



Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

# **Aplicación de Redes Neuronales para la Predicción de Vacantes por Industria en el Mercado Laboral Español**

Autor: Lucía Nárdiz Izuzquiza

Director: María José Martín Rodrigo

MADRID | JUNIO 2023

## **Resumen**

El mercado laboral es uno de los factores más relevantes en la economía y desarrollo de un país, especialmente cuando se trata de los jóvenes que se incorporan a él. En este contexto, el uso del Big Data se ha vuelto cada vez más relevante, ya que brinda la posibilidad de obtener conocimiento de los datos pasados y obtener predicciones para el futuro próximo. A través de un análisis de la situación actual del mercado laboral, y en especial el de los jóvenes, se ha comprobado el estado en el que se encuentra, sus peculiaridades y principales amenazas.

Este trabajo se apoya de un modelo de redes neuronales artificiales que brinda la posibilidad de llevar a cabo predicciones sobre el número de vacantes para cada uno de los grupos de actividades considerados por el Instituto Nacional de Estadística. Los resultados se muestran optimistas con el crecimiento de la oferta de trabajo, especialmente en los sectores de las Industrias extractivas, las de Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación. Se recomienda a las instituciones públicas facilitar este tipo de información, que ayudará a los jóvenes a orientar su formación hacia un sector con amplia oferta de empleo.

**Palabras clave:** redes neuronales, aprendizaje supervisado, mercado laboral, empleo juvenil, vacantes.

## **Abstract**

The labor market is one of the most relevant factors in the economy and development of a country, especially when it comes to young people entering it. In this context, the use of Big Data has become increasingly relevant, as it provides the possibility of gaining knowledge from past data and obtaining future predictions. Through an analysis of the current situation of the labor market, and especially that of young people, the state in which it is, its peculiarities and main threats have been verified.

This work is supported by a model of artificial neural networks that provides the possibility of carrying out predictions on the number of vacancies for each of the groups of activities considered by the National Institute of Statistics. The results are optimistic about the growth of the labor supply, especially in the sectors of extractive industries, water supply, sanitation activities, waste management and decontamination. It is

recommended that public institutions provide this type of information, which will help young people to orient their training towards a sector with a wide range of employment opportunities.

**Keywords:** neural networks, supervised learning, labor market, youth employment, vacancies.

## Índice

1. Introducción.....	6
1.1. Objetivo de estudio .....	6
1.2. Metodología.....	7
1.3. Estructura .....	8
2. Estado de la cuestión .....	8
3. Descripción del mercado laboral juvenil en España.....	9
3.1. Situación actual y características .....	9
3.1.1 Análisis estadístico y tendencias de los principales indicadores del empleo en la actualidad .....	11
3.1.2 Detalle de las diferencias sectoriales .....	13
3.2 Comparación con la situación del empleo juvenil en Europa .....	19
4. Factores que dificultan la inserción de los jóvenes al mercado laboral.....	20
5. Principales indicadores y predictores del mercado laboral .....	23
6. Explicación y nociones del modelo .....	26
6.1. Nociones sobre modelos de redes neuronales.....	26
6.2. Selección y descripción de variables .....	28
7. Metodología.....	31
7.1. Descarga de los datos.....	31
7.2. R Studio y tratamiento de los datos .....	31
7.3. Análisis exploratorio.....	33
7.4. Modelo de redes neuronales .....	35
7.5. Predicciones .....	39
8. Conclusiones y recomendaciones.....	42

## Índice de Figuras

Figura 1: Tasa de empleo. Fuente: Elaboración propia con datos obtenidos del INE....	10
Figura 2: Tasa de desempleo. Fuente: Elaboración propia con datos obtenidos del INE .....	12
Figura 3: Evolución salario medio. Fuente: Elaboración propia con datos obtenidos del INE .....	13
Figura 6: Tabla de resumen de las variables numéricas del modelo. Datos obtenidos del INE .....	34
Figura 7: Matriz de correlación de las variables del modelo. Elaboración propia a partir de datos obtenidos en el INE. ....	34
Figura 8: Tabla de resumen de las variables numéricas normalizadas del modelo. Datos obtenidos del INE. ....	36
Figura 9: Resultados del modelo de redes neuronales. Elaboración propia .....	38
Figura 11: Evolución de la variable “Vacantes” para cada uno de los grupos de actividad (20013-2023). Elaboración propia. ....	40
Figura 12: Resultados del modelo de redes neuronales en la predicción de valores de vacantes por industria para el año 2023. Elaboración propia. ....	41
Figura 13: Tasa de variación del número de vacantes interanual. Elaboración propia. .	42

## **1. Introducción**

### **1.1. Objetivo de estudio**

El objetivo principal de este trabajo de fin de grado es investigar y analizar la eficacia de las redes neuronales en la predicción de vacantes en las diversas industrias del mercado laboral español. El empleo y la disponibilidad de oportunidades laborales son aspectos fundamentales para el desarrollo económico y social de un país, especialmente para los más jóvenes, que son el grupo más castigado por la incertidumbre y contracciones del mercado. En este contexto, el uso de técnicas avanzadas de análisis de datos, como las redes neuronales, ofrece una gran oportunidad para comprender y predecir de manera más precisa la demanda y la oferta de empleo en diferentes sectores.

El interés en este tema radica en la necesidad de mejorar la toma de decisiones en materia de políticas laborales y la gestión del mercado laboral en general. La predicción de vacantes puede proporcionar información valiosa para los responsables de políticas, los planificadores del mercado laboral y los profesionales de recursos humanos, permitiéndoles anticipar las necesidades de empleo y adaptar estrategias y programas de formación en consecuencia. Por otro lado, esta información también es de gran utilidad para los propios jóvenes, a quienes les aporta mayor visibilidad y entendimiento de las proyecciones a futuro, obteniendo mayor margen de maniobra.

En resumen, este trabajo tiene como objetivo aplicar y evaluar el uso de redes neuronales en la predicción de vacantes en el mercado laboral español por industria, con un enfoque especial en el mercado laboral juvenil. Se busca contribuir al conocimiento científico en el campo de la predicción laboral, así como proporcionar información práctica y perspectivas para la mejora de las políticas y estrategias relacionadas con la empleabilidad de los jóvenes.

Por otro lado, los objetivos específicos del trabajo son los siguientes:

- Hacer una revisión de la literatura existente sobre la predicción laboral y el estado actual del mercado laboral juvenil.
- Recopilar y analizar información relevante sobre los factores que dificultan la inserción de los jóvenes en el mercado laboral juvenil.
- Desarrollar un modelo de redes neuronales que permita la correcta predicción de vacantes en el mercado laboral español por industrias.

- Aportar recomendaciones de utilidad para instituciones públicas y jóvenes con el fin de lograr una optimización de la oferta y demanda de empleo en España. Una de las formas de lograr este objetivo es a través de la orientación laboral de los jóvenes, con la elección de formaciones profesionales o estudios universitarios que les ayuden a ajustarse a la oferta de empleo en España.

## 1.2. Metodología

En este trabajo de fin de grado primero se llevará a cabo un estudio cualitativo que consistirá en la revisión bibliográfica de la literatura existente acerca del tema tratado. Esta parte del estudio permite comprender y analizar la información ya escrita sobre el mercado laboral juvenil y las vacantes que existen en las industrias en España, facilitando el enfoque del análisis.

La búsqueda de información se ha realizado utilizando diversas fuentes académicas, incluyendo artículos científicos, revistas especializadas, informes de consultoría e informes anuales de instituciones reconocidas. Estas fuentes han sido obtenidas de plataformas como Google Scholar, Dialnet, páginas web oficiales, Statista y el repositorio de la biblioteca de la Universidad Pontificia Comillas. Asimismo, este análisis se verá complementado por un estudio cualitativo de los principales indicadores del empleo en el mercado español.

El enfoque principal de este trabajo es cuantitativo y gira alrededor de la elaboración de un modelo de redes neuronales artificiales para predecir el número de vacantes por industria dependiendo de los valores de las diferentes variables del modelo. Este análisis se lleva a cabo en la herramienta RStudio, y los datos utilizados para el entrenamiento y prueba del modelo están extraídos en su totalidad del Instituto Nacional de Estadística (en adelante INE). La elección de este tipo de red se debe a su adaptabilidad y flexibilidad, así como su capacidad de aprendizaje no lineal. El modelo pasará por un proceso de entrenamiento y validación por el método de *cross-validation*. A través del modelo, se llevará a cabo el cálculo de proyecciones para el año 2023 con las predicciones de los valores de ese año

Finalmente, mediante un proceso de comparación de los análisis realizados en las primeras dos partes del trabajo, se procederá al planteamiento de una serie de recomendaciones y conclusiones. Será imprescindible en este último proceso el contraste

de resultados y desarrollo de *insights* extraídos a partir de la fusión de conocimiento de los dos análisis.

### **1.3. Estructura**

Mediante la metodología previamente explicada, este estudio se estructura en tres bloques fundamentales. En primer lugar, se realizará una revisión exhaustiva de la literatura existente sobre el tema abordado, que conformará el marco teórico. Las fuentes consultadas incluirán principalmente trabajos académicos, libros, normativas y publicaciones periodísticas

Con el objetivo de lograr un análisis completo y detallado, se abordarán los siguientes temas de manera secuencial, ya que son esenciales para una comprensión integral del objetivo de investigación:

- Descripción y análisis del mercado laboral juvenil en España
- Factores que dificultan la inserción de los jóvenes en el mercado laboral juvenil
- Principales indicadores y predictores del mercado laboral juvenil

Una vez analizados y comprendidos los ya enumerados ámbitos, se procederá con la elaboración de un modelo de redes neuronales con el objetivo de desarrollar predicciones para el año 2023. Se explicará el proceso de selección de variables, la extracción y construcción de la base de datos sobre la que se apoya el modelo. Además, antes de presentar los resultados del modelo se expondrá la explicación del proceso de validación y corrección del modelo. Finalmente, se utilizarán los resultados para presentar recomendaciones a instituciones públicas y privadas en el último bloque del trabajo.

## **2. Estado de la cuestión**

La salud del empleo es componente fundamental del desarrollo de cualquier economía, es por eso por lo que la comprensión del funcionamiento del mercado laboral en cualquier país es crucial. A través del aumento en la disponibilidad de datos durante los últimos años y los avances en técnicas de aprendizaje y predictivas, se han abierto nuevas posibilidades para el análisis y la predicción del mercado laboral. En este contexto, las redes neuronales artificiales, una rama del aprendizaje profundo, han demostrado ser herramientas muy efectivas para la predicción y el análisis de datos complejos.



El acceso a predicciones del número de vacantes por industria en el mercado laboral español puede ser de inmensa utilidad para varias partes interesadas, como los profesionales de recursos humanos, las autoridades gubernamentales y las empresas. Sin embargo, son los jóvenes los más interesados en esta información, ya que les podría facilitar la orientación de su formación hacia aquellas industrias con más disponibilidad de empleo.

La capacidad de anticipar la demanda de empleo en diferentes sectores y comprender las tendencias del mercado laboral puede tener un impacto significativo en la asignación de recursos, la planificación de la formación y la toma de decisiones estratégicas. Sin embargo, a pesar de la importancia de este tema, en la actualidad hay escasez de estudios e investigaciones que aborden la predicción de empleo por industrias. Por lo general, las investigaciones en este ámbito están orientadas a los factores que afectan al empleo de forma general, sin entrar en detalles por industria y sin predecir valores.

### **3. Descripción del mercado laboral juvenil en España**

Con el objetivo de llevar a cabo un análisis preciso para la predicción de vacantes por industria, es necesario conocer el mercado laboral español, sus fortalezas, carencias y la situación que atraviesa. Por esta razón se analizan en este apartado sus características a nivel cuantitativo, así como su tendencia a lo largo de los años y su posición en comparación con Europa.

#### **3.1. Situación actual y características**

El mercado laboral juvenil en España está conformado por aquellas personas activas laboralmente desde los 16 hasta los 25 años de edad, siendo 16 la edad mínima para ocupar un puesto de trabajo. Este será el marco sobre el que se llevará a cabo el estudio estadístico a continuación, destacando que en ocasiones la agrupación se extenderá hasta personas de 29 años, debido a las limitaciones de la información disponible.

