Ekaia, ale berezia 2025, 167-184
https://doi.org/10.1387/ekaia.26347

ekaia
ZIENTZIA eta TEKNOLOGIA ALDIZKARIA

ISSN 0214-9001 – eISSN 2444-3255

Ikaskuntza autogainbegiratuaren ahalmena azaleratzen

(Unleashing the power of self-supervised learning)

*Aitor Sánchez-Ferrera**[1], *Borja Calvo Molinos*[1], *Usue Mori Carrascal*[1], *Jose A. Lozano*[1,2]

[1] Konputazio Zientzia eta Adimen Artifiziala Saila (UPV/EHU)
[2] Basque Center for Applied Mathematics (BCAM)

LABURPENA: Ikaskuntza automatikoak nabarmen egin du aurrera azken urteotan, eta horren fruitu da garatutako algoritmo sorta izugarria, zeinek ahalmena ematen baitigute ataza mota ugari egiteko. Literaturaren arabera, algoritmo gehienak ikaskuntza gainbegiratuan oinarritzen dira. Hala ere, hainbat atazatan emaitza onak lortu arren, ikaskuntza-paradigma horren eragozpen nagusia etiketetiko mendekotasuna da; izan ere, etiketatze-prozesua oso garestia da. Gainera, ikaskuntza automatikoan erabiltzen diren ereduek joera handia dute alborapen okerrak bultzatzen dituzten bide laburrak ikasteko, eta, ondorioz, atazek porrot egiten dute. Arazo horiek saihesteko, ikaskuntza autogainbegiratuak arreta bereganatu du azkenaldian ikaskuntza-paradigma gisa. Lan honetan, ikaskuntza autogainbegiratuari dagokion literaturari buruzko sarrera bat eskaintzen dugu, eta ikaskuntza-paradigma horren barruan bereizten diren metodo motak biltzen eta azaltzen ditugu, zenbait datu motatan aplikatzeko oinarrizko prozedurak aztertuz.

HITZ GAKOAK: Adimen artifiziala, ikaskuntza automatikoa, errepresentazio-ikaskuntza, ikaskuntza auto-gainbegiratua.

ABSTRACT: *Machine learning has made significant progress in recent years, resulting in the development of an impressive array of algorithms that enable us to perform a wide variety of tasks. According to the literature, most algorithms rely on supervised learning. However, despite achieving good results in different tasks, the main drawback of this learning paradigm is its dependency on manually created human labels, as the labeling process is very costly. Moreover, machine learning models tend to learn shortcuts that promote incorrect biases, leading to failures in the tasks they aim to accomplish. To avoid these issues, self-supervised learning has recently gained attention as a learning paradigm. This work provides an introduction to the literature on self-supervised learning. It explains the different types of methods distinguished within this learning paradigm and examine the basic procedures for applying them to different types of data.*

KEYWORDS: Artificial intelligence, machine Learning, representation learning, self-supervised learning.

*** *Harremanetan jartzeko* / *Corresponding author:*** Aitor Sánchez-Ferrera, Konputazio Zientzia eta Adiment Artifiziala, Informatika Fakultatea (20018 Donostia, Gipuzkoa, Euskal Herria). – aitor.sanchezf@ehu.eus – https://orcid.org/

Nola aipatu* / *How to cite: Sánchez-Ferrera, Aitor; Calvo Molinos, Borja; Mori Carrascal, Usue; Lozano, Jose A. (2025). «Ikaskuntza autogainbegiratuaren ahalmena azaleratzen». *Ekaia*, ale berezia 2025, 167-184. (https://doi.org/10.1387/ekaia.26347).

Jasotze-data: 2024, maiatzak 22; Onartze-data: 2024, uztailak 12.

ISSN 0214-9001 - eISSN 2444-3255 / © UPV/EHU Press 2025

Lan hau *Creative Commons Aitortu-EzKomertziala-PartekatuBerdin 4.0 Nazioartekoa* lizentzia baten mende dago

1. **SARRERA**

Azken urteotan, adimen artifizialak aurrerapen izugarriak izan ditu arlo ugaritan. Gure egunerokotasunean, etengabe egiten dugu topo adimen artifiziala erabiltzen duten sistemekin: etxeko laguntzaile birtual gisa jarduten duten *chatbot*ak erabiltzean, musika- eta bideo-plataformek gure gustuen araberako gomendioak eskaintzean, edo web-orrialdeetan gure interesen araberako iragarkiak jasotzean. Adimen artifizialak erabiltzen dituen teknologia guztien artean, ikaskuntza automatikoa (*machine learning*, ingelesez) da nabarmenetarikoa [1]. Datuak oinarri dituelarik eta hainbat ikaskuntza-paradigmatan oinarrituz, ikaskuntza automatikoa indarrez sartu da hainbat eremutan; besteak beste, ordenagailu bidezko ikusmenarekin [2] eta hizkuntzaren prozesamenduarekin [3] erlazionatutako atazetan, edo industriaren [4] eta osasungintzaren [5] arloetan.

Esan bezala, ikaskuntza automatikoan, ereduek hainbat ataza egiten ikasten dute datuak oinarri gisa erabiliz. Ikaskuntza hori egiteko, hainbat ikaskuntza-paradigma bereizten dira, baina ikaskuntza gainbegiratua (*supervised learning*, ingelesez) da hedatuena. Ikaskuntza gainbegiratuan, ereduak ataza ebazteko entrenatzeko, aldez aurretik etiketatutako datu-base erraldoiak erabili ohi dira. Nahiz eta ikaskuntza gainbegiratuan oinarritutako ereduek emaitza onak eman hainbat esparrutan, paradigma bere botila-lepora heltzen ari da. Arrazoi nagusietako bat etiketen beharra da; normalean, etiketa horiek lortzea garestia izaten da, bai denborari dagokionez, eta, askotan, baita ekonomikoki ere. Bestalde, askotan, datuak etiketatzeko jakintza espezializatua behar da [6]. Are gehiago, batzuetan ezinezkoa izan daiteke datu etiketatuak lortzea (adibidez, istripuen etiketak nahi izanez gero, ezin dira istripuak eragin etiketa horiek lortzeko).

Horretaz gain, atazak ebazten era azkarrago eta errazagoan ikasteko ereduek «bide laburrak» erabili ohi dituzte. Bide labur horiek datuen eta ikasi nahi den atazaren arteko ageriko korrelazio «faltsuetan» oinarritzen dira; hau da, nahiz eta ataza ebazteko pistak eman, alborapen handia sortzen dute, eta ereduen suposizio okerrak bultzatzen dituzte [7]. Fenomeno hori hobeto ulertzeko, demagun katuen eta txakurren irudiak sailkatzen dituen eredu bat entrenatu nahi dugula. Normalean, katuak etxeetan egon ohi direnez, katuen irudi gehienen atzealdeek altzariak eta etxeko gelak erakutsiko dituzte. Aldiz, txakurrak askotan paseatzen edo parkeetan egon ohi direnez, haien irudien atzeko planoek zuhaitzak, belarra eta kaleko beste elementu batzuk izango dituzte. Kasu horretan, txakurren eta katuen irudiak desberdintzeko entrenatutako eredu batek irudien atzealdea ikas dezake bide labur gisa, eraikin baten barruan egindako argazkiak katuei eta kalean egindako irudiak txakurrei esleituz. Ondorioz, ereduak sailkapen-ataza bide labur horren bitartez ikasten badu, porrot egingo du kalean dauden katuen eta etxean dauden txakurren irudiak sailkatzean.

Muga horiek gainditzeko estrategia bat errepresentazio-ikaskuntza (*representation learning*, ingelesez) da [8]. Ikaskuntza automatikoko eredu askok dimentsio txikiagoko espazio batean kodetzen dituzte datuak, eta, hala, atazen ikaskuntza eraginkorrari lagundu diezaioketen datuen errepresentazio berriak sortzen dituzte. Errepresentazio-ikaskuntza datuen errepresentazio «didaktikoak» ikastean datza, hau da, datuen informazio erabilgarria arakatzea eta egin nahi diren atazak ikastea errazten dutenak. Horrela, ereduek ataza bat ikasteko behar duten etiketa kopurua handia izatea eta bide laburrak ikastea saihestu daiteke. Ildo horretatik, ikaskuntza autogainbegiratuak (*self supervised learning*, ingelesez) arreta handia jaso du, datuetatik errepresentazio esanguratsuak ikasteko berezkoa duen eraginkortasunagatik eta haietatik abiatuta datuetatik orokortzeko duen gaitasunagatik [9].

