

# **UNIVERSIDAD PERUANA UNION**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

**Escuela Profesional de Ingeniería de Sistemas**



*Una Institución Adventista*

## **Clasificación de la Prioridad de Atención a Reclamos Presentados por Clientes Utilizando Machine Learning.**

Por:

Joselin Jimena Aliaga Ancco  
Lesly Chahuara Flores

Asesor:

Dr. Jorge Alejandro Sánchez Garcés

**Juliaca, diciembre de 2019**

DECLARACIÓN JURADA  
DE AUTORÍA DEL TRABAJO DE  
INVESTIGACIÓN

Dr. Jorge Alejandro Sánchez Garcés, de la Facultad de Ingeniería y Arquitectura,  
Escuela Profesional de Ingeniería de Sistemas, de la Universidad Peruana Unión.

DECLARO:

Que el presente trabajo de investigación titulado: "Clasificación de la Prioridad de Atención a Reclamos Presentados por Clientes Utilizando Machine Learning" constituye la memoria que presentan los estudiantes Joselin Jimena Aliaga Ancco y Lesly Chahuara Flores, para aspirar al grado de bachiller en Ingeniería de Sistemas, cuyo trabajo de investigación ha sido realizado en la Universidad Peruana Unión bajo mi dirección.

Las opiniones y declaraciones de este trabajo de investigación son de entera responsabilidad del autor, sin comprometer a la institución.

Y estando de acuerdo, firmo la presente constancia en Juliaca, a los 5 día del mes de diciembre del año 2019.

  
Dr. Jorge Alejandro Sánchez

Clasificación de la Prioridad de Atención a Reclamos Presentados por  
Clientes Utilizando Machine Learning

# TRABAJO DE INVESTIGACIÓN

Presentada para poder optar el grado de bachiller de Ingeniería de  
Sistemas

## JURADO CALIFICADOR



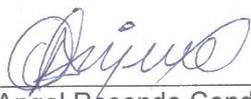
Mtro. Lennin Henry Centurión Julca

Presidente



Mg. Abel Angel Sullon Macalupu

Secretario



Ing. Angel Rosendo Condori  
Coaquira

Vocal



Dr. Jorge Alejandro Sánchez  
Garces

Asesor

Juliaca, 2 de diciembre de 2019



# Clasificación de la Prioridad a Reclamos Presentados por Clientes Utilizando Machine Learning

Joselin Jimena Aliaga Ancco <sup>1</sup>, Lesly Chahuara Flores <sup>2</sup>

*<sup>1</sup>EP. Ingeniería de Sistemas, Facultad de Ingeniería y Arquitectura, Universidad Peruana Unión*

---

## Resumen

En esta investigación se evalúa el desempeño del árbol de decisión, los árboles de decisión son un modelo de clasificación utilizado en la inteligencia artificial, cuya principal característica es su aporte visual a la toma de decisiones. En el dataset se tiene un registro de 706 reclamos de los usuarios que usan los servicios de telecomunicaciones, estos reclamos al no ser solucionados generan pérdidas de clientes, la automatización de clasificación e identificación de los reclamos reduce carga a los trabajadores. Se busca determinar la pertinencia del árbol de decisión, es decir, si puede ser una herramienta de apoyo para la predicción de la prioridad del reclamo.

Palabras clave: Árboles de Decisión, Algoritmos, Clasificación, Reclamos, Telecomunicaciones, Machine Learning.

## Abstract

In this investigation the performance of the decision tree is evaluated, the decision trees are a classification model used in artificial intelligence, whose main characteristic is its visual contribution to decision making. In the data set there is a record of 706 claims of users who use telecommunications services, these claims not being automatically resolved customer losses, classification automation and identification of claims reduce the burden on workers. It seeks to determine the relevance of the decision tree, that is, if it can be a support tool for predicting the priority of the claim.

Keywords: Decision Trees, Algorithms, Classification, Claims, Telecommunications, Machine Learning.

---

## 1. Introducción

En la actualidad existe una alta competencia en la industria de las telecomunicaciones en el Perú, especialmente en diferentes empresas que ofrecen servicios de precios más bajos por planes mejores que otros, planes de voz, de datos, con una atención personalizada. En los últimos años ha habido muchos cambios y mejoras en la industria de las telecomunicaciones, como nuevos servicios, nuevas tecnologías y la liberalización del mercado que se abren a la competencia en el mercado (Huang, Huang, & Kechadi, 2010).

(Barrientos Inostroza, 2012) indica que los principales problemas en cuanto al servicio antes mencionado son: calidad de señal, cobertura, precios, incumplimiento de beneficios, incumplimiento de acuerdos, incumplimiento de tiempos de respuesta. Por lo tanto, la fuga de clientes ocurre por cualquiera de los problemas mencionados. Dichos clientes pertenecen a empresas y personas naturales.

