

CLASIFICACIÓN DE NIVEL DE PODER EJECUTIVO SEGÚN LOS PREÁMBULOS CONSTITUCIONALES.

Stanley Gómez
Edwin Rodríguez
Wilson Gómez

INTRODUCCIÓN

El conjunto de datos utilizado para este proyecto es “Constitution Dataset.csv” creado por Prasad Patil. Este dataset proviene del proyecto Comparative Constitutions Project (CCP), una iniciativa académica reconocida que recopila información detallada sobre las constituciones nacionales de todo el mundo y se utiliza frecuentemente para análisis exploratorios, estudios políticos y experimentos con procesamiento de lenguaje natural (NLP).



INTRODUCCIÓN

1- TIPO DE TAREA: CLASIFICACIÓN

2- ¿CUÁL ES NUESTRO OBJETIVO?

En este trabajo nos propusimos investigar si es posible predecir el nivel de poder ejecutivo de un país (es decir, si es alto o bajo) únicamente a partir del texto del preámbulo de su constitución.

Lo que buscamos no es solo que el modelo acierte, sino también entender qué tanto influye el lenguaje de los preámbulos en este tipo de características estructurales de los gobiernos.

INTRODUCCIÓN

3- ¿QUÉ ES EL PODER EJECUTIVO?

El poder ejecutivo es una de las tres ramas del poder del Estado, junto al legislativo y al judicial. Es el encargado de hacer cumplir las leyes, administrar el país y tomar decisiones políticas y administrativas.

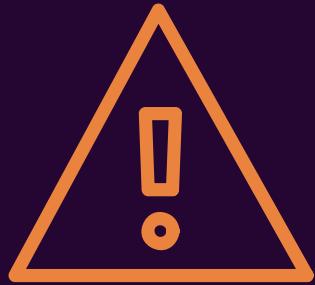
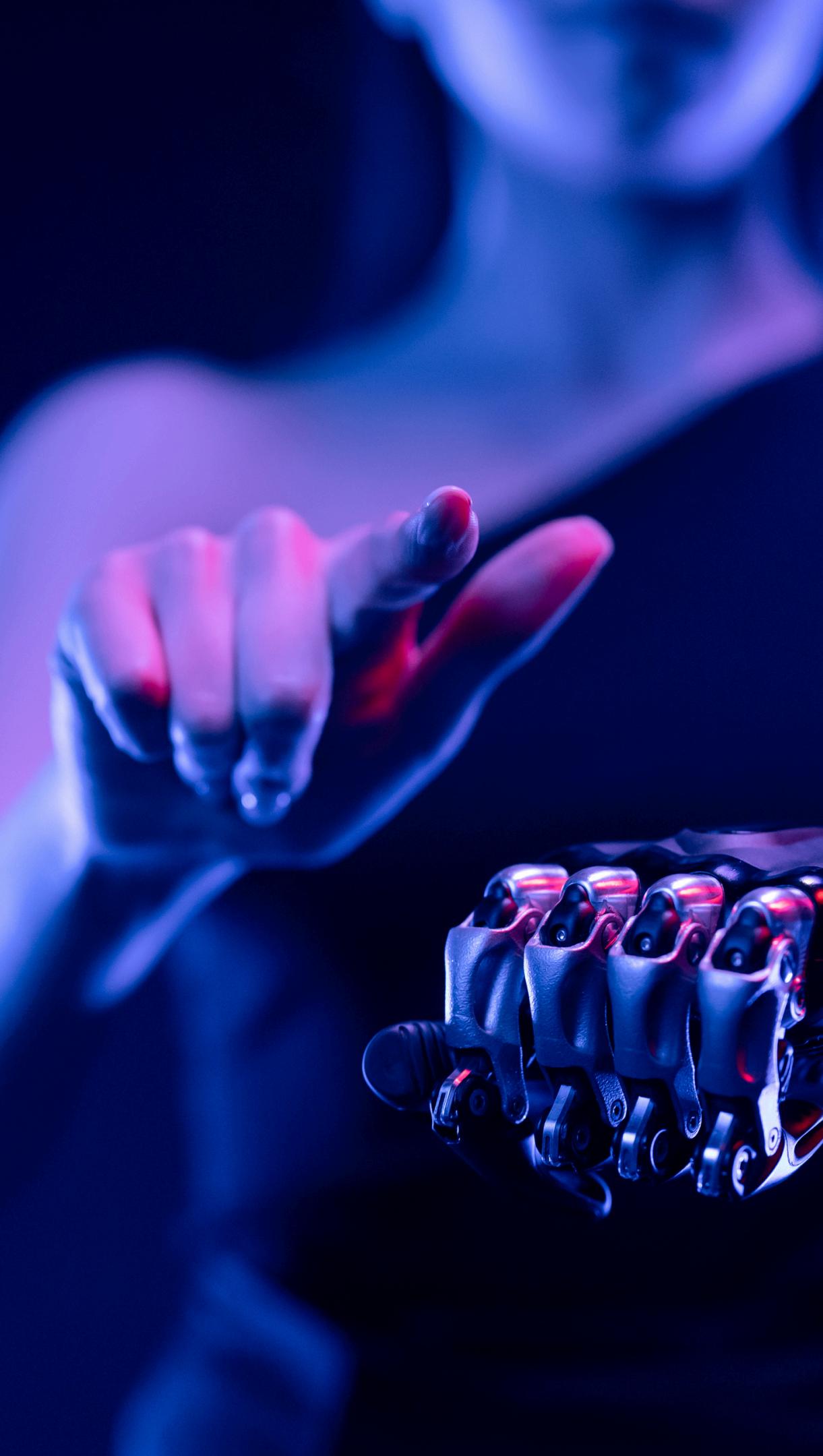
En la mayoría de los países está representado por el presidente, el primer ministro o el gobierno en general.

4- ¿PARA QUÉ SIRVE ESTO?

Este tipo de análisis sirve para entender cómo el lenguaje refleja y representa estructuras políticas, ya que, si los modelos logran identificar el nivel de poder ejecutivo solo con leer el preámbulo, eso indica que el texto constitucional transmite señales claras sobre la distribución del poder en ese país.

TABLA COMPLETA

Parametro	Descripcion	dato
Scope	Mide el porcentaje de 701 temas principales de la encuesta del PCCh incluidos en una constitución dada.	float
Length (in Words)	número total de palabras de la Constitución	int
Executive Power	un índice aditivo extraído del documento de trabajo "Restricciones Constitucionales a la Legislación Ejecutiva".	int
Legislative Power	Captura el grado formal de poder que la Constitución asigna a la legislatura.La puntuación del índice es simplemente la media de los 32 elementos binarios, donde los números más altos indican mayor poder legislativo y los números más bajos, menor.	float
Judicial Independence	un índice aditivo que va de 0 a 6 y que captura la presencia o ausencia constitucional de seis características que se cree que mejoran la independencia judicial.	int
Number of Rights	El índice de derechos indica la cantidad de estos derechos presentes en cada constitución.	int
Preamble	Contiene el texto del preámbulo de la Constitución de cada nación.	text



RETO #1

El dato más representativo que podíamos analizar era el preámbulo. Sin embargo, el problema es que este texto corresponde únicamente a una descripción general de la constitución actual de cada país.

