

# Određivanje stupnjeva slobode pomoću dubokog učenja

Matej Pavlović

Mentor: Davor Horvatić

*Fizički odsjek, Prirodoslovno-matematički fakultet, Bijenička 32, Zagreb*

# Motivacija

- Strojno učenje se pokazalo korisnim u mnogim područjima (računalni vid, NLP, Go)
- Modeli strojnog učenja dobro pronalaze stupnjeve slobode
- Princip sličan renormalizacijskim grupama u fizici
- Primjena strojnog učenja na probleme iz područja fizike u svrhu boljeg razumijevanja istih
- Proučavamo 2D Isingov model

# Renormalizacijske grupe (RG)

- Binarni sistem opisan vektorom  $\mathbf{v}$ ,  $\{v_i\} = \pm 1$
- Particijska funkcija i distribucija vjerojatnosti:

$$Z = \sum_{v \in \text{states}} e^{-H(v)} \quad P(v) = \frac{e^{-H(v)}}{Z}$$

- Općenito:  $H(v) = - \sum_i K_i v_i - \sum_{i,j} K_{i,j} v_i v_j - \sum_{i,j,k} K_{i,j,k} v_i v_j v_k + \dots$
- RG je mapiranje originalnog Hamiltoniana u novi Hamiltonian s drukčijim konstantama vezivanja ( $K'_i, \dots$ ) i u drugu skalu
- Na novoj skali sistem je opisan vektorom  $\mathbf{h}$

- Hamiltonian u novoj skali:  $H^{RG}(h) = - \sum_i K'_i h_i - \sum_{i,j} K'_{i,j} h_i h_j - \sum_{i,j,k} K'_{i,j,k} h_i h_j h_k + \dots$
- Nakon mnogo iteracija RG-a neke konstante vezivanja će nestati
- One koje prežive nam daju važne informacije o sustavu
- RG transformacija uvodi operator  $T(v,h)$  koji ovisi o parametrima  $\lambda$
- Hamiltonian nakon RG transformacije:

$$e^{-H_\lambda^{RG}(h)} = \sum_{v \in \text{states}} e^{T_\lambda(v,h) - H(v)}$$

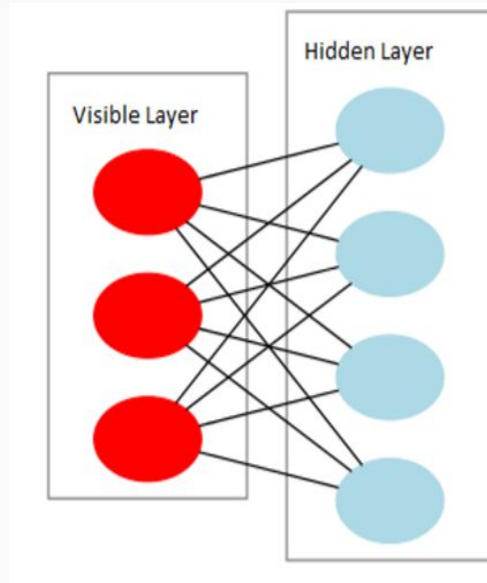
- $T(v,h)$  variramo po  $\lambda$  kako bi minimizirali

$$\log\left(\sum_{v \in \text{states}} e^{-H(v)}\right) - \log\left(\sum_{h \in \text{states}} e^{-H_\lambda^{RG}(h)}\right)$$

# Ograničeni Boltzmanovi strojevi (RBM)

- Jedan od najpopularnijih algoritama dubokog učenja
- Jedan vidiljivi sloj ( $v$ ) i jedan skriveni sloj ( $h$ )
- $v$  i  $h$  povezani matricom težina  $w_{ij}$  te slobodnim parametrima  $b_j$  i  $c_i$
- $\lambda = \{b, w, c\}$
- Energija mreže

$$E(v, h) = \sum_j b_j h_j + \sum_i c_i v_i + \sum_{i,j} w_{ij} v_i h_j$$



Slika 1: Shematski prikaz RBM-a.

