

Clasificación de datos con IBM Granit

LECCIÓN 1: CLASIFICACIÓN DE DATOS BASADOS EN TEXTO CON IA GENERATIVA

Introducción a la clasificación de datos

- Clasificar datos permite identificar patrones o problemas fácilmente.
- **Clasificación de datos:** proceso de organizar la información en categorías para facilitar su análisis y extracción de información.
- Es fundamental para transformar datos dispersos en conocimiento útil y tomar decisiones informadas.

¿Qué es la clasificación de texto?

- Clasificación de texto: organizar datos textuales en categorías específicas.
- Ejemplos comunes: filtrar correos spam o clasificar películas por género.
- Beneficios: permite automatizar la organización del texto y obtener información sin necesidad de clasificar manualmente.

Tipos de clasificación de texto

Existen tres tipos principales:

- **Binario:** el texto se clasifica en una de dos categorías (por ejemplo, spam o no spam).
- **Multiclase:** el texto se asigna a una sola categoría entre varias posibles.
- **Multietiqueta:** el texto puede pertenecer a múltiples categorías al mismo tiempo.

Inteligencia artificial generativa y clasificación de texto

- La IA generativa usa aprendizaje automático para detectar patrones en grandes volúmenes de datos y automatizar la clasificación de texto.
- Es ideal para tareas a gran escala por su precisión y consistencia al clasificar textos.

Proceso de clasificación con IA generativa:

1. **Ingesta de datos:** se recopilan textos (ej. solicitudes de empleo).

2. **Extracción de características:** se identifican datos clave usando procesamiento de lenguaje natural (PLN).
3. **Entrenamiento del modelo:** la IA aprende a clasificar según los datos.
4. **Resultado:** la IA clasifica automáticamente, ayudando a tomar mejores decisiones.

La automatización agiliza las operaciones

- Clasificar manualmente ya no es viable con grandes volúmenes de texto (ej. 900 millones de perfiles en LinkedIn).
- Automatizar permite:
 - Ahorrar tiempo
 - Mejorar la precisión
 - Escalar el proceso

Casos de uso de la clasificación automatizada:

- **Reseñas de productos:** detectar satisfacción/insatisfacción.
- **Atención al cliente:** analizar grandes volúmenes de feedback.
- **Monitoreo de redes sociales:** identificar tendencias y problemas rápidamente.

Resultado: mayor eficiencia, decisiones más rápidas y mejor uso de recursos.

Optimización de la clasificación

- La IA generativa optimiza la clasificación gracias a su velocidad, escalabilidad y consistencia.
- Permite clasificar textos más rápido, preciso y a gran escala, reduciendo errores humanos.

Ventajas clave:

- **Velocidad:** procesa grandes volúmenes rápidamente.
- **Escalabilidad:** se adapta al crecimiento de datos.
- **Consistencia:** mantiene resultados uniformes.

Cómo abordar los desafíos de la clasificación de texto con IA generativa

La IA generativa enfrenta algunos retos comunes. Sin embargo, puede adaptarse para superarlos:

1. Lenguaje ambiguo:

- Palabras con significados múltiples según el contexto.
- La IA mejora su comprensión contextual entrenándose con grandes volúmenes de datos diversos.

2. Variabilidad del tono y el sentimiento:

- El tono puede cambiar según la persona o situación.
- La IA identifica matices emocionales complejos con entrenamiento adecuado.

3. Estructura de datos inconsistente:

- Textos no estructurados o con formatos variables (ej. abreviaturas, errores).
- La IA es capaz de entender diferentes formatos y normalizarlos para clasificarlos correctamente.

LECCIÓN 2: MODELOS DE IBM GRANITE PARA CLASIFICAR DATOS

Modelos de Granite para la clasificación de textos

- **Modelos IBM Granite:** familia de modelos de lenguaje (LLM) diseñados para aplicar la IA generativa a tareas prácticas como la clasificación de texto.
- Hacen la IA más accesible y útil para aplicaciones del mundo real.

Ventajas:

- Clasificación más rápida, precisa y escalable en distintos contextos.

Seleccionar el modelo adecuado

- La efectividad depende de elegir el modelo Granite apropiado según el tipo de tarea.

