



**UNIVERSIDAD ESTATAL PENÍNSULA  
DE SANTA ELENA  
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES  
INSTITUTO DE POSTGRADO**

**TITULO DEL TRABAJO DE TITULACIÓN**

Modelo Predictivo del tráfico de Internet: caso puntos digitales gratuitos Zona 5

**AUTOR**

**Ing. Ramírez Borbor, Abel Fabricio**

**TRABAJO DE TITULACIÓN**

Previo a la obtención del grado académico en  
**MAGISTER EN TECNOLOGÍAS DE LA INFORMACIÓN**

**TUTOR**

**Quirumbay Yagual, Daniel, Mgtr.**

**Santa Elena, Ecuador**

**Año 2024**



**UNIVERSIDAD ESTATAL PENÍNSULA  
DE SANTA ELENA  
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES  
INSTITUTO DE POSTGRADO**

**TRIBUNAL DE SUSTENTACIÓN**



Firmado electrónicamente por:  
**DANIEL IVAN  
QUIRUMBAY YAGUAL**

---

**Ing. ALICIA ANDRADE VERA, Mgtr.  
COORDINADORA DEL  
PROGRAMA**

---

**Lsi. DANIEL IVAN QUIRUMBAY  
YAGUAL, Mgtr.  
TUTOR**



Firmado electrónicamente por:  
**JORGE LUIS RODAS  
SILVA**

---

**Ing. JORGE LUIS RODAS SILVA,  
Ph.D  
DOCENTE  
ESPECIALISTA**

**HUGO  
GIOVANN  
Y VERA  
FLORES**

Firmante digital:HUGO  
GIOVANNY VERA FLORES  
DN: C=EC, O=SECURITY  
DATA S.A. 2, OU=ENTIDAD  
DE CERTIFICACION DE  
INFORMACION,  
SERIALNUMBER=1811211028  
11, CN=HUGO GIOVANNY  
VERA FLORES  
Fecha: 14/5/2024 10:41:45 -  
05:00

---

**Ing. HUGO GIOVANNY VERA  
FLORES, Mgtr-  
DOCENTE  
ESPECIALISTA**

---

**Abg. MARIA RIVERA, Mgtr.  
SECRETARIO GENERAL  
UPSE**



**UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA**  
**FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**  
**INSTITUTO DE POSTGRADO**

**CERTIFICACIÓN**

Certifico que luego de haber dirigido científica y técnicamente el desarrollo y estructura final del trabajo, este cumple y se ajusta a los estándares académicos, razón por el cual apruebo en todas sus partes el presente trabajo de titulación que fue realizado en su totalidad por RAMÍREZ BORBOR ABEL FABRICIO, como requerimiento para la obtención del título de Magister en Tecnologías de la Información.

**TUTOR**



Firmado electrónicamente por:  
**DANIEL IVAN**  
**QUIRUMBAY YAGUAL**

---

**Lsi. Daniel Iván Quirumbay Yagual, Mgrt.**

**18 días del mes de marzo del año 2024**



**UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA**  
**FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**  
**INSTITUTO DE POSTGRADO**  
**DECLARACIÓN DE RESPONSABILIDAD**

**Yo, RAMÍREZ BORBOR ABEL FABRICIO**

**DECLARO QUE:**

El trabajo de Titulación, “Modelo Predictivo del tráfico de Internet: caso puntos digitales gratuitos Zona 5” previo a la obtención del título en Magister en Tecnologías de la Información, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

**Santa Elena, a los 18 días del mes de marzo del año 2024**

**EL AUTOR**

---

**ABEL FABRICIO RAMIREZ BORBOR**

**IV**



UPSE

**UNIVERSIDAD ESTATAL PENÍNSULA  
DE SANTA ELENA  
FACULTAD DE CIENCIAS DE LA INGENIERÍA INSTITUTO  
DE POSTGRADO**

**CERTIFICACIÓN DE ANTIPLAGIO**

Certifico que después de revisar el documento final del trabajo de titulación denominado “Modelo Predictivo del tráfico de Internet: caso puntos digitales gratuitos Zona 5”, presentado por el estudiante, RAMÍREZ BORBOR ABEL FABRICIO fue enviada al Sistema Antiplagio Compilatio, presentando un porcentaje de similitud correspondiente 5%, por el que se aprueba el trabajo para que continúe con el proceso de titulación.

INFORME DE ANÁLISIS  
registro

MaestriaTIC\_AbelRamirez

5%  
Textos sospechosos

4- Similitudes  
1- 1% similitudes entre consultas  
2- entre las fuentes manifiestas  
3- 1% idiomas no reconocidos

Nombre del documento: MaestriaTIC_AbelRamirez.docx	Deposítante: DANIEL IVAN QUIRUMBAY YAGUAL	Número de palabras: 15.820
ID del documento: 14ca0c776707d9303818426888000a4c4ac2b4f73	Fecha de depósito: 15/3/2024	Número de caracteres: 100.873
Tamaño del documento original: 1,62 MB	Tipo de carga: Inscrito	Fecha de fin de análisis: 15/3/2024

**TUTOR**



Firmado electrónicamente por:  
**DANIEL IVAN  
QUIRUMBAY YAGUAL**

**Lsi. Daniel Iván Quirumbay Yagual, Mgrt.**



**UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA**  
**FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**  
**INSTITUTO DE POSTGRADO**

**AUTORIZACIÓN**

Yo, **RAMÍREZ BORBOR ABEL FABRICIO**

Autorizo a la Universidad Estatal Península de Santa Elena, para que haga de este trabajo de titulación o parte de él, un documento disponible para su lectura consulta y procesos de investigación, según las normas de la Institución.

Cedo los derechos en línea patrimoniales de artículo profesional de alto nivel con fines de difusión pública, además apruebo la reproducción de este artículo académico dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor

**Santa Elena, a los 18 días del mes de marzo del año 2024**

**EL AUTOR**

---

**ABEL FABRICIO RAMIREZ BORBOR**

## **AGRADECIMIENTO**

Agradezco primero a Dios, a mis Padres y mis hermanos que fueron que día a día me apoyaron en esta etapa de estudio.

Agradezco a mi amiga Lorena que me dio ánimos para continuar con el desarrollo de mi tesis.

Agradezco a mi tutor que en la última etapa me guio para desarrollar mi proyecto final de estudio.

Agradezco a la Universidad Estatal Península de Santa Elena que me dio la oportunidad de seguir superándome.

*Abel Fabricio Ramírez Borbor*

## **DEDICATORIA**

Dedico este logro a mis Padres y mis hermanos, quienes son mi ejemplo y mentores que me han inspirado a seguir adelante pese a las dificultades que se han presentado en mi camino a lo largo de mi carrera académica.

*Abel Fabricio, Ramírez Borbor*



# INDICE GENERAL

CERTIFICACIÓN.....	III
CERTIFICACIÓN DE ANTIPLAGIO .....	V
AGRADECIMIENTO .....	VII
DEDICATORIA.....	VIII
ÍNDICE GENERAL .....	IX
ÍNDICE DE TABLAS.....	XII
ÍNDICE DE FIGURAS.....	XIII
RESUMEN.....	XVI
ABSTRACT.....	XVII
INTRODUCCIÓN .....	2
CAPÍTULO 1. MARCO TEÓRICO REFERENCIAL.....	7
1.1.....	Revisión de literatura7
1.2. Desarrollo teórico y conceptual.....	12
1.2.1 Redes de comunicación.....	12
1.2.2 Tráfico de redes.....	14
1.2.3 Factores que influyen en el consumo de internet.....	15
1.2.4 Tipos de tráfico de red .....	16
1.2.5 Ingeniería del tráfico de red .....	19
1.2.6 Machine Learning .....	21
1.2.7 Técnicas para los modelos predictivos .....	22

1.2.8 Evaluación de los modelos.....	27
1.2.9 Matriz de confusión (Confusion Matrix) .....	29
1.2.10 Minería de datos.....	31
1.2.11 Lenguaje de programación Python .....	32
1.2.12 Flask.....	37
<b>CAPÍTULO 2. METODOLOGÍA.....</b>	<b>39</b>
2.1. Contexto de la investigación .....	39
2.2. Diseño y alcance de la investigación.....	39
2.3. Tipo y métodos de investigación.....	40
2.4. Población y muestra .....	40
2.5. Técnicas e instrumentos de recolección de datos.....	41
2.6. Procesamiento de la evaluación: Validez y confiabilidad de los instrumentos aplicados para el levantamiento de información. ....	42
<b>CAPÍTULO 3. RESULTADOS Y DISCUSIÓN .....</b>	<b>43</b>
3.1. Análisis de resultados.....	43
3.2. Fases para el desarrollo del modelo predictivo .....	48
3.3. Extracción de la información .....	49
3.3.1 Definición de la información a extraer .....	49
3.3.2 Definición de temporalidad de los datos.....	51
3.3.3 Datos extraídos.....	51
3.3.4 Herramientas de desarrollo .....	52
3.4. Filtrado y limpieza de los datos.....	52
3.5. Entrenamiento del modelo .....	54
3.5.1 Diseño del Dataset .....	54

3.5.2 Carga de información al DataSet .....	57
3.5.3 Definición de registros para pruebas y entrenamiento de los modelos	58
3.5.4 Modelo 1: Activaciones relu, relu y sigmoid.....	58
3.5.5 Modelo 2: Activaciones softmax, selu y tanh.....	61
3.5.6 Modelo 3: Activaciones exponencial, hard_sigmoid y mish.....	65
3.5.7 Exactitud de las redes.....	69
3.5.8 Pérdidas del sistema.....	70
3.7. Resultados .....	72
<b>CONCLUSIONES.....</b>	<b>75</b>
<b>RECOMENDACIONES .....</b>	<b>77</b>
<b>REFERENCIAS.....</b>	<b>79</b>
<b>ANEXOS.....</b>	<b>2</b>
Anexo 1: Encuesta.....	2

## ÍNDICE DE TABLAS

<b>Tabla 1:</b> <i>Tipos de ingeniería de tráfico de datos</i> .....	20
<b>Tabla 2:</b> <i>Información a extraer</i> .....	50
<b>Tabla 3:</b> <i>Total de registros extraídos</i> .....	51
<b>Tabla 4:</b> <i>Total de registros después del filtrado y limpieza de datos</i> .....	53
<b>Tabla 5:</b> <i>Estructura del Dataset</i> .....	55
<b>Tabla 6:</b> <i>Resultados de la matriz de confusión del modelo 1</i> .....	60
<b>Tabla 7:</b> <i>Resultados de la matriz de ganancias del modelo 1</i> .....	61
<b>Tabla 8:</b> <i>Resultados de la matriz de confusión del modelo 2</i> .....	64
<b>Tabla 9:</b> <i>Resultados de la matriz de ganancias del modelo 2</i> .....	64
<b>Tabla 10:</b> <i>Resultados de la matriz de confusión del modelo 3</i> .....	67
<b>Tabla 11:</b> <i>Resultados de la matriz de ganancias del modelo 3</i> .....	68
<b>Tabla 12:</b> <i>Resumen de la evaluación de los 3 modelos</i> .....	72
<b>Tabla 13:</b> <i>Resultados del modelo predictivo</i> .....	74

# ÍNDICE DE FIGURAS

<b>Figura 1:</b> <i>Redes de telecomunicaciones</i> .....	13
<b>Figura 2:</b> <i>Modelo de árbol de decisiones</i> .....	23
<b>Figura 3:</b> <i>Modelo de máquina de vectores de soporte</i> .....	24
<b>Figura 4:</b> <i>Modelo de vecinos más cercanos</i> .....	25
<b>Figura 5:</b> <i>Modelo de regresión logística</i> .....	26
<b>Figura 6:</b> <i>Matriz de confusión</i> .....	30
<b>Figura 7:</b> <i>Minería de Datos</i> .....	31
<b>Figura 8:</b> <i>Python</i> .....	32
<b>Figura 9:</b> <i>Logo de Numpy</i> .....	33
<b>Figura 10:</b> <i>Logo de Pandas</i> .....	34
<b>Figura 11:</b> <i>Logo de TensorFlow</i> .....	35
<b>Figura 11:</b> <i>Logo de Matplotlib</i> .....	36
<b>Figura 13:</b> <i>Logo de Scipy</i> .....	36
<b>Figura 9:</b> <i>Flask</i> .....	38
<b>Figura 15:</b> <i>Ubicación de CNT</i> .....	39

<b>Figura 16:</b> <i>Caracterización de la población consultada</i> .....	43
<b>Figura 17:</b> <i>Frecuencia de acceso al punto del encuentro</i> .....	44
<b>Figura 18:</b> <i>Nivel de satisfacción con la velocidad del internet</i> .....	44
<b>Figura 19:</b> <i>Se conecta al internet del punto digital mediante wifi</i> .....	45
<b>Figura 20:</b> <i>Se experimenta cortes del servicio de internet</i> .....	46
<b>Figura 21:</b> <i>Se experimenta cortes del servicio de internet</i> .....	46
<b>Figura 22:</b> <i>Nivel de recomendación del servicio</i> .....	47
<b>Figura 23:</b> <i>Fases del modelo predictivo</i> .....	48
<b>Figura 24:</b> <i>Arquitectura del modelo 1</i> .....	59
<b>Figura 25:</b> <i>Entrenamiento del modelo 1</i> .....	60
<b>Figura 26:</b> <i>Arquitectura del modelo 2</i> .....	62
<b>Figura 27:</b> <i>Entrenamiento del modelo 2</i> .....	63
<b>Figura 28:</b> <i>Arquitectura del modelo 3</i> .....	66
<b>Figura 29:</b> <i>Entrenamiento del modelo 3</i> .....	66
<b>Figura 30:</b> <i>Exactitud de las redes</i> .....	69
<b>Figura 31:</b> <i>Pérdidas del sistema</i> .....	71

**Figura 32:** *Sitio web desarrollado* ..... 73

**Figura 33:** *Funcionalidad para la predicción de ancho de banda* ..... 73

## RESUMEN

La investigación sobre el “Modelo Predictivo del tráfico de Internet: caso puntos digitales gratuitos Zona 5” se centra en el diseño y evaluación de modelos de redes neuronales para proponer un modelo predictivo utilizando técnicas de inteligencia artificial. A lo largo del estudio, se analizan diversas arquitecturas de redes neuronales en entornos dinámicos y cambiantes, explorando diferentes combinaciones de funciones de activación, con el propósito de determinar la configuración óptima que maximicen los resultados. Lo obtenido durante la investigación tienen importantes implicaciones para la optimización de recursos y la mejora de la calidad del servicio en los puntos digitales gratuitos.

El modelo predictivo con mejores resultados predictivos cuenta con una arquitectura de activaciones exponencial, hard\_sigmoid y mish convirtiéndose en una herramienta efectiva para anticipar y gestionar el tráfico de Internet, lo que puede ayudar a planificar de mejor forma el ancho de banda en cada uno de los puntos digitales.

**Palabras claves:** Redes Neuronales, Modelos Predictivos, Tráfico de Internet.



## **ABSTRACT**

The research on the “Predictive Model of Internet traffic: case of free digital points Zone 5” focuses on the design and evaluation of neural network models to propose a predictive model using artificial intelligence techniques. Throughout the study, various neural network architectures are analyzed in dynamic and changing environments, exploring different combinations of activation functions, with the purpose of determining the optimal configuration that maximizes results. What was obtained during the research has important implications for the optimization of resources and the improvement of the quality of service in free digital points.

The predictive model with the best predictive results has an architecture of exponential, hard\_sigmoid and mish activations, becoming an effective tool to anticipate and manage Internet traffic, which can help to better plan the bandwidth in each of the digital points.

**Keywords:** Neural Networks, Predictive Models, Internet Traffic.

# INTRODUCCIÓN

En la era moderna, Internet ha transformado la interacción de las personas con el mundo, ofreciendo un extenso y variado espacio donde se pueden inmiscuir en una amplia gama de actividades, partiendo desde la comunicación y el entretenimiento hasta la educación y el comercio (Calcaneo, 2021). El acceso a Internet a nivel global se encuentra cursando un crecimiento significativo a partir de la pandemia; para el año 2023, se estima que el 64,4% de la población alrededor del mundo se encuentran conectados al Internet, lo que se traduce en un incremento del 1,9% en comparación con el año anterior (BNEW, 2023). Este aumento se atribuye a diversos factores, como la reducción en el costo de la tarifa mensual del servicio, el incremento de usuarios con teléfonos inteligentes, la ampliación de la cobertura de redes móviles y el incremento en la alfabetización digital.

Para el caso de América Latina el acceso a internet se ha incrementado significativamente en los últimos años, para el año 2023, se estima que el 71,9% de la población se encuentra conectada a internet, representando un aumento del 2,2% en comparación al año 2022. Sin embargo, aún existen importantes diferencias en el acceso a internet entre países de la región, en los países más desarrollados, como Chile y Uruguay, el acceso a internet es casi universal, mientras que, en los países menos desarrollados, como Haití y Bolivia, el acceso a internet es mucho más limitado (CEPAL, 2023).

En Ecuador, el acceso a Internet alcanza al 76% de la población, mostrando notables disparidades entre las zonas urbanas y rurales. Además, entre los años 2019 y 2022, se observó un incremento del 32,7% en el número de hogares con conexión a la red (INEC, 2023). Esta creciente

dependencia de Internet ha resultado en una generación masiva de datos digitales, lo que abre oportunidades significativas para comprender y analizar el tráfico en línea (Tejada et al., 2023). Dado que el acceso a Internet es de gran importancia para el desarrollo económico y social, permitiendo el acceso a información y educación, la comunicación, la incorporación en la economía digital y las transacciones comerciales, resulta esencial estudiar su dinámica y patrones.

El Ministerio de Telecomunicaciones y de la Sociedad de la Información (MINTEL) desde su creación ha enfocado su accionar en gestionar acciones para la reducción de la brecha digital; a través, de la implementación de los Puntos Digitales Gratuitos (PDG) a nivel nacional (MINTEL, 2023). Estos PDG permiten que todos los ciudadanos puedan acercarse y acceder a Internet; por lo cual, se han establecido en áreas estratégicas para brindar servicios digitales a comunidades que podrían no tener fácil acceso a la tecnología y la conectividad. En los puntos digitales gratuitos implementados por el gobierno nacional en la Zona 5 de planificación que comprende los cantones de Guayas.

