

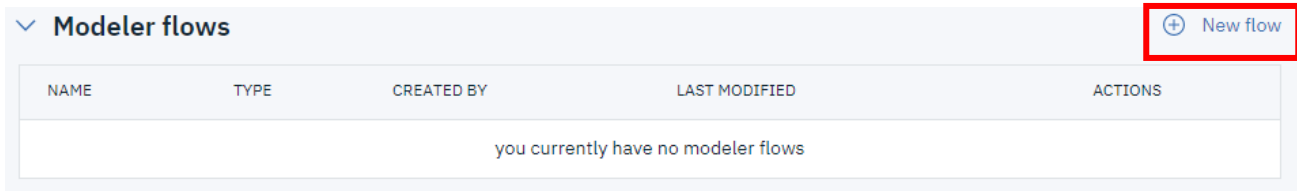


IBM Watson Studio Hands-On

Samir Char Iglesias – IBM Global Entrepreneur Program

Crear modelo SPSS

1. En su proyecto diríjase a “Modeler flows” y seleccione “New flow”.



2. Configure la instancia de spss de la siguiente manera:

New From file From example

Name*

Churn_spss

40

Description

Type description here.

500

Select flow type

☒ Modeler Flow ☐ Neural Network Modeler BETA

Runtime

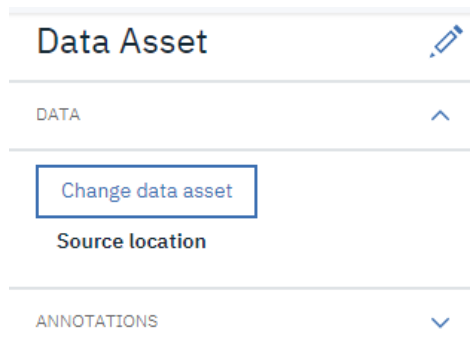
☒ IBM SPSS Modeler ☐ Scala Spark 2.1 BETA

Cree la instancia.

3. Notará que a la izquierda hay un panel con diferentes pestañas. En estas se encuentran los bloques que iremos arrastrando al canvas.

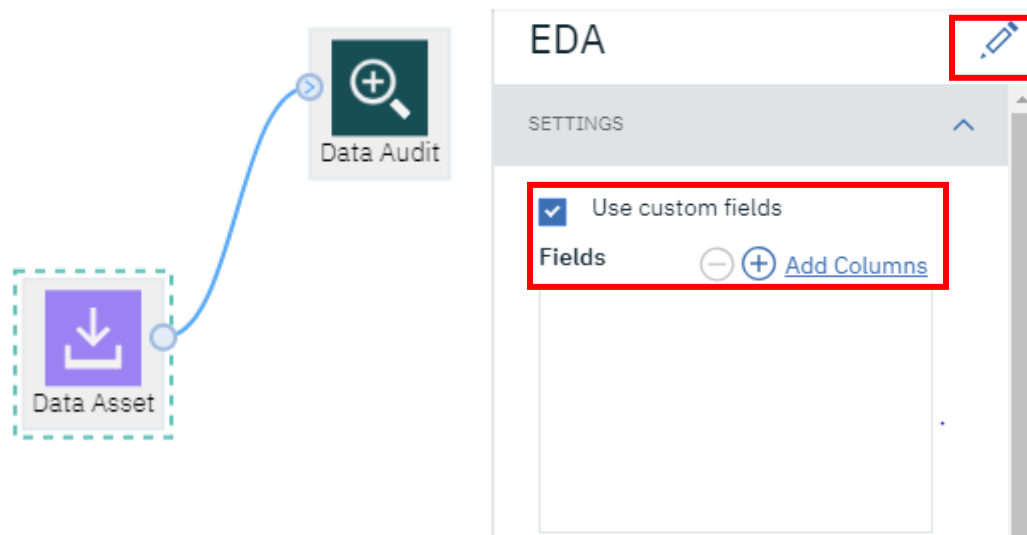


En la pestaña “Import” arrastre el bloque “Data Asset” al canvas. De doble click en este bloque y a la derecha deberá abrirse un panel. En este escoja la opción “Change data asset”



Seleccione el archivo “Churn.csv” y de click en “add”.

