

Pontificia Universidad Javeriana Cali
Facultad de Ingeniería.
Ingeniería de Sistemas y Computación.
Anteproyecto de Grado.

Técnicas de Aprendizaje Automático Aplicadas a la Predicción del Riesgo de Inversión en Criptomonedas

Ana María García Posso
Andrés Felipe Delgado Chamorro

Director: Dr. Diego Luis Linares
Directora: Dra. Gloria Inés Álvarez

02/11/2020



Santiago de Cali, 02/11/2020.

Señores

Pontificia Universidad Javeriana Cali.

Dr. Gerardo Mauricio Sarria

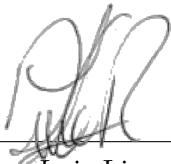
Director Carrera de Ingeniería de Sistemas y Computación.

Cali.

Cordial Saludo.

Por medio de la presente nos permito informarle que los estudiantes de Ingeniería de Sistemas y Computación Ana María García Posso (cod: 8935867) y Andrés Felipe Delgado Chamorro (cod: 8936235) trabajan bajo nuestra dirección en el proyecto de grado titulado “Técnicas de Aprendizaje Automático Aplicadas a la Predicción del Riesgo de Inversión en Criptomonedas”.

Atentamente,



Dr. Diego Luis Linares



Dra. Gloria Inés Álvarez

Santiago de Cali, 02/11/2020.

Señores

Pontificia Universidad Javeriana Cali.

Dr. Gerardo Mauricio Sarria

Director Carrera de Ingeniería de Sistemas y Computación.

Cali.

Cordial Saludo.

Nos permitimos presentar a su consideración el anteproyecto de grado titulado “Técnicas de Aprendizaje Automático Aplicadas a la Predicción del Riesgo de Inversión en Criptomonedas” con el fin de cumplir con los requisitos exigidos por la Universidad para llevar a cabo el proyecto de grado y posteriormente optar al título de Ingeniero de Sistemas y Computación.

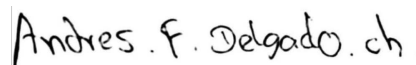
Al firmar aquí, damos fe que entendemos y conocemos las directrices para la presentación de trabajos de grado de la Facultad de Ingeniería aprobadas el 26 de Noviembre de 2009, donde se establecen los plazos y normas para el desarrollo del anteproyecto y del trabajo de grado.

Atentamente,



Ana María García Posso

Código: 8935867



Andrés Felipe Delgado Chamorro

Código: 8936235

Resumen

Una de las monedas con volatilidad más alta del mercado son las criptomonedas, esto quiere decir que su precio puede variar mucho en períodos de tiempo muy cortos. Esta volatilidad es una cualidad bastante marcada en las criptomonedas, sin embargo representa una desventaja para aquellas personas que deseen invertir en este campo. Debido a sus precios tan variables, no es fácil para los inversionistas saber si una inversión realizada en un determinado momento representará una buena o una mala inversión en el futuro.

El riesgo de inversión es un concepto muy importante cuando se habla de volatilidad, pues estos son directamente proporcionales. Este riesgo indica la probabilidad de perder dinero en un período de tiempo determinado. En este trabajo de grado se desarrollará una herramienta que utilizará técnicas de Aprendizaje Automático para realizar una estimación del riesgo de inversión en criptomonedas. El proyecto tiene como objetivo principal limitar la incertidumbre de los inversionistas en este mercado.

Palabras Clave: Riesgo de Inversión, Criptomonedas, Técnicas de Aprendizaje Automático, Volatilidad, Inversionistas.

