

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE CIENCIAS**

### **DESARROLLO DE UN SISTEMA DE ALERTA TEMPRANA SOBRE RIESGO INSTITUCIONAL FINANCIERO: ESTUDIO DE CASOS EN ECUADOR.**

#### **TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO MATEMÁTICO**

#### **PROYECTO DE INVESTIGACIÓN**

**ANDRÉS DAVID BOLAÑOS TOLEDO**

andres.bolanos@epn.edu.ec

**DIRECTORA: DRA. ADRIANA UQUILLAS**

adriana.uquillas@epn.edu.ec

**Quito, enero 2022**

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por Andrés David Bolaños Toledo, bajo mi supervisión.

---

**Dra. Adriana Uquillas**  
**DIRECTORA DE PROYECTO**

## **DECLARACIÓN**

Yo, Andrés David Bolaños Toledo, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

**Andrés David Bolaños Toledo**

## **DEDICATORIA**

Redactar este párrafo no ha sido un gesto fácil imaginando vivir nuevamente los recuerdos que vienen a mi mente y corazón sobre cada instante de apoyo incondicional que he recibido para llegar hasta este momento. El resultado de este episodio de mi vida como estudiante universitario, al igual que muchos otros, es una ofrenda a las personas que han sido mi sostén, mi protección y mis guías durante todo este tiempo, a la Flor, a la Andrea y al Víctor con todo mi amor les dedico este logro el cual hemos construído juntos en nuestra casa que es hogar de alegría, nuestra casa que es hogar de luz, nuestra casa que es hogar de libertad, porque así luchamos para sea todos los días, nada ni nadie nos quitará el amor construido familia.

## AGRADECIMIENTOS

Las palabras de agradecimiento siempre quedarían cortas para expresar la gratitud que tengo con mi amada familia que son los pilares de esta meta cumplida **¡Dios les pague por todo!**

A mis profes que me han sabido guiar y educar en esta magnífica carrera, muchas gracias profe Adriana por compartirme su conocimiento y gracias su paciencia.

A los camaradas de juventud gracias por el tiempo de risa, buen humor, diversión y alegría que intentamos combinar con el esfuerzo, dedicación, fatiga y disciplina que exige la universidad.

Papito gracias por forjarme como hombre, gracias por enseñarme a pescar y no darme la comida en la boca.

Ñañita hermosa gracias por hacerme una persona libre, gracias por hacerme entender que tengo que conquistar el tiempo.

Madre mía, ni sirviéndole de rodillas es resto de mi vida podría compensar todo lo que ha hecho por mí, es usted lo más cerca al cielo que he podido llegar a estar, muchas gracias por todo.

# CONTENIDO

<b>Resumen</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>1 INTRODUCCIÓN</b>	<b>3</b>
1.1 Planteamiento del problema . . . . .	3
1.1.1 Vulnerabilidad Financiera . . . . .	6
1.2 Justificación . . . . .	8
<b>2 MARCO TEÓRICO</b>	<b>13</b>
2.1 Definiciones Financieras . . . . .	14
2.2 Modelo de predicción de eventos raros . . . . .	18
2.3 Machine Learning K-means . . . . .	22
2.3.1 Método El Codo (ElBow) . . . . .	24
2.4 Modelo de Markov . . . . .	26
2.4.1 Ecuación de Chapman-Kolmogorov . . . . .	29
2.5 Modelo Oculto de Markov . . . . .	29
2.5.1 Problemas de interés en los Modelos Ocultos de Markov . . . . .	33
2.5.2 Algoritmo Forward . . . . .	35
2.5.3 Algoritmo Viterbi . . . . .	37
<b>3 METODOLOGÍA Y RESULTADOS</b>	<b>40</b>
3.1 Obtención y tratamiento de la información . . . . .	40
3.2 Actualización/adaptación del modelo estadístico. . . . .	42
3.3 Agrupación de los resultados obtenidos del modelo . . . . .	45
3.4 Predicción de estados de fragilidad bancaria . . . . .	56
3.4.1 Algoritmo Forward . . . . .	58
3.4.2 Algoritmo Viterbi . . . . .	64
<b>4 CONCLUSIONES Y RECOMENDACIONES</b>	<b>76</b>
<b>5 REFERENCIAS BIBLIOGRÁFICAS</b>	<b>78</b>
<b>6 ANEXOS</b>	<b>I</b>

## RESUMEN

El presente Trabajo de Titulación propone establecer un sistema de calificación de riesgo bancario a nivel nacional que permita pronosticar el estado de fragilidad bancaria en el cual las entidades financieras se encontrarán durante períodos futuros, esto en base a índices tanto micro como macroeconómicos a nivel nacional que serán las señales financieras de alerta a ser utilizadas por cada banco para predecir una posible desestabilización financiera.

Dentro del estudio que se llevó a cabo se aplicó la teoría y metodología de los Modelos Ocultos de Markov (HMM por sus siglas en inglés) con mira a decodificar los niveles de riesgo que tendrán las entidades financieras en un cierto período de tiempo en base a una probabilidad inicial de transición de un estado a otro equiprobable. Esto se efectuó mediante el uso del software estadístico R-STUDIO.

**Palabras claves:** Procesos estocástico, Hidden Markov Model (Modelos Ocultos de Markov), Riesgo Financiero, Fargilidad Bancaria, Machine Learning K-means, Programación R-STUDIO.

## **ABSTRACT**

This Degree Project proposes to establish a bank risk rating system at the national level that allows forecasting the state of bank fragility in which financial entities will find themselves during future years based on signals, which is the model to be used, They will be financial indices of each bank and in turn macroeconomic indices at the national level; This is based on an extension of the Markov models called Hidden Markov Models.

Within this study to be carried out, the theory and methodology of Hidden Markov Models will be applied in order to decode the risk levels that financial institutions will have in a certain period of time based on an initial probability of transition of a state to another probable team. This will be done through the use of R-STUDIO statistical software.

**Keywords:** Stochastic Processes, Hidden Markov Model, Financial Risk, Banking Fargility, R-STUDIO Programming.



# 1 INTRODUCCIÓN

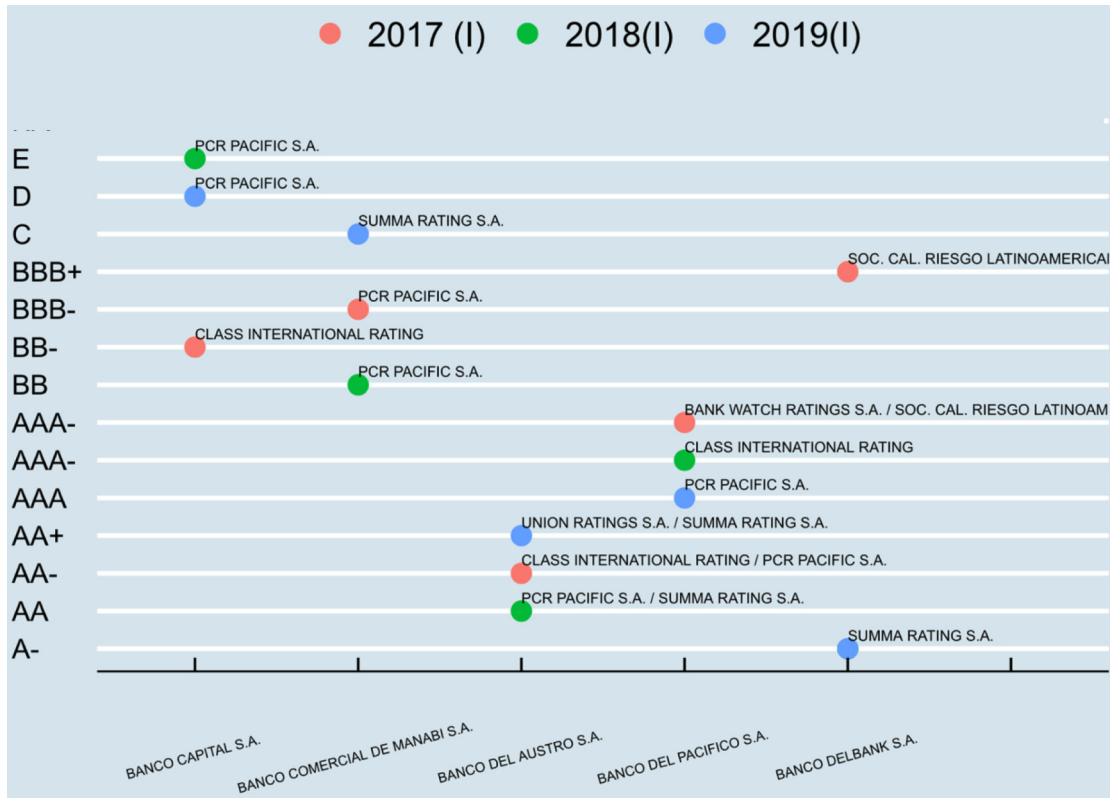
## 1.1 PLANTEAMIENTO DEL PROBLEMA

Ecuador es un país en vías de desarrollo el cual ha experimentado grandes cambios a nivel financiero durante las últimas décadas, un ejemplo claro de ello es la crisis financiera a finales de los años 90, crisis a la cual se acredita gran parte de quiebras bancarias nacionales. Es importante saber que cerca de 30 bancos en los últimos 20 años han quebrado en todo el país, por lo que es considerable tener claro qué tipos de cambios puede soportar el sistema financiero con el fin de que, en caso de existir algún tipo de ‘turbulencia económica’ dentro de los siguientes años, las entidades financieras no corran el riesgo de quebrar pues esto hace que el país experimente problemas socio-financieros como desempleo, migración, entre otros. “Los cambios experimentados por el sistema financiero en general y por el sector bancario en particular en las últimas décadas, a raíz del proceso de globalización, han sido de tal magnitud en el ámbito de las innovaciones financieras que han dado lugar a un nuevo marco financiero, que ha llevado consigo la aparición de la última crisis.” (Pisón, Puime Crespo., 2015 )

La Superintendencia de Bancos es la entidad nacional encargada de regular y fiscalizar el sistema bancario y una de sus labores es de dar a conocer la calificación de riesgo de las instituciones financieras por medio de una escala de niveles ya establecida. Mediante las categorías AAA, AA, A, BBB, BB, B, C, D y E donde a cada una de ellas se pueden asignar

los signos (+) o (-) para indicar su posición relativa dentro de la respectiva categoría, las entidades financieras pueden ver cuán solventes son y observar su capacidad para administrar los riesgos con terceros y para cumplir con sus obligaciones en general. Es así que, nos podemos preguntar si dichos niveles hacen que un banco con mayor calificación sea más competente para afrontar una turbulencia financiera, por ello se genera las siguientes preguntas ¿Las calificaciones publicadas por la Superintendencia de Bancos son suficientes para salvaguardar de una quiebra a una entidad financiera en caso de crisis? ¿Qué factores, tanto micro como macroeconómicos, proporcionan información relevante para que un organismo de control financiero pueda prever un desequilibrio bancario? Tales interrogantes crean un panorama amplio sobre el análisis de fragilidad bancaria, puesto que los índices internos de cada banco no podrían ser suficientes para dar un buen pronóstico de quiebra. Es así que, se deberá analizar qué aspectos macroeconómicos destacan al momento de realizar dicho estudio. Según la Bolsa de Valores de Quito (2019), mediante la Ley de Mercado de Valores afirma: "Las firmas calificadoras de riesgo deberán seguir los siguientes criterios para la evaluación de las instituciones financieras: Acciones, cuotas de fondos de inversión, valores de deuda y titularización."

En la Figura 1.1 se observa un grupo de bancos de los cuales se ha tomado la información obtenida de la Superintendencia de Bancos del primer trimestre de cada año y se evidencia que han cambiado su calificación de riesgo y esto viene de la mano con un cambio de Firma Calificadora en estos períodos de tiempo.



**Figura 1.1:** Cambio de Calificación de riesgos según Firma Calificadora

Si tomamos como ejemplo el *BANCO DEL AUSTRO S.A* podemos ver que en el período de observación (3 años) ha cambiado su calificación de riesgo en tres niveles posibles dentro de la SBS <sup>[1]</sup> tomando en cuenta que en la actualidad la institución que controla los seguros es la *Superintendencia de Compañías, Valores y Seguros*. En la siguiente tabla podemos evidenciar los cambios en su calificación en base a la Firma Calificadora.

Firma Calificadora	Calificación SBS
UNION RATINGS S.A./SUMMA RATINGS S.A	AA+
CLASS INTERNATIONAL RATINGS / PCR PACIFIC S.A	AA-
PCR PACIFIC S.A / SUMMA RATINGS S.A	AA

**Tabla 1.1:** Transición de calificación de riesgo - Banco del Austro S.A.

<sup>[1]</sup> SBS: Super Intendencia de Bancos

Las exigencias hacia las firmas calificadoras de riesgo podrían no ser suficientes al momento de tomar en cuenta factores financieros de carácter macroeconómico, más bien se podría decir que cada una de las empresas abarca diferentes metodologías de evaluación por lo que sería importante tener criterios de evaluación que sean homogéneos para todas las compañías calificadoras de riesgo. Una de las empresas que ha realizado su trabajo en los bancos ecuatorianos llamada Pacific Credit Rating, en su libro Metodología de Calificación de Riesgo de Bancos e Instituciones Financieras Ecuador (2015), afirma: “Conviene advertir que la información y cifras que se utilizan en los análisis, de ninguna manera son auditadas por la calificadora, por lo que ésta no es responsable de errores u omisiones, o por los resultados derivados del uso de dicha información”.

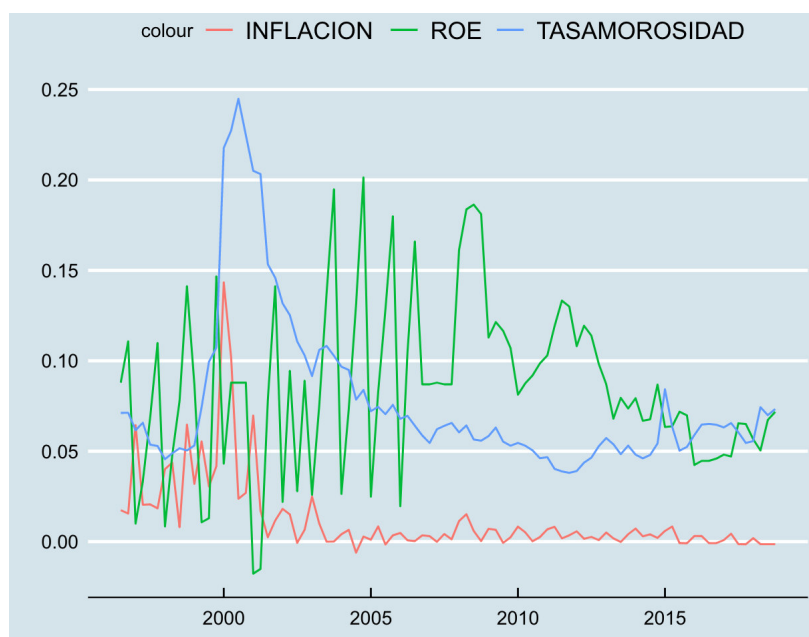
Por ello, la importancia de contar con índices financieros de conocimiento público que sean homogéneos entre cada empresa se hace menester con el fin de garantizar resultados unánimes entre todos los bancos de estudio.

Para validar y verificar el histórico actual de los resultados asignados por las firmas calificadoras hacia los bancos véase **Anexo 6: "Resultados de Calificación de Riesgo vs. Firma calificadora últimos 4 años"** .

### **1.1.1 Vulnerabilidad Financiera**

"El Índice de Vulnerabilidad Financiera es una herramienta que permite medir y analizar el riesgo potencial que puede presentarse por el deterioro en las variables financieras, las cuales en una situación de crisis, están estrechamente vinculadas con las variables reales."(SBS, 2018).

En la Figura 1.2 podemos observar el comportamiento que tienen 3 variables de carácter macro y micro-económico a nivel nacional. Tanto la INFLACIÓN<sup>[2]</sup>, el ROE<sup>[3]</sup> y la TASA DE MOROSIDAD<sup>[4]</sup> del Ecuador en el período 1996-2019 mantienen una tendencia similar entre ellas tomando en cuenta los períodos previo y posterior al Feriado Bancario del año 1999, donde los picos más altos se encuentran en los siguientes años (2000-2001) y posterior a ello buscan una tendencia de estabilidad de manera semejante.



**Figura 1.2:** Tendencias índices vulnerabilidad financiera

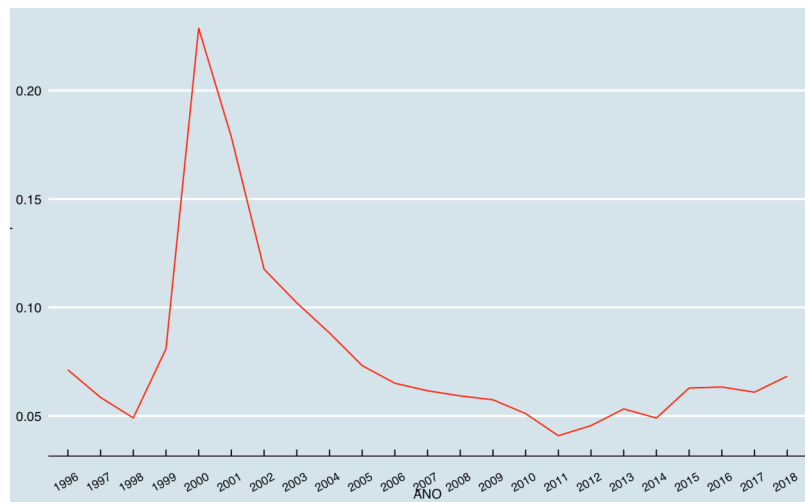
Esto se contrasta de una manera análoga con la tendencia que tiene la TASA DE QUIEBRA a nivel nacional en el mismo período, de esta forma podemos tener un buen indicio de que el comportamiento similar de diferentes variables económicas en el ámbito financiero pueden influir en la quiebra de una entidad financiera. En la siguiente figura podemos ver el

<sup>[2]</sup> INFLACIÓN: Aumento o incremento del nivel general y/o promedio de los bienes y servicios que aportan al crecimiento de la economía.

<sup>[3]</sup> ROE: Indicador utilizado para determinar el nivel de rentabilidad de un banco y de cualquier otra empresa.

<sup>[4]</sup> TASA DE MOROSIDAD: Nivel de riesgo que adquiere una entidad financiera al otorgar un crédito a un individuo y este no cumpla con sus obligaciones de pago

comportamiento histórico de la tasa de quiebra bancaria experimentada de 1996 a 2019 en el Ecuador:



**Figura 1.3:** Tasa de Quiebra Sistema Financiero Ecuatoriano

Entonces, la selección de variables financieras que aporten directamente al análisis de una posible quiebra bancaria es de vital importancia ya que esto generará un modelo económico adecuado para el estudio del comportamiento de los bancos a nivel nacional. “En economía, la dependencia de una variable  $Y$  (la variable dependiente) respecto de otra u otras variables  $X$  (las variables explicativas) pocas veces es instantánea. Con frecuencia  $Y$  responde a  $X$  en un lapso, el cual se denomina rezago...” (Gujarati y Porter, 2010, p. 618)

## 1.2 JUSTIFICACIÓN

Dentro de estudios realizados sobre fragilidad bancaria de Ecuador, no existe uno que tenga un modelo con tal alcance de utilidad para todas las entidades financieras nacionales, pues pronosticar la situación en la que un banco se encontrará en un cierto período de tiempo no es tarea fácil más aun con la falta de datos históricos con los que cuenta el país, por lo que realizar un modelo de esta magnitud es muy importante a nivel nacional. “Es fundamental encontrar indicadores básicos que sirvan de alerta de las crisis bancarias incipientes y

comprender la dinámica de éstas.” (Gonzalez B., 1999). Entre los trabajos que destacan, tanto dentro como fuera del país son:

- \* Barrezueta M. (2010). Aplicación de un modelo de alerta temprana al sistema de bancos privados del Ecuador en el período 2003-2010 (Tesis de maestría). Universidad Andina Simón Bolívar Sede Ecuador. Quito-Ecuador.
- \* Quintanilla C., Tschorne A. Parisi F. (2002). Modelo predictivo de quiebra bancaria, el caso Ecuador (Seminario de título). Universidad de Chile. Santiago de Chile-Chile.
- \* Brito Y. (2001). La administración bancaria: su efecto sobre la crisis financiera del Ecuador (1998-2000), (Tesis de grado). Escuela Politécnica Nacional. Quito- Ecuador.
- \* Mancero J. (2001). La crisis bancaria ecuatoriana en el contexto de las nuevas tendencias financieras internacionales (Tesis de maestría). Universidad Andina Simón Bolívar Sede Ecuador. Quito-Ecuador.

