

# ANDALUCÍA, ¿CAMBIO ESTRUCTURAL O CASO AISLADO?

*Memoria del Trabajo de Fin de Master*

Autor:  
Adrián González Chaminade

19 de mayo de 2021

# Índice

<b>1. Introducción</b>	<b>2</b>
1.1. El caso andaluz . . . . .	3
1.2. Objetivos del trabajo . . . . .	3
<b>2. Datos</b>	<b>4</b>
<b>3. Herramientas y metodología</b>	<b>4</b>
<b>4. Resultados</b>	<b>5</b>

# 1. Introducción

La política, o el arte de administrar la libertad de los demás, ha estado presente a lo largo de toda la historia de la humanidad. La necesidad de organizar sociedades humanas cada vez más complejas y crear marcos de convivencia acordes a las necesidades del momento ha permitido el nacimiento y desarrollo de gran cantidad de sistemas políticos en los que la praxis era siempre la misma. El enfrentamiento entre dos visiones enfrentadas del mundo.

Desde el último cuarto del siglo XX, el sistema que rige España es la democracia liberal, o representativa, basada en que los ciudadanos ceden la defensa de sus intereses políticos a unos representantes elegidos en procesos electorales.

Por la propia naturaleza del sistema, la capacidad de los representantes para influir en el devenir de su país o región depende de la fuerza de su alianza con otros representantes, agrupados en grupos hasta cierto punto homogéneos con ideas presuntamente similares, los partidos políticos, que se convierten con el tiempo, por la Ley de Hierro, en grandes organizaciones que son un fin en sí mismas.

Así, la existencia de los partidos, que implica el modo de vida de la élite y los cuadros medios que lo integran, pasa a depender de su desempeño en los procesos electorales, y del poder y dinero que de ellos adquieren. En consecuencia, la competencia electoral se vuelve dura y descarnada.

En cuanto a los electores, están, en su mayoría, poco interesados en el trazo fino de la praxis política. Un sistema multipartidista y proporcional como es el Español, en que gran cantidad de partidos ofrecen distintas visiones del mundo y distintas soluciones a diversos problemas, genera la necesidad de estudiar las propuestas de cada uno. Una pérdida de tiempo en muchos casos inasumible, y que es evitada mediante el empleo de la ideología, un prejuicio que permite identificar de forma rápida al más cercano a la tribu moral de cada uno, de forma que, con pocas declaraciones, pueda elegirse entre varias opciones sin necesidad de un examen riguroso. Cómo, además, la mente humana es binaria en cuestiones de moral, permite catalogar rápidamente a los distintos partidos en bloques monolíticos con ideas de trazo grueso similares. Lo que, por motivos históricos, se ha denominado *derecha e izquierda política*.

De esta manera, a cada partido, y luego a cada bloque, se le asocia una serie de conceptos ideológicos y morales que lo definen. Esto permite diluir el coste emocional de que una persona asociada a una ideología concreta vote otra opción del mismo bloque si su partido le ha decepcionado.

A partir de lo anterior, el trabajo partirá de las siguientes hipótesis:

1. La capacidad de un partido de gobernar depende de la fuerza de su bloque.
2. Los transvases de votantes entre partidos de un mismo bloque son fenómenos comunes.
3. Por el contrario, los transvases regulares de electores entre partidos de distinto bloque son fenómenos marginales.
4. En consecuencia a lo anterior, el elector que cambie de bloque lo hará generalmente de forma duradera.
5. Puntualmente, los cambios de mayoría entre dos elecciones se producen más por la movilización/abstencionismo de los bloques, a largo plazo, por trasvase de votantes de un bloque a

otro.

6. Los cambios de bloque entre electores deben tener como base circunstancias tangibles, medibles y objetivas.

### 1.1. El caso andaluz

Andalucía es una de las regiones más interesantes, políticamente hablando, de España. Por un lado, es una región que ha tenido desde el comienzo de la democracia un partido claramente dominante, el Partido Socialista Obrero Español de Andalucía, como ha ocurrido también en otras regiones como Galicia (PP de G), Euskadi (EAJ-PNV), Cataluña (CiU) o Extremadura (PSOE-E). Por otro, se da el caso de que, al contrario que en Galicia, Euskadi o Cataluña, no sólo ha habido un partido dominante, sino que éste ha sido apoyado por un bloque claramente hegemónico.

