**Rīgas Tehniskā universitāte**

Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

**Atskaite par otro praktisko darbu**

Studiju kurss “Mākslīgā intelekta pamati”

Komandas numurs: F\_7

Darba izpildītāji:

Andris Smirnijs 231DDB022,

Maksims Kaļiņins 231DDB029,

Natalja Veļičko 231DDB018,

Vladislavs Jegorovs 211DDB016,

Ēriks Kubanovs 201REB600

Mācībspēks: Alla Anohina-Naumeca

Saite uz projektu: [[saite]](https://github.com/Kladbicscze/Pd.2-MIP)

Saite uz datu kopu: [[saite]](https://www.kaggle.com/datasets/valakhorasani/gym-members-exercise-dataset)

2024./2025.studiju gads

# Mākslīgā intelekta rīku izmantošanas paziņojums

< ja darba izstrādes gaitā tika izmantoti mākslīgā intelekta rīki, ir jādod saraksts ar izmantotajiem rīkiem (nosaukums, gads, versija, saite) un atskaites lapaspušu numuri, kurās ir ievietots teksts, kas tika iegūts no konkrētā rīka; ja rīki netika izmantoti, jādod skaidra šī fakta atruna>

# Orange rīka darbplūsma

<ir jādod Orange rīka izveidotās darbplūsmas ekrānuzņēmums visam darbam>

# I daļa

Informacija tika ņemta

## Datu kopas apraksts

### Datu kopas nosaukums: **Gym Members Exercise Dataset**

### Datu kopas avots: Datu kopa bija ģenerēta, izmantojot simulētus datus, kuru tiek pamatoti uz reālistisko vingrojumu modeļi sporta zālē, ņemot vērā: Publiski pieejamus fitnesa pētījumus; Nozares ziņojumus par vingrojumu un veselības tendencēm; Aptaujas, kas saistītas ar treniņu paradumiem, sirdsdarbības datiem un kaloriju patēriņu.

### Datu kopas izveidotājs un/vai īpašnieks: Seyed Vala Khorasani

Datu kopas problēmsfēras apraksts: Šis datu kopums sniedz detalizētu pārskatu par sporta zāles apmeklētāju vingrojumu režīmu, fiziskiem īpašībām un fiziskiem rādītājiem. Tas satur 973 ierakstus, tostarp galvenos snieguma rādītājus, piemēram, sirdsdarbības ātrumu, sadedzinātās kalorijas un treniņa ilgumu. Katrā ierakstā tiek iekļauti demogrāfiskie dati un pieredzes līmeņi, kas ļauj veikt visaptverošu fiziskās sagatavotības modeļu, sportista progresa un veselības trendu analīzi. [[saite]](https://www.kaggle.com/datasets/valakhorasani/gym-members-exercise-dataset)

Datu kopas licencēšanas nosacījumi:Apache 2.0[[saite]](https://www.apache.org/licenses/LICENSE-2.0)

Informācija par datu kopas savākšanas veidu vai procedūru:

Lai izveidotu datu kopu:

Galvenie mainīgie: Definēti, pamatojoties uz fitnesa pētījumos izplatītiem rādītājiem, tostarp sirdsdarbības ātrumu (BPM), sesijas ilgumu, sadedzinātās kalorijas, ķermeņa izmērus (vecumu, svaru, augumu) un pieredzes līmeni.

Dalībnieku profili: Ģenerēti, lai pārstāvētu dažādu sporta zāles apmeklētāju loku, ar treniņu intensitāti un pieredzes līmeni, kas nejaušināti sadalīts reālistiskos diapazonos.

Vingrinājumu modeļi: Simulēti, izmantojot datu vidējos rādītājus no fitnesa ziņojumiem, ar: Treniņu biežumu no 3 līdz 6 dienām nedēļā atkarībā no pieredzes līmeņa, Sesiju ilgumu no 0,5 līdz 2 stundām, Sadedzinātās kalorijas un ūdens uzņemšana tiek aprēķināta, pamatojoties uz treniņa veidu, BPM un demogrāfiskajiem faktoriem.

Datu strukturēšana: Dati tika strukturēti CSV formātā ar skaidri definētām kolonnām, lai nodrošinātu analīzes vienkāršību.

## Datu kopas satura apraksts

Datu objektu skaits datu kopā: 973

Datu kopas pazīmju (atribūtu) atspoguļojums kopā ar to lomām Orange rīkā:

A screenshot of a computer

AI-generated content may be incorrect.

