UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO
CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

TEMA:

Análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador

AUTOR:

MOSCOSO MENDOZA, FRANCIS ARTURO

Trabajo de Titulación previo a la obtención del título de:

INGENIERO EN TELECOMUNICACIONES

TUTOR:

M. Sc. BASTIDAS CABRERA, TOMAS GASPAR

Guayaquil, Ecuador

02 de marzo de 2020

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO

CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

CERTIFICACIÓN

Certificamos que el presente trabajo fue realizado en su totalidad por **Francis Arturo Moscoso Mendoza**, como requerimiento parcial para la obtención del Grado Académico de **Ingeniero en Telecomunicaciones.**

TUTOR

_____

M. Sc. BASTIDAS CABRERA, TOMAS GASPAR

DIRECTOR DE LA CARRERA

_____

Heras Sánchez, Miguel Armando

Guayaquil, 2 de marzo del 2020

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO

CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

DECLARACIÓN DE RESPONSABILIDAD

Yo, Moscoso Mendoza, Francis Arturo

DECLARO QUE:

El Trabajo de Titulación: **"Análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador",** previa a la obtención del **Grado Académico de Ingeniero en telecomunicaciones,** 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.

Guayaquil, 2 de marzo del 2020

AUTOR

_____

Moscoso Mendoza, Francis Arturo

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO

CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

AUTORIZACIÓN

Yo, Moscoso Mendoza, Francis Arturo

Autorizo a la Universidad Católica de Santiago de Guayaquil, la **publicación** en la biblioteca de la institución del Trabajo de Titulación: **"Análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador",** cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

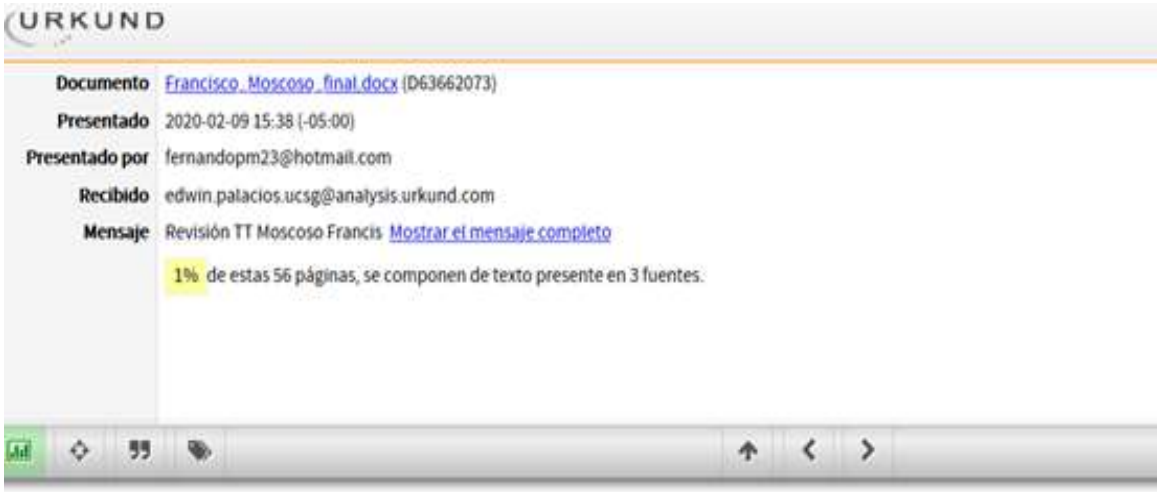
Guayaquil, 2 de marzo del 2020

AUTOR

_____

Moscoso Mendoza, Francis Arturo

REPORTE DE URKUND


URKUND

Documento	Francisco. Moscoso. final.docx (D63662073)
Presentado	2020-02-09 15:38 (-05:00)
Presentado por	fernandopm23@hotmail.com
Recibido	edwin.palacios.ucsg@analysis.urkund.com
Mensaje	Revisión TT Moscoso Francis Mostrar el mensaje completo

1% de estas 56 páginas, se componen de texto presente en 3 fuentes.

UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

TEMA: Análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador

AUTOR: Moscoso Mendoza, Francis Arturo.

Trabajo de Titulación previo a la obtención del título de: INGENIERO EN TELECOMUNICACIONES

TUTOR: XXXXXXXXX

Guayaquil, Ecuador

Xx de diciembre de 2019

UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

CERTIFICACIÓN

Certificamos que el presente trabajo fue realizado en su totalidad por Francis Arturo Moscoso Mendoza, como requerimiento parcial para la obtención del

Grado Académico

DEDICATORIA

Este trabajo de titulación se lo dedico a mis padres por haberme forjado como la persona que soy en la actualidad; muchos de mis logros se los debo a ustedes, incluyendo la presente. También se la dedico a mi novia que constantemente me ayudaba en lo anímico y académico para así poder culminar mi trabajo de titulación.

AUTOR

_____

Moscoso Mendoza, Francis Arturo

AGRADECIMIENTO

En primera instancia agradezco a mis formadores, personas de gran sabiduría quienes se han esforzado por ayudarme a llegar al punto en el que estoy ahora.

Sencillo no ha sido el proceso, por eso les agradezco por haberme transmitido sus conocimientos y dedicación, gracias a ellos he logrado importantes objetivos como culminar el desarrollo de mi trabajo de titulación y obtener así una digna carrera profesional.

AUTOR

Moscoso Mendoza, Francis Arturo

UNIVERSIDAD CATÓLICA

DE SANTIAGO DE GUAYAQUIL

FACULTAD DE EDUCACIÓN TÉCNICA PARA EL DESARROLLO

CARRERA DE INGENIERÍA EN TELECOMUNICACIONES

TRIBUNAL DE SUSTENTACIÓN

f._____

M. Sc. ROMERO PAZ, MANUEL DE JESÚS

DECANO

f._____

M. Sc. PALACIOS MELÉNDEZ, EDWIN FERNANDO

COORDINADOR DE ÁREA

f._____

M. Sc. ROMERO ROSERO, CARLOS BOLIVAR

OPONENTE

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

ÍNDICE DE ANEXOS

Resumen

El presente estudio se ha realizado para analizar el fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador, desde un enfoque que promueva el mejoramiento de la confianza en sus clientes y segmento de mercado. Sobre ello, se ha considerado que la mayoría de los sistemas de detección de fraude (FDS) monitorean los flujos de transacciones de tarjetas de crédito por medio de clasificadores que devuelven alertas para los pagos más riesgosos. La detección de fraude es notablemente un problema difícil debido al concepto *drift* (CD) (es decir, los hábitos de los clientes evolucionan) y el desequilibrio de clase (es decir, las transacciones genuinas superan ampliamente a los fraudes). Además, los FDS difieren de la clasificación convencional porque, en una primera fase, el investigador solo puede proporcionar un pequeño conjunto de muestras supervisadas que tienen tiempo para evaluar solo un número reducido de alertas. La metodología utilizada es cualitativa – cuantitativa, debido a que se cualifica la información proporcionada en el estudio, a través de una revisión de la literatura adecuada a la investigación y fundamentada con datos de la organización (CLARO) para conocer el escenario de fraude en sistemas de pago electrónico. Finalmente, los resultados se llevaron a cabo con el diseño de dos FDS sobre la base de un conjunto y un enfoque de ventana deslizante y se mostró que la estrategia ganadora consiste en entrenar a dos clasificadores separados (en retroalimentaciones y etiquetas retrasadas, respectivamente), y luego agregar los resultados.

Palabras clave: Telecomunicaciones, pago electrónico, fraude en sistema de pago, flujo de datos, anomalías, sector empresarial.

Abstract

The present study has been carried out to analyze fraud in electronic payment systems in the Ecuadorian telecommunications business sector, from an approach that promotes the improvement of trust in its customers and market segment. In this regard, it has been considered that most fraud detection systems (SDS) monitor the flows of credit card transactions through classifiers that return alerts for the riskiest payments. Fraud detection is remarkably a difficult problem due to the drift (CD) concept (that is, customer habits evolve) and class imbalance (that is, genuine transactions far outweigh frauds). In addition, the SDS differ from the conventional classification because, in the first phase, the researcher can only provide a small set of supervised samples that have time to evaluate only a small number of alerts. The methodology used is qualitative - quantitative, because the information provided in the study is qualified, through a review of the literature appropriate to the research and based on data from the organization (CLARO) to know the system fraud scenario Electronic payment Finally, the results were carried out with the design of two SDSs based on a set and a sliding window approach and it was shown that the winning strategy is to train two separate classifiers (in feedback and delayed labels, respectively), and then add the results.

Keywords: Telecommunications, electronic payment, payment system fraud, data flow, anomalies, business sector.

CAPÍTULO 1: INTRODUCCIÓN

1.1. Introducción

El rápido avance de la infraestructura tecnológica en telecomunicaciones e información global durante las últimas décadas, incluida la tecnología de la información y las redes informáticas (sistemas de Internet y telecomunicaciones) ha permitido el desarrollo del comercio electrónico a nivel mundial, lo que motiva a las empresas a interactuar más eficazmente con sus clientes y otras corporaciones dentro y fuera de sus industrias. Por tanto, el comercio electrónico integra servicios de telecomunicaciones, gestión de datos y seguridad, de modo que las aplicaciones comerciales proporcionan un intercambio de información rápido y flexible, con el fin de servir a los clientes y lograr la ventaja competitiva empresarial.

La industria en general, como muchos sectores empresariales, utiliza la tecnología de la información y la comunicación (TIC) para ofrecer a sus clientes servicios de valor agregado y conveniencia. El sistema de banca electrónica para realizar pagos en línea facilita la interacción entre clientes e industrias, con el fin de facilitar muchos servicios a los consumidores. En consecuencia, el sistema de sistema de pago electrónico se denomina frecuentemente banca electrónica, banca virtual o banca en línea.

Es por ello, que, con muchas etiquetas reportadas en la literatura, todas denotan la utilización de la tecnología para telecomunicaciones, con el propósito de realizar transacciones de cobros y/o pagos a través de la web. Siendo así, la banca electrónica incluye la provisión de diversas actividades de pago electrónico, que virtualmente desde cualquier lugar, en cualquier momento fuera de las instalaciones físicas de los bancos o empresas, pueden hacer de manera general los clientes, uso de estas plataformas para realizar cobros y/o pagos determinados a sus actividades.

Sin embargo, este panorama de interacción tecnológica cambiante, desde la banca tradicional a la banca electrónica, ha traído consigo nuevos desafíos.

Dichos desafíos no solo están relacionados con la gestión empresarial, sino también con las autoridades de supervisión y regularidad nacionales e internacionales. Los principales desafíos surgen del aumento de las transacciones en línea desde cualquier latitud y de la dependencia de la tecnología para proporcionar servicios electrónicos a través de internet, con la seguridad necesaria desde una ubicación remota. Estos desafíos incluyen retos de regularidad, legalidad, operatividad, reputación y de seguridad.

La obtención de un entorno seguro de telecomunicaciones en la tecnología informática es la preocupación más importante para todas las organizaciones que brindan servicios a través de internet. La seguridad de las transacciones en línea es uno de los desafíos más importantes para el sector empresarial nacional e internacional, debido a que miles de millones de transacciones de datos financieros se realizan en línea todos los días, y los *cyberdelitos* se llevan a cabo de manera frecuente por piratas informáticos expertos mediante la manipulación del sistema de información en línea de cualquier institución, sea pública o privada.

Por tal razón, todas las empresas que realizan actividades a través de la web, deben considerar escenarios posibles de amenazas que pueden venir desde dentro o fuera del sistema, lo que pone en riesgo a la información y las transacciones de los clientes, donde los administradores empresariales de las plataformas de internet deben asegurarse que tengan las prácticas en línea apropiadas para garantizar la confidencialidad y cuidado de los datos de los clientes, así como la integridad de los sistemas de pagos y/o cobros electrónicos y las transacciones realizadas.

En síntesis, en esta investigación se presenta una situación actual sobre los riesgos existentes en los sistemas de pago electrónico en el sector empresarial de telecomunicaciones ecuatoriano, que además tiene repercusiones en las transacciones fuera del país, debido a que cualquier persona con acceso a internet podría vulnerar la seguridad de estos medios informáticos. Por tanto, se presentan aspectos relacionados con la seguridad informática, fraude y ataques de banca electrónica; abordando los motivos e

importancia de proteger la seguridad de los datos institucionales y de sus clientes, así como los tipos comunes de ataques a los cuales podría verse expuesta la empresa y/o institución.

1.2. Antecedentes

La prevención del fraude electrónico en sistemas de pago en línea, tiene sus inicios con el desarrollo de la tecnología, telecomunicaciones e internet en general, debido a que los modelos de negocio de banca electrónica comenzaron en la década de 1980 en Nueva York, donde fue ofrecido por los principales bancos de la ciudad como Citibank y Chase Manhattan (Santos, 2018). Este no era un servicio completo de transacciones bancarias, sino un conjunto muy básico de servicios, como ver extractos bancarios y pagar facturas y/o cobranzas en línea.

Sin embargo, a través de ello, se allanó el camino para los servicios de banca electrónica más completos y sofisticados que existen actualmente y de los cuales las empresas han integrado en sus plataformas en línea, para brindar un servicio personalizado mediante internet, con el cual sus clientes pueden realizar pagos o diversas transacciones en línea. Es por ello, que desde sus inicios en la banca electrónica, se presentó el avance empresarial para el uso de cajeros automáticos y transacciones telefónicas, que en los últimos tiempos, ha facilitado las transacciones tanto para clientes como para empresas, mediante el uso de Internet, y ha permitido a las organizaciones escalar fronteras, cambiar sus tácticas estratégicas y abrir nuevas posibilidades comunicacionales entre ellos y sus clientes, no solo en sus lugares de desarrollo de actividades, sino en todo el mundo.

Las pioneras iniciales han sido las industrias bancarias, que en países de todo el mundo en este siglo XXI se han transformado o cambiado para operar mejor en el nuevo entorno complejo y competitivo (el entorno o plataforma electrónicos), en el que el clima económico también ha evolucionado. La tecnología de la información es el eje de estos importantes cambios, debido a que con la nueva era de la revolución tecnológica y de telecomunicaciones globales, las industrias y empresas en general, ahora son capaces de ofrecer

innumerables servicios, a través de medios electrónicos, a varios clientes, independientemente de su lugar, tiempo y distancia.

Es así que, en la década de 1990, las telecomunicaciones con el uso de internet evolucionaron cuando más personas poseían computadoras y se conectaron a la red, a través de su conexión de acceso telefónico desde cualquier lugar de sus hogares. Esta evolución tecnológica y la difusión del uso de internet en el hogar permitieron a los clientes disfrutar de servicios de banca electrónica, donde podían realizar sus solicitudes y/o pagos las 24 horas, los 7 días de la semana.

Pero los clientes durante la década de 1990 no confiaban completamente en el *e-banking* lo suficiente como para realizar transacciones monetarias serias y sustanciales. Esto desencadenó un gran esfuerzo e inversión por parte de los bancos para desarrollar más funciones de seguridad para sus servicios de banca en línea. A lo largo de la década de 2000, la banca en línea comenzó a crecer y a ser más aceptable para los clientes y así cubrir la mayor parte de la gama de servicios bancarios, que en la actualidad han sido democratizados hacia todo el sector empresarial, especialmente en telecomunicaciones.

Con ello, existen muchos tipos diferentes de servicios que pueden proporcionar las plataformas electrónicas de servicios empresariales. Los servicios populares incluyen cajeros automáticos, tarjetas de crédito, tarjetas de débito, tarjetas inteligentes, sistemas de transferencia electrónica de fondos, banca móvil, etc. En el nivel de transacción, la banca electrónica incluye acceso a la cuenta, transferencia de saldo, pago de facturas, presentación de facturas, préstamos hipotecarios, servicio al cliente y administración, venta cruzada, etc. Desde el punto de vista del banco, el uso de Internet ha reducido significativamente los costos físicos de las operaciones, incluidos los costos de procesamiento y transmisión de información de cliente hacia la empresa.

En síntesis, las plataformas web de servicios electrónicos, pueden verse como la extensión de las empresas físicas existentes a través de telecomunicaciones globalizadas, donde el uso de la informática desempeña un rol en la recuperación y procesamiento de datos, para iniciar transacciones directa y remotamente con una organización a través de redes de telecomunicaciones. Con ello, se aborda varias tendencias emergentes, como la demanda de los clientes de servicio en cualquier momento y en cualquier lugar, los imperativos del producto al mercado y los desafíos cada vez más complejos de integración para prevención de fraudes en línea, donde se ha centrado la presente investigación.

1.3. Planteamiento del problema

Las compañías de telecomunicaciones del sector empresarial nacional (e internacional) presentan un problema fundamental, referido a la seguridad de sus actividades en internet, que involucra a la relación comercial con los clientes. Este problema se centra de manera específica en las páginas web empresariales y demás derivados de ello, donde se presenta una opción de pago en línea, a través de lo cual, se ingresan los datos de los clientes sobre sus tarjetas de crédito y demás información financiera sensible, que es vulnerable al no existir un sistema de prevención de fraude para el pago electrónico.

Es por ello, que existe la necesidad de que el sector empresarial (especialmente las compañías de telecomunicaciones) observen sus sistemas de detección de fraude electrónico, para prevenir daños económicos, tanto para sí mismas como para sus clientes, lo cual puede generar una afectación irreparable a la confianza del consumidor en el ejercicio de sus transacciones en línea. Por tanto, para abordar el problema de investigación, se han definido las variables dependientes e independientes, sobre lo cual se establece:

- Variable dependiente: Sector empresarial de telecomunicaciones del Ecuador
- Variable independiente: Fraude en sistemas de pago electrónico

Siendo así, el presente estudio busca analizar los sistemas de detección de fraude en internet, donde el sector empresarial mantiene operaciones electrónicas que facilitan a sus clientes el acceso a información y pagos en línea. Razón por lo que, la problemática se fundamenta en la operatividad de las tecnologías informáticas que las empresas locales y nacionales puedan desempeñar para la protección de los datos internos y externos.

En consecuencia, para la ejecución de la investigación, se revisarán los datos sobre los sistemas de prevención de fraude electrónico, específicamente en referencia al pago electrónico de una empresa de telecomunicaciones que opera en el Ecuador y específicamente en la ciudad de Guayaquil, evaluando el sistema de detección de fraudes por suscripción utilizando redes de inteligencia artificial para su uso en el sector empresarial que utiliza plataformas de pago en línea, considerando que a nivel mundial, el desarrollo de la industria de las telecomunicaciones está aumentando rápidamente con una innovación para hacer frente a los retos del mercado y la evolución tecnológica actual.

1.4. Justificación

El presente trabajo de titulación se justifica debido a la gran importancia que día a día ejerce el sector de las telecomunicaciones en la población en general. El sector de las telecomunicaciones es un sector amplio con millones de usuarios en el Ecuador y el mundo, y por tanto esto expone la razón de su importancia para abordar una temática cada día más vigente en la vida de todas las personas que realizan actividades empresariales y/o comerciales a través de plataformas de pago electrónico.

Este sector se clasifica en términos generales en dos categorías según sus usuarios: usuarios domésticos y usuarios comerciales. Generalmente el uso de estas plataformas proporciona a los usuarios domésticos, conexiones a una tasa asequible, mientras que a los usuarios comerciales se le proporcionan conexiones a una tasa comparativamente más alta a medida que su escala de uso es mayor. Pero existen casos en que la suscripción a nivel comercial se incluye de manera fraudulenta, lo que causa una pérdida

significativa para el sector, lo cual presenta que este tipo de suscripciones, si se incluyeran en la categoría correcta, hubieran generado mayores ingresos para el sector empresarial de telecomunicaciones previsto.

Por tanto, sobre tal justificativo práctico y literario, el fraude se define como el acto deliberado y premeditado perpetrado para obtener ganancias en terreno falso, pero también se ve como cualquier transmisión de datos de voz a través de una red de telecomunicaciones, donde la intención del remitente es evitar o reducir los cargos legítimos. El fraude de telecomunicaciones es el robo de servicios o el abuso deliberado de redes de voz o datos (de personas naturales y/o jurídicas) que se encuentran normalmente encriptados en información de tarjetas bancarias y/o empresariales.

El fraude en las telecomunicaciones en el uso de información y datos personales transmitidos a través de plataformas electrónicas se ha abordado desde el punto de vista del fraude de suscripción, constituyendo un fraude contractual. En este tipo de fraude, los ingresos se generan a través del uso normal de un servicio sin tener que pagar; en este escenario, el estafador opera a nivel de números telefónicos donde todas las transacciones de este número son fraudulentas y todas las actividades en tales casos son aún más anormales a lo largo del período activo de la cuenta. En consecuencia, el fraude de suscripción se puede dividir en dos categorías a) fraude de suscripción para uso personal por parte del estafador, y b) fraude de suscripción con fines de lucro.

En síntesis, los problemas de detección de fraude se encuentran en muchos sectores de la vida y especialmente en el sector de las telecomunicaciones no es una excepción. Por lo tanto, la detección de fraude para este estudio de análisis de fraude en medios de pago electrónico se presenta como un intento de descubrir el uso ilegítimo de una red de comunicación al identificar el fraude lo más rápido posible una vez que se ha cometido, y con ello llegar a prevenir una afectación mayor, tanto para el sector empresarial local y nacional.

1.5. Objetivos

1.5.1. Objetivo general

Analizar el modelo de prevención de fraude en sistemas de pago electrónico para el sector empresarial de telecomunicaciones de una empresa local.

1.5.2. Objetivos específicos

- Fundamentar los preceptos literarios y normativos que prevenga el fraude en sistemas de pago electrónico para empresas de telecomunicaciones.
- Evaluar una configuración realista de detección de fraude como medio de manejo de las etiquetas de prevención por parte de la empresa.
- Diseñar dos FDS sobre la base de un enfoque conjunto y de ventana deslizante como estrategia para la prevención de fraude en medios de pago electrónico.

1.6. Hipótesis

El sistema de prevención de fraude en sistemas de pago electrónico ayudaría al sector empresarial ecuatoriano de telecomunicaciones al mejoramiento de la confianza en sus clientes y mercado.

1.7. Metodología de la investigación

La metodología a utilizar en la presente investigación es descriptiva, diseñada para representar a los participantes del proceso de recolección de información de una manera precisa. Es decir, se trata de describir al sector de telecomunicaciones y su incidencia en la prevención del fraude en sistemas de pago electrónico. El alcance de estudio es transversal, debido a que se realiza mediante datos recopilados en un periodo determinado de tiempo en una empresa de telecomunicaciones local, por tanto, el estudio, se ha de desarrollar en la empresa Claro de la ciudad de Guayaquil, para la recolección de información y posterior procesamiento que lleve a su consecución.

El propósito de tener una mejor perspectiva de la situación actual, donde las empresas de telecomunicaciones y los clientes, se encuentran expuestos a un riesgo en la seguridad de datos proporcionados al momento de ser ingresados y/o encriptados en las plataformas respectivas de las instituciones corporativas de la ciudad y del país, lo cual no solo se limita a este sector geográfico, sino que a través de internet tiene un alcance nacional e internacional.

CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA

2.1. Bases teóricas

La prevención de fraudes en sistemas de pago electrónico en el sector empresarial, exponen la perspectiva global de la importancia del comercio electrónico como un sistema complejo que involucra tecnología y seres humanos, razón por lo que, se ha convertido en una herramienta fundamental para las telecomunicaciones e intercambio de datos en línea en el Ecuador y el mundo interconectado (Martínez, 2015). La tecnología como producto del conocimiento tiene vulnerabilidades, por lo que su desarrollo se lleva a cabo continuamente; por ello, la investigación del comercio electrónico utilizada por los clientes se relaciona principalmente con aspectos psicológicos como la percepción, la confianza y la conveniencia, principalmente en transacciones electrónicas (a través de internet).

El presente capítulo tiene como objetivo fundamentar los preceptos literarios y normativos para la prevención de fraude en sistemas de pago electrónico en el Ecuador (Arahuetes & Sequera, 2015). Por tanto, inicia discutiendo el comercio electrónico en general, involucrando su definición y concepto, así como también cómo se realizan las transacciones a través del comercio electrónico en prevención de fraudes en plataformas web, sistemas de pagos electrónicos y demás transacciones bidireccionales entre el cliente y la empresa, que son fundamentales para poder comprender la importancia de la prevención en este tipo de relación comercial en línea, que en la actualidad es de uso frecuente para todas las instituciones y personas (naturales y/o jurídicas).

Asimismo, se revisan otras teorías y construcciones, que están adaptadas a la investigación para definir un modelo teórico – práctico que permita analizar el modelo de prevención de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones. Estas construcciones y teorías incluyen confianza, fraude, riesgo percibido, miedo al delito y teoría del comportamiento planificado (Buku & Mazer, 2017). Además, la discusión sobre la confianza se centra en la confianza del vendedor y la confianza del

medio (Internet). Mientras tanto, la discusión sobre el fraude se centra en las percepciones de fraude cibernético que se derivan de la noción de riesgo percibido y miedo al delito que son abordados por el sector empresarial en telecomunicaciones, específicamente en sus sistemas de pago.

2.2. Prevención de fraude en comercio electrónico

La literatura sobre sistemas de información ha definido el comercio electrónico (*e-commerce*) desde diferentes perspectivas (Ortega & Ramos, 2016). Existen al menos cinco perspectivas que se pueden aplicar para definir el comercio electrónico, incluido el proceso comercial (especialmente en sistemas de pagos), el servicio, el aprendizaje, la colaboración y la comunidad (Arqué, 2014). El comercio electrónico se centra en las transacciones comerciales entre organizaciones y personas que se realizan utilizando tecnología digital. La mayoría de las transacciones que utilizan el comercio electrónico se realizan a través de Internet.

