



GRADO EN CIENCIA DE DATOS

Aplicación de la Ciencia de Datos en compras online de productos alimenticios de proximidad

Presentado por
Julia Monfort Peña

Dirigido por
Joaquín Santos
Victor Yeste

CURSO ACADÉMICO 2023-2024

Agradecimientos

A mis padres y hermanos.

A mis tutores y a todos los involucrados en este proyecto.

Resumen

Este Trabajo de Fin de Grado (TFG) se fundamentará en las experiencias curriculares del proyecto ‘El mercat a un bot(ó)’¹ de Meraki Projectes.

La problemática de los mercados locales y los productos frescos radica en varios aspectos clave. Dado que los precios en los mercados fluctúan diariamente, se ha determinado que la actualización diaria de estos precios requiere de conocimientos en ciencia de datos, como el uso de *computer vision* para obtener, a través de imágenes o videos, el nombre de cada producto junto con su precio.

Los mercados locales enfrentan una competencia feroz con las cadenas de supermercados. Además, la industria alimentaria contribuye a los gases de efecto invernadero especialmente en el transporte de productos.

También se busca obtener métricas y patrones de compras online, para en el futuro implementar un sistema de recomendación de productos, locales, etc. Este TFG busca integrar tecnologías como *computer vision* para abordar los desafíos ambientales de la industria alimentaria mientras promueve la sostenibilidad y la economía local.

Este TFG comienza con la introducción, en la que se comentarán las motivaciones y los objetivos del mismo. En la siguiente sección se desarrolla el estado del arte: exponiendo el reconocimiento de texto a mano, las herramientas y aplicaciones consideradas para el desarrollo de la aplicación, los códigos QR, las aplicaciones de la visión por computadora en productos frescos y finalmente, se discuten las herramientas utilizadas para el análisis y visualización de datos.

En la metodología se describen los pasos llevados a cabo durante el proyecto, desde el desarrollo de la aplicación hasta el análisis de datos, junto con la obtención de los videos y los cambios realizados.

Finalmente se encuentran la discusión de los resultados y las conclusiones del

¹<https://merakiprojectes.eu/el-mercat-a-un-boto/>

proyecto, comentando el alcance de los objetivos y las mejoras a futuro.

Palabras clave: Visión por computadora, escritura a mano, visualización, patrones, análisis.

Abstract

This Final Degree Project will be based on the curricular experiences in the project ‘El mercat a un bot(ó)’² by Meraki Projectes.

The issues faced by local markets and fresh products lie in several key aspects. Given that market prices fluctuate daily, it has been determined that updating these prices daily requires knowledge in data science, such as the use of computer vision to obtain, through images or videos, the name of each product along with its price.

Local markets face fierce competition from supermarket chains. Additionally, the food industry contributes to greenhouse gases, especially in the transportation of products.

The project also aims to obtain metrics and patterns of online purchases to implement a recommendation system of products, local shops, etc., in the future. This Final Degree Project seeks to integrate technologies such as computer vision to address the environmental challenges of the food industry while promoting sustainability and the local economy.

This Final Degree Project begins with an introduction, which will discuss its motivations and objectives. The next section develops the state of the art, presenting handwritten text recognition, the tools and applications considered for the development of the application, QR codes, computer vision applications in fresh products, and finally, the tools used for data analysis and visualization.

The methodology section describes the steps taken during the project, from the development of the application to data analysis, along with the acquisition of videos and the changes made.

Finally, there is a discussion of the project’s results and the conclusion commenting on the scope of the objectives and future improvements.

²<https://merakiprojectes.eu/el-mercat-a-un-boto/>

Keywords: Computer vision, handwriting, visualization, patterns, analysis.

Índice general

1. Introducción	2
1.1. Motivación	2
1.1.1. ODS	2
1.2. Objetivos	4
1.3. Estructura del TFG	5
2. Estado del arte	8
2.1. Reconocimiento de texto	8
2.1.1. Reconocimiento Óptico de Caracteres	9
2.1.2. Reconocimiento de Texto a Mano	11
2.2. Aplicaciones y herramientas	12
2.2.1. Herramientas OpenSource	12
2.2.2. Aplicaciones de uso	15
2.3. Códigos de respuesta rápida	17
2.4. Computer Vision en Productos frescos	19
2.4.1. Detección de productos frescos	19
2.4.2. Clasificador de productos	20
2.5. Análisis y visualización de datos	21
2.5.1. Herramientas	21
3. Metodología	25
3.1. Objetivos e hipótesis	25
3.2. Herramientas	26
3.2.1. Aplicación	27
3.2.2. Análisis y visualización	28
3.3. Diseño	29
3.3.1. Computer Vision	29
3.3.2. Análisis de Datos	33
3.4. Procedimiento	36

4. Resultados y discusión	40
4.1. Testeo aplicación	40
4.1.1. Siguiendo las recomendaciones	40
4.1.2. No siguiendo las recomendaciones	43
4.2. Análisis	44
5. Conclusiones	48
5.1. Conclusiones del proyecto	48
5.2. Mejoras futuras	50
Bibliografía	52

Índice de figuras

2.1. Ejemplo de OCR de Microsoft.	9
2.2. Comparación QR y código de barras.	18
3.1. Detección de bordes.	30
3.2. Resultado del programa.	30
3.3. Archivo CSV final.	32
3.4. PowerBI página 1.	35
3.5. PowerBI página 2.	35
3.6. PowerBI página 3.	36
3.7. Diagrama de Gantt simplificado.	38
4.1. Imagen original cebolla dulce.	41
4.2. Error de carteles de fondo.	42
4.3. Resultado del video en CSV.	43
4.4. Resultado del video con fallos.	43
4.5. Resultado sin recomendaciones.	43
4.6. Resultado sin recomendaciones en CSV.	44
4.7. Gráficos de edad y género.	45
4.8. Mapa coroplético.	45
4.9. Comparación de importe y usuarios semanales.	46
4.10. Comparación de importe y usuarios mensuales.	46
4.11. Media de compra mensual.	47

1. Introducción

El presente Trabajo de Fin de Grado (TFG) se enfoca principalmente en la obtención de los precios de cada producto a partir de un video tomado desde el teléfono móvil del propio tendero del mercado.

1.1. Motivación

Actualmente, el calentamiento global es un tema preocupante para muchas personas. La cadena de producción alimentaria emite contaminantes a la atmósfera, así como emisiones de carbono y gases de efecto invernadero. El proyecto 'El mercat a un bot(ó)', junto con otras colaboraciones, ha optimizado una web-app adaptada a las necesidades de mercados y clientes. Además, se ha lanzado una campaña de sensibilización orientada a desmontar mitos entorno a la compra de productos frescos para atraer a la ciudadanía a los mercados locales, donde se ha comprobado que se produce un impacto positivo en consecuencia del principal propósito de este proyecto: reducir la huella de carbono cuidando del planeta y promoviendo las compras locales de productos frescos en los mercados.

Dado que los precios en los mercados fluctúan diariamente, se ha determinado que la actualización diaria de estos precios en plataformas digitales puede resultar pesada para los vendedores y se requiere de conocimientos en ciencia de datos, como el uso de la visión por computadora para obtener, a través de imágenes o videos, el nombre de cada producto junto con su precio, y así actualizarlos de manera automática.

1.1.1. ODS

Estoy interesada en este proyecto sobretodo debido a que contribuye a la reducción de las emisiones de Gases de Efecto Invernadero, y refleja una perspectiva

innovadora y sostenible en cuanto al cambio en la cadena de producción alimentaria. Además, aborda de manera clara los ODS (Objetivos de Desarrollo Sostenible)(Gamez, 2015), que es un tema que se ha tratado en clase y me ha interesado mucho.

Teniendo en cuenta el contexto del tema de desarrollo del TFG, los ODS a los que podría contribuir son, por orden, los siguientes.

Primeramente el ODS 2 (Hambre Cero). Aunque no contribuiría directamente al fin del hambre en el mundo, sí que ayudaría a algunas de las metas de este objetivo, como podrían ser la 2.4 y la 2.c. La primera quiere asegurar la sostenibilidad de los sistemas de producción de alimentos y aplicar prácticas que aumenten la producción, así manteniendo los ecosistemas y fortaleciendo la capacidad de adaptación al cambio climático, los fenómenos meteorológicos extremos, las sequías, etc. La 2.c quiere facilitar el acceso oportuno a información sobre los mercados, en particular sobre las reservas de alimentos, a fin de ayudar a limitar la extrema volatilidad de los precios de los alimentos.

Le sigue el ODS 8 (Trabajo Decente y Crecimiento Económico), así como el ODS 9 (Industria, Innovación e Infraestructura). El ODS 8 quiere la revitalización de los mercados municipales para generar empleo local y apoyar el crecimiento económico en la comunidad. Y el ODS 9 busca construir infraestructuras resilientes, promover la industrialización inclusiva y fomentar la innovación. La aplicación propuesta impulsa la digitalización del comercio local al facilitar transacciones y mejorar la visibilidad de los productos locales. Esta digitalización es una forma de innovación que puede fortalecer la competitividad de los negocios locales y contribuir al crecimiento económico sostenible, por eso junto estos dos ODS.

El 4º ODS considerado importante para este proyecto es el ODS 11 (Ciudades y Comunidades Sostenibles), esto se debe a que al incentivar la compra local y promover prácticas sostenibles en el sistema alimentario urbano, el proyecto contribuye a la construcción de comunidades más sostenibles y al ODS 11, que busca lograr ciudades inclusivas, seguras, resilientes y sostenibles.

Además, ahora se desarrollarán los 2 ODS más importantes y tratados en este proyecto.

Primeramente, el ODS 12 (Producción y Consumo Responsables). ‘El Mercat a un Bot(ó)’ contribuye al ODS 12, ya que se centra en patrones de producción y consumo sostenibles, mediante la promoción de la compra local, la reducción de desperdicios alimentarios, la sostenibilidad en la cadena de suministro y la concientización sobre prácticas sostenibles. La *webapp* conecta a los usuarios con mercados locales, fomentando la producción local, reduciendo desperdicios y optimizando la cadena de suministro. Además, educa a los usuarios sobre prácticas sostenibles, apoyando así la consecución de objetivos relacionados con la sostenibilidad y la responsabilidad en el consumo.

Y para finalizar el ODS 13 (Acción por el Clima), por la reducción de las emisiones de Gases de Efecto Invernadero mediante la promoción de la compra local y la minimización de la dependencia de cadenas de suministro extensas. Además, al destacar la importancia de productos de proximidad, se apoya indirectamente prácticas agrícolas más sostenibles, moderando la necesidad de técnicas agrícolas intensivas para la producción masiva y promoviendo la acción climática y la sostenibilidad ambiental.

En resumen, ‘El Mercat a un Bot(ó)’ se alinea con diversos ODS para tener un impacto positivo en la sociedad al promover prácticas de producción y consumo responsable que abordan los desafíos por el cambio climático. Todo esto ha resultado en una gran motivación para el desarrollo de este Trabajo de Fin de Grado.

1.2. Objetivos

El objetivo principal de este TFG es la realización de una solución para que los comerciantes tengan la posibilidad de actualizar los precios de los productos frescos en una plataforma de compras online. Esto se haría mediante la subida diaria, a la plataforma, de un vídeo de la exposición de los productos. Además, la misma plataforma permite la realización de un análisis de datos y patrones para la realización futura de un sistema de recomendaciones. Esto contribuiría también al aumento de la transparencia y la accesibilidad de la información para los consumidores, y además, da soporte a la competitividad de los mercados locales.

A continuación se comentarán los objetivos específicos de este Trabajo de Fin de Grado.

1. **Estudio:** Entender las distintas formas en las que se podría abordar este trabajo. También la comprensión y estudio de las distintas herramientas y aplicaciones con las que se puede trabajar.
2. **Aplicación:** Crear una aplicación utilizando el lenguaje de programación Python que detecte el nombre del producto, su precio y la unidad del precio.
3. **Automatización:** Automatizar la extracción de esos *items* a partir de un vídeo.
4. **Métricas y análisis:** Analizar y obtener métricas y gráficos a partir de los datos recogidos de compras online.
5. **Testear:** Testear la aplicación con videos grabados en el propio mercado.

