

Estudio Longitudinal de la estacionalidad turística en España usando Redes Neuronales

Longitudinal Study of Tourism Seasonality in Spain using Neural Networks

Jimmy Patricio Torres-Bastidas^a, Susana Magdalena Cobeña-Cobeña^b, Bertha Haydee Vásquez-Guevara^c, Jenny Fabiola Montes-Párraga^d, Tanya Negrete-Ontaneda^e, Fabricio Rolando Marcillo-Vera^f

Abstract:

Spain has become a benchmark in terms of tourism, since the number of visitors has multiplied extraordinarily, and has gone from a uniform tourism model, throughout the peninsula, to a nuclear model. The economic variations associated with the economic cycle associated with tourism have an indisputable economic impact on autonomous communities. In this work, the seasonality of tourists in pre-pandemic periods was studied. For this purpose, artificial neural networks were used to predict the type of tourist associated with summer periods. Hence, the artificial neural network was used as a technique that achieved an 86.90% accuracy rate. The relevant results showed that domestic tourists had a greater share of seasonality throughout the year than foreign tourists, while foreigners are more significant in summer periods.

Keywords:

tourism, neural networks, autonomous communities, COVID-19, tourism seasonality.

Resumen:

España se ha convertido en un referente en cuanto al turismo ya que el número de visitantes se ha multiplicado de forma extraordinaria, se ha pasado de un modelo turístico uniforme, a lo largo y ancho de la península, a un modelo nuclear. Las variaciones coyunturales asociadas al ciclo económico asociado al turismo tienen un indiscutible impacto económico en las comunidades autónomas. En este trabajo se estudió la estacionalidad de los turistas en épocas prepandémica. Para este propósito se usaron las redes neuronales artificiales para predecir el tipo de turista que está vinculado con los periodos estivales. De ahí que fue usada cómo técnica la red neuronal artificial que consiguió un 86,90% de acierto. Los resultados relevantes se direccionaron a que los turistas nacionales frente a los extranjeros tuvieron mayor participación de estacionalidad a lo largo del año, mientras que los extranjeros son mayormente significativos en periodos estivales.

Palabras Clave:

Turismo, redes neuronales, comunidades autónomas, COVID-19, estacionalidad turística.

^a Instituto Superior Tecnológico "Japón", <https://orcid.org/0000-0001-7690-4936>, Email: ttorres@itsjapon.edu.ec

^b Instituto Superior Tecnológico "Japón", <https://orcid.org/0000-0002-6896-0054>, Email: scobena@itsjapon.edu.ec

^c Instituto Superior Tecnológico "Japón", <https://orcid.org/0009-0009-4706-8890>, Email: hvasquez@itsjapon.edu.ec

^d Instituto Superior Tecnológico "Japón", <https://orcid.org/0000-0002-7012-9956>, Email: jmontes@itsjapon.edu.ec

^e Instituto Superior Tecnológico "Japón", <https://orcid.org/0009-0000-7427-8185>, Email: tnegrete@itsjapon.edu.ec

^f Autor de Correspondencia, Instituto Superior Tecnológico "Japón", <https://orcid.org/0000-0003-2628-9167>, Email: fmarcillo@itsjapon.edu.ec

Fecha de recepción: 23/08/2023, Fecha de aceptación: 6/10/2023, Fecha de publicación: 05/01/2024

Introducción

La Organización Mundial del Turismo define al turismo como “las actividades que realizan las personas durante sus viajes y estancias en lugares distintos a su entorno habitual por un período de tiempo consecutivo inferior a un año, con fines de ocio, negocios u otros”. Es una actividad que realizamos desde tiempos inmemorables, sin embargo, conforme la sociedad ha ido avanzando a lo largo del tiempo, las maneras de hacer turismo han ido cambiando a la par [1]. En los últimos 40 años se ha vivido una revolución en la dinámica del turismo mundial: la globalización ha multiplicado el turismo internacional (organismos como la Unión Europea que fomenta la diversidad cultural), y la mejora de los medios de transporte que permite recorrer mayores distancias a menor coste [2]. Específicamente en España, donde se ha multiplicado el número de turistas extraordinariamente, se ha pasado de un modelo turístico uniforme, a lo largo y ancho de la península, a un modelo nuclear en el que unas pocas urbes están masificadas (y en algunos casos, generando aversión en la población local) [3].

Además, estos cambios han forzado a los negocios turísticos a modificar sus estrategias. Ahora los turistas son exigentes, buscan lugares únicos, con actividades diferenciadoras, lo que fuerza al sector a ser más competitivo e innovador. En este aspecto España está especialmente bien posicionada, de hecho, el Foro Económico Mundial sitúa España en el puesto número uno mundial en competitividad turística [41], por primera vez, gracias a sus avances en el desarrollo digital del turismo [5]. Con independencia de las variaciones coyunturales asociadas al ciclo económico, el turismo tiene una importancia indiscutible como generador de renta y de empleo lo cual justifica el interés de su análisis, tanto desde el punto de vista académico como político. En este sentido, y en el caso particular de España, se viene especulando, en los últimos años, con la presencia de algunos síntomas de cambio estructural que estarían inducidos por alteraciones en la demanda [11, 45].

En cualquier caso, aunque existen nuevos competidores, España sigue ocupando las primeras posiciones, junto

con Estados Unidos y Francia, como destino turístico, tanto en número de visitantes como en ingresos recibidos [12, 41]. No obstante, la preocupación por la competitividad ha ganado peso entre los factores que podrían estar condicionando la evolución de la actividad turística, lo cual está animando la aparición de distintas aportaciones sobre el tema. Cabría citar, por ejemplo, algunos trabajos que aplicaron el enfoque hedónico (placentero o gustoso) para discutir la competitividad relacionando la calidad y el precio de los hoteles en España [24, 42, 28]. Asimismo, y más recientemente, el Instituto de Estudios Turísticos ha publicado un índice de competitividad referido al periodo 2003-2008, construido a partir de la evolución de los precios de la oferta y de los tipos de cambio y que utiliza para comparar la posición de España con otros países [25, 28, 31, 36].

