

## Laboratorio 1: Transformación y Análisis Enriquecido de CVs

### Objetivo General

Desarrollar una herramienta que automatice la conversión de CVs a la plantilla estandarizada de SOAINT y que, adicionalmente, genere resúmenes personalizados y extraiga insights estratégicos valiosos de los CVs procesados.

### Componentes del Laboratorio

1. **Transformación Automática de CVs a Plantilla de SOAINT:** Crear un sistema que transforme eficientemente CVs de cualquier formato al estándar de SOAINT, garantizando la uniformidad y facilitando la revisión.
2. **Generación de Resúmenes Personalizados:** Implementar una funcionalidad que produzca resúmenes de los CVs, destacando información crítica alineada con requisitos específicos de posiciones o proyectos.
3. **Extracción de Insights Deseables:** Desarrollar mecanismos para identificar y reportar insights estratégicos basados en la información contenida en los CVs, utilizando técnicas avanzadas de procesamiento de texto.

### Insights Deseables

- **Perfiles de Tecnología y Experiencia:** Identificar las tecnologías y herramientas más frecuentes mencionadas en los CVs para construir un mapa del stack tecnológico predominante entre los candidatos.
- **Clientes y Proyectos Destacados:** Extraer menciones de trabajos realizados para clientes B2C, incluyendo el tipo de proyectos y las tecnologías utilizadas, para identificar oportunidades comerciales y entender mejor las necesidades del mercado.
- **Análisis de Competencias y Habilidades:** Analizar las competencias y habilidades mencionadas para identificar tendencias y posibles áreas de desarrollo o capacitación necesarias.
- **Experiencia Sectorial:** Destacar la experiencia de los candidatos en diferentes sectores o industrias, ofreciendo una visión de la diversidad y especialización del talento disponible.

### Tecnologías a Utilizar

- **Python** para el desarrollo general.

- **OpenAI GPT** para el procesamiento y análisis avanzado de texto, y extracción de información de imágenes.
- **Librerías de Python** (`python-docx`, `PyPDF2`) para la manipulación de documentos.

## Estructura del Laboratorio

1. **Preparación Inicial:** Configurar un entorno de desarrollo con las herramientas necesarias y recolectar una muestra representativa de CVs.
2. **Desarrollo de la Funcionalidad de Transformación:** Codificar la lógica para leer CVs en distintos formatos y reestructurar la información en la plantilla de SOAINT.
3. **Implementación de Generación de Resúmenes:** Utilizar OpenAI GPT para desarrollar una funcionalidad que genere resúmenes enfocados, basados en criterios específicos suministrados por el usuario.
4. **Extracción y Presentación de Insights:** Diseñar y programar un sistema que, mediante el análisis de texto, identifique y presente insights valiosos extraídos de los CVs.
5. **Documentación y Demostración:** Documentar el proceso de desarrollo y preparar una presentación o demostración de la herramienta, destacando su utilidad y los insights generados.

# Laboratorio 2: Análisis Avanzado y Dashboard Interactivo de Reporte del CRM

## Objetivo General

Desarrollar una solución integral que automatice la extracción de insights significativos a partir de un reporte del CRM, utilizando técnicas avanzadas de análisis de datos con Python, la API de OpenAI para enriquecer el análisis, y Streamlit para la visualización interactiva.

## Descripción

Este laboratorio desafía al equipo a crear una herramienta que no solo automatice la generación de reportes a partir de datos de Excel sino que también utilice capacidades avanzadas de procesamiento de lenguaje natural (NLP) y visión computacional para extraer y analizar información, culminando en la creación de un dashboard interactivo para presentar los resultados.

## Tecnologías a Utilizar

- **Python:** Como lenguaje principal de desarrollo.
- **Pandas y Numpy:** Para manipulación y análisis de datos.
- **OpenAI API:** Para generar insights avanzados y procesar texto o imágenes.
- **Streamlit:** Para desarrollar un dashboard interactivo que presente los análisis y resultados.
- **Matplotlib/Seaborn:** Para visualizaciones dentro del dashboard.