Lan honetan, ikaskuntza autogainbegiratuari buruzko sarrera bat eta ulermena errazteko berrikuspen bat dakargu. Zehazki, honako ekarpen hauek egin nahi ditugu:

- Literaturaren berrikuspena: ikaskuntza autogainbegiratuan oinarritutako metodoen azterketa aurkeztuko dugu dokumentu honetan. Lan hau gaiaren sarrera bat izan daiteke ikerkuntza-lerro honetan murgildu nahi duten ikerlarientzat.
- Metodoen sailkapena: arloko literaturan aurkitu daitezkeen algoritmo autogainbegiratu motak bereiziko ditugu. Gainera, aztertuko dugu ea zer metodo diren ohikoenak datu mota bakoitzerako.
- Aplikazioen azterketa: ikaskuntza autogainbegiratua erabiltzen duten metodoen aplikazio errealak bilduko ditugu.
- Ikerkuntza-lerroen proposamena: aztertutako lanak oinarritzat hartuz, ikaskuntza autogain-begiratuaren arlorako ikerkuntza-lerro berriak proposatuko ditugu.

2. ZER DA IKASKUNTZA AUTOGAINBEGIRATUA?

Ikaskuntza automatikoan, hainbat ikaskuntza-paradigma erabiltzen dira ataza jakin bat egiteko gai diren ereduak eraikitzeko. Egin nahi dugun ataza hori, *zeregin-ataza* (*downstream task*, ingelesez), sailkapena, erregresioa edo anomaliak detektatzea izan daiteke, besteak beste, eta halakoak ebazteko erabiltzen den ikaskuntza-paradigmarik ohikoena gainbegiratua da. Atal honetan, ikaskuntza gainbegiratuaren nondik norakoak azalduko ditugu, eta, haren desabantailetan oinarrituta, alternatiba gisa, ikaskuntza autogainbegiratua aurkeztuko dugu.

2.1. **Ikaskuntza gainbegiratua**

Ikaskuntza gainbegiratuaren helburua sarrerako datu bakoitzari irteerabalio bat esleituko dion eredu bat eraikitzea da. Eredu hori eraikitzeko, ikaskuntza gainbegiratuan, entrenamendurako aldez aurretik etiketatutako datu-base bat erabiltzen da, hau da, datu-base bat zeinetan sarrera bakoitzaren irteera —hots, etiketa— ezaguna den.

Zehatzago esanda, izan bedi $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{N}$ entrenamendurako sarrera-datu multzo bat, non $\mathbf{x}_i$ multzoko *i*-garren datu-lagina eta *N* multzoko lagin kopurua diren. Ikaskuntza gainbegiratuan, zeregin-atazaren arabera, $\mathbf{Y}^{(z)} = \{y_i^{(z)}\}_{i=1}^{N}$ zeregin-atazari lotutako irteerako datu multzo bat izango dugu, non $y_i^{(z)}$ entrenamendu multzoko *i*-garren datu-laginaren zeregin-atazari lotutako etiketa den. Hobeto ulertzeko, demagun aurreko atalean deskribatutako zeregin-ataza egin nahi dugula, hau da, katuen eta txakurren irudien sailkapena. Kasu horretan, $\mathbf{x}_i$ datu-lagin bakoitza irudi bat izango da, eta $y_i^{(z)}$ etiketak «katua» klasearen balioa hartuko du $\mathbf{x}_i$ irudia katu batena bada, eta «txakurra» balioa, berriz, txakur baten irudia bada.

Zeregin-ataza egiten ikasteko, helburua sarrerako datuak eta etiketak mapatuko dituen $f^{(z)}(\cdot)$ eredu bat entrenatzea da, kasu idealean ereduaren zeregin-atazerako iragarpenak $\hat{y}_i^{(z)} = f^{(z)}(\mathbf{x}_i) = y_i^{(z)}$ izanik. Hemendik aurrera, ereduak neurona-sareak direla asumituko dugu, sarrera-datuen errepresentazio berri bat sortuko duen $\phi(\cdot)$ kodetzaile batez eta errepresentazio horien barnean zeregin-atazerako iragarpenak itzuliko dituen $g^{(z)}(\cdot)$ zeregin-modulu batez osatuak.[1] Zehazki, $\phi(\cdot)$ kodetzaileak sarrerako $\mathbf{x}_i$ datu-lagin bakoitzerako $\mathbf{z}_i = \phi(\mathbf{x}_i)$ errepresentazio berri bat sortuko du dimentsio txikiagoko espazio batean. Ondoren, $g^{(z)}(\cdot)$ zeregin-moduluak $\mathbf{z}_i$ errepresentazio berria hartuko du sarrera gisa zeregin-ataza egiteko, hots, $\hat{y}_i^{(z)} = g^{(z)}(\mathbf{z}_i)$. Aurreko adibidearen oinarrian, neurona-sare baten $\hat{y}_i^{(z)} = f^{(z)}(\mathbf{x}_i) = g^{(z)}(\phi(\mathbf{x}_i))$ iragarpenak $\mathbf{x}_i$ datu-lagineko txakur baten edo katu baten irudia izateko probabilitatea adieraziko du.

Ikaskuntza gainbegiratuan, $f^{(z)}(\cdot)$ eredua zeregin-ataza zuzenean ikasteko entrenatzen da, sarrerako $\mathbf{x}_i$ datu-lagin bakoitzari lotutako ereduaren $\hat{y}_i^{(z)}$ iragarpenen eta $y_i^{(z)}$ etiketen arteko berdintasuna helburu ideala izanik. Hala ere, aurreko atalean komentatu dugun bezala, ikaskuntza gainbegiratuak hainbat muga dauzka; besteak beste, entrenamendu multzoko etiketak lortzeak kostu handia du, nahiz errepresentazio-ikaskuntzaren bitartez

[1] Arloko literaturan gehien erabiltzen diren ereduak sakonak dira, berezkoak dituzten propietateak egokienak direlako. Hori dela eta, paradigma ondo ulertzeko, lagungarriak izango dira sare neuronalen arloko terminologia eta kontzeptuak.

arindu daitezkeen. Esplizituki, errepresentazio-ikaskuntzaren helburua zera da, $f^{(z)}(\cdot)$ eredu baten $\phi(\cdot)$ kodetzaileak sortzen dituen errepresentazioek sarrera-datuen berezko ezaugarriak ahalik eta hobekien atzematea. Helburu hori lortzeko modu bat ikaskuntza autogainbegiratua da.

