



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

Aplicación de Redes
Neuronales Artificiales en la
predicción de la demanda
turística de cuartos de hotel
en México

TESIS

Que para obtener el título de
Ingeniero Industrial

PRESENTA

Jaime Yael Pérez Díaz

DIRECTOR DE TESIS

Dr. Wulfrano Gómez Gallardo



Ciudad Universitaria, Cd. Mx., 2019



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

Introducción	5
Planteamiento del problema	6
Justificación.....	6
Objetivos	6
Objetivo general	6
Objetivos específicos	6
Hipótesis	7
Resumen de los capítulos	9
Capítulo 1: Turismo como factor importante para el desarrollo del país y factores que influyen en la demanda turística.....	10
Capítulo 2: Descripción de modelos convencionales de pronóstico	19
Capítulo 3: Antecedentes y aplicación de las redes neuronales artificiales	28
Definición.....	28
Neurona biológica vs Neurona artificial	29
Historia.....	31
Aplicaciones	33
Diseño de una red neuronal artificial	35
Elementos	35
Arquitecturas.....	37
Entrenamiento, Validación y Desarrollo.....	39
Aprendizaje: Propagación hacia atrás (Backpropagation)	43
Ventajas y desventajas.....	45
Capítulo 4: Construcción del modelo con redes neuronales artificiales y su comparación con las técnicas convencionales de pronóstico	47
Conclusiones	69
Bibliografía.....	71
Anexos y código	74
Código.....	85

Índice de tablas

<i>Tabla 1: Factores para determinar la demanda turística de un país</i>	14
<i>Tabla 2: Principales medidas de error</i>	27
<i>Tabla 3: Aplicación de las redes neuronales por área de conocimiento</i>	34
<i>Tabla 4: Arquitecturas de redes neuronales artificiales</i>	38
<i>Tabla 5: Tipos de aprendizaje para redes neuronales</i>	41
<i>Tabla 6: Funciones de estandarización de datos</i>	42
<i>Tabla 7: Etiqueta de las variables del modelo</i>	53
<i>Tabla 8: Error de los modelos HW y SARIMA</i>	59
<i>Tabla 9: Constantes de suavizado</i>	60
<i>Tabla 10: Coeficientes del modelo HW</i>	61
<i>Tabla 11: Coeficientes de método SARIMA</i>	62
<i>Tabla 12: Coeficientes del modelo de regresión lineal múltiple</i>	64
<i>Tabla 13: Comparación de errores</i>	66
<i>Tabla 14: Principales funciones de activación en el diseño de redes neuronales</i>	76
<i>Tabla 15: Resultado de los pesos de la red neuronal artificial por tipo de interacción</i>	79
<i>Tabla 16: Coeficientes de suavizado HW para las variables de entrada</i>	79
<i>Tabla 17: Parámetros del modelos SARIMA para las variables de entrada</i>	80
<i>Tabla 18: Pronósticos de las variables de entrada para el año 2019</i>	80
<i>Tabla 19: Valores puntuales de los pronósticos con diferentes métodos</i>	82

Índice de ilustraciones

Ilustración 1: Metodología	8
Ilustración 2: Índice de competitividad turística y de viajes	11
Ilustración 3: Posición internacional de los países por ingresos turísticos	12
Ilustración 4: Posición de México a nivel mundial por número de llegadas de turistas	12
Ilustración 5: Principales variables turísticas dependientes	16
Ilustración 6: Principales variables turísticas independientes	16
Ilustración 7: Número de hoteles por entidad federativa	18
Ilustración 8: Componentes gráficos de una serie de tiempo	20
Ilustración 9: Anatomía simple de la neurona biológica	29
Ilustración 10: Comparación de la neurona biológica con la neurona artificial	30
Ilustración 11: Neurona de McCulloch-Pitts	31
Ilustración 12: Elementos básicos de una red neuronal artificial	35
Ilustración 13: Relación de los elementos de la neurona artificial	37
Ilustración 14: Elementos de un Perceptrón	37
Ilustración 15: Redes multicapa feedforward y feedback	38
Ilustración 16: Diagramas de flujo de los procesos de aprendizaje	41
Ilustración 17: Entrenamiento, verificación y desarrollo de una red neuronal	43
Ilustración 18: Representación esquemática de backpropagation	44
Ilustración 19: Relación de entradas y salidas en minería de datos	47
Ilustración 20: Proceso de minería de datos	48
Ilustración 21: Cuartos de hotel ocupados en la República Mexicana	51
Ilustración 22: Factores estacionales	52
Ilustración 23: Proceso de validación del modelo	55
Ilustración 24: Arquitectura del modelo de la red neuronal artificial	56
Ilustración 25: Pronóstico con Redes Neuronales Artificiales	58
Ilustración 26: Proyección de cuartos con RNA	60
Ilustración 27: Pronóstico con método Holt-Winters	62
Ilustración 28: Pronóstico con SARIMA	63
Ilustración 29: Pronóstico con Regresión Lineal Múltiple	65
Ilustración 30: Comparación de los métodos de pronóstico	67
Ilustración 31: Proyección de cuartos ocupados con diferentes métodos	68
Ilustración 32: Índice de competitividad turística y de viajes de México	74
Ilustración 33: Desglose de puntaje del índice de competitividad turística y de viajes de México	75
Ilustración 34: Efecto multiplicador del turismo	76
Ilustración 35: Aplicaciones de la minería de datos	77
Ilustración 36: Comportamiento anual de los cuartos de hotel ocupados	78
Ilustración 37: Descomposición de la serie temporal de cuartos de hotel ocupados	78
Ilustración 38: Comparación de pronóstico RNA vs HW	81
Ilustración 39: Comparación de pronóstico RNA vs SARIMA	81
Ilustración 40: Comparación de pronóstico RNA vs MRLM	82

Agradecimientos

Dedico este trabajo de investigación a mis padres, los cuales siempre me han apoyado en todas las decisiones importantes de mi vida y me han enseñado a ser un buen ser humano. Asimismo, dedico este trabajo a mi hermana, ya que es una de mis fuentes de inspiración y una prueba viviente de que los sueños se pueden realizar.

Agradezco a mi familia y a mis amigos por su apoyo incondicional y se extiende un agradecimiento al Departamento de Estadísticas de la Organización Mundial del Turismo (OMT) por facilitar el acceso a la información necesaria para la validación de este trabajo.

Introducción

La formación académica de un ingeniero industrial comprende el conocimiento y aplicación de los modelos de pronóstico, ya que son una herramienta muy poderosa que permiten tener una noción de las situaciones futuras y con ello liderar una toma de decisiones mucho más eficiente y acertada. Conforme se cursan las asignaturas del plan curricular, se profundiza en dicha comprensión y aplicación, pero es en los últimos semestres de la carrera donde se hace evidente de que los pronósticos se relacionan con muchas áreas de conocimiento o investigación que requieren de una planeación estratégica. En particular, en el sector turístico, la estimación de la demanda de infraestructura turística es de vital importancia porque ésta repercute directamente en la economía del país. Además, su conocimiento le permite a las autoridades correspondientes contar con las herramientas adecuadas para afrontar la demanda creciente de turistas e incentivar este sector.

México es un país considerado como potencia turística a nivel mundial (Crotti & Misrahi, 2018) ya que cuenta con una riqueza cultural, natural, gastronómica y de atractivos turísticos muy vasta. Asimismo, es uno de los lugares de Latinoamérica más socorrido por los turistas. Además, este sector, por su efecto multiplicador, es capaz de contribuir no solo en el aspecto económico de un país, sino también en la creación de empleos, el incremento de infraestructura y el fomento de la identidad nacional (Octavio, 2008). Es por ello que, tomando como referente estos aspectos clave, el conocimiento de la demanda turística que tendrá México en los siguientes años, es crucial para que las autoridades correspondientes tengan la información necesaria para tomar las mejores decisiones en aras de incrementar este sector.

Como parte de la propuesta de método de pronóstico más preciso, se presentan las Redes Neuronales Artificiales (ANNs, por sus siglas en inglés), las cuales son una herramienta basada en el funcionamiento del cerebro humano, en el cual se modelan las relaciones existentes en las neuronas biológicas. De acuerdo con la revisión de la literatura, las redes neuronales como método de predicción son capaces de modelar cualquier tipo de función no lineal y adaptarse lo mejor posible a su comportamiento. En este trabajo de investigación se pretende utilizar esta herramienta de predicción para pronosticar la demanda turística del número de cuartos de hotel en la República Mexicana de los principales países que mantienen una relación económica con México, en un horizonte de 6 años, ya que es el tiempo en que el poder ejecutivo permanece sin cambios en nuestro país. Para lo cual se utilizarán los datos históricos recabados de los organismos federales, tales como el INEGI, la SECTUR, la Organización Mundial del Turismo (UNWTO, por sus siglas en inglés), Banco Mundial, entre otros, para el diseño del modelo.

Asimismo, se comparará el desempeño del algoritmo de Redes Neuronales Artificiales contra el desempeño de algunas técnicas convencionales de pronóstico, como lo son las series de tiempo de Holt-Winters, los modelos SARIMA y el modelo de regresión lineal múltiple, para determinar si es conveniente emplear este algoritmo en dicha aplicación.

Planteamiento del problema

Las preguntas que se buscan responder con el desarrollo de este trabajo de investigación son: ¿La infraestructura hotelera y de servicios turísticos de la República Mexicana con los que se cuenta actualmente, son suficientes para satisfacer dicha demanda? ¿Cuál es el pronóstico para un horizonte de 6 años? ¿El desempeño de las redes neuronales artificiales se sobrepone a las técnicas convencionales de predicción?

Justificación

La importancia de este trabajo de investigación radica, principalmente, en tres aspectos. En primer lugar, resulta indispensable conocer la capacidad de infraestructura hotelera y de servicios instalada en la República Mexicana con el objeto de determinar si dichos medios son suficientes para suministrar el servicio a la creciente demanda, puesto que el número de personas que han buscado atracciones turísticas en los últimos años ha aumentado considerablemente. Además, se debe tener en cuenta que los actores en este proceso son los países, Estados Unidos y Canadá, que representan más del 85% del número total de turistas que visitan México, por lo es ampliamente recomendable su satisfacción en el entendido de mantener la relación de turismo con el país. Además, son los principales países con los que México mantiene relaciones comerciales.

Por otro lado, es importante investigar e identificar esta problemática en el sentido de incentivar aún más el turismo mexicano, pues de acuerdo con las cifras de la cuenta satélite del turismo en México de la Secretaría de Turismo, SECTUR (Secretaría de Turismo, 2018), el turismo representa el 7.1% del Producto Interno Bruto del país. Además, es una fuente de empleo que, por su efecto multiplicador, afecta positivamente a la economía del país en diferentes sectores.

Por último, la incorporación de técnicas de “machine learning” y “minería de datos” en el pronóstico de los datos turísticos, aumentará la precisión de los mismos porque dichas técnicas son capaces de adaptarse al comportamiento de los datos históricos con el menor error posible, con lo cual se permitirá que la planeación y designación de los recursos presupuestales sean distribuidos adecuadamente y existan mejores y mayores beneficios. Además, se podrá contar con una herramienta adicional para calcular pronósticos, capaz de reducir el margen de error generado entre las predicciones y los datos reales.

Objetivos

Objetivo general

Pronosticar el número de cuartos de hotel que serán demandados en el sector turístico por el mercado estadounidense y canadiense, mediante la aplicación del algoritmo de redes neuronales artificiales para comprobar las ventajas que éste método pueda tener sobre los métodos convencionales de pronóstico.

Objetivos específicos

- Conformar la base de datos que se utilizará para el pre-procesamiento de los datos para el modelo de regresión a partir de la información disponibles es sitios oficiales como: Secretaría de Turismo, Organización Mundial del Turismo, Banco de México, entre otros.
- Realizar una prueba de linealidad a la base de datos conformada para asegurar el comportamiento no lineal, y con ello la aplicación de las redes neuronales artificiales.
- De las técnicas convencionales de predicción, emplear la de menor error y mayor ajuste para alimentar las entradas del modelo de regresión no lineal no paramétrico de redes neuronales artificiales.
- Comparar las técnicas de predicción convencionales con el algoritmo de redes neuronales artificiales, a partir de las medidas de error, para evaluar su desempeño y proponer su uso en este tipo de análisis.

Hipótesis

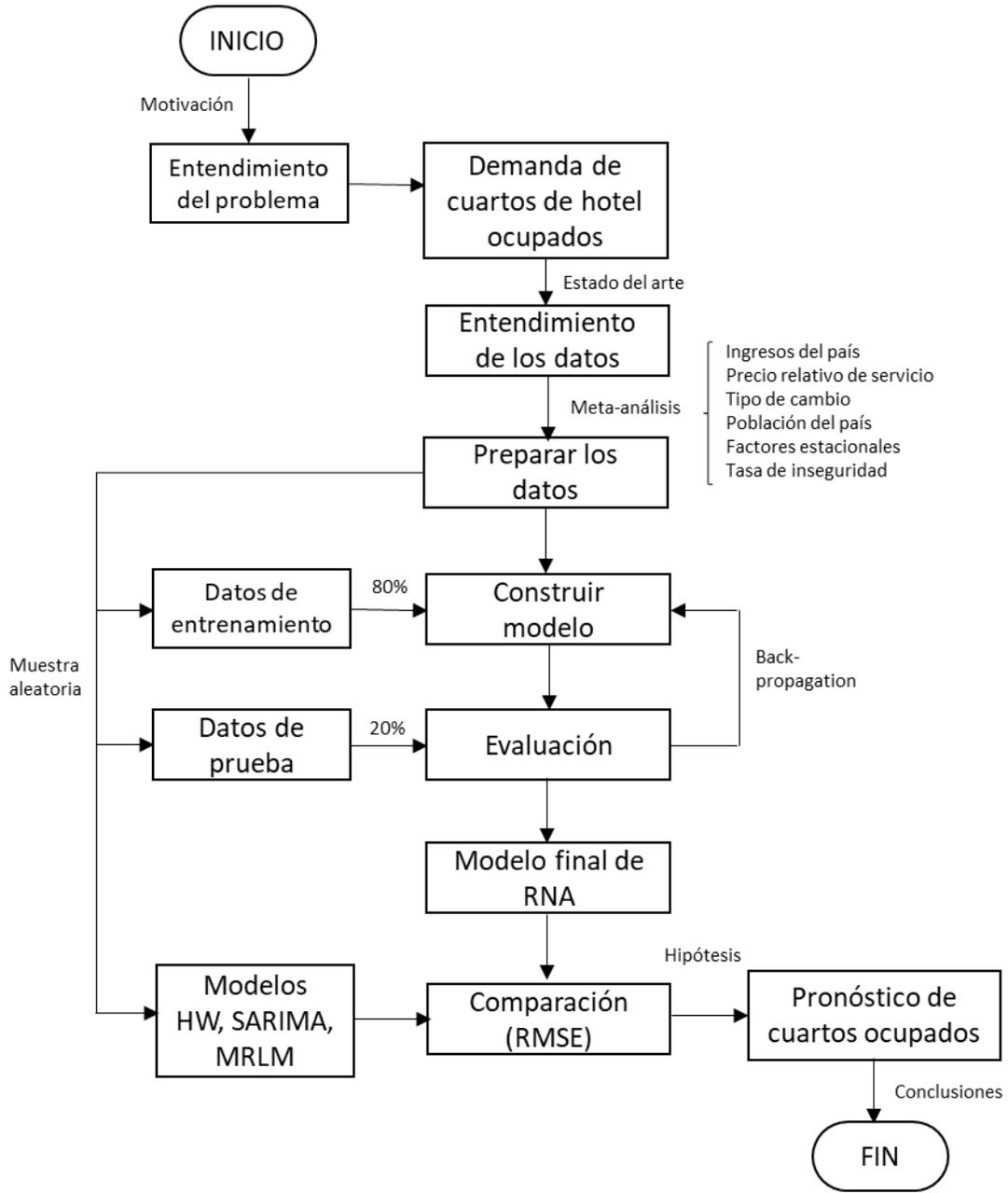
H1: El algoritmo de redes neuronales artificiales tiene un mejor desempeño que las técnicas convencionales de predicción.

H2: La infraestructura hotelera y turística de la República Mexicana es suficiente para abastecer la demanda creciente de turistas de los principales países que mantienen relaciones económicas con México.

A continuación se presenta la metodología que se empleó para el desarrollo de este trabajo de investigación. Se puede observar que todo comienza con la definición y entendimiento del problema, el cual es una consecuencia de la motivación por comprender los métodos de pronóstico. A raíz de ello, se tomó como variable de interés al número de cuartos de hotel ocupados en la República Mexicana, con el objeto de predecir su demanda. Se eligió esta variable ya que con ella se puede tipificar gran parte del sector turístico del país. Enseguida, fue determinar las variables explicativas, con las que es posible modelar el fenómeno.

Una vez que se logró preparar la base de datos, se continuó con el desarrollo del modelo para el algoritmo de Redes Neuronales Artificiales, para lo cual es importante contar con un subconjunto de datos de la base, con el que se entrena al modelo, y el resto de los datos para la validación del mismo. Se hace lo propio con los modelos convencionales de pronóstico, para que al final el desempeño de las Redes Neuronales Artificiales sea comparado con el desempeño de las series de tiempo de los métodos de Holt-Winters, SARIMA y el Modelo de Regresión Lineal Múltiple, mediante el valor de la raíz cuadrada del error cuadrático medio, ya que representa una medida justa del error. Finalmente se obtienen las predicciones de la demanda de cuartos de hotel ocupadas y se concluye sobre los resultados.

Ilustración 1: Metodología



Resumen de los capítulos

En el *Capítulo 1: Turismo como factor importante para el desarrollo del país y factores que influyen en la demanda turística*, se da una breve justificación del porqué el sector turístico tiene un peso considerable en el PIB de la nación. Se mencionan algunos indicadores con los cuales se compara el desempeño de los países en materia turística, así como la participación de México en este sector. Además, se habla de los factores que influyen en la demanda turística y la manera en que las variables que se utilizarán en el modelo fueron definidas. Se retoman publicaciones anteriores y un meta-análisis para enlistar las diferentes variables que han sido utilizadas en diferentes problemas de regresión para predecir la demanda turística de una país. Por último, se menciona cuál es la infraestructura total nacional de hoteles y servicios con la que el país dispone actualmente.

En el *Capítulo 2: Descripción de modelos convencionales de pronóstico*, se habla de los modelos utilizados para el pronóstico de la demanda turística de manera tradicional, que serán comparados con el desempeño de las redes neuronales; y la efectividad que éstos tienen en sus predicciones. Se explican tanto modelos de series de tiempo como el Holt-Winters multiplicativo y el modelo auto-regresivo integrado de media móvil con estacionalidad (SARIMA), así como el modelo de regresión lineal múltiple.

En el *Capítulo 3: Antecedentes y aplicación de las redes neuronales artificiales*, se presentan los antecedentes de las redes neuronales artificiales, así como su descripción general y la metodología de su funcionamiento. Se comienza con la definición de redes neuronales, así como la semejanza que existe entre las redes neuronales artificiales y las biológicas, se determina el porqué de su diseño y algunas características muy particulares que hay que tomar en consideración en el momento de seleccionar alguna de ellas. Se concluye con un listado de ventajas y desventajas del algoritmo que permiten tener un criterio más objetivo de la aplicación de esta metodología.

En el *Capítulo 4: Construcción del modelo con redes neuronales artificiales y su comparación con las técnicas convencionales de pronóstico*, se describe la metodología que se siguió para diseñar y construir el modelo con redes neuronales artificiales. Además, se presenta la forma en que las predicciones se llevan a cabo, se validan y son comparados con las técnicas convencionales de pronóstico.

Por último, en la parte de las *Conclusiones* se comparan los resultados obtenidos con la validación de las hipótesis que se plantearon, así como el cumplimiento total o parcial de los objetivos. Se presentan las áreas de oportunidad del trabajo, así como futuras investigaciones que pueden generar un valor agregado al trabajo de investigación.

Capítulo 1: Turismo como factor importante para el desarrollo del país y factores que influyen en la demanda turística

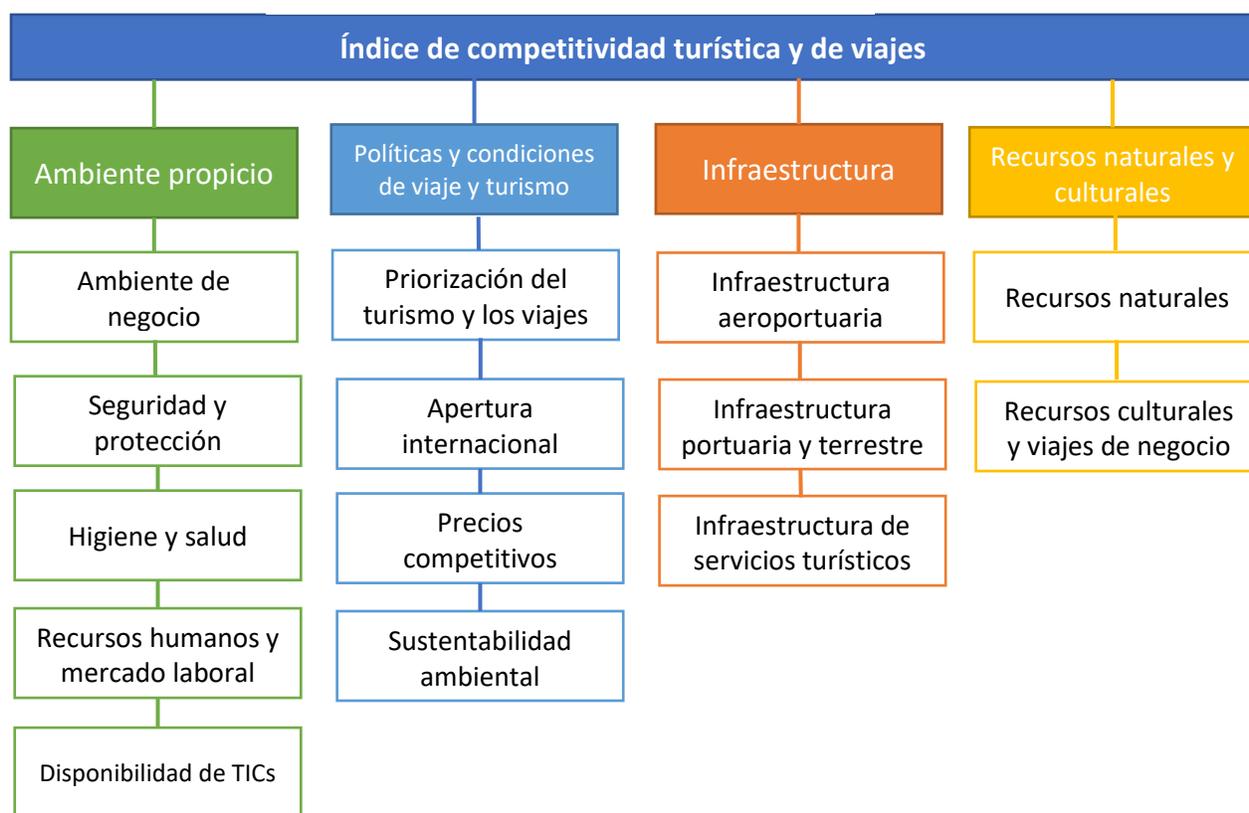
De acuerdo con la Secretaría de Turismo, organismo federal de la República Mexicana, el concepto “turismo” se entiende como el conjunto de actividades que realizan las personas durante sus viajes y estancias, en lugares distintos al de su entorno habitual. Se sabe que son variadas las actividades que una persona puede realizar en dichas estancias, ya que los motivos por los que una persona decide visitar otros lugares fluctúan desde el ocio, vacaciones, recreación, visita a amigos o familiares; hasta viajes profesionales, de negocios o religión (Secretaría de Turismo, 2018). Es dicha variación de actividades la que permite que el turismo sea una actividad versátil, con un mercado objetivo creciente.

Dado que el turismo ha adquirido una participación importante en los últimos años como una de las actividades terciarias o de servicios en cada uno de los países, ya que se ha posicionado como una guía para el crecimiento económico, es importante tener un indicador que muestre el desempeño de cada uno de los países en este sector. Es por ello que el índice de competitividad turística y de viajes, indicador desarrollado por la organización mundial del turismo (UNWTO, por sus siglas en inglés), es uno de los indicadores que mide el conjunto de factores y políticas que permiten el desarrollo sustentable del sector turístico, el cual contribuye al desarrollo y competitividad de un país. Es por ello que la organización mundial del turismo expone este tipo de análisis cada dos años, teniendo como primer ejemplar en el año 2007 (World Tourism Organization , 2018).

Es importante utilizar este índice en la medida que se deseen comparar países en el mismo sector, así como los resultados que se obtienen conjuntamente. Por ejemplo, para 2017 la industria turística mundial contribuyó con 10.4 % del PIB global y fomentó la creación de uno de cada 10 empleos del planeta (World Tourism Organization , 2018).

El índice de competitividad, tal como se puede observar en la siguiente imagen, se conforma de cuatro pilares principales: un ambiente propicio; políticas y condiciones propicias de viaje y turismo; infraestructura; y recursos naturales y culturales, los cuales a su vez se subdividen en 14 elementos que integran la totalidad de los rubros. La metodología de estimación del índice consiste en asignar una calificación del uno al siete a cada uno de los factores, en donde uno es una calificación pobre y siete es la mejor calificación que se puede obtener. De manera que para que el indicador sirva como comparador entre los diferentes países, se genera un valor ponderado promedio de los 14 elementos.

Ilustración 2: Índice de competitividad turística y de viajes



Traducción propia. Información recuperada de (Crotti & Misrahi, 2018)

Así, por ejemplo, de acuerdo con el reporte de competitividad de viajes y turismo de 2017, España es el país que se mantiene en el primer lugar del índice de competitividad, con un indicador promedio de 5.4, marcando la pauta para las demás economías. De igual manera, se sabe que el continente americano es la macro región con el segundo mejor incremento del índice de competitividad, justo después de la región de Asia y el pacífico, la cual presenta una tendencia positiva para los siguientes años. En la macro región del continente americano, se registró que el 80% de los visitantes fue en América del norte y Centroamérica; mientras que el 20% restante fue en América Latina. Esta diferencia abismal se debe, principalmente, al subdesarrollo de la infraestructura y la valuación de los recursos culturales que se cuentan en cada una de las regiones mencionadas (Crotti & Misrahi, 2018).

Bajo esta sintonía, es necesario precisar que México se encuentra posicionado en el lugar 22 del ranking mundial del índice de competitividad de turismo y viajes del 2017, en donde obtuvo un puntaje de 4.5, lo que se deriva de un incremento de 8 lugares desde el informe elaborado en 2015. Sin embargo, aun cuando se ha catalogado a México como uno de los mejores 30 países con un índice de competitividad turística superior, el país se tiene que concentrar en aspectos muy puntuales que garanticen que ese crecimiento sea constante y que los turistas opten por él como una de las mejores alternativas para vacacionar. Asimismo, de acuerdo con (Guzmán & otros, 2011) se determinó que la demanda turística en México responde inelásticamente a los cambios del costo de vida del país (-0.0277), lo que refleja la

competitividad del sector en comparación a los destinos turísticos de centro y sur del continente americano. Además, que mantiene una relación directa con el comportamiento de la actividad económica de Estados Unidos y Canadá (78%). En la sección de anexos, se pueden observar el desempeño de México en cada uno de los rubros que conforman el índice de competitividad turística y de viajes.

Por otro lado, es importante mencionar que a pesar de la existencia del indicador de competitividad turística, éste no garantiza, necesariamente, que el país sea dominante en ese sector. Ya que los parámetros como el número de llegadas turísticas, los ingresos reportados por dichas acciones, la inversión extranjera directa, entre otros; no se encuentra contenido en la evaluación del índice. Es por ello que se debe utilizar el indicador como una referencia de las actividades que los países deben desarrollar en la medida que deseen obtener mejores resultados.

De manera que, si se toma como referencia ‘los acontecimientos destacados del turismo’ proporcionados por la Organización Mundial del Turismo (World Tourism Organization , 2018), México queda posicionado en el sexto lugar de los países que recibieron un mayor número de turistas en el año 2017. Sin embargo, queda al descubierto que México no figura entre los primeros diez países que tienen el mayor ingreso por cuestiones de turismo. Ya que para el año 2017, se registró un ingreso de 21,333 millones de dólares, cerca de una décima parte del ingreso total que tuvo el mejor país evaluado.

Ilustración 4: Posición de México a nivel mundial por número de llegadas de turistas



Imagen recuperada de (World Tourism Organization , 2018)

Ilustración 3: Posición internacional de los países por ingresos turísticos



Imagen recuperada de (World Tourism Organization , 2018)

Asimismo, el consejo mundial de turismo y viajes (WTTC por sus siglas en inglés) presenta el impacto económico que este sector ha dejado en México el pasado año 2017. Se registró que el 7.1% del total del PIB fue contribución directa del turismo y se estima que para 2028 esa participación crezca hasta un 7.9% del total del PIB. De igual manera, se tiene el registro que el sector turístico aportó un 7.5% del total de empleos generados en 2017, el cual se ha pronosticado que crezca hasta un 8.5% de la generación directa total de empleos del país para 2028 (Guevara M., 2018). Esto deja al descubierto la importancia que tiene el turismo en México y el mundo. Las estadísticas y los pronósticos reflejan el crecimiento de un sector, que si recibe la atención debida, puede ser sustento principal de un país.

Es por ello que se puede asegurar que el turismo, además de ser medido como un índice entre las economías mundiales, representa un sector creciente y ventajoso. De acuerdo con (World Tourism Organization , 2018) el turismo es más de lo que nos podemos imaginar, ya que involucra aspectos tanto económicos, culturales, y sociales. El turismo tiene una importancia tal en el desarrollo de un país ya que promueve la preservación cultural, la protección del ambiente, paz y seguridad. Así como incentivar el crecimiento de la economía mediante la creación de empleos. Este sector constituye, a nivel global, con 1.6 billones de dólares en exportaciones, un poco más del 10% del PIB global y un aporte al 30% de exportación de servicios. La importancia del sector turismo dentro del aspecto económico radica en que se convierte en el eslabón final de varias cadenas productivas, por ejemplo: transporte, alimento y hospedaje. Es dicho crecimiento directo que induce el crecimiento en los aspectos indirectos. Este fenómeno es conocido como el efecto multiplicador del turismo (Octavio, 2008), un diagrama que lo representa se encuentra en la sección de anexos.

Otro punto a favor del turismo es la comparación entre sectores económicos elaborado por 'Oxford Economics' (World Travel & Tourism Council, 2012), en donde se ha determinado que el sector turismo tiene beneficios directos, indirectos e inducidos mayores que la mayoría de los demás sectores evaluados: sector minero, educativo, industria química, industria automotriz, sector de comunicaciones y servicios financieros. Se demostró que el turismo, en ocasiones, provee de ventajas dobles, ya que al mismo tiempo en que genera empleos e ingresos al país, promueve el patrimonio cultural y las tradiciones del mismo. Al día de hoy, el turismo es visto como un incentivo para fomentar la inversión extranjera directa, así como el detonante para incrementar la inversión en infraestructura pública y privada, que se refleja en el incremento sostenido del desarrollo económico local.

Sin embargo, el crecimiento de un sector, independientemente de cual se trate, está relacionado fuertemente con el contexto económico, social y cultural en el que se encuentre inmerso. El sector turístico no es la excepción. Existen ciertos fenómenos que obstaculizan el crecimiento sustentable de este sector. Se habla de cuestiones ambientales, como el cambio climático, la sobreexplotación de los recursos naturales, fauna y flora en peligro de extinción; cuestiones sociales como la inseguridad de los países, guerras, problemas étnicos, escasez de recursos, e incluso abuso y explotación de personal; cuestiones económicas como la competitividad de los precios que se ofrecen, y el aumento de la globalización.

Dadas todas las áreas de oportunidad que se tienen para que el sector turístico, en donde se pretende que el sector, además de sostenible sea sustentable, se cita una frase presente en el artículo "Comprendiendo los problemas críticos para el futuro del turismo y los viajes" (World Travel & Tourism Council, 2017) <<Con crecimiento, vienen responsabilidades- una responsabilidad de construir un crecimiento resiliente>>. Con ella se pretende dar a conocer que el turismo sustentable es el futuro del sector.

A pesar de ello, hasta este momento se ha demostrado que el sector turístico ha alcanzado un crecimiento sustancial, no solo en los países industrializados, sino también en aquellos que tienen un alto patrimonio cultural. También es necesario mencionar que los costos de transporte, de unos años a la fecha han ido

disminuyendo, lo que abre la posibilidad que la interacción entre los países sea cada vez mayor (González Alatorre & Conde Pérez, 2011).

Tal es el impacto del turismo en la sociedad contemporánea, que resulta imperante determinar la demanda turística en cualquier destino. Es por ello que debe ser una variable conocida, medida y analizada. Su conocimiento y control ayudará a que los actores económicos, instituciones, y entes de influencia en los países, tengan las herramientas necesarias para tomar las decisiones adecuadas en aras de promover e incrementar el turismo nacional e internacional. Tal como mencionan (González Alatorre & Conde Pérez, 2011), el concepto de la demanda turística puede variar de acuerdo al enfoque que se le haya asignado. Una de ellas es el número de personas que viajan o desean viajar, y la segunda se entiende como el consumo y los ingresos que se generan en la zona turística. Para efectos de este trabajo de investigación, se considera la demanda turística como una combinación de ambos conceptos, esto es, el número de personas que viajan a un lugar diferente a su residencia habitual y que generan ingresos al lugar destino que hayan elegido. Es muy importante esclarecer esta definición, ya que es a partir de ella que las variables de interés toman lugar.

Asimismo, en la publicación de (González Alatorre & Conde Pérez, 2011) se menciona una serie de factores que condicionan la decisión de viajar de los turistas. En el artículo se pueden identificar cuatro grupos principales:

1. Factores económicos,
2. Factores relativos a las y los turistas,
3. Factores del sistema de comercialización, y
4. Satisfacción del cliente

A continuación se presenta una tabla con el concentrado de los factores, los subfactores que los conforman, así como una breve descripción de cada uno.

