

Lze simulovat fyziku pomocí neuronových sítí?

D. Šafrata¹, O. Matyšek², L. Hradečný³ a J. Fišerová⁴

¹GEVO Sázavská; david@klan.cz

²G ALTIS, Praha; ondramat007@gmail.com

³G Cihelní, Frýdek-Místek; hradečnýlukas@gmail.com

⁴G Olomouc Hejčín; fiserovaj@gytool.cz

Garant: Bc. Filip Bartoš; KM FJFI ČVUT

Abstrakt

Asi každý už slyšel o neuronových sítích, jak generují text, obrázky, rozpoznávají předměty a další. Ale lze je použít i pro simulaci fyzikálních jevů? Odpověď je ano. My jsme se rozhodli nasimulovat distribuci teploty v homogenní čtvercové desce, což je problém z oblasti parciálních diferenciálních rovnic.

1 Úvod

S rozvojem neuronových sítí v posledních dekáдах přirozeně vyvstává otázka, jestli je možné je použít k simulování fyzikálních jevů. Jeden z problémů řešitelných tímto způsobem je např. šíření tepla. My jsme zvolili šíření tepla ve 2D. To je charakterizované touto rovnicí (ve 2D tvaru):

$$\frac{\partial T}{\partial t}(t, x, y) = \alpha \left(\frac{\partial^2 T}{\partial x^2}(t, x, y) + \frac{\partial^2 T}{\partial y^2}(t, x, y) \right), \quad t \in [t_0, t_k], \quad x, y \in [0, L], \quad (1)$$

kde T je teplota, která je závislá na čase a na x -ové a y -ové souřadnici desky. Proměnná t značí čas a x, y jsou prostorové souřadnice. Tato závislost je přirozená – teplota se nemusí nutně rovnat ve všech bodech desky v jednom čase, stejně jako se nemusí rovnat v jednom bodě v jiných časech. Jinými slovy, teplota se mění v prostoru i v čase, což odpovídá konceptu distribuce tepla.

Naše hodnoty spadají do určitých intervalů, hraniční hodnoty času jsou t_0 , tedy počáteční čas, a t_k , tedy koncový čas. Prostorové proměnné x, y musí ležet uvnitř naší čtvercové desky, takže obě spadají do výše uvedeného intervalu, kde L je délka strany čtverce, který představuje naši desku. Koeficient α závisí na vybraném materiálu v simulaci, je to tzv. tepelná difuzivita. Fyzikálně se řídí vztahem $\alpha = \frac{\lambda}{\rho c}$, tedy podíl tepelné vodivosti a součinu měrné tepelné kapacity a hustoty daného materiálu.

Jelikož se jedná o parciální diferenciální rovnici (PDE), je nutné mít určité podmínky. Podmínky máme dvě, počáteční a okrajovou. První je tzv. Dirichletova okrajová podmínka, která má tvar

$$T(t, x_b, y_b) = 0,$$

kde $t \in [t_0, t_k]$ a (x_b, y_b) jsou body z okraje čtverce $[0, L] \times [0, L]$ představujícího naši desku. Říká tedy, že ve všech okrajových bodech bude vždy teplota zafixována na nule.

Druhá podmínka nám říká, jak vypadá počáteční stav naší simulace, tedy hodnoty pro všechna přípustná x, y v čase t_0 . Jako počáteční podmínku jsme zvolili

$$T(t_0, x, y) = \sin(\pi x) \sin(\pi y), \quad x, y \in [0, L].$$

2 Neuronové sítě

Na konstrukci architektury neuronové sítě jsme použili knihovnu PyTorch, která slouží pro práci a vývoj neuronových sítí a modelů se strojovým učením a je již optimalizovaná pro paralelizaci výpočtů na grafické kartě.

