

V. Neironu tīkls kā matemātisks modelis un tā novērtēšana

V1. Modelis

Neironu tīkli ir definējami dažādos līmeņos, tomēr jebkurā līmenī tas ir saistīts ar jēdzienu matemātiskais modelis.

Matemātisks modelis ir kāda procesa vai objekta vispārīgs raksturojums, kas nodrošina relatīvi vienkāršu manipulāciju ar mainīgajiem, lai noteiktu šī procesa vai objekta uzvedību dažādās situācijās.

Tabula V1. Neironu tīkla detalizācijas līmeņi.

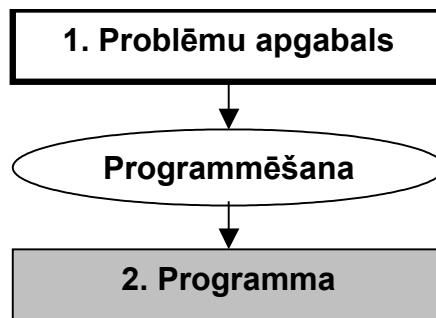
Līmenis un definīcija	Apraksts
1. Neironu tīkli kā skaitļošanas paradigma <u>Neironu tīkli</u> ir skaitļošanas paradigma, kuras ietvaros izmantojamus modeļus raksturo divas galvenās iezīmes: (1) relatīvi liels relatīvi vienkāršu skaitļošanas elementu kopums, kas ir savstarpēji savienoti, (2) apmācības process kā līdzeklis modeļa darbības iegūšanai.	Neironu tīkli kā alternatīva algoritmiskai problēmu risināšanai vai manipulācijām ar relāciju datu bāzēm
2. Neironu tīkla modelis (veids) <u>Neironu tīkla modelis</u> (veids) ir vispārīgs matemātisks modelis, kas kalpo kā ietvars konkrētākiem modeļiem, kas būtu iegūstami, noteiktā veidā veicot vispārīgā modeļa konfigurēšanu un apmācību	Noteiktu modeļu kopums (neironu tīklu paradigmas ietvaros (līmenis nr. 1)) ar precizētiem uzbūves un darbības principiem (arhitektūra, apmācības algoritms utt.). Piemēram, daudzslāņu perceptrons ar atgriezeniskās kļūdu izplatīšanas apmācības metodi vai Kohonena tīkls ar kvadrātisko topoloģiju
3. Konfigurēts, bet neapmācīts neironu tīkls <u>Neironu tīkls</u> ir skaitļošanas modelis, kas sastāv no relatīvi liela skaita savstarpēji savienotu skaitļošanas elementu un kas ir <u>apmācāms</u> noteiktas problēmas risināšanai.	Neironu tīkls (kas atbilst kādam no modeļiem līmenī nr. 2) ar fiksētu arhitektūru, apmācības procesa stratēģiju un dažādiem parametriem. Piemēram, radiālo bāzes funkciju tīkls ar 12 ieejām, 1 izeju, 4 RBF neironiem, fiksētu klāsterizācijas algoritmu, noteiktu apmācības koeficienta maiņas stratēģiju izejas slānī un citiem fiksētiem parametriem.
4. Apmācīts neironu tīkls	Neironu tīkls (no līmeņa nr. 3), kas ir izgājis

<u>Neironu tīkls</u> ir skaitļošanas sistēma, kas sastāv no relatīvi liela skaita savstarpēji savienotu skaitļošanas elementu un kas ir <u>apmācīts</u> noteiktas problēmas risināšanai.	apmācības procesu un validāciju ir gatavs noteiktas problēmas risināšanai
--	---

No praktiskā viedokļa, galvenais ir iegūt apmācītu neironu tīklu (līmenis nr. 4), lai risinātu kādu konkrētu problēmu. Tomēr pirms tam ir jāizvēlas vai jāizveido noteikts neironu tīkla modelis (līmenis nr. 2) un tas jānokonfigurē atbilstoši risināmajai problēmai (līmenis nr. 3), tikai tad var veikt apmācību, lai iegūtu ekspluatēšanai gatavu neironu tīklu.

