



Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores

Evaluación de Etapa Intermedia ITC

Análisis del contexto y la normatividad.

Rodolfo Sandoval Schipper - A01720253

Se estuvo utilizando un conjunto de datos relacionados a los precios de teléfonos móviles. Este dataset se utilizó para la entrega del Módulo 2 Análisis del Modelo y para la construcción de un modelo utilizando random forest classifier de Scikit Learn (Portafolio Análisis). El objetivo fue analizar la información utilizando la herramienta de Python para hacer un programa y considerar que atributo de la tabla afecta más el rango de precios de los nuevos teléfonos móviles. El archivo de python se puede ubicar en Módulo 2 Análisis y Reporte sobre el desempeño del modelo. Ahora, estaremos investigando la normativa asociada al tipo de datos que se utilizó en este momento de retroalimentación.

Ética y Consideraciones Legales en el Análisis de Datos de Precios Móviles utilizando python se adentra en el mundo del análisis de datos de móviles con el uso de una data set donde se proporciona información sensible de varios celulares el cual se clasifican por el tipo de hardware, y surge la necesidad de abordar cuestiones éticas y legales que son fundamentales para un análisis informado y responsable. En este ensayo, explicaré las consideraciones clave que debemos tener en cuenta al utilizar R y Python para analizar datos de automóviles, garantizando así que nuestros resultados sean sólidos y éticos.

Seguir el estándar de IEEE para desarrollar sistemas consistentes y éticos, el estándar incluye:

IEEE P7003: Este estándar proporciona un marco para abordar la ética en la tecnología de IA y la toma de decisiones autónomas. Fuente: The world 's largest technical professional organization dedicated to advancing technology for the benefit of humanity. (s/f). [ieee.org](https://www.ieee.org/). Recuperado el 29 de agosto de 2023, de <https://www.ieee.org/>

IEEE P7001: Enfocado en la gobernanza de los sistemas de IA y la ética en la toma de decisiones autónomas, este estándar busca garantizar que los sistemas de IA sean diseñados, desplegados y operados de manera ética y responsable. Ejemplo, no modificar los resultados y las predicciones de los rangos en los precios por móvil. Fuente: The world 's largest technical professional organization dedicated to advancing technology for the benefit of humanity. (s/f). [ieee.org](https://www.ieee.org/). Recuperado el 29 de agosto de 2023, de <https://www.ieee.org/>

ISO/IEC 27001: Aunque no es exclusivamente un estándar de ética, se centra en la gestión de la seguridad de la información. Dado que la ética en el análisis de datos también implica la protección de la información y la privacidad, esta norma es relevante para garantizar prácticas seguras y éticas. Fuente: ISO - International Organization for Standardization. (2022). ISO. <https://www.iso.org/home.html>

ISO/IEC 38506: Este estándar se enfoca en la gestión de la tecnología de la información y la ética. Ofrece directrices para la implementación y el uso ético de la tecnología de la información, incluido el análisis de datos. Fuente: ISO - International Organization for Standardization. (2022). ISO. <https://www.iso.org/home.html>

ISO/IEC 2382-37: Proporciona una definición del término "ética de la información" y establece un vocabulario relacionado. Aunque no es específico para el análisis de datos, puede ser

relevante para comprender los conceptos éticos asociados. Fuente: ISO - International Organization for Standardization. (2022). ISO. <https://www.iso.org/home.html>

IEEE P7002: Este estándar se centra en la ética de la recopilación y el uso de datos por parte de los sistemas de inteligencia artificial. Busca proporcionar pautas para una recopilación ética y un manejo adecuado de datos. Fuente: ISO - International Organization for Standardization. (2022). ISO. <https://www.iso.org/home.html>

ISO/IEC 21001: Si bien se centra en sistemas de gestión para organizaciones educativas, también aborda la privacidad y la seguridad de la información, lo cual es relevante para el análisis de datos en el ámbito educativo. Fuente: ISO - International Organization for Standardization. (2022). ISO. <https://www.iso.org/home.html>

La precisión es la base de cualquier análisis de datos significativo. En el contexto de los datos de móviles, incluso un pequeño error podría llevar a conclusiones incorrectas que afecten negativamente a la industria, los consumidores o la seguridad vial. Al seguir el estándar de exactitud, te aseguras de que tus resultados sean confiables y reflejan con precisión la realidad.

