

II užduotis

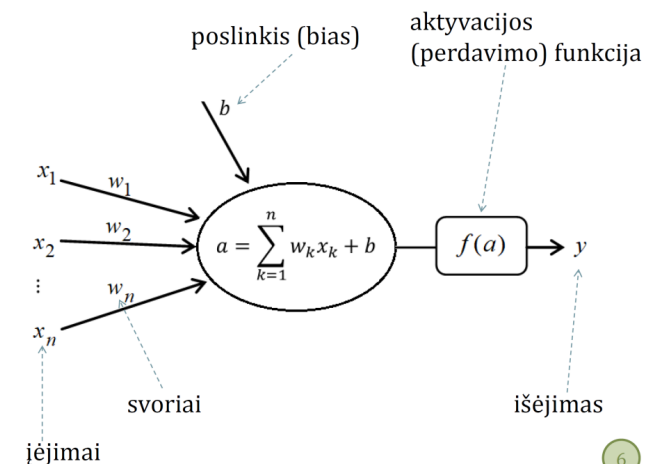
Vieno neurono (perceptrono) mokymas
sprendžiant klasifikavimo uždavinį

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

Užduoties tikslas – apmokyti vieną neuroną spręsti dviejų klasių uždavinį, atlikti tyrimą su dviem duomenų aibėmis.

Šiai užduočiai reiks naudoti dvi duomenų aibes (Irisų ir Krūties vėžio):

- **Irisų duomenų aibę** galima parsisiųsti iš <https://archive.ics.uci.edu/ml/datasets/iris>. Šiuose duomenyse yra trys klasės Setosa, Versicolor ir Virginica. Analizei reikia imti tik dvi: Versicolor ir Virginica.
- **Krūties vėžio duomenų aibę** galima parsisiųsti iš [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)). Šiuose duomenyse yra dvi klasės: 2 – nepiktybinis navikas, 4 – piktybinis navikas.





Index of /ml/machine-learning-databases/iris

- [Parent Directory](#)
- [Index](#)
- [bezdekIris.data](#)
- [iris.data](#)
- [iris.names](#)

Iris Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Famous database; from Fisher, 1936



Data Set Characteristics:	Multivariate	Number of Instances:	150	Area:	Life
Attribute Characteristics:	Real	Number of Attributes:	4	Date Donated	1988-07-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	4221600

```
4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
_ _ _ _ _
```

<https://archive.ics.uci.edu/ml/datasets/iris>



Index of /ml/machine-learning-databases/breast-cancer-wisconsin

- [Parent Directory](#)
- [Index](#)
- [breast-cancer-wisconsin.data](#)
- [breast-cancer-wisconsin.names](#)
- [unformatted-data](#)
- [wdbc.data](#)
- [wdbc.names](#)
- [wpbc.data](#)
- [wpbc.names](#)

Geriau naudoti šį failą

Breast Cancer Wisconsin (Diagnostic) Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Diagnostic Wisconsin Breast Cancer Database



Data Set Characteristics:	Multivariate	Number of Instances:	569	Area:	Life
Attribute Characteristics:	Real	Number of Attributes:	32	Date Donated	1995-11-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	1807653

```

1079304,2,1,1,1,2,1,2,1,1,2
1080185,10,10,10,8,6,1,8,9,1,4
1081791,6,2,1,1,1,1,7,1,1,2
1084584,5,4,4,9,2,10,5,6,1,4
1091262,2,5,3,3,6,7,7,5,1,4
1096800,6,6,6,9,6,?,7,8,1,2
1099510,10,4,3,1,3,3,6,5,2,4
1100524,6,10,10,2,8,10,7,3,3,4
1102573,5,6,5,6,10,1,3,1,1,4
1103608,10,10,10,4,8,1,8,10,1,4
1103722,1,1,1,1,2,1,2,1,2,2

```

Gali būti praleistų reikšmių

[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic))

1. Parsisiųsti ir paruošti duomenis. Klasijų žymes (*label*) pakeiskite 0 arba 1. Ištrinkite nereikalingus atributus (stulpelius), pvz., ID.
2. Sukurti programą (parašyti kodą), kuri įgyvendintų vieno neurono (perceptrono) mokymo ir testavimo procesą, sprendžiant klasifikavimo uždavinį. Galima naudoti bet kurią programavimo kalbą. Turi būti įgyvendintas neurono mokymas ir testavimas, todėl turimi duomenys turi būti padalinami į mokymo ir testavimo aibes. Įprastai santykis tarp mokymo ir testavimo aibių yra 80:20 arba 70:30.

