

RĪGAS TEHNISKĀ UNIVERSITĀTE
Datorzinātnes un informācijas tehnoloģijas fakultāte
Informācijas tehnoloģijas institūts

Andrejs Bondarenko

Doktora studiju programmas “Informācijas tehnoloģija” doktorants

**METODOLOĢIJAS IZSTRĀDE ZINĀŠANU
IZGŪŠANAI NO APMĀCĪTIEM
MĀKSLĪGAJIEM NEIRONU TĪKLIEM**

Promocijas darba kopsavilkums

Zinātniskie vadītāji:

profesors *Dr. habil. sc. comp.*

ARKĀDIJS BORISOVS

profesore *Dr. sc. ing.*

LUDMILA ALEKSEJEVA

RTU Izdevniecība
Rīga 2020

Bondarenko, A. Metodoloģijas izstrāde zināšanu izgūšanai no apmācītiem mākslīgajiem neironu tīkliem. Promocijas darba kopsavilkums. Rīga: RTU Izdevniecība, 2020. 45 lpp.

Iespiests saskaņā ar 2019. gada 25. novembra Informācijas tehnoloģijas institūta padomes sēdes lēmumu, protokols Nr. 12100-2/8.

ISBN 978-9934-22-006-7 (print)

ISBN 978-9934-10-940-9 (pdf)

**PROMOCIJAS DARBS IZVIRZĪTS
ZINĀTNES DOKTORA GRĀDA IEGŪŠANAI
RĪGAS TEHNISKAJĀ UNIVERSITĀTĒ**

Promocijas darbs zinātnes doktora grāda (*Ph. D.*) iegūšanai tiek publiski aizstāvēts 2020. gada 18. maijā Rīgas Tehniskās universitātes Datorzinātnes un informācijas tehnoloģijas fakultātē, Sētas ielā 1, 202. auditorijā.

OFICIĀLIE RECENZENTI

Profesors *Dr. habil. sc. ing.* Jānis Grundspenķis,
Rīgas Tehniskā universitāte, Latvija

Profesors *Dr. sc. ing.* Pēteris Grabusts,
Rēzeknes Tehnoloģiju akadēmija, Latvija

Profesors *Dr. habil. sc. ing. Yevgeniy Bodyanskiy,*
Harkovas Nacionālā radioelektronikas universitāte, Ukraina

APSTIPRINĀJUMS

Apstiprinu, ka esmu izstrādājis šo promocijas darbu, kas iesniegts izskatīšanai Rīgas Tehniskajā universitātē zinātnes doktora grāda (*Ph. D.*) iegūšanai. Promocijas darbs zinātniskā grāda iegūšanai nav iesniegts nevienā citā universitātē.

Andrejs Bondarenko (paraksts)

Datums:

Promocijas darbs ir uzrakstīts angļu valodā, tajā ir ievads, piecas nodaļas, secinājumi, literatūras saraksts, 37 attēli, 22 tabulas, pieci pielikumi, kopā 158 lappuses, ieskaitot pielikumus. Literatūras sarakstā ir 150 nosaukumu.

SATURS

DARBA VISPĀRĪGS APRAKSTS	5
Ievads	5
Tēmas aktualitāte	5
Darba mērķis un uzdevumi	5
Pētījuma objekts un priekšmets	6
Pētījuma hipotēzes	6
Pētījuma metodes	6
Darba zinātniskais jaunieguvums un vērtība	6
Darba praktiskais nozīmīgums	7
Aprobācija	7
Promocijas darba struktūra un saturs	10
DARBA NODAĻU KOPSAVILKUMS	11
1. NEIRONU TĪKLI UN ZINĀŠANU IZGŪŠANA	11
1.1. Neironu tīklu un zināšanu izgūšanas pārskats	11
1.2. Zināšanu izgūšanas metožu veidi	12
2. NEIRONU TĪKLU APGRIEŠANA	13
2.1. Jūtīguma analīzē balstīts apgriešanas algoritms	13
2.2. Izstrādātā apgriešanas algoritma validācija	14
3. LĒMUMU KOKA IZGŪŠANA NO DAUDZSLĀŅU PERCEPTRONA	17
3.1. Zināšanu izgūšanas pārskats	17
3.2. Zināšanu izgūšanas algoritms	18
3.3. Izstrādātā algoritma validācija	19
4. OPTIMIZĀCIJĀ BALSTĪTAS PIEEJAS LIKUMU IZGŪŠANAI	22
4.1. Eliptisko likumu izgūšana no RBF neironu tīkliem	22
Optimizācijas problēma	22
Eksperimenti un rezultāti	23
4.2. Likumu izgūšana no apgabaliem, kas ierobežoti ar hiperplaknēm	23
Pieejas pārskats	24
Ja–tad likumu izgūšanas algoritms	24
Eksperimenti un rezultāti	25
5. ZINĀŠANU IZGŪŠANAS METODOLOĢIJA	27
5.1. Metodoloģijas izstrāde	27
5.2. Precīzi vai saprotami likumi	29
REZULTĀTI UN SECINĀJUMI	32
IZMANTOTĀS LITERATŪRAS SARAKSTS	34

DARBA VISPĀRĪGS APRAKSTS

Ievads

Mākslīgie neironu tīkli (MNT) tiek plaši izmantoti mašīnmācībā. Tie ir jaudīgi nelineāri modeļi, kurus var apmācīt uzraudzītā, daļēji uzraudzītā un neuzraudzītā veidā. Universāls mašīnmācīšanās klasifikators, ko var izmantot visos scenārijos, neeksistē, taču MNT bieži pārspēj citus klasifikatorus. Tomēr ir grūti izskaidrot, kā MNT pieņem klasifikācijas lēmumu. Mākslīgais neironu tīkls ir “melnā kaste”, un nav skaidrs, kā šāds klasifikators darbojas, un tas nopietni ierobežo MNT lietojamību. Promocijas darbs ir veltīts tādu pieeju izstrādei, kas ļauj izgūt zināšanas no apmācīta MNT klasifikatora.

Tēmas aktualitāte

Klasifikācijas modeļa saprotamība ir būtiska prasība tādās svarīgās jomās kā kodolenerģija, medicīna, finanses un citās. Papildus eksistē arī likuma prasības, piemēram, Eiropas Savienības GDPR 2018 likums [118], kas nosaka, ka visiem kritiskiem dzīves ietekmējošiem algoritmiskajiem lēmumiem jābūt izskaidrojamiem. Izskaidrojamība ļauj validēt klasifikatoru, nodrošināt to, ka neeksistē diskriminācija, un izgūt jaunas zināšanas. Zināšanu izgūšanas jomā publikācijas ir, taču nav pieejami lietošanai gatavi algoritmi un risinājumi. Turklāt, kā tika atklāts, zinātnisku rezultātu reproducējamība ir liela problēma, tāpēc instrumentu izstrāde MNT klasifikatoru izskaidrošanai var ievērojami uzlabot un paplašināt to lietojamību.

Darba mērķis un uzdevumi

Pētījuma **mērķis** ir izstrādāt apmācītas MNT apgriešanas un zināšanu izgūšanas (ZI) algoritmus un apvienot tos zināšanu izgūšanas metodoloģijā. Šādai metodoloģijai jāapraksta apmācītus mākslīgus neironu tīklus, izmantojot ja–tad likumu kopu, bināru klasifikācijas koku vai vienādojumu kārtu kopu. Pētījuma mērķa sasniegšanai ir jāatrisina šādi **uzdevumi**.

1. Analizēt zinātniskajā literatūrā un citos pētījumos izmantotās, kā arī aprakstītās pieejas zināšanu reprezentācijas un izguves pieejas uzdevumu risināšanai.
2. Izpētīt eksistējošās MNT apgriešanas pieejas pozitīvās un negatīvās īpašības, izstrādāt uzlabotu apgriešanas pieeju un aprobēt to.
3. Izstrādāt un aprobēt pieeju zināšanu izgūšanai no MNT.
4. Formulēt un aprobēt optimizācijas uzdevumus, kuru risinājums ļaus izgūt ja–tad un eliptiskos likumus no MNT un no radiālo bāzes funkciju neironu tīkla (RBFNT).
5. Izstrādāt un aprobēt vispārinātu metodoloģiju zināšanu izgūšanai no MNT.

Pētījuma objekts un priekšmets

Pētījuma **objekts** ir apmācīts mākslīgā neironu tīkla klasifikācijas lēmumu skaidrojums, pētījuma **priekšmets** – mašīnmācīšanās metodes un zināšanu izgūšanas pieejas.

Pētījuma hipotēzes

Mākslīgo neironu tīklu apgriešanas un likumu izgūšanas metodes izstrādes gaitā pārbaudei ir izvirzītas vairākas hipotēzes.

1. Uzlabota MNT apgriešanas pieeja, kas balstīta uz jutīguma analīzi, izvairās no lokāliem minimumiem un ļauj kontrolēt klasifikācijas kļūdas palielināšanos.
2. Ieejas datu diskretizācija, kas izgūta no visu neironu izejas vērtībām, ļauj izgūt klasifikācijas koku, kas aproksimē daudzslāņu MNT ar kontrolējamu precizitāti.
3. Izliekta optimizācijas uzdevuma risināšanā izgūtie nosacījuma likumi ļauj aproksimēt ar hiperplaknēm ierobežotu apgabalu.
4. Neizliekta optimizācijas uzdevuma risināšanā izgūtie eliptiskie likumi ļauj aproksimēt RBFMNT.

Pētījuma metodes

Promocijas darba teorētiskajā izstrādē ir izmantotas šādas metožu klases: mašīnmācīšanās, matemātiskā statistika, optimizācijas teorija un eksperimentālo pētījumu metodoloģijas. Informācijas apkopošanai par esošajām pieejām priekšmetu jomā izmantots literatūras apskats un analīze.

Darba zinātniskais jauniegums un vērtība

Pētījuma **zinātniskā novitāte** balstās uz esošo zināšanu izgūšanas metožu pārskatīšanu un jauno algoritmu izstrādi. Ir izstrādātas četras metodes, ko var izmantot, klasifikācijas modeļa validācijai, izpratnei un izskaidrojumam. Zinātniskā novitāte un sasniegumi ir šādi.

1. Izstrādāts uz jutīguma analīzes balstīts mākslīga neironu tīkla apgriešanas algoritms ar vairākām modifikācijām, kas ļauj tam izvairīties no lokāliem minimumiem. Veikts svaru anulēšanas un neironu apgriešanas salīdzinājums ar ieteikumiem pieejas izvēlei.
2. Izstrādāta metode bināra klasifikācijas lēmumu koka izgūšanai no apmācīta MNT klasifikatora. Piedāvātais risinājums ir eksperimentāli pārbaudīts.
3. Izstrādāta pieeja ja–tad likumu izgūšanai no apgabala, kas ierobežots ar hiperplaknēm. Lai gan tika parādīts, ka pieejai ir problēmas, apstrādājot lieldimensiju datu kopas, to var izmantot datu kopām ar nelielu ieejas datu dimensiju skaitu.
4. Izstrādāta pieeja eliptisko likumu izgūšanai no divdimensiju vai trīsdimensiju RBF neironu tīkla. Lai gan šī pieeja parādīja, ka tā ir pakļauta dimensiju lāstam, izstrādātā optimizācijas problēma bija pareiza pieeja eliptisko likumu izgūšanai divu vai trīs dimensiju gadījumā. Lielāku dimensiju skaitu var atbalstīt, uzlabojot algoritmu.

5. Balstoties uz veiktajiem pētījumiem un eksperimentiem, tika izstrādāta likumu izgūšanas metožu izmantošanas metodoloģija ar ieteikumiem gadījumos, kad viena pieeja būtu jāizvēlas citas vietā.

Darba praktiskais nozīmīgums

Praktiskais jaunieguvums ietver izstrādāto algoritmu programmatūras realizācijas, to eksperimentālo validāciju un novērtēšanu. Pilns praktisko sasniegumu saraksts ir šāds.

1. Izstrādāts apraksts dažādām zināšanas reprezentācijas shēmām un dotas rekomendācijas to izvēlei.
2. Izstrādāta metodoloģija, kas ļauj izgūt precīzus vai vienkāršus likumus.
3. Salīdzinātas neironu apgriešana un svaru anulēšana, dotas rekomendācijas pieeju izvēlei.
4. Darbā aplūkoto klasifikatoru (MNT, RBFMNT un ar hiperplaknēm ierobežotu apgabalu klasifikators) lietojamība palielinājās dažādām vidēm, jo tagad ir iespējams tos validēt, izprast modeļu piemērotību, saprast klasifikācijas lēmumu un atklāt jaunas zināšanas.
5. Izstrādāta programmatūra *Matlab* vidē, kā arī *Lua* balstītā *Torch7* dziļās mācīšanās (*DM*) bibliotēka un *Python* balstītie *PyTorch DL* bibliotēkas paplašinājumi. *PyTorch* versija ir piemērojama vidēja lieluma datu kopu izpētei.

Aprobācija

Promocijas darbā veiktie pētījumi un to galvenie rezultāti ir prezentēti 13 starptautiskās zinātniskās konferencēs.

1. RTU 60th International Scientific Conference. Latvia, Riga, 10–11 October 2019.
2. RTU 59th International Scientific Conference. Latvia, Riga, 10–12 October 2018.
3. RTU 57th International Scientific Conference. Latvia, Riga, 17–21 October 2016.
4. 10th International Scientific and Practical Conference “Environment. Technology. Resources”. Latvia, Rezekne, 18–20 June 2015.
5. RTU 55th International Scientific Conference. Latvia, Riga, 14–16 October 2014.
6. The 6th International Conference “Applied Information and Communication Technology”. Latvia, Jelgava, 25–26 April 2013.
7. RTU 53rd International Scientific Conference. Latvia, Riga, 11–12 October 2012.
8. International Conference “Information Intelligent Systems”. Ukraine, Kharkiv, 17–19 April 2012.
9. 10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases (AIKED’11). United Kingdom, Cambridge, 20–22 February 2011.
10. 17th International Conference on Soft Computing MENDEL. Czech Republic, Brno, 15–17 June 2011.
11. 15th International Conference on Knowledge-Based and Intelligent Information and

- Engineering Systems (KES 2011). Germany, Kaiserslautern, 12–14 September 2011.
12. 16th International Conference on Soft Computing MENDEL'10. Czech Republic, Brno, 23–25 June 2010.
13. RTU 51st International Scientific Conference, Latvia, Riga, 11–15 October 2010.

Promocijas darbā veiktie pētījumi un to rezultāti ir atspoguļoti 15 publikācijās starptautiskos zinātniskos izdevumos.

1. Bondarenko, A. Controlling Complexity and Accuracy of Classification Decision Tree Extracted from Trained Artificial Neural Network. In: *60th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), 2019*. Available from: doi:10.1109/ITMS47855.2019.8940739. **Indexed in: Scopus.**
2. Bondarenko, A., Aleksejeva, L. Methodology for Knowledge Extraction from Trained Artificial Neural Networks. *Information Technology and Management Science*. 2018, vol. 21, pp. 6–14. Available from: doi:10.7250/itms-2018-001.
3. Bondarenko, A., Aleksejeva, L. Workflow for Knowledge Extraction from Neural Network Classifiers. In: *59th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), 2018*. Available from: doi:10.1109/ITMS.2018.8552964. **Indexed in: Scopus.**
4. Bondarenko, A., Aleksejeva, L., Jumutcs, V., Borisovs, A. Classification Tree Extraction from Trained Artificial Neural Networks. *Procedia Computer Science*, 2017, vol. 104, pp. 556–563. Available from: doi:10.1016/j.procs.2017.01.172. **Indexed in: Scopus, Web of Science. Cited: 9.**
5. Bondarenko, A., Borisovs, A., Aleksejeva, L. Neurons vs Weights Pruning in Artificial Neural Networks. In: *Environment. Technology. Resources: Proceedings of the 10th International Scientific and Practical Conference, Latvia, Rezekne, 18–20 June 2015*. Vol. 3. Rezekne: Rezekne Higher Education Institution, 2015, pp. 22–28. Available from: doi:10.17770/etr2015vol3.166, **Indexed in: Scopus. Cited: 2.**
6. Bondarenko, A., Borisovs, A. Artificial Neural Network Generalization and Simplification via Pruning. *Information Technology and Management Science*. 2014, vol. 17, pp. 132–137. Available from: doi:10.1515/itms-2014-0020.
7. Bondarenko, A., Borisovs, A. Elliptical Rule Extraction from a Trained Radial Basis Function Neural Network. In: *The 6th International Conference “Applied Information and Communication Technology” (CD-ROM)*, Latvia, Jelgava, LUA Faculty of Information Technology, 25–26 April 2013. **Indexed in: Web of Science. Cited: 1.**
8. Bondarenko A., Borisov A. Research on the Classification Ability of Deep Belief Networks on Small and Medium Datasets. *Scientific Journal of Riga Technical University, Information Technology and Management Science*, 2013, vol. 16, pp. 60–65. Available from: doi:10.2478/itms-2013-0009. **Indexed in: EBSCO, De Gruyter, Google Scholar, ResearchGate. Cited: 2.**
9. Bondarenko, A., Borisovs, A. Knowledge Extraction from Piecewise-Linear Approximation of Multi-Surface Classifier. In: *International Conference “Information Intelligent Systems”, Kharkov, Ukraine, 17–19 April 2012*. Vol. 6, pp. 5–6.