En un contexto general, dentro del análisis de cada uno de los trabajos en mención podríamos decir que carecen de datos históricos para poder obtener resultados robustos, no abarcan las suficientes variables de carácter macroeconómicas, microeconómicas e idiosincráticas o por otro lado su nivel predictivo no alcanza períodos de más de un año.

Como lo afirma Girón A.,(2005): "La solidez de un sistema bancario es resultado de políticas macroeconómicas y estructurales que permiten una estrecha correlación con el sistema financiero." Por ello, el presente estudio abarca el análisis, no solo de los aspectos internos de un banco, si no de los aspectos nacionales de conocimiento público que influyen tanto de manera positiva como negativa sobre la economía de la entidad financiera, a esto se le suma los aspectos idiosincráticos que cada uno de los bancos tiene. "Se puede afirmar la estrecha relación que existe entre estos dos conceptos (crisis financiera y crisis bancaria),

admitiendo que las principales consecuencias de una crisis financiera son la disminución de la inversión productiva y el aumento del desempleo.” Pisón F., Puime F. y Crespo M., (2015).

Los factores económicos a nivel nacional juegan un papel sumamente importante ya que los diferentes riesgos que tiene la banca están estrechamente relacionados con el buen manejo de los recursos nacionales y la correcta manipulación de la economía. “Es necesario analizar la dimensión de la capacidad de transmisión del riesgo, especialmente el riesgo sistemático, que es el que proviene de los factores macroeconómicos y no puede ser eliminado mediante la diversificación, a diferencia del no sistemático y el riesgo sistémico, derivado de los peligros del sistema financiero causados por movimientos en la economía y sujeto a la interdependencia de los mercados financieros.” (Pisón et al., 2015).

Una metodología estable, sólida y con variables de conocimiento público sería una herramienta importante para los análisis internos de los bancos con el fin de garantizar los recursos necesarios en miras de una posible desestabilización interna o externa que pueda generar un riesgo de quiebra bancaria. Por otro lado, debe ser factible poder identificar un evento de desestabilización financiera con anticipación mediante señales de alerta que permitan entender la conducta del banco, es por ello que la finalidad del modelo es poder analizar el comportamiento de las transiciones entre los estados de fragilidad de los diferentes tipos de bancos y sus propiedades probabilísticas.

La estrategia a llevar a cabo consta de 4 secciones principales dentro de todo el procedimiento. Primero el abastecimiento de datos estadísticos financieros obtenidos en plataformas de alcance público a nivel nacional como son las páginas web de la Superintendencia de Bancos y el Banco Central del Ecuador; segundo, se busca tener un modelo que ex-



plique la quiebra en los bancos ecuatorianos, es por ello que un punto de arranque es el trabajo realizado por Uquillas y Flores (2020), donde se establece un modelo econométrico para explicar la quiebra en los bancos ecuatorianos. Se demuestra empíricamente la utilidad de combinar determinantes macroeconómicos, financieros e idiosincráticos y se aborda adecuadamente la problemática existente de tener una muestra pequeña y eventos raros a través del método de estimación de Firth en una regresión Logística. En este trabajo replicamos el modelo propuesto por Uquillas y Flores (2020) en un histórico más actual y con los resultados obtenidos se pasará a un proceso de agrupamiento; tercero, los resultados obtenidos en el modelo econométrico de la sección anterior serán tratados mediante uno de los métodos más conocidos y funcionales de Machine Learning llamado K-Means <sup>[5]</sup> esto es, los resultados de cada una de las observaciones que se han obtenido en el modelo entran en un proceso de agrupación para así asignarlos a grupos de similares características (Estado de Fragilidad Bancaria) los cuales dentro del Método de Clusterización K-means se denotan como Centroides, de esta forma se busca grupos que brinden información sobre el comportamiento de los bancos con semejantes cualidades; cuarto, se hará uso de los Modelos Ocultos de Markov en base a los grupos obtenidos en la sección anterior los cuales los denotaremos como estados, se introduce en este punto información llamada *señal* la misma que dentro del Modelo de Markov tiene una comunicación directa con los resultados del modelo econométrico y que será utilizada como una indicación o signo de aviso para la predicción de los Estados de Fragilidad Bancaria futuros en donde los bancos se encontrarán en períodos futuros.

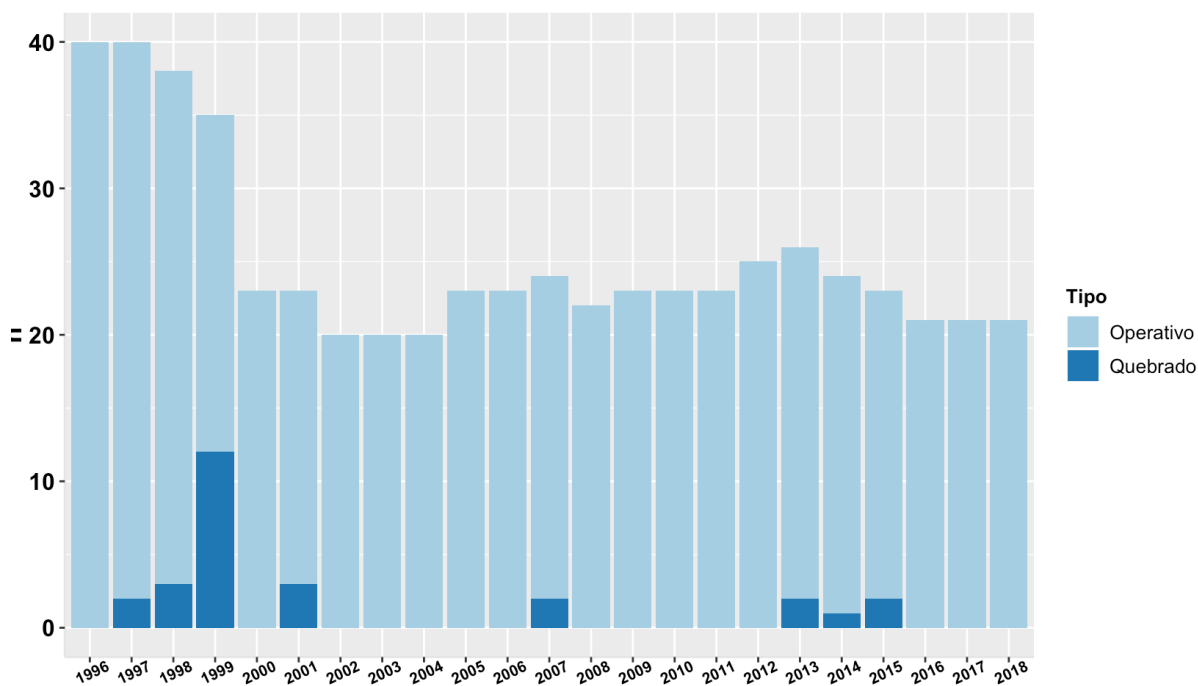
Visto de manera global esta estrategia o metodología busca crear un modelo de alerta temprana que permita ver los estados de fragilidad financiera en los cuales cada uno de

---

<sup>[5]</sup> Método K-Means: Algoritmo de clasificación no supervisada que agrupa datos u objetos en k grupos basándose en sus características.

los bancos nacionales se encuentran en la actualidad y predecir donde se ubicarán en los períodos posteriores, en base a datos de conocimiento público de carácter micro y macro-económicos. En el Capítulo 3 **METODOLOGÍA Y RESULTADOS** se detalla más a fondo los procedimientos a seguir para alcanzar los objetivos que busca el presente trabajo.

El objetivo es poder mitigar el hecho de que las instituciones financieras no cuenten con los recursos necesarios para mantenerse operativas y quiebren, para tener una idea histórica de cuales han sido las consecuencias del manejo inadecuado de la economía en el Ecuador podemos ver de mejor manera la cantidad de bancos que han quebrado durante el periodo de tiempo de estudio versus los bancos que se han mantenido operativos. Se evidencia el golpe financiero que vivió el país entre el año 1999 y 2000 en el llamado “Feriado Bancario”.



**Figura 1.4:** Bancos Operativos vs Bancos Qubrados por año

## 2 MARCO TEÓRICO

En este capítulo se revisan los fundamentos y teorías necesarias para comprender los procesos y metodologías a seguir dentro del presente trabajo el mismo que pretende, mediante el uso de datos de conocimiento público, actualizar el modelo propuesto por Uquillas y Flores (2020), en donde cada uno de los bancos en estudio obtenga una calificación de riesgo con miras a una posible quiebra bancaria, con este conjunto de calificaciones obtenidas por el modelo se crea grupos que tengan mismas cualidades y con ello se busca segmentar las puntuaciones alcanzadas en niveles de riesgo que van desde un nivel manejable (bueno), en términos de indicadores financieros, hasta uno crítico (malo) en el cual se deben generar planes de acción inmediatos y claros con el fin de garantizar la estabilidad financiera del banco. Finalmente, gracias al modelo estadístico, los grupos de estados e información adicional como índices financieros que están fuera del modelo (señales), se pretende predecir para varios períodos futuros si la entidad financiera en estudio se encontrará en el mismo nivel de riesgo o disminuirá/incrementará su riesgo de quiebra en un período de tiempo. A lo largo de este capítulo podremos verificar las diferentes teorías a ser aplicadas dentro de la metodología utilizada.

## 2.1 DEFINICIONES FINANCIERAS

"Es preciso resaltar que durante la historia financiera, la quiebra de bancos ha sido un evento no deseado que ocurre de manera no frecuente en el sistema bancario, donde se define como un acontecimiento raro la ocurrencia de un evento siempre que éste tenga probabilidad de suceder menor al 5 %."(King Zeng 2001)

Entonces lo que se busca en el presente capítulo es un modelo matemático que permita tener el menor sesgo al momento de identificar las variables a ser utilizadas. "El debate real en la literatura sobre el desempeño de los modelos de quiebra está en tono a dos cuestiones: La primera cuestión relacionada a la utilidad de las variables contables utilizadas frente a la variables de previsión del mercado en la predicción de la quiebra. La segunda cuestión es sobre la rigurosidad del modelo matemático utilizado. "(Xu, 2008, p. 535)

El objetivo es garantizar que todas las variables utilizadas en la práctica cumplan con todas las pruebas estadísticas para ser parte de un modelo matemático sostenible en el tiempo, así como lo afirma Pisón (2015) "Todos los estudios de investigación sobre crisis proponen como objetivo final el planteamiento de modelos empíricos que, a partir de las variables citadas, permitan prever o anticipar de algún modo los episodios de desequilibrio." Como se muestra en el presente capítulo, los indicadores en los que se basa el modelo han sido garantizados por pruebas estadísticas que verifican su utilización dentro del modelo, entre estos indicadores de se encuentran:

### **Tasa de morosidad**

La morosidad financiera se refiere al nivel de riesgo que adquiere una entidad financiera al otorgar un crédito a un individuo y este no cumpla con sus obligaciones de pago, de esta

forma se tiene que la tasa de morosidad es nada más que el cociente entre los créditos que están impagos (en mora) y el total de créditos concedidos a los clientes de la entidad financiera.

Así, podemos ver que la expresión matemática aplicada al concepto es:

$$TasaMorosidad = \frac{CarteraEnMora}{TotalCartera}$$

La cartera en mora es nada más que todos los préstamos que nos devengan intereses, ingresos o renta a la entidad financiera.

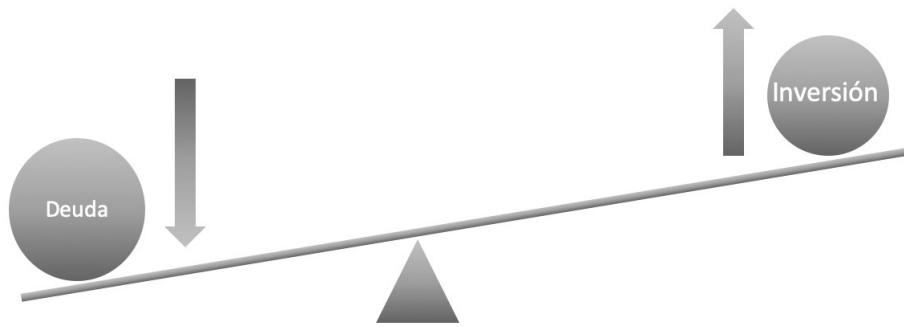
### **Inflación**

La inflación es el aumento o incremento del nivel general y/o promedio de los bienes y servicios que aportan al crecimiento de la economía, en otras palabras la inflación es la reducción del poder adquisitivo de la moneda de intercambio en un ámbito de global.

### **Apalancamiento**

El término apalancamiento se refiere a la potenciación que la variable llamada resultante o de salida (VS) tiene al ser accionada mediante la variable motora o de entrada (VE). El concepto es idéntico a la noción física de palanca, donde la acción de una fuerza motriz es multiplicada por un factor generado por la palanca obteniendo una fuerza resultante mucho mayor. (Buenaventura G. 2002, p 71)

La Figura 2.1 muestra una ilustración de lo que provoca que una empresa o institución financiera incremente su deuda direccionando estos recursos a realizar inversiones a plazo corto o largo dentro del sistema financiero.



**Figura 2.1:** Ilustración Apalancamiento

**Elaboración:** El autor

Dentro del ámbito financiero el apalancamiento se fundamenta en utilizar el endeudamiento para así financiar la operación en contexto de nuevas inversiones y/o proyectos enfocados en el crecimiento institucional. La fórmula matemática que expresa este concepto es:

$$\text{Apalancamiento} = \frac{\text{Pasivo}}{\text{Patrimonio}}$$

### **Spread Financiero**

El Spread está definido como la diferencia que genera un activo financiero entre su valor de adquisición y su valor de venta, es decir el margen que deja la tasa de interés activa versus la tasa de interés pasiva.

$$\text{Spread} = \text{TAR} - \text{TPR}$$

donde,

TAR: Tasa Activa Referencia

TPR: Tasa Pasiva Referencia.

Es útil entender que el *spread* es inversamente proporcional a la liquidez del valor del bien

es decir, a mayor diferencia entre el precio de compra y el precio de venta, la liquidez del valor es menor.

## **ROE**

El ROE (Return On Equity por sus siglas en inglés) es tradicionalmente el indicador más utilizado para determinar el nivel de rentabilidad de un banco y de cualquier otra empresa. Como lo indica el banco BBVA en su boletín *Educación financiera* (2019): "Cuanto mayor sea el ROE, mayor será la rentabilidad que una compañía puede generar en relación con los recursos propios que emplea para financiarse y este se calcula dividiendo el beneficio neto de una empresa entre los recursos propios." La fórmula matemática con la que se calcula el ROE está definida de la siguiente forma:

$$ROE = \frac{UtilidadNeta}{Patrimonio}$$

## **Dolarización**

"Un concepto amplio de dolarización es el que se refiere a cualquier proceso donde el dinero extranjero reemplaza al dinero doméstico en cualquiera de sus tres funciones. En efecto, la dolarización de una economía es un caso particular de sustitución de la moneda local por el dólar estadounidense como reserva de valor, unidad de cuenta y como medio de pago y de cambio." (BCE, 2019)

## 2.2 MODELO DE PREDICCIÓN DE EVENTOS RAROS

### Modelo de Firth

El modelo de regresión logística bajo estimación propuesta por Firth y usado en el trabajo de Uquillas y Flores (2020), es una adaptación del modelo común de regresión logística que busca modificar la función general del score <sup>[1]</sup> de manera que al utilizarla se encuentre un estimador insesgado en la predicción de eventos raros o atípicos en una muestra finita de observaciones.

"Los eventos raros son generalmente de baja posibilidad de ocurrencia, mientras que son extremadamente grandes (o extremadamente pequeños) en amplitud, generalmente tienen el potencial de causar efectos gigantes. Desafortunadamente, es bastante difícil entender el mecanismo detrás de estos comportamientos. En cada uno de estos sistemas también muchos factores interactúan de forma no lineal entre sí, lo que hace que sus resultados sean erráticos, inestables o caóticos "(Zhao, Shang, Lin 2016: 132).

El enfoque y análisis de estos eventos lo aclara King Zeng (2001) al afirmar que:

Entre los problemas que genera la predicción de eventos atípicos podemos citar al método *logit* que es generalmente usado y es sub-óptimo en muestras finitas de datos de eventos raros, lo que lleva a errores en la misma dirección que los sesgos en los coeficientes. Estos problemas serán inofensivos en algunas aplicaciones más aún no siempre pueden ser omitidos. Una segunda fuente de dificultades para analizar eventos raros radica en la recopilación de datos, en base a una fuente de recursos fija siempre existe una compensación

---

<sup>[1]</sup> Score: Herramienta o instrumento para medir el comportamiento financiero a través de una calificación.



entre recopilar más observaciones e incluir mejores variables o adicionales. En los datos de eventos raros, el miedo a recopilar conjuntos de datos sin eventos (y, por lo tanto, sin variación en  $Y$ ) ha llevado a los investigadores a elegir un gran número de observaciones con pocas variables explicativas y, en la mayoría de los casos, mal medidas. Esta es una opción razonable, dadas las limitaciones percibidas, pero resulta que existen estrategias de recopilación de datos mucho más eficientes. Por ejemplo, los investigadores pueden recopilar todos los datos (o todos los disponibles) y una pequeña muestra aleatoria de ellos y no perder consistencia o incluso mucha eficiencia en relación con la muestra completa. Este resultado cambia drásticamente el equilibrio óptimo entre más observaciones y mejores variables, lo que permite a los académicos concentrar los esfuerzos de recopilación de datos donde más importan. (p 138)

Para entrar en la teoría del modelo a ser utilizado de regresión logística de Firth, es importante dar a conocer y tener los fundamentos teóricos sobre el modelo de regresión logística estándar con variable dependiente binaria. Así bien, al ser una variable binaria estamos hablando que sigue una distribución de Bernoulli con su media en el intervalo  $[0, 1]$ , i.e.,

$$E(Y) = P(Y = 1)$$

De esta forma tenemos que la función de  $Y$  está dada de la siguiente forma:

$$Y_i = \beta_0 + \beta_1 X_{i,1}, \dots, \beta_k X_{i,k}, \forall i = 1, \dots, n.$$

Ahora bien, tomaremos el modelo  $Y_i^* = Y_i + \mu_i$ , donde  $\mu_i$  es el error del modelo y de esta

forma tenemos el modelo dicotómico de la variable  $Y$  expresado de la siguiente forma:

$$Y_i = \begin{cases} 1, & \text{si } Y_i^* = X_i\beta + \mu_i > 0 \\ 0, & \text{si } Y_i^* = X_i\beta + \mu_i < 0 \end{cases}$$

Y con ello, el modelo probabilístico viene dado por:

$$\begin{aligned} E(Y_i|X_i) &= P(Y_i = 1|X_i) \\ &= P(Y_i^* > 0) \\ &= F(\beta_0 + \beta_1 X_{i,1}, \dots, \beta_k X_{i,k}) = F(Z_i), \end{aligned}$$

de donde podemos verificar que mientras aumenta el valor de  $Y_i^*$ , la probabilidad de que ocurra el evento se elevará. Finalmente, la función de distribución logística se expresa de la siguiente forma, en donde la probabilidad de éxito es:

$$F(Z_i) = \frac{e^{Z_i}}{1 + e^{Z_i}}$$

### Estimación por Máxima Verosimilitud

La estimación de máxima verosimilitud es análoga a la dada para la distribución de Bernuolli debido a sus estados de éxito o fracaso que en nuestro caso viene a ser quiebra o no quiera, así la función de densidad conjunta de  $y_i, \forall i = 1, \dots, n$  está expuesta por:

$$P(Y_i = y_i | p_i) = \prod_{i=1}^n (F(X_i\beta))^{Y_i} (1 - F(X_i\beta))^{1-Y_i},$$

donde al aplicar el logaritmo se obtiene la función de verosimilitud para una muestra de

tamaño  $n$ ,

$$L(\beta) = \sum_{i=1}^n l_i(\beta)$$

donde  $l_i(\beta) = \sum_{i=1}^n [Y_i * \log(F(X_i\beta)) + (1 - Y_i)\log(1 - F(X_i\beta))]$

Finalmente, es importante recalcar que los estimadores obtenidos por este método son consistentes pero no asintóticamente insesgados, por lo que tendrán falencias al trabajar con muestras pequeñas y más aún analizar eventos atípicos.