Esto limita el análisis profundo, ya que no contamos con el contenido completo de cada constitución, solo con su introducción.



www.reallygreatsite.com



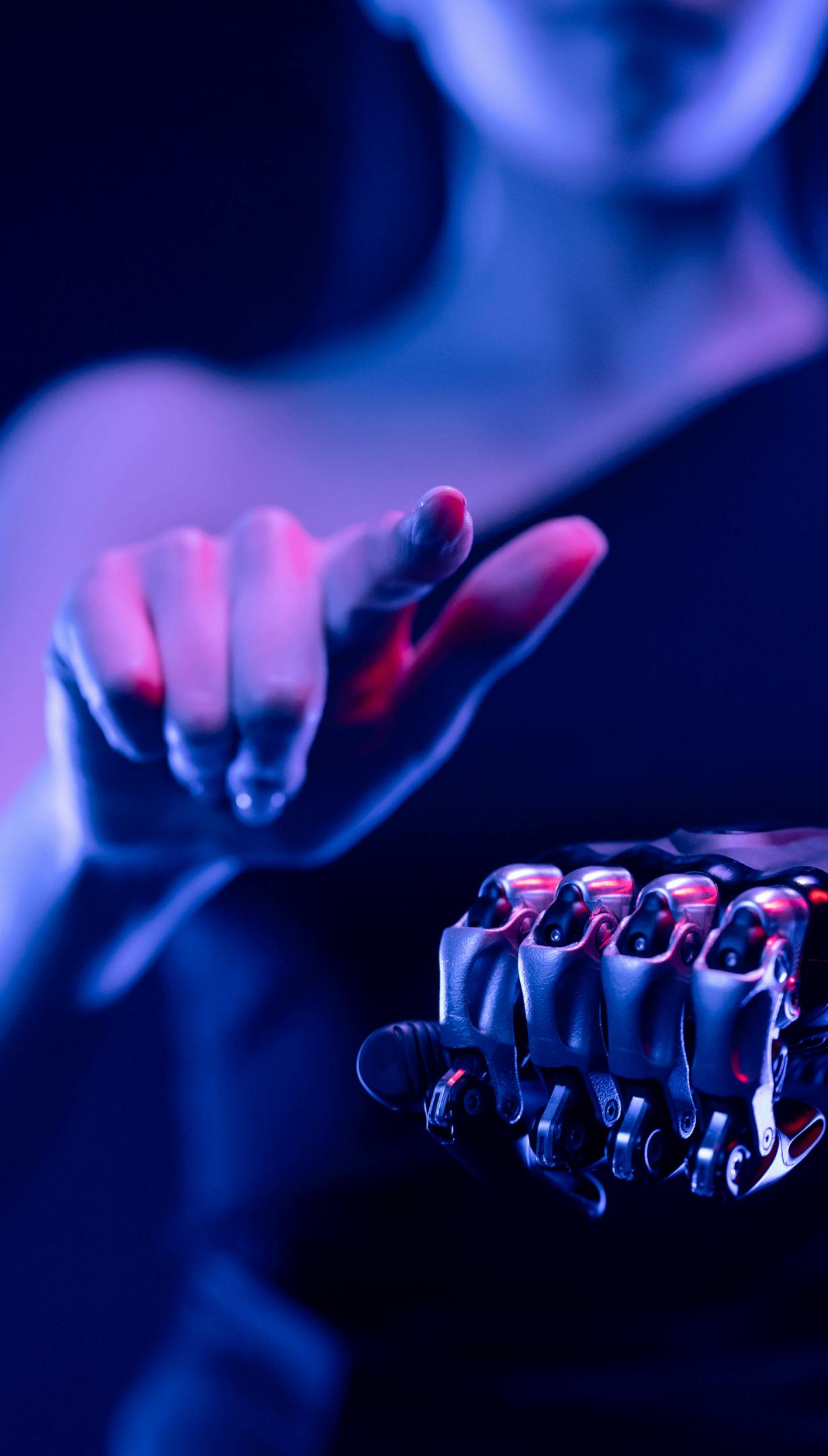


RETO #2

¿Por qué no usamos las constituciones completas para hacer las clasificaciones?

La razón principal es que algunas constituciones tienen más de 50,000 palabras. Analizar un solo documento de ese tamaño ya requiere bastante tiempo, y mucho más si pensamos en procesar más de 100 constituciones como teníamos planeado.

Por eso optamos por trabajar solo con los preámbulos.





SOLUCIÓN

Lo que hicimos fue buscar una variable existente en el conjunto de datos que el preámbulo pudiera ayudarnos a predecir.

En nuestro caso, notamos que había una cantidad casi equilibrada de preámbulos con un nivel de poder ejecutivo mayor que 4 y menor o igual a 4.

A partir de ahí, decidimos enfocarnos en investigar si el contenido del preámbulo podía servir para clasificar el nivel de poder ejecutivo en esos dos grupos.



Hay varios casos donde, al comparar directamente los preámbulos de países con un nivel bajo de poder ejecutivo con los de países con un poder ejecutivo más alto, se pueden notar diferencias en el tono, las palabras utilizadas y la intención general del mensaje.

COMPARACIÓN DE PREÁMBULOS

España(7)	Suiza(1)
<p>"En nombre de Dios Todopoderoso, Padre, Hijo y Espíritu Santo, Autor y Supremo Legislador de la sociedad: Las Cortes Generales y Extraordinarias de la nación española, convencidas, tras un minucioso examen y madura deliberación, de que las antiguas leyes fundamentales de la monarquía, acompañadas de las oportunas disposiciones y precauciones que aseguren su pleno cumplimiento de forma estable y permanente, lograrán dignamente el gran objetivo de promover la gloria, la prosperidad y el bienestar de toda la nación, decretan la siguiente constitución política para el buen gobierno y la justa administración del Estado.."</p>	<p>"¡En nombre de Dios Todopoderoso! Nosotros, el pueblo suizo y los cantones, conscientes de nuestra responsabilidad hacia la creación; resueltos a renovar nuestra alianza para fortalecer la libertad y la democracia, la independencia y la paz, en solidaridad y apertura al mundo; decididos a vivir nuestra diversidad en unidad, respetándonos mutuamente; conscientes de nuestros logros comunes y de nuestra responsabilidad hacia las generaciones futuras; y conscientes de que solo quienes ejercen su libertad son libres, y que la fuerza de un pueblo se mide por el bienestar de sus miembros más débiles; por lo tanto, adoptamos la siguiente Constitución."</p>



COMPARACIÓN DE PREÁMBULOS

Otro ejemplo interesante es la comparación entre países como Uzbekistán y Ghana, donde todavía se notan diferencias importantes en el lenguaje y la intención del preámbulo.