Tabla 1: Factores para determinar la demanda turística de un país

Factor	Subfactor	Características
Factores económicos	Nivel de renta disponible	Es la cantidad de dinero disponible por el consumidor para gastar en bienes de turismo, tras haber pagado sus necesidades básicas.
	Precios del destino	Precios de transporte, alojamiento, alimentación. Un aumento en los precios impacta negativa o positivamente en el destino.
	Precios de los destinos competitivos	El precio de los destinos sustitutos impacta en la decisión de destino original.
Factores relativos al turista	Gustos y preferencias	Usos y costumbres de los turistas: nivel de visa, creencias, nivel cultural, etc.
	Motivación	Se acude a un destino por el deseo de satisfacer una necesidad

	Condiciones socio-culturales	Atributos como la edad, el sexo, la profesión, el empleo, estado civil, nacionalidad, educación; son elementos que determinan un destino específico.
	Tiempo de ocio	Es el tiempo disponible para viajar: vacaciones, días feriados, puentes, etc.
	Factores geográficos	Componentes como la cercanía, el clima, etc.; que propician el turismo.
Factores de sistemas de comercialización	Gastos de publicidad	Es la inversión de los lugares destinos para promocionarse en las diferentes ciudades o países, con el fin de encontrar clientes potenciales.
Satisfacción del cliente	Calidad	Se pretende cumplir con la expectativa del turista, para repetir la experiencia.

Elaboración propia con información de (González Alatorre & Conde Pérez, 2011)

Por otro lado, el meta-análisis desarrollado por Christine Lim de la universidad del Oeste de Australia (Lim, 1997) en donde se analizaron 100 publicaciones citadas en la literatura de la demanda turística internacional de los años 1970-1997, presenta una comparación de las variables empleadas en los modelos de regresión, modelos econométricos y demás técnicas empleadas para pronosticar la demanda turística. Apoyándose del modelo general de demanda turística internacional, se determinaron las variables que tenían una mayor frecuencia en los artículos citados.

El modelo general de demanda turística internacional, en modelos econométricos de una sola ecuación, es el siguiente:

$$DT_{ij} = f(Y_j, TC_{ij}, RP_{ij}, ER_{ij}, QF_i) \text{ donde}$$

DT_{ij} es la demanda de viajes internacionales del origen *j* al destino *i*

Y_j son los ingresos en el origen *j*

TC_{ij} son los costos de transporte entre el destino *i* y el origen *j*

RP_{ij} son los precios relativos (la proporción de los precios en el destino *i* en relación con los precios del origen *j* y destinos alternativos)

ER_{ij} es el tipo de cambio, medido como unidad de tipo de cambio del destino *i* por unidad de tipo de cambio del origen *j*

QF_i son los factores de calidad en el destino *i*

Ilustración 5: Principales variables turísticas dependientes

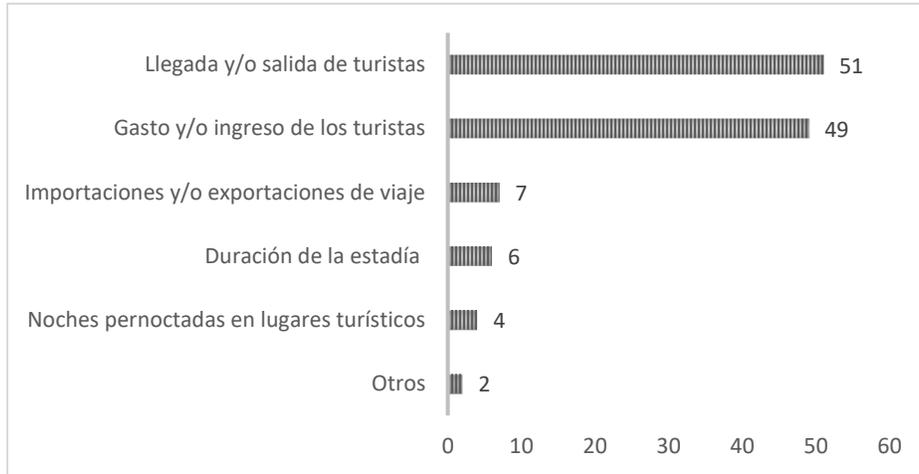


Ilustración tomada de (Lim, 1997). Traducción propia

De manera que, comparando las estadísticas del meta-análisis con el modelo general de la demanda turística internacional, se encuentra que la mayoría de los autores utilizaron como variable dependiente, o en el caso de redes neuronales (la salida), el número de llegadas y/o salidas de turistas, y el gasto e ingresos de los turistas. En la *Ilustración 5: Principales variables turísticas dependientes* se puede apreciar que la llegada y/o salida de los turistas fue empleada 51 veces, mientras que los gastos e ingresos de los turistas 49 veces.

Ilustración 6: Principales variables turísticas independientes

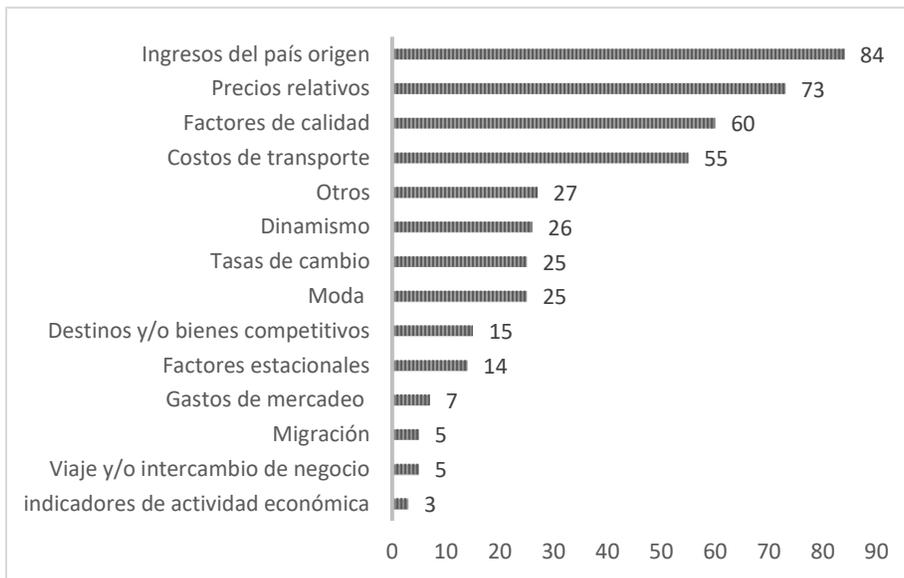


Ilustración tomada de (Lim, 1997). Traducción propia.

Por otra parte, tal como se puede observar en la *Ilustración 6: Principales variables turísticas independientes*, las variables explicativas, independientes, o en nomenclatura de redes neuronales (las

entradas), que tienen una mayor participación en el pronóstico de la demanda turística internacional de los 100 ejemplares analizados son: el ingreso del país origen, los precios relativos, los factores de calidad y los costos de transporte.

Sin embargo, es una realidad que se debe considerar, que a pesar de tener una metodología o un meta-análisis de las variables determinadas, las entradas que se considerarán en el modelo dependerán de la disponibilidad de los datos y la confiabilidad de las fuentes.

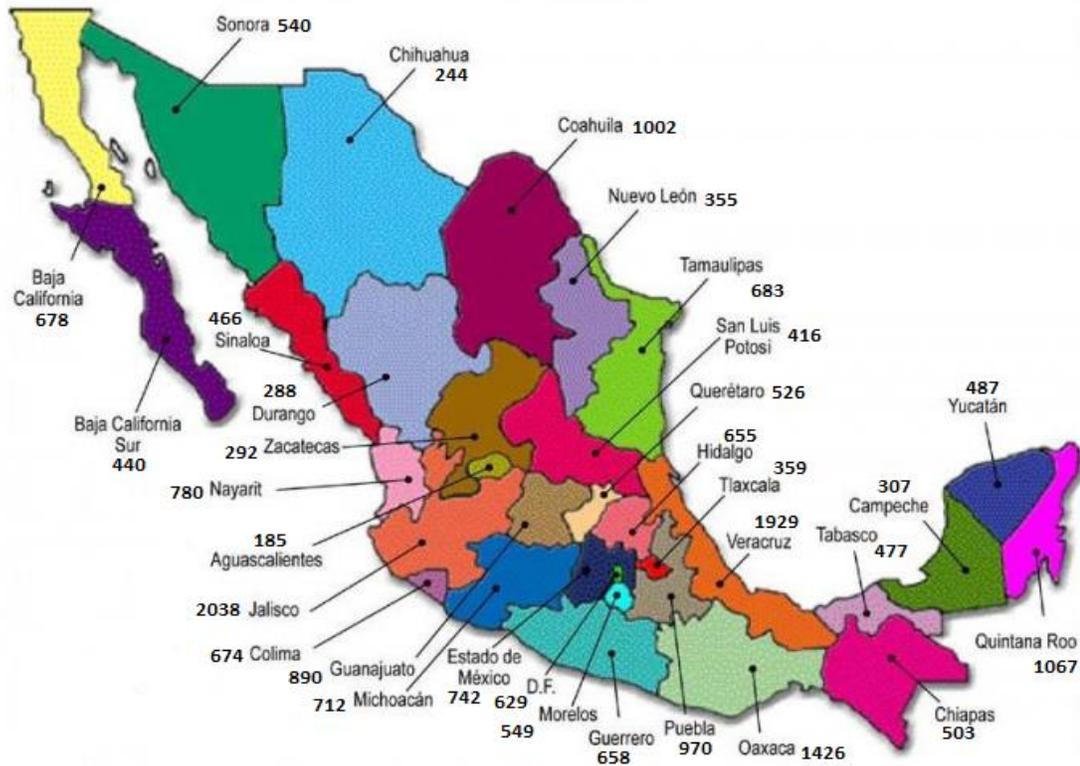
Por último, es importante reconocer que el artículo de (Law & Au, 1999) revela que además de las variables consideradas en el modelo general de demanda turística internacional y las propuestas por el meta-análisis, se deben considerar variables como la población del lugar origen y el costo de vida del lugar destino, la seguridad de los países destinos, así como los factores estacionales.

De manera que, después de realizar una búsqueda exhaustiva de información, se llegó a la conclusión de que las variables explicativas que se emplean para este trabajo de investigación son: el ingreso por país de origen, el precio relativo de servicio, la relación del tipo de cambio, la población del país de origen, la tasa de inseguridad y los factores estacionales. Una descripción más detallada de las variables se encuentra en capítulos posteriores.

Actualmente, México cuenta con una infraestructura hotelera total nacional de 21, 967 hoteles de una a cinco estrellas y algunos sin categoría, lo que resulta en un total de 18, 996,447 cuartos disponibles (Secretaría de Turismo, 2018). En la siguiente imagen se puede observar la distribución del número de hoteles por entidad federativa, se aprecia que los estados con el mayor número de hoteles hasta este momento son: Jalisco, Veracruz, Oaxaca, Quintana Roo y Coahuila; lo cual es consistente, debido a que son los principales estados que son visitados por los turistas. En cambio, los estados que tienen el menor número de infraestructura hotelera son: Aguascalientes, Chihuahua, Durango y Zacatecas. De manera general, se deberá tomar en consideración los pronósticos de la demanda turística de los cuartos de hotel ocupados en cada uno de los estados, con el objeto de determinar si la oferta actual es suficiente.

Asimismo, de acuerdo con los datos del compendio de estadísticas de turismo, recuperado de la Organización Mundial del Turismo (World Tourism Organization , 2018), se sabe que para el año 2017 se registraron 39, 184, 191 turistas en México, de los cuales 30, 826, 053 son de Estados Unidos de América y 1, 985, 084 tienen nacionalidad canadiense. Estos datos son importantes, en la medida de la delimitación del alcance del modelo que se diseñará para pronosticar la demanda turística en México.

Ilustración 7: Número de hoteles por entidad federativa



Elaboración propia con información de (Secretaría de Turismo, 2018)

RESUMEN DEL CAPÍTULO

El turismo es un sector importante para el crecimiento de cualquier país. En México, este sector ha cobrado tal relevancia que ha llegado al grado de aportar el 7.1% del PIB nacional directo. Es una fuente valiosa de empleos y un incentivo para la inversión extranjera directa. A pesar que existen algunos factores que detienen su crecimiento, México cuenta con el potencial necesario para que el sector turismo sea uno de los cimientos más importantes del país. Con la diversidad cultural, de fauna y flora, y las medidas pertinentes para mitigar los problemas de seguridad, se espera que el crecimiento del turismo al 7.9% del PIB total directo para 2028 sea una garantía. Por otro lado, este capítulo presenta también los estudios previos realizados para la determinación de las variables relevantes en modelos de regresión, modelos econométricos y otras técnicas de pronóstico, para predecir la demanda turística de un país. Se encontró que hay un número vasto de variables numéricas y alfanuméricas que se deberán incorporar al modelo en función de la disponibilidad de los datos y la confiabilidad de las fuentes. En este trabajo de investigación se ha tomado en consideración la propuesta de las variables descritas en el capítulo para conformar un modelo que se ajuste mejor a la realidad. No obstante, la disponibilidad de la información en páginas oficiales de secretarías encargadas a estos fines no es la adecuada. Por lo tanto, se han seleccionado las variables disponibles más significativas.

Capítulo 2: Descripción de modelos convencionales de pronóstico

En este capítulo se hablará de los diferentes tipos de modelos cuantitativos que se han utilizado, convencionalmente, para pronosticar o predecir el comportamiento de las variables en el futuro. Se explicarán los modelos en series de tiempo de Holt-Winters multiplicativo, los modelos auto-regresivos integrados de media móvil con estacionalidad (SARIMA por sus siglas en inglés) y, por último, los modelos econométricos o de regresión en especial los que involucran más de una variable independiente.

En primer lugar se debe entender que un pronóstico es la predicción de los hechos y condiciones futuras, es por ello que cualquier organización, institución o sistema deben ser capaces de realizar predicciones para poder tomar decisiones inteligentes, ya que se pretende adelantar a los sucesos futuros y planificar una solución. Las aplicaciones de los pronósticos son tan vastas que permiten la integración de muchos campos de conocimiento, desde la administración de negocios hasta una planificación industrial.

La clasificación de los métodos de pronósticos se realiza de acuerdo con las características de los datos, esto es: cualitativos y cuantitativos. Una de las propiedades de los métodos cualitativos es que se utilizan cuando hay pocos datos o los datos son nulos para el pronóstico. Por lo que se debe recurrir a un experto en el tema o grupo de expertos que guíen el proceso de predicción. En cambio, los métodos cuantitativos son aquellos que utilizan datos numéricos para realizar las predicciones y se dividen a su vez en: univariados y causales. De donde se sabe que los modelos univariados predicen valores futuros de la variable de interés con base en el patrón de esa misma variable, es decir, se comportan como una serie de tiempo. Mientras que los causales predicen valores futuros de la variable de interés tomando en consideración la relación estadística de la variable dependiente con las variables independientes, es decir, un modelo de regresión o también conocidos como econométricos (Bowerman, O'Connell, & Koehler, 2007).

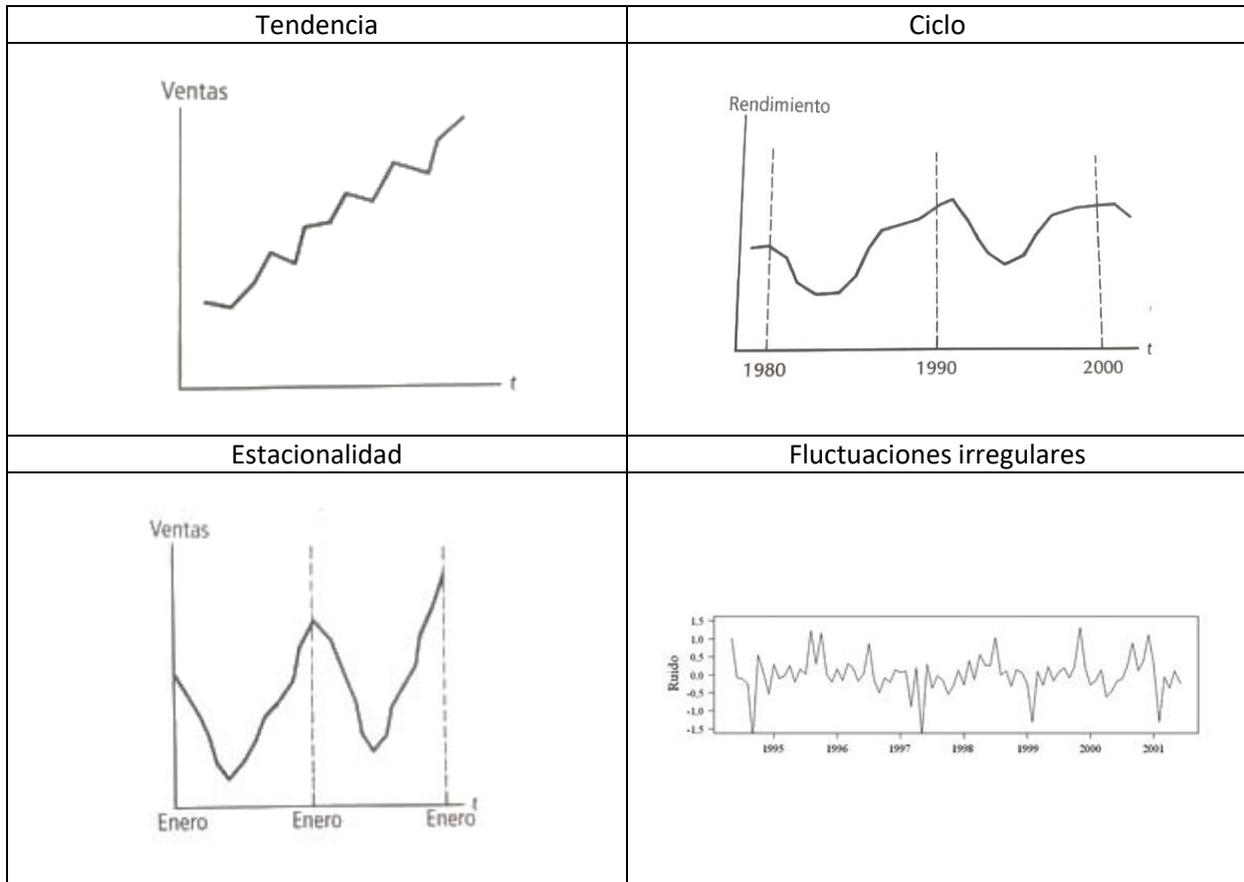
Por un lado, las series de tiempo son una sucesión cronológica de observaciones de una variable particular; es decir, muestran el comportamiento de una variable de interés a lo largo de un periodo específico. De acuerdo con los autores del libro "Pronósticos, Series de tiempo y regresión, un enfoque aplicado" (Bowerman, O'Connell, & Koehler, 2007) y (Hanke, 2010), una serie de tiempo consta de cuatro componentes principales:

1. Tendencia: es el movimiento hacia arriba o hacia abajo los datos con respecto a un periodo. Refleja el crecimiento o declinación de un fenómeno, es decir, la variable de interés.
2. Ciclo: son movimientos hacia arriba y hacia abajo alrededor de los niveles de tendencia. Un ciclo es medido de máximo a máximo o de mínimo a mínimo, de dichos movimientos alrededor de la tendencia.
3. Variaciones estacionales: son patrones periódicos que se completan dentro de un año civil y se repiten cada año. Los principales están marcados por las estaciones del año, los meses o trimestres.

4. Fluctuaciones irregulares: son movimientos erráticos (aleatorios) que siguen un comportamiento irregular. Se debe a errores del analista o incidentes fortuitos en el periodo en el cual se registraron los datos.

Las siguientes imágenes muestran, gráficamente, los componentes de una serie de tiempo. Para ello se tomó como base las gráficas generadas por (Bowerman, O'Connell, & Koehler, 2007) en su libro "pronósticos, series de tiempo y regresión: un enfoque aplicado".

Ilustración 8: Componentes gráficos de una serie de tiempo



Elaboración propia con ilustraciones de (Párraga, Cubillos, & Correa-Ramirez, 2010) y (Bowerman, O'Connell, & Koehler, 2007).

Por otro lado, un modelo econométrico o de regresión identifica las variables independientes que podrían afectar significativamente el valor de la variable objetivo o dependiente, así como la relación que existe entre dichas variables (Bowerman, O'Connell, & Koehler, 2007). Se pueden identificar los modelos de regresión lineal simple, regresión lineal múltiple y regresiones no lineales, por ejemplo la regresión polinomial, logarítmica, entre otras.

A continuación se presentan y explican los métodos seleccionados de regresión y series de tiempo que serán comparados con el modelo de redes neuronales artificiales. Dentro de los métodos de regresión, se considerará únicamente la regresión lineal múltiple. Para los modelos de series de tiempo se expone el suavizado exponencial de Holt-Winters y los modelos SARIMA. Se han seleccionado estas técnicas por la

similitud que tienen con el algoritmo de redes neuronales, ya que se puede considerar a una red neuronal artificial como un modelo de regresión no paramétrica no lineal, donde bastará comparar su desempeño para demostrar su precisión. Además, debido al comportamiento de la variable de interés, el número de cuartos de hotel ocupados, se pueden identificar los componentes de tendencia y estacionalidad, por lo que estos métodos seleccionados son los más adecuados para dicho pronóstico.

Comenzando con los métodos de series de tiempo, el análisis iniciará con el *suavizado exponencial multiplicativo de Holt-Winters*, el cual es un método que se emplea cuando la tendencia y la estacionalidad de los datos de una serie cambian en el tiempo. Las ecuaciones de suavización que definen el comportamiento de una serie de tiempo modelada con el método de Holt-Winters son las siguientes:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)[L_{t-1} + T_{t-1}]$$

$$T_t = \beta[L_t - L_{t-1}] + (1 - \beta)T_{t-1}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}$$

En donde se tiene que L_t es la estimación del nivel actual de la serie suavizada, T_t es la estimación para la tasa de crecimiento o tendencia, S_t es la estimación para el factor estacional de la serie temporal en el periodo t . Y_t es el dato observado de la serie en el tiempo t , y las letras griegas α , β y γ representan las constantes de suavización para el nivel, la tendencia y la estacionalidad, respectivamente. Estas constantes pueden tomar cualquier valor dentro del rango $[0,1]$.

Observando las constantes de suavizado en las ecuaciones de suavización, se puede concluir que mientras más pequeñas sean las constantes, cercanas a cero, la ponderación de las estimaciones estará conformada de los datos históricos, mientras que si las constantes tienen un valor cercano a 1, solo se considerará la participación de los datos más actuales. Para la constante de estacionalidad, se pueden comparar los coeficientes entre sí para determinar cuál estación, mes o trimestre es el que tiene una participación mayor en el diseño del modelo. Esta aseveración se retomará más adelante para analizar los meses en los que se tiene una mayor incidencia de cuartos de hotel ocupados (Hanke, 2010).

Una vez que se tienen las ecuaciones de suavización, es posible obtener el modelo de pronóstico puntual. La expresión matemática que representa el fenómeno se escribe a continuación:

$$\hat{y}_{t+p}(t) = (L_t + pT_t)S_{t+p-s} \text{ con } p = 1, 2, \dots$$

Donde $\hat{y}_{t+p}(t)$ es el pronóstico para el periodo p en el futuro, p es el número de periodos futuros que se desean estimar y s es la longitud de la estacionalidad, ya sea en cantidad de trimestres, meses o por las estaciones. Por lo que, el valor de S_{t+p-s} es el factor estacional más reciente que corresponde con el periodo que se desea pronosticar.

Una manera de iniciar el algoritmo de Holt-Winters es asignar los valores iniciales de L_t , T_t y S_t con diferentes técnicas, ya sea por el método del último valor para encontrar el nivel inicial, igualando la

tendencia con cero y asignar un uno a los coeficientes estacionales. Esto dependerá del paquete de cómputo que se utilice, así como el criterio del analista (Hanke, 2010). Para este caso, se empleará el paquete de cómputo “The R Project for statistical computing”, el cual tiene su método de iniciación de las variables, así como del cálculo del algoritmo. Sin embargo, es necesario precisar que sea cual sea el criterio de iniciación de las constantes de suavización y de los estimadores, el objetivo principal es minimizar la suma de los errores cuadrados (SSE, por sus siglas en inglés) y con ello determinar el valor óptimo de las constantes.

Como se sabe, todos los modelos de pronósticos, por más acertados que se puedan modelar, siempre existirá un margen de error. Es por ello que los intervalos de predicción ofrecen, como su nombre indica, un intervalo en el cual es probable que el pronóstico puntual se encuentre contenido. Para el modelo de Holt-Winters, los intervalos de predicción con un 95% de confianza calculado en el periodo t se calculan de la siguiente manera:

$$[\hat{y}_{t+p}(t) \pm z_{[0.025]}s_r(\sqrt{c_p})(S_{t+p-s})]$$

Donde $\hat{y}_{t+p}(t)$ es el pronóstico para el periodo p en el futuro, $z_{[0.025]}$ es el número de desviaciones estándar que los datos se alejan de la media con una confianza del 95%, s_r es el error estándar relativo; ya que como se trata del modelo Holt-Winters multiplicativo, es necesario calcular la suma de los errores cuadrados relativos. Por otro lado, c_p corresponde a la interacción de la tendencia con el nivel, en el periodo dado. Por último, S_{t+p-s} es el factor estacional más reciente que corresponde al periodo que se quiere estimar. A continuación se presentan las expresiones matemáticas que permiten calcular el error estándar relativo, así como el valor de c_p .

$$s_r = \sqrt{\frac{\sum_{j=1}^t [y_j - \hat{y}_j(j-1)]^2}{\hat{y}_j(j-1)^2}}{t-3}$$

$$c_p = \sum_{j=1}^{p-1} \alpha^2 (1 + [p-j]\beta)^2 (L_t + jT_t)^2 + (L_t + pT_t)^2$$

El segundo modelo de series de tiempo que se analizará, y que será sujeto de comparación con el método de redes neuronales artificiales es el *modelo auto-regresivo integrado con meda móvil y estacionalidad (SARIMA)*. Este modelo, como su nombre lo indica, se compone de dos teorías principales, por un lado las medias móviles y por otro, los modelos auto-regresivos. Cuando se dice que el modelo es auto-regresivo, se debe entender como la relación que existe entre los datos pasados de la variable independiente. Es decir, la autocorrelación que hay entre las observaciones pasadas con las observaciones más recientes. Tiene la apariencia de un modelo de regresión con valores retrasados de la variable dependiente en la posición de las variables independientes (Hanke, 2010).

Por otro lado, cuando se habla del método de medias móviles, se debe entender como el desplazamiento de los errores hacia adelante en función del tiempo. O bien, como la combinación lineal de los errores actuales y pasados de la desviación de la variable de respuesta en contraste con su media. Se puede hacer

el símil con un modelo de regresión, con el cual se intenta adecuar el comportamiento de la variable dependiente a partir de sus errores. (Hanke, 2010) (Pindyck & Rubinfeld, 1998)

Este tipo de modelo “mixto” funciona a la perfección en series de tiempo estacionarias, es decir, cuando los datos no varían alrededor de un nivel fijo, cuando la media es un valor constante a lo largo del tiempo. Sin embargo, la mayoría de las series de tiempo que se presentan en la vida cotidiana, o los problemas que se deben resolver, no son series estacionarias, su media varía a lo largo del tiempo. Cuando esto sucede, es posible aplicar la metodología ARIMA o SARIMA, dependiendo si los datos presentan estacionalidad o no, si y solo si se transforma la serie en una serie estacionaria. Para transformar una serie no estacionaria en una serie estacionaria, se debe diferenciar la serie tantas veces como sea necesario para fijar el nivel. La serie se convierte en una serie estacionaria de orden d . Asimismo, se logra transformar una serie no estacionaria en una serie estacionaria aplicando una función de transformación, tal como el logaritmo natural.

El modelo matemático general ARIMA es el siguiente:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

O bien, después de unos arreglos matemáticos, haciendo que

$$1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p = \phi(B)$$

$$1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_p B^p = \theta(B)$$

El modelo se expresa como:

$$\phi(B)y_t = \delta + \theta(B)\varepsilon_t$$

En donde la primera parte de los términos se refiere al modelo auto-regresivo y la segunda a los promedios móviles. Dado que se asume que el proceso de la serie de tiempo es estacionario, se tiene una sola media

$$\mu = \phi_1 \mu + \dots + \phi_p \mu + \delta \rightarrow \mu = \frac{\delta}{1 - \phi_1 - \dots - \phi_p}$$

De la expresión anterior, es de suponer que la condición necesaria de estacionariedad de la serie ocurre cuando

$$\phi_1 + \phi_2 + \dots + \phi_p < 1$$

Asimismo, para que la serie sea considerada estacionaria, $\phi^{-1}(B)$ debe converger, lo que implica que las raíces de la ecuación característica $\phi(B) = 0$ estén afuera del círculo unitario.

De manera simbólica, el modelo ARIMA se representa como $ARIMA(p, d, q)$ en donde p es el número de términos auto-regresivos regulares, d son el número de diferencias regulares (número de veces en que la serie se diferencia) y q es el número de términos regulares de media móvil. Cuando el término p vale cero,

se dice que el modelo es explicado exclusivamente por los errores de la serie, es decir por el modelo de medias móviles. Cuando el término q vale cero, entonces la serie es explicada solamente por el modelo auto-regresivo, es decir, por la historia de los datos de la variable dependiente. Por otro lado, cuando los términos p y q valen cero simultáneamente, el modelo se explica como una caminata aleatoria, es decir, por los valores del último valor que sucedió.

Para las series temporales que presentan un patrón estacional, se adapta el modelo ARIMA en el modelo SARIMA, en donde se incluye la correlación de los periodos estacionales con los términos de auto-regresión y de media móvil. Por lo que, la representación simbólica es $SARIMA(p, d, q)(P, D, Q)x$, con P como los términos auto-regresivos estacionales, D como la diferencia estacional en el retraso de la estación, ya sea mensual, trimestral, etc. y Q como los términos estacionales de promedio móvil. Para el caso de x , es el indicador que muestra la periodicidad de la estacionalidad. Para este trabajo de investigación la estacionalidad que se analiza es de periodicidad mensual, por lo que el valor de x es 12.

Es importante identificar los términos que conforman el modelo SARIMA para evaluar su desempeño y con ello lograr pronósticos más acertados.

Por último, el tercer modelo que se explicará como parte de los métodos convencionales de pronóstico que se han empleado para realizar predicciones está el *Modelo de Regresión Lineal Múltiple*. De acuerdo con (Bowerman, O'Connell, & Koehler, 2007) se tiene que un modelo de regresión lineal múltiple es aquel que utiliza más de una variable independiente para aproximar la relación con la variable dependiente. Es decir, el modelo que relaciona y con x_1, x_2, \dots, x_k . Matemáticamente, se representa como:

$$y = \mu_{y|x_1, x_2, \dots, x_k} + \varepsilon = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

En donde

1. $\mu_{y|x_1, x_2, \dots, x_k} + \varepsilon = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$ es el valor medio de la variable dependiente y cuando los valores de la variable independientes son x_1, x_2, \dots, x_k .
2. $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ son parámetros de regresión que relacionan el valor medio y con x_1, x_2, \dots, x_k .
3. ε es un término de error que describe los efectos sobre y de todos los otros factores que no son los valores de las variables independientes x_1, x_2, \dots, x_k .

Es necesario precisar que los modelos de regresión están sujetos a una serie de supuestos en los que se basa su teoría (Bowerman, O'Connell, & Koehler, 2007) (Hanke, 2010). A continuación se enlistan dichos supuestos:

- En cualquier combinación de valores de x_1, x_2, \dots, x_k , la media de la población de los valores potenciales del término de error es igual a cero.
- Suposición de la varianza constante, es decir, que la dispersión de los datos con respecto a la media no cambie. Los datos tienen un comportamiento homocedástico, lo cual se verifica graficando los errores en función de las variables independientes, la predicción y del tiempo.
- Suposición de normalidad. Establece que los errores están distribuidos de manera normal, lo cual se verifica, gráficamente, realizando un histograma de los residuos o aplicando una serie de pruebas estadísticas, tal como la prueba Jarque-Bera.

- Y por último, la suposición de independencia, la cual busca que las variables independientes no estén correlacionadas entre sí, es decir, cuando existe una multicolinealidad entre las variables. Se deben realizar pruebas estadísticas. Gráficamente, se verifica con una gráfica de los residuos contra el tiempo en donde se debe observar un patrón aleatorio, no cíclico ni alternante.

Cuando se haya determinado que los errores de los datos reales comparados con los datos pronosticados cumplan con los supuestos antes mencionados, se puede asegurar que el modelo de regresión lineal múltiple es una buena opción para pronosticar. Por lo que se puede aplicar el siguiente modelo matemático

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k$$

Si bien $\beta_0, \beta_1, \dots, \beta_k$ son parámetros de regresión, es necesario aclarar que β_0 es la ordenada al origen del modelo, es decir, el valor que el modelo adquiere cuando todos los demás parámetros valen cero. En la mayoría de los casos prácticos, la interpretación de la ordenada al origen no tiene mucho sentido, debido a que, prácticamente, el asignar un valor de cero a todas las variables explicativas al mismo tiempo es muy poco probable. Además, puede darse el caso en el cual alguna de las variables, en los datos históricos disponibles, no haya tomado o estado cerca del valor cero, por lo que no se tiene registro de su comportamiento en dicho escenario (Hanke, 2010).

Por otra parte, los parámetros de regresión representan el cambio que sufre la variable dependiente cuando se aumenta en una unidad la variable predictiva en cuestión y las demás permanecen constantes. Es decir, el valor de β_1 indica que por cada incremento unitario de la variable x_1 , cuando x_2, x_3, \dots, x_k permanecen constantes, el modelo o la variable de salida y aumentará o disminuirá en proporción y concordancia con el valor y signo de β_1 . Este razonamiento es válido para la interpretación de los parámetros de regresión restantes.

De manera similar que el modelo de regresión lineal simple, el modelo múltiple utiliza el principio de mínimos cuadrados para determinar el valor de los parámetros de regresión, con lo cual se garantiza que la distancia entre los valores reales y los ajustados, o bien el error ocurrido sea el mínimo. Es por ello que, con el principio de mínimos cuadrados, se busca minimizar la suma de los errores al cuadrado.

Para los modelos de regresión es común realizar una prueba de significancia de las variables independientes, para determinar que su aportación al modelo sea relevante. Ya que, de caso contrario, se puede omitir y hacer que el modelo sea más sencillo, tal como lo expresa el principio de parsimonia. El proceso de encontrar la significancia de las variables se conoce como regresión por pasos, ya que en cada uno de los pasos en los cuales se valida el modelo de regresión, se verifica la significancia de las variables explicativas. Dicha significancia se valida mediante una prueba de hipótesis, para determinar si existe suficiente evidencia para aceptar o no la participación de la variable independiente en el modelo mediante la comparación del estadístico de prueba con un nivel particular de confianza (Hanke, 2010).

