

PERFIL DE RIESGO Y APRENDIZAJE AUTOMÁTICO: HERRAMIENTAS QUE APOYAN LA TOMA DE DECISIONES EN LA GESTIÓN DE CRISIS

Risk profile and Machine Learning (ML): Tools that support decision making during crisis management

Mayor Alejandro Arancibia Valdivia*

Resumen: Las relaciones entre Estados se producen en condiciones generadas por las características del sistema internacional, fluctuando entre la cooperación y el conflicto, siendo la crisis la forma más común de relación conflictual. La gestión de la crisis internacional, se encuentra asociada a un proceso de toma de decisiones, el cual puede ser apoyado por instrumentos como un perfil de riesgo que guíe el análisis, el entendimiento del problema y sustente un juicio informado. El propósito de este artículo es discutir sobre la aplicabilidad del uso de perfiles de riesgo basados en algoritmos de aprendizaje automático y estadística, para proyectar escenarios que permitan apoyar el proceso de toma de decisiones, durante la gestión de la crisis internacional.

Palabras clave: Crisis internacional, perfil de riesgo, toma de decisiones, aprendizaje automático, inteligencia artificial

Abstract: Relations between states occur under conditions generated by characteristics of the international system, fluctuating between cooperation and conflict, with the crisis being the most common form of conflict relationship. The management of the international crisis is associated with a decision-making process which can be supported by instruments such as a risk profile that guides analysis, understanding the problem and sustain informed judgment.

The purpose of this article is to discuss the applicability of the use of risk profiles based on machine learning and statistical algorithms to project scenarios and support the decision-making process during the management of the international crisis

* Oficial de Ejército con el grado de Mayor, del arma de Artillería. Licenciado en Ciencias Militares. Magíster en Ciencias Políticas mención Relaciones Internacionales, Pontificia Universidad Católica de Chile. Alumno de III año del Curso Regular de Estado Mayor de la Academia de Guerra (2018-2020). ✉ alejandro.arancibia @acague.cl.

Keywords: International crisis, risk profile, decision making, machine learning, artificial intelligence.

INTRODUCCIÓN

El presente artículo aborda la crisis empleando variables asociadas al sistema internacional, considerando al Estado como unidad de análisis y principal actor en las relaciones internacionales, por consiguiente, se centra en la crisis internacional entre actores estatales.

El estudio de la gestión de crisis internacionales constituye un área de relevancia para la contribución al conocimiento de las relaciones entre Estados, la comprensión de las causas que subyacen al conflicto y los efectos desestabilizadores de la crisis en el sistema internacional, los que constituyen un riesgo para la seguridad. En este contexto la definición de un perfil de riesgo para definir el fenómeno y comprender cómo se gestiona la crisis, proporciona un marco de referencia para identificar actores clave, acciones ejecutadas, impacto de los organismos internacionales involucrados y recursos en juego, permitiendo evaluar el proceso de toma de decisiones y el resultado de la gestión.

La crisis internacional supone potenciales riesgos frente a la adopción de distintos cursos de acción durante su gestión, pudiendo tener desenlaces que escalen a conflicto armado. Así, el riesgo subyace a la crisis internacional, y se asocia a la probabilidad de ocurrencia de eventos o externalidades negativas que impacten a los actores. Por consiguiente, la capacidad de poder determinar y evaluar el riesgo, permite reducir la incertidumbre y preparar acciones tendientes a una mejor respuesta y gestión una vez que la crisis ha estallado.

Esto constituye el punto de partida para determinar la interacción entre las variables asociadas al perfil propuesto, y más importante, proyectar escenarios en los cuales la crisis escala o distiende. Lo anterior construye conocimiento útil que sustenta el proceso de toma de decisiones en juicios informados, y se basa en el uso de datos históricos sobre el desarrollo de las crisis anteriores. Así, los escenarios proyectados a través del perfil de riesgo propuesto, emplean datos

cuantitativos, perfilados con el último avance en Big Data y Aprendizaje Automático (Machine Learning) que los sustentan.

Para cumplir con el propósito, se comienza este artículo conceptualizando el riesgo en la crisis internacional y la toma de decisiones, posteriormente se propone un modelo de perfil de riesgo en función de variables del sistema internacional basados en datos del centro de pensamiento *International Crisis Behaviour* (ICB), finalmente se concluye sobre su utilidad en función del desempeño predictivo del modelo, basado en técnicas estadísticas y de aprendizaje automático.

