

SEP

TNM

INSTITUTO TECNOLÓGICO DE CULIACÁN



PRONÓSTICO A CORTO PLAZO DE LA VELOCIDAD DEL
VIENTO EMPLEANDO REDES NEURONALES
CONVOLUCIONALES EN UNA TARJETA ACELERADORA
GRÁFICA (GPU)

TESIS

PRESENTADA ANTE EL DEPARTAMENTO ACADÉMICO DE ESTUDIOS DE POSGRADO
DEL INSTITUTO TECNOLÓGICO DE CULIACÁN EN CUMPLIMIENTO PARCIAL DE LOS
REQUISITOS PARA OBTENER EL GRADO DE

MAESTRO EN CIENCIAS DE LA COMPUTACIÓN

POR:

JESÚS ARMANDO GUERRA GARCÍA
INGENIERO EN SISTEMAS COMPUTACIONALES

DIRECTOR DE TESIS:

DR. RODRÍGUEZ RANGEL HÉCTOR

CULIACÁN, SINALOA

AGOSTO, 2018

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

Culiacán, Sin., 27 de Agosto del 2018

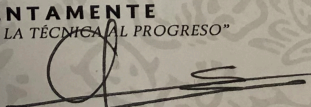
OFICIO: DEPI-340/VIII/2018
ASUNTO: **Autorización Impresión**

ING. JESÚS ARMANDO GUERRA GARCÍA
ESTUDIANTE DE LA MAestrÍA EN CIENCIAS DE LA COMPUTACIÓN
PRESENTE.


Por medio de la presente y en virtud de que ha completado los requisitos para el examen de grado de la **Maestría en Ciencias de la Computación**, se concede autorización para la impresión de la tesis titulada: **"PRONÓSTICO A CORTO PLAZO DE LA VELOCIDAD DEL VIENTO EMPLEANDO REDES NEURONALES CONVOLUCIONALES EN UNA TARJETA ACCELERADORA GRÁFICA (GPU)**, bajo la dirección del(a) **Dr. Héctor Rodríguez Rangel**

Sin otro particular reciba un cordial saludo.

ATENTAMENTE
"CON LA TÉCNICA AL PROGRESO"



M.C. MARÍA ARACELY MARTÍNEZ AMAYA
JEFE(A) DE LA DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN



SEP TecNM
Instituto Tecnológico
de Culiacán
División de Estudios
de Posgrado e Investigación

C.c.p. archivo

MAMA/lucy *

**“PRONÓSTICO A CORTO PLAZO DE LA
VELOCIDAD DEL VIENTO EMPLEANDO REDES
NEURONALES CONVOLUCIONALES EN UNA
TARJETA ACELERADORA GRÁFICA (GPU)**

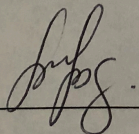
Tesis presentada por:

ING. JESÚS ARMANDO GUERRA GARCÍA

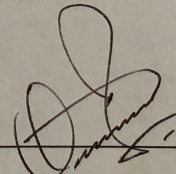
Aprobada en contenido y estilo por:



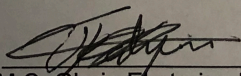
Dr. Héctor Rodríguez Rangel
Director de Tesis



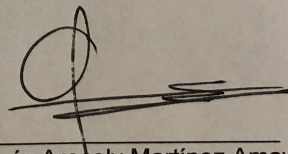
Dra. María Lucía Barrón Estrada
Secretario



Dr. Ramón Zatarain Cabada
Vocal -1



M.C. Gloria Ekaterine Peralta Peñuñuri
Vocal -2



M.C. María Aracely Martínez Amaya
Jefe(a) de la División de Estudios de
Posgrado e Investigación

Dedicatoria

Esta tesis se la dedico a mis padres Eliud García Coronel y Jesús Armando Guerra García que durante mis estudios como en otros aspectos importantes de la vida siempre me han apoyado.

Agradecimientos

Agradezco al Instituto Tecnológico de Culiacán por haberme permitido realizar mis estudios en esta gran institución, a todos los doctores del departamento de posgrado que durante estos dos años me transmitieron sus conocimientos de la mejor manera, al Consejo Nacional de Ciencia y Tecnología por el apoyo económico recibido durante mis estudios.

Declaración de autenticidad

Por la presente declaro que, salvo cuando se haga referencia específica al trabajo de otras personas, el contenido de esta tesis es original y no se ha presentado total o parcialmente para su consideración para cualquier otro título o grado en esta o cualquier otra Universidad. Esta tesis es resultado de mi propio esfuerzo, trabajo, dedicación y no incluye nada que sea resultado de algún trabajo realizado en colaboración, salvo que se indique específicamente en este documento.

Jesús Armando Guerra García. Culiacán, Sinaloa, México, 2018

Resumen

Actualmente la energía eléctrica se genera de diferentes maneras como lo son el gas natural, el carbón natural y el petróleo pero la mayoría proviene de medios fósiles, ocasionando daños irreversibles al medioambiente. Esto ha provocado que en los últimos años se haya desarrollado la necesidad de reducir este tipo de energías y aumentar el uso de energías renovables. Es necesario de alguna manera promover el uso de este tipo de energía. Se considera que el realizar un pronóstico acertado de algunas variables relacionadas con este tipo de energía ayudará a exista una mayor adopción de las energías renovables como la eólica. Además, el pronóstico favorecerá el que exista una disminución en los gastos de producción y será de gran ayuda en centrales eléctricas para la toma de decisiones debido a que se conocerá con anticipación la cantidad de materia prima con la que se contará, entre otras cosas.

En esta Tesis se trabajó con el pronóstico a corto plazo de la velocidad del viento. Se desarrolló una metodología para la obtención de bancos de modelos predictivos de la velocidad del viento, los cuales constan de 24 modelos independientes el uno del otro, obteniendo un pronóstico a 24 pasos (horas hacia adelante).

Para el entrenamiento de los distintos modelos se utilizó una técnica de aprendizaje profundo, Redes Neuronales Convolucionales (RNC). Se utilizó la técnica Gramian Angular Field para realizar la transformación del vector de observaciones de la serie a estructura de matriz. Se creó un banco de imágenes como entrada a la RNC para su entrenamiento. Se implementó un Algoritmo Genético Compacto (AGc) para realizar la optimización de los diferentes hiperparámetros de la RNC. Los resultados obtenidos fueron comparados con la técnica de predicción naïve. En estos se puede observar que con la metodología propuesta se obtuvieron mejores resultados en los 4 series de tiempo utilizadas que con naïve.

Palabras clave

Energías Renovables.

Series de Tiempo.

Pronóstico.

Inteligencia Artificial.

Redes Neuronales.

Algoritmo Genético Compacto.

Redes Neuronales Convolucionales.

Gramian Angular Field.

Banco de Imágenes.

Banco de Modelos.

Medidas de Aptitud.

Índice general

Índice de figuras	X
Índice de tablas	XI
1. Introducción	1
1.1. Definición del Problema	3
1.2. Objetivo	3
1.3. Objetivos Particulares	4
1.4. Justificación	4
1.5. Estructura de la Tesis	6
2. Marco teórico	7
2.1. Energía	7
2.1.1. Energías Sucias	8
2.1.2. Energía Renovable	8
2.1.3. Energía Eólica	9
2.2. Pronóstico	9
2.2.1. Serie de Tiempo	10
2.3. Pronóstico Mediante Inteligencia Artificial	11
2.3.1. Aprendizaje Máquina	12
2.3.1.1. Redes Neuronales Artificiales	13
2.3.2. Aprendizaje Profundo	15
2.3.2.1. Redes Neuronales Convolucionales	15
2.4. Algoritmo Genético	16
2.4.1. Algoritmo Genético Compacto	17
2.5. Medidas de Aptitud	18
3. Estado del Arte	20
3.1. Métodos de Pronóstico	20
3.2. Pronóstico Mediante Inteligencia Artificial	21
3.2.1. Aprendizaje Máquina	21
3.2.2. Aprendizaje Profundo	22

4. Diseño del experimento	25
4.1. Pre-Procesamiento de Datos	26
4.2. Serie de Tiempo a Imagen	29
4.3. Banco de Imágenes	30
4.4. Proceso de Serie de Tiempo a Imagen	32
4.4.1. Escaneo Lineal	32
4.4.2. Gramian Angular Field	34
4.4.3. Proceso de Entrenamiento	36
4.5. Banco de modelos	40
4.6. Tecnologías Utilizadas	43
4.6.1. Python	44
4.6.2. Keras	44
4.6.3. Tensorflow	44
5. Análisis de Resultados	45
6. Conclusiones	50
Bibliografía	52

Índice de figuras

1.1. Crecimiento anual de energías renovables 2002-2012 (Observ, 2013)	6
2.1. Representación gráfica de una serie de tiempo de la velocidad del viento	11
2.2. Neuron artificial básica (Anderson and McNeill, 1992)	14
2.3. Arquitectura de una RNA con 4 neuronas en la capa de entrada, 5 neuronas en la capas ocultas y una sola neurona de salida	14
2.4. Arquitectura general de una Red Neuronal Convolutacional. (Pérez-Espinosa et al., 2016)	16
2.5. Representación del cromosoma (Sivanandam and Deepa, 2007)	17
4.1. Proceso completo para la generación de un modelo predictivo de la velocidad del viento a un paso.	26
4.2. Ejemplo de dato atípico en un conjunto de observaciones	27
4.3. Ejemplo de serie de tiempo de la velocidad del viento suavizada al presentar datos atípicos.	28
4.4. Serie de tiempo normalizada en un rango de valores entre [0, 1]	29
4.5. Deslizamiento de la ventana w y asignación de valor esperado en el paso uno	30
4.6. Imagen a serie de tiempo por Escaneo Lineal.	33
4.7. Serie de tiempo a imagen por Escaneo Lineal.	34
4.8. Serie de tiempo codificada a una imagen utilizando GAF	36
4.9. Arquitectura utilizada para la RNC	37
4.10. Representación de la interacción entre AGc y RNC.	40
4.11. Deslizamiento de la venta y asignación de valor esperado en el paso dos	41
4.12. Ciclo de generación 24 bancos de modelos predictivos	42
4.13. Representación gráfica de la generación del banco de modelos y la interacción entre el AGc en el CPU y la RNC en el GPU	43
5.1. Relación de las posiciones del cromosoma con los diferentes parámetros proporcionados por el AGc	48

Índice de tablas

2.1. Tipos de energías renovables	8
2.2. Clasificación de las escalas de tiempo en pronóstico del viento	10
3.1. Métodos Estadísticos	21
3.2. Tabla comparativa de los diferentes métodos de predicción relacionados a este trabajo de investigación	24
4.1. Ejemplo de generación del banco de imágenes para un paso definido de uno. . .	31
5.1. Número de imágenes en el banco de imágenes utilizando Escaneo Lineal . . .	46
5.2. Número de imágenes en el banco de imágenes utilizando Gramian Angular Field	47
5.3. Hiperparámetros generales y sus posibles valores	47
5.4. Hiperparámetros convolucionales y sus posibles valores	47
5.5. El Fresno	48
5.6. La Piedad	48
5.7. La Palma	48
5.8. Aristeo Mercado	48

Capítulo 1

Introducción

A lo largo de los años la energía mundial se ha generado utilizando diferentes recursos limitados. Ejemplos de estos recursos pueden ser el carbón natural, el petróleo, gas natural, entre otros combustibles fósiles. La producción de energía mediante medios fósiles afecta directamente a nuestro medioambiente y al planeta en el que vivimos.

La producción de energía ha sido uno de los principales problemas ambientales y económicos en el mundo. Aunado a esto, a medida que la demanda de energía aumenta, si ésta es producida por medio fósiles, el daño ambiental también aumenta (Rodríguez et al., 2017a).

Las energías renovables son cada vez más utilizadas debido a los alarmantes incrementos en la contaminación del medioambiente, ya que el producir este tipo de energía no genera daños considerables al medioambiente. Fenómenos como el calentamiento global, incendios forestales, sequías, entre otros son cada vez más comunes debido al daño que el ambiente esta sufriendo con el uso de energías generadas por medio fósiles. Entre las fuentes renovables más utilizadas se encuentran la energía eólica y solar debido a la facilidad y disponibilidad de estos recursos naturales. Además, la energía eólica es considerada una de las más atractivas debido a su alta eficiencia, baja contaminación y al ser la energía renovable de mayor crecimiento en los últimos años (Chang, 2013).

Actualmente, los precios bajos y la disponibilidad del petróleo ponen a la energía eólica en desventaja ante la economía mundial a pesar de las ventajas que puede ofrecernos. Sin embargo, el petróleo al ser un recurso no renovable no durará para siempre (Mohandes et al., 2004).

Con el aumento del uso de las energías renovables es cada vez más común que surjan

problemas relacionados a estas, como el conocer con antelación la demanda a satisfacer o la cantidad de materia prima a disponer, entre otros. Estas son cuestiones que permiten mejorar los costos de producción, la planificación en las plantas de energía y/o la administración de la planta. Por lo que, para aumentar el uso de estas es necesario dar soluciones óptimas a estas problemáticas.

La producción de energía eólica requiere como materia prima el viento, por lo que es necesario contar con un pronóstico acertado sobre el comportamiento del viento en las próximas horas, días o meses. De acuerdo a la literatura (Chang, 2014), existen 4 diferentes horizontes de predicción, el *Ultra corto* que va de un minuto a una hora, *corto* que consta de una hora a varias horas en adelante, *mediano* que va de muchas horas hasta una semana y *largo* que abarca de una semana a un año o más.

Con un pronóstico adecuado, las centrales eólicas podrán tomar decisiones más acertadas respecto a la cantidad de energía a generar para satisfacer la demanda, tiempos de trabajo de los aerogeneradores, mantenimientos de los aerogeneradores, entre otras actividades.

En este trabajo, se aborda el problema de la predicción a corto plazo de la velocidad del viento, en particular, 24 horas hacia adelante. Para esto, se utilizaron técnicas de aprendizaje profundo como las Redes Neuronales Convolucionales (RNC) para realizar el entrenamiento de los modelos predictivos de la velocidad del viento. A su vez se utilizó un Algoritmo Genético Compacto (AGc) para realizar la optimización de los hiperparámetros de la RNC.

El proceso de entrenamiento que consta de el AGc y la RNC es ejecutado 24 (horas a futuro) veces con el propósito de generar un banco de modelos predictivos de la velocidad del viento. Este banco de modelo permite obtener un pronóstico más preciso de la velocidad del viento debido a que cada uno de los modelos es entrenado y optimizado de manera separada.

La metodología propuesta permite saber con antelación cual será la velocidad del viento en las siguientes 24 horas, permitiendo conocer la cantidad de materia prima a disponer por parte de la compañía generadora de energía, mediante la implementación del esquema de Banco de Modelos (Lopez Farias et al., 2018), permitiendo conocer a más de un paso (hora) el pronóstico.

1.1. Definición del Problema

La demanda mundial de energía es cada vez mayor debido al crecimiento de la población, la industrialización de procesos, y a que cada vez se es más dependiente de la energía eléctrica. Esto no quiere decir que esta dependencia sea el problema si no que actualmente un 75 % (aproximadamente) de la energía generada es sucia o producida por medios fósiles. Debido a esto, existe un interés de sustituir la producción de energía generada por medios fósiles (sucia) por las llamadas energías limpias. La energía eólica es una de las energías renovables (limpias) más avanzadas y comercialmente más económicas (Kirbas and Kerem, 2016). Además, entre los diferentes tipos de energía renovable es una de las más utilizadas por la industria debido a la facilidad para obtenerla.

Actualmente, los costos de producción de energía limpia no son equiparables a los costos de la energía generada por medios fósiles (Observ, 2013), esto debido a los diversos retos que han surgido. Dentro de los retos principales, se encuentra el conocer con antelación la cantidad de materia prima que se va a disponer. Permitiendo saber de manera anticipada si la planta generadora de energía podrá satisfacer la demanda. Dado el caso de no satisfacerla, buscar los mecanismos para lograrlo. Es por eso que para ayudar a que exista una adopción más rápida de la energía eólica surge la necesidad de predecir el comportamiento de la velocidad del viento.