Uzbekistan(5)

"El pueblo de Uzbekistán, declarando solemnemente su adhesión a los derechos humanos y a los principios de la soberanía estatal, consciente de su responsabilidad última ante las generaciones presentes y futuras, apoyándose en la experiencia histórica en el desarrollo del Estado uzbeko, afirmando su compromiso con los ideales de la democracia y la justicia social, reconociendo la prioridad de las normas generalmente aceptadas del derecho internacional, aspirando a una vida digna para los ciudadanos de la República, estableciendo la tarea de crear un estado de derecho humano y democrático, con el objetivo de garantizar la paz civil y la concordia nacional, representado por sus diputados plenipotenciarios, adopta la presente Constitución de la República de Uzbekistán."

Ghana(3)

"EN NOMBRE DEL DIOS TODOPODEROSENosotros, el pueblo de Ghana, EN EJERCICIO de nuestro derecho natural e inalienable de establecer un marco de gobierno que asegure para nosotros y la posteridad las bendiciones de la libertad, la igualdad de oportunidades y la prosperidad; EN UN ESPÍRITU de amistad y paz con todos los pueblos del mundo; Y EN SOLEMNE declaración y afirmación de nuestro compromiso con; la libertad, la justicia, la probidad y la rendición de cuentas, el principio de que todos los poderes del gobierno surgen de la voluntad soberana del pueblo; el principio del sufragio universal de los adultos; el estado de derecho; la protección y preservación de los derechos humanos y las libertades fundamentales, la unidad y la estabilidad para nuestra nación; POR LA PRESENTE ADOPTAMOS, PROMULGAMOS Y NOS DAMOS A NOSOTROS MISMOS ESTA CONSTITUCIÓN".



Grecia(3)	Slovenia(6)	Reino Unido(0)
In the name of the Holy and Consubstantial and Indivisible Trinity	<p>Proceeding from the Basic Constitutional Charter on the Sovereignty and Independence of the Republic of Slovenia, and from fundamental human rights and freedoms, and the fundamental and permanent right of the Slovene nation to self-determination; and from the historical fact that in a centuries-long struggle for national liberation we Slovenes have established our national identity and asserted our statehood, the Assembly of the Republic of Slovenia hereby adopts</p>	<p>Magna Carta 1297 THE GREAT CHARTER OF THE LIBERTIES OF ENGLAND, AND OF THE LIBERTIES OF THE FOREST; CONFIRMED BY KING EDWARD, IN THE TWENTY-FIFTH YEAR OF HIS REIGN. EDWARD by the Grace of God King of England, Lord of Ireland, and Duke of Guyan, to all Archbishops, Bishops, &c. We have seen the Great Charter of the Lord Henry sometimes King of England, our Father, of the Liberties of England in these words: HENRY by the Grace of God King of England, Lord of Ireland, Duke of Normandy and Guyan, and Earl of Anjou, to all Archbishops, Bishops, Abbots, Priors, Earls, Barons, Sheriffs, Provosts, Officers, and to all Bailiffs, and other our faithful Subjects, which shall see this present Charter, Greeting: Know Ye, that We, unto the honour of Almighty God, and for the salvation of the souls of our Progenitors and Successors [Kings of England,] to the advancement of Holy Church and amendment of our Realm, of our meer and free will, have given and granted to all Archbishops, Bishops, Abbots, Priors, Earls, Barons, and to all [Freemen] of this our Realm, these Liberties following, to be kept in our Kingdom of England for ever.</p>

¿DÓNDE SE PUEDE EQUIVOCAR?

Primero, de los 190 países incluidos en el conjunto de datos, hay 31 que no tienen preámbulo, por lo tanto, no pueden ser considerados en el análisis.

Además, hay ciertos preámbulos que presentan dificultades:

- Algunos son muy largos, con estructuras complejas.
- Otros tienen un tono demasiado religioso o ceremonial, como en el caso de Grecia o Reino Unido.
- Y hay también preámbulos que, aunque contienen información útil, usan un lenguaje muy distinto, lo que puede confundir al modelo al momento de clasificar.

METODOLOGÍA



Preprocesamiento

Limpiando los
preámbulos.



Creando la columna del nivel poder ejecutivo

Agregando una nueva columna
para los valores binarios del nivel
de poder ejecutivo.



Dividiendo los datos

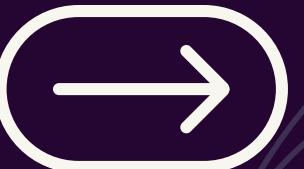
Dividendo los datos en
dataframes separados
para pruebas diferentes.



Selección de Modelo

Seleccionamos y entrenamos
algoritmos de ML para realizar
nuestras pruebas.

PREPROCESAMIENTO DE LOS DATOS



Al leer varios preámbulos presentes en el conjunto de datos, observamos que habían varias inconsistencias en cómo las palabras fueron escritas, como:

- Uso de letras mayúsculas para la palabra completa
- Uso inconsistente (y posiblemente errónea) de signos dentro de los datos

Antes que nada, hay que limpiar los textos en la columna “Preamble” para asegurar que los datos usados en los entrenamientos/pruebas de los modelos son más justos.

CÓDIGO

```
df_const = pd.read_csv("Constitution Dataset.csv")
df_const
```

	Country	Year Enacted	Scope	Length (in Words)	Executive Power	Legislative Power	Judicial Independence	Number of Rights	Preamble
0	Afghanistan	2004	0.67	10,227	6	0.38	2	37	In the name of Allah, the Most Beneficent, the...
1	Albania	1998	0.61	13,826	5	0.43	5	77	We, the people of Albania, proud and aware of ...
2	Algeria	1996	0.61	10,038	7	0.29	1	36	The Algerian people are a free people; and the...
3	Andorra	1993	0.51	8,740	6	0.19	3	51	The Andorran People, with full liberty and ind...
4	Angola	2010	0.80	27,181	7	0.19	2	80	We, the people of Angola, through its lawful r...
5	Antigua and Barbuda	1981	0.56	38,464	3	0.14	0	39	WHEREAS the People of Antigua and Barbuda pro...
6	Argentina	1983	0.64	12,514	3	0.43	6	51	We, the representatives of the people of the A...
7	Armenia	1995	0.66	13,786	6	0.29	4	82	The Armenian People, accepting as a basis the ...

```
def clean_text(text):
    text = text.lower()
    text = re.sub(r'@\w+|\#\w+', '', text)
    text = re.sub(r'^\w\s]', '', text)
    text = re.sub(r'\d+', '', text)
    text = re.sub(r'\s+', ' ', text).strip()
    return text

# eliminamos las filas que no tienen preambulos (31/190)
df_const = df_const.dropna()
df_const.loc[:, "Preamble"] = df_const["Preamble"].apply(clean_text)
```

¿QUÉ ES EL PODER EJECUTIVO?

