ISSN 2992-7315
Año 2 No. 3 enero-junio 2024

RICT

Revista de Investigación Científica, Tecnológica e Innovación

Edición Semestral volumen 2, número 3, enero-junio 2024

Investigación científica, desarrollo tecnológico e innovación multidisciplinaria en los Centros de Cooperación Academia Industria para el desarrollo y la generación de soluciones a los problemas prioritarios en la sociedad, la salud, el medio ambiente, la seguridad y el bienestar en las organizaciones, la industria, la academia y la sociedad en general.

CCAI
Centro de Cooperación Academia Industria

ISSN 2992-7315
9 772992 731002

DESDE 2023

https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024)

ISSN: 2992-7315

Volumen 2 – No. 3 – 2024

Semestral

ISSN: 2992-7315

Centro de Cooperación Academia Industria – TESE

Investigación científica, desarrollo tecnológico e innovación multidisciplinaria en los Centros de Cooperación Academia Industria para el desarrollo y la generación de soluciones a los problemas prioritarios en la sociedad, la salud, el medio ambiente, la seguridad y el bienestar en las organizaciones, la industria, la academia y la sociedad en general.

Comité Editorial

RICT Revista de Investigación Científica, Tecnológica e Innovación, año 2, No. 3, Enero – Junio 2024, es una Publicación semestral, indexada y arbitrada a doble ciego editada Leonardo Miguel Moreno Villalba. Av. Tecnológico sn, Col. Valle de Anahuac Sección Fuentes, Ecatepec de Morelos, Estado de México, C.P. 55210, Tel. (01) 55 55002322, https://ccaitese.com, https://revista.ccaitese.com, ccai@tese.edu.mx. Editor responsable: Leonardo Miguel Moreno Villalba. Reserva de Derechos al Uso Exclusivo No. 04-2023-072413552000-102, ISSN: 2992-7315, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este Número Leonardo Miguel Moreno Villalba, Av. tecnológico sn, Col. Valle de Anahuac Sección Fuentes, Ecatepec de Morelos, Estado de México, C.P. 55210, fecha de última modificación 29 de septiembre de 2023.

DESDE 2023

https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024)

ISSN: 2992-7315

Directorio

Consejo Editorial (Editorial Board)

Comité Técnico Revisor (Reviewer Board)

DESDE 2023

https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024)

ISSN: 2992-7315

M. en C. Edgar Corona Organiche
Tecnológico Nacional de México / TES Ecatepec
Dr. Abraham Jorge Jiménez Alfaro
Tecnológico Nacional de México / TES Ecatepec
Dra. María de la Luz Delgadillo Torres
Tecnológico Nacional de México / TES Ecatepec
Dra. Adolfo Meléndez Ramírez
Tecnológico Nacional de México / TES Ecatepec
M. en ISC. Leonardo Miguel Moreno Villalba
Tecnológico Nacional de México / TES Ecatepec
Dra. Francisco Jacob Ávila Camacho
Tecnológico Nacional de México / TES Ecatepec
Dr. Derlys Hernández Lara
Tecnológico Nacional de México / TES Ecatepec
M. en C Israel Marcos Santillan Méndez
Tecnológico Nacional de México / IT de la Paz

Editorial

En la vanguardia de la investigación científica, tecnológica e innovación, la Revista RICT se enorgullece de presentar su más reciente edición: Vol. 2, No. 3, correspondiente al periodo enero-junio de 2024. Este número regular no solo refleja el compromiso continuo de nuestra revista con la excelencia académica y la divulgación del conocimiento, sino que también marca un hito en nuestra travesía para iluminar los avances más recientes y significativos en el ámbito de la ciencia y la tecnología.

A través de una cuidadosa selección y revisión, este número integra cinco investigaciones que destacan por su innovación, relevancia y contribución al conocimiento científico y tecnológico. Estos artículos representan el espíritu de nuestra revista: impulsar el progreso tecnológico y fomentar un espacio para el diálogo científico.

Esta nueva colección de investigaciones que no solo destacan por su rigor científico y relevancia tecnológica sino que también reflejan el compromiso inquebrantable de nuestra comunidad académica con el avance del conocimiento. La selección de artículos de este número encarna nuestro objetivo de ser un foro líder para el intercambio de ideas innovadoras y avances significativos en el ámbito de la ciencia y la tecnología.

La diversidad y profundidad de los temas tratados en este número son un testimonio de la riqueza de la investigación científica y tecnológica actual. Desde la aplicación de modelos predictivos avanzados en el ámbito de la salud hasta la exploración de soluciones de seguridad informática mediante criptografía asimétrica, cada artículo contribuye de manera significativa a su campo respectivo.

"Predicción de Hospitalización por COVID-19" ofrece un análisis exhaustivo de las tendencias de hospitalización, utilizando modelos predictivos avanzados para mejorar la gestión de recursos en tiempos de pandemia. Esta investigación no solo aporta al campo de la salud pública sino que también establece un precedente para el uso de la inteligencia artificial en la predicción y manejo de enfermedades.

El artículo "Protocolo Criptográfico de Firma Digital" propone una innovadora arquitectura de seguridad para la firma de documentos digitales, utilizando criptografía asimétrica para garantizar intercambios de información

seguros. Este estudio es fundamental para el avance de la seguridad informática en un mundo cada vez más digitalizado.

"Método Numérico de Cuadratura para la Resolución de Ecuaciones Diferenciales" se sumerge en la matemática aplicada para ofrecer una solución eficaz a problemas complejos de ecuaciones diferenciales, con implicaciones significativas para la ingeniería y la física.

Con "Arquitectura de Gestión de Red para la Monitorización y Control de Información", se aborda el desafío de la gestión eficiente de redes informáticas, presentando una arquitectura que facilita la monitorización y control en tiempo real de los nodos de red. Esta contribución es crucial para el desarrollo de infraestructuras de TI más robustas y seguras.

"Análisis del Sistema de Renta de Bicicletas" evalúa la sostenibilidad y eficiencia de los sistemas de transporte compartido, ofreciendo perspectivas valiosas sobre la movilidad urbana y su impacto ambiental.

El proceso de revisión por pares es fundamental para mantener la integridad y calidad de nuestra revista. En este sentido, deseamos expresar nuestro más sincero agradecimiento a los revisores que, con su tiempo, esfuerzo y pericia, han contribuido inmensamente a la excelencia de los artículos publicados en este número. Su dedicación no solo mejora la calidad de la investigación publicada sino que también fortalece la comunidad científica al fomentar un diálogo constructivo y riguroso. La labor de revisión, a menudo realizada de manera anónima, es un acto de generosidad académica que merece nuestro mayor reconocimiento y aprecio.

Esta edición de la Revista RICT se erige como una plataforma para la diseminación de conocimiento que, esperamos, inspire a investigadores, académicos y profesionales a seguir explorando, innovando y contribuyendo a la ciencia y tecnología. Extendemos nuestro agradecimiento a los autores por su dedicación y a nuestros lectores por su constante apoyo y curiosidad científica. Juntos, continuamos avanzando hacia un futuro más informado y tecnológicamente enriquecido.

A través de la publicación de este número, la Revista RICT continúa su misión de ser un catalizador para el intercambio de conocimiento y un puente hacia el futuro de la innovación tecnológica y científica. Invitamos a nuestros lectores a explorar estos artículos, reflexionar sobre sus implicaciones y participar en el diálogo científico que cada uno promueve.

Con cada publicación, nos inspiramos en la pasión y dedicación de nuestra comunidad de autores, revisores y lectores. Juntos, estamos trazando el camino hacia descubrimientos que modelarán el futuro de nuestra sociedad.

Equipo Editorial, Revista RICT

Francisco Jacob Ávila Camacho

Editor Asociado

DESDE 2023

https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024)

ISSN: 2992-7315

Contenido

Artículos de Investigación

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

ISSN: 2992-7315

Publicación Semestral RICT Vol. 2 No. 3 (2024) Páginas 1 – 5

Predicción de la condición de hospitalización para pacientes Covid-19 utilizando modelos de clasificación
Prediction of hospitalization condition for Covid-19 patients using classification models

Alberto Bautista-Loaiza[a], Francisco-Jacob Ávila-Camacho[b]

[a] *Maestría en Ingeniería en Sistemas Computacionales, Tecnológico de Estudios Superiores de Ecatepec, 55210, Ecatepec, Estado de México, México.*
[b] *División de Ingeniería en Sistemas Computacionales, Tecnológico Nacional de México / TES Ecatepec, 55210, Ecatepec, Estado de México, México*

Resumen

Se propone utilizar modelos de inteligencia artificial para predecir si un paciente con COVID-19 requerirá hospitalización basado en síntomas. Con el propósito de apoyar a los servicios de salud que sobrepasan su capacidad de atención en la pandemia. Se utiliza 13,757,682 registros de pacientes de México considerando 12 características que influyen en la evolución de la enfermedad, la recuperación de la información fue a través de la publicación de la dirección de epidemiología del 25 enero del 2022, el entrenamiento de los modelos se lleva a cabo con algoritmos de regresión logística y redes neuronales. Al término de diversos ajustes ambos modelos obtuvieron una precisión cercana al 80%. Se concluye que estos modelos deben considerarse para apoyo de diagnósticos médicos para determinar la necesidad de hospitalización de pacientes con COVID-19 en México. La metodología consistió en la recolección de datos, preprocesamiento, entrenamiento y evaluación del desempeño. Se propone que los modelos entrenados se incorporen en aplicaciones web para facilitar su uso en las áreas de salud.

Palabras clave: Covid-19, regresión logística, clasificación, predicción, machine learning.

Abstract

It is proposed to use artificial intelligence models to predict whether a patient with COVID-19 will require hospitalization based on symptoms. With the purpose of supporting health services that exceed their capacity to care in the pandemic. 13,757,682 patient records from Mexico are used, considering 12 characteristics that influence the evolution of the disease, the recovery of the information was through the publication of the Epidemiology Directorate on January 25, 2022, the training of the models is carried out carried out with logistic regression algorithms and neural networks. After various adjustments, both models obtained an accuracy close to 80%. It is concluded that these models should be considered to support medical diagnoses to determine the need for hospitalization of patients with COVID-19 in Mexico. The methodology consisted of data collection, preprocessing, training and performance evaluation. It is proposed that the trained models be incorporated into web applications to facilitate their use in health areas.

Keywords: Covid-19, logistic regression, classification, prediction, machine learning

1. Introducción

El campo de salud se ha visto beneficiadas a lo largo de la existencia de la inteligencia artificial, la cual ha logrado identificar patrones ocultos en el análisis de estudios de áreas como cardiología, genética, neurología, radiología, oncología, etc. (Ardakani, 2020).

La enfermedad COVID-19 causada por el coronavirus SARS COV 2 ha afectado a millones de personas al rededor del mundo, dado que es una enfermedad de la cual no se tienen mucha información por su corto tiempo de estudio se desconoce de manera específica como reaccionara en diversas personas, ya que no todas las personas tienden a tener los mismos síntomas ni en la misma magnitud lo que ha ocasionado saturación de hospitales con pacientes con

*Autor para la correspondencia: albertoblmsc@gmail.com
Correo electrónico: albertoblmsc@gmail.com (Alberto Bautista-Loaiza), fjacobavila@tese.edu.mx (Francisco-Jacob Ávila-Camacho).

Historial del manuscrito: recibido el 03/11/2023, última versión-revisada recibida el 13/12/2023, aceptado el 15/01/2024, en línea (postprint) desde el 02/02/2024, publicado el 09/04/2024. **DOI:** https://doi.org/10.2992/rict.v2i3.33

síntomas graves (Martínez-Ortega, 2019). Para apoyar en esta situación se considera desarrollar un modelo eficaz de clasificación para realizar predicciones sobre si un paciente con COVID-19 evolucionara de tal forma que requiera ser hospitalizado basado en algoritmos de aprendizaje automático entrenado con casos documentados en la república mexicana. para tal propósito se deberán tener en cuenta los siguientes puntos.

La predicción que llevará a cabo el modelo computacional será solo de apoyo para tomar precauciones, no sustituye la observación médica durante el padecimiento.

Cabe resaltar que el modelo se creara con casos de COVID-19 de la población mexicana por lo tanto un factor será la localidad, ya que el modelo debe considerar las características regionales y con ello una parte de su estilo de vida la cual debe ser lo más similar posible para la etapa de aprendizaje, por lo tanto, el modelo obtenido no podrá aplicarse a personas que no sean residentes de México (Epidemiología, 2022).

La creación de herramientas de este tipo tiene sus orígenes desde la antigüedad, No obstante, fue hasta a mediados del siglo XX que aparecieron herramientas tangibles que se podrían considerar maquinas con aprendizaje. (Ardakani, 2020)

Vaishya, Javaid, Khan, & AbidHaleemb (2020) publicaron un trabajo sobre aplicaciones con inteligencia artificial para la pandemia de COVID-19 puntualizan los principales usos de estas tecnologías en tiempos de pandemia de COVID-19, con estas herramientas se puede evaluar rápidamente síntomas anormales y alertar al servicio médico y con ello tomar decisiones en menor tiempo implementando Support Vector Machine (SVM). (Vaishya, 2020).

Una tarea para estas aplicaciones en una pandemia es apoyar en la creación de medicamentos y vacunas específicas para el virus, ya que en los últimos años la inteligencia artificial ha tomado popularidad en las investigaciones de fármacos mediante análisis de datos con algoritmos de reducción de dimensión junto con el análisis de la información disponible de COVID-19 (Escudero, 2021).

Otra implementación de esta tecnología es la que se emplea en la fase de pruebas en nuevos fármacos donde estas pruebas se realizan en tiempo real, cuando las pruebas tradicionales requerían mucho tiempo y trabajo esto a través de big data y la ciencia de datos (Díaz, 2020). Con la aparición de estas herramientas ayudaron a recortar el tiempo de pruebas de una forma significativa, trabajo que no podría ser realizado de la misma forma por una persona. Con el análisis de datos en tiempo real, los modelos brindan resultados que ayudan a prevenir la propagación de la enfermedad (Chávez Martínez, 2019). La publicación concluye que la inteligencia artificial se convertirá en un aliado indispensable en el futuro contra epidemias y pandemias, así como un gran apoyo en medidas preventivas y correctivas contra muchas otras enfermedades (Vaishya, 2020).

2. Materiales y Método

Para poder crear un modelo de predicción de pacientes COVID-19 se necesita una fuente de información confiable que contenga las principales características que influyen sobre la evolución de los pacientes con esta enfermedad y poder clasificar si el paciente será ambulatorio o requerirá hospitalización. Para llegar a los objetivos se tiene que seguir los siguientes puntos:

- Recabar información acerca de los casos asociados a COVID-19 en México.
- Elaborar una metodología para la exploración de datos con algoritmos de aprendizaje automático.
- Desarrollar y comparar el desempeño de clasificadores desarrollados con diferentes algoritmos de aprendizaje automático.
- Determinar si algún clasificador es adecuado para realizar un pronóstico confiable de que el paciente con Covid-19 requiera ser hospitalizado.

Desde los primeros casos de COVID-19 en México la Dirección General de Epidemiología emitió desde el 12 de abril del 2020 como materia de datos abiertos documentos CVS con casos asociados a personas atendidas por sospechas de COVID-19 con el propósito de facilitar a todos los usuarios que la requieran, el acceso, uso, reutilización y redistribución de la misma (Epidemiología, 2022).

Donde podemos encontrar múltiples características de los pacientes atendidos incluido Tipo Paciente el cual especifica si el paciente fue hospitalizado o fue un caso ambulatorio, Por tanto, podemos usar esta información para entrenar un modelo de predicción sobre qué tipo de paciente será la persona infectada, actualmente ya se tienen vacunas con lo que se concluye que la información de la base de datos contiene personas con y sin vacuna dependiendo la fecha del registro, Por lo tanto se tomarán en cuenta solo los datos a partir de que el sector salud notifico un esquema completo de vacunación en la población adulta mexicana el día 29 de octubre de 2021 y hasta a la fecha 25 de enero del 2022.

Dado que la predicción mantendrá una relación de dependencia entre las características del conjunto de datos es necesario identificarlas como variables dependientes e independientes para determinar que pacientes necesitarán hospitalización y que pacientes serán ambulatorios.

Las variables independientes son los factores de riesgo que posee una persona, es lo que lo hace susceptible a desarrollar síntomas graves, el sector salud ha mencionado cuales son dichos factores que determinan la evolución del COVID-19 (Gutiérrez, 2022), con esta información se identifican los campos del conjunto de datos que se emplean como variables independientes, serían los siguientes.

- ASMA
- CARDIOVASCULAR
- DIABETES
- EDAD
- EMBARAZO
- EPOC
- HIPERTENSION
- INMUSUPR

- OBESIDAD
- RENAL_CRONICA
- SEXO
- TABAQUISMO

Se tomaran para el entrenamiento las 12 variables independientes, para las variables tipo dependientes solo TIPO_PACIENTE ya que si una persona requirió cuidados intensivos así como si presentó un cuadro de neumonía se concluye que dicha persona requirió hospitalización , es por eso que solo la variable tipo paciente es la cual nos indica si la persona contagiada requirió hospitalización o no. Los posibles valores de cada una de las características del conjunto de datos se observan en la tabla 1.

Tabla 1. Diccionario de datos

Característica	Valor	Descripción	Tipo
TIPO_PACIENTE	0	Paciente no requiere ser hospitalizado	Dependiente
	1	Paciente requiere ser hospitalizado	
SEXO	0	Mujer	Independiente
	1	Hombre	
EDAD	0 a 99	Edad numérica del paciente	Independiente
EMBARAZO	0	Si embarazada	Independiente
	1	No embarazada	
DIABETES	0	Padece diabetes	Independiente
	1	No padece diabetes	
EPOC	0	Padece EPOC	Independiente
	1	No padece EPOC	
ASMA	0	Padece asma	Independiente
	1	No padece asma	
INMUSUPR	0	Padece INMUSUPR	Independiente
	1	No padece INMUSUPR	
HIPERTENSION	0	Padece hipertensión	Independiente
	1	No padece hipertensión	
CARDIOVASCULAR	0	Padece enfermedad Cardiovascular	Independiente
	1	No Padece enfermedad Cardiovascular	
OBESIDAD	0	Tiene obesidad	Independiente
	1	No tiene obesidad	
RENAL_CRONICA	0	Padece alguna enfermedad renal	Independiente
	1	No padece enfermedades renales	
TABAQUISMO	0	Es fumador	Independiente
	1	No es fumador	

Al observar la cantidad de datos de entrenamiento podemos notar que nos encontramos con información desbalanceada ya que 860,989 registros son de personas que no requirieron hospitalización y 36,403 de personas que si requirieron hospitalización.

Lo cual equivale a un 95.94 % No hospitalizadas y 4.05 % hospitalizadas del total de los datos.

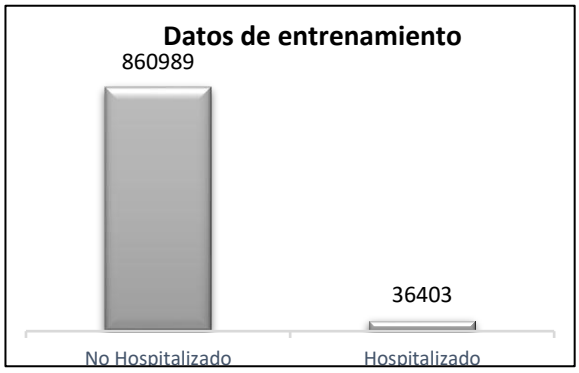

Figura 1. Tipo de pacientes.

Con el IDE Spyder que nos permite utilizar las librerías scikit-learn en Python se crean modelos de clasificación con el conjunto de datos

Se realiza un submuestreo 1 a 1 aleatorio sobre los datos de personas no hospitalizadas que representan el 95.94 % de la información total para obtener un conjunto de datos balanceado posteriormente el 75 % de la información para entrenamiento y el 25 % restante para realizar pruebas.

Para aquellos registros que tienen valores faltantes se les considera la media del campo y se ajustan los datos de entrenamiento como de pruebas a una escala similar para terminar el preprocesamiento de datos.