Índice general

1. Descripción del Problema	11
1.1. Planteamiento del Problema	11
1.1.1. Formulación	12
1.1.2. Sistematización	12
1.2. Objetivos	12
1.2.1. Objetivo General	12
1.2.2. Objetivos Específicos	12
1.3. Justificación	12
1.4. Delimitaciones y Alcances	13
1.4.1. Entregables	13
2. Desarrollo del Proyecto	15
2.1. Marco de Referencia	15
2.1.1. Áreas Temáticas	15
2.1.2. Marco Teórico	15
2.1.3. Trabajos Relacionados	16
2.2. Metodología	17
2.2.1. Tipo de Estudio	17
2.2.2. Actividades	17
2.3. Resultados Esperados	18
2.4. Cronograma	19
2.5. Recursos	20
2.5.1. Humanos	20
2.5.2. Técnicos	20
2.5.3. Presupuesto	20
3. Bibliografía	23

Introducción

En el mundo financiero, un campo de estudio muy interesante y amplio ha sido el análisis del comportamiento de las monedas. Este proyecto busca aportar un grano de arena en este campo. Siendo conscientes de que existe un trabajo muy extenso en el mercado clásico, se pretende incursionar en un nuevo mercado que ha surgido en el siglo XXI, las criptomonedas.

Una de las principales características de las criptomonedas es su gran volatilidad, la cual describe en gran parte el comportamiento de estas monedas. Este proyecto busca hacer uso de técnicas de Aprendizaje Automático para analizar dicha volatilidad y realizar una estimación del riesgo de inversión en criptomonedas. Este proyecto se llevará a cabo con una metodología experimental y se espera obtener resultados interesantes, de cara a la predicción del riesgo de inversión.

Descripción del Problema

1.1. Planteamiento del Problema

A lo largo de los años, la humanidad ha tenido distintos tipos de moneda con la cual realizan sus transacciones y obtienen productos o servicios. Estas han ido evolucionando junto con la tecnología y necesidades de cada época. Tomamos como ejemplo las monedas de oro, las cuales evolucionaron a billetes de papel, estos a su vez evolucionaron a tarjetas de plástico manejadas por bancos y estos posteriormente crearon dinero digital. El tipo de moneda más reciente son las criptomonedas, las cuales son monedas digitales seguras y eficientes que no están reglamentadas por ningún banco ni ningún ente gubernamental, es una moneda global manejada por y para las personas.

En la actualidad, las criptomonedas son una realidad a nivel global. Tan sólo en Colombia en el 2017 las transacciones crecieron en un 1.200 % en comparación al año 2016 [1], ocupando así el tercer puesto a nivel mundial. Sin embargo, para al año 2018, tan sólo un año después, Colombia bajó hasta el puesto número 13 [2].

Este cambio entre un año y otro se debe, entre muchas otras causas, a la disminución de inversores en criptomonedas, pues no se puede negar que, aunque las criptomonedas tienen múltiples beneficios a nivel financiero, también presentan un riesgo para aquellos que buscan invertir, pues hay características inherentes a las criptomonedas que reducen la confianza en el inversor.

La principal característica negativa de las criptomonedas es la volatilidad. Dicha volatilidad está directamente relacionada con el riesgo de inversión, el cual se define como la probabilidad de perder dinero, pues a mayor volatilidad, mayor riesgo de inversión. Este factor está presente también en el mercado general y es la base sobre la que sustentan sus decisiones la mayoría de los inversores.

Por instinto humano, siempre buscamos reducir los riesgos de nuestras acciones y más aún si hay dinero implicado. Por este motivo, proponemos una herramienta que permita predecir una estimación del riesgo de inversión en criptomonedas, lo cual daría información adicional a las personas para tomar sus decisiones de inversión.

1.1.1. Formulación

¿Cómo implementar un modelo de aprendizaje automático para predecir el riesgo de inversión en criptomonedas?

1.1.2. Sistematización

- ¿Qué datos utilizar para el análisis de riesgo de inversión en criptomonedas y cómo procesarlos?
- ¿Qué técnicas de aprendizaje automático serán evaluadas para predecir el riesgo de inversión en criptomonedas?
- ¿Cómo evaluar los modelos de aprendizaje automático y comparar los resultados obtenidos?

1.2. Objetivos

1.2.1. Objetivo General

Desarrollar 3 modelos de aprendizaje automático para predecir el riesgo de inversión en criptomonedas, utilizando Bitcoin como caso de estudio.