Executive Power: Este es un índice aditivo extraído del documento de trabajo "Restricciones Constitucionales a la Legislación Ejecutiva". El índice oscila entre 0 y 7 y refleja la presencia o ausencia de siete aspectos importantes de la legislación ejecutiva:

- (1) la facultad de iniciar leyes;
- (2) la facultad de emitir decretos;
- (3) la facultad de iniciar reformas constitucionales;
- (4) la facultad de declarar estados de emergencia;
- (5) el poder de veto;
- (6) la facultad de impugnar la constitucionalidad de la legislación;
- (7) la facultad de disolver la legislatura.

El puntaje del índice indica el número total de estos poderes otorgados a cualquier ejecutivo nacional (presidente, primer ministro o asignado al gobierno) en su conjunto.

CREANDO LA COLUMNA DEL NIVEL PODER EJECUTIVO

Creamos los datos de esta columna de la forma siguiente:

Si el nivel de “Executive Power” de un pais es mas que 4, tiene un valor de 1.

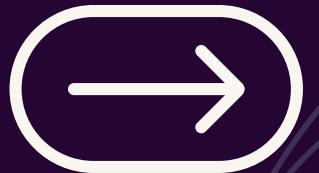
Si el nivel de “Executive Power” de un pais es igual o menor que 4, tiene un valor de 0.

```
[23]: df_const.loc[:, "nivel_poder_ejecutivo"] = df_const["Executive Power"].apply(lambda x: 1 if x > 4 else 0)

df_const.loc[:, "nivel_poder_ejecutivo"].value_counts(dropna=False)
```

```
[23]: nivel_poder_ejecutivo
      1    81
      0    78
Name: count, dtype: int64
```

DIVIDIENDO LOS DATOS



Como mencionado antes, 159 de los 190 datos originales tienen preámbulos, eso significa que los países que no tienen preámbulos tienen un valor nulo “NaN” en su lugar. Esos datos hay que eliminarlos para evitar inconsistencias/errores en los documentos.

Después de eso, tomamos la decisión de dividir los 159 datos de la forma siguiente.

- 100 serán usados para el entrenamiento de los modelos (70%)
 - 50 preámbulos cuyo nivel de poder ejecutivo es 1
 - 50 preámbulos cuyo nivel de poder ejecutivo es 0
- 44 serán usados para las pruebas iniciales de los modelos (30%)
 - 22 preámbulos cuyo nivel de poder ejecutivo es 1
 - 22 preámbulos cuyo nivel de poder ejecutivo es 0
- 10 serán usados para la prueba en vivo.
 - 5 preámbulos cuyo nivel de poder ejecutivo es 1
 - 5 preámbulos cuyo nivel de poder ejecutivo es 0

En total, hay una división total de 77:77 de datos cuyos niveles de poder ejecutivo es 1 o 0. Esto es suficiente para nuestras pruebas ya que la división original de datos de preámbulos eran 81:78.

CÓDIGO

```
# Mejor usar datos con 50/50 probabilidad exacta de ser bajo o alto nivel de poder ejecutivo para nuestros modelos

# Paso 1: separar por clase
df_0 = df_const[df_const["nivel_poder_ejecutivo"] == 0]
df_1 = df_const[df_const["nivel_poder_ejecutivo"] == 1]

# Paso 2: seleccionar primeros 50 de cada clase para el primer conjunto
df_0_train = df_0.iloc[:50]
df_1_train = df_1.iloc[:50]

# Paso 3: seleccionar los siguientes 22 de cada clase para el segundo conjunto
df_0_test = df_0.iloc[50:72]
df_1_test = df_1.iloc[50:72]

# Paso 4: concatenar conjuntos
df_train = pd.concat([df_0_train, df_1_train])
df_test = pd.concat([df_0_test, df_1_test])

# Paso 5: unir todo en un solo DataFrame
df_const_complete = pd.concat([df_train, df_test]).reset_index(drop=True)

df_const_complete.count()

Country          144
Year Enacted     144
Scope            144
Length (in Words) 144
Executive Power  144
Legislative Power 144
Judicial Independence 144
Number of Rights 144
Preamble         144
nivel_poder_ejecutivo 144
dtype: int64

# Para pruebas de los datos completos
preamble = df_const_complete["Preamble"]
poder = df_const_complete["nivel_poder_ejecutivo"]

vectorizador = TfidfVectorizer()
X = vectorizador.fit_transform(preamble)
y = poder

#X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, stratify=y)

split_index = int(0.7 * X.shape[0])
X_train = X[:split_index]
X_test = X[split_index:]

y_train = y.iloc[:split_index]
y_test = y.iloc[split_index:]
```

ESTRUCTURA DE ENTRENAR/PROBAR MODELO

1. Crear Modelo

2. Entrenar modelo
con la función fit()

3. Probar el modelo
con los resultados
con la función
predict()

4. Dar resultados de
reporte de
clasificación

5. Crear matriz de
confusión

```
# Entrenando el modelo Naive Bayes que es uno de los más usados para clasificación de texto
nb_model = MultinomialNB()
nb_model.fit(X_train, y_train)
```

```
MultinomialNB()
MultinomialNB()
```

```
# Haciendo predicciones con los datos de prueba
y_pred_nb = nb_model.predict(X_test)
```

```
# Imprimiendo el accuracy y otras métricas importantes como precision, recall y f1-score
print("Accuracy Naive Bayes:", accuracy_score(y_test, y_pred_nb))
print(classification_report(y_test, y_pred_nb))
```

MODELOS UTILIZADOS

En esta etapa del proyecto decidimos aplicar distintos algoritmos de clasificación con el fin de predecir si una constitución pertenece a un país con poder ejecutivo alto o bajo.:

- Naive Bayes
- Regresión Logística
- Árbol de Decisión
- Linear SVC
- Clasificador de Bosque Aleatorio



¿POR QUÉ FUERON ELEGIDOS?

Estos fueron elegidos porque varian en metodos y queriamos evaluar cuales modelos serian o nos darian mejores resultados en nuestro caso. Las ventajas de los modelos vienen si los estan aplicados a los datos siguientes:

- Naive Bayes
 - Bueno para: Texto, spam, documentos.
- Regresión Logistica
 - Bueno para: Clasificación binaria, problemas simples y explicables.
- Arbol de Decisión
 - Bueno para: Problemas tabulares, decisiones jerárquicas.
- Linear SVC
 - Bueno para: Clasificación en texto, alta dimensionalidad, datos bien escalados.
- Clasificador de Bosque Aleatorio
 - Bueno para: Datos estructurados/tabulares, clasificación general.