Por tanto, el comercio electrónico puede definirse como el uso de tecnología de la información, que incluye Internet, computadora y otros dispositivos electrónicos, para comprar, vender, transferir e intercambiar productos, servicios o información (Ibáñez, 2018). En relación a ello, se ha categorizado según las partes involucradas en las transacciones y en este sentido, se pueden resumir en negocio a consumidor (B2C), negocio a negocio (B2B), consumidor a consumidor (C2C), *peer-to-peer* (P2P), comercio móvil (*M-commerce*) y empresa a gobierno (B2G). Entre esas formas de comercio electrónico, la evidencia reciente muestra que B2C tiene un papel emergente en la economía digital y es la principal transacción involucrada entre los sistemas de pago electrónico entre el sector empresarial y sus clientes.

Por lo tanto, esta investigación se centra en las transacciones B2C, en sus operaciones, el comercio electrónico comprende infraestructura y otro tipo de soporte que incluye personas, políticas públicas, marketing y publicidad, servicios de soporte, así como asociaciones comerciales (Pérez, Pacheco, & Salazar, 2016). Todos los componentes trabajan juntos e influyen entre sí, y

como consecuencia, las empresas que aplican el comercio electrónico en sistemas de pago electrónico deben crear una estrategia para organizar sus recursos y con ello prevenir fraudes. Por lo tanto, el comercio electrónico implica buenas prácticas de gestión porque funciona como un subsistema dentro del sistema de comercialización y venta de una organización.

En consecuencia, la prevención de fraude en sistemas de pago es un elemento central para mantener la confianza entre el sector empresarial y sus clientes, promoviendo así seguridad y fiabilidad en las transacciones en línea o en uso de las plataformas web de las empresas que a través de ello, brindan servicios diversos a sus clientes (Fernández, Alonso, Sotomayor, Adib, & Schultze-Kraft, 2017). Actualmente existen muchas definiciones de fraude y actividades fraudulentas, por lo que principalmente en el mundo de las telecomunicaciones y tecnología, se ha tomado al fraude como el uso de la ocupación de uno para el enriquecimiento personal a través del mal uso deliberado o la aplicación incorrecta de los recursos o activos de la organización empleadora (Asociación de Examinadores de Fraude Certificados, 2018).

Casi todo el sistema tecnológico que involucra dinero y servicios puede verse comprometido por actos fraudulentos; por ejemplo, la tarjeta de crédito, las telecomunicaciones, el seguro de salud, el seguro de automóviles y el sistema de subastas en línea y en general el sistema de pagos/cobros electrónico, que pueden realizarse a través de las plataformas web de las organizaciones, por lo que, su seguridad es fundamental para mantener la confianza entre la empresa y sus clientes al momento de interactuar comercialmente (Roa, García, Frías, & Correa, 2017). Por lo tanto, los fraudes en estos sistemas de pago/cobro en línea, se consideran delitos cibernéticos, causando una gran cantidad de pérdidas financieras, razón por lo cual múltiples instituciones, en todos los tamaños y sectores empresariales, han buscado mecanismos y/o métodos que ayuden a incrementar la seguridad tanto para la empresa como para sus clientes al momento de realizar las transacciones electrónicas.

2.2.1. Detección de fraudes

El sistema de detección de fraude es importante en varios sectores o áreas esenciales y sensibles para que las empresas, de todos los tamaños y segmentos de mercado, puedan mantenerse a través del tiempo en el ejercicio comercial (Almagro, Urrutia, & August-Treppel, 2018). Por lo tanto, la detección de fraudes ha sido el tema de varias encuestas y artículos de revisión incluidas y/o fundamentadas en esta investigación (Cardenas & Rosales, 2017). Es por tal razón, que pueden basarse en temas como áreas de fraude, tipos de fraude, enfoques y técnicas de detección de fraude, detección de fraude realizada en diferentes áreas basadas en base de datos y técnicas estadísticas de consecución de fraude en uso de plataformas web de empresas.

Por ello, se estima que la detección de fraude ha utilizado técnicas inspiradas en la naturaleza. Las técnicas inspiradas en la naturaleza, como su nombre lo indica, son técnicas de inteligencia artificial que se fundamentan en el funcionamiento de los sistemas naturales (Mock & Lupini, 2017). Por ejemplo, la red neuronal está inspirada en el sistema nervioso central de un animal (particularmente el cerebro) que es capaz de aprender y reconocer, y a es a partir de ello, que los sistemas de seguridad utilizados por el sector empresarial (microempresas, pyme y grandes empresas), consideran esencial contar con un sistema eficiente, efectivo y real para salvaguardar datos e información bidireccional entre clientes y la organización.

Desde este aspecto, se presentan los diferentes tipos de fraudes, especialmente los más comunes con tarjetas de crédito, fraude de bancarrota, fraude de falsificación, fraude de robo, fraude de aplicaciones y fraude conductual (Cricco, 2017). Estos tipos de fraude se han desempeñado sobre técnicas apropiadas para combatirlos; tales como el emparejamiento sabio de pares, árboles de decisión, técnicas de agrupamiento, redes neuronales y algoritmos genéticos, verificación de dirección, verificación de sistemas biométricos, geolocalización, firma digital, identificación de dispositivos, detección de proxy, seguridad *tokens*, tarjetas inteligentes, etc.

En la misma área, la literatura ha analizado diferentes tipos de métodos que se utilizan para detectar el fraude con tarjetas de crédito. Estos sistemas de detección se clasifican en métodos supervisados y no supervisados basados en reglas y diferentes técnicas de extracción de datos utilizadas para detectar el fraude en sistemas de pagos/cobros u otros (Zayas, 2016). Por tanto, a partir de ello, la tecnología y telecomunicaciones presentan en la literatura, extensas fuentes sobre los tipos de fraude en los sistemas de pago de sectores empresariales, donde se promueve hacer una descripción estructurada e integral de la investigación sobre detección de fraude (Jiménez, 2018).

Esto se hace cubriendo los tipos de fraude, los enfoques de detección de fraude, las técnicas de detección de fraude, así como los problemas y desafíos de detección de fraude en las áreas identificadas: tarjeta de crédito, telecomunicaciones, seguros, subastas, pagos/cobros electrónicos, etc. (Yuste, 2016).

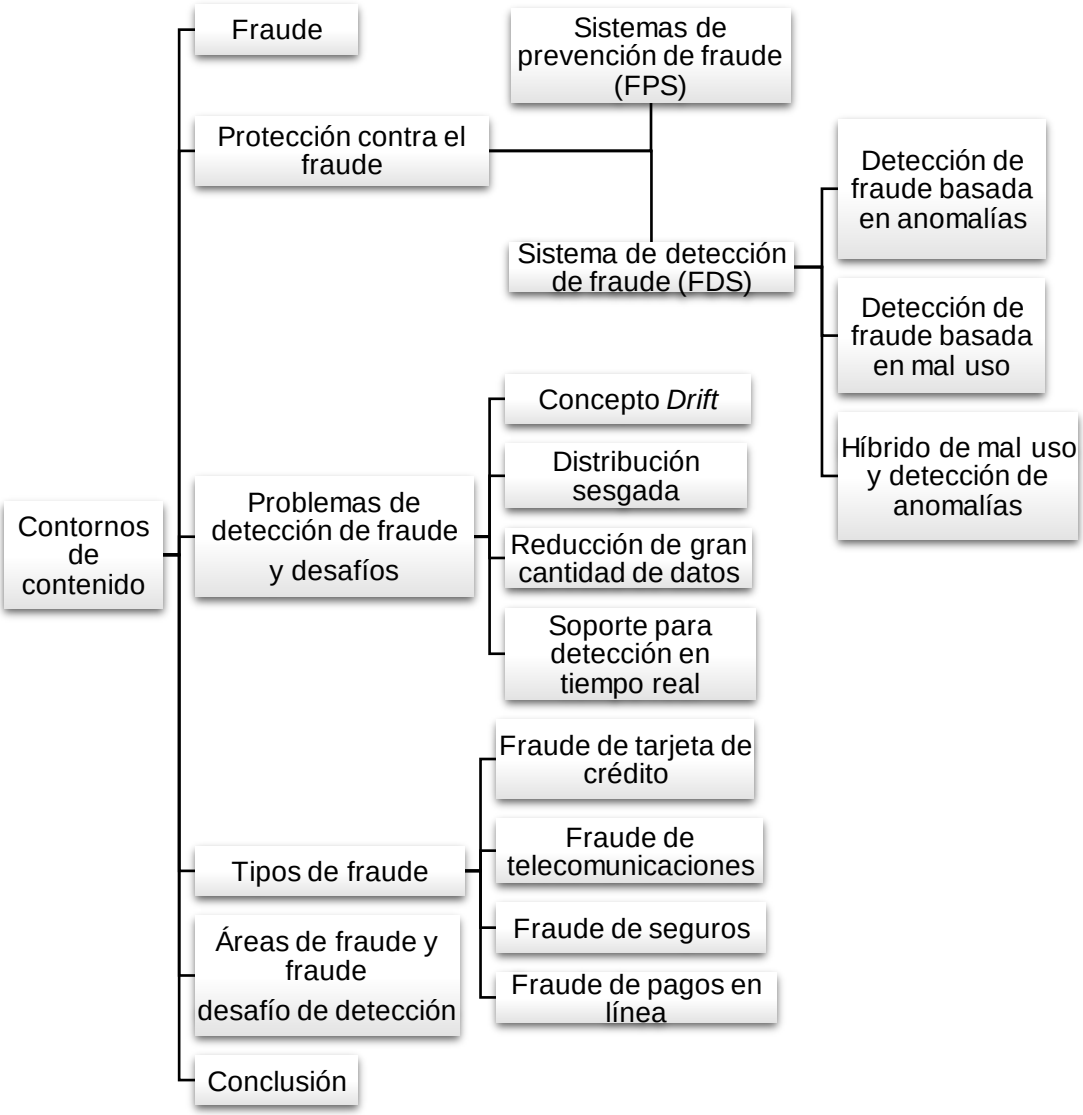
Tabla 2.1 Revisión del sistema de detección de fraude y artículos de revisión

Referencia	***Técnica***	***Área de fraude***
Mónica Jiménez, Encuesta Global Crimen Económico 2018, Fraude y delitos en línea	Los sistemas inteligentes: redes neuronales, inteligencia difusa, algoritmos genéticos, programación evolutiva, programación genética, estrategias de evolución y optimización de enjambres de partículas. Técnicas inspiradas en la naturaleza. Técnicas de minería de datos temporales espaciales.	Telecomunicaciones, seguros, auditoría, atención médica, transacciones con tarjeta de crédito, comercio electrónico, precios de oferta y verificación de identidad. Correo electrónico, spam, *phishing* e intrusión de redes. Seguro de salud.

Alberto Gallego Yuste, 2016, Delitos informáticos: Malware, Fraudes y Estafas a través de la red y cómo prevenirlos.	Métodos basados en reglas, supervisados y no supervisados. Minería de datos y técnicas estadísticas.	Redes de voz sobre IP (VoIP). Detección de fraude financiero. Seguro de hogar, seguro de vida, seguro de automóvil Fraude de tarjeta de crédito.
Leyre Zayas Mariscal, 2016, Señales de alerta para la detección de frauda en las empresas.	Tipos de fraude, enfoques de detección de fraude, técnicas de detección de fraude y problemas y desafíos de detección de fraude.	Fraude de tarjeta de crédito, fraude de telecomunicaciones, fraude de seguro médico, fraude de seguro de automóvil, fraude de subasta en línea

Elaborado por: Autor

Figura 2.1 Esquema de sistema de detección de fraude electrónico



Elaborado por: Autor

La figura 2.1 resume el contenido del esquema de sistema de detección de fraude electrónico y la tabla 2.1, expone estudios literarios sobre técnicas utilizadas y las áreas de fraude estudiadas que se adecuan a la prevención de fraude en sistemas de pago electrónico en el sector empresarial del Ecuador (y del mundo) (Franco, 2018). Es así que las bases de datos implican técnicas estadísticas, matemáticas, de inteligencia artificial y de aprendizaje automático para extraer e identificar información útil y el conocimiento posterior de grandes bases de datos (sistemas de soporte de decisiones y sistemas inteligentes), que son obtenidas de técnicas de detección y prevención de fraude en métodos de prueba de identidad para aplicaciones tecnológicas.

Tabla 2.2 Técnicas de detección y prevención de fraude

Técnicas de detección y prevención de fraude	*Métodos de prueba de identidad*	Sistemas de verificación de direcciones
		Verificación avanzada de direcciones
		Verificación de edad
		Esquemas de seguridad de la tarjeta
		Sistema de verificación de carga
		Sistema de verificación de historial
		Sistema de autenticación del consumidor
		Sistema de verificación de crédito
		Sistema de verificación de propiedad de cuenta
		Verificación de Email
		Autenticación de identidad electrónica
		Sistema de verificación fuera de la cuenta
		Sistema de búsqueda inversa
		Devolver correo electrónico
	Aplicaciones tecnológicas	Biometría
		Geolocalización
		Firma digital
		Identificación del dispositivo
		Detección de proxy
		Fichas seguras
		Tarjetas inteligentes
		Métodos para compartir datos

Fuente: (Herrera-Semenets & Prado, 2014)

Por tanto, las técnicas de detección y prevención de fraude tienen como objetivo mejorar la comprensión de los métodos de prueba de identidad y aplicaciones tecnológicas para la detección de fraude, con la intención de identificar qué problemas y desafíos se deben considerar para un sistema eficiente, de acuerdo a elementos concordantes para su prevención la fraude en sistemas de pago electrónico y demás actividades web (Salazar & Flores, 2016). Estos sistemas tienen varias ventajas principales:

a. El patrón de fraude se obtiene automáticamente de los datos;
b. Especificación de probabilidad de fraude para cada caso, en consecuencia, los esfuerzos en la investigación de casos sospechosos pueden ser priorizados; y
c. Revelación de nuevos tipos de fraude que no se definieron antes.

Los métodos de base de datos constan de categorías principales que son clasificadas en:

1. Agrupación,
2. Regresión,
3. Detección de valores atípicos,
4. Visualización y
5. Predicción

Cada uno de estos métodos está respaldado por técnicas específicas. Por ejemplo, la técnica de red neuronal y la técnica de máquina de vectores de soporte se utilizan para el método de clasificación de base de datos (Sánchez, 2019). La técnica *K-means* se utiliza para el método de agrupación; además, la base de datos ha incorporado muchas técnicas de otros dominios, como estadísticas, aprendizaje automático, reconocimiento de patrones, bases de datos y sistemas de almacenamiento de datos, recuperación de información, visualización, algoritmos, computación de alto rendimiento y muchos dominios de aplicación, sobre ello, recientemente, la detección de fraude integra un enfoque de detección basado en anomalías y un enfoque de detección basado en el mal uso mediante técnicas de minería de datos.

2.2.1.1. Detección de fraude basada en anomalías

El método de detección de anomalías o valores atípicos es utilizado por FDS y se basa en métodos de perfil de comportamiento, donde modela el patrón de comportamiento de cada individuo, monitoreándolo para detectar cualquier desviación de la norma (Vásconez, 2017). Numerosos autores de estudios en tecnología y telecomunicaciones adoptan el FDS basado en anomalías en diferentes áreas de fraude. Los FDS basados en anomalías tienen el potencial de detectar nuevos fraudes, por lo tanto, es utilizado principalmente por la literatura de FDS y este método puede clasificarse en tres tipos; detección de anomalías, semisupervisada, supervisada y sin supervisión.

Supervisada

Las técnicas de aprendizaje supervisado requieren un conjunto de datos que ha sido etiquetado como fraude y no fraude e implica la capacitación de un clasificador (Escobar, 2015). Este es el enfoque de aprendizaje más común, puesto que la principal ventaja del aprendizaje supervisado es que todas las salidas de clases manipuladas por el algoritmo de este enfoque son significativas para los humanos, y pueden usarse fácilmente para la clasificación de patrones discriminativos y la regresión de datos; sin embargo, el aprendizaje supervisado tiene varias limitaciones.

El primero es causado por la dificultad de recopilar supervisión o etiquetas; debido a que cuando hay un gran volumen de datos de entrada, es prohibitivamente costoso, si no imposible, etiquetarlos a todos. En segundo lugar, en algún momento es extremadamente difícil encontrar una etiqueta distintiva, existen incertidumbres y ambigüedades en la supervisión o las etiquetas. Estas limitaciones pueden obstruir la implementación de los enfoques de aprendizaje supervisado en algunos casos; por lo tanto, el aprendizaje no supervisado y el aprendizaje semisupervisado se utilizan para superar estas desventajas. El aprendizaje supervisado abarca muchos algoritmos que incluyen lo siguiente:

a. Algoritmos de clasificación

Por ejemplo, red neuronal artificial, vecinos más cercanos a *K*, árboles, regresión logística, *Naïve-Bayes* y técnicas de máquina de vectores de soporte (SVM).

b. Algoritmos de regresión

Por ejemplo, regresión lineal, regresión simple y regresión logística.

c. Sin supervisión

Las técnicas de aprendizaje no supervisadas detectan acciones fraudulentas en un conjunto de datos de prueba no etiquetados bajo el supuesto de que la mayoría de las instancias en el conjunto de datos no es fraude. A diferencia de la técnica supervisada, sin supervisión significa que no hay una etiqueta de clase para la construcción del modelo (García, Cantoral, Enríquez, & Navas, 2016). El principal beneficio de utilizar un enfoque no supervisado es que no se basa en una identificación precisa de los datos de la etiqueta, que a menudo es escasa o inexistente. Dos algoritmos clásicos simples empleados en el aprendizaje no supervisado son:

- Algoritmos de agrupamiento como técnicas de *K-means.*
- Algoritmos de reducción de dimensionalidad como: Análisis de componentes principales (PCA)
- Semisupervisado

El aprendizaje semisupervisado se encuentra entre el aprendizaje supervisado y el no supervisado, ya que involucra una pequeña cantidad de muestras etiquetadas y una gran cantidad de muestras no etiquetadas. El objetivo principal del enfoque semisupervisado es capacitar a un clasificador a partir de datos etiquetados y no etiquetados (Gilman & Joyce, 2014). Por tanto, tiene más ventajas en comparación con el aprendizaje supervisado porque logra un mejor rendimiento al utilizar datos etiquetados y no etiquetados, pero con menos instancias etiquetadas; además, el aprendizaje semisupervisado también proporciona un modelo computacional para comprender el aprendizaje de categoría humana, donde la mayor parte de la entrada es evidentemente no etiquetada.

Detección de fraude basada en mal uso

En el enfoque de detección de mal uso, sobre los cuales los comportamientos fraudulentos se definen primero mediante el uso de firmas de estafadores, y luego otros comportamientos se definen como comportamientos normales. El enfoque de mal uso adoptado por FDS utiliza estadísticas basadas en reglas o métodos heurísticos correspondientes para revelar la ocurrencia de transacciones sospechosas específicas.

La detección de mal uso es un sistema experto que se considera como un mecanismo de detección simple y rápida (Almagro, Urrutia, & August-Treppel, 2018). Pero tiene una limitación importante porque no es posible detectar todos los tipos diferentes de fraudes porque solo busca patrones conocidos de mal uso.

Híbrido de mal uso y detección de anomalías

Algunos investigadores han propuesto un enfoque híbrido en el que los modelos de detección de anomalías y mal uso se combinan para obtener resultados óptimos (Salazar & Flores, 2016). Esto se debe a la incapacidad de esa detección de mal uso para detectar nuevos fraudes; mientras tanto, la detección de anomalías adolece de la falta de capacidad de generalización y la presencia de altas tasas de falsas alarmas. Sin embargo, según la literatura, los FDS basados en anomalías son el enfoque más utilizado.

2.2.2. Problemas y desafíos de detección de fraude

La detección de fraude es un dominio complejo; donde se puede encontrar que un sistema de detección de fraude es propenso a fallar, tiene una baja tasa de precisión o da muchas falsas alarmas (Guillén & Biarge, 2019). Es extremadamente difícil para los sistemas de comercio electrónico manejar un problema de fraude que los obliga a incurrir en grandes pérdidas. Esto sucede debido a que los sistemas de detección de fraude deben enfrentar múltiples desafíos o retos a tener en cuenta, principalmente para el sector empresarial nacional.

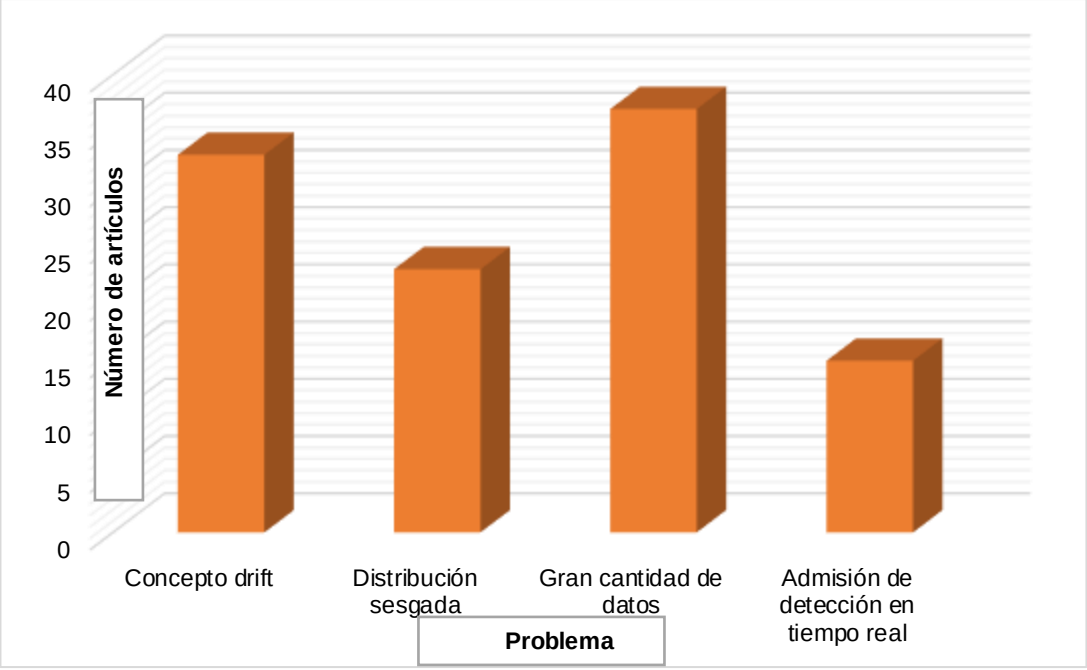


Figura 2.2 Distribución de artículos FDS basados en problemas y desafíos entre 1998 y 2018.

Fuente: (Cámara de Comercio de Guayaquil, 2019)

En esta sección se presentarán varias propiedades de desafíos que debe enfrentar la detección de fraude; por ello, la figura 2.2 muestra la distribución de artículos de Sistema de Detección de Fraude (FDS), por sus siglas en inglés, basados en problemas y desafíos (Arcenegui, Castilla, & Lozano, 2015). Las estadísticas se basan en el número de documentos publicados entre 1998 y 2018, en el sector empresarial, con un enfoque relacionado a las telecomunicaciones, donde se toma los tipos más frecuentes de fraudes electrónicos, tales como: tarjetas de crédito, telecomunicaciones, seguro médico, seguro de automóviles y fraudes de subastas en línea.

2.2.2.1. Concepto drift

Existen varias definiciones para el concepto *drift* en la literatura; por ello, en la minería de datos, se refiere al fenómeno de que el modelo (o concepto) subyacente está cambiando con el tiempo. Los FDS funcionan en un entorno dinámico donde el comportamiento del usuario legítimo o estafador cambia continuamente, y se denomina concepto de fenómeno derivado (Burke, 2016). Por ejemplo, en el área de tarjetas de crédito, el comportamiento del titular de la tarjeta puede estar sujeto a cambios debido a una variedad de causas

externas; tales como el monto y la frecuencia de la transacción están estrechamente relacionados con los hábitos de gasto de una persona que en realidad están influenciados por los ingresos, la disponibilidad de recursos y el estilo de vida de una persona, que pueden cambiar con el tiempo.

Además, los trucos del estafador evolucionan continuamente y la detección tiene que adaptarse a estos nuevos tipos de fraude. Asimismo, este aspecto se refiere principalmente a un escenario de aprendizaje supervisado en línea cuando la relación entre los datos de entrada y la variable objetivo cambia con el tiempo. Mientras que, en el aprendizaje supervisado, el objetivo es predecir una variable objetiva y dada un conjunto de características de entrada X (Riera & Ruano, 2016). En la instancia de entrenamiento que se usa para la construcción de modelos, tanto X como Y corresponden a los datos de entrada y la variable objetivo, respectivamente.

En la nueva instancia en la que se aplica el modelo predictivo, se conoce X, pero no se conoce Y en el momento de la predicción, y la relación entre los datos de entrada y la variable objetivo puede cambiar (Tracker & González, 2019). Por tanto, esto es una gran preocupación, particularmente en el aprendizaje en línea, donde el modelo de detección se actualiza inmediatamente, pero se basa en datos obsoletos, por lo que cuando llegan nuevos datos, el modelo puede ser engañoso y genera muchas falsas alarmas para la detección oportuna y eficiente de posibles ataques o fraudes a las plataformas informáticas.

Siendo así, se ha dedicado atención en la literatura para tratar el comportamiento no estacionario y actualizar dinámicamente el modelo de detección de fraude; sobre lo que se requiere el uso de algoritmos de aprendizaje adaptativo para manejar el problema presentado en un potencial fraude electrónico que puede afectar de manera importante los intereses de la organización como de sus clientes, socavando su fiabilidad y credibilidad en el ejercicio de actividades empresariales (González, Romero, Cruz, & Ortiz, 2018). Los algoritmos de aprendizaje adaptativo pueden verse como algoritmos avanzados de aprendizaje incremental que pueden actualizar el

modelo de detección para transmitir datos en evolución a lo largo del tiempo, sobre lo que se expresa la definición de aprendizaje incremental de la siguiente manera:

Proceso de aprendizaje incremental en cada momento t, donde hay datos históricos disponibles, llega una instancia objetivo X_{t+1}, la tarea es predecir la etiqueta Y_{t+1}, para eso, la persona L_t se construye en la fase de entrenamiento mediante el uso de todos o la selección de datos históricos etiquetados $X_{histórico}$ = ($X1, ..., Xt$). Esto es como se ilustra en la figura 2.3, mediante el proceso de aprendizaje incremental, la etiqueta y_{t+1} estará disponible con X_{t+1}, y será parte de la historia a predecir X_{t+2}.