1.3. Estructura del TFG

En la próxima sección se desarrolla el estado del arte o marco teórico de este trabajo. Se empieza exponiendo lo que es el reconocimiento de texto, primero de forma general y más tarde especificando en el reconocimiento de texto a mano. A continuación, se comentan las distintas herramientas y aplicaciones de código abierto más conocidas y que se han tenido en cuenta a la hora del desarrollo de la aplicación. También se habla en esta sección de los códigos de respuesta rápida o QRs, ya que se plantearon como plan B del proyecto, además se habla de las distintas aplicaciones de la visión por computadora en productos frescos. Y para finalizar, se habla de las herramientas utilizadas para el análisis y visualización de datos.

En el siguiente apartado se desarrolla la metodología y pasos llevados a cabo durante el proyecto. Pasando del desarrollo de la aplicación al análisis de datos final, junto con la obtención de los videos y los distintos cambios que se han ido realizando.

Para finalizar, los dos últimos apartados han sido reservados para la discu-

sión de los resultados y las conclusiones obtenidas de este proyecto. Comentando el alcance de los distintos objetivos y las mejoras a futuro de este proyecto.

2. Estado del arte

En este apartado se describirán las distintas soluciones para el desarrollo de este proyecto, especialmente dentro de la visión por computadora, que hace referencia a un grupo de tecnologías o herramientas que permiten captar imágenes, procesarlas y generar u obtener información a través de ellas (Robotics, [2022](#)).

En primer lugar, se presentarán las diferentes formas de uso del Reconocimiento Óptico de Caracteres (Optical Character Recognition u OCR) y el Reconocimiento de Texto a Mano (Handwriting Text Recognition o HTR), que será el que realmente se utilizará en este proyecto. En segundo lugar, se describirán los distintos tipos de aplicaciones y herramientas de reconocimiento de texto. En tercer lugar, se comentará una posible segunda opción con el uso de códigos de respuesta rápida (Quick Response o QR). En cuarto lugar, se hablará sobre los distintos usos de la visión por computadora en productos frescos, y finalmente se presentarán las distintas herramientas para el análisis y visualización de datos.

2.1. Reconocimiento de texto

El reconocimiento de texto es un área significativa dentro de la visión por computador, y se entiende como el proceso por el cual se convierte una imagen de texto en un formato de texto que pueden leer las máquinas como se puede ver en la Figura 2.1. Dentro del mismo podemos encontrar varios tipos, de los cuales se explicarán los dos mencionados anteriormente: OCR y HTR.

Use one of your own files or choose from a sample below.



Sample form #3



Detected attributes JSON

```
Nutrition Facts Amount Per Serving
Serving size: 1 bar (40g)
Serving Per Package: 4
Total Fat 13g
Saturated Fat 1.5g
Amount Per Serving
Trans Fat 0g
Calories 190
Calories from Fat 110
Percent Daily Values are based on
a diet of Vitamin A 50%
calorie diet.
```

Figura 2.1: Ejemplo de OCR de Microsoft.

2.1.1. Reconocimiento Óptico de Caracteres

El Reconocimiento óptico de caracteres se trata del área más general de lo que es una familia de algoritmos diseñados para la conversión de imágenes a texto, estas imágenes pueden ser fotografías, documentos escaneados o incluso generadas por una computadora (Cheng et al., 2017).

Los orígenes del OCR se remontan a 1914, cuando Emanuel Goldberg desarrolló una máquina que leía caracteres y los convertía en código telegráfico, y en años siguientes, continuó su investigación en documentos en microfilm. Más adelante, en 1974, creó una compañía llamada Kurzweil Computer Products Inc. y se enfocó en una máquina de lectura para ciegos, como explica Boesch (2023).

Al cabo de los años, y con la aparición del *deep learning*, llegaron muchos cambios en el campo de *computer vision* y, por consiguiente, al OCR. Algunos de ellos fueron durante los años 90, donde esto fue más prominente debido a la digitalización de periódicos históricos, así como la aparición de los teléfonos móviles y distintos documentos electrónicos. Como consecuencia, Tesseract es uno de los motores de OCR más populares y utilizados, contando con un código fuente Open Source desde 2005, que sigue en continuo desarrollo y crecimiento gracias a las distintas técnicas

deep learning.

El proceso general del OCR puede ser dividido en 5 etapas. (Equipo de BDRs, 2023)

1. Se empieza con la adquisición de la imagen, donde se captura la imagen. Se debe asegurar la calidad de la imagen, para así conseguir buenos resultados, por ello se deben tener en cuenta: la resolución, el enfoque y la iluminación o brillos que puedan encontrarse. Tener en cuenta estos aspectos proporcionará mejores resultados en el OCR.
2. Por consiguiente, la siguiente etapa o fase es el preprocesamiento de la imagen. Aquí se realizan una serie de operaciones para la mejora de la imagen y su preparación para el OCR. Algunas de las técnicas serían el ajuste del contraste y la eliminación de interferencias, como la iluminación y los brillos. También se pueden aplicar distintos tipos de filtros y algoritmos para la mejora de la claridad de los bordes, para así obtener una imagen nítida y clara. El objetivo principal de esta fase es la optimización de la calidad de la imagen y la eliminación del ‘ruido’ que pueda afectar a la precisión del algoritmo de OCR.
3. La siguiente fase es la segmentación, donde se divide la imagen en zonas más pequeñas, como podrían ser palabras o incluso caracteres individuales, lo que es esencial para procesarlos correctamente utilizando distintos algoritmos y técnicas de segmentación de caracteres. Una segmentación precisa es muy importante para obtener buenos resultados, dependiendo del idioma o el estilo de la fuente la segmentación puede ser un desafío.
4. La cuarta fase es el reconocimiento de los caracteres. Aquí se utilizan distintas combinaciones de algoritmos de aprendizaje automático y técnicas para identificar y asignar los patrones de cada carácter en cada segmento. Se analizan cada una de las características de cada carácter, como el tamaño, forma, contorno, espaciado, y se comparan con una base de datos de caracteres ya conocidos. Los algoritmos de clasificación y las redes neuronales son utilizados para el reconocimiento de los patrones y así realizan la correspondencia de los caracteres con la base de datos. Se le asigna a cada segmento el carácter más probable a medida que se realiza el reconocimiento. Hay que tener en cuenta

que la elección del algoritmo y modelo dependen tanto del idioma como del contexto en el que se va a aplicar.

5. La última fase es el postprocesamiento, donde se analizan y se corrigen los posibles errores del algoritmo, para así conseguir mejorar la precisión del texto reconocido. Para ello se pueden aplicar algoritmos o reglas adicionales, como la corrección de errores ortográficos, la coherencia gramatical o la validación de fechas o números. El objetivo de esta fase es perfeccionar la calidad del texto reconocido, asegurando la mayor precisión y confiabilidad.

2.1.2. Reconocimiento de Texto a Mano

El reconocimiento de texto a mano es un tipo de OCR creado para descifrar texto escrito a mano de forma automática utilizando la computadora (Dietrich, 2021). Esto proporciona una gran oportunidad a la automatización de muchos negocios.

OCR se centra en texto impreso y da relativamente muy buenos resultados. No obstante, estos resultados son frecuentemente debidos a la poca diferencia que se encuentra entre los distintos tipos de fuentes de escritura. Con HTR esto es más complicado: cada persona tiene distinta caligrafía e incluso puede cambiar su forma de escribir con el paso del tiempo. La variabilidad de estos patrones es muy alta, por lo tanto se debe entrenar un modelo lo suficientemente amplio para que no haya errores demasiado grandes. La diferencia que encontramos es que para HTR dentro de la fase de reconocimiento se deben tener en cuenta muchos factores, como la proporción de los trazos de cada letra, utilizando Hidden Markov Models (HMM) para determinar su secuencia (Calvo-Zaragoza et al., 2016).

A día de hoy, la mayoría de algoritmos implementan redes neuronales convolucionales (Convolutional Neural Networks o CNN) para la extracción automática de las características de estos trazos, esto seguido de redes neuronales recurrentes (Recurrent Neural Networks o RNN) o incluso memorias a largo-corto plazo (Long-Short Term Memories o LSTM) con una consecuente fase de decodificación para obtener los resultados del reconocimiento de texto (IBM, 2023).

Se puede ver así que HTR es una evolución del OCR y que todavía se encuentra

en una fase temprana de su potencial desarrollo.

2.2. Aplicaciones y herramientas

En este apartado se explicarán y comentarán las distintas herramientas de tipo OpenSource creadas para el reconocimiento de caracteres, además de las distintas aplicaciones web o móvil que existen. Entre ellas se pueden encontrar: Tesseract, TensorFlow, Keras, Document AI, API Cloud Vision, Amazon Textract, Google Lens, One Note, PDFelement y PentoPrint.

2.2.1. Herramientas OpenSource

Tesseract

Tesseract¹ es un proyecto open source con una trayectoria desde 1985, al cual han contribuido empresas como Hewlett-Packert y Google. Esta herramienta facilita las tareas de OCR, ya que cuenta con modelos tipo LSTM (Long short term memory) para la identificación de caracteres y dígitos. También es usada para la identificación de escritos a mano alzada. Tesseract soporta múltiples formatos de archivo como texto plano, hOCR (HTML), PDF, entre otros.

Esta herramienta ha sido utilizada en diversos estudios y proyectos, como podría ser la implementación de una red neuronal para la detección de matrículas de vehículos (Ortega Lobo, 2020).

TensorFlow

TensorFlow² se trata de una librería de código libre para Machine Learning (ML). Fue desarrollado por Google para satisfacer las necesidades a partir de redes neuronales artificiales. TensorFlow permite construir y entrenar redes neuronales para detectar patrones y razonamientos usados por los humanos. TensorFlow juega un

¹<https://pypi.org/project/pytesseract/>

²<https://www.tensorflow.org/?hl=es-419>

papel muy importante, para el desarrollo de la inteligencia artificial y el aprendizaje automático.

Además de trabajar con redes neuronales, TensorFlow es multiplataforma. Trabaja con GPUs y CPUs e incluso con las unidades de procesamiento de tensores (TPUs). Esta herramienta ha sido utilizada para la realización de un clasificador de productos agrícolas para control de calidad (González y A, 2020) junto con Keras, de la que se habla a continuación.

Keras

Keras³ es una Interfaz de Programación de Aplicaciones (Application Programming Interface o API) de redes neuronales escrita en lenguaje Python. Se trata de una biblioteca de código abierto que se ejecuta sobre frameworks como TensorFlow. Diseñada para ser modular, rápida y fácil de usar, Keras ofrece una forma sencilla e intuitiva de crear modelos de Deep Learning. Facilita enormemente la creación de capas para las redes neuronales o la configuración de arquitecturas complejas.

Un modelo Keras consiste en una secuencia o un gráfico independiente. Hay varios módulos totalmente configurables que pueden combinarse para crear nuevos modelos. Una de las ventajas de esta modularidad es que resulta muy fácil añadir nuevas funcionalidades como módulos independientes. Keras es, por tanto, muy flexible y adecuada para la investigación y la innovación. Y ha sido empleada, junto con TensorFlow, para la creación de un clasificador de productos agrícolas, como ha sido comentado con anterioridad (González y A, 2020).

Document AI

Document AI⁴ es una solución de Procesamiento de Lenguaje Natural (NLP) de Google Cloud Platform basado en la tecnología OCR. Este servicio puede reconocer textos, imágenes y caracteres en más de 200 idiomas.

Con Document AI se pueden extraer datos no estructurados a partir de la

³<https://keras.io/>

⁴<https://cloud.google.com/document-ai?hl=es>

imagen de un documento: factura, formulario, ticket, etc. El servicio interpreta la información utilizando Machine Learning y la estructura en formato ‘clave:valor’ o en una tabla.

También es posible la localización de documentos, a través de los datos que contenga, la identificación de patrones o tendencias y la creación de estadísticas. Además, ha sido aplicada para traducción automática (Ubillús Sánchez, 2022).

