

Clasificación y detección de contaminantes usando IA

Ingeniería | Ingeniería ambiental

Descripción del Curso

Unidad 8: Interpretación de resultados y comunicación para mitigación de contaminantes forma parte del curso Ingeniería Ambiental y está dirigida a estudiantes de dicha disciplina, sin restricción de edad (entre 17 años y más). Este módulo se centra en desarrollar competencias para interpretar resultados de modelos de IA aplicados a contaminantes y comunicar hallazgos a audiencias técnicas y no técnicas, proponiendo recomendaciones de mitigación basadas en evidencia para apoyar la toma de decisiones ambientales.

El enfoque se orienta a la interpretación de resultados, la evaluación de incertidumbres y la traducción de información compleja en conclusiones accionables. Se promueve la claridad comunicativa, la adaptación del lenguaje y el formato a diferentes públicos, así como la consideración de impactos ambientales y sociales. Se emplearán modelos de IA para sustentar recomendaciones de mitigación y para apoyar procesos de toma de decisiones en contextos públicos y privados.

La unidad fomenta el desarrollo de habilidades de pensamiento crítico y de presentación efectiva, mediante casos prácticos, lectura de resultados de modelos y elaboración de recomendaciones viables que contemplen aspectos técnicos, económicos y sociales. Asimismo, se enfatizan principios de ética, transparencia y responsabilidad en la interpretación de modelos y la comunicación de hallazgos a audiencias diversas, incluyendo comunidades afectadas y tomadores de decisiones.

Competencias

- Interpretar resultados de modelos de IA y traducirlos en conclusiones accionables para la mitigación de contaminantes.
- Comunicar hallazgos de IA de forma clara y adaptada a audiencias técnicas y no técnicas.
- Proponer recomendaciones de mitigación basadas en evidencia de IA, considerando impactos ambientales y sociales.
- Evaluar incertidumbres y limitaciones de los modelos de IA y comunicar esas limitaciones a las partes interesadas.
- Aplicar principios de comunicación científica y toma de decisiones basada en evidencia para influir en políticas y prácticas ambientales.
- Desarrollar habilidades de trabajo en equipo interdisciplinario y presentar resultados de forma oral y escrita.
- Utilizar herramientas de IA de manera responsable, transparente y auditable en el contexto de la mitigación de contaminantes.
- Integrar consideraciones de equidad, costo-efectividad y antecedentes sociales en las propuestas de mitigación.

Requerimientos

- Conocimientos previos: fundamentos de estadística, química ambiental y fundamentos de modelado y análisis de datos.
- Software y herramientas: manejo de herramientas de análisis y visualización (por ejemplo, Excel, Python o R) y plataformas de IA con capacidades de interpretación; entornos de cuadernos de lenguaje (Jupyter, etc.).
- Recursos técnicos: computadora con acceso a internet y capacidad para ejecutar software de análisis; cuenta institucional para plataformas educativas si aplica.
- Actividades de aprendizaje: participación en foros, trabajos en equipo, presentaciones orales y escritas, y desarrollo de un informe de hallazgos con recomendaciones de mitigación basadas en IA.
- Idioma y normativas: impartición en español; cumplimiento de normas de citación y de ética en la divulgación de resultados.
- Carga de trabajo: la unidad combina clases teóricas, prácticas y un proyecto de mitigación; se espera commit

Unidades del Curso

Unidad 1: Unidad 1: Conceptos clave de clasificación y detección de contaminantes en IA

Objetivos de Aprendizaje

- Identificar los tipos de datos usados en IA ambiental (sensores, imágenes, datos satelitales, series temporales, datos de calidad ambiental).
- Describir las características relevantes que influyen en el rendimiento de modelos para contaminantes (ruido, escalas, correlaciones, variables confusoras).
- Explicar las métricas de rendimiento (precisión, recall, F1, ROC-AUC) y su interpretación en tareas de clasificación y detección.

Contenidos Temáticos

1. Tipos de datos en IA ambiental — descripción breve de sensores, imágenes, series temporales y datos tabulares.
2. Características relevantes y extracción de características — cómo identificar variables útiles y gestionar el ruido.
3. Clasificación vs. detección de contaminantes — diferencias conceptuales y ejemplos prácticos.
4. Métricas de rendimiento y evaluación — interpretación de métricas y su relación con objetivos ambientales.

