



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo Fin de Grado

**Grado en Administración y Dirección de
Empresas**

La Inteligencia Artificial en la Investigación de Mercados. El caso de la aplicación Wine Quality

Presentado por:

Adriana Juez Pérez

Tutelado por:

Víctor Temprano García

Valladolid, 16 de junio de 2023

AGRADECIMIENTOS

Agradecer en primer lugar a mi tutor de TFG, sin el cual no hubiera sido posible desarrollar este trabajo, por apoyar siempre mis ideas.

Dar las gracias a mis profesores del grado en Administración y Dirección de Empresas por sus enseñanzas, no solo del ámbito académico sino también de la vida en general, las cuales me acompañarán siempre.

Por último, agradecer a mi familia y amigos por apoyarme desde el inicio del programa de estudios conjuntos y animarme en todo momento, representando un pilar fundamental para el desarrollo de este trabajo.

RESUMEN

La Inteligencia Artificial y el Machine Learning están cobrando relevancia en la sociedad en los últimos años y son cada vez más utilizadas como herramientas para el tratamiento de datos en las empresas. En este trabajo se ha querido comprobar la importancia y el papel de estas tecnologías en la función de marketing bajo las lentes de la Teoría de Recursos y Capacidades. Para ello, se ha desarrollado una aplicación que, mediante Inteligencia Artificial, es capaz de predecir la puntuación que le daría un consumidor a una botella de vino, siendo capaz de reducir la incertidumbre para el proceso de toma de decisiones. De esta forma, se ha podido estudiar la posible implementación de la aplicación en la empresa y se han recabado las diversas opiniones de varios stakeholders del sector del vino respecto a este tipo de innovaciones que incorporan Inteligencia Artificial, entre ellos bodegas, restaurantes y consumidores particulares.

PALABRAS CLAVE

Inteligencia Artificial, Investigación de Mercados, Teoría de Recursos y Capacidades, Vino

ABSTRACT

Artificial Intelligence and Machine Learning have gained relevance in society in recent years and are increasingly being used as tools for data processing in companies. This study aims to verify the importance and role of these technologies in the marketing function through the lens of the Resource-Based View theory. To achieve this, an application has been developed that utilizes Artificial Intelligence to predict the rating a consumer would give to a bottle of wine, thus reducing uncertainty in the decision-making process. In this way, the potential implementation of the application in the company has been examined, and various opinions from stakeholders in the wine industry, including wineries, restaurants, and individual consumers, regarding these types of Artificial Intelligence-driven innovations have been identified.

KEY WORDS

Artificial Intelligence, Market Research, Resource Based View, Wine

ÍNDICE

1. INTRODUCCIÓN	8
1.1. Objetivos y motivación	8
1.2. Fases y metodología	9
1.3. Estructura del trabajo	9
2. MARCO TEÓRICO: TEORÍA DE RECURSOS Y CAPACIDADES	11
2.1. Recursos y capacidades en marketing	13
2.2. Big Data, Inteligencia Artificial y Machine Learning	14
2.2.1. Inteligencia Artificial en la empresa	15
2.2.2. Inteligencia Artificial en la función de marketing	17
3. CASO PRÁCTICO: APLICACIÓN WINE QUALITY	20
3.1. Aplicación Wine Quality	20
3.2. Demostración del funcionamiento de la aplicación	21
3.2.1. Introducción de los parámetros en la aplicación	21
3.2.2. Prueba de la aplicación	23
3.3. Estudio teórico de la utilidad de la aplicación Wine Quality	26
3.4. Estudio cualitativo: opinión de los diferentes stakeholders del sector vitivinícola	27
3.4.1. Bodegas	28
3.4.2. Dueño de un restaurante	28
3.4.3. Consumidor de vino particular	29
4. CONCLUSIONES Y LÍNEAS FUTURAS	30
4.1. Conclusiones	30
4.2. Líneas futuras de investigación	31
BIBLIOGRAFÍA	33
ANEXOS	36
Anexo I: Análisis estadístico de la base de datos	36

Desarrollo y elaboración de la base de datos	36
Análisis estadístico de la base de datos.....	37
Tratamiento de la base de datos	41
Anexo II: Análisis cualitativo	46

ÍNDICE DE FIGURAS

Figura 2.1: Características que deben presentar los recursos de la empresa para convertirse en fuentes de ventaja competitiva (Barney, 1991).....	12
Figura 2.2: Esquema de la visualización de la actuación empresarial desde el punto de vista del marketing (Morgan, 2012)	13
Figura 3.1: Ventana principal aplicación (fuente: elaboración propia)	21
Figura 3.2: Ventana etiqueta (fuente: elaboración propia)	22
Figura 3.3: Cambio de etiqueta (fuente: elaboración propia).....	23
Figura 3.4: Botella no conocida por la aplicación diseñada obtenida de Vivino (Vivino)	24
Figura 3.5: Prueba de funcionamiento con botella desconocida por la aplicación (fuente: elaboración propia).....	24

ÍNDICE DE TABLAS

Tabla 2.1 Datos sobre el Big Data y la Inteligencia Artificial en las empresas (fuente: elaboración propia).....	16
Tabla 3.1 Métricas de error de la aplicación para los datos de prueba (fuente: elaboración propia).....	25

1. INTRODUCCIÓN

1.1. Objetivos y motivación

A lo largo de este trabajo se pretende mostrar la relevancia que pueden llegar a tener la Inteligencia Artificial y el *Machine Learning* en la empresa y en la función de marketing. Para ilustrar esto, se plantea un caso práctico en forma de aplicación informática que, mediante el uso de herramientas de Inteligencia Artificial, es capaz de predecir la puntuación de la calidad que un consumidor daría a una botella de vino, medida en una escala del 0 al 5.

Cabe destacar que la autora de este trabajo ha cursado el programa de estudios conjuntos en Ingeniería de Telecomunicaciones y Administración y Dirección de Empresas, es por ello por lo que se desarrolló en un primer momento la aplicación como Trabajo Fin de Grado del grado en Telecomunicaciones y, posteriormente, se estudió la posible implementación en la empresa de la herramienta como comunión de ambos grados.

La motivación original de iniciar el desarrollo de este trabajo fue el gran potencial de la Inteligencia Artificial y el *Machine Learning*, pero, sobre todo, la relevancia que están tomando en la actualidad. Son temas que se escuchan en todos los ámbitos de la sociedad, desde la medicina, arqueología, deportes, hasta el mundo empresarial del que se va a hablar.

Posteriormente, cuando surgió el interés por este tema, fue necesario analizar el propósito del proyecto, ya que los usos de la Inteligencia Artificial son muy variados. Resultó interesante, por la gran tierra vinícola que es la comunidad de Castilla y León, hacer una propuesta sobre el tema del vino.

Las aplicaciones existentes sólo ofrecían la puntuación de la botella de vino si esta estaba registrada en sus bases, es por ello por lo que se planteó el desarrollo de esta aplicación. Se consiguió que obtuviera una puntuación de la valoración del usuario sobre la botella sin necesidad de que esta estuviera registrada en ninguna base de datos. Esto es posible gracias al uso de algoritmos de aprendizaje automático o *Machine Learning*, como pueden ser las redes neuronales.

Más tarde, durante la elaboración de la aplicación, se detectó el posible potencial de la herramienta en la empresa y por ello se empezó el estudio que

se desarrolla en este trabajo, planteando un caso práctico para visualizar cómo puede ser un perfecto complemento a la hora de reducir la incertidumbre que envuelve el proceso de toma de decisiones.

1.2. Fases y metodología

En primer lugar, se desarrolló la aplicación, para lo que fue necesario elaborar una extensa base de datos. Después, se desarrolló el código de programación pertinente y se implementaron los diferentes algoritmos de Inteligencia Artificial y *Machine Learning* como son las redes neuronales para obtener la predicción. Una vez finalizada la aplicación y realizadas varias pruebas de funcionamiento, se puso en marcha el estudio de implementación en la empresa.

Este trabajo se observa desde las lentes de la Teoría de Recursos y Capacidades, y pretende mostrar la relevancia que está adquiriendo la Inteligencia Artificial como capacidad para gestionar recursos de la empresa. Por ello, se comienza presentando el marco teórico bajo el que se engloba este trabajo. A continuación, se realizó un análisis cualitativo¹ de varios stakeholders del sector del vino para determinar si la Inteligencia Artificial puede ser una buena herramienta para tratar la información entendida como recurso. A su vez, este estudio sirvió para conocer en profundidad la opinión de los stakeholders a cerca de la aplicación elaborada por la alumna.

1.3. Estructura del trabajo

En este apartado se describen las secciones en las que se divide este trabajo y el contenido de cada una de ellas.

En primer lugar, se desarrolla el marco teórico sobre el que se fundamenta el trabajo. Por ello, se comienza presentando la Teoría de Recursos y Capacidades desarrollada por Barney (1991). A continuación, se relaciona esta teoría con la función de marketing gracias al estudio de Morgan (2012). Posteriormente, se introducen algunos términos relevantes como son Inteligencia Artificial o *Machine Learning* y se comenta la importancia que tienen para la empresa y la función marketing.

A continuación, se plantea un caso práctico para visualizar una posible implementación de la Inteligencia Artificial en la empresa. Para ello, se realiza

¹ Consultar Anexo II para más información sobre el estudio cualitativo realizado

una descripción de la aplicación elaborada por la alumna y se expone el funcionamiento de la herramienta, el cual se mostró también a diversos stakeholders del sector vitivinícola para que conocieran el funcionamiento de la aplicación. Seguidamente, se presenta una argumentación del posible potencial de la herramienta y se finaliza el capítulo con el análisis cualitativo, resumiendo las diferentes opiniones obtenidas de los stakeholders sobre esta.

Por último, se recogen las diversas conclusiones a las que se llega después de haber realizado este trabajo y se exponen las diferentes líneas futuras de investigación, fruto muchas de ellas de las opiniones recogidas sobre la aplicación.

Al final del trabajo, se incluyen dos anexos para complementar la información. El primero de ellos muestra todo el trabajo y análisis realizado de la base de datos, el cual se considera interesante para este trabajo por su contenido estadístico y las conclusiones sacadas de los datos. El segundo recoge el estudio cualitativo realizado a los diversos stakeholders del sector vitivinícola.

2. MARCO TEÓRICO: TEORÍA DE RECURSOS Y CAPACIDADES

La **Teoría de Recursos y Capacidades (TRC)**, elaborada por Barney (1991), desarrolla los conceptos de recurso y capacidad relacionándolos con las estrategias empresariales como posibles formas de alcanzar ventajas competitivas. Se entiende como **recurso** todo aquel activo de una empresa que le permite diseñar e implementar estrategias que mejoren su eficiencia y efectividad. Por otro lado, **capacidad** se refiere a toda habilidad de la empresa que le permite gestionar y utilizar sus recursos para llegar a alcanzar ventajas competitivas (Barney, 1991).

