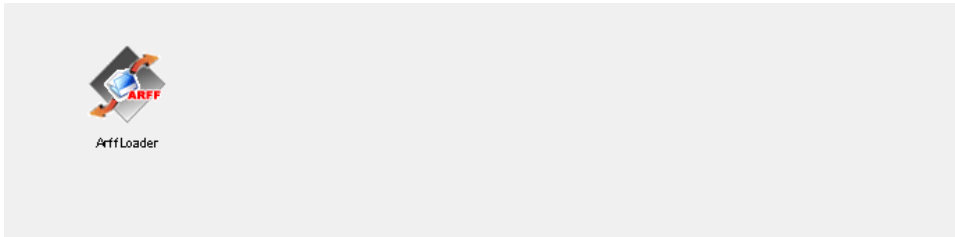


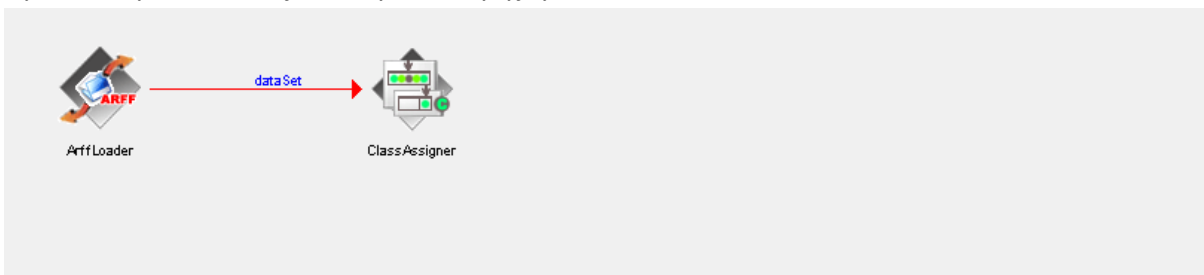
Ocena jakości klasyfikatorów

Niniejsze ćwiczenie demonstruje znaczenie różnych metod oceny jakości klasyfikatorów budowanych przez różne algorytmy, oraz stosowalność różnych algorytmów dla różnych problemów.

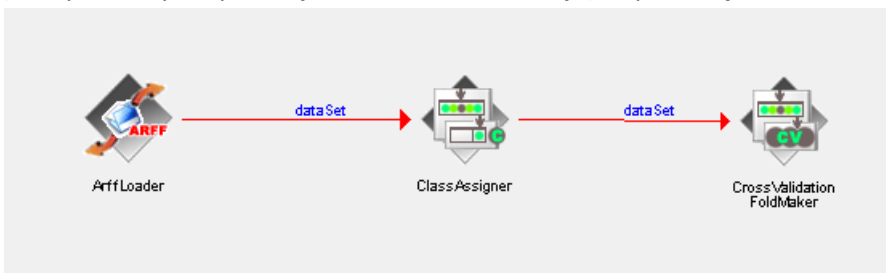
1. Uruchom aplikację „Knowledge Flow”.
2. Wstaw źródło danych „ARFF loader” i skonfiguruj je do odczytu pliku iris.arff.



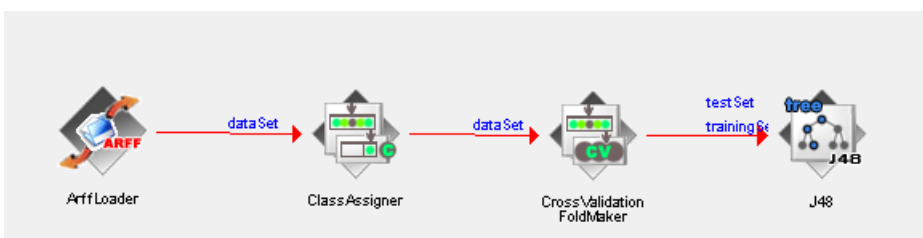
3. Wstaw narzędzie „Class assigner” z zakładki „Evaluation”, połącz je ze źródłem danych za pomocą „data set” i wybierz atrybut „Class” jako atrybut decyzyjny.