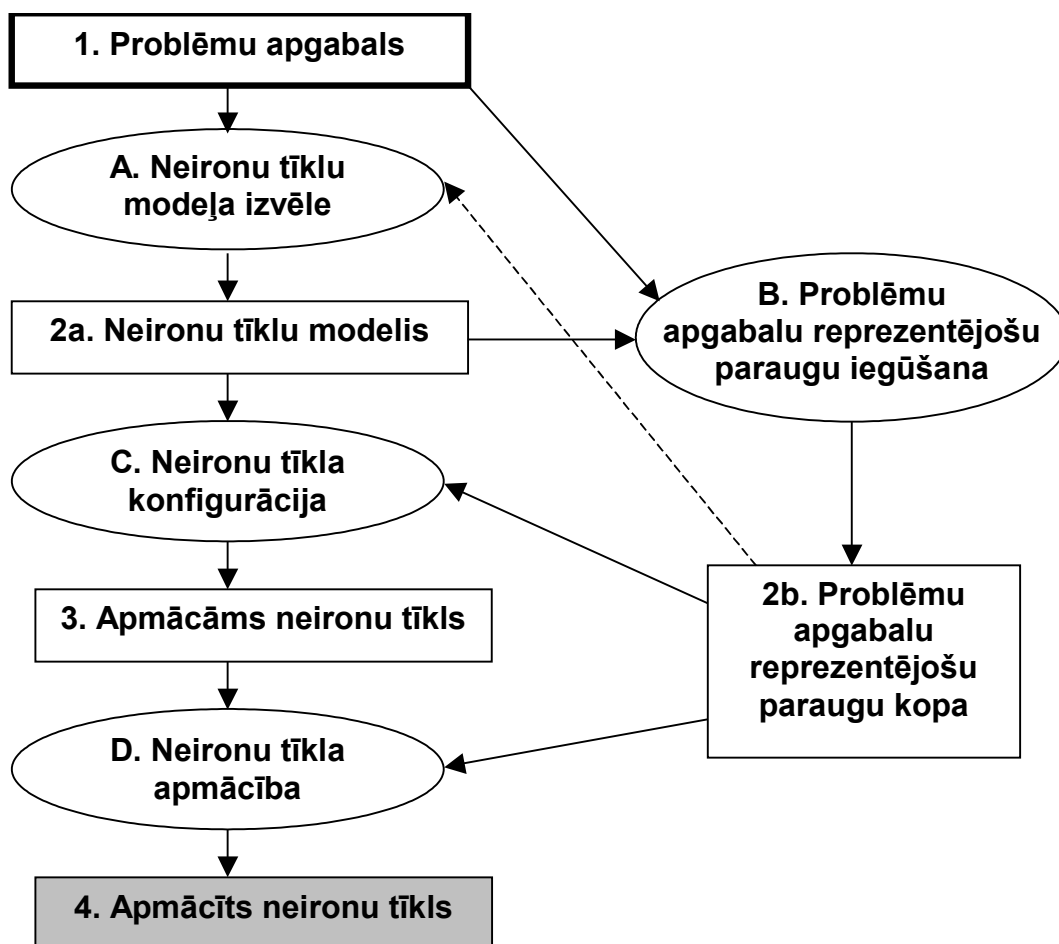
Neirona tīkla jēdziena daudzie līmeņi nosaka arī neirona tīkla vairāklīmeņu izveidošanas shēmu, kas būtiski atšķiras no tradicionālas datorprogrammu izveides.

