
Predicción de golpes de estado en el siglo xxi

José Saint Germain
joesg998@gmail.com

- 1 **Especialización en exploración de datos y descubrimiento del co-**
- 2 **nocimiento**
- 3 **Facultad de Ciencias Exactas y Naturales**
- 4 **Universidad de Buenos Aires**



5 Índice

6	1. Introducción	3
7	1.1. Motivación	3
8	1.2. Estructura del documento	3
9	2. Marco Teórico y estado del arte	3
10	2.1. La Teoría de la Modernización y sus variantes	3
11	2.2. Teoría de la Dependencia y el Subdesarrollo	4
12	2.3. Estado burocrático autoritario	5
13	2.4. Enfoques empíricos	6
14	2.5. Conclusión	6
15	3. Metodología	7
16	3.1. Algoritmos de predicción	7
17	3.2. Métrica de evaluación	7
18	3.3. Optimización de hiperparámetros	7
19	3.4. Block-time-series cross-validation	8
20	3.5. Valores Shapley	8
21	3.6. Ingeniería de atributos	8
22	3.7. Análisis Exploratorio de Datos	8
23	4. Resultados y discusión	12
24	4.1. Performance de los modelos	12
25	4.2. Análisis de variables	14
26	4.3. Análisis de casos particulares	21
27	4.4. Discusiones	23
28	5. Conclusiones	23
29	6. Anexo	23
30	6.1. Código	23
31	6.2. Golpes	23
32	6.3. Gráficos y tablas adicionales	24

33 1. Introducción

34 El objetivo de este trabajo es lograr entrenar un modelo de aprendizaje automático que logre predecir
35 la presencia de golpes de estado durante los años 2020 a 2022 en todos los países del mundo a
36 partir de la utilización de la base de datos provista por la fundación Varieties of Democracy (V-Dem)
37 (Coppedge, Gerring, Knutsen, Lindberg et al., 2024). A su vez, se busca comparar la performance
38 de los algoritmos entrenados con los utilizados por el artículo realizado por Cebotari et al (Cebotari
39 et al., 2024), el cual utilizó una mayor variedad de fuentes. Adicionalmente, este trabajo busca tener
40 una noción acabada de las variables más importantes que los algoritmos utilizan para la predicción de
41 la variable objetivo, aprovechando que la base de datos cuenta mayoritariamente solo con variables
42 ligadas a lo político-institucional.

43 1.1. Motivación

44 La motivación de este trabajo es dialogar con el artículo recientemente realizado por el Fondo
45 Monetario Internacional (FMI) Cebotari et al., 2024. En el mismo se aborda el mismo objeto de
46 estudio utilizando diversas metodologías, siendo una de ellas la utilización de algoritmos de aprendizaje
47 automático. En este trabajo se replicó la metodología utilizada en esa sección; comparando los mismos
48 modelos, sus respectivos hiperparámetros y la métrica a maximizar durante su entrenamiento

49 La principal diferencia entre el paper del organismo y este trabajo radica en el origen de los datos.
50 Por un lado, el artículo del FMI utilizan 14 fuentes provenientes de diferentes organismos, de manera
51 de cubrir 5 grupos de variables sobre diferentes ámbitos (Desarrollo y demografía, Inclusión y
52 gobernanza, macroestabilidad, políticas públicas, estabilidad sociopolítica). En cambio, este trabajo
53 utilizará solamente la base de datos v-dem por dos motivos: en primer lugar, para abarcar solamente
54 variables que estén directamente ligadas a la situación política e institucional de los países, excluyendo
55 en la medida de lo posible atributos ajenos a este ámbito. En segundo lugar, para realizar una
56 comparación con las nutridas y variadas fuentes del artículo citado. De esa manera, podemos tener
57 una noción del poder predictivo de atributos puramente político-institucionales frente a un abanico
58 más diverso de variables.

59 1.2. Estructura del documento

60 - Explicación de la estructura del trabajo (se realizará una vez que esté completado)

61 2. Marco Teórico y estado del arte

62 El estudio de los golpes de estado, así como los procesos de democratización han sido una preocupa-
63 ción central para la ciencia política moderna durante el siglo xx. Diversas teorías se han desarrollado
64 de manera de aprehender los causales de la democratización de un país así como de su proceso
65 inverso, ya sea una erosión democrática gradual o un golpe de estado autoritario; así como los
66 elementos sociales, culturales e institucionales que pueden evitar o disminuir la probabilidad de que
67 se produzcan estos fenómenos.

68 2.1. La Teoría de la Modernización y sus variantes

69 Uno de los primeros marcos para comprender la inestabilidad política que llevaba a un golpe
70 institucional fue la teoría de la modernización, popularizada a mediados del siglo xx. Entre los
71 exponentes de esta teoría se encuentra Seymour Martin Lipset quien con su artículo "*Some social
72 requisites of democracy: economic development and political legitimacy*" (1959). Desde un enfoque
73 sociológico, argumenta que el grado de desarrollo económico de una sociedad es una condición
74 necesaria para el nacimiento y consolidación de un régimen democrático, principalmente porque
75 una sociedad dividida entre una masa empobrecida y una élite rica es más propensa a generar una
76 oligarquía (dictadura del estrato superior de la sociedad) o una tiranía (dictadura basada en el estrato
77 inferior).

78 Para medir el desarrollo económico, Lipset analiza y desagrega cuatro variables: el nivel de riqueza,
79 medido por pbi per cápita y por la cantidad de personas con vehículos de motor, radios, teléfonos y
80 diarios cada mil personas; el grado de industrialización, medido por el porcentaje de trabajadores

81 hombres en la agricultura y el nivel de energía utilizado per cápita (en toneladas de carbón); el
82 nivel de urbanización, medido en índices realizados previamente; así como el nivel educativo de la
83 población, del cual toma principalmente la tasa de alfabetización. El autor subraya este último factor,
84 exponiendo que si no es una condición suficiente para la democracia, es una condición necesaria.

85 A su vez, Lipset describe cambios subyacentes en los diversos estratos sociales producto del desarrollo
86 económico. En primer lugar, se desarrolla una suerte "lucha de clases" por parte de la clase baja, ya
87 que mayores tasas de alfabetización y bienestar económico genera una visión más largoplacista y
88 compleja de la política, desarrollando una ideología secular reformista y gradualista en la clase obrera.
89 En segundo lugar, una clase media fortalecida y ensanchada por el crecimiento económico juega un
90 papel mitigador del conflicto, penalizando extremismos y apoyando movimientos más moderados y
91 democráticos. Por último, en una sociedad en donde las diferencias económicas entre clases sociales
92 se moderan, se atenúan las percepciones negativas de las clases altas hacia las bajas, volviéndolas
93 más tolerantes a compartir el poder y a otorgar derechos al resto de la sociedad. Por último, en una
94 sociedad con mayor riqueza económica se expande la presencia de organizaciones intermedias e
95 instituciones como fuentes de contrapeso al poder.

96 Si bien el desarrollo económico, caracterizado en los párrafos anteriores, se torna una condición
97 mínima para la consolidación democrática, Lipset subraya dos condiciones suficientes para lograr
98 su estabilidad en el tiempo: la efectividad del sistema político -entendida como la performance del
99 sistema político para resolver problemas- y la legitimidad -es decir, la capacidad de lograr la creencia
100 de que la existencia de instituciones políticas es deseable para el conjunto de la sociedad. Una crisis de
101 legitimidad, por lo tanto, es contemplada como un factor de inestabilidad para un sistema democrático.
102 Este tipo de crisis, según el autor, pueden surgir de determinados cambios en la estructura social:
103 cuando todos los grupos mayoritarios no se aseguran el acceso al sistema político de manera temprana
104 en un período de transición, o cuando el estatus de las instituciones conservadoras es amenazado.

105 Una variante de la teoría de la modernización fue planteada por Samuel Huntington en *Political*
106 *Order in Changing societies* (1968), quien mueve el foco de lo social hacia lo político. Para el
107 autor, el crecimiento económico acelerado puede generar tensiones y conflictos que desafían la
108 estabilidad política. En el contexto de Guerra Fría en que Huntington escribe esto, sostiene que esta
109 inestabilidad puede ser aprovechada por la política revolucionaria impulsada por los comunistas.
110 Por eso, considera necesaria una intervención (generalmente a través de las Fuerzas Armadas)
111 para controlar esa inestabilidad y lograr construir instituciones políticas que manejen las tensiones
112 asociadas al proceso de modernización. En este sentido es crítico a la teoría de Lipset, puesto que no
113 piensa que la estabilidad política es una consecuencia natural e inevitable del desarrollo económico y
114 de las reformas sociales. Esto se logrará si están combinadas con oportunidades de movilidad social y
115 económica ascendente e instituciones políticas flexibles por las cuales se canalice el aumento de la
116 participación.

117 **2.2. Teoría de la Dependencia y el Subdesarrollo**

118 Como contraposición a la teoría de la modernización, para analizar las tendencias de desarrollo y
119 autocratización de naciones del tercer mundo, se desarrolló la denominada teoría de la dependencia.
120 En sus distintos enfoques, la teoría de la dependencia explica que el atraso relativo de América Latina
121 y el desarrollo de las economías centrales (fundamentalmente Estados Unidos y Europa Occidental)
122 no son independientes sino complementarios. Estos procesos están vinculados por su inserción en
123 la economía mundial, el cual desfavorece a los exportadores de materias primas e importadores de
124 productos manufacturados, favoreciendo la extracción de sus recursos e inhibiendo el desarrollo de
125 sus economías.

126 La variante más extendida de esta teoría fue formulada por Fernando Henrique Cardoso y Enzo
127 Faletto en *Dependencia y Desarrollo en América Latina* (1979). Allí, matizan las aseveraciones de la
128 teoría, indicando que la inserción de las economías latinoamericanas en la economía internacional
129 no determina su trayectoria sino que incide a través de la estructura social y económica asociada a
130 un tipo de actividad de exportación (en América Latina: agrícola, ganadera y minera). Esta relación
131 de dependencia está conformada por una red de intereses y de coacciones que ligan unos grupos
132 sociales a otros. Allí, el puente de las sociedades latinoamericanas con el capital extranjero es el
133 sector exportador de materias primas. En diversas medida y forma, este sector logra insertarse en el
134 mercado mundial a la vez que logra mantener el control sobre la sociedad local, ya sea imponiéndose
135 o bien negociando con sectores mercantiles internos.

136 En los casos donde los sectores internos lograron cierto espacio de desarrollo, se generaron nuevos
137 grupos sociales (artesanos, pequeños comerciantes, profesionales, sectores vinculados a los servicios,
138 entre otros). En función de ese mercado, se constituyen los primeros núcleos industriales, y se forman,
139 en consecuencia, tanto una burguesía urbana como sectores obrero-populares; así, en un primer
140 momento, los grupos sociales urbano-industriales se constituyen siguiendo la expansión del sector
141 exportador y sin que sus intereses económicos se opongan a los de éstos, sino que, por el contrario,
142 pasan a ser un sector complementario de aquél. En cambio, en los países con predominio de economía
143 de enclave, en donde los sectores exportadores tuvieron primacía total sobre los sectores internos, no
144 se generaron sectores medios. Allí, la relación de subordinación política de los grupos dominantes y, a
145 partir dellos, de las empresas extranjeras se da de manera más directa sobre los obreros y campesinos.
146 En este tipo de países, se logró una preocupación sobre políticas centradas en el mercado interno
147 cuando ya existía una clase media previa a la inserción en el mercado mundial (como en Chile) o
148 cuando los sectores medios lograron insertarse de manera revolucionaria (mediante golpes de estado)
149 dentro del aparato del Estado y lo utilizaron para crear una economía nacional (México y Venezuela).

