

## **II užduotis (Vieno neurono (perceptrono) mokymas sprendžiant klasifikavimo uždavinį)**

**Užduoties tikslas** – apmokyti vieną neuroną spręsti nesudėtingą dviejų klasių uždavinį, atlikti tyrimą.

### **Užduoties punktai:**

1. Parinkti duomenis, kurių klasės būtų žinomos. Kadangi vienas neuronas geba spręsti tik dviejų klasių klasifikavimo uždavinį, tai duomenys turi būti dviejų klasių. Reikia naudoti irisų duomenis (<https://archive.ics.uci.edu/ml/datasets/iris>), tačiau jie yra trijų klasių: Setosa, Versicolor ir Virginica. Reikia pasidaryti du duomenų rinkinius:
  - (1) vieną klasę sudaro Setosa rūšis (50 duomenų įrašų), kitą klasę – Versicolor ir Virginica rūšys (100 duomenų įrašų).
  - (2) vieną klasę sudaro Versicolor rūšis (50 duomenų įrašų), kitą klasę – Virginica rūšys (50 duomenų įrašų).
2. Klasių žymės (*label*) turi būti 0 arba 1.
3. Sukurti programą, kuri apmokytų vieną neuroną (perceptroną) spręsti nesudėtingą klasifikavimo uždavinį (galima naudoti bet kurią programavimo kalbą). Šioje užduotyje turi būti tik mokymas (testavimo, kryžminės patikros įgyvendinti nebūtina).
4. Programoje turi būti:
  - Duomenys nuskaityti iš failo.
  - Galimybė keisti tokius parametrus, kaip mokymo greitis (*learning rate*), epochų/iteracijų skaičius.
  - Įgyvendintos dvi aktyvacijos funkcijos: ir slenkstinė, ir sigmoidinė.
5. Programos rezultatas turi būti:
  - Gauti svoriai (analizuojant irisų duomenis, būtų penki svoriai  $w_0, w_1, w_2, w_3, w_4$ ).
  - Gauta paklaida,
  - Klasifikavimo tikslumas.

Klasifikavimo tikslumas gali būti santykis tarp teisingai klasifikuotų ir visų duomenų. Norint jį gauti, reikia kiekvienam duomenų įrašui paskaičiuoti klasę pagal gautus neuronų svorius. Naudojant slenkstinę funkciją, neurono išėjime gaunamas reikšmės 0 arba 1, tad duomenų priskyrimas klasėms yra akivaizdus. Naudojant sigmoidinę funkciją, neurono išėjimo reikšmės yra intervale (0, 1), tad vertinant klasifikavimo rezultatą, šias reikšmes reiktų suapvalinti arba nustatyti ribas, kad bus priskirta vienai ar kitai klasei.
6. Atlikti tyrimą naudojant du duomenų rinkinius (aprašytus pirmame punkte), kiekvieno rezultatus pateikiant lentelėse arba grafikuose su atitinkamais komentarais:
  - Kaip rezultatai priklauso nuo mokymo greičio parametro?
  - Kaip rezultatai priklauso nuo to, kuri aktyvacijos funkcija yra naudojama?
  - Kaip rezultatai priklauso nuo epochų/iteracijų skaičiaus?

P.S. rezultatus vertinti klasifikavimo tikslumo mato prasme. Šioje užduotyje pakanka klasifikavimo tikslumą vertinti mokymo duomenims, t. y. testavimo etapas nėra būtinas.

**Užduoties ataskaitoje:**

- Aprašyti, kokie duomenys buvo naudojami, kaip jie suskirstyti į dvi klases, kiek yra duomenų įrašų (eilučių), požymių (stulpelių). Pateikti kitą, jūsų manymu, svarbią informaciją;
- Paaiškinti kas yra epocha ir iteraciją.
- Paaiškinti kaip buvo parenkamos pradinių svorių reikšmės.
- Pateikti programos tekstą su išsamiais komentarais;
- Detaliai aprašyti atlikto tyrimo rezultatus (žr. 5 punktą);
- Nustačius geriausią variantą, pateikti gautus rezultatus: svorius, paklaidas, epochų/iteracijų skaičius, klasifikavimo tikslumo įverčius. Taip pat pateikti duomenų klasifikavimo rezultatus (kokias klases nustatė neuronas);
- Pateikti visos užduoties ir tyrimo išvadas (mažiausiai penki sakiniai).

**P. S.** Ataskaitoje turi būti aprašytas kiekvienas atliekamas veiksmas, pateikti žymėjimų aprašymai ir kita, jūsų manymu, svarbi informacija.