Attēls V1 parāda klasisko problēmas risināšanas metodi – programmētājs (plašākā nozīmē) izanalizē problēmu apgabalu un izstrādā programmu, kas spēj veikt noteiktus uzdevumus. Problēmu apgabala analīze un abstrakcija ir programmētāja uzdevums un vispārīgā gadījumā nav automatizējama.



Attēls V1. Problēmu risinošas programmas izveidošanas klasiskā shēma

Neironu tīkla gadījumā problēmu apgabala analīze un abstrakcija tiek daļēji novirzīta uz neironu tīkla apmācības fāzi, kas prasa sarežģītāku un vairākpakāpju shēmu tā iegūšanai (attēls V2, kas atbilst tabulai V1).



Attēls V2. Neirona tīkla izveides shēma

Izstrādātāji, kas savās sistēmās izmanto neironu tīklus, bieži neveic pilnu neirona tīkla izveides shēmu, kāda ir attēlota attēlā V2 – pateicoties pieejamajiem rīkiem, praktiski pietiek ar soļiem B un D (paraugu iegūšana un apmācība), savukārt neironu tīklu pētnieki bieži nodarbojas tieši ar soļiem A un C (modeļa izvēle un konfigurācija, vai pat jauna modeļa izstrāde). Neironu tīkla izveides shēma ir pietiekoši komplicēta ne tikai kopumā, bet arī atsevišķas tās daļas prasa pietiekoši radošu un bieži vien arī resursu ietilpīgu darbu, kas sarežģītības ziņā var neatpalikt no programmēšanas. Problēmu apgabalu reprezentējošas paraugu kopas iegūšana ir viena no darbietilpīgākajām neironu tīkla izstrādē un daudzkārt ir atslēga uz veiksmīgu iznākumu.

Tabula V2. Neironu tīkla izveidošanas piemērs – neironu tīkla iegūšana, kas izrēķina divu skaitļu reizinājumu (objektu numerācija atbilstoši attēlam V2).

Solis	Rezultāts
Problēma	1. Divu skaitļu reizināšana
A. Neironu tīkla modeļa izvēle	2a. Vairākslāņu perceptrons ar kļūdu atgriezeniskās izplatīšanās apmācības algoritmu (<i>backpropagation</i>) un loģistisko aktivitātes funkciju

B. Paraugu iegūšana	2b. Paraugi, kas reprezentē divu skaitļu reizināšanu: 1, 2, 2 2, 2, 4 2, 3, 6 4, 5, 20 ...
C. Neironu tīkla konfigurācija	3. Konfigurēts, bet neapmācīts neironu tīkls, kurš pretendē uz to, ka, pareizi apmācīts, spēs risināt divu skaitļu reizināšanas problēmu noteiktā datu apgabalā. • izvēlēta datu reprezentācija neironu tīklam (normalizācija, denormalizācija), • fiksēts slāņu un neironu skaits, • fiksēts apmācības koeficients (<i>learning rate</i>) vai tā izmaiņas funkcija, • fiksēti citi parametri: maksimālais epohu skaits, maksimālā kļūda, inerces faktors (<i>momentum</i>), aktivitātes funkcijas slīpuma koeficients (<i>gain</i>) utt.
D. Neironu tīkla apmācība un validācija	4. Apmācīts un darbam gatavs neironu tīkls, kas spēj risināt divu skaitļu reizināšanas problēmu (arī uz paraugiem, kas nav ietilpuši apmācības kopā)

V2. Neironu tīkla novērtēšana

Neironu tīkla izveidošanas pamatmērķis ir iegūt **darbam gatavu neironu tīklu**, kas ir spējīgs korekti risināt doto problēmu ne tikai uz piemēriem, kas tikuši izmantoti apmācībā, bet arī uz citiem, respektīvi, lai neironu tīklam piemistu spēja vispārināt (*generalize*).

Vispārināšana (*generalization*) ir apmācāma modeļa spēja darboties ar paraugiem, kas nav tikuši izmantoti apmācības procesā.