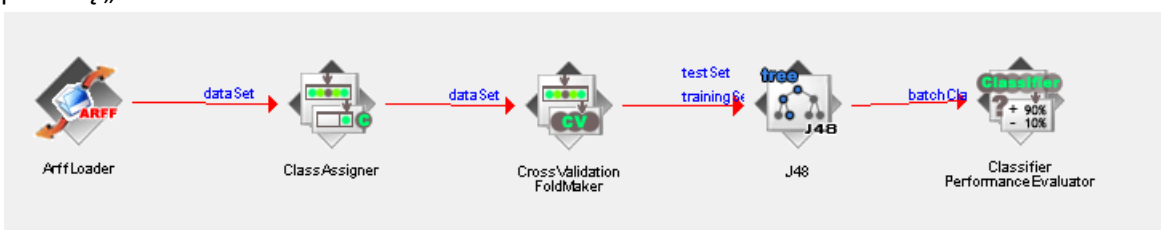
4. Wstaw narzędzie „Crossvalidation Fold Maker” z zakładki „Evaluation” i połącz je z narzędziem „Class Assigner” (domyślnie wykonywana jest 10-krotna walidacja) za pomocą „data set”.



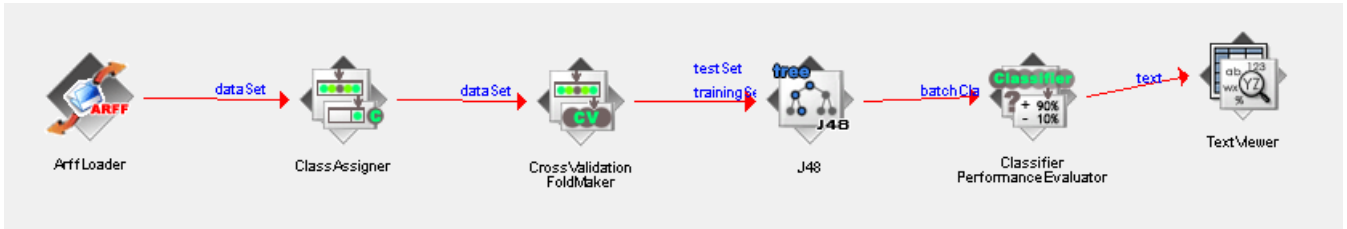
5. Wstaw klasyfikator J48 i połącz go z „Crossvalidation Fold Maker” za pomocą dwóch źródeł: „training set” i „test set”.



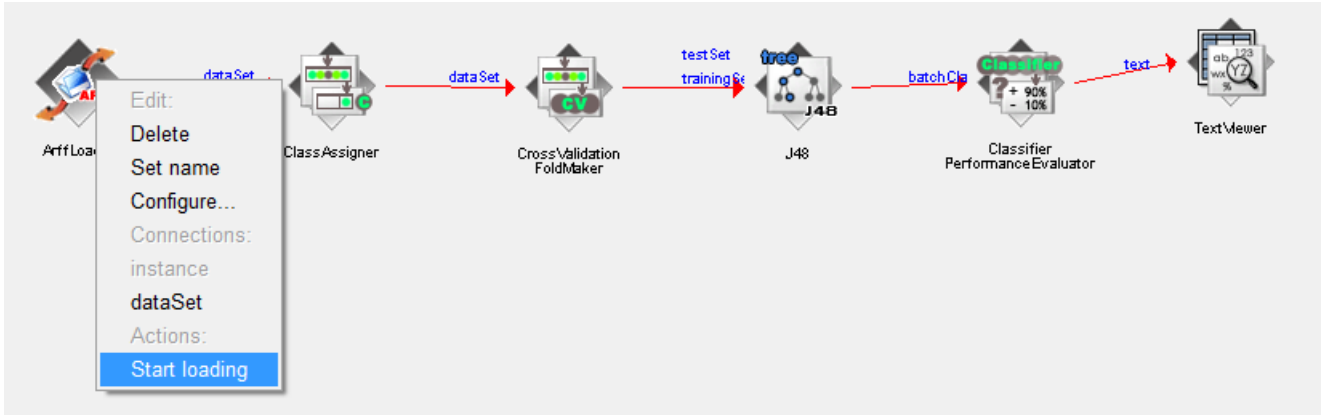
6. Wstaw narzędzie „Classifier Performance Evaluator” z zakładki „Evaluation” i połącz je z klasyfikatorem J48 za pomocą „batch classifier”.