### **Modelo de regresión Logística de Firth**

El modelo de Firth se basa en el método de máxima verosimilitud del modelo estándar pero, para disminuir el sesgo, este adhiere un *elemento penalizador* a la ecuación que en la práctica viene dado de la siguiente forma

$$L^*(\beta) = L(\beta) |i(\beta)|^{\frac{1}{2}}$$

Así como lo afirma Leitgöb (2013): El término penalizador en la función de verosimilitud tiene dependencia del tamaño de la muestra tomada y la cantidad de eventos atípicos, por ello la estimación a través de la introducción del término penalizador produce estimadores insesgados aún cuando tenemos una muestra pequeña y muy pocos eventos atípicos.

De esta forma, notemos que si la dimensión de  $i((\beta))$  es  $p$ , se concluye que la matriz de información de Fisher es  $i((\beta))$ ; recordemos que el determinante de información de Fisher se alcanza su máximo cuando  $p_i = \frac{1}{2}$ , i.e., cuando  $\beta_i = 0$  donde  $p_i$  es la probabilidad de éxito de la  $i$ -ésima observación.

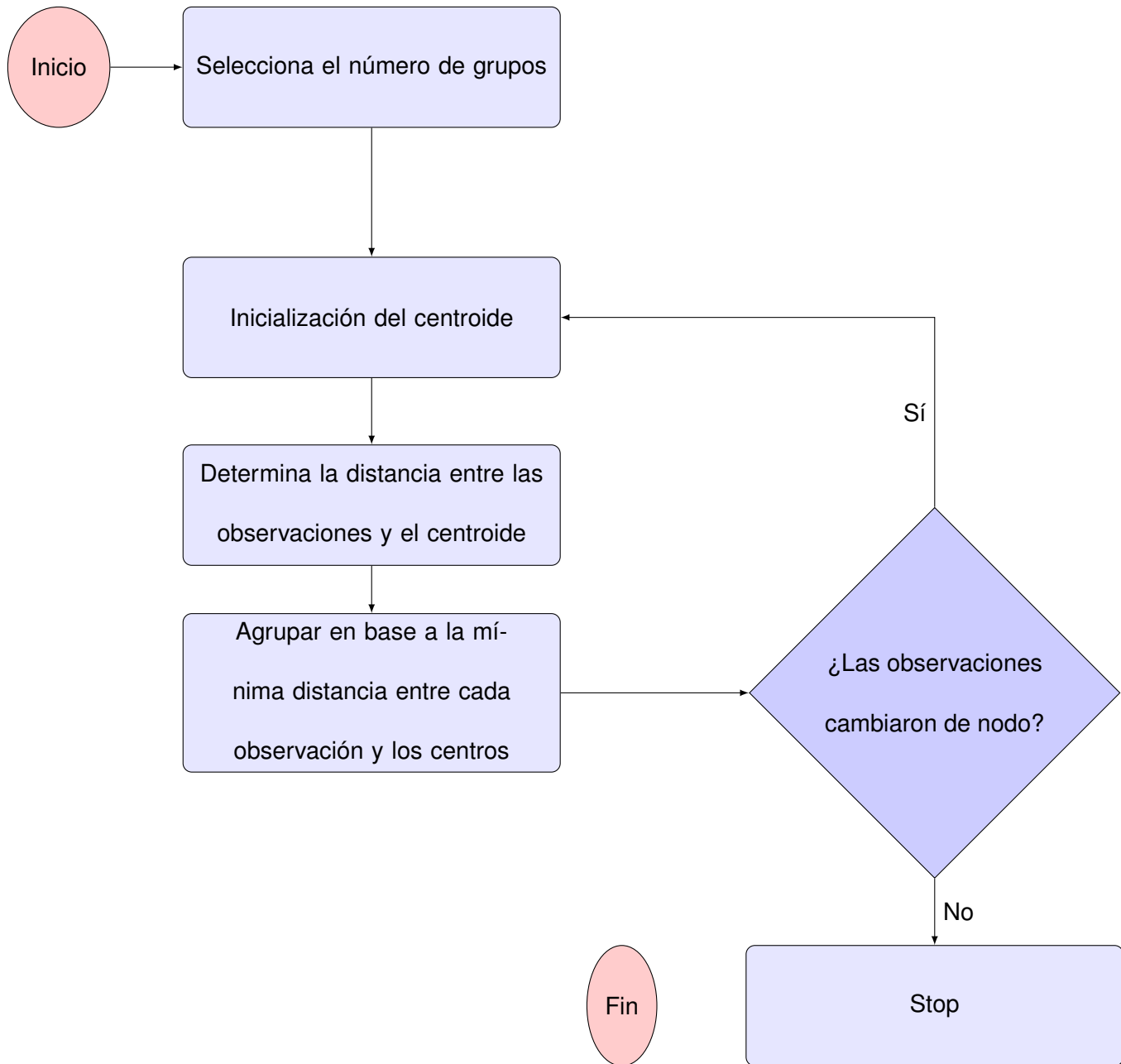
## 2.3 MACHINE LEARNING K-MEANS

El procedimiento de agrupación *K-mens* es el algoritmo de aprendizaje automático no supervisado más utilizado para dividir un conjunto de datos en un determinado número de grupos de datos sin etiquetar, el método requiere que el investigador seleccione el número deseado de grupos o centros y éste genera iterativamente movimientos de los centros para situarse en el conjunto donde se minimice la varianza entre todos sus elementos y donde los objetos de diferentes grupos sean lo más diferentes posibles, es decir los clasifica en grupos que son representados por su centro que corresponde a la media de puntos asignados al grupo.

Para entender el proceso que sigue el algoritmo K-means, se detalla los siguientes pasos dado un conjunto inicial de centros:

- \* Para cada centro identificamos el subconjunto de puntos de entrenamiento (su grupo) que está más cerca de él que cualquier otro centro.
  
- \* Se calculan las medias de los puntos de datos de cada grupo y este vector medio se convierte en el nuevo centro de ese grupo.
  
- \* Estos pasos se repiten hasta la convergencia.

El siguiente flujograma presenta los pasos que sigue el algoritmo de optimización K-means para encontrar los grupos:



**Figura 2.1.1:** Flujograma de Algoritmo K-means

**Elaboración:** El autor

El algoritmo estándar es el algoritmo de Hartigan-Wong (1979), que define la variación total dentro del conglomerado como la suma de las distancias al cuadrado, las distancias euclidianas entre los elementos y el centroide correspondiente, por ello se define:

$$W(C_k) := \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

donde,

- \*  $x_i$  es una observación de los datos que pertenecen al clúster  $C_k$
- \*  $\mu_k$  es la media de los puntos asignados al clúster  $C_k$

y para cada uno de los  $x_i$  es asignado a un grupo donde se garantice que la distancia de suma de cuadrados de la observación con el centro del grupo  $\mu_k$  se minimice.

Para la selección de la cantidad de centros en los que se deben agrupar los datos se tiene varios métodos entre ellos destaca el método conocido como *El codo* o *El Bow* por su nombre conocido del inglés.

### 2.3.1 Método El Codo (ElBow)

Recordemos que la idea principal para los métodos de particionamiento, como lo es el K-means, es la definición de los grupos de tal manera que sus elementos tengan la menor varianza con respecto al total de observaciones. El proceso que sigue el método está dado por los siguientes pasos:

- \* Se ejecuta el algoritmo de agrupación (para el ejercicio K-means) variando la cantidad de centros entre 2 a  $k$ ,  $k \in N$
- \* Para cada  $k$  se calcula la suma total del cuadrado dentro del grupo o WSS<sup>[2]</sup> (Por sus

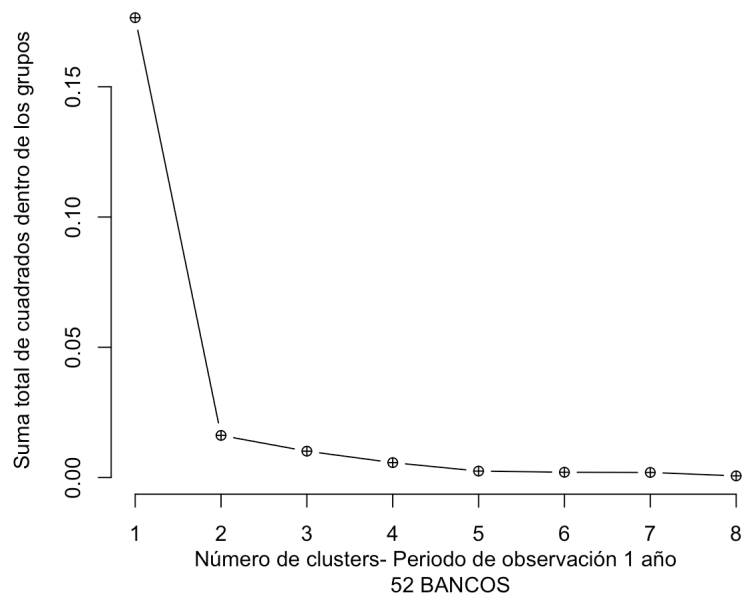
---

<sup>[2]</sup> WSS: La suma de las desviaciones cuadradas de cada observación y el centroide del grupo (With in Sum of Squares) por sus siglas en inglés

siglas en inglés)

- \* Se grafica la curva de WSS en base al número de  $k$  grupos.
- \* Se verifica el ángulo menos obtuso dentro de la línea donde se pueda visualizar el cambio de la gráfica, de donde se obtiene el número heurístico de grupos a ser utilizados.

En la siguiente figura se puede ver la forma empírica el resultado del método, teniendo en cuenta que en el eje de las abscisas están la cantidad de centros y en las ordenadas la Suma Total de Cuadrados dentro de cada uno de los grupos.



**Figura 2.2:** Ejemplo Método El Bow para 52 bancos en un período de 1 año

**Elaboración: El Autor**

## 2.4 MODELO DE MARKOV

Las principales referencias utilizadas en la siguiente sección son Ross Sheldon (2010), Rabiner, L. R. (1989), Z. Ghahramani (2001), V. Kulkarni (2001).

Considere un proceso que tiene un valor en cada período de tiempo. Sea  $X_n$  su valor en el período de tiempo  $n$  y suponga que queremos hacer un modelo de probabilidad para la secuencia de valores sucesivos  $X_0, X_1, X_2, \dots$ . El modelo más simple probablemente sería asumir que las  $X_n$  son variables aleatorias independientes, pero empíricamente esto no siempre sucede o más bien tienen muy pocas posibilidades de suceder. Por ejemplo, a partir de algún momento suponga que  $X_n$  representa el precio de una acción de algún valor al final de  $n$  días de negociación adicionales, entonces ciertamente parece irrazonable suponer que el precio al final del día  $n + 1$  es independiente de los precios en los días  $n, n - 1, n - 2$  y así sucesivamente hasta el día 0. Sin embargo, podría ser razonable suponer que el precio al final del día de negociación  $n + 1$  depende de los precios al final del día anterior solo a través del precio al final del día  $n$ . Es decir, podría ser razonable suponer que la distribución condicional de  $X_{n+1}$  dados todos los precios pasados al final del día  $X_n, X_{n-1}, \dots, X_0$  depende de estos precios pasados solo a través del precio al final del día  $n$ . Tal supuesto define una cadena de Markov, las cadenas y modelos de Markov las revisaremos en la presente sección.

### Definición 2.4.1 (Proceso estocástico)

Un proceso estocástico es un modelo de probabilidad que describe el desarrollo de un sistema que evoluciona aleatoriamente en el tiempo. Si observamos el sistema en un conjunto de momentos discretos, digamos al final de cada día o cada hora, obtenemos un proceso estocástico de tiempo discreto. Por otro lado, si observamos el sistema continuamente en



todo momento, obtenemos un proceso estocástico de tiempo continuo.

Consideremos un sistema que evoluciona aleatoriamente en el tiempo y supongamos que este sistema se observa en los tiempos  $n = 0, 1, \dots$ . Sea  $X_n$  el estado (aleatorio) del sistema en el tiempo  $n$ . La secuencia de variables aleatorias  $(X_0, X_1, X_2, \dots)$  se denomina proceso estocástico (en tiempo discreto) y se escribe  $\{X_n, n \geq 0\}$ . Sea  $S$  el conjunto de valores que  $X_n$  puede tomar para cualquier  $n$ . Entonces  $S$  se denomina espacio de estados del proceso estocástico  $\{X_n, n \geq 0\}$ . (Kulkarni, 2001)

**Definición 2.4.1 (Cadena de Markov)**

Un proceso estocástico a tiempo discreto  $X_n : n = 1, 2, \dots$  en el espacio de estados contables  $S$  se dice Cadena de Markov a Tiempo Discreto si, para  $i, j \in S$

$$P(X_{n+1} = x_{n+1} | X_1 = x_1, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n) = P_{ij} \quad (2.1)$$

donde en base al estado presente del proceso se puede determinar el estado futuro, de ello se siguen las siguientes propiedades :

$$\begin{aligned}
 P_{ij} &\geq 0, i, j \geq 0 \\
 \sum_{i=0}^{\infty} P_{ij} &= 1, i = 0, 1, \dots
 \end{aligned} \quad (2.2)$$

La expresión **2.1** se la conoce como Propiedad de Markov de Procesos Estocásticos. Ahora bien, de esta forma podemos calcular la distribución conjunta de las variables  $X_1, \dots, X_n$  así:

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) * P(X_2 = x_2 | X_1 = x_1) * \dots * P(X_n = x_n | X_{n-1} = x_{n-1}) \quad (2.3)$$

Y finalmente en base a las propiedades de **2.2**, se define **P** la matriz de Probabilidad de transición a un paso, así

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & \vdots & \vdots & \dots \\ p_{i0} & p_{i1} & p_{i2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (2.4)$$

#### **Definición 2.4.2 (Distribución Inicial)**

Como lo afirma Ross (2012 p.26): Una Cadena de Markov se determina completamente por la distribución de probabilidad de estado inicial  $\Pi = \{\Pi_i\}$  y por las probabilidades de transición  $P_{ij}$ . La distribución inicial de la cadena es expresada en forma de vector, en el que cada uno de sus componentes indica la probabilidad de que la cadena se encuentre en el estado  $i$  en el instante inicial y se expresa de la siguiente forma:

$$\Pi_i = (\Pi_1, \Pi_2, \dots, \Pi_m), i \in S$$

En donde se cumplen las propiedades:

$$\begin{aligned} \Pi_i &\geq 0, \forall i = 1, \dots, m. \\ \sum_{i=1}^m \Pi_i &= 1 \end{aligned}$$

para el caso que el conjunto S sea finito.

## 2.4.1 Ecuación de Chapman-Kolmogorov

En base a las definiciones de probabilidades de transición de un paso denotadas como  $P_{ij}$ , se define las probabilidades de transición de  $n$  pasos como  $P_{ij}^n$  que sería la probabilidad de que un proceso en el estado  $i$  estará en el estado  $j$  después de  $n$  transiciones consecutivas. Esto es,

$$P_{ij}^n = P(X_{n+k} = j | X_k = i) \quad (2.5)$$

con  $n, i, j \geq 0$  donde  $P_{ij}^1 = P_{ij}$ .

La Ecuación de Chapman-Kolmogorov nos proporciona un método para calcular las probabilidades de transición de  $n$ -pasos en donde se tiene que:

$$P_{ij}^{m+n} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m, \forall n, m \geq 0 \quad (2.6)$$

Formalmente se tendría,

$$\begin{aligned} P_{ij}^{m+n} &= P(X_{n+m} = j | X_0 = i) \\ &= \sum_{k=0}^{\infty} P(X_{n+m} = j, X_n = k | X_0 = i) \\ &= \sum_{k=0}^{\infty} P(X_{n+m} = j | X_n = k, X_0 = i) P(X_n = k | X_0 = i) \\ &= \sum_{k=0}^{\infty} P_{ik}^n * P_{kj}^m \end{aligned} \quad (2.7)$$

## 2.5 MODELO OCULTO DE MARKOV

Las principales fuentes de referencias serán tomadas de Ross Sheldon (2010), Rabiner, L. R. (1989), Jurafsky D., James H., (2018).

Un modelo de Markov oculto es una herramienta para representar distribuciones de probabilidad sobre secuencias de observaciones donde una observación  $X_t$  en el momento  $t$  es producido por un proceso estocástico, pero el estado  $S_t$  de este proceso no se puede observar directamente, es decir, es **oculto**. Este proceso oculto se supone que satisface la Propiedad de Markov . Este de hecho se denomina Modelo Oculto de Markov de primer orden.

Importante resaltar que al hablar de *transcurso de tiempo* en el modelo se hace referencia a que las observaciones ocurren en "pasos de tiempo" discretos dentro de la línea temporal.

Sea  $X_n, n = 1, 2, \dots$  una cadena de Markov con probabilidades de transición  $P_{i,j}$  y probabilidades de estado inicial  $P_i = P(X_1 = i), i \geq 0$ . Supongamos dos sucesos principales:

- 1) Existe un conjunto finito  $\mathcal{V}$  de señales en donde se emite una señal de  $\mathcal{V}$  cada vez que la cadena de Markov entra en un estado.
- 2) Cuando la cadena de Markov entra en el estado  $j$ , entonces, independientemente de los estados y señales anteriores de la cadena de Markov, la señal emitida es  $p(v|j)$ , donde

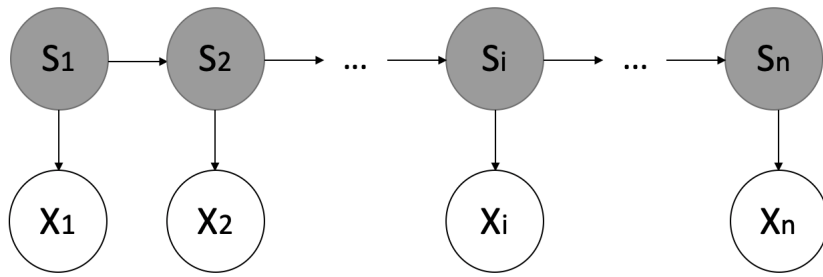
$$\sum_{v \in \mathcal{V}} p(v|j) = 1$$

Es decir, si  $V_n$  representa la n-ésima señal, se sigue

$$P(V_1 = g | X_1 = j) = p(v|j) \tag{2.8}$$

$$P(V_1 = g | X_1, V_1, \dots, X_{n-1}, V_{n-1}, X_n = j) = p(v|j)$$

En el modelo anterior la secuencia de señales  $V_1, V_2, \dots$  son observables, mientras que la cadena de estados de Markov  $X_1, X_2, \dots$  no es observable, a esto se denota como **Cadena de Markov Oculta**. La siguiente figura muestra la estructura de una Cadena Oculta de Markov, donde las circunferencias grises son los estados ocultos:



**Figura 2.3:** Estructura Cadena Oculta de Markov de primer orden

**Elaboración:** El autor

“Un modelo oculto de Markov (HMM) nos permite hablar sobre los eventos observados y los eventos ocultos que consideramos como factores causales en nuestro modelo probabilístico.” (Jurafsky D., James H., 2018).

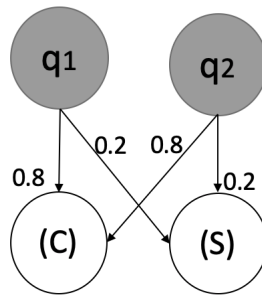
Consideremos un ejemplo de dos monedas no justas, es decir que sus probabilidades de que salga cara(C) o sello(S) son diferentes a 50%. Una persona lanza de manera secuencial las dos monedas, pero el hecho es que no se observa cuál de ellas lanzó, sólo podemos observar el resultado.

