Privačios informacijos išsaugojimas taikant dirbtinio intelekto technologijas

Paulius Milmantas
Darbo vadovas: dr. Linas Petkevičius

Vilniaus Universitetas
Matematikos ir informatikos fakultetas

Bakalauro darbo gynimas

Tiriama sritis

Mašininis mokymas yra dirbtinio intelekto sritis, kuri pasitelkia statistinius algoritmus, kad apibrėžtų duomenų generavimo mechanizmą, ar egzistuojančius sąryšius, priklausomybes.

Problematika

1. Turint sukurtą modelį, neturi būti galima atgaminti duomenų, pagal kuriuos jis buvo mokomas, bei negali būti identifikuoti asmenys.
2. Trečios šalys neturi matyti įvedamų duomenų. Tai gali būti tinklo saugumo spragos, duomenų surinkimo aplikacijų spragos ir t.t. . .
3. Modelio išvesties neturi matyti asmenys, kuriems šie duomenys nepriklauso.
4. Sukurtas modelis negali būti niekieno pasisavintas.

Modelių duomenų lyginimas (1)

$$atvirumas(s[r])_\theta = log_2|r| - log_2 rangas_\theta(s[r]) \quad (1)$$

Naudojama teorijoje, dėl sunkiai apskaičiuojamo rango.
s - duomenų rinkinys.
r $\in$ **R**, parenkamas atsitiktinai.

Modelių duomenų lyginimas (2)

$$atvirumas(s[r])_\theta = -log_2 \int_0^{Px_\theta(s[r])} \rho(x)dx \tag{2}$$

Dėl grafinės interpretacijos naudojama praktikoje.
Px - logaritminis entropijos matas.
s[**r**] - entropija yra $\rho(.)$ pasiskirstymo distribucijos.

Pasiūlyta tyrimo metodika

$$DMDK = \sum_{n=0}^{m}(\sum_{k=0}^{h}(max((|\epsilon| + D_{n,k}) : \epsilon \in R))/h)/m \quad (3)$$

DMDK - Didžiausias maksimalus duomenų nuokrypis.
$\mathbf{D_{eilut:n,stulp:k}}$ - duomenys n eilutėje ir k stulpelyje.
ϵ - ieškomas didžiausias galimas kintamasis, su kuriuo modelis nepakeičia išvesties rezultatų.
m - duomenų eilučių skaičius.
h - parametrų skaičius (stulpeliai).