Aktivno učenje interatomskih potencijala za predikciju kristalne strukture

Jurdana, Janko

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Janko Jurdana

Aktivno učenje interatomskih potencijala za predikciju kristalne strukture

Diplomski rad

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

INTEGRIRANI PREDDIPLOMSKI I DIPLOMSKI SVEUČILIŠNI STUDIJ
FIZIKA; SMJER ISTRAŽIVAČKI

Janko Jurdana

Diplomski rad

Aktivno učenje interatomskih potencijala za predikciju kristalne strukture

Voditelj diplomskog rada: dr. sc. Ivor Lončarić

Ocjena diplomskog rada: _____

Povjerenstvo: 1. _____

2. _____

3. _____

Datum polaganja: _____

Zagreb, 2023.

Sažetak

Predikcija kristalne strukture predstavlja vrlo zahtjevan zadatak u modernoj fizici čvrstog stanja. Većina dosadašnjih metoda zasnovana je na teoriji funkcionala gustoće. Međutim, takve metode postaju računalno prezahtjevne kada su promatrani sustavi kristali organskih molekula. Razvojem metoda strojnog učenja, moguće je ubrzati izračune svojstava takvih sustava što otvara vrata proučavanju široke klase materijala. U ovom radu su takve tehnike iskorištene za predikciju energija osnovnog stanja i sila na atome treniranjem strojno naučenih potencijala. Svojstva izračunata tim tehnikama su potom uspoređena sa svojstvima dobivenim korištenjem metoda teorije funkcionala gustoće.

Ključne riječi: teorija funkcionala gustoće, strojno učenje, predikcija kristalne strukture, deskriptori

Active learning of force fields for crystal structure prediction

Abstract

Crystal structure prediction poses a very demanding task in modern day solid state physics. Most of the existing methods are based on density functional theory. However, these methods become computationally too demanding when the studied systems are crystals of organic molecules. With the development of machine learning techniques, the calculations of the properties of such systems were drastically accelerated and thus opened the door to the study of a wide class of materials. In this paper, such methods were used for predicting ground state energies and forces on the atoms by training machine learned force fields. The properties calculated by these techniques were then compared with the properties obtained using density functional theory.

Keywords: density functional theory, machine learning, crystal structure prediction, descriptors

Sadržaj

1 Uvod **1**

2 Teorija funkcionala gustoće **4**

2.1 Aproksimacija fiksiranih jezgara 6

2.2 Aproksimacija nezavisnih elektrona 7

2.3 Hohenberg-Kohnovi teoremi 9

2.4 Funkcionali 12

3 Strojno naučeni interatomski potencijali **14**

3.1 Aktivno učenje 18

3.2 Deskriptori 19

3.3 Jezgrene metode 21

3.4 Bayesovska linearna regresija 23

4 Metoda **27**

4.1 Podaci 27

4.1.1 Izvor 27

4.1.2 Analiza i priprema podataka 28

4.2 Model 30

5 Rezultati i diskusija **33**

5.1 Hiperparametri 33

5.2 Težinski faktor energija i sila 34

5.3 Težinski faktor deskriptora 37

5.4 Dekompozicija na singularne vrijednosti (engl. Singular Value Decomposition - SVD) 40

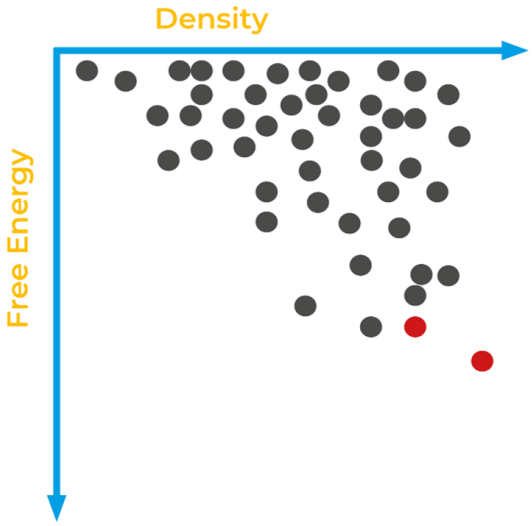
6 Zaključak **45**

Literatura **47**

1 Uvod

Većina krutina koje se pronalaze u prirodi na mikroskopskoj razini sadrže periodičnu kristalnu rešetku. Eksperimentalnim metodama poput kristalografije X-zrakama može se rekonstruirati oblik takve rešetke. Međutim, ispravna identifikacija tipa rešetke i njenih svojstava na temelju poznavanja tipova atoma (iona) koji tvore rešetku je puno složeniji problem. Taj problem je zahtjevan i kada se radi o jednostavnim kristalima, a dodatno se komplicira kada se promatraju tzv. molekulski kristali, odnosno kristali čije su građevne jedinice molekule. Proučavanje njihovih svojstava otvara mogućnosti njihove primjene u industriji i akademiji. Primjerice, velik broj farmaceutika i pigmenata su molekulski kristali. Proučavanjem njihove kristalne strukture, moguće je doći do novih saznanja o njihovoj sintezi, otkrivanju stabilnijih struktura i uštede pri priozvodnji. Osim činjenice da su molekule kompleksnije strukture od atoma, što povećava broj prostornih konformacija, izvor dodatnih komplikacija je fenomen polimorfizma. U kontekstu molekulskih kristala, polimorfizam označava pojavu pri kojoj identične molekule kristaliziraju u međusobno različite kristalne strukture ovisno o uvjetima pri kojima se kristalizacija odvija. Čak i kada se energija osnovnog stanja takvih sustava može izračunati relativno brzo i bez potrošnje velike količine memorije, ne zna se ništa o početnim uvjetima ili kinetičkim procesima koji dovode do formiranja takve strukture. Pri identifikaciji potencijalno stabilnih polimorfa u obzir se uzima ovisnost slobodne energije kristalnog sustava o gustoći pakiranja sastavnih molekula. Primjer grafa koji prikazuje tu ovisnost dan je na Slici 1.1. Takav graf se naziva pejzažnim dijagramom (engl. landscape diagram) i daje generalnu procjenu istraživačima koji pokušavaju sintetizirati nove polimorfe.

Disciplina koja pokušava odgonetnuti kristalnu strukturu čvrstih tijela polazeći od prvih principa i poznavanja tipova molekula koje tvore kristalnu rešetku naziva se predikcijom kristalne strukture (engl. Crystal Structure Prediction - CSP). Jedno od bitnijih svojstava kristalnih struktura potrebnih za ispravnu identifikaciju kristalne rešetke je energija osnovnog stanja kristalne strukture. U modernoj fizici čvrstog stanja, izračun energije osnovnog stanja kristalne strukture se izvršava raznim metodama poput teorije funkcionala gustoće (engl. Density Functional Theory - DFT). Glavna ideja iza DFT-a je činjenica da je energija osnovnog stanja funkcional elektronske gustoće. Time se drastično smanjuje broj varijabli potrebnih za izračun energije



Slika 1.1: Primjer pejzažnog dijagrama. Crvenim točkama označene su kandidati konformacija koje bi se mogle sintetizirati. Preuzeto iz [5].

osnovnog stanja. Međutim, čak i uz takvo pojednostavljenje, izračuni zahtjevaju veliku količinu memorije i vremena zbog kompleksne prirode molekulskih kristala. Taj se problem nastoji riješiti korištenjem metoda strojnog učenja, odnosno tzv. strojno naučenih interatomskih potencijala (engl. Machine Learned Force Fields - MLFF).

Strojno učenje je grana umjetne inteligencije koja se bavi stvaranjem modela koji, na temelju jednog skupa podataka, mogu predvidjeti svojstva nekog drugog, sličnog skupa podataka. Ideja je da se algoritmu strojnog učenja kao ulaz da velik broj kristalnih struktura te da on kao izlaz producira interatomski potencijal pomoću kojeg je moguće u vrlo kratkom vremenu izračunati energiju osnovnog stanja ili sile koje atomi u kristalnoj rešetki osjećaju, iz čega se mogu dobiti razna svojstva materijala.

Podatke o kristalnim sustavima podijelio je CCDC (engl. Camebridge Crystallographic Data Centre) u sklopu sedmog CSP slijepog testa kristalne strukture [27]. CCDC slijepi testovi kristalne strukture su natjecanja koja se provode od 1999. u svrhu boljeg razumijevanja načina na koji se kristalne strukture formiraju, što otkriva bolje i brže načine razvoja novih materijala. Natjecatelji dobiju podatke koji se sastoje od 2D kemijskih struktura spojeva koji su eksperimentalno ostvareni i čija su svojstva eksperimentalno određena. Cilj je, na temelju tih podataka, pronaći tipove struktura koje nastaju kristaliziranjem promatranih molekula i rangirati ih prema stabilnosti. Tu u igru ulaze metode CSP-a. Tipično, postupak se odvija u nekoliko koraka. Najprije se, na temelju 2D modela molekule, generira 3D model molekule. Nakon toga

se, utilizacijom naprednih tehnika pretraživanja prostora konformacija baziranih na slobodnoj energiji i gustoći pakiranja, predviđaju moguće kristalne strukture. Idući korak je rangiranje dobivenih struktura po stabilnosti. Najčešće su najvjerojatnije i najstabilnije strukture u koje će molekule kristalizirati one s niskom energijom osnovnog stanja i visokom gustoćom pakiranja. Zbog toga su pejzažni dijagrami, kao onaj na Slici 1.1, vrlo korisni prikazi podataka. Kako bi se ispravno rangirale strukture, u ovom radu iskorišten je strojno naučeni model koji, na temelju informacija i položaja atoma u strukturi, može izračunati energiju strukture i sile na pojedine atome u strukturi.

2 Teorija funkcionala gustoće

Za cjelokupno razumijevanje kristalnih sustava potrebno je riješiti višečestičnu Schrödingerovu jednadžbu. Rješenje te jednadžbe u principu postoji, ali u stvarnosti za sustave s više od nekoliko čestica moraju se koristiti aproksimativne metode. Stabilnost materijala uvjetovana je uspostavljanjem ravnoteže između odbojnih kulonskih interakcija između elektrona, odbojnih interakcija između atomskih jezgara i privlačnih interakcija između elektrona i jezgara. Energije tih interakcija su opisane jednadžbama klasične elektrodinamike [7]:

$$E_{ee} = \frac{e^2}{4\pi\varepsilon_0 d_{ee}} \tag{2.1}$$

$$E_{nn} = \frac{Z^2 e^2}{4\pi\varepsilon_0 d_{nn}} \tag{2.2}$$

$$E_{en} = -\frac{Ze^2}{4\pi\varepsilon_0 d_{en}}, \tag{2.3}$$

gdje E_{ee}, E_{nn} i E_{en} predstavljaju energiju međuelektronskih interakcija, energiju međujezgrenih interakcija i energiju interakcija elektrona i jezgara, d_{ee}, d_{nn} i d_{en} predstavljaju udaljenosti između pojedinačnih elektrona, udaljenosti između pojedinačnih jezgara i udaljenosti između elektrona i jezgara. ε_0 je permitivnost vakuuma, a Z atomski broj. Kada bi se promatrao jedan izolirani elektron, njegovo ponašanje bilo bi u potpunosti opisano jednočestičnom Schrödingerovom jednadžbom. Ukoliko je sustav u ravnoteži, energetsko stanje koje elektron okupira je osnovno stanje opisano valnom funkcijom ψ_0, što znači da je distribucija naboja dana s $|\psi_0(\mathbf{r})|^2$. Dodavanjem još jednog elektrona u sustav mijenja se distribucija naboja. Uzimajući u obzir Paulijev princip isključenja, drugi elektron može zaposjesti to isto stanje ukoliko je njegov spin suprotan spinu prvog elektrona. U ovakvoj konfiguraciji, distribucija naboja postaje $2|\psi_0(\mathbf{r})|^2$. Međutim, uzimajući u obzir kulonsku interakciju među elektronima, postaje jasno da ona modificira oblik ψ_0, ali i oblik potencijalnog člana V u Schrödingerovoj jednadžbi:

$$\left[\frac{\mathbf{p}^2}{2m_e} + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{2.4}$$

Pri razmatranju mnoštva elektrona i jezgara koristi se višečestična valna funkcija Ψ koja je funkcija N elektronskih koordinata $\mathbf{r}_1, ..., \mathbf{r}_N$ i M nuklearnih koordinata

$\mathbf{R}_1, ..., \mathbf{R}_M$:

$$\Psi = \Psi(\mathbf{r}_1, ..., \mathbf{r}_N; \mathbf{R}_1, ..., \mathbf{R}_M). \tag{2.5}$$

Isto kao što $|\psi(\mathbf{r})|^2$ predstavlja vjerojatnost nalaženja jednog elektrona na poziciji $\mathbf{r}$, $|\Psi(\mathbf{r}_1, ..., \mathbf{r}_N; \mathbf{R}_1, ..., \mathbf{R}_M)|^2$ predstavlja vjerojatnost istovremenog nalaženja prvog elektrona na poziciji $\mathbf{r}_1$, drugog elektrona na $\mathbf{r}_2$ itd. Međutim, u nekim slučajevima dovoljno je znati elektronsku gustoću, odnosno vjerojatnost nalaženja bilo kojeg elektrona na poziciji $\mathbf{r}$, definiranu kao:

$$\int n(\mathbf{r}) d\mathbf{r} = N, \tag{2.6}$$

gdje je N broj elektrona i gdje je iskorištena činjenica da je višečestična valna funkcija normirana:

$$\int |\Psi(\mathbf{r}_1, ..., \mathbf{r}_N; \mathbf{R}_1, ..., \mathbf{R}_M)|^2 d\mathbf{r}_1 ... d\mathbf{r}_N d\mathbf{R}_1 ... d\mathbf{R}_M = 1. \tag{2.7}$$

Koristeći jednadžbe 2.1, 2.2, 2.3, višečestična Schrödingerova jednadžba se zapisuje kao:

$$\Bigg[\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{e^2}{4\pi\varepsilon_0} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} \frac{e^2}{4\pi\varepsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \Bigg] \Psi = E_{tot} \Psi. \tag{2.8}$$

Varijabla E_{tot} predstavlja ukupnu energiju sustava određenu višečestičnom valnom funkcijom Ψ. Ova jednadžba sadrži kinetičke doprinose elektrona i jezgara, kao i elektron-elektron, jezgra-jezgra i elektron-jezgra potencijalne doprinose. Problem leži u tome da je ova jednadžba analitički nerješiva, osim za mali broj najjednostavnijih slučajeva. Zbog toga se pribjegava aproksimacijama koje tu jednadžbu transforimiraju u jednostavniji oblik. U modernoj fizici čvrstog stanja postoje brojne aproksimativne metode za račun svojstava kristalnih sustava. Jedna od njih naziva se teorijom funkcionala gustoće (DFT). DFT se koristi za modeliranje realističnih materijala zbog povoljnog omjera preciznosti i računalne složenosti. Također, teorija funkcionala gustoće pruža odgovor na sljedeće probleme: pronalaženje jednočestičnih elektronskih valnih funkcija i njihovo povezivanje s ukupnom višečestičnom elektronskom valnom funkcijom, pronalazak jednadžbi koje jednočestične elektronske valne funkcije zadovoljavaju te pronalazak načina za izračun ukupne energije elek-

tronskog sustava [1]. Ovi zadaci predstavljaju vrlo složene probleme čija zahtjevnost leži u činjenici da je energija proizvoljnog elektronskog stanja funkcional ukupne višečestične elektronske valne funkcije. Drugim riječima, energija nekog stanja elektronskog sustava u kristalu je funkcija višečestične valne funkcije. Takva vrsta ovisnosti podrazumijeva brojne nelinearne efekte koji onemogućavaju pronalazak egzaktnog rješenja Schrödingerove jednadžbe. Međutim, Hohenberg i Kohn su 1964. pokazali da, ukoliko je promatrana energija elektronskog sustava energija osnovnog stanja tog sustava, tada je energija funkcional isključivo elektronske gustoće [2]. Ova opservacija je krucijalna zbog toga što je broj varijabli potrebnih za definiciju višečestične valne funkcije $3N$, gdje je N broj elektrona u sustavu, dok je broj varijabli potrebnih za definiciju elektronske gustoće stanja jednak 3. Dakle, broj varijabli potrebnih za izračun energije osnovnog stanja je smanjen s $3N$, koliko je bilo potrebno za energiju pobuđenih stanja, na 3. Ključna posljedica ove opservacije je da je za izračun energije osnovnog stanja dovoljno izračunati elektronsku gustoću stanja.