### Klašu skaits datu kopā: 3

Klašu apraksts:

**Experience Level – 1** – (Beginner) Tas ir cilvēks, kurš piedalās sportzalā mazāk, nekā sešus mēnešus. Viņš nevar pacelt lielu svaru, ir samērā mazs. Parasti, par Iesācējiem tiek parskatīti tie cilvēki, kuram nav pieredzes sportām. [[saite]](https://bonytobeastly.com/beginner-intermediate-advanced-lifter/)

**Experience Level – 2 –** (Intermediary) Tas ir cilvēks, kurš piedalās sportzalā lielāk, nekā sešus mēnešus, bet mazāk, nekā divi gadi. Parasti, par Starpniekiem tiek parskatīti tie cilvēki, kuri zino savu rutīnu un pie tās pieturas. [[saite]](https://bonytobeastly.com/beginner-intermediate-advanced-lifter/)

**Experience Level – 3 –** (Advanced) Tas ir cilvēks, kurš piedalās sportzalā lielāk, nekā divi gadi. Parasti, par Ekspertiem tiek parskatīti tie cilvēki, kuri atrodas sportzalē sacensību dēļ, kuri grīb sasniegt jaunus augstumus un rekordus. [[saite]](https://bonytobeastly.com/beginner-intermediate-advanced-lifter/)

Datu objektu skaits, kas pieder katrai klasei:

|  |  |
| --- | --- |
| **Pazīmes apzīmējums/nosaukums** | **Datu Objektu Skaits** |
| **Experience Level - 1** | 376 |
| **Experience Level - 2** | 406 |
| **Experience Level - 3** | 191 |

### Pazīmju apraksts:

**Age** – Vecums, gados.

**Gender** – Dzimums, (Male/Female).

**Weight (kg)** – Svars, kilogrammos.

**Height (m)** – Augstums, metros.

**Max\_BPM** – Maksimālī sasniegta sirdsdarbība minutē.

**Avg\_BPM** – Videja sirdsdarbība minutē.

**Resting\_BPM** - Atpūtas sirdsdarbība minutē.

**Session\_Duration (Hours)** – Sessijas Ilgstums, stundos.

**Calories\_Burned** – Sadedzinātas Kalorijas vienā sessijā.

**Workout\_Type** – vingrinājumu komplekss (Yoga/HIIT/Cardio/Strength).

**Fat\_Percentage** - Tauku daļa ķermeņa svarā, procentos.

**Water\_Intake (Liters**) - Ūdens uzņemšana, litros.

**Workout\_Frequency (days/week)** - Treniņu biežums dienās/nedēļā.

**Experience\_Level** - Pieredzes līmenis (Beginner/Intermediary/Expert).

**BMI** – Ķermeņa Massas indekss, svars/augstums².

|  |  |  |  |
| --- | --- | --- | --- |
| **Klases iezīme** | **Pazīmes skaidrojums** | **Vērtību tips** | **Vērtību diapazons** |
| Age | Vecums | Numeric | [18 - 59] |
| Gender | Dzimums | Categorical | Male / Female |
| Weight (kg) | Svars | Numeric | [40 - 129.9] |
| Height (m) | Augstums | Numeric | [1.5 - 2] |
| Max\_BPM | Maksimālī sasniegta sirdsdarbība minutē. | Numeric | [160 - 199] |
| Avg\_BPM | Videja sirdsdarbība minutē. | Numeric | [120 - 169] |
| Resting\_BPM | Atpūtas sirdsdarbība minutē. | Numeric | [50 - 74] |
| Session\_Duration (Hours) | Sessijas Ilgstums, stundos. | Numeric | [0.5 - 2] |
| Calories\_Burned | Sadedzinātas Kalorijas vienā sessijā. | Numeric | [303 - 1783] |
| Workout\_Type | vingrinājumu komplekss. | Categorical | Yoga / HIIT / Cardio / Strength |
| Fat\_Percentage | Tauku daļa ķermeņa svarā. | Numeric | [10 - 35] |
| Water\_Intake (Liters) | Ūdens uzņemšana. | Numeric | [1.5 - 3.7] |
| Workout\_Frequency (days/week) | Treniņu biežums dienās/nedēļā. | Numeric | [2 - 5] |
| Experience\_Level | Pieredzes līmenis (Beginner/Intermediary/Expert). | Categorical | [1 - 3] |
| BMI | Ķermeņa Massas indekss, svars/augstums². | Numeric | [12.32 - 49.84] |

### Datu faila struktūra:

A screenshot of a computer

AI-generated content may be incorrect.