Este problema es conocido como “churn” y está presente en casi todas las empresas, industrias, banca y telecomunicaciones; los autores (Barrientos & Ríos, 2013) (Meza, 2018) mencionan que el “churn” es la fuga de clientes, cuyos índices actualmente alarman a los empresarios especialmente a los de telecomunicaciones siendo que tiene incremento de hasta 30% por año. A este problema se suma la competencia en la industria de las telecomunicaciones, y particularmente está en diversas empresas que ofrecen servicios de Televisión, hasta marzo de 2017 hubo un incremento de 10.3% más suscriptores que en el 2014. Constituye de 5 empresas principalmente “(VTR, 35%; Telefónica, 21.6% DIRECTV, 17.5%, Claro, 15%, Entel S.A 3.1% y otras compañías, 7.6%), Chile (Contreras Morales, Ferreira Correa, & Valle, 2017)”.

Con el pasar de los años el “Organismo Superior de Inversión Privada en Telecomunicaciones (OSIPTTEL)” tuvo un alto crecimiento en el número de quejas y reclamos en los servicios brindados de telefonía móvil, telefonía fija, internet y televisión por cable (EC, 2018)

Los servicios de telecomunicaciones que se brindan en Perú van incrementando, pero también incrementan las observaciones que tienen, van perdiendo el enfoque de calidad que sus servicios que deben a sus clientes. Es de vital importancia conocer de cerca las necesidades, percepciones e incomodidades de los clientes.

Las personas que estipulan un servicio de telecomunicaciones en Perú, tienen el derecho de realizar una queja o reclamo, en caso de que el servicio que le brindaron no funciona al 100%. En este caso el reclamo se genera en primera instancia en las empresas operadoras y luego, en segunda instancia se eleva ante el OSIPTTEL. La mayoría de veces el cliente no sabe cómo hacer un reclamo y deja pasar el tiempo o en otras ocasiones aceptan el mal servicio que le ofrecen (Perú21, 2019).

Machine learning, en la última década ha marcado un rápido y significativo crecimiento del mercado global de automatización de almacenes de datos. El mayor desafío radica en la identificación y manejo de las variables. Se requiere un conjunto de datos grande para reducir el sobreajuste y aumentar la precisión de la validación. Cuando se alcanzan suficientes precisiones de clasificación, se pueden implementar estrategias de selección inteligente para manejar la data de manera eficiente.

¿Por qué agregar aprendizaje automático a la mezcla? Con los modelos de aprendizaje automático apropiados, las organizaciones tienen la capacidad de predecir continuamente cambios en el negocio para que son más capaces de predecir qué sigue. Como los datos son constantemente agregado, los modelos de aprendizaje automático aseguran que la solución sea constantemente actualizada. El valor es sencillo: si usa las fuentes de datos más apropiadas y en constante cambio en el contexto del aprendizaje automático, tienes la oportunidad de predecir el futuro (Langley & Carbonell, 1984).

El agrupamiento es una técnica que es comúnmente conocido en el dominio del aprendizaje automático como un no supervisado método, tiene como objetivo construir a partir de un conjunto de objetos algunos grupos diferentes que sean lo más homogéneos posible (Elaidi, Elhaddar, Benabbou, & Abbar, 2018).

Los árboles de decisión y los conjuntos son muy populares en el aprendizaje automático y, a menudo, logran un rendimiento de vanguardia en las tareas de predicción y clasificación. Tomamos un enfoque probabilístico en el que proyectamos las estructuras del árbol de decisión y los parámetros asociados con los nodos como modelo probabilístico; con ejemplos etiquetados, podemos entrenar el modelo probabilístico usando una variedad de enfoques (Lakshminarayanan, 2016).

Ofrece un mecanismo basado en principios para obtener predicciones probabilísticas que es crucial para aplicaciones donde la cuantificación de la incertidumbre es importante.

Un primer trabajo corresponde a (Saripalli, Tirumala, & Chimmad, 2017), quien realizó la “Evaluación del riesgo de rechazo de reclamos de atención médica mediante el aprendizaje automático”, en esta investigación se hicieron algoritmos para clasificar reclamos de atención médica, que automatiza la identificación de los reclamos, con datos recolectados con campos como: razón, fuente, causa y otros factores decisivos para reducir la carga de trabajo. Con el fin de identificar los reclamos frecuentes para una mejora de la atención brindada.