Para que los recursos se conviertan en una verdadera fuente de ventaja competitiva, el modelo VRIO sugiere que deben cumplir una serie de requisitos, concretamente ser valioso, raro, inimitable y difícil de sustituir (Kozlenkova et al., 2014). **Recurso valioso** se refiere a que sólo puede convertirse en una verdadera ventaja competitiva si es valioso para la organización, es decir, si implementa estrategias que mejoren la eficiencia y efectividad de la empresa. **Recurso raro** hace referencia a que, si muchas empresas del sector o entorno pueden poseer y explotar dicho recurso, no podrá convertirse en una ventaja competitiva a priori para una empresa concreta. Si el recurso no es escaso, un gran número de empresas podrá implementar la estrategia en cuestión y no podrán llegar a convertirse en ventajas competitivas. No obstante, resulta complejo medir cuan escaso debe ser un recurso para convertirse verdaderamente en una ventaja. **Recurso inimitable**, requisito relacionado con el anterior, hace referencia a que un recurso debe ser difícil de imitar por el resto de las empresas. **Recurso difícil de sustituir** se refiere a que sea difícil de reemplazar dentro de la empresa debido a, por ejemplo, su alta especialización (Barney, 1991; Arbelo y Pérez, 2001).

En la *Figura 2.1*, se presenta el modelo VRIO, en el que se observan las relaciones de mediación de los recursos VRIO entre los recursos de la empresa y la obtención de una ventaja competitiva sostenible.

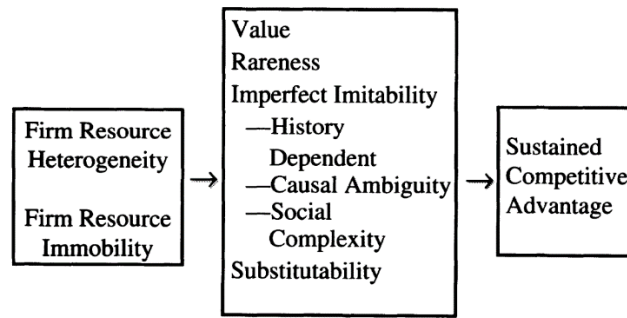


Figura 2.1: Características que deben presentar los recursos de la empresa para convertirse en fuentes de ventaja competitiva (Barney, 1991)

Además, la teoría organizacional ofrece dos enfoques relacionados con la forma en que las empresas organizan sus recursos y desarrollan sus capacidades para alcanzar una ventaja competitiva sostenible. Por un lado, se encuentra la elección estratégica, y por otro, el determinismo contextual (Gopalakrishnan y Dugal, 1998). Desde la elección estratégica se entiende la función directiva como aquella que permite anticipar y prevenir dinámicas externas (Parker et al., 2010). Mientras que desde el determinismo contextual se sugiere que la función directiva se basa en adaptar la organización a las dinámicas externas, centrándose en reaccionar a dichas dinámicas (Bateman y Crant, 1993). La elección estratégica y el determinismo contextual se operacionalizan mediante un enfoque proactivo o reactivo. El enfoque proactivo permite anticiparse a eventos futuros, adelantándose a aquellos cambios o acontecimientos que puedan ocurrir en el entorno de la empresa. El enfoque reactivo permite hacer frente a eventos después de que estos se hayan ocurrido. Se analiza el pasado de la empresa y de su entorno, y con los datos conseguidos se diseñan e implementan estrategias para adaptarse al contexto. Este trabajo se centrará en el desarrollo de un sistema de apoyo a la toma de decisiones basado en inteligencia artificial para dar apoyo a aquellas empresas que han desarrollado un enfoque proactivo, de forma que buscan anticiparse a posibles acontecimientos.

2.1. Recursos y capacidades en marketing

Dentro de los recursos y capacidades con los que cuenta una empresa, esta tiene unos recursos y capacidades específicos de la función de marketing.

En la *Figura 2.2* se presenta el modelo de Morgan (2012), que presenta un enfoque holístico de la Teoría de Recursos y Capacidades y la función de Marketing de la empresa.

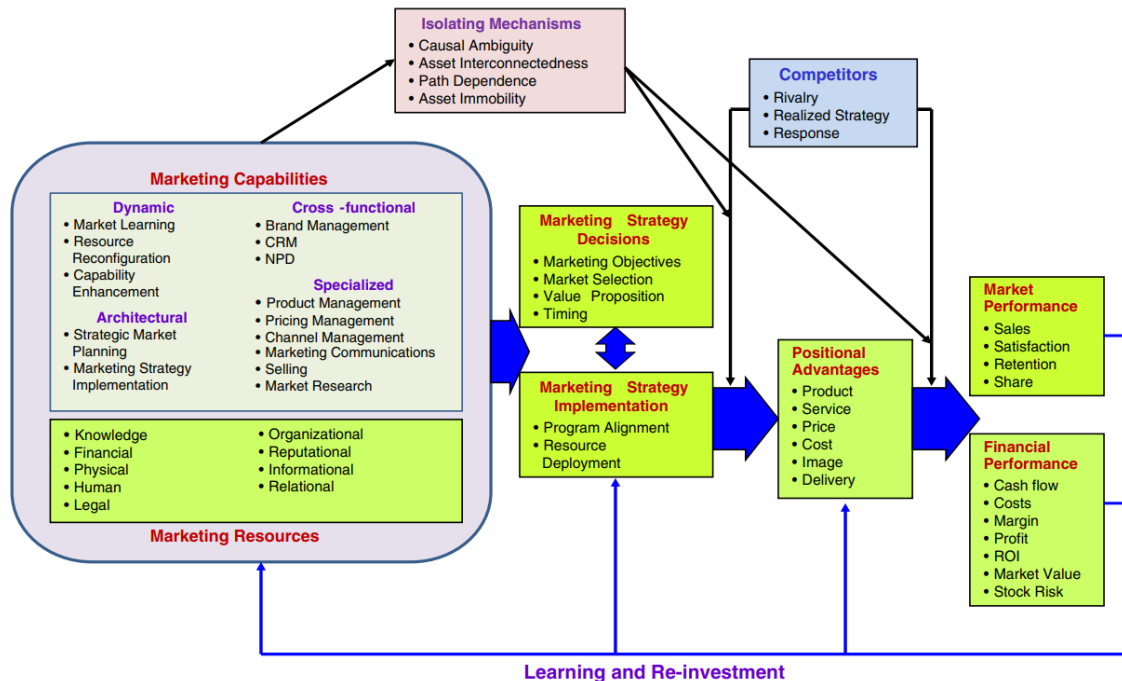


Figura 2.2: Esquema de la visualización de la actuación empresarial desde el punto de vista del marketing (Morgan, 2012)

Entre ellos, en el presente TFG se pone el foco en aquellos recursos que sirven de apoyo para la toma de decisiones. Concretamente, en la **información** como recurso de la empresa. Los datos pueden ser utilizados para tomar mejores decisiones en las empresas y mejorar la eficiencia. Con los datos que obtiene la empresa se pueden identificar patrones o tendencias para conocer mejor a los consumidores o al entorno en general y obtener información valiosa.

No obstante, es necesario que la empresa cuente con un buen sistema de gestión de la información para asegurarse de que esta se recopila, almacena y utiliza de manera efectiva. Aquí es donde entran en juego las **capacidades** de la empresa. Estas capacidades se desarrollan cuando los individuos y los grupos aplican repetidamente sus conocimientos y habilidades para combinar y

transformar los recursos de manera que contribuyan a lograr los objetivos de la empresa.

Además, se incluye un tercer pilar dentro de la función de marketing, que es la estructura, concretamente el **Sistema de Información de Marketing (SIM)**. Kotler en 1966 define el SIM como: *“Un Sistema de Información de Marketing es una estructura permanente e interactiva compuesta por personas, equipos y procedimientos, cuya finalidad es obtener, clasificar, analizar, evaluar y distribuir información pertinente, oportuna y precisa que servirá a quienes toman decisiones de marketing para mejorar la planificación, ejecución y control”*. Gracias al SIM, la empresa puede proporcionar respuestas a problemas específicos y coyunturales en relación con el mercado.

2.2. Big Data, Inteligencia Artificial y Machine Learning

A lo largo de los años, el volumen de datos que manejan las empresas, denominado como **Big Data**, es cada vez mayor. En ocasiones, esta gran cantidad de datos se puede llegar a hacer intratable si no se tienen las herramientas adecuadas, puesto que es fundamental, al igual que se obtienen los datos, saber sacar resultados y conclusiones de estos que ayuden al buen funcionamiento de la empresa. La Inteligencia Artificial y en concreto el Aprendizaje Automático o *Machine Learning* se han vuelto herramientas indispensables en muchas de las grandes empresas a la hora de hacer el tratamiento de datos. Antes que nada, se considera relevante introducir y definir estos conceptos, que serán importantes para el desarrollo del trabajo.

La **Inteligencia Artificial (IA)** es un campo de la informática que se enfoca en desarrollar sistemas y programas que poseen la capacidad de realizar tareas que requieren inteligencia humana. Se basa en la idea de que las máquinas pueden aprender, razonar, tomar decisiones y resolver problemas de manera similar a los seres humanos. La IA involucra la utilización de algoritmos y modelos matemáticos para procesar y analizar grandes volúmenes de datos, identificar patrones, tomar decisiones autónomas y anticipar situaciones. Los enfoques comunes de la IA incluyen el aprendizaje automático (*machine learning*), el procesamiento del lenguaje natural (*natural language processing*), la visión por ordenador (*computer vision*) y la robótica, entre otros. El objetivo de la Inteligencia Artificial es mejorar la eficiencia, la precisión y la

automatización en diversas áreas, como la medicina, la empresa, la atención al cliente y la investigación científica, entre otros (Rouhiainen, 2018). En este trabajo se explorará su capacidad para el Aprendizaje Automático (*Machine Learning*) en el contexto empresarial.

El ***Machine Learning*** es una rama de la Inteligencia Artificial gracias a la cual se consigue automatizar procesos de aprendizaje. Esto tiene numerosas aplicaciones como, por ejemplo, obtener sugerencias o predicciones. La misión de este campo es desarrollar algoritmos capaces de automatizar el proceso de aprendizaje del ordenador y hacer que sea capaz de encontrar patrones en los datos. Es necesario tener en cuenta que, para el buen funcionamiento del *Machine Learning*, hay que alimentar al sistema con grandes cantidades de datos. Los algoritmos de *Machine Learning* consiguen crear modelos que dan valor a los datos y facilitan la toma de decisiones (Maisueche Cuadrado, 2019).

En la *Figura 2.3* se presenta un esquema de la relación existente entre la Inteligencia Artificial y el *Machine Learning*.

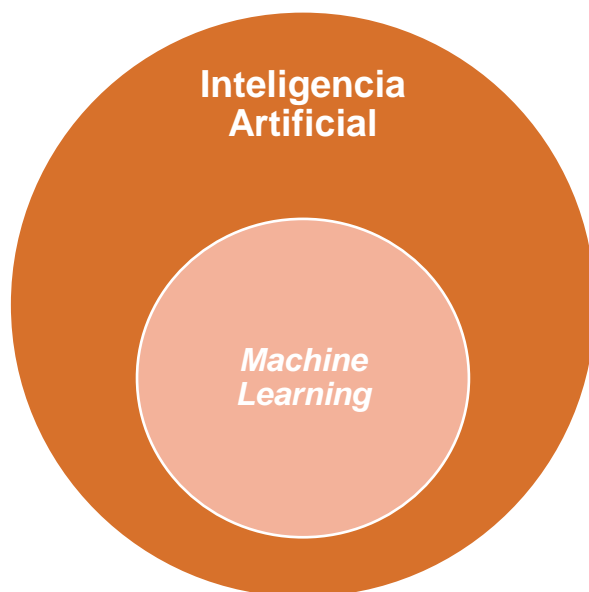


Figura 2.3 Relación entre la Inteligencia Artificial y el Machine Learning (fuente: elaboración propia)

2.2.1. Inteligencia Artificial en la empresa

La implantación de la Inteligencia Artificial en la empresa ha ido evolucionando a lo largo de los años. Al principio, la Inteligencia Artificial se utilizaba para realizar las operaciones rutinarias como, por ejemplo, la gestión de inventarios o facturas, pero dejando al capital humano las actividades que pudieran ser

más complejas y que requirieran de un mayor componente de raciocinio, como tomar decisiones (Rahwan et al., 2022). En la actualidad, se ha evolucionado a modelos de negocio en los cuales la participación de la Inteligencia Artificial en la toma de decisiones es alta, combinándose con el recurso humano (Daugherty y Wilson, 2018). La Inteligencia Artificial puede ser utilizada como capacidad para conseguir ventajas competitivas, ya que otorga a la organización la habilidad de tratar la información y disipar la incertidumbre. Esto es así porque los algoritmos de tratamiento de datos que ofrece esta herramienta permiten anticiparse y predecir resultados futuros, por lo que se cree que puede ser útil para preparar el proceso de toma de decisiones. No obstante, se debe tener en cuenta que con la Inteligencia Artificial no se pretende sustituir el papel del recurso humano en la empresa, sino que con ella se pretende dotar de una herramienta complementaria, suponiendo un apoyo para el trabajador y ampliando sus capacidades (Quinto et al., 2021).