Desde las primeras elecciones, el bloque de izquierdas ha sido el mayoritario en todas las elecciones generales, y todas las elecciones autonómicas. En el caso más extremo, durante el periodo 1989/1990, el bloque llegó a amasar más del 70 % de los votos.

Sin embargo, pese a esta historia de éxito, en las elecciones autonómicas del año 2018, el bloque de la derecha consiguió una victoria inesperada, alcanzando el 51 % de los votos y desalojando por primera vez en la historia al PSOE-A de la Junta de Andalucía.

Los resultados tomaron por sorpresa a la mayor parte de los analistas del país, lo que hace pensar que fue un caso puntual no necesariamente reproducible, consecuencia de circunstancias como el hartazgo por la corrupción (las elecciones se produjeron en medio del juicio por el caso E.R.E.) o el hastío tras 36 largos años de gobierno socialista. Pero, ¿fue realmente así? ¿O, tal vez, es un resultado que llevaba gestándose desde hacía tiempo?

### 1.2. Objetivos del trabajo

Los objetivos del trabajo van a ser, básicamente, tres, relacionados con lo dicho anteriormente.

1. Deducir si los resultados del año 2018 responden a un hecho aislado fruto de circunstancias puntuales o, por el contrario, responden a una tendencia a largo plazo.
2. Independientemente del resultado del punto anterior, encontrar las causas de dicho resultado.
3. En política, lo fundamental es gobernar, para lo que hace falta ganar elecciones. ganar elecciones implica tomar las decisiones adecuadas en los momentos adecuados. Para ello, no basta con explicar lo pasado, sino que es necesario ser capaz de predecir el futuro. Por ello, el último objetivo será tratar de construir un modelo solvente capaz de predecir con acierto los resultados electorales de las siguientes elecciones, generales o autonómicas y que, además, sea consistente con el análisis anterior.

## 2. Datos

Para construir el dataset sobre el que se aplicará el análisis y el modelado, se recurrirá a datos públicos, tanto de organismos públicos como de webs privadas. Los links para acceder a las fuentes se incluyen en los notebooks en los que se tratan. Para el trabajo, se emplearán los siguientes datos:

- Los resultados de las elecciones, tanto a nivel municipal como a nivel provincia, para autonómicas y generales.
- Los datos de población por municipio. Se dispone de serie histórica desde 1998, aunque se dispone de datos para todos los municipios para las elecciones generales.
- Los datos de paro por municipios. Se dispone de serie mensual, por sectores (agricultura, construcción, industria y servicios) y en total, desde 2005.
- Los datos de renta por municipio. Se dispone de serie anual entre 2013 y 2018.
- Los datos de inmigración por municipio. Se dispone de serie anual desde 1998, e incluye distintos desagregados por año. Por motivos de uniformidad, se ha decidido agrupar por continente de procedencia.
- Los datos de delincuencia por provincias. Se dispone de serie anual desde 2013, e incluye distintos desagregados según año. Por motivos de uniformidad, se ha decidido reducir al mínimo, quedando los datos desagregados por continente (fuera de Europa), España (con nacionalidad), y europea, distinguiendo entre comunitarios y extracomunitarios.
- Los datos de educación. Se dispone de serie anual desde 2014, aunque para toda la comunidad. Aún así se considera de utilidad.
- Además de los anteriores, se han incluido otros datos que podrían tener influencia, tales como si el municipio es o no costero, el número de partidos grandes presentados por bloque, eventualidades sucedidas que afectaron a las elecciones o, incluso a las tendencias, datos meteorológicos (temperatura y precipitaciones) y el bloque que ostenta, para cada elección, los gobiernos nacional y autonómico.

## 3. Herramientas y metodología

El proyecto ha sido realizado casi totalmente en python, con excepción de unas pocas líneas de código destinadas a la visualización de resultados a nivel provincial, escritas en R.