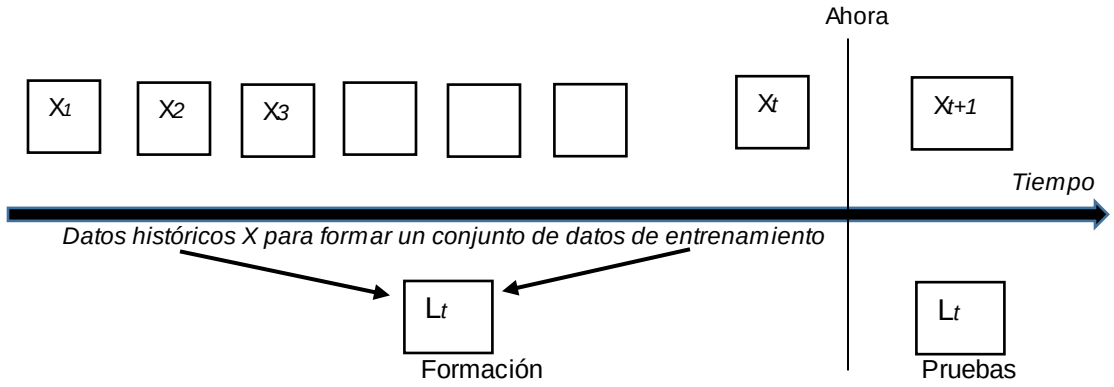


Figura 2.3 Aprendizaje incremental en el tiempo t.

Fuente: (Ruiz-Capillas & Fernández, 2014)

La taxonomía común que responden a los FDS adaptativos se clasifica en dos grupos según cuándo se activa la función adaptativa, esto puede basarse en la evolución o regularse como se muestra en la figura 2.4. El enfoque basado en la evolución es cuando el alumno adapta automáticamente su comportamiento para mantenerse actualizado con la dinámica de la transmisión (Porlles & Luján, 2017). Mientras tanto, el enfoque basado en la regulación es cuando el concepto *drift* y la clasificación se manejan como problemas separados. El concepto diseñado de detectores se marcará cuando se produzcan cambios, y luego se deben tomar algunas reacciones.

En general, la ventaja de la base regulada no solo se debe a su adaptación conceptual, sino también a proporcionar la información necesaria sobre ello; por ejemplo, cuando FDS detecta un fraude con tarjeta de crédito, se requiere

que tome más medidas de investigación sobre el comportamiento de los estafadores. El método basado en regla no utiliza con frecuencia un método basado en evolución para manejar el concepto en el área de fraude.

2.2.2.2. Distribución sesgada de clases

La distribución sesgada (clase desequilibrada) se considera uno de los problemas más críticos que enfrenta el FDS. En general, el problema de la clase desequilibrada es la situación en la que hay muchas menos muestras de instancia fraudulenta que la instancia normal (González & Morán, 2017). En un enfoque de aprendizaje supervisado, el problema del desequilibrio de clase ocurre cuando la clase minoritaria es muy pequeña, lo que lleva a numerosos problemas, como la discapacidad de los alumnos para descubrir patrones en los datos de la clase minoritaria. Además, la clase de desequilibrio tiene un grave impacto en el rendimiento de los clasificadores que tienden a ser abrumados por la clase mayoritaria e ignoran a la clase minoritaria.

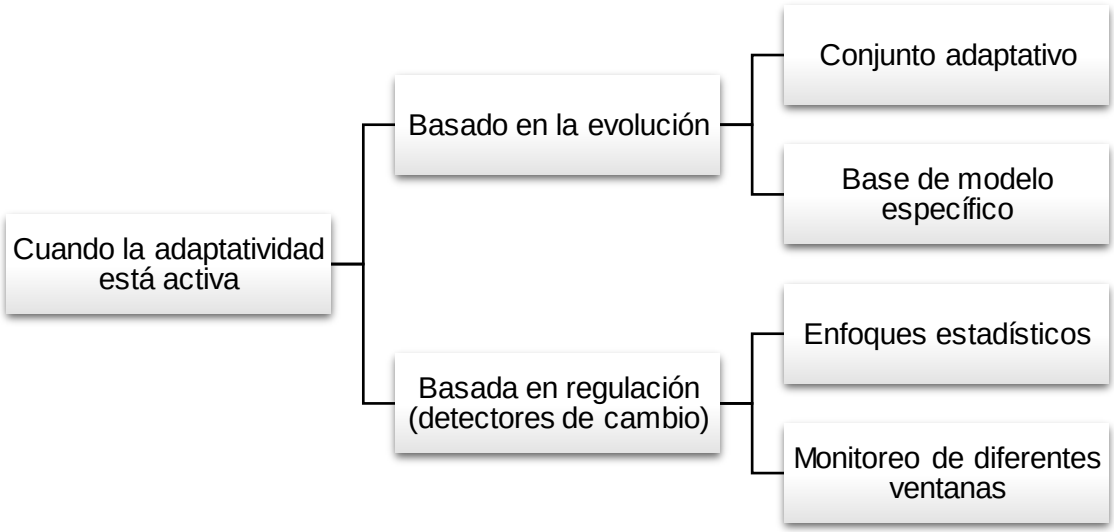


Figura 2.4 Enfoques de aprendizaje adaptativo

Fuente: (Buku & Mazer, 2017)

Para aclarar, el conjunto de datos presentado por la Cámara de Comercio de Guayaquil en la base de datos utilizó el problema de la clase desequilibrada, como un conjunto de datos real de transacciones de comercio electrónico que se utiliza para detectar transacciones anómalas (tal como se presenta en la figura 2.2 El conjunto de datos de capacitación contiene

100,000 transacciones de 73,729 clientes que abarcan un período de 98 días (del año 2018); y el conjunto de prueba consta de 50,000 transacciones (Cámara de Comercio de Guayaquil, 2019).

Los datos están altamente desequilibrados, que consisten en 97,346 transacciones normales (clase mayoritaria) y solo 2654 transacciones fraudulentas (clase minoritaria), como se muestra en la figura 2.5 el porcentaje de transacciones normales (clase mayoritaria) es de alrededor del 97,3% fraudulento transacciones (clase minoritaria). Por lo tanto, se requiere un mecanismo de equilibrio para equilibrar estos datos con una relación de 1:1 entre la clase normal y la fraudulenta para manejar el desequilibrio de clase.

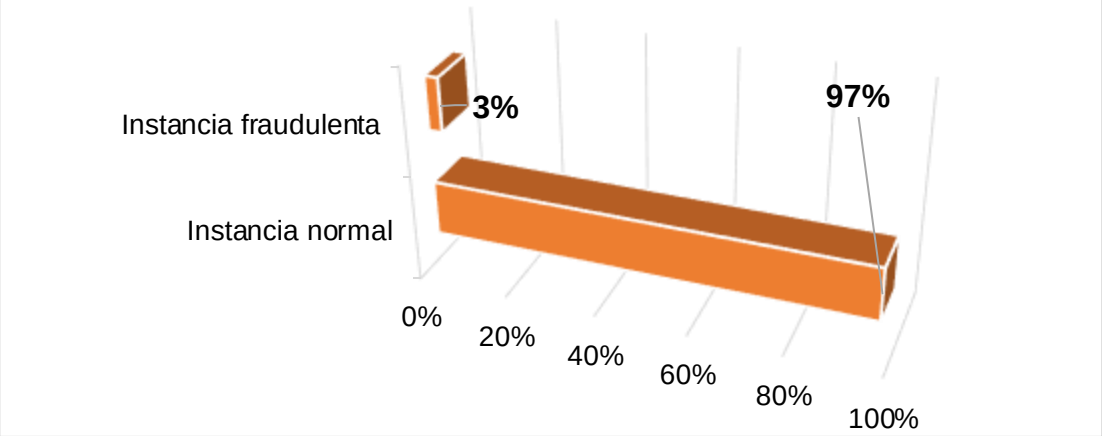


Figura 2.5 Desequilibrio del conjunto de datos

Fuente: (Cámara de Comercio de Guayaquil, 2019)

Esto hará que el fraude sea extremadamente fácil de detectar, porque la diferencia entre muestras minoritarias y de clase mayoritaria se puede reconocer de manera efectiva. Por tanto, los enfoques de equilibrio de datos se pueden clasificar en dos niveles diferentes, nivel de datos y nivel algorítmico, para poder estimar de manera adecuada el estado y/o posibilidad del fraude en la transacción, tal como se pudo apreciar en la figura 6 en referencia a los enfoques y técnicas equilibrados.

Métodos de nivel de datos

Las técnicas de equilibrio a nivel de datos se utilizan como un paso de

preprocesamiento para reequilibrar el conjunto de datos o eliminar el ruido antes de la aplicación de otros algoritmos de clasificación. En la literatura de FDS, la mayoría de los investigadores emplean técnicas de equilibrio de nivel de datos, por ejemplo, enfoques de submuestreo o sobremuestreo (Espino, Fontes, & Daradoumis, 2017). El enfoque de muestreo bajo elimina parte de los datos en la clase mayoritaria, por tanto, una gran cantidad de los sistemas de detección de fraude propuestos utilizaron el enfoque de submuestreo para equilibrar los datos de capacitación sobre la prevención de fraude en sistemas de pago.

El enfoque de sobremuestreo replica los datos en la clase minoritaria y rara vez se usa porque causa un sobreajuste de un modelo, particularmente con la existencia de datos ruidosos (Donayre & Santos, 2018). Además, el muestreo excesivo no da como resultado que se incluya más información en el conjunto de capacitación, lo que conduce a un modelo muy complejo. Alternativamente, SMOTE (Técnica de sobremuestreo de minorías sintéticas) se utiliza para la detección de fraude electrónico en sistemas, que es la mejor alternativa al enfoque de sobremuestreo para la utilización mayoritaria en el sector empresarial (de telecomunicaciones).

Por tanto, para la ejecución efectiva de un sistema de prevención de fraude electrónico en el sector empresarial, SMOTE sobremuestra la clase minoritaria generando ejemplos minoritarios sintéticos en la vecindad de los elementos y/o datos observados (Tafra, 2016). Por ello, generalmente en este ámbito (sistemas de pagos electrónicos) se realizan experimentos sobre diferentes tipos de técnicas de nivel de datos equilibrado para descubrir la técnica o métodos (antes descritos) más eficiente para el sistema de detección de fraude de en sistemas de cobros y/o pagos, ya sea a través de la utilización en línea de tarjetas de crédito/debito u otros sistemas de pago que pueden ser vulnerados en las plataformas web de los programas informáticos de las empresas en general.

Métodos de nivel algorítmico

Este tipo de algoritmo de clasificación se ocupa de la clase fraudulenta. El método de nivel algorítmico empleado más frecuente se fundamenta sobre un aprendizaje sensible al costo para lidiar con la distribución sesgada de la clase (Roa, García, Frías, & Correa, 2017). El aprendizaje sensible al costo pone una variable de costo a la clasificación errónea de las diferentes clases al suponer que hay una matriz de costos disponible para los diferentes tipos de errores en matriz de costos creada para sesgar el modelo para minimizar el costo o maximizar el beneficio. En la formulación de la matriz de costos, los costos están asociados con esas predicciones, tales como:

a) Falso negativo (si la etiqueta verdadera es fraude y se clasifica como normal),
b) Falso positivo (si la etiqueta verdadera es normal y se clasifica como fraude),
c) Verdadero negativo (si la etiqueta verdadera es normal y se clasifica como normal) y,
d) Verdadera positiva (si la etiqueta verdadera es fraude y el FDS se clasifica como fraude).

La clasificación para cada instancia puede dar solo dos entradas (falso positivo, verdadero negativo) o (falso negativo, verdadero positivo) para la denominación de los resultados obtenidos en el ejercicio de prevención de fraude (Crespo & Toscano, 2015). Por tanto, en la literatura de FDS, se han propuesto dos enfoques principales para utilizar el aprendizaje sensible al costo para la clase desequilibrada, tales como umbrales de metacostos o el uso de estudiantes que no son sensibles al problema del desequilibrio de clase para su revisión estructurada, basada en la formulación de la matriz de costos estimada en los métodos de nivel algorítmico, sobre lo cual, se espera la presentación o resultados de las etiquetas que darán una definición sobre el nivel de riesgo.

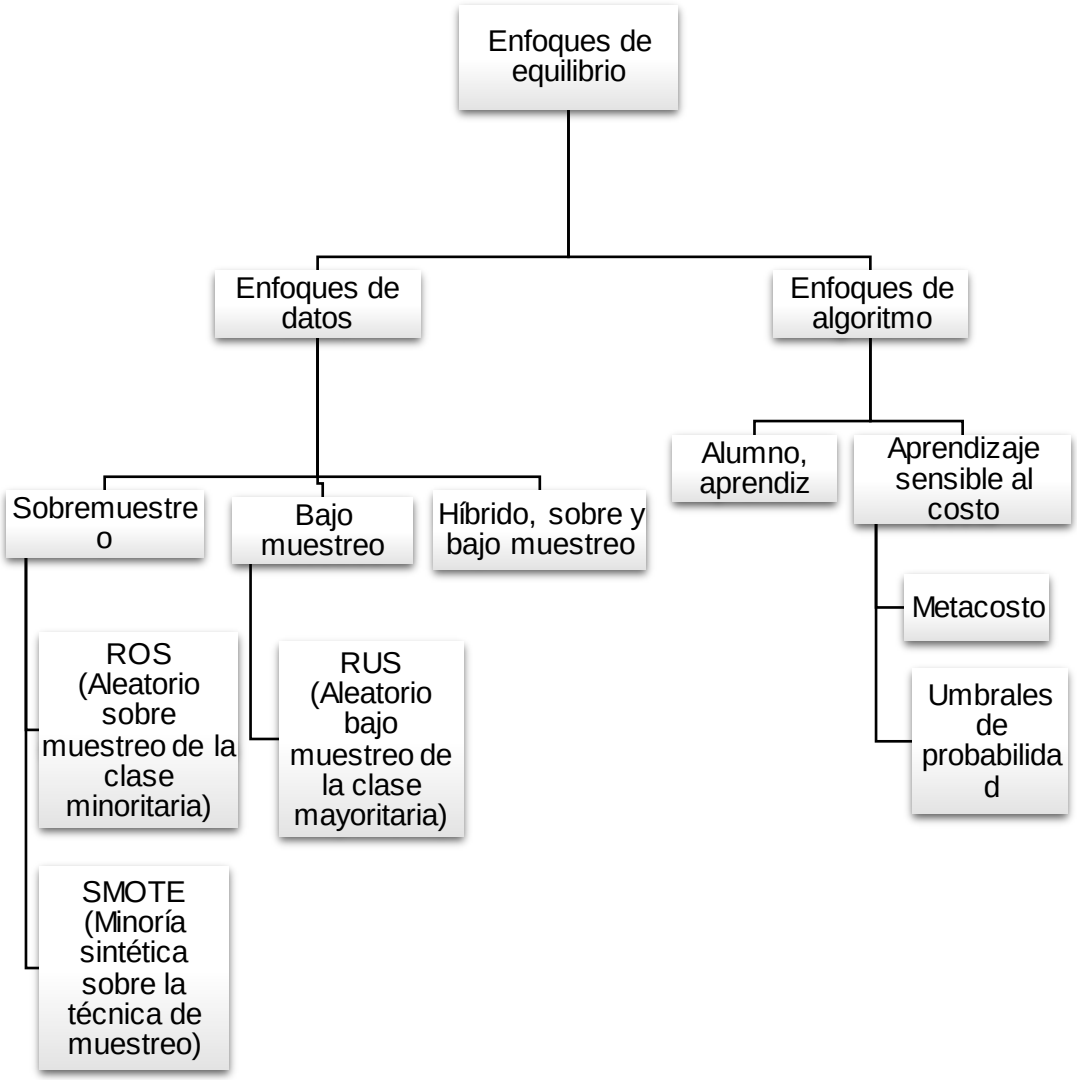


Figura 2.6 Enfoques de equilibrio

Fuente: (Donayre & Santos, 2018)

Estos métodos se utilizan con frecuencia en el sistema de detección de fraude para equilibrar los datos de capacitación en sistemas de detección de fraude en el comercio electrónico a través de internet (Almagro, Urrutia, & August-Treppel, 2018). Usando aprendices para manejar la distribución sesgada, es otro método algorítmico utilizado en la literatura de FDS. Estos alumnos (aprendices del sistema en el ámbito empresarial) son resistentes al problema de desequilibrio de clase a través de las propiedades inherentes del alumno, como en el caso del algoritmo de poda incremental repetida para producir reducción de errores.

En general, los métodos de datos funcionan mejor que los métodos de algoritmo; esto se debe al hecho de que los métodos de datos son más fáciles de implementar y no conducen al aumento del tiempo de capacitación o los recursos necesarios (Guillén & Biarge, 2019). Por lo tanto, la mayoría de la literatura de FDS utiliza técnicas de equilibrio de nivel de datos y es sobre la cual se centra el presente estudio para la prevención de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones en el Ecuador.

2.2.2.3. Reducción de gran cantidad de datos

El conjunto de datos de fraude a gran escala y alta y la presencia de un número de características/atributos/entradas/variables hacen que el proceso de extracción y detección de datos sea extremadamente difícil y complicado (Arcenegui, Castilla, & Lozano, 2015). Además, esta situación también ralentiza el proceso de detección; por lo tanto, los FDS existentes utilizan enfoques de reducción de datos para reducir el tamaño del conjunto de datos, produciendo un tamaño de modelo pequeño que puede ser útil con respecto al procesamiento en tiempo real para la detección y prevención del fraude electrónico en el sector empresarial, aplicado a todos los tamaños de organizaciones, en ejecución de sus actividades comerciales bidireccionales (cliente – empresa), a través de la web.

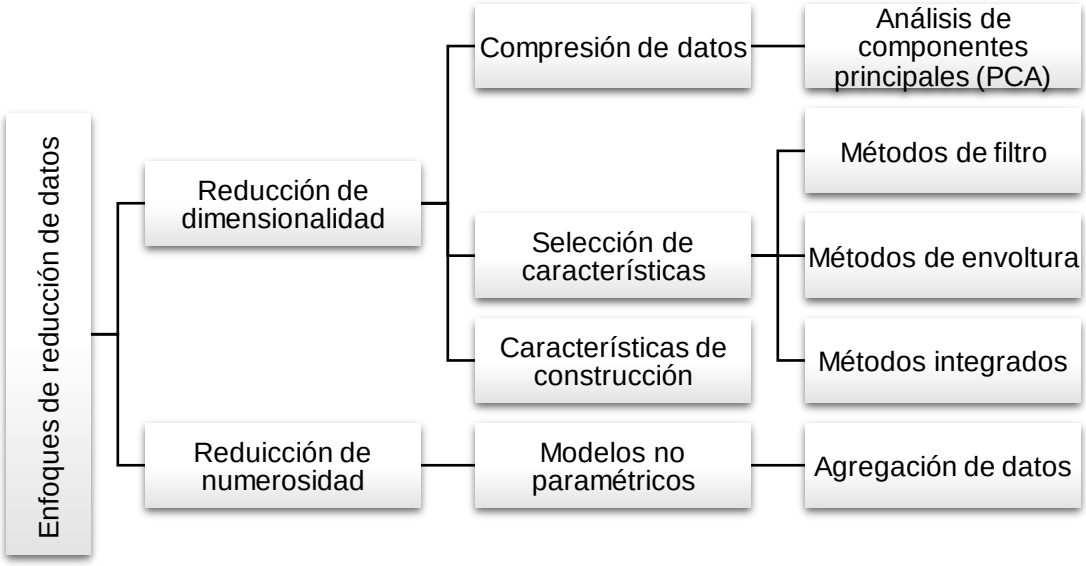


Figura 2.7 Estrategias de reducción de datos

Fuente: (Gilman & Joyce, 2014)

Sobre lo presentado en la figura 2.7, se puede colegir que, además, los datos pequeños reducirán el tamaño del modelo, reduciendo en consecuencia el tiempo de cálculo (Gilman & Joyce, 2014). Los enfoques de reducción de datos incluyen reducción de dimensionalidad y reducción de numerosidad, por lo que, la reducción de la dimensionalidad incluye muchas estrategias, es decir, la compresión de datos, la selección de características y la construcción de características son las estrategias más comunes y de uso frecuente en los FDS.

La estrategia de compresión de datos comprime la representación de los datos originales mediante el uso de técnicas de compresión de datos. Mientras tanto, la selección de características es otra estrategia de reducción de dimensionalidad, las características más significativas y relevantes se seleccionan para ser utilizadas en la construcción del modelo (Herrera-Semenets & Prado, 2014). Por tanto, se utilizan tres métodos de selección de características en FDS tales como: métodos de filtro, métodos de envoltura y métodos integrados. Los métodos de filtro actúan como algoritmo de preprocesamiento para clasificar las características en las que las características altamente clasificadas se seleccionan y aplican a un predictor.

En los métodos de envoltura, el criterio de selección de características es el rendimiento del predictor, es decir, el predictor está envuelto en un algoritmo de búsqueda que encontrará un subconjunto que proporciona el mayor rendimiento del predictor (Pérez, Pacheco, & Salazar, 2016). Los métodos integrados incluyen la selección de variables como parte del proceso de capacitación sin dividir los datos en conjuntos de capacitación y prueba para la ejecución eficiente del proceso de detección de fraude en sistemas de pago electrónico destinado a fortalecer e incrementar su fiabilidad.

A continuación, la construcción de características es donde se deriva un pequeño conjunto de características más útiles del conjunto original. Mientras tanto, en la reducción de numerosidad, los datos se reemplazan por representaciones más pequeñas, como el uso de la agregación de datos, por

lo que, los enfoques de reducción de datos incluyen reducción de dimensionalidad y reducción de numerosidad para la determinación de las principales áreas o segmentos donde se establecen los focos mayoritarios de riesgo de fraude electrónico.

2.2.3. Principales áreas de fraude en sistemas de pago electrónico

Casi cualquier sistema tecnológico de telecomunicaciones que involucre dinero y servicios puede verse comprometido por actos fraudulentos, por ejemplo, sistema de tarjeta de crédito, sistema de telecomunicaciones, sistema de seguro de salud, etc. La figura 2.8 muestra las áreas más comunes de fraudes, sobre lo cual se deriva el sistema de pagos electrónico en el sector empresarial de telecomunicaciones del Ecuador. Por tanto, a partir de ello, se abordará el fraude que ocurre en las áreas más frecuentes, que son las tarjetas de crédito, las telecomunicaciones, sistemas de pago de seguros, etc., (principalmente páginas web de compra – venta).

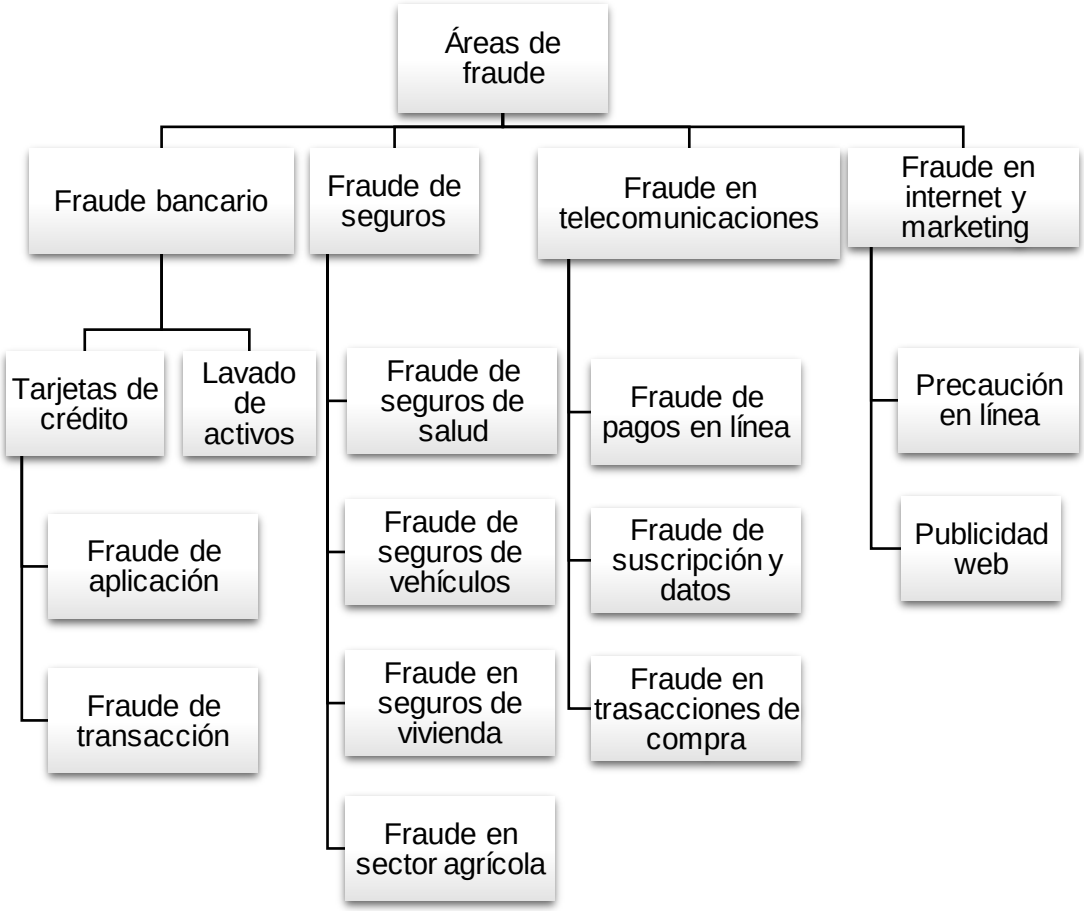


Figura 2.8 *Taxonomía* de las áreas más comunes de fraudes
Elaborado por: Autor

Posteriormente a ello, la figura 2.9 muestra una estadística del trabajo publicado relacionado con las cinco áreas de fraude de 1998 a 2018 tomada de la CCG, como referencia en 20 años, para la revisión de la evolución de fraude en el sector empresarial nacional. Sobre ello, prevalece que el fraude bancario es el área más investigada en el área de tecnología y telecomunicaciones. El fraude de seguros es la tercera área popular, que ha sido el tema principal de varios estudios, ya que puede incluir otras áreas como el fraude de seguros de salud, fraude de seguros de automóviles, fraude de seguros de hogar y fraude de seguros de agricultura (principalmente en el sector bananero y cacaotero del país).

