

Przetwarzanie i rozpoznawanie obrazów

Wstęp

**Dr hab. Inż. Zapotoczny Piotr, prof. UWM
Uniwersytet Warmińsko-Mazurski w Olsztynie**

❖ **Barwa**

❖ **Geometria**

❖ **Tekstura**

powierzchni

❖ **Eksploatacja danych**

Eksploracja danych

Eksploracja danych i terminy pokrewne



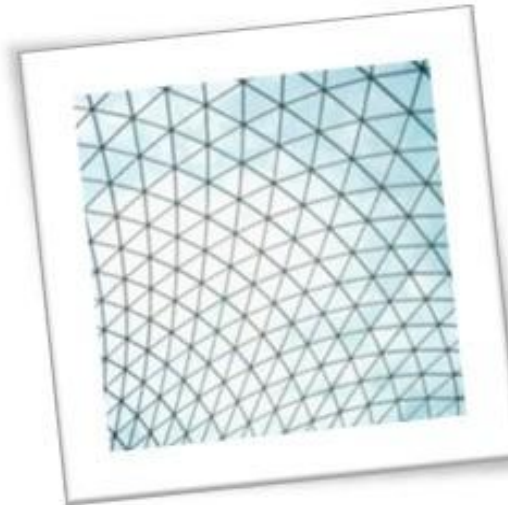
Data mining



Knowledge discovery



Machine learning



Pattern recognition

Pojęcie wiedzy

Encyklopedia Powszechna PWN

- ogół wiarygodnych informacji o rzeczywistości wraz z umiejętnością ich wykorzystywania.

Wiedza a priori

- jest niezależna od zmysłów i dotyczy prawd "absolutnych" lub uniwersalnych jakimi są prawa logiki, prawa matematyki.

Wiedza a posteriori

- jest wiedzą nabytą poprzez zmysły i jej prawdziwość może być obalona poprzez następne obserwacje.

W psychologii w szerokim rozumieniu wiedza to np.

- ogół treści utrwalonych w umyśle ludzkim w wyniku kumulowania doświadczenia oraz uczenia się. (...) W węższym znaczeniu wiedza stanowi osobisty stan poznania człowieka w wyniku oddziaływania na niego obiektywnej rzeczywistości. Często zamiast pojęcia „wiedza” używa się terminu "pamięć".

Czym jest eksploracja danych?

Eksploracja danych obejmuje zespół metod pozwalających na wydobycie **informacji użytecznej** nt. ludzi, przedmiotów bądź zjawisk lub ich reprezentacji (np. obrazów) na podstawie zgromadzonych cech tych obiektów i z zastosowaniem metod **uczenia maszynowego**. Uzyskana informacja ma umożliwiać m.in. klasyfikację danych, aproksymację funkcji, predykcję zachowań. Wynik tych działań ma **wzbogacać wiedzę** o przedmiocie zainteresowania.

Pojęcie wiedzy w uczeniu maszynowym

W badaniach nad sztuczną inteligencją wiedzę traktuje się jako materiał wejściowy albo efekt działania algorytmów sztucznej inteligencji.

Część metod sztucznej inteligencji zajmuje się próbami formalizacji ludzkiej wiedzy celem automatycznego wnioskowania (np. systemy ekspertowe).

Znane jest też odwrotne podejście. Dzięki zastosowaniu tych metod w dużych zbiorach danych można odnaleźć wiedzę o ich naturze (np. sieci neuronowe, algorytmy genetyczne, ekstrakcja cech).

Eksploracja danych



Czym różnią się te obrazy?
Jaka regularność występuje w powyższym zestawie?

Ze względu na kształt

Ze względu na kolor



Co to są dane?

Dane tworzą opis obiektów rzeczywistości wyrażony za pomocą ich cech (atrybutów, parametrów).

Obiektami mogą być ludzie, wszelkie przedmioty lub zjawiska fizyczne oraz ich reprezentacje, np. obrazy.

Cechy jakościowe to np. kolor, smak, stan pogody; pochodzą najczęściej z obserwacji.

Cechy ilościowe to np. waga, rozmiar, wartość funkcji; pochodzą najczęściej z pomiaru oraz obliczeń.

Zbiór cech danego obiektu stanowi odpowiadający mu **wektor danych**.

Kolekcja wektorów danych wyznaczonych dla pewnej serii obiektów stanowi **zbiór danych**.

Czym jest uczenie się maszyn?

