Određivanje stupnjeva slobode pomoću dubokog učenja

Matej Pavlović
Mentor: Davor Horvatić
Fizički odsjek, Prirodoslovno-matematički fakultet, Bijenička 32, Zagreb

Kako duboke neuronske mreže (DNN) izvlače bitne značajke iz slike je još uvijek nejasno, ali vjeruje se da se izvlačenje značajki odvija hijerarhijski. To dosta podsjeća na koncept renormalizacijske grupe (RG) iz statističke fizike. Kako bi istražili povezanost RG i DNN primijenili smo ograničene Boltzmannove strojeve na 2D Isingov model kako bi generirali tok parametara (temperature) sustava. Pokazalo se da RBM koji je treniran na temperaturama od 0 do 6 K konvergira u sustav blizu kritične temperature od $T_C = 2.27$ K, tj. T_C se ponaša kao atraktor toka.

I. UVOD

Strojno učenje je područje koje u zadnjih par godina dobiva sve veću i veću pozornost. Pokazalo se korisnim u područjima preko računalnog vida, obrade prirodnog jezika, otkrića novih lijekova pa sve do pobijede nad svjetskim prvakom u igri Go. Neuronske mreže daju tako dobre rezultate jer su u mogućnosti kvalitetno izvući stupnjeve slobode sustava, a s dobrim stupnjevima slobode sustav se puno lakše i kvalitetnije može opisati. Presjek fizike i strojnog učenja nalazimo u činjenici da se oba područja bave proučavanjem složenih sustava s mnogo stupnjeva slobode. U ovom istrživanju primijenit ćemo alate strojnog učenja na poznate probleme iz fizike kako bi nam dali neki bolji uvid u sustav. Bavit ćemo se proučavanjem faznog prijelaza 2D Isingovog modela koristeći neuronske mreže. Konkretno, proučavat ćemo ponašanje sustava za tok parametara u prostoru značajki koje neuronska mreža odredi da su relevantne.

II. 2D ISINGOV MODEL[1][2]

Isingov model je matematički model feromagnetizma u statističkoj fizici. 2D sustav modeliramo rešetkom $N \times N$ na čijim se čvorovima nalaze spinovi koji mogu poprimiti vrijednost $\sigma = \{-1, 1\}$. Na stanje čvora utječu stanja njegovih susjeda, a spinovi se rasporede tako da minimiziraju energiju na danoj temperaturi. Hamiltonijan sustava (kada nema vanjskog magnetskog polja) dan je jednadžbom:

$$H(\sigma) = -\sum J_{ij}\sigma_i\sigma_j, \tag{1}$$

gdje suma ide preko parova susjednih spinova, svaki par se broji jednom, J_{ij} je jačina interakcije između dva susjedna čvora i i j. Particijska funkcija sustava na temperaturi T je dana izrazom:

$$Z = \sum_{\sigma^i} e^{-H(\sigma^i)/(kT)}, \tag{2}$$

gdje je k Boltzmannova konstanta, σ^i označava jedno od stanja cijelokupnog sustava, a suma ide po svi mogućim stanjima sustava. Vjerojatnost da se sustav nađe u stanju σ^S dana je jednadžbom:

$$P(\sigma^S) = \frac{e^{-H(\sigma^S)/(kT)}}{Z}, \tag{3}$$

dobivena distribucija se naziva Boltzmannova distribucija.

III. OGRANIČENI BOLTZMANNOV STROJ[3][4]

Ograničeni Boltzmannovi strojevi (Restricted Boltzmann Machine - RBM) su generativni stohastički energetski modeli koji uče distribuciju ulaznih podataka. Najjednostavniji RBM se sastoji od jednog vidljivog i jednog skrivenog sloja (Slika 1[5]), mreža nije usmjerena što znači da informacija može ići od vidljivog do skrivenog sloja i obrnuto.
Mreža ima skup parametara $\lambda = \{a, w, b\}$, gdje je a vektor slobodnih parametra vidljivog sloja, b vektor slobodnih parametra skrivenog, a w matrica težina koja povezuje skriveni i vidljivi sloj. Energetska funkcija[6] RBM-a je dana izrazom:

$$E(v, h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i w_{i,j} h_j \tag{4}$$

$$E(v, h) = -a^T v - b^T h - v^T W h, \tag{5}$$

gdje je v vidljivi (ulazni) vektor, a h je skriveni vektor. Cilj RBM-a je naučiti distribuciju $p_\lambda(v)$ što bližu stvarnoj distribuciji $P(v)$. Potrebno je naglasiti da je ulazni i skriveni vektori su binariziran $v_i,\ h_i \in \{0, 1\}$. Skriveni vektor se računa kao

$$h(v) = S(v^T W + b),\ S(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

izlaz iz funkcije S je ograničen intervalom $[0, 1]$ i predstavlja vjerojatnost da se dani čvor postavi na vrijednost 1 (upravo zbog toga se ovakvi modeli nazivaju stohastičkim).