2.1 Aproksimacija fiksiranih jezgara

Pogodno je jednadžbu 2.8 transformirati na pregledniji oblik koristeći tzv. Hartreejeve atomske jedinice. Jednadžba 2.8 tada poprima oblik:

$$\left[-\sum_i \frac{\nabla_i^2}{2} - \sum_I \frac{\nabla_I^2}{2M_I} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \right] \Psi = E_{tot}\Psi. \tag{2.9}$$

Aproksimacija fiksiranih jezgara (engl. clamped nuclei approximation) primjenjiva je na čvrsta tijela. Ona iskorištava činjenicu da se, u čvrstim tijelima, atomske jezgre ne gibaju na velikim udaljenostima, nego su efektivno fiksirane na određenim pozicijama. Ova pretpostavka se u jednadžbu 2.9 uvodi postavljanjem mase jezgara na $M_I = \infty$, što je opravdano jer su jezgre nekoliko redova veličine masivnije od elektrona. Posljedično, nuklearni kinetički doprinos energiji iščezava i kulonska odbojna interakcija između jezgara postaje konstantna. Uz redefiniciju energije:

$$E = E_{tot} - \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{2.10}$$

i uvođenje kulonskog potencijala kojeg osjećaju jezgre zbog prisustva elektrona:

$$V_n(\mathbf{r}) = -\sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}, \tag{2.11}$$

jednadžba 2.9 poprima oblik [1]:

$$\left[-\sum_i \frac{\nabla_i^2}{2} - \sum_i V_n(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \Psi = E\Psi. \tag{2.12}$$

Jednadžba 2.12 predstavlja temeljnu jednadžbu teorije elektronske strukture. Jednadžbu 2.12 se također može zapisati kao:

$$\hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_N)\Psi = E\Psi, \tag{2.13}$$

gdje je

$$\hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_i \hat{H}_0(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.14}$$

višeelektronski Hamiltonijan, a

$$\hat{H}_0(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_n(\mathbf{r}) \tag{2.15}$$

jednoelektronski Hamiltonijan.

2.2 Aproksimacija nezavisnih elektrona

Ukoliko se pretpostavi da promatrani elektron ne osjeća prisustvo ostalih elektrona, jednadžba 2.12 se dodatno pojednostavljuje i poprima oblik:

$$\sum_i \hat{H}_0(\mathbf{r}_i)\Psi = E\Psi. \tag{2.16}$$

Ovakav model se naziva aproksimacijom nezavisnih elektrona. Kako su elektroni nezavisni, vjerojatnost simultanog pronalaska elektrona 1 na poziciji $\mathbf{r}_1$, elektrona 2 na poziciji $\mathbf{r}_2$ itd. dana je produktom pojedinačnih vjerojatnosti pronalaska elektrona i na poziciji $\mathbf{r}_i$ $|\phi_i(\mathbf{r}_i)|^2$. Tada je sigurno pretpostaviti da višečestična valna funkcija

koja opisuje sve elektrone poprima oblik:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1) \ldots \phi_N(\mathbf{r}_N). \tag{2.17}$$

Nadalje, ako su jednočestične valne funkcije $\phi_i(\mathbf{r}_i)$ rješenja jednočestične Schrödingerove jednadžbe:

$$\hat{H}_0(\mathbf{r})\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}), \tag{2.18}$$

te se iskoristi činjenica da $\hat{H}_0(\mathbf{r}_i)$ djeluje samo na valnu funkciju $\phi_i(\mathbf{r}_i)$, dolazi se do procjene za energiju:

$$E = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_N. \tag{2.19}$$

Ovaj rezultat govori da je konfiguracija najniže energije unutar aproksimacije nezavisnih elektrona ona koja popunjava najniža stanja dana jednočestičnom Schrödingerovom jednažbom točno jednim elektronom. Međutim, aproksimacija nezavisnih elektrona ne uzima u obzir Paulijev princip isključenja jer se, generalno, predznak višečestične valne funkcije $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ ne mijenja zamjenom jednog para argumenata $(\mathbf{r}_i, \mathbf{r}_j)$. Također, kulonska interakcija među elektronima koja je zanemarena je istog reda veličine kao i ostale interakcije, pa njeno isključivanje iz višečestične Schrödingerove jednadžbe nije opravdano. Paulijev princip isključenja je zadovoljen ako je višečestična valna funkcija dana preko Slaterove determinante [6]:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \ldots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \ldots & \phi_N(\mathbf{r}_2) \\ \vdots & \ddots & & \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \ldots & \phi_N(\mathbf{r}_N) \end{vmatrix}. \tag{2.20}$$

Elektronska gustoća je tada dana kao:

$$n(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2, \tag{2.21}$$

zbog toga što je, ako su elektroni nezavisni, vjerojatnost nalaženja elektrona i na poziciji $\mathbf{r}_i$ jednaka $|\phi_i(\mathbf{r}_i)|^2$.

2.3 Hohenberg-Kohnovi teoremi

Energija osnovnog stanja E dana je jednadžbom

$$E = \langle \Psi | \hat{H} | \Psi \rangle , \tag{2.22}$$

a Hamiltonijan $\hat{H}$ koji se u njoj pojavljuje je dan jednadžbom 2.14. Struktura tog Hamiltonijana ne ovisi o vrsti promatranog materijala, što znači da je bilo kakva promjena u energiji u jednadžbi 2.13 uzrokovana isključivo promjenom višečestične valne funkcije Ψ [1]. Drugim riječima, energija E je funkcional višečestične valne funkcije Ψ:

$$E = F[\Psi]. \tag{2.23}$$

TEOREM 1. Neka je $n_g(\mathbf{r})$ elektronska gustoća osnovnog stanja sustava N elektrona. Tada $n_g(\mathbf{r})$ ne definira samo broj elektrona:

$$N = \int d\mathbf{r} n_g(\mathbf{r}), \tag{2.24}$$

nego i vanjski nuklearni potencijal $V_n(\mathbf{r})$, a posljedično i Hamiltonijan $\hat{H}$. Dakle, $n_g(\mathbf{r})$ određuje višečestične valne funkcije svih stanja [8].

Dokaz

Neka je Ψ_{n_g} proizvoljna normalizirana antisimetrična valna funkcija koja određuje elektronsku gustoću osnovnog stanja n_g. Vrijedi varijacijski princip:

$$\langle \Psi_{n_g} | \hat{H} | \Psi_{n_g} \rangle = \langle \Psi_{n_g} | (\hat{T} + \hat{V}_{ee}) | \Psi_{n_g} \rangle + \int d\mathbf{r} n_g(\mathbf{r}) V_n(\mathbf{r}) \geq E_g, \tag{2.25}$$

gdje je E_g najniža svojstvena vrijednost Hamiltonijana $\hat{H}$, a $\hat{T}$ i $\hat{V}_{ee}$ su operatori kinetičke energije i kulonske energije definirani kao:

$$\hat{T} = -\sum_i \frac{1}{2} \nabla_i^2 \tag{2.26}$$

$$\hat{V}_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.27}$$

Dodatno, neka je $\Psi_{n_g}^{min}$ normalizirana antisimetrična valna funkcija koja koja određuje elektronsku gustoću i minimizira $\langle \Psi_{n_g} | (\hat{T} + \hat{V}_{ee}) | \Psi_{n_g} \rangle$. Tada je $\Psi_{n_g}^{min}$ jednaka valnoj funkciji osnovnog stanja Ψ_{n_g} i rezultat Levy-ograničenog [9] pretraživanja preko svih

valnih funkcija koje definiraju elektronsku gustoću osnovnog stanja $n_g(\mathbf{r})$. Dakle, $n_g(\mathbf{r})$ određuje $\Psi_{n_g}^{min}$, a posljedično i Ψ_{n_g}. ■

TEOREM 2. Postoji univerzani funkcional elektronske gustoće $F[n]$, takav da je za proizvoljnu elektronsku gustoću sustava N čestica $n(\mathbf{r})$ funkcional energije jednak:

$$E[n] = F[n] + \int d\mathbf{r} n(\mathbf{r}) V_n(\mathbf{r}) \geq E_g. \tag{2.28}$$

Jednakost 2.28 vrijedi ako je $n(\mathbf{r})$ elektronska gustoća osnovnog stanja za vanjski potencijal $V_n(\mathbf{r})$.

Dokaz

Neka je univerzalni funkcional definiran kao:

$$F[n] = \langle \Psi_n^{min} | \left(\hat{T} + \hat{V}_{ee}\right) | \Psi_n^{min} \rangle \,, \tag{2.29}$$

gdje je $n(\mathbf{r})$ elektronska gustoća nekog proizvoljnog energetskog stanja, a Ψ_n^{min} valna funkcija koja definira $n(\mathbf{r})$ i minimizira univerzalni funkcional $F[n]$. Po varijacijskom principu, funkcional energije je tada:

$$E[n] = \langle \Psi_n^{min} | \, \hat{H} \, | \Psi_n^{min} \rangle = F[n] + \int d\mathbf{r} n(\mathbf{r}) V_n(\mathbf{r}) \geq E_g. \tag{2.30}$$

Iz ove jednakosti je očito da je Ψ_n^{min} jednaka valnoj funkciji osnovnog stanja Ψ_g ako je $n(\mathbf{r})$ elektronska gustoća osnovnog stanja $n_g(\mathbf{r})$. ■

Hohenberg-Kohnovi teoremi govore da je ukupna energija mnoštva elektrona u osnovnom stanju funkcional elektronske gustoće. Međutim, oni ne pružaju način da se takav funkcional konstruira. Iako egzaktna forma ovog funkcionala još uvijek nije poznata, razvijene su brojne aproksimacije.

Iz jednadžbe 2.28 vidljivo je da funkcional ukupne energije $E[n]$ ovisi eksplicitno o elektronskoj gustoći n preko člana $\int d\mathbf{r} n(\mathbf{r}) V_n(\mathbf{r})$, ali i implicitno preko člana $F[n] = \langle \Psi[n] | \left(\hat{T} + \hat{V_{ee}}\right) | \Psi[n] \rangle$. Kohn i Sham su 1965. predložili model [10] unutar kojeg se članovi implicitno ovisni o n u funkcionalu ukupne energije mogu zapisati kao suma kinetičke i kulonske energije u aproksimaciji nezavisnih elektrona i dodatnog člana E_{xc} koji u sebi sadrži sve efekte koji ne potječu od vanjskog nuklearnog potencijala,

kinetičke i kulonske potencijalne energije nezavisnih elektrona:

$$E[n] = \int d\mathbf{r} n(\mathbf{r}) V_n(\mathbf{r}) - \sum_i \int d\mathbf{r} \phi_i^*(\mathbf{r}) \frac{\nabla^2}{2} \phi_i(\mathbf{r}) + \frac{1}{2} \int \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n]. \tag{2.31}$$

Član E_{xc} naziva se energijom izmjene i korelacije. Kada bi njegova struktura bila poznata, mogla bi se direktno izračunati ukupna energija osnovnog stanja koristeći elektronsku gustoću n. Kako je elektronska gustoća osnovnog stanja n_g ona koja minimizira ukupnu energiju, može se pisati:

$$\left. \frac{\delta E[n]}{\delta n} \right|_{n_g} = 0. \tag{2.32}$$

Ovo svojstvo se naziva Hohenberg-Kohnovim varijacijskim principom i iz njega slijede tzv. Kohn-Shamove jednadžbe:

$$\left[-\frac{1}{2}\nabla^2 + V_n(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc(\mathbf{r})} \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}). \tag{2.33}$$

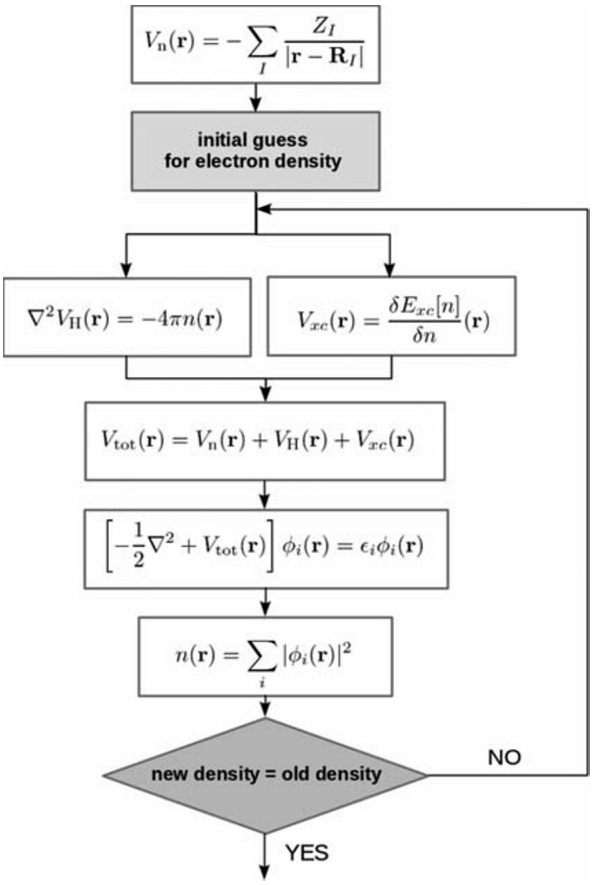
Član

$$V_{xc}(\mathbf{r}) = \left. \frac{\delta E_{xc}[n]}{\delta n} \right|_{n_g} = 0 \tag{2.34}$$

se naziva potencijalom izmjene i korelacije. U njemu su sadržani doprinosi energiji koji dolaze od zamjene pozicija elektrona identičnog spina unutar degeneriranog nivoa i doprinosi koji dolaze od interakcije pojedinačnih elektrona sa svim ostalim elektronima. Procedura rješavanja Kohn-Shamovih jednadžbi je sljedeća:

1. Specificikacija nuklearnih koordinata i izračun nuklearnog potencijala $V_n(\mathbf{r})$
2. Odabir probne funkcije elektronske gustoće $n(\mathbf{r})$
3. Određivanje ukupnog potencijala $V_{tot}(\mathbf{r})$ izračunom Hartreejeva potencijala $V_H(\mathbf{r})$ i potencijala izmjene i korelacije $V_{xc}(\mathbf{r})$ pomoću $n(\mathbf{r})$
4. Rješavanje Kohn-Shamovih jednadžbi
5. Određivanje nove elektronske gustoće $n(\mathbf{r})$.

Koraci 3.-5. se ponavljaju dok se ne postigne željena preciznost za $n(\mathbf{r})$. Shema postupka rješavanja Kohn-Shamovih jednadžbi prikazana je na Slici 2.2.



Slika 2.1: Shema postupka rješavanja Kohn-Shamovih jednadžbi. Dobivena rješenja su samosuglasna. Preuzeto iz [1].

