

Primjena neuronskih mreža u modeliranju svojstava atomskih jezgara

Šura, Krešimir

Master's thesis / Diplomski rad

2023

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:217:519531>

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

Download date / Datum preuzimanja: **2025-01-04**



Repository / Repozitorij:

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



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Krešimir Šura

PRIMJENA NEURONSKIH MREŽA U
MODELIRANJU SVOJSTAVA ATOMSKIH
JEZGARA

Diplomski rad

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

INTEGRIRANI PREDDIPLOMSKI I DIPLOMSKI SVEUČILIŠNI STUDIJ
FIZIKA; SMJER ISTRAŽIVAČKI

Krešimir Šura

Diplomski rad

**Primjena neuronskih mreža u
modeliranju svojstava atomskih
jezgara**

Voditelj diplomskog rada: prof. dr. sc. Nils Paar

Ocjena diplomskog rada: _____

Povjerenstvo: 1. _____

2. _____

3. _____

Datum polaganja: _____

Zagreb, 2023.

Hvala mentoru, prof. Paaru na prijedlogu teme i pomoći oko diplomskog rada.

Hvala kolegama, prijateljima i obitelji na podršci tokom studija.

Sažetak

Energija vezanja temeljno je svojstvo atomske jezgre. Može se mjeriti eksperimentalno, ali ne za sve jezgre koje su potrebne za modeliranje nukleosinteze. Razvijeni su mnogi teorijski modeli koji opisuju energije vezanja kroz cijelu mapu nuklida, a najbolji među njima postižu srednje kvadratno odstupanje RMS od eksperimentalnih vrijednosti od oko 500 keV-a. U ovome radu proučen je pristup određivanja energije vezanja atomskih jezgara pomoću strojno naučenih neuronskih mreža koristeći eksperimentalne podatke iz baze nuklearnih podataka AME2020. Prediktivna moć strojno naučenih modela uspoređena je s eksperimentalnim podacima, kao i s drugim metodama određivanja energija vezanja. Ključni korak razvoja neuronskih mreža za primjenu u nuklearnim istraživanjima je odabir značajki modela, kojima unosimo relevantne fizikalne informacije u proces strojnog učenja. Neuronska mreža s deset odabranih značajki postigla je $RMS = 248$ keV. Ovakav rezultat bolji je od tradicionalnih modela nuklearne fizike.

Ključne riječi: nuklearna struktura, model energija vezanja atomske jezgre, strojno učenje, neuronske mreže, duboko učenje, odabir značajki, fizikalno informirane neuronske mreže

Application of neural networks in modelling the properties of atomic nuclei

Abstract

Binding energy of an atomic nucleus is one of its fundamental properties. It can be measured in experiment, however not for all nuclei necessary in nucleosynthesis modeling applications. Many theoretical models that calculate binding energies for all nuclei have been developed, the best of which describe binding energies with a root mean square RMS deviation from experimental data of 500 keV. In this work a novel approach to modelling binding energies has been investigated: machine learning using neural networks utilizing experimental data from the AME2020 mass evaluation. Predictive power of machine learning models has been compared with experiment, as well as other nuclear models. Crucial step of developing neural networks for use in nuclear physics has been feature engineering, a way to introduce and encode physically important information in the process of machine learning. Best performing neural network had ten selected features, achieving an $\text{RMS} = 248 \text{ keV}$, a result surpassing traditional nuclear physics models.

Keywords: nuclear structure, nuclear models ,binding energy, machine learning, neural networks, deep learning, feature engineering, physics informed neural networks

Sadržaj

1	Uvod	1
1.1	Nuklearna struktura	2
1.1.1	Model kapljice	4
1.1.2	Weizsacker-ova formula mase	4
1.1.3	Model ljusaka	7
1.1.4	Nukleonski faktor promiskuiteta	8
1.1.5	Efekti sparivanja	9
1.1.6	Wignerova energija	10
2	Metode	12
2.1	Strojno učenje	12
2.2	Neuronske mreže	13
2.2.1	Aktivacijske funkcije	14
2.3	Šum	16
2.4	Odabir modela	18
3	Rezultati	22
3.1	ANN2	27
3.2	ANN4	28
3.3	ANN5	32
3.4	ANN7	34
3.5	DNN7	37
3.6	DNN8	40
3.7	DNN9	44
3.8	DNN10	48
4	Zaključak	59
	Literatura	61

1 Uvod

Atomske jezgre složeni su mnogočestični kvantni sustavi, sastavljeni od nukleona - protona i neutrona. Nukleoni su i sami složeni sustavi kvarkova vezanih jakim nuklearnom silom. Ostatak jake interakcije odgovorne za hadronizaciju kvarkova, odgovoran je za vezanje samih nukleona u jezgri. Stvaranje atomskih jezgara je egzoterman proces te se prilikom vezanja protona i neutrona oslobađa energija vezanja (E_B) [1]. Ona je energija koja je potrebna kako bi se svi sastavni nukleoni jezgre iz nje izdvojili, razlika masa slobodnih nukleona i jezgre vodika,

$$E_B(Z, N) = Zm_H + Nm_n - m_{atom}. \quad (1.1)$$

Pouzdani modeli atomskih jezgara kao i kvalitetna mjerenja njihovih svojstava potrebni su kako bi se produbilo shvaćanje prirode nukleon-nukleon (N-N) interakcija. Precizne vrijednosti energija vezanja, vremena poluživota, radijusa jezgara i drugih opservabli vrlo su važna za mnoga područja istraživanja. Energije vezanja do sada neopaženih jezgara koje teoretski mogu postojati, od onih bogatim neutronima do onih bogatim protonima, bitne su u astrofizici, istraživanjima stvaranja teških elemenata r-procesom [2] i u eksplozijama supernova te istraživanjima svojstava neutronskih zvijezda [3]. Od interesa je i potraga za novim, superteškim, kemijskim elementima [4]. Nuklearna svojstva poput energija vezanja i udarnih presjeka od interesa su i u nuklearnom inženjerstvu [5].

U ovom radu energije vezanja atomskih jezgara modelirane su strojno naučenim umjetnim neuronskim mrežama (engl. *artificial neural network* - ANN). Koristeći se dostupnim eksperimentalnim podacima trenirane su plitke i duboke neuronske mreže s fizikalno relevantnim ulaznim varijablama (engl. *physically informed neural networks* - PINN) [6]. Neuronske mreže iznimno su korisne za rješavanje problema regresije. Uz dovoljno veliki skup podataka za učenje i dobro izabrane ulazne varijable ANN bi savršeno mogao reproducirati podatke. U realnosti smo ograničeni dostupnim eksperimentalnim podacima, a kako ne znamo točnu prirodu interakcija u atomskim jezgrama možemo empirijski potražiti najbolji skup ulaznih varijabli.

Za početak će se razviti bitni fizikalni koncepti, krucijalni za konstrukciju neuronskih mreža. Ukratko će se uvesti strojno učenje te proći kroz korištene metode i alate. Konačno, bit će predstavljene neke od najbolje istreniranih neuronskih mreža

te prodiskutirana njihova učinkovitost, prednosti i nedostaci. Dobiveni rezultati biti će uspoređeni s drugim dostupnim modelima.

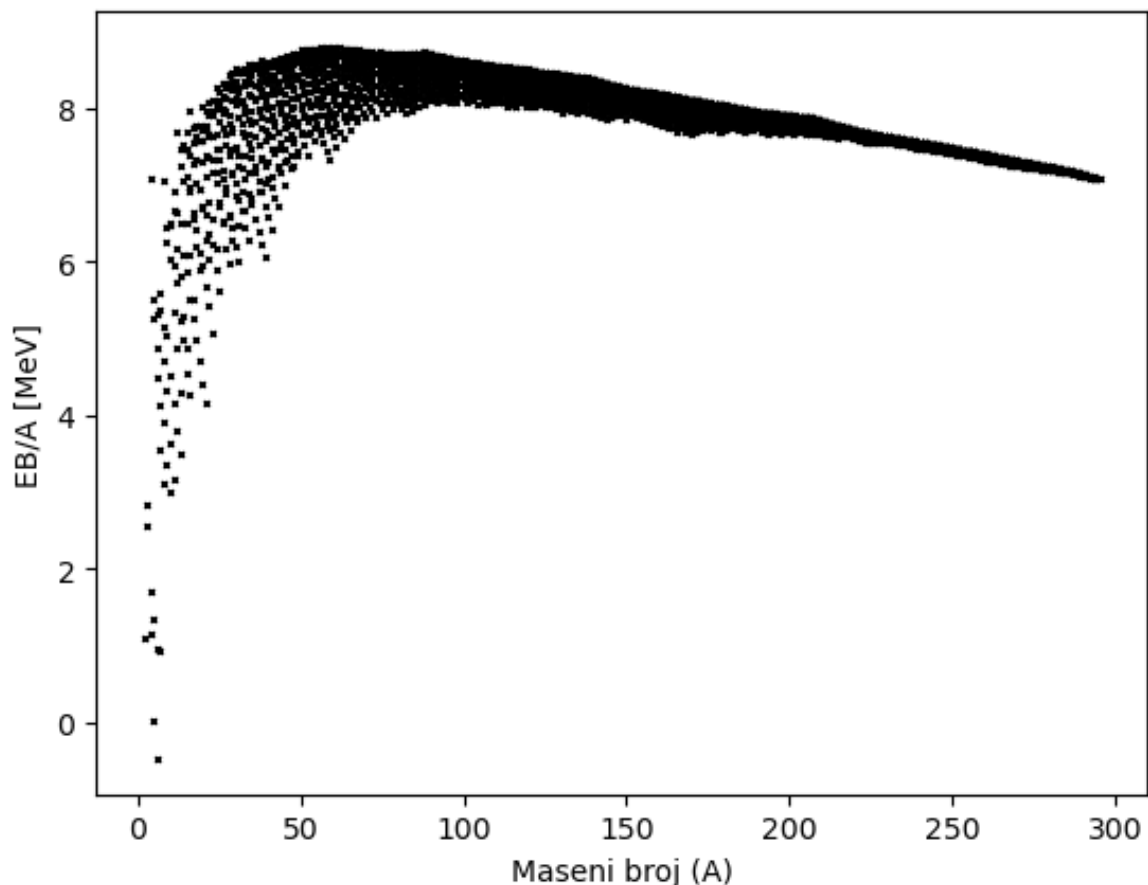
1.1 Nuklearna struktura

Poznato je da N-N međudjelovanje ovisi o udaljenosti, spinu, orbitalnom momentu, izospinu. Prisutne su dvočestične ali i tročestične interakcije, kao i efekti sparivanja [7]. Na skalama energija atomskih jezgara problemu nije moguće pristupiti perturbativnim računom kvantne kromodinamike (QCD). Čak i najjednostavniji nuklearni sustav - deuteron - ima 16 unutarnjih stupnjeva slobode (svaki nukleon ima spin i izospin), nuklearni potencijal također sadrži 16 članova [1], za rješavanje ovako složenog problema potrebno je mnogo računalne snage.

Do danas su razvijene mnoge metode rješavanja sustava atomskih jezgara, *ab initio* metode [13], Monte-Carlo varijacijske metode [8], model kapljice, parametrizirane Skyrme sile [9], model ljusaka, koji dobro opisuju jezgre duž doline stabilnosti, metode samosuglasnog srednjeg polja poput Hartree-Fock (HF) i Hartree-Fock-Bogoljubov (HFB) [10], relativistička aproksimacija srednjeg polja [11], koje daju bolje rezultate za teže jezgre $Z > 20$, te algebarski modeli poput modela interagirajućih bozona (engl. *interacting boson model* - IBM) [12]. Spomenute metode oslanjaju se na numeričko rješavanje problema dijagonalizacijom nuklearnog hamiltonijana te iteriranjem do konvergencije. Pojedine metode izvrsne su za specifične slučajeve, ali od eksperimenta odstupaju za ostale. Različite nuklearne modele možemo uspoređivati srednjim kvadratnim odstupanjem izračunatih vrijednosti od eksperimentalno poznatih vrijednosti, RMS (engl. *root-mean-square*),

$$\text{RMS} = \sqrt{\sum_i^N \frac{(E_{Bi} - E_{Bi}^{EXP})^2}{N}}. \quad (1.2)$$

Sustavi lakih jezgara su riješeni *ab initio* metodama i efektivnom teorijom polja, no takav pristup je previše zahtjevan za opis težih jezgara [13]. Za lake jezgre energija vezanja po nukleonu E_B/A raste linearno s brojem nukleona zbog nuklearne sile, dok za teže jezgre, zbog odbojne električne sile između protona, ona otprilike linearno opada. Graf energije vezanja po nukleonu prikazan je na slici 1.1. Maksimum E_B po nukleonu postiže se za maseni broj 56, elemente željezo Fe i nikal Ni te iznosi oko 8.75 MeV.

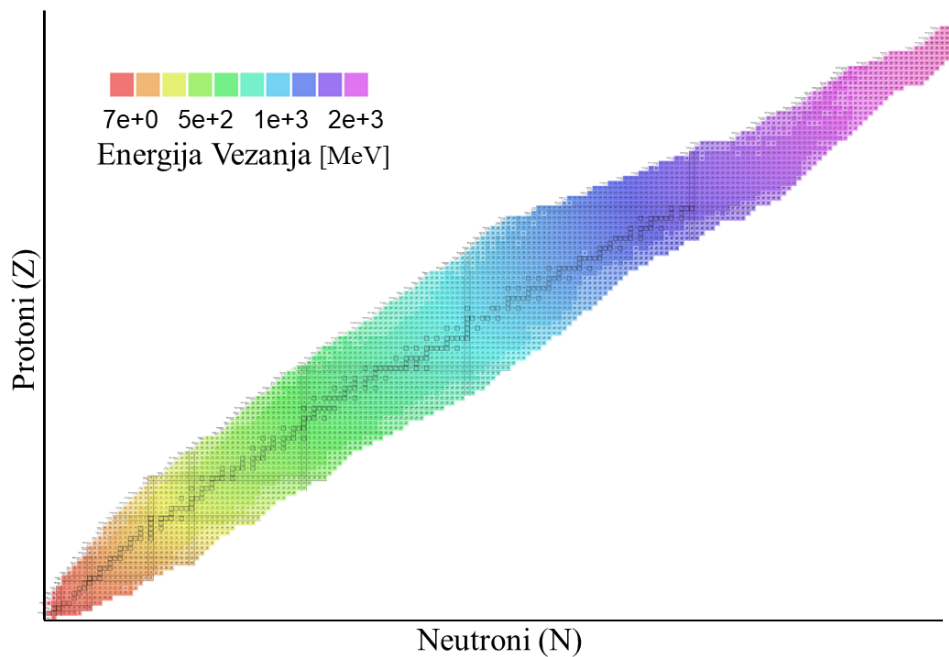


Slika 1.1: Graf energija vezanja po nukleonu

Slika 1.2 je prikaz mape nuklida s eksperimentalnim podacima iz baze masa atomskih jezgara AME2020 [14, 15]. Na slici su naznačene linije za brojeve protona i neutrona koje označavaju na 2, 8, 20, 28, 50, 82, 126, takozvane *magične brojeve*. Stabilne jezgre su posebno istaknute. Vidljivo je da se lakše stabilne jezgre nalaze oko $N = Z$ linije, teže su jezgre bogatije neutronima kako bi se N-N sila suprotstavila električnoj. Stabilne jezgre nalaze se u tzv. *dolini stabilnosti* u mapi nuklida. Oko magičnih brojeva protona i neutrona, atomske jezgre su posebno stabilne. Jezgre izvan doline stabilnosti imaju sve kraća vremena poluživota i nuklearnim raspadima se gibaju prema dolini stabilnosti. Upravo te nestabilne jezgre izvan doline stabilnosti su posebno bitne u astrofizičkim procesima i kozmologiji kako bi se razumjelo podrijetlo težih elemenata u svemiru [2].

Daleko od doline stabilnosti, u području velikog broja nukleona, pretpostavlja se da postoje *otoci stabilnosti* na kojima bismo mogli pronaći, do sada, neotkrivene jezgre kemijskih elemenata [4].

U AME2020 bazi podataka nalaze se informacije za 3827 danas poznatih atom-



Slika 1.2: Mapa nuklida, broj neutrona N raste duž x-osi, broj protona Z raste duž y-osi, nijanse crvene označavaju jezgre niže energije vezanja, nijanse plave više.

skih jezgara 118 kemijskih elemenata od kojih je samo 251 stabilna. Pretpostavlja se da bi u svemiru moglo postojati oko 10000 atomskih jezgara.

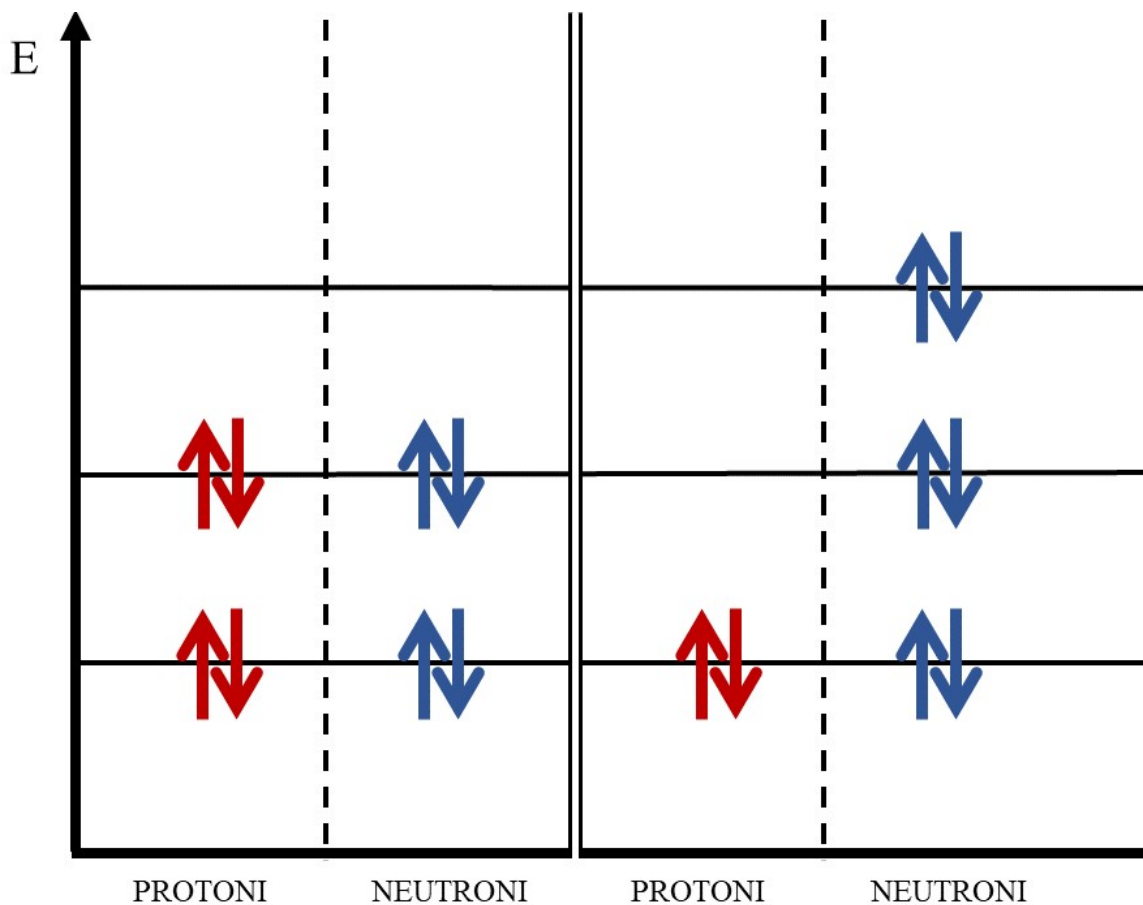
1.1.1 Model kapljice

Energija vezanja E_B je, više-manje, glatka funkcija broja nukleona (A) i makroskopskih stupnjeva slobode. Razvijeni su mnogi semi-empirijski modeli koji E_B izražavaju kao funkciju broja protona (Z), broja neutrona (N) i masenog broja (A). Ovakvi modeli nisu direktno vezani s nukleon-nukleon interakcijom već su određeni parametrima namještenim prema eksperimentalnim podacima [1].

1.1.2 Weizsacker-ova formula mase

Atomsku jezgru tretiramo kao kapljicu nestlačive tekućine, njen volumen raste s brojem nukleona. Energija vezanja također raste s volumenom jezgre, odnosno brojem nukleona. Analogno kapljici tekućine, nukleoni na površini jezgre slabije su vezani od onih u unutrašnjosti jer imaju manje susjeda i time manje interakcija zbog ograničenog dometa nuklearne interakcije. Energija vezanja, stoga je manja što je jezgra veće površine. Potrebno je uvesti i korekciju radi odbojne električne interakcije između protona. Iako atomske jezgre općenito nisu sferno simetrične, u gruboj aproksimaciji možemo jezgru tretirati kao uniformo nabijenu sferu kako bi dobili

jednostavnu ovisnost o Z , N i A . Energija vezanja umanjena je za elektrostatsku energiju nabijene sfere radijusa $R = 1.2A^{1/3}$. Stabilne jezgre teže prema ravnoteži broja protona i neutrona $N \approx Z$, ovaj fenomen posljedica je kvantnog efekta Paulije- vovog principa isključenja koji u istom energetsom stanju dopušta samo dva protona ili neurona suprotnih spinova. Ishod takvog uvjeta popunjavanja energetskih stanja je da su jezgre jednakog masenog broja za koje vrijedi $Z = N$ jače vezane. Kada bi- smo zamijenili protone za neutrone, ili obrnuto, morali bismo ih staviti u stanje više energije. Razmak između energetskih stanja smanjuje se s brojem nukleona, skica ovog efekta prikazana je na slici 1.3.



Slika 1.3: Skica prikaza Paulije- vovog principa isključenja i njegovog utjecaja na ener- giju vezanja atomske jezgre. Kada bismo zamijenili protone za neutrone, ili obrnuto, a da pri tome ostanemo na istom masenom broju, morali bismo ih staviti u stanje više energije zbog ograničenja kvantne mehanike. Ovo smanjuje energiju vezanja za jezgre koje imaju nesimetrične strukture protona i neutrona.

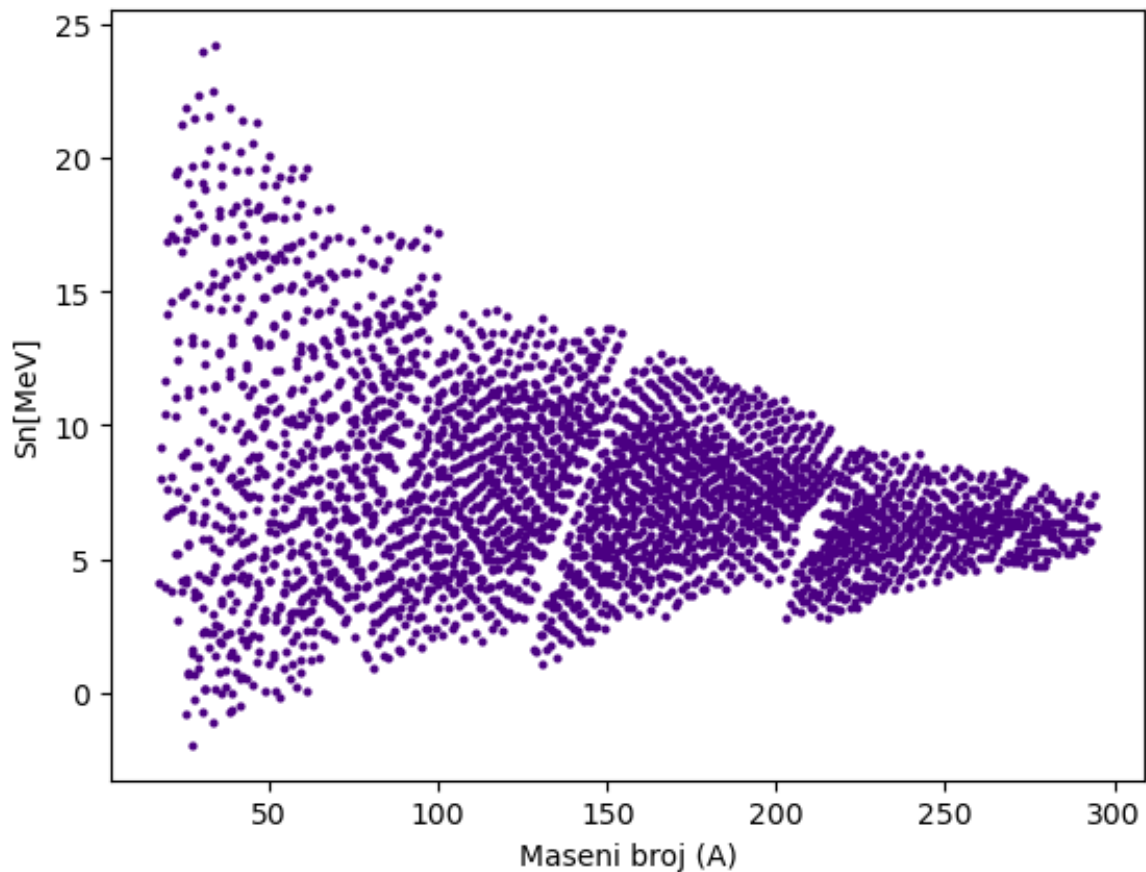
Konačno, dobije se izraz koji se naziva semi-empirijska formula mase

$$E_B(Z, N, A) = c_{vol}A - c_{sur}A^{2/3} - c_{el}\frac{Z(Z-1)}{A^{1/3}} - c_{asy}\frac{(Z-N)^2}{A}. \quad (1.3)$$

Konstante u izrazu (1.3) određene su namještanjem slobodnih parametara prema eksperimentu. U atomskim jezgrama prisutne su i interakcija sparivanja koje su razmotrene u kasnijem poglavlju.

Energije vezanja računane semi-empirijskom formulom generalno prate eksperimentalne podatke, no primjećuju se odstupanja za neke jezgre koje su jače vezane od susjednih. Ovi skokovi E_B nalaze se oko popunjenih ljusaka. Efekt ljusaka može se vidjeti i u energijama separacije neutrona iz jezgre - Slika 1.4. Energije separacije izračunate su iz eksperimentalnih podataka u bazi AME2020, koristeći :

$$S_n = E_B(Z, N) - E_B(Z, N - 1). \quad (1.4)$$



Slika 1.4: Graf energija separacija jednog neutrona iz jezgre kao funkcija masenog broja A . Na grafu se vide diskontinuiteti pri brojevima neutrona 28, 50, 82 koji odgovaraju popunjenim ljuskama.

1.1.3 Model ljusaka

Jezgra je skup A nerelativističkih čestica bez unutarnje strukture. Ovakva aproksimacija izvrsna je za proučavanje nuklearne strukture pri niskim energijama [16]. Schrödingerova jednačba egzaktno je rješiva za $A < 10$. Kako bismo riješili složenije mnogočestične probleme elegantno rješenje je aproksimacija srednjeg polja pri kojoj tretiramo svaki nukleon kao da se nalazi u vanjskom potencijalu kojeg stvaraju ostali nukleoni

$$H = T + V = T + V_{MF} + V_{RES}. \quad (1.5)$$

V_{MF} je efektivni potencijal srednjeg polja, a V_{RES} je rezidualna interakcija, koja je puno manja od početne interakcije V ($V_{RES} \ll V$). Rješenje mnogočestičnog problema sada je svedeno na rješavanje jednočestičnog gibanja u vanjskom potencijalu.