Uczenie się systemu

- Każda **autonomiczna zmiana** w systemie zachodząca na podstawie **doświadczeń**, która prowadzi do **poprawy** jakości jego działania.

zmiana — każdy uczący się system zmienia się w czasie;

poprawa – zwiększenie skuteczności, sprawności systemu biorąc pod uwagę preferencje konstruktora lub użytkownika lub obu na raz.

autonomiczność — system, który się uczy, *sam* zmienia się na lepsze (a nie jest zmieniany przez kogoś).

doświadczenia —korzystna zmiana następuje pod wpływem czynników zewnętrznych, które można traktować jako doświadczenia zdobywane przez system.

Motywacja do uczenia się maszyn

Złożoność środowiska

- Złożone zadania nie posiadają prostych rozwiązań
- Praca w warunkach niedookreślonych
- Trudność w sformułowaniu modelu rzeczywistości

Potrzeba autonomii

- Programy uczące się posiadają zdolność adaptacji do zmiennych warunków środowiska pracy.

Potrzeba automatyzacji

- Bazy danych zawierają ogromne zasoby informacji.
- Wydobycie z nich informacji użytecznej poprzez zwykłe przeszukiwanie stanowi bardzo czasochłonne zadanie.

Zadania eksploracji danych

Klasyfikacja

Przypisywanie wektorów danych do kategorii

Regresja

Aproksymacja funkcji rzeczywistej

Predykcja

Przewidywanie klasy, zmiany stanu systemu lub trendu

Modelowanie

Matematyczny model zjawisk i obiektów

Wizualizacja

Prezentacja wektorów danych wielowymiarowych

Optymalizacja

Inteligentne rozwiązywanie skomplikowanych zagadnień matematycznych (polioptymalizacyjnych, zwłaszcza w przestrzeni wielowymiarowej)

