



UNIVERSIDAD LA SALLE

FACULTAD DE NEGOCIOS

Con Reconocimiento de Validez Oficial de Estudios de la Secretaría de Educación Pública según acuerdo número No 954315 De fecha 30 de Octubre de 1995

TESIS

“Análisis de factibilidad de las redes neuronales artificiales en la predicción de flujos de inversionistas no residentes en bonos gubernamentales mexicanos”

**QUE PARA OBTENER EL GRADO DE
MAESTRO EN INGENIERÍA ECONÓMICA Y FINANCIERA**

PRESENTA:

Guadalupe Cruz Colula

Asesor de Tesis:

Dr. Roberto Antonio Vázquez Espinoza de los Monteros

Ciudad de México, septiembre de 2022

Ciudad de México, a 08 de septiembre de 2022

ESP. GABRIEL NUÑEZ GONZÁLEZ
DIRECTOR DE GESTIÓN ESCOLAR
UNIVERSIDAD LA SALLE
P R E S E N T E

Le informo que el (la) C.

GUADALUPE CRUZ COLULA

Pasante de la Escuela o Facultad: **FACULTAD DE NEGOCIOS**

de la **UNIVERSIDAD LA SALLE**, de la Maestría en:

INGENIERÍA ECONÓMICA Y FINANCIERA

con reconocimiento de validez oficial de estudios de la Secretaría de Educación

Pública según acuerdo número **954315** De fecha **30 de octubre de 1995**

Ha elaborado el trabajo de tesis titulado: “Análisis de factibilidad de las redes neuronales artificiales en la predicción de flujos de inversionistas no residentes en bonos gubernamentales mexicanos”.

De conformidad con la modalidad para la obtención de grado aprobada para esta Maestría de acuerdo a lo establecido en el Reglamento General de las Universidades La Salle Integrantes del Sistema Educativo de las Universidades la Salle.

Cumplió con todos los requisitos y el trabajo que fue elaborado bajo la conducción del Dr. Roberto Antonio Vázquez Espinoza de los Monteros que fungió como asesor, tiene la calidad suficiente para ser la base de sustentación de su Examen de Grado por lo que se le autoriza presentarlo.



Mtro. José Ramón Barreiro Iglesias
Director Facultad de Negocios

A mi compañero de vida, mi esposo, por su paciencia, ánimo, apoyo, contención y amor.

A mi familia y amigos por su apoyo inagotable.

A mis jefes y amigos del trabajo que me impulsan y me motivan a dar lo mejor de mí todos los días.

A mis profesores,
director de tesis y revisores de La Salle, por su guía y enseñanza.

Contenido

Introducción	3
Objetivo general	4
Objetivos particulares	5
Hipótesis.....	5
Estado del arte.....	7
1. Mercado de renta fija en México.....	10
1.1 Historia y relevancia.....	10
1.2 Instrumentos de renta fija	12
1.3 Tenedores de los bonos gubernamentales.....	14
1.4 Índices de referencia de inversión en renta fija	15
1.5 El mercado secundario de los bonos gubernamentales y las condiciones de operación	17
1.6 Variables económicas y financieras que intervienen en la dirección de los flujos de entrada o salida de los bonos gubernamentales mexicanos	19
1.7 Panorama económico post COVID en el mercado de renta fija	28
2 Redes neuronales artificiales y clasificadores lineales	30
2.1 Características	30
2.2 Entrenamiento	33
2.3 Tipos de redes neuronales	34
2.3.1 Redes neuronales <i>feedforward</i>	35
2.3.2 Redes neuronales de base radial	35
2.2.3 Máquina de soporte de vectores (<i>Support Vector Machine, SVM</i>).....	38
2.4 Clasificadores lineales y no lineales.....	42
2.4.1 Clasificación de patrones por funciones de distancia.....	42
2.4.2 Clasificador por distancia euclidiana.....	43
2.4.3 Clasificador por distancia Mahalanobis.....	44
2.5 Clasificador bayesiano	44
2.6 <i>K-means</i>	45
3 Propuesta de metodología	46
3.1 Diseño de la base de datos.....	48
3.1.1 Variables para la capa de entrada de la red neuronal	48
3.1.2 Capa de salida de la red neuronal: semáforo de flujos de entradas y salidas	51
3.2 Procesamiento y limpieza de datos	56
3.2.1 Herramientas para el procesamiento de la serie de datos.....	59
3.2.2 Normalización de datos	59
3.3 Selección de variables relevantes para la capa de entrada	60
3.4 Diseño de las arquitecturas de las redes neuronales	63
3.4.1 Diseño de red neuronal <i>feedforward</i>	63
3.4.2 Diseño de red neuronal de base radial	66
3.4.3 Diseño de red neuronal SVM.....	68

3.5 Entrenamiento	70
3.5.1 Entrenamiento de red neuronal <i>feedforward</i>	71
3.5.2 Entrenamiento de red neuronal de base radial	71
3.5.3 Entrenamiento de red neuronal <i>SVM</i>	72
3.6 Validación	72
4 Análisis de los resultados de la aplicación de redes neuronales en la predicción de flujos	73
4.1 Protocolo experimental	74
4.1.1 Capa de entrada y de salida para las redes neuronales	74
4.1.2 Clasificadores lineales	77
4.1.3 Configuración red neuronal <i>feedforward</i>	79
4.1.4 Configuración red neuronal de base radial	81
4.1.5 Configuración red neuronal <i>SVM</i>	83
4.1.6 Validación de experimentos	85
4.2 Análisis descriptivo de los resultados	89
4.3 Discusión de resultados	97
5 Conclusiones y trabajo futuro	101
Referencias	105
Índice de figuras	111
Índice de ecuaciones	113
Índice de tablas	114
Anexos	115
Anexo 1 Gráficas de los flujos acumulados predichos por las redes neuronales	115
Anexo 2 Mapas de calor de flujos predichos y reales de la red neuronal <i>SVM</i>	120

Introducción

El mercado de renta fija en México es uno de los mercados emergentes con mayor operación en el mundo (*Trade Association for the Emerging Markets*, 2021). A partir de 2010 la tenencia de bonos gubernamentales mexicanos en el portafolio de inversionistas extranjeros creció por más de un lustro. En 2017 alcanzó un porcentaje máximo de tenencia con respecto al total colocado, sin embargo, desde el 2018 se observa un decremento constante en las tenencias de este activo en manos de los extranjeros y en 2020 los flujos de salida fueron más acentuados debido al incremento en la aversión al riesgo que desencadenó la contingencia sanitaria global por COVID 19, generando un flujo histórico de salidas en bonos soberanos mexicanos. Las salidas no se detuvieron en 2021, incluso superaron las de 2020, y no parece que se detendrán en el corto plazo (Banco de México, 2021).

Al poseer un porcentaje tan importante de la tenencia de la deuda gubernamental mexicana, el inversionista extranjero se vuelve un participante muy importante en el mercado de renta fija del país y los flujos de salidas o entradas que éste genere influyen en el deterioro o mejora de las condiciones de operación de dicho mercado. El desarrollo eficiente del mercado de deuda en cualquier país, combinado con una política monetaria que cumpla con los objetivos de inflación y una política fiscal eficaz, mejoran las condiciones de demanda de deuda del gobierno (Jeanneau & Pérez, 2005, p. 97) . Por lo tanto, las tasas de financiamiento del gobierno federal pueden ser más accesibles para los diferentes proyectos de desarrollo de un país y en consecuencia para varios sectores económicos. Es conveniente resaltar la importancia de tener un indicador que pronostique la dirección de flujos futuros radica en la prevención del deterioro de las condiciones de operación del mercado de renta fija mediante acciones focalizadas que disminuyan el impacto, si éste se considera fuerte o grave. De esta forma, el mercado de renta fija en México se mantendrá estable, lo que le permitirá al gobierno federal colocar su deuda a tasas de rendimiento competitivas para cada uno de sus plazos y, por lo tanto, mantener la estabilidad de la economía mexicana.

Diversos estudios econométricos se han realizado para intentar predecir variables y encontrar la causalidad de una variable respecto a otra, por ejemplo, variables financieras como tasas de interés, tipo de cambio, inflación, precios de acciones, etcétera, cuyo análisis puede

contribuir al proceso de toma de decisiones para los inversionistas, empresas, bancos o cualquier otro actor de los mercados financieros. Los modelos lineales tradicionales han presentado algunas limitaciones debido a la estructura y a las suposiciones hechas dentro de las distribuciones de las variables, que en los problemas reales no siempre se cumplen. Por muchos años el método de regresión lineal fue una herramienta poderosa para realizar predicciones, sin embargo, algunos problemas que se presentaron al utilizar este modelo como encontrar el tamaño adecuado de la muestra y ciertas suposiciones asumidas en las distribuciones (por ejemplo, una varianza homogénea en los datos) resultaron en predicciones pobres o débiles.

Por otro lado, los algoritmos de inteligencia artificial surgieron como una alternativa de análisis ante la complejidad y diversidad de información que ha venido surgiendo en diferentes ámbitos. Estos modelos han sido implementados en la detección de valores atípicos, extracción de características, clasificación, regresión de datos y predicción de indicadores [(Lahmiri, 2016), (Goodell, et al., 2021), entre otros]. El poder de la inteligencia artificial sobre los modelos econométricos clásicos despertó interés en la investigación y aplicación en temas como detección de fraudes, prevención y gestión de riesgo financiero, análisis de sentimientos en textos financieros y predicción de variables económicas. Las redes neuronales artificiales han sido aplicadas ampliamente en investigación financiera que involucra pronósticos y predicciones de series de tiempo, ya que son capaces de desarrollar modelos no lineales sin un conocimiento previo de las relaciones entre las variables, además de poseer flexibilidad para aprender las relaciones entre los datos, aunque estas no sean lineales.

Objetivo general

Realizar la predicción de flujos de entradas y salidas de los inversionistas extranjeros en los bonos gubernamentales mexicanos a través del diseño e implementación de varias arquitecturas de redes neuronales artificiales y a partir de la identificación y evaluación de variables económicas y financieras que intervienen en dichos flujos.

Objetivos particulares

- Identificar las variables de entrada para la red neuronal; variables económicas, financieras, medidas de liquidez, profundidad, volumen, volatilidad y sensibilidad que alimenten la red neuronal.
- Obtener la serie histórica de cada una de las variables identificadas, realizando procesos de limpieza y transformación para homogeneizar la base de datos de variables de entrada y salida.
- Diseñar e implementar una o varias arquitecturas de redes neuronales artificiales que hagan la predicción más eficiente de los flujos de entrada o salida de bonos gubernamentales por parte de inversionistas extranjeros.

Hipótesis

La implementación de la arquitectura de redes neuronales artificiales es una alternativa metodológica de pronóstico y de análisis de tendencia de flujos de inversión de no residentes en los bonos gubernamentales mexicanos, con un grado de confiabilidad mayor al 60%.

En el primer capítulo de esta investigación se presenta el marco teórico que incluye los antecedentes, evolución e importancia del mercado de renta fija mexicano, los tipos de instrumento en este mercado y sus características, quiénes son los principales tenedores de la deuda gubernamental mexicana, las particularidades del mercado primario y secundario en donde se operan, las métricas que calculan las condiciones de operación de estos instrumentos de deuda y la descripción de las variables económicas y financieras que pueden intervenir en la toma de decisiones de los inversionistas extranjeros. En el capítulo dos se presentan los conceptos básicos de las redes neuronales, sus características, funcionamiento, tipos de redes neuronales y algoritmos de aprendizaje que servirán como base para la implementación de este enfoque metodológico alternativo que se propone para atender el problema. Además, este capítulo describe las características de los clasificadores lineales que se utilizaron en este trabajo de investigación para comparar la eficiencia de los modelos tradicionales contra los modelos no lineales, específicamente las redes neuronales.

En el capítulo tres se desarrolla la propuesta de la metodología de redes neuronales artificiales, el diseño de la base de datos de las variables de entrada para la red neuronal y la arquitectura e implementación de las redes neuronales que predigan de forma más eficiente la tendencia en los flujos de inversión. El cuarto capítulo detalla las configuraciones en las arquitecturas de redes neuronales, además describe y analiza los resultados obtenidos con los modelos implementados y, finalmente, el capítulo cinco expone las conclusiones respecto a la viabilidad de emplear la metodología de las redes neuronales artificiales para la predicción de los flujos de inversionistas no residentes en bonos gubernamentales mexicanos, así como las recomendaciones y comentarios para la realización de trabajos futuros con la misma línea de investigación.

Estado del arte

La importancia de estimar el comportamiento de cierto instrumento o indicador financiero en el futuro permite generar ventajas competitivas, tomar mejores decisiones o tomar precauciones para reducir riesgos. En algunos casos, la presencia de fenómenos atípicos en los mercados financieros y en la economía impiden que la predicción se ajuste al comportamiento real, por lo que siempre existirán discrepancias entre el modelo y el dato real. Para pronosticar existen diferentes técnicas y modelos, pueden ser resumidos de forma general en: cualitativos, cuantitativos y sistemas inteligentes (Valenzuela, 2017), como se muestra en la Figura 1.

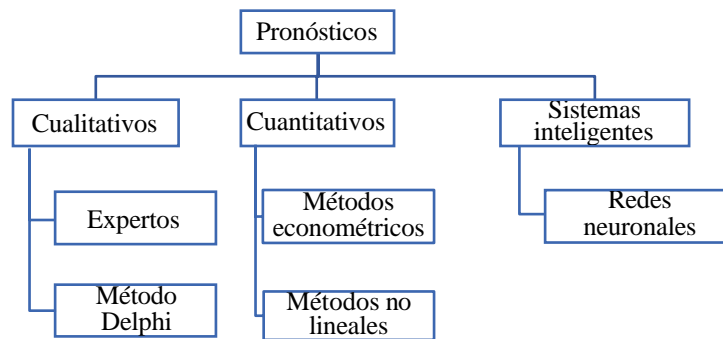


Figura 1 Métodos de pronóstico, basado en (Valenzuela, 2017), p. 5

En un modelo cualitativo un conjunto de expertos a través de su intuición, conocimiento y experiencia realiza un pronóstico. Por el lado de los métodos cuantitativos se han propuesto modelos basados en teorías económicas para predecir el comportamiento de indicadores relevantes en los mercados, sin embargo, han presentado fallas al predecir el indicador en épocas de crisis debido a las limitaciones de las estructuras lineales y a suposiciones en la distribución de las variables que en el mundo real no se observan. Hay también varios modelos no lineales que se han propuesto para mejorar el pronóstico de sistemas como el bilineal, el modelo de umbral autorregresivo, el modelo autorregresivo heterocedástico (ARCH) y el modelo autorregresivo generalizado condicional heterocedástico (GARCH), entre otros. Qiao, Yang y Li utilizaron un modelo de tipo GARCH para predecir el índice de volatilidad VIX, el cual es calculado con base en el precio de las opciones del S&P 500 (Qiao, et al., 2020), en este estudio encontraron que la predicción con su modelo debe considerarse cautelosa si el mercado cambia significativamente. En el mercado cambiario

también se han aplicado estos modelos en el pronóstico de la volatilidad del tipo de cambio, aunque se encontró que los modelos GARCH no son útiles en periodos de alta volatilidad (Pilbeam & Langeland, 2015). Los problemas de este tipo de modelos son, por un lado, que se desarrollaron para problemas específicos que son difíciles de aplicar para situaciones más generales, además de que algunos de estos modelos no alcanzan a capturar todas las relaciones no lineales entre los datos.

Otro tipo de modelo que se usa para la predicción de flujos de inversión se ubica dentro del enfoque de los sistemas inteligentes que se basa en algoritmos de inteligencia artificial, en donde el aprendizaje de máquina es una implementación específica de éstos que desarrollan técnicas que permiten reconocer patrones en conjuntos de datos. El aprendizaje profundo (*deep learning*) es un subconjunto del aprendizaje de máquina que crea técnicas para resolver problemas complejos. Dentro de los algoritmos de inteligencia artificial se encuentran las redes neuronales artificiales que surgieron como una propuesta con un enfoque alternativo y prometedor para pronosticar series de tiempo y clasificar conjuntos de datos. La idea de usar una red neuronal artificial se planteó por primera vez a mediados de los sesenta, en donde se proponía utilizar una red adaptativa lineal para predecir el clima. Más tarde, a principios se presentó una teoría matemática de una clase de máquina conocida como perceptrón, aunque con capacidades teóricas y limitaciones. Debido a las restricciones en la investigación en los algoritmos de aprendizaje para las redes neuronales de esa época, no fue hasta 1986 que se publicó el artículo que mostraba cómo usando un algoritmo de aprendizaje llamado *backpropagation* ayudaría a la red neuronal a auto ajustar sus parámetros para aprender comportamientos. Aunado a esto, la evolución de poder de cómputo ayudó al apogeo en la investigación y aplicación de las redes neuronales en diferentes ámbitos.

Las redes neuronales artificiales han demostrado aprender relaciones complejas entre series de datos de diferente índole, incluso relaciones no lineales que han sido un reto para modelos de regresión y de series de tiempo tradicionales. El uso de los algoritmos de inteligencia artificial, aprendizaje de máquina y las redes neuronales en temas financieros es un fenómeno reciente, a partir de 2016 se incrementan los artículos científicos dedicados a esta área de conocimiento. Las aplicaciones se han implementado en varios temas financieros y

económicos, (Goodell, Kumar, Lim, & Pattnaik, 2021) los clasifica de forma general en tres áreas: construcción de portafolios, valuación y comportamiento del inversionista; fraudes y crisis financieros e inferencia de sentimientos, pronósticos y planeación

Una de las arquitecturas más comunes en las redes neuronales son las *feedforward*, Lahmiri la implementa en la predicción de la tasa de interés a un día, mezclando un conjunto de técnicas para optimizar los pesos iniciales en las neuronas y comparando los resultados con un modelo autorregresivo de promedio móvil (ARMA), mostrando menores errores de predicción en la red neuronal con respecto al modelo convencional (Lahmiri, 2016). El tipo de red neuronal conocida como soporte de vectores (*Support Vector Machine SVM*) ha sido utilizado en varias investigaciones del ámbito financiero, Zhang, Shi y otros autores lo aplicaron en la predicción de riesgo de crédito usando como función de optimización el algoritmo de simulación de comportamiento de las luciérnagas (Zhang, et al., 2019), encontrando una mejora en la precisión de la predicción para indicar si una empresa es candidata para un financiamiento de acuerdo al nivel de su riesgo de crédito. Cao y Tay examinaron en su experimento contratos de futuros en el índice S&P 500, la predicción realizada por la SVM resultó ser más eficiente que un algoritmo de tipo *back progression* (Cao & Tay, 2000).

Los trabajos de implementación de redes neuronales en el campo financiero abundan para el mercado accionario y el cambiario, sin embargo, para el mercado de renta fija hay muy pocos y la mayoría se enfoca en la predicción de tasas de interés, dado el rol que juega este indicador en la economía y en las finanzas. El presente trabajo de investigación pretende explorar la aplicación de las redes neuronales artificiales en el mercado de renta fija y en uno de los instrumentos más importantes en este mercado, los bonos gubernamentales a tasa fija a través de la implementación de varias arquitecturas que ya han sido probadas en temas financieros como lo son las redes neuronales de tipo *feedforward* y las de tipo SVM, cuya eficiencia ha sido probada en varias investigaciones, sus características serán explicadas en capítulos posteriores.

1. Mercado de renta fija en México

En esta sección se revisarán los conceptos básicos y el contexto del mercado de renta fija en México, se conocerá la relevancia y el impacto que éste tiene en la economía del país. Además, se describirá el detalle de los instrumentos que se negocian en este mercado, los tipos de inversionistas que participan en él y los canales de operación de los títulos de deuda soberanos. Por otro lado, es necesario estudiar no sólo el impacto local de su operación sino los beneficios que involucra un buen funcionamiento y que conllevan a que los títulos de deuda sean incluidos en índices de referencia de inversión globales, seguidos por millones de inversionistas. Además, se introduce al lector a los conceptos de las variables económicas, financieras y de mercado que se utilizarán en la propuesta de metodología de esta investigación, descrita en el capítulo 3. Finalmente se describe el entorno económico iniciado en 2020 por la crisis sanitaria por COVID-19 como un factor relevante para explicar la evolución de los flujos en los bonos gubernamentales mexicanos.

1.1 Historia y relevancia

El mercado de renta fija es el mercado más antiguo en la historia de la humanidad, se define como el lugar físico o electrónico en donde se encuentran inversionistas con exceso de liquidez, en búsqueda de rendimiento y emisores de deuda con necesidades de financiamiento. En este tipo de mercado se cuenta con un documento o prospecto de inversión que indica principalmente cuánto se compromete a pagar el emisor, la forma de generar rendimientos (a descuento o mediante el pago de intereses en varios periodos de tiempo, también conocidos como cupones) y el plazo de inversión. El emisor más importante que participa en este mercado es el gobierno federal, además de otras instituciones como los bancos comerciales, corporativos, etcétera. Los gobiernos emiten deuda para obtener recursos para el financiamiento de proyectos, además, estos títulos se consideran como referencia o base de los mercados de instrumentos de renta fija en un país y fungen como activos subyacentes en varios productos derivados. De acuerdo con datos de *Bloomberg* a agosto de 2020 el total en circulación del mercado de deuda global rebasa los 128 trillones de dólares, más de la mitad la emiten gobiernos soberanos.

En México el mercado de valores gubernamentales surgió en 1978, los inversionistas únicamente estaban dispuestos a prestar a plazos cortos debido a la poca confianza en el emisor soberano. Fue hasta inicios de los noventa cuando se emitió el primer instrumento con plazo de un año. En esa misma década surgió una de las crisis financieras más severas de la historia que ha experimentado el país, conocida mundialmente como la “crisis del tequila”, la cual tuvo como origen un conflicto en la balanza de pagos provocada por un déficit público, que obligó al gobierno a emitir más deuda en moneda extranjera, principalmente en dólares y a corto plazo. A principios de 1994 esta deuda representaba la mayor parte del financiamiento del gobierno federal. La confianza de los inversionistas de mercados internacionales se deterioró y se generó la incertidumbre de una posible devaluación e incumplimiento de las obligaciones de deuda, por lo que los tenedores de ésta optaron por deshacerse de los instrumentos mexicanos. Las liquidaciones provocaron que las reservas internacionales del país se agotaran en su totalidad, el peso se depreció y a su vez se incrementó el nivel de la deuda mexicana emitida en dólares.

A raíz de aquella crisis financiera, se llevaron a cabo una serie de reformas y cambios en las políticas de administración de deuda enfocadas a fortalecer la situación macroeconómica del país. Una de las medidas fue sustituir la deuda en moneda extranjera por deuda denominada en moneda local y con tasa de interés fija. En el terreno fiscal se reformó la Ley del Seguro Social, se crearon cuentas individuales para los fondos de retiro de los trabajadores con aportaciones definidas, manejadas por Administradoras de Fondos para el Retiro (Afores), con lo que se eliminó una fuente importante de gasto corriente y por lo tanto esto ayudó a sanear las finanzas públicas. La política cambiaria mexicana también se modificó y adoptó un régimen cambiario de libre flotación, donde el comportamiento de la divisa lo determina el mercado. En el 2000 se creó la figura de Formador de Mercado de valores gubernamentales que ayudó a aumentar la liquidez y el volumen de transacciones de los títulos en el mercado nacional.

Como resultado de las medidas implementadas, paulatinamente los plazos de colocación de la deuda soberana se incrementaron con vencimientos de 3 y 5 años y finalmente en el 2006 se emitió el plazo más largo de la curva de vencimientos de los bonos emitidos por el gobierno federal: bonos de 30 años. La implementación de las medidas mencionadas por

parte de las autoridades financieras en el país y la ejecución de reformas estructurales ayudó al desarrollo del mercado de bonos. México se volvió atractivo para la inversión extranjera.

En la década posterior a la gran crisis financiera mundial de 2008 se observaron cambios significativos de volumen y de composición de los flujos globales de capital, con un incremento de la participación extranjera particularmente en las economías emergentes (García & Stracca, 2021). La incertidumbre global post crisis, con un escenario de debilidad económica y políticas monetarias acomodaticias de los principales bancos centrales en el mundo propiciaron entre los inversionistas una búsqueda de rendimientos atractivos, lo cual favoreció a varias economías emergentes, especialmente a México. La inversión de no residentes puede actuar como un catalizador en el desarrollo de los mercados de bonos locales porque ayuda a la diversificación de la base de inversionistas institucionales, genera mayor demanda por los instrumentos de deuda en un país, lo cual incrementa la liquidez de los activos, por lo tanto, genera tasas de interés bajas que a su vez propicia mejores condiciones en el crédito y por consiguiente más proyectos de inversión en un país. En un efecto contrario, la inversión de extranjeros en el mercado local de bonos también resulta en mayor volatilidad en dicho mercado cuando existe un incremento en la aversión al riesgo ante situaciones globales (Peiris, 2010), ya que la salida de capitales extranjeros puede generar incluso graves crisis económicas en un país.

1.2 Instrumentos de renta fija

El gobierno federal emite cuatro tipos diferentes de instrumentos de deuda en el mercado local. Estos cuentan con la descripción de sus características y cálculo de rendimientos en los documentos publicados en la página del Banco Central de México (Banco de México & Secretaría de Hacienda Crédito Público, 2021):

- **Cetes:** bonos cupón cero (no pagan intereses en el transcurso de su vida y liquidan su valor nominal en la fecha de vencimiento), (Banco de México, s.f.). Estos instrumentos son regularmente emisiones de corto plazo (menores a un año). El principal tenedor de estos títulos son los bancos.

- **Bon-des:** son instrumentos a tasa revisable que pagan intereses en plazos. A partir de octubre de 2021 se colocan los llamados Bon-des F que pagan intereses en pesos cada mes. La tasa de interés se calcula componiendo diariamente la tasa de interés interbancaria de equilibrio de fondeo a un día publicada por el Banco Central (Banco de México & Secretaría de Hacienda Crédito Público, 2021). Los inversionistas extranjeros no tienen mucho interés en este tipo de instrumento, siendo las sociedades de inversión y la banca de desarrollo los principales tenedores.
- **Udibonos:** son instrumentos que protegen de la inflación a su tenedor. Pagan intereses cada seis meses en función a una tasa de interés real fija. Devengan intereses en UDIS (unidades de inversión, que se actualizan de acuerdo con el Índice Nacional de Precios al Consumidor) que son pagaderos en pesos (Banco de México, 1996). Estos títulos son considerados de largo plazo ya que la curva de vencimientos va de los 3 a los 30 años. Las Afores son los más importantes tenedores ya que cubren muy bien sus necesidades de inversión de largo plazo.
- **Bonos:** los Bonos de Desarrollo del Gobierno Federal con tasa de interés fija pagan intereses cada seis meses a una tasa de interés que se mantiene fija durante toda la vida del instrumento y que se determina desde su emisión, actualmente son colocados a plazos de 3, 5, 10, 20 y 30 años. Su valor nominal es de 100 pesos. Las operaciones de compraventa de estos títulos se pueden realizar cotizando su precio o su tasa de rendimiento al vencimiento, la convención actual del mercado es cotizarlos a través de su rendimiento al vencimiento. El precio de los bonos está compuesto por el valor presente de los cupones, el valor presente del principal y los intereses devengados del cupón vigente (Banco de México, s.f.)

De acuerdo con datos publicados por el Banco Central a noviembre de 2021 (Banco de México, 2021), los bonos representan el 38% de monto total en circulación de la deuda gubernamental mexicana, con poco más de tres billones de pesos emitidos y colocados. Los inversionistas extranjeros son los principales tenedores de este título, seguido por el sector bancario y las Siefomes (Sociedades de Inversión Especializadas

en Fondos para el Retiro). La curva de vencimientos de los bonos gubernamentales a tasa fija, a enero de 2022, incluye 17 nodos de vencimientos distribuidos en plazos entre 1 y 30 años, lo cual la hace atractiva para distintos perfiles de inversión. El Gobierno Federal, a través de la Secretaría de Hacienda y Crédito Público (SHCP) se encarga de informar a los inversionistas las fechas de colocación trimestral para cada plazo de la curva. Esto brinda información relevante a los inversionistas y proporciona transparencia al mercado.

De toda la curva de vencimientos de bonos, el mercado presta especial atención en los nodos que son llamados “de referencia” que incluyen los vencimientos con plazo de 1, 3, 5, 10 y 30 años. Estos títulos son utilizados para comparar su rendimiento con bonos de otros países, incluyendo los de Estados Unidos, y, por lo tanto, el comportamiento del volumen, profundidad, liquidez y volatilidad es diferente al del resto de los bonos. En distintas literaturas, es común que este tipo de nodos son conocidos como títulos *on the run*.

