

03-009

**SELECTION OF LEVELS FOR THE PROPERTIES OF INDUSTRIAL PRODUCTS IN THE DESIGN
PHASE USING MACHINE LEARNING TECHNIQUES**

Oviedo Sifontes, Francisco Javier ⁽¹⁾; Zamora-Polo, Francisco ⁽¹⁾; De las Heras García de Vinuesa,
Ana ⁽¹⁾; Luque Sendra, Amalia ⁽¹⁾

⁽¹⁾ Universidad de Sevilla

In the constant race to achieve organizational efficiency, companies have taken on the task of looking for tools that allow the productive improvement of processes within a business value chain. Industrial Product Design Engineering has not been left behind, it has been evolving incorporating technological tools that will allow it to increase its performance, reduce operational waste and aim towards the objectives, satisfying the specific needs of customers. The fusion of traditional processes and cutting-edge technologies, such as the application of Machine Learning Techniques for the prediction of levels in the properties of products in their initial phase. The data obtained for its processing, derived directly from the Kansei Engineering and Neuroscience, which demonstrate the behavior in the neuro-sensory field of the products with respect to their consumers. However, it seeks to create a tool applicable to companies, so that the implementation of these applications is as friendly and efficient as possible for the environment of any Product Development Department.

Keywords: Machine Learning; Kansei Engineering; neurosciences; data; product development

**SELECCIÓN DE NIVELES PARA LAS PROPIEDADES DE PRODUCTOS INDUSTRIALES EN LA FASE
DE DISEÑO MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO**

En la constante carrera para alcanzar la eficiencia organizacional, las empresas se han dado a la tarea de buscar herramientas que permitan el mejoramiento productivo de los procesos dentro de una cadena de valor empresarial. La Ingeniería del Diseño de Productos Industriales, no se ha quedado atrás, ha estado evolucionando incorporando herramientas tecnológicas que permitirán elevar su rendimiento, disminuyendo desperdicios operacionales y apuntando hacia los objetivos, satisfacer las necesidades puntuales de los clientes. La fusión de los procesos tradicionales y las tecnologías de vanguardia, como lo es, la aplicación de Técnicas de Aprendizaje Automático para la predicción de niveles en las propiedades de productos en su fase inicial. Los datos obtenidos para su procesamiento, derivados directamente de los procesos de las Neurociencias e Ingeniería Kansei, que demuestran el comportamiento en el ámbito neuro sensorial de los productos con respecto a sus consumidores. No obstante, se busca crear una herramienta aplicable en las empresas, para que la puesta en marcha de estas aplicaciones sean lo mas amigable y eficiente para el entorno de cualquier Departamento de Desarrollos de Productos.

Palabras clave: Aprendizaje Automático; Ingeniería Kansei; neurociencias; datos; desarrollo de productos.



© 2023 by the authors. Licensee AEIPRO, Spain. This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introducción

La visión y el oído son sentidos naturales que los seres humanos poseen desde su nacimiento. Sus cerebros están diseñados para desarrollar rápidamente habilidades que les permitan soportar estímulos visuales y auditivos, utilizando herramientas que se han desarrollado incluso antes de su nacimiento (Elliot, 2000). Por otro lado, el aprendizaje del lenguaje requiere meses de desarrollo y años para dominarlo. Muchas personas dan por sentado el desarrollo de su visión y audición, pero todos hemos tenidos que entrenar activamente nuestros cerebros para comprender y utilizar el lenguaje.

Curiosamente, el aprendizaje automático sigue una dinámica inversa, ya que se basa en algoritmos de autoaprendizaje que derivan conocimiento a partir de datos para realizar predicciones (Raschka & Mirjalili, 2018). Surge la pregunta, de si todos estos avances tecnológicos han surgido de la nada. Se debe entender que, desde tiempos inmemoriales, los seres vivos han evolucionado con un objetivo primordial: sobrevivir en entornos desafiantes. La inteligencia se ha desarrollado a través de estímulos emocionales, sociales y culturales, impulsando el aprendizaje y la madurez neurosensorial, lo que ha permitido crear a los seres vivos herramientas y productos para facilitar su vida diaria. Por ejemplo, las civilizaciones desde la Edad del Bronce, con su avanzada industria metalúrgica capaz de fundir cobre y estaño para crear bronce.

Sin embargo, hace más de una década, resultó inimaginable para los seres humanos que una cámara digital pudiera extraer información de un objeto al tomar una foto y luego utilizarla para fabricarlo, o que se pudieran extraer características de productos para su rediseño mediante técnicas como la fabricación aditiva. No obstante, la sociedad continúa evolucionando y las industrias van más allá, interactuando con máquinas capaces de predecir los gustos de los usuarios mediante el uso del aprendizaje automático, fusionando así la inteligencia artificial con las neurociencias.

2. Objetivos

La propuesta se basa en un proyecto de investigación que busca diseñar y crear un Manual de Procedimientos para la implementación de técnicas de Machine Learning en la detección de características de productos industriales existentes y su relación con las sensaciones neurosensoriales de los usuarios. Con este fin, la investigación se ha estructurado con los siguientes objetivos específicos: identificar las técnicas más utilizadas en la extracción de propiedades de productos y describir las técnicas de Machine Learning que podrían permitir la predicción de las preferencias de los usuarios mediante el análisis de datos recopilados a partir del historial de uso del producto por parte de los consumidores.