### A white background with black lines AI-generated content may be incorrect.

### Informācija par trūkstošajām vai izlecošajām vērtībām:

Mūsu datasetā nav trūkstošās vai izlecošās vērtības.

## Datu kopas vizuālais un statistiskais atspoguļojums

A screenshot of a computer

AI-generated content may be incorrect.

Fat Percentage – Workout Type Izkliedes diagrammas ekrānuzņēmums

A screen shot of a computer screen

AI-generated content may be incorrect.

BMI – Height (m) Izkliedes diagrammas ekrānuzņēmums

A screenshot of a computer

AI-generated content may be incorrect.

Calories Burned - Histogrammas ekrānuzņēmums

A screenshot of a computer

AI-generated content may be incorrect.

Session Duration - Histogrammas ekrānuzņēmums

A screen shot of a computer

AI-generated content may be incorrect.

Session Duration – Fat Percentage Pazīmju sadalījuma ekrānuzņēmums

A screen shot of a graph

AI-generated content may be incorrect.

Session Duration – Calories Burned Pazīmju sadalījuma ekrānuzņēmums

A screenshot of a computer

AI-generated content may be incorrect.

Ekrānuzņēmums ar statistiskajiem rādītājiem

## Atbildes uz jautājumiem

<jādod atbildes uz zemāk norādītājiem jautājumiem, atsaucoties uz augšā esošajiem ekrānuzņēmumiem un sniedzot iegūto rezultātu analīzi>

### Vai klases datu kopā ir līdzsvarotas, vai dominē viena klase (vai vairākas klases)?

### Vai datu vizuālais atspoguļojums ļauj redzēt datu struktūru?

### Cik datu grupējumus ir iespējams identificēt, pētot datu vizuālo atspoguļojumu?

### Vai identificētie datu grupējumi atrodas tuvu viens otram vai tālu viens no otra?

## Secinājumi, kas izriet no statistisko rādītāju analīzes

<jāveic statistisko rādītāju aprēķina analīze, atsaucoties uz konkrētajām vērtībām>

## Secinājumi par atlasītajām pazīmēm

<jāsniedz informācija par to, kādas pazīmes tiks izmantotas tālākajā analīzē>

# II daļa

<šajā apakšnodaļā ir jāsniedz apraksts par nepārraudzītās mašīnmācīšanās algoritmu lietojumu, to pavadot ar ekrānuzņēmumiem un atsaucēm uz izmantotajiem informācijas avotiem>

## Hierarhiskā klasterēšana

### Orange rīkā pieejamo hiperparametru apraksts:

<pievienojiet tabulai rindas pēc nepieciešamības>

|  |  |
| --- | --- |
| **Hiperparametrs** | **Apraksts** |
|  |  |
|  |  |

<ekrānuzņēmums ar algoritmam uzstādītajām hiperparametru vērtībām>

### Eksperimentu apraksts

<ekrānuzņēmums 1.eskperimentam, kurā tiek mainīts horizontālās atdalošās līnijas novietojums>

<ekrānuzņēmums 2.eksperimentam, kurā tiek mainīts horizontālās atdalošās līnijas novietojums>

<ekrānuzņēmums 3.eksperimentam, kurā tiek mainīts horizontālās atdalošās līnijas novietojums>

### Secinājumi no eksperimentiem:

<apraksts, kā mainās klasteru skaits un saturs atbilstoši horizontālās atdalošās līnijas novietojumam, un secinājumi par to, vai izdodas sasniegt klašu atdalamību, atsaucoties uz augšā esošajiem ekrānuzņēmumiem un veicot to analīzi>

## K-vidējo algoritms

### Orange rīkā pieejamo hiperparametru apraksts:

<pievienojiet tabulai rindas pēc nepieciešamības>

|  |  |
| --- | --- |
| **Hiperparametrs** | **Apraksts** |
|  |  |
|  |  |

<ekrānuzņēmums ar algoritmam uzstādītajām hiperparametru vērtībām>

### Eksperimentu apraksts

<ekrānuzņēmums Silueta koeficientam ar vismaz 5 dažādām k vērtībām>

<ekrānuzņēmums izkliedes diagrammai atbilstoši vislabākajai Silueta koeficienta vērtībai>

### Secinājumi no eksperimentiem:

<apraksts, vai klases ir atdalāmas, atsaucoties uz augšā esošajiem ekrānuzņēmumiem un veicot to analīzi>