Przeszukiwanie

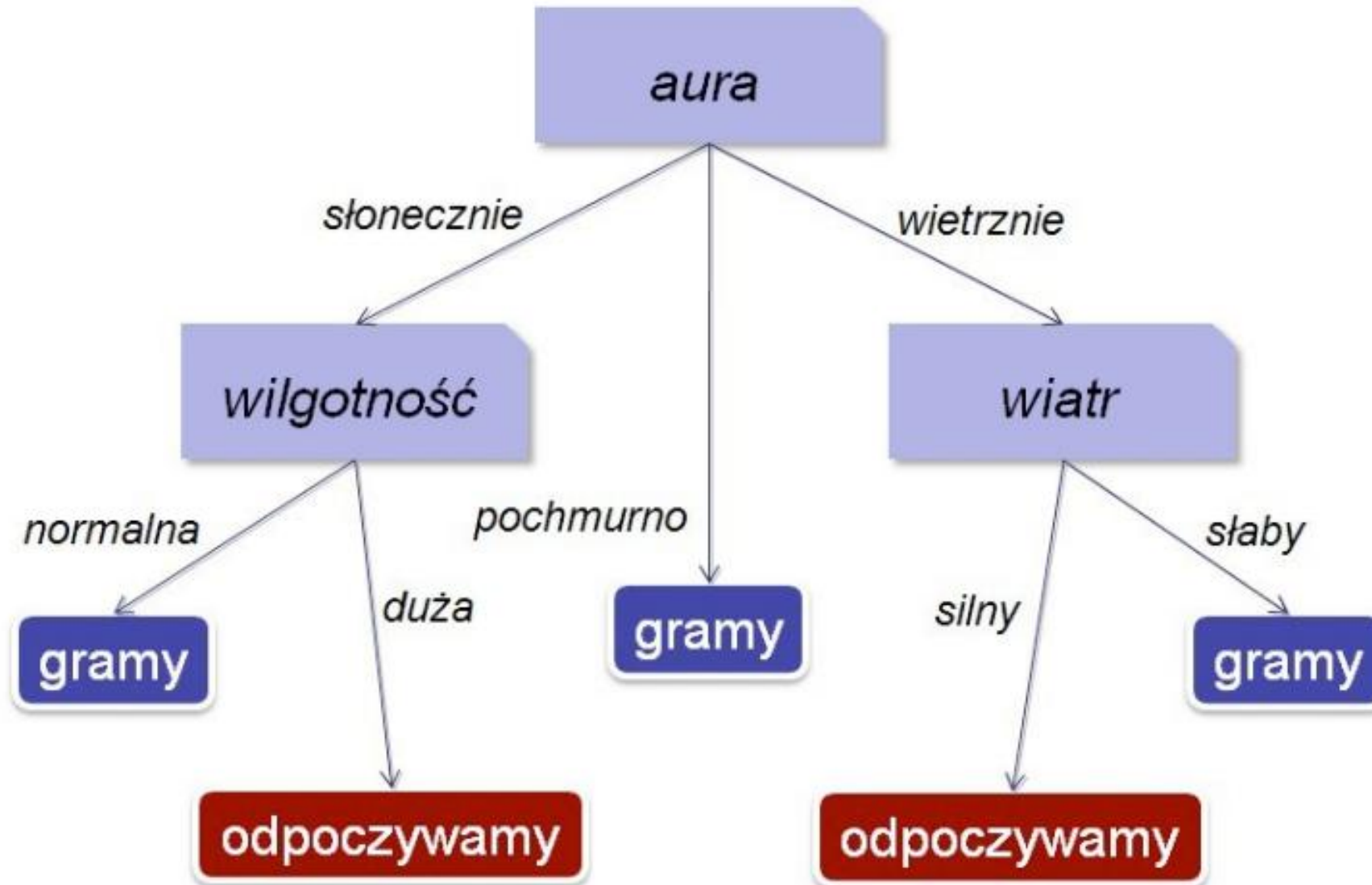
Wydobywanie informacji

Odnajdywanie właściwych dokumentów lub obrazów

Metody reprezentacji wiedzy — tabele decyzyjne

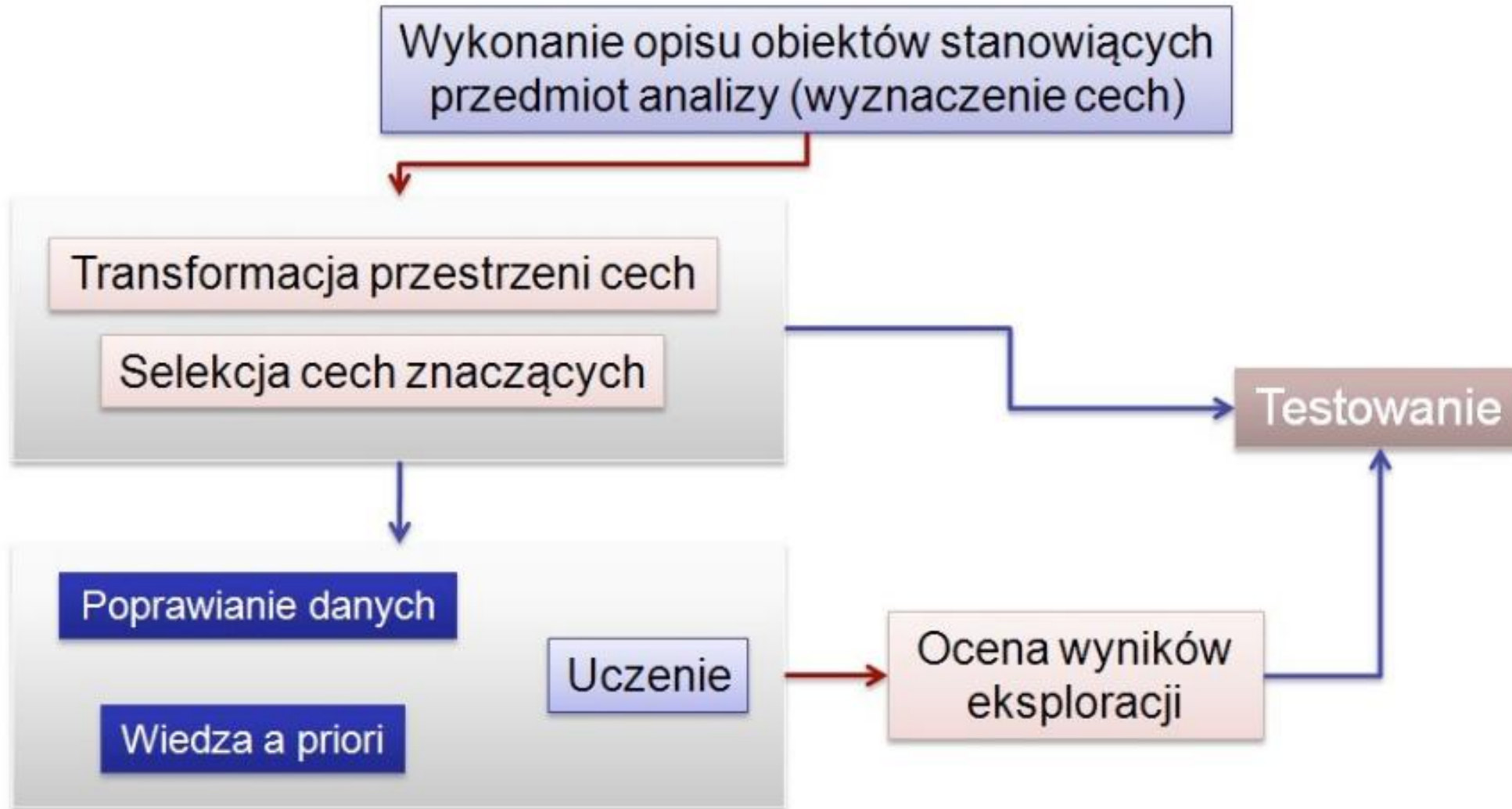
<i>x</i>	<i>aura</i>	<i>wilgotność</i>	<i>wiatr</i>	<i>decyzja</i>
1	słoneczna	duża	słaby	odpoczywamy
2	słoneczna	duża	silny	odpoczywamy
3	pochmurna	duża	słaby	gramy
4	deszczowa	duża	słaby	gramy
5	deszczowa	normalna	słaby	gramy
6	deszczowa	normalna	silny	odpoczywamy
7	pochmurna	normalna	silny	gramy
8	słoneczna	normalna	silny	gramy
9	słoneczna	normalna	słaby	gramy
10	pochmurna	duża	silny	gramy