Siguiendo estas directrices, se busca una herramienta que permita programar algoritmos inteligentes, utilizando códigos de Machine Learning, y finalmente se plantea la creación de un documento que reúne procedimientos de fácil interpretación para su aplicación.

3. Metodología

La investigación en cuestión ha seguido un enfoque documental, el cual ha sido dividido en tres secciones. En la primera parte se han examinado y analizado los trabajos previos relacionados con este proyecto y su contribución a la investigación en cuestión. En segundo lugar, se han revisado los enfoques teóricos que se consideran relevantes para el desarrollo de un marco teórico apropiado en este contexto documental. Por último, se ha realizado un estudio exhaustivo de las distintas legislaciones y marcos legales que resultan fundamentales para comprender el problema objeto de investigación.

3.1. Antecedentes de la Investigación

Ruso, C. et al. (2016) en su obra “Tratamiento masivo de datos utilizando técnicas de Aprendizaje Automático”, expresan que, “...ésta es un área de la inteligencia artificial que engloba un conjunto de técnicas que hacen posible el aprendizaje automático a través del entrenamiento con grandes volúmenes de datos o al área en la que se aplica”. En consecuencia, resulta imperativo examinar los proyectos previos a este trabajo con el fin de obtener una comprensión global del campo en el cual podría ser aplicado. A continuación, se presentan los proyectos relevantes que preceden a esta investigación:

En los últimos años se han llevado a cabo investigaciones relacionadas con el desarrollo y aplicación de técnicas de Machine Learning. Un ejemplo es el trabajo de Artacho, G.C. (2021), quien desarrolló técnicas como la Regresión Lineal Múltiple y la Regresión Polinómica de segundo y tercer grado, para la predicción del grado de contagio de COVID-19, en España. En su estudio, se utilizaron mediciones diarias sectorizadas por provincias como datos de entrada. El autor pudo observar resultados significativos a través de este enfoque, como la capacidad de predecir infecciones para el día siguiente y también valoraciones para los siete días siguiente, con definiciones necesarias según las exigencias del algoritmo utilizado.

Artacho, G.C. (2021), “...ha utilizado distintas variables de entrada tales como: temperaturas máximas y mínimas, población, densidad y extensión de la provincia, o el número de pasajeros que se desplazan por medio aéreo en una provincia determinada”. Todo esto con el objetivo de evaluar que parámetros garantizarán resultados más precisos en términos de pronóstico en la cantidad de contagios.

El autor plantea que Python, en combinación con las técnicas de Machine Learning, permite procesar una gran cantidad de conjuntos de datos, también conocidos como Big Data, con el propósito de alcanzar el objetivo de la investigación, que consiste en ayudar a las empresas a obtener información sobre la situación epidemiológica local y así gestionar de manera efectiva la planificación adecuada a su sector.

Tras la implementación del modelo propuesto por el autor, se llega a la conclusión de que el diseño algorítmico se ajusta de manera efectiva a los objetivos planteados en la investigación. Se observa una clara sinergia entre las variables consideradas y los resultados obtenidos al aplicar las técnicas mencionadas con anterioridad. Sin embargo, se destaca que ciertos elementos, como las temperaturas registradas, la población y densidad de la provincia, así como el número de pasajeros transportados por diferentes medios, (públicos o privados), son factores determinantes en los cálculos utilizados para las predicciones de contagios.

Ramos Cueli, J.M., et al (2019), precisan, “...que la recogida de residuos es uno de los servicios más importantes que las ciudades proporcionan. Las posibilidades de optimización a partir de la explotación de los datos recogidos son muy interesantes, si bien, deben considerarse todas las restricciones que aplican”. Por lo que proponen la creación de modelos matemáticos con el fin de desarrollar una herramienta que pueda predecir el nivel de llenado de contenedores en una población específica, con el objetivo de mejorar la planificación logística para la recolección de residuos. En consecuencia, los autores sugieren que el algoritmo a desarrollar debe ser flexible y escalable para su aplicación en diferentes áreas geográficas.

Como resultado, se concluye que, debido a la alta variabilidad de los datos de entrada del algoritmo, es necesario reconsiderar el enfoque y la aplicación de la técnica desarrollada. Según los autores, resulta mas apropiado predecir la tasa de producción de residuos derivada en lugar del nivel de llenado directo de los contenedores.

García, D. (2021), En esta investigación, se destaca la importancia de mantener estándares de calidad en los procesos de producción de elementos metalmecánicos. Debido a la complejidad de estos procesos, se desarrolla una función denominada C(.) utilizando tecnología de Machine Learning, que tiene como objetivo detectar la probabilidad de ocurrencia de defectos en las piezas a fabricar, mejorando así la productividad de las líneas de producción relacionadas.