Vispārināšanas spēja ir viena no svarīgākajām neironu tīklu īpašībām, bez kuras neironu tīkliem būtu daudz mazāka jēga uz pastāvēšanu.

Tai pat laikā daudzi mākslīgā intelekta pētnieki uzskata, ka neironu tīkli nemāk vispārināt, bet tikai atcerēties (*memorize*). Viens no piemēriem ir neironu tīklu nespēja tikt apmācītiem **paritātes** (*parity*) **problēmas** risināšanai. Šāda tipa diskusijas izsauc arī dažādas izpratnes par jēdzienu **vispārināšana**.

Lai iegūtu šādu, vispārināt spējīgu, neironu tīklu, jāizpildās šādiem priekšnosacījumiem:

- Neironu tīklam modeļa līmenī (līmenis 3 tabulā V3) jābūt spējīgam vispārināt, t.i., neironu tīklam jābūt tā konfigurētam, lai tas vispār būtu spējīgs vispārināt,
- Neironu tīklam (līmenis 4 tabulā V3) jābūt atbilstoši apmācītam (reprezentatīvi piemēri, korekts apmācības process).

Šie divi priekšnosacījumi tālāk izvērsti un paplašināti tabulā V3.

Pretēja īpašība vispārināšanas spējai ir **pārmērīga pielāgošanās** (*overfitting*), kam gan lielākas iespējas notikt ir tieši lineāros modeļos (piemēram, vienslāņa perceptronam).

Pārmērīgas pielāgošanās galvenais iemesls parasti ir pārāk liels neironu tīkla elementu skaits, salīdzinot ar apmācības paraugu skaitu. Bez tam pārmērīgu pielāgošanos var izsaukt nekorekta apmācības procesa organizācija – pārāk ilga apmācība ar pārāk nelielu apmācības koeficientu (*learning rate*).

Cita starpā, no pārmērīgas pielāgošanās cēloņiem izriet arī metodes, kas ļauj to novērst:

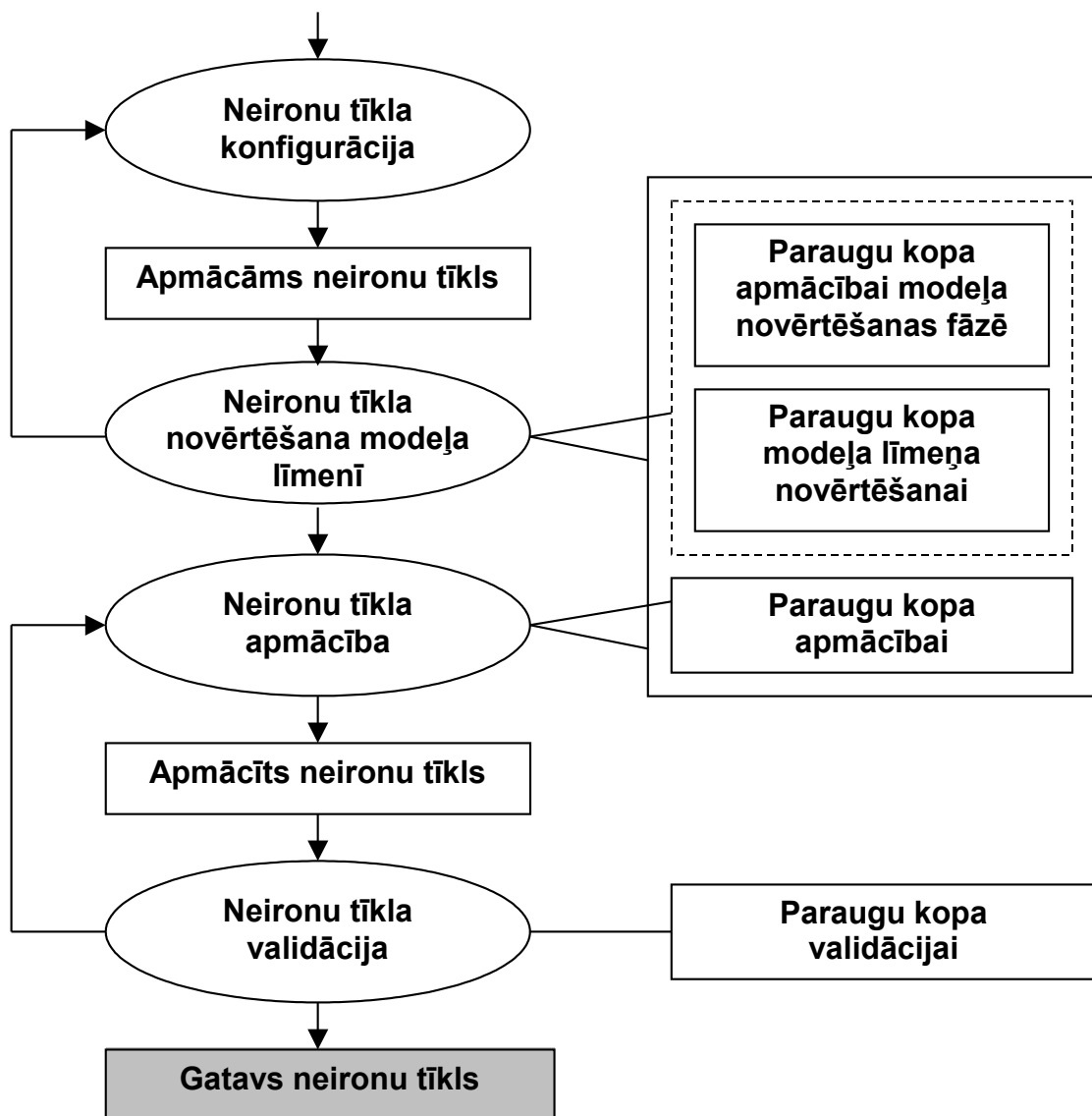
- korekta modeļa izvēle un konfigurēšana (t.sk., pietiekoši **mazs** neironu tīkla apjoms),
- apmācībai pieejamās paraugu kopas palielināšana,
- apmācības procesa ātrāka pārtraukšana (*early stopping*),
- vairāku neironu tīklu kombinēšana.

Tabula V3. Neironu tīkla novērtēšanas līmeņi (salīdzināt ar attēlu V2 un tabulām V1, V2)

Līmenis	Novērtēšanas iespēja
1. Neironu tīkli kā skaitļošanas paradigma	Novērtēšana filozofiskā līmenī : Iegūtie pozitīvie rezultāti pārējos līmeņos dod iespēju apgalvot, ka neironu tīkli kā paradigma ir izmantojami dažādu problēmu risināšanai
2. Neironu tīkla modelis (veids) Parādīt, ka šāda tipa neironu tīkli ir spējīgi darboties	Novērtēšana vispārīgā līmenī : • teorētiski, izmantojot matemātiskas metodes; • eksperimentāli, līmeņos 3 un 4
3. Konfigurēts, bet neapmācīts neironu tīkls	Neironu tīkla novērtēšana modeļa līmenī : Dažādas paraugu pārdalīšanas metodes,

Pierādīt, kā šādi konfigurēts neironu tīkls ir apmācāms <u>uz pieejamajiem paraugiem</u> tādējādi, lai būtu spējīgs risināt noteikta veida problēmas	piemēram, <i>cross-validation</i> (dod iespēju novērtēt, kā modelis spēj strādāt uz datiem, kas nav izmantoti apmācībā)
4. Apmācīts neironu tīkls Apmācīts un darbam gatavs neironu tīkls, kas spēj risināt doto problēmu	Neironu tīkla testēšana (validācija) skaitļošanas sistēmas līmenī: Neironu tīkls tiek pārbaudīts uz datiem, kas nav izmantoti apmācībā (un vēlams, nav bijuši izmantoti modeļa līmeņa novērtēšanā)

Faktiski, runājot par neironu tīkla novērtēšanu, parasti domā līmeņus 3 un 4 (tabula V3). Neironu tīkla izveidošana un novērtēšana šo divu līmeņu ietvaros shematiski parādīta attēlā V3.