4. Bajo la pestaña “Outputs” arrastre el bloque de “Data Audit” y conéctelo con los datos. De doble click a este bloque que acaba de añadir, cambie el nombre a “EDA” (Exploratory Data Analysis) dando click en el ícono del lápiz, seleccione la casilla “Use custom fields” y de click en “Add Columns”.



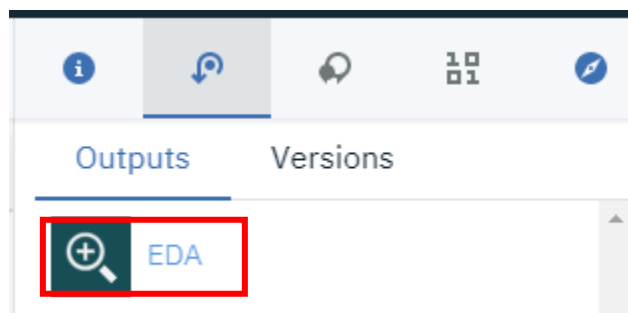
Deberá ver la siguiente imagen en donde debe dar click a la casilla enmarcada en rojo.

Select Fields for 11 Fields

Search in column Field name Filter: ☐ ☐ ☐ [Reset](#)

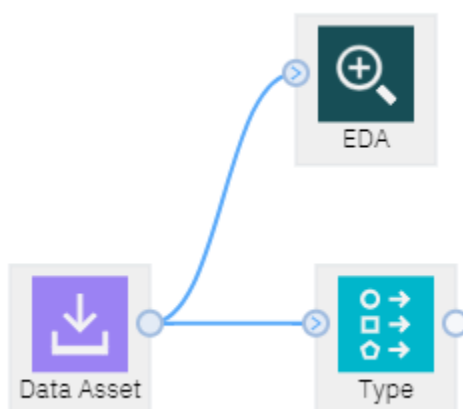
<input type="checkbox"/>	Field name ^	Data type ^
<input type="checkbox"/>	CreditScore	<input type="checkbox"/> integer
<input type="checkbox"/>	Geography	<input checked="" type="checkbox"/> string
<input type="checkbox"/>	Gender	<input checked="" type="checkbox"/> string
<input type="checkbox"/>	Age	<input type="checkbox"/> integer
<input type="checkbox"/>	Tenure	<input type="checkbox"/> integer
<input type="checkbox"/>	Balance	<input type="checkbox"/> double
<input type="checkbox"/>	NumOfProducts	<input type="checkbox"/> integer
<input type="checkbox"/>	HasCrCard	<input type="checkbox"/> integer
<input type="checkbox"/>	IsActiveMember	<input type="checkbox"/> integer
<input type="checkbox"/>	EstimatedSalary	<input type="checkbox"/> double
<input type="checkbox"/>	Exited	<input type="checkbox"/> integer

Observará como se seleccionan todas las casillas. De click en “OK” y luego en “Save”. De click derecho al bloque de “EDA” y luego “Run”. Se deberá abrir el panel derecho en donde observará el siguiente ícono:



De doble click sobre este y observe los resultados.

5. En el panel izquierdo bajo la pestaña “Field operations” arraste el bloque “Type” y conéctelo al bloque de datos.



Doble click sobre el bloque “Type” y bajo la pestaña “Settings” presione en “Configure types”. En la ventana emergente presione en “Read values”. Después de cargar haga las siguientes modificaciones: Para la variable “Exited” cambia el “Role” a “Target” y cambiar su “Measure” a “Flag”. Luego, cambiar el “Measure” de “Gender” a “Nominal”.