2.4 Funkcionali

Kohm-Shamova teorija je formalno egzaktna jer kaže da mora postojati egzaktan funkcional energije izmjene i korelacije $E_{xc}[n]$ preko kojeg su, u principu, uključeni svi višečestični efekti. Kako forma tog funkcionala nije poznata, razvijene su brojne aproksimacije. Najjednostavnija od njih naziva se aproksimacijom lokalne gustoće (engl. Local Density Approximation - LDA). Ona pretpostavlja da funkcional energije izmjene i korelacije ovisi isključivo o elektronskoj gustoći:

$$E_{xc}^{LDA}[n] = \int d\mathbf{r} n(\mathbf{r}) e_{xc}^{hep}(n(\mathbf{r})), \tag{2.35}$$

gdje je e_{xc}^{hep} energija izmjene i korelacije po elektronu za homogeni elektronski plin. Ovakav pristup rezultira premalim iznosima konstanta rešetka, relativno dobrim iznosima površinskih energija za jednostavne metale (uz veliko poništenje grešaka energije izmjene i energije korelacije) i previsokim iznosima atomizacijskih energija za molekule [8]. Prirodno poopćenje LDA aproksimacije je aproksimacija generalizi-

ranog gradijenta (engl. Generalized Gradient Approximation - GGA). U GGA aproksimaciji funkcional energije izmjene i korelacije ovisi o elektronskoj gustoći i njenom gradijentu:

$$E_{xc}^{GGA}[n] = \int d\mathbf{r} e_{xc}^{GGA}(n(\mathbf{r}), \nabla n(\mathbf{r})). \tag{2.36}$$

Dodavanjem gradijentnog člana GGA model uključuje semilokalne informacije o $n(\mathbf{r})$, za razliku od LDA aproksimacije. Nadogradnja na GGA funkcionale su metaGGA funkcionali. U njima se dodatno uključuje ovisnost elektronske gustoće o gradijentu do kvadratnog člana i ovisnost elektronske gustoće o Kohn-Shamovoj gustoći kinetičke energije:

$$\tau = \frac{1}{2} \sum_i |\nabla \phi_i|^2. \tag{2.37}$$

Sumacija ide po svim zauzetim Kohn-Shamovim orbitalama, odnosno svim ϕ_i koje zadovoljavaju Kohn-Shamove jednadžbe. Uključivanjem ovog člana, potencijal izmjene i korelacije postaje ovisan o orbitalnom gibanju elektrona.

3 Strojno naučeni interatomski potencijali

Iako metode teorije funkcionala gustoće znatno ubrzavaju određivanje energije osnovnog stanja sustava, složenost algoritma raste s trećom potencijom broja elektrona u sustavu. To znači da se za složene sustave kao što su molekulski kristali algoritam izvršava predugo i zahtjeva velike količine memorije. Kako bi se taj problem zaobišao, koriste se metode strojnog učenja. Strojno učenje je grana umjetne inteligencije koja se bavi stvaranjem modela koji, na temelju jednog skupa podataka, mogu predvidjeti svojstva nekog drugog, sličnog skupa podataka. Ulazni podaci algoritma strojnog učenja nazivaju se primjerima (engl. example, instance). Oni su podatkovne točke u promatranom skupu podataka. Svaki primjer karakteriziraju značajke (engl. features), odnosno svojstva koja su bitna za promatrani problem. Dakle, svaki primjer $\mathbf{x}$ može se smatrati vektorom značajki koji pripada vektorskom prostoru nazvanom prostorom primjera ili ulaznom prostoru (engl. instance space, input space) $\mathcal{X}$ [3]:

$$\mathbf{x} = (x_1, ..., x_n), \tag{3.1}$$

gdje je n ukupan broj značajki. Ukoliko su značajke numeričke, prostor primjera je definiran kao $\mathcal{X} = \mathbb{R}^n$. Generalno, strojno učenje se dijeli na nadzirano i nenadzirano. Nadzirano strojno učenje karakteriziraju označeni podaci. Kod nadziranog učenja, svaki primjer ima svoju oznaku (engl. label). Ako se radi o problemu klasifikacije, oznaka predstavlja pripadnost klasi, a ako se radi o problemu regresije, oznaka predstavlja ciljnu brojčanu vrijednost. Skup označenih primjera označava se s $\mathcal{Y}$, a skup označenih primjera definiran je kao skup parova:

$$\mathcal{D} = \{(\mathbf{x}^i, y_i)\}_{i=1}^{N} \subseteq \mathcal{X} \times \mathcal{Y}, \tag{3.2}$$

gdje je N ukupan broj primjera. Poželjno je da broj primjera bude puno veći od broja značajki. Alternativno, skup označenih podataka se može prikazati u matričnom

zapisu preko dvije komponente: matrice neoznačenih primjera $\mathbf{X}$ i vektora oznaka $\mathbf{y}$:

$$\mathbf{X} = \begin{pmatrix} x_1^{(1)} x_2^{(1)} \dots x_n^{(1)} \\ x_1^{(2)} x_2^{(2)} \dots x_n^{(2)} \\ \vdots \\ x_1^{(N)} x_2^{(N)} \dots x_n^{(N)} \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{pmatrix}.$$

Matrica $\mathcal{D}$ se tada može zapisati kao $\mathcal{D} = (\mathbf{X}, \mathbf{y})$. Matrica $\mathbf{X}$ naziva se matricom dizajna. Cilj strojnog učenja je naučiti funkciju h koja primjerima iz $\mathcal{X}$ pridružuje oznaku iz $\mathcal{Y}$:

$$h : \mathcal{X} \mapsto \mathcal{Y}. \tag{3.3}$$

Takva funkcija se naziva hipotezom. Cilj algoritma strojnog učenja je podesiti parametre modela, koji je funkcija definirana do na parametre. Ako je vektor parametara zadan kao $\boldsymbol{\theta}$, tada je:

$$h = h(\mathbf{x}; \boldsymbol{\theta}). \tag{3.4}$$

Pojam modela u strojnom učenju zapravo označava skup hipoteza parametriziranih s $\boldsymbol{\theta}$. Njega se označava s $\mathcal{H}$. Formalno, model je definiran na sljedeći način:

$$\mathcal{H} = \{h(\mathbf{x}; \boldsymbol{\theta})\}_{\boldsymbol{\theta}}. \tag{3.5}$$

Svaki vektor parametara $\boldsymbol{\theta}$ jednoznačno određuje jednu funkciju (hipotezu) iz $\mathcal{H}$. Učenje modela predstavlja pretraživanje skupa $\mathcal{H}$ u potrazi za najboljom hipotezom h. Kod regresije, najbolja hipoteza je ona koja daje najmanje odstupanje od stvarnih vrijednosti, što znači da je strojno učenje zapravo optimizacijski problem. Numerička metrika koja određuje koliko je neka hipoteza dobra u opisivanju podataka iz skupa označenih podataka naziva se empirijskom pogreškom i označava se s $E(h|\mathcal{D})$. Ova notacija pojašnjava da je empirijska pogreška funkcija hipoteze za fiksirani skup podataka $\mathcal{D}$. Iznos pogreške na pojedinačom primjeru naziva se funkcijom gubitka (engl. loss function) i označava se s $L(y, h(\mathbf{x}))$. Funkcija gubitka $L : \mathcal{Y} \times \mathcal{Y} \mapsto \mathbb{R}^+$ kao argumente prima dvije oznake, onu točnu i onu dobivenu kao izlaz hipoteze, te računa odstupanje među njima, koje može biti proizvoljno definirano, dokle god je nenegativno. Funkcija pogreške je stoga očekivana vrijednost funkcije gubitka. Cilj algoritma strojnog učenja je naći hipotezu h^* koja minimizira

empirijsku pogrešku. Dakle, svaki algoritam strojnog učenja se sastoji od tri komponente: modela $\mathcal{H} = \{h(\mathbf{x}; \boldsymbol{\theta})\}$, funkcije gubitka $L(y, h(\mathbf{x}))$ i optimizacijskog postupka $h^* = argmin_{(h \in \mathcal{H})} E(h|\mathcal{D})$. Hipoteza koja producira empirijsku pogrešku jednaku nuli gotovo nikad ne postoji. Kako bi se pronašla hipoteza koja minimizira empirijsku pogrešku, potrebno je manipulirati tzv. složenošću modela. Složenost ili kapacitet modela određuje stupanj kompleksnosti funkcije koju je potrebno naučiti. Odabere li se model premale složenosti, model neće dobro funkcionirati ni na skupu za učenje ni na skupu za testiranje (skupu modelu nepoznatih primjera). Tada se kaže da je model podnaučen. Suprotno, ako je model prevelike složenosti, on će dobro funkcionirati na skupu za učenje jer će se previše prilagoditi šumu u podacima, ali neće biti dobar na skupu za testiranje (neće dobro generalizirati). Tada se kaže da je model prenaučen. Stupanj složenosti modela definiran je preko tzv. hiperparametara modela. Kod regresije oni najčešće predstavljaju stupanj nelinearnosti hipoteze (npr. stupanj polinomijalne funkcije). Razlika između hiperparametara i parametara modela je sljedeća: parametri su one varijable po kojima algoritam strojnog učenja radi minimizaciju empirijske pogreške i bira najbolju hipotezu $h^* \in \mathcal{H}$, dok su hiperparametri varijable koje određuju model $\mathcal{H}_i$ iz familije modela $\{\mathcal{H}_1, \ldots, \mathcal{H}_k\}$ koji se koristi za optimizaciju parametara. Dakle, parametri određuju hipotezu, a hiperparametri određuju model. Optimalan izbor modela (hiperparametara) provodi se postupkom unakrsne provjere (engl. cross validation). Zadatak koji se unakrsnom provjerom želi ispuniti je pronalazak modela koji najbolje radi na neviđenim primjerima. U tu svrhu, skup označenih primjera se dijeli na dva disjunktna skupa: skup za učenje i skup za testiranje:

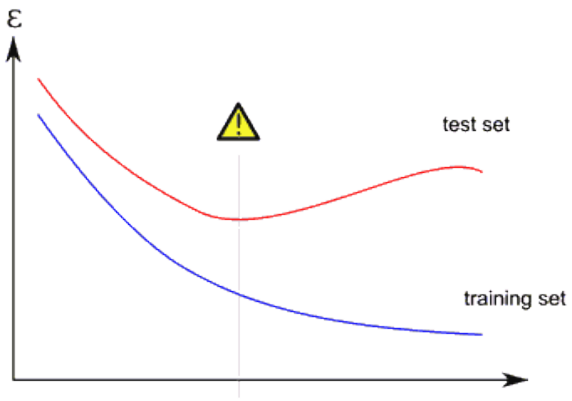
$$\mathcal{D} = \mathcal{D}_{train} \cup \mathcal{D}_{test}. \tag{3.6}$$

Potom se računa empirijska pogreška na oba skupa. Najbolji model je onaj koji minimizira pogrešku na skupu za testiranje $E(h|\mathcal{D}_{test})$, kao što je prikazano na Slici 3.1.

Poznavanje vremenske evolucije sustava zahtjeva poznavanje plohe potencijalne energije (engl. potential energy surface - PES). Ploha potencijalne energije definirana je kao funkcija [13]

$$f : \{Z_i, \mathbf{r}_i\}_{i=1}^N \mapsto E_i, \tag{3.7}$$

koja preslikava nuklearni naboj Z_i i poziciju $\mathbf{r}_i$ atoma i u njegovu potencijalnu ener-



Slika 3.1: Shematski prikaz odabira najboljeg modela primjenom križne validacije. Preuzeto iz [4].

giju E_i. Njeno je postojanje implicirano parametrizacijom ovisnosti između energije, pozicije i naboja jezgre preko Born-Oppenheimerove (BO) aproksimacije. BO aproksimacija odvaja nuklearno i elektronsko gibanje unutar Schödingerove jednadžbe. Pretpostavlja se da se elektroni instantno prilagođavaju nuklearnom gibanju, što ima smisla jer su jezgre nekoliko redova veličine masivnije od elektrona. Posljedično, gledano iz sustava elektrona, pozitivno nabijene jezgre su stacionarne, tako da parametarski ulaze u elektronsku Schrödingerovu jednadžbu. Odnosno, energija elektrona je potpuno određena vanjskim nuklearnim potencijalom, koji je potpuno određen nuklearnim pozicijama i nabojima [13]. Čak i s ovakvim pojednostavljenjem, Schrödingerova jednadžba može biti računalno skupa za riješiti. *Ab initio* metode daju dobre predikcije za male sustave periodičnih materijala ili male sustave materijala u plinskoj fazi [13]. Velike sustave modelira se pomoću empirijskih funkcija interatomskih potencijala (engl. force fields - FF) - parametrizacijama potencijalne energije. Takav pristup žrtvuje preciznost nauštrb vremena potrebnog za izvršavanje algoritma. Većina konvencionalnih interatomskih potencijala potencijalnu energiju tretira kao sumu vezanih i nevezanih članova [13]:

$$E_{tot} = E_{vezani} + E_{nevezani}, \tag{3.8}$$

gdje E_{vezani} uključuje efekte koji dolaze od kovalentnih veza među atomima i može se izraziti kao jednostavna funkcija udaljenosti između kovalentno vezanih atoma. $E_{nevezani}$ uključuje efekte uzrokovane dugodosežnim kulonskim i van der Waalsovim interakcijama. Kako su svi prethodno opisani članovi lako izračunljivi, klasični in-

teratomski potencijali su primjenjivi na velikim sustavima te mogu dati kvalitativno dobar opis kemijskih interakcija. Međutim, kvaliteta molekularno dinamičkih simulacija je ograničena preciznošću korištenog interatomskog potencijala. Strojno naučeni interatomski potencijali daju most između preciznih i računalno zahtjevnih *ab initio* metoda i ne toliko preciznih, vremenski efikasnih FF metoda. Glavna prednost strojno naučenih interatomskih potencijala leži u njihovoj mogućnosti da u podacima identificiraju funkcionalne ovisnosti bez rješavanja fizikalnih jednadžbi koje određuju ponašanje sustava. Prikladni parametri koji se mogu koristiti za učenje funkcionalne ovisnosti koja povezuje strukturu i njena svojstva su *ab initio* izračunate energije, sile i tenzori naprezanja.

3.1 Aktivno učenje

Aktivno učenje (engl. active learning) je grana nadziranog strojnog učenja koja je posebno korisna u slučajevima kada u podacima postoji puno neoznačenih primjera. Glavna ideja aktivnog učenja je da model, umjesto manualnog označavanja podataka (koje može biti računalno vrlo složen problem), može postići puno veću razinu preciznosti ukoliko mu se dopusti da sam bira podatke nad kojima će se trenirati. Način na koji model sam odabire podatke za treniranje je postavljanje upita za označavanjem podataka (engl. query) korisniku ili nekom drugom izvoru informacija. Prilikom svake iteracije i, ukupan skup podataka T se dijeli na tri podskupa [11]: podskup $T_{K,i}$ u kojemu se nalaze primjeri čije su oznake poznate, podskup $T_{U,i}$ u kojemu su primjeri neoznačeni i podskup $T_{C,i}$ u kojemu se nalazi dio primjera iz $T_{U,i}$ koji su odabrani za označavanje. Postoji nekoliko metoda odabira skupa $T_{C,i}$ [12]. Jedna od njih naziva se sintezom upita članstva (engl. membership query synthesis). Nju karakterizira to da model daje upite za označavanjem bilo kojeg primjera iz ulaznog prostora, uključujući primjere koji su nanovo generirani na temelju neke distribucije. Još jedna metoda odabira skupa $T_{C,i}$ je selektivno uzorkovanje temeljeno na struji podataka (engl. stream-based selective sampling). Njena glavna pretpostavka je da je dobivanje neoznačenog primjera računalno jeftino (ili besplatno), tako da se može uzorkovati primjer iz stvarne distribucije podataka i onda model odlučuje hoće li upitom zatražiti njegovo označavanje na temelju neke prethodno definirane mjere informativnosti. Takvim pristupom neoznačeni primjeri se odabiru jedan po

jedan, odnosno dolaze u struji, te od tu dolazi ime metode. Zadnja od najpopularnijih metoda odabira $T_{C,i}$ je uzorkovanje na skupu neoznačenih primjera (engl. pool-based sampling). Glavna pretpostavka je da postoji mali skup označenih primjera i veliki skup neoznačenih primjera. Način na koji radi je sličan prethodnoj metodi, ali umjesto da se upiti nad neoznačenim primjerima rade jedan po jedan, oni se izvršavaju nad nekim podskupom neoznačenog skupa primjera, te se na kraju odabire onaj podskup nad kojim je upit rezultirao najvećom vrijednosti mjere informativnosti. Neki od načina izračuna mjere informativnosti su: nasumično uzorkovanje, odabir uzoraka upitom komisiji, redukcija očekivane pogreške i dr. [12]

