

UNIVERSIDAD TÉCNICA DE MACHALA

Maestría en Software

Asignatura:
Titulación II

Tema:

**Taller N° 6: Estructura del Trabajo de
Titulación: Materiales y Métodos.**

Docente:

Walter Fuertes Díaz, PhD

Estudiante:

Ing. Jimmy Fernando Castillo Crespín

2021-2022

CAPÍTULO II: MATERIALES Y MÉTODOS

2.1 Tipo de investigación seleccionada.

Dado a la revisión sistemática de literatura y al planteamiento de aspectos como el problema, objetivos y variables realizados anteriormente, se determinó que el nivel de profundidad para esta investigación sea de tipo correlacional debido a que se desea determinar qué tipo de correlación es (negativa o nula) entre la variable independiente que son las tecnologías de registros distribuidos en arquitectura de microservicios cloud y la dependiente que son los delitos informáticos por estafas y fraudes en aplicaciones Fintech, es decir, si se implementa los DLT en microservicios cloud, ayudaría a disminuir o no afectar la tasa de delitos informáticos por estafas y fraudes en aplicaciones Fintech. En la figura 17 se ilustra la correlación negativa entre las variables de la presente investigación.

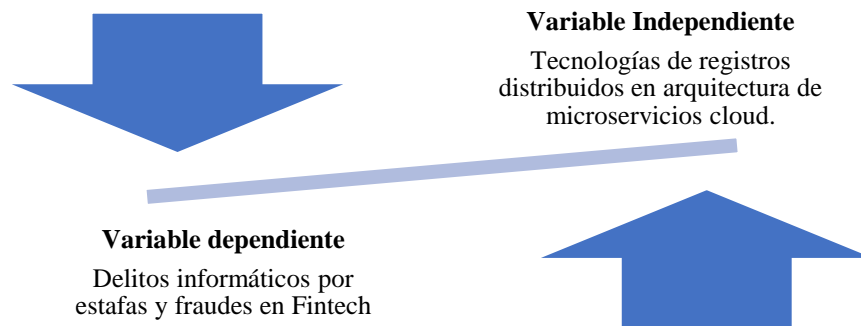


Figura 1: Correlación entre variables de la investigación

Fuente: Elaboración propia.

2.2 Paradigma de investigación realizada.

Esta investigación se la realizó bajo un enfoque cuantitativo debido principalmente a que se manipularán valores numéricos como latencia, cantidad de transacciones entre otros aspectos (ver figura 18) para su posterior medición con técnicas estadísticas y también esta investigación se ajusta al proceso cuantitativo propuesto por Sampieri en su libro de metodología de la investigación científica (ver figura 19) la misma que será utilizada para el cumplimiento de los objetivos propuestos.

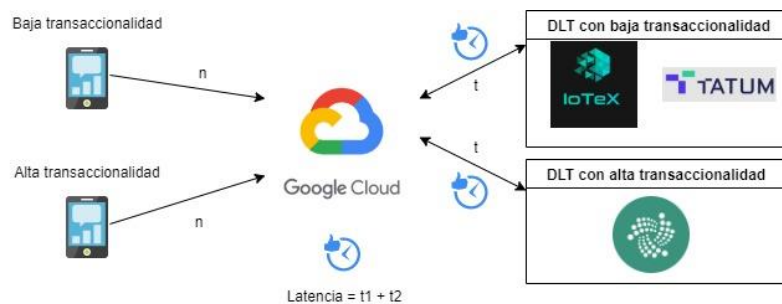


Figura 2: Enfoque cuantitativo a nivel arquitectónico

Fuente: Elaboración propia.