Actividades

- **Actividad 1: Exploración de un conjunto de datos ambiental** — Los estudiantes identificarán tipos de datos, variables clave y posibles valores atípicos en un conjunto de datos real o simulado; discutirán impacto de la calidad de datos en el rendimiento de modelos y presentarán hallazgos en equipo.
- **Actividad 2: Cálculo de métricas básicas** — Se utilizará una matriz de confusión para calcular precisión, recall, F1 y ROC-AUC; se discutirán interpretaciones en problemas de contaminación ambiental.

- **Actividad 3: Debate sobre sesgo y robustez** — Discusión estructurada sobre sesgos de datos, representatividad y efectos en la toma de decisiones ambientales basadas en IA.

Evaluación

La evaluación se alinea con el objetivo general de la unidad y comprende:

- Examen corto de conceptos clave (20%).
- Actividad práctica de métricas y análisis de datos (40%).
- Participación y reflexión en el debate sobre sesgo y ética (20%).
- Informe breve de conceptos clave y ejemplos prácticos (20%).

Unidad 2: Unidad 2: Flujo de trabajo de IA para clasificación y detección de contaminantes

Objetivos de Aprendizaje

- Describir las fases del flujo de trabajo de IA: definición del problema, recopilación de datos, preprocesamiento, modelado, validación e implementación.
- Identificar entregables y hitos típicos de proyectos de IA ambiental (plan de datos, protocolo de experimentación, informe de resultados).
- Señalar consideraciones de gobernanza de datos, ética, trazabilidad y reproducibilidad a lo largo del ciclo de vida del proyecto.

Contenidos Temáticos

1. Definición del problema y requerimientos — cómo traducir necesidades ambientales en objetivos de IA.
2. Recopilación y gestión de datos — fuentes, calidad, gobernanza y trazabilidad.
3. Pipeline de IA: preprocesamiento, modelado y validación — etapas y buenas prácticas.
4. Despliegue responsable y gobernanza de datos — consideraciones éticas, regulaciones y mantenimiento.

Actividades

- **Actividad 1: Mapa del flujo de trabajo** — En equipos, diseñarán un diagrama de flujo para un proyecto de clasificación de contaminantes, identificando entradas, procesos y entregables.
- **Actividad 2: Escenarios de recopilación de datos** — Análisis de fuentes de datos y definiciones de criterios de calidad y gobernanza; elaboración de una lista de verificación.
- **Actividad 3: Discusión de despliegue y trazabilidad** — Debate sobre requisitos de reproducibilidad, versión de datos y seguimiento de experimentos.

Evaluación

Evaluación centrada en la capacidad de describir el flujo de trabajo y proponer entregables y prácticas de gobernanza:

- Ejercicio de diseño de flujo de trabajo y lista de entregables (30%).
- Análisis crítico de gobernanza de datos y ética en IA ambiental (20%).
- Participación en debates y calidad de la discusión (20%).
- Informe de caso de estudio breve (30%).

Unidad 3: Unidad 3: Análisis de conjuntos de datos ambientales y preprocesamiento

Objetivos de Aprendizaje

- Realizar análisis exploratorio de datos (EDA) en conjuntos ambientales y detectar anomalías.
- Aplicar técnicas de ingeniería de características y selección de características para reducir dimensionalidad y mejorar la señal.
- Implementar preprocesamiento adecuado (limpieza, manejo de valores faltantes, normalización/escala) y evaluar su impacto en el rendimiento.

Contenidos Temáticos

1. Exploración y limpieza de datos ambientales — interpretación de estadísticas y visualización básica.
2. Ingeniería y selección de características — métodos simples de filtrado y envoltura.
3. Preprocesamiento para IA — manejo de valores faltantes, normalización, estandarización y reducción de dimensionalidad.
4. Calidad de datos y sesgo — control de variabilidad y verificación de supuestos.

