

Base de datos: EcoSorter

Descripción del Sistema Clasificador de Residuos

Propósito del Sistema

El **Sistema Clasificador de Residuos** es una solución tecnológica diseñada para optimizar el proceso de reciclaje en un centro especializado mediante la automatización de la identificación, clasificación y procesamiento de materiales reciclables. Utilizando técnicas de machine learning, el sistema analizará datos provenientes de sensores, imágenes y registros manuales para clasificar los residuos de manera eficiente, reducir errores humanos, mejorar la calidad del reciclaje y generar predicciones que optimicen las operaciones del centro.

Funcionalidades Principales

1. Registro de Recolecciones

- El sistema permitirá a los usuarios (ciudadanos, empresas o empleados) registrar las recolecciones de residuos a través de la tabla Recolecciones.
- Datos como id_usuario, fecha_recoleccion y estado se almacenarán para rastrear cada entrega.

2. Clasificación Automática de Materiales

- Utilizando un modelo de machine learning entrenado (almacenado en Modelos_ML), el sistema analizará los residuos recibidos para identificar su tipo (id_material) y calidad (calidad).
- **Entradas:**
 - Imágenes de los residuos capturadas por cámaras.
 - Datos de sensores (ej. densidad, color) registrados en Procesos_Reciclaje bajo datos_sensor.
- **Salida:**
 - Predicciones almacenadas en Predicciones (resultado_prediccion, confianza), como "Material **identificado:** plástico PET, Confianza: 0.95".
- Los resultados se vinculan a Detalle_Recoleccion para especificar qué materiales componen cada recolección.

3. Asignación de Categorías y Valor Económico

- Una vez clasificados los materiales, el sistema consultará la tabla Materiales (relacionada con Categorías) para asignar la categoría correspondiente (ej. "plástico", "vidrio") y determinar su valor_reciclaje.
- Esto permitirá calcular el valor económico total de la recolección y ofrecer incentivos a los usuarios si el centro lo implementa.

4. Gestión de Procesos de Reciclaje

- El sistema registrará los pasos de procesamiento (clasificación, trituración, lavado) en Procesos_Reciclaje, asociándolos a cada id_recoleccion.
- Usará las predicciones del modelo para decidir el mejor proceso según la calidad y tipo de material (ej. "lavado intensivo" para plásticos sucios).
- Los datos_sensor generados durante el procesamiento (temperatura, humedad) se almacenarán para retroalimentar el modelo de machine learning.

5. Entrenamiento y Mejora del Modelo de Machine Learning

- El sistema recopilará datos históricos de Detalle_Recoleccion, Procesos_Reciclaje y Predicciones para reentrenar periódicamente los modelos en Modelos_ML.
- Las métricas de desempeño (precisión, recall) se actualizarán en metricas para evaluar la efectividad del clasificador.
- Con el tiempo, el sistema aprenderá a identificar materiales más complejos o a predecir problemas en el procesamiento (ej. fallos en máquinas).

6. Reportes y Análisis

- Generará reportes para los administradores del centro, como:
 - Cantidad de materiales reciclados por categoría (Categorías).
 - Eficiencia de las recolecciones (Recolecciones y Detalle_Recoleccion).
 - Tasa de acierto del clasificador (Predicciones).
- Estos reportes ayudarán a tomar decisiones estratégicas, como ajustar incentivos o mejorar equipos.

Flujo de Trabajo Ejemplificado

1. Un usuario entrega 10 kg de residuos (Recolecciones).

2. Cámaras y sensores capturan imágenes y datos del material (Procesos_Reciclaje).
3. El modelo de machine learning analiza los datos y predice: "6 kg de plástico PET, 4 kg de vidrio" (Predicciones).
4. Los resultados se registran en Detalle_Recoleccion, vinculando los materiales a Materiales y Categorías.
5. El sistema calcula el valor económico basado en valor_reciclaje y decide el proceso óptimo (ej. trituración para plástico).
6. Los datos generados retroalimentan el modelo para mejorar futuras clasificaciones.