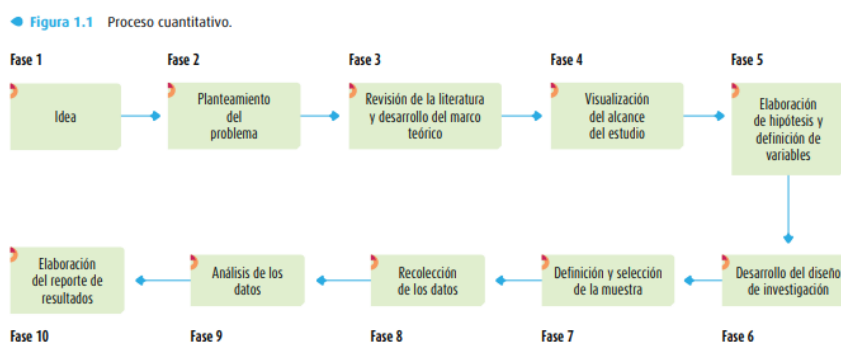


Figura 3: Fases del enfoque cuantitativo

Fuente: [1]

Finalmente, según el grado de uso de la variable, la investigación será de tipo cuasiexperimental por las razones de que en esta investigación no será posible seleccionar la población al azar sino que existirá un grupo determinado de transacciones etiquetadas potencialmente como fraudulentas o estafas, también porque la variable independiente será manipulada en un ambiente de producción con un grupo específico de sujetos obtenidos de las transacciones financieras realizadas por los usuarios de la aplicación de Pagar es Fácil para posteriormente verificar el comportamiento de los DLT ante distintas funcionalidades de la aplicación Fintech donde se han detectado casos de estafas y fraudes. En la tabla 5 se muestra las dos fases de experimentación que se utilizarán para esta investigación, con sus respectivas descripciones y aplicación.

Fases	Descripción	Aplicación
Baja transaccionalidad	Son aquellas transacciones que tomarán un tiempo en concatenarse a los DLT.	Se aplicarán en aquellas funcionalidades que requieran de la utilización de los smart contracts.
Alta transaccionalidad	Son aquellas transacciones que se concatenarán a los DLT ni	Se aplicarán en aquellas funcionalidades que

	bien termine la operación del usuario.	requieran de la utilización de IOTA y NFT de Tatum.
--	--	---

Tabla 1: Fases de experimentación

Fuente: Elaboración propia.

2.3 Población y muestra de la investigación.

La población con la cual se trabajará en la presente investigación serán las transacciones detectadas como potencialmente fraudulentas o estafas realizadas por los usuarios de Pagar es Fácil en los meses de enero, febrero y marzo obtenidos a través de un proceso ETL y debido a que ya se tiene una cantidad fija de transacciones realizadas, se determinó que la población será finita.

El tipo de muestreo será probabilístico simple debido a que todas las muestras tendrán las mismas oportunidades para ser seleccionadas, para eso se utilizó la fórmula de tamaño de muestra para poblaciones finitas mostrado en la figura 20, donde se estableció una población de 202 transacciones, un nivel de confianza del 99% con su respectivo valor z-score de 2.580, un margen de error de 3% y probabilidad de éxito y fracaso del 50%, los resultados de la aplicación de la fórmula se muestran a continuación:

n= muestra = ?

N= población= 202

p= probabilidad a favor = 50%

q= probabilidad en contra = 50%

z= nivel de confianza = 2.580 (99%)

e= error de muestra = 3%

$$n = \frac{N * z^2 * p * q}{e^2 * (N - 1) + z^2 * p * q} = \frac{202 * 2.580^2 * 0.50 * 0.50}{0.03^2 * (202 - 1) + 2.580^2 * 0.50 * 0.50} = 182$$

Figura 4: Cálculo de la muestra

Fuente: Elaboración propia.

Obteniendo un resultado de 182 transacciones que deberán ser analizadas para la comprobación de la hipótesis.

2.4 Método teórico utilizado.

El método teórico seleccionado fue el hipotético-deductivo, debido a que esta investigación plantea una hipótesis y se requiere verificarla o refutarla. Entonces, analizando los pasos que conlleva este método, las mismas serán útiles para cumplir con el objetivo general de la investigación. En la tabla 6 se ilustra el método hipotético-deductivo propuesto por el autor Rodríguez J. [2] el mismo que fue adaptado para esta investigación.