El proceso de obtención y tratamiento de datos ha sido recopilado en cuatro Jupyter Notebooks suficientemente explicados y listos para ejecutarse directamente. Por último, además de los Notebooks, y las conclusiones expuestas en ellos, se ha creado una aplicación web mediante la librería Dash. Ésta ha sido subida, aunque de forma incompleta, debido a limitaciones de memoria, a internet mediante la plataforma heroku.com. El enlace ha sido adherido al *readme.md* del repositorio. La aplicación completa puede verse ejecutando localmente el programa *index.py*.

Los Notebooks, que se exponen a continuación con una pequeña descripción, deberán ser ejecutados secuencialmente, ya que los resultados de cada uno son empleados por el siguiente.

#### 1. Recolección y limpieza de los datos.

- Se importan los datos de las distintas fuentes, tanto los que han sido descargados como los que son obtenidos directamente de la web por webscraping. Tras ello, se realizan la limpieza y transformaciones necesarias. También se realizan varias visualizaciones para llevar a cabo un análisis exploratorio previo al estudio.

#### 2. Fusión de los datos.

- Recibe los datos generados en el anterior notebook y los fusiona en un único conjunto de datos. A éste se le añaden algunos datos, principalmente los del último punto de la sección de datos.

#### 3. Análisis de los datos.

- Recibe el conjunto de datos completo y crea visualizaciones que sirven de base para el análisis, exponiendo las conclusiones del mismo.

#### 4. Modelado de los datos.

- Recibe el conjunto de datos completo y lo preprocesa para construir modelos de clasificación y regresión para predecir el resultado electoral en cada municipio.

## 4. Resultados

El primer objetivo del proyecto era deducir si el triunfo de la derecha en las elecciones autonómicas de 2018 era un caso aislado en la historia electoral de la región o respondía a una tendencia a largo plazo. Observando las gráficas con los resultados electorales a nivel comunidad,

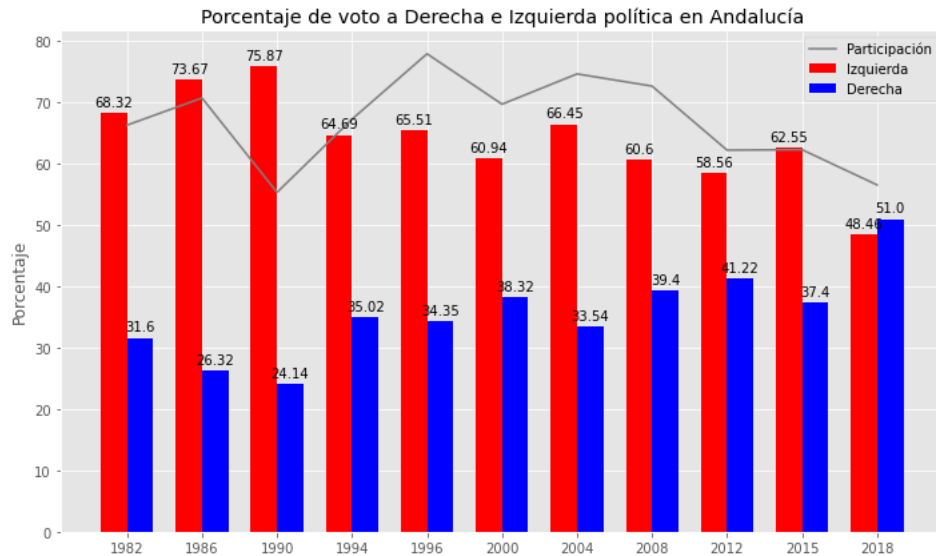


Figura 1: Resultados electorales por bloque y participación en las elecciones autonómicas de Andalucía.