Beneficios

- **Eficiencia:** Automatiza la clasificación, reduciendo el tiempo y el esfuerzo humano.
- **Precisión:** Mejora la identificación de materiales con machine learning, minimizando errores.
- **Escalabilidad:** La base de datos normalizada y los modelos entrenados permiten manejar grandes volúmenes de residuos.
- **Sostenibilidad:** Optimiza el reciclaje, aumentando la recuperación de materiales valiosos.

Integración con la Base de Datos

- **Usuarios:** Identifica quién entrega los residuos.
- **Categorías y Materiales:** Proporciona el catálogo de materiales y su valor.
- **Recolecciones y Detalle_Recoleccion:** Registra y detalla las entregas.
- **Procesos_Reciclaje:** Gestiona el procesamiento y genera datos para el modelo.
- **Modelos_ML y Predicciones:** Almacena los modelos y sus resultados, permitiendo aprendizaje continuo.

Proceso de normalización:

Paso 1: Primera Forma Normal (1NF)

Regla:

- Todos los datos deben ser atómicos (sin valores multivaluados o listas en una celda).
- Cada tabla debe tener una clave primaria única.
- No debe haber filas duplicadas.

Revisión:

1. Usuarios:

- Todos los campos (id_usuario, nombre, tipo_usuario, etc.) son atómicos.
- id_usuario es la clave primaria única.

2. Materiales:

- Campos como id_material, nombre, descripcion son atómicos.
- id_material es la clave primaria.

3. Recolecciones:

- ubicacion podría ser un problema si contiene datos no atómicos (ej. "Calle 123, Ciudad"), pero lo asumimos como un solo valor por simplicidad. Si se necesitara desglosar (calle, ciudad, país), se crearía una tabla adicional.
- id_recoleccion es la clave primaria.

4. Detalle_Recoleccion:

- Todos los campos son atómicos.
- id_detalle es la clave primaria.

5. Procesos_Reciclaje, Reportes, Modelo_ML, Predicciones:

- Similarmente, todos tienen campos atómicos y claves primarias únicas (id_proceso, id_reporte, id_modelo, id_prediccion).
- Cumplen 1NF.

Resultado: Todas las tablas ya están en 1NF porque no hay valores multivaluados y tienen claves primarias.

Paso 2: Segunda Forma Normal (2NF)

Regla:

- Todos los atributos no clave deben depender completamente de la clave primaria (no debe haber dependencias parciales).

Revisión:

1. Usuarios:

- **Clave primaria:** id_usuario.

- Atributos como nombre, email, telefono dependen completamente de id_usuario.

2. Materiales:

- **Clave primaria:** id_material.
- nombre, descripcion, categoria, valor_por_kg dependen de id_material.

3. Recolecciones:

- **Clave primaria:** id_recoleccion.
- id_usuario, fecha_recoleccion, ubicacion, estado, peso_total dependen de id_recoleccion.
- No hay dependencias parciales (ningún atributo depende solo de una parte de la clave, ya que es simple).

4. Detalle_Recoleccion:

- **Clave primaria:** id_detalle.
- id_recoleccion, id_material, cantidad, observaciones dependen de id_detalle.

5. Procesos_Reciclaje:

- **Clave primaria:** id_proceso.
- id_recoleccion, id_material, tipo_proceso, etc., dependen de id_proceso.

6. Reportes, Modelo_ML, Predicciones:

- Similarmente, todos los atributos dependen de sus respectivas claves primarias (id_reporte, id_modelo, id_prediccion).

Resultado: Todas las tablas están en 2NF porque no hay claves compuestas con dependencias parciales, y los atributos no clave dependen completamente de las claves primarias.

Paso 3: Tercera Forma Normal (3NF)

Regla:

- No debe haber dependencias transitivas: los atributos no clave no deben depender de otros atributos no clave (solo de la clave primaria).

Revisión:

1. Usuarios:

- nombre, email, telefono, direccion, fecha_registro dependen directamente de id_usuario.
- No hay dependencias transitivas (ej. email no depende de nombre).