10. Bondarenko A., Borisov A. The Extraction of Elliptical Rules from the Trained Radial Basis Function Neural Network. *Information Technology and Management Science*. 2012, vol. 15, pp. 161–165. Available from: doi:10.2478/v10313-012-0027-2.
11. Bondarenko A., Jumutc V. Extraction of Interpretable Rules from Piecewise-Linear Approximation of a Nonlinear Classifier using Clustering-Based Decomposition. *Proceedings of the 10th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases (AIKED'11), United Kingdom, Cambridge, 22–22 February 2011*. Cambridge: 2011, pp.145–149. **Indexed in: Scopus.**
12. Bondarenko, A., Zmanovska, T., Borisovs, A. Piece-Wise Classifier Application to RBF Neural Network Rules Extraction. In: *17th International Conference on Soft Computing (MENDEL'11), Czech Republic, Brno, 15–17 June 2011*. Brno: Brno University of Technology, 2011, pp. 170–176. **Indexed in: Scopus, Web of Science.**
13. Jumutcs, V., Bondarenko, A. Polytope Classifier: A Symbolic Knowledge Extraction from Piecewise-Linear Support Vector Machine. In: *Knowledge-Based and Intelligent Information and Engineering Systems: 15th International Conference (KES 2011): Proceedings, Part 1, Germany, Kaiserslautern, 12–14 September 2011*. Berlin: Springer Berlin Heidelberg, 2011, pp. 62–71. Available from: doi:10.1007/978-3-642-23851-2_7. **Indexed in: Scopus, Web of Science, ResearchGate, SpringerLink.**
14. Bondarenko A., Borisov A. Decompositional Rules Extraction Methods from Neural Networks. In: *Proceedings of the 16th International Conference on Soft Computing MENDEL'10, Czech Republic, Brno, 23–25 June 2010, Brno: University of Technology, 2010*, pp. 256–262. ISBN 9788021441200. **Indexed in: Scopus, Web of Science.**
15. Bondarenko A., Borisov A. Research of Artificial Neural Networks Abilities in Printed Words Recognition. *Scientific Journal of Riga Technical University, Information Technology and Management Science*. 2010, vol. 44, issue 5, pp. 124–129.

Promocijas darba rezultāti izmantoti trīs projektos.

1. Latvijas Zinātnes padomes finansēts projekts Nr. 09.1564 “Simulācijas un skaitļošanas intelekta metodes loģistikas un e-biznesa optimizācijai” (2010.–2012.). Projekta vadītājs *Dr. habil. sc. ing.* prof. Y. Merkurjevs.
2. Latvijas Zinātnes padomes finansēts projekts “Ātrā mikroorganismu aktivitātes noteikšana ar optisko bezkontakta metodi” Nr. LZP-2018/2-0051 (2018.–2020.). Projekta vadītājs *Dr. phys.* A. Lihačovs.
3. Latvijas Zinātnes padomes finansēts projekts “Ādas vēža agrīnas diagnostikas precizitātes uzlabošana ar neironu tīkliem” Nr. LZP-2018/2-0052 (2018.–2020.). Projekta vadītāja *Dr. phys.* I. Lihačova.

Promocijas darba struktūra un saturs

Promocijas darbā ir ievads, piecas nodaļas, secinājumi, literatūras avotu saraksts un pielikumi. Darbs ir rakstīts angļu valodā.

Ievadā ir pamatota izvēlētās tēmas aktualitāte, formulēts pētījumu mērķis un uzdevumi, izvirzītas hipotēzes, uzskaitītas promocijas darba izstrādē izmantotās zinātniskās metodes, aprakstīts pētījumu zinātniskais jaunieguvums un darbā izgūto rezultātu praktiskā vērtība, kā arī sniegts darba aprobācijas raksturojums.

Pirmajā nodaļā ir aprakstīti mākslīgie neironu tīkli un piedāvāts pārskats par zināšanu izgūšanas pieejām.

Otrajā nodaļā aprakstīts pirmais solis zināšanu izgūšanai no mākslīgajiem neironu tīkliem (MNT) – MNT apgriešana. Nodaļā ir apkopoti eksistējošie apgriešanas algoritmu tipi un prezentēts izstrādātais algoritms.

Trešajā nodaļā ir parādīta izstrādātā pieeja bināro klasifikācijas lēmumu koku izgūšanai no apmācīta daudzslāņu perceptrona. Šajā nodaļā aprakstīts izstrādātais algoritms, sniegts tā pseidokods un eksperimentāla validācija.

Ceturtajā nodaļā ir parādītas izstrādātās optimizācijā balstītās metodes ja–tad un eliptisko likumu izgūšanai apgabaliem, kas ierobežoti ar hiperplaknēm, un radiālās bāzes funkcijas neironu tīkla (RBFNT).

Piektajā nodaļā ir aprakstīta izstrādātā metodoloģija, kas ļauj izvēlēties vienu no aprakstītajām metodēm un nosaka, kā rīkoties gadījumos, kad izgūtie likumi nav pietiekami precīzi vai nav pietiekami saprotami (tos daudzums ir pārāk liels).

Rezultātu un secinājumu nodaļā ir ietverti mērķi, uzdevumi un hipotēzes, publicēti secinājumi, apkopots darba zinātniskais jaunieguvums un praktiskā novitāte.

DARBA NODAĻU KOPSAVILKUMS

1. NEIRONU TĪKLI UN ZINĀŠANU IZGŪŠANA

Ņemot vērā definēto mērķi un uzdevumus, pirmajā nodaļā veikts mašīnmācīšanās nozares un mākslīgo neironu tīklu (MNT) kā izpētes objekta pārskats. Šajā nodaļā apskatīti MNT veidi, to apmācība un lietošana. Tiek apskatīti zināšanu izgūšanas (ZI) algoritmi un zināšanu reprezentācijas shēmas.

1.1. Neironu tīklu un zināšanu izgūšanas pārskats

Padziļināts ievads bioloģiskajos un mākslīgajos neironu tīklos ir dots [90]. Autori [40] ir izvirzījuši ticamu skaidrojumu, aprakstot, kā neironu tīkli var darboties un tuvināt vienkāršas lineāras funkcijas. Vēlāk pētnieki, kurus iedvesmojuši bioloģiskie neironu tīkli, ir ierosinājuši dažādas mākslīgo neironu tīklu arhitektūras.

Pašiem datiem ir maza vērtība. Savukārt informācija ir dati, kas saistīti ar konkrētu kontekstu, kas šiem datiem piešķir zināmu nozīmi un ļauj mums redzēt sakarības. Zināšanas ir informācija, kas ir sakārtota tā, lai ļautu redzēt konkrētus modeļus. Gudrība ir vēl abstraktāka un ļauj izprast vispārējos principus. Vispārējo mākslīgā intelekta (MI) sistēmu ietvaros zināšanām bija jābūt attēlotām tā, lai tām varētu lietot spriešanas mehānismus. Zināšanas ir līmenis, kas ir aprakstīts šajā darbā. Darbā izmantota jēdziena “likumu izgūšana” definīcija, kas ierosināta [73]. Šī definīcija ir plašāka nekā citas, piemēram, tā, kas dota [96], un uzsver faktu, ka izgūtie likumi var būt reprezentēti ne tikai leksiski, bet arī dažādos citos veidos:

“Ņemot prognozējošo modeli, kas ir “melna kaste”, un datus, uz kuriem tas tika apmācīts, sagatavot modeļa hipotēzes aprakstu, kas ir saprotams, bet precīzi aproksimē prognozējošā modeļa uzvedību.”

Likumu izgūšanas algoritmus atkarībā no konteksta var noregulēt tā, lai izgūtu saprotamākus, tātad kompakus, vai precīzākus (un sarežģītākus) likumus.

Saskaņā ar [76] ierosināto zināšanu izgūšanas taksonomiju pastāv trīs veidu zināšanu izgūšanas algoritmi. Ceturtā zināšanu izgūšanas algoritmu grupa (kompozīcijas) tika ierosināta [139]. Šie veidi ir: dekompozicionāls, pedagoģisks, eklektisks, kompozicionāls.

Atkarībā no izmantošanas scenārija jāpieņem lēmums par kompromisu starp lasāmību un augstu klasifikācijas precizitāti. Šis lēmums ļaus izvēlēties vispiemērotāko likumu veidu. Likumu veidi (sīkāku aprakstu skat. [22], [73]) ir šādi: ja–tad / ja–tad–vēl likumi; M no N likumi; vienādojuma likumi; izplūdušie likumi.

Ja–tad likumi un lēmumu koki ir plaši izplatīti, jo tie ir viegli iebūvējami, ar labu saprotamību un tiem piemīt izteismīgums un kompakums. Vienādojumu likumi ir izteismīgāki, bet tos ir grūti saprast un interpretēt. Divas pēdējās grupas – izplūdušie likumi un M no N likumi – ir mazāk izplatīti, jo tie vai nu ir saistīti ar izplūdušajiem neironu tīkliem, kas ir mazāk izplatīti un nav tik viegli iebūvējami, vai tiem ir zems izteismīguma līmenis.

Šis darbs ir vērsts uz ja–tad likumu un bināro klasifikācijas lēmumu koku izgūvi no daudzslāņu perceptrona, kā arī uz eliptisko likumu izgūšanu no radiālo bāzes funkciju neironu tīkla (RBFNT).

Kā tika atzīmēts [49], lai izgūtu kompakto MNT, eksistē vairākas pieejas. Šis darbs izmanto apgriešanas pieeju. Ja cilvēkam nav apmācīta neironu tīkla, to pašu apgriešanas pieeju var izmantot pārāk sarežģīta tīkla apmācībai un apgriešanai, lai izmestu liekos neironus. Samazinot neironu skaitu, var samazināties likumu skaits. Darbā izstrādātais apgriešanas algoritms ir apskatīts otrajā nodaļā.

1.2. Zināšanu izgūšanas metožu veidi

Ir četri galvenie ZI algoritmu veidi, to stiprās un vājās puses ir apkopotas 1.1. tabulā. Šis darbs koncentrējas uz dekompozīcijas un kompozīcijas ZI algoritmiem. Dekompozīcijas algoritmi, kā parādīts [39], darbojas labāk nekā pedagoģiskas pieejas. Kā svarīgākās īpašības tika izvēlētas klasifikācijas precizitāte, pārnēsamība un spēja ietekmēt izgūto likumu sarežģītību un precizitāti.

Pedagoģiskā pieeja tika lietota apmācītam RBFNT, lai izgūtu eliptiskus likumus to lielā izteismīgā spēka dēļ. Dekompozīcijas pieejas ietvaros kā zināšanu reprezentācijas forma tika izmantoti ja–tad likumi un klasifikācijas lēmumu koks. Šīs reprezentācijas formas ir visbiežāk izplatītas, tās var viegli iebūvēt jebkurā esošajā informācijas sistēmā, un tām ir “iebūvēts” secināšanas mehānisms, vienlaikus tās ir viegli saprotamas.

1.1. tabula

Zināšanu izgūšanas algoritmu tipu salīdzināšana

Īpašība	Algoritma klase			
	Dekompozīcijas	Eklektiskais	Pedagoģiskais	Kompozīcijas
Klasifikācijas precizitāte	++	+	+	n/d
Pārnēsamība (nepiesaiste klasifikācijas algoritmam)	–	+–	+	–
Regulējamība	++	+	+	–
Algoritma pastāvīgums (atkārtota izmantošana – tāds pats rezultāts)	–	+	+–	n/d
Ātrums	–	+	+	n/d
Zināšanu reprezentācijas shēmu daudzveidība	++	+	+	–
Mērogošana (lielie dati)	+	+	+	n/d
Algoritma skaitliskā sarežģītība	–	+	+	+

Nodaļā sniegts vispārīgs pārskats par mašīnmācīšanās jomu un MNT vispārināšanas teoriju. Aprakstītas atšķirības starp datiem, informāciju, zināšanām un gudrību. Tiek apskatītas zināšanu reprezentācijas shēmas un tipiskā zināšanu izgūšanas darbplūsma. Līdz ar to pirmais pētījuma uzdevums ir izpildīts.

2. NEIRONU TĪKLU APGRIEŠANA

Ir zināms, ka ir vieglāk izgūt zināšanas no mākslīgajiem neironu tīkliem (MNT) ar mazāku neironu skaitu, tādā gadījumā izgūtajiem likumiem ir zemāka sarežģītība. Apgriešanas solis vienmēr ir ieteicams, jo tas var pozitīvi ietekmēt MNT vispārināšanas spējas. Papildus apgriešanas process kontrolē izgūto likumu saprotamību un precizitāti, kontrolējot apstrādājamo neironu skaitu. Rakstos [3], [6], [25] sniegts pārskats par apgriešanas algoritmiem. Lai izvēlētos apgriešanas pieeju, pamatojoties uz literatūras pārskatu, tika izveidota 2.1. tabula, izmantojot autora novērtējumu skalā 0–5 balles (jo augstāka vērtība, jo labāk). Eksistē daudz apgriešanas pieeju. Lieluma analīzes gadījumā tiek pieņemts, ka lielākie svāri ir vissvarīgākie, kas nav pilnīgi pareizi. Citās pieejās tiek izmantoti sarežģītāki algoritmi. Pieejamo apgriešanas algoritmu realizācijas trūkums un zema uz lieluma analīzes balstītas pieejas klasifikācijas precizitāte radīja vajadzību pēc jauna apgriešanas algoritma.

2.1. tabula

Apgriešanas pieeju salīdzināšana

Kritērijs / apgriešanas tips	Jutīguma analīze	Jutīguma analīze II OBD/OBS	Lieluma analīze	Svaru samazināšana	Savstarpējās informācijas analīze	Nozīmīguma analīze	Interaktīvā apgriešana
Vienkāršums	4	1	5	2	3	1	2
Izpildes laiks	0	1	3	2	3	1	1
Atmiņas apjoma prasības	3	0	3	3	1	2	2
Nav vajadzīga speciāla apmācība	2	1	2	0	2	2	2
Klasifikācijas precizitāte / vispārināšana	3	2	0	n/d	n/d	3	n/d
Apgrieztu neironu / anulēto svaru skaits	3	3	0	n/d	n/d	4	n/d

Vienkāršība un klasifikācijas precizitāte bija noteicošie faktori, lai izvēlētos jutīguma analīzē balstītu apgriešanas metodi. Metodes ideja ir izņemt vienu neironu (vai svaru) un novērtēt MNT klasifikācijas precizitātes izmaiņas, tādējādi var atrast un izņemt vismazāk jutīgo (svarīgo) neironu vai anulēt svaru.

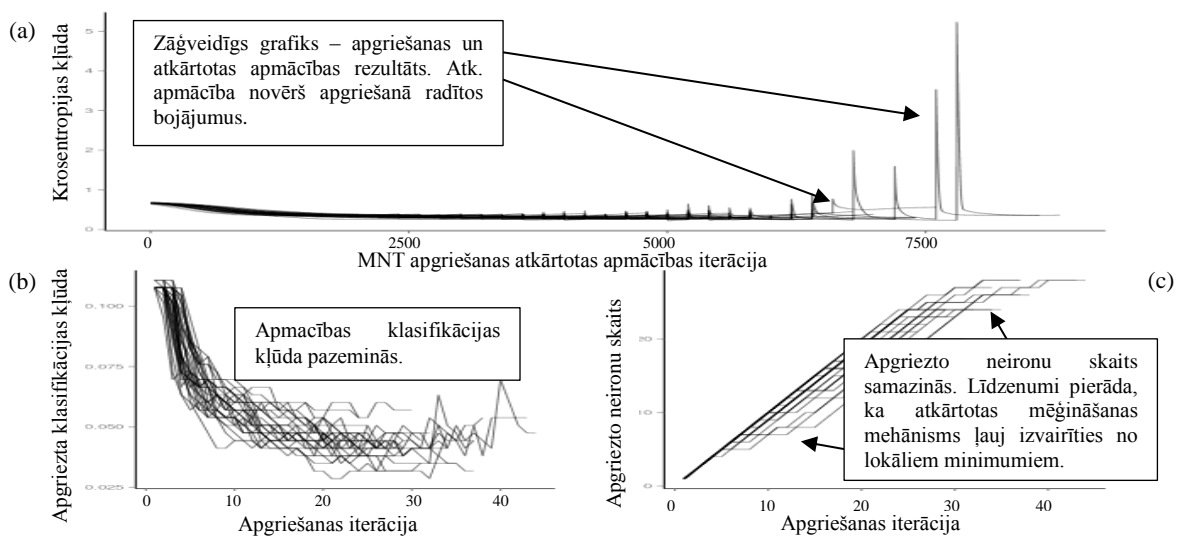
2.1. Jutīguma analīzē balstīts apgriešanas algoritms

Eksperimenti ar jutīguma analīzē balstīto apgriešanas metodi parādīja, ka tai ir tendence iestrēgt lokālajos minimumos, tāpēc tika izstrādāta uzlabota algoritma versija. Lai izvairītos no lokāliem minimumiem, algoritms [26] tika papildināts ar trim uzlabojumiem. Izstrādātā algoritma galvenie soļi ir šādi:

- 1) saglabāt MNT stāvokli;

- 2) atrast vismazāk jutīgo neironu (vai svaru), izņemot to (anulēt);
- 3) atkārtoti apmācīt MNT;
- 4) novērtēt klasifikācijas precizitātes pasliktināšanos; ja tā ir pieņemama, turpināt apgriešanu: pretējā gadījumā – atcelt pēdējā neirona (vai svara) izņemšanu; ja ir notikušas vairākas secīgas atcelšanas, jāpārtrauc apgriešana un jāatgriež pēdējais labākais MNT.

Pirmkārt, trešais un ceturtais solis ir uzlabojumi, kas ir piedāvāti darba gaitā. Tie ļāva apgriezt vairāk neironu (vai svaru) un kontrolēt klasifikācijas pasliktināšanos. Piedāvāto uzlabojumu apstiprināšanai tika izveidota apgriešanas procesa vizualizācija, skat. 2.1. attēlu.



2.1. att. Neironu apgriešanas eksperimenti *Ionosphere* datu kopā.

2.2. Izstrādātā apgriešanas algoritma validācija

Izstrādātā algoritma validēšanai ir izstrādāts eksperimenta plāns. Mērķis bija apstiprināt izstrādāto algoritmu un salīdzināt neironu un svaru apgriešanu. Pirmajās eksperimentu sērijās (2.2. tab.) neironu apgriešanas algoritms tika piemērots neironiem tikai apslēptajos slāņos. Ir redzams, ka visos gadījumos, izņemot trīs, apgriezta MNT testēšanas kļūda bija mazāka nekā neapgriezta MNT kļūda. Izstrādātais apgriešanas algoritms ir parādījis labus rezultātus gan attiecībā uz MNT vienkāršošanu, gan tā vispārināšanas uzlabošanu.

Nākamie eksperimenti tika veikti, lai apstiprinātu izstrādāto algoritmu neironu apgriešanai un svaru anulēšanai [25] ieejas neironiem un apslēptiem neironiem (2.3. tab.).