Un segundo trabajo corresponde a (De Araujo, Santana, & De Alcantara Dos Santos Neto, 2015), quien realizó la “Evaluación de clasificadores basada en el árbol de decisión para aprender el proceso de reclamo médico”, en esta investigación se utilizó técnicas de clasificación basadas en Árbol de decisión para analizar el proceso de reclamo médico, porque para ello se requiere de un revisor médico, el cual se encarga de aconsejar si las solicitudes médicas deberían o no ser autorizadas, y tener un revisor médico requiere de las 24 horas del día y esto genera un alto costo. La finalidad de la investigación fue acelerar el proceso de reclamo médico de autorización o no autorización de solicitudes, con árboles de decisión.

Entonces, antes de realizar el proceso de aprendizaje, la calidad de los datos necesita ser mejorada, en este sentido, el objetivo de esta investigación es usar datos históricos de reclamos de la Industria de Telecomunicaciones de Osiptel diseñar un modelo de aprendizaje automático de la prioridad de reclamos, para ayudar en la mejor decisión al escoger la prioridad de atención.

## 2. Materiales y Métodos

### 2.1. Participantes

El diseño del trabajo de investigación aplicada (descriptiva). Se escogieron los algoritmos de DM y se analizó el conocimiento generado para la toma de decisiones. Para el desarrollo del proceso de DM se utilizó como fuente de información un Data Warehouse (DW), Azure Notebooks, el lenguaje de Python y sus librerías como se muestra en la Tabla 1.

Tabla 1

*Materiales y Herramientas utilizadas para el desarrollo del algoritmo.*

Materiales-Herramienta Tecnológica	Versiones	Descripción
Sistema Osiptel (funde de datos)		Sistema de registros de los reclamos presentados por clientes
Python	3.7	Python es un lenguaje de programación de alto nivel orientado a objetos con semántica dinámica integrada principalmente para el desarrollo web y de aplicaciones. Es extremadamente atractivo en el campo del desarrollo rápido de aplicaciones porque ofrece opciones de tipoo dinámico y enlace dinámico (ThePSF, 2019).
Pandas	0.25.3	En Computación y Ciencia de datos, pandas es una biblioteca de software escrita como extensión de NumPy para manipulación y análisis de datos para el lenguaje de programación Python ("Pandas Basics - Learn Python - Free Interactive Python Tutorial," 2019).
Numpy	1.17.4	NumPy es el paquete fundamental para la computación científica con Python. Contiene entre otras cosas: un poderoso objeto de matriz N-dimensional (NumPy, 2019).
Matplotlib	3.1.1	Matplotlib es una biblioteca de trazado 2D de Python que produce cifras de calidad de publicación en una variedad de formatos impresos y entornos interactivos en todas las plataformas. Matplotlib se puede usar en scripts de Python, los shells de Python e IPython, el cuaderno Jupyter, los servidores de aplicaciones web y cuatro kits de herramientas de interfaz gráfica de usuario (John Hunter, 2012).
DecisionTreeClassifier	0.21.3	Un árbol de decisión es una estructura de árbol similar a un diagrama de flujo donde un nodo interno representa una característica (o atributo), la rama representa una regla de decisión y cada nodo hoja representa el resultado. El nodo superior en un árbol de decisión se conoce como el nodo raíz (Navlani, 2018).
Excel (.csv)	2016	CSV es un formato de archivo simple que se utiliza para almacenar datos tabulares, como una hoja de cálculo o una base de datos. Los archivos en formato CSV se pueden importar y exportar desde programas que almacenan datos en tablas, como Microsoft Excel u OpenOffice Calc. CSV significa "valores separados por comas"(Hope, 2018).
Azure Notebooks		Azure Notebooks es un servicio gratuito para que cualquiera pueda desarrollar y ejecutar código en su navegador con Jupyter. Jupyter es un proyecto de código abierto que permite combinar prosa de rebajas, código ejecutable y gráficos en un solo lienzo (Nbhelp, 2019).

