



IKER
GAZTE
NAZIOARTEKO
IKERKETA EUSKARAZ

VI. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2025eko maiatzaren 28, 29 eta 30a
Bilbo, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)



Aitortu-PartekatuBerdin 4.0

INGENIARITZA ETA ARKITEKTURA

**Neuroeboluzio Informazioaren
Ustiapena: Iragana Aztertu
Etorkizun Eraginkor Baterako**

*Unai Garciarena, Penousal
Machado, Nuno Lourenço,
Alexander Mendiburu
eta Roberto Santana*

85-92 or.

<https://dx.doi.org/10.26876/ikergazte.vi.03.10>

ANTOLATZAILEA



BABESLEAK



EUSKO JAURLARITZA
GOBIERNO VASCO



LAGUNTZAILEAK



Neuroeboluzio Informazioaren Ustiapena: Iragana Aztertu Etorkizun Eraginkor Baterako

Unai Garciarena¹, Penousal Machado², Nuno Lourenço², Alexander Mendiburu¹, Roberto Santana¹

¹ Intelligent Systems Group, Informatika Fakultatea, Euskal Herriko Unibertsitatea, Donostia

² CISUC, Universidade de Coimbra, Coimbra

unai.garciarena@ehu.eus

Laburpena

Neuroeboluzio algoritmoak, neurona sareen egiturak optimizatzeko teknika ebolutiboak erabiltzen dituzten bilaketa automatikoak, prozedura konputazionalak garestiak dira. Hala ere, lortzen diren arkitekturen errendimendua bikainak direla eta, metodo horiek oso erabiliak dira. Neuroeboluzio-prozesuen azken emaitzat hartzen dena aurkitutako egiturarik onena izan ohi da, sortzen den gainerako informazio guztia baztertuz. Bilaketa horietan *informazio erresidual* ugari ekoizten da, eta horrek balio handiko ezagutza eskaini dezake. Lan honetan, informazio hori neuroeboluzio-prozesuetatik erauzteko eta meta-eredu bat eraikitze hurbilketa bat proposatzen da, etorkizuneko neurona sare bilaketetan eragin positiboa izan dezakeena. Eredu sortzaile baten neuroeboluzioan aurkitutako egiturarik onenak aztertuz, sare Bayesiar batean oinarritutako meta-eredu bat proposatzen da. Eredu honek neurona sare sendoak arinago aurkitzeko, etorkizuneko arkitektura optimizazio prozedurak hasieratzeko, edo edozein motatako egituren optimizazio prozesuei laguntzeko erabil daiteke.

Hitz gakoak: Sare antagoniko sortzailea, eredu grafiko probabilistikoak, neuroeboluzioa

Abstract

Neuroevolutionary algorithms, using evolutionary techniques to optimize neural network structures, are computationally costly but widely applied due to their strong performance. Typically, the best-found structure is the only outcome of the algorithm, while valuable residual information from the search is overlooked. This paper proposes a method to extract and leverage this information to build a metamodel that enhances future neural architecture searches. By analyzing optimal generative adversarial network structures, a Bayesian network-based meta-model is proposed, which can improve search initialization, and overcome local optima in structural optimization.

Keywords: Generative adversarial networks, probabilistic graphical models, neuroevolution

1 Sarrera eta motibazioa

Neurona sare sakon (NSS) ereduak azken hamarkadan irudiak sailkatzeko erakutsi zuten gaitasunengatik (Krizhevsky *et al.*, 2012) berreskuratu zirenetik, bertze hainbat atazetan arrakastaz aplikatuak izan dira. Hala nola, sekuentzia datuen tratamendua (Hochreiter eta Schmidhuber, 1997) edo sare antagoniko sortzailea (SAS) (Goodfellow *et al.*, 2014). Ikerlarienezko baliabide konputazionalak zabaldu diren heinean, NSS egitura gero eta konplexuagoak proposatu dira azken urteetan (Krizhevsky *et al.*, 2012), eta nahiz eta NSS onenek oraindik eskuz diseinatuta egon, (Szegedy *et al.*, 2016), honek eredu horien eskuzko diseinua ezinezkoa bihurtzea eragin du. Honek, aldi berean, neurona sareen bilaketa (NSB) (Elsken *et al.*, 2017) ikerketa arloaren sorkuntza ekarri du, prozesu hori automatizatzea biltzen dituen. Nahiz eta ikasketa indartuan eta bilaketa lokalean oinarrituta dauden hurbilketa batzuk proposatu diren, eboluzio-algoritmoen (EA) erabilerak ikerketa eremu honen zati handia hartu du (Miikkulainen *et al.*, 2019). Honek neuroeboluzio (NE) terminoa sortzera eraman gaitu.

EAek optimizazio-problema konplexuei epe motzean soluzio onak lortzeko metodo eraginkorrak direla jakina da. Normalean, populaziotan oinarritutako algoritmoak dira, eta lortu dezaketen arrakasta maila hainbat diseinu-erabakiren arabera izan daiteke: kodifikazioa, aldaketa operadoreak, lehen populazioa osatzen duten indibiduen hasieraketa (Kazimipour *et al.*, 2014), etab.

