Vilniaus Universitetas

Tiesioginio sklidimo DNT naudojant sistemą WEKA

Dirbtinio intelekto pagrindai

3 užduotis

Darbą atliko:

Dovydas Martinkus

Duomenų Mokslas 4 kursas 1gr.

Vilnius, 2022

**Turinys**

[1 Tikslas ir uždaviniai 3](#_Toc114422146)

[2 Duomenys 4](#_Toc114422147)

[3 Užduoties ataskaita 5](#_Toc114422148)

[3.1 Programinis kodas 5](#_Toc114422149)

[3.2 Tyrimo rezultatai 8](#_Toc114422150)

[4 Išvados 17](#_Toc114422151)

[Priedas 18](#_Toc114422152)

# Tikslas ir uždaviniai

Tikslas: Išmokyti neuroninį tinklą teisingai klasifikuoti duomenis naudojant sistemą WEKA.

Uždaviniai:

Analizuojamų duomenų paruošimas ir aprašymas.

Užduočių sekų sudarymas ir jų vaizdinis pateikimas su komentarais kam skirta kiekviena naudojama komponentė.

Neuroninio tinklo parametrų parinkimas: optimalių paslėptų neuronų skaičiaus, mokymo greičio parametro bei momentum reikšmių radimas pateikiant klasifikavimo tikslumo įverčius visiems tirtiems atvejams.

Naujų duomenų klasifikavimas naudojant išsaugotą tinklo modelį.

Neuronų išėjimo reikšmių perskaičiavimas „rankiniu“ būdu pasirinktoje programoje. Išvadų apie rezultatų sutapimą padarymas.

# Duomenys

Užduotyje naudoti 2 duomenų rinkiniai:

Irisų duomenų aibę sudaro 150 stebėjimų, iš kurių kiekvienas turi po 4 skaitinius požymius ir klasę, kuriai priklauso. Šiuose duomenyse yra trys klasės Setosa, Versicolor ir Virginica. Analizei naudotos tik dvi: Versicolor ir Virginica. Prieiga per internetą: <https://archive.ics.uci.edu/ml/datasets/iris>.

Krūties vėžio duomenų aibę sudaro 6873 stebėjimai, turintys po 10 skaitinių požymių ir klasės kintamąjį, kurį sudaro dvi klasės: 2 – nepiktybinis navikas , 4 – piktybinis navikas. Prieiga per internetą: <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic).>

Abi duomenų aibes pasirinkta dalinti į mokymo ir testavimo aibes naudojant santykį 80-20.

Kai kurie ataskaitoje naudojami terminai:

Epocha – vienas algoritmo pilnas perėjimas pro visus turimus mokymo duomenis.

Iteracija – tai mokymo proceso dalis, kai į perceptroną (arba į dirbtinį neuroninį tinklą) pateikiamas vienas mokymo aibės duomuo.

# Užduoties ataskaita

**from** scipy.io **import** arff **as** sciarff

**import** pandas **as** pd

**from** sklearn.model\_selection **import** train\_test\_split

**from** collections **import** Counter

data **=** sciarff**.**loadarff('iris.arff')

df **=** pd**.**DataFrame(data[0])

In [3]:

df**.**columns **=** ["sepal\_length","sepal\_width","petal\_length","petal\_width","class"]

df **=** df**.**replace({b'Iris-setosa': 'setosa', b'Iris-virginica': 'virginica', b'Iris-versicolor' : 'versicolor'})