- Vjerojatnost da je mreža u stanju  $\mathbf{v}, \mathbf{h}$ :  $p_\lambda(v, h) = \frac{e^{-E(v, h)}}{Z}$
- Možemo sumirati po jednoj varijabli i dobiti da vjerojatnost ovisi samo o drugoj

$$p_\lambda(v) = \frac{\sum_{h \in \text{state}} e^{-E(v, h)}}{Z} = \frac{e^{-H_\lambda^{\text{RBM}}(v)}}{Z}$$

$$p_\lambda(h) = \frac{\sum_{v \in \text{state}} e^{-E(v, h)}}{Z} = \frac{e^{-H_\lambda^{\text{RBM}}(h)}}{Z}$$

- Minimiziramo Kullback-Leibler divergenciju (sličnost dvije distribucije)

$$D_{KL}(P(v) || p_\lambda(v)) = \sum_{v \in \text{state}} P(v) \log \frac{P(v)}{p_\lambda(v)}$$

# Mapiranje između RG i RBM

- Izaberimo:  $T_\lambda(v, h) = -E(v, h) + H(v)$

- Slijedi:

$$\frac{e^{-H_\lambda^{RG}(h)}}{Z} = \frac{\sum_{v \in \text{states}} e^{T_\lambda(v, h) - H(v)}}{Z} = \frac{\sum_{v \in \text{states}} e^{-E(v, h)}}{Z} = p_\lambda(h) = \frac{e^{-H_\lambda^{RBM}(h)}}{Z}$$

- $\Rightarrow H_\lambda^{RG}(h) = H_\lambda^{RBM}(h)$
- Hamiltonian RG-a je isti kao i Hamiltonian skrivenog sloja RBM-a!
- Varijacija RG uvjeta je isto što i minimizacija  $D_{KL}$

# 2D Isingov model

- Atomi složeni u ćeliji mogu imati pozitivan i negativan magnetski moment
- Matematički model feromagnetizma
- $N \times N$  rešetka sa vrijednostima  $\sigma = \{-1, 1\}$
- $H(\sigma) = - \sum J_{ij} \sigma_i \sigma_j$ , suma po susjedima,  $J_{ij}$  je jačina interakcije
- Particijska funkcija je suma po svim stanjima tj:

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

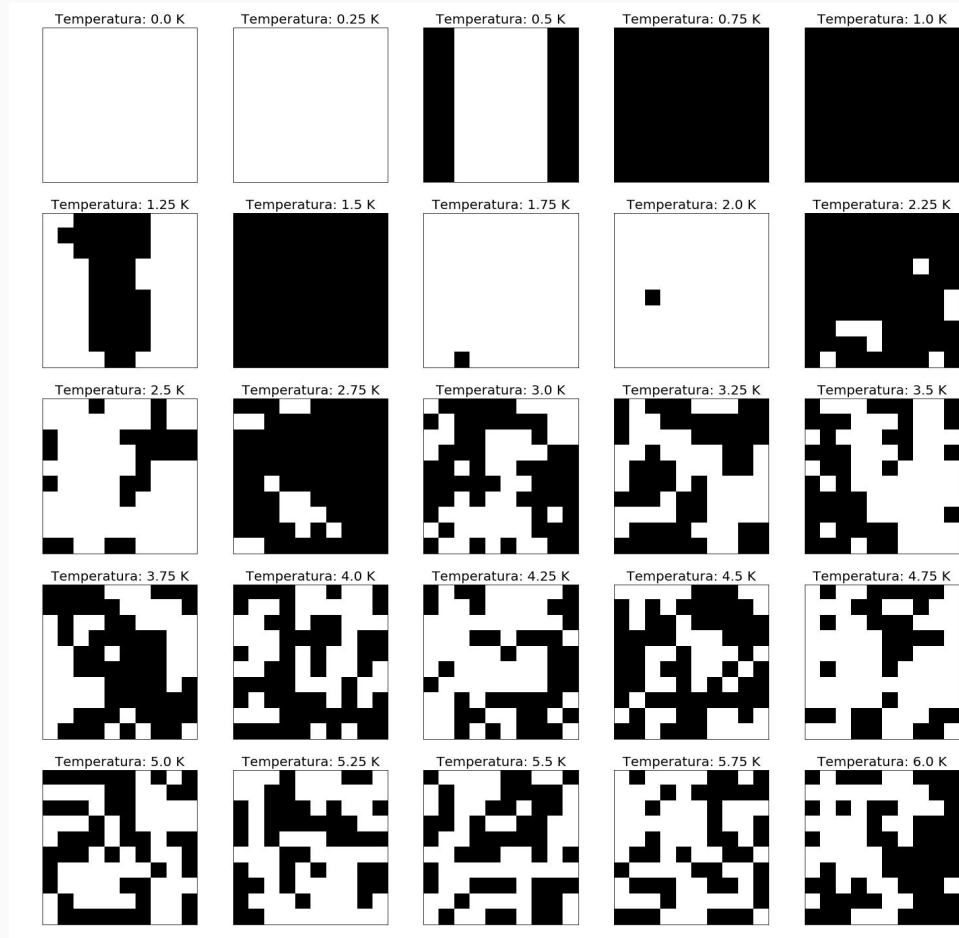
- Gdje  $\sigma^i$  predstavlja jedno od mogućih stanja sustava