1.2.2. Objetivos Específicos

- Identificar los datos a utilizar para el análisis de riesgo de inversión en criptomonedas y procesarlos.
- Seleccionar y experimentar con 3 técnicas de aprendizaje automático y ajustar los modelos para predecir el riesgo de inversión.
- Evaluar los 3 modelos de aprendizaje automático y comparar los resultados obtenidos.

1.3. Justificación

Las inversiones son un tema importante cuando de criptomonedas se habla, debido a que de estas depende en gran parte el éxito o fracaso de esta moneda. Actualmente, debido a la volatilidad que las criptomonedas presentan, las inversiones son riesgosas pues al realizarlas no hay manera cuantitativa de saber si a futuro será una buena inversión o no.

En este trabajo de grado se llevará a cabo una herramienta dirigida principalmente a los inversores en criptomonedas. El fin de esta herramienta será predecir el riesgo de inversión basándose en datos de dominio público con valores históricos de esta moneda para analizar y predecir su comportamiento.

Una herramienta de este tipo puede ayudar a estimar el riesgo que el usuario tendrá y de esa forma proveer información adicional para la toma de decisiones de inversión en Bitcoin.

1.4. Delimitaciones y Alcances

- Se utilizarán bases de datos gratuitas como insumo para realizar el entrenamiento de los modelos
- El tiempo estimado para llevar a cabo la investigación, experimentación de los modelos de aprendizaje automático, comparación de resultados y reporte escrito es de 6 meses.
- Se experimentará con 3 modelos de aprendizaje automático.

1.4.1. Entregables

- Documento de trabajo de grado con información respectiva a los modelos de aprendizaje automático, conceptos introductorios de criptomonedas, resultados obtenidos y comparaciones.
- Software con la experimentación de los modelos de aprendizaje automático utilizados.

Desarrollo del Proyecto

2.1. Marco de Referencia

2.1.1. Áreas Temáticas

Este proyecto de grado está enfocado en las siguientes áreas temáticas:

- Aprendizaje Automático Supervisado
- Aprendizaje Automático con Datos Estructurados
- Riesgo de Inversión
- Volatilidad
- Criptomonedas

2.1.2. Marco Teórico

- Criptomonedas:

Las criptomonedas son monedas digitales las cuales permiten realizar transacciones de manera rápida y segura. Utilizan el sistema de seguridad de Blockchain para proteger y controlar las transacciones realizadas. Estas monedas digitales son descentralizadas, es decir, que no están reglamentadas ni son manejadas por ningún ente en específico.

La primer criptomoneda creada y la más popular hasta el momento es Bitcoin, por este motivo será utilizada en este proyecto.

- Riesgo de Inversión:

“El valor de riesgo, o VaR (en inglés, Value at Risk), cuantifica el riesgo de mercado para una cartera. Más precisamente, el VaR mide la pérdida máxima de la cartera, en un periodo dado, con una fiabilidad dada.”[3]

Supongamos un ejemplo: Si se toman como parámetros un día y una fiabilidad del 97 %, el valor de riesgo calcularía la pérdida máxima de dinero en un día y con una fiabilidad del 97 %.

- Volatilidad:

La volatilidad mide el comportamiento de los valores de un activo financiero a lo largo del tiempo. En el caso de las criptomonedas, la volatilidad es muy grande, lo que quiere decir que el comportamiento de estas es muy oscilante.

2.1.3. Trabajos Relacionados

Los modelos de Heterocedasticidad Condicional Autorregresiva Generalizada (GARCH) son los más utilizados en la literatura para modelar la volatilidad y estimar el Valor de Riesgo (VaR).

En el artículo “Volatility estimation for Bitcoin: A comparison of GARCH models”[4] del año 2017, se explora entre una gran variedad de modelos GARCH para definir qué modelo produce una mejor estimación de volatilidad para Bitcoin. Se encontró que el modelo AR-CGARCH es el que destaca, ya que incluye un componente de corto y largo plazo de la varianza condicional.