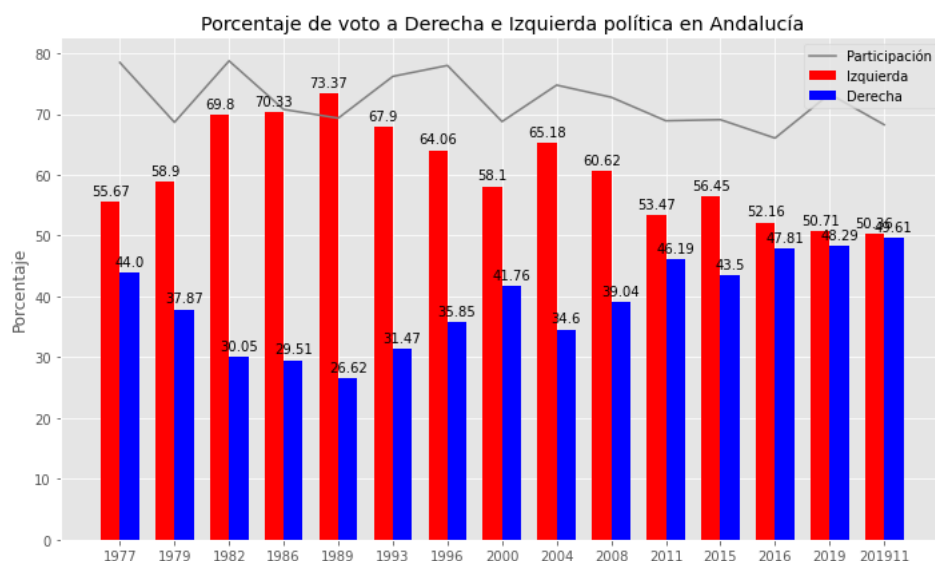


Figura 2: Resultados electorales por bloque y participación en las elecciones generales en Andalucía.

Se observa claramente, aunque de forma más clara en las elecciones generales, que la izquierda política crece de forma consistente desde los primeros comicios hasta el periodo 1989/1990, alcanzando hasta casi el 76 % de los votos en 1990, y cómo, a partir de ese momento, comienza a descender en favor de la derecha, con marcadas excepciones: 2004 y 2015. 2004 es un resultado que es necesario tomar con mucho cuidado debido a las circunstancias en las que se produjo.

En cualquier caso, viendo los resultados como diferencias entre bloque,

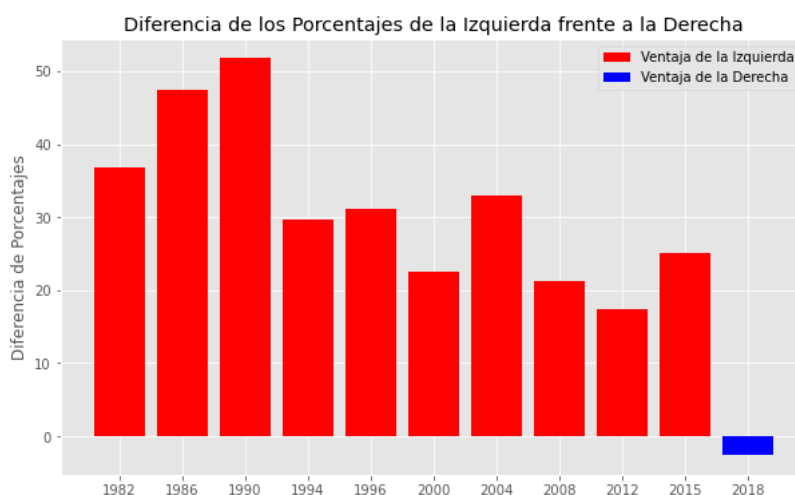


Figura 3: Resultados electorales por diferencia entre bloques para las elecciones autonómicas

se observa claramente que el resultado de 2018 es absolutamente excéntrico y no se explica por si sólo con las tendencias anteriores. Por ello, se analizaron los datos, observando

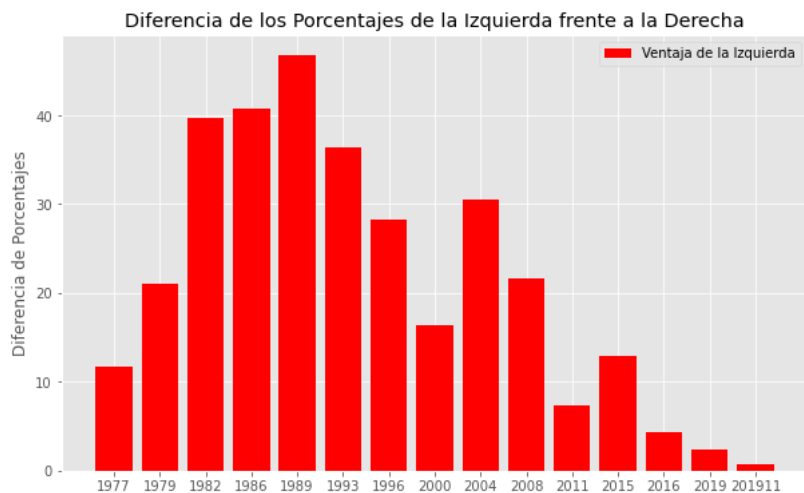


Figura 4: Resultados electorales por diferencia entre bloques para las elecciones generales