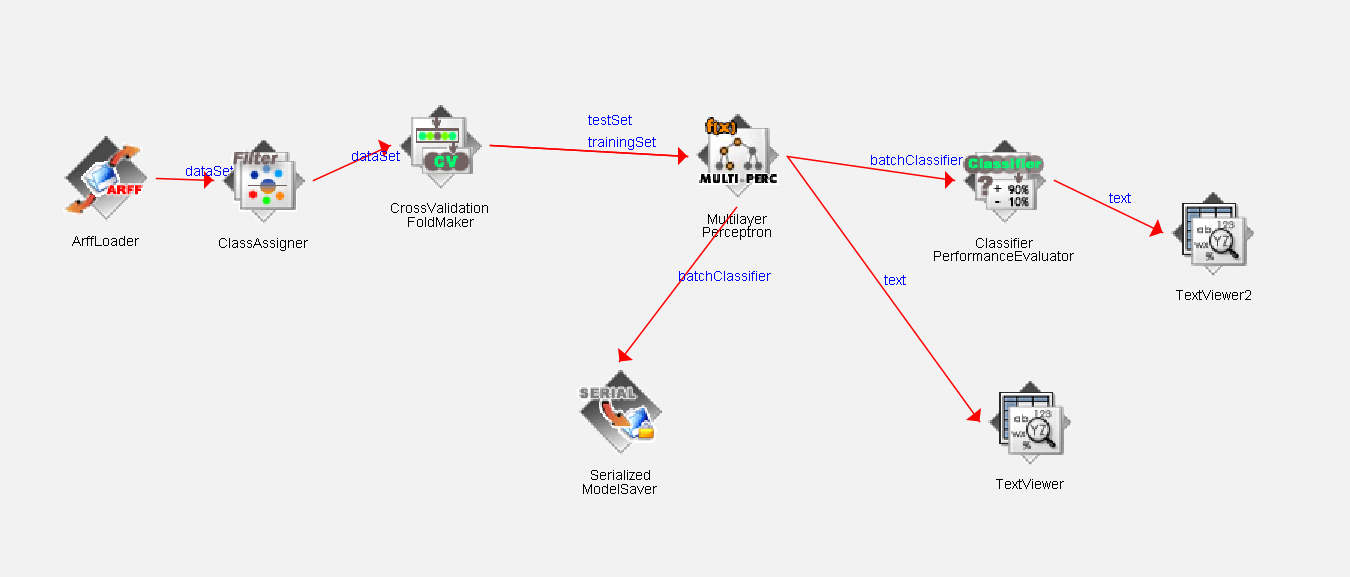
train, test **=** train\_test\_split(df,train\_size**=**0.8,test\_size**=**0.2,stratify**=**df["class"])

In [5]:

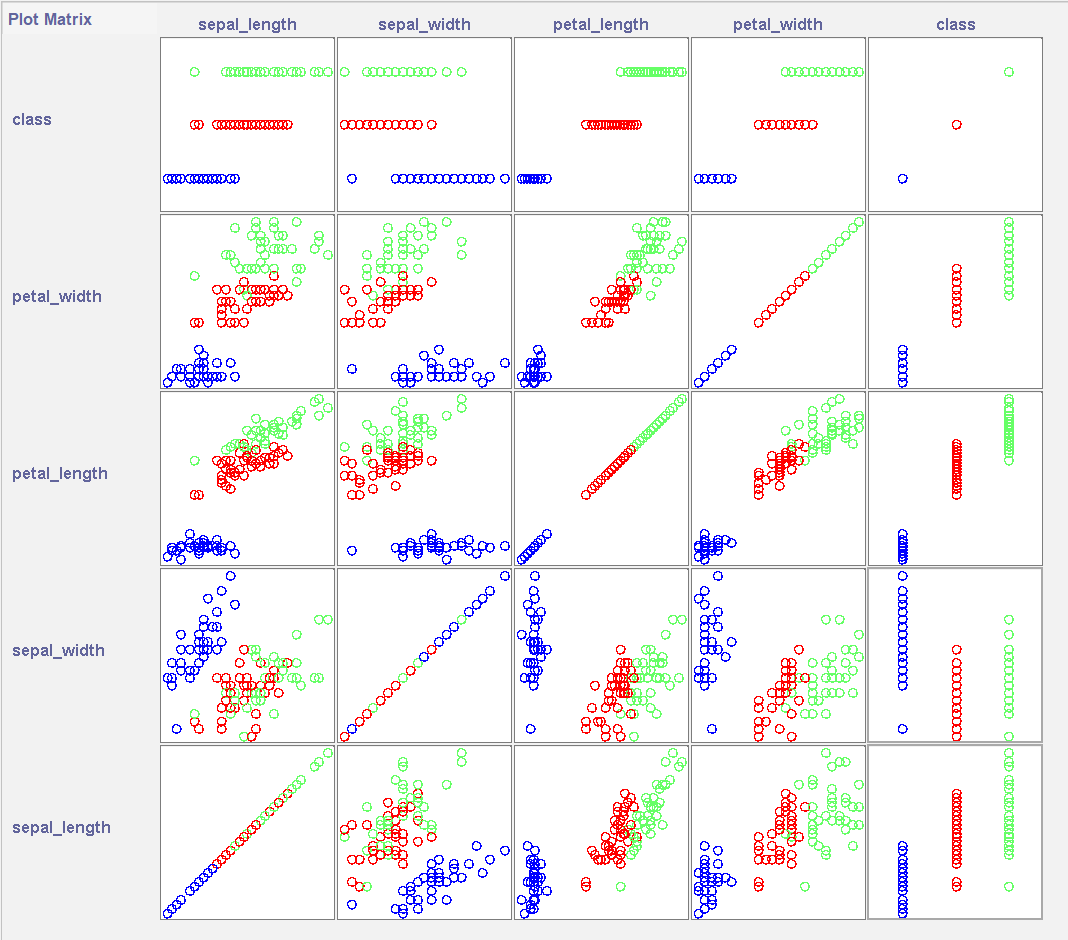
**from** pandas2arff **import** pandas2arff