Učení neuronové sítě funguje na bázi postupného upravování vah na základě tzv. ztrátové funkce, která udává chybu rozdílů mezi predikcí a analytickou hodnotou, přičemž lze najít lokální minimum této funkce, a tudíž snížit chybovost neuronové sítě. Fyzikou informovaná neuronová síť (PINN) se liší v tom, že do ztrátové funkce jsou zakomponované fyzikální zákony, a proto lze naučit neuronovou síť jak simulovat danou situaci bez porušení fyzikálních zákonů. Na výpočet ztrátové funkce využíváme tzv. mean squared error, která pracuje s kvadráty odchylek od správných hodnot. Ztrátová funkce L je definovaná vztahem

$$L = \sum_{i=1}^{N_{rov}} \left(\frac{\partial \hat{T}(t_i, x_i, y_i)}{\partial} - \alpha \left(\frac{\partial^2 T}{\partial x^2}(t, x, y) + \frac{\partial^2 T}{\partial y^2}(t, x, y) \right) \right)^2 + \sum_{j=1}^{N_{poč}} \left(\hat{T}(t_0, x_j, y_j) - \sin(\pi x_j) \sin(\pi y_j) \right)^2 + \sum_{k=1}^{N_{okr}} \hat{T}(t_k, x_k, y_k)^2$$

Simulace PDEs pomocí neuronových sítí má oproti numerickým přístupům jednu podstatnou výhodu a tou je škálování problému do vyšších dimenzí, tedy přidávání nezávislých proměnných, na kterých je zkoumaná proměnná závislá. I při vyvíjení našeho programu jsme nejdříve vyřešili problém v 1D. Následné přidání dimenze, tedy přesun problému do 2D, již nepředstavovalo velký problém.

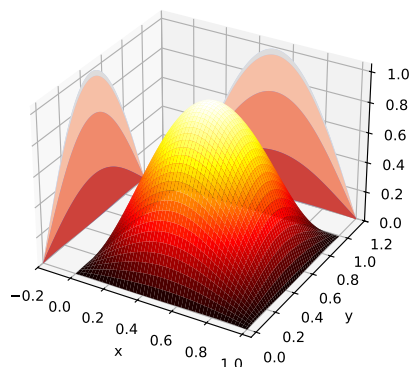
Neuronové sítě ale mají samozřejmě mnoho parametrů, od kterých se odvíjí jejich náročnost, ale i přesnost jejich výsledků. Mezi tyto parametry patří počet vrstev neuronů v naší síti a i počet neuronů v jednotlivých vrstvách sítě, jinými slovy velikost naší sítě. Dále tu máme parametry týkající se samotného trénování, tedy učení, naší neuronové sítě. Mezi ně patří tzv. koeficient trénovacího kroku, který udává, jak rychle budeme měnit váhy neuronové sítě, a počet trénovacích iterací, tedy kolikrát budeme upravovat váhy neuronové sítě.

3 Výsledky

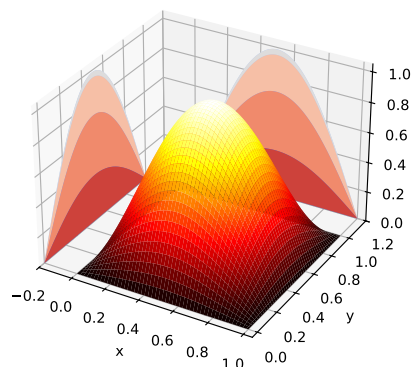
Náš model je plně propojená neuronová síť a obsahuje 3 skryté vrstvy o 50 neuronech a jeden výstupní neuron. Používáme aktivační funkci hyperbolický tangens na aktivaci neuronů. Pro optimalizaci ztrátové funkce jsme využili optimizer Adam s rychlostí učení 10^{-3} a s 10000 epochami. Dohromady jsme na trénování použili 8000 kolokačních bodů, 400 bodů pro okrajové podmínky a 400 bodů pro počáteční podmínky. Hodnotu pro difuzní konstantu jsme použili $\alpha = 0.05$.

Z těchto výsledků je zřejmé, že neuronová síť dokáže aproximovat teplotu desky v závislosti na pozici. Zároveň se aproximace liší více při vyšších hodnotách t . Ztrátová funkce rovněž dosahuje relativně nízkých hodnot vůči počátečním hodnotám ztrátové funkce.