Actividades

- **Actividad 1: Análisis exploratorio guiado** — Realizar EDA en un dataset ambiental, generar visualizaciones y reportar hallazgos clave y posibles problemas de calidad.
- **Actividad 2: Ingeniería y selección de características** — Aplicar métodos simples de selección de características y justificar las elecciones para un problema de contaminantes.
- **Actividad 3: Preprocesamiento y su impacto** — Implementar limpieza y normalización, comparar rendimiento de un modelo con y sin preprocesamiento.

Evaluación

La evaluación se centra en la capacidad de analizar datos y preparar entradas de calidad para modelos:

- Informe de análisis exploratorio y decisiones de preprocesamiento (35%).
- Documento de ingeniería/selección de características y justificación (25%).
- Ejercicio práctico de preprocesamiento con evaluación de impacto (20%).
- Participación y presentaciones breves (20%).

Unidad 4: Unidad 4: Algoritmos de aprendizaje automático supervisado básicos

Objetivos de Aprendizaje

- Implementar modelos supervisados simples para clasificación/detección de contaminantes en conjuntos de datos preparados.
- Comparar y justificar la elección de modelos en función de la complejidad de la frontera, tamaño de datos y presencia de ruido.
- Interpretar resultados para inferir conclusiones ambientales y posibles limitaciones de los modelos.

Contenidos Temáticos

1. Soporte vectorial (SVM) — fundamentos, kernel y cuando usarlo.
2. Random Forest — ensembles, manejo de datos ruidosos y interpretabilidad.
3. k-Nearest Neighbors (k-NN) — simplicidad, sensibilidad a escala y tamaño de datos.
4. Selección de modelos y consideraciones prácticas — cuando priorizar cada algoritmo.

Actividades

- **Actividad 1: Entrenamiento de modelos básicos** — Construcción e entrenamiento de SVM, Random Forest y k-NN en un conjunto de datos de contaminantes y comparación de rendimientos.
- **Actividad 2: Evaluación de supuestos** — Analizar cómo la escala de características y la dimensionalidad afectan el rendimiento de cada algoritmo.
- **Actividad 3: Informe de justificación de modelo** — Elegir un modelo para un escenario ambiental específico y justificar la elección con criterios técnicos y prácticos.

Evaluación

La evaluación se centrará en la aplicación y justificación de modelos supervisados:

- Ejercicio práctico de implementación y comparación de modelos (40%).
- Informe de justificación de la selección de modelo (30%).
- Cuestionario corto de conceptos clave (15%).
- Participación en actividades de clase (15%).

Unidad 5: Evaluación de modelos de clasificación y detección de contaminantes

Objetivos de Aprendizaje

- Calcular y comparar métricas de rendimiento para distintos modelos y umbrales.
- Interpretar curvas ROC y la curva de precisión-recall para evaluar trade-offs.
- Traducir métricas técnicas a conclusiones y recomendaciones ambientales.

Contenidos Temáticos

1. Precisión, recall y F1 — interpretación en contextos de contaminación.
2. ROC-AUC y curvas de rendimiento — análisis de umbrales.
3. Interpretabilidad y límites de las métricas — sesgos y escenarios desequilibrados.
4. Comunicación de resultados para audiencias técnicas y no técnicas.

Actividades

- **Actividad 1: Cálculo y análisis de métricas** — Calcular métricas en diferentes modelos y umbrales; discutir implicaciones para decisiones ambientales.
- **Actividad 2: Interpretación de curvas** — Interpretar ROC y precisión-recall en escenarios con desequilibrio de clases (p. ej., detección de contaminantes raros).
- **Actividad 3: Informe de interpretación** — Redactar un informe que comunique hallazgos a una audiencia técnica y a una audiencia no técnica.

Evaluación

La evaluación se centrará en la capacidad de evaluar y comunicar resultados de modelos:

- Ejercicio práctico de métricas y curvas (40%).
- Proyecto de interpretación y reporte para audiencias distintas (35%).
- Cuestionario de conceptos y definiciones (15%).
- Participación y discusión (10%).

Unidad 6: Unidad 6: Diseño y ejecución de un experimento reproducible

Objetivos de Aprendizaje

- Diseñar un experimento con control de variables, conjuntos de entrenamiento/prueba y criterios de éxito.
- Registrar de forma clara y reproducible datos, preprocesamiento y configuración de modelos (hiperparámetros).
- Garantizar trazabilidad mediante gestión de versiones y documentación de resultados.