### **Planteamiento de la investigación**

El problema investigativo aborda el desarrollo de un modelo capaz de predecir con precisión el tráfico de Internet, con el fin de permitir la toma de decisiones que permitan mejorar el uso de los recursos de la red (Herrera y otros, 2022). Esta tarea es crucial debido al crecimiento exponencial del tráfico de Internet, lo que demanda una gestión eficiente para evitar congestiones y garantizar la disponibilidad del servicio a los usuarios. Por lo tanto, el modelo predictivo debe considerar una extensa variedad de factores, como el tipo de tráfico, el horario, la ubicación geográfica y otros datos relevantes, para realizar pronósticos precisos. Esto tendría un impacto

positivo en la calidad del servicio de Internet y contribuiría a la disminución de los costos operativos (Norambuena y otros, 2022).

Las organizaciones, desde proveedores de servicios de Internet hasta empresas y universidades, necesitan comprender cómo se utiliza su red para gestionarla de manera efectiva, considerando que el análisis del tráfico proporciona información sobre el uso de la red, los cuellos de botella y los posibles puntos de congestión (Spitsina y otros, 2022); en este sentido, desarrollar un Modelo Predictivo del tráfico de Internet para su optimización es crucial debido a la creciente importancia de un servicio de Internet eficiente para la actual sociedad de la información; ya que, el aumento constante de usuarios, dispositivos y aplicaciones conectados genera una demanda cada vez mayor de ancho de banda y recursos de red (Martínez y otros, 2018). Un modelo predictivo preciso permitiría anticipar patrones de tráfico, picos de demanda y tendencias de uso, lo que a su vez facilitaría una asignación más eficiente de recursos, una planificación de capacidad más precisa y una mejor respuesta a situaciones de tráfico inusual. Esto no solo mejoraría la experiencia del usuario al garantizar una conectividad fluida, sino que también reduciría costos operativos y energéticos al evitar la subutilización o sobrecarga de recursos. Para lo cual, el método con mejores características para cumplir este objetivo es el aprendizaje automático supervisado.

### **Formulación del problema de investigación**

Ante la problemática descrita se formulan los siguientes lineamientos de investigación:

- El machine learning está en la capacidad de aprender patrones complejos de datos del tráfico de internet para realizar predicciones precisas que permitan optimizar los recursos de red de los puntos digitales gratuitos Zona 5

- La definición de los métodos y estrategias más adecuadas para el etiquetado de los distintos conjuntos de datos aportan en mejorar la precisión del modelo de machine learning.
- A través del uso de técnicas de aprendizaje automático se podrá analizar y clasificar de manera precisa las tendencias del tráfico de internet en tiempo real
- Mediante la aplicación de minería de datos se podrá identificar patrones y tendencias predominantes en el uso de Internet en los puntos digitales gratuitos de la Zona 5.
- La matriz de confusión permitirá evaluar el rendimiento del modelo de aprendizaje automático, identificando los factores a mejorar.

**Objetivo General:**

Implementar un modelo predictivo del tráfico de Internet para los puntos digitales gratuitos en la zona 5 que permita evitar congestiones y garantizar el servicio a los usuarios.

**Objetivos Específicos:**

1. Evaluar los modelos de machine learning que permita el análisis del tráfico de Internet.
2. Preparar un conjunto de datos etiquetados adecuadamente que representen información sobre uso del internet en la red de los Puntos digitales Gratuitos Zona 5.
3. Desarrollar un modelo de aprendizaje automático que permita identificar la orientación del uso del internet en los puntos digitales gratuitos de la zona 5 a través del análisis de los resultados obtenidos.

### **Planteamiento hipotético**

El uso de un modelo predictivo sobre el tráfico de internet permite mejorar el servicio de los Puntos Digitales Gratuitos de la Zona 5 y mejorar la experiencia de sus usuarios.

# CAPÍTULO 1. MARCO TEÓRICO REFERENCIAL

## 1.1. Revisión de literatura

López et al., (2019) en su documento “Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red” centran su investigación en la aplicación de modelos predictivos capaces de estimar el comportamiento de ciertos indicadores relacionados con la respuesta de la infraestructura de red. Los autores abordan el problema de la complejidad y la dinámica en la gestión de redes de telecomunicaciones, proponiendo un enfoque fundamentado en técnicas de aprendizaje automático para predecir el comportamiento de variables críticas. Utilizan datos históricos de una infraestructura de red real para entrenar y validar el modelo predictivo, evaluando su rendimiento mediante métricas de precisión y eficacia. Los resultados que los investigadores obtuvieron demuestran la viabilidad y la utilidad del modelo propuesto en la predicción precisa del comportamiento de variables clave, lo que podría tener importantes implicaciones en la gestión y optimización de infraestructuras de red.

Aguirre (2022), en su investigación “Modelo algorítmico para alta disponibilidad en transporte de volúmenes crecientes de tráfico variable en redes ópticas” presenta una propuesta para abordar el transporte de grandes volúmenes de tráficos de manera eficiente y confiable en redes ópticas elásticas (EON). Este modelo se basa en técnicas de aprendizaje automático para mejorar diferentes aspectos lo que permitió que la implementación del modelo EON con aprendizaje automático se lleva a cabo como resultado de la investigación, y su desempeño se evalúa mediante simulaciones que consideran diversos escenarios de tráfico. Los resultados

obtenidos le permitieron al autor demostrar que el modelo propuesto ofrece un transporte eficiente y adaptable a diferentes condiciones, lo que permite mejorar el uso de los recursos de la red y mejorar su eficiencia en el transporte de datos.

Hernández et al., (2021) en su investigación “Redes neuronales y predicción de tráfico” desarrollan un modelo de tráfico fundamentado en redes neuronales, utilizando la información disponible de los dispositivos (routers y switches) sobre el tráfico de una red de datos. Los resultados resaltan la flexibilidad y precisión de las redes neuronales en el modelado del tráfico de redes Ethernet, especialmente cuando se dispone de suficientes datos de entrenamiento. La utilización de sistemas inteligentes, como las redes neuronales artificiales (ANN), se muestra como una herramienta indispensable cuando otros modelos matemáticos no logran representar con exactitud el comportamiento de un sistema. Si bien las redes neuronales son útiles para identificar sistemas lineales y no lineales, es preponderante contar con tecnología de alto rendimiento para acelerar el proceso de entrenamiento con mayor precisión. Este enfoque subraya la importancia del uso de las redes neuronales para en la gestión del tráfico y otros campos donde se requiera modelado y predicción de datos.

Osorio et al. (2023) en su investigación “Algoritmos supervisados para la predicción del ancho de banda de las aplicaciones en Amazon Web Service desde una pyme rural” proponen una metodología destinada a medir y anticipar el comportamiento del ancho de banda en pequeñas y medianas empresas situadas en la ruralidad caracterizado por ser zonas de difícil acceso en Colombia. El propósito de la investigación es de optimizar la calidad del servicio en las aplicaciones web; para lo cual, se compara entre tres algoritmos de redes neuronales para configurar una red neuronal multicapa, eligiendo el modelo que presente una mayor precisión. Este



algoritmo se entrena con datos para prever el tráfico de red hacia la nube. Dada la importancia de asegurar el funcionamiento adecuado de las aplicaciones web, especialmente aquellas con alto tráfico y operación en tiempo real, el análisis del comportamiento de la red resulta esencial. La planificación de la capacidad de la red adquiere una relevancia crítica, sobre todo en empresas con recursos limitados, como aquellas ubicadas en áreas rurales.

Chávez (2021) en su investigación denominada “Predicción del tráfico de una red inalámbrica basada en redes neuronales artificiales mediante el algoritmo de Levenberg-Marquardt” tiene como objetivo identificar la mejor ruta de tráfico en una red virtual utilizando análisis de datos mediante el desarrollo de modelos de aprendizaje automático para predecir el tráfico. Al igual que en las redes físicas, el diseño eficiente de una red de comunicación es crucial en las redes virtuales. Sin embargo, optimizar una red de datos es un desafío complejo debido a la multitud de factores que deben considerarse para satisfacer la creciente demanda del servicio desde los usuarios y stakeholders. El análisis del tráfico de red es fundamental para definir la arquitectura tecnológica y ha sido objeto de gran interés en años recientes.

Hernández et al. (2018) en su investigación “Modelamiento y pronósticos de tráfico correlacionado” posiciona como factores de suma importancia la necesidad de implementar una planificación y gestión efectiva en las redes de comunicación al interior de las organizaciones, considerando que la capacidad de predecir los niveles de tráfico de datos futuros es un factor crucial en estas labores de planificación y gestión de redes. Por lo tanto, la investigación se enfoca en introducir al lector en modelos de inteligencia artificial basados en series de tiempo, los cuales permiten pronosticar escenarios futuros de tráfico en las redes al aprovechar la relación existente

entre los valores de tráfico previos en una red de datos. Este enfoque de modelado en series de tiempo se fundamenta en observaciones recopiladas de manera periódica sobre eventos en la red.

Gildardo (2019) en su investigación “Predicción de Tráfico en Redes de Telecomunicaciones basado en Técnicas de Inteligencia Analítica”, el autor presenta una metodología que se fundamenta en técnicas de inteligencia artificial con el propósito de identificar el impacto de los altos consumos de internet sobre los enlaces que brindan el servicio. Asimismo, se introduce un modelo de redes neuronales en minería de datos diseñado para prever el grado de uso de los enlaces en una red informática. La información relacionada con el estado de la red se adquiere mediante un simulador especializado en redes de telecomunicaciones, luego se somete a un riguroso análisis mediante la aplicación de técnicas analíticas adaptadas a las particularidades del estudio. Los resultados obtenidos en los experimentos subrayan las ventajas de la metodología propuesta al identificar con precisión las variables más relevantes, así como la efectividad del modelo predictivo al anticipar la carga de tráfico en un entorno de red de datos.

Barba et al (2018) en su investigación “Aplicación de Perceptrón Multicapa en la predicción de tráfico para redes académicas empresariales” se presenta una metodología basada en técnicas de inteligencia analítica con el fin de identificar y prever las variables de red que influyen en la utilización de los enlaces. Además, se introduce un modelo de redes neuronales en minería de datos diseñado para anticipar el nivel de utilización de los enlaces en una red de datos. La información sobre el estado de la red se adquiere mediante un simulador de redes de telecomunicaciones y se analiza utilizando técnicas analíticas adaptadas. Los resultados experimentales evidencian las ventajas de la metodología propuesta al identificar las variables más

relevantes y la precisión del modelo predictivo al anticipar la carga de tráfico en un entorno de red de datos.

Barrezueta (2018) en su investigación “Desarrollo de un algoritmo de redes neuronales artificiales aplicado a la predicción de tráfico de la infraestructura de comunicaciones de redes corporativas” se desarrolló un algoritmo con el propósito de prever el tráfico de red en entornos corporativos mediante el uso de redes neuronales artificiales. Para recopilar los datos de tráfico generados por la red de la Facultad de Informática y Electrónica, se estableció un servidor de recolección de datos utilizando TCPdump y un script que facilitaba la captura automática de datos durante períodos de monitoreo de 4 semanas. Los datos recolectados incluyeron información sobre protocolos específicos seleccionados debido a su relevancia en cada paquete, lo que permitió llevar a cabo un análisis del tamaño de cada uno de los protocolos capturados. Posteriormente, los datos se almacenaron en una base de datos creada en Excel y se exportaron a Matlab, donde se implementó una Red Neuronal Artificial mediante un algoritmo de retropropagación (backpropagation) para realizar la predicción de datos.

Abdelhadi y Guy (2018) en su artículo científico “Traffic prediction based on LSTM recurrent neural networks in SDN” presentan un enfoque innovador para predecir el tráfico de redes definidas por software (SDN) utilizando redes neuronales recurrentes LSTM (Long Short-Term Memory). Los autores proponen un modelo de predicción que aprovecha la capacidad de las LSTM para capturar las dependencias temporales en los datos de tráfico de red. Utilizan conjuntos de datos reales para el entrenamiento y evaluación del modelo, calculando los principales indicadores que den cuenta de su nivel de rendimiento en términos de precisión y eficiencia. Los resultados experimentales muestran que el modelo propuesto logra una precisión significativa en

la predicción del tráfico de red en entornos SDN, lo que demuestra su utilidad potencial en la gestión y optimización de redes. En conclusión, el artículo destaca la promesa de las redes neuronales recurrentes, especialmente LSTM, como herramientas efectivas para abordar problemas de predicción de tráfico en redes SDN.

Por lo antes descrito, en este contexto dinámico de la sociedad digital actual, el crecimiento exponencial del tráfico de Internet representa un desafío crucial para la gestión eficiente de las redes de comunicación, con el objetivo de anticipar patrones, tendencias y fluctuaciones en la utilización de la red; razón por la cual, mediante un riguroso análisis teórico, se explorarán conceptos fundamentales de redes de comunicación, teoría de colas, modelos estadísticos y técnicas de aprendizaje automático, con el fin de establecer un marco conceptual sólido que respalde el desarrollo y la efectividad de un modelo predictivo destinado a enfrentar los desafíos actuales y futuros del tráfico de Internet.

## **1.2. Desarrollo teórico y conceptual**

### **1.2.1 Redes de comunicación**

Las redes de comunicación son sistemas complejos que facilitan la transmisión eficiente y el intercambio de información entre varios dispositivos, permitiendo la conectividad y la comunicación en tiempo real a través de diversos medios. Estas redes pueden abarcar desde entornos locales, como redes de área local (LAN), hasta vastas infraestructuras globales, como Internet. En su esencia, las redes de comunicación se basan en la interconexión de nodos mediante protocolos y tecnologías específicas, tales como el protocolo TCP/IP, que establece la base para la comunicación en la red mundial. Estas redes pueden ser cableadas o inalámbricas, y su diseño y gestión involucran aspectos cruciales como la topología, la seguridad, el rendimiento y la manera

para satisfacer las demandas de transmisión de datos en constante evolución (Fernández et al., 2020).

**Figura 1:** *Redes de telecomunicaciones*



**Fuente:** (Audacia Comunicación, 2019).

El contante avance y evolución de las redes de comunicación ha sido fundamental en la transformación de la sociedad moderna, permitiendo una variedad de servicios y aplicaciones, desde la comunicación interpersonal hasta el acceso a recursos digitales y la colaboración global. La creciente complejidad de estas redes, junto con el surgimiento de nuevas tecnologías como la 5G y la Internet de las cosas (IoT), continúa definiendo la forma en que las personas, empresas y dispositivos se conectan y comparten información en la era digital (Medina et al., 2020).

### **1.2.2 Tráfico de redes**

El tráfico de redes se refiere al flujo de datos que circula a través de una red de computadoras, este flujo de datos puede incluir una variedad de tipos de información, como archivos, correos electrónicos, mensajes instantáneos, solicitudes de páginas web, transmisiones de video y voz, entre otros, este tráfico se produce cuando los dispositivos conectados entre sí, como computadoras, servidores, enrutadores y dispositivos móviles, intercambian datos entre sí, este intercambio de datos puede ocurrir tanto dentro de una red local (LAN) como en redes más amplias, como la Internet. El tráfico puede ser de naturaleza bidireccional, donde los datos se envían y reciben entre dispositivos, o unidireccional, donde los datos se transfieren en una sola dirección, la cantidad y el tipo de tráfico en una red pueden variar ampliamente dependiendo de factores como la cantidad de dispositivos conectados, la actividad de los usuarios, el tipo de aplicaciones utilizadas y la capacidad de la infraestructura de red (Borrero y Quintero, 2018).

La gestión efectiva del tráfico de redes es esencial para asegurar un rendimiento óptimo e incrementar la experiencia de usuario, esto implica la supervisión y el monitoreo del flujo de datos para garantizar la disponibilidad de todos los recursos de la red, como el ancho de banda y la capacidad de procesamiento, se utilicen de manera eficiente y se prioricen según las necesidades. Además, la seguridad del tráfico de redes es una preocupación importante, ya que se deben implementar medidas para proteger los datos en tránsito contra amenazas como el acceso no autorizado, el robo de información y los ataques cibernéticos (Ríos y Fermin, 2019).

### 1.2.3 Factores que influyen en el consumo de internet

Abarcan una serie de elementos clave que tienen estrecha relación con la manera en que los usuarios interactúan y utilizan los servicios en línea, por lo que comprenderlos y gestionarlos eficazmente es esencial para diseñar y mantener infraestructuras de red robustas y eficientes para satisfacer los requerimientos de los usuarios en el cambiante entorno digital actual.

- **Ancho de banda:** Se refiere a la disponibilidad de ancho de banda disponible en una red, que tiene directa relación sobre la velocidad y la capacidad de transferencia de datos. En los momentos en los que el ancho de banda es mayor, la red cuenta con más capacidad para manejar un alto consumo de datos.
- **Tecnología de red:** La tecnología utilizada en la infraestructura de red, como el tipo de conexión (cable, fibra óptica, inalámbrica), la generación de redes móviles (4G, 5G), y la topología de red (estrella, malla, bus), influye en la velocidad y confianza de la conexión a Internet.
- **Dispositivos y aplicaciones:** El tipo de dispositivos utilizados por los usuarios, como teléfonos inteligentes, computadoras portátiles, tabletas, y las aplicaciones y servicios que utilizan, como streaming de video, redes sociales, juegos en línea, afectan el volumen y el tipo de datos consumidos.
- **Patrones de uso:** Los hábitos y comportamientos de los usuarios, como el tiempo de conexión, las actividades en línea y los picos de uso durante ciertos períodos del día, influyen en la demanda de datos y en la carga de la red.

- **Ubicación geográfica:** La ubicación física de los usuarios y de los servidores de contenido afecta la latencia, la velocidad y la disponibilidad de la conexión a Internet, especialmente en áreas con infraestructura de red limitada o en regiones remotas.
- **Calidad de servicio:** La calidad del servicio proporcionada por los proveedores tiene estrecha relación con la estabilidad de la conexión, la consistencia de la velocidad y la atención al cliente influye en la satisfacción del usuario y en su propensión a consumir más datos.
- **Políticas de gestión de tráfico:** Las políticas implementadas por los proveedores y los operadores de red, como la priorización de ciertos tipos de tráfico o la limitación de la velocidad de conexión, pueden afectar el consumo de datos y la experiencia del usuario.

#### **1.2.4 Tipos de tráfico de red**

El tráfico típicamente proviene de tres fuentes principales: los dispositivos conectados a la red, los usuarios que acceden a ella y los servicios que la red ofrece a estos usuarios. Los dispositivos de red están programados para mantener una sincronización continua, gestionar la configuración y mantener interacciones con otros dispositivos, lo que implica, entre otras cosas, la emisión de alarmas y la gestión de la comunicación dentro de la red.