Metody reprezentacji wiedzy — drzewa decyzyjne



Eksploracja danych

Etapy eksploracji danych



Formy uczenia się

Wykład



Nauczyciel przekazuje wiedzę uczniowi

Laboratorium



Uczeń wykorzystuje wiedzę z wykładu do rozwiązywania zadań i sam uczy się nowych zjawisk

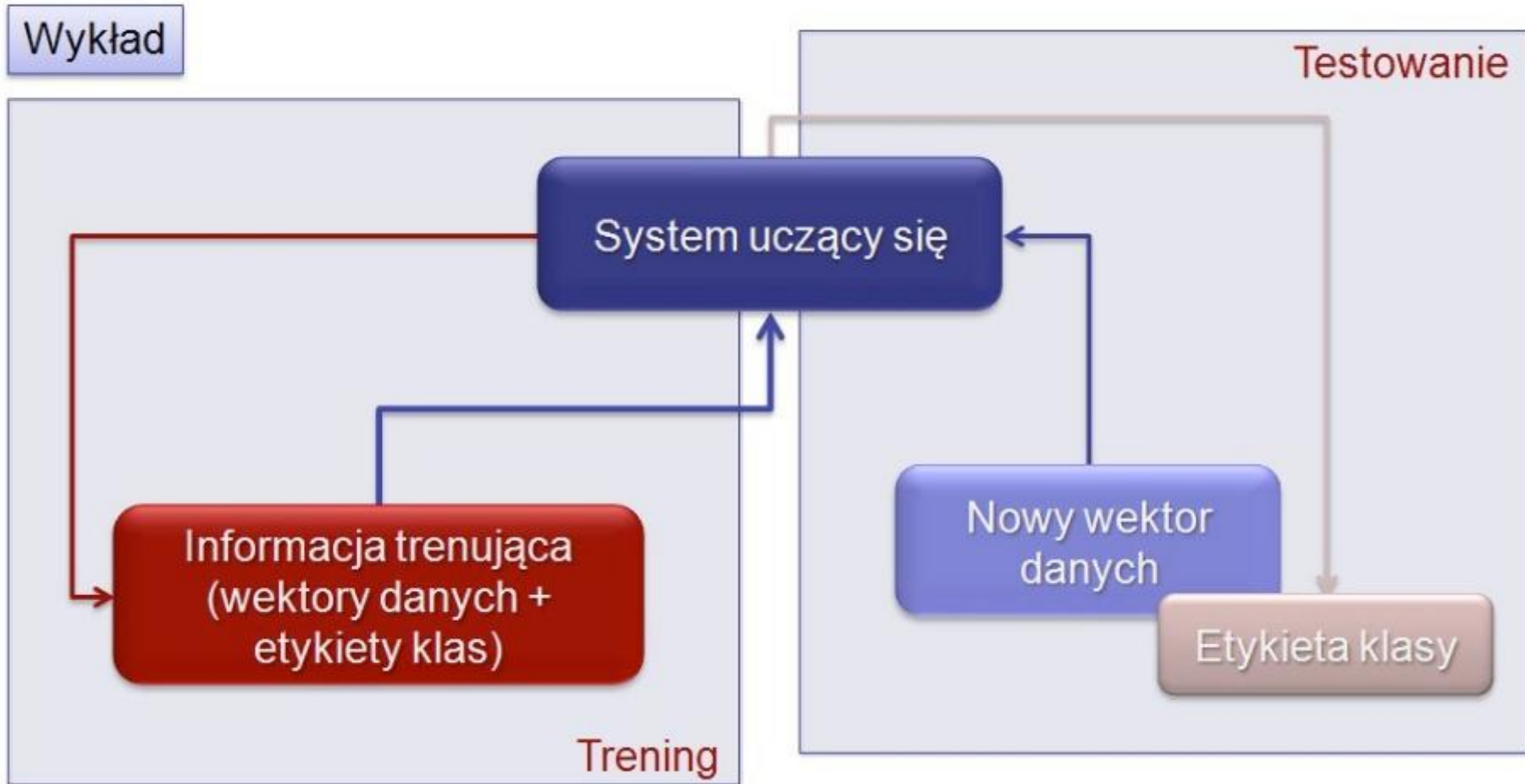
Projekt



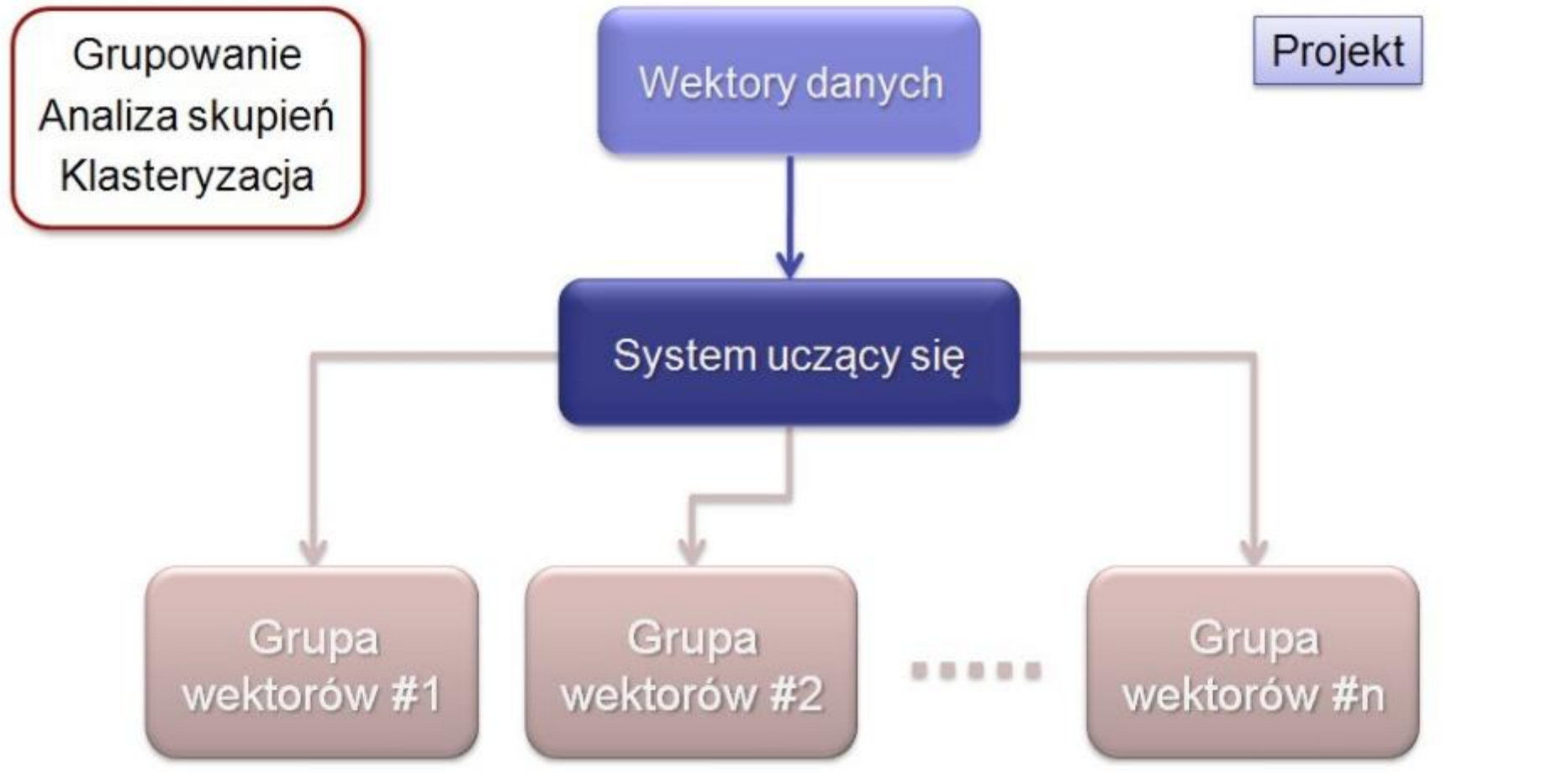
Uczeń samodzielnie zdobywa wiedzę

Eksploracja danych

Uczenie nadzorowane



Uczenie nienadzorowane



Teoria decyzji

Uczenie maszynowe pozwala **wnioskować**, np. o przynależności wektorów danych do klas

Jak na podstawie informacji uzyskanej na drodze automatycznego wnioskowania **podejmować optymalne decyzje** dotyczące obiektów stanowiących przedmiot badania?