En otras palabras, el objetivo final, según lo expuesto por García, D. (2021), es mejorar la clasificación de defectos en piezas complejas y generalizar su uso para reducir los tiempos de ajuste o entrenamiento en nuevas piezas con geometrías diferentes, pero de la misma naturaleza que se incorporan en los procesos de producción de los clientes.

En la metalmecánica, la fundición es una de las fases más importantes, ya que genera productos para la industria de la automoción, aeronáutica y la armamentística. Dado el interés económico global de estas organizaciones, se busca evitar defectos y mejorar la calidad y los estándares aplicando técnicas de inteligencia artificial para minimizar casi por completo el error humano.

La exposición de estos trabajos tiene como finalidad resaltar la amplia gama de aplicaciones en las que se puede utilizar el Machine Learning. En un mundo que se está transformando digitalmente, sectores como el financiero, comercial, climático, tecnológico y de telecomunicaciones están adoptando estas tecnologías. En el caso central de esta investigación, se utiliza la Ingeniería de Clasificación (Feature Engineering) como una nueva estrategia para la selección de propiedades y niveles aplicados en la fase de diseño y reingeniería de productos industriales existentes.

Aunque el aprendizaje automático ha existido como campo científico durante décadas, solo unas pocas organizaciones en el mundo han aprovechado su potencial al máximo. A pesar de la disponibilidad de bibliotecas, paquetes y marcos modernos de aprendizaje automático de código abierto respaldados por organizaciones líderes y amplias comunidades de científicos e ingenieros de software, la mayoría de las organizaciones aún enfrentan dificultades al aplicar el aprendizaje automático para resolver problemas comerciales prácticos.

Según Brukov (2020), uno de los desafíos radica en la escasez de talento en este campo. Incluso cuando se cuenta con ingenieros y analistas de datos talentosos en aprendizaje automático, muchas organizaciones aún enfrentan obstáculos en la implementación de modelos. En 2020, la mayoría de las empresas tardaron entre 31 y 90 días en implementar un modelo, mientras que el 18% de las empresas tardaron más de 90 días, e incluso algunas más de un año para poner un modelo en producción. Por lo tanto, los desafíos principales que enfrentan las organizaciones al desarrollar programas de Machine Learning incluyen la disponibilidad de tecnologías de aplicación y el personal con las habilidades necesarias para llevar a cabo los proyectos deseados.

3.2. Marco Teórico

3.2.1. Inteligencia vs. Inteligencia Artificial

Ardila, R. (2011), se hace referencia a que los seres humanos comparten muchas características, pero también tienen diferencias. Existen características universales que son comunes a todos los seres humanos, sin importar su cultura, época, etnia, género o edad y se conoce como naturaleza humana. En general, estos rasgos distintivos permiten categorizar a los seres humanos y unirlos en un solo grupo, basado en el conocimiento: la capacidad mental para aprender, tomar decisiones y dar forma a ideas sobre la realidad. Sin embargo, la definición más precisa y completa también es la más vasta: "...la inteligencia es la capacidad de adaptarse." definido por Piaget, (1963).

La inteligencia ha sido tradicionalmente atribuida exclusivamente a los seres humanos, quienes han buscado destacar sobre los animales y cualquier aspecto que los diferencia de ellos. Sin embargo, según, Mathivet, V. (2018), la definición de inteligencia como la capacidad de adaptarse permite considerar numerosos comportamientos presentes en la vida animal y en otros seres vivos.

Según la afirmación de McCarthy, J. (1956), la capacidad de cualquier sistema para adaptarse y proporcionar respuestas adecuadas a su entorno puede ser considerada como inteligencia. La naturaleza exhibe numerosos casos de inteligencia, tanto biológicos como no biológicos, siendo estos últimos conocidos como inteligencia artificial.

El campo de la inteligencia artificial abarca una amplia gama de técnicas y enfoques. Gracias al aumento constante de la capacidad de cálculo de las computadoras, así como a un mejor entendimiento de los procesos naturales relacionados con la inteligencia y los avances en ciencias fundamentales, se han logrado grandes progresos en esta área. Sin embargo, no todas las capacidades que posee una computadora pueden considerarse como parte de la inteligencia artificial. Por ejemplo, resolver ecuaciones complejas en poco tiempo no se considera un indicador de inteligencia.

Dado esto, resulta difícil definir de manera precisa el término "Inteligencia Artificial". McCarthy, J. (1956), la define como la capacidad de una máquina para dar la impresión de ser inteligente al resolver un problema, ya sea imitando el comportamiento humano o implementando estrategias más flexibles que las permitidas por la programación convencional. Incluso en esta definición se encuentra presente la noción de adaptabilidad.

3.2.1. Aprendizaje Automático

Burkov, A. (2020) define el aprendizaje automático como un subcampo de la inteligencia artificial que construye algoritmos útiles que provienen de la creación por parte humana proveniente de la naturaleza de cualquier fenómeno o derivados de otros desarrollos algorítmicos. Además, se involucran en la resolución de grandes cantidades para dar respuesta a una problemática en sí.

Se utilizan los términos "aprendizaje" y "aprendizaje automático" indistintamente. Por la misma razón, a menudo en la vida diaria se utilizan términos no tan específicos que denotan la complejidad de un algo, por ejemplo, "modelo" para referirse a un modelo estadístico, o para denotar un modelo algorítmico, simplemente hay que situarse en el contexto donde se este trabajando.