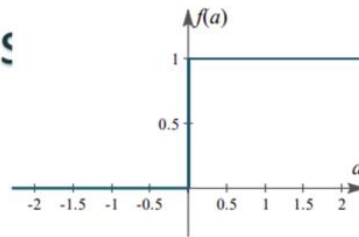
Programoje turi būti įgyvendinta:

- Duomenų nuskaitymas iš failo.
- Galimybė keisti tokius hiperparametrus, kaip mokymo greitis (*learning rate*), epochų skaičius.
- Dvi aktyvacijos funkcijos: ir slenkstinė, ir sigmoidinė.

Aktyvacijos funkcijos

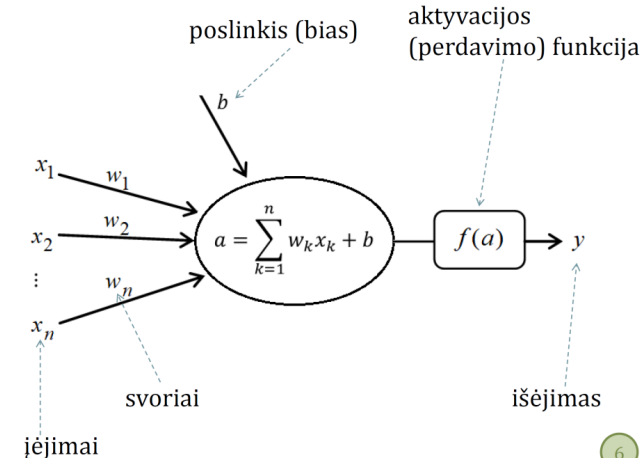
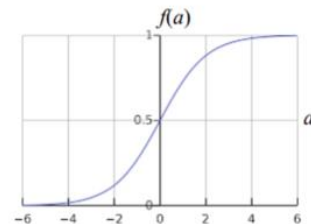
- Slenkstinė

$$f(a) = \begin{cases} 1, & \text{if } a \geq 0 \\ 0, & \text{if } a < 0 \end{cases}$$



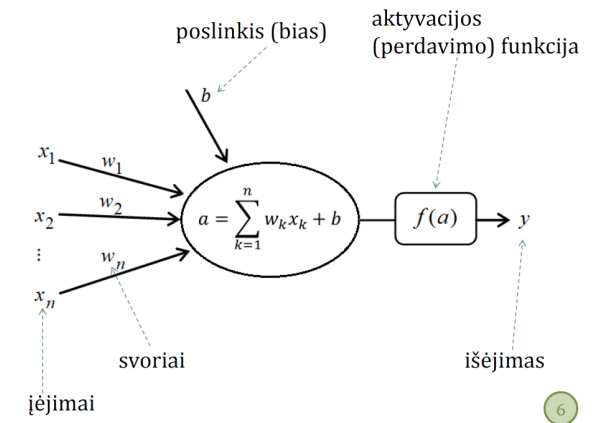
- Sigmoidinė

$$f(a) = \frac{1}{1 + e^{-a}}$$



Perceptrono mokymas

- Galimybė **mokytis** yra **esminė intelekto savybė**.
- Tinkamų svorių radimas vadinamas neurono (perceptrono) **mokymu**.
- Duomenų aibė, kuri bus naudojama neuronui mokyti, vadinama **mokymo aibe**.
- Tegul turime m mokymo aibės vektorių $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$, $i = 1, \dots, m$, kuriuos vadinsime **įėjimų vektoriais**.
- Šie vektoriai yra susieti su **norima reikšme** t_i (*target*). Tai norima reakcija į vektorių X_i .
- Sprendžiant klasifikavimo uždavinį, **norimos reikšmės yra klasių numeriai**.