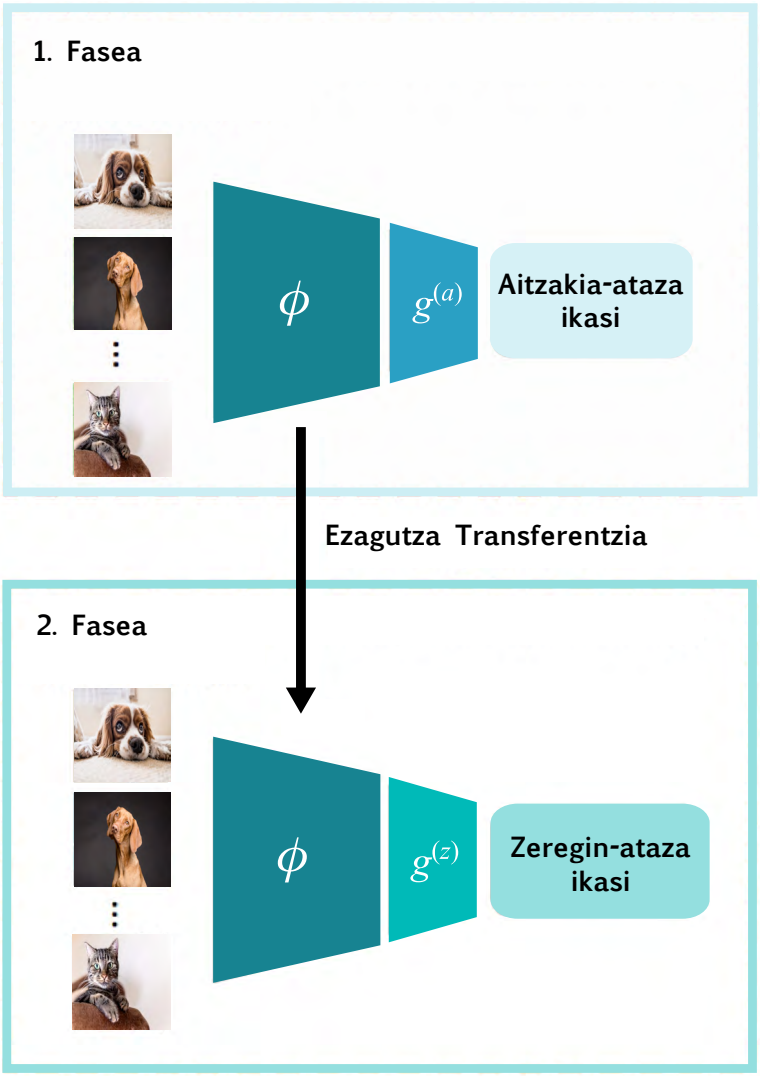
2.2. **Ikaskuntza autogainbegiratua**

Ikaskuntza autogainbegiratuak hainbat definizio jaso ditu denboran zehar. Gaur egun, ikaskuntza autogainbegiratua era honetan definitzen da: «Datuen errepresentazio esanguratsuak ikastea datuen zenbait ezaugarri, erlazio edo aldaketa iragarriz, etiketetan oinarritu gabe» [9]. Zehazki, ikaskuntza autogainbegiratuan, zeregin-atazarekin informazioa partekatzen duten ataza osagarriak ikasiz, ereduek zeregin-ataza ebazteko lagungarriak izan daitezkeen errepresentazioak atzemango dituztela asumitzen da, zeregin-atazaren ikaskuntza errazagoa, azkarragoa eta eraginkorragoa izan dadin. Literaturan, datuen errepresentazio esanguratsuak atzemateko lagungarriak izan daitezkeen ataza osagarriei *aitzakia-atazak* (ingelesez *pretext task*) deitzen zaie [10].

Izan bedi aurreko atalean adierazitako testuinguru bera: $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{N}$ entrenamendurako sarrera-datu multzo bat, egin nahi dugun zeregin-atazaren araberako $\mathbf{Y}^{(z)} = \{y_i^{(z)}\}_{i=1}^{N}$ etiketa multzo bat, eta zeregin-ataza egiteko erabiliko dugun $f^{(z)}(\cdot)$ eredu bat, $\phi(\cdot)$ kodetzaile batez eta $g^{(z)}(\cdot)$ zeregin-modulu batez osatua. Azalpenak errazteko, aurreko adibidearekin jarraituko dugu, non gure zeregin-ataza katuen eta txakurren irudien sailkapena egitea den. Aitzakia-ataza osagarriak ikastearen helburua $\phi(\cdot)$ kodetzaileak sortutako datuen errepresentazioak ahalik eta esanguratsuenak eta deskriptiboenak izatea da. Horretarako, jatorrizko $\mathbf{Y}^{(z)}$ etiketetan oinarritzen ez den aitzakia-ataza bat defini daiteke, datuen berezko zenbait ezaugarri, erlazio edo eraldaketa iragartzean oinarrituta [9].

Literaturan, irudien errepresentazio esanguratsuak ikasteko sarritan erabiltzen den aitzakia-ataza mota bat autokodetzaileak dira, non sarrerako datu-lagin bakoitza dimentsio txikiagoko espazio batean kodetzen den eta, ondoren, jatorrizko sarrerako datu-lagina berreraikitzen den [11]. Zehazki, gure adibidean, aitzakia-atazaren entrenamendurako sarrera datu multzoa, $\mathbf{X}$, zeregin-atazaren berdina izango da: katuen eta txakurren irudi sorta. Autokodetzaileen kasuan, aitzakia-ataza jatorrizko datuak berreraikitzean oinarritzen denez, aitzakia-atazaren irteera-datu multzoa $\mathbf{Y}^{(a)} = \mathbf{X}$ izango da. Aitzakia-ataza ebazteko, $f^{(a)}(\cdot)$ ereduari $\phi(\cdot)$ kodetzaileak sortutako errepresentazioak sarrera gisa erabiliko dituen $g^{(a)}(\cdot)$ aitzakia-modulu bat erantsiko diogu; kasu honetan, deskodetzaile bat izango da. Aitzakia-modulua, $\phi(\cdot)$ kodetzaileak sortutako $\mathbf{x}_i$ sarrerako datu-lagin bakoitzaren errepresentaziotik abiatuta, $\mathbf{z}_i = \phi(\mathbf{x}_i)$, $y_i^{(a)} = \mathbf{x}_i$ berreraikitzen saiatuko da,

kasu idealean $\hat{y}_i^{(a)} = f^{(a)}(\mathbf{x}_i) = g^{(a)}(\phi(\mathbf{x}_i)) = \mathbf{x}_i$ izanik. Aitzakia-ataza hau ikasiz, $\phi(\cdot)$ kodetzaileak sarrerako argazkien testuingurua eta edukia hobeto deskribatzen dituzten errepresentazio baliagarriak eraikiko ditu, eta horrek zeregin-atazaren ikaskuntza erraztuko du [12].



1. irudia. Aitzakia-atazaren aurre-entrenamendua. Lehenengo fasean, aitzakia-ataza egiten ikasteko entrenatzen da eredua, kodetzaile bat eta aitzakia-modulu bat erabiliz. Behin aitzakia-ataza ikasita, bigarren fasean, aurre-entrenatutako kodetzaileak sortutako errepresentazioak erabiltzen dira zeregin-modulu baten sarrera gisa, zeregin-ataza egiten ikasteko.

Hori guztia kontuan hartuz, metodologia honen gakoa ondoriozta daiteke: zeregin-ataza ebazteko baliagarriak diren errepresentazioen ikaskuntza bultzatzen duten aitzakia-ataza aproposak diseinatzea [13]. Hala ere, nahiz eta aitzakia-atazaren helburua ulertu, beharrezkoa da azaltzea nola aprobetxatzen duen ereduak aitzakia-ataza ikastearen bitartez lortutako jakintza zeregin-ataza ikasteko. Jakintza-transferentzia gauzatzeko, bi prozedura nagusi desberdintzen dira literaturan: i) ereduen aurre-entrenamendua aitzakia-atazan eta ii) aitzakia-atazaren eta zeregin-atazaren ikaskuntza bateratua.

