

## 25. Uczenie indukcyjne metody zastosowanie.

### [Wstęp]

Uczenie indukcyjne prowadzi do znajdowania hipotez najlepiej wyjaśniających i uogólniających obserwowane fakty. W przypadku uczenia indukcyjnego fakty te często nazywane są przykładami.

Wyróżniamy następujące rodzaje indukcyjnego uczenia się, które jednocześnie warunkują jego zastosowania:

- **uczenie się pojęć** - pojęcia służą klasyfikacji obiektów do kategorii, np. pojęcie krzesła umożliwia nam wskazanie go pośród innych mebli (klasyfikacja)
- **tworzenie pojęć** - uczeń na podstawie zaobserwowanych przykładów, grupuje je samodzielnie w kategorie zgodnie z pewnymi kryteriami podobieństwa zależnymi od algorytmu (klasteryzacja)
- **uczenie się aproksymacji funkcji** - gdzie zbiorem wartości podlegających uczeniu nie jest zbiór kategorii jak w uczeniu się pojęć, a zbiór liczb rzeczywistych, przykłady trenujące złożone są z argumentu funkcji (wektora) i jej wartości dla tego argumentu (regresja?)

### Tryby uczenia:

- **inkrementacyjny** - przykłady dostarczane pojedynczo, udoskonalenie hipotezy po każdym z nich
- **epokowy** - cykl uczenia podzielony na epoki, w którym hipotezę udoskonala się po przetworzeniu danej liczby przykładów
- **wsadowy** - uczeń otrzymuje wszystkie przykłady, po przetworzeniu których podaje gotową hipotezę, nie można odczytać hipotez w trakcie procesu uczenia. W przypadku zmiany zbioru przykładów uczenie przeprowadza się od nowa.
- **korekcyjny** - odmiana inkrementacyjnego, w którym uczeń najpierw przedstawia swoją odpowiedź na podany przykład, a następnie otrzymuje informację korygującą, powstałą na podstawie różnicy wartości otrzymanej od oczekiwanej

### [Odpowiedz]

**Uczenie pojęć (concept learning)** to pozyskanie definicji ogólnej kategorii mając próbkę pozytywnych i negatywnych przykładów. Uczenie pojęć można zdefiniować jako przeszukiwanie przestrzeni predefiniowanych hipotez w celu znalezienia hipotez najbardziej pasujących do przykładów uczących. Uczenie pojęć nazywane bywa aproksymacją wartości funkcji logicznej z przykładów. Automatyczne wnioskowanie ogólnych definicji pojęć jest postrzegane jako etykietowanie danych przykładów jako należących bądź nie należących do uczonego pojęcia. Zatem, uczenie pojęć może być zdefiniowane jako: Wnioskowanie wartości funkcji logicznej z przykładów zawierających wejścia jak i wyjścia. Reprezentacja hipotezy zaprojektowana w systemie składa się z koniunkcji ograniczeń na wartości atrybutów przykładów.

Generowanie hipotez od-ogółu-do-szczegółu (General-to-Specific Ordering of Hypotheses) pozwala na tworzenie algorytmów uczących bez konieczności specyfikacji explicite wszystkich możliwych hipotez.

- **Uczenie drzew decyzyjnych** (Decision Tree Learning) Jest to popularna metoda wnioskowania indukcyjnego. Służy do aproksymacji funkcji dyskretnej, jest odpowiednia dla zaszumionych danych i ma możliwości uczenia wyrażeń dysjunkcyjnych (alternatyw). Do tej rodziny należą ID3, ASSISTANT, C4.5. Ich 'inductive bias' jest preferencja małych drzew. Uczona funkcja jest reprezentowana jako drzewo decyzyjne. Drzewo decyzyjne może być przedstawione w postaci reguł if-then dla poprawienia jego czytelności.