Neuroeboluzioa problema konplexuen kategorian sartzen da, NSS bat eraikitzeke diseinu-aukerak izugarri anitzak direlako, konbinazio kopurua esponentzialki handitzea eragiten duena. Hori dela eta, ausazko abiapuntua duen eta jakintza gabeko operadoreetan oinarritzen den optimizazio-prozesu batek urrats ugari beharko ditu soluzio onak aurkitu ahal izateko. Algoritmoari emandako informazio gehigarria funtsezkoa izan daitekeen testuinguru batean, lan honen helburua dagoeneko burututako eboluzio-bilaketetatik ahalik eta etekin handiena ateratzeko moduak aztertzea da. Horretarako, aurretik helburu desberdinekin eboluzionatutako SAS-en bi erabilera-kasu erabiliko ditugu; irudiak sortu (Costa *et al.*, 2019) eta Pareto multzoaren (PM) hurbilketak lortzea (Garciaena *et al.*, 2020). Bilaketa ebolutibo horietan aurkitutako SAS onenen ezaugarriak erregistratzen ditugu. Konfigurazio horiek modelatuz, errendimendu handiko SAS-etan desiragarriak diren atributuak zehaztu daitezke, eta hori baliagarria izan daiteke NSS berriak sortzeko (adibidez, NE algoritmoetako indibiduoak hasieratzean, edo funtzionatuko dutela aurreikusten diren egiturak diseinatzean). Gainera, eredu gai bada emaitza onak emango dituzten SAS egiturak identifikatzeko, ebaluazio ugari saihestu litezke, etorkizuneko bilaketa estrukturaletan merketzen. Horrek NE algoritmoen kostu konputazionala (daukaten desabantaila nagusietako bat (White *et al.*, 2020)) murriztuko luke.

Eredu grafiko probabilistikoak (EGP) (Pearl, 1988) lan hau burutzeko tresna egokia da, SAS baten egitura osatzen duten osagaien arteko menpekotasun handiak direla eta. Zehazki, sare Bayesiarrak (SB) (Castillo *et al.*, 1997) erabiliko ditugu. Aurreko bi abantailaz gain, SB erabiltzeak hirugarren abantaila bat ere badu: grafo moduko egitura bat sortzen dute, eta hori ikusmenez interpreta daiteke, SAS batek kalitate handiko emaitzak sortzeko behar dituen ezaugarriak hobeto ulertzeko aukera emanez. Eredu hori hainbat helburutarako erabil daiteke: dagoeneko ezagutzen den SAS egitura baten kalitatearen estimazioa lortzeko, probabilitatearen bidez errendimendu handiko egiturak sortzeko, edo NSB bilaketak gidatzeko.

2 Arloko egoera eta ikerketaren helburuak

2.1 Sare Antagoniko Sortzailea

SAS tradizional bat bi NSS kontrajarritan oinarritzen da: sorgailu bat (S) eta diskriminatzaile bat (D). Azken honek benetako datuak eta faltsuak bereizten ikasi nahi duen bitartean, sorgailuaren helburua diskriminatzailea engainatzea da, benetako datuekin ahalik eta antzekotasun handiena duten laginak sortuz.

Sortu nahi diren datuen arabera, NSS-ek arkitektura mota ezberdinak har ditzakete. Datuak denboran zehar aldatzen badira, SAS-ak neurona sare errepikakorretan oinarritu daitezke (Yoon *et al.*, 2019). Datuen ezaugarriak espazioan erlazionatuta badaude (aurpegi baten irudian bezala), (iraulitako) neurona sare konboluzionalak erabili ohi dira (Karras *et al.*, 2018). Azkenik, datuen egitura-menpekotasunik ezagutzen ez bada, geruza anitzeko perzeptroiak (GAP) aukera lehenetsi gisa erabiltzen dira (Cybenko, 1989). Arkitekturak guztiz ezberdinak izan daitezkeen arren, SAS baten bi azpi-ereduek elkarrekin lan egiteko era ez da aldatzen. S sorgailua (bere parametroak θ_S izanda) ausazko zarata $\mathbf{z}$ batetik abiatuta laginak sortzen ditu datu errealeen espazioan: $S(\mathbf{z}, \theta_S) \rightarrow \hat{\mathbf{x}}$.

Bestalde, diskriminatzaileak (θ_D parametroekin) benetako datu-multzotik $\mathbf{x}$ edo sorgailuak sortutako datuetatik $\hat{\mathbf{x}}$ sarrerak jasotzen ditu eta laginak benetazkoak izateko daukaten probabilitateak kalkulatzeko dituzte.

SAS-ak entrenatzeko erabil daitezkeen *minmax* joko funtzioa ekuazio honetan agertzen da:

$$\min_{\theta_S} \max_{\theta_D} = \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log(D(\mathbf{x}))] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(S(\mathbf{z})))] \quad (1)$$

$\mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})}$ eta $\mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})}$ jatorrizko datu p_{data} eta latentearen banaketa $p_{\mathbf{z}}(\mathbf{z})$ -rekiko itzaropenak izanda.

2.2 Sare Bayesiarrak

SB-ak aldagai-multzoen arteko (in)dependentziak adierazteko erabiltzen diren eredu grafikoak dira. Eredu horiek ziklorik gabeko zuzendutako grafoen (ZGZG) bidez irudikatzen dira (Pearl, 1988). ZGZGaren nodo bakoitzak aldagai bat adierazten du, eta bi nodoen arteko arku baten bidez, haien arteko dependentzia (edo independentzia) adierazten da. SB baten grafo egitura aditu batek ematen duenean, parametroak soilik ikasten dira. Hori horrela ez bada, egitura ere datuetatik ondorioztatu behar da teknika desberdinak erabiliz.