El Banco Central de México es el agente financiero del gobierno federal, se encarga de colocar la deuda gubernamental en el mercado primario a través de mecanismos de colocación conocidos como subastas: ventas públicas al mejor postor. El mercado primario es donde se emiten los valores por primera vez ya sea por medio de subastas primarias semanales o usando la colocación por sindicación, en donde los valores se venden a un grupo específico de instituciones financieras que se encarga de distribuirlos en el mercado. Una vez que los bonos se colocan, las instituciones que recibieron asignación en la subasta pueden efectuar operaciones de compraventa, préstamo o reporto con otros inversionistas, a esto se le conoce como mercado secundario.

1.3 Tenedores de los bonos gubernamentales

La presencia de inversionistas nacionales y extranjeros asegura una demanda constante de los valores gubernamentales ayudando con esto a disminuir el costo de la deuda pública, se promueve la innovación financiera y el desarrollo de los mercados financieros (Abreu et al., 2014). Contar con una curva de rendimientos tan amplia como la de los instrumentos gubernamentales, le permite a las diferentes necesidades y horizontes de inversión de los participantes posicionarse en el segmento de la curva que las satisfaga. Así, por ejemplo, las

Siefores que requieren títulos de largo plazo a tasa nominal o real, pueden contar en su portafolio con bonos o Udibonos en plazos entre 10, 20 o 30 años, mientras que las sociedades de inversión que necesitan de mayor liquidez para cubrir las solicitudes de sus clientes pueden posicionarse en los cetes o bonos de menor plazo.

Contar con inversionistas extranjeros como tenedores de la deuda soberana ayuda a profundizar y dar liquidez a los instrumentos, lo cual mejora el desarrollo del mercado. Por otra parte, los flujos de entrada de inversionistas extranjeros a países emergentes hacen más interdependiente y susceptible a la economía de estos países a la volatilidad de los mercados financieros internacionales, como sucede con la crisis sanitaria por COVID 19, donde la aversión a activos riesgosos aumentó a niveles históricos y los inversionistas buscaron salirse abruptamente de este tipo de instrumentos para mover su portafolio hacia activos de menor riesgo.

1.4 Índices de referencia de inversión en renta fija

Varias instituciones financieras internacionales se encargan de publicar índices de inversión referenciados a instrumentos de renta fija o renta variable que los inversionistas pueden replicar. Cuando la deuda de un país es incluida a estos índices, los inversionistas que los replican incluyen en su portafolio los títulos de deuda de dicho país. Cada proveedor de los índices establece los criterios de selección para que un país entre o salga de su índice, esos criterios generalmente incluyen el tamaño del mercado, calificación crediticia del país y que el acceso a los mercados sea libre y sin políticas de restricción la participación de inversionistas extranjeros (Banco de México , 2021).

El *World Government Bond Index* (WGBI) es un índice de referencia para el mercado de deuda gubernamental con grado de inversión y emitidos en moneda local a tasa fija, es administrado por *Financial Times Stock Exchange* (FTSE) Russell. Los países miembros son Australia, Austria, Bélgica, Canadá, Dinamarca, Finlandia, Francia, Alemania, Irlanda, Israel, Italia, Japón, Malasia, México, Holanda, Noruega, Polonia, Singapur, Suecia, España, Reino Unido, Estados Unidos y China (incluido en 2021). Cada país tiene asignado un peso en el índice, el cual es re balanceado a final de mes (Banco de México , 2021).

Los bonos gubernamentales mexicanos están incluidos en varios de los índices internacionales como el *Government Bond Index Broad-JP Morgan*, *Global Emerging Markets-Markit Group* y el *Global Government Bond Index-Merrill Lynch* entre otros. A inicios de 2010 se anunció la entrada de los bonos mexicanos denominados en pesos al índice de referencia de inversión WGBI. Los bonos mexicanos fueron los primeros en Latinoamérica en formar parte de este índice. Para entrar a este tipo de referencias los activos de renta fija deben cumplir con un conjunto de requisitos, como ya se mencionó anteriormente, una vez que los bonos soberanos los cumplen, son incluidos al índice y se les asigna un peso en el portafolio de inversión. Como resultado los inversionistas que siguen el índice compran dichos bonos en el mercado del país en cuestión. A partir de la inclusión en el WGBI, el apetito de los inversionistas extranjeros por los bonos de México creció de manera importante, llegando a concentrar en su portafolio el 66% respecto al total colocado (Figura 2), de acuerdo con los datos publicados por Banco de México (marzo de 2021), convirtiéndose en el participante más importante en este instrumento.

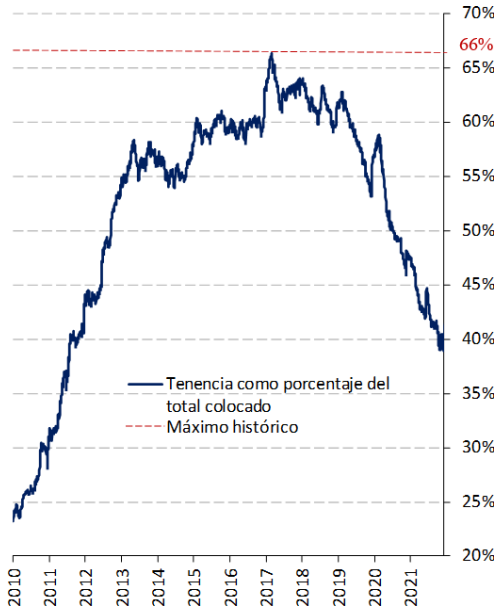


Figura 2 Porcentaje de tenencia de bonos gubernamentales con respecto al total colocado por parte de inversionistas extranjeros. Fuente: Banco de México.

1.5 El mercado secundario de los bonos gubernamentales y las condiciones de operación

El desarrollo de un mercado secundario eficiente en un país es fundamental para darle sustento al mercado de renta fija. Para lograr esta eficiencia se debe contar con: una estructura de mercado competitiva, bajos costos de transacción, bajos niveles de fragmentación, una infraestructura segura y robusta y un alto nivel de heterogeneidad entre participantes (Banco Mundial y Fondo Monetario Internacional, 2001). De acuerdo con datos que calcula Banco de México, en junio de 2021 (Banco de México, 2021) se operó un promedio diario en el mercado secundario local de 137,857 millones de pesos en donde los bonos representan un 56% de la operación total. Los principales actores que participan en el mercado secundario son:

1.5.1 Proveedores de precios

Son instituciones que se encargan de calcular el precio diario actualizado de valuación de cada uno de los títulos en circulación en el mercado. En México hay dos entidades que fungen este papel: Proveedor Integral de Precios (PIP) y Valuación Operativa y Referencia de Mercado (Valmer) y son regulados por la Comisión Nacional Bancaria y de Valores. Su importancia radica en la credibilidad que tienen sus clientes en el cálculo del precio diario de los títulos de deuda, de tal forma que los inversionistas pueden medir el valor de mercado de su portafolio.

1.5.2 Casas de corretaje o “brokers”

Son empresas que ofrecen servicios de intermediación de operaciones en los mercados financieros a través de plataformas electrónicas y a través del teléfono, sus ingresos son generados a través del cobro de comisiones (Abreu et al., 2014). Los inversionistas pueden ingresar a estas plataformas para ofrecer o pujar por un precio para la compra o venta de un título, si otro participante está interesado, la transacción se pacta. Regularmente las operaciones se liquidan con un promedio de dos días hábiles después de que fueron pactadas, pero también pueden ser liquidadas el mismo día o con varios días después de que se concertaron.

1.5.3 Métricas de las condiciones de operación del mercado secundario

Si se necesita medir la “salud” con algún tipo de termómetro, las condiciones de operación del mercado de renta fija en un país son un conjunto de indicadores que nos revelan qué tan adecuado o deteriorado se encuentra dicho mercado. Es en estas condiciones de operación en donde se ve reflejado el efecto de los flujos de entrada y salida de inversionistas extranjeros. A continuación, se describe con mayor detalle los indicadores que miden este comportamiento en el mercado secundario de bonos en México.

1.5.3.1 Liquidez

Mide la habilidad de comprar o vender cierta cantidad de un instrumento, a un precio deseado y en un tiempo determinado sin impactar el precio de dicho instrumento (International Organization of Securities Commissions, 2017). Generalmente esta medida se asocia a bajos costos de transacción y la inmediatez de ejecución de la orden. Existen varias maneras de medir la liquidez de un instrumento, una de las más utilizadas es el diferencial entre el precio de compra y el precio de venta. El diferencial promedio de compraventa de un bono gubernamental se calcula como el promedio de las diferencias de las posturas de venta y las posturas de compra, de todas las transacciones realizadas durante un día hábil de operación en las plataformas de los *brokers*.

1.5.3.2 Profundidad

Esta medida está muy ligada a la liquidez. Si un inversionista necesita vender su posición de algún instrumento y además esta posición es de gran tamaño, el precio promedio al que se ejecutará dicha orden depende de la elasticidad del mercado para ejecutarla y en qué medida se mueven los precios o tasas de interés con esta operación. El índice de Amihud (Amihud & Mendelson, 1991) mide el cambio en las tasas por cada millón de pesos operado.

1.5.3.3 Volumen de operación

Se define como el volumen promedio diario operado en las plataformas electrónicas de las casas de corretaje y en los sistemas de negociación con la clientela para cada uno de los instrumentos de deuda. En México estas operaciones son reportadas al Banco de México por Enlace Int, Euro Brokers México, Remate Lince, SIF-ICAP, Tradition, GFI Group México,

Bull&Bear, Mercado Electrónico Institucional (MEI), Sistemas Var y Sistema de Información de Posturas (SIPO).

1.5.3.4 Volatilidad en los precios

La volatilidad indica qué tanto varían los precios de un bono gubernamental en un periodo de tiempo. Se calcula con los cambios logarítmicos de las posturas concretadas (hechos) de operaciones realizadas con bonos de forma diaria en el mercado secundario. Se asume que X_t es el cambio logarítmico del hecho a tiempo t y \bar{X} es la media de los cambios logarítmicos de los hechos.

1.6 Variables económicas y financieras que intervienen en la dirección de los flujos de entrada o salida de los bonos gubernamentales mexicanos

Desde los noventa (Calvo, et al., 1993) se ha pretendido explicar los flujos de entrada o salida de inversionistas a ciertas regiones como resultado de las condiciones económicas y financieras externas o globales y la importancia de que los cambios negativos en éstas pueden conducir a futuras salidas de capitales. El estado de las condiciones globales de la economía aunado a los factores locales de cada región contribuye al direccionamiento de los capitales hacia economías desarrolladas y emergentes. Los determinantes o factores responsables de los flujos de capitales dependen del tipo de flujo de capital (acciones, bonos, etc.) y de un conjunto de características que se han agrupado en: factores globales o sistemáticos (*push factors*) y factores locales o idiosincráticos (*pull factors*). De forma más reciente se han identificado también los factores de infraestructura (*pipes*).

Con base a los factores mencionados anteriormente, se seleccionaron un conjunto de variables financieras, económicas y de flujos que de acuerdo con varias investigaciones realizadas son relevantes en el comportamiento de los flujos de capital; trabajos como los de Eguren que examina cómo el efecto de los cambios en los factores locales es más persistente que el de los factores globales en economías emergentes (Eguren-Martin, et al., 2020). Por otro lado, hay trabajos que examinan los flujos de capitales en riesgo de países emergentes y la dependencia con los cambios en las condiciones financieras de economías avanzadas y la política monetaria de Estados Unidos (Norimasa, et al., 2021, p. 2). Las tasas de interés más altas en países desarrollados, principalmente Estados Unidos, aunado con un dólar

estadounidense fuerte, se asocian con flujos futuros más débiles hacia economías emergentes en el corto y mediano plazo.

1.6.1 Variables *push factors*

Los factores globales o de empuje (*push factors*) dependen del apetito por riesgo global y las condiciones financieras. Dentro de estos factores se encuentran las perspectivas de política monetaria y de estabilidad financiera de Estados Unidos, el crecimiento esperado y el nivel de tasas relativo ajustado por riesgo (*carry to risk*).

- ***Carry to risk***

Se define como el diferencial de tasas (tasa local y tasa foránea) ajustado por riesgo. La forma de calcularlo es a través de la tasa implícita en el forward cambiario dólar estadounidense- peso mexicano con plazo de 3 meses (debido a que este plazo es el más operado y líquido en el mercado) dividido por la volatilidad implícita o esperada con plazo de 3 meses.

$$\frac{\ln \left(\frac{MXNUSD + Fwd3m}{\frac{1000}{MXNUSD}} \right) * 360/90}{VolUSDMXN3m/100}$$

Ecuación 1 Cálculo carry to risk

- ***Carry to risk relativo a otras economías emergentes***

Para esta métrica se calcula el *carry to risk* para un conjunto de países emergentes, utilizando la forma de cálculo de la Ecuación 1 pero sustituyendo el tipo de cambio con el correspondiente al país seleccionado. Una vez obtenido el indicador para cada país, se obtiene un promedio y finalmente se divide el *carry to risk* de México entre dicho promedio. La canasta de países emergentes considerados son Brasil, Chile, Colombia, Turquía, Sudáfrica, Corea del Sur, Polonia y Rusia.

- **Índice de aversión al riesgo global**

Este índice calculado por Citigroup mide la aversión al riesgo de todo tipo de activos, se pondera de igual forma con los diferenciales soberanos entre mercados desarrollados y emergentes y los bonos de Estados Unidos. También considera la volatilidad implícita en divisas, acciones y *swaps*. El índice es medido por el número de desviaciones estándar con la media de la distribución.

- **Índice de aversión al riesgo de economías emergentes**

Este indicador también es calculado por Citigroup, sin embargo, solo considera activos de países emergentes. Generalmente en épocas de crisis económicas mundiales, se encuentra en niveles elevados, lo cual significa que los inversionistas prefieren activos denominados de refugio en lugar de activos de países emergentes que son considerados de riesgo.

- **Índice de condiciones financieras de Estados Unidos**

El estado actual de aquellas variables financieras que influyen en el comportamiento de empresas, consumidores e inversionistas y por lo tanto afectan el estado futuro de la actividad económica se toma en cuenta para condensar dicha información en un indicador puntual. Este permite capturar diferentes episodios de estrés financiero, para el caso de Estados Unidos, es relevante conocer dichas condiciones para la toma de decisiones de inversión. El índice tomado en cuenta en este trabajo de investigación es el que calcula Bloomberg a través de una suma ponderada de indicadores del mercado de dinero, renta variable, tasas de hipotecas, tipo de cambio efectivo y precios de energía en Estados Unidos.

- **Tasas de rendimiento de los bonos del tesoro (*treasuries*) de 3, 5,10 y 30 años**

El nivel de las tasas de rendimiento para los nodos de referencia de los bonos del tesoro de Estados Unidos es un elemento importante que influencia de manera favorable a los flujos de los bonos de economías emergentes cuando se encuentra en niveles bajos, ya que genera un comportamiento de búsqueda por rendimiento. También incorporan las expectativas de política monetaria de la Reserva Federal, por lo que, en periodos de alza de tasas, provocan flujos menos acelerados de entradas en países emergentes, como México.

- **Producto Interno Bruto (PIB) de Estados Unidos**

El PIB es un indicador económico que mide la riqueza que genera un país a través del cálculo de la producción de bienes y servicios. En el caso del PIB de Estados Unidos, está estrechamente ligado con el PIB de nuestro país.

- **DXY**

El índice dólar estadounidense (DXY por sus siglas en inglés) mide el valor del dólar estadounidense en relación con una canasta de otras divisas. Cuando este índice sube, significa que el dólar está ganando fuerza en comparación con las otras divisas de la canasta, en la mayoría del tiempo cuando esto sucede, se observan depreciaciones en monedas de países emergentes.

- **Índice de materias primas; subíndice industrial de metales, subíndice de agricultura, subíndice del petróleo WTI y subíndice de energía**

Se incluyen índices de las principales materias primas debido al impacto que pueden tener las fluctuaciones de las cotizaciones de estas sobre los precios internos al consumidor local y a un posible impacto sobre el proceso de formación de precios del resto de los bienes y servicios del país y que por lo tanto pueden generar presiones a la inflación.

- **Índice de sorpresas inflacionarias para economías emergentes**

Una variable importante en contextos como el que se vive actualmente en el mundo post COVID y donde la inflación general ha rebasado los rangos objetivos de varias economías en el mundo, revela información relevante respecto a los niveles esperados con respecto a los que realmente sucedieron. Este índice es calculado por Citigroup y se define como la desviación estándar de la diferencia entre las cifras publicadas y la mediana de los estimados correspondientes. Un número positivo en este índice significa que la inflación fue más alta respecto a lo esperado, mientras que un número negativo indica que fue más baja.

1.6.2 Variables *pull factors*

Por su lado, los factores locales o de atracción (*pull factors*) se miden a través de las condiciones nacionales que afectan qué tan atractivo es un país para la inversión extranjera. Las variables que miden estos factores son el régimen de inflación creíble, el régimen cambiario flexible, la sostenibilidad fiscal, las herramientas macro prudenciales eficaces y la profundidad en el mercado de capital interno.

- **Tasas de rendimiento de bonos mexicanos *on the run* de 3, 5, 10, 20 y 30 años**

Los niveles de las tasas de rendimientos de los bonos indican, por una parte, qué tan atractivo puede ser para un inversionista tener este instrumento en su portafolio y por otra, las expectativas del país. Se incluyen las valuaciones realizadas por PIP para los nodos de referencia (*on the run*).

- **Pendiente de la curva de rendimientos de bonos gubernamentales mexicanos medido por el diferencial 2-10 años y 3-30 años**

El diferencial entre los bonos gubernamentales de corto plazo (2 o 3 años) y los bonos de largo plazo (10 o 30 años) normalmente se comporta con forma de pendiente positiva ascendente, es decir, las tasas de rendimiento de corto plazo con más bajas que las de largo plazo. Cuando los inversionistas muestran desconfianza sobre el futuro económico, la curva de rendimientos suele tener una pendiente descendente, los instrumentos de corto plazo poseen mayores tasas de rendimiento que las de largo plazo. Conocer la pendiente de la curva de rendimientos es importante para la gestión activa en los instrumentos de renta fija a través de diferentes estrategias basadas en la curva de tipos de interés.

- ***Breakeven* de 3, 10 y 30 años**

Permite medir las expectativas de inflación en el mercado de bonos a través de la diferencia de las tasas nominales y las tasas reales para cada plazo mencionado.

- **Reservas internacionales más línea de crédito flexible**

Son activos financieros que el banco central invierte en el exterior, son líquidos y sirven para contribuir a la estabilidad de precios cuando se presenta una disminución de los flujos comerciales o de capital de la balanza de pagos, debido a desequilibrios macroeconómicos y/o financieros (internos o externos) (Banco de México, s.f.). En este sentido el nivel de reservas internacionales que posee un país indica la capacidad para adoptar determinadas políticas u otras, respalda la confianza en la divisa nacional, garantizando un respaldo monetario. Por otro lado, la línea de crédito flexible (LCF) complementa las reservas internacionales para enfrentar condiciones externas adversas. El Fondo Monetario Internacional es quien se encarga de facilitar este instrumento a los países que cumplen con ciertos requisitos. México la obtuvo por primera vez en 2009.

- **Deuda de México como porcentaje del PIB**

La deuda pública que tiene el gobierno con los ciudadanos (deuda interna) y con otros países (deuda externa) y la sostenibilidad de esta deuda es una de las condiciones necesarias para que un país pueda potenciar su crecimiento económico, reforzar las bases de un mayor nivel de desarrollo en el largo plazo y por lo tanto generar expectativas de crecimiento para las inversiones externas. Un nivel alto de deuda puede provocar un desbalance en las finanzas públicas de un país que obligaría a implementar medidas para pagar. Cuando los prestamistas confían en que los gobiernos son fiscalmente responsables, es más sencillo financiar déficits mayores y reestructurar la deuda. Este indicador se pondera como porcentaje del PIB del país para los cálculos subsecuentes.

- **Volatilidad implícita en opciones cambiarias dólar norteamericano – peso mexicano con plazo de un mes**

Indica la expectativa del mercado para la volatilidad a un plazo determinado, para este cálculo se utiliza el precio al que cotizan las opciones en el mercado. Cuando la volatilidad implícita aumenta, los mercados de opciones experimentan una tendencia a la baja, una volatilidad más alta indica una expectativa de mayor movimiento del precio de la opción en el futuro. El

dato es tomado directamente de la plataforma de *Bloomberg* para opciones pactadas a un mes entre el dólar estadounidense y el peso mexicano.

- **Sesgo implícito en opciones cambiarias dólar norteamericano – peso mexicano con plazo de tres meses**

La forma resultante de la distribución de opciones pactadas con la misma fecha de vencimiento en el mercado a menudo muestra un sesgo o “*smile*” que indica que los inversionistas prefieren tener opciones dentro o fuera del dinero que opciones dentro del dinero.

- **Curtosis implícita en opciones cambiarias dólar norteamericano – peso mexicano con plazo de tres meses**

El valor teórico de las opciones fuera del dinero se explica en gran parte por el efecto de los momentos superiores, en particular la curtosis. El dato se obtiene directamente de Bloomberg con opciones de plazo de tres meses dada la liquidez observada en éstas.

- **Tipo de cambio *spot* del peso mexicano – dólar estadounidense**

El tipo de cambio es una variable presente en las inversiones de participantes extranjeros, ya que regularmente cambian sus dólares a pesos mexicanos para poder comprar bonos gubernamentales y obtener ganancias de las estrategias que implementen, por lo que las fluctuaciones en éste afectaran o beneficiaran dichas ganancias. El dato se toma del tipo de cambio *spot* entre el peso mexicano y el dólar norteamericano proporcionado por Bloomberg.

- **Producto Interno Bruto (PIB) de México**

Como ya se explicó anteriormente, el PIB indica la riqueza que genera un país, por lo que el nivel del PIB a lo largo del tiempo de México es de interés para los inversionistas extranjeros para proyectar las expectativas de crecimiento hacia adelante.

- **Posición en *forwards* y *swaps* cambiarios con el tipo de cambio peso mexicano - dólar estadounidense pactados en México**

Se refiere a la posición pactada por inversionistas extranjeros en *forwards* y *swaps* cambiarios con el tipo de cambio peso mexicano y dólar norteamericano que fueron operados en territorio nacional.

- **Razón de cobertura cambiaria para instrumentos de renta fija (RCC)**

Este índice se calcula por *State Street Global Markets*, de la siguiente manera:

$$RCC = \frac{\text{Posición neta MXN/USD}}{\text{Tenencia de instrumentos que no son cambiarios}}$$

Ecuación 2 Cálculo de razón de cobertura de State Street

Cuando la posición neta es menor que cero significa que es largo dólar, corto peso. Cuando la posición neta es mayor que cero indica que es largo peso, corto dólar. Un aumento en el índice revela una posición cubierta por parte de los inversionistas, un nivel de 0% indica una posición no cubierta y finalmente niveles por debajo de cero indican una posición larga en el peso mexicano.

- **Calificación promedio de la deuda mexicana**

Existen agencias de calificadoras de riesgo que se dedican a evaluar, analizar y determinar el riesgo crediticio de una entidad o gobierno. La calificación otorgada indica la probabilidad de impago, cada agencia cuenta con escalas distintas al asignar la calificación de la deuda emitida por la entidad. Los instrumentos de deuda se pueden clasificar en tres niveles principales:

- Grado de inversión: significa que la entidad cuenta con capacidad para cumplir las obligaciones de pago
- Especulativo: existe incertidumbre en la capacidad de cumplimiento de pago

- Alto riesgo crediticio: la entidad se clasifica como de alto riesgo, con una alta probabilidad de que ocurra un impago en sus obligaciones.

La calificación otorgada a la deuda de un emisor con nivel de grado de inversión es requisito para algunos índices de referencia (como el WGBI), por lo que un descenso del nivel de calificación de los bonos mexicanos podría provocar flujos de salidas en estos instrumentos. Dado que se cuenta con tres agencias calificadoras; *Fitch Ratings*, *Moody's* y *Standard and Poor's* y con diferentes escalas de calificación entre ellas, se realiza una ponderación de las calificaciones históricas para la deuda soberana mexicana, para finalmente realizar un promedio de las tres calificaciones.

- **Indicador Global de la Actividad Económica**

Es una métrica calculada por el Instituto Nacional de Estadística y Geografía (INEGI) que permite conocer y dar seguimiento a la evolución del sector real de la economía mexicana, en el corto plazo.

1.6.3 Otras variables consideradas (flujos hacia economías emergentes)

- **Flujo en activos de renta fija y renta variable dirigidos hacia China**

Como se ahonda más en el siguiente subapartado de este capítulo, China ha atraído grandes volúmenes de flujos en renta fija y renta variable dada su inclusión en varios índices de referencia, por lo que la evolución de estos juega un papel relevante en los flujos de salida de bonos gubernamentales mexicanos observados durante 2020.

- **Índice de bonos gubernamentales en países emergentes (EMBI) JP Morgan**

Es un índice de referencia que calcula JP Morgan y que contiene títulos de deuda de mercados emergentes. Este índice goza de gran popularidad entre los inversionistas del mundo y sirve como referencia para estimar la evolución del mercado de deuda emergente, además muestra el rendimiento de estos bonos y su diferencial con respecto a una canasta de bonos denominados de libre de riesgo.

- **Índice de activos de renta variable MSCI**

Este índice de referencia selecciona un conjunto de acciones de renta variable que es representativo siguiendo una metodología definida.

1.7 Panorama económico post COVID en el mercado de renta fija

A inicios de 2021 el desarrollo de vacunas contra COVID 19, aperturas de varias economías, políticas monetarias acomodaticias en los bancos centrales y estímulos históricos en algunos países, sobre todo en Estados Unidos, generaron altas expectativas de la recuperación económica a nivel mundial. México, por ser uno de los principales socios comerciales de Estados Unidos y con la ventaja de la cercanía geográfica, probablemente sería uno de los países mayormente beneficiado por la creciente cifra al alza de la economía del país norteamericano. Aún con expectativas altas de crecimiento, fluctuación estable del tipo de cambio peso-dólar estadounidense en el segundo trimestre de 2021, se observaron flujos de salida de los extranjeros en los bonos mexicanos, en lo que va del 2021 en montos alrededor de 237 mil millones de pesos (Banco de México, 2021).

Aunado a lo anterior, las presiones inflacionarias en varios países desarrollados y emergentes que parecen ser todo menos transitorias (en varios de ellos se han acelerado las políticas monetarias restrictivas de los bancos centrales) y por último la expectativa de un inicio anticipado en la normalización de la política monetaria de la Reserva Federal de Estados Unidos y su impacto en los mercados internacionales (el llamado *tapering*), es difícil predecir la dirección de los flujos de entrada o salida en los activos de estos mercados en desarrollo y mantienen a los inversionistas con un sentimiento de cautela, reflejado en los flujos neutrales en renta fija para países emergentes. Otro factor importante es la inclusión de China en el índice WGBI, anunciada en septiembre de 2020, con lo que los bonos de este país asiático empezaron a recibir crecientes flujos de capital respecto a otras economías emergentes (Banco de México , 2021), el país asiático concentró el interés de los inversionistas en este activo de deuda. De acuerdo con estimaciones de varios analistas, los flujos potenciales hacia este mercado asiático continuarán hasta 2024, hasta llegar a los cien mil millones de dólares. Esta inclusión y los ajustes en los ponderadores del índice puede provocar mayores salidas de capital en los demás países emergentes, dado el tamaño y el grado de desarrollo del

mercado chino, además de tasas de interés atractivas en los activos de deuda más competitivas que en otros países en desarrollo y por si no fuera poco, la rápida recuperación que ha tenido esta economía ante la crisis sanitaria de 2020.

En cuanto a riesgos locales, sobresalen el posible debilitamiento de la actividad económica en México, la incertidumbre ocasionada por políticas públicas y su impacto en las perspectivas de inversión. La pérdida de grado de inversión de la deuda gubernamental permanece como un riesgo, la materialización de esto podría generar que la deuda gubernamental mexicana salga de los principales índices de inversión globales y esto a su vez genere flujos de salidas más pronunciados de los inversionistas extranjeros en la deuda mexicana.