1.2. Objetivo

Este trabajo tiene como objetivo obtener modelos predictivos de la velocidad del viento a corto plazo mediante la técnica de aprendizaje profundo llamada Redes Neuronales Convolucionales que permitan conocer con antelación cual será el comportamiento del viento (velocidad) en las próximas 24 horas.

1.3. Objetivos Particulares

1. Investigar y analizar el estado del arte relacionado con pronóstico de la velocidad del viento a corto plazo utilizando técnicas de aprendizaje profundo.
2. Diseñar el proceso para obtener el pronóstico de la velocidad del viento a corto plazo, el cual tendrá un enfoque multi-modelo convolucional evolutivo.
3. Implementar una RNC en una tarjeta gráfica GPU.
4. Optimizar los hiperparámetros de la RNC utilizando un AGc.
5. Obtener bancos de modelos para 4 diferentes series de tiempo de la velocidad de viento.
6. Comparar los resultados obtenidos mediante el uso de la técnica de aprendizaje profundo llamada RNC y la técnica de clasificación y predicción Naïve.

1.4. Justificación

La obtención de un banco de modelos predictivos de la velocidad del viento será de gran ayuda para las plantas generadoras de energía eólica en su planeación estratégica, toma de decisiones, optimización en la generación de energía y mantenimiento de los aerogeneradores. A su vez, se sabrá cual sera el comportamiento del viento (velocidad) en las siguientes 24 horas, permitiendo conocer con antelación con cuanta materia prima se contará en las próximas 24 horas, pudiendo realizar una estimación anticipada de la energía que se necesitará generar para satisfacer la demanda. Además, saber con anticipación los costos de generar dicha cantidad de energía, entre otras cosas.

En la actualidad a pesar de que cada vez son más utilizadas las energías renovables aun es poco su porcentaje de uso (Lund, 2007). Las energías renovables son consideradas un importante suplemento de energía en muchos países alrededor del mundo pero en una escala global menos del 15 % de la energía es renovable al año 2005. Para el año 2015 el porcentaje aumentó a un 19.3 % y su capacidad y producción continua creciendo para 2016 (Sawin et al., 2017).

De acuerdo con el censo conducido por Observ'ER & Foundation Énergies pour le Monde en 2012, la producción dominante de energía fue la generada por medio de combustibles fósiles. La energía producida por estos medios cubre un 68 % del total de la producción de energía (Observ, 2013). Este tipo de energías, además de que son no renovables, al consumirla, generan daños irreversibles al medioambiente.

Un mecanismo que busca contrarrestar el daño ambiental y reducir el costo de la producción de energía es el uso de energías renovables, también llamadas energías limpias. Existe un amplio rango de opciones de energías renovables como: energía solar, viento, biomasa, geotérmica, desperdicios no renovables, hidrógeno, entre otros. De los anteriores la energía solar y de viento han presentado un crecimiento anual entre 2002 y 2012 del 50 % y 26.1 % respectivamente (Observ, 2013).

Tomando en cuenta el crecimiento que han tenido las energías renovables en los últimos años, muchos retos han surgido, como lo son el conocer la demanda, reducir los precios de producción de la energía, mejorar la planeación de las plantas de energía, etc. Para lograr una adopción completa de este tipo de energías es necesario encontrar soluciones efectivas para estos retos. Una de las más importantes es la necesidad de realizar predicciones de las variables involucradas en la producción de energía. Esto es, tener la certeza de la cantidad de energía que será producida en los próximos minutos, horas, días o meses. Algunas de las variables involucradas en la producción de energía limpia pueden ser las siguientes: flujo de agua, mareas de los océanos, el calor interno de la tierra, vapor producido por géiseres, etc (Rodriguez et al., 2017b).

Un estimado reciente menciona que para el año 2030 un pronóstico perfecto será evaluado en 3 billones de dólares anualmente (Piwko and Jordan, 2010). Para el sistema de energía de los Estados Unidos la estimación fue realizada por el departamento de energía (Barber et al., 2010). De esta manera el pronóstico de la velocidad del viento tiene un rol importante en el suministro de energía.

La Figura 1.1 muestra el crecimiento anual promedio de distintos tipos de energía renovables. La energía eólica es la segunda con mayor crecimiento solo después de la energía solar.

Esta claro que las energías renovables son cada año más utilizadas pero su adopción es

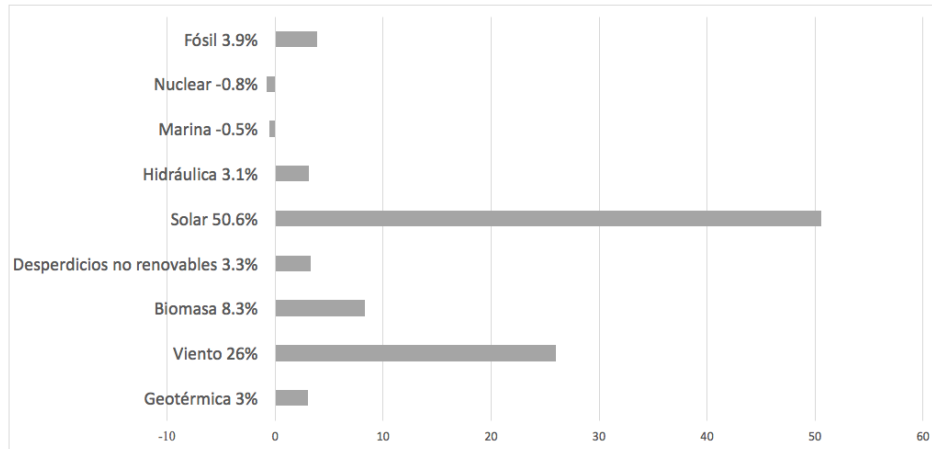


Figura 1.1: Crecimiento anual de energías renovables 2002-2012 (Observ, 2013)

lenta aun. Por lo cual este trabajo tiene como objetivo el mejorar la estimación de la velocidad del viento, provocando la reducción de los costos de producción como de mantenimiento de la energía eólica.

1.5. Estructura de la Tesis

Este trabajo está dividido en 6 capítulos y se encuentra organizado de la siguiente manera. En el capítulo 2 se describe la teoría fundamental. En el capítulo 3 se presenta la literatura actual relacionada con este trabajo. En el capítulo 4 se muestra la metodología. En el capítulo 5 se muestran los resultados obtenidos y por último las conclusiones son presentadas en el capítulo 6.

Capítulo 2

Marco teórico

En este capítulo se abordarán conceptos utilizados en el desarrollo de este documento. Esto con la intención de proporcionar al lector más información acerca del tema desarrollado.

Se partirá de conceptos básicos relacionados a la energía con el propósito de conocer los distintos tipos que existen, como las energías renovables o energías limpias y energías no renovables o sucias. Posteriormente se abordarán conceptos relacionados con el pronóstico de la velocidad del viento. Por último serán definidos conceptos asociados con la inteligencia artificial como los distintos métodos utilizados, medidas de aptitud y conceptos relacionados con aprendizaje máquina y aprendizaje profundo.

2.1. Energía

La energía es un ingrediente indispensable para el desarrollo económico de un país. Su importancia incrementa en relación a la industrialización y la necesidad de modernización. El incremento en su producción se traslada en una mejor calidad de vida y la creación de riqueza (Observ, 2013). La energía es definida como la capacidad de un cuerpo para realizar o producir un trabajo, ya sea en forma de movimiento, luz, etc. La energía no es un estado físico real, sino sólo una magnitud escalar que se le asigna al cuerpo, es decir que es una herramienta o abstracción matemática de una propiedad de los sistemas físicos (en este caso, su capacidad de efectuar un trabajo) (Gutiérrez and Calderón, 2013). Existe una amplia variedad de métodos para la creación de energía y cada uno de estos tiene sus ventajas y desventajas. Existen las energías sucias o energías creadas a través de medios fósiles que a pesar de ser

las más utilizadas son las que más impacto tienen sobre el medioambiente de nuestro planeta. También existen las energías renovables o limpias, las cuales son menos utilizadas por el costo que supone su adopción pero tienen una gran ventaja al no impactar el medioambiente.

2.1.1. Energías Sucias

Las energías sucias son todas aquellas que causan daños al medioambiente de nuestro planeta. Estas energías provienen de medios fósiles como el carbón, el petróleo o el gas natural. La quema de combustibles fósiles causa los llamados gases de invernadero derivando en cambios climáticos. Está claro que la quema de combustible fósil tiene un impacto sobre el medioambiente, pero hay combustibles que son más dañinos que otros. Es importante dejar de utilizar este tipo de recursos no renovables y utilizar energías limpias o en su defecto utilizar los combustibles que causen un menor daño en el medioambiente como el gas natural, ya que según (Greenpeace, 2015) una tercera parte de las emisiones de dióxido de carbono (CO_2) en el mundo proviene de la quema de carbón natural.

2.1.2. Energía Renovable

De acuerdo con (Casas et al., 2007) se les llama energías renovables a las obtenidas de fuentes naturales inagotables. Unas porque tienen un gran potencial para extraer energía de ellas y otras son capaces de generarse por medios naturales, i.e. la generada por medio de biomásas, además estas son fuentes de energía que se llevan bien con el medioambiente.

La Tabla 2.1 muestra algunas de las opciones más conocidas de energías renovables.

Tabla 2.1: Tipos de energías renovables

Tipo	Descripción
Hidráulica	Producida por corrientes de agua en centrales hidroeléctricas
Eólica	Producida por viento a través de aerogeneradores
Solar	Proporcionada por el sol en forma de radiación electromagnética
Geotérmica	Obtenida del calor almacenado en la superficie terrestre
Mareomotriz	Producida por el movimiento de las mareas y corrientes marinas
Biomasa	Producida a partir materia orgánica animal o vegetal

Existe una amplia variedad de tipos de energías renovables y distintos métodos para su producción. Este proyecto solo tiene relación con la energía eólica al tratarse de la obtención de pronóstico de la velocidad del viento a corto plazo. A continuación se presenta una breve descripción de lo que es la energía eólica.

2.1.3. Energía Eólica

Se le llama energía eólica a la generada por aerogeneradores, los cuales utilizan corrientes de viento para producirla. Este tipo de energía es una excelente alternativa a las producidas por medios fósiles debido a que es energía limpia, renovable y no existe emisión de gases durante su operación que produzcan un efecto invernadero o dañe el medioambiente. Además de la ventaja de no generar un daño al medioambiente, otra ventaja es que los parques eólicos no resultan tan invasivos para las personas ya que las tierras donde se coloquen también pueden ser utilizadas para la agricultura o parques recreativos. Además, estos parques eólicos también pueden ser colocados en el océano (Fthenakis and Kim, 2009).

2.2. Pronóstico

Cuando se habla de pronóstico generalmente se refiere al proceso de realizar una predicción o estimación. Durante muchos años se ha tenido la necesidad de contar con un pronóstico o predicción de distintas variables como el clima, ventas en una empresa, resultados de algún evento deportivo, etc. Para las empresas el contar con una predicción acertada es de mucha importancia debido a que se puede reducir drásticamente la incertidumbre dentro de la cual se toman decisiones importantes para el rumbo de la empresa. A pesar de esto el pronóstico no sustituye el juicio de un experto, es simplemente un apoyo para la toma de decisiones.

El pronóstico es un problema importante que abarca muchos campos incluyendo negocios, industria, gobierno, economía, ciencias ambientales, medicina, ciencias sociales, política, entre otros. En la literatura no existe un estándar para clasificar los horizontes de predicción. De acuerdo con (Chang, 2014) el horizonte de predicción de la velocidad del viento se puede clasificar de 4 diferentes maneras ultra corto, corto, mediano y largo plazo. Donde los

rangos para cada uno de los plazos se describe en la siguiente Tabla 2.2.

Tabla 2.2: Clasificación de las escalas de tiempo en pronóstico del viento

Escala de Tiempo	Rango
Ultra corto	De pocos minutos a una hora
Corto	De una hora a unas horas
Mediano	De horas hasta una semana
Largo	De una semana a un año

El pronóstico a corto plazo es requerido para actividades regulatorias, pronóstico a corto y mediano plazo son requeridos para actividades como manejo de operación para realizar presupuestos, seleccionar nuevas investigaciones y desarrollo de proyectos. Pronóstico a largo plazo impacta en temas como la planeación estratégica (Montgomery et al., 2015).

Las predicciones se realizan utilizando información recopilada de eventos pasados y/o información actual. Para la obtención del pronóstico se utilizaron observaciones pasadas de la velocidad del viento recopilada en forma de series de tiempo.

2.2.1. Serie de Tiempo

Una serie de tiempo es un conjunto de observaciones obtenidas en un tiempo específico (Brockwell and Davis, 2013). Se puede ver a las series de tiempo como un modo estructurado de representar datos. Por ejemplo, las ventas diarias de un producto pueden representarse como una serie de tiempo. Existen distintos tipos de series de tiempo como lo son:

1. Serie de tiempo de valores discretos, por ejemplo, 1, 2, 3, A, B, C.
2. Serie de tiempo de valores continuos, por ejemplo, 0.01, 0.02, 0.03.

Muchos conjuntos de datos se encuentran representados como series de tiempo. Ejemplos de series de tiempo son una secuencia mensual de la calidad de las entregas en una fábrica, una serie semanal del número de accidentes automovilísticos, cantidad de lluvia diaria, etc.

La tasa de tiempo con la que se colectan las observaciones en una serie de tiempo tiene que ser del mismo espaciado para cada observación, lo cual se realiza así típicamente en la mayoría de las aplicaciones de pronóstico. La tasa de tiempo de recolección de observaciones

puede ser diaria, semanal, quincenal e incluso anual (Montgomery et al., 2015). En campos como economía, negocios, ingeniería, ciencias naturales; Hay muchos ejemplos de series de tiempo (Box et al., 2015).

Un ejemplo de una serie de tiempo puede observarse en la Figura 2.1. donde se muestra una gráfica de un conjunto de puntos ordenados que tienen una magnitud. Siendo el eje de las y donde se representa la magnitud de la observación (velocidad del viento), y el eje x para representar el tiempo que se tomó la observación. Formando así una serie de tiempo de la velocidad del viento con una separación entre observación de una hora.

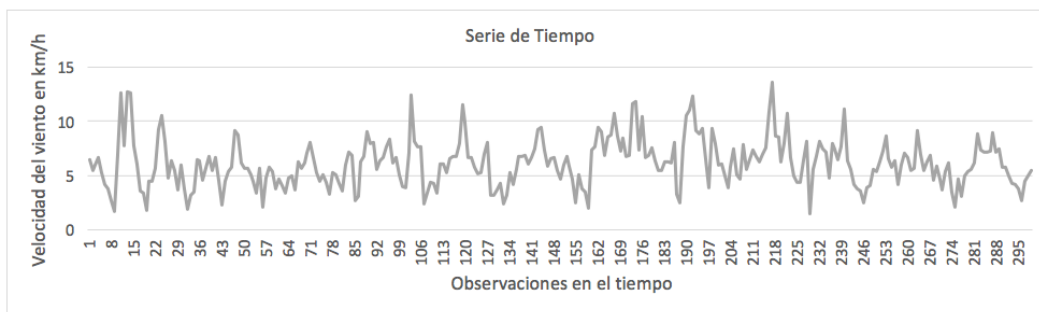


Figura 2.1: Representación gráfica de una serie de tiempo de la velocidad del viento

2.3. Pronóstico Mediante Inteligencia Artificial

Desde la antigüedad los filósofos han tratado de entender y obtener la respuesta a dos grandes preguntas, ¿cómo es que funciona la mente? y ¿las máquinas pueden tener la capacidad de pensar o tener inteligencia?, sin embargo éstas son preguntas que difícilmente pueden contestarse. Algunos filósofos han tomado el enfoque computacional y aceptado la idea de que las máquinas pueden hacer todo lo que el humano puede hacer. Otros han negado rotundamente y dicho que comportamientos tan sofisticados como el amor y la moral siempre estarán más allá del alcance de las máquinas. Entonces ¿que significa inteligencia?, algunos lo definen como :

1. La habilidad de entender y aprender cosas
2. La habilidad de pensar y entender en lugar de hacer cosas solo por instinto

Por lo que de acuerdo con (Negnevitsky, 2002) se puede definir la inteligencia como “La habilidad de aprender y entender para resolver problemas y tomar decisiones”. La meta de la Inteligencia Artificial (IA) como ciencia es hacer que las máquinas realicen cosas que podrían requerir inteligencia si éstas fueran hechas por humanos.