En el artículo “The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression”[5] del año 2018, el modelo GARCH clásico se combinó con el enfoque de aprendizaje automático, de modo que las ecuaciones de media y volatilidad se estimaron mediante la regresión de vectores de soporte. El modelo generado se llamó SVR-GARCH y al evaluar su desempeño predictivo de la volatilidad, se concluyó que la incorporación de una técnica basada en el aprendizaje automático produce un mejor poder de pronóstico de la volatilidad sobre los puntos de referencia GARCH.

En el artículo “A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis”[6] del año 2018, se propuso un modelo híbrido de heterocedasticidad condicional autorregresiva generalizada de red neuronal artificial (ANN-GARCH) con preprocesamiento para pronosticar la volatilidad del precio del Bitcoin. Se concluyó que la combinación de varios enfoques puede llevar a resultados muy interesantes que, en este trabajo en específico, fueron aplicados para Bitcoin. Se sugiere como futuros trabajos la creación de modelos ANN-GARCH para otras criptomonedas y comparar su rendimiento.

En el artículo “Modelling volatility of cryptocurrencies using Markov-Switching GARCH models”[7] del año 2018, se muestra cómo modelar la volatilidad de las 4 criptomonedas más populares en la actualidad: Bitcoin, Ethereum, Ripple y Litecoin. Por medio de modelos GARCH, usando más de mil de ellos para determinar cuál genera una mejor predicción del Valor de Riesgo. Se pudo concluir en este relevante estudio que los modelos GARCH estandar pueden producir predicciones incorrectas, dando como resultado una gestión de riesgo ineficaz.

En el artículo “Forecasting Value-at-Risk of Cryptocurrencies with RiskMetrics type models”[8] del año 2020, se analizó empíricamente si el Valor de Riesgo de las criptomonedas se puede pronos-

ticar con modelos de Función de Medias Móviles Ponderadas Exponenciales (EWMA), las cuales tienen como objetivo principal la estimación de la volatilidad de un periodo en una serie de tiempo y rastrearla muy de cerca a medida que cambia, siendo una alternativa a los modelos GARCH. Se concluyó que el Valor de Riesgo de las criptomonedas se puede pronosticar con éxito con modelos EWMA Parsimoniosos.

Después de la búsqueda de los artículos relacionados, pudimos comprobar que la gran mayoría de la documentación encontrada tiene como epicentro modelos matemáticos y estadísticos que, en algunas ocasiones, son combinados con técnicas basadas en Aprendizaje Automático.

Nuestra propuesta pretende abarcar este problema haciendo uso únicamente de modelos de Aprendizaje Automático, por lo tanto, podríamos decir que nos encontramos en un área poco explorada y se podrían generar resultados interesantes para futuros estudios en este campo.

2.2. Metodología

2.2.1. Tipo de Estudio

Este proyecto será desarrollado con un tipo de estudio experimental, y se enfoca en el desarrollo de una herramienta para la solución de un problema en específico. En este caso, el problema se centra en la estimación del riesgo de inversión en criptomonedas, de modo que se logre limitar la incertidumbre en los inversionistas.

2.2.2. Actividades

A continuación se enumeran las 7 actividades principales, junto con las subactividades de cada una. Cabe resaltar que el proceso de la realización de las actividades será iterativo y no en cascada, pues a lo largo del proyecto se realizarán ciclos para la revisión, ajuste y/o corrección de actividades anteriores.

1. Identificar los datos a utilizar para el análisis del riesgo de inversión en criptomonedas
 - Listado de los datos que se necesitan
 - Búsqueda de los datos necesarios
 - Obtención de los datos
 - Conversión de los datos a un formato de fácil manipulación en caso de ser necesario
 - Verificación del tipo de datos (serie temporal, muestral, geográfico, etc.)
2. Procesamiento de los datos
 - Corrección o eliminación de valores atípicos
 - Completar los valores faltantes (con cero, media, mediana, etc.) o eliminar sus filas o columnas