Beraz, SB baten ikasketa automatikoa bi fasetan banatzen da: i) sarearen topologia diseinatzea, hau da, zein konexio egon behar diren eta zein ez, eta ii) aldagai bakoitzaren baldintzapeko edo probabilitate marjinalak ikastea

(Larrañaga eta Lozano, 2001). Lan honetan, SB-ak NSS-ak eraikitze erabilitako parametroen arteko harreman probabilitistoa kodifikatzeko erabiltzen ditugu, hala nola aktibazio funtzioak, geruzen tamaina edo pisuen hasieratze prozedurak. SB bat NSS desberdinen eboluzioan bildutako datuetatik ikasi ondoren, inferentzia egingo dugu konfigurazio probabilitaterik handiena dutenak bilatzeko edo NSS zehatz baten probabilitatea kalkulatzeko.

2.3 SAS-en NSB

(Garciaarena *et al.*, 2020) lanean GAPetan oinarritutako SAS-en eboluzioa proposatzen da. Nahiz eta proposamen honek goian aipatutakoarekin antzekotasun kontzeptual batzuk partekatu, hala nola NSS-en kodifikazioa (pisuak ez dira eboluzionatzen, eta NSS-ak geruzen zerrenden bidez zehazten dira, haien tamaina eta aktibazio funtzioaren arabera kodetuta), kasu honetan NSS-ak ez dira elkarrengandik isolatuta eboluzionatzen. Indibiduo bakoitzak SAS funtzional oso bat osatzen du, bi NSS-ekin.

Aldaketa operadoreek ere zenbait ezaugarri amankomunean dituzte: mutazio operadoreek geruza gehitzea, ezabatzea eta aldatzea barne hartzen dute (azken operadore honek geruza osoa aldatzeaz gain, geruza baten zati batzuk ere alda ditzake). Horrez gain, gurutzaketa operadore bat definitzen da, SAS bateko bi NSS-ak elkartrukatzuz, banakoen atal ezberdinen arteko interakzioa ahalbidetuz.

SAS-en gaitasunak alderantzizko sorkuntza-distantzia (ASD) (Coello Coello eta Reyes Sierra, 2004) neurtuz ebaluatzen dira. Horretarako, bi helburu dituen optimizazio problema batzuen ((Li eta Zhang, 2008)) Pareto multzoan dauden soluzioak, eta SAS-ak sortutako soluzioen arteko distantzia kalkulatu da, SAS-ak puntu horiek sortzeko entrenatu ondoren. Lan originalean ((Li eta Zhang, 2008)) eskuragarri dauden 9 funtzioetatik 8 erabili ziren. Gainera, problemek aldagaien kopuru arbitrarioa daukenez, bi tamaina desberdin aztertu ziren: 10 eta 784.

Eboluzionatutako SAS-en kalitatea egiaztatzeko, lortutako egiturak 2D 8-Gaussiarren hurbilketa-probleman (Metz *et al.*, 2016) aplikatu ziren, arrakasta maila altuarekin.

3 Ikerketaren muina

Problema bakoitzarentzat NSS arkitektura egokiena desberdinela dela, eta eredu neuronaletako osagai desberdinen artean menpekotasunak daudela jakina da. Ezaugarri hori bereziki aplikagarria da SAS-etan, bi NSS-z osatuta baitaude (kasu orokorrean), eta horrek parametro kopurua bikoizten du. Zehatzago esanda, SAS bateko NSS-ak hainbat ezaugarriren arabera definitzen dira: entrenamendu-maiztasuna, NSS bakoitzaren geruza kopurua eta haien ezaugarriak (aktibazio-funtzioak eta hasierako pisuak, GAPetan neuronak zenbat diren, konboluziozko neurona sareen iragazkien kopurua, tamainak eta urratsak, etab.). Gainera, bertze parametro *global* batzuk ere badaude, hala nola, diskriminatzaileari elikatzen zaion zaratak jarraitzen duen banaketa edo ereduaren galera funtzioa.

3.1 Meta-ereduaren aukeraketa

Kalitate handiko SAS-en ezaugarriak adierazteko ereduak SAS-en egituretan dauden menpekotasunak islatu behar ditu. EGP-ek baldintza hori betetzen dute, eta lan honetan eredu hori hautatu da hiru arrazoi nagusirengatik.

Lehenik, berez modularrak direlako, eta horri esker, menpekotasun-egitura konplexuak elementu sinpleen konbinazio egokien bidez deskribatu eta kudeatu daitezke. Bigarrenik, bisualki adierazgarriak direlako, eta horrek azken produktua —laginak— sortzeko prozesua ulertzen lagundu dezake (Lauritzen, 1996). Azkenik, eredu honetan *aditu ezagutza* sartzeko modu desberdinak daudelako.

NSB algoritmo gehienek sakonera aldakorrek NSS-ak garatzeko aukera ematen dute, eta NSS batek zenbat eta geruza gehiago izan, orduan eta osagai-zerrenda luzeagoa behar da bere ezaugarriak zehazteko. Honengatik, NSS-en ezaugarrien bektoreak dinamikoak izan behar dira. Aldagai kopuru aldakorra duen datu-multzoekin SB-ak kudeatzea ez da erraza.

Hori dela eta, lan honetan proposatutako meta-eredua bi mailatan diseinatzea erabaki dugu. Lehenengo mailan, *goi-ereduan*, sorgailuaren eta diskriminatzailearen geruza kopurua finkatzen da. Bigarrenik, SB bat ikasten da NSS sakonera bakoitzerako. SAS bat entitate bakar gisa hartzen bada, bi sareen sakonera-konbinaketa bakoitzerako *azpi-eredu* bat ikasten da. Aldiz, sorgailuak eta diskriminatzaileak banaka eboluzionatzen badira, sare bakoitzaren sakonera bakoitzerako SB bakarra behar da. 3.1. Irudian meta-ereduaren irudikapen grafikoa erakusten da.