Esto estaría definido, por ejemplo, con una matriz de de probabilidades de la siguiente forma:

	$m_1$	$m_2$
C	0,8	0,2
S	0,2	0,8

**Tabla 2.1:** Ejemplo HMM Monedas no justas

El problema de interés es cómo construimos un Modelo Oculto de Markov para explicar (modelar) la secuencia observada de Caras y Sellos. La estructura del juego para cada una de las monedas viene dado como se presenta a continuación:



**Figura 2.4:** Estructura Ejemplo monedas no justas HMM

**Elaboración:** El autor

Como resultado del lanzamiento de las monedas que se muestra en la figura anterior, hay 2 estados en el modelo y cada estado corresponde al lanzamiento de una moneda diferente y sesgada. Cada estado se caracteriza por una distribución de probabilidad de caras y sellos y las transiciones entre estados se caracterizan por una matriz de transición de estados. El mecanismo físico que explica cómo se seleccionan las transiciones de estado podría ser en sí mismo un conjunto de lanzamientos de monedas independientes o algún otro evento probable.

Finalmente, en base a las especificaciones que debe cumplir un Modelo de Markov Oculto definimos los elementos propios de este modelo:

$V = (v_1, v_2, \dots, v_m)$ : Los símbolos de observación en cada estado, para la práctica de nuestro modelo tenemos dos en base al índice de liquidez baja y alta.

$O_t$ : Observación en el tiempo  $t$ , es decir uno de los símbolos de  $V$ ,

$S = (S_1, \dots, S_n)$ : Cada uno de los estados individuales que en el ejercicio de la práctica son *verde, amarillo, naranja, rojo*

$q_t$ : El estado al tiempo  $t$ .

$A = \{a_{ij}\}$ : Matriz de transición entre los estados, donde

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], 1 \leq i, j \leq n$$

$B$ : La distribución de probabilidad del símbolo de observación en el estado  $B = \{b_j(k)\}$ ,  
donde

$$b_j(k) = P[V_{k,t} | q_t = S_j]$$

$1 \leq j \leq n$  y  $1 \leq k \leq m$ .

$\Pi$ : Matriz de Probabilidad inicial, donde

$$\pi_i = P[q_1 = S_i], 1 \leq i \leq n$$

Es así que, en base a las definiciones anteriores de los parámetros del modelo usaremos la notación compacta del **Modelo de Markov Oculto**

$$\lambda = (A, B, \Pi).$$

### 2.5.1 Problemas de interés en los Modelos Ocultos de Markov

Según Lawrence R. Rabiner 1989, los Modelos Ocultos de Markov para que sean útiles en aplicaciones dentro del mundo real deben resolver 3 principales problemas:

**Problema 1:** Dada la secuencia de observación  $(O_1, O_2, \dots, O_n)$ , un modelo  $\lambda = (A, B, \Pi)$  ¿Cómo calculamos de manera eficiente  $P(O|\lambda)$ , es decir la probabilidad de la secuencia de observación dado el modelo?

Este es un problema de evaluación, es decir, dada una secuencia de observaciones

necesitamos ver cómo calculamos la probabilidad de que la secuencia observada haya sido producida por el modelo  $\lambda$ . Desde otro punto de vista podemos decir que se necesita ver qué tan bien un modelo dado coincide con una secuencia de observación dada, este último punto de vista será de mucha utilidad ya que por ejemplo, si consideramos el caso en el que estamos tratando de elegir entre varios modelos, la solución sobre el problema 1 nos direcciona a elegir el modelo que mejor se adapta a las observaciones.

**Problema 2:** Dada la secuencia de observación  $(O_1, O_2, \dots, O_n)$  y el modelo  $\lambda = (A, B, \Pi)$  ¿Cómo elegimos una secuencia de estado correspondiente  $Q = (q_1, q_2, \dots, q_n)$  que sea óptima es decir, que mejor explique las observaciones?

En este problema intentamos descubrir la parte oculta del modelo, es decir, encontrar la secuencia de estados "correcta". Es importante resaltar que para todos los modelos, excepto para el caso de los modelos degenerados <sup>[3]</sup>, no hay una secuencia de estado "correcta" que se pueda encontrar. Por lo tanto, en la práctica se suele usar un criterio de optimalidad para resolver este problema lo mejor posible.

**Problema 3:** El problema se basa en ¿Cómo se debe ajustar los parámetros del modelo  $\lambda = (A, B, \Pi)$  para maximizar  $P(O|\lambda)$ ? En este problema se busca optimizar los parámetros del modelo para describir de mejor manera cómo se produce una determinada secuencia de observación. La secuencia de observación utilizada para ajustar los parámetros del modelo se denomina secuencia de entrenamiento, ya que se usa para entrenar al HMM<sup>[4]</sup>.

El enfoque del presente trabajo es dar solución a 2 de los 3 problemas. Esto es, lo primero

---

<sup>[3]</sup> Modelos Degenerados: Modelos que tienen al menos una restricción redundante.

<sup>[4]</sup> HMM: Hidden Markov Model (Modelo Oculto de Markov)



es buscar de manera eficiente la probabilidad de ocurrencia de las observaciones (Estados de Fragilidad Bancaria) dado el modelo  $\lambda$ , para lo que se requiere la aplicación del método Forward [Véase a continuación]. Segundo, se busca resolver el problema 3 antes expuesto, es decir encontrar en base a la matriz de probabilidad de transición de estados y la probabilidad inicial de los estados, la sucesión de señales que maximicen la probabilidad de predecir la secuencia de estados de fragilidad bancaria en un período de tiempo futuro. Para ello se requiere la aplicación del método Viterbi [Véase a continuación].

## 2.5.2 Algoritmo Forward

Como solución al problema 1 que se ha revisado previamente, para nuestro problema de fragilidad bancaria tomamos una secuencia de observaciones entre las opciones: *Liquidez baja* o *Liquidez alta*. Para así ver qué probabilidad tiene cada uno de los estados de suceder en base a la secuencia de observaciones. Para ello tomamos una secuencia de estados fijos dada  $Q = q_1, q_2, \dots, q_T$ , de esta forma veamos que:

$$P(O|Q, \lambda) = \prod_{t=1}^T P(O_t|q_t|\lambda) \quad (2.9)$$

que es la probabilidad de que ocurra la secuencia de observaciones  $O$  definida como  $O = (O_1, O_2, \dots, O_T)$ , dada la secuencia de estados  $Q$ . Y como asumimos independencia de las observaciones se sigue que

$$P(O|Q, \lambda) = b_{q_1}(O_1)b_{q_2}(O_2), \dots, b_{q_T}(O_T)$$

$$P(O|Q, \lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3}, \dots, a_{q_{T-1} q_T}$$

De esta forma, dado que  $P(O, Q|\lambda) = P(O|Q, \lambda)P(Q, \lambda)$  se tiene que  $P(O|\lambda)$  está definida como

$$\begin{aligned} P(O|\lambda) &= \sum_{Q \in \hat{Q}} P(O|Q, \lambda)P(Q, \lambda) \\ &= \sum_{q_1, \dots, q_T} (\pi_{q_1} b_{q_1}(O_1))(a_{q_1 q_2} b_{q_2}(O_2)), \dots, (a_{q_{T-1} q_T} b_{q_T}(O_T)) \end{aligned} \quad (2.10)$$

Donde  $\hat{Q}$  es el conjunto de todas las posibles secuencias de estados.

Esta última definición de  $\hat{Q}$  genera un alto nivel de dificultad dentro de los cálculos para llegar al objetivo principal que es la solución del *Problema 1*. Es así que se define

$$\alpha_t(i) := P(O_1 O_2, \dots, O_t, q_t = S_i | \lambda)$$

Es decir, tomamos una secuencia parcial de observaciones hasta el tiempo  $t$  y de igual manera el estado  $i$ , entonces para resolver  $\alpha_t(i)$  vamos por inducción de tal manera que:

**1. Inicialización:** Para  $1 \leq i \leq N$

$$\alpha_1(i) = \pi_1 b_1(O_1)$$

**2. Inducción:** Para  $1 \leq t \leq T - 1$  y  $1 \leq j \leq N$

$$\alpha_{t+1}(j) = \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1})$$

### 3. Terminación:

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i)$$

A este algoritmo de desarrollo se le conoce como **Algoritmo Forward** de Modelos Ocultos de Markov.

### 2.5.3 Algoritmo Viterbi

Con el objetivo de solucionar el problema 2 propuesto previamente, que en la práctica de nuestro problema sobre *Fragilidad Bancaria* radica en buscar una secuencia de señales entre Liquidez Baja o Liquidez Alta que permitan determinar estados futuros en los cuales, en base a las señales emitidas, se encontrarán los bancos o en sí un banco que cumpla con una cierta secuencia de señales tiene una cierta probabilidad de situarse en los estados pronosticados ya sea *verde, amarillo, naranja, rojo*. Es así que procedemos a definir las variables necesarias para implementar la solución de dicho problema.

Definamos la probabilidad de estar en el estado  $S_i$  en el tiempo  $t$  dado el modelo  $\lambda$  y la secuencia de observaciones  $O$

$$\varphi_t(i) = P(q_t = S_i | \lambda, O) \quad (2.11)$$

donde se cumple que

$$\sum_{i=1}^N \varphi_t(i) = 1 \quad (2.12)$$

Ahora bien, utilizando ecuación (2.12) podemos solucionar la búsqueda del estado más

probable (de manera individual), esto es:

$$q_t = \arg \max_{1 \leq i \leq N} [\varphi_t(i)], 1 \leq t \leq T \quad (2.13)$$

La ecuación (2.13) nos brinda la solución óptima para el estado más probable en cada instante pero no tiene en cuenta la probabilidad de ocurrencia de secuencias de estados. La solución de este problema se basa en buscar y encontrar la mejor secuencia de estados, esto es, se busca maximizar  $P(Q|O, \lambda)$  que finalmente se resumen el maximizar  $P(Q, 0|\lambda)$ . Ahora, con el fin de encontrar la mejor secuencia de estados  $Q = (q_1, \dots, q_T)$  dada la secuencia de observaciones  $O = (O_1, \dots, O_T)$ , vamos a definir la variable  $\delta_t(i)$  que simboliza la mayor probabilidad obtenida en una ruta única de estados en el tiempo  $t$ , que representa las primeras  $t$  observaciones y termina en el estado  $S_i$ .

$$\delta_t(i) := \max_{q_1, q_2, \dots, q_t} P(q_1 q_2 \dots q_{t-1} = i, O_1, O_2, \dots, O_t | \lambda) \quad (2.14)$$

es decir,  $\delta_t(i)$  es la puntuación más alta (probabilidad más alta) a lo largo de una ruta única, en el tiempo  $t$ , que representa las primeras  $t$  observaciones y termina en el estado  $S_i$ . Por inducción tenemos:

$$\delta_{t+1}(i) = [\max_j (\delta_t(j) a_{ij})] b_i(O_{t+1}) \quad (2.15)$$

Vamos a dar seguimiento a cada uno de los argumentos que maximizaron la ecuación (2.15) para así tener la secuencia de estados de manera recursiva para cada  $j$  y  $t$ . Finalmente para encontrar la secuencia de estados obtenida por (2.15) mediante los argumentos que la maximizaron, se sigue el siguiente procedimiento:

Se define la matriz  $\psi_t(i)$ ,

### 1. 1. Inicialización:

$$\begin{aligned}\delta_1(i) &= \pi_i b_i(O_1), 1 \leq i \leq N \\ \psi_1(i) &= 0\end{aligned}\tag{2.16}$$

### 2. 2. Recursividad:

$$\begin{aligned}\delta_t(j) &= \max_{1 \leq i \leq N} (\delta_{t-1}(i) a_{ij}) * b_i(O_t), 2 \leq t \leq T \\ &1 \leq j \leq N\end{aligned}\tag{2.17}$$

$$\begin{aligned}\psi_t(j) &= \arg \max_{1 \leq i \leq N} \delta_{t-1}(i) a_{ij}, 2 \leq t \leq T \\ &1 \leq j \leq N\end{aligned}\tag{2.18}$$

### 3. 3. Terminación:

$$\begin{aligned}P^* &:= \max_{1 \leq i \leq N} [\delta_T(i)] \\ q_T^* &= \arg \max_{1 \leq i \leq N} [\delta_T(i)]\end{aligned}\tag{2.19}$$

### 4. 4. Retroceso (Seguimiento secuencia de estados):

$$q_t^* = \psi_{t+1} q_{t+1}^*, t = T - 1, T - 2, \dots, 1\tag{2.20}$$

A este proceso que busca encontrar una predicción de secuencia de estados en base a una secuencia de observaciones dadas se le llama **Algoritmo Viterbi**.

### **3 METODOLOGÍA Y RESULTADOS**

Dentro del presente capítulo se detalla el proceso realizado para obtener un sistema de alerta temprana sobre riesgo institucional financiero para los bancos ecuatorianos, para ello el procedimiento se subdivide en 4 pasos principales. En primera instancia se busca obtener la mayor la información sobre variables micro y macroeconómicas a nivel nacional. Con estos datos se actualiza el modelo desarrollado en el trabajo " Quiebra bancaria: Un análisis econométrico para el Ecuador " Uquillas y Flores (2020) y se detalla varias de las pruebas estadísticas que el modelo debe cumplir, dicho proyecto ha tendido un desarrollo exitoso en el ámbito académico por los resultados sostenibles obtenidos. El tercer paso es agrupar los resultados del modelo de cada uno de los bancos estudiados para buscar umbrales entre cada grupo a los que llamaremos *estados* y finalmente se incluye una variable no utilizada en el modelo como señal para que forme parte del Modelo Oculto de Markov con el cual se realizará la predicción de estados en un período de tiempo.

#### **3.1 OBTENCIÓN Y TRATAMIENTO DE LA INFORMACIÓN**

Mediante la plataforma web del Banco Central del Ecuador en su Micrositio de Información Económica y de la Superintendencia de Bancos mediante los boletines financieros mensuales emitidos en el Portal Estadístico y de Estudios, se obtienen datos de los factores del entorno macroeconómico del país y a su vez de las variables del sistema financiero que

permitan tener la información suficiente para crear un sistema de calificación interbancario y gracias a estos datos obtenidos podemos calcular diferentes índices financieros que puedan ser de utilidad para el modelo de fragilidad financiera.

El período de análisis de la información se encuentra desde julio de 1996 hasta marzo de 2019 con un total de 2501 observaciones de donde se extrajo la información de las posibles variables a utilizar en el modelo para cada banco operativo/quebrado en el período observado. Definimos la matriz  $D := \{d_{ijk}\}$  de tal manera que las posibles variables a ser utilizadas en el modelo estén de manera vertical y las observaciones de todos los bancos de manera horizontal, de esta forma se tendría una matriz de la siguiente forma:

<b>VARIABLE</b>	<i>Banco<sub>1k</sub></i>	<i>Banco<sub>2k</sub></i>	...	<i>Banco<sub>jk</sub></i>
APALANCAMIENTO	$d_{11k}$	$d_{12k}$	...	$d_{1jk}$
DOLARIZACIÓN	$d_{21k}$	$d_{22k}$	...	$d_{2jk}$
ÍNDICE DE LIQUIDEZ				
INFLACIÓN				
PIB	...	....	...	...
QUIEBRA				
ROE				
SPREAD FINANCIERO				
TASA DE MOROSIDAD				
VALOR PETRÓLEO	$d_{10.1k}$	$d_{10.2k}$	...	$d_{10jk}$

**Tabla 3.1.1:** Variables Obtenidas o Calculadas de las fuentes SBS o BCE

donde  $i \in \{1, \dots, 10\}$  es el número de variables obtenidas,

$j$  es el número de bancos operativos en el período  $k$ ,

$k$  la cantidad de períodos en análisis.

De esta forma se tiene una matriz con un total de 25010 datos.

Con esta base de información se procede con la depuración de los datos debido a inconsistencias en las plataformas web o falta de información para períodos antiguos. Posterior a ello se obtiene una base de 2443 observaciones con las cuales se procede a trabajar.

### **3.2 ACTUALIZACIÓN/ADAPTACIÓN DEL MODELO ESTADÍSTICO.**

Con toda la información obtenida en el punto 3.1, el siguiente paso se enfoca en buscar y desarrollar un modelo que garantice el uso de las variables macro y micro económicas principales del sistema financiero para generar una calificación de bancos en miras de predecir una quiebra bancaria. Como ya se mencionó el proceso parte en actualizar el modelo propuesto por Uquillas y Flores (2020) ya que este modelo presenta variables representativas en el ámbito macroeconómico del país al igual que variables propias de cada entidad financiera, por ello el modelo es un perfecto punto de inicio para alcanzar los objetivos planteados.

En el modelo de regresión logística de Firth que se hizo mención, se han utilizado las siguientes variables financieras, las cuales se valida que cumplan con las pruebas estadísticas. En la siguiente tabla se muestra el nivel predictivo de cada variable utilizada en el modelo mediante la prueba IV.<sup>[1]</sup>

---

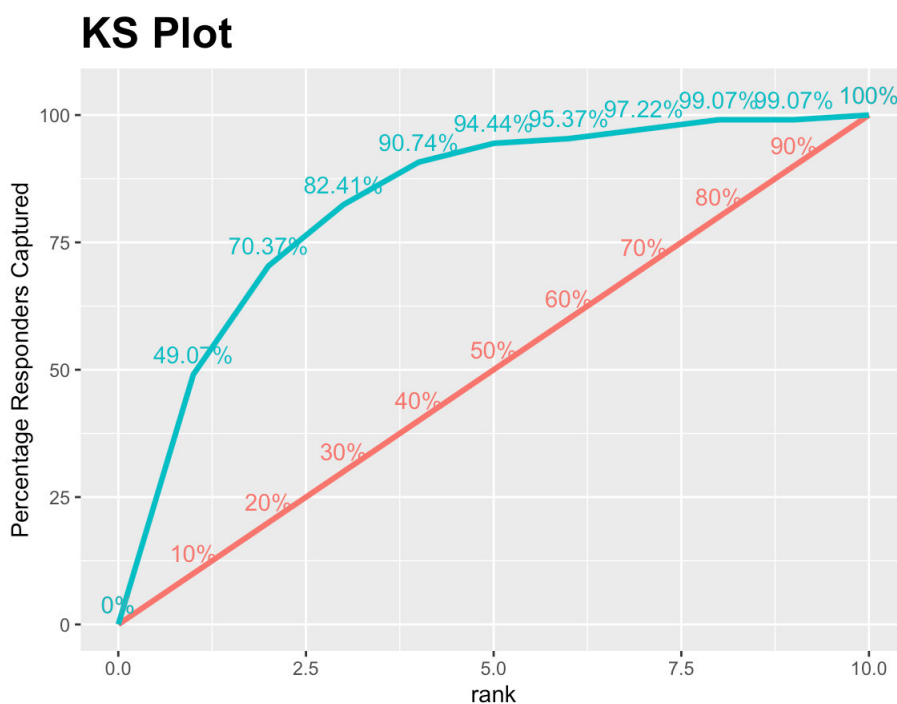
<sup>[1]</sup> Information Value: El estudio del Valor de la Información es una técnica de exploración de datos que ayuda a determinar qué columnas de un conjunto de datos tienen poder predictivo o influyen el valor de una variable dependiente especificada.



N	Variable	Information Value
1	DOLARIZACIÓN	Highly Predictive
2	INFLACIÓN	Highly Predictive
3	TASA DE MOROSIDAD	Highly Predictive
4	APALANCAMIENTO	Highly Predictive
5	SPREAD	Highly Predictive
6	ROE	Highly Predictive

**Tabla 3.1:** Variables utilizadas en el Modelo Logístico de Firth

Se detalla a continuación una de las principales pruebas conocida como *distancia de Kolmogorov-Smirnov* o *KS-Test* en la cual podemos entender el trecho entre la distribución empírica de la muestra analizada y la distribución teórica, en la Figura 3.1 podemos evidenciar que entre las dos líneas roja y verde existe una distancia **vertical** que supera el 0.50 %.



**Figura 3.1:** Prueba Kolmogórov-Smirnov del modelo de regresión logística de Firth

**Elaboración:** El autor

Según Mays (2004), como una medida que tiene la capacidad de clasificación, el valor de KS-Test va desde el 20 % donde se debería cuestionar el valor y la capacidad del modelo y hasta el 70 % donde "probablemente sea demasiado bueno para ser verdad". Para nuestro caso se tiene el valor de 58.44 %.