En el análisis de datos de móviles, es crucial que los resultados sean entendibles y transparentes. Las decisiones basadas en el análisis pueden afectar la seguridad vial, la industria y los consumidores. Al seguir el principio de explicabilidad, garantizas que tus resultados puedan ser comprendidos por expertos y no expertos por igual, lo que facilita la toma de decisiones informadas y éticas.

Estándares utilizados en el análisis de clasificación de precios de móviles, y en el resto de proyectos involucrando tal análisis y desarrollo.

¿Por qué cumplimos con las normas y la ética?

Las normativas y regulaciones son cruciales para comprender los reglamentos que se deben estar utilizando y que están asociadas al conjunto de información. Aunque el conjunto de datos que presentamos aquí es una simulación, es vital cerciorarse de que los datos reales cumplan con las regulaciones establecidas por la industria móvil y las leyes de protección de datos.

Si los datos contienen información personal identificable, debemos priorizar la privacidad de los individuos. La técnica de anonimización se convierte en nuestro aliado en este caso. Eliminar información como nombres e identificadores únicos garantiza que la privacidad de las personas no se vea comprometida durante el análisis.

En este caso estamos utilizando datos de fuentes externas, es imperativo revisar los términos de la licencia de uso. Cumplir con los términos de licencia, incluyendo la debida atribución a la fuente original, es esencial para respetar los derechos y restricciones establecidos por los proveedores de datos.

El análisis de datos es libre de sesgos y discriminación. Ya que es nuestra responsabilidad evaluar y abordar posibles sesgos en los datos y resultados. Un análisis libre de sesgos es fundamental para tomar decisiones justas e informadas basadas en nuestros resultados.

La ventaja de utilizar Python es su transparencia. Podemos compartir el código que utilizamos para el análisis, lo que permite a otros revisar y probar nuestros métodos y resultados. Esto aumenta la integridad y la confianza del análisis.

Posibles Faltas Éticas: Debemos considerar tanto la malicia como la negligencia en nuestro análisis. La malicia implica manipular los resultados deliberadamente para obtener beneficios injustos, mientras que la negligencia radica en una interpretación incorrecta de los resultados debido a una comprensión insuficiente de los métodos.

Cumplimiento de Normativas y Regulaciones: es crucial asegurarse de que los datos utilizados cumplan con las regulaciones establecidas por la industria móvil y las leyes de protección de datos. Esto demuestra un compromiso con el cumplimiento de las normativas legales y éticas relacionadas con la recopilación y el uso de datos.

Privacidad de los Individuos: Se resalta la importancia de preservar la privacidad de las personas si los datos contienen información personal identificable. La técnica de anonimización se menciona como una medida para garantizar que la privacidad de los individuos no se vea comprometida durante el análisis. En el caso de nuestro conjunto estos datos no corresponden al llenado.

Cumplimiento de Términos de Licencia: Se enfatiza la necesidad de revisar y cumplir con los términos de licencia de uso de datos de fuentes externas, incluyendo la atribución adecuada a la fuente original. Esto asegura el respeto por los derechos y restricciones establecidos por los proveedores de datos. Para nuestro análisis, utilizamos la referencia del conjunto que se encuentra en Fuente: <https://www.kaggle.com/datasets/iabhishekofficial/mobile-price-classification?select=train.csv>

Análisis Libre de Sesgos: Se reconoce la responsabilidad de evaluar y abordar posibles sesgos en los datos y resultados. Un análisis libre de sesgos es considerado fundamental para tomar decisiones justas e informadas basadas en los resultados.