- Vjerojatnost da se sustav nalazi u stanju S (Boltzmanova distribucija):

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

- Kod faznog prijelaza, sustav mijenja ponašanje
- Za 2D Isingov model fazni prijelaz je na  $\sim 2.27$  K
- Podaci su generirani Monte-Carlo simulacijom
- Izgenerirali smo po 1000 primjera za temperature od 0 K do 6 K s inkrementom od 0.25 K



Slika 2: Primjeri iz skupa podataka.

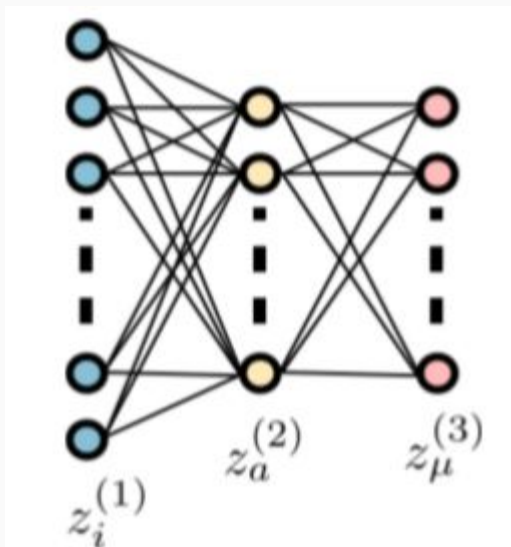
# Temperaturni model

- Za dano stanje sustava predviđa distribuciju vjerojatnosti temperature
- Neuronska mreža s jednim skrivenim slojem (100-64-25)
- Aktivacijska funkcija skrivenog sloja je  $\tanh(x)$
- Izlazna funkcija softmax:

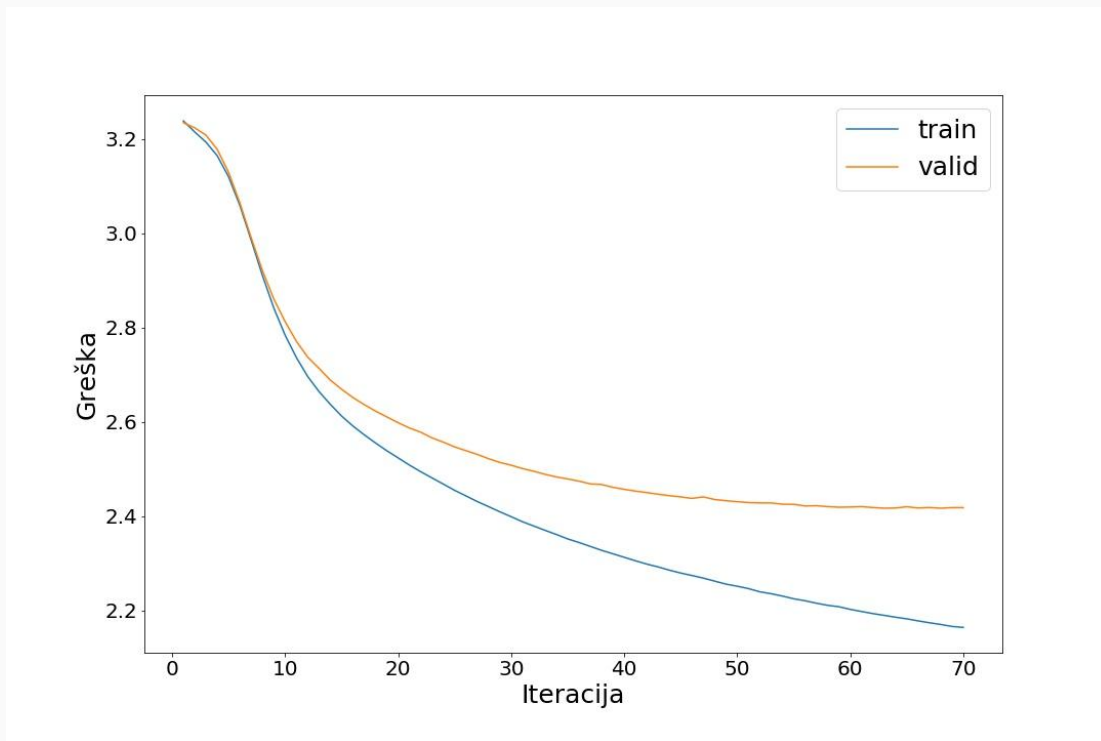
$$f(\vec{X})_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

- Minimizira se unakrsna entropija:

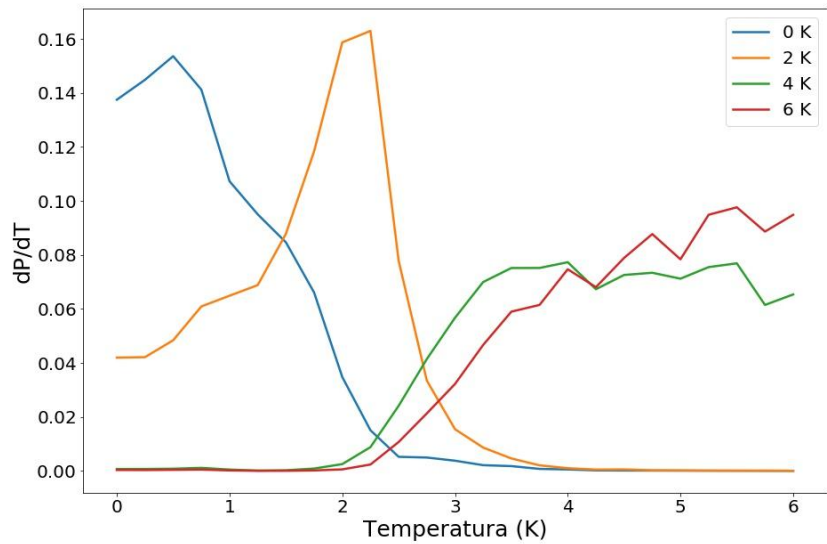
$$L = - \sum_{c=1}^M y_i^c \log(p_i^c)$$