RIESGO EN LA ADMINISTRACIÓN DE LA CRISIS

El riesgo en relaciones internacionales puede ser definido como "una posibilidad que ocurra una situación desfavorable durante la implementación de los planes presupuestados por el Estado. Son las potenciales consecuencias negativas no deseadas de un evento o acción" (Ostrowska & Mazur, 2015, p. 4). Crisis, riesgo y conflicto constituyen conceptos interrelacionados que generan un entorno complejo de definir e impacta en el proceso de toma de decisiones para la gestión de la crisis internacional.

Durante la gestión de la crisis, contar con un modelo que proyecte escenarios asociados a un perfil de riesgo profundiza la comprensión de la interacción de distintas variables, tales como valores o recursos en disputa y sus características, el impacto del involucramiento de potencias y otros factores de preponderancia, además de ayudar a la definición del contexto en el cual se desarrolla la crisis y guiar el enmarcamiento o *framing* del problema.

Estas variables constituyen vehículos de acción que podrían determinar los posibles cursos de acción emprendidos por un Estado para gestionar la crisis. La lógica subyacente al empleo de un perfil de riesgo es que, si se logran relacionar las variables asociadas al riesgo, es factible proyectar escenarios en los cuales su interacción pudiese escalar la crisis a conflicto armado. Como consecuencia, se

puede comprender cómo la elección de distintos cursos de acción aumenta o disminuye las probabilidades de escalada basado en evidencia empírica.

Saber bajo qué circunstancias aumentan las probabilidades que un Estado ejecute acciones que escalen la crisis es información relevante. Proyectar un escenario probable ayuda a planificar, preparar y ejecutar acciones que permitan enfrentar las acciones de la contraparte en una mejor posición, ya sea de negociación, coerción y eventualmente coacción. Aquí radica la utilidad y significación de un modelo de perfil de riesgo que sustente el proceso de toma de decisiones.

Una forma de abordar la interacción de las variables antes descritas, es construir un modelo de perfil de riesgo que permita identificar, contextualizar y asociar la toma de decisiones a la administración de la crisis, incluyendo variables del sistema internacional y las características de los actores involucrados, permitiendo de este modo determinar su contribución o utilidad en el proceso de toma de decisiones durante la gestión de la crisis.

Según Brecher & Wilkenfeld (2000) existen dos condiciones para la aparición de una crisis internacional: (1) cambios en el tipo y/o forma de la intensidad de la hostilidad verbal o física en las interacciones entre dos o más Estados, elevando por ende la probabilidad del empleo de hostilidad militar; y (2) la desestabilización de la relación entre actores y la estructura del sistema internacional.

El manejo o administración de la crisis constituye la última oportunidad para realizar acciones tendientes a su minimización, control o solución, antes de la escalada hacia un conflicto armado. Requiere de una combinación de elementos constitutivos de poder duro (poder como coacción) y blando (poder como cooperación), además de capacidad de negociación, como plantea Joseph Nye en su obra "La naturaleza cambiante del poder mundial" (Nye, 1990).

La articulación de los elementos del poder y la capacidad de negociación debe permitir a los tomadores de decisiones generar condiciones favorables para la consecución de sus objetivos evitando la escalada, la que trae aparejada consigo de manera indefectible, un aumento en los costos; ergo una disminución en las

utilidades. Para lo anterior, deben considerarse alternativas de salida pacífica de cara a la contraparte, las que deben ser honrosas o decorosas debido a que el estatus posee importancia en materia de relaciones internacionales, asociada al porte estratégico por parte de un Estado. A su vez debe mantenerse el equilibrio entre el uso del poder militar, manteniendo una capacidad disuasiva creíble, acompañada de la demostración de voluntad en su uso (Verdugo, 2011).

El concepto de manejo de crisis se refiere a buscar una solución que resulte aceptable para ambas partes, evitando llegar a la guerra. Posee tres etapas: el desafío, la reacción y el acuerdo. La reacción es la forma en cómo se conduce la crisis empleando un conjunto de acciones para anular al oponente (Gallardo, 2005).

En el contexto de la crisis, el riesgo constituye una amenaza a valores o intereses para un actor estatal, supone una condición de inestabilidad, un evento disruptivo que se interpone entre las acciones planificadas y los objetivos propuestos. Esto pone en marcha una serie de acciones tendientes a responder a las circunstancias, las que podrían ser vehículos constitutivos de una crisis. La dificultad de poder prever el desenlace o el resultado de las acciones a implementar generan condiciones de riesgo.

La falta de información y la incertidumbre que la crisis provoca, la importancia de los intereses amenazados, la dificultad de definir y comprender el problema sumado al corto tiempo de respuesta, son vehículos de riesgo, interactúan y se influyen mutuamente para crear condiciones que complejizan el proceso de toma de decisiones y la implementación de acciones.