3.2.1.1. Aprendizaje Supervisado y No Supervisado

En aprendizaje supervisado, el arquitecto de datos trabaja con bibliotecas de características clasificadas con identificadores $\{(\mu_1, \beta_1), (\mu_2, \beta_2), \dots, (\mu_N, \beta_N)\}$.

Cada elemento μ_i entre N es llamada característica vectorial. En informática, un vector es una matriz unidimensional. Una matriz unidimensional, a su vez, es una secuencia de valores ordenada e indexada. La longitud de esa secuencia de valores, D , se llama el *vector de dimensionalidad*.

Un vector de característica es un vector en el que cada dimensión j desde 1 a D contiene un valor que describe el ejemplo. Cada uno de estos valores se denominan características y se denota como $\mu(j)$. Es decir, si cada etiqueta μ en la colección representa a un individuo, posteriormente cada una de las subclases se orienta a cada dato único de esa persona: inclinación sexual, origen étnico, razón social, entre otros. A su vez la etiqueta β puede ser un elemento que categoriza los conjuntos finitos de clases, como un vector, una matriz o un gráfico.

En resumen, un algoritmo de aprendizaje automático tiene como objetivo el análisis en tiempo real de datos para diseñar un modelo matemático computarizado, que a su vez los utiliza como *input* para generar resultados probabilísticos. Es decir, se crea un modelo que utilice los datos de entrada de una infraestructura en reforma y obtener la probabilidad de que materiales o sistemas son necesarios para dicho mejoramiento.

A diferencia del aprendizaje supervisado, en este tipo de aprendizaje se utilizan datos sin identificar, $\{y_1, y_2, \dots, y_N\}$. A su vez “ y ” es una característica vectorial, donde el objetivo del algoritmo del aprendizaje no supervisado es desarrollar un modelo que utilice la característica vectorial “ y ” como entrada, para luego ser transformado en otro factor que pueda dar soluciones a otros problemas. Un ejemplo práctico, resultaría la implementación de *clustering* o agrupación, el modelo desarrollado devuelve el identificador del grupo para cada etiqueta identificativa del conjunto total inicial. La agrupación en grupo es útil para encontrar grupos de objetos similares en una gran colección de objetos, como imágenes o documentos de texto.

3.2.2. ¿Qué es la ingeniería de aprendizaje automático?

La ingeniería de aprendizaje automático (*Machine Learning Engineering: M.L.E.*), es el uso de principios científicos, herramientas y técnicas de aprendizaje automático e ingeniería de software tradicional para diseñar y construir sistemas informáticos complejos. MLE abarca todas las etapas, desde la recopilación de datos hasta la capacitación del modelo y la puesta a disposición del modelo para que lo use el producto o los clientes.

En la práctica, el aprendizaje automático se implementa como una canalización que contiene etapas encadenadas de transformación de datos, desde la partición de datos hasta la imputación de datos faltantes, el desequilibrio de clases y la reducción de dimensionalidad, hasta el entrenamiento de modelos. los hiperparámetros de datos del canal suelen estar optimizados, todo el canal se puede implementar y usar para predicciones.

El uso del aprendizaje automático en problemas comerciales es apropiado cuando el problema es demasiado complejo para ser abordado mediante codificación tradicional, cuando el problema está en constante cambio, cuando involucra percepciones o fenómenos estudiados, cuando tiene un objetivo simple y cuando es económicamente factible. Sin embargo, hay situaciones en las que el uso del aprendizaje automático puede no ser recomendable: cuando se requiere una explicación clara de los resultados, cuando los errores son inaceptables, cuando la ingeniería de software convencional resulta más económica, cuando todas las entradas y salidas se pueden enumerar y almacenar en una base de datos, y cuando los datos son difíciles de obtener o tienen un costo prohibitivo.

3.2.3. Ingeniería de Características

La diferenciación entre datos sin procesar y características es fundamental y conlleva a tomar decisiones clave en una exitosa ingeniería de características. La ingeniería de características se define como el proceso de representar el dominio de un problema de manera adecuada para las técnicas de aprendizaje. Este proceso implica descubrir inicialmente las funciones relevantes y mejorarlas gradualmente en base al conocimiento del dominio y el rendimiento observado de un algoritmo de machine learning en conjuntos de datos de entrenamiento específicos.

En palabras de Bengio, Y. (2009): “Las buenas características de entrada son esenciales para el éxito del aprendizaje de machine learning. La ingeniería de características representa cerca del 90% del esfuerzo en el aprendizaje automático industrial”. Después de haber definido los términos relacionados con el tema de investigación, es esencial comprender su interacción con los procedimientos y herramientas de extracción de características. Para llevar a cabo un proyecto de extracción de datos, se deben seguir ciertos procesos. En primer lugar, es necesario analizar el problema y evaluar el rendimiento del modelo entrenado. La forma en

que se evalúa el modelo tiene un impacto significativo en la selección final de la Ingeniería de Características. No es recomendable basar la elección de la métrica de evaluación únicamente en las disponibles en las herramientas existentes, ya que muchos grupos de técnicas permiten utilizar sus propias métricas.