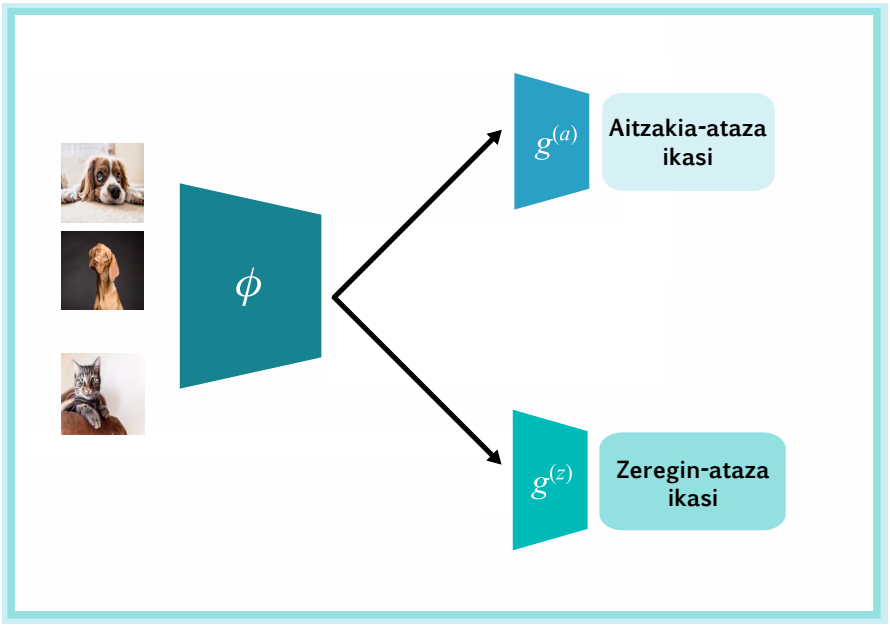
Lehenengo metodoan, aitzakia-ataza ikasteko aurre-entrenatzen da eredua, eta zeregin-ataza ebazteko erabiltzen dira lortutako errepresentazioak, sarrera gisa. Zehatz-mehatz, bi azpiprozesu bereiz daitezke (ikusi 1. irudia):

1. **Aitzakia-atazaren aurre-entrenamendua**. Fase honetan, $f^{(a)}(\cdot)$ eredua erabiliko da aitzakia-ataza ikasteko. Sarrerako $\mathbf{X}$ datu multzoa aitzakia-atazaren irteerako $\mathbf{Y}^{(a)}$ datu multzoarekin mapatzen ikasiko du $f^{(a)}(\cdot)$ ereduak, $\phi(\cdot)$ kodetzailea eta $g^{(a)}(\cdot)$ aitzakia-modulua erabiliz, era idealean $\hat{y}_i^{(a)} = f^{(a)}(\mathbf{x}_i) = g^{(a)}(\phi(\mathbf{x}_i)) = \mathbf{x}_i$ izateko. Momentu horretan ez da $f^{(z)}(\cdot)$ ereduaren $g^{(z)}(\cdot)$ zeregin-modulua erabiliko.
2. **Zeregin-atazaren afinazioa**. Behin aitzakia-atazaren aurre-entrenamendua bukatuta, $\phi(\cdot)$ kodetzaileak datuen errepresentazio esanguratsuak atzematen dituela asumitzen da. Bere- ganatutako jakintza hori zeregin-ataza ikasteko aprobetxatzeko, $g^{(a)}(\cdot)$ aitzakia-modulua deuseztatu eta $g^{(z)}(\cdot)$ zeregin-moduluarekin ordezkatuko da. Horrela, fase honetan, sarrerako $\mathbf{X}$ datu multzoa zeregin-atazaren irteerako $\mathbf{Y}^{(z)}$ etiketekin mapatzen ikasiko du ereduak (aurreko pausoan ikasitako) $\phi(\cdot)$ kodetzailea eta $g^{(z)}(\cdot)$ erabiliz, bere zeregin-atazarako iragarpenak $\hat{y}_i^{(z)} = f^{(z)}(\mathbf{x}_i) = g^{(z)}(\phi(\mathbf{x}_i))$ izanik. Bigarren fase honetan, lehenengo fasean $\phi(\cdot)$ kodetzaileak atzemandako errepresentazio esanguratsuak aprobetxatzen dira errepresentazio horiek $g^{(z)}(\cdot)$ zeregin-moduluaren sarrera gisa erabiliz.

Bigarren metodoan, aldiz, ereduak aldi berean ikasten ditu aitzakia-ataza eta zeregin-ataza (ikusi 2. irudia). Horretarako, $\phi(\cdot)$ kodetzaileak sarrerako $\mathbf{x}_i$ datu-lagin bakoitzerako $\mathbf{z}_i = \phi(\mathbf{x}_i)$ errepresentazio berri bat sortuko du. Ondoren, ataza-modulu bakoitzak $\mathbf{z}_i$ errepresentazio berri hori ataza bakoitzeko irteera-datuekin mapatuko du, $\hat{y}_i^{(a)} = g^{(a)}(\phi(\mathbf{x}_i))$ aitzakia-atazarako eta $\hat{y}_i^{(z)} = g^{(z)}(\phi(\mathbf{x}_i))$ zeregin-atazarako iragarpenak konputatuz.

Helburua bera da bi kasuetan: aitzakia-atazaren bitartez lortutako ezagutza aprobetxatzea zeregin-ataza bat (edo gehiago) ikasteko. Hala ere, aitzakia-atazaren ikaskuntza eta lortutako ezagutzaren transferentzia era

desberdinean egiten dira. Lehenengo metodoaren helburua datuen errepresentazio baliagarriak lortzen dituzten aitzakia-atazen diseinuan datza. Baldintza idealetan, errepresentazio horiek behar adina orokorrak izango dira zeregin-ataza desberdinetan baliagarriak izateko. Bigarren metodoan, berriz, berariazko zeregin-ataza bat ikasteko onuragarriak izan daitezkeen aitzakia-atazak diseinatzean datza helburua. Azken kasu horretan, aitzakia-ataza eta zeregin-ataza aldi berean entrenatuz, ikasitako errepresentazioek ez dute zertan hain orokorrak izan, beste zeregin-ataza batzuetan baliagarriak izateko. Azpimarratzekoa da aurreko bi metodoetan eredu bat entrenatzeko aitzakia-ataza bat baino gehiago konbina daitekeela, eskuratutako errepresentazioen kalitatea eta egokitasuna hobetzeko eta, horrela, ereduaren errendimendu orokorra zeregin-atazetan nabarmen hobetzeko [14].



2. irudia. Aitzakia-atazaren eta zeregin-atazaren ikaskuntza bateratua. Kodetzaileak sortutako errepresentazioak aitzakia-modulu baten eta zeregin-modulu baten sarrera gisa erabiltzen dira, aitzakia-ataza eta zeregin-ataza aldi berean ikasiz.

Bi kasuetan, aitzakia-ataza aurrez entrenatuta edo aldi berean, ikaskuntza autogainbegiratuari buruzko lanak aitzakia-atazen arabera sailka daitezke. Literaturan, bi aitzakia-ataza mota nagusi bereizten dira: i) ikaskuntza autoprediktiboan oinarritutakoak, eta ii) ikaskuntza kontrastatzailea erabiltzen dutenak [9]. Hurrengo bi ataletan, ataza mota bakoitzaren ezaugarriak eta adibide esanguratsuenak deskribatuko ditugu.

3. IKASKUNTZA AUTOPREDIKTIBOA

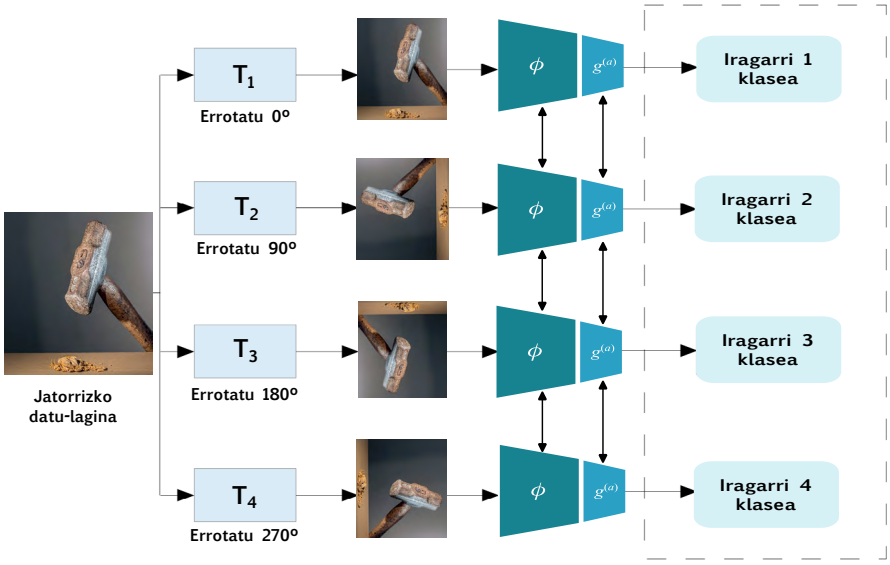
Ikaskuntza autoprediktiboan (ingelesez, *self-predictive learning*) oinarritutako aitzakia-atazetan, datuen zati bat erabili ohi da beste zati batzuk iragartzeko. Horretarako, ohikoa da transformazioak aplikatzea datuei, ondoren haiekin aitzakia-atazak ikasteko. Ikaskuntza autoprediktiboan oinarritutako aitzakia-atazak ikasiz eta ondorioz atzemandako errepresentazioez baliatuz, zeregin-ataza ikasteko erabili beharreko etiketa kopurua murriztu daiteke.