API Cloud Vision

API Cloud Vision⁵ contiene diversas características como la detección de rostros, detección de objetos, detección de contenido inapropiado, y la relacionada con este proyecto, reconocimiento óptico de caracteres.

La API también cuenta con la capacidad de extraer texto de imágenes mediante el reconocimiento óptico de caracteres. Esto es especialmente útil en situaciones en las que se necesita extraer información de documentos, letreros, carteles u otros contenidos escritos presentes en una imagen. El OCR de Google Cloud Vision API es compatible con varios idiomas, lo que amplía aún más su utilidad en aplicaciones y servicios globales.

También se ha empleado en la detección y traducción de objetos físicos para personas con discapacidad visual (Mendoza Escobar, 2018).

Amazon Textract

Amazon Textract⁶ es un servicio de escritorio que utiliza aprendizaje automático avanzado para extraer escritura a mano y texto impreso de cualquier documento o imagen. Este software OCR puede extraer datos de tablas, identificaciones, facturas, pasaportes y otros documentos en minutos. Puede extraer texto editable y procesable de imágenes y documentos. Utiliza inteligencia artificial y aprendizaje automático para escanear y extraer con precisión texto de formularios, tablas, imágenes, archivos PDF, etc. También funciona con documentación profesional como

⁵<https://cloud.google.com/vision/docs?hl=es-419>

⁶<https://aws.amazon.com/es/textract/>

recibos y facturas. Además utiliza respuestas de consulta para analizar y especificar el tipo de datos que desea extraer.

Otra característica interesante de Amazon Textract es la revisión humana incorporada. Después de extraer el texto impreso y la escritura a mano de un documento, este software de OCR te permite agregar reseñas y comentarios para mostrar tus pensamientos fácilmente. Curiosamente, utiliza IA para dar la retroalimentación correcta sin ninguna entrada manual.

Un posible uso es la extracción de información a partir de imágenes de facturas (Sánchez Duarte et al., 2022).

2.2.2. Aplicaciones de uso

Google Lens

Google Lens⁷ es un servicio de Google que funciona como herramienta de búsqueda visual, identificando cosas que aparecen en tus fotos. La funcionalidad de Google Lens ha aumentado recientemente, ahora al abrir una fotografía que contenga algún tipo de texto relativamente legible, la foto mostrará un pequeño icono de copiar. Al pulsarlo Google Lens comenzará a analizar más en profundidad la imagen y más tarde aparecerá la opción de copiar el texto de la imagen en una barra lateral. De momento, Google Lens en escritorio se limita a la función de reconocer texto. No funciona con texto escrito a mano, no identifica objetos ni tampoco resuelve ecuaciones.

Otro uso diferente que se le ha dado a esta herramienta ha sido el reconocimiento de tipos de serpientes en Panamá (Santamaria et al., 2022), esto da a entender la diversidad de usos que se le pueden dar a esta aplicación.

⁷<https://lens.google/intl/es/>

OneNote

OneNote⁸ es un bloc de notas digital, que es compatible con Reconocimiento óptico de caracteres, permite copiar el texto de una imagen o copia impresa de archivo y pegarlo en las notas para que pueda realizar cambios en las palabras. Esta es una buena manera de hacer cosas como copiar información de una tarjeta de presentación que haya digitalizado en OneNote. Después de extraer el texto, se puede pegar en otro lugar de OneNote o en otro programa, como Outlook o Word.

Esta herramienta sobretodo da pie a su uso en la enseñanza, tanto por parte de docentes como de estudiantes (Bailon Lourido et al., 2021).

PDFelement

Wondershare PDFelement⁹ es una potente herramienta PDF que ayuda a los usuarios de Windows a trabajar con archivos PDF más fácilmente y les permite producir documentos y formularios PDF de forma rápida, asequible y segura. Esto permite desbloquear el texto dentro de las imágenes utilizando la tecnología OCR, así utilizando campos específicos definidos por el propio usuario dentro del archivo PDF.

Esta aplicación puede ser utilizada para la gestión de documentos electrónicos (Pascual Pérez y Vásquez Grande, 2022).

Pen to Print

Pen to Print¹⁰ es tanto una aplicación como una API, está diseñado específicamente para reconocer la escritura a mano en imágenes escaneadas y archivos PDF. Está capacitado para extraer y convertir diferentes tipos de escritura a mano en texto: letras mayúsculas, cursivas y regulares, ordenadas o desordenadas o finas.

Pen to Print está diseñado específicamente para computadoras de escritorio para convertir escritura a mano en texto en el entorno de oficina o en casa, como se

⁸<https://www.microsoft.com/es-es/microsoft-365/onenote/digital-note-taking-app>

⁹<https://pdf.wondershare.es/>

¹⁰<https://www.pen-to-print.com/>

comenta es su propia web. Permite una conversión rápida de documentos grandes como imágenes o archivos PDF a Word o texto, haciéndolos fácilmente editables, para leer escritura a mano desordenada y notas en cursiva en mensajes de texto en línea.

Se puede acceder al OCR de Pen to Print desde cualquier dispositivo, móvil o de escritorio, sin necesidad de descargar ningún software previamente. El motor HTR de Pen to Print se encuentra disponible como un servicio API, por tanto se puede aprovechar para la creación de este proyecto, encargándose del reconocimiento y conversión de imagen a texto haciendo también un preprocesado inicial para dar mejores resultados.

Esta aplicación ha sido empleada para la transformación de textos manuscritos de idiomas como el Swahili (Wanjawa et al., 2023).

2.3. Códigos de respuesta rápida

Los códigos QR (Quick Response) son representaciones de datos digitales que se pueden imprimir para más tarde escanear desde un teléfono móvil u otros dispositivos. El uso de códigos QR ha experimentado un notable aumento desde el inicio de la pandemia de COVID-19, debido a que ayudan a evitar el contacto con objetos o incluso personas.

Estos tienen una amplia gama de usos, entre ellos compartir menús de restaurantes, redireccionar a una página web, perfil de redes sociales o formulario, o puede simplemente contener texto escrito.

Los códigos QR pueden ser dinámicos o estáticos, la principal diferencia es que los dinámicos pueden ser editados una vez se han imprimido. Esto implica que se puede modificar la funcionalidad y el contenido del mismo en cualquier momento y las veces necesarias, contienen una URL (Uniform Resource Locator) corta que redirige a la página de destino o al contenido. Además ofrece ver las distintas estadísticas, incluyendo el número de escaneos, la fecha, la hora o la ubicación del escaneo, incluso el sistema operativo que se ha utilizado para escanearlo. Por otro

lado, los estáticos no ofrecen ninguna de estas opciones, no pueden modificarse una vez impresos y tampoco es posible ver la información adicional (Bitly, 2024).



(a) Ejemplo QR.



(b) Ejemplo Código de barras.

Figura 2.2: Comparación QR y código de barras.

Los códigos QR son muy similares a los códigos de barras en los productos de un supermercado o paquetes de correos (Figura 2.2). Estos códigos se pasan por escáneres infrarrojos para obtener la información que contienen, esta información se encuentra almacenada a lo largo de un solo eje, de forma horizontal. Los códigos de barras registran el conjunto de números que especifican la identificación de un producto. Sin embargo, se considera que los datos almacenados son redundantes, lo que permite que incluso si parte del código de barras se destruye u oscurece, aún es posible que un dispositivo lea la identificación del producto. Los QRs por otro lado contienen información tanto en el eje horizontal como en el vertical, esto les permite contener mucha más información, lo que hace que sean tan versátiles en su uso (Geurts, 2021).

Los datos en un código QR son una serie de puntos en forma de cuadrícula, cada punto representa un 1 y cada espacio en blanco representa un 0 en código binario. Estos patrones codifican conjuntos de números y letras, formando así texto o URL. También cuentan con los marcadores de posición, que son los cuadrados colocados en las esquinas: ambas superiores y la inferior izquierda. Esto permite que las cámaras orienten el código QR al escanearlo. Y al igual que los código de barras, los QRs cuentan con redundancia de datos para que se sigan pudiendo escanear incluso siendo difíciles de leer. Por esta razón, cuando se quiere añadir un logotipo

al código QR, estos datos tapados por el logotipo se pueden recuperar fácilmente. Cuando se escanea un código QR, el lector de QR en la cámara del teléfono descifra el código, y la información resultante activa una acción en el dispositivo. Si el código QR contiene una URL, esta se mostrará para redirigir desde el navegador. Por otro lado si contiene un texto este se puede copiar para pegarlo en otro lugar.

Una alternativa menos eficiente al reconocimiento de texto mediante *computer vision* es la utilización de QRs para contener la información relevante de cada cartel del mercado. En cambio, la aplicación de computer vision para el reconocimiento de escritura en el ámbito del mercado local es mucho más versátil ya que no requiere de códigos prefijados y no requiere ningún conocimiento técnico por parte de los vendedores.

2.4. Computer Vision en Productos frescos

El mercado minorista se encuentra experimentando una transformación significativa, impulsada por la innovación tecnológica. La IA (Inteligencia Artificial) es una de las fuerzas impulsoras de esta transformación. A continuación se comentará algún ejemplo de lo que se ha realizado en este ámbito.

2.4.1. Detección de productos frescos

Grabit está introduciendo un software de visión por computador para productos como frutas y verduras. Trabajando junto con Dibal, empresa que desarrolla soluciones de pesaje y etiquetado de productos, han llevado esta tecnología a los supermercados. Este proyecto busca ahorrar tiempo a los compradores y los empleados integrando balanzas inteligentes que puedan identificar los productos que están siendo pesados. Además de esto, Grabit también está tratando de crear soluciones autoservicio, que permitan a los clientes escanear y pagar sus productos previniendo pérdidas y asegurando que el producto escaneado coincida con el código de barras (Grabit, 2023).

Sin embargo, Grabit no es el primero en realizar algo similar, en 2018 la School

of Information Technology (Escuela de Tecnología de la Información) de Suecia realizó este mismo proyecto, utilizando 10 productos diferentes para su estudio (Femling et al., 2018). Contando con un nivel de precisión variable según el tipo de fruta y su color, además del tipo de redes neuronales convolucionales utilizadas (Inception y MobileNet), además crearon una interfaz de uso con el objetivo de ser lo más sencilla posible para el usuario.

2.4.2. Clasificador de productos

Otro ejemplo de esta transformación sería el trabajo realizado en la Escuela Superior Politécnica de Chimborazo en Ecuador (González y A, 2020).

Este trabajo pretende crear un sistema capaz de supervisar, identificar y clasificar la calidad de productos del sector agrícola. Esto es debido a que las técnicas existentes en ese momento no eran lo suficientemente avanzadas para alcanzar los estándares de calidad y cubrir la demanda del mercado internacional.

Este sistema se desarrolló mediante la aplicación de técnicas de visión por computadora y algoritmos de aprendizaje automático que contribuyen a la identificación de imágenes en tiempo real. La investigación permitió implementar algoritmos de clasificación de K vecinos más cercanos para etiquetar los productos según su calidad y enviar los reportes en tiempo real a una aplicación web. Los productos utilizados en este estudio fueron: bananas, naranjas, plátanos y manzanas (González y A, 2020).

El impacto de estas soluciones de IA puede crear un efecto dominó para el desarrollo de más tecnologías útiles tanto para los minoristas como para los clientes. Por un lado, los minoristas cuentan con que la automatización no sólo acelera el proceso de pago, sino que minimiza los errores humanos, que a su vez reduce el número de quejas y devoluciones de los clientes. Además, esto permite que el personal centre sus esfuerzos en otras tareas, como servicio al cliente y gestión de inventario. Por otro lado, los clientes también cuentan con el beneficio de un proceso rápido de compra y facturación sin errores, significando una entrada y salida de clientes más eficiente.

La IA está transformando y seguirá transformando la forma en la que operan tanto supermercados como minoristas con soluciones como el reconocimiento visual y el autoservicio, se está allanando el camino hacia un futuro automatizado y eficiente.

2.5. Análisis y visualización de datos

En este apartado se hará mención de las distintas herramientas de posible uso para la visualización de los datos obtenidos de las compras online.