3.3. Marco Legal

Lozano, P. (2021), Los avances tecnológicos han tenido un impacto revolucionario en la sociedad actual, convirtiendo a los seres humanos en una fuente constante de generación de datos a través de diversas plataformas en línea, como Instagram, Facebook, WhatsApp, LinkedIn, Twitter, Pinterest, entre otras. Cada día se genera una enorme cantidad de información en estos medios. Estos antecedentes recopilados por las organizaciones se almacenan en repositorios de información digital que, en su conjunto y debido a su complejidad, se conoce como "Big Data".

En el ámbito de la competencia global entre empresas líderes en la fabricación de productos industriales, se ha vuelto crucial aplicar herramientas que permitan extraer características de los productos manufacturados existentes. Esto se logra a través de la ingeniería de características, también conocida como Features Engineering, mediante la aplicación de la reingeniería. Estas prácticas permiten a las empresas ahorrar costos en investigación, desarrollo e innovación.

Las patentes también desempeñan un papel fundamental en este contexto. Las patentes son derechos otorgados por organismos gubernamentales que otorgan al inventor el derecho exclusivo de explotar el producto, servicio o desarrollo en cuestión.

Sin embargo, cuando se trata de extraer características utilizando herramientas tecnológicas derivadas de la Inteligencia Artificial, como el aprendizaje automático o Machine Learning, esta información puede verse afectada por la inseguridad jurídica en cuanto a su protección, así como por los crecientes riesgos derivados de la globalización, la externalización de servicios y el aumento en el uso de herramientas tecnológicas que brindan acceso a la información. Esto contribuye al aumento de casos de plagio y violación de derechos de autor.

La inseguridad jurídica en cuanto a la protección de estos datos se debe, en parte, a la actividad regulatoria reciente de los legisladores y a las diferencias que existen en la protección de cada uno de los componentes involucrados, como el Big Data y los algoritmos.

Lozano, P. (2021) cita: "El 01 de abril de 1998 entró en vigor la Ley 5/1998, de 06 de marzo, de incorporación al Derecho español de la Directiva 96/9/CE del Parlamento Europeo y del Consejo de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos. En sus artículos, 7 a 11, soluciona un vacío existente en todas las legislaciones de los países comunitarios. El Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba la Ley de Propiedad Intelectual y que resultaría, consiguientemente, modificado por la Ley de 5/1998. En la mencionada normativa se prevén dos formas de protección para las bases de datos: el derecho de autor y el derecho "sui generis". Según BOE-A (1998 -5568)".

En relación a las dos formas de protección mencionadas, cuando se hace referencia al derecho de autor, el legislador aplica la protección proporcionada por este derecho a las bases de datos en sí, sin importar su contenido, en base a su selección y organización, que constituye una creación intelectual original del autor. En consecuencia, se protege la forma personal en la que el autor dispone del contenido de la base de datos, pero no el contenido en sí ni tampoco el software que permite su utilización.

4. Resultados y Discusión

4.1. Identificación del grado de madurez de las distintas técnicas de extracción de las propiedades y sus niveles en productos industriales existentes.

Existen herramientas que pueden realizar la extracción de características físicas o digitales, dependiendo del objetivo principal del proyecto en el que se esté trabajando. Estas extracciones se realizan mediante transformaciones de datos que convierten información no procesada en características numéricas que pueden ser analizadas, preservando al mismo tiempo información de la matriz inicial de datos. El propósito de estos procedimientos es complementar la aplicación de algoritmos de aprendizaje automático, lo que resultaría en mejores resultados.

La extracción de características implica identificar y describir patrones relevantes para un problema específico y encontrar una manera de extraerlos. En muchas ocasiones, contar con un buen conocimiento del contexto o dominio puede ayudar a tomar decisiones informadas sobre qué características podrían ser útiles.

En el caso de la extracción automática, se utilizan algoritmos o redes neuronales profundas para extraer características automáticamente de señales o imágenes sin intervención humana. Estas técnicas pueden ser útiles cuando se desea avanzar rápidamente desde datos sin procesar hacia el desarrollo de algoritmos de aprendizaje automático. De manera simplificada, se explica el siguiente:

4.1.1. Recuperación de imágenes utilizando bolsas de características.

La técnica de recuperación de imágenes utilizando bolsas de características en MATLAB consiste en representar las imágenes mediante un conjunto de descriptores visuales, conocidos como características. Estas características capturan información relevante de las imágenes, como texturas, formas o colores, y se agrupan en una estructura llamada bolsa de características.

En el proceso de recuperación de imágenes, se extraen las características de un conjunto de imágenes de referencia y se construye la bolsa de características, que actúa como una representación compacta de las imágenes. Luego, cuando se realiza una consulta de búsqueda de imágenes, se extraen las características de la imagen de consulta y se comparan con las características de la bolsa utilizando medidas de similitud, como la distancia euclidiana o la similitud del coseno. Las imágenes de referencia que sean más similares a la imagen de consulta se consideran resultados relevantes y se devuelven al usuario. En la figura 1, se muestra un elemento objetivo en la ejecución del código del algoritmo en cuestión.

