
Entrega 3: Introducción y Resultados

José Saint Germain
joesg998@gmail.com

Índice

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

27 1. Introducción

ser más
específico

28 El objetivo de este trabajo es lograr entrenar un modelo de aprendizaje automático que logre predecir
29 de manera aceptable la presencia de golpes de estado durante los años 2020 a 2022 en todos los
30 países del mundo a partir de la utilización de la base de datos provista por la fundación Varieties
31 of Democracy (V-Dem) (Coppedge, Gerring, Knutsen, Lindberg et al., 2024), así como tener una
32 noción acabada de las variables más importantes que los algoritmos utilizan para la predicción de la
33 variable objetivo.

34 1.1. Motivación

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

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

50 1.2. Estructura del documento

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

52 2. Marco Teórico

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

59 2.1. La Teoría de la Modernización

60 - Explicación sobre la teoría de la modernización y el trabajo de Lipset

61 - Trabajo alternativo de Barrington Moore, Jr.

62 2.2. Teoría de la Dependencia y el Subdesarrollo

63 - Teorías de la dependencia y el subdesarrollo: Cardoso y Faletto

64 2.3. Enfoques políticos y sociales

65 - Análisis institucionales: Huntington, Linz y Stepan

66 - Análisis sociales: elites y movimientos sociales (O'donnell y Schmitter)

67 2.4. Datos Empíricos y Nuevas Perspectivas

68 - Enfques empíricos y cuantitativos: Powell y Thyne; Przeworski.

- 69 - Machine learning en la ciencia política. Usos recientes
- 70 - Síntesis y cierre del marco teórico / estado del arte

71 3. Metodología

**replicar es llegar a los mismos resultados con el mismo
procedimiento y mismos datos
reproducir es usar los mismos métodos con otros datos**

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

75 3.1. Algoritmos de predicción

76 Los algoritmos que se utilizarán serán Random Forest (Breiman, 2001) y XGBoost (Chen y Guestrin,
77 2016). Ambos algoritmos son modelos de ensamble basados en múltiples árboles de decisión. Un
78 árbol de decisión individual es un modelo predictivo que divide los datos en subconjuntos cada vez
79 más pequeños basándose en una serie de decisiones binarias sobre las características de los datos. En
80 cada nodo del árbol, se selecciona una característica y un umbral para dividir los datos en dos grupos:
81 aquellos que cumplen la condición y aquellos que no. Este proceso se repite de manera recursiva hasta
82 que se alcanza una condición de parada, ya sea un mínimo de muestras en un nodo o una profundidad
83 máxima del árbol.

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

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

98 3.2. Métrica de evaluación

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

107 3.3. Optimización de hiperparámetros

108 Con respecto al ajuste de hiperparámetros se utilizará la optimización bayesiana. La misma consistirá
109 en 100 iteraciones en donde se buscará el valor óptimo de los siguientes hiperparámetros:

- 110 ■ Random Forest: profundidad máxima de los árboles (max_depth) y la submuestra del ratio
111 de columnas a considerar cuando se construye cada árbol (max_features).
- 112 ■ XGBoost: la tasa de aprendizaje (learning_rate) y el término de regularización L2 en los
113 pesos (reg_lambda).

114 Adicionalmente el parámetro que establece la cantidad de árboles creados (n_estimators) quedará
115 fijado en 1000.

116 3.4. Block-time-series cross-validation 117 definir

117 Para evitar el **data leakage**, en cada iteración de la optimización bayesiana se utilizará la validación
118 cruzada. Sin embargo, como se trabajará con una base de datos de panel, conviene utilizar una versión
119 adaptada: el método *block- time-series cross-validation*, basado en Burman et al., 1994 y Racine,
120 2000. El método aplicado en este caso consiste en generar 5 pares de entrenamiento y validación:
121 1970 - 2009, 2010 - 2011; 1970 - 2011, 2012 - 2013; 1970 - 2013, 2014 - 2015; 1970 - 2015, 2016 -
122 2017; 1970 - 2017, 2018- 2019. Por lo tanto, cada set de entrenamiento consiste en observaciones
123 desde 1970 hasta un año de corte (2009, 2011, 2013, 2015, 2017) y el set de validación contempla los
124 dos años siguientes del mismo. Una vez realizada la optimización bayesiana, se toman los valores de
125 hiperparámetros que lograron maximizar el AUC y se entrena el modelo con el set de entrenamiento
126 para intentar predecir los golpes de estado entre 2020 y 2022.