2.5.1. Herramientas

Las herramientas BI (Business Intelligence) pueden ayudar a las diversas empresas en la toma de decisiones, proporcionando información y análisis detallados. Algunas herramientas que vamos a abordar son más conocidas, como PowerBI, Tableau o QlikSense, sin embargo, también hablaremos de otras dos importantes como lo son Sisense y Tibco Spotfire (Emilio, 2021).

PowerBI

Microsoft PowerBI¹¹ ofrece funciones de tipo *self-service* permitiendo a los usuarios explorar la herramienta sin necesidad de conocimientos especializados. Integra otros servicios como Excel, SharePoint o Visual Studio, por lo que ayuda a las empresas a buscar una visibilidad completa desde todos los departamentos de la organización.

PowerBI además ofrece una serie de funciones diseñadas para ayudar a las organizaciones en la colaboración entre equipos, además proporcionando un acceso seguro a la información confidencial. Entre las mejores funcionalidades de PowerBI podemos encontrar: visualizaciones interactivas, consultas en lenguaje natural, conectividad de datos robusta e información y alertas automatizadas. Además de ser una de las herramientas de más fácil uso y entendimiento.

¹¹<https://www.microsoft.com/en-us/power-platform/products/power-bi>

Power BI fue la herramienta principal en la implementación de Data Mart para el análisis de ventas a clientes, y el uso de *Business Intelligence* en el inventario en los Ecomercios ‘Gransol’ (Bermeo-Pérez y Campoverde-Molina, 2020) y (Bermeo Moyano y Campoverde Molina, 2020). Además, se realizó un análisis comparativo entre Power BI y Qlik Sense, del que se hablará más adelante (Collaguazo et al., 2024).

Tableau

Tableau¹² es una de las herramientas de BI más fáciles de utilizar junto con PowerBI. Su interfaz facilita la creación de cuadros de mando interactivos, además de análisis predictivos para que los usuarios puedan anticipar tendencias futuras. Proporciona una interfaz intuitiva, integrándose con bases de datos como Oracle, SQL Server, PostgreSQL, entre otras.

Entre sus mejores funcionalidades se pueden observar: conectividad de datos robusta, generación de información rápida, colaboración con otras herramientas e integración de inteligencia artificial y *machine learning* (García-Vellido Santías, 2020).

QlikSense

QlikSense¹³ es otra de las soluciones BI diseñada de forma específica para clientes que necesitan generar informes, pero no tienen intención en invertir en infraestructura IT (Information Technology) o que no tienen experiencia técnica como para utilizar otras herramientas. Proporciona una interfaz de usuario con funcionalidad de arrastrar y soltar gráficos interactivos, ofrece además análisis predictivo, procesamiento del lenguaje natural, algoritmos de aprendizaje automático y más.

Las mejores funcionalidades de QlikSense son: su escalabilidad, análisis de datos automatizado, herramientas avanzadas de visualización y, seguridad y privacidad (Collaguazo et al., 2024).

¹²<https://www.tableau.com/es-es>

¹³<https://www.qlik.com/es-es/products/qlik-sense>

Sisense

Sisense¹⁴ se distingue por su capacidad para el manejo de grandes volúmenes de datos, tanto estructurados como no estructurados, de forma rápida y eficaz con su tecnología In-Chip, que ha revolucionado la extracción de valor de los conjuntos de Big Data. Sisense ofrece potentes capacidades analíticas que permiten crear fácilmente informes y cuadros de mando, además de contar con funciones que permiten a los usuarios crear imágenes y gráficos a partir de sus datos.

Las mejores funcionalidades de Sisense son: Una interfaz fácil de usar, *self-service* BI, capacidades analíticas avanzadas, automatización y escalabilidad e información en tiempo real.

Se ha realizado un análisis de datos de la banca corporativa y de inversión utilizando esta herramienta (Rodríguez Merchán y Martínez Guzmán, 2018).

Tibco Spotfire

Tibco Spotfire¹⁵ tiene una gran velocidad a la hora de procesar números, combinando visuales personalizables que permiten explorar las relaciones entre las variables, correlaciones, etc. Además, sus servicios proporcionan una visión más profunda de las tendencias y comportamientos de los clientes, cuenta con un modelo de seguridad para que la información confidencial permanezca a salvo de intrusiones.

Sus mejores funcionalidades son: una arquitectura robusta y flexible, capacidad de integración completa, seguridad a nivel empresarial, facilidad de uso y alto rendimiento. Además, se ha realizado un estudio de la optimización de la aplicación web de Industria 4.0 utilizando Tibco Spotfire (González Luna, 2021).

¹⁴<https://www.sisense.com/>

¹⁵<https://www.spotfire.com/>

3. Metodología

En este capítulo se discute el funcionamiento del proyecto, su desarrollo y problemáticas que se han ido encontrando. Empieza explicando los distintos objetivos del proyecto y las posibles rutas que se podrían tomar para su desarrollo. La segunda parte de este capítulo trata sobre de las herramientas que finalmente se han utilizado para el desarrollo de la aplicación y el análisis. Continúa abordando del propio diseño de la aplicación y el preprocesado y análisis de datos. Finalmente se hablará de los pasos que se han seguido durante el proyecto.

Las imágenes de los productos alimenticios que aquí aparecen han sido autorizadas para su uso.

3.1. Objetivos e hipótesis

Los objetivos generales y específicos de este TFG se encuentran en la introducción del trabajo. Este apartado dentro de la metodología se centra en los objetivos propios del desarrollo de la aplicación, que ahora se explicarán uno a uno.

1. **Busqueda de información.** Buscar información sobre proyectos o herramientas que pudiesen hacer el desarrollo del proyecto más sencillo y de fácil uso.
2. **Creación de la aplicación.** La creación de la aplicación para su uso a partir de imágenes, es decir, que fuese funcional pasando una imagen. Empezando por la lectura de una imagen hasta poder indagar en la zona de interés de la imagen.
3. **Procesamiento de imágenes.** A partir del anterior, que pudiese procesar diversas imágenes distintas de forma seguida.
4. **Detección en vídeo.** Realizar el cambio de la detección en imágenes a la detección en vídeos.

5. **Transformación de resultados.** El último objetivo en relación a la parte de visión por computador fue la transformación de esos datos recogidos a un archivo para su uso posterior.
6. **Preprocesado.** El preprocesado de los datos recibidos de las compras online.
7. **Análisis y visualización.** El objetivo final fue la realización de un análisis y visualización de esos datos.
8. **Uso de códigos QR.** Como objetivo secundario se encontraba la posible utilización de códigos QR que contuviesen la información de cada producto.

Al comienzo del proyecto se comentó la posibilidad de realizar el entrenamiento de un modelo de detección de carteles o de precios, sin embargo, se llegó a la conclusión de que con el poco tiempo con el que se contaba y la dificultad que esto podría conllevar, debido a la gran cantidad de datos e imágenes que recopilar y el tiempo de modelado y entrenamiento, no iba a ser posible contar con un buen resultado al finalizar. Además, ya existían soluciones disponibles que realizaban esta tarea de forma eficiente.

Otras de las opciones balanceadas fue el uso de los códigos QR, que como se encuentra explicado en los objetivos, debían contener la información de cada producto, es decir, el nombre y su precio. Esto podría ser una buena opción debido a que los códigos QR se han estandarizado mucho los últimos años y son de fácil creación y fácil detección, a través de la cámara propia del teléfono móvil.

En cuanto a la fase de análisis de datos, se quiso obtener la información necesaria y los patrones relevantes relacionados con las compras en los mercados. Así como obtener relaciones sobre si un negocio puede llevar a la gente a otros negocios próximos a él.

3.2. Herramientas

En este subapartado se explica que herramientas han sido utilizadas para el desarrollo del proyecto y por qué, para así asegurar un mejor entendimiento del proceso.

3.2.1. Aplicación

En el segundo capítulo del estado del arte explicaba las distintas herramientas y aplicaciones de código abierto más utilizadas y con más opciones de funcionalidad para este proyecto. Todas ellas fueron puestas en práctica y testadas de forma simple al inicio del proyecto para tomar la decisión de cuál llevaría a un mejor desarrollo y resultados.

La mayoría de las herramientas presentadas constan de mejores resultados utilizando imágenes con texto a máquina, cuando lo que realmente requiere este proyecto es el reconocimiento de texto escrito a mano. Por ello, se pensó en la posibilidad de realizar una detección de escritura a mano únicamente de los números que pudiesen aparecer en el cartel, es decir, el precio de cada producto, de esta manera el nombre del producto podría ser impreso como texto a máquina o incluso sacado a partir de los códigos QR que se planteaban como segunda opción. Al llegar a esta conclusión, se intentó hacer pruebas de detección de números con Tesseract, TensorFlow y Keras utilizando el dataset MNIST (Modified National Institute of Standards and Technology), una base de datos de números escritos a mano con distintas caligrafías que consta de 60000 imágenes de entrenamiento y 10000 imágenes de prueba. Esto no dio buenos resultados debido a que no solamente se necesitan números si no también comas o puntos (, '.') para adaptarse a cada puesto del mercado.

Las aplicaciones OneNote y Google Lens contaban con la problemática de que se debe hacer de forma manual desde el propio teléfono móvil. OneNote también se encuentra para uso desde el ordenador, sin embargo no había posibilidad de que fuese a través de una imagen, si no que transforma la escritura dentro de la aplicación en texto que después puede ser pegado en otros documentos. Por otro lado, PDFelement cuenta con más funciones dirigidas a la detección de texto dentro de un documento ya escrito a máquina, por lo tanto no tiene tan buenos resultados como se esperaba.

Finalmente, este apartado acaba desarrollando el uso que se le ha dado y la funcionalidad de la aplicación seleccionada para el desarrollo del proyecto, Pen to Print. Esta aplicación tiene su uso principal a través del teléfono móvil, es decir, se entra en la aplicación, se saca una fotografía, esa imagen es procesada y la aplicación

devuelve todo el texto que se pueda ver en la imagen separado en filas. También cuenta con una fase en la que se pueden realizar correcciones sobre el texto que ha detectado en caso de que se vea un error. Después de testarla diversas veces con distintos tipos de letras, iluminaciones y ángulos, se llegó a la conclusión de que era la que mejores resultados daba.

Una vez tomada la decisión de que Pen to Print sería la aplicación con la que se trabajaría el proyecto, llegó el momento de ver cómo se integraría de manera funcional. El uso de la aplicación no era factible debido a que llevaría tiempo explicarle a cada tendero su uso, y les costaría un tiempo diario que podría no parecerles rentable, que es justo una de las cosas que se quieren evitar con este proyecto. Se trató de buscar si la aplicación contaba con código abierto o un modelo del que se pudiese partir con el proyecto, finalmente a través de la web Rapid API se descubrió que Pen to Print contaba con una API que podría utilizarse para el desarrollo de la aplicación. Y por esa razón y después de ver los resultados se decidió que esta sería la herramienta que se utilizará como lectora de texto a mano.

3.2.2. Análisis y visualización

En esta sección se hablará de las dos herramientas utilizadas para tanto el preprocesado de datos como la visualización de los mismos, Jupyter Notebook y PowerBI, respectivamente.

Jupyter Notebook es una aplicación en la que se puede trabajar tanto en el lenguaje de programación Python como en R. En este proyecto se ha trabajado en Python, ya que es uno de los lenguajes de programación más extendidos y populares actualmente debido a su alta usabilidad y simpleza (Peña, 2021). Jupyter como se ha comentado se utilizará solamente durante el preprocesado de los datos, que luego serán enviados a PowerBI para mostrarlos de forma más visual. La decisión de utilizar PowerBI se basó en su versatilidad, alta variedad de funciones y su fácil uso tanto para el desarrollador como para el usuario final.

3.3. Diseño

3.3.1. Computer Vision

Esta fase del proyecto tiene como objetivo automatizar la actualización de precios en páginas web a partir de imágenes o vídeos de un puesto de compra en un mercado local o similar. Aquí no se discutirá realmente la actualización de los precios en la aplicación, sino la extracción de esos precios a partir de las imágenes o videos. Se empezará explicando la versión inicial del código centrada en la extracción a partir de imágenes y, más tarde, la versión que extrae la información desde un vídeo.

