
TOOL WEAR PREDICTION WITH DEEP LEARNING MODELS

Gorka Dabó Aizpurua

Univ: University of the Basque Country
City: San Sebastián
gdabo001@ikasle.ehu.eus

Mikel Oscoz Rivas

Univ: University of the Basque Country
City: Irún
moscoz002@ikasle.ehu.eus

1 Proiektuaren deskribapena

Proiektu honek mekanizazio prozesuetan erreminten higadura aurreikusteko deep learning estrategia ikertzen du. Erreminta bat higatuta dagoen edo fabrikatutako piezen kalitatean eta ekoizpenaren eraginkortasunean eragin esanguratsurik izan daiteke, horregatik makina bat higatuta degoen ala ez aurreikusteko gaitasuna oso erabilgarria da.

Arazo horri aurre egiteko, CNC Mill Tool Wear datu-multzoa erabiltzen da, 18 mekanizazio prozesuri buruzko informazioa duena. Helburua neurona-sareen arkitekturak implementatzea eta ebaluatzea da, tresna higatuta dagoen ala ez zehazteko. Datuen kopuru mugatua dela eta, "utzi bat kanpo" (leave-one-out) baliozkotze gurutzatua eta "sliding-window" teknikak erabiliko dira ereduaren ebaluazioa neurtzeko.

Proiektuaren barruan sartzen dira, besteak beste, datuak egituratzeko ikuspegi egokia diseinatzea, deep learning arkitektura desberdinak alderatzea eta ebaluazio metodologia sendoa ezartzera. Emaitzak eta aurkikuntzak Colab koaderno batean implementatuko dira, garatutako ereduaren funtzionamendua frogatzeko.

2 Arazoa

Mekanizazio industrialeko prozesuetan, erreminten higadura faktore kritikoa da, eta fabrikatutako piezen kalitatean eta enpresaren produktibilitatean eragiten dezake. Higatutako tresna bat behar denenen detektatzen ez bada, akatsak sor ditzake produktuetan, atzerapenak produkzioan eta eragiketa kostuak handitu dezake.

Gaur egun, erreminten higaduraren identifikazioa eskuzko edo aurrez definitutako arauetan oinarritutako prozesu bat da, eta horrek detekzioaren zehaztasuna eta azkartasuna mugatzen ditu, bereziki automatizazio eta eraginkortasun maila handiak eskatzen dituzten industria eredu modernoetan, 4.0 Industria bezala. Errebisio periodikoak egiten dituzten enpresa batean, makinaren higadura ez dute antzematzeko teknikoa datorren hurrengora arte eta hain konpetitiboa den sektore batean honek galera handiak sor ditzake.

Proiektu honek lantzen duen arazoan deep learning-ean oinarritutako eredu bat garatzea da, tresna bat higatuta dagoen ala ez automatikoki aurreikustea ahalbidetuko duena, mekanizazio prozesuan erregistratutako parametroak erabiliz. Arazo hori sailkapen bitarreko ataza bat da, eta hainbat erronkari egin behar zaie aurre: datu erabilgarrien kopuru mugatuari edo ebaluazioaren fidagarritasunari beste batzuen artean.

3 Datuen deskribapena

Proiektu honetan erabilitako datuak Michigan-eko Unibertsitateko CNC mekanizazio ingurune batean egindako 18 esperimentuetatik dator. Esperimentuak egiteko, argizarizko blokeak erabili ziren oinarrizko material gisa, eta CNC fresatzeko makina bat, mekanizazio-prozesuaren parametroak erregistratzeko sentsoreekin ekipatuta. Helburua erreminten baldintzak aztertzea abiaduraren eta presioaren konfigurazio desberdinatan izan zen.

Datuak bi zatitan banatzen dira, alde batetik, esperimentu guztien datu orokorrak ditugu eta bestetik esperimentu bakoitzaren datu zehatzak.

