

ILP & BioData

ILP pro hledání vztahů mezi genotypem, fenotypem a snášenlivostí léčby

- 1) Popis dat (tabulka)
- 2) Motivace pro ILP
- 3) Transformace dat do formátu pro ILP
- 4) Doménová znalost
- 5) Příklady výstupu ILP

ILP & BioData

Popis dat

Lékařská data

Vliv léku na rychlost metabolizace určitého proteinu (MR)

Několik částí:

- Údaje o pacientovi

- Dotazník vedlejších příznaků

- Genotyp pacienta

- Detail genotypu (více možných hodnot)

- MR před a po léčbě

- Fenotyp pacienta (diskretizované MR)

ILP & BioData

Motivace pro ILP

Obsahují struktury špatně dělitelné do tabulky

U některých atributů chybí hodnoty (ale atribut je důležitý)

Klasické metody strojového učení neuspěly

Jde spíše o získání informací o datech, než o prediktivní model

ILP & BioData

Převod dat

Formát dat pro Aleph

Aleph – implementace v SWI prologu

Relační model (Prologovské klauzule)

Sloupce ~ Predikáty

ILP & BioData

Doménová znalost

Doménová znalost = prologovské pravidlo

Př: `stary(pacient1) :- vek(X), X>80.`

Přidává do dat další informaci

ILP & BioData

Příklad výstupu

[Rule 1] [Pos cover = 7 Neg cover = 0]

fenpo(A) :-

anxieta(A, 2), unava(A, 1).

[Rule 2] [Pos cover = 10 Neg cover = 0]

fenpo(A) :-

somnol(A, 0), fenpred(A, B), genetl(A, 1).