Efektivni potencijal koji dobro opisuje sustav je Woods-Saxon [16],

$$V(r) = -\frac{V_0}{1 + \exp\left(\frac{r-R_0}{a}\right)}, \quad (1.6)$$

no za njega Schrödingerova jednačba nije analitički rješiva. Drugi dobar odabir je potencijal harmoničkog oscilatora $V(r) = m\omega^2 r^2/2$, uz takav odabir radijalni dio Schrödingerove jednačbe ima oblik:

$$-\frac{\hbar^2}{2m} \frac{d^2 u}{dr^2} + \left(\frac{l(l+1)\hbar^2}{2mr^2} + \frac{m\omega^2 r^2}{2} \right) u(r) = Eu(r). \quad (1.7)$$

Uz uvjet da valne funkcije iščezavaju u beskonačnosti, pripadne svojstvene energije sustava su:

$$E = \hbar\omega \left(N + \frac{3}{2} \right) = \hbar\omega \left(2k + l + \frac{3}{2} \right). \quad (1.8)$$

N je glavni kvantni broj i sva stanja s jednakim N su degenerirana te imaju jednake energije, l je orbitalni kvantni broj, a k broj čvorova valne funkcije. Ovakav potencijal reproducira magične brojeve za niže ljuske : 2, 8, 20. Ostale brojeve dobijemo uključivanjem rezidualne spin-orbit interakcije, koja je relativistički efekt. Spin-orbit uvodimo kao perturbaciju u sustav te vidimo da ona cijepa neperturbirana stanja na dva za svaki l . Sada imamo nešto drugačiju strukturu ljusaka koja odgo-

vara eksperimentalnim opažanjima i reproducira sve magične brojeve: 2, 8, 20, 28, 50, 82, 126 - Slika 1.5. Atomske jezgre s magičnim brojem protona i(li) neutrona imaju popunjene ljuske, sferno su simetrične i stabilnije. Njihova energija vezanja je veća. Zbog električnog odbijanja među protonima atomske jezgre teže od kalcija ^{40}Ca moraju imati više neutrona kako bi jezgra ostala vezana. Iz tog razloga manje je "dvostruko magičnih" jezgara s više protona, iako se primjećuje više stabilnih izotopa za jezgre s magičnim brojem protona, poput kositra Sn ($Z = 50$) i olova Pb ($Z = 82$).

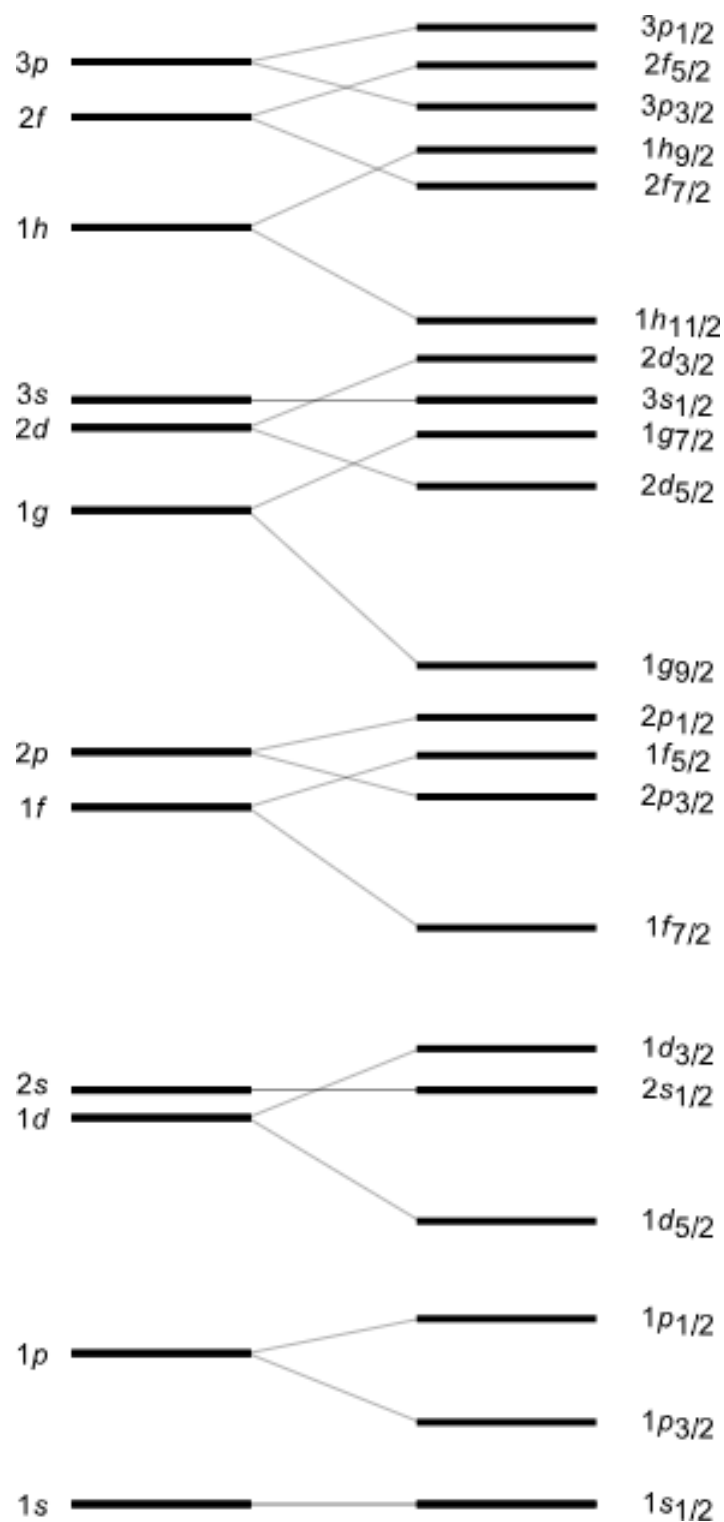
1.1.4 Nukleonski faktor promiskuiteta

Nukleoni koji nisu u popunjenoj ljusci su aktivni ili valentni nukleoni i oni su ti koji određuju svojstva atomske jezgre u modelu ljusaka, a ostale smatramo inertnom sredicom. Ovo je, naravno, aproksimacija koja uvelike olakšava račun. Možemo uvesti faktor nukleonskog promiskuiteta (engl. *nucleonic promiscuity factor*) [17, 18] koji je mjera prosječnog broja proton-neutron interakcija valentnih nukleona,

$$P = \frac{N_p N_n}{N_p + N_n}, \quad (1.9)$$

N_p je broj valentnih protona, a N_n valentnih neutrona.

Ako nemamo valentne nukleone imamo sferno simetričnu jezgru te faktor iščezava ($P = 0$). Za veće vrijednosti faktora $P > 4$ jezgra je više deformirana. Faktor promiskuiteta pokazao se kao korisna veličina u opisu radijusa i energija vezanja srednje teških i teških jezgara.



Slika 1.5: Jednočestični energetska spektar modela ljusaka uz uključenu spin-orbit interakciju od 1s do 3p. Stanja čiji orbitalni angularni moment ne iščezava rascijepana su na stanje više i stanje niže energije. Veći energetska procijepi na desnoj strani definiraju zatvorene ljuske.

1.1.5 Efekti sparivanja

Spin parno-parnih atomskih jezgara bez iznimke iščezava. Ovaj fenomen posljedica je fundamentalnog svojstva nuklearne interakcije – sparivanja. Spareni nukleoni

tvore $J^\pi = 0^+$ parove, ekvivalentne Cooperovim parovima BCS teorije supravodljivosti [16]. Od ranije spomenutih rezidualnih interakcija V_{RES} efekti sparivanja su najvažniji [10]. Interakcije sparivanja bitne su samo između jednakih nukleona (dva protona ili dva neutrona), s obzirom da su spinovi neparno-neparnih jezgri generalno različiti od nula, sparivanje je prisutno samo u $T = 1$ izospinskom stanju dva nukleona. Možemo uvesti novi kvantni broj – senioritet (s) (engl. *seniority*) koji broji nesparene nukleone.

Pobuđena stanja parno-parnih jezgara moguće je postići ili razbijanjem para ili prebacivanjem para u više energetske stanje, u oba slučaja potrebno je oko 2 MeV energije što je više nego za jezgre bez sparenih nukleona. Ova energija definira procijep sparivanja (engl. *pairing gap*) atomske jezgre, kojega ne možemo direktno mjeriti eksperimentom zbog drugih efekata, ali možemo izračunati takozvanom formulom 5 točaka (engl. *five-point formula*) [19] koja je simetrična i dobro odvaja energiju procijepa od efekata srednjeg polja :

$$\Delta^{(5)} = -\frac{\pi N_0}{8} [E_B(N_0 - 2) - 4E_B(N_0 - 1) + 6E_B(N_0) - 4E_B(N_0 + 1) + E_B(N_0 + 2)], \quad (1.10)$$

uz $\pi_{N_0} = (-)^{N_0}$ paritet.

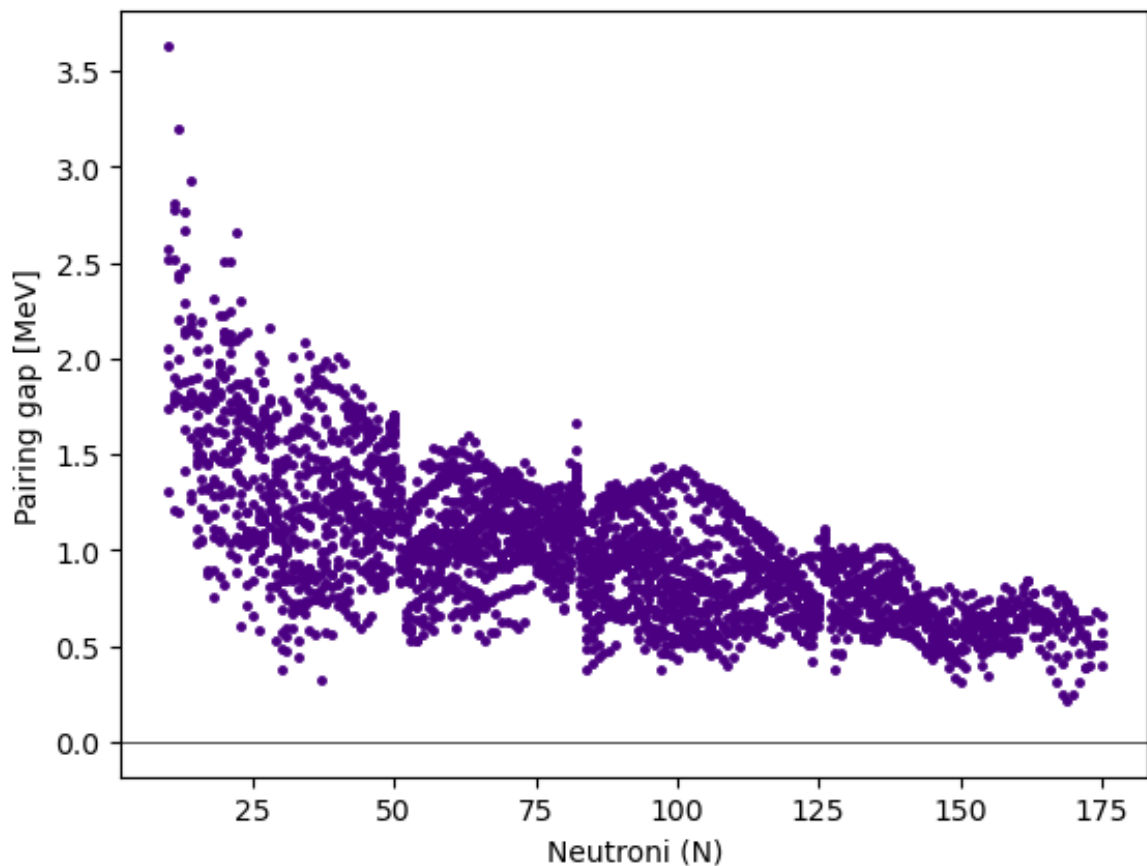
Zbog efekta sparivanja parno-neparne, a posebno parno-parne jezgre su jače vezane od neparno-neparnih.

1.1.6 Wignerova energija

Eksperimentalna mjerenja energija parno-parnih i neparno-neparnih atomskih jezgara pokazuju diskontinuitet kada je broj protona jednak broju neutrona ($N = Z$). Ove jezgre jače su vezane zbog dodatnog efekta sparivanja protona i neutrona koji se nalaze u istim orbitalama modela ljusaka. Ovo sparivanje tvori parove slične deuteronu, izospina $T = 0$. Dodatna energija u sustavu naziva se Wignerovom energijom [20, 21] te ima oblik :

$$W(Z, N) = a_W \frac{|N - Z|}{N + Z} + a_d \frac{\delta_{N,Z} \pi_{N,Z}}{N + Z}, \quad (1.11)$$

$\delta_{N,Z}$ iščezava za sve kombinacije N i Z osim $N = Z$, a $\pi_{N,Z}$ ne iščezava samo za neparno-neparne atomske jezgre, a_W i a_d su konstante određene ekperimentom. Izraz u apsolutnoj vrijednosti naziva se još izospinska asimetrija i podsjeća na četvrti član u semi-empirijskoj formuli (1.3). Wignerova energija proizlazi is izospinske ovisnosti nukleon-nukleon interakcije, zato što vezana stanja različitog izospina, iako je jaka sila invarijantna na izospin, imaju različite energije. Doprinos vezanju jezgre Wignerovog člana opada što je veća izospinska asimetija, daleko od $N = Z$ linije, dok doprinos Pauli člana simetrije raste [22].



Slika 1.6: Graf energija procijepa sparivanja neutrona kao funkcija broja neutrona N izračunatih five-point formulom iz eksperimentalnih podataka baze AME2020.

2 Metode

Posljednjih godina, umjetne neuronske mreže postale su kvalitetno rješenje traženja uzoraka u golemim količinama podataka. Pronašle su svoje mjesto u podatkovnoj znanosti (engl. *data science*), primjenjivane su u biologiji, kemiji, a sve više i u nuklearnoj fizici. Neuronske mreže i slične tehnike strojnog učenja pokazale su se korisnim alatom za modeliranje globalnih svojstava atomskih jezgara s obzirom na veliku količinu dostupnih eksperimentalnih podataka. ANN-ovi su iskorišteni za računanje vremena poluživota α [23] i β raspada; [24] radijusa atomskih jezgara, [25] udarnih presjeka, [26] te za poboljšavanje postojećih metoda [27].

2.1 Strojno učenje

Strojno učenje je jedan vid umjetne inteligencije pri kojem računalo prolazeći kroz skup podataka za koje je poznat rezultat pokušava naučiti uzorke ulaznog skupa kako bi se što bolje reproducirali očekivani rezultati. Upravo zbog toga ANN su veoma obećavajuće u nuklearnoj fizici jer postoji mogućnost da dobro istrenirana mreža ima prediktivne mogućnosti koje pokrivaju cijelu mapu nuklida.

Algoritam strojnog učenja sastoji se od tri komponente: modela, funkcije gubitka te optimizacijskog postupka [28]. Algoritam ima svoj ulaz i izlaz. Ulazni sloj sastoji se od primjera koje prikazujemo kao vektore značajki (engl. *feature*). Značajke su relevantne za problem koji se rješava algoritmom, u konkretnom slučaju značajke su neka svojstva atomske jezgre. Svaki primjer jedan je vektor u ulaznom prostoru (engl. *input space*), dimenzija tog prostora jednaka je broju značajki. Pri nadziranom strojnom učenju primjeri imaju oznake (engl. *label*). Ako se radi o klasifikaciji ta oznaka je klasa primjera, ako se radi o regresiji, kao u ovom slučaju, svaki primjer ima pridruženu brojčanu vrijednost. Oznake primjera sačinjavaju svoj prostor. Skup označenih primjera (engl. *labeled dataset*) je skup uređenih parova iz ulaznog prostora i prostora oznaka. Bit nadziranog strojnog učenja jest pustiti računalo da nauči funkciju – hipotezu - koja će primjerima iz ulaznog prostora pridruživati oznake iz prostora oznaka.

Zadatak algoritma strojnog učenja je podesiti parametre modela koji je jednostavno skup parametriziranih funkcija. Treniranje modela svodi se na pretraživanje cijelog skupa hipoteza kako bi se pronašla najbolja. Problem strojnog učenja se stoga

svodi na postupak optimizacije pretraživanja skupa funkcija po nekom odabranom kriteriju dobrote.

2.2 Neuronske mreže

Neuronska mreža se sastoji od neurona podijeljenih u slojeve koji su međusobno povezani, a informacije teku samo unaprijed. Ulazni i izlazni slojevi su obavezni, a između njih se može nalaziti, po potrebi, proizvoljan broj skrivenih slojeva s jednako proizvoljnim brojem neurona. Mreže s više od jednog skrivenog sloja nazivaju se duboke neuronske mreže (engl. *deep neural networks* - DNN) i s njima se asocira pojam dubokog učenja (engl. *deep learning*). Vektor značajki ide u ulazni sloj, a aktivnosti neurona u sljedećim slojevima ažuriraju se redom jedan po jedan. Nakon što je ANN inicijaliziran s hiperparametrima (dimenzija ulaznog prostora, broj slojeva, broj neurona...), ponašanje istog je potpuno određeno jačinama veza između neurona koje su na početku nasumično postavljene. Kako se mreža trenira - izlaže podacima - parametri veza se mijenjaju tako da se minimizira funkcija gubitka (engl. *loss function*). Ona govori koliko je model izgubio na točnosti na jednom primjeru.

ANN je nelinearna funkcija koja mapira ulazni vektor značajki i parametre kroz slojeve linearnih kombinacija - neurona. Za mrežu s jednim skrivenim slojem funkcija je dana kao

$$f(x) = \alpha + \sum_n^N \beta_n \phi(\gamma_n + \sum_m \delta_{nm} x_n). \quad (2.1)$$

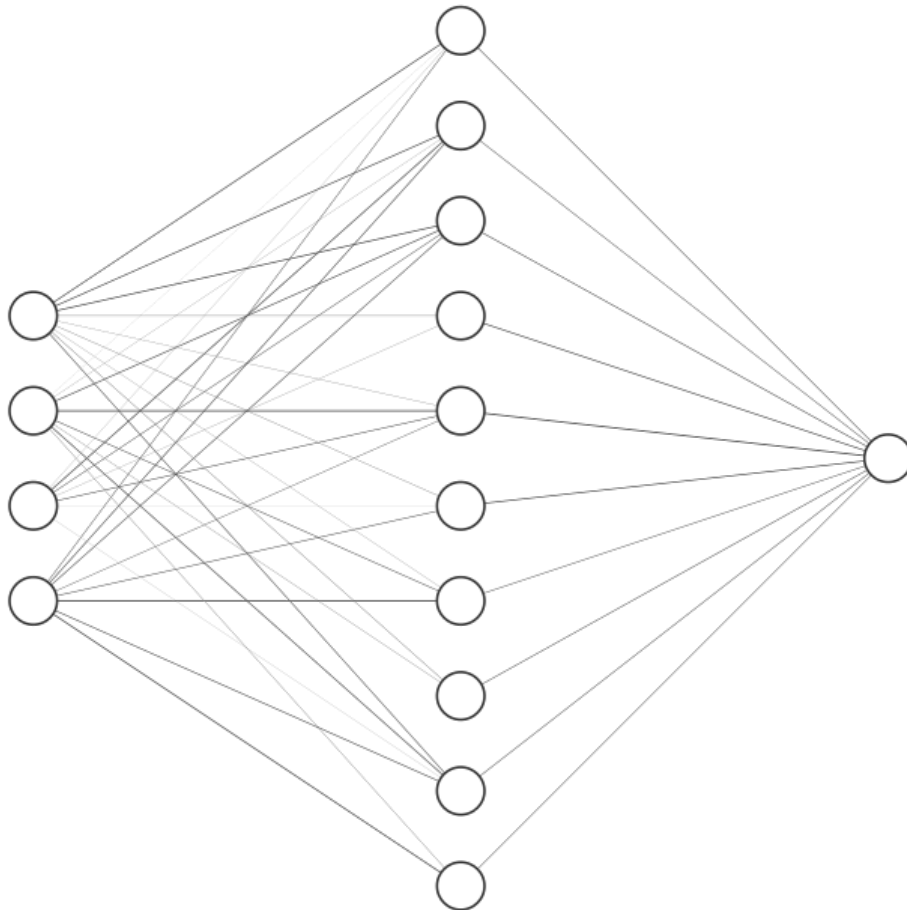
N je broj skrivenih neurona, $\alpha, \beta, \gamma, \delta$ su parametri modela koje treniramo, $\phi()$ je aktivacijska funkcija neurona, njena vrijednost govori mreži je li neuron u aktivnom stanju ili ne. Broj parametara veza jedna je od glavnih značajki dizajna neuronske mreže i ključan je za njenu uspješnost. Za mrežu s L skrivenih slojeva po N neurona, s I ulaznih vrijednosti i O izlaznih imamo ukupno P parametara,

$$P = (1 + I) * N + (N(N + 1))^{L-1} + (N + 1) * O. \quad (2.2)$$

S prikladno odabranim ulaznim parametrima neuronska mreža s prijenosom unaprijed i propagacijom unatrag (engl. *feedforward backpropagation*) je u stanju naučiti iz dostupnih podataka i predviđati svojstva za nove nuklide. Predviđanja dobro

naučenih ANN-a su bila usporediva s rezultatima tradicionalnih mikroskopskih i fenomenoloških modela.

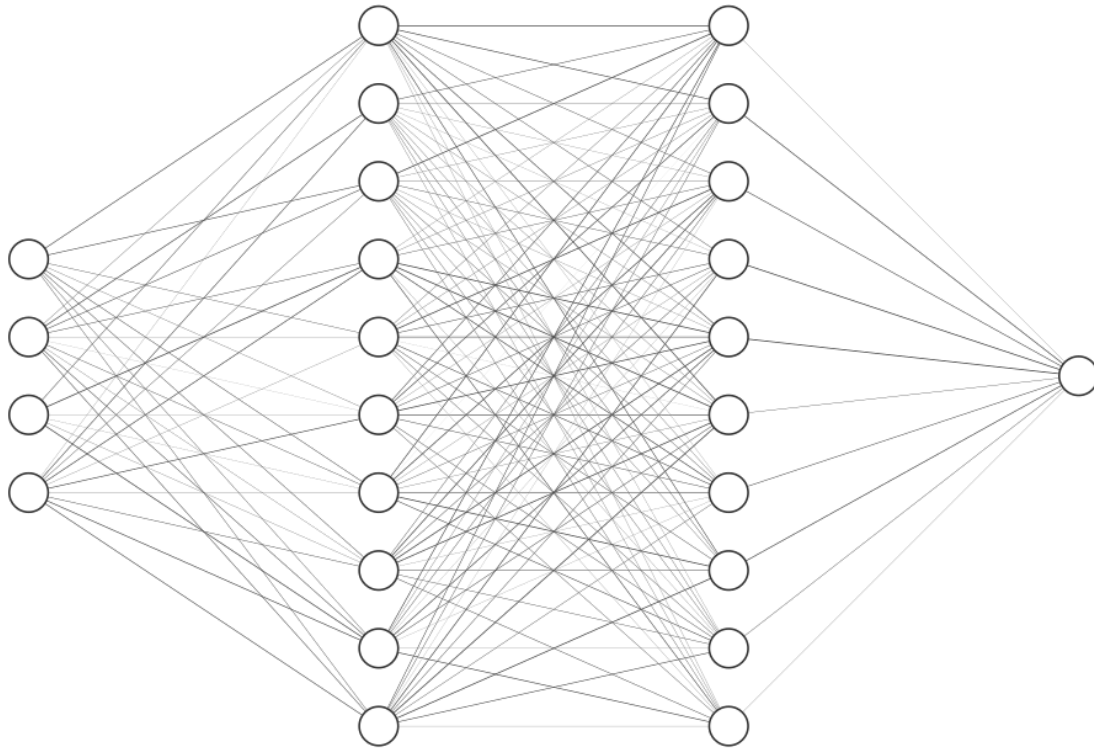
Za ANN na slici 2.1 imamo ukupno 61 parametar koje treba treniranjem prilagoditi podacima. Kada bismo dodali još jedan skriveni sloj, kao na slici 2.2, u mrežu bismo dodali još 110 novih veza, i skoro utrostručili broj parametara koje treba podesiti.



Slika 2.1: Shematski prikaz neuronske mreže s 4 neurona u ulaznom, 10 u skrivenom te jednim u izlaznom sloju - ANN (4-10-1), debljina linije koja spaja neurone predstavlja jačinu veze.

2.2.1 Aktivacijske funkcije

Svaki neuron mreže može postojati u dva stanja, aktivan ili ne, 1 ili 0. Aktivacijska funkcija neurona odgovorna je za izračun njegovog konačnog stanja na osnovu svih veza koje su spojene s tim neuronom, one omogućavaju strojno učenje propagacijom unatrag jer se mijenjanjem parametara modela jačina veza među neuronima, mijenja i njihova aktivnost. Druga zadaća aktivacijske funkcije je dodati nelinearnost u



Slika 2.2: Shematski prikaz neuronske mreže s 4 neurona u ulaznom, dva skrivena sloja s po 10 neurona te jednim neuronom u izlaznom sloju - ANN (4-10-10-1), debljina linije koja spaja neurone predstavlja jačinu veze.

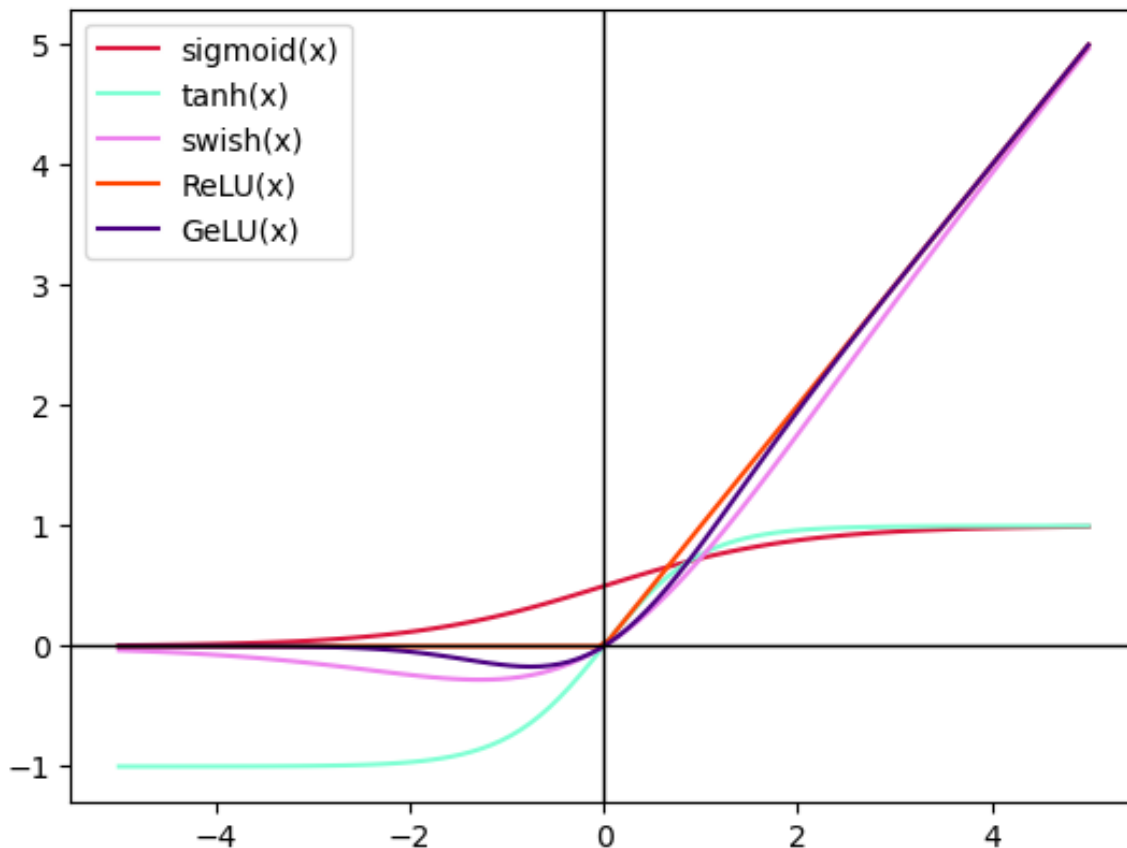
ponašanje neurona, a time i cijelog modela.