El hecho de que la Inteligencia Artificial está siendo cada vez más relevante en las empresas viene respaldado por el estudio elaborado por NewVantage Partners titulado *“How Big Data and AI are Accelerating Business Transformation”* en 2019. En este estudio se hace una revisión de la adopción del *Big Data* y la Inteligencia Artificial en las organizaciones, habiendo encuestado a 60 empresas de la lista *Fortune1000* (NewVantage Partners, 2019). Se recogen a continuación los datos más relevantes obtenidos en ese estudio:

Tabla 2.1 Datos sobre el Big Data y la Inteligencia Artificial en las empresas (fuente: elaboración propia)

Empresas que consideran necesario invertir en Inteligencia Artificial	87,80%
Empresas con más de \$50MM invertidos en Inteligencia Artificial	55,00%
Empresas que ven resultados tras invertir en Inteligencia Artificial	62,20%
Empresas que han creado un departamento dedicado a los datos	47,60%
Empresas que tienen la figura de un director de datos	67,90%

Con esta información se puede apreciar cuantitativamente el nivel de implantación del *Big Data* y de la Inteligencia Artificial. Esto apunta a que la Inteligencia Artificial está teniendo cada vez más relevancia en las empresas en su proceso de toma de decisiones y que puede constituir en los próximos años una pieza clave en el éxito empresarial, complementando al conocimiento y experiencia que respaldan las decisiones gerenciales.

2.2.2. Inteligencia Artificial en la función de marketing

Una de las tareas principales de la función de marketing es el análisis de las preferencias del consumidor. La empresa cuenta con la información como recurso, y dependerá de sus capacidades, para que sea capaz para generar información que permita reducir la incertidumbre en la toma de decisiones estratégicas de la compañía. Para ello, además debe contar con una estructura adecuada. Así, es necesario disponer de un SIM adecuado para recoger y analizar los datos de cara a obtener información valiosa. El SIM de una empresa cuenta con cuatro subsistemas:

- a) Subsistema de datos internos: gestiona los datos internos de la empresa que se generan en su actividad. Por ejemplo, los datos de compras o de ventas.
- b) Subsistema de investigación de mercados: gestiona los registros de información de momentos concretos en la empresa para resolver problemas puntuales. Este es el caso de la información necesaria para sacar un producto nuevo al mercado o si se lanza una nueva campaña publicitaria, por ejemplo.
- c) Subsistema de inteligencia de marketing: gestiona los datos que provienen del entorno de la empresa. Estos registros pueden ser de la economía política, sociales o de aspectos legales entre otros.
- d) Subsistema de análisis de datos: gestiona el tratamiento tanto cuantitativo como cualitativo de los datos, de cara a encontrar información de utilidad para la empresa y reducir la incertidumbre en el proceso de toma de decisiones.

Los tres primeros subsistemas tienen como papel fundamental la recogida de datos, mientras que el cuarto subsistema tiene como tarea el tratamiento de los datos y la extracción de información relevante (López-Bonilla et al., 2015). La

Inteligencia Artificial puede ser una buena herramienta de gestión de la información para los cuatro subsistemas del SIM por diversos motivos:

- La Inteligencia Artificial realiza un papel fundamental a la hora de realizar la recogida de datos en cualquiera de los subsistemas, ya que se automatizan todos los procesos y son más eficientes (Abashidze y Dąbrowski, 2016).
- En el **subsistema de datos internos y de inteligencia de marketing**, se ha detectado la capacidad de la Inteligencia Artificial de automatizar las tareas repetitivas, como, por ejemplo, el análisis de encuestas, la gestión de inventarios o la automatización de procesos con proveedores (Cepeda y Otálora, 2020).
- En cuanto al **subsistema de investigación de mercados**, la Inteligencia Artificial constituye una capacidad muy valiosa a la hora de ofrecer una experiencia personalizada al cliente, puesto que facilita la segmentación haciendo que esta sea más precisa (Sánchez, 2021). Además, gracias a los algoritmos de la rama de *Machine Learning*, se pueden obtener predicciones a partir de grandes conjuntos de datos.
- Para el **subsistema de análisis de datos** también puede resultar interesante, puesto que ayuda a la depuración de las bases, esto incluye aspectos como eliminar datos duplicados o encontrar valores atípicos. De esta forma, se permite al capital humano centrarse en tareas que generen un mayor valor añadido (Cepeda y Otálora, 2020). Además, la Inteligencia Artificial es una herramienta muy potente para el tratamiento de datos y la identificación de patrones en estos. Gracias a la rama de *Machine Learning* se puede conseguir que el ordenador aprenda de los datos y sea capaz de extraer información valiosa para la empresa. Esto es útil, por ejemplo, a la hora de identificar diferentes tipos de consumidores o patrones de conducta y preferencias (Maisueche Cuadrado, 2019).

En este sentido, se encuentra el estudio realizado por *Salesforce* en 2019, realizado a más de 4100 líderes del sector del marketing de todo el mundo. Los resultados indicaban que el 40% de ellos habían utilizado Inteligencia Artificial en 2018 y que el uso de los profesionales del marketing de esta herramienta

había crecido un 44% desde 2017 (Salesforce, 2019). Esto está en línea con la idea de que la Inteligencia Artificial está cada vez más arraigada en las empresas y tomando peso en la función de marketing.

3. CASO PRÁCTICO: APLICACIÓN WINE QUALITY

En este capítulo se presenta un caso práctico para ilustrar una posible implementación de la Inteligencia Artificial en la empresa y, concretamente, en la función de marketing, mediante una aplicación informática para ordenador. En primer lugar, se realiza una descripción de la aplicación diseñada, la cual es capaz de predecir la valoración que un usuario daría a una botella de vino concreta. En segundo lugar, se presenta una demostración práctica de la herramienta, la cual fue mostrada a varios stakeholders del sector del vino. Más tarde, se valorará su posible utilidad en el SIM y, por último, se expondrán las conclusiones obtenidas del análisis cualitativo realizado a varios stakeholders a los que se cree que les puede aportar utilidad la aplicación.

3.1. Aplicación Wine Quality

La aplicación, denominada *Wine Quality*, tiene la capacidad de predecir la valoración, medida en una nota de 0 a 5, que un consumidor daría a una botella de vino. Es decir, es capaz de anticipar la respuesta de los consumidores ante un determinado producto, sin necesidad de que este lo haya probado o que el producto se haya desarrollado aún. Esto es así porque la aplicación ha sido diseñada con tecnología de Inteligencia Artificial y *Machine Learning*, las cuales mejoran y aprenden con el tiempo y los datos gracias a complejos algoritmos como, por ejemplo, las redes neuronales. La aplicación analiza y compara grandes cantidades de datos de diferentes botellas de vino para identificar patrones y características de forma que el resultado sea un identificador sintético de la valoración de la calidad del vino.

Cuando el usuario introduce ciertas características del vino en la aplicación como, por ejemplo, el precio de la botella o el tipo de uva, la aplicación utiliza las redes neuronales de *Machine Learning* para hacer una predicción de la calidad de la botella. La herramienta tiene en cuenta estos datos introducidos y los compara con los de otros vinos de características similares para hacer su predicción.

Se cree que esta herramienta puede ser útil a la hora de obtener y tratar la información entendida como recurso para aquellas empresas que lo utilicen. Por tanto, para argumentar esto, se realizará en primer lugar un estudio teórico de las implicaciones de la aplicación desarrollada. Posteriormente, se

observarán las ventajas o inconvenientes que se han visualizado en casos reales a la hora de implementar la aplicación en la parte final del capítulo.

3.2. Demostración del funcionamiento de la aplicación

En esta sección se va a explicar el funcionamiento de la aplicación. Esta demostración que aparece a continuación es la que se les presentó a los stakeholders del sector del vino para explicarles la herramienta antes de conocer su opinión. El usuario podrá introducir varias características del vino que se explicarán más adelante y la aplicación le ofrecerá la valoración que otorgaría el usuario a esa botella en una escala de 0 a 5 puntos.

Para comenzar, se muestra en la *Figura 3.1* la ventana principal que aparece cuando el usuario inicia la aplicación.

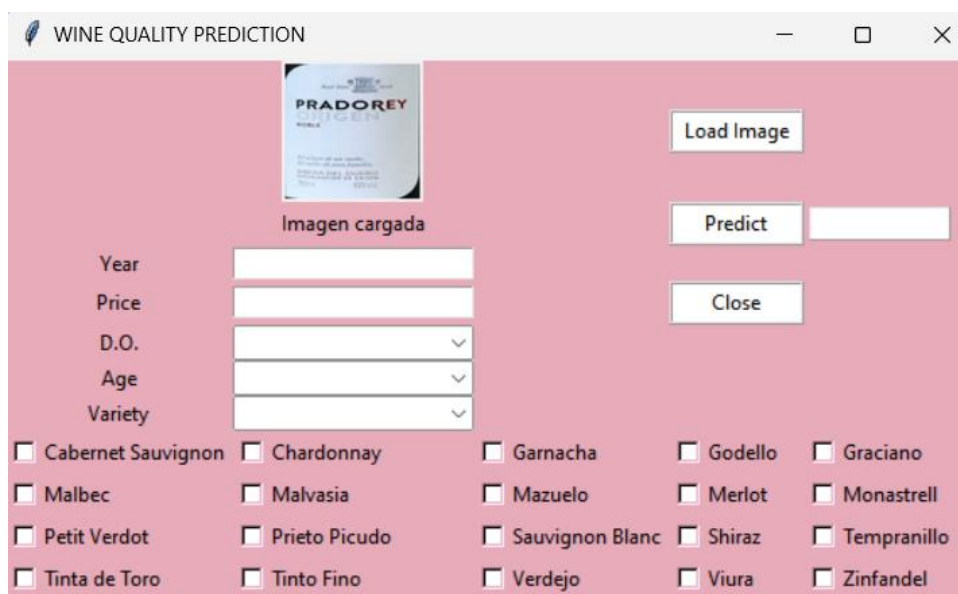


Figura 3.1: Ventana principal aplicación (fuente: elaboración propia)

A continuación, se va a explicar cómo se introducen las diversas características de la botella de vino de la que se quiere saber la calidad.