Las telecomunicaciones y el marketing en Internet son las áreas menos estudiadas durante el período de tiempo especificado. Por tal razón, este documento se centrará principalmente en el análisis de fraude de sistemas de pago enfocado en la utilización de tarjetas de crédito que se encuentra en el fraude electrónico, para determinar en el sector empresarial de telecomunicaciones realizado a través de internet.

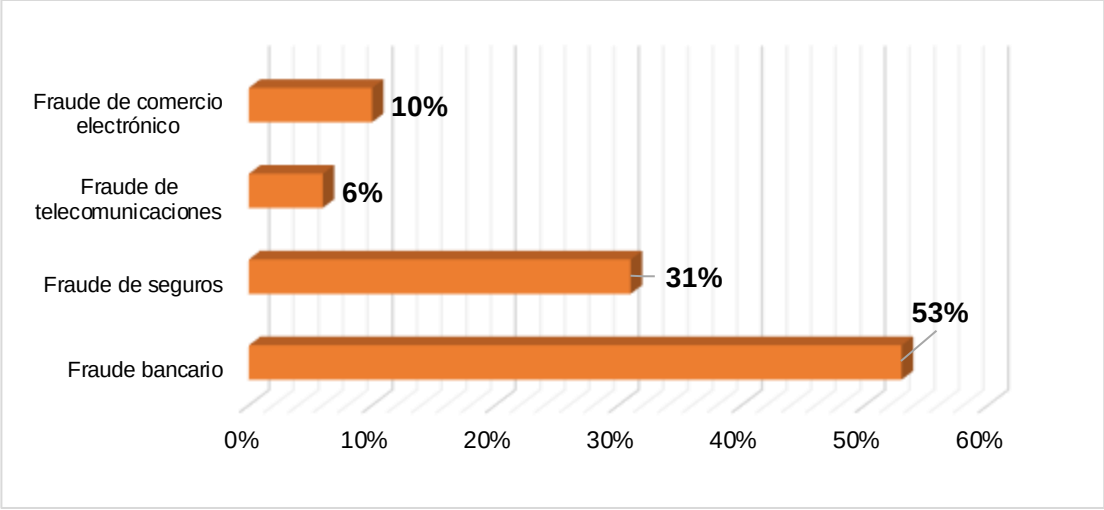


Figura 2.9 Resumen de la cantidad de área de fraude más investigada

Fuente: (Cámara de Comercio de Guayaquil, 2019)

2.2.3.1. Fraude en tarjetas de crédito

El término crédito se usa para describir el método de compra y venta de bienes sin tener dinero; por lo que, la tarjeta de crédito es una pequeña tarjeta

de plástico para proporcionar el servicio de crédito al cliente (Osorio, 2017). La tarjeta de crédito es muy popular y juega un papel importante en el comercio electrónico y el área de transacciones de dinero en línea que crece cada año. Como resultado del uso creciente de la tarjeta de crédito, los estafadores intentan encontrar más oportunidades para cometer fraudes que pueden causar grandes pérdidas a los titulares de tarjetas y bancos. El fraude con tarjetas de crédito se considera un tipo común de fraude crediticio y toma muchas formas, las publicaciones existentes las clasifican en varias categorías comunes en función de:

a) Fraude de tarjeta de crédito fuera de línea: sucede cuando los estafadores roban la tarjeta de plástico, usándola en las tiendas como el propietario real. Este es un tipo de fraude poco común porque las instituciones financieras bloquearán inmediatamente la tarjeta perdida cuando los titulares de tarjetas denuncien el robo.

b) Fraude de tarjetas de crédito en línea: un fraude popular y muy peligroso, los estafadores roban la información de las tarjetas de crédito para usarla más adelante en transacciones en línea por Internet o por teléfono. Otro nombre para este tipo de fraude es el titular de la tarjeta no está presente; mientras que los detalles de la tarjeta se proporcionan a través de uno de los siguientes métodos: clonación de sitios, generadores de tarjetas de crédito o *phishing.*

Hay otra clasificación para el fraude de tarjetas de crédito, es decir existen dos tipos: fraude de aplicaciones y fraude conductual (Arcenegui, Castilla, & Lozano, 2015). Esta clasificación se basa en la estrategia del estafador para cometer fraude; por lo que, el fraude de aplicaciones ocurre cuando los estafadores ingresan información y detalles incorrectos en el formulario de solicitud para abrir una nueva tarjeta de crédito. Los estafadores pueden usar la información de otras personas para obtener tarjetas de crédito u obtener sus nuevas tarjetas de crédito mediante el uso de información personal falsa con la intención de nunca pagar las compras.

Por otro lado, el fraude conductual ocurre cuando los estafadores obtienen los detalles del titular de la tarjeta para usarlos más tarde en las ventas que se realizan en base al presente del titular de la tarjeta que es considerado como el cliente o consumidor de la empresa a la que se ha referido la transacción electrónica (Pérez, Pacheco, & Salazar, 2016). Estas ventas incluyen ventas telefónicas y transacciones de comercio electrónico, donde solo se requieren los detalles de la tarjeta, y se clasifican en:

a) Tarjetas perdidas y tarjetas robadas.
b) Tarjetas falsificadas.
c) Robo de tarjetas por correo o no recibo de emisión (NRI).
d) Fraude de pedidos por correo/teléfono. En tales casos, el comprador no está físicamente presente ante el comerciante en el momento de la transacción, y no hay impresión de tarjeta que se pueda obtener como un registro de la transacción.
e) El fraude comercial puede implicar el lavado de activos de comerciantes falsos, obteniendo grandes sumas de dinero para transacciones que nunca ocurrieron.
f) El titular de la tarjeta realiza compras en la tarjeta por las cuales no tiene intención de pagar. Esto se llama fraude de bancarrota.

Además de ello, se desprende una subcategorización de fraude con tarjetas de crédito, que se divide en tres categorías, que son: fraudes tradicionales relacionados con tarjetas, fraudes relacionados con comerciantes y fraudes por Internet.

Detección de fraude con tarjetas de crédito

La detección de fraude con tarjeta de crédito es uno de los dominios más explorados de detección de fraude; por lo que, se utilizan numerosas técnicas de autorización para evitar fraudes con tarjetas de crédito, como firmas, número de tarjeta de crédito, número de identificación, dirección del titular de la tarjeta, fecha de vencimiento, etc., (González, Barrero, & Bohórquez, 2016) Sin embargo, estas técnicas no son suficientes para obstaculizar el fraude con

tarjeta de crédito. Por lo tanto, existe la necesidad de utilizar enfoques de detección de fraude que analicen los datos que pueden detectar y eliminar el fraude con tarjetas de crédito.

Principalmente, la estrategia de detección de fraudes con tarjetas de crédito es el reconocimiento de patrones mediante el análisis automático del comportamiento del gasto del usuario. El comportamiento del gasto del cliente contiene información sobre el monto de la transacción, el intervalo de tiempo desde la última compra, el día de la semana, la categoría del artículo, la dirección del cliente, etc., (Noguera, 2014). La detección de fraude basada en anomalías se usa principalmente para el sistema de detección de fraude de tarjeta de crédito en el que se compone el perfil del titular de la tarjeta mediante el análisis del patrón de comportamiento de gasto del titular de la tarjeta. Al hacerlo, cualquier transacción entrante que sea inconsistente con el perfil del titular de la tarjeta se considerará sospechosa.

Los enfoques de perfil se pueden hacer en función del enfoque del propietario o del enfoque de la operación. El enfoque de propietario es donde cada usuario de tarjeta de crédito se perfila en función de su historial de uso de tarjeta de crédito (Muguiro & Lafuente, 2014). Cualquier nueva transacción que se realice se compara con el perfil del usuario y se sospecha que es un fraude si no coincide con el perfil. Mientras tanto, el enfoque operativo detecta transacciones fraudulentas de transacciones que tienen lugar en una ubicación geográfica específica.

Por otro lado, el enfoque de detección de uso indebido siempre se ha llevado a cabo para detectar fraudes con tarjetas de crédito, donde los patrones de huellas digitales de personas fraudulentas se utilizarán para aprender sobre el FDS de la tarjeta de crédito (González, Barrero, & Bohórquez, 2016). En general, el enfoque de detección de mal uso rara vez se utiliza para FDS de tarjetas de crédito u otras áreas de fraude.

Tabla 2.3 Enfoques y técnicas en tarjetas de crédito FDS

Tipo de detección	***Herramienta de detección***	***Enfoque de aprendizaje***	***Categoría de minería de datos***	***Técnicas***
Métodos de detección de fraude basados en anomalías.	Minería de datos	Supervisado	Clasificación	Árboles de decisión Redes neuronales artificiales
		Sin supervisar	Agrupamiento	Modelo oculto de Markov Sistemas de inmunidad artificial *K*-vecinos más cercanos Máquinas de vectores soporte Algoritmo genético Borroso Algoritmo de mapa autoorganizado

Fuente: (Arqué, 2014)

Esto se debe a su incapacidad para detectar nuevos fraudes; por lo tanto, en la tabla 2.3 solo se enfoca en el ámbito basado en anomalías, así como en otros tipos de fraude. Como se puede observar en la tabla 2.3, los métodos de minería de datos más utilizados para crear FDS de tarjetas de crédito es el método de clasificación, en particular las técnicas de redes neuronales bajo aprendizaje supervisado.

Problemas y desafíos de detección de fraude con tarjetas de crédito

El área de la tarjeta de crédito está siendo abordada principalmente por investigadores en tecnología y telecomunicaciones, que se enfocan en todas las áreas de fraude con respecto al comercio electrónico (específicamente a través de internet) (González, Barrero, & Bohórquez, 2016). Esto se debe a

que los titulares de tarjetas de crédito siempre cambian su comportamiento que puede deberse a circunstancias específicas (por ejemplo, vacaciones, navidad), y en este período, aumentará el poder adquisitivo de los usuarios y/o consumidores.

Si FDS no considera esto como un cambio normal, se considerará un comportamiento fraudulento y se activarán las alarmas, bloqueando la transacción del titular de la tarjeta, lo que conducirá a una disminución de la reputación bancaria en el futuro (Guillén & Biarge, 2019). Por lo tanto, el FDS de la tarjeta de crédito debe discriminar y clasificar las transacciones fraudulentas y legítimas de manera efectiva. Además, el FDS de la tarjeta de crédito debe ser capaz de capturar y adaptar el comportamiento *drift* de los tarjetahabientes, actualizando el modelo de detección a ese comportamiento con el tiempo.

En consecuencia, el FDS de la tarjeta de crédito debe tener alertas falsas bajas pero alta precisión de detección. Por esa razón, se presentan varios enfoques para manejar el sistema de tarjeta de crédito (Salazar & Flores, 2016). Estos enfoques se clasificarán según su taxonomía, es decir, el enfoque basado en evolución y el enfoque basado en regulaciones. La mayoría de los FDS adaptativos existentes que manejan el concepto *drift* están utilizando un enfoque basado en la evolución que incluye conjuntos adaptativos y técnicas específicas del modelo base.

Por ejemplo, a través del diseño de un marco efectivo de detección de fraudes con tarjetas de crédito que extraiga flujos de datos utilizando clasificadores de conjuntos ponderados (Mock & Lupini, 2017). Por lo que, a partir de fragmentos secuenciales de datos de tarjetas de crédito, se estima que un modelo adaptativo para el sistema de detección de fraude con tarjetas de crédito, aborde la fluctuación o evolución, así como el comportamiento del titular de la tarjeta regular o el estafador. Por tanto, de ello se desprenden tres enfoques diferentes que manejan el problema desequilibrado en un entorno cambiante, con un enfoque estático, enfoque de actualización y enfoque de olvido.

Esto promueve que la detección basado en un método de conjunto de árbol de decisión de validación cruzada simple, maneja varias áreas; uno de ellos es la detección de fraude con tarjeta de crédito (Pérez, Pacheco, & Salazar, 2016). Por lo que, además, el enfoque de aprendizaje en evolución puede aplicarse mediante el uso de un modelo base específico, en el que la adaptación se logra mediante la gestión de parámetros o diseños de modelos específicos.

La evolución de las arquitecturas permite a las redes neuronales artificiales (ANN) adaptar sus topologías a la deriva de los tarjetahabientes sin intervención humana y, por lo tanto, proporciona un enfoque para el diseño automático de ANN (Roa, García, Frías, & Correa, 2017). Por otro lado, sobre ello, se adopta un enfoque regulado que maneja la deriva del concepto al monitorear dos técnicas de perfiles diferentes proponiendo FDS adaptativo, utilizando métodos de detección de mal uso y anomalías con algoritmo incremental BOAT para descubrir el fraude y cambiar el comportamiento de gasto del usuario legítimo.

2.2.3.2. Fraude en telecomunicaciones

El fraude de telecomunicaciones es un problema que ha crecido dramáticamente en los últimos 10 años. El fraude en las telecomunicaciones móviles es un problema complejo y dinámico para los operadores de telecomunicaciones; esto se debe a que estos fraudes amenazan los servicios prepagos y pospagos (Guillén & Biarge, 2019). Además, se puede cometer fraude con líneas telefónicas fijas y móviles; se comete fraude de línea fija contra compañías telefónicas; esto a medida que los estafadores obtienen acceso a la central de la empresa de telecomunicaciones y venden la capacidad de otras personas para hacer llamadas a través de ella.

El fraude de telecomunicaciones es el uso no autorizado, la manipulación o manipulación de un teléfono celular o servicio (tal como sucede con el sistema de pagos en línea de las empresas de telecomunicaciones a través de sus plataformas web) (Salazar & Flores, 2016). En general, el objetivo

principal de cometer fraude en ambos tipos de telecomunicaciones (telefonía e internet) es obtener servicios medios ilegales.

Según la encuesta global de pérdida por fraude anunciada por la Asociación de Control de Fraude de Comunicaciones (CFCA), en 2018, se registró una pérdida por fraude en telecomunicaciones de US$ 46.3 mil millones de dólares, un aumento del 15% desde 2015. Como porcentaje de los ingresos mundiales de telecomunicaciones, las pérdidas por fraude son aproximadamente 2.09%, con un aumento del 0,21% desde 2011. Esto se debe a la gran cantidad de fraude de telecomunicaciones registrado en diferentes categorías, por ello, se agrupa el fraude de telecomunicaciones en cuatro categorías:

a. Fraude contractual: el estafador utiliza servicios de telecomunicaciones sin intención de pagar el cargo por el servicio; por ejemplo, fraude de suscripción y fraude de tarifa premium a través de internet o medios digitales.
b. *Hackeo* de fraude: los estafadores en esta categoría violaron los sistemas comerciales y aprovecharon los recursos disponibles ilegalmente.
c. Fraude técnico: los estafadores en esta categoría aprovechan las debilidades existentes en la tecnología de sistemas móviles. Tal fraude necesita un alto conocimiento técnico. Ejemplos de tal fraude son la clonación y el fraude técnico interno.
d. Fraude procesal: los fraudes en este grupo implicaron ataques contra los procedimientos implementados para reducir el riesgo de exposición al fraude y, a menudo, atacan las debilidades en los procedimientos comerciales utilizados para otorgar acceso al sistema.

Por otro lado, la literatura también clasifica el fraude de telecomunicaciones de acuerdo con tres áreas que son:

1. Motivo: la razón principal detrás de cometer fraude.

2. Medios: la naturaleza o forma del fraude utilizado para satisfacer el motivo.
3. Métodos: las instalaciones y herramientas que se utilizan para cometer fraude.

Existen muchos tipos de fraudes que amenazan a los sectores de telecomunicaciones, que se consideran el área de fraude más popular. Se estima que existen más de 200 variantes de fraude de telecomunicaciones en la industria de las telecomunicaciones (Herrera-Semenets & Prado, 2014). En la actualidad, la literatura en tecnología y telecomunicaciones ha enumerado los fraudes de telecomunicaciones más estudiados, los cuales son presentados como fraude de suscripción y el fraude superpuesto, que son los tipos más frecuentes de fraudes de telecomunicaciones. Por lo tanto, se abordan con frecuencia en la literatura en comparación con otros tipos de fraude de telecomunicaciones con enfoque al sector empresarial.

Telecomunicaciones FDS

La detección de fraude basada en anomalías generalmente se utiliza para FDS de telecomunicaciones. El perfil extraído de cada suscriptor basado en sus patrones de CDR se utiliza para detectar comportamientos anormales (Otero & Bustamante, 2012). Estos perfiles se basan en CDR (por ejemplo, número de llamadas, duración de la llamada, tipo de llamada) o propiedades demográficas del suscriptor (por ejemplo, edad, sexo, región) o ambos. CDR es muy útil para extraer el comportamiento del usuario, por tanto, el FDS de telecomunicaciones basado en el enfoque de perfil se basa en una comparación de historiales de comportamiento recientes y a largo plazo derivados de los datos obtenidos.

Por tanto, si hay un cambio significativo en el patrón, se activarán las alarmas, razón por lo que, en la tabla 2.4 se muestra los enfoques y técnicas utilizados en el FDS de telecomunicaciones para especificar el tipo de detección, herramienta, enfoque, categoría y técnicas (García & Aller, 2017). A partir de ello, se ha utilizado una hibridación de técnicas supervisadas y

técnicas no supervisadas para obtener los mejores resultados. Por tanto, se puede utilizar una red neuronal de retroalimentación basada en el aprendizaje supervisado para el funcionamiento de la ejecución discriminativa y a través de ello, proceder a la clasificación a los suscriptores mediante el uso de estadísticas resumidas que son fundamentales para la detección del fraude en medios de pago electrónicos.

Tabla 2.4 Enfoques y técnicas en telecomunicaciones FDS

Tipo de detección	***Herramienta de detección***	***Enfoque de aprendizaje***	***Categoría de minería de datos***	***Técnicas***
Métodos de detección de intrusiones basados en anomalías.	Minería de datos	Supervisado	Clasificación	Árboles de decisión Redes neuronales artificiales
		Sin supervisar	Agrupamiento	Modelo oculto de Markov Sistemas de inmunidad artificial Razonamiento basado en casos Clasificación bayesiana Máquina de vectores de soporte Basada en reglas Lógica difusa PCA Distancia basada Mezcla gaussiana Aglomerativo jerárquico Agrupamiento Análisis discriminante Visualización de datos
	Minería estadística Minería visual y de datos		Método estadístico Visualización	

Fuente: (Burke, 2016)

Luego, el modelo de mezcla gaussiana se usa para modelar la densidad de probabilidad del comportamiento pasado de los suscriptores, de modo que la probabilidad del comportamiento actual se pueda calcular para detectar cualquier anomalía del comportamiento pasado (Ruiz-Capillas & Fernández, 2014). Por último, las redes bayesianas se utilizan para describir las

estadísticas de un usuario en particular y las estadísticas de diferentes escenarios de fraude. Aparte de eso, hay casos en los que se han combinado múltiples técnicas supervisadas, que usaba reglas difusas y redes neuronales.

2.2.4. Síntesis literaria

Los datos de información de los clientes se clasifican de forma permanente en las plataformas web del sector empresarial (especialmente en las páginas relacionadas a pagos en línea), debido a muchas razones, una de ellas es que los clientes cambian sus comportamientos debido a la introducción de nuevos servicios agregados por las empresas. Las consecuencias de ignorar los cambios en la seguridad para la prevención del fraude en las transacciones electrónicas pueden ser catastróficas.

En el área de telecomunicaciones, los FDS adaptativos más actuales se basan en el enfoque evolutivo para detectar los cambios; especialmente utilizando la técnica específica del modelo base. El sistema de detección de fraude adaptativo revisado se ha aplicado al problema (respecto a la prevención del fraude en sistemas de pago electrónico) de detectar el fraude de clonación de datos basado en una base de datos de registros. Este documento en la revisión de la literatura presenta un marco que se ha utilizado el método basado en reglas, que funciona mediante la generación de indicadores (características) para que los monitores (modelos) los utilicen automáticamente a fin de detectar un fraude y activar alarmas.

Por ello, el diseño del modelo de detección de fraude adaptativo utiliza un sistema de inferencia *neurofuzzy* adaptativo en redes de telecomunicaciones; por lo que se estima que logre una alta precisión del sistema de clasificación basado en difusos y la propiedad de adaptabilidad (retropropagación) de las redes neuronales en la clasificación de datos. Siendo así, se puede desarrollar un modelo adaptativo de detección de fraude de telecomunicaciones mediante el uso de una combinación de tecnología neural, basada en reglas y basada en casos, utilizando la clasificación Naive-Bayesiana para calcular la probabilidad y una versión adaptada de la divergencia KL para identificar la

diferencia significativa entre un usuario normal y un usuario sospechoso sobre la base de la suscripción.

CAPÍTULO 3: METODOLOGÍA Y RESULTADOS

3.1. Metodología de investigación

La metodología a utilizar en la presente investigación es descriptiva, diseñada para representar a los participantes del proceso de recolección de información de una manera precisa. Es decir, se trata de fundamentar un análisis al sector empresarial de telecomunicaciones y su incidencia en el fraude en sistemas de pago electrónico con el que cuentan sus plataformas web para interactuar con el cliente.

A partir de ello, la investigación descriptiva establece los pasos necesarios para llegar a la consecución del estudio, basado en: revisión de la situación del problema, definición de hipótesis, elección de fuentes literarias, selección de técnicas e instrumentos de recolección de información y finalmente análisis e interpretación de datos, con los que se espera llegar a una discusión del estudio acertada para el análisis objeto de la investigación.

Por tanto, para el presente estudio se realizará una encuesta para analizar el escenario de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones en la empresa Claro de la ciudad de Guayaquil. Para la detección de fraude en sistemas de pago electrónico, se ha identificado los métodos de comprobación que utiliza la organización en la web. El robo de identidad es una vía importante para el fraude en línea y, por lo tanto, la mayoría de las técnicas de prevención y detección de fraude se basan en verificar la identidad correcta del usuario en cada transacción en línea.

La identidad del usuario puede confirmarse de muchas maneras utilizando información diferente del usuario, como dirección, edad y fecha de nacimiento, etc. Estos detalles del usuario pueden confirmarse mediante llamadas telefónicas, correo electrónico o por medios electrónicos. Esta sección analiza metodológicamente el uso de los detalles del usuario para confirmar la identidad en prevención de fraude en sistemas de pago electrónico en la empresa de telecomunicaciones objeto de estudio.

3.2. Tipo de investigación

El tipo de investigación es cualitativo – cuantitativo, debido a que se cualifica la información proporcionada en el estudio, a través de una revisión de la literatura adecuada a la investigación y fundamentada con datos de la organización para conocer el escenario de fraude en sistemas de pago electrónico.

Además, es cuantitativo debido a que se cuantifica la información mediante la consecución de una encuesta dirigida de manera directa a la empresa de telecomunicaciones que realiza sus actividades comerciales y de mercado en la ciudad de Guayaquil. Por lo tanto, asigna resultados a determinada población a través de la técnica de muestreo para analizar de primera mano el fraude electrónico y el sistema de protección institucional.

3.3. Enfoque

El enfoque de investigación es interpretativo, de acuerdo al objetivo de análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones local. Con ello, la investigación se adecua a la realización de una encuesta que ayude a comprender su contexto desde su ejecución natural y cotidiana en el ejercicio de las actividades corporativas que desarrolla en el mercado. Este enfoque se ha seleccionado, debido a que es adecuado a la investigación cualitativa – cuantitativa de metodología descriptiva, cuyo enfoque lleva a interpretar los datos a obtener mediante este método.

Por tanto, este enfoque ayuda a tener una mejor perspectiva de la situación actual del sector empresarial de telecomunicaciones, específicamente desde la ciudad de Guayaquil, donde los clientes se encuentran expuestos a un riesgo en la seguridad de sus datos e información al momento de ser ingresados en las plataformas respectivas de la empresa donde se ha desarrollado la investigación; con lo que, además no solo se limita a este sector geográfico, sino que a través de internet se alcanza a todo el mundo mediante sus plataformas web para interacción con el cliente.

3.4. Técnicas e instrumentos

La técnica utilizada para esta investigación es la encuesta, que será aplicada a través del instrumento seleccionado, que es un cuestionario de 10 preguntas realizadas aplicando escala Likert para cuantificar la información obtenida y sobre ello al ser cuantitativa, contrastará y cuantificará la información sobre datos operativos en el departamento de telecomunicaciones y comercio electrónico de la empresa, sobre su sistema de detección de fraude electrónico. Esta herramienta que utiliza una escala psicosomática ayudará al estudio a la consecución de datos adecuados al escenario del problema presentado que pretende analizar las técnicas de detección y prevención de fraude en medios de pago electrónico en una empresa de telecomunicaciones en la ciudad de Guayaquil.