##### **1.2.3.1 Tráfico de red por DHCP**

El tráfico de red por DHCP consiste en las comunicaciones entre dispositivos y servidores DHCP dentro de una red, este tráfico se encarga de asignar direcciones IP automáticamente a los equipos autorizados que se conectan a la red, así como de proporcionarles información de configuración, como la dirección del servidor DNS y la puerta de enlace predeterminada. Este tipo



de tráfico implica mensajes intercambiados entre clientes DHCP y servidores DHCP, donde los clientes envían solicitudes para obtener una dirección IP y otros detalles de configuración, mientras que los servidores responden ofreciendo direcciones disponibles y proporcionando la información necesaria. La gestión eficiente del tráfico de red por DHCP es fundamental para garantizar una asignación rápida y confiable de direcciones IP, así como para mantener la conectividad y la estabilidad de la red en su conjunto (Castillo y Forero, 2020).

### **1.2.3.2 Tráfico de red por autenticación**

El tráfico de red por autenticación se refiere a las comunicaciones que tienen lugar durante el proceso de comprobación de la identidad de los usuarios o dispositivos que intentan acceder a la red. Este proceso implica la transferencia de datos entre el cliente y el servidor de autenticación, donde el cliente envía credenciales de inicio de sesión al servidor para su verificación. El servidor, a su vez, valida estas credenciales y determina si el usuario o dispositivo tiene permiso para acceder a la red. Si las credenciales son válidas, el servidor envía una respuesta de autenticación positiva al cliente, permitiéndole el acceso a la red. Por el contrario, si las credenciales no son válidas, el servidor envía una respuesta de autenticación negativa y deniega el acceso. Durante este proceso, se pueden intercambiar diferentes tipos de datos, como nombres de usuario, contraseñas, certificados digitales u otros tokens de autenticación, dependiendo de los métodos de autenticación utilizados en la red. La gestión eficiente del tráfico de red por autenticación es fundamental para incrementar la seguridad y la integridad de la red, protegiéndola contra accesos no autorizados y posibles amenazas cibernéticas (Mejía et al., 2020).

### **1.2.3.3 Tráfico de red por inicio de sesión**

Se refiere a las comunicaciones que tienen lugar cuando un dispositivo o usuario intenta acceder a un sistema o servicio en línea que requiere autenticación. Durante este proceso, el dispositivo o usuario envía credenciales de inicio de sesión al servidor correspondiente, que pueden incluir el nombre de usuario y contraseña u otros mecanismos de identificación. El servidor, a su vez, verifica estas credenciales para determinar si son válidas y si el acceso debe ser permitido. Si las credenciales son correctas, el servidor responde con una confirmación de inicio de sesión exitoso y otorga acceso al sistema o servicio solicitado. En caso contrario, el servidor emite un mensaje de error indicando que las credenciales proporcionadas no son válidas. Este intercambio de datos durante el inicio de sesión es crucial para garantizar la seguridad de los sistemas informáticos y servicios en línea, ya que ayuda a protegerlos contra accesos no autorizados y posibles ataques cibernéticos. La gestión eficiente del tráfico de red que se genera durante el inicio de sesión de los distintos dispositivos conectados a la red es esencial para garantizar de seguridad del usuario y sin problemas en entornos en línea (Torres et al., 2023).

### **1.2.3.4 Tráfico de red por navegación en internet**

El tráfico de red por navegación en Internet es el flujo de datos que se produce cuando los dispositivos conectados a una red acceden a diferentes páginas web y recursos en línea a través de un navegador web. Este proceso comienza cuando un usuario ingresa una dirección web en el navegador o hace clic en un enlace, lo que genera una solicitud de datos desde el dispositivo hacia el servidor que aloja el sitio web deseado. Esta solicitud incluye información como la dirección URL y otros detalles necesarios para recuperar el contenido solicitado, como cookies y encabezados HTTP.

Una vez que el servidor recibe la solicitud, responde enviando los datos solicitados de vuelta al dispositivo del usuario, lo que permite que el navegador web renderice y muestre el contenido en la pantalla del usuario, estos datos pueden incluir texto, imágenes, videos, archivos multimedia y otros tipos de contenido web. Durante el tiempo que el usuario interactúa con el sitio web, como hacer clic en enlaces o llenar formularios, se generan nuevas solicitudes y respuestas de datos, lo que resulta en un flujo continuo de tráfico de red.

El tráfico de navegación en internet puede variar ampliamente en volumen y tipo dependiendo de los sitios web visitados, las actividades en línea realizadas y el número de equipos conectados a la red. La gestión eficiente de este tráfico es crucial para garantizar una experiencia de usuario fluida, optimizar el rendimiento de la red y proteger la privacidad de los datos durante la navegación en línea.

### **1.2.5 Ingeniería del tráfico de red**

La ingeniería de redes, como campo especializado dentro de las Tecnologías de la Información y la Comunicación (TIC), se enfoca en el diseño, implementación y mejora de redes informáticas para asegurar un flujo óptimo y confiable del tráfico de datos. Este ámbito comprende una amplia gama de aspectos, desde la capacidad de la red y el enrutamiento de datos, hasta la gestión del ancho de banda, la calidad de servicio (QoS), la seguridad de la red y la optimización del rendimiento (Donoso et al., 2021).

En la ingeniería del tráfico de red, los ingenieros de redes utilizan diversas herramientas y técnicas para diseñar y gestionar redes de datos que puedan satisfacer los requerimientos de rendimiento de los sistemas informáticos y servicios que se ejecutan en ellas. Esto puede incluir la implementación de políticas de enrutamiento para dirigir el tráfico de manera eficiente, la

configuración de mecanismos de control de congestión para impedir que la red se sature, y la implementación de tecnologías de optimización de ancho de banda, como el empaquetamiento de los datos y el almacenamiento en caché de contenido (Torres et al., 2019).

La gestión del tráfico de red es esencial para garantizar el funcionamiento efectivo de las infraestructuras de comunicación, como redes empresariales, proveedores de servicios de Internet (ISP) y centros de datos. Al optimizar el flujo de datos e incrementar el rendimiento de la red, los ingenieros de redes pueden contribuir significativamente a la productividad, la reducción de costos operativos e incrementar la experiencia del usuario. Actualmente, existen diversas técnicas de gestión del tráfico, cada una dirigida a aspectos específicos del funcionamiento de la red y la mejora de su rendimiento (Donoso et al., 2021). Algunos de los tipos más comunes incluyen:

**Tabla 1:** *Tipos de ingeniería de tráfico de datos*

Tipo	Descripción
Ingeniería del Tráfico de Ancho de Banda	El enfoque se dirige hacia la gestión y mejora del ancho de banda disponible en la red, con el propósito de asegurar una distribución justa y eficaz de los recursos de red entre distintos usuarios y aplicaciones.
Ingeniería del Tráfico de Enrutamiento	Está relacionada con la configuración y gestión de los protocolos de enrutamiento en la red para dirigir el tráfico de manera eficiente y evitar congestiones y cuellos de botella.

Ingeniería del Tráfico de Calidad de Servicio (QoS)	Se enfoca en garantizar que ciertos tipos de tráfico, como voz o video, reciban prioridad y ancho de banda garantizado para conservar la calidad de la experiencia del usuario.
Ingeniería del Tráfico de Seguridad	Está orientada a implementar medidas de seguridad para proteger el tráfico de red contra amenazas como intrusiones, ataques de denegación de servicio (DDoS) y malware.
Ingeniería del Tráfico de Optimización de Rendimiento	Busca mejorar la eficiencia y el rendimiento general de la red mediante la implementación de técnicas como la compresión de datos, la aceleración de la entrega de contenido y la optimización de la latencia.

---

**Fuente:** *Elaboración propia.*

### **1.2.6 Machine Learning**

También conocido como aprendizaje automático, constituye una rama de la inteligencia artificial orientada en el desarrollo de algoritmos y modelos que capacitan a las computadoras para aprender y perfeccionarse de manera autónoma, utilizando datos y experiencias previas como referencia. En lugar de programar explícitamente a una computadora para que realice una tarea específica, en el aprendizaje automático se le proporcionan datos y se le permite aprender patrones

y relaciones dentro de esos datos para hacer predicciones o tomar decisiones en el futuro (Forero y Negre, 2023).

El proceso de aprendizaje automático generalmente implica varios pasos, que incluyen la recopilación y preparación de datos, la selección y entrenamiento de modelos, la evaluación del rendimiento del modelo y, finalmente, la implementación y ajuste del modelo en un entorno de producción. Los algoritmos de aprendizaje automático pueden clasificarse en diferentes categorías según el tipo de aprendizaje que realizan, como aprendizaje no supervisado, supervisado, semi supervisado y por refuerzo (Urrea y Ramos, 2023).

El aprendizaje automático tiene un amplio abanico de aplicaciones para las diferentes industrias, incluyendo la medicina, las finanzas, la tecnología, el comercio electrónico y más. Se utiliza para realizar tareas relacionadas con la identificación de patrones, la predicción de resultados, la optimización de procesos, la detección de anomalías y la toma de decisiones automatizadas. El crecimiento explosivo en la disponibilidad de datos y la capacidad computacional ha impulsado el rápido avance y la adopción generalizada del aprendizaje automático en los últimos años (Forero y Negre, 2023).

### **1.2.7 Técnicas para los modelos predictivos**

Existen diversas técnicas que se utilizan para descubrir patrones, tendencias y conocimientos ocultos en conjuntos de datos. Estas técnicas se aplican según el tipo de problema y el objetivo de análisis.

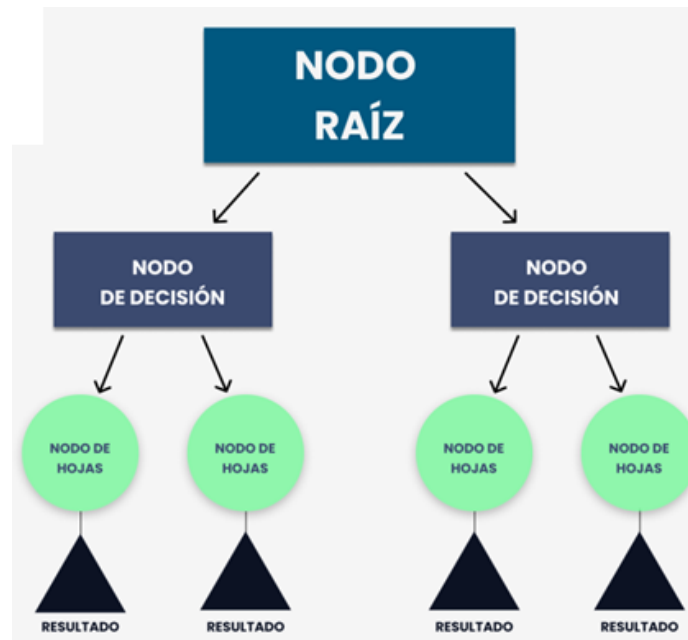
### **1.2.7.1 Naïve Bayes**

El algoritmo se fundamenta en el teorema de Bayes entrenando el modelo con datos previamente etiquetados, se aprenden las probabilidades condicionales de cada clase y características, lo que permite clasificar nuevas instancias en función de la probabilidad más alta. Naïve Bayes se utiliza comúnmente en problemas de clasificación de textos debido a su simplicidad, eficiencia y buen rendimiento incluso con conjuntos de datos grandes (Rico et al., 2019).

### **1.2.7.2 Árbol de decisiones**

Son modelos clasificadores que se representan mediante una estructura de árbol que se construyen mediante la división recursiva del conjunto de datos en función de las propiedades más notables, buscando maximizar la pureza de las clases en cada subdivisión. A medida que se desciende por el árbol, las instancias se asignan a una clase final en las hojas. Los árboles de decisión son fáciles de interpretar y visualizar, y pueden manejar tanto características numéricas como categóricas (Atamari et al., 2022).

**Figura 2:** *Modelo de árbol de decisiones*



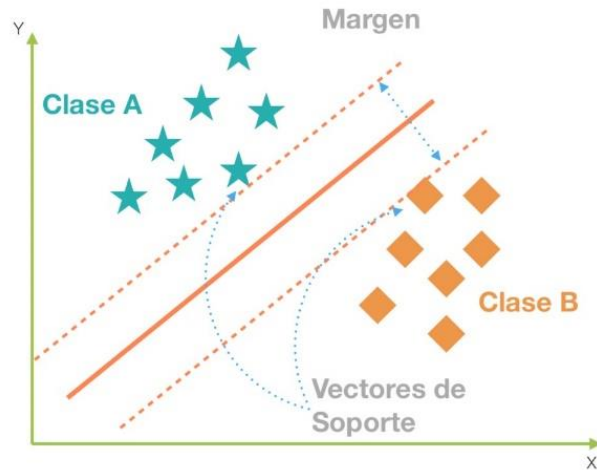
Fuente: Rico y otros (2019).

### 1.2.7.3 Máquinas de vectores de soporte (SVM)

Son modelos de aprendizaje automático cuyo objetivo es maximizar el margen entre las instancias de las diferentes clases y el hiperplano de separación. Las SVM son efectivas en la clasificación de conjuntos de datos complejos y no lineales mediante el uso de funciones de kernel. Este modelo es reconocido por su capacidad para generalizar bien y manejar problemas con conjuntos de datos de tamaño moderado a grande (Aparicio et al., 2023).

**Figura 3:** *Modelo de máquina de vectores de soporte*





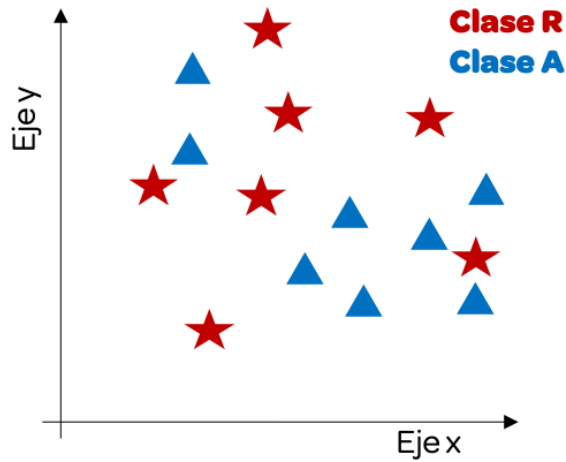
**Fuente:** Aparicio y otros (2023).

#### 1.2.7.4 Vecinos más cercanos (KNN)

Es un algoritmo utilizado para la clasificación y reconocimiento de patrones, se fundamenta en clasificar las instancias en función de las clases de sus vecinos más cercanos en el espacio de características. El valor de "K" representa el número de vecinos más cercanos considerados para la clasificación. El KNN busca instancias similares en términos de distancia y asigna la clase más común entre los vecinos seleccionados.

Este enfoque es especialmente útil en conjuntos de datos donde la separación entre clases es clara. El KNN es fácil de entender e implementar; razón por la cual, es usado en aplicaciones desde reconocimiento de imágenes hasta sistemas de recomendación. Sin embargo, es importante elegir adecuadamente el valor de "K" y tener en cuenta el impacto de las características y la métrica de distancia utilizada en el rendimiento del modelo. (Contreas et al., 2023).

**Figura 4:** *Modelo de vecinos más cercanos*

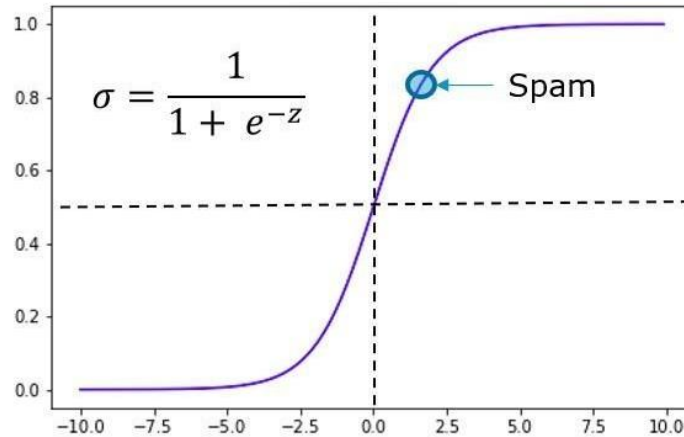


**Fuente:** Contreras y otros (2023).

### 1.2.7.5 Modelo de regresión logística

Permite predecir la probabilidad de ser parte de una clase en lugar de predecir un valor numérico, este modelo se utiliza una función logística para transformar una combinación lineal de las características en una probabilidad, que representa la posibilidad de pertenecer a la clase positiva. Mediante el ajuste de los coeficientes el modelo busca encontrar la mejor línea de separación entre las dos clases en el espacio de características. Para clasificar nuevas instancias, se utiliza un umbral de probabilidad para decidir a qué clase pertenecen. La regresión logística es ampliamente utilizada debido a su simplicidad, interpretabilidad y eficiencia computacional (Mazuera et al., 2021).

**Figura 5:** *Modelo de regresión logística*



**Fuente:** Contreras y otros (2023).

## 1.2.8 Evaluación de los modelos

Existen diversas métricas que permiten medir el desempeño, la calidad y el rendimiento de los clasificadores, su importancia radica en su capacidad de evaluar y comparar el desempeño relacionado con la capacidad de clasificación, la exactitud y la capacidad de detectar correctamente instancias positivas y negativas. Es importante considerar estas métricas en conjunto para obtener una evaluación completa del rendimiento del clasificador en diferentes aspectos.

### 1.2.8.1 Precisión (Accuracy)

La precisión, también conocida como accuracy, es una métrica fundamental que permite medir la efectividad de los modelos de clasificación, cuantificando la proporción de predicciones correctas en relación con el total de instancias evaluadas. Se calcula dividiendo el total de predicciones correctas (verdaderos positivos más verdaderos negativos) entre el universo de instancias evaluadas. La precisión proporciona una visión global del rendimiento del modelo, evaluando su capacidad para clasificar de manera acertada tanto las instancias positivas como las negativas. Si bien es una métrica útil, es importante considerar su interpretación en contextos

donde la distribución de clases es desigual, ya que la precisión podría no reflejar de manera adecuada el rendimiento en situaciones con clases desbalanceadas (Aparicio et al., 2023).