3.2 Deskriptori

Deskriptori su funkcije koje opisuju lokalnu atomsku strukturu materijala i znatno ubrzavaju račune [15]. Parametri koji definiraju deskriptore se optimiziraju tako da se reproduciraju rezultati *ab initio* računa. Unutar modela strojno naučenih interatomskih potencijala, potencijalna energija strukture koja se sastoji od N_a atoma je aproksimirana izrazom [16]:

$$U = \sum_{i}^{N_a} U_i, \tag{3.9}$$

gdje su U_i lokalne energije. Funkcija gustoće vjerojatnosti nalaženja atoma j u točki $\mathbf{r}$ koja se nalazi unutar sfere radijusa R_{cut} centrirane oko atoma i se može zapisati kao:

$$\rho_i(\mathbf{r}) = \sum_{i}^{N_a} = f_{cut}(|\mathbf{r}_{ij}|)g(\mathbf{r} - \mathbf{r}_{ij}), \tag{3.10}$$

gdje je $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ relativna udaljenost atoma i i atoma j, f_{cut} je funkcija koja osigurava da su sve informacije o atomskoj strukturi izvan radijusa R_{cut} uklonjene, a $g(\mathbf{r})$ je delta funkcija [16]. Pošto su delta funkcije neprikladne za numeričke izračune, koriste se aproksimacije. Jedna od najpoznatijih je glatko preklapanje atomskih pozicija (engl. smooth overlap of atomic positions - SOAP), u kojem funkcija $g(\mathbf{r})$ poprima oblik:

$$g(\mathbf{r}) = \frac{1}{\sqrt{2\sigma_{atom}\pi}} \exp\left(- \frac{|\mathbf{r}|^2}{2\sigma_{atom}^2} \right). \tag{3.11}$$

Kako su lokalne energije U_i funkcionali gustoće vjerojatnosti $U_i = F[\rho_i(\mathbf{r})]$, najjednostavniji numerički pristup bi uključivao razvijanje $\rho_i(\mathbf{r})$ u nekoj konačnoj bazi i izražavanje F kao funkcije koeficijenata razvoja. Međutim, tako definiran F ne bi

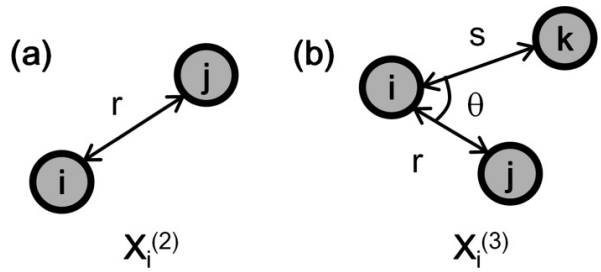
bio rotacijski invarijantan [14, 16]. Zbog toga se uvode pomoćne rotacijski i translacijski invarijantne funkcije ovisne o $\rho_i(\mathbf{r})$ - deskriptori. Najjednostavniji deskriptor naziva se radijalnim deskriptorom i definiran je kao:

$$\rho_i^{(2)}(r) = \frac{1}{4\pi} \int \rho_i(r\hat{\boldsymbol{r}}) d\hat{\boldsymbol{r}}. \tag{3.12}$$

On mjeri radijalnu udaljenosti od atoma i do atoma j u sferi radijusa R_{cut} centrirane u atomu i. Radijalni deskriptor sam nije dovoljan da uspješno producira plohu potencijalne energije. Razlog tome je što dvije različite gustoće vjerojatosti ρ_i mogu korespondirati s istim deskriptorom $\rho_i^{(2)}$, odnosno s istom energijom U_i. Taj se problem rješava uvođenjem dodatnog deskriptora koji sadrži angularnu ovisnost [16]:

$$\rho_i^{(3)}(r, s, \theta) = \int \int \delta(\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{s}} - \cos\theta) \rho_i(r\hat{\boldsymbol{r}}) \rho_i^*(s\hat{\boldsymbol{s}}) d\hat{\boldsymbol{r}} d\hat{\boldsymbol{s}}. \tag{3.13}$$

Sheme za definiciju radijalnog i angularnog deskriptora prikazane su na Slici 3.2.



Slika 3.2: **(a)** Radijalni deskriptor **(b)** angularni deskriptor. Preuzeto iz [16].

Gustoća vjerojatnosti ρ_i se može razviti na sljedeći način:

$$\rho_i(\mathbf{r}) = \sum_{l=1}^{L_{max}} \sum_{m=-l}^{l} \sum_{n=1}^{N_R^l} c_{nlm}^i \chi_{nl}(r) Y_{lm}(\hat{\boldsymbol{r}}), \tag{3.14}$$

gdje su c_{nlm}^i koeficijenti razvoja, $\{\chi_{nl} | n = 1, \ldots, N_R^l, l = 0, \ldots, L_{max}\}$ radijalne bazne funkcije koje zadovoljavaju relaciju ortonormiranosti:

$$4\pi \int_0^\infty \chi_{nl}(r) \chi_{n'l}(r) r^2 dr = \delta(n - n'), \tag{3.15}$$

a Y_{lm} kugline funkcije. Koristeći taj razvoj, moguće je radijalne i angularne deskriptore napisati kao:

$$\rho_i^{(2)}(r) = \frac{q}{\sqrt{4\pi}} \sum_{n=1}^{N_R^0} c_n^i \chi_{nl}(r) \tag{3.16}$$

$$\rho_i^{(3)}(r,s,\theta) = \sum_{l=1}^{L_{max}} \sum_{n=1}^{N_R^l} \sum_{\nu=1}^{N_R^l} \sqrt{\frac{2l+1}{2}} p_{n\nu l}^i \chi_{nl}(r) \chi_{\nu l}(s) P_l(\cos\theta), \tag{3.17}$$

gdje $\chi_{\nu l}$ predstavljaju normalizirane sferične Besselove funkcije, a P_l Legendreove polinome stupnja l. Član razvoja $p_{n\nu l}^i$ ima oblik [14, 16]:

$$p_{n\nu l}^i = \sqrt{\frac{8\pi^2}{2l+1}} \sum_{m=-l}^{l} c_{nlm}^i c_{\nu lm}^{i*}. \tag{3.18}$$

3.3 Jezgrene metode

Lokalna energija atoma U_i je funkcional radijalnih i angularnih deskriptora:

$$U_i = F\left[\rho_i^{(2)}, \rho_i^{(3)}\right]. \tag{3.19}$$

U praktičnim računima, svi koeficijenti c_n^i i $p_{\nu lm}^i$ se stavljaju u jedan vektor $\mathbf{x}_i$, tako da se može pisati:

$$U_i = F(\mathbf{x}_i). \tag{3.20}$$

Postoje brojne metode [18, 19] prilagodbe funkcije F na podatke dobivene *ab initio* računima. Sve te procedure rezultiraju lokalnom potencijalnom energijom atoma i oblika [17]:

$$U_i^\alpha = \sum_{i_B}^{N_B} w_{i_B} K(\mathbf{X}_i^\alpha, \mathbf{X}_{i_B}), \tag{3.21}$$

gdje α predstavlja indeks promatrane strukture, a N_B definira dimenziju baze, odnosno broj referentnih konfiguracija. Vektori $\boldsymbol{X}_i^\alpha$ i $\boldsymbol{X}_{i_B}$ sadrže sve informacije o koeficijentima razvoja gustoće vjerojatnosti, $\boldsymbol{X}_i^\alpha$ oko atoma i za promatranu lokalnu konfiguraciju α, a $\boldsymbol{X}_{i_B}$ za referentnu konfiguraciju izračunatu nekom *ab initio* metodom. Funkcija $K(\boldsymbol{X}_i^\alpha, \boldsymbol{X}_{i_B})$ naziva se jezgrenom funkcijom (engl. kernel function) i ona daje mjeru sličnosti između promatrane lokalne konfiguracije i referentne lokalne konfiguracije. Generalno, argumenti jezgrene funkcije ne moraju biti vektori,

nego mogu pripadati nekom apstraktnom prostoru $\mathcal{X}$. Jezgrena funkcija je tada definirana kao realna funkcija koja kao argumente prima dva primjera iz prostora $\mathcal{X}$ i kao izlaz daje realan broj:

$$K : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}. \tag{3.22}$$

Ako je kodomena jezgrene funkcije K ograničena na interval [0,1], K se naziva mjerom sličnosti. Tipično, što su primjeri sličniji, vrijednost K se približava jedinici, a nuli što su primjeri različitiji. Jezgrena funkcija K je također nenegativna i simetrična na zamjenu argumenata:

$$K(\boldsymbol{X}, \boldsymbol{X'}) \geq 0$$
$$K(\boldsymbol{X}, \boldsymbol{X'}) = K(\boldsymbol{X'}, \boldsymbol{X}).$$

Kada se ulazni primjeri mogu prikazati kao vektori, postoji više izbora za jezgrenu funkciju. Neki od njih su [20]:

$$\begin{aligned}
&\text{Linearna jezgra:} && K(\mathbf{X}, \mathbf{X}') = \boldsymbol{X}^T \boldsymbol{X'} \\
&\text{Gaussova jezgra:} && K(\mathbf{X}, \mathbf{X}') = \exp\left(- \frac{\left\|\mathbf{X} - \mathbf{X}'\right\|^2}{2\sigma^2} \right) \\
&\text{Eksponencijalna jezgra:} && K(\mathbf{X}, \mathbf{X}') = \exp\left(- \gamma \left\|\mathbf{X} - \mathbf{X}'\right\| \right).
\end{aligned}$$

Jezgrena funkcija koja je korištena u ovom radu je polinomijalna funkcija dana izrazom [21]:

$$K(\mathbf{X}_i, \mathbf{X}_{i_B}) = \left[\beta \mathbf{X}_i^{(2)} \cdot \mathbf{X}_{i_B}^{(2)} + (1 - \beta) \hat{\mathbf{X}}_i^{(3)} \cdot \hat{\mathbf{X}}_{i_B}^{(3)} \right]^{\zeta}, \tag{3.23}$$

gdje vektori $\mathbf{X}_i^{(2)}$ i $\mathbf{X}_i^{(3)}$ sadrže koeficijente c_n^i i $p_{n\nu l}^i$, vektor $\hat{\mathbf{X}}_i^{(3)}$ označava normalizirani vektor $\mathbf{X}_i^{(3)}$, β je težinski parametar koji regulira kolika se prednost daje radijalnom ili angularnom deskriptoru, a ζ je parametar koji kontrolira izoštrenost jezgrene funkcije K. Prvi član u izrazu 3.23 predstavlja linearni interakcijski član po parovima koji je pogodan za opis dugodosežnih interakcija (npr. kulonska) [16]. Drugi član predstavlja nelinearne interakcije više čestica [15, 16].

Uz energije, u fazu učenja modela se uključuju sile na atome i elementi tenzora stresa, koji se oboje isto mogu zapisati kao linearna funkcija koeficijenata w_{i_B} [16]. Kombiniranjem energija, sila i elemenata tenzora stresa za strukturu α dobiva se matrična

jednadžba:

$$\mathbf{y}^{\alpha} = \Phi^{\alpha}\boldsymbol{w}. \tag{3.24}$$

Vektor $\mathbf{y}^{\alpha}$ sadrži sve *ab initio* vrijednosti energija, sila i tenzora stresa te je dimenzije $m^{\alpha} = 1 + 3N_a^{\alpha} + 6$. Φ^{α} je $m^{\alpha} \times N_B$ matrica koja sadrži jezgrene funkcije pomoću kojih se računaju predikcije energije, sila i tenzora stresa, a $\mathbf{w}$ je N_B-dimenzionalan vektor koji sadrži sve koeficijente w_{i_B}. Spajanjem informacija svih struktura u jednu jednadžbu, dobiva se izraz:

$$\mathbf{Y} = \Phi\mathbf{w}, \tag{3.25}$$

gdje je $\mathbf{Y}$ supervektor dimenzije $N_{st} \times (1+3N_a^{\alpha}+6) \times N_B$ koji se sastoji od svih vektora $\mathbf{y}^{\alpha}$ za svaku strukturu α. N_{st} predstavlja broj struktura. Matrica Φ je matrica dizajna, podijeljena u blokove od kojih svaki predstavlja jednu strukturu α. Prvi redak svakog bloka sadrži jezgrene funkcije potrebne za predikciju potencijalne energije, idućih $3N_a^{\alpha}$ redaka se sastoji od derivacija jezgrene funkcije po atomskim koordinatama koje su potrebne za predikciju sila, a posljednjih 6 redaka svakog bloka sadrži derivacije jezgrene funkcije po koordinatama jedinične ćelije koje su potrebne za izračun tenzora stresa [21].

3.4 Bayesovska linearna regresija

Za rješavanje jednadžbe 3.25 postoji više metoda. Jedna od njih naziva se maksimizacijom funkcije izglednosti (engl. maximum likelihood estimation - MLE). Ona je bazirana na procjeni parametara neke pretpostavljene distribucije vjerojatnosti na temelju nekih opaženih podataka. Ako je ta pretpostavljena distribucija gausijan, tada je rješenje jednadžbe 3.25 [22]:

$$\mathbf{w}_{ML} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{Y}. \tag{3.26}$$

Veličina $(\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T$ naziva se Moore-Penroseovim pseudoinverzom [23] i može se smatrati poopćenjem matričnog inverza na nekvadratne matrice. Unutar ovakvog pristupa krucijalno je odabrati ispravan stupanj složenosti modela. Međutim, taj problem je vrlo težak zbog toga što, kada se u logaritam funkcije izgledosti uključi regularizacijski član, maksimizacija (logaritma) izglednosti vodi na prekompleksne i prenaučene modele [22]. Taj se problem može riješiti korištenjem druge metode

rješavanja jednadžbe 3.25 - bayesovske linearne regresije.

Bayesovski tretman linearne regresije započinje definiranjem apriorne distribucije gustoće vjerojatnosti po parametrima modela **w**. Pošto je funkcija izglednosti eksponencijalna funkcija kvadrata parametara **w** [22], njoj pridružena konjugirana apriorna distribucija vjerojatnosti je dana izrazom [21]:

$$p(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \sigma_w^2\mathbf{I}), \tag{3.27}$$

gdje je σ_w regularizacijski parametar modela, a **I** jedinična matrica. Funkcija izglednosti koja govori kolika je vjerojatnost opažanja strukture s *ab initio* vrijednostima energija, sila i stresa **Y**, uz opaženi set parametara **w** također je dana gausijanom [22]:

$$p(\mathbf{Y}|\mathbf{w}) = \mathcal{N}(\mathbf{Y}|\mathbf{\Phi w}, \sigma_v^2\mathbf{I}), \tag{3.28}$$

gdje je σ_v regularizacijski parametar modela. U konačnici se želi dobiti vjerojatnost $p(\mathbf{y}|\mathbf{Y})$ opažanja nove strukture **y**, uz opaženi set podataka **Y** i maksimizirati istu. U tu svrhu, potrebno je prvo izračunati aposterionu distribuciju vjerojatnosti $p(\mathbf{w}|\mathbf{Y})$. Nju se može dobiti iz Bayesovog teorema [24]:

$$p(\mathbf{w}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{Y})}, \tag{3.29}$$

gdje je

$$p(\mathbf{Y}) = \int p(\mathbf{Y}|\mathbf{w})p(\mathbf{w})d\mathbf{w}. \tag{3.30}$$

Koristeći sve ove pretpostavke, aposteriorna distribucija vjerojatnosti i njeni parametri dani su izrazima [16]:

$$p(\mathbf{w}|\mathbf{Y}) = \mathcal{N}(\bar{\mathbf{w}}, \mathbf{\Sigma}) \tag{3.31}$$

$$\bar{\mathbf{w}} = \frac{1}{\sigma_v^2}\mathbf{\Sigma\Phi}^T\mathbf{Y} \tag{3.32}$$

$$\mathbf{\Sigma}^{-1} = \frac{1}{\sigma_w^2}\mathbf{I} + \frac{1}{\sigma_v^2}\mathbf{\Phi}^T\mathbf{\Phi}. \tag{3.33}$$

Parametri $\bar{\mathbf{w}}$ predstavljaju centre multivarijatne Gaussove distribucije i oni maksimiziraju aposteriornu distribuciju vjerojatnosti, a matrica $\mathbf{\Sigma}$ sadrži kovarijance između značajki i naziva se kovarijacijskom matricom. Mutivarijatna Gaussova distribucija

dana je izrazom [21]:

$$\mathcal{N}(\bar{\mathbf{w}}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^{N_B} \det \boldsymbol{\Sigma}}} \exp\left[-\frac{(\mathbf{w}-\bar{\mathbf{w}})^T \boldsymbol{\Sigma}^{-1} (\mathbf{w}-\bar{\mathbf{w}})}{2} \right]. \tag{3.34}$$