2.1. Entrenamiento y predicción con regresión logística

Para la correcta elección de algoritmos en la clasificación binaria se evalúan los algoritmos Regresión Lineal, Regresión Logística, Arboles de Decisión, SVR, Clustering, Redes Neuronales y Naive Bayes , que se usan comúnmente para este tipo de aplicaciones, en base a los resultados se observa que los algoritmos regresión logística y redes neuronales obtienen un mejor desempeño por lo cual son estos dos algoritmos a los que se les da continuidad en su configuración.

La Regresión Logística es un Algoritmo Supervisado y se utiliza para la clasificación binaria (González Segoviano, 2023), partiendo de nuestros datos de entrenamiento y prueba procedemos a crear nuestro modelo con la función LogisticRegression que nos proporciona sklearn (AWS, 2023). Para poder visualizar el desempeño del algoritmo de regresión logística obtenemos la matriz de confusión que nos permitirá observar las predicciones correctas e incorrectas de cada clase. En la siguiente tabla se observa la matriz de confusión con la relación entre valores de predicción y valores reales

Tabla 2. Matriz de confusión de predicción aplicando regresión logística

Predicción/Reales	Verdaderos	Falso
Verdadero	7550	1467
Falso	2762	6423

2.2. Entrenamiento y predicción con redes neuronales

Las redes neuronales artificiales son un algoritmo que se puede implementar para la clasificación supervisada binaria, dada su buena fama para aplicaciones médicas implementamos una predicción con este algoritmo (Céspedes, 2021), partiendo de nuestros datos de entrenamiento y con ayuda de Keras la cual es una librería de redes neuronales escrita en de Python (APD, 2019).

Se construye una red neuronal profunda de 3 capas, se declara con el método dense, 12 neuronas en la capa de entrada, 24 neuronas en la capa oculta y una neurona de salida, la función de activación relu para las capas de entrada como oculta y sigmoid para la capa de salida.

Se realiza la compilación de la red neuronal con *binary_crossentropy* como función de perdida que se utiliza para evaluar el grado de error entre salidas calculadas y las salidas deseadas de los datos de entrenamiento. un optimizador Adam y se realiza el entrenamiento con una cantidad de iteraciones de 250. En la siguiente tabla se muestra la matriz de confusión del modelo con redes neuronales una vez realizada la predicción.

Tabla 3. Matriz de confusión aplicando redes neuronales

Predicción/Reales	Verdaderos	Falso
Verdadero	7735	1282
Falso	2801	6384

3. Resultados

En la tabla 4 compararemos los resultados de la matriz de confusión, de cada uno de los algoritmos de regresión logística y redes neuronales.

Tabla 4.Comparación de matriz de confusión de ambos algoritmos

Predicción	Regresión logística	Redes neuronales
Verdaderos positivos	7550	7735
Verdaderos negativos	6423	6384
Falsos positivos	2762	2801
Falsos negativos	1467	1282

Las mediciones obtenidas de ambos modelos se observan en la tabla 5.

Tabla 5. Comparación de las curvas ROC de ambos modelos

Indicador	Regresión logística	Redes neuronales
Precisión	0.8113541534960172	0.8327680667884164
Exactitud	0.7659597846390507	0.7756839907702451
Sensibilidad	0.6986390854654327	0.6950462710941753
Puntaje f1	0.7507897507897509	0.7576998397721203

Se crean las gráficas ROC de ambos modelos que se puede apreciar en la figura 2 las cuales se calculan a partir de los resultados obtenidos y con ayuda de la librería matplotlib.

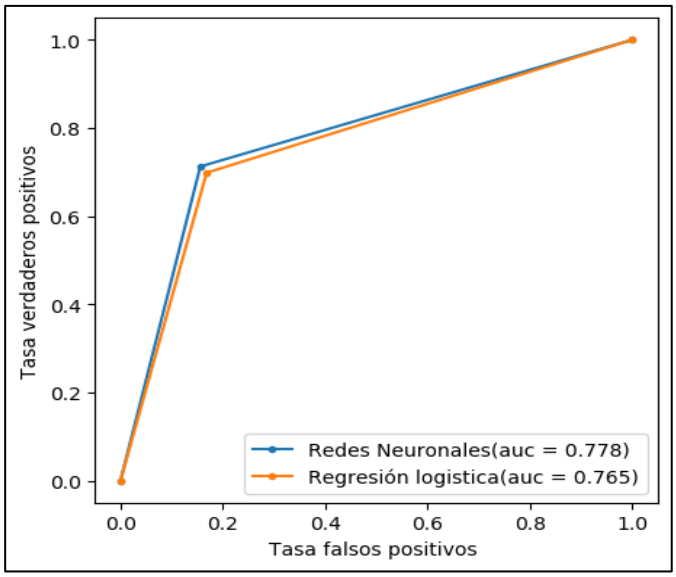

Figura 2. Comparación de las curvas ROC de ambos modelos.

4. Discusión

En el algoritmo de regresión logística no hubo cambios significativos al modificar algún parámetro de entrenamiento. Para el algoritmo de redes neuronales, se realizaron diversas pruebas, la configuración con el mayor desempeño fue con las siguientes características, la capa de entrada en 12, la capa oculta en 24, y la cantidad de interacciones 250, con esta configuración se obtuvo el resultado para redes neuronales mostrado en la tabla 5.

Para esta implementación obtenemos que el algoritmo de clasificación de redes neuronales tiene un mejor desempeño al predecir si un paciente requerirá hospitalización, aunque la ventaja de este algoritmo sobre el de regresión logística no es significativo, siendo ambos óptimos para el objetivo ya que como se aprecia en la Figura 2 la diferencia del área bajo la curva de ambos algoritmos regresión logística y redes neuronales es mínima. Esta razón fue calculada en base al desempeño de cada modelo, a partir del resultado de la razón de verdaderos positivos que se obtiene a partir de la formula:

$$VPR = \frac{VP}{P} = \frac{VP}{VP+FN}$$

Y de la razón de falsos positivos de la formula:

$$FPR = \frac{FP}{N} = \frac{FP}{FP+VN}$$

Con el resultado de las operaciones se puede dibujar la curva ROC y con ello el área bajo la curva de los modelos (Darlington, 2020). Los resultados son los siguientes, para regresión logística el AUC = 765 y para redes neuronales AUC = 778 lo cual es levemente mayor al área bajo la curva de clasificador entrenado con el algoritmo de regresión logística.

5. Conclusiones

Con la generación de clasificadores con ambos algoritmos se determina que se alcanza los objetivos propuestos ya que con la información recolectada se tuvo la capacidad para crear modelos de clasificación de personas infectadas con la enfermedad de COVID-19, con el conjunto de datos creado a partir del preprocesamiento junto con las técnicas de inteligencia artificial lograron generar clasificadores con un 80% de precisión lo cual se considera un modelo que puede ser probado en pacientes futuros al igual como en la literatura expuesta cuando se aplicaron las herramientas con un 70% de efectividad en diversos diagnósticos en la aplicación de algoritmos de regresión y clasificación.

Para incrementar la efectividad de los modelos sería necesario conocer otras características de los pacientes por ejemplo desde cuando padece alguna enfermedad de las mencionadas o días transcurridos desde el primer síntoma de COVID-19, con esta nueva información complementaria se tendrá un incremento en la precisión de los modelos.

Como se aprecia en los resultados todos los clasificadores mostraron un rendimiento similar en las diversas métricas evaluadas, esto por un lado se debe a la naturaleza del conjunto de datos la cual no mostró problemas extremos aun cuando la totalidad de la información presenta un claro desbalanceo, pero sin inexistencia de valores atípicos o alguna dominancia marcada de alguna clase. En base a los resultados obtenidos se determina que la implementación de clasificadores desarrollados con los algoritmos de regresión logística y redes neuronales son confiables para determinar la evolución de pacientes, siempre y cuando la información recolectada sea suficiente ya que para un buen aprendizaje los clasificadores basan su efectividad dependiendo en gran medida de la información con la cual se les alimente y por supuesto a la incorporación de los avances tecnológicos en la inteligencia artificial que se tienen día con día, Esto se reflejara en herramientas de diagnóstico cada vez más innovadoras enfocadas en ayudar al sector salud y enfrentar cualquier tipo de enfermedad.

6. Trabajos futuros

Como continuación de este trabajo de tesis en un futuro próximo a fin de realizar predicciones se plantea la posibilidad de incorporar el modelo clasificación en una aplicación web los cuales son sencillos de utilizar y no requieren conocimientos avanzados de informática, además de ser personalizables a gusto y adaptarse a la mayoría de formas de trabajo actuales donde lo único que tendría que realizar el operador es capturar las características del paciente para obtener el resultado de hospitalizado o ambulatorio sin necesidad de algún conocimiento extra.

Un trabajo que se debe realizar periódicamente es actualizar la información del conjunto de datos con las diversas publicaciones de la dirección general de epidemiología sobre la base de datos publica de COVID-19 ya que recordemos que cada día hay nueva información incluso durante el desarrollo de este trabajo ha surgido información importante de diversas fuentes sobre COVID-19 en México, una continuación al proyecto es buscar mejorar el desempeño de los modelos de clasificación con una retroalimentación sobre el conjunto de datos con más registros o incluso incorporando más características que ayuden a determinar la evolución de la enfermedad en el paciente cada vez con más precisión.

7. Referencias

APD. (04 de 04 de 2019). *APD.* Obtenido de https://www.apd.es/algoritmos-del-machine-learning/

Ardakani, A. A. (30 de 03 de 2020). *National library of medicine.* Obtenido de https://www.sciencedirect.com/science/article/pii/S0010482520301645

AWS. (01 de 01 de 2023). *AWS.* Obtenido de https://aws.amazon.com/es/what-is/logistic-regression/#:~:text=La%20regresi%C3%B3n%20log%C3%ADstica%20es%20una,factores%20bas%C3%A1ndose%20en%20el%20otro.

Céspedes, F. A. (28 de 12 de 2021). *Facultad de Ingeniería de Sistemas e Informática - UNMSM.* Obtenido de https://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/download/21862/17543/75813

Chávez Martínez, R. (2019). *Universidad de LIma*. Obtenido de https://repositorio.ulima.edu.pe/handle/20.500.12724/8401

Cid, A. S. (29 de 10 de 2021). *EL PAÍS.* Obtenido de https://elpais.com/mexico/2021-10-29/mexico-concluye-su-plan-de-vacunacion-un-83-de-la-poblacion-mayor-de-edad-tiene-al-menos-la-primera-dosis.html

Darlington, K. (22 de 05 de 2020). *BBVA Open Mind.* Obtenido de https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/esta-ayudando-la-inteligencia-artificial-contener-la-pandemia-covid-19/

Díaz, J. E. (29 de 07 de 2020). *Universidad de Cundinamarca*. Obtenido de https://revistes.ub.edu/index.php/RBD/article/view/31643

Epidemiología, D. G. (25 de 01 de 2022). *Secretaría de Salud.* Obtenido de https://www.gob.mx/salud/documentos/datos-abiertos-152127

Escudero, X. (24 de 03 de 2021). *Archivo de cardeologia de México.* Obtenido de https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-99402020000500007

Etecé, E. (05 de 08 de 2021). *Editorial Etecé.* Obtenido de https://concepto.de/diagnostico/

González Segoviano, L. J. (30 de 9 de 2023). Regresión logística vs árboles de decisión en el riesgo crediticio. (T. e. RICT Revista de Investigación Científica, Ed.) *RICT Revista de Investigación Científica, Tecnológica e Innovación*, *1*(2), 32-37. Obtenido de https://revista.ccaitese.com/index.php/ridt/article/view/21

Gutiérrez, V. F. (2022). *Ecografía en el manejo del paciente crítico con infección por SARS-CoV-2 (COVID-19): una revisión narrativa.* Obtenido de https://www.sciencedirect.com/science/article/pii/S0210569120301558?via%3Dihub

Martínez-Ortega, A. G. (22 de 12 de 2019). *Pro Sciences.* Obtenido de https://d1wqtxts1xzle7.cloudfront.net/87614928/287162093-libre.pdf?1655404808=&response-content-disposition=inline%3B+filename%3DTecnologias_en_la_inteligencia_artificia.pdf&Expires=1695936710&Signature=HNx9yigjuH3nIQY0ZJx8G5kahLtl7svfwgOCVc72rlUBspJf0K9h

Novás, J. D. (2006). *Revista Cubana de Medicina General Integral.* Obtenido de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252006000100007#:~:text=Cuando%20unimos%20los%20s%C3%ADntomas%20y,o%20de%20otra%2C%20cu%C3%A1les%20son

Shibly, K. H. (2020). *sciencedirect.* Obtenido de https://www.sciencedirect.com/science/article/pii/S2352914820305554

Vaishya, R. J. (4 de 08 de 2020). *Facultad de Ingeniería de Sistemas e Informática - UNMSM.* Obtenido de https://www.sciencedirect.com/science/article/pii/S1871402120300771

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

ISSN: 2992-7315

Publicación Semestral RICT Vol. 2 No. 3 (2024) Páginas 6 – 11

Protocolo criptográfico de firma digital para el signado de documentos digitales con criptografía asimétrica para el intercambio seguro de información en la empresa CDS, S.C.

Cryptographic digital signature protocol for signing digital documents with asymmetric cryptography for the secure exchange of information in the company CDS, S.C.

Abraham-Jorge Jiménez-Alfaro [a], Edgar Corona-Organiche [a], Griselda Cortés-Barrera [a], Irving-Cardiel Alcocer-Guillermo [b]

[a]Ingeniería en Sistemas Computacionales, TECNM/Tecnológico de Estudios Superiores de Ecatepec, Valle de Anáhuac, 55210 Ecatepec de Morelos, Estado de México.
[b]Ingeniería en Tecnologías de Información y Comunicaciones , TECNM/Instituto Tecnológico Gustavo A. Madero, Calle 608 No. 300 y Av. 412, Col. San Juan de Aragón, 07470 Alcaldía. Gustavo A. Madero, Ciudad de México.

Resumen

La firma digital corresponde a la versión computarizada de la firma personal manuscrita o firma ológrafa. Estas se utilizan ampliamente como prueba de autoría o acuerdo de una parte, entre otros usos, siempre en referencia a un documento. Sin embargo para la empresa CDS, S.C., pueden resultar inseguras en tanto que pueden ser, con relativa facilidad, aplicadas de manera deshonesta. Por ejemplo, pueden ser tomadas desde una pieza de papel para pasarlas a otra, o los documentos podrían ser modificados luego de la aplicación de la firma. La firma digital y los protocolos criptográficos como metodología proporcionan tecnología que apoya para resolver estos problemas en documentos digitales. El artículo presenta la arquitectura de un protocolo criptográfico de firma digital para signar documentos en la empresa CDS, S.C., así como, los aplicativos de encriptado y desencriptado que tendrán los mandos de alta dirección y medios de la organización para garantizar la prueba de autoría de un documento digital.

Palabras clave: Firma Digital, Cifrado Clave Publica, Protocolo de Firma Digital

Abstract

The digital signature corresponds to the computerized version of the personal handwritten signature or holographic signature. These are widely used as proof of authorship or agreement of a party, among other uses, always in reference to a document. However, for the company CDS, S.C., they can be unsafe since they can be, with relative ease, applied dishonestly. For example, they can be taken from one piece of paper to another, or documents could be modified after the signature has been applied. Digital signature and cryptographic protocols as a methodology provide technology that supports solving these problems in digital documents. The article presents the architecture of a cryptographic digital signature protocol to sign documents in the company CDS, S.C., as well as the encryption and decryption applications that the organization's senior management and media will have to guarantee proof of authorship. of a digital document.

Keywords: Digital Signature, Public Key Encryption, Digital Signature Protocol

1. Introducción

Los protocolos son una serie de pasos que envuelven a dos o más partes, diseñados para realizar una tarea. En particular, si usan un algoritmo criptográfico se denominan protocolos criptográficos (Stallings, 2017).

Para Stallings (2019) los algoritmos criptográficos por sí solos no cumplen con la función de resolver los problemas de seguridad; deben formar parte de un protocolo criptográfico.

Por otra parte, supone también que:

- Las partes involucradas deben conocer el protocolo.

*Autor para la correspondencia: ajimenez@tese.edu.mx
Correo electrónico: ajimenez@tese.edu.mx (Abraham-Jorge Jiménez-Alfaro), ecorona@tese.edu.mx (Edgar Corona-Organiche), gcortes@tese.edu.mx (Griselda Cortés-Barrera), irving.ag@gamadero.tecnm.mx (Irving-Cardiel Alcocer-Guillermo)

Historial del manuscrito: recibido el 25/01/2024, última versión-revisada recibida el 03/04/2024, aceptado el 04/04/2024, en línea (postprint) desde el 08/04/2024, publicado el 09/04/2024. **DOI:** https://doi.org/10.2992/rict.v2i3.43

- Las partes lo aceptan y concuerdan en aplicarlo.
- El protocolo no tiene ambigüedades.
- El protocolo es completo, para toda situación se contempla una acción determinada.

Para Craig (2003) Todas estas características son fundamentales para permitir que:

- Las comunicaciones entre computadoras sean seguras, al tener queseguir un protocolo formal para el intercambio de mensajes.

- Al especificar los pasos que se deben seguir, sea posible examinar en detalle si existen puntos débiles en cuanto a la seguridad; y por otro lado, evitar realizar acciones fuera del protocolo con intenciones delictivas.

Para Stallings (2019) los problemas de seguridad de las redes pueden dividirse de forma general en cuatro áreas interrelacionadas:

1.-**El secreto**, encargado de mantener la información fuera de las manos de usuarios no autorizados.

2.-**La validación de identificación**, encargada de determinar la identidad de la persona o computadora con la que se esta hablando.

3.-**El control de integridad**, encargado de asegurar que el mensaje recibido fue el enviado por la otra parte y no un mensaje manipulado por un tercero.

4.-**El no repudio**, encargado de asegurar la "firma" de los mensajes, de igual forma que se firma en papel una petición de compra/venta entre empresas.

Las firmas manuales en los documentos se usan desde tiempos inmemoriales como prueba de autoría, o al menos de consentimiento con el contenido del documento. Para Stallings (2019) existen algunas características que las hacen tan confiables:

- **La firma es inolvidable e irrepudiable**. El firmante no puede aducir que nosabe si es su firma, o negarla.
- **La firma es auténtica**. El que recibe el documento está convencido de que elfirmante deliberadamente firmó el documento.
- **La firma no es reusable**. Es parte del documento y no se puede mover o copiar a otro.
- **La firma es inalterable**.

Se asume que una firma realizada por otro medio distinto al manual, pero quecumpla con estas características, es confiable y puede ser aceptada por las partes. Es el caso de la firma efectuada en un medio digital.

2. Materiales y Método

2.1.- Criptografía y Criptoanálisis

El criptoanálisis (Maiorano, 2009) se encarga de descifrar los mensajes, la criptografía busca métodos más seguros de cifrado, la criptografía viene del griego KRYPTOS = oculto y GRAPHE = escrito. Para Stallings (2017) se clasifica en:

- Criptografía clásica. Algoritmo secreto. Cifrados por sustitución y transposición, entre otros.
- Criptografía moderna. Algoritmo público. Cifrados en base a claves que se mantienen secretas.

Para Maiorano (2019) el cifrado y descifrado constan de una serie de etapas para que un texto normal se cifre y descifre para garantizar la seguridad de información, ver figura 1.

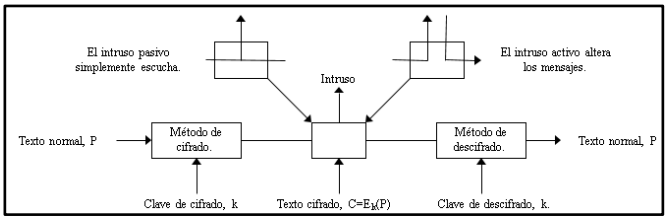

Figura 1.- La encriptación o cifrado.

- El texto normal (P) se transforma (cifra) mediante una función que tiene como parámetro una clave k.
- C = Ek(P) es el texto cifrado (C) obtenido a partir de P, usando la clave k y la función matemática Ek para codificar.
- P = Dk(C) es el descifrado de C para obtener el texto normal P.