VALIDACIÓN CRUZADA

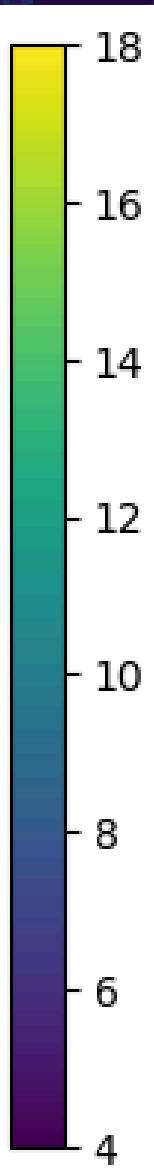
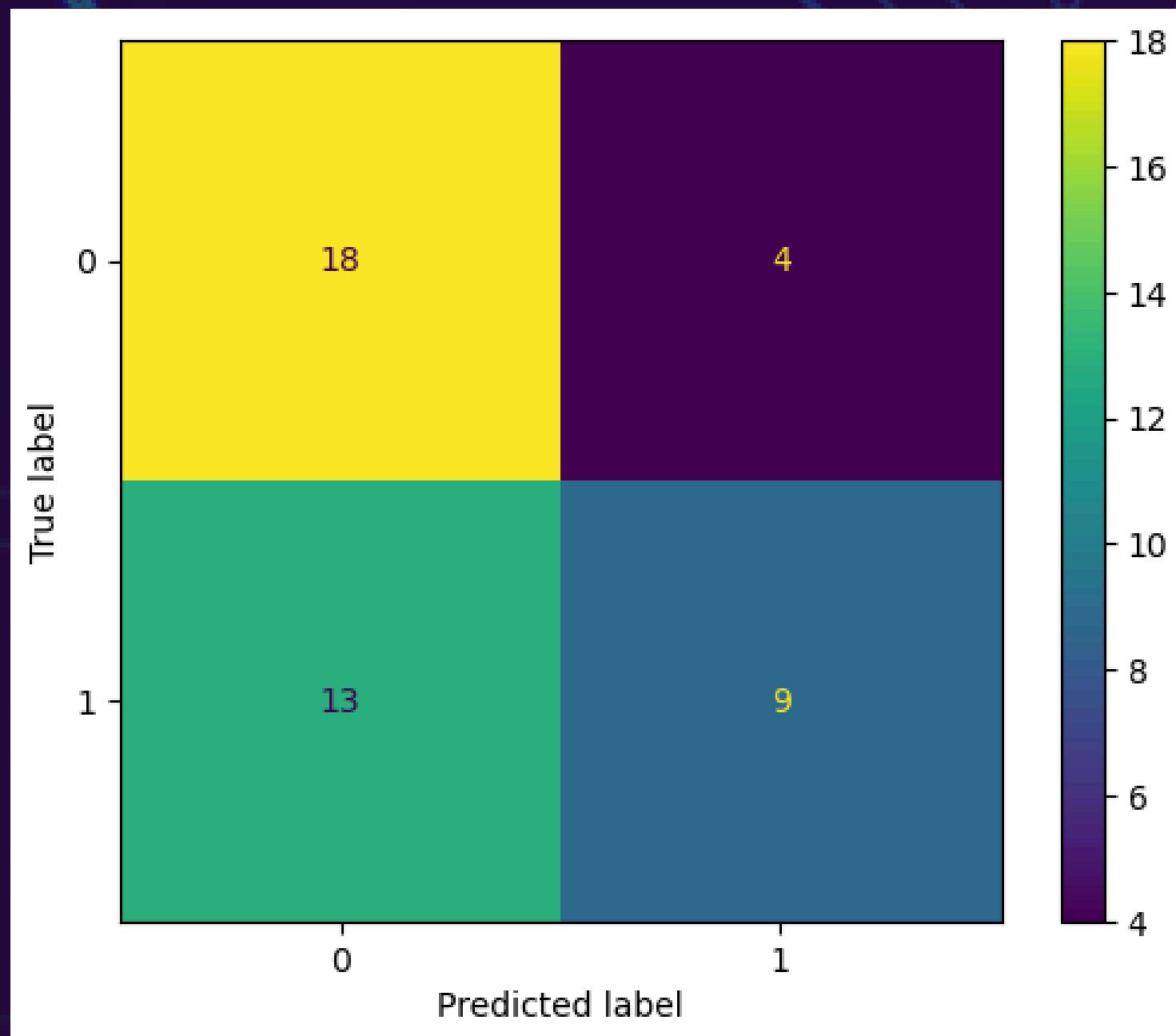
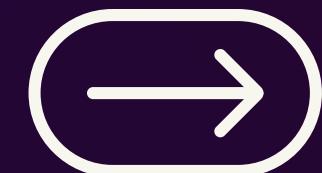
Implementamos validación cruzada estratificada con 10 folds como metodología principal de evaluación por varias razones críticas específicas a nuestro proyecto:

- Limitaciones del Dataset y Necesidad de Máximo Aprovechamiento
- Validación Robusta de Generalización

También ayuda a entender mejor los resultados iniciales de los modelos después de las pruebas iniciales. Como nuestro conjunto de datos es más limitado en comparación con otros, aplicando esta clasificación es ideal para llegar a una conclusión más clara.