Contenidos Temáticos

1. Planeación de experimentos — definición de hipótesis, métricas y criterios de éxito.
2. Gestión de datos y reproducibilidad — versiones, registros y notebooks reutilizables.
3. Control de hiperparámetros y configuración experimental — técnicas básicas.
4. Documentación y reporte de resultados — formato y trazabilidad.

Actividades

- **Actividad 1: Protocolo de experimento** — Crear un protocolo reproducible para un experimento de clasificación de contaminantes, asociando datos, preprocesamiento y parámetros.

- **Actividad 2: Notebook reproducible** — Construir un notebook con código claro, dependencias y pasos de ejecución, para replicar resultados.
- **Actividad 3: Registro de resultados** — Documentar resultados, gráficos y conclusiones con trazabilidad de versiones de datos y código.

Evaluación

Evaluación basada en la calidad de la reproducibilidad y la claridad de la documentación:

- Protocolo y plan de experimento (30%).
- Notebook reproducible y código bien documentado (40%).
- Informe de resultados con trazabilidad (20%).
- Participación en revisión por pares (10%).

Unidad 7: Validación y pruebas de robustez ante condiciones ambientales variables

Objetivos de Aprendizaje

- Aplicar validación cruzada y evaluación en conjuntos separados para estimar la generalización.
- Diseñar pruebas de robustez ante ruido, cambios de sensores o variabilidad estacional.
- Interpretar resultados de validación para decidir sobre mejoras o despliegue seguro.

Contenidos Temáticos

1. Validación externa y generalización — evaluación en datos no vistos.
2. Robustez ante ruido y variabilidad de sensores — estrategias de mitigación.
3. Pruebas ante escenarios ambientales variables — escenarios de estrés y simulaciones.
4. Estimación de incertidumbre y estabilidad — técnicas básicas.

Actividades

- **Actividad 1: Validación y generalización** — Realizar validación cruzada y evaluar desempeño en un conjunto de datos diferente al de entrenamiento; discutir resultados y límites.
- **Actividad 2: Pruebas de robustez** — Introducir perturbaciones (ruido, cambios de sensor) y medir el impacto en las predicciones.
- **Actividad 3: Informe de escenarios** — Preparar un informe que describa pruebas de robustez y recomendaciones para mantenimiento del modelo.

Evaluación

La evaluación valorará la capacidad de garantizar generalización y robustez:

- Ejercicio de validación y reporte (40%).
- Pruebas de robustez y análisis (30%).
- Participación y discusión (15%).
- Informe final de escenarios y recomendaciones (15%).

Unidad 8: Interpretación de resultados y comunicación para mitigación de contaminantes

Objetivos de Aprendizaje

- Interpretar resultados de modelos y traducirlos en conclusiones accionables.
- Comunicar hallazgos de IA de forma clara y adaptada a diferentes audiencias (técnica y no técnica).
- Proponer recomendaciones de mitigación de contaminantes basadas en evidencia de IA y considerar impactos ambientales y sociales.

Contenidos Temáticos

1. Interpretabilidad y explicabilidad — enfoques y límites.
2. Comunicación de resultados a audiencias técnicas y no técnicas — formatos y lenguaje.
3. Recomendaciones de mitigación basadas en IA — diseño de acciones y evaluación de impacto.
4. Riesgos, gobernanza y toma de decisiones ambientales — consideraciones éticas y regulatorias.

Actividades

- **Actividad 1: Presentación ejecutiva** — Preparar una presentación para una audiencia no técnica que resuma hallazgos, límites y recomendaciones.
- **Actividad 2: Informe técnico para audiencia técnica** — Elaborar un informe detallado con metodología, resultados y análisis de incertidumbre dirigido a especialistas.
- **Actividad 3: Caso de mitigación** — Proponer un conjunto de medidas para reducir contaminantes basadas en los modelos aprendidos y analizar su impacto potencial.

Evaluación

La evaluación enfatiza la capacidad de interpretación y comunicación, así como la viabilidad de las recomendaciones:

- Presentación ejecutiva (25%).
- Informe técnico (35%).
- Caso de mitigación y análisis de impacto (25%).
- Participación y comentarios en clase (15%).