

1 Normativa asociada al tipo de datos utilizados

La normativa asociada al tipo de datos utilizados en este proyecto se refiere principalmente a la protección de datos personales y la ética en el procesamiento de información. En este caso, se utilizó un conjunto de datos de mensajes de texto (SMS) para desarrollar un modelo de clasificación de spam. A continuación, se explicará cómo se aseguró el cumplimiento de la normativa:

1.1 Tipo de Datos y Normativa Aplicable

Los datos utilizados en este proyecto consisten en mensajes de texto de SMS y están disponibles públicamente en el conjunto de datos SMS Spam Collection Dataset proporcionado por la Universidad de California, Irvine (UCI) en Kaggle. Al trabajar con datos personales, es fundamental cumplir con regulaciones de privacidad de datos, como la Ley Federal de Protección de Datos Personales en Posesión de los Particulares en México.

1.2 Uso de Datos y Cumplimiento Normativo

1. **Consentimiento:** La normativa de privacidad de datos generalmente requiere el consentimiento de los individuos para el uso de sus datos personales. En este caso, como los mensajes de texto son de acceso público y no se recopilaron datos personales sensibles, no es necesario obtener el consentimiento explícito de los usuarios.

2. **Anonimización:** Los datos se procesaron de manera que no se pudiera identificar a ninguna persona específica. Se eliminaron números de teléfono y cualquier información personal identificable.

3. **Eliminación de Datos Sensibles:** Como se trata de mensajes de texto, no se recopilaron datos sensibles, como información financiera o médica.

4. **Cumplimiento de Reglas de Kaggle:** Si bien el conjunto de datos se obtuvo de Kaggle, se respetaron las reglas y políticas de Kaggle con respecto a la recopilación y el uso de datos.

5. **Ética en el Modelado de Datos:** Se aplicó una metodología ética al desarrollar el modelo de clasificación de spam, garantizando que no se introdujera ningún sesgo injusto en el proceso. Se utilizaron técnicas de preprocesamiento de texto estándar y se evitó sesgo en la selección de características o en la evaluación del modelo.

1.3 Cumplimiento con la Normativa de la Industria

Para garantizar que la solución cumple con la normativa establecida en la industria, se siguieron prácticas recomendadas en el procesamiento de datos y modelado de machine learning. Además, se aseguró que el modelo no introdujera sesgos éticos o discriminación injusta en las predicciones.

1.4 Escenarios de Falta Ética

A pesar de tomar medidas para garantizar el cumplimiento normativo y la ética en el proyecto, existen posibles escenarios de falta ética si la herramienta se mal utiliza, como:

1. **Uso Malicioso:** Si alguien utiliza el modelo para enviar spam a gran escala, esto violaría las regulaciones sobre envío de mensajes no solicitados y podría tener consecuencias legales.

2. **Uso Negligente:** Si el modelo se usa para filtrar mensajes importantes que no son spam, como comunicaciones críticas de servicios de salud o seguridad, esto podría causar problemas graves y ser considerado un uso negligente.

1.5 Artículo 10 de la Ley Federal de Protección de Datos Personales

El Artículo 10 de la Ley Federal de Protección de Datos Personales en Posesión de los Particulares establece el concepto de "Fuente de acceso público", que se refiere a bases de datos cuya consulta puede ser realizada por cualquier persona, generalmente sin requerir consentimiento, excepto en casos específicos establecidos por el Reglamento de la Ley. Este artículo es relevante en el contexto de este proyecto, ya que los mensajes de texto en el conjunto de datos son de acceso público y no contienen información confidencial.

En resumen, el proyecto cumple con las normativas de protección de datos al garantizar la anonimización de datos y el cumplimiento de las reglas de Kaggle. Además, se aplicaron prácticas éticas en el desarrollo del modelo para evitar sesgos injustos. Sin embargo, es importante recordar que la falta ética puede ocurrir si la herramienta se usa de manera maliciosa o negligente. Por lo tanto, es esencial que se utilice de manera responsable y ética.

1.6 Referencia

Ley Federal de Protección de Datos Personales en Posesión de los Particulares. (2010). Diario Oficial de la Federación, México, [CÁMARA DE DIPUTADOS DEL H. CONGRESO DE LA UNIÓN], [05-07-2010].