Para el descifrado se necesita la inversa de la función matemática descrita como:
Dk(Ek(P)) = P dónde :

- E y D son sólo funciones matemáticas parametrizadas con la clave k
- Estas funciones E() y D( ) son conocidas por el criptoanalista, pero no la clave.

2.2.- Algoritmos del Protocolo Criptográfico

2.2.1.- Algoritmos Simétricos

Llamados algoritmos de una clave. Estos algoritmos

usan para cifrar y descifrar mensajes con la misma clave. En algunos casos pueden tener una clave para cada operación, pero se puede deducir una clave de la otra (Stallings, 2019).

Algunos protocolos utilizan la figura de un árbitro; es una tercera parte desinteresada y confiable, que garantiza a las partes involucradas el cumplimiento de un protocolo. Es desinteresada pues no tiene intereses particulares para intervenir en el protocolo, y es confiable pues las partes toman como honestas las acciones que realiza.

El árbitro otorga mayor seguridad en algunos protocolos. Sin embargo, también genera algunos problemas:

- Demoras en la ejecución del protocolo, ya que se agrega la transmisión de datos al árbitro, al total de mensajes transmitidos entre las partes.

- Cuello de botella en el computador del árbitro.

El árbitro es otro punto factible de atacar para alguien que desee quebrar elprotocolo.
El emisor y el receptor deben acordar la clave que usarán. El emisor la usa para encriptar el mensaje plano y el receptor para desencriptarlo, ver figura 2.

El protocolo es:

1. **A y B concuerdan un algoritmo simétrico.**
2. **A y B concuerdan una clave.**
3. **A encripta el mensaje con el algoritmo y la clave seleccionados.**
4. **A envía el mensaje cifrado a B.**
5. **B desencripta el mensaje cifrado con el algoritmo y la clave seleccionados.**

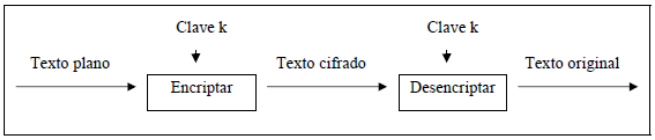

Figura 2.- Cifrado Simétrico.

2.2.2- Algoritmos de Clave Pública

Para Liu (2000) Estos algoritmos usan una clave para encriptar el mensaje claro, llamada clave pública, y otra para desencriptarlo, llamada clave privada o secreta. Esta clave no se puede deducir de la clave pública, ver figura 3.

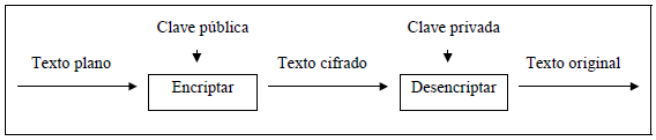

Figura 3.- Cifrado Asimétrico.

Este método se puede comparar con una casilla de correos; cualquier personapuede dejar una carta en la casilla de A, pero sólo A tiene la llave para abrirla y leer las cartas.

El protocolo es:

1. **A y B concuerdan un algoritmo de clave pública.**
2. **A recibe la clave pública de B (k_B).**
3. **A encripta el mensaje con el algoritmo seleccionado y la clave pública de B(k_B).**
4. **A envía el mensaje cifrado a B.**
5. **B desencripta el mensaje cifrado con su clave privada**.

2.3.- Algoritmo Híbrido del Protocolo de Cifrado

En algunos casos, es conveniente combinar los dos tipos de algoritmos; para encriptar mensajes largos los simétricos son más rápidos, y para distribuir las claves que usa el algoritmo asimétrico, los algoritmos de clave pública son más seguros (Stallings, 2017).

El protocolo es:

1. **A y B concuerdan un algoritmo simétrico y uno de clave pública.**
2. **B recibe la clave pública de A (k_A).**
3. **B encripta la clave que usará en el algoritmo simétrico (k_B) con k_A.**
4. **B envía a A k_B encriptada.**
5. **A desencripta con su clave privada a k_B.**
6. **A encripta el mensaje con k_B.**
7. **B desencripta el mensaje con k_B.**

Para Maiorano(2009) desde el punto de vista práctico los pasos son:

1.- Se conoce solo el texto cifrado, ver figura 4.

Figura 4.- Cifrado Asimétrico.- Texto Cifrado.

2.- Conoce un texto cifrado y el texto normal al que pertenece, ver figura 5.

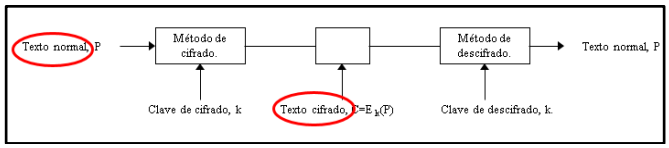

Figura 5.- Cifrado Asimétrico.- Textos.

3.- Dispone del sistema de cifrado. Puede escoger un texto normal y cifrarlo, ver figura 6.

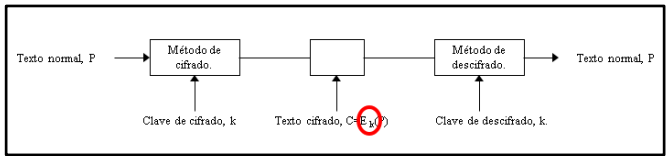

Figura 6.- Cifrado Asimétrico.- Sistema de Cifrado.

3. Resultados

Para Stallings (2019) La idea es usar el mecanismo de clave pública y clave privada para firmar documentos electrónicos. Algunos algoritmos de clave pública, como el RSA, sirven tanto para encriptar como para firmar mensajes.

RSA es el algoritmo de cifrado asimétrico más popular en la actualidad. Creado por Ron Rivest, Adi Shamir y Leonard Adleman –notar que son las iniciales de los apellidos las que forman el nombre del algoritmo– fue publicado en el año 1977.Actualmente el algoritmo es considerado seguro, en tanto sean utilizadas llaves de longitud suficientemente seguras (se siguen utilizando llaves de 1 024 bits, pero ya se recomienda al menos una longitud de 2 048). El algoritmo sirve tanto para encriptar y desencriptar, como para la generación de firmas digitales. Es, en la actualidad, ampliamente utilizado en protocolos de comercio electrónico, entre otras aplicaciones (Maiorano, 2009).

Es decir que para firmar sólo encripta mensajes (con su clave privada), y cualquiera puede desencriptarlos con la clave pública. El mensaje encriptado es la firma del mensaje, ver figura 7, ya que sólo el dueño de la clave privada pudo generarlo (Stallings, 2019).

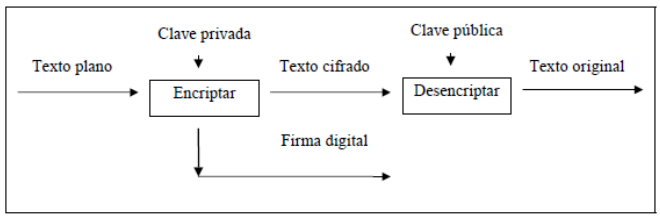

Figura 7.- Protocolo Criptográfico para le generación de firma Digital.

El protocolo es:

1. A concuerdan un algoritmo de firma digital con clave pública.

2. A firma el mensaje con su clave privada kA.

3. A envía a B el mensaje firmado.

4. B desencripta el mensaje con la clave pública de A y verifica la firma.

Se puede observar que se cumplen las características que tiene una firma manual:

- **La firma es inolvidable e irrepudiable, pues A y sólo A conoce la clave privada, y B demuestra que A lo firmó con la clave pública.**
- **La firma es auténtica. B lo verifica con la clave pública de A.**
- **La firma no es reusable, ya que es función del mensaje.**
- **La firma es inalterable, pues si cambia el mensaje, ya no concuerda con la firma.**

La firma digital:

- **Debe ser fácil de generar.**
- **Será irrevocable, no rechazable por su propietario con el acuse de recibo.**
- **Será única, sólo posible de generar por su propietario.**
- **Será fácil de autenticar o reconocer por su propietario y los usuarios receptores.**
- **Debe depender del mensaje (por compendio) y del autor (por certificado).**

Para Maiorano (2009) y Stallings (2019) la firma Digital se genera entonces como sigue:

Supongamos los algoritmos públicos tal que:

E(D(P)) = P

D(E(P)) = P

El algoritmo de cifrado E( ), descifrado D( ) y la clave de cifrado, se hacen públicos (de ahí el nombre de criptografía de clave pública), **pero se mantiene secreta la clave de descifrado , ver figura 8. E_A es clave pública y D_B es clave secreta** (Stallings, 2019).

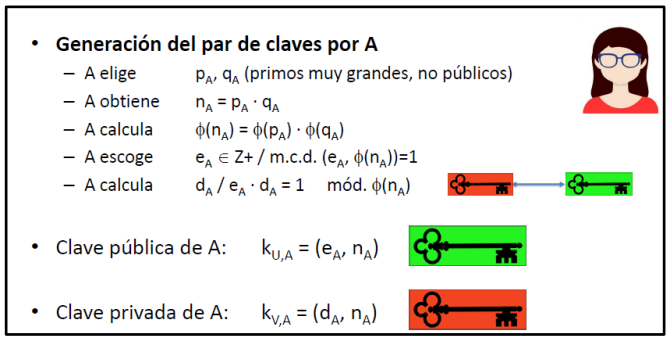

Figura 8.- Etapas del Protocolo Criptográfico para la generación de firma Digital.

Basándose en la figura 8, el algoritmo del **Protocolo Criptográfico** se basa en factorizar números grandes:

1. Seleccionar dos números primos grandes, p y q (generalmente mayores que 10^{100} → 1024 bits).
2. Calcular:
   i. ***n = p * q***
   ii. ***z = (p-1) * (q-1) la* función multiplicativa de Euler.**
3. Seleccionar un número d primo relativo con z (sin ningún factor común).
4. Encontrar e tal que ***((e * d) modulo z) = 1.***
5. **Los datos que serán públicos son el par *(e,n)* y privados *(d,n).***

La figura 9 presenta la captura del texto que firmara el documento de acuerdo al algoritmo del **Protocolo Criptográfico: Gerente Hugo Jimenez 737 ID 05**.

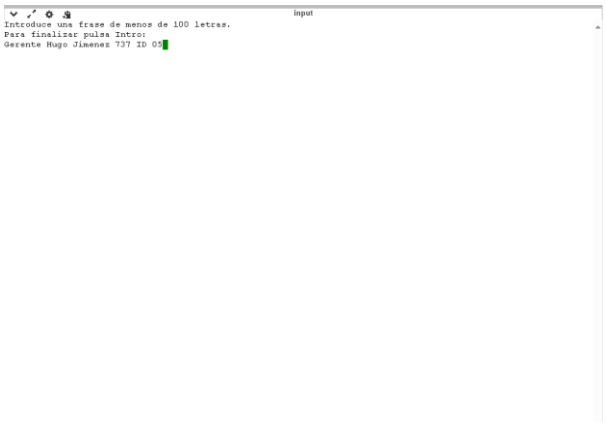

Figura 9.- Texto a firmar en las Etapas del Protocolo Criptográfico para generar el cifrado de firma Digital.

La figura 10 presenta el cifrado y generación de la firma digital de acuerdo con el algoritmo del protocolo criptográfico: **D-A-24-A-1C-28-A-35-F-2A-E-1E-35-13-12-1A-A-1C-A-34-35-3D-39-3D-35-11-7-35-36-3B.**

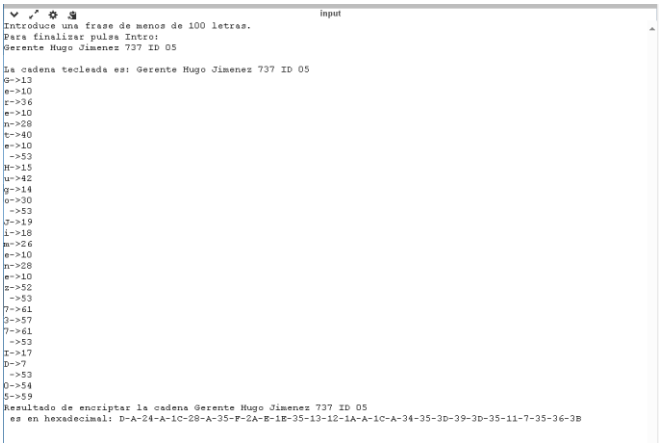

Figura10.- Ejecución de las Etapas del Protocolo Criptográfico para generar el cifrado de firma Digital.

Si el criptoanalista pudiera factorizar *n* (conocido públicamente), podría encontrar *p* y *q*, y a partir de éstos, *z*. Equipado con el conocimiento de *z* y de *e*, que es pública, puede encontrar *d* usando el algoritmo extendido de Euclides:

d = ((Y*z) +1)/e para Y=1,2,3,... hasta encontrar un d entero.

La figura 11 presenta el descifrado de firma digital de acuerdo al algoritmo del protocolo criptográfico: **Gerente Hugo Jimenez 737 ID 05**.

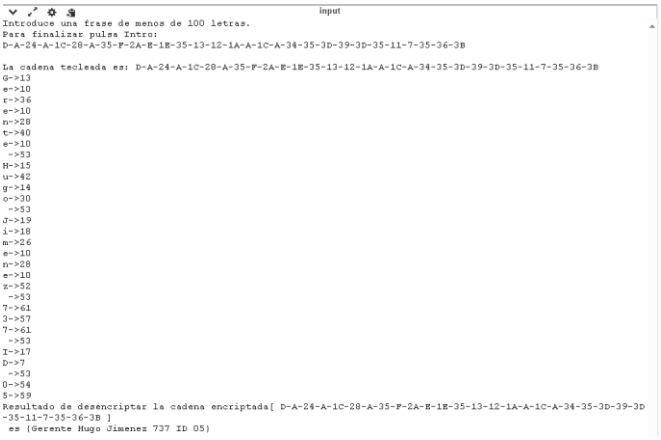

Figura11.- Ejecución de las Etapas del Protocolo Criptográfico para generar el descifrado de firma Digital.

4. Discusión

Para Maiorano (2009) la firma digital corresponde a la versión Computarizada de la firma personal manuscrita o firma ológrafa; esto es, se usará para probar la autoría de, o el acuerdo a, la información contenida en un documento electrónico. Existen diferentes protocolos para implementar esta funcionalidad criptográfica. La implementación más utilizada involucra la utilización de funciones hash junto con el protocolo de cifrado con criptografía asimétrica. Básicamente, la parte autora o firmante del documento firmará el hash resultante (Stalling, 2019).

Esto es: La parte "A" producirá el hash del documento, lo encriptará con su llave privada y enviará esto, junto con el documento, a la parte "B". "B" computará por su cuenta el hash sobre el documento. Luego, mediante la llave pública de "A", desencriptará el hash que "A" ha computado. Entonces, "B" podrá comparar ambos hashes y verificar la firma.

5. Conclusiones

El protocolo criptográfico de firma digital al emplear la encriptación y las firmas en el documento lograr mayor seguridad; el protocolo criptográfico permite segurar el signado de documentos en forma segura al establecer un protocolo entre las partes involucradas considerando las siguientes fases:

1. **A y B concuerdan un algoritmo de firma digital (con las claves pública f_A y laprivada F_A) y otro de clave pública (con las claves pública k_B y privada K_B).**
2. **A firma el mensaje con su clave privada (F_A).**
3. **A encripta el mensaje y la firma con la clave pública de B (k_B) y se los envía B.**
4. **B desencripta el mensaje y la firma con su clave privada (K_B).**
5. **B genera el hash a partir del mensaje desencriptado.**
6. **B desencripta la firma recibida con la clave pública de A (f_A), y lo comparacon el hash generado anteriormente. Si son iguales, la firma es válida.**

6. Agradecimientos

A la empresa CD, S.C., por el apoyo técnico y documental, así como, al área de Recursos Humanos por el acceso a las claves con las que signan cada uno de los miembros de la organización. A Cada uno de los integrantes de artículo por el aporte computacional y matemático para realizar los aplicativos.

7. Referencias

Craig H.(2003). Networking Personal Computers, whit TCP/IP. O'Relly Associates, Inc.Sebastopol, CA 95472.

Liu C.(2000). Managing Internet Information Services. O'Relly Associates, Inc.Sebastopol, CA 95472

Maiorano, A. (2009). Criptografía: técnicas de desarrollo para profesionales. México: Alfaomega.

Stallings, W. (2017). Fundamentos de Seguridad en Redes: Aplicaciones y Estandáres. México: Pearson, Prentice Hall.

Stallings, W. (2019). Criptografía y Seguridad de Red: Principios y Práctica. México: Pearson, Prentice-Hall.

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024) Páginas 12-18

ISSN: 2992-7315

Utilización del método numérico de la cuadratura de Carl Friedrich Gauss en conducción de calor.
Use of the numerical method of Carl Friedrich Gauss's quadrature in heat conduction.

Áyax Saúl Martínez Magaña[1], Esiquio Martin Gutiérrez Armenta[2], Marco Antonio Gutiérrez Villegas[3], Israel Isaac Gutiérrez Villegas[4], Javier Norberto Gutiérrez Villegas[5], María de Lourdes Sánchez Guerrero[6], Josué Figueroa González[7], Nicolás Domínguez Vergara[8], Alfonso Jorge Quevedo Martinez[9], Julio Lara García[10], Minerva del Mar Gutiérrez Armenta[11], Juan Manuel Figueroa Flores[12]

[1,2,3,6,7,8]Departamento de Sistemas, UAM-AZC, México, Ciudad de México.
[4,5]División de Ingeniería en Sistemas Computacionales, TESE- TecNM, Estado de México.
[9]Departamento de Administración, UAM-AZC, México, Ciudad de México.
[10]División de Ingeniería Mecánica, Universidad Politécnica de Tecámac (UPT), Estado de México.
[11]Sección de Estudios de Posgrado e Investigación, ESIME-Zacatenco, México D.F., México.
[12]Departamento de Ingeniería y Ciencias Sociales, ESFM-IPN, México, Ciudad de México.
Teléfono 55 1339-1343 Fax (55) 5729-55015 E-mail: esiqv11@hotmail.com

Resumen

En este artículo tiene la finalidad de mencionar la gran aportación del matemático Riemann Georg Friedrich Bernhard, que dio el paso a la integración analítica a lo que hoy se llama cálculo integral, y al teorema fundamental de cálculo, para encontrar la solución de una función de manera analítica pero en general hay una infinidad de las cuales no se puede encontrar esta , se utilizara un método que cualquiera puede utilizar, con estos métodos se realizan aplicaciones en la cual una si se tiene buenos resultados, pero en otra pasa lo contrario los resultados no son buenos, también se calcula la primitiva para compararlos con esta , en este trabajo se utilizara el método numérico de la cuadratura de Carl Friedrich Gauss..

Palabras Clave: Seno, Coseno, tangente, Derivada

Abstract

The purpose of this article is to mention the great contribution of the Riemann mathematician Georg Friedrich Bernhard, who gave way to analytical integration to what is now called integral calculus, and to the fundamental theorem of calculus, to find the solution of a function of analytical way but in general there are an infinity of which this cannot be found, a method that anyone can use will be used, with these methods applications are made in which one does have good results, but in another the opposite happens the results they are not good, the primitive is also calculated to compare them with this one, in this work the numerical method of the quadrature of Carl Friedrich Gauss will be used.