Lan honetan, alde zuzeneko esperimentazio baten ondoren, ARACNE algoritmoa erabiltzea erabaki dugu (Chow-Liu algoritmoaren (Chow eta Liu, 1968) luzapena) (Margolin *et al.*, 2006), SB egitura ikasteko. SB-ekin lotutako inplementazio guztia **bnlearn** R liburutegian (Scutari eta Denis, 2014) oinarrituta dago.

3.2 Esperimentuak

Atal esperimental honen helburua proposamenaren baliozko-tasuna hainbat mailatan aztertzea da, hain zuzen ere, meta-ereduak SAS *egitura onak* eta *txarrak* bereizteko duen gaitasuna, egitura sendoak lagintzeko duen ahalmena, eta laginketen NE prozedura baten abiapuntu gisa izateko egokitasuna.

Atal esperimental hau (Garciaarena *et al.*, 2020) lanean gauzatutako optimizazio prozeduretan oinarritzen da (ikus 2. Atala). Orokorrean, atal honetarako, guztira $8 \text{ funtzio} \times 2 \text{ aldagai} - \text{tamaina} \times 30 \text{ exekuzio} = 480 \text{ exekuzio}$ daude. (Li eta Zhang, 2008) lanean aurkeztzen diren funtzioen artean, lan honetan erabiltzen direnak honakoak dira: $F1$, $F2$, $F3$, $F4$, $F5$, $F7$, $F8$, eta $F9$.

3.3 SAS aukeraketa

Meta-eredua SAS-en ezaugarriekin elikatzeke eta ondoren ahalik eta erabilgarriena izan dadin, aurreko lanean ebaluatutako SAS multzo osoa hiru azpi-multzotan banatu dugu. Lehenengo multzoa, *Lehen*, exekuzio bakoitzean hobekien funtzionatu duten n indibiduoek osatzen dute. Bigarren multzoa, *Bigarren*, bigarren n indibiduoek osatuta dago, hau da, hautatutako kalitate-metrikaren arabera ordenatutako rankingean n -garrenetik $2n$ -garrenera daudenak. Azkenik, *Ausaz* multzoak exekuzio bakoitzeko n ausazko indibiduo biltzen ditu. *Lehen* multzoa meta-eredua ikasteko erabiliko da, *Bigarren* multzoak bere orokortze-gaitasuna probatzeko balio du, eta *Ausaz* multzoak kontrol-multzo gisa jokatu du, meta-eredua potentzial handiko indibiduoak arruntetatik bereizteko gai den ala ez egiaztatzeke. Kasu honetan, $n = 5$ indibiduo erabiltzen ditugu.

Hurbilketa honetan, SAS baten bi NSS-ak batera eboluzionatzen dira, gehienez 11 geruza ezkutu izan ditzakete eta gutxienez bat, 100 konbinazio posible emanez. Ikusi dugu *Lehen* multzoko GAP-SAS-en %75 baino gehiago hiru geruza arteko sakonera duten sorgailu eta lau geruza baino gehiago ez dituzten diskriminatzaileekin osatuta daudela. Eraikitako ereduaren konplexutasuna eta interpretagarritasuna murrizteko, SAS-ak hamabi sakonera-konbinaziotara mugatzea erabaki dugu ($3 \times 4 = 12$), guztira izan zitezkeen 100 konbinazioetatik ($10 \times 10 = 100$).

Entrenamendu-maiztasuna eta geruza bakoitzeko neuronak SAS bat deskribatzen duten aldagai ordinal baka-rrak dira, eta bost balioetara diskretizatu dira. Honen luzapena etorkizunean ikertu daiteke.

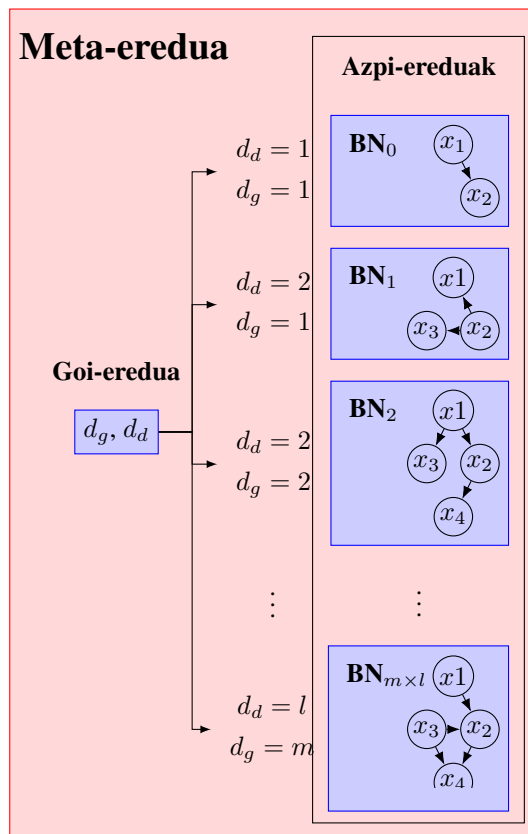
3.4 Meta-eredua frogatzen

Lehenik eta behin, gure hurbilketa GAP-SAS domeinuan garatu eta baliozkotzen dugu. Proposatutako meta-ereduak SAS-en osagaien arteko menpekotasunak zein neurritan harrapatu dituen zehazteko bi metodo ezberdin definitu ditugu. Lehen urratsa meta-eredua kalitate handiko SAS-ak *identifikatzeko* gai den egiaztatzea da, eta behin hori eginda, halako eredu sortzaileak *sortzeko* gai den aztertzea.