2. Materiales:

- nombre, descripcion, categoria, valor_por_kg dependen de id_material.
- No hay dependencias transitivas (ej. valor_por_kg no depende de categoria).

3. Recolecciones:

- **Atributos:** id_usuario, fecha_recoleccion, ubicacion, estado, peso_total.
- id_usuario es una clave foránea, pero no introduce dependencias transitivas porque es parte del diseño relacional.

4. Detalle_Recoleccion:

- id_recoleccion, id_material, cantidad, observaciones dependen de id_detalle.
- No hay dependencias transitivas.

5. Procesos_Reciclaje:

- id_recoleccion, id_material, tipo_proceso, fecha_inicio, fecha_fin, costo, resultado dependen de id_proceso.
- No hay dependencias transitivas.

6. Reportes:

- id_usuario, fecha_generacion, tipo_reporte, contenido dependen de id_reporte.
- No hay dependencias transitivas.

7. Modelo_ML:

- nombre, tipo, fecha_entrenamiento, parametros, ruta_archivo dependen de id_modelo.
- No hay dependencias transitivas.

8. Predicciones:

- id_modelo, id_recoleccion, fecha_prediccion, entrada, resultado, precision dependen de id_prediccion.

- No hay dependencias transitivas.

1. Usuarios

- **Propósito:** Almacena la información de las personas o entidades que interactúan con el centro de reciclaje (residentes que entregan materiales, empleados que gestionan procesos, empresas asociadas).
- **Funcionamiento en el sistema:**
 - Cuando un usuario se registra, se crea una entrada con su id_usuario, nombre, email, etc.
 - Este id_usuario se usa en Recolecciones para rastrear quién entrega o solicita una recolección, y en Reportes para identificar quién genera o solicita un informe.
 - Ejemplo: Un residente solicita una recolección su id_usuario vincula esa solicitud a su perfil.
- **Relación con ML:** Podría usarse para analizar patrones de comportamiento (ej. cuántos kg recicla un usuario promedio) y alimentar modelos predictivos.

2. Materiales

- **Propósito:** Catalogo de los tipos de materiales reciclables que el centro acepta y procesa (plástico, vidrio, papel, etc.).
- **Funcionamiento en el sistema:**
 - Cada material tiene un id_material único y detalles como nombre y valor_por_kg.
 - Se referencia en Detalle_Recoleccion para especificar qué materiales se recolectaron y en Procesos_Reciclaje para indicar qué se procesó.
 - Ejemplo: Si se recolecta "Plástico PET" (id_material = 1), este ID aparece en las tablas relacionadas.
- **Relación con ML:** Un modelo de clasificación podría usar imágenes o datos de sensores para identificar automáticamente el id_material y registrar su entrada.

3. Recolecciones

- **Propósito:** Registra las operaciones de recolección de materiales, ya sea por usuarios o por el centro.
- **Funcionamiento en el sistema:**

- Una recolección se crea con un `id_recoleccion`, asociada a un `id_usuario`, con datos como `fecha_recoleccion` y `ubicacion`.
- Su `peso_total` resume la cantidad recolectada (puede calcularse desde `Detalle_Recoleccion` o ingresarse manualmente).
- Sirve como "puente" entre el usuario y los detalles de lo recolectado.
- Ejemplo: Un camión recoge 50 kg de basura el 20/03/2025; se registra aquí y se detalla en `Detalle_Recoleccion`.
- **Relación con ML:** Un modelo predictivo podría estimar el `peso_total` futuro o predecir la demanda de recolecciones en ciertas ubicaciones, guardando resultados en `Predicciones`.

4. Detalle_Recoleccion

- **Propósito:** Desglosa los materiales específicos recolectados en cada recolección.
- **Funcionamiento en el sistema:**
 - Cada entrada tiene un `id_detalle` y vincula un `id_recoleccion` con un `id_material`, especificando la cantidad (en kg).
 - Permite un análisis granular de lo recolectado.
 - Ejemplo: En una recolección (`id_recoleccion` = 10), se registran 20 kg de plástico (`id_material` = 1) y 10 kg de vidrio (`id_material` = 2).
- **Relación con ML:** Podría alimentar un modelo para clasificar materiales automáticamente (ej. usando imágenes) o predecir proporciones de materiales en futuras recolecciones.

