

## **Caso 1: Digitalización de Facturas Manuscritas en formato JSON**

### **1. Propuesta de solución basada en IA**

Para llevar a cabo la digitalización de facturas manuscritas se propone como tecnología principal el uso de **Azure AI Document Intelligence** (anteriormente conocido como **Azure Form Recognizer**). Según la documentación oficial, este servicio incluye modelos preentrenados que permite extraer texto de diferentes formatos y devuelve contenido estructurado de calidad.

El workflow de la propuesta incluye:

1. Cargar la factura (imagen o PDF) en el servicio de reconocimiento.
2. Aplicar modelo OCR preentrenado para convertir la escritura manuscrita en texto digital, concretamente el modelo **Prebuilt Layout** (llamada a la API de Azure AI Document Intelligence).
3. Extraer y clasificar automáticamente los campos clave de la factura (Fecha, Total, Número de Factura, Cantidad, etc.).
4. Crear una función que devuelva como resultado final un JSON estructurado con la información extraída, tal y cómo se describe en la petición.
5. Aplicación de métricas, observación de resultados y valoración de uso de técnicas y/o tecnologías extra para mejorar rendimiento si necesario.

### **2. Técnicas/modelos a implementar y justificación**

- **Azure AI Document Intelligence**: Esta tecnología parece ofrecer la mejor solución en comparación con otras que han sido valoradas (Google Vision AI, Amazon Textract o OCR Tesseract), ya que se encuentra optimizada en la detección de alta precisión de textos manuscritos, por lo que sería necesario entrenar un modelo ni realizar fine-tune. Además, es fácilmente escalable, ya que se puede integrar con sistemas como Power Automate si así se estimara.

Enlace 1: [Documentación Azure AI Document Intelligence](#)

- **Modelo Prebuilt Layout**: La tecnología de Azure comentada ofrece diversos modelos disponibles, pero el que puede ofrecer mejores resultados en nuestro caso es Prebuilt Layout. Según la documentación, este modelo permite, además de reconocer escritura manuscrita, detectar la estructura del documento y devolver las posiciones de cada campo en la página, lo que resulta esencial en el caso de facturas (datos organizados en tablas o bloques).

De esta manera, no sólo obtenemos el texto extraído, si no que nos permite la identificación de valores específicos como “Total”, “Fecha” o “Cantidad”. Todas estas características facilitan enormemente la posterior estructuración en formato JSON.

Enlace 2: [Documentación Layout Models](#)

En el caso de no obtener los resultados esperados después de evaluar el rendimiento de la solución (manuscrito poco reconocible, oscuridad, ruido en la imagen, etc.), podemos valorar añadir técnicas extra que podrían mejorar nuestros resultados:

- **Preprocesamiento de imagen:** conversión a **escala de grises**, ajustes de **brillo**, aplicar **filtros de ruido** para mejorar la detección de caracteres o corrección de **inclinación** (si necesario).
- **Postprocesamiento:** Si el texto extraído contiene algún error que se repite de forma continuada en la extracción, podríamos valorar utilizar **expresiones regulares (Regex)** a modo de postprocesamiento, o incluso alguna técnica NLP, como puede ser **Azure Text Analytics** en el caso de que tengamos problemas con el reconocimiento de categorías.

### 3. Posible fine-tuning según necesidades

Si después de aplicar técnicas de pre y postprocesamiento los resultados siguen sin ser óptimos, es posible que en algunos casos pueda sea necesario la aplicación de un entrenamiento o fine-tuning (facturas con formato no estándar, facturas de distintos proveedores con estructura variable, excesivo ruido en la imagen, etc.). En este caso, Azure permite entrenar modelos personalizados con **pocos ejemplos**, por lo que podría ser una solución sin necesidad de invertir un tiempo excesivo. El proceso sería el siguiente:

1. Recolección y subida de 10-15 facturas de ejemplo (incluir la mayor variabilidad posible) a **Azure AI Document Intelligence Studio**.
2. Etiquetado manual de los campos clave en cada factura.
3. **Entrenamiento supervisado** con etiquetas (llegado a este punto del proceso, buscaremos obtener la mayor precisión posible con training supervisado, dado que las estrategias utilizadas anteriormente no han dado buenos resultados).
4. Evaluación de resultados y ajuste del modelo.

### 4. Métricas

En cada checkpoint, se tendrán en cuenta la siguientes métricas para evaluar el rendimiento del modelo:

- **Accuracy:** Porcentaje de campos extraídos correctamente.
- **Confidence Score:** Nivel de confianza en cada predicción.
- **Field Extraction Rate:** Porcentaje de campos detectados en comparación con los esperados.

Obtendremos un dataset de valores reales de las facturas manuscritas realizando una primera extracción con Prebuilt Layout y corrigiendo errores.

En caso de necesitar fine-tuning, se procederá a realizar una división 80/20 para datos de entrenamiento y testeo. Por último, ejecutaremos el modelo prebuilt layout con los mismos datos de test para hacer una comparación de métricas entre ambos modelos.

## **5. Justificación del posible éxito**

El éxito de la solución radica en la combinación de tecnología avanzada y sin necesidad de entrenar un modelo desde cero. Azure AI Document Intelligence (Prebuilt Layout) ha sido seleccionado por su capacidad para extraer texto manuscrito con alta precisión y estructurar la información en tablas, superando a otras herramientas alternativas (gratuitas o soluciones de pago).

Para mejorar la extracción en facturas de difícil manejo, se contemplan ajustes en preprocesamiento y postprocesamiento antes de recurrir a un fine-tuning. Además, la solución se evalúa con métricas clave en diferentes, asegurando que el modelo alcanza niveles de precisión superiores al 90%.

Más allá de la precisión, esta propuesta garantiza escalabilidad y facilidad de integración, ya que permite procesar volúmenes variables de documentos sin infraestructura adicional y se adapta a flujos de trabajo automatizados o sistemas ERP. Su implementación optimiza costos, reduce la intervención manual y la inversión de tiempo adicional, asegurando una digitalización eficiente de facturas manuscritas.