TOMA DE DECISIONES EN LA GESTIÓN DE LA CRISIS

Considerando las condiciones que la administración de la crisis impone, existen variables asociadas al proceso de toma de decisiones y diversas formas de enmarcar el problema a fin de comprender desde una perspectiva teórica cómo se gesta una decisión. El modelo propuesto en este artículo asume los supuestos

considerados en el paradigma de actor racional (Rational Choice)¹, es decir, preferencias relativamente estables en el tiempo, capacidad de cálculo sobre posibles desenlaces, transitabilidad entre las alternativas y la elección del curso de acción que maximiza utilidades o beneficios. Es relevante plantear que la Teoría de Racionalidad Limitada (Bounded Rationality) dentro del paradigma de actor racional, es la que más se acerca al contexto de la crisis, dado que descarta el supuesto de información completa o perfecta y de cálculo ilimitado por parte del actor en crisis.

De igual modo, el perfil de riesgo constituye un apoyo relevante para las decisiones bajo riesgo o incertidumbre. El proceso de toma de decisiones en crisis internacionales se caracteriza por plantear escenarios de incertidumbre o riesgo. Es muy difícil que exista certeza, debido a la existencia de objetivos o intereses contrapuestos, y de dos o más actores que luchan por imponer su voluntad y maximizar ganancias. Por ende, es un escenario competitivo, de información incompleta, escaso tiempo para resolver y de altos costos en juego.

Las decisiones bajo riesgo suponen la elección de alternativas en las cuales los resultados no pueden ser proyectados con certeza. A razón de ello, es factible listarlos y asignarles probabilidades de ocurrencia. No es factible tener certeza, pero se pueden proyectar probabilidades.

En el otro extremo, se encuentran las decisiones bajo incertidumbre. La incertidumbre significa la incapacidad de efectuar suposiciones plausibles sobre las condiciones futuras en las que se desarrollarán los cursos de acción escogidos, dada la dificultad de asignar probabilidades. Dentro de la clasificación de incertidumbre, se consideran dos posibles estados: la incertidumbre estructurada y la no estructurada. En la incertidumbre estructurada se conocen los estados del ambiente, pero no las probabilidades. En la incertidumbre no estructurada, se desconocen los estados del ambiente y las probabilidades de ocurrencia, por tanto, también se denomina contexto de decisión bajo ignorancia.

¹ Principal paradigma económico para la descripción y entendimiento de la toma de decisiones de un individuo.

El escenario de incertidumbre ya sea estructurado o no, constituye el escenario más común en crisis internacionales, se caracteriza por la complejidad de listar todos los posibles resultados y asignarles probabilidades. Un problema complejo y/o mal estructurado, requiere de un encuadramiento o “framing”, que permita la comprensión de las alternativas existentes y otorgue un contexto.

La construcción del marco del problema requiere de un ordenamiento de preferencias, y la construcción de una función de utilidad asociada al contexto o ambiente en el cual el problema se presenta, lo que también se conoce como “estado de la naturaleza” (Resnik, 1987). Esta permitirá ordenar las preferencias respecto de los resultados y las acciones disponibles frente al problema, pudiendo determinar los criterios de selección de la mejor alternativa.

El escenario deseable es pasar de un contexto de incertidumbre estructurada o no estructurada, a un contexto de decisión bajo riesgo, lo que sería posible a través del empleo de un modelo de perfil de riesgo que proyecte la probabilidad de ocurrencia de escenarios en función de las alternativas escogidas.