2 Redes neuronales artificiales y clasificadores lineales

Como parte de los conceptos básicos que serán de utilidad para el análisis de la metodología seleccionada en este trabajo de investigación, en esta sección se revisarán las características de las redes neuronales artificiales, cómo funcionan y cuál es el proceso que siguen para “aprender” comportamientos en los datos. Por otra parte, se describen los tipos principales de redes neuronales que se manejan en diferentes trabajos de investigación y que fueron utilizados en la implementación de la metodología para atacar el problema en este trabajo, por lo que es importante que el lector comprenda dichos conceptos para analizar el diseño de las arquitecturas propuestas e implementadas.

2.1 Características

Una red neuronal es una técnica computacional inspirada en los estudios del funcionamiento del cerebro y el sistema nervioso. Es un sistema de procesamiento de información que imita la capacidad del cerebro humano para comprender relaciones y patrones a través del procesamiento entre células del cerebro: neuronas. Una neurona es una unidad básica que se encarga de procesar información a partir de valores de entrada, está formada por:

- Parámetros de entrada (X)
- Pesos para cada valor de entrada que indican con qué intensidad afecta a la neurona la variable de entrada (w).
- Parámetro de sesgo (o *bias*). Incrementa o decrementa la entrada neta a la función de activación, dependiendo si es positivo o negativo.
- Función de activación $\Phi()$ se encarga de transformar la salida de la neurona con el objetivo de acotar los valores de salida de la neurona y determina la relación entre las entradas y las salidas. Introduce un grado de no linealidad a la resolución del problema, lo cual es valioso para la mayoría de las aplicaciones en redes neuronales.

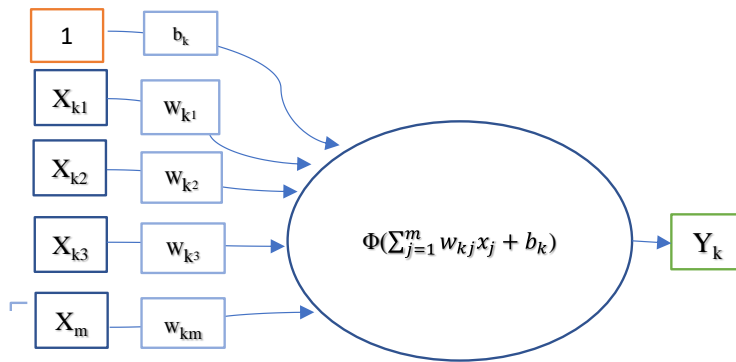
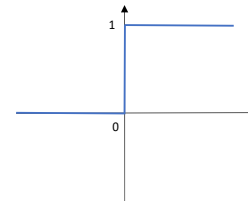


Figura 3 Neurona artificial. Basado en (Ince, et al., 2017), p.6

Se han identificado tres funciones de activación principales que se ocupan en las redes neuronales:

1. **Función escalón.** La salida de la neurona es expresada como:

$$(Y_k) = \begin{cases} 1 & \text{si } v_k \geq 0 \\ 0 & \text{si } v_k < 0 \end{cases}$$



Ecuación 3 Función escalón

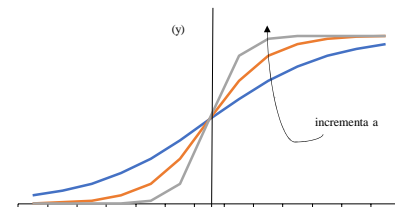
Donde v_k es el cálculo local de la suma ponderada de las variables de entrada:

$$v_k = \sum_{j=1}^m w_{kj}x_j + b_k$$

Ecuación 4 Ponderación de variables de entrada

2. **Función sigmoide.** Es una de las más utilizadas en la construcción de redes neuronales. Está definida como una función de incremento que muestra un balance entre un comportamiento lineal y uno no lineal. Un ejemplo de una función sigmoide es la función logística, definida como:

$$(Y_k) = \frac{1}{1 + \exp(-av)}$$

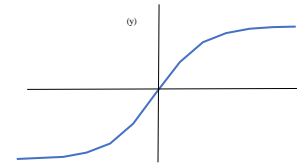


Ecuación 5 Función sigmoide

Donde a es el parámetro de la pendiente de la función sigmoide. Variando este parámetro de pueden obtener diferentes pendientes en la función.

3. **Función tanh o tangente hiperbólica.** Está relacionada con la función sigmoide logística por $\tanh(y) = 2\sigma(a) - 1$. Puede definirse como:

$$(Y_k) = \tanh(y) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$



Ecuación 6 Función tanh

La red neuronal se forma por nodos conectados entre sí, los nodos con características similares se encuentran en una capa (Figura 4). Existen generalmente tres tipos de capas:

- **Capa de entrada.** Es la primera capa de acceso, se encuentran en ella las variables de entrada de la red que serán procesadas.
- **Capas ocultas.** Son las capas intermedias. Para procesar problemas complejos, la red puede tener una o varias capas ocultas. Una capa oculta generalmente puede mapear cualquier función no lineal.
- **Capa de salida.** Es la última capa conectada y contiene la o las variables de salida. El número de nodos en esta capa es relativamente sencillo de especificar ya que está directamente relacionado con el problema de estudio.

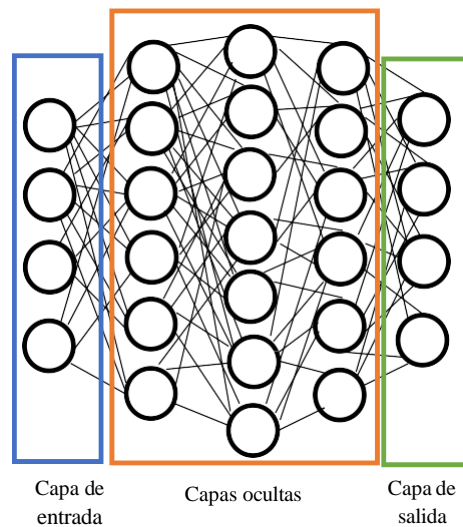


Figura 4 Tipos de capas en una red neuronal

La información se transmite entre las capas a través de conexiones entre los nodos o neuronas. El número de capas ocultas, número de neuronas en cada capa y las funciones de activación entre las neuronas determinan la arquitectura de la red. Cada conexión se configura con un peso que multiplica la señal transmitida. La manera en que la red determina los pesos en las conexiones se conoce como entrenamiento o aprendizaje.

2.2 Entrenamiento

La red neuronal aprende a través de un proceso interactivo de ajuste en los pesos y los niveles de los *bias*, idealmente en cada iteración la red mejora su conocimiento. Si tomamos como ejemplo una red neuronal con una capa de entrada con n parámetros de entrada, con sus correspondientes pesos $w_k(n)$ y una neurona de salida $y_k(n)$, el error que se produce entre la salida de neurona y la salida objetivo del problema $d_k(n)$ se denota como $e_k(n)$.

$$e_k(n) = d_k(n) - y_k(n)$$

Ecuación 7 Error en la red neuronal

El proceso para corregir el error para que el valor de salida de la red neuronal esté lo más cerca posible al valor real del problema, se encargará de modificar los pesos. El proceso de entrenamiento de la red neuronal tiene por objetivo minimizar los errores de salida haciendo actualizaciones de los pesos. Las reglas que definen las actualizaciones constituyen el algoritmo de aprendizaje de la red. Existen varios métodos de optimización que proveen de opciones al entrenamiento de la red. Uno de los más populares es el algoritmo de *backpropagation* que es utilizado para explorar qué tan bien ha aprendido la red.

Por otra parte, la **tasa de aprendizaje** define los tamaños de los pesos que se toman en cada iteración, es decir, cuánto se afecta la actualización de los parámetros en cada iteración. Una tasa demasiado pequeña implicará muchas iteraciones y el algoritmo podría quedarse atrapado en un mínimo local, sin llegar a la solución. Por el contrario, con una tasa demasiado grande los pasos podrían ser tan largos que el algoritmo no converja en ningún punto y se quede atrapado en un ciclo infinito.

Otro aspecto relevante en el proceso de entrenamiento es la especificación de la **función de costo**. La suma de errores cuadrados o el error cuadrático medio son utilizados generalmente (Ecuación 8 y Ecuación 9). Otras funciones de error que también son utilizadas son la desviación absoluta media, la distancia media cuadrática mínima y el porcentaje de error medio.

$$\sum(e_t)^2$$

Ecuación 8 Sumatoria de errores al cuadrado

$$\sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2}$$

Ecuación 9 Error cuadrático medio

Comúnmente el total de datos disponibles son divididos en dos conjuntos; uno de entrenamiento con el que la red neuronal realizará el proceso de aprendizaje y el restante de los datos se utilizan para medir la habilidad de la red para generalizar. Uno de los problemas a los que se enfrenta el proceso de aprendizaje en las redes neuronales es el llamado *overfitting* o sobre entrenamiento en donde el modelo ajusta casi exactamente los datos en el entrenamiento, cuando esto sucede no se generaliza el conocimiento para nuevos datos porque el modelo comienza a aprender el “ruido” o información irrelevante del conjunto de datos o bien, memoriza ese ruido. En estos casos el modelo no podrá realizar predicciones confiables para datos futuros y las decisiones tomadas basada en él podrían incrementar la probabilidad de error. Si los datos de entrenamiento tienen una tasa de error baja y, por el contrario, los datos de prueba tienen una tasa alta de error, es una señal de sobre entrenamiento.

2.3 Tipos de redes neuronales

Cada arquitectura de red neuronal tiene un comportamiento intrínseco en la forma en la que distribuye las diferentes capas, las funciones de activación que utilizan las neuronas y la dirección en que se comunican las conexiones entre neuronas. En este apartado se describen los principales tipos de arquitectura que existen y que han sido utilizadas en otros trabajos con enfoque financiero.

2.3.1 Redes neuronales *feedforward*

También es conocida como red neuronal multi capa. Está compuesta por conexiones unidireccionales entre las capas de la red, es decir, el flujo de conexión viene solo de la capa de entrada a la capa de salida, pasando por las diferentes capas ocultas de la red. La función de activación comúnmente presente es la sigmoidea. Una característica importante de este tipo de red es el aprendizaje supervisado, en donde son utilizados los algoritmos del gradiente descendiente, Levenberg-Marquardt, etc. La velocidad de aprendizaje y el desempeño de este tipo de red depende del número de capas y de neuronas contenidas en el modelo. La complejidad del problema a resolver influye en determinar:

- Número de neuronas en la capa de entrada
- Número de neuronas en la capa de salida
- La cantidad de capas ocultas
- El número de neuronas a utilizar en cada capa

2.3.2 Redes neuronales de base radial

En la forma básica de esta red, se construye con tres capas de neuronas. La primera capa incluye las variables de entrada, la segunda es la única capa oculta de la red que aplica la transformación no lineal a los parámetros de entrada. El proceso de aprendizaje consiste en dos fases; en la fase no supervisada se obtiene el valor de los centroides y la varianza, en la fase supervisada se calculan los pesos de las neuronas de la capa de salida. La capa de salida es de tipo lineal y es la última capa en este tipo de red. La función de transferencia utilizada por la capa oculta es una función de base radial, este tipo de funciones acepta como argumento la distancia entre el vector de entrada y un *centroide* μ preestablecido (Figura 5).

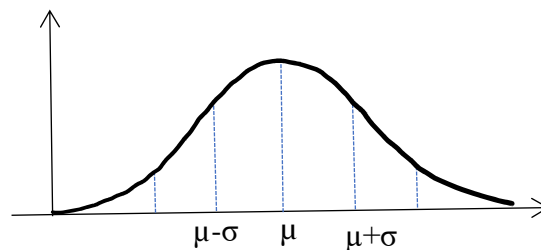


Figura 5 Campana de Gauss

Cuanto más cerca del centroide se encuentre el vector de entrada, mayor es el valor de la función. La función gaussiana puede ser utilizada como de base radial (Ecuación 10). Otras funciones utilizadas son la multi cuadrática y el *spline* poli armónico.

$$y = \exp - \frac{(S - R)^2}{2\sigma^2}$$

Ecuación 10 Función gaussiana

Donde $S = ||X - C||$ es la distancia entre el centroide C y el vector de entrada X . El parámetro escalar σ determina la tasa de descomposición de la función cuando el vector está alejado del centroide. El parámetro R determina el desplazamiento de la función de activación.

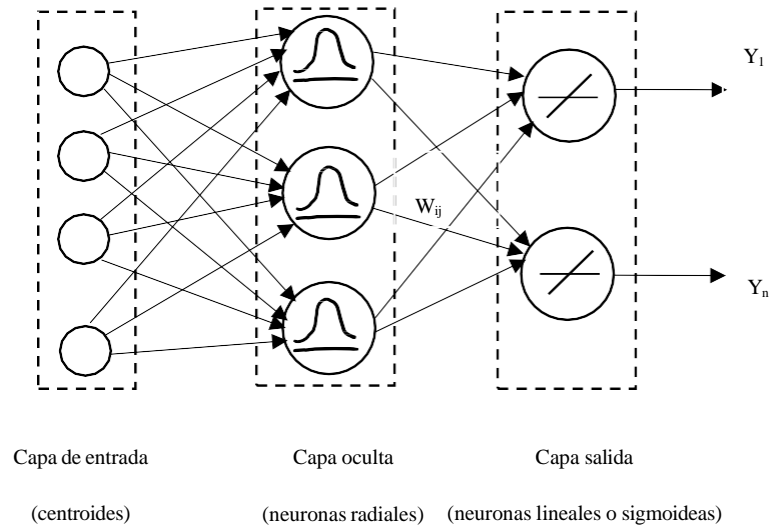


Figura 6 Red neuronal de base radial. Basado en (Martyniuk & Bis, 2021), p. 172

Pueden utilizarse diferentes métodos para calcular la distancia entre los vectores, comúnmente es utilizada la distancia euclidiana:

$$\sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2}$$

Ecuación 11 Distancia euclidiana

Donde X_j es el j-ésimo componente de la capa de entrada y C_j es el j-ésimo elemento del vector que determina la posición del centroide de la función de transferencia.

Cada neurona de la capa oculta determina la distancia entre el vector de entradas y su centroide y realiza una transformación no lineal sobre él. Existe un área llamada radio alrededor del centroide que se relaciona con la sensibilidad de la red y se ajusta usando un vector de coeficientes suavizados $(\sigma_1, \dots, \sigma_m)$. La función de transformación gaussiana toma valores en el rango de 0 a 1 para determinar la salida de la capa oculta. El ajuste del peso determina la salida de la red (Martyniuk & Bis, 2021).

Las ventajas que presenta este tipo de red neuronal son:

- Solo tienen una capa oculta lo cual simplifica la arquitectura de la red
- Velocidad de aprendizaje rápida
- Habilidad de entrenamiento en muestras de datos heterogéneas
- Posibilidad de modelar sistemas con alto grado de inestabilidad
- Capacidad de modelar y predecir procesos aleatorios

Para la fase de aprendizaje supervisado es común utilizar el método de Regla Delta o bien el método de matriz pseudo inversa para calcular los pesos w y el umbral Θ de las neuronas de la capa de salida.

Método Regla Delta

Debido a que las operaciones matriciales son costosas de realizar, en términos computacionales, es recomendable utilizar el método de Regla Delta. Los pesos se inicializan con valores aleatorios pequeños, el umbral de la neurona de salida $\theta=1$, la tasa de aprendizaje y el número de épocas. La salida de la red neuronal se obtiene a partir de:

$$y = \sum_{i=1}^n (w_i * \varphi_i) - \theta$$

Ecuación 12 Salida red neuronal por la regla Delta

La corrección de los pesos w depende del error cometido por la red neuronal dada una muestra i . Esto se hace de forma iterativa hasta encontrar el error mínimo:

$$e_i = (d_i - y_i)$$

Ecuación 13 Error de los pesos

Donde d_i es la salida deseada y y_i la salida de la red.

Para evaluar el error global de la red, se calcula:

$$e_{medio} = \frac{1}{2} (e_i)^2$$

$$E = \frac{1}{p} \sum_{k=1}^p e_{medio}$$

Ecuación 14 Error global

Los pesos se ajustan de acuerdo con el error calculado obteniendo nuevamente la salida de la red. Este proceso se repite hasta encontrar el error mínimo.

2.2.3 Máquina de soporte de vectores (*Support Vector Machine, SVM*)

Es un algoritmo de clasificación inicialmente desarrollado para clasificación binaria pero su aplicación se extendió para otros problemas de clasificación múltiple y regresión. El fundamento de este tipo de red neuronal es el Clasificador de Margen Máximo que se basa en el concepto de hiperplano. En un espacio de n dimensiones, un hiperplano es un subespacio de $n-1$ dimensiones. En el caso de un espacio de dos dimensiones, el hiperplano corresponde a una recta, con ecuación:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p = 0$$

Ecuación 15 Ecuación del hiperplano en un espacio bidimensional

En este caso el hiperplano divide un espacio n dimensional en dos mitades. El objetivo de una máquina de soporte de vectores es encontrar un hiperplano particular para el cual el

margen de separación p entre este hiperplano y los datos más cercanos es maximizado, un hiperplano óptimo:

$$x = x_p + r \frac{w_0}{\|w_0\|}$$

Ecuación 16 Hiperplano óptimo

Donde x_p es una proyección de x en el hiperplano óptimo, r es la distancia algebraica deseada; si es positiva x se encuentra en el lado positivo del hiperplano óptimo. Las observaciones que se encuentran justo en el margen se conocen como vectores de soporte (Figura 7).

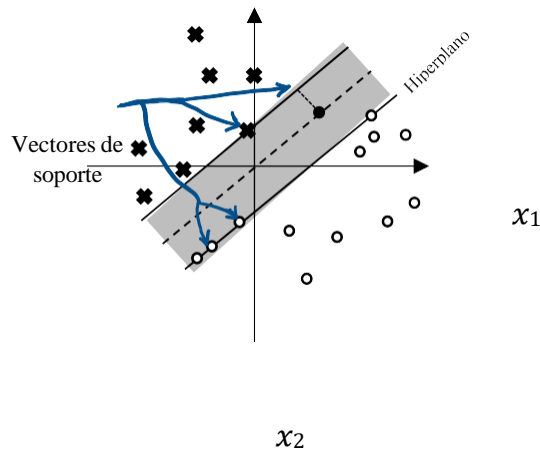


Figura 7 Hiperplano óptimo para clases separables linealmente. Basado en (Haykin, 1998), p. 342

En la práctica es difícil encontrar casos en los que se puedan separar a través de un hiperplano de forma perfecta, el hiperplano es muy sensible a las variaciones en los datos por una parte y por otra un ajuste perfecto puede conllevar a problemas de *overfitting*, como ya se describió en la sección anterior. Por estas razones es mejor crear un clasificador basado en un hiperplano que no separe perfectamente las clases en el que algunas observaciones no estén en el lado correcto de clasificación. El proceso de optimización para obtener el hiperplano óptimo incluye al hiper parámetro o estimador C , que se encarga de controlar el número y severidad de las violaciones en la clasificación que se toleran en el proceso de ajuste. Si $C = \infty$, no se permite ninguna violación del margen, cuanto más se aproxime a cero dicho hiper parámetro, menos penalización tienen los errores y se permite que más observaciones estén

del lado incorrecto del margen o del hiperplano. El valor óptimo de C se identifica mediante validación cruzada, un método que utiliza diferentes partes del conjunto de datos para probar y entrenar el modelo en varias iteraciones. El valor de este hiper parámetro es muy específico para los datos de cada problema.

El método de Máquinas de Soporte de Vectores aumenta la dimensión de los datos en donde es posible conseguir separarlos en clases, de esta forma la separación de datos es posible, aunque de primera vista los datos no sean separables (Figura 8).

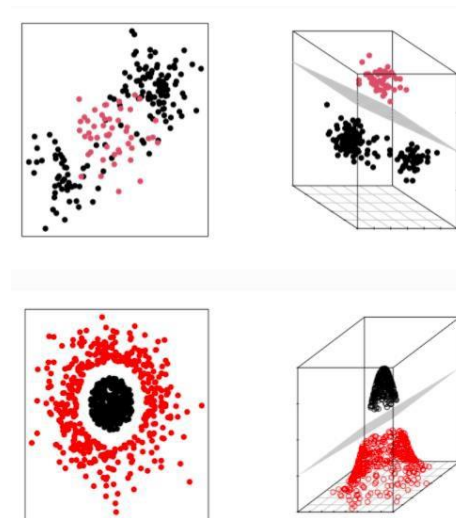


Figura 8 Separación de datos en espacios no bidimensionales. Basado en (Kowalczyk, 2017) , p.23

La dimensión de una muestra de datos se puede transformar a través de modificar o combinar cualquiera de sus dimensiones. Para obtener la función de conversión más adecuada se utiliza la función $kernel(K)$ que de forma general devuelve el resultado del producto escalar entre dos vectores en un nuevo espacio dimensional distinto al espacio original de los vectores para resolver el problema de optimización. Cada $kernel$ tiene un conjunto de hiper parámetros que pueden ser optimizados a través de la validación cruzada. Existen diferentes tipos de $kernels$, depende de la naturaleza de los datos el tipo de $kernel$ a utilizar, entre los más utilizados están:

Kernel lineal: Es el más simple de los *kernels*, se utiliza comúnmente para clasificación de textos. Cuando se utiliza este *kernel* solo se necesita optimizar el hiper parámetro C.

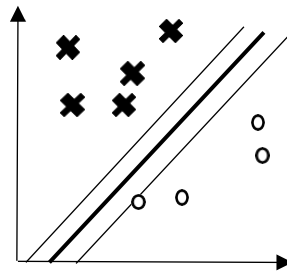


Figura 9 Kernel lineal

Kernel polinómico: Puede obtener la forma de cualquier ecuación de grado 2 en adelante.

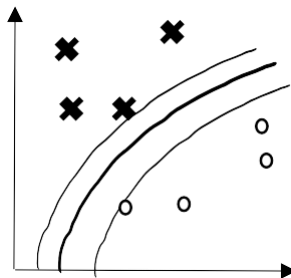


Figura 10 Kernel polinómico

Kernel gaussiano (RBF): Cuando se tiene un conjunto de datos complejo, los *kernels* polinomiales no son lo suficientemente sofisticados para el trabajo, por lo que este tipo de *kernel* puede ser utilizado. También se le llama RBF, que significa función de base radial, que como se describió en el tipo de red neuronal con base radial, el valor de este tipo de funciones depende solo de la distancia al origen o a algún punto.

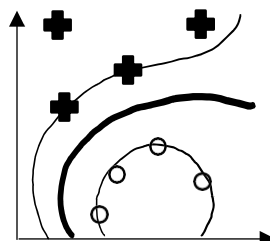


Figura 11 Kernel gaussiano

2.4 Clasificadores lineales y no lineales

Uno de los primeros cuestionamientos para empezar a aplicar el modelo de redes neuronales artificiales en la serie de datos histórica que se procesó fue evaluar de alguna forma si las relaciones entre los datos eran realmente tan complejos para necesitar un modelo de la inteligencia artificial en lugar de uno tradicional, por lo que se implementaron modelos de clasificación lineal para inicialmente justificar el uso de las redes neuronales y posteriormente para comparar la eficiencia entre ambos tipos de modelos. Este subcapítulo abordará las características del clasificador lineal para comparar los resultados de la implementación de un modelo como éste contra un modelo no lineal como el de las redes neuronales. Se explican los conceptos y características básicas para que el lector tenga claras las diferencias entre ambos tipos de modelos.

2.4.1 Clasificación de patrones por funciones de distancia

El proceso de clasificación de un conjunto de puntos en dimensión n , dadas M clases ajenas consiste en asignar a cada punto en alguna de las M clases. Las clases son formadas a partir de una muestra de puntos para posteriormente asignar nuevos puntos a la clase que le corresponde. Uno de los métodos para clasificar consiste en asignar un punto a la clase más cercana. Dicha cercanía puede ser medida con una función de distancia o medida de similitud. Dos de las funciones de distancia más utilizadas son la distancia euclidiana y la de Mahalanobis.

Uno de los problemas a los que nos podemos enfrentar cuando se realiza el proceso de clasificación es encontrar casos en los que la distancia de un punto es la misma con relación a dos puntos, en este caso es necesario establecer criterios para solucionar casos particulares. El desbalance en los datos y el traslape entre las clases de datos usados en el entrenamiento son dos problemas en los que aún se está investigando, por lo que no hay una solución definitiva para resolverlos.

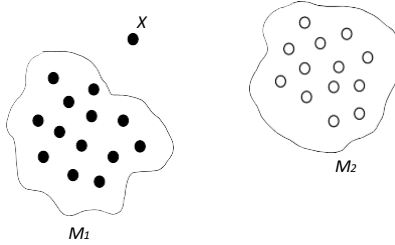


Figura 12 Patrones clasificables por concepto de proximidad

Independiente del clasificador que se seleccione, el proceso de clasificación consiste en los siguientes pasos:

- 1) Obtener la muestra de objetos de clases conocidas y calcular las características de los objetos (prototipos)
- 2) El conjunto de datos de la muestra se utiliza para entrenar el clasificador y se calculan las fronteras entre las clases
- 3) Se extraen las mismas características de los objetos desconocidos a clasificar
- 4) Se utilizan las fronteras calculadas por el clasificador para decidir a qué clases pertenecen los vectores de características de los objetos que se quieren reconocer

Es importante mencionar que el aprendizaje supervisado es aquel que supone que partimos de un conjunto de datos etiquetado o clasificado previamente mientras que el aprendizaje no supervisado parte de datos no etiquetados previamente. Los clasificadores lineales descritos en este apartado corresponden al aprendizaje supervisado.

2.4.2 Clasificador por distancia euclidiana

La distancia euclidiana es la distancia en línea recta entre dos puntos. Si los dos puntos están en un plano de dos dimensiones la fórmula utilizada es la Ecuación 17. Es una medida muy utilizada que tiene en cuenta la variabilidad de cada componente del vector característico. Su uso es ampliamente utilizado en la economía, por ejemplo (Duranton & Overman, 2005) y (Partridge, et al., 2008). Se define como:

$$d(A, B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}$$

Ecuación 17 Distancia euclidiana

2.4.3 Clasificador por distancia Mahalanobis

La distancia de Mahalanobis es una métrica de distancia más robusta que calcula una representación precisa de cuanto dista un punto de una distribución. La Ecuación 18 detalla cómo se realiza el cálculo. Cuando el vector de datos no es homogéneo puede suceder que las distintas componentes están en diferentes escalas por lo que surge la necesidad de usar distancias normalizadas.

$$d_M(x, \mu) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}$$

Ecuación 18 Distancia Mahalanobis

Donde:

El punto $x \in R^n$, generado desde una distribución P

$$\mu = E_P(x)$$

$\Sigma = E_P(x - \mu)^2$, la matriz de covarianza

2.5 Clasificador bayesiano

Es un clasificador probabilístico que se basa en el teorema de Bayes (Ecuación 19). El clasificador bayesiano original asume que todos los atributos de las muestras son independientes y no se considera la información de dependencia entre los atributos. Es uno de los más utilizados por su simplicidad y rapidez. La técnica de clasificación y predicción que construye este modelo predice la probabilidad de posibles resultados con base en el teorema de Bayes, también conocido como teorema de la probabilidad condicionada.

$$p(A|B) = \frac{p(A, B)}{p(B)} = \frac{p(A)p(B|A)}{p(B)} = \frac{p(A)p(B|A)}{\sum_{A'} p(A')p(B|A')}$$

Ecuación 19 Teorema de Bayes

2.6 K-means

Este algoritmo forma parte de los algoritmos de clasificación no supervisada. Agrupa objetos en k clases con base en sus características. De forma general, los pasos que realiza este modelo son:

1. Inicialización: se eligen k centros o centroides en los datos, pueden elegirse de forma aleatoria
2. Asignar objetos a los centroides más cercanos
3. Actualizar centroides de cada grupo tomando un nuevo centroide de la posición del promedio de los objetos pertenecientes a cada grupo

Se repiten los pasos 2 y 3 hasta que los centroides no se mueven o se mueven por debajo de una distancia límite en cada paso (Deng & Gao, 2020). Dados un conjunto de k clústeres y un conjunto de centroides $Z_f(I), j = 1, 2, \dots$, se calcula la distancia de cada objeto al centro del clúster.