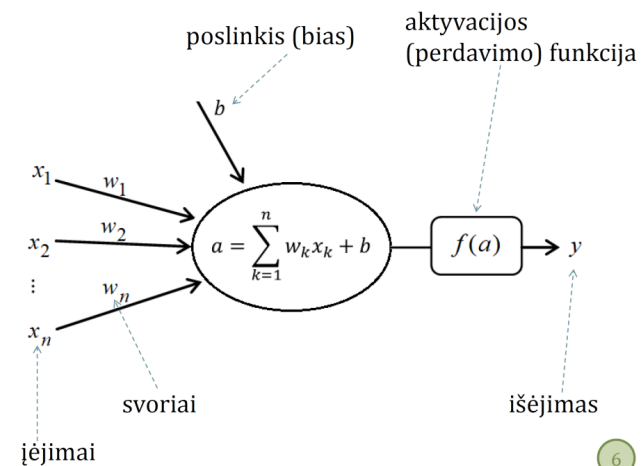


4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
- - - - -

Perceptrono mokymas

- **Mokymo procese** svoriai $W = (w_0, w_1, w_2, \dots, w_n)$ keičiami taip, kad tinklo išėjimo reikšmė y_i , gauta į įėjimą pateikus vektorių X_i , būtų kiek galima **artimesnė norimai reikšmei** t_i .
- Tai yra perceptrono veikimo **paklaida** būtų kiek galima **mažesnė**.
- Ši **paklaida** $E(W)$ gali būti apibrėžiama, kaip skirtumų tarp neurono išėjime gautų reikšmių ir norimų reikšmių kvadratinės sumos funkcija

$$E(W) = \sum_{i=1}^m (y_i - t_i)^2$$



```
4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
- - - - -
```

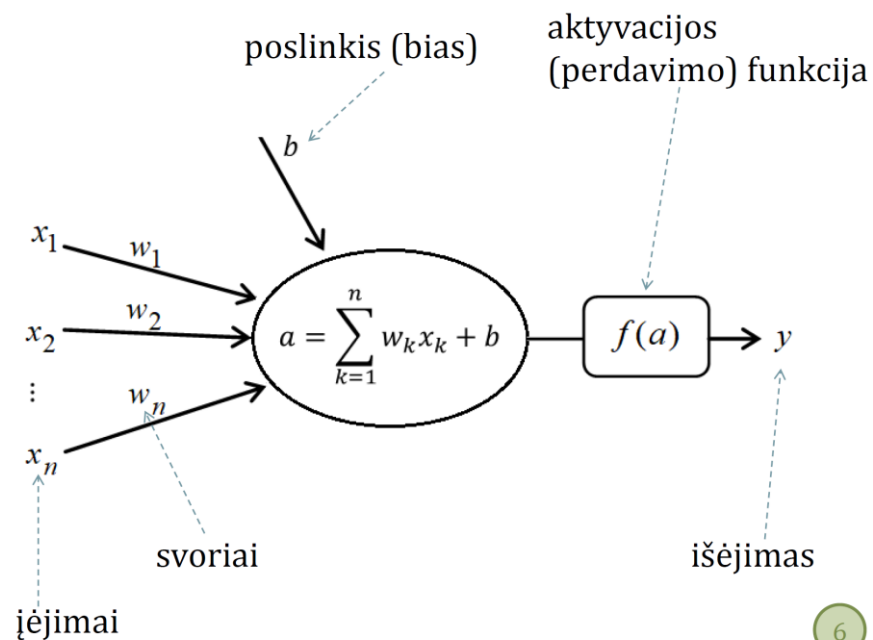
A. $W = a * \text{random}(-y, y)$

B. $W = a * \text{random}(\text{in_num}, \text{y_num})$

C. $W = a * \text{random}(\text{in_num}, \text{y_num}); a = 1 / \sqrt{\text{in_num}}$ – *Xavier initialization (*)*

D. $W = a * \text{random}(\text{in_num}, \text{y_num}); a = 1 / \sqrt{\text{in_num} / 2}$ – *He initialization (**)*

...