### **1.2.8.2 Sensibilidad (Recall) y Especificidad (Specificity)**

La sensibilidad, también conocida como recall, y la especificidad son medidas fundamentales en la evaluación de modelos de clasificación, la sensibilidad cuantifica la capacidad del modelo para identificar correctamente las instancias positivas, calculando la cantidad de verdaderos positivos con respecto al total de positivos reales. Por otro lado, la especificidad mide la habilidad del modelo para descartar adecuadamente las instancias negativas, evaluando la proporción de verdaderos negativos con respecto al total de negativos reales. Ambas métricas ofrecen perspectivas cruciales sobre el rendimiento del clasificador: la sensibilidad se enfoca en la capacidad de capturar positivos, mientras que la especificidad se centra en evitar falsos positivos. En conjunto, proporcionan una valoración completa de la capacidad discriminativa del modelo en situaciones de clasificación binaria, permitiendo a los evaluadores sopesar adecuadamente los errores de tipo I y tipo II según el contexto de la aplicación. (Rosso et al., 2022).

### **1.2.8.3 Valor F1 (F1 Score)**

Se refiere a un indicador de evaluación comúnmente utilizado en trabajos de clasificación, que combina la precisión y la exhaustividad en una sola medida, su cálculo es conocido como la media armónica de precisión y exhaustividad, el F1 Score proporciona una evaluación equilibrada del rendimiento del modelo al considerar tanto los falsos positivos como los falsos negativos. Su utilidad radica en situaciones donde ambas clases, positiva y negativa, son de igual importancia, y se busca un equilibrio entre la capacidad del modelo para identificar instancias positivas y evitar falsos positivos. Un F1 Score más alto indica un mejor rendimiento del clasificador, siendo

especialmente útil en escenarios donde la asimetría en la distribución de clases podría afectar negativamente la interpretación de otras métricas de evaluación. (Atamari et al., 2022).

#### **1.2.8.4 Curva ROC (Receiver Operating Characteristic)**

Es una herramienta gráfica ampliamente empleada en la evaluación de modelos de clasificación, sobre todo para modelos de inteligencia artificial y de aprendizaje automático. Esta curva representa la relación entre la sensibilidad (tasa de verdaderos positivos) y la tasa de falsos positivos (1 - especificidad) a medida que se ajusta el umbral de decisión del modelo. Proporciona una visualización intuitiva del rendimiento del modelo en distintas configuraciones de umbral, permitiendo a los analistas y desarrolladores evaluar y comparar la capacidad discriminativa de un clasificador en diferentes condiciones. Un valor del área bajo la Curva ROC (AUC-ROC) más cercano a 1 indica un rendimiento superior del modelo, mientras que un valor cercano a 0.5 sugiere un rendimiento similar al azar. En resumen, la Curva ROC es una herramienta invaluable para comprender y mejorar la capacidad predictiva de los modelos de clasificación en una variedad de aplicaciones (Contreas et al., 2023).

#### **1.2.9 Matriz de confusión (Confusion Matrix)**

Es un instrumento esencial que permite evaluar los modelos de clasificación, proporcionando resultados detallados de su rendimiento al comparar las predicciones con los resultados reales. Organizada en filas y columnas que representan las clases predichas y las clases verdaderas, respectivamente, la matriz de confusión desglosa la clasificación de instancias en verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos. Esta matriz de 2 x 2 permite calcular métricas clave como la precisión, la sensibilidad, la especificidad y el F1 Score, proporcionando una visión comprehensiva de la capacidad del modelo para clasificar

correctamente instancias positivas y negativas. Su interpretación es esencial para entender el origen y el impacto de los errores cometidos por el modelo, lo que facilita ajustes y mejoras en el proceso de entrenamiento y evaluación. (Brito y Brito, 2018).

**Figura 6:** *Matriz de confusión*

		Actual Values	
		Yes	No
Predicted Values	Yes	True Positive	False Positive
	No	False Negative	True Negative

**Positivo (P):** La observación es positiva (por ejemplo, es un perro).

**Negativo (N):** La observación no es positiva (por ejemplo, no es un perro).

**Verdadero Positivo (TP):** Resultado en el que el modelo predice correctamente la clase positiva.

**Verdadero Negativo (TN):** Resultado donde el modelo predice correctamente la clase negativa.

**Falso Positivo (FP):** También llamado error de tipo 1, resultado donde el modelo predice incorrectamente la clase positiva cuando en realidad es negativa.

**Falso Negativo (FN):** También llamado error de tipo 2, un resultado en el que el modelo predice incorrectamente la clase negativa cuando en realidad es positiva.

**Fuente:** DataSource (2023)

### 1.2.10 Minería de datos

La minería de datos juega un papel crucial en el desarrollo y perfeccionamiento de modelos predictivos al brindar las herramientas y técnicas necesarias para analizar y explorar amplios conjuntos de datos en busca de patrones y relaciones ocultas. La minería de datos permite identificar insights valiosos que son fundamentales para la creación de modelos predictivos precisos y eficaces. Para ello, se requiere una cuidadosa selección y preparación de los datos pertinentes, así como la identificación de características clave, junto con la elección adecuada de los algoritmos de aprendizaje automático para abordar el problema específico en cuestión. Además, la minería de datos facilita la evaluación y validación de los modelos predictivos, lo que posibilita ajustes y refinamientos continuos para mejorar su rendimiento y capacidad de generalización (Orellana y Cedillo, 2020).

**Figura 7: Minería de Datos**



**Fuente:** Elaboración propia.

Además, la minería de datos desempeña un papel fundamental en la aplicación práctica de modelos predictivos, ya que ayuda a interpretar y aprovechar los resultados obtenidos para tomar decisiones informadas. Los modelos predictivos desarrollados a través de la minería de datos pueden ser utilizados en una vasta variedad de aplicaciones y sectores, desde la predicción de ventas y la gestión de inventarios hasta la detección de fraudes y la personalización de servicios. Al comprender los patrones y tendencias presentes en los datos, las organizaciones pueden anticipar eventos futuros, identificar oportunidades y mitigar riesgos de manera más efectiva. En última instancia, la combinación de minería de datos y modelos predictivos permite aprovechar al máximo el vasto conjunto de información disponible en el mundo actual para incrementar la toma de decisiones y lograr el cumplimiento de los objetivos comerciales y estratégicos (Martínez y Palencia, 2021).

### **1.2.11 Lenguaje de programación Python**

Python es ampliamente requerido para la implementación de aplicaciones de IA debido a su facilidad de uso, flexibilidad y gran variedad de bibliotecas especializadas, desde sus orígenes se ha posicionado como un lenguaje de alto nivel, interpretado y de propósito general que ofrece una sintaxis clara y amigable, lo que lo convierte en la primera opción entre los desarrolladores independientemente de su nivel de conocimientos.

**Figura 8:** *Python*





**Fuente:** Python (2023).

Python es especialmente conocido por su ecosistema de bibliotecas específicas para la IA y el aprendizaje automático, estas bibliotecas proporcionan herramientas poderosas y eficientes que permiten el procesamiento de datos, el desarrollo y entrenamiento de modelos, y la evaluación de resultados, lo que facilita la implementación de algoritmos de IA y la experimentación con diferentes enfoques. Además, la naturaleza de código abierto de Python fomenta la colaboración y la innovación en la comunidad de IA, lo que ha contribuido a su prominencia en este campo (Python, 2023). A continuación, se describen las librerías que formarán parte del desarrollo del presente modelo predictivo.

- **Numpy:** proporciona una gran cantidad de herramientas y estructuras de datos para la manipulación eficiente de arreglos multidimensionales, así como operaciones matemáticas y estadísticas avanzadas. Esta biblioteca es ampliamente utilizada por su capacidad para realizar cálculos numéricos rápidos y eficientes, lo que la convierte en la base de muchas otras bibliotecas de análisis de datos en como pandas, scikit-learn y TensorFlow. NumPy también ofrece funcionalidades para la generación de números aleatorios, la transformada de Fourier, el álgebra lineal y la optimización, lo que la hace indispensable para una variedad de aplicaciones en ciencia de datos, ingeniería, investigación académica y más.

**Figura 9:** *Logo de Numpy*



**Fuente:** Numpy (2023).

- **Pandas:** suministra estructuras de datos flexibles y eficientes, especialmente la serie y el marco de datos, que permiten la organización y manipulación de datos tabulares de una manera intuitiva y potente. Esta biblioteca es ampliamente utilizada por su capacidad para cargar datos de diversas fuentes, limpiarlos, transformarlos y realizar análisis estadísticos y exploratorios de manera efectiva. Pandas también ofrece funciones para la gestión de fechas y horas, agrupación de datos, fusiones y concatenaciones de conjuntos de datos, entre otras operaciones útiles.

**Figura 10:** *Logo de Pandas*



**Fuente:** Pandas (2023).

- **Os:** brinda una interfaz fácil de usar para acceder a funcionalidades del sistema operativo, como la manipulación de archivos y directorios, la navegación por rutas de archivos, la obtención de información sobre el entorno del sistema y la ejecución de comandos del sistema desde el propio código Python. Con os, los científicos de datos pueden automatizar tareas relacionadas con la gestión de archivos, organizar datos en el sistema de archivos, acceder a recursos externos y garantizar la portabilidad de sus scripts en diferentes plataformas y sistemas operativos.

- **Tensorflow:** permite el desarrollo de modelos complejos de inteligencia artificial, permitiendo a los científicos de datos implementar un sinnúmero de algoritmos de aprendizaje automático y realizar tareas como clasificación, regresión, segmentación y generación de datos. Con su arquitectura escalable y su soporte para la ejecución distribuida en múltiples dispositivos y plataformas, TensorFlow facilita la experimentación con modelos complejos y la optimización de su rendimiento. Además, TensorFlow ofrece herramientas integradas para la visualización de datos, la depuración de modelos y la producción de código optimizado para su implementación en producción, lo que la transforma en una opción versátil y poderosa para proyectos de análisis de datos y aprendizaje automático a gran escala. En resumen, TensorFlow es una biblioteca fundamental que permite a los desarrolladores construir, entrenar y desplegar modelos de aprendizaje automático con eficiencia y escalabilidad en entornos de producción.

**Figura 11:** *Logo de TensorFlow*



**Fuente:** Tensorflow (2023).

- **Matplotlib.pyplot:** proporciona una interfaz amigable que permite crear gráficos de alta calidad y visualmente atractivos, que son fundamentales para explorar y comunicar patrones y tendencias en los datos. Con esta biblioteca, los científicos de datos pueden

generar una amplia variedad de gráficos, incluyendo gráficos de líneas, histogramas, diagramas de dispersión, diagramas de barras y mucho más, con una sintaxis simple y flexible. Además, Matplotlib.pyplot ofrece opciones de personalización avanzadas para ajustar el estilo, color y formato de los gráficos, así como herramientas para agregar etiquetas, títulos y leyendas, lo que permite crear visualizaciones informativas y profesionales.

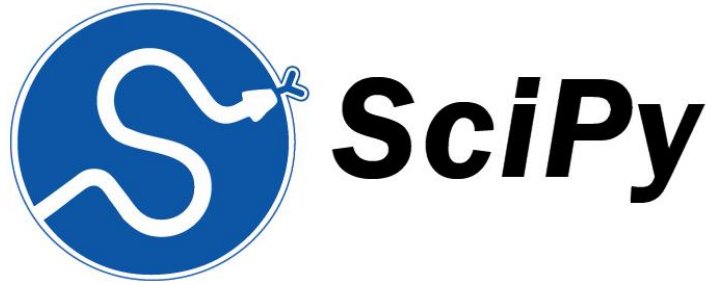
**Figura 12:** *Logo de Matplotlib*



**Fuente:** *Matplotlib (2023).*

- **Scipy:** proporciona una amplia gama de funcionalidades y algoritmos de alto nivel diseñados para abordar problemas comunes en diversas disciplinas científicas, como la optimización, la interpolación, el álgebra lineal, la estadística, el procesamiento de señales y la integración numérica, entre otros. Esta biblioteca se basa en la eficiente manipulación de matrices y arrays multidimensionales de NumPy, extendiendo su funcionalidad con herramientas adicionales y más especializadas. Con SciPy, los científicos de datos pueden acceder a una gran cantidad de algoritmos robustos y probados que les permiten resolver una variedad de problemas complejos de manera eficiente y efectiva.

**Figura 13:** *Logo de Scipy*



**Fuente:** *Scipy* (2023).

- **Tensorflow.keras:** ofrece una interfaz de alto nivel y fácil de usar que simplifica significativamente el proceso de construcción, entrenamiento y despliegue de redes neuronales. Esta biblioteca proporciona una amplia variedad de capas predefinidas, funciones de activación, optimizadores y métricas que permiten a los desarrolladores crear modelos complejos con solo unas pocas líneas de código. Además, TensorFlow.keras integra perfectamente con TensorFlow, lo que proporciona acceso a las capacidades de computación de GPU y TPU para acelerar el entrenamiento de modelos en grandes conjuntos de datos. Con TensorFlow.keras, los científicos de datos pueden experimentar con diferentes arquitecturas de red y optimizar el rendimiento de sus modelos con facilidad, lo que la convierte en una opción necesaria para proyectos de análisis de datos y aprendizaje automático a gran escala.

### **1.2.12 Flask**

Flask es un microframework de Python diseñado para construir aplicaciones web de forma rápida y sencilla, se destaca por su simplicidad y minimalismo, lo que ha permitido que sea una opción popular para proyectos pequeños y medianos donde se requiere flexibilidad y agilidad en el desarrollo. Este framework proporciona un conjunto básico de herramientas y funcionalidades

para crear aplicaciones web, pero también es altamente extensible, lo que permite a los desarrolladores agregar funcionalidades adicionales según sea necesario.

**Figura 14:** *Flask*



**Fuente:** Flask (2023).

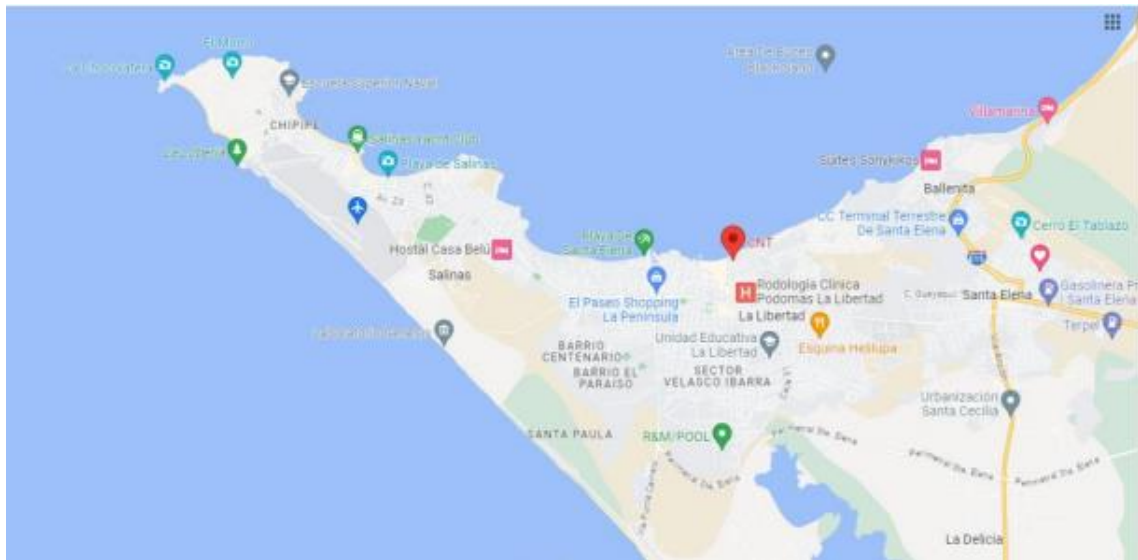
Flask sigue la arquitectura Modelo-Vista-Controlador (MVC), aunque de manera más flexible que otros frameworks. Permite a los desarrolladores organizar su código de manera modular y escalarlo fácilmente a medida que el proyecto crece. Además, Flask es conocido por su gran comunidad de usuarios y su amplia documentación, lo que facilita a los desarrolladores encontrar ayuda y recursos adicionales cuando lo necesitan.

## CAPÍTULO 2. METODOLOGÍA

### 2.1. Contexto de la investigación

La ubicación se realizará en ciudad de la Libertad Provincia de Santa Elena, lugar desde donde se administra el servicio de internet de los puntos digitales gratuitos de la zona 5.

**Figura 15:** *Ubicación de CNT*



**Fuente:** (Google Maps, 2023).

### 2.2. Diseño y alcance de la investigación

Se aplicará el diseño experimental, considerando que este tipo de diseño puede establecer grupos de control y grupos de tratamiento para evaluar el impacto de diferentes técnicas o algoritmos (Rodríguez et al., 2021) de machine learning para el modelo predictivo. En cuanto al alcance será del tipo relacional, ya que se centra en identificar y cuantificar relaciones de correlación entre diferentes variables relacionadas con el tráfico de Internet y sus patrones de uso. En este enfoque, el objetivo principal es comprender cómo las variables se comportan juntas y

cómo los cambios en una variable pueden estar relacionados con cambios en otras variables, sin necesariamente establecer una relación causal directa.

### **2.3. Tipo y métodos de investigación**

La investigación cuantitativa se caracteriza por mostrar un enfoque metodológico centrado en la recolección y análisis de datos numéricos y estadísticos que permiten describir y explicar fenómenos, relaciones y patrones. En el contexto del desarrollo de un Modelo Predictivo del tráfico de Internet para su optimización, se llevaría a cabo utilizando este tipo de investigación para recopilar y analizar datos numéricos relacionados con el tráfico de Internet y factores influyentes.

El método de investigación analítico se enfoca en descomponer un fenómeno complejo en partes más manejables para comprender mejor sus componentes individuales y sus relaciones. En el contexto del desarrollo de un Modelo Predictivo del tráfico de Internet para su optimización, el enfoque analítico implicaría desglosar el problema en sus elementos constituyentes y analizarlos en detalle para llegar a conclusiones más sólidas.

### **2.4. Población y muestra**

En cuanto a la información relacionada con el modelo del aprendizaje automático, la población estará dada por el total de tráfico generado por los beneficiarios de los puntos digitales gratuitos de la Zona 5 entre los meses de julio a diciembre de 2023, no se aplicará una muestra pues en el modelo de aprendizaje se requiere el análisis de toda la información.