Apskatot 2.2. un 2.3. tabulas, ir redzams, ka apgrieztam MNT trijos gadījumos ir neliela klasifikācijas kļūdas palielināšanās, bet vairumā gadījumu kļūdu līmenis ir samazinājies. Četros gadījumos (2.3. tab.) neironu svaru apgriešana uzrādīja labākus rezultātus nekā neironu apgriešana. Atlikušajās piecās eksperimentu kopās neironu apgriešana izrādījās labāks risinājums. Kļūda pēc svaru anulēšanas *Monks-1* datu kopā ir 1,81 %, un tas ir mazāk nekā 13,22 % pēc neironu apgriešanas. Šis ir vienīgais gadījums, kurā ir redzamas tik ievērojamas klasifikācijas precizitātes atšķirības.

2.2. tabula

Apgriešanas eksperimentu rezultāti

Datu kopa	MNT apm. kļūda	MNT testa kļūda	Apgr. MNT apm. kļūda	Apgr. MNT testa kļūda	Arhitektūra pirms/pēc apgriešanas (neironi divos apslēptajos slāņos)
<i>Ionosphere</i>	10,83 %	10,83 %	5,39 %	10,44 %	15–15 / 5,4–3,8
<i>Monks-1</i>	20,16 %	29,68 %	18,47 %	24,35 %	15–15 / 5–3
<i>Monks-2</i>	36,82 %	36,55 %	31,83 %	32,58 %	15–10 / 5–3,1
<i>Monks-3</i>	6,64 %	2,80 %	5,98 %	2,85 %	15–15 / 1,7–1,1
<i>WPBC</i>	0,00 %	0,00 %	0,00 %	0,00 %	10–10 / 1–1
<i>WDBC</i>	3,89 %	4,04 %	3,03 %	3,69 %	30–30 / 23,2–17,1
<i>Pima</i>	23,02 %	23,56 %	25,94 %	26,81 %	10–10 / 2,8–3,1
<i>Haberman</i>	26,13 %	26,57 %	28,50 %	28,20 %	15–15 / 2,3–3,7
<i>Parkinsons</i>	24,62 %	24,61 %	16,29 %	15,82 %	30–30 / 26,8–28

2.3. tabula

Klasifikācijas kļūda un vidējais anulēto svaru / apgriezto neironu skaits

Datu kopa	MNT apm. kļ.	MNT testa kļ.	Anulēto. svaru apm. kļ.	Anulēto. svaru testa kļ.	Apgriezto neironu apm. kļ.	Apgriezto neironu testa kļ.	Anulēto svaru / apgriezto neironu skaits
<i>Haberman</i>	25,99 %	26,78 %	24,39 %	24,91 %	24,98 %	26,17 %	54,9/23,8
<i>Ionosphere</i>	10,83 %	10,83 %	4,21 %	10,25 %	4,55 %	9,22 %	34,1/34,3
<i>Monks-1</i>	21,51 %	32,74 %	0,83 %	1,81 %	6,83 %	13,22 %	45,9/22,4
<i>Monks-2</i>	38,46 %	36,04 %	12,47 %	12,21 %	11,26 %	10,25 %	16,8/20,1
<i>Monks-3</i>	6,56 %	2,88 %	5,16 %	3,45 %	3,33 %	5,76 %	32,4/29,3
<i>Parkinsons</i>	24,58 %	24,61 %	14,83 %	16,38 %	14,30 %	15,57 %	10,5/8,3
<i>Pima</i>	23,93 %	24,56 %	21,64 %	23,74 %	22,12 %	23,05 %	56,0 / 22,7
<i>WDBC</i>	4,16 %	4,33 %	1,83 %	2,63 %	1,77 %	2,93 %	23,3/18,5
<i>WPBC</i>	0,00 %	0,00 %	0,00 %	0,17 %	0,00 %	0,00 %	153,6/50,0

Ir pierādīts, ka apgriešana ir noderīgs apmācīta MNT vispārināšanas spēju uzlabojošs solis. Tālāk tekstā ir minēti nodaļā aprakstītie ieguldījumi.

- Nodaļā tiek parādīts promocijas darba autora izstrādātais apgriešanas algoritms; piedāvātās pieejas pamatā ir jutīguma analīze, apvienota ar MNT atkārtotu apmācību, klasifikācijas precizitātes pasliktināšanās sliekšnis un pēdējā labākā MNT kopijas saglabāšana, kas ļauj apgriešanas procedūrai veiksmīgi izvairīties no lokāliem minimumiem.

- Veiktajos eksperimentos ir pierādīta izstrādātā algoritma lietderība neironu tīkla struktūras vienkāršošanā un tā vispārināšanas spēju paaugstināšanā.
- Ieejas un apslēpto neironu apgriešana ir parādījusi zemākas klasifikācijas kļūdas, salīdzinot ar tikai apslēpto slāņu neironu apgriešanu.
- Eksperimenti parādīja, ka neironu apgriešana ir vēlamāka par svaru anulēšanu, jo tas prasa mazāk aprēķinu. Vienīgais izņēmums būtu nepieciešamība saņemt MNT ar pēc iespējas zemāku klasifikācijas kļūdu, tad jālieto svaru anulēšana.

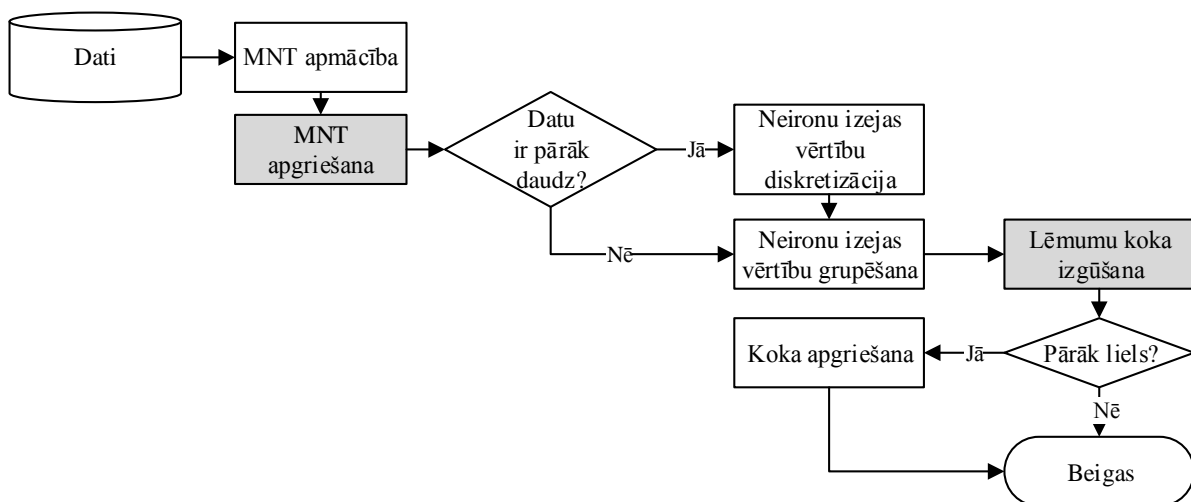
Šajā nodaļā tika izpildīts otrais pētniecības uzdevums – MNT apgriešanas metožu izpēte. Tika izvēlēta apgriešanas pieeja, kā arī izstrādāts apgriešanas algoritms.

3. LĒMUMU KOKA IZGŪŠANA NO DAUDZSLĀŅU PERCEPTRONA

Mākslīgie neironu tīkli (MNT) bieži pārspēj citus klasifikatoru veidus, bet pēc būtības šis klasifikators ir “melnā kaste”, kas stipri ierobežo to izmantošanu un lietojamību. Tāpēc zināšanu izgūšana (ZI) no apmācīta daudzslāņu perceptrona (DSP) var palīdzēt atklāt jaunas zināšanas, validēt vai iegult MNT klasifikatorus citā sistēmā.

3.1. Zināšanu izgūšanas pārskats

Dažādu MNT tipu un arhitektūru apraksts atrodams [12], [38], [56], [70], [92], [98], [120]. Sakarā ar to, ka pilnībā savienoti neironu slāņi ir plaši izmantoti, šī nodaļa koncentrējas uz zināšanu izgūšanu no priekšplūsmas neironu tīkliem (PPNT). Nelinearitāte, ko ievieš slēpti slāņi, ir tas, kas padara PPNT par jaudīgu modeli, bet tas arī neļauj saprast, kā tiek pieņemts klasifikācijas lēmums. Setiono [96] bija viens no pirmajiem pētniekiem, kurš ierosināja strādāt nevis ar neironu tīkla svāriem, bet ar neironu “statistiku” – neironu izejas vērtībām, kas iegūtas, izmantojot apmācības datu kopu. Kā sākotnējais ZI solis tiek ierosināts apgriezt neironus, lai samazinātu apstrādājamo neironu izejas vērtību skaitu (skat. 2. nodaļu) – tas potenciāli var samazināt izgūto likumu skaitu. Nākamais posms ietver neironu izejas vērtību diskretizāciju, lai atrastu neironu izejvērtību grupas robežas, kurās var izmainīties klasifikācijas lēmums. Vairāku neironu gadījumā – ieejas telpas reģionus, kuros visi šim reģionam piederošie ieejas vektori tiek klasificēti kā piederīgi vienai klasei. Darbā izstrādāto algoritmu (3.1. att.) vēlāk var izmantot, lai no šādiem ieejas slāņa diskrētiem neironiem izgūtu induktīvo klasifikācijas lēmumu koku.



3.1. att. Zināšanu izgūšanas process.

Izstrādātais algoritms ir dekompozicionāls, jo tas izmanto raksturīgās zināšanas par neironu izejām. Saistībā ar zināšanu izgūšanas algoritma specifiku binārais klasifikācijas lēmumu koks bija izvēlēts kā likumu reprezentācijas shēma. Tās vizualizācija ļauj skaidri

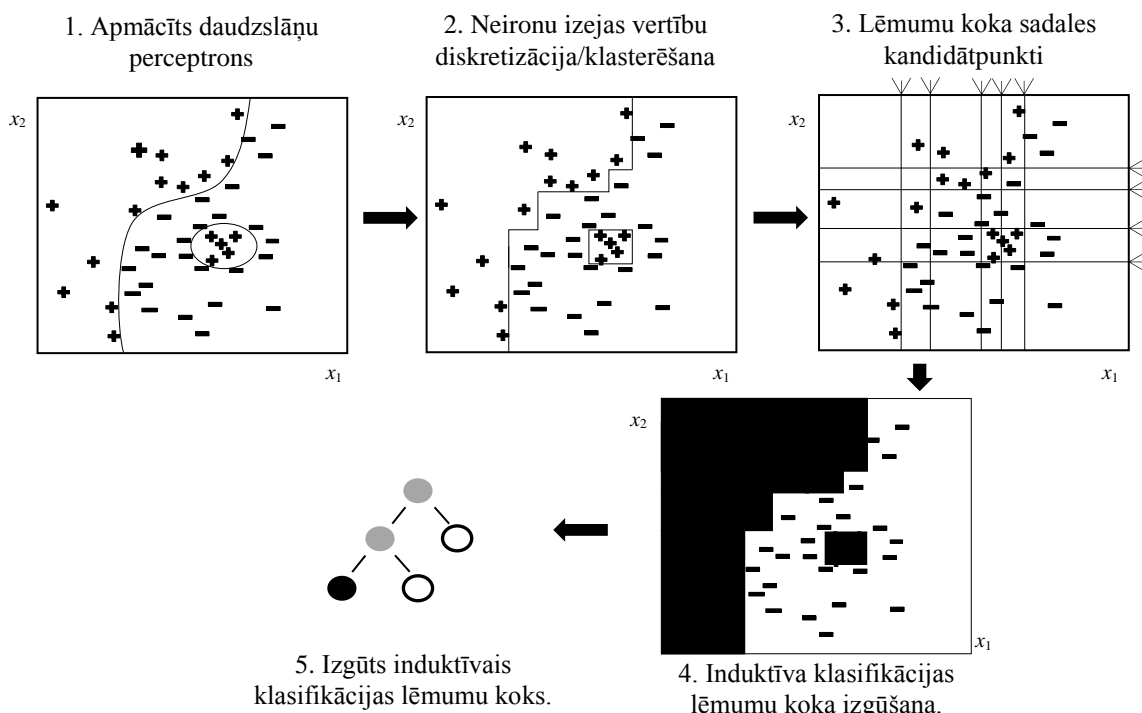
izprast klasifikācijas procesu. Arī ja–tad likumus var viegli izgūt no izgūta induktīva klasifikācijas lēmumu koka.

Zināšanu izgūšanas process ir redzams 3.1. attēlā. Izceltie soļi parāda šī pētījuma jaunieguvumu. Pirms zināšanu izgūšanas posma ir svarīgi vienkāršot MNT; tas samazinās izgūtā lēmumu koka sarežģītību. Gadījumā, ja ir jāsauglabā augsta klasifikācijas precizitāte vai MNT nav pārmācīts, apgriešanas posmu var izlaist.

Gadījumos, kad ieejas datu apjoms ir pārmērīgi liels, ir iespējams: 1) strādāt ar datu apakškopu; 2) veikt neironu izejas vērtību diskretizāciju (noapaļot katra neirona izejas vērtības tā, lai klasifikācijas kļūda nepalielinātos).

3.2. Zināšanu izgūšanas algoritms

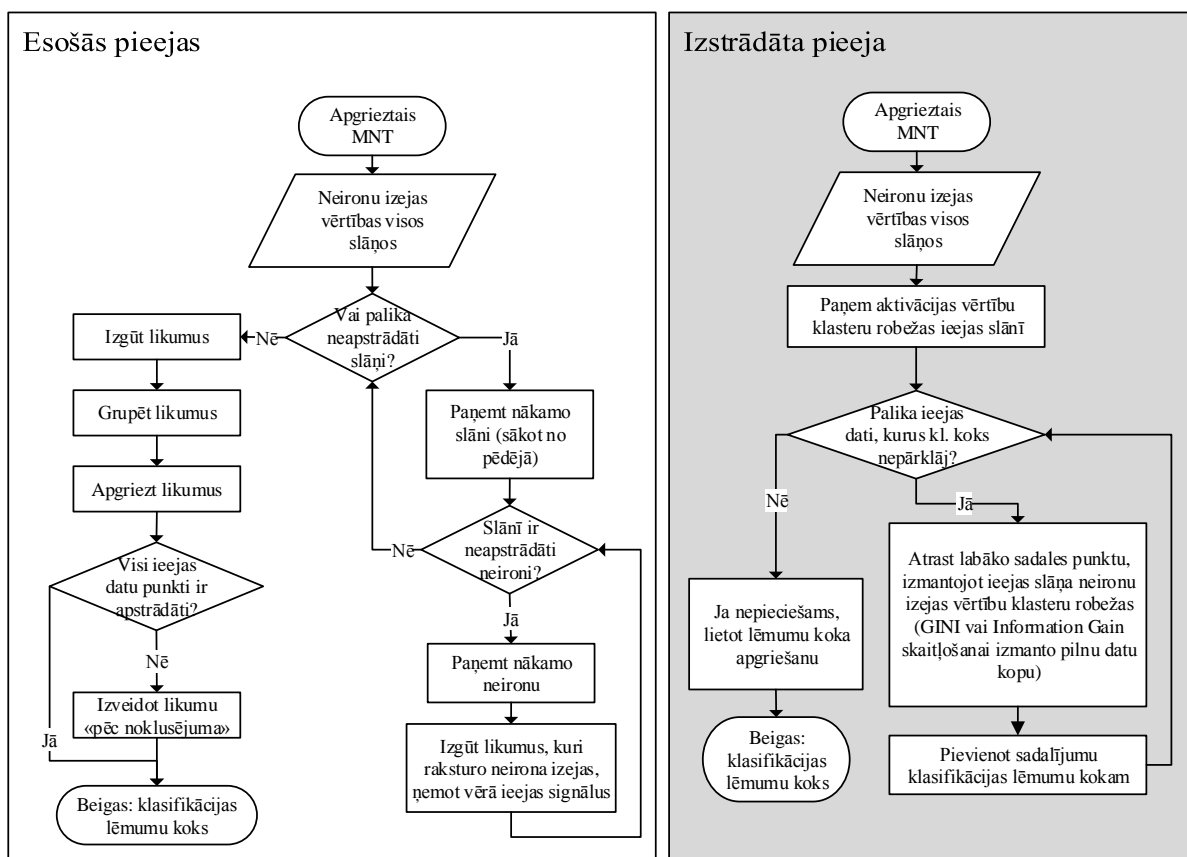
Daudzi no literatūrā aprakstītajiem ZI algoritmiem paplašina sākotnējo *NeuroRule* algoritmu [68]. *NeuroRule* salauž neironu izejas vērtības nelinearitātes, izmantojot grupēšanu, pēc tam visi neironi, sākot no izejas slāņa, tiek aizstāti ar ja–tad likumu kopām, visbeidzot – visi izgūtie likumi tiek apvienoti un apgriezti. Darbā izstrādāta pieeja, kas ļauj izgūt induktīvo klasifikācijas lēmumu koku, izmantojot tikai ieejas slāņa neironu grupas robežas. Piedāvātajā algoritmā lēmumu koka izgūšanai tiek izmantotas tikai neironu izejas vērtības grupas robežas, lai novērtētu informācijas ieguvumu vai Džini indeksa vērtību, izmantojot pilnu apmācības datu kopu. Šis uzlabojums samazina algoritma sarežģītību. Lai pārlicinātos, ka lēmumu koks precīzi apraksta sākotnējo MNT, kā apmācības kopu izmanto tikai vektorus, kas atrodas vistuvāk kvantētai MNT klasifikācijas robežai (skat. trešo soli 3.2. att.). Pieeju salīdzinājums ir redzams 3.3. attēlā.



3.2. att. Izveidotais lēmumu koks, izmantojot ieejas neironu izejas vērtību grupēšanu.

Piedāvātajā klasifikācijas lēmumu koka izgūšana algoritmā ir trīs galvenie soļi (3.2. att., 2.–4. solis):

- 1) MNT nelinearitāte tiek salauzta, izmantojot neironu izejas vērtību grupēšanu;
- 2) tiek atrasti potenciālie dalīšanas punkti no ieejas slāņa neironu kvantēšanas tabulām;
- 3) modificēts indukcijas lēmumu koku izguves algoritms izmanto atrasto kandidātu kopu robežas no ieejas slāņa neironiem kā sadales punktus (vienlaikus izmantojot pilnu apmācības datu kopu *Džini* vai *Information-Gain* aprēķināšanai), lai izveidotu klasifikācijas lēmumu koku.



3.3. att. Esošās un izstrādātā pieejas likumu izgūšanai.