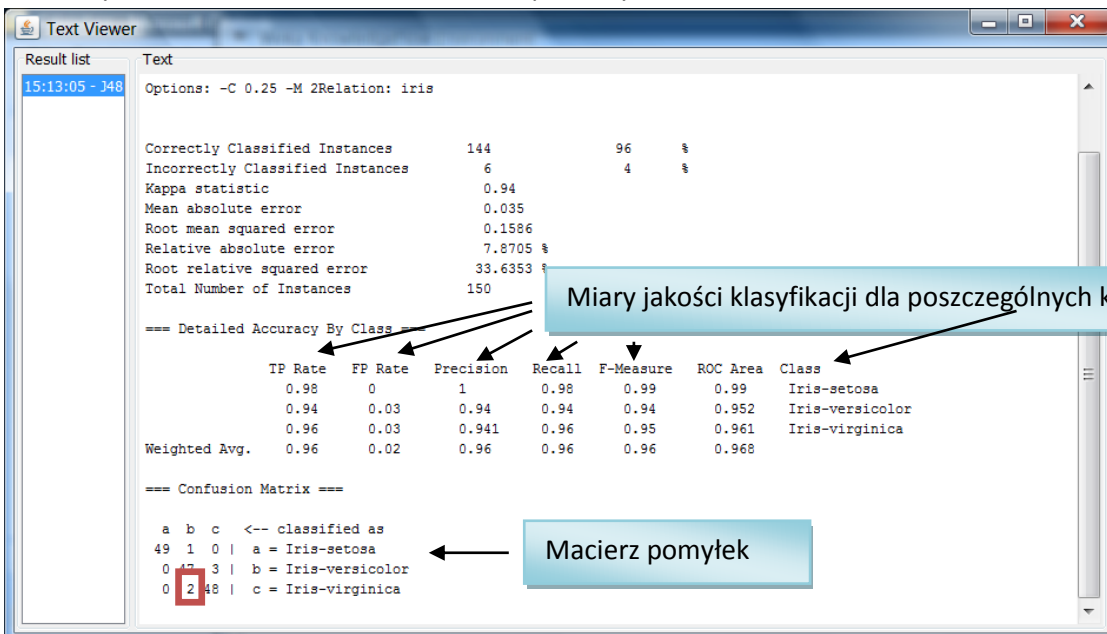
7. Wstaw „Text Viewer” z zakładki „Visualization” i połącz z „Classifier Performance Evaluator” za pomocą „text”.



8. Uruchom proces za pomocą opcji „Start loading” źródła danych.



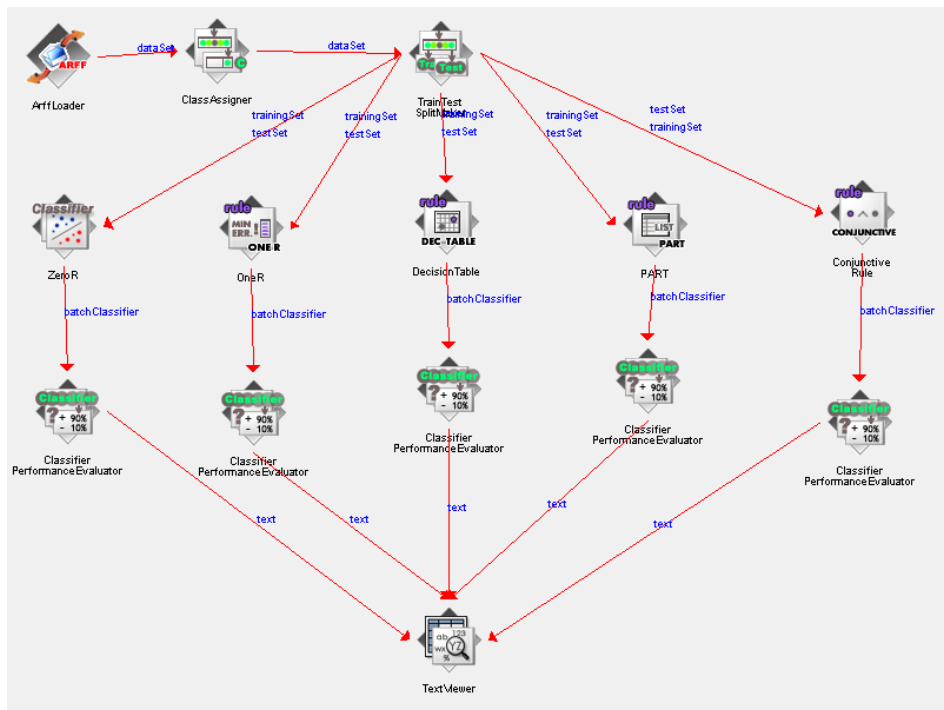
9. Obejrzyj rezultat oceny klasyfikatora za pomocą polecenia „Show Results” narzędzia „Text Viewer”. Zastanów się co znaczy liczba zaznaczona na czerwono na poniższym screenshocie.