Brojnik u eksponentu predstavlja poopćenje euklidske udaljenosti u prostoru parametara i naziva se Mahalanobisovom udaljenošću [25]. Koristeći relaciju:

$$p(\mathbf{y}|\mathbf{Y}) = \int p(\mathbf{y}|\mathbf{w}) p(\mathbf{w}|\mathbf{Y}) d\mathbf{w} \tag{3.35}$$

dolazi se do konačnog izraza za vjerojatnost opažanja strukture **y** uz opaženi set *ab initio* podataka **Y** [16]:

$$p(\mathbf{y}|\mathbf{Y}) = \mathcal{N}(\boldsymbol{\phi}\bar{\mathbf{w}}, \boldsymbol{\sigma}) \tag{3.36}$$

$$\boldsymbol{\sigma} = \sigma_v^2 \mathbf{I} + \boldsymbol{\phi}^T \boldsymbol{\Sigma} \boldsymbol{\phi}. \tag{3.37}$$

Sve što je preostalo je pronaći optimalne regularizacijske parametre σ_v i σ_w kako bi se sprječila prenaučenost. To se postiže korištenjem tzv. generalizirane maksimalne izglednosti (engl. generalized maximum likelihood, 2 maximum evidence, evidence approximation) [22, 26]. Ideja je maksimizacija funkcije dokaza [21]:

$$p(\mathbf{Y}|\sigma_w^2, \sigma_v^2) = \left(\frac{1}{\sqrt{2\pi\sigma_v^2}} \right)^M \left(\frac{1}{\sqrt{2\pi\sigma_w^2}} \right)^{N_B} \exp[-E(\mathbf{w})] d\mathbf{w} \tag{3.38}$$

$$E(\mathbf{w}) = \frac{1}{2\sigma_v^2} \|\boldsymbol{\Phi}\mathbf{w} - \mathbf{Y}\|^2 + \frac{1}{2\sigma_w^2} \|\mathbf{w}\|^2, \tag{3.39}$$

gdje je M dimenzija supervektora **Y**. Funkcija $p(\mathbf{Y}|\sigma_w^2, \sigma_v^2)$ korespondira s vjerojatnošću da model s parametrima regularizacije σ_w i σ_v kao izlaz da referentne *ab initio* podatke **Y**. Optimalan model je tada onaj koji maksimizira tu vjerojatnost. Jednadžbe koje maksimiziraju $p(\mathbf{Y}|\sigma_w^2, \sigma_v^2)$ se dobiju iz uvjeta $\frac{\partial p}{\partial \sigma_v^2} = \frac{\partial p}{\partial \sigma_w^2} = 0$ [22]:

$$\sigma_w^2 = \frac{|\bar{\mathbf{w}}|^2}{\gamma} \tag{3.40}$$

$$\sigma_v^2 = \frac{|\mathbf{Y} - \boldsymbol{\phi}\bar{\mathbf{w}}|^2}{M - \gamma} \tag{3.41}$$

$$\gamma = \sum_{k=1}^{N_B} \frac{\lambda_k}{\lambda_k + \frac{1}{\sigma_w^2}}, \tag{3.42}$$

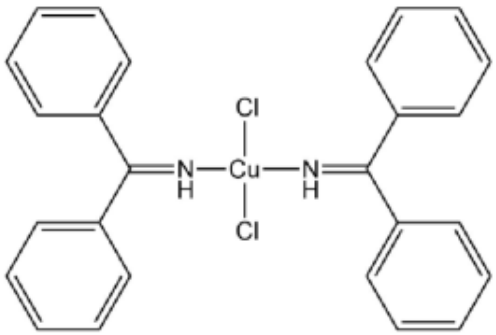
gdje su λ_k svojstvene vrijednosti matrice $\frac{\mathbf{\Phi}^T\mathbf{\Phi}}{\sigma_v^2}$. Jednadžbe 3.40-3.42 se rješavaju iterativno do konvergencije.

4 Metoda

4.1 Podaci

4.1.1 Izvor

Sirovi podaci koje je podijelio CCDC sadrže 500 različitih konformacija koje odgovaraju polimorfima metalno-organske molekule koja se sastoji od jednog atoma četverovalentnog bakra na koji su vezana dva atoma klora i dva atoma dušika [28]. Svaki atom dušika je vezan jednostrukom vezom na jedan atom vodika te dvostrukom vezom na jedan atom ugljika, koji je vezan za dva benzenska prstena. Molekula je prikazana na Slici 4.1.



Slika 4.1: Molekula koja čini strukture koje se koriste pri treniranju modela strojnog učenja. Slika preuzeta iz paketa podataka dobivenih od CCDC-a.

Kako bi se dobio skup podataka koji će poslužiti kao ulaz u model strojnog učenja, napravljeni su DFT izračuni za skup podataka koji je osigurao CCDC. Ti izračuni su rezultirali skupom podataka koji sadrži informacije o položajima i vrsti atoma, energijama osnovnog stanja pojedinih konformacija, kao i silama na pojedinačne atome u svakoj konformaciji. Programski paket korišten za DFT izračune je VASP (engl. Vienna Ab initio Simulation Package) [29]. VASP se koristi za modeliranje materijala na atomskoj skali, npr. za računanje elektronske strukture, polazeći od prvih principa. Njegov glavni zadatak je nalaženje aproksimativnih rješenja za višečestičnu Schrödingerovu jednadžbu, u ovom slučaju korištenjem metoda teorije funkcionala gustoće, odnosno rješavanjem Kohn-Shamovih jednadžbi. Centralne vrijednosti kao

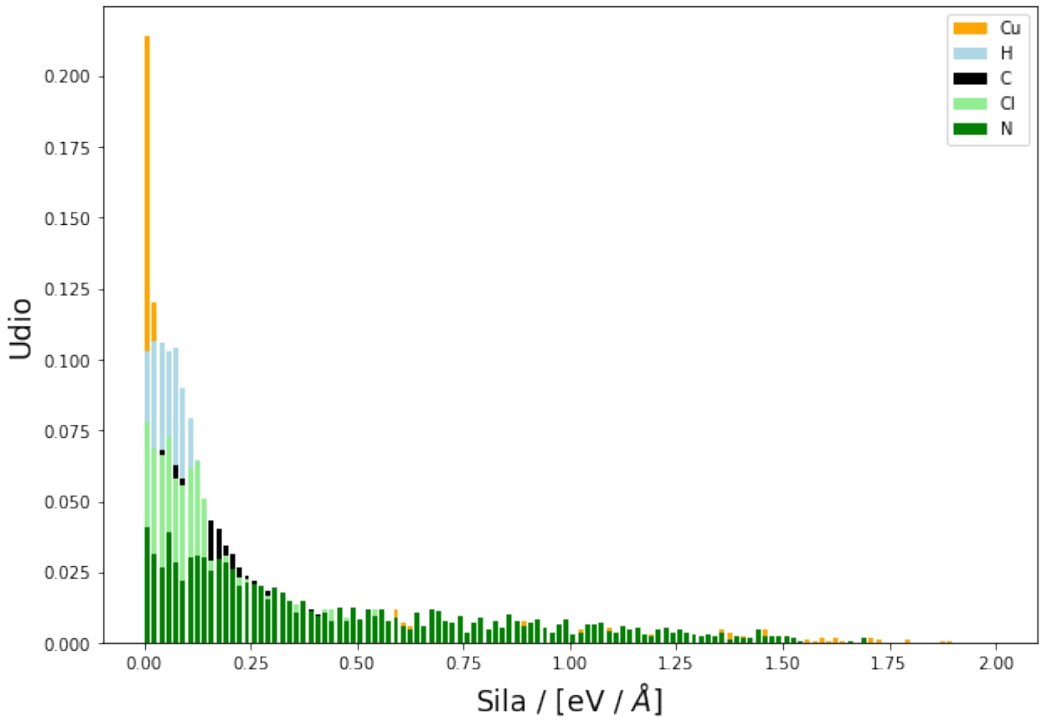
jednoelektronske orbitale, elektronske gustoće naboja i lokalni potencijali izraženi su u bazi ravnih valova. Za računanje elektronskog osnovnog stanja koriste se tehnike iterativne matrične dijagonalizacije. Baza reprezentacije valnih funkcija koja je korištena pri provedbi DFT računa odgovara ravnim valovima kinetičke energije do 420 eV. Integracija preko prve Brillouinove zone koja je potrebna za izvršavanje računa zamjenjuje se težinskom sumom po odabranim točkama u recipročnom prostoru (*k*-točkama) koje su međusobno razmaknute za 0.2 $Å^{-1}$. Kod metalnih spojeva se javlja dodatan problem u vidu integracije step funkcije centrirane u Fermijevom nivou unutar prve Brillouinove zone. Kako se integral aproksimira sumom preko *k*-točaka, prisutnost step funkcije znatno smanjuje brzinu konvergencije algoritma zbog toga što zaposjednutost stanja pada s 1 na 0 na Fermijevom nivou. Taj se problem zaobilazi metodom Gaussovog mrljanja (engl. Gaussian smearing). Ideja Gaussovog mrljanja je da se step funkcija zamijeni nekom glatkom funkcijom sličnom gausijanu koja ubrzava konvergenciju bez uništavanja preciznosti sume. U ovom slučaju, step funkcija je zamijenjena funkcijom oblika:

$$f\left(\frac{\varepsilon-\mu}{\sigma}\right)=\frac{1}{2}\left(1-erf\left(\frac{\varepsilon-\mu}{\sigma}\right)\right), \tag{4.1}$$

gdje je *erf(x)* funkcija pogreške (engl. error function), μ je energija Fermijeva nivoa, a σ je parametar koji određuje širinu termalnog prozora. Vrijednost parametra σ postavljena je na 0.03 eV. Za funkcional izmjene i korelacije odabran je R2SCAN [30], funkcional iz klase metaGGA.

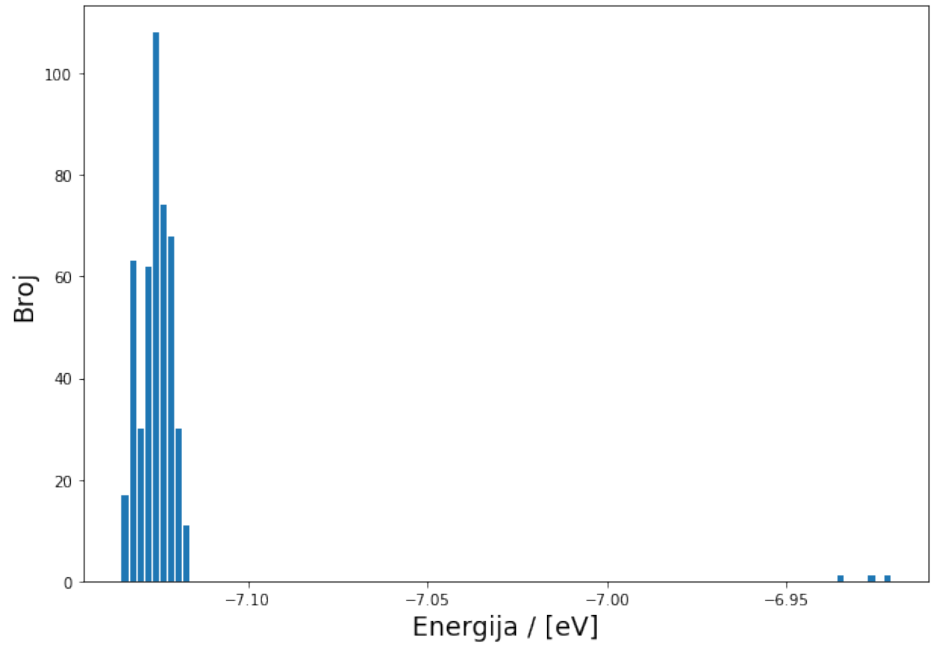
4.1.2 Analiza i priprema podataka

U svrhu boljeg razumijevanja skupa ulaznih podataka i smanjivanja vjerojatnosti pojave grubih grešaka, dobra je praksa napraviti vizualnu analizu podataka. Za zadanu strukturu bilo je zadano 500 različitih konformacija kristalnog uređenja te su za svaku konformaciju bili dostupni podaci o pozicijama svih atoma. Međutim, DFT izračuni su uspjeli za 464 strukture jer VASP ima problema s većim strukturama, pa je to bio cjelokupni skup podataka korišten za treniranje i evaluaciju algoritma strojnog učenja. Distribucija apsolutnih iznosa sila na atome prikazana je na Slici 4.2.



Slika 4.2: Distribucija apsolutnih iznosa sila na atome za sve konformacije.

Na Slici 4.2 je raspon apsolutnih iznosa sila odabran kao 0-2 eV / Å, iako vrijednosti sila na neke atome dosežu vrijendosti od 12.5 eV / Å. One su izuzete pošto bi se njihovim uključivanjem smanjila čitljivost histograma. Na Slici 4.3 prikazana je distribucija energija osnovnog stanja normirana na broj atoma pripadajuće konformacije. Iz Slike 4.3 vidi se da je raspon energija reda veličine 0.25 eV po atomu, a raspoređene su u grupe u kojima je raspon energija reda veličine 0.05 eV po atomu.

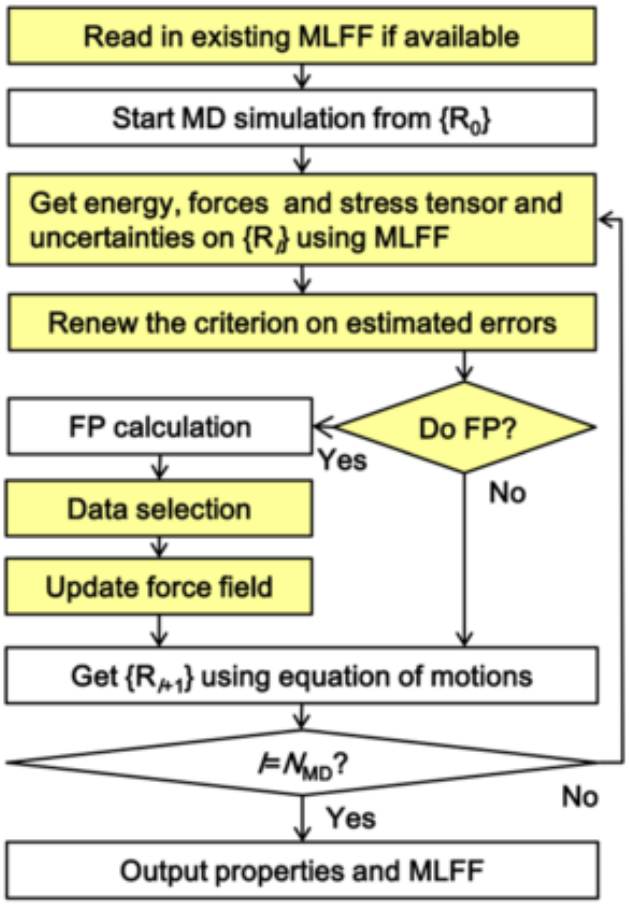


Slika 4.3: Distribucija energija osnovnog stanja svih konformacija. Energije su normirane na broj atoma pojedine konformacije.