5. Procesos_Reciclaje

- **Propósito:** Registra los pasos de procesamiento aplicados a los materiales recolectados (clasificación, trituración, etc.).
- **Funcionamiento en el sistema:**
 - Cada proceso tiene un `id_proceso` y está vinculado a una `id_recoleccion` y un `id_material`.
 - Campos como `tipo_proceso`, `fecha_inicio`, `fecha_fin` y `costo` rastrean el flujo de trabajo.
 - **Ejemplo:** El plástico recolectado en `id_recoleccion` = 10 se tritura el 21/03/2025; se registra aquí con `resultado` = "Material triturado".

- **Relación con ML:** Un modelo podría optimizar los procesos (ej. sugerir el orden más eficiente) o predecir costos/tiempos, almacenando resultados en Predicciones.

6. Reportes

- **Propósito:** Genera informes sobre las operaciones del centro, ya sea manualmente o automáticamente.
- **Funcionamiento en el sistema:**
 - Cada reporte tiene un id_reporte, vinculado a un id_usuario, con detalles en tipo_reporte y contenido.
 - Puede resumir datos de Recolecciones, Detalle_Recoleccion o Procesos_Reciclaje.
 - **Ejemplo:** Un empleado genera un reporte mensual el 20/03/2025 con el total de kg reciclados se guarda aquí.
- **Relación con ML:** Un modelo podría generar automáticamente el contenido del reporte basado en patrones detectados en otras tablas.

7. Modelo_ML

- **Propósito:** Almacena información sobre los modelos de machine learning usados en el sistema.
- **Funcionamiento en el sistema:**
 - Cada modelo tiene un id_modelo, con detalles como nombre, tipo (regresión, clasificación), y ruta_archivo (donde se guarda el modelo entrenado).
 - Registra cuándo se entrenó (fecha_entrenamiento) y sus parametros.
 - **Ejemplo:** Un modelo "Clasificador de materiales" se entrena el 15/03/2025 y se usa para identificar materiales en Detalle_Recoleccion.
- **Relación con ML:** Esta tabla es el núcleo de la integración de ML, gestionando los modelos que generan Predicciones.

8. Predicciones

- **Propósito:** Almacena los resultados de las predicciones generadas por los modelos de machine learning.
- **Funcionamiento en el sistema:**

- Cada predicción tiene un id_prediccion, vinculada a un id_modelo y opcionalmente a una id_recoleccion.
- Campos como entrada (datos usados), resultado (predicción) y precision registran el output.
- Ejemplo: El modelo predice que el 25/03/2025 se recolectarán 60 kg de plástico en una zona; se guarda aquí con resultado = "60 kg".
- **Relación con ML:** Es la salida práctica del sistema de machine learning, utilizada para optimizar recolecciones, procesos o reportes.

Diccionario De Datos:

Usuarios

Campo	Tipo de Dato	Descripción	Restricciones
id_usuario	INT	Identificador único del usuario	PK, Auto Increment
nombre	VARCHAR(100)	Nombre completo del usuario	NOT NULL
tipo_usuario	ENUM('residente', 'empleado', 'empresa')	Tipo de usuario	NOT NULL
email	VARCHAR(100)	Correo electrónico del usuario	UNIQUE, NOT NULL
telefono	VARCHAR(15)	Número de teléfono del usuario	NULLABLE
direccion	VARCHAR(255)	Dirección física del usuario	NULLABLE
fecha_registro	DATE	Fecha de registro en el sistema	NOT NULL