Para recabar información relacionada con la percepción de la calidad del servicio de internet brindado en los puntos digitales gratuitos, la población se estimará considerando el número



de personas que entre el 04 al 08 de diciembre se acercaron a alguno de los centros de la Zona 5, el cálculo de la muestra se lo realizará aplicando la siguiente fórmula:

$$TM = \frac{Z_{\alpha}^2 * N * p * q}{e^2 * (N - 1) + (Z_{\alpha}^2 * p * q)}$$

Donde:

- TM: Tamaño de la muestra
- $Z_{\alpha}$  (nivel de confianza del 95%): 1,96
- N (tamaño del universo): 2.548
- p (proporción de individuos de la población con las características de estudio): 0,5
- q (proporción de individuos de la población sin las características de estudio): 0,5

$$TM = \frac{1,96^2 * 2.548 * 0,5 * 0,5}{5\%^2 * (2.548 - 1) + (1,96^2 * 0,5 * 0,5)}$$

$$TM = \frac{2447,10}{7,33}$$

$$TM = 333,94$$

$$TM = 334 \text{ ENCUESTAS}$$

## 2.5. Técnicas e instrumentos de recolección de datos

La consolidación mensual del tráfico de internet se la realizará mediante la consolidación mensual de la información generada en los sistemas de información de CNT y MINTEL, mediante archivos de Excel que serán procesado y cargados al Dataset diseñado para el efecto.

Para la recolección de información sobre la percepción de la calidad del servicio se internet se aplicará una encuesta estructurada con preguntas cerradas, mismas que ofrecen varios beneficios significativos durante la recopilación y análisis de datos, al proporcionar respuestas específicas

facilitando la recopilación y el análisis de datos de manera eficiente y precisa al limitar las opciones de respuesta.

Finalizada la recolección de la información utilizando alguna de los softwares existentes para el efecto todas las respuestas serán procesadas, tabuladas y analizadas. Durante la tabulación de datos se utilizarán los análisis cuantitativo y porcentual:

- **Análisis Cuantitativo:** a través de este tipo de análisis se cuantificará en valores absolutos (numéricos) la frecuencia de respuestas por cada una de las preguntas planteadas.
- **Análisis Porcentual:** posterior al análisis cuantitativo se realizará un análisis porcentual de la frecuencia de respuestas por cada una de las preguntas planteadas.

Los resultados obtenidos permitirán contar con información del tiempo que toma el proceso manual para su posterior comparación con el proceso automatizado.

## **2.6. Procesamiento de la evaluación: Validez y confiabilidad de los instrumentos aplicados para el levantamiento de información.**

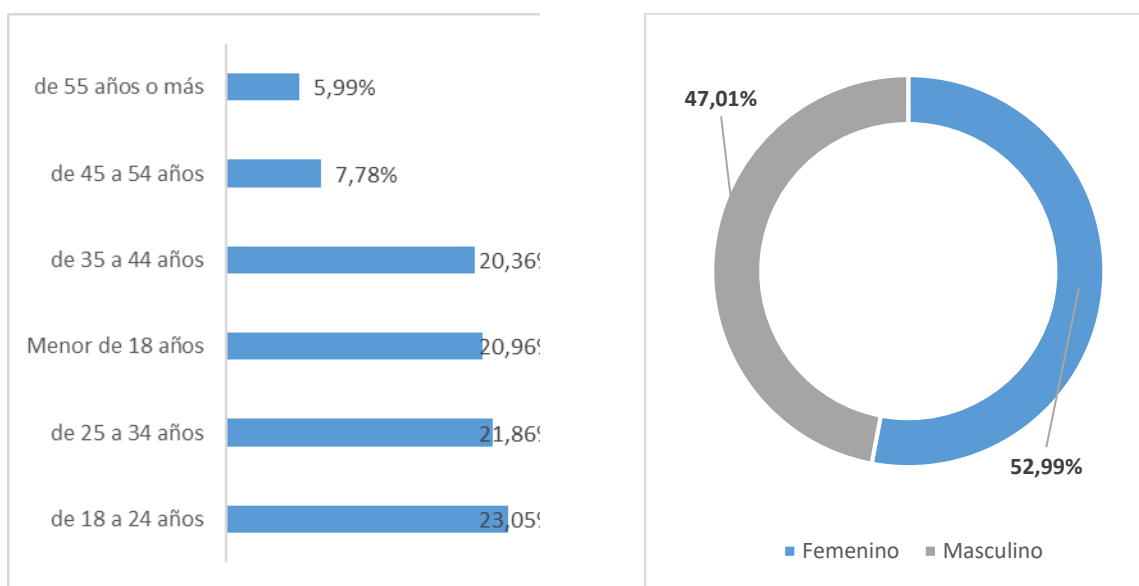
Considerando que el modelo predictivo se alimenta de información que proviene de los sistemas de información de CNT y MINTEL no aplica la evaluación de validez y confiabilidad de los instrumentos. Para el caso de la encuesta, durante su diseño se garantizó que la misma tenga validez de criterio y de contenidos, y para el caso de las respuestas al definir que sean cerradas se garantiza la validez de la respuesta en los diferentes contextos.

## CAPÍTULO 3. RESULTADOS Y DISCUSIÓN

### 3.1. Análisis de resultados

Para la recopilación de la información se elaboró un Google Forms con un acceso directo en cada computador a través del cual los usuarios manera anónima y voluntaria llenaron la información solicitada. Acorde a lo descrito en la Figura 16, el 86,23% de las personas que acuden a los puntos digitales gratuitos son desde adolescentes hasta personas de 44 años; así también, se determinó que el 52,99% de los asistentes son de sexo masculino.

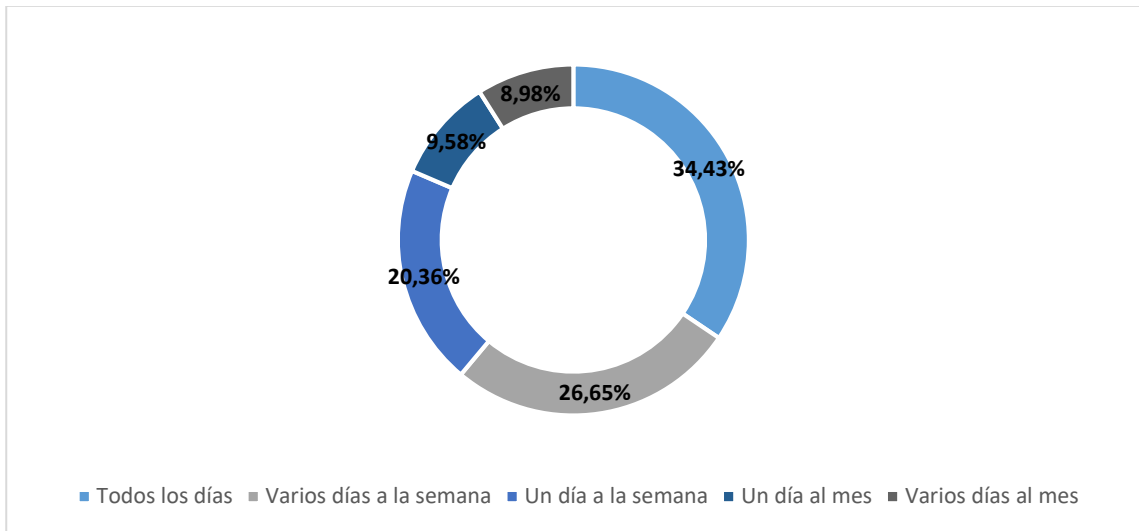
**Figura 16:** *Caracterización de la población consultada*



**Fuente:** Elaboración propia.

Sobre la frecuencia en la que las personas acuden al centro, la Figura 17 demuestra que el 81,44% de las personas acude al menos una vez a la semana, lo que da cuenta de la efectividad del proyecto que permite que las personas que se encuentran excluidas del acceso a la internet puedan acceder de manera gratuita para realizar sus actividades digitales.

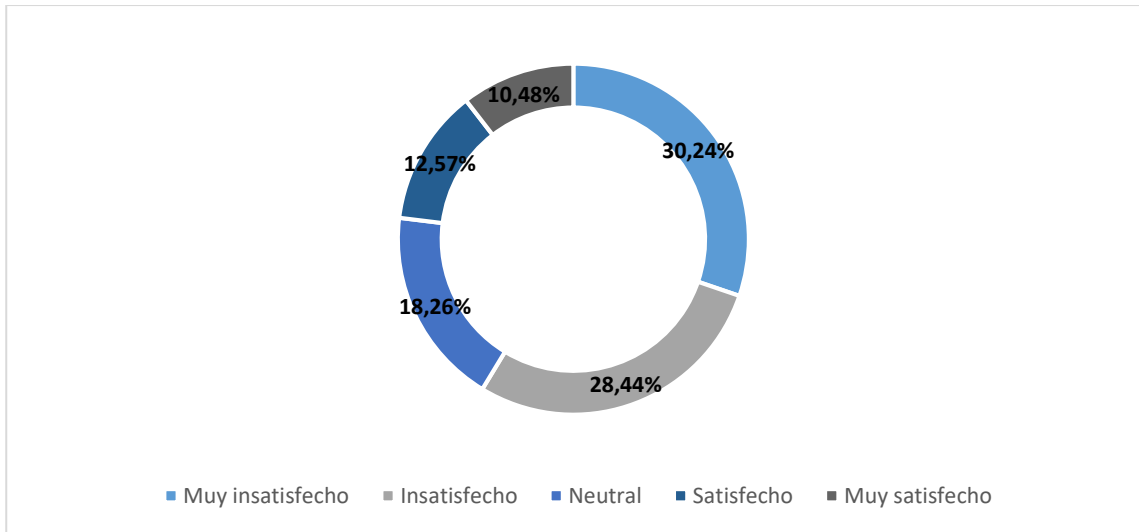
**Figura 17:** *Frecuencia de acceso al punto del encuentro*



**Fuente:** Elaboración propia.

Sobre la satisfacción de la velocidad del internet, la Figura 18 demuestra que el 58,68% de los usuarios del servicio están entre insatisfechos o muy insatisfechos, generándose una alerta que debe ser atendida por los proveedores del servicio para mejorar la calidad del servicio.

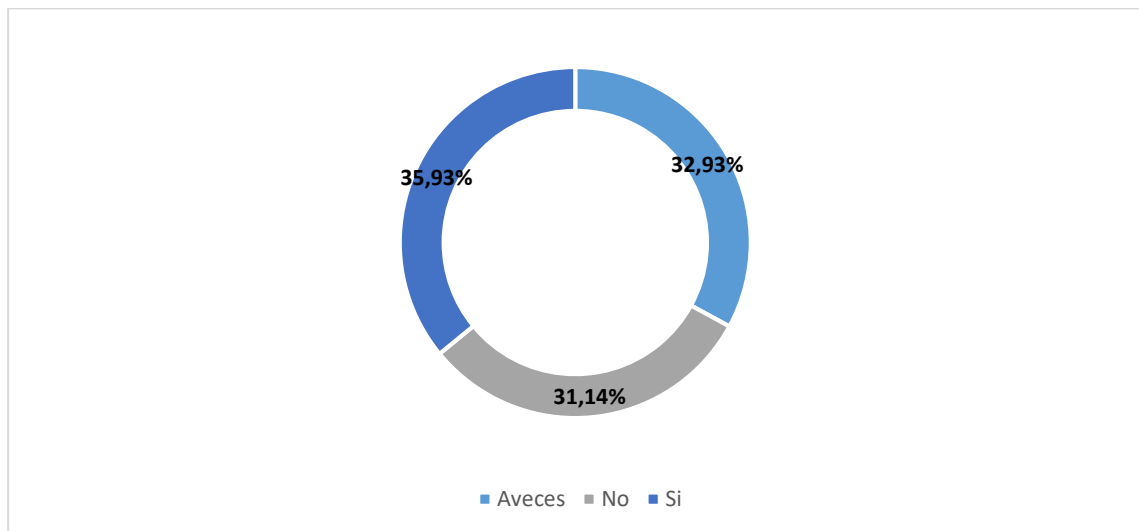
**Figura 18:** *Nivel de satisfacción con la velocidad del internet*



**Fuente:** Elaboración propia.

Sobre el uso del internet mediante la red de wifi del punto digital, la Figura 19 demuestra que hay una distribución bastante homogénea en las respuestas donde un 35,93% de la población si la utiliza, lo que da cuenta la importancia de garantizar todos los medios disponibles para que la población acceda a los servicios.

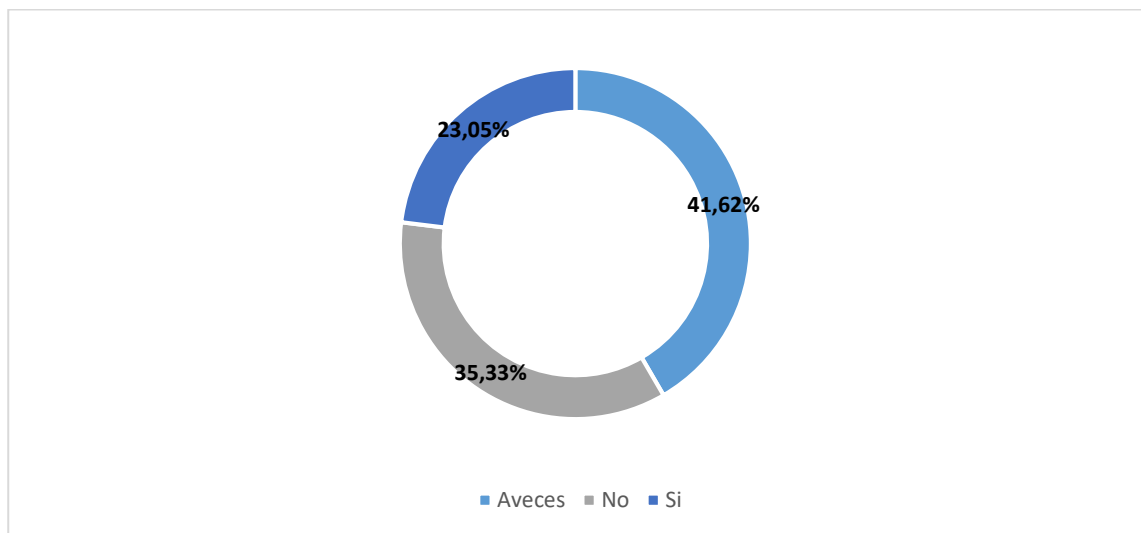
**Figura 19:** *Se conecta al internet del punto digital mediante wifi*



**Fuente:** Elaboración propia.

Así también, la Figura 20 demuestra que el 58,38% de los usuarios afirman que durante el tiempo que se encuentran utilizando el servicio han experimentado cortes del servicio, lo que no les permite concluir de manera exitosa sus tareas extendiendo por más tiempo la estancia del usuario en el centro.

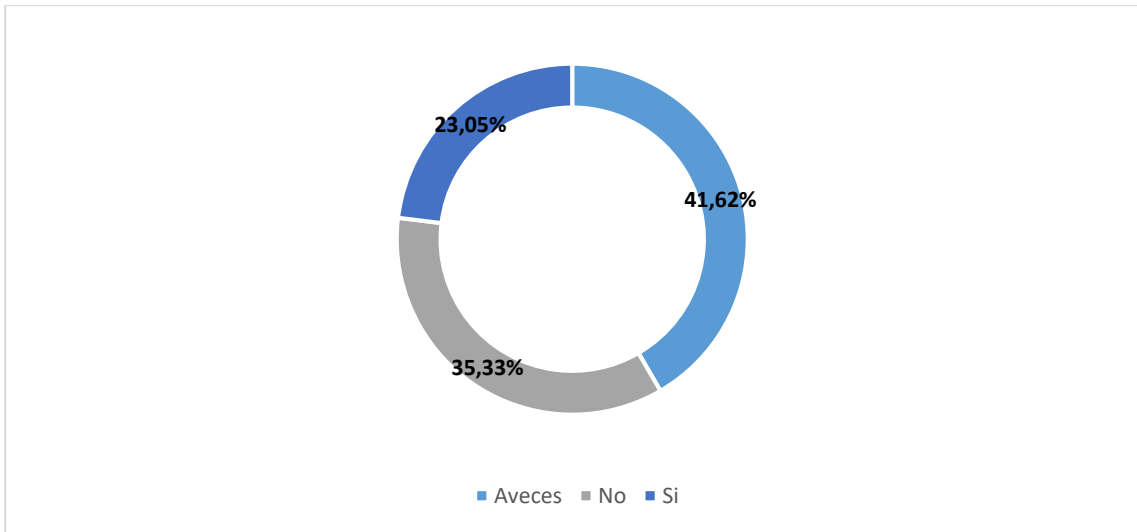
**Figura 20:** *Se experimenta cortes del servicio de internet*



**Fuente:** Elaboración propia.

Corroborando lo indicado en el análisis anterior, la Figura 21 da cuenta que el 35,33% de los usuarios no consideran que la calidad del servicio de internet les permite realizar sus actividades, mientras que un 41,62% indica que a veces si es suficiente y otras veces no lo es.

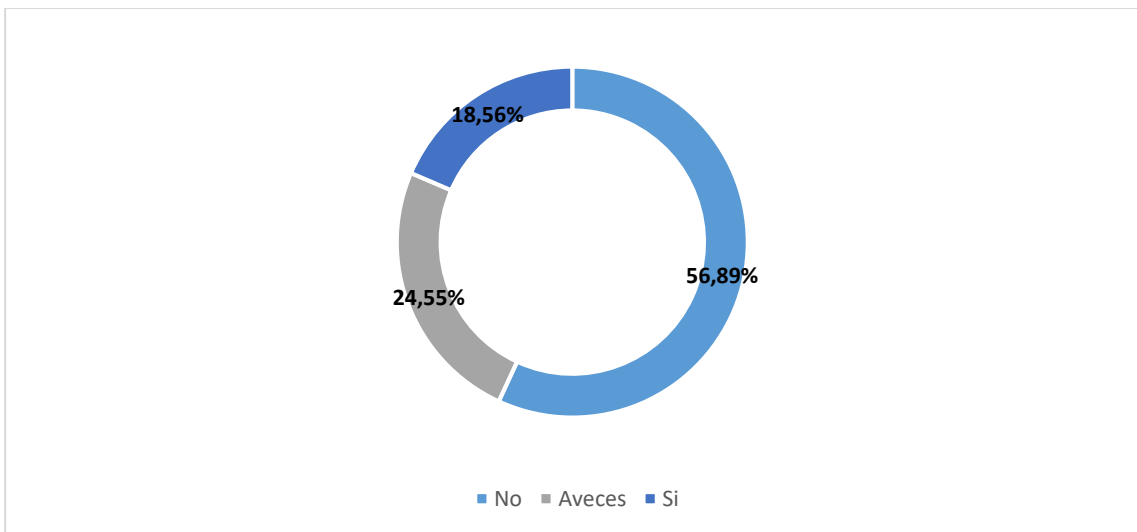
**Figura 21:** *Se experimenta cortes del servicio de internet*



**Fuente:** Elaboración propia.

Finalmente, como resultado de los antes descrito, la Figura 22 muestra que apenas el 18,56% de los usuarios recomendaría el servicio de los puntos digitales gratuitos considerando fundamentalmente la calidad del servicio de internet.

