Explore model explanations, error analysis, counterfactuals, and causal analysis by creating a Responsible AI dashboard. You'll create and run the pipeline in Azure Machine Learning using the Python SDK v2 to generate the dashboard.

**Learning objectives**

In this module, you learn how to:

* Understand Azure Machine Learning's built-in components for responsible AI.
* Create a Responsible AI dashboard.
* Explore a Responsible AI dashboard.

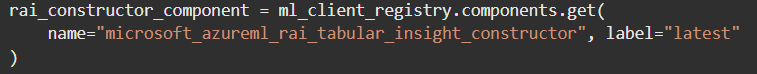
--- Intro

El **Responsible AI Dashboard** de Azure Machine Learning facilita la implementación de prácticas de IA responsables al ofrecer una visión completa de la interpretabilidad, equidad y rendimiento de tus modelos. Con este panel, puedes detectar problemas de sesgo, comprender los errores y tomar decisiones informadas sobre el despliegue seguro y ético del modelo.

--- Para crear un dashboard de IA Responsable (RAI) en Azure Machine Learning, puedes seguir estos pasos utilizando el SDK de Python. Este dashboard te permitirá evaluar si tu modelo cumple con los principios de IA Responsable, como equidad, transparencia y seguridad.

**Paso a Paso para Crear un Dashboard de IA Responsable**

1. **Registrar los Componentes Necesarios**
   * Primero, debes registrar tus datasets de entrenamiento y prueba como activos MLtable y tu modelo en el registro de Azure Machine Learning.
   * Luego, obtén los componentes predefinidos que Azure Machine Learning proporciona para construir el dashboard.
2. **Crear el Constructor del Dashboard**
   * El pipeline comienza con el componente de constructor del dashboard, que crea la base del dashboard de RAI.



1. **Agregar Componentes de Análisis de IA Responsable**

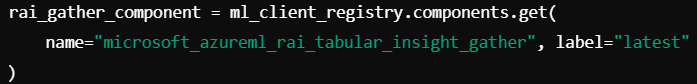
* Puedes elegir entre varias herramientas de análisis, como explicaciones y contrafactuales, para obtener una visión más profunda de cómo funciona el modelo.
* Por ejemplo, para agregar explicaciones al dashboard:

Una captura de pantalla de un celular con texto e imagen

Descripción generada automáticamente con confianza media

1. **Agregar el Componente de Recolección**

* Finaliza el pipeline con el componente de recolección que junta todos los insights generados en un solo dashboard interactivo.