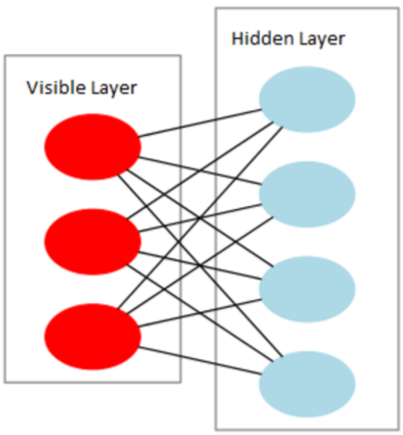
Vjerojatnost da se mreža poprimi stanje zadano s v i h je:

$$p_\lambda(v,h) = \frac{e^{-E(v,h)}}{\sum_{v^i,h^i \in \text{stanje}} e^{-E(v^i,h^i)}} \tag{7}$$

gdje suma ide po svim mogućim stanjima vektora v i h. Također možemo sumirati po svim stanjima vektora $v(h)$ i dobiti vjerojatnost vektora h (v) neovisno o vektoru $v(h)$, jednadžbe[7] (8) i (9).

$$p_\lambda(v) = \mathrm{Tr}_h(p_\lambda(v,h)) \tag{8}$$

$$p_\lambda(h) = \mathrm{Tr}_v(p_\lambda(v,h)) \tag{9}$$



Slika 1: Shematski prikaz RBM-a.

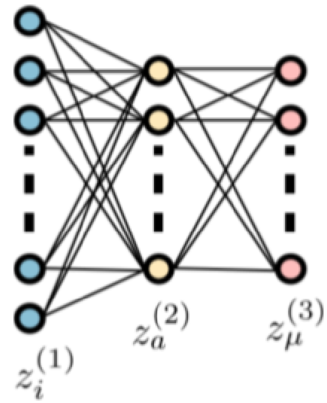
Minimiziranjem Kullback-Leibler divergencije (D_{KL}) pronalazimo optimalne parametre λ^{opt}, D_{KL} je metrika sličnosti dvije distribucije.

$$D_{KL}(P(v)||p_\lambda(v)) = \sum_v P(v)\log(\frac{P(v)}{p_\lambda(v)}) \tag{10}$$

Parametri λ se treniraju metodom gradijentog spusta. Nakon treniranja za dani ulaz v dobiveni vektor h predstavlja latentne varijable, također možemo iz vektora h generirati novi vektor v (stanje sustava) koji će biti nešto drukčiji od početnog.
Kako bi mogli odrediti temperaturu dobivenog sustava implementirali smo neuronsku mrežu za mjerenje temperature. Ulaz u mrežu su spinovi sustava dok je izlaz distribucija vjerojatnosti za temperaturu. Mreža se sastoji od jednog skrivenog sloja (Slika 2[5]) i trenirana je da minimizira unakrsnu entropiju izlaza i stvarne vrijednosti temperature. Temperatura je binarizirana u 25 binova. Parametri mreže su trenirani gradijentnim spustom kao i kod RBM-a.

$$L = -\sum_{c=1}^{M} y_i^c \log(p_i^c) \tag{11}$$

Jednadžba (11) je izraz za unakrsnu entropiju gdje je M broj klasa (25 u našem slučaju), a $y_i^c \in \{0,1\}$ je oznaka pripada li i-ti primjer klasi c, dok p_i^c je predviđa vjerojatnost da i i-ti primjer pripada klasi c.



Slika 2: Shematski prikaz modela za temperaturu, u svakom čvoru se nalazi aktivacijska funkcija.