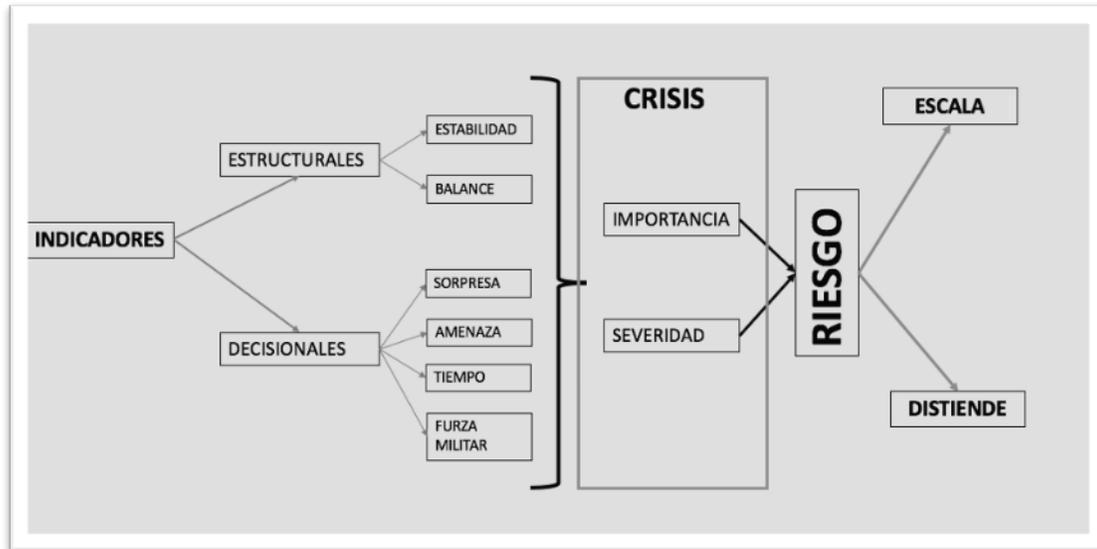
HACIA UN MODELO DE PERFIL DE RIESGO

Articulando las condiciones bajo las cuales se puede desarrollar una crisis, desde la perspectiva de la estructura del sistema internacional, es factible conceptualizar el riesgo, y a partir de ello crear un modelo de perfil de riesgo para proyectar escenarios. La construcción del modelo propuesto emplea indicativos de crisis estructurarles, debido a que las decisionales a nivel del individuo tomando decisiones, constituyen variables de contexto las que si bien son importantes para la comprensión de la crisis, son difíciles de medir dada las condicionantes que la rodean.

Las variables propuestas se centran en el nivel de la estructura del sistema internacional y asumen al Estado como actor con capacidad de agencia, es decir tiene voluntad y racionalidad, utilizando todos sus medios para la consecución de

sus objetivos. De este modo, las variables constitutivas del modelo propuesto se presentan en la Figura 1.

Figura 1. Indicadores de crisis

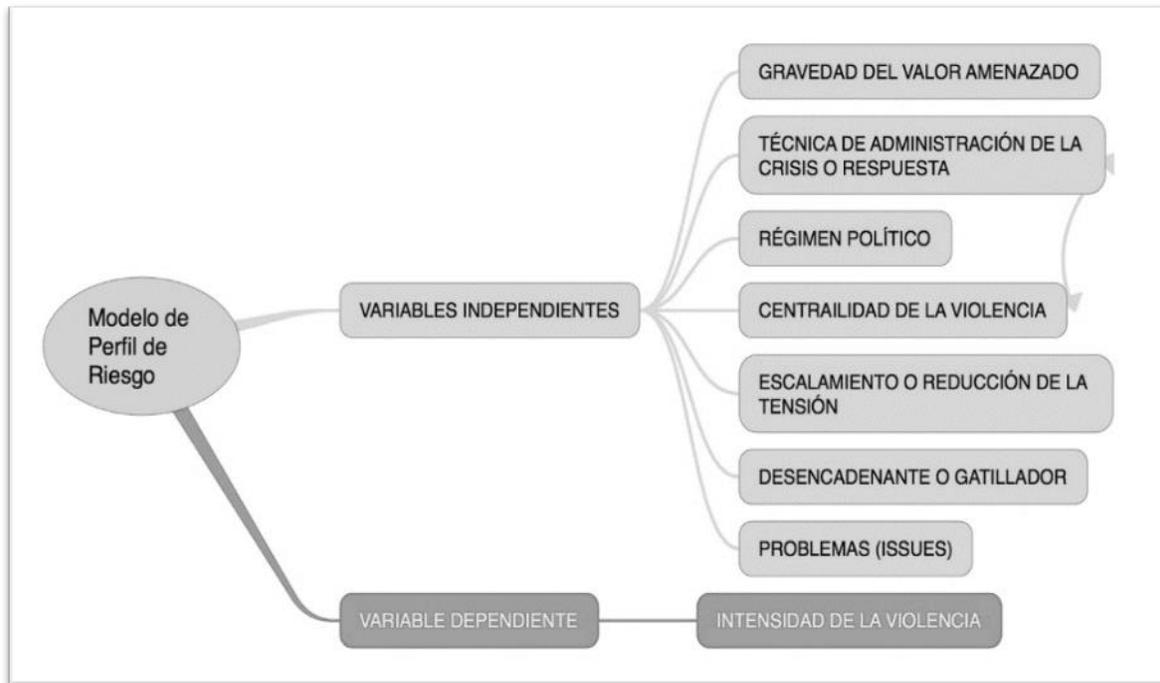


Fuente: Elaboración propia basado en la definición de la crisis por Brecher y Wilkinfred.

El modelo de perfil de riesgo se basa en variables de nivel actor en el sistema internacional (Estado). Se han definido en función de su aporte a priori en la generación de una proyección de crisis internacional. Este modelo de variables constituye la base del perfil de riesgo propuesto y, para ello, se procedió a determinar el grado de correlación entre ellas.