3.1 Datu orokorrak (train.csv)

Fitxategi honek 18 esperimentuetako informazio laburtua dauka, datu guztien artean informazio garrantzitsuena hau da:

- Esperimentuaren zenbakia
- Erabilitako materiala (cera)
- Abiadura (feed rate)
- Presioa (clamp pressure)
- Erremintaren egoera (tool condition), "unworn" (ez-higatua) edo "worn" (higatua)
- Tresnaren ikuskapen bisualaren emaitza

3.2 Esperimentuan datuak (experimentxx.csv)

Fitxategi bakoitzak 100 ms-ko tarteetan erregistratutako neurketak ditu mekanizazio-prozesuetan. Neurketa horietan sartzen dira:

- Posizioa X, Y eta Z. ardatzetan
- Torlojuaren biraketa (spindle)
- Uneko abiadura (current feed rate)
- Exekutatzen ari den programa zenbakia

4 Diseinua eta implementazioa

4.1 Sare arkitektura

Proiektu honetan implementatutako sare neuronalak, HybridCNNRNN izenekoak, sare konboluzionalen (CNN) eta sare rekurrenteen (RNN) ezaugarriak konbinatzen ditu, mekanizazioaren prozesuan datuetan dauden ezaugarri lokalak eta mendekotasun sekuentzialak modu optimo batean aprobetxatzeko.

4.1.1 Geruza Konboluzionalak

Geruza konboluzionalak ezaugarri lokal esanguratsuenak ateratzeko diseinatuta daude. Lehen geruza konboluzionalan sarrera-ezaugarrien kopurua 47 da, 32 iragazki ditu eta padding 2 tamaina ertaineko tokiko patroiak harrapatzeko. Bigarren geruza konboluzionalan 64ra igotzen da iragazkien kopurua, ezaugarri abstraktuagoak ateratzeko. Iragazki txikiagoa erabiltzen du (kernel size = 3), neurriak gordetzen dituen padding batekin. Geruza konboluzional bakoitzaren ondoren batch normalizazioa dago entrenamendua egonkortzeko, ReLU aktibazio bat, eta maxpooling eragiketa bat dimentsionaltasuna murrizteko. Azkenik, Dropout geruza bat aplikatzen da 0.3 probabilitatearekin overfitting saihesteko.

4.1.2 Geruza errekurrenteak (LSTM)

Bi norazko LSTM sareak datuetan epe luzerako erlazioak sortzeko balio du. Sarreran CNNek prozesatutako ezaugarriak daude. Gero, bi LSTM geruza daude, noranzko bakoitzeko 128 unitate ezkuturekin (256 guztira). Dropout geruza errekurrenteen artean ere erabiltzen da.

4.1.3 Fully Connected geruzak

Azken geruzek ateratako ezaugarriak iragarpen bitar batean bihurtzen dituzte. Lehen geruzak dimentsionaltasuna 256tik 64ra murrizten du, ondoren ReLU aktibazioa eta dropouts. Bigarren geruzak irteera eskalarra sortzen du, eta, ondoren, Sigmidea aktibazio-funtzioa, tresna higatuta dagoen ala ez adierazten duen probabilitatea sortzeko.

Laburtuz, Sarrerako datuak berrantolatu egiten dira geruza konboluzionalek prozesatzeko. CNNek ateratako ezaugarri lokalak LSTM geruzetara pasatzen dira, denbora-erlazio zabalagoak atzemateko. LSTMen azken irteera fully connected geruzen bidez prozesatzen da, iragarpen bitarra sortzeko.

4.2 Eeduaren entrenamendua

Entrenamendu prozesua train loop funtziaren bidez implementatzen da. Funtzio honek sare neuronalaren parametroak optimizatzen ditu eta beraien garrantzia dev multzo batean ebaluatzen du. Funtzio horren optimizaziorako, ebaluaziorako eta datuak kudeatzeko teknika aurreratuak konbinatzen ditugu, eredu sailkapen bitarreko zereginera egokitzen dela ziurtatzeko.

Funtzioak sarrera hauek hartzen ditu: entrenatu beharreko eredu, entrenamendu eta validazio multzoetarako data loaderrak, gailua (CPU edo GPU), epoka kopurua, eta validazio gurutzatua erabiltzeko aukerako adierazlea (leave-one-out cross-validation). Optimizazioa Adam algoritmoa erabiliz egiten da, 5×10^{-4} -ko learning rate batekin eta L_2 erregularizazioa overfitting-a ekiditzeko.