Cuando se tiene el valor de las variables predictivas que se esperan pronosticar se evalúan en el modelo

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k ,$$

y se obtiene la estimación puntual del valor medio y . Sin embargo, tal como se mencionó en la parte de las series de tiempo, ningún método de pronóstico está exento de cometer errores, por lo que lo más conveniente es calcular un intervalo de confianza, con el cual se asegure que la estimación puntual esté contenida. La expresión matemática que permite calcular dicho intervalo de confianza es la siguiente:

$$\hat{y} \pm t_{\left[\frac{\alpha}{2}\right]}^{n-(k+1)} s_{\hat{y}}$$

En donde \hat{y} es el valor pronosticado, $s_{\hat{y}}$ es el error estándar de estimación, la cual se puede calcular como la desviación estándar por la raíz del valor de la distancia ($s\sqrt{\text{Valor de la distancia}}$) y $t_{\left[\frac{\alpha}{2}\right]}^{n-(k+1)}$ es el estadístico t de student con $n - (k + 1)$ grados de libertad y una significancia el $\frac{\alpha}{2}$ en cada cola.

Por último, es de vital importancia hacer notar un par de consideraciones en los modelos de regresión. En primer lugar, la función que se encuentra para describir el comportamiento de los datos solo aplica para predecir valores que estén contenidos en la región experimental, también conocido como el plano de mínimos cuadrados, el cual se compone de las combinaciones de los valores de las variables x_1, x_2, \dots, x_k . Esto se debe a que, como no se tienen datos del comportamiento futuro de la función no se sabe con certeza si la función se seguirá comportando como hasta ese momento lo había hecho. Se pueden predecir datos fuera de la región experimental, si y solo si se asume que el comportamiento de la función se mantiene (Bowerman, O'Connell, & Koehler, 2007).

En segundo lugar, se debe tener cuidado con el sobreajuste u overfitting de la prueba, la cual se refiere al hecho de agregar variables independientes a la función de regresión que dan cuenta de todas las excentricidades de los datos muestrales, es decir, cuando se tiene un mayor número de variables explicativas que datos totales, lo cual provoca que el desempeño del pronóstico sea bajo. Los especialistas, menciona (Hanke, 2010), dicen que deben existir al menos 10 observaciones por cada variable predictiva que se tenga. Además, es buena práctica dividir el conjunto de datos en dos porciones; una para ajustar los valores y la otra para validar el pronóstico y calcular los errores. Es justamente este procedimiento el que se emplea para realizar el pronóstico con el modelo de redes neuronales artificiales, ya que de primera instancia se desea evitar un sobreajuste de los datos.

De manera general es importante mencionar que el método de pronóstico que se debe utilizar para la solución de cualquier problema es aquel que cumpla las necesidades de la situación al menor costo y con la mínima inconveniencia. Ya sean del tipo cuantitativo, cualitativo o una mezcla de ambos, ya que la experiencia de los analistas y la seguridad de los números están estrechamente relacionados. Con la comparación de las técnicas de pronóstico que se realizará en este trabajo de investigación, se pretende validar esta aseveración al observar que el desempeño de unas técnicas se sobrepone al de otras, aun cuando algunas tengan un mayor consumo de recursos.

Para poder observar el desempeño de las técnicas de pronóstico propuestas, existen las medidas de error o residuos, las cuales son las diferencias que existen entre los datos reales observados y su valor pronosticado para su ajuste. En la siguiente tabla se observan las principales medidas de error que se pueden calcular para medir el desempeño de un modelo de pronóstico.

Tabla 2: Principales medidas de error

Nombre	Fórmula
Error de pronóstico o residuo	$e_t = Y_t - \hat{Y}_t$
Desviación media absoluta	$MAD = \frac{1}{n} \sum_{t=1}^n Y_t - \hat{Y}_t $
Error cuadrático medio	$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$
Raíz cuadrada del error cuadrático medio	$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$
Error porcentual absoluto medio	$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{ Y_t - \hat{Y}_t }{ Y_t }$
Error porcentual medio	$MPE = \frac{1}{n} \sum_{t=1}^n \frac{(Y_t - \hat{Y}_t)}{Y_t}$

Elaboración propia con información de (Hanke, 2010)

Para este trabajo de investigación se empleará la medida de la raíz cuadrada del error cuadrático medio estandarizado con los grados de libertad de cada uno de los métodos, ya que con esa medida de error los métodos de pronóstico pueden ser comparados, es decir, es una medida de error más justa o fidedigna, puesto que los valores son ajustados de acuerdo con el número de grados de libertad que incorpore.

RESUMEN DEL CAPÍTULO

En este capítulo se abordaron algunas técnicas de pronóstico convencionales de series de tiempo y regresión que serán comparadas con el desempeño de las redes neuronales artificiales. Se mencionan las principales características del modelo de series de tiempo Holt-Winters multiplicativo, los modelos autorregresivos integrales de media móvil con estacionalidad (SARIMA) y los modelos de regresión lineal múltiple.

Capítulo 3: Antecedentes y aplicación de las redes neuronales artificiales

Definición

Es la necesidad de los seres humanos de comprender el funcionamiento de los sistemas que lo rodean, que se han dado a la tarea de explorar, o al menos intentar de manera significativa, cada parte singular del cuerpo humano. El cerebro, al ser un músculo conformado de células capaces de intercambiar, procesar, analizar y retener información basada en la experiencia, se ha posicionado como el centro de atención para todos los avances en tecnología artificial y de computación. Es a partir de los estudios de Alan Turing, enfocados en visualizar el cerebro como una forma de ver el mundo de la computación, que los avances en este campo crecieron de manera exponencial. Pero no fue sino hasta 1943, cuando Warren McCulloch y Walter Pitts, neurofisiólogo y matemático de profesión respectivamente, concibieron los fundamentos de la computación neuronal, dando pauta al inicio de una nueva era computacional (Matich, 2001).

De tal forma que las redes neuronales artificiales resultan del esfuerzo humano de emular las características del cerebro en actividades computacionales. Entre la literatura no existe una definición común para este tipo de algoritmos, ya que desde diferentes perspectivas se puede asemejar a diferentes métodos estadísticos. Sin embargo, a continuación se presentan una serie de definiciones del concepto de acuerdo con algunos autores.

1. Las redes neuronales artificiales son algoritmos computacionales o programas de cómputo que se derivan de los intentos por modelar el funcionamiento de las células cerebrales, los cuales han sido aplicados a campos como reconocimiento de patrones, clasificación y problemas de optimización en ciencias físicas y químicas (Stern, 1996).
2. Las redes neuronales son métodos auto adaptativos basados en datos, que se ajustan a los problemas cuya solución es difícil de especificar, pero donde hay suficientes datos u observaciones con las que se pueda encontrar la relación existente entre los datos. [...] En este sentido, las redes neuronales pueden ser tratadas como un método estadístico multivariado no paramétrico y no lineal. (Zhang, Patuwo, & Hu, 1998)
3. Las redes neuronales artificiales son redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso (Matich, 2001).
4. Una red neuronal es un procesador distribuido masivamente en paralelo hecho para procesar unidades que tienen la naturaleza de almacenar conocimiento empírico y ponerlo a disposición para su uso. Se asemeja al cerebro en dos aspectos: A) El conocimiento es adquirido por la red desde el entorno a través de un proceso de aprendizaje. B) Las fuerzas de conexión entre

neuronas, conocidas como pesos sinápticos, son usadas para almacenar el conocimiento adquirido (Haykin, 2009).

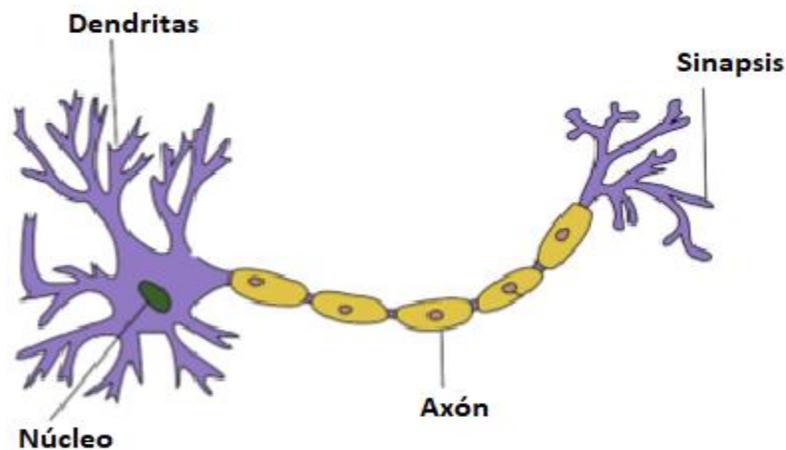
De acuerdo con las definiciones antes mencionadas, se puede llegar a la conclusión que las redes neuronales artificiales son algoritmos que emulan el comportamiento de las neuronas biológicas. Utiliza la información obtenida del entorno y, mediante procesos de aprendizaje, logra retener dicha información y utilizarla para adecuarse lo mejor posible a la estructura de los datos originales. Su aplicación es variada, ya que en los campos donde los métodos estadísticos convencionales son empleados, se ha demostrado que las redes neuronales tienen un mejor desempeño. En los siguientes capítulos se desarrolla de manera más fehaciente este argumento.

Neurona biológica vs Neurona artificial

Para comenzar a describir el funcionamiento de las redes neuronales artificiales, se hará la comparación obligada con las neuronas biológicas. De esta manera se entenderá mucho mejor el proceso y las similitudes que existen entre cada una de ellas. Es necesario mencionar que las arquitecturas de las redes neuronales tanto biológicas como artificiales son variadas, por cuestiones de simplicidad y ejemplificación, se ha optado por mostrar las arquitecturas más sencillas. Más adelante se hace un recuento de las arquitecturas o topologías más utilizadas para las aplicaciones con redes neuronales artificiales.

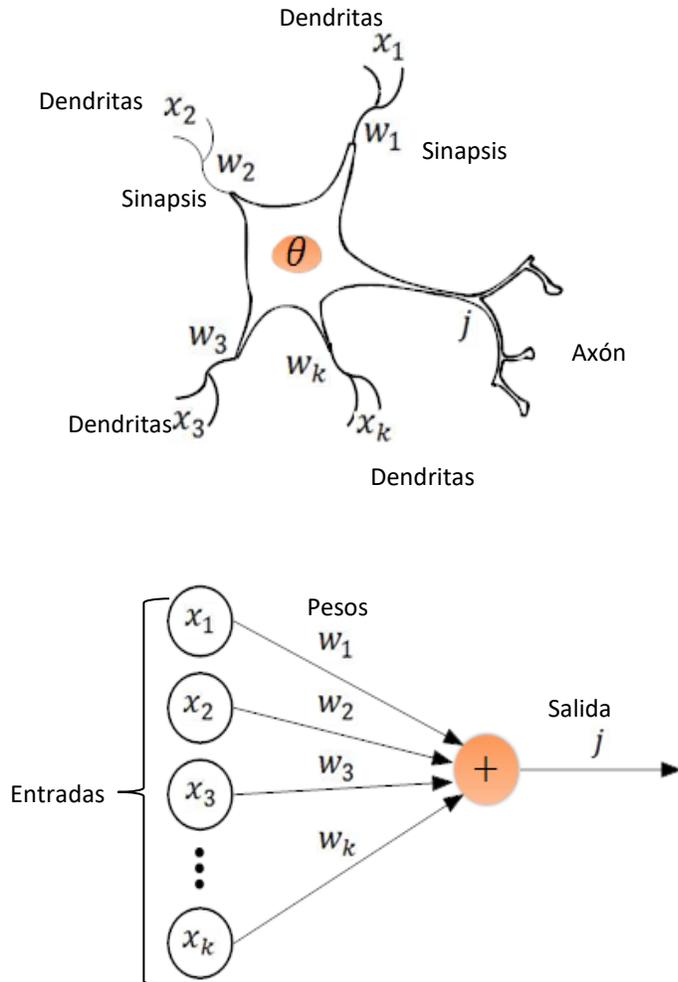
Por parte de la medicina se sabe que la unidad de células principales en el sistema nervioso es la neurona. Se estima que existen alrededor de 100 mil millones de neuronas en el cerebro humano y que están conectadas entre sí para formar este órgano tan importante. Ellas permiten la transmisión de información mediante señales eléctricas y químicas. Dicho intercambio se denomina sinapsis y se realiza cuando dos neuronas se conectan. Una neurona consiste, anatómicamente, en un cuerpo celular, con una estructura delgada que es conocida como dendritas y una extensión celular lineal llamada axón. El Axón de una neurona se conecta con las dendritas de otra neurona a través del proceso de sinapsis. Se estima que hay cerca de 100 billones de procesos sinápticos en el cerebro humano (Kotu & Deshpande, 2015). A continuación se muestra la anatomía de una neurona biológica.

Ilustración 9: Anatomía simple de la neurona biológica



Elaboración propia, modificando original de (Kotu & Deshpande, 2015)

Ilustración 10: Comparación de la neurona biológica con la neurona artificial



Elaboración propia con ilustraciones de recuperadas (Guerra Fernández, Lomaña Padilla, Guzmán Obregón, & Pérez Arbella, 2013)

Por otra parte, las redes neuronales artificiales se componen, principalmente, de los datos de entrada, los pesos y las salidas. Tal como se puede observar en la ilustración 9, las entradas de datos de las neuronas artificiales son el símil con las dendritas de una neurona biológica, que es precisamente el punto en dónde la célula recibe entradas de otras neuronas en forma de impulsos eléctricos. Los pesos o la ponderación que se le da a los datos en la neurona artificial es equivalente al proceso sináptico que ocurre en las neuronas biológicas, en donde se lleva a cabo el intercambio de información. Por último, la salida de la neurona artificial es totalmente comparable con el axón de una neurona biológica, ya que es la respuesta o estímulo final que se realiza en la célula y que es transportado a otro número finito de neuronas (Ciaburro & Venkateswaran, 2017).

Asimismo, tal como las neuronas biológicas lo hacen, las redes neuronales artificiales son capaces de conectarse con otras neuronas, es decir, que las salidas de una neurona sean las entradas de otras, lo que permite que la relación entre los datos y el proceso de aprendizaje sea mucho más acertado. También

genera que la topología o arquitectura de las redes sea cada vez más compleja, debido a que se contarán con un mayor número de neuronas interconectadas entre sí.

Historia

La historia de las redes neuronales se remonta, a los primeros estudios que realizó Alan Turing sobre el cerebro humano y la analogía que se podía hacer con las ciencias computacionales. Esto debido a que, como cualquier avance científico, se debe investigar sobre un concepto para después aterrizar la teoría en modelos matemáticos que permitan emular el funcionamiento de los procesos físicos, o en este caso biológicos, como lo es el cerebro.

Fue hasta 1943 cuando el neurofisiólogo Warren McCulloch y el matemático Walter Pitts lanzaron una teoría sobre la forma de trabajar de las neuronas, modelaron una red neuronal simple mediante circuitos eléctricos. Esta red mostró, en principio, que las redes neuronales pueden calcular cualquier función lógica o aritmética. Por lo que su trabajo es reconocido como el origen de las redes neuronales (Matich, 2001). La neurona recibe una suma ponderada de las entradas conectadas y la salida es un (1) si la suma es mayor que el umbral propuesto, de otra manera la salida es cero. Matemáticamente se representa como:

$$y_i = \theta\left(\sum_j w_{ij}x_j - \mu_i\right)$$

Donde y_i es la salida de la neurona i , w_{ij} es el peso de la neurona j a la neurona i , x_j es la salida de la neurona j , μ_i es el umbral de la neurona i y θ es la función de activación; que se define como (Warner & Misra, 1996):

$$\theta(\text{entrada}) = \begin{cases} 1 & \text{si entrada} \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

En seguida se puede observar esquemáticamente la arquitectura de la neurona propuesta por McCulloch y Pitts, en donde el círculo representa la neurona y la expresión que contiene es la función de activación (estos conceptos serán explicados más adelante).

Ilustración 11: Neurona de McCulloch-Pitts

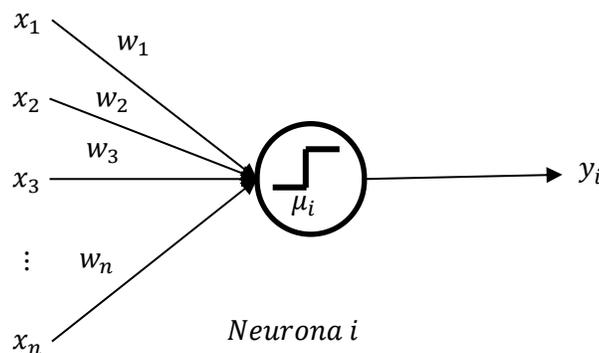


Ilustración recuperada de (Warner & Misra, 1996)

Posteriormente, los estudios de McCulloch y Pitts fueron seguidos por la propuesta de explicar los procesos de aprendizaje de las neuronas por Donald Hebb en 1949. Su idea fue que el aprendizaje ocurría cuando ciertos cambios en una neurona eran activados. Trató de asemejar el aprendizaje de las neuronas con la actividad nerviosa. Es por ello que su trabajo forma las bases de la Teoría de las Redes Neuronales, lo que después se convertiría en el mecanismo de aprendizaje conocido como aprendizaje Hebbiano (Matich, 2001), (Hagan, Demuth, Hudson, & De Jesús, 2018).

Sin embargo, la primera aplicación práctica de las redes neuronales artificiales se llevó a cabo alrededor de 1960 cuando Frank Rosenblatt desarrolló el “Perceptrón”. La forma más primitiva de red neuronal. La cual consistía en el modelo de McCulloch-Pitts con dos capas, entrada y salida (Warner & Misra, 1996). Con ella se demostraba la habilidad de reconocimiento de patrones y después generalizarlos. Es decir, reconocer patrones similares aun cuando no se le hayan presentado previamente en el entrenamiento. Además, el algoritmo era capaz de demostrar que para los problemas de clasificación en donde la solución existe, el modelo converge en un número finito de pasos. Este desarrollo generó mucho interés en el campo de las redes neuronales, por lo que su estudio creció exponencialmente en esa década. (Matich, 2001).

Contemporáneamente, Bernard Widrow y Ted Hoff desarrollaron un nuevo algoritmo utilizado para entrenar redes neuronales lineales adaptativas (ADALINE, por sus siglas en inglés), el cual era similar en estructura y capacidad al Perceptrón de Rosenblatt. Su primera aplicación fue en filtros adaptativos para eliminar ecos en las líneas telefónicas (Hagan, Demuth, Hudson, & De Jesús, 2018), así como para pronosticar el estado del tiempo (Zhang, Patuwo, & Hu, 1998). Sin embargo, tanto el Perceptrón de Rosenblatt como el algoritmo ADALINE sufrían de ciertas limitaciones, ya que no eran capaces de resolver problemas que no sean linealmente separables. Dichas limitaciones fueron publicadas en el libro de Minsky y Papert, con el cual se demostraba que los algoritmos hasta entonces desarrollados eran muy débiles y que suponían una limitante para los problemas reales de la sociedad. Esta publicación llevó a un decremento sustancial en la investigación de este campo, al grado de suponer el final de las redes neuronales artificiales (Matich, 2001).

Después de una década, aproximadamente, se publicaron trabajos importantes que dieron vida nuevamente a este campo de investigación. Kohonen y Anderson desarrollaron de manera separada e independiente una red neuronal capaz de funcionar como una memoria. Stephen Grossberg implementó la teoría de resonancia adaptada (TRA, por sus siglas en inglés), la cual es una propuesta de arquitectura de red que simula tanto la memoria de largo plazo como la de corto plazo del cerebro. A pesar de ello, con la publicación del libro “Computación neuronal de decisiones en problemas de optimización” de John Hopfield fue que se dio por renovada la teoría de las redes neuronales. Con ello se proponía que el acceso a la memoria de las neuronas se realizaba por medio de direcciones focalizadas, que si bien su uso no fue el de una memoria ya que se incurría en el problema de quedarse en un mínimo local, su aplicación fue para la solución de problemas de optimización, tal como el problema del agente viajero (Matich, 2001), (Warner & Misra, 1996).

Junto con el trabajo de la memoria asociativa usado para explicar la operación de ciertas redes recurrentes propuesta por Hopfield, el desarrollo del algoritmo de propagación hacia atrás o “backpropagation”

fueron los conceptos que hicieron posible el renacer de las redes neuronales. El algoritmo de propagación hacia atrás fue desarrollado primeramente por Paul Werbos pero no fue apreciado en su momento. Fue hasta 1986 que David Rumelhart y James McClelland retomaron el algoritmo y todas las críticas que se habían escrito en las publicaciones de Papert y Minsky fueron aclaradas. Esto es, el algoritmo de propagación hacia atrás es capaz de resolver problemas linealmente no separables (Warner & Misra, 1996).

La literatura coincide en que a partir de 1986, el panorama de la investigación de las redes neuronales fue alentador. Se desarrollaron diferentes arquitecturas de redes, además de la red multicapa hacia adelante, se implementó la máquina de Boltzmann basada en unidades estocásticas empleada para completar patrones faltantes. Asimismo se han atacado problemas de series de tiempo con redes neuronales recurrentes, tales como: la red Elman, el aprendizaje recurrente en tiempo real o la red Jordan. También existen arquitecturas que ejecutan análisis de componentes principales, creación de prototipos, codificación y arreglo de grupos o “clustering” (Warner & Misra, 1996). Existen, además, una serie de publicaciones en donde se compara la metodología de las redes neuronales artificiales con el desempeño que tienen los métodos convencionales estadísticos de pronóstico como Box-Jenkins y otros modelos no lineales, lo cual será, de manera indirecta, tema importante para abordar en este trabajo de investigación ya que se ratificará el desempeño de las redes neuronales artificiales sobre técnicas de pronóstico convencionales.

En la actualidad, son numerosos los trabajos que se realizan y publican cada año, las aplicaciones nuevas que surgen crecen a razón de la exploración misma del cerebro humano, así como la innovación del software y hardware de los equipos de cómputo. Sin embargo, son las necesidades de automatizar, eficientizar y agilizar los procesos y métodos, que han resultado en un crecimiento exponencial de las áreas de inteligencia artificial, acompañadas del análisis masivo de datos (Data Science), machine learning y Deep learning, campos en donde las redes neuronales artificiales tienen una participación predominante.

Aplicaciones

Las aplicaciones de las redes neuronales artificiales son variadas y se están expandiendo debido a que son buenas herramientas para resolver problemas, no solo de ingeniería, ciencia o matemáticas, sino también en medicina, negocios, finanzas e incluso literatura. Este amplio campo de aplicación es lo que las convierte en una herramienta atractiva para cualquier tipo de problema. Se han reportado casos en los cuales las redes neuronales artificiales han contribuido significativamente, por ejemplo: clasificar el RNA (ácido ribonucleico) para documentar la pureza del aceite de oliva, han traducido un discurso del inglés al chino, también se han utilizado para la identificación de imágenes con palabras clave, e incluso han mejorado la tasa de trasplantes de corazón al identificar los donadores óptimos (Hagan, Demuth, Hudson, & De Jesús, 2018).

A continuación se listan una serie de aplicaciones de las redes neuronales artificiales, de acuerdo con el área de estudio.

Tabla 3: Aplicación de las redes neuronales por área de conocimiento

Área de aplicación	Ejemplos
Aeroespacial	Simulación de rutas de vuelos, sistemas de control de aviones, mejora en piloto automático, detección de fallas.
Automotriz	Control de inyectores, sistemas de frenos, sensores virtuales de emisión
Bancos	Evaluación de aspirantes a crédito, pronósticos de efectivo, clasificación de empresas, pronósticos de tipo de cambio, riesgo de crédito
Defensa	Reconocimiento facial, procesamiento de imágenes y radar, identificación de señales, armas inteligentes, optimización de los recursos escasos.
Electrónica	Predicción de secuencia de código, control de proceso, análisis de falla de chips, modelado no lineal
Entretenimiento	Animación, efectos especiales, pronóstico de mercado
Finanzas	Hipotecas, consejos de préstamos, bonos corporativos, análisis de finanzas corporativas, predicción de precios corrientes, portafolio de inversión, identificación de falsificaciones.
Seguros	Evaluación de aspirantes, optimización de productos
Manufactura	Control de proceso de manufactura, análisis y diseño del producto, modelado dinámico, análisis de mantenimiento de máquinas.
Medicina	Análisis de células cancerígenas, diseño de prótesis, optimización de trasplantes, mejora de la calidad hospitalaria, predicción de reacciones adversas en los medicamentos.
Petróleo y gas	Exploración, sensores inteligentes, decisiones de tratamiento, interpretación sísmica
Robótica	Controladores, vehículos autónomos, control de trayectorias
Discursos	Reconocimiento de discursos, clasificación de vocales, síntesis de textos
Mercado accionario	Análisis de mercado, clasificación de bonos, venta de activos
Telecomunicaciones	Compresión de datos e imágenes, servicios de información automáticos, procesamiento de pagos de clientes, exploración de base datos.
Transporte	Calendarización de vehículos, sistema de rutas, temporización de los semáforos.

Elaboración propia con información de (Hagan, Demuth, Hudson, & De Jesús, 2018) y (Matich, 2001).

La aplicación de las redes neuronales artificiales también se puede abordar desde su área de estudio, es decir: como asociación y clasificación; regeneración de patrones; regresión y generalización; y optimización. Como parte de este trabajo de investigación se da mayor importancia a la aplicación de regresión y generalización; o bien, predicción de datos, ya que como objetivo principal se tiene la estimación de la demanda turística en México. Además, parte de la literatura menciona que las aplicaciones de las redes neuronales en predicción caen en campos como: el pronóstico del desempeño del mercado de valores, el uso como herramienta para predecir estructuras proteínicas secundarias en el cuerpo humano, así como el pronóstico para aumentar el éxito de un trasplante de corazón (Warner & Misra, 1996). Como se puede observar, no todas las aplicaciones son exclusivamente técnicas o industriales, sino que hay un amplio espectro en el cual esta herramienta forma parte.

Diseño de una red neuronal artificial

Elementos

Los elementos principales que componen una red neuronal artificial son: las capas; los pesos; las funciones de entrada, de activación, y de salida.

La distribución de las neuronas dentro de una red se realiza por medio de capas; es decir, de un conjunto de neuronas a un mismo nivel. De manera que, de acuerdo a su arreglo se pueden distinguir entre:

1. Capas de entrada: es la que recibe directamente la información proveniente de fuentes externas.
2. Capas ocultas: son internas a la red y no tienen contacto directo con la información del exterior. Estas capas pueden variar desde cero a un número muy alto, y sus conexiones pueden ser simples o muy complejas, todo de acuerdo con la topología de la red. Los científicos y estudiosos del tema no han logrado determinar el número exacto de capas ocultas y el número de neuronas por capa oculta que optimice el proceso de minería de datos que se desee realizar, por lo que se debe verificar su número a prueba y error (Matich, 2001).
3. Capas de salida: son las capas que transfieren la información hacia el exterior de la red.

En la siguiente ilustración se puede apreciar el arreglo de las capas, así como su ubicación dentro de la red.

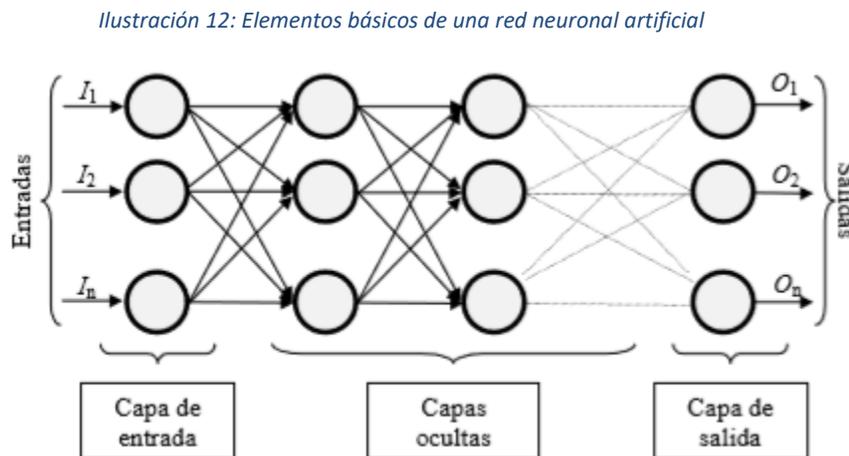


Ilustración tomada de (Matich, 2001)

Por otra parte, los pesos son un factor muy importante dentro de los componentes de la red neuronal, ya que son los responsables de ponderar una entrada. Son parámetros numéricos que determinan que tan fuerte es el efecto o relación que una neurona tiene con otra. Es similar a la pendiente de un modelo de regresión lineal, y de manera biológica es la activación o desactivación de una neurona en el proceso sináptico. Dentro del rubro de los pesos, es necesario mencionar al "bias", el cual es un parámetro adicional que se usa para ajustar la salida a lo largo de la suma ponderada de las entradas. Haciendo la comparación con un modelo de regresión, el bias sería la ordenada al origen de la ecuación lineal. (Ciaburro & Venkateswaran, 2017) En ocasiones se considera al bias como el peso de la neurona cero, ya que es solo un elemento de ajuste.

Dentro de los elementos de las funciones, la función de entrada se conoce como la entrada global al sistema, la cual se calcula a través de un vector entrada:

$$entrada_i = (in_{i1} \cdot w_{i1}) * (in_{i2} \cdot w_{i2}) * \dots * (in_{in} \cdot w_{in})$$

Donde el operador “*” representa la operación adecuada (máximo, suma, producto, etc.) dependiendo de las características de los datos que se utilicen, con n refiriéndose al número de entradas a la neurona N_i y w_i al peso. Como los pesos se multiplican por la entrada de los valores, éstos son los que ponderan la influencia de cada una de las entradas (Matich, 2001).

Las funciones de entrada más comunes son:

1. Suma de las entradas pesadas

$$\sum_j (n_{ij} w_{ij}) \quad \forall j = 1, 2, \dots, n$$

2. Producto de las entradas pesadas

$$\prod_j (n_{ij} w_{ij}) \quad \forall j = 1, 2, \dots, n$$

3. Máximo de las entradas pesadas

$$Max_j (n_{ij} w_{ij}) \quad \forall j = 1, 2, \dots, n$$

Para este caso, se utiliza la función de entrada de la suma de las entradas pesadas, en donde se obtendrá la suma ponderada de las entradas más el bias en cada una de las neuronas, la que a su vez será la entrada para la función de activación, en las siguientes ilustraciones se puede apreciar con mucho más detalle los elementos que componen a las redes neuronales artificiales.

La función de activación es una función matemática que convierte la entrada en una salida, sin ella el trabajo de la red neuronal sería similar a funciones lineales. Su importancia radica en que los problemas para los cuales se utilizan son no lineales. Existe una variedad de funciones de activación que se pueden emplear, sin embargo, las principales son: la función lineal, el escalón unitario, la función sigmoid, la función tangente hiperbólica y la función unidad lineal rectificada. (Ciaburro & Venkateswaran, 2017)

Una pregunta obligada que surge al momento de hablar de las funciones de activación es saber que función utilizar, dado que existe una variedad de ellas. Los autores de “Neural Networks with R” (Ciaburro & Venkateswaran, 2017) sugieren que no importa la función que se utilice siempre y cuando se cumplan tres características: la función debe ser diferenciable, debe ser de rápido procesamiento y simple, y no debe estar centrada en cero. Por lo tanto, debido a las sugerencias, se ha llegado a la conclusión que la función de activación óptima para el uso de las redes neuronales es la función ReLU o unidad lineal rectificada. Las gráficas y características de cada una de las funciones se pueden apreciar en la sección de anexos.

Por último, la función de salida es el valor resultante de todas las operaciones. Determina el valor que será transferido a las demás neuronas conectadas. Si la función de activación está por debajo de un umbral determinado (valor que se utiliza como referencia, generalmente propuesto por el diseñador), ninguna

salida pasa a la neurona siguiente. Es por ello que los valores de salida, normalmente se restringen a rangos $[0,1]$, $[-1,1]$ o valores binarios $\{0,1\}$, $\{-1,1\}$. Los tipos de funciones de salida más comunes son la función identidad y la binaria. (Matich, 2001)

Ilustración 13: Relación de los elementos de la neurona artificial

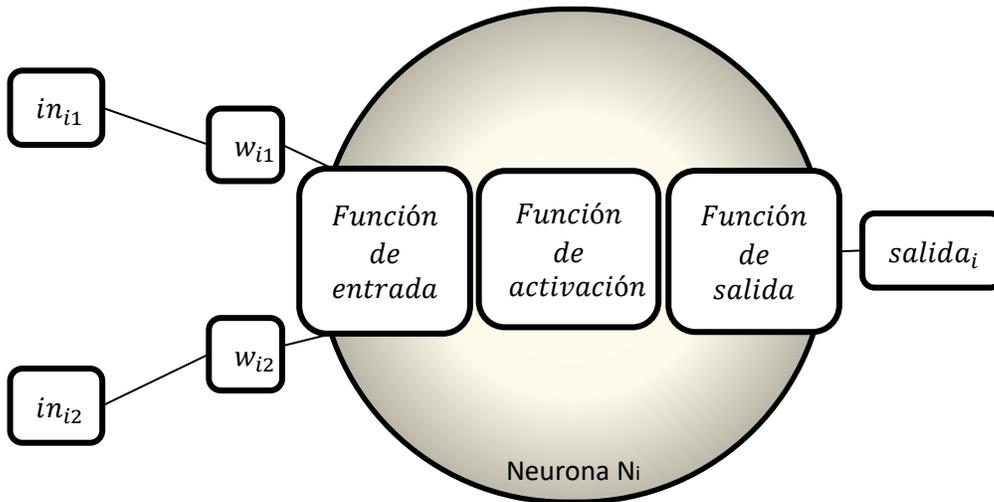


Ilustración tomada de (Matich, 2001)

Ilustración 14: Elementos de un Perceptrón

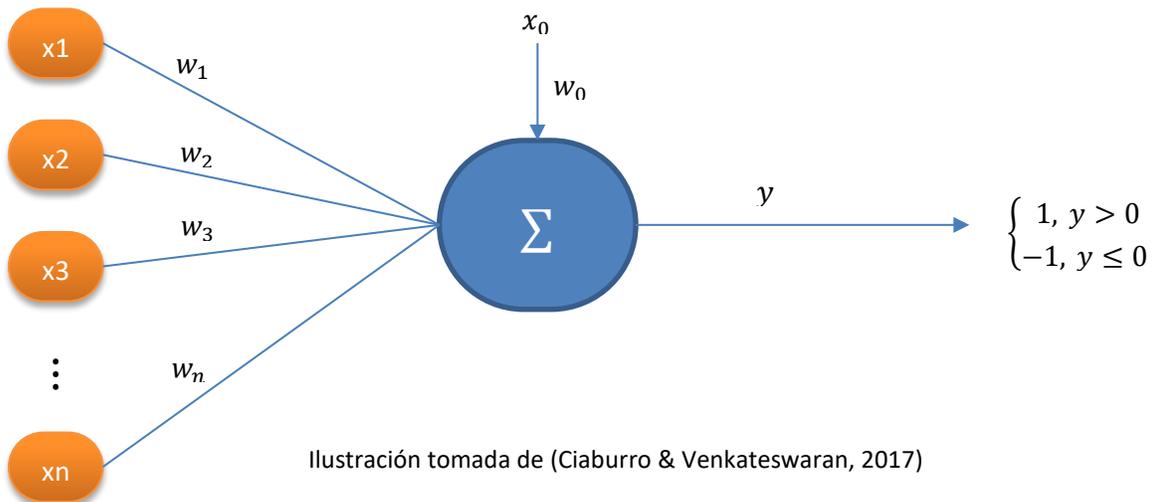


Ilustración tomada de (Ciaburro & Venkateswaran, 2017)

Arquitecturas

La topología o arquitectura de una red es la organización y disposición de las neuronas dentro de las capas que conforman la red. Los parámetros fundamentales son: el número de capas, el número de neuronas

por capas, el grado de conectividad y el tipo de conexión entre ellas. En la siguiente tabla se presentan los tipos de arquitecturas más socorridos por los teóricos en el diseño de una red neuronal artificial.