Posteriormente se efectuó una regresión logística con la finalidad de determinar la factibilidad de proyectar escenarios en los cuales la crisis escala o distiende. Finalmente se empleó un algoritmo "Random Forest" para determinar la existencia de un mayor ajuste del modelo y una mejora en la proyección de escenarios. Las variables dependientes e independientes se presentan en la Figura 2.

Figura 2. Modelo de perfil de riesgo



Fuente: Elaboración propia basada en la definición de la crisis por Brecher y Wilkinfred.

ANÁLISIS Y DISCUSIÓN DE DATOS

Para determinar la relación existente entre el perfil de riesgo propuesto y el proceso de toma de decisiones se procedió a efectuar un test de independencia Chi cuadrado o tabla de contingencia entre la variable dependiente “intensidad de la violencia” y las variables independientes o explicativas del modelo propuesto.

Las pruebas de Chi cuadrado son un test de hipótesis que compara la distribución observada de los datos con una distribución esperada de los datos. Se centran en determinar el grado de asociación o relación entre las variables constitutivas del modelo. No buscan determinar causalidad, sino relación entre variables. Esto permite definir qué variables son factibles de permanecer en el modelo y ser empleadas para una regresión logística en este caso.

Se esperaba encontrar relaciones estadísticamente significativas y con magnitudes que fuesen de relaciones moderadas a fuertes según los valores obtenidos por el estadístico Chi cuadrado. En caso de existir esas relaciones, se

ganaría confianza en que el perfil de riesgo propuesto apoya el proceso de toma de decisiones, y se constituiría en una herramienta que permitiría operacionalizar el riesgo para así poder ordenar preferencias y seleccionar alternativas que guíen la decisión.

Lo anterior se sustenta en la lógica que si es factible saber que el método de gestión de crisis escogido se relaciona con algún posible resultado de ésta, ya sea escalando o distendiendo, se pueden ordenar preferencias y tomar decisiones informadas. Esto acercaría un proceso de toma de decisiones bajo incertidumbre o incertidumbre estructurada a uno de riesgo.

Posterior a las correlaciones entre las variables de interés y condicionado a los resultados de estas, se procedió a efectuar una regresión logística, la que buscaba determinar la factibilidad de poder proyectar el resultado de la crisis, es decir, en función del perfil de riesgo determinar cuándo una crisis podría escalar o distender. Este procedimiento se efectuó mediante una regresión logística binominal y el empleo de un algoritmo de minería de datos (o data mining en inglés) ya señalado. La intención fue determinar cuál de los dos métodos genera un modelo que se ajuste mejor a los datos y, por ende, aumente la posibilidad de proyección de la variable dependiente, en este caso, intensidad de la violencia.

Al respecto, aplica lo señalado por Muchlinski et al. (2015) quienes concluyen sobre la mayor robustez y ajuste del algoritmo Random Forest por sobre la regresión logística clásica para problemas de clasificación y predicción, indicando que los resultados demuestran la precisión predictiva superior de Random Forest, ya que la flexibilidad que ofrecen los métodos no paramétricos -como Random Forest- permiten predicciones más precisas de eventos.

Las correlaciones presentaron significación estadística entre las variables que constituyen el modelo propuesto por sobre 1.000 casos de crisis tabuladas desde 1918 hasta el año 2015, constituidas por la base de datos del ICB. A continuación se presentan esos datos.

Tabla 1. Pruebas de Chi cuadrado

| Variables en pruebas de Chi Cuadrado | Chi-cuadrado de Pearson | Significación asintótica (bilateral) |
|---|--------------------------------|---|
| Intensidad de violencia / Gravedad del valor amenazado | 138,757 ^a | ,000 |
| Intensidad de violencia con Régimen | 27,250 ^a | ,007 |
| Intensidad de violencia con Escalamiento o distención | 60,263 ^a | ,000 |
| Intensidad de violencia con Forma de administración de crisis o respuesta | 578,876 ^a | ,000 |
| Intensidad de violencia con Centralidad de la violencia | 1341,541 ^a | ,000 |
| Intensidad de violencia con Problema | 21,117 ^a | ,012 |
| Intensidad de violencia con Evento desencadenante | 174,459 ^a | ,000 |
| N de casos válidos | 1.000 | |

Fuente: Elaboración propia basado en datos del ICB.

REGRESIÓN LOGÍSTICA

La regresión logística (RL) constituye una herramienta estadística que permite predecir el resultado de una variable dependiente o explicada en función de los datos de una o más variables independientes o explicativas. Otorga mayor flexibilidad para el análisis debido a que puede emplear variables cualitativas que pueden ser dicotomizadas, es decir asignadas a solo dos categorías a predecir.