Aprakstītie algoritmi realizēti, paplašinot *Torch7* bibliotēkas *nn* moduli *Lua* programmēšanas valodā [31]. Paplašināta algoritma versija realizēta, paplašinot *PyTorch* dziļās mācīšanās bibliotēku *Python* programmēšanas valodā.

3.3. Izstrādātā algoritma validācija

Izstrādātā algoritma validēšanai tika veikti eksperimenti ar deviņām labi zināmām *UCI* repozitorija datu kopām. Eksperimentu rezultāti ir apkopoti 3.1. tabulā; C4.5 algoritmam rezultāti tika ņemti no [2], [10], [17], [40], [74]. Apgriezta neironu tīkla klasifikācijas robežas piemērs ir parādīts 3.4. a attēlā, izgūta binārā induktīvā lēmumu koka klasifikācijas robeža ir redzama 3.4. b attēlā. *Ripley* un *Iris* datu kopām izgūtie lēmumu koku paraugi ir redzami

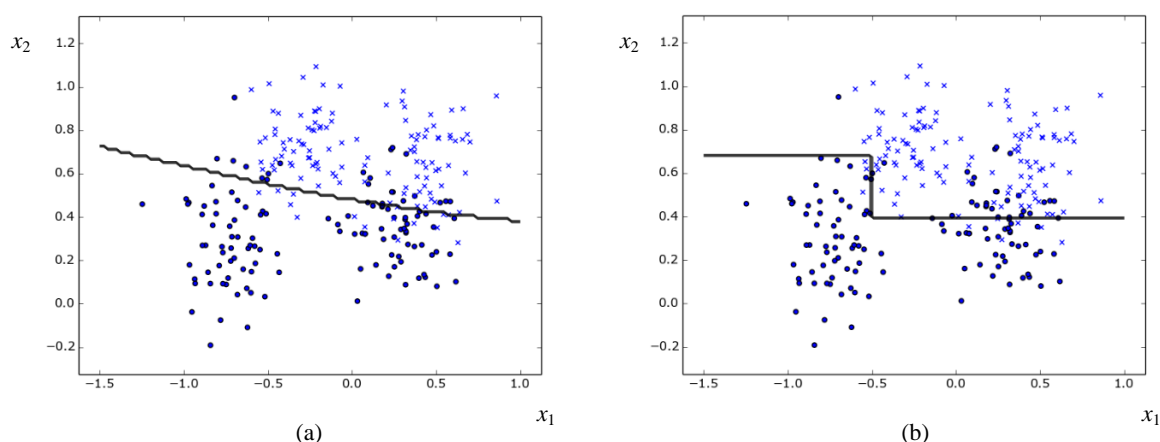
3.5. a un b attēlā. Šie attēli ilustrē minimālos un maksimālos binārās klasifikācijas lēmumu kokus, kas izgūti, izmantojot izstrādāto algoritmu *Iris* datu kopas trīs klašu klasifikācijas problēmai.

3.1. tabula

Precizitāte un lapu skaits DSP, apgrieztam DSP, izgūtam kokam un C4.5

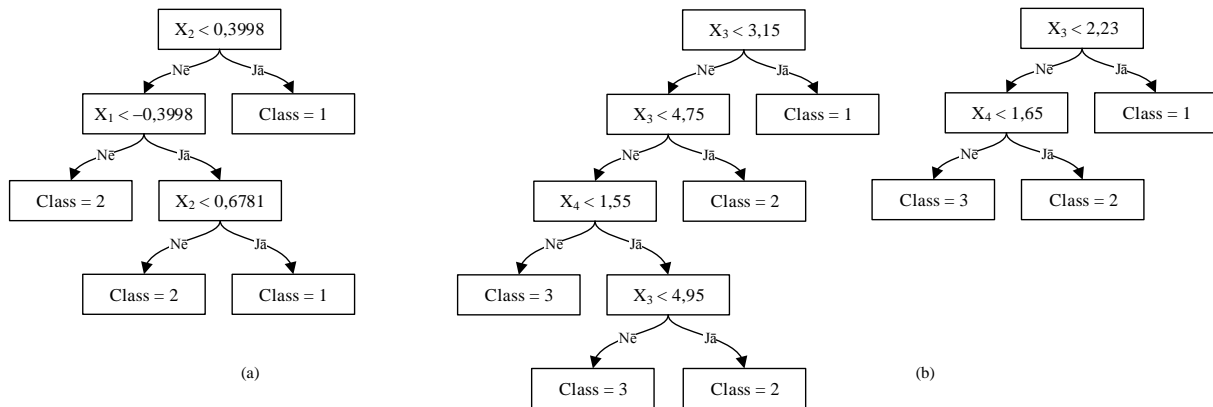
Datu kopa	DSP apm./test.	Piedāvātās metodes			C4.5/J48 test	C4.5 lapu skaits
		Apgriezts DSP apm./test.	Izgūta koka apm./test.	Izgūtā koka lapu skaits _{min} ^{maks}		
<i>Iris</i>	0,9911/0,9667	0,9652/0,9667	0,9689/ 0,9533	4,2 ₅ ³	0,9400	6
<i>Pima diabetis</i>	0,7319/0,7332	0,7253/0,7279	0,7433/ 0,7423	2 ₂ ²	0,7210	14
<i>Ionosphere</i>	1,0000/0,9087	0,9552/0,9544	0,9546/ 0,9059	21,6 ₂₆ ¹²	0,8971	17
<i>Ripley</i>	0,8560/0,8920	0,8640/0,8982	0,8631/ 0,8946	12,6 ₂₀ ⁴	0,8941	6
<i>Haberman</i>	0,7509/0,7381	0,7542/0,7547	0,7567/ 0,7446	2 ₂ ²	0,7190	n/d
<i>Monks-1</i>	1,0000/1,0000	1,0000/1,0000	1,0000/ 1,0000	8 ₈ ⁸	1,0000	28
<i>Monks-2</i>	0,6923/0,6736	0,7160/0,6435	0,7041/0,5949	14 ₁₄ ¹⁴	0,6700	1
<i>Monks-3</i>	0,9754/0,9259	0,9508/0,9722	0,9508/ 1,0000	6 ₆ ⁶	0,9440	14
<i>Parkinsons</i>	0,8051/0,8001	0,8006/0,8099	0,8092/0,8102	2 ₂ ²	0,9261	n/d

Ja ir neliels skaits kopu robežu, kas jāizmanto kā sadalīšanas punkti DT veidošanai, tad skaitļošanas sarežģītība ir diezgan maza. Eksperimentālie rezultāti parāda, ka izgūtā koka klasifikācijas precizitāte ir tieši saistīta ar neironu tīkla precizitāti, kas tiek izmantots kā zināšanu avots, kā arī ar ieejas neironu izejas vērtību skaitu klasteru robežām, kas tiek izmantotas, lai izveidotu koku. Autori [74] ir paziņojuši, ka viņi varēja iegūt MLP precizitāti 0,9876, kas ir daudz augstāka nekā uzrādītajos eksperimentos. Tas nozīmē, ka labāk apmācīta DSP gadījumā izgūtajam kokam būtu augstāka klasifikācijas precizitāte. Ja izgūtā lēmumu koka saprotamība un vienkāršība ir svarīgāka, tad var izmantot agresīvāku apgriešanu.



3.4. att. *Ripley* testa datu kopas sadalījums divās klasēs ar apgrieztu neironu tīklu (a) un izgūto klasifikācijas lēmumu koku (b).

Apskatot 3.1. tabulu, ir redzams, ka izstrādātais lēmumu koka izgūšanas algoritms pārspēj C4.5 algoritmu klasifikācijas precizitātes ziņā. Gadījumi, kad klasifikācijas precizitāte ir zemāka, ir tie, kuros MNT ir parādījis zemu sākotnējo klasifikācijas precizitāti. Aplūkojot likumu skaitu, redzams, ka abi algoritmi ir līdzīgi, bet augstāka precizitāte ļauj vēl vairāk samazināt likumu skaitu, kas izstrādātajai pieejai dod priekšrocības, salīdzinot ar C4.5 algoritmu.



3.5. att. Izgūtais *Ripley* datu kopai parauga lēmumu koks (a) un izgūtie no *Iris* datu kopas dažāda dziļuma lēmumu koki (b).

Trešajā nodaļā ir apskatītas MNT zināšanu izgūšanas pieejas; pamatota zināšanu izgūšana no MNT binārā klasifikācijas lēmumu koka veidā; pamatota *Torch7* bibliotēkas izvēle *Neural Network Knowledge eXtraction (NNKX)* [31] bibliotēkas izveidošanai; prezentēts un eksperimentāli apstiprināts izstrādātais algoritms klasifikācijas lēmumu koka izgūšanai. Piedāvātajam algoritmam ir zemāka klasifikācijas kļūda nekā C4.5 visās pārbaudītajās datu kopās, kur apmācītajā MNT kļūda bija zemāka nekā C4.5 klasifikatoram. Eksperimenti apstiprināja izstrādāto pieeju. Izgūtajam lēmumu kokam ir augsta klasifikācijas precizitāte un zema sarežģītība, līdz ar to trešais pētījuma uzdevums ir paveikts.

4. OPTIMIZĀCIJĀ BALSTĪTAS PIEEJAS LIKUMU IZGŪŠANAI

Trešajā nodaļā prezentēta induktīvo klasifikācijas lēmumu koku izgūšanas pieeja, kas var tikt lietota apmācītam pilnībā savienotam mākslīgam neironu tīklam (MNT). Tomēr, runājot par klasifikācijas modeļa apmācību, var būt, ka izvēlētais MNT neuzrādīs labākos rezultātus. Tāpēc šīs nodaļas mērķis ir izstrādāt divas alternatīvas optimizācijā balstītas pieejas, no kurām viena ļauj izgūt eliptiskus likumus no radiālo bāzes funkciju (RBF) neironu tīkliem, otra – ļauj izgūt no apgabaliem, kas ierobežoti ar hiperplaknēm (kas ir binārs klasifikators), ieejas datu telpā un to aproksimēt, izmantojot ja–tad likumus. Hiperplakņu klasifikatora ieguve ir atsevišķa problēma, un tā iegūšanai tiek piedāvātas divas pieejas.

4.1. Eliptisko likumu izgūšana no RBF neironu tīkliem

RBF neironu tīklu (RBFNT) var izmantot kā alternatīvu klasifikatoru papildus pilnībā savienotiem MNT ar sigmoīdām aktivācijas funkcijām. Eliptiski likumi ir izteismīgāki nekā ja–tad likumi. Tāpēc tika izstrādāta un novērtēta optimizācijā balstīta pedagoģiskā pieeja, kas ļauj izgūt elipsveida likumus no RBFNT [21], [27].

Optimizācijas problēma

Eliptisko likumu izgūšana no apmācītā RBFNT var tikt uzskatīta par neizliektu optimizācijas problēmu, lai atrastu maksimālā apjoma elipsoīdus, kas ierakstīti ieejas telpā, ko nosaka RFBNN klasifikācijas lēmuma robeža. Apzīmēsim elipsoīdu kā:

$$\varepsilon = \{Bu + d \mid \|u\|_2 \leq 1\}, \quad (4.1.)$$

kur B – simetriski pozitīva noteikta matrica; u – vienības sfēra (punktu attālums ir 1 no fiksēta centrālā punkta); d – vektors, kas apzīmē elipsoīdu centru; $\|u\|_2$ – apzīmē Eiklīda normu, t. i., $\|u\|_2 = (u^T u)^{1/2}$.

Elipsoīds ε ir vienības sfēra, kas tiek pārveidota ar afīnu pārveidošanu. Tad elipsoīda tilpums ir proporcionāls $\det B$ [32]. Tādējādi optimizācijas problēmu iespējams formulēt kā

$$\max \log(\det(B)), \text{ s. t. RBFNT} \supseteq \varepsilon. \quad (4.2.)$$

Aprakstītā problēma ļauj atrast pirmo elipsoīdu, kas ierakstīts RBFNT klasifikācijas lēmuma robežās. Izmantojot meklēšanu no vairākiem sākumpunktiem, iespējams tikt galā ar lokālajiem optimumiem, jo optimizācijas problēma nav izliekta. Vairumā gadījumu RBFNT aproksimēšanai ar vienu elipsoīdu nebūs pietiekami; tādējādi ir nepieciešama papildu elipsoīdu meklēšana. Lai atrastu papildu elipsoīdus, ir pietiekami atrast ierakstīto elipsoīdu (kas potenciāli pārklājas ar iepriekš atrastajiem elipsoīdiem) ar maksimālo tilpumu, ko nepārklāj jau atrastie elipsoīdi:

$$\max(\varepsilon_{\text{vol}} - E_{\text{vol}}) - P, \text{ s. t. RBFNN} \supseteq \varepsilon, \quad (4.3.)$$

kur ε_{vol} – jaunatrastā elipsoīda tilpums; E_{vol} – jau esošo (iepriekš atrasto) elipsoīdu tilpums; P – soda parametrs. Ar ieviesto soda parametru P aprēķina minimālo attālumu (4.4. izteiksme) starp kandidāta elipsoīda centru un kopas robežu, ko veido visu iepriekš atrasto elipsoīdu krustojums.

$$P = \min(\text{dist}(\varepsilon_{\text{center}}, E_{\text{surf}}): E_{\text{surf}} \in (E_1 \cup \dots \cup E_N)), \quad (4.4.)$$

kur $\varepsilon_{\text{center}}$ – jaunatrastā elipsoīda kandidāta centrs; E_{surf} – virsma, ko veido jau atrastu elipsoīdu krustojums. P termins nodrošina, ka katrā iterācijā optimizācijas mērķis atrod jaunu elipsoīdu, kas aptvers pēc iespējas lielāku tilpuma daļu, ko vēl neaptver esošie elipsoīdi.

Ekspierimenti un rezultāti

Ekspierimenti veikti ar sintētisko divdimensiju *Ripley* datu kopu (lai atvieglotu vizuālo analīzi), ko var atrast *UCI* datu kopas repozitorijā [49]. RBFNT inicializācijā tika izmantots [37] aprakstītais algoritms, lai izveidotu vairākus neironu tīklus, kas satur mainīgu skaitu neironu. Ekspierimentos tika novērotas tikai slēgtas RBFNT definētās klasifikācijas robežas, ko var redzēt attēlos. Aplūkojot algoritmu, var pamanīt mainīgo *maxEllipsoidsCount*. Tas aplūkotajā RBFNT tika inicializēts ar vairākiem neironu skaitiem, vienīgais izņēmums bija neironu tīkls ar deviņiem neironiem, kuriem maksimālais izgūto elipsoīdu skaits bija iestatīts uz 7. RBFNT neironu skaits tika izvēlēts – 2, 6, 7 un 9. Vizuālā analīze apstiprina, ka algoritms darbojas kā paredzēts, savukārt ekspierimentālie rezultāti (4.1. tab.) parāda augstu izgūto elipsveida likumu klasifikācijas precizitāti.

4.1. tabula

RBFNT precizitāte, izgūto eliptisko likumu precizitāte un skaits

RBFNT neironu skaits	RBFNT apm. precizitāte	RBFNT testa precizitāte	Elipsoidu apm. precezitāte_{stand.nov.}	Elipsoidu testa precezitāte_{stand.nov.}	Elipsoidu skaits vid._{min maks}
2 neironi	0,852	0,911	0,8400 _{0,000}	0,8870 _{0,000}	2 ₂
6 neironi	0,868	0,905	0,8680 _{0,002}	0,9032 _{0,005}	4,4 ₅
7 neironi	0,876	0,905	0,8760 _{0,000}	0,9031 _{0,002}	5,1 ₇
9 neironi	0,868	0,905	0,8728 _{0,005}	0,9039 _{0,001}	6,8 ₇

Ekspierimenta rezultāti rāda, ka izgūtajiem elipsoīdiem ir gandrīz identiska precizitāte, salīdzinot ar sākotnējiem RBFNT, bet tiem ir vienāds vai mazāks elipsoīdu skaits (salīdzinot ar RBF neironu skaitu). Svarīgs faktors, kas jāmin, ir skaitļošanas aprēķinu sarežģītība, jo algoritms izmanto RBFNT, lai pārbaudītu, vai elipsoīds pilnībā atrodas RBFNT lēmuma robežās. Šī operācija ir intensīva skaitļošanas ziņā, taču tā ir paralēlizējama.

Apakšnodaļas ieguldījumu saraksts ir šāds: noformulēta optimizācijas problēma, ieskaitot īpašu soda parametru; definēta mērķa funkcija; piedāvāta algoritma ekspierimentālā validācija un realizācija. Tika parādīts, ka neliels atrastu elipsoīdu skaits var veikt klasifikāciju ar nelielu klasifikācijas precizitātes kritumu. Tas pierāda, ka ierosinātā pieeja ir realizējama un lietojama, it īpaši RBFNT, kas darbojas ar mazdimensiju datu kopām.