Entrenamendu epoka bakoitzean, modeloa entrenamendu moduan (train) konfiguratzen da, eta batchez batch prozesatzen da sarrera datuen multzoa. Batch bakoitza hautatutako gailura eramatzen da, eta ereduaren iragarpenean kalkulatzeko erabiltzen da. Ondoren, errorea ebaluatzen da entropia gurutzatu bitarraren galera funtzioa erabiliz (Binary Cross-Entropy Loss); funtzio hori egokia delako sailkapen bitarreko arazoetarako. Sarearen parametroak Adam optimizatzialearen bidez eguneratzen dira. Prozesu horretan zehar, entrenamendu-multzoaren galera metatzen da, eta metrikak kalkulatzeko iragarpenean sortzen dira.

Entrenamendu-multzoaren iterazioa amaitu ondoren, eredu dev multzoan ebaluatzen da, eta ebaluazio-moduan konfiguratzen da (eval). Etapa honetan, gradienteen kalkulua desaktibatzen da kostu konputazionala murrizteko. Validazio multzoko iragarpenean benetako etiketekin alderatzen dira F1-score metrika kalkulatzeko, bai entrenamendu multzorako, bai validazio multzorako.

F1-score onena lortzen duen eredu, validazio multzoan, eredu onena bezala gordetzen da. Iterazioak amaitzean, funtzioak eredu optimizatu hori itzultzen du, iragarpenean erabiltzeko prest dagoena.

4.3 Eeduaren ebaluazioa

Eeduaren ebaluazioa evaluate model funtziaren bidez egiten da. Funtzio honi esker, entrenatutako ereduaren eraginkortasuna aztertu daiteke, leihoa mailan zein esperimentu mailan. Esan beharra dago, test garaian ezin denez "sliding-window" erabili, esperimentuak leihoen tamainako azpisekuentietan banatu ditugula entrenamendu garaian erabilitako sekuentzia luzer antzekoak erabiltzearen. Ikuspegi honek atazaren ezaugarriak hartzen ditu kontuan, non esperimentu bakoitza datuen leihoa ugariz osatuta dagoen. Funtzioak bozketa (majority voting) bat erabiltzen du azken iragarpenean deribatzeko, leihoa indibidualetarako egindako iragarpenean oinarrituta, esperimentu bakoitza klasifikatzeko.

Prozedura hasterakoan, eredu ebaluazio moduan konfiguratzen da, eta horrek gradienteen kalkulua desaktibatzen du karga konputazionala murrizten. Jarraian, ereduaren iragarpenean batchez batch sortzen dira, proba guztiak kontuan hartuta. Batch bakoitzerako, sarrerako leihoa hautatutako gailura mugitzen dira, eta ereduak leihoa bakoitzari lotutako probabilitateak kalkulatzentzen dira. Probabilitate horiek iragarpenean bitar batera bihurtzen dira, 0.5eko atalasea erabiliz. Iragarpenean eta benetako etiketak zerrendatan biltegiratzen dira, ondoren leihoa-mailan metrikak errazago kalkulatzeko.

Leihoa zein esperimentu mailako iragarpenean sortu ondoren, metrika estandarrak kalkulatzentzen dira, hala nola zehaztasuna edo F1-score. Metrika horiek bi mailetarako kalkulatzentzen dira (leihoa eta esperimentu), eta ereduak test multzoan duen errendimendua ikuspegi zehatzera ematen dute. Azkenik, funtziok hiztegi bat itzultzen du, kalkulatzeko metrikak, leihoa eta esperimentu mailako iragarpenean eta dagozkien benetako etiketak barne hartzen dituena.

4.4 Leave One Out Cross Validation Bariantea

Gure lehenengo implementazioan, Leave One Out Cross Validation teknika erabili genuen, baina azkenean "sliding-window" erabili genuenez, data kopuruarekin ez zen beharrezko. Horregatik, bariante modura jarri dugu Colab-aren bukaeran eta konparaketak egiteko balio dezake.

4.4.1 Leave One Out Cross Validation Arkitektura (no window)

HybridCNNRNNNoWindow klaseak sekuentzia osoak "sliding-window"-etan banatu gabe prozesatzeko diseinatutako ereduaren aldaera bat da. Geruza konboluzionalak (CNN) konbinatzen ditu ezaugarri lokalak ateratzeko, bi norabideko LSTM geruzak epe luzerako mendekotasunak harrapatzeko eta sigmoide funtzio batekin fully connected irteera bat sailkapen bitarrerako; aurreko bertsioak egiten zuen moduan.

Eredu honek sarrera doitzen du sekuentzia osoak zuzenean maneiatzeko, leihonetan oinarritutako aurreprozesamenduaren beharra ezabatuz. Arkitektura hau egokia da datuak testuinguru globalean aztertzeko, leihon irristakorrik erabiltzen dituen eredu nagusiarekin alderatuz.