Neuronska mreža bez aktivacijske funkcije, odnosno s identitetom kao aktivacijskom funkcijom je model koji provodi jednostavnu linearnu regresiju. Za rješavanje netrivialnih problema potrebne su nelinearne funkcije. Kada se koriste nelinearne funkcije može se pokazati da su višeslojne neuronske mreže univerzalni aproksimatori [32].

Neke aktivacijske funkcije mogu se vidjeti na grafu 2.3.

Domena aktivacijskih funkcija je $< -\infty, \infty >$ kako bi preslikali sve moguće ulaze. Izlazne vrijednosti su, u principu, u rasponu $[0,1]$, no to ne mora biti slučaj kako se vidi na grafu 2.3 za funkcije swish [29] ili GeLU [30]. Kada je raspon funkcije konačan, treniranje traženjem minimuma spustom niz gradijent stabilnije je te je moguće „brže učenje“. Za beskonačne aktivacijske funkcije treniranje je učinkovitije, ali je potrebno koristiti manje vrijednosti stope učenja. Poželjno je da su funkcije kontinuirane i kontinuirano derivabilne [30, 31].

Teoretski je moguće da u odabranom modelu postoji funkcija čija je pogreška nula, koja savršeno opisuje naše podatke. Model u kojemu ne postoji takva funkcija



Slika 2.3: Graf korištenih aktivacijskih funkcija.

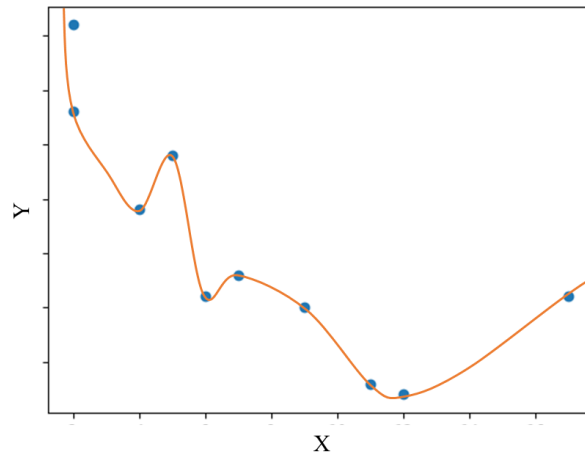
nije dovoljno složen i u realnim primjenama ne postoji model koji bi imao pogrešku jednaku nuli jer je prisutan šum.

2.3 Šum

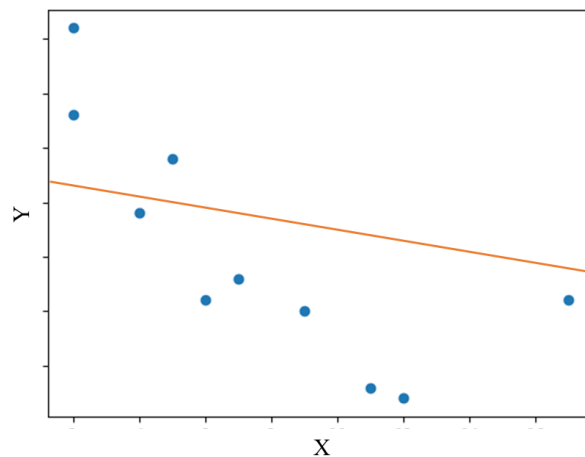
Šum je neželjena anomalija u skupu podataka zbog kojega mjerene ili opažene vrijednosti odstupaju od pravih, on može biti prisutan u nekim značajkama ili u oznakama podataka. Šum može biti uzrokovan nepreciznim mjerenjima značajki ili oznaka, što je lako moguće s obzirom da se u ovom slučaju radi o eksperimentalnim mjerenjima energija vezanja stotina jezgara. Moguće je da postoje i latentne varijable – značajke koje nam nisu poznate ali utječu na konačni rezultat. Bez određene značajke podaci bi izgledali besmisleni, puni šuma, no zapravo nisu, već nam nedostaju ključne informacije. Ovo također može biti slučaj u slučaju modeliranja globalnih svojstava atomskih jezgara jer, kako je spomenuto ranije, točne interakcije unutar jezgara nisu nam posve poznate.

Ako postoji šum, jednostavniji modeli neće moći opisati podatke, a ako koristimo

složenije modele postoji mogućnost da će se oni previše prilagoditi podacima.



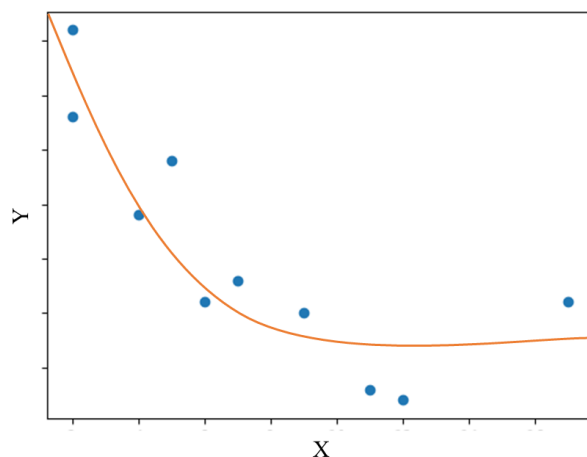
Slika 2.4: Hipoteza na ovome grafu rezultat je prenaučnosti modela. Pogođene su točke poznatih podataka, ovakav model loše generalizira.



Slika 2.5: Hipoteza prikazana na grafu nema dovoljnu složenost da opiše ni poznate podatke, a pogotovo neviđene, radi se o podnačenom modelu.

Prejednostavni modeli imati će visoku pogrešku i općenito rade loše čak i na podacima za treniranje. S druge strane, presloženi model se previše prilagođava šumu podataka za treniranje koji nije nužno jednak onome u neviđenim podacima. U oba slučaja modeli loše generaliziraju.

Plitke ANN lakše su za dizajniranje i treniranje, prikladne su za rad s manje dostupnih podataka. Općenito jednostavniji modeli dovoljne složenosti imaju veću šansu dati bolje rezultate na neviđenim podacima, osim što se brže treniraju, njihove parametre lakše je tumačiti. Složeniji model na kraju treniranja funkcionira kao crna



Slika 2.6: Graf koji prikazuje dobro naučenu hipotezu s potencijalom da dobro generalizira na neviđenim podacima.

kutija (lakše je shvatiti model s nekolicinom parametara koji modelira, recimo, polinom, za razliku od neuronske mreže s tisućama parametara). Duboke mreže mogu modelirati složenije funkcije i tražiti uzorke u podacima na više načina istovremeno te bolje generaliziraju, no imaju problem koji se javlja kada istrenirana mreža zapamti uzorke prisutne u skupu za treniranje i nema nikakvu moć predviđanja dalje od njega - prenaučenosť (engl. *overfitting*). Isti problem se javlja i ako ANN predugo treniramo s istim podacima - Slika 2.4.

Suprotnu situaciju imamo kada mreža nije dovoljno kompleksna da modelira funkciju ili je ne treniramo dovoljno dugo, kažemo da je model podnaučen (engl. *underfitting*). Ovakav model ima loše predikcije i na viđenim i neviđenim podacima - Slika 2.5.

Na slici 2.6 prikazan je rezultat dobro istreniranog modela, takav model najbolje opisuje dane podatke, bez da gubi na moći generaliziranja.

2.4 Odabir modela

Očito treba izabrati model koji najbolje odgovara složenosti problema. Razuman kompromis između kompleksnosti mreže i vremena treniranja rezultira najboljim slaganjem predviđanja neuronske mreže s podacima. Treniranjem dobrog modela možemo, osim kvalitetne reprodukcije energija vezanja, dobiti intuiciju koje značajke su bitne i koliko, analizom parametara samog modela.

Od svih modela određenih kombinacijama hiperparametara želimo odabrati onaj

koji najbolje generalizira. U procesu potrage za modelom koristimo empirijski postupak unakrsne provjere (engl. *cross-validation*).

Iz ukupnog skupa označenih primjera izdvojiti ćemo dio, a model ih neće vidjeti pri treniranju, već će se istrenirani model testirati koristeći te „neviđene primjere“. Sada imamo dva disjunktna skupa, skup za treniranje (engl. *training set*) i skup za testiranje (engl. *test set*). Ako istrenirani model dobro generalizira na skupu za testiranje, možemo pretpostaviti da će dobro generalizirati i na pravim neviđenim podacima.

Skup označenih podataka sastoji se od energija vezanja jezgara iz baze podataka AME2020 [14, 15]. 10% ukupnih podataka je nasumično odabrano i tvori skup za testiranje. Preostalih 90% podjeljeno je na dva djela nasumičnim odabirom u omjeru 80% : 20% na skup za treniranje i skup za validaciju (engl. *validation set*). Skup za treniranje se koristi za učenje ANN-a i ugođavanje parametara veza između neurona tako da se minimizira funkcija gubitka s skupom za validaciju. Nakon treniranja mreže koristimo novi skup podataka, koji ANN nije susreo, kao skup za testiranje kako bismo ocijenili učinkovitost istrenirane neuronske mreže.

S obzirom da je skup podataka s kojim treniramo mrežu reda veličine nekoliko tisuća jezgara, nepoželjno je imati mrežu s jednako velikim brojem parametara. Dodavanjem više neurona povećavamo vrijeme treniranja mreže bez ikakvog dobitka na njejoj učinkovitosti, a postoji i rizik da će suviše složena mreža biti prenaučena.

Kao funkciju gubitka biramo pogrešku srednjeg kvadratnog odstupanja – MSE (engl. *mean squared error*). Više je razloga za takav odabir. MSE je nenegativna i veća odstupanja od točnih vrijednosti rezultiraju većim gubitkom što poboljšava trening modela. Također minimiziranjem MSE minimiziramo i RMS koji se standardno koristi za uspoređivanje različitih nuklearnih modela. Budući da svi podaci koje koristimo prate istu raspodjelu vjerojatnosti, te su sve oznake međusobno nezavisne, možemo govoriti o IID (engl. *independently and identically distributed*) varijablama. U tome slučaju vjerojatnost skupa oznaka jednaka je umnošku vjerojatnosti pojedinačnih oznaka,

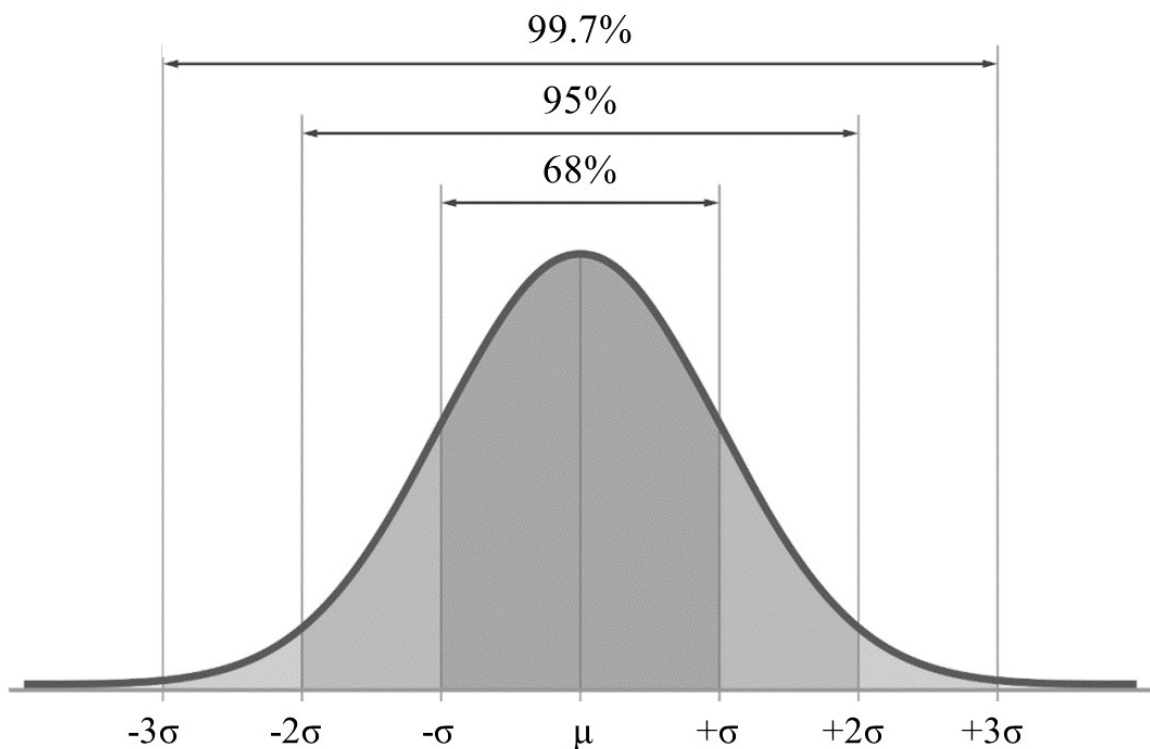
$$p(y) = \prod_{i=1}^N p(y_i|y_i) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - f(x_i))^2}{2\sigma^2}\right). \quad (2.3)$$

U svakoj točki x_i imamo jednu Gaussovu raspodjelu sa središtem u vrijednosti

funkcije $f(x_i)$ koja opisuje energije vezanja, funkcije koje pokušavamo modelirati. Oznake y_i mogu odstupati od predviđanja našeg modela i to odstupanje je modelirano normalnom raspodjelom. Ukupna gustoća vjerojatnosti skupa oznaka jest umnožak gustoća vjerojatnosti pojedinih oznaka y za svaku točku x . Vjerojatnost je veća što se više oznaka y nalazi bliže maksimumima svojih Gaussijana, odnosno biti će veća što bolje naš model pogodi temeljnu funkciju f i tako prođe što bliže oznakama y . Cilj je pronaći hipotezu koja maksimizira vjerojatnost oznaka, tj. naše podatke čini najvjerojatnijima. Logaritmiramo li izraz za vjerojatnost dobijamo:

$$\ln(p(y)) = -N \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - h(x_i))^2. \quad (2.4)$$

Prvi član je konstanta, a drugi član je pogreška kvadratnog odstupanja [33]. Želimo li maksimizirati vjerojatnost, jasno je da drugi član u izrazu valja minimizirati.



Slika 2.7: Prikaz normalne, Gaussove, raspodjele srednje vrijednosti μ i varijance σ^2 .

Standardni algoritam za treniranje zove se engl. „*vanilla backpropagation*”. On optimizira traženje minimuma prvim derivacijama. U drugim derivacijama sadržano je još više informacija te je u principu moguće minimizirati funkciju ranije, no u više dimenzija račun viših derivacija je zahtjevniji. Adam algoritam aproksimira druge

derivacije i mnogo je brži za optimiziranje neuronskih mreža [34].

Kako bi se dobila mreža koja najbolje modelira podatke, valja dobro ugoditi hiperparametre ANN-a (stopa učenja, broj neurona, broj epoha, veličina serija, stopa opadanja). Stopa učenja (engl. *learning rate*) je hiperparametar neuronske mreže u rasponu od 0 do 1. Vrijednost stope učenja utječe na iznos promjene parametara veza među neuronima prilikom svakog ažuriranja. Broj epoha definira koliko će puta algoritam za treniranje mreže iskoristiti cijeli skup podataka, odnosno koliko će puta podaci proći kroz treniranje. Treniranje neuronskih mreža podrazumijeva velike vrijednosti broja epoha, od par tisuća do nekoliko stotina tisuća. Jasno je da će treniranje ANN-a biti duže za više epoha. Prilikom treninga ne koriste se svi podaci odjednom, već se skup za treniranje podjeli u serije (engl. *batch*). Veličina serije određuje koliko je algoritmu za trening potrebno iteracija da završi jednu epohu. Stopa propadanja (engl. *decay*) neobavezni je hiperparametar algoritma Adam. Propadanje se odnosi na stopu učenja i očituje se njenim smanjenjem što dulje traje treniranje. Na taj način, što je bliže minimumu funkcije gubitka, algoritam radi sve manje promjene parametara modela, učenje se penalizira kako se minimum ne bi preskočio. Ovdje nailazimo na posljednji razlog izbora MSE kao funkcije gubitka : ona je kvadratna funkcija s jednim minimumom. Ovo je bitno jer se algoritam neće moći „zaglaviti” u nekom lokalnom minimumu.

3 Rezultati

U ovom radu neuronske mreže izvedene su u programskom jeziku Python korštenjem Alphabetovog open source paketa umjetne inteligencije za strojno učenje TensorFlow uz Keras za duboko učenje [35]. Korišten je Adam algoritam za treniranje, MSE kao funkcija gubitka, a praćena je i veličina MAPE (engl. *mean absolute percentage error*) kao mjera točnosti modela:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{E_{Bi}^{EXP} - E_{Bi}}{E_{Bi}^{EXP}} \right|. \quad (3.1)$$

Mreže su trenirane na serveru fizičkog odsjeka PMFa te računalu osposobljenom s nVidia grafičkom karticom, kako bi se moglo paralelno trenirati koristeći tensorflowgpu paket.

Isprobano je više različitih aktivacijskih funkcija - tablica 3.1. Za plitke mreže se najbolja pokazala sigmoid(x), a za duboke mreže najbolji rezultati postignuti su s funkcijom swish(x) (graf na slici 2.3). Duboke mreže s miješanim aktivacijskim funkcijama nisu se pokazale efikasnim.

Tablica 3.1: Usporedba RMS-a skupa za testiranje za plitki ANN sa 128 skrivenih neurona treniranih 5000 epoha uz stopu učenja 0.045.

aktivacijska funkcija	RMS _{test} [MeV]
tanh(x)	1.763
swish(x)	0.893
ReLU(x)	1.591
GeLU(x)	1.144
sigmoid(x)	0.879

Kako bi se dobila mreža koja najbolje modelira podatke, isprobane su različite kombinacije hiperparametara. Plitke mreže trenirane su s između 16 i 256 neurona u skrivenom sloju, te su granice uzete kao razumna pretpostavka s obzirom na dostupnu količinu podataka [36]. Male mreže s manje od 30 neurona prejednostavne su da bi uspješno modelirale kompleksnije funkcije i zbog toga nisu korisne. Istrenirana mreža nije uspjela pratiti eksperimentalne podatke za lake i teške jezgre. Predviđanja nekih manjih mreža odgovaraju samo dijelovima mape nuklida, dok druge uopće ne mogu postići slaganje čak ni s puno treninga.

Velike mreže s više od 200 neurona prevelike su da bi naučile generalizirati. Naj-

bolji rezultati postignuti su za između 128 i 144 neurona, što ovisno o broju ulaznih značajki odgovara oko 700 parametara koje je potrebno ugoditi treniranjem, za sve konstruirane mreže broj parametara je bio oko 1000. Duboke mreže trenirane su s najviše 4 skrivena sloja, a isprobane su različite arhitekture, s jednakim i različitim brojem neurona po sloju.

Početna stopa učenja postavljena je na 0.00001 što algoritmu daje dosta fleksibilnosti pri učenju, s obzirom da se parametri iz epohe u epohu ne mijenjaju jako. Kako je stopa učenja već postavljena nisko, stopa propadanja je na početku bila postavljena na 0. Kako su se isprobavale ostale kombinacije hiperparametara i uvodile dodatne značajke stopu učenja se postepeno povećavalo dok se nisu uočili negativni rezultati. Stopa je povećavana za oko jednog reda veličine, ili barem pola, jer između 0.0001 i 0.00011 nema prevelike razlike. Za suviše malene vrijednosti čak i nakon velikog broja epoha treniranja nema značajnih promjena funkcije gubitka, nema konvergencije ka minimumu. S druge strane, ako je vrijednost parametra previsoka, primjećuju se oscilacije u funkciji gubitka iz epohe u epohu koje mogu nekontrolirano rasti i model divergira. Opimalna stopa učenja za plitke mreže pokazala se 0.045, a za one duboke, s dva sloja i više 0.0003.

Treniranje počinjemo sa stotinjak epoha prateći kako se pri tome mijenja funkcija gubitka. Ako nisu vidljive značajne promjene iz epohe u epohu model treniramo duže, koliko duže ovisi o rezultatima kraćeg treniranja. Neki modeli brzo postignu $RMS < 10 \text{ MeV}$, za nekoliko stotina epoha, a većinu vremena treniranja provedu daljnim optimiziranjem kako bi RMS smanjili ispod MeV granice, par tisuća epoha. Što dulje treniramo, povećavamo rizik od prenaučavanja, a uz visoke stope učenja i divergencije modela. Oba problema riješena su povećanjem hiperparametra propadanja stope učenja za treniranja od više tisuća epoha. Stope propadanja koje su korištene su 0.1, 0.9 i 0.999.

Kratko trenirane mreže, također, nemaju dobra prediktivna svojstva jer nemaju dovoljno vremena da pronađu uzorke među podacima i ne mogu modelirati energije vezanja za sve atomske jezgre. U najboljim slučajevima to uspijevaju za one lakše. Problem je istaknutiji ako koristimo više značajki na ulazu.

Treniranje je provedeno na tri načina, u serijama veličine 1, tako da se parametri modificiraju nakon svakog označenog primjera, alternativno su parametri ažurirani nakon svake epohe jer su svi podaci korišteni odjednom. Najefikasniji pristup je bio

sa serijama veličine 200, na ovaj način po epohi se parametri modela ažuriraju 13 puta.

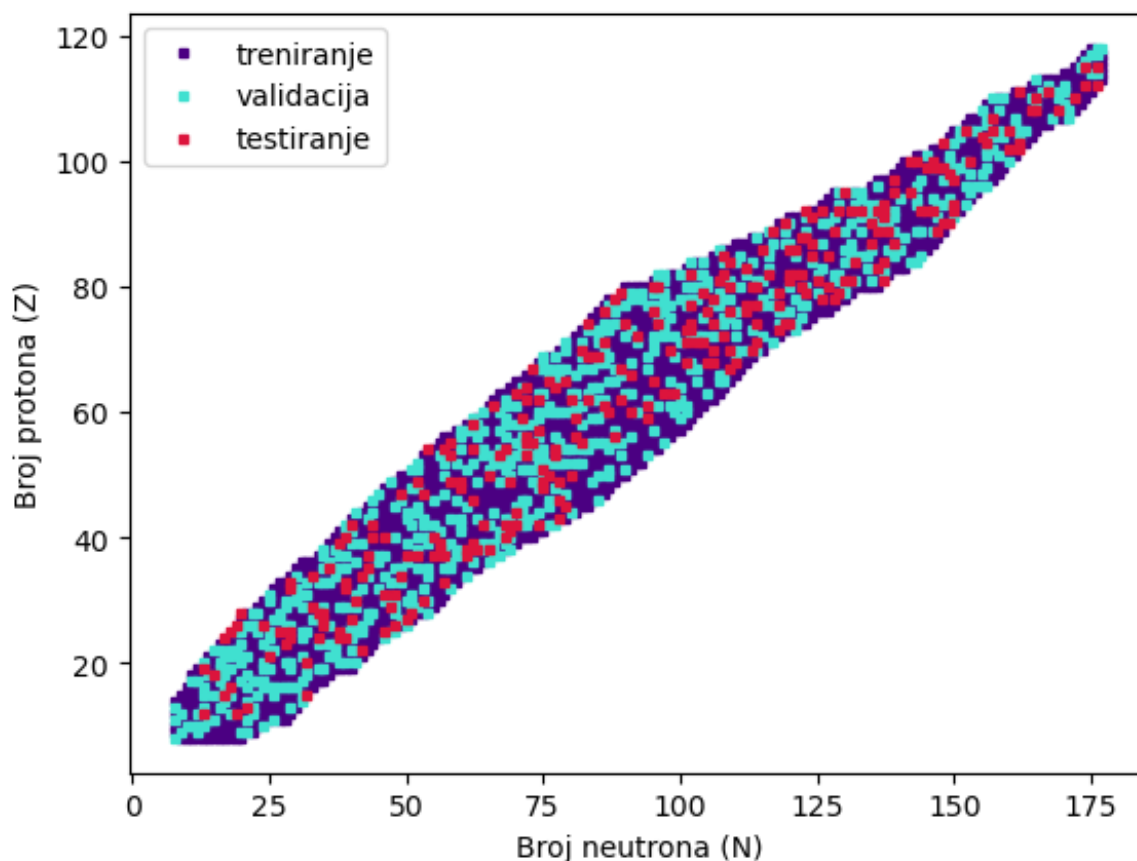
Idealan pristup bio bi koristiti sve podatke odjednom i trenirati dulje vremena, tako se parametri neće često mijenjati i funkcija gubitka će konstantno težiti prema minimumu, bez oscilacija. Ovakav pristup najzahtjevniji je po računalne performanse, a postoji mogućnost da se u sustavu s mnogo parametara nađemo u nekom lokalnom minimumu koji nije najbolje rješenje problema. Korištenje jednog po jednog primjera za ažuriranje parametara unosi puno šuma u proces treniranja pa isti traje puno dulje, no zato što učestalo modificiramo parametre moguće je da će minimum prije biti pronađen. Pokazalo se da veće serije ubrzavaju treniranje po cijenu konvergencije funkcije gubitka ka minimumu.

Trening mreža može se odigravati paralelno, ovisno o računalnom hardwareu. Kako se ne bi gubilo vrijeme s modelima koji su u početku loši ili počinju divergirati, uveden je sustav ranog zaustavljanja treniranja (engl. *early stopping*) koji prati funkciju gubitka kroz treniranje te ako ne vidi poboljšanje iste nakon određenog broja epoha, zaustavlja treniranje. Sustav također pamti najbolju konfiguraciju parametara modela te nakon završenog treninga dostavlja te parametre kao najbolje naučeni model.

Na slici 3.1 prikazana je mapa nuklida na kojoj su pojedine atomske jezgre označene ovisno kojem skupu pripadaju. Ukupni skup podataka nisu sve jezgre iz AME2020 već samo one kojima su protonski i neutronski broj veći od 8. Lake jezgre izostavljene su jer su rješive drugim metodama uz zadovoljavajuće preciznosti. Njihovo prisustvo u skupu podataka za strojno učenje je kontraproduktivno radi raznih specifičnih efekata prisutnih u lakim atomskim jezgrama. Kada oduzmemo te lake jezgre ostaju nam skupovi veličina danih u tablici. Na slici 3.2 prikazan je graf energija vezanja atomskih jezgara, a jezgre su ponovno označene ovisno o tome kojem skupu pripadaju, podjela jezgara po skupovima prikazana je i u tablici 3.2.