4.2. Likumu izgūšana no apgabaliem, kas ierobežoti ar hiperplaknēm

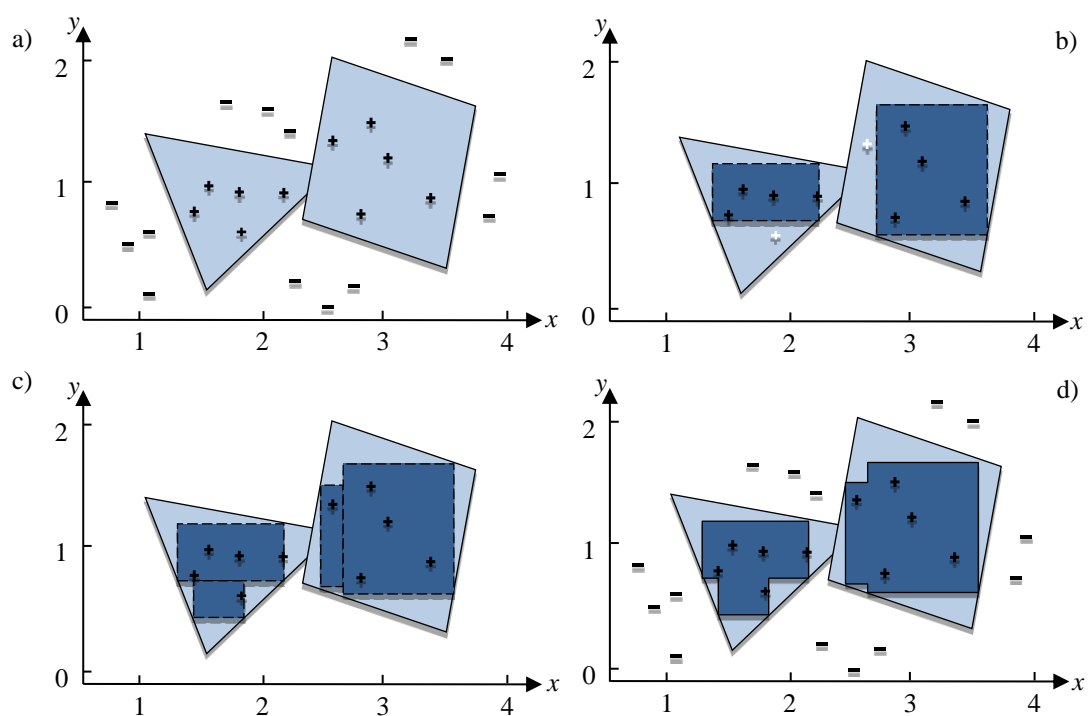
Šajā apakšnodaļā ir piedāvāta optimizācijā balstīta pieeja ja–tad likumu izgūšanai no apgabaliem, kas ierobežoti ar hiperplaknēm (kas ir binārs klasifikators). Šī pieeja tika izvēlēta,

jo, kā tas jau tika parādīts [127], nelineāro sigmoidālo neironu lēmumu robežu var tuvināt ar gabalos sadalāmām lineārām funkcijām. Izstrādātais algoritms ir iepriekš izstrādātā algoritma [80] vispārinājums.

Pieejas pārskats

Lai pārbaudītu ja–tad likumu izgūšanas algoritmu, tika izstrādāts gabalveida politopu klasifikators. Izstrādātā klasiskā klasifikatora pamatideja ir veidot izliektus politopus, ko definē hiperplaknes. Pēc tam no šādiem izliektiem politopiem tiek izgūti ja–tad likumi. Kā politopu avotu var izmantot arī *MSM-T* klasifikatoru [28]–[30], [79]. Izstrādātais algoritms ir rekursīvs, tā galvenie soļi ir šādi.

1. Hiperpolitopu klasifikatora iegūšana (4.1. a att.).
2. Labākā ja–tad likuma izgūšana (pēc tilpuma vai segto punktu skaita; 4.1. b att.).
3. Rekursijas sākums: telpas sadalīšana neaptvertās apakštelpās. Labākā ja–tad likuma izgūšana katrā no atrastajām apakštelpām (4.1. c att.).
4. Pārbaudot rekursijas dziļuma ierobežojumu, ja tas nav sasniegts, atkārtot trešo soli, pretējā gadījumā – pārtraukt (4.1. d att.).



4.1. att. Ja–tad likumu izgūšanas algoritma iterācijas.

Ja–tad likumu izgūšanas algoritms

Izstrādātais algoritms seko definētajai lineāras programmēšanas (LP) problēmai [28]–[30], [79] un definē divas virsotnes, kas reprezentē atrastā hiperparalēlskalda *apakšējās* un *augšējās* robežas. Klasifikācijas precizitātes ziņā viens ierakstīts hiperkubs var būt nepietiekams. Lai novērstu šo nevēlamo rezultātu, ir jāveic rekursīva meklēšana, lai atrastu papildu hiperkubus (ja–tad likumus), kas ierakstīti atlikušajos reģionos. Šo procesu var atkārtot rekursīvi, lai, meklējot mazākos neaptvertos reģionus, izgūtu vairāk likumu, kas

aproximēs sākotnējos ar hiperplaknēm ierobežotos apgabalus (klasifikatoru) ar vēlamo klasifikācijas precizitātes līmeni. Šādus atlikušos apakšreģionus var definēt kā 4.5. izteiksmēs:

$$I_i^l = \begin{cases} x \in R^n, \text{ s. t. ,} \\ l_j^* < x_j \leq u_j^* \forall 1 \leq j \leq i, \\ x_i \leq l_i^*, \end{cases} \quad I_i^u = \begin{cases} x \in R^n, \text{ s. t. ,} \\ l_j^* \leq x_j < u_j^* \forall 1 \leq j \leq i, \\ x_i \geq u_i^*, \end{cases} \quad (4.5.)$$

kur I_i^l, I_i^u – politopu reģioni, kas atrodas apkārt, iegūtajam likumam i -tā dimensijā; l, u – pašlaik apstrādātā hiperkuba augšējā un apakšējā robeža (likums). Kā parādīts 4.5. izteiksmēs, vienādības ir izpildītas x pirmajās $i - 1$ dimensijās, nevienādība, kas attiecas uz i -to dimensiju, neizpildās, un pārējās dimensijas ir brīvas, un to nevajag ierobežot. Lai atbalstītu rekursīvo meklēšanu, ir svarīgi garantēt, ka rekursīvi jaunizgūtie likumi (hiperkubi) nekrustojas savā starpā. Tādēļ dimensijām i, j ir spēkā $j > i$. Katram $x \in I_j$, mums ir $l_i^* < x_i < u_i^*$, un katram $x \in I_i$, mums ir $x_i \leq l_i^*$ vai $x_i \geq u_i^*$. Tādējādi I_i nekrustojas, un likumi, kas tiek izgūti katram I_i , atšķiras aproksimēta politopa reģiona ziņā. Jāatzīmē, ka ar hiperplaknēm ierobežotiem apgabaliem varētu būt krustojumi vienam ar otru, tādējādi izgūtie likumi (hiperkubi) var krustoties. Izgūto likumu optimizēšana (apgriešana) nav mērķis, tādēļ tā nav apskatīta.

Eksperimenti un rezultāti

Izstrādātās likumu izgūšanas pieejas pārbaudei tika izvēlētas vairākas publiski pieejamas *UCI* datu kopas [49]. Piedāvātās metodes pārbaude ir apskatīta [28]–[30], [79]. Eksperimentu datu kopas tika izvēlētas, pamatojoties uz popularitātes kritērijiem.

Pirms likumu izgūšanas jāizveido politopi. Lai iegūtu politopus, no kuriem tiks izgūti ja-tad likumi, ir iespējams izmantot *MSM-T* klasifikatoru, kura ideja ir līdzīga induktīvam klasifikācijas lēmumu kokam, bet tiek izmantoti optimālie datu sadalījumi, kas nav paralēli koordinātu asīm. Otra iespēja – izmantot piedāvāto lineāro klasifikatoru (ar hiperplaknēm ierobežotie apgabali), skat. [79]. Datu kopām, kas sākotnēji netika sadalītas validācijas un apmācības kopās, tika veikta 10 likumu šķērsvalidācija un aprēķināta vidējā klasifikācijas precizitāte. *Monks* datu kopas apmācības un validācijas gadījumā datu kopas jau bija sadalītas, un tika veikti 10 eksperimenti.

Ir redzams (4.2. tab.), ka visas datu kopas, izņemot “*Balance-Scale*”, nav tik labi atdalāmas, izmantojot lineāro SVM. Tomēr nelineāras lēmumu virsmas tuvināšana dod vajadzīgo uzlabojumu klasifikācijas precizitātei politopu klasifikatoram un likumiem, izgūtiem no tā. Vājus klasifikācijas rezultātus parādīja C4.5 algoritms. Vairāku virsmu metodes koka (*MSM-T*) metode atpaliek no SVM metodēm un politopu klasifikatora. Likumi, kas izgūti no *MSM-T*, parāda augstu klasifikācijas kļūdu. Empīriski tika atklāts, ka rekursijas dziļuma palielināšana likumu izgūšanai no *MSM-T* palīdz samazināt klasifikācijas kļūdu.

Pelēkā krāsā izceltās rindās ir rezultāti, ko parāda izstrādātie klasifikatori. Treknraksts norāda labāko kopējo klasifikācijas rezultātu testa komplektā visās datu kopās (kolonnās). Pelēkā krāsa norāda izskaidrojamus klasifikatorus. Zaļā, dzeltenā un sarkanā krāsa norāda labāku, vidēju un sliktāku rezultātu izskaidrojamiem klasifikatoriem. Šeit var redzēt, ka likumiem, kas izgūti no politopu klasifikatora, ir augstāka precizitāte nekā likumiem, kas

izgūti no *MSM-T*. Kopumā SVM ar RBF kodolu pārspēj citus klasifikatorus. Starp visiem klasifikatoriem ir izskaidrojami tikai trīs klasifikatori (C4.5, likumi no *MSM-T* un likumi no politopiem). Starp tiem ir visaugstākā precizitāte likumiem, kas izgūti no (hiper)politopiem (apakšējā rindā). Tas pierāda, ka izvēlēto pieeju var veiksmīgi izmantot, lai izgūtu precīzus likumus.

4.2. tabula

Klasifikācijas precizitāte, %

Klasifikators	<i>Monks-1</i>	<i>Monks-2</i>	<i>NDCC</i>	<i>Balance</i>
<i>SVMlinear</i>	65,509	67,130	73,333	93,730
<i>SVMrbf</i>	86,806	80,556	95,000	98,082
<i>MSM-T</i>	83,565	79,630	88,667	87,526
Likumi (<i>MSM-T</i>)	69,444	54,861	75,333	68,524
C4.5	75,690	65,050	74,000	70,780
Politopi	99,537	80,324	93,667	97,913
Likumi (politopi)	96,296	74,537	89,738	96,857

Šajā nodaļā tiek piedāvātas divas pieejas, kas izstrādātas optimizācijā balstītā zināšanu izgūšanas jomā. Šīs pieejas papildina un kalpo kā alternatīvas zināšanu izgūšanas metodes gadījumos, kad klasifikācijas lēmumu koka izgūšana no apmācīta MNT parāda sliktus rezultātus. Gadījumos, kad izmantotais MNT ir RBFNT un mēs vēlamies samazināt likumu skaitu, labāks risinājums ir eliptisko likumu izgūšana. Gadījumā, ja lietotājam jau ir atrasti hiperpolitopi, tad no tiem ir iespējams izgūt ja–tad likumus. Izstrādātā pieeja nodrošina vēl vienu alternatīvu klasifikācijas lēmumu koka izgūšanai.

Eksperimentālā validācija ir parādījusi, ka ja–tad likumus var izgūt no hiperpolitopiem, izmantojot izstrādāto algoritmu kā izliektas optimizācijas problēmas risinājumu. Tika parādīts, ka izgūtiem ja–tad likumiem ir augsta precizitāte, tāpēc tos var izmantot, lai aprakstītu sākotnējo klasifikatoru, tie var kalpot kā vēl viens mašīnmācīšanās instruments. Tika izvirzīta arī neizliekta optimizācijas problēma, kā arī izstrādāts jauns algoritms, kas ļauj izgūt eliptiskus likumus. Tas ļauj secināt, ka ir paveikts uzdevums, kas saistīts ar optimizācijas metožu izstrādi un novērtēšanu ja–tad un eliptisku likumu izgūšanai no apmācītiem MNT un RBFNT.

Līdz ar to ir paveikts pētījuma ceturtais uzdevums, kas saistīts ar optimizācijā balstītu metožu izstrādi un novērtēšanu ja–tad un eliptisko likumu izgūšanai no apmācītiem MNT un RBFNT.

5. ZINĀŠANU IZGŪŠANAS METODOLOĢIJA

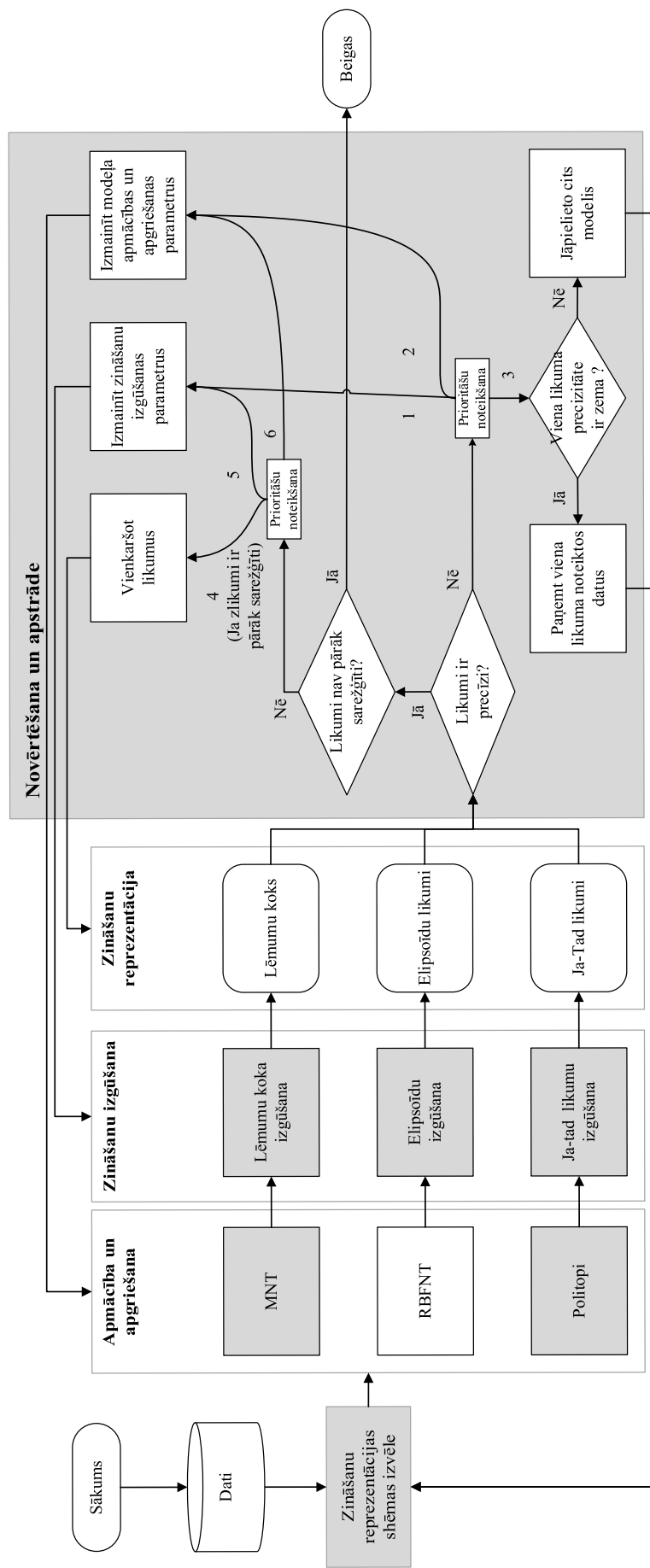
Balstoties uz paveikto darbu un iegūtajiem rezultātiem, šajā nodaļā ir izstrādāta un eksperimentāli apstiprināta likumu izgūšanas metodoloģija. Tiek prezentēta vienota zināšanu attēlošanas un modeļa atlases, zināšanu izgūšanas, novērtēšanas un pilnveidošanas darbplūsma. Pārskats par zināšanu reprezentāciju un izguvi no mākslīgiem neironu tīkliem un cita veida klasifikatoriem, kas ietverti pirmā nodaļā, ir pamatā metodoloģijai [18]–[20], kurā ir aprakstīts apgriešanas algoritms (skat. 2. nodaļu). Jauns algoritms binārā klasifikācijas lēmumu koka izgūšanai no pilnībā savienota daudzslāņu MNT aprakstīts trešajā nodaļā. Visbeidzot, optimizācijā balstītas metodes likumu izgūšanai, izmantojot izliektu optimizācijas problēmu, ir iekļautas 4. nodaļā. Tajā pašā nodaļā ir aprakstīta jauna izstrādāta (neizliektas optimizācijas) metode eliptisko likumu izgūšanai no RBFNT klasifikatora. Šajā nodaļā piedāvāta darbplūsma, kas ir eksperimentāli validēta. Darbplūsmā sniegtas vadlīnijas, kā izgūt zināšanas, kas reprezentētas, izmantojot precīzus vai vienkāršus un saprotamus likumus.

5.1. Metodoloģijas izstrāde

Izstrādātā metodoloģija zināšanu izgūšanas (ZI) procesu iedala vairākos secīgos posmos. Daži no tiem nav obligāti (ja jau ir apmācīts klasifikators, kas jāapraksta kā likumu kopums).

1. Zināšanu reprezentācijas shēmas un klasifikatora tipa izvēle. Šis solis paredz zināšanu shēmas izvēli. Eliptiski likumi ir izteismīgāki nekā lēmumu koks vai ja–tad likumi, bet mazāk saprotami. Tātad, ja kopējais mērķis ir saprast, kā tiek veikta klasifikācija, eliptiski likumi varētu nebūt pietiekami optimāla izvēle. No otras puses, lēmumu koku var izgūt tikai no MNT, tātad, ja klasifikators ar labāko precizitāti ir balstīts uz hiperpolitopiem, tad zināšanu reprezentācijai būs jāizmanto ja–tad likumus.
2. Klasifikatora apmācība un sagatavošanās ZI posmiem. Ja ir pieejams MNT, to var apgriezt. Ja klasifikators jau ir iepriekš apmācīts, ZI var sākties uzreiz.
3. Zināšanu izgūšana.
4. Izgūto zināšanu novērtēšana un pilnveidošana. Šis solis ir visu iesniegto un pārskatīto algoritmu izpētes un attīstības rezultāts. Tas ļauj novērtēt izgūtās zināšanas un sniedz norādījumus par to precizitātes palielināšanu vai vienkāršošanu. Ja izgūtie likumi ir sarežģīti, var veikt vienu vai vairākus no šiem soļiem:
 - likumus var apvienot un apgriezt;
 - lietot agresīvāku MNT apgriešanu vai izmantot mazāku elipsoīdu daudzumu vai mazāku ja–tad likumu izgūšanas rekursijas dziļumu;
 - klasifikatoru var atkārtoti apmācīt ar izmainītiem apmācības parametriem un veikt ZI soli atkārtoti;
 - apakštelpas reģions, kurā klasifikatora precizitāte ir slikta, var tikt izmantots, lai apmācītu atsevišķu klasifikatoru, lai tajā lietotu ZI.

Visi izstrādātie algoritmi ir apvienoti vienotā zināšanu izgūšanas metodoloģijā, kas parādīta 5.1. attēlā.



5.1. att. Zināšanu izgūšanas metodoloģija. Promocijas darba ieguldījums ir izcelts pelēkā krāsā.