IV. REZULTATI I ANALIZA

IV.1. Generiranje podataka

Podatke 2D Isingovog modela generiramo pomoću Monte-Carlo simulacije. Prvo generiramo 2D mrežu dimenzija $L \times L$ s nasumičnim vrijednostima $\sigma_{x,y} = \pm 1$, bitno je napomenuti da zahtijevamo periodične rubne uvijete:

$$\sigma_{L,y} = \sigma_{0,y},\ \sigma_{x,L} = \sigma_{x,0}. \tag{12}$$

Generiranje spinske konfiguracije sustava na temperaturi T vršimo pomoću Monte-Carlo simulacije. Nakon generiranja početne mreže nasumično odaberemo spin $\sigma_{x,y}$ te ga okrenemo s vjerojatnošću[5]:

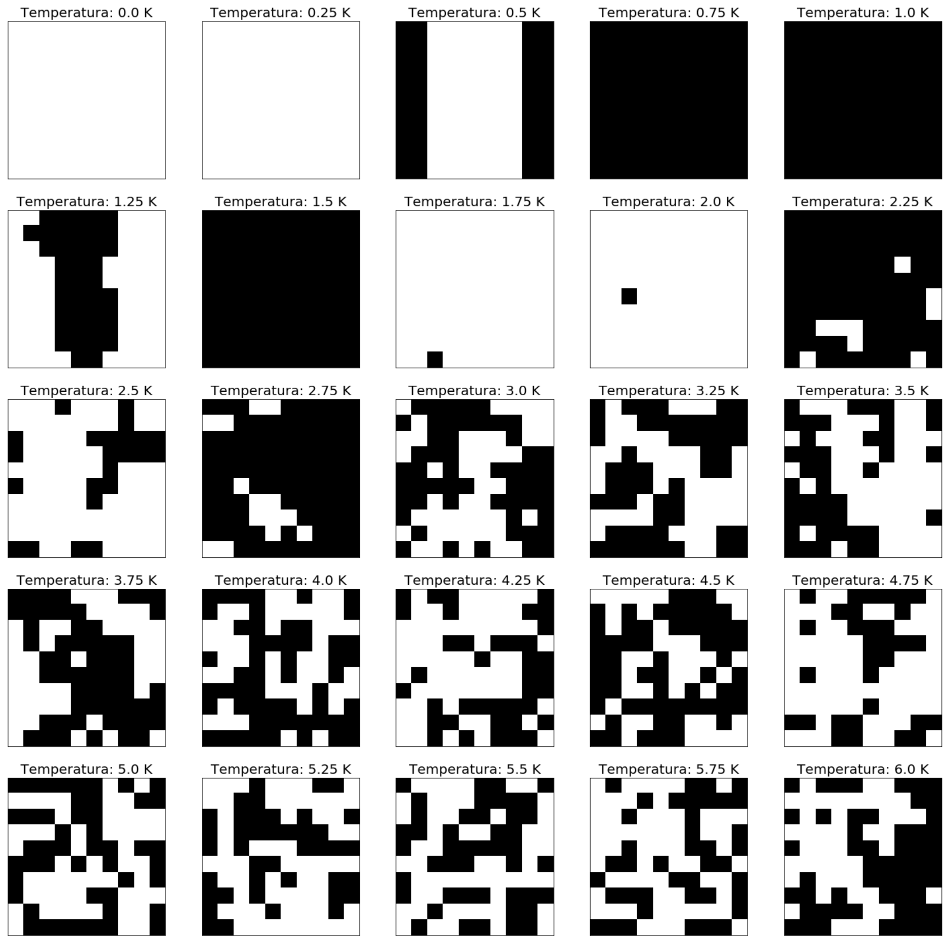
$$p_{x,y} = \begin{cases} 1, & dE_{x,y} < 0 \\ e^{-dE_{x,y}/k_B T}, & dE_{x,y} > 0 \end{cases} \tag{13}$$

gdje je $dE_{x,y}$ promjena energije sistema kada okrenemo spin na poziciji x,y. Iz energije interakcije između elektrona možemo izraziti $dE_{x,y}$ kao:

$$dE_{x,y} = 2J\sigma_{x,y}(\sigma_{x+1,y} + \sigma_{x-1,y} + \sigma_{x,y+1} + \sigma_{x,y-1}). \tag{14}$$

Bez gubitka općenitosti možemo postaviti konstante J i k_B na 1.

Veličinu mreže smo postavili na $L \times L = 10 \times 10$, te Monte-Carlo postupak ponavljamo $100L^2 = 10000$ puta, tj. 10000 puta nasumično odaberemo elektron te ga okrenemo s vjerojatnošću danom jednadžbom 13. Sustav generiramo na temperaturama od 0 K do 6 K s inkrementom od 0.25 K, za svaku temperaturu izgeneriramo 1000



Slika 3: 25 nasumično izvučenih primjera iz skupa za treniranje, za svaku temperaturu prikazujemo jedan primjer.