Tablica 3.2: Sadržaj skupova za strojno učenje nakon eliminacije lakih jezgara.

skup	broj nuklida
treniranje	2487
validacija	692
testiranje	277
ukupno	3456

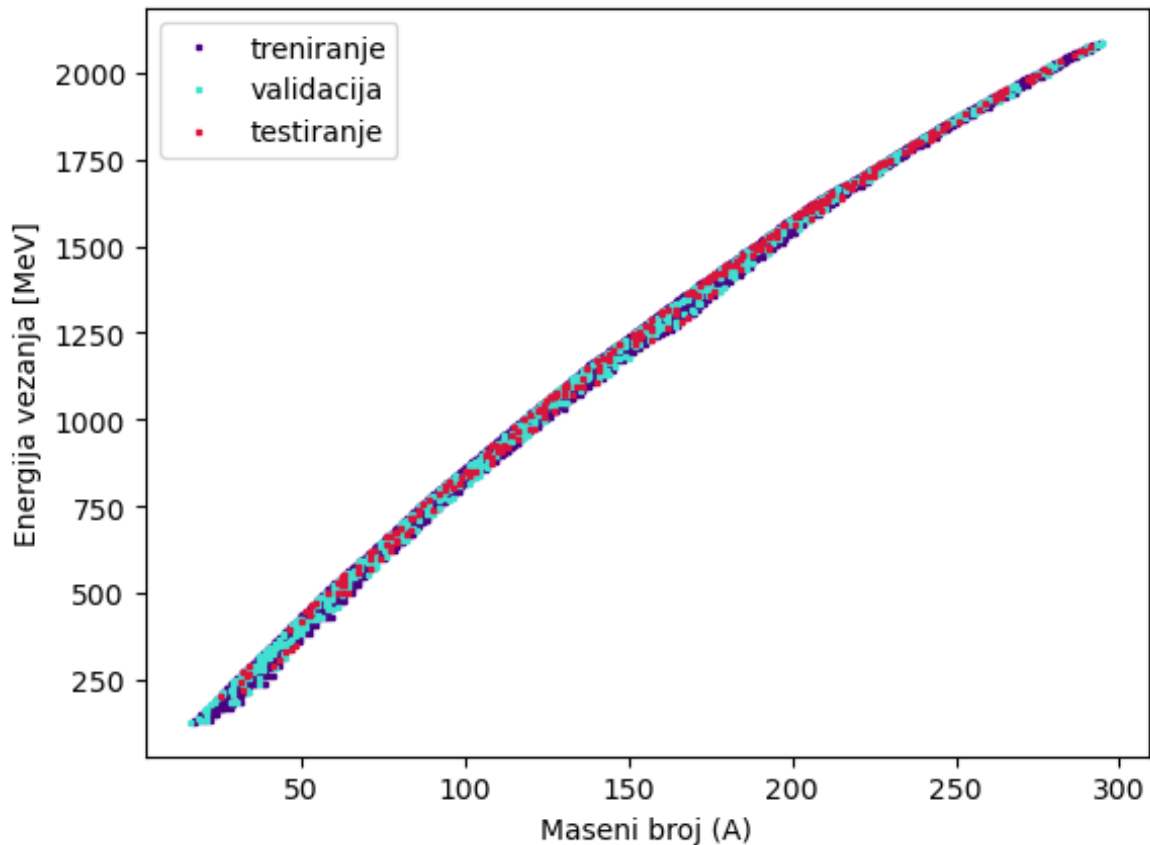


Slika 3.1: Prikaz mape nuklida iz baze podataka AME2020 na kojoj su označene atomske jezgre korištene u skupu za treniranje - ljubičastom, validaciju - plavom i testiranje - crvenom bojom.

Upoznati sa optimalnim intervalima za postavke hiperparametara modela, sada je ključan korak odabrati značajke po kojima će modeli učiti. Bit je izabrati fizikalno relevantne veličine za koje je poznato da utječu na energiju vezanja, ali i one za koje slutimo da bi mogle igrati ulogu. Na učinkovitosti naučenog modela je da otkrije, ako je odabir bio dobar ili ne. Jasno, postoji mnoštvo veličina koje bismo mogli uključiti u vektor značajki, od općih svojstava atomskih jezgara, do informacija o električnim i jakim silama prisutnima u samoj jezgri.

Za početak pokušavamo trenirati mrežu s dva ulazna parametra, broj protona (Z) i broj neutrona (N), kao minimalni skup vrijednosti o kojima znamo da ovisi energija vezanja.

U suštini i najbolje istrenirana neuronska mreža ne zna ništa o fizici koju modelira, i dok ona iz uzoraka u eksperimentalnim podacima može naslutiti postojanje, npr. strukture ljusaka, ona ne prepoznaje njihovu relevantnost. Zbog toga u drugom koraku model proširimo s više ulaznih parametara, odnosno obogatimo ulazni



Slika 3.2: Graf energija vezanja u ovisnosti o masenom broju (A) za atomske jezgre iz AME2020. Jezgre su označene ovisno u kojem skupu za strojno učenje se nalaze. Ljubičastom bojom - skup za testiranje, plavom - skup za validaciju te crvenom - skup za testiranje.

prostor s više značajki. Osim broja protona i neutrona, ANN sada prima još dva parametra koji su mjera koliko protona te neutrona je viška od popunjenih ljusaka - ΔZ i ΔN broje valentne protone, tj. neutrone. Na ovaj način smo obogatili mrežu sa znanjem nuklearne fizike i strukture ljusaka. Iako su eksperimentalni podaci za treniranje jednaki, koristimo isti skup oznaka, dva dodatna ulaza promijeniti će se kako, sada nova, mreža uči. Cilj je pronaći neki minimalni najbolji skup značajki kako bi što bolje opisali sve energije vezanja. Više značajki ne garantira bolji model, zato odabiru značajki pristupamo sistematski i pažljivo [38].

Istrenirane plitke mreže označene su ANNX, mreže trenirane dubokim učenjem označene su DNNX, gdje je X broj značajki korištenih u ulaznom sloju. Za potrebe ovoga rada istrenirano je nekoliko tisuća neuronskih mreža, što za potrebe optimizacije hiperparametara, što u potrazi za optimalnim značajkama.

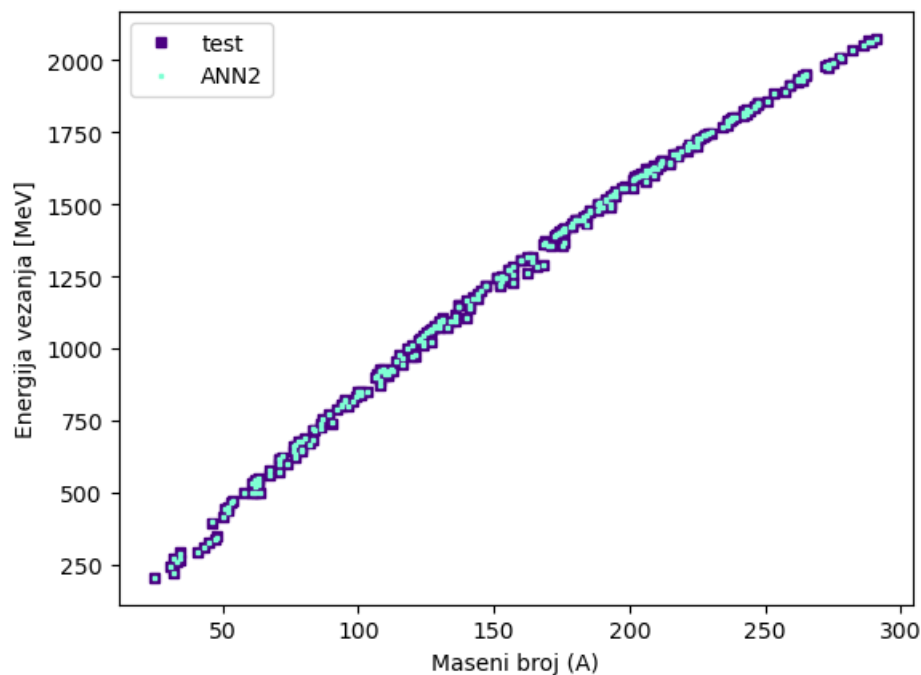
Svojstva istrenirane mreže ocjenjujemo RMS odstupanjem predviđenih vrijednosti od eksperimentalnih vrijednosti skupa za testiranje, jednadžba (1.2). Ako su RMS

vrijednosti skupa za testiranje i skupa za treniranje bliski možemo reći da je model dobro istreniran. U slučaju da je RMS skupa za treniranje mnogo manji od onoga skupa za testiranje, možemo naslutiti da je model prenaučten te da ne generalizira. Ako je model dobar izračunamo njegov ukupni RMS te rezidualne predviđenih vrijednosti i eksperimenta. Naravno, želimo da te vrijednosti budu što manje. Na isti način uspoređujemo učinkovitost strojnog učenja i tradicionalnih metoda, uspoređujući njihove RMS vrijednosti.

U nastavku rada slijedi pregled najuspješnije istreniranih modela.

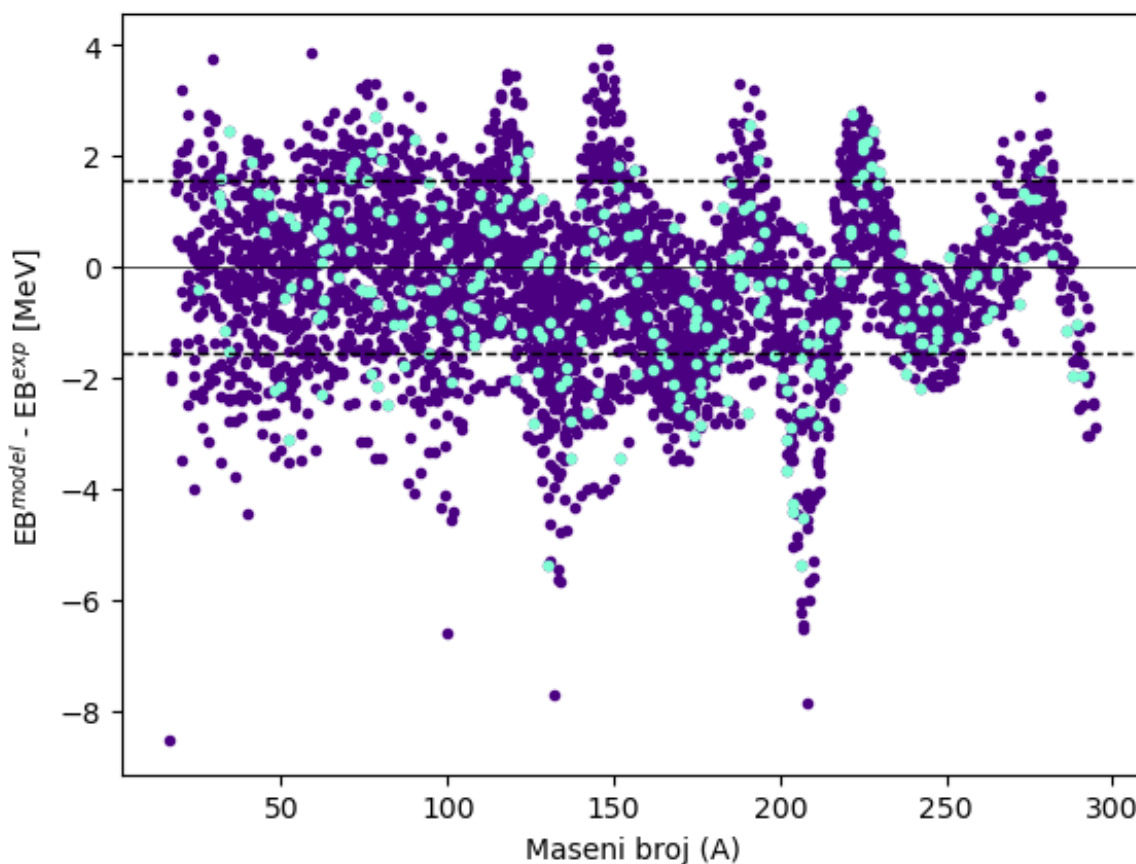
3.1 ANN2

Najjednostavnija plitka neuronska mreža koja na ulazu prima dvije značajke, broj protona (Z) i broj neutrona (N). Model sa 128 skrivenih neurona je treniran kroz 10000 epoha u serijama od 200 primjera. Stopa učenja postavljena je na 0.045 s stopom opadanja od 0.9. Predviđanja istrenirane neuronske mreže ANN2 prikazana su na slici 3.3.



Slika 3.3: Grafički prikaz predviđanja neuronske mreže ANN2 sa 128 skrivena neurona, trenirane kroz 10000 epoha uz stopu učenja 0.045. Graf usporede skupa za testiranje i predviđanja neuronske mreže, $RMS = 1.5434$ MeV.

$RMS_{test} = 1.5644$ MeV , a ukupni RMS modela je $RMS_{tot} = 1.5434$ MeV. MAPE vrijednost iznosi 0.15.



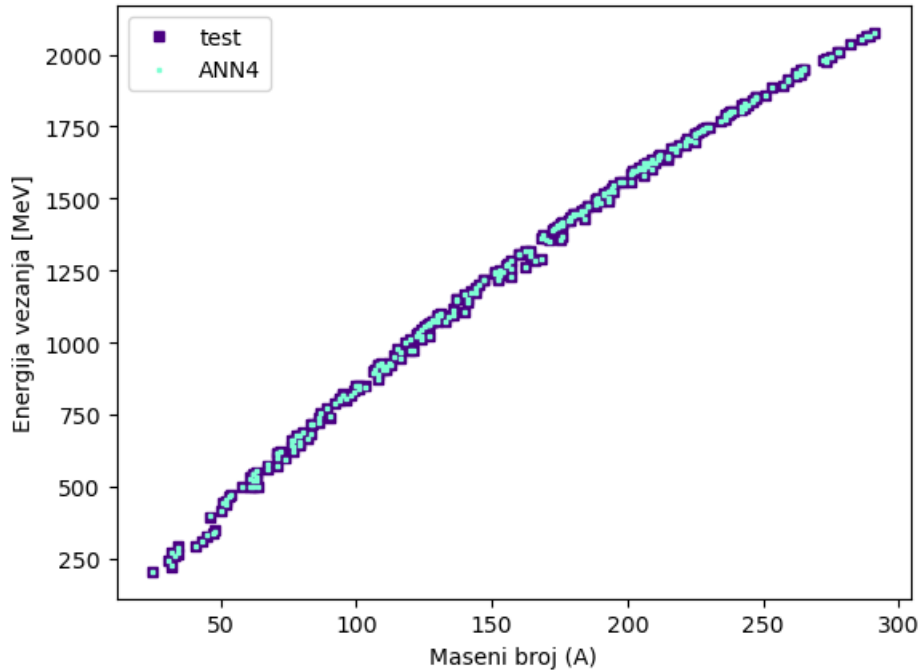
Slika 3.4: Razlika energija vezanja dobivenih ANN2 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 1.5 MeV.

Na slici 3.4 prikazan je grafički prikaz razlika vrijednosti energija vezanja dobivenih istreniranim ANN2 modelom i poznatih eksperimentalnih podataka za sve jezgre iz ukupnog skupa podataka. Posebno su istaknute one jezgre koje su bile dio skupa za testiranje. Vidi se da većina upada u interval RMS-a najbolje istrenirane ANN2 mreže. Odstupanja se mogu primijetiti oko magičnih brojeva protona i neutrona, što je za očekivati jer su te jezgre jače vezane. Dodavanje značajke masenog broja (A) nije doprinijelo poboljšavanju modela, jer su informacije već sadržane u Z i N . Također su nekorisne bile različite potencije masenog broja.

3.2 ANN4

Ovaj plitki model obogaćen je s informacijama o nuklearnoj strukturi ljusaka u očekivanju da će to znanje popraviti predviđanje jednostavnijeg modela oko magičnih brojeva. ANN4 prima četiri značajke Z , N , ΔZ i ΔN , tj. broj valentnih protona i neutrona.

Magični brojevi za protone i neutrone su različiti, posljednja zatvorena ljuska za protone je sa 126, dok za neutrone imamo još dva magična broja 184 i 196 [39, 40]. Mreža sadrži 128 skrivenih neurona i trenirana je 10000 epoha u serijama od 200 primjera. Početna stopa učenja bila je 0.045 uz propadanje od 0.9. Predviđanja ove, poboljšane, neuronske mreže prikazana su na slici 3.5.

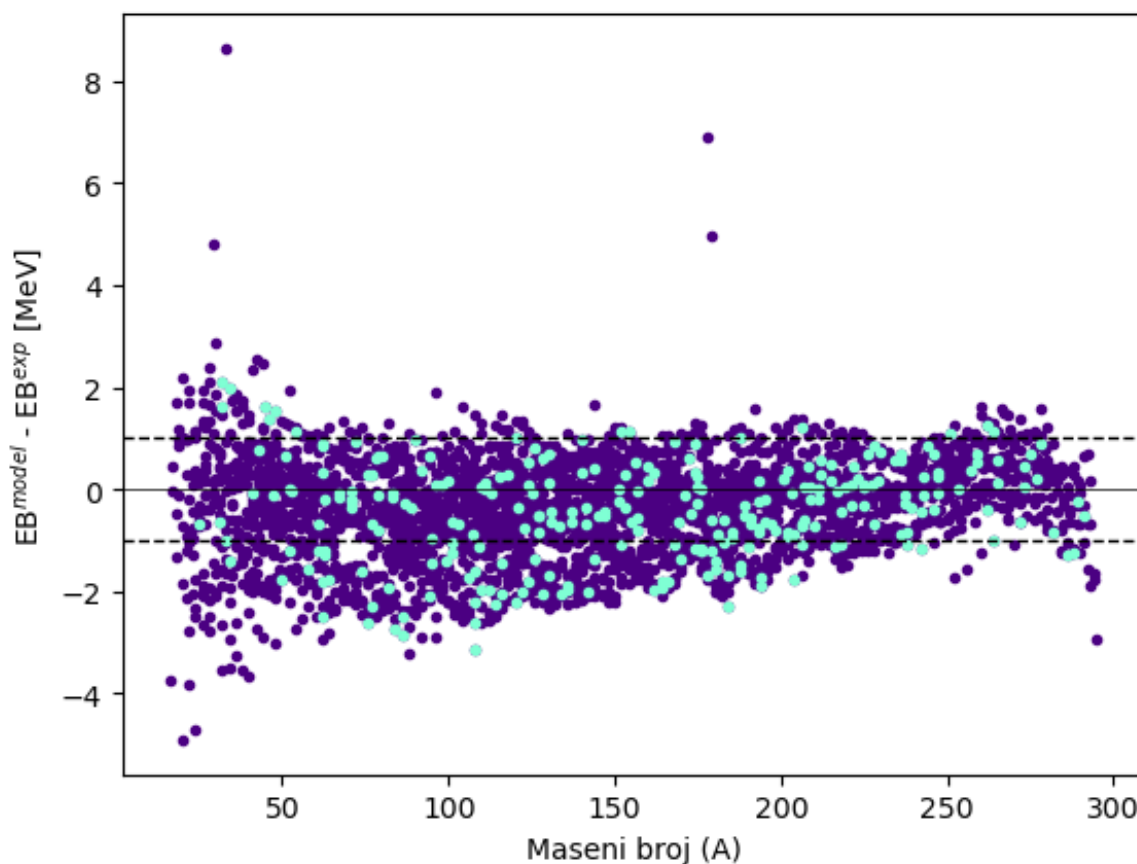


Slika 3.5: Grafički prikaz predviđanja neuronske mreže ANN4 sa 128 skrivenih neurona, trenirane kroz 10000 epoha uz stopu učenja 0.045. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, RMS = 1.0046 MeV.

$RMS_{test} = 1.1027$ MeV , a ukupni RMS modela je $RMS_{tot} = 1.0046$ MeV. MAPE vrijednost iznosi 0.11.

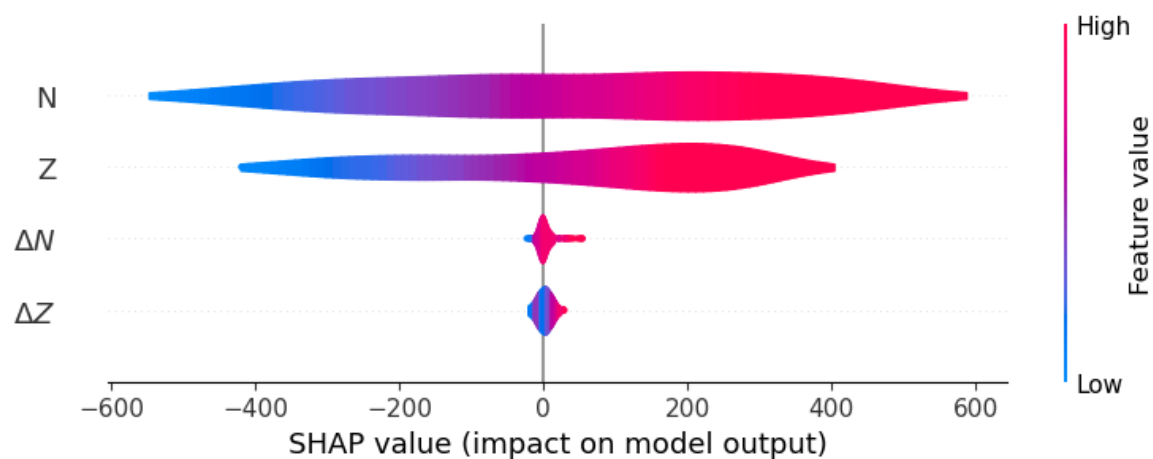
Na slici 3.6 prikazan je grafički prikaz reziduala energija vezanja dobivenih ANN4 modelom i eksperimentalnih podataka za sve jezgre iz ukupnog skupa podataka. Vrijednosti koje su posebno istaknute su jezgre iz skupa za testiranje. Vidljivo je da ANN4 bolje modelira energije vezanja, a razlika je pogotovo vidljiva oko magičnih brojeva. Većina reziduala upada u interval RMS-a najbolje istrenirane mreže. Odstupanja su prisutna za atomske jezgre daleko od zatvorenih ljusaka kada efekti deformacija utječu na vezanje. Jezgre koje najviše odstupaju su ^{20}Ne čija je energija vezanja podcjenjena za 4.9 MeV-a i ^{33}Ca čija je energija vezanja precjenjena za 8.6 MeV-a.

Na slici 3.7 prikazana je SHAP (engl. *SHapley Additive exPlanations*) analiza značajki modela. SHAP je metoda koja nastoji objasniti kako strojno naučeni mo-



Slika 3.6: Razlika energija vezanja dobivenih ANN4 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 1 MeV.

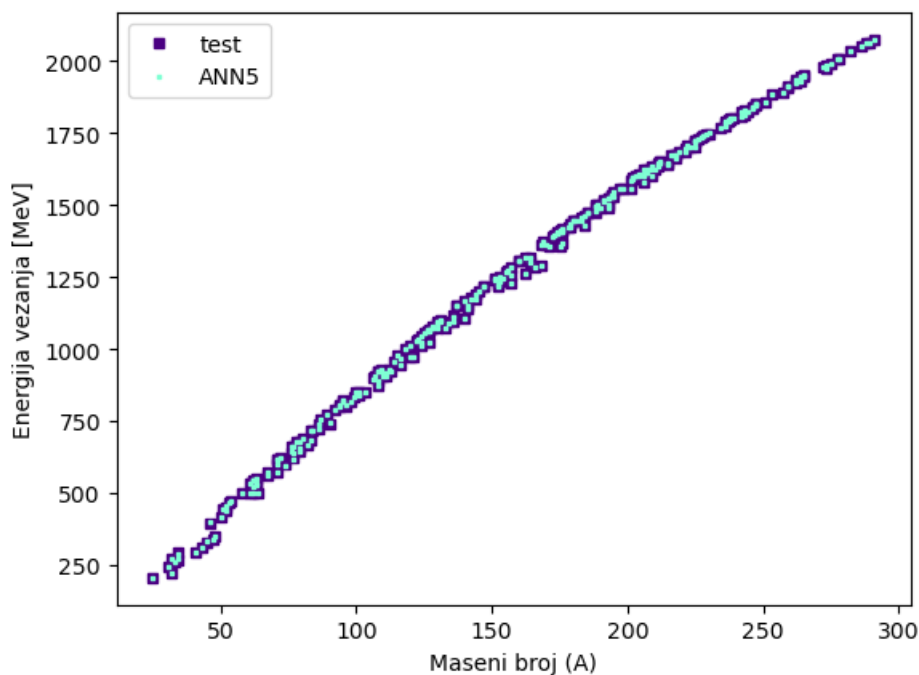
del predviđa vrijednosti te kvantificirati koliko je pojedina značajka modela bitna za predviđanja. Ovakva analiza korisna je provjera kako bi se uvjerali da se model ponaša onako kako bismo to željeli. Rezultati SHAP analize daju uvid ne samo u važnosti pojedinih značajki, već otkrivaju odnose značajki i predviđene vrijednosti neuronske mreže. Što je SHAP vrijednost (na grafu engl. *SHAP value*) značajke na grafu veća to je energija vezanja viša. Također vrijedi i obratno. Numerička vrijednost značajke (na grafu engl. *feature value*) označena je bojom, plavo za niže, crveno za više vrijednosti. Iz konkretne analize možemo vidjeti kako broj protona (Z) i neutrona (N), dominiraju među značajkama, te što je vrijednost Z ili N veća to je veća i energija vezanja.



Slika 3.7: Rezultat SHAP analize utjecaja značajki na svojstva modela ANN4. Vidi se da je broj neutrona N značajka koja najviše doprinosi, također se vidi da jezgre s više protona i neutrona imaju veće energije vezanja. Valentni nukleoni slabije pridonose modelu, ali to je zato što su vrijednosti značajki ΔZ i ΔN malene u usporedbi s N i Z .

3.3 ANN5

Prošli model ANN4, obogatili smo koristeći nukleonski faktor promiskuiteta P , definiran kao (1.9). Uz Z , N , ΔZ i ΔN ANN5 prima dodatnu značajku P . Faktor P može poslužiti kao mjera deformacije atomske jezgre. Model je treniran 15000 epoha u serijama od 200, uz početnu stopu učenja od 0.045 s propadanjem iste od 0.9. Mreža sadrži 136 skrivena neurona. Na slici 3.8 prikazana su predviđanja neuronske mreže ANN5 na skupu za testiranje.

