

Učení se k-term DNF pomocí k-CNF

Návod k jedenáctému cvičení

ZS 2011/2012

Na přednáškách jste se seznámili s postupem, pomocí něhož se lze učit k-term DNF formule tak, že se nejprve naučíme k-CNF formuli a tu pak převedeme na DNF formuli (která bohužel nemusí být k-DNF). V tomto cvičení implementujete tuto metodu a zkusíte se pomocí ní naučit DNF formuli pro malý dataset.

Data

1. V souboru `data.m` jsou definována trénovací data pro učení (nejedná se o skutečná data). Data popisují pacienty: jejich věk, typ předepsaných brýlí, zda mají astigmatismus a jestli mají redukované nebo normální množství slz. Naším cílem je naučit se pravidla (v podobě diskunkce konjunkcí, tj. DNF), která nám řeknou, komu mají být předepsány kontaktní čočky (pozitivní příklady) a komu ne (negativní příklady). Trénovací příklady jsou uloženy v proměnné `examples`. Zařazení příkladů do tříd (pozitivní ... 1, negativní ... -1) je uloženo v proměnné `labels`. Názvy proměnných jsou uloženy v proměnné `header`.

Nastudujte si...

2. Z přednáškových slajdů si nastudujte metodu učení se DNF formulí pomocí učení se k-CNF formulí. **Stručně:** Dostaneme multi-množinu pozitivních trénovacích příkladů. Rozšíříme ji přidáním "odvozených atributů" (odpovídajících všem k-klauzulím vytvořeným z atributů původního datasetu) - těch je obecně $\mathcal{O}(n^k)$. Na takto transformovaných datech se naučíme monotónní konjunkci. Pokud taková konjunkce existuje, máme k-CNF formuli (dosadíme do naučené monotónní konjunkce za odvozené atributy). Formuli v k-CNF převedeme do DNF formy. Pokud jsme měli štěstí, dostaneme k-term DNF formuli, pokud ne, dostaneme větší DNF formuli obsahující více termů.

Implementujte...

3. Implementujte učení se k-CNF formulí do funkce `k_cnf_learn`. K dispozici máte funkci `transform`, kterou můžete použít pro rozšíření trénovacích

příkladů o odvozené atributy (odpovídající všem k-konjunkcím). Kromě toho máte připravenou funkci `cnf2str`, která umožňuje přehledně vypisovat formule v CNF.

4. Použijte funkci `k.cnf.learn` a funkci `cnf2dnf`, kterou máte k dispozici, k vytvoření funkce pro učení se DNF formulí pomocí k-CNF formulí (ideálně bychom chtěli dostat k-term DNF formuli, ale to není vždy možné). Pro vypisování DNF formulí můžete použít funkci `dnf2str`.

Otestujte...

5. Použijte funkci pro učení se DNF-formulí na datasetu popisujícího pacienty. **Očekávaný výsledek:** $(\text{tear-prod-rate_normal} \ \& \ \neg \text{age_presbyopic} \ \& \ \neg \text{age_pre-presbyopic} \ \& \ \text{age_young}) \ \text{or} \ (\text{tear-prod-rate_normal} \ \& \ \neg \text{age_presbyopic} \ \& \ \neg \text{astigmatism_yes} \ \& \ \neg \text{age_young} \ \& \ \text{age_pre-presbyopic}) \ \text{or} \ (\text{tear-prod-rate_normal} \ \& \ \neg \text{age_presbyopic} \ \& \ \text{spectacle-prescrip_myope} \ \& \ \neg \text{age_young} \ \& \ \text{age_pre-presbyopic}) \ \text{or} \ (\text{tear-prod-rate_normal} \ \& \ \neg \text{spectacle-prescrip_myope} \ \& \ \neg \text{astigmatism_yes} \ \& \ \neg \text{age_pre-presbyopic} \ \& \ \neg \text{age_young} \ \& \ \text{age_presbyopic}) \ \text{or} \ (\text{tear-prod-rate_normal} \ \& \ \text{astigmatism_yes} \ \& \ \text{spectacle-prescrip_myope} \ \& \ \neg \text{age_pre-presbyopic} \ \& \ \neg \text{age_young} \ \& \ \text{age_presbyopic})$
6. Jak byste zjednodušili naučenou formuli s použitím doménové znalosti?