1. **Wymagania**
   1. **Narzędzie powinno umożliwiać wczytanie algorytmów decyzyjnych w dwóch tekstowych formatach:**

- z warunkami zapisanymi w postaci słabych nierówności w przypadku wartości rzeczywistych atrybutów

- z warunkami zapisanymi w postaci równości w przypadku wartości nominalnych atrybutów

* 1. **Algorytmy powinno dać się przeglądać przy różnych parametrach wyświetlania i każdą z wyświetlanych wersji powinno dać się zapisać w postaci pliku**

- pokazanie tylko reguł z warunkami na określone atrybuty

- pokazanie reguł o określonych wartościach wsparcia

- pokazanie reguł o określonej długości

- pokazanie reguł o określonych wartościach miary zdefiniowanej w oparciu o parametry reguł (takich jak wsparcie, długość, uwzględnione w nich atrybuty)

- we wszystkich kombinacjach tych parametrów

* 1. **Algorytmy powinno dać się testować dla dowolnego wczytanego pliku testowego (lub plików – więcej niż jednego na raz bez konieczności oddzielnego definiowania parametrów testu) o określonym formacie, z możliwością generowania serii testów oraz dla określonych parametrów testowania**

- wyniki testów pokazane poprzez podanie etykiet przypisanych w ramach klasyfikacji dla poszczególnych przykładów lub poprzez podanie macierzy błędu (? confusion matrix)

- testowanie powinno umożliwiać różne podejście do konfliktów: z głosowaniem ważonym, z głosowaniem większościowym, z odrzuceniem konfliktów (traktowanie jak zła klasyfikacja)

- wszystkie wyniki powinno się dać zapisywać do plików tekstowych

* 1. **Serie testów to wyniki klasyfikacji dla algorytmów powstających w punkcie 2 przy tych różnych warunkach przeglądania**

- przy dodaniu ograniczenia na wielkość wsparcia reguł, z przejściem przez cały zakres wartości tego wsparcia,

- to samo względem długości reguł,

- z ograniczeniem reguł do tych odnoszących się do określonych atrybutów, zorganizowanych w postaci uporządkowanej listy, w kolejnych krokach liczba atrybutów się zwiększa, a przez to liczba reguł też

- względem wartości zdefiniowanych wag

* 1. **Jeśli warunki występujące w regułach odwołują się do liczb rzeczywistych, powinno dać się je modyfikować poprzez zastosowanie danych z wczytanego pliku z wynikami dyskretyzacji**

- z zastąpieniem wartości rzeczywistej oznaczeniem numerycznych przedziału, w którym ta wartość występuje

- z zastąpieniem przez wartości brzegowe tego przedziału (z uwzględnieniem zwrotu nierówności)

- z zastąpieniem przez wartość średnią przedziału

* 1. **Wyniki testów zapisane czy wygenerowane powinno się dać agregować dla tego samego pliku testowego, czyli uzyskać na podstawie zdefiniowanego podejścia do konfliktów (głosowanie zwykłe, ważone, z odrzuceniem konfliktów) ostatecznego rezultatu klasyfikacji każdej próbki przez kilka algorytmów decyzyjnych, w postaci etykiety i w postaci macierzy błędu**
* **ITERACJA 1, 29.07 – 12.08.2017**

**CELE:**

1. Struktura projektu: 29.07.2017
2. Struktury danych dla przechowywania reguł oraz zbiorów danych
3. Parser plików z regułami dla formatu RSES