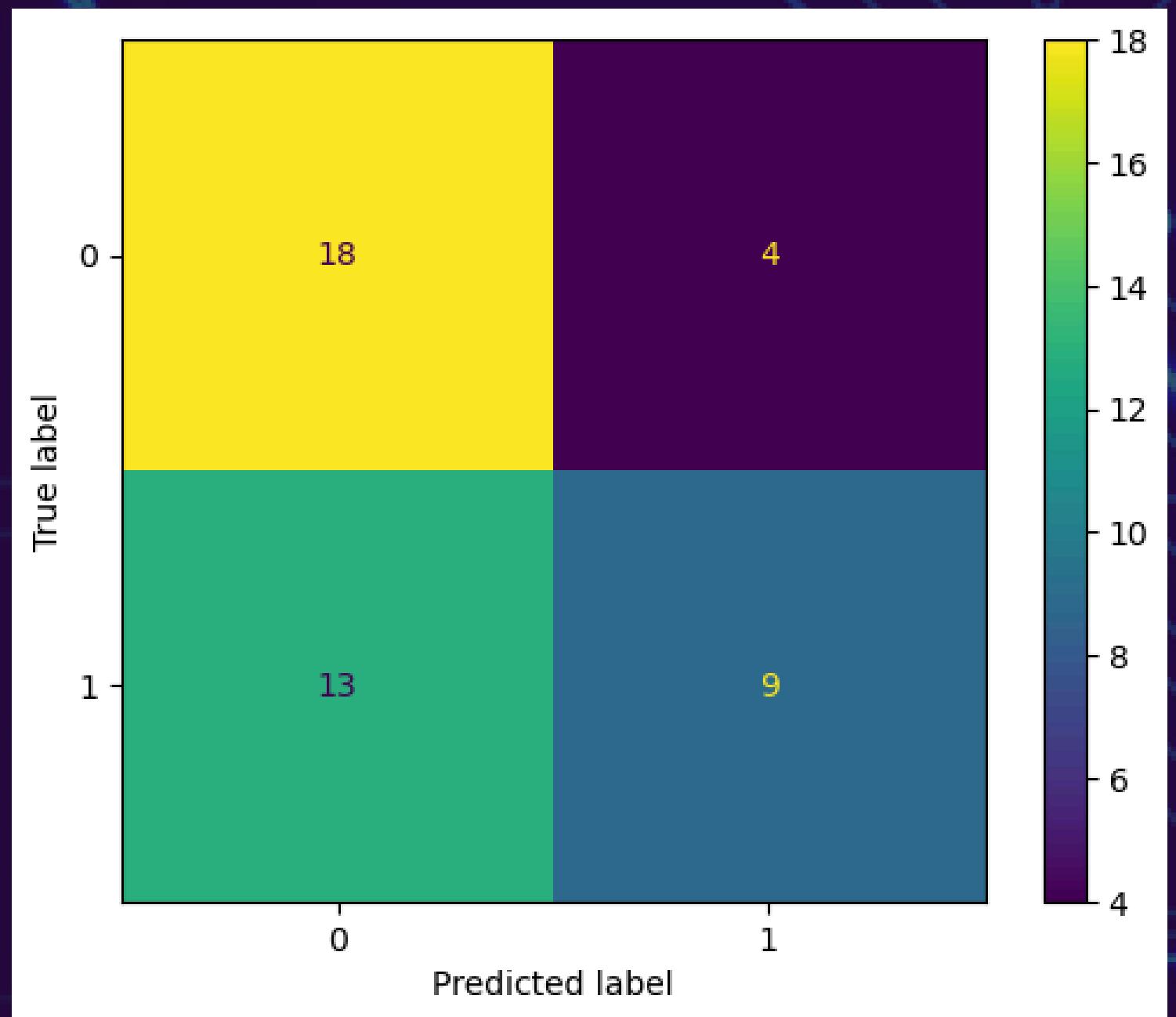
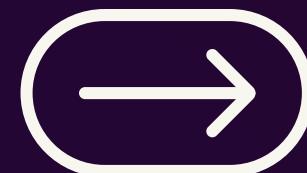
RESULTADOS (MODELOS)



		Accuracy	Naive Bayes: 0.6136363636363636	precision	recall	f1-score	support
	0	0.58	0.82	0.68	22		
	1	0.69	0.41	0.51	22		
	accuracy			0.61	44		
	macro avg			0.64	0.61	0.60	44
	weighted avg			0.64	0.61	0.60	44

Naive Bayes

RESULTADOS (MODELOS)

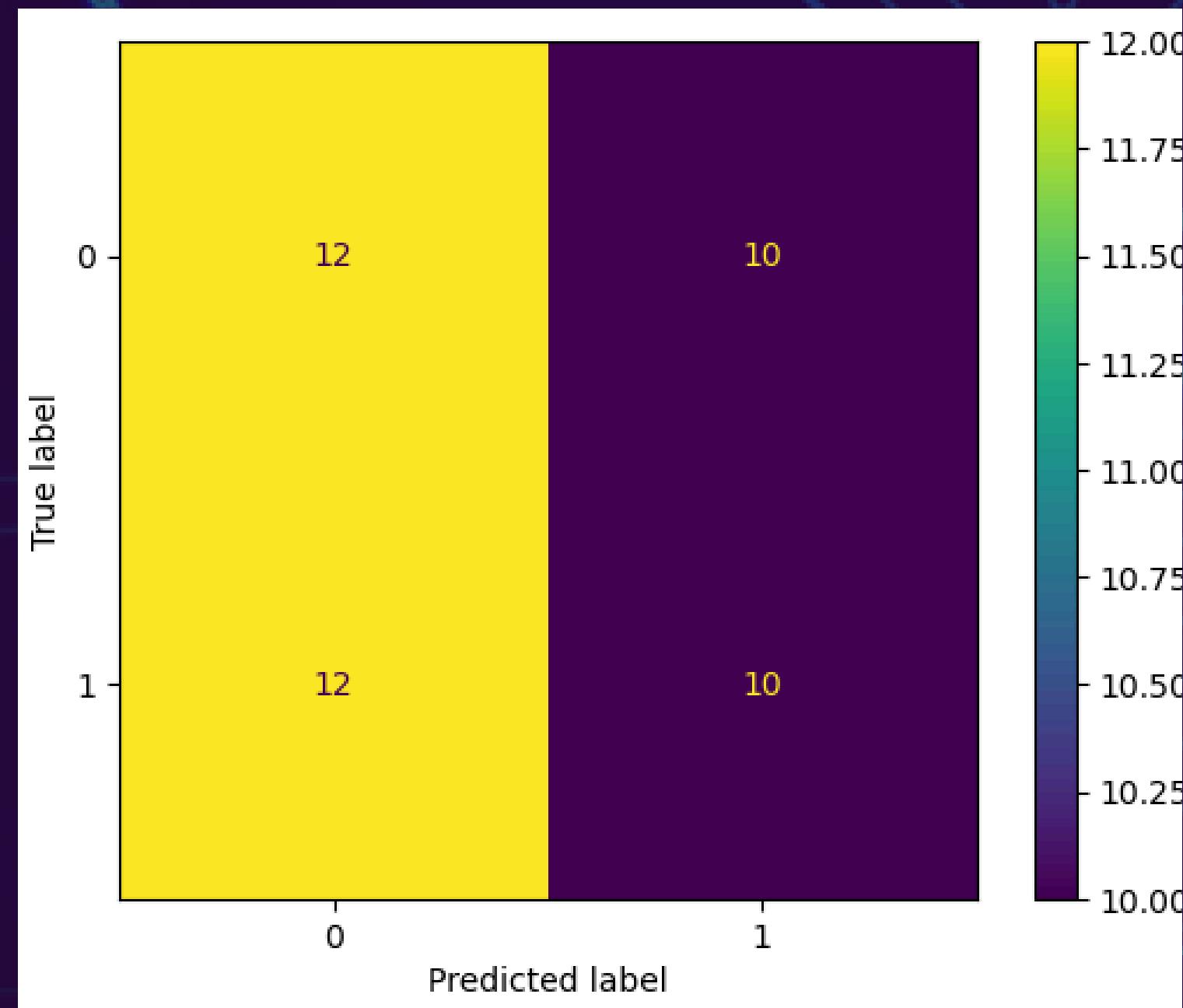
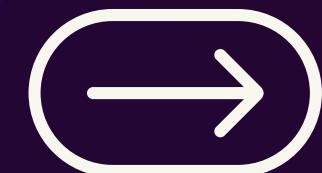


Accuracy Regresion Logistica: 0.6136363636363636

	precision	recall	f1-score	support
0	0.58	0.82	0.68	22
1	0.69	0.41	0.51	22
accuracy			0.61	44
macro avg	0.64	0.61	0.60	44
weighted avg	0.64	0.61	0.60	44