En lo que respecta a pronóstico con inteligencia artificial, existen distintas técnicas o métodos dentro del área de aprendizaje máquina con los cuales se pueden obtener excelentes resultados. Dentro de las técnicas que se pueden utilizar están las Redes Neuronales Artificiales (Cadenas and Rivera, 2010), Máquinas de Soporte Vectorial (Du et al., 2008), k-Vecinos cercanos (Yesilbudak et al., 2013), entre otros.

2.3.1. Aprendizaje Máquina

El aprendizaje máquina actualmente se encuentra en muchos aspectos de la sociedad, desde buscadores web para filtrar contenido así como recomendaciones en páginas de comercio electrónico (e-commerce), también está presente en productos tales como cámaras y Smartphones. Los sistemas con Aprendizaje Máquina (Machine Learning) son usados para identificar objetos en imágenes, pasar el habla a texto y seleccionar resultados relevantes en búsquedas (LeCun et al., 2015). Además, el aprendizaje máquina permite dotar a un equipo de cómputo de la habilidad de aprender y mejorarse conforme a ejemplos que se le proporcione sin la necesidad de realizar una programación explícita de este conocimiento. El proceso de aprendizaje consta de dar a la máquina una serie de ejemplos o datos históricos para que esta pueda detectar patrones en los datos que le permita tomar una mejor decisión con los nuevos datos. Existen diferentes tipos de aprendizaje máquina como lo son :

1. **Aprendizaje Supervisado:** En el aprendizaje supervisado se tiene una serie de observaciones asociadas a un valor esperado, en otras palabras se le dice a la máquina que es lo que debe aprender.
2. **Aprendizaje No supervisado:** Al contrario del aprendizaje supervisado, en el aprendizaje no supervisado no hay valor esperado asociado, se realiza una búsqueda de cuáles son los grupos que mejor se apeguen a las características que buscamos, en otras palabras se realiza una agrupación y no una clasificación (Harrington, 2012).

En términos generales, un sistema automático debe ser capaz de proporcionar una respuesta apropiada cuando se introduce información al mismo. En el aprendizaje supervisado, el núcleo del sistema es un modelo predictivo que asigna uno o varios valores de salida a cada elemento de entrada, con base al conocimiento adquirido a partir de un conjunto de datos cuyas salidas son conocidas. Sin embargo, la información disponible no siempre se puede o se debe utilizar directamente para alimentar al modelo. Por ello, suele ser necesaria una fase denominada de pre-procesamiento para adecuar las entradas en aras de maximizar el rendimiento del modelo. Asimismo, la salida que ofrece el modelo puede no ser satisfactoria o quizás resulte difícil de interpretar. Por lo tanto, en una última fase de post-procesamiento, se evalúa el rendimiento del modelo y, en la medida de lo posible, se adaptan las salidas a las necesidades del usuario (Sampedro and García, 2012).

2.3.1.1. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) son metodologías computacionales inspiradas en una red de neuronas biológica, estas pueden contener capas de nodos que realizan operaciones simples, estos nodos están altamente interconectados y cada conexión cuenta con un peso, el cual es ajustado cuando datos son presentados a la red mediante un proceso de entrenamiento. Un entrenamiento correcto puede resultar en un modelo de RNA que pueda realizar tareas como predecir un valor de salida, clasificar un objeto, aproximar una función y completar un patrón conocido, entre otras (Dayhoff and DeLeo, 2001).

La unidad de procesamiento fundamental en una RNA es la neurona. Básicamente una neurona biológica recibe entradas de otras fuentes, las combina y realiza generalmente operaciones no lineales y emite un resultado final (Anderson and McNeill, 1992). La Figura 2.2 muestra la representación de una neurona artificial. En la cual cada entrada a la red es representada por x_i , cada una de las entradas es multiplicada por un peso; Estos pesos son representados por w_{ij} . En el caso más simple, este producto son simplemente sumados. Pasados a una función de activación o transferencia para generar un resultado y/o una salida.

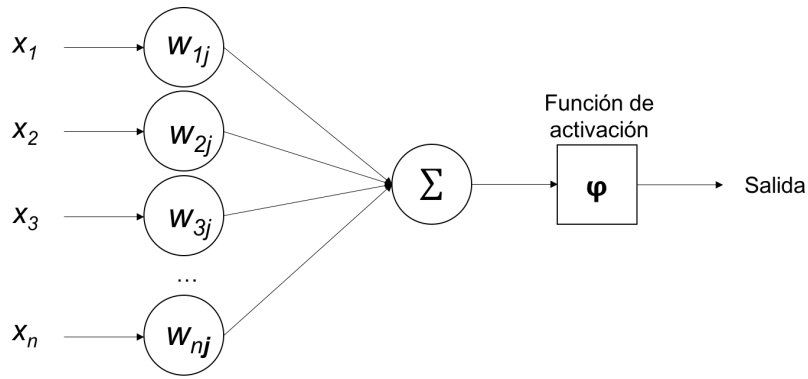


Figura 2.2: Neurona artificial básica (Anderson and McNeill, 1992)

En la Figura 2.3 se muestra la arquitectura clásica de una red neuronal, la cual consta de 3 capas (capa de entrada, capa oculta y capa de salida). En la capa de entrada no se realiza cómputo alguno, esta solo sirve como entrada a la red neuronal. En la capa oculta es donde se realiza el mayor cálculo y la capa de salida provee la salida de la red. Una RNA como un aproximador universal puede aprender cualquier función dadas suficientes neuronas (Haykin, 1994).

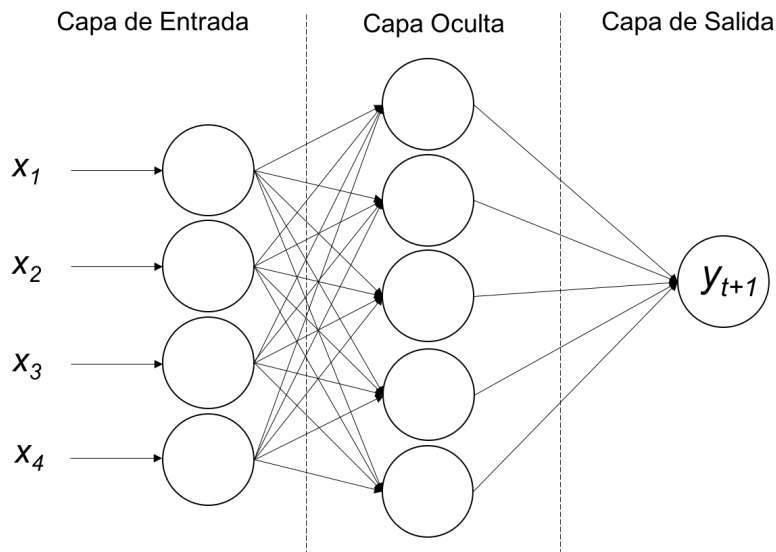


Figura 2.3: Arquitectura de una RNA con 4 neuronas en la capa de entrada, 5 neuronas en la capas ocultas y una sola neurona de salida .

2.3.2. Aprendizaje Profundo

El aprendizaje profundo permite a los modelos computacionales que están compuestos de múltiples capas de procesamiento aprender de los datos con múltiples niveles de abstracción. Estos métodos han mejorado el estado del arte dramáticamente en el reconocimiento del habla, reconocimiento de objetos y en muchos otros dominios tales como el descubrimiento de drogas. El aprendizaje profundo descubre estructuras intrínsecas en un conjunto de datos amplio usando el algoritmo de propagación hacia atrás, de acuerdo con (LeCun et al., 2015). Dentro del aprendizaje profundo es posible encontrar una variedad de diferentes técnicas, entre las cuales se encuentran las Redes Neuronales Recurrentes específicamente las LSTM, Autoencoders, Redes Neuronales Convolucionales, entre otros (LeCun et al., 2015).

2.3.2.1. Redes Neuronales Convolucionales

Las Redes Neuronales Convolucionales (RNC) son un modelo jerárquico cuya arquitectura consiste en capas de convolución y sub-muestreo. La principal ventaja de las RNC es que estas pueden capturar características locales dentro de filtros y pesos compartidos (Yu et al., 2016). Las RNC son un tipo de RNA inspirada por la estructura del sistema visual, éstas en lugar de tener una pila de capas de matrices que se multiplican cuentan con capas de convolución. Actualmente, los sistemas de reconocimiento de patrones basados en RNC están entre los sistemas con mejor rendimiento. Las RNC usualmente están estructuradas por un conjunto de capas de procesamiento. Al final de la estructura está la capa completamente conectada la cual es una Red Neuronal Artificial y por último un clasificador, esto de acuerdo a (Pérez-Espinosa et al., 2016).

La Figura 2.4 muestra la arquitectura que usualmente se usa para las RNC. En la arquitectura son representadas en forma de flujo las diferentes capas de la RNC. En ella se muestran seis diferentes capas de procesamiento como lo son la capa de convolución, la capa activación, la capa de agrupación o pooling, la capa de aplanamiento o Reshape, una capa de RNA completamente conectada y la capa de reducción de sobre entrenamiento o dropout. Al principio cuenta con la capa de entrada la cual solo sirve como entrada de los datos y en ella no se realiza ningún tipo de procesamiento.

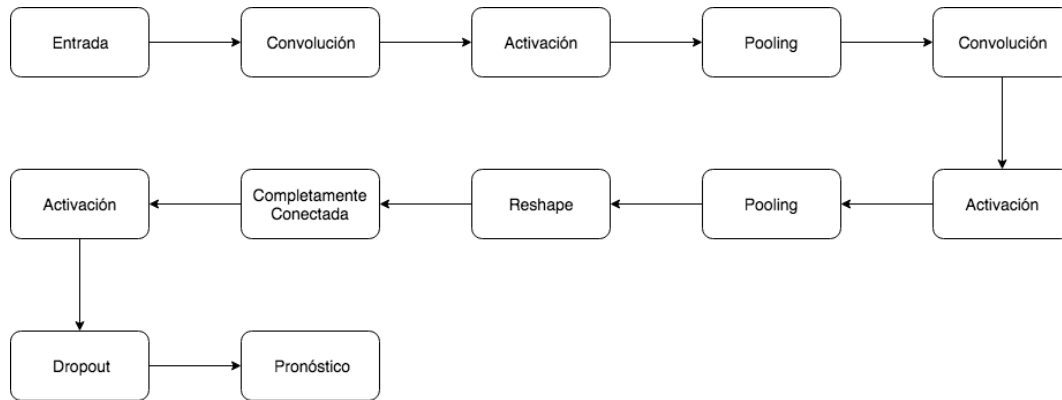


Figura 2.4: Arquitectura general de una Red Neuronal Convolutiva. (Pérez-Espinosa et al., 2016)

Las RNC cuentan con una amplia variedad de parámetros que pueden ser optimizados, los cuales se encuentran a través de todas sus capas. Estos parámetros pueden ir desde el algoritmo de entrenamiento, número de épocas utilizadas para el entrenamiento, la función de activación, entre otros.

2.4. Algoritmo Genético

La computación evolutiva simula la teoría de la evolución en una computadora, esta simulación puede ser realizada utilizando una serie de algoritmos de optimización. Esta optimización iterativamente mejora la calidad de las soluciones hasta que se encuentra una solución óptima. El enfoque evolutivo está basado en modelos de la selección natural y la genética. Estos modelos son basados en computación evolutiva y esta formada por una amplia gama de términos combinando algoritmos genéticos, estrategias evolutivas y programación genética. Todas estas técnicas simulan la evolución usando los procesos de selección, mutación y reproducción de acuerdo con (Negnevitsky, 2002).

Los algoritmos genéticos son una clase algoritmos de búsqueda estocástica basados en la evolución biológica. Esta es la técnica más popular en la computación evolutiva. la representación del problema utilizada (Cromosoma) es una cadena de bits de longitud fija, la cual es mostrada en la Figura 2.5. Cada posición en la cadena representa una característica de un individuo (Gen). Por lo que, el cromosoma representa una propuesta de solución al problema dado.

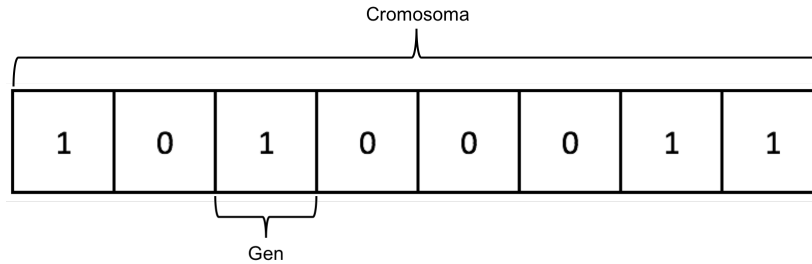


Figura 2.5: Representación del cromosoma (Sivanandam and Deepa, 2007)

El cromosoma está constituido por genes, los cuales son representados por bits con valores en un rango de $[0,1]$. Estos pueden ser vistos como una analogía de los genes de un organismo biológico. Las operaciones principales realizadas en la cadena de bit son el cruzamiento (crossover) y mutación. En el cruzamiento dos cadenas son usadas como padres y nuevos individuos son formados (Sivanandam and Deepa, 2007).

2.4.1. Algoritmo Genético Compacto

El Algoritmo Genético Compacto (AGC) es una técnica de búsqueda/optimización probabilística. Está relacionado con el algoritmo genético y otros algoritmos evolutivos que están inspirados en la teoría de la evolución por medio de la selección natural. El objetivo AGC es simular el comportamiento del algoritmo genético con mucho menos requerimiento de memoria (Sin requerir que una población sea mantenida en memoria). Esto se logra manteniendo un vector de probabilidades en lugar de la población entera. Las soluciones candidatas son probabilísticamente generadas desde la cadena de características. Las características que proveen una mejor solución se utilizan para realizar pequeños cambios en un vector de probabilidades (Brownlee, 2011). El Algoritmo 2.1 muestra el Algoritmo Genético Compacto (Harik et al., 1999).

Algoritmo 2.1 Algoritmo Genético Compacto.

```

1: procedure AGC( $n, l$ )      ▷  $n$  como tamaño de población,  $l$  longitud del cromosoma
2:   Inicializar el vector de probabilidad
3:   for  $i := 1$  to  $l$  do  $p[i] := 0.5$ ; do
4:     Generar 2 individuos desde el vector
5:      $a := \text{generate}(p)$ ;
6:      $b := \text{generate}(p)$ ;
7:      $\text{Evaluate}(a)$ ;
8:      $\text{Evaluate}(b)$ ;
9:      $\text{winner}, \text{loser} := \text{compete}(a, b)$ ;
10:    //Actualizar el vector utilizando el mejor
11:    for  $i := 1$  to  $l$  do
12:      if  $\text{winner}[i] \neq \text{loser}[i]$  then
13:        if  $\text{winner}[i] = 1$  then
14:           $p[i] := p[i] + 1/n$ 
15:        else
16:           $p[i] := p[i] - 1/n$ ;
17:        end if
18:      end if
19:    end for
20:    // Revisar si el vector a convergido
21:    for  $i := 1$  to  $l$  do
22:      if  $p[i] > 0$  &  $p[i] < 1$  then
23:        return to step 4;
24:      end if
25:    end for
26:  end for
27:  return  $p$ 
28: end procedure

```

El Algoritmo genético inicia definiendo el vector de probabilidad en la línea 2. Enseguida, desde ese vector dos individuos se generan y evalúan. El proceso de evaluación de los individuos prueba los hiperparámetros propuestos por el AGc. Los individuos generados compiten y de acuerdo al ganador el vector de probabilidad es actualizado. Este proceso se repite hasta que el criterio de convergencia se cumple.

