

Strojové učenie ako riadený proces

Kristína Machová

Úvod

Strojové učenie ako vedná disciplína sa vyčlenilo z oblasti umelej inteligencie [1]. Medzi strojovým učením [2] a riadiacim procesom existuje niekoľko paralel. Prvá sa týka delenia metód strojového učenia. Tieto metódy delíme na **kontrolované a nekontrolované**. V niektorých zdrojoch sa hovorí priamo o **riadenom a neriadenom učení**. Rozdiel medzi týmito dvomi prístupmi sa dá charakterizovať rozličným spôsobom. Jeden uhol pohľadu charakterizuje kontrolované (riadené) učenie ako **učenie s okamžitou spätnou väzbou**, zatiaľ čo nekontrolované (neriadené) učenie je charakterizované ako **učenie bez spätnej väzby**.

Učiacie algoritmy rozlišujeme aj podľa toho, či majú riešiť učiacu úlohu klasifikačného, alebo sekvenčného typu. Klasifikačná úloha predstavuje rozpoznanie nejakého objektu na základe jeho vlastností (hodnôt atribútov) a následné zaradenie daného objektu do správnej klasifikačnej triedy. Sekvenčná úloha hľadá riešiacu cestu, t. j. postupnosť krokov od známeho počiatočného stavu do známeho cieľového stavu. Príkladom takejto úlohy je šachová hra.

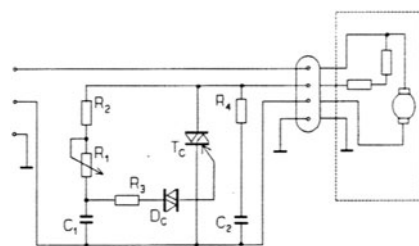
Keď sa teraz vrátíme ku spätnej väzbe, môžeme tvrdiť nasledovne: Pri riešení úlohy klasifikačného typu kontrolovaným učením je spätná väzba daná informáciou o triede v tréningovej množine. Tréningová množina je množinou objektov, ktoré sú charakterizované svojimi vlastnosťami (hodnotami atribútov). Posledný atribút obsahuje informáciu o triede, ak takáto

informácia je dostupná. Príklad takejto tréningovej množiny je uvedený v tab. 1.

Ak informácia o triede v tréningovej množine nie je uvedená (neexistuje spätná väzba), potom ide o učenie nekontrolované, ktoré sa môže realizovať zhľukovaním objektov napríklad na základe podobnosti.

Pri riešení úlohy sekvenčného typu kontrolovaným učením je okamžitá spätná väzba predstavovaná informáciou o vhodnosti vykonania každého kroku. Príkladom spomínaného učenia je napríklad to, keď sa učíme hrať šachy a o každom kroku sa poradíme s učiteľom. Ak sa pri hre nemáme s kým radiť a jediným ohodnotením každého kroku je konečné víťazstvo alebo prehra, ak teda nemáme okamžitú spätnú väzbu, ide o riešenie úlohy sekvenčného typu nekontrolovaným učením, napríklad „učení odmenou a trestom“. Pri „učení odmenou a trestom“ žiadanejšie stavy a akcie, teda častejšie navštevované, kumulujú vyššiu odmenu. Cesta s najvyššími hodnotami odmiern v stavoch, ktoré sa na nej nachádzajú, predstavuje riešenie úlohy.

Uvedená tréningová tabuľka opisuje pozorovania (objekty), ktoré sa vyskytli pri po-



Obr.1 Schéma triakového regulátora vyhrievacej jednotky bez spätnej väzby

ruchách triakového regulátora vyhrievacej jednotky bez spätnej väzby. Schéma regulátora je na obr. 1. Túto tabuľku treba chápať ako ilustračnú. Skutočná tréningová množina by obsahovala tisíce príkladov, pričom aj početnosť výskytov každej triedy by bola omnoho vyššia.