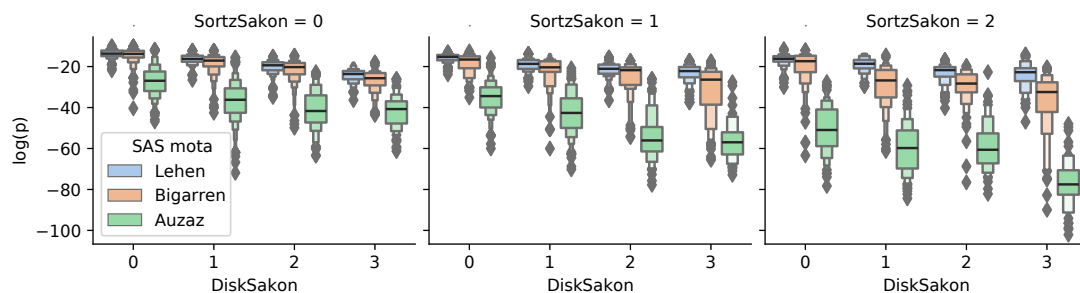
3.5 SAS probabilitatea

Meta-ereduak ikasitakoa baliagarria izan daitekeen ala ez egiaztatzeke modu eraginkor bat, meta-ereduak *Lehen* (meta-eredua ikasteko erabilitako SAS multzoa), *Bigarren* eta *Ausaz* multzoetako SAS-ei esleitutako probabilitateak alderatzea da. Meta-ereduak *Lehen* eta *Bigarren* multzoetako SAS-ei antzeko probabilitate-balioak ematen

1. Irudia: Meta-ereduaren errepresentazio grafikoa. SB adibideek ez dute benetako eszenatoki bat errepresentatzen nodo kopuruari dagokionez. l eta m balioek diskriminatzaileen eta sorgailuaren geruza ezkutu gehienezko kopurua adierazten dute, hurrenez hurren.



2. Irudia: Meta-ereduak hiru multzo definituetan dauden SAS-ei esleitutako probabilitateak (eskala logaritmikoan).



badizkie, eta *Bigarren* eta *Ausaz* multzoen artean balio horiek desberdinak badira, orduan ondoriozta daiteke meta-eredua gai dela bai *kalitate oneko* egitura ezezagunetara orokortzeko, bai *kalitate txarreko* egiturak bereizteko.

3.5 Irudian hiru kaxa-irudi erakusten dira (sorgailuaren sakoneraren arabera), non meta-ereduak multzo bakoitzeko SAS-ei esleitutako probabilitateak bistaratzen diren (y ardatzean, eskala logaritmikoan). X ardatzak diskriminatzailearen sakonera adierazten du. Irudian ikus daitekeen moduan, meta-ereduak antzeko probabilitate-balioak esleitzen dizkie *First* eta *Second* multzoetako SAS-ei, eta, aldi berean, gai da *Random* multzoko SAS-ak bereizteko, horiei probabilitate nabarmen baxuagoak esleitzuz.

Probabilitate-multzoen arteko alde esanguratsuak Kruskal-Wallis test estatistikoaren bidez baieztatu dira (Kruskal eta Wallis, 1952). Lortu den p-balio altuena $\sim 10^{-27}$ -koa izan da, aztertutako $4diskr. \times 3sorgailusakonerak = 12$ konbinaketa desberdinetarako. Dunn-en test estatistikoak (Dunn, 1964) bikoteko alderaketa guztiak desberdinak direla erakutsi du (aurkitutako gehieneko p-balioa ~ 0.04 -koa izan da). Hala ere, Dunn testaren p-balioen arteko aldeak oso aldakorak izan dira alderatutako bikoteen arabera. *Lehen* eta *Bigarren* multzoen artean, p-balioak ~ 0.04 eta $\sim 3 \times 10^{-26}$ artean ibili dira, baina edozein multzo horietatik *Ausaz* multzora egindako konparazioek p-balioak $\sim 10^{-16}$ eta $\sim 10^{-161}$ bitartean egon dira. Honek erakusten du *Lehen* eta *Bigarren* multzoetako SAS-en probabilitateen arteko *tartea* askoz estuagoa dela, edozein multzo horietatik *Random* multzora dagoen distantziarekin alderatuta.

3.6 Laginak sortzeko gaitasuna

Meta-ereduaren laginketa-gaitasunak ebaluatzeko, lehenik eta behin, *Lehen* eta *Ausaz* multzoen bi aldaera definitzen dira. $Lehen_E$, *Lehen* multzoaren bertsio mugatu bat da, multzoaren 480 egiturak erabili beharrean, soilik 300 kontuan hartzen dira, 8 funtzioetatik 5 funtzio (F1-F4 eta F7) osatzen duten multzoari dagozkionak. $Ausaz_E$ era berean osatzen da, baina aurkitutako onenak izan beharrean, n ausazko SAS ditu. Multzo bakoitzean funtzioak hautatzeko irizpidea (Garciaarena *et al.*, 2020) lanean ondorioztatutako funtzioen antzekotasunean oinarritzen da.