3.2.1. Introducción de los parámetros en la aplicación

En primer lugar, se presenta una breve explicación de las características de la botella que se pueden introducir en la aplicación, en orden de aparición, en la ventana principal de la *Figura 3.1*:

- Etiqueta (*Label*): se trata de una imagen de la etiqueta de la botella de vino.

- Año (*Year*): hace referencia al año en el que se recolectó la uva con la que se elaboró el vino.
- Precio (*Price*): precio de la botella de vino.
- Denominación de origen (*D.O.*): denominación de origen a la que pertenece el vino.
- Edad (*Age*): hace referencia a la edad del vino (crianza, roble, reserva, etc.).
- Variedad (*Variety*): variedad del vino (tinto, rosado, blanco, etc.).
- Uva (*Grape*): tipo de uva de la que está compuesto el vino. Se puede seleccionar más de un tipo.

Estas características de la botella pueden ser fácilmente introducidas en la aplicación. Para el caso de la etiqueta, se puede seleccionar cualquier imagen que el usuario tenga almacenada en su ordenador. Bastará con pinchar en el botón **Load Image**. Si se presiona dicho botón se despliega una nueva ventana que llevará al usuario a su explorador de archivos para que elija la foto de la etiqueta que desee. Esto se puede observar en la *Figura 3.2*.

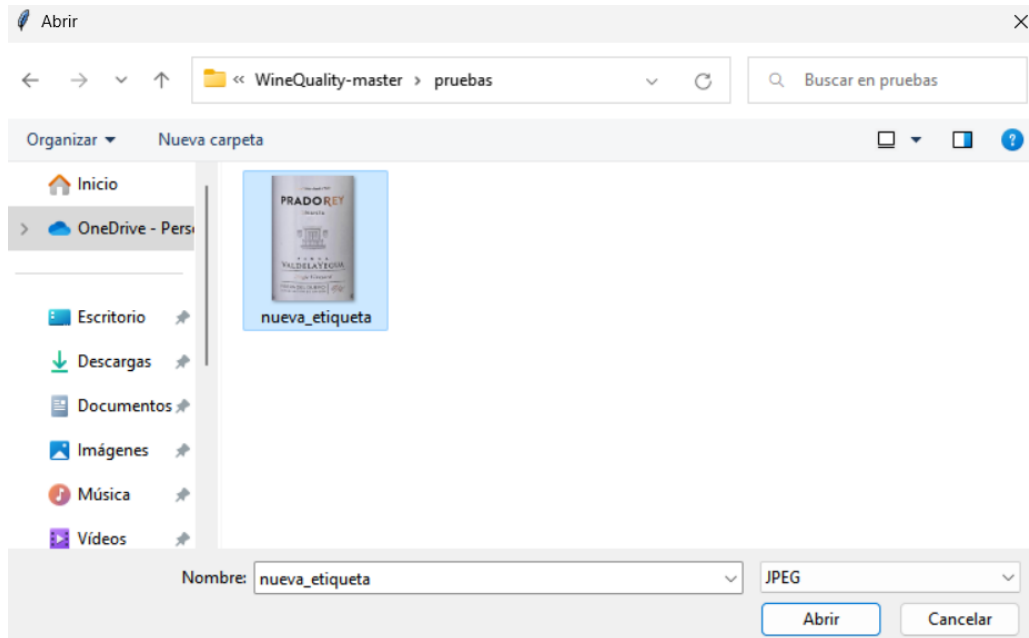


Figura 3.2: Ventana etiqueta (fuente: elaboración propia)

Si se selecciona una nueva etiqueta se podrá observar el cambio de esta en la ventana principal de la aplicación, véase la *Figura 3.3*.

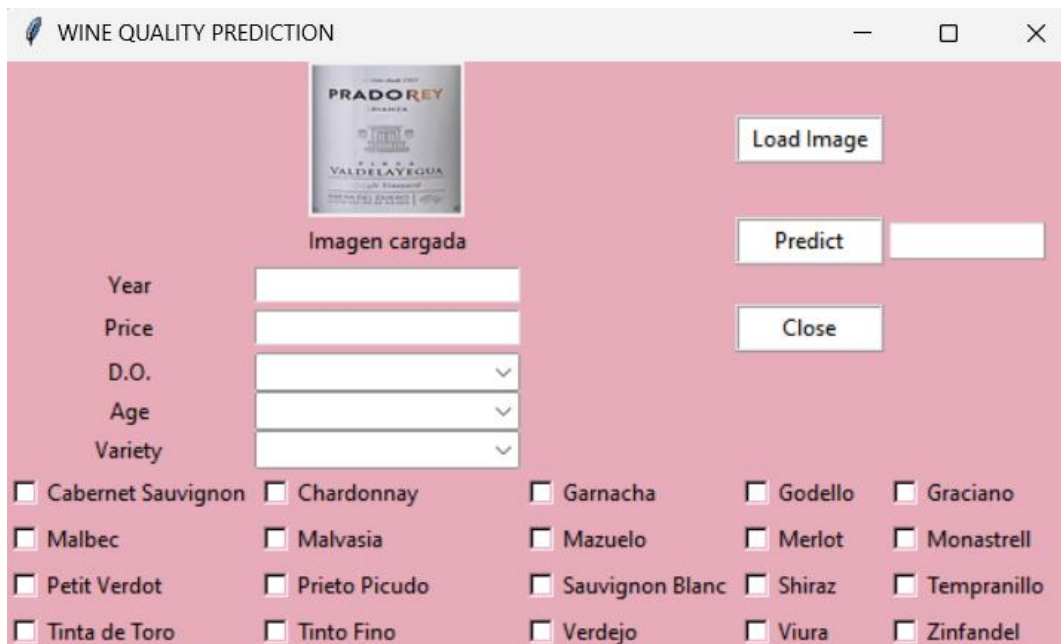


Figura 3.3: Cambio de etiqueta (fuente: elaboración propia)

El resto de los parámetros se pueden introducir fácilmente tecleando la selección en cada casilla o utilizando los desplegados, dependiendo de la característica en cuestión. Para el caso del tipo de uva, se seleccionarán las uvas que desee el usuario marcando las casillas. La aplicación permite la selección de más de un tipo de uva, marcando a la vez varias casillas.

Para obtener la predicción se pulsará el botón **Predict** una vez que se hayan introducido todos los parámetros de la botella de vino. Si se deja alguno de los campos en blanco, la aplicación dará mensaje de error y no ofrecerá la predicción obtenida al usuario. A continuación, se demostrará el funcionamiento de la aplicación.

3.2.2. Prueba de la aplicación

Para probar la aplicación se va a comparar el resultado que se obtendría utilizando la herramienta diseñada por la alumna con el resultado obtenido en la aplicación Vivino que ya existe en el mercado.

La botella escogida para la prueba de la aplicación aparece en la *Figura 3.4*.



Figura 3.4: Botella no conocida por la aplicación diseñada obtenida de Vivino (Vivino)

Se puede apreciar que, para una botella Tomás Postigo 2019 con un precio de 28,85€, con uva Malbec, denominación Ribera de Duero tiene una puntuación de 4,4 sobre 5 puntos en dicha aplicación. Esta calificación se obtiene calculando la media de todas las puntuaciones registradas por los usuarios de esta botella concreta en la aplicación Vivino.

Ahora se van a introducir estas mismas características en la aplicación *Wine Quality* para observar su resultado tal y como aparece en la *Figura 3.5*.

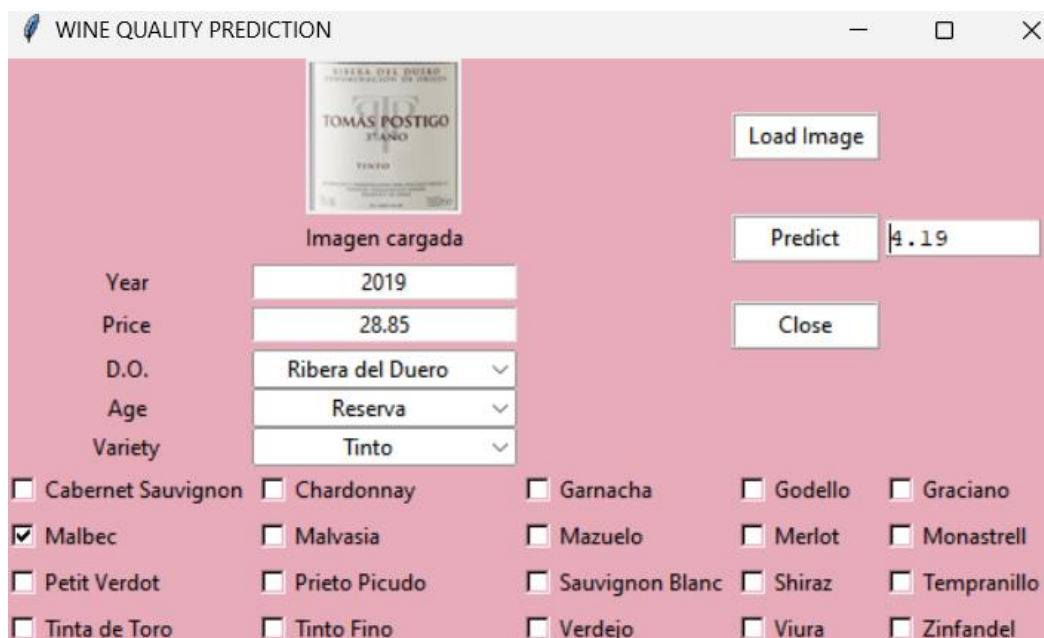


Figura 3.5: Prueba de funcionamiento con botella desconocida por la aplicación (fuente: elaboración propia)

Una vez completada la información, se presiona el botón **Predict** para obtener la predicción de la calidad del vino. En el ejemplo se observa que se obtiene la puntuación concreta de 4,19.

Se puede ver cómo los resultados difieren de los de Vivino, debido a que estos se calculan como la media de las valoraciones de los usuarios y los de la aplicación desarrollada se obtienen a partir de la Inteligencia Artificial de un modelo que utiliza un gran número de muestras y que está aprendiendo continuamente de los datos gracias a las redes neuronales de *Machine Learning*.

Para ver el poder predictor del sistema, se extrajeron varias métricas de error:

Tabla 3.1 Métricas de error de la aplicación para los datos de prueba (fuente: elaboración propia)

MÉTRICA DE ERROR	RESULTADO
Error cuadrático medio	0.012
Raíz del error cuadrático medio	0.11
Coefficiente de determinación	0.6

En la *Tabla 3.1* se pueden apreciar las diferentes métricas de error de la aplicación obtenidas con un conjunto de datos de prueba. En primer lugar, se calculó el error cuadrático medio, obteniéndose un resultado de 0.012 y la raíz de este, siendo igual a 0.11. Para valorar estos resultados, es necesario recordar que la valoración que ofrece la aplicación toma valores entre 0 y 5, por lo que tener un error de centésimas se considera pequeño en comparación con la magnitud de la variable “puntuación de la calidad”. A su vez, se calculó el coeficiente de determinación de la herramienta, obteniéndose un 0.6. Se considera que el poder explicativo del sistema es bueno si se obtienen valores cercanos a 1. En este caso, puede considerarse un valor bajo, pero hay que tener en cuenta que el coeficiente de determinación puede no ser del todo exacto a medida que se van introduciendo un mayor número de variables al sistema, tal y como ha ocurrido en esta aplicación, en la que para hacer la predicción de la puntuación de la calidad se han incluido variables como el precio de la botella, el tipo de uva o el año, sin haber hecho previamente un

análisis del nivel de significación de estas con la variable dependiente “puntuación de la calidad”.