Orokorrean, aitzakia-ataza autoprediktiboak hiru multzo nagusitan bereiz daitezke: *sailkapen autogainbegiratua* (*self-supervised classification*), *berreraikitze autogainbegiratua* (*self-supervised reconstruction*) eta *iragarpen autogainbegiratua* (*self-supervised forecasting*) [9].

3.1. Sailkapen autogainbegiratua

Kategoria honetan sartzen diren aitzakia-atazetan, $\mathcal{T} = \{T_1, ..., T_K\}$ transformazio familia bat aplikatzen zaio entrenamendu multzoko $\mathbf{x}_i$ lagin bakoitzari. Horren emaitza gisa, lagin originalaren $\mathcal{T}(x_i) = \{T_1(\mathbf{x}_i), ..., T_K(\mathbf{x}_i)\}$ ikuspegi multzo bat sortzen da, non $T_k(\mathbf{x}_i)$ ikuspegiari atxikitako sasietiketa k den. Sailkapen autogainbegiratuaren zeregina $f^{(a)}(\cdot)$ eredu bat entrenatzea da, $\mathbf{x}_i$-ren ikuspegi bakoitza gauzatzeko honi aplikatutako transformazioa iragarriko duena, egoera idealean $f^{(a)}(T_k(\mathbf{x}_i)) = k$ izanik ($1 \leq k \leq K$). Teknika honetan, proposatutako transformazioak aplikatzean, datuek izango duten portaera iragartzen da. Horrela, ereduek datuen berezko ezaugarriei buruzko informazio baliagarria ikasiko dutela asumitzen da. Multzo honetan aurki ditzakegun aitzakia-atazen arrakasta problema bakoitzerako onuragarriak izan daitezkeen transformazio egokiak aukeratzearen menpekoa da.

Teknika hau hobeto ulertzeko, [15] lanean irudien errepresentazio esanguratsuak ikasteko proposatutako aitzakia-ataza azalduko dugu (ikusi 3. irudia). Zeregin-ataza zenbait tresnaren irudien sailkapena izanik, lau errotazio-transformaziotan oinarritutako sailkapen autogainbegiratua proposatzen da ereduaren aurre-entrenamendua egiteko. Sarrerako datu-lagin bakoitzari lau transformazioak aplikatuz, datu-lagin bakoitzaren lau ikuspegi berri sortzen dira. Ondoren, kodetzaile batez eta sailkatzaile-lana egiten duen aitzakia-modulu bat erabiliz, ikuspegi bakoitza sortzeko jatorrizko datu-laginari aplikatutako transformazioaren etiketa iragartzen da. Kasu horretan, egileek asumitzen dute errotazio-transformazioak bereizteko ereduak tresna bakoitzaren berezko ezaugarriak ikasten dituela (tresnaren ohiko posizioa, hain zuzen). Behin eredua aurre-entrenatuta, aitzakia-modulua deuseztatu eta sailkapenerako zeregin-modulu batez ordezkatzen da. Zeregin-moduluak aitzakia-atazan aurretik entrenatutako kodetzailearen errepresentazioak erabiliko ditu zeregin-ataza egiten ikasteko; tresnen irudien sailkapena, hain zuzen.



3. irudia. Sailkapen autogainbegiratua. Jatorrizko datu-laginari transformazio sorta bat aplikatzen zaio, eta transformazio bakoitzeko laginaren ikuspegi berri bat sortzen da. Ondoren, transformazioen bitartez sortutako ikuspegi bakoitza sarrera gisa erabiliz, ikuspegi bakoitzari lotutako transformazioaren etiketa iragartzen da.

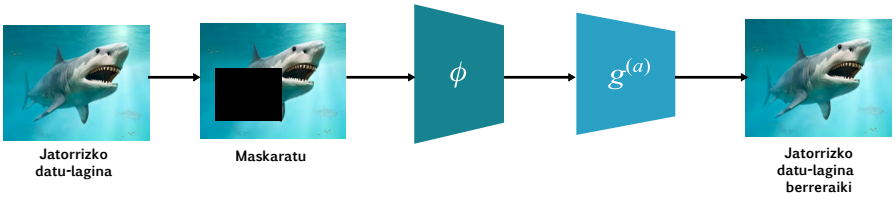
Kontuan hartzekoa da hautatutako transformazioak eskura daukagun datu motaren menpekoak direla, hau da, beharbada datu mota batean aplika daitezkeen transformazioak ez dira bideragarriak beste datu mota batzuetan aplikatzeko. Irudien kasuan, transformazio geometrikoak dira ohikoenak [16, 15]. Era antzekoan, hainbat lanek transformazio afinen erabilera proposatzen dute denbora-segiden sailkapen autogainbegiratua gauzatzeko [17, 18]. Hizkuntzaren prozesamenduaren arloan, aldiz, testu bateko esaldien ordena iragartzea proposatzen da sailkapen autogainbegiratua egiteko [19]. Azken kasu horretan, nabarmentzekoa da ez dela transformaziorik erabiltzen aitzakia-ataza eraikitzeko, testuaren koherentzia erabiltzen baita seinale autogainbegiratu gisa. Modu horretan, sailkapen autogainbegiraturako sasietiketa esaldien ordenak definitzen du.

3.2. **Berreraikitze autogainbegiratua**

Kategoria honetako aitzakia-atazek T transformazio bat aplikatzen diote $\mathbf{x}_i$ sarrera-lagin bakoitzari, eta lagin horren $T(\mathbf{x}_i)$ ikuspegi berri bat sortzen da. Berreraikitze autogainbegiratuaren esparruan, $f^{(a)}(\cdot)$ eredu bat entrenatzen da, $T(\mathbf{x}_i)$ sarrera gisa hartuta $f^{(a)}(T(\mathbf{x}_i)) = \mathbf{x}_i$ jatorrizko lagina berreraikitzen saiatuko dena. Horrela, ereduaren irteeraren eta jatorrizko laginaren arteko errorea minimizatzea da entrenamenduaren helburua.

Multzo horretan sartzen diren aitzakia-atazen arteko desberdintasun nagusia proposatutako transformazioan datza. Aukerarik hedatuena aurreko atalean azaldutako autokodetzaileen erabilera da [20]. Horretaz aparte, badaude beste metodo batzuk autokodetzaileak beste transformazio batzuekin konbinatzen dituztenak. Halakoetan, datu-lagin bakoitzari hautatutako transformazioa aplikatzen zaio, eta, ondoren, autokodetzailearen bitartez jatorrizko lagina berreraikitzen da.

Adibide gisa, [21] lanean proposatutako aitzakia-atazan, irudien errepresentazioak ikasteko sarrera-datuen zati bat maskaratzea (zati bat ezabatzea) proposatzen da transformazio gisa (ikusi 4. irudia). Aitzakia-ataza zati bat maskaratua duten datu-laginak berreraikitzea da. Horretarako, kodetzaile batez eta aitzakia-modulua den deskodetzaile batez osatuta egongo da $f^{(a)}$ eredua; sarreran maskaratutako lagina hartuko du, eta jatorrizko lagina $\mathbf{x}_i$ itzuliko du egoera idealean. Egileek asumitzen dute ezen, maskaratutako zatiak iragartzeko, ereduak irudi osoen testuingurua ulertu behar duela, eta, hala, kodetzailearen errepresentazio esanguratsuak atzemateko ahalmena bultzatzen da. Ondoren, aitzakia-modulua sailkapen zeregin-modulu batez ordezkatzen da, eta irudien sailkapenean oinarritutako zeregin-ataza bat ikasten da aurre-entrenatutako kodetzailearen errepresentazioak oinarritzat hartuz.