5.2. Precīzi vai saprotami likumi

Lai validētu piedāvāto algoritmu un novērtētu rekomendācijas precīzu vai vienkāršu likumu izgūšanai, ir izstrādāts un izpildīts eksperimentu plāns vidēja un maza izmēra datu kopām. Lai pārvarētu lielas skaitļošanas prasības, kas nepieciešamas reālas vidēja lieluma (~50 000 ierakstu) *Adult Census* datu kopas (ņemta no *UCI* repozitorija) apstrādei, *ZI* algoritmā tika iekļautas divas modifikācijas:

- neironu izejas diskretizācija, izmantojot visu neironu izejas, noapaļojot līdz n cipariem, kļūst par obligātu soli, tas var negatīvi ietekmēt klasifikācijas precizitāti;
- neironu izejas vērtību grupēšana tika veikta datu apakškopā (eksperimentos tas bija 15 % no visas apmācības kopas).

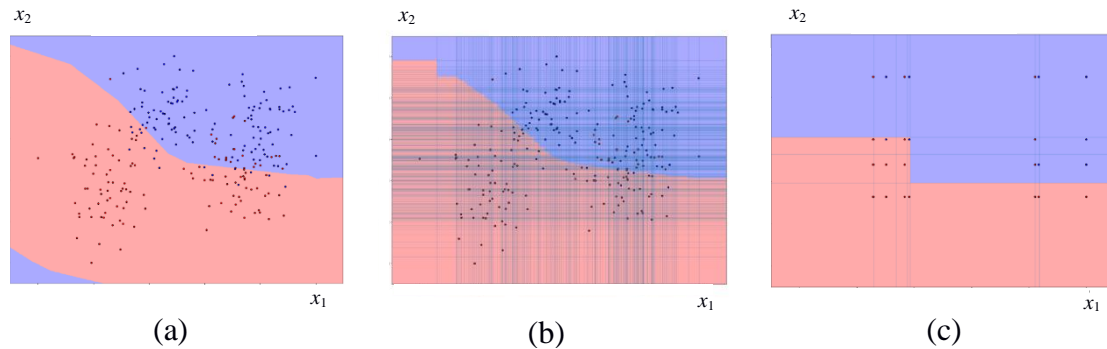
Eksperimenta mērķis ir pierādīt, ka parametrs, kas kontrolē pieļaujamo klasifikācijas precizitātes degradāciju neironu izejas vērtību grupēšanas fāzē, ļauj kontrolēt izgūto lēmumu koku sarežģītību un klasifikācijas galīgo precizitāti. Svarīgi bija saprast, kā neironu apgriešana ietekmē izgūtā lēmumu koka sarežģītību un klasifikācijas precizitāti. 5.1. tabulā sniegtie rezultāti apkopo 10 eksperimentus ar katru no divām datu kopām un parāda, ka agresīva apgriešana kopā ar lielu neironu izejas vērtību grupēšanas klasifikācijas degradācijas sliekšni rada zemāku klasifikācijas precizitāti un mazāku lēmumu koku. Viegla apgriešana un zems sliekšnis ļauj izgūt lielākus, bet precīzākus lēmumu kokus.

5.1. tabula

Zināšanu izgūšanas parametru ietekme uz izgūto likumu sarežģītību un precizitāti

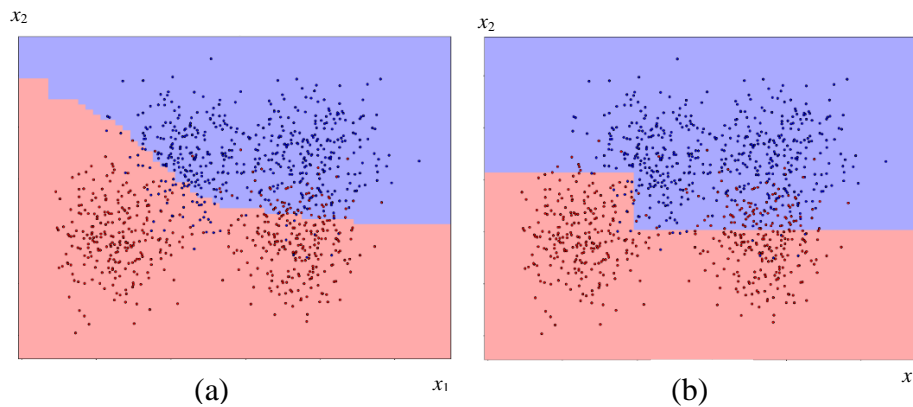
MNT apgriešanas līmenis	Neironu izejas vērtējumu grupēšana	Pazīmes	Datu kopa			
			Adult		Ripley	
			Apm.	Testa	Apm.	Testa
Agresīva apgriešana	MNT		82,27	82,15	83,84	87,92
	Mazs grup. klasifikācijas pasliktināšanās sliekšnis	Diskretizēta MNT precizitāte, %	82,09	81,95	83,84	87,89
		Izgūtā lēm. koka precizitāte, %	79,74	79,81	83,80	88,00
		Izgūtā lēm. koka likumu skaits / lēm. koka dziļums	212,7/16,2		55,6/28,7	
	Vidējs grup. klasifikācijas pasliktināšanās sliekšnis	Diskretizēta MNT precizitāte, %	80,89	80,99	83,28	87,76
		Izgūtā lēm. koka precizitāte, %	79,80	79,74	83,28	87,76
		Izgūtā lēm. koka likumu skaits / lēm. koka dziļums	20,1/7,7		16,8/12,3	
Minimāla apgriešana	MNT		83,41	83,22	84,44	89,52
	Mazs grup. klasifikācijas pasliktināšanās sliekšnis	Diskretizēta MNT precizitāte, %	83,39	83,21	84,44	89,62
		Izgūtā lēm. koka precizitāte, %	82,12	81,95	84,44	89,52
		Izgūtā lēm. koka likumu skaits / lēm. koka dziļums	1562/21,8		137/91,6	
	Vidējs grup. klasifikācijas pasliktināšanās sliekšnis	Diskretizēta MNT precizitāte, %	82,27	82,00	83,4	88,4
		Izgūtā lēm. koka precizitāte, %	79,25	78,87	83,4	88,4
		Izgūtā lēm. koka likumu skaits / lēm. koka dziļums	161,1/12,5		11,3/8,1	

Ripley datu kopas eksperimenti palīdz saprast, kā grupēšanas sliekšņa parametrs ietekmē izgūto lēmumu koku. 5.2. a attēlā parādīts, kā izskatās sākotnējā grupēšanas robeža, ko definē apmācīts un nedaudz apgriezts MNT. Diskretizācijas un grupēšanas rezultāts, izmantojot mazu (5.2. b att.) un lielu (5.2. c att.) pieļaujamo klasifikācijas pasliktināšanās sliekšni, redzams 5.2. attēlā. Vertikālās un horizontālās līnijas parāda kopu robežas.



5.2. att. MNT klasifikācijas robežas *Ripley* datu kopai.

Neliels punktu skaits 5.2. c attēlā ir saistīts ar grupēšanu (kvantēšanas tabulas izmantošanu), kas aizstāj neironu izejas vērtības. 5.3. attēlā parādīti induktīva lēmumu koka klasifikācijas robežas, kas izgūti no kvantēta MNT, izmantojot neironu izejas vērtību grupēšanu.



5.3. att. Lēmumu koka klasifikācijas robeža izgūta no MNT ar mazu (a) un lielu (b) grupēšanas parametru klasifikācijas pasliktināšanos kontrolei.

Eksperimenti parādīja, ka pastāv vairāki parametri, kas ļauj kontrolēt izgūtā lēmumu koka sarežģītību (lielumu) un precizitāti.

1. Neironu izejvērtību grupēšanas klasifikācijas pasliktināšanās sliekšnis. Vidējām un lielām datu kopām grupēšanas fāze paredz izmantot nejauši izvēlētu apmācības datu kopas apakškopu. Jo mazāka apakšgrupa, jo lielāka būs klasifikācijas kļūda. Tas, iespējams, var radīt vienkāršākus un mazāk precīzus likumus, kas izgūti, izmantojot mazāk aprēķinu.
2. Diskretizācija – cik skaitļu aiz komata jāņem neironu izvadē. Tas ļauj samazināt

aprēķināšanas izmaksas grupēšanas posmā (uz dažu klasifikāciju pasliktināšanās rēķina).

3. Apgriešanas procedūras agresivitāte, kas ļauj kontrolēt izgūto likumu sarežģītību un klasifikācijas precizitāti.

Ir vērts atzīmēt, ka apgriešanai ir liela nozīme, jo tā ļauj novērst pārāpmācību. Jebkurā gadījumā, kad tiek piemērota neagresīva apgriešana, kā arī neagresīva neironu izejvērtību klasifikācija (kvantēšana), izgūtajam kokam būs klasifikācijas precizitāte, kas līdzīga MNT, no kura tas tiek izgūts.

Šajā nodaļā apkopots veiktais darbs un sniegta vienota darbplūsma ar ieteikumiem ZI. Šī darbplūsma un ieteikumi uzsver iepriekšējās nodaļās paveikto pētniecības darbu un sniedz vadlīnijas zināšanu izgūšanai ar eksperimentālās validācijas rezultātiem. Ir sniegtas vadlīnijas klasifikatoru un atbilstošo zināšanu izgūšanas algoritmu izvēlei, ņemot vērā datu kopas raksturlielumus. Tiek prezentēta un eksperimentāli apstiprināta darbplūsma pārāk sarežģītu vai pārāk vienkāršu izgūto likumu apstrādei. Pašreizējo darba ieguldījumu var apkopot šādi.

- Piedāvāta metodoloģija – vispārēja zināšanu izgūšanas darbplūsma no apmācīta MNT vai hiperpolitopu klasifikatora.
- Metodoloģijā formulēti ieteikumi klasifikācijas modeļa izvēlei, pamatojoties uz datu kopas raksturlielumiem.
- Metodoloģijā, pamatojoties uz eksperimentālo validāciju, tiek piedāvāti ieteikumi apgriešanas un zināšanu izgūšanas parametru izvēlei, aprakstīto parametru ietekme uz izgūto zināšanu sarežģītību un klasifikācijas precizitāti.
- Noteikta novērtēšanas procedūra, kurā ieteikts, kā zināšanas jānovērtē un kādas korekcijas var veikt konstatēto problēmu novēršanai (ja tādas ir).

Klasifikācijas lēmumu koku izgūšana ir visizplatītākais veids, kā izgūt zināšanas no neironu tīkla. Tomēr saskaņā ar “Nav bezmaksas pusdienu” teoriju nav vienas metodes, kas būtu vienlīdz laba visām datu kopām. Tādējādi tika izstrādāti un pierādīti izmantojami alternatīvi algoritmi eliptisku un ja–tad likumu izgūšanai. Visi uzskaitītie uzlabojumi ir kļuvuši par daļu no vispārējās zināšanu izgūšanas darbplūsmas. Šajā nodaļā aprakstītais darbs izpilda piekto pētījuma uzdevumu.

REZULTĀTI UN SECINĀJUMI

Promocijas darbs ir veltīts zināšanu izgūšanai no apmācītiem mākslīgajiem neironu tīkliem (MNT). Šajā darbā analizēta esošā pieeja zināšanu reprezentācijai, MNT apgriešanai un zināšanu izgūšanai. Šīs analīzes rezultātā izstrādāta zināšanu izgūšanas metodoloģija, kurā izskatītas tipiskās zināšanu reprezentācijas shēmas, sniegtas vadlīnijas labākās zināšanu reprezentācijas izvēlei. Gadījumā, kad zināšanas tiek izgūtas no MNT, nemot vērā to, ka nav brīvi pieejamas algoritmu realizācijas, tika izstrādāta jauna jutīguma analīzē balstīta MNT apgriešanas pieeja. Izstrādāts klasifikācijas lēmumu koka izgūšanas algoritms, kas darbojas ar apmācītu MNT. Lai pārklātu gadījumus, kad MNT nav lietojams, tika izstrādātas alternatīvas uz optimizācijas balstītas zināšanu izgūšanas pieejas. Izstrādātās pieejas ļauj izgūt ja–tad likumus no klasifikatora, ko apraksta hiperpolitopi, un eliptiskos likumus no radiālo bāzes funkciju neironu tīkla (RBFNT). Visas izstrādātās pieejas tiek apvienotas vienā darbplūsmā kopā ar ieteikumiem attiecībā uz īpašu darbplūsmas soļu izvēli. Lēmumu koku klasifikatoram doti eksperimentāli validēti ieteikumi, kas ļauj lietotājam izgūt vai nu sarežģītus un precīzus, vai saprotamākus un mazāk precīzus induktīvus klasifikācijas lēmumu kokus.

Promocijas darba mērķis bija izstrādāt algoritmus MNT apgriešanai un zināšanu izgūšanai no apmācītā MNT un apvienot tos zināšanu izgūšanas metodoloģijā. Promocijas darba sasniegumi ir šādi.

1. Veikts esošās zinātniskās literatūras pārskats un analīze, kas aptver teoriju un zināšanu reprezentācijas un izgūšanas algoritmus. Tā rezultātā pētījumā pārskatīti esošo pieeju plusi un mīnusi, kas ļāva definēt prasības zināšanu izgūšanas darbplūsmai.
2. Pārskatīti mākslīgo neironu tīklu apgriešanas algoritmi, izstrādāts un validēts jauns jutīguma analīzē balstīts MNT apgriešanas algoritms ar pārāpmācību un “kabatas” atmiņu. Ir izstrādāti ieteikumi svaru anulēšanas vai neironu apgriešanas izvēlei. Algoritma spēja pārvarēt lokālus minimumus pierādīta eksperimentāli.
3. Izstrādāta jauna dekompozīcijas pieeja klasifikācijas lēmumu koku izgūšanai no apmācīta daudzslāņu perceptrona (DSP). Algoritms paplašina *Torch7* balstītu *nn* dziļās mācības bibliotēku (*Lua* valodā) ar papildu neironu tīklu slāņiem un funkcijām.
4. Izstrādātas un novērtētas optimizācijā balstītas pedagoģiskās pieejas ja–tad un eliptisko likumu izgūšanai no izliektas hiperpolitopu kopas un RBF neironu tīkla.
5. Izstrādāta un novērtēta vispārināta zināšanu izgūšanas metodoloģija. Metodoloģijā iekļauti ieteikumi modeļa izvēlei (DSP, RBFNT un izliekti politopi), kas ietekmē zināšanu reprezentācijas un izgūšanas pieeju. Metodoloģija satur arī darbplūsmu, kas virza zināšanu izgūšanu, novērtēšanu un pilnveidošanu atkarībā no izvēlēta klasifikācijas modeļa. Metodoloģijas eksperimentālā validācija apliecina secinājumus.

Visu izstrādāto algoritmu klasifikācijas precizitāte tika novērtēta un analizēta, lai pārbaudītu izvirzītās hipotēzes. Balstoties uz veikto pētījumu, var izdarīt vairākus secinājumus.

- Izstrādātais apgriešanas algoritms, kas balstīts jutīguma analīzē, veiksmīgi izvairās no lokāliem minimumiem un ļauj kontrolēt klasifikācijas kļūdu pieaugumu. Atsevišķos

gadījumos jutīguma analīzē balstīta svāra anulēšana var dot labākus rezultātus nekā neironu apgriešana. Lai gan neironu apgriešana ir vēlamāka, jo tā ir mazāk skaitļošanas intensīva metode un kopumā sniedz rezultātus, kas līdzinās svāra apgriešanai, šie rezultāti apstiprina **pirmo hipotēzi**.

- Pārāpmācības izmantošana un “kabatas” atmiņas modifikācijas ir efektīvi algoritmiski uzlabojumi, kas, lietojot jutīguma analīzē balstītu apgriešanas algoritmu, sniedz labus rezultātus, un tas apstiprina **pirmo hipotēzi**.
- Lēmumu koks, kas izgūts, izmantojot daudzslāņu perceptrona ieejas slāņa neironu izejas vērtību klasifikācijas lēmumu robežu (kas ir izgūta, izmantojot neironu izejas vērtību grupēšanu) visu neironu izejas vērtību vietā, aproksimē oriģinālo DSP. Šāda metode ir vienkāršāka (skaitļošanas resursu ziņā), salīdzinot ar visu neironu apstrādi, izmantojot likumu kopas ar secīgu likumu apvienošanu un apgriešanu (lai izgūtu nepieciešamos likumus ieejas slānim). Turklāt izstrādātā pieeja ļauj kontrolēt izgūtā koka sarežģītību un klasifikācijas precizitāti. Šie rezultāti apstiprina **otro hipotēzi**.
- Balstoties uz eksperimentu rezultātiem, var izdarīt secinājumu, ka optimizācijā balstītu pieeju var izmantot ja–tad likumu izgūšanai no izliektiem politopiem. Strādājot ar mazāk jaudīgu personīgo datoru, šī metode ir piemērojama datu kopām ar mazāk nekā 11 atribūtiem. Tādējādi šī metode ir izmantojama kā alternatīva metode, lietojot to tikai ieejas datu kopu apakškopās (izmantojot samazināto ieejas dimensiju skaitu), ja, izgūstot klasifikācijas lēmumu koku no DSP, tiek iegūts slikts klasifikācijas lēmumu koks noteiktam apakšreģionam. Eksperimenti pierāda, ka izgūtie ja–tad likumi aproksimē ieejas telpas reģionus, ko ierobežo hiperpolitopi. Tas apstiprina **trešo hipotēzi**.
- Balstoties uz eksperimentiem, kas saistīti ar izstrādāto algoritmu, var izdarīt secinājumu, ka optimizācijā balstītu pieeju var izmantot eliptisko likumu izgūšanai no RBFNT datu kopām ar ne vairāk ka četrām dimensijām (lietojot parastus personīgos datorus). Šī metode ir izmantojama kā veids, kā aizstāt lielus apakškokus lēmumu kokā ar izteismīgākiem elipsveida likumiem. Eksperimenti pierāda, ka izgūtie eliptiskie likumi aproksimē RBFNT un sasniedz līdzīgu klasifikācijas precizitāti. Tas apstiprina **ceturto hipotēzi**.