A través de ello, se pretende poder conocer de primera mano su ejecución y detalle para la comprobación de la hipótesis de si el sistema de prevención de fraude en sistemas de pago electrónico ayudaría al sector empresarial ecuatoriano de telecomunicaciones al mejoramiento de la confianza en sus clientes y mercado, basados en métodos de prueba de identidad y aplicaciones tecnológicas. La premisa de dicha hipótesis aplicada a las técnicas e instrumentos se basa en la variedad del fraude en línea y las diferentes técnicas utilizadas por los estafadores en línea, lo cual también han requerido la invención de diferentes técnicas de prevención y detección para detener el efecto negativo del fraude electrónico.

Sin embargo, las técnicas adoptadas para combatir el fraude en línea dependen de las innovaciones tecnológicas y su aceptación y uso efectivo en la empresa. Si bien algunas técnicas son fáciles de usar y efectivas en algunas áreas, es posible que algunas técnicas sean difíciles e inapropiadas de usar en algunas áreas y culturas. Por lo tanto, las diferentes técnicas están diseñadas para fines específicos, tales como verificaciones de identidad, verificación de datos y verificación de dirección, y pueden aplicarse manualmente o con el uso de tecnologías.

3.5. Población y muestra

Para la definición de la población y muestra en la presente investigación, se ha recurrido a lo suscrito por la delimitación de estudio en el sector empresarial de telecomunicaciones del Ecuador, específicamente se realizará la encuesta en los departamentos de telecomunicaciones y comercio electrónico de la empresa CLARO, en la sede de la ciudad de Guayaquil, en el horario: lunes 21 de octubre a viernes 25 de octubre de 2019, de 9H00 a 17H00. Por tanto, la población total de los departamentos constituye un personal de 17 colaboradores (8 en el primer departamento y 9 en el segundo).

Debido a la cantidad exacta de colaboradores que desempeñan labores en los departamentos descritos de la empresa de telecomunicaciones CLARO, no ha sido necesaria una formula estadística que permita conocer la muestra; siendo así, la población y muestra constituye el mismo elemento de estudio en ambos departamentos, sobre los cuales se realizará el análisis de estudio, para conocer de primera mano el desempeño de la institución en el sistema de pagos de su plataforma electrónica, y a través de ello, analizar la incidencia de fraude electrónico en sus sistemas de pago. Por tanto, se detalla:

Tabla 3.1 Descripción de población y muestra

Población y muestra
Unidades de Análisis: Departamentos de Telecomunicaciones y Comercio electrónico.

Descripción	**No.**
Población	17
Muestra	17

Fuente: (Consorcio Ecuatoriano de Telecomunicaciones (CONECEL), CLARO, 2019)

En consecuencia, las innovaciones tecnológicas han hecho posible que los datos personales de la identidad y otra información relevante de clientes y clientes potenciales se almacenen y verifiquen en línea utilizando algunas formas y plataformas de base tecnológica, sobre las cuales se procede a la revisión y análisis de resultados. Estas plataformas y dispositivos ahora se

usan comúnmente en las secciones de telecomunicaciones y comercio electrónico de la empresa y otras transacciones relacionadas con la seguridad, basado especialmente en aplicaciones tecnológicas tales como:

Tabla 3.2 Técnicas de detección y prevención de fraude

Aplicaciones tecnológicas para detección y prevención de fraude	*Biometría*
	Geolocalización
	Firma digital
	Identificación del dispositivo
	Detección de proxy
	Fichas seguras
	Tarjetas inteligentes
	Métodos para compartir datos

Elaborado por: Autor

Por tanto, en relación con estas aplicaciones tecnológicas para detección y prevención de fraude en sistemas de pago electrónico, se procede a revisar la encuesta.

3.6. Análisis de los resultados

3.6.1. Resultados de encuesta

3.6.1.1. Primera pregunta

¿Considera usted que el sistema de pago electrónico que utiliza la plataforma web de la empresa CLARO, debe ser observado continuamente para prevenir el fraude en la interacción comercial en línea?

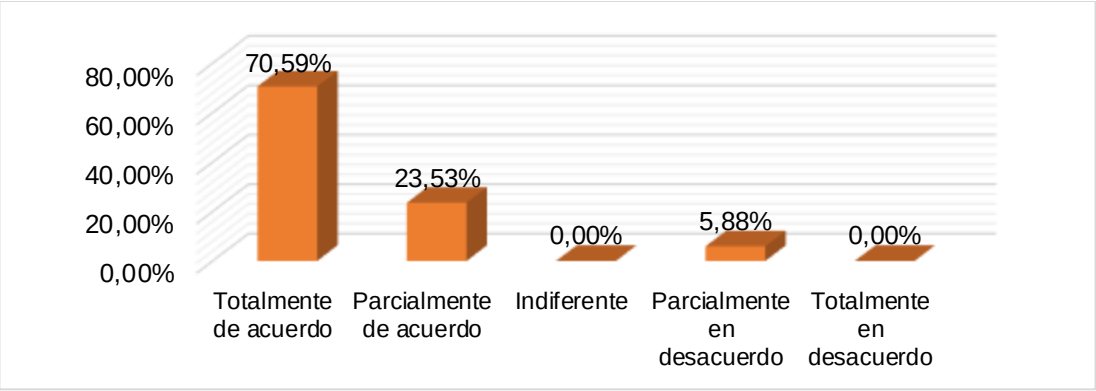


Figura 3.1 Observaciones en sistema de pago

Elaborado por: Autor

Tabla 3.3 Datos de observación en sistemas de pago

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*12*	*70,59%*
Parcialmente de acuerdo	*4*	*23,53%*
Indiferente	-	-
Parcialmente en desacuerdo	*1*	*5,88%*
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 70,59% (12) de los encuestados, dijeron estar totalmente de acuerdo en que la empresa debe observar de manera continua el sistema de pago electrónico para prevenir fraude en la interacción comercial en línea que realizan los clientes al momento de acceder a un pago de producto o servicio.

3.6.1.2. Segunda pregunta

¿Cree usted que el sector empresarial de telecomunicaciones del Ecuador debe trabajar en conjunto con el Ministerio de Telecomunicaciones y de la Sociedad de la Información para prevenir fraudes en sistemas de pago electrónicos?

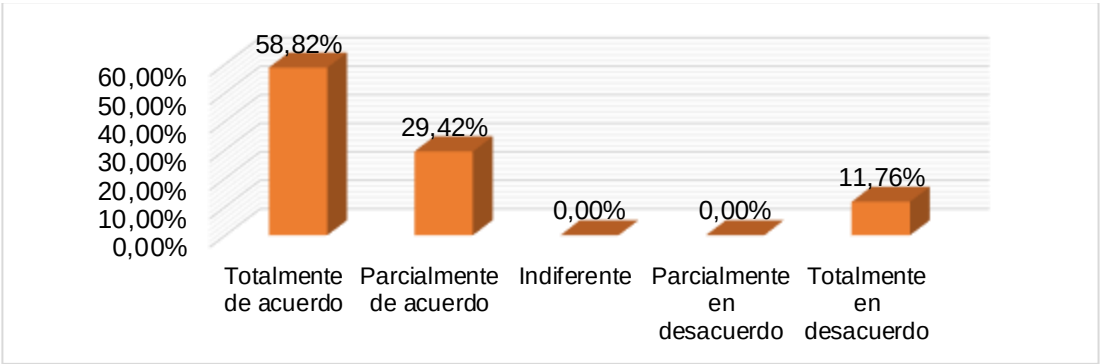


Figura 3.2 Trabajo conjunto, sectores público – privado

Elaborado por: Autor

Tabla 3.4 Datos de cooperación entre el Estado y la empresa privada

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*10*	*58,82%*
Parcialmente de acuerdo	*5*	*29,42%*

Indiferente	-	-
Parcialmente en desacuerdo	*2*	*11,76%*
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 58,82% (10) de los encuestados, están totalmente de acuerdo en colaborar de manera conjunta con el Ministerio de Telecomunicaciones y de la Sociedad de la Información para prevenir el fraude electrónico, como una medida de coordinación y cooperación entre el sector público y el sector privado, y con ello, adherirse a las políticas del Estado para la protección de datos e información bidireccional entre cliente – empresa.

3.6.1.3. Tercera pregunta

¿Estaría usted de acuerdo en que la empresa de telecomunicaciones CLARO implemente una mayor previsión tecnológica (verificaciones avanzadas, autenticación de identidad) en su plataforma de pagos en línea, que eleve la confianza de sus clientes al momento de realizar sus transacciones comerciales?

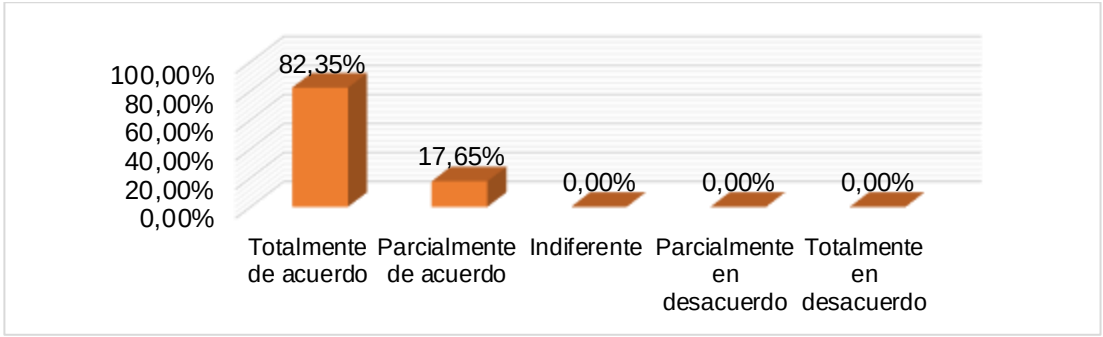


Figura 3.3 Implementación de verificaciones avanzadas y autenticación de identidad

Elaborado por: Autor

Tabla 3.5 Datos de implementación de verificaciones avanzadas

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*14*	*82,35%*
Parcialmente de acuerdo	*3*	*17,65%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	-	-

Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 82,35% (14) de los encuestados, dijeron estar totalmente de acuerdo en que la empresa CLARO implemente una mayor previsión tecnológica en sistemas de prevención tales como verificaciones avanzadas y autenticación de identidad para las transacciones comerciales en línea que realizan sus clientes, con el propósito de salvaguardar los datos e información proporcionada y garantizar el pago u otras actividades derivadas de la relación comercial.

3.6.1.4. Cuarta pregunta

¿En la actualidad, estima que el sistema de prevención de fraudes en los medios electrónicos para transacciones comerciales en línea del sector empresarial de telecomunicaciones del Ecuador, es seguro para los clientes?

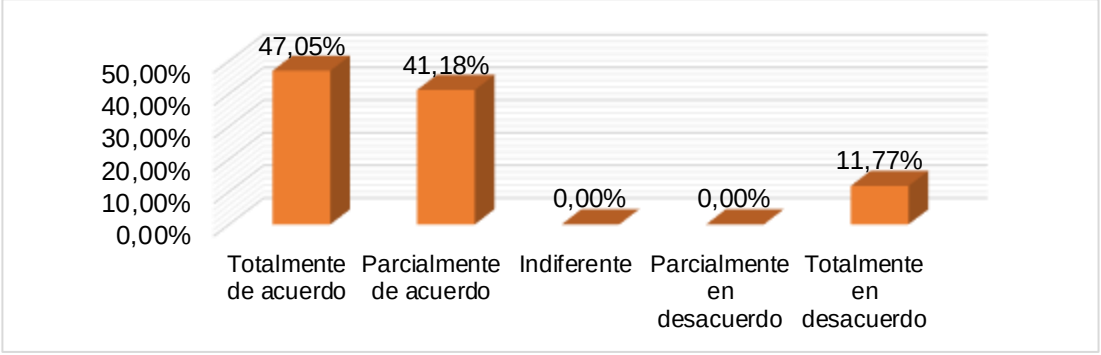


Figura 3.4 Seguridad en sistemas de prevención de fraudes

Elaborado por: Autor

Tabla 3.6 Datos de estimación de seguridad en medios electrónicos de transacciones web

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*8*	*47,05%*
Parcialmente de acuerdo	*7*	*41,18%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	*2*	*11,77%*
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 47,05% (8) de los encuestados, consideran que el sistema de prevención de fraudes en los medios electrónicos de pagos electrónicos del sector de telecomunicaciones es seguro para los clientes, mientras que un 41,18% dijo estar parcialmente de acuerdo con ello; por tanto, esto demuestra que los criterios estar divididos, razón por lo que, se debe fortalecer el sistema de transacciones comerciales electrónicas para que tanto las empresas como los clientes puedan estar seguros de la información personal y económica que intercambian.

3.6.1.5. Quinta pregunta

¿Está usted de acuerdo en que la empresa CLARO debe implementar aplicaciones tecnológicas en todas las transacciones donde interactúa comercialmente el cliente a través de la plataforma web?

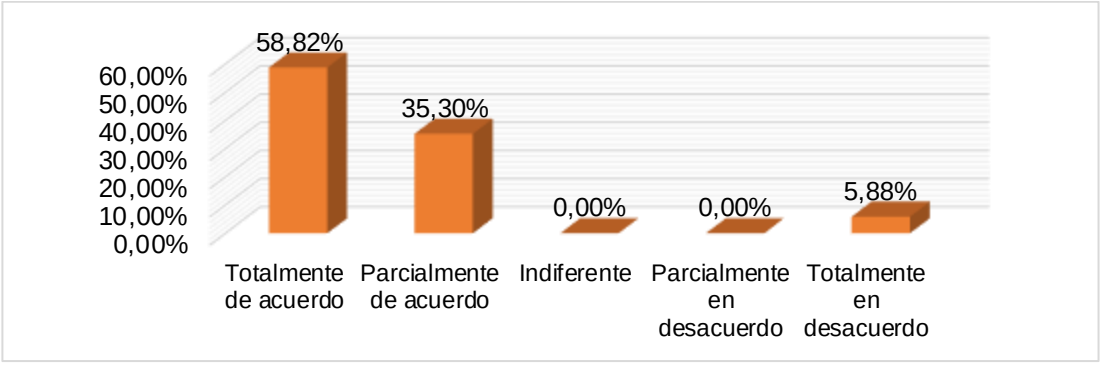


Figura 3.5 Implementación de aplicaciones tecnológicas para prevención de fraudes
Elaborado por: Autor

Tabla 3.7 Datos de implementación de aplicaciones tecnológicas para prevención de fraude electrónico

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*10*	*58,82%*
Parcialmente de acuerdo	*5*	*35,30%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	*1*	*5,88%*
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 58,82% (10) de los encuestados, dijeron estar de acuerdo en que la empresa implemente aplicaciones tecnológicas en todas las transacciones

comerciales donde interactúa el cliente al momento de contratar un servicio o pagar por un producto a través de la plataforma web. Un 35,30% dijo estar parcialmente de acuerdo, debido a que consideran que sí debería ser implementado, pero no para todas las transacciones comerciales, debido a que estiman que solo deber aplicado en aquellas más sensibles a la transmisión de datos personales e información financiera o económica.

3.6.1.6. Sexta pregunta

¿Considera usted que la autentificación biométrica utilizada por la empresa, ha tomado de manera adecuada las conductas de los clientes para ser identificados en verificación de transacción?

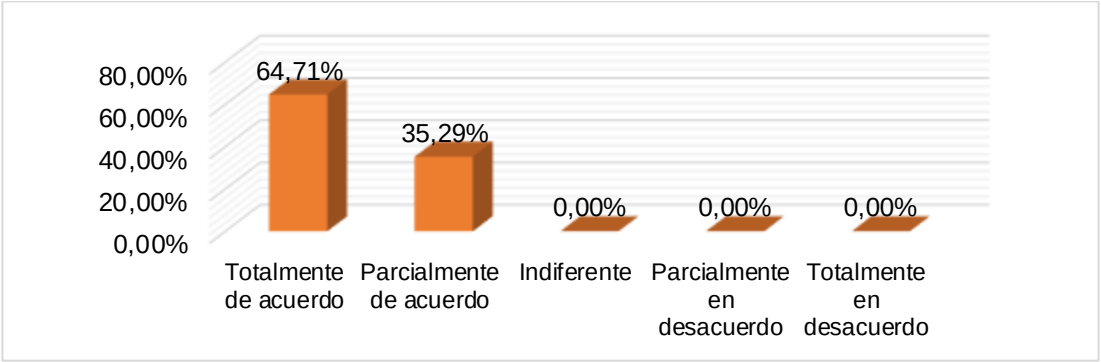


Figura 3.6 Autentificación biométrica para verificación de transacciones

Elaborado por: Autor

Tabla 3.8 Datos de autentificación biométrica en verificación de datos

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*11*	*64,71%*
Parcialmente de acuerdo	*6*	*35,29%*
Indiferente	-	-
Parcialmente en desacuerdo	-	
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 64,71% (11) de los encuestados, se pronunciaron como totalmente de acuerdo en que la autentificación biométrica utilizada por la empresa sí toma de manera adecuada las conductas de los clientes al momento de ser identificados para verificar su transacción comercial. Mientras que el 23,53%

considera estar parcialmente de acuerdo, debido a que estiman que el sistema aún tiene falencias que deben ser tomadas en cuenta al momento de verificar la identidad del cliente y así prevenir fraude en las transacciones comerciales en línea que se realizan en la plataforma web.

3.6.1.7. Séptima pregunta

¿Estima que la aplicación de geolocalización para la identificación de clientes constituye un sistema adecuado para prevenir el fraude en la interacción comercial en línea entre el cliente y la empresa?

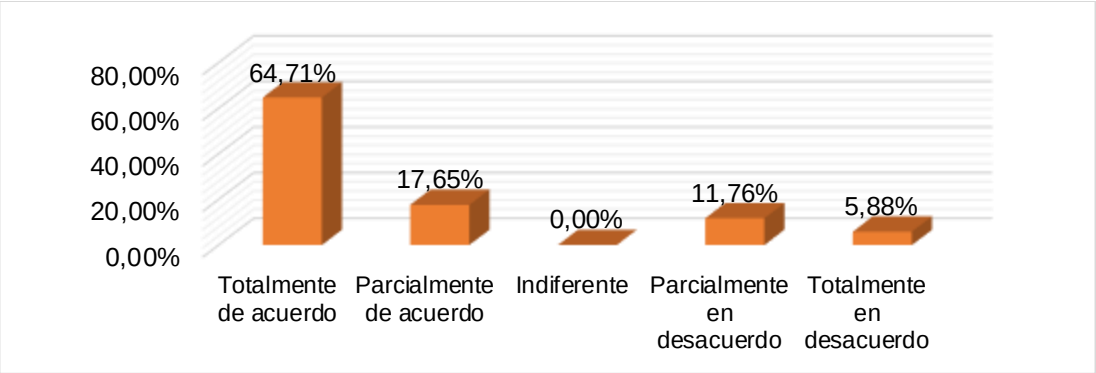


Figura 3.7 Aplicación de geolocalización para identificación de conductas de clientes

Elaborado por: Autor

Tabla 4.9 Datos de autenticación a través de geolocalización

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*11*	*64,71%*
Parcialmente de acuerdo	*3*	*17,65%*
Indiferente	-	-
Parcialmente en desacuerdo	*2*	*11,76%*
Totalmente en desacuerdo	*1*	*5,88%*
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 64,71% (11) de los encuestados, se pronunciaron totalmente de acuerdo en que la aplicación de geolocalización para la identificación de clientes, constituye un sistema adecuado para prevenir el fraude en la interacción comercial electrónica, debido a que este elemento es uno de los verificadores de conducta en características usuales desde donde realiza las transacciones sus clientes, por tanto, una anomalía o cambio en la

geolocalización de la actividad podría representar una alarma para el sistema y así prevenir el fraude.

3.6.1.8. Octava pregunta

¿En concordancia con la pregunta anterior, está usted de acuerdo en que la empresa identifique el dispositivo y realice la detección de proxy que utiliza el cliente al momento de realizar el pago electrónico u otra interacción comercial, con el fin de prevenir el fraude en línea?

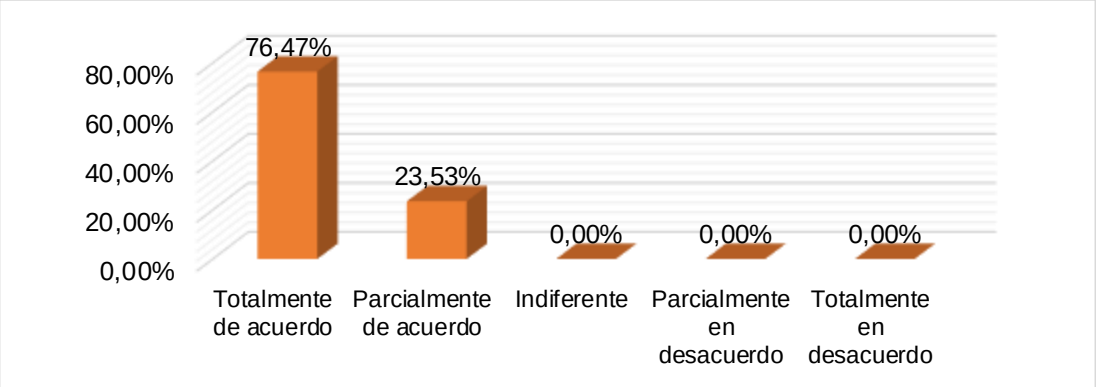


Figura 3.8 Identificación de dispositivo y detección de proxy
Elaborado por: Autor

Tabla 3.10 Datos de identificación de dispositivo y detección de proxy

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*13*	*76,47%*
Parcialmente de acuerdo	*4*	*23,53%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 76,47% (13) de los encuestados, dijeron estar totalmente de acuerdo en que la empresa identifique el dispositivo y realice la detección de proxy, como medida de verificación de datos e información para la ejecución de la transacción comercial en línea, ya que, a partir de la secuencia de conductas de transacciones, la empresa puede detectar cualquier cambio o anomalía en la interacción y así poder prevenir el fraude en el sistema de pagos electrónico.

3.6.1.9. Novena pregunta

¿Estaría de acuerdo en que la empresa implemente un sistema de prevención con firma digital y fichas seguras, para reducir significativamente el fraude en línea y aumentar la confianza de sus clientes en la plataforma web?

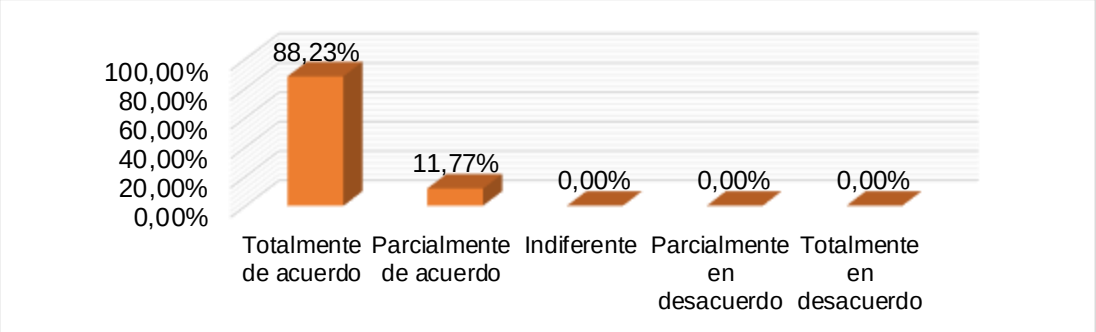


Figura 3.9 Implementación de firma digital y fichas seguras

Elaborado por: Autor

Tabla 3.11 Datos de implementación de firma digital y fichas seguras para identificación de cliente

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*15*	*88,23%*
Parcialmente de acuerdo	*2*	*11,77%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 88,23% (15) de los encuestados, se pronunció como totalmente de acuerdo en que la empresa implemente un sistema de prevención a través de la verificación con firma digital y fichas (tokens) seguras, ya que es otra forma de prevenir el fraude mediante la autenticación de usuarios finales. Sin embargo, no está claro qué tan efectivos pueden ser los sistemas para combatir el fraude en el propósito de detectar la identidad de los estafadores y requiere que los usuarios realicen una tarea adicional con su firma digital; también necesita soporte adicional para validar el número único (clave).

3.6.1.10. Décima pregunta

¿Considera usted que la empresa CLARO debe asumir las observaciones de sus departamentos de telecomunicaciones y comercio electrónico para elevar el nivel de seguridad en el comercio electrónico que realizan a través de su plataforma web?

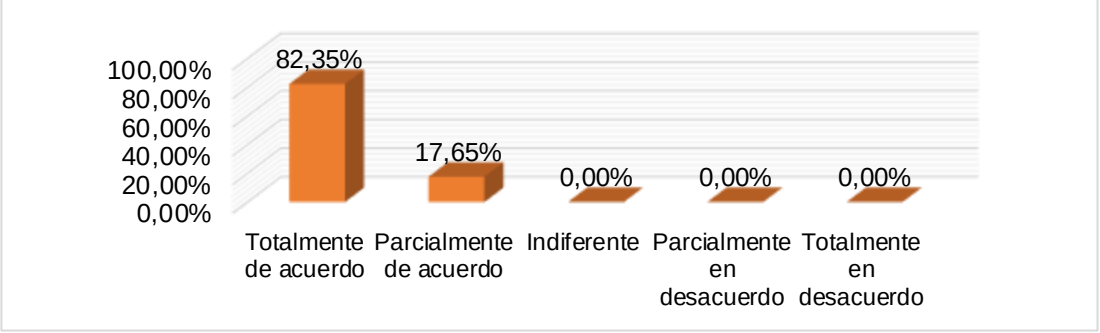


Figura 3.10 Observaciones de departamentos para fortalecer la confianza de clientes
Elaborado por: Autor

Tabla 3.12 Datos de observaciones departamentales para prevención de fraude electrónico

Opción	Descripción	Frecuencia (Fr=17) %
Totalmente de acuerdo	*14*	*82,35%*
Parcialmente de acuerdo	*3*	*17,65%*
Indiferente	-	-
Parcialmente en desacuerdo	-	-
Totalmente en desacuerdo	-	-
Total	***17***	***100%***

Elaborado por: Autor

Análisis

El 82,35% (14) de los encuestados, dijeron estar totalmente de acuerdo en que la empresa CLARO debe asumir las observaciones de sus departamentos de telecomunicaciones y comercio electrónico, debido a que constituyen las áreas principales para solventar y garantizar la protección de datos e información en el sistema de pagos electrónico de la plataforma web, por lo que considerar el criterio departamental podría fortalecer la seguridad web y la seguridad en las transacciones comerciales bidereccionales (cliente –empresa).