Figura 1: Elemento objetivo a procesar



El plan de este estudio en la búsqueda de similitud en la “*bag of features*” utilizada como “*input*”, con referencia a la característica (color) de la imagen consultada. Una característica de imagen simple basada en el diseño espacial del color es un buen lugar para iniciar el proceso. En la tabla 1, se diseña un código de procesamiento de bolsas de características, que al programarse busca similitudes en todos los datos previamente entrenados y cargados en el sistema.

Tabla 1: Código de programación recuperación de imágenes utilizando bolsas de características

```
% Mostrar una de las imágenes de flores
figure
I = imread(flowerImageSet.Files{1});
imshow(I);
doTraining = false
if doTraining
    %Elija un subconjunto aleatorio de las imágenes de flores.
    trainingSet = splitEachLabel(flowerImageSet, 0.6, 'randomized');
    % Especifique el número de niveles y el factor de ramificación del
    % vocabulario
    % árbol utilizado en bagOfFeatures. Se requiere un análisis empírico para
    % elige valores óptimos.
    numLevels = 1;
    numBranches = 5000;
    % Cree una bolsa personalizada de entidades utilizando la opción
    'CustomExtractor'.
    colorBag = bagOfFeatures(trainingSet,... 'CustomExtractor',
    @exampleBagOfFeaturesColorExtractor,... 'TreeProperties',
    [numLevels numBranches]);
else
    % Cargue un bagOfFeatures previamente entrenado.
    load('savedColorBagOfFeatures.mat','colorBag');
end
% Muestra de resultados
figure
montage(flowerImageSet.Files(imageIDs),'ThumbnailSize',[200 200])
```

En la figura 2, se presentan los resultados obtenidos después de procesar los datos utilizando la técnica de bolsa de características. Se puede observar que, de las 200 imágenes originales, se han extraído aquellas que presentaban similitud en términos de color, que era la característica objetivo en este caso.

Figura 2: Muestra de resultados



En la tabla 2, se presenta una comparación concisa de diversas técnicas de extracción de características en objetos, junto con su campo de aplicación y su relevancia al considerar su implementación.

Tabla 2: Técnicas de extracción de características en objetos: Comparación y campo de aplicación

Técnicas	Diferencias	Importancia y campo de aplicación
Clasificación de dígitos utilizando características HOG	Se enfoca en clasificar dígitos escritos a mano mediante el cálculo de distribución de gradientes en una imagen.	Importante para reconocimiento de caracteres y sistemas de reconocimiento de escritura.
Ensamblaje panorámico de imágenes basado en funciones	Se centra en combinar múltiples imágenes para crear una panorámica mediante la alineación y función de características comunes,	Importante para fotografía panorámica y realidad virtual.
Localización de rotación y escala de imágenes mediante coincidencia automatizada	Se utiliza para identificar la rotación y escala óptimas para imágenes alineales en diferentes orientaciones y tamaños.	Importante para reconocimiento de objetos y su seguimiento cuando este en movimiento.

El desarrollo de productos y servicios en la actualidad requiere un enfoque centrado en las necesidades, expectativas y preferencias de los usuarios. Con el objetivo de crear productos que se ajusten de manera efectiva a las demandas del mercado, es fundamental comprender la importancia de cada propiedad y características en el proceso de diseño y manufactura. Si bien tradicionalmente se han utilizado métodos estadísticos y conceptos de marketing para analizar los requisitos de los clientes, la implementación de nuevas técnicas audaces ha demostrado ser cada vez más beneficiosa en la comprensión de las necesidades reales de los consumidores.

La aplicación de técnicas como la Ingeniería Kansei, las neurociencias y la antropología sensorial ha brindado un enfoque más profundo en la comprensión de las preferencias y emociones de los usuarios. Estas disciplinas permiten explorar los aspectos subjetivos y emocionales asociados a la interacción con productos y servicios, ayudando a evitar el desperdicio o la improductividad en el diseño de nuevos productos. Al estudiar la respuesta neurofisiológica y emocional de los usuarios, se obtiene información valiosa para diseñar experiencias más satisfactorias y personalizadas.

En este contexto, los arquitectos de datos desempeñan un papel crucial. Son responsables de la recopilación y análisis de información relevante para elaborar modelos de aprendizaje automático capaces de predecir las características y preferencias interpretadas de los usuarios. Estos modelos proporcionan una base sólida para que los desarrolladores e ingenieros de productos puedan crear diseños iniciales que se alineen con las expectativas de los consumidores.

La sinergia entre las neurociencias y las técnicas de aprendizaje automático (ML) es especialmente poderosa. Al combinar la comprensión de los procesos cognitivos y emocionales con la capacidad de procesamiento y análisis de grandes cantidades de datos que ofrece el aprendizaje automático, se pueden obtener conocimientos profundos sobre el comportamiento humano en la elección de productos. Esto permite a los departamentos de

marketing y producción de las industrias tomar decisiones más informadas y desarrollar estrategias más efectivas para satisfacer las necesidades y preferencias de los clientes.

