

1.El hecho de querer importar archivos implica la existencia de una base de datos?

R: La pregunta es que debe existir una base de datos centralizada con el tema de captación de la información . Como explique en la exposición puede haber capturas de información de diferentes fuentes estructuras como datos no estructurados.

Cuando no existe una base de datos las posibilidades pueden ser las siguientes (data no estructurada):

- Archivos como Excel, CSV, JSON o XML pueden ser utilizados temporalmente para análisis o generación de reportes.
- Los datos pueden ser procesados en memoria utilizando herramientas como Python, R o software de análisis.
- En este caso, los archivos actúan como la fuente principal de datos, pero no se almacenan de forma estructurada ni persistente

En el caso de que exista una base de datos estructurada:

- Archivos que contienen datos históricos, climáticos, de plagas o de mercado que se cargan en tablas de la base de datos. (<https://agrometeorologia.cl/>)
(<https://www.agromet.cl/>) <https://www.agromet.cl/datos-historicos>

2) Si la fuente es online se hace una copia a una base de datos o se examina en tiempo real?

R: El grupo debe definir si esta data es estructurada o no estructurada que va a utilizar ya que puede ser un modelo combinado

Si te colocas en un escenario de copia de un dbms debería tener considerado estas variables:

- **Acceso rápido y eficiente:** Los datos están disponibles localmente, lo que permite consultas rápidas sin depender de la conexión a la fuente en línea.
- **Historial y trazabilidad:** Se puede mantener un registro histórico de los datos, lo que es útil para análisis a largo plazo o auditorías.
- **Optimización de consultas:** Los datos pueden ser organizados y optimizados para consultas específicas, como índices o estructuras de datos personalizadas
- **Integración con otros sistemas:** Los datos almacenados pueden ser combinados con otras fuentes o utilizados en sistemas internos.

Si te pones en el escenario de trabajar datos en tiempo real debe tener considerado (realizar es web-scraping)

Datos siempre actualizados: Se accede a la información más reciente directamente desde la fuente.

Sin necesidad de almacenamiento: No se requiere espacio adicional para guardar los datos.

Menor mantenimiento: No es necesario preocuparse por sincronizar o actualizar una base de datos local.

3) En caso de hacer una copia de una fuente que es online, cada cuanto tiempo es necesario verificar que esa fuente sigue al día?

R: Las publicaciones de reportes son cada 3 meses aproximadamente

<https://opia.fia.cl/601/w3-propertyvalue-148969.html>

lo que se recomendaría es que fuera mensual dicha verificación.

4. El usuario va a interactuar directo con la IA/red neuronal o se añaden los datos a una base de datos que es la que interactúa con un algoritmo?

R: Como se explicó ese día en la exposición se interactúan con algunas ia para la recopilación de información más precisa como se dijo tiene sus ventajas y desventajas las cuales paso a explicar a continuación.

Ventajas:

- **Respuestas en tiempo real:** El usuario recopilador obtiene resultados inmediatos basados en los datos proporcionados. Según la temática propuesta
- **Simplicidad en el flujo de datos:** No es necesario almacenar los datos en una base de datos antes de procesarlos. Esta actualmente se trabaja en un repositorio Fileserver de la unidad donde se deja por temática encontradas como fuentes para el boletín o futuros boletines.

Desventajas:

- **Falta de historial:** Si los datos no se almacenan, no se puede realizar un análisis posterior o mantener un registro de las interacciones. (Mejora futura) o modelo de datos a proponer por parte del equipo de desarrollo USM.
- **Dependencia del modelo:** Todo el procesamiento depende de la IA, lo que puede ser un problema si el modelo no está optimizado o si hay problemas de rendimiento. Es puede producir un desgaste ya que existen varios modelos de ia y estos evolutivos en el mercado en su precisión.

Ejemplo o temáticas:

Modelos de Predicción de Precios

(Regresión Lineal y Polinómica, Redes Neuronales Recurrentes (RNN) y LSTM, Modelos ARIMA)

Modelos de Análisis de Demanda

Redes Neuronales Artificiales (ANN), Modelos de Clustering (K-Means, DBSCAN).