Modelos destacados:

- **Granite-13b-chat-v2:** versátil, confiable y rentable. Ideal para tareas generales.
- **Granite-13b-instruct-v2:** enfocado en instrucciones precisas, útil en tareas especializadas. Ofrece precisión, comprensión de contexto y eficiencia.

Factores clave para elegir un modelo:

1. **Tipo de datos:** estructurado, desorganizado, variado, etc.
2. **Complejidad de la clasificación:** tareas simples o complejas.
3. **Idioma del contenido:** hay modelos optimizados por idioma (ej. inglés, japonés, multilingüe).

Modelos de traducción de idiomas de IBM Granite

- **Granite-13b-japanese:** diseñado para tareas en japonés, con capacidades de clasificación y traducción.
- **Granite-20b-multilingual:** admite múltiples idiomas, ideal para empresas globales que trabajan con textos en distintos idiomas.

Evaluación de las necesidades de tu proyecto

- Para elegir el modelo correcto, se deben tener en cuenta:
 - Objetivos del proyecto
 - Complejidad de los datos
 - Idioma y precisión requerida

Con una buena evaluación, es posible maximizar el rendimiento del modelo y aplicar la IA de manera efectiva en clasificación de textos.

LECCIÓN 3: REALIZAR INDICACIONES A IBM GRANITE PARA LA CLASIFICACIÓN DE TEXTO

Introducción a la realización de indicaciones

- En **IA generativa**, una **indicación (prompt)** es una instrucción que guía al modelo para generar una respuesta.
- Las indicaciones claras y específicas mejoran la precisión de los resultados.
- Las mejores indicaciones dan contexto y establecen el formato esperado del resultado.

Cómo escribir indicaciones eficaces para clasificación de textos

- Una buena indicación debe ser:
 - **Clara:** indicar exactamente qué se espera.
 - **Contextualizada:** incluir información relevante del contenido.
 - **Ejemplificada:** ofrecer ejemplos guía.

Ejemplo eficaz:

Indicación clara: "Clasifica esta reseña del cliente como Positiva o Negativa. Tu respuesta debe incluir solamente la categoría de clasificación."

Modificación de indicaciones para clasificación de texto

Si la indicación original no produce buenos resultados, se puede refinar con los siguientes pasos:

1. **Divide tareas complejas:** simplifica la indicación en pasos más pequeños y manejables.
2. **Refina las categorías:** define mejor las clases a clasificar para evitar ambigüedades.
3. **Especifica el formato del resultado:** aclara si se espera una palabra, frase, número, etc.

Estas mejoras ayudan a obtener respuestas más precisas, útiles y alineadas con los objetivos del proyecto.

Resumen de datos con IBM Granite

LECCIÓN 1: RESUMEN DE DATOS ESCRITOS CON IA GENERATIVA

La IA generativa y el resumen de texto

- **Resumen automático de texto:** método que usa la IA generativa para condensar uno o varios textos en un resumen coherente y conciso, manteniendo los puntos clave.
- Se basa en procesamiento de lenguaje natural (PLN) y modelos entrenados para entender y generar resúmenes útiles.

Tipos de resumen de texto

1. **Resumen extractivo:** selecciona frases clave directamente del texto original.
2. **Resumen abstractivo:** genera nuevas frases que condensan el significado global del texto de manera más natural y original.

Agilización del resumen de texto

La IA generativa agiliza la creación de resúmenes gracias a tres ventajas:

- **Precisión:** identifica las ideas clave con más exactitud.
- **Coherencia:** mantiene un hilo lógico entre frases.
- **Escalabilidad:** procesa grandes volúmenes de texto rápidamente.

Abordar los desafíos del resumen con IA generativa

Los principales retos y cómo la IA los enfrenta:

- **Sobrecarga de información:** filtra lo relevante.
- **Contenido específico del dominio:** adapta el resumen al tema concreto.
- **Lenguaje complejo:** traduce o simplifica el contenido técnico.
- **Resultados incoherentes:** mejora la fluidez y lógica del resumen.
- **Variabilidad del lenguaje:** maneja diferencias lingüísticas y culturales.