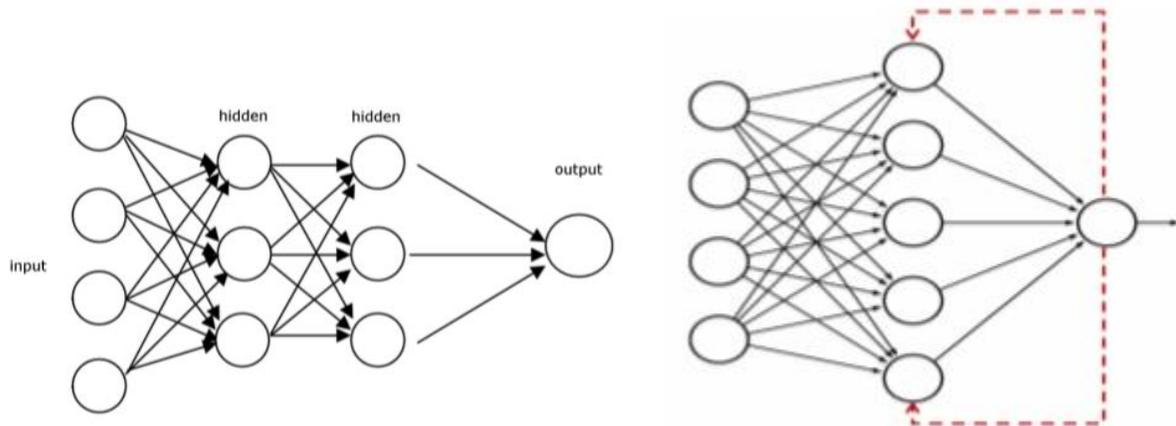
Tabla 4: Arquitecturas de redes neuronales artificiales

Tipo de red	Descripción
Monocapa (un solo nivel)	Se establecen conexiones entre neuronas que pertenecen a la única capa de la red. Se utiliza para tareas de autoasociación, regenerar información de entrada que está incompleta o distorsionada.
Multicapa (diferentes niveles)	Conexiones autorecurrentes: la señal puede ser la entrada de la misma neurona
	Conexiones hacia adelante o Feedforward: las salidas solo son entradas de capas posteriores.
	Conexiones hacia atrás o Feedback: cuando las salidas pueden ser entradas de niveles previos, cuando tienen lazos cerrados, se denominan sistemas recurrentes.
Heteroasociativa	Se refiere al caso en que la red aprende parejas de datos. De tal forma que cuando la información de entrada sea A_i , deberá responder generando la salida asociada B_i .
Asociativa	La red aprende ciertas informaciones A_1, A_2, \dots de tal forma que cuando se presenta la información de entrada realizará una autocorrelación, respondiendo con uno de los datos almacenados, más parecido a la entrada.

Elaboración propia con información recuperada de (Ciaburro & Venkateswaran, 2017)

Enseguida se presentan dos ilustraciones que ejemplifican las redes neuronales multicapa. La ilustración de la izquierda con una conexión hacia adelante o Feedforward, y la de la derecha con una conexión hacia atrás o Feedback.

Ilustración 15: Redes multicapa feedforward y feedback



Ilustraciones tomadas de (Ciaburro & Venkateswaran, 2017)

Debido al tipo de conexión que se maneja entre las neuronas, es la complejidad que se le adjudica a la red. Para este trabajo de investigación se optará por una topología con conexión hacia adelante. Solo en

el proceso de aprendizaje se utilizará un algoritmo recursivo con propagación hacia atrás, el cual será detalladamente explicado en la siguiente sección.

Por último, en cuanto a los tipos de topologías que se pueden manejar en el diseño de una red neuronal artificial, se presenta a continuación una serie de modelos especiales que conforman la línea de investigación de este campo.

En primer lugar se puede mencionar al Perceptrón, el cual es una sola neurona de una sola capa. Luego, el elemento lineal adaptativo (ADALINE), el cual es un simple Perceptrón que puede resolver solo problemas lineales pero con la característica de ser adaptativo. Esto, a su vez, nos lleva a la estructura MADALINE, la cual es una conformación de múltiples ADALINEs, es decir, una red multicapa de unidades ADALINE. Entre las topologías con conexión hacia adelante o Feedforward, se identifican aquellas que se han utilizado como un autoencoder, con fines probabilístico, aquellas que tienen un retraso de tiempo y las que utilizan una función de convolución. Por último, dentro del diseño de redes neuronales se identifica a las redes recurrentes, dentro de las que destacan: la estructura de Hopfield, la máquina de Boltzmann, los mapas auto organizativos (SOM), la memoria asociativa bidireccional (BAM) y la memoria de largo y corto plazo (LSTM) (Matich, 2001).

No cabe duda que a pesar de existir una cantidad diferente de estructuras y formas de conexiones entre neuronas, los estudios seguirán ampliándose cada vez más.

Entrenamiento, Validación y Desarrollo

Después que se ha definido la estructura de la red neuronal y se han identificado las entradas y las salidas, es necesario diseñar el algoritmo de entrenamiento de la red, la validación de los resultados que se han obtenido hasta ese momento y, con ello, generar el modelo final.

Es por ello que el proceso de entrenamiento o acondicionamiento es el algoritmo de aprendizaje con el cual una red neuronal debe aprender a calcular la salida correcta para cada arreglo de entradas de un conjunto de ejemplos. Debido a que es muy complejo y poco práctico cambiar la arquitectura de la red, así como los vectores de entrada y salida; el aprendizaje está enfocado a cambiar los pesos entre las conexiones de las neuronas, es decir, una adaptación de los pesos. Al igual que ocurre en las neuronas biológicas, el cambio de las conexiones entre ellas en el proceso sináptico produce modificaciones en todo el sistema.

Es importante aclarar que antes de comenzar con el entrenamiento, es necesario escoger un conjunto inicial de pesos para las diversas conexiones, de manera que el algoritmo pueda tener un valor de referencia y posteriormente, con el algoritmo seleccionado, los pesos se irán ajustando. Se puede otorgar un valor aleatorio entre dos números naturales $[-n, n]$, sin embargo se tendrá que ser muy precavido al seleccionar estos valores iniciales, ya que el tiempo de convergencia del algoritmo depende indudablemente de los valores iniciales que se hayan asignado. (Matich, 2001)

Se dice que una red neuronal ha aprendido cuando la modificación de los pesos ha terminado, es decir, su variación respecto al tiempo es cero ($\frac{dw_{ij}}{dt} = 0$).

La manera en que las neuronas cambian los pesos en el proceso de aprendizaje, se basa principalmente en dos metodologías: en un aprendizaje supervisado y en un aprendizaje no supervisado (Ciaburro & Venkateswaran, 2017) (Matich, 2001).

En primer lugar, el aprendizaje supervisado se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo que determina la respuesta que debería generar la red a partir de una entrada determinada, es decir, comparar la salida con los valores reales. En caso que la salida no sea la esperada, se procede a modificar los pesos en las conexiones. Dentro de este primer bloque se pueden identificar tres tipos de entrenamiento:

El aprendizaje por corrección de error consiste en ajustar los pesos en función de la diferencia entre los valores reales y los pronosticados. Algunos ejemplo de este tipo de aprendizaje son la regla de aprendizaje del perceptrón, la regla del aprendizaje delta o regla del mínimo error cuadrado (LMSE) y la regla de aprendizaje de propagación hacia atrás o “Backpropagation”, el cual se considera un método LMSE multicapa. En él, el error es propagado hacia atrás desde la capa de salida, lo cual permite que los pesos de las conexiones de las capas ocultas también se ajusten, lo que aumenta la exactitud del modelo. Es por ello que este último algoritmo de aprendizaje es preferido sobre los demás métodos de entrenamiento.

Por otro lado, se encuentra el *aprendizaje por refuerzo*, el cual se basa en la idea de no comparar un comportamiento deseado, sino solo indicar mediante una señal si la salida obtenida en la red se ajusta a la deseada, es decir, si la respuesta se encuentra dentro de un rango establecido, aun cuando no sea el valor buscado, el algoritmo tendrá una señal positiva; pero si la respuesta no está dentro del rango, la señal será negativa. Esto solo ayuda al modelo a guiar su aprendizaje, más no le permite captar toda la esencia de los datos.

Por último, el *aprendizaje estocástico* consiste en realizar cambios aleatorios en los valores de los pesos de las conexiones y evaluar su efecto a partir del objetivo deseado y distribuciones de probabilidad. De manera que, mediante una simulación se evalúan diferentes escenarios hasta obtener la combinación de pesos de menor variación y que resulte con el menor error posible. Este método tiene la desventaja en el tiempo de convergencia, debido a que si los pesos inicialmente asignados no se encuentran próximos a los valores óptimos, el algoritmo tardará en encontrar la solución.

El segundo gran bloque de la metodología de entrenamiento de una red neuronal es el aprendizaje no supervisado, el cual tiene la característica de no requerir una influencia externa para ajustar los pesos de las conexiones. El algoritmo se centra en la tarea de encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los términos de la base de datos. Para este tipo de entrenamiento, se distinguen dos tipos: el aprendizaje hebbiano y el competitivo-comparativo.

En primer lugar, el *aprendizaje hebbiano* intenta encontrar la familiaridad entre neuronas. Si dos neuronas toman el mismo estado (activo o inactivo) de manera simultánea, el peso entre las neuronas incrementa. En el caso que los estados de dos neuronas sean distintos, los pesos decrecen. Es así como se ajustan los

pesos, mediante incrementos o decrementos de los mismos. El segundo tipo, el *aprendizaje competitivo-comparativo*, se orienta a la clasificación de los datos de entrada. Esto es, si un patrón es reconocido, los pesos se comportan de una forma diferente a si el patrón no es reconocido. Estos tipos de entrenamiento no supervisado son más socorridos en problemas de clasificación, ya que no se tiene ninguna referencia con la cual se pueda comparar los resultados de salida.

Algunos autores mencionan otro tipo de clasificación de los métodos de aprendizaje de las redes neuronales, el aprendizaje on-line y el aprendizaje off-line (Matich, 2001). En realidad este tipo de entrenamientos son formas en las cuales se lleva a cabo el proceso de aprendizaje, ya que en el aprendizaje on-line la red aprende durante su funcionamiento habitual, mientras que en el off-line se debe desconectar o inhabilitar la red hasta que el proceso de entrenamiento termine. Sin embargo, los tipos principales de algoritmos que se emplean para que una red aprenda son los dos grandes bloques que se mencionaron anteriormente. En la siguiente tabla se presenta el concentrado de la información de los tipos de aprendizajes para las redes neuronales.

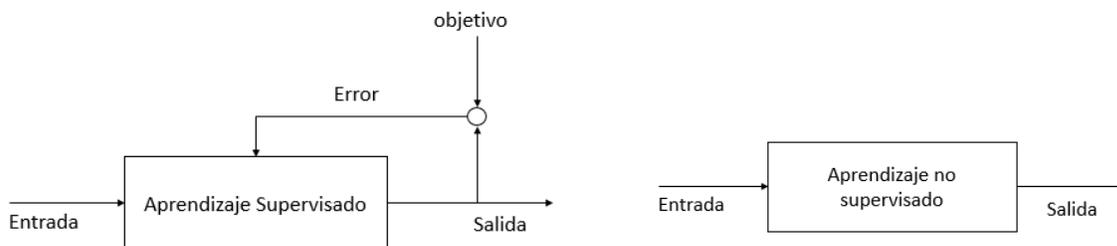
Tabla 5: Tipos de aprendizaje para redes neuronales

Aprendizaje supervisado	Corrección por error	Aprendizaje del perceptrón
		Regla del mínimo error cuadrado (LMSE)
		Backpropagation (LMSE multicapa)
	Por refuerzo	
	Estocástico	
Aprendizaje no supervisado	Hebbiano	
	Competitivo y comparativo	
Aprendizaje on-line		
Aprendizaje off-line		

Elaboración propia con información de (Matich, 2001), (Ciaburro & Venkateswaran, 2017).

En la siguiente ilustración se pueden observar los dos tipos de aprendizajes que se pueden emplear para entrenar una red neuronal. En la figura de la izquierda se presenta el aprendizaje supervisado, el cual, tal como se explicó, compara los resultados de las salidas con los valores objetivos y se va ajustando de manera que se obtenga el menor error posible. Por otro lado, en la figura de la derecha se muestra el aprendizaje no supervisado, el cual no tiene ninguna influencia externa o valores sobre los cuales comparar.

Ilustración 16: Diagramas de flujo de los procesos de aprendizaje



Elaboración y traducción propia con información de (Ciaburro & Venkateswaran, 2017)

De manera general, la forma de calcular los cambios de los pesos es:

$$\text{Peso Nuevo} = \text{Peso Viejo} + \text{Cambio de Peso}$$

En donde la representación matemática es:

$$w_{ij}(t + 1) = w_{ij}(t) + \Delta w_{ij}(t)$$

Sin embargo, existe una teoría detrás de esta expresión que fundamenta la variación de los pesos con respecto al tiempo. Es tarea de la siguiente sección explicar detalladamente el proceso de aprendizaje supervisado de propagación de los errores hacia atrás para tener un mejor entendimiento del trabajo de las redes neuronales y no pensar todo como un proceso de caja negra.

Tanto en la parte de entrenamiento como en cualquier otra fase de la minería de datos que se realice, es indispensable tener en consideración que las redes neuronales emplean datos numéricos como entradas, de manera que es necesario hacer una codificación de los atributos numéricos para que el valor nominal, no tenga una injerencia en el resultado, o en su defecto, convertir variables alfanuméricas en variables numéricas. Esto se realiza mediante una función de transformación (Anysz, Zbiciak, & Ibadov, 2016), las cuales permiten hacer comparables los valores, ya que son escalados. Las expresiones matemáticas de las principales funciones de transformación se presentan a continuación.

Tabla 6: Funciones de estandarización de datos

Nombre de la función de estandarización	Expresión matemática
Estandarización Vectorial	$A_i = \frac{A_{0i}}{\sqrt{\sum_{i=1}^n (A_{0i})^2}}$
Estandarización Manhattan	$A_i = \frac{A_{0i}}{\sum_{i=1}^n A_{0i} }$
Estandarización lineal máxima	$A_i = \frac{A_{0i}}{\max A_{0i}}$
Estandarización lineal de Weitendorf	$A_i = \frac{A_{0i} - \min A_{0i}}{\max A_{0i} - \min A_{0i}}$
Estandarización no lineal de Peldschus	$A_i = \left(\frac{A_{0i}}{\max A_{0i}}\right)^2$
Estandarización logaritmica de Zavadskas y Turskis	$A_i = \frac{\ln(A_{0i})}{\ln \prod_1^n A_{0i}}$

Elaboración propia con información de (Anysz, Zbiciak, & Ibadov, 2016)

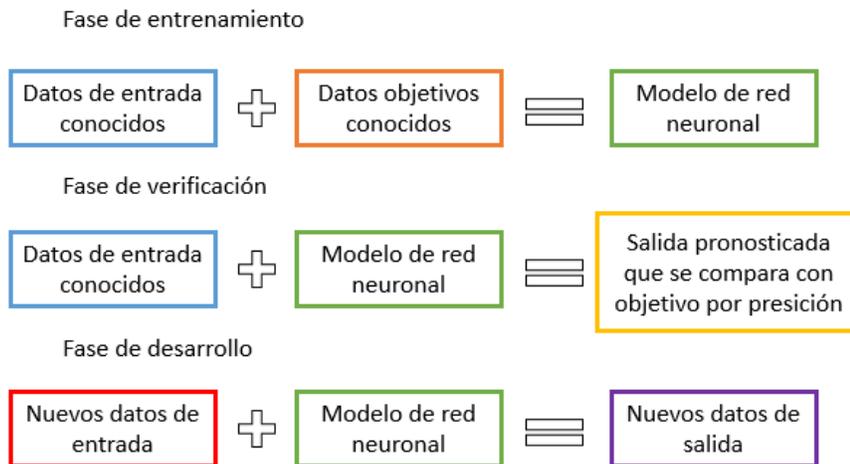
En donde A_i es el elemento i de los datos después de la estandarización, A_{0i} es el elemento i de los datos antes de la estandarización.

Con el estudio realizado por (Anysz, Zbiciak, & Ibadov, 2016) se sabe que no hay una mejor forma de estandarización, debido a que la arquitectura de la red influye mucho en el cálculo de los errores. De manera que el criterio de selección de la técnica de estandarización se hará por simplicidad, es decir, la expresión matemática que sea más sencilla de utilizar.

Una vez que se ha concluido con el proceso de entrenamiento, es necesario verificar que la red resuelva los problemas para los cuales fue creado, es por ello que se requiere de otro conjunto de datos denominado conjunto de validación o testeo. Con el cual se verifica que los pronósticos, para este caso particular de redes neuronales, sean iguales o muy próximos a los datos originales. Para determinar cuándo se tiene que detener el algoritmo, es necesario establecer una condición de detención: ya sea cuando el error cuadrado de los entrenamientos haya encontrado un mínimo; cuando se esté por debajo de cierto valor (umbral); o se hayan completado cierto número de ciclos (“epoch”: corresponde a una iteración o paso a través del proceso de proveer a la red con una entrada y actualizar los pesos). Una vez que se alcanzó la condición de detención, los pesos ya no cambian. Se dice, entonces, que se ha encontrado el modelo que logra emular el comportamiento de los datos. (Ciaburro & Venkateswaran, 2017)

A continuación se presenta un diagrama con el cual se comprende la razón de ser de cada una de las fases de entrenamiento, verificación y desarrollo.

Ilustración 17: Entrenamiento, verificación y desarrollo de una red neuronal



Elaboración propia con información de (Ciaburro & Venkateswaran, 2017)

Aprendizaje: Propagación hacia atrás (Backpropagation)

El aprendizaje supervisado de propagación hacia atrás es un algoritmo que, como su nombre lo indica, se fundamenta en propagar los errores hacia atrás. Se puede hacer la analogía con un dicho popular: “aprender de los errores”, ya que aprendemos de nuestros errores en cada iteración, hasta que se alcanza la convergencia. El objetivo es ajustar los pesos en cada capa y minimizar la variación. En la siguiente

imagen se puede observar esquemáticamente el proceso de backpropagation. Así mismo, se enlistan una serie de pasos con los cuales se pretende explicar el funcionamiento de este método de aprendizaje.

Ilustración 18: Representación esquemática de backpropagation

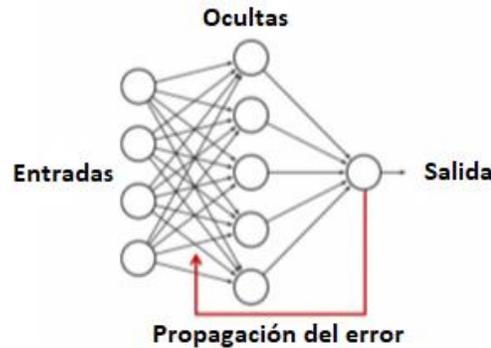


Ilustración editada de original (Ciaburro & Venkateswaran, 2017)

Los pasos con los cuales se inicia el proceso de entrenamiento de la red neuronal mediante un aprendizaje supervisado, con la propagación hacia atrás de los errores es la siguiente:

1. Iniciar la red neuronal, asignando los pesos y las biasses de cada neurona en las capas ocultas
2. Obtener la entrada de cada neurona (suma(peso*entrada) + bias)
3. Aplicar la función de activación
4. Tomar la salida y pasarla a la siguiente capas de neuronas
5. Si la capa es la de salida, aplicar (suma(peso*entrada) + bias) a cada neurona
6. De nuevo aplicar la función de activación a la capa de salida
7. Esa es la salida de la red
8. Con los datos de entrenamiento, identificar los errores en cada salida de las neuronas, restando la salida real
9. El total de los errores se calcula con alguna expresión matemática de medida de error, por ejemplo: $E = \frac{1}{2}(t - y)^2$
10. La técnica del gradiente descendiente requiere del cálculo de una derivada parcial del error respecto con el peso. Esto es posible gracias a la regla de la cadena.

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}}$$

11. El primer factor es la derivada parcial del error con respecto a la salida de la neurona j

$$\frac{\partial E}{\partial o_j} = \frac{\partial E}{\partial y} = \frac{\partial}{\partial y} \frac{1}{2}(t - y)^2 = y - t$$

12. El segundo factor es la derivada parcial de la salida de la neurona j con respecto a sus entradas, la cual es la derivada parcial de la función de activación

$$\frac{\partial o_j}{\partial net_j} = \frac{\partial}{\partial net_j} \varphi(net_j) = \varphi(net_j)(1 - \varphi(net_j))$$

13. El tercer término es simplemente la neurona j

$$\frac{\partial net_j}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \left(\sum_{k=1}^n w_{kj} o_k \right) = \frac{\partial}{\partial w_{ij}} w_{kj} o_k = o_i$$

14. Combinando los pasos 11,12 y 13

$$\frac{\partial E}{\partial w_{ij}} = o_i \delta_j \quad \text{donde}$$

$$\delta_j = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} = \begin{cases} (o_j - t_j) o_j (1 - o_j), & \text{si } j \text{ es una neurona de salida} \\ \left(\sum_{\ell} \delta_{\ell} w_{j\ell} \right) o_j (1 - o_j), & \text{si } j \text{ es una neurona interna} \end{cases}$$

15. Los pesos w_{ij} de cada neurona son actualizados con esta derivada parcial, combinada con la tasa de aprendizaje.

De manera que una vez que se alcanzó la actualización de los pesos y se aplicó la regla de paro del algoritmo, se puede continuar con el proceso de verificación y desarrollo que se abordaron anteriormente. Es importante mencionar que la propagación hacia atrás de los errores está basada en la derivada parcial de los pesos con respecto al tiempo. De ahí la importancia que la función de activación sea diferenciable.

Ventajas y desventajas

Finalmente, es necesario precisar que como todo tipo de método de modelado de datos, las redes neuronales tienen ventajas y desventajas que lo caracterizan. (Matich, 2001) (Warner & Misra, 1996) (Ciaburro & Venkateswaran, 2017)

Entre las ventajas se pueden mencionar:

1. Una red neuronal es capaz de aproximar cualquier función continua a cualquier grado de precisión. Es por ello que también son conocidas como estimadores universales de funciones. Para ello no es necesario tener una noción del comportamiento de los datos.
2. Se puede implementar un aprendizaje adaptativo; esto es, la capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial. Son capaces de adaptarse a las nuevas condiciones, bajo el patrón que hayan aprendido. Por lo que es tarea del diseñador de la red obtener la arquitectura adecuada y el algoritmo de aprendizaje óptimo, ya que el proceso de discriminación lo realiza la red por sí sola.
3. Son capaces de tener una auto-organización. Una red neuronal puede crear su propia organización de la información que recibe mediante una etapa de entrenamiento, esto permite que la red tenga la facultad de responder ante situaciones a las que no ha sido entrenada previamente.
4. Una red neuronal es tolerante a fallos. Gracias a que la información que se almacena en las redes neuronales es no localizada, se puede tener información distribuida en las conexiones de la

misma, es por ello que a pesar de una destrucción parcial de la estructura de la red, puede continuar con su función con cierto grado de alteración. Además es capaz de reconocer patrones con ruidos, distorsionados o incompletos.

5. Una red neuronal artificial opera en tiempo real, es decir, los cálculos realizados son ejecutados en paralelo.
6. Debido a sus características presenta una fácil inserción dentro de la tecnología existente, ya que el proceso de entrenamiento, verificación e instalación son relativamente cortos, se pueden emplear para tareas específicas en dispositivos particulares.

Entre las desventajas de las redes neuronales se mencionan:

1. Las redes son difíciles de interpretar. Debido a que mientras más capas ocultas se tengan, el sistema se comienza a entender como un sistema de caja negra, ya que solo se pueden apreciar las entradas y las salidas, y no la influencia de las variables independientes.
2. Dependiendo de las variables iniciales (pesos y bias) es la convergencia del algoritmo, puede llegar a ser muy lenta.
3. Depende de muchos datos. Esto lleva a un problema de sobre ajuste (overfitting) y generalización.

RESUMEN DEL CAPÍTULO

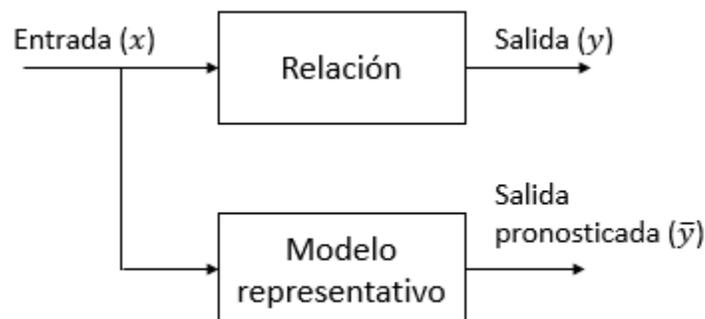
En este capítulo se habló de los antecedentes de las redes neuronales artificiales, para que cualquier persona que esté relacionada con el tema o no, pueda comprenderlo sin dificultad. Se comienza con la definición de red neuronal y se hace la comparación entre las neuronas biológicas y las artificiales para comprender las similitudes que dieron origen a este campo de investigación. Se continúa con la historia de las redes neuronales, es decir, estudios cronológicos que se han realizado en este campo y las aplicaciones de esta herramienta. Posteriormente se presenta una sección en la cual se especifican las partes, elementos y formas de diseñar una red neuronal, ya que dependiendo de la finalidad será la arquitectura que se deberá emplear. Por último se presenta una serie de ventajas y desventajas para dejar en claro que, como cualquier modelo que busca emular la realidad, existirán tanto beneficios como inconvenientes.

Capítulo 4: Construcción del modelo con redes neuronales artificiales y su comparación con las técnicas convencionales de pronóstico

El método científico que se utiliza para este trabajo de investigación, pertinente a la elaboración del modelo de pronóstico, está basado en el diseño de proceso de la minería de datos. Por lo que resulta imperante definir esta metodología, sus características y pasos. De acuerdo con los autores de “Predictive Analytics and Data mining” (Kotu & Deshpande, 2015), se sabe que la minería de datos es una herramienta que permite encontrar patrones en una base de datos existente, se refiere al descubrimiento de conocimiento, machine learning y análisis predictivo. De acuerdo con (Kotu & Deshpande, 2015), todos estos métodos computacionales han sido derivados de los campos de la estadística, machine learning y la inteligencia artificial, por lo que resulta pertinente utilizar este método científico para abordar el diseño y desarrollo de la red neuronal artificial para predicción de los cuartos de hotel ocupados en México. En la sección de anexos se puede observar la imagen que muestra las actividades que se pueden realizar con la herramienta de minería de datos, en donde destacan temas como series de tiempo, regresión, pronósticos y clasificación.

El objetivo principal de la minería de datos es encontrar conclusiones útiles que puedan ser utilizadas por los analistas. Como parte del análisis predictivo y las redes neuronales como herramienta para predecir valores, se ha encontrado que la minería de datos es el proceso con el cual se construye un modelo representativo que se ajusta a los datos originales. La siguiente imagen muestra como las entradas y salidas de un proceso están vinculadas por una relación intrínseca entre cada parte. El objetivo que se tiene es, a partir de los mismos datos de entrada, generar una salida que mejor se ajuste a las observaciones reales, todo esto mediante un modelo único que sea capaz de emular la relación entre los datos.

Ilustración 19: Relación de entradas y salidas en minería de datos



Traducción propia con información de (Kotu & Deshpande, 2015)

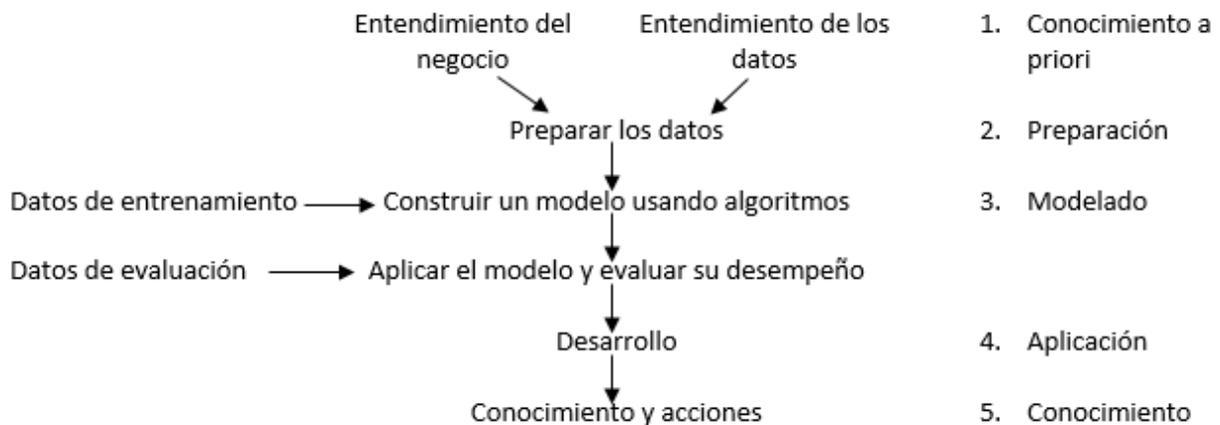
Las redes neuronales artificiales, como parte de las aplicaciones de la minería de datos, enfoca el modelado del problema en un algoritmo que mejor predice las variables de salida a partir de las entradas, al “aprender” de los errores predictivos de pasos previos. Tal como se mencionó en el *Capítulo 3: Antecedentes y aplicación de las redes neuronales artificiales*, en el apartado de *Diseño de una red neuronal artificial* es mediante el proceso de aprendizaje supervisado o no supervisado que una red neuronal logra aprender de sus errores (Kotu & Deshpande, 2015). Para este trabajo de investigación se opta por emplear el aprendizaje supervisado para entrenar la red neural, debido a que es posible tener las variables de salida que se espera pronosticar y poder compararlas con una medida de error, para evaluar su desempeño.

El proceso estándar de la metodología de minería de datos, utilizado en este caso para redes neuronales específicamente, consiste en cinco pasos (Ciaburro & Venkateswaran, 2017):

1. Entendimiento del problema
2. Preparación de muestras de datos
3. Desarrollo del modelo
4. Aplicación del modelo en el conjunto de datos para observar su comportamiento en el mundo real
5. Despliegue de producción o generación de conocimiento

En la siguiente imagen se puede observar el proceso completo de la metodología de minería de datos, que comienza con el conocimiento a priori de lo que se quiere lograr, hasta terminar con la generación adicional de conocimiento.

Ilustración 20: Proceso de minería de datos



Traducción propia con información de (Kotu & Deshpande, 2015)

En breve se presenta la metodología aplicada a las redes neuronales artificiales.

Para el primer paso, conocimiento a priori, que comprende el entendimiento tanto del problema como de los datos; se inició con una lluvia de ideas para definir la aplicación de las redes neuronales como herramienta de pronóstico para la solución de problemas. Se llegó a la conclusión de utilizar dicha técnica de pronóstico para predecir la demanda turística de México debida a los principales países socios económicos, porque tal como se explicó en el *Capítulo 1: Turismo como factor importante para el desarrollo del país y factores que influyen en la demanda turística*, el turismo es un pilar para el desarrollo de un país, ya que tiene un impacto positivo y multiplicador en muchos sectores económicos.

En cuanto al segundo paso, la preparación de los datos, se realizó una investigación de las variables relevantes que impactan en la demanda turística de cualquier país y que deben ser analizadas si se desea tener una predicción fidedigna de la realidad. Tal como se explica en la segunda parte del primer capítulo, se enlistan una serie de variables que se recomiendan emplear para el diseño del modelo de regresión no paramétrica y no lineal, en este caso, de redes neuronales artificiales. Las variables independientes o entradas que se emplearán para el desarrollo del modelo son:

1. Ingresos del país origen
2. Precio relativo de servicio
3. Relación del tipo de cambio
4. Población del país origen
5. Factores estacionales y
6. Tasa de inseguridad

Mientras que la variable de salida o dependiente es el número de cuartos de hotel ocupados en la República Mexicana, variable indispensable para poder comparar la capacidad instalada y determinar si es suficiente o no para abastecer la creciente demanda turística de los siguientes años.

Se han seleccionado dichas variables debido a su participación en diferentes modelos de pronóstico y regresión de la demanda turística de un país, en donde la literatura respalda su uso, tal como se fundamenta con el meta-análisis expuesto en el primer capítulo. No obstante, se vio en la necesidad de seleccionar las variables condicionadas a la disponibilidad de la información, ya que por desgracia, la información histórica requerida de todas las variables que se consideraban relevantes no está disponible en los organismos y secretarías federales.

Los países que serán analizados con la herramienta de redes neuronales artificiales para predecir su demanda turística para los próximos seis años son Estados Unidos de América y Canadá, ya que son los principales países socios de relaciones económicas con México, son los países que en conjunto generan más del 85% de los turistas en México, y porque son los países de los cuales se logró reunir la mayor cantidad de información histórica de las variables seleccionadas. Hay que puntualizar que para algunas variables seleccionadas (ingresos del país origen y el precio relativo de servicio), se utilizó un indicador similar que revela la información requerida, debido nuevamente a la ausencia de los datos, la imprecisión de las variables o su dificultad para ser medida.