Proceso del método hipotético-deductivo	
Observación	Uso del instrumento de registro anecdótico en las muestras seleccionadas en la investigación obtenidas de las bases de datos de firebase y mysql.
Elaboración de hipótesis	La implementación de tecnologías de registros distribuidos (DLT) en una arquitectura de microservicios cloud disminuye casos de estafas y fraudes en transacciones financieras de una aplicación Fintech.
Deducción de consecuencias	Los casos de fraudes y estafas serán menores o nulos con la implementación de los DLT.
Experimentación	<ul style="list-style-type: none">- Pruebas: Ejecución de los artefactos de software para estudiar la incidencia de una muestra obtenida de la población, en este caso las transacciones financieras, antes y después de la implementación de los DTL.- Encuestas: Aplicados a los usuarios que realicen compras/ventas en los marketplaces de la aplicación Fintech para conocer el nivel de satisfacción del producto obtenido y deducir si se produjeron estafas durante el proceso de compra/venta. El formato de la encuesta se ve reflejado en el anexo 2.
Refutación o verificación	Se muestran los resultados obtenidos a través de estadística inferencial para comprobar o no la hipótesis planteada inicialmente.

Tabla 2: Proceso sistemático del método teórico utilizado

Fuente: Adaptado de [2]

2.5 Métodos empíricos utilizados.

Para esta investigación, se utilizaron los siguientes métodos empíricos:

- **Experimento**

Como se explicó en el punto 2.2, el tipo de investigación seleccionada es cuasiexperimental, por tal motivo se hará uso de este método en dos escenarios de experimentación planteadas en la tabla 5 con la población y muestra explicada en el punto 2.3. Para cada uno de estos escenarios, se diseñará e implementará diferentes arquitecturas de software donde se manipulará la variable independiente, es decir, se hará uso de distintos microservicios con DLT para posteriormente verificar el comportamiento de la variable dependiente.

- **Artefactos**

Se desarrollará en total tres artefactos entregables. El primer artefacto son todos los diseños de arquitecturas de microservicios en Google Cloud siguiendo la metodología ABCDE. El segundo artefacto serán los códigos fuentes correspondientes a las funcionalidades más destacadas donde involucren smart contracts, NFT y registros con IOTA. Finalmente, el tercer artefacto serán la aplicación web y móvil desarrollados para testear las implementaciones.

- **SLR**

Se desarrollará un SLR usando la guía metodológica de B. Kitchenham para posteriormente elaborar un cuadro comparativo donde se seleccionará las tecnologías de registros distribuidos a utilizarse en la investigación.

- **Herramientas utilizadas**

Las herramientas seleccionadas para esta investigación se encuentran dividido en tres grupos tal y como se ilustra en la figura 21, el primer grupo será utilizado para el análisis de datos, siendo seleccionada la herramienta IBM SPSS y validándolo con R Studio. El segundo será utilizado para la recolección de datos, en este grupo se encuentran bases de datos como Firebase y Mysql de donde se obtendrán los registros transaccionales y Google Forms para obtener los resultados de las encuestas aplicados a los usuarios de la aplicación Fintech. El tercer grupo está enfocado a la realización de pruebas donde se encuentra JMeter que será utilizado para realizar pruebas funcionales, testeo de aplicaciones y servicios web; Postman para testeo de endpoints; Mythrill para el análisis de vulnerabilidades en Smart contracts y Kiuwan para el análisis de calidad del código fuente.