- Eliminación de los atributos que no proporcionan información útil para la tarea
 - Normalizar atributos si es necesario
 - Realización de otras funciones necesarias para el preprocesamiento de los datos
3. Experimentar múltiples técnicas de aprendizaje automático, seleccionar las 3 mejores y ajustar los modelos
- Entrenamiento de múltiples modelos de diferentes categorías como SVM, Random Forest, redes neuronales, etc.
 - Comparación de los modelos utilizando métricas de desempeño como Accuracy, F1-Score, Recall, etc.
 - Selección de los 3 modelos con mejor desempeño
 - Estimación de parámetros de los 3 modelos para el ajuste con los datos.
4. Comparación de los resultados obtenidos en los 3 modelos
- Documentación y comparación del desempeño de los modelos y los resultados de cada uno
5. Documentación y presentación del proyecto

2.3. Resultados Esperados

- Experimentación con distintos modelos de Aprendizaje Automático para la predicción del riesgo de inversión en criptomonedas.
- Selección de los 3 mejores modelos para realizar la estimación del riesgo.
- Análisis de los resultados obtenidos.

2.4. Cronograma

[illegible]

2.5. Recursos

2.5.1. Humanos

- Diego Luis Linares
 - Doctorado en Ingeniería de La Programación E Inteligencia Artificial, Universidad Politécnica de Valencia
 - Especialización en Sistemas Gerenciales de Ingeniería
 - Pregrado en Ingeniería de Sistemas y Computación, Pontificia Universidad Javeriana Cali
- Gloria Inés Álvarez
 - Doctorado en Reconocimiento de Formas e Inteligencia Artificial, Universidad Politécnica de Valencia
 - Maestría en Sistemas y Computación, Universidad de Los Andes
 - Pregrado en Ingeniería de Sistemas, Universidad Autónoma de Manizales
- Andrés Felipe Delgado
 - Estudiante de Pregrado en Ingeniería de Sistemas y Computación, Pontificia Universidad Javeriana Cali
- Ana María García
 - Estudiante de Pregrado en Ingeniería de Sistemas y Computación, Pontificia Universidad Javeriana Cali

2.5.2. Técnicos

- Computador de mesa o portátil con las siguientes especificaciones como mínimo:
 - 8 GB de Memoria RAM
 - 100 GB de Almacenamiento Interno libre
- Conexión a internet que cuente con las siguientes especificaciones como mínimo:
 - 5 MBps de Download
 - 1 MBps de Upload

2.5.3. Presupuesto

Para el cálculo del presupuesto se tendrá en cuenta lo siguiente:

- El salario mensual promedio de los directores de proyecto es de \$8.000.000 y agregando los parafiscales que equivalen a un 1.5 del valor de salario, los costos mensuales de los directores suman \$12.000.000
- Los directores trabajan un promedio de 40 horas semanales, al igual que los estudiantes.

- El valor de la matrícula de cada estudiante es de \$8.874.000. A lo largo del semestre se estudian 18 semanas lo que equivale a 4.5 meses, por ende el costo mensual del trabajo de un estudiante es de \$1.972.000
- El computador con las especificaciones mínimas requeridas tiene un costo de \$2.000.000 en promedio. Suponiendo una vida útil de 3 años, el costo mensual del computador es de \$56.000 aproximadamente.
- El costo mensual de un servicio de internet con las especificaciones mínimas requeridas es de \$42.000

Recursos	Tipo de Recurso	Costo Por Mes (Incluyendo Parafiscales)	Costo Por Hora	Cantidad de horas dedicadas semanalmente	Cantidad de Semanas dedicadas al Proyecto	Presupuesto del Recurso durante el Proyecto
Director Diego Linares	Humano	12.000.000	70.000	2	24	3.360.000
Directora Gloria Álvarez	Humano	12.000.000	70.000	2	24	3.360.000
Estudiante Ana García	Humano	1.972.000	12.325	10	24	2.958.000
Estudiante Andrés Delgado	Humano	1.972.000	12.325	10	24	2.958.000
Computador	Técnico	56.000	350	10	24	84.000
Internet	Técnico	42.000	263	10	24	63.000
Presupuesto Total del Proyecto						12.783.000