Keywords: Sine, Cosine, Tangent, Derivative

1. Introducción

*Autor para la correspondencia: al2182003847@azc.uam.x
Correo electrónico: al2182003847@azc.uam.x (Áyax-Saúl Martínez-Magaña)

Historial del manuscrito: recibido el 07/09/2023, última versión-revisada recibida el 09/11/2023, aceptado el 30/11/2023, en línea (postprint) desde el 08/04/2024, publicado el 09/04/2024. **DOI:** https://doi.org/10.2992/rict.v2i3.23

La técnica sobre las series de Taylor de Winslow Taylor a principios del siglo XVII, se consideraron estas series infinitas, para el desarrollo del seno, coseno, tangente, logaritmo natural
ln(1+x), que fueron obtenidas por técnicas geométricas por Leibniz, para más detalles consultarse el artículo de Frédéric Barbaresco and Jean-Pierre Gazeau (2019)., en análisis armónico conmutativo lo aborda Rahim Kouki , Barry J. Griffiths, (2020), de utilizando curvas sinusoidales para describir la temperatura, este fue pasando como una idea matemática que dada Cualquier función se podría descomponerse en una combinación lineal de funciones seno. Esta se aplica en resolver ecuaciones diferenciales parciales que se originaron a partir de la física y matemáticas, en espacios Euclideanos, Hanwen Guan, Haoyi Song, Yang and Jiayuan Zhang, (2021), una demostración rigorosa del teorema de sobre superposición no lineal, se construirle a partir del álgebra de Lie en espacios vectoriales. Este es argumento más convincente que apoya el uso de esta definición alternativa de la regla de superposición, se muestra que esta definición permite una generalización inmediata del Teorema de Lie para el caso de sistemas de ecuaciones diferenciales parciales, Josué F. Cariñena, (2006). El método de separación de variables y el principio de superposición para resolver el problema de conducción de calor en estado estable para un rectángulo finito con las siguientes dimensiones $0 \le x \le a$ y $0 \le y \le b$. Las condiciones de frontera conocidas de primer tipo utilizadas por Peter Gustav Lejeune Dirichlet que resolvió el problemas de potencial en una región cerrada con condiciones de frontera, para asegurar la existencia y unicidad de la solución, la ecuación de Laplace cumple con esta condición, a este se le conoce como problema de Peter Gustav Lejeune Dirichlet, al cual se tenía que encontrar una función armónico con la condición de contorno derivada normal $\frac{d\phi}{dn} = g(x).x\grave{o}\Gamma$ frontera donde n es la normal exterior a la superficie de la frontera**.** se llama condición de frontera del segundo tipo, de Carl Gottfried Neumann (1832-1925), Alexander H.-D. Chenga, Daisy T. Cheng. (2005), También en su artículo Pushpendra Singh , Amit Singhal, Binish Fatimah, Anubha Gupta and Shiv Dutt Josh.(2021).realizan una demostración completa cuando se obtiene una única solución y no tiene.
Las condiciones de frontera del Tipo Dirichlet, con temperaturas fijas, o indeterminadas para paredes aisladas (o también llamadas paredes adiabáticas).

En este trabajo se mantienen tres lados del rectángulo (cuadrado) tipo Dirichlet $T\big|_{\Gamma_{1,2,3}}$ temperaturas constantes las cuales son funciones de su posición a lo largo de la cara donde se encuentra.

2. Metodologia a desarrollar

Se hará uso del teorema del principio de superposición para descomponer el problema en cuatro para este caso se usará la figura 1. Donde se muestra este principio para facilitar el desarrollo, con esta se obtendrán cuatro problemas en los cuales se utilizará el método de variables separable para cada uno de estos (método de Fourier). Planteamiento del problema, sea la ecuación de Laplace en dos dimensiones dada por la ecuación 1.

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0$$

$$0 \le x \le a = L \qquad 0 \le y \le b = W \qquad (1)$$

Con las condiciones de frontera.

$$T = f_1(y) \qquad x = 0 \qquad (1.a)$$

$$T = f_2(y) \qquad y = a = L \qquad (1.b)$$

$$T = f_3(x) \qquad y = 0 \qquad (1.c)$$

$$T = f_4(x) \qquad y = b = W \qquad (1.d)$$

Donde Utilizando superposición se separa la ecuación 1, con sus condiciones de frontera (1 a, b, c, d) en cuatro problemas simples, que la solución general es de la forma:

$$T(x,y) = T_1(x,y) + T_2(x,y) + T_3(x,y) + T_4(x,y) \qquad (2)$$

La figura 1 muestra al sistema de la ecuación 1 con sus condiciones de frontera (1 a, b, c, d), en cuatro problemas simples. Estos se resolverán por el método de separación de variables (Fourier).

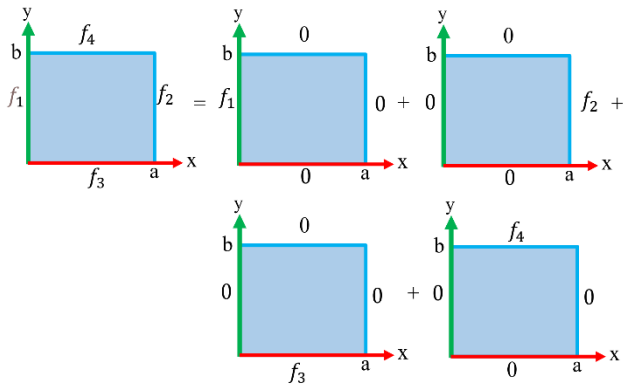

Figura 1. Muestra como un problema de Laplace

En un cuadrado se descompone en cuatro problemas más simples utilizando el principio de superposición basándose en la formulación descrita en Jambrina, L (2023).

Aplicando este principio se expresan en cuatro figuras las cuales tienen su expresión matemática con sus condiciones de frontera. Se resolverán cuatro casos.

Caso 1

Donde la función $T_1(x,y)$ se representa como

$$\frac{\partial^2 T_1}{\partial x^2} + \frac{\partial^2 T_1}{\partial y^2} = 0$$

$$0 \le x \le a = L \qquad 0 \le y \le b = W \qquad (1.1)$$

Utilizando las condiciones de frontera son 1a, 1b, 1c y 1d. Las ecuaciones diferenciales de segundo orden asociadas al sistema.

$$T_1 = f_1(y) \qquad y = 0 \qquad (1.1\ a)$$

$$T_1 = 0 \qquad x = a = L \qquad (1.1\text{ b})$$
$$T_1 = 0 \qquad y = 0 \qquad (1.1\text{ c})$$
$$T_1 = 0 \qquad y = b = W \qquad (1.1\text{ d})$$

Caso 2

Donde la función $T_2(x,y)$ se representa como

$$\frac{\partial^2 T_2}{\partial x^2} + \frac{\partial^2 T_2}{\partial y^2} = 0$$
$$0 \le x \le a = L \qquad 0 \le y \le b = W \qquad (2.1)$$

Utilizando las condiciones de frontera son 1a, 1b, 1c y 1d. Las ecuaciones diferenciales de segundo orden asociadas al sistema.

$$T_2 = 0 \qquad x = a = L \qquad x = 0 \qquad (2.1\text{ a})$$

$$T = f_2(y) = \left.\frac{\partial \boldsymbol{T}_2}{\partial \boldsymbol{y}}\right|_{\Gamma_2} = (\boldsymbol{T_s} - \boldsymbol{T}_\infty) \qquad x = a = L = w \qquad (2.1\text{ b})$$

Donde $\boldsymbol{T_s}$ la temperatura de la superficie del cuero ya no tiene influencia sobre el material**.** $\boldsymbol{T}_\infty$ = es la temperatura del medio ambiente.

$$T_2 = 0 \qquad y = 0 \qquad (2.1.\text{c})$$
$$T_2 = 0 \qquad y = b = W \qquad (2.1.\text{d})$$

Caso 3.

Para $T_3(x,y)$ como:

$$\frac{\partial^2 T_3}{\partial x^2} + \frac{\partial^2 T_3}{\partial y^2} = 0$$
$$0 \le x \le a = L \qquad 0 \le y \le b = W \qquad (3.1)$$

Las Condiciones de frontera:

$$T_3 = 0 \qquad x = 0 \qquad (3.1.\text{a})$$
$$T_3 = 0 \qquad x = a = L \qquad (3.1.\text{b})$$

$$T_3 = f_3(x) \qquad y = 0 \qquad (3.1.\text{c})$$
$$T_3 = 0 \qquad y = b = W \qquad (3.1.\text{d})$$

Caso 4

Resolviendo el caso 4.

$$\frac{\partial^2 T_4}{\partial x^2} + \frac{\partial^2 T_4}{\partial y^2} = 0$$
$$0 \le x \le a = L \qquad 0 \le y \le b = W \qquad (4.1)$$

Las Condiciones de frontera:

$$T_4 = 0 \qquad x = 0 \qquad (4.1.\text{a})$$
$$T_4 = 0 \qquad x = a = L \qquad (4.1.\text{b})$$
$$T_4 = 0 \qquad y = 0 \qquad (4.1.\text{c})$$
$$T_4 = f_4(x) \qquad y = b = W \qquad (4.1.\text{d})$$

Resolviendo los cuarto casos para estos se propone una solución de la forma:

$$T_4(x,y) = X(x)Y(y) \qquad (4.\text{a})$$

Obteniendo las segundas derivadas parciales de la ecuación 4.a con respecto a x, y sustituyendo en la ecuación 4 se obtiene:

$$-\frac{1}{X}\frac{d^2X(x)}{dx^2} = \frac{1}{Y}\frac{d^2Y(y)}{dy^2} = \lambda^2 \quad , \quad \lambda^2 > 0 \qquad (a5)$$

La cuales solo tiene solución para $\lambda^2 > 0$ de la ecuación a5 se obtienen dos ecuaciones diferenciales ordinarias de segundo orden:

$$\frac{d^2X(x)}{dx^2} + \lambda^2 X = 0 \qquad (a6)$$
$$\frac{d^2Y(y)}{dy^2} - \lambda^2 Y = 0 \qquad (a7)$$

Donde sus soluciones dadas por:

$$X = c_1\cos(\lambda x) + c_2 sen(\lambda x) \qquad (a8)$$
$$Y = c_3 e^{-\lambda y} + c_4 e^{\lambda y} \qquad (9)$$

Sustituyendo las ecuaciones (a8- a9) en la ecuación (4a):

$$T_4(x,y) = (c_1\cos(\lambda x + c_3 \text{sen}(\lambda x)\left[c_3 e^{\lambda y} + c_4 e^{-\lambda y}\right] \qquad (a10)$$

Aplicando la condición de frontera en la ecuación 4a:

$$T_4(0,y) = [c_1(1) + c_2(0)]\left[c_3 e^{-\lambda y} + c_4 e^{\lambda y}\right] = 0 \qquad (a11)$$

se tiene:

$$c_1 = 0 \qquad (a12)$$
$$T_4(x,y) = c_2 sen(\lambda x)\left[c_3 e^{-\lambda y} + c_4 e^{\lambda y}\right] \qquad (a13)$$
$$(a14)$$

de la ecuación (4c):

$$T_4(x,0) = c_2 sen(\lambda x)\left[c_3 e^{-\lambda(0)} + c_4 e^{\lambda(0)}\right] \qquad (a15)$$

$$T_4(x,0) = c_2 sen(\lambda x)[c_3 + c_4] = 0 \qquad (a16)$$
$$c_3 + c_4 = 0 \qquad (a17)$$

Se obtiene la ecuación (a18):

$$c_3 = -c_4 \qquad (a18)$$

utilizando la condición de frontera a la ecuación (4b)

$$T_4(L,y) = c_2 c_4 sen(\lambda L)\left[e^{\lambda y} - e^{-\lambda y}\right] = 0 \qquad (a19)$$

Todas las constantes deben ser diferente de cero. debido a que la función exponencial es diferente de cero en todo $\mathbb{R}$, así que se tiene la única posibilidad es que:

$$sen(\lambda L) = 0 \qquad (a20)$$

Esto sólo puede tenerse si de la ecuación (a20) el argumento sea cero esto debe cumplir (a21):

$$\lambda = \frac{n\pi}{L}, \quad \text{para} n = 1,2,\ldots \qquad (a21)$$

Por el principio de superposición para las constantes C_n , así utilizando la identidad $2senh(\lambda x) = e^{\lambda x} - e^{-x\lambda}$ se tiene:

$$T_4(x,y) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi y}{L}\right) \qquad (a22)$$

Utilizando la condición de la ecuación (4d):

$$T_4(x,W) = f_4(x) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi W}{L}\right) \qquad (a23)$$

Usando el principio de ortogonalidad de la serie de Taylor:

$$c_n sen h\left(\frac{n\pi W}{L}\right)=\frac{2}{L}\int_0^L f_4(x)sen\left(\frac{n\pi x}{L}\right)dx \tag{a24}$$

Despejando a C_n

$$c_n=\frac{2}{L\,sen h\left(\frac{n\pi W}{L}\right)}\int_0^L f_4(x)sen\left(\frac{n\pi\, x}{L}\right)dx \tag{a25}$$

Sustituyendo en la ecuación (a25-a22)

$$T_4(x,y)=\frac{2}{W}\sum_{n=1}^{\infty}\frac{senh\left(\frac{n\pi y}{L}\right)}{senh\left(\frac{n\pi L}{W}\right)}\left[\int_0^L f_4(x')sen\left(\frac{n\pi\, x'}{L}\right)dx'\right] \tag{a26}$$

La ecuación (a26) es la solución de la ecuación 4 sujeta a las condiciones de frontera dadas por las ecuaciones (4a -4d).

SOLUCIÓN PARA EL CASO 1.

$$\frac{\partial^2 T_1}{\partial x^2}+\frac{\partial^2 T_1}{\partial y^2}=0$$

$$0\le x\le a=L \qquad 0\le y\le b=W \qquad 1$$

Condiciones de frontera para el caso 1

$$T_1=f_1(y) \qquad x=0 \tag{1a}$$
$$T_1=0 \qquad x=a=L \tag{1b}$$
$$T_1=0 \qquad y=0 \tag{1c}$$
$$T_1=0 \qquad y=b=W \tag{1d}$$

Se vuelve a proponer una solución de la forma

$$T_1(x,y)=X(x)Y(y) \tag{a5}$$

Derivando parcialmente a la ecuación (b5) con respecto a x, y y sustituyendo en la ecuación (1) se obtiene dos ecuaciones diferenciales ordinarias de segundo orden que son

$$\frac{1}{X}\frac{d^2X(x)}{dx^2}=-\frac{1}{Y}\frac{d^2Y(y)}{dy^2}=\lambda^2 \quad , \quad \lambda^2>0 \tag{a6}$$

de la ecuación (a6) se obtienen dos ecuaciones diferenciales de segundo orden

$$\frac{d^2Y(y)}{dy^2}+\lambda^2 Y=0 \tag{a7}$$

$$\frac{d^2X(x)}{dx^2}-\lambda^2 X=0 \tag{a8}$$

Donde las soluciones para ésta están dadas por

$$Y(y)=c_1\cos(\lambda y)+c_2 sen(\lambda y) \tag{a9}$$

$$X(x)=c_3e^{-\lambda x}+c_4e^{\lambda x} \tag{a10}$$

Sustituyendo las ecuaciones (a9) y (b10) en la ecuación (a5)

$$T_1(x,y)=\left[c_1\cos(\lambda y)+c_2 sen(\lambda y)\right]\left[c_3e^{-\lambda x}+c_4e^{\lambda x}\right] \tag{a11}$$

Aplicando la condición de frontera (3b) .

$$T_1(x,0)=\left[c_1\cos(\lambda(0))+c_2 sen(\lambda(0))\right]\left[c_3e^{-\lambda x}+c_4e^{\lambda x}\right]=0 \tag{a12}$$

$$T_1(x,0)=\left[c_1\cos(0)+c_2 sen(0)\right]\left[c_3e^{-\lambda x}+c_4e^{\lambda x}\right]=0 \tag{a13}$$

$$T_1(x,0)=\left[c_1(1)+c_2(0)\right]\left[c_3e^{-\lambda x}+c_4e^{\lambda x}\right]=0 \tag{a14}$$

De donde

$$c_1=0 \tag{a15}$$

$$T_1(x,y)=c_2 sen(\lambda y)\left[c_3e^{-\lambda x}+c_4e^{\lambda x}\right] \tag{a16}$$

Aplicando la condición de frontera dada por la ecuación (1b)

$$T_1(L,y)=c_2 sen(\lambda y)\left[c_3e^{-\lambda L}+c_4e^{\lambda L}\right]=0 \tag{a17}$$

Para este caso se tiene la siguiente suposición $e^{\lambda L}\to senh(\lambda L)$ para $x>>0$ (grandes), $e^{-\lambda L}\to senh(\lambda L)$ para x pequeños, entonces la ecuación (a17) se escribe de la siguiente manera $senh(\lambda L)\left[c_3+c_4\right]=0$. Como $senh(\lambda L)\neq 0$ así $c_3+c_4=0$ de esta $c_3=-c_4$ Sustituida en la ecuación (a14)

$$T_1(x,y)=c_2c_4\ sen(\lambda y)senh(\lambda x) \tag{a18}$$

Aplicando la condición de frontera dada por la ecuación (1 d)

$$T_1(x,W)=c_3c_2 sen(\lambda W)senh(\lambda x)=0 \tag{a18}$$

Donde las constantes deben ser diferente de cero así que la única posibilidad es que

$$sen(\lambda W)=0 \tag{a19}$$

Esto sólo puede tenerse si

$$\lambda=\frac{n\pi}{W}, \quad \text{para } n=1,2,... \tag{a20}$$

$$T_1(x,y)=\sum_{n=1}^{\infty}c_n sen\left(\frac{n\pi y}{W}\right)senh\left(\frac{n\pi x}{W}\right) \tag{a21}$$

Utilizando la condición (1a)

$$T_1(0,y)=f_1(y)=\sum_{n=1}^{\infty}c_n sen\left(\frac{n\pi y}{L}\right)senh\left(\frac{n\pi(0)}{W}\right)$$

$$T_1(0,y)=f_1(y)=\sum_{n=1}^{\infty}c_n sen\left(\frac{n\pi y}{L}\right) \tag{a22}$$

$$T_1(0,y)=f_1(y)=\sum_{n=1}^{\infty}c_n sen\left(\frac{n\pi y}{L}\right)$$

Usando la ortogonalidad

$$c_n=\frac{2}{W}\int_0^W f_1(y)sen\left(\frac{n\pi\, y}{W}\right)dy \tag{a23}$$

Sustituyendo la ecuación (b23) en la ecuación (b21)

$$T_1(x,y)=\frac{2}{W}\sum_{n=1}^{\infty}senh\left(\frac{n\pi x}{W}\right)sen\left(\frac{n\pi y}{W}\right)\int_0^w f_1(y')sen\left(\frac{n\pi\, y'}{W}\right)dy' \tag{a24}$$

La ecuación (b24) es la solución de la ecuación (1) sujeta a las condiciones de frontera dadas por las ecuaciones (2a, 2b, 2c, 2d)

SOLUCIÓN PARA EL CASO 2

$$\frac{\partial^2 T_2}{\partial x^2}+\frac{\partial^2 T_2}{\partial y^2}=0$$

$$0\le x\le a=L \qquad 0\le y\le b=W \tag{2}$$

Condiciones de frontera:

$$T_2=0 \qquad x=0 \tag{2a}$$

$$T_2=f_2(y)==\left.\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{y}}\right|_{\Gamma_2}=(\boldsymbol{T_s}-\boldsymbol{T_\infty}) \qquad x=a=L \tag{2b}$$

$$T_2=0 \qquad y=0 \tag{2 c}$$
$$T_2=0 \qquad y=b=W \tag{2 d}$$

Realizando un proceso análogo al anterior

$$T_2(x,y) = X(x)\square Y(y) \tag{b5}$$

Calculando las segundas derivadas parciales de la ecuación b5 con respecto a x, y sustituyendo en la ecuación 2 se obtiene:

$$\frac{1}{X(x)}\frac{d^2X(x)}{dx^2} = \frac{1}{Y(y)}\frac{d^2Y(y)}{dy^2} = \lambda^2, \lambda > 0 \tag{b6}$$

De la ecuación (b6) se obtienen dos ecuaciones diferenciales ordinarias de segundo orden.

$$\frac{d^2Y(y)}{dy^2} + \lambda^2 Y(y) = 0 \tag{b7}$$

$$\frac{d^2X(x)}{dy^2} - \lambda^2 X(x) = 0 \tag{b8}$$

Donde las soluciones para ésta están dadas por:

$$Y(y) = c_1\cos(\lambda y) + c_2\text{sen}(\lambda y) \tag{b9}$$

$$X(x) = c_3 e^{\lambda x} + c_3 e^{-\lambda x} \tag{b10}$$

Sustituyendo las ecuaciones (b9-b10) en la ecuación (b5)

$$T_2(x,y) = \left[c_1\cos(\lambda y) + c_2\text{sen}(\lambda y)\right]\left(c_3 e^{\lambda x} + c_4 e^{-\lambda x}\right) \tag{b11}$$

Derivando $\frac{\partial T_2}{\partial y} = -\lambda\cos(\lambda y) + c_2\lambda\cos(\lambda y)$ aplicando la condición de frontera de la ecuación (2b) para esto se tiene que derivar parcialmente con respecto a y la ecuación b11:

$$\frac{\partial T_2(x,y)}{\partial y} = \left[-\lambda c_1\text{seno}(\lambda y) + \lambda c_2\cos(\lambda y)\right] = (\boldsymbol{T_s} - \boldsymbol{T_\infty}) \tag{12b}$$

Utilizando la condición de frontera ecuación (2b) la siguiente teoría se utiliza para $N \in \mathbb{N}, N > 1$ se pueden tomar las siguientes sumatorias para las series seno y coseno. $\sum_{i=1}^{N}\cos(\frac{2k\pi}{\boldsymbol{N}}) = 0$, $\sum_{i=1}^{N}\text{sin}(\frac{2k\pi}{\boldsymbol{N}})$ para una ampliación sobre el tema ver Michael P. Knapp Loyola, S. Greitzer, la demostración del teorema la realiza

Many cheerful Facts Arbelos 4 (1986), no. 5, 14–17. 2. J. Holdener, Math bite: Sums of Sines and Cosines, this Magazine, 82 (2009), 126 utilizando números complejos.