127 3.5. Valores Shapley

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

137 3.6. Ingeniería de atributos

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

150 3.7. Análisis Exploratorio de Datos

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

- 159 ■ Estimador del modelo (Variable sin sufijo): es la medida recomendada para su análisis.
160 Corresponde a obtener la mediana del valor de la variable entre los expertos, reescalado a
161 valores entre -5 a 5.
- 162 ■ Medidas de incertidumbre (*_codelow y *_codehigh): corresponden a un desvío estándar
163 por encima y por debajo del estimador del modelo. Usadas conjuntamente, construyen un
164 intervalo de confianza del 95 %.
- 165 ■ Escala original (*_osp): mediana de la variable, pero sin reescalar. Esta versión también
166 cuenta con sus medidas de incertidumbre correspondientes.

Media entre las estimaciones de todos los expertos?

- 167 ■ **Media simple** (_mean): mediana de la variable, pero sin reescalar.
- 168 ■ Desvío estándar (_sd): desvío estándar de la variable.
- 169 ■ Media simple (_mean): media de la variable.
- 170 ■ Cantidades de expertos (_nr): cantidad de expertos que respondieron por país, año y variable.

se contruyen 4607 variables por país?

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

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

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

- 181 ■ ca: Espacio cívico y académico
- 182 ■ cl: Libertad civil
- 183 ■ cs: Sociedad civil
- 184 ■ dd: Democracia directa
- 185 ■ de: Demografía
- 186 ■ dl: Deliberación
- 187 ■ el: Elecciones
- 188 ■ ex: Ejecutivo
- 189 ■ exl: Legitimación
- 190 ■ ju: Poder judicial
- 191 ■ leg: Legislatura
- 192 ■ lg: Legislatura
- 193 ■ me: Medios de comunicación
- 194 ■ pe: Igualdad política
- 195 ■ ps: Partidos políticos
- 196 ■ sv: Soberanía
- 197 ■ st: Estado
- 198 ■ x: Índice (calculado a partir de variables que también se incluyen en la base de datos)
- 199 ■ zz: Cuestionario posterior a la encuesta
- 200 ■ ws: Encuesta de sociedad digital

201 A la base original obtenida desde la librería de V-Dem, se le realizaron los siguientes filtros: en primer
202 lugar, se removieron todas las variables que no sean las principales, es decir, que no cuenten con
203 sufijo. De esa manera, se busca reducir el tamaño de la base y así poder agregar nuevas columnas
204 mediante ingeniería de atributos. En segundo lugar, se filtraron los años superiores a 1950, para
205 adecuarnos al periodo utilizado en el artículo del FMI. De esa manera, la base filtrada cuenta con
206 12208 filas y 1460 columnas. Por último, se remueven todas las variables de fuentes externas (cuyo
207 agrupador comienza con 'e'), las variables pertenecientes a la base histórica (agrupador 'hist') y las
208 de la encuesta de sistema de partidos políticos; en parte debido a que provienen de fuentes ajenas
209 a V-Dem que pueden comprometer la completitud futura de los datos y en parte porque algunas de
210 estas variables cuentan con alta tasa de nulos.