Figura 2.1: Presupuesto del Proyecto

Bibliografía

- [1] “Cómo en Colombia creció el mercado de bitc  in.” [Online]. Available:
<https://www.dinero.com/economia/articulo/como-en-colombia-crecio-el-mercado-de-bitcoin/256116>.
- [2] “Pa  ses con mayor volumen de transacciones de bitcoin - ProEconom  a.” [Online]. Available:
<https://proeconomia.net/paises-volumen-transacciones-bitcoin/>.
- [3] “Value at Risk | Estrategias de Inversi  n.” [Online]. Available:
<https://www.estrategiasdeinversion.com/herramientas/diccionario/mercados/value-at-risk-t-1157>.
- [4] P. Katsiampa, “Volatility estimation for Bitcoin: A comparison of GARCH models,” *Econ. Lett.*, vol. 158, pp. 3–6, 2017, doi: 10.1016/j.econlet.2017.06.023.
- [5] Y. Peng et al., “The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression,” *Expert Syst. Appl.*, vol. 97, pp. 177–192, 2018, doi: 10.1016/j.eswa.2017.12.004.
- [6] W. Kristjanpoller and M. C. Minutolo, “A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis,” *Expert Syst. Appl.*, vol. 109, pp. 1–11, 2018, doi: 10.1016/j.eswa.2018.05.011.
- [7] G. M. Caporale and T. Zekokh, “Modelling volatility of cryptocurrencies using Markov-Switching GARCH models,” *Research in International Business and Finance*, vol. 48, pp. 143–155, 2019, doi: 10.1016/j.ribaf.2018.12.009.
- [8] W. Liu, A. Semeyutin, C. Keung, M. Lau, and G. Gozgor, “Forecasting Value-at-Risk of Cryptocurrencies with RiskMetrics type models,” 2020, doi: 10.1016/j.ribaf.2020.101259.
- [9] “Invierte en bitcoin con m  nimo riesgo | TyN Magazine.” [Online]. Available:
<https://www.tynmagazine.com/invierte-en-bitcoin-con-minimo-riesgo/>.
- [10] “La volatilidad de bitcoin explicada.” [Online]. Available:
<https://www.criptonoticias.com/mercados/volatilidad-bitcoin-explicada/>.
- [11] “Riesgos de invertir en bitcoin.” [Online]. Available:
<https://www.dinero.com/edicion-impresa/caratula/articulo/riesgos-de-invertir-en-bitcoin/247208>.

-
- [12] “Conozca los pro y contras de invertir en Bitcoin o criptomonedas en este 2020.” [Online]. Available: <https://www.larepublica.co/economia/conozca-los-pro-y-contras-de-invertir-en-bitcoin-o-criptomonedas-en-este-2020-2957002>.
- [13] “Coin Dance | Volumen de LocalBitcoins (Colombia).” [Online]. Available: <https://coin.dance/volume/localbitcoins/COP>.
- [14] “Cómo se mide el riesgo financiero.” [Online]. Available: <https://estrategafinanciero.com/como-se-mide-riesgo-financiero/>.
- [15] “Indicadores de volatilidad analizados en detalle | InviertenBitcoin.” [Online]. Available: <https://inviertenbitcoin.com/indicadores-de-volatilidad/>.
- [16] “Volatilidad — Indicadores técnicos — TradingView.” [Online]. Available: <https://es.tradingview.com/ideas/volatility/>.
- [17] P. Hrytsiuk, T. Babych, and L. Bachyshyna, “Cryptocurrency portfolio optimization using Value-at-Risk measure,” 2019, pp. 385–389, doi: 10.2991/smtesm-19.2019.75.
- [18] J. Chu, S. Chan, S. Nadarajah, and J. Osterrieder, “Risk and Financial Management GARCH Modelling of Cryptocurrencies,” doi: 10.3390/jrfm10040017.
- [19] “¿Qué es criptomoneda?” [Online]. Available: <https://es.cointelegraph.com/bitcoin-for-beginners/what-are-cryptocurrencies>.
- [20] “Volatilidad - Qué es, definición y concepto | Economipedia.” [Online]. Available: <https://economipedia.com/definiciones/volatilidad.html>.