Modelos de Optimización de la Producción y Distribución

Redes Neuronales Profundas (DNN), Modelos de Aprendizaje por Refuerzo (Reinforcement Learning)

Modelos de Análisis de Tendencias del Mercado

Procesamiento de Lenguaje Natural (NLP), Modelos de Clustering y Análisis de Componentes Principales (PCA)

Modelos de Predicción Climática y su Impacto en el Mercado

Modelos de Series Temporales (LSTM, GRU), Modelos de Simulación Basados en Agentes: Para simular el impacto del clima en la oferta y demanda de productos agrícolas.

Modelos de Detección de Anomalías:

Máquinas de Soporte Vectorial (SVM), Modelos de Detección de Cambios (Change Point Detection)

Modelos de Recomendación

Sistemas de Filtrado Colaborativo: Para recomendar productos agrícolas a compradores o mercados específicos.

Basados en Contenido: Para sugerir cultivos o prácticas agrícolas a los agricultores según sus condiciones locales.

Modelos de Clasificación y Segmentación

K-Means y Clustering Jerárquico, Clasificación Supervisada (SVM, Random Forest)

5) Cada requisito del 1 al 4 necesita su propio portal/subdominio dentro del sitio web o existen requisitos que van lado a lado con otros?

Requisito 1: Recopilación de datos

Requisito 2: Procesamiento de información

Requisito 3: Generación de boletines

Requisito 4: Publicación y distribución

Para poder abordar los requisitos del proyecto de vigilancia agrícola desde un enfoque de ingeniería de software, es necesario estructurar el proceso en etapas claras, utilizando principios de diseño, desarrollo y gestión de software. Se procede a explicar cómo se puede organizar el proceso para cumplir con los requisitos planteados (1-2-3-4)

Recopilación de datos:

1. Análisis de requisitos de datos (fuentes)
2. Diseño del módulo de recopilación (Como Crear, diseño de artefacto integración)
3. Tecnologías sugeridas
4. Resultado esperado (datos recopilados en forma estructurada, almacenados en repositorio temporal o final) esto debe definirlo el equipo de desarrollo

Procesamiento de información:

1. Diseño del flujo de procesamiento (**pipeline de datos**)
2. Implementación modular (procesamiento en etapas independientes ETL)
3. Integración de modelos de IA (Entrenar e integrar modelos de predicción, frameworks como TensorFlow, PyTorch o Scikit-learn)
4. Gestión de errores (manejo de excepciones para datos corruptos o fuentes no disponibles).
5. Resultado esperado (Información procesada y almacenada en una base de datos estructurada, lista para ser utilizada en la generación de boletines)

Generación de boletines:

1. Diseño del formato de boletines (Definir plantillas estándar para los boletines)
2. Automatización de la generación (Crear un módulo que tome los datos procesados y los inserte en las plantillas).
3. Personalización (Permitir la generación de boletines personalizados según la región, tipo de cultivo o necesidades del usuario, Permitir la generación de boletines personalizados según la región, tipo de cultivo o necesidades del usuario)
4. Pruebas de calidad (Validar que los boletines generados sean correctos, legibles y contengan toda la información necesaria)
5. Resultado esperado (Boletines generados automáticamente en formatos como PDF, HTML o correo electrónico).

Publicación y distribución

1. Diseño del sistema de distribución (Identificar los canales de distribución, Implementar un sistema de notificaciones para alertar a los usuarios sobre nuevos boletines).
2. Automatización de la publicación (Configurar un sistema de envío automatizado)
3. Gestión de usuarios (Registro y autenticación, Configuración de preferencias (por ejemplo, frecuencia de boletines, temas de interés)
4. Monitoreo y retroalimentación (implementar herramientas de análisis para monitorear la distribución , Recopilar retroalimentación de los usuarios para mejorar el contenido y la distribución)
5. Resultado esperado (Boletines distribuidos de manera eficiente a los usuarios finales a través de los canales seleccionados).

6. El boletín es un resumen? si no, se guardan por separado los boletines y resúmenes?

R: Los boletines no son resúmenes, sino artículos enfocados en una temática específica. Para obtener más información, puedes visitar el siguiente enlace.

<https://opia.fia.cl/601/w3-propertyvalue-148969.html>

<https://opia.fia.cl/601/w3-propertyvalue-147024.html>