4. irudia. Berreraikitze autogainbegiratua. Jatorrizko datu-laginari transformazio bat aplikatzen zaio, eta haren ikuspegi berri bat sortzen da. Ondoren, eredu bat entrenatzen da ikuspegi berritik jatorrizko datu- lagina berreraikitzeko.

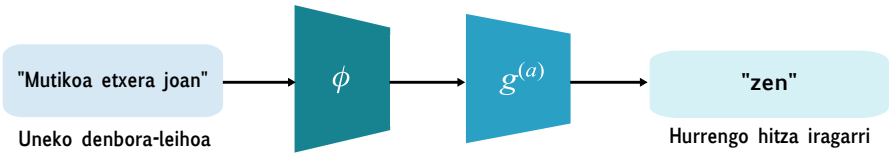
Literaturan, badira maskaratze bidezko berreraikitze autogainbegiratua askotariko datu motetan aplikatzen duten lanak, hala nola denborazko serieetan [22, 23] eta hizkuntzaren prozesamenduan [24]. Bestalde, badaude beste lan batzuk zarata gausstarraren injekzioa erabiltzen dutenak aitzakia-atazetarako transformazio gisa, askotariko datu moten errepresentazio-ikaskuntzarako [25].

3.3. **Iragarpen autogainbegiratua**

Iragarpen autogainbegiratuan, datu mota batzuek berezkoak dituzten denbora-dependentziak erabiltzen dira aitzakia-atazak eraikitzeko. Izan

bedi denboraren menpekoa den $\mathbf{x} \in \mathbb{R}^{D \times T}$ lagin bat, non D laginaren dimentsio kopurua eta T denboran zehar egindako laginaren neurketa kopuru osoa diren. Hala, $\mathbf{x}_t \in \mathbb{R}^D$-k laginak t unean duen balioa adierazten du. Multzo horretan biltzen diren aitzakia-atazetan, jatorrizko laginaren $\mathbf{x}_{t-\omega,t} = (\mathbf{x}_{t-\omega}, \mathbf{x}_{t-\omega+1}, ..., \mathbf{x}_t)$ denbora-leiho bat oinarritzat hartuz, lagin horren hurrengo p puntuz osatutako denbora-leihoa $\mathbf{x}_{t+1,t+p} = (\mathbf{x}_{t+1}, ..., \mathbf{x}_{t+p})$ iragartzen da.

Hizkuntzaren prozesamendurako lan batean proposatutako aurre-entrenamendurako aitzakia-ataza daukagu (ikusi 5. irudia) horren adibide gisa [26]. Lan horretan, hitz-sekuentzia bat erabiltzen da ereduaren sarrera gisa (denbora-leihoa), kodetzailearen eta aitzakia-moduluaren bitartez sekuentziaren hurrengo hitza iragartzeko. Haren bitartez, kodetzaileak sortzen dituen errepresentazioek kontzeptuen testuingurua eta erlazioa atzemango dituzte. Behin aurre-entrenatuta, aitzakia-modulua sailkapenerako zeregin-modulu batez ordezkatzen da eta sentimendu-analisian oinarritutako zeregin-ataza bat ikasteko erabiltzen dira ikasitako errepresentazioak, hau da, testu baten emozioak eta iritziak automatikoki identifikatzeko eta sailkatzeko.



5. irudia. Iragarpen autogainbegiratua. Uneko denbora-leiho bat sarrera gisa erabiliz, ereduak laginaren hurrengo balioak iragartzen ikasi behar du.

Multzo horretako aitzakia-atazak bideoen errepresentazio-ikaskuntzan [27] eta denborazko serieen arloan [28] ere aplikatu izan dira.

4. IKASKUNTZA KONTRASTATZAILEA

Ikaskuntza kontrastatzailea (*contrastive learning*, ingelesez) datu-laginen arteko antzekotasunak nabarmentzean oinarritzen da. Testuinguru honetan erabiltzen diren ereduek sarrerako datuak dimentsio txikiagoko espazio berri batean kodetzen dituzte kodetzaile baten bitartez. Ondoren, espazio horretan ikasitako errepresentazioak datuen arteko antzekotasunaren araberakoak izatera bultzatzen dira. Zehazki, bi datu-lagin zenbait irizpideren arabera antzekoak badira (literaturan pare positiboa osatzea deritzona), eredua haien errepresentazioak elkarrengandik hurbil mantentzen saiatuko da. Aldiz, laginak antzekoak ez badira (hau da, pare negatiboa

osatzen badute), haien errepresentazioak ahal den heinean aldenduko ditu ereduak. Horrela, ereduak datuen arteko antzekotasun esanguratsuak atzemango dituela asumitzen da. Ikusi da ezagutza horrek balio handia duela, ereduen errendimendua hobetu baitezake zeregin-ataza anitzetan, eta bide laburrak saihestu [29].

Kategoria honetan sartzen diren aitzakia-atazen arteko bereizketa nagusia pare positiboak eta negatiboak osatzeko hartutako irizpideetan oinarritzen da. Bereizketa horren arabera, nahiz eta literaturan aitzakia-ataza mota ugari bereizten diren, bi dira oinarrizkoenak: *transformazio-kontrastea* (*transformation contrast*) eta *testuinguru-kontrastea* (*sampling contrast*).

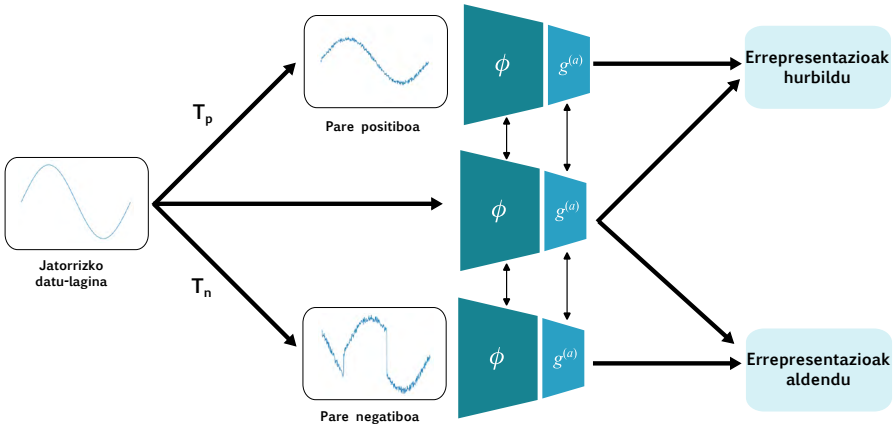
4.1. **Transformazio-kontrastea**

Transformazio-kontrastea pare kontrastatzaileak eratzeko teknika ohikoena da, non sarrerako datuei transformazioak aplikatzen zaizkien pare positiboak eta negatiboak sortzeko. Zehazki, sarrerako $\mathbf{x}_i$ lagin bakoitzari, aingura deritzonari, T_p transformazio positibo bat aplikatzen zaio, eta $\mathbf{x}_i^{(p)}$ pare positiboa osatzen da. Hori oinarritzat hartuz, $\mathbf{x}_i$ ainguraren eta $\mathbf{x}_i^{(p)}$ pare positiboaren errepresentazioak hurbiltzen saiatuko da eredua; ainguraren eta $\mathbf{x}_i^{(n)}$ pare negatiboen errepresentazioak, aldiz, aldendu egingo ditu. $\mathbf{x}_i^{(n)}$ pare negatiboa osatzeko aukera ohiko bat $\mathbf{x}_i$ aingurari T_n transformazio negatibo bat aplikatzea da.