pandas2arff(train, "iris\_train\_test.arff", wekaname**=**"iris\_train\_test")

pandas2arff(test, "iris\_new.arff", wekaname**=**"iris\_new")



1 pav. Sudarytos užduočių sekų vaizdas



2 pav. Duomenų požymių sklaidos diagramų matrica

Paveikslėlis, kuriame yra žinutė

Automatiškai sugeneruotas aprašymas

3 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė

Automatiškai sugeneruotas aprašymas

4 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė

Automatiškai sugeneruotas aprašymas

5 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė

Automatiškai sugeneruotas aprašymas

6 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė

Automatiškai sugeneruotas aprašymas

7 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė

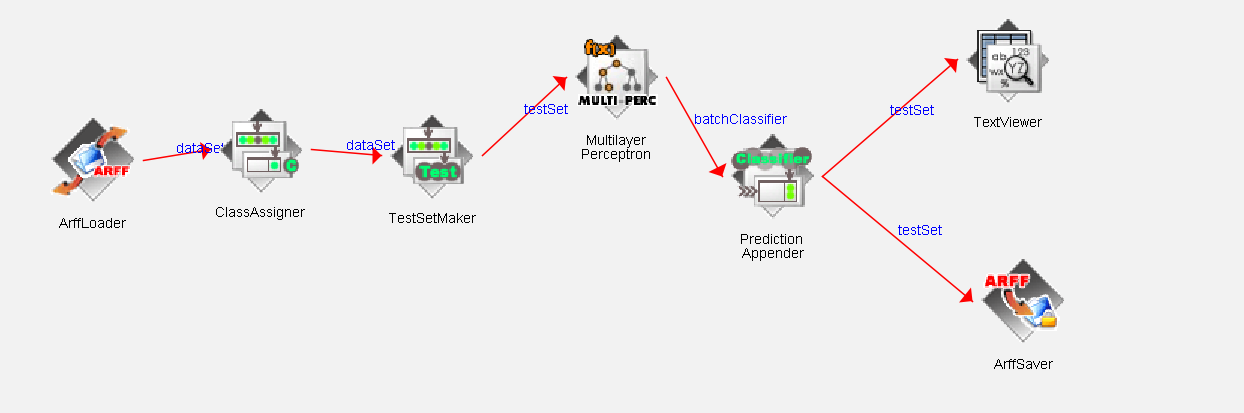
Automatiškai sugeneruotas aprašymas

8 pav. Klasifikavimo rezultatai

Paveikslėlis, kuriame yra žinutė, žemėlapis

Automatiškai sugeneruotas aprašymas

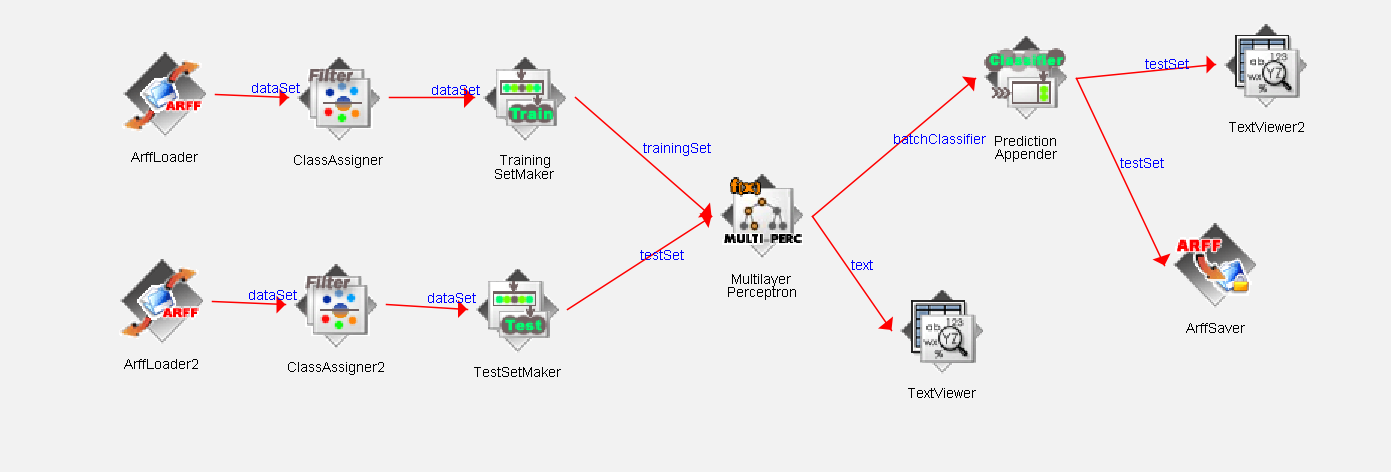
9 pav. Sudaryto dirbtinio neuroinio tinklo vaizdas



10 pav. Sudarytos užduočių sekų vaizdas

1 lentelė Tikros ir prognozuotos klasės testavimo aibei naudojant išsaugotą modelį

|  |  |
| --- | --- |
| Tikra klasė | Prognozuota klasė |
| setosa | setosa |
| setosa | setosa |
| versicolor | versicolor |
| versicolor | versicolor |
| versicolor | versicolor |
| versicolor | versicolor |
| setosa | setosa |
| virginica | virginica |
| setosa | setosa |
| versicolor | versicolor |
| virginica | virginica |
| versicolor | versicolor |
| versicolor | versicolor |
| virginica | virginica |
| setosa | setosa |
| versicolor | versicolor |
| setosa | setosa |
| virginica | virginica |
| virginica | virginica |
| virginica | virginica |
| setosa | setosa |
| virginica | virginica |
| setosa | setosa |
| setosa | setosa |
| virginica | virginica |