Empleando la RL se pretende determinar la *probabilidad* que ocurra el fenómeno de estudio a través de variables asumidas como relevantes o influyentes. Así, la

RL consiste en determinar una *función logística* de las variables independientes que logre clasificar a los casos muestreados en una de las dos posibles categorías de la variable dependiente.

La función logística constituye la que para cada caso encuentra según los valores de las variables independientes (x) y la probabilidad (p) de que el efecto observado se produzca en la variable dependiente (y). Esta técnica utiliza una transformación logarítmica de la función logística denominada *logit*, convirtiendo la probabilidad (p) en odds (Keohane et al., 1994).

El modelo de regresión logística trabajado presenta un robustez y ajuste que permite predecir el valor de la variable dependiente, determinando si una crisis escala o no. La utilidad del modelo para la toma de decisiones durante la gestión de la crisis requiere de la probabilidad de ocurrencia de una escalada y, en función de ello, escoger cursos de acción que busquen minimizar esas probabilidades.

Aquí radica la utilidad de un modelo de perfil de riesgo, ya que no se busca definir las causas de la crisis mediante una demostración estadística, sino que entregar una herramienta que permita tomar una decisión informada asociada a la contextualización del riesgo durante la crisis.

La variable de control no presentó significación estadística y fue incluida para aislar la posibilidad que el origen del problema tuviese interacción en la escalada o distención de la crisis. De este modo se puede inferir que el problema que origina la crisis no tiene incidencia en si esta escala o no, existiendo mayor relevancia en la forma de conducción de la crisis o del hecho que la desencadena, más que el problema en sí.

El “odds ratio” dado por el coeficiente Exp (b) indica que para la variable “detonante de la crisis” el aumento de una unidad de esta variable incrementa en 1,2 veces la probabilidad que la crisis escale.

El modelo de regresión logística logrado en esta investigación es estadísticamente significativo $\chi^2 (5) = 747,718$, $p < .0005$. El modelo explicó el 70.0% (Nagelkerke R²) de la variación en la escalada de la crisis y clasificó correctamente el 86,4%

de los casos. Del modelo inicial propuesto, existen tres variables predictoras presentadas que son estadísticamente significativas: detonante, técnica de administración de crisis y régimen.

La varianza es explicada en el pseudo R² Cox y Snell, ello implica que el modelo posee una bondad de ajuste que permite afirmar que entre el 53% y el 70% de la variación de la variable dependiente es explicada por las variables independientes en el modelo.

La tabla de clasificación presenta un 86% de precisión en la calificación de casos que escalan o distienden, mejorando el 52% inicial en el cual no se ingresaban las variables independientes, aumentando en veinticuatro puntos porcentuales.

ALGORITMO DE CLASIFICACIÓN RANDOM FOREST

El algoritmo Random Forest es utilizado en el aprendizaje automático e inteligencia artificial, se utiliza para predecir eventos basados en gran número de datos y se usa de manera extensiva en minería de datos.

El procedimiento es crear un bosque de árboles de clasificación de datos para inferir sobre los resultados del árbol. El árbol de clasificación para la generación de un resultado utiliza una muestra de entrenamiento de N cantidad de casos. Estos se emplean para construir una regla de clasificación estructurada en el árbol. El algoritmo divide la muestra de entrenamiento en grupos cada vez más homogéneos para lograr la separación de datos y clasificarlos. El resultado es una clasificación que maximiza la heterogeneidad de los datos, es decir, identifica las semejanzas para centrarse en las diferencias (Muchsilnky et al., 2015).

Es una colección de “n” cantidad de árboles de decisión distribuidos de manera idéntica, donde cada árbol se construye utilizando el algoritmo de clasificación descrito anteriormente, usando para ellos los datos de entrenamiento. Así, se forma un árbol de clasificación completo para cada muestra.

El poder predictivo superior del Random Forest por sobre la regresión, se demuestra a través de los resultados del índice AUC (área bajo la curva) y la

puntuación F1, que es la media armónica de precisión (precisión) y recuperación o recuerdo (recall). La precisión hace referencia a cuántas clases han sido clasificados correctamente, en este caso en la asignación de si el conflicto escala o distiende.

El área bajo la curva AUC mide la discriminación, es decir, la capacidad de la prueba para clasificar correctamente si la crisis escala o distiende (según el modelo presentado). Es el porcentaje de pares clasificados aleatoriamente de forma correcta.

Como aclaración, el AUC es un resumen numérico de la probabilidad que un clasificador dado clasifique una observación positiva elegida aleatoriamente más alta que una observación negativa elegida aleatoriamente.