Všetky atribúty v danej tréningovej množine sú dvojhodnotové. Hodnota „a“ znamená, že sa príznak definovaný v názve atribútu vyskytuje pri danej poruche. Hodnota „n“ znamená, že sa nevyskytuje. Jednotlivé atribúty predstavujú nasledovné príznaky:

- A1 prepálená poistka špirály,
- A2 chyba vo vyhrievacej časti,
- A3 chyba ochrany triaku Tc,
- A4 poškodenie triaku Tc,
- A5 problém v triaku Tc,
- A6 triak Tc je zatvorený,
- A7 chyba v regulačnej časti,
- A8 problém v diaku Dc,
- A9 diak Dc je zatvorený,
- A10 skrat v nabíjacom kondenzátore C1,
- A11 na nabíjacom kondenzátore C1 nie je napätie,
- A12 na diaku Dc nie je napätie,
- A13 odpor R1 nie je v poriadku,
- A14 odpor R2 nie je v poriadku,
- A15 odpor R3 nie je v poriadku,
- A16 špirála sa nezohrieva,
- A17 ventilátor nie je v poriadku,
- A18 funguje iba ventilátor.

Daná tréningová tabuľka obsahuje aj informáciu o triede, teda umožňuje riadené učenie. V tejto tabuľke sa vyskytujú nasledovné triedy, t. j. diagnostikované poruchy triakového regulátora:

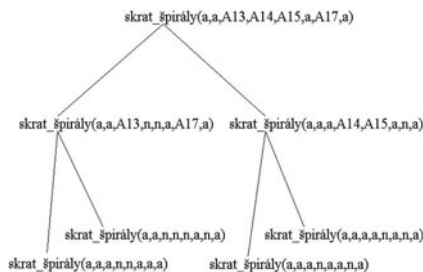
- T1 skrat špirály,
- T2 skrat svorkovnice špirály,
- T3 chybná poistka špirály,
- T4 prerušená špirála,
- T5 nefunkčný odpor R4,
- T6 nefunkčný kondenzátor C2,
- T7 chybný triak Tc,
- T8 chybný diak Dc,
- T9 chybný nabíjací kondenzátor C1.

1. Prehľadávanie priestoru pojmov

Prehľadávanie priestoru pojmov (version space search) [4] predstavuje implementáciu induktívneho učenia nad týmto priestorom. Tomuto prehľadávaniu predchádza usporiadanie priestoru pojmov napríklad podľa všeobecnosti. Preto hovoríme o riadenom prehľadávaní, pričom na riadenie

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	T
a	a	n	n	n	n	n	n	n	n	n	n	a	a	n	a	n	a	T1
a	a	n	n	n	n	n	n	n	n	n	n	n	a	n	a	n	a	T2
a	a	n	n	n	n	n	n	n	n	n	n	n	n	a	a	n	a	T3
n	a	n	n	n	n	n	n	n	n	n	n	n	n	n	a	n	a	T4
a	a	n	n	n	n	n	n	n	n	n	n	a	n	n	a	a	a	T1
a	a	n	n	n	n	n	n	n	n	n	n	n	n	n	a	n	a	T1
a	a	n	n	n	n	n	n	n	n	n	n	a	n	a	a	n	a	T1
n	a	n	n	n	n	n	n	n	n	a	n	n	n	n	a	n	a	T4
n	n	n	n	a	a	a	a	n	n	n	n	n	n	n	a	n	a	T7
n	n	a	a	a	a	a	a	n	n	n	n	n	n	n	a	n	a	T5
n	n	a	a	a	a	a	a	n	n	n	n	n	n	n	a	n	a	T6
n	n	a	a	a	a	a	a	n	n	a	n	n	n	a	a	n	a	T6
n	n	n	n	n	n	a	a	n	a	n	n	n	n	n	a	n	a	T9
n	n	n	n	n	n	a	a	n	a	a	a	n	n	n	a	n	a	T9
n	n	n	n	n	n	a	a	a	a	n	n	n	n	n	a	n	a	T8
n	n	n	n	n	n	a	a	a	a	n	n	n	a	n	a	n	a	T8
n	n	n	n	n	n	a	a	a	a	n	n	n	n	a	a	n	a	T8

Tab.1 Množina tréningových príkladov



Obr.2 Fragment usporiadaného priestoru pojmov

prehľadávania sa používa práve spomínané usporiadanie podľa všeobecnosti.