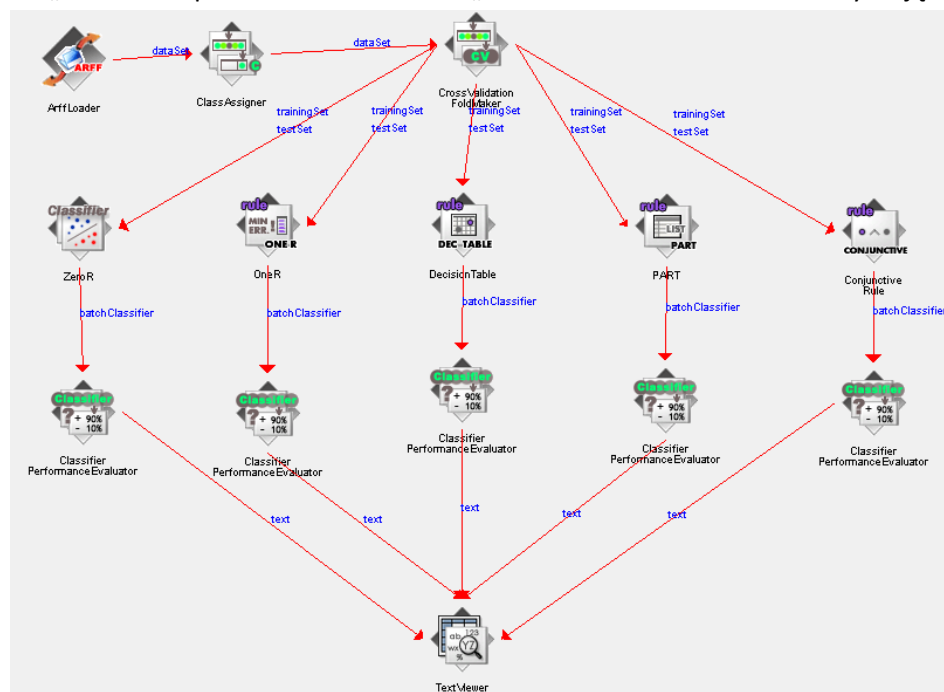
10. Zwróć uwagę na wyniki „Correctly Classified Instances”, „Incorrectly Classified Instances”, czy też „Confusion Matrix”. Jak łatwo zauważyć dotyczą one wszystkich 150 opisów kwiatów znajdujących się w pliku. Zastanów się, który z 10 klasyfikatorów powstałych podczas procedury walidacji skrośnej został użyty do wygenerowania tych wyników na całym zbiorze danych (pytanie podchwytliwe)?
11. Obejrzyj drzewo decyzyjne wygenerowane przez J48 dla zbioru iris.arff. Można je uzyskać podłączając element „Graph Viewer” do elementu J48 za pomocą przepływu „graph”. Jaka jest cecha charakterystyczna irysu setosa?
12. Obejrzyj drzewa decyzyjne wygenerowane przez J48 dla zbioru zoo.arff. Powiedz, czy jesteś w stanie na podstawie jednego z powyższych drzew zaproponować zbiór reguł, które pozwalają odróżnić poszczególne rodzaje zwierząt i które będą wymagać mniej testów logicznych niż zaproponowane drzewa?

13. Ocena klasyfikatorów dla zestawów danych o atrybutach nominalnych.

- a. Zbuduj proces konstrukcji i oceny klasyfikatorów ZeroR, OneR, DecisionTable, PART i Conjunctive Rule wykorzystujący prosty podział zbioru danych, wzorując się na poniższym rysunku. Podziel dane następująco: 1/3 na zbiór testujący i 2/3 na zbiór uczący.