Antes de empezar con la explicación del código se debe poner en contexto que esto tiene como objetivo ponerse en funcionamiento dentro de una aplicación/*webapp*.

Imágenes

El código se desarrolló con el pensamiento de que las imágenes se obtendrían de un carpeta, como si las imágenes se hubiesen subido a la aplicación y a partir de ahí se pusiese en funcionamiento el programa.

Se decidió utilizar la API de Pen to Print para el desarrollo del proyecto después de varios intentos de utilizar las distintas herramientas y aplicaciones de las que se ha hablado anteriormente, y observar que esta era la que contaba con mejores resultados.

En principio la API de Pen to Print ha sido creada para enviar imágenes sin la necesidad de preprocesarlas, sin embargo, el código se hizo con el objetivo de encontrar mejores resultados debido a las distintas iluminaciones y calidad de imágenes que se podían encontrar.

El programa empieza obteniendo cada una de las imágenes que se encuentran en la carpeta con las imágenes originales. Para cada imagen primeramente la pasa a escala de grises para poder hacer la detección de los bordes de la misma, esto quiere decir que detecta las zonas en las que se ve un cambio brusco de tonalidad, en este

caso en la escala de grises, y las señala como zonas en las que hay un borde. Una vez delineados los bordes o contornos de la imagen, encuentra los que están cerrados y tienen una forma similar a un polígono rectangular o cuadrado, es decir, que tenga 4 vértices.



Figura 3.1: Detección de bordes.

La figura 3.1 muestra un ejemplo de la detección de ese polígono rectangular en una imagen a color. Una vez esto ocurre, se encuentra el borde más grande de la imagen y se centra en esa zona para recortarla un poco y guardarla en otra carpeta para enviarla más tarde a la API de Pen to Print y así obtener el mejor resultado posible.

```
{'result': '1', 'subScans': [], 'value': "ARROZ\nREDONDO\n3'50k"}  
{'result': '1', 'subScans': [], 'value': "ARROZ\nALBUFERA\n5'35 K"}  
{'result': '1', 'subScans': [], 'value': "ARROZ\nPROHIBIDO\n5'35 K"}  
{'result': '1', 'subScans': [], 'value': "ARROZ\nINTEGRAC\n3'50K"}  
["ARROZ\nREDONDO\n3'50k", "ARROZ\nALBUFERA\n5'35 K", "ARROZ\nPROHIBIDO\n5'35 K", "ARROZ\nINTEGRAC\n3'50K"]
```

Figura 3.2: Resultado del programa.

El resultado obtenido una vez se envía la imagen a la API aparecería como un archivo JSON (JavaScript Object Notation) como se puede observar en la Figura 3.2.

Vídeos

A partir del código centrado en las imágenes se empezó a trabajar en el de vídeo. Este funciona de manera similar, la diferencia principal es en cómo se consiguen las imágenes para enviar a la API.

En esta ocasión en vez de obtener imágenes de la carpeta, se obtienen vídeos. Los videos son considerados una sucesión de imágenes a gran velocidad, que se llaman frames. Cada uno de estos frames pueden ser procesados o leídos como imágenes individuales.

El programa tiene el siguiente funcionamiento, se obtiene el vídeo y se guardan los tres primeros 3 frames en distintas variables, de forma que cuando se pasa al cuarto frame el primero que ha sido guardado se sustituye por el segundo, ese segundo se sustituye por el tercero y el tercero es sustituido por este nuevo frame, el cuarto, y así sucesivamente. Cada vez que se reemplazan, los frames guardados se compararán entre ellos, si el programa considera que se parecen lo suficiente entre ellos se guarda el frame guardado en la variable de en medio, es decir, el segundo.

Una vez guardados los frames en un carpeta nueva, se pone en marcha la versión original del código que se ha explicado con anterioridad que estaba centrada en las imágenes.

Sin embargo, también se añadió al programa una nueva parte que transforma ese archivo JSON a un archivo CSV (Comma Separated Values) con los valores limpios, es decir, separando en columnas el nombre del producto, el precio y la unidad (Figura 3.3). Esto se llevó a cabo con la utilización de reemplazos de caracteres y separadores o *splits*.

```
precios_29-04-2024.csv
1 ,Producto,Precio,Unidad
2 0,ARROZ REDONDO,3.50,k
3 1,ARROZ ALBUFERA,5.35,K
4 2,ARROZ PROHIBIDO,5.35,K
5 3,ARROZ INTEGRAC,3.50,K
```

Figura 3.3: Archivo CSV final.

Para que esto funcione bien, se deben seguir ciertas pautas tanto en la escritura del cartel como a la hora de grabar el video.

1. El cartel debe estar escrito en el orden 'producto', 'precio' y 'unidad', por ejemplo, 'Patatas 0.49 Euro/Kg'. Esto es debido a que si se encuentra escrito de otra manera las columnas no coinciden con lo escrito, por ejemplo, '0'49 Euro/kg patatas' puede hacer que todas las columnas tengan mal los datos.
2. Si el nombre del producto cuenta con un número este debe ser escrito en letra, por ejemplo, 'arroz tres delicias'.
3. Se recomienda que los carteles estén escritos en letra mayúscula y clara para una mejor detección de los nombres.
4. Al grabar en video no es importante si el móvil se encuentra de forma horizontal o vertical. Pero si es importante no cambiarlo mientras se graba el video.
5. No es necesario grabar todos los productos del puesto, solamente aquellos productos nuevos o a los que se les ha cambiado el precio.
6. Asegurarse de que cuando se pare el vídeo solo salga un producto.
7. El video debe enfocarse y pararse en cada producto que se quiera actualizar, sin embargo, no se debe parar más de 1 segundo porque puede detectar la misma imagen varias veces.
8. El video debe ser estable, ya que si al parar el video se mueve mucho la cámara es muy probable que no detecte ningún producto.

Si se quiere probar u observar el código, este ha sido subido a GitHub¹.

3.3.2. Análisis de Datos

En la sección de análisis de datos se facilitaron los datos de compras realizadas de forma online. Aquí se tratarán los datos de productos del mercado d'Algirós, específicamente productos frescos, sin embargo, se cuenta con gran variedad de datos desde gastos en bares hasta compras de ropa o bolsos.

Preprocesado

Los datos se proporcionaron en un archivo excel donde se podía encontrar gran variedad de datos, como podrían ser la fecha de compra, el ID del usuario que ha realizado la compra, el importe gastado en esa compra, el comercio, la edad del usuario, la fecha de registro en la aplicación, el género, el código postal del usuario, el ticket de compra medio del comercio, el número de compras mensuales del usuario, la cantidad de compras que ha hecho el usuario, el dinero que le queda en su monedero de la aplicación, sus bonificaciones y la cantidad de gente que ha invitado a la aplicación, entre otros.

Se le tuvo que hacer un preprocesado a los datos debido a la cantidad de datos faltantes y columnas vacías. En otras columnas se sustituyeron los valores nulos por la palabra 'Indefinido' o por un cero si se consideraba óptimo. Estas columnas en las que se sustituye el valor nulo por cero contaban con una gran cantidad de nulos y con ningún cero, el número mínimo que aparecía en todas ellas era 1. Por tanto, se consideró que esto podría significar que por ejemplo, al no haber invitado a nadie a la aplicación no se ha guardado ese dato, y solamente se guarda una vez se invita a alguien, esto podría haber sucedido con todas estas columnas, y por ello es que se decidió hacer este cambio.

Además se tuvo que modificar la columna 'CP' (Código Postal) debido a que en el excel se la consideraba numérica y a los valores de código postal que contaban

¹<https://github.com/JuliaMPTP/MercatAUnBoto.git>

con un 0 como primer número, este les desaparecía, convirtiéndolos en otro código postal diferente.

Visualización

Una vez realizado el preprocesamiento de los datos, para tenerlos de forma más visual, se trasladaron a Power BI, un servicio de analítica de datos de Microsoft que proporciona visualizaciones interactivas para la inteligencia empresarial con una interfaz simple en la que los usuarios finales pueden crear por sí mismos sus propios informes y paneles o dashboards.

A continuación, se presentarán los diferentes tipos de gráficos utilizados en el dashboard:

1. **Gráficos de líneas:** Utilizados para mostrar tendencias a lo largo del tiempo, permitiendo visualizar cambios y patrones en los datos de manera continua.
2. **Gráficos de barras:** Empleados para comparar cantidades, proporcionando una visión clara y precisa de las diferencias y similitudes en los datos.
3. **Gráficos de pastel o *pie charts*:** Utilizados para representar la proporción de diferentes partes en relación a un todo, facilitando la comprensión de la distribución porcentual de los datos.
4. **Mapas coropléticos:** Utilizados para mostrar variaciones en un indicador a lo largo de una región geográfica, donde las diferencias en los colores representan distintas magnitudes de un valor específico.
5. ***Tree maps*:** Empleados para visualizar datos jerárquicos y proporciones en un espacio limitado, permitiendo identificar rápidamente las relaciones y tamaños relativos de las categorías.
6. **Tarjetas:** Utilizadas para resaltar métricas clave y proporcionar una visión rápida y concisa de los valores más importantes en el dashboard.

Una vez los datos se encontraron en Power BI se empezaron a realizar distintas visualizaciones de los mismos, como podrían ser el recuento de las personas por edad o género, la zona de residencia utilizando el código postal (Figura 3.4) o la

comparación de gastos de compra y recuento de usuarios semanal y mensualmente (Figura 3.5).

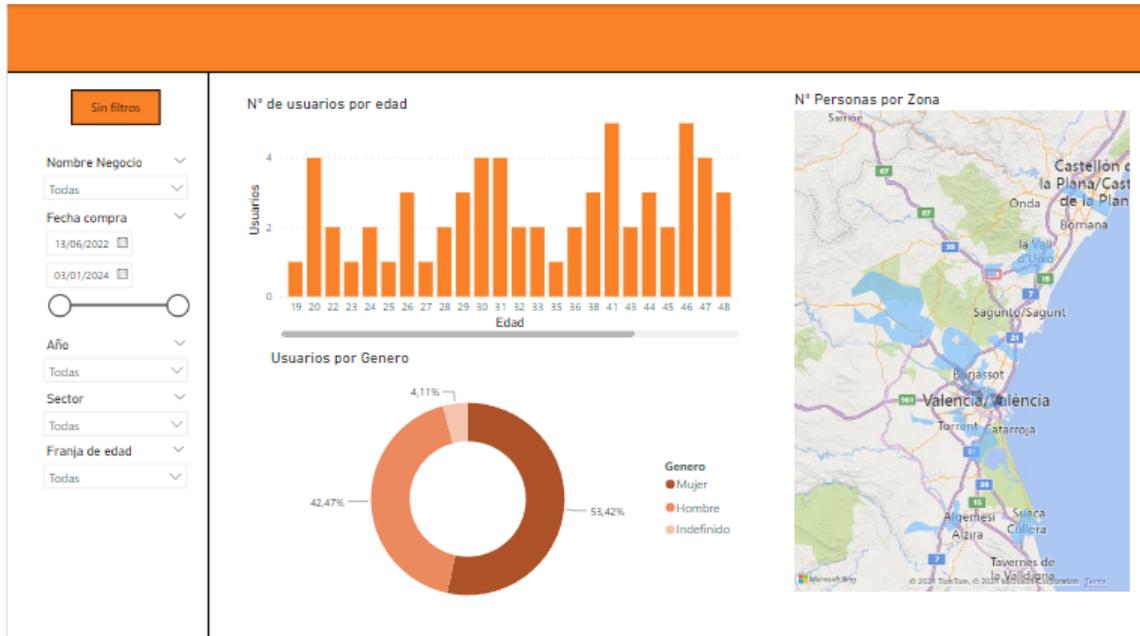


Figura 3.4: PowerBI página 1.



Figura 3.5: PowerBI página 2.