En este sentido, sorprende un tanto la poca atención que la literatura ha dedicado precisamente al enfoque placentero, ya que pocas veces se mide la estacionalidad y sus implicaciones en función de las medidas y posibilidades [26, 32, 37-39]. En España, por ejemplo, con los trabajos de [14, 17, 18, 37, 40, 43, 44] o [45], los cuáles básicamente utilizan el coeficiente de Gini para medir la estacionalidad para diferentes destinos, desde las Islas Baleares hasta la Costa del Sol o las Comunidades Autónomas. Fundamentalmente, el punto de interés es que las medidas utilizadas pueden dar resultados distintos en función de los datos en cada caso.

El objetivo principal de esta investigación es realizar una aplicación empírica para las comunidades españolas en el período 2002-2019 que estudie la estacionalidad y sus implicaciones. ¿Qué comunidades son mayormente atractivas para la estacionalidad? ¿Qué relación existe entre la estacionalidad y el tipo de turista? ¿Cuál es el turista demanda mayor estacionalidad? Los datos de la actividad turística utilizados provienen de las series de estacionalidades hoteleras del INE. Como es bien sabido, estas estacionalidades son sólo una parte de la demanda. En todo caso, implican la mayor parte de las estacionalidades regladas año a año en todas las comunidades, a falta de datos rigurosos y suficientemente completos sobre la demanda en el formato de alojamiento

residencial. El enfoque del trabajo es general y, por tanto, no va a explorar los patrones y situaciones de las diferentes comunidades con detenimiento. Adicionalmente, se repasa tentativamente la significatividad de algunas variables en relación con la estacionalidad comparada a través de diversos análisis correlativos preliminares. El trabajo, pues, se estructura de la siguiente manera: En la sección segunda se repasan los principales aspectos metodológicos relativos a la medición de la estacionalidad turística, cuando se concibe ésta en términos del desequilibrio existente cada año en la distribución temporal de la actividad.

En la sección tercera se realiza un cálculo de la concentración temporal de la actividad turística en todas las provincias españolas a lo largo del período 2002-2019. En la sección cuarta se sintetizan algunos análisis correlativos con el objeto de establecer una evaluación preliminar sobre factores que pueden estar asociados a los niveles y las diferencias en la concentración temporal. Finalmente, se recogen las principales conclusiones obtenidas del trabajo.

Trabajos relacionados

Diferentes aspectos metodológicos de la estacionalidad

Según Duro [15] fue en Baron [7] uno de los primeros autores en subrayar la estacionalidad turística como uno de los principales desequilibrios de la actividad turística. Han surgido en la literatura diferentes definiciones de estacionalidad. A pesar de la diversidad, bien es verdad que una amplia mayoría de autores hace referencia a la aparición en el destino de una regularidad sistemática dentro del año en los flujos. El propio Baron [7], por ejemplo, habla de ésta como los efectos que ocurren cada año con mayor o menor intensidad en el destino. En [45] definen concretamente en términos de la concentración de los flujos de turistas en cortos períodos de tiempo del año. En todo caso, la definición que más nos satisface es la de los autores [10, 46], que es más abierta que las anteriores: «sería el desequilibrio temporal en el fenómeno del turismo, que puede ser expresado en términos de dimensiones de tales elementos como el número de

visitantes, el gasto de los visitantes, el tránsito en autopistas y otras formas transporte, empleo y entradas de atracciones». Tomando, pues, la definición de Butler como referencia, la estacionalidad debería medirse a través de indicadores de desequilibrio temporal intra-año o, si se quiere, de desigualdad. En este sentido, una vía inmediata consistiría en extraer algunos de los aprendizajes surgidos de la literatura sobre medición de la desigualdad. Autores como [13, 47] o [48] han ofrecido interesantes revisiones al respecto. En esencia, esta literatura ha derivado ciertos axiomas básicos que toda medida debe satisfacer y ha analizado las características asociadas al uso de diferentes medidas. Entre ellas, algunas de las más conocidas son el coeficiente de Gini, los índices de la familia de Theil o el Coeficiente de Variación. El punto esencial aquí es que la aplicación de cada una de ellas puede generar resultados distintos en función de cómo pondera cada una las distancias de las diferentes observaciones. Vamos a revisar en esta sección las características peculiares de la lista de medidas anterior.

En este sentido, las observaciones, que son las unidades básicas para el análisis, pueden ser variadas. Una opción razonable sería considerar los meses del año como unidades básicas, lo cual daría al análisis un grado de detalle interesante, con los datos disponibles, y coincidiría con la unidad de seguimiento temporal típico de las empresas turísticas y las administraciones. A este respecto, sabemos que no es posible, de manera homogénea en España, detallar aún más la actividad turística. En todo caso, podría elegirse una base temporal más agregada para el análisis, por trimestres, cuatrimestres o temporadas. En este sentido, se perdería precisión analítica, pero, se podrían evitar las consecuencias del efecto calendario en algunos de los ejercicios, en lo que respecta a la ubicación de la semana santa. El punto clave sería, entonces, establecer los límites de estos subperíodos, lo cual no es trivial. Por tanto, supongamos, porque la concentración temporal de la actividad va a tomar el mes como unidad de referencia. Entonces, las medidas de desigualdad en realidad serían medidas de desigualdad intermensual de la actividad.

Probablemente la medida más conocida destinada a medir de manera sintética el grado de desigualdad (dispersión) sea el coeficiente de Gini [22, 49, 50, 52], que se relaciona de manera gráfica con la conocida curva de Lorenz. Una de las fórmulas más convencionales de esta medida es la siguiente:

$$G = \frac{1}{2u} \sum_i \sum_j P_i P_j |y_i - y_j|$$

donde P_i y P_j son los pesos de las observaciones (meses o temporadas); y_i es, en nuestro caso, la demanda hotelera mensual y u es la media anual que P_i y P_j pueden ser dos meses cualesquiera del año. Asimismo, nótese que las ponderaciones de cada observación, en su aplicación al análisis de estacionalidad, son iguales para todos los períodos comparados y equivalentes a $1/n$, donde «n» es el número de períodos considerados. Así las cosas, si el análisis se centra en una comparación mensual «n» es igual a 12 y si, por ejemplo, es cuatrimestral sería $1/3$. En particular, se han ensalzado tres características respecto a esta medida [34, 53, 55-57, 60, 62]. Primero, su reducida dependencia en relación con los cambios en meses altos (peak months); segundo, su elevada estabilidad comparada con otras medidas y tercero, su reducida sensibilidad a los valores (meses) extremos. Sin embargo, y más allá de compartir estas aseveraciones, este índice tiene una curiosa propiedad: atribuye una mayor importancia a los cambios distributivos que tienen lugar en el centro (moda) de la distribución (en nuestro caso, meses de demanda media anual), y da un peso simétrico a los cambios en los extremos (i.e. meses de mayor y menor demanda). En este sentido, parecería razonable no estar necesariamente de acuerdo con este funcionamiento automático. Como alternativa es posible que los investigadores aprecien las ventajas de medidas que son especialmente sensibles a los cambios que ocurren en la parte baja de la distribución y, por tanto, en nuestro caso a los meses de menor demanda.