Una vez analizadas las métricas de error se puede afirmar que el poder de predicción de la aplicación diseñada es alto.

3.3. Estudio teórico de la utilidad de la aplicación Wine Quality

La aplicación Wine Quality, al utilizar Inteligencia Artificial, puede ser de gran utilidad para obtener y tratar información que constituya un recurso que cumpla los criterios VRIO, puesto que no todas las empresas tienen aún la capacidad de desarrollar y aplicar las tecnologías de Inteligencia Artificial y ponerlas en marcha en el seno de su organización. Además, el desarrollo de esta aplicación lleva implícito la construcción de una base de datos² sobre botellas de vino y un análisis estadístico de la misma, lo que se considera también un recurso valioso. La implementación de esta aplicación se puede relacionar con la visión proactiva en la toma de decisiones y actuación de una empresa, ya que da la posibilidad de anticiparse, disipar la incertidumbre y disponer de información sobre las preferencias de los consumidores, incluso antes de que haya probado la botella de vino.

Desde el punto de vista de la función de marketing, se piensa que la aplicación puede ayudar en los cuatro subsistemas del SIM. Se recuerda que estos eran: subsistema de datos internos, de investigación de mercados, de inteligencia de marketing y de análisis de datos. La herramienta puede ser un buen complemento para realizar tanto la recogida como el análisis de los datos, pero se cree que sería de especial utilidad dentro del subsistema de investigación de mercados. La aplicación obtiene la predicción de los consumidores acerca de una botella de vino, lo que es muy útil a la hora gestionar la información relativa a momentos concretos en la empresa para solventar problemas puntuales, puesto que se anticipa a las preferencias del cliente y puede complementar el trabajo de una investigación de mercados. Es importante tener en cuenta que, para este tipo de mercados, el producto debe estar finalizado antes de lanzarlo y poder conocer la opinión de los consumidores. La implementación de la aplicación da la posibilidad de ofrecer un servicio innovador y útil para, por

² Consultar Anexo I para más información sobre la base de datos y su tratamiento estadístico

ejemplo, el desarrollo o adquisición de nuevos productos, lo que puede suponer una ventaja competitiva sostenible a largo plazo.

Por todos estos motivos, se ha considerado desde un punto de vista teórico que la aplicación puede llegar a generar ventajas competitivas, siendo una valiosa herramienta para obtener información que constituya un recurso VRIO. Concretamente, se ha detectado que la aplicación puede aportar beneficios a diversos stakeholders del sector del vino, como son bodegas, restaurantes y consumidores particulares.

Las bodegas podrían verse beneficiadas a la hora de desarrollar nuevos tipos de vino, ya que, gracias a la aplicación, pueden conocer la opinión del consumidor antes de tener incluso que plantar las diferentes cepas de la uva.

Los restaurantes, como grandes distribuidores del vino, podrían utilizar la herramienta para guiarse en la compra de vino a las bodegas. Gracias a ella se puede saber la valoración del comensal antes de adquirir grandes cantidades de cajas de vino a una bodega guiándose tan solo de su propia opinión al probarlo o de la de su sumiller.

Por último, los consumidores podrían utilizar la herramienta para decidir qué vino comprar, independientemente de que este se encuentre o no registrado en una base de datos. Hasta ahora, se encontraban con el problema de que, si un vino era muy nuevo o no estaba registrado aún en la base de datos de aplicaciones concretas, no podían consultar la opinión de otros usuarios antes de comprar un vino para determinada ocasión.

3.4. Estudio cualitativo: opinión de los diferentes stakeholders del sector vitivinícola

Mediante la realización de un estudio cualitativo³ se obtuvo información de diversos stakeholders del sector vitivinícola que fue de gran utilidad para contrastar el estudio teórico realizado sobre la utilidad de la aplicación. En este apartado se recogen las diversas opiniones de stakeholders del sector vitivinícola extraídas del estudio. Para su realización se contactó con dos bodegas de Castilla y León, con el dueño de un restaurante que cuenta con un

³ Consultar Anexo II para más información acerca del estudio cualitativo realizado

gran repertorio de vino para sus comensales y con un consumidor de vino particular.

3.4.1. Bodegas

Bodega 1:

El departamento de comunicación y marketing de la bodega 1 opina que podrían llegar a utilizar la herramienta como complemento a una investigación de mercados, más que como un sustitutivo a la misma, puesto que creen que no recoge del todo ciertos comportamientos más sentimentales a la hora de elegir un vino. Por ello, sugieren que sería de gran ayuda poder introducir diferentes variables algo más subjetivas a la aplicación para que aporte mayor valor a la empresa. A la hora de tomar decisiones, comentan que podría aportarles mayor utilidad si pudiera aplicarse a grupos de consumidores concretos, más que a una visión general de los mismos.

Bodega 2:

Tanto el CEO como el director del departamento de tecnología de la bodega 2 están muy interesados en la aplicación. Les parece un recurso muy valioso para los planes de futuro que tienen en la bodega. Tanto es así, que han propuesto realizar una colaboración ofreciendo su base de datos para alimentar a la aplicación y poder obtener más información y predicciones. Como sugerencia, comentaron que estaría bien que la aplicación discriminara según clústeres de clientes, puesto que esa es la nueva línea que van a seguir en la bodega 2 para realizar la segmentación de los clientes.

3.4.2. Dueño de un restaurante

Para el dueño del restaurante la aplicación sería de gran utilidad de cara a ofrecer nuevos productos de calidad a sus comensales. Considera que le aportaría mucho valor poder ver la puntuación que darían los clientes a vinos que no ha probado nunca en el restaurante antes de comprarlos. Sugiere que sería interesante introducir a la aplicación una variable de segmentación que discrimine por zonas geográficas, puesto que considera que los gustos de los consumidores varían según el área. Esta opinión la fundamenta en que tiene restaurantes tanto en Valladolid como en Madrid y los vinos que vende son diferentes en cada uno. También considera que sería valioso incluir una opción

de recomendación de maridaje a la aplicación, de esta forma podría visualizar también los diferentes vinos que encajarían mejor con las comidas que ofrece.

3.4.3. Consumidor de vino particular

El consumidor particular de vino considera que esta herramienta puede ser muy interesante y de gran ayuda en el momento de la compra del vino. En la mayoría de las ocasiones no puede probar el vino antes de comprarlo y considera que la aplicación le ayudaría a tener una estimación de la valoración de la botella. De esta forma, puede comparar o detectar la mejor botella antes de comprarla. Como sugerencia, comenta que sería interesante añadir valores más subjetivos que pueden hacer diferente la calidad de un vino dependiendo del consumidor, como el aroma o la entrada en boca.

4. CONCLUSIONES Y LÍNEAS FUTURAS

4.1. Conclusiones

A lo largo del trabajo, se sigue la línea principal que destaca la relevancia de implementar estrategias en la empresa para utilizar sus capacidades, explotando los recursos existentes y de esta forma alcanzar ventajas competitivas. El estudio se ha centrado en los recursos y capacidades de la función de marketing y cómo la Inteligencia Artificial y la aplicación diseñada pueden constituir una capacidad para obtener información que suponga un recurso VRIO para alcanzar dichas ventajas. Habiendo realizado el estudio, incluyendo el análisis cualitativo, se obtienen las siguientes conclusiones:

- Está incrementando la importancia de la IA en las empresas y cada vez son más las organizaciones que se ayudan de la IA para complementar las decisiones directivas, tanto a nivel general de la empresa como concretamente en el área la función de marketing.
- La implementación en las empresas de herramientas que utilicen IA ayuda a identificar y aprovechar oportunidades del mercado y a utilizar la información recopilada para elaborar nuevos productos que satisfagan las necesidades de los clientes.
- La IA ayuda a las empresas a adaptarse a los cambios del mercado y a las necesidades de los clientes, mejorando su capacidad de competir a largo plazo. Así mismo, la aplicación de la IA es adecuada para obtener información valiosa para las empresas, lo que aumentaría su valor y su capacidad para crear una ventaja competitiva.
- De la implementación concreta de la aplicación, en general, se observa cómo los stakeholders encuentran interés en la herramienta y consideran que puede aportar valor en sus respectivos procesos de toma de decisiones. En empresas como bodegas y restaurantes, consideran que puede mejorar significativamente la toma de decisiones en términos de selección de vinos y la oferta de opciones para los clientes. Estos stakeholders observan que la implementación de esta aplicación puede permitirles destacar en un mercado cada vez más saturado, sobre todo debido a que ven cómo con la herramienta consiguen anticiparse a los gustos y preferencias de los consumidores.

Por otro lado, también consideran que se podrían introducir modificaciones a la aplicación para que esta les aporte mayor beneficio, como, por ejemplo, la segmentación o variables más subjetivas, contemplado todo esto como líneas futuras del trabajo. También se ha detectado que la aplicación aporta valor a consumidores particulares, sobre todo en el momento de la compra.

Como conclusión final, se ha podido ver la utilidad que puede llegar a tener la aplicación en la empresa, que era el principal objetivo de este trabajo. De esta forma, se ha ilustrado mediante un caso práctico que la tecnología y el marketing pueden ir de la mano, siendo buenos complementos el uno del otro.

4.2. Líneas futuras de investigación

La elaboración de este estudio ha sido muy interesante a la hora de plantear líneas futuras en la investigación y en el desarrollo de la aplicación. Gracias a las aportaciones realizadas por los diferentes stakeholders participantes en el estudio cualitativo, se han podido recoger posibles mejoras para la aplicación de cara a incrementar el valor que aporta esta herramienta tanto a la empresa como a particulares. Algunas de estas mejoras pueden ser:

- La implementación de una variable de segmentación, lo que podría ser por ejemplo por comunidad autónoma o incluso por provincia. También la introducción de variables más subjetivas en la aplicación, como podrían ser el sabor, aroma o textura del vino. A su vez, se tiene en cuenta la sugerencia de poder segmentar la predicción según clústeres de clientes. Por último, se podría incorporar también un algoritmo para los diversos maridajes.
- Siguiendo con la investigación realizada para estudiar el valor de la Inteligencia Artificial y la aplicación en la empresa y en la función de marketing, se plantea probar la aplicación de forma empírica. Para ello, sería necesario realizar una investigación de mercados y obtener las valoraciones de varios consumidores de botellas de vino y comparar estos resultados con los que ofrece la aplicación. También sería interesante alimentar la base de datos de la aplicación con los datos que ofrecería la investigación de mercados, de manera que se llegaría a una

total fusión entre la tecnología y la investigación de mercados. Tal y como se expuso anteriormente, una de las bodegas con la que se realizó el estudio cualitativo, ofreció la posibilidad de establecer esta línea colaborativa.

- A su vez, se podría implementar esta tecnología en diferentes áreas y utilizar la aplicación para obtener predicciones en otro tipo de mercados, como por ejemplo para recomendar hoteles o restaurantes según los parámetros de búsqueda de los clientes. También se pueden obtener predicciones de calidad de otros productos como alimentos, adaptando las variables que se introducen en la aplicación.
- Una última línea futura sería realizar un estudio respecto a qué procedimiento tendría más validez a la hora de realizar una investigación y extrapolar los resultados, si la herramienta que aplica Inteligencia Artificial con enormes conjuntos de datos que están constantemente actualizándose o una investigación de mercados convencional con una selección del tamaño de la muestra basado en métodos probabilísticos.