Además se añadieron filtros de fecha o de sector de compra, además de un filtro por negocio, para saber la misma información para negocios en específico, aunque estos estarían disponibles únicamente de forma anónima.



Figura 3.6: PowerBI página 3.

La Figura 3.6 es la tercera página del PowerBI en la que se pueden encontrar en la parte superior distintas tarjetas con valores entre los que se encuentran el promedio de los bonos de consumo que tienen los usuarios, el promedio de bonos pendientes de gastar, el promedio de bonos de prescripción, el dinero promedio que siguen teniendo los usuarios dentro de la aplicación y para finalizar la media de devoluciones realizadas por los usuarios. Además cuenta con un Treemap, que muestra de forma jerárquica cuales son los sectores que más dinero recogen. Siendo el primero los bares y el mercado el quinto. Los gráficos serán explicados detalladamente en el apartado de resultados.

3.4. Procedimiento

Este trabajo se encuentra enmarcado como un proyecto de investigación. Esto implica un mayor tiempo de desarrollo debido a las diversas reuniones que se tuvieron que llevar a cabo, en las que se encontraban tanto tutores como profesores implicados en el proyecto, llegando a incluir comunicación con los tenderos en el mercado para el mejor desarrollo del mismo.

Han existido reuniones semanales con uno de los tutores al cual se le comentaba

lo que había ido haciendo durante la semana y se le preguntaba sobre las dudas que habían ido surgiendo. También se intercambió número de teléfono con la tutora de la empresa para comunicarnos de forma más rápida y directa, de esta manera se pudo quedar para realizar las distintas pruebas en el mercado y aclarar dudas.

Así como hubo reuniones semanales, también las hubo mensuales, en las que asistían más personas relacionadas con el proyecto y discutimos el proceso y avance realizado, así como las posibles mejoras a futuro.

Las visitas al mercado fueron varias, en las que se aprovechó tanto para informar a los tenderos sobre el proyecto, como para pedir permiso para realizar fotos y vídeos para la realización de los testeos de la aplicación. También se habló con algunos de ellos para ver si ellos detectaban algún patrón de compra o similar. Haciendo esto se vió que es cierto que se vende más los sábados, y centrándonos en el sector de Mercado en los gráficos se puede ver que se vende mucho más en los meses fríos que en el resto del año.

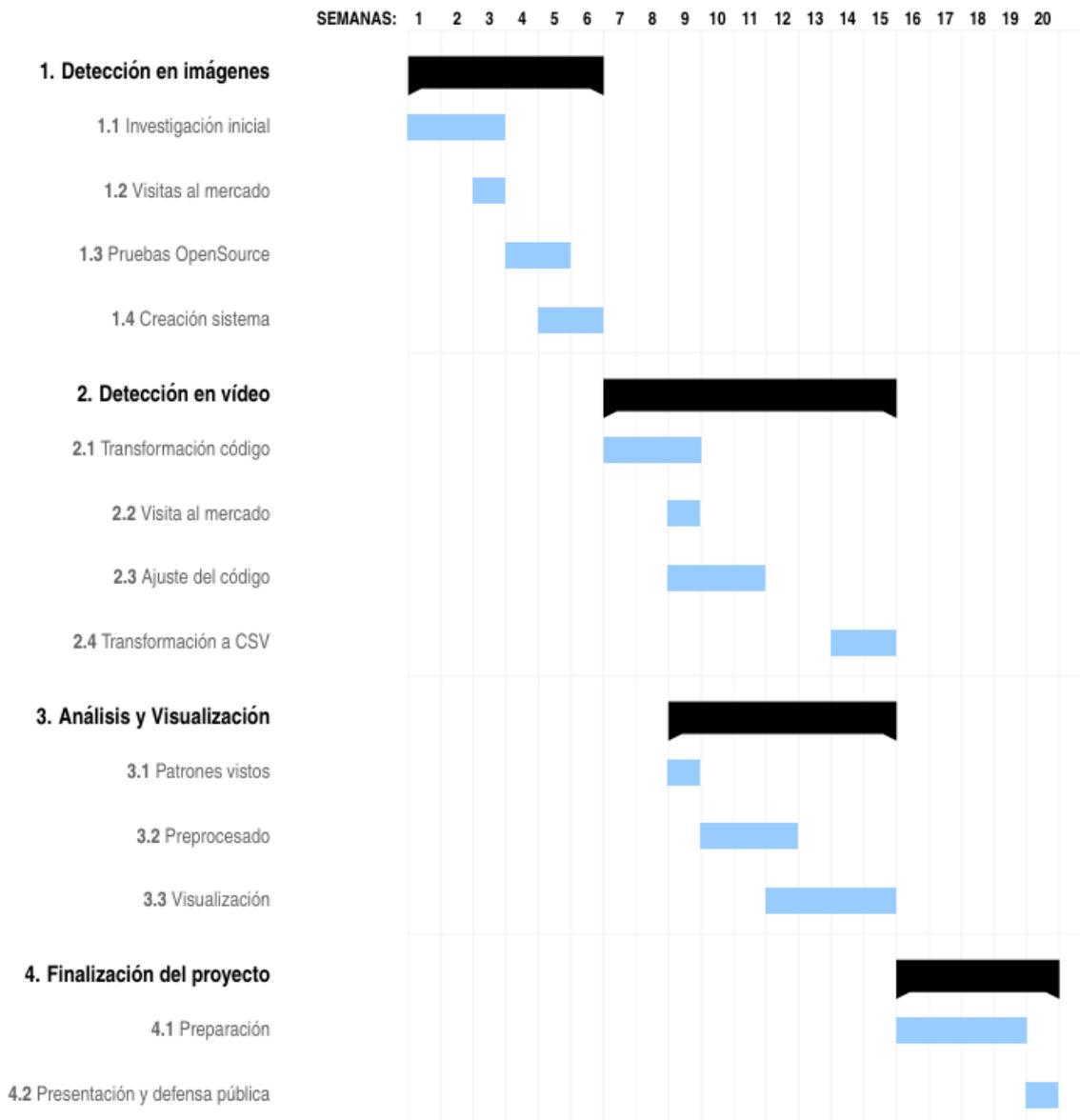


Figura 3.7: Diagrama de Gantt simplificado.

En la Figura 3.7 podemos ver el procedimiento general que se ha seguido durante la realización del trabajo representado en forma de diagrama de Gantt. Además, en este se añade la etapa de finalización del proyecto, en la cual se llevaron a cabo los preparativos para la presentación del mismo en Las Naves, el centro de innovación social y urbana, de Valencia.

4. Resultados y discusión

Para finalizar el proyecto se debe realizar la fase de testeo para asegurar que el programa funciona correctamente y que los resultados sean los esperados.

4.1. Testeo aplicación

Se realizaron distintos tipos de videos para ver los diferentes resultados que podrían salir, sobre todo dependiendo de si se siguieron las recomendaciones indicadas en el capítulo de metodología o no.

El testeo fue realizado con 15 vídeos distintos, variando desde 2 hasta 6 productos en cada uno de ellos, 8 de los vídeos siguen las recomendaciones y los otros 7 no. Primeramente se comentarán los resultados obtenidos siguiendo las recomendaciones y secundamente cuando no han sido seguidas.

4.1.1. Siguiendo las recomendaciones

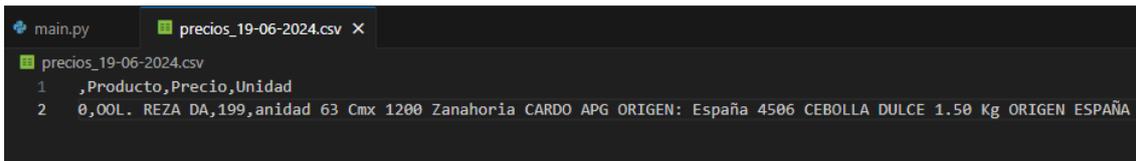
Calculando el nivel de precisión con videos muy cuidados, siguiendo las recomendaciones, en los distintos aspectos que se han comentado con anterioridad, se ha detectado un nivel de precisión de 0.94, siendo este bastante elevado.

Sin embargo, hay que tener muchas cosas en cuenta, como puede ser la letra y la forma en la que enfocas cada cartel. Un buen ejemplo de esto sería repasar las distintas figuras que hay en el proyecto, como la Figura 3.2, en la que como ejemplo se han puesto varios tipos de arroz, entre ellos se puede ver que hay ‘ARROZ INTEGRAL’ sin embargo debido al tipo de letra el programa lo ha detectado como ‘ARROZ INTEGRAC’, no entendiendo que la letra final es una ‘L’. Este sería uno de los errores más repetidos en el testeo, a continuación se comentarán con más detalle los tipos de errores encontrados.

1. Uno de los posibles errores es que el cartel no salga bien, puede ser que aparezca borroso o incluso recortado en algún lateral. Esto podría causar muchos errores de lectura, ya sea que no lea nada del cartel o que lea solamente una zona y por tanto, no sea una lectura correcta. Sin embargo, este error es poco común si se siguen las recomendaciones marcadas.
2. Otro error podría ocurrir si se pueden observar otros carteles de fondo en las imágenes tomadas, esto podría causar que el programa leyese varios carteles al mismo tiempo y por tanto causaría un problema a la hora de separarlo en las columnas correctas. Esto puede comprobarse en las siguientes figuras (Figura 4.1 y Figura 4.2).



Figura 4.1: Imagen original cebolla dulce.



```
main.py | precios_19-06-2024.csv x
precios_19-06-2024.csv
1 ,Producto,Precio,Unidad
2 0,00L. REZA DA,199,anidad 63 Cmx 1200 Zanahoria CARDO APG ORIGEN: España 4506 CEBOLLA DULCE 1.50 Kg ORIGEN ESPAÑA
```

Figura 4.2: Error de carteles de fondo.

3. El tercer error posible es que no se tome ninguna imagen del vídeo, esto podría ocurrir si el vídeo se toma sin parar en cada cartel o si no se espera el tiempo suficiente en cada uno de los carteles. Esto no debería ocurrir si se han seguido las recomendaciones.
4. El error más común es la mala lectura de caracteres, esto podría ser debido a una mala escritura o a que el programa interpreta que es una letra similar a la escrita. Como se ha comentado con anterioridad en la Figura 3.2.
5. El último error es una consecuencia de los anteriores, este sería el que no se guardasen los productos en el archivo CSV al no poder separarse en 3 columnas. Como se podrá ver a continuación con el vídeo realizado por una tendera en su parada del mercado.

Video realizado por una tendera

Uno de los vídeos con los que ha sido hecho el testeo fue tomado directamente por una de las tenderas del mercado. Esto se realizó con el objetivo de mostrar lo sencillo que realmente puede ser este proceso.

Los resultados de este vídeo fueron buenos. A pesar de ello, como se puede ver en la figura 4.3, al segundo producto detectado ‘Pak Choi’ le falta la letra ‘I’ al final porque en el cartel no se encontraba con la misma claridad que las otras letras. Además, uno de los productos que salía en el vídeo no se transcribió al CSV debido a que ocurrió lo mismo con la unidad del precio, y por tanto, no se pudieron dividir en las columnas correspondientes (Figura 4.4).

```
main.py precios_18-06-2024.csv X
precios_18-06-2024.csv
1 ,Producto,Precio,Unidad
2 0,COIRABE,6.00,/K
3 1,PAK CHO,4.50,/k
```

Figura 4.3: Resultado del video en CSV.

```
['COIRABE', '6.00', '/K']
['PAK CHO', '4.50', '/k']
['BOBBY', '2.50']
[['COIRABE', '6.00', '/K'], ['PAK CHO', '4.50', '/k']]
```

Figura 4.4: Resultado del video con fallos.