1. **Definir el Pipeline con el SDK de Python**

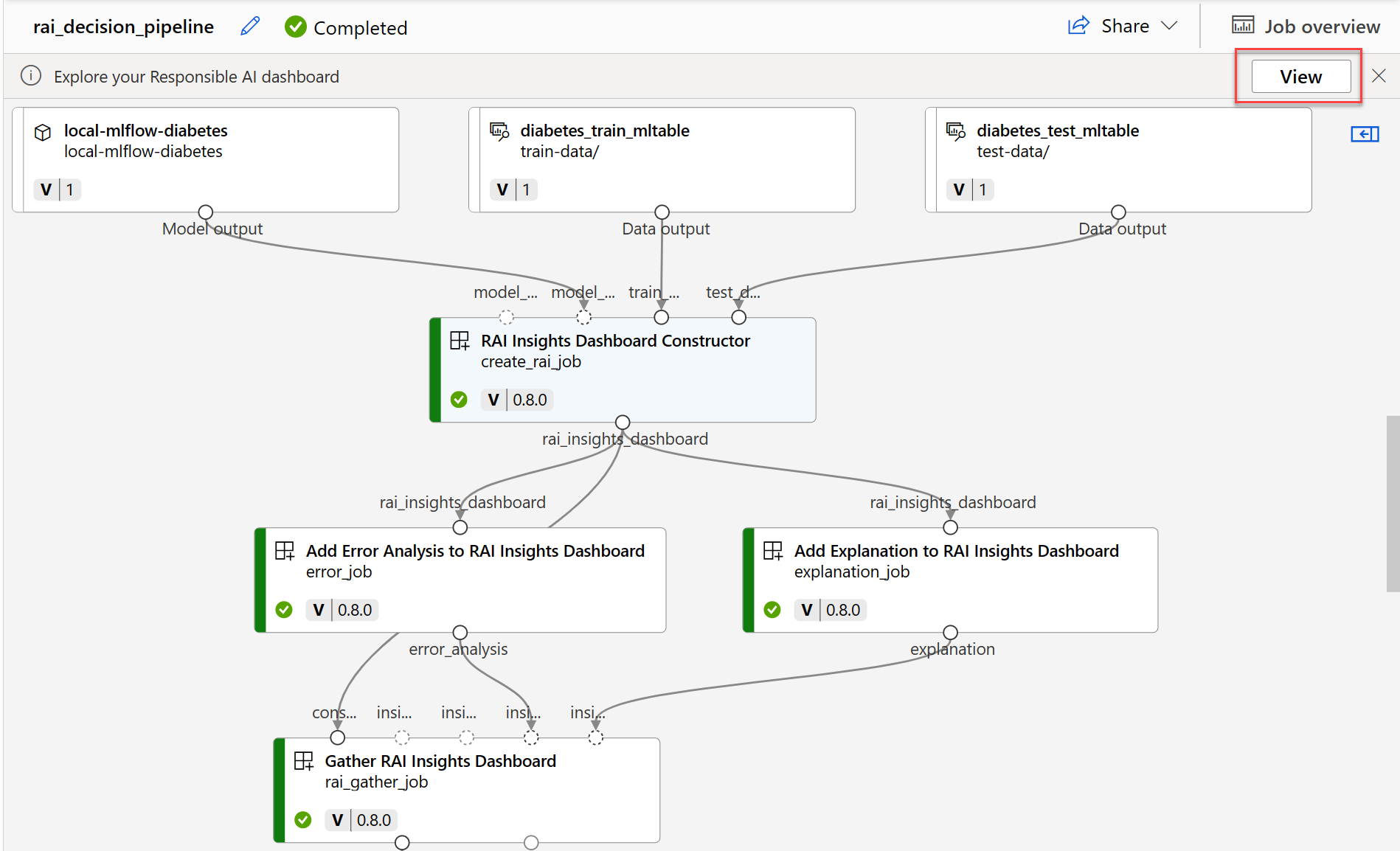
* Texto

  Descripción generada automáticamenteAhora puedes definir el pipeline que utilizará estos componentes.

