

Kategorizacija podataka iz mjerenja $12C(n,p)$ i $12C(n,d)$ reakcija primjenom metoda strojnog učenja

Barović, Tadej

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:366014>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-29**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



Kategorizacija podataka iz mjerenja $12C(n,p)$ i $12C(n,d)$ reakcija primjenom metoda strojnog učenja

Barović, Tadej

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:366014>

Rights / Prava: [In copyright](#)

Download date / Datum preuzimanja: **2022-09-12**



Repository / Repozitorij:

[Repository of Faculty of Science - University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Tadej Barović

KATEGORIZACIJA PODATAKA IZ
MJERENJA $^{12}\text{C}(n,p)$ I $^{12}\text{C}(n,d)$
REAKCIJA PRIMJENOM METODA
STROJNOG UČENJA

Diplomski rad

Voditelj rada:
doc. dr. sc. Petar Žugec

Zagreb, 2022.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

| | |
|--|------------|
| Sadržaj | iii |
| 1 Uvod | 1 |
| 2 Eksperiment | 2 |
| 2.1 n_TOF postrojenje | 2 |
| 2.2 Motivacija eksperimenta | 3 |
| 2.3 Eksperimentalni postav | 4 |
| 2.4 Analiza podataka eksperimenta | 5 |
| 3 Kategorizacija čestica | 8 |
| 3.1 Uloga simulacije | 8 |
| 3.2 Neuronske mreže | 9 |
| 3.3 Implementacija GEANT4 simulacija | 14 |
| 3.4 Obrada podataka | 18 |
| 3.5 Implementacija neuronske mreže | 22 |
| 4 Zaključak | 28 |
| Bibliografija | 29 |

Poglavlje 1

Uvod

U mnogim istraživanjima šireg znanstvenog područja susrećemo se s problemom statističkog zaključivanja. Posebno možemo naglasiti područje nuklearne fizike i fizike elementarnih čestica gdje su mjerenja podložna raznim pozadinskim i kompetitivnim utjecajima. Jedan primjer takvog problema jest identifikacija reakcija $^{12}\text{C}(n, p)$ i $^{12}\text{C}(n, d)$ preko mjerenih varijabli. Odvajanje skupova koji predstavljaju pojedinu reakciju u koordinatnom prostoru mjerenih varijabli može se ostvariti ručnim određivanjem granica. Takva primitivna metoda, osim što zahtijeva puno vremena i napora, ima ograničene sposobnosti klasificiranja reakcija. Razvojem tehnologije i računalnih sposobnosti otvorila se mogućnost korištenja metoda strojnog učenja u svrhu rješavanja takvih problema. Implementacija alata strojnog učenja zahtijeva velike skupove podataka. Nadalje, kako bi se osigurala maksimalna učinkovitost alata strojnog učenja, potrebno je takve podatke detaljno analizirati i obraditi određenim metodama. Proces prikupljanja podataka, njihova obrada te implementacija alata strojnog učenja kod problema identifikacije reakcija $^{12}\text{C}(n, p)$ i $^{12}\text{C}(n, d)$ cilj je ovog diplomskog rada. Sama identifikacija, tj. klasifikacija spomenutih reakcija, uvjet je za daljnju rekonstrukciju značajnih veličina u istraživanju [1] reakcija $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$. Spomenuto istraživanje kontekst je ovog rada. Ovdje ćemo naglasiti da će se u radu razmatrati reakcije $^{\text{nat}}\text{C}(n, p)^{12}\text{B}$ i $^{\text{nat}}\text{C}(n, d)^{11}\text{B}$, no zbog zanemarivog udjela izotopa ^{13}C u prirodnom ugljiku rezultate pripisujemo reakcijama nad izotopom ^{12}C .

U 2. poglavlju opisana je motivacija istraživanja, n_TOF postrojenje CERN-a u kojemu se istraživanje provodi te sama matematička analiza mjerenih veličina i veličina od interesa za istraživanje. U 3. poglavlju iznesen je glavni dio – simulacija i obrada podataka te implementacija neuronske mreže kao alata strojnog učenja. U zadnjem, 4. poglavlju, zapisat ćemo glavne zaključke diplomskog rada.

Poglavlje 2

Eksperiment

2.1 n_TOF postrojenje

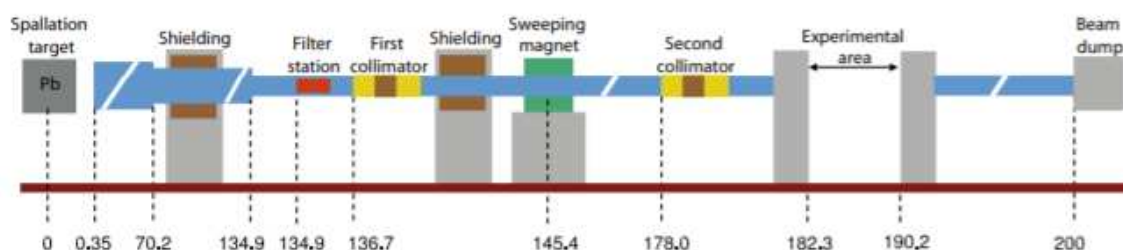
CERN, Europska organizacija za nuklearna istraživanja, najveći je laboratorij za nuklearnu fiziku i fiziku elementarnih čestica na svijetu. U n_TOF postrojenju neutronske vremena proleta s CERN-a, opširnije opisanom u [2], primarni predmet istraživanja su neutronske inducirane reakcije. Neutroni su proizvedeni protonskom indukcijom u masivnom olovnom bloku. Protonski snop energije 20 GeV stvoren je protonskim sinkrotronom CERN-a i sadrži u prosjeku $7 \cdot 10^{12}$ protona po 7 ns širokom pulsu perioda od 1,2 s. Takav snop u sudaru s jezgama olova proizvede oko 300 neutrona po protonu. Neutronske snop se otprilike prostire na intervalu energija od ≈ 10 meV pa sve do ≈ 10 GeV, a takav se nastavlja gibati kroz vakuumske cijevi prema dvjema eksperimentalnim prostorijama.

Kako bismo odredili energiju pojedinog neutrona, koristimo se tehnikom vremena proleta, što je i poveznica s imenom postrojenja. Naime, vrijeme proleta, tj. gibanja neutrona, dovodimo u vezu s njegovom kinetičkom energijom preko relativističke relacije:

$$E_n = m_n c^2 \left(\frac{1}{\sqrt{1 - \left(\frac{L}{cT}\right)^2}} - 1 \right), \quad (2.1)$$

gdje je m_n masa neutrona, c brzina svjetlosti u vakuumu, a veličina L predstavlja prijeđenu udaljenost neutrona u vremenskom intervalu T , tj. tijekom vremena proleta neutrona.

Eksperimentalna prostorija 1 (eng. experimental area 1, EAR1) na horizontalnoj je udaljenosti od 185 m od spalacijskog bloka (slika 2.1). Velika udaljenost eksperimentalne prostorije – preko relacije 2.1 – osigurava dobru rezoluciju u rekonstrukciji energija neutrona. Snop se inicijalno filtrira od raznih onečišćenja; borom obogaćena voda preko reakcije $^{10}\text{B}(n, \alpha)^7\text{Li}$ sprječava nastanak gama zraka energije 2,2 MeV nastalih uhatom neutrona na jezgama vodika (prisutnih u rashladnom sustavu spalacijskog bloka). Nada-



Slika 2.1: Skica linije puta neutronske snopove do Eksperimentalne prostorije 1

lje, 145 m od spalacijskog bloka iz snopove se uklanjaju nabijene čestice jakim magnetskim poljem. Konačno se kolimatorima snop oblikuje te takav dolazi do eksperimentalne prostorije. Eksperimentalna prostorija 2 (eng. experimental Area 2, EAR2) nalazi se 20 m iznad spalacijskog bloka. Blizina EAR2 omogućava veće tokove neutrona te je samim time namijenjena za istraživanje reakcija malog udarnog presjeka. Ovdje definiramo tok neutrona kao:

$$\phi(E) \equiv \frac{dN(E)}{dE}, \quad (2.2)$$

gdje je $dN(E)$ ukupan broj neutrona na energiji E (unutar beskonačno uskog intervala dE) po nominalnom protonskom snopu od $7 \cdot 10^{12}$ protona, integriran po cijeloj površini presjeka snopove. Mjerenje raspodjele toka neutrona po energijama detaljno je opisano u [3]. Broj neutrona mjeri se preko reakcija koje su prihvaćene kao standard na određenom intervalu energija. Jedna takva reakcija je $^{235}\text{U}(n,f)$ za energije od 0,15 MeV do 200 MeV, što obuhvaća interval [15 MeV, 30 MeV] koji će nam biti relevantan za eksperiment opisan u sljedećim poglavljima.

2.2 Motivacija eksperimenta

Ugljik je primarni element organskih spojeva i zastupljenost njegovih atoma u ljudskom tijelu od $\approx 12\%$ [4] čini ga trećim najbrojnijim elementom u sastavu ljudskog tijela. Konkretno, izotop ^{12}C čini 98,9% zastupljenosti u prirodnom ugljiku te kao takav čini reakcije $^{12}\text{C}(n,p)^{12}\text{B}$, $^{12}\text{C}(n,d)^{11}\text{B}$ i $^{12}\text{C}(n,\alpha)^9\text{Be}$ ¹ relevantnima u hadronskoj terapiji. Doprinos dozi u zračenom tkivu reakcije ostvaruju kroz produkt reakcije (proton, deutron, α -čestica) i elektron visoke energije iz β raspada nestabilne jezgre bora. Poznavanje udarnog presjeka reakcija i raspodjele kutova emisije produkata reakcije, osim za izračun doze u tkivu, za svrhu ima i omogućavanje pravilnog dizajniranja zaštite i kolimatora nuklearnih postrojenja u medicinskim, znanstvenim, energetskim i drugim sektorima. U dosadašnjim is-

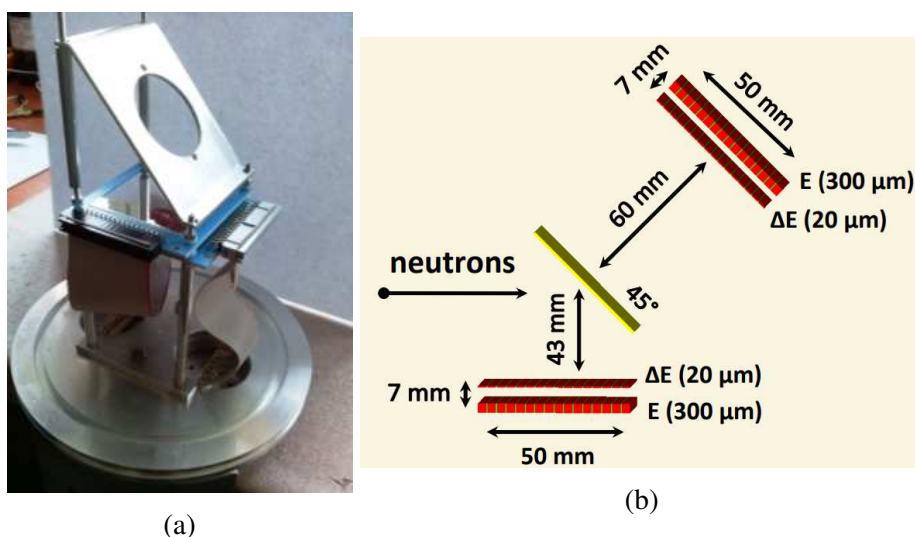
¹Iako je reakcija $^{12}\text{C}(n,\alpha)^9\text{Be}$ relevantna u hadronskoj terapiji, te je njezino istraživanje poželjno, u našem radu nećemo ju razmatrati.

traživanjima i u dostupnim bazama podataka informacije o navedenim reakcijama veoma su oskudne te su većinom u međusobnom neslaganju. Eksperiment [5] upravo je i došao do zaključka da je integralni udarni presjek $^{12}\text{C}(n, p)^{12}\text{B}$ reakcije od praga reakcije do 10 GeV podcijenjen u većini dostupnih baza podataka. No baze podataka se slažu da se maksimum udarnog presjeka nalazi oko 20 MeV. Imajući u vidu takva (ne)slaganja iznimno je poželjna provedba novih i pouzdanijih mjerenja.