Para un análisis global del comportamiento del modelo, en la Tabla 3.2 se resume la clasificación de aciertos:

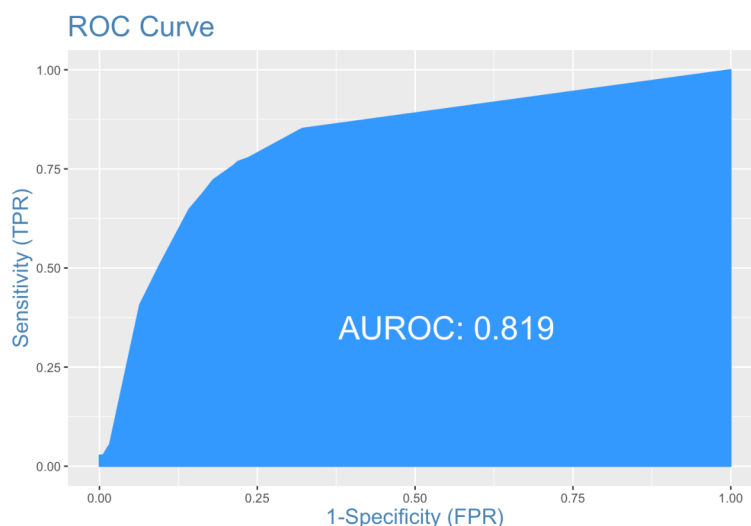
<b>Aciertos Modelo de Firth</b>	
<b>Indicador</b>	<b>Resultado</b>
Precisión	75 %
Especificidad	99.9 %
Sensibilidad	2.77 %
Tasa de Error	4.85 %

**Tabla 3.2:** Clasificación de aciertos del modelo de Firth

Entonces al revisar los resultados obtenidos, podemos identificar que el modelo actualizado tiene una exactitud para predicción de 75 %; en el caso de la especificidad del modelo, que es el porcentaje de entidades financieras en quiebra que predice el modelo, alcanza el 99.9 %. Finalmente podemos revisar que la probabilidad de quiebra con relación a las que no se encuentran en quiebra alcanza un porcentaje de 4.85 %.

En la Figura 3.2 se muestra la curva ROC<sup>[2]</sup> por sus siglas en inglés (Receiver Operating Characteristic Curve), en donde se evidencia que un 81.9 % es clasificado de manera correcta entre los bancos que se encuentran en quiebra o no quiebra. En otras palabras hay aproximadamente un 82 % de acierto en la clasificación de verdaderos y falsos positivos.

<sup>[2]</sup> ROC: Área bajo la curva que muestra la relación sensibilidad y especificidad del modelo



**Figura 3.2:** Curva ROC del modelo estadístico

**Elaboración:** El autor

En los **Anexos 2, 3, 4 y 5** podemos verificar el cumplimiento de las pruebas estadísticas de Multicolinealidad, Choque Estructural, Variables Omitidas y Heteroscedasticidad respectivamente las cuales el modelo cumple para garantizar la robustez del mismo.

### **3.3 AGRUPACIÓN DE LOS RESULTADOS OBTENIDOS DEL MODELO**

Uno de los puntos más trascendentales dentro del trabajo es obtener los estados de fragilidad bancaria con éxito. Esto es, en base a la calificación obtenida en el modelo, obtener de manera adecuada los umbrales dentro de los que cada banco se encontrará en un período de tiempo. Por ello, entre la gran variedad de métodos de agrupación es importante resaltar el método K-means que es un método de aprendizaje no supervisado que dividen los datos observados (resultados del modelo) en k grupos según sus diferentes características, de esta forma dicho método será un gran elemento de ayuda para realizar una buena agrupación de datos en base a sus propiedades o patrones que genera las variables utilizadas.

Lo que se desea con ello es crear  $k$  grupos con lo cual se obtendrá los niveles de alerta llamados estados y de esta forma ubicar a los bancos según sus características financieras en cada uno de los niveles de alerta.

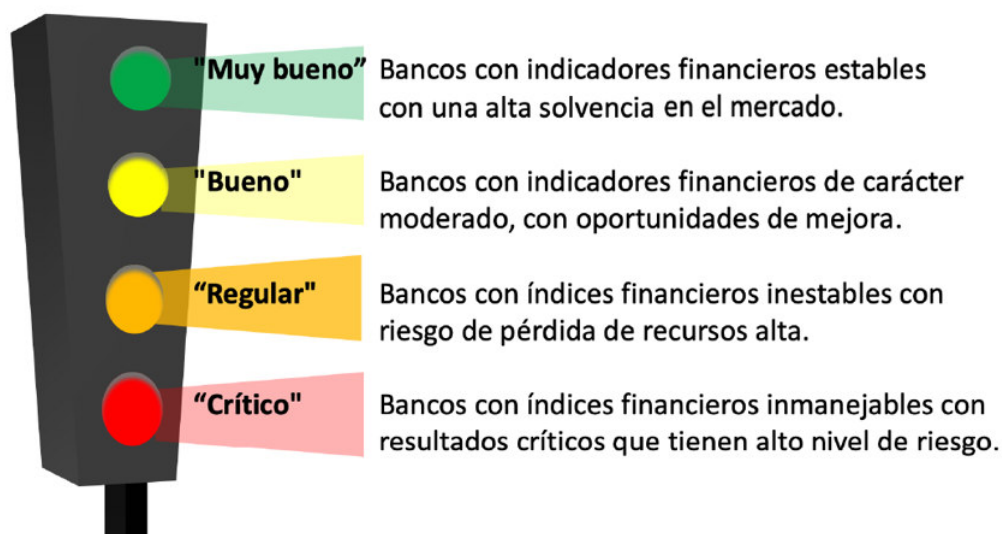
El algoritmo de agrupación *K-means* lo que se busca es poder encontrar la cantidad de grupos apropiada con respecto a la calificación obtenida en el modelo de Firth para la creación de niveles en donde los grupos mantengan las características más similares entre sí, es decir que los bancos que se encuentren en estos niveles tengan variables financieras de mismas cualidades o comportamiento, por ejemplo desde un plano a corto plazo podemos ver en las siguientes figuras (3.4, 3.5, 3.6) la calificación de riesgo de los bancos analizados en un periodo de tiempo, donde mientras más se aleje el resultado de 0 en las ordenadas existe una mayor probabilidad que este banco haya quebrado, mientras que las observaciones de los bancos que se encuentren más cercanas a 0 quiere decir que dichas entidades financieras tienen un comportamiento más estable y con baja probabilidad de quiebra. Lo que se busca es crear un "**Semáforo**" de alertas donde los niveles aumentan su riesgo en función del color al que son asignados, para el problema en gestión se definen 4 niveles de la siguiente forma:

\* **Verde:** Muy bueno.

\* **Amarillo:** Bueno.

\* **Naranja:** Regular.

\* **Rojo:** Crítico.

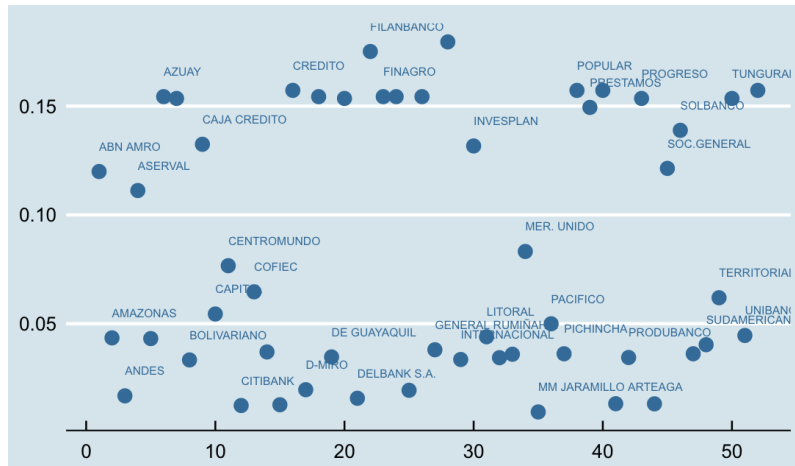


**Figura 3.3:** Semáforo de estados de Fragilidad Bancaria

**Elaboración:** El autor

En la figura anterior podemos ver 4 niveles dentro del semáforo con colores estándar y una explicación intuitiva de cada uno de ellos, el objetivo de la construcción de un semáforo es poder tener una visión maquinal al momento de obtener los resultados y que estos se vean representados por estos grupos o niveles a los cuales se denotan como **ESTADOS DE FRAGILIDAD**.

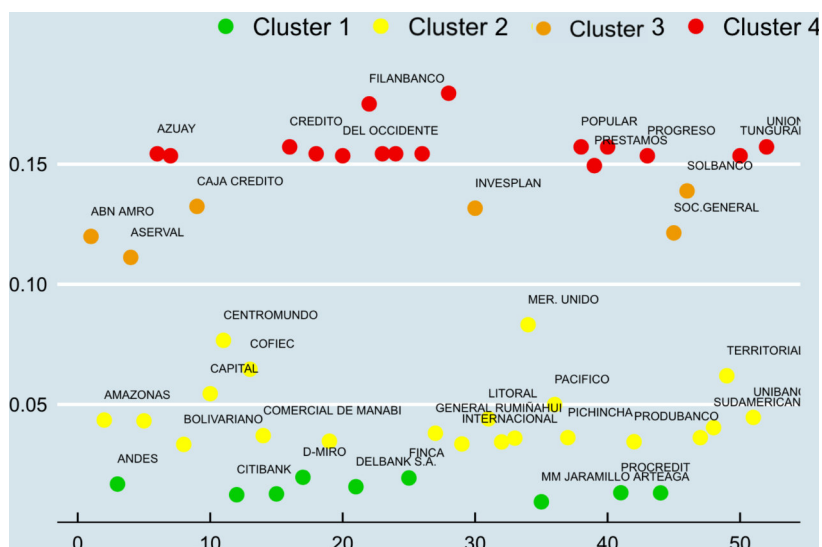
En la Figura 3.4 podemos ver los resultados del score de Firth para todos los bancos ya sea que han quebrado o que se mantienen operativos haciendo un promedio de sus calificaciones dentro del período 1996-2019, esto con el fin tener una percepción de la ubicación en la que los bancos con buena o mala puntuación de score, se encuentran en un plano visual  $(X, Y)$ . Donde podemos ver que los bancos que han quebrado, por ejemplo *FILANBANCO* o el *BANCO POPULAR*, tienen una de las mayores calificaciones en el Modelo de Firth, es decir se encuentran alejadas del valor 0 en el eje Y.



**Figura 3.4:** Calificación de Riesgo de los Bancos a corto plazo período 1996-2019

**Elaboración:** El autor

Ahora bien, el siguiente paso es verificar la cantidad de grupos en los cuales se dividen las observaciones e identificar a qué grupo el método de K-means asignó a cada banco. Para el ejemplo de visualización anterior se asigna la cantidad de 4 centroides al algoritmo. En el eje de las abscisas se detalla la cantidad de bancos en estudio mientras que en el eje de las ordenadas la calificación que estos bancos obtuvieron en el score de Firth.



**Figura 3.5:** Calificación de Riesgo de los Bancos a corto plazo período 1996-2019 ejecutado K-means

**Elaboración:** El autor

Es importante conocer y entender qué tipo de bancos se encuentran en cada uno de los grupos a los cuales fueron asignados, como se puede ver en la Figura 3.5 la distribución de la cantidad de bancos tiene diferente concentración dependiendo de cada grupo, por ejemplo:

**Clúster 1:** El estado verde está integrado por 9 bancos que cumplen con las mejores características en función de sus resultados obtenidos en el periodo de análisis.

**Clúster 2:** El estado amarillo aloja 21 bancos que tienen un resultado aceptable basado en sus indicadores financieros. En el ejemplo podemos evidenciar que este estado tiene la mayor cantidad de bancos agrupados.

**Clúster 3:** El estado naranja contiene 6 bancos con oportunidades altas de mejora con respecto a los resultados obtenidos en su score que viene directamente conectado con los resultados de sus índices financieros mensuales.

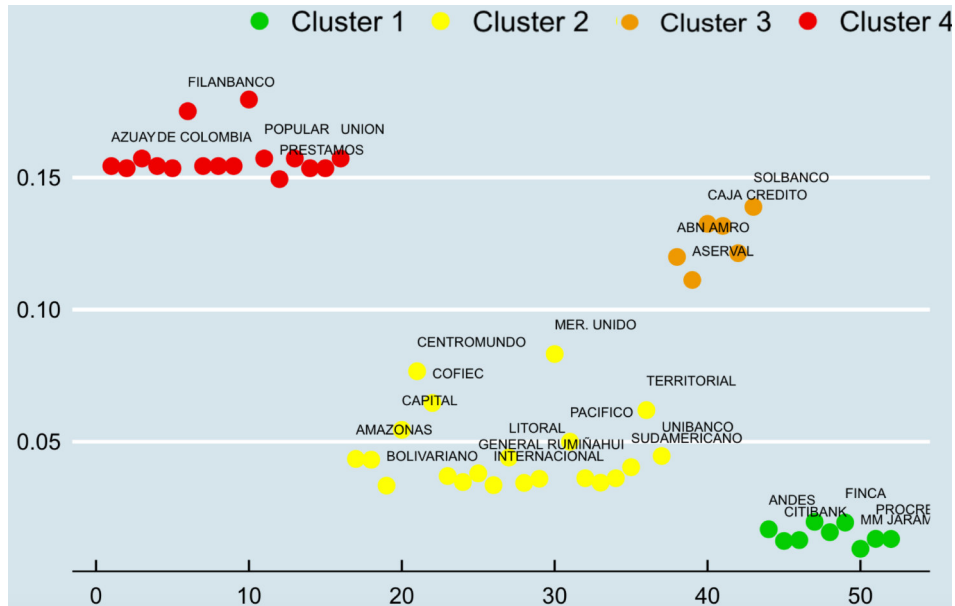
**Clúster 4:** El estado rojo, que debido al alto número de bancos quebrados en el período 1999-2001, tiene alojados 16 bancos lo que significa aproximadamente 31 % del total de bancos observados entre bancos operativos y quebrados; estos bancos tienen uno o más indicadores inestables a nivel micro y macro económico.

Indicador	Resultado
Verde	17.3 %
Amarillo	40.4 %
Naranja	11.5 %
Rojo	30.8 %

**Tabla 3.3:** Concentración de bancos agrupados en 4 centroides

Como se puede evidenciar en la Tabla 3.3, los dos grupos con mayor concentración son el amarillo y el rojo, en el primer caso del color amarillo la idea es que varios de los bancos no han sido estables en el tiempo al momento de conseguir resultados a nivel económico buenos, más aún se mantienen en un nivel bueno con oportunidades de mejora; por otro

lado la concentración en el nivel rojo, como ya se dijo, se debe al alto índice de bancos que tuvieron indicadores que los llevaron a la quiebra o que es su defecto están con resultados críticos en la actualidad.



**Figura 3.6:** Calificación de Riesgo de los Bancos en el período 1996-2019, ejecutado K-means AGRUPADO

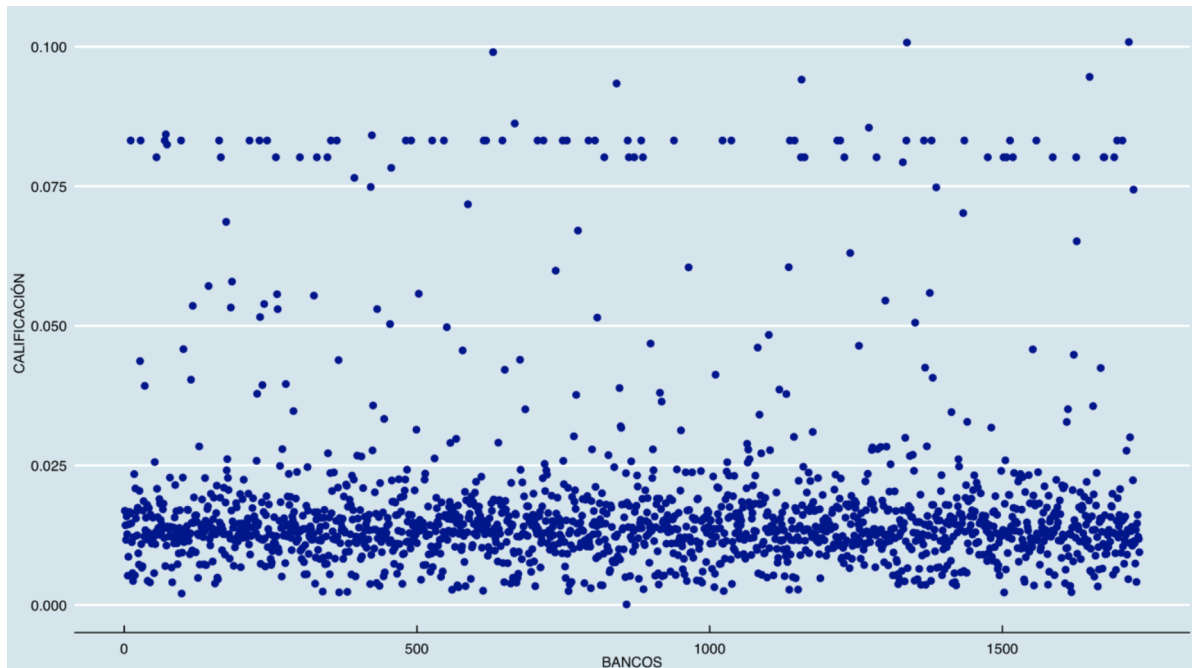
**Elaboración:** El autor

Con el fin de evidenciar la concentración de cada uno de los grupos *verde*, *amarillo*, *naranja*, *rojo* la idea es poder agruparlos dentro de los ejes para así evidenciar la tendencia de cada banco dependiendo de su resultado en el Score de Firth.

Con esta representación de cómo trabaja el Algoritmo K-means y cuál es la idea de tener grupos de bancos con similares características, el objetivo es analizar cada una de las observaciones de manera independiente en el tiempo, es decir cada una de las calificaciones del modelo sin tener en cuenta la etiqueta del banco al que pertenece. Para esto la siguiente figura detalla el comportamiento de cada una de las calificaciones obtenidas en el modelo, en donde el puntaje mientras esté más alejado de 0 en el eje de las ordenadas quiere decir que esa observación tuvo una alta probabilidad de quiebra o en su defecto quebró. En el eje



de las abscisas se encuentran enumeradas las observaciones (Bancos en el tiempo) para una mejor visualización:

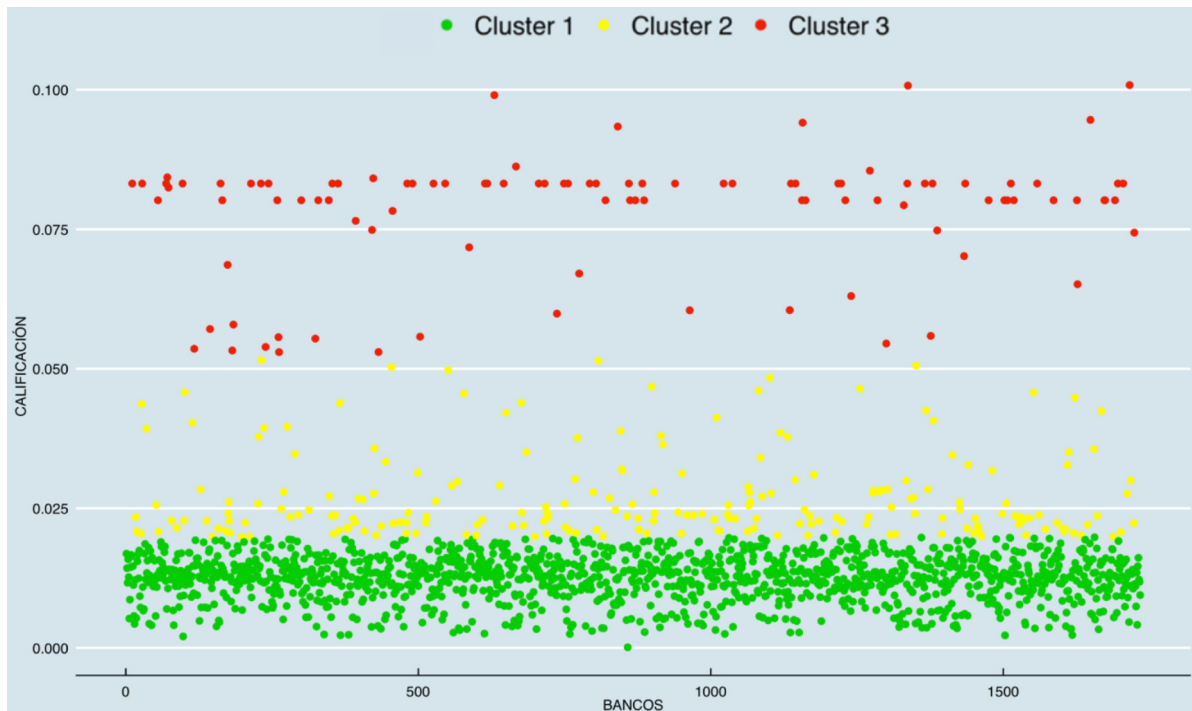


**Figura 3.7:** Calificación de Bancos en el período de tiempo 1996-2019

**Elaboración:** El autor

Se evidencia una aglomeración de resultados cerca de 0 en su calificación y muy pocos en la parte superior lo que visualmente nos rectifica que el objetivo se basa en pronosticar un evento raro( con baja probabilidad de ocurrencia) en un conjunto de datos.