En España se considera población activa a aquellas personas de 16 años en adelante que ocupan un puesto de trabajo proporcionando mano de obra, o simplemente se encuentran a disposición de incorporarse a realizarlo. Este segmento de la población se divide entre ocupados, aquellas personas que trabajan al menos 1 hora semanal a

cambio de una retribución, y los parados, quienes no poseen trabajo, pero se encuentran disponibles para incorporarse y buscan activamente un empleo.

Los ocupados se pueden categorizar entre los trabajadores por cuenta propia y asalariados, y a su vez entre aquellos que se hayan empleados a tiempo completo y los que trabajan a tiempo parcial. Esta última clasificación se lleva a cabo a partir del número de horas que componen la jornada laboral semanal del empleado, siendo 35 horas semanales el punto de corte entre ambos. La Figura 1 indica la evolución de la tasa de empleo desde el año 2006 hasta los últimos datos disponibles de 2022. Esta tasa es un cálculo del ratio entre el total de personas ocupadas y el total de la población nacional.

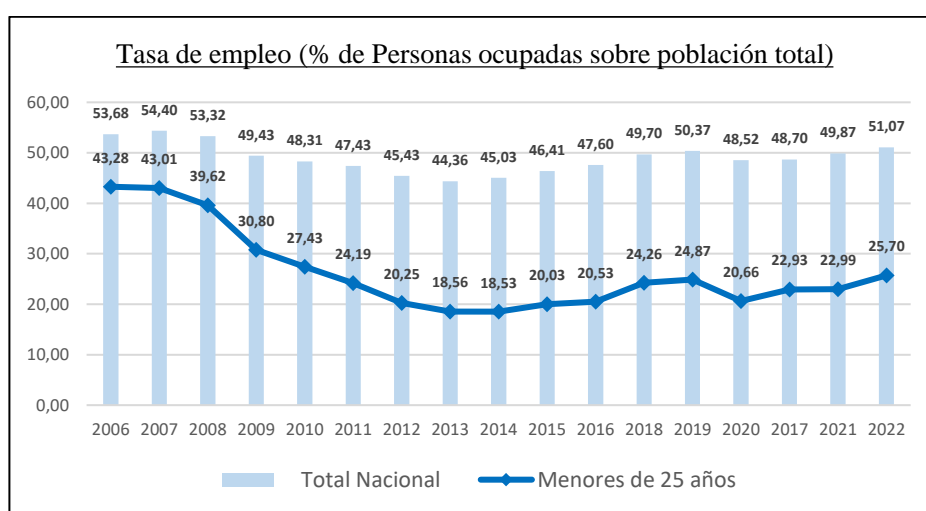


Figura 1: Tasa de empleo. Fuente: Elaboración propia con datos obtenidos del INE

Es notorio como en la actualidad, la tasa de empleo para los menores de 25 se encuentra alrededor de 20 puntos porcentuales (en adelante p.p.) por debajo del valor de la tasa a 2006 (43,28%). Esta drástica caída puede deberse a mayor imposibilidad por parte de los jóvenes de encontrar empleo o a otros factores relacionados con la prolongación de los años de formación, lo cual retrasa la incorporación al mercado laboral.

La última categoría que se considera es la de la población inactiva, la cual está compuesta por aquellas personas que no están clasificadas en ninguna de las categorías previamente expuestas. Es decir, aquellas personas que no se hayan ocupadas ni paradas en búsqueda activa de empleo. Esta falta de actividad laboral se puede dar por muchas razones, como por ejemplo aquellas personas ocupadas por los estudios, los jubilados, o las personas incapacitadas.

Durante las últimas décadas, es notoria la tendencia alcista en los años de formación de la población a nivel global, lo cual retrasa la edad media de entrada en el mercado laboral. Este es uno de los factores de mayor relevancia en la evolución del empleo, ya que en la formación de una persona vienen determinados muchos de los aspectos de su vida laboral (Martin, 2009).

Sin embargo, siendo la formación un eslabón clave en el desarrollo profesional de las personas, sigue siendo una decisión altamente condicionada por la influencia de factores como la familia, el entorno, la clase social o la situación del mercado laboral. También se ha vinculado las carencias formativas o el abandono escolar con la escasa empleabilidad y altas tasas de desempleo en el mercado juvenil (Moreno Mínguez, 2015).

En la actualidad, los jóvenes tienen que afrontar numerosas adversidades a la hora de incorporarse al mercado laboral. No solamente es complejo el proceso de encontrar trabajo, sino también el que ese empleo coincida con sus cualificaciones y sea de calidad. Además, el factor de la temporalidad del empleo indica también una situación alarmante en cuanto a la estabilidad (Arrazola, Galán & De Hevia, 2018).

Siendo este el panorama que se encuentran los jóvenes al adentrarse en el mercado laboral es comprensible la escasa motivación y tendencia a la emigración de talento que azota a España en la actualidad.

### **3.1.1 Análisis estadístico y tendencias de los principales indicadores del empleo en la actualidad**

Una vez delimitado el marco de estudio, tomaremos los principales indicadores que considera el INE para el análisis exploratorio y explicativo del mercado laboral juvenil en España. A continuación, se detallan y definen las variables seleccionadas para la caracterización de la situación actual. Se comparará el valor para cada una de las variables mencionadas del mercado laboral total con el mercado laboral juvenil. Es importante señalar que la evolución de estas variables ha sido afectada por la crisis sanitaria y económica que ha atravesado España desde el estallido de la pandemia de la COVID-19 en marzo de 2020.

- Personas activas: tomando como referencia el último cuatrimestre de 2022, el total de personas activas era de 4.350.000 (35,2% sobre el total del grupo de edad). Esta

cifra se dividía entre 1.958.200 personas ocupadas (24,9% sobre el total del grupo de edad) y 449.300 parados (10,3% sobre el total del grupo de edad) (INE, 2023).

- La tasa de empleo juvenil ha sido ciertamente más baja que la del mercado laboral total, pero esta diferencia es justificable al encontrarse gran parte de la población de ese rango de edad todavía en periodos de formación o terminando los estudios. Sin embargo, es apreciable como a lo largo de los años, esta cifra ha disminuido a un ritmo y en dimensiones proporcionalmente mayores a las del total nacional. Sobre todo, durante periodos de estallido de crisis como 2008, 2012 o 2020, esta fracción del mercado es duramente castigada. Este fenómeno es apreciable en la Figura 2, donde se ilustra la evolución de la tasa de desempleo de ambos grupos.

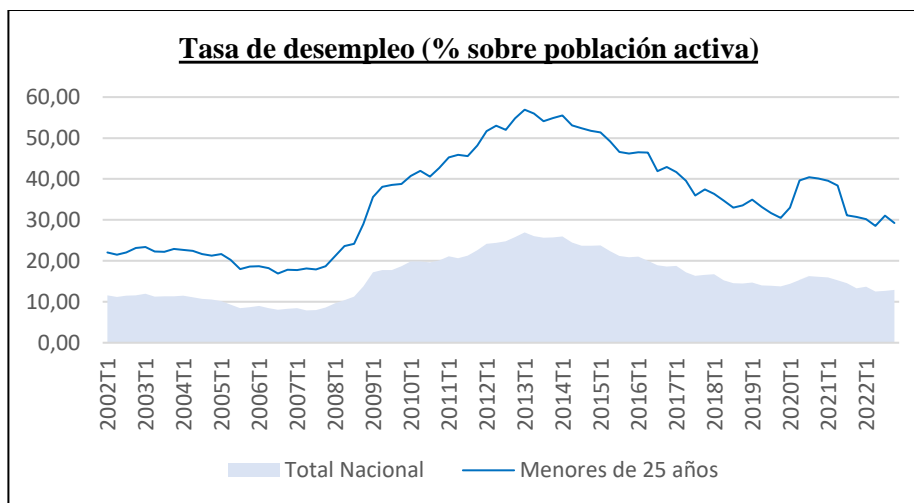


Figura 2: Tasa de desempleo. Fuente: Elaboración propia con datos obtenidos del INE

De nuevo, los valores son significativamente más altos para los jóvenes y escalan de forma alarmante durante los años de la crisis financiera de 2012. Durante los últimos años, vemos un pico que hace que la tasa de paro juvenil roce el 40%, siendo consecuencia de la pandemia de la COVID.

- Manteniendo el estudio en un plano generalista, es relevante el estudio del salario medio de la población activa segregando y comparando el total nacional a la media de los menores de 25. En este caso partimos de una base 100, lo que permite entender un mismo punto de partida para ambas cifras y analizar su evolución de una forma comparable. Esta forma de analizar los datos permite un plano más sencillo de análisis, ya que es complicado comparar directamente ambas cifras al natural.

Como es observable en la Figura 3, el ritmo de crecimiento del grupo más joven es significativamente menor ya que en el año 2022 la diferencia es de 18 p.p.. Durante los años 2010-2018 vemos una acentuada diferencia entre las cifras, hasta el punto en el que el salario para la población joven cae por debajo del punto de partida.

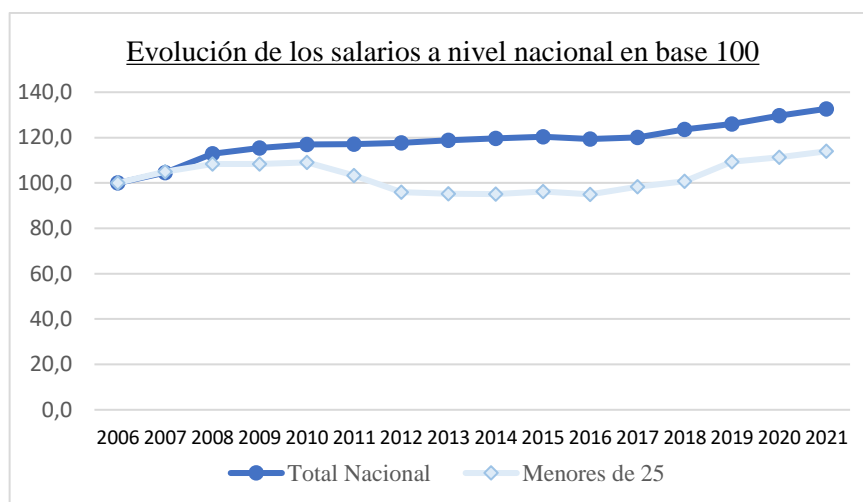


Figura 3: Evolución salario medio. Fuente: Elaboración propia con datos obtenidos del INE

### 3.1.2 Detalle de las diferencias sectoriales

La situación del mercado juvenil no es uniforme, es decir, dependiendo del género, grupo de edad, e incluso industria, las condiciones e indicadores pueden variar. Siguiendo el objetivo de la comprensión global del mercado de trabajo en España, es imprescindible abordar estas brechas o diferencias. A continuación, se describen las principales diferencias entre distintos grupos para el salario, desempleo e IPC.

#### 3.1.2.1 Salario

La brecha de género sin ajustar se refiere a la diferencia bruta en salarios entre hombres y mujeres, sin considerar las diferencias subyacentes en términos de características socioeconómicas y puestos de trabajo. Esta medida está más relacionada con el concepto de igualdad. Por otro lado, la brecha salarial ajustada busca medir la diferencia en salarios controlando estos factores, con el objetivo de obtener una medida de "diferencias en remuneración por el mismo trabajo". Esta medida está más relacionada con la idea de equidad. Para lograr esta medida de "mismo pago por el mismo trabajo", es fundamental controlar tantas diferencias relevantes entre hombres y mujeres en la determinación del salario como sea posible (Anghel et al., 2019).