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

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

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

Haciendo foco en la **variable** objetivo, es importante aclarar que en este trabajo no estamos contando
la cantidad precisa de golpes de estado sucedidos en un período de tiempo, sino que simplemente
relevamos si al menos un golpe de estado sucedió en un país y año determinado. Por lo tanto, si
un país sufrió más de un golpe de estado en un año, el mismo será contabilizado una sola vez.
Adicionalmente, en este trabajo también se consideran los golpes de estado que no fueron exitosos,
es decir, que no lograron derrocar al gobierno en cuestión. De allí se desprende que países como
Argentina, que en total ha tenido seis golpes de estado exitosos, figure con el doble de golpes en la
figura 1.

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

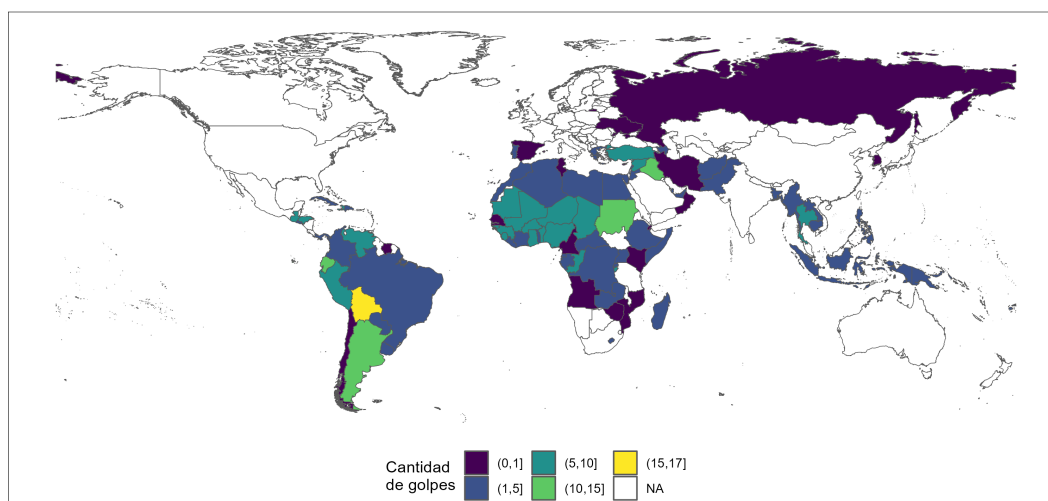


Figura 1: Golpes de estado período (1950-2023) Fuente: Powell y Thyne, 2011

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

247 Desagregando por década se observan algunos cambios, así como la persistencia en algunas regiones.
 248 La región del Sahel y varias naciones circundantes fueron persistentemente afectadas por golpes
 249 de estado desde los años 60. En América del Sur, en cambio, la presencia casi total de situaciones
 250 golpistas en la región se fue acotando a partir de los años 80 hasta finalmente desaparecer en el siglo
 251 xxi. Para observar con más detalle y discriminado por años y países se puede ver la figura 5.