Las pruebas de estas dos identidades son un buen ejercicio de álgebra compleja: usando el teorema de DeMoivre y la fórmula de Euler, $e^{i(2\pi)} = \left(e^{i\left(\frac{2\pi}{N}\right)}\right)^N = 1$ se derivan las identidades de la parte real e imaginaria se obtiene la siguiente ecuación

$$\sum_{i=1}^{N}(\cos(\frac{2\pi k}{n})) + i\,sin\left(\frac{2\pi k}{n}\right)) = \sum_{K0=}^{N-1}\frac{\left(e^{\frac{2\pi i}{N}}\right)^N - 1}{\left(e^{\frac{2\pi i}{N}}\right)^N - 1} = 0\,.$$

$$\frac{\partial T_2(0,w)}{\partial y} = \left[-\lambda c_1\text{seno}(\lambda w) + \lambda c_2\cos(\lambda w)\right] = (\boldsymbol{T_s} - \boldsymbol{T_\infty}) \tag{b13}$$

Como el segundo término de la ecuación (b13) no puede ser cero entonces se debe de tener:

$$\text{seno}(\lambda w) = \cos(\lambda w) \tag{b14}$$

$\cos(\theta) = \cos\left(\frac{\pi}{2} + \theta\right)$; $\sin(\theta) = \sin\left(\theta - \frac{\pi}{2}\right)$; en estos casos al evaluar es cero.

También ser puede ser utilizado en este caso.

$$T_2(x,y) = c_4\left[c_1\cos(\lambda y) + c_2 sen(\lambda y)\right] \tag{b15}$$

Con la condición de la ecuación (2c):

$$T_2(x,0) = \left[c_1\cos(\lambda(0)) + c_2 sen(\lambda(0))\right] = 0 \tag{b16}$$

$$T_2(x,0) = \left[c_1(1) + c_2(0)\right] = 0 \tag{b17}$$

De aquí $c_1 = 0$:

$$T_2(0,y) = c_2 sen(\lambda y) \tag{b17.1}$$

Aplicando las condiciones de frontera (2d)

$$T_4(x,W) = c_2 sen(\lambda W) = 0 \tag{b18}$$

Donde las constantes deben ser diferente de cero así que la única posibilidad es que

$$sen(\lambda W) = 0 \tag{b19}$$

Esto sólo puede tenerse si

$$\lambda = \frac{n\pi}{W}, \quad \text{para } n = 1, 2, \ldots \tag{b20}$$

Utilizando el principio de superposición para las soluciones y sustituyendo la identidad $2senh(\lambda y) = e^{\lambda x} - e^{-\lambda x}$,las constantes c_n.

$$T_2(x,y) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi y}{W}\right) senh\left(\frac{n\pi x}{W}\right)(\boldsymbol{T_s} - \boldsymbol{T_\infty}) \tag{b21}$$

De la condición tomando la ecuación (2b)

$$T_2(L,y) = (\boldsymbol{T_s} - \boldsymbol{T_\infty})\sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi y}{L}\right) senh\left(\frac{n\pi L}{W}\right) \tag{b22}$$

Usando la ortogonalidad

$$c_n senh\left(\frac{n\pi L}{W}\right) = \frac{2(\boldsymbol{T_s} - \boldsymbol{T_\infty})}{W}\int_0^W sen\left(\frac{n\pi\, y}{W}\right)dy \tag{b23}$$

Despejando c_n de la ecuación (b23)

$$c_n = \frac{2(\boldsymbol{T_s} - \boldsymbol{T_\infty})}{W senh\left(\frac{n\pi L}{W}\right)}\int_0^W sen\left(\frac{n\pi\, y}{W}\right)dy \tag{b24}$$

Sustituyendo la ecuación (b24) en la ecuación (b21)

$$T_2(x,y) = \frac{2}{W}\sum_{n=1}^{\infty}\frac{senh\left(\frac{n\pi x}{W}\right)}{senh\left(\frac{n\pi L}{W}\right)}\int_0^w sen\left(\frac{n\pi\, y'}{W}\right)dy' \tag{b25}$$

La ecuación (b25) es la solución de (2) sujeta a las ecuaciones (2a, 2b, 2c, 2d)

SOLUCIÓN DEL CASO 3

$$\frac{\partial^2 T_3}{\partial x^2} + \frac{\partial^2 T_3}{\partial y^2} = 0$$

$$0 \le x \le a = L \qquad 0 \le y \le b = W$$

$$T_3 = 0 \qquad x = 0 \tag{3a}$$
$$T_3 = 0 \qquad x = a = L \tag{3b}$$
$$T_3 = f_3(x) \qquad y = 0 \tag{3c}$$
$$T_3 = 0 \qquad y = b = W \tag{3d}$$

Se propone una solución de la forma

$$T_3(x,y) = X(x)Y(y) \tag{a5}$$

Calculando las segundas derivadas parciales de la ecuación (a5) con respecto a x, y y sustituyendo en la ecuación (4) se obtiene

$$-\frac{1}{X}\frac{d^2X(x)}{dx^2} = \frac{1}{Y}\frac{d^2Y(y)}{dy^2} = \lambda^2 \quad , \quad \lambda^2 > 0 \tag{a6}$$

De la ecuación (a6) se obtienen dos ecuaciones diferenciales de segundo orden

$$\frac{d^2X(x)}{dx^2} + \lambda^2 X = 0 \tag{a7}$$

$$\frac{d^2Y(y)}{dy^2} - \lambda^2 Y = 0 \tag{a8}$$

Donde las soluciones están dadas por

$$X = c_1\cos(\lambda x) + c_2 sen(\lambda x) \tag{a9}$$

$$Y = c_3 e^{-\lambda y} + c_4 e^{\lambda y} \tag{a10}$$

Sustituyendo la ecuación (a9) y (a10) en la ecuación (a5)

$$T_3(x,y) = \left[c_1\cos(\lambda y) + c_2 sen(\lambda y)\right]\left[c_3 e^{-\lambda x} + c_4 e^{\lambda x}\right] \tag{a11}$$

Aplicando la condición de frontera dada por la ecuación (3a)

$$T_3(0,y) = \left[c_1\cos(\lambda 0) + c_2 sen(\lambda 0)\right]\left[c_3 e^{-\lambda x} + c_4 e^{\lambda y}\right] = 0 \tag{a12}$$

$$T_3(0,y) = \left[c_1(1) + c_2(0)\right]\left[c_3 e^{-\lambda x} + c_4 e^{\lambda}\right] = 0 \tag{a14}$$

De donde se tiene que

$$c_2 = 0 \tag{a15}$$

$$T_3(x,y) = c_1 cos(\lambda y)\left[c_3 e^{-\lambda x} + c_4 e^{\lambda x}\right] \tag{a15') (a16}$$

Utilizando la condición (3d)

$$T_3(x,W) = c_2 sen(\lambda x)\left[c_3 e^{-\lambda W} + c_4 e^{\lambda W}\right] = 0 \tag{a17}$$

Para este caso se tiene la siguiente suposición $e^{\lambda W} \to senh(\lambda W)$ para $x >> 0$ (grandes), $e^{-\lambda W} \to senh(\lambda W)$ para x pequeños, entonces por la ecuación (a17) se escribe de la siguiente manera

$$T_3(x,W) = c_2 sen(\lambda y) senh(\lambda W)\left[c_3 + c_4\right] = 0 \tag{a17) (a18}$$

Como $senh(\lambda W), sen(\lambda x) \neq 0$ entonces:

$$c_3 + c_4 = 0 \tag{a19}$$

$$c_3 = -c_4 \tag{a20}$$

Sustituyendo en las ecuaciones (a20) y en (a16)

$$T_3(x,y) = c_2 c_4 sen(\lambda y) senh(\lambda y)$$

Aplicando la condición de frontera dada por la ecuación (3b)

$$T_3(L,y) = c_2 c_4 sen(\lambda y) senh(\lambda x) = 0 \tag{a21}$$

Donde las constantes deben ser diferente de cero y las funciones exponenciales también son diferentes de cero, así que la única posibilidad es que

$$sen(\lambda L) = 0 \tag{a22}$$

Esto sólo puede tenerse si

$$\lambda = \frac{n\pi}{L}, \quad \text{para } n = 1,2,... \tag{a23}$$

Utilizando el principio de superposición para las soluciones y sustituyendo c_1, c_2 por c_n.

$$T_3(x,y) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi y}{L}\right) \tag{a24}$$

Utilizando la condición (3c)

$$T_3(x,0) = f_3(x) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right)$$
$$T_3(x,0) = f_3(x) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi(0)}{L}\right) \tag{a25}$$
$$T_3(x,0) = f_3(x) = \sum_{n=1}^{\infty} c_n sen\left(\frac{n\pi x}{L}\right)$$

Usando la ortogonalidad

$$c_n = \frac{2}{L}\int_0^L f_4(x) sen\left(\frac{n\pi x}{L}\right) dx \tag{a26}$$

Sustituyendo la ecuación (a26) en la ecuación (a24)

$$T_3(x,y) = \frac{2}{L}\sum_{n=1}^{\infty} sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi y}{L}\right)\left[\int_0^L f_4(x') sen\left(\frac{n\pi x'}{L}\right) dx'\right] \tag{a27}$$

La ecuación (a27) es la solución de (3) sujeta a las condiciones de frontera dadas por las ecuaciones (3a, 3b, 3c, 3d).
La solución de la ecuación (5) bajo las condiciones de frontera (5a, 5b, 5c, 5d) está dada por la ecuación(a28):

$$T(x,y) = \frac{2}{W}\sum_{n=1}^{\infty} senh\left(\frac{n\pi x}{W}\right) sen\left(\frac{n\pi y}{W}\right)\int_0^w f_1(y') sen\left(\frac{n\pi y'}{W}\right) dy' +$$
$$\frac{2}{W}\sum_{n=1}^{\infty}\frac{senh\left(\frac{n\pi x}{W}\right)}{senh\left(\frac{n\pi L}{W}\right)}\int_0^w sen\left(\frac{n\pi y'}{W}\right) dy' + \tag{a28}$$

$$\frac{2}{L}\sum_{n=1}^{\infty} sen\left(\frac{n\pi x}{L}\right) senh\left(\frac{n\pi y}{L}\right)\left[\int_0^L f_4(x') sen\left(\frac{n\pi x'}{L}\right) dx'\right] +$$
$$\frac{2}{W}\sum_{n=1}^{\infty}\frac{senh\left(\frac{n\pi y}{L}\right)}{senh\left(\frac{n\pi L}{W}\right)}\left[\int_0^L f_4(x') sen\left(\frac{n\pi x'}{L}\right) dx'\right]$$

La ecuación 28 es la solución del problema utilizando el principio de super posición que viene dada en la ecuación 2, el procedimiento puede ser utilizado para resolver con otro tipo de condiciones de frontera.

3. Conclusiones

Una de las formas sencilla pero procedimiento laboriosa de cálculo para este tipo de problemas, es el método separación de variables de Fourier, se puede utilizar la transformada de Laplace, pero este tiene un problema que cuando ya se tiene despejada la función transformada, se debe encontrar la función inversa de esta, por lo general hay dificultades ya que se requiere una buena habilidad en el cálculo transformada inversa, esta puede ser complicada para encontrarla, otra

alternativa seria las técnicas numéricas por ejempló diferencias finita, elemento finito o elemento frontera. Pero estas soluciones obtenidas por estos métodos solo dan solución en unos puntos llamados nodos, y los anteriores son soluciones analíticas que sirven para obtener la solución en cualquier parte del modelo.

El resultado de este procedimiento se utilizará para realizar la validación de los métodos numéricos mencionados. Otro problema a resolver es cuando una de las caras de la frontera tiene condiciones de Neumann, esta condición es la derivada normal que no es otra cosa que la derivada direccional en la dirección normal a la superficie de la frontera correspondiente.

4. Agradecimientos

Los autores agradecen a la escuela Universidad Nacional Autónoma de México, al Tecnológico Nacional de México, al Tecnológico de Estudios Superiores de Ecatepec, y a la revista RICT Revista de Investigación Científica, Tecnológica e Innovación por su apoyo de la publicación del manuscrito.

5. Referencias

[1]. Arora, G., Joshi, V., & Garki, I. S. (2020). Developments in Runge–Kutta method to solve ordinary differential equations. En *Recent Advances in Mathematics for Engineering* (pp. 193–202). CRC Press.

[2]. *Grafiati. (2021) Journal articles on the topic '120219 - Ecuaciones diferenciales ordinarias*. (s/f).

[3]. Griffiths, B. J., & Kouki, R. (2019). Introducing Taylor series and local approximations using a historical and semiotic approach. *International electronic journal of mathematics education*, *15*(2). https://doi.org/10.29333/iejme/6293

[4]. Hubbard, J. H., Habre, S. S., & West, B. H. (2001). The convergence of an Euler approximation of an initial value problem is not always obvious. *The American mathematical monthly: the official journal of the Mathematical Association of America*, *108*(4), 326. https://doi.org/10.2307/2695239.

[5]. Jambrina, L. F., (2023), Capítulo 7 Ecuación de Laplace Departamento de Matemática e Informáticas aplicadas a las ingenierías civil y naval. Pp. 166-167.

[6]. Kamruzzaman, M. C. (2018). *A Comparative Study on Numerical Solution of Initial Value problem by Using Euler´s Method and Ruge-Kutta*.

[7]. Nurujjaman, M. (2020). *Enhanced Euler's Method to Solve First Order Ordinary Differential Equations with Better Accuracy*.

[8]. Youssef, I. K., & El-Arabawy, H. A. (2007). Picard iteration algorithm combined with Gauss–Seidel technique for initial value problems. *Applied Mathematics and Computation*, *190*(1), 345–355. https://doi.org/10.1016/j.amc.2007.01.058

[9]. Zhang Lijuan, G. (2018). comparison of several Numerical Algorithms for Solvin ordinary Differential Equation initial Value problem. *Advances in Computer Science Research*, *78*.

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 2 No. 3 (2024) Páginas 19-25

ISSN: 2992-7315

Comparativa de diferentes técnicas de Minería de Datos para la predicción del uso de la bicicleta de acuerdo con las condiciones climáticas y estacionales en Washington
Comparison of different Data Mining techniques for predicting bicycle use according to climatic and seasonal conditions in Washington

Ana Rosa Velázquez Cordero[a,*], Francisco-Jacob Ávila-Camacho[b]

[a] *Maestría en Ingeniería en Sistemas Computacionales, Tecnológico de Estudios Superiores de Ecatepec, 55210, Ecatepec, Estado de México, México.*
[b] *División de Ingeniería en Sistemas Computacionales, Tecnológico Nacional de México / TES Ecatepec, 55210, Ecatepec, Estado de México, México*

Resumen

En este artículo se aborda la importancia de los sistemas de bicicletas compartidas como una evolución significativa en el paradigma del alquiler de bicicletas, destacando su impacto positivo en el tráfico, el medio ambiente y la salud, así como la extensa red de programas de bicicletas compartidas a nivel mundial junto con la valiosa cantidad de datos generados por estos sistemas, en especial en las estaciones de Washington. Se realizo una análisis de datos histórico del sistema Capital Bikeshare en Washington D.C. durante los años 2011 y 2012, lo que proporciona una base sólida para capturar variaciones estacionales y tendencias en los patrones de alquiler de bicicletas y con ello realizar la formulación de modelos predictivos precisos. Finalmente se destaca la importancia de los sistemas de bicicletas compartidas, la disponibilidad de datos y la relación entre el proceso de alquiler de bicicletas, así como diversos factores ambientales para comprender y predecir la demanda de bicicletas en entornos urbanos, para el caso de la predicción se compararon 2 técnicas de minería de datos las cuales son "Árboles de regresión y Regresión Simple" con el histórico y elegir la que presento la menor desviación.

Palabras Clave: Predicción, Árboles de Regresión, redes neuronales, regresión, error cuadrático medio(MSE) y machine learning.

Abstract

This article addresses the importance of bike-sharing systems as a significant evolution in the bike rental paradigm, highlighting their positive impact on traffic, the environment, and health, as well as the extensive network of bike-sharing programs. globally along with the valuable amount of data generated by these systems, especially at the Washington stations. A historical data analysis of the Capital Bikeshare system in Washington D.C. was performed. during the years 2011 and 2012, which provides a solid basis for capturing seasonal variations and trends in bicycle rental patterns and thereby formulating accurate predictive models. Finally, the importance of bicycle sharing systems, the availability of data and the relationship between the bicycle rental process, as well as various environmental factors to understand and predict the demand for bicycles in urban environments, in the case of prediction, are highlighted. Two data mining techniques were compared, which are "regression trees and Simple Regression" with the historical data and the one that presented the least deviation was chosen.

Keywords: Prediction, Regression Trees, Neural networks, regression, mean square error (MSE) and machine learning.

1. Introducción

*Autor para la correspondencia: 202222@tese.edu.mx
Correo electrónico: 202222@tese.edu.mx (Ana Rosa Velázquez Cordero), fjacobavila@tese.edu.mx (Francisco Jacob Ávila Camacho).

Historial del manuscrito: recibido el 26/01/2024, última versión-revisada recibida el 02/04/2024, aceptado el 03/04/2024, en línea (postprint) desde el 08/04/2024, publicado el 09/04/2024. **DOI:** https://doi.org/10.2992/rict.v2i3.44

Los sistemas de bicicletas compartidas representan una evolución significativa en el paradigma de alquiler de bicicletas tradicional. Estos sistemas automatizan todo el proceso, desde la membresía hasta el alquiler y la devolución, facilitando a los usuarios la posibilidad de alquilar una bicicleta desde una ubicación específica y devolverla en otra.(Shaheen et al., 2013). La singularidad de estos sistemas radica en que se registra la duración del viaje y las posiciones de salida y llegada(Rosales-Asensio et al., 2019). A nivel mundial, existen alrededor de 500 programas de bicicletas compartidas, con más de medio millón de bicicletas en circulación (Eren & Uz, 2020). Esta extensa red de alquiler de bicicletas no solo proporciona una alternativa de transporte eficiente, sino que también genera grandes cantidades de datos que resultan valiosos para este artículo.

La importancia de los sistemas de bicicletas compartidas se destaca en su impacto positivo en cuestiones cruciales como el tráfico, el medio ambiente y la salud pública. Al fomentar el uso de la bicicleta como medio de transporte, estos sistemas contribuyen a la reducción de la congestión vehicular en las ciudades, disminuyendo las emisiones de gases contaminantes (Gong et al., 2024). Además, promueven un estilo de vida activo y saludable al integrar el ejercicio físico en las rutinas diarias de los usuarios (Zhou et al., 2022). Desde una perspectiva medioambiental, la reducción de la dependencia de los vehículos motorizados también tiene implicaciones positivas para la calidad del aire y la sostenibilidad urbana (Zhou et al., 2022).