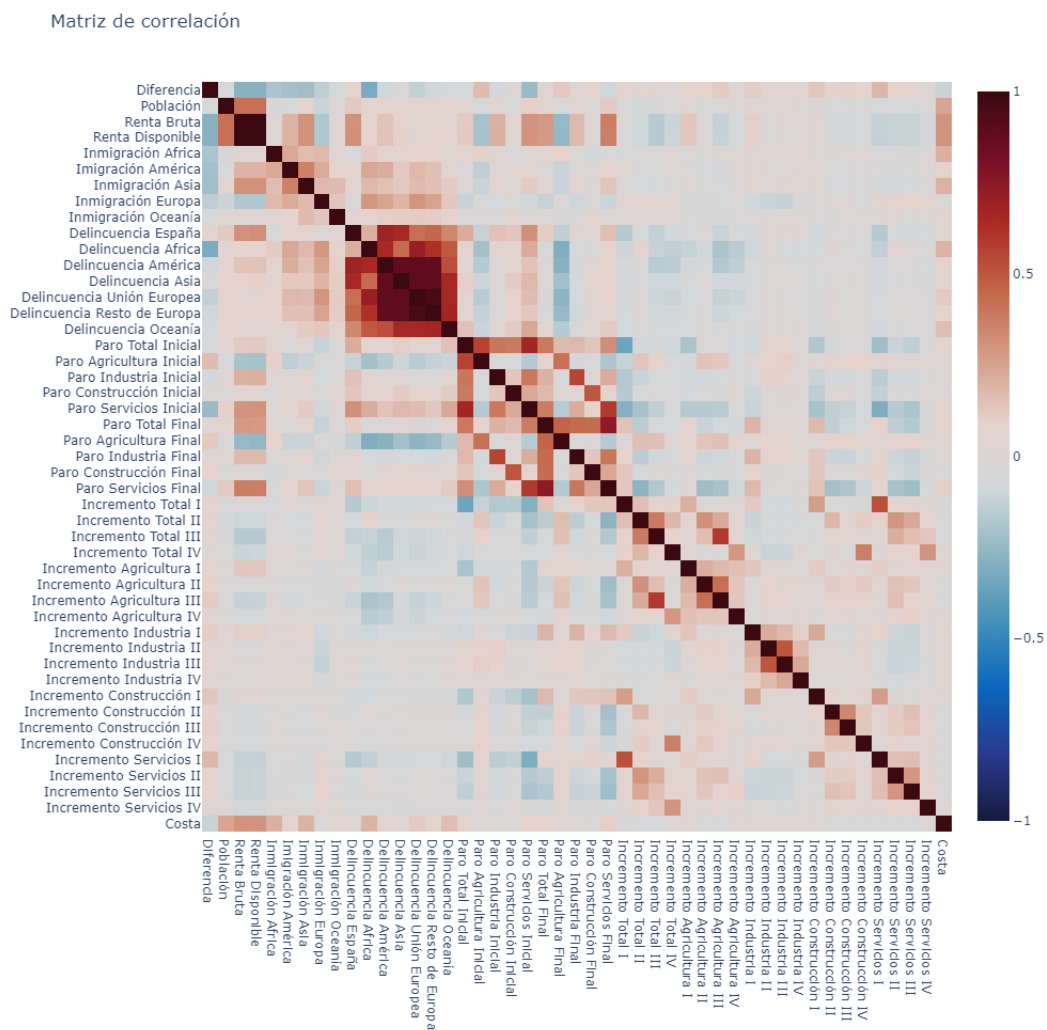


Figura 5: Matriz de correlación entre una serie de variables elegidas.



se ve que no existe una gran correlación entre las distintas variables y los resultados electorales, aunque sí existen algunas relaciones importantes a tener en cuenta, a saber