Regresion Logistica

RESULTADOS (MODELOS)

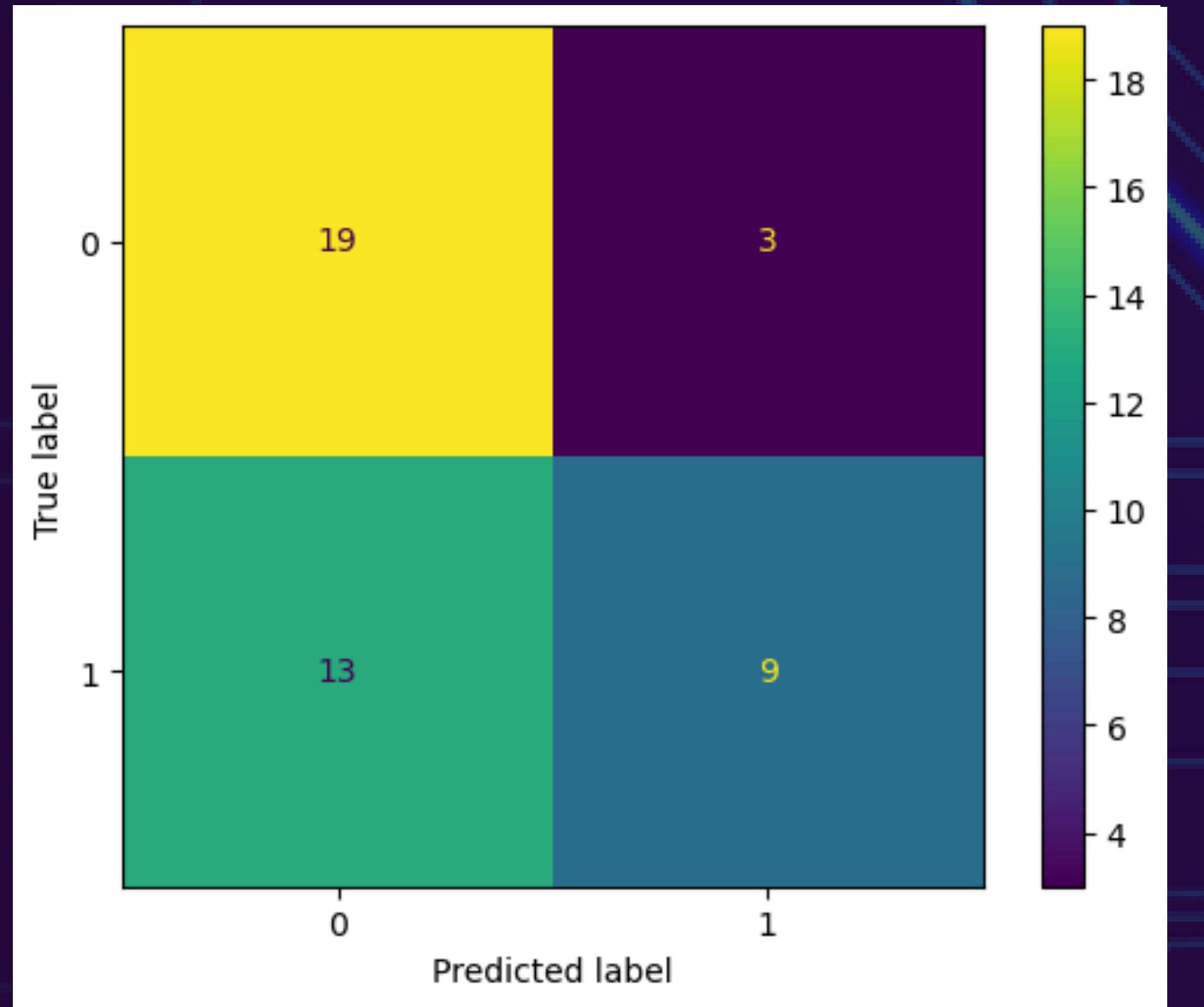
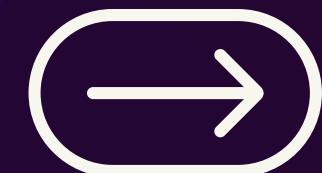


Arbol de Decision

Accuracy Árbol de Decisión: 0.5

	precision	recall	f1-score	support
0	0.50	0.55	0.52	22
1	0.50	0.45	0.48	22
accuracy			0.50	44
macro avg	0.50	0.50	0.50	44
weighted avg	0.50	0.50	0.50	44

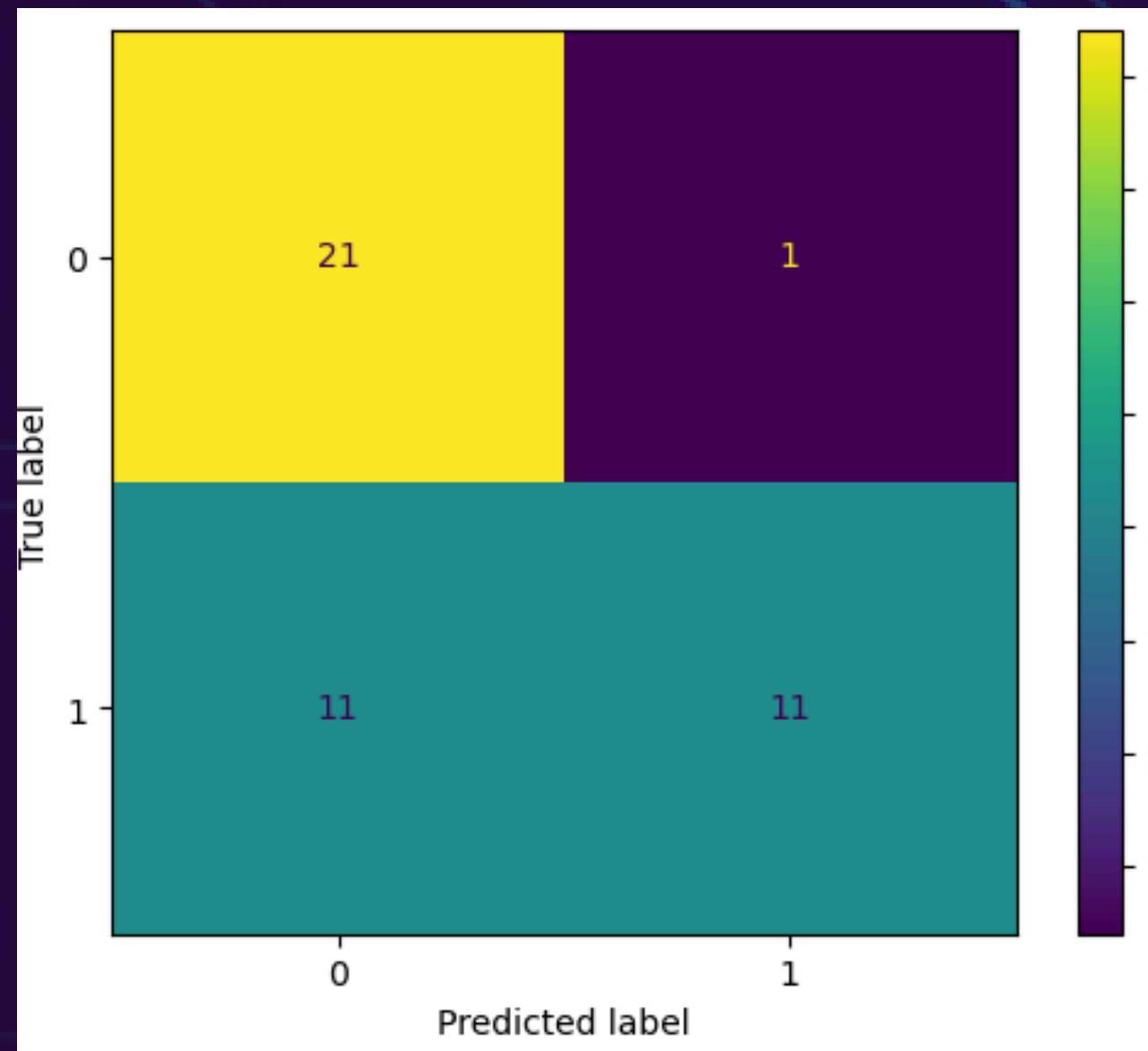
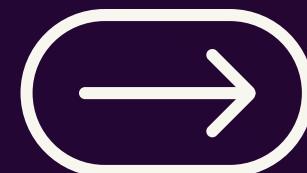
RESULTADOS (MODELOS)



