

# Odpowiedzi na pytania do egzaminu z APU

## Wykłady 1–3

### 1 Wykład

#### 1. Definicja procesów uczenia

Nie ma jednej definicji procesów uczenia się:

- “Uczenie się oznacza zmiany w systemie, które mają charakter adaptacyjny w tym sensie, że pozwalają systemowi wykonać za następnym razem takie same zadanie lub zadania podobne bardziej efektywnie” - Herbert Simon (1983)
- “System uczący się wykorzystuje zewnętrzne dane empiryczne w celu tworzenia i aktualizacji podstaw dla udoskonalania działania na podobnych danych w przyszłości oraz wyrażania tych podstaw w zrozumiałej i symbolicznej postaci” - Donald Miche (1991)
- “Uczenie się to konstruowanie i zmiana reprezentacji doświadczanych faktów. W ocenie konstruowanych reprezentacji bierze się pod uwagę:
  - (a) wiarygodność - określa stopień w jakim reprezentacja odpowiada rzeczywistości;
  - (b) efektywność - charakteryzuje przydatność reprezentacji do osiągnięcia danego celu;
  - (c) poziom abstrakcji - odpowiada zakresowi szczegółowości i precyzji pojęć używanych w reprezentacji; określa on tzw. moc opisową reprezentacji. Reprezentacja jest rozumiana jako np. opisy symboliczne, algorytmy, modele symulacyjne, plany obrazy.” - Ryszard Michalski (1986)
- Elementem wspólnym tych definicji są: Wejście (dane empiryczne), miara oceny (Zmiany i poprawa działania) oraz postulat zdobywania wiedzy, reprezentowania jej wewnątrz systemu i stosowania jej do wykonania zadania (nacisk na zrozumiałość reprezentacji)

#### 2. Przykłady problemów rozwiązywanych przez systemy uczące się

- Uczenie się rozpoznawania mowy
- Uczenie się kierowania pojazdem (np. ALVINN)
- Uczenie się klasyfikacji obiektów astronomicznych (NASA Sky Survey)
- Uczenie się rozgrywania pewnych gier
- Uczenie się rozpoznawania chorób na podstawie symptomów
- Uczenie się rozpoznawanie pisma na podstawie przykładów
- Uczenie się klasyfikowania tekstów do grup tematycznych
- Uczenie się aproksymacji nieznanej funkcji na podstawie próbek
- Uczenie się odnajdowania drogi w nieznanym środowisku
- Automatyczne odkrywanie zależności funkcyjnych w danych
- Przewidywanie trendów w danych finansowych

### 3. Motywacje dla budowy systemów uczących się

- Zadania eksploracji i analizy danych, gdzie duże rozmiary zbiorów danych uniemożliwiają ich analizę w sposób nieautomatyczny (np. ekonomiczne lub medyczne bazy danych)
- Środowiska gdzie system musi się dynamicznie dostosowywać do zmieniających się warunków (np. systemy sterowania)
- Problemy które są złożone, trudne do opisu i często nie posiadają wystarczających modeli teoretycznych albo ich uzyskanie jest bardzo kosztowne lub mało wiarygodne.

### 4. Klasyfikacja metod maszynowego uczenia się

- Uczenie indukcyjne - na podstawie znanych faktów i obserwacji tworzona jest uogólnienie, które próbuje dopasować wszystkie znane fakty do hipotezy
- “Nabywanie umiejętności” - optymalizacja zastosowania już posiadanej wiedzy w sposób zwiększający jej efektywność

### 5. Tworzenie modelu uczenia maszynowego

- (a) Przygotowanie danych (zebranie danych, uzupełnienie braków, normalizacja)
- (b) Zdefiniowanie zadania (regresja, klasyfikacja, grupowanie, inne)
- (c) Wybór metody (regresja liniowa, logistyczna, drzewa decyzyjne, sieci neuronowe)
- (d) Strojenie parametrów
- (e) Ocena modelu

### 6. Język R. Wykonywanie instrukcji

Sekwencyjne, możliwe wpisywanie kolejnych poleceń z konsoli lub z pliku skryptu o rozszerzeniu \*.R. W konsoli występuje znak zachęty >. Komentarz tworzony jest przez #. Zapisywanie ma postać `zmienna <- wartość`.

### 7. Korzystanie z pomocy R

- `help(max)` - klasyczna pomoc
- `?max` - skrócona wersja
- `example(max)` - przykłady użycia
- `RSiteSearch("max function")` - przeszukiwanie forum
- `apropos("max", mode = "function")` - wyszukiwanie z funkcji z nazwą "max"
- `data()` - nazwy obiektów w pakiecie
- `vignette()` - pdf z dokumentacją

### 8. Zarządzanie obszarem roboczym R

### 9. Pakiety rozszerzające R

### 10. Skalary i wektory R

### 11. Ramka danych R

### 12. Przegląd wykresów

### 13. Język R i uczenie maszynowe

## 2 Wykład

1. Analiza eksploracyjna i analiza potwierdzająca
2. Czym są dane w uczeniu maszynowym?
3. Wnioskowanie o typach danych w kolumnach
4. Podsumowania liczbowe w R
5. Średnie, mediany i dominanty w R
6. Kwantyle w R
7. Odchylenia standardowe i wariancje w R
8. Eksploracyjne wizualizacje danych
9. Wizualizowanie powiązań pomiędzy kolumnami
10. Klasyfikacja; Zdefiniowanie zadania
11. Trening i testowanie klasyfikacji
12. Kryteria porównawcze metod klasyfikacji
13. Metody klasyfikacji
14. Drzewa decyzyjne
15. Funkcje testu w celu konstruowania drzew decyzyjnych
16. Konstrukcja drzew decyzyjnych
17. Problem brakujących wartości przy konstruowaniu drzew decyzyjnych
18. Analiza ROC jakości klasyfikacji
19. Krzywe ROC
20. Czułość, a specyficzność klasyfikacji binarnej
21. Konstruowanie krzywych ROC
22. Pakiet ROCR

## 3 Wykład

1. Wieloatrybutowe problemy decyzyjne
2. Proces analitycznej hierarchizacji problemu decyzyjnego
3. Kroki rozwiązywania problemu AHP
4. Podstawy wieloatrybutowej teorii użyteczności
5. Agregacja ocen z wykorzystaniem macierzy porównań parami
6. Skala preferencji względnej
7. Ocena spójności macierzy porównań parami
8. Krok V – obliczenie priorytetów AHP

9. Obliczanie przybliżonego wektora własnego macierzy porównań parami
10. Inne metody rozwiązywania problemu AHP
11. Przykłady zastosowań metody AHP