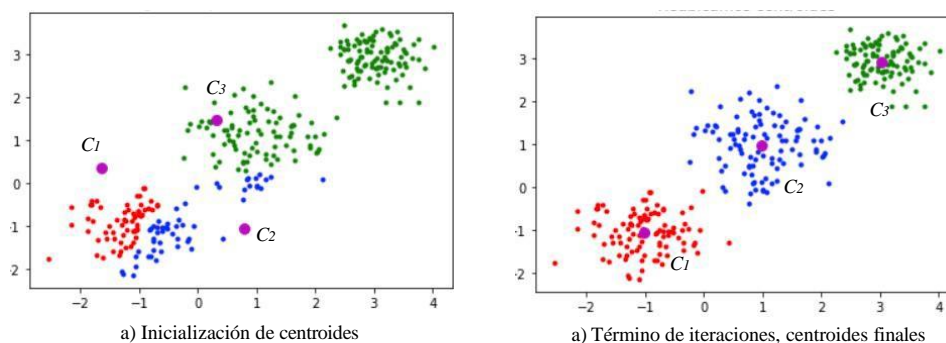


Figura 13 Ejemplo Asignación de centroides k-means

3 Propuesta de metodología

En esta sección se describen los pasos que se siguieron para implementar la arquitectura de las redes neuronales artificiales en la predicción de flujos de entrada o salida de inversionistas extranjeros en los bonos gubernamentales mexicanos. El flujo general de la metodología que se siguió para este trabajo se describe en la Figura 14, cada fase se verá con detalle en este apartado, a manera de resumen para cada uno de los pasos mostrados en el diagrama:

1. Diseño de la base de datos histórica para todas las variables financieras y económicas que pudieran ser factores de consideración por parte de inversionistas no residentes para decidir comprar o vender bonos gubernamentales mexicanos. Además, se diseñó el semáforo de flujos de entradas y salidas para indicarle a la red neuronal la magnitud o el impacto que cada flujo tiene en las condiciones de operación del mercado de renta fija mexicano.
2. Una vez hecho el análisis de la base de datos de estudio, se aplicaron un conjunto de técnicas para procesar los datos atípicos (*outliers*), para completar la historia faltante para algunas de las variables y homogeneizar el periodo de estudio.
3. En una tercera etapa se estudiaron las variables financieras y económicas que se seleccionaron en la primera fase reducir la dimensionalidad e identificar las variables más relevantes para el modelo.
4. El diseño y la implementación de varios modelos de redes neuronales se realizaron en el paso cuatro para encontrar la arquitectura más eficiente para la predicción de los flujos.
5. La selección de los algoritmos de aprendizaje para cada red neuronal se estudia en la etapa cinco del proceso.

6. Finalmente, los resultados de la predicción de flujos a los que llega cada modelo y su eficiencia con respecto a los datos reales se explican en la fase seis.

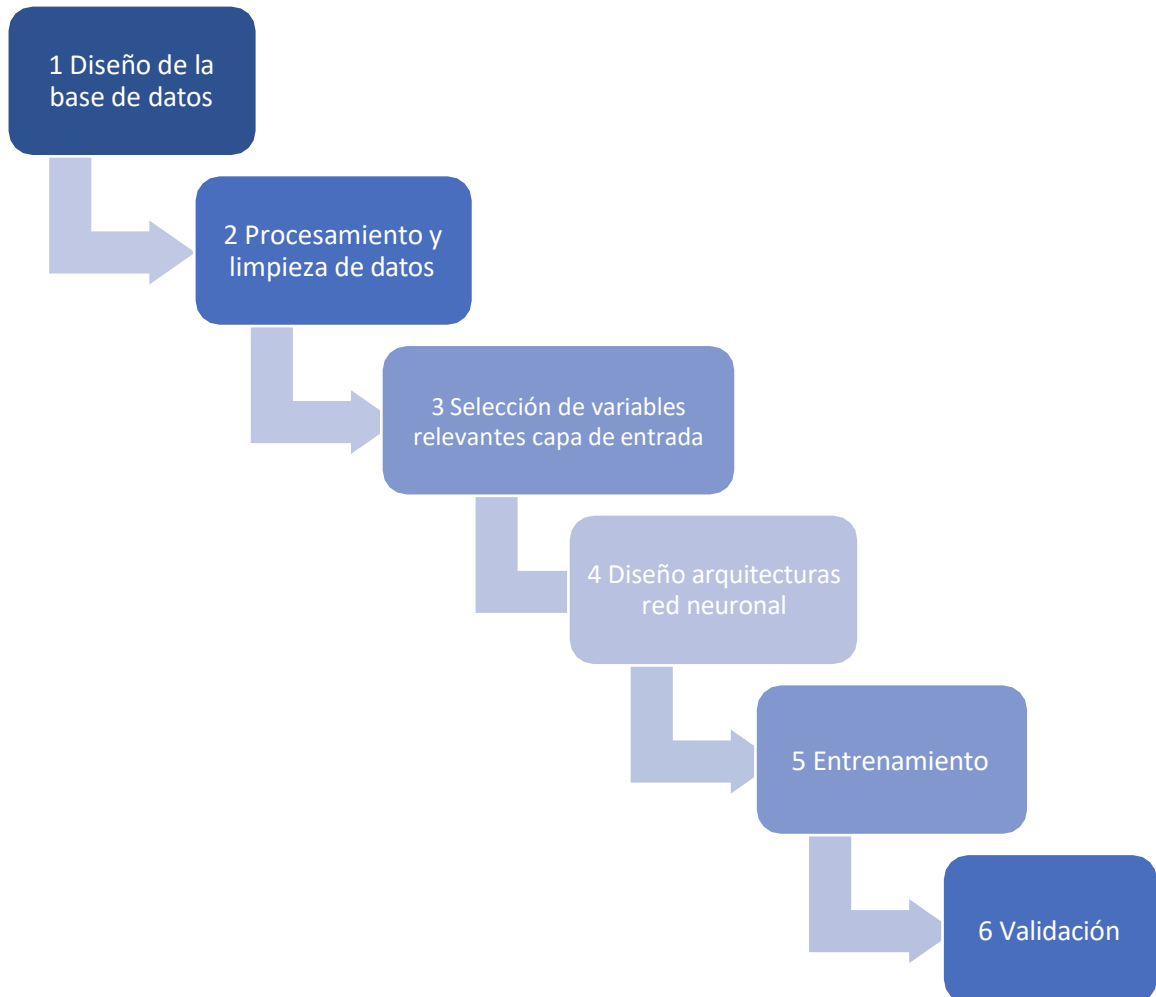


Figura 14 Diagrama de flujo general de la metodología

3.1 Diseño de la base de datos

3.1.1 Variables para la capa de entrada de la red neuronal

La primera fase de la metodología consistió en buscar y enlistar las variables financieras y económicas que pudieran influir en las decisiones de inversionistas extranjeros para comprar o vender bonos gubernamentales mexicanos con base en los factores que se detallan en la sección 1.6. Algunas variables seleccionadas pudieran no tener ninguna relación directa en primera instancia (por ejemplo, las que se relacionan con el mercado cambiario o de renta variable), pero tomando en cuenta que las redes neuronales son capaces de encontrar relaciones no lineales, que a primera vista son inexistentes, se agregaron a la lista de candidatas para formar parte de la capa de entrada de las arquitecturas que más adelante se implementaron. En total se tomaron en cuenta 31 variables, de las cuales 13 corresponden al factor *push factors*, 14 al factor *pull factors* y 4 al factor flujos hacia economías emergentes.

La Tabla 1 muestra un resumen de los factores y las variables que se seleccionaron para la arquitectura del modelo propuesto (en la sección 1.6 se explica en qué consiste cada una de las variables).

Tabla 1 Resumen de variables para la capa de entrada

Factores	Variable	Unidad	Fuente
<i>Push factors</i>	<i>Carry to risk</i>	Índice	Bloomberg
	<i>Carry to risk</i> relativo a otras economías emergentes	Porcentaje	Bloomberg
	Índice de aversión al riesgo global	Índice	Citigroup RAMI (Risk Aversion Index)
	Índice de aversión al riesgo de economías emergentes	Índice	Citigroup RAMI (Risk Aversion Index)
	Índice de condiciones financieras de Estados Unidos	Índice	Bloomberg
	Tasas de rendimiento de los bonos del tesoro (<i>treasuries</i>) de 3, 5, 10 y 30 años	Porcentaje	Bloomberg
	Producto Interno Bruto (PIB) de Estados Unidos	Porcentaje	Bloomberg
	DXY	Índice	
	Índice de materias primas, subíndice industrial de metales	Dólares	Bloomberg
	Índice de materias primas, subíndice de agricultura	Dólares	Bloomberg
	Índice de materias primas, subíndice del petróleo WTI	Dólares	Bloomberg
	Índice de materias primas, subíndice de energía	Dólares	Bloomberg
	Índice de sorpresas inflacionarias para economías emergentes	Índice	Citigroup
	Tasas de rendimiento de bonos mexicanos <i>on the run</i> de 3, 5, 10, 20 y 30 años	Porcentaje	Proveedor Integral de Precios (PIP)
	Pendiente de la curva de rendimientos de bonos gubernamentales mexicanos medido por el diferencial 2-10 años y 3-30 años	Porcentaje	Proveedor Integral de Precios (PIP)
	<i>Breakeven</i> de 3, 10 y 30 años	Porcentaje	Proveedor Integral de Precios (PIP)
	Reservas internacionales más línea de crédito flexible	Millones de dólares	Proveedor Integral de Precios (PIP)
	Deuda de México como porcentaje del PIB	Porcentaje	

<i>Pull factors</i>	Volatilidad implícita en opciones cambiarias dólar norteamericano – peso mexicano con plazo de un mes	Porcentaje	Bloomberg
	Sesgo implícito en opciones cambiarias dólar norteamericano – peso mexicano con plazo de tres meses	Porcentaje	Bloomberg
	Curtosis implícita en opciones cambiarias dólar norteamericano – peso mexicano con plazo de tres meses	Porcentaje	Bloomberg
	Tipo de cambio <i>spot</i> del peso mexicano – dólar estadounidense	Peso/dólar	Bloomberg
	Producto Interno Bruto (PIB) de México	Porcentaje	Banco Mundial
	Posición en <i>forwards</i> y <i>swaps</i> cambiarios con el tipo de cambio peso mexicano - dólar estadounidense pactados en México	Millones de dólares	Banco de México
	Razón de cobertura cambiaria para instrumentos de renta fija (RCC)	Porcentaje	State Street
	Calificación promedio de la deuda mexicana	Calificación promedio	Moody's, Fitch, Standard & Pooors
	Indicador Global de la Actividad Económica	Porcentaje	INEGI
Flujos hacia economías emergentes	Flujo en activos de renta fija dirigidos hacia China	Millones de dólares	Emerging Portfolio Fund Research (EPFR)
	Flujo de activos de renta variable dirigidos hacia China	Millones de dólares	Emerging Portfolio Fund Research (EPFR)
	Índice de bonos gubernamentales en países emergentes (EMBI) JP Morgan	Dólares	JP Morgan
	Índice de activos de renta variable MSCI	Dólares	Bloomberg

3.1.2 Capa de salida de la red neuronal: semáforo de flujos de entradas y salidas

Los datos utilizados para la capa de salida de la red neuronal son los flujos de tenencias a valor nominal, en millones de pesos, de bonos gubernamentales mexicanos en manos de inversionistas extranjeros publicados por el Banco de México de forma diaria y con un rezago de siete días con respecto al día actual. La información se encuentra disponible en el portal de internet del banco central desde 1978, sin embargo, por cuestiones de homogeneización de esta variable con las de la capa de entrada de la red neuronal y además considerando que las condiciones financieras en los años noventa eran muy diferentes a las que se desarrollaron dados los cambios estructurales realizados en el pasado y mencionados en la sección 1.1, se recopiló información a partir de 2004 para el análisis de esta investigación.

Otra consideración que se tomó en cuenta fue la diferencia que existe en cuestión de operación, liquidez y tenencia en los bonos de corto, mediano y largo plazo, por lo que los nodos de vencimiento de los bonos soberanos se clasificaron en canastas, con respecto al número de años resultado de la diferencia entre la fecha de vencimiento del título y la fecha de consulta del bono:

- Corto plazo: bonos con hasta tres años de plazo vigente.
- Mediano plazo: bonos con más de tres años y hasta diez años de plazo vigente.
- Largo plazo: bonos con plazo vigente mayor a diez años.

Una vez clasificado el nodo en las canastas anteriores, se suma el saldo de la tenencia para el día t de los bonos que cumplan con cada uno de los plazos descritos arriba. Finalmente, se calculan los flujos para cada plazo como la tenencia en el día t (T_t) menos la tenencia del día anterior (T_{t-1}):

$$T_t - T_{t-1}$$

Ecuación 20 Cálculo de flujos de bonos gubernamentales

Con base en el objetivo de la red neuronal que es predecir la tendencia de los flujos de entrada o salida de bonos gubernamentales mexicanos por parte de inversionistas extranjeros, el diseño de la capa de salida debe contener un semáforo que no solo indique si habrá una entrada o salida, sino que indique la magnitud de estos flujos, para poder tomar acciones preventivas en el mercado. Para resolver este planteamiento, se realizó un análisis de la distribución de los flujos de entrada y salida y se clasificó en 7 clases de acuerdo con el número de desviaciones estándar del flujo, además de utilizar el criterio y experiencia del autor en el análisis diario y mensual de los flujos de inversión y su implicación en el comportamiento del mercado de renta fija. Para aplicar la metodología de las redes neuronales artificiales se decidió hacerlo en dos fases; la primera consistió en entrenar una red neuronal solo indicando flujos de entrada y de salida para posteriormente en la fase 2, entrenar redes neuronales que se especializaran en la magnitud de los flujos de entrada y en los flujos de salida. El detalle de este diseño se explica a continuación:

1. **Fase 1: binario, arquitectura de red neuronal general**, con capa de salida con dos clases:
 - Flujos de entrada → Clase 1
 - Flujos de salida → Clase -1

2. **Fase 2: magnitud de flujos, arquitectura de red neuronal especializada en flujos de entradas y arquitectura de red neuronal especializada en flujos de salidas**
 - Arquitectura para flujos de entrada
 - Flujos de entrada neutros → Clase 1
 - Flujos de entrada medios → Clase 2
 - Flujos de entrada fuertes → Clase 3

 - Arquitectura para flujos de salida
 - Flujos de salida neutros → Clase -1
 - Flujos de salida medios → Clase -2
 - Flujos de salida fuertes → Clase -3

3.1.2.1 Fase 1: binario, arquitectura de red neuronal general

Se clasificaron los flujos de cada canasta de plazos en dos clases:

- Clase 1: flujos mayores a cero (Si se incrementa la tenencia de los bonos el día $t+1$ con respecto al día t , se considera que hubo una entrada)
- Clase -1: flujos menores a cero (Si disminuye la tenencia de los bonos el día $t+1$ con respecto al día t , se considera que hubo una salida)

El siguiente paso fue analizar que la distribución de la clasificación realizada en la base de datos fuese equiprobable, es decir, que la cantidad de observaciones dentro de cada clase no difiera demasiado entre una y otra y por lo tanto sean igualmente probables. La Figura 15 muestra la cantidad de registros en la base de datos que se clasificó en la clase con identificador -1 (flujos negativos o salidas en las tenencias) y el número de observaciones para la clase 1 (flujos positivos o entradas en la tenencia de bonos por parte de extranjeros). Como segunda clasificación en los datos, se identificó el plazo del bono en el que un inversionista extranjero decide entrar o salir, en la gráfica mencionada se identifica en barras de colores la concentración en la cantidad de registros que se seleccionó para cada clase. Se puede observar que cada clase cuenta con un rango entre 1,150 y 1,597 registros, con lo que se puede tener mayor certeza de que no exista algún sesgo entre las clases de la capa de salida de la red neuronal.

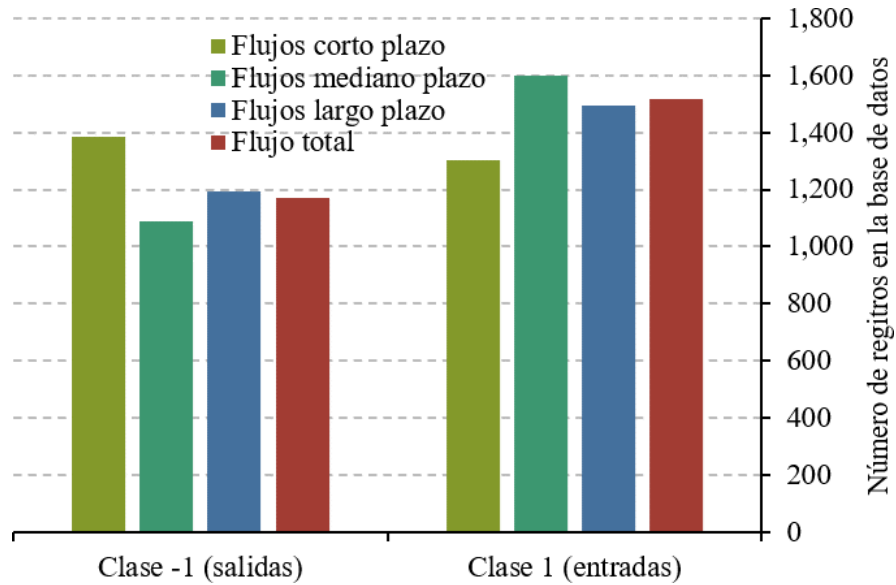


Figura 15 Flujos de entradas y salidas por plazo con dos clases (Número de registros de la muestra). Elaboración propia con datos de las tenencias de bonos gubernamentales en manos de extranjeros, obtenidos del Banco de México

3.1.2.2 Fase 2: magnitud de flujos, arquitectura de red especializada

Para la segunda etapa se construyó una clasificación especializada para la capa de salida de la red neuronal con base a la magnitud del flujo de entrada o de salida. Para clasificar los flujos en cada clase se tomó el equivalente a las desviaciones estándar de la distribución de flujos especificados en la Tabla 2. Esta metodología además es congruente con lo que el autor observa en los análisis diarios sobre los flujos de las tenencias de los inversionistas extranjeros y su implicación en el comportamiento del mercado de bonos.

Tabla 2 Semáforo de flujos especializado en entradas y salidas

Tipo de flujo	Clase	Desviaciones estándar (σ)	Intervalos (millones de pesos)			
			Flujos de corto plazo	Flujos de mediano plazo	Flujos de largo plazo	Flujos totales
Entrada	3 (flujos fuertes)	3 σ	> 3,000	> 1,400	> 1,400	> 3,000
	2 (flujos medios)	2 σ	[700 a 3,000)	[700 a 1,400)	[700 a 1,400)	[700 a 3,000)
	1 (flujos neutros)	1 σ	[0 a 700)	[0 a 700)	[0 a 700)	[0 a 700)
Salida	-1 (flujos neutros)	1 σ	[-0.01 a -700)	[-0.01 a -700)	[-0.01 a -700)	[-0.01 a -700)
	-2 (flujos medios)	2 σ	[-700 a -3,000)	[-700 a -1,400)	[-700 a -1,400)	[-700 a -3,000)
	-3 (flujos fuertes)	3 σ	< -3,000	< -1,400	< -1,400	< -3,000

3.2 Procesamiento y limpieza de datos

El siguiente paso fue obtener la serie histórica de las variables de la capa de entrada de las diferentes fuentes de información para posteriormente realizar un análisis estadístico de cada serie con el fin de identificar comportamientos extraños en algunos periodos de tiempo, valores erróneos en los cálculos o valores extremos (*outliers*) y darles algún tipo de tratamiento. Esta identificación requiere el registro y análisis de algunos eventos económicos mundiales que han afectado visiblemente a varios mercados, incluyendo los emergentes, por lo que a continuación se enlistan los principales eventos que se tomaron en cuenta para explicar valores atípicos debido a éstos:

- La gran crisis financiera de 2008 en la que el mercado de créditos hipotecarios colapsó y hubo un contagio en los mercados accionario, de renta fija y de derivados.
- Crisis griega 2011
- *Taper tantrum*, se le conoce de esta forma al episodio que inició cuando la Reserva Federal de Estados Unidos anunció en 2013 una posible reducción anticipada de sus compras de bonos (medida derivada de las acciones para combatir la crisis de 2008), esto provocó caídas en varios mercados como el de bonos y el de renta variable de economías emergentes.
- Desaceleración económica China en 2015
- Elecciones de Estados Unidos 2016
- Contingencia sanitaria por COVID-19, 2020

En cuanto a eventos locales, algunos eventos que han ocurrido en el pasado y han afectado los mercados financieros nacionales son:

- Tensiones comerciales en 2018 por la renegociación del TLCAN entre México y Estados Unidos
- Cancelación del aeropuerto internacional de Texcoco, 2019

En el análisis gráfico que se realizó para identificar posibles datos erróneos, los indicadores con comportamientos extremos que no pudieran ser explicados por los eventos arriba mencionados, fueron analizados a detalle para identificar la cifra y decidir quitarla, suavizarla

o conservarla en caso de que su existencia estuviese justificada por la fuente de información (por ejemplo, por cambios en metodologías). A continuación, se presentan algunos ejemplos de datos con valores correctos explicados por eventos económicos (Figura 16) y datos encontrados con algún error en la base de datos original (Figura 17).

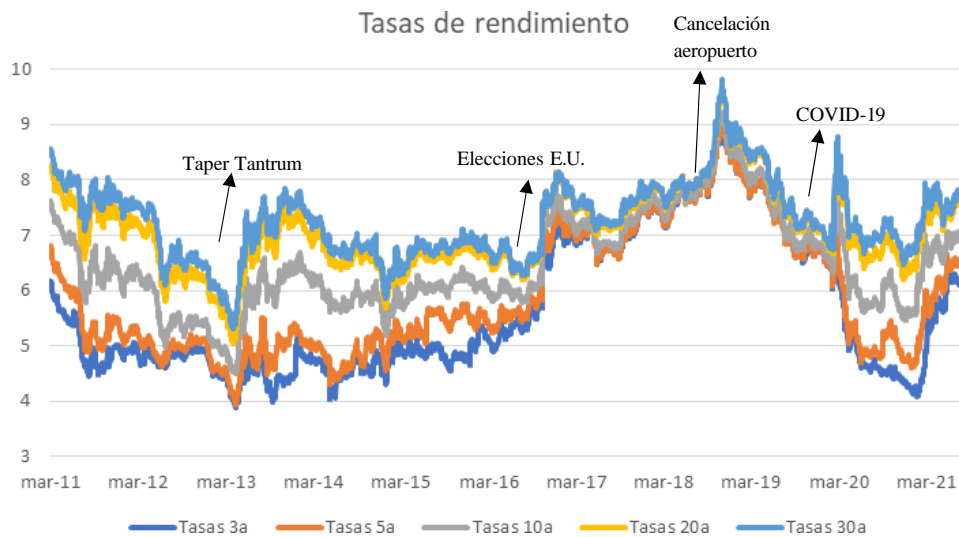


Figura 16 Tasas de rendimiento de los bonos gubernamentales por plazo con eventos seleccionados

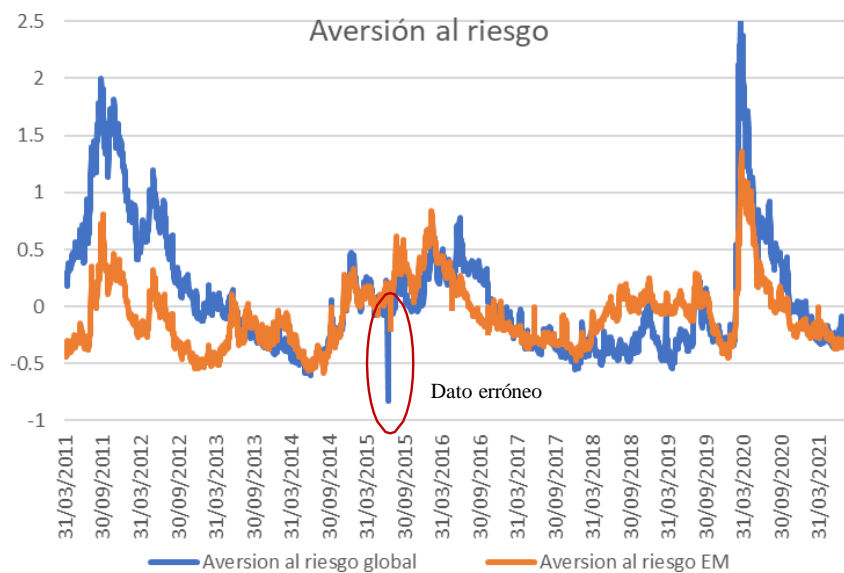


Figura 17 Índice de aversión al riesgo con dato erróneo encontrado

Otro de los problemas que se encontró fue huecos de datos para algunas variables debido a problemas con la fuente original de información o por errores en los cálculos para fechas históricas muy antiguas. Las técnicas que se aplicaron para esos huecos de información fueron:

- Calcular promedios móviles de los últimos 30 días disponibles
- Interpolación lineal, para aproximar el valor desconocido mediante la fórmula:

$$Y = Y_1 + \left[\left(\frac{X - X_1}{X_2 - X_1} \right) (Y_2 - Y_1) \right]$$

Ecuación 21 Interpolación lineal

Una vez completados los huecos de información, se realizó nuevamente un análisis gráfico para detectar comportamientos extraños en las series históricas, en este caso, no se identificaron datos atípicos.

3.2.1 Herramientas para el procesamiento de la serie de datos

El lenguaje de programación Python es uno de los más populares en la actualidad para implementar trabajos de aprendizaje de máquina y de redes neuronales porque contiene varios paquetes de funciones que tienen implementadas las tareas necesarias para ejecutarlos. En esta investigación se utilizó la versión 3.6.5 de Python, el código fuente se escribió en el *notebook* de *Jupyter*. La librería *scikit-learn* contiene un conjunto de herramientas que permiten preprocesar los datos a utilizar; hacer transformaciones lineales, estandarizar, normalizar, entre otras. La estandarización de datos es un requisito común para varios estimadores de aprendizaje de máquina, la mayoría de ellos asumen ciertas características (por ejemplo, datos centrados alrededor de cero y con varianza en el mismo orden) en la serie de datos por lo que es importante aplicar las funciones de preprocesamiento de ese paquete de funciones para darle un tratamiento adecuado a las variables de la capa de entrada de la red neuronal.

3.2.2 Normalización de datos

Debido a la naturaleza de cada variable considerada para la capa de entrada de la red neural (detallada en la Tabla 1) fue importante estandarizar o normalizar la serie de datos construida, recordando que este proceso de ajuste tiene por objetivo contar con uniformidad en la información y por otra parte los algoritmos de aprendizaje de las redes neuronales hacen énfasis en la importancia de realizar este procesamiento en los datos. Se utilizó *StandardScaler* para hacer esta transformación en los datos, *StandardScaler* transforma los datos de manera que su distribución tenga un valor medio μ de cero y una desviación estándar σ de uno, Ecuación 22.

$$z = \frac{x - \mu}{\sigma}$$

Ecuación 22 Estandarización

Con media:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

Ecuación 23 Media

Y desviación estándar:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

Ecuación 24 Desviación estándar

3.3 Selección de variables relevantes para la capa de entrada

El análisis de componentes principales (*Principal Component Analysis PCA*) es un método que se utiliza para reducir la dimensionalidad en un conjunto de datos y por lo tanto simplificar la complejidad de múltiples dimensiones, sin que los datos originales pierdan información (Figura 18) (Rodrigo, 2022). Debido a la cantidad de variables candidatas a ser parte de la capa de entrada de la red neuronal, se utilizó este método para ayudar en el rendimiento de las arquitecturas de redes neuronales.