		Accuracy LinearSVC: 0.6363636363636364				
		precision	recall	f1-score	support	
True label	Predicted label	0	0.59	0.86	0.70	
		1	0.75	0.41	0.53	
		accuracy		0.64	44	
		macro avg	0.67	0.64	0.62	
		weighted avg	0.67	0.64	0.62	

LinearSVC

RESULTADOS (MODELOS)

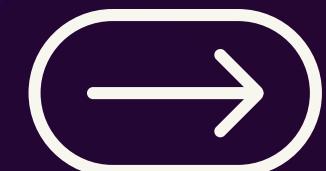


		Accuracy Random F: 0.7272727272727273				
		precision	recall	f1-score	support	
True	Predicted	0	0.66	0.95	0.78	
		1	0.92	0.50	0.65	
		accuracy		0.73	44	
		macro avg		0.79	44	
		weighted avg		0.79	44	

Clasificador de Bosque Aleatorio

RESULTADOS

(VALIDACIÓN CRUZADA)



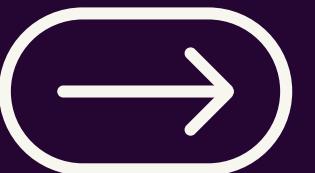
```
Naive Bayes Accuracy en cada fold: [0.66666667 0.66666667 0.6  
0.71428571 0.64285714 0.78571429 0.64285714]  
Naive Bayes Accuracy promedio: 0.6952
```

```
Regresión Logística Accuracy en cada fold: [0.73333333 0.66666667 0.6  
0.5 0.71428571 0.71428571 0.64285714]  
Regresión Logística Accuracy promedio: 0.6438
```

```
Árbol de Decisión Accuracy en cada fold: [0.6  
0.85714286 0.42857143 0.64285714 0.64285714]  
Árbol de Decisión Accuracy promedio: 0.6395
```

```
Linear Vector Machine Accuracy en cada fold: [0.6  
0.5 0.57142857 0.71428571 0.64285714]  
Linear Vector Machine Accuracy promedio: 0.6533
```

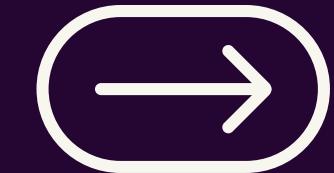
```
Random Forest Classifier Accuracy en cada fold: [0.86666667 0.73333333 0.73333333 0.53333333 0.64285714 0.42857143  
0.57142857 0.71428571 0.85714286 0.78571429]  
Random Forest Classifier Accuracy promedio: 0.6867
```



CONCLUSIONES

Este estudio demostró la viabilidad de predecir automáticamente el nivel de poder ejecutivo constitucional mediante análisis de preámbulos usando técnicas de procesamiento de lenguaje natural. Los resultados de los cinco algoritmos evaluados con entrenamiento/prueba de los datos son los siguientes: Bosque Aleatorio logró la mayor precisión promedio (73%), los modelos Linear SVC(64%), Naive Bayes(61%) y Regresión Logística(61%) mostraron resultados intermedios mientras que el modelo de Árbol de Decisión(50%) tuvo la calificación más baja. Las matrices de confusión evidenciaron un sesgo sistemático hacia la identificación correcta de constituciones de bajo poder ejecutivo (negativos verdaderos), con los modelos mostrando mayor dificultad para clasificar correctamente constituciones de alto poder ejecutivo (positivos verdaderos). Esta asimetría sugiere que las características textuales distintivas de preámbulos son más pronunciadas en constituciones que limitan el poder ejecutivo que en aquellas que lo concentran. La evaluación de validación cruzada (10 casos) mostró que Naive Bayes y Bosque Aleatorio alcanzaron aproximadamente 69% de precisión, mientras que Regresion Logistica, Árbol de Decisión y Linear SVC lograron 64-65%, confirmando que las características textuales de preámbulos contienen información predictiva significativa sobre estructura de poder ejecutivo. Cuando aplicadas a algoritmos (y después a modelos) especializados en analizar textos, hay una mejor probabilidad de precisión en los resultados.

FUTUROS TRABAJOS



La primera expansión más ideal del dataset es usar las constituciones actuales en lugar de los preámbulos. Con eso hay una ventaja significativa de poder usar datos de cada país (como que tuvimos que quitar 31 de 190 para poder realizar la investigación), y también puede garantizar mejor estabilidad de datos al entrenar/probar los modelos. Además, el análisis de constituciones completas puede facilitar la implementación de técnicas de segmentación textual para identificar secciones relevantes automáticamente, como capítulos sobre el ejecutivo, procedimientos de emergencia, y mecanismos de control del poder. Por tener la mayor cantidad de texto disponible para cada país, se mejoraría significativamente el entrenamiento de modelos de machine learning al proporcionar más características textuales para la vectorización TF-IDF.



MUCHAS GRACIAS
POR SU ATENCIÓN

DATASET ORIGINAL:
[HTTPS://WWW.KAGGLE.COM/DATASETS/PRASAD22/CONSTITUTION-DATASET](https://www.kaggle.com/datasets/prasad22/constitution-dataset)