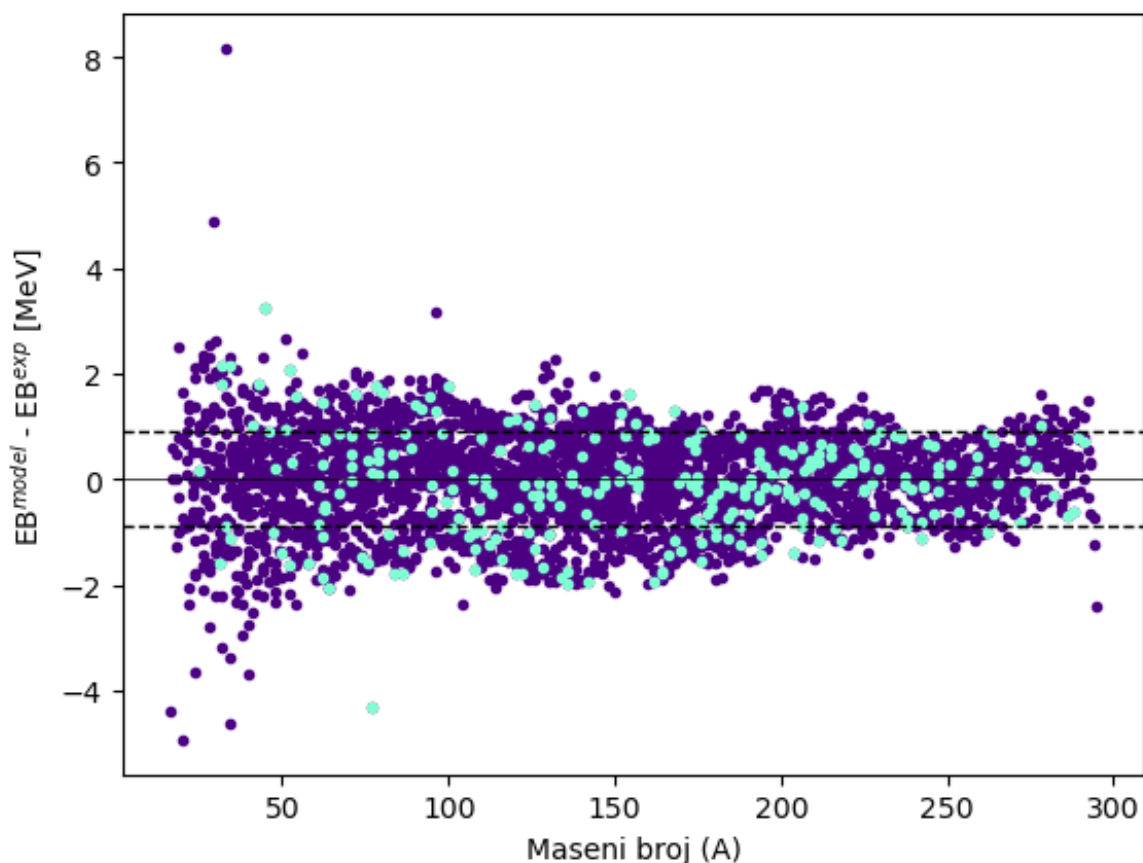


Slika 3.8: Grafički prikaz predviđanja neuronske mreže ANN5 sa 136 skrivena neurona, trenirane kroz 15000 epoha uz stopu učenja 0.045. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, $RMS = 0.8912$ MeV.

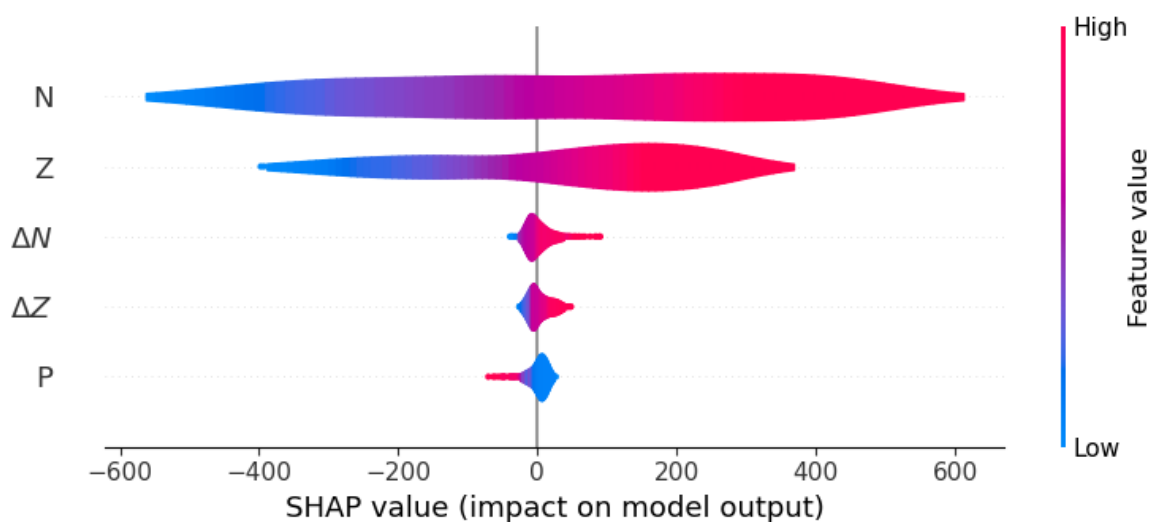
$RMS_{test} = 0.9683$ MeV, a ukupni RMS modela je $RMS_{tot} = 0.8912$ MeV. MAPE vrijednost iznosi 0.10.

Dodavanjem faktora P kao značajke dodatno smo poboljšali model te se riješili većih odstupanja za srednje teške i poboljšali grupiranje za teške jezgre - slika 3.9.

Slika 3.10 prikazuje rezultate analize utjecaja značajki na trenirani model. Novo uvedeni faktor P smanjuje energiju vezanja što je veći, što fizikalno odgovara sustavima deformiranih jezgara čije energije vezanja jesu manje od onih sferno-simetričnih, oko potpunjenih ljusaka.



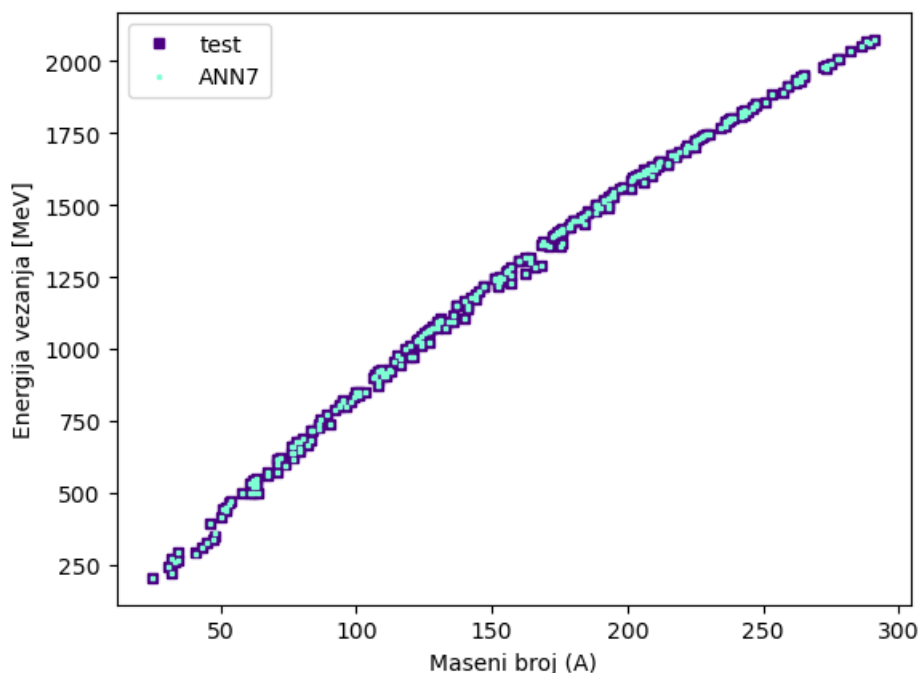
Slika 3.9: Razlika energija vezanja dobivenih ANN5 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.9 MeV.



Slika 3.10: Rezultat SHAP analize utjecaja značajki na svojstva modela ANN5. Broj neutrona i protona, N i Z , dominiraju među značajkama. Vidi se da manje vrijednosti faktora promiskuiteta P vode ka većoj energiji vezanja.

3.4 ANN7

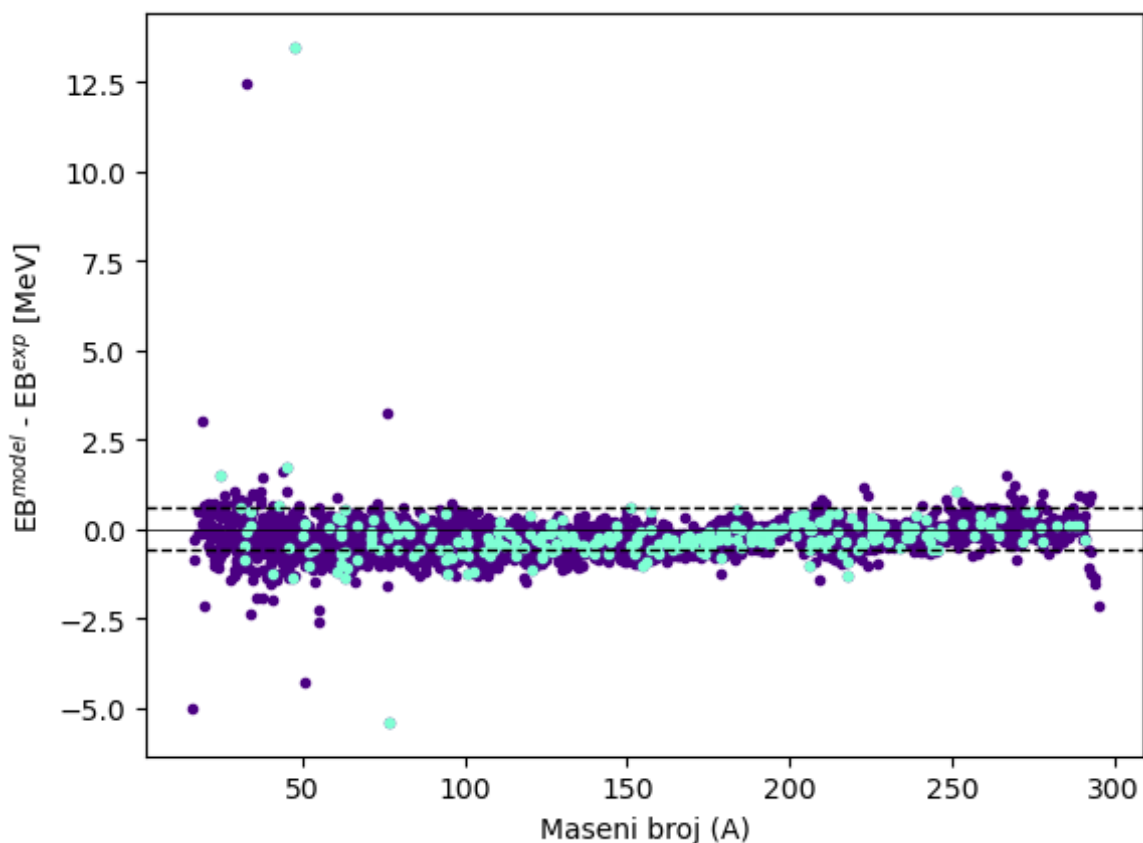
Nakon što je u ANN4 i ANN5 uključena struktura ljusaka, ovaj model ANN7 proširujemo znanjem o efektima sparivanja. Prvi pokušaj bio je izračunati energije procijepa za protone i neutrone te ih uvesti u vektor značajki. Takav pristup pokazao se lošim te se učinkovitost istrenirane ANN7 mreže smanjila u odnosu na jednostavnije modele. Razlog tome je što procjep sparivanja nije prisutan za sve jezgre, već samo za parne što rezultira u lošim podacima za treniranje. Umjesto energija procijepa dodane su značajke senioriteta za protone i neutrone, s_Z, s_N . ANN7 prima vektor značajki $Z, N, \Delta Z$ i $\Delta N, P, s_Z, s_N$. Model sa 144 skrivena neurona treniran je 15000 epoha u serijama od 200 primjera, uz stopu učenja 0.035 s propadanjem iste postavljenim na 0.9. Rezultat treniranja ove mreže vidi se na slici 3.11.



Slika 3.11: Grafički prikaz predviđanja neuronske mreže ANN7 sa 144 skrivena neurona, trenirane kroz 20000 epoha uz stopu učenja 0.03. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, $RMS = 0.5828$ MeV.

$RMS_{test} = 0.7145$ MeV, a ukupni RMS modela je $RMS_{tot} = 0.5828$ MeV. MAPE vrijednost iznosi 0.06.

Odstupanja od eksperimentalnih podataka prikazana su na grafu - slika 3.12. Model ima najviše problema s lakšim jezgrama, za koje se primjećuje najviše i najveća odstupanja. Ta odstupanja upravo ruše RMS modela. Najveće odstupanje zabilježeno je za jezgru nikla ^{48}Ni čiju energiju vezanja je model podcijenio za 13.74 MeV-a, što

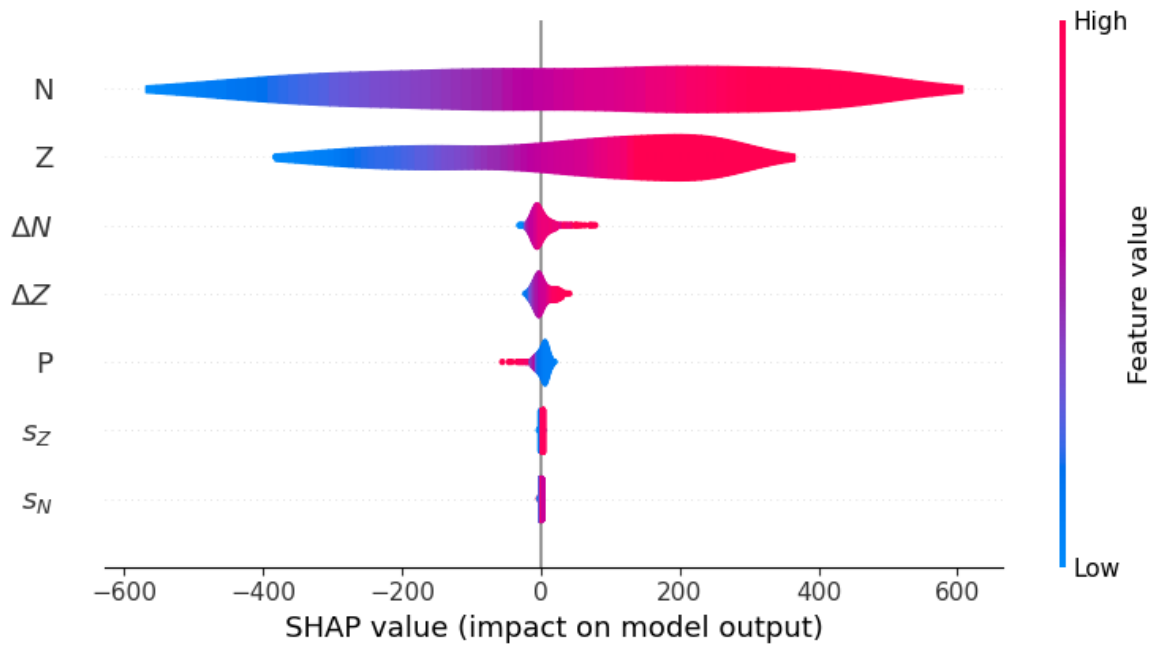


Slika 3.12: Razlika energija vezanja dobivenih ANN7 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.6 MeV.

je posebno interesantno s obzirom da se radi o dvostruko magičnoj, parno-parnoj jezgri.

Analiza značajki modela prikazana je na slici 3.13. Naizgled se čini da nove značajke ne doprinose mnogo, no to nije tako. Budući da kvantni broj senioriteta broji neparne nukleone, u slučaju značajki odabranih za ovaj model, s_Z , s_N mogu iznositi ili 0 ili 1, što je za SHAP analizu mala numerička vrijednost. Ove značajke određuju radi li se o parno-parnoj, neparnoj ili neparno-neparnoj atomskoj jezgri, a kako vidimo na slici 3.12 te nove informacije su dodatno poboljšale model. Uvođenjem značajki senioriteta protona i neutrona model je obogaćen znanjem o efektima sparivanja.

Daljnje dodavanje značajki u plitke mreže nije značajno poboljšavalo rezultate. Dodatne značajke poput Coulombovog člana, čak i uz korekcijske parametre po uzoru na Skyrme [9] ili FRDM modele [41] više su štetile predviđanjima modela.



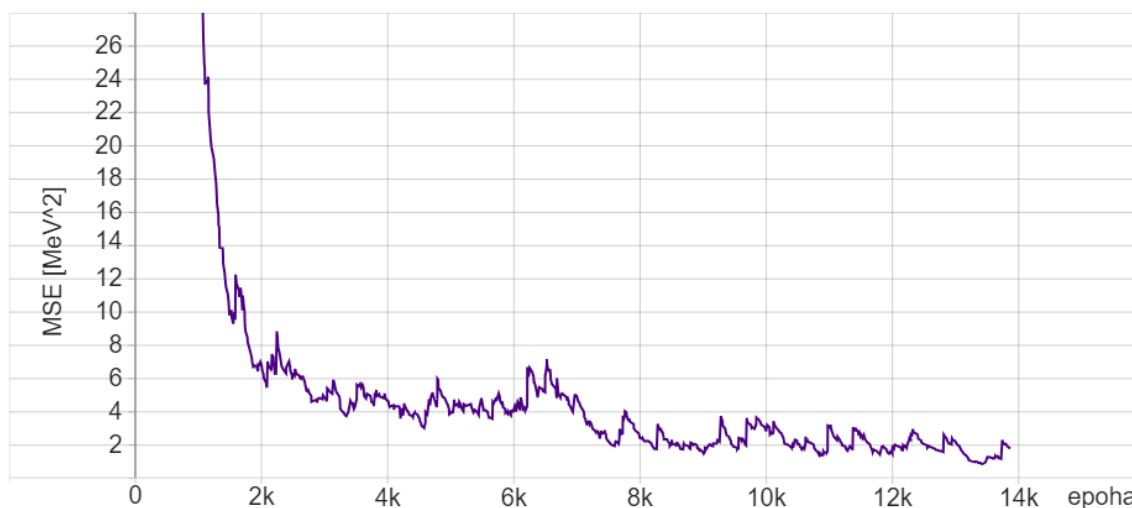
Slika 3.13: Rezultat SHAP analize utjecaja značajki na svojstva modela ANN7.

Očito je dostignuta granica korisnosti plitkih modela i stoga nadalje diskusiju fokusiramo na razvijene duboke modele. Iskušani su duboki modeli DNN4, DNN5, analogni ANN4 i ANN5 no njihova učinkovitost nije bila bolja od ekvivalentnog plitkog modela. ($\text{RMS}_{test}^{DNN4} = 0.9786 \text{ MeV}$, $\text{RMS}_{tot}^{DNN4} = 1.0896 \text{ MeV}$ i $\text{RMS}_{test}^{ANN4} = 1.1027 \text{ MeV}$, $\text{RMS}_{tot}^{ANN4} = 1.0046 \text{ MeV}$) Poboljšanje se primjećuje za modele s više značajki, jer upravo to je prednost dubokoga učenja.

Prvobitni pristup konstrukciji dubokih mreža bio je razviti arhitekturu koja bi imala zadovoljavajući broj parametara, ne više od 1000, te raspoređivati neurone u skrivene slojeve da se to postigne na različite načine. Ovakav pristup pokazao se sporim i češće se sveo na intuiciju koju su omogućili prošli pokušaji. Bolji pristup razvijen je korištenjem regularizacije pri procesu strojnog učenja. Složenost modela ovisi o jačinama veza između značajki i neurona. Poznato je da presloženi modeli imaju tendenciju prenaučavanju i takvi modeli imati će vrlo jake veze - visoke vrijednosti za mnoge parametre modela koji će dominirati. Ovo možemo spriječiti ako penaliziramo one parametre veza koje postanu previše jake [35]. U procesu učenja moguće je efektivno izbaciti pojedine neurone iz mreže tako da nikada nisu aktivni. U suštini treniranje je započeto s modelom koji je vrlo složen kako bi se izbjegla podnaučenos, a za vrijeme učenja regulariziramo model kako bi ga ukrotili i spriječili prenaučenos.

3.5 DNN7

Prvi duboki model koji je pokazao značajnije poboljšanje od plitkih modela je DNN7, duboka mreža s značajkama jednakima ANN7: $Z, N, \Delta Z$ i $\Delta N, P, s_Z, s_N$. Ova mreža započela je treniranje s po 30 neurona u 4 skrivena sloja (30x30x30x30), uz stopu učenja 0.0003 s propadanjem 0.999. Tok treniranja mreže možemo pratiti promatranjem funkcije gubitka iz epohe u epohu, što se vidi na grafu - slika 3.14. Primjeri su bili u serijama od 200, a treniranje je postavljeno na 25000 epoha, no sustav ga je prekinuo nakon 14000. Rezultati predviđanja istreniranog modela na skupu za testiranje prikazani su na slici 3.15.

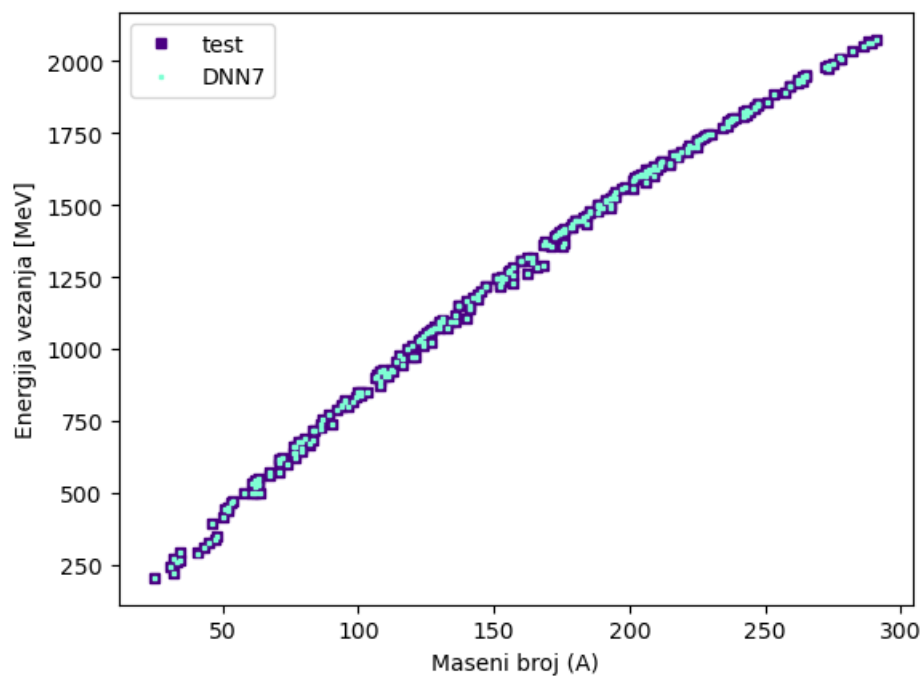


Slika 3.14: Graf funkcije gubitka po epohi treniranja neuronske mreže DNN7. Skokovi funkcije gubitka vidljivi su kada se za vrijeme treninga pojave velike vrijednosti parametara modela, penalizacijom istih gubitak se smanjuje. Treniranje je obustavljeno kada nije primijećeno poboljšanje modela, nakon otprilike 14000 epoha.

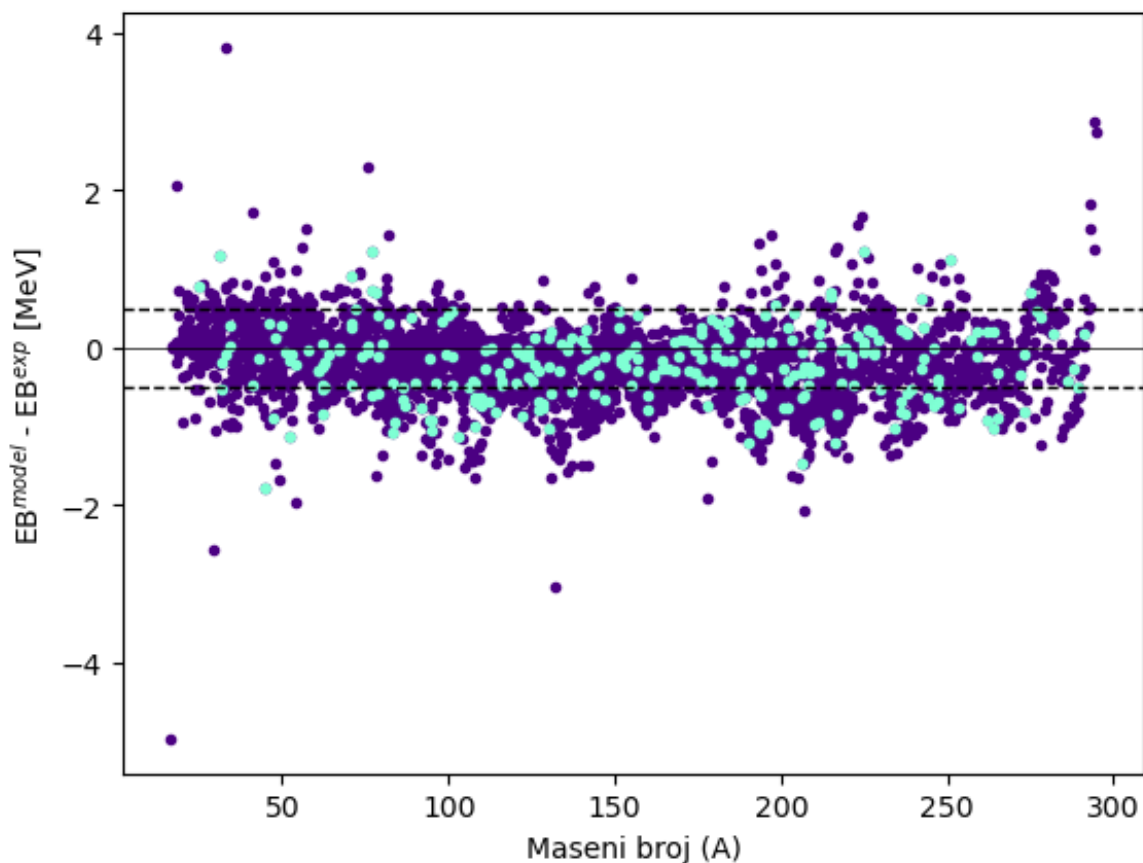
$RMS_{test} = 0.5202$ MeV , a ukupni RMS modela je $RMS_{tot} = 0.5063$ MeV. MAPE vrijednost iznosi 0.05.

Odstupanja od eksperimentalnih podataka prikazana su na grafu - slika 3.16. Kao i ranije lakše jezgre predstavljaju naviše problema pri modeliranju energije vezanja, iako su za ovaj model odstupanja mnogo manje nego za ekvivalentnu plitku mrežu, što među između ostalog poboljšava RMS modela.

Zanimljivo je primijetiti razliku između utjecaja značajki za plitki ANN7 model (slika 3.13) te duboki DNN7 model (slika 3.17). Duboko naučenom modelu dominantna značajka je atomski broj (Z), dok sve ostale značajke, osim broja neutrona (N), više doprinose modelu nego u slučaju plitke mreže.



Slika 3.15: Grafički prikaz predviđanja neuronske mreže DNN7 s 30x30x30x30 skrivenih neurona, trenirane kroz 25000 epoha uz stopu učenja 0.0003. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, RMS = 0.5063 MeV.



Slika 3.16: Razlika energija vezanja dobivenih DNN7 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.5 MeV.