Attēls V3. Neironu tīkla novērtēšana modeļa un skaitļošanas sistēmas līmenī

Attēlā V3 parādīts vienkāršota neironu tīkla izveides un novērtēšanas shēma. Tā nosaka, ka validācijas kopai (*validation set*) noteikti jābūt nošķirtām no pārējam paraugu kopām. Paraugu kopas novērtēšanai modeļa līmenī var būt nošķirtas (*holdout* metode, sk. zemāk) vai apvienotas, veicot paraugu dalīšanu (*K-fold cross-validation*).

Nākošā nodaļa apraksta populārāko metodi modeļa līmeņa novērtēšanai – *cross-validation* jeb validāciju ar krustošanu.

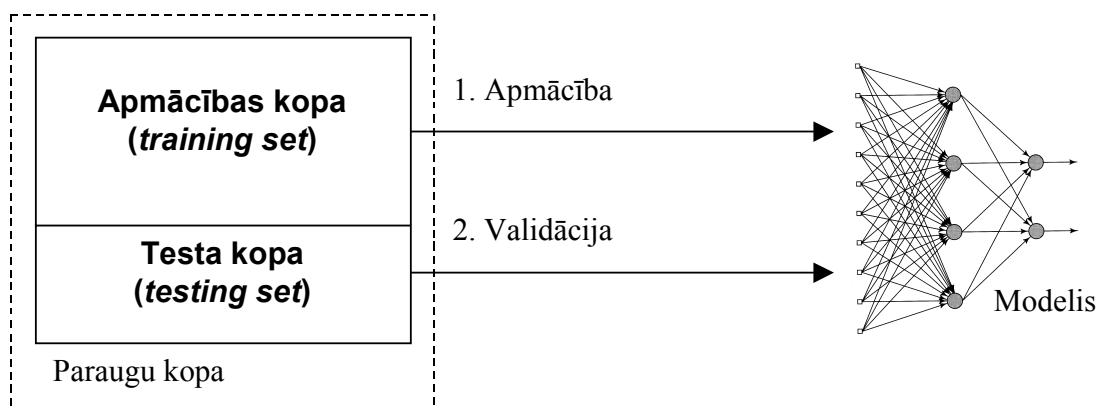
V3. Cross-validation (validācija ar krustošanu) – novērtēšana modeļa līmenī

Viena no labākajām modeļu novērtēšanas metodēm ir **validācija ar krustošanu** (*cross-validation*). Būtībā tā ir nevis novērtēšanas metode, bet gan **paraugu dalīšanas metode** apakškopās tā, lai viena no apakškopām tiktu izmantota neironu tīkla apmācībā, bet atlikusī daļa – tā derīguma novērtēšanā (validācijā).

V3.1. Holdout method (metode ar noturēšanu)

Metode ar noturēšanu (*the holdout method*) ir vienkāršākais *cross-validation* paveids, kas pēc būtības nemaz nav *cross-validation*, jo nekāda krustošana šeit nenotiek. Paraugi (*data set*) uz labu laimi tiek sadalīti divās kopās – apmācības kopā (*training set*) un testa kopā (*testing set*). Neironu tīkla apmācība notiek tikai, izmantojot apmācības kopu, bet pēc tam tas tiek pārbaudīts, izmantojot testa kopu. Parasti testa kopā ir ne vairāk kā viena trešdaļa no kopējā paraugu skaita.

Metodes trūkums ir vērtējuma augstā dispersija (*variance*) – vērtējums lielā mērā atkarīgs no tā, kādi paraugu ir izvēlēti apmācības kopai un kādi testa kopai.

