625
Macaroni_and_cheese	0.82	0.87	0.84	470
cannelloni	0.73	0.67	0.69	135
hot_dog	0.83	0.81	0.82	403
tagliatelle	0.74	0.81	0.77	130
french_toast	0.87	0.85	0.86	493
pizza	0.87	0.89	0.88	419
sausage	0.76	0.74	0.75	34
burrito	0.84	0.80	0.82	468
potato_salad	0.76	0.72	0.74	47
tortellini	0.78	0.72	0.75	147
Greek_salad	0.87	0.90	0.88	443
fried_shrimp	0.88	0.83	0.86	36
fish_and_chips	0.89	0.87	0.88	120
accuracy			0.83	9302
macro avg	0.83	0.79	0.81	9302
weighted avg	0.83	0.83	0.83	9302

### Modelo - INCEPTION V3:

El modelo INCEPTION V3 logra superar los resultados del modelo Resnet50, este modelo a diferencia del anterior logra en 15 épocas superar el 80% de precisión en la clasificación de ellos alimentos, logrando su punto máximo en la época 27 con un total de 86% de precisión:



	precision	recall	f1-score	support
Ice_cream	0.85	0.90	0.87	430
brisket	0.82	0.74	0.77	163
Ravioli	0.80	0.80	0.80	452
Chicken_wing	0.90	0.92	0.91	759
lasagna	0.74	0.85	0.79	513
Fried_calamari	0.89	0.91	0.90	434
Caesar_salad	0.90	0.95	0.93	501
meatball	0.90	0.79	0.84	165
hamburger	0.90	0.84	0.87	439
Grilled_cheese_sandwich	0.86	0.89	0.87	447
steak	0.83	0.83	0.83	351
french_fries	0.90	0.93	0.92	435
roast_chicken	0.83	0.69	0.75	35
penne	0.78	0.82	0.80	122
chicken_nugget	0.84	0.76	0.80	34
spaghetti	0.87	0.88	0.88	52
Ribs	0.85	0.85	0.85	625
Macaroni_and_cheese	0.86	0.90	0.88	470
cannelloni	0.76	0.67	0.71	135
hot_dog	0.88	0.87	0.87	403
tagliatelle	0.91	0.82	0.86	130
french_toast	0.90	0.83	0.86	493
pizza	0.90	0.89	0.90	419
sausage	0.83	0.71	0.76	34
burrito	0.89	0.85	0.87	468
potato_salad	0.87	0.70	0.78	47
tortellini	0.85	0.73	0.78	147
Greek_salad	0.91	0.91	0.91	443
fried_shrimp	0.97	0.89	0.93	36
fish_and_chips	0.91	0.86	0.88	120
accuracy			0.86	9302
macro avg	0.86	0.83	0.85	9302
weighted avg	0.87	0.86	0.86	9302

#### 4.6 RESULTADOS DE LOS MODELOS

	VGG16	ResNet50	InceptionV3
<b>Acuracy</b>	0,5	0,83	0,86
<b>Recall</b>	0,4	0,79	0,83
<b>F1-score</b>	0,42	0,8	0,84

**Tabla 7:** Resultados de los modelos entrenados VGG16, ResNet50, InceptionV3.

**Fuente:** Elaboración propia.

Los mejores resultados se obtienen con el modelo Inception V3 en el cual la precisión en la clasificación de las imágenes supera en 3 puntos porcentuales al modelo ResNet50, así también con respecto al recall, la detección de imágenes que si pertenecen a la categoría definida, Inception V3 estos 4 puntos porcentuales por encima del modelo ResNet50, igualmente ocurre con la métrica de F1-Score.

#### 4.7 SELECCIÓN DEL MODELO FINAL

De acuerdo con la métrica de rendimiento definida se selecciona el modelo Inception V3 para el cual se obtuvieron los siguientes resultados por categoría:

La época en la que mejores resultados obtiene el modelo es en la época 33, de acuerdo con la comparación gráfica, con este modelo en la época 33 se logra una precisión por encima del 80 %, adicionalmente en esta época se obtiene la menor pérdida.