Análisis del turismo español, 2002–2019.

La demanda de turismo nacional hace referencia a los viajes que realizan los residentes de un país, en este caso España. A su vez podemos dividir esta demanda en dos bloques, el primero de ellos la demanda de turismo interna compuesta por los residentes que realizan viajes dentro de las fronteras del propio país, y en segundo lugar, la demanda emisora que engloba a las personas residentes en un país que viajan fuera de las fronteras de este [19, 54]. El turismo interior presenta una cifra significativamente superior, en el año 2019 supuso un 89,62% del total. Esto muestra la tendencia de los españoles a viajar dentro de las propias fronteras del país, y se pueden encontrar múltiples razones por las que esto es así, una de ellas puede ser la existencia de segundas residencias dentro del territorio nacional y otra que, a diferencia del turismo receptor el cual principalmente demanda turismo de sol y playa, los turistas nacionales optan en mayor medida por otros tipos de turismo como el rural, cultural o viajes cortos de fin de semana.

El año 2002, fue de incertidumbre y crisis global y al igual que las otras variables analizadas, el turismo nacional también presenta comportamientos negativos, en especial el turismo emisor que se ve más afectado por la incertidumbre comentada, disminuye más de un 7%. Durante los años siguientes el turismo nacional presenta ya tasas de crecimiento positivas, alcanzando una cifra récord en el año 2005 [63]. Los efectos de la crisis que vive el país en los años 2007 y 2008 se hacen latentes en el turismo nacional unos años más tarde, de hecho, en el año 2008 se alcanza la mayor cifra para el turismo nacional hasta la fecha [58, 64]. Pero a partir de este año, el turismo nacional se ve realmente afectado y cae prácticamente todos los años, con una tasa de variación negativa del 10,33% hasta el año 2014.

Una vez superada esta crisis y hasta el año 2019, el turismo de las personas residentes en España ha crecido casi de manera anual, incrementándose un 28% en cinco años o lo que es lo mismo 42 millones de turistas más. Si desglosamos en tipos de turismo, el emisor ha crecido en gran medida llegando a tasas de crecimiento cercanas al 12% en el año 2018, sin embargo, el turismo dentro de

las fronteras ha ido ralentizando su crecimiento ya que desde del año 2015 su crecimiento no supera el 1% ningún año, a pesar de ello sigue teniendo un peso muy superior al turismo emisor [19] representando en el año 2019 prácticamente el 90% del total del turismo nacional. Los datos básicos para el análisis provienen de la extracción de información a partir de la Encuesta de Ocupación Hotelera, realizada periódicamente por el INE. La variable de demanda turística seleccionada han sido las estacionalidades hoteleras y cubren el período 2002-2019, máximo disponible en la actualidad. De esta manera, el período nos permite examinar la evolución a medio plazo de la estacionalidad en España, más allá de los posibles vaivenes anuales. En este sentido, estas estacionalidades suponen la mayoría de las producidas en las provincias. Se han seleccionado todas las comunidades españolas. En particular, se pretende realizar un análisis global de la concentración temporal de la actividad turística en España a nivel territorial, sin detallar en exceso el análisis para cada una de ellas, en primer lugar, para salvar espacio y, en segundo lugar, para concentrarnos en las observaciones de patrones generales.

Metodología

Las técnicas de Inteligencia Artificial se basan en el aprendizaje a partir de los datos y de su semejanza con un pensamiento estructurado similar al comportamiento humano. Existen varias técnicas sugeridas dentro de esta rama de la Inteligencia Artificial como son redes neuronales, máquinas de soporte vectorial (SVM), algoritmos genéticos, sistemas de inducción de reglas, árboles de decisión o teoría de rough set [23, 53, 55, 56]. Específicamente la Red Neuronal Artificial Perceptrón Multicapa (RNAPM) ha demostrado ser una de las arquitecturas más útiles en la resolución de problemas de pronóstico. Esto es debido, fundamentalmente, a su capacidad como aproximador universal. La arquitectura de esta red se caracteriza porque tiene sus neuronas agrupadas en capas de diferentes niveles. Cada una de las capas está formada por un conjunto de neuronas y se

distinguen tres tipos de capas diferentes: la capa de entrada, la capa de salida y la capa oculta [11, 12, 59].

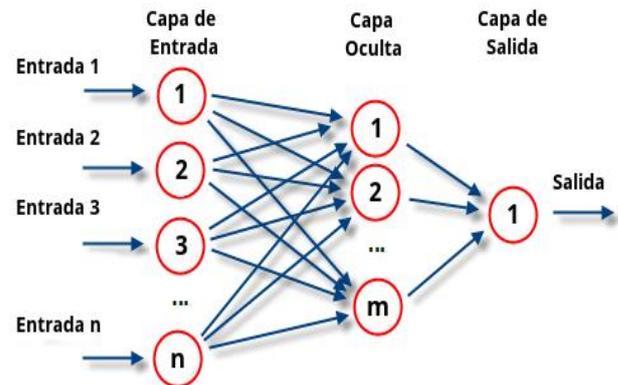


Figura 1. Arquitectura de una Red Neuronal Artificial perceptrón multicapa.