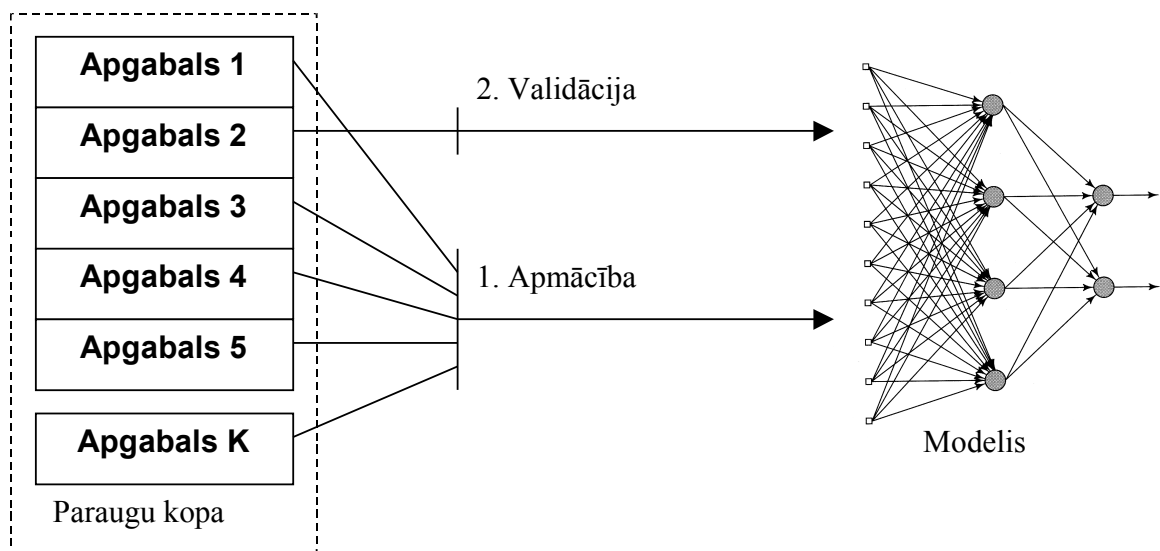


Attēls V4. Metode ar noturēšanu (holdout method)

V3.2. K-fold cross-validation (K-apgabalu validācija ar krustošanu)

K-apgabalu validācijā ar krustošanā (*K-fold cross-validation*) paraugu kopā tiek dalīta K daļās (apgabalos). Tad K reizes tiek veikta novērtēšana pēc *holdout* metodes – viens datu apgabals tiek izvēlēts par testa kopu, bet pārējie $K-1$ apgabali – par apmācības kopu. Lai iegūtu kopējo novērtējumu, visi K sesiju rezultāti tiek noteiktā veidā kombinēti, piemēram, ņemta to vidējā vērtība.

Izmantojot **K-apgabalu validāciju ar krustošanu**, katrs paraugs tieši vienu reizi atrodas kādā testa kopā un tieši $K-1$ reizi – kādā apmācības kopā. Metodes priekšrocība ir mazāka dispersija nekā *holdout* metodei un tā var vēl tikt samazināta, palielinot K vērtību. Metodes trūkums ir liels nepieciešamais skaitļošanas apjoms – K reizes lielāks nekā *holdout* metodei. Vēl viens šīs metodes variants ir tāds, ka sākotnēji paraugu dalīšana apgabalos nenotiek (apgabali netiek fiksēti), bet apmācības kopas un testa kopas izvēle pēc nejaušības principa notiek **pirms katras** novērtēšanas sesijas.