Meta-eredu berria FirstE multzo murriztua erabiliz ikasten da, eta ondoren 100 aldiz lagintzen da. Hori, probabilitatezko logika laginketa (Henrion, 1988) erabiliz lortzen da, aldagaiak haien arbaso-ordena jarraituz lagintzen dituen metodoa erabiliz. Laginketa horiek osatzen dute *Lagin* multzoa. Bertze 100 SAS ausaz hautatzen dira FirstE eta AusazE multzoetatik. Egitura horiek guztiak entrenatzen dira Pareto multzoaren hurbilketa erreproduzitzeko, prozedura honetatik isolatu diren funtzioetarako (F5, F8 eta F9), bi aldagaien tamainetarako. SAS-ak eboluzionatzeko erabilitako egokitasun-funtzioa, SAS-ek lortutako laginen eta benetako Pareto Frontearen (PF) arteko ASD balioak, eta egitura horiek arazo desberdinetan lortutako emaitzak 3.6. Irudian erakusten dira.

Irudian ikus daitekeen bezala, emaitzak aldatu egiten dira erabiltzen den funtzioaren arabera, eta koherenteak dira problemaren tamainarekin. F5 funtzioari dagokionez, non ASD balio baxuak erregistratu ziren, lagindutako SAS-ek emaitza askoz hobeak lortu zituzten, ausazko SAS-ak nabarmen gaitutuz, eta *Lehen* multzotik hautatutako SASek baino hobeto aritu zirela argudiatu daiteke. Hori ez da gertatzen F8 funtzio zailagoarekin, zeinerako SAS onenek emaitzarik onenak lortu zituzten, baina ausazko SASek emaitza antzekoak eskaini zituzten lagindutakoekin alderatuta. F9 funtzioarekin egindako probek ez zuten *Lehen* eta lagindutako SAS-en artean hobia zein den argi eta garbi bereizi, baina bi multzo horiek nabarmen nagusitu ziren ausazko SAS-en aurrean.

Kruskal-Wallis probaren emaitzek desberdintasun esanguratsuak erakutsi zituzten ASD balioen sei hirukote guztietan. Dunn probaren emaitzak 3×10^{-7} baino txikiagoak izan ziren F5 funtziorako egindako bikoteka alderaketa guztietan, *Lehen* eta *Lagin* arteko konparaketan izan ezik (10 aldagai dituzten kasuan), non p-balioa 0.12 izan zen. Honek erran nahi du kasurik txarrean ere lagindutako SAS-ek onenekin pareko emaitzak lortu zituztela, hobeak ez badira. Ausazko SAS-ek, aldiz, ez zuten emaitza lehiakorrik eskaini.

F8 funtzioari dagokionez, ausazko eta lagindutako SAS-en arteko konparaketek p-balioak 0.36 eta 0.96 izan zituzten, hurrenez hurren, 10 eta 784 aldagaie-tarako. Gainerako konparaketek 0.002-ko p-balioak edo txikiagoak eman zituzten. Kasu honetan, lagindutako SAS-ek ez zuten ausazkoek baino emaitza hobeak eskaini, eta bi multzo horiek maila bat beherago kokatu ziren SAS onenekin alderatuta.

Azkenik, F9 funtzioari dagokionez, SAS onenen eta lagindutakoen arteko konparaketek ez zuten diferentzia argirik erakutsi (~ 0.2 p-balioekin), baina gainerako konparaketenak 5×10^{-8} baino txikiagoak izan ziren. F5 funtzioarekin ateratako ondorio berdinak atera daitezke kasu honetan ere.

Oro har, analisi honen ondorio nagusia ASD metrikaren arabera (hau da, *Lehen* multzoko SAS-en eboluzioaren minimizatutako metrika), lagindutako SAS-ek batzuetan ausazko SAS-ek baino emaitza okerragoak eskaintzen dituztela (F8, 784 aldagaiekin) da. Hala ere, kasu gehienetan, lagindutako SAS-en errendimendua *Lehen* multzoko SAS-ena bezalakoa da (F9 edo F8, 10 aldagaiekin), edo are hobe (F5). Interesgarria da ausazko egitura txiki batzuek sofistikatuenekin antzeko errendimendua eskaini dutela ikustea. Horrek SAS egitura inolako ezagutzarik gabe eraikitzean ere ikasketa-prozedura eraginkorrak burutu daitezkeela adierazten du.

3.7 Meta-eredua SAS arkitektura-bilaketak hobetzeko

Aurreko atalean, proposatutako meta-ereduaren gaitasun handiko SAS-ak sortzeko ahalmena erakutsi da. Orain, gaitasun hori proban jartzen dugu guztiz desberdina den problema batean: bertze NE algoritmo baten hasierako populazioak hasieratzeko gaitasuna. Horretarako, SAS egituren bilaketa bat diseinatu dugu 2D 8 eta 25-Gaussian hurbilketa problemarendako (Metz *et al.*, 2016). Problema hori bereziki SAS-en *moda kolapso* (eredu sortzaile horien arazo esanguratsuenetako bat) fenomenoaren agerian uzteko diseinatuta dagoena. Hori dela eta, (Garcia-ena *et al.*, 2020) lanean proposatutako eboluzio-prozesuaren bertsio murritzitako bat exekutatu dugu, 20 indibiduoko populazioak 20 belaunaldiz eboluzionatzen dituen.

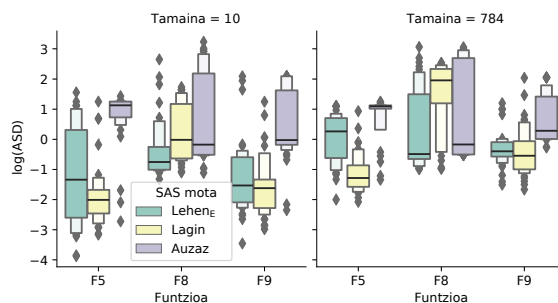
Hiru NE algoritmo bertsio exekutatu dira, eta haien arteko desberdintasuna hasierako populazioa osatzeko era da. Aurreko atalean definitutako multzoen antzekoak erabili dira (*Lehen*, *Lagin* eta *Ausaz*), baina oraingoan meta-eredua ikasteko exekuziorik ez da baztertu. NE algoritmoaren aldaera bakarra bakoitzak hasierako populazioa betetzeko, multzo desberdinetako indibiduak erabili dituztela da.