Para este efecto se menciona, en primer lugar, que dado que el ingreso del país origen es una variable subjetiva y no precisamente medible, ya que no se puede simplemente obtener un valor medio de los

ingresos de todos los habitantes del país, se sustituye por la aproximación que se tiene con el PIB per cápita, ya que su disponibilidad es mucho más precisa que cualquier variable de tipo subjetivo. En cuanto al precio relativo de servicio, que comprende los costos de bienes y servicios tales como: hospedaje, transporte local, comida y entretenimiento; ya que, de igual forma que la variable anterior, es complicado tener un índice o indicador de precios turísticos de manera histórica, se puede aproximar esta variable con el índice de precios al consumidor (IPC) del origen y del destino que refleje dicha proporción.

Cuando se utiliza el IPC como factor para medir los precios de bienes y servicios, se debe ajustar por efectos de inflación y tipo de cambio, tal como recomienda (Lim, 1997) resultando en:

$$RP_{ij} = \frac{CPI_{it}}{CPI_{jt}} ER_{it}$$

RP_{ij} es la variable de precios relativos del destino i en el periodo t

CPI_{it} es el índice de precios al consumidor del destino i en el periodo t

CPI_{jt} es el índice de precios al consumidor en el origen j en el periodo t

ER_{it} índice de precios de tipo de cambio en terminos del tipo del cambio del destino i en el periodo t

Para los datos del IPC de Estados Unidos, Canadá y México, se emplearon los valores actualizados en dólares ajustados al 2010. Para la variable “Relación del tipo de cambio” se obtuvo los datos de la paridad del dólar americano y el dólar canadiense con el peso mexicano. Asimismo, se investigó la cantidad de personas registradas en Estados Unidos y Canadá como una serie temporal histórica.

La variable de los factores estacionales es empleada como variable “dummy” o “ficticia”, ya que se asigna un valor de (1) al mes en el cual se tenga una incidencia, un (0) en caso contrario. Esto se realiza de esa manera, ya que no hay una variable *per se* que sea capaz de absorber todo el comportamiento estacional de la serie. Por último, es importante hacer notar que para la variable de tasa de inseguridad, se recurrió a la Encuesta Nacional de Victimización y Percepción Sobre Seguridad Pública proporcionada por el INEGI, en donde se muestra la cantidad de incidencias delictivas que se registraron durante un periodo, a nivel estatal y nacional. Esta variable no forma parte del meta-análisis que realizó (Lim, 1997), sin embargo debe considerarse en el modelo; ya que, por desgracia, la última década de México se ha vivido bajo un ambiente violento que impacta negativamente en la percepción de otros países.

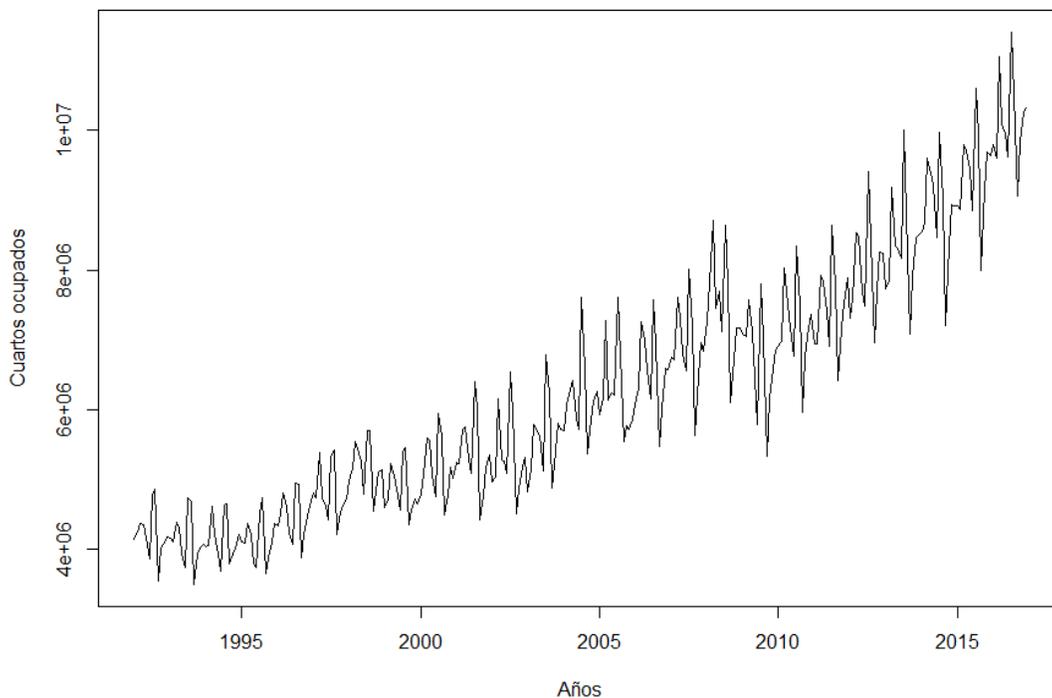
La variable independiente o salida del sistema, el número de cuartos de hotel ocupados, fue recuperada de los datos históricos del portal del Sistema Nacional de Información Estadística y Geográfica de Turismo (SNIETG), organismo federal que colabora en conjunto con la Secretaría de Turismo. En esta información se contabilizan los establecimientos registrados, hoteles de 1 a 5 estrellas, hoteles sin clasificación, clase económica, o cualquier tipo de alojamiento. Lo importante es considerar que son establecimientos registrados ante la Secretaría de Turismo, y por consecuencia ante la Secretaría de Hacienda y Crédito Público.

Para este trabajo de investigación la no incorporación de entes que ofrecen servicios de alojamiento como Airbnb, hostales o cualquier tipo de establecimiento no registrado públicamente, no representa un problema sustancial, debido a que esta investigación pretende, indirectamente, aumentar las ganancias del sector turístico a favor de los establecimientos públicos. Se fundamenta en el hecho de que los hoteles de 1 a 5 estrellas son los principales captadores de turistas de alta gama, es decir, aquellos cuya derrama económica es importante para el sector turístico.

De manera que, una vez que se definieron las variables explicativas y la variable predictiva del modelo, con sus consideraciones pertinentes, se tuvo que homogeneizar la base de datos para que todas las variables tuvieran la misma periodicidad. Los datos que se tienen comprenden desde el año 1992 hasta el 2016, con una frecuencia mensual. Se ha considerado la inclusión de los datos en ese rango de años, por su disponibilidad en sitios oficiales, las actualizaciones de las bases de datos, o para el caso de la tasa de inseguridad, por la fecha en que se comenzó con la encuesta. A continuación se presenta la gráfica de los datos históricos del número de cuartos de hotel ocupados en la República Mexicana durante dicho periodo.

Ilustración 21: Cuartos de hotel ocupados en la República Mexicana

Cuartos de hotel ocupados en la República Mexicana de 1992 a 2016



Elaboración propia a partir de (Team, 2017)

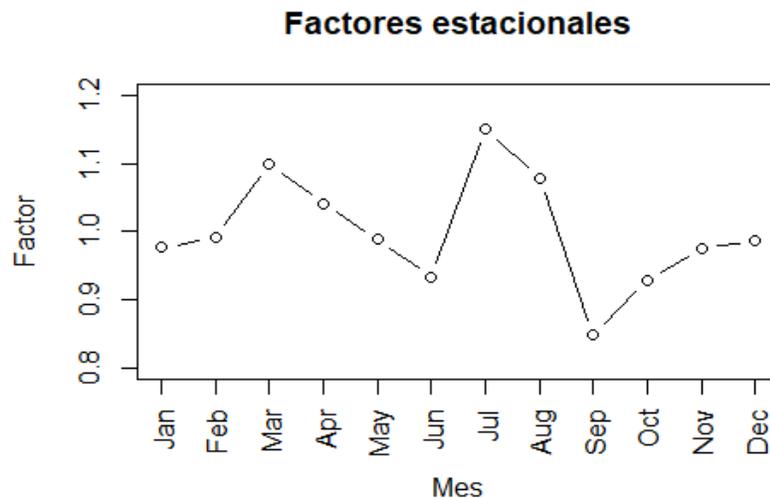
De manera general, se puede observar que la serie de datos temporal tiene un comportamiento de patrón tendencial creciente, con una estacionalidad marcada. En la sección de anexos se puede observar la gráfica de la serie de tiempo anterior, con la división estacional marcada, debida a los periodos estacionales considerados, es decir, los doce meses del año civil. Asimismo, en la sección de anexos se exhibe la descomposición de la serie en sus patrones: tendencial, estacional y aleatorio.

Si bien desde un inicio se puede observar que la serie temporal tiene un “buen comportamiento” (no se observan cambios bruscos de nivel, tendencia ni estacionalidad) y que al parecer no es necesario incorporar la técnica de Redes Neuronales Artificiales, ya que ello implica un mayor consumo de recurso (principalmente de cómputo) y de especialización de los analistas, es necesario reconocer que la variable que se está analizando es una variable macroeconómica, en donde la exactitud de los pronósticos debida a la magnitud de las cifras de los mismos, exige una mayor precisión. Además, con base en los resultados obtenidos, que se discuten más adelante, se puede mencionar que las Redes neuronales Artificiales son más sensibles a los cambios del comportamiento de los datos, por lo que arroja resultados más confiables.

Por otro lado, en la siguiente imagen se pueden observar los factores estacionales calculados para cada uno de los meses del año. Estos factores indican la participación de cada uno de los meses en la regularidad de cuartos de hotel ocupados, y por ende, la llegada de turistas de Estados Unidos y de Canadá.

Con base en la gráfica, se puede inferir que los meses con los factores estacionales más grandes (julio, marzo y agosto, en orden descendente) tienen dichos valores, ya que son meses en los cuales hay vacaciones, de semana santa (pascua) o de verano, y días de asueto; por lo que los turistas tienen una posibilidad más alta de viajar y elegir lugares turísticos. Por otro lado, septiembre, octubre y junio, en orden descendente, son los meses que tienen los factores estacionales más pequeños, es decir, son los meses cuya participación en llegadas de turistas y ocupación de cuartos de hotel tienen un registro mucho menor.

Ilustración 22: Factores estacionales



Elaboración propia a partir de (Team, 2017)

Es importante aclarar que se ha supuesto que el comportamiento de los factores estacionales es constante para todos los años analizados. Esto es, a partir de los datos históricos, con los cuales se calculan los factores, se espera que dicho comportamiento se conserve para los periodos siguientes que serán pronosticados.

Una vez que se logró tener la base de datos preparada, depurada y lista para manipular, se continúa con el tercer paso de la metodología de la minería de datos, pertinente al diseño del modelo, tanto en los datos de entrenamiento como en los de evaluación de desempeño. El código empleado para el diseño del modelo de redes neuronales artificiales, realizado con ayuda del software “The R project for statistical computing” (Team, 2017), se puede visualizar en la sección de anexos y código. En él se puede observar cada línea de código comentada para tener una mejor comprensión del algoritmo.

De manera general, el código comienza con la declaración de la base de datos que se va a utilizar, en donde se tienen las observaciones de las variables independientes antes descritas. Para las variables que son propias de cada país, se asignaron valores únicos. Esto es, por ejemplo, para la variable de la relación de tipo de cambio, se asignó una variable para el tipo de cambio de Canadá y otra para el tipo de cambio de Estados Unidos. Esto con la finalidad de hacer la distinción entre países, aun cuando se trate de la misma variable explicativa. A continuación se presenta el nombre de cada una de las variables ocupadas en el modelo con su correspondiente etiqueta, para que no existan problemas de congruencia con los resultados del modelo. Se realizó esta asignación de etiquetas para que la programación del algoritmo tuviera una secuencia y orden lógico en la nomenclatura de las variables.

Tabla 7: Etiqueta de las variables del modelo

Nombre de la variable	Etiqueta
Número de cuartos de hotel ocupados	CO
Tasa de inseguridad	TI
Relación del tipo de cambio de Canadá	TC_CAN
Relación del tipo de cambio de Estados Unidos	TC_USA
Población de Canadá	POB_CAN
Población de Estados Unidos	POB_USA
Precio de servicio relativo de Canadá	PS_CAN
Precio de servicio relativo de Estados Unidos	PS_USA
Ingreso de Canadá	IP_CAN
Ingreso de Estados Unidos	IP_USA
Estacionalidad en el mes de enero	ENE
Estacionalidad en el mes de febrero	FEB
Estacionalidad en el mes de marzo	MAR
Estacionalidad en el mes de abril	ABR
Estacionalidad en el mes de mayo	MAY
Estacionalidad en el mes de junio	JUN
Estacionalidad en el mes de julio	JUL
Estacionalidad en el mes de agosto	AGO
Estacionalidad en el mes de septiembre	SEP
Estacionalidad en el mes de octubre	OCT
Estacionalidad en el mes de noviembre	NOV

Elaboración propia a partir de (Team, 2017)

Como se puede observar en la definición de las variables, las variables ficticias o dummies que se consideran para contener la estacionalidad de los datos, solo consideran los primeros once meses del año civil. Esto se debe a que el doceavo mes, diciembre, absorbe el valor de la ordenada al origen y evita que

se presente una multicolinealidad entre las variables, es decir, una correlación entre las variables explicativas. Tanto para el modelo de regresión lineal múltiple como el de redes neuronales artificiales, la incorporación de diciembre como variable explicativa solo resulta en problemas de convergencia. Además, de acuerdo con los valores de los factores estacionales, el doceavo mes no tiene una participación tan sobresaliente para no justificar su inserción en el modelo.

En seguida, lo siguiente que se realizó en el tercer paso de la metodología de minería de datos fue la transformación de los datos de la base para que estuvieran escalados y dentro del rango [0,1]. Esto con la finalidad, tal como se explicó en el tercer capítulo, de que los datos que tenga un valor absoluto mayor (comparado con los demás datos de la base) no tengan una preferencia en el modelo o que puedan sesgar los resultados. De manera que con una transformación de escala, los datos pueden ser comparados sin ningún problema. De las opciones de expresiones de estandarización que se presentaron, se optó por emplear la técnica de estandarización lineal de Weitenorf, la cual relaciona los valores máximos y mínimos con el dato objetivo de la siguiente manera:

$$A_i = \frac{A_{0i} - \min A_{0i}}{\max A_{0i} - \min A_{0i}}$$

En donde el dato objetivo es centrado con el valor mínimo de la variable y se divide entre el rango. Se ha escogido esta técnica de estandarización por la simplicidad de su aplicación, ya que como se explicó en el tercer capítulo, de acuerdo con (Anysz, Zbiciak, & Ibadov, 2016), se sabe que no hay una mejor forma de estandarizar los valores, debido a que la arquitectura de la red influye mucho en el cálculo de los errores.

De manera que lo primero fue obtener el valor máximo y mínimo de cada una de las variables. Posteriormente, centrar cada uno de los datos con el dato mínimo y obtener así los datos escalados en un rango [0,1]. En seguida, dado que dentro del diseño del algoritmo se encuentran incluidas la parte de prueba y verificación, se tuvo que obtener una muestra aleatoria de los datos, de manera que se formaran dos conjuntos: por un lado el conjunto de datos que se utilizará para la prueba o ajuste y el segundo para la verificación y obtención de los errores. Para tal finalidad se utilizó una función del software con la que se pueden generar números aleatorios para determinar los datos que serán parte de cada uno de los conjuntos.

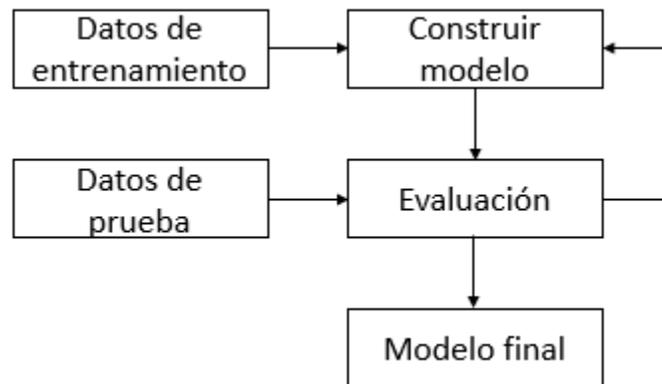
Lo elementos que están contenidos en cada uno de los conjuntos son seleccionados aleatoriamente, debido a que de esa manera, se reduce la probabilidad de que en el modelo exista un sobreajuste, es decir, que existan más variables que observaciones y que el modelo aprenda solamente el comportamiento de ciertos datos. Cuando se le presentan datos aleatorios, el comportamiento cambia y el modelo se tiene que esforzar por encontrar el mejor ajuste.

Para la aplicación de este trabajo de investigación, se ha seleccionado que el 80% del total de los datos (240 de 300) sean parte del conjunto de prueba y que los 60 datos restantes sean parte del conjunto de verificación y cálculo de los errores, esto para darle la oportunidad al algoritmo de tener la mayoría de los datos para aprender su comportamiento. De acuerdo con la metodología de las redes neuronales artificiales, mientras mayor número de observaciones se tengan, el modelo tendrá una mayor oportunidad de adecuarse al comportamiento de los datos. Por desgracia, debido a la disponibilidad de

las trescientas observaciones ha provocado que el porcentaje de dichas observaciones se vea desproporcional.

De manera gráfica, el proceso de validación del modelo de redes neuronales artificiales se define de acuerdo con la información del siguiente diagrama, propuesto por (Kotu & Deshpande, 2015).

Ilustración 23: Proceso de validación del modelo



Traducción propia con información de (Kotu & Deshpande, 2015)

Se puede apreciar, que es un proceso recursivo con el cual se pretende generar el mejor modelo de ajuste partiendo de los dos conjuntos de datos.

Luego, cuando los conjuntos de datos que se utilizarían para cada uno de los procesos estaban definidos, se continuó con la declaración de la función que se utilizará. Como se mencionó en el capítulo tres, el algoritmo de redes neuronales artificiales se puede considerar como un modelo de regresión no paramétrico no lineal, de manera que la ecuación que relaciona las variables independientes o entradas con la variable dependiente o salida es la siguiente:

$$CO \sim TI + TC_CAN + TC_USA + POB_CAN + POB_USA + PS_CAN + PS_USA + IP_CAN + IP_USA + ENE + FEB + MAR + ABR + MAY + JUN + JUL + AGO + SEP + OCT + NOV$$

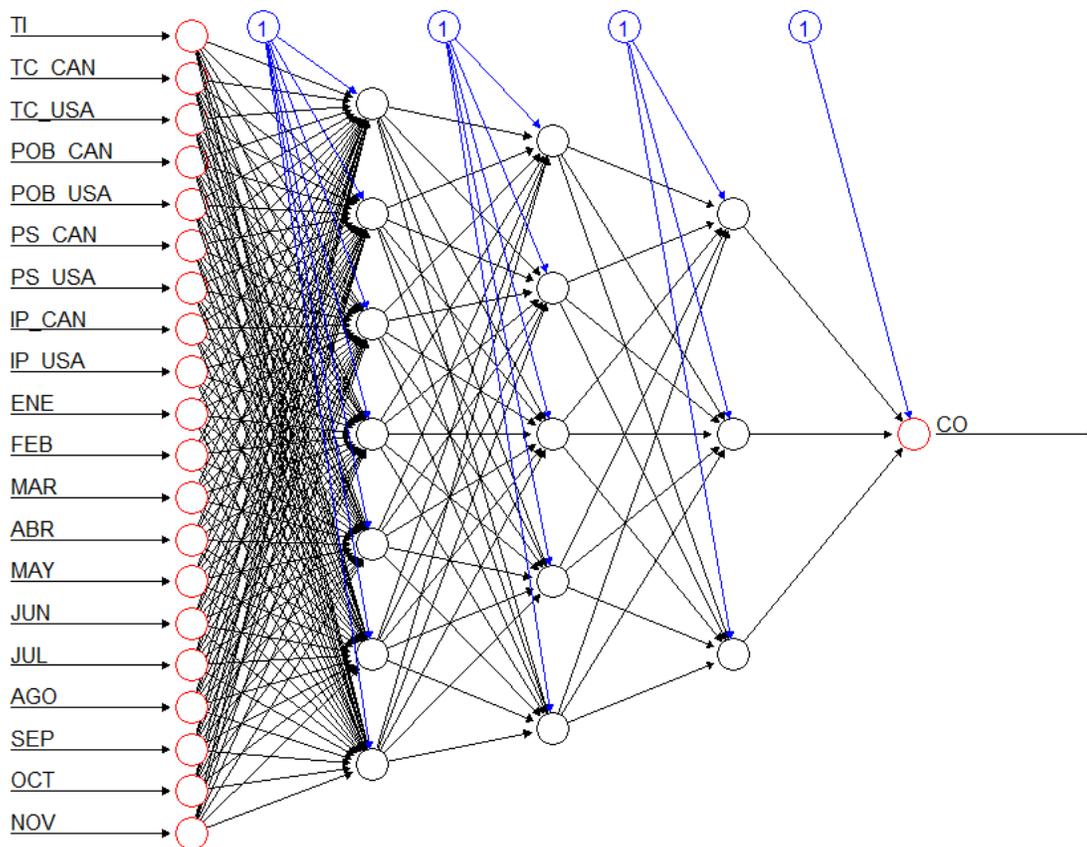
Posteriormente, se continuó con el diseño de la función de redes neuronales artificiales. Para ello, se recurrió a la paquetería “neuralnet” del software “The R Project for statistical computing” y la función “neuralnet” (Fritsch & Guenther, 2016). Con ellos, es posible crear una red neuronal artificial con los parámetros y especificaciones que se requieran. En la parte de anexos y código se puede observar, de manera más detallada, la función empleada, así como sus atributos. Para los datos que se analizan en este trabajo de investigación fue necesario considerar tres capas ocultas con 7, 5 y 3 neuronas cada una. La asignación del número de capas ocultas, así como del número de neuronas por capa se realizó de manera arbitraria a prueba y error, en donde se eligió una combinación de capas y neuronas que reducía el error de ajuste de los datos. En la literatura se menciona que no existe, hasta ahora, un procedimiento o metodología exacta que determine el número preciso de capas ocultas y número de neuronas por capa,

ya que la arquitectura de una red neuronal artificial dependerá, en gran medida, de la aplicación que se le dé, así como de las características de los datos.

Es necesario mencionar que, dentro de las características de la función “neuralnet”, se tiene que la función de activación que se emplea es la función “Logistic” o “Sigmoid”. Además, el algoritmo que utiliza para el aprendizaje supervisado es “backpropagation”, en donde los errores son propagados hacia atrás para mejorar los ajustes.

A continuación se presenta la arquitectura de la red neuronal artificial que se utiliza para la validación de la hipótesis de este trabajo de investigación. En primer lugar se pueden observar 20 flechas que apuntan a 20 círculos de color rojo, esas son las entradas del sistema, con las etiquetas que corresponden a cada una de las variables explicativas. Enseguida se puede apreciar que cada una de las neuronas de color rojo, que representan las entradas, se relacionan con cada una de las neuronas de la primera capa oculta. Esto es así, debido a que es una red multicapa con conexiones hacia adelante o feedforward.

Ilustración 24: Arquitectura del modelo de la red neuronal artificial



Elaboración propia a partir de (Team, 2017)

Luego, las salidas de cada una de las neuronas de la primera capa oculta, son las entradas de las neuronas de la segunda capa oculta. De igual forma, las salidas de las cinco neuronas de la segunda capa oculta son las entradas de las tres neuronas de la última capa oculta. Por último, todas las relaciones que se tuvieron

a partir de las veinte entradas, concluyen en la salida general de la red que es la neurona de salida con la etiqueta del número de cuartos de hotel ocupados (CO). Se puede observar, también, que hay 4 círculos de color azul que se relacionan con las salidas de las capas. Esa es la representación de los biases o las ordenadas al origen, si se desea hacer un símil con el modelo de regresión convencional. Esos valores, tal como se puede corroborar en el tercer capítulo, son un valor de ajuste para la suma ponderada de los pesos de las variables.

Por cuestiones de estética, la gráfica que representa la arquitectura de la red neuronal artificial se expone sin los pesos asociados en cada una de las interacciones con las neuronas. Sin embargo en la sección de anexos se pueden observar algunos de los valores de los pesos asociados a cada una de las interacciones. Se puede hacer una comparación con el modelo de regresión múltiple, diciendo que los pesos de las interacciones son un símil con el valor de los coeficientes de los parámetros de regresión. Sin embargo, los pesos no tienen una explicación concreta o una interpretación específica. Es por ello que el algoritmo de redes neuronales se puede asimilar como un sistema de caja negra, en donde las variables explicativas entran al sistema y al final resulta una salida o respuesta.

Hasta este momento, todas las interacciones que se registraron como producto del aprendizaje de la red neuronal artificial, están guardadas en una variable. Dichos valores serán utilizados para la segunda parte del diseño del modelo, la cual consiste en evaluar el desempeño del algoritmo mediante la comparación de los datos pronosticados con el conjunto de datos de verificación, y con los datos reales que se tienen en la base de datos. Las interacciones que el modelo logró aprender de las características de la base de datos, son las que permiten ajustar el modelo lo mejor posible.

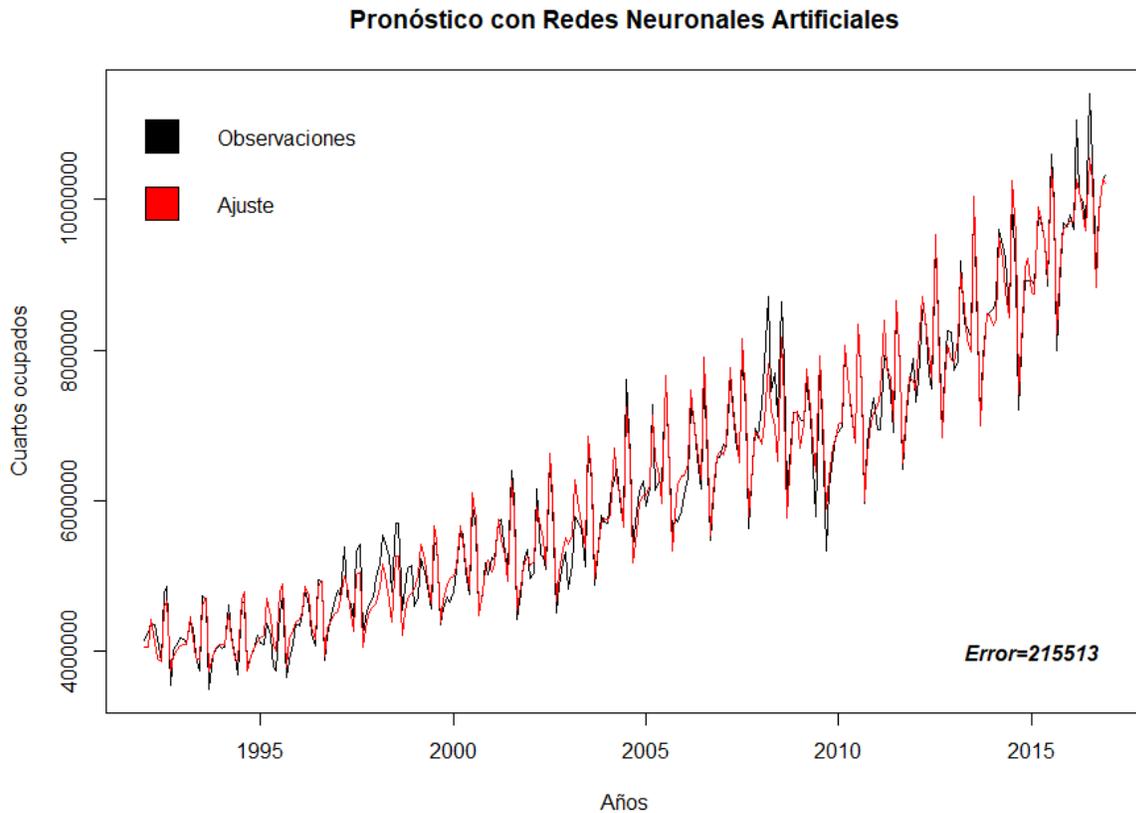
En la siguiente gráfica se pueden observar los resultados de la verificación del modelo. Se grafica la serie original del número de cuartos de hotel ocupados en la República Mexicana de color negro, así como el ajuste o pronóstico que el algoritmo de redes neuronales arrojó, de color rojo. Bajo una inspección rápida, se puede inferir que los patrones, tanto estacionales como de tendencia, son muy similares entre las dos series. Sin embargo, el punto clave para determinar si el ajuste de este modelo es satisfactorio o no, es la medida de error. De acuerdo con la raíz cuadrada del error cuadrático medio con k grados de libertad, en este caso 20, se encontró que el algoritmo de redes neuronales artificiales tiene un error de 215513 cuartos. Tal como se mencionó en el segundo capítulo, esta medida de error, con la inclusión de los grados de libertad de los métodos seleccionados, permite que los errores de los métodos sean comparables.

El valor del error anterior puede parecer demasiado grande, lo cual es cierto. Pero es debido a la naturaleza de la medida del error, ya que las diferencias entre los errores se elevan al cuadrado, lo cual penaliza enormemente cuando el error varía ciertas unidades. Además, su comparación con los datos originales corresponde con una fracción del 3.4%.

Hasta ahora, la metodología del diseño del modelo de redes neuronales artificiales se cumple al pie de la letra; sin embargo, es necesario hacer la aclaración que para la obtención del error debido al algoritmo se realizó una simulación de los errores, con el objetivo de encontrar un valor mínimo de estos. Esto es posible, ya que la asignación de los datos en los conjuntos de prueba y verificación son variables aleatorias.

Con la simulación se busca encontrar la combinación de los valores en los conjuntos que logre minimizar el error. Dicha simulación permite obtener un valor mínimo aceptable, pero no se sabe con certeza si se trata del valor óptimo. Será tema de investigaciones futuras el proponer la solución a la optimización del error mínimo.

Ilustración 25: Pronóstico con Redes Neuronales Artificiales



Elaboración propia a partir de (Team, 2017)

El cuarto paso de la metodología de minería de datos que se aplica al diseño del algoritmo de redes neuronales artificiales corresponde a la aplicación o desarrollo del modelo. En esta etapa se pretende poner en acción el algoritmo entrenado con valores reales, es decir, observaciones diferentes de los datos utilizados para el aprendizaje y verificación del modelo. Para ello, es necesario explicar que para hacer los pronósticos de los siguientes periodos, las entradas al modelo deberán ser los pronósticos de las variables explicativas, ya que por razones triviales no se cuenta con dicha información. De manera que la predicción de los periodos con el algoritmo de redes neuronales artificiales es un pronóstico alimentado de pronósticos.

Para determinar los valores de las variables que serán las entradas para el modelo, se utilizaron las técnicas de series de tiempo de Holt-Winters multiplicativo y los modelos auto-regresivos integrados de

medias móviles con estacionalidad (SARIMA). Se hizo una comparación del desempeño de cada uno de los métodos y se eligió al mejor. En primer lugar se revisaron los coeficientes de suavizado que el modelo de Holt-Winters multiplicativo arrojó, asimismo se analizaron los parámetros del modelo SARIMA. La tabla con los resultados de los coeficientes y los parámetros del modelo SARIMA se encuentran en la sección de anexos y código.

Debido a que la revisión de los coeficientes y de los parámetros resultó inconclusa en la mayoría de las variables, debido a que los coeficientes de suavizado eran muy cercanos a 1, o que los términos autorregresivos y de media móvil tomaban el valor de cero simultáneamente, se optó por un tercer criterio de selección. El pronóstico de cada una de las variables explicativas se realizó con base en la comparación de los errores de ambos métodos. En la siguiente tabla se puede apreciar el concentrado de los errores que resultaron de la ejecución de cada uno de los métodos para predecir el valor de cada una de las variables de entrada. En la cuarta columna se encuentra el criterio que se ha elegido para pronosticar la variable en cuestión, debido a que dicho método ofrece el mejor ajuste con el menor error. Por ejemplo, para la variable de la tasa de inseguridad se utilizó el modelo de Holt-Winters multiplicativo (HW), pero para la variable de la relación del tipo de cambio de Estados Unidos se empleó el método SARIMA. En la sección de anexos se pueden observar los pronósticos de las variables para el año 2019.

Tabla 8: Error de los modelos HW y SARIMA

Error de los modelos			
	Holt.Winters	SARIMA	Mejor Ajuste
TI	312.4150964	327.7515898	HW
TC_CAN	0.265266019	0.265826243	HW
TC_USA	0.366853132	0.362861467	SARIMA
Pob_CAN	4494.630647	2567.650625	SARIMA
Pob_USA	38464.79947	193452.8036	HW
PS_CAN	0.004306576	0.003737289	SARIMA
PS_USA	0.003831679	0.003366621	SARIMA
IP_CAN	17.25942985	18.62461387	HW
IP_USA	16.81569506	16.32305646	SARIMA

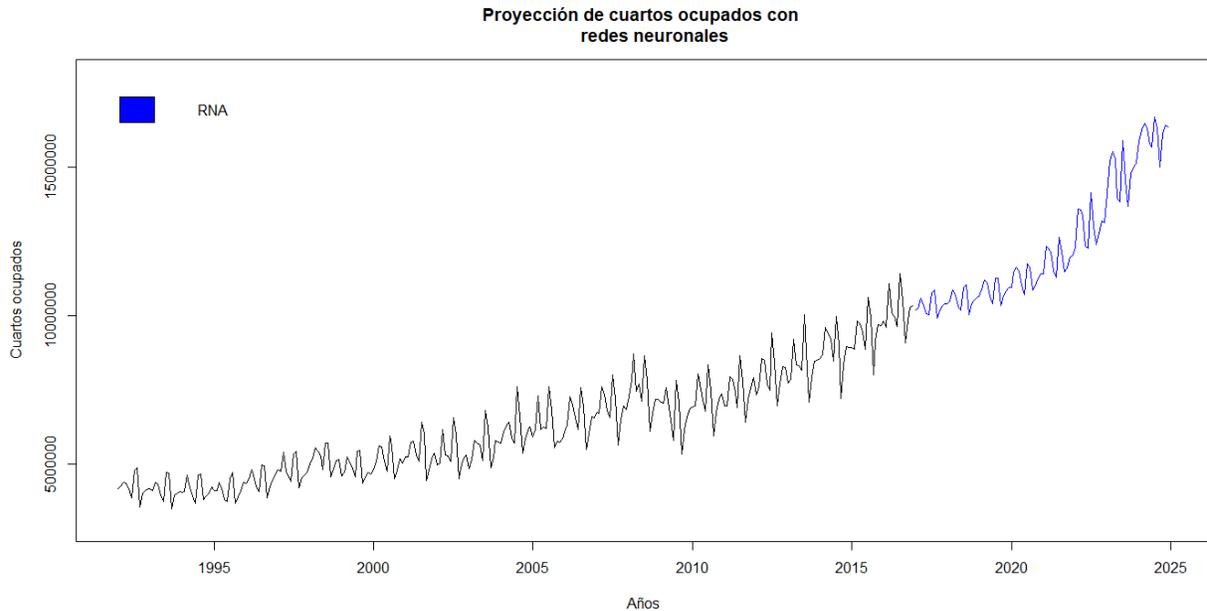
Elaboración propia a partir de (Team, 2017)

De manera que, una vez que se logró conformar la base de datos con los valores respectivos a los pronósticos de las variables que funcionan como las entradas del algoritmo de redes neuronales, se procedió a realizar las predicciones. Para ello, igual que con los datos de prueba y verificación, se comenzó con la estandarización de los datos. Luego, con los resultados de los pesos del proceso de aprendizaje, que estaban almacenadas en una variable dentro del código, se ponderan los valores de las entradas y se obtiene el resultado final. La siguiente gráfica muestra, en color azul, las predicciones realizadas con el algoritmo de redes neuronales artificiales.