BIBLIOGRAFÍA

- Abashidze, I., y Dąbrowski, M. (2016). Internet of Things in marketing: opportunities and security issues. *Management Systems in Production Engineering*, 217-221.
- Arbelo, A., y Pérez, P. (2001). La reputación empresarial como recurso estratégico: un enfoque de recursos y capacidades. *XI Congreso Nacional de Acede*. Zaragoza.
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 99-120.
- Bateman, T. S., y Crant, J. M. (1993). The proactive component of organizational behavior: A measure and correlates. *Journal of organizational behavior*, 103-118.
- Boyce, C., y Neale, P. (2006). *Conducting In-depth Interviews: A Guide for Designing and Conducting In-depth Interviews for Evaluation Input*. Watertown: Pathfinder International.
- Castro-Morales, L. G., Arias-Collaguazo, W. M., Arciniegas-Paspuel, O. G., y Maldonado-Gudiño, C. W. (2022). La imputación de datos vacíos en las matemáticas y su modelo estadístico. *Revista Conrado*, 18, 272-280.
- Cepeda, M. J., y Otálora, G. (2020). Modernización de la administración de justicia a través de la inteligencia artificial.
- Daugherty, P. R., y Wilson, H. J. (2018). Human+ machine: Reimagining work in the age of AI. *Harvard Business Press*.
- Gopalakrishnan, S., y Dugal, M. (1998). Strategic choice versus environmental determinism: A debate revisited. *The International Journal of Organizational Analysis*.
- Kozlenkova, I. V., Samaha, S. A., y Palmatier, R. W. (2014). Resource-based theory in marketing. *Journal of the academy of marketing science*, 42, 1-21.

- López-Bonilla, J. M., López-Bonilla, L. M., y Peña Vincés, J. D. C. (2015). Sistemas de Información de Marketing: una visión integradora. *Tourism & Management Studies*, 11 (2), 197-203.
- Maisueche Cuadrado, A. (2019). Utilización del Machine Learning en la industria 4.0.
- Morgan, N. (2012). Marketing and business performance. *Journal of the academy of marketing science*, 102-119.
- NewVantagePartners. (2 de Enero de 2019). *Business Wire*. Recuperado el 15 de mayo de 2023, de How Big Data and AI are Accelerating Business Transformation:
<https://www.businesswire.com/news/home/20190102005073/en/NewVantage-Partners-Releases-2019-Big-Data-and-AI-Executive-Survey>
- Okada, S., Ohzeki, M., y Taguchi, S. (2019). Efficient partition of integer optimization problems with one-hot encoding. *Scientific reports*, 9, 13036.
- Parker, S. K., Bindl, U. K., y Strauss, K. (2010). Making things happen: A model of proactive motivation. *Journal of management*, 827-856.
- Quinto, N. M. D., Villodas, A. J. C., Montero, C. P. C., Cueva, D. L. E., y Vera, S. A. N. (2021). La inteligencia artificial y la toma de decisiones gerenciales. *Revista de Investigación Valor Agregado*, 52-69.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., y Wellman, M. (2022). Machine behaviour. *Machine Learning and the City: Applications in Architecture and Urban Design*, 143-166.
- Rodríguez, P., Bautista, M. A., Gonzalez, J., y Escalera, S. (2018). Beyond one-hot encoding: Lower dimensional target embedding. *Image and Vision Computing*, 75, 21-31.
- Rouhiainen, L. (2018). *Inteligencia artificial*. Madrid: Alienta Editorial.
- Roy, B. (10 de julio de 2020). *Datasource.ai*. Recuperado el 10 de febrero de 2023, de Todo Sobre El manejo De Datos Faltantes:

<https://www.datasource.ai/es/data-science-articles/todo-sobre-el-manejo-de-datos-faltantes>

Sánchez, C. A. (2021). Efectos de la inteligencia artificial en las estrategias de marketing: Revisión de literatura. *ADResearch: Revista Internacional de Investigación en Comunicación*, 24, 26-41.

Selesforce. (14 de diciembre de 2019). *Salesforce press and news*. Recuperado el 13 de junio de 2023, de La personalización basada en Inteligencia Artificial y la confianza del cliente, claves para los profesionales del marketing: <https://www.salesforce.com/es/company/news-press/press-releases/2018/12/1812141/>

Vivino. (s.f.). *Tomás Postigo 3er Año Ribera del Duero Tinto*. Recuperado el 17 de mayo de 2023, de Vivino: <https://www.vivino.com/ES/es/toms-postigo-three-er-ano-ribera-del-duero-tinto/w/1187752>

ANEXOS

Anexo I: Análisis estadístico de la base de datos

Desarrollo y elaboración de la base de datos

El primer paso para llevar a cabo la aplicación consistió en la construcción de la base de datos. En esta parte hubo que tomar ciertas decisiones relevantes que determinaron el desarrollo del resto de la aplicación.

En primer lugar, hay que decidir con qué parámetros se quiere trabajar antes de empezar a recoger información. Para ello, se tiene en cuenta lo que se ha comentado anteriormente. La idea principal de la aplicación es que los parámetros que utilice sean sencillos de introducir por el usuario con un simple vistazo a la botella o a su etiqueta. Por tanto, parámetros como el grado de acidez del vino, que son difíciles de conocer, quedan descartados.

Para elaborar la base de datos se llevó a cabo la recogida de estos mediante dos fuentes secundarias:

- a) Consulta y utilización de la base de datos de la aplicación Vivino.
- b) Consulta de varias grandes superficies y vinotecas para la recogida manual de datos de las botellas de vino que allí se vendían.

A continuación, se incluyeron los parámetros que se van a describir y desarrollar. Todos estos parámetros se introducen por cada botella de vino de la base de datos.

- Identificador (ID): se le ha añadido a cada botella un identificador por una cuestión de orden y fácil identificación.
- Nombre (*Name*): correspondiente al nombre de la botella de vino, que en ocasiones suele coincidir con el nombre de la bodega.
- Etiqueta (*Label*): imagen en formato JPG de la etiqueta de cada botella.
- Uva (*Grape*): tipo de uva de la que está compuesta el vino, puede ser más de un tipo.
- Año (*Year*): hace referencia al año en el que se recolectó la uva con la que se elaboró el vino.
- Denominación de origen (*AO*): denominación de origen a la que pertenece la botella de vino.

- Edad (*Age*): hace referencia a la edad del vino (crianza, roble, reserva, etc.).
- Variedad (*Variety*): variedad del vino (tinto, rosado, blanco, etc.).
- Calidad (*Quality*): esta valoración de la calidad del vino se ha obtenido de la aplicación Vivino. Se trata de una puntuación de 0 a 5 que los usuarios de la aplicación otorgaron a cada botella. Aunque este va a ser el parámetro que va a tener como salida la aplicación, es necesario dárselo al sistema para que aprenda de ello.
- Precio (*Price*): precio de cada una de las botellas de vino en euros en el momento de realización de la base de datos.

A continuación, aparece una pequeña muestra de los 5 primeros registros de la base.

ID	Name	Label	Grape	Year	AO	Age	Variety	Quality	Price	
0	1	Casa Castillo	0001.JPG	Monastrell	2006	Jumilla	NaN	Tinto	3.8	7.90
1	2	Casa Castillo	0002.JPG	Monastrell	1996	Jumilla	NaN	Tinto	3.7	26.37
2	3	Casa Castillo	0003.JPG	Monastrell	2010	Jumilla	NaN	Tinto	3.9	6.95
3	4	Tomás Postigo	0004.JPG	Merlot,Tinto Fino,Tempranillo,Cabernet Sauvign...	2017	Ribera del Duero	Reserva	Tinto	4.4	29.16
4	5	PradoRey	0005.JPG	Tempranillo	2019	Ribera del Duero	Roble	Tinto	3.5	7.10

Primeros registros de la base de datos (fuente: elaboración propia)

Al final, después de la recogida de datos, queda una base con un total de 506 botellas de vino. Es necesario puntualizar que no es una base de datos muy extensa comparada con las bases de datos que se suelen utilizar para las aplicaciones de *Machine Learning*, pero a efectos prácticos y académicos, con esta extensión es suficiente para mostrar el funcionamiento de la aplicación.

Para algunos datos la celda está vacía porque no se encontró información de estos. Más adelante se explicará qué se hizo con los datos vacíos.

Análisis estadístico de la base de datos

En esta sección se ha procedido a hacer un análisis estadístico de los datos recogidos y las relaciones entre ellos.

Para comenzar, se presenta el análisis descriptivo de los datos que se denominan de tipo numérico.

	ID	Year	Quality	Price
count	506.000000	506.000000	506.000000	506.000000
mean	253.500000	2013.061265	4.010277	92.228735
std	146.213884	8.770913	0.375346	218.662129
min	1.000000	1964.000000	2.800000	3.450000
25%	127.250000	2012.000000	3.700000	9.900000
50%	253.500000	2016.000000	4.000000	19.385000
75%	379.750000	2018.000000	4.300000	55.865000
max	506.000000	2021.000000	4.700000	1289.440000

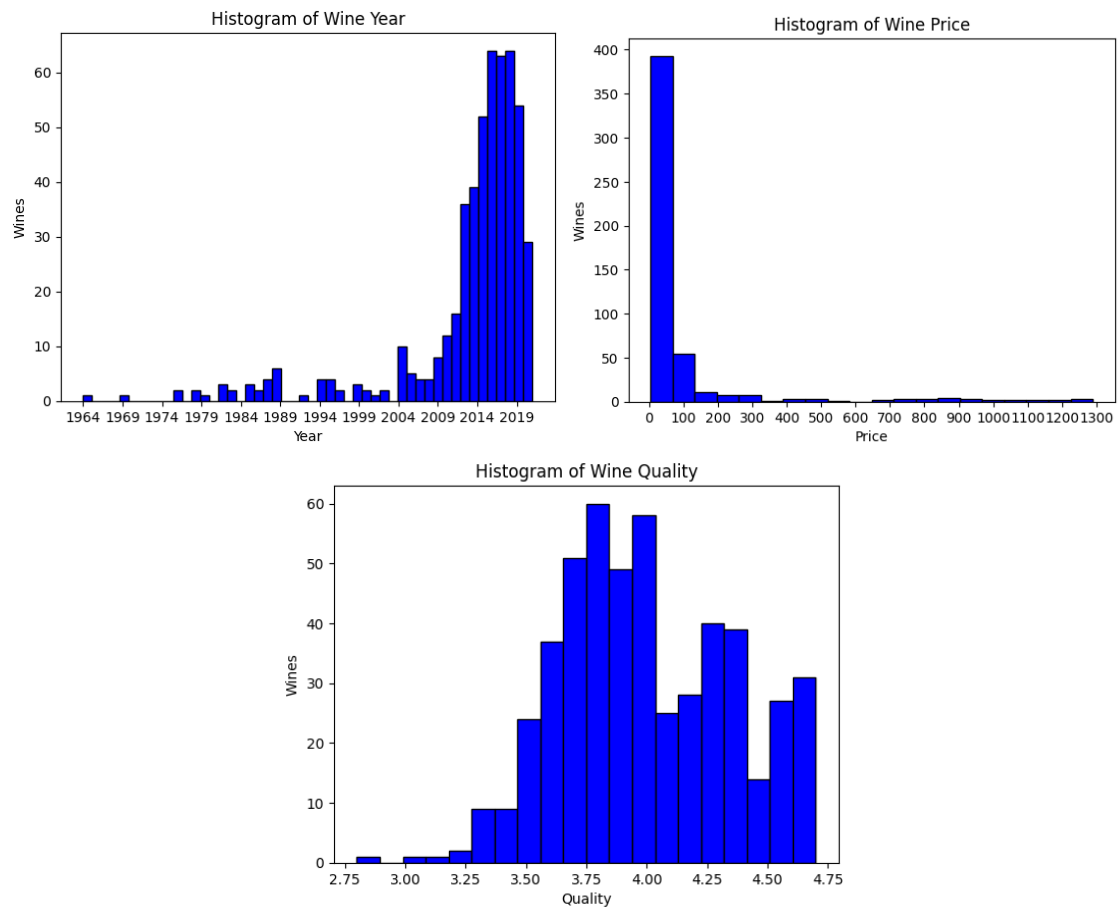
Análisis descriptivo datos numéricos (fuente: elaboración propia)

Aquí se han calculado ciertos parámetros que pueden ser relevantes. La media de los parámetros sirve por ejemplo para visualizar en torno a qué año se distribuyen las botellas o la calidad media que en general los usuarios les dan a las mismas, que se ve que está en torno a un 4. También se puede ver que el precio medio de la botella es de 92€, esto es así puesto que hay botellas en la base de datos con alto precio que hacen que suba la media, pero esto se visualizará más adelante.