Sin embargo, este concepto es complejo, ya que incluso los cálculos más complejos pueden omitir variables relevantes que no se pueden observar directamente, como la motivación, la competitividad o las actitudes hacia el riesgo. A pesar de estas limitaciones, la brecha de género ajustada es la medida más confiable para determinar si hombres y mujeres reciben una remuneración similar por tareas similares. En este artículo, se encontró que una vez que se controlan todas las características observables, la brecha salarial es desfavorable para las mujeres, lo que indica que potencialmente están siendo discriminadas en el mercado laboral. Además del salario, las mujeres también enfrentan mayores tasas de desempleo, temporalidad y empleo parcial no deseado, lo que indica que la discriminación de género abarca varias dimensiones (Anghel et al., 2019).

En España, se produjo un proceso de transformación social en un período de tiempo bastante breve en comparación con otros países de nuestro entorno. En tan solo 30 años, la tasa de participación de las mujeres aumentó en un 30%, pasando de ser una de las más bajas de la Unión Europea a mediados de los años ochenta, a estar por encima del promedio de la región en años más recientes. Este cambio es aún más notable si consideramos que, durante el mismo período, la tasa de participación en edades más tempranas disminuyó debido a la mayor inversión de las mujeres jóvenes en educación, lo que amplió su tiempo dedicado a la formación. Este mayor compromiso de las mujeres con la educación también refuerza el cambio cultural hacia su inserción en el mercado laboral y, lo que es aún más importante, su permanencia en él (Cuerva, 2018).

Las mujeres han logrado superar gradualmente a los hombres en términos de nivel educativo, especialmente entre los grupos de edad menores de 50 años y de manera más pronunciada en las cohortes más jóvenes (Conde-Ruiz y Marra, 2016a). Este progreso educativo ha tenido un impacto significativo en el ámbito laboral. El porcentaje de mujeres empleadas con educación terciaria alcanzó casi el 40% en 2008, un incremento de casi 10 puntos porcentuales en comparación con el año 2000, y situándose por encima del promedio europeo que era del 31%. Al mismo tiempo, el porcentaje de mujeres empleadas con niveles de educación más bajos disminuyó en proporción similar (Cuerva, 2018).

Por otro lado, la brecha de género no es el único problema que azota el mercado laboral en la actualidad. Por ejemplo, durante la crisis, la brecha salarial se acentuó principalmente entre los trabajadores mayores de 50 años, aquellos con empleos

permanentes y, en menor medida, entre los que tenían jornadas laborales a tiempo completo. Sin embargo, se produjo una disminución notable de la brecha salarial entre la población más joven y los empleados a tiempo parcial. Durante la etapa de recuperación, la brecha salarial disminuyó de manera generalizada, afectando a todos los tipos de contratos, jornadas laborales y grupos de edad, excepto en la población de mayor edad, donde continuó aumentando de manera significativa (Cuerva, 2018).

### *3.1.2.2 Desempleo*

Indudablemente, todas las formas de desempleo constituyen un problema de importancia social y económica en cualquier economía. Sin embargo, el desempleo juvenil plantea un desafío adicional en comparación con el desempleo en otros grupos, ya que el futuro social y económico de cualquier economía depende de los individuos jóvenes. Por lo tanto, si el mercado laboral no logra incorporarlos de manera efectiva y en buenas condiciones a la actividad económica y social, el futuro de esa economía se encuentra en riesgo (Arrazola et al., 2018).

El empleo juvenil es uno de los problemas globales que requiere atención en la actualidad y es motivo de preocupación tanto para los gobiernos como para las organizaciones internacionales, debido a sus efectos más allá del mercado laboral. La falta de oportunidades o las escasas perspectivas laborales que enfrentan los jóvenes después de completar la educación obligatoria tienen un impacto significativo en la cohesión social, especialmente cuando se produce el abandono temprano de la educación. En esta etapa, la armonía entre lo personal y lo profesional es de gran importancia tanto a nivel individual como social (Olaguibe, 2021).

Según la Organización Internacional del Trabajo (OIT), hay alrededor de 70 millones de jóvenes en todo el mundo en busca de empleo. Además, más de 160 millones de jóvenes trabajan, pero viven en la pobreza, y 2 de cada 5 jóvenes se encuentran desempleados o trabajan en condiciones de pobreza. En comparación con los adultos, la tasa de desempleo juvenil es tres veces más alta. En la actualidad, como resultado de la crisis sanitaria y económica causada por la pandemia de COVID-19, la OIT prevé que el problema del empleo juvenil se agravará en tres aspectos principales: interrupciones en la educación, la formación y el aprendizaje en el trabajo; mayores dificultades para los

jóvenes en busca de empleo y aquellos que se incorporan por primera vez al mercado laboral; y la pérdida de empleos e ingresos, junto con la disminución de la calidad del empleo (Olaguibe, 2021).

Si el hecho de tener una mayor probabilidad de desempleo durante la juventud en comparación con la edad adulta se originara únicamente por causas friccionales y, además, el desempleo juvenil no tuviera ningún impacto negativo a largo plazo para los individuos, se podría considerar como un pequeño obstáculo sin importancia significativa en la trayectoria laboral, algo por lo que muchos jóvenes pasan debido a las dificultades naturales de pasar de ser estudiantes sin experiencia laboral a estar plenamente integrados en el mercado de trabajo. Si esto fuera así, desde una perspectiva de política económica, el desempleo juvenil no requeriría un tratamiento diferenciado ni una atención especial en comparación con el desempleo en general (Arrazola et al., 2018).

Sin embargo, si el desempleo juvenil tuviera un componente diferencial propio, más allá de las fricciones normales, y/o si generara consecuencias económicas y sociales duras y significativas, es decir, "cicatrices" reales en la vida laboral de los individuos que perduran en el tiempo, entonces la actitud hacia el desempleo juvenil desde una perspectiva de política económica debería ser distinta: se requeriría una respuesta proactiva y persistente en la lucha contra esta problemática (Arrazola et al., 2018).

La evidencia empírica a nivel internacional parece sugerir la existencia de estos efectos permanentes de los episodios de desempleo juvenil, esas "cicatrices" en la vida laboral de los individuos que se reflejan en diferentes aspectos, como los salarios, las futuras situaciones de empleo y desempleo, e incluso la felicidad y el bienestar de las personas. Aunque limitada, en el caso de España también existe cierta evidencia de la existencia de estas "cicatrices" debido al impacto del desempleo juvenil. Precisamente, uno de los objetivos principales de este estudio es proporcionar una primera evidencia sobre el posible impacto que tiene haber experimentado episodios de desempleo durante la juventud en las posibilidades de sufrir desempleo en la edad adulta en el caso de España (Bell & Blanchflower, 2011; Dolado et al., 2013).

En la actualidad, el fomento del empleo juvenil se considera una expresión de la política social europea y es un objetivo prioritario dentro de la estrategia europea de empleo. Dado que existen diferentes realidades nacionales y regionales en el espacio



europeo, abordar el problema del empleo juvenil resulta difícil desde una perspectiva unificada y homogénea en la Unión Europea (Ahumada, 2021).

No obstante, se han formulado recomendaciones a los sistemas educativos y a los mercados laborales con el fin de mejorar la empleabilidad de los jóvenes. Este desafío depende en gran medida de lograr un mayor nivel de cualificación profesional y facilitar la transición de la escuela al trabajo (Ahumada, 2021).

En este sentido, se han implementado diversas iniciativas nacionales que cumplen con el requisito comunitario de garantía laboral juvenil. En términos generales, esto implica que todos los jóvenes menores de veinticinco años reciban una oferta de empleo o formación al finalizar su educación o cuando se encuentren en situación legal de desempleo. Esta meta también ha sido respaldada por una alianza europea para la formación de aprendices, que busca combatir el desempleo juvenil mediante la mejora de la calidad y la oferta formativa para los aprendices en la Unión Europea (Ahumada, 2021).

### *3.1.2.3 IPC*

El Índice de Precios al Consumidor (IPC) es una medida estadística que muestra cómo varían los precios de los bienes y servicios que consume la población de un país. Es el indicador más conocido y relevante de la inflación en una economía. Los agentes económicos utilizan este índice para formar sus expectativas sobre la inflación, lo cual tiene un impacto significativo en los medios de comunicación y en la vida política, especialmente cuando la inflación es alta. Para calcular el IPC, se observa la variación de precios de los bienes comunes que adquieren los habitantes de un país. Dado que los patrones de consumo cambian con el tiempo, el IPC se debe ajustar para reflejar de manera precisa la realidad que se pretende medir. Estos cambios en los patrones de consumo alteran las ponderaciones que el índice asigna a las diferentes categorías de bienes (Perni Llorente & Toribio Muñoz, 2022).

El IPC se presenta como una serie de números que indican cómo los precios han variado a lo largo del tiempo, tomando como referencia un año en el que el índice tiene un valor de 100. Si el IPC de un año específico es superior a 100, significa que los precios son más altos que en el año de referencia, mientras que si es inferior a 100, indica que los precios han disminuido. Además, si tomamos una magnitud expresada en euros

corrientes, la dividimos por el IPC del año base y luego la multiplicamos por 100, obtendremos el valor de esa magnitud en euros constantes (Perni Llorente & Toribio Muñoz, 2022)

Por lo general, la composición del IPC varía entre países, ya que los hábitos de consumo suelen ser diferentes. Sin embargo, con la creación de la Unión Monetaria Europea, y con el objetivo de que la inflación de los países miembros se acerque, surgió el Índice de Precios al Consumo Armonizado (IPCA). Este índice utiliza una misma canasta de bienes de consumo para todos los países y, en general, el IPC y el IPCA difieren solo en algunas décimas (Perni Llorente & Toribio Muñoz, 2022).

Los economistas especializados en macroeconomía hacen una distinción entre las curvas de Phillips a corto plazo y a largo plazo. En la Figura 5, se presenta un ejemplo de una curva de Phillips típica, donde el eje horizontal representa la tasa de desempleo, mientras que en el eje vertical izquierdo se muestra la tasa anual de inflación de precios, y en el eje vertical derecho se representa la tasa de inflación salarial. A medida que la curva se desplaza hacia la izquierda y el desempleo disminuye, la tasa de precios y salarios se incrementa, y por lo tanto, la curva se eleva. En el corto plazo, una curva de Phillips muestra una relación inversa entre la inflación y el desempleo, y además, la escala del eje vertical derecho es mayor que la del lado izquierdo debido a una tasa de crecimiento promedio supuesta del 1% en la productividad laboral (Liquitaya Briceño, 2011).

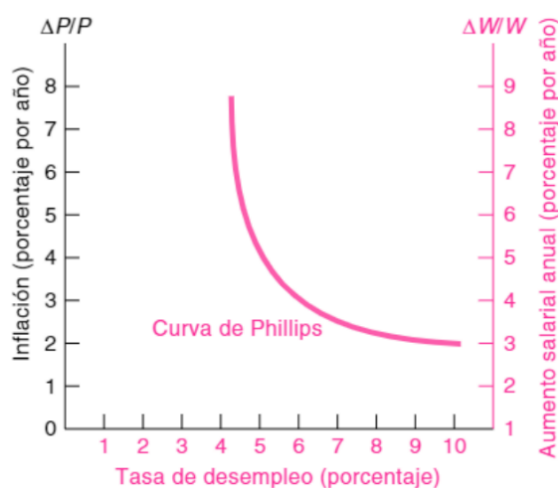


Figura 5. La curva de Phillips (Samuelson, n.d.)

A largo plazo, la curva de Phillips se vuelve vertical, sin tener una pendiente negativa. Esto significa que existe una tasa de desempleo mínima que es consistente con una inflación estable, conocida como la tasa de desempleo no aceleradora de la inflación (NAIRU). En este punto, las fuerzas que afectan los precios y salarios están en equilibrio, lo que implica que la inflación no tiende a cambiar. La NAIRU representa la tasa de desempleo más baja sostenible sin ejercer presión al alza sobre la inflación (Liquitaya Briceño, 2011).

En 2024, se espera que la inflación disminuya, con un crecimiento del deflactor del consumo del 3,5%, superando aún la meta establecida por el Banco Central Europeo. Además, se pronostica un crecimiento del 3,2% en términos de deflactor del PIB (Torres & Fernández, 2023).