150 En ambas situaciones, en los momentos en que los intereses de los sectores exportadores fueron
151 puestos en tela de juicio fue cuando la inestabilidad política se hizo presente, derivando en algunos
152 casos en golpes de estado por parte de las fuerzas armadas. En las naciones con economías de enclave
153 sucedió, por ejemplo, con la crisis económica de 1930 ante la falta de respuestas del modelo al
154 aumento del desempleo y a la falta de respuestas por parte del Estado (como si ocurrió en con sectores
155 internos más robustos). En cambio, en los países con sectores medios más fuertes, la inestabilidad
156 política emergió varios años después. Allí gobiernos de corte populista lograron utilizar el estado
157 para fortalecer la industria nacional y los sectores medios, y los golpes de estado se expresaron como
158 una búsqueda del sector agroexportador de volver a imponer su modelo vinculado estrechamente con
159 el mercado global.

160 En definitiva, Cardoso y Faletto aportan una teoría más compleja al incluir factores históricos,
161 coyunturales y productivos en la trayectoria de cada uno de los países de América Latina; evidenciando
162 que sus momentos de inestabilidad están fuertemente ligados a los procesos previos de incorporación
163 al mercado mundial. Una de las principales críticas es su excesivo foco en América Latina, puesto
164 que hay casos de países que lograron un desarrollo exitoso rompiendo el ciclo de dependencia, siendo
165 el mayor contra ejemplo los llamados "tigres asiáticos".

166 2.3. Estado burocrático autoritario

167 Desde un ángulo diferente, Guillermo O'donnell también propició algunas críticas a la teoría de
168 la modernización observando los procesos en países de América del Sur. En *Modernización y*
169 *autoritarismo* (1972), O'donnell sostiene que la modernización económica no necesariamente lleva a
170 la democratización política. En su lugar, puede llevar a la consolidación de regímenes autoritarios
171 burocráticos, en los cuales el poder político está concentrado en las fuerzas armadas y en la burocracia
172 estatal. En estos regímenes, la participación política está restringida y la oposición es reprimida, pero
173 a diferencia de los regímenes totalitarios, la sociedad civil y la economía pueden ser relativamente
174 autónomas.

175 Estos estados burocráticos autoritarios, cuyos ejemplos más claros encontró en los regímenes militares
176 de Brasil desde 1964 y Argentina entre 1966 y 1973, surgen a partir de los límites macroeconómicos
177 encontrados por la industrialización por sustitución de importaciones, impulsado por una coalición
178 social formada por una burguesía industrial focalizada en el sector de bienes de consumo y los
179 sectores medios urbanos. La industrialización impulsada por la crisis del 30 apuntó fundamentalmente
180 a satisfacer la demanda de bienes de consumo faltantes por la depresión y la posterior guerra mundial.
181 Este desarrollo no trajo consigo una ampliación en la producción de bienes intermedios y de capital,
182 lo cual produjo que esta etapa de la industrialización venga aparejada de una fuerte necesidad de
183 divisas para importación de bienes de capital, lo cual lleva en el mediano plazo a una crisis en la
184 balanza de pagos, una estructura productiva distorsionada y altas expectativas de consumo.

185 Este proceso llevó a la conclusión de que para poder continuar con un desarrollo industrial era
186 necesario la postergación de las demandas de participación en el consumo así como en el poder
187 político del sector popular urbano. Esta conclusión generó un péndulo en la configuración de las
188 coaliciones sociales: el sector empresario industrial y los sectores agroexportadores se unieron para
189 bloquear la participación de sectores populares mediante la instauración de un régimen autoritario
190 que sea administrado por una tecnocracia capacitada.

191 El trabajo de O'donnell significó un contrapunto importante a las teorías precedentes y contempo-
192 ráneas a su tiempo, logrando evidenciar como un acelerado proceso de modernización económica
193 puede derivar a golpes de estado que generen autoritarismos; aunque, a diferencia de la teoría de la
194 dependencia, no se da directamente por la influencia del mercado mundial en las élites sino por la
195 búsqueda de las mismas de generar un desarrollo económico por vía autoritaria. A pesar de la enorme
196 precisión para describir los procesos autoritarios de Brasil y Argentina en la década de 1960, la teoría
197 de O'donnell encontró serios problemas para explicar procesos de autoritarismo futuros, tanto en la
198 misma Argentina en 1976 como la de Chile en 1973, en dónde la coalición y el enfoque económico
199 de los golpes de estado estaban más relacionado con un enfoque neoliberal y de libre mercado más
200 que con un desarrollismo por vía autoritaria.

201 **2.4. Enfoques empíricos**

202 Casi medio siglo desde las primeras publicaciones de Lipset, Przeworski et al (2000) a partir de
203 una exhaustiva recolección de datos sobre el devenir histórico de los distintos países entre 1950 y
204 1990, los autores llegan a una serie de conclusiones sobre la relación entre desarrollo económico y
205 democracia. En primer lugar, consideran que si bien el desarrollo económico per se no conduce a una
206 democratización, si es verdad que las democracias en países en vías de desarrollo son mucho más
207 frágiles e inestables que en los países desarrollados. En segundo lugar, las democracias no producen
208 una disminución de la inversión en el país; en especial si es un país pobre, ya que para los autores no
209 hay mucho que el Estado pueda hacer para modificar esa tendencia. Por lo tanto, no hay evidencia que
210 indique que haya que sacrificar la democracia para alcanzar el desarrollo: los países que lo lograron
211 podrían haberlo hecho tanto en una democracia como en una dictadura.

212 Por último, los autores utilizan la información recolectada para predecir la situación de las democra-
213 cias y autoritarismos para el año 2030. Por un lado, afirman que tanto el pbi per cápita va a aumentar
214 (2,5 veces mayor al de 1990) como que las dictaduras van a ser casi inexistentes. De todas formas,
215 algunas pocas dictaduras prevalecerán en algunos países pobres, especialmente en África; así como
216 también predicen para este continente que será el único lugar donde seguirán aconteciendo conflictos
217 bélicos.

218 **2.5. Conclusión**

219 A lo largo de la historia de la ciencia política del siglo xx y xxi se observa una preeminencia de
220 la temática asociada a la democratización y al desarrollo. Cada autor desde su perspectiva teórica
221 así como del análisis de experiencias históricas ha intentado rescatar los factores que habilitan y
222 fortalecen un sendero democrático, así como aquellos que impiden su consolidación. Haciendo un
223 recorrido por todos los autores tratados en esta sección podemos rescatar que la eficiencia del Estado
224 para resolver problemas políticos así como su legitimidad son factores relevantes para la estabilidad
225 de un régimen democrático. Adicionalmente, comprender el contexto socioeconómico y la trayectoria
226 histórica de un país se vuelve fundamental para comprender los límites que encuentran los países
227 periféricos para alcanzar el desarrollo económico y político. Finalmente, los aportes de Przeworski
228 et al son críticos para no dejarse llevar por asentados axiomas que determinan la postergación
229 democrática en pos de un desarrollo económico rápido.

230 Como cierre, es importante destacar que si bien los golpes de estado rondan permanentemente en el
231 desarrollo de estas teorías, no figuran estudios relevantes que se hayan concentrado exclusivamente
232 en su estudio, así como en sus motivadores. Es por eso que el artículo realizado por Cebotari et al
233 (2024) nos ha llamado la atención, ya que es un punto de partida importante para el desarrollo del
234 estudio de golpes de estado que esté prudentemente separado de las teorías antes descritas; no solo
235 para poder concentrar los esfuerzos en el estudio del hecho en sí sino también para abarcar no solo
236 los golpes de estado hacia regímenes democráticos, sino también hacia regímenes autoritarios en
237 todas sus variantes. En la siguiente sección se desarrollará la metodología con la que se guiará el
238 trabajo, utilizando en buena medida las técnicas expuestas por el artículo del Cebotari et al.

3. Metodología

Puesto que buscamos reproducir el mismo trabajo realizado por el FMI (Cebotari et al., 2024) con diferentes datos, vamos a utilizar las mismas técnicas de optimización de hiperparámetros, así como los mismos algoritmos de entrenamiento y de interpretación de resultados.

3.1. Algoritmos de predicción

Los algoritmos que se utilizarán serán Random Forest (Breiman, 2001) y XGBoost (Chen y Guestrin, 2016). Ambos algoritmos son modelos de ensamble basados en múltiples árboles de decisión. Un árbol de decisión individual es un modelo predictivo que divide los datos en subconjuntos cada vez más pequeños basándose en una serie de decisiones binarias sobre las características de los datos. En cada nodo del árbol, se selecciona una característica y un umbral para dividir los datos en dos grupos: aquellos que cumplen la condición y aquellos que no. Este proceso se repite de manera recursiva hasta que se alcanza una condición de parada, ya sea un mínimo de muestras en un nodo o una profundidad máxima del árbol.

El algoritmo Random Forest (bosque aleatorio) busca combinar múltiples árboles de decisión con características disímiles, combinando sus predicciones mediante un promedio (en regresión) o mediante votación (en clasificación). La variedad de árboles se logra mediante una selección aleatoria de un subconjunto de los datos con remplazo, así como seleccionando una proporción aleatoria de atributos del dataset. De esa manera, se reduce la varianza del modelo, se evita el sobreajuste y se mejora la capacidad predictiva.

Por otro lado, XGBoost (Extreme Gradient Boosting) es un algoritmo de boosting que mejora las predicciones combinando múltiples árboles de decisión débiles (de menor capacidad predictiva) de manera secuencial. A diferencia de Random Forest, donde los árboles se entrenan de forma independiente, en el boosting los árboles se entrenan uno tras otro, cada uno tratando de corregir los errores cometidos por los árboles anteriores. Particularmente, XGBoost utiliza la técnica de gradient boosting, donde cada árbol nuevo se ajusta a los residuos (errores) del modelo anterior utilizando el gradiente del error. Adicionalmente, XGBoost incluye algunas mejoras como la regularización y el manejo eficiente de datos faltantes.

3.2. Métrica de evaluación

Adicionalmente, para la evaluación de performance se utilizará el área bajo la curva ROC (AUC). La curva ROC es construida trazando la tasa de verdaderos positivos (la sensibilidad) frente a la tasa de falsos positivos (especificidad) en diferentes umbrales de decisión. El área total de esta curva es la que se utilizará para evaluar la performance del modelo. Esta métrica toma valores entre 0.5 y 1. Un valor de AUC de 0.5 indica que el modelo no tiene mayor capacidad predictiva que el puro azar, mientras que un valor cercano a 1 indica que el modelo es un excelente predictor. Las ventajas de esta métrica son que es insensible al desbalance de clases y que proporciona una evaluación única del rendimiento del modelo en distintos umbrales de decisión.

3.3. Optimización de hiperparámetros

Con respecto al ajuste de hiperparámetros se utilizará la optimización bayesiana. Este algoritmo goza de varias ventajas con respecto a sus contrapartes grid search y random search. Mientras que estos últimos recorren los hiperparámetros sin prestar atención al puntaje que obtiene la métrica a optimizar, la optimización bayesiana sí lo hace (en este caso, el área bajo la curva ROC) e intenta mejorar ese puntaje en cada iteración. Así, recorre el espacio de hiperparámetros de manera más eficiente, probando combinaciones de hiperparámetros que permitan minimizar o maximizar la métrica. Además, por cómo está desarrollado este algoritmo, en los hiperparámetros con decimales, puede explorar cualquier tipo de valores entre un rango provisto por el usuario, obteniendo valores precisos de los hiperparámetros. Esta optimización bayesiana consistirá en 100 iteraciones en donde se buscará el valor óptimo de los siguientes hiperparámetros:

- Random Forest: profundidad máxima de los árboles (max_depth) y la submuestra del ratio de columnas a considerar cuando se construye cada árbol (max_features).

288 ■ XGBoost: la tasa de aprendizaje (`learning_rate`) y el término de regularización L2 en los
289 pesos (`reg_lambda`).

290 Adicionalmente el parámetro que establece la cantidad de árboles creados (`n_estimators`) quedará
291 fijado en 1000.