Prehľadávanie pojmov v priestore sa uskutočňuje pomocou operátorov zovšeobecnenia, resp. operátorov špecifikácie, ktoré umožňujú pohyb v tomto priestore. Za primárne operácie zovšeobecnenia používané v strojovom učení sa považuje: nahradenie konštanty premennou, vypustenie podmienky z konjunktívneho výrazu a pridanie disjunkcie do výrazu. Fragment usporiadaného priestoru pojmov, ilustrovaný na príklade triakového regulátora vyhrievacej jednotky, je uvedený na obr. 2.

Na tomto obrázku je pojem *skrat_špirály* opisovaný pomocou atribútov: A1, A2, A13, A14, A15, A16, A17 a A18. Opis je zjednodušený vynechaním niektorých atribútov, ktoré sú pre daný prípad nevýznamné. Opis pojmu predstavuje konjunkciu hodnôt atribútov. Najšpecifickejšími pojmami sú jednotlivé pozorovania, resp. trénovacie príklady. Tieto pojmy sa v priestore pojmov nachádzajú na najspodnejšej úrovni. Takýmto pojmom je napríklad pozorovanie *skrat_špirály (a, a, n, n, n, a, n, a)*, ktorý predstavuje konjunkciu:

$A1 = a \ \& \ A2 = a \ \& \ A13 = n \ \& \ A14 = n \ \& \ A15 = n \ \& \ A16 = a \ \& \ A17 = n \ \& \ A18 = a.$

Na najvyššej úrovni všeobecnosti v priestore pojmov sa v tomto prípade nachádza zovšeobecnený pojem, ktorý predstavuje riešenie – diagnostikovanú chybu:

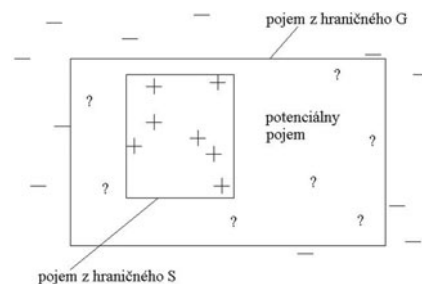
$A1 = a \ \& \ A2 = a \ \& \ A16 = a \ \& \ A18 = a.$

Na hodnotách A13, A14, A15 a A17 v tomto prípade nezáleží.

V princípe existujú tri spôsoby prehľadávania priestoru pojmov a to: v smere od špecifického k všeobecnému, v smere od všeobecného ku špecifickému a kombinácia týchto dvoch prístupov do obojsmerného prehľadávania. My sa budeme podrobnejšie venovať iba obojsmernému prehľadávaniu, ktoré realizuje algoritmus eliminácie kandidátov pojmov. Tento algoritmus je riadený údajmi, tzv. trénovacími príkladmi. Podrobnejší opis algoritmu možno nájsť v [3]. Algoritmus buduje dve množiny kandidátov pojmov:

- **G** – množinu maximálne všeobecných kandidátov pojmov a
- **S** – množinu maximálne špecifických kandidátov pojmov.

Algoritmus špecifikuje obsah **G** a zovšeobecňuje obsah **S**, až kým obidve neskonvergujú k cieľovému pojmu. Prácu algoritmu eliminácie kandidátov pojmov ilustruje obr. 3.



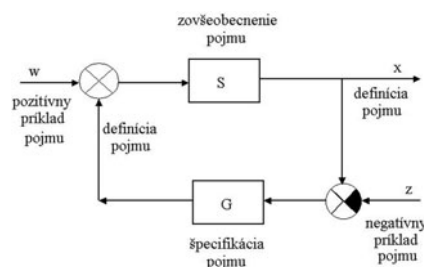
Obr.3 Hranice pokrývania množín G a S algoritmu eliminácie kandidátov

Algoritmus „sťahuje“ hraničné **G** natolko, nakoľko je to nutné v záujme vylúčenia všetkých negatívnych trénovacích príkladov. Zároveň „rozťahuje“ hraničné **S** do tej miery, aby zahrnul všetky pozitívne trénovacie príklady. Ak **G** a **S** skonvergujú do toho istého pojmu, daný pojem je riešením a algoritmus môže ukončiť svoju prácu. Ak **G** a **S** ostávajú prázdne, potom neexistuje pojem, ktorý by bol konzistentný s trénovacími príkladmi, teda neexistuje riešenie. To sa môže stať, ak sú trénovacie údaje zašumené alebo ak sa cieľový pojem nedá vyjadriť v danom reprezentačnom jazyku.