En cuanto a los resultados por categoría, se observa en la matriz que los principales clasificadores de cada alimento se encuentran en la diagonal lo que indica que el modelo está categorizando correctamente los alimentos dentro del set de pruebas. Es decir, por ejemplo para la categoría 1, el modelo clasificó 25 veces correctamente el alimento y tuvo 9 errores de clasificación, y para la categoría 2, el modelo clasifico correctamente 369 alimentos.

```

Ice_cream [[386  2  7  2  2  0  2  1  2  0  4  5  0  0  0  0  2  2
1  2  0  3  1  1  4  0  1  0  0  0]
brisket [  1 120  0  3  2  0  0  0  2  1  7  0  1  2  0  0 19  1
0  0  1  1  1  0  1  0  0  0  0  0]
Ravioli [  3  0 360  5 36  6  4  2  1  3  2  1  0  1  0  1  2  5
2  1  0  0  2  0  3  0  8  4  0  0]
Chicken_Wing [  6  0  2 698  5  8  1  1  3  1  3  1  0  0  0  0 16  2
2  1  0  5  2  1  0  0  0  0  0  1]
lasagna [  2  1 13  1 438  1  1  1  1  4  1  0  1  3  0  0  2  8
11  0  0  4 13  0  5  0  1  1  0  0]
Fried_calamari [  4  0  3  6  0 394  3  1  0  0  3  5  0  2  3  0  0  2
0  1  2  0  0  0  2  0  1  2  0  0]
Caesar_salad [  1  0  4  0  1  1 477  0  0  1  0  0  0  0  0  0  0  0
0  1  0  0  0  0  0  1  12  0  1]
meatball [  1  0  2  7  4  0  0 130  1  1  0  0  0  0  0  1  3  4
1  4  0  1  2  0  2  0  1  0  0  0]
hamburger [  5  0  0  1  4  2  2  0 369 13  9  3  0  1  0  0  5  3
0 12  0  1  3  0  3  0  0  3  0  0]
Grilled_cheese_sandwich [  1  1  2  1  2  0  2  0 5 398  0  3  0  0  0  0  3  2
0 10  0 14  0  0  1  0  0  2  0  0]
steak [  2  2  4  7  3  1  1  0  5  2 290  1  1  0  0  1  22  0
0  1  0  2  2  0  3  0  0  1  0  0]
french_fries [  3  0  1  1  1  1  0  0  3  1  1 404  0  8  0  0  1  2
0  2  1  1  1  0  0  0  0  1  0  2]
roast_chicken [  0  0  1  6  0  0  0  0  0  1  2  0  24  0  0  0  0  0
0  0  0  1  0  0  0  0  0  0  0]
penne [  0  0  0  1  2  0  1  1  0  0  0  3  0 100  0  1  0  5
1  1  2  0  1  0  0  2  0  0  1]
chicken_nugget [  0  0  0  2  0  4  0  0  0  0  1  0  0  0  26  0  0  0
0  0  0  0  0  0  1  0  0  0  0]
spaghetti [  0  0  0  0  1  0  0  0  0  0  0  0  0  0  46  0  3
0  0  1  0  0  0  1  0  0  0  0]
Ribs [  7 16  4 19  5  3  3  4  4  2 15  1  0  0  0  0 529  2
0  0  0  3  0  0  4  0  0  2  0  2]
Macaroni_and_cheese [  4  0  3  1 13  2  2  0  2  2  1  5  0  2  0  1  1 421
1  2  0  1  2  0  1  1  2  0  0  0]
cannelloni [  0  0  4  0 25  0  2  0  0  0  0  0  0  0  0  0  3  2
91  1  2  1  1  0  3  0  0  0  0]
hot_dog [  3  1  4  0  1  3  1  0  8  3  2  7  0  0  1  0  1  2
0 349  0  2  4  2  9  0  0  0  0]
tagliatelle [  2  1  2  0  1  0  1  0  1  0  0  1  4  0  2  0  1
0  2 107  1  0  0  1  1  2  0  0]
french_toast [  5  1  3  9  9  3  2  0  0 25  2  4  1  0  0  0  7  3
4  1  0 409  2  0  2  0  0  0  1]
pizza [  4  0 12  1 17  1  2  0  0  1  0  1  0  0  0  0  2  2
1  0  0 0 374  0  0  0  1  0  0]
sausage [  0  0  0  0  0  0  2  0  0  1  1  0  0  0  1  0  0  0
0  2  0  2  0 24  0  0  1  0  0]
burrito [ 11  1  8  3 13  3  2  0  1  3  2  0  0  3  0  0  2  4
4  2  0  2  3  0 398  0  2  0  1]
potato_salad [  2  0  2  0  1  1  1  0  0  0  2  0  0  0  0  0  0  3
0  0  0  0  0  0  33  0  1  1  0]
tortellini [  0  0  6  2  4  6  0  2  0  0  0  0  2  0  0  0  8
0  0  2  0  0  0  1 107  6  0  1]
Greek_salad [  1  1  2  1  3  1 18  0  2  1  0  0  0  1  0  0  0  0
1  0  0  0  1  1  3  0  1 405  0  0]
fried_shrimp [  1  0  0  0  0  1  0  1  0  0  0  0  0  0  0  0  1  0
0  0  0  0  0  0  0  0  32  0]
fish_and_chips [  1  0  0  2  1  1  0  0  1  0  0  4  0  0  0  0  1  2
0  0  0  2  0  0  2  0  0  0  103]]