A diferencia de otros servicios de transporte, la información detallada registrada en cada viaje, como la duración, la posición de salida y llegada, se convierte en un conjunto de datos robusto que puede considerarse una red de sensores en sí misma. Esta red virtual ofrece una oportunidad única para monitorear la movilidad en la ciudad en tiempo real (Zheng & Li, 2020). El análisis de estos datos puede revelar patrones de comportamiento, preferencias de rutas y cambios en la demanda de bicicletas, lo que tiene implicaciones importantes para la planificación urbana y la optimización de la infraestructura de transporte. En este sentido, la convergencia de la movilidad urbana y la generación de datos en tiempo real abre nuevas posibilidades para abordar desafíos contemporáneos en el diseño de ciudades inteligentes y sostenibles (Guo et al., 2022),.(Ricci, 2015).

1.1. Características de los Datos Generados por los Sistemas

Los sistemas de bicicletas compartidas se distinguen notablemente de otros servicios de transporte en la cantidad y tipo de datos recopilados (Garcia-Gutierrez et al., 2014). A diferencia de servicios como autobuses o metro, donde los datos suelen limitarse a la cantidad de pasajeros o las horas de operación, los sistemas de bicicletas compartidas registran explícitamente detalles específicos de cada viaje. Esta característica única proporciona un conjunto de datos detallado y completo que incluye la duración del viaje, así como la posición exacta de salida y llegada de cada bicicleta alquilada. Mientras que otros servicios de transporte suelen generar datos agregados, los sistemas de bicicletas compartidas ofrecen una granularidad sin precedentes en la información, permitiendo un análisis más preciso de los patrones de movilidad urbana (Eren & Uz, 2020).

El registro explícito de la duración del viaje, la posición de salida y llegada en los sistemas de bicicletas compartidas no solo representa una ventaja en términos de cantidad de datos, sino que también ofrece una ventana directa a la dinámica de la movilidad urbana. La duración del viaje permite no solo evaluar la eficiencia del sistema en términos de tiempo de desplazamiento, sino también comprender patrones de uso y preferencias de los usuarios. La posición de salida y llegada de cada bicicleta brinda información detallada sobre las rutas más frecuentes, puntos de interés y la distribución geográfica de la demanda (Gámez-Pérez et al., 2020). Estos datos específicos son esenciales para diseñar estrategias efectivas de ubicación de estaciones, mejorar la infraestructura de ciclovías y responder de manera ágil a las fluctuaciones en la demanda de bicicletas en diferentes áreas de la ciudad.

La diferencia en la cantidad y tipo de datos recopilados por los sistemas de bicicletas compartidas resalta su potencial como una valiosa red de sensores virtuales. Esta riqueza de información no solo beneficia a los usuarios en términos de comodidad y accesibilidad, sino que también abre nuevas oportunidades para investigadores y planificadores urbanos (Fin De Máster et al., 2019.).

La capacidad de monitorear y analizar detalladamente la movilidad en tiempo real a través de estos datos posiciona a los sistemas de bicicletas compartidas como herramientas fundamentales para comprender la dinámica urbana y mejorar la eficiencia de los servicios de transporte sostenible en nuestras ciudades (Shaheen et al., 2013).

2. Conjunto de Datos del Sistema Capital Bikeshare

El conjunto de datos histórico abarca un periodo significativo de dos años, comprendidos entre 2011 y 2012, del sistema Capital Bikeshare en Washington D.C. son los siguientes:

Fecha	Estación	Clima	Temp	Humedad	Reg Usua	Cnt Bici
2011-01-01	1	6	0.3441	0.8058	654	985
2011-01-02	1	0	0.3634	0.6960	670	801
2011-01-03	1	1	0.1963	0.4372	1229	1349
2011-01-04	1	1	0.2	0.5904	1454	1562

Tabla 1 Datos recolectados del sistema Capital Bikeshare 2011 y 2012

En la tabla anterior se describen las variables que se van a utilizar las cuales son:

- Fecha
- Estación: Corresponde a la temporada del año (1 Primavera, 2 Verano, 3 Otoño y 4 Invierno)
- Clima: 1 Despejado, 2 Niebla, 3 Nieve y 4 Lluvia
- Temperatura
- Humedad
- Req. Usua: Cantidad de usuarios registrados

- Cnt. Bici: Recuento total de bicicletas de alquiler

De acuerdo con las variables descritas en el texto anterior, la variable dependiente que se utilizará para predecir será la columna “Conteo de bicicletas” que corresponde al total de bicicletas requeridas, la cual dependerá del resto de las variables de la tabla anterior.

La extensión de dos años permite un análisis completo de cómo las condiciones climáticas, la estacionalidad y otros factores pueden afectar los patrones de alquiler de bicicletas. Además, la inclusión de información detallada de estos años específicos permite contextualizar los resultados en un marco temporal específico, considerando posibles eventos externos o cambios en la infraestructura que podrían influir en los datos (Capital Bikeshare DC, 2020.; Fin De Máster et al., 2023.).

La disponibilidad de datos en el portal de http://capitalbikeshare.com/system-data agrega un componente crucial para determinar que técnica de minería de datos permite realizar una mejor predicción. La accesibilidad a la información facilita la replicación de estudios, la validación de resultados y el fomento de la transparencia en la investigación de este artículo. La disponibilidad en línea no solo beneficia a los académicos y personas que analizan los datos, sino que también puede ser utilizada por responsables políticos, urbanistas y cualquier persona interesada en comprender la movilidad urbana. Este enfoque abierto contribuye a la democratización del conocimiento y a la colaboración en la búsqueda de soluciones sostenibles y eficientes para los desafíos urbanos (Ricci, 2015b).

La frecuencia de datos, registrada diariamente, aporta gran detalle temporal al conjunto de datos. Esta alta frecuencia de muestreo permite capturar variaciones diarias en los patrones de alquiler de bicicletas. El análisis a nivel de día ofrece una visión más granular de cómo factores como la fecha, temporada del año, mes y condiciones meteorológicas pueden influir en la demanda de bicicletas y la evaluación de tendencias a lo largo de jornadas completas. Esta riqueza temporal es esencial para comprender la dinámica de uso de las bicicletas compartidas y representa una ventaja significativa en la formulación de modelos predictivos precisos (Vogel et al., 2011).

La relación entre el proceso de alquiler de bicicletas y diversos factores ambientales es un aspecto crucial para comprender los patrones de movilidad en los sistemas de bicicletas compartidas. Las condiciones climáticas ejercen una influencia significativa, ya que los usuarios pueden ser más propensos a utilizar bicicletas en días con climas agradables. La presencia de precipitación, por otro lado, puede disminuir la demanda, ya que muchas personas prefieren evitar andar en bicicleta bajo la lluvia. Además, el día de la semana y la temporada también juegan un papel crucial. Por ejemplo, los patrones de alquiler pueden variar en días laborables con respecto a los fines de semana, así como en verano e invierno. La hora del día añade otra capa de complejidad, ya que los comportamientos de alquiler pueden diferir significativamente entre las horas pico y las horas no pico (Fin De Máster et al., 2020).

El impacto de las condiciones climáticas, la precipitación, el día de la semana, la temporada y la hora del día en los comportamientos de alquiler de bicicletas destaca la interconexión entre la movilidad urbana y el entorno. Las preferencias de los usuarios están estrechamente vinculadas a las condiciones ambientales y estacionales. Por ejemplo, la primavera puede ver un aumento en la demanda de bicicletas debido al clima agradable, mientras que el invierno puede influir en la disminución de los alquileres. La variabilidad en la demanda a lo largo de las horas del día también puede relacionarse con patrones de viaje diarios y la conveniencia percibida del uso de bicicletas en momentos específicos. Este análisis detallado de los factores ambientales proporciona una base sólida para el diseño de modelos predictivos precisos que tengan en cuenta la complejidad de las influencias externas en la movilidad urbana (Alvarez-Valdes et al., 2016).

3. Estado del Arte

La revisión de artículos previos sobre la predicción del alquiler de bicicletas proporciona una perspectiva valiosa sobre los enfoques y métodos utilizados en investigaciones similares. En ellos se han abordado esta temática desde diversas disciplinas, incluyendo la ciencia de datos, el machine learning y la estadística. Se han empleado métodos predictivos tradicionales, como regresiones lineales y series temporales, para entender la relación entre las variables ambientales y estacionales y la demanda de bicicletas (Wang et al., 2019). Asimismo, se ha observado un creciente interés en la aplicación de modelos de machine learning, como regresiones polinómicas, árboles de regresión y redes neuronales, para capturar la complejidad de los patrones de alquiler. Estas investigaciones previas han arrojado luz sobre la importancia de considerar múltiples variables y la necesidad de modelos flexibles y adaptables para la predicción precisa.

Los métodos y enfoques utilizados en investigaciones similares han evolucionado con la expansión de la disponibilidad de datos y el desarrollo de técnicas avanzadas de machine learning. Algunos estudios han incorporado técnicas de agrupamiento para identificar patrones de comportamiento de usuarios similares, mientras que otros han explorado la aplicación de modelos de aprendizaje profundo para capturar relaciones no lineales en los datos (Zheng & Li, 2020). La inclusión de variables interactivas y la consideración de la multicolinealidad entre los factores ambientales han sido áreas de mejora identificadas en la literatura (Garcia-Gutierrez et al., 2014). Además, la selección adecuada de características y la evaluación robusta del rendimiento del modelo son aspectos críticos abordados por investigadores previos. La revisión de estos enfoques proporciona un marco contextual esencial para la formulación de un modelo efectivo en el contexto específico de predicción del alquiler de bicicletas según condiciones ambientales y estacionales.

4. Metodologías de Predicción

En la selección de algoritmos de inteligencia artificial y minería de datos para la predicción del alquiler de bicicletas, se deben considerar cuidadosamente las características del conjunto de datos y los objetivos específicos de la investigación (Implementación de Modelo Supervisado de Aprendizaje de Máquinas Para La Predicción de Alquiler de Bicicletas. | by Javier Ochoa | Medium, 2020.). Algunos de los algoritmos comúnmente utilizados son:

- Regresión lineal
- Regresión polinómica
- Árboles de Regresión
- Redes neuronales

La elección del algoritmo adecuado dependerá de la naturaleza de las relaciones entre las variables y la complejidad del problema. Los modelos de machine learning (Thirumalai & Koppuravuri, 2018.), al permitir la identificación de patrones no lineales, pueden ser particularmente efectivos para capturar la variabilidad en los comportamientos de alquiler asociados con las condiciones ambientales y estacionales.

En este caso se van a elegir 2 técnicas de predicción de minería de datos las cuales son: Árboles de Regresión y la técnica de Regresión Simple.

Los cálculos y modelos matemáticos que se dan entre las técnicas de Árboles de Regresión y Regresión Simple (Series de Tiempo) consideran las características de los datos y la naturaleza del problema de predicción de la demanda de alquiler de bicicletas, debido a que los Árboles de Regresión (Villafañez et al., 2022.) son modelos de aprendizaje supervisado que dividen repetidamente el conjunto de datos en subconjuntos más pequeños, basándose en las características de las variables predictoras, y luego aplican regresión en cada subconjunto para realizar predicciones. Matemáticamente, estos modelos se basan en algoritmos de partición recursiva como es el algoritmo CART (Árboles de Clasificación y Regresión). Por otro lado, la Regresión Simple en Series de Tiempo (Villafañez et al., 2022.)se basa en la relación entre una variable dependiente y una o más variables independientes en una serie temporal. En este enfoque, se utiliza el histórico de la variable dependiente (en este caso, el alquiler de bicicletas) y otras variables independientes (como las condiciones climáticas) para predecir valores futuros. Matemáticamente, la regresión simple se basa en el método de mínimos cuadrados para encontrar la línea de mejor ajuste a los datos observados.

Técnica Minería Datos	Características
Árboles de Regresión	Modelos predictivos formados por reglas binarias que segmentan el espacio de predictores. Fácil interpretación y manejo de interacciones entre variables. No requieren cumplir distribuciones específicas.
Regresión Simple	Basada en la relación entre una variable dependiente y variables independientes en una serie temporal. Utiliza histórico de la variable dependiente para predecir valores futuros.

Tabla 2 Características de la técnica árboles de regresión y regresión simple

Para realizar el comparativo de estas 2 técnicas se va a utilizar Python 3.11, junto con las librerías numpy, pandas, matplotlib y seaborn.

Los árboles de regresión son una herramienta fundamental en el campo de la minería de datos para predecir valores numéricos basados en un conjunto de variables predictoras. Funcionan mediante la segmentación progresiva de los datos en subconjuntos más pequeños, donde se ajustan modelos de regresión en cada subdivisión. Este proceso se lleva a cabo de manera iterativa, dividiendo el conjunto de datos en función de la variable que mejor separa los datos en términos de la variable de respuesta. Cada división se realiza buscando minimizar la varianza del valor de respuesta dentro de cada subconjunto, lo que conduce a un árbol que representa un conjunto de reglas de decisión basadas en las variables predictoras (López Pedraza et al., 2019).

La operación de un árbol de regresión comienza en la raíz, donde se encuentran todos los datos, y luego se ramifica en nodos más pequeños a medida que se realizan divisiones sucesivas. Estas divisiones se basan en criterios que buscan maximizar la homogeneidad de los valores de respuesta dentro de cada nodo resultante. Los criterios comunes incluyen la reducción de la varianza, el coeficiente de correlación o el error cuadrático medio. Cada nodo del árbol representa una regla de decisión que divide el espacio de características en subconjuntos más simples y homogéneos en términos de la variable de respuesta. Este proceso de crecimiento del árbol continúa hasta que se alcanza algún criterio de parada, como un número máximo de niveles en el árbol o un número mínimo de observaciones en cada nodo (López Pedraza et al., 2019).

La regresión simple es un método estadístico utilizado para analizar la relación entre dos variables, donde una variable, llamada variable dependiente o respuesta, se predice en función de una variable independiente o predictora. En este enfoque, se supone que la relación entre las dos variables puede ser modelada mediante una función lineal. Es decir, se busca una línea recta que mejor se ajuste a los datos observados, de manera que pueda predecir el valor de la variable dependiente para diferentes valores de la variable independiente. Esta línea de mejor ajuste se determina minimizando la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo (Carrasquilla-Batista et al., 2016).

La regresión simple utiliza la ecuación de $Y = a + bX$ en donde:

- (Y) representa la variable dependiente.
- (X) representa la variable independiente.

- (a) es la intersección en el eje Y, es decir, el valor de (Y) cuando (X) es 0.
- (b) es la pendiente de la línea, que representa el cambio en (Y) por unidad de cambio en (X).

En el caso del modelo de predicción del uso de la bicicleta de acuerdo a las condiciones climáticas la variable dependiente es el recuento total de las bicicletas de alquiler y las variables dependientes son la temporada del año, el clima, temperatura, humedad y cantidad de usuarios registrados.

Una vez cargado los datos históricos se inicia con el preprocesamiento de datos, es fundamental abordar cuestiones como los valores nulos y la normalización para garantizar la calidad y la eficacia de los modelos. Según López Malca (Implementación de Modelo Supervisado de Aprendizaje de Máquinas Para La Predicción de Alquiler de Bicicletas. | by Javier Ochoa | Medium, 2018.)la gestión de valores nulos puede implicar la eliminación de observaciones incompletas o la imputación de valores faltantes utilizando técnicas como la media o la mediana. Además, la normalización de variables es esencial para llevar a todas las características a una escala comparable, evitando que algunas variables dominen sobre otras. Este proceso facilita la convergencia del modelo durante el entrenamiento y mejora la interpretación de los coeficientes en algoritmos lineales. El preprocesamiento adecuado es un paso crítico para garantizar la robustez y la generalización del modelo.

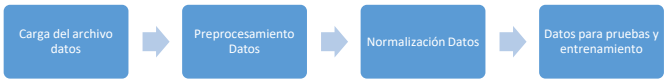

Figura 1 Proceso del procesamiento de datos

La división del conjunto de datos en conjuntos de entrenamiento y prueba es esencial para evaluar la capacidad predictiva del modelo de manera objetiva. Usualmente, se reserva un porcentaje del conjunto de datos para la evaluación del rendimiento del modelo después de entrenarlo. De acuerdo con Zhang (Ma et al., 2022), la elección del tamaño del conjunto de prueba dependerá de la cantidad total de datos y la complejidad del modelo. Un enfoque común es utilizar un 80% de los datos para entrenamiento y el 20% restante para prueba. Esta división permite verificar la capacidad del modelo para generalizar a datos no vistos y brinda una evaluación realista de su rendimiento en situaciones del mundo real.

5. Modelado y Evaluación

La implementación de modelos predictivos en el alquiler de bicicletas es un paso crucial que requiere la traducción de algoritmos seleccionados a un código ejecutable. La elección de herramientas y lenguajes de programación dependerá de la complejidad del modelo y la disponibilidad de bibliotecas especializadas. De acuerdo con Zheng (Zheng & Li, 2020) eneste proceso, se deben cargar los datos de entrenamiento, ajustar los parámetros del modelo y realizar validaciones cruzadas para evitar el sobreajuste. La implementación eficaz garantiza la capacidad del modelo para aprender patrones a partir de los datos de entrenamiento y realizar predicciones precisas en nuevas instancias.

La evaluación de la precisión y rendimiento de los modelos es un paso crítico para validar la eficacia de las predicciones. Se utilizan diversas métricas, como el error cuadrático medio (MSE) o el coeficiente de determinación (R^2), para medir la calidad de las predicciones en relación con los valores reales. Además, se realiza una evaluación del modelo en el conjunto de prueba, que representa datos no vistos durante el entrenamiento. Este proceso ayuda a identificar posibles problemas de sobreajuste o subajuste y proporciona una estimación realista del rendimiento del modelo en situaciones del mundo real. La interpretación de estas métricas no solo permite comparar diferentes modelos, sino que también guía ajustes adicionales para mejorar la precisión predictiva. De acuerdo con Fanaeee (Fanaee-T & Gama, 2014), el error cuadrático medio (MSE) es una medida comúnmente utilizada para evaluar la precisión de los modelos de predicción de bicicletas. Por otro lado, Xu (Xu et al., 2020) sugiere que el coeficiente de determinación (R^2) es una métrica útil para evaluar la calidad de las predicciones en relación con los valores reales. Además, Jelic (Jelic & Roncaglia, 2021) menciona que la evaluación del modelo en el conjunto de prueba es una técnica comúnmente utilizada para identificar posibles problemas de sobreajuste o subajuste y proporcionar una estimación realista del rendimiento del modelo en situaciones del mundo real.

La iteración entre la implementación de modelos predictivos y la evaluación de su rendimiento es una parte fundamental del desarrollo de un sistema predictivo robusto. A medida que se analizan los resultados, se pueden realizar ajustes en los parámetros del modelo o considerar la incorporación de nuevas variables para mejorar la precisión (Thirumalai & Koppuravuri, 2017.). Este ciclo iterativo es esencial para afinar el modelo y garantizar su capacidad de generalización en condiciones diversas. De acuerdo con Jelic (Jelic & Roncaglia, 2021), la implementación y evaluación efectivas se complementan mutuamente en el proceso de construcción de modelos predictivos robustos y aplicables al contexto específico de la predicción del alquiler de bicicletas según condiciones ambientales y estacionales.

La proyección de la información respecto a la renta de bicicletas implica evaluar cómo cada técnica modela y predice la demanda de alquiler de bicicletas en función de las condiciones ambientales y estacionales. Con los Árboles de Regresión, se espera que el modelo capture relaciones no lineales y complejas entre las variables predictoras y la variable objetivo, lo que podría ser beneficioso cuando hay interacciones no lineales o efectos de umbral en los datos. Por otro lado, la Regresión Simple en Series de Tiempo se centra en modelar la tendencia y la estacionalidad en los datos de alquiler de bicicletas a lo largo del tiempo, lo que puede ser útil

para capturar patrones temporales y estacionales en la demanda. La elección entre estas técnicas dependerá de la naturaleza de los datos y de la complejidad del problema de predicción.