4.2 Model

Model korišten u ovom radu je zasnovan na bayesovskoj linearnoj regresiji, a implementiran je u sklopu VASP programskog paketa [21]. Strojno naučeni interatomski potencijali tipično zahtjevaju manualnu konstrukciju skupa za učenje koji se sastoji od velikog broja *ab initio* dobivenih podataka. To često rezultira smanjenom primjenjivosti modela nad neviđenim strukturama [32]. Zbog toga je odabran model koji prilikom svake iteracije algoritma odlučuje hoće li prihvatiti predikcije generirane trenutnim interatomskim potencijalom ili pokrenuti *ab initio* račun koji će rezultirati novom konformacijom koja će se dodati skupu za treniranje modela. Zbog ove metode odabira podataka za treniranje model spada u klasu modela aktivnog učenja, specifično letećeg učenja (engl. on-the-fly learning). Ulazni podaci i podaci za evaluaciju modela su sva kristalna uređenja za koje su DFT izračuni uspjeli. U njima su sadržane informacije o Bravaisovoj rešetki, atomskim pozicijama, ukupnoj energiji osnovnog stanja, silama na pojedine atome i tenzor naprezanja. Algoritam iz tih podataka kreira lokalnu konfiguraciju (radijalnu i angularnu distribuciju atoma) oko nekog referentnog atoma i te informacije pohranjuje unutar deskriptora. Točan slijed operacija unutar algoritma prikazan je na Slici 4.4. Algoritam najprije izračuna predikciju energija, sila i tenzora naprezanja i njihove neodređenosti za pojedinačnu

strukturu na temelju trenutnog interatomskog potencijala. Nakon toga odlučuje hoće li pokrenuti *ab initio* račun kojim se generiraju nove strukture kandidati za dodavanje skupu za treniranje. To je uvjetovano iznosom bayesovske greške sila nad pojedinim atomima, odnosno *ab initio* računi se pokreću ako je greška veća od nekog predefiniranog praga. U tom slučaju, novogenerirane strukture se dodaju skupu za treniranje i pokreće se proces učenja koji rezultira ažuriranim interatomskim potencijalom. Koristeći taj novi potencijal, ažuriraju se atomske pozicije i njihove brzine. Ovaj proces se ponavlja dok se ne postigne zadani broj ionskih koraka. Razlog zašto je korišten ovaj model je taj što mu ne treba puno početnih podataka da uspije precizno replicirati *ab initio* rezultate, za razliku od npr. modela baziranih na neuronskim mrežama. Za potrebe naknadne obrade podataka korišten je paket ASE [33] (engl. Atomic Simulation Enviroment). On se sastoji od skupa alata i Python modula za manipulaciju, vizualizaciju i analizu atomskih simulacija. Unutar ASE paketa moguće je definirati kristalnu strukturu kao objekt s raznim svojstvima koja definiraju kristalnu rešetku poput vrste atoma, atomskih položaja, brzina, naboja, temperature itd. Također, ASE paket nudi mogućnost definicije ”kalkulatora” pomoću kojih se mogu izračunati energije osnovnog stanja struktura, sile na pojedine atome i tenzori stresa.



Slika 4.4: Shema generiranja interatomskog potencijala korištenjem letećeg (engl. on-the-fly) učenja. Slika preuzeta iz [21].

5 Rezultati i diskusija

5.1 Hiperparametri

Prilikom treniranja bilo kojeg modela strojnog učenja, krucijalno je pronaći vrijednosti hiperparametara modela koje minimiziraju pogreške. Kako su ciljne vrijednosti koje se predviđaju modelom energije osnovnog stanja sustava i sile na atome, potrebno je u model uvesti induktivnu pristranost u vidu težinskog faktora koji daje prednost energijama nauštrb sila. To se radi zbog diskrepancije u broju podataka za energije i sile. Naime, broj podataka za energije je N_{st}, dok podataka za sile ima $3 \times N_{at} \times N_{st}$, gdje je N_{st} broj konformacija, a N_{at} broj atoma unutar jedinične ćelije pojedine konformacije. Dakle, podataka za sile ima dva reda veličine više od podataka za energije. Također, pošto jezgrena funkcija koja služi kao mjera sličnosti između lokalnih konfiguracija iz skupa za treniranje i baznog skupa sadrži informacije o deskriptorima, potrebno je ugoditi težinski faktor koji regulira relativni stupanj zastupljenosti radijalnog i angularnog deskriptora. Kod samih deskriptora je potrebno ugoditi parametre radijusa rezanja. Oni definiraju iznos volumena okoline oko promatranog atoma, odnosno lokalne konfiguracije. Kako oni implicitno određuju broj susjednih atoma koji ulaze u razmatranje, povećavanjem tog volumena uključuje se sve više značajki koje doprinose točnosti predikcija. Međutim, preveliki volumeni lokalnih konfiguracija znatno usporavaju račune zbog količine memorije koju zahtjevaju, tako da je nužno pronaći iznos za koji će vrijeme izvršavanja biti razumno, a preciznost predikcija sačuvana. Hiperparametri modela koji su ugođavani u ovom radu uključuju:

- Težinski faktor energija i sila (ML_WTOTEN) - određuje stupanj zastupljenosti energija nauštrb sila.
- Težinski faktor deskriptora (ML_W1) - određuje stupanj zastupljenosti radijalnog desktiptora nauštrb angularnog deskriptora.
- Radijus rezanja (engl. cutoff radius) radijalnog deskriptora (RCUT1) - određuje domet radijalnog deskriptora u Å.
- Radijus rezanja angularnih deskriptora (RCUT2) - određuje domet angularnog deskriptora u Å.

- Širina Gausijana koji se koristi za proširivanje atomskih distribucija radijalnog deskriptora (ML_SION1) - implicitno određuje iznos jezgrene funkcije, odnosno sličnost lokalnih konfiguracija iz skupa za treniranje i baznog skupa. Mjerna jedinica je Å.
- Broj baznih funkcija za razvoj radijalnog deskriptora u red (ML_MRB1) - određuje stupanj razvoja u red preko radijalnih baznih funkcija.
- Metoda rješavanja problema minimizacije $\| \mathbf{Y} - \mathbf{\Phi w} \| \rightarrow \min$ (ML_IALGO_LINREG) - određuje hoće li se koristiti obična bayesovska linearna regresija ili dekompozicija matrice dizajna na singularne vrijednosti.

Ukupni skup podataka podijeljen je na skup za treniranje i skup za testiranje u omjeru 316:148.

5.2 Težinski faktor energija i sila

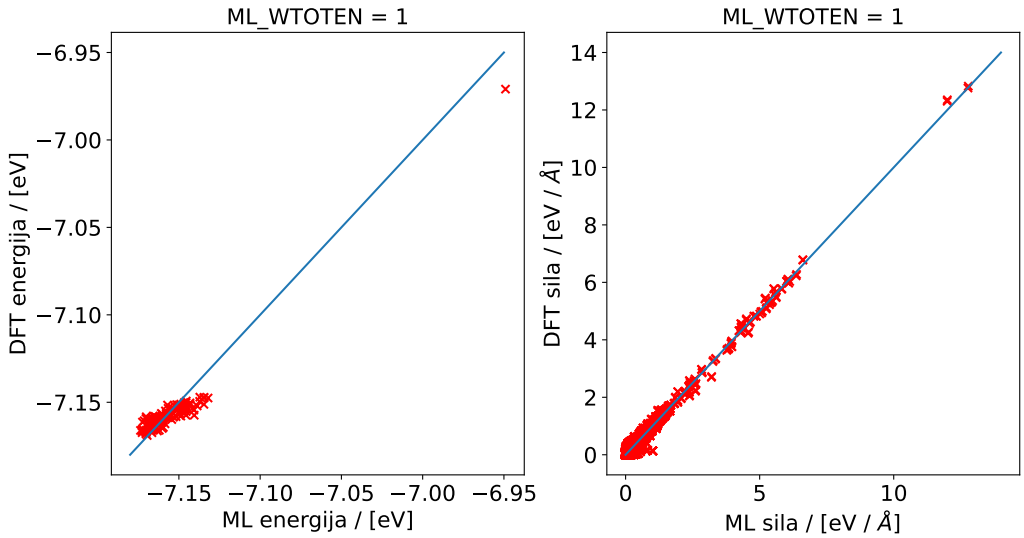
Kako bi se postigli optimalni rezultati, potrebno je napraviti normalizaciju ulaznih podataka. Nadalje, ponekad je poželjno favorizirati neke fizikalne veličine nauštrb drugih, npr. ako je potrebno dobiti što preciznije rezultate za energiju, njoj se može pridodati više pažnje nego silama na atome. Zbog toga je vrlo bitno ugoditi hiperparametar koji regulira magnitudu favoriziranja energija, odnosno težinski faktor energija. U ovom radu, energije, sile i elementi tenzora stresa normalizirani su na način da se skup za treniranje podijeli na podskupove te se potom zasebno izračunaju standardne devijacije energija, sila i tenzora stresa unutar spomenutih podskupova. Nakon toga se izračuna prosjek standardnih devijacija po svim podskupovima te se energije, sile i tenzor stresa normaliziraju koristeći te prosjeke. Potom se energije množe s faktorom definiranim labelom ML_WTOTEN. U Tablici 5.1 prikazane su srednje kvadratne pogreške za energije normirane na broj atoma i apsolutnog iznosa sile za različite vrijednosti težinskog faktora energije. Iznosi ostalih hiperparametara iz poglavlja 5.1 su: ML_W1 = 0.1, ML_RCUT1 = 6 Å, ML_RCUT2 = 3 Å, ML_SION1 = 0.3, ML_MRB1 = 12, ML_IALGO_LINREG = 1.

Da bi dani model bio upotrebljiv, potrebna je greška reda veličine 0.05 eV / Å za sile i 0.01 eV za energije [34]. Iz Tablice 5.1 može se vidjeti da najbolje su najbolji rezultati i za energije i za sile dobiveni za vrijednost težinskog faktora 16. Najgori rezultati dobiveni su korištenjem težinskog faktora 1, što je i očekivano zbog prethodno

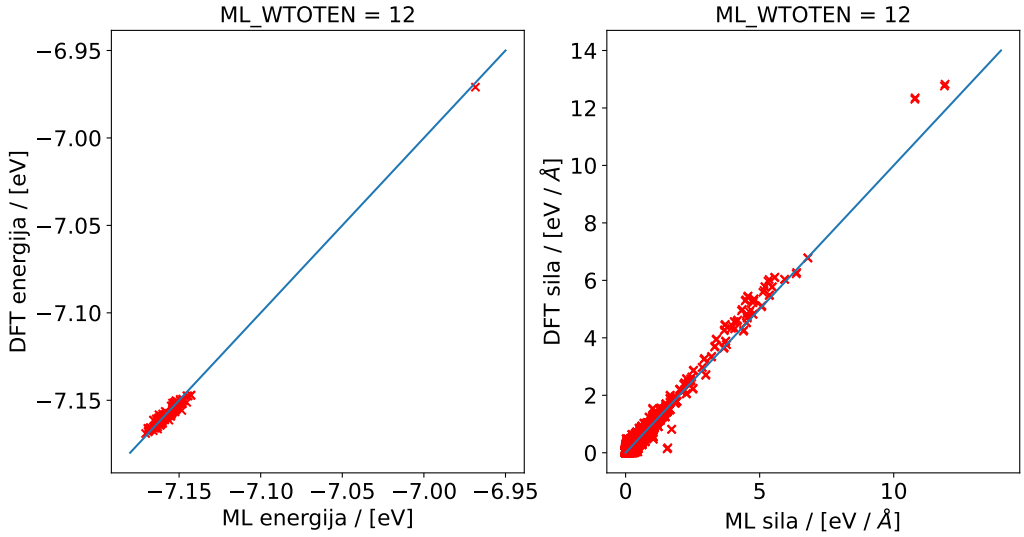
ML_WTOTEN	RMSE energija / [eV]	RMSE sile / [eV / Å]
1	0.00616	0.05233
12	0.00211	0.06161
14	0.00209	0.05649
16	0.00170	0.05028
18	0.00193	0.05726

Tablica 5.1: Tablični prikaz rezultata regresije interatomskih sila i energija za zadane konformacije s varijacijom težinskog faktora energija. Model je treniran nad 316 DFT konformacija.

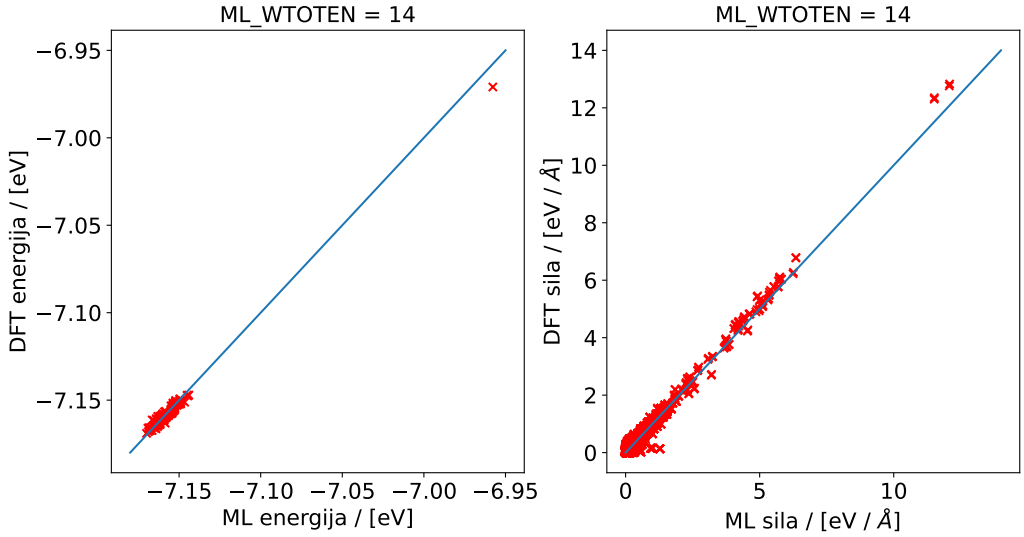
spomenute diskrepancije u broju ulaznih podataka za energije i sile. Grafički prikazi rezultata regresije energija normiranih na broj atoma i iznosa sila na pojedine atome nalaze sa na Slikama 5.1 - 5.5. Podaci na tim grafovima pripadaju skupu za testiranje, odnosno onima koje model nije vidio tijekom treninga. Poanta ovih grafova je da se vidi koliko dobro trenirani model predviđa energije i sile na pojedine atome, te da se vidi koliko se preciznost može povećati dodatnim učenjem. Svaka točka na grafu predstavlja predviđenu energiju/silu na pojedini atom neke of konformacija (x-os) u odnosu na energije/sile dobivene DFT računom (y-os). Također, prikazan je pravac y = x iz razloga što bi u idealnom slučaju sve točke ležale na njemu.