```

**Figura 24:** Predicción por las 30 categorías del modelo en cada uno de los tipos de alimentos seleccionados.

**Fuente:** Elaboración propia.

#### 4.8 EVALUACIÓN

Después de decidir el modelo con el cual se pondrá en marcha un servidor web para CAOBA, Anexo 1, este modelo será objeto de ajuste de los hiperparámetros con el fin de determinar la mejor configuración del modelo en el reconocimiento de imágenes que presente la mayor precisión con imágenes nuevas provenientes de Redes Sociales.

#### 4.8.1 VALIDACIÓN DE RESULTADOS:

Dados los buenos resultados obtenidos en la evaluación del modelo, se pondrá a prueba con las imágenes que circulan en redes sociales, para esto se seleccionaron cuentas de Twitter y de Instagram relacionadas con alimentos y se implementará el modelo anteriormente entrenado.

Se garantizará que el reconocimiento de las imágenes provenientes de las cuentas de redes sociales se realice de acuerdo con las categorías seleccionadas para el entrenamiento, con el fin de testear el modelo sobre imágenes nuevas. Se seleccionaron 14 cuentas de Twitter inicialmente y se propondrá incluir imágenes descargadas de Instagram para complementar el mercado de imágenes que circulan a través de redes sociales.

Se espera que el modelo, dado los resultados obtenidos, logre una buena clasificación de las imágenes y ayuden a comprender y generar conocimiento asociado a los consumidores, relacionado con las imágenes compartidas en estas redes.

#### 4.8.2 ETL DE ALIMENTOS EN REDES SOCIALES

En colaboración con el equipo del proyecto CAOBA de la Universidad de la Universidad Javeriana, Equipo de ETL's, se obtuvo los sets de datos con imágenes nuevas para poner a prueba los resultados del modelo seleccionado. El equipo de ETL realizó la extracción de más de 50 mil imágenes de Twitter e Instagram, redes sociales con alta frecuencia de imágenes relacionadas con alimentos, con el fin de garantizar el correcto uso de las imágenes, se seleccionaron aquellas que correspondían a las 30 categorías seleccionadas para el entrenamiento del modelo, esta selección trae consigo una larga búsqueda manual de las imágenes, dado que la búsqueda del nuevo set de datos para la prueba, realiza el procedimiento de extracción de todas las imágenes publicadas por las cuentas de usuarios seleccionadas.

Para la selección de las cuentas, se tuvo en cuenta aquellas de más popularidad en Colombia y que estaban en cierta medida relacionadas con las 30 categorías determinadas para el modelamiento. Las cuentas de Twitter e Instagram correspondientes fueron las siguientes:

##### Twitter:

1. IFOOD [App Domicilios](#)
2. Delicious Martha [Bloguera](#)

3. Laura Lopez [Bloguera](#)
4. Quenomeladenconqueso [Blog](#)
5. Archies [Restaurante](#)
6. Tulio Recomienda [Blog](#)
7. PPC Colombia [Restaurante](#)
8. Presto [Restaurante](#)
9. Teriyaki Sushi Bar Bogotá [Restaurante](#)
10. Al Agua Patos [Restaurante](#)
11. Santa Costilla [Restaurante](#)
12. La Taquería [Restaurante](#)
13. El Gourmet [TV Chanel](#)
14. McDonalds [Restaurante](#)