Prevención de Faltas Éticas: Se advierte sobre posibles faltas éticas, como la malicia (manipulación deliberada de resultados) cómo precios incorrectos e información falsa en los resultados del modelo y la negligencia (interpretación incorrecta debido a una comprensión insuficiente de los métodos). Este reconocimiento anticipado de posibles faltas éticas muestra una actitud proactiva hacia la prevención de comportamientos inapropiados.

Caso de falta ética

A continuación se especifican cómo se pueden manipular los datos de manera inapropiada:

Manipulación de Datos de Entrada: Los datos de entrada utilizados en el análisis pueden ser manipulados antes de ser procesados por Python. Esto podría incluir la eliminación o la modificación de registros para influir en los resultados en los precios.

La elección de modelos de análisis puede influir en los resultados. Se podría seleccionar un modelo específico que presente resultados más favorables para una determinada agenda, incluso si ese modelo no es el más apropiado desde un punto de vista estadístico, en nuestro caso el modelo de random forest es perfecto para predecir con de acuerdo a la clasificación en los datos. En muchas herramientas de análisis, también existen parámetros que pueden ajustarse para obtener resultados deseados. Modificar estos parámetros de manera inapropiada podría llevar a resultados sesgados este proceso se lleva a cabo en el análisis de modelo. El concepto de la interpretación engañosa involucra que los resultados del análisis puedan ser presentados de manera engañosa o sesgada, destacando ciertos hallazgos y omitiendo otros para influir en las decisiones de compra o venta de automóviles. Debemos de tomar en cuenta que al recopilar datos para el análisis, se pueden seleccionar fuentes de datos sesgadas que refuercen una narrativa particular o que favorezcan ciertos resultados.

Con de acuerdo a las herramientas también podemos inferir falta de ética en la transparencia en el Código: Si el código utilizado en Python no se comparte de manera transparente o si se oculta la lógica detrás del análisis, se puede dificultar la revisión y la detección de posibles manipulaciones.

Para prevenir estas manipulaciones con falta de ética, es importante establecer buenas prácticas en el desarrollo y el uso de herramientas de análisis de datos:

Primero necesitamos fomentar la transparencia al compartir el código y la lógica detrás del análisis para que otros puedan revisarlo y verificar. Es necesario validar y limpiar datos para realizar una limpieza con un pequeño proceso de ETL y validación rigurosa de los datos de entrada para asegurarse de que sean precisos y confiables. Otro concepto importante es utilizar modelos adecuados y seleccionar modelos de análisis basados en criterios estadísticos y no en objetivos sesgados.

La revisión independiente también se considera cómo un factor importante para permitir que se pueda revisar el análisis para identificar posibles sesgos o manipulaciones. Debemos asegurarnos de cumplir con las leyes y regulaciones de privacidad de datos y protección del consumidor. Y finalmente podemos promover una cultura ética en el análisis de datos, enfatizando la importancia de la integridad y la honestidad en el proceso.

Conclusiones:

Al utilizar Python para el análisis de datos de móviles, entramos en un terreno que va más allá de la mera técnica. La responsabilidad ética y legal se convierte en un pilar central de nuestro trabajo. Cumplir con las regulaciones, asegurar la privacidad, abordar los sesgos y mantener la transparencia son imperativos para generar resultados sólidos y éticos. Nuestro análisis no solo

debe ser acertado, sino también ético, con la responsabilidad de tomar decisiones informadas y justas en base a nuestros resultados. En última instancia, la combinación de Python y la consideración ética forman un conjunto que impulsa análisis donde existe la confiabilidad y responsabilidad. En el proceso de análisis de datos de móviles utilizando herramientas como Python, se destaca la importancia primordial de la ética y el cumplimiento de normativas y regulaciones. Esto implica comprometerse con la precisión de los resultados y la transparencia en todo el proceso. Además, se deben cumplir las normativas legales y garantizar la privacidad de los individuos cuando los datos contienen información personal. La revisión independiente por expertos y la prevención de faltas éticas, como la manipulación deliberada de datos, son prácticas cruciales. En última instancia, el análisis de datos de móviles no solo debe ser técnico sino también ético, con la responsabilidad de tomar decisiones informadas y justas basadas en resultados confiables y responsables.