1. Set Role: thiết lập vai trò cho các thuộc tính.
2. input ports:cổng vào: exemple set(exa) là đầu vào cho ExampleSet và là đầu ra của Retrieve điều hành trong suốt ví dụ của tôi. output của các operators cũng có thể sử dụng để làm input. điều quan trọng là meta data phải được gắn với dữ liệu cho input bởi vai trò của 1 thuộc tính là quan trọng trong meta data của ExampleSet. Retrieve operator cũng cấp meta data cùng với dữ liệu.
3. Output Porst: cổng ra ví dụ thiết lập ExampleSet với các kiểu vai trò là đầu ra của cổng này. original (nguyên) ExamplrSet được đưa ra như là đầu vào là quy cách thông qua sự thay đổi để có output. điều này thường được dùng để tái sử dụng ExampleSet trong Operators hoặc hiển thị ExampleSet trong Results Workspace
4. Parameter:

Name: tên thuộc tính muốn thay đổi vai trò.

Target role: vai trò mới được lựa chọn. Có thể có các giá trị sau:

* reguler: Các thuộc tính liên tục không có vai trò đặc biệt, nó chỉ mô tả các ví dụ có thuộc tính liên tục, chỉ có vai trò trong một vài trường hợp. các thuộc tính liên tục được sử dụng như các biến đầu vào trong tập huấn luyện.
* ID: đây là một vai trò đặc biệt, nó hoạt động như id thuộc tính của một ExampleSet và nó thường duy nhất trong mỗi ví dụ của ExampleSet.Vai trò ID được sử dụng để xác định rõ ràng các ví dụ của ExampleSet có liên quan.
* cluster:phan cum co vai tro dac biet no cho thay su phan cum cu the cua mot thanh vien trong mot vi du cu the trong du lieu mau. vi du dau ra của nhà điều hành K-Mean khi them mot cot thi co vai tro phan cum
* Weight : Đây là một vai trò đặc biệt, nó chỉ ra trọng lượng của examples bằng xem cho nhãn hiệu này. Weight được sử dụng trong quá trình huấn luyện để cung cấp sự khác nhau quan trọng giữa các Weight examples. Weight cũng có thể được sử dụng trong đánh giá hiệu quả của các mô hình xem xét trọng lượng của examples trong quá trình đánh giá hiệu suất.
* batch:đây là một vai trò đặc biệt,nó chỉ ra thành viên nào là dữ liệu mẫu.
* lable: (nhãn) Đây là một vai trò đặc biệt, nó hoạt động như một thuộc tính mục tiêu cho các nhà khai thác học. ví dụ :Decision Tree operator xác định nhãn trong bất kỳ ví dụ nào và họ phải dự đoán cho các ví dụ mới mà chưa được đặc trưng theo cách như vậy. Nhãn cũng được gọi là 'biến mục tiêu'.
* prediction: Đây là một vai trò đặc biệt, nó hoạt động thuộc tính như dự đoán của một sơ đồ học. Ví dụ như khi một mô hình dự đoán được học qua bất kỳ nhà điều hành học và sau đó nó được áp dụng sử dụng nhà khai thác Apply Model, trong đầu ra chúng ta có một thuộc tính mới với vai trò dự đoán nắm giữ các giá trị của nhãn được dự đoán của mô hình đã cho. Nhãn và các thuộc tính dự đoán đều được sử dụng để đánh giá hiệu quả của mô hình.

1. Decision Tree: tạo ra cây quyết định để phân lớp cho cả dữ liệu không xác định(nominal) và dữ liệu số.
   1. Input Ports training set : Cổng đầu vào này là ExampleSet . Nó là đầu ra của các nhà điều hành lấy trong quá trình ví dụ kèm theo. Đầu ra của nhà điều hành khác cũng có thể được sử dụng như đầu vào
   2. Output Ports Model: cây quyết định được gửi từ cổng đầu ra này. Mô hình classi-cation này có thể được áp dụng trên dữ liệu vô hình đặt ra cho các dự đoán của các thuộc tính nhãn example set: ExampleSet đã được đưa ra như đầu vào được thông qua với sự thay đổi để đầu ra thông qua cổng này.Điều này thường được sử dụng để tái sử dụng cùng một ExampleSet.
   3. Parameter:
2. Criterion:( tiêu chí chọn )là các thuộc tính được chọn để tách, nó có thể là một trong những giá trị sau:

* information-gain (tăng thông tin): là entropy của tất cả các thuộc tính đc tính toán. các thuộc tính có entropy nhỏ nhất được lấy để phân chia. phương pháp này có xu hướng thiên về việc lựa chọn các thuộc tính với số giá trị là lớn nhất.

- gain\_ratio: là một biểu thức khác của information-gain. nó điều chỉnh các information-gain cho mỗi thuộc tính để thống nhất độ rộng mẫu và giá trị thuộc tính.

- gini index(chỉ số Gini): Đây là thước đo của độ bất tịnh của một tập mẫu. Được tách trên thuộc tính được lựa chọn cho việc giảm chỉ số Gini trung bình của các kết quả tập con.)

- acurracy:Một thuộc tính như vậy được chọn để phân chia tối đa hoá sự chính xác của toàn bộ cây.

b. No prepruning (boolean):Theo mặc định, cây quyết định được tạo ra với prepruning. Đặt tham số này để vô hiệu hóa prepruning và cung cấp một cây mà không cần bất kỳ prepruning. không cắt tỉa (boolean) Theo mặc định, cây quyết định được tạo ra với cắt tỉa. Đặt tham số này để vô hiệu hóa sự cắt tỉa và cung cấp một cây unpruned.

c. Minimal size for split: Kích thước của một nút là một số ví dụ trong tập hợp con của nó. Kích thước của nút gốc là tương đương với tổng số ví dụ trong ExampleSet. Chỉ những nút được tách ra có kích thước là lớn hơn hoặc bằng kích thước tối thiểu của tham số tách.

d. Minimal leaf size: Kích thước của một nút lá là một số ví dụ trong tập hợp con của nó.Cây được tạo ra theo cách mà mỗi tập con nút lá có ít số lượng kích thước tối thiểu lá trong các trường hợp.

e. Minimal gain (real):Gain của một nút được tính toán trước khi chia nhỏ nó. Các nút được chia nếu Giain của nó là cao hơn mức Gain tối thiểu. Giá trị Gain cao hơn tối thiểu cho kết quả chia tách ít hơn và cho một cây nhỏ. Một giá trị quá cao sẽ hoàn toàn ngăn chặn việc chia nhỏ và một cây với một nút duy nhất được tạo ra.