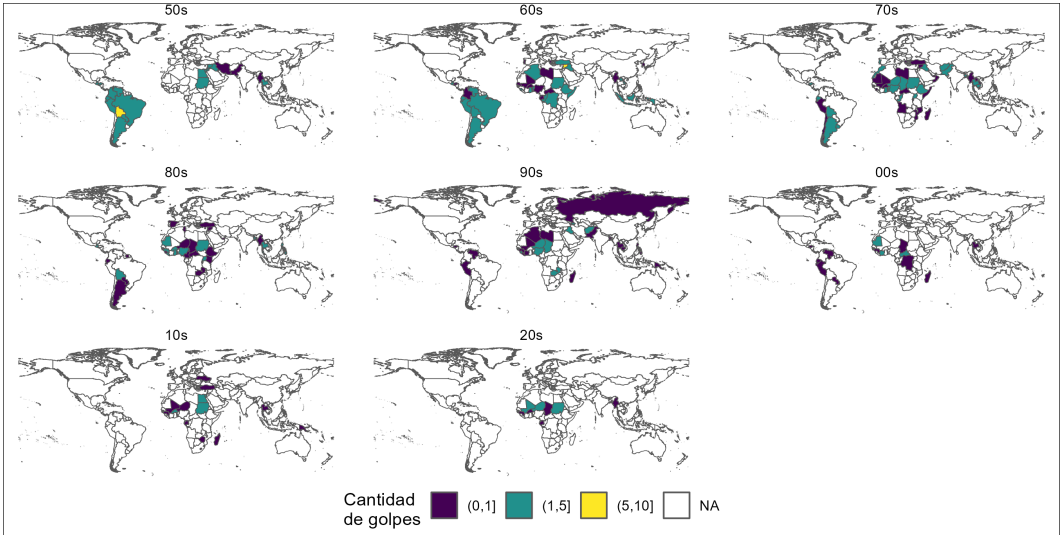


Figura 2: Conteo de golpes por década Fuente: Powell y Thyne, 2011

252 4. Resultados y discusión

253 4.1. Performance de los modelos

254 En primer lugar, se **realizó** la optimización bayesiana de ambos modelos según lo indicado en la
 255 metodología. En el caso de XGBoost se pudo realizar las 100 iteraciones sin mayores inconvenientes,
 256 tomando los valores óptimos de hiperparámetros para el entrenamiento final. Con respecto a Random
 257 Forest, en cambio, se alcanzaron 53 iteraciones, debido a que cada iteración consumía una gran
 258 cantidad de tiempo (en promedio una hora) y no se observaban mejoras significativas en el AUC.
 259 De las iteraciones generadas, se tomó los hiperparámetros del segundo mejor AUC, puesto que la
 260 diferencia con el ganador en el score era insignificante, pero el tiempo de cómputo era menos de la
 261 mitad.

262 Una vez seleccionado los mejores hiperparámetros, se procede a entrenar los modelos en el conjunto
 263 de entrenamiento final, el cual abarca los registros desde 1950 hasta 2019; así como a evaluar el
 264 desempeño del mismo en los años 2020, 2021 y 2022 para emular el trabajo realizado por el FMI.

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

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

Año	XGBoost		Random Forest	
	AUC	AUC acumulada	AUC	AUC acumulada
2020	1.000000	1.000000	1.000000	1.000000
2021	0.750000	0.785714	0.830443	0.855718
2022	0.666667	0.750000	0.666667	0.799051

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

Lo primero que podemos observar es que ambos modelos logran una performance perfecta para el año 2020, lo cual resulta esperable ya que cuentan con información del año inmediatamente anterior. También esperable, la performance decae en los años siguientes, lo cual impacta en el valor del AUC acumulada. Lo más destacable es que Random Forest logra una mejor performance que XGBoost en el resto de años, alcanzando un AUC de casi 0.8 y 0.75, respectivamente. Con esta información, se tomó la decisión de continuar el análisis de resultados con Random Forest.

Focalizando solamente en Random Forest, y puesto que los casos negativos (527 en los tres años de evaluación) fueron predichos de manera perfecta, podemos aprovechar para visualizar los casos positivos que son apenas diez casos, tantos los verdaderos positivos como los falsos negativos (Cuadro 2). En primer lugar, podemos notar que la predicción perfecta en el año 2020 se debe a que el modelo logró predecir correctamente el único golpe de ese año en Malí, en la región del Sahel. Después, en el año 2021 esta performance disminuye, al no lograr predecir los golpes de nuevo en Mali y en Niger, aunque si predice golpes en Sudán, Guinea y Chad. A simple vista, no parece haber datos geográficos o históricos que permitan establecer por qué logra predecir algunos golpes y en otros no, en especial porque son países relativamente similares, de pocos años de independencia y dentro de la misma región. Adicionalmente, en este año también logró predecir el golpe de estado en Burma/Myanmar, una nación ubicada en una región alejada de África, en el sudeste asiático.