La desaceleración económica a nivel internacional, especialmente en Europa, afectará negativamente las cuentas externas, lo que provocará un deterioro en la balanza por cuenta corriente en 2023. Sin embargo, esta tendencia cambiará en 2024 gracias a la mejora global pronosticada, respaldada por el FMI, así como a la competitividad de las empresas españolas. A pesar de esto, teniendo en cuenta los fondos europeos Next Generation, se espera que el saldo total se mantenga en terreno positivo (capacidad de financiación) durante todo el período de previsión (Torres & Fernández, 2023).

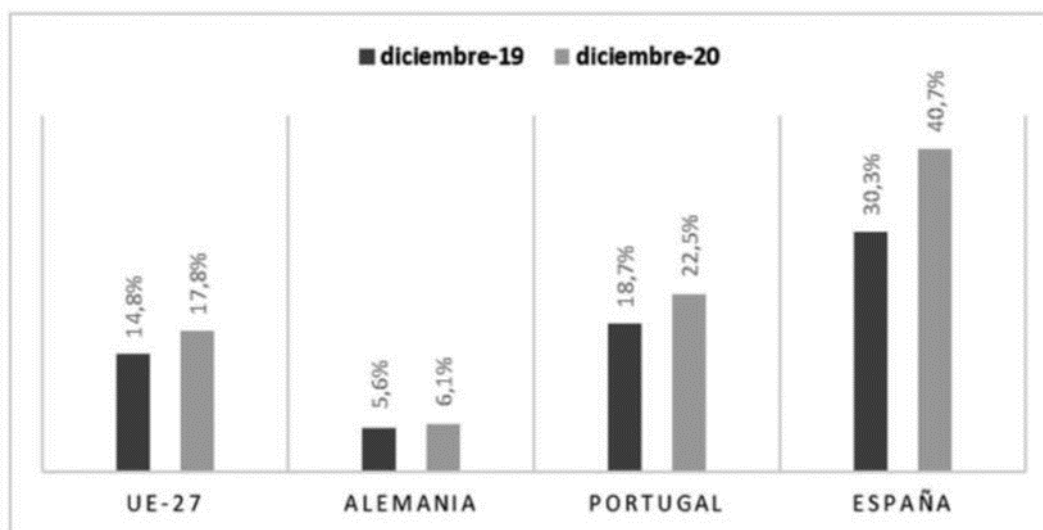
El mercado laboral se verá afectado por la desaceleración, pero no se revertirá la recuperación experimentada en los últimos meses. Se proyecta la creación de casi 100.000 empleos netos en el presente año y otros 250.000 adicionales en 2024 (equivalentes a tiempo completo). Todo esto podría contribuir a una ligera reducción de la tasa de desempleo, situándola alrededor del 11,5% al final del período de previsión, aunque aún sería el peor resultado dentro de la Unión Europea (Torres & Fernández, 2023).

### **3.2 Comparación con la situación del empleo juvenil en Europa**

Los jóvenes se enfrentan a diversos problemas durante la transición entre el sistema educativo y el mercado laboral, y estos desafíos no se limitan únicamente a la crisis económica y sanitaria de COVID-19 sufrida por países de todo el mundo. Estos problemas incluyen el desempleo, el abandono prematuro de la educación, la sobre cualificación y la precariedad laboral. Es evidente que esta situación requiere una

reevaluación más amplia, donde el enfoque no se centre únicamente en medidas sociales, sino en el sistema laboral actual (Olaguibe, 2021).

En la Unión Europea, la tasa de desempleo ajustada por estacionalidad en diciembre de 2019 fue del 14,8%, y en diciembre de 2020 aumentó al 17,8%. España se destaca por tener la tasa de desempleo juvenil más alta, alcanzando el 40,7% como se muestra en la figura 4. Por otro lado, Alemania registra la tasa de desempleo juvenil más baja, con un 6,1%. Sin embargo, también se ha visto afectada por la crisis sanitaria y económica, como se refleja en el gráfico. Portugal se encuentra en una situación intermedia, con una tasa de desempleo juvenil del 22,5% en 2020. Los demás países miembros de la UE se sitúan en diferentes puntos entre estos extremos. En el gráfico que se adjunta, solo se han representado los casos de Alemania, Portugal y España, por ser los más relevantes para la región de Iberoamérica, con Alemania mostrada como el punto de referencia de la UE con el nivel más bajo de desempleo juvenil.



**Figura 4.** Desempleo juvenil desestacionalizado (menores de 25 años) en la UE, Alemania, Portugal, España.  
Fuente: Eurostat, 2021

#### **4. Factores que dificultan la inserción de los jóvenes al mercado laboral**

En la literatura económica se han propuesto diversas explicaciones para abordar esta cuestión. En primer lugar, los jóvenes generalmente poseen un capital humano menos específico, lo que implica que tienen habilidades menos especializadas para desempeñar actividades productivas concretas en el entorno empresarial. Este capital humano específico se adquiere principalmente a través de la experiencia laboral, la cual los jóvenes poseen en menor medida en comparación con los adultos. Las empresas,

conscientes de esta falta de capital humano específico, deben evaluar el costo económico que implica contratar a jóvenes que requerirán tiempo para adquirir dichas habilidades. Además, debido a su menor experiencia laboral, los jóvenes pueden tener dificultades para demostrar su potencial productividad a las empresas (Arrazola et al., 2018; Bell & Blanchflower, 2011).

Por lo tanto, las empresas suelen tener más incertidumbre sobre el desempeño adecuado de los jóvenes en tareas productivas, lo que dificulta su contratación. Para mitigar este efecto, es común que las empresas ofrezcan salarios más bajos y contratos con condiciones menos favorables a los jóvenes en comparación con los adultos, como contratos temporales, lo que les permite evaluar la potencial productividad de los jóvenes. Esto podría explicar, al menos en parte, por qué los jóvenes a menudo tienen condiciones contractuales menos favorables que los adultos (Arrazola et al., 2018; Bell & Blanchflower, 2011).

En el análisis del empleo juvenil contemporáneo, a menudo se pasa por alto la mención del empleo informal o irregular, a pesar de su importancia cuantitativa y cualitativa, especialmente en los países del sur de Europa. En el caso de España, aunque parece que la relevancia del empleo irregular ha aumentado, no siempre se aborda como un fenómeno específico. Se da por sentado que la informalidad está asociada con la precariedad laboral en las trayectorias laborales de los jóvenes españoles, pero no se examina de manera sistemática. La informalidad en el empleo juvenil se manifiesta en trabajos no declarados, pasantías o prácticas no remuneradas fuera del currículo y trabajos esporádicos de diversa índole, que a veces se alternan con períodos de desempleo, inactividad o, en el mejor de los casos, se combinan con actividades de formación. Sin embargo, hay escasos estudios que enfoquen el empleo irregular desde una perspectiva longitudinal secuencial. Por lo tanto, se desconoce el papel que desempeña el empleo irregular en el desarrollo de las trayectorias laborales juveniles y en qué medida se alterna con situaciones de temporalidad y desempleo (Verd & Yepes-Cayuela, 2021).

Además, si el sistema educativo no forma adecuadamente a los jóvenes, no logra evitar el abandono escolar de aquellos que carecen de competencias mínimas o no atiende las nuevas necesidades formativas que surgen en el ámbito productivo, resultará difícil para las empresas encontrar candidatos adecuados entre los jóvenes sin experiencia para ocupar sus puestos de trabajo. Por lo tanto, se señala con frecuencia como una de las

causas de la mala situación de los jóvenes en los mercados laborales un funcionamiento deficiente de los sistemas educativos, especialmente en países como España, con altas tasas de desempleo juvenil y abandono escolar temprano (Dolado et al., 2013; García, 2014).

Otro factor que puede explicar en parte las altas tasas de desempleo juvenil es que, lógicamente, los jóvenes tienen poca experiencia en afrontar el proceso de búsqueda activa de empleo. Han tenido poco o ningún contacto previo con entrevistas de trabajo, pruebas de selección, entre otros, lo que constituye una desventaja en comparación con los adultos. Además, carecen de contactos en el mundo laboral, los cuales pueden ser cruciales para conseguir empleo, y pueden desconocer las mejores opciones de empleo que se ajusten a sus cualidades. En los países más desarrollados, muchos jóvenes pueden tener pocos incentivos para realizar una búsqueda activa de empleo, ya que cuentan con el apoyo económico de sus padres, quienes no tienen inconvenientes en mantenerlos financieramente sin importar su edad o circunstancias. A esto se suma el hecho de que es común que las empresas prefieran contratar a personas con experiencia, aunque sea mínima, lo que introduce a muchos jóvenes en un ciclo vicioso de "no me seleccionan para un puesto de trabajo porque no tengo experiencia, pero no puedo adquirir experiencia porque no tengo empleo" (Arrazola et al., 2018; García, 2014).

Entre los elementos que tienen un impacto negativo en las oportunidades laborales de los jóvenes, además de las condiciones externas y contextuales del mercado de trabajo, se destacan el abandono prematuro de la educación y el nivel educativo alcanzado. Ambos factores plantean desafíos específicos que deben abordarse mediante políticas sociales relacionadas con el empleo juvenil. Además, el desempleo, la sobre cualificación y la precariedad laboral también son aspectos que caracterizan esta transición que los jóvenes en España han enfrentado durante décadas (Olaguibe, 2021).

Los impactos negativos de la gran recesión se han visto amplificados por las graves consecuencias de la crisis económica y social provocada por el Covid-19. En general, se ha observado un retraso en el proceso normal de inserción laboral de los jóvenes. La prolongación de los períodos de formación, junto con la disminución del empleo y el consiguiente aumento del desempleo desde la perspectiva de la Seguridad Social, ha creado una situación anómala y atípica para el fomento del empleo juvenil (Ahumada, 2021).

A esto se suman otros aspectos estructurales, como la frecuente utilización de contratos temporales y la existencia de relaciones laborales precarias que impiden la planificación de un futuro concreto. Todo esto genera un contexto de incertidumbre y dificulta la atención a los intereses personales en términos de acceso al empleo y la integración social. Este retraso en la entrada al mercado laboral está relacionado con el impacto de las coyunturas económicas de crisis. Esta situación ha acentuado la problemática social del empleo juvenil, convirtiéndose en un problema estructural, arraigado y crónico en nuestra sociedad (Ahumada, 2021).

## **5. Principales indicadores y predictores del mercado laboral**

Los efectos del desempleo son más perjudiciales cuando persisten a lo largo del tiempo. El desempleo de larga duración, que se refiere a estar sin trabajo durante más de un año, ha experimentado un aumento constante en los últimos años en España. Según el INE, en el año 2008, era del 2%, pero en 2016 llegó al 11%, alcanzando un pico del 17,8% en 2013. Este tipo de desempleo afecta en mayor medida a las mujeres jóvenes (11,2% en 2016) en comparación con los hombres jóvenes (10,8%). El desempleo de larga duración se considera una de las principales causas que desmotivan la búsqueda de empleo y generan frustración ante un mercado laboral que ofrece pocas oportunidades, lo que ha llevado a muchos jóvenes a buscar oportunidades en otros países (López, 2018; Piqué et al., 2017).

Otro aspecto característico del empleo juvenil es la temporalidad. En el año 2016, la tasa de empleo temporal para el grupo de edad de 15 a 29 años en España alcanzó el 57,4%, según los datos de Eurostat. Esta tasa mide el número de personas contratadas bajo un contrato temporal, frente al total de personas empleadas. Esto representa la tasa más alta de Europa, siendo el promedio de la UE-28 del 32,5%. A lo largo de todo el periodo de crisis, la temporalidad en el empleo de los jóvenes no ha dejado de aumentar. En términos de género, las mujeres jóvenes se ven más afectadas por la temporalidad en España (58,2% en 2016) en comparación con los hombres jóvenes (56,6%), y esta diferencia se ha mantenido e incluso aumentado en los años de la crisis (Benedicto, 2017; López, 2018).

