

Proyecto de Iniciación a la Investigación

Título descriptivo: **Predicción de las elecciones parlamentarias con aprendizaje automático**

Apellidos, Nombre del estudiante: **Ibáñez González, Miguel**

Apellidos, Nombre del Tutor: **Dominguez Carreta, David Renato**

Curso: **1º Master en Ingeniería Informática**

Fecha: (27/02/2025)

1 Antecedentes, motivación y objetivos

Antecedentes

Las elecciones parlamentarias son muy importantes para comprender las dinámicas sociales y políticas de España, ya que reflejan las preferencias y prioridades de la población. El estudio de los resultados electorales es una oportunidad para identificar patrones y predecir la orientación política de la población. Con el avance de las técnicas de aprendizaje automático y la disponibilidad de herramientas computacionales, se puede abordar este problema de manera sistemática y generar modelos capaces de realizar predicciones precisas.

Motivación

La predicción de orientación política a partir de características sociales y demográficas no sólo tiene aplicaciones prácticas en campañas electorales y estudios de mercado político, sino que también contribuye a la comprensión de cómo las características de la población influyen en sus decisiones electorales. En este trabajo se pretende aprovechar las bases de datos de resultados históricos para explorar estas relaciones y construir modelos predictivos. Este estudio permite poner en práctica técnicas avanzadas de aprendizaje automático, como algoritmos de clasificación y regresión y redes neuronales.

Objetivos

El principal objetivo de este trabajo es desarrollar modelos de aprendizaje automático que puedan predecir la orientación política (izquierda, derecha, regionalista) de una población en base a un conjunto de características socio-demográficas. Para alcanzar este objetivo, se plantean los siguientes sub-objetivos: recolectar y procesar datos, explorar y analizar datos, implementar modelos de predicción, evaluar el rendimiento, optimización de modelos y aplicabilidad.

2 Actividades y cronograma

Actividades:

Recolección de datos (14 días)

Búsqueda de bases de datos de elecciones parlamentarias desde 1996 hasta 2024.

Obtención de datos demográficos de fuentes oficiales (INE, CIS, etc.).

Limpieza y preprocesamiento de los datos para crear una única base de datos que se usará en los modelos.

Exploración y análisis de datos (14 días)

Análisis exploratorio de datos (EDA) utilizando herramientas como Python (Pandas, Matplotlib, Seaborn).
Identificación de patrones, correlaciones y distribuciones.
Análisis de outliers y tratamiento de datos faltantes.

Diseño e implementación de modelos (35 días)

Implementación de modelos: regresiones lineales, KNN, redes neuronales, ajustes cuadráticos SVM, regresión logística, árboles de decisión, random forest y modelo de Bayes.
Uso de bibliotecas como Scikit-Learn, TensorFlow o PyTorch.

Evaluación de modelos (21 días)

Comparar los modelos en términos de métricas de rendimiento. Los modelos de aprendizaje para tareas de clasificación, como Regresión Logística, Modelo de Bayes, Árboles de Decisión, Random Forest, SVM, KNN y Redes Neuronales, serán evaluados mediante precisión, recall, F1-score, matriz de confusión y AUC-ROC. Para tareas de regresión, como las abordadas por Árboles de Decisión, Random Forest, SVM, Regresión Lineal, KNN y Redes Neuronales, se utilizarán métricas como MAE, RMSE y R^2 .

Optimización de modelos (14 días)

Ajuste de hiperparámetros mediante Grid Search y Random Search.
Validación cruzada para asegurar un rendimiento consistente.

Análisis de resultados y aplicabilidad (14 días).

Análisis de la aplicabilidad de los modelos seleccionados en contextos reales.
Proporcionar un análisis detallado de cómo estos modelos pueden aplicarse en situaciones prácticas y discutir sus limitaciones. Discusión de limitaciones y posibles mejoras.

Elaboración del informe final (21 días)

Redacción de la memoria del proyecto incluyendo antecedentes, metodología, resultados, discusión y conclusiones. Revisión y corrección del documento.

Cronograma

Proceso	Fecha Inicio	Fecha Fin	13/01 - 19/01	20/01 - 26/01	27/01 - 02/02	03/02 - 09/02	10/02 - 16/02	17/02 - 23/02	24/02 - 02/03	03/03 - 09/03	10/03 - 16/03	17/03 - 23/03	24/03 - 30/03	31/03 - 06/04	07/04 - 13/04	14/04 - 20/04	21/04 - 27/04	28/04 - 04/05	05/05 - 11/05	12/05 - 18/05	19/05 - 23/05
Recolección de Datos	13/01/2025	26/01/2025																			
Exploración y Análisis de Datos	27/01/2025	09/02/2025																			
Diseño e implementación de modelos	10/02/2025	16/03/2025																			
Evaluación de modelos	17/03/2025	06/04/2025																			
Optimización de modelos	07/04/2025	20/04/2025																			
Análisis de resultados y aplicabilidad	21/04/2025	04/05/2025																			
Elaboración del informe final	05/05/2025	23/05/2025																			

3 Bibliografía inicial

- [1] <https://infoelectoral.interior.gob.es/es/elecciones-celebradas/area-de-descargas/>
- [2] <https://www.ine.es/>
- [3] https://es.wikipedia.org/wiki/Anexo:Presidencias_auton%C3%B3micas_espa%C3%B1olas
- [4] H. K. Tripathy, M. Paliwal, A. Alkhayyat and N. Mishra, "Predicting Election Outcome and Voting Trend with Generative Machine Intelligence," *2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC)*, Guntur, India, 2024, pp. 1-5, doi: 10.1109/ICEC59683.2024.10837206.
- [5] Cedano, D., Picon, M., Ticona, W. (2024). Sentiment Analysis Based on Twitter Comments Using Artificial Intelligence Techniques to Predict Peruvian Presidential Election Results. In: Silhavy, R., Silhavy, P. (eds) Artificial Intelligence Algorithm Design for Systems. CSOC 2024. Lecture Notes in Networks and Systems, vol 1120. Springer, Cham. https://doi.org/10.1007/978-3-031-70518-2_17
- [6] Cameron, M. P., Barrett, P., & Stewardson, B. (2015). Can Social Media Predict Election Results? Evidence From New Zealand. *Journal of Political Marketing*, 15(4), 416–432. <https://doi.org/10.1080/15377857.2014.959690>

4 Seminarios de investigación

Adicionalmente a las actividades anteriormente detalladas asistiré a un mínimo de 4 Seminarios de Posgrado de la EPS.¹

Firma del/de la estudiante:



VºBº del tutor/res

¹ En caso de que, por motivos debidamente justificados, resulte imposible la asistencia a dichos seminarios o cursos de posgrado, el estudiante deberá presentar a la Comisión del Máster MUII una solicitud en la que informe sobre dichos motivos, los justifique documentalmente, y proponga alternativas que sean equiparables a esta actividad. La Comisión evaluará la idoneidad de las alternativas propuestas y emitirá un informe favorable o desfavorable. Una vez aprobadas las actividades alternativas por parte de la Comisión del Máster MUII el estudiante deberá justificar documentalmente la realización de dichas actividades.