Zaujímavým aspektom tohto algoritmu je jeho inkrementálna povaha, čo znamená, že akceptuje trénovacie príklady postupne, jeden za druhým. Formuje použiteľné, aj keď možno neúplné zovšeobecnenie, resp. špecifikáciu po každom prípade.

2. Regulovaný proces učenia

Ako ilustračný príklad procesu učenia bol vyššie použitý regulátor vyhrievacej jednotky. Avšak aj na samotný proces učenia je možné nazerať ako na regulovaný, resp. riadený proces. V oblasti strojového učenia existuje množstvo kognitívnych algoritmov, ktoré používajú rôzne riadiace mechanizmy: riadenie údajmi (trénovacími príkladmi), riadenie chybou (chybnou klasifikáciou jednotlivého trénovacieho príkladu) a pod. Aj algoritmus eliminácie kandidátov pojmov, ktorý bol opísaný v predchádzajúcej časti, je možné vidieť ako proces riadený údajmi, teda trénovacími príkladmi. Túto skutočnosť ilustruje obr. 4.



Obr.4 Riadiaci obvod eliminácie kandidátov pojmov

Na tomto obrázku vidíme, ako príchod negatívneho príkladu daného pojmu, teda poruchy, vyvolá špecifikáciu pojmov v množine **G**. Účelom tejto špecifikácie je zúžiť definície pojmov natoľko, aby nepokrývali negatívne príklady. Na druhej strane, príchod pozitívneho príkladu daného pojmu, teda riadiacej veličiny, vyvolá zovšeobecnenie pojmov v množine **S**. Účelom tohto zovšeobecnenia je rozšíriť definície pojmov v množine **S** natoľko, aby pokryli aj nový pozitívny príklad. Definícia pojmu, na ktorej sa ustáli veličina **x** po spracovaní všetkých trénovacích príkladov, predstavuje výsledok regulácie.

Záver

V tomto článku je načrtnutá paralela strojového učenia s riadiacim procesom. Zameriava sa na riešenie úloh klasifikačného typu pomocou prehľadávania priestoru pojmov. Klasifikačnú úlohu a jej riešenie ilustruje na príkladoch diagnostikovania poruchy technického zariadenia. Prehľadávanie priestoru pojmov sa uskutočňuje pomocou algoritmu eliminácie kandidátov pojmov. Je to kognitívny algoritmus, ktorý slúži na riešenie klasifikačnej úlohy. Algoritmus teda vygeneruje definíciu pojmu, ktorá predstavuje klasifikačné pravidlo. Ak sa na technickom zariadení, napríklad na triakovom regulátore vyhrievacej jednotky, vyskytne záhada a je následne opísaná svojimi príznakmi (hodnotami atribútov trénovacej tabuľky), klasifikačné pravidlo priradí túto záhadu k jej príčine. Klasifikačná úloha v tomto prípade predstavuje diagnostiku poruchy technického zariadenia.

Literatúra

- [1] KUBÁT, M.: Strojové učení. In: Mařík, V., Štěpánková, O., Lažanský, J.: *Umělá inteligence*. ACADEMIA, Praha 1993. ISBN 80-200.0502-3
- [2] LANGLEY, P.: *Elements of Machine Learning*. Morgan Kaufmann Publishers, Inc. San Francisco, California 1996, pp. 419.
- [3] MACHOVÁ, K.: *Strojové učenie. Princípy a algoritmy*. ELFA, s. r. o., Košice 2002, 117 s. ISBN 80-89066-51-8
- [4] MITCHELL, T. M.: *Machine Learning*. The McGraw-Hill Companies, Inc. New York, 1997 b, pp. 414.

40

Ing. Kristína Machová, CSc.

**Katedra kybernetiky a umelej inteligencie FEI TU
Letná 9, 042 00 Košice
e-mail: Kristina.Machova@tuke.sk**