3.3.1.6. Síntesis de resultados

A continuación, se presentan los resultados más destacados obtenidos de la encuesta anteriormente revisada, sobre lo cual se fundamenta la discusión y análisis del estudio.

Tabla 3.13 Principales resultados de encuesta

No.	*Pregunta*	**Opción**	**Descripción**	**Frecuencia (Fr=17) %**
1	¿Considera usted que el sistema de pago electrónico que utiliza la plataforma web de la empresa CLARO, debe ser observado continuamente para prevenir el fraude en la interacción comercial en línea?	Totalmente de acuerdo	12	70,59%
2	¿Cree usted que el sector empresarial de telecomunicaciones del Ecuador debe trabajar en conjunto con el Ministerio de Telecomunicaciones y de la Sociedad de la Información para prevenir fraudes en sistemas de pago electrónicos?	Totalmente de acuerdo	10	58,82%
3	¿Estaría usted de acuerdo en que la empresa de telecomunicaciones CLARO implemente una mayor previsión tecnológica (verificaciones avanzadas, autenticación de identidad) en su plataforma de pagos en línea, que eleve la confianza de sus clientes al momento de realizar sus transacciones comerciales?	Totalmente de acuerdo	14	82,35%
4	¿En la actualidad, estima que el sistema de prevención de fraudes en los medios electrónicos para transacciones comerciales en línea del sector	Totalmente de acuerdo	8	47,05%

	empresarial de telecomunicaciones del Ecuador es seguro para los clientes?			
5	¿Está usted de acuerdo en que la empresa CLARO debe implementar aplicaciones tecnológicas en todas las transacciones donde interactúa comercialmente el cliente a través de la plataforma web?	Totalmente de acuerdo	10	58,82%
6	¿Considera usted que la autentificación biométrica utilizada por la empresa ha tomado de manera adecuada las conductas de los clientes para ser identificados en verificación de transacción?	Totalmente de acuerdo	11	64,71%
7	¿Estima que la aplicación de geolocalización para la identificación de clientes constituye un sistema adecuado para prevenir el fraude en la interacción comercial en línea entre el cliente y la empresa?	Totalmente de acuerdo	11	64,71%
8	¿En concordancia con la pregunta anterior, está usted de acuerdo en que la empresa identifique el dispositivo y realice la detección de proxy que utiliza el cliente al momento de realizar el pago electrónico u otra interacción comercial, con el fin de prevenir el fraude en línea?	Totalmente de acuerdo	13	76,47%
9	¿Estaría de acuerdo en que la empresa implemente un sistema de prevención con firma digital y fichas seguras, para reducir significativamente el fraude en línea y aumentar la	Totalmente de acuerdo	15	88,23%

	confianza de sus clientes en la plataforma web?			
10	¿Considera usted que la empresa CLARO debe asumir las observaciones de sus departamentos de telecomunicaciones y comercio electrónico para elevar el nivel de seguridad en el comercio electrónico que realizan a través de su plataforma web?	Totalmente de acuerdo	14	82,35%

Elaborado por: Autor

Habiendo revisado los resultados de la encuesta, es claro que, dentro de la empresa, todos los días se lleva a cabo un número enorme y creciente de pagos con tarjetas de crédito mientras son objeto de actividades fraudulentas, por tal razón, la organización debe detectar rápidamente cualquier comportamiento fraudulento para preservar la confianza de los clientes y la seguridad de su propio negocio. La mayoría de los sistemas de detección de fraude (FDS) emplean algoritmos de aprendizaje automático para aprender los patrones de fraudes y detectarlos a medida que se producen flujos de datos de transacciones.

En particular, este estudio se ha centrado en los FDS que tienen como objetivo detectar fraudes mediante clasificadores que etiquetan las transacciones como fraudulentas o genuinas. En consecuencia, la detección de fraude es particularmente difícil por dos razones: los fraudes representan una pequeña fracción de todas las transacciones diarias y su distribución evoluciona con el tiempo debido a la estacionalidad y las nuevas estrategias de ataque.

Esta situación generalmente se conoce como concepto *drift* y es de extrema relevancia para los FDS que deben actualizarse constantemente, ya sea explotando las muestras supervisadas más recientes u olvidando la información desactualizada que podría no ser más útil pero no engañosa. En

un entorno del mundo real, es imposible verificar todas las transacciones; ya que el costo del trabajo humano limita seriamente la cantidad de alertas, devueltas por el FDS, que pueden ser validadas por el personal de la empresa.

De hecho, el personal departamental verifica las alertas llamando a los titulares de las tarjetas, y luego proporcionan al FDS comentarios que indican si las alertas estaban relacionadas con transacciones fraudulentas o genuinas. Estos comentarios, que se refieren a una pequeña fracción del monto de las transacciones diarias, son la única información en tiempo real que se puede proporcionar para capacitar o actualizar clasificadores. Es así, que estos clasificadores para la prevención en la empresa CLARO, se han estimado en dos etapas (métodos de prueba de identidad y aplicaciones tecnológicas) de sistema de prevención de fraude electrónico:

1. Métodos de prueba de identidad:

    a. Sistemas de verificación de direcciones (AVS)
    b. Verificación avanzada de direcciones (AAV+)
    c. Verificación de edad
    d. Esquemas de seguridad de la tarjeta (CVV)
    e. Sistema de verificación de carga
    f. Sistema de verificación de historial
    g. Sistema de autenticación del consumidor
    h. Sistema de verificación de crédito
    i. Sistema de Verificación de Propiedad de Cuenta
    j. Verificación de Email

2. Aplicaciones tecnológicas

    a. Biometría
    b. Geolocalización
    c. Firma digital
    d. Identificación del dispositivo
    e. Detección de proxy

f. Fichas seguras
g. Tarjetas inteligentes

En base a ello, la encuesta ha demostrado que se puede suponer que las etiquetas del resto de las transacciones se conocen varios días después, una vez que ha pasado cierto tiempo de reacción para los clientes: todas las transacciones que los clientes no informan como fraudes se consideran genuinas. En la etapa de discusión y análisis se discutirá entre muestras de retroalimentación inmediata (es decir, transacciones anotadas con la retroalimentación del investigador) y muestras retrasadas, cuyas etiquetas se obtienen solo después de un tiempo. Esta distinción es crucial para el diseño de un FDS preciso, aunque la mayoría de los FDS en la literatura asumen un etiquetado inmediato y preciso después del procesamiento de cada transacción. Esta suposición demasiado simplificada ignora la interacción alerta – retroalimentación, lo que hace que las supervisiones dependan del desempeño del FDS.

Otra diferencia sustancial entre los entornos del mundo real y los ideales considerados en la literatura es que la principal preocupación de cualquier FDS debe ser devolver un pequeño número de alertas muy precisas, y luego reducir el número de transacciones genuinas (falsos positivos) controlado por investigadores. En la práctica, el FDS óptimo debe ser el que maximice el número de fraudes detectados dentro del presupuesto de alertas que se pueden informar. No obstante, las métricas de rendimiento clásicas consideradas en la literatura son el área bajo la curva (AUC), el costo (es decir, las pérdidas financieras derivadas de una clasificación errónea) y las métricas basadas en la matriz de confusión (por ejemplo, la medida F), que no son necesariamente significativos para la precisión de la alerta.

En este trabajo se demostrará en el capítulo de discusión y análisis que, en un escenario de detección de fraude real, es conveniente manejar los comentarios inmediatos por separado de las muestras supervisadas retrasadas. Las primeras, de hecho, se seleccionan como las transacciones más riesgosas según el FDS en sí, mientras que las segundas se refieren a

todas las transacciones ocurridas. El reclamo se ilustra mejor en el próximo capítulo, donde se analizan dos enfoques de aprendizaje tradicionales para los FDS, es decir, un enfoque de ventana deslizante donde un clasificador se entrena todos los días en las muestras supervisadas más recientes y un enfoque de conjunto donde, todos los días, un nuevo componente reemplaza al más antiguo del conjunto. A partir de ello, se diseña y evalúan dos soluciones diferentes para cada enfoque: en el primero, las retroalimentaciones y las muestras supervisadas retrasadas se agrupan, mientras que en el segundo se entrenan dos clasificadores distintos, basados en retroalimentaciones y muestras retrasadas respectivamente, y luego se agregan los resultados.

La ejecución y análisis que se muestran en el capítulo siguiente en dos conjuntos de datos de tarjetas de crédito del mundo real indican que manejar las retroalimentaciones por separado de las muestras de entrenamiento retrasado puede mejorar sustancialmente la precisión de la alerta en la prevención del fraude electrónico. Por tanto, se motiva este resultado como el hecho de que esta solución garantiza una reacción rápida basados en experimentos adicionales en conjuntos de datos que han sido manipulados para introducir el concepto *drift* en días específicos, que confirman la prevención de fraude en sistemas de pago electrónico.

CAPÍTULO 4: DISCUSIÓN Y ANÁLISIS

4.1. Discusión y análisis

Los FDS se enfrentan a dos desafíos principales: a) manejar flujos de transacciones no estacionarios, es decir, un flujo donde las propiedades estadísticas de fraudes y transacciones genuinas cambian con el tiempo; b) manejar el desequilibrio de clase, ya que las transacciones legítimas generalmente superan ampliamente a las fraudulentas. A continuación, se ofrece una descripción general de los FDS de última generación con un enfoque específico en soluciones para flujos de datos en evolución y desequilibrados.

En la literatura sobre detección de fraude se han propuesto soluciones supervisadas y no supervisadas. Los métodos no supervisados no se basan en etiquetas de transacciones (es decir, genuinas o fraudulentas) y asocian comportamientos fraudulentos a transacciones que no se ajustan a la mayoría. Los métodos no supervisados aprovechan los algoritmos de agrupamiento para agrupar a los clientes en diferentes perfiles e identificar fraudes como transacciones que se apartan del perfil del cliente. Este capítulo se centra en métodos supervisados, los cuales explotan las etiquetas que se asignan a las transacciones para capacitar a un clasificador y, durante la operación, detectar fraudes clasificando cada transacción en la secuencia entrante.

La detección de fraude a menudo se ha considerado como un escenario de aplicación para varios algoritmos de clasificación como redes neuronales, máquinas de vectores de soporte, árboles de decisión y bosque aleatorio. Por tanto, aprender sobre el flujo de transacciones de crédito es un problema desafiante, debido a que las transacciones evolucionan y cambian con el tiempo. El comportamiento de los clientes cambia en las temporadas de vacaciones y pueden aparecer nuevas actividades de fraude; este problema se conoce como concepto *drift* y los algoritmos de aprendizaje que operan en entornos no estacionarios generalmente dependen solo de la información supervisada que está actualizada (por lo tanto, relevante), y eliminan cualquier

muestra de entrenamiento obsoleta. Con mayor frecuencia, la adaptación conceptual se logra entrenando a un clasificador sobre una ventana deslizante de las muestras supervisadas recientes o mediante un conjunto de clasificadores donde se utilizan datos supervisados recientes para entrenar a un nuevo clasificador mientras que las obsoletas se descartan.

Los flujos de transacciones con tarjeta de crédito presentan un desafío adicional, las clases están extremadamente desequilibradas, ya que los fraudes suelen ser inferiores al 1% de las transacciones genuinas. El desequilibrio de clase generalmente se aborda mediante métodos de remuestreo, que equilibran el conjunto de entrenamiento eliminando muestras de la clase mayoritaria (submuestreo) o replicando la clase minoritaria (sobremuestreo). En la práctica, la adaptación conceptual en un entorno desequilibrado a menudo se logra combinando métodos de conjunto y técnicas de remuestreo.

El problema del desequilibrio de clase se aborda mediante la propagación de muestras de entrenamiento en clases minoritarias y submuestreando la clase mayoritaria. Sobre esto, se estima propagar solo ejemplos de la clase minoritaria que pertenecen al mismo concepto utilizando un algoritmo de *K* más cercanos. A través de ello, se crea múltiples conjuntos de entrenamiento equilibrados a partir de un lote utilizando submuestreo, luego aprende un clasificador en cada subconjunto equilibrado y combina todas las predicciones del clasificador para propagar no solo los positivos, sino también las observaciones de la clase negativa que se clasifican erróneamente en el lote anterior para aumentar la definición de límites entre las dos clases.

Todos los marcos de aprendizaje mencionados anteriormente exigen un conjunto de capacitación de instancias recientes con su propia etiqueta de clase de verdad básica. Sin embargo, en un FDS de ejecución real, esto a menudo no es posible porque solo se proporcionan pocas parejas supervisadas recientes de acuerdo con la interacción alerta – retroalimentación. El único FDS que maneja explícitamente el concepto *drift* en los flujos de transacciones, ignora la interacción alerta – retroalimentación,

por lo que, vale la pena señalar que esta interacción alerta – retroalimentación podría recordar un escenario de aprendizaje activo en el que el alumno puede consultar para requerir parejas supervisadas informativas de un gran conjunto de observaciones no marcadas. Desafortunadamente, en un escenario de FDS, esta solución no es factible ya que una fase de exploración, donde los investigadores deben verificar una gran cantidad de transacciones (posiblemente sin interés), no se consideraría aceptable.

4.2. Descripción de escenario

Aquí se formula la descripción del escenario de detección de fraude como una tarea de clasificación binaria donde cada transacción está asociada a un vector de características x, y una etiqueta y. Las características en x pueden ser el monto de la transacción, la identificación de la tienda, la identificación de la tarjeta, la marca de tiempo o el país, así como las características extraídas del perfil del cliente.

Debido a la naturaleza que varía en el tiempo del flujo de transacciones, por lo general, los FDS entrenan (o actualizan) un clasificador K_t todos los días (t). El clasificador K_t: R n → {+, -} se asocia a cada vector de características x ∈ R n, una etiqueta K_t (x) ∈ {+, -}, donde + denota un fraude y - una transacción genuina. Como los fraudes representan una fracción insignificante del número total de transacciones, la clase positiva también se llama clase minoritaria y la clase negativa la clase mayoritaria. En la figura 4.1 se puede apreciar el método de muestras supervisadas en donde el día t incluye:

En general, los FDS operan en un flujo continuo de transacciones porque los fraudes deben detectarse en línea, sin embargo, el clasificador se actualiza una vez al día, para reunir una cantidad suficiente de transacciones supervisadas. Las transacciones que llegan al día t, es decir, T_t, son procesadas por el clasificador K_{t-1} entrenado en el día anterior ($t-1$). Las k transacciones más riesgosas de T_t se informan al investigador, donde k> 0 representa el número de alertas que los investigadores pueden validar. Las alertas informadas At se determinan clasificando las transacciones de T_t de

acuerdo con la probabilidad posterior PK_{t-1} (+ | x), que es la estimación, devuelta por K$_{t-1}$, de la probabilidad de que x sea un fraude. El conjunto de alertas informadas en el día t se define como:

$$A_t = \{x \text{ s.t. } r(x) \leq k\} \quad (1)$$

Donde r (x) ∈ {1, ..., # T$_t$} es el rango de la transacción x de acuerdo con PK_{t-1} (+ | x), y # (·) denota la cardinalidad de un conjunto. En otros términos, la transacción con la probabilidad más alta ocupa el primer lugar (r (x) = 1) y la que tiene la probabilidad más baja ocupa el último lugar (r (x) = #T_t). Luego, los investigadores proporcionarán retroalimentaciones Ft sobre las alertas en At, definiendo un conjunto de k parejas supervisadas (x, y)

$$Ft = \{(x,y),\ x \in At\}, \quad (2)$$

Que representa la única información inmediata que recibe el FDS. En el día t, también recibimos las etiquetas de todas las transacciones procesadas en el día t - δ, proporcionando un conjunto de parejas supervisadas retrasadas D$_{t-\delta}$ = {(x, y), x ∈ $T_{t-\delta}$}, (ver figura 4.1). Aunque el investigador no ha verificado personalmente estas transacciones, se supone por defecto que son genuinas después de δ días, siempre que los clientes no denuncien fraudes. Como resultado, las etiquetas de todas las transacciones anteriores a δ días se proporcionan en el día t.

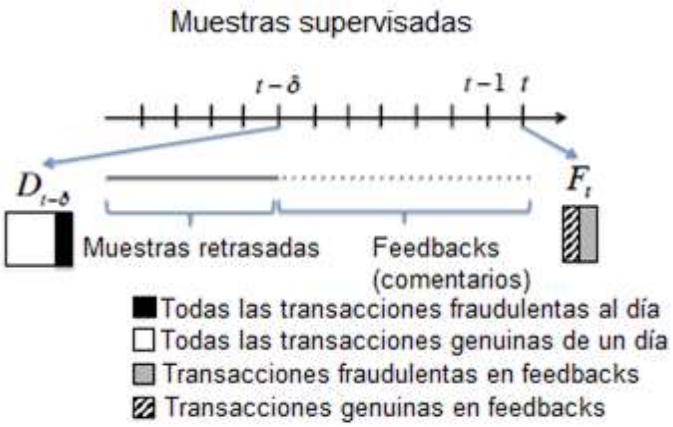


Figura 4.1 Muestras supervisadas disponibles en el dia t.

Elaborado por: Autor

El problema de recibir etiquetas retrasadas también se conoce como latencia de verificación. Vale la pena señalar que esta sigue siendo una descripción simplificada de los procesos que regulan a las compañías que analizan las transacciones con tarjetas de crédito como CLARO. Por ejemplo, normalmente no es posible extraer las alertas At clasificando todo el conjunto T*t*, ya que las transacciones deben pasarse inmediatamente a los investigadores; de manera similar, las parejas supervisadas retrasadas D*t* − *δ* no vienen todas a la vez, sino que se proporcionan con el tiempo. No obstante, se considera que los aspectos más importantes del problema (es decir, la interacción alerta-retroalimentación y la naturaleza del flujo que varía en el tiempo) ya están contenidos en la formulación y que los detalles adicionales harían innecesariamente complejo el establecimiento del problema.

Los comentarios F*t* pueden referirse a fraudes (alertas correctas) o transacciones genuinas (alertas falsas): las alertas correctas son los verdaderos positivos (TP), mientras que las falsas alertas son los falsos positivos (FP). De manera similar, D*t* − *δ* contiene tanto fraude (falso negativo) como transacciones genuinas (verdaderos negativos), aunque la gran mayoría de las transacciones pertenecen a la clase genuina. La figura 4.2 ilustra los dos tipos de pares supervisados que se proporcionan todos los días. El objetivo de un FDS es devolver alertas precisas: cuando se informan demasiados FP, el investigador podría decidir ignorar las próximas alertas. Por lo tanto, lo que realmente importa es lograr la máxima precisión en At. Esta precisión se puede medir por la cantidad. Todos los días se presenta un nuevo conjunto de retroalimentaciones (Ft, Ft − 1, ..., Ft− (δ − 1)) desde los primeros δ días y se produjo un nuevo conjunto de transacciones retrasadas en el día δ (Dt − δ). En esta figura se asume que δ = 7 y los colores se refieren a la notación en la Figura 4.1.

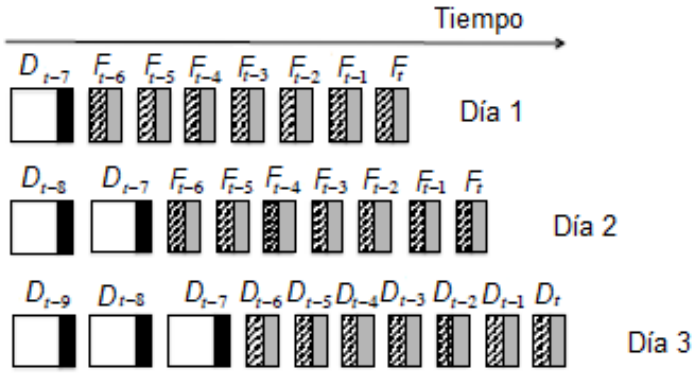


Figura 4.2 Conjunto de retroalimentaciones.

Elaborado por: Autor

$$p_k(t) = \frac{\#\{(x,y) \in F_t \text{ s.t. } y = +\}}{k} \quad (3)$$

Donde $p_k(t)$ es la proporción de fraudes en las principales transacciones k con la mayor probabilidad de ser fraudes.

4.3. Estrategia de aprendizaje

El escenario de detección de fraude descrito sugiere que aprender de las retroalimentaciones F_t es un problema diferente que aprender de muestras demoradas en $D_t - \delta$. La primera diferencia es evidente: F_t proporciona información reciente y actualizada, mientras que $D_t - \delta$ ya puede estar obsoleto una vez que llegue. La segunda diferencia se refiere al porcentaje de fraude en F_t y $D_t - \delta$. Si bien está claro que la distribución de clases en $D_t - \delta$ siempre está sesgada hacia la clase genuina (ver figura 4.2), el número de fraudes en F_t en realidad depende del rendimiento del clasificador K_{t-1}: valores de $p_k(t) \sim 50\,\%$ proporciona retroalimentación F_t donde los fraudes y las transacciones genuinas están equilibrados, mientras que los valores de alta precisión pueden incluso resultar en F_t sesgado hacia los fraudes.

La tercera diferencia, y probablemente la más sutil, es que las parejas supervisadas en F_t no son sorteadas de forma independiente, sino que son seleccionadas por K_{t-1} entre las transacciones que tienen más

probabilidades de ser fraudes. Como tal, un clasificador capacitado en F_t aprende a etiquetar las transacciones que tienen más probabilidades de ser fraudulentas y, en principio, podrían no ser precisas en la gran mayoría de las transacciones genuinas.

Por lo tanto, además del hecho de que F_t y $D_t - \delta$ pueden requerir diferentes métodos de remuestreo, F_t y $D_t - \delta$ también son representativos de dos problemas de clasificación diferentes y, como tales, deben manejarse por separado. A continuación, se presentan dos enfoques tradicionales de detección de fraude (sección 4.3.1.), y se desarrollan más para manejar retroalimentaciones por separado y parejas supervisadas retrasadas (sección 4.3.2.). Los experimentos en la sección 4.3.2. muestran que esta es una estrategia valiosa, que mejora sustancialmente la precisión de la alerta.

4.3.1. Enfoques de clasificación convencionales en FDS

Durante la operación, las retroalimentaciones F_t y las muestras supervisadas retrasadas $D_t - \delta$ pueden explotarse para entrenar o actualizar el clasificador K_t. En particular, se entrena el FDS considerando los comentarios de los últimos días δ (es decir, $\{F_t, F_{t-1}, ..., F_{t-(\delta-1)}\}$) y los pares supervisados retrasados de los últimos días α antes de los comentarios, es decir, $\{D_{t-\delta}, ..., D_{t-(\delta+\alpha-1)}\}$ (ver figura 4.2). A continuación, se presentan dos soluciones convencionales para la adaptación de concepto *drift*, basadas en un algoritmo de clasificación que demuestra una estimación de la probabilidad $P(+|x)$.

a. W_t: un clasificador de ventana deslizante que se actualiza diariamente sobre las muestras supervisadas recibidas en los últimos días $\delta + \alpha$, es decir, $\{F_t, ..., F_{t-(\delta-1)}, D_{t-\delta}, ..., D_{t-(\delta+\alpha-1)}\}$ (ver figura 22).

b. E_t: un conjunto de clasificadores $\{M_1, M_2, ..., M_\alpha, F\}$, donde M_i está entrenado en $D_{t-(\delta+i-1)}$ y F_t está entrenado en todos los comentarios de los últimos δ días $\{F_t, ..., F_{t-(\delta-1)}\}$. La estimación

de probabilidad posterior PE_t (+|x) se estima promediando las probabilidades posteriores de los clasificadores individuales, PM_i (+|x), $i = 1, ..., \alpha$ y PF_t (+|x). Teniendo en cuenta que se utiliza un solo clasificador para aprender del conjunto de comentarios, ya que su tamaño es típicamente pequeño. Todos los días, F_t se vuelve a entrenar teniendo en cuenta los nuevos comentarios, mientras que se capacita a un nuevo clasificador en las nuevas parejas supervisadas retrasadas proporcionadas ($D_{t-\delta}$) e incluidas en el conjunto. Al mismo tiempo, el clasificador más obsoleto se elimina del conjunto.

Estas soluciones implementan dos enfoques básicos para manejar el concepto *drift* que pueden mejorarse aún más mediante la adopción de ventanas dinámicas deslizantes o tamaños de conjuntos adaptativos.

4.3.2. Separación de muestras supervisadas retrasadas de los *feedbacks*

La intuición es que las retroalimentaciones y las transacciones retrasadas deben tratarse por separado porque, además de requerir diferentes herramientas para manejar el desequilibrio de clase, se refieren a diferentes problemas de clasificación.

Por lo tanto, en el día t se entrena a un clasificador específico F_t en las retroalimentaciones de los últimos δ días $\{F_t, ..., F_{t-(\delta-1)}\}$ y denotamos por PF_t (+ | x) su probabilidad posterior. Luego se entrena un segundo clasificador en las muestras retrasadas por medio de una ventana deslizante o un mecanismo de conjunto (ver figura 4.3). Se denota por $W_{D\,t}$ el clasificador entrenado en una ventana deslizante de muestras retrasadas $\{D_{t-\delta}, .., D_{t-(\delta+\alpha-1)}\}$ y por $PW_{D\,t}$ (+|x) sus probabilidades posteriores, mientras que t denota el conjunto de clasificadores α $\{M1, M2, ..., M\alpha\}$ donde cada clasificador individual M_i está entrenado en $D_{t-\delta-i}, i = 1, ..., \alpha$. Entonces, la

probabilidad posterior $P_{ED\,t}(+ \mid x)$ se obtiene promediando las probabilidades posteriores de los clasificadores individuales.