A continuación se procede a correr el algoritmo *K-means* con 3 centroides para el conjunto de observaciones de la Figura 3.7, de donde se obtienen los siguientes resultados:

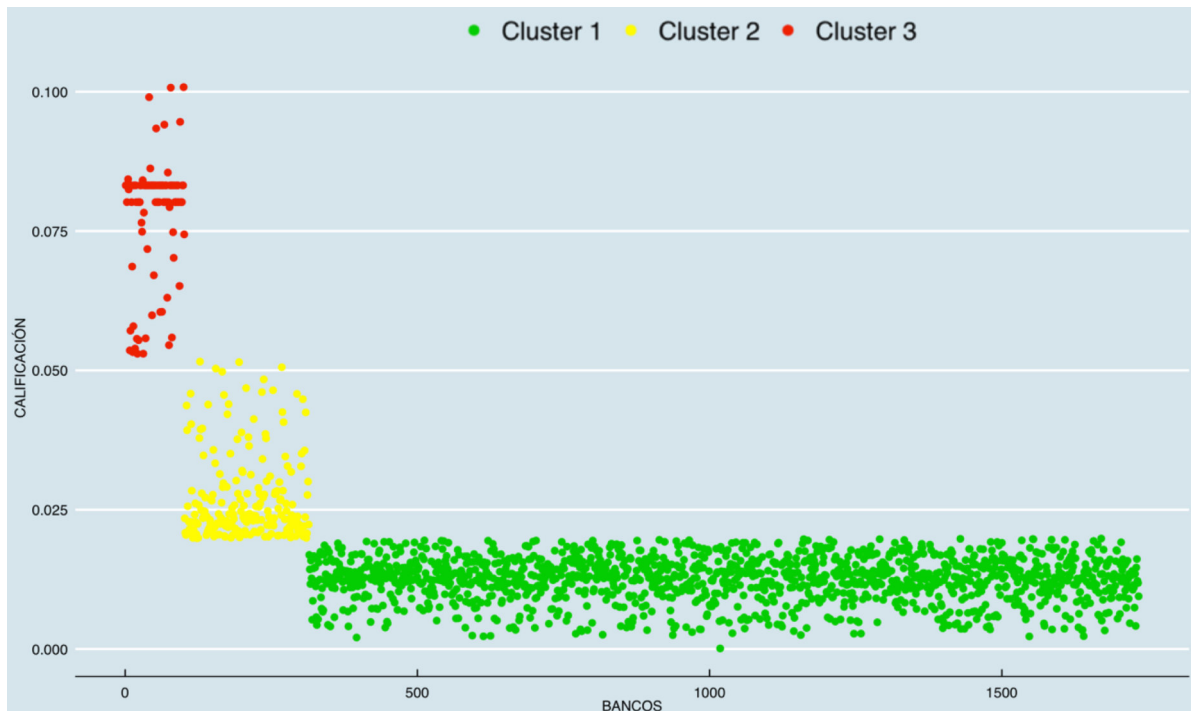


**Figura 3.8:** Algoritmo Kamean con 3 Centroides

**Elaboración:** El autor

En la Figura 3.8 se presenta de color *verde* (Clúster 1) las observaciones que cumplen con un buen desempeño en base a sus indicadores financieros, mientras que los bancos que han obtenido un desempeño por debajo de lo regular se subdividen en dos grupo *amarillo* y *rojo*. Los elementos que se encuentra de color rojo tienen el peor desempeño a lo largo del tiempo y muchos de ellos fueron los bancos que en cierto período quebraron o tuvieron un alto riesgo de quiebra.

La Figura 3.9 crea una mejor perspectiva visual de la cantidad de bancos que se encuentran agrupados en cada uno de los estados, lo que levanta la pregunta sobre cuál es el número correcto de centroides a ser utilizados para la muestra en estudio.



**Figura 3.9:** Algoritmo K-means con 3 Centroides Agrupados

**Elaboración:** El autor

Dentro de cada uno de los clústers que han sido creados existe diferentes comportamientos de las variables utilizadas para el modelo, un visión general de ello se puede encontrar en el **Anexo 7:** Gráfico de araña para las principales variables diferenciadas por grupo.

**Método ElBow:**

El método *El Codo* o *ElBow* es uno de los procedimientos más conocidos dentro del análisis de conglomerados pues consiste en graficar la variación en función de la cantidad de conglomerados a fin de encontrar el ángulo obtuso de menor valor entre todas las variaciones y con esto identificar el número de grupos óptimos a ser seleccionados. Para nuestro caso de estudio se realiza el método *ElBow* de manera gráfica para verificar la cantidad correcta de estados de fragilidad a ser utilizada para los datos obtenidos en base de calificaciones los bancos en el modelo de Firth.

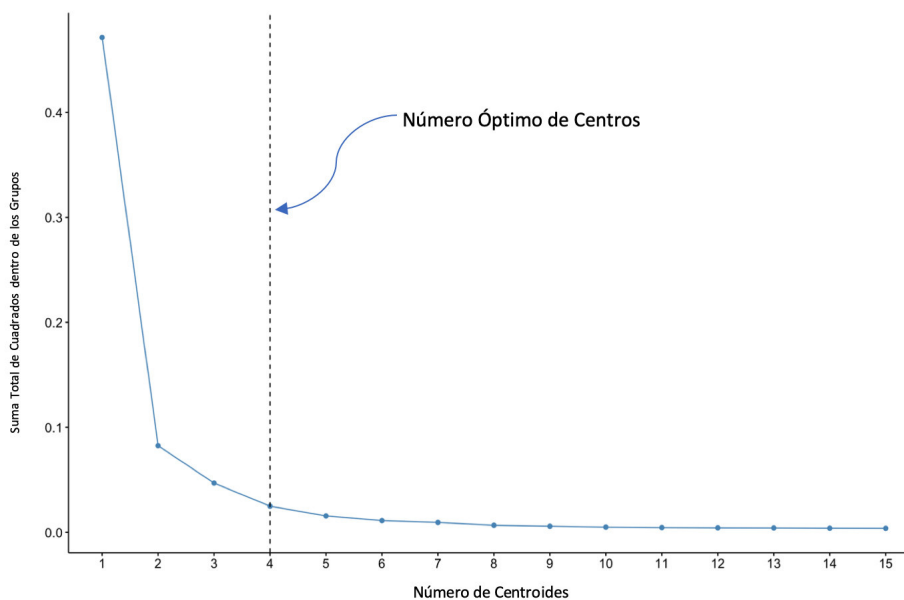
La idea dentro de la Figura 3.10 es determinar el número óptimo de conglomerados, me-

diante definición de los clústers de tal manera que la variación total dentro del clúster se minimice, esto es:

$$\text{minimizar} \left[ \sum_{k=1}^K W(C_k) \right]$$

Donde  $K$  es el número máximo de clústers asignados,  $C_k$  es el  $k$ -ésimo clúster y  $W(C_k)$  es la variación dentro de cada conglomerado. "El método del codo es un método que analiza el porcentaje de varianza explicada en función del número de conglomerados. Este método existe sobre la idea de que uno debe elegir un número de centroides para que, si se agrega otro centroide, esta variación no genere un mejor resultados de los datos." (Bholowalia y Kumar, 2014).

El porcentaje de varianza explicado  $WSS$  por los conglomerados se grafica contra el número de conglomerados.

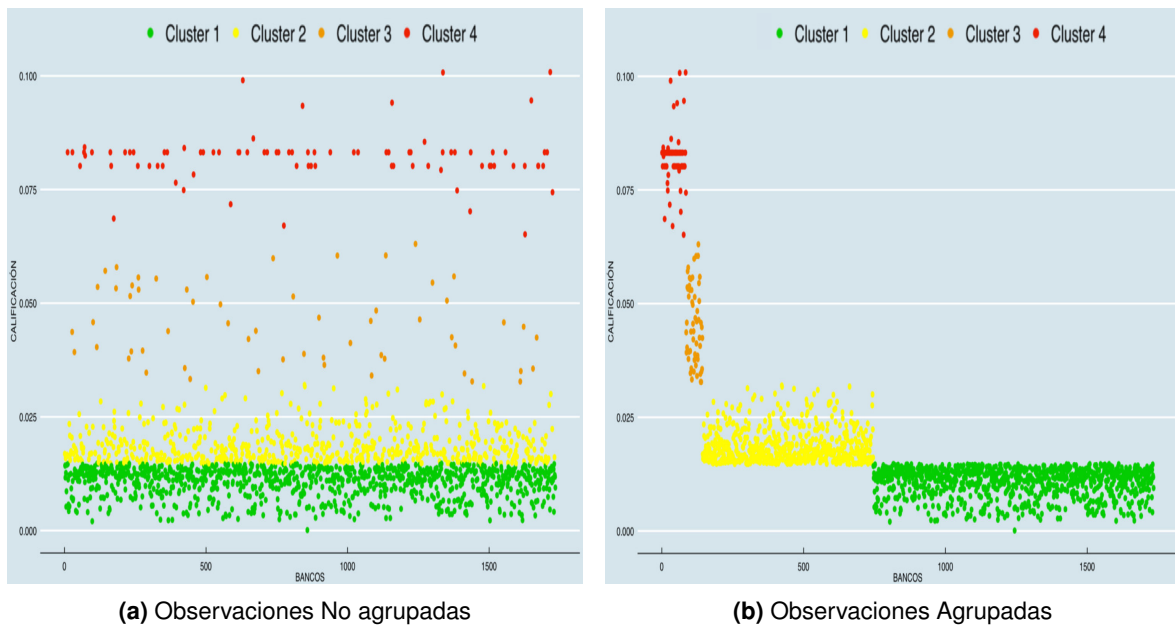


**Figura 3.10:** Método ElBow de las calificaciones obtenidas del modelo

**Elaboración:** El autor

Finalmente, a partir de la curva obtenida podemos ver cómo a medida que se aumenta la

cantidad de centroides, el valor de  $WSS$  disminuye de tal forma que la gráfica adopta una forma de codo. Para la correcta selección del valor óptimo de  $k$ , se toma el punto en donde ya no se dejan de producir variaciones importantes del valor de  $WSS$  al aumentar  $k$ . Para el caso de estudio, podemos ver que esto se produce a partir de  $k > 4$ .



**Figura 3.11:** Agrupación de Bancos con 4 centroides.

**Elaboración:** El autor

Gracias a la Figura 3.11 podemos ver que existe una mayor aglomeración de observaciones de los grupos 1 y 2 es decir en los estados de fragilidad bancaria *verde* y *amarillo*, dicho resultado era un pronóstico esperado pues el modelo de Firth lo que busca es predecir valores atípicos, que en nuestro caso son las observaciones del grupo o estado de color *rojo* en donde se ve la menor aglomeración de observaciones.

### 3.4 PREDICCIÓN DE ESTADOS DE FRAGILIDAD BANCARIA

Finalmente, ya que tenemos nuestro ‘semáforo’ de estados, es decir, los umbrales de alerta o de seguridad donde cada uno de los bancos operativos se encuentra al momento, lo que se busca es pronosticar si un banco que se encuentra en uno de estos niveles estará en el mismo nivel o en otro dentro de los siguientes periodos. Para ello, una muy útil metodología con la que se trabaja son los Modelos de Markov, en particular los Modelos Ocultos de Markov o HMM (por sus siglas del inglés, Hidden Markov Model), donde se requiere la información de cada nivel obtenido en el proceso anterior y también un conjunto de datos que llamaremos señales, que para nuestro caso serán los datos ocultos que requiere el HMM. Des esta forma podemos estudiar y entender las transiciones de estado y las diferentes propiedades probabilísticas del sistema financiero nacional.

#### Construcción de la Señal

Una de las brechas más relevantes que dejó la formulación y creación del modelo estadístico fue la inclusión de todas las variables primarias del mercado financiero entre las cuales la **Liquidez Bancaria**<sup>[3]</sup> quedó excluida del modelo por falta de significancia. En primera instancia se espera que una variable con alta influencia en el mercado global pueda ser incluida en el modelo, por ello el índice **Liquidez** será tomada como la señal dentro del análisis y desarrollo del Modelo Oculto de Markov.

Así bien, para el uso de una señal en el Modelo Oculto de Markov esta debe tomar valores

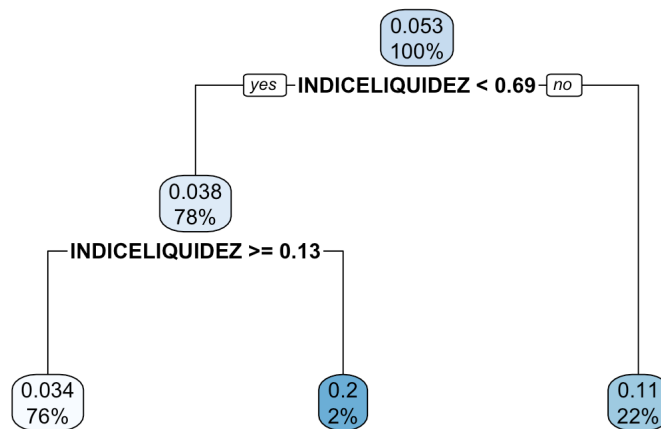
---

<sup>[3]</sup> Liquidez Bancaria: El índice de liquidez indica cuantas unidades monetarias (USD) tiene la entidad financiera en bienes y derechos del activo corriente por cada unidad monetaria(USD) que tiene de deuda a corto plazo. Su fórmula es: (activo circulante – stock) /pasivo circulante

de carácter discreto por ejemplo:

- i) **Liquidez Buena:** Cuando el índice de Liquidez es superior a 1. Esto es, el activo circulante es mayor que el pasivo corriente o en otras palabras que la entidad financiera sea capaz de pagar sus deudas sin comprometer las inversiones.
  
- ii) **Liquidez Mala:** Cuando el índice de liquidez es cercano a 0. Esto quiere decir que un banco tiene problemas para hacer frente a sus obligaciones contraídas a corto plazo (menor a un año).

Se realiza el análisis mediante un árbol de decisión para saber el corte de quiebre de la liquidez (Buena-Mala) en base a las observaciones de las entidades financieras para definir si se encuentra con Liquidez Buena o Mala.



**Figura 3.12:** Árbol de decisión índice de Liquidez Bancaria

**Elaboración:** El autor

El índice de liquidez dentro de todas las observaciones en estudio tiene un comportamiento donde su punto de corte es 0.69, esto quiere decir que segmentamos los bancos de la

siguiente forma:

- \* Si el banco en el tiempo  $t_n$  se encuentra sobre el 69 % de liquidez, será tratado como un banco liquidez alta que. Por facilidad se le denotará como  $L.A$
- \* Si el banco en el tiempo  $t_n$  se encuentra debajo del 69 % de liquidez, será tratado como un banco de liquidez baja que por facilidad se le denotará como  $L.B$

Importante notar que en base al árbol de decisión se debería tomar 3 tipos de señales por ejemplo liquidez alta, media y baja; pero debido a que el porcentaje de bancos que se encuentran en uno de los niveles es no significativo para el análisis, lo unificamos para tomar sólo las primeras dos ramas del árbol de decisión.

### 3.4.1 Algoritmo Forward

En base al proceso explicado en la sección 2.5.2 vamos a obtener los resultados gracias a los datos del modelo.

#### Datos Algoritmo Forward

Para el modelo  $\lambda$  se requiere 3 datos necesarios, que al obtenerlos en base a los datos del modelo son:

$$\pi = \begin{pmatrix} 0.30082 \\ 0.21703 \\ 0.35348 \\ 0.12866 \end{pmatrix} \quad (3.1)$$

$$A = \begin{pmatrix} 0.68812 & 0.2000 & 0.015775 & 0.015205 \\ 0.12375 & 0.5000 & 0.052817 & 0.013835 \\ 0.08437 & 0.2125 & 0.718028 & 0.165753 \\ 0.16594 & 0.0625 & 0.213239 & 0.805068 \end{pmatrix} \quad (3.2)$$



$$B = \begin{pmatrix} 0.41228 & 0.58771 \\ 0.77674 & 0.22325 \\ 0.79115 & 0.20884 \\ 0.69456 & 0.30543 \end{pmatrix} \quad (3.3)$$

Ahora vamos a revisar los resultados obtenidos en base a las posibles secuencias de señales emitidas, tomando en cuenta que el cálculo de la variable **señal** es dicotómica de la forma:

$$Liquidez = \begin{cases} LB & \text{si índice de liquidez} < 0.69 \\ LA & \text{si índice de liquidez} \geq 0.69 \end{cases} \quad (3.4)$$

*LB: Liquidez Baja*

*LA: Liquidez Alta*

Es importante resaltar que para la lectura adecuada de los resultados del algoritmo, la periodicidad que fueron tomados los datos es trimestral, es decir que para todas las predicciones de *n-periodos* es equivalente a analizar los siguientes *n-trimestres*. La primera secuencia que se analizará el comportamiento es:

**1era:** (LB, LB, LB):

LB-LB-LB	
Estado Semáforo	Resultado
Verde	11.00 %
Amarillo	29.37 %
Naranja	38.05 %
Rojo	21.56 %

**Tabla 3.4:** Probabilidad de ocurrencia de estados con secuencia de señales LB-LB-LB

De donde se interpreta que de darse una secuencia de señales *LB-LB-LB*, el 38.05 % de los bancos probablemente estarían en estado naranja.

**2da:** (LB,LA,LB):

<b>LB-LA-LB</b>	
<b>Estado Semáforo</b>	<b>Resultado</b>
Verde	22.89 %
Amarillo	28.72 %
Naranja	27.44 %
Rojo	20.93 %

**Tabla 3.5:** Probabilidad de ocurrencia de estados con secuencia de señales LB-LA-LB

Se comprende que el 28.7 % de los bancos, en caso de tener una secuencia de señales *LB-LA-LB*, estarían en estado amarillo.

**3era:** (LB,LB,LA):

<b>LB-LB-LA</b>	
<b>Estado Semáforo</b>	<b>Resultado</b>
Verde	35.93 %
Amarillo	19.33 %
Naranja	23.01 %
Rojo	21.72 %

**Tabla 3.6:** Probabilidad de ocurrencia de estados con secuencia de señales LB-LB-LA

Podemos ver que en caso de que una sucesión de señales de la forma *LB-LB-LA* preceda

a un banco, este tiene un 21.7 % de estar en estado rojo.

**4ta:** (LB,LA,LA):

LB-LA-LA	
Estado Semáforo	Resultado
Verde	59.91 %
Amarillo	14.40 %
Naranja	12.63 %
Rojo	16.06 %

**Tabla 3.7:** Probabilidad de ocurrencia de estados con secuencia de señales LB-LA-LA

Si tenemos una secuencia de señales de la forma *LB-LA-LA* la probabilidad de que un banco se encuentre en estado Amarillo es de 14.4 % y en estado Naranja es de 13.63 %.





**5ta:** (LA,LA,LA):

LA-LA-LA	
Estado Semáforo	Resultado
Verde	71.85 %
Amarillo	12.68 %
Naranja	6.25 %
Rojo	9.21 %

**Tabla 3.8:** Probabilidad de ocurrencia de la con secuencia de señales LA-LA-LA

Si en los últimos 3 períodos de tiempo tenemos una secuencia tal que se repite (*LA-LA-LA*) entonces un 71.85 % de los bancos se encontrarán en estado verde.

En la siguiente tabla de calor en base a las 5 de las 8 posibles secuencias de señales que podemos obtener al revisar 3 periodos anteriores, podemos analizar con una visión global la probabilidad de ocurrencia de un estado basado en una secuencia de señales.