292 **3.4. Block-time-series cross-validation**

293 Para evitar el data leakage -es decir, incorporar en el conjunto de entrenamiento información que no
294 vamos a tener en futuro- en cada iteración de la optimización bayesiana se utilizará la validación
295 cruzada. Sin embargo, como se trabajará con una base de datos de panel, conviene utilizar una versión
296 adaptada: el método *block-time-series cross-validation*, basado en Burman et al., 1994 y Racine,
297 2000. El método aplicado en este caso consiste en generar 5 pares de entrenamiento y validación:
298 1970 - 2009, 2010 - 2011; 1970 - 2011, 2012 - 2013; 1970 - 2013, 2014 - 2015; 1970 - 2015, 2016 -
299 2017; 1970 - 2017, 2018- 2019. Por lo tanto, cada set de entrenamiento consiste en observaciones
300 desde 1970 hasta un año de corte (2009, 2011, 2013, 2015, 2017) y el set de validación contempla los
301 dos años siguientes del mismo. Una vez realizada la optimización bayesiana, se toman los valores de
302 hiperparámetros que lograron maximizar el AUC y se entrena el modelo con el set de entrenamiento
303 para intentar predecir los golpes de estado entre 2020 y 2022.

304 **3.5. Valores Shapley**

305 Para interpretar las variables más importantes en la predicción de golpes de estado, se utilizarán
306 los valores Shapley (Strumbelj y Kononenko, 2010; Lundberg y Lee, 2017). Basado en la teoría de
307 juegos, los valores Shapley consideran todas las posibles coaliciones de características y calculan
308 la contribución promedio de cada característica a través de todas las permutaciones posibles. En
309 otras palabras, determinan cuánto contribuye cada característica al valor de predicción del modelo,
310 considerando la interacción entre las características y evitando atribuciones injustas o redundantes.
311 Los valores Shapley proporcionan una forma intuitiva y sólida de interpretar y entender cómo las
312 características individuales afectan las decisiones del modelo, lo que los hace valiosos para explicar
313 modelos de aprendizaje automático complejos.

314 **3.6. Ingeniería de atributos**

315 Para dotar de mayor información a los algoritmos a la hora de predecir la variable objetivo, se crearon
316 nuevas variables a partir de las ya existentes. Fundamentalmente, se generaron variables llamadas
317 "lag" que toman el valor que obtuvo cierto país una x cantidad de años atrás. En este caso específico
318 se generaron lags para 1, 5 y 10 años anteriores. De esa manera, los algoritmos tienen algo más de
319 información sobre la tendencia temporal de las variables. Adicionalmente, se agregaron variables
320 binarias que informan sobre la región a la que pertenecen los países, especulando con que estas
321 variables pueden llegar a tener importancia si una región específica cuenta con muchos golpes de
322 estado en un momento determinado. Finalmente, se excluyeron todos los grupos de variables que
323 provengan de fuentes externas, con el objetivo de tener la certeza de contar con la mayoría de las
324 mismas en caso de querer repetir este experimento en años futuros; así como también se excluyeron
325 variables que no cuentan con información para ningún país en cierto punto de la serie (por ejemplo,
326 las variables históricas, que trabajan con datos anteriores al siglo xx).

327 **3.7. Análisis Exploratorio de Datos**

328 Como primera aproximación a la base de datos de Varieties of Democracy o V-Dem (Coppedge,
329 Gerring, Knutsen, Lindberg et al., 2024), pasaremos a explicar la manera en que se construye la misma.
330 Las variables centrales se obtienen a partir de encuestas suministradas a expertos sobre los distintos
331 países. Inicialmente, se busca que cada país cuente con al menos cinco expertos. Actualmente, la
332 institución cuenta con 22 expertos promedio por país y 7,1 expertos por combinación de variable
333 y país. Una vez obtenida las respuestas de los expertos, se pasa al proceso de agregación para así
334 conformar una base de datos donde cada fila corresponda a un país en un año específico. De esta
335 agregación obtienen diferentes versiones de la misma variable:

- 336 ■ Estimador del modelo (Variable sin sufijo): es la medida recomendada para su análisis.
337 Corresponde a obtener la mediana del valor de la variable entre los expertos, reescalado a
338 valores entre -5 a 5.
- 339 ■ Medidas de incertidumbre (*_codelow y *_codehigh): corresponden a un desvío estándar
340 por encima y por debajo del estimador del modelo. Usadas conjuntamente, construyen un
341 intervalo de confianza del 95 %.
- 342 ■ Escala original (*_osp): mediana de la variable para todos los expertos, pero sin reescalar.
343 Esta versión también cuenta con sus medidas de incertidumbre correspondientes.
- 344 ■ Desvío estándar (*_sd): desvío estándar de la variable.
- 345 ■ Media simple (*_mean): media de la variable para todos los expertos.
- 346 ■ Cantidades de expertos (*_nr): cantidad de expertos que respondieron por país, año y
347 variable.

348 Podemos mencionar que la base cuenta con 27734 filas y 4607 columnas. Como es una base de datos
349 de panel, se tiene información de 202 países durante 235 años. Las variables cuentan con un tipo de
350 codificación particular que permite identificar el origen de la variable. En primer lugar, el primer
351 prefijo es indicativo de si fue producido por V-Dem o no:

- 352 ■ v2: variables de V-Dem.
- 353 ■ v3: variables pertenecientes a la base V-Dem histórica.
- 354 ■ v2x_: Índices principales e índices componentes.
- 355 ■ v2x[indicador de dos letras]: Índices específicos de ciertas áreas (ver más abajo).
- 356 ■ e_: variables no generadas por V-Dem y variables V-Dem en versión ordinal.

357 El nombre de la variable también permite identificar la área temática a la que pertenece:

- 358 ■ ca: Espacio cívico y académico
- 359 ■ cl: Libertad civil
- 360 ■ cs: Sociedad civil
- 361 ■ dd: Democracia directa
- 362 ■ de: Demografía
- 363 ■ dl: Deliberación
- 364 ■ el: Elecciones
- 365 ■ ex: Ejecutivo
- 366 ■ exl: Legitimación
- 367 ■ ju: Poder judicial
- 368 ■ leg: Legislatura
- 369 ■ lg: Legislatura
- 370 ■ me: Medios de comunicación
- 371 ■ pe: Igualdad política
- 372 ■ ps: Partidos políticos
- 373 ■ sv: Soberanía
- 374 ■ st: Estado
- 375 ■ x: Índice (calculado a partir de variables que también se incluyen en la base de datos)
- 376 ■ zz: Cuestionario posterior a la encuesta
- 377 ■ ws: Encuesta de sociedad digital

378 A la base original obtenida desde la librería de V-Dem, se le realizaron los siguientes filtros: en primer
379 lugar, se removieron todas las variables que no sean las principales, es decir, que no cuenten con
380 sufijo. De esa manera, se busca reducir el tamaño de la base y así poder agregar nuevas columnas
381 mediante ingeniería de atributos. En segundo lugar, se filtraron los años superiores a 1950, para
382 adecuarnos al periodo utilizado en el artículo del FMI. De esa manera, la base filtrada cuenta con
383 12208 filas y 1460 columnas. Por último, se remueven todas las variables de fuentes externas (cuyo
384 agrupador comienza con 'e'), las variables pertenecientes a la base histórica (agrupador 'hist') y las
385 de la encuesta de sistema de partidos políticos; en parte debido a que provienen de fuentes ajenas
386 a V-Dem que pueden comprometer la completitud futura de los datos y en parte porque algunas de
387 estas variables cuentan con alta tasa de nulos.

388 Realizando un análisis generalizado de los distintos grupos de variables de la base de datos, podemos
389 aprehender ciertos patrones sobre la presencia de nulos: En primer lugar, observamos variables que,
390 anteriormente a un año puntual, no cuentan con información. En este ejemplo caen las variables
391 sobre gobernanza otorgadas por el banco mundial (e7), las preguntas pertenecientes a la encuesta
392 de sociedad digital (wsmcio), variables referentes a la libertad en medios digitales (wsmdmf), las
393 referentes a la polarización en medios online (wsmpomp) y las referentes a clivajes sociales (wsmsc).

394 En segundo lugar, figuran casos contrarios, en donde a partir de determinado año la cantidad de datos
395 faltantes salta a la totalidad de los casos. En este grupo figuran las variables asociadas a instituciones
396 y eventos políticos (e13), cuya fuente es un artículo de Przeworski de 2013; las variables cuya
397 fuente es la base de datos polity V (e14); las variables sobre educación (aumentan los nulos en
398 algunas variables) (eb1); las variables sobre recursos naturales (eb5), cuya fuente tiene datos hasta
399 2006; las variables sobre infraestructura (eb6); y las relacionadas a conflictos (eb8). En general, esta
400 discontinuidad sucede debido a que la información de estas variables provienen de fuentes externas
401 no gestionadas por V-Dem, las cuales finalizaron su serie en un año puntual. Por último figuran los
402 grupos de variables asociados a la base de datos histórica de v-dem (las que comienzan con hist), lo
403 cual es lógico puesto que esta base busca tomar datos previos a 1900.

404 Consecuentemente, quitaremos del grupo de variables a utilizar aquellas que sean de fuentes externas,
405 ya que de esa manera podemos asegurarnos que contaremos con todas las variables para predecir
406 golpes de estado en años futuros. También quitamos las variables provenientes de fuentes históricas y
407 de las encuestas de sociedad digital, ya que no cuentan con información para toda la serie.

408 Haciendo foco en la variable objetivo, es importante aclarar que en este trabajo no estamos contando
409 la cantidad precisa de golpes de estado sucedidos en un período de tiempo, sino que simplemente
410 relevamos si al menos un golpe de estado sucedió en un país y año determinado. Por lo tanto, si
411 un país sufrió más de un golpe de estado en un año, el mismo será contabilizado una sola vez.
412 Adicionalmente, en este trabajo también se consideran los golpes de estado que no fueron exitosos,
413 es decir, que no lograron derrocar al gobierno en cuestión. De allí se desprende que países como
414 Argentina, que en total ha tenido seis golpes de estado exitosos, aparezca con el doble de golpes en la
415 figura 1. Esta variable fue obtenida (al igual que en el artículo del FMI) de la base de datos realizada
416 por Powell y Thyne (2011). En el artículo que respalda esta base, definen a un intento de golpe de
417 Estado como *intentos ilegales y manifestos por parte de las fuerzas armadas u de otras elites dentro*
418 *del aparato estatal de derrocar al gobierno en ejercicio* (Powell y Thyne, 2011, p. 252).

419 Para realizar un paneo general de la variable objetivo, es decir, la presencia de golpes de estado a
420 lo largo de los años, generamos un conteo y lo visualizamos en un planisferio. Destacamos que la
421 mayor presencia de golpes se encuentra en el continente africano, en América del Sur y parte del
422 Caribe, Medio Oriente y el Sudeste Asiático, con algunos casos de apenas un golpe en España, Rusia,
423 Ucrania y Corea del Sur; así como dos y tres golpes en Grecia y Portugal, respectivamente.

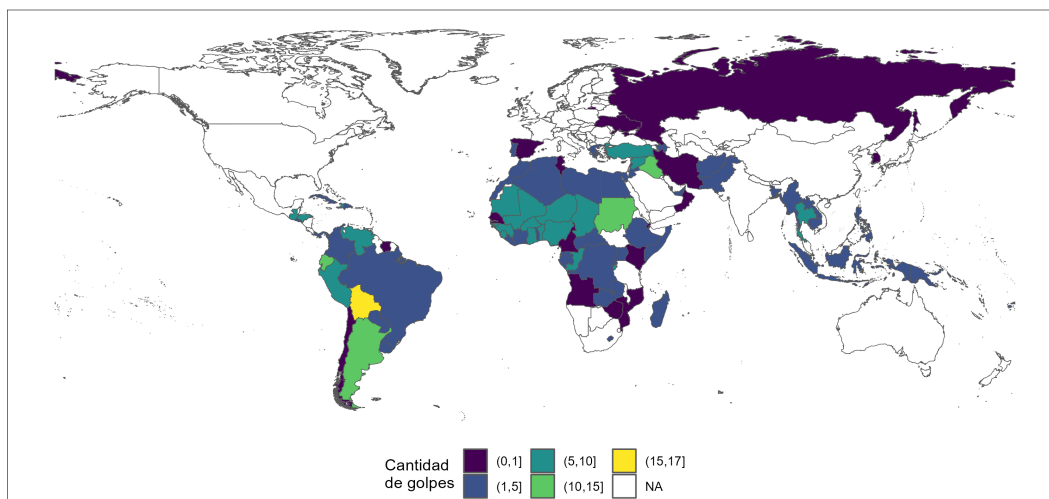


Figura 1: Golpes de estado período (1950-2023) Fuente: Powell y Thyne, 2011

424 Con mayor precisión, observamos que la región del Sahel se destaca con respecto a sus vecinos
 425 africanos. Los países en donde más golpes de estado se han producido son Bolivia (17), Sudán (14),
 426 Argentina (13), Ecuador (11), Iraq (11), Siria(11), Guatemala (10) y Tailandia (10).