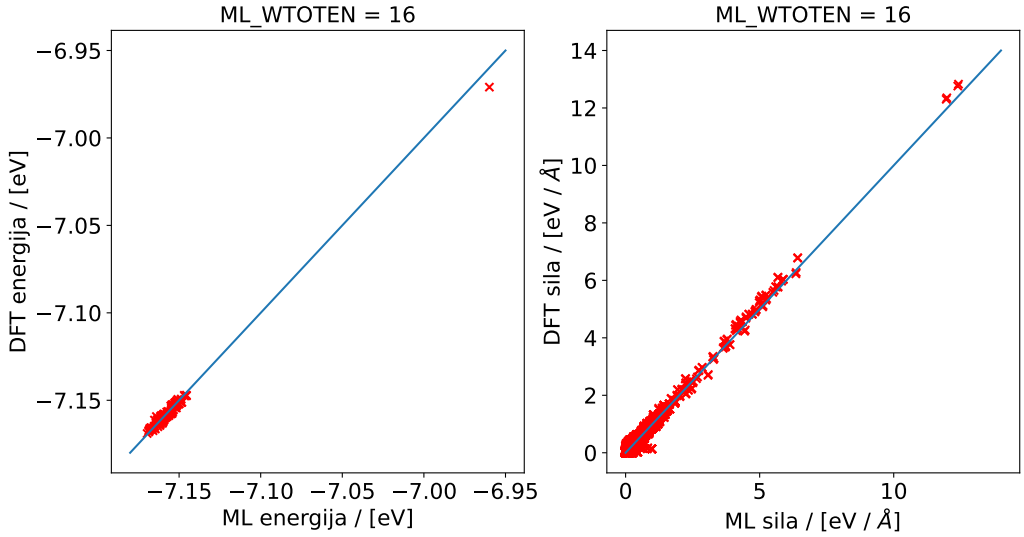
Slika 5.1: Grafički prikaz rezultata za težinski faktor energija iznosa 1. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



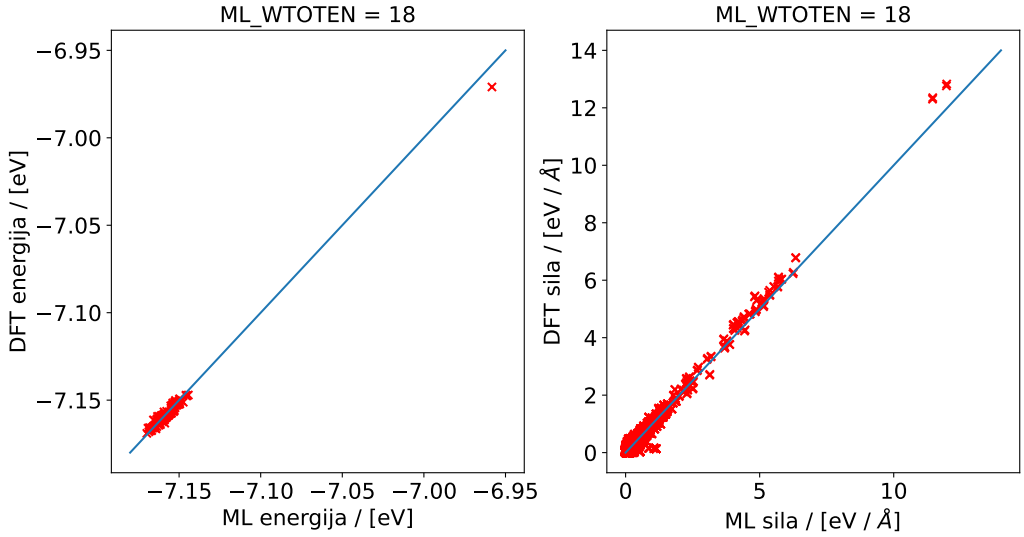
Slika 5.2: Grafički prikaz rezultata za težinski faktor energija iznosa 12. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.3: Grafički prikaz rezultata za težinski faktor energija iznosa 14. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.4: Grafički prikaz rezultata za težinski faktor energija iznosa 16. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.5: Grafički prikaz rezultata za težinski faktor energija iznosa 18. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.

5.3 Težinski faktor deskriptora

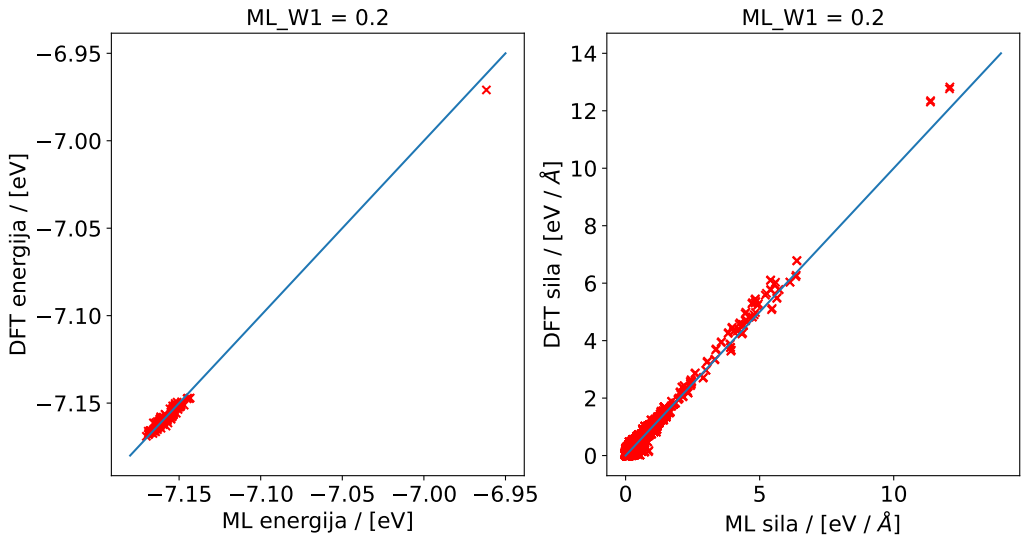
U pokušaju da se rezultati s težinskim faktora energije iznosa 16 dodatno poboljšaju napravljeni su izračuni interatomskog potencijala s varijacijom težinskog faktora des-

kriptora ML_W1. U Tablici 5.2 prikazane su srednje kvadratne pogreške za energije normirane na broj atoma i apsolutnog iznosa sile za različite vrijednosti ML_W1.

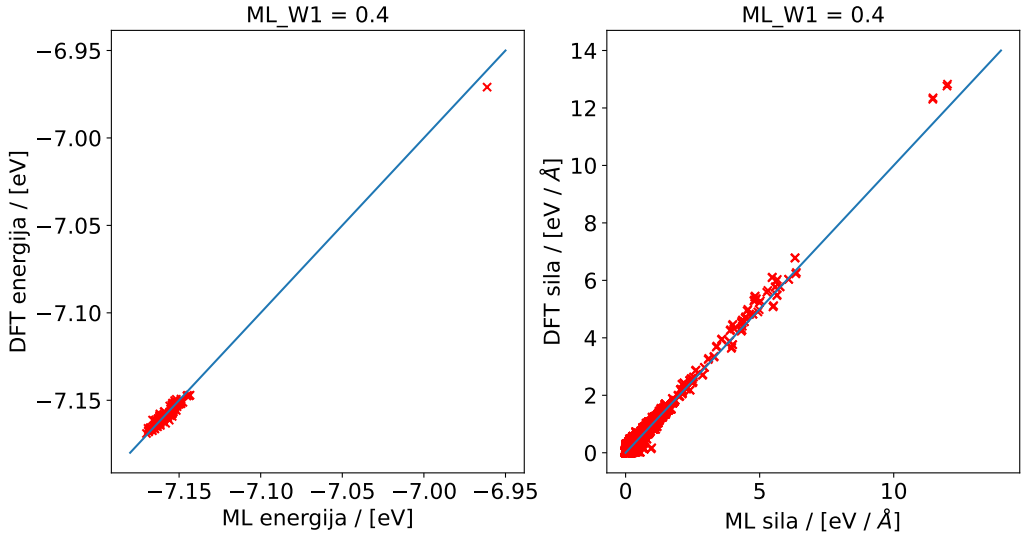
ML_W1	RMSE energija / [eV]	RMSE sile / [eV / Å]
0.2	0.00207	0.04891
0.4	0.00213	0.05120
0.6	0.00205	0.05369
0.8	0.00199	0.05825

Tablica 5.2: Tablični prikaz rezultata regresije interatomskih sila i energija za zadane konformacije. Model je treniran nad 316 DFT konformacija.

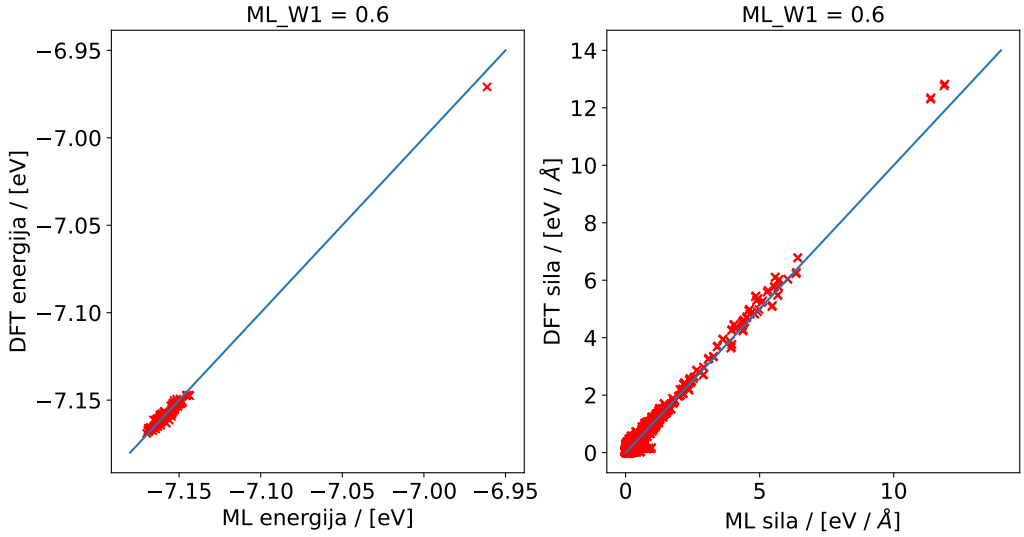
Iz Tablice 5.2 vidljivo je da su iznosi srednjih kvadratnih pogrešaka za energije lošiji za 0.3 do 0.4 meV po atomu od onih pri izboru ML_W1 = 0.1, dok je za sile iznos srednje kvadratne pogreške smanjen pri izboru ML_W1 = 0.2. Također, daljnjim povećavanjem ML_W1 greške za sile rastu, što može značiti da radijalni deskriptori produciraju iste gustoće vjerojatnosti nalaženja atoma na više različitih pozicija koje su jednako udaljene od promatranog atoma. Grafički prikazi rezultata regresija za različite vrijednosti hiperparametra ML_W1 dani su na Slikama 5.6 - 5.9.



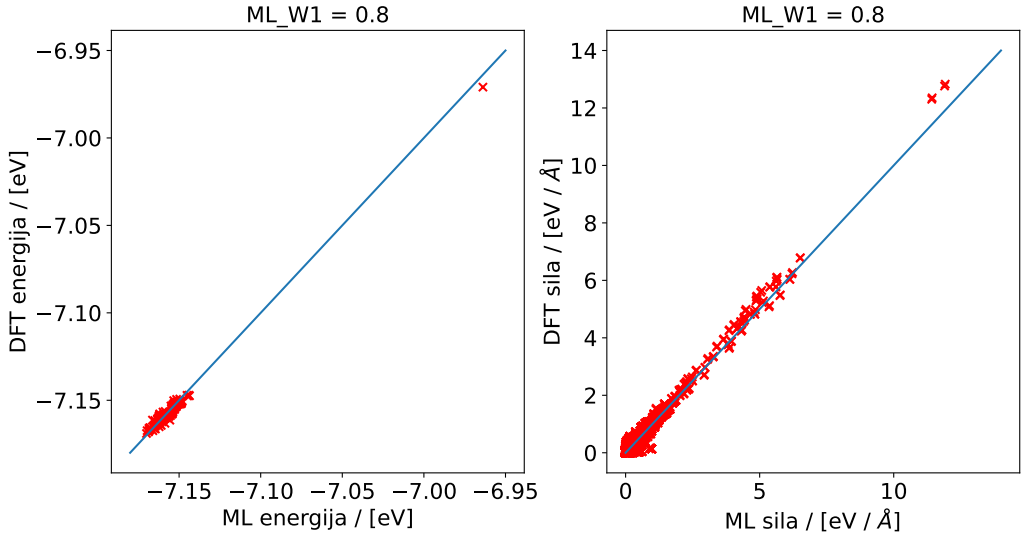
Slika 5.6: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.2. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.7: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.4. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.8: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.6. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.9: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.8. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.

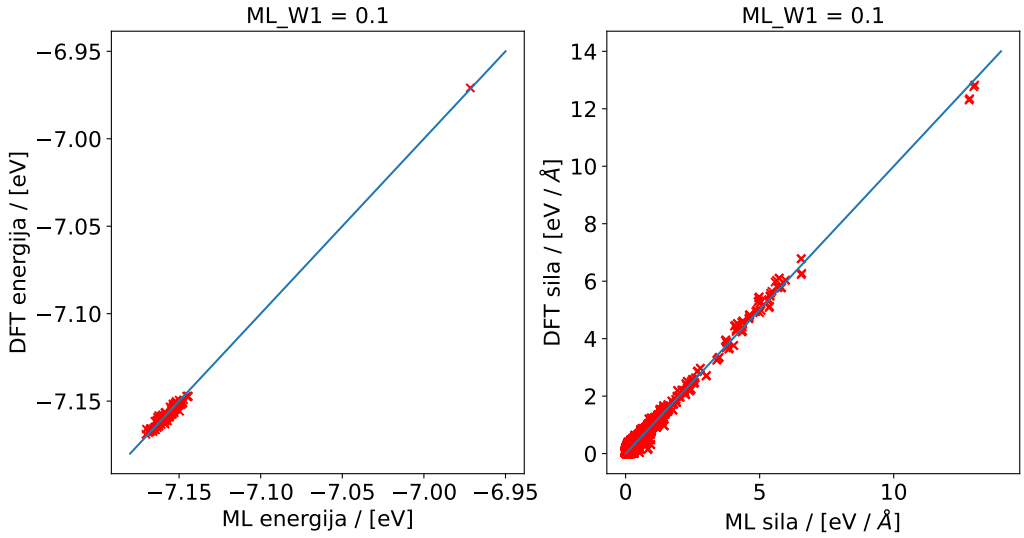
5.4 Dekompozicija na singularne vrijednosti (engl. Singular Value Decomposition - SVD)

Model bayesovske linearne regresije može biti računalno zahtjevan zbog činjenice da je prilikom koraka učenja potrebno kvadrirati problem, odnosno pojavljuje se matrična operacija množenja matrice dizajna same sa sobom. Posljedično, to dovodi do kvadriranja kondicijskog broja matrice, što znači da će predikcije biti osjetljivije na perturbacije u ulaznim podacima. Da se ti efekti umanje, nakon što je model istreniran po prvi put, VASP sadrži opciju dotreniravanja dodatnim iteracijama (engl. continuation run) u kojima se predikcije dobivaju korištenjem dekompozicije matrice dizajna na singularne vrijednosti. Ovim pristupom se može dodatno fino ugoditi naučeni interatomski potencijal. Računi s ovom tehnikom provedeni su za iste vrijednosti težinskog faktora deskriptora kao i u 5.3 i povećanim radijusom rezanja angularnog deskriptora ML_RCUT2 = 4. Rezultati su prikazani u Tablici 5.3.

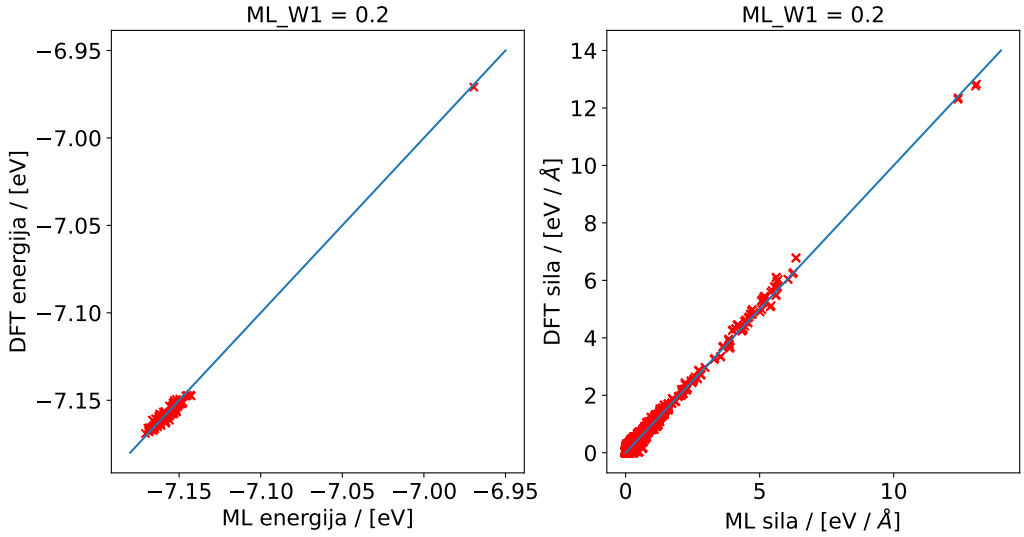
ML_W1	RMSE energija / [eV]	RMSE sile / [eV / Å]
0.1	0.00187	0.05138
0.2	0.00216	0.04473
0.4	0.00218	0.04626
0.6	0.00225	0.04842
0.8	0.00219	0.05160

Tablica 5.3: Tablični prikaz rezultata regresije interatomskih sila i energija za zadane konformacije. Model je treniran nad 316 DFT konformacija koristeći SVD.