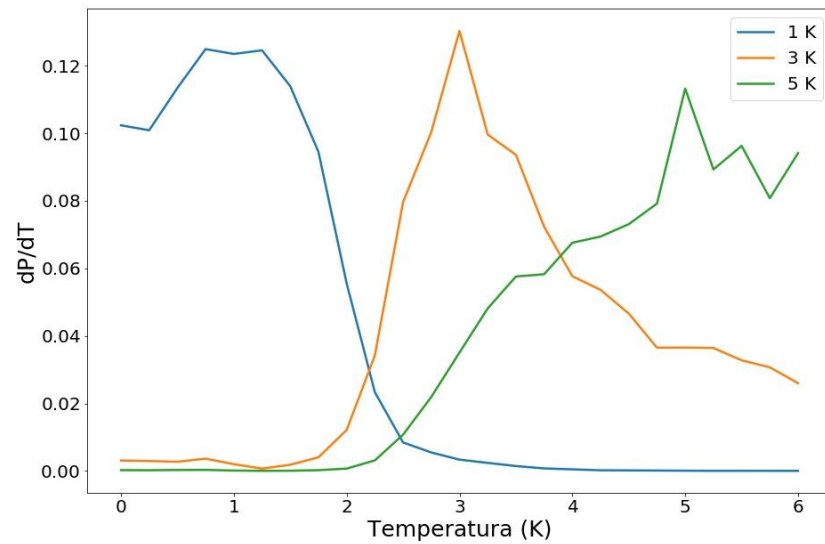
Slika 3: Shematski prikaz modela za temperaturu.



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



Slika 5



Slika 6

Slike 5 i 6 prikazuju distribucije vjerojatnosti temperature za razne temperature sustava.

# Tok parametara

- Nakon što se RBM istrenira, koristimo ga za generiranje sekvence distribucije sustava (tok temperature)

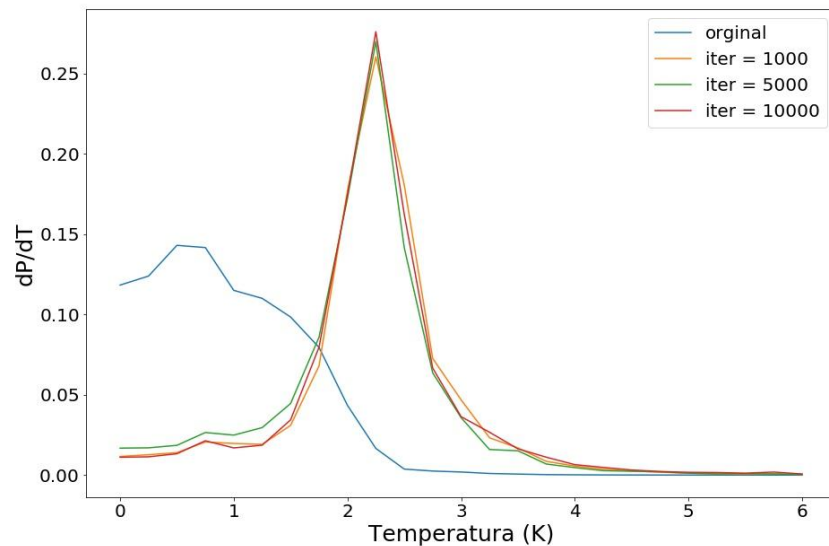
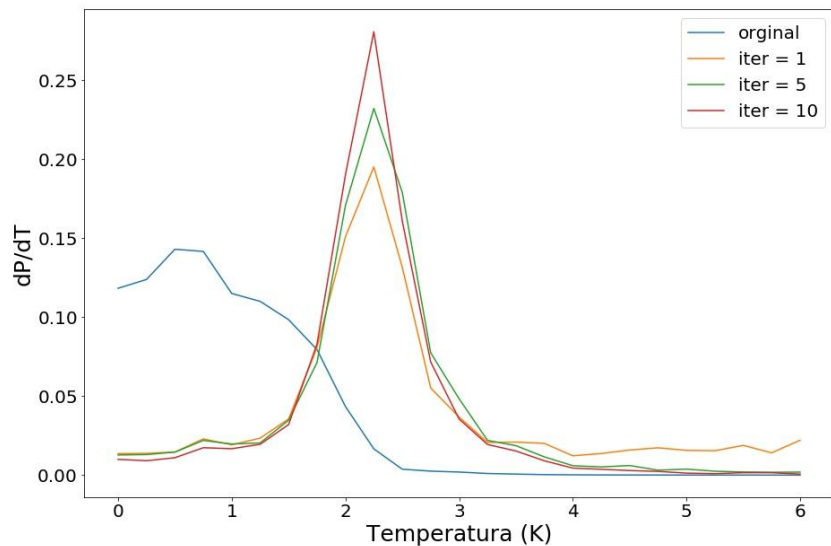
- Ekvivalentno iterativnoj primjeni renormalizacijske grupe