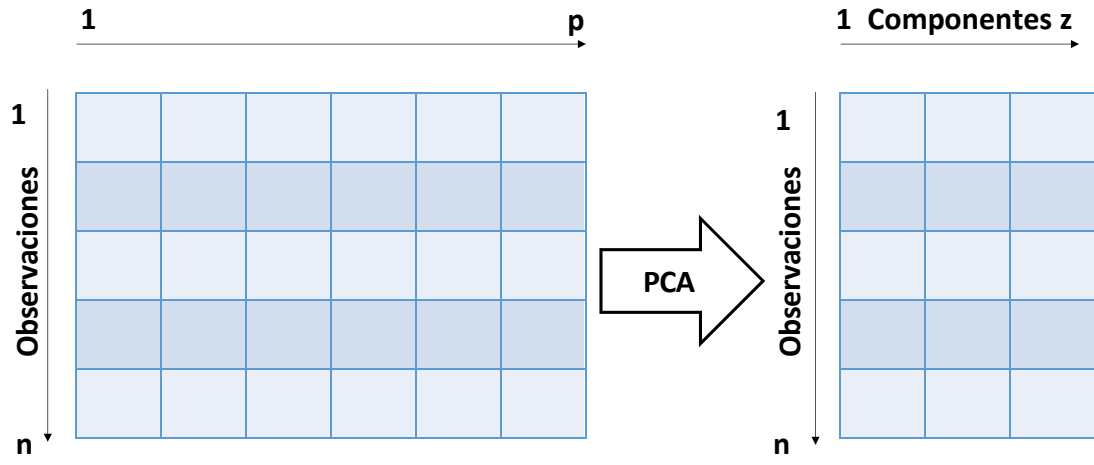


Figura 18 Dimensionalidad PCA

Es importante recordar que, para aplicar esta técnica, los datos deben estar normalizados, como ya se explicó en la sección anterior, de lo contrario aquellas variables cuya escala sea mayor dominarán al resto. La varianza explicada por cada componente principal y la varianza acumulada son valores de gran utilidad para decidir el número de componentes principales a utilizar en análisis posteriores, si se utilizan todos los componentes principales, se está almacenando toda la información presente en los datos originales. La suma de la varianza acumulada de todas las componentes es siempre 1. Es común seleccionar la cantidad de componentes a través de la evaluación de la proporción de la varianza explicada acumulada y seleccionar el número de componentes mínimo a partir del cual el incremento deja de ser significativo.

Para el caso de los componentes principales calculados para las variables de la Tabla 1, se muestra el detalle del porcentaje de varianza explicada por cada componente en la Figura 19 y el porcentaje de varianza acumulada en la Figura 20. A partir del análisis realizado con estas dos gráficas, se decidió tomar los primeros 11 componentes principales para la capa de entrada de la red neuronal, aunque también se hizo el experimento de utilizar todas las variables candidatas para posteriormente hacer un análisis de rendimiento y eficiencia. Los primeros 11 componentes explican el 95.32% de la varianza de los datos (como se observa en la Figura 20) los componentes subsecuentes a éstos explican un porcentaje agregado muy pequeño por lo que se decidió tomar sólo los primeros 11.



Figura 19 Porcentaje de varianza explicada por componente principal

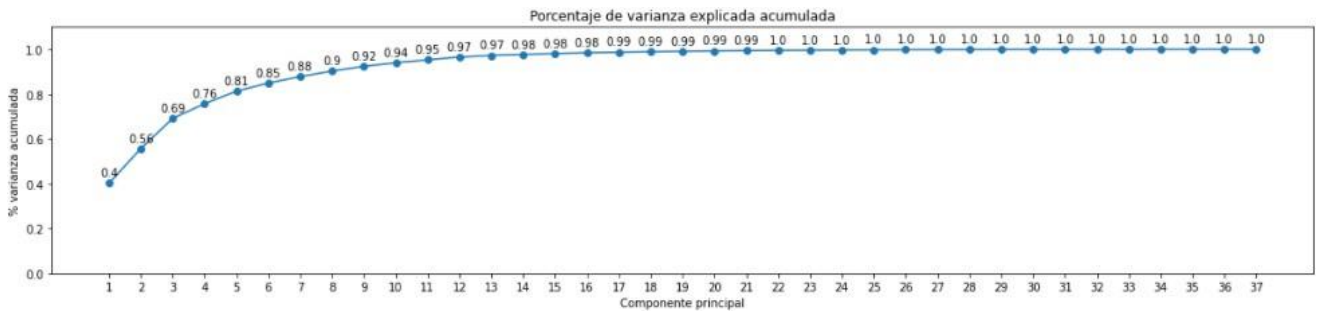


Figura 20 Porcentaje de varianza explicada acumulada

3.4 Diseño de las arquitecturas de las redes neuronales

Se diseñaron e implementaron diferentes arquitecturas de los tipos de red neuronal que se describen en el marco teórico de la sección 2.3.1 para obtener los resultados con mayor eficiencia en la predicción de los flujos de inversión. A continuación, se describen los pasos y parámetros implementados para discutir los resultados en el capítulo posterior.

3.4.1 Diseño de red neuronal *feedforward*

Se diseñaron varias arquitecturas de este tipo de red neuronal, considerando y variando los parámetros:

- Número de neuronas en la capa de entrada
- Número de neuronas en la capa de salida
- Número de capas ocultas (mínimo 1)
- Número de neuronas en cada capa oculta

Se implementó una clase en Python con un constructor de inicialización de variables que recibiera como parámetros los descritos arriba. La clase inicializa en números aleatorios los pesos sinápticos de las neuronas de la capa de entrada y los valores del *bias* se igualan a cero. El siguiente paso es definir la neurona sigmoidea que realizará en primera instancia la ponderación de las variables de entrada, generalizada por la Ecuación 4, por ejemplo, para la variable 1:

$$a_1 = w_1 * x_1 + w_2 * x_2 + b_1$$

Ecuación 25 Ponderación de variable de entrada 1

Lo siguiente fue aplicar la función sigmoidea al resultado obtenido para cada variable de la capa de entrada (Ecuación 5). La siguiente capa de la red se encarga de recibir los resultados ponderados de la capa de entrada y calcular las derivadas parciales de esos pesos con respecto a la función de pérdida. De forma análoga a lo anterior, se hizo una iteración a través de cada punto de datos, se calculan las derivadas parciales guardando los resultados en un conjunto

de variables para ser actualizados en siguientes iteraciones. Por último, se tiene la función de predicción para la serie de datos.

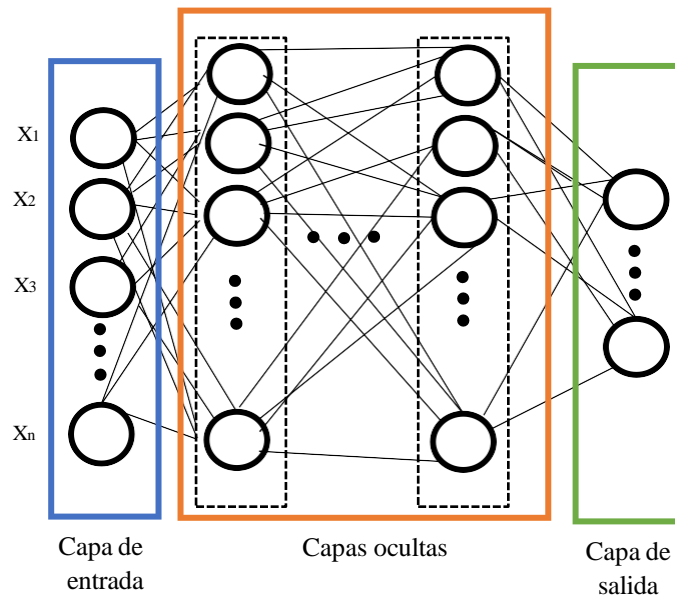


Figura 21 Diseño de red neuronal feedforward

Con base al semáforo de flujos diseñado, fue necesario implementar varias arquitecturas de red neuronal *feedforward* para modelar 4 casos especificados en Tabla 3.

Tabla 3 Modelos de red neuronal *feedforward*

Núm. modelo	Capa de entrada	Capas ocultas	Capa de salida
Modelo 1	11 componentes principales	Mínimo 1 capa oculta con una neurona	Entradas (1) Salidas (-1)
Modelo 2	31 variables normalizadas	Mínimo 1 capa oculta con una neurona	Entradas (1) Salidas (-1)
Modelo 3	11 componentes principales	Mínimo 1 capa oculta con una neurona	Entradas: Fuertes (3) Medias (2) Neutras (1)
Modelo 4	31 variables normalizadas	Mínimo 1 capa oculta con una neurona	Salidas: Fuertes (-3) Medias (-2) Neutras (-1)

Para encontrar la arquitectura de red neuronal *feedforward* más eficiente combinando los parámetros comentados para cada modelo mencionado, se implementaron un conjunto de ciclos para mover dichos parámetros con base al incremento en la eficiencia de la red y con un número específico de vueltas en el ciclo y tiempo de respuesta.

3.4.2 Diseño de red neuronal de base radial

La etapa 1 de esta red neuronal consistió en obtener el valor de los centroides μ y la varianza σ a partir de las variables de entrada, esta etapa es conocida como entrenamiento no supervisado. Es importante recordar que en la primera capa de la red no hay pesos sinápticos, solo están los centroides. Para este paso se implementó el algoritmo de agrupamiento *k-means* para entrenar los valores $\{\mu_1, \mu_2, \dots, \mu_k\}$. La etapa 2, también conocida como la fase de aprendizaje supervisado, es la que se encarga de procesar el resultado de la fase anterior con neuronas radiales de la capa oculta. A la capa de salida se le agregó una función de activación sigmoide ya que se necesita clasificar la salida en las clases del semáforo de flujos de entradas y salidas. La Figura 22 detalla de forma gráfica las dos fases utilizadas en el diseño.

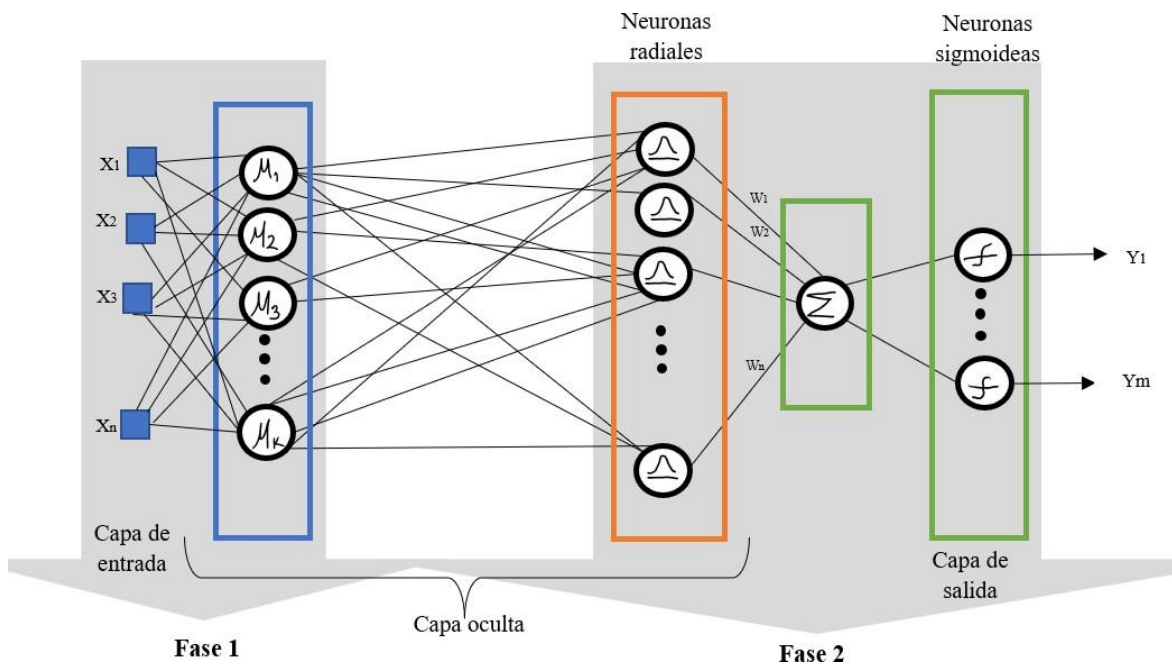


Figura 22 Diseño de red neuronal de base radial

Para construir la salida del semáforo de flujos, se continuó utilizando cuatro modelos para esta arquitectura, descritos en la Tabla 4.

Tabla 4 Modelos de red neuronal de base radial

Núm. modelo	Variables de entrada	Capa oculta	Salida (Y)
Modelo 1	11 componentes principales	Neuronas radiales	Entradas (1) Salidas (-1)
Modelo 2	31 variables normalizadas	Neuronas radiales	Entradas (1) Salidas (-1)
Modelo 3	11 componentes principales	Neuronas radiales	Entradas: Fuertes (3) Medias (2) Neutras (1)
Modelo 4	31 variables normalizadas	Neuronas radiales	Salidas: Fuertes (-3) Medias (-2) Neutras (-1)

3.4.3 Diseño de red neuronal SVM

Se implementó la red neuronal SVM intercambiando *kernels* de tipo lineal, polinómico y radial, para los cuatro modelos descritos en la Tabla 5 y representados en la Figura 23.

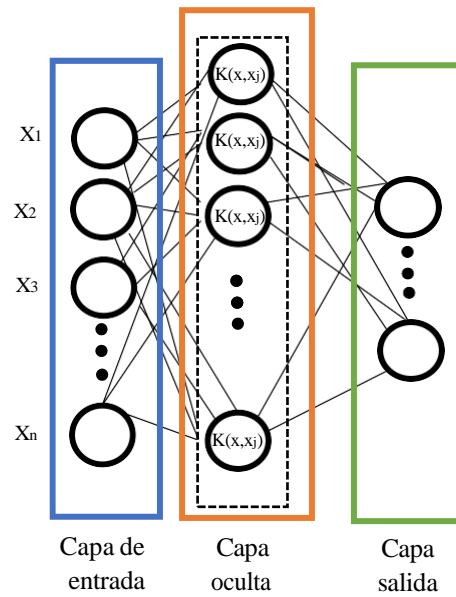


Figura 23 Modelo de red neuronal SVM

Tabla 5 Modelos de red neuronal SVM

Núm. modelo	Variables de entrada	Capa oculta	Salida
Modelo 1	11 componentes principales	<i>Kernel (K):</i> lineal <i>Kernel (K):</i> polinómico <i>Kernel (K):</i> radial	Entradas (1) Salidas (-1)
Modelo 2	31 variables normalizadas	<i>Kernel (K):</i> lineal <i>Kernel (K):</i> polinómico <i>Kernel (K):</i> radial	Entradas (1) Salidas (-1)
Modelo 3	11 componentes principales	<i>Kernel (K):</i> lineal <i>Kernel (K):</i> polinómico <i>Kernel (K):</i> radial	Entradas: Fuerzas (3) Medias (2) Neutras (1)
Modelo 4	31 variables normalizadas	<i>Kernel (K):</i> lineal <i>Kernel (K):</i> polinómico <i>Kernel (K):</i> radial	Salidas: Fuerzas (-3) Medias (-2) Neutras (-1)

3.5 Entrenamiento

El tamaño de la base de datos utilizada fue de 4,389 registros, se tomó el 70% para entrenamiento y el 30% para validación en todos los modelos de las arquitecturas presentadas. La Tabla 6 muestra un resumen de los conjuntos de datos utilizados para entrenar y probar las arquitecturas implementadas. En las siguientes subsecciones se describe el proceso de entrenamiento para cada arquitectura de red neuronal.

Tabla 6 Conjuntos de datos para la red neuronal

Nombre del periodo	Número de variables (columnas)	Número de registros	Fecha inicio	Fecha fin
Total, base de datos	31	4,389	01/07/2004	03/12/2021
Conjunto de entrenamiento	31	3,071	01/07/2004	03/12/2021
Conjunto de prueba	31	1,318	01/07/2004	03/12/2021

3.5.1 Entrenamiento de red neuronal *feedforward*

El entrenamiento de esta red neuronal involucró una combinación entre los hiper parámetros:

- Número de épocas: la cantidad de vueltas en que la red neuronal ajusta sus parámetros para reducir el error
- Tasa de aprendizaje: indica cuánto se avanza en cada paso, define el comportamiento del algoritmo

El algoritmo de aprendizaje que se utilizó fue el de *backpropagation* con el descenso del gradiente para minimizar el costo del error, se programó una función que calcula el vector del gradiente, es decir, las derivadas parciales de los parámetros con respecto al costo.

3.5.2 Entrenamiento de red neuronal de base radial

La primera fase de la red neuronal de base radial, la de aprendizaje no supervisado se implementó con el algoritmo de clúster *K-means*, que intenta separar los datos en k grupos de igual varianza, minimizando la inercia (medida de la coherencia interna de los clústeres). Los parámetros que requiere este algoritmo son:

- Número de clústeres inicial
- Inicialización de los centroides: aleatoria o predeterminada

Los hiper parámetros que necesitan ser configurados para esta fase en la red neuronal de base radial son los mismos que para un *feedforward* para la etapa de aprendizaje supervisado: número de épocas y tasa de aprendizaje, adicionalmente se envían como parámetros los centroides calculados en la fase de aprendizaje no supervisado y la varianza de éstos. El algoritmo de aprendizaje utilizado consistió en calcular los pesos w y el umbral Θ de las neuronas de capa de salida, para esto se utilizó el método de Regla Delta.

3.5.3 Entrenamiento de red neuronal SVM

Para el caso de la red neuronal de soporte de vectores se realizó una búsqueda de la combinación de parámetros óptima (*GridSearchCV*) entre los valores de los hiperparámetros:

- C : controla la tolerancia del número y severidad de las violaciones en la clasificación, es común para todos los tipos de *kernel*. Un valor pequeño de este parámetro suaviza la severidad, un valor alto apunta a clasificar correctamente todos los ejemplos de la muestra de datos
- Γ (Gamma): nivel de ajuste del conjunto de datos, aplica para el *kernel* radial

Se realizó el procedimiento anterior intentando con secuencias de crecimiento exponencial para C y γ , ($C=2^{-5}, 2^{-3}, \dots, 2^{15}$, $\Gamma=2^{-15}, 2^{-13}, \dots, 2^3$) como recomienda (Hsu, et al., 2016). Con el conjunto de valores encontrados por el optimizador, se entrena la red neuronal.

3.6 Validación

Para validar los resultados experimentales obtenidos y evaluar el poder de predicción de cada arquitectura y modelo se utilizó la métrica del error cuadrado medio (*Root Mean Square Error*, *RMSE*, detallado en la Ecuación 9), recordando que éste mide la calidad de predicción de un modelo a través de las diferencias entre los valores predichos y los reales. En el caso de que el *RMSE* del conjunto de prueba sea más alto que el de los datos de entrenamiento podría implicar una alta probabilidad de que el modelo tenga *overfitting* (Syriopoulos, et al., 2021). Se realizó un total de 30 experimentos por modelo y arquitectura de red neuronal para verificar estadísticamente la eficiencia para cada uno.

4 Análisis de los resultados de la aplicación de redes neuronales en la predicción de flujos

Este capítulo presenta las configuraciones implementadas en las diferentes arquitecturas de redes neuronales que se describieron en el diseño y que tuvieron los mejores resultados en el porcentaje de eficiencia de los pronósticos y menor error cuadrado medio. Se detalla el proceso de esta configuración, además, se presentan los resultados de los clasificadores lineales para comparar ambos tipos de modelos en la propuesta de solución. Los porcentajes de eficiencia de cada modelo y la comparación entre ellos se menciona en la subsección dos de este apartado, correspondiente al análisis descriptivo de los resultados y finalmente la discusión de los resultados obtenidos se detalla en la subsección tres.

Cabe mencionar que en este apartado se utiliza el término hiper parámetro para hacer referencia a las configuraciones realizadas en las arquitecturas de redes neuronales y diferenciar con el concepto de parámetro o variable de entrada de la capa de entrada de estas redes neuronales.

4.1 Protocolo experimental

4.1.1 Capa de entrada y de salida para las redes neuronales

Para la capa de entrada fue necesario considerar el número de neuronas óptimo que fuera capaz de describir adecuadamente el problema y que optimizara el rendimiento del modelo con respecto a recursos computacionales y resultados de entrenamiento y prueba en cada arquitectura. La primera prueba realizada consistió en utilizar los datos de todas las variables sin normalizar como parámetros de entrada, con esto se obtuvieron resultados con niveles de error alto y porcentajes de eficiencia de pronóstico bajos y volátiles en las etapas de entrenamiento y prueba. Aunado a esto, la mayoría de las arquitecturas de redes neuronales artificiales recomiendan que estas variables de entrada estén normalizadas para no generar sesgos en las ponderaciones para ningún parámetro. Además, en la Tabla 1 se observa la diferencia entre las unidades de las variables utilizadas, por lo que la normalización de los indicadores fue un factor importante para el rendimiento y eficiencia en los resultados. Dado lo anterior, los resultados de los experimentos realizados para las arquitecturas de redes neuronales desarrollados en este trabajo se presentan tomando en consideración en la capa de entrada valores normalizados de las variables para un grupo de experimentos y los 11 primeros componentes principales calculados con base a estos datos normalizados para otro grupo de pruebas y aplicados para cada una de las arquitecturas detalladas en el diseño.

Por su parte, el diseño de la capa de salida de las redes neuronales fue igual de importante, en primera instancia se realizaron experimentos utilizando seis clases que describían la magnitud de las salidas y de las entradas, el resultado para todas las arquitecturas no superó el 20% de eficiencia, por lo que se decidió replantear la estrategia con la generación de dos fases de experimentación: la primera fase con una capa de salida binaria de entradas y salidas y una fase posterior con tres clases solo para entradas y tres clases únicamente para flujos de salida, esta metodología se explica con detalle en la sección 3.1.2. Utilizar esta técnica de dividir las tareas en una red neuronal binaria que exclusivamente aprendiera las relaciones entre salidas y entrada y posteriormente especializar otras redes neuronales en la magnitud de entradas y salidas mejoró notablemente los resultados.

En la Tabla 7 se resume el conjunto de experimentos realizados con los dos tipos de capa de entrada (variables normalizadas y PCAs) y con las dos fases para la capa de salida (binaria y magnitud). Como se puede observar, se realizaron 6 grupos de experimentos combinando las capas de entrada con los diferentes tipos de redes neuronales (*feedforward*, de base radial y SVM) y con cada diseño de la capa de salida (fase binaria y de magnitud de flujos de entradas y salidas).

Tabla 7 Configuración de la capa de entrada y de salida

Grupo de experimentos	Capa de entrada	Tipo de red neuronal	Capa de salida
1	<ul style="list-style-type: none"> • Variables normalizadas (31 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 1: binaria (2 neuronas) <ul style="list-style-type: none"> ○ Entradas ○ Salidas
2	<ul style="list-style-type: none"> • Variables normalizadas (31 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 2: magnitud <ul style="list-style-type: none"> → Entradas (3 neuronas) ○ Entradas neutras ○ Entradas medias ○ Entradas fuertes
3	<ul style="list-style-type: none"> • Variables normalizadas (31 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 2: magnitud <ul style="list-style-type: none"> → Salidas (3 neuronas) ○ Salidas neutras ○ Salidas medias ○ Salidas fuertes
4	<ul style="list-style-type: none"> • PCAs (11 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 1: binaria (2 neuronas) <ul style="list-style-type: none"> ○ Entradas ○ Salidas
5	<ul style="list-style-type: none"> • PCAs (11 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 2: magnitud <ul style="list-style-type: none"> → Entradas (3 neuronas) ○ Entradas neutras ○ Entradas medias ○ Entradas fuertes
6	<ul style="list-style-type: none"> • PCAs (11 neuronas) 	<ul style="list-style-type: none"> • <i>Feedforward</i> • Base radial • SVM 	<ul style="list-style-type: none"> • Fase 2: magnitud <ul style="list-style-type: none"> → Salidas (3 neuronas) ○ Salidas neutras ○ Salidas medias ○ Salidas fuertes

A continuación, se describen las configuraciones realizadas para los clasificadores lineales y para cada modelo de redes neuronales que presentaron los mejores resultados en la eficiencia de pronóstico, cada tipo de red neuronal también presentó resultados diferentes respecto a rendimiento y eficiencia en la predicción de flujos con base al tipo de capa de entrada y de salida utilizada.

4.1.2 Clasificadores lineales

Cada clasificador lineal se aplicó en la base de datos con las 31 variables de entrada, se consideraron 3 clases para los flujos de entrada y 3 para los flujos de salida, tal como se diseñó la fase 2 del semáforo de flujos (sección 3.1.2.2). Se tomaron en consideración los plazos en los vencimientos de los bonos de los flujos de las tenencias de inversionistas extranjeros: corto, mediano, largo plazo y el flujo total sin considerar el plazo. Se tomó, de forma aleatoria, el 70% de la muestra total para entrenar al modelo y el resto para la validación. Finalmente, el porcentaje de eficiencia se calculó como:

$$\frac{\sum v}{n}$$

Ecuación 26 Cálculo del porcentaje de eficiencia de los clasificadores lineales

Donde:

v = Cantidad de veces que el modelo acertó a la clase correcta

n = Número total de la muestra para entrenar o para validar

La Tabla 8 muestra los resultados obtenidos al aplicar los clasificadores lineales en la base de datos. Como prueba experimental inicial se consideraron los datos de las variables de entrada sin normalizar y normalizados para observar posibles cambios en el comportamiento de estos modelos. Como se puede observar, la máxima eficiencia obtenida fue del 52% al utilizar el clasificador bayesiano con los PCAs de los flujos de largo plazo, este porcentaje de eficiencia no resulta significativo para poder confiar en un modelo del tipo clasificador lineal como metodología de pronóstico de flujos.

Tabla 8 Resultados del porcentaje de eficiencia de los clasificadores lineales

Tipo de clasificador	Plazo de los flujos de entrada y salidas	Fase de entrenamiento		Fase de prueba	
		Eficiencia datos normalizados (%)	Eficiencia con PCAs (%)	Eficiencia datos normalizados (%)	Eficiencia con PCAs (%)
Distancia euclidiana	Corto plazo	8	51	9	47
	Mediano plazo	11	48	10	45
	Largo plazo	14	37	12	37
	Total, sin plazo	13	33	11	34
Distancia Mahalanobis	Corto plazo	8	51	9	47
	Mediano plazo	9	11	11	10
	Largo plazo	10	37	37	37
	Total, sin plazo	12	23	26	23
Bayesiano	Corto plazo	16	45	42	51
	Mediano plazo	19	40	46	48
	Largo plazo	17	38	39	52
	Total, sin plazo	15	30	41	39

Una vez realizado el experimento de clasificar los flujos con un modelo lineal y analizar sus resultados, es justificable aplicar el modelo de redes neuronales para buscar un mejor porcentaje de eficiencia en el pronóstico del comportamiento de los flujos de entrada y de salida.

4.1.3 Configuración red neuronal *feedforward*

Dado que en el punto inicial de esta arquitectura de red neuronal los pesos se asignan de forma aleatoria, los pronósticos que se obtienen pueden diferir entre los modelos. Las iteraciones con la base de datos implicaron un costo computacional alto. La mejor configuración para este tipo de red neuronal fue al utilizar la base de datos de componentes principales, cinco capas ocultas con tres neuronas sigmoideas en cada capa, el promedio del porcentaje de eficiencia para la red neurona general en el entrenamiento fue de 24.6%, mientras que en la etapa de validación fue de 25.5% (Tabla 14). A partir de colocar ocho capas ocultas, el tiempo de ejecución incrementó exponencialmente y el porcentaje de eficiencia mejoró marginalmente. También es posible visualizar en la Figura 24 que, a partir de 150 épocas configuradas, la tasa de error del modelo no disminuye de manera considerable, sin embargo, se establecieron hasta 400 épocas para intentar que la tasa de error disminuyera lo más posible. Por otra parte, se observó un porcentaje de eficiencia menor cuando se utilizaron todas las variables normalizadas, además de que el costo computacional fue más alto que la ejecución con la base de datos de los PCAs para la capa de entrada. El tiempo promedio que se tardó el algoritmo en correr la mejor configuración fue de 7 minutos por iteración. La Tabla 9 detalla el valor de cada hiper parámetro configurado en esta red neuronal.