2.3 Eksperimentalni postav

Kako bismo iz mjerenih podataka mogli izračunati udarni presjek, eksperimentalni postav mora zadovoljavati određene uvjete. U eksperimentu [5, 6] provedeno je mjerenje integralnog udarnog presjeka. Broj takvih reakcija mjerio se indirektno, preko β raspada jezgre kćeri ^{12}B . Izlazna čestica se tada nije dovodila u vezu s energijom ulaznih neutrona zbog vremenske zadržke u β raspadu. Sada je produkte reakcije potrebno dovesti do izravne veze s energijom neutrona, tj. produkte reakcije klasificirati po intervalima energija neutronske snop. Nadalje, detektor ne smije biti izravno izložen neutronske snopu, gdje će biti podvrgnut neutronske reakcijama. Također, produkti mjerene reakcije su teške nabijene čestice te takve iznimno brzo deponiraju energiju u materiji kroz koju prolaze. Tada je mjerenje potrebno provesti u vakuumu, kako bismo spriječili gubitke energije između uzorka i detektora te tako uspjeli detektirati česticu. Naposljetku, mjerenje mora osigurati dovoljno informacija za kategorizaciju relevantnih produkata reakcije i odbacivanje ostalih. Zadovoljavajuća rješenja takvih uvjeta pronađena su u već prije korištenom teleskopu silicijskih detektora (eng. multi-channel silicon telescope, SITE) na n-TOF-u. Radi se o konfiguraciji koja koristi korelacije $\Delta E - E$ između energija deponiranih u dvama slojevima silicijskog teleskopa u svrhu klasifikacije čestica. Naime, svaki produkt reakcije deponirat će dio energije u prvom ΔE sloju debljine $20 \mu\text{m}$, a potom u drugom E sloju SITE-a debljine $300 \mu\text{m}$. Slojevi su udaljeni 7 mm , a svaki od slojeva veličine je $5 \text{ cm} \times 5 \text{ cm}$. Slojevi su podijeljeni u 16 traka veličine $5 \text{ cm} \times 3 \text{ mm}$ odvojenih neaktivnim silicijem. Svaka čestica deponirat će energiju u određenom paru traka, ovisno o kutu svoje emisije. U eksperimentu korištena su dva teleskopa, jedan je postavljen ispred uzorka, paralelan s uzorkom (uzorak zatvara kut od 45° s osi upadnog snop), a drugi iza njega te je paralelan sa snopom, pokrivajući tako veći kutni prostor oko uzorka (slika 2.2). Sama reakcija događa se u uzorku prirodnog ugljika. Dimenzije i položaj uzorka dizajnirani su tako da on pokriva čitav neutronske snop.

Općeniti je cilj u eksperimentu imati što veći broj detektiranih (opaženih) produkata uz što manju količinu pozadinskih događaja. Tomu će najviše doprinostiti izbor debljine uzorka. Veća debljina uzorka doprinosit će većem broju reakcija, kako relevantnih tako i pozadinskih. No osim te karakteristike, deblji uzorak ujedno će biti skloniji apsorpciji nabijenih produkata. Najbolji kompromis između korisnih i kompetitivnih, nerelevantnih



Slika 2.2: a) Slika silicijskog detektora. Ovdje je prikazan slučaj sa samo jednim detektorom, onim paralelnim s neutronske snopom. b) Skica eksperimentalnog postava korištenog u mjerenjima $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$ reakcija. Takav sadrži dva silicijska detektora: jedan ispred, a drugi iza uzorka ugljika.

informacija bit će debljina pločice od 0,25 mm. S obzirom na to da je zakrenuta za 45° od osi upadnog snopa, put neutrona kroz ugljik će iznositi 0,35 mm.

2.4 Analiza podataka eksperimenta

Za početak, spomenimo da je cilj sljedeće analize dovesti mjerene veličine – broj pojedinih reakcija– i veličine od interesa – udarni presjek reakcije– u matematičku vezu. Iako su u eksperimentu [1] razmatrane $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$ reakcije, radi jednostavnosti u sljedećoj analizi govorit ćemo samo o reakciji $^{12}\text{C}(n, p)^{12}\text{B}$.

Definirajmo prvo veličine koje su u mjerenju relevantne. Neka x predstavlja energijsko stanje jezgre kćeri ^{12}B nastale u reakciji. Što je energijska razina jezgre kćeri veća, to je veći udio kinetičke energije neutrona predan samoj jezgri. E_n predstavlja energiju neutrona. $\chi = \cos \theta$ je kosinus kuta raspršenja čestice produkta u sustavu centra mase. Radi jednostavnosti izražavanja na χ referirat ćemo se samo na kao kut čestice.

Osnovni je korak analize rastav veličine broja *proizvedenih* produkata reakcije. Definirajmo veličine o kojima prethodna ovisi. Neka n predstavlja gustoću atoma uzorka n iskazanu u broju atoma po jedinici površine. Veličine σ_r i σ_t su udarni presjek razmatrane reakcije i ukupni udarni presjek. $\Phi(E)$ ukupan je neutronske tok kojemu je uzorak izložen.

Uvedimo sada broj produkata reakcije u ovisnosti o prethodnim veličinama:

$$d^2\mathcal{N}(x, E, \chi) = \Phi(E) \left(1 - e^{-n\sigma_t(E)}\right) \frac{\varrho_r(x, E, \chi)}{\sigma_t(E)} \mu(E) dE d\chi. \quad (2.3)$$

Ovdje $1 - e^{-n\sigma_t(E)}$ predstavlja vjerojatnost bilo koje reakcije neutrona s uzorkom u prvom koraku. Primjetimo faktor $\mu(E)$; on opisuje korekciju višestrukog raspršenja neutrona. Za tanki uzorak – koji ćemo definirati kao onog sa svojstvom $n\sigma_t(E) \ll 1$ – faktor $\mu(E)$ postaje zanemariv, tj. $\mu(E) \approx 1$. Također, za tanki uzorak član $1 - e^{-n\sigma_t(E)}$ može se aproksimirati kao $n\sigma_t(E)$ zadržavanjem najnižih članova Taylorovog razvoja. Parcijalni udarni presjek reakcije izgleda ovako:

$$\varrho_r(x, E, \chi) = \sigma_r(E) \rho(x, E) A(x, E, \chi). \quad (2.4)$$

Možemo ga rastaviti po sljedećim faktorima: $\sigma_r(E)$ je udarni presjek relevantne reakcije, $\rho(x, E)$ omjer je grananja reakcije po pobuđenim stanjima jezgre kćeri x , a $A(x, E, \chi)$ kutna je raspodjela emitiranih produkata reakcije. Veličine neovisne o detaljima reakcije koju proučavamo možemo izdvojiti kao:

$$w(E) \equiv \frac{1 - e^{-n\sigma_t(E)}}{\sigma_t(E)} \Phi(E) \mu(E), \quad (2.5)$$

što s prethodno opravdanim aproksimacijama postaje:

$$w(E) \approx n\Phi(E), \quad (2.6)$$

Sada jednadžbu 2.3 preko jednadžbi 2.4 i 2.5 zapisujemo:

$$d^2\mathcal{N}(x, E, \chi) = w(E) \sigma_r(E) \rho(x, E, \chi) dE d\chi. \quad (2.7)$$

Jednadžba 2.7 predstavlja teorijsku relaciju za broj *proizvedenih* produkata reakcije s parametrima x , E i χ . Broj *detektiranih* produkata s druge strane ovisi o samom eksperimentalnom postavu i njegovoj osjetljivosti te je tada potrebno uvesti veličinu koja opisuje to svojstvo eksperimentalnog postava. U tu svrhu definirat ćemo vjerojatnost koincidence detekcije produkta reakcije $\varepsilon_{ij}(x, E, \chi)$. Pod koincidence detekcijom ovdje mislimo na onu pri kojoj se energija produkta deponira u oba sloja silicijskog teleskopa. Neka je $d^2N_{ij}(x, E, \chi)$ broj takvih detekcija u i -toj vrpici ΔE i j -toj vrpici E sloja silicijskog teleskopa. Vjerovatnost koincidence detekcije tada je:

$$\varepsilon_{ij}(x, E, \chi) = \frac{d^2N_{ij}(x, E, \chi)}{d^2\mathcal{N}(x, E, \chi)}. \quad (2.8)$$

Do ovih vjerojatnosti tehnički je veoma teško doći eksperimentalno te se u tu svrhu koriste podatci iz GEANT4 simulacija. Sada možemo preko jednadžbi 2.7 i 2.8 dobiti vezu između broja detektiranih čestica i traženih parametara reakcije:

$$d^2 N_{ij}(x, E, \chi) = \varepsilon_{ij}(x, E, \chi) w(E) \sigma(E) \rho(x, E) A(x, E, \chi) dE d\chi, \quad (2.9)$$

Podatci će biti analizirani u energijskim intervalima konačne širine. Definirajmo tada proizvoljni interval energije neutrona:

$$\mathbb{E} \equiv [E_{min}, E_{max}]. \quad (2.10)$$

Sumiranje po svim energijskim razinama te integracija po svim kutovima i energijama neutrona u intervalu 2.10 rezultiraju ukupnim brojem čestica detektiranih u pojedinom paru vrpca detektora:

$$N_{ij}^{(\mathbb{E})} = \sum_x \int_{\mathbb{E}} dE \int_{-1}^1 d\chi \varepsilon_{ij}(x, E, \chi) w(E) \sigma(E) \rho(x, E) A(x, E, \chi), \quad (2.11)$$

Prethodna jednadžba teorijska je veza između mjerenih veličina i veličina od interesa. U članku [7] detaljnije je opisana prethodna analiza kojoj je cilj iz mjerenih podataka rekonstruirati udarni presjek, omjere grananja i kutnu distribuciju emitiranih produkata reakcija (protona i deuteronu). Kako bismo iz detektiranih podataka odredili broj pojedinih reakcija N_{ij} potrebno je prethodno podatke klasificirati. Zadatak ovog diplomskog rada je implementirati alat kojim ćemo iz mjerenih podataka rekonstruirati broj detektiranih čestica pojedine reakcije.

Poglavlje 3

Kategorizacija čestica

3.1 Uloga simulacije

Priroda eksperimenata koji se provode u n_TOF-u takva je da zahtijeva podosta aproksimacija nepoznatih veličina. Nepoznate veličine su ili intrinzični dio rastava parametara reakcije iz mjerenih podataka ili su to veličine uvjetovanje eksperimentalnim postavom. U one veličine koje su intrinzični dio rastava ubrajamo omjere grananja po energijskim razinama jezgre kćeri $\rho(x, E)$ i kutne raspodjele $A(x, E, \chi)$ opisane u prethodnom poglavlju. Primjer veličina koje su uvjetovana eksperimentalnim postavom je vjerojatnost detekcije 2.8, pozadinska zračenja okoline ili raspršenja neutrona o eksperimentalni postav. Kako je mjerenje navedenih veličina tehnički veoma teško, lakše i idealnije rješenje jest uvesti ih iz vanjskih izvora. Osim tih veličina, za uspješnu obradu podataka potrebni su nam i alati kojima ćemo iz mjerenih podataka dobiti relevantne informacije. Jedan takav alat, koji ćemo implementirati u ovom poglavlju, jest klasifikator samih reakcija. Takav će omogućiti postavljanje veze između informacija detektiranih silicijskim teleskopom s reakcijom od značaja. Tek tim uvjetom možemo doći do broja detekcija pojedine reakcije 2.9.

Kako bismo došli do podataka potrebnih za implementaciju klasifikatora te kako bismo implementirali spomenuti klasifikator, provest ćemo simulacije u softverskim paketima kao što je GEANT4. GEANT4 (skraćeno od *geometry and tracking*) omogućava geometrijsku implementaciju eksperimentalnog postava te simulaciju prolaska čestica kroz isti. Simulacije u svrhu određivanja vjerojatnosti detekcije 2.8 opisane su u članku [7]. U nastavku poglavlja opisujemo proces simulacije reakcija u svrhu njihove klasifikacije te konačno implementiramo sami klasifikator.