Año	País	¿Hubo golpe?	Predicción	Resultado
2020	Mali	Sí	Sí	Verdadero positivo
2021	Burma/Myanmar	Sí	Sí	Verdadero positivo
2021	Sudan	Sí	Sí	Verdadero positivo
2021	Guinea	Sí	Sí	Verdadero positivo
2021	Chad	Sí	Sí	Verdadero positivo
2022	Burkina Faso	Sí	Sí	Verdadero positivo
2021	Mali	Sí	No	Falso negativo
2021	Niger	Sí	No	Falso negativo
2022	Guinea-Bissau	Sí	No	Falso negativo
2022	Sao Tome and Principe	Sí	No	Falso negativo

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

Finalmente, el año 2022 expone la peor performance del modelo: si bien logra predecir un golpe de estado en Burkina Faso, falla al predecir golpes en Guinea-Bissau y Sao Tome y Principe, todos países en la misma región del continente africano. De manera general, podemos asociar esta baja en la performance a que el modelo deja de contar con información del año inmediatamente anterior al del conjunto de evaluación.

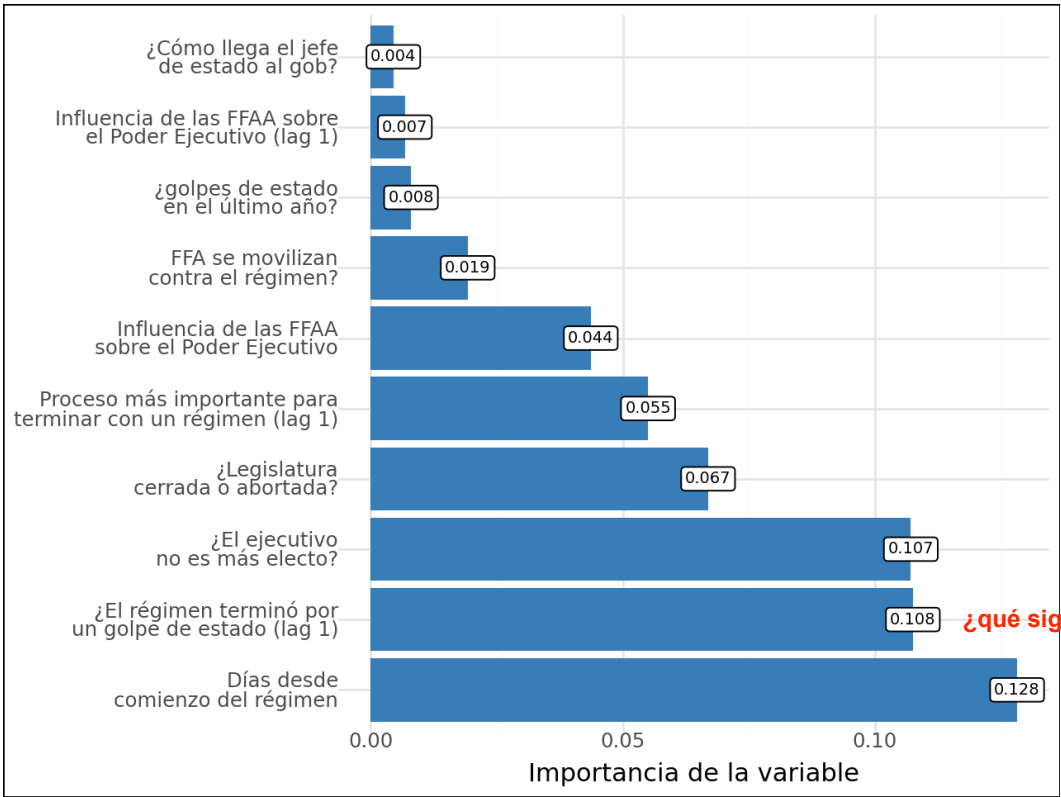
4.2. Análisis de variables

A continuación, pasaremos a evaluar la relevancia de las distintas variables del dataset para la predicción del modelo. De esa manera, podremos extraer elementos para determinar o reforzar los

301 posibles causales de un golpe de estado en un territorio determinado. En primer lugar, utilizaremos
302 la importancia de las variables según Random Forest, la cual se puede observar en la figura 3 (los
303 nombres de las variables fueron traducidas y resumidas del libro de códigos de la base de datos para
304 una vista amigable. Se puede verificar el nombre codificado y original de las variables en la tabla 3).