Instagram:

1. dogger1996
2. elpuestotradicionurbana
3. labiferia
4. elcarnaldisfrutadetodo
5. burritosandco
6. burritoscompany.co
7. ensaladasdiferentes
8. ensaladas\_saludables.seibo
9. kfc.colombia
10. calimioco
11. buffalowingscol
12. wingznbeer
13. paisawings
14. greeksownsalad
15. greeksaladmemes
16. greeksaladfestival
17. greeksalad.2016
18. emiliagracerestaurante
19. ensalada\_de\_papas
20. companiadelsabororiginal
21. zenuoficial
22. pietran\_oficial
23. rancheraoficial

#### 4.8.3 GRIDSEARCH MANUAL

Dado que la función DataImageGenerator del modelo Inception V3 permite ejecutar el ajuste automático de parámetros sin embargo por la complejidad del problema el cual envuelve miles de parámetros y gran cantidad de datos hace de este sea un procedimiento difícil de

ejecutar, por tanto se procede a realizar un modelamiento de con un ajuste manual de cuatro parámetros presentes del modelo:

- ❖ Fine Tune
- ❖ Optimizador
- ❖ Learning Rate
- ❖ Drop Out

Fine Tune	Optimizador	Learning Rate	Drop Out	Accuracy Validación	Accuracy Prueba
0 capas entrenables	SGD	0,001	0,4	0,78	0,55
10 capas no entrenables	SGD	0,001	0,4	0,86	0,72
20 capas no entrenables	SGD	0,001	0,4	0,87	0,73
30 capas no entrenables	SGD	0,001	0,4	0,86	0,68
20 capas no entrenables	Adam	0,001	0,4	0,78	0,62
20 capas no entrenables	RMSprop	0,001	0,4	0,73	0,57
20 capas no entrenables	Adamax	0,001	0,4	0,83	0,65
20 capas no entrenables	SGD	0,01	0,4	0,84	0,68
20 capas no entrenables	SGD	0,001	0,4	0,87	0,69
20 capas no entrenables	SGD	0,0001	0,4	0,86	0,71
20 capas no entrenables	SGD	0,00001	0,4	0,84	0,68
20 capas no entrenables	SGD	0,0001	0,3	0,86	0,71
20 capas no entrenables	SGD	0,0001	0,5	0,86	0,72

**Tabla 8:** Resultados con el "Tuneo" de los hyperparametros del modelo Inception V3.

**Fuente:** Elaboración propia.

Después de la modificación de los parámetros seleccionados y comparando los resultados tanto en imágenes de validación como en imágenes nuevas para la prueba de los modelos, se obtuvo la mayor precisión con imágenes nuevas de redes sociales de 73%, los parámetros utilizados para llegar a este resultado serán usados para el diseño del servidor flask, (Flask, s.f.) el cual se encuentra en el Anexo 1, en donde se implementara el modelo en una pequeña aplicación en la cual el usuario pueda subir una imagen dentro de las categorías entrenadas del modelo y este le informe en probabilidades a que categoría puede pertenecer. Este

constituye el entregable del proyecto para la compañía Nutresa y su equipo de trabajo en CAOBA, el cual se espera se integre al proyecto de entendimiento de las tendencias de mercados e interpretación de contextos de las imágenes compartidas por los usuarios de redes sociales vinculando sus reacciones y sentimientos respecto a un tema de análisis definido.