Visi definētie teorētiskie jautājumi ir eksperimentāli pierādīti. Darbā izstrādāta un eksperimentāli novērtēta zināšanu izgūšanas metodoloģija, ļauj veikt zināšanu reprezentācijas shēmas un klasifikācijas metodes izvēli, kā arī zināšanu izgūšanu, novērtēšanu un pārveidošanu. Turpmākajos pētniecības virzienos var ietilpt pētījumi par veidiem, kā ieviest reproducējamību MNT apgriešanā un likumu izgūšanā, kā arī neironu izejas vērtību grupēšanas paātrināšanu.

IZMANTOTĀS LITERATŪRAS SARAKSTS

1. Abadi, M., Agarwal, A., Barham, P., et al. *TensorFlow: Large-scale Machine Learning on Heterogeneous Systems* [online]. 2015 [viewed 1 June 2019]. Available from: <https://www.tensorflow.org/extras/tensorflow-whitepaper2015.pdf>.
2. Abdelhalim, A., Traore, I., Sayed, B. RBDT-1: A New Rule-Based Decision Tree Generation Technique. In: G. Governatori, J. Hall, A. Paschke, eds. *Rule Interchange and Applications. RuleML 2009*. Berlin, Heidelberg: Springer, 2009, pp. 108–121 (Lecture Notes in Computer Science. Vol. 5858).
3. Alshahrani, M., Soufan, O., Magana-Mora, A., Bajic, V. B. DANNP: An Efficient Artificial Neural Network Pruning Tool. *PeerJ Computer Science*. 2017. Available from: <https://doi.org/10.7717/peerj-cs.137>.
4. Anderson, J. A. *An Introduction to Neural Networks*. MIT Press, 1995. 672 p.
5. Augasta, M. G., Kathirvalavakumar, T. A Novel Pruning Algorithm for Optimizing Feedforward Neural Network of Classification Problems. *Neural Processing Letters*. 2011, vol. 34, no. 3, pp. 241–258.
6. Augasta, M. G., Kathirvalavakumar, T. Pruning Algorithms of Neural Networks – A Comparative Study. *Central European Journal of Computer Science*. 2013, vol. 3, no. 3, pp. 105–115.
7. Augasta, M. G., Kathirvalavakumar, T. Reverse Engineering the Neural Networks for Rule Extraction in Classification Problems. *Neural Processing Letters*. 2012, vol. 35, no. 2, pp. 131–150.
8. Babaeizadeh, M., Smaragdis, P., Campbell, R. H. *NoiseOut: A Simple Way to Prune Neural Networks* [online]. 2016 [viewed 1 June 2019]. Available from: <https://arxiv.org/pdf/1611.06211.pdf>.
9. Baesens, B., Gestel, V., Viaene, T. S., et al. Benchmarking State-of-the-art Classification Algorithms for Credit Scoring. *Journal of the Operational Research Society*. 2003, vol. 54, no. 6, pp. 627–635.
10. Belacel, N., Al-Obeidat, F. Learning PROAFTN with a Comparative Study with Decision Trees. In: C. J. Butz and P. Lingras, eds. *Advances in Artificial Intelligence: 24th Canadian Conference on Artificial Intelligence, Canadian AI 2011, St. John's, Canada, 25–27 May 2011, Proceedings*. Berlin: Springer, 2011, pp. 56–61. (Lecture Notes in Artificial Intelligence. Vol. 6657).
11. Barakat, N., Diederich, J. Eclectic Rule-Extraction from Support Vector Machines. *International Journal of Computational Intelligence*. 2005, vol. 2, no. 1, pp. 59–62.
12. Bengio, Y. Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*. 2009, vol. 2, no. 1, pp. 1–127.
13. Beltramelli, T. *pix2code: Generating Code from a Graphical User Interface Screenshot* [online]. Copenhagen: UIzard Technologie, 2017 [viewed 1 June 2019]. Available from: <https://uizard.io/research#pix2code>.

14. Bilbao, I., Bilbao, J. Overfitting Problem and the Over-training in the Era of Data: Particularly for Artificial Neural Networks. In: *The 8th International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt. 2017, 5–7 December 2017*. IEEE, 2017. Available from: doi: 10.1109/INTELCIS.2017.8260032.
15. Bishop, C. M. B6. Neural Networks: A Pattern Recognition Perspective. In: E. Fiesler, R. Beale, eds. *Handbook of Neural Computation*. Oxford: Oxford University Press. 1997, pp. 190–212.
16. Bodyanskiy, Ye. V., Rudenko, O. G. *Artificial Neural Networks: Architectures, Learning, Applications*. Kharkiv: TELETEH, 2004. 372 p. (in Russian).
17. Bondarenko, A., Aleksejeva, L., Jumutcs, V., Borisovs, A. Classification Tree Extraction from Trained Artificial Neural Networks. *Procedia Computer Science*, 2017, vol. 104, pp. 556–563.
18. Bondarenko, A., Aleksejeva, L. Methodology for Knowledge Extraction from Trained Artificial Neural Networks. *Information Technology and Management Science*. 2018, vol. 21, pp. 6–14.
19. Bondarenko, A., Aleksejeva, L. Workflow for Knowledge Extraction from Neural Network Classifiers. In: *59th Intern. Scientific Conf. on Information Technology and Management Science of Riga Technical University (ITMS), 2018*. Available from: doi:10.1109/ITMS.2018.8552964.
20. Bondarenko, A., Aleksejeva, L. Controlling Complexity and Accuracy of Classification Decision Tree Extracted from Trained Artificial Neural Network. In: *60th International Scientific Conference, 10–11 October 2019*. Available from doi:10.1109/ITMS47855.2019.8940739.
21. Bondarenko, A., Borisov, A. The Extraction of Elliptical Rules from the Trained Radial Basis Function Neural Network. *Information Technology and Management Science*. 2012, vol. 15, pp. 161–165. Available from: doi:10.2478/v10313-012-0027-2.
22. Bondarenko, A., Borisov, A. Decompositional Rules Extraction Methods from Neural Networks. In: *Proceedings of the 16th International Conference on Soft Computing MENDEL'10, Brno, Czech Republic, 23–25 June 2010*. Brno: University of Technology, 2010, pp. 256–262.
23. Bondarenko, A., Borisov, A. Research of Artificial Neural Networks Abilities in Printed Words Recognition. *Information Technology and Management Science*. 2010, vol. 44, issue 5, pp. 124–129.
24. Bondarenko, A., Borisov, A. Research on the Classification Ability of Deep Belief Networks on Small and Medium Datasets. *Scientific Journal of Riga Technical University, Information Technology and Management Science*. 2013, vol. 16, pp. 60–65.
25. Bondarenko, A., Borisovs, A., Aleksejeva, L. Neurons vs Weights Pruning in Artificial Neural Networks. In: *Environment. Technology. Resources: Proceedings of the 10th International Scientific and Practical Conference, Rezekne, Latvia, 18–20 June 2015*. Vol. 3. Rezekne: Rezekne Higher Education Institution, 2015, pp. 22–28.

26. Bondarenko, A., Borisovs, A. Artificial Neural Network Generalization and Simplification via Pruning. *Information Technology and Management Science*. 2014, vol. 17, pp. 132–137.
27. Bondarenko, A., Borisovs, A. Elliptical Rule Extraction from a Trained Radial Basis Function Neural Network. In: *The 6th International Conference “Applied Information and Communication Technology”, Jelgava, Latvia, 25–26 April 2013*. Jelgava: LUA Faculty of Information Technology, 2013 (CD-ROM).
28. Bondarenko, A., Borisovs, A. Knowledge Extraction from Piecewise-Linear Approximation of Multi-Surface Classifier. In: *International Conference “Information Intelligent Systems”, Kharkov, Ukraine, 17–19 April, 2012*. Vol. 6, pp. 5–6.
29. Bondarenko, A., Jumutc, V. Extraction of Interpretable Rules from Piecewise-Linear Approximation of a Nonlinear Classifier Using Clustering-Based Decomposition. In: *Proceedings of the 10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases (AIKED'11), Cambridge, United Kingdom, 22–22 February 2011*. Cambridge: 2011, pp. 145–149.
30. Bondarenko, A., Zmanovska, T., Borisovs, A. Piece-Wise Classifier Application to RBF Neural Network Rules Extraction. In: *17th International Conference on Soft Computing (MENDEL'11), Brno, Czech Republic, 15–17 June 2011*. Brno: Brno University of Technology, 2011, pp. 170–176.
31. Bondarenko, A. *NNKX – Neural Networks Knowledge eXtraction* [online]. Publisher: bitbucket.org, 2018 [viewed 1 June 2019]. Available from: <https://bitbucket.org/bondtnt/nnkx>.
32. Boyd, S., Vandenberghe, L. *Convex Optimization*. Cambridge: Cambridge University Press, 2004. 727 p.
33. Brockman, J. *Thinking: The New Science of Decision-Making, Problem-Solving, and Prediction*. Harper Perennial, 2013. 432 p. (Best of Edge Series).
34. Buhmann, M. D. *Radial Basis Functions: Theory and Implementations*. Cambridge University, 2003. 272 p.
35. Castro, J. L., Mantas, C. J., Benítez, J. M. Interpretation of Artificial Neural Networks by Means of Fuzzy Rules. *IEEE Transactions on Neural Networks*. 2002, vol. 13, no. 1, pp. 101–116.
36. Chaudhary, V., Ahlawat, A. K., Bhatia, R. S. Growing Neural Networks using Soft Competitive Learning. *International Journal of Computer Applications*. 2011, vol. 21, no. 3, pp. 1–6. Available from: doi: 10.5120/2495-3372.
37. Chen, S., Hong, X., Luk, B. L., et al. Construction of Tunable Radial Basis Function Networks using Orthogonal Forward Selection. *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*. 2009, vol. 39, no. 2, pp. 457–466.
38. Chung, J., Gulcehre, C., Cho, K. H., et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *NIPS 2014 Workshop on Deep Learning*, December 2014. Available from: <https://arxiv.org/abs/1412.3555>.

39. Craven, M. W. *Extracting comprehensible models from trained neural networks*. Ph.D. thesis. Madison: University of Wisconsin–Madison, 1996. 199 p. Available from: <https://www.biostat.wisc.edu/~craven/papers/thesis.pdf>.
40. De Backer, M., Haesen, R., Martens, D., et al. A Stigmergy Based Approach to Data Mining. In: S. Zhang, R. Jarvis, eds. *AI 2005: Advances in Artificial Intelligence: 18th Australian Joint Conference on Artificial Intelligence, Sydney, Australia, 5–9 December 2005*. Berlin Heidelberg Springer-Verlag 2005, p. 975–978. (Lecture Notes in Artificial Intelligence. Vol. 3809).
41. *Deeplearning4j: Open-source, Distributed Deep Learning for the JVM* [online]. Deeplearning4j Team, [viewed 1 June 2019]. Available from: <https://deeplearning4j.org/>.
42. Deng, L. Deep Learning: From Speech Recognition to Language and Multimodal Processing. *APSIPA Transactions on Signal and Information Processing*. 2016, vol. 5, e1, pp. 1–15. Available from: doi:10.1017/ATSIP.2015.22.
43. Denil, M., Bazzani, L., Larochelle, L., et.al. 2011. Learning where to Attend with Deep Architectures for Image Tracking. CoRR. [Online]. Available from: <https://arxiv.org/abs/1109.3737> [Accessed: 1 June 2019].
44. Dienes, Z., Perner, J. A Theory of Implicit and Explicit Knowledge. *Behavioral and Brain Sciences*. 1999, vol. 22, pp. 735–808.
45. Domingos, P. A Unified Bias-Variance Decomposition and Its Applications. In: *Proceedings of the 17th International Conference on Machine Learning, ICML'2000, Stanford, CA, USA, 29 June – 2 July 2000*. San Francisco: Morgan Kaufmann, 2000, pp. 231–238.
46. Dong, X., Chen, S., Pan, S. J. *Learning to Prune Deep Neural Networks via Layer-wise Optimal Brain Surgeon* [online]. 2017 [viewed 1 June 2019]. Available from: <https://arxiv.org/pdf/1705.07565.pdf>.
47. Dua, D. and Graff, C. *UCI Machine Learning Repository* [online]. Irvine, CA: University of California, School of Information and Computer Science, 2019 [viewed 10 June 2019]. Available from: <http://archive.ics.uci.edu/ml>.
48. Erhan, D., Bengio, Y., Courville, A., et al. Why Does Unsupervised Pre-training Help Deep Learning? *J. of Machine Learning Research*. 2010, vol. 11, pp. 625–660.
49. Fiesler, E., Cios, K. C1.7 Supervised ontogenic networks. In: E. Fiesler, R. Beale, eds. *Handbook of Neural Computation*. Oxford University Press, 1997, pp. 320–337.
50. Fukushima, K. Cognitron: A Self-organizing Multilayered Neural Network. *Biological Cybernetics*. 1975, vol. 20, no. 3–4, pp. 121–136. Available from: DOI:10.1007/BF00342633.
51. Fukushima, K. Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biological Cybernetics*. 1980, vol. 36, no. 4, pp. 193–202. Available from: DOI: 10.1007/BF00344251.

52. Fung, G., Sandilya, S., Bharat Rao, R. Rule Extraction from Linear Support Vector Machines. In: *KDD'05 Proceedings of 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, Chicago, Illinois, USA, 21–24 August 2005*. New York: ACM, 2005, pp. 32–40.
53. Ganon, Z., Keinan, A., Ruppin, E. Evolutionary Network Minimization: Adaptive Implicit Pruning of Successful Agents. In: W. Banzhaf, et al. *Advances in Artificial Life: 7th European Conference, ECAL 2003. Dortmund, Germany, 14–17 September 2003*. Springer-Verlag, 2003, pp. 319–327. (Lecture Notes in Artificial Intelligence, vol. 2801).
54. Glorot, X., Bordes, A., Bengio, Y. Deep Sparse Rectifier Neural Networks. In: *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS), 2011, Fort Lauderdale, FL, USA, 11–13 April 2011*. Vol 15 of JMLR: W&CP 15, pp. 315–323.
55. Golik, P., Doetsch, P., Ney, H. Cross-Entropy vs. Squared Error Training: a Theoretical and Experimental Comparison. In: *Proc. of 14th Annual Conf. of the Intern. Speech Communication Association (Interspeech 2013), 25–29 August 2013, Lyon, France*. Lyon, 2013, pp. 1756–1760.
56. Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*, MIT Press, 2016, 785 p. Available from: <http://www.deeplearningbook.org>.
57. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. Generative Adversarial Nets. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc. 2014, vol. 27, pp. 2672–2680. Available from: <http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>.
58. Graves, A., Wayne, G., Reynolds, M., et. al. Hybrid computing using a neural network with dynamic external memory. *Nature*. 2016, vol. 538, pp. 471–476.
59. Grossberg, S. Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*. 2012, Special Issue. Available from: <https://doi.org/10.1016/j.neunet.2012.09.017>.
60. Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L., eds. *Feature Extraction, Foundations and Applications*. Series Studies in Fuzziness and Soft Computing, Physica-Verlag, Springer, 2006. 781 p.
61. Grünbaum, B. *Convex Polytopes*. 2nd ed. Springer–Verlag, 2003. 471 p. (Series: Graduate Texts in Mathematics, vol. 221)
62. Hahnloser, R. H. R., Sarpeshakar, R., Mahowald, M. A., et al. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*. 2000, vol. 405, pp. 947–951.
63. Hailesilassie, T. Rule Extraction Algorithm for Deep Neural Networks: A Review. *International Journal of Computer Science and Information Security*. 2016, vol. 14, no. 7, pp. 376–381.
64. Hammer, B., Rechten, A., Strickert, M., et al. Rule Extraction from Self-Organizing Networks. In: *Artificial Neural Networks — ICANN 2002: Proceedings of International*

- Conference, 28–30 August 2002, Madrid, Spain.* Berlin, Heidelberg: Springer, 2002, pp. 877–883. (Lecture Notes in Computer Science. Vol. 2415/2002).
65. Hammer, B., Villmann, T. Estimating Relevant Input dimensions for Self-organizing Algorithms. In: N. Allison, H. Yin, L. Allison, J. Slack, eds. *Advances in Self-Organizing Maps*. London: Springer-Verlag, 2001, pp. 173–180.
 66. Han, S., Mao, H., Dally, W. J. *Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding* [online]. ArXiv.org, 2016 [viewed 1 June 2019]. Available from: <https://arxiv.org/abs/1510.00149>.
 67. Hawkins, J., Ahmad, S., Cui, Y. A Theory of How Columns in the Neocortex Enable Learning the Structure of the World. *Frontiers in Neural Circuits*. 2017, vol. 11 (81). Available from: <https://doi.org/10.3389/fncir.2017.00081> [Accessed 1 June 2019].
 68. Hayashi, Y., Setiono, R., Yoshida, K A Comparison between Two Neural Network Rule Extraction Techniques for the Diagnosis of Hepotibiliary Disorders. *Artificial Intelligence in Medicine*. 2000, vol. 20, no. 3, pp. 205–216.
 69. He, K., Zhang, X., Ren, S., et al. *Deep Residual Learning for Image Recognition* [online]. arXiv.org, 2015 [viewed 1 September 2017]. Available from: <https://arxiv.org/abs/1512.03385>.
 70. Hochreiter, S., Schmidhuber, J. Long Short-Term Memory. *Neural Computation*. 1997, vol. 9, no. 8, pp. 1735–1780.
 71. Hruschka, E. R., Ebecken, N. F. F. Extracting Rules from Multilayer Perceptrons in Classification Problems: A Clustering-based Approach. *Neurocomputing*. 2006, vol. 70, no. 1–3, pp. 384–397.
 72. Hu, Y. H., Hwang, J-N. Ch.1. Introduction to Neural Networks for Signal Processing. In: Y. H. Hu, J-N. Hwang, eds. *Handbook of Neural Network Signal Processing*. CRC Press, 2002, pp. 12–41.
 73. Huysmans, J., Baesens, B., Vanthienen, J. *Using Rule Extraction to Improve the Comprehensibility of Predictive Models*. Leuven, Belgium: Katholieke Universiteit Leuven, 2006. 55 p.
 74. Hyontai, S. More Balanced Decision Tree Generation for Imbalanced Data Sets including the Parkinson’s Disease Data. *Intern. J. of Biology and Biomedical Engineering*. 2016, vol. 10, pp. 115–123.
 75. Iqbal, R. A. Eclectic Rule Extraction from Neural Networks Using Aggregated Decision Trees. In: 7th Intern. Conf. on Electrical & Computer Engineering (ICECE 2012), Dhaka, Bangladesh, 20–22 December 2012. IEEE, 2013, pp. 1–5.
 76. Jacobsson, H. Rule Extraction from Recurrent Neural Networks: A Taxonomy and Review. *Neural Computation*. 2005, vol. 17, no. 6, pp. 1223–1263.
 77. Jifeng, D., Yang, L., Nian, W.Y., Generative Modeling of Convolutional Neural Networks. In: 3rd Intern. Conf. on Learning Representations (ICLR), New Orleans, USA, 6–9 May 2015. Available from: <https://arxiv.org/abs/1412.6296#> [Accessed 1 June 2019].
 78. Ji, C., Snapp, R. R., Psaltis, D. Generalizing Smoothness Constraints from Discreet Samples. *Neural Computation*. 1990, vol. 2, no. 2, pp. 188–197.