- Krenemo od sustava sa stanjem  $v_i^{(0)}$

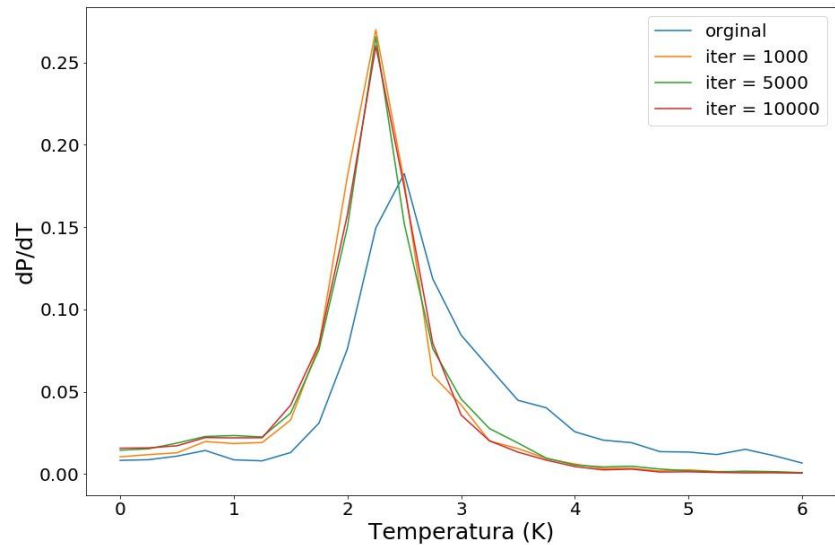
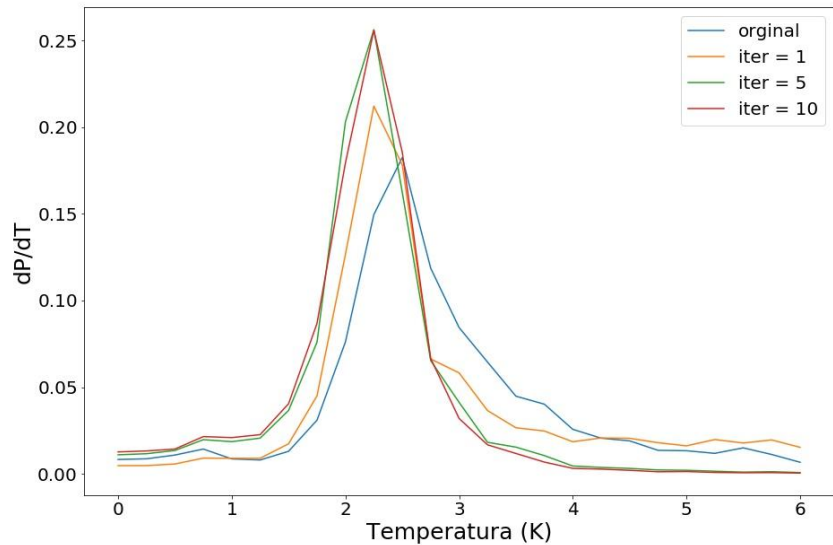
$$v_i^{(0)} \rightarrow h_i^{(1)} \rightarrow v_i^{(1)} \rightarrow h_i^{(2)} \rightarrow v_i^{(2)} \rightarrow \dots$$

- Stohastički dio RBM-a nam osigurava da je stanje  $v_i^{(k)}$  različito od  $v_i^{(k-1)}$
- Nakon svakog koraka tražimo distribuciju temperature stanja  $v_i^{(k)}$