Como se muestra en la figura 24, algunas categorías son más difíciles de identificar que otras, es el caso de los tagliatelles con precisión del 46 % que por su estructura, el modelo los confunde con los spaghetti, situación similar ocurre entre otras categorías dentro del modelo.

	precision	recall	f1-score	support
Caesar_salad	0.31	0.76	0.44	21
Chicken_Wing	0.42	0.90	0.58	21
Fried_calamari	0.42	0.79	0.55	24
Greek_salad	0.74	1.00	0.85	20
Grilled_cheese_sandwich	0.76	0.74	0.75	50
Ice_cream	0.67	0.76	0.71	50
Macaroni_and_cheese	0.47	0.94	0.63	18
Ravioli	0.56	0.56	0.56	34
Ribs	0.79	0.70	0.74	79
brisket	0.84	0.98	0.91	50
burrito	0.55	0.82	0.65	22
cannelloni	0.74	0.61	0.67	23
chicken_nugget	1.00	0.30	0.47	23
fish_and_chips	0.90	0.90	0.90	21
french_fries	0.52	0.85	0.64	20
french_toast	0.52	0.44	0.48	25
fried_shrimp	0.67	0.08	0.14	25
hamburger	0.95	0.88	0.91	264
hot_dog	0.50	0.92	0.65	50
lasagna	0.39	0.47	0.43	19
meatball	0.90	0.82	0.86	22
penne	1.00	0.38	0.55	21
pizza	0.94	0.82	0.88	139
potato_salad	0.78	0.37	0.50	49
roast_chicken	1.00	0.30	0.46	20
sausage	0.92	0.43	0.59	51
spaghetti	0.78	0.76	0.77	123
steak	0.69	0.72	0.71	50
tagliatelle	0.46	0.21	0.29	28
tortellini	0.62	0.90	0.73	20
accuracy			0.73	1382
macro avg	0.69	0.67	0.63	1382
weighted avg	0.77	0.73	0.72	1382

*Figura 25: Resultados finales del modelo por categoría*

## CONCLUSIONES

- En este proyecto, hacemos una introducción a un protocolo de implementación de modelamiento de arquitecturas de reconocimiento de imágenes, por tanto se entrega a Nutresa una implementación web del modelo evaluado siendo en este ejercicio Inception V3 el modelo con los mejores resultados en las métricas de medición Accuracy, Recall y F1-Score. Sin embargo esto no limita el hecho de poder usar modelos con requerimiento computacionales más robustos las cuales permitan entrenar un número mayor de categorías de alimentos.
- Debido a las limitaciones computacionales se seleccionaron algunas categorías de alimentos y comidas provenientes del contexto colombiano o sus alrededores. Por lo tanto, para robustecer el modelo y garantizar la clasificación adecuada de los alimentos y comidas, se propone alimentar las bases de datos de entrenamiento con nuevas clases de comidas atípicas (Comida Hindú, griega o comidas fuera de contexto evaluado).
- A partir de modelo implementado es posible identificar los alimentos en las imágenes compartidas en redes sociales con una exactitud de 73% basados en la taxonomía de alimentos. Lo cual permite entender el contexto y generar acciones de negocio
- El uso de las redes sociales para la extracción de información, en este caso para la industria de alimentos, requiere un minucioso ejercicio debido a la cantidad de variedad de imágenes compartidas a través de las plataformas por los usuarios, por tanto, poder extraer la información de la categoría del alimento correctamente es el desafío a seguir por la cantidad de imágenes que circulan en las redes sociales.

## RECOMEDACIONES

- El entrenamiento de una CNN para el reconocimiento de imágenes depende de una adecuada categorización de los alimentos y así mismo contar con imágenes que reduzcan el sesgo en el proceso de entrenamiento, además existen imágenes que contienen múltiples alimentos, esto podría generar ruido al modelo al momento de realizar el entrenamiento.
- Dentro del proceso de prueba de los modelos se requieren procesos de ETL para la extracción de imágenes de redes sociales acorde con el ejercicio del modelamiento, extraer imágenes de cuentas de usuarios de las múltiples redes sociales puede aumentar el sesgo de reconocimiento de imágenes, dado que las imágenes no son propiamente de un alimento en particular, en la mayoría de los casos las imágenes

posteadas por los usuarios en muchos no traen una imagen de primer plano del alimento. La recomendación en este sentido es trabajar con un proceso de ETL de extracción de imágenes de redes sociales basado en hashtags (#), esto reduciría la varianza de las imágenes y el ruido existente que pueda presentarse en las mismas para el proceso de reconocimiento de los alimentos entrenados por el modelo. El hashtag es utilizado en redes sociales para hablar de un tema específico o relacionar la información posteada con este tema, lo cual sería de mucha utilidad si se busca una categoría de alimento que haga parte del objetivo de negocio de la compañía Nutresa.