Como parte de la presentación gráfica de los datos pronosticados con el algoritmo de redes neuronales artificiales, es necesario aclarar que el algoritmo necesita un inicio o referencia con la cual se pueda comenzar con las predicciones. Dado que no se especificó ningún punto de referencia al comienzo del

diseño del algoritmo, es pertinente que cuando se desee realizar cualquier pronóstico, se indique el inicio de dichos pronósticos.

Ilustración 26: Proyección de cuartos con RNA



Elaboración propia a partir de (Team, 2017)

Para este caso, se ha tomado como inicio el primero de enero de 2017 y tomando como referencia el último dato de la serie temporal original, es decir, la observación número 300 que corresponde al 1 de diciembre de 2016. Es importante tener en consideración que también se pudo haber elegido la diferencia entre el último pronóstico y el primer pronóstico del algoritmo para fijar el inicio de las predicciones, ya que no hay un solo punto que pueda ser considerado como el óptimo. Este hallazgo también será sujeto de análisis en investigaciones futuras, en donde se pretenda determinar el punto de inicio de los periodos pronosticados más adecuado.

Por último, la quinta parte de la metodología de minería de datos se refiere a las acciones posteriores al modelo, con el cual se pretende desplegar una generación de conocimiento. En este sentido, dicha generación de conocimiento se atribuye a la comparación de las técnicas de pronóstico convencionales, series de tiempo (Holt-Winters y SARIMA) y modelos de regresión, con el algoritmo de redes neuronales artificiales, para analizar su desempeño. Como primer análisis, se presenta el modelo y pronósticos realizados con el método Holt-Winters multiplicativo. Se encontró que las constantes de suavizado tienen los siguientes valores:

Tabla 9: Constantes de suavizado

Constante de suavizado	Valor
<i>Alfa</i> (α)	0.4577648744
<i>Beta</i> (β)	0.01416656154
<i>Gamma</i> (γ)	0.5228548352

Elaboración propia a partir de (Team, 2017)

A partir de las constantes de suavizado se puede inferir que el nivel (α) y el patrón estacional (γ) son causados, prácticamente, por la mitad de los datos más actuales y la mitad por los datos históricos. En cambio, el parón tendencial (β) se explica mejor con base en los datos históricos de la serie temporal, es decir, dado que el coeficiente es muy cercano a cero, su comportamiento solo se ve afectado por el último valor de tendencia calculado. Los coeficientes del modelo se presentan a continuación:

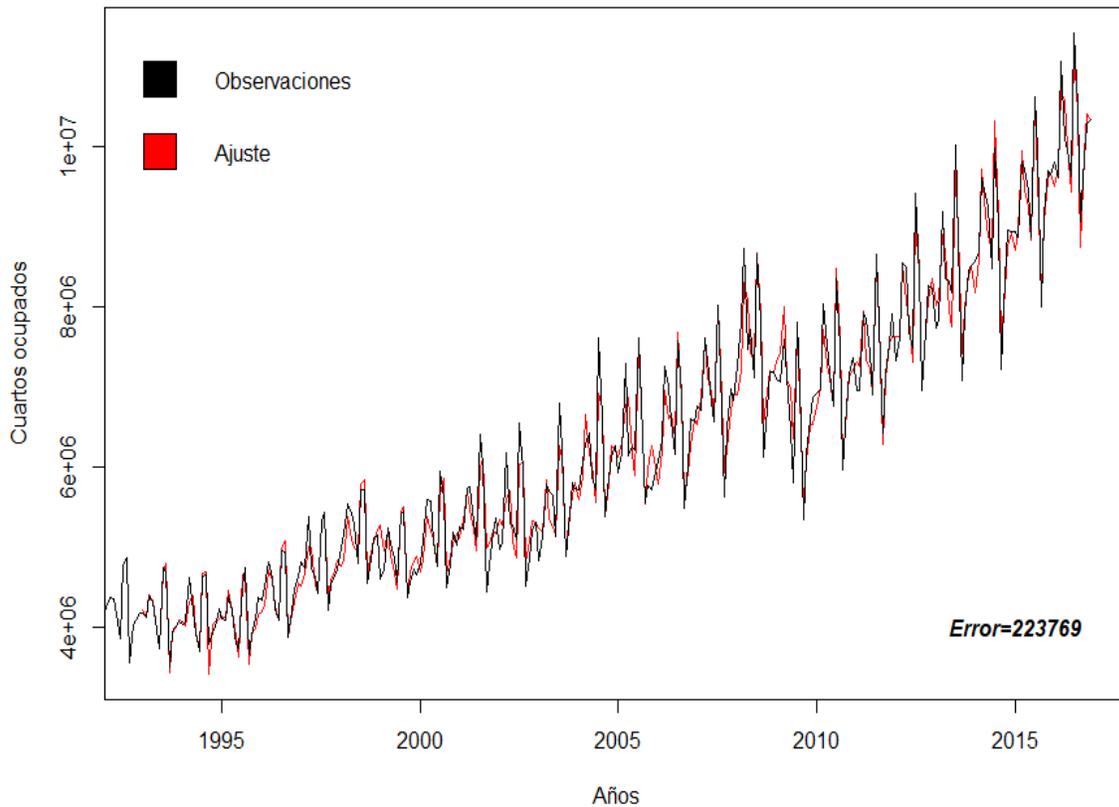
Tabla 10: Coeficientes del modelo HW

Coeficiente	Valor
Nivel (α)	10235708.417411
Tendencia (β)	33500.149609
S1	4844.124908
S2	-71632.849114
S3	1052216.674662
S4	488192.033012
S5	225949.280572
S6	-289418.932276
S7	1381645.339635
S8	435756.984261
S9	-1252454.588222
S10	-365581.996149
S11	100168.373297
S12	94791.960203

Elaboración propia a partir de (Team, 2017)

Asimismo, a continuación se presenta la gráfica que representa el ajuste que el método de Holt-Winters desempeñó. Se observa que a pesar de tener un comportamiento muy similar con la serie temporal original, el error registrado es de 223, 769 cuartos.

Pronóstico con método Holt-Winters



Elaboración propia a partir de (Team, 2017)

El segundo método que se analiza y que participa en el proceso de comparación de técnicas de pronóstico es el modelo auto-regresivo integral de medias móviles con estacionalidad. Los parámetros que el software arrojó, en función de las características de los datos, son $ARIMA(0,1,1)(0,1,1)$ [12] en donde se puede apreciar que el modelo se explica únicamente por la autocorrelación parcial de los errores, tanto de la serie original como de la estacionalidad, es decir, se basa en el componente de medias móviles exclusivamente. Además, para que la serie temporal sea estacionaria, se debe derivar solo una vez. Los coeficientes que corresponden al modelo son los siguientes, en donde $ma1$ corresponde al coeficiente del componente de media móvil y el factor $sma1$ al componente estacional.

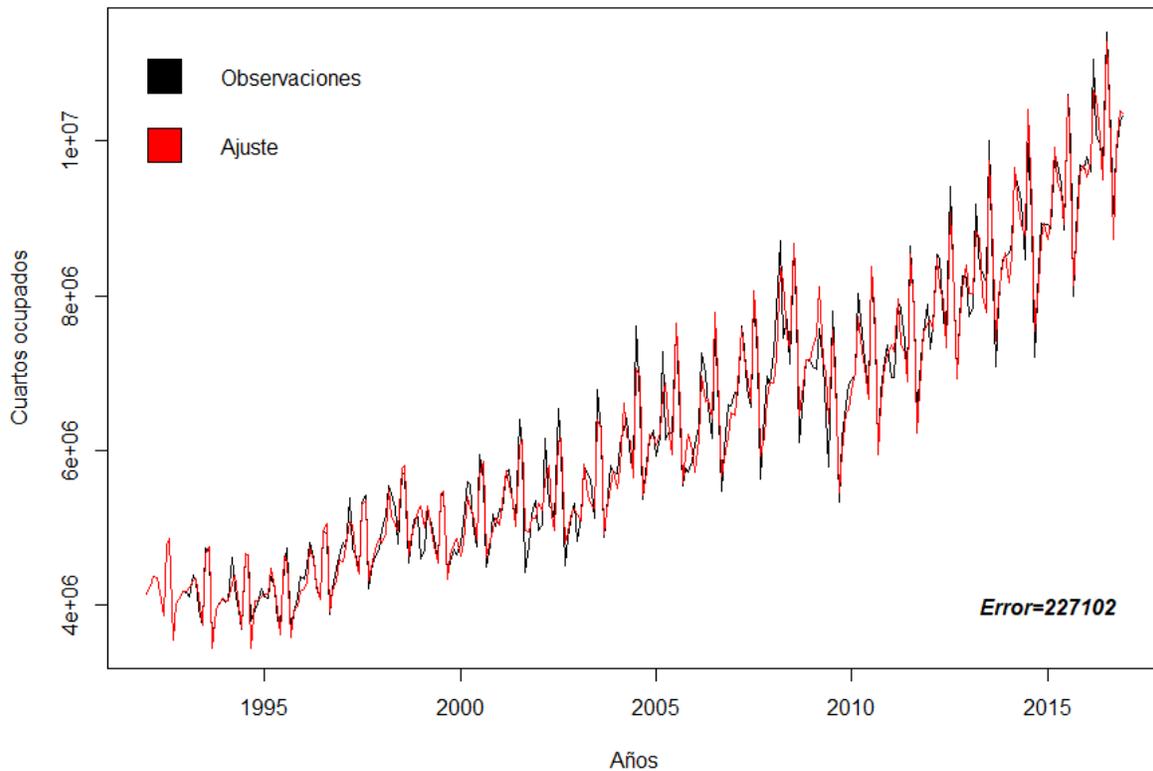
Tabla 11: Coeficientes de método SARIMA

Coeficientes	Valores
$ma1$	-0.4822438
$sma1$	-0.6483724

Elaboración propia a partir de (Team, 2017)

En la siguiente gráfica se plasman tanto las observaciones de los datos reales, como los dato ajustados por el modelo. Se puede observar como el ajuste del modelo se acopla con los patrones de tendencial y de estacionalidad de los datos; sin embargo, se encontró que el error del modelo es de 227102 cuartos.

Pronóstico con SARIMA



Elaboración propia a partir de (Team, 2017)

Para finalizar, el último modelo que fue comparado con el desempeño de las redes neuronales artificiales es el Modelo de Regresión Lineal Múltiple. Tal como se puede observar en la gráfica de los datos originales del número de cuartos de hotel ocupados en la República Mexicana, se pueden apreciar los patrones estacionales y de tendencia. Es por ello que, para poder implementar la técnica de regresión lineal múltiple con mínimos cuadrados, se debe transformar la serie no estacionaria en una serie estacionaria para asegurar que su media sea constante y no exista heterocedasticidad, es decir, una varianza que cambia con el tiempo.

Tal es la razón que el primer paso para la implementación del modelo de regresión múltiple es la transformación de la serie temporal con la función logaritmo natural. Se puede comprobar en el código que se elaboró para esta sección, que se tuvo que realizar una regresión por etapas para obtener el modelo que contiene las variables significativas. Asimismo, se aplicó la prueba de normalidad de Jarque-Bera a los residuos de los datos para verificar que los errores se comportan de manera normal, lo cual es una condición necesaria para la aplicación de la técnica de regresión lineal múltiple.

La hipótesis nula H_0 indica que los errores se distribuyen de manera normal, mientras que la hipótesis alternativa H_a expresa lo contrario, es decir, que los errores del modelo no se distribuyen de manera normal. En este sentido, el estadístico de la prueba de normalidad que se emplea es el de Jarque-Bera

(JB), el cual es un estadístico que se comporta asintóticamente como una Chi-cuadrada (χ^2) con dos grados de libertad.

Para este caso, el valor de χ^2 de la prueba JB es de 1.974, el cual tiene una probabilidad acumulada desde infinito al valor del estadístico de $p=0.3726974$, que comparado con el valor de alfa, $\alpha = 0.05$, el valor P es mayor que alfa; por lo que existe suficiente evidencia para no rechazar la hipótesis nula de normalidad en los errores.

El modelo, que se presenta a continuación, se concluyó en una segunda etapa de la regresión a pasos:

$$\log(CO) \sim \text{Fecha} + POB_USA + PS_CAN + PS_USA + IP_CAN + IP_USA + MAR + ABR + JUN + JUL + AGO + SEP + OCT$$

Asimismo en la siguiente ecuación se muestra el modelo de regresión lineal múltiple con los coeficientes (β) de regresión calculados.

$$\log(CO) = 23.263 + 4.623 \times 10^{-9} * \text{Fecha} - 1.421 \times 10^{-7} * POB_USA + 2.152 * PS_CAN - 1.884 * PS_USA - 2.969 \times 10^{-4} * IP_CAN + 5.351 \times 10^{-4} * IP_USA + 9.6 \times 10^{-2} * MAR + 6.386 \times 10^{-2} * ABR - 6.326 \times 10^{-2} * JUN + 1.719 \times 10^{-1} * JUL + 9.475 \times 10^{-2} * AGO - 1.555 \times 10^{-1} * SEP - 4.359 \times 10^{-2} * OCT$$

Este modelo tiene un coeficiente de determinación ajustado de 0.9787142, lo cual indica que las variables independientes pueden explicar el 97.87% del comportamiento del logaritmo natural del número de cuartos de hotel ocupados y que el resto se debe a factores aleatorios o eventos fortuitos que no pueden ser calculados. Enseguida se pueden observar los valores de los parámetros de regresión de manera aislada:

Tabla 12: Coeficientes del modelo de regresión lineal múltiple

Variable	Valor
(Intercept)	23.2631483
Fecha	4.6234E-09
POB_USA	-1.4214E-07
PS_CAN	2.15195099
PS_USA	-1.88365791
IP_CAN	-0.00029688
IP_USA	0.00053512
MAR	0.09600114
ABR	0.06385972
JUN	-0.06325613
JUL	0.17193364
AGO	0.09474738
SEP	-0.15545609
OCT	-0.04359398

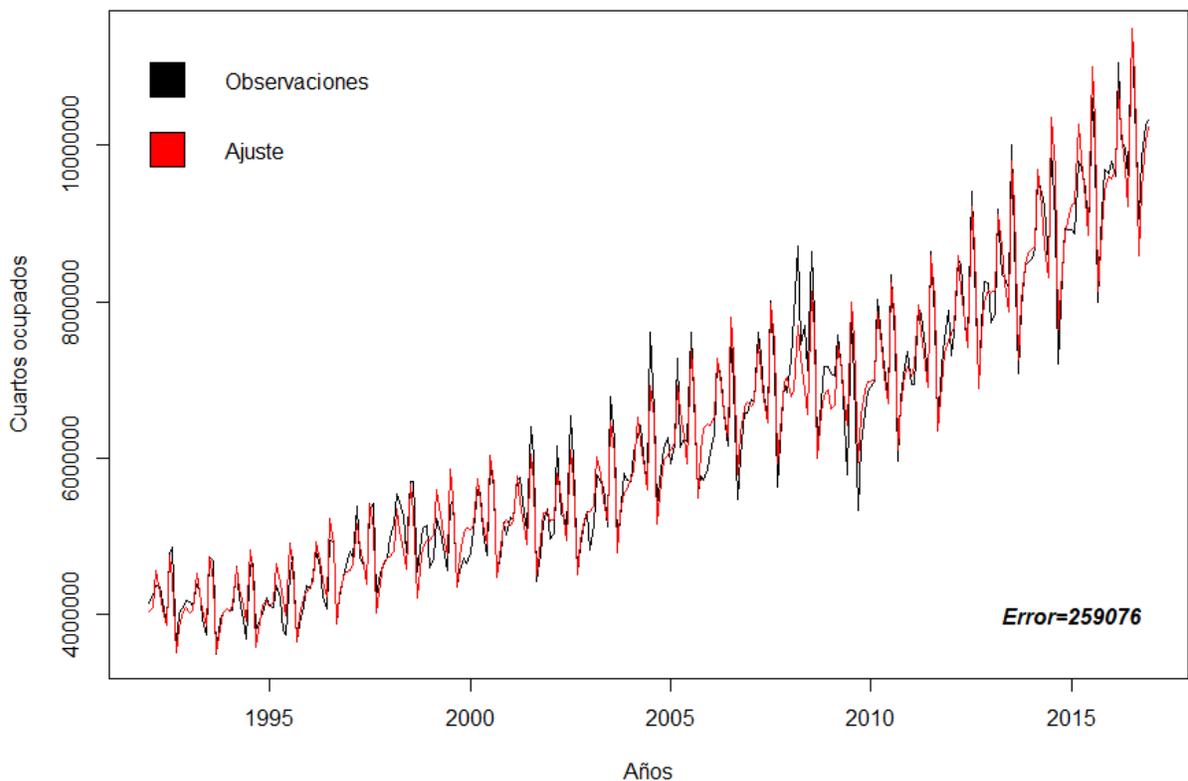
Elaboración propia a partir de (Team, 2017)

Una vez que se calcularon los parámetros de regresión, es posible realizar el ajuste de los datos y la determinación del error. Es importante recordar que, dado que la serie de datos se transformó con la función logaritmo natural, después de realizar el pronóstico con la regresión se le tienen que regresar los atributos de tendencia y estacionalidad. Para ello, al resultado se le tiene que aplicar la función exponencial, dado que es la función inversa del logaritmo natural. Bajo esta lógica, la siguiente gráfica presenta los valores de la serie temporal original en color negro y los datos ajustados por la regresión lineal múltiple en color rojo. Se puede observar, a simple vista, que los patrones de tendencia y estacionalidad son muy similares. Sin embargo el error que presenta dicho ajuste con mínimos cuadrados es de 259076 cuartos.

El pronóstico de los datos correspondientes a los siguientes periodos de interés, se realiza de manera similar a la del método de redes neuronales artificiales. Ya que se deberán emplear los pronósticos de las variables obtenidos anteriormente con los métodos de Holt-Winters y SARIMA, para obtener el resultado final de cada año.

Ilustración 29: Pronóstico con Regresión Lineal Múltiple

Pronóstico con Regresión Lineal Múltiple



Elaboración propia a partir de (Team, 2017)

Como resumen de la quinta parte de la metodología de minería de datos, aplicado al algoritmo de redes neuronales artificiales, se presenta el siguiente cuadro con la comparación de los errores en cada uno de los métodos analizados. Se puede observar que el valor de error, calculado como la raíz de la suma de los errores cuadráticos en cada una de las técnicas, es diferente. En la tercera columna se simboliza con *** a la técnica de pronóstico que tiene el menor error de ajuste de los datos.

Tabla 13: Comparación de errores

Comparación de los errores			
	Error	Mínimo	Porcentaje base RNA
Holt-Winters	223769	-	3.83%
SARIMA	227102	-	5.38%
MRLM	259076	-	20.21%
RNA	215513	***	-

Elaboración propia a partir de (Team, 2017)

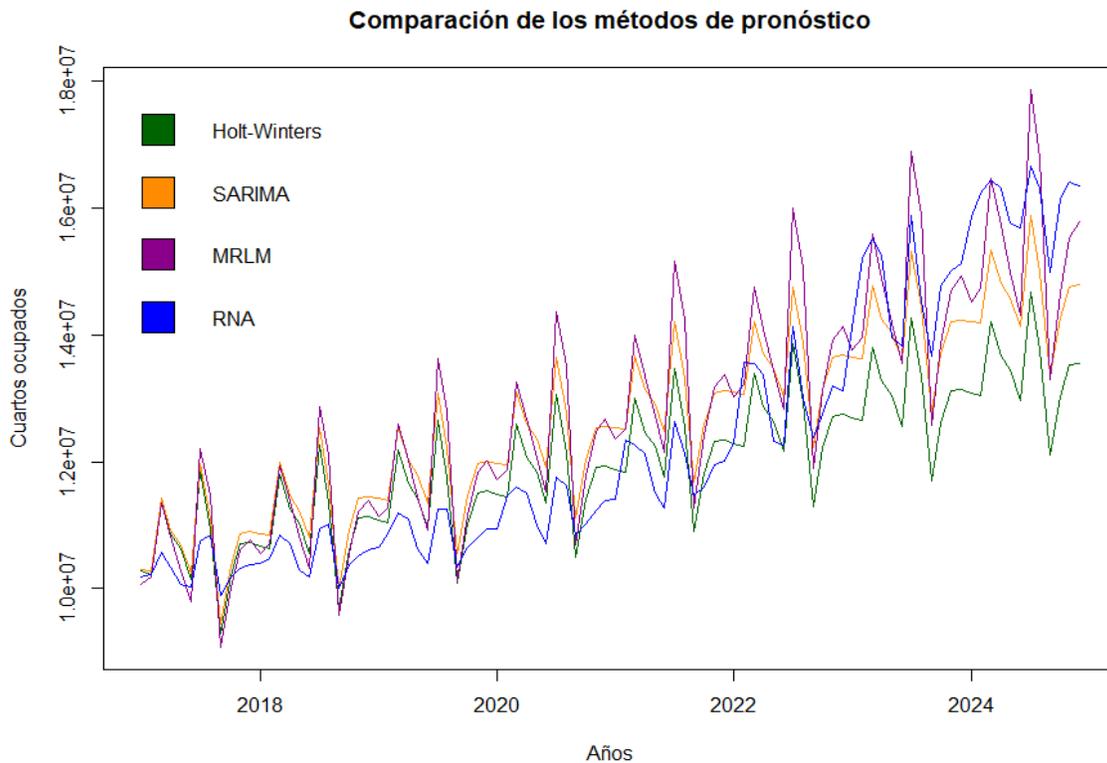
De acuerdo con el cuadro de comparación de los errores, el método de pronóstico de redes neuronales artificiales resulta tener el menor valor, es decir, tiene el mejor ajuste para los datos. En segundo lugar se encuentra el método de serie de tiempo de Holt-Winters multiplicativo, seguido por el método SARIMA y por último el modelo de regresión lineal múltiple. En la cuarta columna se puede observar el porcentaje de variación que tiene cada uno de los métodos tomando como referencia al valor de error que se obtuvo con el método de RNA. Se observa que para el método de Holt-Winters solo hay una variación del 3.83%, lo cual puede parecer muy poco e incluso despreciable, por lo que un analista podría ser indiferente en emplear uno u otro algoritmo. Sin embargo, es importante mencionar que para la estimación de variables macroeconómicas, el tamaño de las cifras si resulta representativo. Además de que el resultado obtenido con las Redes Neuronales está condicionado a un sistema de paro en la simulación de sus errores, tal como se menciona más adelante, por lo que es posible encontrar un valor de error mucho más pequeño, capaz de incrementar la brecha entre las dos técnicas.

Las siguientes gráficas muestran el desempeño de los métodos empleados, comparándolos de manera gráfica. En la primera gráfica se observan los pronósticos realizados de los periodos siguientes, que comprenden del primero de enero de 2017 al primero de diciembre de 2024. Como se puede inferir, se realizaron las predicciones para 8 años, correspondientes a 96 periodos mensuales, dado que uno de los objetivos de este trabajo de investigación es proporcionar la información pronosticada para todo un sexenio presidencial. Es por ello que se vio en la necesidad de incluir los dos últimos años del periodo presidencial pasado para disponer de la totalidad de observaciones requeridas. De color verde oscuro se muestra la trayectoria de datos que se generó mediante el método de Holt-Winters multiplicativo, de color naranja oscuro se presentan los datos pronosticados de acuerdo con el método SARIMA, en color morado se pueden apreciar los resultados del modelo de regresión lineal múltiple y, por último de color azul, los pronósticos realizados con el método de redes neuronales artificiales.

Los resultados que se presentan en la gráfica corresponden a observaciones o predicciones puntuales, ya que de esta forma resulta más evidente la comparación entre los métodos. Es importante mencionar que

para las técnicas convencionales de pronóstico, llámese Holt-winters, modelo SARIMA y el modelo de regresión, existe la posibilidad de generar un intervalo de confianza a un nivel de significancia, con el cual se genera un rango en el que se asegura que el pronóstico este contenido. Para el caso de las redes neuronales artificiales, se puede generar un intervalo de confianza pero para el valor de los pesos, es decir, se genera un rango del cual los pesos pueden tomar cualquier valor, con lo que el resultado final del pronóstico puede cambiar. Será tema de investigaciones futuras la propuesta de un intervalo de confianza para el modelo de redes neuronales artificiales, ya que como se mencionó anteriormente, los pesos están en función de la generación aleatoria de la muestra que se divide en los conjuntos de prueba y verificación.

Ilustración 30: Comparación de los métodos de pronóstico

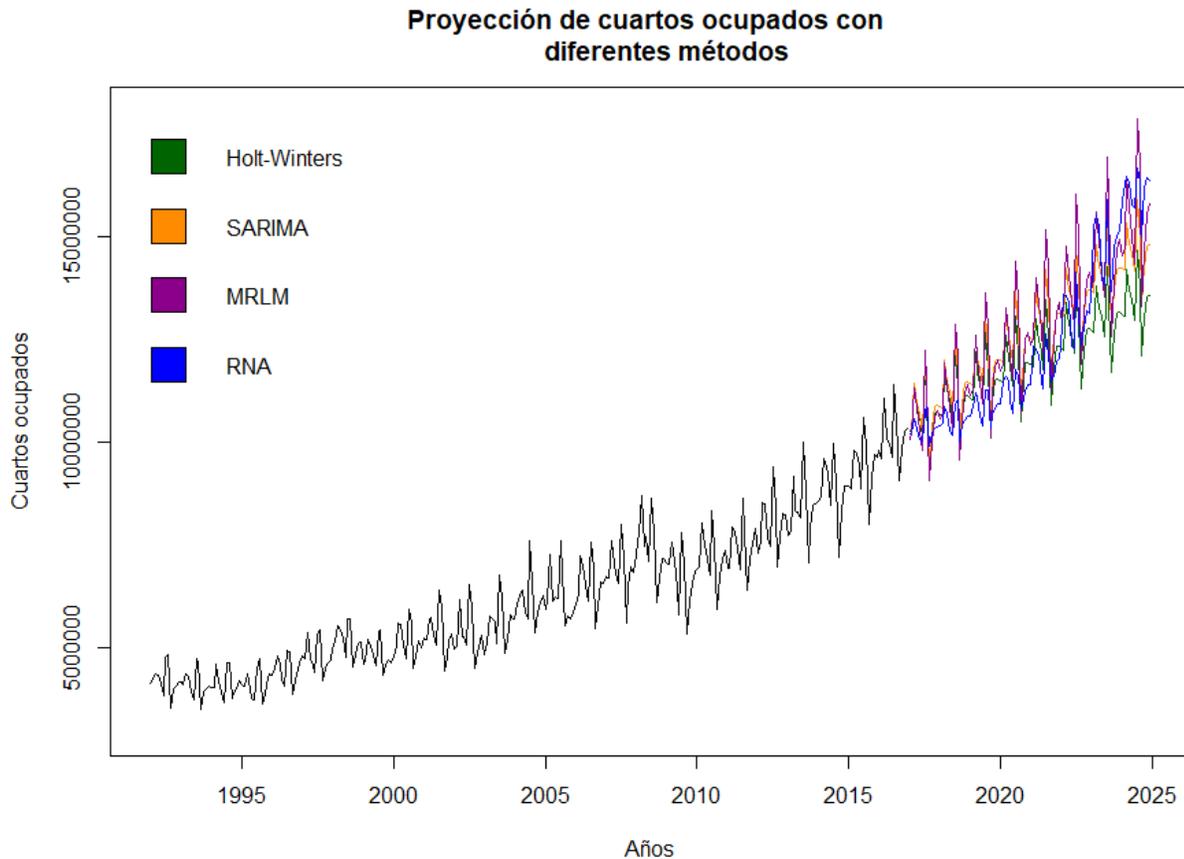


Elaboración propia a partir de (Team, 2017)

De manera general se puede observar que todos los pronósticos comienzan relativamente del mismo punto y conservan los patrones estacionales y de tendencia, cada uno de la mejor manera en que su metodología lo permite. Se puede apreciar, además, que tanto el método de Holt-Winters y SARIMA tienen un comportamiento muy similar, solo que la trayectoria del modelo Auto-regresivo se encuentra siempre por arriba del modelo de suavizado exponencial. Asimismo, la trayectoria de los pronósticos con el modelo de regresión lineal múltiple se encuentra por arriba de todos los demás pronósticos, con una inclinación a disparar su valor, es decir, presenta rasgos de volatilidad en la medida en que los periodos pronosticados se alejan del inicio. No así, el modelo de redes neuronales artificiales comienza sutilmente debajo de las demás técnicas y conforme avanza el tiempo logra cruzar las trayectorias de las otras técnicas, manteniendo los patrones estacionales y de tendencia. Es decir, se puede apreciar un comportamiento no lineal, mucho más adaptativo que las tres técnicas adicionales analizadas.

La segunda gráfica representa la proyección de los cuartos ocupados con las diferentes técnicas, tomando como punto de inicio el último dato de la serie temporal original. Se observa la trayectoria de los datos históricos de los cuartos de hotel ocupados hasta el año 2016 y, a partir de ahí, como continuaría la serie en función de cada uno de los modelos propuestos. En la sección de anexos y código se presentan las gráficas de los pronósticos comparando el modelo de redes neuronales artificiales con cada uno de los modelos adicionales, de manera individual. Asimismo, se agrega una tabla con los valores numéricos de las predicciones generadas por cada uno de los métodos.

Ilustración 31: Proyección de cuartos ocupados con diferentes métodos



Elaboración propia a partir de (Team, 2017)

Por lo tanto, dado que se demostró la superioridad de la herramienta de pronóstico del modelo de redes neuronales artificiales, se utilizarán sus resultados para concluir sobre este trabajo de investigación.

RESUMEN DEL CAPÍTULO

En este capítulo se explica detalladamente la metodología que se siguió para el diseño, construcción, validación y comparación del modelo de Redes Neuronales Artificiales con las demás técnicas de pronóstico convencionales. Se muestra el desempeño de cada uno de los modelos y las predicciones que se obtienen con la aplicación de cada uno de los algoritmos.

Conclusiones

De acuerdo con los resultados obtenidos, el objetivo general de este trabajo de investigación se satisfizo completamente. Esto es, se logró pronosticar el número de cuartos de hotel que serán demandados en el sector turístico por el mercado estadounidense y canadiense, mediante la aplicación del algoritmo de redes neuronales artificiales. Asimismo, se realizó una comparación de esta técnica de predicción con las técnicas convencionales de pronóstico, llámense: las series de tiempo como Holt-Winters y el modelo SARIMA, y el modelo de regresión lineal múltiple, para determinar su desempeño y comprobar su ventaja en este campo de conocimiento.

Se comprobó que para el pronóstico del número de cuartos de hotel ocupados en la República Mexicana, con las variables explicativas y las variables ficticias seleccionadas, el modelo de redes neuronales artificiales tiene un mejor desempeño. El valor del error calculado con la raíz cuadrada de la suma de los errores cuadráticos, resultó menor que el error de las demás técnicas evaluadas.

De manera que, de acuerdo con la primera hipótesis formulada, en la que asevera que “el algoritmo de redes neuronales artificiales tiene un mejor desempeño que las técnicas convencionales de predicción”, se puede asegurar que se cumplió cabalmente. Para este análisis es importante mencionar que las características del modelo de redes neuronales artificiales le brindan una ventaja competitiva sobre las demás técnicas de pronóstico convencionales, ya que con su diseño se es capaz de emular las relaciones de los datos y replicar con ello su comportamiento. Es por ello que se considera a la técnica de regresión no lineal no paramétrica una buena herramienta para predecir variables relacionadas con el sector turístico. Así pues, retomando la importancia de tener un buen pronóstico de las variables turísticas y una buena planeación para la asignación de los recursos, se propone la incorporación de esta técnica como opción de herramienta para calcular los pronósticos en los organismos y secretarías federales.

Si bien es cierto que la mejor herramienta de pronóstico es aquella que ofrece los mejores resultados con un modelo sencillo, siguiendo el principio de parsimonia, es necesario aclarar que en ocasiones la precisión de un resultado justifica la inversión en los recursos. De esta forma que los gastos de implementación que puedan estar asociados con la incorporación de una nueva técnica de pronóstico, tales como capacitación del personal, equipo de cómputo más potente o paquetería especial; serán subsanados con los beneficios que resultarán. Se logrará que la cantidad de pérdida por faltante (cuartos insuficientes) o que la sobreoferta de los hoteles se vea reducida a su capacidad óptima, de manera que las utilidades del país por materia del sector turístico se vean beneficiadas.

Asimismo, hay que recordar que la globalización crece a un ritmo acelerado, lo que obliga a las personas, empresas y países a tomar decisiones cada vez más rápido y con un menor margen de error. Además de que estamos inmersos en la cuarta revolución industrial, en donde conceptos como la minería de datos, machine learning y data science, tienen una participación importante. De manera que mientras más involucrados se esté con este tipo de herramientas, la transición que se presente no será tan abrupta.

Por otro lado, de acuerdo con la segunda hipótesis que se formuló, en la que se plantea que “la infraestructura hotelera y turística de la República Mexicana es suficiente para abastecer la demanda creciente de turistas de los principales países que mantienen relaciones económicas con México”, se

puede concluir que se cumplió parcialmente para el horizonte de predicción establecido, esto es, a partir del primero de enero de 2017 al primero de diciembre de 2024.