4.1.2. Describir las técnicas de Machine Learning para la predicción de preferencias de los usuarios con respecto a las características y propiedades en el diseño de los productos.

El propósito de este objetivo es analizar las técnicas de Machine Learning aplicadas en la predicción de las preferencias de los usuarios en relación con las características y propiedades de los productos durante su proceso de diseño. Estas técnicas permiten recopilar y analizar datos relevantes sobre las preferencias de los usuarios, lo que ayuda a comprender mejor sus necesidades y expectativas. Al utilizar el Machine Learning de manera efectiva, se puede generar información valiosa para orientar el diseño de productos sostenibles, que sean tanto deseables para los usuarios como respetuosos con el medio ambiente.

La tabla 3, muestra una comparativa de aplicativos para el desarrollo de códigos de programación en técnicas de análisis de datos.

Tabla 3: Comparación de Aplicativos para el Desarrollo de Códigos de Programación en Técnicas de Análisis de Datos

Técnica Machine Learning	Descripción	Ventajas	Desventajas
Regresión Lineal	Método que busca establecer una relación lineal entre variables para predecir preferencias de usuarios.	<ul style="list-style-type: none"> Fácil interpretación de resultados. Eficiente para problemas con pocos atributos. 	<ul style="list-style-type: none"> Limitado en la capacidad de modelar relaciones no lineales. Sensible a valores atípicos.
Árboles de decisión	Estructura de decisión basadas en reglas que dividen los datos en ramas, permitiendo predecir preferencias.	<ul style="list-style-type: none"> Fácil interpretación y visualización. Capacidad de manejar atributos categóricos y numéricos. 	<ul style="list-style-type: none"> Propensos a "overfitting" si no se controla la profundidad del árbol. No es adecuado para problemas con alta dimensionalidad.
Métodos de Conjunto	Combinación de múltiples modelos de aprendizaje para predicción más robusta y precisa.	<ul style="list-style-type: none"> Mejora la generalización y reduce el sobreajuste. Capacidad de manejar datos desequilibrados 	<ul style="list-style-type: none"> Mayor capacidad y consumo de recursos computacionales. Difícil interpretación comparado con modelos individuales.

Estas técnicas de Machine Learning son fundamentales para comprender las preferencias de los usuarios en el diseño de productos y servicios. Permiten analizar datos de manera objetiva y precisa, identificar patrones ocultos y realizar predicciones que pueden influir en la toma de decisiones durante el proceso de diseño y desarrollo de productos.

4.1.3. Analizar las aplicaciones en las que se han utilizado técnicas de machine learning para la predicción de las preferencias de los usuarios a las variables de diseño de un producto.

La tabla 4, proporciona una visión general de los diferentes aplicativos utilizados en el desarrollo de códigos de programación para técnicas específicas. Estos aplicativos desempeñan un papel crucial en la implementación y ejecución de algoritmos y modelos de aprendizaje automático, permitiendo a los desarrolladores trabajar de manera eficiente y efectiva en la aplicación de las técnicas descritas. A continuación, se presentan seis variables clave que se han tenido en cuenta para evaluar los aplicativos, brindando información relevante sobre su funcionalidad, versatilidad y usabilidad en el contexto de desarrollo de códigos para las técnicas mencionadas. Esto ayudará a los profesionales e investigadores a tomar decisiones informadas al seleccionar la herramienta adecuada para proyectos específicos.

Tabla 4: Comparación de aplicativos para el desarrollo de códigos de programación en técnicas de análisis de datos

Aplic.	L.P.	Ventajas	Desventajas	Flexibilidad	Rendimiento	Comunidad
Python	Phyton	Amplia comunidad de usuarios y soporte. Gran cantidad de bibliotecas y paquetes para el aprendizaje automático. Fácil sintaxis y legibilidad del código.	Puede ser más lento en comparación con otros lenguajes de programación. Requiere la instalación de diferentes paquetes y dependencias. No es tan eficiente en tareas de procesamiento intensivo.	Alta	Moderado	Alta
MatLab	MatLab	Amplio conjunto de funciones y herramientas específicas para el procesamiento de señales e imágenes. Interfaz gráfica amigable para el desarrollo y la visualización de resultados.	No es un software libre y su licencia puede ser costosa. Algunas funciones y paquetes avanzados pueden requerir conocimientos adicionales.	Moderada	Moderado	Moderada
Java	Java	Alto rendimiento y escalabilidad. Amplia compatibilidad y portabilidad en diferentes plataformas.	Requiere conocimientos mas avanzados de programación en comparación con otros. La	Alta	Alto	Moderado

		Puede ser utilizada tanto para aplicaciones web como de escritorio. Buen soporte y documentación.	sintaxis puede ser mas verbosa y compleja. Menos orientado al aprendizaje automático en comparación con Python, R o MatLab.			
--	--	---	--	--	--	--

En resumen, la elección del aplicativo adecuado dependerá de las necesidades específicas del proyecto, el dominio de aplicación y las preferencias personales del usuario. Es importante considerar factores como la facilidad de uso, la disponibilidad de bibliotecas y recursos, el rendimiento y las capacidades requeridas para tomar una decisión informada.