2.5. Medidas de Aptitud

Se utilizaron diferentes medidas de aptitud como lo son el Error Cuadrático Medio (MSE por sus siglas en inglés), Error Medio Absoluto (MAE por sus siglas en inglés), la Raíz

Cuadrada del Error Cuadrático Medio (RMSE por sus siglas en inglés) y el Error Porcentual Absoluto Medio (MAPE por sus siglas en inglés). Estas medidas de aptitud permiten saber que tan alejado se encuentra el valor pronosticado \hat{y} de los datos reales y . La definición de estas medidas estadísticas son :

$$MSE = \frac{1}{N} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (2.1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (2.2)$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{\hat{y}_i} \quad (2.3)$$

$$RMSE = \sqrt{MSE} \quad (2.4)$$

Capítulo 3

Estado del Arte

Se investigó en la literatura distintos métodos de pronóstico de la velocidad del viento a corto plazo. Se consultaron trabajos sobre series tiempo, pronóstico de series de tiempo utilizando técnicas de aprendizaje máquina y técnicas de aprendizaje profundo, optimización de hiperparámetros, entre otros.

3.1. Métodos de Pronóstico

En el Survey de (Chang, 2014) se presentan los métodos más utilizados con respecto al pronóstico de la velocidad del viento. Además, se categorizan diferentes enfoques que lidian con el problema de pronóstico estableciendo los métodos por persistencia, físicos, estadísticos, de correlación espacial, de inteligencia artificial e híbridos.

Los métodos de persistencia (Naïve) son métodos de pronóstico básicos donde $Y_{t+\Delta} = Y_t$ (Zhao et al., 2011). Estos métodos son particularmente usados como un método base y son comúnmente utilizados en una escala de tiempo ultra corto donde estos son más exactos que otros métodos de pronóstico (Chang, 2014). Los métodos físicos son predicciones numéricas del clima hechas por meteorólogos para predicciones de clima a gran escala (Lei et al., 2009). Estos métodos físicos necesitan ser ejecutados en súper computadoras ya que estos requieren de mucho poder de cómputo (Chang, 2014). Por otro lado, los métodos estadísticos usan datos históricos para encontrar relaciones dentro de las series de tiempo. Algunos de los métodos estadísticos son los mostrados en la tabla 3.1 (Flores et al., 2009).

Tabla 3.1: Métodos Estadísticos

1	Auto Regressive (AR)
2	Auto Regressive Moving Average (ARMA)
3	Regressive Integrate Moving Average (ARIMA)

Estos métodos estadísticos son buenos para periodos de tiempos cortos, y pueden ser utilizados para resolver problemas de ingeniería, economía y ciencias naturales. Ling-ling et al., 2011 presentan un modelo predictivo ARMA de la velocidad del viento que a su vez utiliza la transformada Wevelet. Esta transformada Wavelet es utilizada para captar los vientos de baja frecuencia, mientras que el modelo ARMA es utilizado para realizar el pronóstico de la velocidad del viento con los datos captados. Palomares-Salas et al., 2009 utilizan un modelo ARIMA para realizar el pronóstico de la velocidad del viento. El presenta un proceso de validación del modelo junto con un análisis de regresión. Miranda and Dunn, 2006 se centran en la predicción de la velocidad del viento un paso adelante (una hora en el futuro), utilizando un enfoque Bayesiano. Los resultados de este trabajo presentan que el enfoque bayesiano puede ser útil tanto para el pronóstico de la velocidad del viento como para el pronóstico de la fuerza del viento.

3.2. Pronóstico Mediante Inteligencia Artificial

Se revizó literatura sobre métodos de aprendizaje máquina y aprendizaje profundo para la obtención de un modelo predictivo. Con respecto a aprendizaje máquina podemos encontrar distintos métodos como las Redes Neuronales, Máquinas de Soporte Vectorial, K-Vecinos más Cercanos, Sistemas Difusos, entre otros. Con respecto a aprendizaje profundo en la literatura podemos encontrar métodos cómo las Redes Neuronales Recurrentes (LSTM), Redes Neuronales Convolucionales, entre otros.

3.2.1. Aprendizaje Máquina

Cadenas and Rivera, en 2007 realizaron una comparación entre dos técnicas, ARIMA y Redes Neuronales Artificiales para el pronóstico de la velocidad del viento con datos de la

costa sur del estado de Oaxaca. Ambas técnicas fueron aplicadas con datos correspondientes a 7 años de observaciones de la velocidad del viento, donde se utilizaron 6 años para el proceso de entrenamiento y el último año para la validación de los modelos. Se llega a la conclusión que el modelo ARIMA presentan mejores resultados con respecto al caso de uso presentado, sin embargo ambas técnicas proporcionan una manera viable de obtener un pronóstico de la velocidad del viento acertada.

Rodríguez et al., en 2017 presentaron un trabajo el cual tiene la peculiaridad de que las series de tiempo de la velocidad del viento utilizadas tienen datos faltantes, los cuales pueden deberse a diferentes problemas durante la recolección de los datos como el error humano, problema de los sensores, etc.. El trabajo está enfocado en realizar la reconstrucción de las series de tiempo de la velocidad de viento. Esta reconstrucción se realizó utilizando dos técnicas de Aprendizaje Máquina, k vecinos más cercanos y redes neuronales artificiales. Para obtener el mejor modelo para la reconstrucción de las series de tiempo, con ambas técnicas se realizó una optimización de los distintos parámetros. Para la Red Neuronal Artificial se realizó una búsqueda estocástica de la topología con la que se obtuviera un modelo que mejor se apegara a los datos de entrenamiento. Con k vecinos más cercanos se realizó una búsqueda de fuerza bruta de los mejores parámetros como lo son el tamaño del vector de características, el número de vecinos cercanos, entre otros. De las series reconstruidas se obtuvieron modelos predictivos con ambas técnicas, obteniendo mejores resultados con k vecinos más cercanos.

En el trabajo Khashei et al., 2008 se propone un enfoque híbrido el cual combina los modelos de regresión, difuso y RNA. Se utilizaron con datos incompletos con los cuales los modelos difusos son capaces de trabajar, obteniendo buenos resultados. En este trabajo no se realiza pronóstico de la velocidad del viento si no que el modelo propuesto se aplica a series de tiempo de mercados financieros.

3.2.2. Aprendizaje Profundo

Dentro de la inteligencia artificial se puede encontrar una serie de métodos que pertenecen a una sub-rama del aprendizaje máquina llamado aprendizaje profundo con el que también es posible obtener un pronóstico de la velocidad del viento. Sergio and Ludermir, 2015 presentan un trabajo en el que se ataca el problema del pronóstico de la velocidad del viento,

específicamente la velocidad promedio del viento capturada cada hora en una región del Noroeste de Brasil. En el realiza una breve explicación de las técnicas utilizadas, las cuales son las Redes de Creencia Profunda (Deep Belief Network). Se utilizaron dos enfoques de pre-entrenamiento para el desarrollo del trabajo. En el primer caso cada capa de la red utilizada fue considerada a ser un Autoencoder agregando ruido a los datos de entrada, representando el modelo conocido como Stacked Denoising Autoencoders(SADE por sus siglas en inglés). En el segundo enfoque cada capa de la red fue una Máquina Restringida de Boltzmann, representando el modelo conocido como Deep Belief Network(DBN por sus siglas en inglés).

Otra técnica que pertenece al aprendizaje profundo y la que fue utilizada para el desarrollo de este proyecto son las Redes Neuronales Convolucionales. A la fecha en la literatura no se encuentra mucho material sobre esta técnica aplicada al problema específico de la predicción de la velocidad del viento pero se pueden encontrar trabajos como el presentado por (Yang et al., 2015) en el cual se utiliza las RNC para obtener un pronóstico, utilizando series de tiempo multi-canal de la actividad humana. En el se propone un sistema de selección de características eficiente y una manera de realizar una clasificación de la actividad humana. El método propuesto utiliza una Red Neuronal Convolutacional para realizar el aprendizaje de características automáticamente desde los datos de entrada. Se realizó una comparación del método propuesto con otras técnicas tanto de aprendizaje máquina como de aprendizaje profundo, entre ellas las Maquinas de Soporte Vectorial, 1-Vecino más cercano, Redes de Creencia Profunda, entre otros. De los dos casos de uso con los que se realizaron experimentos con las diferentes técnicas se obtuvieron mejores resultados con las Redes Neuronales Convolucionales.

Debido a que los datos con los que se contó eran series de tiempo de la velocidad del viento, en el proceso de entrenamiento se tuvo que realizar un proceso para la transformación de estos datos a formato de matriz, representando una imagen o estampa de una sección dada de la serie de tiempo para que los datos pudieran ser utilizados por la Red Neuronal Convolutacional. Para ello se utilizó una técnicas para transformar un subconjunto de la serie de tiempo en imagen y otra técnica para a partir de una serie de tiempo de la velocidad del viento completa, generar un banco de imágenes, el cual serviría para alimentar a la RNC y realizar el proceso de entrenamiento. El proceso de transformación de un segmento de la

serie de tiempo a imagen está basado en el artículo publicado por (Wang and Oates, 2015) en el se presentan dos técnicas distintas para transformar una serie de tiempo en imagen. Las técnicas son Gramian Angular Field(GAF) y Markov Transitions Fields(MTF). Las cuales permiten el uso de técnicas de visión computación para problemas de clasificación. Para realizar el proceso de entrenamiento (Wang and Oates, 2015) utilizaron una variante de las Redes Neuronales Convolucionales llamada Redes Neuronales Convolucionales en Mosaico (Tiled CNN). La tabla 3.2 resume algunos trabajos mostrados en esta sección de estado del arte.

Tabla 3.2: Tabla comparativa de los diferentes métodos de predicción relacionados a este trabajo de investigación

		Estado del Arte	
	Métodos	Descripción	
Aprendizaje Máquina	Redes Neuronales K - Vecinos más Cercanos	Cadenas y Rivera realizaron una comparación entre las técnicas: ARIMA y Redes Neuronales Artificiales para el pronóstico de la velocidad del viento	Ambas técnicas utilizadas fueron aplicadas con datos correspondientes a 7 años de observaciones de la velocidad del viento, donde se utilizaron los datos de 6 años para el proceso de entrenamiento y el último año para la validación de los modelos.
	Máquinas de Soporte Vectorial Sistemas Difusos	Rodríguez et al., presentaron un trabajo donde las series de tiempo de la velocidad del viento contienen datos faltantes	Este trabajo esta enfocado en realizar la reconstrucción de las series de tiempo de la velocidad de viento utilizando técnicas de aprendizaje Máquina, k vecinos más cercanos y redes neuronales artificiales.
Aprendizaje Profundo	Redes Neuronales Recurrentes	Khashei et al., proponen un enfoque híbrido el cual combina modelos de regresión difuso y RNA.	Este trabajo utiliza series de tiempo de mercados financieros.
		Sergio and Ludermi presentan un problema del pronóstico de la velocidad del viento, específicamente la velocidad promedio del viento capturada cada hora en una región del Noroeste de Brasil.	Este trabajo utiliza la técnica de Redes de Creencia Profunda (Deep Belief Network)
	Redes Neuronales Convolucionales	Wang and Oates presentan dos técnicas distintas para transformar una serie de tiempo en imagen	El uso de las técnicas de Gramian Angular Field(GAF) y Markov Transitions Fields(MTF) permiten el uso de técnicas de visión computación para problemas de clasificación

Capítulo 4

Diseño del experimento

En esta sección se presenta a detalle como se diseñó el enfoque Multi-modelo Convolutivo Evolutivo. A su vez se describen los distintos experimentos que se realizaron para la obtención de un banco de modelos para el pronóstico a corto plazo de la velocidad de viento.

La Figura 4.1 muestra el proceso completo para la obtención de un modelo predictivo a 1 paso hacia adelante de la velocidad de viento a corto plazo a partir de una serie de tiempo. El proceso completo consta de 3 sub-procesos principales, los cuales son *Pre-Procesamiento de datos*, *Generación del banco de imágenes* y *proceso de entrenamiento*. El Pre-Procesamiento de los datos toma como entrada la serie de tiempo y en él se realizan tareas como revisar por datos atípicos (Outliers) en la serie de tiempo, la normalización de los datos, entre otras cosas. Seguido del Pre-Procesamiento de los datos se realiza la generación del banco de imágenes necesarias para entrenar el modelo predictivo. En el se generan las imágenes y estas son la entrada del proceso de entrenamiento. El proceso de entrenamiento es el más extenso, debido a que en el se realiza la optimización de los hiperparámetros de la RNC utilizando un AGc. El algoritmo genético genera una serie de posibles soluciones que a medida que pasan las generaciones estas soluciones serán cada vez mejores. Cada individuo del AGc es una posible solución a nuestro problema, en este caso la serie hiperparámetros para RNC. Con esta serie de hiperparámetros la RNC es entrenada retornando al AGc la aptitud del modelo entrenado. El AGc realiza la optimización de los parámetros obteniendo un modelo óptimo.

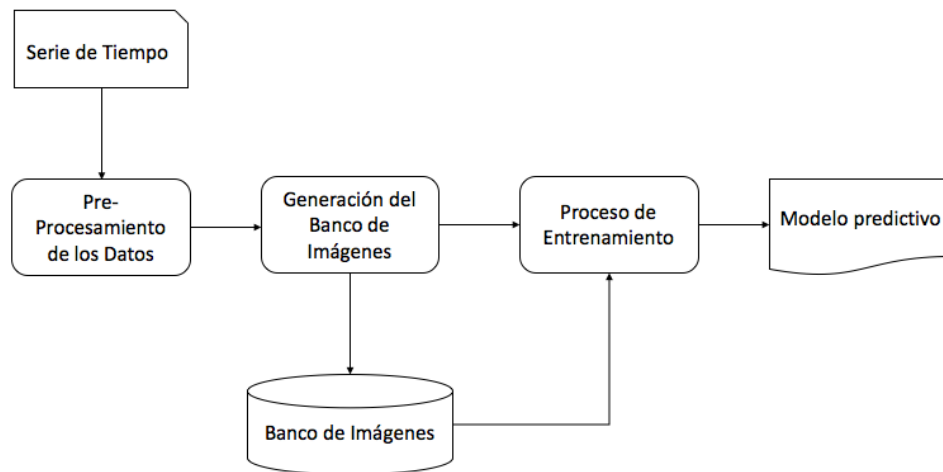


Figura 4.1: Proceso completo para la generación de un modelo predictivo de la velocidad del viento a un paso.

Se presenta un enfoque multi modelo el cual consta de la creación de bancos de modelos (de 24 modelos) para cada una de las series de tiempo de la velocidad del viento. Estos bancos de modelos se crearon realizando una iteración a 24 pasos p donde $1 \leq p \leq 24$. Para cada uno de estos pasos un banco de imágenes fue creado, asociando el valor esperado basado en este paso y es ejecutado el proceso de entrenamiento con el banco de imágenes como entrada. El proceso de generación del banco de modelos se describirá en la sección 4.5.

4.1. Pre-Procesamiento de Datos

En el proceso de Pre-Procesamiento de datos se realizan distintas tareas que preparan los datos para que puedan ser procesados correctamente. Entre estas tareas está el revisar por datos atípicos (Outliers) en la serie de tiempo y la normalización de los datos.

Un dato atípico es una observación que cae fuera del patrón general de la distribución de los datos. Usualmente, la presencia de un dato atípico indica algún tipo de problema. Esto puede ser un caso en el cual el dato no quepa en el modelo que se estudia, o un error en la medición. Los datos atípicos en los datos pueden sesgar significativamente los resultados de los datos que se están procesando (Wolfram, 1999). La Figura 4.2 muestra un ejemplo de un dato atípico, se observa un conjunto de observaciones donde una de ellas se encuentra fuera

del comportamiento de las observaciones. A este dato de llamamos un dato atípico.