Potencjalne problemy przy podejmowaniu decyzji:

Różne typy błędów mogą pociągać za sobą różne konsekwencje.

Różne, niezależne cechy obiektów mogą prowadzić do odmiennych wniosków.

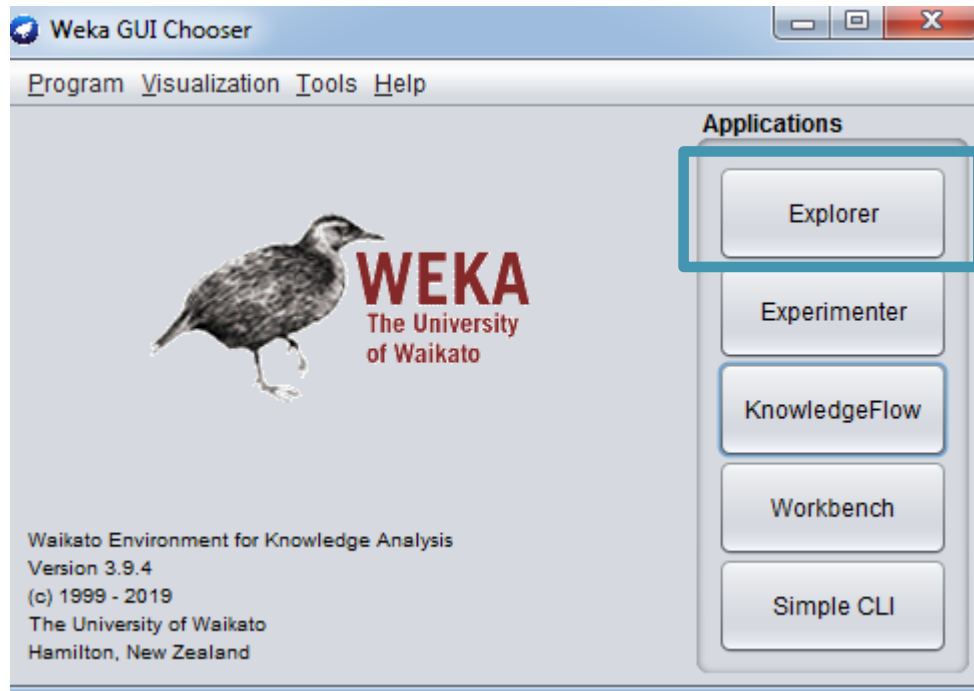
Mała próba wektorów danych może być niewystarczająca, aby predykcja była wiarygodna.