El resultado de elegir una de las dos técnicas dependerá de varios factores, incluida la calidad de los datos, la interpretabilidad del modelo y el rendimiento predictivo. Los Árboles de Regresión pueden proporcionar modelos más flexibles y capaces de manejar relaciones complejas entre variables, pero pueden ser propensos a sobreajuste si no se controlan adecuadamente. Por otro lado, la Regresión Simple en Series de Tiempo puede proporcionar modelos más interpretables y fácilmente interpretables, pero pueden ser menos flexibles para capturar relaciones no lineales en los datos.

6. Resultados

La presentación de los resultados obtenidos en la predicción del alquiler de bicicletas basada en condiciones ambientales y estacionales es un momento crucial en el proceso de este artículo. Los resultados se expresarán en términos de métricas de evaluación, como el error cuadrático medio (MSE) o el coeficiente de determinación (R^2) (Thirumalai & Koppuravuri, 2017.), que proporcionarán una medida cuantitativa de la precisión del modelo. Se presentarán visualizaciones gráficas, como gráficos de dispersión entre las predicciones y los valores reales, para una comprensión intuitiva de la calidad del modelo. Además, se examinarán las características más influyentes identificadas (Xu et al., 2020) por el modelo para entender cómo las condiciones ambientales y estacionales impactan la demanda de alquiler de bicicletas.

Para la técnica de árboles regresión se utilizó la librería de from sklearn.tree import DecisionTreeRegressor en Spyder y se utilizó el 25% de los 731 registro, donde nos mostró un score de predicción de 0.8362 para la técnica de regresión se utilizó el 25% de los 731 registro, donde nos mostró un score de predicción de 0.8804 quedando de la siguiente forma:

	Árboles Regresión	Regresión Simple
Total Registros	731	731
Datos Prueba	183	183
Datos Entrenamiento	548	548
Score Predicción	0.8362	0.8804

Figura 2 Comparación de los datos con las 2 técnicas de predicción

La interpretación de la relación entre condiciones ambientales, estacionales y las predicciones de alquiler de bicicletas implica analizar la importancia relativa de cada variable en el modelo. Se buscarán patrones y tendencias en las predicciones para entender cómo factores como la temperatura, la temporada o la hora del día afectan la demanda de bicicletas (Jelic & Roncaglia, 2021). Este análisis permitirá una comprensión más profunda de los comportamientos de los usuarios y cómo responden a diferentes condiciones. Además, se explorarán posibles interacciones entre variables para identificar relaciones complejas que podrían no ser evidentes de manera aislada. La interpretación de estos resultados contribuirá al conocimiento general sobre la influencia de factores ambientales y estacionales en la movilidad urbana a través de sistemas de bicicletas compartidas.

La coherencia entre la presentación de resultados y la interpretación de la relación entre condiciones ambientales, estacionales y predicciones de alquiler de bicicletas es esencial para proporcionar una visión integral y significativa del estudio. Este análisis final no solo valida la efectividad del modelo propuesto, sino que también contribuye al entendimiento más profundo de cómo los factores ambientales y estacionales impactan la dinámica de la movilidad en entornos urbanos.

Al comparar las 2 técnicas, la que se acercó a una mayor predicción fue “Predicción” ya que tuvo un mayor grado de eficiencia:

7. Conclusiones

La comparación entre la utilización de la técnica de minería de regresión y árboles de regresión para la predicción del alquiler de bicicletas son fundamentales para entender la eficacia de ambos enfoques en este contexto específico. La minería de regresión, al basarse en la modelización de relaciones lineales entre variables, puede ofrecer resultados precisos cuando la relación entre condiciones ambientales y estacionales es predominantemente lineal. Su capacidad para cuantificar la influencia de cada variable de manera directa puede ser una ventaja en la interpretación de los resultados y la identificación de factores más influyentes. Sin embargo, puede no capturar de manera óptima relaciones no lineales complejas presentes en conjuntos de datos más intrincados.

Por otro lado, la utilización de árboles de regresión ofrece una aproximación más flexible y adaptable. Estos modelos pueden manejar relaciones no lineales de manera más efectiva, permitiendo capturar patrones complejos y no evidentes en un primer análisis. Su capacidad para segmentar el espacio de características de manera jerárquica los hace adecuados para conjuntos de datos con interacciones no lineales. Sin embargo, la interpretación de árboles de regresión puede ser más compleja, ya que la relación entre las variables se expresa en términos de condiciones de bifurcación.

Puede ser beneficioso explorar y comparar ambos enfoques, incluso combinándolos en un modelo conjunto para aprovechar las fortalezas de cada uno. La elección entre minería de regresión y árboles de regresión dependerá de la complejidad subyacente de las relaciones en los datos y los objetivos específicos de la predicción del alquiler de bicicletas según las condiciones ambientales y estacionales.

Referencias

Alvarez-Valdes, R., Belenguer, J. M., Benavent, E., Bermudez, J. D., Muñoz, F., Vercher, E., & Verdejo, F. (2016). Optimizing the level of service quality of a bike-sharing system. *Omega*, *62*, 163–175. https://doi.org/10.1016/j.omega.2015.09.007

Capital Bikeshare DC. (n.d.). Retrieved January 23, 2024, from https://capitalbikeshare.com/

Carrasquilla-Batista, A., Chacón-Rodríguez, A., Núñez-Montero, K., Gómez-Espinoza, O., Valverde-Cerdas, J., & Guerrero-Barrantes, M. (2016). Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal. *Revista Tecnología En Marcha*, *29*(8), 33. https://doi.org/10.18845/tm.v29i8.2983

Eren, E., & Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, *54*, 101882. https://doi.org/10.1016/J.SCS.2019.101882

Fanaee-T, H., & Gama, J. (2014). Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, *2*(2–3), 113–127. https://doi.org/10.1007/s13748-013-0040-3

Fin De Máster, T., Beltrante, A., & Santana, A. E. (n.d.). *Predicción del uso de bicis compartidas dependiendo de las condiciones climáticas del día*.

Gámez-Pérez, K., López, P. E. A., & Iniestra, J. G. (2020). Supporting the strategic design of public bicycle sharing systems: The experience of a large Mexican city. *Contaduria y Administracion*, *65*(3). https://doi.org/10.22201/FCA.24488410E.2020.2192

Garcia-Gutierrez, J., Romero-Torres, J., & Gaytan-Iniestra, J. (2014). Dimensioning of a Bike Sharing System (BSS): A study case in Nezahualcoyotl, Mexico. *Procedia - Social and Behavioral Sciences*, *162*, 253–262. https://doi.org/10.1016/j.sbspro.2014.12.206

Gong, W., Rui, J., & Li, T. (2024). Deciphering urban bike-sharing patterns: An in-depth analysis of natural environment and visual quality in New York's Citi bike system. *Journal of Transport Geography*, *115*, 103799. https://doi.org/10.1016/J.JTRANGEO.2024.103799

Guo, Y., Yang, L., & Chen, Y. (2022). Bike Share Usage and the Built Environment: A Review. In *Frontiers in Public Health* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fpubh.2022.848169

Implementación de modelo supervisado de aprendizaje de máquinas para la predicción de alquiler de bicicletas. | by Javier Ochoa | Medium. (n.d.). Retrieved January 23, 2024, from https://medium.com/@javier8amoreno/implementaci%C3%B3n-de-modelo-supervisado-de-aprendizaje-de-m%C3%A1quinas-para-la-predicci%C3%B3n-de-alquiler-de-d504b046e09b

Jelic, A., & Roncaglia, P. (2021). *Predicting bike sharing demand with machine learning*.

López Pedraza, F. J., Macias González, Ma. D. C., & Sandoval García, E. R. (2019). Minería de datos: identificando causas de deserción en las instituciones públicas de Educación Superior de México. *TIES, Revista de Tecnología e Innovación En Educación Superior*, *2*, 1–14. https://doi.org/10.22201/dgtic.26832968e.2019.2.4

Ma, X., Zhang, S., Jin, Y., Zhu, M., & Yuan, Y. (2022). Identification of metro-bikeshare transfer trip chains by matching docked bikeshare and metro smartcards. *Energies*, *15*(1). https://doi.org/10.3390/en15010203

Ricci, M. (2015a). Bike sharing: A review of evidence on impacts and processes of implementation and operation. *Research in Transportation Business & Management*, *15*, 28–38. https://doi.org/10.1016/J.RTBM.2015.03.003

Ricci, M. (2015b). Bike sharing: A review of evidence on the impacts and processes of implementation and operation. *Research in Transportation Business & Management*, *15*, 28–38. https://doi.org/10.1016/j.rtbm.2015.03.003

Rosales-Asensio, E., Borge-Diez, D., Blanes-Peiró, J. J., Pérez-Hoyos, A., & Comenar-Santos, A. (2019). Review of wind energy technology and associated market and economic conditions in Spain. *Renewable and Sustainable Energy Reviews*, *101*, 415–427. https://doi.org/10.1016/J.RSER.2018.11.029

Shaheen, S., Cohen, A., & Martin, E. (2013). Public bikesharing in North America. *Transportation Research Record*, *2387*, 83–92. https://doi.org/10.3141/2387-10

Thirumalai, C., & Koppuravuri, R. (n.d.). *Bike Sharing Prediction using Deep Neural Networks*.

Villafañez, F., Escudero, G., & Ángel, L. (n.d.). *MÉTODOS DE REGRESIÓN Y CLASIFICACIÓN BASADOS EN ÁRBOLES*.

Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns. *Procedia Social and Behavioral Sciences*, *20*, 514–523. https://doi.org/10.1016/j.sbspro.2011.08.058

Wang, J., Huang, J., & Dunford, M. (2019). Rethinking the utility of public bicycles: The development and challenges of station-less bike sharing in China. *Sustainability (Switzerland)*, *11*(6). https://doi.org/10.3390/su11061539

Xu, M., Liu, H., & Yang, H. (2020). A Deep Learning Based Multi-Block Hybrid Model for Bike-Sharing Supply-Demand Prediction. *IEEE Access*, *8*, 85826–85838. https://doi.org/10.1109/ACCESS.2020.2987934

Zheng, L., & Li, Y. (2020). The development, characteristics and impact of bike-sharing systems: A literature review. In *International Review for Spatial Planning and Sustainable Development* (Vol. 8, Issue 2, pp. 37–52). SPSD Press. https://doi.org/10.14246/irspsd.8.2_37

Zhou, J., Guo, Y., Sun, J., Yu, E., & Wang, R. (2022). Review of bike-sharing system studies using bibliometrics method. *Journal of Traffic and Transportation Engineering (English Edition)*, *9*(4), 608–630. https://doi.org/10.1016/J.JTTE.2021.08.003

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

ISSN: 2992-7315

Publicación Semestral RICT Vol. 2 No. 3 (2024) Páginas 26 – 32

Arquitectura de gestión de red para la monitorización y control de información de nodos de red de la subdirección de informática de la empresa CDS, S.C.

Network management architecture for the monitoring and control of information from network nodes of the IT department of the company CDS, S.C.

Abraham-Jorge Jiménez-Alfaro[a], Edgar Corona-Organiche[a], Claudia-Teresa González-Ramírez[b],Jhacer-Kharen Ruiz-Garduño[b], Griselda Cortes-Barrera[a]

[a]Ingeniería en Sistemas Computacionales, TECNM/Tecnológico de Estudios Superiores de Ecatepec, Valle de Anáhuac, 55210 Ecatepec de Morelos, Estado de México.
[b]Ingeniería en en Sistemas Computacionales , TECNM/Instituto Tecnológico de Zitácuaro, Av. Tecnológico No.186, 61534, Zitácuaro, Michoacán.

Resumen

La gestión de red se suele centralizar en un centro de gestión, donde se controla y vigila el correcto funcionamiento de todos los equipos integrados en las distintas redes de la empresa en cuestión. Un centro de gestión de red dispone de tres tipos principales de recursos: Métodos de gestión, recursos humanos y herramientas de apoyo que facilitan las tareas de gestión a los operadores de red. La Situación actual de la empresa CDS, S.C. es que los nodos de red ante eventos adversos quedan deshabilitados provocando que sea presencial la solución del evento y la puesta en marcha del nodo de red; esto provoca que no se pueda mantener la funcionalidad de la red de forma inmediata ya que existe el retraso considerable ya que es: el traslado del personal al sitio, la detección del evento de red y la solución de configuración. Ante esta situación se desarrolla una arquitectura y aplicativo de Gestión de Red para la monitorización, configuración de información y control necesarios para operar de manera efectiva la red de comunicaciones de la subdirección de informática de la empresa CDS, S.C. las 24 Horas del día los 7 días de la semana. Estas tareas están distribuidas sobre diferentes nodos de la red, lo cual requiere repetidas acciones de recogida de datos, configuración y análisis cada vez que sucede un nuevo evento y mantener la funcionalidad de la red. Dada esta situación se hace necesario la gestión de la red para mantener la operación desde cualquier punto de acceso a la red. Para ello se emplean agentes del protocolo de administración simple de Red (SNMP), el acceso a la base de datos jerárquica de variables de información gestionada del dispositivo (MIB) y la estructura del paradigma gestor-agente.

Palabras clave: Arquitectura de Gestión, Agentes SNMP, Base de Información Gestionada MIB.

Abstract

Network management is usually centralized in a management center, where the correct functioning of all the equipment integrated into the different networks of the company in question is controlled and monitored. A network management center has three main types of resources: Management methods, Human resources and Support tools that facilitate management tasks for network operators. The current situation of the company CDS, S.C.is that the network nodes in the event of adverse events are disabled, causing the solution of the event and the start-up of the network node to be in person; This means that network functionality cannot be maintained immediately since there is a considerable delay in: the transfer of personnel to the site, the detection of the network event and the configuration solution. Given this situation, a Network Management architecture and application is developed for the monitoring, configuration of information and control necessary to effectively operate the communications network of the IT subdirectorate of the company CDS, S.C. 24 hours a day, 7 days a week. These tasks are distributed over different nodes of the network, which requires repeated actions of data collection, configuration and analysis every time a new event happens and maintain the functionality of the network. Given this situation, network management is necessary to maintain operation from any network access point. For this purpose, Simple Network Management Protocol (SNMP) agents are used, access to the device's hierarchical database of Managed Information Variables (MIB) and the structure of the manager-agent paradigm.

Keywords: Management Functional architecture, SNMP Agents, Managed Information Base MIB.

* Autor para la correspondencia: josejuan@dominio1.edu.mx
Correo electrónico: ajimenez@tese.edu.mx @tese.edu.mx (Abraham-Jorge Jiménez-Alfaro), ecorona@tese.edu.mx x (Edgar Corona-Organiche)
Historial del manuscrito: recibido el 01/02/2024, última versión-revisada recibida el 03/04/2024, aceptado el 04/04/2024, en línea (postprint) desde el 08/04/2024, publicado el 09/04/2024. **DOI:** https://doi.org/10.2992/rict.v2i3.45

1. Introducción

La información de configuración describe la naturaleza y estado de los recursos que forman la red (tanto físicos como lógicos). Esta información incluye una especificación del recurso y de los atributos de ese recurso (p.e., nombre, dirección, número de identificación, estados, características operacionales, versión del software, entre otros). Esta información (en realidad, toda la información de gestión) se puede estructurar de diversas formas desde los agentes SNMP y una base a objetos o variables de red (Stallings, 2015). El emplear agentes SNMP permite que la información sea accesible por agentes de administración y agentes clientes asociadas a los nodos o dispositivos a gestionar. Esta función permite al administrador especificar el rango y el tipo de valores que puede tomar un determinado atributo de un agente, el rango puede ser una lista de todos los valores posibles o los valores superior e inferior permitidos; así como, modificar los objetos o variables del dispositivo de red. Para accesar a estos objetos o variables on-line, en línea, se deben de crear los correspondientes agentes de administración y clientes empleando el protocolo de administración simple de red (SNMP) junto al paradigma gestor-agente. Una vez modificadas las variables es necesario inicializar el dispositivo o nodo para que se incorporen los nuevos parámetros de configuración.

Para el caso de la Subdirección de Informática de la empresa CDS, S.C., se hace necesario que los recursos de red estén activos y funcionales las 24 horas del día los 7 días de la semana para atender los diversos servicios intrínsecos de la actividad de la organización, por lo que ante un evento en la red, esta debe de restablecerse de forma inmediata sin necesidad de la presencia del personal de redes para reestablecer el servicio, ya que esto provocaría pérdidas cuantiosas en términos monetarios.

2. Materiales y Método

2.1.- Arquitectura de Gestión de Red

Para Stallings (2017) La gestión de red es el conjunto de tareas de monitorización, configuración, información y control, necesarias para operar de manera efectiva una red. Estas tareas pueden estar distribuidas sobre diferentes nodos de la red, lo cual puede requerir repetidas acciones de recogida de datos y su análisis, cada vez que sucede un nuevo evento en la red. Las redes son cada vez más importantes en empresas y organizaciones tomando en consideración: la tendencia a redes más grandes, más complejas, más heterogéneas; la red y las aplicaciones distribuidas se hacen imprescindibles; los costos de gestión de la red aumentan; la gestión de la red no se puede hacer manualmente, se requieren herramientas de gestión de red automatizadas.

La Arquitectura de Gestión de red desarrollada se fundamenta en el estándar ISO de gestión de redes que clasifica las tareas de los sistemas de gestión en cinco áreas funcionales. **La tarea del encargado de gestionar la red empresarial de CDS, S.C., será emplear el aplicativo de gestión de red considerando cada una de estas áreas: Gestión de configuración, Gestión de rendimiento, Gestión de contabilidad, Gestión de fallos y Gestión de seguridad (Stallings, 2019).**

Gestión de configuración

El objetivo de la gestión de configuración es obtener datos de la red y utilizarlos para incorporar, mantener y retirar los distintos componentes y recursos a integrar. Consiste en la realización de tres tareas fundamentales: Recolección automatizada de datos sobre el inventario y estado de la red, tales como versiones software y hardware de los distintos componentes; cambio en la configuración de los recursos y almacenamiento de los datos de configuración.

Gestión de rendimiento

La gestión de prestaciones o del rendimiento tiene como objetivo principal el mantenimiento del nivel de servicio que la red ofrece a los usuarios, asegurándose de que está operando de manera eficiente en todo momento. La gestión de prestaciones se basa en cuatro tareas: Recogida de datos, variables indicadoras de rendimiento, tales como el rendimiento de la red; los tiempos de respuesta o latencia y la utilización de la línea.

Análisis de los datos para determinar los niveles normales de rendimiento.

Establecimiento de umbrales, como indicadores que fijan los niveles mínimos de rendimiento que pueden ser tolerados y determinación de un sistema de procesado periódico de los datos de prestación de los distintos equipos, para su estudio continuado.

Gestión de contabilidad

La gestión de contabilidad tiene como misión la medida de parámetros de utilización de la red que permitan preparar las correspondientes facturas a sus clientes. Entre las tareas que se deben realizar en esta área, están: Recolección de datos sobre la utilización de los recursos.

Establecimiento de cuotas.

Cobro a los usuarios con las tarifas derivadas de la utilización de los recursos.

Gestión de fallos

La gestión de fallos tiene por objetivo fundamental la localización y recuperación de los problemas de la red. La gestión de problemas de red implica las siguientes tareas:

- Determinación de los síntomas del problema.
- Aislamiento del fallo.
- Resolución del fallo.
- Comprobación de la validez de la solución en todos los subsistemas importantes de la red.
- Almacenamiento de la detección y resolución del problema.

Gestión de seguridad

La misión de la gestión de seguridad es ofrecer mecanismos que faciliten el mantenimiento de políticas de seguridad (orientadas a la protección contra ataques de intrusos). Entre las funciones realizadas por los sistemas de gestión de seguridad, están:

- Identificación de recursos sensibles en la red, tales como archivos o dispositivos de comunicaciones.
- Determinación de las relaciones entre los recursos sensibles de la red y los grupos de usuarios.
- Monitorización de los puntos de acceso a los recursos sensibles de red.
- Almacenamiento de los intentos de acceso no autorizados a estos recursos, para su posterior análisis.