Tabla 9 Configuración de hiper parámetros de la red neuronal general feedforward

Red neuronal feedforward		
Hiper parámetros	PCAs	Variables normalizadas
Tasa de aprendizaje	0.001	0.001
Número de épocas	400	400
Número de nodos en la capa de entrada	11	31
Número de nodos en la capa de salida	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3
Número de capas ocultas	5	7
Número de neuronas en las capas ocultas	3	5
Función de activación	Sigmoidea	Sigmoidea

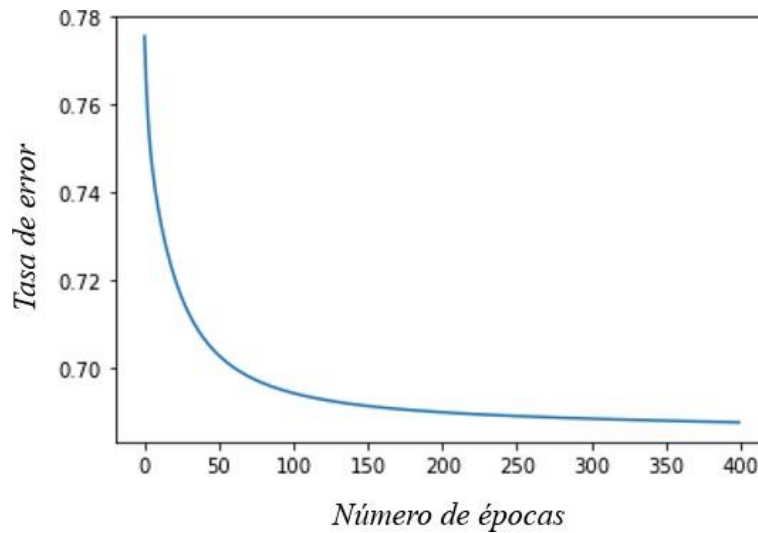


Figura 24 Convergencia de la tasa de error en la red neuronal feedforward

4.1.4 Configuración red neuronal de base radial

La estructura del diseño fue más simple para este tipo de red neuronal además de una baja complejidad computacional y desempeño adaptativo superior que una red neuronal *feedforward*. Sin embargo, un problema presentado con esta arquitectura fue determinar el número de centroides, no hay algún método específico para la asignación inicial y la red neuronal depende de ellos y de los pesos obtenidos. El método utilizado para definir estos centroides fue iniciar con un aleatorio.

Tabla 10 Configuración de hiper parámetros red neuronal de base radial

Red neuronal base radial

Hiper parámetros	PCAs	Variables normalizadas
Tasa de aprendizaje	0.001	0.001
Número de épocas	500	500
Número de nodos en la capa de entrada	11	31
Número de nodos en la capa de salida	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3
Número de capas ocultas	1	1
Número de neuronas en las capas ocultas	6	7
Función de activación	Gaussiana	Gaussiana

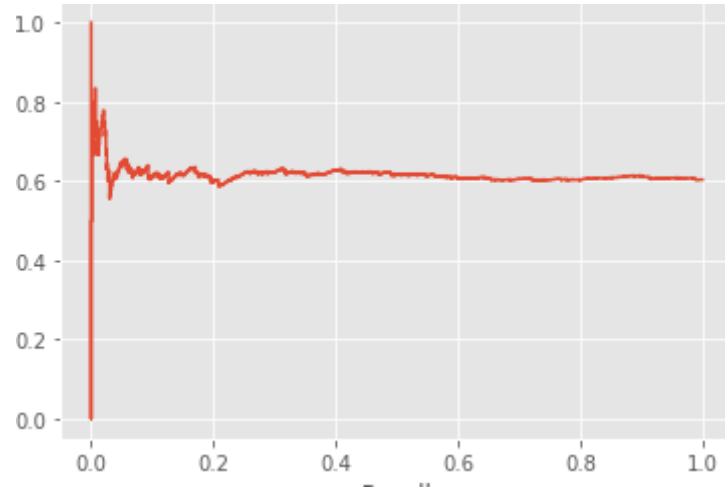


Figura 25 Convergencia de la tasa de error en la red neuronal de base radial

4.1.5 Configuración red neuronal SVM

Este tipo de red neuronal tuvo varias ventajas en su implementación como el número de hiper parámetros a configurar y la velocidad al ejecutar el algoritmo, lo complejo fue encontrar la mejor combinación posible para los valores de C y γ para lo cual la librería *sklearn* de Python fue de gran utilidad para encontrar esa combinación y para ejecutar el algoritmo de SVM sobre los flujos de las tenencias de extranjeros. Los experimentos realizados con un *kernel* lineal, polinómico y sigmooidal resultaron con porcentajes de eficiencia mucho menores que al utilizar uno de base radial. El tiempo de ejecución promedio para encontrar la mejor combinación de hiper parámetros fue de 3 minutos.

Tabla 11 Configuración de hiper parámetros de la red neuronal SVM

Red neuronal SVM		
Hiper parámetros	PCAs	Variables normalizadas
Número de nodos en la capa de entrada	11	31
Número de nodos en la capa de salida	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3	Fase 1 (red neuronal general): 2 Fase 2 (red neuronal especializada): 3
Valor de C	10	10
Valor de γ	0.001	0.001
<i>kernel</i>	Base radial	Base radial

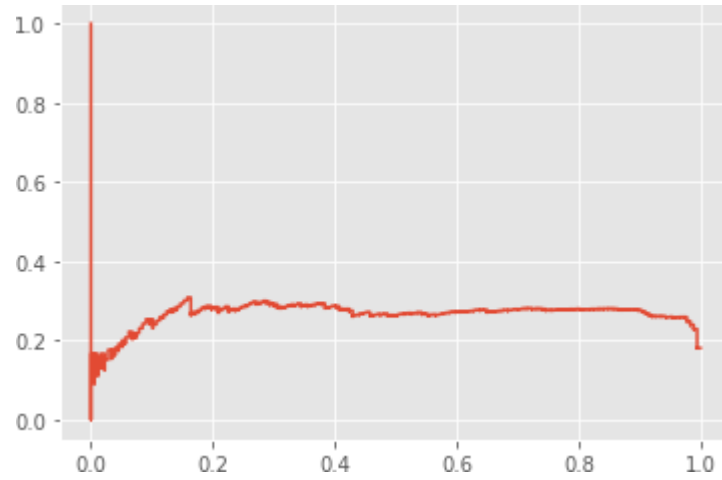


Figura 26 Convergencia de la tasa de error en la red neuronal SVM

4.1.6 Validación de experimentos

Se realizó un grupo de experimentos para la fase 1 de la capa de salida en la que se identifican solo dos clases: entradas y salidas y otro conjunto de pruebas para la fase 2 en la que se clasifica la magnitud de entradas por una parte y por otra la magnitud de las salidas a través de los niveles del semáforo detallado en la Tabla 2. A continuación, se presentan los resultados del error cuadrático medio para cada arquitectura de red neuronal con la configuración descrita en las secciones correspondientes al protocolo experimental. Cabe mencionar que los resultados mostrados de las redes neuronales se calcularon con los datos de prueba generados de manera aleatoria, para recordar, se utilizó el 70% de la base de datos, seleccionados de forma aleatoria, para entrenar a cada modelo, mientras el 30% restante se usó para validarlos, los datos precisos del número de registros que se probó en cada etapa de entrenamiento y prueba se pueden consultar la Tabla 6.

Tabla 12 Resultado del error cuadrático medio de las arquitecturas de redes neuronales

Arquitectura	Fase											
	1 (binaria)				2 (magnitud) Entradas				2 (magnitud) Salidas			
	Flujos CP	Flujos MP	Flujos LP	Flujos Totales	Flujos CP	Flujos MP	Flujos LP	Flujos Totales	Flujos CP	Flujos MP	Flujos LP	Flujos Totales
<i>Feedforward</i>	2.76	2.76	2.76	2.68	0.92	1.41	1.10	1.26	1.00	1.10	1.32	1.30
Base radial	2.36	2.36	2.36	2.09	0.57	0.76	0.89	0.70	0.97	0.88	0.76	0.58
SVM	1.76	1.55	1.69	0.62	0.29	0.66	0.43	0.49	0.31	0.49	0.44	0.39

Como se observa en la Tabla 12 la arquitectura con la red neuronal SVM, en la fase 2 con el semáforo de flujos de entradas en el corto plazo obtuvo el menor error cuadrático medio respecto a todos los experimentos realizados (0.29). Por otra parte, el modelo con el mayor error cuadrático fue la arquitectura *feedforward* en la fase 1 para los datos divididos por plazo. En general la arquitectura de red neuronal SVM obtuvo los menores errores cuadráticos en los experimentos realizados, mientras que las redes neuronales *feedforward* obtuvieron los errores más grandes en las predicciones para flujos de inversionistas extranjeros en bonos mexicanos. Las redes neuronales de base radial fueron mejores que las *feedforward*, pero no superaron a las SVM.

Para ver el comportamiento de las predicciones realizadas por las diferentes arquitecturas y fases se realizó el siguiente tratamiento:

1. Para la cada una de las fases del semáforo de la capa de salida de la red neuronal (1 binaria y 2 magnitud de flujos) se tomó cada registro de la muestra de prueba para identificar la correspondencia del valor que predijo la red neuronal y el valor real del flujo con los intervalos descritos en la Tabla 2. Por ejemplo, para la fase 2 del semáforo aplicado a entradas, de corto plazo, los valores de los flujos pueden ser: 1 para flujos de entrada neutros, 2 para flujos de entrada de mediana magnitud y 3 para flujos de entrada fuertes. Si el valor predicho fue igual a 2, el intervalo correspondiente a éste es [700 a 3,000) .
2. A continuación, se suma el valor del límite inferior más el límite superior del intervalo encontrado
3. Finalmente se acumula el resultado del paso anterior, iniciando en 0 para el primer registro y se va sumando el valor del paso dos.

En la Tabla 13 se presenta un ejemplo de cómo se aplica el procedimiento arriba descrito en los flujos de entradas de corto plazo para el semáforo de fase 2. El valor real para el número de registro 3 equivale a 2, lo cual significa que el intervalo de este flujo de entrada va de 700 (límite inferior [d]) a 3,000 (límite superior [e]) millones de pesos, es un flujo de entrada medio. El flujo acumulado de este valor real es 4,400, que es la suma de los valores anteriores.

Esta forma de graficar los resultados permite ver la tendencia de las predicciones y la tendencia de los valores observados, además es posible comparar la diferencia entre estas tendencias y el nivel del diferencial entre los datos predichos y los reales.

Tabla 13 Ejemplo de la generación de flujos acumulados para entradas de corto plazo

Número de registro	Flujos de Entradas de corto plazo						Flujos acumulados			
	Predicción	Limite inf (a)	Limite sup (b)	Valor real	Limite inf (d)	Limite sup (e)	c = a + b	Predicción	f = d + e	Valor real
1	2	700	3,000	1	0	700	3,700	0	700	0
2	2	700	3,000	1	0	700	3,700	3,700	700	700
3	1	0	700	2	700	3,000	700	4,400	3,700	4,400
4	2	700	3,000	2	700	3,000	3,700	8,100	3,700	8,100
5	1	0	700	1	0	700	700	8,800	700	8,800
6	1	0	700	3	3,000	20,000	700	9,500	23,000	31,800
7	1	0	700	1	0	700	700	10,200	700	32,500
8	1	0	700	3	3,000	20,000	700	10,900	23,000	55,500
9	1	0	700	1	0	700	700	11,600	700	56,200
10	2	700	3,000	3	3,000	20,000	3,700	15,300	23,000	79,200

Así, en la gráfica de la Figura 27 se puede observar la arquitectura con menor valor de error cuadrático medio (red neuronal SVM para entradas de corto plazo), la tendencia de la predicción y la de los datos observados se ve bastante similar, aunque la serie de los datos predichos por la red neuronal indican una magnitud menor que los datos que realmente pasaron, en este caso para los flujos de corto plazo.

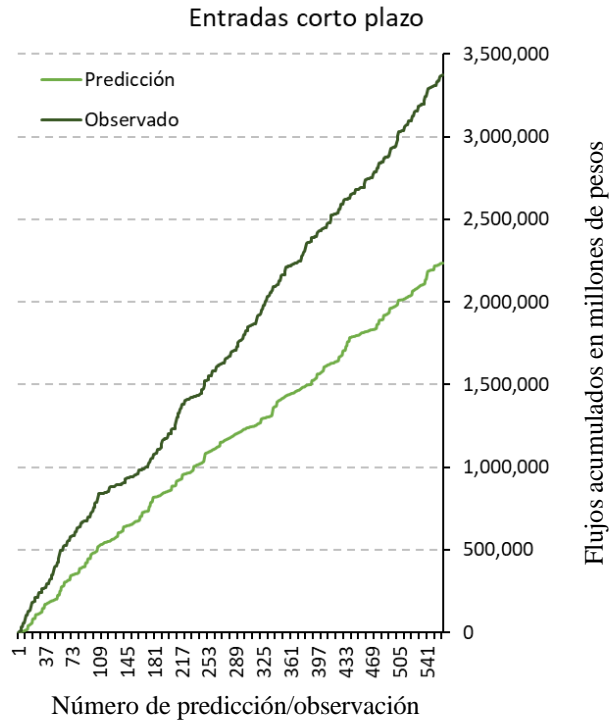


Figura 27 Flujos acumulados de corto plazo resultados de la red neuronal SVM

4.2 Análisis descriptivo de los resultados

Como ya se mencionó anteriormente, se realizaron 30 experimentos con cada modelo con la mejor configuración encontrada, detallada en las tablas del protocolo experimental, y para cada grupo de experimentos detallados en la Tabla 7, de esta forma la eficiencia puede ser estadísticamente validada con base al teorema central del límite. Además de validar las arquitecturas con el nivel del error cuadrático medio como ya se presentó en el punto anterior, también se analizó el porcentaje de eficiencia de cada modelo, el cual puede observarse en la Tabla 14, lo que hace comparable la tabla de resultados del porcentaje de eficiencia con la de los modelos lineales. El promedio del porcentaje de eficiencia del total de experimentos por cada arquitectura y número de fase es el que se menciona en este apartado para la etapa de prueba o validación de los modelos. La eficiencia más alta resultado del grupo de experimentos fue de 84.4% con la arquitectura de red neuronal SVM para los flujos totales, en la fase binaria. Esta arquitectura de red neuronal que identifica un flujo de entrada o de salida de los inversionistas extranjeros en todos los nodos de bonos mexicanos es capaz de predecir un flujo en el futuro con una eficiencia alta, sin embargo, cuando se intenta predecir la magnitud de la entrada o salida, la eficiencia disminuye hasta 28%. Como se puede observar el porcentaje de eficiencia se deteriora considerablemente en la fase 2 para la magnitud de entradas y salidas para la mayoría de los modelos, aunque la misma arquitectura de red neuronal SVM presenta los mejores resultados para los flujos totales y por plazo. Respecto al promedio de eficiencia que logró cada tipo de red neuronal, el de SVM es el único por arriba del 50% y tiene la eficiencia promedio más alta mientras en la red neuronal de tipo *feedforward* es la que presenta los resultados más bajos en la predicción de flujos incluso varios modelos de los clasificadores lineales de la Tabla 8 la superan en porcentaje de eficiencia.

Tabla 14 Porcentaje de eficiencia de las arquitecturas y fases de redes neuronales

Arquitectura	Fase												Promedio (%)
	1 (binaria)				2 (magnitud) Entradas				2 (magnitud) Salidas				
	Flujos CP	Flujos MP	Flujos LP	Flujos Totales	Flujos CP	Flujos MP	Flujos LP	Flujos Totales	Flujos CP	Flujos MP	Flujos LP	Flujos Totales	
Feedforward	25.6	23.4	24.2	29.9	24.6	24.5	26.7	25.7	25.2	23.8	27.1	24.7	25.5
Base radial	50.9	49.2	54.2	50.2	47.2	50.2	48.1	47.7	48.2	47.1	50.4	49.8	49.4
SVM	54.8	60.1	56.1	84.4	58.3	55.7	60.7	66.4	57.6	55.3	62.3	66.2	61.5

El clasificador bayesiano con datos normalizados y flujos de largo plazo obtuvo un 52% de eficiencia, comparable con el promedio de eficiencia de la arquitectura de red neuronal de base radial y fue superado por el promedio de los resultados de la arquitectura de red neuronal SVM. En la Figura 28 se puede observar el porcentaje promedio de los modelos de red neuronal implementados en las diferentes fases para la capa de salida (binaria y de magnitud). Al comparar este porcentaje de eficiencia contra el promedio de eficiencia de los clasificadores lineales (28%) se observa como las redes neuronales de base radial y SVM fueron superiores.

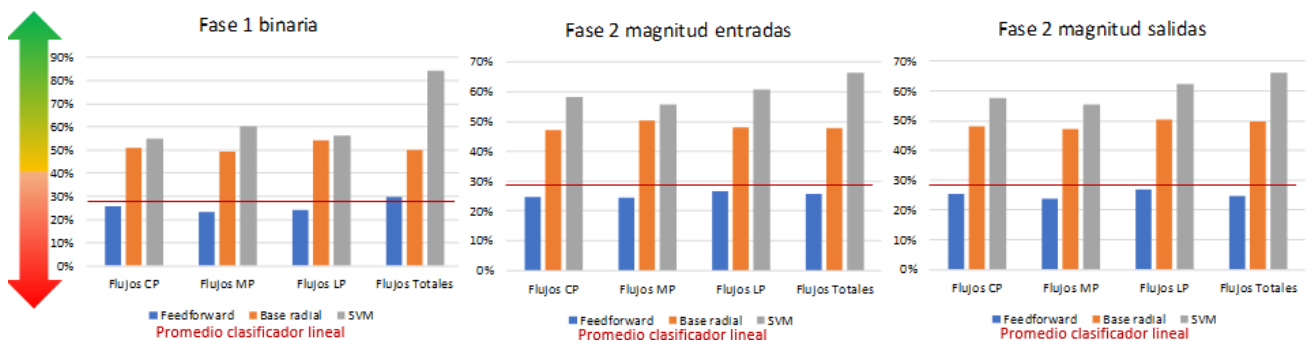


Figura 28 Comparación del porcentaje de eficiencia de los modelos de redes neuronales y los clasificadores lineales

En las gráficas de la Figura 29 se presentan los flujos acumulados de las predicciones realizadas por la arquitectura de red neuronal SVM para la fase binaria, modelo que obtuvo el mejor porcentaje de eficiencia. Por una parte, las predicciones de los flujos de corto plazo son las más volátiles en este modelo y no alcanza a reproducir la tendencia de entradas observada en la última ventana de información probada (línea naranja de la gráfica *a*) en la Figura 29), mientras que las predicciones para los flujos de mediano y largo plazo, así como para los flujos totales, la tendencia es similar, aunque para estos últimos el número de flujos de salida predichos por la red destaca por asemejarse en mayor medida a los datos observados dado que el diferencial entre los datos observados y las predicciones es menor que para los otros modelos (gráfica *d*) en la Figura 29).

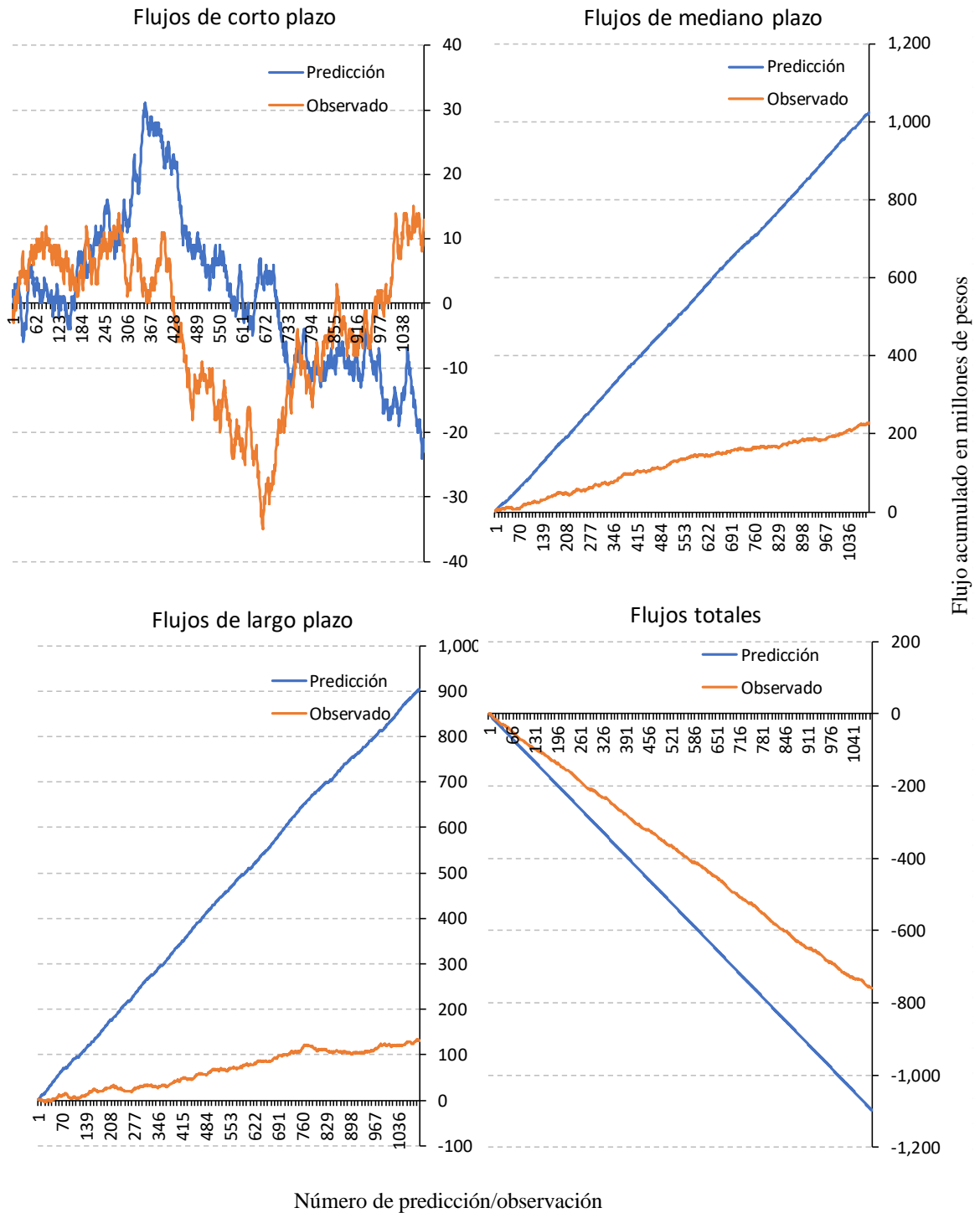


Figura 29 Flujos acumulados de la predicción de la red neuronal SVM, fase binaria

Por otra parte, las predicciones realizadas por la arquitectura de red neuronal de base radial de la fase binaria se presentan en las gráficas de la Figura 30, se puede analizar en la gráfica a) que la tendencia en los flujos de corto plazo en las primeras 600 observaciones de los datos que predijo esta red neuronal (línea morada) tienen una dirección similar a los datos reales observados (línea rosa), pero a partir de la observación 672 la tendencia cambia radicalmente indicando salidas, cuando en los datos reales se observaron entradas. Para las predicciones de los flujos de mediano y largo plazo (gráficas b) y c) de la Figura 30) la tendencia de las entradas hacia estos plazos mejora, aunque para el mediano plazo se ven predicciones de salidas mientras que los datos reales indican entradas para las primeras observaciones de la muestra, en la mitad de los datos de prueba se visualizan entradas constantes a partir de la observación 540. Finalmente, para los flujos totales se observa una tendencia similar entre los datos pronosticados y los observados pero el diferencial entre ellos se nota bastante amplio lo que indica que esta red neuronal predice salidas la mayor parte del tiempo y en la información de prueba hay varias subidas y bajadas en los flujos acumulados de entradas y salidas.

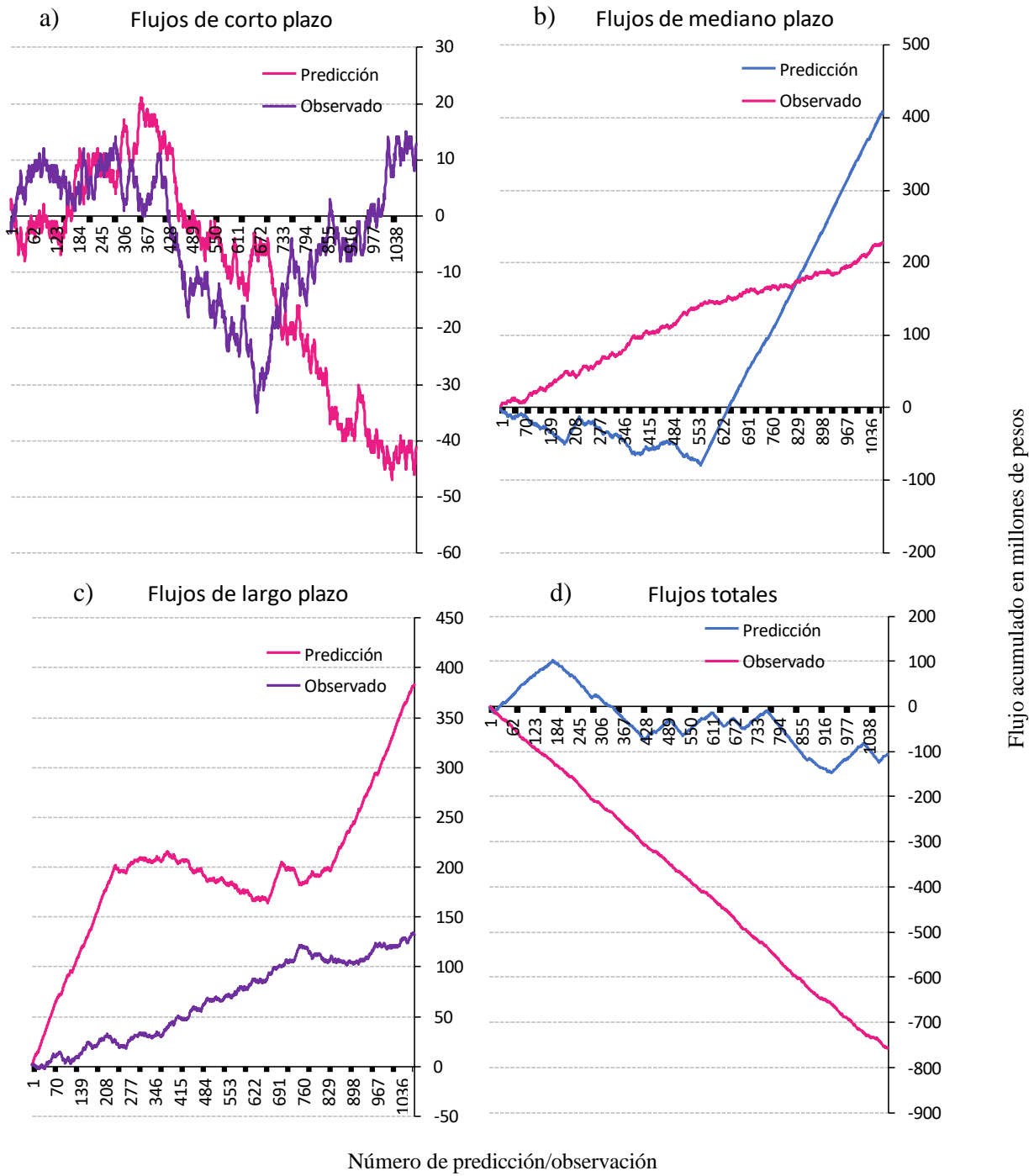
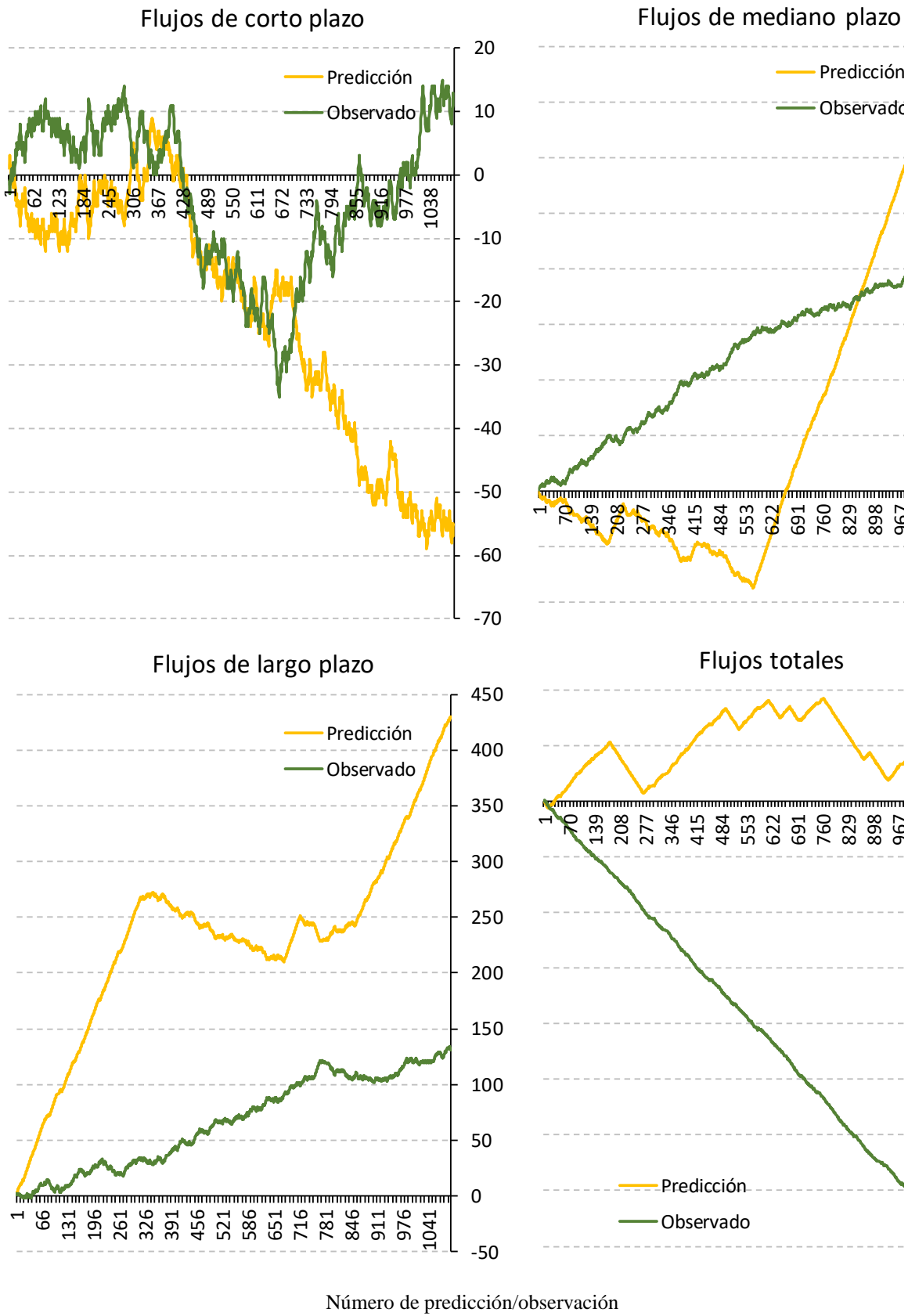


Figura 30 Flujos acumulados de la predicción de la red neuronal de base radial, fase binaria

Los resultados de la arquitectura con el menor porcentaje de eficiencia, la red neuronal *feedforward*, se muestran en la Figura 31, el diferencial entre los datos que este tipo de red neuronal predijo contra los datos reales es mayor que los resultados de las otras redes neuronales. Es interesante notar que las tendencias de las predicciones del mediano y largo plazo, así como las del flujo total para la fase binaria de esta arquitectura y la de la red neuronal de base radial lucen similares, aunque los diferenciales entre los datos observados y los predichos son más amplios para la arquitectura *feedforward*. En el Anexo 1 se puede consultar las gráficas de flujos acumulados para todas las arquitecturas y fases que se probaron.