En la Figura 1, se presenta el funcionamiento de este tipo de redes. A partir de las variables de entrada, se generan con ellas un importante número de funciones no lineales, de modo que los valores de estas funciones sean a su vez valores de entrada en la Red. Esto hace que el número de nodos de entrada aumente extraordinariamente. Sin embargo, si se han utilizado las funciones adecuadas, es muy posible, que exista prácticamente una combinación lineal, entre las entradas y las salidas. Si existe esa quasi-combinación lineal, se puede eliminar la capa oculta de forma, que se pueden atenuar significativamente los problemas derivados de la memorización, o de la necesidad de disponer de un número muy elevado de vectores de aprendizaje. Gráficamente, la arquitectura de la red sería la siguiente:

Para facilitar el análisis y procesamientos de los datos se ha optado por utilizar los datos del Instituto Nacional de Estadística (INE) [25]. Estos datos se componen de varias series temporales con información sobre entradas a España y preferencias de gasto de turistas. La presente investigación es de tipo longitudinal, ya que se han recabado los datos en diferentes puntos del tiempo, para la realización inferencias acerca de la evolución, y de tipo de diseño de tendencia ya que se han analizado los cambios a través del tiempo dentro de la población estudiada. [16, 41]. La muestra está constituida por la

afluencia turística (nacional e internacional), de enero del 2002 hasta diciembre de 2019.

Para seleccionar el modelo de red que más se ajusta a los datos, se analizan las diferentes alternativas de arquitecturas de perceptrón multicapa y se evalúan según su rendimiento, en función del menor error cuadrático medio y mayor coeficiente de determinación. Para ello, los parámetros a varias son: número de capas ocultas, cantidad de neuronas dentro de las capas ocultas, las funciones de activación de las capas ocultas, división de la muestra para el entrenamiento y prueba y el método para optimizar la función de rendimiento en el entrenamiento. El modelo quedará validado si las métricas de la RNAPM son superiores a las del Modelo Lineal utilizado por la comunidad y se procede a pronosticar el por ciento de ocupación de este.

Resultados

Para dar respuesta a las preguntas de este trabajo sobre la estacionalidad y sus implicaciones. Se ha iniciado desde el procesamiento y organización de los datos donde se utilizó el software SPSS. El análisis e interpretación estadística de los datos para la presente investigación. Además, se hizo uso del análisis de redes neuronales artificiales del modelo de perceptrón multicapa.

El modelo está diseñado según los registros de cuatro variables de entrada o independientes (Nacional, Extranjero, PIB y Consumo) y una variable de salida o dependiente (Comunidad). Al considerar los parámetros de la Tabla 1 se obtuvieron distintas arquitecturas de redes neuronales con el objetivo de seleccionar en base al coeficiente de determinación la de mejor desempeño para la predicción. La RNAPM con la cual se obtuvo un mejor rendimiento fue de un 55.2% de entrenamiento, 32.1% de prueba y 12.7% de reserva.

Tabla 1. Resumen del procesamiento de los casos RNA

Conjunto de datos		N	Porcentaje
Muestra	Entrenamiento	170	55,2%
	Prueba	99	32,1%
	Reserva	39	12,7%
Válidos		308	100,0%
Excluidos		0	
Total		308	

Fuente: *Elaboración propia*

La Tabla 2, de información de red muestra información sobre la red neuronal y resulta útil para garantizar que las especificaciones son correctas. Aquí se ve en particular que: El número de unidades en la capa de entrada es el número de covariables más el número total de niveles de factor. Se crea una unidad independiente para cada categoría de Nacional y ninguna de las categorías se considera como una unidad "redundante", como es habitual en muchos procedimientos de creación de modelos. De igual manera, se crea una unidad de resultado independiente para cada categoría de "Previously defaulted", para un total de dos unidades en la capa de resultados. La selección de arquitectura automática ha elegido cuatro unidades en la capa oculta. El resto de la información de red se toma por defecto para el procedimiento.

Tabla 2. Información sobre la configuración de la red neuronal.

Capas	Descripción	Variables
Entrada	Covariables	Nacional
		Extranjero
		PIB_
		Consumo
	Número de unidades	4
	Método de cambio de escala de las covariables	Tipificados
Ocultas	Número de capas ocultas	1
	Número de unidades de la capa oculta 1 ^a	4
	Función de activación	Sigmoide
Salida	Variables dependientes	Comunidad
	Número de unidades	17
	Función de activación	Softmax
	Función de error	Entropía cruzada
^a Sin incluir la unidad de sesgo.		

Fuente: Elaboración propia.

En la Tabla 3. Se muestra el resumen del modelo muestra información sobre los resultados de entrenar y aplicar la red final a la muestra reservada. El error de entropía cruzada 62,92% se muestra porque la capa de resultados usa la función de activación softmax. Ésta es la función de error que la red intenta minimizar durante el entrenamiento. El porcentaje de pronósticos incorrectos de 12,9% se toma de la tabla de clasificación. El algoritmo de estimación se ha detenido ya que se ha alcanzado el número máximo de épocas. Lo ideal es que el entrenamiento se detenga, puesto que el error ha convergido. Esto plantea cuestiones sobre si se ha producido algún error durante el entrenamiento, y se debe tener en cuenta al realizar una inspección adicional de los resultados.

Los resultados del entrenamiento y prueba de la red neuronal mostrados en la Tabla 3 facilitan que se detecte la tasa de aciertos que se calcula restando el número de aciertos incorrectos al total de los datos (100%). De esta

manera, para el entrenamiento de la red neuronal existe un 87,10% de precisión, mientras que en los datos de prueba que se usan para comprobar la efectividad del modelo obtenido en el entrenamiento siendo su tasa de acierto el 86,90%. Por otra parte, la curva ROC proporciona una representación visual de la susceptibilidad y especificidad para todos los cortes posibles en un único gráfico, que es mucho más limpio y potente que una serie de tablas.

Tabla 3. Resumen del modelo

Capa	Descripción	Valores
Entrenamiento	Error de entropía cruzada	62,92
	Porcentaje de pronósticos incorrectos	12,9%
	Regla de parada utilizada	1 pasos consecutivo s sin disminución del error ^a
	Tiempo de entrenamiento	00:00:00,200
Prueba	Error de entropía cruzada ^a	48,136
	Porcentaje de pronósticos incorrectos	13,1%
Reserva	Porcentaje de pronósticos incorrectos	23,1%

Variable dependiente: Comunidad

^a Los cálculos del error se basan en la muestra de prueba.