5. Conclusiones

Después de analizar los conceptos, técnicas y aplicativos relacionados con la investigación Selección de Niveles para las Propiedades de Productos en la fase de diseño mediante técnicas de Machine Learning, se pueden establecer las siguientes conclusiones:

La aplicación de técnicas de Machine Learning en el diseño de productos es una herramienta poderosa para comprender las preferencias de los usuarios y mejorar la satisfacción del cliente. Estas técnicas permiten recopilar y analizar grandes cantidades de datos para identificar patrones y tendencias en las preferencias de los usuarios, lo que facilita la toma de decisiones informadas en el proceso de diseño.

La ingeniería de características desempeña un papel fundamental en la predicción de preferencias de los usuarios. Mediante la extracción y selección adecuada de características relevantes de los productos, es posible crear modelos de Machine Learning más precisos y efectivos. Esto implica la necesidad de un análisis exhaustivo de los datos y una comprensión profunda del dominio del problema.

La integración de las neurociencias y la psicología del consumidor en el análisis de datos permite un enfoque más completo y profundo en la comprensión de las preferencias de los usuarios. Al considerar los aspectos cognitivos y emocionales de los consumidores, se pueden diseñar productos que generen una mayor satisfacción y conexión emocional con los usuarios.

La sostenibilidad juega un papel importante en el diseño de productos. La incorporación de técnicas de Machine Learning en la predicción de preferencias de los usuarios puede contribuir a la creación de productos más sostenibles, que sean tanto deseables para los usuarios como respetuosos con el medio ambiente. Esto implica la consideración de criterios de ecoeficiencia, ciclo de vida del producto y reducción de impactos ambientales.

6. Referencias

- Ardila, R. (2011). El mundo de la psicología / Obras selectas. . *Revista Latinoamericana de Psicología*, 396.
- Artacho Gomez, C. (2021). Trabajo de Fin de Grado. *Desarrollo y aplicación de técnicas de Machine Learning para la predicción de contagios por Covid-19*. Sevilla, España: Universidad de Sevilla.
- Bengio, Y. (2009). *Learning Deep Architectures fot Artificial Intelligence*. S/I: Now Publishers.

- Burkov, A. (2020). *Machine Learning Engineering*. S/I: Google Cloud Platform.
- Burkov, A. (2020). *The Hundred-Page Machine Learning Book*.
- Elliot, J. (2000). *La Investigación acción en educación*. Ediciones Morata
- García, D. (2021). Trabajo de Fin de Máster. *Diseño e Implementación de Técnicas de Machine Learning para la detección de defectos superficiales en piezas sometidas a procesos de estampado o fundición*. Oviedo, España: Universidad de Oviedo.
- Lozano, P. (2019). Trabajo de Fin de Grado. *Mecanismos Jurídicos para la Protección de Soluciones de Machine Learning*. Zaragoza, España: Universidad de Zaragoza.
- MacCarthy, J. (1956). *Automata Studies*. Princeton, New Jersey: Preinceton University Press.
- Mathivet, V. (2018). *Inteligencia Artificial para Desarrolladores - Conceptos e implementación en C#*. Barcelona: Ediciones ENI.
- Piaget, J. (1963). *The Origen of Intelligence in Children*. Michigan: International Universities Press.
- Ramos Cueli, J. M., Luque Sendra, A., Larios Martín, D. F., & Barbancho Concejero, J. (2019). Avances en la investigación en ciencia e ingeniería. *Machine learning para la optimización inteligente de la recogida de residuos*. Alicante, España: Editorial Área de Innovación y Desarrollo, S.L.
- Raschka, S., & Mirjalili, V. (2018). *Python: Machine Learning*. España: Marcombo.
- Russo, C., Ramon, H., Alonso, N., Cicerchia, B., Esnaola, L., & Tessore, J. (14 de 04 de 2016). *Tratamiento masivo de datos utilizando técnicas de aprendizaje automático*. Obtenido de Universidad Nacional de Entre Ríos: <http://repositorio.unnoba.edu.ar:8080/xmlui/handle/23601/107>

7. Comunicación alineada con los Objetivos de Desarrollo Sostenible



1. Consumo y producción responsables: El objetivo es desarrollar técnicas de análisis de datos que permitan comprender las preferencias de los usuarios y diseñar productos que sean sostenibles y respetuosos con el medio ambiente. Al tener en cuenta las características y propiedades deseadas por los usuarios, se pueden evitar productos innecesarios y reducir el consumo excesivo, fomentando un enfoque de consumo más responsable.

2. Acción por el clima: Al utilizar técnicas de Machine Learning para predecir las preferencias de los usuarios y diseñar productos sostenibles, se contribuye a mitigar los impactos ambientales. Al desarrollar productos que se ajusten a las necesidades de los usuarios y sean respetuosos con el medio ambiente, se reducen las emisiones de gases de efecto invernadero y se promueve la transición hacia una economía baja en carbono. Esto apoya los esfuerzos para abordar el cambio climático y lograr una acción sostenible para el clima.