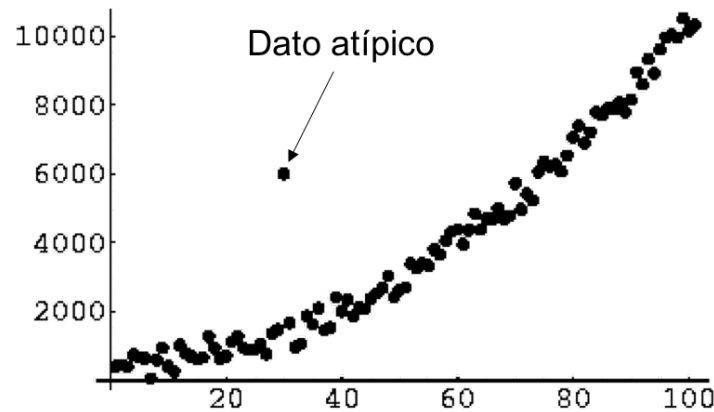


Figura 4.2: Ejemplo de dato atípico en un conjunto de observaciones

Un dato atípico puede presentarse por distintas razones. Algunas veces debido a fallo de los sensores, fallo humano, falta de mantenimiento, falta de comunicación, entre otras razones. Los sensores que se utilizan en la estación meteorológica son llamados anemómetros, los cuales se utilizan para medir la velocidad del viento y así realizar predicciones del clima o en este caso para realizar la medición de la velocidad del viento. Cuando se presentaron datos atípicos en las series de tiempo utilizadas se empleó una técnica estadística para suavizar dichos datos atípicos. Esta técnica es llamada Moving Average o en español Media Móvil, la cual se utiliza para realizar suavizados de un conjunto de datos. En estadística esta técnica es utilizada para realizar un análisis de los datos u observaciones creando una serie a partir de promedios. De esta manera, una media móvil es una lista de observaciones donde cada una de estas resulta del promedio de un subconjunto de los datos originales.

La técnica Media Móvil es comúnmente utilizada para el suavizado de los datos. Se desliza una ventana a través de los datos, computando la media de un subconjunto de observaciones de la serie. Ayudando a eliminar significativamente variaciones de un punto de datos a otros (Wolfram, 1999). Esta técnica es efectiva para el suavizado de datos cuando la serie de tiempo esta contaminada con datos inusuales o Outliers. La Media Móvil es definida de la siguiente manera (Montgomery et al., 2015).

$$M_T = \frac{y_T + y_{T-1} + \dots + y_{T-N+1}}{N} = \frac{1}{N} \sum_{t=T-N+1}^T y_t \quad (4.1)$$

La Figura 4.3 muestra el resultado de aplicar la técnica Media Móvil en una serie de tiempo de la velocidad del viento con datos atípicos.

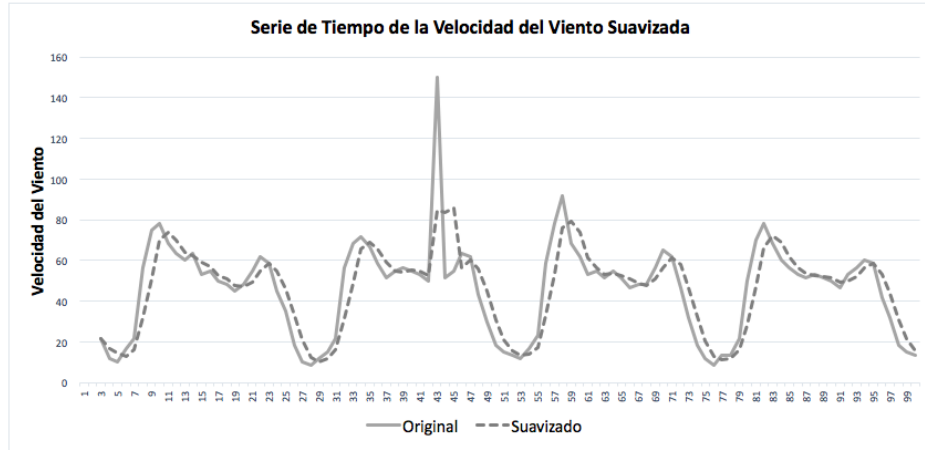


Figura 4.3: Ejemplo de serie de tiempo de la velocidad del viento suavizada al presentar datos atípicos.

Además del análisis de datos atípicos en la serie de tiempo se realizó una normalización de los datos en la serie de tiempo, lo cual significa que las observaciones fueron pasadas a un rango más pequeño de valores entre $0 \leq x_i \leq 1$ donde x_i es una observación en la serie de tiempo X en la posición i .

La Equación 4.2 muestra la manera en la que se llevó a cabo la normalización de los datos. Las funciones X_{max} y X_{min} obtienen el valor máximo y mínimo respectivamente de X .

$$X_{normalizado} = \frac{x_i - X_{min}}{X_{max} - X_{min}} \quad (4.2)$$

La Figura 4.4 muestra un ejemplo de la normalización de una serie de tiempo, a la izquierda están los datos como fueron medidos originalmente por los sensores sin pasar por

un proceso de normalización. Sus valores se encuentran en un rango no estandarizado. A la derecha se muestra la serie de tiempo normalizada entre un rango de valores de $[0, 1]$ y en el centro se muestran los valores Máximo y Mínimo respectivamente de la serie de tiempo completa.

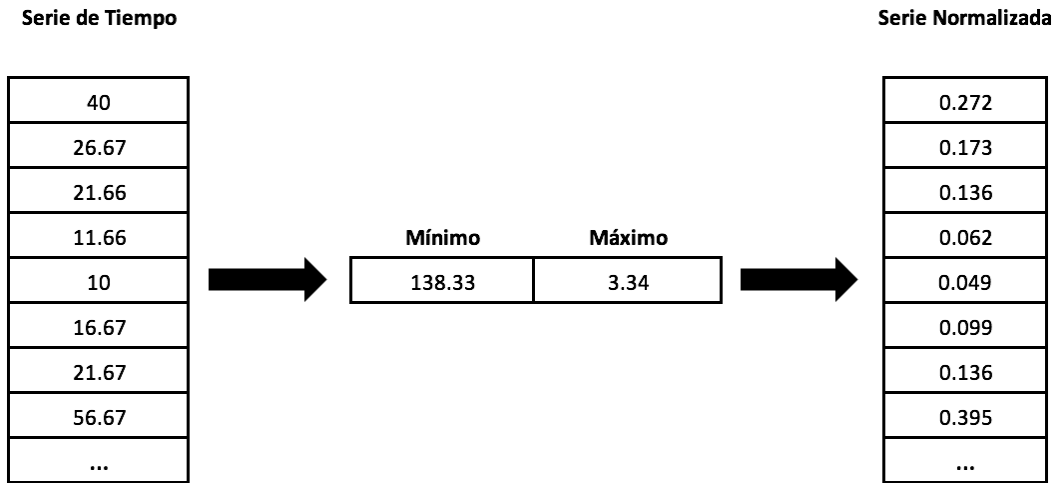


Figura 4.4: Serie de tiempo normalizada en un rango de valores entre $[0, 1]$

4.2. Serie de Tiempo a Imagen

Debido a que se trabajó con el pronóstico de la velocidad del viento utilizando la técnica de aprendizaje profundo llamada RNC, la cual está pensada para tareas como la visión artificial, detección de objetos, clasificación de imágenes, entre otras. Las RNC reciben como entrada la información estructurada en un vector de matrices donde cada una de estas matrices representan los datos de una imagen.

Debido a esto se buscó en la literatura una técnica que permita realizar el proceso codificación de una serie de tiempo en una imagen con la cual la RNC pueda ser alimentada para el proceso de entrenamiento. Para la codificación de la serie de tiempo a imagen se desarrolló e implementó una técnica basada en Gramian Angular Field, la cual se describe la sección 4.4.2. La técnica presentada en este trabajo presenta una variación importante, para la implementación solo se consideró a las observaciones de la velocidad del viento y no así la marca de tiempo de la serie.

4.3. Banco de Imágenes

Para este proyecto se trabajó en la realización de bancos de imágenes a partir de una serie de tiempo completa. El proceso para convertir una serie de tiempo a banco de imágenes consta de definir un tamaño de imagen w por ejemplo 24 con el cual las imágenes resultantes tendrán una medida final de 24×24 pixeles. Además se debe definir el número de paso p , con el cual se calcula el valor que será asociado y_{t+p} a la imagen resultante.

La Figura 4.5 muestra como la ventana definida w se va recorriendo a través de toda la serie de tiempo. Cada sub-serie resultante \tilde{X} es convertida a imagen posteriormente.

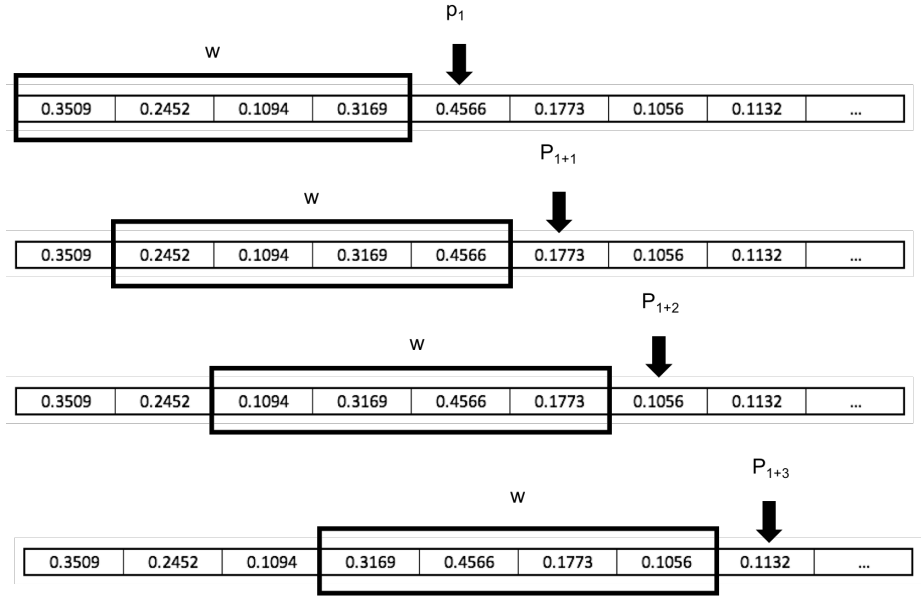


Figura 4.5: Deslizamiento de la ventana w y asignación de valor esperado en el paso uno

Una vez tomados los w datos de la serie de tiempo estos pasan por un proceso para convertir una serie de tiempo a imagen, el cual está descrito en el Algoritmo 4.2. A la matriz resultante se le asocia un valor esperado \hat{y}_{t+p} .

La Tabla 4.1 ejemplifica como es que los datos son tomados de la serie de tiempo. Se extraen w características y son asociadas a un valor de salida. Cada registro dentro de la tabla representa las últimas w observaciones en un tiempo definido (t), asociado a una salida esperada \hat{y}_{t+p} , las observaciones correspondientes a un registro son tomadas para generar la imagen junto con el valor esperado.

#Imagen	w					$Esperado$	
1	$\tilde{x}_1,$	$\tilde{x}_2,$	$\tilde{x}_3,$	$\tilde{x}_4,$	\dots	$\tilde{x}_w,$	\hat{y}_{t+1}
2	$\tilde{x}_2,$	$\tilde{x}_3,$	$\tilde{x}_4,$	$\tilde{x}_5,$	\dots	$\tilde{x}_{w+1},$	\hat{y}_{t+2}
3	$\tilde{x}_3,$	$\tilde{x}_4,$	$\tilde{x}_5,$	$\tilde{x}_6,$	\dots	$\tilde{x}_{w+2},$	\hat{y}_{t+3}
4	$\tilde{x}_4,$	$\tilde{x}_5,$	$\tilde{x}_6,$	$\tilde{x}_7,$	\dots	$\tilde{x}_{w+3},$	\hat{y}_{t+4}
5	$\tilde{x}_5,$	$\tilde{x}_6,$	$\tilde{x}_7,$	$\tilde{x}_8,$	\dots	$\tilde{x}_{w+4},$	\hat{y}_{t+5}
6	$\tilde{x}_6,$	$\tilde{x}_7,$	$\tilde{x}_8,$	$\tilde{x}_9,$	\dots	$\tilde{x}_{w+5},$	\hat{y}_{t+6}
7	$\tilde{x}_7,$	$\tilde{x}_8,$	$\tilde{x}_9,$	$\tilde{x}_{10},$	\dots	$\tilde{x}_{w+6},$	\hat{y}_{t+7}
8	$\tilde{x}_8,$	$\tilde{x}_9,$	$\tilde{x}_{10},$	$\tilde{x}_{11},$	\dots	$\tilde{x}_{w+7},$	\hat{y}_{t+8}
9	$\tilde{x}_9,$	$\tilde{x}_{10},$	$\tilde{x}_{11},$	$\tilde{x}_{12},$	\dots	$\tilde{x}_{w+8},$	\hat{y}_{t+9}
10	$\tilde{x}_{10},$	$\tilde{x}_{11},$	$\tilde{x}_{12},$	$\tilde{x}_{13},$	\dots	$\tilde{x}_{w+9},$	\hat{y}_{t+10}
			\dots				
n	$\tilde{x}_{n-w-1},$	$\tilde{x}_{n-w},$	$\tilde{x}_{n-w+1},$	$\tilde{x}_{n-w+1},$	\dots	$\hat{y}_{n-1},$	\hat{y}_n

Tabla 4.1: Ejemplo de generación del banco de imágenes para un paso definido de uno.

El Algoritmo 4.1 muestra el proceso de creación del banco de imágenes, inicia recibiendo como parámetros la serie de tiempo, el tamaño de la imagen o las w observaciones que representarán las imágenes y el número de paso. En la línea 4 se realiza una iteración recorriendo la ventana definida (tamaño de imagen) a través de toda la serie de tiempo. En la línea 5 con el tamaño de ventana, se extrae una sub-serie que se convertirá en la matriz que representa la imagen para posteriormente (siguiente línea) ser normalizada en un rango de $[-1, 1]$. Con el paso definido se obtiene el valor esperado que se asocia a la imagen en la línea 7. Una vez que se obtiene el vector de características o sub-serie se le da el formato de matriz que representa la imagen en la línea 8. De esta manera e iterando la serie de tiempo completa es como se obtienen el banco de de imágenes.

Algoritmo 4.1 Banco de imágenes

```
1: procedure BIMAGENES(serieTiempo, tamañoImagen, paso)
2:   bancoDeImágenes := []
3:   valoresEsperados := []
4:   for  $i := 0$ ; ( $\text{len}(\text{serieTiempo}) - \text{tamañoImagen} - \text{paso} + 1$ );  $i := i + 1$  do
5:     serieImagen := serieTiempo[  $i$ :  $i + \text{tamañoImagen}$  ]
6:     normalizar(serieImagen) // -1 a 1
7:     valorEsperado = serieTiempo[  $i + \text{tamañoImagen} + \text{paso} - 1$  ]
8:     matrizImagen = serieAImagen(serieImagen)
9:     bancoDeImágenes.append(matrizImagen)
10:    valoresEsperados.append(valorEsperado)
11:  end for
12:  return bancoDeImágenes, valoresEsperados
13: end procedure
```

4.4. Proceso de Serie de Tiempo a Imágen

En la literatura se puede encontrar diversas técnicas para la transformación de una serie de tiempo a imagen. Entre ellas están las técnicas llamadas Gramian Angular Field y Escaneo Lineal. Se utilizaron ambas técnicas para crear bancos de imágenes con diferentes tamaño de imágenes. Se observó que en los bancos de imágenes creados con la técnica de escaneo lineal el tamaño del banco de imágenes disminuye significativamente. En la Sección 4.4.1 y 4.4.2 se explica como fueron implementadas las diferentes técnicas.

4.4.1. Escaneo Lineal

La técnica por escaneo lineal está basada en la publicación de (Izbicki, 2011). Esta técnica realiza un escaneo de la imagen pixel por pixel pasando los valores a un vector (donde un vector se puede ver como una serie de tiempo) de números reales. La Figura 4.6 muestra el proceso de realizar el Escaneo Lineal para generar una serie de tiempo a partir de una imagen.



Figura 4.6: Imagen a serie de tiempo por Escaneo Lineal.

Debido a que es necesario transformar una serie de tiempo a imagen se realizó la operación contraria. Se define una ventana o número de características que se necesitan para representar la imagen. Cada una de estas observaciones se normalizan en valores de un rango de $[0, 255]$ para representar los valores de pixel. Cada uno de estas observaciones ya normalizada representa un valor de pixel de la imagen. La Figura 4.7 muestra el proceso para convertir una serie de tiempo a imagen con esta técnica.