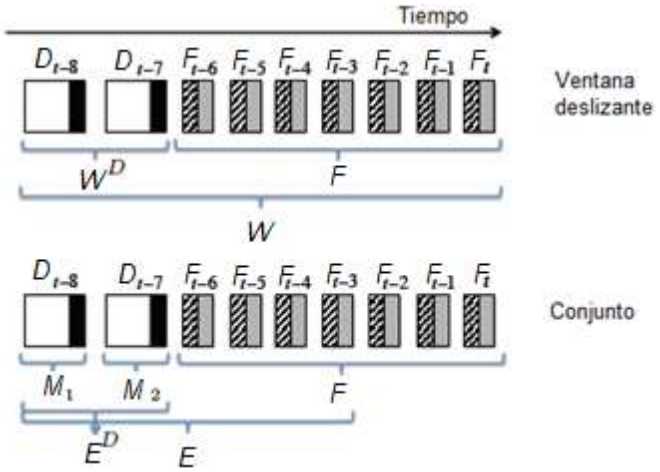


Figura 4.3 Información supervisada utilizada por clasificadores.

Elaborado por: Autor

Cada uno de estos dos clasificadores debe agregarse con F_t para explotar la información proporcionada por los comentarios. Sin embargo, para generar alertas, no se está interesados en los métodos de agregación a nivel de etiqueta sino más bien en el nivel de probabilidad posterior. En aras de la simplicidad, se adopta el enfoque de combinación más directo basado en el promedio de las probabilidades posteriores de los dos clasificadores (F_t y uno entre $W_{D\,t}$ y $E_{D\,t}$). Denotemos por $A_{E\,t}$ la agregación de F_t y $E_{D\,t}$ donde $P_{AE\,t}(+|x)$ se define como:

$$P_{AE\,t}(+|x) = \frac{P_{Ft}(+|x) + P_{ED\,t}(+|x)}{2} \qquad (4)$$

Una definición similar es válida para la agregación de F_t y $W_{D\,t}$ ($A_{W\,t}$), por lo que se debe tener en cuenta que F_t y $W_{D\,t}$ usan conjuntamente el conjunto de entrenamiento de W_t y, de manera similar, los dos clasificadores F_t y $E_{D\,t}$ usan conjuntamente las mismas muestras de entrenamiento de E_t (ver figura 4.3). Sin embargo, los comentarios de W_t representan una pequeña porción de las muestras supervisadas utilizadas para el entrenamiento, por lo tanto, tienen poca influencia en $P_{Wt}(+|x)$, mientras que en la agregación $A_{W\,t}$ su

contribución se vuelve más prominente. De manera similar, F_t representa uno de los clasificadores del conjunto E_t, por lo tanto, tiene en principio la misma influencia que todos los otros clasificadores α entrenados en muestras retrasadas para determinar $P_{Et}(+|x)$.

Los experimentos presentados en la sección 4.4. muestran que manejar las retroalimentaciones por separado de las muestras supervisadas retrasadas proporciona alertas mucho más precisas, y que los FDS que dependen de clasificadores entrenados exclusivamente en retroalimentaciones y muestras supervisadas retrasadas (como AW_t y AE_t) superan sustancialmente a los FDS entrenados en retroalimentaciones y supervisados retrasados muestras agrupadas (como W_t y E_t). A continuación, como ejemplo práctico de la separación de retroalimentaciones de parejas supervisadas retrasadas, se detallan las soluciones específicas basadas en bosques aleatorios que se utilizaron en los experimentos.

4.3.3. Dos FDS específicos basados en bosque aleatorio

Como algoritmo base, los FDS presentados en la sección anterior se usa un bosque aleatorio con 100 árboles. En particular, para WD_t, W_t y para todos los M_i, $i = 1, ..., \alpha$, se utiliza un Bosque aleatorio equilibrado (BRF) donde cada árbol se entrena en una muestra de arranque equilibrada, obtenida submuestreando al azar la clase mayoritaria mientras se conservan todas las muestras de clases minoritarias en el conjunto de entrenamiento correspondiente. Cada árbol de BRF recibe una muestra aleatoria diferente de las transacciones genuinas y las mismas muestras de la clase de fraude en el conjunto de entrenamiento, produciendo un conjunto de entrenamiento equilibrado. Esta estrategia de submuestreo permite aprender árboles con distribución equilibrada y explotar muchos subconjuntos de la clase mayoritaria.

Al mismo tiempo, este método de remuestreo reduce los tamaños de entrenamiento y mejora la velocidad de detección. Un inconveniente del submuestreo es que potencialmente se está eliminando muestras de capacitación relevantes del conjunto de datos, sin embargo, este problema se

mitiga por el hecho de que se aprende de 100 árboles diferentes. El uso de submuestreo permite reequilibrar los lotes sin propagar observaciones de clases minoritarias a lo largo de las secuencias. Propagar fraudes entre lotes debe evitarse siempre que sea posible, ya que requiere acceso a lotes anteriores que no se podría almacenar cuando los datos lleguen a las secuencias. Por el contrario, para F_t que está capacitado en retroalimentaciones, se adapta un bosque aleatorio estándar (RF) donde no se realiza un nuevo muestreo.

4.4. Experimento de estudio

Se consideran dos conjuntos de datos de transacciones con tarjeta de crédito de titulares de tarjetas; el primero (denominado 2018) se compone de transacciones diarias del 5 de septiembre del 2018 al 18 de enero del 2019, según datos tomados de las transacciones de clientes de la empresa CLARO en la ciudad de Guayaquil, el segundo (denominado 2019) contiene transacciones del 5 de agosto al 9 de septimbre del 2019. En el conjunto de datos del 2018 hay un promedio de 160k transacciones por día y alrededor de 304 fraudes por día, mientras que en el conjunto de datos del 2019 hay un promedio de 173k transacciones y 380 fraudes por día. La tabla 4.1 informa algunos detalles adicionales sobre estos conjuntos de datos y muestra que también están muy desequilibrados.

Tabla 4.1 Conjunto de datos

Id	*Día inicial*	*Día final*	*Número de días*	*Número de instancias*	*Número de características*	*% Fraudes*
2018	2018-09-05	2019-01-18	136	21,830,330	51	0.19%
2019	2019-08-05	2019-10-09	44	7,619,452	51	0.22%

Elaborado por: Autor

En los primeros experimentos se procesan ambos conjuntos de datos para evaluar la importancia de separar las retroalimentaciones de las muestras

supervisadas retrasadas. Aunque se espera que estas corrientes se vean afectadas por el concepto *drift* (CD), ya que abarcan un rango de tiempo bastante largo, no se tiene ninguna verdad fundamental para investigar la reacción a la deriva del concepto del FDS propuesto. Para este propósito, se diseña el segundo experimento donde se yuxtapone lotes de transacciones adquiridas en diferentes épocas del año para introducir artificialmente CD en un día específico en el flujo de transacciones.

En ambos experimentos se prueba los FDS construidos en bosques aleatorios presentados, por lo que se considera tanto la ventana deslizante como los enfoques de conjunto y se compara la precisión de la combinación de retroalimentaciones y muestras supervisadas retrasadas juntas (W_t y E_t) contra el aprendizaje de clasificadores separados (F_t, $W_{D\,t}$ y $E_{D\,t}$) que luego se agregan ($A_{W\,t}$ y $A_{E\,t}$). Se recuerda que las alertas son generadas por cada clasificador probado; esto significa que las devoluciones a los clasificadores pueden ser diferentes. Esto debe tenerse en cuenta al comparar diferentes clasificadores, por ejemplo, al comparar W_t y $W_{D\,t}$, la información supervisada proporcionada no es la misma porque, en el primer caso, W_t genera alertas basadas en el segundo W_{Dt}.

Se supone que después de $\delta = 7$ días se proporcionan todas las etiquetas de transacciones (información supervisada retrasada) y que se tiene un presupuesto de $k = 100$ alertas que pueden ser verificadas; por lo tanto, F_t está capacitado en una ventana de 700 retroalimentaciones. Se establece $\alpha = 16$ para que la $W_{D\,t}$ sea entrenada en una ventana de 16 días y la $E_{D\,t}$ sea un conjunto de 16 clasificadores.

Cada experimento se repite 10 veces para reducir la variabilidad de los resultados debido al arranque de los conjuntos de entrenamiento en los bosques aleatorios. El rendimiento del FDS se evalúa por medio del p_k promedio en todos los lotes (cuanto más alto, mejor) y utiliza una prueba t pareada para evaluar si las brechas de rendimiento entre cada par de clasificadores probados son significativas o no.

Se calcula la prueba t pareada en los rangos resultantes de la prueba, por lo que, en la práctica, para cada lote, se clasifican las estrategias de menor a mejor desempeño y luego comparando con cada estrategia con las otras por medio de una prueba t asociada basada en los rangos. Luego se suman los rangos sobre todos los lotes. Más formalmente, sea $r_{s,j} \in \{1, ..., S\}$ el rango de la estrategia s el día j y S el número de estrategias para comparar. La estrategia con la mayor precisión en j tiene $r_{s,j} = S$ y la que tiene la menor tiene $r_{s,j} = 1$. La prueba t pareada compara los rangos de la estrategia a contra b por medio de $r_{a,j} - r_{b,j}$, $j \in \{1, ..., J\}$, donde J es el número total de lotes. Entonces, la suma de los rangos para la estrategia s se define como $PJ_{j=1} r_{j,s}$. Cuanto mayor es la suma, mayor es el número de veces que una estrategia es superior a las demás.

4.4.1. Experimentos en conjuntos de datos de 2018 y 2019

Para evaluar el beneficio de aprender sobre las retroalimentaciones y las muestras retrasadas por separado, primero se compara el rendimiento del clasificador W_t contra F_t, $W_{D\,t}$ y la agregación AW_t. La tabla 4.2 muestra el p_k promedio de todos los lotes para los dos conjuntos de datos por separado. En los conjuntos de datos de 2018 y 2019, AW_t supera a los otros FDS en términos de paquete. Los gráficos de barras de la figura 4.4 muestran la suma de los rangos para cada clasificador y los resultados de las pruebas t emparejadas.

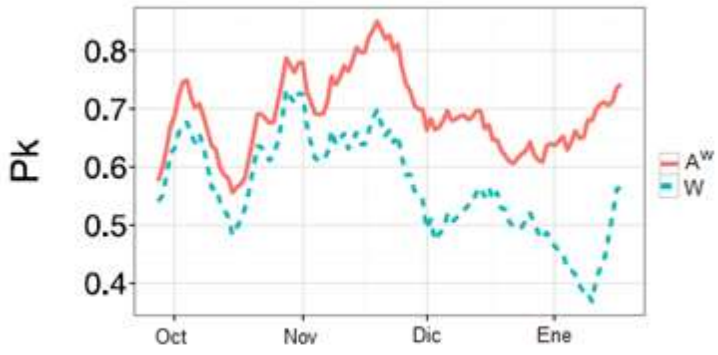
Tabla 4.2 Promedio P_K entre todo el lote

	Conjunto de datos 2018		*Conjunto de datos 2019*	
Calsificador	Media	Estándar	Media	Estándar
F	0.609	0.250	0.596	0.249
W_D	0.540	0.227	0.549	0.253
W	0.563	0.122	0.559	0.256
AW	0.697	0.212	0.657	0.236

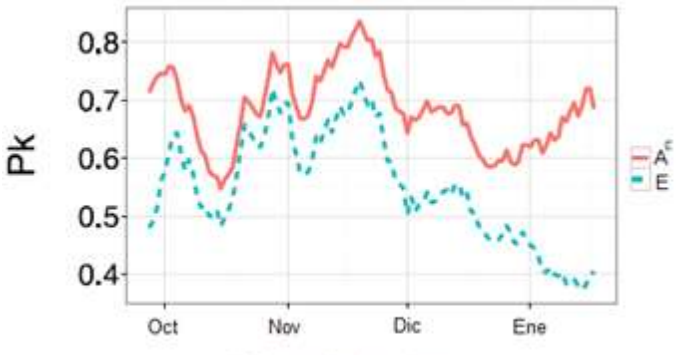
Elaborado por: Autor

La figura 4.5 indica que en ambos conjuntos de datos (figuras 4.4 (a) y 5 (b)) AW_t es significativamente mejor que todos los demás clasificadores. F_t

logra un promedio mayor de pk y un mayor nivel de esfuerzo que W_t y Wt, esto confirma que las retroalimentaciones son muy importantes para aumentar el pk. La figura 4.4 muestra el valor de pk para AW_t y Wt en cada día, promediado en un vecindario de 15 días. Durante diciembre hay una caída sustancial en el rendimiento, que puede verse como un concept *drift* (CD) debido a un cambio en el comportamiento del titular de la tarjeta antes de navidad. Sin embargo, AW_t domina a lo largo de todo el conjunto de datos de 2018, lo que confirma que un AW_t clasificador que aprende sobre las retroalimentaciones y las transacciones retrasadas supera por separado al clasificador Wt entrenado en toda la información supervisada reunida (retroalimentaciones y transacciones retrasadas).



(a) Estrategias de ventana deslizante



(b) Estrategias de conjunto

Figura 4.4 Promedio de Pk por día.

Elaborado por: Autor

Las figuras 4.5 (c), (d) y la tabla 4.3 confirman esta afirmación también cuando los FDS implementan un conjunto de clasificadores. En particular, la

figura 4.4 (b) muestra el pk promedio suavizado de los clasificadores AE_t y E_t. Para todo el conjunto de datos, AE_t tiene mejor p_k que E_t.

Tabla 4.3 Promedio P_K entre todo el lote

	Conjunto de datos 2018		*Conjunto de datos 2019*	
Calsificador	Media	Estándar	Media	Estándar
F	0.603	0.258	0.596	0.271
E_D	0.459	0.237	0.543	0.242
E	0.555	0.239	0.516	0.252
A_E	0.683	0.220	0.634	0.239

Elaborado por: Autor

4.4.2. Experimentos en un conjunto de datos artificial con concepto *drift*

En esta sección, se introduce artificialmente un concepto *drift (*CD*)* abrupto en días específicos por medio de transacciones de postulación adquiridas en diferentes épocas del año. La tabla 4.2 informa los tres conjuntos de datos que se han generado mediante la concatenación de lotes del conjunto de datos 2018.

Tabla 4.4 Base de datos con CD introducción artificial

Id	*Inicio 2018*	*Fin 2018*	*Inicio 2019*	*Fin 2019*
CD1	2018-09-05	2018-09-30	2019-08-05	2019-08-31
CD2	2018-10-01	2018-10-31	2019-09-01	2019-12-09
CD3	2018-11-01	2018-11-30	2019-08-05	2019-08-31

Elaborado por: Autor

La tabla 4.5 (a) muestra los valores de p_k promediados en todos los lotes en el mes anterior al cambio para el enfoque de ventana deslizante, mientras que la tabla 4.5 (b) muestra p_k en el mes posterior al CD. AW_t informa el paquete más alto antes y después del CD. Se obtienen resultados similares con el enfoque de conjunto.

Tabla 4.5 Paquete promedio en el mes antes y después del CD para el enfoque de ventana deslizante

(a) Antes de CD

	CD1		**CD2**		**CD3**	
Clasificador	**Media**	**Estándar**	**Media**	**Estándar**	**Media**	**Estándar**
F	0.411	0.142	0.754	0.270	0.690	0.252
W_D	0.291	0.129	0.757	0.265	0.622	0.228
W	0.332	0.215	0.758	0.261	0.640	0.227
A_W	0.598	0.192	0.788	0.261	0.768	0.221

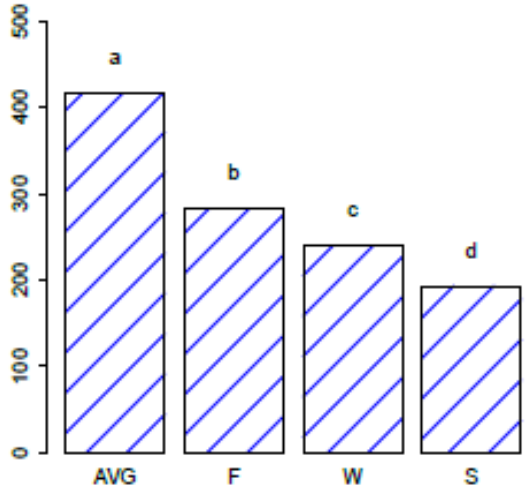
(b) Después de CD

	CD1		***CD2***		***CD3***	
Clasificador	***Media***	***Estándar***	***Media***	***Estándar***	***Media***	***Estándar***
F	0.635	0.279	0.511	0.224	0.599	0.271
W_D	0.536	0.335	0.374	0.218	0.515	0.331
W	0.570	0.309	0.391	0.213	0.546	0.319
A_W	0.714	0.250	0.594	0.210	0.675	0.244

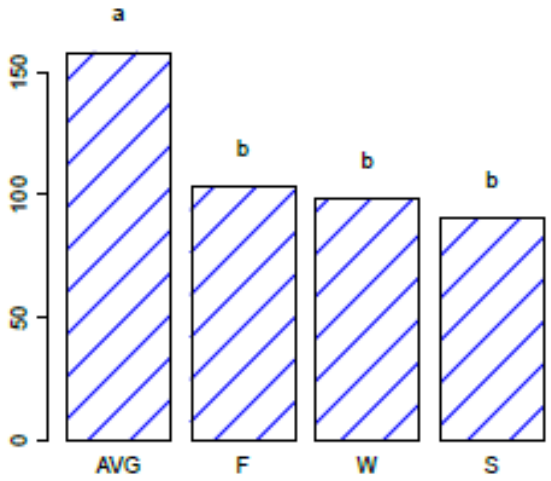
Elaborado por: Autor

La tabla 4.5 (a), (b)), presentan en todos estos experimentos, AE_t también es más rápido que los clasificadores estándar E_t y W_t para reaccionar en presencia de un CD (ver figura 4.6).

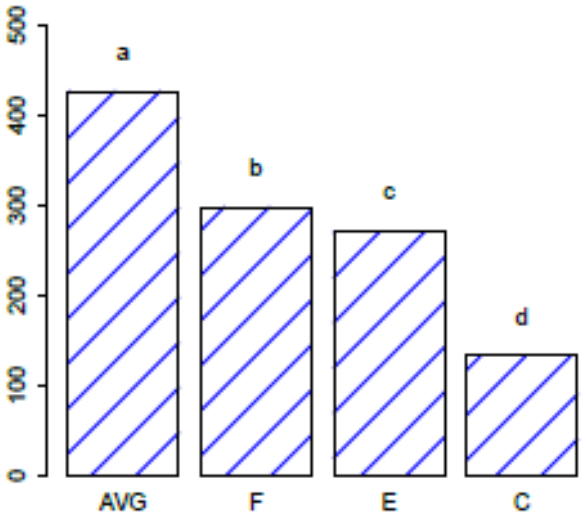
La figura 4.5 muestra la comparacion de estrategias en cunato a su clasificacion, utilizando la suma de los rangos en todos los lotes y la prueba t pareada basada en los rangod de cada lote (los calsificadores que tienen la misma letra en su barra no son significativamente diferentes con un nivel de confianza de 0.95). en ambos conjuntos de datos (2018 y 2019), los clasificadores AW t y AE son significativamente mejores que los demas.



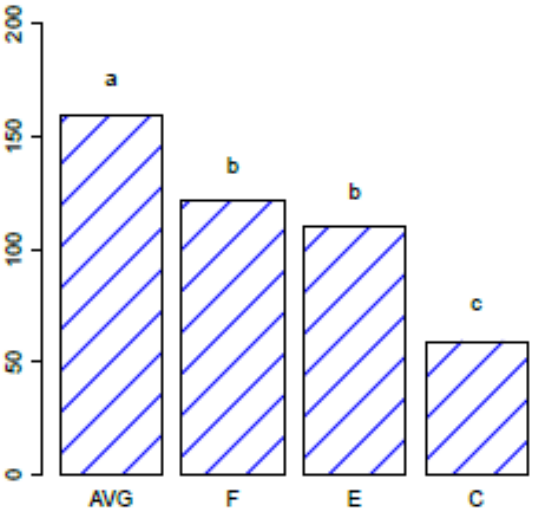
(a) Estrategias de ventanas deslizantes en el conjunto de datos 2018



(b) Estrategias de ventanas deslizantes en el conjunto de datos 2019



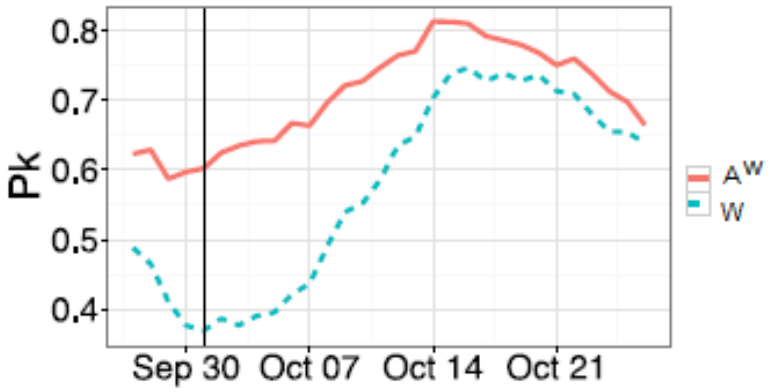
(c) Conjunto de estrategias en el conjunto de datos 2018



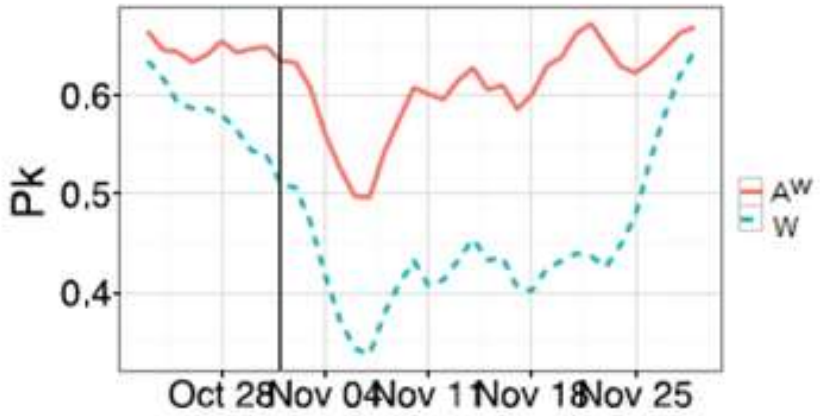
(d) Conjunto de estrategias en el conjunto de datos 2019

Figura 4.5 Comparación de las estrategias de clasificación.

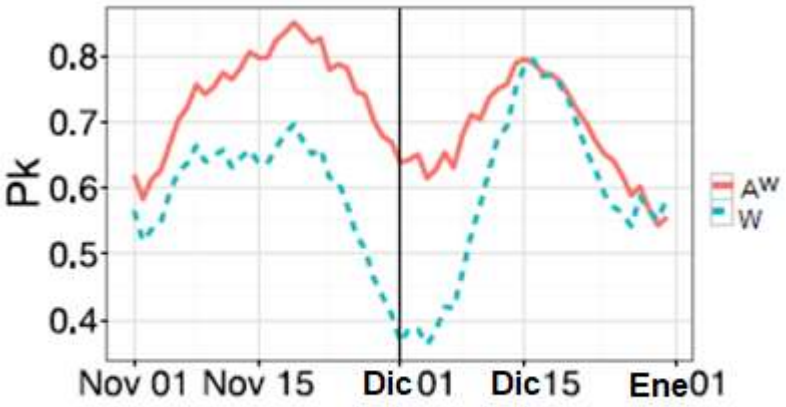
Elaborado por: Autor



(a) Estrategias de ventana deslizante en el conjunto de datos CD1



(b) Estrategias de ventana deslizante en el conjunto de datos CD2



(c) Estrategias de ventana deslizante en el conjunto de datos CD3

Figura 4.6 *Promedio de pk por día.*

Elaborado por: Autor

De acuerdo a los promedio presentados en la figura 4.6, para los clasificadores en conjuntos de datos con deriva de concepto artificial (CD1, CD2 y CD3) suavizados usando un promedio móvil de 15 días. En todos los conjuntos de datos, AW t tiene un pk más alto que Wt. Para el enfoque de conjunto, se muestra solo el conjunto de datos CD3, donde AE t domina a Et para todo el conjunto de datos (se obtienen resultados similares en CD1 y CD2, pero no se incluyen para la compacidad). La barra vertical indica la fecha de la deriva del concepto.

La gran variación de *pk* a lo largo del tiempo refleja la no estacionariedad del flujo de datos. Por lo que, para el conjunto de datos CD1, se tiene un *pk* menor en promedio después del concepto *drift* (CD).

4.5. Discusión

En esta sección se analizan las mejoras de precisión logradas por los clasificadores *AW* t y *AE t* descritas en párrafos anteriores. En primer lugar, se puede notar que el clasificador aprendido en las retroalimentaciones recientes es más preciso que el aprendido en las muestras retrasadas. Esto se hace explícito en las tablas 4.2 y 4.3 que muestran que *Ft* a menudo supera a *WD t* (y *ED t*), y *Wt* (y *Et*). Se considera que *Ft* supera a *WD t* desde *WD t* (resp. *ED t*), por lo que están capacitados en parejas supervisadas menos recientes. En

lo que respecta a la mejora con respecto a W_t (y E_t), la interpretación es que esto se debe al hecho de que W_t (y E_t) están capacitados en todo el conjunto de datos supervisados, lo que debilita la contribución específica de los comentarios (*feedbacks*).

En cambio, los resultados muestran que la agregación evita que la gran cantidad de muestras supervisadas retrasadas dominen el pequeño conjunto de retroalimentaciones inmediatas. Esto se reduce a asignar pesos más grandes a las muestras más recientes que a las antiguas, lo cual es una regla esencial cuando se aprende en entornos no estacionarios. La agregación $A_{W\,t}$ es de hecho una forma efectiva de atribuir mayor importancia a la información incluida en los comentarios. Al mismo tiempo, $A_{E\,t}$ es una forma de equilibrar la contribución de F_t y los modelos α restantes de E_t.