427 Desagregando por década se observan algunos cambios, así como la persistencia en algunas regiones.
 428 La región del Sahel y varias naciones circundantes fueron persistentemente afectadas por golpes
 429 de estado desde los años 60. En América del Sur, en cambio, la presencia casi total de situaciones
 430 golpistas en la región se fue acotando a partir de los años 80 hasta finalmente desaparecer en el siglo
 431 xxi. Para observar con más detalle y discriminado por años y países se puede ver la figura 13.

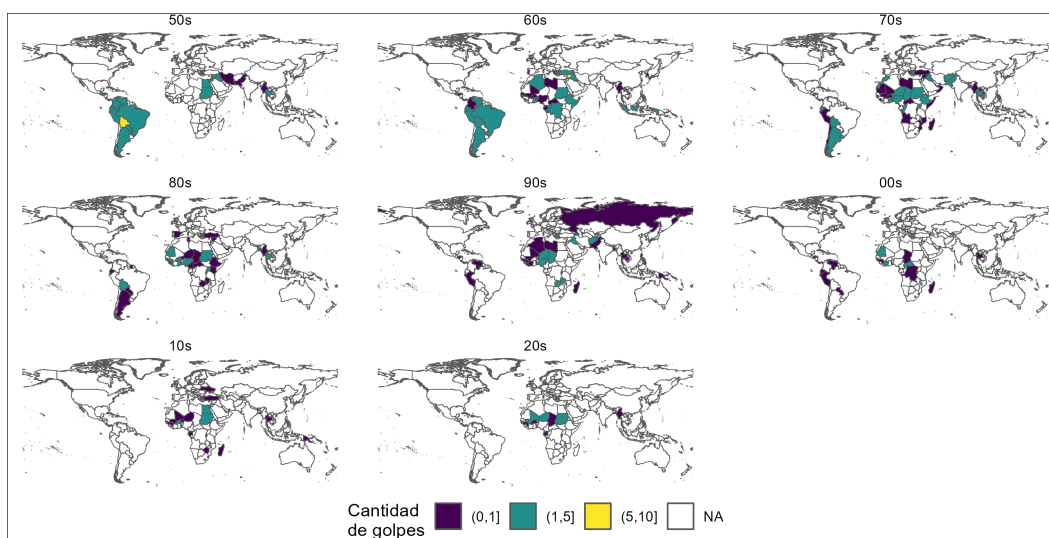


Figura 2: Conteo de golpes por década Fuente: Powell y Thyne, 2011

432 **4. Resultados y discusión**

433 **4.1. Performance de los modelos**

434 En primer lugar, se realizó la optimización bayesiana de ambos modelos según lo indicado en la
435 metodología. En el caso de XGBoost se pudo realizar las 100 iteraciones sin mayores inconvenientes,
436 tomando los valores óptimos de hiperparámetros para el entrenamiento final. Con respecto a Random
437 Forest, en cambio, se alcanzaron 44 iteraciones, debido a que cada iteración consumía una gran
438 cantidad de tiempo (en promedio una hora) y no se observaban mejoras significativas en el AUC.
439 De las iteraciones generadas, se tomaron los hiperparámetros que daban mejor resultado y tardaban
440 menos minutos en entrenar el modelo.

441 Una vez seleccionado los mejores hiperparámetros, se procede a entrenar los modelos en el conjunto
442 de entrenamiento final, el cual abarca los registros desde 1970 hasta 2019; así como a evaluar el
443 desempeño del mismo en los años 2020, 2021 y 2022 para emular el trabajo realizado por el FMI.

444 Es importante destacar que existen dos enfoques para evaluar el modelo en los años de testeo: por un
445 lado se pueden evaluar todos los años en su conjunto utilizando como datos de entrenamiento los
446 registros hasta el año anterior del primer año de validación. Una opción alternativa es ir entrenando el
447 modelo hasta el año anterior al de validación para cada año individualmente, de manera de poder
448 utilizar todos los años anteriores y no perder performance. Para este trabajo utilizamos el primer
449 enfoque, es decir que entrenamos los modelos hasta 2019 y los evaluamos en todos los años de
450 evaluación a la vez, de manera de aprehender de manera general la importancia de cada variable en la
451 predicción de la variable objetivo.

452 En el cuadro 1 observamos el desempeño de los modelos en los años de testeo. Por un lado figura
453 el AUC individual de cada año por separado y por el otro observamos el AUC acumulada es decir,
454 evaluando en ese año junto con los anteriores.

Año	XGBoost		Random Forest	
	AUC	AUC acumulada	AUC	AUC acumulada
2020	1.000	1.000	1.000	1.000
2021	0.750	0.786	0.747	0.784
2022	0.667	0.750	0.667	0.749

Cuadro 1: Área bajo la curva ROC por año puntual y acumulado (XGBoost y Random Forest)

455 Lo primero que podemos observar es que ambos modelos logran una performance perfecta para el
456 año 2020, lo cual resulta esperable ya que cuentan con información del año inmediatamente anterior
457 y a que sólo deben predecir un golpe de estado (en Mali). También esperable, la performance decae
458 en los años siguientes, lo cual impacta en el valor del AUC acumulada. Comparando ambos modelos,
459 observamos que XGBoost fue ligeramente superior en el año 2021 con respecto a Random Forest.
460 Como no son diferencias significativas, vamos a analizar ambos modelos para observar si utilizan
461 diferentes variables para sus predicciones.

462 Puesto que los casos positivos son apenas diez, aprovecharemos para visualizar la performance de los
463 modelos en cada uno de los casos. Adicionalmente, incorporamos el único falso positivo predicho
464 por Random Forest (Afganistan en el año 2021). La lista figura en la tabla (Cuadro 2).

país	año	¿Hubo golpe?	Random Forest	XGBoost
Malí	2020	Sí	Verdadero positivo	Verdadero positivo
Birmania/Myanmar	2021	Sí	Verdadero positivo	Verdadero positivo
Malí	2021	Sí	Falso negativo	Falso negativo
Sudán	2021	Sí	Verdadero positivo	Verdadero positivo
Afganistán	2021	No	Falso positivo	Verdadero negativo
Níger	2021	Sí	Falso negativo	Falso negativo
Guinea	2021	Sí	Verdadero positivo	Verdadero positivo
Chad	2021	Sí	Falso negativo	Falso negativo
Burkina Faso	2022	Sí	Verdadero positivo	Verdadero positivo
Guinea-Bisáu	2022	Sí	Falso negativo	Falso negativo
Santo Tomé y Príncipe	2022	Sí	Falso negativo	Falso negativo

Cuadro 2: Falsos negativos y verdaderos positivos (Random Forest)

465 Ambos modelos predicen correctamente el único golpe del año 2020 en Mali. A su vez, fallan en
466 predecir los golpes de Mali, Níger y Chad de 2021; así como fallan en predecir casi todos los golpes
467 del 2022, a excepción del de Burkina Faso. La única discrepancia entre los algoritmos surgió con
468 respecto a Afganistán en el año 2021: fue el único caso de falso positivo predicho por Random Forest,
469 mientras que XGBoost logró categorizarlo como negativo.

470 Profundizando en el análisis de la performance, vamos a observar la probabilidad asignada a cada
471 caso por el algoritmo. Mirando la probabilidad de Random Forest (figura 3), llama la atención como
472 el caso de Afganistán en 2021 figura tan arriba en la lista, superando incluso a cuatro casos positivos
473 ¿Por qué sucedió esto? Resulta útil remitirnos nuevamente a la definición de golpes de Estado de
474 Powell y Thyne: *intentos ilegales y manifiestos por parte de las fuerzas armadas u de otras elites*
475 *dentro del aparato estatal de derrocar al gobierno en ejercicio* (Powell y Thyne, 2011, p. 252). En el
476 caso de Afganistán, la toma del poder no fue perpetrada por sectores internos del Estado afgano, sino
477 por un ejército externos (en este caso, el movimiento Talibán).

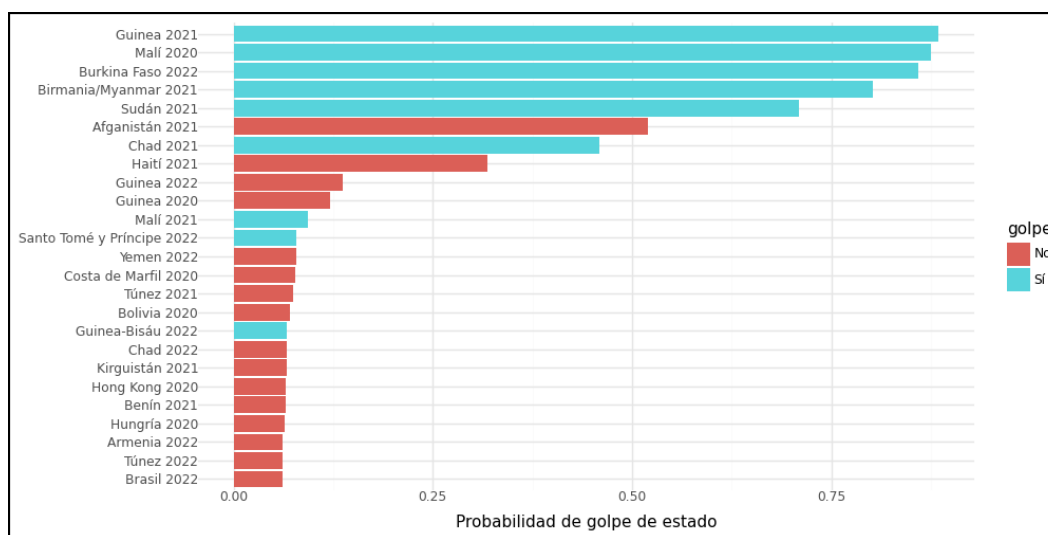


Figura 3: Predicción probabilística años 2020-2022 (Random Forest)

478 En cambio, Xgboost separa más consistentemente los casos positivos de los negativos en su asignación
479 de probabilidades (figura 4). Incluso posiciona en sexto lugar a Chad en 2021, si bien no lo predice
480 como positivo. De todas formas, en ambos casos Sao Tome y Príncipe en 2022 así como Guinea-
481 Bissau en 2022 figuran bastante abajo en el ranking lo cual hace esperable que fallen en lograr
482 predecir sus respectivos golpes de Estado. Por último, existe un único caso positivo que ni siquiera
483 alcanzó las primeras 25 posiciones: nos referimos a Níger en 2021. Random Forest le asignó una

484 probabilidad del 4,46 por ciento, posicionándolo en el puesto 36. Aún más interesante, XGBoost
 485 le asignó apenas una probabilidad de 0,033 por ciento (puesto 44). Profundizando en este caso,
 486 descubrimos que lo que sucedió tanto en Níger como en Guinea-Bisáu y en Santo Tomé y Príncipe
 487 fueron intentos fallidos de golpe de Estado. Esto nos puede dar una pista de qué tan efectivos son
 488 estos modelos para predecir golpes de Estado que no llegan a concretarse.