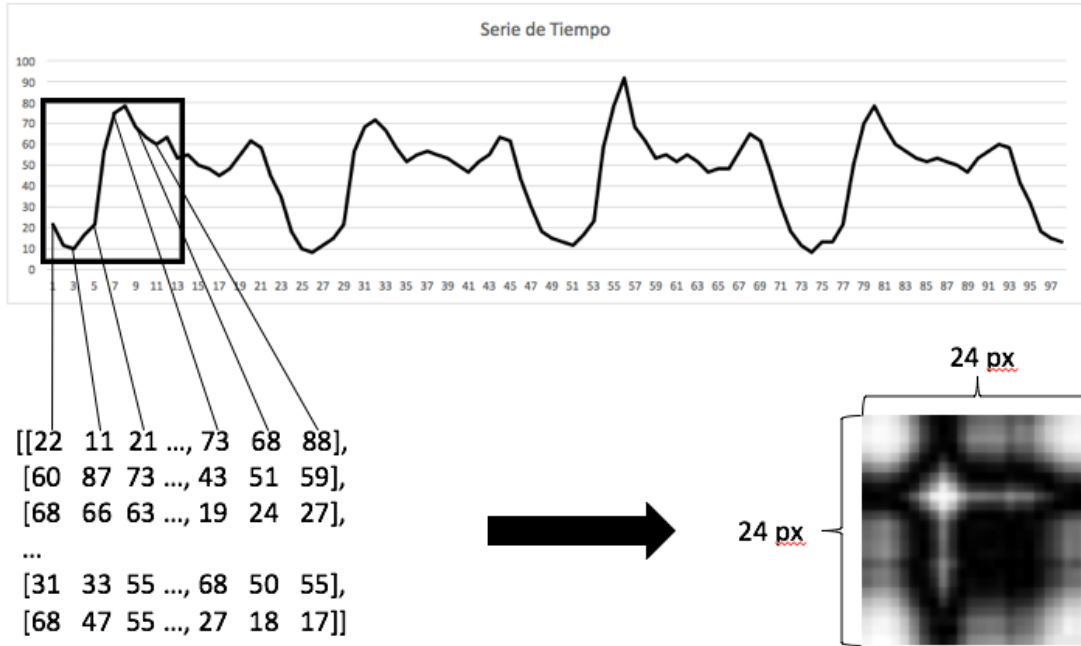


Figura 4.7: Serie de tiempo a imagen por Escaneo Lineal.

4.4.2. Gramian Angular Field

Esta técnica llamada Gramian Angular Field (GAF) permite transformar un subconjunto de la serie de tiempo en una imagen. La técnica empleada está basada en el trabajo publicado por (Wang and Oates, 2015).

Para la generación de las imágenes, dada la serie de tiempo $X = \{x_1, x_2, \dots, x_n\}$ de k observaciones, donde :

$$\{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n \quad (4.3)$$

se toma un **subconjunto** de esta donde $\tilde{X} \subset X$, cuyos valores se normalizan en un rango de $[-1, 1]$ de la siguiente manera :

$$x_{i\text{normalizado}} = \frac{(x_i - \tilde{X}_{max}) + (x_i - \tilde{X}_{min})}{\tilde{X}_{max} - \tilde{X}_{min}} \quad (4.4)$$

Donde x_i representa una observación en la posición i del vector \tilde{X} , las funciones \tilde{X}_{max} y \tilde{X}_{min} representan la obtención del valor máximo y mínimo respectivamente del vector \tilde{X} y

x_i *normalizado* representa el dato x_i normalizado. Quedando los valores normalizados en el vector \tilde{X}

(Wang and Oates, 2015) representan de la siguiente manera la sub-serie recalada \tilde{x} en coordenadas polares.

$$\begin{cases} \phi = \arccos(\tilde{x}_i), -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \quad (4.5)$$

Se codifica el valor de la serie de tiempo como el $\arccos \alpha$ siendo α representado por \tilde{x}_i . La marca de tiempo es considerada como el radio de la coordenada polar.

En el trabajo de (Wang and Oates, 2015) la generación de la imagen se realizó tomando un subconjunto serie de tiempo y representándolo en el sistema de coordenadas polares. Se forma una matriz en la cual cada uno de sus elementos consiste en la suma del coseno de ángulos (Wang and Oates, 2015).

Una vez codificados los valores de la serie de tiempo en su valor angular se procede a generar la matriz que represente la imagen, como se muestra en la siguiente matriz:

$$\begin{bmatrix} \cos(\tilde{x}_i + \tilde{x}_i) & \cos(\tilde{x}_i + \tilde{x}_{i+1}) & \cdots & \cos(\tilde{x}_i + \tilde{x}_{i+n}) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\tilde{x}_{i+n} + \tilde{x}_i) & \cos(\tilde{x}_{i+n} + \tilde{x}_{i+1}) & \cdots & \cos(\tilde{x}_{i+n} + \tilde{x}_{i+n}) \end{bmatrix} \quad (4.6)$$

Donde \tilde{x}_i representa una observación en la posición i del vector \tilde{X} . Una vez que se cuenta con la matriz, sus valores se escalan entre $[0, 255]$ de forma que estos representen valores de pixel en una imagen. El tamaño de la matriz resultante es de $w \times w$ donde w representa la longitud del vector \tilde{X} . El Algoritmo 4.2 describe el proceso de codificar una serie de tiempo a imagen usando GAF. El Algoritmo inicia calculando el \arccos de cada uno de los elementos de la serie de tiempo en la línea 2. Enseguida en un ciclo anidado se crea la la matriz resultante, recorriendo cada uno de los valores del vector \tilde{x} y calculando la sumatoria del *coseno* de la sumatoria de $v1$ y $v2$.

Algorithm 4.2 Serie de Tiempo a Imagen

```

1: procedure SERIEIMAGEN(serieTiempo)
2:   for all  $x$  in serieTiempo do
3:      $x := \arccos(x)$ 
4:   end for
5:   imagen := []
6:   for  $v_1$  in serieTiempo do
7:     for  $v_2$  in serieTiempo do
8:        $v_3 := \cos(v_1 + v_2)$ 
9:       imagen.append( $v_3$ )
10:    end for
11:  end for
12:  normalizar(imagen) // 0 a 255
13:  return imagen
13: end procedure

```

La Figura 4.8 muestra un ejemplo de la imagen resultante al transformar una serie de tiempo de la velocidad del viento a imagen utilizando la técnica Gramian Angular Field.

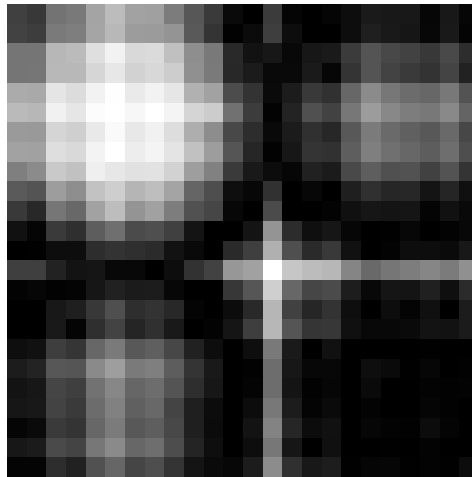


Figura 4.8: Serie de tiempo codificada a una imagen utilizando GAF

4.4.3. Proceso de Entrenamiento

El sub-proceso de entrenamiento es el más largo del proceso completo. Este es un proceso iterativo que consta de un Algoritmo Genético Compacto y una Red Neuronal Convolutiva. El AGc se encarga de realizar una búsqueda estocástica de un conjunto de hiperparámetros. Estos hiperparámetros o posible mejor solución del AGc serán utilizados por la RNC para realizar el entrenamiento de un modelo predictivo de la velocidad del viento. La RNC es

la encargada de realizar el entrenamiento del modelo predictivo. La arquitectura de RNC utilizada para este proyecto está basada en el trabajo realizado por (Pérez-Espinosa et al., 2016). Esta conformada por 10 capas de las cuales son 2 capas de convolución las cuales van transformando la entrada a través de cada una de ellas. En esta arquitectura se optimizaron los diferentes hiperparámetros obtenidos por AGc. La Figura 4.9 muestra la arquitectura de la RNC utilizada para este proyecto.

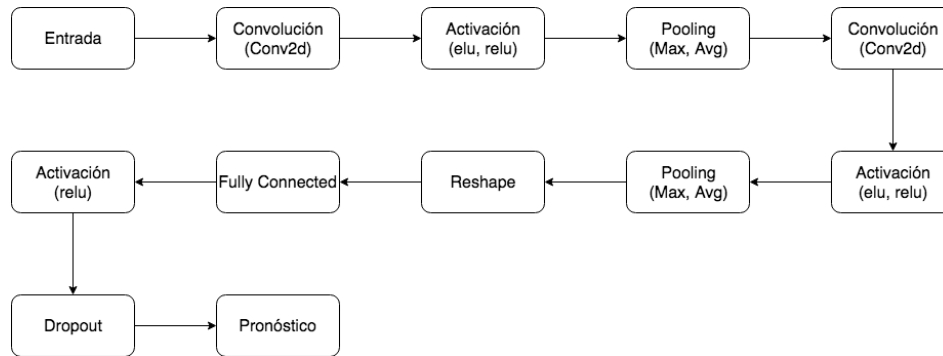


Figura 4.9: Arquitectura utilizada para la RNC

El determinar los mejores hiperparámetros para la Red Neuronal Convolutiva es un problema de optimización. Las Redes Neuronales por su complejidad toman una serie de hiperparámetros que necesitan ser configurados correctamente para obtener un mejor resultado. Normalmente esta serie de hiperparámetros son configurados utilizando un método muy común de prueba y error hasta llegar a un resultado deseado. A pesar de eso este método no garantiza un rendimiento óptimo en el proceso de entrenamiento y el tiempo para llegar a un resultado deseado puede ser demasiado alto (Pérez-Espinosa et al., 2016). En este trabajo se realiza una optimización de los hiperparámetros utilizando un AGc. Esta optimización de hiperparámetros está basada en el trabajo de (Pérez-Espinosa et al., 2016). Donde se define una serie de hiperparámetros a optimizar con una técnica llamada Covering Array, la cual consiste en crear una base de datos con todas las posibles combinaciones de hiperparámetros y con esta base de datos de combinaciones realizar los experimentos. Para la implementación de la RNC se tomaron solo algunos de esos hiperparámetros debido a que se consideró que no era necesario incluirlos todos.

A continuación se describen cada uno de los hiperparámetros considerados, en total son 10 hiperparámetros clasificados entre General y Convolutiva. En la clasificación **General**

se refieren a parámetros que son “normales” en Redes Neuronales Artificiales, mientras en la clasificación **Convolutiva** están los relacionados con las Redes Neuronales Convolutivas.

1. General

- a)* **Épocas:** Es un número entero positivo que limita el número de veces que el conjunto de validación es evaluado.
- b)* **Aprendizaje:** Es un número de punto flotante que representa la magnitud de la actualización por cada época de entrenamiento.
- c)* **Entrenamiento:** Es un número que representa el porcentaje que se considerara para el conjunto de entrenamiento.
- d)* **Optimizador:** Son funciones que computan gradientes para una medida de pérdida y aplican gradiente a variables. Los tres posibles valores de este parámetro son Stochastic gradient descent(SGD), AdamOptimizer(ADAM) y RMSPropOptimizer (RMSProp).
- e)* **Activación:** Funciones que proveen la no linealidad. Este parámetro puede tomar tres valores. una función rectificadora (relu) y una función para suavizar la no linealidad.

2. Convolutiva

- a)* **Tamaño del Filtro:** Es un número que define el tamaño del filtro que recorrerá la matriz que representa la imagen.
- b)* **Strides:** Una lista de números enteros que representa el número de características que serán recorridas de izquierda a derecha en la matriz, moviendo el filtro por cada dimensión en el vector de entrada. En este caso solo se usan imágenes con una sola dimensión.
- c)* **Padding:** Indica si el filtro puede ir más allá de los límites de la matriz. Los valores que puede tomar este parámetro son dos same y valid.

- d) **Pooling:** Reduce la dimensionalidad de la entrada permitiendo hacer asunciones de características contenidas en una región de la entrada. Los valores que puede tomar este parámetro son `max_pooling` y `avg_pooling`.
- e) **Dropout:** Es una técnica de regularización para reducir el sobre entrenamiento en redes neuronales.

Este proceso de entrenamiento recibe como entrada la base de datos formada a partir de la serie de tiempo con el pre-procesamiento de los datos ya realizado. El Algoritmo Genético Compacto 2.1 utilizado fue definido en la sección 2.4.1 del Marco Teórico.

Para la optimización de los hiperparámetros se utiliza un AGc el cual es una versión mejorada del algoritmo genético simple. En el algoritmo AGc cada individuo (Cromosoma) es definido como un vector binario que codifica cada uno de los hiperparámetros listado anteriormente. El AGc es usado para definir el mejor conjunto de hiperparámetros para la RNC.

El AGc básicamente provee cuantas épocas serán las óptimas para el modelo, que factor de aprendizaje utilizar para obtener el mejor resultado, que algoritmo de optimización es el adecuado, etc. Al final del proceso, el mejor individuo (i.e. el que presentó mejor adaptabilidad) encontrado durante la evolución será retornado como la solución del problema (el conjunto de hiperparámetros).

En el proceso de entrenamiento para la generación de una modelo predictivo de la velocidad del viento existe una interacción entre el AGc y la RNC, estos se ejecutan en CPU y GPU respectivamente. El AGc ejecutándose en el CPU genera un individuo o posible solución, el cual es pasado a la RNC e inicia su ejecución en la GPU. Se realiza el entrenamiento con los parámetros que se le provee, obtiene la aptitud del modelo resultante y retorna su aptitud al AGc para posteriormente realizar la optimización de los hiperparámetros basado en esta medida de aptitud. La medida de aptitud utilizada para la optimización de los hiperparámetros fue el error cuadrático medio. La Figura 4.10 muestra la interacción que existe entre el AGc y la RNC. La ejecución de la RNC se realizó utilizando una unidad de procesamiento gráfico (GPU). Mientras que la implementación del AGc se realizó utilizando la unidad central de procesamiento (CPU) de la computadora en la que se realizaron los experimentos.

La RNC cuenta con un enfoque en paralelo al ejecutar todas sus operaciones en la tarjeta GPU el algoritmo genético trabaja de manera secuencial en el CPU. Durante el proceso de entrenamiento el AGc genera un individuo o posible solución (Topología de la RNC) y éste es enviado al GPU para que realice el entrenamiento. Este es un proceso iterativo que finaliza al satisfacerse un criterio de convergencia. Al finalizar el entrenamiento en la RNC esta retorna la aptitud del modelo con la solución (topología) dada por el AGc.

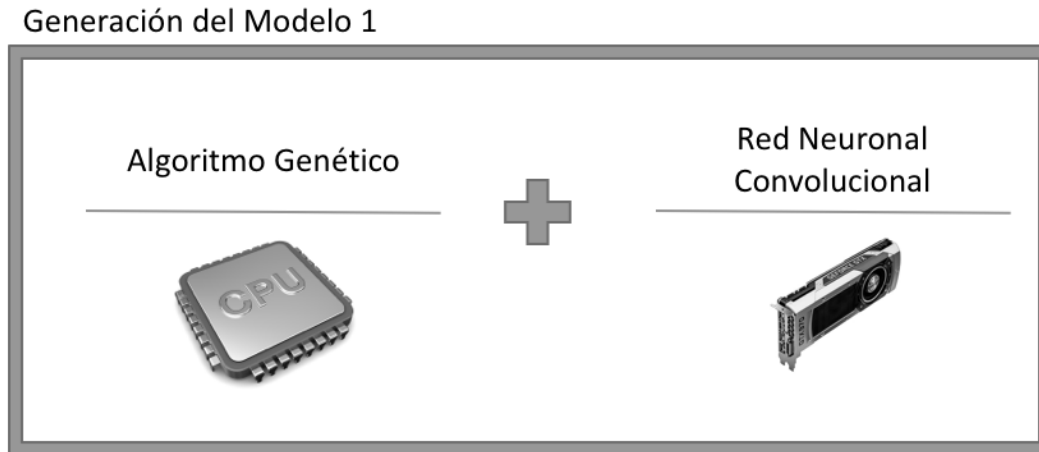


Figura 4.10: Representación de la interacción entre AGc y RNC.