Otra motivación de la mejora de la precisión es que los clasificadores capacitados en retroalimentaciones y muestras retrasadas abordan dos tareas de clasificación diferentes. Por esta razón también, no es conveniente agrupar los dos tipos de muestras supervisadas juntas; finalmente, la agregación presentada en la ecuación 4 proporciona igual peso a las dos probabilidades posteriores P_{Ft} y $P_{WD\,t}$ ($P_{ED\,t}$). Sin embargo, se podrían utilizar esquemas de agregación más sofisticados y eventualmente adaptativos (por ejemplo, no lineales o de apilamiento) para reaccionar al concepto *drift* (CD). De hecho, en un entorno que cambia rápidamente, el peso relativo de P_{Ft} debería aumentar eventualmente, porque $W_{D\,t}$ (o $E_{D\,t}$) podría ser obsoleto y propenso a falsas alarmas.

CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

Se concluye que el análisis de fraude en sistema de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador, ha llevado a cabo la consecución del objetivo de investigación, determinado en analizar el modelo de prevención de fraude en sistemas de pago electrónico, que se realizó en una empresa de telecomunicaciones en la ciudad de Guayaquil (CLARO). Por tanto, en este documento se formalizó un marco que reproduce las condiciones de trabajo de los FDS del mundo real.

El primer objetivo específico, para fundamentar los preceptos literarios y normativos para la prevención de fraude en sistemas de pago electrónico, se determinaron a partir de un escenario de detección de fraude en el mundo real, que es la única información supervisada reciente donde se proporciona sobre la base de las alertas generadas por el FDS y los comentarios proporcionados por el investigador para conocer de primera mano la estrategia utilizada en prevención del fraude electrónico.

El segundo objetivo específico destinado a evaluar una configuración realista de detección de fraude que demuestre el manejo de las etiquetas de prevención por parte de la empresa se llevó a cabo en relación con la consecución del objetivo específico anterior, desarrollando una encuesta en una empresa de telecomunicaciones, donde se determinaron todas las demás muestras supervisadas que reciben un retraso mucho mayor basados en el sistema de prevención de fraude de CLARO. Por ello, la consecución de las encuestas se ha determinado en la interacción alerta – retroalimentación que debe considerarse explícitamente para mejorar la precisión de la alerta y que las retroalimentaciones y las muestras retrasadas deben manejarse por separado al entrenar un FDS realista.

Finalmente, se diseñaron dos FDS sobre la base de un enfoque conjunto y de ventana deslizante como estrategia para la prevención de fraude en medios de pago electrónico. Para este propósito, se consideraron dos

enfoques generales para la detección de fraude: una ventana deslizante y un conjunto de clasificadores. Luego se comparó los FDS que aprenden por separado sobre las retroalimentaciones y las muestras retrasadas contra los FDS que agrupan toda la información supervisada disponible.

Los experimentos realizados en flujos de transacciones del mundo real muestran que la primera estrategia proporciona alertas mucho más precisas que la segunda, y que también se adapta más rápidamente en entornos de concepto *drift* (CD). El trabajo futuro se centrará en investigar mecanismos adaptativos para agregar al clasificador entrenado en retroalimentaciones y el capacitado en muestras retrasadas, para mejorar aún más la precisión de la alerta en flujos no estacionarios.

5.2. Recomendaciones

Las recomendaciones establecidas para el presente estudio, se han considerado en relación a los objetivos de investigación, por lo tanto, el autor recomienda lo siguiente:

- Se recomienda al sector empresarial de telecomunicaciones del Ecuador, específicamente aquellos que desarrollan actividades en la ciudad de Guayaquil, implementar un sistema de prevención de fraude electrónico en sus plataformas web, que considere mecanismos adaptativos para agregar al clasificador entrenado en retroalimentaciones y el capacitado en muestras retrasadas, para mejorar aún más la precisión de la alerta en flujos no estacionarios que constituyan la parte esencial de la funcionalidad en pagos/cobros de la empresa en su plataforma web.

- Se recomienda evaluar de manera continua a los sistemas de prevención de fraude electrónico, con el propósito de mantener actualizada sobre las bases de las nuevas tecnologías y las telecomunicaciones al sector empresarial, y con ello prevenir el fraude en las transacciones comerciales en línea que se realizan

desde el cliente hacia la empresa, con la finalidad de mantener y aumentar la confianza tecnológica en la proporción de datos sensibles (económicos y personales) por parte de los clientes al momento de realizar sus transacciones comerciales.

- Se recomienda observar el presente estudio, como un aporte relacionado al diseño de FDS en base a un enfoque conjunto y de ventana deslizante como estrategia de prevención de fraude en medios de pago electrónica, para fortalecer la confianza bidireccional entre la empresa y el cliente al momento de interactuar en una plataforma web de la empresa.

BIBLIOGRAFÍA

Almagro, L., Urrutia, F. D., & August-Treppel, A. (2018). *Estado de la ciberseguridad en el sector empresarial, bancario en América Latina y el Caribe.* Informe de Seguridad Electrónica y Teconología de la Información, Organización de los Estados Americanos, Programa de Ciberseguridad del Comité Interamericano contra el Terrorismo, San José.

Arahuetes, J. A., & Sequera, J. L. (2015). *Revisión de los Sistemas de Pagos Online en e-commerce.* Investigación científica, Universidad de Alcalá, Programa de Maestría en Sistemas de Información de la Escuela Politécnica Superior, Alcalá de Henares.

Arcenegui, J. A., Castilla, V. O., & Lozano, J. M. (2015). *Propuesta de un modelo para la prevención y gestión del riesgo de fraude interno por banca paralela en los bancos españoles.* Artículo científico, Universidad Loyola Andalucía, Centro de Investigación de Ética y Finanzas del Departamento de Informática , Madrid.

Arqué, G. (Marzo de 2014). Departamento de Análisis Económico, Área de Estudios y Análisis Económico. Nuevos sistemas de pago: un gigante en el horizonte. (L. C. Research, Ed.) *Dossier: Sistemas de pago. Presente y futuro*(3), 2.

Asociación de Examinadores de Fraude Certificados. (2018). *Prevención y análisis de fraude en comercio electrónico en España y América Latina.* Informe sobre seguridad electrónica, ACFE, Austin.

Buku, M. W., & Mazer, R. (11 de Abril de 2017). El fraude en los servicios financieros móviles: Cómo proteger a los consumidores, los proveedores y el sistema. *Revista CGAP, 11*(2), 3.

Burke, E. (2016). *Fraude Interno. Prevención, detección y tratamiento.* Artículo científico, Ágora del Conocimiento, Madrid.

Cámara de Comercio de Guayaquil. (2019). *Distribución de artículos FDS basados en problemas y desafíos entre 1998 - 2018.* Informe anual, CCG, Departamento de Estadística Empesarial, Guayaquil.

Cardenas, C. A., & Rosales, G. (2017). *Fraude Empresarial: Encuesta e Investigación del sector empresarial en prevención de fraude*

electrónico. Investigación científica, Universidad Militar de Nueva Granada, Facultad de Relaciones Internacionales, Seguridad y Defensa, Bogotá.

Consorcio Ecuatoriano de Telecomunicaciones (CONECEL), CLARO. (2019). *Descripción de personal departamental, áreas: Telecomunicaciones y Comercio electrónico.* Informe de Talento Humano, CLARO, Departamento de Administración Corporativa, Guayaquil.

Crespo, G. R., & Toscano, C. G. (2015). *Diseño de una guía para la implementación del proyecto de tarjetas inteligentes aplicado en el sector financiero del Ecuador, basada en el estándar de administración de proyectos desarrollado por el Project Management Institute.* Investigación científica, Universidad de las Américas, Programa de Maestría de Gerencia de Sistemas y Tecnologías de la Información, Quito.

Cricco, M. (2017). *Estructura tarifaria del mercado de pagos electrónico en el Uruguay y otros países de América Latina.* Comisión Económica para América Latina y el Caribe, Oficina de Seguridad Electrónica del Minsterio de Economía y Finanzas de la República Oriental del Uruguay. Santiago: CEPAL.

Donayre, E. F., & Santos, R. E. (2018). *Sistemas de información para la prevención y control de fraude para colaboradores de red de tienda de una entidad financiera del Perú .* Investigación científica, Universidad Tecnológica del Perú, Facultad de Ingeniería de Sistemas y Electrónica, Lima.

Escobar, C. E. (2015). *Estudio de la implementación de una red de dinero electrónico a través de la plataforma celular en el Ecuador, comparándola con el uso de tarjetas de débito.* Pontificia Universidad Católica del Ecuador, Programa de Maestría en Redes de Comunicación. Quito: PUCE.

Espino, C., Fontes, X. M., & Daradoumis, A. (2017). *Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo. Herramientas Open Source que permite su uso.* Investigación científica, Universidad Oberta de Cataluña, Grado de Ingeniería Informática, Barcelona.

Fernández, J. C., Alonso, D., Sotomayor, N., Adib, J., & Schultze-Kraft, J. F. (2017). *Tendencias en medios de pago.* Telecomunicaciones y Energia. Madrid: IndraTecnocom.

Franco, O. C. (2018). *Incidencia de la seguridad en el comercio electrónico de las MiPymes B2C de la Provincia del Guayas y propuesta de un sistema de pago seguro.* Universidad Católica de Santiago de Guayaquil, Sistema de Posgrado. Maestría en Administración de Empresas. Guayaquil: UCSG.

García, L., Cantoral, J., Enríquez, D. A., & Navas, M. A. (2016). *Detección de fraude en transacciones con tarjetas de crédito y débito.* Artículo científico, Universidad del Valle, XVIII Convención Científica de Ingeniería en Telecomunicaciones y Tecnología de la Información, Ciudad de Guatemala.

García, T. R., & Aller, C. F. (2017). *La protección de datos ante el internet de las cosas.* Investigación científica, Universidad Politécnica de Madrid, Programa de Maestría de Ingeniería Técnica en Informática de Gestión , Madrid.

Gilman, L., & Joyce, M. (2014). *Gestionando el riesgo de fraude en el dinero móvil.* Mobile Money for the Unbanked. Nueva York: GSMA.

González, E. F., Romero, G. R., Cruz, D. L., & Ortiz, A. F. (2018). *Detección de fraude en tarjetas de crédito mediante técnicas de minería de datos.* Investigación científica, Universidad Santo Tomás, Trabajo de Grado en Ingeniería Informática, Bogotá.

González, L. J., Barrero, D. S., & Bohórquez, M. P. (2016). *La implementación de la biometría en los medios de pagos electrónicos internacionales realizados con tarjetas de débito y crédito en España y su aplicación en Colombia para mitigar el riesgo reputacional en las entidades financieras.* Investigación científica, Universidad de la Salle de Colombia, Escuela de Finanzas y Comercio Internacional, Bogotá.

González, M. E., & Morán, M. (2017). *Impacto de la implementación del sistema de pagos electrónicos en la competitividad del sistema financiero del Paraguay.* Investigación científica, Universidad Nacional de Asunción, Dirección de Posgrado, San Lorenzo.

Guillén, R. I., & Biarge, V. R. (2019). *Sistemas para detectar fraudes en medios de pago.* Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Informáticos. Madrid: UPM.

Herrera-Semenets, V., & Prado, M. A. (2014). *Análisis de los métodos de detección de fraude en servicios de telecomunicaciones.* Centro de Aplicaciones de Tecnologías de Avanzada, Equipo de Investigación de Minería de Datos. Madrid: CENATAV.

Ibáñez, D. (2018). *Factores que influyen en la aceptación de los pagos móviles en la economía; y un nuevo modelo a probar.* Investigación científica, Universidad de Lérida, Programa Doctoral en Tecnología y Empresa, Lérida.

Jiménez, M. (2018). *Fraude al descubierto. Encuesta Global sobre Crimen Económico.* Informe anual, PwC, Bogotá.

Martínez, M. R. (2015). *Nuevos métodos de pago online, seguridad y confiabilidad.* Investigación científica, Universidad de Cantabria, Máster Oficial en Empresas y Tecnologías de la Información, Santander.

Mock, V., & Lupini, L. (2017). *Servicios de pago: Pago electrónicos más seguros y más innovadores en beneficio de los consumidores.* Boletín sobre Teconologías y Seguridad de Información Electrónica de la Unión Europea, Comisión Europea, Departamento de Normas Técnicas de Regulación de la DSP2 sobre autenticación del cliente, Madrid.

Muguiro, M., & Lafuente, G. (2014). *Sistemas de gestión de pagos electrónicos. Mantenimiento y actualización.* Investigación científica, Universidad Nacional de la Pampa, Escuela de Ingeniería en Sistemas, Buenos Aires.

Noguera, D. F. (2014). *Estudio de las tecnologías de seguridad perimetral informáticas y propuesta de un plan de implementación para la Agencia Nacional de Tránsito.* Pontificia Universidad Católica del Ecuador, Programa de Maestría en Redes de Comunicación. Quito: PUCE.

Ortega, L. P., & Ramos, M. Á. (2016). *E-Commerce y Pago Seguro.* Investigación científica, Universidad Carlos III de Madrid, Departamento de Ingeniería Técnica en Informática de Gestión, Madrid.

Osorio, K. D. (2017). *Detección de fraudes en bodegas de datos basado en los niveles de agregación.* Investigación científica, Universidad

Nacional de Colombia, Departamento de Ciencias de la Computación y de la Decisión, Bogotá.

Otero, R., & Bustamante, J. (2012). *Evoluciín de los mecanismos de control en las telecomunicaciones: Hacia una sociedad tecnificada vigilada*. Investigación científica, Universidad Complutense de Madrid, Departamento de Informática, Madrid.

Pérez, C., Pacheco, B. H., & Salazar, N. (2016). *Beneficios potenciales de un incremento en el uso de los medios de pago electrónicos .* Investigación científica, FEDESARROLLO, Centro de Investigación Económica y Social, Bogotá.

Porlles, D. A., & Luján, L. A. (2017). *Influencia de un sistema integrado de información en la gestión del fraude en una empresa de telecomunicaciones.* Investigación científica, Universidad César Vallejo, Escuela de Postgrado. Programa de Maestría en Gestión de Tecnologías de Información, Lima.

Riera, J., & Ruano, P. (2016). Diseño del sistema organizativo y de control interno para la prevención y detección del fraude. (E. &. Young, Ed.) *Revista de Dirección Empresarial y Tecnología, 23*, 43.

Roa, M. J., García, N., Frías, A., & Correa, L. (2017). *Panorama del dinero móvil en América Latina y el Caribe* (Primera Edición ed.). México D.F., México: Centro de Estudios Monetarios Latinoamericanos; Banco de la República de Colombia.

Ruiz-Capillas, S. H., & Fernández, C. S. (2014). *Random Forest para detección de fraude en medios de pago.* Investigación científica, Universidad Autónoma de Madrid, Escuela Politécnica Superior. Departamento de Ingeniería Informática, Madrid.

Salazar, A. M., & Flores, D. O. (2016). *Análisis de los delitos informático y de telecomunicaciones en el Ecuador bajo las nuevas normas jurídicas.* Universidad de las Fuerzas Armadas, Departamento de Eléctrica y Electrónica. Carrera de Ingeniería Electrónica, Redes y Comunicación de Datos. Quito: ESPE.

Sánchez, J. L. (2019). *Métodos y técnicas de detección temprana de casos de phishing.* Investigación científica, Universidad Oberta de Cataluña,

Programa de Máster en Seguridad de las Tecnologías de la Información y de las Comunicaciones, Barcelona.

Santos, S. G. (2018). *Banca digital y su aplicación en el sistema bancario ecuatoriano.* Universidad San Francisco de Quito, Colegio de Postgrados. Máster en Gerencia Bancaria y Financiera. Quito: USFQ.

Tafra, V. L. (2016). *Evolución de los métodos de pago: Análisis de costos y beneficios de una trasición hacia una sociedad sin dinero en efectivo y diagnóstico de la condición actual de la República de Chile.* Investigación científica, Universidad Técnica Federico Santa María, Escuela de Ingeniería de Industrias, Santiago.

Tracker, S. F., & González, J. (05 de Marzo de 2019). Inteligencia artificial y biometría para la prevención del fraude en la RSA Conference. *Revista Gradiant de Tecnología e Innovación, 3*(2), 4.

Vásconez, T. A. (2017). *Propuesta de rediseño del sistema de pagos interbancarios aplicando arquitectura empresarial en el Banco Central del Ecuador.* Investigación científica, Universidad de las Américas, Facultad de Posgrados. Magister en Gerencia de Sistemas y Tecnologías de la Información, Quito.

Yuste, A. G. (2016). *Delitos informáticos: Malware, Fraudes y Estafas a través de la red y cómo prevenirlos.* Universidad Carlos III de Madrid, Programa de Maestría en Tecnología e Informática. Madrid: UC3M.

Zayas, L. (2016). Señales de alerta para la detección de fraude en empresas. *Revista de Contabilidad y Dirección Empresarial, 23*(7), 63-64.

ANEXOS

Anexo 1. *Encuesta*

UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL
Carrera de Ingeniería en Telecomunicaciones

ENCUESTA PARA INVESTIGACIÓN

DEPARTAMENTOS DE TELECOMUNICACIONES Y COMERCIO ELECTRÓNICO

Instrucciones: Responda las preguntas de encuesta según su criterio respecto a la estructura actual de la empresa en telecomunicaciones y comercio electrónico.

Encuestados: 17 colaboradores

Instrucciones: Elija la opción según su criterio, marcando con una X

No.	PREGUNTAS				
1	**¿Considera usted que el sistema de pago electrónico que utiliza la plataforma web de la empresa CLARO, debe ser observado continuamente para prevenir el fraude en la interacción comercial en línea?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
2	**¿Cree usted que el sector empresarial de telecomunicaciones del Ecuador, debe trabajar en conjunto con el Ministerio de Telecomunicaciones y de la Sociedad de la Información para prevenir fraudes en sistemas de pago electrónicos?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo

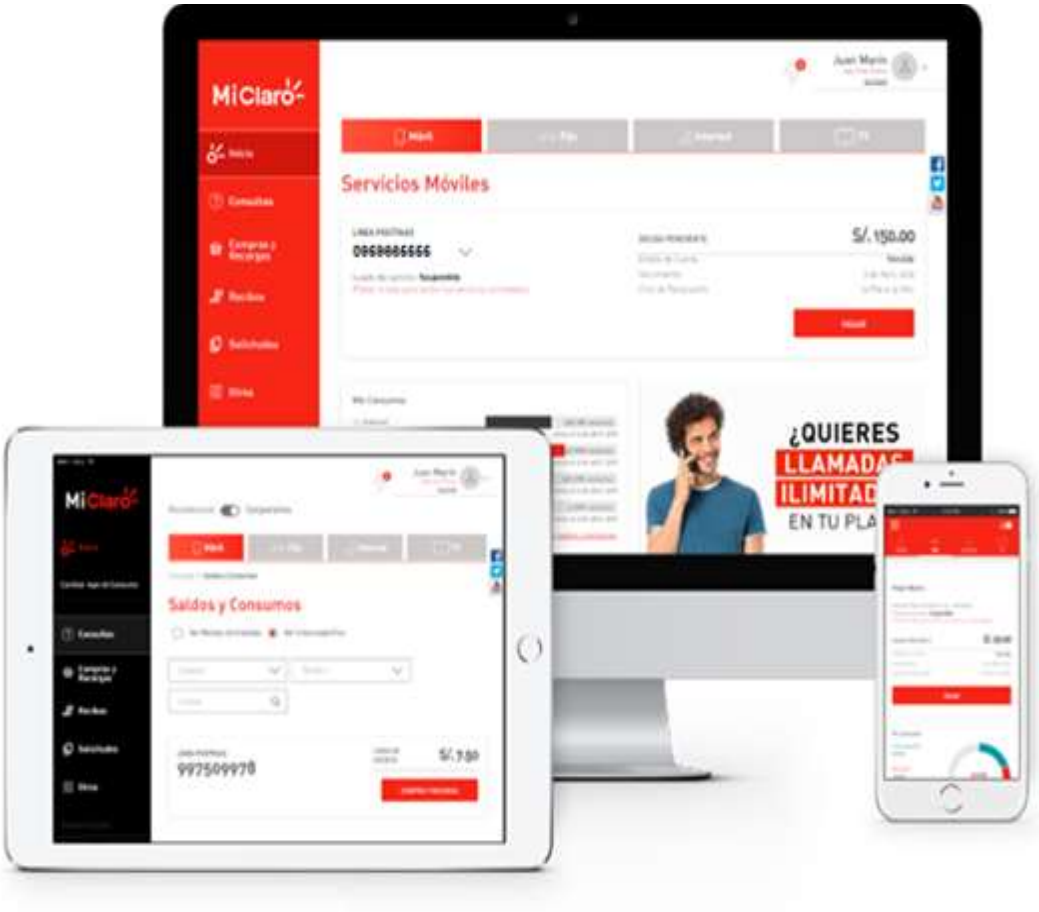
3	**¿Estaría usted de acuerdo en que la empresa de telecomunicaciones CLARO implemente una mayor previsión tecnológica en su plataforma de pagos en línea, que eleve la confianza de sus clientes al momento de realizar sus transacciones comerciales?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
4	**¿En la actualidad, estima que el sistema de prevención de fraudes en los medios electrónicos para transacciones comerciales en línea del sector empresarial de telecomunicaciones es seguro para los clientes?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
5	**¿Está usted de acuerdo en que la empresa debe implementar aplicaciones tecnológicas en todas las secciones donde interactúa comercialmente el cliente con la empresa a través de la plataforma web?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
6	**¿Considera usted que la autentificación biométrica utilizada por la empresa, ha tomado de manera adecuada las conductas de los clientes para ser identificados en verificación de transacción?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
7	**¿Estima que la aplicación de geolocalización para la identificación de clientes constituye un sistema adecuado para prevenir el fraude en la interacción comercial en línea entre el cliente y la empresa?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo

8	**¿En concordancia con la pregunta anterior, está usted de acuerdo en que la empresa identifique el dispositivo y realice la detección de proxy que utiliza el cliente al momento de realizar el pago electrónico u otra interacción comercial, con el fin de prevenir el fraude en línea?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
9	**¿Estaría de acuerdo en que la empresa implemente un sistema de prevención con firma digital y fichas seguras, para reducir significativamente el fraude en línea y aumentar la confianza de sus clientes en la plataforma web?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo
10	**¿Considera usted que la empresa CLARO debe asumir las observaciones de sus departamentos de telecomunicaciones y comercio electrónico para elevar el nivel de seguridad en el comercio electrónico que realizan a través de su plataforma web?**				
	De acuerdo	Parcialmente de acuerdo	Indiferente	Parcialmente en desacuerdo	Totalmente en desacuerdo

Anexo 2. Sistema de pago electrónico, CLARO

Anexo 3. Plataforma web de CLARO, para registro de datos de cliente

Anexo 4. Plataforma web CLARO



Anexo 5. Plataforma web CLARO, para recargas de crédito

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Moscoso Mendoza, Francis Arturo** con C.C: 0925629313, autor del Trabajo de Titulación: **Analisis de fraude en sistemas de pago electronico en el sector empresarial de telecomunicaciones del Ecuador,** previo a la obtención del título de INGENIERO EN TELECOMUNICACIONES en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 2 de marzo del 2020

f. _____

Moscoso Mendoza, Francis Arturo

C.C: 092562931-3

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TÍTULO Y SUBTÍTULO:	Análisis de fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador.		
AUTOR(ES)	Moscoso Mendoza, Francis Arturo		
REVISOR(ES)/TUTOR(ES)	M. Sc. Bastidas Cabrera, Tomas Gaspar		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Educación Técnica para el Desarrollo		
CARRERA:	Ingeniería en Telecomunicaciones		
TITULO OBTENIDO:	Ingeniero en Telecomunicaciones		
FECHA DE PUBLICACIÓN:	2 de marzo del 2020	No. DE PÁGINAS:	101
ÁREAS TEMÁTICAS:	Sistemas Digitales, Telecomunicaciones, Modelos De Seguridad		
PALABRAS CLAVES/ KEYWORDS:	Telecomunicaciones, Pago Electrónico, Fraude En Sistema De Pago, Flujo De Datos, Anomalías, Sector Empresarial		

Resumen:

El presente estudio se ha realizado para analizar el fraude en sistemas de pago electrónico en el sector empresarial de telecomunicaciones del Ecuador, desde un enfoque que promueva el mejoramiento de la confianza en sus clientes y segmento de mercado. Sobre ello, se ha considerado que la mayoría de los sistemas de detección de fraude (FDS) monitorean los flujos de transacciones de tarjetas de crédito por medio de clasificadores que devuelven alertas para los pagos más riesgosos. La detección de fraude es notablemente un problema difícil debido al concepto drift (CD) (es decir, los hábitos de los clientes evolucionan) y el desequilibrio de clase (es decir, las transacciones genuinas superan ampliamente a los fraudes). Además, los FDS difieren de la clasificación convencional porque, en una primera fase, el investigador solo puede proporcionar un pequeño conjunto de muestras supervisadas que tienen tiempo para evaluar solo un número reducido de alertas.

ADJUNTO PDF:	☒ SI	☐ NO
CONTACTO CON AUTOR/ES:	Teléfono: +593980126936	E-mail: francis.mosmen@gmail.com
CONTACTO CON LA INSTITUCIÓN: COORDINADOR DEL PROCESO DE UTE	Nombre: Palacios Meléndez, Edwin Fernando	
	Teléfono: +593-9-67608298	
	E-mail: edwin.palacios@cu.ucsg.edu.ec	

SECCIÓN PARA USO DE BIBLIOTECA	
Nº. DE REGISTRO (en base a datos):	
Nº. DE CLASIFICACIÓN:	
DIRECCIÓN URL (tesis en la web):	