## Estructura del Laboratorio

- Preparación y Exploración de Datos:**
  - Cargar el reporte del CRM utilizando Pandas.
  - Realizar un análisis exploratorio para entender la estructura de los datos, incluyendo tipos de datos, columnas relevantes y limpieza necesaria.
- Análisis de Datos con Pandas y Numpy:**
  - Aplicar técnicas de manipulación de datos para preparar los datos para el análisis.
  - Realizar cálculos estadísticos básicos y avanzados para extraer insights preliminares.
- Generación de Insights con OpenAI:**

- Utilizar la API de OpenAI para procesar y analizar segmentos de texto o datos numéricos, buscando patrones o tendencias no evidentes.
- Explorar la utilización de modelos de visión de OpenAI para analizar cualquier elemento visual o textual dentro de los datos que requiera conversión a texto.

#### 4. **Desarrollo de Dashboards con Streamlit:**

- Crear un dashboard interactivo que muestre tanto los datos procesados como los insights generados.
- Incorporar visualizaciones de datos dinámicas que permitan a los usuarios explorar diferentes aspectos del análisis.

#### 5. **Documentación y Reflexión:**

- Documentar el proceso de desarrollo, incluyendo decisiones tomadas, desafíos enfrentados y cómo se superaron.
- Reflexionar sobre los insights generados y cómo pueden impactar decisiones de negocio o estrategias futuras.

### Guía para el Desarrollo

- **Inicio Rápido:** Asegúrate de tener un entorno Python con todas las bibliotecas necesarias instaladas. Explora brevemente el dataset para familiarizarte con su estructura.
- **Enfoque Iterativo:** Empieza con análisis simples y aumenta la complejidad gradualmente. No intentes abordar todo en un solo paso.
- **Experimentación con OpenAI:** Prueba diferentes prompts con la API de OpenAI para ver cuáles generan los insights más valiosos o interesantes.
- **Visualización y Presentación de Datos:** Considera qué tipo de gráficos o visualizaciones comunican mejor tus hallazgos. Streamlit permite una interacción fácil con los datos, aprovecha esta capacidad para hacer tu dashboard más dinámico.
- **Colaboración y Feedback:** Trabaja en equipo y busca feedback de tus colegas para iterar sobre el análisis y la presentación de datos.

### Insights Potenciales

#### 1. **Distribución de Probabilidades:**

- **Insight:** Analizar la distribución de las oportunidades entre diferentes categorías de probabilidad (No Aprobada, Cancelada, Aplazada, etc.).

	<ul style="list-style-type: none"> <li>○ <b>Impacto:</b> Esto puede ayudar a identificar tendencias en la aceptación o rechazo de propuestas y ajustar las estrategias de venta o enfoques de negociación.</li> </ul>
2.	<b>Rendimiento por Propietario del Negocio:</b> <ul style="list-style-type: none"> <li>○ <b>Insight:</b> Determinar el rendimiento en términos de valor generado por cada propietario del negocio. Identificar quiénes están sobresaliendo y quiénes podrían necesitar apoyo adicional.</li> <li>○ <b>Impacto:</b> Permitiría focalizar esfuerzos de capacitación, redistribuir recursos o reconocer y replicar estrategias exitosas.</li> </ul>
3.	<b>Impacto de la Probabilidad en el Valor de la Divisa de la Empresa:</b> <ul style="list-style-type: none"> <li>○ <b>Insight:</b> Evaluar cómo la probabilidad asignada a las oportunidades influye en el valor final en la divisa de la empresa. Identificar si ciertas categorías de probabilidad tienden a asociarse con valores más altos.</li> <li>○ <b>Impacto:</b> Ayuda a priorizar esfuerzos en las oportunidades con mayor potencial de retorno económico.</li> </ul>
4.	<b>Tasas de Conversión y Cancelación:</b> <ul style="list-style-type: none"> <li>○ <b>Insight:</b> Calcular tasas de conversión de oportunidades aprobadas vs. canceladas o aplazadas, para obtener una visión clara del éxito de las estrategias comerciales actuales.</li> <li>○ <b>Impacto:</b> Identificar áreas de mejora en el proceso de venta para aumentar las tasas de conversión.</li> </ul>
5.	<b>Análisis Temporal de Oportunidades:</b> <ul style="list-style-type: none"> <li>○ <b>Insight:</b> Si los datos temporales están disponibles, analizar tendencias estacionales o mensuales en la creación y cierre de oportunidades.</li> <li>○ <b>Impacto:</b> Permite ajustar estrategias de marketing y ventas basadas en temporadas de alto rendimiento.</li> </ul>