- Parece haber una cierta correlación positiva entre la renta y que el municipio en cuestión sea costero, así como respecto a la población o la inmigración de origen africano y asiático.
- Hay una importante correlación negativa entre la tendencia de voto y la Delincuencia, a nivel general, pero especialmente de origen africano, y después de la Unión Europea. Por el contrario, la delincuencia de origen español parece tener una ligera tendencia positiva.
- Existe, del mismo modo, una obvia correlación negativa entre renta y paro.
- También existe correlación entre renta e inmigración asiática y americana, frente a la de origen europeo o africano, en cuyo caso la correlación se vuelve negativa.
- Existe una clara correlación entre inmigración de todos los orígenes y voto a la derecha, aunque esta se da especialmente respecto a la inmigración de origen asiático, americano y africano, algo menos para la de origen europeo y prácticamente nulo para la de origen oceánico, lo que se puede explicar por lo escaso de esa inmigración.
- Hay, en general, escasa correlación entre la diferencia de votos y el paro, salvo en servicios, donde la correlación es negativa, y en agricultura, donde la correlación es positiva.

Mas detalladamente, realizando un gráfico de dispersión,

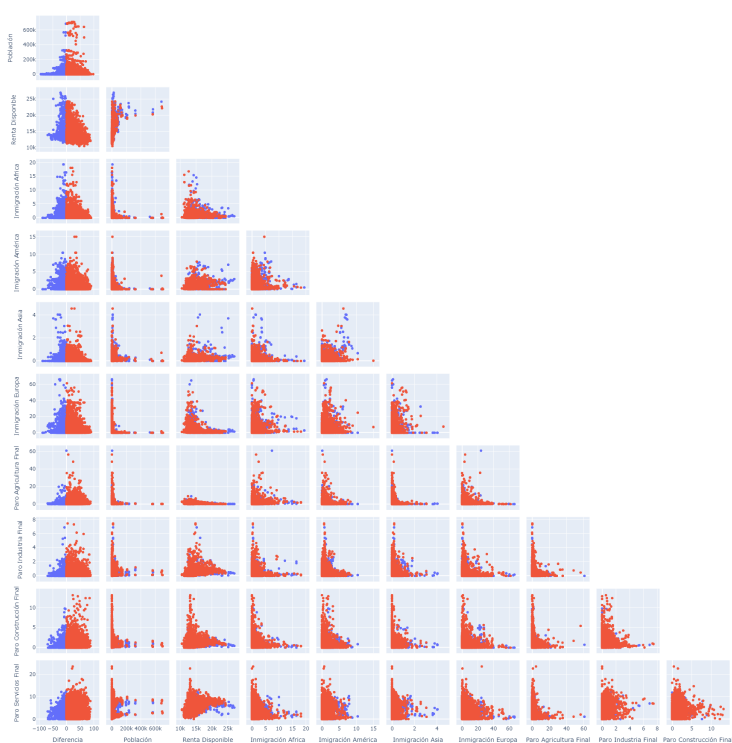


Figura 6: Matriz de dispersión.

pueden extraerse estas conclusiones:

- Existe un claro factor de clase en la distribución del voto.
- Los municipios con mayor tasa de paro tienden a votar más a la izquierda, y, aunque se habla mucho del impacto del P.E.R. y el paro Agrario en las victorias electorales del Partido Socialista, se ve cómo la izquierda se ve beneficiada por las tasas de paro en todos los sectores, de forma similar, por lo que, de ser cierto el caso, no parece determinante.
- Existe una clara relación entre el voto a la derecha y la inmigración de todo origen, aunque es menos evidente para el caso de la inmigración de origen americano, quizá por el idioma y la cultura.
- Parece claro que los municipios más pequeños son, también, los más frentistas en cuanto a posiciones políticas. Las diferencias en ellos tienden a ser mucho más extremas que en los grandes municipios urbanos. Es decir, hay una mayor, y más cerrada tradición de voto en los municipios pequeños, lo que tiende a favorecer a la izquierda, que en los municipios grandes, donde existe más pluralismo y más oportunidad de que muten las mayorías, lo que favorece, últimamente, a la derecha.