Cuando la temporalidad se combina con altas tasas de desempleo juvenil, como ocurre en España, es un claro indicador de precariedad y aumenta la incertidumbre tanto

laboral como personal. La proliferación de contratos temporales conduce a una rotación laboral, con una secuencia más o menos continua de entradas y salidas del empleo y períodos frecuentes de desempleo. Como señala Santos, llegamos a un punto en el que resulta difícil distinguir si estos jóvenes son trabajadores temporales o desempleados intermitentes: son casi sinónimos. Si bien se asociaba la inestabilidad laboral con la juventud como una etapa transitoria que en algunos casos servía como un trampolín hacia empleos más estables, se está demostrando que, para un número significativo de jóvenes y adultos jóvenes, esa inestabilidad inicial difícilmente se corrige con el paso de los años (E. C. Cano, 2007).

Por otro lado, los porcentajes de empleo a tiempo parcial para los jóvenes han ido aumentando progresivamente. En 2008, era del 15,8%, mientras que en 2016 alcanzó el 27,5%, según los datos de Eurostat. Este tipo de jornada laboral presenta una clara disparidad de género. En 2016, más de un tercio (34%) de las mujeres de entre 16 y 29 años en España trabajaban a tiempo parcial, cifra superior a la de los hombres jóvenes (21,8%). Se observa que, a menor edad, la tasa de empleo a tiempo parcial aumenta, mientras que en las cohortes de edad más avanzadas es menor (Benedicto et al., 2017; Pericàs & López-Andreu, 2016).

Estas tasas son considerablemente altas para los jóvenes en comparación con la media de la población española (15,1% en 2016), especialmente considerando que el mercado laboral español no tenía mucha experiencia en empleo a tiempo parcial y solo ha aumentado gradualmente en tiempos recientes. En este sentido, la demanda de empleo a tiempo parcial es escasa y, en la mayoría de los casos, involuntaria, ya que la mayoría de las personas con jornadas a tiempo parcial no encuentran empleo a tiempo completo. Específicamente en España, el incremento de jóvenes de 15 a 29 años con contratos a tiempo parcial que afirman no encontrar empleo a tiempo completo ha aumentado durante la crisis (del 31,9% en 2000 y 37,3% en 2008 al 62,8% en 2016) (Benedicto, 2017; Piqué et al., 2017).

Estas proporciones difieren considerablemente de las existentes en otros países como Alemania, Bélgica o Reino Unido, donde los porcentajes de jóvenes que trabajan a tiempo parcial porque no encuentran empleo a tiempo completo son inferiores al 20% (Benedicto et al., 2017; López, 2018; Piqué et al., 2017).



Otro indicador de precariedad laboral que se refleja en los datos son los bajos salarios para los jóvenes. En general, durante la crisis, los salarios se han devaluado, pero la devaluación salarial ha sido aún mayor para los jóvenes, quienes han experimentado un aumento en la temporalidad y la parcialidad del empleo. Según los datos de la Encuesta de Estructura Salarial del INE en 2015, la remuneración media anual disminuyó un 5,1% para las personas menores de 24 años (11.228 euros al año), un 1,6% para aquellos de entre 24 y 29 años (hasta 16.064 euros) y un 3% para aquellos de entre 30 y 34 años (hasta 19.597 euros) (Piqué et al., 2017).

La precariedad salarial está estrechamente vinculada al tipo de empleo que los jóvenes tienen. Uno de los principales problemas de la temporalidad laboral y del empleo a tiempo parcial, que se consideran normales para los jóvenes, es que los salarios en estos tipos de empleo son más bajos. Cuanto más reducido es el salario, menor es la base de cotización y, por lo tanto, más limitadas son las prestaciones posibles. Además, en este tipo de empleo es más probable experimentar períodos de desempleo y subempleo entre contratos, lo que reduce aún más los ingresos y prestaciones actuales y futuras, aumentando el riesgo de pobreza y exclusión social (trabajadores pobres) (E. C. Cano, 2007; Junco, 2019; Standing, 2014).

Estos indicadores nos permiten ver cómo se ha configurado un modelo de empleo juvenil precario, que no se limita al período de crisis, ya que muchos de estos rasgos precarios ya existían antes de la crisis. Su configuración se ve influenciada claramente por factores estructurales que hacen de la precariedad un elemento central del mercado laboral, pero también por factores institucionales relacionados con la configuración del Estado de Bienestar y la distribución de recursos públicos a través de políticas públicas, especialmente las políticas de empleo (E. Cano et al., 2000; E. C. Cano, 2007; López, 2018).

La pregunta sobre las secuelas o cicatrices que la crisis y el deterioro generalizado del empleo juvenil están dejando a corto y largo plazo es inevitable, pero las respuestas no son nuevas. Se perpetúan señales que ya se estaban manifestando antes de la crisis, como la diversificación de las trayectorias laborales, las dificultades para lograr independencia económica y vivir de forma autónoma, el retraso en la emancipación juvenil, el regreso al hogar familiar, el aplazamiento en la edad para tener descendencia

y la baja tasa de natalidad, entre otros aspectos (López, 2018; Pericàs & López-Andreu, 2016; Piqué et al., 2017).

Algunos estudios indican que la crisis ha aumentado las desigualdades económicas entre diferentes grupos de edad y que los jóvenes han sido los más perjudicados durante este período, considerándolos los "grandes perdedores de la crisis". Queda por verse si la falta de expectativas y oportunidades para tener una vida mejor que las generaciones anteriores representa una brecha generacional o una ruptura del pacto intergeneracional con graves consecuencias para el futuro de la juventud actual (Pericàs & López-Andreu, 2016; Piqué et al., 2017).

## **6. Explicación y nociones del modelo**

### **6.1. Nociones sobre modelos de redes neuronales**

Antes de explicar qué es un modelo de redes neuronales, es necesario comprender el grupo al que pertenecen, los modelos de aprendizaje supervisados. Este tipo de algoritmos utilizan un conjunto de datos y una serie de instrucciones para entrenar y aprender a realizar procesos como predicciones o clasificaciones. Por lo general, el entrenamiento de estos modelos pasa por una primera etapa de entrenamiento en la que el modelo utiliza un conjunto de datos etiquetados para aprender los patrones y relaciones entre las características de entrada y las salidas esperadas. Posteriormente se verifica la efectividad o capacidad de explicación del modelo mediante el uso de un conjunto de datos de prueba (Nasteski, 2017).

Comprendido en contexto en el que se encuentra este tipo de algoritmo, una red neuronal es un modelo inspirado en el funcionamiento del cerebro humano. De igual forma que nuestro cerebro, las redes neuronales están compuestas por una serie de neuronas artificiales que se organizan por capas. En estas capas, cada una de las neuronas recibe una serie de entradas, las cuales son procesadas y transmitidas a las neuronas de la capa siguiente. Aunque sea un modelo de aprendizaje supervisado, es un modelo no interpretable y funciona como una caja negra que se puede aplicar tanto a tareas de clasificación como de regresión.

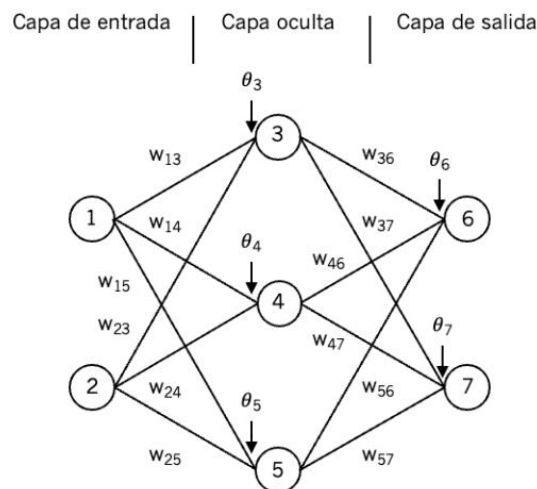
La unidad básica de procesamiento dentro de estos modelos son las ya mencionadas neuronas, las cuales actúan como una función matemática, recibiendo una serie de valores de entrada y generando un valor de salida a través de un procesamiento

interno. El cálculo interno va determinado por las intensidades o pesos que se le atribuyan a cada una de las variables de entrada, los cuales actuarán de palancas para el entrenamiento del modelo.

El siguiente elemento clave en la construcción de las redes neuronales es la arquitectura, es decir, de qué forma se organizan las neuronas para llevar a cabo un proceso que optimice el conocimiento y permita llegar a soluciones no lineales. Las agrupaciones de neuronas se denominan capas, las cuales se organizan de forma secuencial y se dividen en 3 grupos:

- **Capa de entrada:** esta es la primera capa de neuronas, las cuales reciben la información directamente de las variables in haber sido previamente procesadas por otras neuronas.
- **Capas ocultas:** Cada neurona en la capa oculta toma las entradas provenientes de la capa anterior y realiza una combinación lineal ponderada de estas entradas, seguida de la aplicación de una función de activación. Estas funciones se aplican para evitar la linealidad y darle dimensionalidad al modelo de redes y dependiendo del objetivo se pueden aplicar diferentes tipos. La capa oculta permite a la red neuronal aprender y representar características abstractas y de mayor nivel a medida que se profundiza en las capas, lo cual consigue un conocimiento jerarquizado.
- **Capa de salida:** La capa de salida recibe la información del procesamiento de la capa oculta y elabora la solución o respuesta del modelo.

Además de la estructura de las capas de neuronas, hay otros elementos de la arquitectura de la red que son determinantes, como por ejemplo el número de neurona en cada capa, o el peso sinóptico entre ellas (Shmueli et al, 2017).



**Figura 5:** Estructura de una red neuronal (Shmueli et al, 2017)

Finalmente, es necesario comentar la figura del proceso de *backpropagation*, técnica fundamental utilizada en el entrenamiento de redes neuronales. Consiste en la propagación del error desde la capa de salida hacia las capas anteriores de la red, con el objetivo de ajustar los pesos sinápticos de manera que se minimice la función de pérdida. Este proceso es factible gracias a las derivadas parciales de los errores, permitiendo el ajuste de los pesos otorgados a las variables (Jayalakshmi & Santhakumaran, 2011).

## 6.2. Selección y descripción de variables

En línea con el estudio hasta ahora realizado, se ha llevado a cabo la selección de una serie de variables que se consideran relevantes para el desarrollo del modelo. Todas las variables utilizadas en el modelo, y que a continuación se describen, han sido obtenidas del INE y se toman datos observados entre los años 2013 y 2022. Es importante aclarar que la clasificación y organización de las actividades que forman parte de este estudio, vienen dadas por el marco de la Clasificación Nacional de Actividades Económicas (CNAE), la cual asigna un código y sector a cada una de las industrias. En el conjunto de los datos encontraremos variables con valor específico para cada una de las industrias, así como indicadores macroeconómicos nacionales con el mismo valor.

- Fecha: en formato YYYY, indica únicamente el año en el que se toma muestra de la observación.
- Grupo de Actividad: esta variable indica a que rama de actividad pertenece la observación según la clasificación CNAE de 2009. Las ramas que existen dentro de esta clasificación son las siguientes:
  - o B Industrias extractivas

- C Industria manufacturera
  - D Suministro de energía eléctrica, gas vapor y aire acondicionado
  - E Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación
  - F Construcción
  - G Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas
  - H transporte y almacenamiento
  - I Hostelería
  - J Información y comunicaciones
  - K Actividades financieras y de seguros
  - L Actividades inmobiliarias
  - M Actividades profesionales, científicas y técnicas
  - N Actividades administrativas y servicios auxiliares
  - O Administración Pública y defensa; Seguridad Social obligatoria
  - P Educación
  - Q Actividades sanitarias y servicios sociales
  - R Actividades artísticas, recreativas y de entretenimiento
  - S Otros servicios
- Identificador: esta variable creada para el estudio y aísla la primera letra del nombre de cada grupo de actividad para facilitar las labores de clasificación y desarrollo del lenguaje dentro del modelo.
  - Sector: clasifica cada una de las industrias en 3 diferentes categorías que se muestran a continuación, seguidas por el identificador de las industrias que pertenecen a cada una de ellas:
    - Industria: B, C, D y E
    - Construcción: F
    - Servicios: G, H, I, J, K, L, M, N, O, P, Q, R y S
  - Vacantes (valor en miles de personas): esta será la variable objetivo del estudio, es decir, para la que se hará una predicción del valor a futuro. Cada una de las observaciones muestra el número de vacantes para cada grupo de actividad. En el espíritu de llevar a cabo un análisis riguroso y homogéneo, se

han tomado los valores del último trimestre de cada año dentro del periodo de observación.