De acuerdo con las predicciones realizadas con el modelo de redes neuronales artificiales, el número de cuartos de hotel ocupados aumentarán de 10, 184,140 a 16, 352,420. Suponiendo que el porcentaje de turistas de Estados Unidos de América y de Canadá se mantiene en la misma proporción del 85.11% y que la infraestructura con la que actualmente cuenta el país, contabilizada a la última actualización de los datos de la Secretaría de turismo el pasado septiembre de 2018, es de 18, 996, 447 cuartos; la cantidad máxima de cuartos que pueden ser ocupados por estadounidenses y canadienses es de 16, 167, 733. Por lo que, analizando los datos puntuales pronosticados, alrededor del año 2023 y 2024 se espera que la infraestructura actual no sea suficiente para abastecer la demanda de estos países. De manera que sería necesario contemplar la opción de aumentar la infraestructura hotelera y de servicios para el año 2023, ya que una fracción del 85.11% de los turistas no tendrá donde hospedarse. Esto podría traer consecuencias como la incentivación de los servicios de hospedajes alternativos a los públicos federales que repercutan en la economía de la nación.

Además, con la creación de nuevas infraestructuras hoteleras y de servicios, se promueve la creación de empleos, la creación de ingresos mediante la ampliación de la cadena de suministro de este sector. Por el efecto multiplicador del sector turístico, antes mencionado, se esperaría que la infraestructura beneficiara a las comunidades cercanas, reduciendo el índice de pobreza en ese lugar y aumentando, a su vez, los servicios de seguridad y salud. Fortaleciendo, de manera natural, el orgullo nacional y cultural.

Para concluir, se mencionarán algunas propuestas de estudios posteriores que se pueden realizar para aumentar el valor de esta investigación que toma como base a las redes neuronales artificiales para resolver un problema de predicción. En primer lugar, para el diseño del modelo de las redes neuronales, resultaría interesante poder determinar una relación óptima de la cantidad de datos que se deben evaluar en el conjunto de datos para la prueba del modelo y la cantidad de datos que se deben considerar para su verificación. Es decir, encontrar la proporción óptima del tamaño de la muestra que reduce el error del algoritmo. Asimismo, determinar el número adecuado de capas y neuronas ocultas que se deben considerar. Si bien la literatura menciona que éstos dependerán del tipo de arquitectura de la red, así como de los datos que se deseen pronosticar, sería una gran aportación establecer un valor aproximado en función del problema que se presente.

Por otro lado, tal como se mencionó en el cuarto capítulo en la descripción de los resultados, dado que el algoritmo de redes neuronales artificiales comienza su ajuste a partir de la selección de una muestra aleatoria, la combinatoria de posibilidades produce diferentes valores de error. Para este trabajo de investigación se realizó una simulación en donde la condición de paro era una cota inferior para el valor de error; sin embargo, no se sabe con certeza si el valor encontrado es el error mínimo. Se propone realizar un estudio de técnicas heurísticas para la solución de este problema, con el cual se pretende aumentar la probabilidad de encontrar el error mínimo y disminuir el tiempo de convergencia del algoritmo.

Finalmente, se considera pertinente cuestionar y tomar como investigación futura el inicio o punto de referencia del pronóstico con redes neuronales artificiales, después del proceso de entrenamiento y verificación. Ya que, como todo tipo de modelo de pronóstico, se debe especificar el punto a partir del cual se debe comenzar con la trayectoria de las observaciones pronosticadas.

Bibliografía

- Anysz, H., Zbiciak, A., & Ibadov, N. (2016). The influence of input data standardization method on prediction accuracy of artificial neural networks. *Procedia Engineering*, 66-70.
- Bowerman, B., O'Connell, R., & Koehler, A. (2007). *Pronósticos, series de tiempo y regresión. Un enfoque aplicado*. México: THOMSON.
- Ciaburro, G., & Venkateswaran, B. (2017). *Neural Networks with R*. United Kingdom : Packt.
- Crotti, R., & Misrahi, T. (2018). *The Travel & Tourism Competitiveness Report 2017*. Ginebra: World Economic Forum. doi:978-1-944835-08-8
- Fritsch, S., & Guenther, F. (2016). *neuralnet: Training of Neural Networks*. Obtenido de <https://CRAN.R-project.org/package=neuralnet>
- González Alatorre, E. A., & Conde Pérez, E. M. (2011). Procedimiento para medir la demanda turística en un destino. *TURyDES*, 1-25.
- Grolemund, G., & Wickham, H. (2011). Dates and Times Made Easy with {lubridate}. *Journal of Statistical Software*, 40(3), 1-25. Obtenido de <http://www.jstatsoft.org/v40/i03/>
- Guerra Fernández, S., Lomaña Padilla, Y., Guzmán Obregón, O. A., & Pérez Arbella, Y. (2013). Optimización de la Estimación de DOA en Sistemas de Antenas Inteligentes usando criterios de Redes Neuronales. *Ingeniería Electrónica, Automática y Comunicaciones*, 70-86.
- Guevara M., G. (2018). *Economic Impact 2018 México*. Obtenido de World Travel and Tourism Council: <https://www.wttc.org/-/media/files/reports/economic-impact-research/countries-2018/mexico2018.pdf>
- Guzmán, E., & otros. (2011). Factores determinantes de la demanda internacional del turismo en México. *Globalización, Competitividad y Gobernabilidad*, 30-49.
- Hagan, M., Demuth, H., Hudson, M., & De Jesús, O. (2018). *Neural Network Design*. EBook.
- Hanke, J. E. (2010). *Pronósticos en los negocios*. México: PEARSON EDUCACIÓN.
- Haykin, S. (2009). *Neural networks and learning machines*. United States of America: PEARSON Prentice Hall.
- Hill, T., O'Connor, M., & Remus, W. (1996). Neural Networks for Time Series Forecasts. *Management Science*, 1082-1092.
- Hyndman, R. (2017). *forecast: Forecasting functions for time series and linear models*. Obtenido de R package version 8.2: <http://pkg.robjhyndman.com/forecast>
- Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1-22. Obtenido de <http://www.jstatsoft.org/article/view/v027i03>
- Kotu, V., & Deshpande, B. (2015). *Predictive Analytics and Data Mining*. United States of America: Elsevier.

- Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management* , 89-97.
- Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research* , 835-849.
- Matich, D. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Rosario, España.
- Octavio, R. C. (2008). Turismo: Factor de Desarrollo y Competitividad en México. *Centro de Estudios Sociales y de Opinión Pública*, 1-36.
- Párraga, D., Cubillos, L., & Correa-Ramirez, M. (2010). Variaciones espacio-temporales de la captura por unidad de esfuerzo en la pesquería artesanal costera del pargo rayado *Lutjanus synagris*, en el Caribe colombiano y su relación con variables ambientales. *Revista de biología marina y oceanografía*, 77-88.
- Pindyck, R., & Rubinfeld, D. (1998). *Econometric models and economic forecasts*. Singapore : McGraw-Hill.
- Secretaría de Economía. (8 de Diciembre de 2018). *Datos abiertos*. Obtenido de Información estadística de flujos de IED hacia México por país de origen desde 1999:
<https://datos.gob.mx/busca/dataset/inversion-extranjera-directa-de-se-creado-el-2015-10-27-17-37/resource/3ae82565-4bec-4303-a434-a26fdd25bd28>
- Secretaría de Turismo. (9 de Diciembre de 2018). Obtenido de Portal del Sistema Nacional de Información Estadística y Geográfica de Turismo - SNIEGT:
<http://www.datatur.sectur.gob.mx/SitePages/Inicio.aspx>
- Secretaría de Turismo. (2018). *Glosario*. Obtenido de SECTUR:
<http://www.datatur.sectur.gob.mx/SitePages/Glosario.aspx>
- Stern, H. (1996). Neural Networks in Applied Statistics. *Technometrics*, 205-214.
- Team, R. C. (2017). *R: A language and environment for statistical computing*. Obtenido de R Foundation for Statistical Computing: <https://www.R-project.org/>
- Trapletti, A., & Hornik, K. (2018). *tseries: Time Series Analysis and Computational*. Obtenido de R package version 0.10-43: <https://CRAN.R-project.org/package=tseries>
- Warner, B., & Misra, M. (1996). Understanding Neural Networks as Statistical Tools. *The American Statistician* , 284-293.
- Wickham, H., & Bryan, J. (2017). *readxl: Read Excel Files*. Obtenido de <https://CRAN.R-project.org/package=readxl>
- World Tourism Organization . (2 de Diciembre de 2018). Obtenido de <http://www2.unwto.org/>
- World Tourism Organization . (2018). *UNWTO Tourism Highlights*. Obtenido de World Tourism Organization : <https://www.e-unwto.org/doi/pdf/10.18111/9789284419876>
- World Travel & Tourism Council. (Noviembre de 2012). *The Comparative Economic Impact of Travel & Tourism*. Obtenido de World Travel & Tourism Council: <https://www.wttc.org/>

/media/files/reports/benchmark-reports/the_comparative_economic_impact_of_travel__tourism.pdf

World Travel & Tourism Council. (Septiembre de 2014). *Inversión en Viajes y Turismo en las Américas*. Obtenido de World Travel & Tourism Council: https://www.wttc.org/-/media/files/reports/policy-research/americas-investment-report_es-la.pdf

World Travel & Tourism Council. (Agosto de 2016). *Exchange rate trends and travel & tourism performance*. Obtenido de World Travel & Tourism Council: <https://www.wttc.org/-/media/files/reports/special-and-periodic-reports/exchange-rates-august-2016/exchange-rates-and-tt-performance.pdf>

World Travel & Tourism Council. (Agosto de 2016). *The impact of exchange rates on travel & tourism*. Obtenido de World Travel & Tourism Council: <https://www.wttc.org/-/media/files/reports/special-and-periodic-reports/exchange-rates-august-2016/wttc-exchange-rate-analysis-august-2016.pdf>

World Travel & Tourism Council. (2017). *Benchmark Report 2017-Mexico*. Obtenido de World Travel & Tourism Council: <https://www.wttc.org/-/media/files/reports/benchmark-reports/country-reports-2017/mexico.pdf>

World Travel & Tourism Council. (Marzo de 2017). *Understanding the Critical Issues for the Future of Travel & Tourism*. Obtenido de World Travel & Tourism Council: <https://www.wttc.org/-/media/files/reports/special-and-periodic-reports/understanding-the-critical-issues-for-the-future-of-travel-tourism-report-final.pdf>

Zhang, G., Patuwo, E., & Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* , 35-62.

Anexos y código

Ilustración 32: Índice de competitividad turística y de viajes de México

Mexico

22nd / 136

Travel & Tourism Competitiveness Index 2017 edition



Key Indicators

Sources: World Tourism Organization (UNWTO) and World Travel and Tourism Council (WTTTC)

International tourist arrivals	32,093,323	T&T industry GDP	US \$79,673.9 million
International tourism inbound receipts	US \$17,733.7 million	% of total	7.0%
Average receipts per arrival	US \$552.6	T&T industry employment	3,768,030 jobs
		% of total	7.5%

Performance Overview

Key Score Highest score



Mexico is one of the most-improved countries in the index, rising 8 positions to reach 22nd place in the global ranking as it continues to close the gap with the regional leaders. The government's prioritization of the T&T sector (30th) and the effective use of both natural (2nd) and cultural (10th) resources have paid off. To continue enhancing its competitiveness, however, Mexico should continue focusing on making the country safer (113th). While the country has significantly improved security in tourist areas, there are large discrepancies around the country, which may lead tourists not to select Mexico as a destination

due to security concerns related to crime and violence. Environmental sustainability (116th) also requires more attention, as commitment to international treaties has not kept up with progress internationally (having signed only 22 of 32 of them), and a significant share of the local fauna species continues to be threatened (133rd). The new challenge for Mexico will be to develop further, while preserving the environment in some of its most famous destinations and increasing its value proposition now that prices and wages are increasing.

Past performance

Travel & Tourism Competitiveness Edition	2015	2017
Rank	30 / 141	22 / 136
Score	4.4	4.5

Información recuperada de (Crotti & Misrahi, 2018)

Ilustración 33: Desglose de puntaje del índice de competitividad turística y de viajes de México

Enabling environment						
Country/Economy	Global rank	Business environment	Safety and security	Health and hygiene	Human resource and labour market	ICT readiness
NORTH AND CENTRAL AMERICA						
United States	6	5.4	5.2	5.7	5.5	6.0
Canada	9	5.3	6.1	5.6	5.5	5.6
Mexico	22	4.2	4.2	5.3	4.6	4.3
Panama	35	4.9	5.3	5.1	4.4	4.5
Costa Rica	38	4.5	5.4	5.0	4.9	5.2
Barbados	58	4.4	5.6	6.0	4.7	5.2
Jamaica	69	4.8	4.0	4.7	4.7	4.2
Trinidad and Tobago	73	4.4	4.1	5.1	4.5	4.8
Dominican Republic	76	4.2	4.5	4.9	4.1	3.7
Guatemala	86	4.4	3.7	4.6	4.1	3.8
Honduras	90	4.2	3.5	4.6	4.3	3.4
Nicaragua	92	3.5	5.4	4.6	4.2	3.3
El Salvador	105	4.0	3.0	4.9	4.2	3.7
North and Central America Average		4.5	4.6	5.1	4.6	4.4

Country/Economy	T&T policy and enabling conditions				Infrastructure			Natural and cultural resources	
	Prioritization of T&T	International Openness	Price Competitiveness	Environmental Sustainability	Air Transport Infrastructure	Ground and port infrastructure	Tourist Service Infrastructure	Natural Resources	Cultural Resources & Business Travel
NORTH AND CENTRAL AMERICA									
United States	5.3	4.0	4.4	3.6	6.0	4.6	6.6	4.9	4.8
Canada	4.9	3.3	4.5	4.7	6.8	4.0	6.0	4.6	4.1
Mexico	5.1	3.7	4.9	3.6	3.7	3.2	4.7	5.6	5.3
Panama	5.1	4.4	5.2	4.6	4.7	4.2	4.9	4.6	1.7
Costa Rica	5.2	4.2	4.3	4.6	3.2	2.6	5.3	5.4	1.7
Barbados	5.8	2.8	3.0	4.7	3.8	5.2	4.9	2.3	1.2
Jamaica	5.9	4.0	4.0	3.6	2.4	4.4	4.6	3.0	1.4
Trinidad and Tobago	3.7	3.5	4.8	3.8	4.1	4.2	4.1	2.7	1.3
Dominican Republic	5.8	3.3	4.4	3.7	2.8	3.5	4.6	3.2	1.4
Guatemala	4.4	3.9	5.6	3.8	1.9	2.7	3.7	3.7	1.8
Honduras	5.0	4.2	4.7	4.3	2.2	3.0	3.7	3.5	1.5
Nicaragua	4.6	4.3	5.1	4.0	2.0	2.7	3.5	3.3	1.4
El Salvador	4.3	4.5	4.8	4.1	2.1	3.0	3.3	2.4	1.5
North and Central America Average	5.0	3.9	4.6	4.1	3.5	3.6	4.6	3.8	2.2

Información recuperada de (Crotti & Misrahi, 2018)

Ilustración 34: Efecto multiplicador del turismo

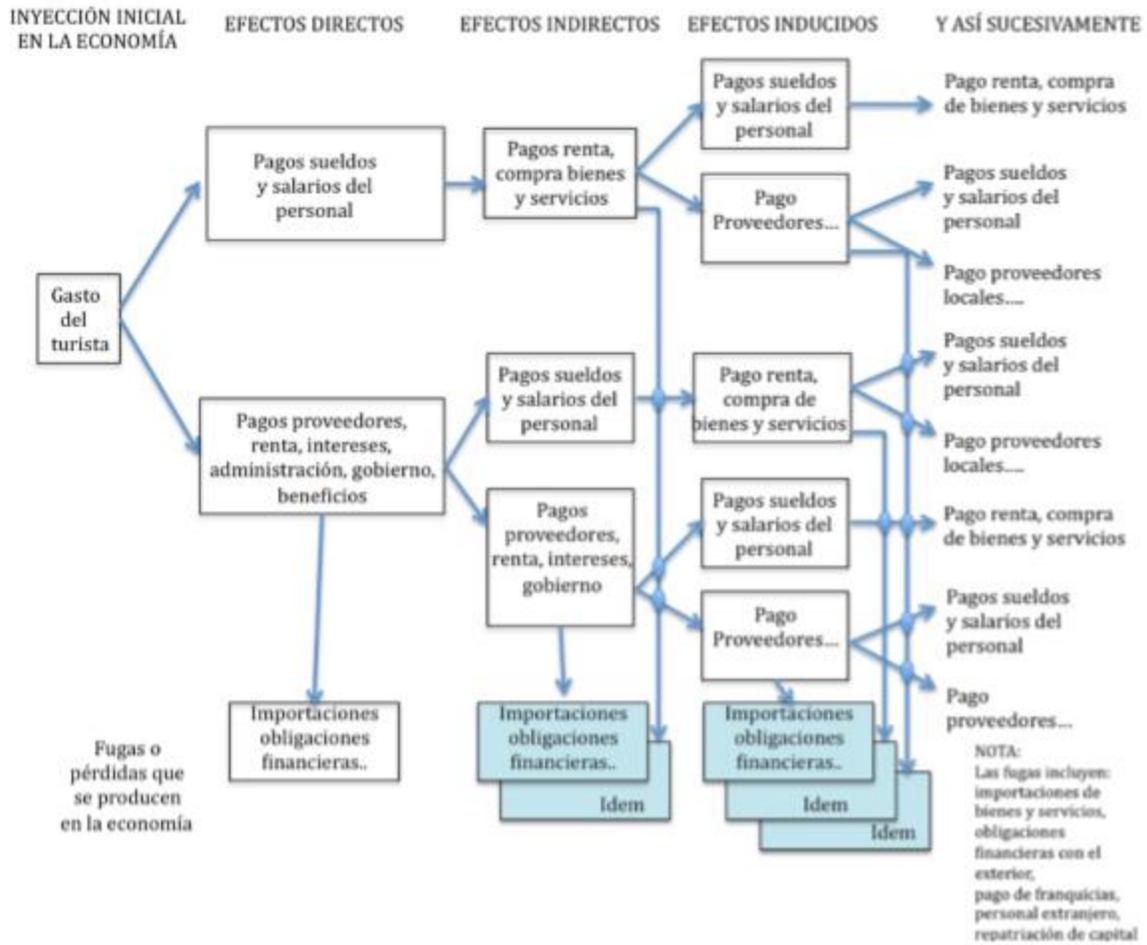
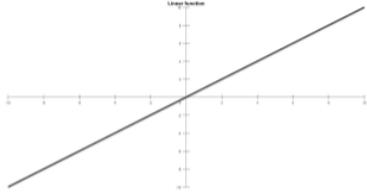
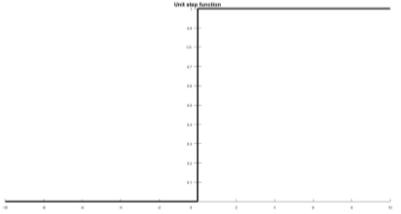
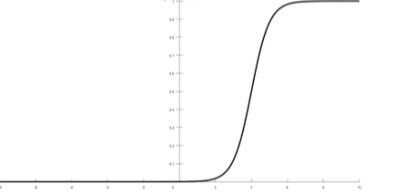
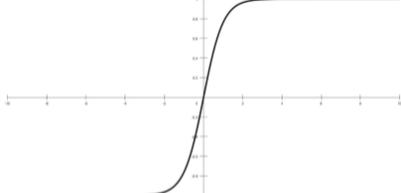
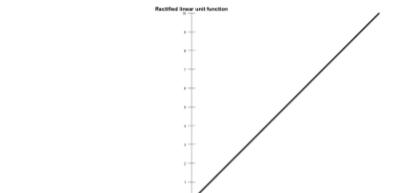


Ilustración tomada de (Octavio, 2008)

Tabla 14: Principales funciones de activación en el diseño de redes neuronales

Función	Descripción	Gráfica
Función Lineal	La entrada es igual que la salida. $y = f(x) = x$	

Función de activación escalón unitario	La salida asume el valor 0 para un argumento negativo y un 1 para un argumento positivo. El rango es (0,1) y la salida es binaria por naturaleza	
Función Sigmoid	Es una función matemática que produce una curva sigmoideal, en forma de S. hace que el modelo sea logístico por naturaleza. $f(x) = \frac{1}{1+e^{-x}}$	
Función tangente hiperbólica	Se puede decir que es una función escalada de la función sigmoid. $f(x) = \tanh(x)$	
Función unidad lineal rectificada (ReLU):	Es la función de activación más utilizada desde 2015. El rango es (0,inf)	

Elaboración propia con información recuperada de (Match, 2001)

Ilustración 35: Aplicaciones de la minería de datos

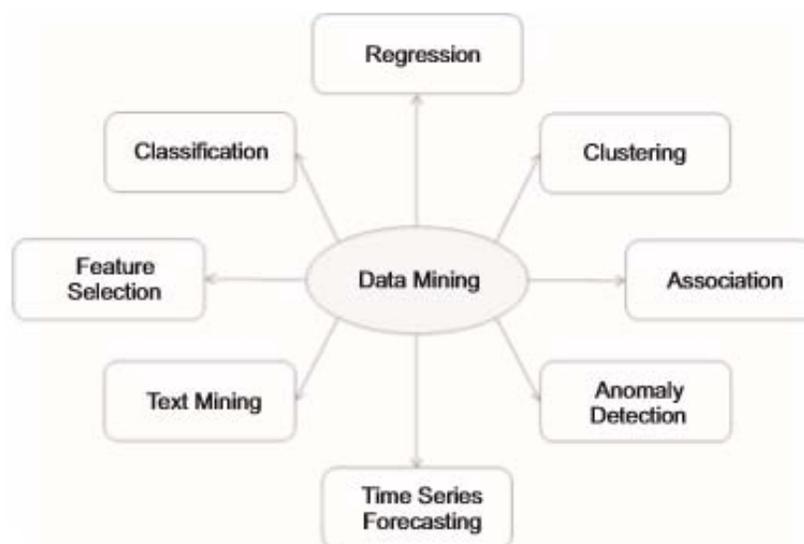
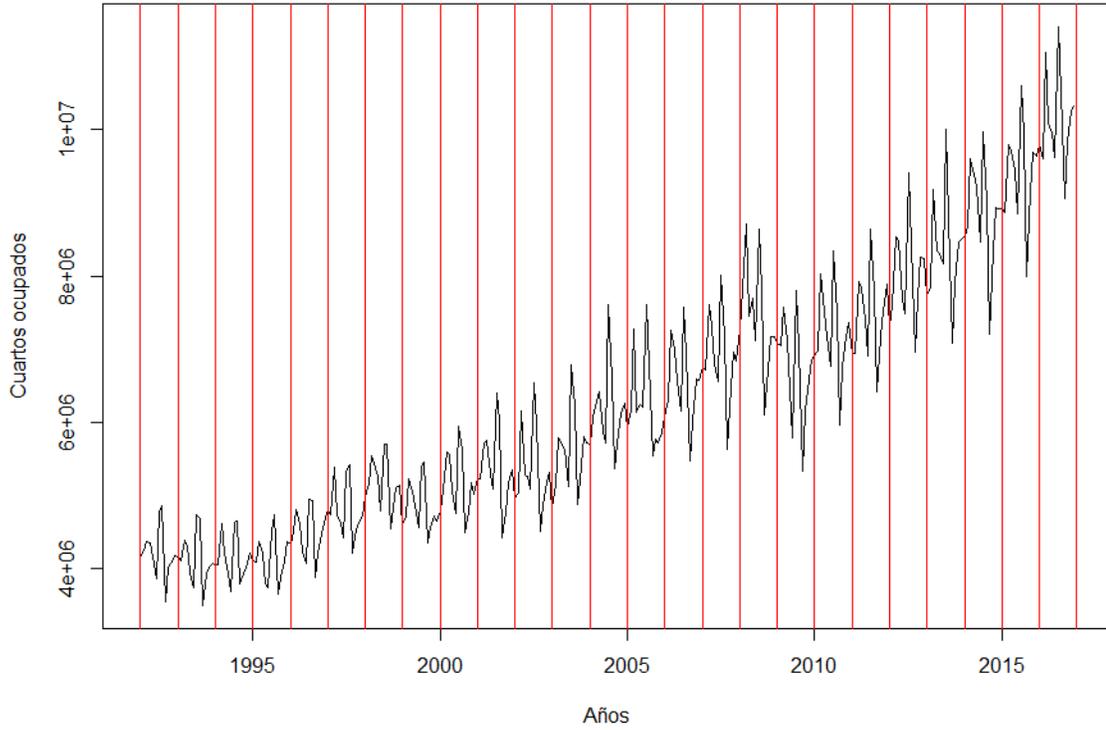


Ilustración tomada de (Kotu & Deshpande, 2015)

Ilustración 36: Comportamiento anual de los cuartos de hotel ocupados

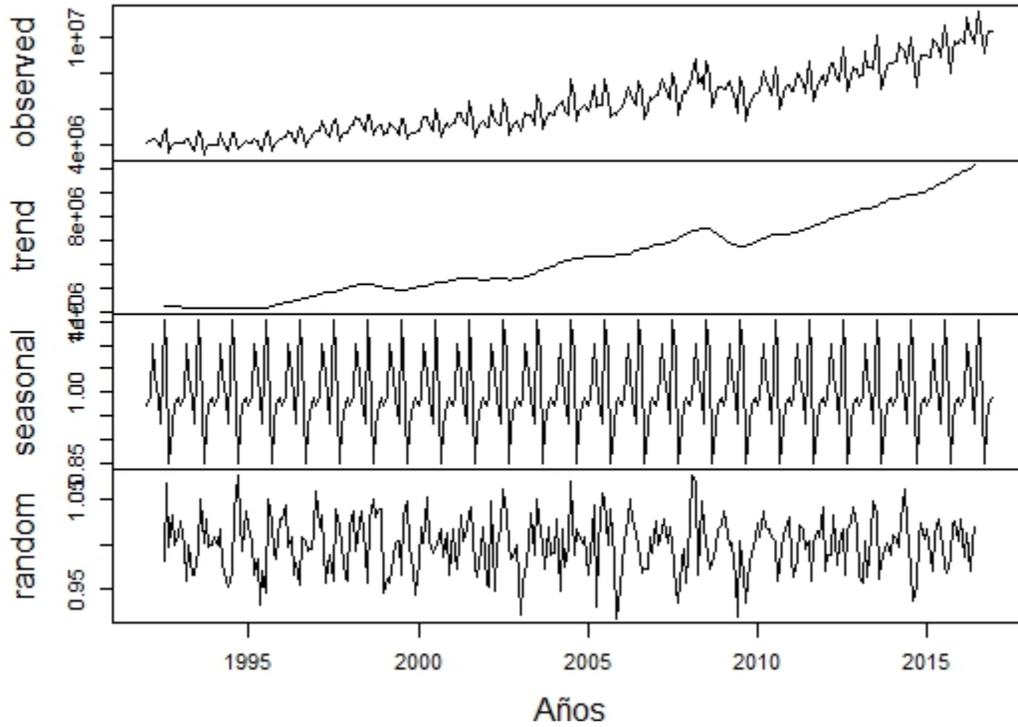
Comportamiento anual de los cuartos de hotel ocupados



Elaboración propia a partir de (Team, 2017)

Ilustración 37: Descomposición de la serie temporal de cuartos de hotel ocupados

Descomposición de la serie temporal



Elaboración propia a partir de (Team, 2017)

Tabla 15: Resultado de los pesos de la red neuronal artificial por tipo de interacción

Interacción	Valor del peso
Intercept.to.1layhid1	-2.20913117
TI.to.1layhid1	0.10773707
TC_CAN.to.1layhid1	-1.72313568
TC_USA.to.1layhid1	-0.35921081
POB_CAN.to.1layhid1	-0.1557009
POB_USA.to.1layhid1	-1.68227919
PS_CAN.to.1layhid1	1.90988989
PS_USA.to.1layhid1	1.49911247
IP_CAN.to.1layhid1	-0.53979078
IP_USA.to.1layhid1	0.36038852
ENE.to.1layhid1	0.80682588
FEB.to.1layhid1	1.10324581
MAR.to.1layhid1	0.3720206
ABR.to.1layhid1	0.81430051
MAY.to.1layhid1	-1.6856815
JUN.to.1layhid1	-1.63197545
JUL.to.1layhid1	-1.55113892
AGO.to.1layhid1	-0.18826529
SEP.to.1layhid1	2.08810968
OCT.to.1layhid1	0.81988558
NOV.to.1layhid1	0.14529775
...	...
Intercept.to.CO	2.37766423
3layhid.1.to.CO	-3.2229823
3layhid.2.to.CO	-0.97703141
3layhid.3.to.CO	-2.92294343

Elaboración propia a partir de (Team, 2017)

Tabla 16: Coeficientes de suavizado HW para las variables de entrada

Coeficientes de suavizado de Holt-Winters			
	Alpha	Beta	Gamma
TI	0.992067254	0.001374602	1
TC_CAN	0.880120896	0.012352846	0.710294222
TC_USA	1	0.012277353	0
Pob_CAN	0.928432316	0	1
Pob_USA	0.890737893	0.038521611	1
PS_CAN	0.868199961	0.09696771	1
PS_USA	0.973030054	1	1
IP_CAN	1	0	0

IP_USA	0.998388456	0	1
---------------	-------------	---	---

Elaboración propia a partir de (Team, 2017)

Tabla 17: Parámetros del modelos SARIMA para las variables de entrada

Parámetros del Modelo SARIMA								
	p	d	q	P	D	Q	m	drift
TI	0	1	0	0	0	1	12	FALSE
TC_CAN	0	1	0	0	0	1	12	TRUE
TC_USA	1	1	0	-	-	-	-	TRUE
Pob_CAN	4	0	3	2	1	1	12	TRUE
Pob_USA	0	2	1	-	-	-	-	FALSE
PS_CAN	2	2	1	0	0	2	12	FALSE
PS_USA	2	2	3	0	0	1	12	FALSE
IP_CAN	0	1	0	2	0	0	12	FALSE
IP_USA	0	1	0	2	0	1	12	FALSE

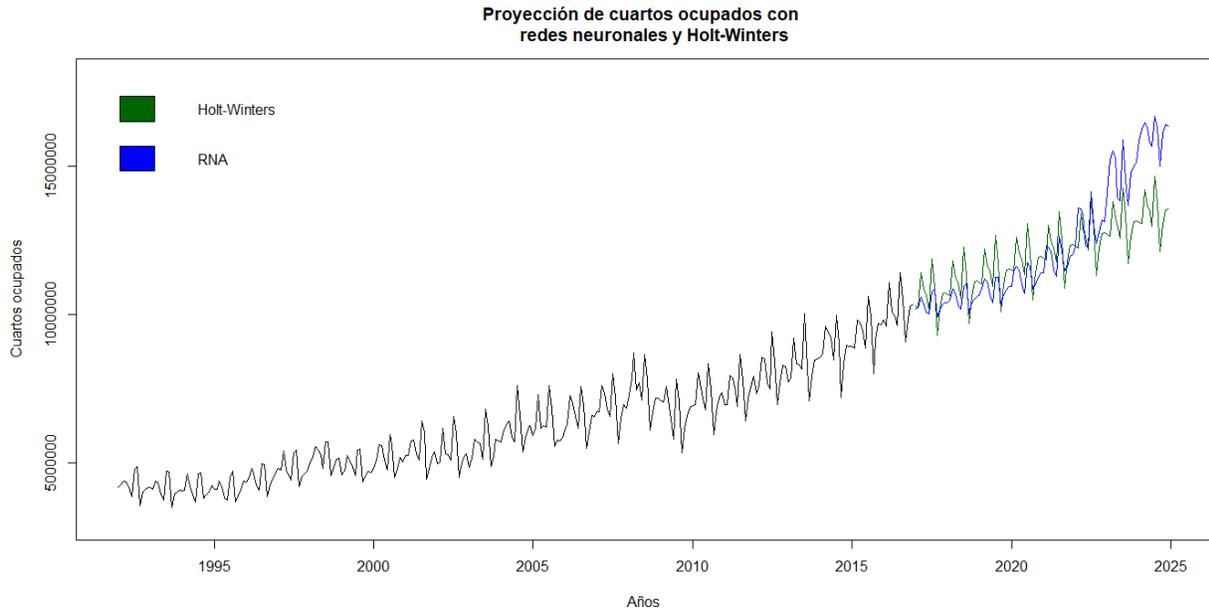
Elaboración propia a partir de (Team, 2017)

Tabla 18: Pronósticos de las variables de entrada para el año 2019

Fecha	TI	TC_CAN	TC_USA	Pob_CAN	Pob_USA	PS_CAN	PS_USA	IP_CAN	IP_USA
1/1/2019	29877.3317	16.526844	22.2040144	12372500.1	109726472	1.07975343	1.06756771	4335.60151	4453.4061
2/1/2019	29858.8895	16.6904167	22.263087	12372514.6	109736366	1.08170263	1.06923113	4335.60151	4456.43571
3/1/2019	29861.7976	16.625804	22.3221595	12372526.9	109733797	1.08361594	1.07092684	4335.60151	4459.46532
4/1/2019	29865.5942	16.8381329	22.3812321	12398807.7	109935945	1.08551225	1.07263698	4335.60151	4462.49492
5/1/2019	29869.4127	16.9484782	22.4403047	12398821.6	109933549	1.08740053	1.07435354	4335.60151	4465.52453
6/1/2019	29873.2316	16.9167632	22.4993773	12398833.5	109921519	1.08928499	1.07607298	4335.60151	4468.55414
7/1/2019	29877.0505	16.9469648	22.5584498	12439919.4	110122035	1.09116764	1.07779371	4335.60151	4471.58375
8/1/2019	29880.8694	17.0806379	22.6175224	12439932.1	110123223	1.09304944	1.079515	4335.60151	4474.61335
9/1/2019	29884.6883	17.2181206	22.676595	12439943	110116250	1.09493084	1.08123656	4335.60151	4477.64296
10/1/2019	29888.5072	17.1327219	22.7356676	12485279.2	110321370	1.09681205	1.08295822	4335.60151	4480.67257
11/1/2019	29892.326	17.2881875	22.7947401	12485290.5	110325540	1.09869316	1.08467994	4335.60151	4483.70217
12/1/2019	29896.1449	17.2792524	22.8538127	12485300.1	110320702	1.10057423	1.08640168	4335.60151	4486.73178