**Figura 22:** Nivel de recomendación del servicio



**Fuente:** Elaboración propia.

### 3.2. Fases para el desarrollo del modelo predictivo

Para el diseño e implementación del modelo predictivo del tráfico de internet de los puntos digitales gratuitos Zona 5, acorde a lo descrito en la Figura 23 se identifican 4 fases a través de las cuáles se garantiza el óptimo funcionamiento.

**Figura 23:** *Fases del modelo predictivo*



**Fuente:** Elaboración propia.



### **3.3. Extracción de la información**

Para los proyectos de machine learning el proceso de extracción de información es de suma importancia, en primer lugar, la calidad y cantidad de datos disponibles son factores críticos que influyen directamente en el rendimiento y la precisión de los modelos de IA. En este sentido, la extracción de información se la realiza directamente desde los diferentes sistemas de información institucionales; para lo cual, se exporta la información en formato cvs que posteriormente permitirá realizar todos los procesos necesarios para el diseño e implementación del modelo predictivo.

#### **3.3.1 Definición de la información a extraer**

Es crucial comprender la importancia de definir con precisión la información a extraer para construir un dataset, esta definición precisa sienta las bases para el éxito del desarrollo, ya que garantiza que los datos recopilados sean relevantes, completos y representativos. Al tener una comprensión clara de los datos necesarios, se optimizará la eficiencia en la recopilación y preparación de datos, lo que posteriormente ahorrará tiempo y recursos; así también, esta definición permite garantizar la calidad de los datos, lo que es esencial para obtener resultados precisos y confiables en los resultados obtenidos mediante los modelos de inteligencia artificial.

En este sentido se han identificado a dos instituciones como proveedoras de la información necesaria: 1) Corporación Nacional de Telecomunicaciones (CNT): proveerá la información relacionada con el tráfico de internet; y, 2) Ministerio de Telecomunicaciones (MINTEL): proveerá información relacionada con capacitaciones, visitas, equipos y velocidad de los puntos digitales gratuitos

En la Tabla 2 se describe la información que se identificó como necesaria para garantizar el óptimo funcionamiento del modelo predictivo planteado en la presente investigación.

**Tabla 2:** *Información a extraer*

<b>Información</b>	<b>Descripción</b>
Equipos de red	Información relacionada con la velocidad de subida y bajada de cada equipo del punto de encuentro y la descripción de la marca y modelo de los dispositivos.
Datos de la población	Total de habitantes de las parroquias en las cuales se brinda cobertura de internet gratuito con los puntos digitales gratuitos.
Visitas (cobertura)	Total de visitas que cada uno de los puntos digitales gratuitos recibió.
Capacitación	Total de capacitaciones que se dieron en cada uno de los puntos digitales gratuitos.
Ancho de banda	Información relacionada con el ancho de banda de subida y de bajada tomado cada media hora; así también, como información de la ubicación geográfica de cada punto del encuentro.

**Fuente:** Elaboración propia.

### 3.3.2 Definición de temporalidad de los datos

Para el diseño e implementación del modelo predictivo, se ha determinado que los datos necesarios a extraer corresponden a aquellos relacionados con el tráfico web generado en los puntos de encuentro desde julio hasta septiembre de 2023; para lo cual, la información será exportada directamente desde cada sistema de información los cinco primeros días de cada mes.

### 3.3.3 Datos extraídos

En la Tabla 3 se describe el total de registros extraídos para cada mes, en las misma se puede evidenciar que se cuenta con un total de 81.720 de registros que serán considerados en la siguiente fase de filtrado y limpieza.

**Tabla 3:** *Total de registros extraídos*

Mes	Total de registros
Julio	2.9016
Ago	2.6784
Sept	2.5920
<b>Tot</b>	<b>81.720</b>

**Fuente:** Elaboración propia.

### **3.3.4 Herramientas de desarrollo**

Es fundamental reconocer la importancia de definir cuidadosamente el lenguaje de programación y las librerías adecuadas para el desarrollo del presente proyecto, ya que estas decisiones influirán significativamente en la eficiencia, flexibilidad y rendimiento del modelo predictivo. Para garantizar su correcto desarrollo se utilizaron las siguientes herramientas, mismas que son las más utilizadas en desarrollo similares alrededor del mundo:

- Lenguaje de programación: Python
- Librerías:
  - Numpy,
  - Pandas,
  - Os,
  - Tensorflow,
  - Matplotlib.pyplot,
  - Scipy; y,
  - Tensorflow.keras.
- Entornos de Desarrollo Integrado (IDE): Jupyter Notebook

### **3.4. Filtrado y limpieza de los datos**

El filtrado y la limpieza de datos constituyen dos etapas fundamentales durante la preparación de los datos para su análisis, durante la fase de filtrado, es necesario identificar y eliminar datos no deseados o irrelevantes, tales como valores atípicos, duplicados o información que no aporta significativamente al análisis; por otro lado, la fase de limpieza de datos se centra

en corregir errores, completar información faltante y estandarizar el formato, para incrementar la calidad general de los datos. Los criterios aplicados fueron:

- **Texto es cambiado a mayúsculas:** permite estandarizar el texto y reducir la variabilidad de las palabras evitando que el modelo considere dos palabras idénticas en minúsculas y mayúsculas como diferentes.
- **Eliminación de tildes:** asegura que las palabras acentuadas se presenten de la misma manera que las no acentuadas.
- **Identificación de información faltante:** asegura que todos los registros cuenten con la información completa para el posterior procesamiento
- **Identificación de información duplicada:** asegura que no se considere información duplicada que alteraría los resultados finales.

Finalmente, como se describe en la Tabla 4 luego del proceso de filtrado y limpieza de datos el modelo contará con un total de 93.891 de registros que permitirán comenzar con el entrenamiento del modelo predictivo.

**Tabla 4:** *Total de registros después del filtrado y limpieza de datos*

Mes	Total de registros
Julio	2.9016
Ago	2.6784

sto

<b>Mes</b>	<b>Total de registros</b>
Sept	2.5920
iembre	
<b>Tot</b>	<b>81.720</b>

**al**

---

**Fuente:** Elaboración propia.

### **3.5. Entrenamiento del modelo**

La importancia del entrenamiento del modelo de inteligencia artificial (IA) radica en su capacidad para aprender patrones y relaciones complejas a partir de los datos disponibles, durante el entrenamiento, el modelo ajusta sus parámetros internos mediante algoritmos de mejora continua para minimizar una función de pérdida, lo que le permite generalizar y realizar predicciones precisas sobre datos nuevos y no vistos. Este proceso es fundamental para garantizar el éxito del modelo, ya que determina su capacidad para capturar la estructura subyacente de los datos y realizar inferencias significativas. Un entrenamiento riguroso y adecuado involucra la selección cuidadosa de algoritmos de aprendizaje, la optimización de hiperparámetros, la validación cruzada y la evaluación constante del rendimiento del modelo para garantizar que esté capacitado de manera óptima y pueda adaptarse a diferentes escenarios y conjuntos de datos.

#### **3.5.1 Diseño del Dataset**

Un dataset bien diseñado es el pilar fundamental sobre la cual se construyen los modelos de inteligencia artificial robustos y precisos, la importancia de este diseño radica en la representatividad y calidad de los datos recopilados, que deben capturar fielmente la diversidad y

complejidad del problema que se está abordando. Un dataset bien diseñado no solo garantiza la efectividad de los modelos de inteligencia artificial, sino que también facilita la interpretación de los resultados.

La generación del dataset implicó la identificación de las variables, atributos y tipos de datos que permitan contar con un modelo estructurado de información que permita realizar el análisis, modelado de datos, entrenamiento del modelo de inteligencia artificial que permita el desarrollo del modelo predictivo propuesto, la Tabla 5 muestra la estructura definida.

**Tabla 5:** *Estructura del Dataset*

<b>Campo</b>	<b>Tipo de datos</b>
CODIGO	String
PROVINCIA	String

<b>Campo</b>	<b>Tipo de datos</b>
CANTON	String
PARROQUIA	String
NOMBRE DEL PUNTO DEL ENCUENTRO	String
DIRECCION IP	String
IDENTIFICADOR	Integer
VELOCIDAD DE SUBIDA (Mbps)	Integer
VELOCIDAD DE BAJADA (Mbps)	Integer
FECHA	Date
HORA	Date
CONEXIONES	Integer
ANCHO DE BANDA SUBIDA (MB)	Double
ANCHO DE BANDA BAJADA (MB)	Double
POBLACION	Integer

**Fuente:** Elaboración propia.

Al momento la construcción del DataSet no es un proceso automático, de manera manual se recopilan los archivos en formato Excel de y se construye un archivó en formato csv único con toda la información necesaria.



### 3.5.2 Carga de información al DataSet

Para garantizar la efectividad de la carga de información al dataset, fue necesario realizar algunas validaciones para asegurarse de que los datos sean correctos y estén en el formato esperado. A continuación, se detallan las validaciones realizadas:

- **Verificación de columnas requeridas:** permite asegurar que el dataset y los archivos de donde proviene la información contengan todas las variables esenciales para el análisis, evitando posibles errores y garantizando la consistencia en las etapas posteriores del proceso.
- **Detección de datos faltantes:** La presencia de datos faltantes puede afectar la calidad y validez de los resultados, introduciendo sesgos y distorsiones en las conclusiones obtenidas; razón por la cual, esta validación permite comprobar que el dataset cuente con toda la información para su posterior procesamiento.
- **Validación de tipos de datos:** permite asegurar que las variables se alineen con los tipos de datos esperados, lo cual fundamental para evitar errores en cálculos, interpretaciones incorrectas y fallos en el código.
- **Identificación de valores atípicos:** la identificación temprana y la gestión adecuada de valores atípicos es esencial para evitar que estos puntos influyan de manera desproporcionada en los resultados, distorsionando interpretaciones y afectando la validez de conclusiones.

Al finalizar la carga de la información se comprobó que se hayan cargado todos los registros al dataset, garantizando de esta manera que la información a ser procesada es la misma que el archivo de origen.

### **3.5.3 Definición de registros para pruebas y entrenamiento de los modelos**

Al definir el porcentaje de registros para las pruebas, es importante encontrar un equilibrio que permita una evaluación rigurosa del modelo mientras se asegura que tenga suficiente capacidad para generalizar a datos no evaluados. Para este modelo predictivo se ha realizado la siguiente división de los datos:

- 10% del total de registros para pruebas (8.172).
- 90% del total de registros para entrenamiento (73.548).

La selección del 10% de registros para pruebas responde a un proceso de selección aleatorio garantizando la máxima representatividad de los datos.

### **3.5.4 Modelo 1: Activaciones relu, relu y sigmoid**

Un modelo de inteligencia artificial (IA) que sigue la secuencia de activaciones relu, relu y sigmoid en sus capas neuronales está en la capacidad de diseñarse para una variedad de tareas de clasificación o regresión. En este sentido, "relu" se refiere a la función de activación rectificadora lineal, que introduce no linealidad en el modelo y ayuda a aprender representaciones más complejas de los datos, su utilización como activación en las primeras dos capas, relu permite que el modelo esté en la capacidad de aprender características generales del conjunto de datos.

Después de la secuencia relu-relu, la activación sigmoid se utiliza típicamente en la capa de salida, especialmente en problemas de clasificación binaria, ya que produce valores en el rango de 0 a 1, lo que puede interpretarse como probabilidades.

**Figura 24:** *Arquitectura del modelo 1*

```
#Arquitectura de la red neuronal 1
lr=0.0001 #learning rate
nn=[6, 100, 40, 2]

#Modelo por capas de la estructura de la red
modelo=kr.Sequential()
#Agregar primera capa oculta
modelo.add(Dense(nn[1],input_dim=nn[0],activation='relu'))
modelo.add(Dense(nn[2],activation='relu'))
modelo.add(Dense(nn[3],activation='sigmoid'))
modelo.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 100)	700
dense_1 (Dense)	(None, 40)	4040
dense_2 (Dense)	(None, 2)	82

```
-----
Total params: 4822 (18.84 KB)
Trainable params: 4822 (18.84 KB)
Non-trainable params: 0 (0.00 Byte)
-----
```

**Fuente:** Elaboración propia.

**Figura 25:** Entrenamiento del modelo 1

```
#Entrenamos La red Neuronal 1
historial=modelo.fit(trainIn, trainOut,
                    epochs=30,
                    batch_size=20,
                    validation_data=(testIn, testOut),
                    verbose=1)
```

**Fuente:** Elaboración propia.

La Tabla 6 describe los resultados de los indicadores que pueden calcularse mediante la matriz de confusión, demostrando que la exactitud de este modelo se encuentra en el 74,66%.

**Tabla 6:** Resultados de la matriz de confusión del modelo 1

Matriz de confusión		Negativos	Positivos	Precisión
Real	Negativo	<b>3040</b> (a)	<b>1046</b> (b)	d/(b+d)
	Positivo	<b>1025</b> (c)	<b>3061</b> (d)	<b>74,53%</b>
		<b>Sensibilidad</b> d/(d+c)	<b>Especificidad</b> a/(a+b)	<b>Exactitud</b> (a+d)/(a+b+c+d)
		<b>74,91%</b>	<b>74,40%</b>	<b>74,66%</b>

**Fuente:** Elaboración propia.

Para la elaboración de la matriz de ganancias para los 3 modelos se utilizarán los siguientes valores constantes que permitirán identificar un aproximado de los beneficios que cada modelo puede generar:

- Ganancia por satisfacer la demanda de los usuarios (TP): \$100
- Pérdida por desperdicio de recursos (FP): -\$100
- Pérdida por falta de recursos asignados (FN): -\$100
- Sin costo ni beneficio (TN): \$0

Con los valores antes descritos y sobre la base de la matriz de confusión, la Tabla 7 muestra el costo que se generan para cada cuadrante de la predicción, determinando que este modelo estaría en la capacidad de generar un ahorro de aproximadamente \$99.000 dólares.

**Tabla 7:** Resultados de la matriz de ganancias del modelo 1

	Predicción	
	Negativa	Positiva
Sin consumo	\$ 0 * 3040	-\$100 * 1046
Consumo real	-\$100 * 1025	\$100 * 3061

**Fuente:** Elaboración propia.

### 3.5.5 Modelo 2: Activaciones softmax, selu y tanh

Un modelo de inteligencia artificial (IA) con activaciones softmax, selu y tanh en sus capas neuronales está diseñado para problemas de clasificación multiclase o regresión, la activación

softmax se usa en la capa de salida para solventar problemas de clasificación multiclase, ya que produce una distribución de probabilidad sobre varias clases, lo que permite asignar una clase a cada instancia de entrada, esta función asegura que las salidas sumen uno, lo que facilita la interpretación como probabilidades.

La activación relu, utilizada en la segunda capa, introduce no linealidad en el modelo y ayuda a aprender características más complejas de los datos, relu es una función de activación popular debido a su eficiencia computacional y su capacidad para prevenir el problema del desvanecimiento del gradiente. Por último, la activación tanh en la tercera capa, que produce valores en el rango  $[-1, 1]$ , se utilizó para regularizar y normalizar las salidas de las neuronas, lo que ayuda a estabilizar el entrenamiento y mejorar la convergencia del modelo. En conjunto, estas activaciones permiten al modelo aprender representaciones complejas de los datos y generar predicciones adecuadas.

**Figura 26:** *Arquitectura del modelo 2*

```

#Arquitectura de la red neuronal 2
nn2=[6, 300, 100, 2]

#Modelo por capas de la estructura de la red
modelo2=kr.Sequential()
#Agregar primera capa oculta
modelo2.add(Dense(nn2[1],input_dim=nn2[0],activation='softmax'))
modelo2.add(Dense(nn2[2],activation='selu'))
modelo2.add(Dropout(0.5))
modelo2.add(Dense(nn2[3],activation='tanh'))
modelo2.summary()

```

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
dense_3 (Dense)	(None, 300)	2100
dense_4 (Dense)	(None, 40)	12040
dropout (Dropout)	(None, 40)	0
dense_5 (Dense)	(None, 2)	82

=====  
 Total params: 14222 (55.55 KB)  
 Trainable params: 14222 (55.55 KB)  
 Non-trainable params: 0 (0.00 Byte)

**Fuente:** Elaboración propia.

**Figura 27:** Entrenamiento del modelo 2

```

#Entrenamos la red neuronal 1
historial=modelo.fit(trainIn, trainOut,
                    epochs=30,
                    batch_size=20,
                    validation_data=(testIn, testOut),
                    verbose=1)

```

**Fuente:** Elaboración propia.

La Tabla 8 **Tabla 6** describe los resultados de los indicadores que pueden calcularse mediante la matriz de confusión, demostrando que la exactitud de este modelo es del 49,87% lo que indica que su uso no brindaría predicciones confiables.

**Tabla 8:** Resultados de la matriz de confusión del modelo 2

<b>Matriz de confusión</b>		<b>Negativos</b>	<b>Positivos</b>	<b>Precisión</b>
Real	Negativo	<b>2025</b> (a)	<b>2061</b> (b)	$d/(b+d)$
	Positivo	<b>2036</b> (c)	<b>2050</b> (d)	<b>49,87%</b>
		<b>Sensibilidad</b> $d/(d+c)$	<b>Especificidad</b> $a/(a+b)$	<b>Exactitud</b> $(a+d)/(a+b+c+d)$
		<b>50,17%</b>	<b>49,56%</b>	<b>49,87%</b>

**Fuente:** Elaboración propia.

Con los valores antes descritos y sobre la base de la matriz de confusión, la Tabla 9 muestra el costo que se generan para cada cuadrante de la predicción, determinando que este modelo estaría en la capacidad de generar pérdidas de aproximadamente \$204.700 dólares.

**Tabla 9:** Resultados de la matriz de ganancias del modelo 2



	Predicción	
	Negativa	Positiva
<b>Sin consumo</b>	\$ 0 * 2025	-\$100 * 2061
<b>Consumo real</b>	-\$100 * 2036	\$100 * 2050

**Fuente:** Elaboración propia.

### 3.5.6 Modelo 3: Activaciones exponential, hard\_sigmoid y mish

Un modelo de inteligencia artificial (IA) con activaciones exponential, hard\_sigmoid y mish en sus capas neuronales tiene la capacidad de adaptarse a problemas específicos que requieren funciones de activación diferentes a las tradicionales. La activación exponential es menos común, pero es muy útil en situaciones donde se necesita una respuesta exponencial a las entradas, como es el caso de este modelo predictivo.

La activación hard\_sigmoid es una versión aproximada de la función sigmoid, que es útil en aplicaciones donde se necesita una activación suave pero que computacionalmente sea más eficiente, al ser una función de activación lineal por partes que produce una salida en el rango [0, 1], similar a la sigmoid pero con una función de activación más simple.

La activación mish es una función de activación relativamente nueva que ha demostrado buenos resultados, al ser una función suave y derivable que ofrece una alternativa interesante a funciones de activación más comunes como relu o sigmoid. Algunos estudios sugieren que mish puede ayudar a mitigar problemas de desvanecimiento del gradiente y mejorar el rendimiento del modelo. En conjunto, estas activaciones proporcionan una diversidad de respuestas a las entradas del modelo, lo que en este modelo es muy útil al requerir una respuesta no lineal y compleja.

**Figura 28:** *Arquitectura del modelo 3*