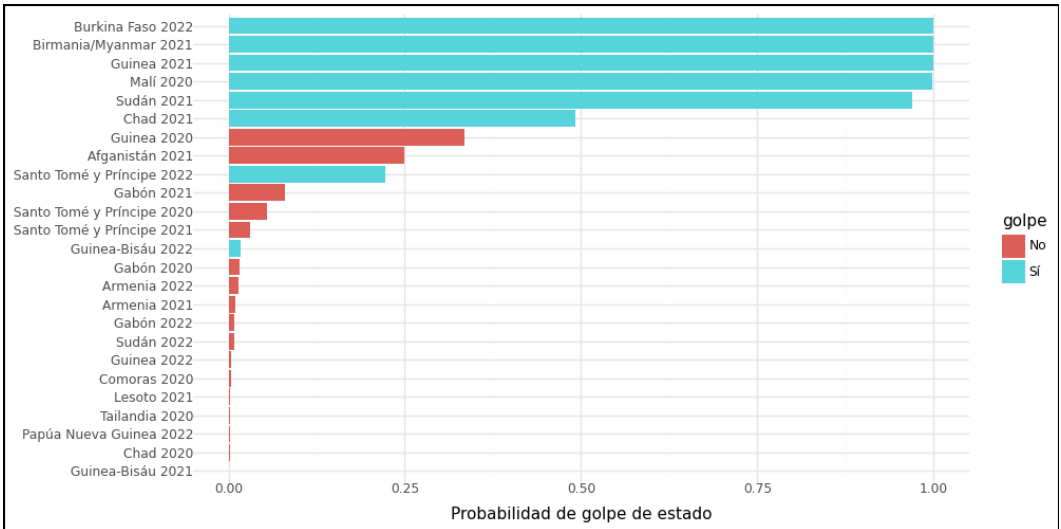


Figura 4: Predicción probabilística años 2020-2022 (XGBoost)

489 Con el objetivo de profundizar los motivos para predecir correcta o incorrectamente los casos
 490 expuestos, así como para enriquecer el análisis sobre los factores explicativos del fenómeno estudiado,
 491 realizaremos un análisis de las variables más relevantes para la predicción de golpes de acuerdo
 492 a los dos algoritmos. Utilizaremos la importancia de las variables, incluido dentro de los modelos
 493 entrenados, así como enfoques externos con los valores Shapley.

494 4.2. Análisis de variables

495 A continuación, pasaremos a evaluar la relevancia de las distintas variables del dataset para la
 496 predicción del modelo. De esa manera, podremos extraer elementos para determinar o reforzar los
 497 posibles causales de un golpe de estado en un territorio determinado. En primer lugar, utilizaremos la
 498 importancia de las variables según Random Forest, la cual se puede observar en la figura 5¹.

499 En el caso de Random Forest la importancia de las variables indica cuánto disminuye la impureza
 500 de los nodos (en este caso, medida con el índice de Gini, establecido por default en el modelo de
 501 Scikit-Learn) al utilizar una variable para dividir los datos de un árbol de decisión, de los cuales se
 502 obtiene un promedio para todo el ensamble de árboles.

503 Las barras indican el porcentaje de importancia de las 10 variables con mayor peso. En total, estas diez
 504 variables representan alrededor del 75 por ciento de la importancia. En general, todas las variables
 505 están relacionadas con la forma de gobierno, con la influencia de las fuerzas armadas en el mismo o
 506 con la misma variable objetivo en años anteriores. Entre el primer y el cuarto lugar figuran variables
 507 que reflejan muy evidentemente una relación con la presencia de golpes de estado, como tener la
 508 legislatura cerrada o abortada o que el ejecutivo no sea más electo.

¹Los nombres de las variables fueron traducidas y resumidas del libro de códigos de la base de datos para una vista amigable. Se puede verificar el nombre codificado y original de las variables en el cuadro 10

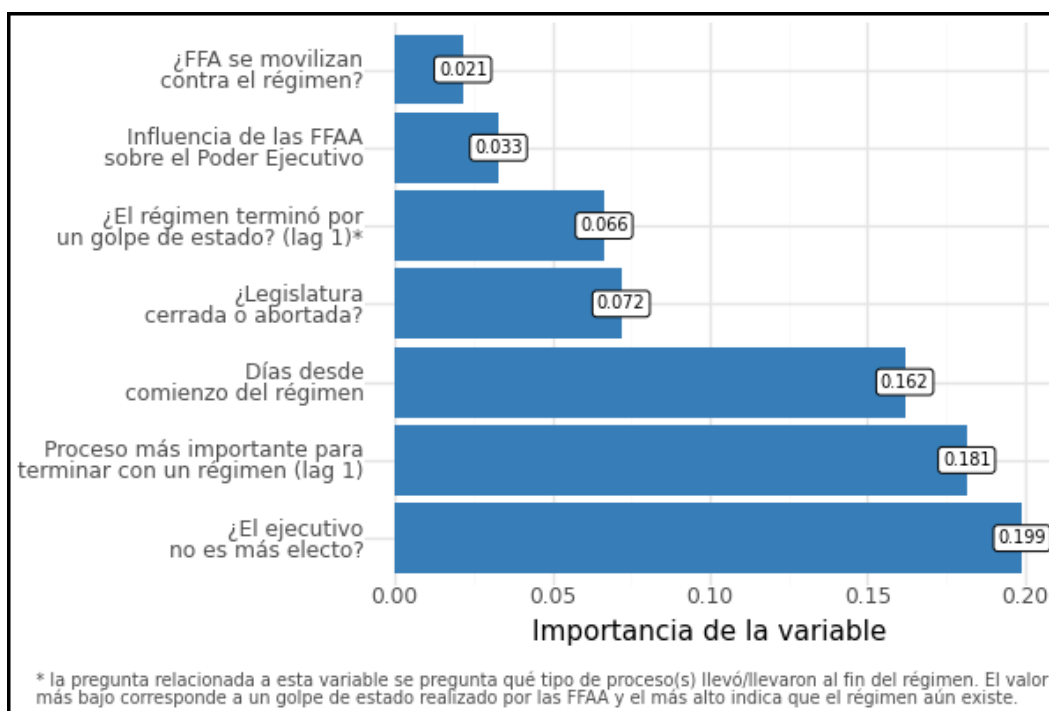


Figura 5: Importancia de las variables para predicción 2020-2022 (Random Forest)

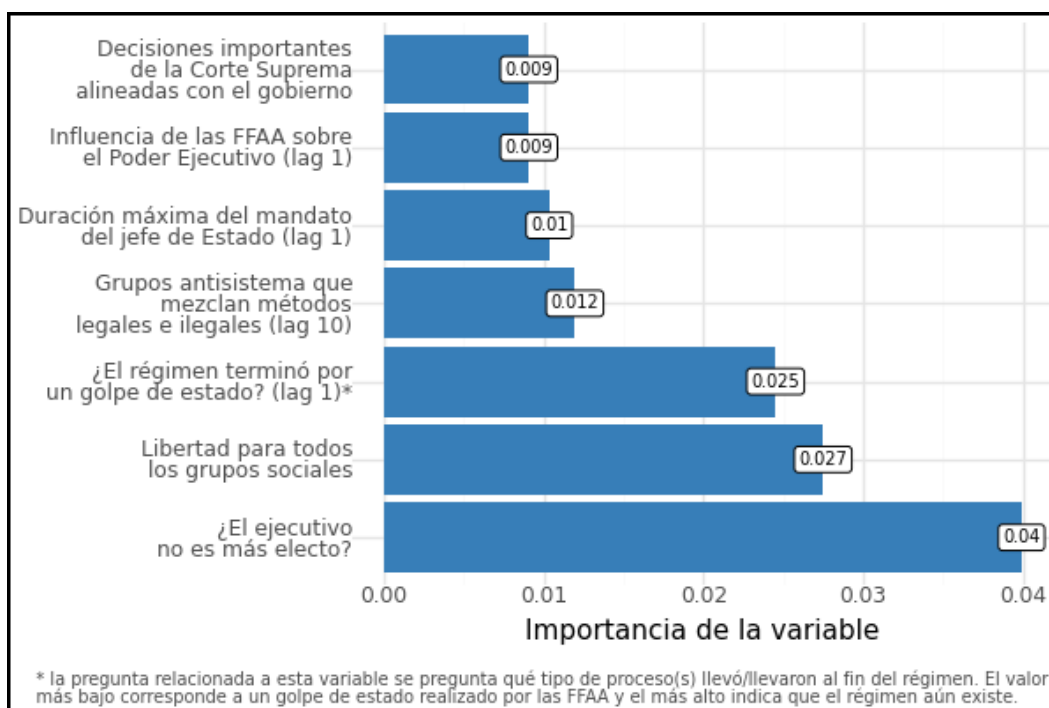


Figura 6: Importancia de las variables para predicción 2020-2022 (XGBoost)

509 Ahora observaremos una variable similar para XGBoost en la figura 6, si bien la importancia de
510 la variable para este algoritmo significa cuántas veces aparece cada variable en cualquier nodo de
511 decisión de los árboles del modelo, normalizado por la cantidad de veces que se realizaron divisiones
512 en todos los árboles. Si bien se encuentran similitudes con Random Forest, XGBoost destaca otras
513 variables que llaman la atención, como la libertad de grupos sociales, las decisiones de la Corte
514 Suprema que puedan ser adversas al gobierno, la duración del mandato del jefe de Estado y los
515 métodos de lucha de los grupos antisistema presentes en el país.

516 Una desventaja de esta métrica es que no permite evaluar el uso de estos atributos comparado con
517 el valor de los mismos. Es por eso que incorporamos un análisis utilizando valores Shapley Values.
518 En el eje Y figuran las primeras 11 variables con mayor valor de Shapley y en el eje X figura el
519 valor Shapley, visualizando la distribución de los casos en forma de violín y los outliers como puntos.
520 Finalmente, el color de los violines y de los puntos indica el valor de la variable en cuestión.

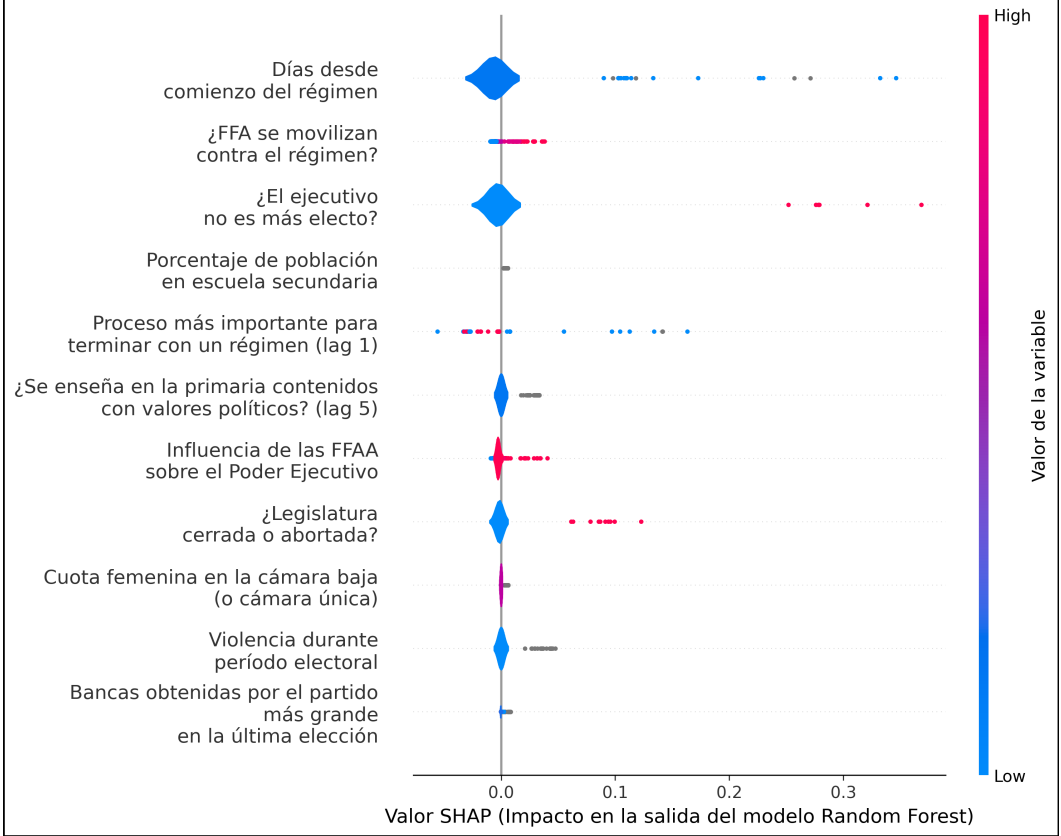


Figura 7: Shapley values para predicción 2020-2022 (Random Forest)

521 Si bien en la figura 7 algunas variables figuran también en el gráfico de la importancia de las variables,
522 podemos destacar algunas diferencias. Primero, figura la enseñanza de valores políticos en la escuela,
523 en cuyo valores nulos tienen alto valor Shapley. Lo mismo sucede con la violencia durante un período
524 electoral.