La RNC es utilizada dentro del AGc al momento de realizar la evaluación. Esta recibe como entrada una base de datos de entrenamiento la cual está conformada del banco de imágenes generado en el proceso mostrado en la Sección 4.3 donde cada una de las imágenes del banco de imágenes está asociada a un valor esperado. Por último el resultado final del proceso de entrenamiento nos da un modelo predictivo de la velocidad de viento optimizado.

4.5. Banco de modelos

Durante la sección 4.4.3 se habló del proceso de entrenamiento. El proceso visto hasta ahora es capaz de generar solo un modelo predictivo de la velocidad del viento. Debido al enfoque orientado a banco de modelos dado es necesario generar más de un modelo predictivo para una serie de tiempo. Esto con el propósito de contar con pronósticos a más de una hora en el futuro.

Este enfoque orientado a banco de modelos busca contar con modelos predictivos que

permiten obtener el pronóstico de la velocidad del viento en las siguientes 24 horas. Para esto es necesario obtener un modelo predictivo para cada una de las horas (también llamados pasos) es decir 24 modelos por serie de tiempo. Dado un número de paso generar un modelo para este paso. Este número de paso representa que dadas w observaciones pasadas, el valor asociado a ellas estará p pasos en el futuro.

En la Figura 4.5 se ejemplifica la manera en la que se genera la base de datos necesaria para el entrenamiento del modelo cuando el número de paso es 1. Se muestra como la ventana w se va recorriendo a través de la serie de tiempo. A su vez se muestra el valor del número de paso p que será asociado a la imagen resultante del subconjunto de observaciones w . A medida que la ventana w se va recorriendo también se recorre el valor asociado en p pasos hacia el futuro.

En la Figura 4.11 también se ejemplifica como se genera la base de datos necesaria para el entrenamiento del modelo solo que en este caso el número de paso es 2.

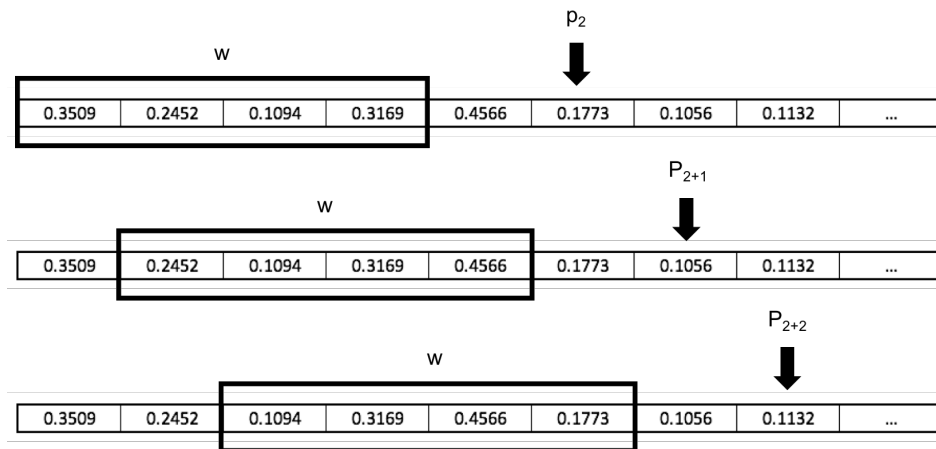


Figura 4.11: Deslizamiento de la venta y asignación de valor esperado en el paso dos

El proceso de generación del banco de modelos es un proceso iterativo, que recibe como entrada la base datos de imágenes. La generación de los bancos de modelos se realiza a través del proceso de entrenamiento. Este es un proceso iterativo que va 1 a 24 pasos, es decir se generan los modelos para 24 pasos por lo cual el proceso de entrenamiento debe ser ejecutado 24 veces. La Figura 4.12 muestra el proceso iterativo para la generación de los 24 modelos.

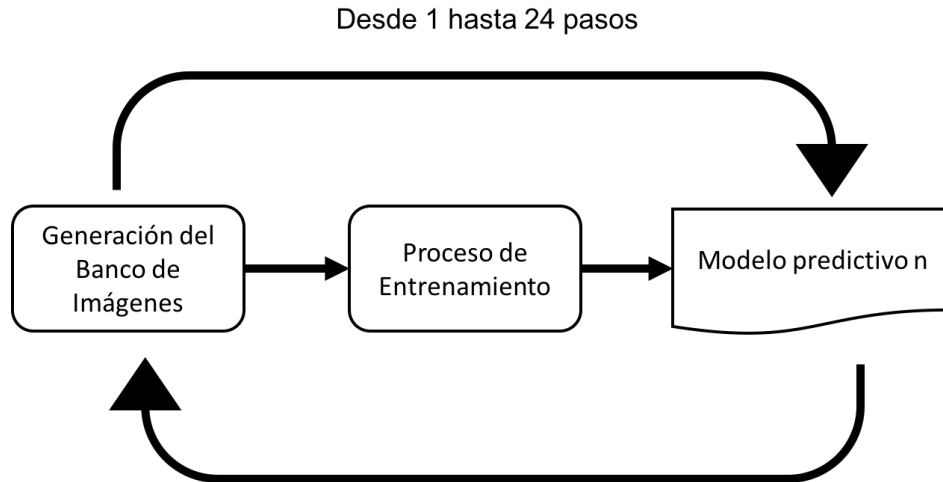


Figura 4.12: Ciclo de generación 24 bancos de modelos predictivos

El siguiente Algoritmo 4.3 muestra como son generados los 24 modelos que conforman el banco de modelos, inicia recibiendo como entrada la serie de tiempo con la que se generará el banco de modelos. En la línea 3 se realiza un recorrido de los 24 pasos, generando el banco de imágenes y ejecutando el proceso de entrenamiento en las siguientes dos líneas respectivamente. El proceso de entrenamiento retorna una modelo entrenado el cual será agregado al banco.

Algoritmo 4.3 Algoritmo para la generación del banco de modelos dada una serie de tiempo

```

1: procedure BANCODEMODELOS(serieTiempo)
2:   tamanoImagen := 24
3:   for paso := 1, i ≤ 24 do
4:     x := bImagenes(serieTiempo, tamanoImagen, paso)
5:     modelo := ProcesoDeEntrenamiento
6:     BM := insertarEnBancoDeModelos(modelo)
7:   end for
   return BM
8: end procedure

```

La imagen 4.13 ejemplifica la interacción que existe entre el CPU y el GPU para la generación del banco de modelos ejecutándose el AGc y la RNC en el CPU y GPU respectivamente.

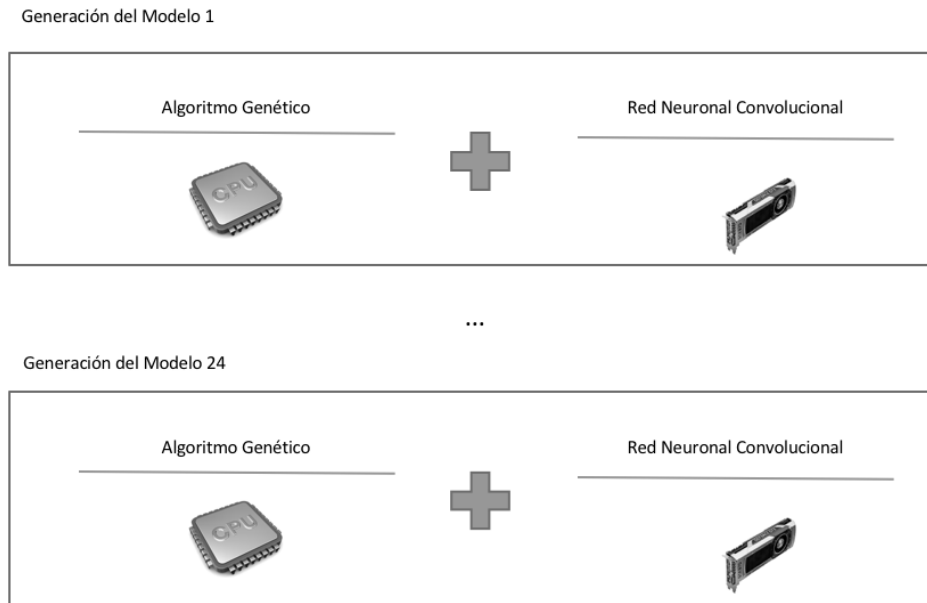


Figura 4.13: Representación gráfica de la generación del banco de modelos y la interacción entre el AGc en el CPU y la RNC en el GPU

Con la realización de este banco de modelos se obtiene una manera más completa de obtener un pronóstico de la velocidad del viento a corto plazo. De esta manera se tiene la posibilidad de contar con un pronóstico de las siguientes 24 pasos o las siguientes 24 horas en el futuro.

4.6. Tecnologías Utilizadas

A lo largo del desarrollo se hizo uso de distintas tecnologías, así como librerías y hardware. El desarrollo se realizó utilizando el lenguaje de programación Python por su sencillez y debido a que este cuenta con un gran soporte por parte de la comunidad de desarrolladores y actualmente cuenta con mucho desarrollo sobre aprendizaje máquina y aprendizaje profundo. Para la implementación se utilizaron las librerías Keras y TensorFlow, las cuales permiten realizar la implementación de la RNC de una manera mucho más sencilla. Además la ejecución de la RNC se realizó utilizando una tarjeta aceleradora gráfica con la finalidad de obtener una reducción importante en los tiempos de ejecución. Empezaremos con la definición del lenguaje de programación Python.

4.6.1. Python

Python es un lenguaje de propósito general de alto nivel cuya filosofía de diseño se centra en la legibilidad. La sintaxis de Python le permite a los programadores expresar sus ideas en pocas líneas de código. Python soporta múltiples paradigmas de programación, incluyendo orientada a objetos, imperativa y funcional (Van Rossum et al., 2007).

4.6.2. Keras

Para la implementación de la RNC en este trabajo se utilizó la librería Keras la cual es una interfaz de programación de aplicaciones (API por sus siglas en inglés) de alto nivel para redes neuronales. Esta escrita en el lenguaje de programación Python y es capaz de correr sobre diferentes Backends como lo son TensorFlow, CNTK o Theano. Fue desarrollada enfocándose en permitir experimentación rápida, permitiendo poder ir de una idea a resultados rápidamente con el menor retraso posible lo cual es una de las claves para realizar una buena investigación (Chollet et al., 2017).

Además, minimiza el número de acciones que un usuario debe seguir para una tarea común. Está diseñada pensando en la modularidad. Un Modelo es entendido como una secuencia y son completamente configurables de manera que pueden ser unidos unos con otros (Chollet et al., 2017).

4.6.3. Tensorflow

TensorFlow fue originalmente desarrollado para los investigadores e ingenieros del equipo Google's Brain con el objetivo de conducir investigaciones sobre aprendizaje máquina y aprendizaje profundo. Este software es actualmente usado tanto para investigación y producción de diferentes equipos en muchos de los productos comerciales de la empresa Google. Productos como reconocimiento del habla, Gmail y Google Fotos. Una importante característica de esta herramienta es su arquitectura la cual es muy flexible y permite implementar software que utilice uno o más CPU's o GPU's en una computadora de escritorio, servidor o dispositivo móvil usando la misma API. Google liberó TensorFlow bajo la licencia de código libre Apache 2.0 en Noviembre de 2015 (Pérez-Espinosa et al., 2016).

Capítulo 5

Análisis de Resultados

Con el propósito de comprobar la efectividad de la metodología propuesta, esta se probó en el caso de estudio concreto del pronóstico a corto plazo de la velocidad del viento. En particular fueron utilizadas 4 series de tiempo de la velocidad del viento que contienen valores continuos capturados por estaciones de viento de distintas regiones del estado de Michoacán México. Cada series de tiempo contiene observaciones cuyo intervalo de captura entre observación es de una hora promedio, es decir cada observación en la serie de tiempo corresponde a un promedio de todos los datos capturados durante un lapso de una hora.

Las mediciones del viento fueron realizadas utilizando un dispositivo llamado anemómetro el cual es un instrumento utilizado para medir la velocidad de una ráfaga de gas (Choon et al., 2012). Las series de tiempo que se utilizaron son sub-series de las series completas. Las sub-series obtenidas no contaron con huecos o datos faltantes que pudieran perjudicar el resultado final. Cada una de las sub-series obtenidas constan de 4,416 observaciones.

Los experimentos fueron divididos en tres fases, la fase de pre-procesamiento de los datos, la fase entrenamiento del modelo (la cual se muestra en la Figura 4.13), y la fase de pronóstico. Una vez que el pre-procesamiento de los datos fue realizado para cada serie se procedió a realizar el entrenamiento. Dentro del proceso de entrenamiento se realiza la generación del banco de imágenes, generando las imágenes con las que este será realizado y asociándolas a un valor esperado. El tamaño del banco de imágenes está directamente relacionado con el tamaño de la serie de tiempo utilizada, el tamaño de las imágenes generadas fue de 24 x 24 pixeles.

La última fase es la de pronóstico en la cual se utiliza el modelo generado en la fase

de entrenamiento para generar el pronóstico, en esta fase también se evalúa que el modelo generado cumpla con la aptitud buscada. El tamaño de las series fue estandarizada en 4,416 observaciones para todas las series utilizadas. El tamaño del banco de imágenes resultante para cada banco de modelos que corresponde al paso uno (1 hora a futuro) fue de 4,392 imágenes. La Tabla 5.2 muestra el tamaño del banco de imágenes para el paso uno utilizando diferentes tamaños de imágenes.

Para la transformación de las series de tiempo a imágenes principalmente se utilizó una técnica basada en Gramian Angular Field presentada por (Wang and Oates 2015). Se realizaron pruebas con dos técnicas, Escaneo Lineal y Gramian Angular Field para crear bancos imágenes. Se observó que en los bancos de imágenes creados con la técnica de escaneo lineal el tamaño (número de imágenes) de este disminuye significativamente.

Las Tablas 5.1 y 5.2 muestran la manera en la que varía el número de imágenes en el banco. La columna **Tamaño** contiene el tamaño de la imagen en pixeles, la columna **Observaciones** contiene el número de observaciones de la serie tiempo necesarias para formar una imagen, esto es el tamaño de la ventana w y la columna **Imágenes** contiene el número de imágenes en el banco de imágenes resultante. En la Tabla 5.1 se muestra que a medida que el tamaño de la imagen va aumentando el número de imágenes por banco de modelo disminuye, utilizando la técnica de Escaneo Lineal. La Tabla 5.2 muestra el comportamiento para la técnica Gramian Angular Field en el cual también disminuye el número de imágenes a medida que aumenta el tamaño de la imagen, pero esta disminución no es tan drástica como en la técnica por escaneo lineal. Se puede observar que en el banco de modelos con imágenes de 24 x 24 pixeles generado con Gramian Angular Field cuenta con 542 imágenes mas que el generado con Escaneo Lineal.

Tabla 5.1: Número de imágenes en el banco de imágenes utilizando Escaneo Lineal

Tamaño	#Observaciones	#Imágenes
12x12	144	4272
18x18	324	4092
24x24	576	3850
32x32	1024	3392
38x38	1444	2972
44x44	1936	2480

Tabla 5.2: Número de imágenes en el banco de imágenes utilizando Gramian Angular Field

Tamaño	#Observaciones	#Imágenes
12x12	12	4404
18x18	18	4398
24x24	24	4392
32x32	32	4384
38x38	38	4378
44x44	44	4372

Se realizó la generación de un banco de modelos por serie de tiempo de la velocidad del viento donde cada banco cuenta con 24 modelos, generando un total de 96 modelos predictivos de la velocidad del viento, correspondientes a las 4 series de tiempo utilizadas.