Fuente: Elaboración propia

Tabla 4. Área debajo de la curva

Comunidad	Área
Andalucía	1,000
Aragón	0,999
Asturias	0,986
Baleares	1,000
Canarias	1,000
Cantabria	0,983
Castilla de la Mancha	0,998
Castilla y León	0,982
Cataluña	1,000
Extremadura	0,995
Galicia	0,986
La Rioja	0,989
Madrid	1,000
Murcia	0,968
Navarra	0,990
País Vasco	0,997
Valencia	1,000

Fuente: Elaboración propia.

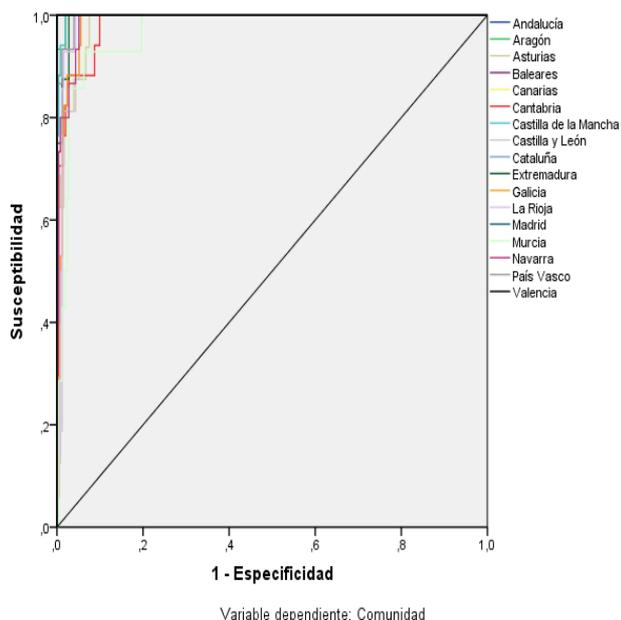


Figura 2. Curva ROC

Fuente: Elaboración propia.

La Figura 2, que se muestra aquí presenta dos curvas, una para la categoría “No” y una para la categoría “Sí”.

Debido a que sólo hay dos categorías, las curvas son simétricas con respecto a una línea de 45 grados (no se muestra) que va de la esquina superior izquierda del gráfico a la esquina inferior derecha. Este gráfico se basa en las muestras de prueba y entrenamiento combinadas. El área bajo la curva (Tabla 4) es un resumen numérico de la curva ROC, y los valores de la tabla representan, para cada categoría, la probabilidad de que la pseudo probabilidad pronosticada de estar en esa categoría sea mayor para un caso elegido aleatoriamente en esa categoría que para un caso elegido aleatoriamente que no pertenezca a esa categoría. Mientras que el área bajo la curva es un resumen de una estadística de la precisión de la red útil, debe poder elegir un criterio específico por el que clasificar a los clientes. El gráfico de pronosticados por observados proporciona un punto de partida visual sobre este proceso.

La importancia de una variable independiente (Tabla 5) es una medida que indica cuánto cambia el valor pronosticado por el modelo de la red para diferentes valores de la variable independiente. La importancia normalizada es el resultado de los valores de importancia divididos por los valores de importancia mayores expresados como porcentajes.

Tabla 5. Importancia de las variables independientes

Variables	Importancia	Importancia normalizada
Nacional	0,336	100,0%
Extranjero	0,286	85,0%
PIB_	0,249	74,2%
Consumo	0,129	38,2%

Fuente: Elaboración propia

El gráfico de importancia es un gráfico de barras de los valores de la tabla anterior, clasificado en valor de importancia descendente.

La Figura 3, de importancia muestra que los resultados están dominados por el procedimiento quirúrgico realizado, seguido de si han surgido complicaciones, seguidos a distancia por otros predictores. La importancia del procedimiento quirúrgico puede apreciarse claramente en los gráficos de Turista nacional; algo menos en Turista Extranjero, aunque el efecto de las complicaciones sobre Duración de la estancia parece ser visible en los turistas con las mayores duraciones de las estancias observadas.

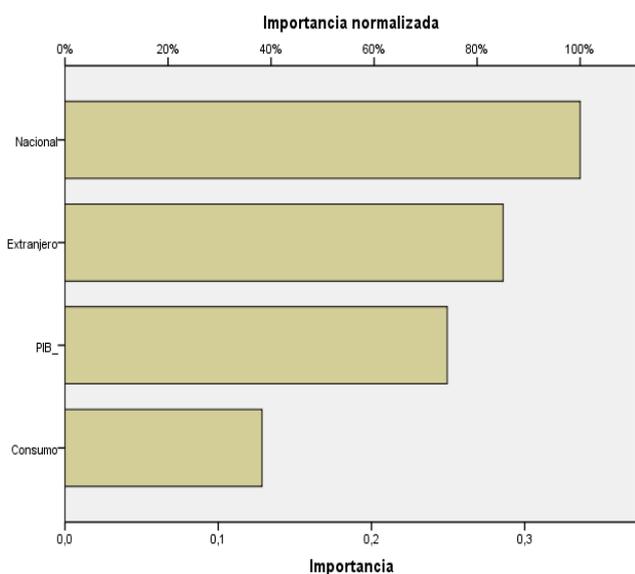


Figura 3. Importancia de la variable independiente

Fuente: Elaboración propia

Discusión

Las cifras de turismo nacional en Cantabria adquieren más relevancia cuando se comparan con las de sus competidores directos: los destinos de la España Verde. Cantabria, Asturias, Galicia y País Vasco se disputan un mercado ofreciendo productos turísticos similares, y teniendo que competir a pesar de no hacerlo con las mismas condiciones de tamaño del territorio y volumen de población. Cantabria, siendo el territorio más pequeño, ofrece unos resultados turísticos, en valores absolutos, comparables a los registrados en el resto de las regiones del norte y que, en términos relativos, superan a las otras tres comunidades en casi todos los análisis. La estancia en sus alojamientos hoteleros significó un valor medio de

2,35 noches/turista, cifra superior a los 2,03 de Asturias, 1,91 de Galicia y 1,90 del País Vasco [32].

En definitiva, el comportamiento del turista es una pieza clave en el estudio de la estacionalidad, ya que, como se aprecia, está vinculado a diversas perspectivas de estudio, siendo complejo su delimitación por separado. El presente trabajo es un ejemplo de ello debido a que abarca el estudio de los efectos que provoca el patrón estacional sobre la percepción del turista experimentado, con objeto de mejorar su satisfacción y, con ello, favorecer el desarrollo sostenible del destino turístico, que debe ser un propósito prioritario y permanente de los gestores de dicho destino [32].

