

Opciones de Clasificación y Columnas Recomendadas

Es importante tener en cuenta que, para que los modelos de clasificación funcionen correctamente, las variables predictoras deben estar en formato numérico, incluso si originalmente eran categóricas. Esto se logra mediante técnicas como Label Encoding o One-Hot Encoding, que transforman texto en números sin perder el significado de las categorías.

Opción 1: Clasificación por nivel de experiencia

La variable 'experience_level' clasifica a los empleados en:

- EN: Entry-level (nivel de entrada)
- MI: Mid-level (nivel medio)
- SE: Senior
- EX: Executive (alta dirección)

Clasificación binaria:

- Grupo 0: EN y MI → poca o mediana experiencia
- Grupo 1: SE y EX → mucha experiencia

¿Para qué sirve?

Este modelo permite predecir si un perfil es de alta experiencia, basándose en variables como:

- employment_type (tipo de contrato)
- job_title o job_group (tipo de cargo)
- company_size (tamaño de empresa)
- remote_ratio o is_remote (nivel de trabajo remoto)
- company_location y employee_residence (ubicación)
- salary_in_usd (opcional como predictor)

Aplicación real:

- - Sistemas de recomendación de empleo
- - Clasificación automática de candidatos por experiencia

Opción 2: Clasificación por tamaño de la empresa

La variable 'company_size' toma los valores:

- S: Small (pequeña)

- M: Medium (mediana)
- L: Large (grande)

Clasificación binaria:

- Grupo 0: S y M
- Grupo 1: L

¿Para qué sirve?

Predecir si alguien trabaja en una empresa grande o no, usando variables como:

- experience_level
- job_group
- salary_in_usd
- employment_type
- remote_ratio o is_remote
- company_location

Aplicación real:

- - Análisis de mercado laboral
- - Inferencia del tipo de empresa en portales de empleo

Opción 3: Clasificación por nivel de salario

La variable 'salary_in_usd' es continua. Se convierte en binaria usando la mediana:

- Grupo 0: Salario bajo
- Grupo 1: Salario alto

¿Para qué sirve?

Predecir si un perfil tendrá un salario alto o no. Variables recomendadas:

- experience_level
- job_group
- company_size
- employment_type
- is_remote
- company_location y employee_residence

Aplicación real:

- - Estimadores salariales en portales de empleo
- - Modelos de recomendación de compensación
- - Estudios sobre brechas salariales