## Noslēguma secinājumi

<pamatojoties uz divu algoritmu darbības salīdzinājumu, ir jādod secinājumi, vai datu kopā esošās klases ir labi vai slikti atdalāmas>

# III daļa

<šajā apakšnodaļā ir jāsniedz apraksts par pārraudzītās mašīnmācīšanās algoritmu lietojumu, to pavadot ar ekrānuzņēmumiem un atsaucēm uz izmantotajiem informācijas avotiem>

## Izvēlēto algoritmu apraksts

<apraksts brīvi izvēlētajiem algoritmiem un to izvēles motivācijai (izņemot mākslīgo neironu tīklu)>

### Pirmā algoritma nosaukums:

### Pirmā algoritma apraksts:

### Otrā algoritma nosaukums:

### Otrā algoritma apraksts:

## Hiperparametru apraksts

<jādod Orange rīkā pieejamo hiperparametru apraksts katram no algoritmiem, pievienojot rindas tabulai pēc nepieciešamības>

|  |  |
| --- | --- |
| **Hiperparameters** | **Apraksts un vērtības** |
| Mākslīgo neironu tīkli | |
| 1. Neironi paslēptos slāņos | Definēts kā i-tais elements, kas apzīmē neironu skaitu i-tajā slēptajā slānī |
| 1. Activation | Slēptā slāņa aktivizēšanas funkcija:  **Identity**: bezoperācijas aktivizēšana, noderīga, lai īstenotu lineāro sašaurinājumu  **Logistic**: loģistikas sigmoīda funkcija  **tanh**: hiperboliskā iedeguma funkcija  **ReLu**: rektificētas lineārās vienības funkcija |
| 1. Solver | Risinātājs svara optimizēšanai:  **L-BFGS-B**: optimizētājs kvaziņūtona metožu saimē  **SGD**: stohastiskā gradienta nolaišanās  **Adam**: uz stohastisku gradientu balstīts optimizētājs |
| 1. Alfa | Apmācības ātrums: L2 soda (regulēšanas termiņa) parametrs (**0 – 1000**) |
| 1. Iterāciju maksimālais skaits | Pašsaprotami |
| kNN | |
|  |  |
|  |  |
|  |  |
| Loģistiskā regressija | |
|  |  |
|  |  |

## Informācija par testa un apmācības datu kopām

A screenshot of a computer

AI-generated content may be incorrect.

*Datu kopas sadalīšana testa un apmācības datu kopās (Att. 3.1)*

### Datu objektu skaits apmācības datu kopā: **682**

### Datu objektu % proporcija apmācības datu kopā: **70%**

|  |  |  |
| --- | --- | --- |
| **Klases iezīme** | **Datu objektu skaits apmācības datu kopā** | **Datu objektu % proporcija apmācības datu kopā** |
| 1 | 266 | 39 % |
| 2 | 274 | 40 % |
| 3 | 142 | 21 % |

### Datu objektu skaits testa datu kopā: **291**

### Datu objektu % proporcija testa datu kopā: **30%**

|  |  |  |
| --- | --- | --- |
| **Klases iezīme** | **Datu objektu skaits testa datu kopā** | **Datu objektu % proporcija testa datu kopā** |
| 1 | 110 | 38 % |
| 2 | 132 | 45 % |
| 3 | 49 | 17 % |

## Eksperimenti ar mākslīgo neironu tīklu

|  |  |
| --- | --- |
| **Eksperiments** | **Hiperparametru vērtības** |
| 1.eksperiments | Neironi: 100; 100  Alfa: 0,1  Iterācijas: 1000 |
| 2.eksperiments | Neironi: 300; 300  Alfa: 0,1  Iterācijas: 1000 |
| 3.eksperiments | Neironi: 100; 100  Alfa: 1  Iterācijas: 1000 |
| 4.eksperiments | Neironi: 100; 100  Alfa: 0,05  Iterācijas: 10.000 |
| 5.eksperiments | Neironi: 500; 500  Alfa: 0,05  Iterācijas: 2000 |

*A screenshot of a computer

AI-generated content may be incorrect.*

*1.eksperimenta hiperparametru vērtības(Att. 3.2)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*1.esperimenta veiktspējas metrikas(Att. 3.3)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*2.eksperimenta hiperparametru vērtības(Att. 3.4)*

A screenshot of a computer

AI-generated content may be incorrect.