- Promatramo tok distribucije temperature za različite početne uvjete

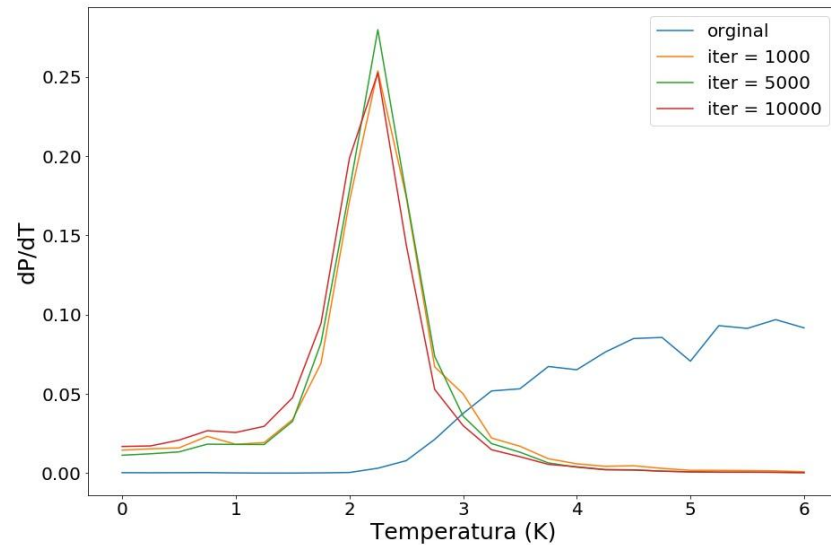
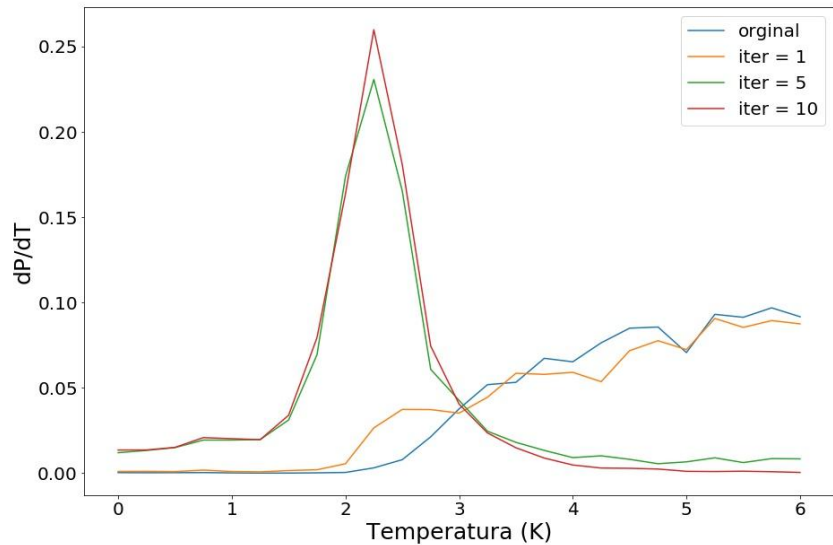


Slika 7: Distribucija temperatura sustava nakon raznog broja iteracija toka. Početni sustav se nalazio na 0 K, a konačni se stabilizirao u blizini  $T_C$ .



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





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

- Kritična temperaturu 2D Isingovog modela je 2.27 K, a sustav nam je konvergirao u 2.25 K (najbliži bin koji imamo)
- Nakon velikoj broj iteracija distribucija temperature ostane skoro pa ista
- RBM-a konvergira u stabilnu kritičnu temperaturu  $T_c$
- Neovisno o početnoj temperaturu sustava

# Zaključak

- Strojno učenje i fizika imaju puno dodirnih točaka
- Alati stajnog učenja se mogu iskoristiti kako bi bolje razumjeli fiziku
- Skiveni sloj RBM-a je ekvivalentan primjeni renormalizacijske grope
- RBM je generirao stanja iz nekog početnog sustava dok smo s drugim modelom mjerili temperaturu dobivenog sustava
- Sustav konvergira u temperaturu  $T_c$
- Ovakva metodologija mogla bi nam dati kritičnu temperaturu sustava koju ne možemo teorijski odrediti
- Moguće preciznije mjerenje temperature uz generiranje podataka u užem intervalu

# 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)

Hvala na pažnji!