U svrhu razlikovanja reakcija definirat ćemo trodimenzionalni parametarski prostor u kojem je svaka detektirana reakcija jedinstveno prikazana koordinatama:

$$(E_n, \Delta E, E). \tag{3.1}$$

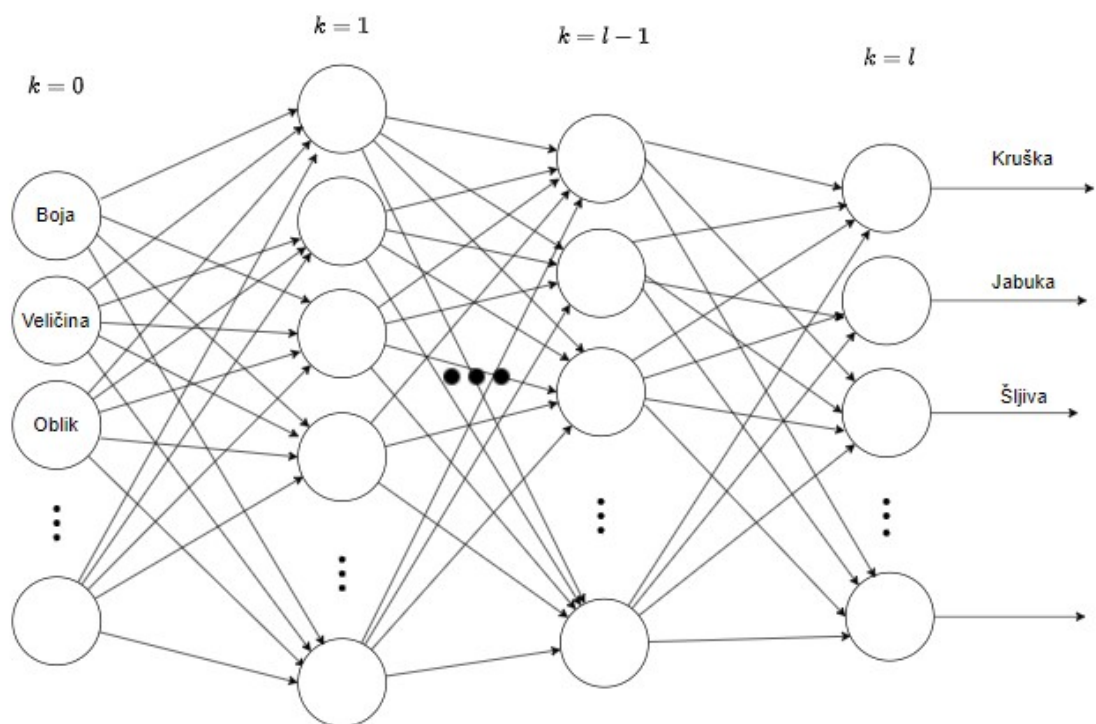
Koordinate predstavljaju energiju upadnog neutrona E_n te deponirane energije pripadne izlazne čestice ΔE i E u prvom ΔE sloju i u drugom E sloju. Upravo te vrijednosti opisuju tip reakcije. Kako bismo uspješno kategorizirali reakcije potrebno je odrediti dijelove parametarskog prostora u kojima će se određena reakcija pojavljivati.

3.2 Neuronske mreže

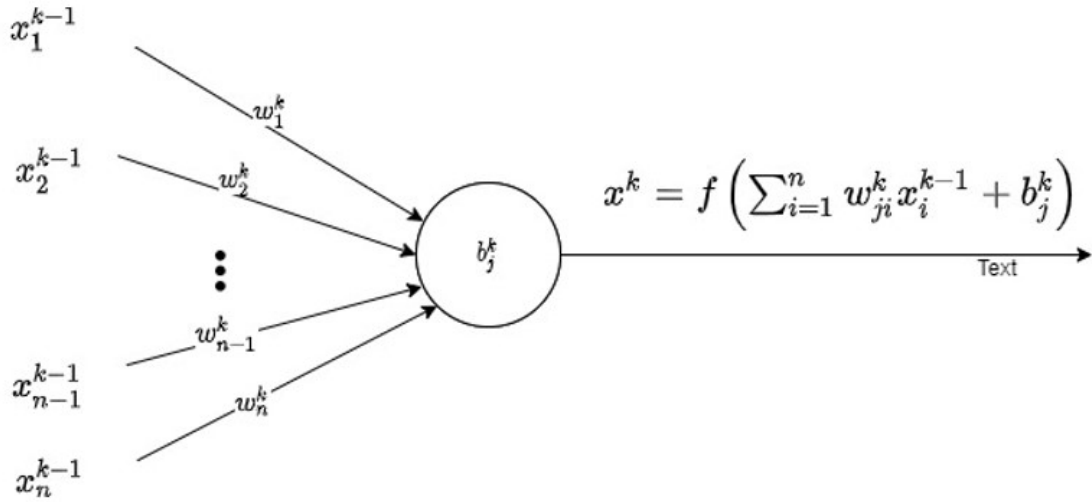
Neuronske mreže, kao što im i ime nalaže, svojim radom oponašaju biološki živčani sustav. Cilj takvih mreža je, analogno biološkom mozgu, točno riješiti postavljeni problem tj. točno odgovoriti na postavljeno pitanje, a u slučaju netočnog ishoda promijeniti komponente svojeg ponašanja s ciljem poboljšanja izvedbe. Mozak može naučiti rješavati probleme uz pomoć jako malog skupa informacija, npr. možemo naučiti razlikovati voće preko samo jednog primjera podatka o vrsti voća (npr. slike voća). Ovakvo se učenje u računalnom sustavu oslikava kroz veoma zahtjevan proces. Osnova učenja u računalnom sustavu je, za razliku od biološkog učenja, velik broj informacija (npr. tisuće primjera slika za pojedino voće). Način funkcioniranja takvog računalnog procesa razvijao se od sredine 20. stoljeća, no tek nedavno razvijena mogućnost računalnog hardvera omogućila je korištenje ovakve tehnologije u mnogim područjima kao što su robotika, analiza umjetnog vida i sluha, meteorologija itd. Naime, sve ove grane imaju pristup te se bave velikim količinama podataka. Ti podatci su digitalizirani te se algoritmima strojnog učenja obrađuju matematičkim operacijama koje zahtijevaju veće količine procesorskih resursa kakvi postaju dostupni tek tijekom zadnjih 10 godina.

Acikličke neuronske mreže (eng. feedforward neural network) sastoje se od niza slojeva, a svaki od njih sastoji se od niza neurona. Informacije u mreži, kao što ime mreže kaže, kreću se isključivo jednosmjerno, od neurona prvog sloja prema zadnjem sloju. Prvi sloj nazivamo ulaznim jer je broj neurona n u njemu zadan brojem parametara ulazne informacije (npr. boja, veličina i oblik voća). Svaki od ulaznih neurona prima numeričku vrijednost tog parametra. Ulazni podatak može se zapisati n -dimenzionalnim vektorom \vec{x} gdje koordinate vektora x_i predstavljaju numeričku vrijednost parametra i . Posljednji sloj nazivamo izlaznim, moguće vektorske vrijednosti njegove neurona bijektivno se preslikavaju u skup mogućih odgovora mreže (vrsta voća). Između ulaznog i izlaznog nalazi se proizvoljan broj slojeva koje nazivamo skrivenim. Svi neuroni u potpuno povezanim neuronskim mrežama (osim onih u ulaznom sloju) primaju informaciju od svakog neurona prethodnog sloja, te svi neuroni (osim onih u izlaznom sloju) šalju informacije svim neuronima sljedećeg sloja (slika 3.1).

Svaka veza između dva neurona ima svoju mjeru relevantnosti: težinu w (eng. weight). Isto tako i svaki neuron ima svoju mjeru relevantnosti: pomak b (eng. bias). Izlazna vrijednost neurona u potpuno povezanim mrežama ovisi o izlaznim vrijednostima svih neurona prethodnog sloja, težinama veza kojima se s prethodnim slojem veže te svojem



Slika 3.1: Potpuno povezane acikličke neuronske mreže: svaki neuron u sloju povezan je sa svakim neuronom sljedećeg sloja



Slika 3.2: Izlazna vrijednost neurona x^k ovisi o izlaznim vrijednostima neurona prethodnog sloja x_i^{k-1} , mjerom relevantnosti veza (težinama) s neuronima prethodnog sloja w_i^k te sa svojom mjerom relevantnosti (pomakom) b_j^k .

pomaku:

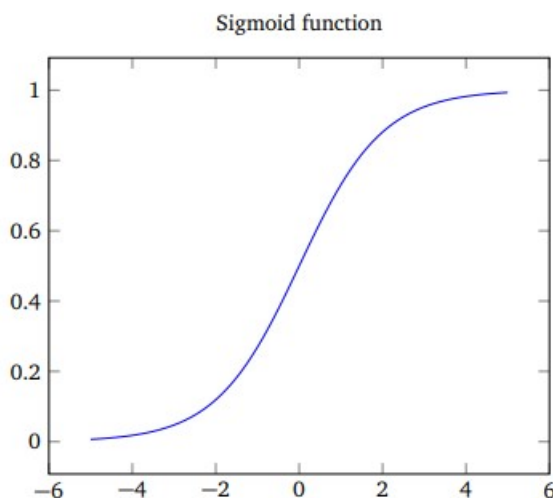
$$x_j^k = f\left(\sum_{i=1}^n w_{ji}^k x_i^{k-1} + b_j^k\right). \quad (3.2)$$

Na slici 3.2 možemo vidjeti da $k \in [1, l]$ predstavlja indeks sloja ispred veze ($k = 1$ je prvi skriveni sloj itd.), i indeks neurona u sloju iza veze, a j indeks neurona ispred veze. Sada vidimo da izlazna funkcija svakog neurona mreže (osim onog u ulaznom sloju) sadržava $n + 1$ stupanj slobode, gdje je n broj neurona $k - 1$ sloja. Upravo ti parametri mreže, w_{ji}^k i b_j^k podložni su učenju. Arhitektura mreže (broj slojeva i neurona u njoj), te aktivacijska funkcija f proizvoljni su i izabire ih korisnik.

f je aktivacijska funkcija neurona. Najčešći primjer je sigmoidalna (logistička) funkcija (slika 3.3) koja izlazne vrijednosti neurona ograničava na interval $\langle 0, 1 \rangle$:

$$f(z) = \frac{1}{1 + e^{-z}}, \quad (3.3)$$

gdje je raspis argumenta z vidljiv iz 3.2. Za učenje mreže prednost sigmoidalne funkcije je derivabilnost na cijeloj domeni. No kao što ćemo vidjeti, brzina učenja pada zajedno s derivacijom aktivacijske funkcije. Udaljavanjem vrijednosti varijable od nule derivacija u toj točki ove funkcije teži nuli. Time je učenje, osim na području oko nule, značajno usporeno.



Slika 3.3: Sigmoidalna funkcija

Algoritam učenja mreže

Pronalaženje optimalnih parametara mreže (težine w_{ji}^k i pomaci b_j^k) za zadanu arhitekturu provodimo učenjem mreže. Ovaj proces je dobio ime zbog svoje analogije s biološkim učenjem. Mreži je zadano pitanje (ulazni podatak) te joj je nakon njenog odgovora (izlaznog podatka) otkriven točan odgovor. S obzirom na veličinu pogreške (metrika vektora izlaznog i točnog podatka) mreža mijenja svoje parametre. Takav postupak iterativno se ponavlja te tako njegova efektivnost očito ovisi o broju podataka koje možemo dati mreži.