En resumen, cuanto mayor sea el puntaje de AUC, mejor será la precisión predictiva del modelo. Las puntuaciones de AUC en los 0,80 puntos se consideran buenas, cualquier AUC superior a 0,90 se considera un excelente clasificador (Muchlinski et al., 2015).

El índice F1 es una medida de la precisión del test, en este caso Random Forest. Considera tanto la precisión (p) como la recuperación o recall (r), de la prueba para calcular la puntuación. Finalmente, la puntuación F1 es la media armónica de la precisión y la recuperación, donde una puntuación F1 alcanza su mejor valor en 1 (precisión y recuperación perfecta) y peor en 0 (Sasaki, 2007).

En la Tabla 2 se presentan los puntajes de dos algoritmos de clasificación, Random Forest y Naive Bayes para efectos de referencia. En este caso el modelo predictivo Random Forest presenta un AUC muy alto 0,988, lo que significa que constituye un buen predictor. El valor F1 se aproxima a 1 con un valor de 0,948, lo que demuestra una media armónica alta, constituyendo un modelo robusto de predicción considerando la precisión y la recuperación.

Tabla 2. Test y puntaje de modelo Random Forest

| Test and Score | | | | | |
|--------------------|-------|-------|-------|-----------|--------|
| Evaluation Results | | | | | |
| Model ▼ | AUC | CA | F1 | Precision | Recall |
| Random Forest | 0.988 | 0.949 | 0.948 | 0.948 | 0.949 |
| Naive Bayes | 0.987 | 0.934 | 0.934 | 0.934 | 0.934 |

Fuente: Elaboración propia.

Asimismo, la prueba de precisión del algoritmo se efectuó utilizando un 70% de los datos como entrenamiento y el 30% de los datos para prueba, empleando una validación cruzada.

Como complemento, se utilizó además una matriz de confusión cuya representación gráfica las clasificaciones correctas y las incorrectas. Permite demostrar la robustez del modelo en la predicción.

Tabla 3. Matriz de Confusión

| | | Predicted | | | | Σ |
|--------|---|-----------|-----|-----|-----|------|
| | | 1 | 2 | 3 | 4 | |
| Actual | 1 | 489 | 0 | 0 | 0 | 489 |
| | 2 | 0 | 115 | 14 | 1 | 130 |
| | 3 | 0 | 14 | 198 | 17 | 229 |
| | 4 | 0 | 0 | 4 | 197 | 201 |
| | Σ | 489 | 129 | 216 | 215 | 1049 |

Fuente: Elaboración propia.

En este caso Random Forest clasifica correctamente el 100% de los casos en que no se presenta violencia (1). En la clasificación de casos de enfrentamiento menor (2) clasifica correctamente 115 casos e incorrectamente 14. Para casos de enfrentamientos mayores (3) clasifica correctamente 198 casos y clasifica erróneamente 14 casos como enfrentamientos menores y 4 casos como guerra.

Los casos de guerra (4) son clasificados correctamente en 197 oportunidades, existiendo 17 casos clasificados erróneamente como enfrentamientos mayores (3) y un caso clasificado como enfrentamiento menor (2).

En adición a lo anterior, la tabla 4 presenta una matriz de confusión con porcentajes de los casos clasificados correctamente (proporción) con un 100% de clasificación para ausencia de violencia (1), un 89,1% de clasificación correcta para enfrentamientos menores (2), un 91,7% de aciertos en enfrentamientos mayores (3) y un 91,6% de casos clasificados correctamente para guerra (4).

Tabla 4. Matriz de confusión proporcional

| | | Predicted | | | | Σ |
|--------|---|-----------|--------|--------|--------|------|
| | | 1 | 2 | 3 | 4 | |
| Actual | 1 | 100.0 % | 0.0 % | 0.0 % | 0.0 % | 489 |
| | 2 | 0.0 % | 89.1 % | 6.5 % | 0.5 % | 130 |
| | 3 | 0.0 % | 10.9 % | 91.7 % | 7.9 % | 229 |
| | 4 | 0.0 % | 0.0 % | 1.9 % | 91.6 % | 201 |
| | Σ | 489 | 129 | 216 | 215 | 1049 |

Fuente: elaboración propia.

Como se aprecia, Random Forest ofrece una mejora considerable en la capacidad de predicción de la variable dependiente en comparación a la regresión logística. Clasifica cerca del 90% de los casos correctamente, lo que otorga confianza en la proyección de escenarios y la contextualización del riesgo asociado a los posibles desenlaces.

REFLEXIONES FINALES

El análisis desarrollado precedentemente ha demostrado el valor de uso que un perfil de riesgo, como instrumento de toma de decisiones en la crisis internacional, genera en beneficio de la conducción de la autoridad política y militar.