305 Las barras indican el porcentaje de importancia de las 10 variables con mayor peso. En total, estas
306 diez variables representan alrededor del 50 % de la importancia. En general, todas las variables están
307 relacionadas con la forma de gobierno, con la influencia de las fuerzas armadas en el mismo o con
308 la misma variable objetivo en años anteriores. Entre el segundo y el cuarto lugar figuran variables
309 que reflejan muy evidentemente una relación con la presencia de golpes de estado, como tener la
310 legislatura cerrada o abortada o que el ejecutivo no sea más electo. El dato más interesante a destacar
311 es que la variable con mayor importancia es la cantidad de días desde que comenzó el régimen. Se
312 puede inferir de esto último que **un régimen joven es más inestable** y, por lo tanto, propensa a sufrir
313 un nuevo cambio de régimen mediante un golpe.

dónde se ve el signo?



¿qué significa esta variable?

Figura 3: Importancia de las variables para predicción 2020-2022 (Random Forest)

314 Ahora incorporaremos los Shapley Values para identificar variables importantes a la hora de predecir
315 la presencia de golpes de Estado, como se expone en la figura 4. En el eje Y figuran las primeras 11
316 variables con mayor valor de Shapley y en el eje X figura el valor Shapley, visualizando la distribución
317 de los casos en forma de violín y los outliers como puntos. Finalmente, el color de los violines y de
318 los puntos indica el valor de la variable en cuestión.