Spomenuta metrika između točnog i izlaznog vektora također je proizvoljna funkcija te je nazivamo funkcijom podatci (eng. cost function). Najjednostavniji i najčešći primjer takve je srednje kvadratno odstupanje:

$$C(w, b) = \frac{1}{2n} \sum_{x=1}^{2n} \|(y(x) - a(x))\|^2, \quad (3.4)$$

Ovdje w i b predstavljaju skupove vrijednosti težina i pomaka u mreži. $y(x)$ predstavlja željeni (točni) vektor mreže za ulazni vektor x , a $a(x)$ dobiveni izlazni vektor za ulazni vektor x . Operator $\|\cdot\|$ jest norma vektora. Osim srednjeg kvadratnog odstupanja koje uglavnom koristimo u regresijskim problemima, kod klasifikacijskih problema (kao na primjeru voća) kao funkciju cijene koristit ćemo kategoričku unakrsnu entropiju (eng. categorical crossentropy).

Konačan cilj učenja mreže je pronaći vrijednosti w i b u kojima funkcija cijene postiže minimum. Zbog prirode aktivacijske funkcije varijabli analitički izračun ovog minimuma

nije moguć, no ono što možemo izračunati jest gradijent funkcije C po varijablama w i b za pojedini podatak x . Iterativnim kretanjem u smjeru suprotnom od gradijenta te njegovim ponovnim izračunom, kretat ćemo se u smjeru minimuma (tzv. gradijentni spust):

$$\begin{aligned}w' &= w - \eta \frac{\partial C}{\partial w}, \\b' &= b - \eta \frac{\partial C}{\partial b},\end{aligned}\tag{3.5}$$

gdje je η korak učenja (eng. learning rate). Ovakav proces s velikim brojem podataka x postaje vremenski zahtjevan. Tada koristimo metodu stohastičnog gradijentnog spusta (eng. stochastic gradient descent) tj. radimo prosjeke gradijenta po skupu podataka:

$$\begin{aligned}w' &= w - \frac{\eta}{m} \sum_{i=1}^m \frac{\partial C_{x_i}}{\partial w}, \\b' &= b - \frac{\eta}{m} \sum_{i=1}^m \frac{\partial C_{x_i}}{\partial b},\end{aligned}\tag{3.6}$$

gdje je m broj podataka u skupu. Problem kod metode gradijentnog spusta je nemogućnost izlaska iz lokalnog minimuma funkcije cijene. Određene metode pomažu u rješavanju tog problema. Jedna od takvih je simulacija momenta inercije, gdje osim samog gradijenta, na smjer promjene varijabli w i b utječe i komponenta brzine koja je pod utjecajem „akceleracije”, tj. gradijenta ∇C .

Računanje samog gradijenta za svaku varijablu funkcije cijene provodimo algoritmom propagacije unazad (eng. backpropagation). Algoritam se zasniva na veličini pogreške neurona j u sloju k :

$$\delta_j^k = \frac{\partial C}{\partial a_j^k} \frac{\partial f}{\partial z_j^k},\tag{3.7}$$

gdje je z_j^k definirano s:

$$z_j^k = \sum_{i=1}^n w_{ji}^k x_i^{k-1} + b_j^k.\tag{3.8}$$

Ova veličina jednostavno se računa za izlazni sloj. No izračun te veličine za neuron i skrivenog sloja zahtijeva težinsko sumiranje svih pogrešaka neurona j sljedećeg sloja:

$$\delta_i^k = \left(\sum_{j=1}^n w_{ji}^{k+1} \delta_j^{k+1} \right) \frac{\partial f}{\partial z_i^k}.\tag{3.9}$$

Izračunom pogreška neurona u mreži dolazimo do komponenti potrebnog gradijenta:

$$\begin{aligned}\frac{\partial C}{\partial b_j^k} &= \delta_j^k, \\ \frac{\partial C}{\partial w_{ji}^k} &= a_i^{k-1} \delta_j^k.\end{aligned}\tag{3.10}$$

Ovdje možemo vidjeti da brzina učenja preko pogreške neurona 3.7 ovisi o derivaciji funkcije f . Prethodni izračun kompaktnije možemo zapisati matricama:

$$\delta^k = ((w^{k+1})^T \delta^{k+1}) \odot f'(z^k).\tag{3.11}$$

gdje je $w^{k+1} = [w_{ji}^{k+1}]$ matrica težina između slojeva k i $k + 1$, $\delta^{k+1} = [\delta_j^{k+1}]$ vektor pogreška neurona sloja $k + 1$ i $z^k = [z_j^k]$ vektor veličina z_j^k neurona sloja k . Ovdje funkcija f' djeluje po komponentama matrice, a operacija množenje \odot nad komponentama matrice. Ovime opravdavamo tvrdnju da provedba ovog algoritma zahtijeva sve značajnije računalne mogućnosti s povećanjem broja neurona i slojeva.

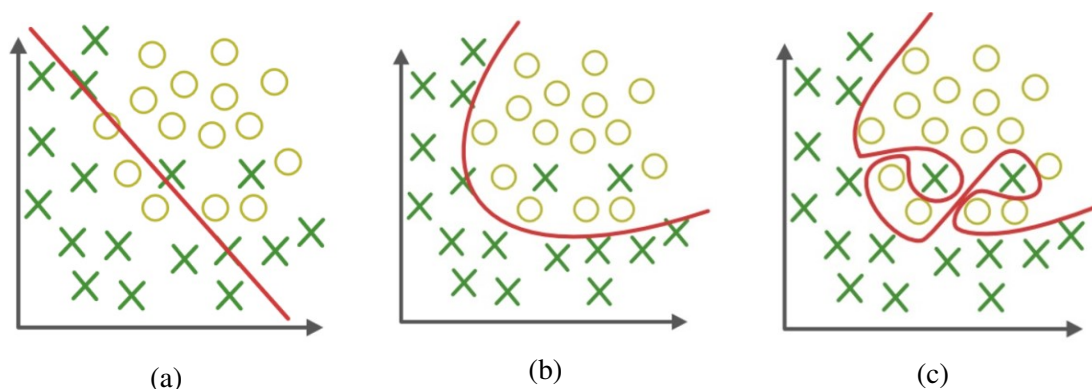
Problem prenaučivosti

Korisnikova intervencija u algoritam učenja mreže obuhvaća izbor broja parametara w i b . Cilj takve intervencije je smanjivanje složenosti algoritma i ubrzavanje njegove provedbe. Povećavanjem parametara mreže povećat će se i preciznost kategorizacije, no dobivena granica će tada točno klasificirati samo podatke na kojima je naučena, tj. smanjit će se mogućnost generalizacije na druge skupove podataka; „naučit će napamet”. S druge strane, premalo parametara neće omogućiti dovoljnu preciznost (slika 3.4).

U svrhu provjere stupnja prenaučivosti skup podataka simulacije podijelili smo na dva podskupa. Prvi je skup za učenje (eng. train set) te će se sastojati od 80% podataka i služiti će samo za učenje mreže, dok će drugi, manji skup za testiranje (eng. test set) biti korišten za testiranje mreže. Ako je sposobnost klasifikacije na skupu za testiranje manji od sposobnosti klasifikacije na skupu za treniranje, zaključujemo prenaučivost mreže – mreža je naučila napamet.

3.3 Implementacija GEANT4 simulacija

Kako je zadaća ovog diplomskog rada sam proces simulacije te analiza simuliranih podataka i implementacija klasifikatora podataka, pretpostavit ćemo pojednostavnijenu verziju eksperimentalnog postava opisanog u poglavlju 2. U tom pojednostavnijem slučaju zane-marujemo podijeljenost slojeva teleskopa na trake. Simulacije reakcija provest ćemo u GE-ANT4 softveru implementiranom objektno orijentiranim tehnikama u C++ programskom jeziku. U sljedećem tekstu opisat ćemo implementaciju simulacija u GEANT4 softveru.



Slika 3.4: (a) Podnaučenost (eng. underfit): funkcija s nedovoljno parametara ne može ostvariti potrebnu preciznost. (b) Primjer idealne funkcije: optimalan omjer preciznosti i broja parametara. (c) Prenaučenost (eng. overfit): s previše parametara funkcija uči nebitne informacije koje negativno utječu na sposobnost generalizacije. [8]

Prvi je korak definiranje geometrije i materijala eksperimentalnog postava te dodjela uloga pojedinim komponentama postava. Ovaj korak ostvarit ćemo sukladno opisu eksperimentalnog postava u poglavlju (2.3), no sada bez podijeljenosti teleskopa na vrpce. Sljedeći korak zahtijeva definiranje parametara generirane čestice, tj. njenog vektora položaja i brzine. Ako bi generirana čestica bila neutron, što nam prirodno nameće eksperiment, tada moramo simulirati samu reakciju koja slijedi, $^{12}\text{C}(n, p)^{12}\text{B}$ ili $^{12}\text{C}(n, d)^{11}\text{B}$. Na implementaciju takve reakcije u GEANT4 paketima ne možemo se osloniti zbog nedostatka adekvatnih eksperimentalnih podataka. Osim toga, zbog tankoće uzorka ($n\sigma_t \ll 1$) relativno mali broj produkata od značaja bio bi stvoren s obzirom na broj upadnih neutrona. Tada bi nam trebala značajna razina procesorskih resursa kako bismo zadovoljili statističku preciznost dovoljnim brojem detektiranih čestica. U tom slučaju odlučujemo se za izravno generiranje produkata reakcije. Generirat ćemo svaku reakciju određeni broj puta. Početni parametri reakcije bit će dodijeljeni s obzirom na (1) pretpostavljene parametre neutronskog snopa – prostorne i energijske raspodjele neutrona – (2) Q-vrijednosti reakcije u uzorku ugljika i (3) energijskoj razini jezgre kćeri E_x . Raspodjele ovih parametara simulirane reakcije najbolje je zadati kao jednostavne. Konačne podatke simulacije tada možemo u budućoj analizi korigirati težinskim faktorima s obzirom na raspodjele stvarnih parametara. Raspodjelu neutrona u poprečnom presjeku snopa tada ćemo aproksimirati normalnom raspodjelom čija je standardna devijacija $\sigma = 0,7$ cm. Početni položaj produkata reakcije pratit će istu raspodjelu kao i snop neutrona, ali unutar uzorka ugljika; tada trebamo imati na umu da je uzorak pod kutom od 45° s obzirom na upadni snop. Također, moramo uzeti u obzir i debljinu uzorka te definirati raspodjelu produkata duž neutronskog snopa unutar njega. Ovu raspodjelu aproksimiramo uniformnom jer zbog tankoće

uzorka možemo zanemariti prigušenje neutronske snage. Što se tiče vektora brzine produkta, njegov smjer će biti dodijeljen izotropnom kutnom raspodjelom u sustavu centra mase neutrona i atoma ugljika, a potom će se transformirati u laboratorijski sustav – sustav mirujuće mete. Iznos vektora brzine ćemo pronaći klasičnim kinematičkim izračunom. Kinetička energija produkta, osim o energiji neutrona, ovisit će i o Q -vrijednosti reakcije te energijskoj razini jezgre kćeri E_x . Reakcije našeg eksperimenta su endotermne. Po definiciji takve reakcije imaju negativnu Q -vrijednost. To znači da će – u slučaju kada je jezgra kćeri u osnovnom energijskom stanju E_x – ukupno Q energije biti reakcijom pretvoreno u masu. Radi lakšeg uvida u predznake koristit ćemo definiciju gdje je Q -vrijednost pozitivna. Spomenimo da je prag reakcije nešto veći nego sama Q -vrijednost. Naime, takav slučaj bi vrijedio u sustavu centra mase, no u laboratorijskom sustavu gdje je ukupna količina gibanja veća od nule reakcija zahtijeva više energije:

$$E_{prag} = \left(1 + \frac{m_p}{m_t}\right) Q. \quad (3.12)$$

Ovdje m_p i m_t predstavljaju masu projektila i mete. Dodamo li Q -vrijednosti i energijsko stanje razine kćeri E_x dolazimo do ukupne kinetičke energije produkata u laboratorijskom sustavu:

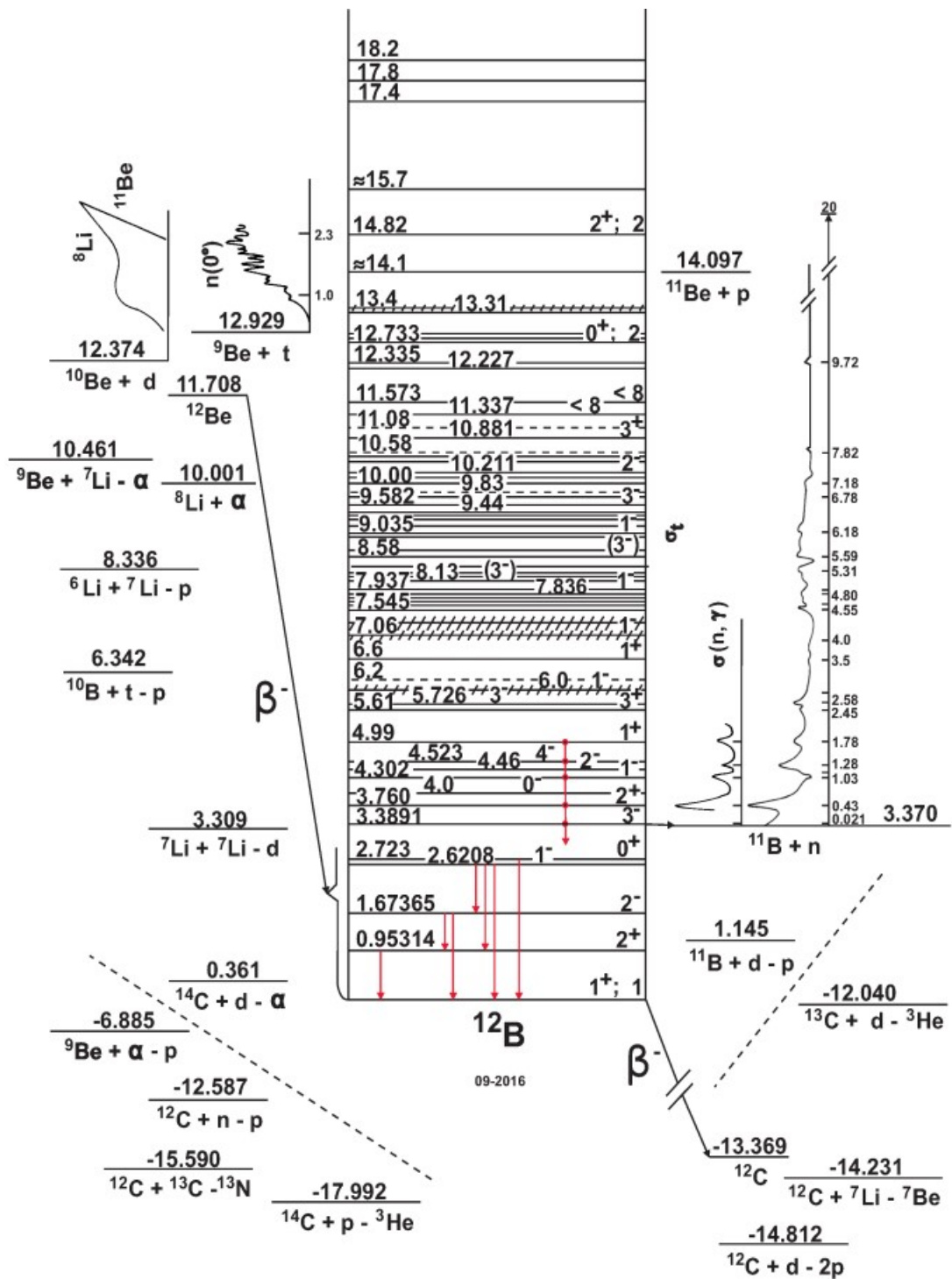
$$E_k = E_n - \left(1 + \frac{m_p}{m_t}\right) (Q + E_x). \quad (3.13)$$

Energijska stanja jezgre, kao i ona cijelog atoma su diskretna, no širine kvantnomehaničkih stanja im se preklapaju. Pogledamo li raspodjelu energijskih razina jezgri nakon reakcija u dostupnim podatcima [9], vidjet ćemo glatke i kontinuirane funkcije (slika 3.5). Promotrimo sada sve udjele grananja $\rho(x, E_n)$ na nekoj energiji E_n . Veličinu $\rho(x, E_n)$ opisujemo kao raspodjelu reakcija po diskretnim energijskim razinama x na energiji E_n . Takvu ćemo u simulacijama, zbog svoje neodređenosti, aproksimirati uniformnom i kontinuiranom raspodjelom. Interval mogućih energijskih razina jezgre definirat ćemo na intervalu deset relevantnih energijskih razina jezgre, dakle: $E_x \in [E_0, E_9]$.

Označimo određenu reakciju s \mathcal{R} . U našem slučaju promatramo šest relevantnih reakcija: $\mathcal{R} \in \{^{12}\text{C}(n,p)^{11}\text{B}, ^{12}\text{C}(n,d)^{10}\text{B}, ^{12}\text{C}(n,t)^9\text{B}, ^{13}\text{C}(n,p)^{12}\text{B}, ^{13}\text{C}(n,d)^{11}\text{B} \text{ i } ^{13}\text{C}(n,t)^{10}\text{B}\}$. Energije praga za iste prikazane su u tablici 3.1. Moramo naglasiti da se u eksperimentu, osim prethodnih reakcija, pojavljuju i (n, np) reakcije. Takve u ovom radu nećemo obuhvatiti. Naime, kod reakcija s tri ili više izlaznih čestica, raspodjele njihovih energija postaju kontinuirane, čak i u sustavu centra mase. Takve su nam nepoznanica i njihovo uvođenje dodatno komplicira račun, a njihovo izostavljanje neće utjecati na cilj ovog rada.

Konačno, nakon zadavanja energije neutrona E_n , Q -vrijednosti i energetske razine jezgre kćeri E_x , te uz pomoć poznatih masa jezgri izotopa ugljika i bora te čestica neutrona, protona, deuteronu i tritona dolazimo do iznosa brzine promatrane izlazne čestice.

Nakon što smo u opisanom procesu definirali i implementirali geometriju i materijale postava, početni položaj i brzinu početnih čestica, daljnji proces simulacije prepuštamo



Slika 3.5: Energijske razine jezgre ^{12}B [9]

| \mathcal{R} | Q / MeV | $E_{\text{prag}} / \text{MeV}$ |
|--|------------------|--------------------------------|
| $^{12}\text{C}(\text{n,p})^{12}\text{B}$ | 12,59 | 13,64 |
| $^{12}\text{C}(\text{n,d})^{11}\text{B}$ | 13,73 | 14,89 |
| $^{12}\text{C}(\text{n,t})^{10}\text{B}$ | 18,93 | 20,52 |
| $^{13}\text{C}(\text{n,p})^{13}\text{B}$ | 12,65 | 13,64 |
| $^{13}\text{C}(\text{n,d})^{12}\text{B}$ | 15,31 | 16,49 |
| $^{13}\text{C}(\text{n,t})^{11}\text{B}$ | 12,42 | 13,36 |

Tablica 3.1: Q-vrijednosti i energije praga za promatrane reakcije u laboratorijskom sustavu.

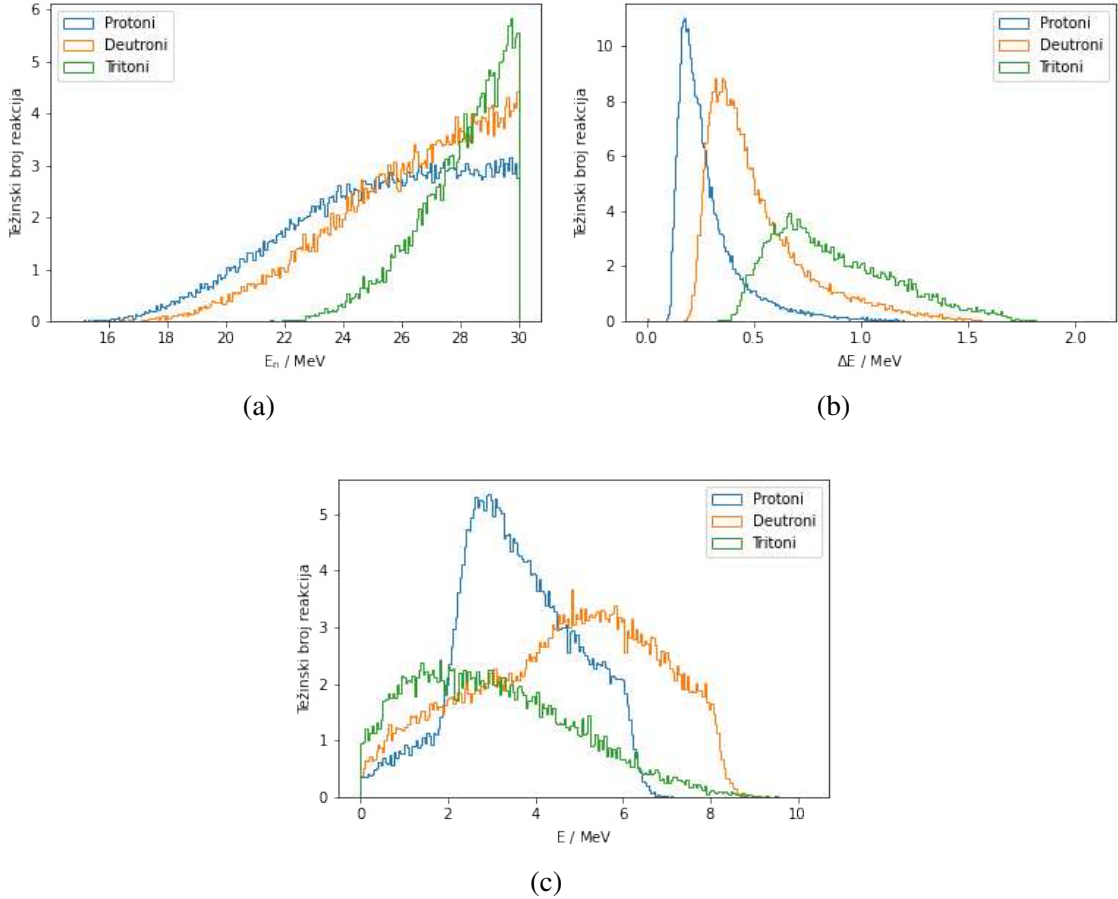
GEANT4 softveru. U našoj implementaciji početna čestica simulacije je izlazni produkt promatrane reakcije (proton, deuteron, triton), a kod takvih simulacija vrijeme provedbe ne predstavlja problem. Tada možemo generirati statistički dovoljnu količinu događaja. Taj broj je izabran okvirno kao 10^5 po vrsti reakcije, što čini ukupno $6 \cdot 10^5$ pojedinih reakcija.

3.4 Obrada podataka

Obradom podataka identificirat ćemo korisne podatke (koicidentalne detekcije u ΔE i E sloju teleskopa), omogućit ćemo jasniji prikaz i interpretaciju podataka te ćemo povećati efikasnost njihove klasifikacije. Kompletnu obradu i prikaz podataka izvršit ćemo preko biblioteka NumPy, Scikit-learn i Matplotlib u programskom jeziku *Python*.