A su vez, se puede ver el máximo de cada parámetro, por ejemplo, como se comentaba en el párrafo anterior, es destacable que el máximo de precio alcanza los 1289€.

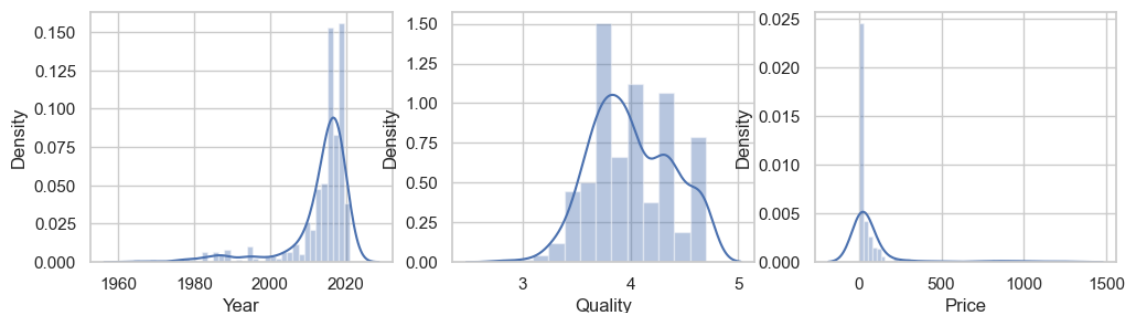
Este análisis permite ver de un vistazo rápido cierta información de los datos la cual se procederá ahora a desarrollar de manera más visual.

A continuación, se presentan los histogramas que muestran las relaciones de los vinos con diferentes parámetros numéricos.



Histogramas datos (fuente: elaboración propia)

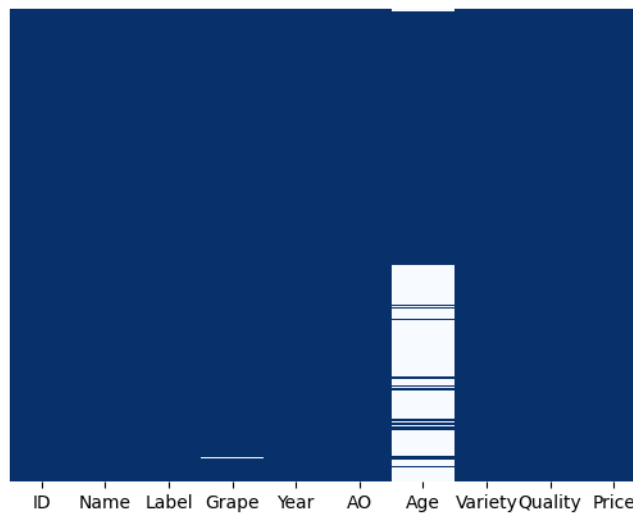
Se puede destacar que los vinos más frecuentes son de los años más recientes. El rango de precios más habituales es entre 1€ y 100€. Las puntuaciones toman valores relativamente altos, concentrándose en las valoraciones más optimistas. A su vez, se presentan a continuación las funciones de densidad de los datos. En la primera gráfica aparece la función de densidad de la variable año, en la segunda la respectiva a calidad y por último la función de densidad de precio.



Funciones densidad datos (fuente: elaboración propia)

Se ha realizado también un estudio de los datos que se encuentran vacíos en la base de datos.

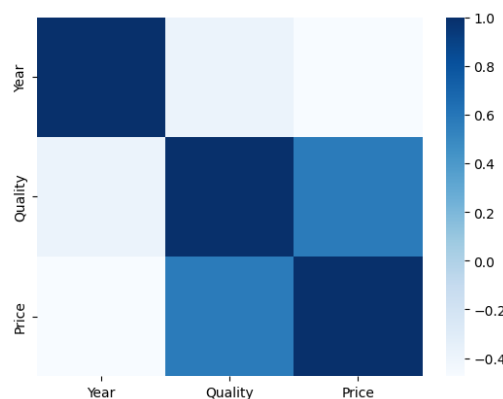
Las variables que tienen datos vacíos son la edad (*Age*) y la uva (*Grape*), siendo un 41,70% el porcentaje de datos vacíos de la primera variable y un 0,20% de la segunda. Esto se puede ver con más detenimiento a continuación.



Relación de parámetros vacíos en la base de datos (fuente: elaboración propia)

En el eje de ordenadas aparecen todas las variables que se pueden encontrar en cada registro de la base de datos. A lo largo del eje de abscisas se observa si existen datos para cada variable concreta del registro. Si hay dato, este aparece en azul y si no lo hay aparece en blanco. De esta forma, se puede ver mejor cómo de completa está en global la base de datos.

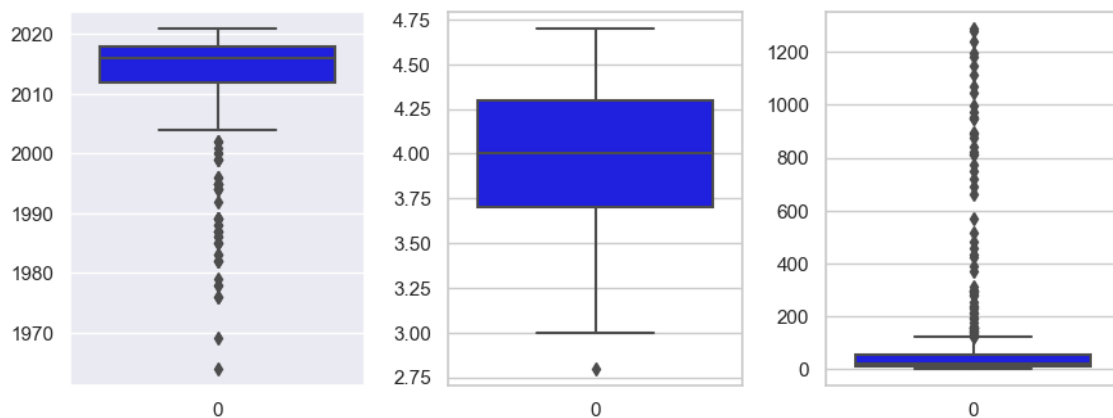
A continuación, se va a presentar la correlación entre los diferentes tipos de datos numéricos (año, calidad y precio).



Correlación entre los parámetros numéricos de la base de datos (fuente: elaboración propia)

Se muestra visualmente la correlación entre los parámetros. Se puede apreciar una relación directa y fuertemente correlacionada entre la calidad del vino y su precio. Además, también existe una relación proporcional entre el año del vino y su precio o calidad.

Seguidamente, se muestran los diagramas de caja de los datos para visualizar la dispersión de estos y los diferentes valores atípicos que aparecen para cada parámetro de la base de datos.



Diagramas de dispersión de los datos (fuente: elaboración propia)

El primer diagrama de caja corresponde a la variable año, el segundo a la calidad y el tercero al precio. En ellos se aprecia una notable dispersión de alguno de los datos en las variables año y precio. La variable año se concentra entre los años 2012 y 2018, mientras que la variable calidad se concentra entre los valores 3.7 y 4.3. Finalmente, la variable precio se concentra entre los valores 10 y 50.

Tratamiento de la base de datos

En esta sección se van a explicar dos tratamientos concretos que hubo que hacer a la base de datos para poder trabajar con ella:

- *One Hot Encoding*

Desarrollando la base de datos surgió un problema para el parámetro tipo de uva (*grape*), puesto que el usuario puede introducir varios tipos de uva para una misma botella y esto hay que codificarlo.

Para el tratamiento de este problema se utilizó el algoritmo de *One Hot Encoding*. Este algoritmo consiste en crear una nueva columna para cada valor

distinto del parámetro que se quiere codificar y marcar con el valor 1 la columna que contenga ese valor y 0 el resto (Rodríguez et al., 2018; Okada et al., 2019).

A continuación, se muestra la base de datos original. Se va a visualizar gráficamente la transformación de esta y así ofrecer una explicación más completa. En la fila número 3 se puede apreciar cómo para una misma botella se introducen varios tipos de uvas.

ID	Name	Label	Grape	Year	AO	Age	Variety	Quality	Price
0	1	Casa Castillo 0001.JPG	Monastrell	2006	Jumilla	NaN	Tinto	3.8	7.90
1	2	Casa Castillo 0002.JPG	Monastrell	1996	Jumilla	NaN	Tinto	3.7	26.37
2	3	Casa Castillo 0003.JPG	Monastrell	2010	Jumilla	NaN	Tinto	3.9	6.95
3	4	Tomás Postigo 0004.JPG	Merlot,Tinto Fino,Tempranillo,Cabernet Sauvign...	2017	Ribera del Duero	Reserva	Tinto	4.4	29.16
4	5	PradoRey 0005.JPG	Tempranillo	2019	Ribera del Duero	Roble	Tinto	3.5	7.10

Primeras líneas de la base de datos sin transformar (fuente: elaboración propia)

Seguidamente, se aplica el algoritmo de *One Hot Encoding*. En este caso se ha utilizado en las variables categóricas que son la uva, la denominación de origen, la añada y la variedad, de esta forma se hace un tratamiento más homogéneo de todos estos datos. Se observa la transformación en la siguiente figura:

ID	Name	Label	Year	Quality	Price	Cabernet Sauvignon	Chardonnay	Garnacha	Godello	...	Crianza	Gran Reserva	Joven	Reserva	Reserva Especial	Roble	Unknown	Blanco	Rosado	Tinto
0	1	Casa Castillo 0001.JPG	2006	3.8	7.90	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
1	2	Casa Castillo 0002.JPG	1996	3.7	26.37	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
2	3	Casa Castillo 0003.JPG	2010	3.9	6.95	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
3	4	Tomás Postigo 0004.JPG	2017	4.4	29.16	1	0	0	0	...	0	0	0	1	0	0	0	0	0	1
4	5	PradoRey 0005.JPG	2019	3.5	7.10	0	0	0	0	...	0	0	0	0	0	1	0	0	0	1

Base de datos transformada mediante el algoritmo One Hot Encoding (fuente: elaboración propia)

De esta forma, será más sencillo codificar más tarde el funcionamiento de la aplicación.

- Tratamiento de datos vacíos

A continuación, se ha estudiado qué hacer con aquellas botellas de vino de las cuales se desconocía información acerca de algún parámetro.