LECCIÓN 2: MODELOS IBM GRANITE PARA EL RESUMEN DE DATOS

Modelos de Granite para el resumen de textos

- Los modelos IBM Granite aplican IA generativa para crear resúmenes precisos, eficientes y fáciles de entender.
- Hacen que la IA sea accesible y útil para aplicaciones reales.

Modelos utilizados para resumen:

- **Granite-13b-chat-v2:** versátil, escalable y confiable.
- **Granite-13b-instruct-v2:** enfocado en coherencia, comprensión del contexto y generación de resúmenes efectivos.

Selección del modelo adecuado para resumen

La elección del modelo depende de:

1. Características clave:

- Granite-13b-chat-v2: mejor para tareas generales de resumen.
- Granite-13b-instruct-v2: ideal para instrucciones detalladas y resultados más precisos.

2. Tipo de datos:

- Estructurados, clínicos, jurídicos, etc.

3. Complejidad del contenido:

- Documentos largos o técnicos requieren modelos más robustos.

Evaluación de las necesidades del proyecto

Antes de seleccionar el modelo, evalúa:

- Precisión
- Legibilidad
- Pertinencia
- Concisión

Así se garantiza que el modelo ofrezca un resumen adaptado al propósito del proyecto, con calidad y claridad.

LECCIÓN 3: REALIZACIÓN DE INDICACIONES A IBM GRANITE PARA EL RESUMEN DE TEXTO

Cómo escribir indicaciones eficaces para resumir textos

- Una buena indicación (prompt) para resumir debe ser clara, concreta y ajustada a la tarea.
- Las mejores indicaciones tienen tres elementos clave:
 - **Claridad:** usar frases directas.
 - **Contexto:** explicar de qué trata el contenido.

- **Formato:** especificar cómo debe ser el resultado.

Indicaciones para varias tareas de resumen

Las indicaciones pueden ajustarse según la tarea que se desee realizar:

- **Agrupación:** combinar ideas de varios textos.
- **Comparación:** analizar similitudes o diferencias entre textos.
- **Extracción:** sacar información clave de documentos extensos.

El arte de dar indicaciones

- Una indicación bien formulada es crucial para que el modelo de IA entienda qué debe hacer.
- La calidad del resultado depende directamente de cómo está estructurada la indicación.

Perfeccionamiento de las indicaciones

Tres maneras de mejorar una indicación para obtener resúmenes más útiles:

1. **Especificar el enfoque deseado:** aclarar si se quiere un resumen general, legal, técnico, etc.
2. **Ajustar la longitud del resumen:** establecer si debe ser corto, medio o largo.
3. **Experimentar con el formato:** pedir lista de puntos, párrafos o frases clave.

Esto mejora la legibilidad, coherencia y adecuación del resultado a los objetivos del proyecto.

Clasificación y resumen de datos con IBM Granite

LECCIÓN 1: FORMULACIÓN DE SOLICITUDES EFICACES PARA CLASIFICAR Y RESUMIR DATOS

Tipos de clasificación

Tres categorías principales:

1. **Clasificación supervisada:** el modelo aprende con datos etiquetados.
2. **Clasificación no supervisada:** el modelo detecta patrones por sí solo.
3. **Clasificación jerárquica:** los datos se agrupan en niveles o subcategorías.

LECCIÓN 2: RESULTADOS DE LAS SOLICITUDES DE CLASIFICACIÓN Y RESUMEN

Resultados de las solicitudes de clasificación y resumen

- Las solicitudes bien formuladas son clave para obtener resultados de calidad en IBM Granite.
- A veces los resultados pueden no coincidir con las expectativas, lo que resalta la necesidad de una evaluación cuidadosa y crítica.
- Ejemplos de errores: clasificar mal un correo importante como “spam” u omitir información clave en un resumen.

Evaluación de los resultados de la clasificación

- Criterios clave:
 - **Precisión:** si la clasificación fue correcta.
 - **Pertinencia:** si la información clasificada es útil para el objetivo.
 - **Claridad:** si el resultado es comprensible.

Ejemplo: evaluación de correos electrónicos según importancia para soporte técnico.