Los factores determinantes de la estacionalidad son variados. Como se señala en [4], el trabajo seminal de Baron [7] distingue entre estacionalidad natural y estacionalidad institucional en función de las causas que la originan estando relacionadas básicamente con el tiempo atmosférico y con los valores culturales, respectivamente. En este sentido [9], identifica cinco factores principales como son: (1) la climatología; (2) las decisiones personales en el ámbito de lo social, religioso o cultural; (3) la presión social o la moda; (4) las temporadas deportivas; y (5) la inercia o la tradición. En [8] incorporan un nuevo factor determinante a las cuatro causas indicadas anteriormente que son las restricciones del lado de la oferta como pueden ser la disponibilidad laboral, es decir, los trabajadores pueden tener un menor interés en estar ocupados en períodos estivales porque coinciden con las vacaciones escolares [33].

Los efectos que provocan las variaciones estacionales también son muy diversos. Se pueden citar los siguientes: (1) efectos económicos que se centran especialmente sobre la pérdida de rentabilidad y la ineficiencia de los recursos [20, 51, 21]; (2) efectos laborales que constituyen una categoría independiente a la anterior ya que adquieren una gran relevancia como para ser considerados de forma autónoma [6, 29]; (3) efectos ecológicos que incluyen por ejemplo el deterioro de la vegetación, los trastornos en la fauna y la erosión física del territorio [35] (4) efectos socioculturales, que están relacionados principalmente con el impacto sobre la

comunidad receptora de turismo, aunque abarcan también a los visitantes [30, 33, 61].

En este trabajo se ha examinado el turismo estacional en las provincias españolas a lo largo del período 2002-2019 y tomando como indicador de referencia las pernoctaciones hoteleras. Para ello en primer lugar se incide en la relevancia fundamental de las medidas de concentración a analizar, habida cuenta de su sensibilidad respecto a los resultados. En este punto, el trabajo recomienda el uso de diferentes medidas satisfactorias al mismo tiempo. En el apartado empírico el análisis muestra diferentes puntos de interés: En primer lugar, los niveles de concentración temporal son muy variados a lo largo de la geografía provincial. En todo caso, se destacan, como las provincias unánimemente más concentradas las Baleares y las costeras catalanas, exceptuando el caso de la provincia de Barcelona. En el polo contrario, se destaca también una baja estacionalidad en las provincias canarias o en Madrid.

En segundo lugar, y respecto a la evolución, el ascenso global de la concentración en el período, y que ha sido, por ejemplo, del 5.7% según el coeficiente de variación, se reparte de manera muy diversa a lo largo del territorio. Provincias de elevada concentración como Baleares y las costeras andaluzas (excepto Granada, presumiblemente por su potente producto patrimonial) han registrado incluso aumentos superiores a la media nacional en el período. En contraste, otras provincias de concentración elevada como Tarragona o Girona, gallegas y cantábricas han revelado un patrón positivo en este sentido. En tercer lugar, la observación con detenimiento de los resultados en función de las medidas de concentración utilizadas refleja algunos detalles a mejorar, aunque en su globalidad no son especialmente significativas.

Finalmente, se ha examinado, como punto de partida para modelos de explicación superiores, el grado de correlación entre una serie de factores potencialmente explicativos y las variaciones provinciales en la concentración temporal de la demanda turística. En particular, se han probado variables relacionadas con la magnitud de la demanda, el tipo cualitativo de producto, el área climática y el peso del mercado nacional. En

síntesis, los resultados iniciales ilustran, por ejemplo, que típicamente el montante de demanda global, ser provincia costera, estar especializada en sol y playa o formar parte de la región climática mediterránea son rasgos asociados a una mayor concentración temporal de la actividad turística. Por el contrario, ser provincia interior, pertenecer a la zona climática Sur, ser provincia canaria o con importante peso del mercado nacional son rasgos que tienden a atenuar este desequilibrio.

Conclusiones

Al hilo de las preguntas establecidas en la investigación se puede concluir que las redes neuronales han obtenido un modelo predictivo con especial enfoque en turistas nacionales debido a la forma latente de hacer turismo ya que tienen mayor presencia durante todos los meses del año. De forma análoga, en [57] se detectó que los turistas autóctonos en ciudades con calidad medioambiental adecuada tuvieron mayor preferencia debido al diseño y la morfología del sitio turístico. El modelo predictivo de redes neuronales además destacó la fuerte aparición de turistas extranjeros en periodos estivales, es así que de forma similar [55] estudió en su modelo de predicción basado en el histórico de apariciones y recomendaciones que efectuó el turista para de esta manera recomendar otros o mejores sitios visitados por turistas con el mismo patrón de comportamiento. Las redes neuronales consiguieron detectar las comunidades localizadas en la costa mediterránea e Islas Baleares y Canarias que fueron mayormente concurridas y preferidas para los turistas. Sin duda, las redes neuronales terminan siendo una herramienta de gran soporte para enfocarse en la predicción de turistas.

Referencias

- [1] Aditya, F., Nasution, S. M., Virgono, A. Traffic Flow Prediction Using SUMO Application with K-Nearest Neighbor (KNN) Method. *International Journal of Integrated Engineering* 2020; 12(7): 98-103.
- [2] Al-Bakri, N. F., Yonan, J. F., Sadiq, A. T., Abid, A. S. Tourism companies assessment via social media using sentiment analysis. *Baghdad Science Journal* 2022; 19(2): 422-429.
- [3] Alkan, T., Dokuz, Y., Ecemiş, A., Bozdağ, A., & Durduran, S. S. Using machine learning algorithms for predicting real estate values in tourism centers. *Soft Computing* 2022; 1-13.