Pareak eratzeko transformazioei dagokienez, T_p transformazio positiboaren helburua ereduak transformazio horrekiko inbariantza lortzea da. T_n transformazio negatiboak proposatzean, aldiz, helburua datuen berezko ezaugarriak aldatzea da, eta, hala, ainguraren aldean desberdina den ikuspegi berri bat sortzen da. Behin pareak eratuta, $f^a(\cdot)$ ereduak $\phi(\cdot)$ kodetzaileak sortutako ainguraren $\mathbf{z}_i = \phi(\mathbf{x}_i)$ eta pare positiboaren $\mathbf{z}_i^{(p)} = \phi(\mathbf{x}i^{(p)})$ errepresentazioak hurbilduko ditu, eta aipatutako T_p transformazioarekiko inbariantza bultzatuko du. Aldi berean, ainguraren $\mathbf{z}_i = \phi(\mathbf{x}_i)$ eta pare negatiboaren $\mathbf{z}_i^{(n)} = \phi(\mathbf{x}_i^{(n)})$ errepresentazioak aldenduko ditu [32].

Multzo honetako lanen desberdintasun nagusia datuen testuinguruaren araberako transformazio positiboen eta negatiboen aukeraketan dago. Adibidez, [33] lanean, denborazko serieen errepresentazioen ikaskuntzarako transformazioetan oinarritutako ikaskuntza kontrastatzailea erabiltzen dute (ikusi 6. irudia). Zarata gausstarraren injekzioa eta balioen eskalatzea konbinatzen dituzte transformazio positibo gisa, perturbazio horiekiko inbariantza bultzatu nahian. Aldi berean, zarata gausstarraren injekzioa eta denbora-segida osatzen duten azpisekuentzien permutazioa (azpisekuentzian ordena aldatzea) konbinatzen dituzte ainguraren pare negatiboak sortzeko. Behin eredua aitzakia-ataza kontrastatzailean entrenatuta, zeregin-modulu bat eransten diote aurre-entrenatutako kodetzaileari, eta sailkapenean oina-

rritutako zeregin-ataza ikasten dute ikasitako errepresentazioak moduluaren sarrera gisa erabiliz.



6. irudia. Transformazio-kontrastea. Transformazio-kontrastea. Jatorrizko datulaginari transformazio positibo bat eta negatibo bat aplikatuz, pare positibo bat eta negatibo bat sortzen dira. Horrela, helburua ainguraren eta pare positiboaren errepresentazioak hurbiltzea izango da; ainguraren eta pare negatiboaren errepresentazioak, aldiz, aldendu egingo dira.

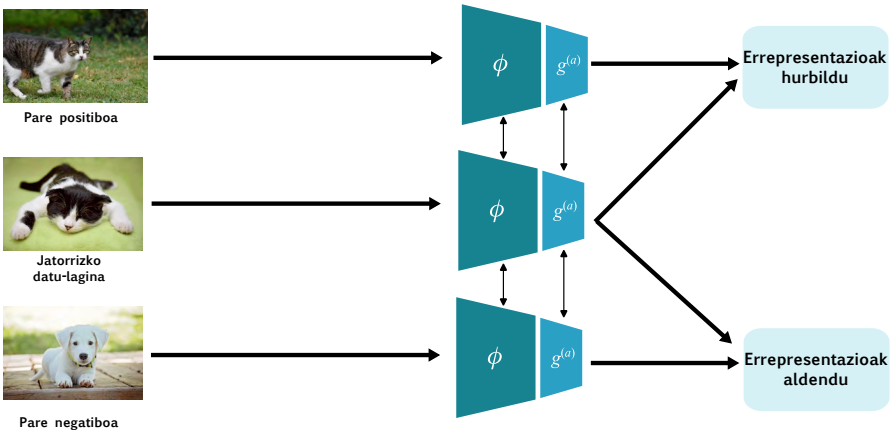
Transformazioetan oinarritutako ikaskuntza kontrastatzailea lan askotan erabiltzen da, hainbat datu motari lotutako errepresentazio esanguratsuak ikasteko, eta testuinguru bakoitzaren araberako transformazio positibo eta negatibo egokiak proposatu dira [29, 30].

4.2. **Testuinguru-kontrastea**

Testuinguru-kontrastean, transformazioak erabili ordez, datuen berezko ezaugarriei buruzko intuizioa erabiltzen da pare kontrastatzaileak eratzeko. Kasu honetan, $\mathbf{x}_i$ ainguraren pare positiboa eratzeko, sarrerako datu multzoko irizpide batzuen arabera horren antzekoa den $\mathbf{x}_i^{(p)}$ datulagin bat aukeratzen da. Aldiz, pare negatiboa osatzeko, $\mathbf{x}_i$ ainguraren aldean desberdina dela asumitzen dugun $\mathbf{x}_i^{(n)}$ datu-lagin bat aukeratzen da. Horrela, $f^{(a)}(\cdot)$ ereduak $\phi(\cdot)$ kodetzaileak sortutako ainguraren $\mathbf{z}_i = \phi(\mathbf{x}_i)$ eta pare positiboaren $\mathbf{z}_i^{(p)} = \phi(\mathbf{x}_i^{(p)})$ errepresentazioak hurbilduko ditu; pare negatiboaren $\mathbf{z}_i^{(n)} = \phi(\mathbf{x}_i^{(n)})$ errepresentazioa, aldiz, aldendu egingo du.

Oso teknika lagungarria izan daiteke bide laburrak ekiditeko. Hori ulertzeko, gogoratu dezagun sarreran azaldutako bide labur posiblea katuen eta txakurren irudien sailkapenean: etxe barruko atzealdea duten irudiei «katua» klasea eta kaleko irudiei «txakurra» klasea esleitzea. Bide labur hori

ekiditeko, klase bereko datu-laginak pare positibo gisa eta klase desberdinekoak pare negatibo gisa erabil ditzakegu (ikusi 7. irudia). Zehazki, gure aingura katu baten irudia bada, katu bat duen argazki bat eta txakur bat duen beste argazki bat har ditzakegu ainguraren pare positibo eta negatibo gisa, hurrenez hurren. Horrela, katuen irudien errepresentazioak hurbilago egongo dira kodetzaileak sortutako dimentsio txikiagoko espazioan, eta txakurren irudien errepresentazioak, aldiz, haietatik aldenduta geratuko dira. Horrek bide laburra ekiditen lagunduko du. Aitzakia-ataza hori eta sailkapenean oinarritutako zeregin-ataza era bateratuan ikas daitezke lehen deskribatutako bide laburra ekiditeko.



7. irudia. Testuinguru-kontrastea. Datuen testuinguruan oinarritutako irizpide bat izanik, jatorrizko datu-laginarekiko pare positiboa eta negatiboa osatuko duten bi datu-lagin aukeratuko dira. Horrela, helburua ainguraren eta pare positiboaren errepresentazioak hurbiltzea izango da; ainguraren eta pare negatiboaren errepresentazioak, aldiz, aldendu egingo dira.

Aurreko adibideaz gain, badaude beste zenbait lan testuinguru-kontrastea erabiltzen dutenak, hainbat datu motaren errepresentazio esanguratsuak ikasteko eta problema bakoitzaren testuinguruaren araberako pare kontrastatzaileak osatzeko irizpideak proposatzen.

5. ONDORIOAK

Ikaskuntza gainbegiratuaren mugak kontuan hartuta, ikaskuntza autogainbegiratua ikaskuntza-paradigma interesgarri moduan aurkeztu da literaturan azken urteotan. Ikaskuntza-paradigma horren helburua zeregin-ataza jakin baterako erabilgarriak izango diren datuen errepresentazioak

ikastea da, eta, horretarako, zeregin-atazako etiketak erabiliko ez dituen aitzakia-ataza bat baliatzen da.

Artikulu honetan, datuen errepresentazio esanguratsuen ikaskuntza bultzatzen duten ikaskuntza autogainbegiratuan oinarritutako aitzakia-atazak laburbildu ditugu. Nahiz eta aniztasun handia egon, aitzakia-atazak bi multzo nagusitan sailkatu ditugu: ikaskuntza autoprediktiboa eta ikaskuntza kontrastatzailea. Multzo bakoitzean aurki daitezkeen ataza motak definitu, eta ataza horiek askotariko datu motetan nola aplika daitezkeen erakusteko adibide argiak eman ditugu.