* Understanding the difficulty of training deep feedforward neural networks,

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.207.2059&rep=rep1&type=pdf>

** Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification

<https://arxiv.org/abs/1502.01852>

Perceptrono mokymas

- Vadinasi perceptrono mokymo eigoje reikia **minimizuoti paklaidos funkciją** $E(W)$, kuri dar vadinama **nuostolių** funkcija (*loss function*).
- Jeigu ši funkcija yra diferencijuojama pagal svorius, jos minimumą galima rasti **gradientiniais optimizavimo metodais**.
- Patogumo dėlei, dažnai minimizuojama **tokios išraiškos funkcija**:

$$E(W) = \frac{1}{2} \sum_{i=1}^m (y_i - t_i)^2$$

Perceptrono mokymas

- Iš pradžių **generuojamos atsitiktinės svorių** w_k reikšmės. Paprastumo dėlei, dažnai intervale $(0, 1)$, bet gali būti taikomos ir kitos pradinių svorių parinkimo strategijos.
- Tada gradientinio nusileidimo algoritmu judama antigradiento kryptimi, **svorių reikšmes keičiant** pagal iteracinę formulę

$$w_k(t + 1) = w_k(t) - \eta \frac{\partial E(W)}{\partial w_k},$$

t – iteracijos numeris,

η – yra teigiamas daugiklis, kuris vadinamas **mokymo greičiu** (*learning rate*) ir kuriuo reguliuojamas gradientinio optimizavimo žingsnio ilgis.

Bendra taisyklė.
Geriau naudoti ADALINE
mokymo taisyklę (žr. kitą
skaidrę)

4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
- - - - -

Paklaidos minimizavimas (ADALINE mokymo taisyklė)

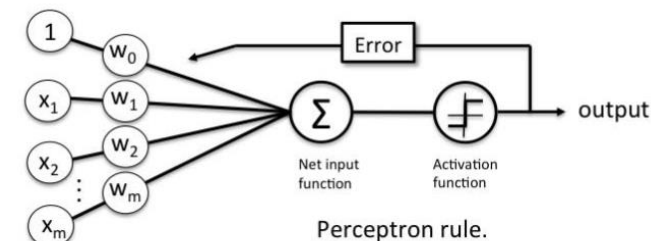
- **Minimizuojama** funkcija $E(W) = \frac{1}{2} \sum_{i=1}^m (t_i - y_i)^2$.
- Tai suma paklaidų **kiekvienam** įėjimų vektoriui:
 $E(W) = \sum_{i=1}^m E_i$.
- Prisiminkime, kad $y_i = f(a_i) = f(\sum_{k=0}^n w_k x_{ik})$.
- Randama funkcijos **išvestinė** pagal w_k :

$$\frac{\partial E_i(W)}{\partial w_k} = (t_i - y_i) \times \frac{\partial f(a)}{\partial w_k} = (t_i - y_i)(-x_{ik}),$$

$$\text{kai } f(\sum_{k=0}^n w_k x_{ik}) = \sum_{k=0}^n w_k x_{ik}.$$

- Todėl bendru atveju neurono mokyme galima taikyti **taisyklę**:

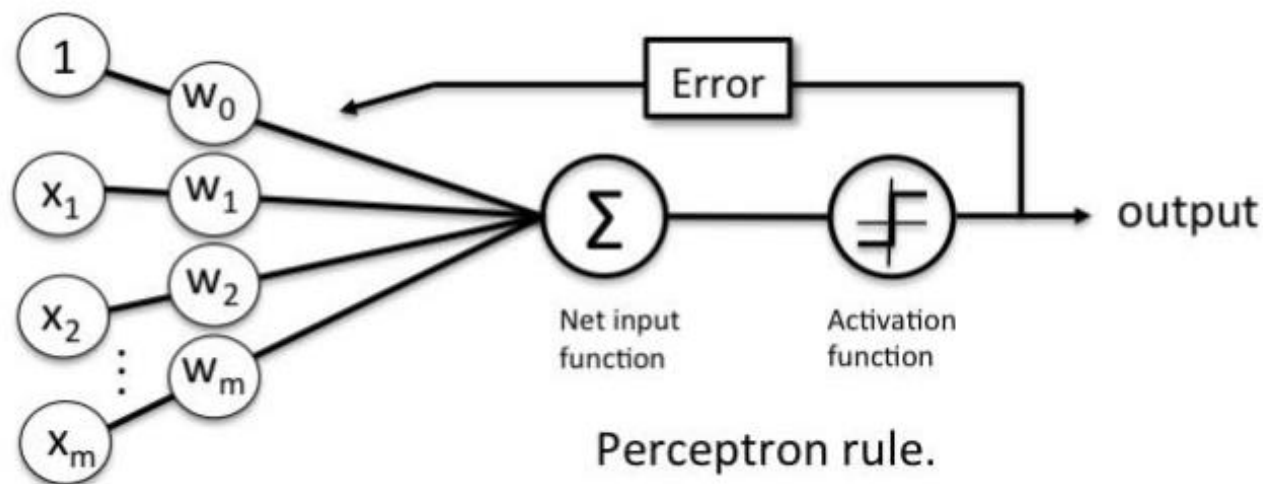
$$w_k(t+1) = w_k(t) + \eta(t_i - y_i)x_{ik}$$



```
4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
- - - - -
```