4.4.2 Evaluazioa (loocv)

Loocv evaluation funtzioko "Leave-One-Out" (LOOCV) balidazio gurutzatuko prozesua implemenatzen du. Prozesu horretan, esperimendu bakotza test multzo gisa erabiltzen da, eta gainerakoek, berriz, entrenamendu multzoa osatzen dute. Lehenik, higatutako eta higatu gabeko esperimentuetan banatutako datuak kargatzen dira, eta zerrenda bakar batean konbinatzen dira iterazioa errazteko. Iterazio bakotzean, esperimendu bat test gisa bereizten da, eta gainerakoak entrenamendu gisa. Datuak ezaugarriak antolatzen dituen funtziobaten bidez prestatzen dira, eta normalizatzen dira, datu guztiak eskala uniformean daudela bermatzeko, eredura pasatu aurretik. Ez da leihon irristakorrik sortzen, sekuentzia osoekin lan egiten delako.

Erabilitako eredu HybridCNNRNNNoWindow bariantea da, entrenamendu begizta espezifiko bat erabiliz entrenatzen dena (train loop) tolestura bakotzeko datuekin. Behin entrenatu ondoren, eredu dagokion test multzoan evaluatzen da, eta iragarpenean benetako etiketekin alderatzen dira metrikak kalkulatzeko, hala nola F1 puntuazioa eta zehaztasuna. Metrika horiek tolestura bakotzerako biltzegiratzen dira.

Esperimentu guztiak iteratu ondoren, lortutako metriken batez bestekoa kalkulatzen da, hala nola doitasuna eta F1 score.

5 Emaitzan analisia

Orain, lortutako emaitzak komentatuko ditugu, HybridCNNRNN ereduarekin hasita.

5.0.1 HybridCNNRNN

Entrenamenduak erakusten du entrenamendu multzoaren galera nabarmen jaisten dela epoch-eten zehar, eta horrek adierazten du eredu datuetatik ikasten ari dela. Entrenamendu multzoaren F1 metrikak gora egiten du pixkanaka: lehen denboraldian 0.7247 izan zen, eta azkenekoan, berriz, 0.9829. Horrek esan nahi du doikuntza ona egin dela multzoan. Hala ere, balidazio multzoaren F1 metrika nahiko egonkor mantentzen da 0.61 inguru, aldaketa txikiiek baina entrenamenduan hobekuntza esanguratsurik gabe. Horrek esan nahi du eredu agian gehiegi egokitzten ari dela entrenamendu datuetara, edo balidazio datuak konplexuagoak edo ez hain adierazgarriak direla.

Leihon-mailako proba guztien emaitzak normalak dira, %41,38ko zehaztasunarekin, %85,71ko recall altuarekin eta %55,81eko F1-score batekin. Horrek esan nahi du ereduak leihon gehienak positiboko sailkatzea joera duela, eta horrek lagundu egiten diola recall altuari, baina zehaztasunak sufritu egiten du positibo faltsuen kopuru handi bat dagoelako.

Esperimentu mailan, emaitzak nabarmen hobetzen dira. Doitasuna %66.67koa da, eta recall-a %100era iristen du, horrek ereduak behar bezala identifikatzen dituela klase positiboko esperimendu guztiak adierazten du. F1-score %80 da, balio ona da, eta erakusten du eredu fidagarriagoa dela esperimendu osoen sailkapenean banakako leihonetan baino. Horren arrazoia izan daiteke esperimendu mailan majority voting banakako leihonetan egindako akatsak leuntzen laguntzen duela.

Azkenik, esperimendu mailako iragarpenek [0, 1, 1, 1] dira. Benetako etiketen [0.0, 0.0, 1.0, 1.0] aurrean erakusten du ereduak zuen identifikatzen dituela esperimendu positibotako bi, baina negatibo batean huts egiten duela. Horrek esan nahi du klase positiborako joera dagoela, eta datuen edo ereduaren diseinuaren desbalantze baten eragina izan liteke.

5.1 HybridCNNRNNNoWindow

Metrikak Leave-One-Out (LOOCV) balidazio gurutzatuko eredu emaitza inkonsisteanteagoak dituela erakusten dute eta, oro har, baxuak direla. Metrika orokorrean batez bestekoak 0.3236 F1-Score eta %43,76ko zehaztasuna erakusten du, ereduak kasu askotan datuetan patroiak behar bezala atzematea lortzen ez duela adieraziz.