## 2.2. Arquitectura de Solución

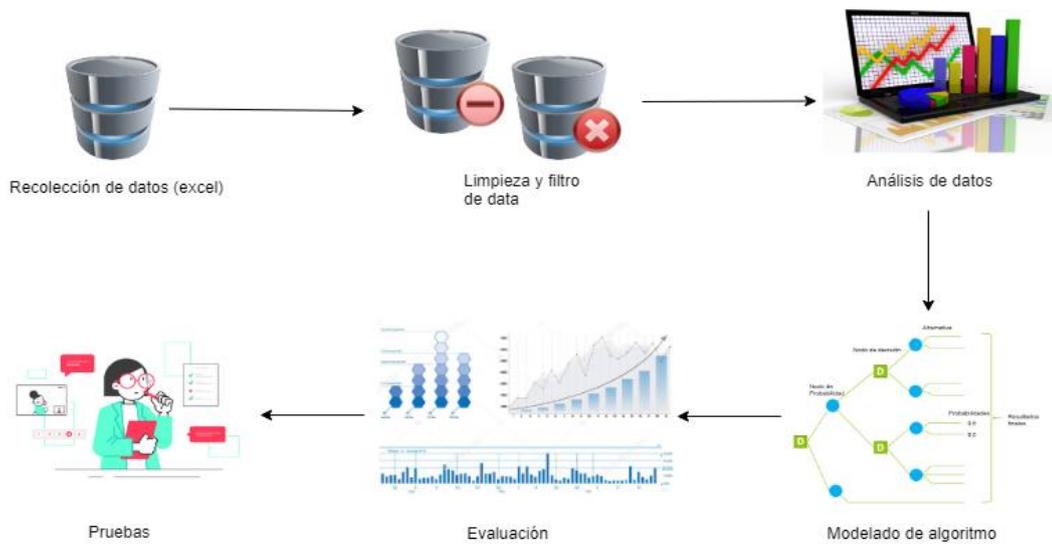


Figura 1. Arquitectura de solución.

## 2.3. Análisis de datos

La metodología CRISP-DM proporciona un enfoque estructurado para planificar un trabajo de minería de datos. Es una metodología robusta y bien probada. No reclamamos ninguna propiedad sobre él. No lo inventamos. Sin embargo, somos evangelistas de su poderosa practicidad, su flexibilidad y su utilidad al usar análisis para resolver problemas comerciales espinosos. Es el hilo dorado que atraviesa casi todos los compromisos del cliente.

Este modelo es una secuencia idealizada de eventos. En la práctica, muchas de las tareas se pueden realizar en un orden diferente y, a menudo, será necesario retroceder a tareas anteriores y repetir ciertas acciones. El modelo no intenta capturar todas las rutas posibles a través del proceso de minería de datos (Europe, 2018).

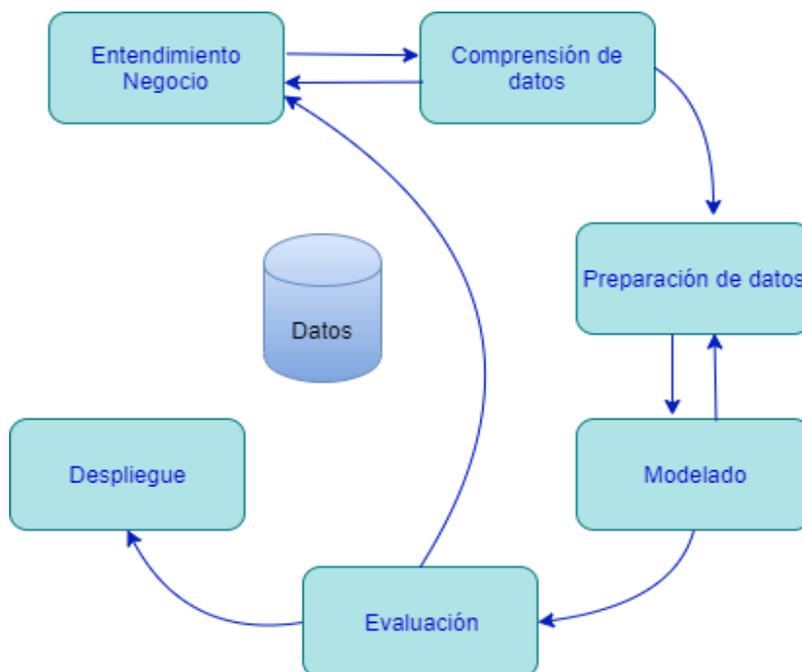


Figura 2. Metodología CRISP-DM.

- Entendimiento negocio
  - Comprender los objetivos y requisitos del proyecto; Definición del problema de minería de datos.
- Comprensión de datos
  - Recolección inicial de datos y familiarización; Identificar problemas de calidad de datos; Inicial, resultados obvios.
- Preparación de datos
  - Selección de registros y atributos; Limpieza de datos.
- Modelado
  - Ejecución de las herramientas.
  - Se separó la data en 80% para entrenamiento y 20% para evaluación, se determinó las variables predictoras y objetivo.
- Evaluación
  - Determinar si los resultados cumplen con los objetivos.
- Despliegue
  - Poner en práctica los modelos resultantes.

### **3. Resultados y Discusión**

Se realizó la investigación para el tratamiento masivo de datos y procesamiento con la ayuda de sistemas inteligentes. Se procesó diversos tipos de reclamos que se presentan en los servicios de telecomunicaciones, para poder ayudar a clasificar todos los reclamos registrados, es necesario usar un algoritmo para encontrar semejanzas entre los datos para poder clasificarlos.