Perceptrono mokymas

- Įprastai svoriai **pakeičiami pateikus vieną** įėjimo vektorių.
- Mokymo procesas kartojamas **daug kartų pateikiant** visus įėjimo vektorius.
- Mokymas **stabdomas** arba atlikus iš anksto nustatytą **iteracijų skaičių**, arba pasiekus norimą **mažą paklaidos reikšmę**.



Perceptrono mokymo algoritmas

WHILE (išėjimo reikšmės nelygios trokštamoms reikšmėms ir iteracijų skaičius neviršija nustatytąjį)

{ FOR (visiems įėjimo vektoriams X_i)

{

$$a_i = w_0x_0 + w_1x_1 + w_2x_2$$

$$y_i = f(a_i)$$

IF ($y_i \neq t_i$)

$$w_k(\text{naujas}) = w_k(\text{senas}) + \eta t_i x_{ik}$$

}

}

Programos rezultatas turi būti:

- Gauti svoriai (analizuojant irisų duomenis, būtų penki svoriai w_0, w_1, w_2, w_3, w_4 , čia w_0 – poslinkis (bias); analizuojant krūties vėžio duomenis, būtų 10 svorių).
- Klasifikavimo tikslumas. Turi būti apskaičiuotas tikslumas ir mokymo, ir testavimo duomenims.

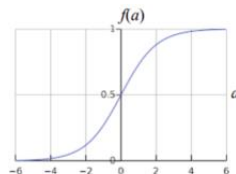
Klasifikavimo tikslumo matas gali būti santykis tarp teisingai klasifikuotų ir visų duomenų. Norint jį gauti, reikia kiekvienam duomenų įrašui paskaičiuoti klasę pagal gautus neuronų svorius. Naudojant slenkstinę funkciją, neurono išėjime gaunamos reikšmės 0 arba 1, tad duomenų priskyrimas klasei yra akivaizdus. Naudojant sigmoidinę funkciją, neurono išėjimo reikšmės yra intervale (0; 1), tad vertinant klasifikavimo rezultatą, šias reikšmes reiktų suapvalinti iki artimiausio sveiko skaičiaus (0 arba 1).

- Gautos paklaidos po kiekvienos epochos mokymo duomenims ir paklaida testavimo duomenims.

Paklaidą reikia skaičiuoti, skirtumą tarp perceptrono gautų ir trokštamų reikšmių pakėlus kvadratu. Skaičiuojant paklaidą, sigmoidinės funkcijos reikšmių apvalinti nereikia.