Program WEKA



<https://www.cs.waikato.ac.nz/~ml/weka/>

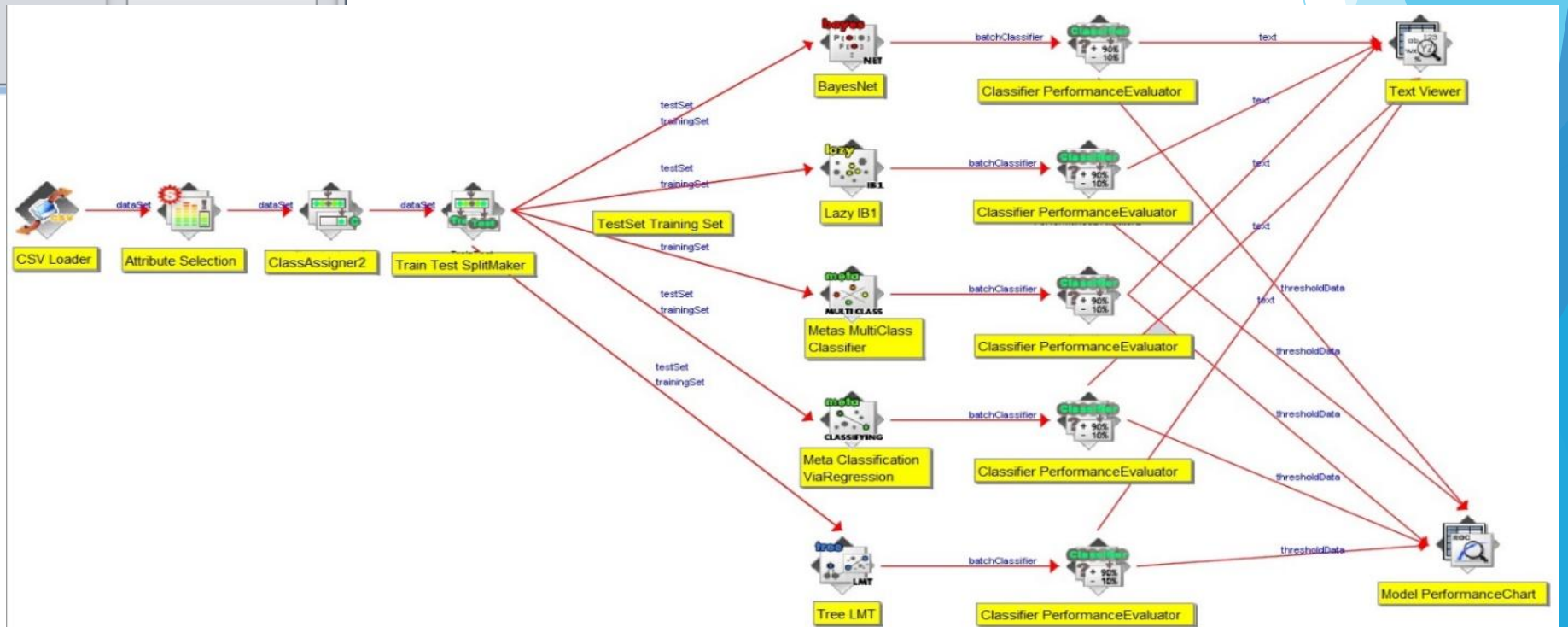
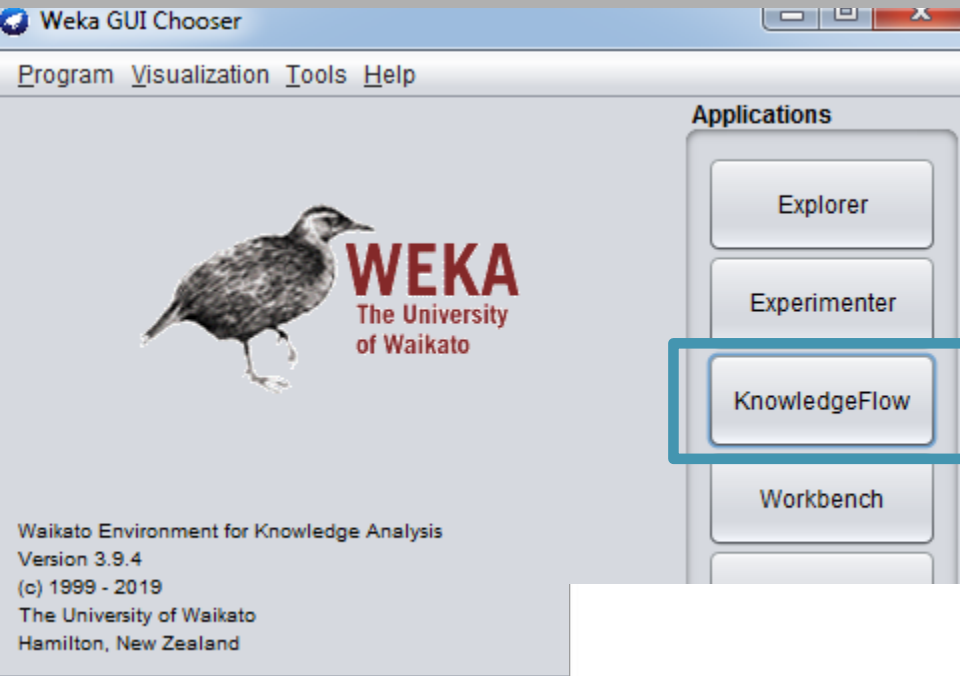
Eksploracja danych



The screenshot shows the 'Weka Explorer' window. At the top, there are several tabs for different visualization and analysis methods: 'Interactive Parallel Coordinates Plot', 'Visualize3D', 'Projection Plot', 'RVines', 'CPython Scripting', 'Preprocess', 'Classify', 'Cluster', 'Associate', 'Select attributes', 'Visualize', and 'Auto-WEKA'. Below these are buttons for 'Open file...', 'Open URL...', 'Open DB...', 'Generate...', 'Undo', 'Edit...', and 'Save...'. A 'Filter' section has a 'Choose' button and a dropdown set to 'None'. The 'Current relation' section shows 'Relation: Zbiorowka_G' with 'Attributes: 31' and 'Instances: 304'. The 'Selected attribute' section shows 'Name: Category', 'Missing: 0 (0%)', 'Distinct: 3', and 'Type: Nominal Unique: 0 (0%)'. Below this is a table with 3 columns: 'No.', 'Label', 'Count', and 'Weight'. The 'Attributes' section has buttons for 'All', 'None', 'Invert', and 'Pattern', and a list of 16 attributes with checkboxes. The 'Class: Category (Nom)' dropdown is set to 'Visualize All'. At the bottom, there are three vertical bars representing the distribution of the 'Category' attribute: a blue bar for 'Zdrowe_G' (count 102), a red bar for 'Grzyb_G' (count 100), and a cyan bar for 'Zadeszczone_G' (count 102). The 'Status' bar at the bottom shows 'OK' and a 'Log' button.