Métodos de evaluación sistemática para clasificación

Dos enfoques clave:

1. **Comparación con la verdad fundamental:** comparar resultados con una clasificación previamente verificada.
2. **Revisión de expertos por pares:** especialistas humanos evalúan la calidad del resultado.

Se puede aplicar, por ejemplo, a solicitudes de soporte TI clasificadas como urgentes o no urgentes.

Evaluación de los resultados del resumen

- Criterios clave:
 - **Coherencia:** estructura lógica y conectada.
 - **Exhaustividad:** que incluya los puntos más importantes.
 - **Pertinencia:** que sea útil para el objetivo del resumen.

Técnicas para evaluar resultados del resumen

- **Análisis comparativo:** comparar resúmenes generados por humanos y por IA.
- Se observan diferencias de claridad, detalle o exactitud en los resúmenes generados para evaluar calidad.

Identificar fortalezas y debilidades

- Evaluar rendimiento ayuda a mejorar los modelos.
- En clasificación: revisar precisión, cobertura y eficacia.
- En resumen: revisar coherencia, inclusión de ideas clave y claridad.

Reconocimiento de áreas de mejora

- Una buena evaluación permite ajustar instrucciones, afinar modelos y optimizar los resultados.
- Se mejora así la utilidad general del sistema para las aplicaciones empresariales.

LECCIÓN 3: MEJORA DE LAS SOLICITUDES PARA UNA MEJOR CLASIFICACIÓN Y RESUMEN

Identificación de problemas en las solicitudes

- Incluso solicitudes bien intencionadas pueden fallar si son ambiguas, faltan datos o ejemplos, o son demasiado generales.
- **Problemas comunes:**
 - **Ambigüedad:** instrucciones poco claras.
 - **Falta de contexto:** no se incluye información relevante.
 - **Ejemplos insuficientes:** el modelo no entiende bien qué se espera.

Evaluación de resultados

- Es clave evaluar los resultados de clasificación y resumen en función de **métricas de calidad:**
 - Precisión y pertinencia
 - Cobertura temática
 - Explicación justificada

Estrategias para refinar las solicitudes

Cinco formas de mejorar una solicitud defectuosa:

1. Aclarar instrucciones
2. Añadir contexto
3. Incorporar ejemplos
4. Especificar el formato deseado
5. Ajustar la complejidad

31. Bucle de retroalimentación iterativo

Método cíclico para mejorar solicitudes mediante 4 pasos:

1. Evaluar los resultados
2. Ajustar las solicitudes
3. Probar y comparar
4. Documentar los resultados

Permite mejorar la eficacia del modelo paso a paso.

LECCIÓN 4: PARÁMETROS DE IBM GRANITE PARA FORMULAR SOLICITUDES Y MEJORAR CALIDAD DE RESULTADOS

Parámetros del modelo IBM Granite

- Los parámetros son opciones configurables que afectan cómo responde el modelo de IA.
- Se dividen en:
 1. **Parámetros de entrenamiento** (ajustan el comportamiento general del modelo).
 2. **Parámetros de inferencia** (afectan la salida del modelo en tareas específicas como clasificación o resumen).
- Un **token** es una unidad de texto procesada por el modelo (puede ser una palabra, parte de una palabra o un carácter).

Parámetros comunes para formular solicitudes

Algunos de los más utilizados:

Parámetro	Función	Ejemplo de uso
Top_k	Controla la diversidad de tokens generados	Bajo: coherente / Alto: diverso
Top_p	Ajusta la probabilidad acumulada de selección	Bajo: preciso / Alto: creativo
Max_tokens	Límite máximo de tokens generados	Bajo: conciso / Alto: detallado

Min_tokens	Mínimo de tokens requeridos	Bajo: breve / Alto: más desarrollado
Random_seed	Fija la aleatoriedad para reproducibilidad	Fijo: consistencia / Aleatorio: variabilidad
Repetition_penalty	Penaliza repeticiones	Bajo: repetitivo / Alto: variado
Stop_sequence	Detiene la generación al encontrar una secuencia específica	Activa la interrupción automática
Stop_criteria	Define condiciones de detención de la respuesta	Permite controlar la longitud y cierre

Aquí tienes el resumen de la parte nueva (lo que no se había cubierto aún):

Evaluación de resultados

Es crucial analizar la calidad del resultado y si cumple con los criterios esperados. Algunas métricas clave:

- Ambigüedad: genera resultados poco precisos.
- Falta de contexto: puede llevar a resultados incompletos.
- Objetivo inadecuado: causa desajuste entre la solicitud y el resultado.