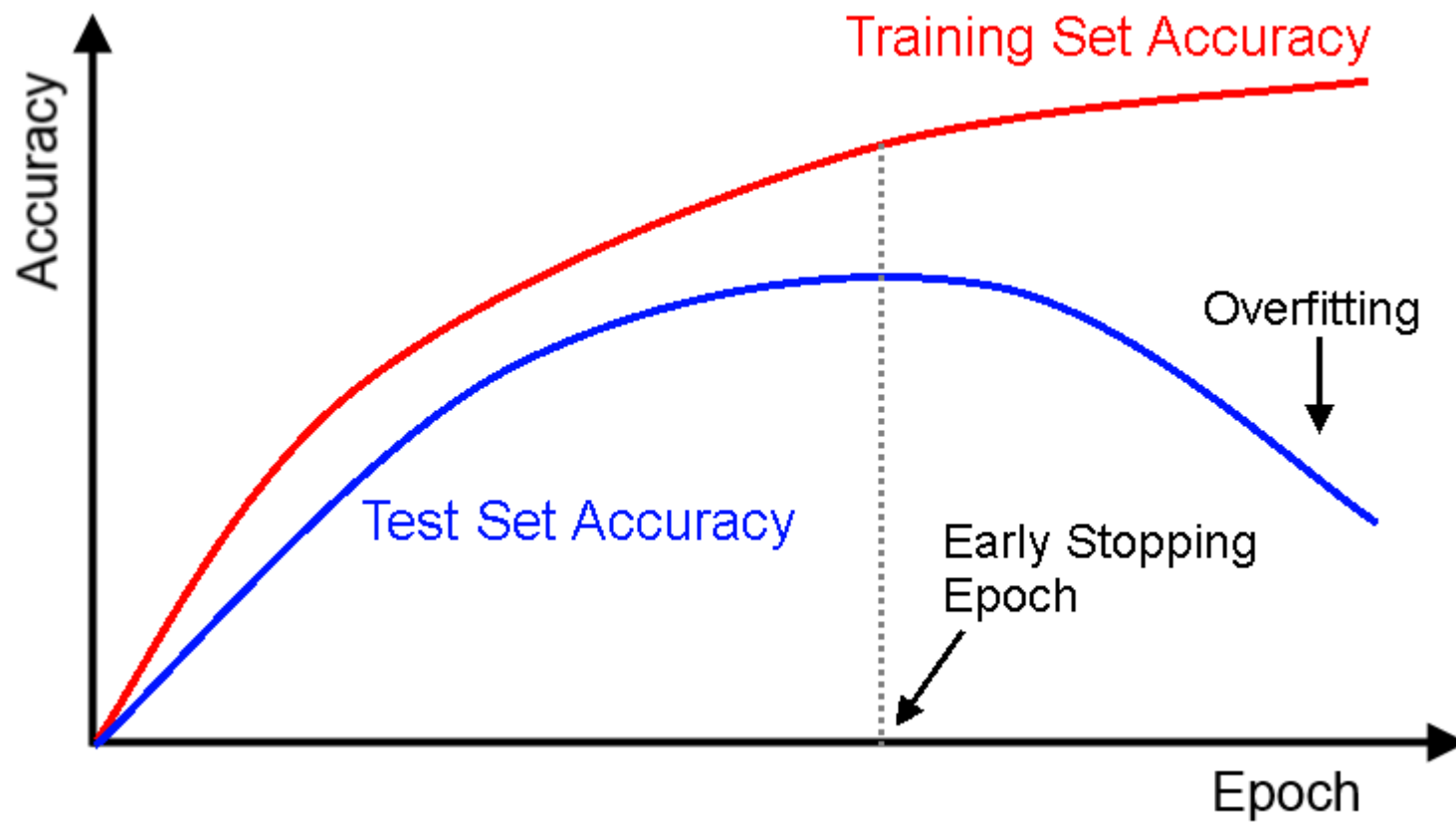
- Sigmoidinė

$$f(a) = \frac{1}{1 + e^{-a}}$$

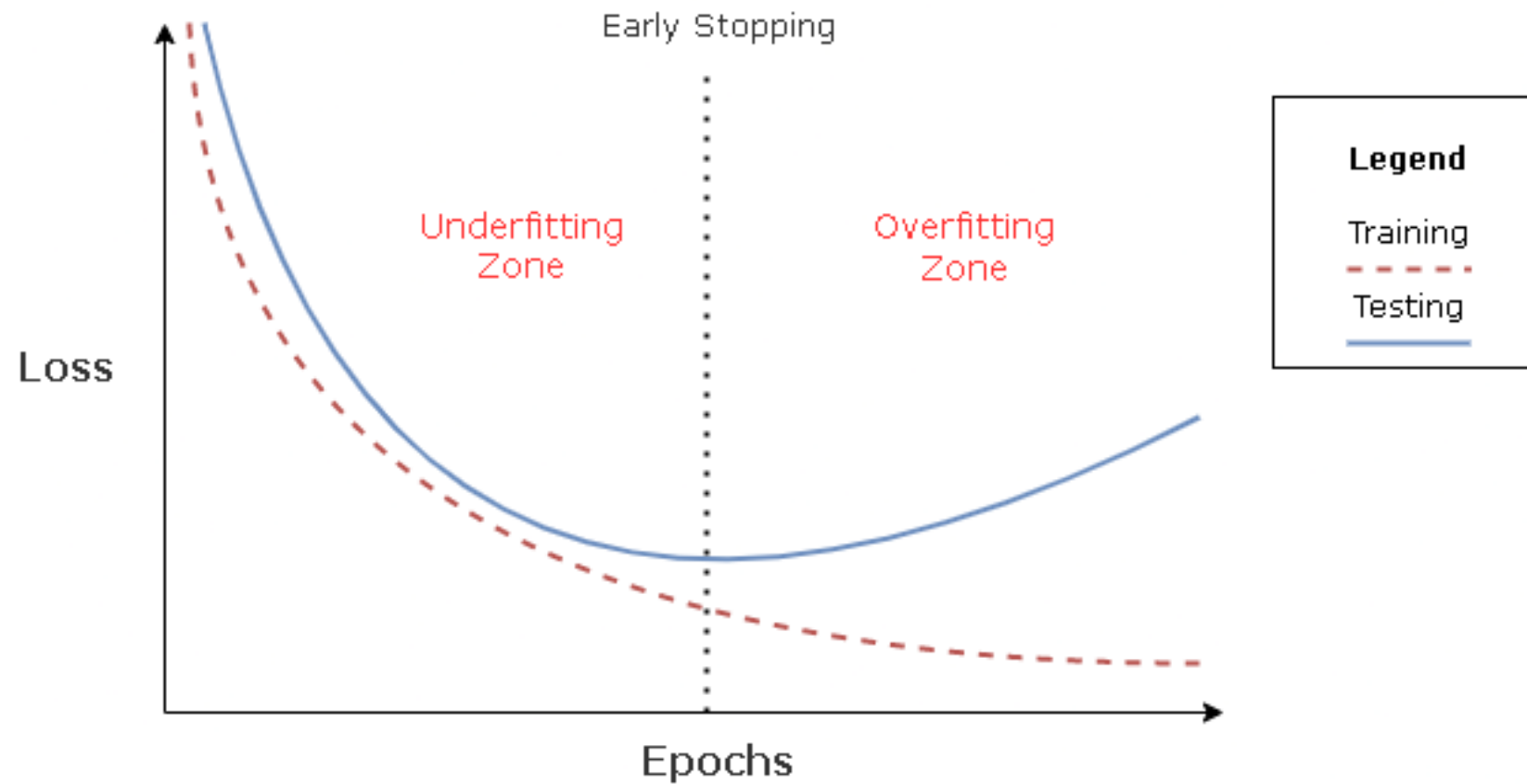


4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.0,1.4,0.1,Iris-setosa
4.3,3.0,1.1,0.1,Iris-setosa
5.8,4.0,1.2,0.2,Iris-setosa
_ _ _ _ _

$$E(W) = \frac{1}{2} \sum_{i=1}^m (t_i - y_i)^2$$



<https://forum.knime.com/t/how-to-generate-accuracy-curve-by-epoch-to-detect-overfitting-with-training-validation-data/11916>



3. Atlikti tyrimus ir vienai, ir kitai duomenų aibei (Irisų ir Krūties vėžio). Kiekvieno tyrimo rezultatus (klasifikavimo tikslumas ir paklaida) pateikti lentelėse arba grafikuose su atitinkamais komentaramis. Tie rezultatai, kuriuos lengviau ir greičiau galima suprasti iš grafikų, jie turi būti pateikti grafikuose, pvz., rezultatų priklausomybė nuo tiriamų hiperparametrų.

Tyrimų metu būtina nustatyti ir aprašyti:

- Kaip klasifikavimo tikslumas priklauso nuo epochų skaičiaus? Pateikti grafiką, kurio x ašyje yra atidėdamos epochos, o y ašyje klasifikavimo tikslumas mokymo duomenims.
- Kaip paklaidos reikšmės priklauso nuo epochų skaičiaus? Pateikti grafiką, kurio x ašyje yra atidėdamos epochos, o y ašyje paklaidų reikšmės mokymo duomenims.
- Kaip rezultatai priklauso nuo skirtingų mokymosi greičio reikšmių? Vaizdavimo būdą (lentelės, grafikai) galima pasirinkti patiems.
- Kaip rezultatai priklauso nuo to, kuri aktyvacijos funkcija yra naudojama? Vaizdavimo būdą (lentelės, grafikai) galima pasirinkti patiems.