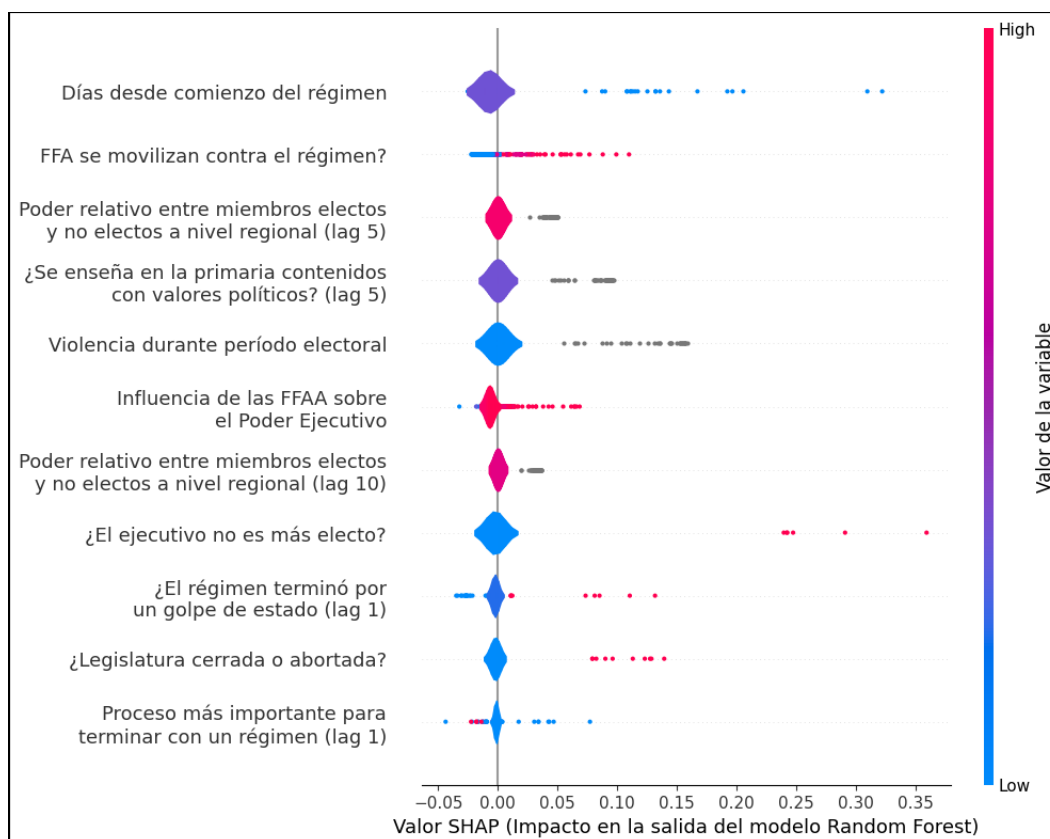


Figura 4: Shapley values para predicción 2020-2022 (Random Forest)

Si bien en este gráfico algunas variables figuran también en el gráfico de la importancia de las variables, podemos destacar algunas diferencias. Primero, figura la enseñanza de valores políticos en la escuela, en cuyo valores nulos tienen alto valor Shapley. También destacan los datos nulos en el poder relativo entre miembros electos y no electos a nivel regional, en su misma versión hace 10 años (lag 10) y en la presencia de violencia durante el período electoral.

Para comprender qué significan estos datos nulos, es de utilidad recurrir al libro de códigos de la base de datos. Por ejemplo, un valor faltante en el poder relativo entre oficiales electos y no electos significa que todos o casi todos de los funcionarios electos son subordinados de algún otro poder que no surgió de las urnas (a nivel regional).

Otros temas a tratar en los resultados en la próxima entrega: - Valores Shapley en cada año individual y/o en cada país - Análisis histórico de las variables destacadas

4.3. Discusiones

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

5. Conclusiones

- Resumen de los hallazgos principales
- Conclusiones generales y su relación con los objetivos del trabajo

337 - Recomendaciones para futuros trabajos

338 **6. Anexo**

339 **6.1. Código**

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

344 **6.2. Golpes**

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

346 **6.3. Gráficos y tablas adicionales**

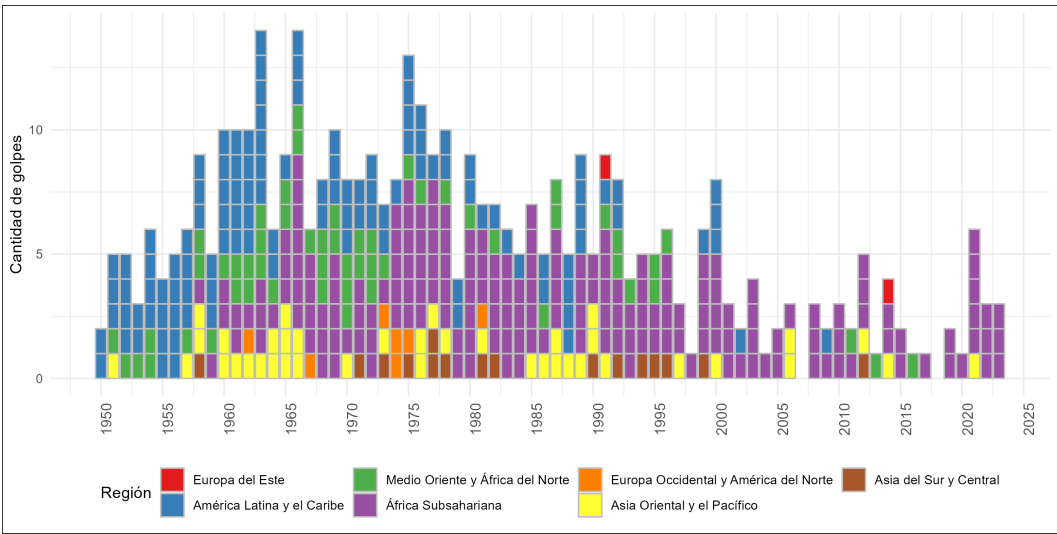


Figura 5: Conteo de golpes por año y región (Powell y Thyne, 2011)