Hona hemen etorkizunerako proposatzen ditugun zenbait ikerkuntza-lerro ikaskuntza autogainbegiratuaren arloan:

- **Oinarri teorikoen garapena.** Nahiz eta ikaskuntza autogainbegiratuak errepresentazio esanguratsuen ikaskuntza bultzatzeko teknika oso baliagarriak biltzen dituen, oso lan gutxitan aztertu da paradigma horren arrakastaren arrazoiketa teorikoa. Ondorioz, beharrezkoa da metodologiaren oinarri teorikoak finkatzen dituzten lanak garatzea.
- **Aitzakia-atazaren aukeraketa.** Zeregin-ataza bakoitza egiten ikasteko baliagarriak izan daitezkeen aitzakia-atazen azterketari buruzko lanen falta dago. Interesgarria izango litzateke zeregin-ataza jakin bakoitza egiten ikasteko balio handia duten aitzakia-atazak diseinatzeko irizpideak esploratzea.
- **Aitzakia-ataza moten konbinazioa.** Ikaskuntza autoprediktiboan eta ikaskuntza kontrastatzailean oinarritutako aitzakia-atazak erabiltzen dituzten lanek emaitza onak ematen dituzte. Hala ere, aitzakia-ataza bat baino gehiago konbinatu daitekeenez lan berean, interesgarria litzateke ataza autoprediktiboak eta kontrastatzaileak konbinatzen dituzten lanak garatzea.

BIBLIOGRAFIA

[1] MICHALSKI, R. S., CARBONELL, J. G., and MITCHELL, T. M. (eds.). 2013. Machine learning: An artificial intelligence approach. Springer Science and Business Media.

[2] VOULODIMOS, A., DOULAMIS, N., DOULAMIS, A., & PROTOPAPADAKIS, E. 2018. Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018.

[3] CHOWDHARY, K., & CHOWDHARY, K. R. 2020. Natural language processing. Fundamentals of artificial intelligence, 603-649.

[4] BERTOLINI, M., MEZZOGORI, D., NERONI, M., & ZAMMORI, F. 2021. Machine Learning for industrial applications: A comprehensive literature review. *Expert Systems with Applications,* **175**, 114820.

[5] MIOTTO, R., WANG, F., WANG, S., JIANG, X., & DUDLEY, J. T. 2018. Deep learning for health- care: review, opportunities and challenges. *Briefings in bioinformatics,* **19(6)**, 1236-1246.

[6] OQUAB, M., BOTTOU, L., LAPTEV, I., & SIVIC, J. 2015. Is object localization for free?-weakly- supervised learning with convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 685-694).

[7] GEIRHOS, R., JACOBSEN, J. H., MICHAELIS, C., ZEMEL, R., BRENDEL, W., BETHGE, M., & WICHMANN, F. A. 2020. Shortcut learning in deep neural networks. *Nature Machine Intelligence,* **2(11)**, 665-673.

[8] BENGIO, Y., COURVILLE, A., & VINCENT, P. 2013. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence,* **35(8)**, 1798-1828.

[9] LIU, X., ZHANG, F., HOU, Z., MIAN, L., WANG, Z., ZHANG, J., & TANG, J. 2021. Self-supervised learning: Generative or contrastive. *IEEE transactions on knowledge and data engineering,* **35(1)**, 857-876.

[10] ZHAI, X., OLIVER, A., KOLESNIKOV, A., & BEYER, L. 2019. S4l: Self-supervised semi- supervised learning. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 1476-1485).

[11] TSCHANNEN, M., BACHEM, O., & LUCIC, M. 2018. Recent advances in autoencoder-based representation learning. *arXiv preprint arXiv:1812.05069.*

[12] GUI, J., CHEN, T., ZHANG, J., CAO, Q., SUN, Z., LUO, H., & TAO, D. 2023. A survey on self-supervised learning: Algorithms, applications, and future trends. *arXiv preprint arXiv:2301.05712.*

[13] MISRA, I., & MAATEN, L. V. D. 2020. Self-supervised learning of pretext-invariant representations. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 6707-6717).

[14] DOERSCH, C., & ZISSERMAN, A.(2017. Multi-task self-supervised visual learning. In Proceedings of the IEEE international conference on computer vision (pp. 2051-2060).

[15] GIDARIS, S., SINGH, P., & KOMODAKIS, N. 2018. Unsupervised representation learning by predicting image rotations. *arXiv preprint arXiv:1803.07728.*

[16] GOLAN, I., & EL-YANIV, R. 2018. Deep anomaly detection using geometric transformations. *Advances in neural information processing systems,* **31**.

[17] BLÁZQUEZ-GARCÍA, A., CONDE, A., MORI, U., & LOZANO, J. A. 2021. *Water leak detection using self-supervised time series classification. Information Sciences,* **574**, 528-541.

[18] ZHANG, J., SALEEBY, K., FELDHAUSEN, T., BI, S., PLOTKOWSKI, A., & WOMBLE, D. 2021, december. Self-supervised anomaly detection via neural autoregressive flows with active learning. In NeurIPS 2021 workshop on deep generative models and downstream applications.

[19] GHOSAL, D., MAJUMDER, N., MIHALCEA, R., & PORIA, S. 2021. Stack: Sentence ordering with temporal commonsense knowledge. *arXiv preprint arXiv:2109.02247.*

[20] BANK, D., KOENIGSTEIN, N., & GIRYES, R. 2023. Autoencoders. Machine learning for data science handbook: data mining and knowledge discovery handbook, 353-374.

[21] PATHAK, D., KRAHENBUHL, P., DONAHUE, J., DARRELL, T., & EFROS, A. A. 2016. Context encoders: Feature learning by inpainting. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2536-2544).

[22] HE, K., CHEN, X., XIE, S., LI, Y., DOLLÁR, P., & GIRSHICK, R. 2022. Masked autoencoders are scalable vision learners. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 16000-16009).

[23] FEICHTENHOFER, C., LI, Y., & HE, K. 2022. Masked autoencoders as spatiotemporal learners. *Advances in neural information processing systems,* **35**, 35946-35958.

[24] DEVLIN, J., CHANG, M. W., LEE, K., & TOUTANOVA, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805.*

[25] TSCHANNEN, M., BACHEM, O., & LUCIC, M. Recent advances in autoencoder-based representation learning. *arXiv 2018. arXiv preprint arXiv:1812.05069.*

[26] DAI, A. M., & LE, Q. V. 2015. Semi-supervised sequence learning. *Advances in neural information processing systems,* **28**.

[27] SCHIAPPA, M. C., RAWAT, Y. S., & SHAH, M. 2023. Self-supervised learning for videos: A survey. *ACM Computing Surveys,* **55(13s)**, 1-37.

[28] ZHANG, K., WEN, Q., ZHANG, C., CAI, R., JIN, M., LIU, Y., ... & PAN, S. 2024. Self-supervised learning for time series analysis: Taxonomy, progress, and prospects. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[29] JAISWAL, A., BABU, A. R., ZADEH, M. Z., BANERJEE, D., & MAKEDON, F. 2020. A survey on contrastive self-supervised learning. *Technologies,* **9(1)**, 2.

[30] CHEN, T., KORNBLITH, S., NOROUZI, M., & HINTON, G. 2020, november. A simple framework for contrastive learning of visual representations. In International conference on machine learning (pp. 1597-1607). PMLR.

[31] LE-KHAC, P. H., HEALY, G., & SMEATON, A. F. 2020. Contrastive representation learning: A framework and review. *Ieee Access,* **8**, 193907-193934.

[32] CAI, T. T., FRANKLE, J., SCHWAB, D. J., & MORCOS, A. S. 2020. Are all negatives created equal in contrastive instance discrimination?. *arXiv preprint arXiv:2010.06682.*

[33] ELDELE, E., RAGAB, M., CHEN, Z., WU, M., KWOH, C. K., LI, X., & GUAN, C. 2021. Time- series representation learning via temporal and contextual contrasting. *arXiv preprint arXiv:2106.14112.*