No.	Label	Count	Weight
1	Zdrowe_G	102	102.0
2	Grzyb_G	100	100.0
3	Zadeszczone_G	102	102.0

Eksploracja danych

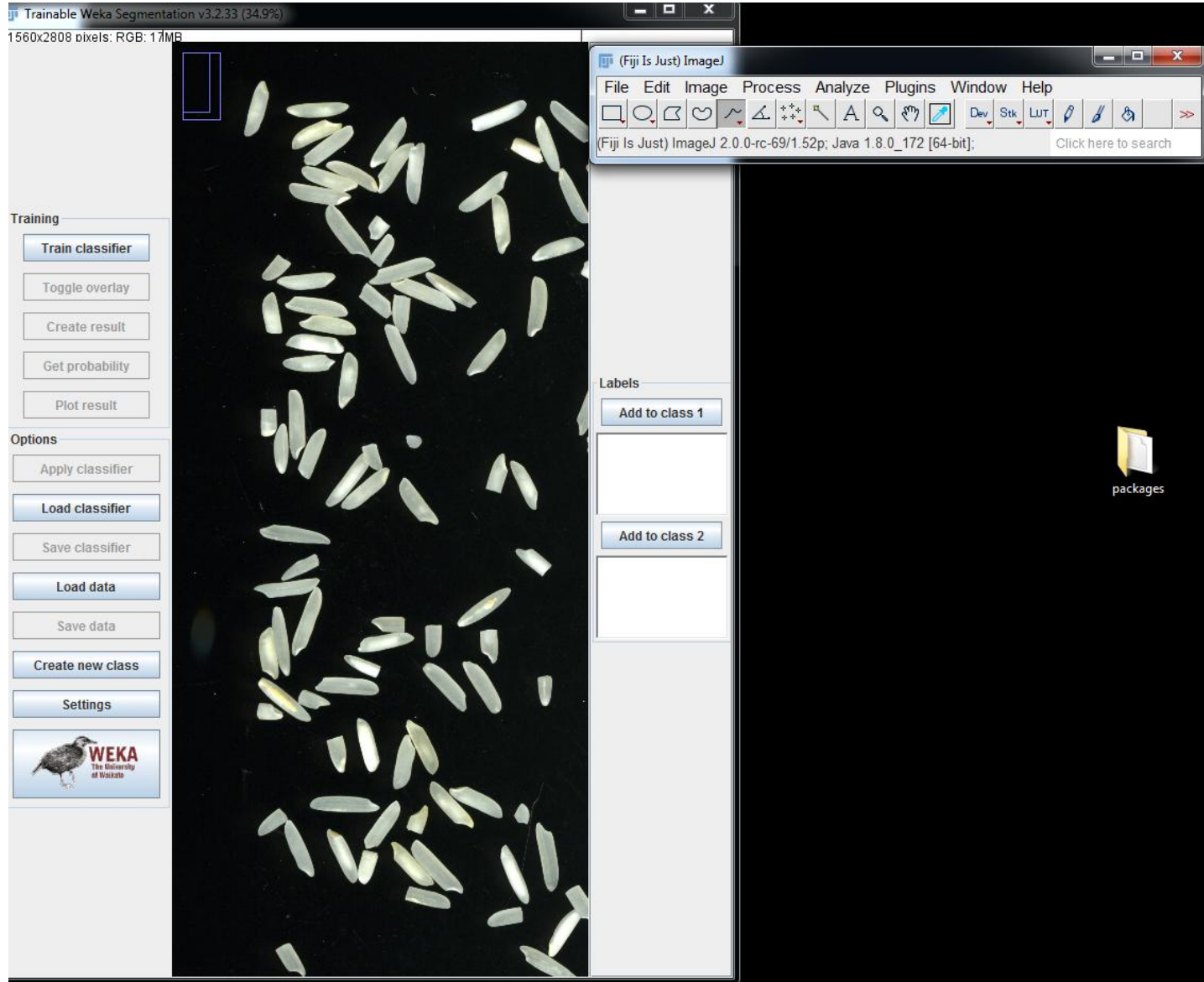


Program ImageJ



<https://imagej.net/Welcome>

Treunable Weka Segmentation



Dziękuję za uwagę