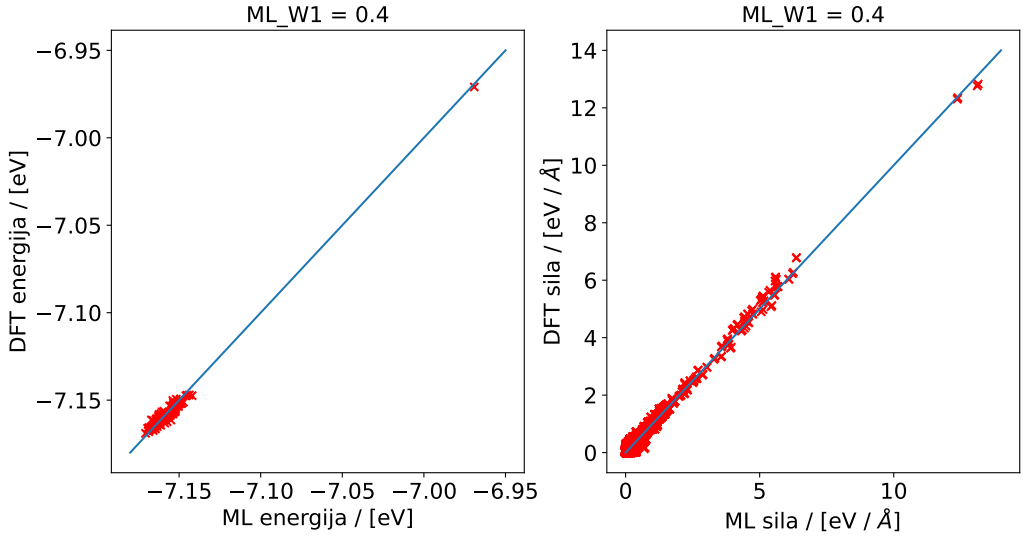
Iz Tablice 5.3 vidi se da je u svim slučajevima preciznost za energije smanjena, dok je preciznost za sile povećana. Grafički prikazi regresija prikazani su na Slikama 5.10 - 5.14.



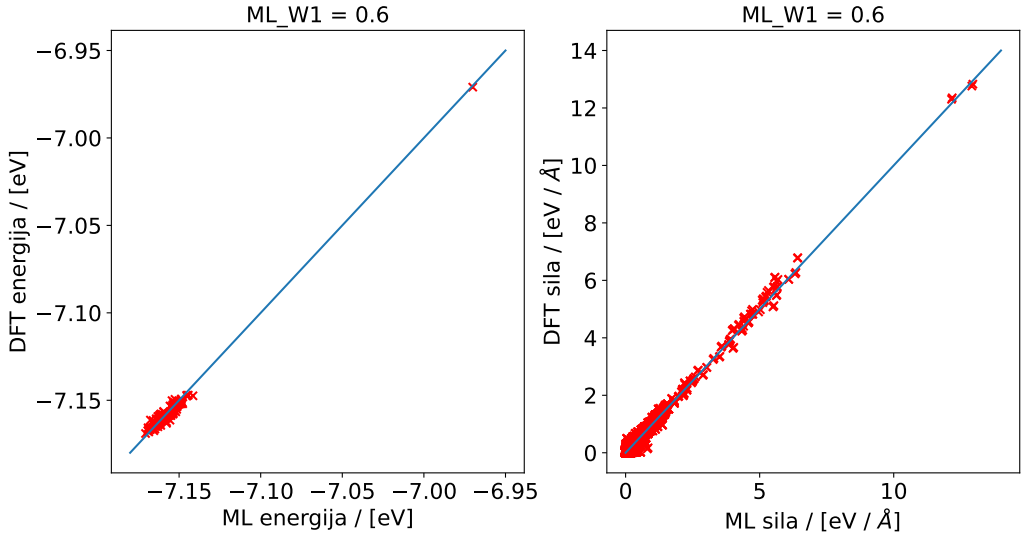
Slika 5.10: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.1 pri korištenju SVD-a. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



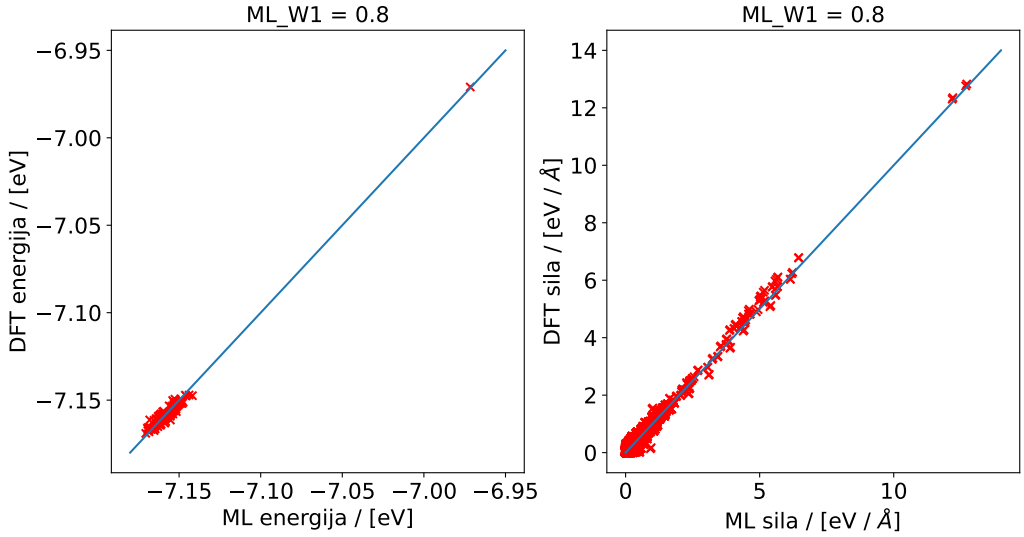
Slika 5.11: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.2. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.12: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.4. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.13: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.6. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.



Slika 5.14: Grafički prikaz rezultata za težinski faktor deskriptora iznosa 0.8. Lijevo: prikaz regresije potencijalnih energija normiranih na broj atoma. Desno: prikaz regresije apsolutnih iznosa sila na atom.

U ovom radu pokazano je da model strojnog učenja rezultira interatomskim potencijalima koji, usprkos relativno malom broju podataka, predviđaju energije i sile na atome za dobivene strukture s točnošću od 2 meV po atomu, odnosno 0.05 eV / Å. Pošto se interatomski potencijal smatra dovoljno preciznim ukoliko predviđa energije

s točnošću od 0.01 eV i sile s točnošću od 0.05 eV / Å [34], model se može smatrati upotrebljivim.

Konačno, potrebno je usporediti izvedbu modela korištenog u ovom radu s izvedbama sličnih modela. Međutim, pošto su drugi modeli primijenjeni na kristalne sustave koji se razlikuju od onih korištenih u ovom radu, objektivna usporedba nije skroz moguća jer su metrike koje su pogodne za neke sustave neprimjenjive za druge. Modeli evaluirani nad manjim molekulama daju bolje predikcije za energije jer je energija ekstenzivna veličina. Model korišten u [35] rezultirao je prosječnim pogreškama od 2.6 meV po atomu za energije, dok su za sile pogreške 0.07 eV / Å, što je gore od rezultata dobivenih u ovom radu, iako se radi o sličnoj klasi spojeva. Model korišten u [36] je producirao dosta veću grešku za energije - 8.9 meV po atomu. Međutim, greška za sile je nešto manja nego u ovom radu - 0.03 eV / Å.

Zaključno, primjenom aktivnog učenja uspješno je treniran model koji producira greške usporedive sa njemu sličnim modelima. Iako ograničen brojem podataka, model je predvidio energije osnovnog stanja i sile na pojedine atome s velikom razinom preciznosti, pritom zahtjevajući puno manje vremena za izvršavanje od DFT računa.

6 Zaključak

Kvantna fizika u principu omogućava potpuno razumijevanje fizike materijala. Ipak, za složene sustave postoje brojne teškoće kao što je računska složenost. Zbog toga je veliki napor uložen u razvitak aproksimativnih metoda koje omogućavaju uvid u fenomene koji proizlaze iz složenosti promatranih sustava. Jedna od najpopularnijih aproksimativnih metoda u modernoj fizici čvrstog stanja naziva se teorijom funkcionala gustoće. Njena snaga leži u tome što može vrlo precizno odrediti svojstva nekog materijala s puno manjom računalnom kompleksnošću u usporedbi s višečestičnom Schrödingerovom jednadžbom. No, čak i uz to smanjenje, algoritmi teorije funkcionala gustoće zahtjevaju veliku količinu računalne memorije i vremena za izvršavanje kada se primjene na sutave s velikim brojem elektrona. S druge strane, postoje i fenomenološke metode koje pokušavaju modelirati materijale korištenjem efektivnih potencijala. Njihova prednost leži u činjenici da su puno brže od *ab initio* metoda, ali s cijenom smanjene preciznosti.

Razvoj strojnog učenja omogućio je cijeli novi spektar metoda koje tvore svojevrsni most između *ab initio* i fenomenoloških metoda. Algoritmi strojnog učenja mogu, uz dovoljnu količinu podataka dobivenih *ab initio* metodama, kreirati efektivne potencijale koji su puno precizniji od onih dobivenih fenomenološkim metodama, ali bez plaćanja cijene prevelikog vremena izvršavanja. Glavni problem kod strojnog učenja je ispravan odabir složenosti modela koji dovodi do minimaliziranih grešaka u predikciji.

Strojno učenje interatomskih potencijala zahtjeva modifikaciju algoritma uzimajući u obzir ograničenja nametnuta fizikalnim simetrijama. Ona se ugrađuju u model putem deskriptora koji opisuju lokalne atomske konfiguracije, a pritom posjeduju sve simetrije potrebne da bi fizikalni zakoni ostali ispoštovani. Također, potrebno je pametno uzorkovati konfiguracijski prostor tako da se u skup za učenje uključe one konfiguracije koje omogućuju modelu da postigne što veću generizabilnost. U ovom radu je u tu svrhu korišteno aktivno učenje, odnosno opetovano labeliranje novih podataka željenim vrijednostima pri treniranju modela.

Proučeno je ponašanje modela bayesovske linearne regresije uz aktivno učenje na strukturama koje tvori metalo-organski spoj koji se sastoji od atoma bakra, ugljika, dušika, klora i vodika. Prednost korištenja bayesovske linearne regresije u sklopu

aktivnog učenja je to što zahtjeva relativno malen broj ulaznih podataka, za razliku od modela baziranih na neuronskim mrežama. Glavni hiperparametri koji su bili ugođavani su težinski faktori energija i deskriptora te metoda koja se koristi za optimizaciju parametara težina u potencijalnoj energiji. Pokazano je da su greške u energijama i silama koje producira ovako odabran model usporedivi s performansama istog tog modela na spojevima hibridnih perovskita. Konačno, pokazano je da je moguće dobiti interatomski potencijal za kompleksne molekulske kristale koji je naučen na relativno malim strukturama.

Literatura

[1] Giustino, F. Materials Modelling using Density Functional Theory. 1st ed. Oxford : Oxford University Press, 2014., str. 19-36, 239

[2] Hohenberg, P. and Kohn, W. Inhomogenous Electron Gas.// Phys. Rev. Vol. 136(1964.)

[3] E. Alpaydin., Introduction to machine learning. MIT press, 2020., str. 17-38

[4] http://mlwiki.org/index.php/Overfitting, 9.8.2022.

[5] https://www.ccdc.cam.ac.uk/Community/blog/what-is-crystal-structure-prediction-csp/, 13.8.2022.

[6] Szabo, A.; Ostlund, N. S. (1996). Modern Quantum Chemistry. Mineola, New York: Dover Publishing, str. 50.

[7] John David Jackson, Classical Electrodynamics, John Wiley and Sons, 1962., str. 12-14

[8] Blinder S.M., House J.E, Mathematical physics in theoretical chemistry, Elsevier, 2019., str. 125-127, 137-141

[9] M. Levy, Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem, Proc. Natl. Acad. Sci. U.S.A. 76, 6062 (1979).

[10] Kohn, W. and Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. // Phys. Rev. Vol. 140 (1965.)

[11] https://en.wikipedia.org/wiki/Active_learning_(machine_learning), 15.5.2023.

[12] Settles, B. Active Learning Literature Survey. // Computer Sciences Technical Report 1648. University of Wisconsin–Madison (2010.), str.8-26.

[13] Unke, O. T.; Chmiela, S.; Sauceda, H. E. et al., Machine learning force fields. // Chemical Reviews. Vol. 121, 16 (2021.), str. 10142–10186.

[14] R. Jinnouchi, F. Karsai, C. Verdi, R. Asahi, and G. Kresse, Descriptors representing two- and three-body atomic distributions and their effects on the accuracy of machine-learned inter-atomic potentials. // The Journal of Chemical Physics Vol. 152, 23 (2020)

[15] A. Glielmo, C. Zeni, A. De Vita, Efficient nonparametric n-body force fields from machine learning. // Phys. Rev. B Vol. 97, 18 (2018)

[16] R. Jinnouchi, F. Karsai, G. Kresse, On-the-fly machine learning force field generation: Application to melting points. // Phy. Rev. B Vol. 100, 1 (2019)

[17] A. P. Bartók, R. Kondor, Gábor Csányi, On representing chemical enviroments. // Phys. Rev. B Vol. 87, 18 (2013)

[18] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. // Phy. Rev. Lett. Vol. 104, 13 (2010)

[19] J. Behler, M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. // Phys. Rev. Lett. Vol. 98, 14 (2007)

[20] J. Shawe-Taylor, N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, 2004., str. 74-82

[21] https://www.vasp.at/wiki/index.php/Machine_learning_force_field:_Theory, 29.8.2022.

[22] C. M. Bishop, Pattern Recognition and Machine Learning, Springer 1st ed. 2006. Corr. 2nd printing, 2006., str. 138-143, 161-172

[23] R. Penrose, A generalized inverse for matrices. // Mathematical Proceedings of the Cambridge Philosophical Society, 51(3), 1955., str. 406-413.

[24] https://en.wikipedia.org/wiki/Bayes%27_theorem, 29.8.2022.

[25] https://en.wikipedia.org/wiki/Mahalanobis_distance, 29.8.2022.

[26] R. Jinnouchi, R. Asahi, Predicting Catalytic Activity of Nanoparticles by a DFT-Aided Machine-Learning Algorithm. // The Journal of Physical Chemistry Letters Vol. 8, 17 (2017)

[27] https://www.ccdc.cam.ac.uk/Community/initiatives/cspblindtests/csp-blind-test-7/, 12.1.2023.

[28] CCDC, , Https://www.ccdc.cam.ac.uk/Community/initiatives/cspblindtests/7-csp-blind-test-targets/, 13.1.2023.

[29] https://www.vasp.at/wiki/index.php/Category:Theory, 13.1.2023.

[30] S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew, J. Gerit Brandenburg , r2SCAN-D4: Dispersion corrected meta-generalized gradient approximation for general chemical applications. // J. Chem. Phys. 154, 061101 (2021) https://doi.org/10.1063/5.0041008

[31] Eike Caldeweyher, Christoph Bannwarth, Stefan Grimme; Extension of the D3 dispersion coefficient model. // J. Chem. Phys. 147, 034112 (2017) . https://doi.org/10.1063/1.4993215

[32] Vandermause, J., Torrisi, S.B., Batzner, S. et al. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. // npj Comput Mater 6, 20 (2020). https://doi.org/10.1038/s41524-020-0283-z

[33] https://wiki.fysik.dtu.dk/ase/, 18.5.2023.

[34] Noé, F.; Tkatchenko, A.; Müller, K.-R. et al. Machine learning for molecular simulation. // Annual review of physical chemistry. Vol. 71 (2020), str. 361–390.

[35] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, M. Bokdam, Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on-the-fly with Bayesian inference. // Phys. Rev. Lett. 122, 225701 (2019), https://doi.org/10.1103/PhysRevLett.122.225701

[36] C. Chen, Z. Deng, R. Tran, H. Tang, I.H. Chu, S. P. Ong, Accurate force field for molybdenum by machine learning large materials data. // Phys. Rev. Mater. 1, 043603 (2017), https://doi.org/10.1103/PhysRevMaterials.1.043603