```
#Arquitectura de la red neuronal 3
nn3=[6, 600, 1024, 2]

#Modelo por capas de la estructura de la red
modelo3=kr.Sequential()
#Agregar primera capa oculta
modelo3.add(Dense(nn3[1],input_dim=nn3[0],activation='exponential'))
modelo3.add(Dropout(0.5))
modelo3.add(Dense(nn3[2],activation='hard_sigmoid'))
modelo3.add(Dropout(0.5))
modelo3.add(Dense(nn3[3],activation='mish'))
modelo3.summary()
```

Model: "sequential\_2"

Layer (type)	Output Shape	Param #
dense_6 (Dense)	(None, 600)	4200
dropout_1 (Dropout)	(None, 600)	0
dense_7 (Dense)	(None, 1024)	615424
dropout_2 (Dropout)	(None, 1024)	0
dense_8 (Dense)	(None, 2)	2050

```
=====  
Total params: 621674 (2.37 MB)  
Trainable params: 621674 (2.37 MB)  
Non-trainable params: 0 (0.00 Byte)
```

**Fuente:** Elaboración propia.

**Figura 29:** *Entrenamiento del modelo 3*

```
### Entrenamos la red Neuronal 3
historial3=modelo3.fit(trainIn, trainOut,
                      epochs=50,
                      batch_size=100,
                      validation_data=(testIn, testOut),
                      verbose=1)
```

**Fuente:** Elaboración propia.

La Tabla 10 **Tabla 6** muestra los resultados de los indicadores que pueden calcularse mediante la matriz de confusión, demostrando que la exactitud de este modelo es del 76,66% lo que indica que su uso brindaría mejores predicciones que los dos modelos antes descritos, considerando adicionalmente que el resto de los indicadores de igual manera tienen resultados muy similares por sobre el 75%.

**Tabla 10:** Resultados de la matriz de confusión del modelo 3

<b>Matriz de confusión</b>		<b>Negativos</b>	<b>Positivos</b>	<b>Precisión</b>
Real	Negativo	<b>3115</b> (a)	<b>971</b> (b)	$d/(b+d)$
	Positivo	<b>936</b> (c)	<b>3150</b> (d)	<b>76,44%</b>
		<b>Sensibilidad</b> $d/(d+c)$	<b>Especificidad</b> $a/(a+b)$	<b>Exactitud</b> $(a+d)/(a+b+c+d)$
		<b>77,09%</b>	<b>76,24%</b>	<b>76,66%</b>

**Fuente:** Elaboración propia.

- **Métrica de Exactitud:** Es una medida que permite evaluar el rendimiento de un modelo de clasificación, el valor obtenido es de un 76,66%, es decir que el modelo está clasificando

de manera acertada la mayoría de las publicaciones en sus respectivas categorías de sentimiento.

- **Métrica de Precisión:** La precisión indica la proporción de las predicciones positivas que son verdaderamente positivas en comparación con el total de predicciones clasificadas como positivas; el valor obtenido es del 76,44% por lo que se puede concluir que el modelo está clasificando de manera acertada las predicciones positivas.
- **Métrica de sensibilidad:** Esta métrica evalúa la proporción de predicciones positivas correctas sobre el total de observaciones que son verdaderamente positivas en los datos, el valor calculado es del 77,09%, es decir que el modelo identifica los sentimientos positivos expresados en los comentarios de Facebook con mucha precisión.
- **Métrica de Especificidad:** Representa la proporción de las predicciones negativas correctas (verdaderos negativos) sobre el número total de observaciones que en realidad son negativas en los datos, el valor calculado es del 76,24%, es decir que el algoritmo identifica las publicaciones con sentimientos negativos de una manera correcta.

Con los valores antes descritos y sobre la base de la matriz de confusión, la Tabla 11 muestra el costo que se generan para cada cuadrante de la predicción, determinando que este modelo estaría en la capacidad de generar ganancias de aproximadamente \$124.300 dólares.

**Tabla 11:** *Resultados de la matriz de ganancias del modelo 3*

	<b>Predicción</b>	<b>Predicción</b>
	<b>Negativa</b>	<b>Positiva</b>

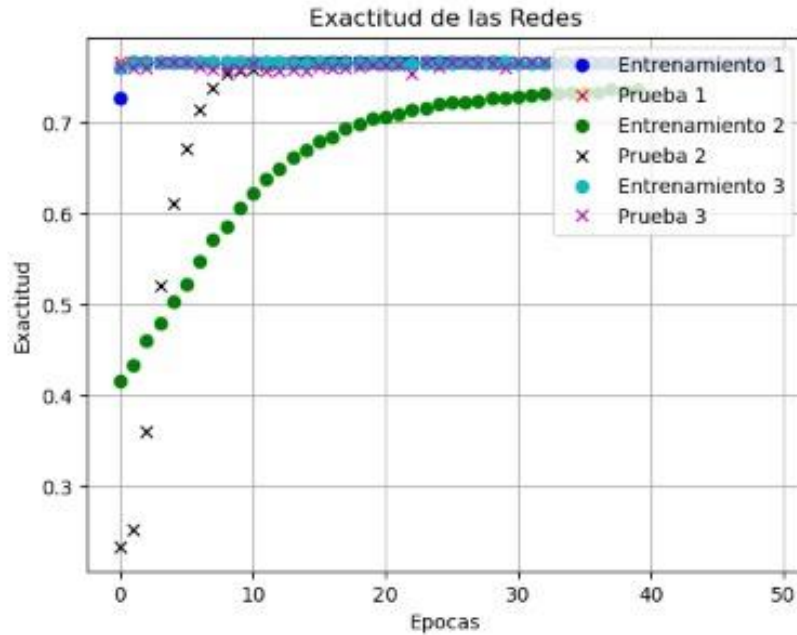
<b>Sin consumo</b>	\$ 0 * 3115	-\$100 * 971
<b>Consumo real</b>	-\$100 * 936	\$100 * 3150

**Fuente:** Elaboración propia.

### 3.5.7 Exactitud de las redes

Es fundamental comprender la importancia de analizar la exactitud de los modelos propuestos en cualquier proyecto de IA. La exactitud es una métrica crucial que evalúa qué tan bien los modelos pueden predecir o clasificar correctamente los datos de prueba. Al analizar la exactitud de los modelos, podemos determinar su capacidad para generalizar y hacer predicciones precisas en datos no vistos.

**Figura 30:** *Exactitud de las redes*



**Fuente:** Elaboración propia.

El análisis de la exactitud nos proporciona información valiosa sobre el rendimiento relativo de diferentes modelos y nos ayuda a identificar cuál de ellos es el más adecuado para nuestro problema específico. Además, nos permite detectar posibles problemas como el sobreajuste (el modelo se ajusta de manera exagerada a los datos de entrenamiento y está en la capacidad de generalizar correctamente a los nuevos datos) o el subajuste (el modelo es demasiado simple para comprender la complejidad de los datos).

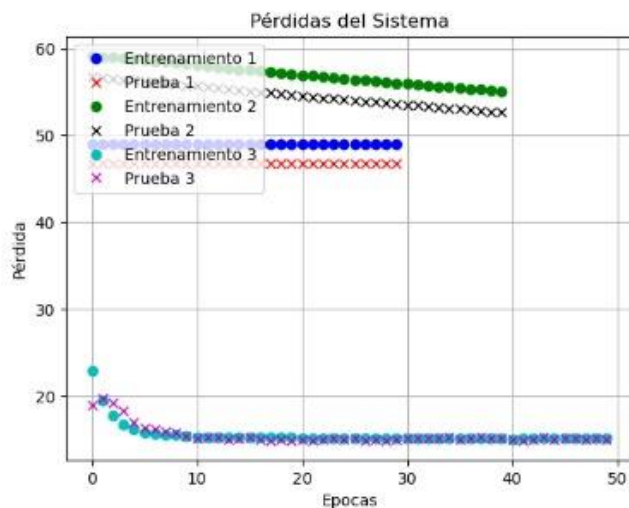
### 3.5.8 Pérdidas del sistema

La evaluación de las pérdidas del sistema en los tres modelos propuestos es esencial para comprender su rendimiento y eficacia en la tarea específica. Las pérdidas, también conocidas como

funciones de pérdida o errores, representan la discrepancia entre las predicciones del modelo y las etiquetas verdaderas en los datos de entrenamiento. Analizar estas pérdidas nos proporciona una medida cuantitativa de qué tan bien los modelos están aprendiendo y generalizando a partir de los datos disponibles.

La importancia de analizar las pérdidas radica en varios aspectos clave del desarrollo de modelos de inteligencia artificial. En primer lugar, nos permite identificar si los modelos están convergiendo durante el entrenamiento y si están mejorando su capacidad para hacer predicciones precisas. Además, nos ayuda a comparar el rendimiento de diferentes modelos y seleccionar aquellos que minimicen las pérdidas de manera más efectiva. Además, el monitoreo continuo de las pérdidas durante el entrenamiento puede ayudar a detectar problemas como el sobreajuste o el subajuste, lo que nos permite realizar ajustes en la arquitectura del modelo o en los hiperparámetros para mejorar su rendimiento general.

**Figura 31:** *Pérdidas del sistema*



**Fuente:** Elaboración propia.

### 3.7. Resultados

A continuación, en la Tabla 12 se muestra un resumen de las distintas evaluaciones realizadas a los 3 modelos propuestos, donde se muestra con un SI el modelo que mejor rendimiento muestre para cada parámetro, evidenciando que el modelo 3 cuya arquitectura cuenta con activaciones exponencial, hard\_sigmoid y mish es el que mejores resultados ha mostrado.

**Tabla 12:** *Resumen de la evaluación de los 3 modelos*

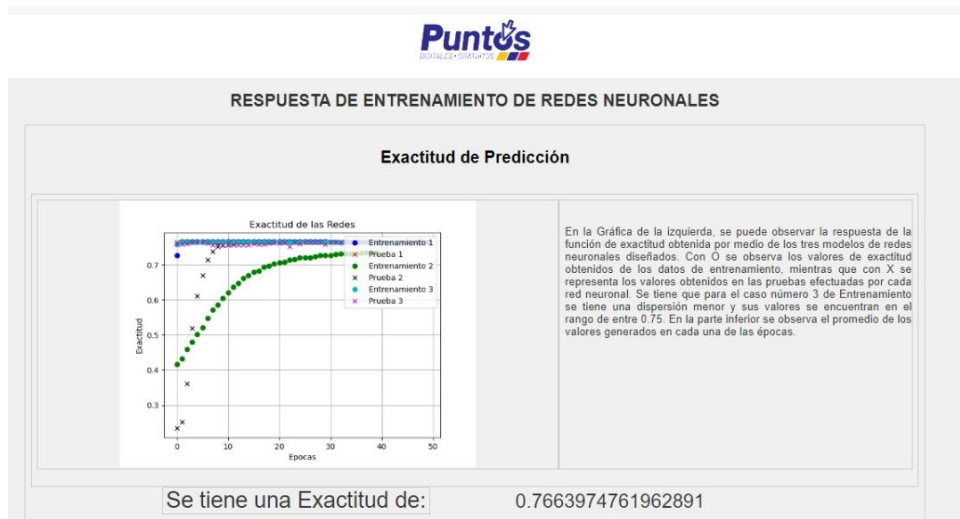
<b>Parámetro</b>	<b>Modelo 1</b>	<b>Modelo 2</b>	<b>Modelo 3</b>
Matriz de confusión			SI
Exactitud de redes			SI
Pérdidas del sistema			SI

**Fuente:** Elaboración propia.

Para presentar los resultados de la predicción de tráfico futuro de internet se desarrolló un sitio web en Flask considerando en primer lugar que proporciona una interfaz intuitiva y accesible para que los usuarios puedan interactuar con los resultados de manera fácil y conveniente, lo que facilita la comprensión de los datos predichos y permite a los usuarios interpretar y aplicar los hallazgos de manera efectiva.



**Figura 32:** Sitio web desarrollado



**Fuente:** Elaboración propia.

Además, un sitio web en Flask ofrece flexibilidad para personalizar la presentación de los resultados según las necesidades específicas del proyecto, lo que permite una mejor visualización y análisis de la información. Finalmente, al integrar el modelo predictivo en un entorno web, se facilita su uso y adopción por parte de un público más amplio, lo que puede conducir a una mayor utilidad y aplicación de los hallazgos del modelo en la toma de decisiones y la planificación de recursos en el ámbito de las redes de Internet.

Para las predicciones del ancho de banda, el sitio web cuenta con la funcionalidad que permite realizar el cálculo con el ingreso de tres parámetros: 1) Dirección IP, 2) Hora de la predicción y 3) Número de usuarios conectados acorde a lo descrito en la Figura 33.

**Figura 33:** Funcionalidad para la predicción de ancho de banda

INGRESO DE DATOS	PREDICCIÓN RESULTANTE
Ingrese la Dirección IP del Punto de Encuentro: <input type="text" value="xxx.xxx.xxx.xxx"/>	El Ancho de Banda de Subida ES: <input type="text" value="5.278167492173446"/> (MB/S)
Ingrese la Hora de la Predicción: <input type="text" value="hh:mm"/>	El Ancho DE Banda de Bajada ES: <input type="text" value="2.9781271015724413"/> (MB/S)
Ingrese la Cantidad de Usuarios Registrado: <input type="text"/>	
<b>INICIAR PREDICCIÓN</b>	

**Fuente:** Elaboración propia.

La Tabla 13 muestra los resultados de las predicciones obtenidas del modelo para cinco direcciones de IP distintas, mismas que acorde a la validación realizada en un ambiente real se ha comprobado su efectividad.

**Tabla 13:** *Resultados del modelo predictivo*

Dirección IP	Hora	Usuarios	Predicción	
			Subida	Bajada
190.214.56.98	15:00	20	6,7	5,59
181.196.246.174	14:00	15	1,72	3,5
181.196.246.54	13:00	30	17,55	5,49
181.196.145.14	10:00	25	13,99	1,03
181.196.244.202	11:00	18	6,94	2,21

**Fuente:** Elaboración propia.

## CONCLUSIONES

- El proyecto de investigación aborda la problemática de la falta de un modelo predictivo para estimar el consumo de red en los Puntos Digitales gratuitos de la zona 5. Esta carencia genera congestiones y lentitud en el servicio, lo que afecta negativamente la experiencia de los usuarios y limita la capacidad de la empresa para tomar decisiones estratégicas relacionadas con la mejora de la gestión administrativa y operativa.
- La preparación meticulosa de un conjunto de datos adecuadamente etiquetados, que representen información sobre el uso de Internet en la red de los Puntos Digitales Gratuitos Zona 5, ha demostrado ser fundamental para el análisis y comprensión de los patrones de uso de Internet en esta área específica. Esta labor garantiza una etiquetación precisa y exhaustiva de los datos recopilados, facilitando así la identificación de tendencias, comportamientos y necesidades específicas de los usuarios en estos puntos digitales. Además de su papel crucial en la generación de información valiosa para la toma de decisiones informadas y el diseño de estrategias efectivas de mejora y optimización de los servicios de Internet en estos espacios comunitarios, la disponibilidad de un conjunto de datos bien etiquetados proporciona una base sólida para la implementación y evaluación de modelos de machine learning y otras técnicas de análisis de datos. Los datos utilizados para construir este dataset fueron extraídos de diversas fuentes de la empresa, incluyendo una selección meticulosa de datos históricos relevantes. Posteriormente, se llevó a cabo la normalización de los datos y la división entre conjuntos de entrenamiento y prueba, que se utilizaron para evaluar el rendimiento de tres arquitecturas de redes neuronales distintas con el objetivo de identificar la que alcanzara una precisión mínima del 75%.

- Es importante destacar que los datos fueron sometidos a rigurosos procesos de filtrado y limpieza para asegurar su calidad y fiabilidad. Esta medida garantiza que los usuarios, particularmente el personal administrativo de Intel, puedan llevar a cabo una revisión y análisis profundo de manera eficiente. Los datos se presentan en forma numérica y se encuentran expresados en MB, lo que facilita su interpretación en el sistema web interactivo desarrollado.
- La implementación de este algoritmo predictivo ha generado un impacto positivo directo en el equipo administrativo y operativo de Intel. Ahora disponen de un modelo actualizado regularmente con datos precisos y una interfaz de usuario intuitiva. Esto les permite detectar tempranamente cualquier congestión de tráfico en los puntos de conexión de red (PDG) de la zona 5 que pueda surgir con el tiempo.

## RECOMENDACIONES

Considerando la experiencia acumulada en el desarrollo de un modelo predictivo se plantea las siguientes recomendaciones:

- A medida que el modelo acumule más información es necesario que constantemente se valide el modelo mediante la matriz de confusión que incluye métricas de evaluación del rendimiento del modelo, como precisión, sensibilidad, especificidad, y curvas ROC. La validación continua garantiza que el modelo siga siendo preciso y efectivo en la predicción del tráfico, incluso a medida que cambian las condiciones y patrones de uso de la red. Además, permite detectar posibles desviaciones o errores en las predicciones y tomar medidas correctivas para mejorar la precisión del modelo. Esta práctica contribuye a mantener la fiabilidad y relevancia del modelo a lo largo del tiempo, lo que es crucial para su aplicación práctica en entornos reales de redes de comunicación.
- Para mejorar la visualización y comprensión de las predicciones se requiere la incorporación de visualizaciones efectivas para presentar los resultados del modelo de manera comprensible, como mapas de calor, curvas de aprendizaje y comparaciones entre los distintos modelos. Para lo cual, es importante adaptar la visualización a las necesidades específicas del público objetivo, como ingenieros de redes, administradores de sistemas o gerentes de TI, para garantizar que la información sea relevante y útil para la toma de decisiones informadas, mediante una visualización mejorada de los datos del modelo predictivo, se facilita la comprensión de los patrones de tráfico de red y se mejora la capacidad de los usuarios para identificar y abordar eficazmente los problemas de rendimiento y congestión en la red.

- Se debe examinar el impacto práctico del modelo en la toma de decisiones y la gestión de los Puntos Digitales Gratuitos, esto implica evaluar cómo las predicciones del modelo están siendo utilizadas por los responsables de la red y cómo estas decisiones están afectando el rendimiento y la eficiencia de la red en general. Además, es importante recopilar retroalimentación de los usuarios y stakeholders para identificar áreas de mejora y realizar ajustes en el modelo según sea necesario. Mediante un análisis continuo del impacto en la toma de decisiones, se puede garantizar que el modelo predictivo esté generando valor real para la organización y contribuyendo de manera significativa a la mejora de la infraestructura de red.

## REFERENCIAS

- Abdelhadi, A., & Guy, P. (2018). Traffic prediction based on LSTM recurrent neural networks in SDN. *NOMS* , 1-5. <https://doi.org/10.1109/NOMS.2018.8406199>