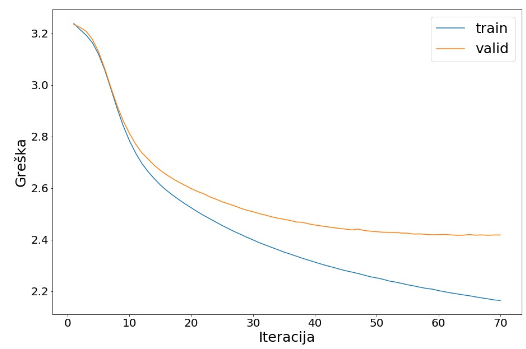
puta, cijeli postupak ponovimo dva puta kako bi imali skup za treniranje i skup za testiranje. Na slici 3 su prikazani primjeri izgeneriranih podataka.

IV.2. Temperaturni model

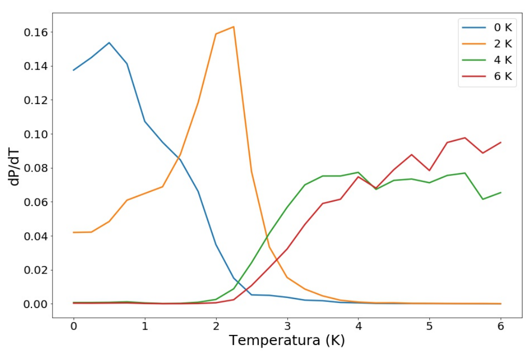
Model za određivanje temperature iz spinske konfiguracije sastoji se od neuronske mreže s jednim skrivenim slojem. Brojevi neurona u svakom sloju dani su redom ulazni sloj 100, skriveni sloj 64 i izlazni sloj 25. Ulazni broj neurona određen je time što nam je rešetka dimenzija 10×10 koju "rastegnemo" u vektor duljine 100, dok je izlazni sloj određen brojem temperatura koje smo generirali. Aktivacijska funkcija u skrivenom sloju je $f(x) = \tanh(x)$. Zbog prirode problema, spinovi mogu biti ± 1 dok je domena tanh od -1 do 1, izabrana je takva aktivacijska funkcija. Aktivacijska funkcija na izlaznom sloju je softmax:

$$f(\vec{X})_i = \frac{e^{x_i}}{\sum_i e^{x_i}}. \tag{15}$$

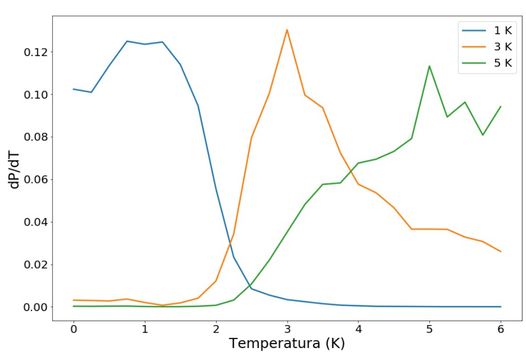
Softmax nam vraća vjerojatnost za svaki "bin", što nam omogućava da promatramo distribuciju po temperaturi. Na Slici 4 je prikazana greška na skupu za treniranje i validaciju ovisno o broju iteracija, da smo treniranje pustili još iteracija model bi se prenaučio, greška na skupu za treniranje bila bi značajno manja nego na skupu za validiranje. Slika 5 i 6 prikazuju temperature koja mreža mjeri na skupu podataka koji nije korišten za treniranja.



Slika 4: Greška na skupu za učenje i validaciju o ovisnosti o broju iteracija kod modela za temperaturu.



Slika 5: Distribucije vjerojatnosti temperature za sustave na temperaturama 0, 2, 4 i 6 K.



Slika 6: Distribucije vjerojatnosti temperature za sustave na temperaturama 1, 3 i 5 K.

Možemo primijetiti da model daje maksimum distribucije na temperaturi sustava. Maksimum nije uvijek oštar vrh ali s grafa možemo očitati o kojoj temperaturi se radi. Ako pogledamo sliku 3 (primjer podataka) vidimo da za npr. sustav na 0.75 K i 1.5K skoro pa i nema razlike (na danom primjeru) pa je za očekivati da će ih model teže razlikovati.