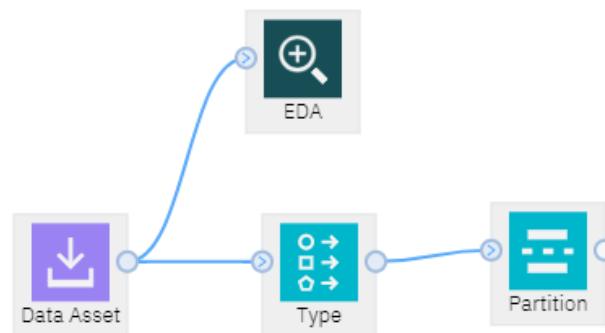
Types

⊖ ⊕ Add Columns

Field^	Measure^	Role^	Value mode^	Values^	Check^
Tenure	Continuous▼	Input ▼	Specify ▼	0, 10	None ▼
Exited	Flag ▼	Target ▼	Specify ▼	0, 1	None ▼
Gender	Nominal ▼	Input ▼	Specify ▼	Female, M...	None ▼
CreditScore	Continuous▼	Input ▼	Specify ▼	350, 850	None ▼

Seleccionar “Target” como “Role” indica que esta será la variable que queremos predecir. La columna “Measure” es para indicar cómo tomar el tipo de dato de las variables. Cuando “Measure” = “Flag” se está indicando que la variable asociada es binaria (Verdadero o Falso). Cuando “Measure”=“Nominal” se está indicando que la variable es categórica con un número definido de categorías.

6. Bajo la misma pestaña del panel izquierdo arrastre el bloque “Partition” y conéctelo después del bloque “Type”.



Doble click sobre el bloque que acaba de arrastrar y bajo la pestaña “settings” asignar el 80% de los datos para entrenar y el 20% para prueba.