4.1.2. No siguiendo las recomendaciones

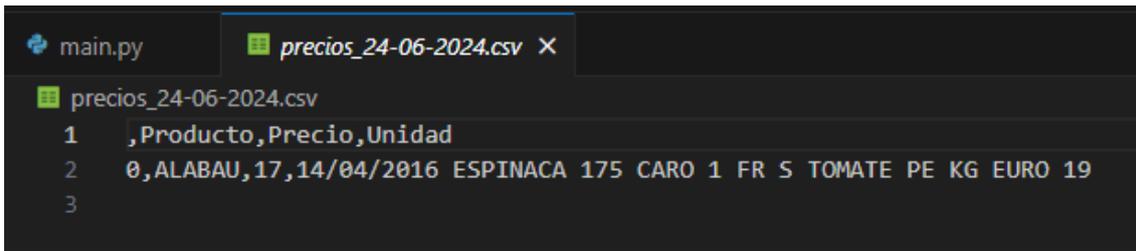
Por otro lado, al no seguir las distintas recomendaciones han podido haber resultados muy dispares, desde que la letra sea peor a que el video se haga demasiado rápido. También se debía tener en cuenta el enfoque y donde paras el video.

Con este gran abanico de posibilidades y procesando una gran cantidad de videos, el nivel de precisión varía en gran medida, desde 0.78 hasta incluso 0.12 si el video era excesivamente rápido y no paraba lo suficiente en los productos.

```
['6917 AU TOMATE KG PERA CURO', '1.98']
['ALABAU', '17', '14/04/2016 ESPINACA 175 CARO 1 FR S TOMATE PE KG EURO 19']
['PIMIENTOS KG DE ASAR']
[['ALABAU', '17', '14/04/2016 ESPINACA 175 CARO 1 FR S TOMATE PE KG EURO 19']]
There has been 2 products not detected.
```

Figura 4.5: Resultado sin recomendaciones.

En las figuras anteriores (Figura 4.5 y Figura 4.6) se puede ver un ejemplo del resultado que tuvo uno de los vídeos realizados, en el que de 3 productos solo se llegó a poder dividir en las columnas uno de ellos, y aun así este no se detectó correctamente.



```
main.py | precios_24-06-2024.csv X
precios_24-06-2024.csv
1 ,Producto,Precio,Unidad
2 0,ALABAU,17,14/04/2016 ESPINACA 175 CARO 1 FR S TOMATE PE KG EURO 19
3
```

Figura 4.6: Resultado sin recomendaciones en CSV.

4.2. Análisis

A continuación, se explicarán los resultados del análisis y visualización de los datos, se comprobarán las hipótesis y patrones planteados por los tenderos con los que se habló en las visitas al mercado y se realizarán observaciones de cada gráfico.

En la Figura 4.7 se pueden observar 2 gráficos, el primero dedicado al estudio de las edades de los usuarios y el segundo al género. Estos como todos los gráficos pueden cambiar drásticamente dependiendo de los filtros que se utilicen. De forma general y sin aplicar ningún filtro, se puede ver que no hay una cantidad muy grande de usuarios en los datos. Y que todas las edades varían desde 0 a 5 personas. Por otro lado, en el gráfico de género se observa que aunque con poca diferencia hay más mujeres que hombres realizando estas compras online.

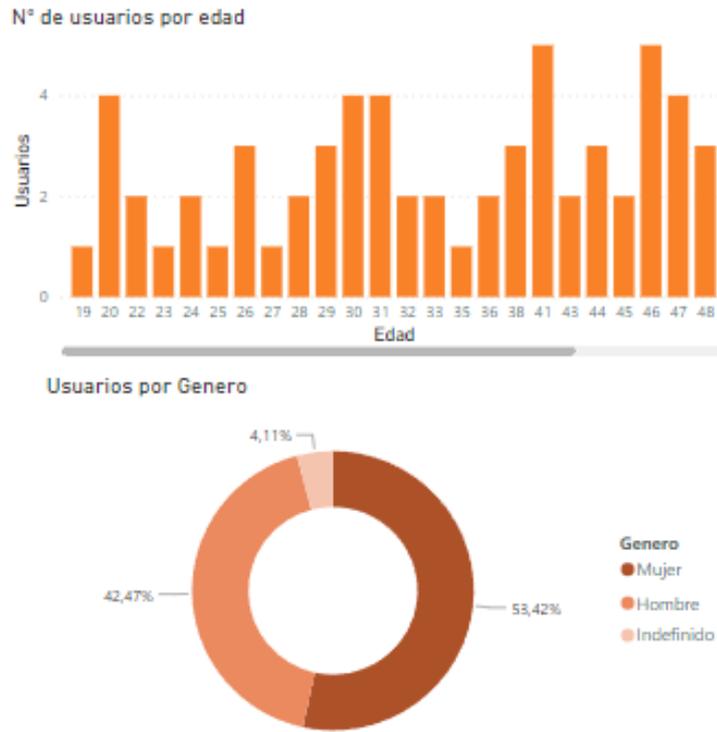


Figura 4.7: Gráficos de edad y género.

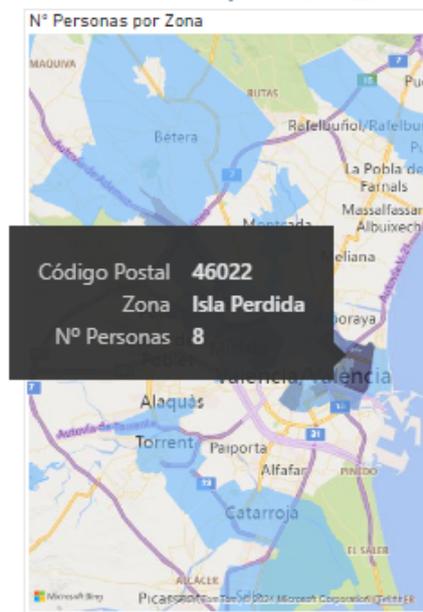


Figura 4.8: Mapa coroplético.

Con respecto al mapa coroplético que se puede ver en la Figura 4.8, se pueden observar las distintas zonas, detectadas a partir del código postal, en las que residen los usuarios de la aplicación. Pasando el ratón por encima de cada zona aparece un

recuadro de color negro en el que sale cierta información, como el código postal y el nombre de la zona, así como el número de personas que residen ahí. Cuantas más personas residan en una zona más oscura será el color de esa zona pintada en el mapa.

En la Figura 4.9 aparecen 2 gráficos, uno lineal en el que aparece el importe semanal, tanto total como medio y un gráfico de barras en el que se hace un recuento de las personas que han comprado ese día. Se puede observar una alta cantidad de usuarios que compran durante los viernes y sábados, sin embargo, puede ver que los sábados es el día que más dinero se gasta la gente con diferencia.



Figura 4.9: Comparación de importe y usuarios semanales.

Al observarlo de forma mensual (Figura 4.10), se puede ver que el mes más prominente en usuarios de compra es julio seguido de octubre, noviembre y diciembre, sin embargo, se ve una gran diferencia en cuanto a los gastos de compra, ya que se gasta mucho más estos últimos tres meses que en julio, aunque este también cuenta con un pico de compras. Esto podría deberse a que los productos se encarecen en esa época.

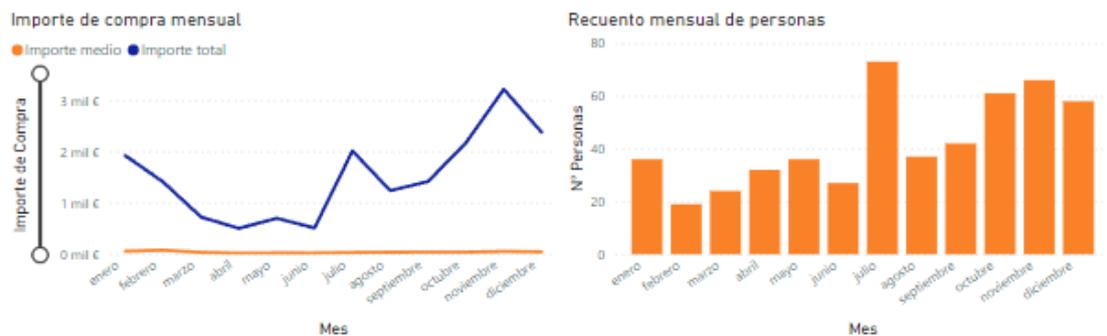


Figura 4.10: Comparación de importe y usuarios mensuales.

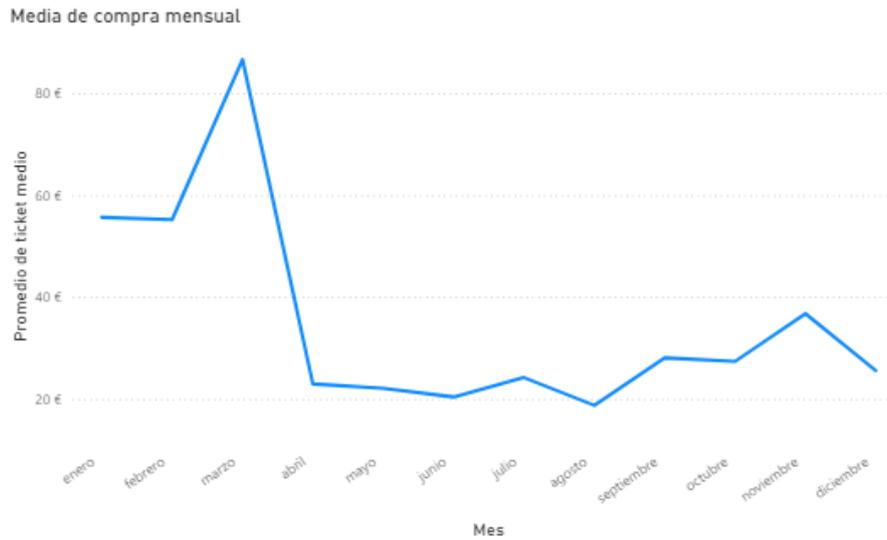


Figura 4.11: Media de compra mensual.

El último gráfico del que hablaremos se trata de la media del tique medio de compra de los negocios de forma mensual (Figura 4.11), se observa un gran pico en marzo, probablemente causado por la época de fallas en la que la gente realiza más actividades fuera de sus casas. En abril hay un descenso muy grande, pasando de alrededor de los 80 Euros a los 20 Euros, y luego continúa oscilando a lo largo del año.

5. Conclusiones

5.1. Conclusiones del proyecto

El proyecto presentado es pionero en introducir *computer vision* en el ámbito del producto fresco en el entorno local. Esta innovación no solo representa un avance tecnológico significativo, sino que también abre nuevas oportunidades para la digitalización y modernización de mercados tradicionales que se han rezagado en la utilización de tecnologías avanzadas.

Una de las principales ventajas de este proyecto es su capacidad para digitalizar e importar nuevas tecnologías en un entorno con un difícil punto de entrada, donde la variabilidad de los precios hasta ahora ha imposibilitado su automatización. Gracias a la implementación de *computer vision*, se pueden actualizar los precios de manera eficiente y precisa, algo que previamente requería un esfuerzo considerable y estaba sujeto a errores humanos.

La solución implementada permite la digitalización y actualización de precios y productos con la simple realización de un vídeo con el móvil. Esto no solo facilita el proceso para los comerciantes, sino que también mejora la experiencia del consumidor al proporcionar información actualizada y precisa sobre los productos disponibles en el mercado. La accesibilidad y facilidad de uso de esta tecnología son aspectos clave que contribuyen a su éxito y adopción en mercados locales.

Además, el proyecto busca obtener métricas y patrones de compras online, lo que permite visualizar la implementación futura de un sistema de recomendación de productos y locales. Esta característica no solo optimiza la experiencia de compra para los consumidores, sino que también apoya a los vendedores al proporcionarles datos valiosos sobre las tendencias y preferencias de los clientes.

En términos de sostenibilidad, el proyecto tiene un impacto muy positivo al promover la compra de productos locales y frescos, reduciendo la huella de carbono

asociada con el transporte de alimentos. Al fomentar las compras locales, se contribuye a la economía local y se apoya a los pequeños productores y comerciantes.

Este proyecto no solo es un avance significativo en el uso de tecnologías avanzadas como computer vision en mercados locales, sino que también promueve la sostenibilidad, la eficiencia y la modernización de un sector tradicional. Las soluciones implementadas ofrecen un camino claro hacia la digitalización y la mejora continua, estableciendo un modelo a seguir para futuras innovaciones en el ámbito del comercio local de productos frescos.