Para cumplir estas premisas las funciones de la arquitectura de gestión de red se agrupan en dos categorías:

Monitorización, funciones de "lectura":

- Observar y analizar el estado y comportamiento de la configuración de red y sus componentes.
- Abarca: prestaciones, fallos y costos.

Control, funciones de "escritura":

- Alterar parámetros de los componentes de la red.
- Abarca: configuración y seguridad.

La información a monitorizar se encuentra en una de las siguientes categorías:

1.- Estática: Caracteriza la configuración actual de la red y de sus elementos (p.e., el número de enlaces de un ruteador o concentrador, entre otros). Esta información cambiará con muy poca frecuencia. **La información estática** es generada y almacenada por el propio elemento de red (p.e., un ruteador almacena su propia configuración).

2.- Dinámica: Información relacionada con eventos en la red (p.e., la transmisión de un paquete por la red). **La información dinámica** puede almacenarla el propio elemento, u otro encargado de ello (p.e., en una red de área local (*LAN*)) cada elemento puede almacenar el número total de paquetes que envía, o un elemento de la LAN puede estar escuchando y recoger esa información.

3.- Estadística: Para Comer (2000) y Craig (2003) Información que puede ser derivada de la información dinámica (p.e., el número medio de paquetes transmitidos por unidad de tiempo por un sistema final).**La información estadística** se genera por cualquier elemento que tenga acceso a la información dinámica en base a dos opciones básicas:

1.- Puede enviarse toda la información dinámica al gestor de red para que realice las estadísticas.

2.- Si el gestor no necesita toda la información, ésta puede ser resumida por el propio elemento antes de enviarla al gestor, ahorrando procesamiento en el gestor y generando menos tráfico en la red.

Los elementos de la arquitectura de gestión de red , bajo el paradigma gestor-agente, se clasifican en dos grandes grupos:

- Los gestores son los elementos del sistema de gestión que interaccionan con los operadores humanos y desencadenan acciones necesarias para llevar a cabo las tareas por ellos invocadas.

- Los agentes, por otra parte, son los componentes del sistema de gestión invocados por el gestor o gestores de la red.

Para Huitema(2001) el principio de funcionamiento reside en el intercambio de información de gestión entre nodos gestores y nodos gestionados. Habitualmente, los agentes mantienen en cada nodo gestionado información acerca del estado y las características de funcionamiento de un determinado recurso de la red. El gestor pide al agente, a través de un protocolo de gestión de red, que realice determinadas operaciones con estos datos de gestión, gracias a las cuales podrá conocer el estado del recurso y podrá influir en su comportamiento, véase figura. 1.

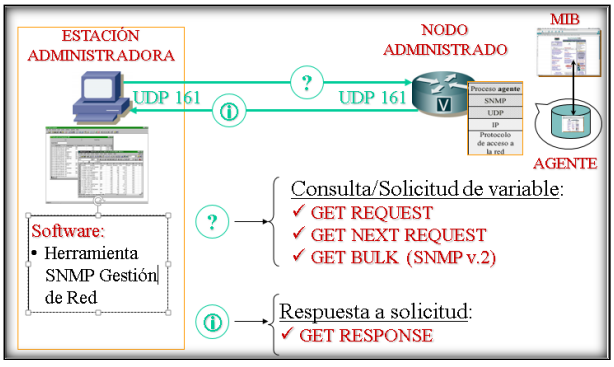

Figura 1.- Arquitectura de gestión de Red, obtención de Información.

3. Resultados

En Kirch (2001) SNMP hace uso de ASN.l (abstract syntax notation one) para describir la estructura de los objetos del MIB. La definición de tipos mediante ASN.l se apoya en las siguientes estructuras (véase tabla. 1 a 5):

- Tipos simples que representan directamente valores numéricos, de cadena, lógicos, etc.
- Tipos estructurados que se apoyan en los tipos simples para construir estructuras más complejas.
- Etiquetas que añaden información adicional a los tipos para facilitar su codificación.
- Módulos que agrupan definiciones de tipos.

Tabla 1.- Tipos Básicos

Etiqueta	Nombre	Descripción
UNIVERSAL 1	Bolean	Verdadero o falso
UNIVERSAL 2	Integer	Numeros enterospositivos o negativos, sin restricción en el numero de digitos
UNIVERSAL 3	Bit String	Secuencia de bits, con posibilidad de hacer referencia a cada bit por su posición
UNIVERSAL 4	Octet String	Secuencia de octetos
UNIVERSAL 9	Real	Numero real
UNIVERSAL 10	Enumerated	Tipo enumerado

Tabla 2.- Cadenas de caracteres

Etiqueta	Nombre	Descripción
UNIVERSAL 22	IA5String	Alfabeto internacional 5
UNIVERSAL 26	VisibleString	Subconjunto de IA5 de 95 caracteres, unicamente se encuentran los caracteres visible y el espacio
UNIVERSAL 19	PrintableString	Subconjunto de IA5 de 74 caracteres, unicamente se encuentran algunos caracteres especiales
UNIVERSAL 18	NumericString	Subconjunto de IS5 compuesto de los digitos de 0 a 9 y el espacio
UNIVERSAL 25	GraphicString	Extensión de IA5 que añade caracteres graficos
UNIVERSAL 27	GeneralString	Extensión de IA5 que añade caracteres graficos y de control
UNIVERSAL 20	TeletexString	Subconjunto teletex del CCITT
UNIVERSAL 21	VideotexString	Subconjunto videotex del CCITT

Tabla 3.-Varios

Etiqueta	Nombre	Descripción
UNIVERSAL 5	NULL	Valor nulo
	ANY	Equivale a cualquier tipo
UNIVERSAL 8	EXTERNAL	Para incorporar tipos externos que no han sido definidos en ANS.1
UNIVERSAL 23	Hora UTC	Fecha en formato YYMMDDHHMMSS
UNIVERSAL 24	Hora General	Fecha en formato YYYYMMDDHHMMSS.S

Tabla 4-Objetos

Etiqueta	Nombre	Descripción
UNIVERSAL 6	OBJECT IDENTIFIER	Cadena de números que identifica a un objeto
UNIVERSAL 7	Object Designer	Texto informativo de un objeto

Tabla 5.-Constructores

Etiqueta	Nombre	Descripción
	CHOICE	Lista de tipos alternativos
UNIVERSAL 16	SEQUENCE SEQUENCE OF	Lista ordenada de identificadores de tipo. En caso de usar los mismo tipos se utiliza SEQUENCE Of sino SEQUENCE
UNIVERSAL 17	SET SET OF	Lista en la que no importa el orden de identificadores de tipos. Si todos tienen el mismo tipo se usa SET OF, en otro caso SET

Con estos elementos se implementan los gestores y agentes de la arquitectura que permitirán gestionar un nodo de red, véase figura 2.

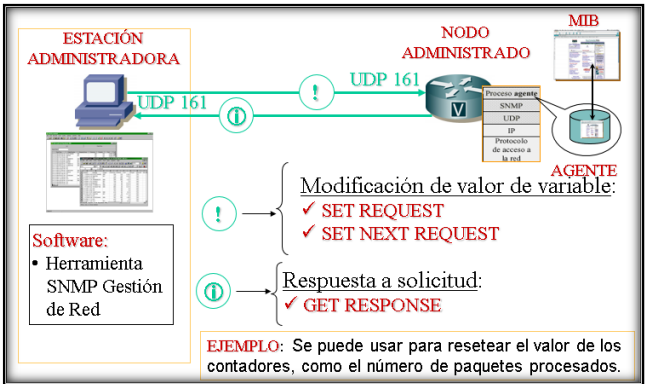

Figura 2.- Elementos gestionados por la arquitectura con gestor y agente *SNMP*, modificación de información.

Por lo que una codificación es como sigue, véase figura 3:

```
Thost ::= SEQUENCE {
        Nombre          VisibleString,
        dirIP           NumericString,
        dirMAC          NumericString,
        Infomacion      VisibleString,
}
```

Figura 3.- Codificación *ASN.1*.

ASN.1 propone una sintaxis de transferencia en la que todos los valores se codifican siguiendo el siguiente formato, véase figura 4.

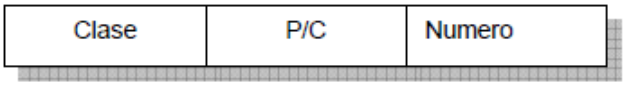

Figura 4.- Identificador *ASN.1*

El campo clase son 2 bits e indica el tipo de etiqueta, su codificación es la siguiente, véase tabla 6:

Tabla 6.-Etiqueta

Codificacion	Descripción
00	Universal
01	Aplicación
10	Contexto
11	Privada

El campo P/C es un bit e indica si el contenido es primitivo, es decir, si se trata de un dato final, o es un constructor. Un valor de 1 indica que es un tipo primitivo, un valor de 0 indica que es un constructor.

El campo Número identifica una etiqueta concreta dentro de la clase indicada. El principio se cuenta con 5 bits para codificar el número de etiqueta, lo que proporciona un rango

del 0 al 63. Si se necesita un número mayor se pueden usar más bytes para codificar el número de etiqueta con el siguiente formato, véase figura 5.

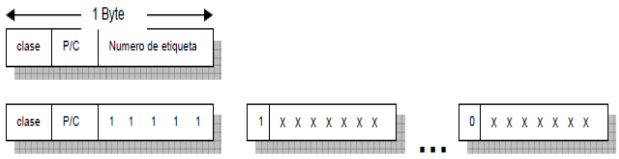

Figura 5.- Campo Numero *ASN.1*.

La longitud de los datos indica cuántos bytes vienen a continuación, es decir, cuánto espacio ocupa la información que se envía (el contenido del campo que se es codificando). Se puede indicar una longitud concreta, o bien se puede indicar una longitud indefinida, marcando el final de los datos con dos bytes a 0, véase figura 6.

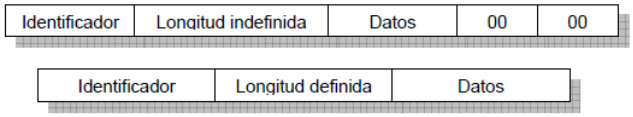

Figura 6.- Campo Longitud *ASN.1*.

El **agente de gestión** se encarga de supervisar un elemento de la red. Se comunica con el sistema gestor para atender las peticiones y para informarle de eventos sucedidos en el objeto gestionado, este agente de gestión es parte del dispositivo; el **gestor** es el software residente en una estación de gestión que se comunica con los agentes y que ofrece al usuario una interfaz a través de la cual comunicarse con los elementos de gestión para obtener información de los recursos gestionados. Además, recibe las notificaciones enviadas por los agentes, este componente se programa y forma parte de del programa de configuración; los **objetos gestionados** son las abstracciones de los elementos físicos de la red que se gestionan (tarjeta de red, concentrador, módem, ruteador, entre otros.). Se pueden manejar los atributos y las operaciones que se puedan realizar sobre el objeto. De la misma forma, las notificaciones que dicho objeto puede generar, así como, las relaciones con otros objetos de la red también son susceptibles de ser controladas. La base de datos de gestión (*MIB*) previamente, mencionada está formada por todos los objetos gestionados, ésta *MIB* es parte del dispositivo gestionado; finalmente **protocolo de gestión,** *SNMP*, especifica como se realiza la comunicación entre los agentes de gestión y el gestor (Stallings, 2019).

Para implementar la arquitectura funcional de gestión se realizó un software modular, que funciona en ambientes Windows y Linux, que permite accesar a las primitivas monitorización, control y configuración asociadas al protocolo SNMP, véase figura 7.

Figura 7.- Estructura modular para implementar la arquitectura funcional de monitorización por *SNMP*.

Una de las claves de la flexibilidad de la arquitectura de funcional de gestión es el empleo de *SNMP y* el uso de "variables" como forma de representación de los recursos, tanto físicos como lógicos, en los sistemas gestionados. En cada nodo gestionado, el agente *SNMP* proporciona una base de datos llamada *MIB* (Management Information Base*)*, que contiene objetos de datos, más conocidos como variables *MIB*. La monitorización del nodo la lleva a cabo la estación gestora, la cual, periódicamente, lee los valores de estas variables. El control del nodo se realiza mediante el cambio de los valores de las variables. Adicionalmente, existe una operación ***trap*** para permitir al nodo gestionado informar a una estación gestora sobre determinadas condiciones o eventos inusuales.

El menú SNMP *(SIMPLE NETWORK MANAGEMENT PROTOCOL)*, tiene las principales primitivas para manipular la MIB *(MANAGEMENT INFORMATION BASE)*, además de un módulo para las estadísticas del sistema, véase figura 8.

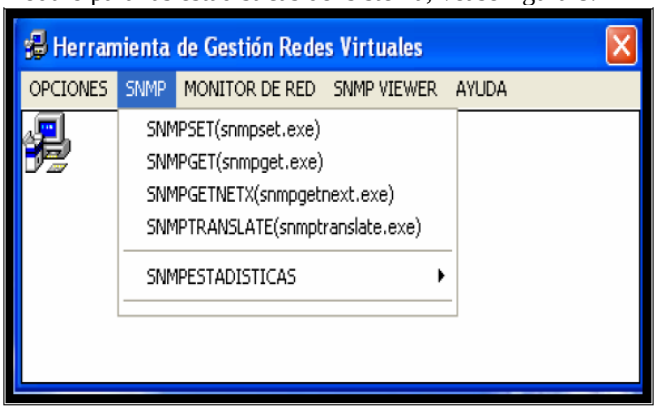

Figura 8.- Modelo de configuración estación Gestora y sistemas Gestionados.

Cada opción que se elija de los mandatos SNMP tiene un editor de ayuda para poder imprimirla o adicionarle nueva información, así como una ventana de consola para poder ejecutar el programa respectivamente, véase figura 9.

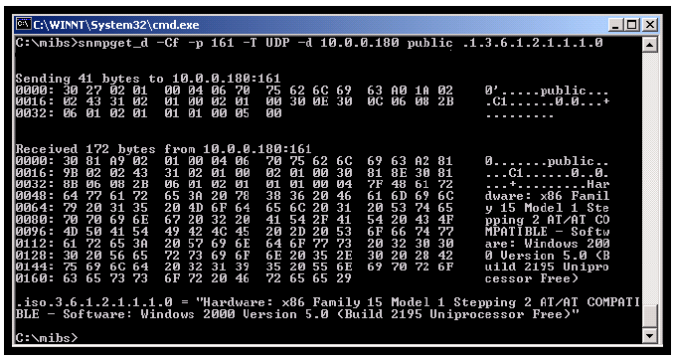

Figura 9.- Información de sistemas Gestionados.

4. Discusión

Los entornos de red, sus configuraciones y funcionamiento forman parte de la gestión de red, algunos de los elementos en los que se hace hincapié de forma determinantes en la configuración y el monitoreo, son las variables que residen en la base de datos del los dispositivos que interviene en la red (Stallings,2015). El protocolo de administración de red simple, *SNMP* (Simple Network Management Protocol) permite el acceso a las bases de datos del dispositivo y modificar parámetros asociados a las mismas, con lo que los cambios se reflejan de manera inmediata. *SNMP* es un protocolo que permite la gestión de los recursos que están disponibles en una red. Dentro de un entorno de red gestionado con *SNMP* habrá un conjunto de nodos de la red que se encarguen de la gestión y un conjunto de componentes de la red (hosts, concentradores, ruteadores, modems, entre otros.) que podrán ser gestionados por estas estaciones. Así como, la base de datos donde se encuentra toda la información que se gestiona. Esta base de datos se denomina *MIB* (Management Information Base).

Para la arquitectura de gestión de red se consideran los siguientes elementos para su implementación:

- Agente de gestión.
- Gestor
- Objeto gestionado
- Protocolo de gestión.

El agente de gestión se encarga de supervisar un elemento de la red. Se comunica con el gestor para atender sus peticiones y para informarle de eventos acaecidos en el objeto gestionado. El agente de gestión suele residir físicamente en el elemento gestionado.

El gestor, es un software residente en una estación de gestión que se comunica con los agentes y que ofrece al usuario una interfaz a través de la cual comunicarse con los agentes de gestión.

Los objetos gestionados, son las abstracciones de los elementos físicos de la red que se gestionan (tarjeta de red, concentrador, módem, ruteador, etc.). Se pueden manejar los atributos y las operaciones que se pueden realizar sobre el objeto. De la misma forma, las notificaciones que dicho objeto puede generar, así como, las relaciones con otros objetos también son susceptibles de ser controladas. La base de datos de gestión (*MIB*) previamente mencionada está formada por todos los objetos gestionados.

Protocolo de gestión, es el protocolo que especifica cómo se realizará la comunicación entre los agentes de gestión y el gestor (Liu,.2000). En nuestro caso este protocolo es el *SNMP*. La comunicación se realiza en base a requerimientos, respuestas y notificaciones, véase figura. 10.

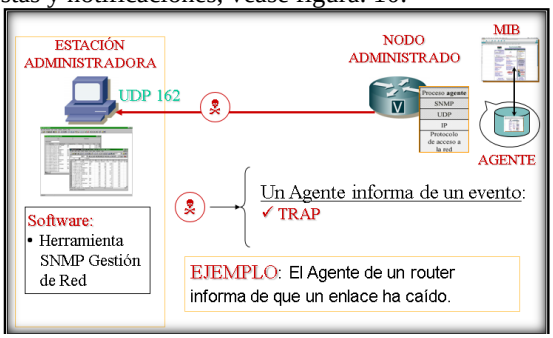

Figura 10.- Elementos de gestión de la arquitectura de gestión de Red, Interrupciones.

Los mensajes *SNMP* implementados son:

1.- **Get Request**: Para leer el valor de una o varias variables del *MIB*.
2.- **Get Next Request**: Para realizar lecturas secuenciales a través del *MIB*.
3.- **Get.Response**: Es el mensaje de respuesta a un Set Request, Get Request o Get Next Request.
4.- **Set Request**: Mensaje enviado para establecer el valor de una variable.
5.- **Trap**: A través de este mensaje se hacen notificaciones de eventos.

Estos cinco tipos de mensajes *SNMP* son encapsulados en datagramas *UDP*. Los mensajes de petición y respuesta son enviados al puerto **161**, mientras que las notificaciones de eventos usan el puerto **162**. El *MIB* se organiza en forma de árbol. Cada nodo del árbol tiene asociado un número entero y una etiqueta de texto. Un nodo se identifica unívocamente con una secuencia de números enteros que identifican los nodos a través de los cuales hay que pasar para llegar desde la raíz al nodo que interese.

5. Agradecimientos

A la empresa CD, S.C., por el apoyo técnico y documental, así como, a la subdirección de informática por apoyo para accesar a las instalaciones, las configuraciones de los nodos, las redes y efectuar las pruebas de configuración con el aplicativo de gestión de redes vía *SNMP*. A Cada uno de los integrantes de artículo por el aporte computacional para realizar el aplicativo.

6. Referencias

Comer, D(2000) . Internetworking With TCP/IP. Prentice Hall: Englewood Cliffs, NJ.

Craig, H(2003). Networking Personal Computers, whit TCP/IP. O'Relly Associates, Inc. Sebastopol, CA 95472.

Huitema, C(2001). Routing in the Internet. Prentice Hall: Englewood Cliffs, NJ.

Kirch, O(2001). The Linux Network Administrators Guide.

Liu, Cricket et all. (2000).Managing Internet Information Services. O'Relly Associates, Inc. Sebastopol, CA 95472.

Stallings W. (2017). Comunicaciones y redes de computadoras. México, Prentice-Hall.

Stallings W. (2019). SNMP, SNMPv2, SNMPv3, RMON 1 y 2. México, Pearson.

RFC´s

RFCs 1155, 1157 y 1212 para SNMPv1.
RFCs 1441 a 1452 para SNMPv2, SNMPv3.
RFC 1213 para MIB y MIB-II.

RICT Revista de Investigación Científica, Tecnológica e Innovación

Edición semestral volumen 2, número 3, abril 2024

ISSN 2992-7315