De acuerdo a la data obtenida se pudo observar que en el año 2017 ocurrieron 3875 reclamos, y para el 2018 la tasa sube a 4223 reclamos por parte del usuario, y se ve que por año va incrementando, como también se observó que hay reclamos que demoran más tiempo en ser atendidos, y otros que son más comunes y se solucionan más rápido, lo que se quiere lograr es clasificar los reclamos para poder predecir semejanzas entre los datos, para ello usaremos algoritmos de aprendizaje supervisado (REUTERS, 2019).

Se desarrolló de la siguiente manera:

- Primero se obtuvo un conjunto 706 datos de reclamos de dos meses del año 2018 que fue más que suficiente para realizar la clasificación.
- Se realizó en pre procesamiento de datos para su normalización y adecuación.
- La selección del algoritmo de Machine Learning apropiados para el tipo de data obtenida y la problemática de resolución.
- Análisis de los resultados obtenidos en el desempeño y fiabilidad.

Después de ejecutar el modelo de clasificación en Azure Notebooks, se obtuvo el diagrama de árboles de decisión en base a lo ejecutado, que nos da facilidad para una mejor interpretación de la data obtenida.

#### **3.1. Resultados 1**

Entendimiento del negocio, en esta fase del proyecto se investigó el procedimiento que realiza un cliente a la hora de hacer un reclamo en servicios de telecomunicaciones.

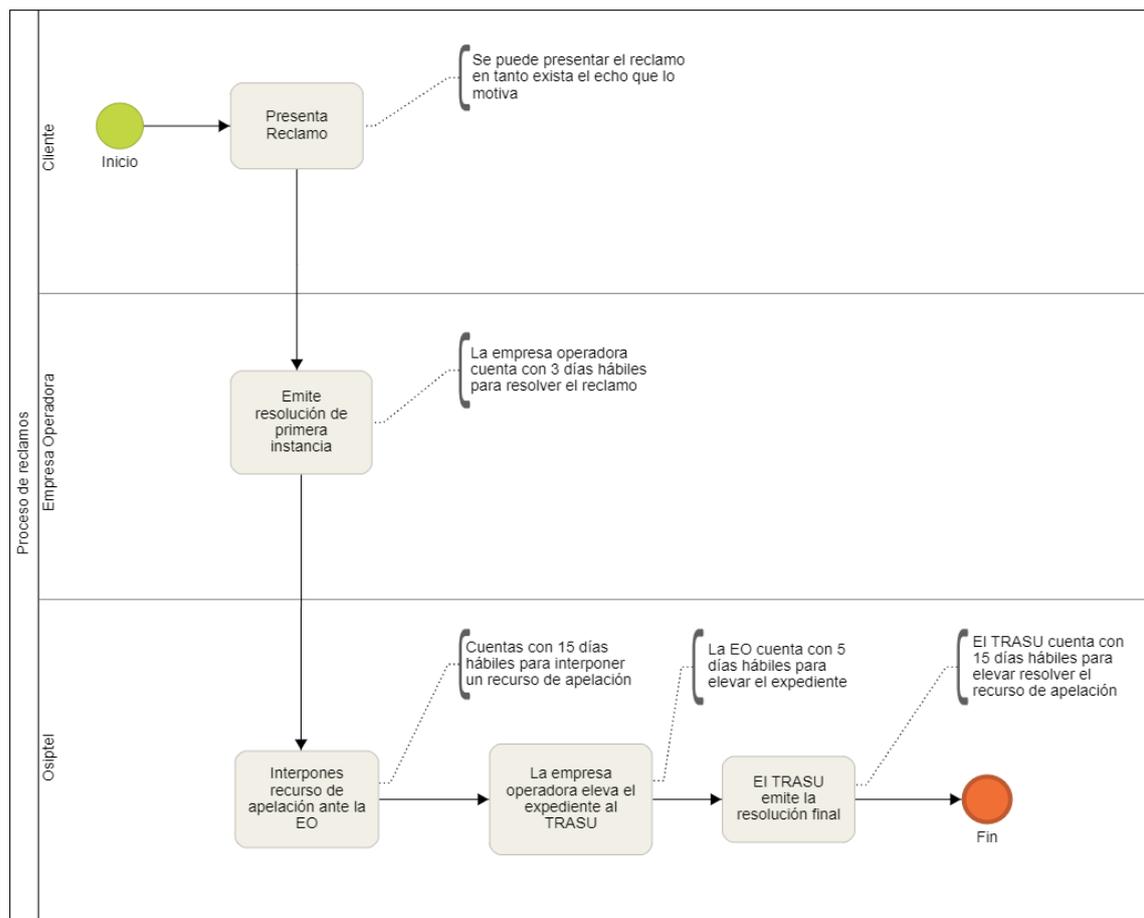


Figura 3. Procedimiento para realizar un reclamo.