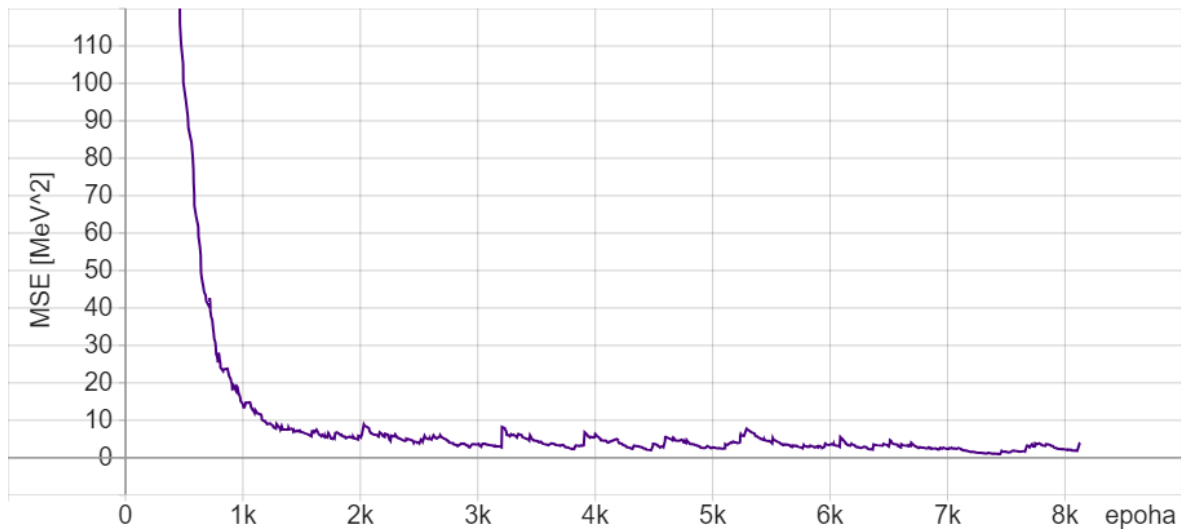
Slika 3.17: Rezultat SHAP analize utjecaja značajki na svojstva modela DNN7.

3.6 DNN8

Nadograđujemo model DNN7, dodavanjem značajke koja odgovara energiji asimetrije uslijed Paulijevog principa iz semi empirijske formule mase - jednačba (1.3)

$$E_{asy} = \frac{(Z - N)^2}{Z + N}. \quad (3.2)$$

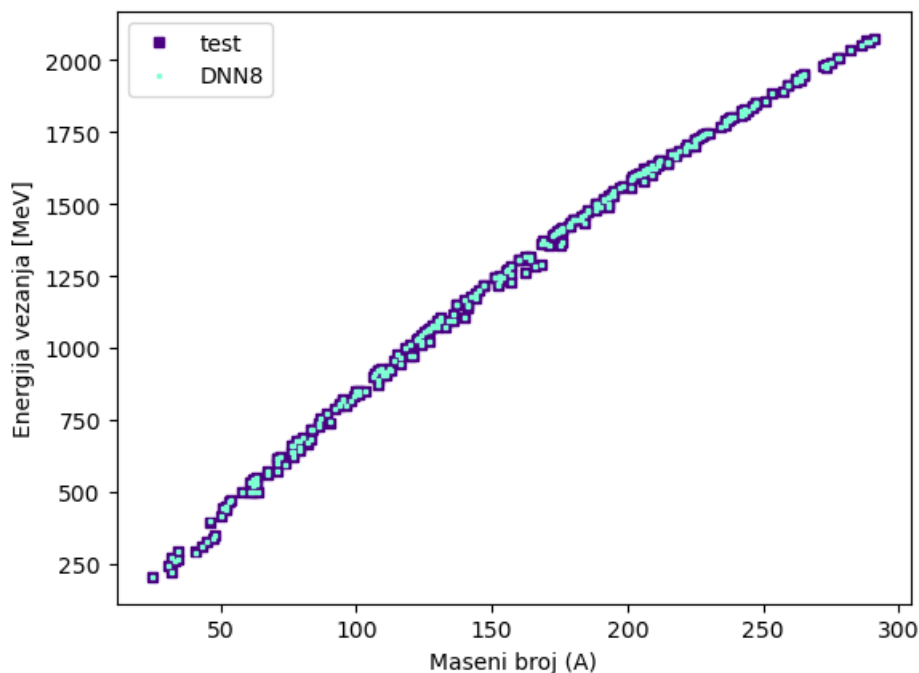
Ova duboka mreža ima 8 značajki: $Z, N, \Delta Z$ i $\Delta N, P, s_Z, s_N$ i Paulijev član. Treniranje je započeto s 4 skrivena sloja po 100 neurona (100x100x100x100), uz stopu učenja 0.0004 s propadanjem 0.999, a tijek treniranja kontroliran je praćenjem funkcije gubitka iz epohe u epohu - slika 3.18. Predviđeno je treniranje od 25000 epoha u serijama od 200 primjera. Algoritam za treniranje isto je prekinuo nakon 8000 epoha. Rezultati predviđanja istreniranog modela na skupu za testiranje prikazani su na slici 3.19.



Slika 3.18: Graf funkcije gubitka po epohi treniranja neuronske mreže DNN8. Treniranje je obustavljeno kada nije primijećeno poboljšanje modela, nakon otprilike 8000 epoha. Veće oscilacije funkcije gubitka primjećuju se otprilike svakih 1000 epoha, uz trend prema manjim vrijednostima.

$RMS_{test} = 0.3769$ MeV , a ukupni RMS modela je $RMS_{tot} = 0.3198$ MeV. MAPE vrijednost iznosi 0.03.

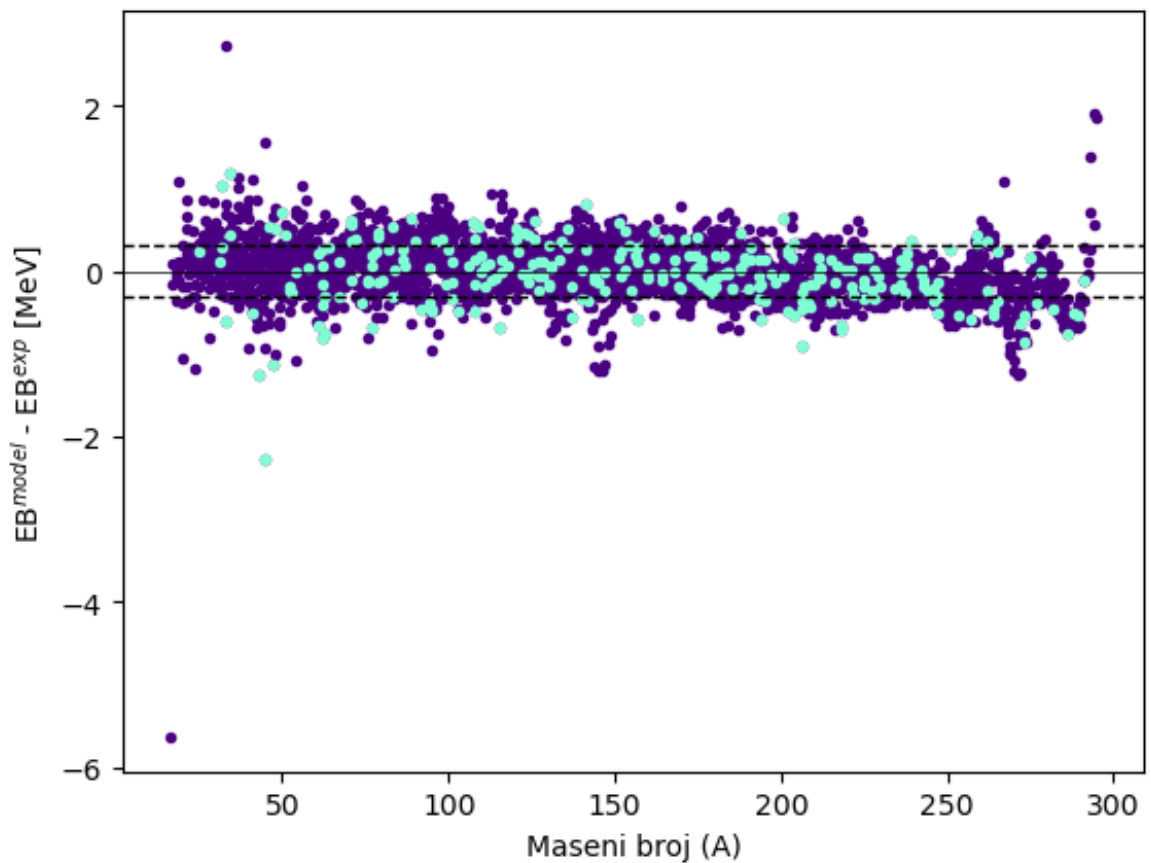
Na slici 3.20 vidimo grafički prikaz reziduala energija vezanja dobivenih DNN8 modelom i eksperimentalnih podataka za sve jezgre iz ukupnog skupa podataka. Posebno su istaknute jezgre iz skupa za testiranje. Vidljivo je da DNN8 bolje modelira energije vezanja. Većina reziduala upada u interval RMS-a najbolje istrenirane mreže. Jezgra koja najviše odstupa od eksperimentalnih podataka je ^{16}O , dvostruko



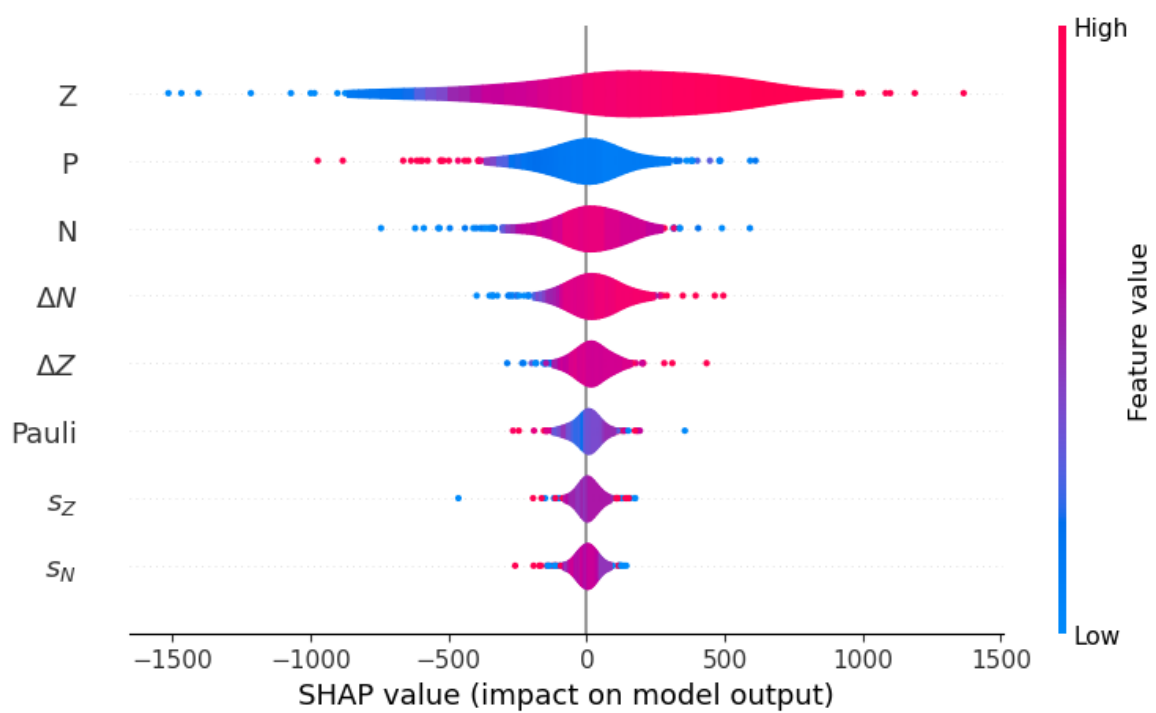
Slika 3.19: Grafički prikaz predviđanja neuronske mreže DNN8 s 100x100x100x100 skrivenih neurona, trenirane kroz 25000 epoha uz stopu učenja 0.0004. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, RMS = 0.3198 MeV.

magična jezgra, čiju je energiju vezanja model podcijenio za 5.6 MeVa.

Na slici 3.21 prikazana je analiza značajki modela. Utjecaj dodane značajke energije asimetrije bitan je za model otprilike koliko i mjera senioriteta protona i neutrona. Iz analiza Paulijeovog člana vidi se da za male vrijednosti, koje se postižu za jezgre sa simetričnom energetsom raspodjelom nukleona, on ne mijenja energiju vezanja. Utjecaj člana raste za ekstreme, kada ima mnogo više protona nego neutrona, ili obratno.



Slika 3.20: Razlika energija vezanja dobivenih DNN8 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.32 MeV.



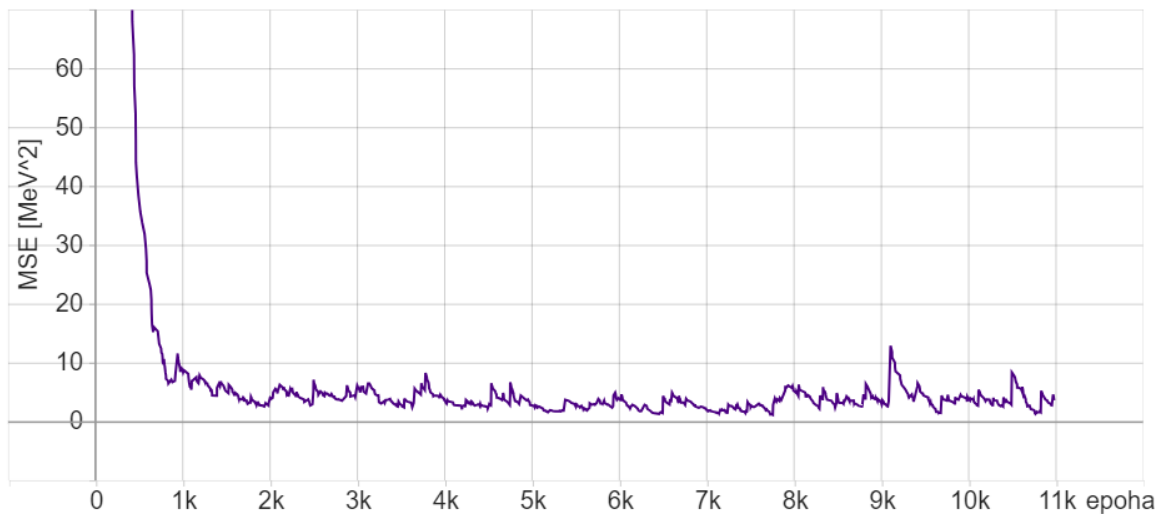
Slika 3.21: Rezultat SHAP analize utjecaja značajki na svojstva modela DNN8. Vidi se da su značajke modela prilično uravnotežene po doprinosu konačnom rezultatu, uz dominantni utjecaj broja protona (Z).

3.7 DNN9

Ranije se u procesu potrage za korisnim značajkama $A^{-2/3}$, član energije površine iz semi-empirijske formule (1.3) pokazao kao nekoristan, no atomske jezgre mogu imati difuzne površine, a nuklearna sila ima konačan doseg. Ovdje je uvedena nova značajka koja to uzima u obzir :

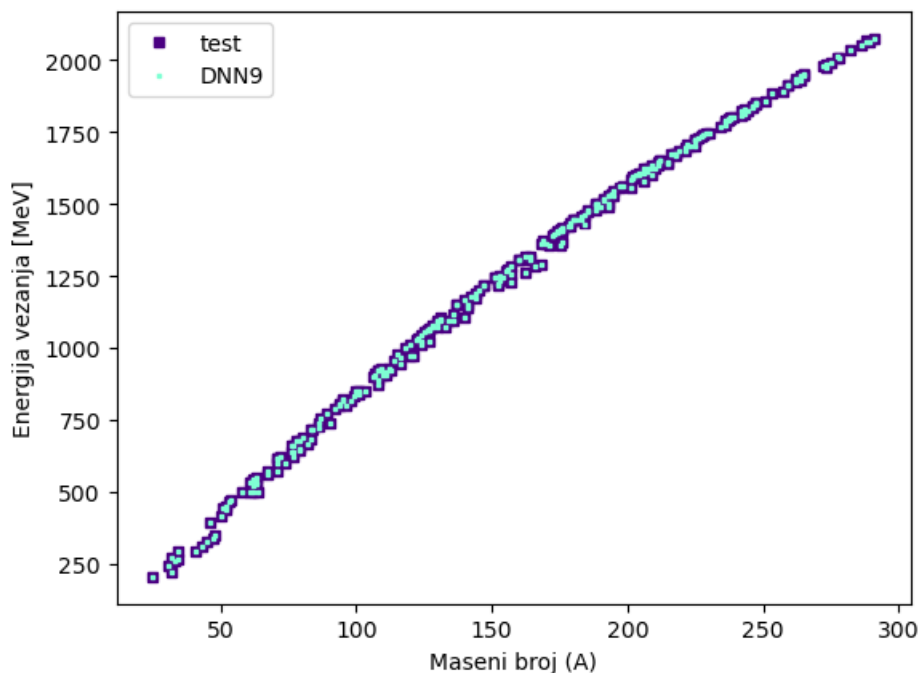
$$E_{sur}^* = \exp(-2A^{1/3})A^{-2/3}. \quad (3.3)$$

Model DNN9 prima devet-dimenzionalni vektor značajki, Z , N , $\exp(-2A^{1/3})A^{-2/3}$, ΔZ i ΔN , P , s_Z , s_N te Paulijev član. Konstruirana je mreža s 50 neurona u 4 skrivena sloja (50x50x50x50), kojoj je predviđeno treniranje kroz 25000 epoha uz stopu učenja 0.0001 te propadanje stope učenja 0.999. Veličina jedne serije bila je 200 primjera. Treniranje je obustavljeno nakon što se stopa učenja prestala smanjivati, nakon 11000 epoha, što se vidi na grafu - slika 3.22. Predviđanje istreniranog modela na skupu za testiranje prikazano je na slici 3.23.



Slika 3.22: Graf funkcije gubitka po epohi treniranja neuronske mreže DNN9. Nakon otprilike 11000 epoha sustav ranog zaustavljanja prekinuo je treniranje. Manje oscilacije funkcije gubitka prisutne su čim vrijednosti padnu ispod 10, no i uz njih se gubitak nastavlja smanjivati. Do većih oscilacija dolazi oko 9000 epoha, a pred kraj treniranja više se ne primjećuje poboljšanje.

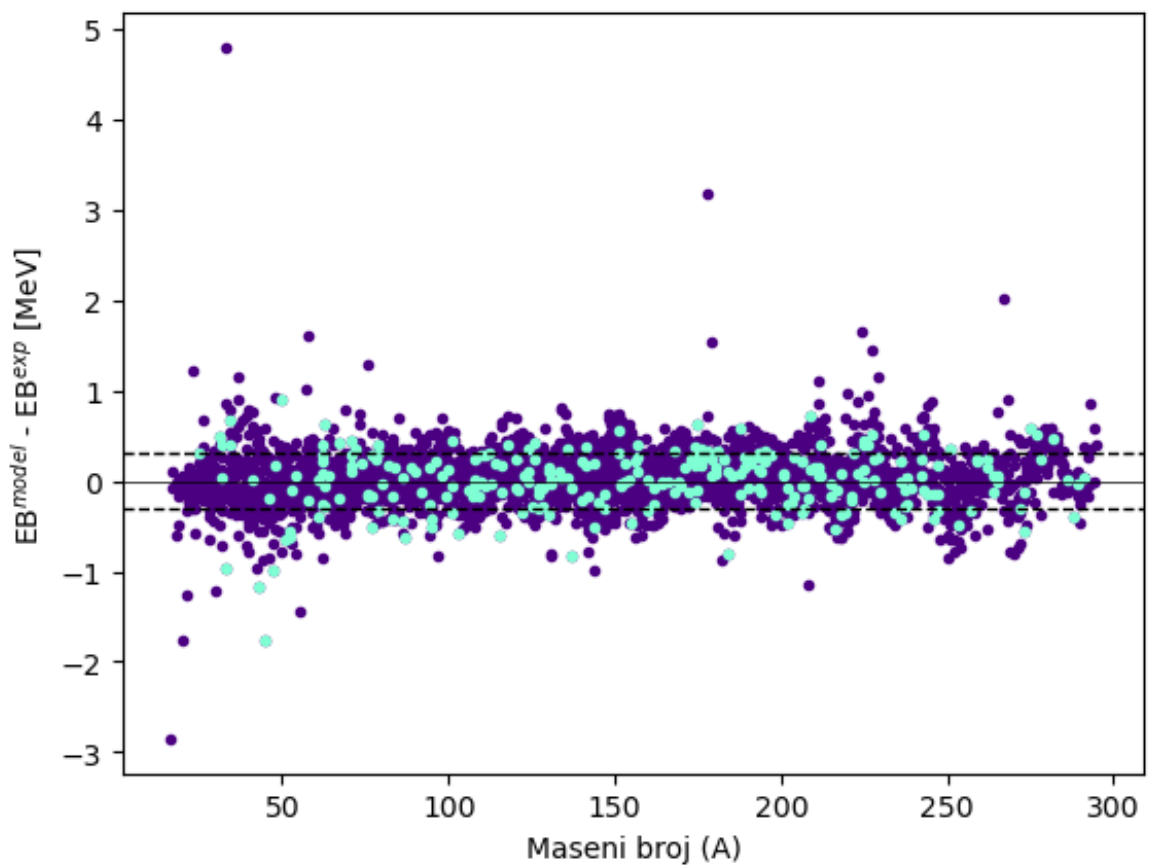
$RMS_{test} = 0.3240$ MeV , a ukupni RMS modela je $RMS_{tot} = 0.3067$ MeV. MAPE vrijednost iznosi 0.03.. Odstupanja od eksperimentalnih podataka prikazana su na grafu - slika 3.24. Kao i do sada lakše jezgre predstavljaju problem pri modeliranju energije vezanja, no odstupanja su manja. Atomske jezgre za koje model dobro



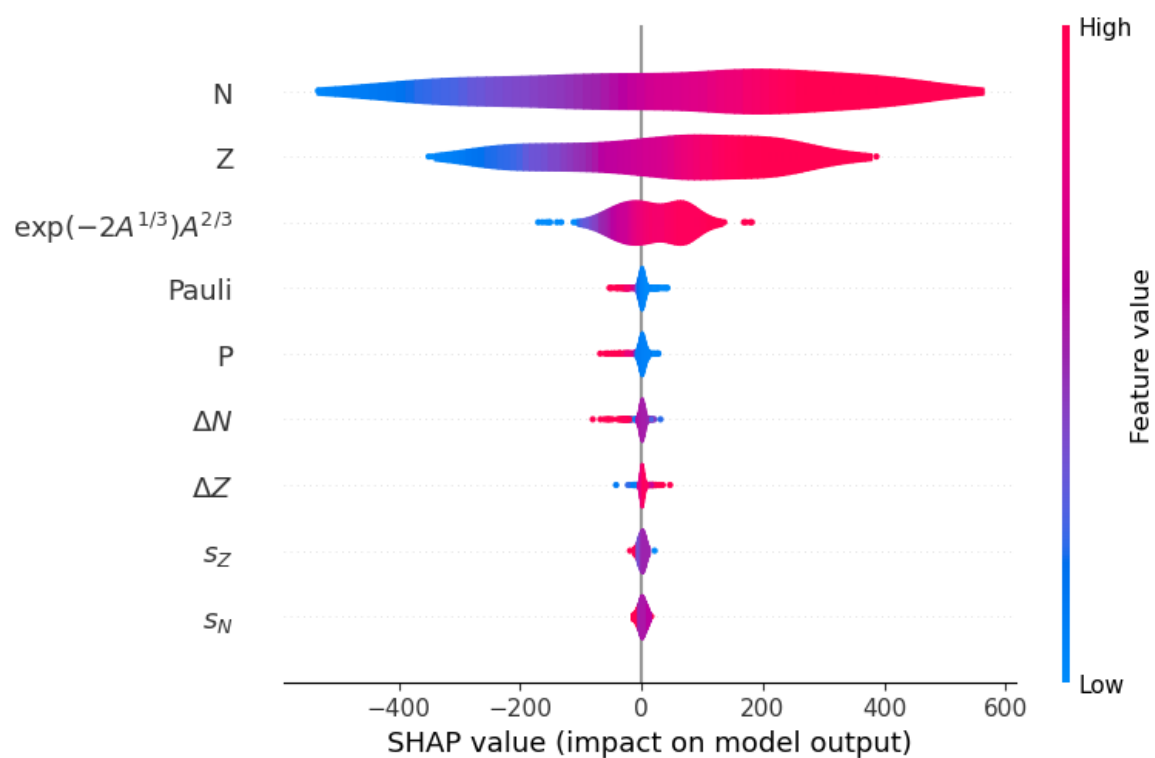
Slika 3.23: Grafički prikaz predviđanja neuronske mreže DNN9 s 50x50x50x50 skrivenih neurona, trenirane kroz 25000 epoha uz stopu učenja 0.0001. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, RMS = 0.3067 MeV.

predviđa energiju vezanja gusto su grupirane s rezidualima manjim od 1 MeV, a većina jezgara iz skupa za testiranje se nalazi unutar 300 keV-a.

Analiza značajki modela DNN9 prikazana je na slici 3.25. Vidi se da je nova značajka modificirane energije površine dobar izbor. Po utjecaju na model jedino su broj protona (N) i neutrona (Z) dominantniji od energije površine, a te su dvije značajke bile dominantne za, gotovo, svaki model. Jasno je da je energija vezanja veća što je energija površine veća. Modificirana energija površine opada eksponencijalno za teže jezgre čiji je radijus veći.



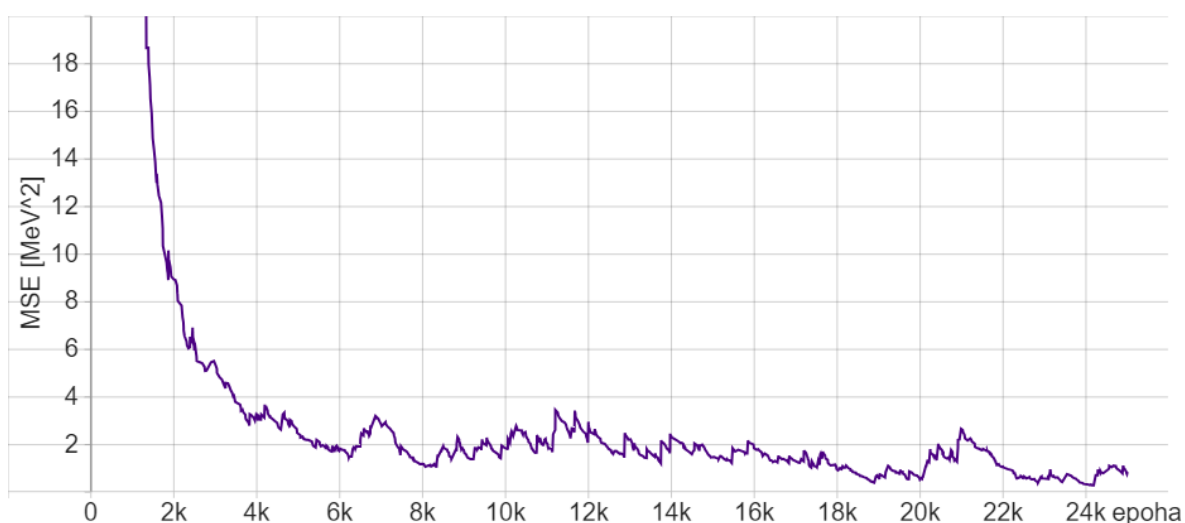
Slika 3.24: Razlika energija vezanja dobivenih DNN9 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.3 MeV.