Dodjeljivanje težinskih faktora

U simuliranim podacima nalaze se i one reakcije čije izlazne čestice zbog nedovoljne energije nisu prošle kroz prvi, ΔE sloj, tj. treća komponenta parametara ($E_n, \Delta E, E$) nije određena. Takve podatke početnom filtracijom izbacujemo. Relevantna područja za veličine $E_n, \Delta E$ i E određena su redom: [15 MeV, 30 MeV], [0 MeV, 2 MeV] i [0 MeV, 10,5 MeV]. Od predhodnih područja mi utječemo samo na širinu E_n vrijednosti. Pretpostavka je da će se na određenim energijama neutrona dogoditi značajno preklapanje različitih reakcija u $\Delta E - E$ prostoru. Naš je zadatak odrediti do koje energije možemo odvojiti reakcije sa značajnom sigurnošću. U tom slučaju maksimalna energija neutrona mora biti iznad točke preklopa reakcija u $\Delta E - E$ području. U [1] pretpostavljeno je da možemo potpuno identificirati reakciju $^{12}\text{C}(\text{n,p})$ na energijama do 25 MeV; u tom slučaju za gornju granicu izabiremo 30 MeV. Što se tiče donje granice vrijednosti E_n , nju izabiremo s obzirom na minimalnu energiju kojom je moguće detektirati produkt reakcije u oba sloja silicijskog teleskopa. Takva je nešto veća no prag reakcije jer zahtjeva dodatnu energiju za probi-



Slika 3.6: Raspodjele parametara po vrsti simulirane reakcije za događaje detektirane u koincidenciji: a) raspodjela energije neutrona E_n , b) raspodjela energije ΔE sloju teleskopa i c) raspodjela energije u E sloju teleskopa.

janje prvog sloja. Određena minimalna vrijednost E_n je 15 MeV. Nadalje, svaki događaj simulacije nije jednako bitan. Naime, broj simuliranih događaja svake pojedine reakcije je gotovo pa jednak, što se ne slaže s brojem eksperimentalno detektiranih reakcija. To rezultira netočnim prikazivanjem granica između skupova točaka prostora $(E_n, \Delta E, E)$ koji se vezuju uz specifični produkt reakcije.

Označimo početnu jezgru s \mathcal{J} , a produkt reakcije s \mathcal{P} . Primijetimo kako je izotop bora jednoznačno određen početnom jezgrom i česticom produkta (proton, deutron, triton). Reakciju tada možemo jedinstveno zapisati uređenim parom $(\mathcal{J}, \mathcal{P})$. Stvaran broj događaja ovisit će o udarnom presjeku reakcije $\sigma_{\mathcal{J}\mathcal{P}}$, zastupljenosti pojedinog izotopa $A_{\mathcal{J}}$ u uzorku prirodnog ugljika i širini intervala energije neutrona $(30 \text{ MeV} - E_{\text{prag}})$ na kojoj se reakcije

ostvaruju. Tako ćemo važnost pojedinog događaja korigirati s obzirom na ukupan broj simuliranih događaja $N_{\mathcal{J},\mathcal{P}}$ preko koeficijenta težine događaja:

$$w_{\mathcal{J},\mathcal{P}} = \frac{A_{\mathcal{J}} \cdot \sigma_{\mathcal{J},\mathcal{P}} \cdot (30 \text{ MeV} - E_{\text{prag}})}{N_{\mathcal{J},\mathcal{P}}}. \quad (3.14)$$

Radi kasnije potrebe prebrojavanja pojedinih događaja indeksirat ćemo ih *jedinstvenim* indeksom. Parametarski zapis događaja preslikavamo surjeksijom:

$$(E_n, \Delta E, E) \longrightarrow i. \quad (3.15)$$

Reakcije na različitim izotopima ugljika, čije su izlazne čestice iste, nismo u mogućnosti raspoznati u eksperimentalnim podatcima. Naime, izlazne čestice ostavljaju gotovo isti $\Delta E - E$ trag, a energijski prag tih reakcija zanemarivo se razlikuje. Iz tog razloga nadalje nećemo razlikovati takve reakcije, te ćemo ih označavati samo s oznakom produkta \mathcal{P} i indeksom i koji predstavlja parametarski zapis događaja. Našim podatcima sada dodajemo još jedan stupac, težinu događaja $w_{\mathcal{P},i}$. Sada skup reakcija u našem skupu podataka postaje $\{\text{}^{nat}\text{C}(n,p), \text{}^{nat}\text{C}(n,d), \text{}^{nat}\text{C}(n,t)\}$. No zbog veoma malog udjela izotopa ^{13}C u prirodnom ugljiku ^{nat}C rezultate klasifikacije možemo pripisati reakcijama nad ^{12}C izotopom.

Transformacija

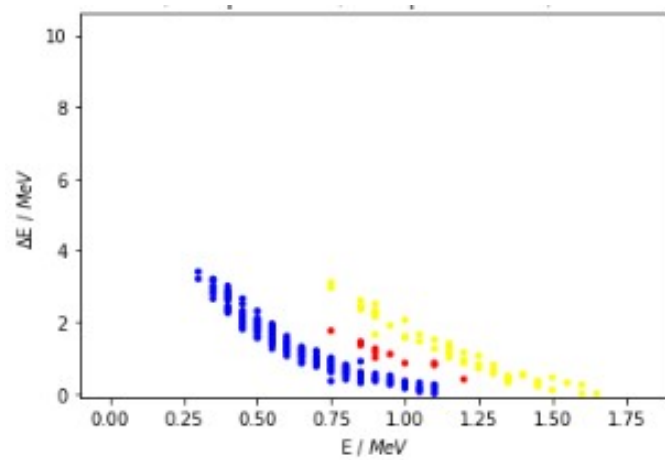
Algoritam učenja neuronskih mreža osjetljiv je na apsolutne vrijednosti ulaznih podataka. To znači da će brzina gradijentnog spuštanja 3.6 biti uvelike određena ulaznim parametrima – mreža će puno brže učiti na događajima većih ulaznih vrijednosti. Energija neutrona doseže 30 MeV-a, dok energija deponirana u prvom sloju doseže tek nekoliko MeV-a. To znači da će energija neutrona biti puno relevantniji podatak za učenje naše mreže. Takav efekt je nepoželjan. Taj problem riješit ćemo normalizacijom podataka ulaznih parametra $X \in \{E_n, \Delta E, E\}$:

$$X'_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}. \quad (3.16)$$

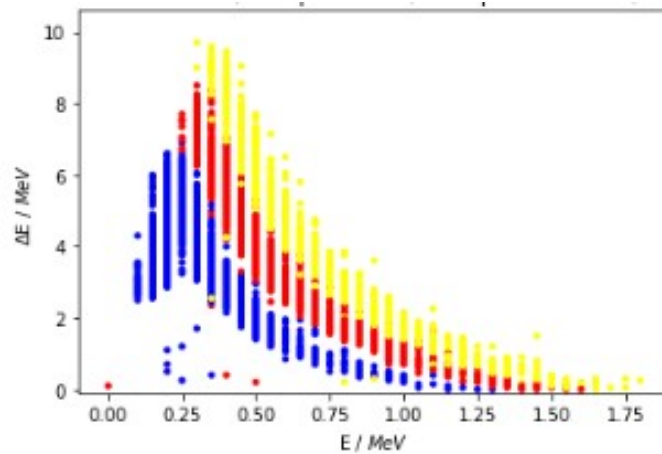
Druga komponenta istog problema dolazi zbog neujednačenih raspodjela po istom parametru, čime manjina elemenata na krakovima raspodjele brže uči mrežu; primjer takve je ΔE i E_n . Kod veličine ΔE , zbog specifičnog oblika njene raspodjele, taj problem možemo umanjiti transformacijom:

$$\Delta E' = \ln(\Delta E + 1).$$

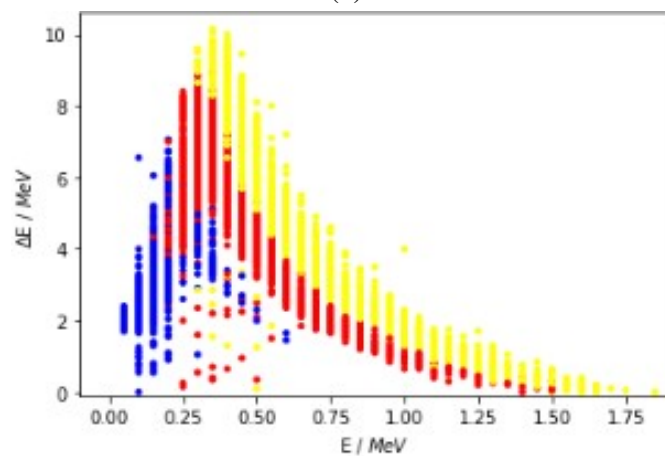
Raspodjela veličine E_n ima drugačiju karakteristiku te se oblikom bitno razlikuje od raspodjele ΔE . Transformacije u svrhu ujednačavanja istih bit će drugačije. Raspodjela veličine E_n je takva da količina podataka raste s veličinom parametra. Mreža će u ovom slučaju gotovo zanemariti podatke na malim energijama neutrona (prvo zbog većeg broja podataka



(a)



(b)



(c)

Slika 3.7: Prikazani su vokseli te njihove vrijednosti pridružene energiji neutrona. E_n -širina je 0,33 MeV, ΔE -širina je 0,05 MeV, E -širina je 0,05 MeV. a) $E_n \in [17 \text{ MeV}, 17,33 \text{ MeV}]$, b) $E_n \in [25 \text{ MeV}, 25,33 \text{ MeV}]$ i c) $E_n \in [29 \text{ MeV}, 29,33 \text{ MeV}]$.

na većim energijama, a potom zbog njihove veće vrijednosti), čime će naš klasifikator biti neupotrebljiv u tom području energija.

Prvi način da riješimo problem je da djelujemo na veličinu E_n čime ćemo povećati ujednačenost razdiobe iste. U slučaju pozitivnih vrijednosti varijable možemo koristiti Box-Cox preslikavanje, čime nova veličina teži normalnoj distribuciji:

$$(E_n)_i^{(\lambda)} = \begin{cases} \frac{(E_n)_i^\lambda - 1}{\lambda} & \text{ako } \lambda \neq 0 \\ \ln[(E_n)_i] & \text{ako } \lambda = 0 \end{cases}, \quad (3.17)$$

gdje se za eksponent $\lambda \in [-5, 5]$ uzima ona vrijednost za koju veličina $E_n^{(\lambda)}$ najbolje aproksimira normalnu razdiobu. U našem slučaju optimalna pronađena vrijednost je $\lambda = 3,78$

Drugi način da riješimo problem raspodjele veličine E_n je konstrukcija novog skupa podataka na temelju onog početnog. Cilj transformacije je smanjenje broja podataka na višim energijama dok broj na nižim ostaje očuvan, a da se pritom ne gube korisne informacije sadržane u njima. Transformacija podataka koju ćemo u tu svrhu koristiti je tzv. vokselizacija. Prostor parametara $(E_n, \Delta E, E)$ podijelit ćemo na podskupove jednake veličine \mathbb{V} – voksele. Svaki voksel je označen vrijednošću $(E_n, \Delta E, E)$ njegove središnje točke u prostoru (parametrizirat ćemo koordinatama njegova središta). Slično kao i u slučaju točkastih podataka, radi jednostavnosti voksele ćemo jedinstveno indeksirati bijekcijom:

$$(E_n, \Delta E, E) \longrightarrow j. \quad (3.18)$$

Pojedinom vokselu pripisat ćemo vrijednost one reakcije \mathcal{P} koja u njemu težinski prevladava te sumu težina istih:

$$W_{\mathcal{P}}(E_n, \Delta E, E) = \sum_{j \in \mathbb{V}} w_{\mathcal{P}, j}. \quad (3.19)$$

Ovakva transformacija, osim rješavanja spomenutog problema, čini podatke vizualno jasnima; učenje mreže višestruko je ubrzano te se dodatno reducira mogućnost prenaučenosti klasifikatora. Sada se može postaviti pitanje zašto ne koristiti vokselizirani prostor kao klasifikator tako da se sve čestice pojedinog vokselu klasificiraju kao one koje u njemu prevladavaju. Takav klasifikator nema mogućnost generalizacije u području gdje nemamo ulaznih podataka.

3.5 Implementacija neuronske mreže

Konceptualno najjednostavnije rješenje problema kategorizacije čestica bilo bi ručno određivanje granica između produkata reakcija u parametarskom prostoru. Problem u tome

je preklapanje granica istih (slika 3.7), tj. granice nisu trivijalne. To povlači da sami podatci parametara $(E_n, \Delta E, E)$ nisu dovoljni za određivanje tipa čestice s potpunom sigurnošću, no takvi su jedini dostupni. Najbolje što možemo napraviti je pronaći optimalnu granicu između skupova. Tip reakcije modeliramo funkcijom u trodimenzionalnom prostoru parametara $(E_n, \Delta E, E)$. Takvu funkciju odredit ćemo alatom strojnog učenja – neuronskom mrežom. Dok je zadatak korisnika pronaći optimalan broj parametara funkcije te tako ostvariti ravnotežu između preciznosti i pouzdanosti klasifikatora, cilj algoritma je pronaći vrijednosti tih parametara.