IV.3. Tok parametara

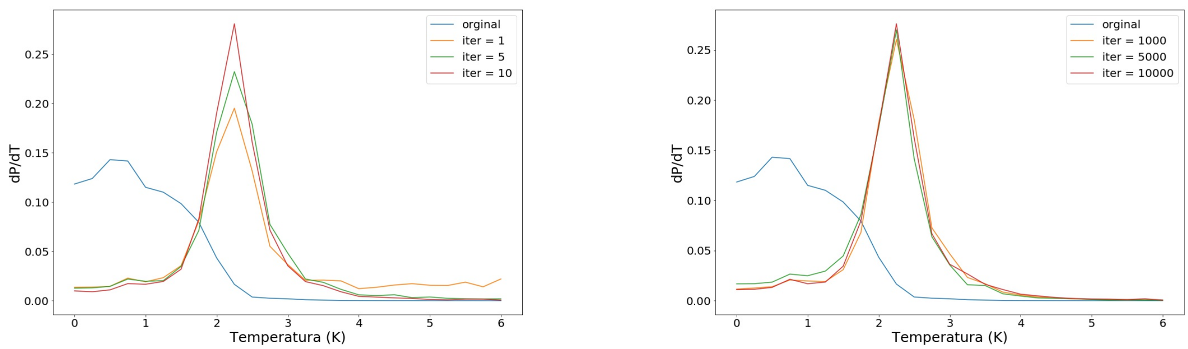
Nakon što se istrenira RBM sa skrivenim slojem od 64 neurona koristimo ga za generiranje sekvence distribucije sustava (tok temperature). Stohastički dio unutar RBM-a nam garantira da se ne izgenerira isti sustav nego sustav koji ima slične latente varijable kao i početni. Krenemo od nekog sustava $v_i^{(0)}$ i generiramo sustave kao:

$$v_i^{(0)} \to h_i^{(1)} \to v_i^{(1)} \to h_i^{(2)} \to v_i^{(2)} \to ... \tag{16}$$

što nam daje tok konfiguracije sustava, tj. kada koristimo model za mjerenje temperature dobijemo tok distribucije temperature sustava. Bitno je napomenuti da je sustav mapiran u 0 i 1 prije ulaska u RBM, tako da trebamo paziti da izgenerirani sustav mapiramo u ± 1 prije prolaska kroz temperaturni model.
Promatramo tok temperature za različite početne uvjete, početni sustavi su oni iz skupa za validaciju tako da mreže se nikad nisu učile na njima. Na slikama 7, 8 i 9 vidimo ponašanje sustava kada su početne temperature 0, 2.5 i 6 K. Grafovi pokazuju distribucije temperature sustava ovisno o iteraciji, na lijevoj slici je distribucija za mali broj iteracija, dok je na desnoj za veliki. Rezultati indiciraju da tok koji generiramo pomoću RBM-a konvergira u stabilnu kritičnu temperaturu T_C, čak i nakon velikoj broj iteracija distribucija temperature ostane skoro pa ista.
Kritična temperaturu 2D Isingovog modela je 2.27 K, a sustav nam je konvergirao u 2.25 K (najbliži bin koji imamo). Također vidimo da sustav uvijek konvergira u T_C neovisno o početnoj temperaturu sustava.