- Aguirre, D. (2022). Modelo algorítmico para alta disponibilidad en transporte de volúmenes crecientes de tráfico variable en redes ópticas. Retrieved 4 de Febrero de 2024, from Sitio web de la Universidad Distrital Francisco José de Caldas: <https://repository.udistrital.edu.co/bitstream/handle/11349/30617/AguirreMorenoDiegoFernando2022.pdf?sequence=1&isAllowed=y>
- Aparicio, L. E., Amaya, O. J., & Devia, P. A. (2023). Aplicación de Deep Learning para la identificación de defectos superficiales utilizados en control de calidad de manufactura y producción industrial: una revisión de la literatura. *Ingeniería*, 28(1). <https://doi.org/https://doi.org/10.14483/23448393.18934>
- Atamari, J., Flores, C., Mamani, J., & Rondon, S. (2022). Uso de árboles de decisión para detectar si una habitación está ocupada usando Python. *Innovación y Software*, 3(1), 58-66. <https://www.redalyc.org/journal/6738/673870840005/673870840005.pdf>
- Audacia Comunicación. (2019). Empresas mexicanas no logran identificar el 51% del tráfico en su red. Sitio web: <https://www.audacia.com.mx/2018/04/24/empresas-mexicanas-no-logran-identificar-el-51-del-trafico-en-su-red-2/>
- Barba, R., Mantilla, C., Arellano, A., Ramos, V., & Barraqueta, P. . (2018). Aplicación de Perceptrón Multicapa en la predicción de tráfico para redes académicas empresariales.

<https://journal.espe.edu.ec/ojs/index.php/Infociencia/article/view/1223/word>

Barrezueta, P. (2018). DESARROLLO DE UN ALGORITMO DE REDES NEURONALES ARTIFICIALES APLICADO A LA PREDICCIÓN DE TRÁFICO DE LA INFRAESTRUCTURA DE COMUNICACIONES DE REDES CORPORATIVAS. Sitio web

web

Universidad

SPOCH:

<http://dspace.esPOCH.edu.ec/bitstream/123456789/8444/1/98T00190.pdf>

BNEW. (2023). El número de usuarios de internet en el mundo crece. Sitio web:

[https://marketing4ecommerce.net/usuarios-de-internet-](https://marketing4ecommerce.net/usuarios-de-internet-mundo/#:~:text=De%20acuerdo%20con%20el%20informe,(7.910%20millones%20de%20personas).)

[mundo/#:~:text=De%20acuerdo%20con%20el%20informe,\(7.910%20millones%20de%20](https://marketing4ecommerce.net/usuarios-de-internet-mundo/#:~:text=De%20acuerdo%20con%20el%20informe,(7.910%20millones%20de%20personas).)

[0personas\).](https://marketing4ecommerce.net/usuarios-de-internet-mundo/#:~:text=De%20acuerdo%20con%20el%20informe,(7.910%20millones%20de%20personas).)

Bolívar, J. (2019). Estudio de algoritmos de Machine Learning para clasificación de tráfico de red de datos TCP/IP. Retrieved 4 de Febreri de 2024, from Sitio web Universidad de las Palmas de Gran Canaria: <http://hdl.handle.net/10553/75971>

Borrero, A., & Quintero, C. (2018). Diagnóstico y caracterización de series autosimilares en trazas de un dispositivo de almacenamiento con aplicaciones financieras. *Ciencia e Ingeniería*, 35(1), 47-55. <https://www.redalyc.org/articulo.oa?id=507550624006>

Britto Montoya, J. A., Pérez Rivera, C. A., & Isaza Echeverry, G. A. (2005). APLICACIÓN DE REDES NEURONALES PARA LA DETECCIÓN DE INTRUSOS EN REDES Y SISTEMAS DE INFORMACIÓN . Pereira, Colombia: Universidad Tecnológica de



- Pereira. Retrieved 4 de Febrero de 2024, from <https://www.redalyc.org/pdf/849/84911698042.pdf>
- Calcano, M. A. (2021). Internet, redes sociales y libertad de expresión. *Cuestiones constitucionales*(44), 35-54. <https://doi.org/https://doi.org/10.14482/INDES.30.1.303.661>
- Castillo, C., & Forero, F. (2020). Caracterización de IPv6. *Tecnura*, 17(36), 111-128. <https://www.redalyc.org/articulo.oa?id=257028093010>
- CEPAL. (2023). Portal de desigualdades en América Latina. Sitio web: <https://statistics.cepal.org/portal/inequalities/>
- Chávez, C. (2021). RUTA ÓPTIMA DE TRÁFICO DE UNA RED VIRTUAL BASADA EN ANÁLISIS DE DATOS Y ALGORITMO DE MACHINE LEARNING. Retrieved 4 de Febrero de 2024, from Sitio web Universidad Técnica de Ambato: <https://repositorio.uta.edu.ec/bitstream/123456789/34043/1/t1916mtel.pdf>
- Contreas, L. E., Nieves, N., & González, K. (2023). Prediction of University-Level Academic Performance through Machine Learning Mechanisms and Supervised Methods. *Ingeniería*, 28(1). <https://doi.org/https://doi.org/10.14483/23448393.19514>
- DataSource. (2023). Comprensión de la Matriz de Confusión y Cómo Implementarla en Python. Sitio web DataSource: <https://www.datasources.ai/es/data-science-articles/comprension-de-la-matriz-de-confusion-y-como-implementarla-en-python>

- Donoso, Y., Galera, I., & Fabregat, R. (2021). Algoritmo de balanceo de carga aplicado a enrutamiento multicast. *Ingeniería y Desarrollo*(15), 31-44. <https://www.redalyc.org/articulo.oa?id=85201503>
- Fernández, E., Fernández, Y., & Crespo, M. (2020). Integración de las tecnologías de la información y la comunicación. *Revista Cubana de Información*, 31(3). <https://www.redalyc.org/journal/3776/377665638015/377665638015.pdf>
- Forero, W., & Negre, F. (2023). Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática. *Estudios e Investigaciones*, 27(1), 209-253. <https://doi.org/https://doi.org/10.5944/ried.27.1.37491>
- Gildardo, M. (2019). Predicción de Tráfico en Redes de Telecomunicaciones basado en Técnicas de Inteligencia Analítica. Sitio web CINVESTAV: <https://www.cs.cinvestav.mx/TesisGraduados/2019/tesisGildardoMillan.pdf>
- Gran Josa, J. M. (2019). Predicción de tráfico en redes móviles mediante Deep Learning. Catalunya, España. Retrieved 4 de Febrero de 2024, from <https://openaccess.uoc.edu/handle/10609/87345>
- Hernández, C., Martínez, F., & Escobar, A. (2018). Modelamiento y pronósticos de tráfico correlacionado. *Tecnura*, 11(22), 124-133. <https://www.redalyc.org/articulo.oa?id=257020603011>
- Hernández, C., Torres, N., & Pedraza, L. (2021). Redes neuronales y predicción de tráfico. *Tecnura*, 15(29), 90-97. <https://www.redalyc.org/articulo.oa?id=257020887009>

- Herrera, R., García, J. G., & Bello, B. (2022). Propuesta de modelo predictivo inteligente. Revista Cubana de Ciencias Informáticas, 144-162.  
<https://www.redalyc.org/journal/3783/378370413010/378370413010.pdf>
- INEC. (2023). Ecuador Estado Digital Junio 2023. Sitio web:  
<https://drive.google.com/file/d/1AoEKgM8x6USXZy2Xx0G0915SCvW85FAs/view>
- López, D. A., García, N. Y., & Herreara, J. (2015). Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red . Bogotá, Colombia: Universidad Distrital Francisco José de Caldas. Retrieved 4 de Febrero de 2014, from [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-07642015000500018](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07642015000500018)
- López, D., García, N., & Herrera, J. (2019). Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red. Información tecnológica, 26(5). <https://doi.org/http://dx.doi.org/10.4067/S0718-07642015000500018>.
- López, J., & Salazar, C. (2020). Desarrollo de un prototipo de software predictivo, sobre el consumo de banda ancha en una red que permita predecir la capacidad necesaria en el futuro. Retrieved 4 de Febrero de 2024, from Sitio web Universitaria Agustiniiana: <https://repositorio.uniagustiniana.edu.co/bitstream/handle/123456789/1550/LopezBarrientos-JeissonArley-2020.pdf?sequence=1&isAllowed=y>
- Martínez, C., & Palencia, O. (2021). Modelo de minería de datos para el análisis de la productividad y crecimiento personal en las mujeres emprendedoras: el caso de la Asociación las Rosas. Suma de Negocios, 12(26), 23-30.  
<https://doi.org/10.14349/sumneg/2021.V12.N26.A3>

- Martínez, J.-M., Mejía, J., Muñoz, M., & García, Y.-M. (2018). La Seguridad en Internet de las Cosas: Analizando el Tráfico de Información en Aplicaciones para iOS. *ReCIBE. Revista electrónica de Computación, Informática*, 6(1), 77-96. <https://www.redalyc.org/articulo.oa?id=512253717005>
- Medina, V., Almansa, A., León, M., & Torres, D. (2020). Gestión del conocimiento y Capital Social en la universidad. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 25, 1-21. <https://doi.org/https://doi.org/10.5007/1518-2924.2020.e72536>
- Mejía, J., Vidal, S., & López, J. (2020). Sistema de acceso seguro a recursos de información para redes. *Científica*, 10(4), 199-205. <https://www.redalyc.org/articulo.oa?id=61410407>
- MINTEL. (Abril de 2023). Gobierno de Ecuador incentiva la inclusión digital de niñas, adolescentes y jóvenes mujeres. Sitio web: <https://www.telecomunicaciones.gob.ec/gobierno-de-ecuador-incentiva-la-inclusion-digital-de-ninas-adolescentes-y-jovenes-mujeres/>
- Montaño Moreno, J. J. (2022). Redes neuronales artificiales aplicadas al análisis de datos. Universitat de les Illes Balears. Departament de Psicologia. Retrieved 4 de Febrero de 2024, from <https://www.tdx.cat/handle/10803/9441>
- Norambuena, J. M., Badilla, M. G., & Angulo, Y. L. (2022). Modelos predictivos basados en uso de analíticas de aprendizaje automático. *Texto livre*, 15. <https://doi.org/https://doi.org/10.35699/1983-3652.2022.36310>
- Orange Data Mining. (2023). Data Mining Fruitful and Fun. Sitio web Orange: <https://orangedatamining.com/>

- Orellana, M., & Cedillo, P. (2020). Detección de valores atípicos con técnicas de minería de datos y métodos estadísticos. *Enfoque UTE*, 11(1), 56-67. <https://doi.org/https://doi.org/10.29019/enfoque.v11n1.584>
- Osorio, R., Segura, M., & Alonso, M. (2023). ALGORITMOS SUPERVISADOS PARA LA PREDICCIÓN DEL ANCHO DE BANDA DE LAS APLICACIONES EN AMAZON WEB SERVICE DESDE UNA PYME RURAL. *RIMCI*, 10(20), 27-38. <https://doi.org/http://dx.doi.org/10.21017/rimci.2023.v10.n20.a138>
- Python. (2023). Python. Sitio web Python: <https://www.python.org/>
- Rico, A., Gaytán, N. D., & Sánchez, D. (2019). Construcción e implementación de un modelo para predecir el rendimiento académico de estudiantes universitarios mediante el algoritmo Naïve Bayes. *Diálogos sobre educación*, 10(19). <https://doi.org/10.32870/dse.v0i19.509>
- Ríos, R., & Fermin, J. (2019). ANÁLISIS DE TRÁFICO DE UNA RED LOCAL UNIVERSITARIA. *Télématique*, 8(2), 15-27. <https://www.redalyc.org/articulo.oa?id=78411787002>
- Rivas Asanza, W., & Mazón Olivo, B. (2017). Redes neuronales artificiales aplicadas al reconocimiento de patrones (UTMACH ed.). Machala, El oro, Ecuador: Universidad Técnica de Machala. Retrieved 4 de Febrero de 2024, from [https://www.researchgate.net/profile/Bertha-Mazon-Olivo/publication/327703478\\_Capitulo\\_1\\_Generalidades\\_de\\_las\\_redes\\_neuronales\\_artificiales/links/5b9fe3c0299bf13e6038a1d8/Capitulo-1-Generalidades-de-las-redes-neuronales-artificiales.pdf](https://www.researchgate.net/profile/Bertha-Mazon-Olivo/publication/327703478_Capitulo_1_Generalidades_de_las_redes_neuronales_artificiales/links/5b9fe3c0299bf13e6038a1d8/Capitulo-1-Generalidades-de-las-redes-neuronales-artificiales.pdf)

- Sangama, E., & Amaya, S. (2019). Redes neuronales artificiales para la mejora de la gestión del consumo de ancho de banda de internet en la Municipalidad Distrital de Belén – 2018. Retrieved 4 de Febrero de 2024, from Sitio web Universidad Científica del Perú: <http://repositorio.ucp.edu.pe/handle/UCP/734>
- Sierra, J., & Donoso, Y. (2009). Modelo de asignación predictivo de longitudes de ondas en redes WDM teniendo en cuenta dispersión residual y tráfico unicast/multicast con QoS. Medellín, Colombia: Universidad de Antioquia. Retrieved 04 de Febrero de 2024, from [http://www.scielo.org.co/scielo.php?pid=S0120-62302009000400015&script=sci\\_arttext](http://www.scielo.org.co/scielo.php?pid=S0120-62302009000400015&script=sci_arttext)
- Spitsina, L., Kretinin, A., & Spitsin, V. (2022). Tráfico de internet y desempeño. RETOS. Revista de Ciencias de la Administración y Economía, 23(12). <https://doi.org/https://doi.org/10.17163/ret.n23.2022.06>
- Torres, E., Valderrama, J., de los Santos, A., & Galvez, E. (2023). Implementación de controles de acceso para un sistema web de gestión de prácticas profesionales. Innovación y Software, 4(1), 37-51. <https://www.redalyc.org/journal/6738/673874721003/673874721003.pdf>
- Torres, N., Hernández, C., & Pedraza, L. (2019). Redes neuronales y predicción de tráfico. Tecnura, 15(29), 90-97. <https://www.redalyc.org/articulo.oa?id=257020887009>
- Urra, C., & Ramos, M. (2023). Un enfoque de machine learning para la predicción de la calidad de tableros contrachapados. Maderas. Ciencia y tecnología, 25(36). <https://doi.org/https://doi.org/10.4067/s0718-221x2023000100436>

# ANEXOS

Se presenta la información consultada. Pueden ser, carta de compromiso, carta aval, instrumentos, evidencia fotográfica, entre otros que el autor considere necesario.

## Anexo 1: Encuesta

	<p align="center"><b>Proyecto de Titulación</b> <b>FACULTAD DE SISTEMAS Y TELECOMUNICACIONES</b> <b>INSTITUTO DE POSTGRADO</b></p>
<p><b>Objetivo de la encuesta:</b> Recopilar información relacionada con la calidad del de Internet en los Puntos del Encuentro.</p>	
<p><b>Marque con una x dentro del recuadro de su respuesta</b></p>	
<p><b>1. Características del usuario</b></p>	
<b>Rango de edad</b>	<input type="checkbox"/> Menor de 18 años <input type="checkbox"/> de 18 a 24 años <input type="checkbox"/> de 25 a 34 años <input type="checkbox"/> de 35 a 44 años <input type="checkbox"/> de 45 a 54 años <input type="checkbox"/> de 55 años o más
<b>Género</b>	<input type="checkbox"/> Masculino <input type="checkbox"/> Femenino <input type="checkbox"/> Prefiere no decirlo
<p><b>2. ¿Con qué frecuencia visita el punto digital?</b> <i>Señale una opción</i></p>	
<input type="checkbox"/>	Todos los días
<input type="checkbox"/>	Varios días a la semana
<input type="checkbox"/>	Un día a la semana
<input type="checkbox"/>	Varios días al mes
<input type="checkbox"/>	Un día al mes

<b>3. ¿Qué tan satisfecho estás con la velocidad de conexión a Internet en el punto digital?</b>	
<i>Señale una opción</i>	
<input type="checkbox"/>	Muy insatisfecho
<input type="checkbox"/>	Insatisfecho
<input type="checkbox"/>	Neutral
<input type="checkbox"/>	Satisfecho
<input type="checkbox"/>	Muy satisfecho
<b>4. ¿Conectas tus dispositivos en el punto del encuentro mediante wifi?</b>	
<i>Señale una opción</i>	
<input type="checkbox"/>	Si
<input type="checkbox"/>	No
<input type="checkbox"/>	Aveces
<b>5. ¿Experimentas frecuentemente cortes o interrupciones en la conexión a Internet durante tu visita al punto digital?</b>	
<i>Señale una opción</i>	
<input type="checkbox"/>	Si
<input type="checkbox"/>	No
<input type="checkbox"/>	Aveces
<b>6. ¿Consideras que la calidad del Internet en el punto digital es suficiente para realizar tus actividades en línea de manera satisfactoria?</b>	
<i>Señale una opción</i>	
<input type="checkbox"/>	Si
<input type="checkbox"/>	No
<input type="checkbox"/>	Aveces
<b>7. ¿Recomendarías este punto digital a otras personas basándote en la calidad de su Internet?</b>	
<i>Señale una opción</i>	
<input type="checkbox"/>	Si
<input type="checkbox"/>	No
<input type="checkbox"/>	Aveces