- [4] Allcock, J. Seasonality. En *Tourism Marketing and Management Handbook*. S. Witt y L. Moutinho (eds.), Prentice Hall 1994, New York pp; 191-208.
- [5] Amato, G., Falchi, F. kNN based image classification relying on local feature similarity. In *Proceedings of the Third International Conference on Similarity Search and Applications 2010*: 101-108.
- [6] Ashworth, J.; Thomas, B. Patterns of seasonality in employment in tourism in the United Kingdom. *Applied Economics Letter* 6 1999; (11): 735-739.
- [7] Baron, R.V. (1975): *Seasonality in Tourism-A Guide to the Analysis of Seasonality and Trends for Policy Making*, Technical Series 1975; 2, Economist Intelligence Unit, London.
- [8] Baum, T.; Hagen, L., Responses to seasonality: the experiences of peripheral destinations. *International Journal of Tourism Research* 1999; 1(5): 299-312.
- [9] Butler, R. Seasonality in tourism: issues and problems. En *Tourism. The State of the Art*. A. Seaton (Edit.) Wiley Chichester 1999: 332-340
- [10] BUTLER, R.W., «Seasonality in tourism: Issues and implications», en BAUM, T. y LUNDTORP, S. (Eds.): *Seasonality in tourism 2001*; Pergamon-Elsevier, Oxford: 5-22.
- [11] Carneiro, J., Meira, J., Novais, P., & Marreiros, G. (2021). Using machine learning to predict the users ratings on tripadvisor based on their reviews. In *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Social Good. The PAAMS Collection: International Workshops of PAAMS 2021*, Salamanca, Spain, October 6-9, 2021, Proceedings 19: 127-138.
- [12] Chadha, S., Mittal, S., & Singhal, V., An Insight of Script Text Extraction Performance using Machine Learning Techniques. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* 2019; 9(1).
- [13] COWELL, F. *Measuring Inequality*, Second edition. New York: Prentice Hall.
- [14] Duro, J.A. (2008): «La concentración temporal de la demanda turística en España y sus regiones: un análisis empírico a partir de índices de desigualdad», *Revista de Análisis Turístico* 2008; 6: 36-48.
- [15] Duro, J. A., Farré, F. X., Estacionalidad turística en las provincias españolas: medición y análisis. *Cuadernos de Turismo* 2015; (36): 157-174.
- [16] Hernández, R., Fernández, C., Baptista, P., *Metodología de la Investigación* 2010. México D.F.: McGraw-HILL / Interamericana Editores, S.A. de C.V.
- [17] Fernández-Morales, A., «Decomposing seasonal concentration», *Annals of Tourism Research* 2003; 30 (4): 942-956.
- [18] Fernández-Morales, A. y Cruz-Mayorga, M.C. «Seasonal concentration of the hotel demand in Costa del Sol: A decomposition by nationalities», *Tourism Management* 2008; 29: 940-949.
- [19] Flores Ruiz, D., Bago Sotillo, E., & Barroso González, M. de la O. Comportamiento del turismo nacional y crecimiento en España. *Revista Investigaciones Turísticas* 2018, 16, 68-86. <https://doi.org/10.14198/inturi2018.16.04>
- [20] Georgantzis, N.C. Cyprus' hotel value chain and profitability. *System Dynamics Review* 2003; 19(3):175-212
- [21] Getz, D.; Nilsson, P.A. Responses of family businesses to extreme seasonality in demand: case of Bornholm, Denmark. *Tourism Management* 2004; 25(1): 17-30.
- [22] GINI, C., *Variabilità e mutabilità*, C.Cuppini, Bologna 1912.
- [23] Gutiérrez Cordero, M. de L., Segovia-Vargas, M. J., y Escamilla, M. R. Análisis del Riesgo de Caída de Cartera en Seguros: Metodologías de “Inteligencia Artificial” vs “Modelos Lineales Generalizados.” *Economía Informa* 2017; 407: 56-86. <https://doi.org/https://doi.org/10.1016/j.ecin.2017.11.004>
- [24] Hapsari, I., & Surjandari, I. Visiting time prediction using machine learning regression algorithm. In *2018 6th International Conference on Information and Communication Technology (ICoICT) 2018*: 495-500.
- [25] Instituto de estudios turísticos (IET). Balance del turismo año 2009 [Balantur] 2010. [http://estadisticas.tourspain.es/es-ES/estadisticas/analisis/turistico/balantur/anales/Balance turismo en España en 2009.pdf](http://estadisticas.tourspain.es/es-ES/estadisticas/analisis/turistico/balantur/anales/Balance%20turismo%20en%20Espa%C3%B1a%20en%202009.pdf)
- [26] KOENIG-LEWIS, N. y BISCHOFF, E. (2005): «Seasonality research: the state of the art», *International Journal of Tourism Research* 2005; 7: 201-219.
- [28] Kokate, S., Gaikwad, A., Patil, P., Gutte, M., Shinde, K., *Traveler's Recommendation System Using Data Mining Techniques*. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA) 2018*: 1-5.
- [29] Krakover, S. Partitioning seasonal employment in the hospitality industry. *Tourism Management* 2000; 2(3):461-471
- [30] Kuvan, Y.; Akan, P. Resident's attitudes toward general and forest-related impacts of tourism: the case of Belek, Antalya. *Tourism Management* 2005; 26(5):691-706.
- [31] Leal, F., Malheiro, B., & Burguillo, J. C. Trust and reputation modelling for tourism recommendations supported by crowdsourcing. In *Trends and Advances in Information Systems and Technologies 2018*;1 (6) :829-838. Springer International Publishing.
- [32] López, J. y López, L. La concentración estacional en las regiones españolas desde una perspectiva de la oferta turística, *Revista de Estudios Regionales* 2006, 77, pp. 77-104.
- [33] López Bonilla, J. M., López Bonilla, L. M. Variabilidad estacional del mercado turístico en Andalucía. *Estudios y perspectivas en turismo* 2007; 16(2): 150-172.
- [34] Lundtorp, S. «Measuring tourism seasonality» en: BAUM, T. y LUNDTORP, S. (Eds.), *Seasonality in tourism 2001*: 23-50, Pergamon-Elsevier, Oxford.
- [35] Lusseau, D.; Higham, J.E.S. Managing the impacts of dolphin-based tourism through the definition of critical habitats: the case of bottlenose dolphins (*Tursiops* spp.) in Doubtful Sound, New Zealand. *Tourism Management* 2004; 25(6): 657-667.
- [36] Mariwa, S. O., Tunduny, T. K. a Web-Based Application for Making Low-Cost Vacation Reservations for Tourists Using the K-Nearest Neighbors Algorithm. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 2021; 48, 67-76.
- [37] Martín, J.M., Jiménez, J. y Molina, V., (2014): Impacts of seasonality on environmental sustainability in the tourism sector based on destination type: an application to Spain's Andalusia region, *Tourism Economics* 2014, 20 (1), pp. 123-142.