Eboluzio-prozesu bakoitza 30 aldiz exekutatu da. 3.7 Irudiak belaunaldi bakoitzean aurkitutako SAS onenaren bilakaera erakusten du Gehieneko Batez Bertzeko Desadostasuna (GBBD) (Gretton *et al.*, 2012) neurriaren arabera. GBBD funtzioa SAS-en belaunaldien kalitatea ebaluatzeko egokitasun-neurri gisa erabiltzen da (bi-helburuko eboluzio-prozesuaren bigarren helburua SAS-en entrenamendu eta laginketa denboraren minimizazioa izanik). $2dimentsio \times 3hasieratze = 6$ exekuzio motetarako GBBD hobekuntza erakusten da.

Irudian ikus daitezkeen bezala, ausazkoak ez diren hasieratzeekin egindako eboluzio-exekuzioek abantaila nabarmena dute eboluzioaren hasierako faseetan. Egia bada ere ausazko hasieratzea duten exekuzioek irabazi handiagoak lortzen dituztela prozeduran zehar bertze bi algoritmoekin alderatuta, hori ez da nahikoa emaitza lehiakorrik lortzeko. Antzeko zerbait gertatzen da *Lehen* eta *Lagin* SAS-ekin hasitako exekuzioekin, izan ere, *Lagin* exekuzioaren hasierako populazioan dauden SAS-ak ez ditu gainditu *Lehen*-dik hasieratutako exekuzioak.

Irudi honek argi erakusten du meta-eredua NE prozedura baten hasierako populazioa sortzeko erabiltzeko dituen abantailak. Zeregin espezifikoetarako eboluzionatutako SAS-ak erabili beharrean, pro-

3. Irudia: ASD balioak *Lehen* multzoko SAS-ek, meta-eredutik lagindutako eta *Ausaz* multzoko SAS-ek lortutakoak.



blema desberdinetarako exekuzio anitzetan eboluzionatutako SAS-en ezaugarri orokorrak ikasteak, meta-eredu bidezko ikuspegia beharrezko orokortze-gaitasunaz hornitzen du, eskuragarri dagoen informazioa modu orekatuan ustiatzeko. Honen azalpen posible bat *Lehen* multzoko egiturek euren eboluzio-problemarekiko gehiegizko egokitzapena garatzen dutela da. Meta-ereduak ezin ditu *Lehen* multzoko SAS-ak berdin-berdin erreproduzitu, baizik eta soilik haien barneko menpekotasunak atzema-tera bideratu. Horrek, ondoren sortutako SAS-ek orokortze-gaitasun hobea izatea ahalbidetzen du.

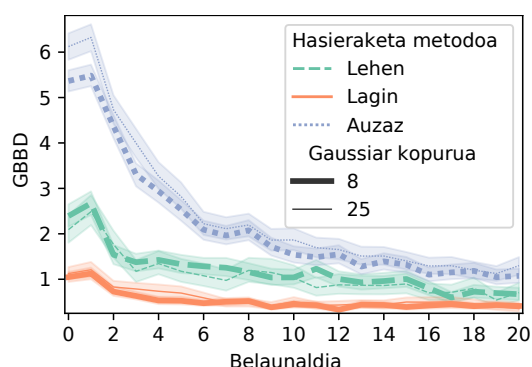
4 Ondorioak

Lan honetan, neurona sareen arkitektura bilaketa batean egindako ahalegin konputazionala ahalik eta gehien aprobetxatzeko metodologia bat proposatu da. Metodologia hori SAS-ak sortzeko neuroeboluzio-algoritmo batean probatu dugu. Proposamena optimizazio-prozeduretan aurkitutako arkitektura onenak modelatzean datza, sare Bayesiarretan oinarritutako meta-eredu baten bidez.

Lehenik eta behin, aurreko eboluzio-prozesuetan aurkitutako indibiduo onenekin entrenatuz, meta-ereduak sareak identifikatzeko gaitasuna lortzen du. Horrela, etorkizuneko neurona sareen bilaketa-prozedurek ahalegin konputazionala arindu dezakete, errendimendu gutxiko indibiduoak baztertuz. Bigarrenik, meta-ereduak optimizaziorik egin gabe ere errendimendu altua eskaintzeko probabilitate handia duten SAS-ak sortzeko gaitasuna du.

Ondoren, PM hurbilketarako eboluzionatutako GAP-SAS-ekin ikasitako meta-eredua populazio bat hasieratzeko erabili da. Populazio hori bigarren mailako neuroeboluzio-prozesu bat burutzeko erabili da, 2-D Gaussiar nahasketa hurbilketarako SAS-ak bilatzeko helburuarekin. Azkenik, ereduak neurona sareen arkitektura optimizatzeko algoritmo bat bizkortzeko duen gaitasuna ere frogatu dugu.

4. Irudia: SAS-ek belaunaldi desberdinetan (x ardatzean) sortutako laginei dagokien GBBD baliorik onena (y ardatzean) eboluzio-prozesuaren zehar. Lerro opakuek belaunaldi bakoitzean kalkulaturako 20 ind. $\times$ bel. $\times$ 30 = 60 GBBDren batezbertzekoa adierazten dute, eta banda erdi-gardenek exekuzio guztien %95eko konfiantza-tartea erakusten dute.