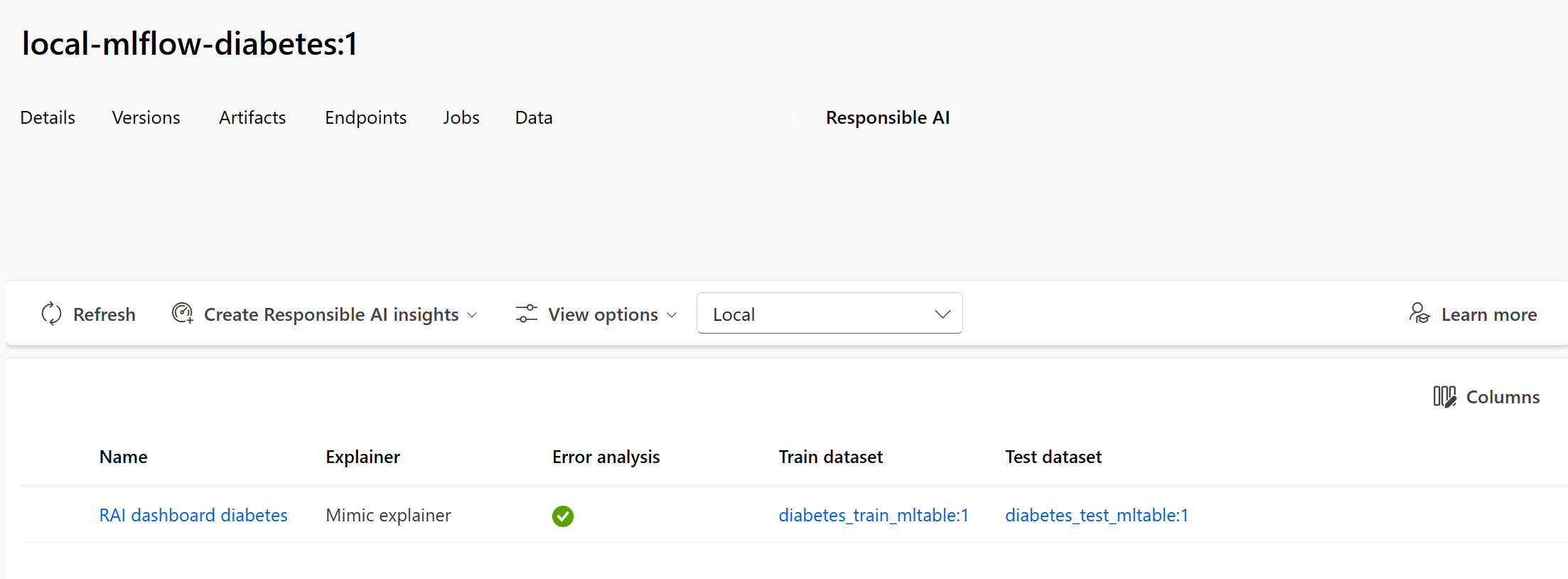
Texto

Descripción generada automáticamente

Después de ejecutar el pipeline y que se complete exitosamente, puedes acceder al **dashboard de IA Responsable** desde el resumen del pipeline en Azure Machine Learning Studio. Aquí, encontrarás las herramientas interactivas necesarias para explorar y evaluar varios aspectos de tu modelo en términos de los principios de IA Responsable.



Alternatively, you can find the Responsible AI dashboard in the **Responsible AI** tab of the registered model.



--- Para evaluar el **dashboard de IA Responsable** en Azure Machine Learning, aquí tienes un desglose de cómo analizar cada sección y los insights clave que puedes obtener:

**1. Análisis de Errores**

* **Árbol de Errores**: Este visual permite identificar combinaciones de subgrupos en los datos que llevan a una mayor tasa de errores en las predicciones. Por ejemplo, podrías descubrir que el modelo tiene más errores en ciertos rangos de edad o niveles de ingreso.
* **Mapa de Calor de Errores**: Proporciona una visión general de los errores en una escala de una o dos características, lo que facilita ver si ciertos valores de características específicas contribuyen a errores más frecuentes.

Estos insights son útiles para identificar si hay sesgos en el modelo y en qué grupos específicos su rendimiento es menos preciso.

**2. Explicaciones**

* **Importancia de Características Global (Aggregate)**: Muestra el peso de cada característica en las predicciones generales del modelo. Esto ayuda a entender qué factores están impulsando la mayoría de las decisiones y es esencial para la transparencia del modelo.
* **Importancia de Características Individual**: Permite ver cómo cada característica afecta una predicción específica, proporcionando una visión detallada de cada caso individual.

La importancia de características ayuda a responder preguntas sobre la lógica del modelo y asegura que las decisiones sean comprensibles y justificables.

**3. Contrafactuales**

* **Análisis de escenarios "qué pasaría si"**: Puedes seleccionar un punto de datos y simular cambios en las características de entrada para observar si el resultado cambia en la dirección deseada. Esto es útil para entender cómo cambios específicos en los datos de entrada podrían llevar a una predicción distinta.

Los contrafactuales son especialmente útiles para aplicaciones de toma de decisiones, ya que te permiten explorar cómo el modelo podría ajustarse en situaciones límite o cambiantes.

**4. Análisis Causal**

* **Efectos Causales Agregados**: Muestra el impacto promedio de características seleccionadas (como tratamientos o intervenciones) en la predicción deseada.
* **Efectos Causales Individuales**: Permite explorar cómo el cambio en características específicas afectaría el resultado para puntos de datos individuales.
* **Política de Tratamiento**: Identifica cuáles subgrupos de datos se beneficiarían más de una intervención específica.

Este análisis es valioso para escenarios donde deseas entender cómo modificar ciertos factores podría mejorar los resultados, proporcionando una base sólida para la toma de decisiones basadas en datos.

**Resumen de Evaluación**

Usando el dashboard, puedes:

* **Identificar y Mitigar Sesgos**: Con el análisis de errores, puedes reconocer sesgos en subgrupos específicos y tomar medidas para mejorar la equidad del modelo.
* **Mejorar la Interpretabilidad**: Las explicaciones globales e individuales aclaran el proceso de toma de decisiones del modelo.
* **Evaluar la Robustez del Modelo**: Los contrafactuales y el análisis causal muestran cómo el modelo responde a cambios en los datos de entrada, lo cual es crucial para garantizar la estabilidad y precisión en escenarios reales.

Evaluar el dashboard regularmente y antes de implementar el modelo en producción asegura que cumpla con los principios de IA Responsable, promoviendo su confiabilidad y equidad.