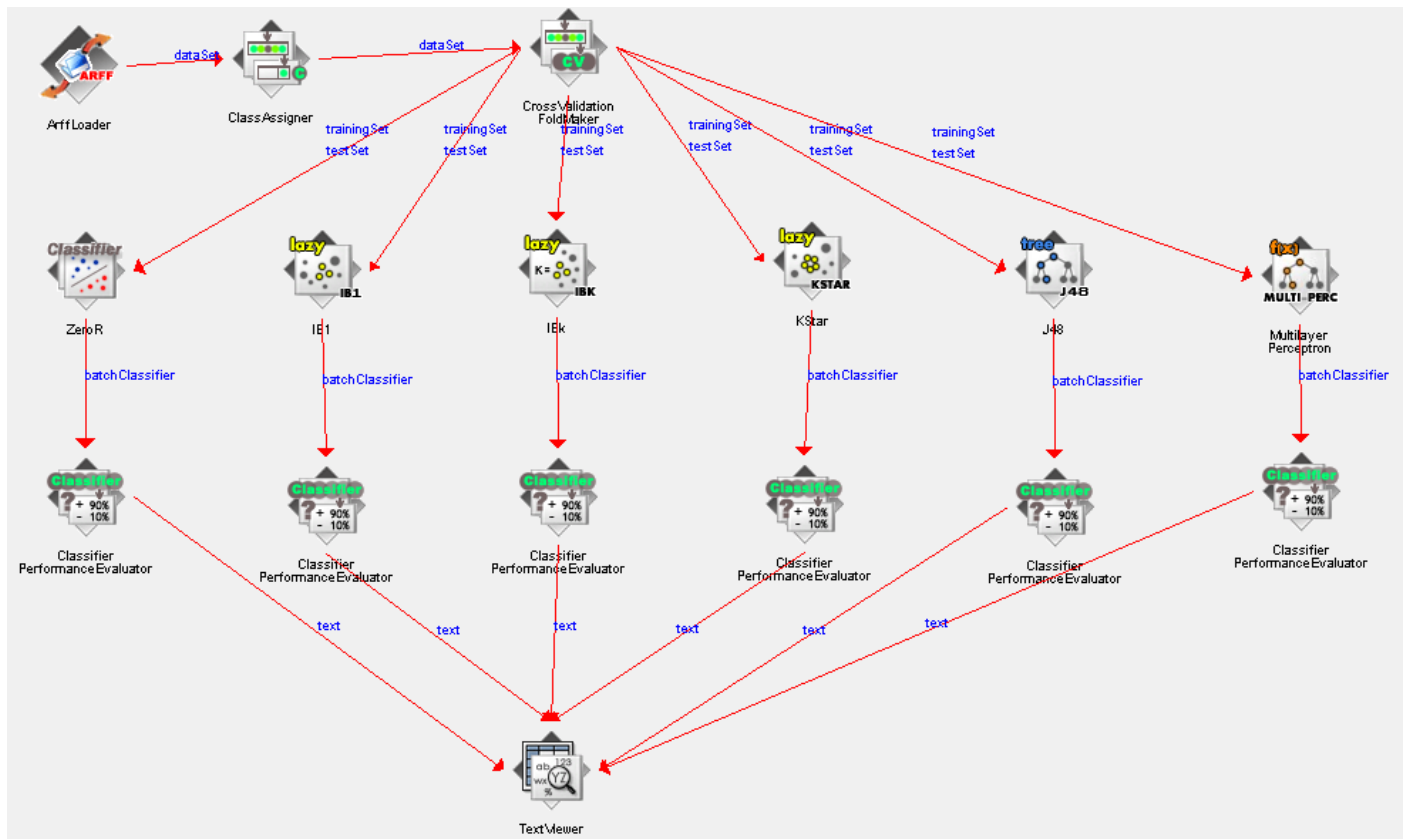
- b. Wygeneruj oceny klasyfikatorów dla zbiorów weather.nominal.arff, car.arff i zoo.arff, labor.arff oraz vote.arff.
 c. Porównaj uzyskane oceny. Który algorytm klasyfikacji jest najlepszy? Czy ranking algorytmów jest zawsze taki sam, dla każdego zestawu danych?
 d. Zastąp element „TrainTest SplitMaker” elementem „CrossValidation FoldMaker” uzyskując poniższy diagram:



- e. Wygeneruj oceny klasyfikatorów dla zbiorów weather.nominal.arff, car.arff i zoo.arff, labor.arff oraz vote.arff.
 f. Porównaj uzyskane oceny. Który algorytm klasyfikacji jest najlepszy? Czy rankingi uzyskane za pomocą crosswalidacji i podziału zbioru są takie same czy różne?

14. Ocena klasyfikatorów dla zestawów danych o atrybutach ciągłych.

- a. Zbuduj proces konstrukcji i oceny klasyfikatorów Multilayer Perceptron, IB1, IBk, KStar, J48 i ZeroR wzorując się na poniższym rysunku.



- b. Zmodyfikuj domyślne ustawienia „Multilayer Perceptron” na następujące:

- training Time=100,
- validationSetSize=10,
- validationThreshold = 10.

- c. Zmodyfikuj ustawienia IBk na następujące:

- KNN=10.

- d. Wygeneruj oceny zbudowanych klasyfikatorów dla zbiorów iris.arff i spam.arff. Uwaga! W przypadku tego drugiego zbioru będzie trzeba trochę poczekać na wyniki (ok. 10 minut).

- e. Porównaj uzyskane oceny. Który algorytm klasyfikacji jest najlepszy? Czy ranking algorytmów jest zawsze taki sam, dla każdego zestawu danych?