- Las métricas de medición de los objetivos de negocio propuestas en este proyecto requieren un procedimiento de aplicación de metodologías para estas mediciones. Primero, el Riesgo de Reputación requiere un ejercicio de aplicación de la metodología del NPS lo cual implica la necesidad de poner en práctica ciertas preguntas de análisis de los usuarios relacionadas con la recomendación de las marcas ofertadas por la compañía. Segundo, la eliminación de marcas debe estar apalancada con cifras de ventas, ingresos y costos de producción de los productos analizados para la compañía Nutresa, con el fin de determinar si unido con las opiniones y comentarios encontrados en redes sociales respecto a algún alimento en específico, requiere su eliminación dentro del portafolio de productos.
- La gran cantidad de productos alimenticios dentro del portafolio de Nutresa implica grandes costos computacionales al momento de entrenar un volumen alto de imágenes y categorías de alimentos, para esto se requerirán capacidades de procesamiento, se debe tener en cuenta estos requerimientos de sistema una vez se desea realizar la implementación del modelo.
- Se puede implementar modelos más complejos dependiendo la capacidad de cómputo con que se cuente para la implementación del proyecto, se recomienda seguir la literatura de análisis de arquitecturas hecha por los autores Bianco, Napoletano y Celona respecto al comparativo entre diferentes arquitecturas de reconocimiento de imágenes, con el fin de encontrar modelos de porcentaje de precisión en el reconocimiento de imágenes, mayores a los probados en este proyecto.

## BIBLIOGRAFÍA

- Bossard, L. (s.f.). *Data Vision*. Obtenido de [https://data.vision.ee.ethz.ch/cvl/datasets\\_extra/food-101/](https://data.vision.ee.ethz.ch/cvl/datasets_extra/food-101/)
- Coss, Y. (s.f.). *Digimind*. Obtenido de <https://blog.digimind.com/es/tendencias/30-cifras-a-retener-sobre-el-universo-food-en-las-redes-sociales>
- Dooley, D., Griffiths, E., Gosal, G., Buttigieg, P., Hoehndorf, R., Lange, M., . . . Hsiao, W. (2018). FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. *Science of Food*.
- Flask. (s.f.). *Flask web development*. Obtenido de <https://flask.palletsprojects.com/en/2.0.x/>
- FoodCam. (s.f.). *Foodcam*. Obtenido de <http://foodcam.mobi/dataset256.html>
- FoodOn. (s.f.). *FoodOn*. Obtenido de <https://foodon.org/>
- Foundry, T. O. (s.f.). *The OBO Foundry*. Obtenido de <http://www.obofoundry.org/>
- García, B. (s.f.). *El Correo Web*. Obtenido de <https://elcorreoweb.es/extra/que-es-el-movimiento-foodie-CY5652532>
- LeCun, Y., Bottou, L., Haffner, P., & Bengio, Y. (25 de Julio de 2020). *Toward Data Science*. Obtenido de <https://towardsdatascience.com/understanding-and-implementing-lenet-5-cnn-architecture-deep-learning-a2d531ebc342>
- Martínez Llamas, J., & De Mingo, L. F. (s.f.). *Reconocimiento de Imágenes Mediante Redes Neuronales Convolucionales*.
- Medina, B. (29 de Junio de 2019). *GitHub*. Obtenido de <https://bryanmed.github.io/conv2d/>
- Murillo, D. (s.f.). 'Foodies', influenciadores que marcan tendencia en Barranquilla. *El Heraldo*.
- Nelson, D. (s.f.). *Stack Abuse*. Obtenido de <https://stackabuse.com/image-recognition-in-python-with-tensorflow-and-keras/>
- Parneet Kaur, Karan Sikka, Weijun Wang, Serge Belongie, & Ajay Divakaran. (s.f.). FoodX-251: A Dataset for Fine-grained Food Classification. 5. Obtenido de <https://deepai.org/publication/foodx-251-a-dataset-for-fine-grained-food-classification>
- Perumanoor, T. J. (23 de Septiembre de 2021). *Medium*. Obtenido de <https://medium.com/@mygreatlearning/what-is-vgg16-introduction-to-vgg16-f2d63849f615>
- Szedy, C., Vanhoucke, V., Ioffe, S., & Shlens, J. (2 de Diciembre de 2015). *Paper With Code*. Obtenido de <https://paperswithcode.com/method/inception-v3>
- TOMÀS, A. (2013). 'Foodies', pasión por la comida. *Diario La Vanguardia*.