Slika 3.25: Rezultat SHAP analize utjecaja značajki na svojstva modela DNN9. Uz broj protona (Z) i broj neutrona (N), novo dodana značajka energije površine najznačajnije doprinosi konačnom rezultatu.

3.8 DNN10

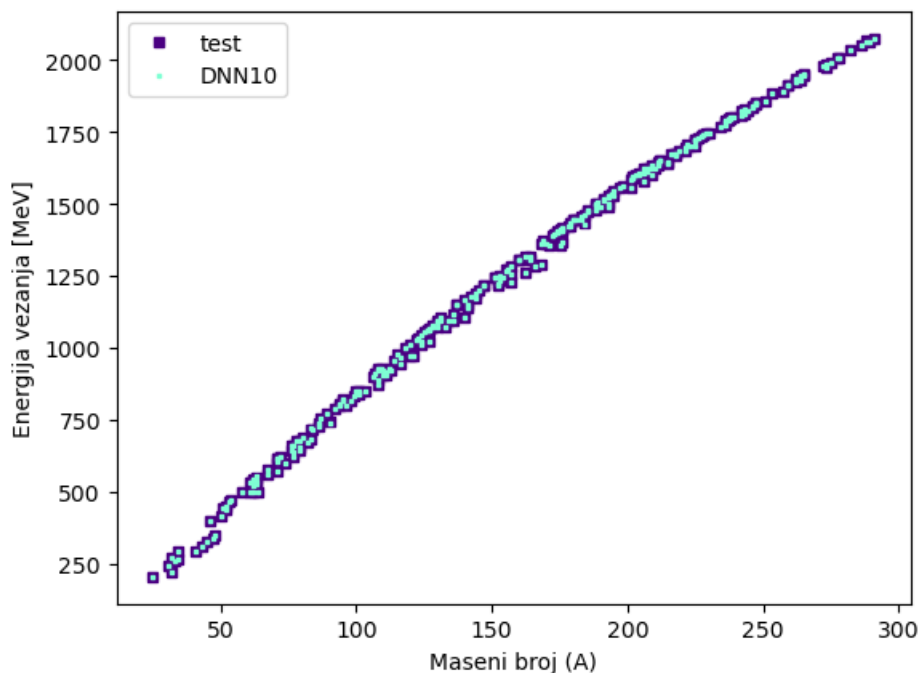
Ovaj duboki model sadrži deset značajki, uz sve uspješne vrijednosti prošlih modela, DNN10 obogaćen je znanjem Wignerove energije, dodana je značajka Wigner koja se računa kao (1.11). Ulazne značajke sada su Z , N , $\exp(-2A^{1/3})A^{-2/3}$, ΔZ i ΔN , P , s_Z , s_N te Paulijev i Wignerov član. Treniranje je započeto s 4 skrivena sloja po 100 neurona (100x100x100x100), uz stopu učenja 0.0003 s propadanjem 0.999. Predviđeno je treniranje od 25000 epoha u serijama od 200 primjera. Tijek treniranja kontroliran je praćenjem funkcije gubitka iz epohe u epohu - slika 3.26. Rezultati predviđanja istreniranog modela na skupu za testiranje prikazani su na slici 3.27.



Slika 3.26: Graf funkcije gubitka po epohi treniranja neuronske mreže DNN10. Vidi se da učenjem model vrlo brzo konvergira ka vrijednostima funkcije gubitka ispod 2, no nakon toga one počinju oscilirati. Kako se model mijenja tokom učenja gubitak raste, ali ne prelazi 4, dok se pred kraj treniranja, oko 24000 epoha ne postigne minimalna vrijednost.

$RMS_{test} = 0.2686$ MeV, a ukupni RMS modela je $RMS_{tot} = 0.2476$ MeV. MAPE vrijednost iznosi 0.02.

Slika 3.28 prikazuje odstupanja predviđanja neuronske mreže i eksperimentalnih vrijednosti. Svega desetak jezgara odstupa za više od 1 MeV-a energije, dok samo dvije odstupaju za više od 2 MeV-a. Jezgra ^{26}Al kojoj je model podcijenio energiju vezanja za 2.65 MeV-a te jezgra ^{33}Ca kojoj je precijenio energiju za 2.5 MeV-a. Sa grafa se vidi da je najviše odstupanja oko jezgara masenog broja $A = 30$, odnosno oko jezgre prije značajnijih magičnih brojeva. Također se vidi da model precjenjuje energije vezanja za teže jezgre, blizu $A = 300$. Model najbolje funkcionira između $A = 100$ i $A = 200$ gdje gotovo svi reziduali upadaju u interval 500 keV-a, a većina je

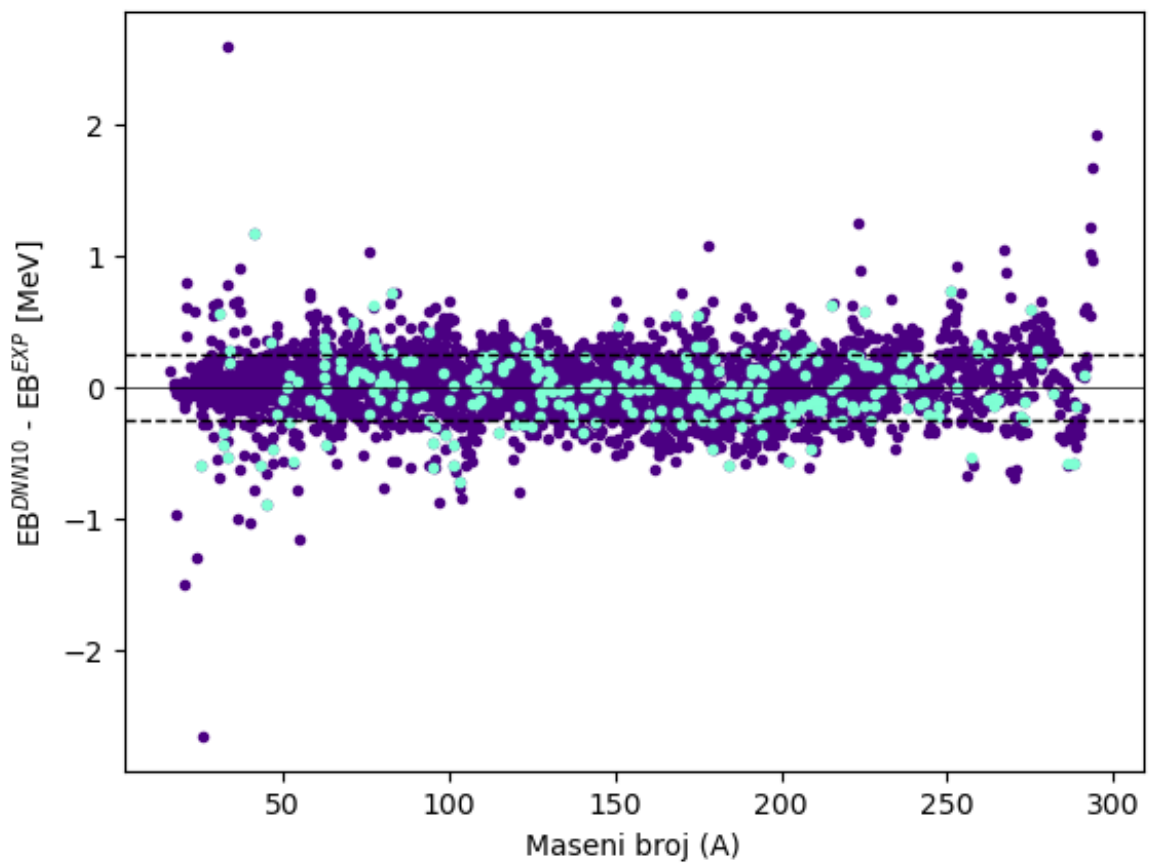


Slika 3.27: Grafički prikaz predviđanja neuronske mreže DNN10 s skrivena neurona, trenirane kroz 25000 epoha uz stopu učenja 0.0003. Graf usporedbe skupa za testiranje i predviđanja neuronske mreže, RMS = 0.2476 MeV.

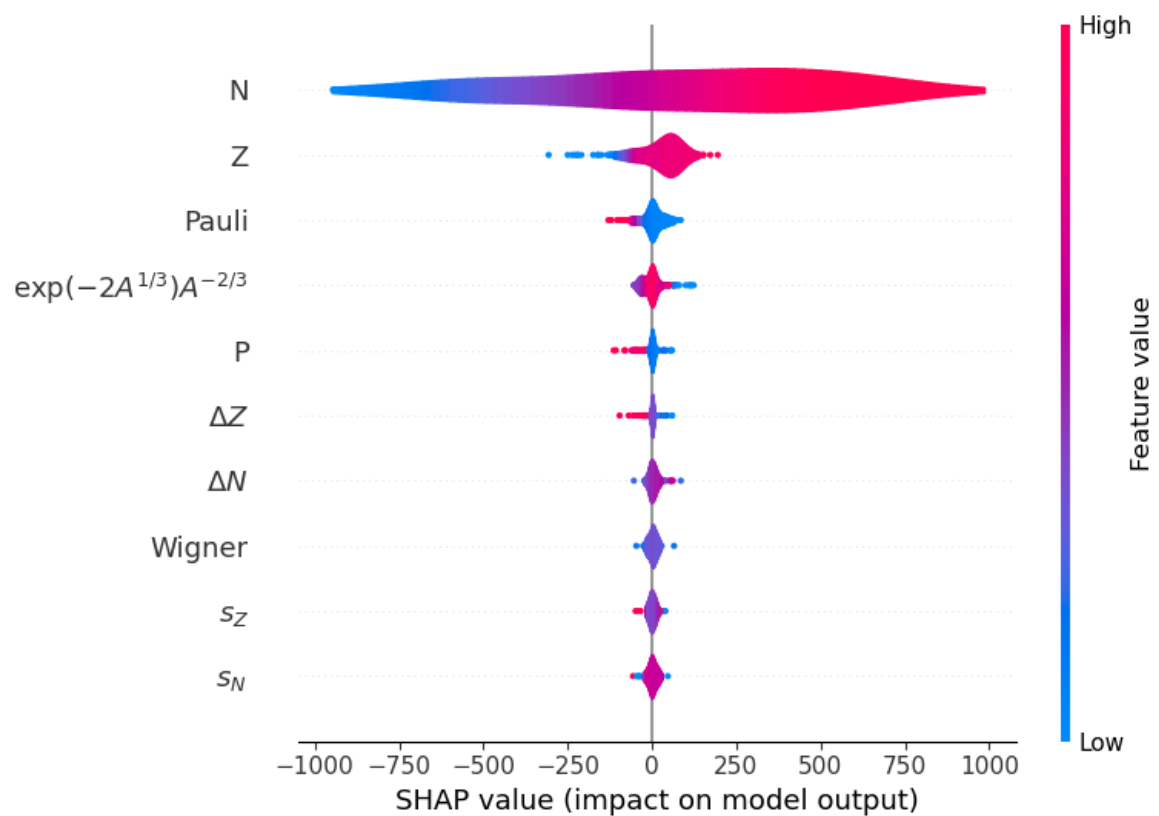
grupirana unutar 250 keV-a.

Provedena SHAP analiza modela pokazala je da broj neutrona (N) dominira među odabranim značajkama, dok ostale doprinose mnogo manje - slika 3.29. Broj protona (Z) je i dalje druga najrelevantnija značajka, a Wignerova energija dodana u ovom modelu doprinosi koliko i broj valentnih neutrona (ΔN). Kada je Wignerov član veći, doprinos energiji vezanja je manji, izospinska asimetrija je veća i nema proton-neutron sparivanja. Za manje vrijednosti Wignerovog člana, blizu smo $N = Z$ i doprinos energiji vezanja je veći.

Ovaj model postigao je najbolje rezultate od svih treniranih. Pregled najuspješnijih istreniranih modela u ovom radu prikazan je u tablici 3.3.



Slika 3.28: Razlika energija vezanja dobivenih DNN10 mrežom i eksperimentalnih podataka, za atomske jezgre iz ukupnog skupa podataka. Posebno su istaknute, plavom bojom, vrijednosti za jezgre iz skupa za testiranje. Crtkane linije na grafu označavaju odstupanje energije vezanja za 0.25 MeV.

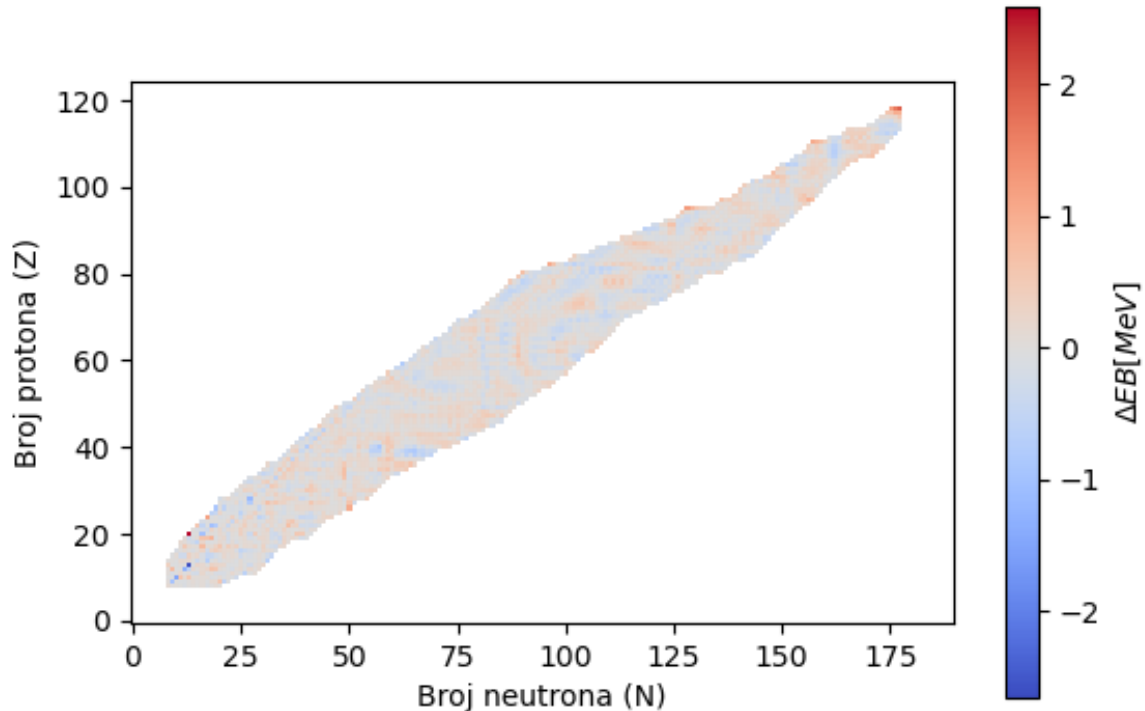


Slika 3.29: Rezultat SHAP analize utjecaja značajki na svojstva modela DNN10. U ovome istreniranom modelu najveći doprinos ima broj neutrona (N), dok ostale značajke, više fino ugađaju predviđanja neuronske mreže.

Tablica 3.3: Pregled treniranih modela uz njihove RMS i MAPE vrijednosti.

model	RMS_{test} [MeV]	RMS_{tot} [MeV]	RMS_{train} [MeV]	MAPE
ANN2	1.5644	1.5434	1.5317	0.15
ANN4	1.1027	1.0046	0.9244	0.11
ANN5	0.9683	0.8912	0.8066	0.10
ANN7	0.7145	0.5828	0.4325	0.06
DNN7	0.5202	0.5063	0.4690	0.05
DNN8	0.3769	0.3198	0.2712	0.03
DNN9	0.3240	0.3067	0.2536	0.03
DNN10	0.2686	0.2476	0.2095	0.02

Predviđanja najbolje istreniranog DNN10 modela kroz cijelu mapu nuklida prikazana su na slici 3.30, a devijacije od eksperimentalnih mjerenja energija vezanja označene su bojom.

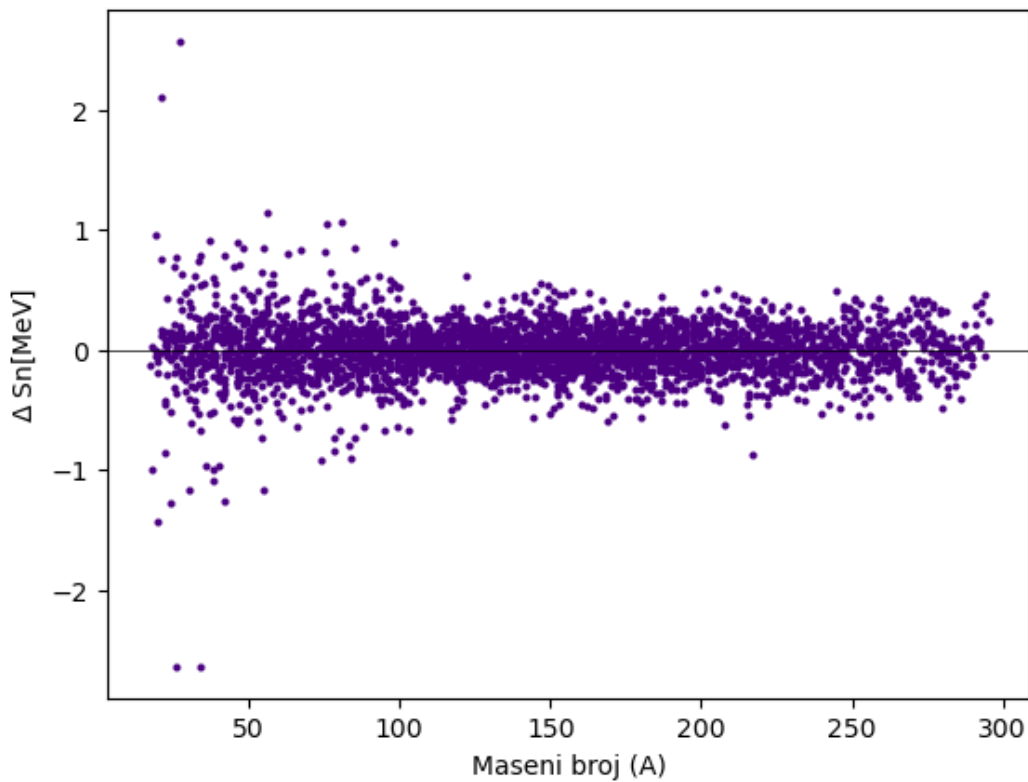


Slika 3.30: Mapa nuklida s atomskim jezgrama obojenima ovisno o odstupanju predviđanja DNN10 modela od eksperimentalnih podataka. Za plavo obojene jezgre model podcjenjuje energije vezanja, a za crveno obojene precjenjuje.

Koristeći istreniranu mrežu DNN10 moguće je izračunati vrijednosti, energetskog procijepa sparivanja i energije separacije nukleona te ih usporediti s poznatim, tabličnim podacima. Podaci su preuzeti iz NNDC(National Nuclear Data Center) - NuDat3 online baze podataka koju održava Brookhaven National Laboratory [42].

Energije separacije neutrona izračunate su korištenjem energija vezanja dobivenih modelom DNN10 i formule (1.4). Razlika tih vrijednosti i poznatih energija

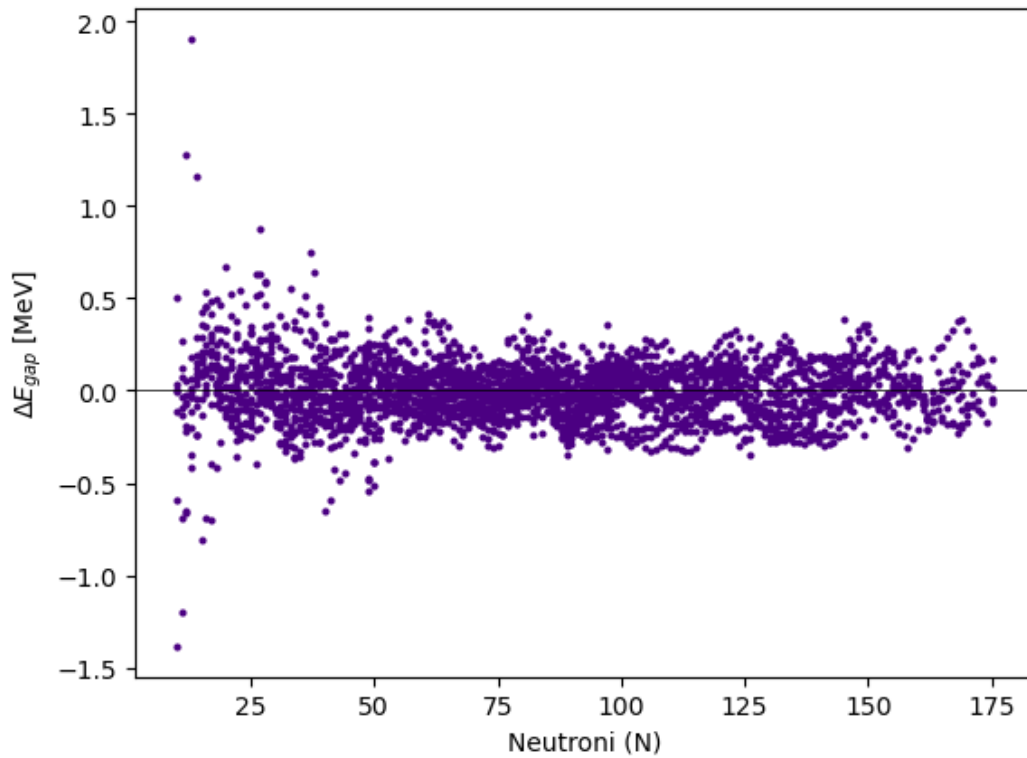
separacije iz NuDat3 vodi na RMS odstupanje od 236 keV-a. Slika 3.31 prikazuje odstupanja predviđanja modela i dostupnih podataka.



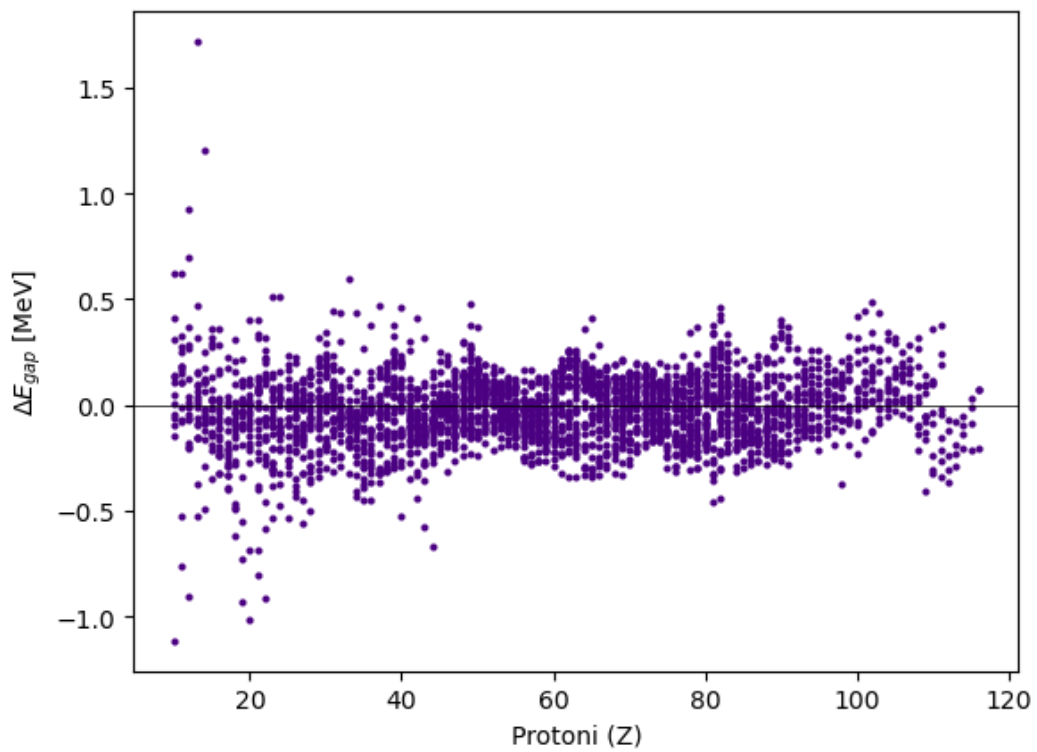
Slika 3.31: Odstupanja energija separacije neutrona dobivena predviđanjem neuronske mreže DNN10 i podataka dostupnih iz NuDat3 baza podataka. $RMS(S_n) = 236$ keV.

U radu [43] korišteno je strojno učenje kako bi se minimizirali reziduali FRDM i eksperimentalnih mjerenja, takav hibridni model postigao je $RMS(S_n) = 350$ keV, za isti skup podataka model DNN10 postiže $RMS(S_n) = 235$ keV.

Vrijednosti energetskog procijepa za protone i neutrone izračunate su five-point formulom (1.10). Izračunato je RMS odstupanje s podacima preuzetima iz NuDat3 te iznosi 166 keV-a za neutrone i 174 keV-a za protone. Graf devijacija za neutrone prikazan je na slici 3.32, a za protone na slici 3.33.



Slika 3.32: Odstupanja energetskeg procijepa za neutrone dobivena predviđanjem neuronske mreže DNN10 i podataka dostupnih iz NuDat3 baza podataka. $\text{RMS}(E_{gap,n}) = 166 \text{ keV}$.



Slika 3.33: Odstupanja energetskeg procijepa za protone dobivena predviđanjem neuronske mreže DNN10 i podataka dostupnih iz NuDat3 baza podataka. $\text{RMS}(E_{gap,p}) = 174 \text{ keV}$.

Usporedba učinkovitosti predviđanja energija vezanja za eksperimentalno dostupne atomske jezgre, iz AME2020 baze podataka, raznih tradicionalnih modela i strojno naučene neuronske mreže prikazana je u tablici 3.4. SkM*, UNEDF0 i UNEDF1 modeli su zasnovani na energetsom funkcionalu gustoće (engl. *energy density functional* - EDF) i funkcioniraju za sve jezgre. DD-ME2 i DD-PC1 relativističke su interakcije dobre za računanje svojstava parno-parnih jezgara pomoću Hartree-Fock-Bogoljubov modela. HFB-31 je Hartree-Fock-Bogoljubov model s uključenim sparivanjem, koji koristi Skyrme interakcije u računu srednjeg polja, primjenjiv na sve atomske jezgre. ANN je strojno naučeni model koji koristi parametre Weizsäcker-Skyrme WS4 modela.

Tablica 3.4: Usporedba RMS vrijednosti istreniranih neuronskih mreža s drugim teorijskim modelima za isti skupove atomskih jezgara. Za usporedbu modela HFB-31 korišten je skup atomskih jezgara iz baze AME2012 [48, 49]. Za ostale modele korišteni su podaci iz AME2020 baze nuklearnih podataka. DD-ME2 i DD-PC1 modeli su uspoređeni samo za parno-parne jezgre.