Flujo acumulado en millones de pesos

Figura 31 Flujos acumulados de la predicción de la red neuronal feedforward, fase binaria

4.3 Discusión de resultados

Un factor sumamente importante en este tipo de modelos fue la definición de la capa de entrada, la sensibilidad en los resultados al utilizar todas las variables normalizadas, sin un tratamiento previo o, por otro lado, usar los componentes principales repercutió en una diferencia notable en los resultados obtenidos, por lo que fue significativo el diseño de esta capa en las diferentes arquitecturas de redes neuronales implementadas. Los algoritmos genéticos se han utilizado en algunas investigaciones para eliminar las variables irrelevantes en la capa de entrada de las redes neuronales. Sexton et al. los utilizaron desde los noventa para optimizar el número de parámetros de la capa de entrada en las redes neuronales y por lo tanto mejorar tiempos de entrenamiento y la eficiencia en las predicciones. Estos modelos fueron retomados más tarde por Kyoung-jae y Hyunchul en 2012 para optimizar la predicción de índices en el mercado accionario. En el caso de este trabajo de investigación, aunque el uso de componentes principales mejoró no solo el porcentaje de eficiencia de las predicciones sino también el rendimiento y tiempo de respuesta de los modelos, un análisis más complejo y exhaustivo, como el cómputo genético, podría mejorar la capacidad de predicción de las arquitecturas de redes neuronales artificiales estudiadas.

En relación con la capa de salida la técnica utilizada de “divide y vencerás” logró mejorar los resultados de predicción de las redes neuronales. El rediseño que se realizó dividiendo las pruebas de las redes neuronales en dos fases en las que, por una parte, los modelos aprendieran a predecir el comportamiento general de entradas y salidas en los flujos de inversión y la fase posterior de especialización para el pronóstico de la magnitud de las entradas y salidas dio buenos resultados y fue posible obtener un porcentaje de eficiencia mayor o igual al 60% en la mayoría de los casos. La eficiencia en la fase de especialización de la magnitud de las entradas y salidas de los flujos de inversión fue menor debido a que la red neuronal tenía más clases para predecir en la capa de resultados.

Hacia adelante la arquitectura de red neuronal de tipo SVM, que fue la que presentó mejores resultados de predicción, podría combinarse con otro enfoque para aumentar la eficiencia en la magnitud de los flujos de entrada y salida de los inversionistas extranjeros. Algunos autores

como Yolcu consideraron en sus investigaciones que es casi imposible que una serie de tiempo sea solamente lineal o puramente no lineal, por lo que propusieron un análisis simultáneo considerando ambas características en los datos y obtuvieron los mejores pronósticos de predicción que utilizar solo uno de los dos enfoques (Yolcu, et al., 2013) . Por otro lado, para Lahmiri fue de utilidad una solución híbrida, mezclando la técnica de optimización del enjambre de partículas para optimizar los pesos iniciales de la red neuronal en la predicción de tasas de interés de bonos corporativos con distintas clasificaciones crediticias en donde los porcentajes de error en los resultados de la red neuronal de tipo *feedforward* fueron menores que los otros modelos de referencia (Lahmiri, 2016). Por otro lado, Petridis et al. midió la eficiencia de fusiones y adquisiciones de empresas con relación a indicadores de negocio a través de un modelo de regresión, en una primera etapa y un modelo SVM como una segunda fase. Modelos híbridos como estos han logrado resultados satisfactorios en el campo de las aplicaciones financieras, por lo que es un campo por explorarse en trabajos futuros para el pronóstico de flujos en el mercado de renta fija.

En lo que se refiere a los resultados obtenidos en la fase de entrenamiento y de prueba éstos fueron muy similares y en promedio solo variaron 1% entre sí, lo cual es un indicador de una probabilidad baja en el nivel de *overfitting*, como lo indica (Syriopoulos, Tsatsaronis, & Karamanos, 2021). Fue importante contar con una cantidad adecuada en la serie de datos histórica, la cantidad total de observaciones fue de poco más de cuatro mil registros, desde cierto punto de vista puede parecer insuficiente pero el tomar datos con mayor antigüedad podría sesgar de alguna manera a la red neuronal debido a que las condiciones en el mercado de renta fija en México eran muy diferentes en lo que respecta a años anteriores a 2004, como se detalló en el capítulo 1, el mercado aún no era maduro en términos de liquidez, profundidad y volumen.

En cuanto a implementar los diferentes modelos de redes neuronales es evidente la necesidad de combinar un conjunto de hiper parámetros de forma óptima para obtener buenos resultados. Se pueden utilizar métodos de selección óptima de hiper parámetros como estrategias evolutivas, cotas efectivas de error de validación o de riesgo empírico. Henao propuso una técnica de estrategias evolutivas denominada adaptación de covarianzas para reducir el tiempo de convergencia de la función objetivo. Aunque el desempeño fue

competitivo con respecto a otras técnicas, la precisión decremента o el *overfitting aumenta* con respecto al número de observaciones utilizadas en el entrenamiento (Henao, 2004). En la implementación del presente trabajo de investigación se utilizaron las librerías de funciones para optimizar la configuración de los hiper parámetros para las redes neuronales de tipo SVM, en otras arquitecturas fue necesario implementar desde cero y hacer algunos experimentos de fuerza bruta para afinar los hiper parámetros. En el caso de la red neuronal SVM que obtuvo mejores resultados se trató de la que implementó un *kernel* de base radial, como menciona Ding et al. en su trabajo, el principal costo computacional de utilizar modelos de entrenamiento de tipo SVM es afinar los hiper parámetros, en su caso implementaron una modificación al *kernel* de base radial con parámetros asignados de forma aleatoria y obtuvieron una SVM más estable.

En lo que respecta a los resultados de esta investigación, los modelos de comparación fueron los clasificadores lineales. En general, los resultados obtenidos por los modelos de redes neuronales fueron mejores que los de los clasificadores lineales que intentan resolver el problema (en promedio 84% contra 28% respectivamente, ver Figura 28), aunque no fue el caso para las redes neuronales *feedforward*, cuyo porcentaje de eficiencia no alcanzó el de los clasificadores lineales (línea roja de las gráficas de la Figura 28). Trabajos como el de (Ince, et al., 2017) que combina un modelo como el VAR(*Vector Autoregressive*) con las redes neuronales *feedforward* para la predicción del precio flexible del modelo monetario para seis tipos de cambio resultó más eficiente en términos de valor económico y menor magnitud en los errores de predicción frente al modelo lineal de caminata aleatoria.

Las redes neuronales SVM lograron obtener una predicción mayor a la que se planteó en la hipótesis para los flujos totales en la fase binaria. Es posible a partir de los resultados obtenidos plantearse preguntas respecto a lo que pasará con el mercado de renta fija mexicano ante el contexto actual de inflaciones altas, políticas monetarias restrictivas, conflictos geopolíticos y los estragos de la pandemia por COVID-19, en la capacidad de los sectores locales para absorber las posibles salidas de inversionistas extranjeros y la afectación en la curva nominal nacional. Por un lado, en la investigación de (Zhang, et al., 2019) de la predicción del riesgo de crédito de varias empresas Pymes utilizando una red neuronal SVM

con un algoritmo del movimiento de las luciérnagas para optimizar los parámetros, se mejora el porcentaje de error en las predicciones respecto al método de descomposición LIBSVM en más del 3%, por el lado de los resultados en el porcentaje de error cuadrático medio de las redes neuronales SVM implementadas en este trabajo, éste es menor en 7%, en promedio, con respecto a las otras arquitecturas de redes neuronales con las que se experimentaron y en porcentaje de eficiencia superó en más de 15% a los resultados de los clasificadores lineales.

Por otra parte, se esperaba obtener un nivel de eficiencia arriba del 60% para la predicción de la magnitud de los flujos de entrada o salida y este porcentaje solo se logró con los flujos de entrada de largo plazo, los flujos de salida de largo plazo y los flujos totales en entradas y salidas. Ahora bien, la predicción de la magnitud de flujos de entradas y salidas es importante para dimensionar el impacto de los flujos de estos inversionistas y tomar acciones al respecto, por lo que sería de utilidad tener un nivel de confiabilidad arriba del 60% para la predicción en la magnitud de los flujos de salida que se esperan dada la aversión al riesgo ante contextos de alta incertidumbre, como el que se vive en 2022, y poder aplicar operaciones extraordinarias en el mercado de dinero o de deuda de forma oportuna y con un alto nivel de confiabilidad predicho por los modelos.

Con base a los resultados obtenidos es recomendable tener cierta cautela para tomar acciones preventivas en el mercado de renta fija mexicano solo basadas en el modelo de redes neuronales, sobre todo en el nivel de la magnitud de dichos flujos. En complemento con otras herramientas o análisis de predicción, este modelo puede ser útil para predecir tendencias en el futuro. Podría ser necesario un modelo que indique la proporción en la relación de las variables de entrada con los flujos de inversionistas extranjeros y que por lo tanto mejore la eficiencia en la predicción del semáforo de entradas o salidas. El trabajo en conjunto de un modelo lineal y uno no lineal de tipo red neuronal puede contribuir a la mejora de los resultados en las predicciones, de acuerdo con la comparación de resultados realizada en este apartado.

5 Conclusiones y trabajo futuro

Como se describió al principio de este trabajo, un modelo no lineal puede resultar de poca utilidad para ajustar datos complejos en donde la relación entre éstos y la variable independiente es a simple vista débil o inexistente, un modelo de red neuronal puede aprender esa relación entre las variables a través del proceso de entrenamiento, pero también implica gestionar más puntos de control y configuración que un modelo lineal. Cada cambio en los hiper parámetros de la red neuronal afecta la sensibilidad en los resultados de las predicciones, en varios experimentos ésta cambiaba radicalmente, por lo que se debe tener cuidado y dedicación en su asignación. Es por esto por lo que para trabajos futuros surge la necesidad experimentar y encontrar el modelo adecuado para localizar los valores óptimos que maximicen la eficiencia de la red neuronal.

Por una parte, para el número de variables de entrada contar con la cantidad idónea que alcance a representar el problema completo y tenga una ponderación significativa en relación con la variable resultado, para futuras investigaciones se podrían considerar otras variables de mercado que esta investigación pudo haber omitido, por ejemplo, medidas para el nivel de intervencionismo en los mercados financieros, el riesgo político en el país o el nivel de correlación entre el mercado local y las consecuencias de las sanciones económicas derivadas del conflicto geopolítico entre Rusia y Ucrania; variables que puedan ser incorporadas de forma dinámica de acuerdo al contexto económico que se esté viviendo y que puedan tener una afectación en el comportamiento de los mercados . El tratamiento de esta capa de entrada es un punto que considerar para que la red neuronal reciba los parámetros más representativos del problema y en la dimensionalidad adecuada para que un elemento no tenga una ponderación mayor que pueda sesgar de alguna forma el comportamiento del aprendizaje de la arquitectura. En este trabajo los componentes principales utilizados para la capa de entrada fueron de gran utilidad para evitar transmitirle a las redes neuronales algún tipo de sesgo en la ponderación de los parámetros de entrada, sin embargo, como se sabe, su cálculo es muy sensible a los datos *outliers* y si estas observaciones anormales no se detectan en la serie histórica, es posible que desde la capa de entrada de la red neuronal ya se tengan sesgos indetectables en el procesamiento posterior. Por otra parte, la reducción en la

dimensionalidad de los parámetros de entrada de la red neuronal (de 31 variables a 11 componentes principales) aportó eficiencia en los resultados de las predicciones y en el tiempo de procesamiento de las arquitecturas, el desempeño del grupo de experimentos realizado con estos elementos fue visiblemente mejor que utilizar el número de parámetros completo en esta capa de inicio en los modelos de redes neuronales.

Aunado a lo anterior, también se debe analizar y optimizar el conjunto de hiper parámetros propio de cada tipo de red neuronal ya que es en este punto es donde la resolución del problema puede verse afectada, por ejemplo, por no poder “escapar” de un mínimo o máximo local y por lo tanto la eficiencia de clasificación del modelo no pueda mejorar, aunque el número de iteraciones se incremente. La combinación óptima en la configuración de estos hiper parámetros es otro punto de inflexión en el desempeño de este tipo de modelos no lineales. El trabajo hacia adelante en este sentido consiste en probar algoritmos de optimización en la combinación de estos hiper parámetros para mejorar la eficiencia de los modelos en la fase de prueba y validación. Aunado a lo anterior, el nivel de *overfitting* es también un tema relevante para que el modelo pueda funcionar con eficiencia con datos reales por lo que también sería necesario explorarlo a detalle en el futuro, de manera general no se observó este problema en las etapas de prueba y validación, sin embargo el reto es validar estos modelos de redes neuronales diseñados en este trabajo de investigación con situaciones con incertidumbre como el que se presenta en el primer trimestre de 2022; con el inicio del ciclo restrictivo de la Reserva Federal, escenarios de inflaciones elevadas y rebasadas en varias regiones del mundo y las consecuencias del conflicto geopolítico entre rusos y ucranianos.

Ahora bien, como se observó en los resultados de los modelos implementados, la predicción de flujos de bonos corto plazo fue limitada, lo que puede deberse al efecto de la reacción de los inversionistas ante eventos de aversión al riesgo, apretamiento de las condiciones financieras globales e implementación de ciclos de política monetaria restrictivas en economías avanzadas y emergentes, en las que la parte corta de la curva de inversión recibe un mayor impacto y por lo tanto puede tener una volatilidad más elevada que el mediano y largo plazo. Por el contrario, los mejores resultados obtenidos fueron los que corresponden a

los flujos totales sin división por tipo de plazo. Esto puede deberse a que el flujo total lleva implícito el comportamiento de los dos tipos de inversionistas: los activos que buscan entrar al corto plazo para obtener rendimientos atractivos y también el inversionista pasivo, que como se estudió en la sección 1.4, los índices de referencia juegan un papel importante en el interés de compra de los inversionistas extranjeros en los bonos soberanos que se encuentran dentro de éstos.

En cuanto al uso de este tipo de herramienta para un banquero central que tiene un punto de vista más tradicional, las redes neuronales podrían ser percibidas como una caja negra que realiza un proceso de aprendizaje oscuro, por lo que puede generar hasta cierto punto desconfianza en los resultados que se obtienen y resistencia a tomar decisiones con base a este tipo de modelos. De modo que puede desarrollarse, como parte del trabajo futuro, un análisis econométrico para evaluar y validar los resultados de las arquitecturas de redes neuronales implementadas que permitan probar su importancia en un contexto dado. Además, pueden implementarse modelos de introspección (también conocidos como forenses) que ayuden a comprender el funcionamiento de los modelos y las relaciones de las variables de entrada y salida, herramientas que apoyen a la obtención de conclusiones más sólidas. Varios bancos centrales en el mundo ya han empezado a utilizar herramientas de *big data* o algoritmos de aprendizaje de máquina (*machine learning*) para cumplir con eficiencia sus funciones de supervisión y regulación de los mercados. Este trabajo de investigación deja como precedente el uso de las arquitecturas de redes neuronales artificiales como herramienta de predicción de flujos de inversión en el mercado de renta fija mexicano y la investigación y profundización en los temas para el trabajo futuro pueden ser explorados para la implementación de estos modelos como parte de las herramientas utilizadas en la toma de decisiones de un banco central para la prevención en el deterioro de las condiciones de operación de los mercados o para el mejoramiento del desarrollo de los mercados financieros del país.

Para concluir, este acercamiento a las redes neuronales artificiales en la predicción de flujos de bonos gubernamentales por parte de inversionistas extranjeros es de utilidad para el análisis de tendencias en las entradas y salidas totales. Hacia adelante los complementos más

importantes que se puede tener en esta línea de investigación son, por una parte, aumentar el porcentaje de eficiencia para conocer la magnitud de esos flujos de entrada y de salida, a través de la combinación de un modelo no lineal y la arquitectura de redes neuronales, por ejemplo, y, por otra parte, contar con modelos de validación o de introspección que validen los resultados obtenidos y que sustenten la toma de decisiones en la implementación de medidas preventivas que sean necesarias para atenuar las implicaciones negativas en las condiciones de operación del mercado de bonos en México.

Referencias

- Abreu, G.G.B., Acosta, A.M.R., Álvarez, T.C., Cortina, M.J. J., Gallardo, G.M. D., García, P.J. R., Jiménez, V.L., Tapia, R.C. & Tegho V.M.S. (2014). *El mercado de valores gubernamentales en México* (pp. 17-128). Banco de México.
http://educa.banxico.org.mx/ebooks_descargas/%7BCD2C4B20-74C9-6BE6-14CA-CAD02563FED9%7D.pdf
- Amihud, Y. & Mendelson, H. (1991). *Liquidity, asset prices and financial policy*. Financial Analysis Journal, 47(6), 58-59. <https://doi:10.2469/faj.v47.n6.56>
- Banco de México & Secretaría de Hacienda y Crédito Público. (s.f.). *Descripción técnica de los bonos de desarrollo del gobierno federal con tasa de interés fija*. Notas técnicas y títulos múltiples de valores gubernamentales y del IPAB.
<https://www.banxico.org.mx/mercados/d/%7B77F7E7D0-8EDF-B4C1-406A-32968005EBC9%7D.pdf>
- Banco de México & Secretaría de Hacienda y Crédito Público. (s.f.). *Descripción técnica de los certificados de la tesorería de la federación*. Notas técnicas y títulos múltiples de valores gubernamentales y del IPAB.
<https://www.banxico.org.mx/mercados/d/%7B0DE0044F-662D-09D2-C8B3-4F1A8E43655F%7D.pdf>
- Banco de México. (s.f.). *Reservas internacionales*. Banxico educa.
http://educa.banxico.org.mx/banco_mexico_banca_central/reservas-internacionales.html
- Banco de México & Secretaría de Hacienda y Crédito Público. (1996). *Descripción técnica de los Udibonos bonos de desarrollo del gobierno federal denominados en unidades de inversión*. Notas técnicas y títulos múltiples de valores gubernamentales y del IPAB. <https://www.banxico.org.mx/mercados/d/%7B52319AD4-4B78-6F95-E313-7AC67498B728%7D.pdf>
- Banco de México. (16 de junio de 2021). *Reporte de estabilidad financiera Junio 2021*. Reportes sobre el sistema financiero. Recuperado el 2 de diciembre de 2021 de <https://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-el-sistema-financiero/%7BFBA9B09-D285-B1CC-33A8-84E35347B970%7D.pdf>

- Banco de México. (28 de julio de 2021). *Operación de valores gubernamentales y del IPAB en directo vía casas de corretaje*. Operaciones en el mercado secundario de valores. Recuperado el 28 de julio de 2021 de https://www.banxico.org.mx/mercsecd/index.html?tabla=N&locale=es_ES
- Banco de México. (28 de julio de 2021). *Tenencia de residentes en el extranjero (Saldo a valor nominal en millones de pesos)*. Sistema de Información Económica. Recuperado el 28 de julio de 2021 de <https://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?sector=7&accion=consultarCuadroAnalitico&idCuadro=CA175&locale=es>
- Banco de México. (30 de noviembre de 2021). *Aplicación de Consultas de la Posición*. Recuperado el 30 de noviembre de 2021 de https://www.banxico.org.mx/valores/PresentaDetallePosicionGub.faces?BMXC_instrumento=2&BMXC_lang=es_MX
- Banco de México & Secretaría de Hacienda y Crédito Público. (2021). *Descripción técnica Bonos de Desarrollo del Gobierno Federal BONDES F*. Notas técnicas y títulos múltiples de valores gubernamentales y del IPAB. <https://www.banxico.org.mx/mercados/d/%7B997516C8-3B76-53A7-75CC-06C4D39F7584%7D.pdf>
- Banco Mundial & Fondo Monetario Internacional. (2001). *Developing Government Bond Markets a handbook (pp. 25-28)*. Autoedición. <https://openknowledge.worldbank.org/handle/10986/13865>
- Calvo, G., Leiderman, L. & Reinhart, C. (1993). *Capital Inflows and Real Exchange Rate Appreciation in Latin America: The Role of External Factors*. IMF Staff Papers, 40(1), 108-112. <https://www.imf.org/en/Publications/WP/Issues/2016/12/30/Capital-inflows-and-Real-Exchange-Rate-Appreciation-in-Latin-America-The-Role-of-External-1824>
- Cao, L. J. & Tay, F. E. H. (2000). *Feature Selection for Support Vector Machines in Financial Time Series Forecasting*. Intelligent Data Engineering and Automated Learning - IDEAL 2000, 1983(1), 268-273. https://doi.org/10.1007/3-540-44491-2_38
- Deng, Y. & Gao, Q. (2020). *A study on e-commerce customer segmentation management*

- based on improved K-means algorithm*. Information Systems and e-Business Management, 18(4), 497-510. <https://doi.org/10.1007/s10257-018-0381-3>
- Ding, X., Liu J., Yang F., Cao J. (2021). *Random radial basis function kernel-based support vector machine*. Journal of the Franklin Institute, 358(s.n.), 10121-10140. <https://doi.org/10.1016/j.jfranklin.2021.10.005>
- Duranton, G. & Overman, H. (2005). *Testing for localization using micro-geographic data*. The Review of Economic Studies, 72(4), 1077-1106. <https://doi.org/10.1111/0034-6527.00362>
- Eguren-Martin, F., O'Neill, C., Sokol, A., & Berge, L. (2020). *Capital flows-at risk: push, pull and the role of policy*. Staff Working Paper Bank of England, s.n., 3-17. <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2020/capital-flows-at-risk-push-pull-and-the-role-of-policy.pdf?la=en&hash=ADDCF2EDF04478194A02698A31DC35ABB8265507>
- García, G. I., & Stracca, L. (2021). *Changing patterns of capital flows*. Committee on the Global Financial System, 66 (s.n.), 1-4. Bank for International Settlements. <https://www.bis.org/publ/cgfs66.pdf>
- Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). *Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis*. Journal of Behavioral and Experimental Finance, 32 (s.n.), 5-10. <https://doi.org/10.1016/j.jbef.2021.100577>
- Haykin, S. (1998). *Neural networks. A comprehensive foundation*. (2^a ed., pp. 278-347). Pearson Prentice Hall.
- Henao, R. (2004). *Selección de hiperparámetros en máquinas de soporte vectorial*. [Tesis de maestría]. Universidad Nacional de Colombia. Repositorio institucional - Universidad Nacional de Colombia. <https://repositorio.unal.edu.co/bitstream/handle/unal/69900/ricardohenao.2004.pdf?sequence=1>
- Hsu, C.W., Chang, C.C., & Lin C.J. (2016). *A Practical Guide to Support Vector Classification*. Department of Computer Science National Taiwan University, s.n., 2-5. <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

- Ince, H., Cebeci, A. F. & Imamoglu, S. Z. (2017). *An Artificial Neural Network Based Approach to the Monetary Model of Exchange Rate*. Computational Economics, 53 (s.n.), 817-831. <https://doi.org/10.1007/s10614-017-9765-6>
- International Organization of Securities Commissions. (2017). *Examination of Liquidity of the Secondary Corporate Bond Markets*. The International Organization of Securities Commissions.
<https://www.iosco.org/library/pubdocs/pdf/IOSCOPD558.pdf>
- Jeanneau, S., & Pérez, V. C. (2005). *Reducing financial vulnerability: the development of the domestic government bond market in Mexico: BIS Quarterly Review* (diciembre 2005). Bank for International Settlements.
https://www.bis.org/publ/qtrpdf/r_qt0512h.pdf
- Kowalczyk, A. (2017). *Support vector machines Succinctly*. Syncfusion. (pp. 23-25).
<https://www.syncfusion.com>
- Kyoung-jae, K. & Hyunchul, A. (2012). *Simultaneous optimization of artificial neural networks for financial forecasting*. Applied Intelligence - APIN. 36(s.n.) 1-12.
<https://doi.org/10.1007/s10489-011-0303-2>.
- Lahmiri, S. (2016). *Interest rate next-day variation prediction based on hybrid feedforward neural network, particle swarm optimization, and multiresolution techniques*. Physica A: Statistical Mechanics and its Applications, 444(s.n.), 386-396.
<https://dx.doi.org/10.1016/j.physa.2015.09.061>
- Martyniuk, V., & Bis, J. (2021). *Artificial Neural Networks with Radial-Basis Transmission Functions for Modelling Economic Phenomena and Processes*. European Research Studies Journal, 24(s.n.), 168-178. <https://doi:10.35808/ersj/2215>
- Norimasa, Y., Kazuki, U., & Watanabe, T. (2021). *Emerging Economies' Vulnerability to Changes in Capital Flows: The Role of Global and Local Factors: Bank of Japan Working Paper Series* (Informe n° 21). Bank of Japan.
https://www.boj.or.jp/en/research/wps_rev/wps_2021/data/wp21e05.pdf
- Partridge, M.D., Rickman, D. S., Ali, K., & Olfert, R. (2008). *Lost in space: population growth in the American hinterlands and small cities*. Journal of Economic Geography, 8(6), 727-757. <https://doi:10.1093/jeg/lbn038>
- Peiris, S. J. (2010). *Foreign participation in emerging markets' local currency bond*

- markets*. IMF Working Paper, 10(88), 4-10.
<https://www.imf.org/external/pubs/ft/wp/2010/wp1088.pdf>
- Petridis, K., Tampakoudis I., Drogalas G. (2022). *A Support Vector Machine model for classification of efficiency: An application to M&A*. International Business and Finance, 61 (101633), 3-18. <https://doi.org/10.1016/j.ribaf.2022.101633>
- Pilbeam, K., & Langeland, K. N. (2015). *Forecasting exchange rate volatility: GARCH models versus implied volatility forecasts*. International Economics & Economic Policy, 12(s.n.), 127-142. <https://doi.org/10.1007/s10368-014-0289-4>
- Qiao, G., Yang, J., & Li, W. (2020). *VIX forecasting based on GARCH-type model with observable dynamic jumps: A new perspective*. North American Journal of Economics and Finance, 53(s.n.), 3-5. <https://doi.org/10.1016/j.najef.2020.101186>
- Rodrigo, J. A. (21 de enero de 2022). *Análisis de componentes principales PCA con Python*. Ciencia de Datos, Estadística, Machine Learning y Programación. <https://www.cienciadedatos.net/documentos/py19-pca-python.htm>
- Syriopoulos, T., Tsatsaronis, M., & Karamanos, I. (2021). *Support Vector Machine Algorithms: An Application to Ship Price Forecasting*. Computational Economics, 57(s.n.), 55-87. <https://doi.org/10.1007/s10614-020-10032-2>
- Trade Association for the Emerging Markets. (2021). *EMTA Survey: First quarter emerging markets debt trading at US\$1.372 trillion (1^{er} trimestre 2021)*. Auto edición. <https://www.emta.org/>
- Valenzuela, P. J. I. (2017). *Predicción de series financieras del mercado latinoamericano mediante redes neuronales artificiales, a través de un algoritmo de colonias de abejas*. [Tesis de licenciatura]. Universidad Técnica Federico Santa María. Repositorio Institucional - Universidad Técnica Federico Santa María. <https://repositorio.usm.cl/handle/11673/23698>
- Yolcu, U., Egrioglu E. & Aladag, C. H. (2013). *A new linear & nonlinear artificial neural network model for time series forecasting*. Decision Support Systems, 54(s.n.), 1340-1347. <http://dx.doi.org/10.1016/j.dss.2012.12.006>
- Zhang, H., Shi, Y., Yang, X., & Zhou, R. (2019). *A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance*. Research in International Business and Finance, 58(s.n.), 1-8.