Materiales

Campo	Tipo de Dato	Descripción	Restricciones
id_material	INT	Identificador único del material	PK, Auto Increment
Nombre_m	VARCHAR(50)	Nombre del material	NOT NULL

descripcion	TEXT	Descripción breve del material	NULLABLE
categoria	VARCHAR(50)	Categoría del material	NOT NULL
valor_por_kg	DECIMAL(10,2)	Valor económico por kilogramo	NULLABLE

Recolecciones

Campo	Tipo de Dato	Descripción	Restricciones
id_recoleccion	INT	Identificador único de la recolección	PK, Auto Increment
id_usuario	INT	ID del usuario asociado	FK (Usuarios), NOT NULL
fecha_recoleccion	DATETIME	Fecha y hora de la recolección	NOT NULL
ubicacion	VARCHAR(255)	Dirección o coordenadas de la recolección	NOT NULL
estado	ENUM('pendiente', 'completada', 'cancelada')	Estado de la recolección	NOT NULL
peso_total	DECIMAL(10,2)	Peso total recolectado (kg)	NULLABLE

Detalle_Recoleccion

Campo	Tipo de Dato	Descripción	Restricciones
id_detalle	INT	Identificador único del detalle	PK, Auto Increment
id_recoleccion	INT	ID de la recolección asociada	FK (Recolecciones), NOT NULL
id_material	INT	ID del material recolectado	FK (Materiales), NOT NULL
cantidad	DECIMAL(10,2)	Cantidad recolectada (kg)	NOT NULL
observaciones	TEXT	Notas adicionales	NULLABLE

		sobre el material	
--	--	-------------------	--

Procesos_Reciclaje

Campo	Tipo de Dato	Descripción	Restricciones
id_proceso	INT	Identificador único del proceso	PK, Auto Increment
id_recoleccion	INT	ID de la recolección asociada	FK (Recolecciones), NOT NULL
id_material	INT	ID del material procesado	FK (Materiales), NOT NULL
tipo_proceso	VARCHAR(50)	Tipo de proceso realizado	NOT NULL
fecha_inicio	DATETIME	Fecha y hora de inicio del proceso	NOT NULL
fecha_fin	DATETIME	Fecha y hora de finalización	NULLABLE
costo	DECIMAL(10,2)	Costo del proceso	NULLABLE
resultado	VARCHAR(100)	Resultado del proceso	NULLABLE

Reportes

Campo	Tipo de Dato	Descripción	Restricciones
id_reporte	INT	Identificador único del reporte	PK, Auto Increment
id_usuario	INT	ID del usuario que generó el reporte	FK (Usuarios), NOT NULL
fecha_generacion	DATETIME	Fecha y hora de generación del reporte	NOT NULL
tipo_reporte	VARCHAR(50)	Tipo de reporte	NOT NULL
contenido	TEXT	Contenido del reporte (JSON o texto)	NOT NULL

Modelo_ML

Campo	Tipo de Dato	Descripción	Restricciones
id_modelo	INT	Identificador único del modelo	PK, Auto Increment
Nombre_modelo	VARCHAR(50)	Nombre del modelo	NOT NULL
tipo	VARCHAR(50)	Tipo de modelo (regresión, clasificación)	NOT NULL
fecha_entrenamiento	DATETIME	Fecha y hora del último entrenamiento	NOT NULL
parametros	TEXT	Configuración del modelo (JSON)	NULLABLE
ruta_archivo	VARCHAR(255)	Ubicación del archivo del modelo	NOT NULL

Predicciones

Campo	Tipo de Dato	Descripción	Restricciones
id_prediccion	INT	Identificador único de la predicción	PK, Auto Increment
id_modelo	INT	ID del modelo que generó la predicción	FK (Modelo_ML), NOT NULL
id_recoleccion	INT	ID de la recolección asociada (si aplica)	FK (Recolecciones), NULLABLE
fecha_prediccion	DATETIME	Fecha y hora de la predicción	NOT NULL
entrada	TEXT	Datos de entrada usados (JSON)	NOT NULL
resultado	TEXT	Resultado de la predicción	NOT NULL
precision	DECIMAL(3,2)	Nivel de confianza de la predicción	NOT NULL