## Estrategias para Extraer Insights

- **Agrupación y Sumarización:** Utilizar Pandas para agrupar datos por categorías relevantes (como probabilidad o propietario del negocio) y sumarizar valores numéricos para identificar tendencias.
- **Visualización:** Crear visualizaciones gráficas (usando Matplotlib/Seaborn) para representar distribuciones, comparaciones entre grupos, y tendencias a lo largo del tiempo.
- **Análisis de Texto con OpenAI:** Para columnas de texto libre o comentarios, usar la API de OpenAI para extraer sentimientos o temas comunes que podrían informar sobre la satisfacción del cliente o áreas de mejora.

# Criterios de Evaluación y Formato de Presentación para Laboratorios

## Criterios de Evaluación

**1. Cumplimiento de Objetivos:** Se evaluará si los entregables satisfacen los objetivos definidos para cada laboratorio, incluyendo la automatización de procesos, generación de resúmenes personalizados y extracción de insights estratégicos.

**2. Calidad Técnica:** La calidad del código será revisada en términos de precisión técnica, eficiencia, y correcta implementación de las tecnologías y librerías especificadas. Se espera código bien organizado y documentado.

**3. Innovación y Creatividad:** Se valorará la aplicación creativa de las herramientas y tecnologías para resolver los problemas planteados, especialmente aquellas soluciones que superen los requisitos básicos y ofrezcan nuevas perspectivas o funcionalidades.

**4. Profundidad del Análisis:** Se considerará la profundidad y complejidad del análisis realizado, evaluando cómo se han utilizado los datos para generar insights valiosos y recomendaciones prácticas para el negocio.

**5. Presentación y Usabilidad:** La usabilidad del dashboard interactivo, la claridad de la documentación y la efectividad en la presentación de resultados serán críticos. Se espera que los participantes demuestren cómo sus proyectos pueden aplicarse en un contexto real de negocio.

## Formato de Entrega y Presentación

### Entrega de Código Fuente:

- **Plazo:** El código fuente de ambos laboratorios debe ser entregado a más tardar el **lunes 1 de abril en la noche**.
- **Método:** Deberán subir el código a un repositorio Git compartido o enviarlo directamente por correo electrónico, según se especifique.

### Presentaciones:

- **Fecha:** Las presentaciones se llevarán a cabo el **martes 2 de abril por la mañana**, asignándose un total de 2 horas para todas las presentaciones.

- **Duración:** Cada participante tendrá **15 minutos como máximo** para presentar su proyecto, incluyendo una demostración del sistema y tiempo para preguntas.
- **Formato:** Se espera que cada presentación cubra:
  - Una breve introducción del enfoque tomado.
  - Demostración de la funcionalidad desarrollada.
  - Discusión de los insights generados y su relevancia para el negocio.
  - Reflexión sobre los desafíos enfrentados y aprendizajes clave.
  - Se valorará la claridad, concisión y la capacidad de conectar el trabajo técnico con impactos prácticos en el negocio.

#### Notas Adicionales para los Participantes:

- Es fundamental que las soluciones presentadas no solo sean técnicamente sólidas, sino también aplicables y valiosas desde la perspectiva del negocio.
- La capacidad de comunicar efectivamente el trabajo realizado y su impacto potencial será tan importante como la calidad técnica del proyecto.
- Se anima a todos los participantes a revisar y practicar sus presentaciones para asegurar que se ajusten al tiempo asignado y cumplan con los objetivos deseados.