Kod implementacije neuronske mreže koristila se biblioteka `Tensorflow` u programskom jeziku *Python*.

Za algoritam učenja izabran je stohastički gradijentni spust u kombinaciji s kategoričkom unakrsnom entropijom kao funkcijom cijene. Hiperparametri algoritma nađeni su ručnim podešvanjem: izabrani korak učenja je $\eta = 6$, a veličina uzorka je 128. Na takvim hiperparametrima naučili smo tri klasifikatora (neuronske mreže) različitim podacima: (1) s netransformiranim podacima, (2) s podacima transformiranim Box-Cox preslikavanjem te (3) s vokseliziranim podacima.

Ulazni podatak neuronske mreže je trodimenzionalni vektor. Prvi, ulazni sloj mreže čine tri neutrona čije vrijednosti predstavljaju koordinate vektora. Izlazna vrijednost je također uređena trojka s tri moguće vrijednosti: $(1, 0, 0)$ za proton, $(0, 1, 0)$ za deutron i $(0, 0, 1)$ za triton. Arhitekturu mreže, tj. broj skrivenih slojeva i neutrona u njima testirali smo ručno te potom odabrali optimalnu – dva skrivena sloja po 6 neurona (slika 3.8).

Za mjeru točnosti klasifikatora definirat ćemo količinu točno klasificiranog događaja (stvarno pozitivnog ishoda) $T_{\mathcal{P},i}$ i netočno klasificiranog događaja $F_{\mathcal{P},i}$ (lažno pozitivnog ishoda). Klasificirani događaj doprinijet će količini točno ili netočno klasificiranih događaja ovisno o svojoj težini:

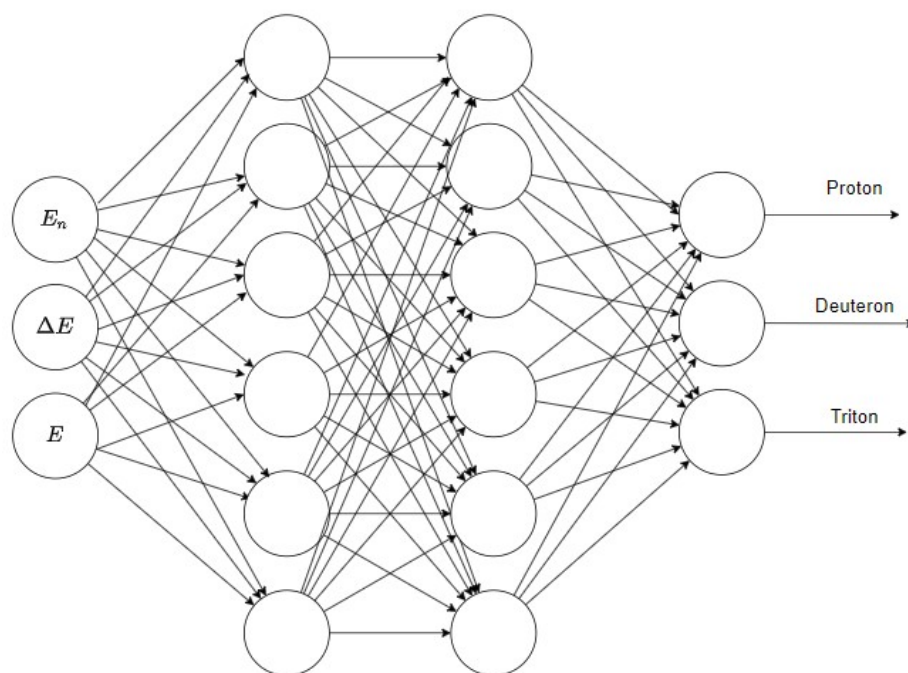
$$T_{\mathcal{P},i} = \begin{cases} w_{\mathcal{P},i} & \text{ako je događaj točno klasificiran} \\ 0 & \text{ako je događaj netočno klasificiran} \end{cases}, \quad (3.20)$$

$$F_{\mathcal{P},i} = \begin{cases} w_{\mathcal{P},i} & \text{ako je događaj netočno klasificiran: } \mathcal{P}' \neq \mathcal{P} \\ 0 & \text{ako je događaj točno klasificiran} \end{cases}. \quad (3.21)$$

Nadalje, kako bismo mogli usporediti točnost klasifikatora na pojedinim energijama, definirat ćemo i udio točno klasificiranih (stopu stvarno pozitivnih) $U_T^{(\mathcal{P})}(\mathbb{E})$ i udio krivo klasificiranih događaja (stopu lažno pozitivnih ishoda) $U_F^{(\mathcal{P})}(\mathbb{E})$ unutar intervala energija neutrona \mathbb{E} :

$$U_T^{(\mathcal{P})}(\mathbb{E}) = \frac{\sum_{i \in \mathbb{E}} T_{\mathcal{P},i}}{\sum_{i \in \mathbb{E}} w_{\mathcal{P},i}}, \quad (3.22)$$

$$U_F^{(\mathcal{P})}(\mathbb{E}) = \frac{\sum_{i \in \mathbb{E}} F_{\mathcal{P},i}}{\sum_{i \in \mathbb{E}} w_{\mathcal{P},i}}. \quad (3.23)$$



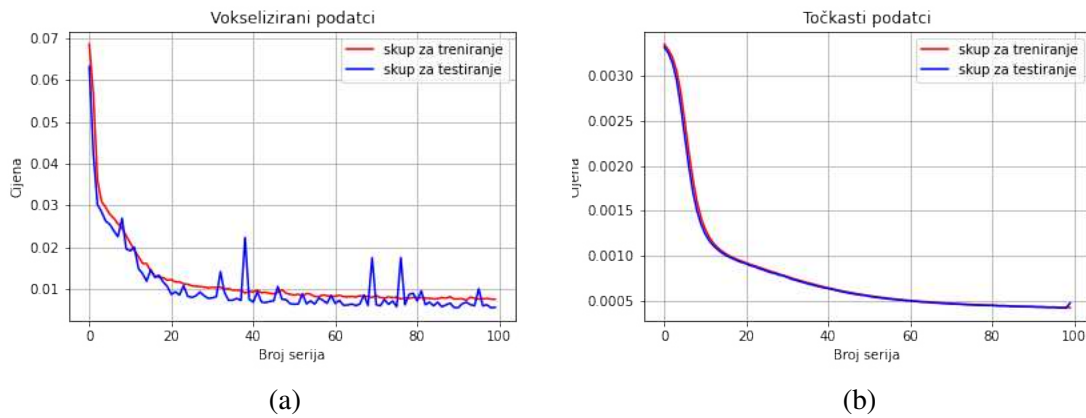
Slika 3.8: Arhitektura optimalne neuronske mreže za kategorizaciju reakcija $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$.

Ovdje u sumi pod $i \in \mathbb{E}$ mislimo na sve događaje čija se E_n vrijednost nalazi u zadanom intervalu \mathbb{E} . U slučaju da su podatci vokselizirani, u zapisu prethodnih definicija koristimo indekse j umjesto i . Također, koristimo težine vokselu $W_{\mathcal{P}}$ umjesto dosadašnjih težina točkastih podataka $w_{\mathcal{P}}$.

Radi lakše analize učenja mreže i uspoređivanja različitih arhitektura mreže definirat ćemo i ukupnu točnost neuronske mreže. Takva će uzeti u razmatranje sve reakcije \mathcal{P} na svim energijama E_n :

$$U = \frac{\sum_i T_{\mathcal{P},i}}{\sum_i w_{\mathcal{P},i}}. \quad (3.24)$$

Brzina učenja mreže nad točkastim podacima tekla je nešto sporije nego ona nad vokseliziranim podacima (slika 3.9), a potonja je počela konvergirati već nakon 20 serija učenja. Kako je učenje vremenski veoma kratko (nekoliko minuta) te nas neće ograničavati, broj serija učenja će biti 80 s korakom učenja $\eta = 6$ te dodatnih 20 sa smanjenim korakom učenja $\eta = 1$. Sve tri mreže postizale su slične rezultate maksimalne točnosti. Već u prvim pokušajima izbora arhitekture vidimo iznenađujuće visoku točnost. Mreža sa samo jednim skrivenim slojem tri neutrona postiže ukupnu točnost veću od 80%. No također, na takvoj mreži vidimo da točnost skupa za testiranje nije stabilna, stoga zaključujemo da

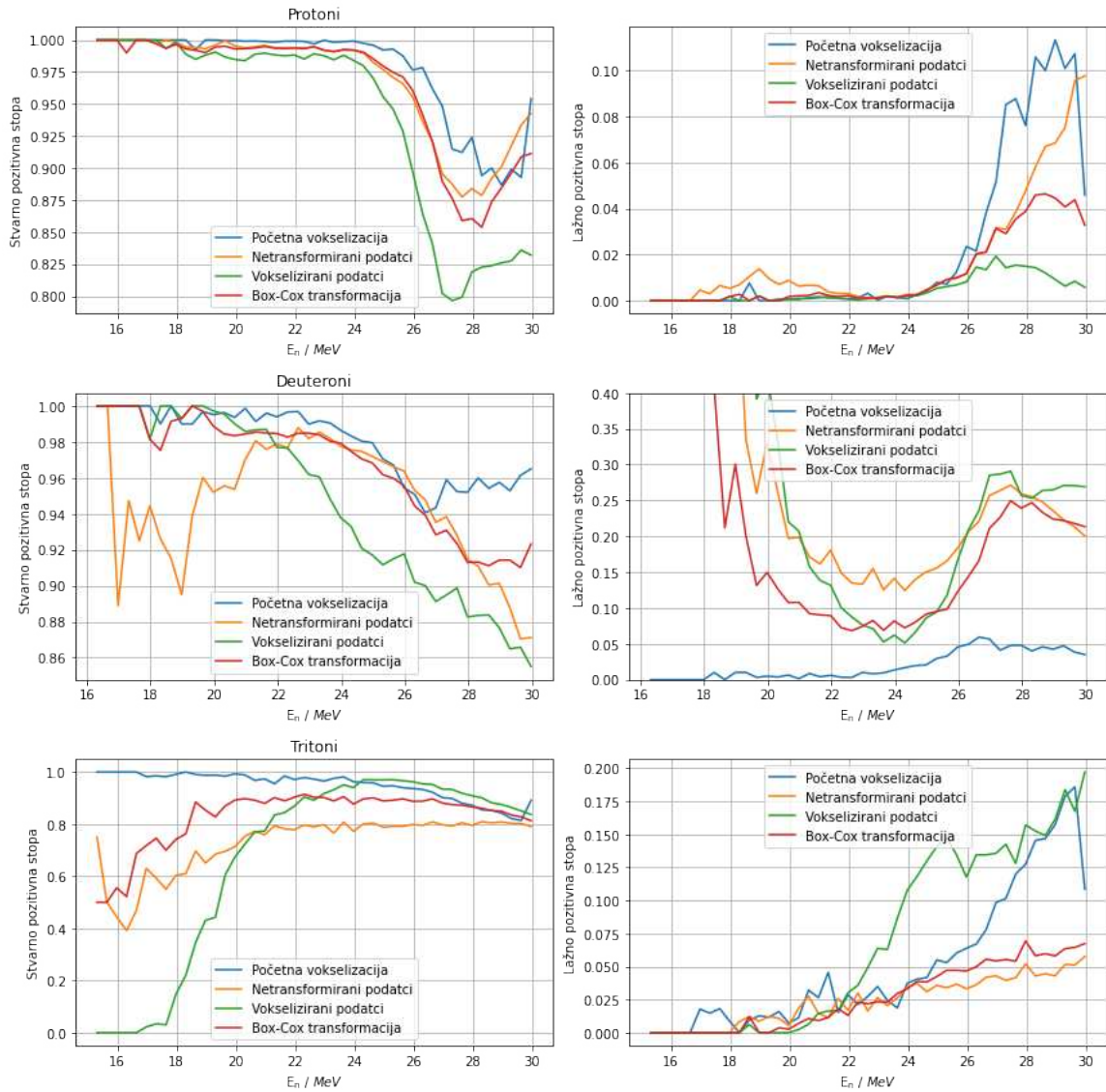


Slika 3.9: Brzina učenja prikazana funkcijom cijene. a) Funkcija cijene na vokseliziranim podacima. b) Funkcija cijene na točkastim podacima.

se radi o krivom modelu. Povećavanjem broja neurona i slojeva povećava se i konačna točnost mreže, no već nakon 15 neurona u svakom od dva sloja mreže rast točnosti zanemarivo raste. Konačni je cilj pronaći arhitekturu mreže sa što manje parametara čime ćemo smanjiti mogućnost prenaučivosti mreže. Smanjivanjem broja neurona u svakom sloju te pomnijom provjerom zaključujemo da je ona s dva skrivena sloja po šest neurona optimalna opcija gdje je točnost od 93% očuvana.