Ahora se comprobará que los objetivos que fueron expuestos en el capítulo de la introducción han sido fructíferos.

1. **Estudio:** Se ha logrado comprender las distintas herramientas con las que se podía abordar este proyecto, así como elegido con la que se recibieron mejores resultados.
2. **Aplicación:** Se ha logrado crear una aplicación que pueda detectar el producto, junto con su precio.
3. **Automatización:** Se ha automatizado la extracción del producto y precio partiendo de un vídeo.
4. **Métricas y análisis:** Se ha logrado procesar y visualizar los datos recibidos de las compras online.
5. **Testear:** Se ha testado la aplicación con videos realizados en el propio mercado.

Junto con esto se ha conseguido evitar la utilización de códigos QR para la detección de los productos. Esto ha resultado en una ventaja para los tenderos, debido a que no necesitarán crear ningún código QR nuevo ni deberán invertir tiempo en imprimirlos y pegarlos en cada cartel.

Por otro lado, el análisis de los datos de las compras online nos ha dado información que les podría ser útil a los distintos negocios. En concreto, en los mercados se puede observar que hay más gente comprando los sábados por tanto, ese día de la semana se deberían abastecer más los negocios del mercado que los demás días.

Además, los últimos meses del año (octubre, noviembre y diciembre) son los que más gasta la gente, esto podría ser debido al tipo de productos más especiales al acercarse navidad o podría deberse a que la gente compra más cantidad de alimentos sin embargo, se puede observar que el tique medio no varía en gran medida a lo largo del año, siendo el mayor gasto durante el mes de mayo, por tanto se puede deducir que el mayor gasto en esos meses es debido a que sube el número de gente que compra y no tanto que suba el precio de los productos.

Esto claro es observando el mercado en general, si nos fijásemos en cada uno de los negocios podrían existir grandes variaciones en los datos de unos a otros, por eso se quiere ofrecer a cada negocio su propio análisis.

5.2. Mejoras futuras

Como en todo proyecto siempre existe margen de mejora al finalizar, en cuanto a este proyecto en específico se pueden pulir diversas cosas.

1. **Automatización.** Se podría automatizar más la detección de los productos y precios a partir de los vídeos, de forma que no sea necesario parar ni enfocar demasiado en cada producto.
2. **Análisis.** Se podría realizar un análisis más exhaustivo sobre los datos obtenidos. Las visualizaciones y información valiosa que se podría sacar de los datos todavía se podría explotar más, con el objetivo de informar y crear esa inteligencia de negocio que puede proporcionar PowerBI y a la que se le podría sacar más partido.
3. **Toma de datos.** Al recopilar el texto sacado de las imágenes es posible que haya distintos resultados dependiendo de cómo ponga cada tendero los carteles. Por ejemplo, cuando se va a comprar frutos secos el precio se escribe de la siguiente manera ‘Almendra 3’00 1/4 Kg’ y esto puede causar una mala transformación de los datos a CSV.

Aún con estas mejoras a futuro, podemos concluir que este proyecto ha cumplido sus objetivos principales, consiguiendo desarrollar una aplicación de detección

de precios a partir de vídeos, junto con la analítica y visualización de datos. Una aplicación con buena perspectiva y de fácil uso.

Bibliografía

- Bailon Lourido, W. A., Arauz Barcia, G. O., & Macías Valencia, D. G. (2021). Utilización de herramientas ofimáticas por parte de docentes y estudiantes universitarios ecuatorianos [Publisher: Polo de Capacitación, Investigación y Publicación (POCAIP) Section: Dominio de las Ciencias]. *Dominio de las Ciencias*, 7(Extra 3), 471-492. Consultado el 21 de junio de 2024, desde <https://dialnet.unirioja.es/servlet/articulo?codigo=8229646>
- Bermeo Moyano, D. M., & Campoverde Molina, M. A. (2020). Implementación de Data Mart, en Power BI, para el análisis de ventas a clientes, en los Econegocios “Gransol” [Publisher: Imprenta y Casa Editora Çoni"Section: Polo del Conocimiento: Revista científico - profesional]. *Polo del Conocimiento: Revista científico - profesional*, 5(1 (Enero 2020)), 647-673. Consultado el 21 de junio de 2024, desde <https://dialnet.unirioja.es/servlet/articulo?codigo=7436046>
- Bermeo-Pérez, S. K., & Campoverde-Molina, M. A. (2020, enero). Vista de Implementación de inteligencia de negocios, en el inventario de la Cooperativa GranSol, con la herramienta Power BI. Consultado el 19 de junio de 2024, desde <https://fipcaec.com/index.php/fipcaec/article/view/169/263>
- Bitly. (2024, abril). ¿Cuál es la diferencia entre los Códigos QR Dinámicos y Estáticos? Consultado el 22 de abril de 2024, desde <https://support.qr-code-generator.com/hc/es/articles/7664170522893--Cu%C3%A1l-es-la-diferencia-entre-los-C%C3%B3digos-QR-Din%C3%A1micos-y-Est%C3%A1ticos>
- Boesch, G. (2023, noviembre). Optical Character Recognition (OCR) - The 2024 Guide. Consultado el 16 de febrero de 2024, desde <https://viso.ai/computer-vision/optical-character-recognition-ocr/>
- Calvo-Zaragoza, J., Toselli, A. H., & Vidal, E. (2016). Early Handwritten Music Recognition with Hidden Markov Models [ISSN: 2167-6445]. *2016 15th In-*

- ternational Conference on Frontiers in Handwriting Recognition (ICFHR)*, 319-324. <https://doi.org/10.1109/ICFHR.2016.0067>
- Cheng, Z., Bai, F., Xu, Y., Zheng, G., Pu, S., & Zhou, S. (2017). Focusing Attention: Towards Accurate Text Recognition in Natural Images, 5076-5084. Consultado el 22 de marzo de 2024, desde https://openaccess.thecvf.com/content_iccv_2017/html/Cheng_Focusing_Attention_Towards_ICCV_2017_paper.html
- Collaguazo, L. R. Q., Paula, G. G. E., Pilaguano, D. D. S., & Alajo, Á. J. L. (2024). Desarrollo de inteligencia empresarial: Análisis comparativo entre Power BI y Qlik Sense en la implementación de tableros estratégicos de indicadores clave. [Number: 1 Publisher: Editorial Kolibri]. *Technology Rain Journal*, 3(1). <https://doi.org/10.55204/trj.v3i1.e30>
- Dietrich, F. (2021). OCR vs. HTR or “What is AI, actually?” Consultado el 7 de febrero de 2024, desde <https://readcoop.eu/insights/ocr-vs-htr/>
- Emilio, N. (2021). Top 5 herramientas y software BI del mercado. Consultado el 3 de mayo de 2024, desde <https://blog.bismart.com/los-5-mejores-software-bi-herramientas-business-intelligence>
- Equipo de BDRs, E. d. B. (2023, julio). OCR (Reconocimiento Óptico de Caracteres): ¿qué es y cómo usarlo? Consultado el 10 de febrero de 2024, desde <https://blog.truora.com/es/ocr>
- Femling, F., Olsson, A., & Alonso-Fernandez, F. (2018). Fruit and Vegetable Identification Using Machine Learning for Retail Applications. *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 9-15. <https://doi.org/10.1109/SITIS.2018.00013>
- Gamez, M. J. (2015). Objetivos y metas de desarrollo sostenible. Consultado el 30 de mayo de 2024, desde <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>
- García-Vellido Santías, A. (2020). Análisis de datos y toma de decisiones utilizando una herramienta de business intelligence: Tableau [Accepted: 2020-09-24T17:37:11Z]. Consultado el 19 de junio de 2024, desde <https://idus.us.es/handle/11441/101457>

- Geurts, S. (2021). Códigos QR versus códigos de barras: ¿Cuál debes utilizar? Consultado el 14 de mayo de 2024, desde <https://rentman.io/es/blog/codigos-qr-versus-codigos-de-barras-cual-debes-utilizar-para-tus-equipos>
- González, C. A. G., & A, V. H. (2020). Clasificador de Productos Agrícolas para Control de Calidad basado en Machine Learning e Industria 4.0 [Number: 2]. *Revista Perspectivas*, 2(2), 21-28. <https://doi.org/10.47187/perspectivas.vol2iss2.pp21-28.2020>
- González Luna, D. (2021). Optimización de la aplicación web de Industria 4.0 TIBCO Spotfire [Accepted: 2021-03-02T14:52:37Z]. Consultado el 21 de junio de 2024, desde <https://uvadoc.uva.es/handle/10324/45454>
- Grabit. (2023). IA transformando el retail: una visión de Grabit – Grabit. Consultado el 29 de abril de 2024, desde <https://grabit.ai/ia-transformando-el-retail-una-vision-de-grabit/?lang=es>
- IBM. (2023, julio). ¿Qué son las redes neuronales convolucionales? | IBM. Consultado el 14 de mayo de 2024, desde <https://www.ibm.com/es-es/topics/convolutional-neural-networks>
- Mendoza Escobar, F. (2018). Cloud visión api para la detección y traducción de objetos físicos - personas con discapacidad visual. Consultado el 12 de junio de 2024, desde <https://repositorio.umsa.bo/handle/123456789/17480>
- Ortega Lobo, I. (2020). Implementación de la red neuronal YOLOv3 para la detección de matrículas de vehículos [Publisher: E.T.S.I. Telecomunicación (UPM)]. Consultado el 15 de junio de 2024, desde <https://oa.upm.es/63082/>
- Pascual Pérez, W. S., & Vásquez Grande, A. B. (2022, noviembre). *Lineamientos para la gestión de documentos electrónicos de la Unidad de Gestión documental y Archivo del Centro Nacional de Registro* [bachelor]. Universidad de El Salvador. Consultado el 21 de junio de 2024, desde <https://oldri.ues.edu.sv/id/eprint/31641/>
- Peña, M. J. (2021, enero). ¿Por qué es Python el lenguaje de programación más popular? [Section: Python]. Consultado el 12 de abril de 2024, desde <https://eiposgrados.com/blog-python/por-que-es-python-es-el-lenguaje-de-programacion-mas-popular/>

- Robotics, E. (2022, enero). Visión por Computador Qué es, Aplicaciones y Objetivos. Consultado el 11 de abril de 2024, desde <https://www.edsrobotics.com/blog/vision-computador-que-es/>
- Rodríguez Merchán, L. V., & Martínez Guzmán, E. M. (2018). Análisis de herramientas del enfoque de inteligencia de negocios: Caso de estudio datos de la banca corporativa y de inversión [Accepted: 2021-02-25T03:51:07Z]. Consultado el 17 de junio de 2024, desde <http://repository.unipiloto.edu.co/handle/20.500.12277/9944>
- Sánchez Duarte, L. C., Martínez Ramos, J. F., & Gallego León, J. S. (2022). Extracción de información de imágenes de facturas [Publisher: Universidad de los Andes]. Consultado el 18 de junio de 2024, desde <http://hdl.handle.net/1992/63385>
- Santamaria, K., Sole, M., & Rodríguez, K. (2022). Reconocimiento de tipos de serpientes en Panamá por medio de visión artificial. *Revista de Iniciación Científica*, 8(1). <https://doi.org/10.33412/rev-ric.v8.1.3513>
- Ubillús Sánchez, C. K. (2022). El uso de la traducción automática de Google Cloud en un proceso de posesión [Accepted: 2022-09-28T02:09:06Z Publisher: Universidad Peruana de Ciencias Aplicadas (UPC)]. *Universidad Peruana de Ciencias Aplicadas (UPC)*. Consultado el 10 de junio de 2024, desde <https://repositorioacademico.upc.edu.pe/handle/10757/661194>
- Wanjawa, B., Wanzare, L., Indede, F., McOnyango, O., Ombui, E., & Muchemi, L. (2023, julio). Kencorpus: A Kenyan Language Corpus of Swahili, Dholuo and Luhya for Natural Language Processing Tasks [arXiv:2208.12081 [cs]]. <https://doi.org/10.48550/arXiv.2208.12081>
- Comment: 24 pages, 6 figures.