Así, tomando todas las consideraciones anteriores, pueden considerarse los siguientes motivos para el resultado de 2018:

1. Paro en mínimos desde el comienzo de la crisis.
2. Renta en crecimiento para la mayor parte de los municipios desde 2014.
3. Inmigración y delincuencia en máximos.
4. La corrupción que había lastrado el resultado de la derecha en 2015 ha dado pasado a nuevos casos que lastran ahora a la izquierda.
5. La distribución de la población favorece un resultado propicio a la derecha, dado que aumenta el peso de los grupos de municipios menos escorados a la izquierda en detrimento del resto.
6. Los casos de corrupción afectaban al Gobierno Autonómico, en unas elecciones autonómicas, lo que penalizó aún más a la izquierda.

Por último, el objetivo final, aunque secundario, era conseguir un modelo capaz de predecir con razonable acierto los resultados para cada municipio y, además fuera consistente con los resultados anteriores.

Se han probado diferentes algoritmos de clasificación y regresión, pero el mejor de los modelos ha surgido del algoritmo regresor de la biblioteca XGBoost, que ha dado como resultados, por un lado prediciendo el ganador de los municipios, un acierto del 94 %, con un recall para la clase minoritaria (la derecha política) del 71 % y un balance del 84 %, mientras que para el cálculo concreto de los resultados ha dado una varianza explicada del 82 %, con un mae de 7.72 y un rmse de 141.77. Teniendo en cuenta que el cálculo se efectúa sobre la diferencia, que va de -100 a 100, un error de menos de 8 parece un buen resultado.

En cuanto a su parecido con los resultados del análisis, se ha estudiado el modelo con la biblioteca SHAP.

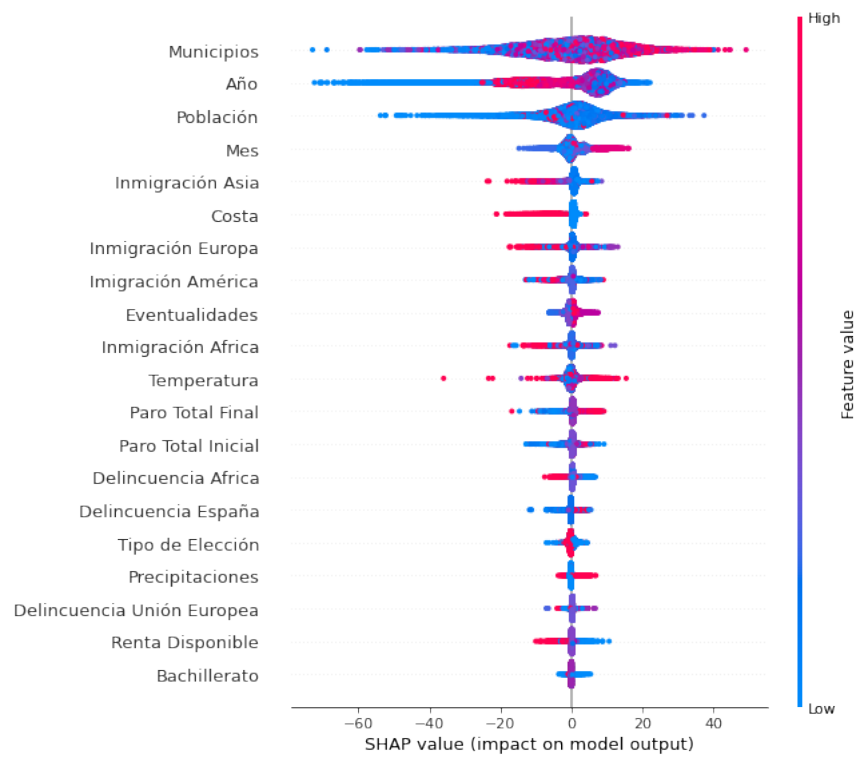


Figura 7: Gráfica de SHAP en la que se incluye el impacto y la importancia de las principales características.