La existencia de correlaciones estadísticamente significativas permite ganar confianza en que las variables que constituyen el modelo son apropiadas para ser empleadas en la predicción de la variable dependiente. Lo anterior significa que es factible proyectar escenarios en los que la crisis escala o distiende, constituyendo esto la mayor utilidad de un modelo de perfil de riesgo.

El algoritmo Random Forest mejora la capacidad predictiva del modelo, esto constituye un aspecto importante y evidencia la relación existente entre el uso del perfil de riesgo y el proceso de toma de decisiones debido a que la predicción de variables da poder de uso al modelo del perfil, ello se traduce en la capacidad de proyectar escenarios y asignar probabilidades pasando de un proceso decisional de incertidumbre en el cual nada se sabe de los posibles efectos o desenlace de las alternativas a uno de bajo riesgo en el cual se pueden asignar probabilidades de ocurrencia. Random Forest aumenta los resultados de la regresión logística sobre la correcta clasificación de casos que escalan o distienden a un valor cercano al 90%, incrementando la confianza que entregan los modelos para proyectar resultados.

Para contextualizar los resultados se puede determinar qué el perfil más propenso al uso de la violencia como forma de gestión de la crisis, es aquel que emplea como técnica principal de administración de la crisis los actos militares violentos, los que incrementan en un 21% la probabilidad de escalada. Los “actos políticos” incrementan en un 5% las probabilidades de escalada, y se encuentran por sobre otras formas de administración tales como acciones militares no violentas.

Por ejemplo, un régimen de gobierno de tipo “**militar indirecto**”, es decir civiles controlados por militares, administrando una crisis que ha sido desencadenada por un “**cambio externo**” a su esfera de acción, dada por cambios en la distribución de poder en el sistema, y que utiliza la “**acción militar violenta**” como forma principal de gestión o administración de la crisis. Frente a este perfil se espera que las probabilidades de escalada estén en el orden del 70%.

Cabe hacer mención que un perfil que posea un régimen militar dual, es decir completo dominio militar, que administre una crisis desencadenada por una acción militar violenta y que emplee una forma de gestión de crisis con violencia múltiple incluida la militar, tiene un 21% de probabilidades de escalar la crisis. Esto es relevante ya que resulta contra intuitivo el que un perfil de características más complejo, con autoritarismos militares en control completo, que utilice formas violentas de gestión, tenga un 49% de probabilidades menos que el perfil más propenso a escalar la crisis descrito en el párrafo anterior. Lo anterior dice mucho de la compleja interacción que existe entre las variables que se asocian a la crisis, y también existe incidencia directa en ese 30% restante que el modelo no es capaz de explicar. Cabe hacer mención que correlaciones no implican causalidad, por lo que el modelo propuesto solo proporciona probabilidades de escalada o distensión, pero bajo ninguna circunstancia implica que las variables propuestas constituyan causales de la escalada o distensión, solo expresan una relación.

Finalmente, la utilidad de un perfil de riesgo, como instrumento de toma de decisiones en la crisis internacional, se identifica en función de poder proyectar escenarios probables a base de las variables que constituyen el modelo. Así, el uso de técnicas estadísticas y algoritmos de aprendizaje automático, constituyen herramientas de gran utilidad para el apoyo al proceso de toma de decisiones, generando un juicio informado y opciones que reducen en alguna medida la incertidumbre propia de las crisis.

REFERENCIAS

Brecher, M., y Wilkenfeld, J. (2000). *A study of crisis*. Ann Harbor: University of Michigan Press.

Gallardo, A. (2005). *Crisis internacionales en Sudamérica*.

King, Gary; Keohane, Robert; Verba, Sidney (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*.

- Muchlinski, David et al. (2015). Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data. *Political Analysis* (2016) 24:87–103. doi:10.1093/pan/mpv024
- Nye, J. (1990). Soft Power. *Foreign Policy*. (80), pp. 153-171. Doi: 10.2307/1148580
- Ostrowska, M., & Mazur, S. (2015). Risk in a crisis situation. *Procedia Economics and Finance*. 23, pp. 1054–1059. [http://doi.org/10.1016/S2212-5671\(15\)00373-1](http://doi.org/10.1016/S2212-5671(15)00373-1)
- Resnik, M. (1987). *Choices: An Introduction to Decision Theory*. University of Minnesota Press. www.jstor.org/stable/10.5749/j.ctttshgd
- Sasaki, Y. (2007). *The truth of the F-measure*. School of Computer Science, University of Manchester MIB, 131 Princess Street, Manchester, M1 7DN.
- Verdugo, J. C. (2011). Una estructura para la asesoría en el manejo de crisis internacionales: Caso Nacional. *Academia Nacional De Estudios Políticos y Estratégicos*, pp. 1-97.