- Salario Mujeres (valor en €): representa el valor del salario medio de una mujer en cada una de las industrias en el año observado. En la actualidad, el INE solo ha presentado datos hasta 2020, por lo que se hará una predicción explicada más adelante para que la variable sea útil hasta el año 2022.
- Salario Hombres (valor en €): representa el valor del salario medio de un hombre en cada una de las industrias en el año observado. De igual forma que con el salario de las mujeres, el INE sólo presenta datos hasta el año 2020, por lo que también es necesaria una predicción simple para que la variable sea útil.
- Ocupados Totales: Esta cifra (en miles de personas) representa el número de personas empleadas en cada una de las industrias del estudio. Esta variable también se divide entre “Total Hombres” y “Total Mujeres”, dividiendo el número de empleados por sexo.
- Activos 16-19 (valor en miles de personas): número de personas activas en cada industria por año, que se encuentran en la franja de edad entre 16-19 años.
- Activos 20-24 (valor en miles de personas): número de personas activas en cada industria por año, que se encuentran en la franja de edad entre 20-25 años
- IPC (valor porcentual): indicador macroeconómico que indica la variación del índice de precios al consumidor de año en año. Esta variable tiene un valor común para todas las industrias cada año.
- Tasa de Paro Juvenil (valor porcentual): indicador macroeconómico que indica el porcentaje de personas activas desempleadas que sean menores de 25 años en el conjunto del país.
- Coste laboral por trabajador (valor en €): el coste laboral por trabajador es común para cada uno de los sectores anteriormente mencionados y se tomará como referencia el valor del indicador en el último trimestre de cada año.
- Tasa de empleo total: indicador macroeconómico que indica el porcentaje de la población activa que tiene trabajo. Es común a todas las industrias y varía de forma anual.

## 7. Metodología

El objetivo de este modelo de redes neuronales es el de hacer una predicción del número de vacantes de cada industria para un año en concreto. Con esta información se pretende poder presentar recomendaciones a instituciones públicas, privadas y a los propios jóvenes para gestionar exitosamente la demanda y oferta de empleo en el país. El periodo de datos utilizado para desarrollar el modelo es el de 2013-2022, viniendo limitado por la disponibilidad de datos para las variables utilizadas.

Como en cualquier modelo de aprendizaje supervisado el proceso de construcción va a comenzar por la descarga de las variables y la creación de la base de datos. Posteriormente se hará un análisis exploratorio de dichas variables para comprender la estructura y captar *insights* de valor. Seguidamente se hará una partición de los datos entre el conjunto de entrenamiento y el de prueba. Finalmente, tras el entrenamiento del modelo se llevarán a cabo predicciones utilizando estimaciones del valor de las variables para los años venideros.

### 7.1. Descarga de los datos

Los datos utilizados para el modelo provienen en su totalidad del INE, ya que es la fuente con mayor abundancia de información para España. Para todas las variables consultadas, las observaciones tienen carácter anual y en caso de que se registren de cualquier otra forma (cuatrimestral, trimestral, mensual...) se toma el último valor registrado ese año con el objetivo de obtener consistencia.

La consulta y descarga de los datos se hace de forma manual a través de la página web oficial del Instituto, la cual permite la personalización del periodo de tiempo, edades, sexo e industrias para la variable deseada. Una vez descargados se procede a la construcción de la base de datos de forma manual quedando 181 filas y 16 variables (ya descritas anteriormente) en un libro de Excel (xlsx).

### 7.2. R Studio y tratamiento de los datos

El desarrollo del modelo se hará con la herramienta R Studio ya que es una de las herramientas paquetes más completos para modelizar redes neuronales, en este caso usaremos el paquete *nnet*. Además, la plataforma permite la integración de

funcionalidades de visualización muy completas, las cuales son de gran utilidad para la ilustración y ejemplificación de este trabajo.

En primer lugar, se descargan las bibliotecas que serán necesarias para el tratamiento de los datos, el análisis exploratorio, el desarrollo del modelo y las visualizaciones. Las bibliotecas de las que se hará uso en el proceso serán las siguientes:

- *Caret* (Classification And Regression Training): biblioteca de RStudio que facilita el entrenamiento y evaluación de modelos de aprendizaje automático.
- *NeuralNetTools*: biblioteca utilizada para la visualización y análisis de redes neuronales.
- *NNet*: esta es una de las librerías de mayor importancia, ya que entrena y ajusta redes neuronales, permitiendo la personalización de la arquitectura de la red y demás parámetros.
- *Readxl*: utilizada para la importación y lectura de los datos.
- *Ggplot2*: es una de las bibliotecas más conocidas y utilizadas en Rstudio ya que permite elaborar gráficos de alta capacidad de personalización y detalle.
- *Reshape2*: este paquete permite transformar y reorganizar conjuntos de datos en diferentes formatos.
- *Corrplot*: la funcionalidad de esta biblioteca es la de elaborar visualizaciones de matrices de correlación, haciendo su lectura y comprensión más accesible.
- *Forecast*: es una biblioteca dedicada al análisis y pronóstico de series de tiempo. Proporciona funciones y modelos para realizar pronósticos, incluyendo métodos clásicos como promedios móviles, modelos ARIMA y modelos de suavizado exponencial.
- *Stats*: proporciona una amplia gama de funciones y algoritmos estadísticos. Incluye métodos para cálculos estadísticos básicos, pruebas de hipótesis, análisis de regresión y distribuciones probabilísticas entre otros.
- *Effects*: Effects es una biblioteca en RStudio que se utiliza para visualizar los efectos de las variables predictoras en un modelo estadístico.

Una vez descargadas y actualizadas las librerías, leemos los datos en crudo que utilizaremos en el modelo y los guardaremos bajo el nombre “datos” con la función *read\_xlsx*. En primer lugar es necesario que obtengamos una predicción para el salario durante los años 2021 y 2022, ya que, como se ha comentado anteriormente, los datos



no estaban disponibles para esos años en el INE. Esta predicción se lleva a cabo con la función *lm* que permite estimar los coeficientes de una relación lineal entre una variable dependiente y otras predictoras. Estas predicciones son incorporadas al resto de los datos para poder desarrollar el modelo sin valores vacíos.

Por último, y con el objetivo de simplificar la estructura de la base de datos, es necesario que hagamos un cálculo simple para estimar el valor del salario medio, ya que hasta ahora teníamos los datos segregados entre hombres y mujeres. El cálculo es el siguiente:

$$\text{Salario Medio} = \frac{\text{Salario Mujeres} * \text{Total Mujeres} + \text{Salario hombre} * \text{Total Hombres}}{\text{Total Hombres} + \text{Total Mujeres}}$$

Con este cálculo obtendremos para cada grupo de actividad y año el salario medio de sus empleados. Además, creamos otra variable llamada *empleados\_jovenes*, la cual es calcula a través de la suma de las variables de *Activos 16-19* y *Activos 20-24*.

### 7.3. Análisis exploratorio

Antes de comenzar a modelar la red neuronal, es necesario comprender los datos, cómo interactúan entre ellos y detectar si hubiera dependencias. Para esto, primero seleccionamos las variables que vamos a utilizar y las agrupamos en un nuevo conjunto de datos llamado *datos\_resumen*. Las variables que se utilizarán en el modelo son: *Grupo de actividad*, *Sector*, *Identificador*, *Vacantes*, *Ocupados Totales*, *empleados\_jovenes*, *IPC*, *Tasa de Paro Juvenil*, *Coste Laboral por Trabajador*, *Tasa de empleo total*.

En primer lugar haremos un resumen de la información de las variables numéricas con la función *summary*, la cual refleja los resultados expuestos en la Figura 6.

Variable	Mínimo	1er Cuartil	Mediana	Media	3er Cuartil	Máximo
Vacantes	0	771,2	2.579,5	4.568,3	6.290,5	51.426,0
Ocupados totales	28,60	378,40	934,60	976,7	1.364,7	3.140,1
Empleados totales	0,20	11,68	40,75	59,2	60,52	281,3
IPC	-0,01	0	0	0,015	0,016	0,065
Tasa de Paro Juvenil	0,29	0,30	0,38	0,39	0,46	0,54
Coste Laboral por Trabajador	2.518	2.558	2.645	2.783	2.914	3449
Tasa de empleo total	0,44	0,46	0,48	0,48	0,49	0,51

Figura 6: Tabla de resumen de las variables numéricas del modelo. Datos obtenidos del INE

Para la matriz de correlación, primero creamos un subconjunto de datos llamado *datos\_numéricos* y a través de la función *cor* y *ggplot* obtenemos la ilustración de la matriz de correlación de la Figura 7. En la que la escala indica que la matriz muestra colores azules para los valores que están estrechamente inversamente correlacionados y colores rojos para los valores muy correlacionados.

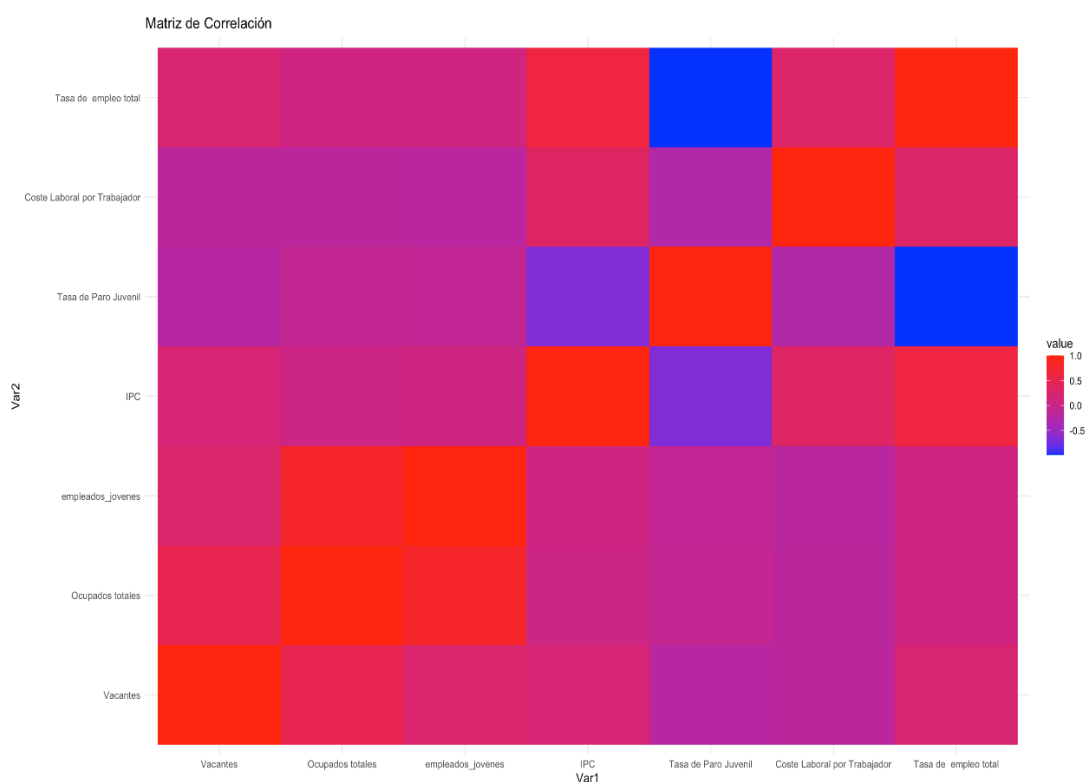


Figura 7: Matriz de correlación de las variables del modelo. Elaboración propia a partir de datos obtenidos en el INE.