### 3.2. Resultados 2

Comprensión y preparación de datos, en esta fase del proyecto se obtuvo una data de 706 registros, registros que tenían una variedad de campos, los cuales se fue escogiendo los features y targets para el entrenamiento del algoritmo, la data obtenida era data categórica, se preparó y se convirtió en data numérica.

Tabla 2  
Comprensión (Dataset)

...	Dirección	Empresa Operadora	Servicio	Tema	Detalle Servicio	...
...	Cercado de Juliaca	OSIPTEL	Telefonía Móvil	Facturación Renta fija	Postpago	...
...	Jr. Colonia Mz C Lt 20	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago	...
...	Juliaca Cercado	OSIPTEL	OSIPTEL	Información general Funciones del OSIPTEL		...
...	Jr. San Marcos Mz A Lote 3	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago Postpago	...
...	Jr. Huanday Mz P Lt 28	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago	...
...	Av. Mariano Nuñez Nro. 1030	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago	...
...	Cercado de Juliaca	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago Postpago	...
...	Jr. Cesar Vallejo Mz N5 Lt 09 Ubr. Néstor Cáceres Velasquez	OSIPTEL	Telefonía Móvil	Procedimiento para consultar número de líneas móviles	Prepago	...
...	...	...	...	...	...	...

Tabla 3  
Preparación de datos (Dataset)

EO	Servicio	Tema	Prioridad
11	5	4	Bajo
11	5	2	Bajo
11	5	13	Medio
12	5	1	Medio
11	5	1	Medio
12	5	1	Medio
11	5	11	Medio
11	5	5	Medio
11	5	1	Medio
1	5	5	Medio
11	5	4	Bajo
11	5	3	Medio
10	2	6	Bajo
11	5	3	Medio
11	5	2	Bajo
11	5	2	Bajo
1	5	13	Medio
10	2	6	Bajo
11	5	5	Medio
1	5	1	Medio
11	5	3	Medio
11	5	1	Medio
11	5	11	Medio
11	5	5	Medio
12	5	5	Medio
11	5	3	Medio
11	5	3	Medio

Tabla 4  
Preparación de datos (Dataset)

Empresa Operativa	Servicio	Tema	
AMERICA MOVIL PERU S.A.C.	1 Internet	1 Mensaje bloqueo IMEI adulterado	1
AMERICATEL PERÚ S.A.	2 OSIPTEL	2 Solicita información sobre situación de recurso	2
CABLE ESTACION S.R.L.	3 Servicios Empaquetados	3 Penalidades	3
CABLE MAS S.A.C.	4 Telefonía Fija	4 Charlas	4
DIRECTV PERU S.R.L	5 Telefonía Móvil	5 IMEI si figura en lista de codigos invalidos	5
ENTEL PERÚ S.A. (antes	6 Televisión por cable	6 Funciones del OSIPTEL	6
INDECOPI	7 Telefonía Pública	7 Consulta sobre significado de la respuesta	7
MINISTERIO DE	8 Otros servicios	8 Empresa no responde a reclamo	8
ORGANISMO SUPERVISOR DE LA INVERSIÓN EN ENERGÍA Y MINERÍA (OSINERGMIN)	9	Migración pendiente	9
OSIPTEL	10	Portabilidad no solicitada por abonado	10
TELEFÓNICA DEL PERÚ S.A.A.	11	Solicitud de activacion no atendida	11
VIETTEL PERU S.A.C.	12	Procedimiento para consultar número de líneas móviles	12
		Otras transgresiones normativas	13

### 3.3. Resultados 3

Modelado, evaluación y despliegue, en esta fase del proyecto se desarrolla el algoritmo de clasificación.