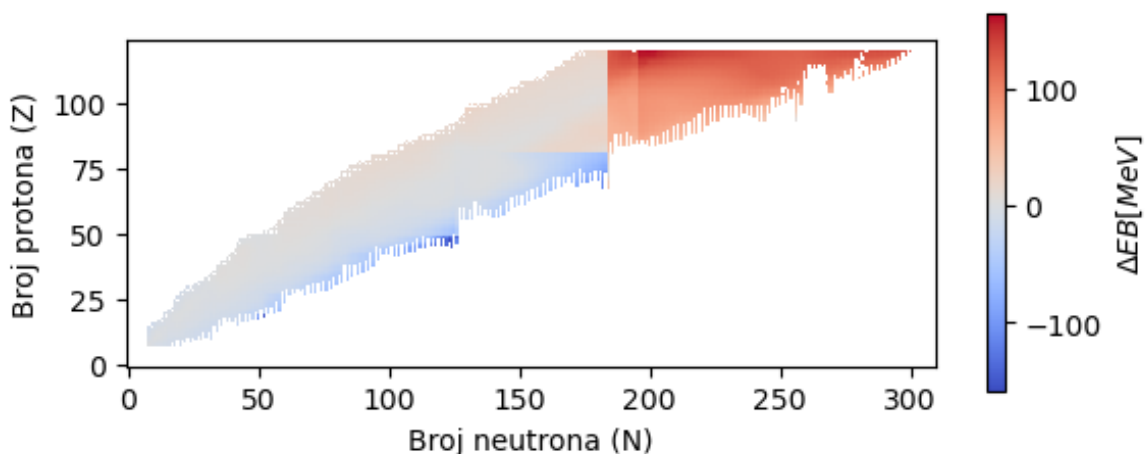
model	RMS [MeV]
DNN10	0.2476
SkM* [9]	7.752
UNEDF0 [45]	1.597
UNEDF1 [46]	2.105
ANN [47]	0.159
HFB-31 [44]	0.561
DNN10 (pp)	0.2319
DD-ME2 [11]	2.458
DD-PC1 [11]	2.198

Osim testiranja na eksperimentalno dostupnim jezgrama, moguće je neuronsku mrežu usporediti s drugim modelima za predviđanje mape nuklida, što je vrlo korisno u nuklearnim istraživanjima. Podaci energija vezanja za gore navedene modele preuzeti su s MassExplorer internetske baze podataka [50]. Svakim modelom izračunate su energije vezanja za što više atomskih jezgara, kako bi se popunila mapa nuklida između granica vezanja (engl. *drip line*) [51]. Veličina skupova jezgara i izračunatih energija vezanja predstavljena je u tablici 3.5. Za sve jezgre iz pripadnog skupa izračunata je energija vezanja modelom DNN10 te su vrijednosti uspoređene i prikazane na mapi nuklida, odstupanja su vizualizirana bojom jezgre. Odstupanja neuronske mreže i EDF modela SkM* prikazan je na slici 3.34. Slike 3.35 i 3.36 prikazuju odstupanja za modele DD-ME2 i DD-PC1 respektivno, a slike 3.37 i 3.38 odstupanja za UNEDF0 i UNEDF1 respektivno.

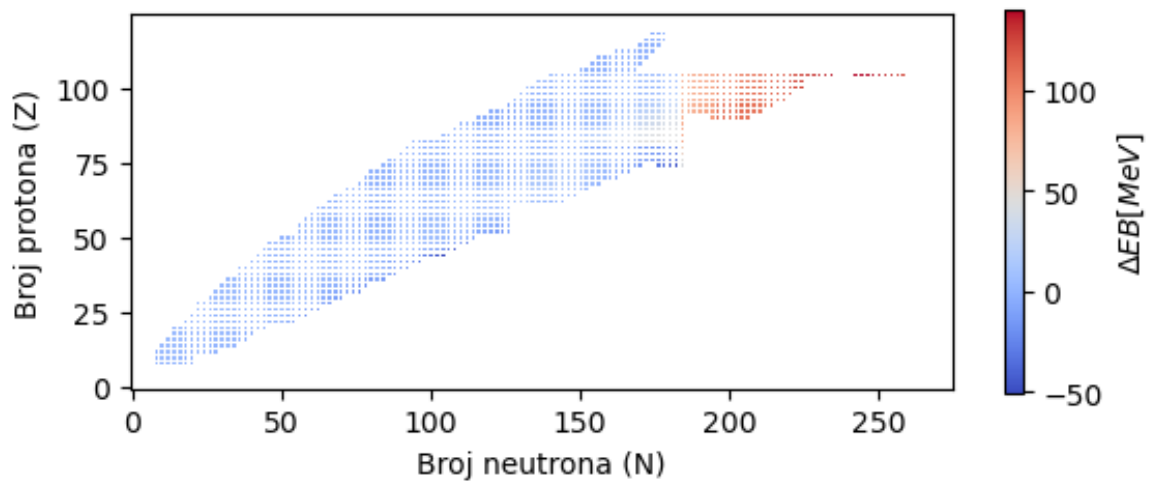
Tablica 3.5: Popis modela i veličina njihovih skupova podataka za energije vezanja. DD-ME2 i DD-PC1 skupovi sadrže samo parno-parne jezgre.

model	skup jezgara
SkM* [9]	8639
DD-ME2 [11]	1684
DD-PC1 [11]	1745
UNEDF0 [45]	8243
UNEDF1 [46]	8243

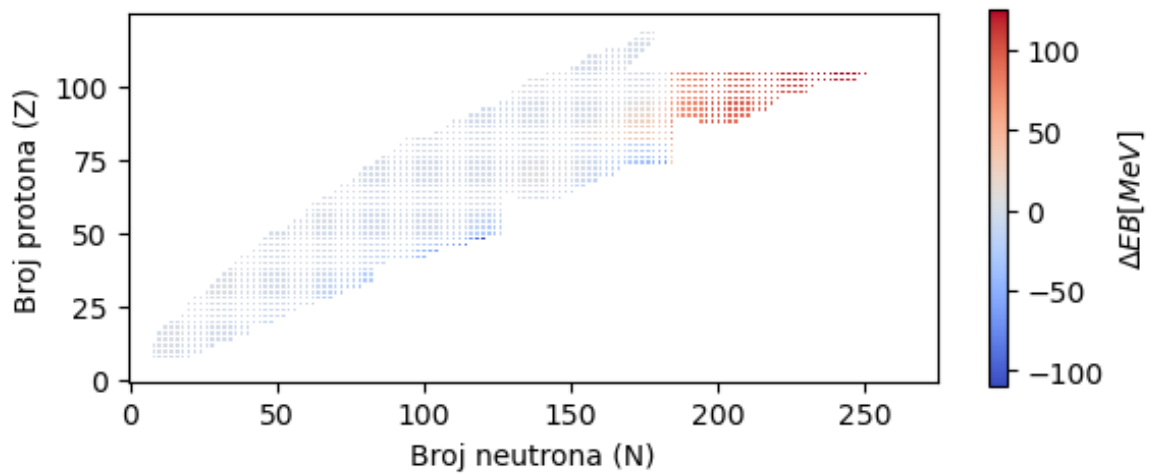
Za sve uspoređene modele vidi se isti problem za jezgre bogate neutronima. Moguće je da tradicionalni modeli podcjenjuju energije vezanja, no vjerojatnije je objašnjenje da neuronska mreža precjenjuje energije jer su u nju ugrađene informacije o strukturi ljusaka ograničene s magičnim brojem 196 za neutrone. Na svim grafovima se vidi da diskontinuitet u energijama vezanja počinje oko $N = 200$. Situacija bi se popravila nadogradnjom neuronske mreže s više magičnih brojeva. Za sve jezgre s manje od 200 neutrona mreža podcjenjuje energije vezanja u odnosu na sve modele. Takva situacija je očekivana. Pogledom na tablicu 3.4 vidimo da je RMS devijacija svakog modela veća od RMS devijacije neuronske mreže. Budući da neuronska mreže bolje predviđa energije vezanja na poznatom području mape nuklida od tradicionalnih modela jasno je da će u usporedbi s njima izgledati kao da mreža podcjenjuje energije vezanja. Ekstremna odstupanja modela i neuronske mreže prisutna su uglavnom na rubovima mape nuklida, blizu protonskih i neutronskih drip linija.



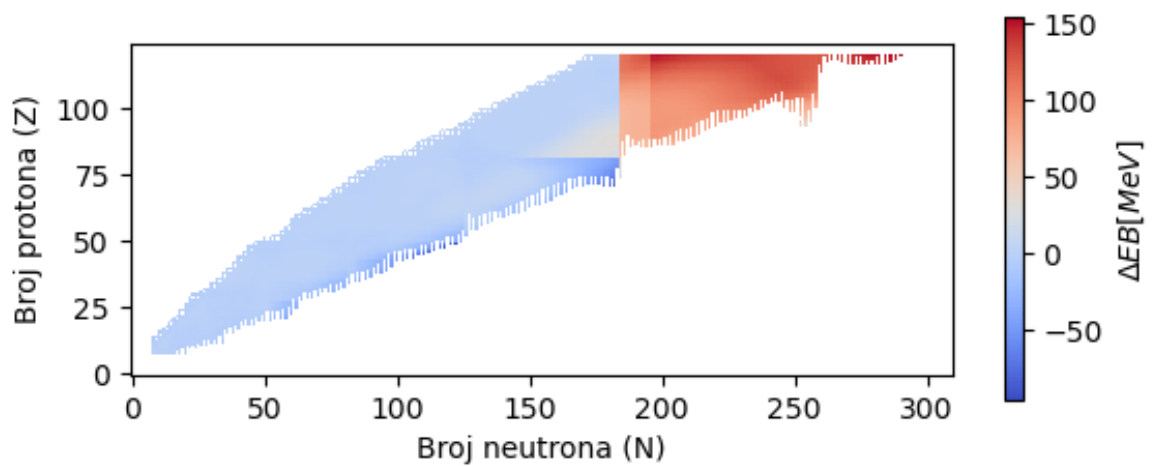
Slika 3.34: Mapa nuklida s atomskim jezgrama obojenima ovisno o odstupanju predviđanja energija vezanja neuronske mreže DNN10 od modela SkM*. Za plavo obojene jezgre neuronska mreža podcjenjuje energije vezanja modela, a za crveno obojene precjenjuje.



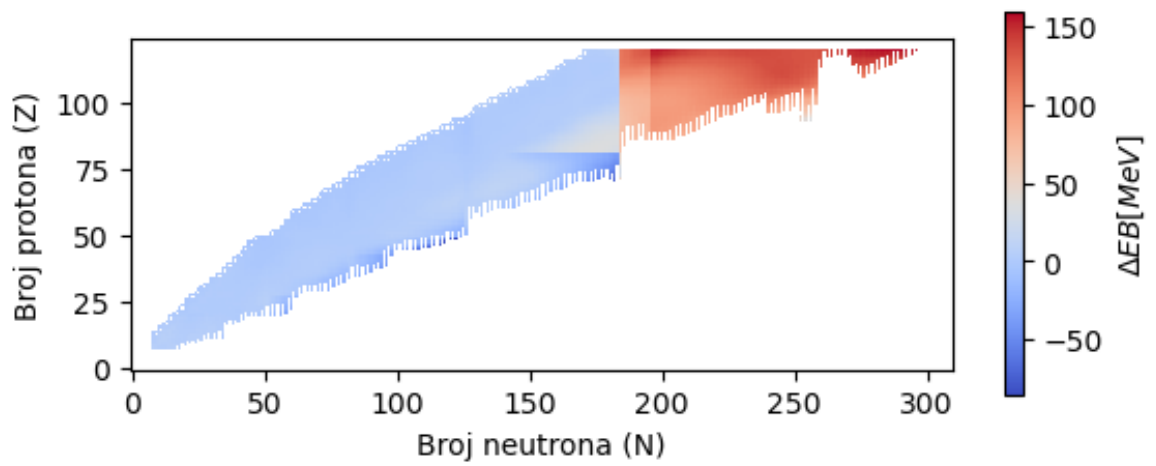
Slika 3.35: Isto kao slika 3.34, ali za odstupanja modela DNN10 i DD-ME2.



Slika 3.36: Isto kao slika 3.34, ali za odstupanja modela DNN10 i DD-PC1.



Slika 3.37: Isto kao slika 3.34, ali za odstupanja modela DNN10 i UNEDF0.



Slika 3.38: Isto kao slika 3.34, ali za odstupanja modela DNN10 i UNEDF1.

4 Zaključak

S obzirom na rezultate dobivene treniranjem neuronskih mreža raznih arhitektura s mnogim kombinacijama hiperparametara jasno je da je strojno učenje vrlo perspektivno područje istraživanja računalne fizike i da su neuronske mreže primjenjive u nuklearnoj fizici. S trenutno razvijenom tehnologijom strojno je učenje u stanju konkurirati tradicionalnim globalnim modelima energija vezanja.

Mreže konstruirane u ovome radu koristile su 2, 4, 5, 7, 8, 9 i 10 ulaznih značajki, koje su pažljivo odabrane i isprobane. Model koji je pokazao najbolje rezultate s RMS odstupanjem od eksperimentalnih vrijednosti energija vezanja koje iznosi 248 keV-a je onaj s 10 ulaznih značajki. One su Z , N , $\exp(-2A^{1/3})A^{-2/3}$, ΔZ i ΔN , P , s_Z , s_N . Redom : broj protona i neutrona, energija površine, broj valentnih protona i neutrona, nukleonski faktor promiskuiteta, senioritet protona i neutrona te Paulijev i Wignerov član. Dodavanjem još dobro odabranih značajki, model bi se potencijalno mogao poboljšati i dalje. Proširivanjem modela na sve jezgre bilo bi poželjno dodati značajke koji uključuju efekte deformacije jezgre te sljedeće, više magične brojeve za protone i neutrone kada se radi o superteškim jezgrama. Pri razvoju modela baziranih na neuronskim mrežama dobar odabir značajke modela je prioritet, pogotovo, jer su one, u sustavima s ograničenim količinama podataka ključne za uspješnost modela.

Promatranjem utjecaja odabira dodatnih značajki na učinkovitost modela, provođenjem SHAP analize za istrenirane modele, možemo utvrditi koliko su te veličine relevantne za nuklearne interakcije. Duboko učenje moglo bi dati uvide u do sada nepoznate čimbenike nuklearne sile što bi pospješilo razvoj novih i boljih nuklearnih modela ili poboljšanje postojećih [52].

Modeli bazirani na neuronskim mrežama, koji dosljedno opisuju atomske jezgre u nepoznatim dijelovima mape nuklida, mogli bi biti iznimno korisni u modeliranju astrofizičkih procesa i simulacijama eksperimenata.

Najbolje istrenirani model DNN10 postigao je bolje rezultate u reproduciranju energija vezanja nego HFB-31 model zasnovan na efektivnoj Skyrme interakciji (vidi tablicu 3.4 za usporedbu RMS-a modela strojnog učenja i tradicionalnih modela). Dobro istrenirana neuronska mreža, idealno, bolje ekstrapolira podatke atomskih jezgara jer se radi o univerzalnom aproksimatoru, dok modeli srednjeg polja mogu

divergirati u neistraženom području.

Na mapi nuklida koja prikazuje odstupanja između tradicionalnih modela i strojno naučene neuronske mreže DNN10, vidi se da u području teških jezgara, bogatih neutronima, imamo velika odstupanja od do 100 MeV-a u energijama vezanja. Budući da taj dio mape nuklida nije još eksperimentalno dostupan nije odmah jasno koji modeli bi bili točniji.

Obične plitke i duboke neuronske mreže kakve su korištene u ovom radu samo su neke od novijih metoda u analizi podataka koje se pokazuju korisne u nuklearnoj fizici. Bolji modeli predviđanja energija vezanja mogli bi iskoristiti prednosti Bayesovog dubokog učenja, koje uz predviđanje očekivane vrijednosti daje i mjeru neodređenosti [53].

Na kraju, bitno je spomenuti da, iako se u ovom radu strojno učenje koristilo za modeliranje energija vezanja atomskih jezgara, princip konstrukcije i treniranja neronskih mreža jednak je za modeliranje drugih opservabli. Uz dovoljno veliki skup podataka za treniranje i inteligentno konstruiranu mrežu s prigodnim odabirom značajki rezultat će biti dobro istrenirana mreža, u stanju da predviđa sva nuklearna svojstva od interesa. U principu je moguće razviti neuronsku mrežu koja bi iz ulaznog vektora značajki predviđala ne samo jednu, već više vrijednosti za atomske jezgre: energiju vezanja, stabilnost, vrijeme poluživota. Ovo bismo postigli provođenjem multivarijantne regresije pomoću dubokog učenja.

Literatura

- [1] Samuel S. M. Wong : Introductory Nuclear Physics, 2nd edition, Weinheim : WILEY-VCH Verlag GmbH Co. KGaA, 2004.
- [2] John J. Cowan; Friedrich-Karl Thielemann: R-Process Nucleosynthesis in Supernovae, *Physics Today*. 57, str. 47-53. 2004.
- [3] A. W. Steiner et al. : Core-Collapse Supernova Equations Of State Based On Neutron Star Observations, <https://arxiv.org/abs/1207.2184>, 2012.
- [4] Yuri T. Oganessian; A. Sobiczewski; G. M. Ter-Akopian : Superheavy nuclei: from predictions to discovery, *Physica Scripta*. 92, 023003, 2017.
- [5] John R. Lamarsh, : Introduction To Nuclear Reactor Theory, Addison-Wesley Publishing Company, 1966.
- [6] Jassem Abbasi; Pål Østebø Andersen : Physical Activation Functions (PAFs): An Approach for More Efficient Induction of Physics into Physics-Informed Neural Networks (PINNs), <https://arxiv.org/abs/2205.14630>, 2023.
- [7] Kenneth S. Krane : Introductory Nuclear Physics, John Wiley & Sons, New York, 1988.
- [8] J. Carlson; S. Gandolfi; F. Pederiva; Steven C. Pieper; R. Schiavilla; K.E. Schmidt; R.B. Wiringa : Quantum Monte Carlo methods for nuclear physics, *Rev. Mod. Phys.* Vol. 87, 1067, 2015.
- [9] J. Bartel; P. Quentin; M. Brack; C. Guet; H. B. Håkansson : Towards a better parametrisation of Skyrme-like effective forces: A critical study of the SkM force, *Nuclear Physics A* 386, str. 79-100. 1982.
- [10] Carlos A. Bertulan : Nuclear Physics in a Nutshell, Princeton University Press, Princeton, New Jersey, 2007.
- [11] T. Nikšić; N. Paar; D. Vretenar; P. Ring : DIRHB—A relativistic self-consistent mean-field framework for atomic nuclei, *Computer Physics Communications*, Vol. 185, Issue 6, str. 1808-1821, 2014.

- [12] K. Nomura; N. Shimizu; T. Otsuka : Mean-Field Derivation of the Interacting Boson Model Hamiltonian and Exotic Nuclei, Phys. Rev. Lett. 101, 142501, 2008.
- [13] Petr Navratil, Ab initio no-core shell model calculations for light nuclei, <https://arxiv.org/abs/0711.2702>, 2007.
- [14] W.J. Huang; Meng Wang; F.G. Kondev; G. Audi; S. Naimi : The AME 2020 atomic mass evaluation (I). Evaluation of input data, and adjustment procedures, Chinese Physics C, Vol. 45, 030002, 2021.
- [15] W.J. Huang; Meng Wang; F.G. Kondev; G. Audi; S. Naimi : The AME 2020 atomic mass evaluation (II). Tables, graphs and references, Chinese Physics C, Vol. 45, 030003, 2021.
- [16] Jouni Suhonen : From Nucleons to Nucleus, Springer-Verlag Berlin, Heidelberg 2007.
- [17] I. Angeli : Effect of valence nucleons on RMS charge radii and surface thickness, J. Phys. G Nucl. Part. Phys. 17, str. 439-454. 1991.
- [18] R. Neugart; A. Wöhr : Nuclei Far From Stability and Atomic Masses and Fundamental Constants 1992, CRC Press, 1993.
- [19] M. Bender; K. Rutz; P.G. Reinhard; J. A. Maruhn : Pairing Gaps from Mean-Field Models, <https://arxiv.org/abs/nucl-th/0005028>, 2000.
- [20] Y. Q. He; J. K. Ge; G. J. Fu; H. Jiang : Isospin dependence of the nuclear binding energy, Chinese Phys. C 45, 014106, 2021.
- [21] N. Wang; M. Liu; X. Wu : Modification of nuclear mass formula by considering isospin effects, Physical Review C 81, 044322, 2010.
- [22] Junlong Tian; Haitao Cui; Teng Gao; Ning Wang : Effect of Wigner energy on the symmetry energy coefficient in nuclei, <https://arxiv.org/abs/1508.05780>, 2016.
- [23] Chen-Qi Li; Chao-Nan Tong; Hong-Jing Du; Long-Gang Pang : Deep learning on nuclear mass and α decay half-lives, <https://arxiv.org/abs/2202.11897>, 2022.

- [24] Z. M. Niu; H. Z. Liang; B. H. Sun; W. H. Long; Y. F. Niu : Predictions of nuclear β -decay half-lives with machine learning and their impact on r-process nucleosynthesis, *Phys. Rev. C* 99, 064307, 2019.
- [25] S. Akkoyun et al. : An artificial neural network application on nuclear charge radii, *J. Phys. G: Nucl. Part. Phys.* 40, 055106, 2013.
- [26] S. Akkoyun : Estimation of fusion reaction cross-sections by artificial neural networks, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, Vol. 462, str. 51-54. 2020.
- [27] Y. L. Yang; P. W. Zhao : Deep-neural-network approach to solving the *ab initio* nuclear structure problem, *Phys. Rev. C* 107, 034320, 2023.
- [28] E. Alpaydin : Introduction to machine learning, MIT press, 2020.
- [29] Prajit Ramachandran; Barret Zoph; Quoc V. Le : SWISH: A Self-gated Activation Function, <https://arxiv.org/abs/1710.05941v1>, 2017.
- [30] Prajit Ramachandran; Barret Zoph; Quoc V. Le : Searching For Activation Functions, <https://arxiv.org/abs/1710.05941>, 2017.
- [31] Shiv R. Dubey; Satish K. Singh; Bidyut B. Chaudhuri : Activation functions in deep learning: A comprehensive survey and benchmark, *Neurocomputing*. Volume 503, str. 92-108. 2022.
- [32] Kurt Hornik; Maxwell Stinchcombe; Halbert White : Multilayer feedforward networks are universal approximators, *Neural Networks*, Vol. 2, str. 359-366. 1989.
- [33] M. H. Kutner; C. J. Nachtsheim; J. Neter; W. Li et al. : Applied linear statistical models, Vol. 5, McGraw-Hill Irwin New York, 2005.
- [34] Diederik P. Kingma; Jimmy Ba : Adam: A Method for Stochastic Optimization, <https://arxiv.org/abs/1412.6980>, 2015.
- [35] Ivan Vasilev; Daniel Slater; Gianmario Spacagna; Peter Roelants; Valentino Zocca : Python Deep Learning: Exploring deep learning techniques, neural network architectures and GANs with PyTorch, Keras and TensorFlow, Packt Publishing, 2019.

- [36] Steve Lawrence; C. Lee Giles; Ah Chung Tsoi : What Size Neural Network Gives Optimal Generalization? Convergence Properties of Backpropagation, Technical Report, Institute for Advanced Computer Studies, University of Maryland, 1996.
- [37] J.L. Friar : The Structure of Light Nuclei and Its Effect on Precise Atomic Measurements, <https://arxiv.org/abs/nucl-th/0207065>, 2002.
- [38] A.E. Lovell; A.T. Mohan; T.M. Sprouse; M.R. Mumpower : Nuclear masses learned from a probabilistic neural network, <https://arxiv.org/abs/2201.00676>, 2022.
- [39] A. M. Nagaraja; H. C. Manjunatha; N. Sowmya; K. N. Sridhar; P. S. Damodara Gupta; S. A. Cecil Raj : Theoretical evidence for neutron magic number 184 from cluster radioactivity studies, International Journal of Modern Physics E, Vol. 31, No. 01, 2250004-247, 2022.
- [40] H. C. Manjunatha; N. Sowmya; N. Manjunath; L. Seenappa : Investigations on the superheavy nuclei with magic number of neutrons and protons, International Journal of Modern Physics E, Vol. 29, No. 05, 2050028-673, 2020.
- [41] P. Moller; A. J. Sierk; T. Ichikawa; H. Sagawa : Nuclear ground-state masses and deformations: FRDM(2012), <https://arxiv.org/abs/1508.06294>, 2015.
- [42] NuDat 3, National Nuclear Data Center (NNDC), Brookhaven National Laboratory, <https://www.nndc.bnl.gov/nudat3/>
- [43] S. Athanassopoulos; E. Mavrommatis; K. A. Gernoth; J. W. Clark : Nuclear mass systematics by complementing the Finite Range Droplet Model with neural networks, <https://arxiv.org/abs/nucl-th/0511088v1>, 2005.
- [44] S. Goriely; N. Chamel; J. M. Pearson : Further explorations of Skyrme-Hartree-Fock-Bogoliubov mass formulas. XVI. Inclusion of self-energy effects in pairing, Phys. Rev. C 93, 034337, 2016.
- [45] M. Kortelainen; T. Lesinski; J. Moré; W. Nazarewicz; J. Sarich; N. Schunck; M. V. Stoitsov; S. Wild : Nuclear Energy Density Optimization, <https://arxiv.org/abs/1005.5145>, 2010.

- [46] M. Kortelainen; J. McDonnell; W. Nazarewicz; P.-G. Reinhard; J. Sarich; N. Schunck; M. V. Stoitsov; S. M. Wild : Nuclear energy density optimization: Large deformations, <https://arxiv.org/abs/1111.4344>, 2012.
- [47] Lin-Xing Zeng; Yu-Ying Yin; Xiao-Xu Dong; Li-Sheng Geng : Nuclear binding energies in artificial neural networks, <https://arxiv.org/abs/2210.02906v1>, 2022.
- [48] G. Audi; W. Meng; A. H. Wapstra; F. Kondev; M. McCormick; X. Xu; B. Pfeiffer : The AME 2012 atomic mass evaluation(I). Evaluation of input data, and adjustment procedures, Chinese Physics C, Vol. 36, 1287, 2012.
- [49] Wang M.; G. Audi; A. H. Wapstra; F. G. Kondev; M. MacCormick; X. Xu; B. Pfeiffer : he AME 2012 atomic mass evaluation(II). Tables, graphs and references, Chinese Physics C, Vol. 36, 1603, 2012.
- [50] MassExplorer, FRIB Facility for Rare Isotope Beams, Michigan State University, <https://massexplorer.frib.msu.edu/>
- [51] A. V. Afanasjev; S. E. Abgemava; D. Ray; P. Ring : Nuclear landscape in covariant density functional theory, <https://arxiv.org/abs/1309.3289v1>, 2013.
- [52] Léo Neufcourt; Yuchen Cao; Witold Nazarewicz; Frederi Viens : Bayesian approach to model-based extrapolation of nuclear observables, Physical Review C 98, 034318, 2018.
- [53] Daniel T. Chang : Bayesian Neural Networks: Essentials, <https://arxiv.org/abs/2106.13594>, 2021.