En la matriz de relación se puede observar que las variables *ocupadas\_totales* y *empleados\_jovenes* están fuertemente correlacionadas con un coeficiente de correlación de 0,80. Sin embargo, las variables *Tasa de empleo total* y *Tasa de Paro Juvenil* están casi perfectamente negativamente correlacionadas, con un coeficiente de -0,99. Es también significativo que las variables *Tasa de empleo total* e *IPC* muestran un coeficiente 0.63, lo que indica que tienen una correlación positiva. Mientras tanto, la variable *IPC* está negativamente correlacionada con la variable *Tasa de Paro Juvenil* en -0,66.

#### **7.4. Modelo de redes neuronales**

Como se viene diciendo a lo largo de todo este trabajo, el objetivo es el de predecir el número de vacantes por industria a través de un modelo de redes neuronales. Por esta razón, el planteamiento del modelo parte de la base de que la variable *Vacantes* será la variable dependiente, mientras que las siguientes son las variables que se han considerado relevantes a través del análisis exploratorio para ser las independientes o predictoras:

- *Grupo de actividad*
- *Sector*
- *Identificador*
- *Ocupados totales*
- *Empleado\_jovenes*
- *IPC*
- *Tasa de Paro Juvenil*
- *Coste Laboral por Trabajador*
- *Tasa de empleo total*

Por lo tanto, las enumeradas variables predictoras serán las entradas a nuestro modelo, mientras que la variable *Vacantes* será la variable de salida.

##### *Partición de los datos*

Con el objetivo de poder entrenar y probar el modelo es necesario partir los datos en un subconjunto de entrenamiento y otro de test. En este caso la partición es aleatoria y se asignará el 75% de los datos al modelo de entrenamiento y el 25% se reservarán para comprobar la efectividad del modelo más adelante.

### Normalización

Los datos utilizados vienen dados en diferentes unidades de medición, así como en magnitudes muy diferentes. Al realizar el modelo con los datos el crudo, se corre el riesgo de que haya variables que lo desequilibren o que adquieran importancia, perjudicando la efectividad del modelo. Por esta razón se lleva a cabo la normalización de los datos, para que todos se encuentren en un rango semejante de valores. La forma utilizada para llevar a cabo la normalización es la siguiente:

$$\text{Valor Normalizado} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

En este caso la X representa el valor de la variable sin normalizar, Xmin el valor mínimo que adquiere la variable en el conjunto de datos y Xmax el valor máximo. A través de esta fórmula, los valores para cada una de las observaciones se mantienen en el rango 0-1. Una vez realizada la normalización el resumen de la información de las variables queda descrito en la Figura 8, omitiendo las variables de texto.

Variable	Mínimo	1er Cuartil	Mediana	Media	3er Cuartil	Máximo
Vacantes	0	0,023	0,077	0,135	0,189	1
Ocupados totales	0	0,115	0,303	0,320	0,447	1
Empleados totales	0	0,0427	0,150	0,214	0,221	1
IPC	0	0,133	0,240	0,334	0,346	1
Tasa de Paro Juvenil	0	0,056	0,424	0,419	0,662	1
Coste Laboral por Trabajador	0	0,042	0,136	0,293	0,425	1
Tasa de empleo total	0	0,305	0,620	0,556	0,812	1

Figura 8: Tabla de resumen de las variables numéricas normalizadas del modelo. Datos obtenidos del INE.

### Parámetros e hiperparámetros del modelo

Otro de los factores necesarios antes de comenzar a modelar la red neuronal es establecer un parámetro de validación y control del modelo. Estos parámetros permiten ajustar la forma en que los datos se dividen en conjuntos de entrenamiento y prueba, así como las métricas utilizadas para evaluar el desempeño del modelo. En este caso se va a utilizar el proceso de *cross-validation*, el cual divide el conjunto de entrenamiento en

múltiples subgrupos que entrenan y evalúan el modelo obteniendo una estimación robusta y confiable del rendimiento del modelo. Asimismo, es una forma de evitar el sobreajuste o *overfitting* del modelo a los datos de prueba.

Los parámetros que necesitaremos establecer el número de *folds* o subgrupos en los que se dividirá el conjunto del entrenamiento, en este caso son 10. Se entrena el modelo de redes neuronales utilizando los datos de entrenamiento y se evalúa su rendimiento utilizando los datos de prueba. Se repite este proceso *k* veces (en este caso 3), de manera que cada subconjunto se utilice una vez como conjunto de prueba una vez, mientras que los restantes *k-1* subconjuntos se utilizan como conjunto de entrenamiento. Una vez se ha llevado a cabo todo el proceso de control y repetición, el modelo seleccionado será el que consiga la mejor precisión en la métrica que se utilice.

Por último se establece un hiperparámetro a través de la función *expand.grind*, la cual genera todas las combinaciones posibles de los valores proporcionados. Se crea una cuadrícula de búsqueda basada en los valores del tamaño proporcionados, en este caso 3,4,5, y 6, indicando el número de neuronas en una capa oculta.

### *Entrenamiento de la red*

Con los parámetros previamente establecidos, se toman los valores del conjunto de entrenamiento y con el modelo de *nnet* se entrena la red neuronal con el conjunto de los datos de entrenamiento. Los resultados expuestos a continuación en la Figura 9 se refieren a la validación cruzada llevada a cabo en el modelo de redes neuronales. En este caso las métricas utilizadas son las siguientes:

- RMSE: mide la raíz cuadrada del promedio de los errores al cuadrado entre los valores observados y los valores predichos. El objetivo es conseguir el mínimo RMSE, lo cual indicará que el modelo es preciso.
- Rsquared: indica qué porcentaje de la varianza total de la variable predicha está explicada por el modelo. En este caso, el objetivo es obtener un valor cercano a 1, lo que indica que el modelo se ajusta perfectamente.
- MAE: Es una métrica de evaluación del modelo que representa el promedio de los valores absolutos de los errores entre los valores observados y los valores predichos. El MAE es una medida de la magnitud promedio de los

errores del modelo. Cuanto menor sea el valor del MAE, mejor será el rendimiento del modelo en términos de precisión.

Además, durante el proceso de entrenamiento de la red se aplica la función de *backpropagation*, que consiste en dos pasos principales: propagación hacia adelante (*forward propagation*) y retro propagación del error (*backward propagation*). En la propagación hacia adelante, los datos de entrenamiento se introducen en la red neuronal y se calculan las salidas de cada neurona hasta llegar a la capa de salida. Estas salidas se comparan con los valores reales para calcular el error. En la retro propagación del error, el error calculado se propaga hacia atrás a través de la red neuronal, ajustando los pesos de las conexiones en sentido contrario. Esto se hace utilizando el algoritmo de descenso del gradiente, que calcula la derivada parcial del error con respecto a cada peso y actualiza los pesos en la dirección que minimiza el error. En relación con las funciones de activación del modelo, se utiliza aquella predeterminada para el paquete *nnet* utilizado, en este caso la función de activación logística.

Tamaño	RMSE	Rsquared	MAE
3	0.05151937	0.9304650	0.03527922
4	0.05169361	0.9289518	0.03516070
5	0.05119412	0.9300564	0.03488265
6	0.05073792	0.9292147	0.03466878

*Figura 9: Resultados del modelo de redes neuronales. Elaboración propia*

La columna de tamaño indica el parámetro *size* seleccionado anteriormente. De todos los tamaños de la red utilizados en *cross-validation*, se selecciona aquel que tenga el menor valor de RMSE, en este caso el modelo de tamaño 6. A través de la función *summary* destacamos la siguiente información del modelo:

- El modelo tiene 6 capas ocultas
  - Capa oculta 1: 42 nodos
  - Capa oculta 2-6: 6 nodos
- La capa de salida tiene un único nodo

Una vez tenemos buenos resultados en la predicción del conjunto de datos de entrenamiento, aplicamos el modelo de redes neuronales al conjunto de datos reservados para probar la efectividad del modelo. En este caso el RMSE del test set es de 0.08.

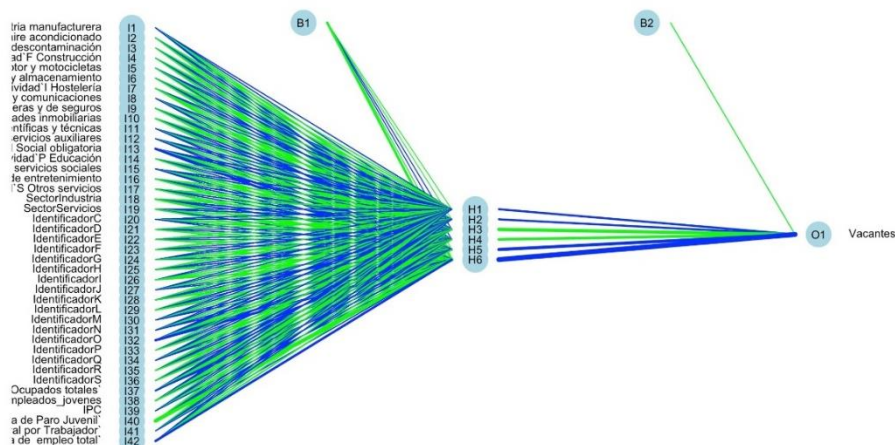


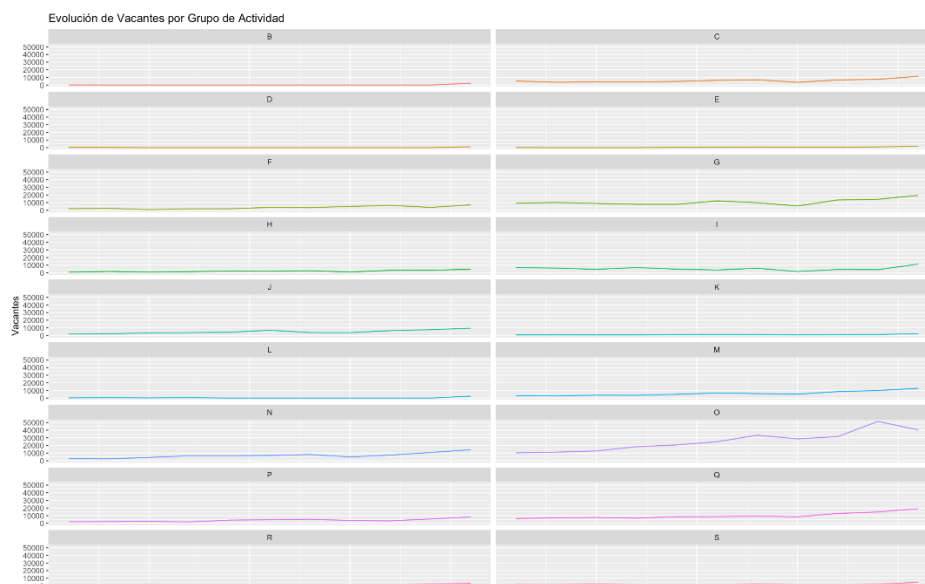
Figura 10: Ilustración del modelo de redes neuronales. Elaboración propia

A parte de el RMSE, el cálculo del error cuadrático medio (MSE) nos dará mucha información a cerca del rendimiento del modelo, ya que es una métrica ideal para los modelos de redes neuronales de regresión. Para el conjunto de test, el valor de MSE se sitúa en 0,0138 y se considerará satisfactorio considerando las características del conjunto de datos utilizados en el modelo.

## 7.5. Predicciones

Con el objetivo de probar la efectividad del modelo y poder presentar recomendaciones, se han introducido los valores de las variables para el primer trimestre del año 2023, de la misma fuente de la que hemos obtenido el resto de los datos. De todas formas, para aplicaciones futuras, se podrían utilizar predicciones de los valores. Una vez aplicado el modelo sobre estos nuevos valores obtenemos una serie de predicciones para el valor de las vacantes durante el año 2023.

Al correr el modelo, obtenemos las predicciones para el total de las vacantes en 2023 en las distintas industrias en el mercado labora español. Estos datos se fusionan en la misma base de datos utilizada para entrenar el modelo, así podremos ver la evolución de las variables para desarrollar las conclusiones. La Figura 11 muestra una visualización de la evolución de las vacantes a lo largo de los años, utilizando la misma escala para que sean comparables entre industrias, mientras que la Figura 12 muestra los resultados específicos de las vacantes por industria para 2023.