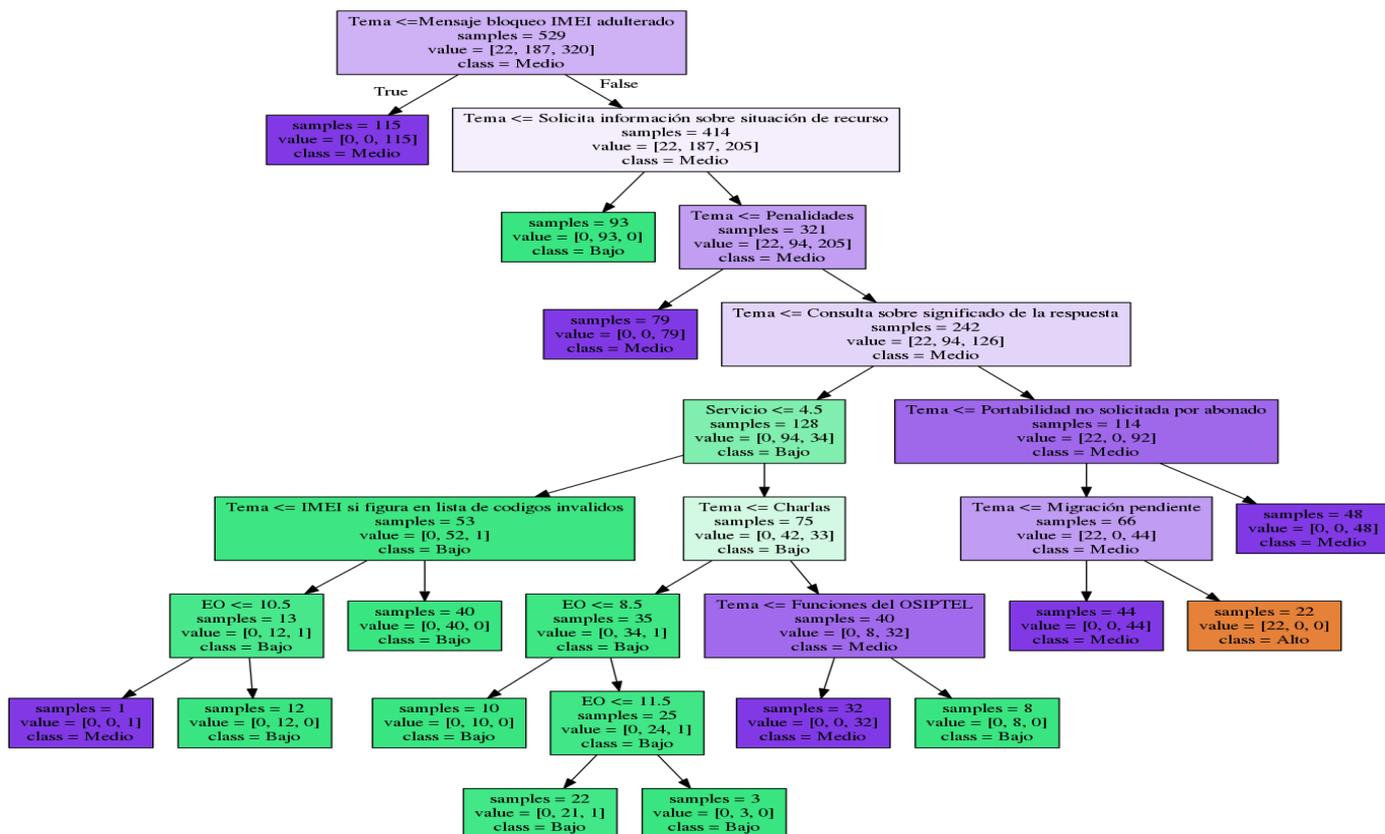


Figura 4. Modelado, Evaluación y Despliegue (Diagrama de árbol de decisión)

En el diagrama obtenido se muestra la prioridad por colores y sus derivados, los colores que sobresalen son el color anaranjado que representa a los reclamos que tienen prioridad ALTO, los de color morado y sus derivados representan los reclamos de prioridad MEDIO y los de color verde y derivados son de prioridad BAJO.

Tabla 5  
Informe de clasificación

	precision	recall	f1-score	support
Alto	1.00	1.00	1.00	6
Bajo	0.97	1.00	0.98	58
Medio	1.00	0.98	0.99	113
micro avg	0.99	0.99	0.99	177
macro avg	0.99	0.99	0.99	177
weighted avg	0.99	0.99	0.99	177

```
In [73]: arbol.score(X_train, y_train)
Out[73]: 0.8979206049149339
```

Figura 5. Porcentaje de precisión del Árbol de decisión.

#### 4. Conclusiones

Se logró obtener una buena cantidad de registros históricos para el procesamiento de la base de datos y un número considerable de variables, los resultados mostrados en el diagrama de árboles de decisión de modo que es más fácil la interpretación de resultados.

Con los resultados obtenidos, se condujo al moderado adicional con el entrenamiento del algoritmo para la mejora de la clasificación de los reclamos futuros que se obtendrán en el área de telecomunicaciones.

#### Agradecimientos

Agradecemos a la empresa OSIPTEL por brindarnos la DATA para realizar la investigación.