Si se visualizan ahora las primeras líneas de la base de datos se observa que hay ciertas botellas de las cuales no se tiene información para ciertos

parámetros como por ejemplo la edad (*Age*). Esto se denomina *NaN*, del inglés “*Not a Number*”.

Llegados a este punto, es conveniente estudiar qué se va a hacer con estas botellas.

ID	Name	Label	Grape	Year	AO	Age	Variety	Quality	Price	
0	1	Casa Castillo	0001.JPG	Monastrell	2006	Jumilla	NaN	Tinto	3.8	7.90
1	2	Casa Castillo	0002.JPG	Monastrell	1996	Jumilla	NaN	Tinto	3.7	26.37
2	3	Casa Castillo	0003.JPG	Monastrell	2010	Jumilla	NaN	Tinto	3.9	6.95
3	4	Tomás Postigo	0004.JPG	Merlot,Tinto Fino,Tempranillo,Cabernet Sauvign...	2017	Ribera del Duero	Reserva	Tinto	4.4	29.16
4	5	PradoRey	0005.JPG	Tempranillo	2019	Ribera del Duero	Roble	Tinto	3.5	7.10

Primeras líneas de la base de datos donde se aprecian datos vacíos (fuente: elaboración propia)

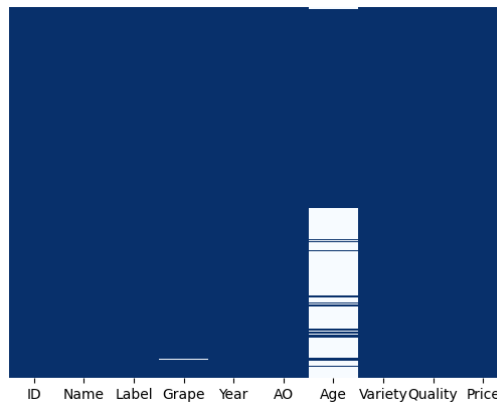
Existen varias alternativas con los datos vacíos de una base de datos. A continuación, se describen brevemente algunas de dichas opciones:

- 1) Descartar datos: uno de los enfoques más utilizados es eliminar el registro en cuestión del que faltan parámetros. En este caso concreto, la solución sería eliminar aquellas botellas de vino para las cuales faltan datos.
- 2) Utilizar media, mediana y moda: mediante esta técnica se sustituyen aquellos datos vacíos por estimaciones estadísticas de los valores. La media, la mediana o la moda pueden ser utilizadas como valor para la sustitución.
- 3) Última observación llevada hacia adelante (*Last Observation Carried Forward*): se sustituye aquel valor que falta mediante el anterior valor observado. Es sencillo de implementar, el problema es que es bastante improbable que el resultado permanezca inalterado en un dato respecto al anterior.
- 4) Siguiendo observación llevada hacia atrás (*Next Observation Carried Back*): se trata de un método equivalente al anterior, consiste en sustituir el dato que falta por el siguiente dato existente observado. Cuenta con los mismos problemas que el anterior.
- 5) Interpolación lineal: se trata de un método puramente matemático que ajusta los datos a una función y extrapola los datos que faltan. El más

simple es la interpolación lineal, pero existen métodos de interpolación más complejos como el cuadrático o el cúbico.

- 6) Añadir una categoría para capturar el *NaN*: este método es de los más utilizados, sobre todo cuando el número de datos faltantes es alto. Consiste en añadir una nueva etiqueta o categoría adicional a la variable para los datos que faltan.
- 7) Imputación de categorías frecuentes: este método consiste en sustituir aquellos datos que faltan por el dato más frecuente de su categoría o variable (Roy, 2020; Castro-Morales et al., 2022).

Para decidir qué método es mejor utilizar, se va a proceder de nuevo a visualizar los datos vacíos.



Visualización de los datos vacíos (fuente: elaboración propia)

En el eje de ordenadas aparecen todas las variables que se pueden encontrar en cada registro de la base de datos. A lo largo del eje de abscisas se observa si existen datos para cada variable concreta del registro. Si hay dato, este aparece en azul y si no lo hay aparece en blanco. De esta forma, se puede ver mejor cómo de completa está en global la base de datos.

Tal y como se había comentado, hay un 41,70% de datos vacíos de la variable *Age* y un 0,20% de la variable *Grape*.

El porcentaje de datos vacíos de la variable *Grape* es muy pequeño (0,20%), en otras palabras, sólo hay una botella de la que no se tiene información de la uva. Por tanto, para este caso, se ha optado por eliminar la botella concreta de la base de datos.

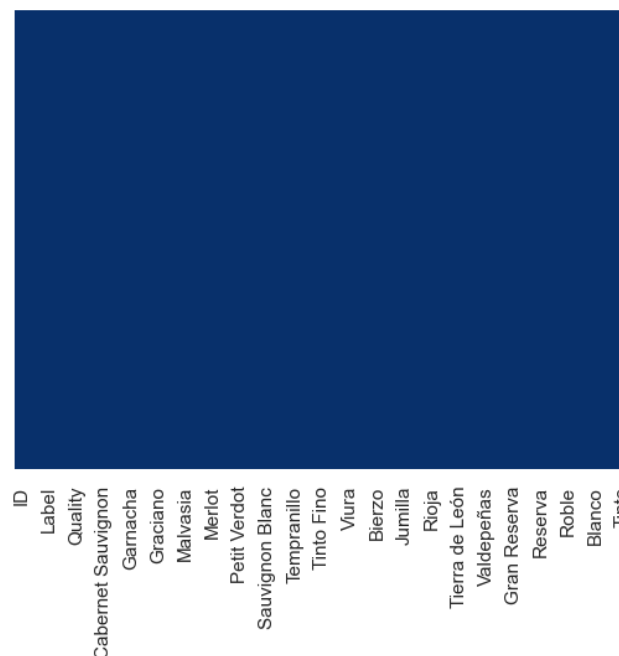
Por otro lado, los datos que faltan de la variable *Age* son mucho más elevados (un 41,7%). Por tanto, aquí resultaría inviable eliminar todas las botellas, puesto que se reduciría mucho el tamaño de la muestra. Por ello, se ha aplicado en este caso el método de añadir una categoría para capturar el *NaN* y se ha creado una nueva categoría que recoge aquellos datos desconocidos.

Se visualiza la base de datos tras aplicar el algoritmo de *One Hot Encoding*, eliminando la botella de la cual falta la uva y añadiendo una nueva categoría *Unknown* (desconocido) para los datos de la variable *Age* desconocidos:

ID	Name	Label	Year	Quality	Price	Cabernet Sauvignon	Chardonnay	Garnacha	Godello	...	Crianza	Gran Reserva	Joven	Reserva	Reserva Especial	Roble	Unknown	Blanco	Rosado	Tinto	
0	1	Casa Castillo	0001.JPG	2006	3.8	7.90	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
1	2	Casa Castillo	0002.JPG	1996	3.7	26.37	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
2	3	Casa Castillo	0003.JPG	2010	3.9	6.95	0	0	0	0	...	0	0	0	0	0	0	1	0	0	1
3	4	Tomás Postigo	0004.JPG	2017	4.4	29.16	1	0	0	0	...	0	0	0	1	0	0	0	0	0	1
4	5	PradoRey	0005.JPG	2019	3.5	7.10	0	0	0	0	...	0	0	0	0	0	1	0	0	0	1

Primeras líneas de la base de datos incluyendo la categoría Unknown (fuente: elaboración propia)

Por último, se aprecia a continuación que ya no queda ningún dato vacío en la base de datos de las variables categóricas, puesto que aparece todo el gráfico en azul, sin ningún espacio en blanco.



Visualización de los datos tratando aquellos que estaban vacíos (fuente: elaboración propia)

Anexo II: Análisis cualitativo

La entrevista en profundidad es una técnica de investigación cualitativa ampliamente utilizada en diversos campos académicos y profesionales. Esta técnica se emplea para obtener información detallada y exhaustiva sobre las experiencias, perspectivas y opiniones de los participantes en relación con un tema específico.

En este trabajo, se ha utilizado la entrevista en profundidad como una herramienta cualitativa para obtener una mejor comprensión sobre la apreciación de los stakeholders del sector del vino sobre la Inteligencia Artificial y la utilización de la herramienta diseñada.

Durante el proceso de investigación, se llevaron a cabo entrevistas en profundidad individualmente con varios stakeholders del sector del vino, concretamente con dos bodegas, el dueño de un restaurante y un consumidor particular. Estas entrevistas se realizaron en un ambiente de confianza y empatía, lo que permitió que los participantes se sintieran cómodos y abiertos al compartir sus pensamientos, experiencias y emociones. A través de preguntas abiertas y no directivas, se les brindó la libertad de expresarse libremente y en sus propias palabras.

La entrevista en profundidad permitió explorar en detalle las diversas opiniones de los encuestados respecto a la aplicación diseñada y obtener una comprensión más profunda sobre la aplicación de la Inteligencia Artificial. A través de esta técnica cualitativa, se pudo capturar no solo la información objetiva, sino también los significados, interpretaciones y matices subjetivos que los participantes atribuyeron a sus respuestas relacionadas con la aplicación. Este enfoque holístico y comprensivo brindó una perspectiva enriquecedora y una visión más completa respecto a la herramienta diseñada y la Inteligencia Artificial.

Una de las principales ventajas de la entrevista en profundidad fue su capacidad para revelar perspectivas y conocimientos profundos que podrían haber pasado desapercibidos en otros métodos de investigación más

estructurados, como por ejemplo la encuesta. Además, la flexibilidad y adaptabilidad inherentes a esta técnica permitieron abordar temas emergentes y explorar áreas de interés específicas en mayor profundidad, como por ejemplo las posibles mejoras o complementos a realizar en la aplicación.

Sin embargo, la utilización de la entrevista en profundidad también presentó desafíos. La recopilación y análisis de los datos de la entrevista requirieron un enfoque minucioso y consumieron mucho tiempo. La transcripción de las grabaciones de las entrevistas y el análisis de los datos cualitativos exigieron atención y rigor para asegurar la fiabilidad y validez de los resultados. Además, la subjetividad y los sesgos tanto del investigador como de los participantes podrían haber influido en los resultados, por lo que se tuvieron en cuenta estrategias para mitigarlos, como la realización de preguntas neutrales y no directivas (Boyce y Neale, 2006).

A continuación, se presentan las preguntas planteadas a los diferentes stakeholders del sector del vino para la realización del estudio cualitativo:

BODEGAS

- a) ¿Qué usos le daría a la aplicación Wine Quality en la bodega?
- b) ¿Cree que sería de utilidad la aplicación Wine Quality en la bodega?
- c) ¿Encuentra algún inconveniente en utilizar la aplicación Wine Quality en la bodega?
- d) ¿Qué cambios o sugerencias consideraría para la aplicación Wine Quality?

RESTAURANTES

- a) ¿Qué usos le daría a la aplicación Wine Quality en el restaurante?
- b) ¿Cree que sería de utilidad la aplicación Wine Quality en el restaurante?
- c) ¿Encuentra algún inconveniente en utilizar la aplicación Wine Quality en el restaurante?
- d) ¿Qué cambios o sugerencias consideraría para la aplicación Wine Quality?

CONSUMIDOR PARTICULAR

- a) ¿Qué usos le daría a la aplicación Wine Quality?
- b) ¿Cree que le sería de utilidad la aplicación Wine Quality en el proceso de compra de una botella de vino?
- c) ¿Encuentra algún inconveniente en utilizar la aplicación Wine Quality?

d) ¿Qué cambios o sugerencias consideraría para la aplicación Wine Quality?