525 Adicionalmente, si bien la cantidad de días del régimen también figura en el gráfico de la importancias
526 de las variables, esta visualización permite observar también cómo influye el valor de la variable
527 en sus valores Shapley. Allí se observa que los casos con menor valor en la variable (es decir, con
528 menores días del régimen) tienen mayor valor Shapley e inclina la balanza hacia la predicción de un
529 caso positivo. Se puede inferir de esto último que un régimen joven es más inestable y, por lo tanto,
530 propensa a sufrir un nuevo cambio de régimen mediante un golpe.

531 Para profundizar el análisis vamos a observar los casos positivos y falsos negativos para saber qué
532 valores obtuvieron en las variables más destacadas de la figura 7.

país	año	v2regdur	shap	golpe	random forest	exitoso
Sudán	2021	0	0.346	Sí	Verdadero positivo	Sí
Birmania/Myanmar	2021	0	0.332	Sí	Verdadero positivo	Sí
Guinea	2021	0	0.230	Sí	Verdadero positivo	Sí
Burkina Faso	2022	0	0.227	Sí	Verdadero positivo	Sí
Malí	2020	0	0.226	Sí	Verdadero positivo	Sí
Afganistán	2021	0	0.173	No	Falso positivo	-
Chad	2021	0	0.090	Sí	Falso negativo	No
Santo Tomé y Príncipe	2022	11230	-0.006	Sí	Falso negativo	No
Níger	2021	3556	-0.007	Sí	Falso negativo	No
Guinea-Bisáu	2022	2749	-0.007	Sí	Falso negativo	No
Malí	2021	136	-0.008	Sí	Falso negativo	Sí

Cuadro 3: Días desde comienzo régimen vs valores SHAP (Random Forest)

533 En el cuadro 3 observamos el valor de los días desde comienzo del régimen (con el nombre original de
534 la base de datos para ahorrar espacio), el valor Shapley asociado de la misma variable, la predicción
535 que hizo Random Forest y, adicionalmente mediante una búsqueda en internet, si esos golpes de
536 Estado fueron exitosos o no². Además el cuadro está ordenada de manera descendente a partir de los
537 valores Shapley. Lo primero que podemos notar es que, a excepción de Chad en 2021, el hecho de
538 tener cero días de duración de un régimen influye en un valor Shapley muy alto, ubicándose entre los
539 más altos de toda la serie si se observa la escala de la figura 7. Eso parece influir en el hecho de que
540 Afganistán sea predicho falsamente como positivo; así como que la mayoría de los golpes de Estado
541 fallidos estén al final de la lista con una falsa predicción como negativo.

542 Ahora observaremos la variable que registra si el ejecutivo no es más electo:

país	año	v2x_hosinter	shap	golpe	random forest	exitoso
Chad	2021	1	0.321	Sí	Falso negativo	No
Guinea	2021	1	0.279	Sí	Verdadero positivo	Sí
Burkina Faso	2022	1	0.278	Sí	Verdadero positivo	Sí
Malí	2020	1	0.276	Sí	Verdadero positivo	Sí
Afganistán	2021	1	0.252	No	Falso positivo	-
Birmania/Myanmar	2021	0	-0.002	Sí	Verdadero positivo	Sí
Sudán	2021	0	-0.002	Sí	Verdadero positivo	Sí
Santo Tomé y Príncipe	2022	0	-0.003	Sí	Falso negativo	No
Malí	2021	0	-0.003	Sí	Falso negativo	Sí
Níger	2021	0	-0.003	Sí	Falso negativo	No
Guinea-Bisáu	2022	0	-0.003	Sí	Falso negativo	No

Cuadro 4: ¿El ejecutivo no es más electo? vs valores SHAP (Random Forest)

543 En este caso (tabla 4) la relación entre la predicción y el valor de la variable y el Shapley no es tan
544 directa, puesto que Chad en 2021 cuenta con el valor Shapley más alto de la lista pero es predicho
545 como negativo. Para el resto, se repite el patrón de predecir como negativo a valores bajos del valor
546 Shapley y como positivo a valores altos.

²Powell y Thyne (2011) cuentan con una base que distingue entre golpes exitosos y fallidos, pero llega hasta el año 2010.

país	año	v2regendtype_lag_1	shap	golpe	random forest	exitoso
Birmania/Myanmar	2021	0	0.163	Sí	Verdadero positivo	Sí
Sudán	2021	0	0.134	Sí	Verdadero positivo	Sí
Guinea	2021	0	0.113	Sí	Verdadero positivo	Sí
Malí	2020	0	0.104	Sí	Verdadero positivo	Sí
Burkina Faso	2022	0	0.097	Sí	Verdadero positivo	Sí
Afganistán	2021	1	0.055	No	Falso positivo	-
Santo Tomé y Príncipe	2022	13	-0.002	Sí	Falso negativo	No
Níger	2021	13	-0.002	Sí	Falso negativo	No
Malí	2021	13	-0.003	Sí	Falso negativo	Sí
Guinea-Bisáu	2022	13	-0.021	Sí	Falso negativo	No
Chad	2021	10	-0.056	Sí	Falso negativo	No

Cuadro 5: Proceso más importante para terminar con un régimen (lag 1) vs valores SHAP (Random Forest)

En el cuadro 5, en cambio, parece percibirse claramente la manera en la que el algoritmo separó casos positivos de negativos. Aquí, el caso discutido de Afganistán figura en la frontera entre los casos positivos y negativos. En cambio, todos los casos cuyo valor fue diez o más fueron asignados como negativos, incluso Chad en 2021 que figuraba como fuerte candidato a una predicción positiva en variables anteriores. Nuevamente, todos los casos fallidos de golpes fueron enviados al fondo de la lista. Retomando el caso de Afganistán, resulta relevante ahondar en qué releva esta variable y cómo se codifican sus variables. Esta variable pregunta por el proceso más importante que el experto considera que lleva al fin de un régimenCoppedge, Gerring, Knutsen, Lindberg et al., 2024. Las respuestas que figuran en el cuadro se codifican de la siguiente manera: 0=Un golpe de estado militar, 1=Un golpe de estado llevado a cabo por otros grupos diferentes a los militares, 13: El régimen aún existe. Afganistán figura con el valor 1, por lo que podemos inferir que los expertos de esta base de datos calificaron el proceso político en este país como un golpe de Estado, discrepando así con el artículo de Powell y Thyne, quienes no detectan un golpe de estado ni en 2021 ni en el año anterior, el cual es evaluado en esta variable (puesto que es lag 1).

país	año	v2xlg_leginter	shap	golpe	random forest	exitoso
Birmania/Myanmar	2021	1	0.123	Sí	Verdadero positivo	Sí
Guinea	2021	1	0.096	Sí	Verdadero positivo	Sí
Burkina Faso	2022	1	0.094	Sí	Verdadero positivo	Sí
Malí	2020	1	0.091	Sí	Verdadero positivo	Sí
Guinea-Bisáu	2022	1	0.086	Sí	Falso negativo	No
Chad	2021	1	0.063	Sí	Falso negativo	No
Afganistán	2021	1	0.062	No	Falso positivo	-
Santo Tomé y Príncipe	2022	0	-0.001	Sí	Falso negativo	No
Níger	2021	0	-0.001	Sí	Falso negativo	No
Malí	2021	0	-0.002	Sí	Falso negativo	Sí
Sudán	2021	0	-0.002	Sí	Verdadero positivo	Sí

Cuadro 6: ¿Legislatura cerrada o abortada? vs valores SHAP (Random Forest)

Por último, en el cuadro 6 analizamos la variable que evalúa si la legislatura fue cerrada o abortada en el país. Esta variable parece no haber tenido tanto peso como las anteriores para la predicción, puesto que figura el caso de Sudán 2021 que tuvo el valor Shapley más bajo de la tabla pero de todas formas fue predicho correctamente como positivo. De todas formas, un valor Shapley mayor a 0,90 parece garantizar que el caso positivo.

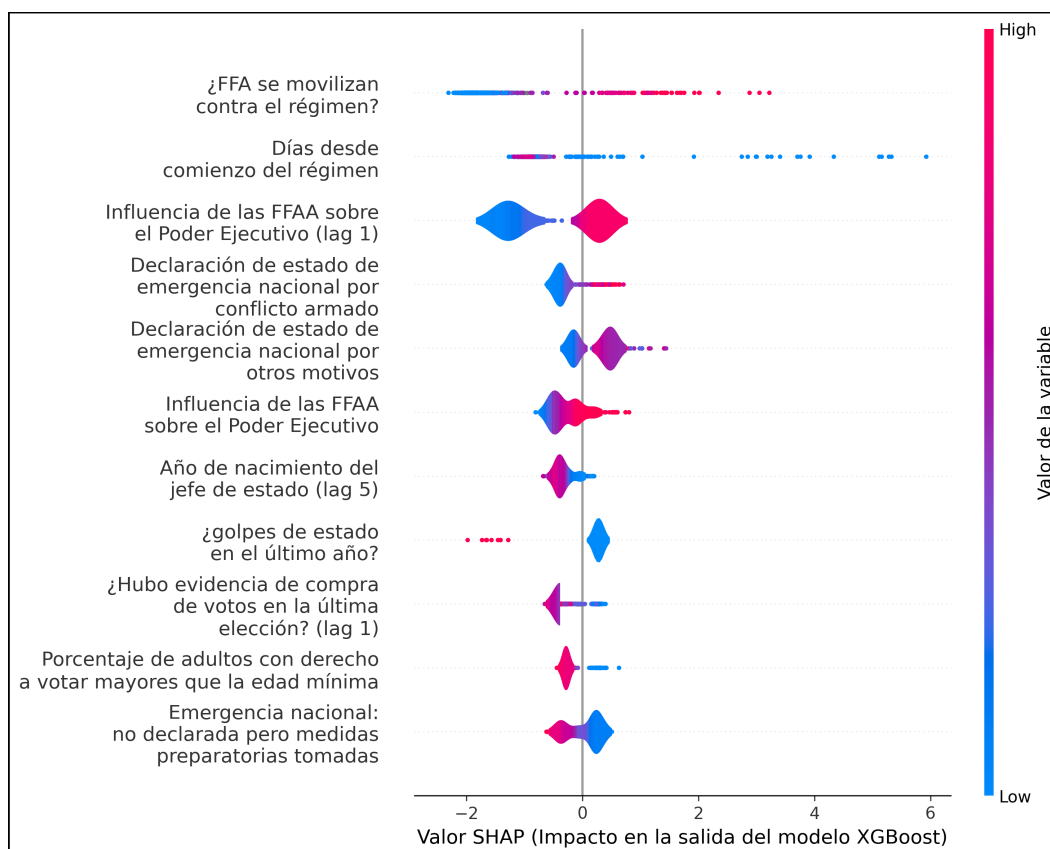


Figura 8: Shapley values para predicción 2020-2022 (XGBoost)

Los valores Shapley de XGBoost (figura 8) nos brinda nuevas perspectivas. Si bien comparte algunas variables con la visualización de Random Forest, figuran muchas más casos con valores Shapley negativos, los cuales influyen para que el caso sea predicho como negativo. Así sucede con la influencia de las FFAA sobre el Poder Ejecutivo en el lag 1 (es decir, el año anterior al del registro), el cual si tiene un valor alto influye para que sea catalogado como positivo, mientras que si cuenta con valores bajos influye para lo contrario.