Variable	Descripción
coup_lag_1	¿golpes de estado en el último año?
v2expathhs	¿Cómo llega el jefe de estado al gob?
v2regdur	Días desde comienzo del régimen
v2regoppgroupsact_5	FFA se movilizan contra el régimen?
v2elrgpwr_lag_5	Poder relativo entre miembros electos y no electos a nivel regional (lag 5)
v2edpoledprim_lag_5	¿Se enseña en la primaria contenidos con valores políticos? (lag 5)
v2elpeace	Violencia durante período electoral
v2x_ex_military	Influencia de las FFAA sobre el Poder Ejecutivo
v2x_ex_military_lag_1	Influencia de las FFAA sobre el Poder Ejecutivo (lag 1)
v2elrgpwr_lag_10	Poder relativo entre miembros electos y no electos a nivel regional (lag 10)
v2x_hosinter	¿El ejecutivo no es más electo?
v2regendtypems_0_lag_1	¿El régimen terminó por un golpe de estado? (lag 1)
v2xlg_leginter	¿Legislatura cerrada o abortada?
v2regendtype_lag_1	Proceso más importante para terminar con un régimen (lag 1)
coup_lag_1	¿Golpes de estado en el último año?

Cuadro 3: 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.
- Moore, B. (1966). *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Beacon Press.
- Gurr, T. R. (1970). *Why Men Rebel*. Princeton University Press.
- Linz, J. J., & Stepan, A. (1978). *The breakdown of democratic regimes*. Johns Hopkins University Press.
- Cardoso, F. H., & Faletto, E. (1979). *Dependencia y desarrollo en América Latina*. Siglo XXI México.
- Skocpol, T. (1979). *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge University Press.
- O'Donnell, G., Schmitter, P. C., & Whitehead, L. (1986). *Transitions from authoritarian rule: Comparative perspectives* (Vol. 3). JHU Press.
- 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>
- Huntington, S. P. (2006). *Political order in changing societies*. Yale university press.
- 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.
- Acemoglu, D., & Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Publishers.
- Mainwaring, S., & Pérez-Liñán, A. (2015). La democracia a la deriva en América Latina. *POSTData*, 20(2), 0-0.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd interntional 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.

384 Garnett, Eds.). 30. [https://proceedings.neurips.cc/paper_files/paper/2017/file/](https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf)
 385 [8a20a8621978632d76c43dfd28b67767-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf)
 386 Cebotari, A., Chueca-Montuenga, E., Diallo, Y., Ma, Y., Turk, R., Xin, W., & Zavarce, H. (2024).
 387 *Political Fragility: Coups d'État and Their Drivers*. IMF Working Paper 24/34. <https://doi.org/https://doi.org/10.23696/mcwt-fr58>
 388
 389 Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Marquardt, K. L., Medzihorsky,
 390 J., Pemstein, D., Fox, L., Gastaldi, L., Pernes, J., Rydén, O., von Römer, J., Tzelgov, E.,
 391 Wang, Y.-t., & Wilson, S. (2024). "V-Dem Methodology v14" *Varieties of Democracy (V-Dem)*
 392 *Project* (Report). <https://v-dem.net/data/reference-documents/>
 393 Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Angiolillo, F.,
 394 Bernhard, M., Borella, C., Cornell, A., Fish, S. M., Fox, L., Gastaldi, L., Gjerløw, H.,
 395 Glynn, A., God, A. G., Grahn, S., Hicken, A., Kinzelbach, K., ... Ziblatt, D. (2024). *V-Dem*
 396 *Dataset v14* Varieties of Democracy (V-Dem) Project* (Report). <https://doi.org/https://doi.org/10.23696/mcwt-fr58>
 397