Attēls V5. K-apgabalu validācijas ar krustošanu (K-fold cross-validation) vienas novērtēšanas sesijas veikšana

V3.3. Leave-one-out cross-validation (vienu elementu izslēdzoša validācija ar krustošanu)

Vienu elementu izslēdzoša validācija ar krustošanu (*leave-one-out cross-validation*) ir parastās k -apgabalu validācijas galējais variants, kad K ir vienāds ar N – paraugu kopējo skaitu. Tādējādi modeļa novērtējums ir izvērsts N sesiju garumā, un katrā sesijā tieši viens paraugs tiek ievietots testa kopā, bet visi pārējie – apmācības kopā.

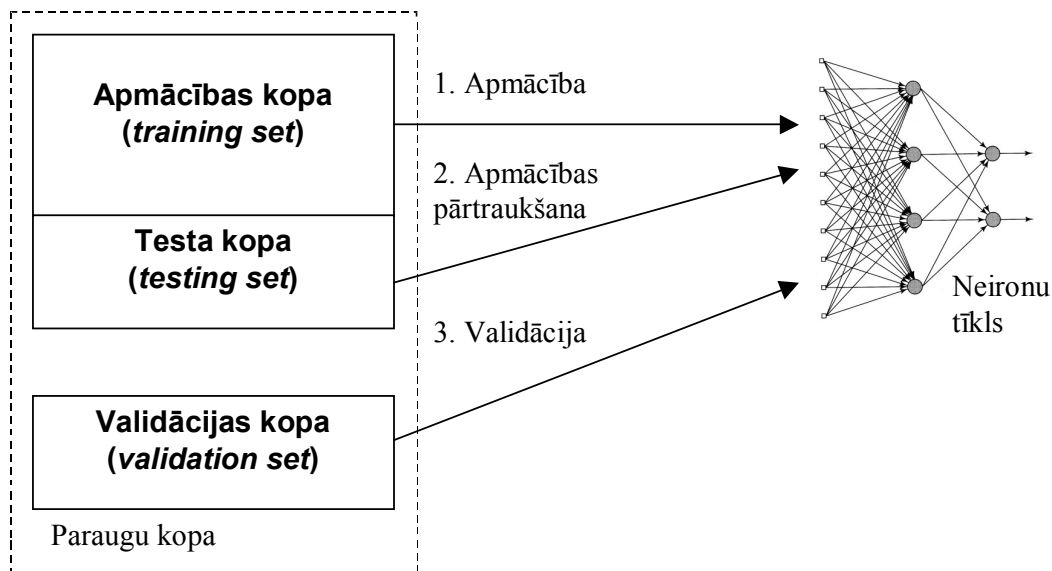
V4. Apmācības procesa ātrāka pārtraukšana

Viena no metodēm neironu tīkla vispārināšanas spējas uzlabošanai ir **apmācības procesa ātrāka pārtraukšana** (*early stopping*), kas nozīmē, ka kontrolēta apmācības process tiek pārtraukts ātrāk nekā neironu aproksimācijas kļūda (jeb kļūda uz ieejas paraugiem) ir samazinājusies zem noteiktas vērtības.

Šim nolūkam no apmācības kopas tiek nodalīta testa kopa, un pēc katras epohas neironu tīkls tiek pārbaudīts arī uz testa kopas. Apmācība tiek pārtraukta, kad neironu tīkla kļūda pret testa kopas paraugiem sāk palikt lielāka. Tādējādi aproksimācijas kļūda kalpo kā svaru vērtību izmaiņas virzītājs apmācības procesā, bet vairs ne kā beigšanas kritērijs.

Līdz ar to neironu tīkla apmācībā un validācijā tiek lietotas trīs dažādas paraugu kopas:

- apmācības kopa – neironu tīkla apmācībai,
- testa kopa – lai noteiktu apmācības procesa beigas,
- validācijas kopa – lai novērtētu iegūto neironu tīklu, izmantojot kādu no modeļu validācijas metodēm (tas pats, kas testa kopa nodaļā V4).



Attēls V6. Apmācības procesa ātrāka pārtraukšana