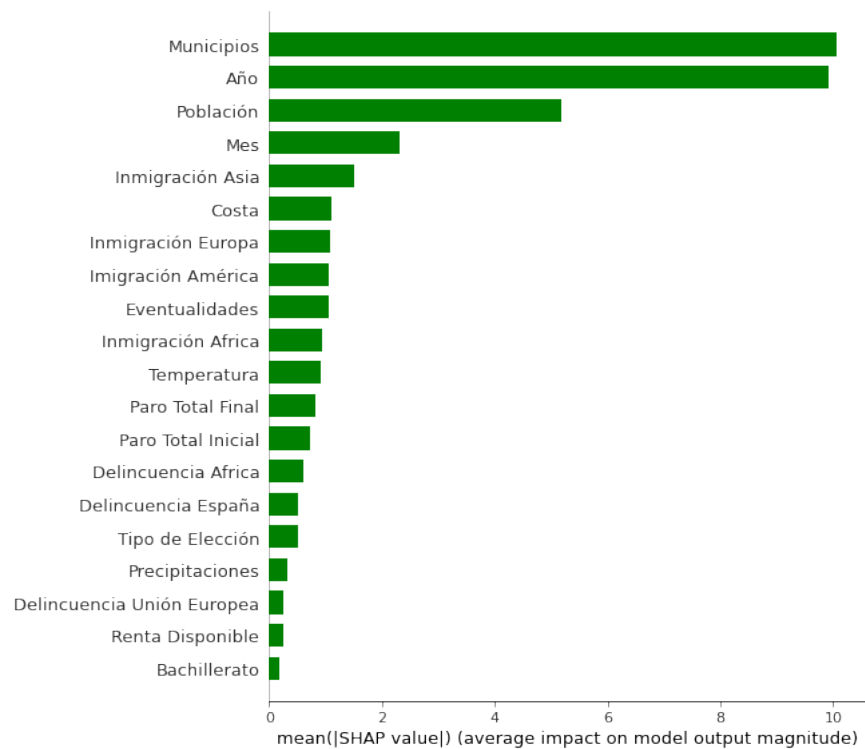


Figura 8: Gráfica con la importancia de las características principales

Aunque a priori podría parecer que sus resultados no coinciden con los de el análisis, realmente son totalmente compatibles, ya que,

- La tradición de voto es importante generalmente, pero es especialmente importante cuando su impacto en la predicción es positivo. Es decir, cuando la predicción da la victoria a la izquierda. Esto es consecuente con lo revelado por el análisis sobre la tradición de voto.
- El ciclo electoral impacta más, y tiene un impacto más importante, cuando afecta de forma negativa. Es decir, cuando la predicción es favorable a la derecha. Esto es razonable, dado que la tendencia general para la derecha es de crecimiento.
- La población tiene un mayor valor cuando su impacto se encuentra en torno a cero, mientras que, a mayor valor, menor impacto. Esto es consecuente con el análisis, que concluye que en las ciudades grandes, donde se produce el mayor impacto de la población, pues tienden a una menor tradición de voto, se producen resultados más ajustados, que benefician a priori a la derecha por la comparación.
- De los diversos orígenes de la inmigración, el que más impacto tiene es el asiático. El que menos, el africano. Sin embargo, tal y como afirmaba el análisis, la inmigración, cuando afecta de forma importante, tiene un impacto más favorable para la derecha, aunque con matices para la inmigración de origen americano.
- El Paro, especialmente a final de la legislatura, tiende a favorecer a la izquierda, tal y como explicaba el análisis.
- La delincuencia que tiene efecto, y reducido, es la de origen africano, principalmente, comunitario y español. La primera beneficia claramente, como ya se había dicho, a la derecha, mientras que la española beneficia a la izquierda y la europea comunitaria se reparte, aunque con ligero beneficio para la izquierda.
- La renta beneficia claramente a la derecha, siendo importante cuando más impacto a favor de ella tiene, perdiendo importancia cuando su impacto es nulo, y no teniendo apenas cuando tiene impacto a favor de la izquierda.

Además de lo anterior, el modelo muestra otras observaciones interesantes:

- El mes en que se convocan las elecciones tiene una gran importancia, y tiende a beneficiar claramente a la izquierda, particularmente a inicios de año.
- Las eventualidades tienen una importancia moderada en la mayor parte de los casos, pero su impacto favorece de forma clara a la izquierda.
- La importancia del tipo de elección se concentra en los momentos en los que su impacto es nulo, aunque muy ligeramente inclinado a favor de la derecha.
- El porcentaje de la población cuyo nivel de estudios se mantiene en el bachillerato, único a tener en cuenta en cuestión de educación, tiene un impacto generalmente nulo.