Estado del Semáforo		Transición de Señales				
		LB-LB-LB	LB-LA-LB	LB-LB-LA	LB-LA-LA	LA-LA-LA
	<i>Verde</i>	11%	23%	36%	60%	72%
	<i>Amarillo</i>	29%	29%	19%	14%	13%
	<i>Naranja</i>	38%	27%	23%	13%	6%
	<i>Rojo</i>	22%	21%	22%	16%	9%

**Figura 3.13:** Comparación de alerta de señales

En la figura 3.12 podemos observar mediante una serie de 3 señales en los tiempos  $n - 2, n - 1, n$  qué porcentaje de los bancos están en uno de los estado de alerta (Verde-Amarillo-Naranja-Rojo) dependiendo de las señales emitidas 3 periodos hacia atrás.

El objetivo es poder discriminar el riesgo de quiebra de los bancos, basados netamente en una secuencia de señales emitidas para así obtener el porcentaje de bancos concentrados en cada uno de los 4 estados.

Por ejemplo, la interpretación sobre las posibles series de señales es:

**LB-LB-LB:** Si se emite por 3 veces consecutivas una señal de LIQUIDEZ BAJA **LB** conllevará a que el:

- \* 11 % de los bancos se encuentren en estado verde.
- \* 29 % de los bancos se encuentren en estado amarillo.
- \* 38 % de los bancos se encuentren en estado naranja.

- \* 22 % de los bancos se encuentren en estado rojo

**LB-LB-LA:** Si se emite por 2 veces consecutivas una señal de LIQUIDEZ BAJA **LB** y finalmente una de LIQUIDEZ ALTA **LA** conllevará a que el:

- \* 36 % de los bancos se encuentren en estado verde.
- \* 19 % de los bancos se encuentren en estado amarillo.
- \* 23 % de los bancos se encuentren en estado naranja.
- \* 22 % de los bancos se encuentren en estado rojo

**LA-LA-LA:** Si se emite por 3 veces consecutivas una señal de LIQUIDEZ ALTA **LA** conllevará a que el:

- \* 72 % de los bancos se encuentren en estado verde.
- \* 13 % de los bancos se encuentren en estado amarillo.
- \* 6 % de los bancos se encuentren en estado naranja.
- \* 9 % de los bancos se encuentren en estado rojo

Podemos analizar de otra forma los resultados, viendo desde una perspectiva propia de cada estado del semáforo de fragilidad bancaria:

- \* Lo que se requiere para tener una estabilidad financiera "*verde*" es que en los últimos 3 periodos, la liquidez bancaria esté consecutivamente alta (sobre el 69 %), para así garantizar que el 72 % de los bancos se encuentren con solidez bancaria.
- \* Al tener dos señales de tiempos de Liquidez Baja en los últimos 3 períodos, un banco corre el riesgo de estar sobre el 20 % de posibilidad de estar en esta naranja con respecto a un quiebra institucional.

- \* Con el fin de evitar encontrarse en estado de posible quiebra (rojo), es necesario que en los último 3 períodos nuestra liquidez como mínimo haya sido ALTA por 2 veces, si no es así pues existe una probabilidad mayor 16% de estar en un estado crítico.

### 3.4.2 Algoritmo Viterbi

Los resultados del Algoritmo Viterbi se basan en el desarrollo de la sección 2.5.3, en donde lo primero es poder contar con la base de datos (matrices) para iniciar el algoritmo.

#### Datos Algoritmo Viterbi

Para el modelo  $\lambda$  se requiere 3 datos  $\pi$ ,  $A$ ,  $B$ , los cuales vienen dados por las siguientes matrices de información.

$$\pi = \begin{pmatrix} 0.30082 \\ 0.21703 \\ 0.35348 \\ 0.12866 \end{pmatrix}$$

$$A = \begin{pmatrix} 0.68812 & 0.2000 & 0.015775 & 0.015205 \\ 0.12375 & 0.5000 & 0.052817 & 0.013835 \\ 0.08437 & 0.2125 & 0.718028 & 0.165753 \\ 0.16594 & 0.0625 & 0.213239 & 0.805068 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.41228 & 0.58771 \\ 0.77674 & 0.22325 \\ 0.79115 & 0.20884 \\ 0.69456 & 0.30543 \end{pmatrix}$$









De esta forma, examinemos por separado varias de las secuencias de señales probables con la finalidad de obtener la predicción de los estados en los siguientes 3 períodos, lo que se busca es analizar de manera independiente la influencia que tienen las señales al momento pronosticar los estados de fragilidad financiera.

Sucesión de señales	Predicción de estados
LB-LB-LB	Naranja - Naranja - Naranja
LB-LA-LB	Naranja - Verde - Rojo
LA-LA-LA	Verde - Verde - Verde

**Tabla 3.9:** Predicción de estados con la secuencia LB-LB-LB

La Tabla 3.9 nos presenta 3 ejemplos de sucesiones consecutivas de tres señales ya sea LB o LA y las predicciones de estados que estas sucesiones implican. Por ejemplo, si tenemos 3 períodos consecutivos donde contamos con Liquidez Alta (LA) esto generará 3 estados futuros consecutivos "muy buenos" es decir verdes.

Con el fin de tener una visibilidad general del comportamiento de los resultados verificaremos que para una secuencia de 3 señales consecutivas podemos obtener 8 tipos de secuencias diferentes en las que se pueden alternar entre Liquidez Baja y Alta, entonces tomando para los periodos  $n - 2, n - 1$  y  $n$  se tiene la siguiente tabla de análisis:

Secuencia de Señales		Estado Pronosticado al tiempo		
		$t_{n+1}$	$t_{n+2}$	$t_{n+3}$
	LB-LB-LB	Naranja	Naranja	Naranja
	LB-LB-LA	Naranja	Naranja	Verde
	LB-LA-LB	Naranja	Verde	Rojo
	LB-LA-LA	Naranja	Verde	Verde
	LA-LB-LB	Rojo	Rojo	Rojo
	LA-LB-LA	Verde	Verde	Verde
	LA-LA-LB	Verde	Verde	Amarillo
	LA-LA-LA	Verde	Verde	Verde

**Figura 3.14:** Estados Pronosticados en base a la secuencia de señales

La figura anterior muestra, en base a 3 períodos de señales, cuáles serán los estados pronosticados en los siguientes 3 ciclos de análisis. De esta forma y con el fin de analizar los resultados basándose propiamente en la influencia que genera la señal para la predicción de los estados de fragilidad financiera tenemos el análisis.



## **Interpretación basada en los resultados obtenidos por el algoritmo Viterbi**

Con el objetivo de resumir el detalle de la interpretación, exclusivamente para siguiente parte denotaremos como:

**V:** *Verde*

**A:** *Amarillo*

**N:** *Naranja*

**R:** *Rojo*

- I Las entidades bancarias que vengan con una señal de Liquidez Baja consecutiva los últimos 3 períodos, seguirán una secuencia de estados N-N-N.
- II Si un banco en sus últimos 3 períodos sigue una secuencia de señales LB-LB-LA, sus estados están directamente relacionados ya que tendrán un pronóstico N-N-V.
- III Se prevé que los bancos que sigan una secuencia de señales LB-LA-LB generarán que en el tercer período de pronóstico se tenga un estado crítico con alta probabilidad de quiebra financiera de la forma N-V-R.
- IV Los bancos que tengan una mejora en su liquidez en los últimos 3 períodos de la forma LB-LA-LA, tendrán una relación directa a los estados que se pronostica en los 3 siguientes ciclos de la forma N-V-V.
- V Si una entidad bancaria tiene un decrecimiento en su liquidez en los últimos tres períodos de la forma LA-LB-LB entonces se pronostica un estado crítico (rojo) durante los 3 períodos y deberá levantar planes de acción urgentes como alerta ante una posible quiebra financiera.
- VI Para un banco que viva una volatilidad de liquidez de la forma LA-LB-LA se pronostica una buena estabilidad bancaria en los siguientes 3 períodos con todos sus estados en verde.

- VII Se prevé que un banco que tenga una estabilidad en su Liquidez por dos períodos y tenga una baja en su Liquidez al tercer período genere que al final de su tercer ciclo de pronóstico tenga un decrecimiento en su estabilidad bancaria, manejable a nivel financiero de la forma V-V-A.
- VIII Las entidades financieras que tengan una Liquidez sobre el 69 % durante 3 períodos consecutivos se estima que en sus siguientes 3 ciclos mantenga una buena estabilidad financiera de la forma V-V-V.

## Validación de Resultados Predicción vs. Real

Vamos a tomar el período  $t_{n+1}$  posterior a los resultados de toda la base en observación para verificar, mediante una muestra de bancos, que cumplan con el pronóstico del algoritmo de Viterbi. Tomemos en cuenta que los resultados del Score obtenidos de un banco que forman parte de un grupo (Estado) están delimitados por los umbrales establecidos por el algoritmo de agrupación K-means y estos son los siguientes:

Umbrales de Estados		
Estado Semáforo	Límite Inferior	Límite Superior
Verde	0.000103	0.0147
Amarillo	0.0147	0.0320
Naranja	0.0321	0.0650
Rojo	0.0651	0.101

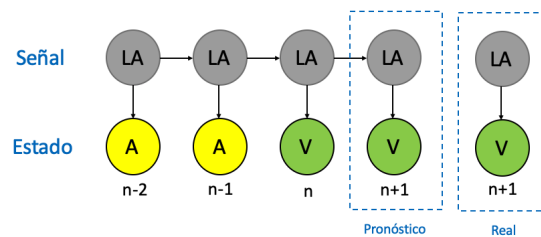
**Tabla 3.10:** Umbrales generados por el Algoritmo K-means de cada estado

Con esta información vamos a verificar para una muestra de 4 bancos el comportamiento en el periodo  $t_{n+1}$  de la base de observaciones para así identificar a qué estado pertenece y a su vez observar el valor de la señal, esto tomando en cuenta la *Figura 3.12* en donde se define el % de quiebre de la variable liquidez para ser tanto Liquidez Alta como Liquidez Baja, con ello tenemos la siguiente información:

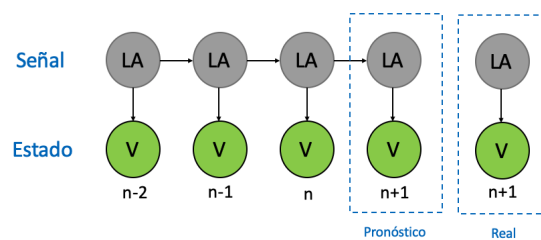
Banco	Calificación Score	Estado	Señal
Pacífico	0.012581194	Verde	81.83 % (LA)
Produbanco	0.00525799	Verde	72.04 % (LA)
Litoral	0.06281953	Naranja	47.40 % (LA)
Capital	0.0910082	Rojo	79.61 % (LB)

**Tabla 3.11:** Verificación de la calificaciones de 4 bancos en el periodo  $t_{n+1}$

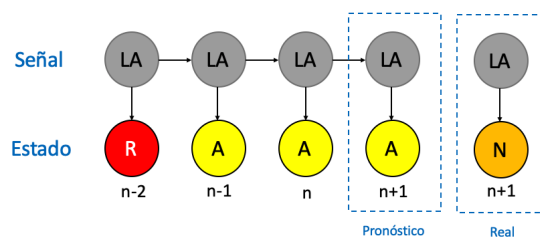
Las siguientes figuras muestran los resultados de los 3 últimos periodos tanto del Score (Grupo de Fragilidad Bancaria) como de la señal (Liquidez) de los bancos en análisis en donde se puede apreciar la diferencia entre el pronóstico versus la realidad obtenida:



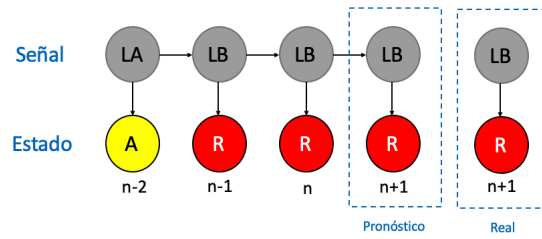
**Figura 3.15:** Resultados Pronóstico Produbanco Señal-Estado



**Figura 3.16:** Resultados Pronóstico Banco del Pacífico Señal-Estado



**Figura 3.17:** Resultados Pronóstico Banco del Litoral Señal-Estado



**Figura 3.18:** Resultados Pronóstico Banco Capital Señal-Estado

Como se puede observar para 3 de los 4 bancos el modelo realizó la predicción correcta del estado en el cual se encontraría en el tiempo  $t_{n+1}$ . Para el Banco del Pacífico, Produbanco y Banco Capital la predicción fue la misma que al tomar los resultados reales de los bancos y por otro lado para el Banco del Litoral se tuvo una brecha entre los resultados de la predicción y los resultados reales del banco debido a la reincidencia de la señal LA en el transcurso del tiempo  $t_n$  al  $t_{n+1}$ .

Ahora bien, para entender el comportamiento global del modelo vamos a realizar el comparativo entre la proyección del modelo netamente basado en los resultados del score de Firth versus la misma proyección pero bajo la influencia de una variable externa que la conocemos como señal.

En la siguiente figura se evidencian los Estados pronosticados por el modelo de Firth para 3 períodos futuros basados propiamente en los resultados del modelo, es decir sin la influencia de una variable externa.

BANCO \ TIEMPO	Señales			Estados		
	n-3	n-2	n-1	n	n+1	n+2
AMAZONAS	LA	LA	LA	V	V	V
AUSTRO	LA	LA	LA	V	V	V
BOLIVARIANO	LB	LB	LA	N	N	V
CAPITAL	LA	LA	LB	R	R	V
CITIBANK	LB	LB	LB	N	N	N
COMERCIAL DE MANABI	LA	LA	LA	V	V	V
COOPNACIONAL	LB	LA	LB	N	V	R
D-MIRO	LB	LA	LA	N	V	V
DE GUAYAQUIL	LA	LA	LA	V	V	V
DELBANK S.A.	LA	LA	LA	V	V	V
FINCA	LA	LB	LA	V	V	V
GENERAL RUMIÑAHUI	LB	LA	LA	N	V	V
INTERNACIONAL	LA	LA	LA	V	V	V
LITORAL	LA	LA	LA	A	N	V
LOJA	LB	LB	LB	N	N	N
MACHALA	LA	LA	LA	V	V	V
PACIFICO	LA	LA	LA	V	V	V
PICHINCHA	LA	LA	LA	V	V	V
PROCREDIT	LA	LB	LA	V	V	V
PRODUBANCO	LA	LA	LA	V	V	V
SOLIDARIO	LB	LB	LB	N	N	N

**Figura 3.19:** Resultados pronosticado por el Score de Firth

Como un dato adicional se incluyó las señales de los períodos anteriores para tener en cuenta la influencia que puede llegar a tener este componente adicional.

Con esto se procede a realizar el mismo pronóstico pero ahora se puede ver el peso que tiene la variable señal dentro del Modelo Oculto de Markov.

BANCO\ TIEMPO	Señales			Señal	Estado	Señal	Estado	Señal	Estado
	n-3	n-2	n-1	n		n+1		n+1	
AMAZONAS	LA	LA	LA	LA->	V	LA->	V	LA->	V
AUSTRO	LA	LA	LA	LA->	V	LA->	N	LA->	V
BOLIVARIANO	LB	LB	LA	LB->	N	LA->	A	LB->	V
CAPITAL	LA	LA	LB	LB->	R	LB->	R	LB->	R
CITIBANK	LB	LB	LB	LB->	V	LB->	A	LB->	V
COMERCIAL DE MANABI	LA	LA	LA	LA->	V	LA->	R	LA->	V
COOPNACIONAL	LB	LA	LB	LB->	V	LB->	A	LB->	V
D-MIRO	LB	LA	LA	LB->	V	LA->	N	LA->	V
DE GUAYAQUIL	LA	LA	LA	LB->	V	LA->	V	LA->	V
DELBANK S.A.	LA	LA	LA	LB->	V	LB->	N	LB->	V
FINCA	LA	LB	LA	LB->	V	LB->	N	LB->	V
GENERAL RUMIÑAHUI	LB	LA	LA	LA->	V	LA->	V	LB->	V
INTERNACIONAL	LA	LA	LA	LA->	V	LA->	A	LA->	V
LITORAL	LA	LA	LA	LA->	N	LA->	N	LA->	A
LOJA	LB	LB	LB	LB->	V	LB->	N	LB->	V
MACHALA	LA	LA	LA	LA->	V	LA->	V	LA->	V
PACIFICO	LA	LA	LA	LA->	V	LA->	V	LA->	V
PICHINCHA	LA	LA	LA	LA->	V	LA->	V	LA->	V
PROCREDIT	LA	LB	LA	LA->	V	LB->	A	LA->	V
PRODUBANCO	LA	LA	LA	LA->	V	LA->	V	LA->	V
SOLIDARIO	LB	LB	LB	LB->	V	LB->	N	LB->	N

**Figura 3.20:** Resultados pronosticados por el Score de Firth bajo la influencia de la señal Índice de Liquidez

La figura anterior muestra los resultados arrojados por el modelo de Firth pero bajo la influencia de la variable externa *índice de liquidez* la misma que fue tomada como señal dentro del Modelo de Markov Oculto. Ahora bien para analizar la influencia que puede llegar a tener una variable externa para el modelo de pronóstico de estados de estabilidad financiera observemos la Figura 3.21 que nos da una información muy valiosa dentro del estudio que se lleva a cabo, debido a que nos proporciona el % de acierto de cada estado, por ejemplo para el estado verde existe una estimación de 20.93 % que fue sobre-estimada es decir el 18.6 % de los bancos que el modelo, bajo influencia de la señal pronosticó como verde, el Modelo de Firth (sin señal) lo pronosticó como Naranja y el 2.33 % lo pronosticó como Rojo dando así un mayor valor de riesgo a este grupo de bancos y de esta forma el 79,07 % no tuvo influencia de la variable liquidez en su pronóstico del estado verde.

Como otro ejemplo podemos ver que el estado Naranja tiene un acierto del 50 % y el otro 50 % indica que la señal experimentada influyó en una sub-estimación del estado verdadero

ya que tanto el 40 % y 10 % que el modelo (con señal) pronosticó como Naranja, el modelo de Firth (sin señal) lo pronosticó como Verde y Amarillo respectivamente dando un menor valor de riesgo a este grupo de bancos.

Comparación modelo de Firth Con Señal vs Sin Señal		Modelo de Firth			
		V	A	N	R
Modelo de Firth + Señal	V	79,07%	0,00%	18,60%	2,33%
	A	66,67%	0,00%	33,33%	0,00%
	N	40,00%	10,00%	50,00%	0,00%
	R	50,00%	0,00%	0,00%	50,00%

**Figura 3.21:** Acierto del Modelo de Firth con señal vs sin señal, por estado

Si bien el modelo de Firth funciona de manera correcta, para alcanzar el objetivo de tener un modelo de alerta temprana sobre riesgo institucional financiero mediante Modelos de Markov Ocultos, es necesario la inclusión de la variable liquidez dentro del modelo, la cual como lo hemos visto en la figura anterior tiene una influencia significativa en cada uno de los niveles ya que ha generado cambios de estados propiamente influenciados por la señal a la que se ha visto sometido el banco en el período analizado generando una sobre-estimación o sub-estimación del riesgo financiero de los bancos en estudio. Como lo afirma el Comité de Supervisión Bancaria de Basilea: " El LCR<sup>[4]</sup> mejorará la capacidad del sector bancario para absorber perturbaciones procedentes de tensiones financieras o económicas de cualquier tipo, reduciendo con ello el riesgo de contagio desde el sector financiero hacia la economía real."

Finalmente, basados en la información que ha proporcionado el Algoritmo Forward, el Algoritmo Viterbi y este último comparativo del pronóstico del Modelo de Firth se tiene como

<sup>[4]</sup> LCR(Liquidity Coverage Ratio): Conocido en su traducción al español como Coeficiente de Cobertura de Liquidez que es la proporción de activos de alta liquidez que las instituciones financieras tienen con el fin de asegurar su capacidad para cumplir con sus obligaciones a corto plazo.



consecuencia un modelo robusto que busca abarcar la mayoría de datos o información para tener una visión amplia del comportamiento del riesgo institucional financiero al cual están sometidos los bancos ecuatorianos basado en la influencia y peso que tiene cada una de las variables macro y micro económicas en estudio.