## Anexo 1

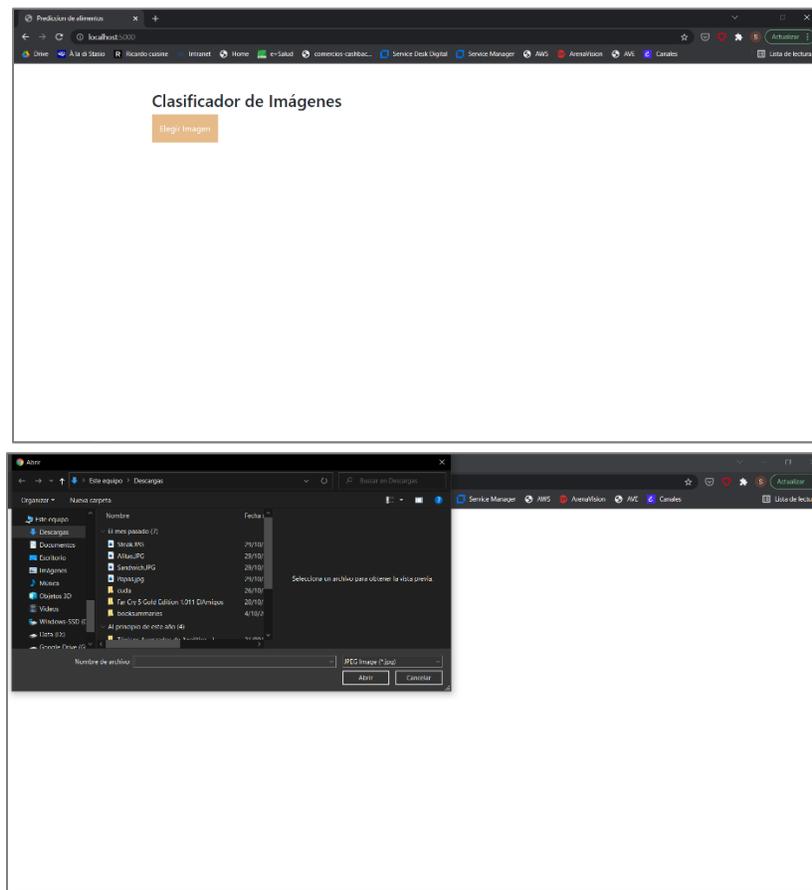
### IMPLEMENTACIÓN DEL MODELO

Como entregable a CAOBA y Nutresa, se ha desarrollado una implementación Web usando Flask en Python.

Flask es un marco web de Python pequeño y ligero que proporciona herramientas y funciones útiles que hacen que crear aplicaciones web en Python sea más fácil. Ofrece a los desarrolladores flexibilidad y un marco más accesible para los nuevos desarrolladores ya que puede crear una aplicación web rápidamente usando únicamente un archivo Python. Flask también es extensible y no fuerza una estructura de directorio concreta ni requiere código estándar complicado antes de iniciarse.

A continuación se presenta el procedimiento de uso del servidor implementado:

1. Seleccionamos la imagen a cargar dentro del servidor, el cual se entrega al equipo de CAOBA implementado en servidor local:



2. Cargue de la imagen seleccionada, antes de ejecutar la predicción confirmamos que la fotografía quede cargada correctamente. A su vez que tenga relación con las categorías usadas para el entrenamiento:



3. Se ejecuta la predicción y se observan los resultados obtenidos en la clasificación de las imágenes nuevas:



Esta implementación utiliza los resultados de la matriz de confusión relacionada con las 30 categorías entrenadas del modelo de reconocimiento de imágenes y define en probabilidades, a cuál categoría corresponde el alimento a identificar. Esta implementación trae las 3 probabilidades mas altas a las que podría pertenecer el alimento.