Partition

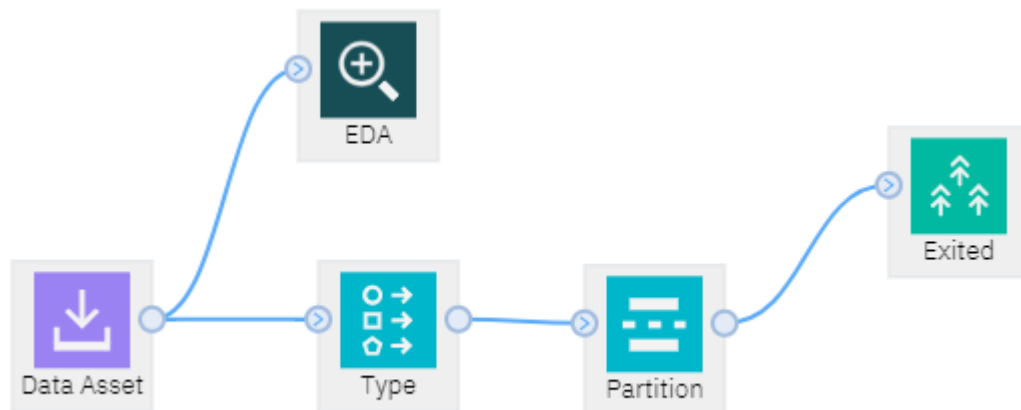
SETTINGS

Derived Field Name
Partition 119

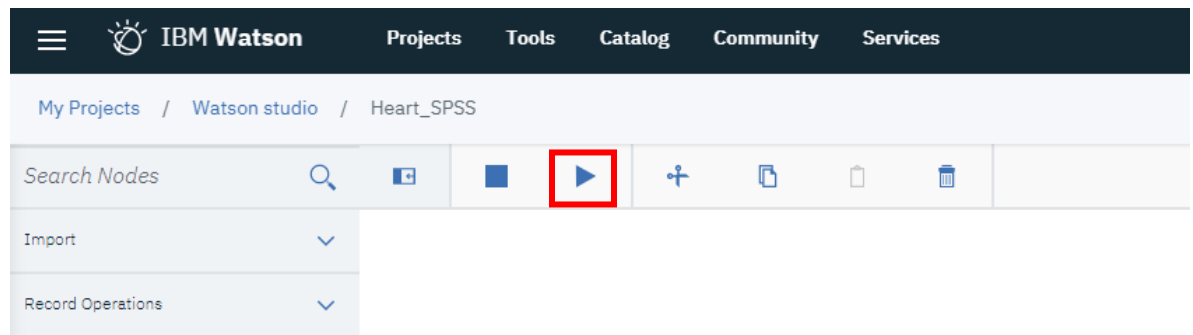
Training Partition
80

Testing Partition
20

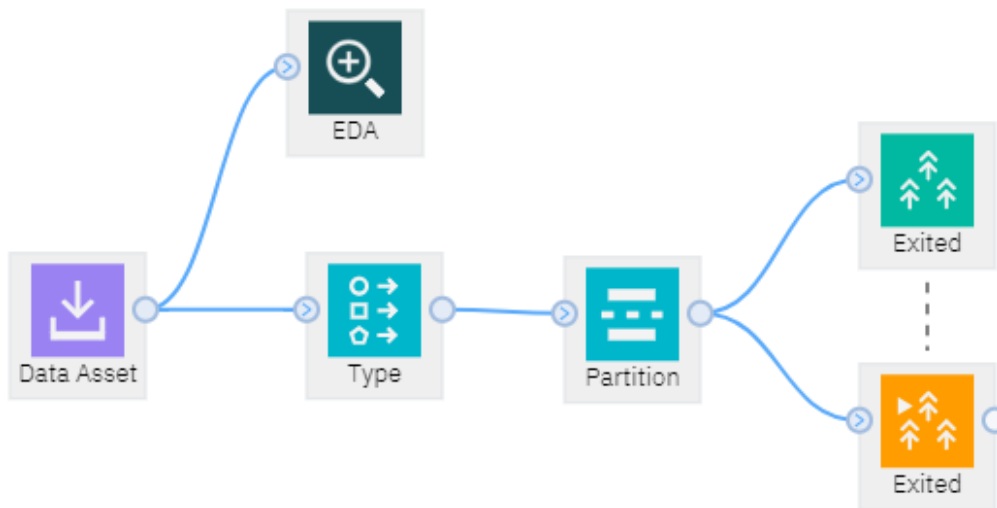
- En el panel izquierdo bajo la pestaña “Modeling” arrastre el bloque “Random trees” y conéctelo al bloque arrastrado en el punto anterior:



- De click en el siguiente ícono para correr el flujo completo:



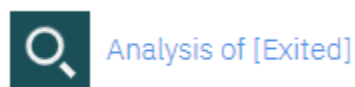
9. Notará que se creó un nuevo bloque:



Este nuevo bloque resume los resultados del modelo. Para observar el modelo damos click derecho "View Model" al bloque de color naranja. Dependiendo del tipo de modelo se observará diferente información.

10. Bajo la pestaña "Outputs" del panel izquierdo arrastre un bloque de "Analysis". Bajo la pestaña "Settings" de seleccione la casilla "Evaluation metric (AUC & Gini, binary classifiers only)". Estos bloques son para ver diferentes métricas pertinentes al desempeño del modelo.

De click derecho "Run" al bloque "Analysis". Se debe abrir el panel lateral donde observará el siguiente ícono. Dar doble click y observe las métricas.



Despliegue del modelo

11. De click derecho sobre el bloque "Analysis of [Exited]" y luego "Save Branch as model". En la nueva ventana nombre el modelo como "Churn_spss_model". Deberá ver seleccionada la instancia de "Machine Learning". De click en save.
12. Salga de SPSS y vaya a su proyecto.