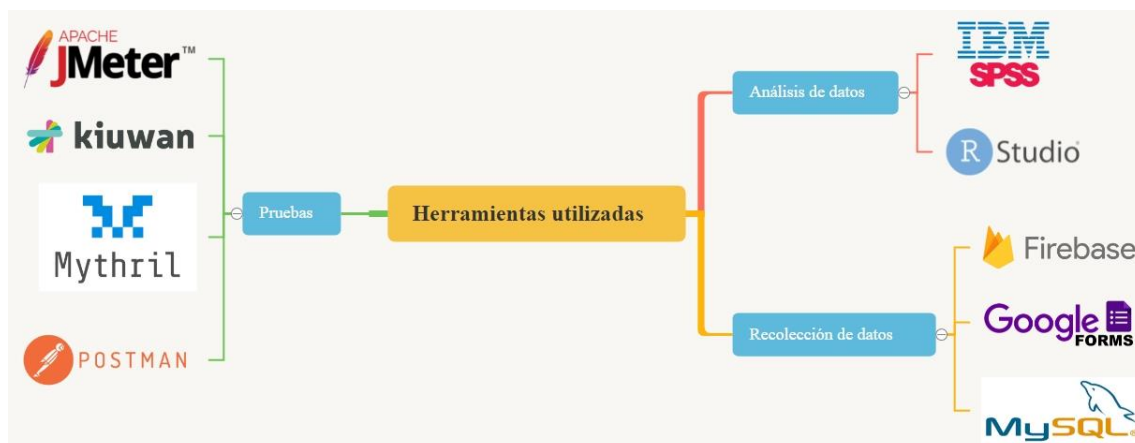


Figura 5: Herramientas utilizadas en la investigación

Fuente: Elaboración propia

2.6 Técnicas estadísticas utilizadas.

Para el análisis de datos cuantitativos se usará la estadística inferencial para probar la hipótesis planteada y debido a que esta investigación es de tipo correlacional, se debe optar por seleccionar alguna técnica correlacional como Pearson, Spearman o Kendall, pero para esto primeramente se debe establecer si la investigación es paramétrica o no paramétrica aplicando pruebas de normalidad.

Las pruebas de normalidad aplicadas fueron las de Kolmogorov-Smirnov y Shapiro-Wilk utilizando el software estadístico SPSS en base a los datos transaccionales obtenidos de los meses de enero y febrero proporcionados por la aplicación Fintech de Pagar es Fácil, en la tabla 7 se ilustra el resultado de esta prueba, donde según los autores Vance & YanYan [3], cuando la población es mayor a 50 se debe optar por seleccionar los resultados de la prueba de Kolmogorov-Smirnov caso contrario seleccionar Shapiro-Wilk y en esta prueba la población o grados de libertad (gl) fue de 59 por tal motivo se seleccionará los resultados obtenidos de la prueba de Kolmogorov-Smirnov. Para conocer si los datos son normales, según la regla de Kolmogorov-Smirnov, si el nivel de significancia es mayor a 0.05 quiere decir que los datos son normales por lo tanto se deberá usar un análisis paramétrico donde se encuentra el coeficiente de correlación de Pearson, caso contrario no son normales y se debe optar por utilizar un análisis no paramétrico como pueden ser el coeficiente de correlación de Spearman o Kendall.

Pruebas de normalidad						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.

Cant_Transacciones	,133	59	,011	,934	59	,003
Tran_Frau_Est	,226	59	,000	,776	59	,000

Tabla 3: Pruebas de normalidad de Kolmogorov-Smirnov y Shapiro-Wilk

Fuente: Elaboración propia

Por tal motivo, según el nivel de significancia de la prueba de Kolmogorov-Smirnov mostrada en la tabla 7 que fue de 0.011 es menor a 0.05 se concluye que la técnica estadística para el análisis de datos cuantitativos que se usará para esta investigación será el coeficiente de correlación de Spearman. Finalmente se realizará una comparación pre y pos testing de los datos y para esto se hará uso del método U Mann-Whitney debido a que esta técnica es usada para análisis de datos no paramétricas.

Bibliografía

- [1] R. Sampieri, Metodología de la investigación, México: McGraw Hill, 2014.
- [2] A. Rodríguez Jiménez y A. O. Pérez Jacinto, «Métodos científicos de indagación y de construcción del conocimiento,» *Revista Escuela de Administración de Negocios*, nº 82, pp. 1-26, 2017.
- [3] V. W. Berger y Y. Zhou, «Kolmogorov–Smirnov Test: Overview,» *Wiley Online Library*, 2014.