- [38] Meira, J., Carneiro, J., Bolón-Canedo, V., Alonso-Betanzos, A., Novais, P., & Marreiros, G. Anomaly detection on natural language processing to improve predictions on tourist preferences. *Electronics* 2022, 11(5), 779.
- [39] Michalska-Dudek, I., Dudek, A. Evaluation of Quality of Neural Network Models and Discriminant Analysis in ROPO Forecasting. In: Jajuga, K., Dehnel, G., Walesiak, M. (eds) *Modern Classification and Data Analysis. SKAD 2021. Studies in Classification, Data Analysis, and Knowledge Organization.* Springer, Cham. https://doi.org/10.1007/978-3-031-10190-8_14
- [40] Nugraha, N. B., & Alimudin, E. (2020). Mobile application development for tourist guide in Pekanbaru City. In *Journal of Physics: Conference Series* 2020; 1430(1): 012038. IOP Publishing.
- [41] UNWTO, *World Tourism Barometer* 2010; 8(2), http://www.unwto.org/facts/eng/pdf/barometer/UNWTO_Barom10_2_en_excerpt.pdf
- [42] Uriel Jiménez, E. Precios y competitividad en el sector turístico, en Uriel JIMÉNEZ, E.; HERNÁNDEZ MARTÍN, R. (Coord.): *Análisis y tendencias del turismo Pirámide* 2004:119-138.
- [43] Ortega Aguaza, B. Determinants of regional labour productivity growth: A study for the hospitality sector in Spain. *INVESTIGACIONES REGIONALES-Journal of REGIONAL RESEARCH* 2013; (25): 89-110.
- [44] Ozaslan, I. N., Degirmenci, A., Karal, O. Tourism Demand Forecasting for Turkey by Using Adaboost Algorithm. In *2022 Innovations in Intelligent Systems and Applications Conference (ASYU) 2022*: 1-5.
- [45] Perelli, O. El turismo en Madrid: el reto de un proyecto colectivo. *Economistas* 2005; 23(104): 326-334.
- [46] Popescu, A., Grefenstette, G., Moëllic, P. A. Mining tourist information from user-supplied collections. In *Proceedings of the 18th ACM conference on Information and knowledge management 2009*: 1713-1716.
- [47] Puspasari, S. Machine Learning for Exhibition Recommendation in a Museum's Virtual Tour Application. *International Journal of Advanced Computer Science and Applications* 2022; 13(4).
- [48] Rahman, M. M., Zaki, Z. B. M., Alwi, N. H. B. M., Monirul Islam, M. (2019). A hybrid approach to improve recommendation system in E-tourism. In *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2018*; 1: 787-797.
- [49] Riswanto, E., RobiIn, B. (2019, May). Mobile Recommendation System for Culinary Tourism Destination using KNN (K-nearest neighbor). In *Journal of Physics: Conference Series* 2019; 1201(1): 012039.
- [50] Roy, P., Setu, J. H., Binti, A. N., Koly, F. Y., Jahan, N. Tourist Spot Recognition Using Machine Learning Algorithms. In *Intelligent Communication Technologies and Virtual Mobile Networks: Proceedings of ICICV 2022*: 99-110.
- [51] Roselló, J.; Riera, A.; Sausó, A. The economic determinants of seasonal patterns. *Annals of Tourism Research* 2004; 31(3):697- 71.
- [52] Rubi, M. A., Bijoy, M. H. I., Chowdhury, S., Islam, M. K. Machine Learning Prediction of Consumer Travel Insurance Purchase Behavior. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT) 2022*: 1-5.
- [53] Saito, N., Ogawa, T., Asamizu, S., Haseyama, M. A tourism category classification method based on estimation of reliable decision. In *2016 IEEE 5th Global Conference on Consumer Electronics 2016*: 1-2
- [54] SECRETARIA GENERAL DE TURISMO. *Plan del Turismo Español. Horizonte 2020* Ministerio de Industria, Comercio y Turismo 2007.
- [55] Shah, S., Thakkar, A., & Rami, S. A novel approach for making recommendation using skyline query based on user location and preference. *Indian Journal of Science and Technology* 2016; 9(30).
- [56] Shrestha, D., Wenan, T., Gaudel, B., Shrestha, D., Rajkarnikar, N., Jeong, S. R., Preliminary analysis and design of a customized tourism recommender system. In *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2021*: 541-561.
- [57] Tapak, L., Abbasi, H., & Mirhashemi, H. Assessment of factors affecting tourism satisfaction using K-nearest neighborhood and random forest models. *BMC research notes* 2019; 12(1): 1-5.
- [58] Valiente, G. C., Forga, J. M. P., & Romero, A. B. (2016). Turismo en España, más allá del sol y la playa. Evolución reciente y cambios en los destinos de litoral hacia un turismo cultural. *Boletín de La Asociación de Geógrafos Españoles* 2016; 71: 431-454. <https://doi.org/10.21138/bage.2289>
- [59] Villada, F., Arroyave, D., y Villada, M. Pronóstico del Precio del Petróleo mediante Redes Neuronales Artificiales. *Información Tecnológica* 2014; 25: 145-154. Recuperado desde https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-07642014000300017&nr=iso
- [60] WANHILL, S., Tackling seasonality: A technical note, *International Journal of Tourism Management* 1980; 1(4): 84-98.
- [61] Waitt, G., Social impacts of Sydney Olympics. *Annals of Tourism Research* 2003; 30 (1): 194-215.
- [62] Wang, S. Hate crime analysis based on artificial intelligence methods. In *E3S Web of Conferences* 2021; 251: 01062.
- [63] Wenan, T., Shrestha, D., Gaudel, B., Rajkarnikar, N., Jeong, S. R., analysis and evaluation of TripAdvisor data: a case of Pokhara, Nepal. In *Intelligent Computing & Optimization: Proceedings of the 4th International Conference on Intelligent Computing and Optimization 2021 (ICO2021)*; 3: 738-750.
- [64] Yuensuk, T., Limpinan, P., Nuankaew, W. S., Nuankaew, P., Information Systems for Cultural Tourism Management Using Text Analytics and Data Mining Techniques. *International Journal of Interactive Mobile Technologies* 2022, 66(8).