"La crisis reafirmó la importancia de la liquidez para el adecuado funcionamiento de los mercados financieros y el sector bancario. Antes de la crisis, los mercados de activos bullían de actividad y era fácil obtener financiación a precios bajos. El súbito deterioro de las condiciones del mercado demostró que la liquidez puede evaporarse rápidamente y que la situación de iliquidez puede prolongarse considerablemente.". Basilea III, 2013.

## 4 CONCLUSIONES Y RECOMENDACIONES

Es menester contar con un modelo de alerta temprana para riesgos institucionales financieros que provea tanto al Estado como a cada uno de los bancos un estatus de comportamiento cronológico de sus indicadores económicos enfocado en la exposición ante una inestabilidad social, económica y/o financiera y de esta forma que brinde las posibles oportunidades a ser tomadas para la mejora y desarrollo de la institución financiera.

El Modelo de Firth sirve muy bien para crear niveles o semáforos de las calificaciones de los bancos pero para que estos sean ubicados en un sistema de alerta temprana se hace necesaria la inclusión de la variable liquidez como una señal para nuestro caso basados en los Modelos Ocultos de Markov, de esta forma se se corrige una posible sobre-estimación o sub-estimación del riesgo financiero pronosticado por el modelo de regresión logística bajo estimación propuesto por Firth y utilizado en el trabajo de Uquillas y Flores (2020).

Dentro del modelo inicial utilizado una de las mayores brechas encontradas fue la exclusión del índice de liquidez bancaria basando en análisis en pruebas estadísticas del modelo, por lo que fue una de las primeras opciones para generar la señal del modelo de Markov y como es una variable relevante a nivel nacional y propia de cada banco, cumplió su función de generar una alerta o señal al modelo sobre los pronósticos de movimiento entre estados de un banco.

La inestabilidad financiera provocada en los años 1999-2001 proporcionan varias alertas de comportamiento bancario para poder controlar de manera preventiva o reactiva varios de los indicadores manejados a nivel interno de cada banco como por ejemplo el Apalancamiento o ROE.

El utilizar una matriz de transición entre cada nivel de alerta, tomada en base a todos los resultados de las observaciones del score, ha proporcionado un mejor desenvolvimiento para pronosticar reportes mayores a un período.

Por el hecho de hacer un enfoque en la ocurrencia de eventos raros ha surgido una combinación de metodologías de diferentes ramas de matemática para alinear recursos con el fin de proveer al sistema bancario una nueva herramienta de gestión ante una posible inestabilidad financiera nacional o local por lo que los resultados obtenidos a lo largo del proyecto han cumplido con el fin las preguntas y objetivos planteados, basado en la solidez que tiene cada proceso, el modelo matemático, la agrupación mediante Machine Learning y los pronósticos basados en Procesos Estocásticos.

Se puede buscar una o varias nuevas señales o a su vez una combinación de ellas para incluirla en el Modelo Oculto de Markov como señal con el fin de ver el contraste e influencia de cada una de ellas para el pronóstico de los nuevos estados.

## 5 REFERENCIAS BIBLIOGRÁFICAS

- Banco Central del Ecuador - BCE (2008). Memorias anuales 1980 - 2007. Quito: Estadísticas Banco Central del Ecuador.
- Bholowalia y Kumar., (2014). EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN. International Journal of Computer Applications. Punjab, India.
- Bolsa de Valores de Quito (2019). Recuperado de <https://www.bolsadequito.com/>
- Buenaventura Vera, Guillermo (2003). La Tasa de Interés: Información con Estructura. Estudios Gerenciales, (86),39-50.
- Comité de Supervisión Bancaria de Basilea (2013). Coeficiente de cobertura de liquidez y herramientas de seguimiento del riesgo de liquidez. Banco de Pagos Internacionales
- Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. Biometrika.
- Flores Francis R., (2018). Quiebra bancaria: Un análisis econométrico para el Ecuador. Escuela Politécnica Nacional.
- Gonzalez Hermsillo B., (1999). Determinants of Ex-Ante Banking System Distress: A Macro-micro Empirical Exploration of Some Recent Episodes. IMF Working Paper.
- Gujarati y Porter (2010), Econometría, 5ta edición.
- Jurafsky Daniel y Martin James H., (2018). Speech and Language Processing.
- King Zeng (2001), Logistic Regression in Rare Events Data.
- Kulkarni V.,(2001). Introduction to Modeling and Analysis of Stochastic Systems.
- Mays, E., Nuetzel, P., (2004). Credit Scoring for Risk Managers: The Handbook for Len-

ders. South-Western Publishing, Mason, OH, Ch. Scorecard Monitoring Reports, pág. 201–217.

- Pacific Credit Rating (2015). Metodología de Calificación de Riesgo de Bancos e Instituciones Financieras Ecuador.
- Pisón Fernández I., Puime Guillén F., y Crespo Cibran M., (2015). Crisis financieras y crisis bancarias: Revisión de literatura y aproximación a elementos de control para el futuro. Vigo: España.
- Rabiner Lawrence R, Fellow, (1989). Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition.
- Ross Sheldon M. (2010), Introduction to Probability Models. 10th Edition.
- Superintendencia de Bancos (2019). Recuperado de <https://www.superbancos.gob.ec/bancos/>
- Xu, M. (2008), Bankruptcy prediction: the case of Japanese listed companies. Springer Science+ Business Media, Japan.
- Z. Ghahramani(2001), "Introducción a los modelos ocultos de Markov y Bayesian Networks," Revista Internacional de Reconocimiento de Patrones e Inteligencia Artificial, vol. 15.
- Zhao, X; Shang, P. and Lin, A. 2016. Universal and non-universal properties of recurrence intervals of rare events. Physica A: Statistical Mechanics and its Applications 448: 132-143.

## 6 ANEXOS

### Anexo 1

Bases principales del código realizado en RStudio para el desarrollo y análisis del problema:

```
1
- MODEL1<-logistf::logistf(FORMULA1,data = BASE1)
- logistf::is.logistf(MODEL1)
-
5 strucchange::sctest(FORMULA1,data=BASE1,asymptotic=F)
- lmtest::resettest(formula=FORMULA1, data = BASE1,power = 2:3)
- car::vif(MODEL1)
- bptest(MODEL1)
-
10 InformationValue::ks_plot(BASE1$QUIEBRA,prob)
- InformationValue::precision(BASE1$QUIEBRA,prob,threshold = 0.5)
- InformationValue::ks_stat(BASE1$QUIEBRA,prob)
- InformationValue::specificity(BASE1$QUIEBRA,prob)
- InformationValue::misClassError(BASE1$QUIEBRA,prob)
15 InformationValue::confusionMatrix(BASE1$QUIEBRA,prob)
- InformationValue::sensitivity(BASE1$QUIEBRA,prob)
- InformationValue::IV(as.factor(BASE1$SPREAD),BASE1$QUIEBRA)
- InformationValue::IV(as.factor(BASE1$INFLACION),BASE1$QUIEBRA)
-
```

```

20 ScoreModel<-kmeans(Result,4)
-
- ggplot(Base2, aes(x=1:nrow(Base2),y=Calificacion))+theme_economist()+
-   geom_point(aes(colour = factor(NomCluster)),size=1.9)+
-   scale_colour_manual(values = c("green3", "yellow1","orange2","red2"))+
25 labs(x = "BANCOS",y = "CALIFICACION", title = "Segmentación de Bancos
-   ",
-   subtitle = " Ecuador Variables: N",caption = "Realizado por: Andrés
-   Bolaños")
-
- wss <- sapply(1:15, function(k){kmeans(Clust1$Result1, k)$tot.withinss})
- plot(1:15, wss,
30   type="b", pch = 10, frame = FALSE,
-   xlab="Número de clusters– Periodo de observación 1 año
-   52 BANCOS",
-   ylab="Suma total de cuadrados dentro de los grupos ")
-
35 x <- dthmm(NULL, Pi, delta, "pois", list(lambda=lambda), discrete=TRUE)
- x <- simulate(x, nsim=100)
-
- Viterbi(x)
40 FViterbi<- function(Pi, P, B, n,m,senal){}

```

**Listing 6.1:** Extracto de código utilizado en R

## Anexo 2 Prueba de Multicolinealidad

<b>Variables</b>	<b>VIF</b>	<b>1/VIF</b>
DOLARIZACION	1.648134	0.6067467
SPREAD	1.371559	0.7290972
INFLACION	1.676466	0.5964930
APALANCAMIENTO	1.146487	0.8722295
DOLARIZACION:TASAMOROSIDAD	1.480203	0.6755829
DOLARIZACION:ROE	1.028404	0.9723805

**Elaborado por:** El autor.

## Anexo 3 Prueba de Choque Estructural

<b>Chow Test</b>
Ho: Constancia de coeficientes antes y después de la dolarización
S = 1.8033
p-value = 4.393e-06

**Elaborado por:** El autor.

## Anexo 4 Prueba de Variables Omitidas RESET

<b>Ramsey Test</b>
Ho: No hay presencia de variables omitidas dentro del modelo
RESET = 6.9513
p-value = 0.0009842

**Elaborado por:** El autor.

## Anexo 5 Prueba de Heteroscedasticidad

<b>Breusch-Pagan Test</b>
Ho: Existe una varianza constante de los residuos
BP = 174.15
p-value < 2.2e-16

**Elaborado por:** El autor.



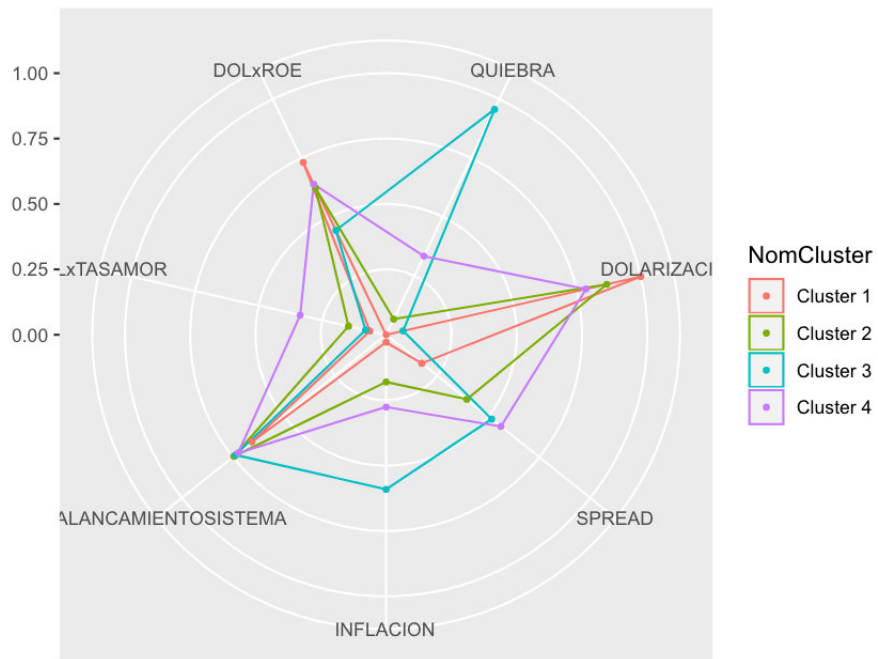
## Anexo 6 Resultados de Calificación de Riesgo vs. Firma calificadoras últimos 4 años.

Fecha	Banco	Calificación SBS	Firma Calificadora
dic-2017	BANCO AMAZONAS S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2017	BANCO DEL AUSTRO S.A.	AA- / AA	CLASS INTERNATIONAL RATING / PCR PACIFIC S.A.
dic-2017	BANCO BOLIVARIANO C.A.	AAA- / AAA-	PCR PACIFIC S.A.
dic-2017	BANCO CAPITAL S.A.	E	CLASS INTERNATIONAL RATING
dic-2017	CITIBANK N.A.	AAA	BANK WATCH RATINGS S.A.
dic-2017	BANCO COMERCIAL DE MANABI S.A.	BBB-	PCR PACIFIC S.A.
dic-2017	BANCO COOPNACIONAL S.A.	AA	SOC.CAL.RIESGO LATINOAMERICANA
dic-2017	BANCO D-MIRO S.A.	A-	PCR PACIFIC S.A.
dic-2017	BANCO DELBANK S.A.	A-	SOC.CAL.RIESGO LATINOAMERICANA
dic-2017	BANCO PARA ASISTENCIA COMUNITARIA FINCA S.A.	BBB+	PCR PACIFIC S.A.
dic-2017	BANCO GENERAL RUMIÑAHUI S.A.	AAA-	PCR PACIFIC S.A.
dic-2017	BANCO DE GUAYAQUIL S.A.	AAA / AAA-	SOC.CAL.RIESGO LATINOAMERICANA / PCR PACIFIC S.A.
dic-2017	BANCO INTERNACIONAL S.A.	AAA- / AAA	BANK WATCH RATINGS S.A./ CLASS INTERNATIONAL RATING
dic-2017	BANCO DEL LITORAL S.A.	A-	SOC.CAL.RIESGO LATINOAMERICANA
dic-2017	BANCO DE LOJA S.A.	AA+	PCR PACIFIC S.A.
dic-2017	BANCO DE MACHALA S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2017	BANCO DEL PACIFICO S.A.	AAA- / AAA-	BANK WATCH RATINGS S.A./ SOC.CAL.RIESGO LATINOAMERICANA *1
dic-2017	BANCO PICHINCHA C.A.	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2017	BANCO PROCREDIT S.A.	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2017	BANCO DE LA PRODUCCIÓN S.A.PRODUBANCO	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2017	BANCO SOLIDARIO S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2018	BANCO AMAZONAS S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2018	BANCO DEL AUSTRO S.A.	AA / AA+	PCR PACIFIC S.A./ SUMMA RATING S.A.*1
dic-2018	BANCO BOLIVARIANO C.A.	AAA- / AAA-	PCR PACIFIC S.A./
dic-2018	BANCO CAPITAL S.A.	E	PCR PACIFIC S.A.
dic-2018	CITIBANK N.A.	AAA	BANK WATCH RATINGS S.A.
dic-2018	BANCO COMERCIAL DE MANABI S.A.	BB	PCR PACIFIC S.A.
dic-2018	BANCO COOPNACIONAL S.A.	AA	PCR PACIFIC S.A.
dic-2018	BANCO D-MIRO S.A.	A / A-	PCR PACIFIC S.A./ CLASS INTERNATIONAL RATING
dic-2018	BANCO DELBANK S.A.	A-	SUMMA RATING S.A.*2
dic-2018	BANCO PARA ASISTENCIA COMUNITARIA FINCA S.A.	BBB+	CLASS INTERNATIONAL RATING
dic-2018	BANCO GENERAL RUMIÑAHUI S.A.	AAA-	PCR PACIFIC S.A.
dic-2018	BANCO DE GUAYAQUIL S.A.	AAA / AAA	SOC.CAL.RIESGO LATINOAMERICANA / PCR PACIFIC S.A.
dic-2018	BANCO INTERNACIONAL S.A.	AAA- / AAA	BANK WATCH RATINGS S.A./ CLASS INTERNATIONAL RATING
dic-2018	BANCO DEL LITORAL S.A.	A-	SOC.CAL.RIESGO LATINOAMERICANA
dic-2018	BANCO DE LOJA S.A.	AA+	PCR PACIFIC S.A.
dic-2018	BANCO DE MACHALA S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2018	BANCO DEL PACIFICO S.A.	AAA-	CLASS INTERNATIONAL RATING
dic-2018	BANCO PICHINCHA C.A.	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2018	BANCO PROCREDIT S.A.	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2018	BANCO DE LA PRODUCCIÓN S.A.PRODUBANCO	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2018	BANCO SOLIDARIO S.A.	AA+ / AA+	CLASS INTERNATIONAL RATING / BANK WATCH RATINGS S.A.

Fecha	Banco	Calificación SBS	Firma Calificadora
dic-2019	BANCO DE GUAYAQUIL S.A.	AAA / AAA	PCR PACIFIC S.A./ SOC.CAL.RIESGO LATINOAMERICANA
dic-2019	BANCO DEL PACIFICO S.A.	AAA	PCR PACIFIC S.A.
dic-2019	BANCO PICHINCHA C.A.	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2019	BANCO DE LA PRODUCCIÓN S.A.PRODUBANCO	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2019	BANCO DEL AUSTRO S.A.	AA+ / AA+	UNION RATINGS S.A./ SUMMA RATING S.A.
dic-2019	BANCO BOLIVARIANO C.A.	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2019	CITIBANK N.A.	AAA	BANK WATCH RATINGS S.A.
dic-2019	BANCO GENERAL RUMIÑAHUI S.A.	AAA-	PCR PACIFIC S.A.
dic-2019	BANCO INTERNACIONAL S.A.	AAA- / AAA	BANK WATCH RATINGS S.A./ CLASS INTERNATIONAL RATING
dic-2019	BANCO DE LOJA S.A.	AAA-	PCR PACIFIC S.A.
dic-2019	BANCO DE MACHALA S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2019	BANCO SOLIDARIO S.A.	AA+	BANK WATCH RATINGS S.A.
dic-2019	BANCO PROCREDIT S.A.	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2019	BANCO AMAZONAS S.A.	AA+	CLASS INTERNATIONAL RATING
dic-2019	BANCO COMERCIAL DE MANABI S.A.	B	SUMMA RATING S.A.
dic-2019	BANCO DEL LITORAL S.A.	A-	SOC.CAL.RIESGO LATINOAMERICANA
dic-2019	BANCO COOPNACIONAL S.A.	AA	PCR PACIFIC S.A.
dic-2019	BANCO CAPITAL S.A.	C	PCR PACIFIC S.A.
dic-2019	BANCO PARA ASISTENCIA COMUNITARIA FINCA S.A.	BBB+	PCR PACIFIC S.A.
dic-2019	BANCO DELBANK S.A.	A-	SUMMA RATING S.A.
dic-2019	BANCO D-MIRO S.A.	A / A	PCR PACIFIC S.A./ SUMMA RATING S.A.
dic-2020	BANCO DE GUAYAQUIL S.A.	AAA	PCR PACIFIC S.A.
dic-2020	BANCO DEL PACIFICO S.A.	AAA	PCR PACIFIC S.A.
dic-2020	BANCO PICHINCHA C.A.	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2020	BANCO DE LA PRODUCCIÓN S.A.PRODUBANCO	AAA / AA+	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2020	BANCO DEL AUSTRO S.A.	AA+ / AA+	UNION RATINGS S.A./ SUMMA RATING S.A.
dic-2020	BANCO BOLIVARIANO C.A.	AAA / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2020	CITIBANK N.A.	AAA	BANK WATCH RATINGS S.A.
dic-2020	BANCO GENERAL RUMIÑAHUI S.A.	AAA-	PCR PACIFIC S.A.
dic-2020	BANCO INTERNACIONAL S.A.	AAA- / AAA	BANK WATCH RATINGS S.A./ CLASS INTERNATIONAL RATING
dic-2020	BANCO DE LOJA S.A.	AAA-	PCR PACIFIC S.A.
dic-2020	BANCO DE MACHALA S.A.	AA+ / AA+	CLASS INTERNATIONAL RATING / PCR PACIFIC S.A.
dic-2020	BANCO SOLIDARIO S.A.	AA+	BANK WATCH RATINGS S.A.
dic-2020	BANCO PROCREDIT S.A.	AAA- / AAA-	PCR PACIFIC S.A./ BANK WATCH RATINGS S.A.
dic-2020	BANCO AMAZONAS S.A.	AA+	SUMMA RATING S.A.
dic-2020	BANCO COMERCIAL DE MANABI S.A.	B	SUMMA RATING S.A.
dic-2020	BANCO DEL LITORAL S.A.	A-	UNION RATINGS S.A.
dic-2020	BANCO COOPNACIONAL S.A.	AA	SUMMA RATING S.A.
dic-2020	BANCO CAPITAL S.A.	C	PCR PACIFIC S.A.
dic-2020	BANCO PARA ASISTENCIA COMUNITARIA FINCA S.A.	BBB+	PCR PACIFIC S.A.
dic-2020	BANCO DELBANK S.A.	A	SUMMA RATING S.A.
dic-2020	BANCO D-MIRO S.A.	A	SUMMA RATING S.A.

Elaborado por: El autor.

## Anexo 7 Gráfico de araña para las principales variables diferenciadas por grupo



**Elaborado por:** El autor.