*2.esperimenta veiktspējas metrikas(Att. 3.5)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.eksperimenta hiperparametru vērtības(Att. 3.6)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.esperimenta veiktspējas metrikas(Att. 3.7)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*4.eksperimenta hiperparametru vērtības(Att. 3.8)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*4.esperimenta veiktspējas metrikas(Att. 3.9)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*5.eksperimenta hiperparametru vērtības(Att. 3.10)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*5.esperimenta veiktspējas metrikas(Att. 3.11)*

Secinājumi no eksperimentiem:

No visiem šajā darbā izmantotajiem algoritmiem neironu tīkls ir visnepiemērotākais mūsu uzdevumam, ņemot vērā tā prasīgo skaitļošanas jaudu. Šis uzdevums viņam ir pārāk vienkāršs. Neironu tīkls labi strādā tikai tad, ja tam ir pietiekami daudz neironu.

### Testēšanai izvēlētais modelis:

<norāde, kura eksperimenta modelis tiks izmantots testēšanas procesā>

## Eksperimenti ar kNN

|  |  |
| --- | --- |
| **Eksperiments** | **Hiperparametru vērtības** |
| 1.eksperiments | Kaimiņu skaits: 5 |
| 2.eksperiments | Kaimiņu skaits: 1 |
| 3.eksperiments | Kaimiņu skaits: 18 |

A screenshot of a computer

AI-generated content may be incorrect.

*1.eksperimenta hiperparametru vērtības(Att. 3.12)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*1.esperimenta veiktspējas metrikas(Att. 3.13)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*2.eksperimenta hiperparametru vērtības(Att. 3.14)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*2.esperimenta veiktspējas metrikas(Att. 3.15)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.eksperimenta hiperparametru vērtības(Att. 3.16)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.esperimenta veiktspējas metrikas(Att. 3.17)*

Secinājumi no eksperimentiem:

<secinājumi par modeļu veiktspēju veiktajos eksperimentos, atsaucoties uz augšā esošajiem ekrānuzņēmumiem un veicot to analīzi>

### Testēšanai izvēlētais modelis:

<norāde, kura eksperimenta modelis tiks izmantots testēšanas procesā>

## Eksperimenti ar loģistisko regressiju

|  |  |
| --- | --- |
| **Eksperiments** | **Hiperparametru vērtības** |
| 1.eksperiments | Regularizācijas veids: L2  Spēks: 1 |
| 2.eksperiments | Regularizācijas veids: L2  Spēks: 0,01 |
| 3.eksperiments | Regularizācijas veids: L2  Spēks: 1000 |
| 4.eksperiments | Regularizācijas veids: L1  Spēks: 1 |
| 5.eksperiments | Regularizācijas veids: None  Spēks: N/A |

*A screenshot of a computer

AI-generated content may be incorrect.*

*1.eksperimenta hiperparametru vērtības(Att. 3.18)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*1.esperimenta veiktspējas metrikas(Att. 3.19)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*2.eksperimenta hiperparametru vērtības(Att. 3.20)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*2.esperimenta veiktspējas metrikas(Att. 3.21)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.eksperimenta hiperparametru vērtības(Att. 3.22)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*3.esperimenta veiktspējas metrikas(Att. 3.23)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*4.eksperimenta hiperparametru vērtības(Att. 3.24)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*4.esperimenta veiktspējas metrikas(Att. 3.25)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*5.eksperimenta hiperparametru vērtības(Att. 3.26)*

*A screenshot of a computer

AI-generated content may be incorrect.*

*5.esperimenta veiktspējas metrikas(Att. 3.27)*

Secinājumi no eksperimentiem:

<secinājumi par modeļu veiktspēju veiktajos eksperimentos, atsaucoties uz augšā esošajiem ekrānuzņēmumiem un veicot to analīzi>

### Testēšanai izvēlētais modelis:

<norāde, kura eksperimenta modelis tiks izmantots testēšanas procesā>

## Apmācīto modeļu testēšanas rezultāti

<ekrānuzņēmums testēšanai izvēlēto modeļu veiktspējas metrikām>

Secinājumi pēc testēšanas:

<salīdzinājums apmācīto modeļu veiktspējai, atsaucoties uz augšā doto ekrānuzņēmumu un veicot tā analīzi >

# Izmantotie informācijas avoti

# Pielikums

<ja darba izstrādei tika izmantoti mākslīgā intelekta rīki, ir jādod ekrānuzņēmumi ar studentu komandas visiem ievadītajiem vaicājumiem (prompt) konkrētajā mākslīgā intelekta rīkā un visām rīka sniegtajām atbildēm; gan vaicājumiem, gan atbildēm ir jābūt redzamiem pilnībā un labā kvalitātē>