<https://doi.org/10.1016/j.ribaf.2021.101482>

Índice de figuras

Figura 1 Métodos de pronóstico, basado en (Valenzuela Pavez, 2017), p. 5	7
Figura 2 Porcentaje de tenencia de bonos gubernamentales con respecto al total colocado por parte de inversionistas extranjeros. Fuente: Banco de México.	16
Figura 3 Neurona artificial. Basado en (Ince, Fehim Cebeci, & Zeki Imamoglu, 2017) p.631	
Figura 4 Tipos de capas en una red neuronal.....	32
Figura 5 Campana de Gauss	35
Figura 6 Red neuronal de base radial. Basado en (Martyniuk & Bis, 2021), p. 172	36
Figura 7 Hiperplano óptimo para clases separables linealmente. Basado en (Haykin, 1998), p. 342	39
Figura 8 Separación de datos en espacios no bidimensionales. Basado en (Kowalczyk, 2017) , p.23	40
Figura 9 Kernel lineal	41
Figura 10 Kernel polinómico	41
Figura 11 Kernel gaussiano	41
Figura 12 Patrones clasificables por concepto de proximidad.....	43
Figura 13 Ejemplo Asignación de centroides k-means.....	45
Figura 14 Diagrama de flujo general de la metodología.....	47
Figura 15 Flujos de entradas y salidas por plazo con dos clases (Número de registros de la muestra). Elaboración propia con datos de las tenencias de bonos gubernamentales en manos de extranjeros, obtenidos del Banco de México	54
Figura 16 Tasas de rendimiento de los bonos gubernamentales por plazo con eventos seleccionados	57
Figura 17 Índice de aversión al riesgo con dato erróneo encontrado.....	57
Figura 18 Dimensionalidad PCA.....	61
Figura 19 Porcentaje de varianza explicada por componente principal	62
Figura 20 Porcentaje de varianza explicada acumulada.....	62
Figura 21 Diseño de red neuronal feedforward	64
Figura 22 Diseño de red neuronal de base radial.....	66
Figura 23 Modelo de red neuronal SVM.....	68
Figura 24 Convergencia de la tasa de error en la red neuronal feedforward	80
Figura 25 Convergencia de la tasa de error en la red neuronal de base radial.....	82
Figura 26 Convergencia de la tasa de error en la red neuronal SVM.....	84
Figura 27 Flujos acumulados de corto plazo resultados de la red neuronal SVM.....	88
Figura 28 Comparación del porcentaje de eficiencia de los modelos de redes neuronales y los clasificadores lineales	90
Figura 29 Flujos acumulados de la predicción de la red neuronal SVM, fase binaria.....	92
Figura 30 Flujos acumulados de la predicción de la red neuronal de base radial, fase binaria	94
Figura 31 Flujos acumulados de la predicción de la red neuronal feedforward, fase binaria	96

Figura 32 Flujos acumulados de las predicciones de la red neuronal SVM para para la fase 2 de magnitud de entradas	116
Figura 33 Flujos acumulados de las predicciones de la red neuronal SVM para la fase 2 de magnitud de salidas.....	117
Figura 34 Flujos acumulados de las predicciones de la red neuronal feedforward para la fase 2 de magnitud de entradas	118
Figura 35 Flujos acumulados de las predicciones de la red neuronal feedforward para la fase 2 de magnitud de salidas.....	119
Figura 36 Mapas de calor de los flujos predichos y reales de la red neuronal SVM.....	121

Índice de ecuaciones

Ecuación 1 Cálculo carry to risk.....	20
Ecuación 2 Cálculo de razón de cobertura de State Street	26
Ecuación 3 Función escalón	31
Ecuación 4 Ponderación de variables de entrada.....	31
Ecuación 5 Función sigmoide.....	31
Ecuación 6 Función tanh.....	32
Ecuación 7 Error en la red neuronal	33
Ecuación 8 Sumatoria de errores al cuadrado.....	34
Ecuación 9 Error cuadrático medio	34
Ecuación 10 Función gaussiana.....	36
Ecuación 11 Distancia euclidiana	36
Ecuación 12 Salida red neuronal por la regla Delta.....	37
Ecuación 13 Error de los pesos.....	38
Ecuación 14 Error global	38
Ecuación 15 Ecuación del hiperplano en un espacio bidimensional.....	38
Ecuación 16 Hiperplano óptimo	39
Ecuación 17 Distancia euclidiana	44
Ecuación 18 Distancia Mahalanobis	44
Ecuación 19 Teorema de Bayes.....	45
Ecuación 20 Cálculo de flujos de bonos gubernamentales	51
Ecuación 21 Interpolación lineal	58
Ecuación 22 Estandarización	59
Ecuación 23 Media	60
Ecuación 24 Desviación estándar	60
Ecuación 25 Ponderación de variable de entrada 1	63
Ecuación 26 Cálculo del porcentaje de eficiencia de los clasificadores lineales.....	77

Índice de tablas

Tabla 1 Resumen de variables para la capa de entrada.....	49
Tabla 2 Semáforo de flujos especializado en entradas y salidas.....	55
Tabla 3 Modelos de red neuronal feedforward	65
Tabla 4 Modelos de red neuronal de base radial.....	67
Tabla 5 Modelos de red neuronal SVM.....	69
Tabla 6 Conjuntos de datos para la red neuronal	70
Tabla 7 Configuración de la capa de entrada y de salida.....	76
Tabla 8 Resultados del porcentaje de eficiencia de los clasificadores lineales.....	78
Tabla 9 Configuración de hiper parámetros de la red neuronal general feedforward.....	80
Tabla 10 Configuración de hiper parámetros red neuronal de base radial	81
Tabla 11 Configuración de hiper parámetros de la red neuronal SVM	83
Tabla 12 Resultado del error cuadrático medio de las arquitecturas de redes neuronales	85
Tabla 13 Ejemplo de la generación de flujos acumulados para entradas de corto plazo.....	87
Tabla 14 Porcentaje de eficiencia de las arquitecturas y fases de redes neuronales	90

Anexos

Anexo 1 Gráficas de los flujos acumulados predichos por las redes neuronales

El siguiente conjunto de gráficas detalla los resultados de todos los experimentos realizados para la fase dos del semáforo de salida y para todos los tipos de red neuronal estudiados. En la sección 4.1.6 Validación de experimentos se describe el procedimiento para realizar las gráficas de los flujos acumulados de las predicciones y de las observaciones reales. En la Figura 32 para la red neuronal SVM se puede observar que también para la fase dos del semáforo de salida, los pronósticos son muy parecidos a las observaciones reales, aunque para el mediano plazo el diferencial entre ambas es más grande que para el corto, largo plazo y los flujos totales de entradas. Por otra parte, en las gráficas de la Figura 33 para los flujos de salidas de largo plazo los pronósticos de este tipo de red presentan el mayor diferencial con respecto a los datos observados.

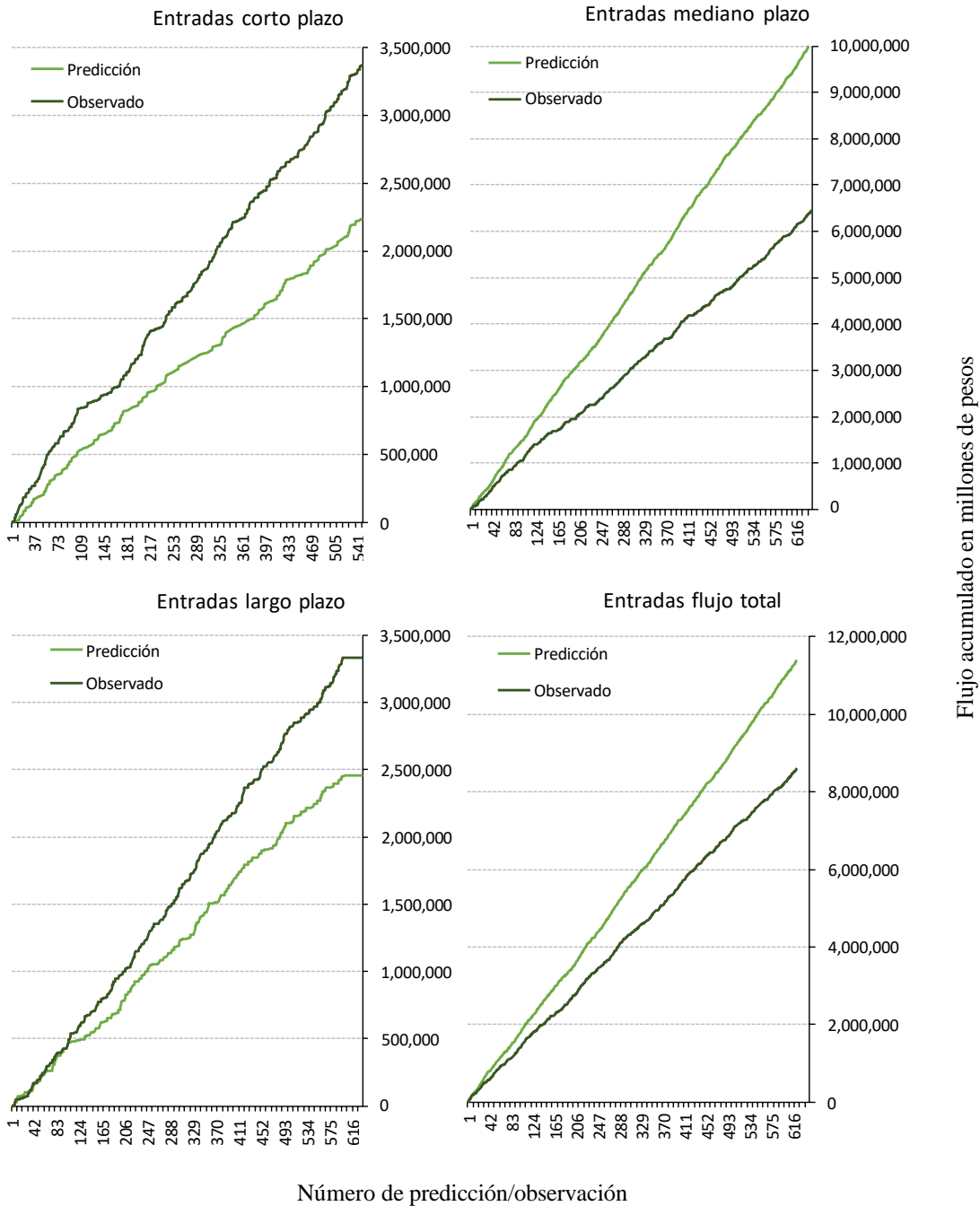


Figura 32 Flujos acumulados de las predicciones de la red neuronal SVM para para la fase 2 de magnitud de entradas

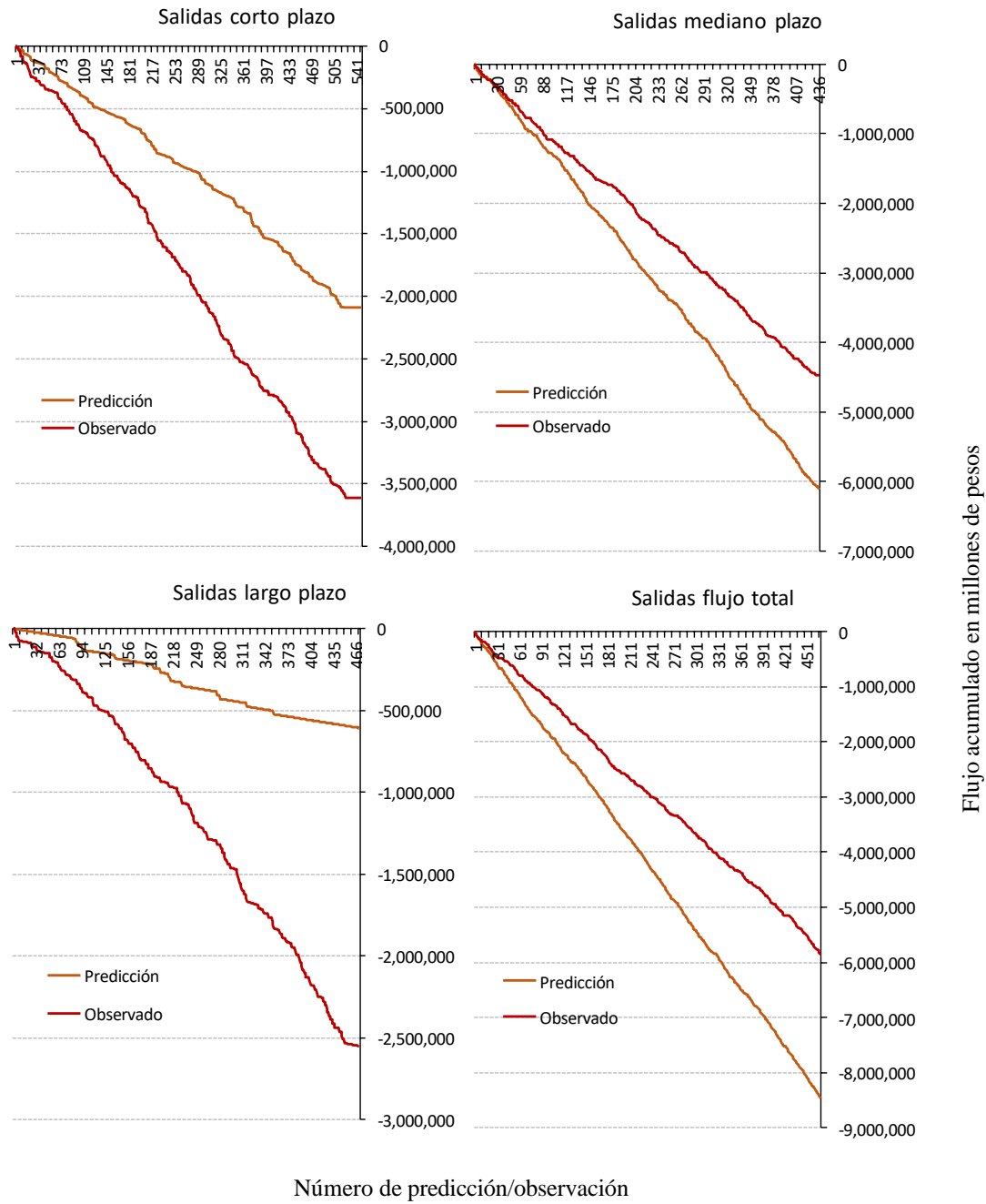


Figura 33 Flujos acumulados de las predicciones de la red neuronal SVM para la fase 2 de magnitud de salidas

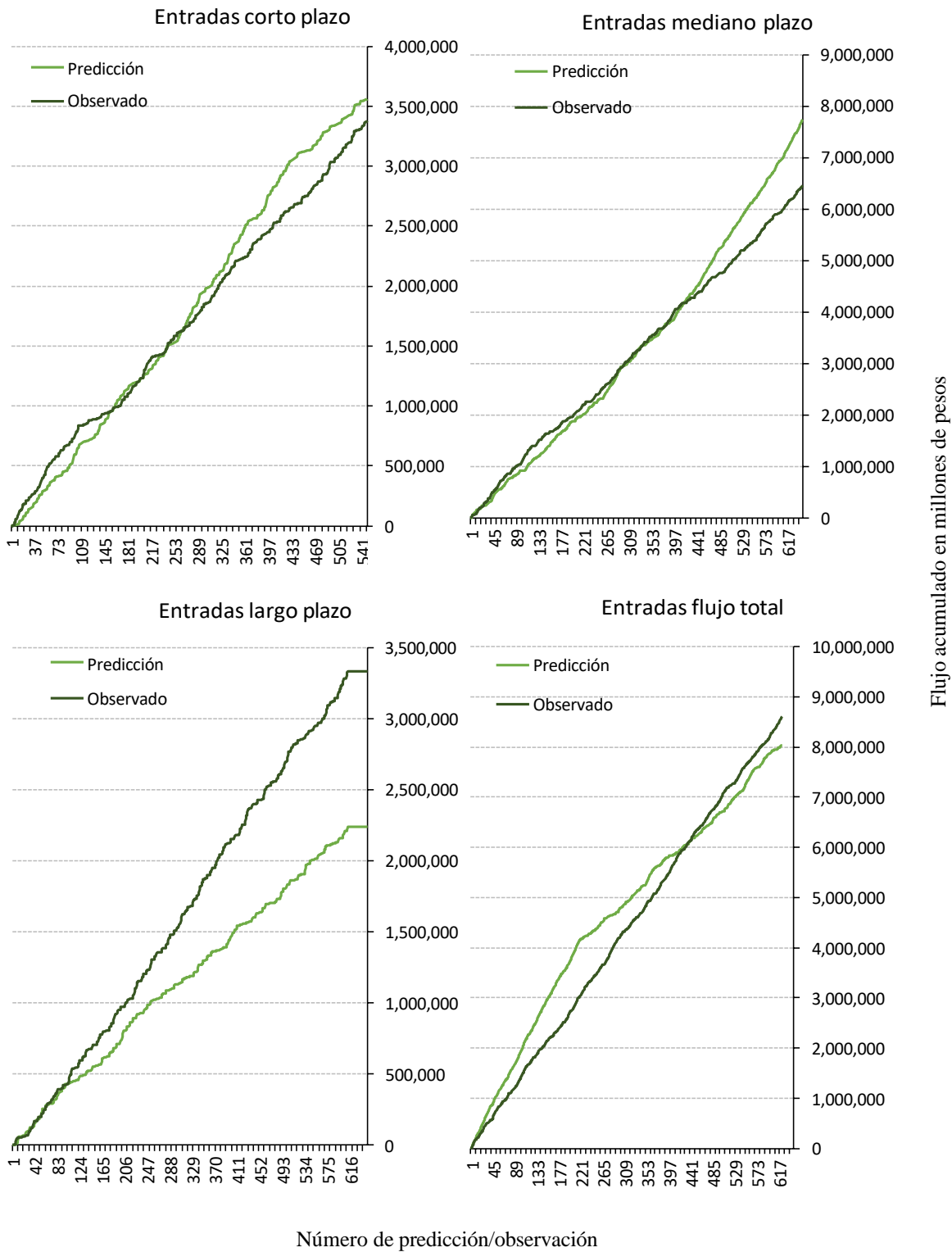


Figura 34 Flujos acumulados de las predicciones de la red neuronal feedforward para la fase 2 de magnitud de entradas

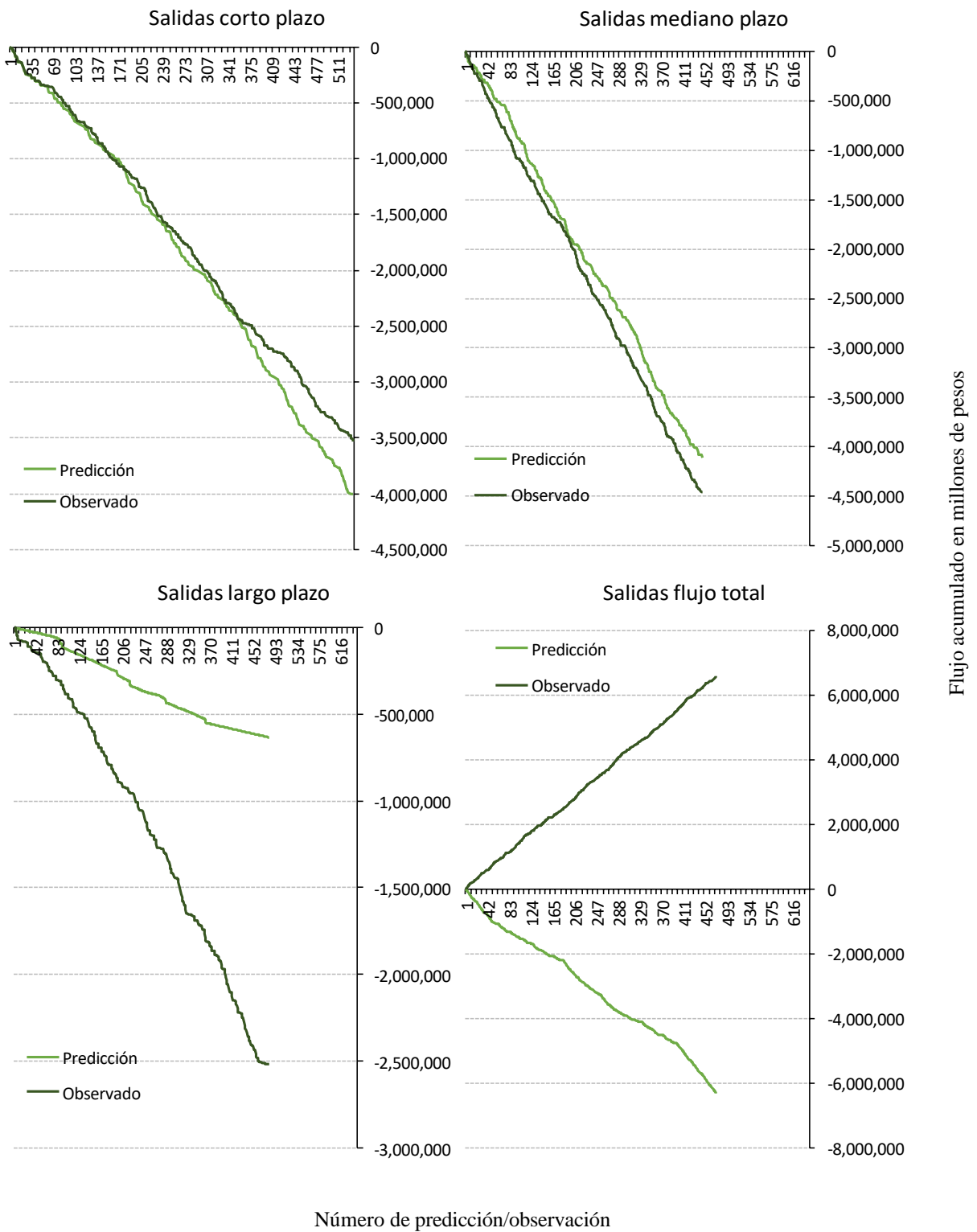


Figura 35 Flujos acumulados de las predicciones de la red neuronal feedforward para la fase 2 de magnitud de salidas

Anexo 2 Mapas de calor de flujos predichos y reales de la red neuronal SVM

Otra manera de presentar los resultados de las redes neuronales es con mapas de calor para identificar los cambios entre los datos observados y los que predijo cada tipo de red. El color verde significa que la observación fue una entrada, el amarillo que fue un flujo neutro y el rojo que es un flujo de salida. Con el objetivo de encontrar un patrón visual entre los datos que se fueron predichos por la red neuronal (primera columna de la Figura 36) comparándolos con los datos observados, se generó este tipo de gráfica para los resultados de la red neuronal SVM. Como se puede observar en la gráfica de mapa de color, no es sencillo identificar algún patrón, sin embargo, se considera que es un complemento a las gráficas de flujos acumulados mostradas en la sección 4 de análisis de resultados.

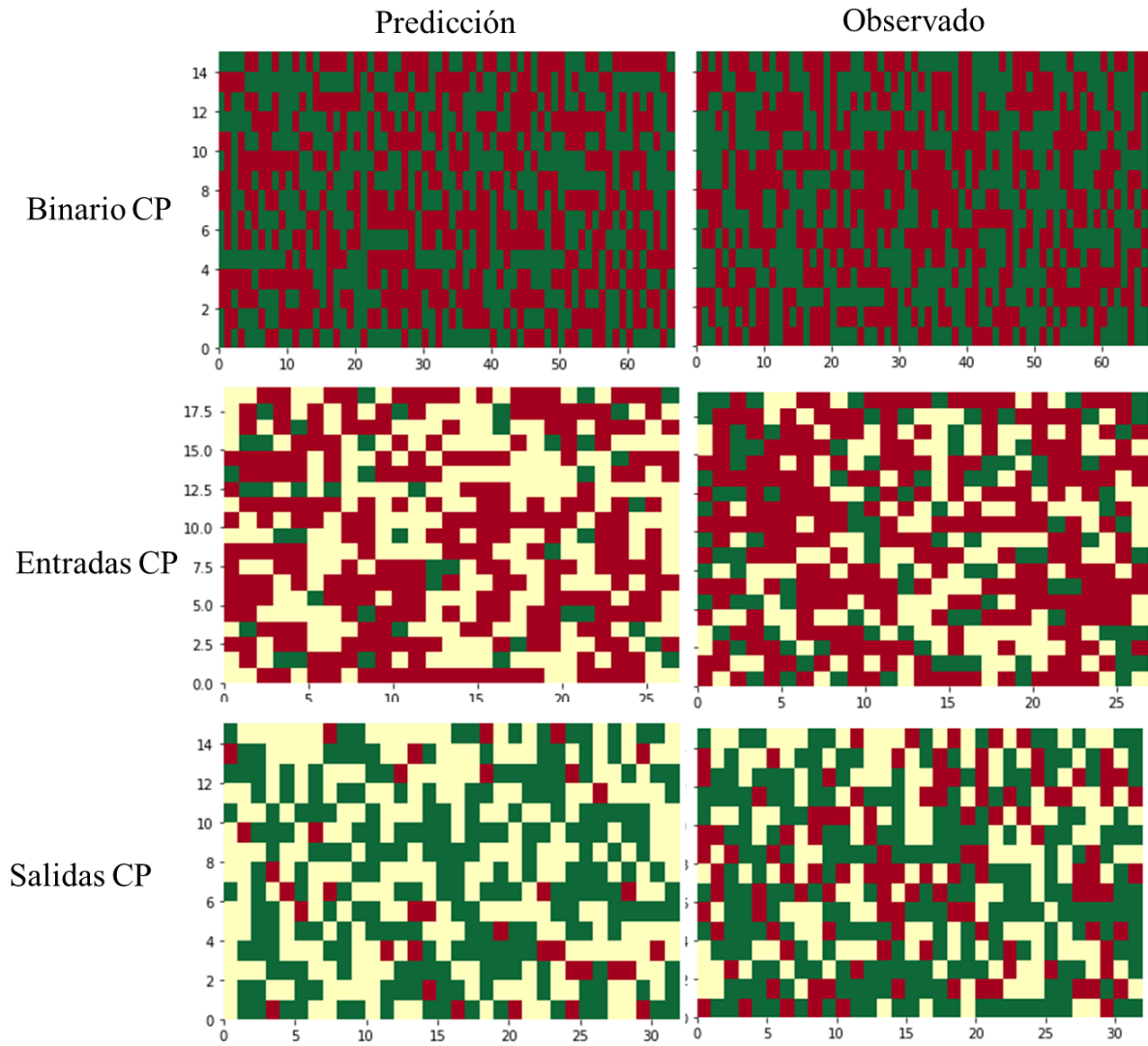


Figura 36 Mapas de calor de los flujos predichos y reales de la red neuronal SVM