Otros datos a destacar surgen de las variables que no figuran en Random Forest. En primer lugar, figuran dos variables que evalúan la declaración de un Estado de emergencia en el país: ya sea si es por un conflicto armado o por otros motivos, tener un valor alto en estas variables inclina la balanza hacia tener un golpe de Estado. También el hecho de haber tenido un golpe de estado en el año anterior parece influir para que no suceda otro en el siguiente. Lo mismo sucede con la evidencia de compra de votos en el año anterior (lag 1) y en el porcentaje de adultos con derecho a votar mayores que la edad mínima. Nuevamente, observaremos los valores Shapley de las variables más relevantes de la figura 8 para así observar con mayor detenimiento la manera en que el algoritmo tomó sus decisiones.

Lo primero que se puede apreciar del cuadro 7 es que asigna valores Shapley muy similares a los que asignó Random Forest en el cuadro 3. También envía casi todos los golpes fallidos al fondo de la lista, fallando en su correcta predicción. Sin embargo, XGBoost si fue exitoso en predecir que no sucedió un golpe de Estado en Afganistán, si bien parece ubicarse en una posición similar al del cuadro 3.

país	año	v2regdur	shap	golpe	xgboost	exitoso
Birmania/Myanmar	2021	0	5.927	Sí	Verdadero positivo	Sí
Sudán	2021	0	5.328	Sí	Verdadero positivo	Sí
Malí	2020	0	5.283	Sí	Verdadero positivo	Sí
Burkina Faso	2022	0	5.159	Sí	Verdadero positivo	Sí
Guinea	2021	0	5.115	Sí	Verdadero positivo	Sí
Afganistán	2021	0	4.333	No	Verdadero negativo	-
Chad	2021	0	2.747	Sí	Falso negativo	No
Malí	2021	136	1.034	Sí	Falso negativo	Sí
Santo Tomé y Príncipe	2022	11230	-0.671	Sí	Falso negativo	No
Níger	2021	3556	-0.845	Sí	Falso negativo	No
Guinea-Bisáu	2022	2749	-0.883	Sí	Falso negativo	No

Cuadro 7: Días desde comienzo régimen vs valores SHAP (XGBoost)

585 Como último análisis de los valores Shapley en XGBoost observaremos dos variables significativas
586 en las tablas 8 y 9. En estas tablas la separación entre casos positivos y negativos no es tan lineal
587 como en el caso anterior, sin embargo, se puede observar que los valores altos de los valores Shapley
588 en la movilización de las FFAA contra el poder ejecutivo influyen para que el caso sea predicho como
589 negativo. Interesantemente, aquí Afganistán figura como última en la lista, lo cual permite hipotetizar
590 que esta variable ayuda a excluir a este país del grupo de positivos.

país	año	v2regoppgroupsact_5	shap	golpe	xgboost	exitoso
Santo Tomé y Príncipe	2022	0.500	3.222	Sí	Falso negativo	No
Níger	2021	0.429	2.020	Sí	Falso negativo	No
Malí	2020	0.686	2.011	Sí	Verdadero positivo	Sí
Malí	2021	0.400	1.682	Sí	Falso negativo	Sí
Guinea-Bisáu	2022	0.250	1.191	Sí	Falso negativo	No
Sudán	2021	0.333	0.863	Sí	Verdadero positivo	Sí
Chad	2021	0.325	0.711	Sí	Falso negativo	No
Birmania/Myanmar	2021	0.271	0.494	Sí	Verdadero positivo	Sí
Burkina Faso	2022	0.250	0.468	Sí	Verdadero positivo	Sí
Guinea	2021	0.081	-1.072	Sí	Verdadero positivo	Sí
Afganistán	2021	0.000	-1.187	No	Verdadero negativo	-

Cuadro 8: ¿Las FFAA se movilizan contra el régimen? vs valores SHAP (XGBoost)

591 En el caso opuesto (tabla 9), un valor alto de los valores Shapley de la influencia de las FFAA sobre
592 el poder ejecutivo parecen influir para que el caso sea predicho como positivo.

país	año	v2x_ex_military	shap	golpe	xgboost	exitoso
Guinea	2021	1.000	0.804	Sí	Verdadero positivo	Sí
Burkina Faso	2022	0.938	0.744	Sí	Verdadero positivo	Sí
Malí	2020	0.917	0.580	Sí	Verdadero positivo	Sí
Malí	2021	1.000	0.554	Sí	Falso negativo	Sí
Birmania/Myanmar	2021	0.500	0.538	Sí	Verdadero positivo	Sí
Afganistán	2021	0.584	0.524	No	Verdadero negativo	-
Chad	2021	0.875	0.515	Sí	Falso negativo	No
Sudán	2021	0.700	0.397	Sí	Verdadero positivo	Sí
Níger	2021	0.286	-0.021	Sí	Falso negativo	No
Guinea-Bisáu	2022	0.125	-0.466	Sí	Falso negativo	No
Santo Tomé y Príncipe	2022	0.083	-0.525	Sí	Falso negativo	No

Cuadro 9: Influencia de las FFAA sobre el poder ejecutivo vs valores SHAP (XGBoost)

En definitiva, si bien las dos últimas variables expuestas parecen ayudar a la decisión final, el algoritmo le da más peso a los días desde el comienzo del régimen como Random Forest, si bien logra establecer un corte claro para excluir a Afganistán en 2021 del grupo de positivos. A continuación abordaremos el análisis de dos casos particulares para evaluar las dificultades que pueden haber tenido los modelos para predecirlos.

4.3. Análisis de casos particulares

En primer lugar, analizaremos el caso de Níger en 2021, en el cual ambos algoritmos fracasaron en predecir la presencia de un golpe de Estado. Comenzaremos observando las primeras cuatro variables con mayores valores Shapley en la predicción de un golpe de Estado en Níger entre 2020 y 2022.

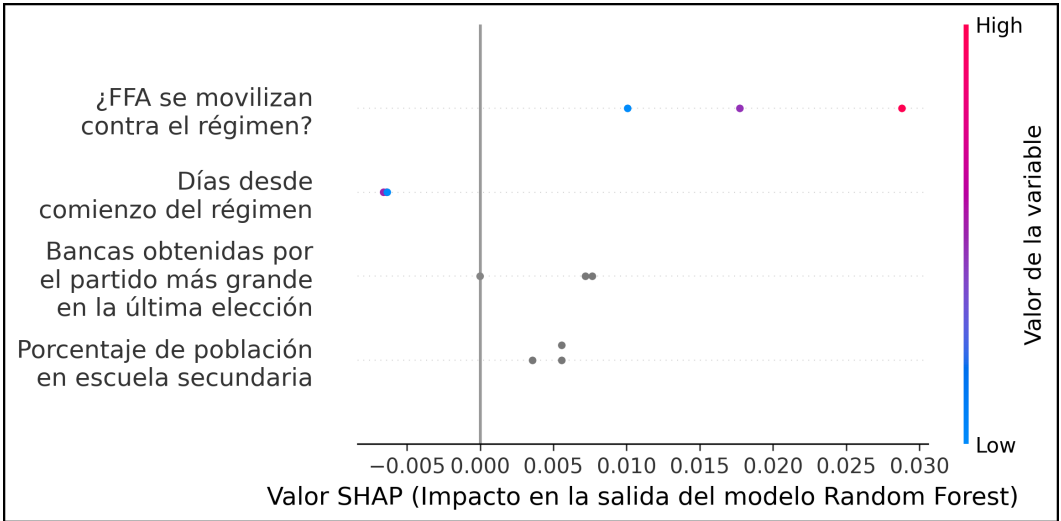


Figura 9: Shapley values para predicción Níger 2020-2022 (Random Forest)

Lo que se puede aprehender de la figura 9 es que en las cuatro variables, los tres casos (Níger en 2020, 2021 y 2023) tienen valores Shapley muy cercanos. Siguiendo esta lógica se vuelve difícil para el algoritmo predecir si un golpe de Estado sucederá o no.

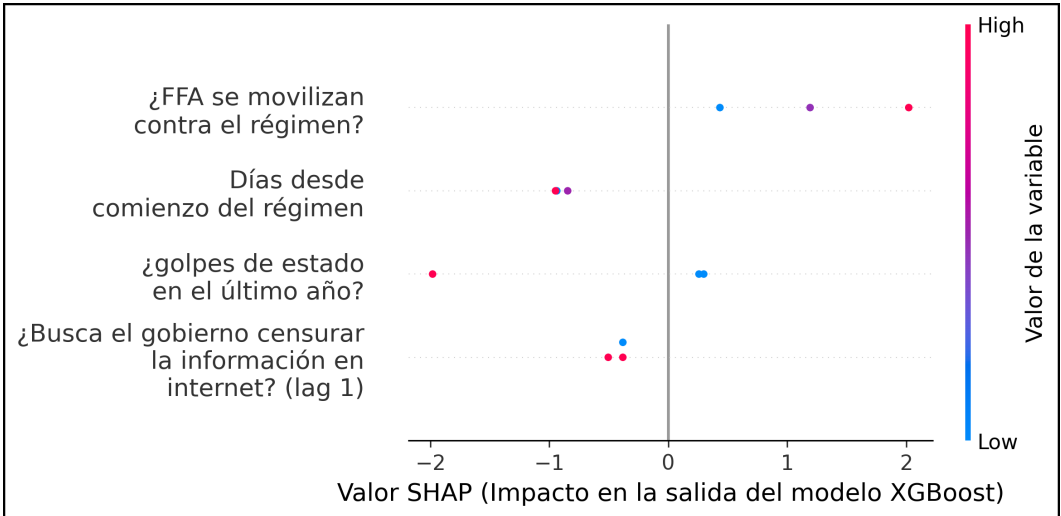


Figura 10: Shapley values para predicción Níger 2020-2022 (XGBoost)

En la figura 10 no parece haber diferencias significativas: en tres de las cuatro variables los tres casos figuran cercanos y con el mismo signo. Sí resulta interesante que el lag 1 de la variable objetivo (¿golpes de estado en el último año?) lógicamente separa al caso de Níger en 2022, pero no parece influir en el resultado final.

El segundo caso que evaluaremos es el de Afganistán en 2021, el cual generó discrepancias entre los datos y la variable objetivo, si bien uno de los modelos logró predecirlo bien.

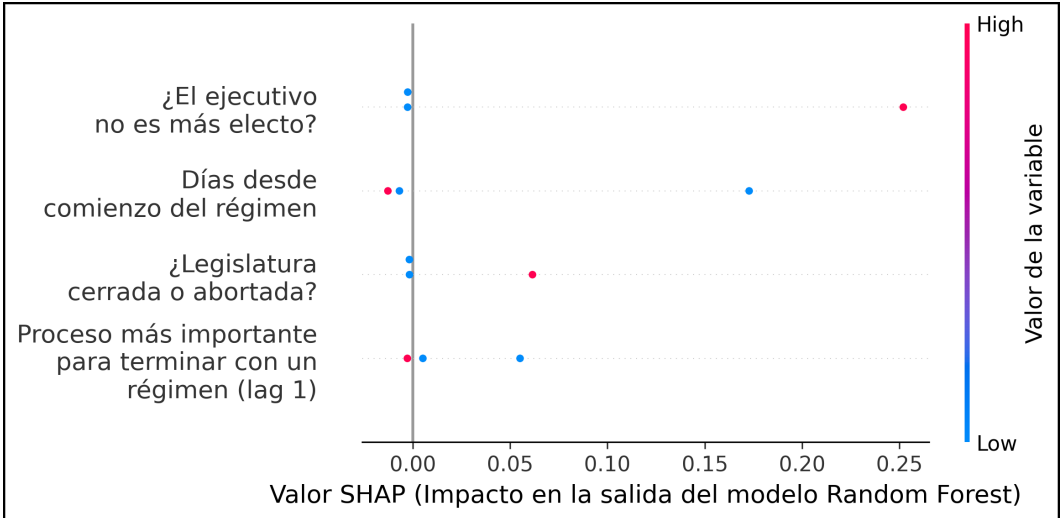


Figura 11: Shapley values para predicción Afganistán 2021 (Random Forest)

Lo que se observa con claridad en la figura 11 es que la cuatro variables separaron sensiblemente el caso de Afganistán en 2021 del resto de los años. Esto llevó a Random Forest a predecir incorrectamente esta año como uno en el que se generó un golpe según Powell y Thyne 2011.