| virginica | virginica |
| virginica | virginica |
| versicolor | versicolor |
| setosa | setosa |
| versicolor | versicolor |



11 pav. Sudarytos užduočių sekų vaizdas

2 lentelė Išvesties sluoksnio svoriai skirtingiems neuronams

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0 | 1 | 2 |
| Poslinkis | 1.032 | -3.560 | -3.690 |
| 3 | 3.657 | -0.823 | -7.646 |
| 4 | -3.694 | 4.242 | 0.244 |
| 5 | -1.770 | -6.863 | 5.331 |
| 6 | -1.841 | -7.726 | 5.833 |
| 7 | -5.342 | 7.392 | 0.609 |

3 lentelė Paslėpto sluoksnio svoriai skirtingiems neuronams

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 3 | 4 | 5 | 6 | 7 |
| Poslinkis | 2.701 | 1.898 | -5.584 | -6.084 | 2.482 |
| sepal\_length | 1.514 | 1.150 | -2.324 | -2.725 | 1.313 |
| sepal\_width | 0.983 | -1.943 | -1.814 | -2.089 | -2.547 |
| petal\_length | -5.153 | 2.916 | 8.838 | 9.825 | 3.782 |
| petal\_width | -2.476 | 2.514 | 4.868 | 5.145 | 3.393 |

4 lentelė Excel atliktais skaičiavimais ir WEKA gautų klasių palyginimas

|  |  |
| --- | --- |
| Skaičiavimais gauta klasė | WEKA gauta klasė |
| setosa | setosa |
| setosa | setosa |
| versicolor | versicolor |
| versicolor | versicolor |
| versicolor | versicolor |
| versicolor | versicolor |
| setosa | setosa |
| virginica | virginica |
| setosa | setosa |
| versicolor | versicolor |
| virginica | virginica |
| versicolor | versicolor |
| versicolor | versicolor |
| virginica | virginica |
| setosa | setosa |
| versicolor | versicolor |
| setosa | setosa |
| virginica | virginica |
| virginica | virginica |
| virginica | virginica |
| setosa | setosa |
| versicolor | versicolor |
| setosa | setosa |
| setosa | setosa |
| virginica | virginica |
| virginica | virginica |
| virginica | virginica |
| versicolor | versicolor |
| setosa | setosa |
| versicolor | versicolor |

# Išvados