Para el proceso de entrenamiento se realizó una división de los datos, separándolos en un conjunto de entrenamiento y un conjunto de validación. Se tomó un 20 % fijo para el conjunto de validación y el porcentaje de datos que se utilizaron para el conjunto de entrenamiento fue decidido por el AGc. Para la obtención de cada uno los modelos se realizó la búsqueda de los diferentes hiperparámetros con el cual se obtuviera el modelo que mejor se apege a los datos de entrenamiento utilizando un AGc para la búsqueda de estos. Las Tablas 5.3 y 5.4 muestran los hiperparámetros utilizados, clasificados en Generales y Convolucionales respectivamente.

Tabla 5.3: Hiperparámetros generales y sus posibles valores

Épocas	Aprendizaje	Entrenamiento	Optimizador	Activación
20	0.0001	0.70	SGD	relu
40	0.0006	0.80	ADAM	elu
60	0.0011	0.90	RMSprop	
80	0.0016	1.00		
100	0.0021			
120	0.0026			
160	0.0031			
	0.0031			

Tabla 5.4: Hiperparámetros convolucionales y sus posibles valores

Tamaño del Filtro	Strides	Padding	Pooling	Dropout
3	2	valid	MaxPooling2D	0.3
4	3	same	AveragePooling2D	0.4
5	4			0.5
6	5			0.6

La Figura 5.1 muestra que posiciones del cromosoma proporcionado por el AGc corresponde a los diferentes hiperparámetros. El cromosoma es un vector de valores binarios, el cual dentro de la función de evaluación del AGc es decodificado, i.e para obtener el hiperparámetro Épocas se toman como referencia las 3 primeras posiciones del vector binario, se codifica a decimal y así se sabe que valor corresponde a la variable Épocas.

Épocas			Aprendizaje			Entrenamiento		Optimizador		Activación	Tamaño del Filtro		Strides		Padding	Pooling	Dropout	
1	0	1	0	0	0	1	1	1	0	1	0	0	0	1	1	0	1	1

Figura 5.1: Relación de las posiciones del cromosoma con los diferentes parámetros proporcionados por el AGc

Para el AGc se utilizó como configuración 100 generaciones, el tamaño del cromosoma de 19 posiciones y 100 individuos.

Se realizó una comparación de resultados obtenidos con la metodología propuesta utilizando RNC y los resultados obtenidos utilizando la técnica de estimación Naïve. Las Tablas 5.5, 5.6, 5.7 y 5.8 muestran una comparación de los resultados de las diferentes implementaciones de las series de tiempo utilizadas de el Fresno, La Piedad, Aristeo Mercado y La Piedad Michoacán.

Tabla 5.5: El Fresno

Método	MSE	RMSE
RNC	0.0369	0.1922
Naïve	0.0397	0.1922

Tabla 5.6: La Piedad

Método	MSE	RMSE
RNC	0.0246	0.1568
Naïve	0.0298	0.1728

Tabla 5.7: La Palma

Método	MSE	RMSE
RNC	0.0194	0.1394
Naïve	0.0327	0.1810

Tabla 5.8: Aristeo Mercado

Método	MSE	RMSE
RNC	0.0148	0.1219
Naïve	0.0169	0.1301

La comparación se realiza utilizando las medidas de aptitud MSE y RMSE. Los valores de las medidas de aptitud son promedios de las 24 modelos.

La implementación de la RNC se realizó utilizando la librería Keras y el backend Tensorflow, se utilizó la versión de Tensorflow para GPU. Todos los experimentos se realizaron

utilizando la plataforma Python (Foundation, 2012). Para la implementación de la RNC se utilizó una tarjeta aceleradora gráfica de NVIDIA GTX 1080 con la cual se obtuvo una reducción en los tiempos de ejecución de la RNC. El CPU con el que contaba la computadora fue un Intel Core i5.

El objetivo principal de implementar la RNC utilizando una tarjeta gráfica fue lograr una reducción considerable en los tiempos de ejecución. Como ejemplo se ejecutó la generación modelo para el paso uno de la serie de tiempo de Aristeo Mercado el tuvo un tiempo de ejecución de una 1 hora 30 minutos sobre la GPU. Se ejecutó el mismo entrenamiento pero utilizando la CPU, superando las 24 horas de ejecución. Por lo que la reducción de tiempo de ejecución fue muy importante para la realización de los experimentos.

Se generaron un total de 24 modelos para cada una de las series de tiempo de la velocidad del viento de diferentes regiones del estado de Michoacán México. Son un total de 4 series de tiempo por lo cual se generaron un total de 192 modelos.

Capítulo 6

Conclusiones

Para la toma de decisiones es de suma importancia el contar con un pronóstico acertado. En el problema específico del pronóstico de la velocidad del viento se pueden tomar decisiones relacionadas con los tiempos de mantenimiento de los aerogeneradores, tener conocimiento de la materia prima con la que se va a disponer en las próximas 24 horas, entre otras cosas. Además con un pronóstico acertado de la velocidad del viento puede ser posible lograr una reducción en los costos de producción de la energía por parte de las centrales eléctricas.

En este trabajo se obtuvo el pronóstico de la velocidad del viento mediante banco de modelos en el cual cada modelo es independiente de los demás y está enfocado en un ahora en particular. Se utilizaron Redes Neuronales Convolucionales para realizar el entrenamiento de los diferentes modelos y se utilizó un Algoritmo Genético Compacto para la búsqueda del mejor conjunto de parámetros de la Red Neuronal Convolutiva, evitando realizar el clásico proceso de búsqueda a prueba y error, el cual podría suponer un mayor esfuerzo para encontrar un conjunto de parámetros con los cuales se obtuvieran los resultados deseados. Al enfoque dado a la metodología desarrollada se le dió el nombre de enfoque Multi-Modelo Convolutivo Evolutivo.

Se propuso una metodología para la obtención de bancos de modelos predictivos de la velocidad del viento, la cual está constituida por diferentes técnicas recopiladas de la literatura. Para el proceso de codificar una serie de tiempo a imagen se desarrolló una técnica basada en el trabajo presentada por (Wang and Oates, 2015), para el proceso de entrenamiento se utilizaron RNC, Para la optimización de los hiperparámetros de la RNC se tomaron como

referencia los trabajos de (Rodríguez et al., 2017) y (Pérez-Espinosa et al., 2016).

Se realizó una comparación de los resultados obtenidos utilizando RNC y la optimización de hiperparámetros de la RNC contra el pronóstico obtenido con el método Naïve, obteniendo mejores resultados con la metodología planteada. Con ambas técnicas o metodologías es posible obtener un pronóstico acertado a corto plazo de la velocidad del viento. En este caso con la metodología Multi-Modelo Convolutiva Evolutiva se obtuvieron mejores resultados.

En la literatura existen muchos trabajos enfocados en resolver el problema del pronóstico de la velocidad del viento empleando diferentes técnicas como las narradas en el capítulo 3. A la fecha de la redacción de este trabajo no se encontraron trabajos enfocados en atacar el problema del pronóstico a corto de la velocidad del viento utilizando el enfoque presentado aquí, por lo se considera este puede ayudar a mejorar el estado del arte en el área de pronóstico.

Como trabajo futuro se le dará al AGc una mayor flexibilidad para decidir diferentes aspectos como el número de capas de la RNC, ya que debido a la complejidad de la implementación, este se implementó con una arquitectura con las capas de forma fija. Esto permitirá además de hacer una búsqueda de los parámetros que mejor se adapten al entrenamiento también hacer una búsqueda de la mejor arquitectura. Para lograr esto se deberá contar con un poder de procesamiento mayor por lo cual se realizará la implementación de una segunda tarjeta gráfica, las cuales estarán realizando las tareas en forma asíncrona en paralelo.

Bibliografía

- Anderson, D. and McNeill, G. (1992). Artificial neural networks technology. *Kaman Sciences Corporation*, 258(6):1–83. x, x, 13, 14
- Barber, C., Bockhorst, J., and Roebber, P. (2010). Auto-regressive hmm inference with incomplete data for short-horizon wind forecasting. In *Advances in Neural Information Processing Systems*, pages 136–144. 5
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons. 11
- Brockwell, P. J. and Davis, R. A. (2013). *Time series: theory and methods*. Springer Science & Business Media. 10
- Brownlee, J. (2011). *Clever algorithms: nature-inspired programming recipes*. Jason Brownlee. 17
- Cadenas, E. and Rivera, W. (2007). Wind speed forecasting in the south coast of Oaxaca, Mexico. *Renewable Energy*, 32(12):2116–2128. 21
- Cadenas, E. and Rivera, W. (2010). Wind speed forecasting in three different regions of Mexico, using a hybrid arima–ann model. *Renewable Energy*, 35(12):2732–2738. 12
- Casas, J. M., Gea, F., Javaloyes, E., Martín, A., Pérez, J., Triguero, I., and Vives, F. (2007). Educación medioambiental. *Alicante: Club universitario*. 8
- Chang, W. Y. (2013). Short-term wind power forecasting using the enhanced particle swarm optimization based hybrid method. *Energies*, 6(9):4879–4896. 1

- Chang, W. Y. (2014). A literature review of wind forecasting methods. *Journal of Power and Energy Engineering*, 2(4). 2, 9, 20
- Chollet, F. et al. (2017). Keras (2015). 44
- Choon, T. W., Prakash, C., Aik, L. E., and Hin, T. T. (2012). Development of low wind speed anemometer. *International Journal on Advanced Science, Engineering and Information Technology*, 2(3):237–240. 45
- Dayhoff, J. E. and DeLeo, J. M. (2001). Artificial neural networks. *Cancer*, 91(S8):1615–1635. 13
- Du, Y., Lu, J.-p., Li, Q., and Deng, Y. (2008). Short-term wind speed forecasting of wind farm based on least square-support vector machine. *Power System Technology*, 32(15):62–66. 12
- Flores, J. J., Loeza, R., Rodríguez, H., and Cadenas, E. (2009). Wind speed forecasting using a hybrid neural-evolutive approach. In *Mexican International Conference on Artificial Intelligence*, pages 600–609. Springer. 20
- Foundation, P. S. (2012). Python language reference. 49
- Fthenakis, V. and Kim, H. C. (2009). Land use and electricity generation: A life-cycle analysis. *Renewable and Sustainable Energy Reviews*, 13(6-7):1465–1474. 9
- Greenpeace (2015). About dirty energy. <https://www.greenpeace.org.uk/what-we-do/climate/energy/dirty-energy/>. Consultado: 2018-04-04. 8
- Gutiérrez, R. and Calderón, M. I. (2013). Estimación de la demanda de energía eléctrica en mendoza. *Trabajo Final, Facultad de Ciencias Económicas, Universidad Nacional de Cuyo*. 7
- Harik, G. R., Lobo, F. G., and Goldberg, D. E. (1999). The compact genetic algorithm. *IEEE transactions on evolutionary computation*, 3(4):287–297. 17
- Harrington, P. (2012). *Machine learning in action*, volume 5. Manning Greenwich, CT. 12

- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR. 14
- Izbicki, M. (2011). Converting images into time series for data mining. Accedido en 19-04-2018 a url <https://izbicki.me/blog/converting-images-into-time-series-for-data-mining.html>. 32
- Khashei, M., Hejazi, S. R., and Bijari, M. (2008). A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. *Fuzzy sets and systems*, 159(7):769–786. 22
- Kirbas, I. and Kerem, A. (2016). Short-term wind speed prediction based on artificial neural network models. *Measurement and Control*, 49(6):183–190. 3
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444. 12, 15
- Lei, M., Shiyang, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4):915–920. 20
- Ling-ling, L., Li, J.-H., He, P.-J., and Wang, C.-S. (2011). The use of wavelet theory and arma model in wind speed prediction. In *Electric Power Equipment-Switching Technology (ICEPE-ST), 2011 1st International Conference on*, pages 395–398. IEEE. 21
- Lopez Farias, R., Puig, V., Rodriguez Rangel, H., and Flores, J. J. (2018). Multi-model prediction for demand forecast in water distribution networks. *Energies*, 11(3):660. 2
- Lund, H. (2007). Renewable energy strategies for sustainable development. *Energy*, 32(6):912–919. 4
- Miranda, M. S. and Dunn, R. W. (2006). One-hour-ahead wind speed prediction using a bayesian methodology. In *Power Engineering Society General Meeting, 2006. IEEE*, pages 6–pp. IEEE. 21
- Mohandes, M., Halawani, T., Rehman, S., and Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Renewable Energy*, 29(6):939 – 947. 1

- Montgomery, D. C., Jennings, C. L., and Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons. 10, 11, 27
- Negnevitsky, M. (2002). *Artificial Intelligence*. 12, 16
- Observ, E. (2013). Worldwide electricity production from renewable energy sources, ninth inventory. *E. Observ.*, x, x, 3, 5, 6, 7
- Palomares-Salas, J., De La Rosa, J., Ramiro, J., Melgar, J., Aguera, A., and Moreno, A. (2009). Arima vs. neural networks for wind speed forecasting. In *Computational Intelligence for Measurement Systems and Applications, 2009. CIMSA'09. IEEE International Conference on*, pages 129–133. IEEE. 21
- Pérez-Espinosa, H., Avila-George, H., Rodriguez-Jacobo, J., Cruz-Mendoza, H. A., Martinez-Miranda, J., and Espinosa-Curiel, I. (2016). Tuning the parameters of a convolutional artificial neural network by using covering arrays. *Research in Computing Science*, 121:69–81. x, x, 15, 16, 37, 44, 51
- Piwko, D. and Jordan, G. (2010). The economic value of day-ahead wind forecasts for power grid operations. In *2010 UWIG Workshop on Wind Forecasting*. 5
- Rodríguez, H., Carrillo, N. A. G., Flores, J. J., Rosales, L. A. M., and Montelongo, G. M. (2017). Pronóstico a corto plazo de velocidad del viento a partir de datos incompletos. 22, 51
- Rodriguez, H., Flores, J., Puig, V., Morales, L., Guerra, A., and Calderon, F. (2017a). Wind speed time series reconstruction using a hybrid neural genetic approach. In *IOP Conference Series: Earth and Environmental Science*, volume 93, page 012020. IOP Publishing. 1
- Rodriguez, H., Puig, V., Farias, R. L., and Flores, J. J. (2017b). Short-term demand forecast using a bank of neural network models trained using genetic algorithms for the optimal management of drinking water networks. *Journal of Hydroinformatics*, 19(1):1–16. 5

- Sampedro, J. and García, J. D. (2012). Estudio y aplicación de técnicas de aprendizaje automático orientadas al ámbito médico: estimación y explicación de predicciones individuales. *EPS-UAM*. 13
- Sawin, J. L., Sverrisson, F., Seyboth, K., Adib, R., Murdock, H. E., Lins, C., Edwards, I., Hullin, M., Nguyen, L. H., Prillianto, S. S., et al. (2017). Renewables 2017 global status report. 4
- Sergio, A. T. and Ludermir, T. B. (2015). Deep learning for wind speed forecasting in northeastern region of brazil. In *Intelligent Systems (BRACIS), 2015 Brazilian Conference on*, pages 322–327. IEEE. 22
- Sivanandam, S. and Deepa, S. (2007). *Introduction to genetic algorithms*. Springer Science & Business Media. x, x, 17
- Van Rossum, G. et al. (2007). Python programming language. In *USENIX Annual Technical Conference*, volume 41, page 36. 44
- Wang, Z. and Oates, T. (2015). Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 40–46. 24, 34, 35, 46, 50
- Wolfram, S. (1999). Mathematica, wolfram media. 26, 27
- Yang, J., Nguyen, M. N., San, P. P., Li, X., and Krishnaswamy, S. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. In *IJCAI*, pages 3995–4001. 23
- Yesilbudak, M., Sagiroglu, S., and Colak, I. (2013). A new approach to very short term wind speed prediction using k-nearest neighbor classification. *energy conversion and management*, 69:77–86. 12
- Yu, G., Li, X., Song, D., Zhao, X., Zhang, P., Hou, Y., and Hu, B. (2016). Encoding physiological signals as images for affective state recognition using convolutional neural networks. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 812–815. IEEE. 15

Zhao, X., Wang, S., and Li, T. (2011). Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia*, 12:761–769. 20