Każdy węzeł w drzewie decyzyjnym specyfikuje test pewnego atrybutu przykładu, każda gałąź wychodząca z węzła odpowiada jednej możliwej wartości tego atrybutu. Przykłady są klasyfikowane poczynając od korzenia drzewa, testując atrybut określony przez korzeń, poruszając się w dół po drzewie zgodnie z wartościami przyjmowanymi przez atrybuty w węzłach. W ogólności drzewo reprezentuje koniunkcje i dysjunkcje ograniczeń nałożonych na wartości atrybutów.

• **Uczenie Bayesowskie.** Cechy takiego podejścia to:

+Każdy obserwowany przykład może stopniowo zwiększać lub zmniejszać estymowane prawdopodobieństwo poprawności hipotezy. Jest to bardziej elastyczne podejście w stosunku do tego, które usuwa hipotezy niezgodne choć z jednym przykładem.

+Wiedza aprioryczna może być łączona z obserwowanymi danymi do określania końcowego prawdopodobieństwa hipotezy. Wiedza aprioryczna jest dostarczana poprzez (1) prawdopodobieństwo apriori dla każdej kandydującej hipotezy, (2) rozkład prawdopodobieństwa na obserwowanych danych dla każdej możliwej hipotezy.

+Metoda może akumulować hipotezy, które tworzą prawdopodobne przesłanki

+Nowe zadanie może być klasyfikowane przez połączenie przesłanek wielu hipotez, ważonych ich prawdopodobieństwami.

+Nawet, jeśli metody są trudne obliczeniowo, mogą dostarczyć standardów optymalnego podejmowania decyzji dla mierzenia innych metod.

+Problemem jest to, że typowo wymagają one początkowej znajomości wielu prawdopodobieństw (jeśli nie są znane, muszą być estymowane z dostępnych wcześniej danych). Druga trudność to duży koszt obliczeniowy.

• **Instance-Based Learning - Uczenie z przykładów**

W odróżnieniu od metod uczenia, które konstruuja ogólny, jawny opis funkcji docelowej, kiedy dostarczane są dane uczące, instance-based uczenie po prostu zapamiętuje uczące przykłady. Uogólnianie nad tymi przykładami jest odwlekane do czasu, aż nowy przykład (zadanie) ma być klasyfikowane. Za każdym razem, kiedy przychodzi nowe zapytanie (przykład), badane są jego powiązania z zapamiętanymi przykładami aby ustalić wartość

docelowej funkcji nowego przykładu. Metody Instance-Based Learning to najbliższy sąsiad, metody lokalnych ważonych regresji. Zakładają one, że przykład (instance) może być reprezentowany jako punkt w przestrzeni Euklidesowej. Obejmują one również metody Case-Based Reasoning (wnioskowanie na bazie przykładów), które stosują bardziej złożoną, symboliczną reprezentację przykładów. Metody Instance-Based Learning są czasami nazywane "leniwymi" metodami uczenia ("lazy") ponieważ opóźniają one proces przetwarzania aż do czasu, kiedy nowe zdarzenie (instance) musi być klasyfikowane.

Najważniejszą zaletą tego opóźnienia jest to, że zamiast estymować docelową funkcję dla przestrzeni wejściowych przykładów, metody te mogą je estymować lokalnie i oddzielnie dla każdego nowego zdarzenia. Metody z tej grupy to k-Nearest Neighbor Learning oraz Locally Weighted Regression. Metody te nie potrzebują uprzedniego formułowania hipotez. Pozwalają na zamodelowanie bardziej złożonych funkcji docelowych poprzez kolekcję mniej złożonych lokalnych aproksymacji oraz fakt, że żadna informacja obecna w ciągu uczącym nie jest tracona (zapamiętywane są wszystkie przykłady). Podstawowa trudność to efektywność tych metod, trudności z ustaleniem i aproksymacją miary odległości dla odzyskiwania "odpowiednich" przykładów, oraz negatywny wpływ niewłaściwych (irrelevant) cech miary odległości.