#### Referencias

- Barrientos, F., & Ríos, S. A. (2013). *Aplicación de minería de datos para predecir fuga de clientes en la industria de las telecomunicaciones*. *Revista Ingeniería de Sistemas*, XXVII, 73–107.
- Barrientos Inostroza. (2012). *Diseño e implementación de una metodología de predicción de fuga de clientes en una compañía de telecomunicaciones*. *OCDE Revisiones de Recursos Escolares : Chile 2017*, 10–13. <https://doi.org/10.1787/9789264287112-3-es>
- Contreras Morales, E. F., Ferreira Correa, F. M., & Valle, M. A. (2017). *Diseño De Un Modelo Predictivo De Fuga De Clientes Utilizando Árboles De Decisión*. *Revista Ingeniería Industrial*, 16(1), 07–23. <https://doi.org/10.22320/s07179103/2017.01>
- Elaidi, H., Elhaddar, Y., Benabbou, Z., & Abbar, H. (2018). *An idea of a clustering algorithm using support vector machines based on binary decision tree*. *2018 International Conference on Intelligent Systems and Computer Vision, ISCV 2018, 2018-May(5)*, 1–5. <https://doi.org/10.1109/ISACV.2018.8354024>
- Huang, Y., Huang, B. Q., & Kechadi, M. T. (2010). *A new filter feature selection approach for customer churn prediction in telecommunications*. *IEEM2010 - IEEE International Conference on Industrial Engineering and Engineering Management*, (2), 338–342. <https://doi.org/10.1109/IEEM.2010.5674306>
- Lakshminarayanan, B. (2016). *Decision Trees and Forests : A Probabilistic Perspective*.
- Langley, P., & Carbonell, J. G. (1984). *Approaches to machine learning*. In *Journal of the American Society for Information Science* (Vol. 35). <https://doi.org/10.1002/asi.4630350509>
- Meza, A. richard. (2018). *Predicción de fuga de clientes en una empresa de telefonía utilizando el algoritmo adaboost desbalanceado y la regresión logística asimétrica*. Retrieved from <http://repositorio.lamolina.edu.pe/bitstream/handle/UNALM/3245/meza-rodriguez-aldo-richard.pdf?sequence=1&isAllowed=y>
- EC. (2018). *El comercio*. Obtenido de *El comercio: 2018*
- Europe. (2018). *sv-europe*. Obtenido de *sv-europe*: <https://www.sv-europe.com/crisp-dm-methodology/>
- Perú21. (23 de 01 de 2019). *Perú 21*. Obtenido de *Perú 21*: <https://peru21.pe/peru/osiptel-reclamo-mal-servicio-telefono-celular-cable-internet-nnda-nnlt-447532-noticia/>
- REUTERS. (25 de 03 de 2019). *RPP*. Obtenido de *RPP*: <https://rpp.pe/economia/economia/cual-es-la-empresa-de-telefonía-movil-con-mas-reclamos-en-el-peru-noticia-1188059>
- De Araujo, F. H. D., Santana, A. M., & De Alcantara Dos Santos Neto, P. (2015). *Evaluation of classifiers based on decision tree for learning medical claim process*. *IEEE Latin America Transactions*, 13(1), 299–306. <https://doi.org/10.1109/TLA.2015.7040662>
- Hope, C. (2018). *How to create a CSV file*. Retrieved November 25, 2019, from <https://www.computerhope.com/issues/ch001356.htm>
- John Hunter. (2012). *Matplotlib: Python plotting — Matplotlib 3.1.1 documentation*. Retrieved November 25, 2019, from <https://matplotlib.org/>
- Navlani, A. (2018). *Clasificación del árbol de decisión en Python (artículo) - DataCamp*. Retrieved November 25, 2019, from <https://www.datacamp.com/community/tutorials/decision-tree-classification-python>
- Nbhelp. (2019). *Microsoft Azure Notebooks*. Retrieved November 25, 2019, from

*<https://notebooks.azure.com/help/introduction>*  
*NumPy. (2019). NumPy — NumPy. Retrieved November 25, 2019, from <https://numpy.org/>*  
*Pandas Basics - Learn Python - Free Interactive Python Tutorial. (2019). Retrieved November 25, 2019, from [https://www.learnpython.org/es/Pandas Basics](https://www.learnpython.org/es/Pandas%20Basics)*  
*Saripalli, P., Tirumala, V., & Chimmad, A. (2017). Assessment of healthcare claims rejection risk using machine learning. 2017 IEEE 19th International Conference on E-Health Networking, Applications and Services, Healthcom 2017, 2017-Decem, 1–6. <https://doi.org/10.1109/HealthCom.2017.8210758>*  
*ThePSF. (2019). What is Python? Executive Summary | Python.org. Retrieved November 25, 2019, from <https://www.python.org/doc/essays/blurb/>*  
*Saripalli, Tirumala, & Chimmad, 2017*