

Curso de Data Science

Modalidad: individual o en grupo (de hasta 4)
Todas las entregas por el campus NO POR MAIL

Trabajo Integrador Final



Objetivos

1. Aplicar técnicas de Data Science: Utilizar y aplicar las técnicas de Data Science aprendidas en el curso, como limpieza y exploración de datos, implementación y evaluación de modelos de machine learning, visualización y análisis de datos, y control de versiones con Git/GitHub.
2. Resolver una problemática real: Seleccionar y abordar una problemática de negocios o industrial utilizando enfoques basados en datos, aplicando técnicas de Data Science para obtener insights o soluciones relevantes y aplicables.
3. Comunicar resultados: Presentar y comunicar de manera efectiva los resultados y hallazgos obtenidos en el proyecto, utilizando habilidades de comunicación oral y escrita, y preparando visualizaciones claras y comprensibles.



Consigna

Metodología:

1. Exploración y limpieza de datos: Familiarizarse con el conjunto de datos, comprender su estructura y contenido, identificar variables relevantes y eliminar inconsistencias o valores faltantes.
2. Selección de características: Identificar las variables del conjunto de datos que serán utilizadas para entrenar el modelo de machine learning. La selección de características adecuadas es crucial para el éxito del modelo.
3. Implementación y entrenamiento del modelo: Seleccionar un algoritmo de machine learning apropiado, entrenar el modelo utilizando los datos preprocesados y evaluar su rendimiento.

4. Evaluación del modelo: Evaluar el rendimiento del modelo utilizando métricas adecuadas, como precisión, sensibilidad, especificidad o error cuadrático medio.
5. Comunicación de resultados: Presentar los resultados y hallazgos del proyecto de manera clara y concisa, utilizando visualizaciones efectivas y una argumentación sólida de las decisiones tomadas.

Fecha límite de entrega: según cronograma de la comisión

Criterios de evaluación

- **Exploración y limpieza de datos (20%)**
 - Comprensión del conjunto de datos (5%)
Demuestra una comprensión clara de la estructura, contenido y características del conjunto de datos. Identifica correctamente las variables relevantes para el problema que se está abordando.
 - Análisis descriptivo (5%)
Realiza un análisis descriptivo exhaustivo de las variables del conjunto de datos. Utiliza visualizaciones efectivas para resumir las características y distribuciones de las variables.
 - Limpieza de datos (10%)
Detecta y elimina inconsistencias, valores faltantes y errores en el conjunto de datos.
Aplica técnicas de preprocesamiento de datos adecuadas para preparar el conjunto de datos para el análisis y el modelado.
- **Selección de características (20%)**
Aplica técnicas de selección de características adecuadas para identificar las variables más relevantes para el problema.
Justifica la selección de características realizadas.
- **Implementación y entrenamiento del modelo (30%)**
 - Elección del algoritmo de machine learning (15%)
Selecciona un algoritmo de machine learning apropiado para el tipo de problema y las características del conjunto de datos. Justifica la elección del algoritmo seleccionado.
 - Preparación de datos para el modelado (15%)
Divide los datos en conjuntos de entrenamiento, validación y prueba.
Aplica técnicas de escalado o normalización de datos si es necesario.
- **Entrenamiento y ajuste del modelo (10%)**

- Entrena el modelo de machine learning utilizando el conjunto de datos de entrenamiento.
- Ajusta los parámetros del modelo para optimizar su rendimiento.
- **Evaluación del modelo (10%)**
 - Métricas de evaluación (5%)

Selecciona métricas de evaluación adecuadas para el tipo de problema y el objetivo del modelo.

Interpreta los resultados de las métricas de evaluación para evaluar el rendimiento del modelo.
 - Análisis de errores (5%)

Analiza los errores del modelo para identificar posibles áreas de mejora. Propone estrategias para mejorar el rendimiento del modelo.
- **Comunicación de resultados (10%)**

Documenta el proceso de desarrollo del proyecto de manera detallada y organizada. Explica claramente la metodología utilizada, los resultados obtenidos y las conclusiones alcanzadas.

Utiliza un lenguaje claro y conciso, y sigue las normas de estilo académico.



Bibliografía utilizada y sugerida

- Material del curso
- Recursos de consulta