Na slici 3.10 prikazani su grafovi udjela točno i netočno klasificiranih događaja za pojedini klasifikator po energijama neutrona E_n . Kod klasifikatora općenito želimo vidjeti što veću stopu stvarno pozitivnih te što manju stopu lažno pozitivnih ishoda. Napomenimo da tritoni prikazani u trećem stupcu nisu razmatrani u mjerenjima [1], no u ovom će radu pomoći kod usporedbe različitih klasifikatora. U grafu je prikazana i točnost početne vokselizacije koja razmatra postotak pravilno raspoređenih čestica po klasificiranim vokselima. Točnosti početne vokselizacije maksimalna je točnost koju naučeni klasifikator može doseći, no ne želimo da se točnost naučenog klasifikatora njoj značajno približi zbog mogućnosti prenaučivosti.

Na slici 3.10 možemo vidjeti da na energiji $E_n \approx 25$ MeV imamo nagli pad točnosti. To smo mogli i očekivati, s obzirom na to da se događaji na većim energijama značajno više preklapaju u parametarskom prostoru (slika 3.7). Takav rezultat potvrđuje pretpostavke u članku [1] da na području od energije praga do 25 MeV možemo s značajnom sigurnošću identificirati $^{12}\text{C}(n, p)$ i $^{12}\text{C}(n, d)$ reakcije. Uspoređujući različite klasifikatore vidimo da kod klasifikatora naučen izvornim podacima ima nisku točnost na nižim energijama u usporedbi s ostalim klasifikatorima. To je upravo bio i jedan od ciljeva vokselizacije i Box-Cox transformacije. Primijetimo nizak broj točno klasificiranih tritona na području energija ispod 20 MeV. Takav rezultat nije posljedica lošeg klasifikatora, već na tom po-



Slika 3.10: Stope stvarno pozitivnih (lijevi stupac) i lažno pozitivnih ishoda (desni stupac) po pojedinoj reakciji. Na svakom grafu uspoređeni su različiti klasifikatori: početna vokselizacija, mreže naučene netransformiranim točkastim podatcima, podatcima transformiranim Box-Cox metodom i vokseliziranim podatcima. U gornjem redu vidimo klasifikaciju protona, srednjem redu deuterona te u donjem redu tritona.

dručju nalazimo zanemariv broj tritona koje su nastale u reakcijama nad izotopom ^{13}C . Zbog istog razloga na tom području energija vidimo i mali broj krivo klasificiranih tritona. Znatno broj krivo klasificiranih $^{12}\text{C}(n, d)^{10}\text{B}$ reakcija pojavljuje se ispod 20 MeV i iznad 26 MeV gdje udio krivo pozitivnih reakcija raste iznad 10%. Općenito možemo vidjeti da Box-Cox transformacija ima najbolje rezultate i u točno i u krivo klasificiranim događajima s obzirom na ostale klasifikatore.

Poglavlje 4

Zaključak

Analiza reakcija $^{12}\text{C}(n,p)$ i $^{12}\text{C}(n,d)$ po energijama neutrona zahtijeva klasifikaciju nuklearnih reakcija. Takva klasifikacija omogućena je informacijom o energijskom tragu produkta reakcije E i ΔE te pripadnom energijom neutrona reakcije E_n . Reakcije na izotopima prirodnog ugljika ^{12}C i ^{13}C koje smo uzeli u razmatranje su (n,p), (n,d) i (n,t). Za metodu klasifikacije izabrali smo alat strojnog učenja – acikličku neuronsku mrežu. Podatke kojima smo učili neuronsku mrežu dobili smo GEANT4 simulacijama. Produkti reakcija u simulacijama direktno su generirani, gdje smo njihove parametre odredili klasičnim dinamičkim računom preko varijabli energije neutrona E_n , Q-vrijednosti reakcije i energijske razine jezgre kćeri E_x . Razdiobe varijabli E_n i E_x pretpostavili smo kao uniformne, a kutnu razdiobu produkata reakcije kao izotropnu u sustavu centra mase. Efikasnost neuronske mreže uspoređivali smo s efikasnošću primitivnog klasifikatora – vokseliziranog prostora: prostor parametara ($E_n, \Delta E, E$) podijelili smo na podskupove jednake veličine – voksele. E_n -širina svakog vokselu je 0,33 MeV, ΔE -širina je 0,05 MeV i E -širina 0,05 MeV. Svakom pojedinom vokselu pridodali smo vrijednost one reakcije koja u njemu prevladava. Odredili smo optimalnu arhitekturu neuronske mreže kao onu s dva skrivena sloja po šest neurona. Efikasnost neuronske mreže definirali smo preko broja točno i netočno klasificiranih reakcija (stvarno i lažno pozitivnih ishoda). Provjera efikasnosti vršili smo za svaku pojedinu reakciju zasebno, a usporedili smo neuronske mreže naučene na: netransformiranim podacima, vokseliziranim podacima i podacima transformiranim Box-Cox transformacijom. Zaključili smo da je Box-Cox optimalna transformacija podataka za učenje. Nijedna transformacija podataka nije omogućila povećanje efikasnosti neuronske mreže iznad 25 MeV, dok je na području od praga reakcije do 25 MeV ostvarena značajna mogućnost klasifikacije. Na tim energijama Box-Cox klasifikator za reakcije $^{12}\text{C}(n,p)^{11}\text{B}$ i $^{12}\text{C}(n,d)^{10}\text{B}$ postiže udio točno klasificiranih iznad 98%. Znatno broj krivo klasificiranih $^{12}\text{C}(n,d)^{10}\text{B}$ reakcija imamo ispod 20 MeV i iznad 26 MeV; na ovom području udio krivo pozitivnih reakcija raste iznad 10%.

Bibliografija

- [1] Žugec, P. et al.: *Measurement of the energy-differential cross section of the $^{12}\text{C}(n,p)$ and $^{12}\text{C}(n,d)$ reactions*, CERN-INTC-2017-081, INTC-P-522 (2017), <http://cds.cern.ch/record/2287421>
- [2] n_TOF, Collaboration: *The neutron Time-Of-Flight facility, n_TOF, at CERN (I): Technical Description* (2013), <https://cds.cern.ch/record/1514680>
- [3] Barbagallo, M. et al.: *High-accuracy determination of the neutron flux at n_TOF*, European physical journal e 1292-8941 (2013)
- [4] Emsley, J.: *Nature's Building Blocks: An A-Z Guide to the Elements* OUP Oxford (2011) <https://books.google.hr/books?id=2EfYXzwPo3UC>
- [5] Žugec, P. et al.: *Integral measurement of the $^{12}\text{C}(n,p)^{12}\text{B}$ reaction up to 10 GeV*, The European Physical Journal A 1434-601X (2016) <https://doi.org/10.1140/epja/i2016-16101-7>
- [6] Žugec, P. et al.: *Measurement of the $^{12}\text{C}(n,p)^{12}\text{B}$ cross section at n_TOF at CERN by in-beam activation analysis*, American Physical Society (2014) <https://link.aps.org/doi/10.1103/PhysRevC.90.021601>
- [7] Žugec, P. et al.: *Study of a data analysis method for the angle resolving silicon telescope*, Journal of Instrumentation (2020) <https://doi.org/10.1088/1748-0221/15/02/p02011>
- [8] *Overfitting*, <https://blog.ineuron.ai/Overfitting-8ikUK4gUAK>
- [9] J.H. Kelley and J.E. Purcell and C.G. Sheu: *Energy levels of light nuclei A=12*, Nuclear Physics A 0375-9474 (2017) <https://www.sciencedirect.com/science/article/pii/S0375947417303330>

Sažetak

Cilj ovog diplomskog rada jest usporediti utjecaj različitih transformacija podataka na efektivnost tj. točnost klasifikatora reakcija $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$. Kao klasifikator koristit ćemo neuronsku mrežu. Kontekst ovog rada je eksperiment proveden u n_TOF postrojenju na CERN-u. Tijekom obrade eksperimentalnih podataka rekonstruirat će se udarni presjek spomenutih reakcije preko podataka mjerenih silicijskim teleskopom. Mjerenje koje se temelji na $\Delta E - E$ tehnici simulirali smo u GEANT4 programskom paketu. Simulirane podatke obradili smo i transformirali Box-Cox transformacijom i vokselizacijom. Neuronske mreže implementirali smo na izvornim i transformiranim podacima. Box-Cox transformacija optimalna je u ovom kontekstu. Svim mrežama točnost naglo pada na energijama neutrona iznad 25 MeV. Takav zaključak u skladu je sa preliminarnim očekivanjima iz prijedloga mjerenja $^{12}\text{C}(n, p)^{12}\text{B}$ i $^{12}\text{C}(n, d)^{11}\text{B}$ reakcija na n_TOF postrojenju.

Summary

In this thesis we aim to compare the effectiveness of various transformation data on the accuracy of the $^{12}\text{C}(n, p)^{12}\text{B}$ and $^{12}\text{C}(n, d)^{11}\text{B}$ reaction classifier – the neural network. The base of our paper is an experiment carried out in the n_TOF facility at CERN. In the experiment, the cross sections of the mentioned reactions are reconstructed using the data measured by a silicon telescope. We simulated the measurement, which is based on the $\Delta E - E$ technique, in the GEANT4 software. We analyzed and transformed the simulated data using Box-Cox transformation and voxelization. The study also implemented neural networks on non-transformed data, along with implementation on each of the transformations. We conclude that the Box-Cox transformation is optimal in this context. For all networks, the accuracy drops sharply above 25 MeV of input neutron energy, confirming the previously set assumptions in the proposal of $^{12}\text{C}(n, p)^{12}\text{B}$ and $^{12}\text{C}(n, d)^{11}\text{B}$ reaction measurement.

Životopis

Rođen sam u Stonu na poluotoku pelješcu 1997. godine, simbolički netom poslije velikog potresa 1996. Moja majka je dio odgoja prepustila mojoj 8 godina starijoj sestri. Tako, na sestrinu žalost, postajem centar njezinog ženskog društva čime razvijam samopouzdan karakter. Osnovnu školu završio sam u rodnom mjestu te upisujem smjer Elektrotehnika u Pomorsko-tehničkoj školi Dubrovnik. 2016. završavam srednju školu i upisujem Nastavnički studij Matematike i fizike na Prirodoslovno-matematičkom fakultetu u Zagrebu. Od prvog dana srednje škole do zadnjeg studentskog dana živim u domu: prvo u Đačkom domu Dubrovnik, potom u Studentskom domu Stjepan Radić. Kroz školske i studentske godine naučio sam dijeliti prostor i vrijeme s drugima čime sam stekao brojna prijateljstva i uspomene koje će me pratiti kroz cijeli život. Majka me je od najmlađih dana poticala u razvijanju njezinih umjetničkih gena. Tako završavam osnovnu glazbenu školu za trombon te sviram u Glazbi Ston. U KUD-u Ponikovska poskočica plešem linđo, sviram lijericu i nastupam kao zdravičar i amaterski glumac. Kroz cijeli studij član sam studentske klape Banderija s kojom 2022. nastupam u finalu Festivala dalmatinskih klapa Omiš.