79. Jumutcs, V., Bondarenko, A. Polytope Classifier: A Symbolic Knowledge Extraction from Piecewise-Linear Support Vector Machine. In: *Knowledge-Based and Intelligent Information and Engineering Systems: 15th Intern. Conf. (KES 2011): Proc., Part 1, Germany, Kaiserslautern, 12–14 September 2011*. 2011, pp. 62–71.
80. Nunez, H., Angulo, C., Catala, A. Rule Extraction from Support Vector Machines. In: *Proceedings of the 10th European Symposium on Artificial Neural Networks, Bruges, Belgium, 24–26 April 2002*. d-side publi., 2002. pp. 107–112. ISBN 2-930307-02-1.
81. Qiang, X., Cheng, G., Wang, Z. An Overview of Some Classical Growing Neural Networks and New Developments. In: *Proceedings of the 2nd International Conference on Education Technology and Computer (ICETC 2010), 22–24 June 2010, Shanghai, China*. Vol. 3. Chengdu, China: IEEE, 2010, pp. V3-351 – V3-355.82. Quinlan, R. *C4.5: Programs for machine learning*. San Manteo, CA: Morgan Kaufman, 1993. 302 p.
83. Kahramanli, H., Allahverdi, N. Rule extraction from trained adaptive neural networks using artificial immune systems. *Expert Systems with Applications: An Intern. J. archive*. 2009, vol. 36, no. 2, pp. 1513–1522.
84. Kamruzzaman, S. M., Islam, M. M. An Algorithm to Extract Rules from Artificial Neural Networks for Medical Diagnosis Problems. *Intern. J. of Information Technology*. 2006, vol. 12, no. 8, pp. 41–59.
85. Karpathy, A., Fei-Fei, L. Deep Visual-Semantic Alignments for Generating Image Descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2014, vol. 39, no. 4. Available from: arXiv:1412.2306.[_](#)[viewed 1 June 2019].
86. Kasabov, N. Evolving Connectionist Systems: From Neuro-Fuzzy-, to Spiking- and Neuro-Genetic. In J. Kacprzyk, W. Pedrycz, eds. *Springer Handbook of Computational Intelligence*. Berlin Heidelberg: Springer, 2015, pp. 771–782. Available from: DOI 10.1007/978-3-662-43505-2.
87. Kingma, D. P., Welling, M. *Auto-encoding Variational Bayes* [online]. arXiv.org, 2013 [viewed 1 June 2019]. Available from: <https://arxiv.org/pdf/1312.6114.pdf>.
88. Kohonen, T. Learning Vector Quantization. In: A.M. Arbib, ed. *The Handbook of Brain Theory and Neural Networks*. 2nd ed. Cambridge, London: The MIT Press, 2002, pp. 631–635.
89. Kohonen, T. *Self-Organizing Maps*. 3rd ed. Berlin: Springer–Verlag, 2001. 528 p.
90. Krishnagopal, D. *The Biology of Thought: A Neuronal Mechanism in the Generation of Thought – A New Molecular Model*. Amsterdam: Academic Press, 2014. 248 p.
91. Krizhevsky, A. *Learning Multiple Layers of Features from Tiny Images*. Technical Report. University of Toronto, 2009. 60 p.
92. Krizhevsky, A., Sutskever, I., Hinton, G. Imagenet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems (NIPS 2012)*. Vol. 25. MIT Press, 2012, pp. 1097–1105.
93. Henderson, E. K., Martinez, T. R. Constructing Low-Order Discriminant Neural Networks Using Statistical Feature Selection. *Journal of*

- Intelligent Systems*. 2007, vol. 16, no. 1, pp. 27–56. Available from DOI: <https://doi.org/10.1515/JISYS.2007.16.1.27>.
94. Langley, P., Laird, J. E., Rogers, S. Cognitive Architectures: Research Issues and Challenges. *Cognitive Systems Research*, 2009, vol. 10, no. 2, pp. 141–160. Available from: <https://doi.org/10.1016/j.cogsys.2006.07.004>.
 95. LeCun, Y., Bottou L., Bengio, Y., et al. Gradient-based Learning Applied to Document Recognition. *Proc. of the IEEE*. 1998, vol. 86, no. 11, pp. 2278–2324.
 96. Liu, H., Setiono, R. Chi2: Feature Selection and Discretization of Numeric Attributes. In: *Proc. of the 7th IEEE Intern. Conf. on Tools with Artificial Intelligence, Herndon, Virginia 5–8 November 1995*. IEEE Computer Society Press, 1995, pp. 388–391.
 97. Liu, H., Tan, S. T. X2R: A Fast Rule Generator. In: *Proc. of IEEE Intern. Conf. on Systems, Man and Cybernetics, Vancouver, BC, Canada, 22–25 October 1995*. IEEE Press, 1995, pp. 1631–1635.
 98. Long, J., Shelhamer, E., Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2015), Boston, MA, 7–12 June 2015*. IEEE, 2015, pp. 3431–3440.
 99. Luger, G. F. *Artificial Intelligence: Structures and Strageies for Complex Problem Solving*. 6th ed. Pearson, 2008. 784 p.
 100. Magdalena, L. Fuzzy Rule-Based Systems. In: J. Kacprzyk, W. Pedrycz, eds. *Springer Handbook of Computational Intelligence*. Berlin Heidelberg: Springer, 2015, pp. 203–218. Available from: DOI 10.1007/978-3-662-43505-2.
 101. Manessi, F., Rozza, A., Bianco, S., Bianco, S., et al. Automated Pruning for Deep Neural Network Compression. In: *24th International Conference on Pattern Recognition (ICPR), Beijung, China, 20–24 August 2018*. IEEE, 2018. Available from: DOI:10.1109/ICPR.2018.8546129.
 102. McCulloch, W., Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 1943, vol. 5, pp. 115–133.
 103. McGarry, K. J., Tait, J., Wermter, S., MacIntyre, J. Rule-Extraction from Radial Basis Function Networks. In: *International Conference on Artificial Neural Networks, Edinburgh, UK, 7–10 September 1999*. Vol. 2. London: Institution of Electrical Engineers, 1999, pp. 613–618.
 104. Merino, E. R., Castrillejo, F. M., Pin, J. D., et al. *Weighted Contrastive Divergence* [online]. ArXiv.org, January 2018. Available from: <https://arxiv.org/abs/1801.02567>.
 105. Minar, M. R., Naher, J. *Recent Advances in Deep Learning: An Overview* [online]. 2018 [viewed 1 June 2019]. Available from: <https://arxiv.org/abs/1807.08169>.
 106. Mitchell, T. *Machine Learning*. McGraw Hill, 1997. 432 p.
 107. Molchanov, P., Tyree, S., Karras, T., et al. Pruning Convolutional Neural Networks for Resource Efficient Inference. In: *Proceedings of 5th International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April, 2017*. [Online]. Available from: <https://arxiv.org/pdf/1611.06440.pdf>.

108. Morissette, L., Chartier, S. The k-means Clustering Technique: General Considerations and Implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*. 2013, vol. 9, no. 1, pp. 15–24. Available from: DOI: 10.20982/tqmp.09.1.p015.
109. Mues, C., Baesens, B., Setiono, R., Vanthienen, J. From Knowledge Discovery to Implementation: A Business Intelligence Approach Using Neural Network Rule Extraction and Decision Tables. In: *Biennial Conference on Professional Knowledge Management*, 2005, pp. 483–495. Available from: DOI: 10.1007/11590019_55.
110. Nowlan, S. J., Hinton, G. E. Simplifying Neural Networks by Soft Weight-Sharing. *Neural Computation*. 1992, vol. 4, no. 4, pp. 473–493.
111. Omlin, C. W., Giles, C. L. Extraction of Rules from Discrete-time Recurrent Neural Networks. *Neural Networks*. 1996, vol. 9, no. 1, pp. 41–52.
112. Pengfei, Z., Qinghua, H. Rule extraction from support vector machines based on consistent region covering reduction. *Knowledge-Based Systems*. 2013, vol. 42, pp. 1–8. Available from: <https://doi.org/10.1016/j.knosys.2012.12.003>.
113. Phil, S. *Too Big to Ignore: The Business Case for Big Data*. Wiley, 2013. 231 p.
114. Prechelt, L. *Adaptive Parameter Pruning in Neural Networks*. International Computer Science Institute, Technical Report TR-95-009. Berkley, California: 1995. 14 p.
115. *Pytorch – Tensors and Dynamic Neural Networks in Python with Strong GPU Acceleration* [online]. PyTorch Team, 2017 [viewed 1 November 2017]. Available from: <http://pytorch.org>.
116. Ranzato, M’A., Poultney, C., Chopra, S., et al. Efficient Learning of Sparse Representations with an Energy-Based Model. In: B. Schölkopf, J. Platt, T. Hoffman, eds. *Advances in Neural Information Processing Systems 19*. Cambridge, MA : MIT Press, 2006, pp. 1137–1144.
117. Reed, R., Marks, R. J. *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*. MIT Press, 1999. 346 p.
118. Regulation (EU) 2016/679 of the European Parliament and of the Council [online]. *Official Journal of European Union*, 2016 [viewed 20 December 2017] Available from: http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf.
119. Rumelhart, D.E., Hinton, G., Williams, R.J. Learning Representations by Back-propagating Errors. *Nature*. 1986, vol. 323, pp. 533–536.
120. Russakovsky, O., Deng, J., Su, H., et al. ImageNet Large Scale Visual Recognition Challenge. *Intern. J. of Computer Vision*. 2015, vol. 115, no. 3, pp. 211–252.
121. Sato, A., Yamada, K. Generalized Learning Vector Quantization. In: D. S. Touretzky, M. C. Mozer, M. E. Hasselmo, eds. *Advances in Neural Information Processing Systems*. Vol. 8. Cambridge, MA, USA: MIT Press, 1996, pp. 423–429.
122. Sato M., Tsukimoto, H. Rule Extraction from Neural Networks via Decision Tree Induction. In: *Proc. IJCNN’01. Intern. Joint Conf. on Neural Networks, Washington, DC, 15–19 July 2001*. Vol. 3. IEEE, 2001, pp. 1870–1875.
123. *Scikit Learn Team, Choosing the Right Estimator* [online]. Scikit [viewed 19 May 2017]. Available from: http://scikit-learn.org/stable/tutorial/machine_learning_map/index.html.

124. Setiono, R., Baesens, B., Mues, C. Recursive Neural Network Rule Extraction for Data With Mixed Attributes. *IEEE Trans. on Neural Networks*. 2008, vol. 19, no. 2, pp. 299–307.
125. Setiono, R., Leow, W. H. Pruned Neural Networks for Regression. In: *Proceedings of PRICAI 2000, The 6th Pacific Rim International Conference on Artificial Intelligence*, Melbourne, Australia, 28 August – 1 September, 2000. Berlin: Springer-Verlag, 2000, pp. 500–509. (Lecture Notes in Artificial Intelligence. Vol. 1886). Available from: doi:10.1007/3-540-44533-1_51.
126. Setiono, R., Leow, W. K., Zurada, J. M. Extraction of Rules from Artificial Neural Networks for Nonlinear Regression. *IEEE Trans. on Neural Networks*. 2002, vol. 13, no. 3, pp. 564–577.
127. Setiono, R., Thong, J. An Approach to Generate Rules from Neural Networks for Regression Problems. *European J. of Operational Research*. 2004, vol. 155, no. 1, pp. 239–250.
128. Sethi, K. K., Mishra, D. K., Mishra, B. KDRuleEx: A Novel Approach for Enhancing User Comprehensibility Using Rule Extraction. In: *3rd International Conference on Intelligent Systems Modelling and Simulation (ISMS), Kota Kinabalu, Malaysia, 8–10 February 2012*. IEEE, 2012, pp. 55–60.
129. Sharma, A., Wolfe, N., Raj, B. *The Incredible Shrinking Neural Network: New Perspectives on Learning Representations through The Lens of Pruning* [online]. 2017 [viewed 1 June 2019]. Available from: <https://arxiv.org/abs/1701.04465#>.
130. Shinde, P. Data Mining Using Artificial Neural Network Tree. *IOSR J. of Engineering*, 2012, pp. 9–12.
131. Siebel, N. T., Botel, J., Sommer, G. Efficient neural network pruning during neuro-evolution. In: *2009 International Joint Conference on Neural Networks Atlanta, GA, USA, 14–19 June 2009*. IEEE, 2009, pp. 2920–2927. Available from: doi:10.1109/IJCNN.2009.5179035.
132. Silver, D., Huang, A., Maddison, C. J., et al. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature*. 2016, vol. 529, pp. 484–489.
133. Simonyan, K., Vedaldi, A., Zisserman, A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. In: *International Conference on Learning Representations, (ICLR 2014), Banff, Canada, 14–16 April 2014, Workshop Proceedings*. Available from: <https://arxiv.org/abs/1312.6034>.
134. Srivastava, N., Hinton, G., Krizhevsky, A., et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. of Machine Learning Research*, 2014, vol. 15, pp. 1929–1958.
135. Taur, J., Kung, S-Y., Lin, S-H. Hierarchical Fuzzy Neural Networks for Pattern Classification. In: Y. H. Hu, J-N. Hwang, eds. *Handbook of Neural Network Signal Processing*. CRC Press, 2002, pp. 348–381.
136. Theano Development Team, *Theano: A Python Framework for Fast Computation of Mathematical Expressions* [online]. arXiv e-prints, 2016 [viewed 1 November 2017]. Available from: <http://arxiv.org/abs/1605.02688>.

137. Thompson, M.-E. *NDCC: Normally Distributed Clustered Datasets on Cubes* [online]. Madison: University of Wisconsin, 2006 [viewed 15 September 2017]. Available from: <http://www.cs.wisc.edu/dmi/svm/ndcc/>.
138. Thrun, S. *The MONKs problems: A Performance Comparison of Different Learning Algorithms*. Technical Report CS-91-197. Pittsburg, PA: 1991.
139. Tickle A. B., Andrews R., Golea M., et al. The Truth will come to Light: Directions and Challenges in Extracting the Knowledge Embedded within Trained Artificial Neural Networks. *IEEE Trans. on Neural Networks*. 1998, vol. 9, no. 6, pp. 1057–1067.
140. Tickle, A. B., Orlowski, M. J., Diederich, J. DEDEC: A Methodology for Extracting Rules from Trained Artificial Neural Networks. In: *Proceedings of the Rule Extraction from Trained Artificial Neural Networks (AISB'96), Brighton, UK, 2 April 1996*. Brighton, University of Sussex, 1996, pp. 90–102.
141. Tsukimoto, H. Extracting Rules from Trained Neural Networks. *IEEE Transactions on Neural Networks*. 2000, vol. 11, no. 2, pp. 377–389.
142. Van der Maaten, L. J. P., Hinton, G. E. Visualizing High-Dimensional Data Using t-SNE. *J. of Machine Learning Research*. 2008, vol. 9, pp. 2579–2605.
143. Wan, W., Mabu, S., Shimada, K., Hirasawa, K. Enhancing the Generalization Ability of Neural Networks through Controlling the Hidden Layers. *Applied Soft Computing J*. 2009, vol. 9, no. 1, pp. 404–414.
144. Weigend, A. S., Rumelhart, D. E., Huberman, B. A. Generalization by Weight-Elimination Applied to Currency Exchange Rate Prediction. In: *Proceedings of IJCNN-91-Seattle International Joint Conference on Neural Networks, Seattle, Canada, 8–12 July 1991*. Vol. 2. IEEE, 1991, pp. 837–841.
145. Xing, H.-J., Hu, B.-G. Two Phase Construction of Multilayer Perceptrons using Information Theory. *IEEE Trans. on Neural Networks*. 2009, vol. 20, no. 4, pp. 715–721.
146. Xu, S., Chen, L. A Novel Approach for Determining the Optimal Number of Hidden Layer Neurons for FNN's and Its Application in Data Mining. In: D. Tien, M. Kavakli, eds. *5th Intern. Conf. on Information Technology and Applications (ICITA 2008), Cairns, Queensland, Australia, 23–26 June 2008*. [electronic resource]. Bathurst, N.S.W.: Macquarie Scientific Publishing, 2008, pp. 683–686. Available from: <https://eprints.utas.edu.au/6995/1/02-au-xu.pdf>.
147. Zeiler, M. D., Fergus, R. Visualizing and Understanding Convolutional Networks. In: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars, eds. *Computer Vision - ECCV 2014: 13th European Conf., Zurich, Switzerland, 6–12 September 2014*. Part I. Springer, 2014, pp. 818–833. Available from: https://doi.org/10.1007/978-3-319-10590-1_53.
148. Zhang, H., Xu, T., Li, H., et al. *StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks* [online]. arXiv.org, 2016 [viewed 7 June 2017]. Available from: <https://arxiv.org/pdf/1612.03242.pdf>.
149. Zhang, Z., Qiao, J. A Node Pruning Algorithm for Feedforward Neural Network Based on Neural Complexity. In: *Proceedings of 2010 International Conference on Intelligent*

Control and Information Processing (ICICIP), Dalian, China, 13–15 August 2010.
IEEE, 2010, pp. 406–410.

150. Zilke, J. R., Mencia, E. L., Janssen, F. DeepRED–Rule Extraction from Deep Neural Networks. In: T. Calders, M. Ceci, D.Malerba, eds. *Discovery Science: Proceedings of the 19th International Conference, DS 2016, Bari, Italy, 19–21 October 2016.* Springer, 2016, pp. 457–473 (Lecture Notes in Artificial Intelligence. Vol. 9956). Available from: http://dx.doi.org/10.1007/978-3-319-46307-0_29.