

**PREDICTION TASK**

¿Cuál es el tipo de aprendizaje? ¿Si es aprendizaje supervisado, indicar qué se predice? ¿Cuáles son los posibles resultados de la tarea de aprendizaje? ¿Cuándo se observan los resultados de esta tarea?

El tipo de aprendizaje de este proyecto es supervisado, debido a que se quiere predecir y clasificar las noticias falsas. Debido a que se trata de una tarea de clasificación, los resultados van a ser un valor binario, 1 si el modelo considera que la noticia es falsa o 0 si es no lo es. El resultado se obtiene en cuestión de segundos o minutos después de ingresar la noticia al sistema, ya que el modelo solo necesita procesar el texto y hacer la predicción con base en su entrenamiento previo.

**DECISIONS**

¿Cómo se convierten los resultados del modelo en recomendaciones o decisiones procesables para el usuario final?

Los resultados del modelo permiten la identificación temprana de noticias falsas en política, lo que se traduce en advertencias a los usuarios, reducción de la difusión de desinformación en redes sociales, generación de alertas para verificadores de datos, monitoreo de campañas de desinformación a nivel nacional y el desarrollo de herramientas para ciudadanos y tomadores de decisiones. Es necesario mencionar que el modelo aprende de los datos, si suponemos que todos los datos son de Colombia, el modelo puede ser malo para clasificar noticias falsas de otros países. Una manera de llevar el modelo al usuario final sería con una extensión de Google Chrome que haga la extracción de la noticia y la clasifique como una noticia falsa o no mediante una API REST.

**VALUE PROPOSITION**

¿Quién es el beneficiario final? ¿De qué empresa es? ¿Qué problemas específicos se abordan? ¿Qué riesgo puede tener para ese beneficiario el uso de este modelo?

El beneficiario final del modelo incluye ciudadanos, periodistas, medios de comunicación, gobiernos, plataformas digitales y organizaciones de verificación de datos. Se abordan problemas como la manipulación de la opinión pública, la polarización política y la desinformación en elecciones y relaciones internacionales. Sin embargo, existen riesgos como falsos positivos, sesgos en el entrenamiento del modelo, posible uso indebido para censura y la evolución de estrategias para evadir la detección.

**DATA COLLECTION**

No aplica

**DATA SOURCES**

¿Qué fuentes de datos se utilizan? (Mencione tablas de bases de datos internas y externas o métodos API). ¿De dónde se toman los datos? ¿Se pueden utilizar para realizar el objetivo del análisis?

Los datos provienen de noticias recopiladas de medios en línea como 'Público', 'La Marea' y 'El Común', almacenados en bases de datos. Estos datos son adecuados para entrenar el modelo, aunque es clave mitigar sesgos y mantenerlos actualizados para mejorar la detección de desinformación. En este caso aplica el ejemplo de lo que sucede con un sistema de spam de emails, si de alguna manera las personas que escriben las noticias falsas saben que palabras se usan para clasificar la misma se evitara usar estas mismas, lo cual produce que el modelo pierda la capacidad predictiva a lo largo del tiempo. Se pueden utilizar otras fuentes de datos como datasets públicos sobre noticias falsas. Con el grupo encontramos dos de estos en Kaggle, uno sobre noticias falsas en España y otro dataset en inglés sobre noticias falsas en estados unidos.

**IMPACT SIMULATION**

¿Cuáles son los valores de costo/beneficio de las decisiones (in)correctas? ¿Cuáles son los criterios de éxito del modelo para su posterior despliegue? ¿Existen restricciones de equidad?

El costo de una decisión incorrecta incluye la posible censura de medios legítimos (falsos positivos) o la propagación de desinformación política (falsos negativos). El modelo será exitoso si logra alta precisión y recall, es transparente y se adapta a nuevas estrategias de desinformación, es necesario recordar que la capacidad predictiva del modelo se desgasta a medida que pasa el tiempo por lo que se explicó anteriormente. Existen restricciones de equidad relacionadas

**MAKING PREDICTIONS**

¿El uso del modelo es por lotes o en tiempo real? ¿Con qué frecuencia se usa?

El modelo puede operar en tiempo real, analizando noticias conforme se publican, o por lotes, procesando grandes volúmenes de datos periódicamente. En tiempo real se puede usar para monitorear y advertir sobre noticias falsas en redes sociales o plataformas de noticias. En cuanto al procesamiento por lotes se emplea para entrenar el modelo con nuevos datos y estudiar tendencias de desinformación. Por la similitud con el problema de los emails anteriormente y por facilidad se opta por un procesamiento por lotes con una vida útil del modelo de 6 a 12 meses. Lo ideal sería tener un modelo en tiempo real, pero esto


**BUILDING MODELS**

¿Cuántos modelos se necesitan? ¿Cuándo deben actualizarse? ¿De cuánto tiempo se dispone para generar el modelo (incluido el proceso de ingeniería de características y el análisis o evaluación de este)?

Se puede usar un solo modelo de clasificación binaria para empezar, pero se pueden agregar modelos auxiliares según las necesidades. El modelo debe actualizarse cada 3-6 meses. El desarrollo del modelo, incluyendo la recolección de datos, ingeniería de características, entrenamiento y evaluación, puede tomar entre 3 y 4 meses, dependiendo de los recursos disponibles.

**FEATURES**

¿Qué variables/características se utilizan en el modelo? ¿Qué agregaciones o transformaciones se aplican a las fuentes de datos originales – incluir las más importantes--?

<p>con sesgos ideológicos, la representación equitativa de fuentes y la transparencia en los criterios de verificación</p>	<p>puede incurrir en costos mayores. La frecuencia de uso varía puede llegar a ser desde continuo en aplicaciones online hasta mensual o trimestral en análisis retrospectivos.</p>			
	<div><div><div>MONITORING</div><div>No aplica</div></div><div></div></div>			