5 Etorkizunerako planteatzen den norabidea

Lan honetan aurkeztutako emaitzek meta-eredu horiek duten ahalmen handia iradokitzen dute. Izan ere, datu aniztasun nahiko xume batetik ikasitako bertsio batek emaitza interesgarriak sortzeko gai izan da. Mundu mailan egiten ari diren neuroeboluzio-prozeduren kopurua kontuan hartuta, artikulua honetan proposatutako ereduaren bertsio konplexuago batek, datu sofistikatu eta ugariagotik ikasita, hain maila altuak lor litzake non edozein sare neuronalaren bilaketari aplikatu ahal izango litzaikeen.

Erreferentziak

- Castillo, E., J. M. Gutierrez, eta A. S. Hadi. 1997. *Expert Systems and Probabilistic Network Models*. Springer.
- Chow, C. K., eta C.Ñ. Liu. 1968. Approximating discrete probability distributions with dependence trees. *IEEE Transactions on Information Theory* 14.462–467.
- Coello Coello, Carlos A., eta Margarita Reyes Sierra. 2004. A Study of the Parallelization of a Coevolutionary Multi-objective Evolutionary Algorithm. In *MICAI 2004: Advances in Artificial Intelligence*, ed. by Raúl Monroy, Gustavo Arroyo-Figueroa, Luis Enrique Sucar, eta Humberto Sossa, 688–697. Springer Berlin Heidelberg.
- Costa, Victor, Nuno Lourenço, eta Penousal Machado. 2019. Coevolution of Generative Adversarial Networks. In *International Conference on the Applications of Evolutionary Computation*, 473–487. Springer.
- Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems* 2.303–314.
- Dunn, Olive Jean. 1964. Multiple comparisons using rank sums. *Technometrics* 6.241–252.
- Elsken, Thomas, Jan-Hendrik Metzen, eta Frank Hutter. 2017. Simple and efficient architecture search for convolutional neural networks. *arXiv preprint arXiv:1711.04528*.

- Garciarena, Unai, Alexander Mendiburu, eta Roberto Santana. 2020. Analysis of the transferability and robustness of GANs evolved for Pareto set approximations. *Neural Networks* 132.281–296. Publisher: Elsevier.
- Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, eta Yoshua Bengio. 2014. Generative Adversarial Nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2672–2680, Cambridge, MA, USA. MIT Press.
- Gretton, Arthur, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, eta Alexander Smola. 2012. A Kernel Two-Sample Test. *Journal of Machine Learning Research* 13.723–773.
- Henrion, M. 1988. Propagating uncertainty in Bayesian networks by probabilistic logic sampling. In *Proceedings of the Second Annual Conference on Uncertainty in Artificial Intelligence*, ed. by John F. Lemmer eta Laveen N. Kanal, 149–164. Elsevier.
- Hochreiter, Sepp, eta Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9.1735–1780.
- Karras, Tero, Timo Aila, Samuli Laine, eta Jaakko Lehtinen. 2018. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In *International Conference on Learning Representations*, 1–26.
- Kazimipour, Borhan, Xiaodong Li, eta A Kai Qin. 2014. A review of population initialization techniques for evolutionary algorithms. In *2014 IEEE Congress on Evolutionary Computation (CEC)*, 2585–2592. IEEE.
- Krizhevsky, Alex, Ilya Sutskever, eta Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, 1097–1105.
- Kruskal, William H, eta W Allen Wallis. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47.583–621.
- Larrañaga, Pedro, eta Jose A Lozano. 2001. *Estimation of distribution algorithms: A new tool for evolutionary computation*, volume 2. Springer Science & Business Media.
- Lauritzen, S. L. 1996. *Graphical Models*. Oxford Clarendon Press.
- Li, H., eta Q. Zhang. 2008. Multiobjective Optimization Problems with Complicated Pareto Sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation* 13.284–302.
- Margolin, Adam A, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Dalla Favera, eta Andrea Califano. 2006. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. In *BMC bioinformatics*, volume 7, p. S7. Springer.
- Metz, L, B Poole, D Pfau, eta J Sohl-Dickstein. 2016. Unrolled generative adversarial networks. *arXiv preprint arXiv:1611.02163*.
- Miikkulainen, Risto, Jason Liang, Elliot Meyerson, Aditya Rawal, Daniel Fink, Olivier Francon, Bala Raju, Hormoz Shahrzad, Arshak Navruzyan, Nigel Duffy, eta others. 2019. Evolving deep neural networks. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, 293–312. Elsevier.
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, California: Morgan Kaufmann.
- Scutari, Marco, eta Jean-Baptiste Denis. 2014. *Bayesian Networks: With Examples in R*. CRC press.
- Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, eta Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818–2826.
- White, Colin, Sam Nolen, eta Yash Savani. 2020. Local Search is State of the Art for NAS Benchmarks. *arXiv preprint arXiv:2005.02960*.
- Yoon, Jinsung, Daniel Jarrett, eta Mihaela van der Schaar. 2019. Time-series generative adversarial networks. In *Advances in Neural Information Processing Systems*, 5508–5518.

6 Eskerrak eta oharrak

Lan honek Eusko Jaurlaritzatik (IT1504-22, eta Elkartek programaren bitartez KK-2024/00030, SEGURH2), Espainiako Zientzia eta Berrikuntza Ministerioa (PID2023-152390NB-I00, PID2023-149195NB-I00) eta Portugalako CISUC - (UID/CEC/00326/2020) proiektuen bidez finantzatu da. Joana Garciarenari testua zuzentzea ere eskertzen zaio.