Estrategias para refinar solicitudes

Se proponen cinco estrategias:

1. Añadir instrucciones
2. Añadir contexto
3. Incluir ejemplos
4. Especificar el formato deseado
5. Ajustar la complejidad

Bucle de retroalimentación iterativo

Proceso para mejorar solicitudes:

1. Evaluar resultados
2. Ajustar solicitudes
3. Probar y comparar
4. Documentar resultados

Ajustes de parámetros en práctica

Problemas comunes que se pueden ajustar:

- Respuestas repetitivas → aumentar diversidad (**top_p**)
- Respuestas demasiado generales → mayor contexto o ejemplos
- Respuestas largas/incompletas → ajustar **max_tokens** y **min_tokens**

Evaluación del impacto de los ajustes

Es importante medir cómo los cambios de parámetros afectan resultados.

- En clasificación: se evalúa **precisión, pertinencia y claridad**
- En resumen: se evalúa **coherencia, exhaustividad y pertinencia**

LECCIÓN 5: MEJORA DE RESULTADOS DEL MODELO AJUSTANDO PARÁMETROS DE IBM GRANITE PARA SOLICITUDES

Identificación de problemas en las solicitudes

- Las solicitudes mal formuladas pueden generar resultados pobres o irrelevantes.
- **Problemas comunes:**
 - **Ambigüedad:** falta de claridad.
 - **Falta de contexto:** sin suficiente información.
 - **Especificidad insuficiente:** muy general o vaga.

- Evaluar resultados es clave para detectar estos errores y ajustar parámetros o reformular.

Evaluación de resultados

- Las métricas permiten medir la calidad del modelo.
- **Problemas comunes evaluados:**
 - **Ambigüedad** → afecta precisión y pertinencia.
 - **Falta de contexto** → resultados localmente lógicos pero incorrectos globalmente.
 - **Especificidad baja** → omisiones o errores en resumen o clasificación.

Estrategias para refinar solicitudes

Para mejorar los resultados, puedes aplicar estrategias como:

- Aclarar instrucciones.
- Añadir contexto.
- Incluir ejemplos.
- Especificar formato deseado.
- Ajustar la complejidad.

Se recomienda usar un **bucle de retroalimentación iterativo**:

1. Evaluar resultados.
2. Ajustar solicitudes.
3. Probar y comparar.
4. Documentar mejoras.

Parámetros del modelo IBM Granite

- **Parámetros** = ajustes del modelo que influyen en su comportamiento.
- **Tokens** = unidades de texto (palabras, partes o caracteres).

Tipos de parámetros:

1. Controlan longitud o estructura de salida (p.ej., `max_tokens`).
2. Ajustan creatividad, diversidad o aleatoriedad (p.ej., `top_k`, `top_p`, `random_seed`).

Parámetros más usados y sus funciones

Parámetro	Controla...
<code>top_k</code>	Diversidad de palabras
<code>top_p</code>	Variedad en selección de tokens
<code>max_tokens</code>	Longitud máxima de respuesta
<code>min_tokens</code>	Longitud mínima de respuesta
<code>random_seed</code>	Reproducibilidad (control del azar)
<code>repetition_penalty</code>	Penaliza repeticiones
<code>stopping_criteria</code>	Condiciones para detener la generación
<code>stopping_sequences</code>	Palabras o secuencias que marcan el fin de salida

Aplicación práctica de ajustes

- Los problemas comunes como repetición o respuestas largas pueden resolverse ajustando parámetros.
- Se debe equilibrar claridad vs. creatividad, dependiendo del objetivo de la tarea.
- Se recomienda ajustar parámetros de forma iterativa (evaluar, cambiar, probar).

Evaluación del impacto de los ajustes

- Antes de modificar parámetros, se debe evaluar su impacto real en precisión, pertinencia y claridad.
- Para resúmenes: evaluar también coherencia y exhaustividad.