Esperimentu bakotzeo banakako emaitzek aldakortasun handia dute. Esperimentu batzuek, 10ek adibidez, metrika oso altuak lortzen dituzte (F1-Score 0.9808koa eta %96,23ko zehaztasuna), eta horrek adierazten du eredu oso ondo bete daitekeela datuak adierazgarriagoak edo sailkatzeo errazagoak direnean. Hala ere, hainbat esperimendu, hala nola 18, 1, 2, eta 3, 0.0000 F1-Score bat dute, eta horrek esan nahi du ereduak ez duela lortzen instantzia positiborik behar bezala identifikatzea kasu horietan.

Zehaztasuna handiagoa da 10 eta 9 bezalako esperimentuetan, baina kasu gehienetan baxua edo moderatua izaten da, eta horrek iradokitzen du ereduak huts egiten duela bai klase positiboa detektatzerakoan, bai benetako positiboen eta negatiboen arteko balantze egokia mantentzeraoan.

Aipatu beharra dago, argi ez dugun arrazoiren batengatik, 15º esperimentua errorea ematen zuela LOOCV erabiltzean, uste dugu daturen batek NaN balioa duela baina ez dugunez arrazoia aurkitu, entrenamendu garaian try-except egitura bat implementatu dugu eta 15º experimentua ez da erabiltzen. Hori ere kontuan izaten da batazbesteko metrikak kalkulatzerakoan.

6 Ondorioak

Ikerlan honetan, neurona-sare konboluzionaleko eta errekurrenteko (HybridCNNRNN) eredu hibrido bat implementatu eta ebaluatu da, baldintza espezifikoak dituzten esperimentuetako datuak sailkatzea. Azterketan zehar, ereduaren errrendimenduan alde nabarmenak ikusi ziren. Horrek zereginaren konplexutasuna eta datuak emaitzetan duten eragina nabarmenzen ditu.

Lehen fasean, eredu emaitza nahiko onak lortu zituen metrika globalei dagokienez, batez besteko 0.8-ko F1-Score batekin, gehiengoaren bozketa aplikatu ondoren. Hala ere, ereduak orokortze gaitasun mugatua erakutsi zuen, leihor mailan %45 inguruko zehaztasuna lortu zuelako. Portaera hori esperimentuen arteko heterogeneotasunarenagatik izan daiteke, baita klaseen arteko balizko desbalantze bat ere, ereduak hain adierazgarriak ez diren datuetan patroiak detektatzeko duen gaitasunari eragin dezakeena.

Beste alde batetik, ereduaren bariantea, ebaluazio zehatzagoa eta doituagoa egiteko implementatuak, LOOCV validazioan oinarritua, emaitza erratikoagoa erakutsi zuen. Eredauak errrendimendu bikaina lortu zuen esperimentu batzuk gorabehera, hala nola 10 esperimentua 0.98ko F1-Score batekin, beste kasu batzuetan, emaitzak oso baxuak izan ziren, eta 0.0000 F1-Score batera ere iritsi ziren hainbat esperimentutan. Portaera horrek sarrerako datuen kalitatearen aldakortasuna eta datuen azterketa sakonagoaren garrantzia azpimarratzen ditu. Esperimentuen arteko puntuazioen desberdintasun handiak erakusten du ereduak datuen ezaugarri espezifikoekiko duen sentikortasuna.

Ondorio orokorra da, eredu hibridoak datu esperimentalen sailkapenerako irtenbide oso ona eskaintzen duen arren, bere errrendimendua, neurri handi batean, datuen kalitatearen eta sendotasunaren mende dagoela. Gainera, klaseak ekilibratzeko tekniken implementazioa eta esperimentu mota bakotzerako ikuspegia pertsonalizatuagoa hartu behar dira kontuan, baina daukagun datu multzoa erreduzituarekin normala da.

Azkenik, ikerketa honetan etorkizunean hobekuntza gehiago egiteko bide librea ustean du. Teknika asko daude zeinek ereduaren errrendimendua hobetu dezakete, adibidez arreta mekanismoen implementazioa beste batzuen artean.

7 Colab Notebook-aren esteka

Hau da Colab koadernoaren esteka