**Figura 11:** Evolución de la variable “Vacantes” para cada uno de los grupos de actividad (20013-2023). Elaboración propia.

<b>Industrias</b>	<b>Identificador</b>	<b>Vacantes</b>	<b>% sobre total para 2023</b>	<b>Tasa de crecimiento anual compuesto (2013-2023)</b>
Industrias extractivas	B	2.309	1,3%	46%
Industria manufacturera	C	11.384	6,4%	8%
Suministro de energía eléctrica, gas, vapor y aire acondicionado	D	1.130	0,6%	13%
Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	E	1.764	1,0%	30%
Construcción	F	7.030	4,0%	13%
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	G	19.619	11,1%	8%
Transporte y almacenamiento	H	4.719	2,7%	19%
Hostelería	I	11.373	6,4%	5%
Información y comunicaciones	J	9.471	5,4%	18%
Actividades financieras y de seguros	K	2.214	1,3%	13%
Actividades inmobiliarias	L	2.486	1,4%	24%
Actividades profesionales, científicas y técnicas	M	12.724	7,2%	15%



Actividades administrativas y servicios auxiliares	N	14.260	8,1%	19%
Administración pública y defensa, Seguridad Social Obligatoria	O	40.309	22,8%	15%
Educación	P	8.304	4,7%	15%
Actividades Sanitarias y de servicios sociales	Q	19.117	10,8%	12%
Actividades artísticas, recreativas y de entretenimiento	R	3.381	1,9%	24%
Otros servicios	S	4.929	2,8%	11%
	<b>Total</b>	<b>176.525</b>	<b>100%</b>	<b>13%</b>

*Figura 12: Resultados del modelo de redes neuronales en la predicción de valores de vacantes por industria para el año 2023.*

*Elaboración propia.*

En la Figura 12 es apreciable como las industrias de la Administración Pública, las Actividades Sanitarias y el Comercio al por mayor y por menor van a ser las más relevantes dentro del listado estudiado. Entre ellas concentrarían alrededor del 45% del total de la oferta de empleo de nuestro país durante el año 2023. Mientras tanto, las actividades relacionadas con el Suministro de energía, el suministro y saneamiento del agua y las Actividades financieras y de seguros, serán las que menos oferta de empleo proporcionarían.

Otro de los aspectos más relevantes es el crecimiento del total de las vacantes con respecto al año 2022, que en este caso es de un 27%. Viendo la evolución presentada en la Figura 13, vemos como se mantiene más o menos en la tendencia de el año anterior. Las cifras de los últimos años son significativamente más altas a las de años anteriores, este crecimiento se puede comprender a través del marco de la situación económica del país durante este periodo. Es decir, son años de recuperación tras la crisis sanitaria, social y económica del Covid-19. Asimismo, en el año 2015 podemos apreciar un crecimiento bastante significativo como consecuencia de la recuperación económica de España tras la crisis de comienzo de esa década.

<b>Año</b>	<b>Variación de vacantes con respecto al año anterior</b>
2014	4%
2015	4%
2016	14%
2017	11%
2018	22%

2019	11%
2020	-25%
2021	44%
2022	29%
2023	27%

*Figura 13: Tasa de variación del número de vacantes interanual. Elaboración propia.*

## 8. Conclusiones y recomendaciones

A modo de cierre del trabajo de fin de grado, es de vital importancia la presentación de una serie de recomendaciones y conclusiones extraídas del estudio realizado. En este caso, comenzando con las conclusiones extraídas de forma directa de las cifras obtenidas del modelo desarrollado se plantean las siguientes:

- **Actividades que presentan el mayor crecimiento:** entre las actividades que destacan por su incremento en vacantes en el año 2023, encontramos las Actividades Artísticas, recreativas y de entretenimiento (24%), las Industrias extractivas (46%), las de Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación (30%) y las Inmobiliarias (30%). Sin embargo, esto no significa que sean aquellas que concentran un mayor número de vacantes en el total del mercado.
- **Actividades que con el mayor número de vacantes:** los grupos de actividades con mayores concentraciones de vacantes son las actividades de Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas (11% del total), las actividades de Administración pública y defensa, Seguridad Social Obligatoria (22,8% del total) y el de las actividades Sanitarias y de servicios sociales (10,8%).

Las conclusiones que se pueden extraer del análisis completo es que habría un notorio crecimiento del número de vacantes en España. Ninguno de los grupos de actividad se ha contraído durante los últimos 10 años y el crecimiento tasa anual compuesta para el total de vacantes supera el 10%. El crecimiento con respecto al año 2022, es positivo y está por encima del 25%, considerándose un crecimiento alto y positivo para el mercado laboral español. Sin embargo, alrededor del 23% de las vacantes dependen de la administración pública (Grupo O), lo que no siempre tiene por qué ser positivo.

Con esta serie de conclusiones y la optimización del mercado laboral juvenil como objetivo del trabajo, se plantean una serie de recomendaciones a instituciones públicas, privadas y los propios miembros de la fuerza laboral española.

Dentro del ámbito de las instituciones públicas, es necesaria una planificación estratégica a largo plazo para abordar el crecimiento en las vacantes laborales. Esto implica identificar sectores clave y promover políticas y programas permitan la incorporación de personal cualificado y preparado en las áreas que estás demostrando crecimiento y las que ofrecen el mayor número de vacantes. Por otro lado, es de vital importancia asegurar que existen programas de formación adecuados para formar a los jóvenes en las universidades y centros de formación profesional públicos. Casi un cuarto del total de las vacantes a nivel nacional depende directamente de la Administración pública y defensa, Seguridad Social Obligatoria, lo que hace que estas instituciones tengan gran responsabilidad en la gestión del empleo en España.

Las instituciones privadas, sobre todo en sectores no tan populares, son una pieza fundamental para llevar a cabo programas que anuncien y den a conocer las oportunidades que hay en el mercado. Por ejemplo, esto sería de gran utilidad para sectores con gran crecimiento como puede ser el de la Industria de la Extracción o las Actividades Inmobiliarias. Visto el exceso de talento comparado con la oferta de empleo, las empresas con altas cifras de vacantes pueden aprovechar para la captación de jóvenes antes de comenzar su formación y así optimizar la demanda en el mercado.

Por su lado, los jóvenes deben ser conocedores de las tendencias del mercado laboral, realizar investigaciones y estar al tanto de las tendencias y oportunidades laborales en los sectores con mayor crecimiento. Estando informados, podrán adquirir habilidades relevantes y demandadas, aumentando su empleabilidad. Por ejemplo, basándonos en los resultados obtenidos las actividades sanitarias y las relacionadas con Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas tienen una tasa de crecimiento elevada y concretan gran numero de vacantes a la vez.

## Bibliografía

- Ahumada, J. E. L. (2021). Reflexiones sobre la situación del empleo juvenil y el fomento de su empleabilidad: mercado de trabajo y políticas de empleo en la Unión Europea y España. *Estudios Latinoamericanos de Relaciones Laborales y Protección Social*, 12, 21–37.
- Anghel, B., Conde-Ruiz, J. I., & De Artíñano, I. M. (2019). Brechas salariales de género en España. *Hacienda Pública Española*, 229, 87–119.
- Arrazola, M., Galán, S., & de HEVIA, J. (2018). Desempleo juvenil en España: situación, consecuencias e impacto sobre la vida laboral de los adultos. *Papeles de economía española*, (156), 62-75.
- Arrazola, M., Galán, S., & de HEVIA, J. (2018). Desempleo juvenil en España: situación, consecuencias e impacto sobre la vida laboral de los adultos. *Papeles de Economía Española*, 156, 62–75.
- Bell, D. N. F., & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, 27(2), 241–267.
- Benedicto, J. (2017). Informe Juventud en España 2016. INJUVE. Recuperado de: [Http://Www. Injuve. Es/Observatorio/Demografia-e-Informacion-General/Informe-Juventud-En-Espana-2016](http://www.injuve.es/Observatorio/Demografia-e-Informacion-General/Informe-Juventud-En-Espana-2016).
- Benedicto, J., Echaves, A., Jurado, T., Ramos, M., & Tejerina, B. (2017). Informe Juventud en España 2016. Madrid: Instituto de la Juventud (INJUVE).
- Cano, E. C. (2007). La extensión de la precariedad laboral como norma social. *Sociedad y Utopía: Revista de Ciencias Sociales*, 29, 117–138.
- Cano, E., Bilbao, E., & Standing, G. (2000). Precariedad laboral, flexibilidad y desregulación. Alemania.
- Cuerva, E. C. (2018). La desigualdad salarial de género en España en el contexto de la crisis económica y la recuperación. *Panorama Social*, 27, 65–87.
- Dolado, J. J., Jansen, M., Felgueroso, F., Fuentes, A., & Wölfl, A. (2013). Youth labour market performance in Spain and its determinants: A micro-level perspective.
- García, J. R. (2014). El desempleo juvenil en España. *Información Comercial Española, ICE: Revista de Economía*, 881, 11–28.

- INE, (2023). Informe de Jóvenes y Mercado de Trabajo a Marzo 2023.  
[https://www.mites.gob.es/ficheros/ministerio/sec\\_trabajo/analisis\\_mercado\\_trabajo/jovenes/2023/Informe-Jovenes-Num36-Marzo-2023.pdf](https://www.mites.gob.es/ficheros/ministerio/sec_trabajo/analisis_mercado_trabajo/jovenes/2023/Informe-Jovenes-Num36-Marzo-2023.pdf)
- Junco, J. M. G. (2019). El muro invisible: las dificultades de ser joven en España. *JSTOR*.
- Liquitaya Briceño, J. D. (2011). De la curva de Phillips a la NAIRU: un análisis empírico. *Análisis Económico-Universidad Autónoma Metropolitana*, 26(62), 5–30.
- López, E. S. (2018). Jóvenes, crisis y precariedad laboral: una relación demasiado larga y estrecha. *Encrucijadas: Revista Crítica de Ciencias Sociales*, 15, 3.
- Martin, G. (2009). A portrait of the youth labor market in 13 countries, 1980-2007. *Monthly Lab. Rev.*, 132, 3.
- Moreno Mínguez, A. (2015). La empleabilidad de los jóvenes en España: Explicando el elevado desempleo juvenil durante la recesión económica. *Revista Internacional de Investigación en Ciencias Sociales*, 11(1), 2-24.  
<https://doi.org/10.18004/riics.2015.julio.3-20>
- Olaguibe, J. I. (2021). La transición de los jóvenes hacia el trabajo decente: política social y empleo juvenil en España. *Estudios Latinoamericanos de Relaciones Laborales y Protección Social*, 12, 39–50.
- Pericàs, J. M. V., & López-Andreu, M. (2016). Crisis del empleo y polarización de las trayectorias laborales. El caso de los adultos jóvenes en Cataluña. *Papers. Revista de Sociologia*, 101(1), 5–30.
- Perni Llorente, Á., & Toribio Muñoz, M. del R. (2022). Guía de Estudio de Macroeconomía.
- Piqué, M. À. C., Veà, A. P., & Serés, A. (2017). El modelo de empleo juvenil en España (2013-2016). *Política y Sociedad*, 54(3), 733.
- Samuelson, P. A. (n.d.). Repositorio de Recursos Digitales.
- Standing, G. (2014). Por qué el precariado no es un concepto espurio. *Sociología Del Trabajo*, 82, 7–15.
- Torres, R., & Fernández, M. J. (2023). Perspectivas para la economía española tras la moderación de los precios energéticos, 2022-2024. *Cuadernos de Información Económica*, 292, 1–8.

- Verd, J. M., & Yepes-Cayuela, L. (2021). El empleo irregular en las trayectorias laborales de la población joven. Una aproximación mixta de carácter descriptivo y secuencial. *Cuadernos de Relaciones Laborales*, 39(2), 329–350.
- Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons*. b, 4, 51-62.
- Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl Jr, K. C. (2017). Data mining for business analytics: concepts, techniques, and applications in R. John Wiley & Sons.
- Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical normalization and back propagation for classification. *International Journal of Computer Theory and Engineering*, 3(1), 1793-8201.