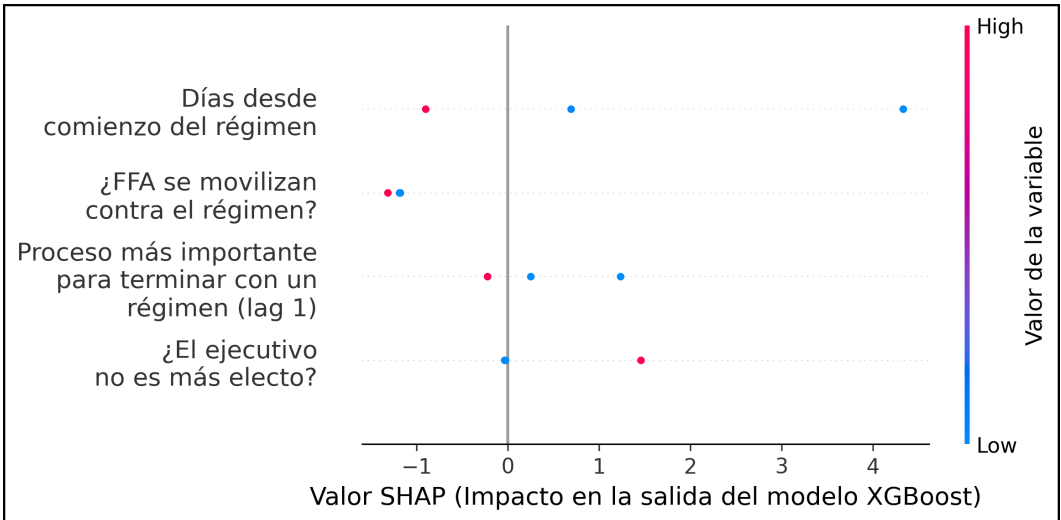


Figura 12: Shapley values para predicción Afganistán 2021 (XGBoost)

El caso de la figura 12 presenta mayores sutilezas. Si bien también los días desde el comienzo del régimen separa al año 2021 del resto. Las últimas dos variables presentan a los tres casos mucho más

616 cercanos. Además, la variable que evalúa si las FFAA se movilizan contra el régimen muestra tres
617 valores muy parecidos (lo cual lo hace incluso imperceptible para la visualización).

618 Como conclusión de esta sección, podemos aseverar que ambos algoritmos logran un piso de
619 performance apreciable para poder predecir la presencia de golpes de estado en un país y año en
620 particular. Sin embargo, XGBoost demostró una mayor capacidad para predecir casos complejos,
621 en los que (como se demostró observando la variable del cuadro 5) grupos de expertos parecen no
622 concordar con los que crearon la base de datos de golpes de Estado que utilizamos para construir
623 la variable objetivo. De todas formas, ambos algoritmos tienen grandes dificultades para predecir
624 golpes de estado que no fueron exitosos. Probablemente se deba a que le dan relevancia a atributos
625 que describen no sólo un golpe de estado sino que también un cambio de régimen. El atributo que
626 más sobresale es la cantidad de días desde que comenzó el régimen, el cual permite evidenciar la
627 idea anterior. Estos algoritmos también recurrieron a variables que describen la interacción entre
628 las fuerzas armadas y el poder ejecutivo, así como variables asociadas a la educación política en las
629 escuelas.

630 En la siguiente sección, compararemos las métricas de ambos modelos con los del artículo de
631 Cebotari et al 2024 para el FMI; así como las valores Shapley de las variables relevantes en ambos
632 casos. También expondremos las limitaciones de este trabajo, tanto en los datos trabajados como
633 en la metodología utilizada para replicar el trabajo del FMI. Por último, vincularemos las variables
634 destacadas por los algoritmos con los enfoques expuestos en el marco teórico.

635 **4.4. Discusiones**

- 636 - Vinculación de resultados con estado del arte y marco teórico
- 637 - Comparación de performance y de variables importantes con el artículo del FMI
- 638 - Limitaciones

639 **5. Conclusiones**

- 640 - Resumen de los hallazgos principales
- 641 - Conclusiones generales y su relación con los objetivos del trabajo
- 642 - Recomendaciones para futuros trabajos

643 **6. Anexo**

644 **6.1. Código**

645 La totalidad del código y entregas en latex y PDF se encuentran en un repositorio abierto de Github
646 de José Saint Germain (Acceso al repositorio). En el mismo se describe la secuencia de códigos a
647 correr para la obtención de datos, la ingeniería de atributos, la optimización bayesiana, la corrida
648 final, el análisis exploratorio de datos y el análisis de resultados de los algoritmos.

649 **6.2. Golpes**

- 650 - Breve descripción de los 10 golpes que se buscaron predecir. Su contexto histórico, político y social.

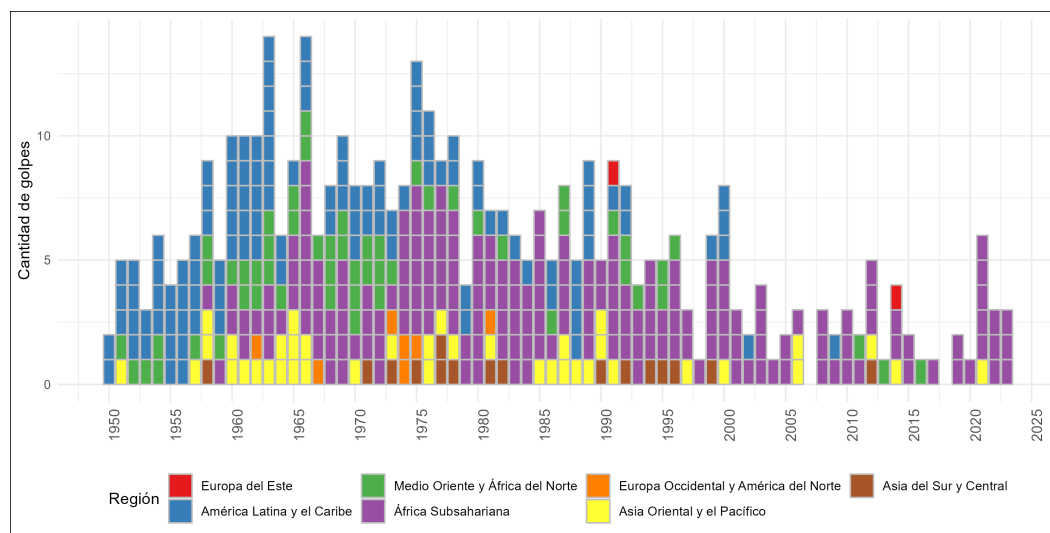


Figura 13: Conteo de golpes por año y región (Powell y Thyne, 2011)

Variable	Descripción
v2regdur	Días desde comienzo del régimen
v2regendtypems_0_lag_1	¿El régimen terminó por un golpe de estado? (lag 1)*
v2x_hosinter	¿El ejecutivo no es más electo?
v2xlg_leginter	¿Legislatura cerrada o abortada?
v2regendtype_lag_1	Proceso más importante para terminar con un régimen (lag 1)
v2x_ex_military	Influencia de las FFAA sobre el Poder Ejecutivo
v2regoppgroupsact_5	FFAA se movilizan contra el régimen?
coup_lag_1	¿golpes de estado en el último año?
v2x_ex_military_lag_1	Influencia de las FFAA sobre el Poder Ejecutivo (lag 1)
v2expathhs	¿Cómo llega el jefe de estado al gob?
v2regendtype_lag_5	¿Qué proceso fue el más importante en el fin del régimen? (lag 5)*
v2clpolcl_lag_1	Libertad para todos los grupos sociales
v2csanmvch_2_lag_10	Grupos antisistema que mezclan métodos legales e ilegales (lag 10)
v2juhcind	Decisiones importantes de la Corte Suprema alineadas con el gobierno
v2exfxtmhs_lag_1	Duración máxima del mandato del jefe de Estado (lag 1)
v2edpoledprim_lag_5	¿Se enseña en la primaria contenidos con valores políticos? (lag 5)
v2elpeace	Violencia durante período electoral
v2pesecsch	Porcentaje de población en escuela secundaria
v2lgqugent	Cuota femenina en la cámara baja (o cámara única)
v2ellostlg	Bancas obtenidas por el partido más grande en la última elección
v2casoe_1	Emergencia nacional: no declarada pero medidas preparatorias tomadas
v2casoe_4	Declaración de estado de emergencia nacional por conflicto armado
v2casoe_6	Declaración de estado de emergencia nacional por otros motivos
v2exagehos_lag_5	Año de nacimiento del jefe de estado (lag 5)
v2elvotbuy_lag_1	¿Hubo evidencia de compra de votos en la última elección? (lag 1)
v2asuffrage	Porcentaje de adultos con derecho a votar mayores que la edad mínima

Cuadro 10: Nombre original de variables y su descripción

Referencias

- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American political science review*, 53(1), 69-105.
- Huntington, S. P. (1968). *Political order in changing societies*. Yale university press.
- O'donnell, G. (1972). Modernización y autoritarismo.
- Cardoso, F. H., & Faletto, E. (1979). *Dependencia y desarrollo en América Latina*. Siglo XXI México.
- Burman, P., Chow, E., & Nolan, D. (1994). A Cross-Validatory Method for Dependent Data. *Biometrika*, 81(2), 351-358. Consultado el 1 de mayo de 2024, desde <http://www.jstor.org/stable/2336965>
- Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press.
- Racine, J. (2000). Consistent cross-validatory model-selection for dependent data: hv-block cross-validation. *Journal of Econometrics*, 99(1), 39-61. [https://doi.org/https://doi.org/10.1016/S0304-4076\(00\)00030-0](https://doi.org/https://doi.org/10.1016/S0304-4076(00)00030-0)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. <https://doi.org/http://doi.org/10.1023/A:1010933404324>
- Strumbelj, E., & Kononenko, I. (2010). An Efficient Explanation of Individual Classifications using Game Theory. *The Journal of Machine Learning Research*, 11, 1-18.
- Powell, J. M., & Thyne, C. L. (2011). Global Instances of Coups from 1950 to 2010: A New Dataset. *Journal of Peace Research*, 48(2), 249-259.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd international conference on knowledge discovery and data mining*, 785-794. <https://doi.org/https://doi.org/10.48550/arXiv.1603.02754>
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan & R. Garnett, Eds.). 30. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- Cebotari, A., Chueca-Montuenga, E., Diallo, Y., Ma, Y., Turk, R., Xin, W., & Zavarce, H. (2024). *Political Fragility: Coups d'État and Their Drivers*. IMF Working Paper 24/34. <https://doi.org/https://doi.org/10.23696/mcwt-fr58>
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Marquardt, K. L., Medzihorsky, J., Pemstein, D., Fox, L., Gastaldi, L., Pernes, J., Rydén, O., von Römer, J., Tzelgov, E., Wang, Y.-t., & Wilson, S. (2024). "V-Dem Methodology v14" *Varieties of Democracy (V-Dem) Project* (Report). <https://v-dem.net/data/reference-documents/>
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Angiolillo, F., Bernhard, M., Borella, C., Cornell, A., Fish, S. M., Fox, L., Gastaldi, L., Gjerløw, H., Glynn, A., God, A. G., Grahn, S., Hicken, A., Kinzelbach, K., ... Ziblatt, D. (2024). *V-Dem Dataset v14* Varieties of Democracy (V-Dem) Project* (Report). <https://doi.org/https://doi.org/10.23696/mcwt-fr58>