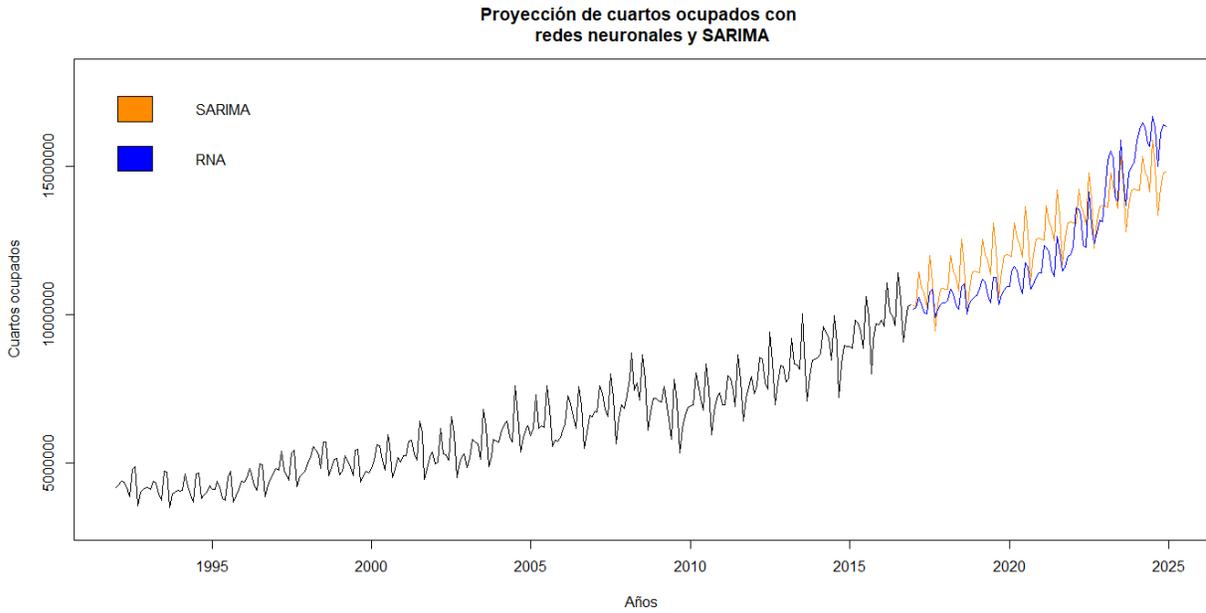
Elaboración propia a partir de (Team, 2017)

Ilustración 38: Comparación de pronóstico RNA vs HW



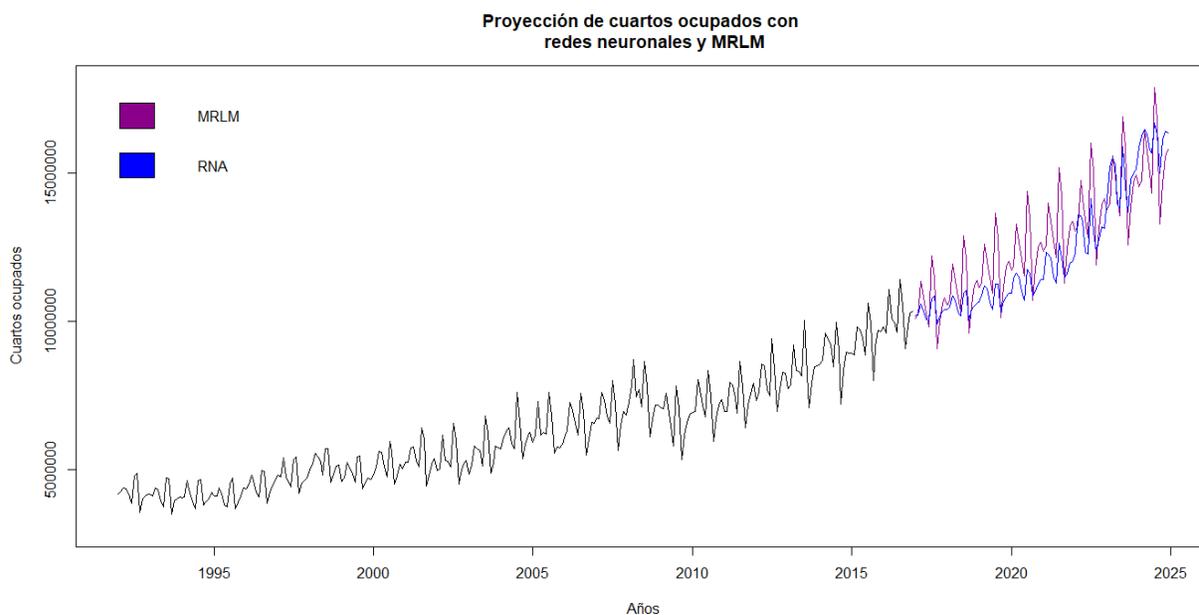
Elaboración propia a partir de (Team, 2017)

Ilustración 39: Comparación de pronóstico RNA vs SARIMA



Elaboración propia a partir de (Team, 2017)

Ilustración 40: Comparación de pronóstico RNA vs MRLM



Elaboración propia a partir de (Team, 2017)

Tabla 19: Valores puntuales de los pronósticos con diferentes métodos

Fecha	Holt-Winters	SARIMA	MRLM	RNA
1/1/2017	10274053	10306800	10069561	10184140
2/1/2017	10231076	10280223	10191267	10224639
3/1/2017	11388426	11430415	11342880	10565193
4/1/2017	10857901	10913485	10826362	10333767
5/1/2017	10629158	10684323	10276250	10058544
6/1/2017	10147290	10240325	9799440	10022432
7/1/2017	11851855	11975228	12209962	10754713
8/1/2017	10939467	11092923	11473244	10852855
9/1/2017	9284755	9448064	9083886	9900664
10/1/2017	10205128	10359599	10000215	10186083
11/1/2017	10704378	10861733	10604292	10317081
12/1/2017	10732502	10893430	10773595	10380211
1/1/2018	10676054	10863572	10557557	10391933
2/1/2018	10633078	10836995	10708792	10471137
3/1/2018	11790427	11987186	11936634	10852147
4/1/2018	11259903	11470257	11382043	10704512
5/1/2018	11031160	11241095	10819313	10297315
6/1/2018	10549292	10797096	10315968	10176642

7/1/2018	12253857	12531999	12869570	10944697
8/1/2018	11341468	11649695	12101312	11022872
9/1/2018	9686757	10004836	9578845	10011911
10/1/2018	10607130	10916370	10553869	10385955
11/1/2018	11106380	11418504	11207377	10522494
12/1/2018	11134504	11450202	11379452	10616055
1/1/2019	11078056	11420343	11125583	10640388
2/1/2019	11035079	11393766	11278596	10867718
3/1/2019	12192429	12543958	12591253	11188182
4/1/2019	11661905	12027028	12025052	11088232
5/1/2019	11433162	11797866	11449066	10631606
6/1/2019	10951294	11353868	10926410	10391173
7/1/2019	12655858	13088771	13630188	11247104
8/1/2019	11743470	12206467	12803823	11260787
9/1/2019	10088759	10561607	10128383	10337276
10/1/2019	11009132	11473142	11161374	10652355
11/1/2019	11508382	11975276	11825637	10805126
12/1/2019	11536506	12006973	12005529	10944968
1/1/2020	11480058	11977115	11714110	10942816
2/1/2020	11437081	11950538	11874794	11469488
3/1/2020	12594431	13100730	13263494	11609957
4/1/2020	12063906	12583800	12669169	11506907
5/1/2020	11835164	12354638	12064690	11005983
6/1/2020	11353296	11910640	11516339	10706308
7/1/2020	13057860	13645543	14369193	11752529
8/1/2020	12145472	12763239	13500971	11613165
9/1/2020	10490761	11118379	10682205	10837795
10/1/2020	11411133	12029914	11774275	11014288
11/1/2020	11910384	12532048	12477768	11236490
12/1/2020	11938508	12563745	12670377	11387801
1/1/2021	11882060	12533887	12350985	11402552
2/1/2021	11839083	12507310	12522225	12335234
3/1/2021	12996433	13657502	13983086	12257566
4/1/2021	12465908	13140572	13358459	12128965
5/1/2021	12237166	12911410	12722941	11499770
6/1/2021	11755298	12467412	12146438	11279500
7/1/2021	13459862	14202315	15157584	12628401
8/1/2021	12547474	13320010	14243796	12133586
9/1/2021	10892762	11675151	11271580	11473334
10/1/2021	11813135	12586686	12425709	11598056
11/1/2021	12312386	13088820	13170038	11948129

12/1/2021	12340509	13120517	13375278	12010726
1/1/2022	12284062	13090659	13029904	12290222
2/1/2022	12241085	13064082	13211821	13572119
3/1/2022	13398435	14214273	14754542	13549090
4/1/2022	12867910	13697344	14096802	13374623
5/1/2022	12639167	13468181	13427442	12325861
6/1/2022	12157299	13024183	12820243	12267576
7/1/2022	13861864	14759086	15999959	14133856
8/1/2022	12949476	13876782	15036826	13000085
9/1/2022	11294764	12231923	11900269	12372903
10/1/2022	12215137	13143457	13120026	12750010
11/1/2022	12714387	13645591	13907276	13193949
12/1/2022	12742511	13677288	14125356	13122989
1/1/2023	12686063	13647430	13754909	14106396
2/1/2023	12643087	13620853	13947826	15201664
3/1/2023	13800436	14771045	15577470	15510937
4/1/2023	13269912	14254115	14883981	15261758
5/1/2023	13041169	14024953	14178137	13945366
6/1/2023	12559301	13580955	13537842	13813056
7/1/2023	14263866	15315858	16896603	15870469
8/1/2023	13351477	14433554	15880496	14521693
9/1/2023	11696766	12788694	12568747	13667328
10/1/2023	12617139	13700229	13857894	14761299
11/1/2023	13116389	14202363	14690344	14994202
12/1/2023	13144513	14234060	14921642	15112314
1/1/2024	13088065	14204202	14526348	15853622
2/1/2024	13045088	14177625	14730695	16235553
3/1/2024	14202438	15327817	16459063	16448612
4/1/2024	13671914	14810887	15726978	16305081
5/1/2024	13443171	14581725	14981777	15762244
6/1/2024	12961303	14137727	14305782	15677897
7/1/2024	14665867	15872630	17855809	16660485
8/1/2024	13753479	14990326	16782713	16301743
9/1/2024	12098768	13345466	13283365	14982370
10/1/2024	13019140	14257001	14646415	16138541
11/1/2024	13518391	14759135	15526875	16405886
12/1/2024	13546515	14790832	15771998	16352420

Elaboración propia a partir de (Team, 2017)

Código

```
library(readxl) #Librería para Leer archivos con extensión .xlsx
library(tseries) #Librería para generar series de tiempo
library(forecast) #Librería para realizar pronósticos
library(lubridate) #Librería para manipular fechas
library(neuralnet) #Librería para generar al algoritmo de redes neuronales

BF <- read_excel("datos_tesis.xlsx",sheet=1) #Leer la base de datos
BF2 <- read_excel("datos_tesis.xlsx",sheet = 2)

CO<-ts(BF$CO, start=c(1992,1), end=c(2016,12), frequency = 12) #Convertir la
variable de cuartos a serie de tiempo
plot(CO,main="Cuartos de hotel ocupados en la República Mexicana de 1992 a
2016",ylab="Cuartos ocupados",xlab="Años") #Graficar
plot(CO,type="l",main="Comportamiento anual de los cuartos de hotel ocupados"
,xlab="Años",ylab="Cuartos ocupados")
abline(v=c(1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,
2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017), col="red")

#Convertir a serie de tiempo todas las variables explicativas

TI<-ts(BF2$TI, start=c(2010,1), end=c(2016,12), frequency = 12)
TCCAN<-ts(BF$TC_CAN, start=c(1992,1), end=c(2016,12), frequency = 12)
TCUSA<-ts(BF$TC_USA, start=c(1992,1), end=c(2016,12), frequency = 12)
PobCAN<-ts(BF$POB_CAN, start=c(1992,1), end=c(2016,12), frequency = 12)
PobUSA<-ts(BF$POB_USA, start=c(1992,1), end=c(2016,12), frequency = 12)
PSCAN<-ts(BF$PS_CAN, start=c(1992,1), end=c(2016,12), frequency = 12)
PSUSA<-ts(BF$PS_USA, start=c(1992,1), end=c(2016,12), frequency = 12)
IPCAN<-ts(BF$IP_CAN, start=c(1992,1), end=c(2016,12), frequency = 12)
IPUSA<-ts(BF$IP_USA, start=c(1992,1), end=c(2016,12), frequency = 12)

#Pronóstico HW de cuartos ocupados CO

COHW<-HoltWinters(CO) #Aplicar la función Holt-Winters
CP1<-((COHW$SSE/(297))^0.5) #medida de error con 3 grados de Libertad
plot(COHW,main="Pronóstico con método Holt-Winters",#Graficar
      xlab="Años",ylab="Cuartos ocupados")
legend("topleft",legend=c("Observaciones","Ajuste"),fill=c("black","red"),
      bty="n")
text(x=2015,y=4000000,label="Error=223769",font = 4)
FCOHW<-forecast(COHW, h=96) #Calcular el pronóstico

#Pronóstico SARIMA de cuartos ocupados CO

COSAR<-auto.arima(CO) #Aplicar la función SARIMA
CP2<-round((COSAR$sigma2)^0.5,0) #medida de error
plot(COSAR$x,main="Pronóstico con SARIMA",xlab="Años",
```

```

ylab="Cuartos ocupados") #Graficar
lines(COSAR$fitted,col="red")
legend("topleft",legend=c("Observaciones","Ajuste"),fill=c("black","red"),
bty="n")
text(x=2015,y=4000000,label="Error=227102",font = 4)
FCOSAR<-forecast(COSAR, h=96) #Calcular el pronóstico

#Comparación de método SARIMA y HW para selección de entradas
#Tasa de inseguridad
TIHW<-HoltWinters(TI)
FTI<-forecast(TIHW, h=96)
ETI1<-(TIHW$SSE/(length(TI)-3))^0.5
TISAR<-auto.arima(TI)
TISARf<-forecast(TISAR, h=96)
ETI2<-(TISAR$sigma2)^.5

#Tipo de cambio de Canadá
TCCANHW<-HoltWinters(TCCAN)
FTC_CAN<-forecast(TCCANHW, h=96)
ETCCAN1<-(TCCANHW$SSE/(297))^.5
TCCANSAR<-auto.arima(TCCAN)
TCCANSARf<-forecast(TCCANSAR, h=96)
ETCCAN2<-(TCCANSAR$sigma2)^.5

#Tipo de cambio de Estados Unidos
TCUSAHW<-HoltWinters(TCUSA)
ETCUSA1<-(TCUSAHW$SSE/(297))^.5
TCUSASAR<-auto.arima(TCUSA)
FTC_USA<-forecast(TCUSASAR, h=96)
ETCUSA2<-(TCUSASAR$sigma2)^.5

#Población de Canadá
PobCANHW<-HoltWinters(PobCAN)
EPobCAN1<-(PobCANHW$SSE/(297))^.5
PobCANSAR<-auto.arima(PobCAN)
FPob_CAN<-forecast(PobCANSAR, h=96)
EPobCAN2<-(PobCANSAR$sigma2)^.5

#Población de Estados Unidos
PobUSAHW<-HoltWinters(PobUSA)
FPob_USA<-forecast(PobUSAHW, h=96)
EPobUSA1<-(PobUSAHW$SSE/(297))^.5
PobUSASAR<-auto.arima(PobUSA)
PobUSASARf<-forecast(PobUSASAR, h=96)
EPobUSA2<-(PobUSASAR$sigma2)^.5

#Precio de servicio de Canadá
PSCANHW<-HoltWinters(PSCAN)

```

```

EPSCAN1<- (PSCANHW$SSE/(297))^.5
PSCANSAR<-auto.arima(PSCAN)
FPS_CAN<-forecast(PSCANSAR, h=96)
EPSCAN2<- (PSCANSAR$sigma2)^.5

#Precio de servicio de Estados Unidos
PSUSAHW<-HoltWinters(PSUSA)
EPSUSA1<- (PSUSAHW$SSE/(297))^.5
PSUSASAR<-auto.arima(PSUSA)
FPS_USA<-forecast(PSUSASAR, h=96)
EPSUSA2<- (PSUSASAR$sigma2)^.5

#Ingreso de Canadá
IPCANHW<-HoltWinters(IPCAN)
FIP_CAN<-forecast(IPCANHW, h=96)
EIPCAN1<- (IPCANHW$SSE/(297))^.5
IPCANSAR<-auto.arima(IPCAN)
IPCANSARf<-forecast(IPCANSAR, h=96)
EIPCAN2<- (IPCANSAR$sigma2)^.5

#Ingreso de Estados Unidos
IPUSAHW<-HoltWinters(IPUSA)
EIPUSA1<- (IPUSAHW$SSE/(297))^.5
IPUSASAR<-auto.arima(IPUSA)
FIP_USA<-forecast(IPUSASAR, h=96)
EIPUSA2<- (IPUSASAR$sigma2)^.5

#Concentración de Los coeficientes de suavizado de HW
CHW<-matrix(c(TIHW$alpha, TIHW$beta, TIHW$gamma, TCCANHW$alpha, TCCANHW$beta, TCCANHW$gamma, TCUSAHW$alpha, TCUSAHW$beta, TCUSAHW$gamma, PobCANHW$alpha, PobCANHW$beta, PobCANHW$gamma, PobUSAHW$alpha, PobUSAHW$beta, PobUSAHW$gamma, PSCANHW$alpha, PSCANHW$beta, PSCANHW$gamma, PSUSAHW$alpha, PSUSAHW$beta, PSUSAHW$gamma, IPCANHW$alpha, IPCANHW$beta, IPCANHW$gamma, IPUSAHW$alpha, IPUSAHW$beta, IPUSAHW$gamma), nrow=9, byrow=T)
colnames(CHW)<-c("Alpha", "Beta", "Gamma")
rownames(CHW)<-c("TI", "TC_CAN", "TC_USA", "Pob_CAN", "Pob_USA", "PS_CAN", "PS_USA", "IP_CAN", "IP_USA")

#Concentración de Los parámetros del modelo SARIMA
CSAR<-matrix(c(0,1,0,0,0,1,12,F,
               0,1,0,0,0,1,12,T,
               1,1,0,"-","-","-","-",T,
               4,0,3,2,1,1,12,T,
               0,2,1,"-","-","-","-",F,
               2,2,1,0,0,2,12,F,

```

```

        2,2,3,0,0,1,12,F,
        0,1,0,2,0,0,12,F,
        0,1,0,2,0,1,12,F),nrow = 9,byrow = T)
colnames(CSAR)<-c("p","d","q","P","D","Q","m","drift")
rownames(CSAR)<-c("TI","TC_CAN","TC_USA","Pob_CAN","Pob_USA","PS_CAN",
"PS_USA","IP_CAN","IP_USA")

#Concentración de Los errores de HW y SARIMA
CE<-matrix(c(ETI1,ETI2,ETCCAN1,ETCCAN2,ETCUSA1,ETCUSA2,EPobCAN1,EPobCAN2,EPob
USA1, EPobUSA2,EPSCAN1,EPSCAN2,EPSUSA1,EPSUSA2,EIPCAN1,EIPCAN2,EIPUSA1,
EIPUSA2),nrow=9,byrow=T)
colnames(CE)<-c("Holt-Winters","SARIMA")
rownames(CE)<-c("TI","TC_CAN","TC_USA","Pob_CAN","Pob_USA","PS_CAN","PS_USA",
"IP_CAN","IP_USA")
MP<-c()
for (i in 1:9) { #Selección del mejor modelo con menor error
  if(CE[i,1] < CE[i,2]){
    MP[i]<-"HW"
  }else{
    MP[i]<-"SARIMA"
  }
}
format(data.frame(CE,MP),scient=F)

#Modelo con REDES NEURONALES ARTIFICIALES
entradas<- data.frame(FTI$mean,FTC_CAN$mean,FTC_USA$mean,FPob_CAN$mean,
FPob_USA$mean,FPS_CAN$mean,FPS_USA$mean,FIP_CAN$mean,
FIP_USA$mean)
names(entradas)<-c("TI","TC_CAN","TC_USA","Pob_CAN","Pob_USA","PS_CAN",
"PS_USA","IP_CAN","IP_USA")

ENE<-rep(c(1,0,0,0,0,0,0,0,0,0,0,0), times=25) #variables ficticias
FEB<-rep(c(0,1,0,0,0,0,0,0,0,0,0,0), times=25)
MAR<-rep(c(0,0,1,0,0,0,0,0,0,0,0,0), times=25)
ABR<-rep(c(0,0,0,1,0,0,0,0,0,0,0,0), times=25)
MAY<-rep(c(0,0,0,0,1,0,0,0,0,0,0,0), times=25)
JUN<-rep(c(0,0,0,0,0,1,0,0,0,0,0,0), times=25)
JUL<-rep(c(0,0,0,0,0,0,1,0,0,0,0,0), times=25)
AGO<-rep(c(0,0,0,0,0,0,0,1,0,0,0,0), times=25)
SEP<-rep(c(0,0,0,0,0,0,0,0,1,0,0,0), times=25)
OCT<-rep(c(0,0,0,0,0,0,0,0,0,1,0,0), times=25)
NOV<-rep(c(0,0,0,0,0,0,0,0,0,0,1,0), times=25)

meses<-data.frame(ENE, FEB, MAR, ABR, MAY, JUN, JUL, AGO, SEP,OCT, NOV)

dmax<-apply(BF[,2:11],2,max) #Calcular el valor máximo
dmin<-apply(BF[,2:11],2,min) #Calcular el valor mínimo

```

```

escala<-scale(BF[,2:11],center=dmin,scale=dmax-dmin) #Escalar Los valores
BFR<-data.frame(escala, meses)

muestra<-sample(1:nrow(BF),round(0.8*nrow(BF))) #Calcular el muestra
entrena<-as.data.frame(BFR[muestra,]) #Datos para La prueba
prueba<-as.data.frame(BFR[-muestra,]) #Datos para La verificación
nombres<-names(entrena) #Nombres de Las variables explicativas
Formu<-as.formula(paste("CO ~",paste(nombres[!nombres %in% "CO"],
collapse= " + "))) #Generación de La fórmula
Formu

## CO ~ TI + TC_CAN + TC_USA + POB_CAN + POB_USA + PS_CAN + PS_USA +
##      IP_CAN + IP_USA + ENE + FEB + MAR + ABR + MAY + JUN + JUL +
##      AGO + SEP + OCT + NOV

redneu<-neuralnet(formula = Formu,data=entrena,hidden = c(7,5,3),
linear.output = F) #Generar La red neuronal con La fórmula y Los datos de
prueba, tres capas ocultas y La salida no es lineal
pred<-compute(redneu,prueba[,c(2:length(prueba))])$net.result #Obtener el
ajuste con Los datos de verificación
pred_co<-pred*(max(BF$CO)-min(BF$CO))+min(BF$CO) #Desescalar Los valores
valores_reales<-prueba$CO*(max(BF$CO)-min(BF$CO))+min(BF$CO)
error1<-((sum((pred_co-valores_reales)^2)/(nrow(prueba-20)))^0.5 #Calcular el
error del modelo
error1

while (error1>223768) { #Simulación de Los errores hasta que encuentre un va
lor mínimo arbitrario, menor al error del modelo HW
  muestra<-sample(1:nrow(BF),round(0.8*nrow(BF)))
  entrena<-as.data.frame(BFR[muestra,])
  prueba<-as.data.frame(BFR[-muestra,])
  nombres<-names(entrena)
  Formu<-as.formula(paste("CO ~",paste(nombres[!nombres %in% "CO"],
collapse= " + ")))
  redneu<-neuralnet(formula = Formu,data=entrena,hidden = c(7,5,3),
linear.output = F)
  pred<-compute(redneu,prueba[,c(2:length(prueba))])$net.result
  pred_co<-pred*(max(BF$CO)-min(BF$CO))+min(BF$CO)
  valores_reales<-prueba$CO*(max(BF$CO)-min(BF$CO))+min(BF$CO)
  error<-((sum((pred_co-valores_reales)^2)/(nrow(prueba-20)))^0.5
  error1<-error
}
CP4<-error

plot(redneu,rep = "best",show.weights = F,arrow.length = 0.15,dimension = 8,
col.entry = "red",col.out = "red") #Graficar La red neuronal

COP<-compute(redneu,BFR[,c(2:length(BFR))])$net.result
COP2<-COP*(max(BF$CO)-min(BF$CO))+min(BF$CO)
COP22<-ts(COP2,start=c(1992,1), end=c(2016,12), frequency = 12)
plot(CO,type="l", main="Pronóstico con Redes Neuronales Artificiales",xlab="A

```

```

ños",ylab="Cuartos ocupados") #Graficar el ajuste del modelo
lines(COP22,type = "l",col="red")
legend("topleft",legend=c("Observaciones","Ajuste"),fill=c("black","red"),
bty="n")
text(x=2010,y=4000000,label="Error=215513",font = 4)

#Predicción con modelo de REDES NEURONALES ARTIFICIALES

entramax<-apply(entradas[,1:9],2,max)
entramin<-apply(entradas[,1:9],2,min)
entra_escal<-scale(entradas[,1:9],center=entramin,scale=entramax-entramin)

ENE<-rep(c(1,0,0,0,0,0,0,0,0,0,0,0), times=8)
FEB<-rep(c(0,1,0,0,0,0,0,0,0,0,0,0), times=8)
MAR<-rep(c(0,0,1,0,0,0,0,0,0,0,0,0), times=8)
ABR<-rep(c(0,0,0,1,0,0,0,0,0,0,0,0), times=8)
MAY<-rep(c(0,0,0,0,1,0,0,0,0,0,0,0), times=8)
JUN<-rep(c(0,0,0,0,0,1,0,0,0,0,0,0), times=8)
JUL<-rep(c(0,0,0,0,0,0,1,0,0,0,0,0), times=8)
AGO<-rep(c(0,0,0,0,0,0,0,1,0,0,0,0), times=8)
SEP<-rep(c(0,0,0,0,0,0,0,0,1,0,0,0), times=8)
OCT<-rep(c(0,0,0,0,0,0,0,0,0,1,0,0), times=8)
NOV<-rep(c(0,0,0,0,0,0,0,0,0,0,1,0), times=8)

mes<-data.frame(ENE, FEB, MAR, ABR, MAY, JUN, JUL, AGO, SEP,OCT, NOV)
entraF<-data.frame(entra_escal,mes)
predt<-compute(redneu,entraF[,1:length(entraF)])$net.result
predt_co<-predt*(max(BF$CO)-min(BF$CO))+min(BF$CO) #Calcular Los pronósticos

#Modelo de Regresión Lineal Múltiple

BML<-data.frame(BF,meses)
ML1<-lm(log(CO)~.,data = BML) #Calcular el MRLM con todas Las variables

ML2<-lm(log(CO)~Fecha+POB_USA+PS_CAN+PS_USA+IP_CAN+IP_USA+MAR+ABR+JUN+JUL+AGO
+SEP+OCT,data=BML) #Calcular el MRLM con Las variables significativas
jarque.bera.test(ML2$residuals) #Prueba de normalidad
summary(ML2) #Resumen del modelo
coefficients(ML2) #Coeficientes del modelo

Ent <- read_excel("entradas.xlsx", sheet=1)
Ent<-Ent[,-5] #Eliminar columnas que no son necesarias
Ent<-Ent[,-2]

MAR<-rep(c(0,0,1,0,0,0,0,0,0,0,0,0), times=8)
ABR<-rep(c(0,0,0,1,0,0,0,0,0,0,0,0), times=8)
JUN<-rep(c(0,0,0,0,0,1,0,0,0,0,0,0), times=8)
JUL<-rep(c(0,0,0,0,0,0,1,0,0,0,0,0), times=8)

```

```

AGO<-rep(c(0,0,0,0,0,0,0,1,0,0,0,0), times=8)
SEP<-rep(c(0,0,0,0,0,0,0,0,1,0,0,0), times=8)
OCT<-rep(c(0,0,0,0,0,0,0,0,0,1,0,0), times=8)

DF<-data.frame(Ent, MAR, ABR, JUN, JUL, AGO, SEP, OCT)
DF<-DF[, -3]
DF<-DF[, -2]

Pronl<-c()#Calcular el pronóstico aplicando función exponencial
for (i in 1:96){
  Pronl[i]<-exp(sum(as.numeric(coefficients(ML2)[2:14])*as.numeric(DF[i,]))
+coefficients(ML2)[1])
}
BML<-BML[, -22]
BML<-BML[, -16]
BML<-BML[, -13]
BML<-BML[, -12]
BML<-BML[, -6]
BML<-BML[, -5]
BML<-BML[, -4]
BML<-BML[, -3]
BML<-BML[, -2]
DR<-BML
Orig<-c()
for (i in 1:300){
  Orig[i]<-exp(sum(as.numeric(coefficients(ML2)[2:14])*as.numeric(DR[i,]))+
coefficients(ML2)[1])
}

plot(CO,type="l",main="Pronóstico con Regresión Lineal Múltiple",xlab="Años",
yylab="Cuartos ocupados") #Graficar el ajuste del modelo

Orig2<-ts(Orig, start=c(1992,1), end=c(2016,12), frequency = 12)
lines(Orig2,type = "l",col="red")
legend("topleft",legend=c("Observaciones", "Ajuste"),fill=c("black", "red"),bty
="n")
text(x=2010,y=4000000,label="Error=259076",font = 4)
CP3<-((sum((BF$CO-Orig)^2)/(300-15))^0.5

#Comparación

CP<-matrix(c(CP1,CP2,CP3,CP4),byrow=T,nrow = 4)
colnames(CP)<-c("Error")
rownames(CP)<-c("Holt-Winters", "SARIMA", "MRLM", "RNA")

mcp<-c(rep("-",4))
mcp[which.min(CP)]<- "****"
compa<-data.frame(CP,mcp)

```

```
names(compa)<-c("Error", "Mínimo")
```

#Gráficas

```
H<-c(BF$CO,rep(NA,96))
H1<-ts(H,start=c(1992,1), end=c(2024,12), frequency = 12)
plot(H1,ylim=c(3000000,18000000),main="Proyección de cuartos ocupados con
diferentes métodos",xlab="Años",ylab="Cuartos ocupados")
```

```
Hj<-c(rep(NA,300),as.numeric(FCOHW$mean))
Hj1<-ts(Hj,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hj1,type="l",col="darkgreen")
Hs<-c(rep(NA,300),as.numeric(FCOSAR$mean))
Hs1<-ts(Hs,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hs1,type="l",col="darkorange")
Hp<-c(rep(NA,300),Pron1)
Hp1<-ts(Hp,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hp1,type="l",col="darkmagenta")
r3<-FCOHW$mean[1]-pred_co[1]
Hrs<-predt_co+r3
Hr<-c(rep(NA,300),Hrs)
Hr1<-ts(Hr,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hr1,type="l",col="blue")
legend("topleft",legend=c("Holt-Winters", "SARIMA", "MRLM", "RNA"),
fill=c("darkgreen","darkorange","darkmagenta","blue"),bty="n")
```

```
H<-c(BF$CO,rep(NA,96))
H1<-ts(H,start=c(1992,1), end=c(2024,12), frequency = 12)
plot(H1,ylim=c(3000000,18000000),main="Proyección de cuartos ocupados con
redes neuronales",xlab="Años",ylab="Cuartos ocupados")
r3<-FCOHW$mean[1]-pred_co[1]
Hrs<-predt_co+r3
Hr<-c(rep(NA,300),Hrs)
Hr1<-ts(Hr,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hr1,type="l",col="blue")
legend("topleft",legend=c("RNA"),
fill=c("blue"),bty="n")
```

```
H<-c(BF$CO,rep(NA,96))
H1<-ts(H,start=c(1992,1), end=c(2024,12), frequency = 12)
plot(H1,ylim=c(3000000,18000000),main="Proyección de cuartos ocupados con
redes neuronales y Holt-Winters",xlab="Años",ylab="Cuartos ocupados")
Hj<-c(rep(NA,300),as.numeric(FCOHW$mean))
Hj1<-ts(Hj,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hj1,type="l",col="darkgreen")
r3<-FCOHW$mean[1]-pred_co[1]
Hrs<-predt_co+r3
Hr<-c(rep(NA,300),Hrs)
Hr1<-ts(Hr,start=c(1992,1), end=c(2024,12), frequency = 12)
```

```

lines(Hr1,type="l",col="blue")
legend("topleft",legend=c("Holt-Winters","RNA"),
      fill=c("darkgreen","blue"),bty="n")

H<-c(BF$CO,rep(NA,96))
H1<-ts(H,start=c(1992,1), end=c(2024,12), frequency = 12)
plot(H1,ylim=c(3000000,18000000),main="Proyección de cuartos ocupados con
      redes neuronales y SARIMA",xlab="Años",ylab="Cuartos ocupados")
Hs<-c(rep(NA,300),as.numeric(FCOSAR$mean))
Hs1<-ts(Hs,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hs1,type="l",col="darkorange")
r3<-FCOHW$mean[1]-pred_co[1]
Hrs<-predt_co+r3
Hr<-c(rep(NA,300),Hrs)
Hr1<-ts(Hr,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hr1,type="l",col="blue")
legend("topleft",legend=c("SARIMA","RNA"),
      fill=c("darkorange","blue"),bty="n")

H<-c(BF$CO,rep(NA,96))
H1<-ts(H,start=c(1992,1), end=c(2024,12), frequency = 12)
plot(H1,ylim=c(3000000,18000000),main="Proyección de cuartos ocupados con
      redes neuronales y MRLM",xlab="Años",ylab="Cuartos ocupados")
Hp<-c(rep(NA,300),Pron1)
Hp1<-ts(Hp,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hp1,type="l",col="darkmagenta")
r3<-FCOHW$mean[1]-pred_co[1]
Hrs<-predt_co+r3
Hr<-c(rep(NA,300),Hrs)
Hr1<-ts(Hr,start=c(1992,1), end=c(2024,12), frequency = 12)
lines(Hr1,type="l",col="blue")
legend("topleft",legend=c("MRLM","RNA"),
      fill=c("darkmagenta","blue"),bty="n")

CPT<-data.frame(FCOHW$mean,FCOSAR$mean,Pron1,Hrs)
names(CPT)<-c("Pronóstico HW","Pronóstico SARIMA","Pronóstico MRLM",
"Pronóstico RNA")

```