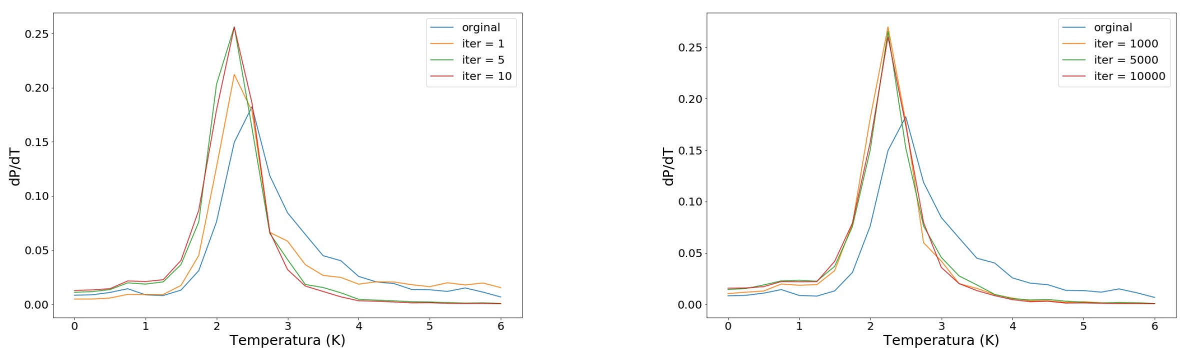
V. ZAKLJUČAK

Strojno učenje i fizika imaju puno dodirnih točaka, te se alati stojnog učenja mogu iskoristiti kako bi bolje razumjeli fiziku. Upotrijebili smo ograničene Boltzmannove stojeve kako bi dobili stupnjeve slobode Isingovog modela.
Prvo smo izgenerirali stanja 2D Isingovog modela pomoću Monte-Carlo simulacije. Koristili smo po 1000 stanja za temperature od 0 K do 6 K, s korakom od 0.25 K, jedan takav skup je bio za učenje, a drugi za validaciju. Zatim smo natrenirali model za mjerenje temperature sustava koji nam je vraćao distribuciju po temperaturama. Dva modela smo spojili tako da je RBM generirao stanja iz nekog početnog dok smo s drugim modelom mjerili temperaturu dobivenog sustava.
Dobili smo da generirani sustavi konvergiraju u sustav



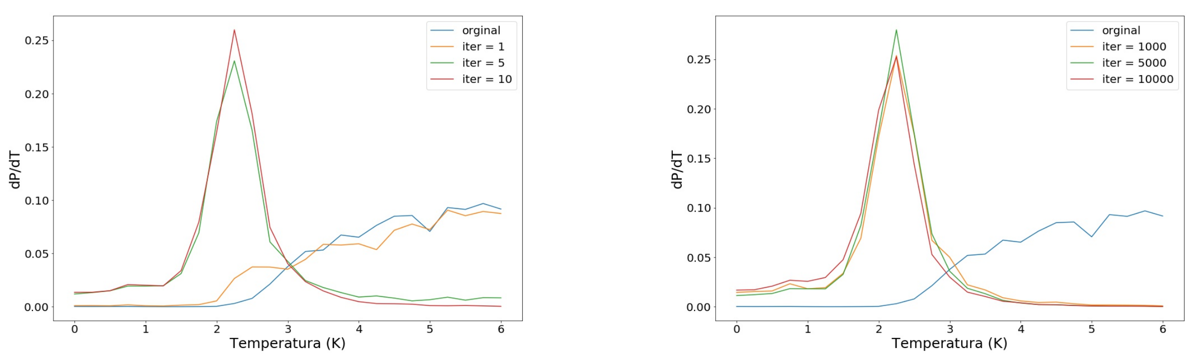
(a) (b)

Slika 7: Distribucija temperatura sustava nakon raznog broja iteracija toka. Početni sustav se nalazio na 0 K, a konačni se stabilizirao u blizini T_C



(a) (b)

Slika 8: Distribucija temperatura sustava nakon raznog broja iteracija toka. Početni sustav se nalazio na 2.5 K, a konačni se stabilizirao u blizini T_C



(a) (b)

Slika 9: Distribucija temperatura sustava nakon raznog broja iteracija toka. Početni sustav se nalazio na 6 K, a konačni se stabilizirao u blizini T_C

koji ima temperaturu $T_C = 2.25$ K, također smo pokazali da temperatura početnog sustava nema utjecaj na temperaturu konvergencije, tj. T_C se ponaša kao atraktor toka RBM-a.

Ovakva metodologija mogla bi nam dati kritičnu temperaturu sustava koje ne možemo teorijski odrediti, te proučavanje skrivenog vektora RBM-a nam sigurno može dati još korisnih informacija o sustavu.

LITERATURA

[1] Barry M. McCoy and Tai Tsun Wu, "The Two-Dimensional Ising Model". Harvard University Press, Cambridge Massachusetts (1973), ISBN 0-674-91440-6
[2] Baxter, Rodney J., "Exactly solved models in statistical mechanics", London: Academic Press Inc. (1982), ISBN 978-0-12-083180-7
[3] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science 313 (2006) 504.
[4] G. E. Hinton, "A principal guide to training restricted Boltzmann machines," *Neural networks: Tricks of the Trade*, Springer (2012) pp.599-619.
[5] Satoshi Iso at al. "Scale-invariant Feature Extraction of Neural Network and Renormalization Group Flow", In: *ArXiv* (Jan. 2018). arxiv: 1801.07172
[6] LeCun, Yann at al. "A Tutorial on Energy-Based Learning" . In Predicting Structured Data. Neural Information Processing series. MIT Press (2006). ISBN 978-0-26202617-8. [7] Luo Di, "Machine Learning, Renormalization Group and Phase Transition", (2017)