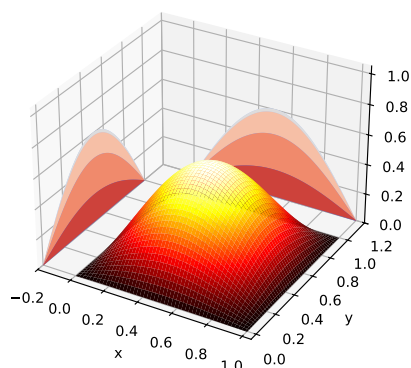
3D plot teploty v čase $t = 0.0$



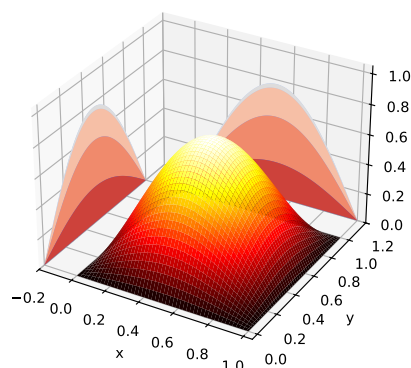
3D plot teploty v čase $t = 0.0$



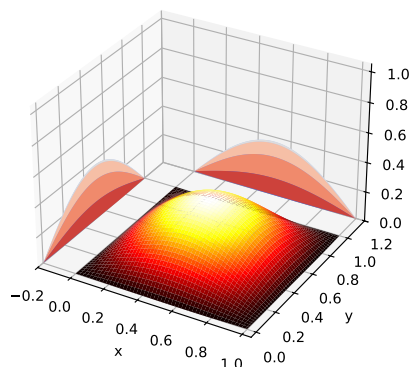
3D plot teploty v čase $t = 0.5$



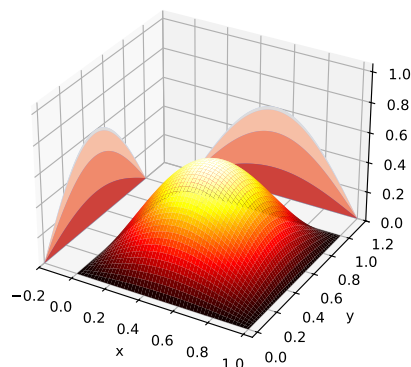
3D plot teploty v čase $t = 0.5$



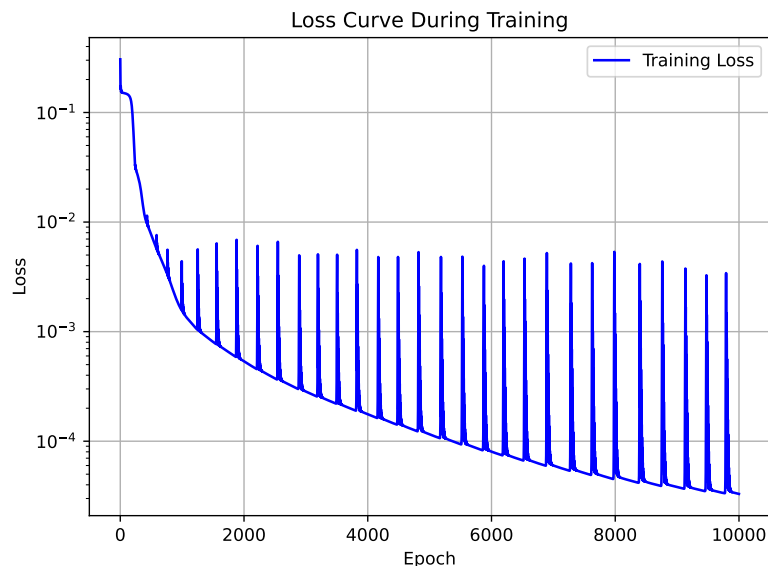
3D plot teploty v čase $t = 1$



3D plot teploty v čase $t = 1$



Porovnání aproximace teploty neuronové sítě (vlevo) a správné hodnoty (vpravo) v časech $t = 0, 0.5$, a 1 v závislosti na pozici.



Hodnoty ztrátové funkce v závislosti na počtu iteračních kroků.

4 Závěr

Naše práce spočívala v simulaci neuronovou sítí vedení tepla v jednodimenzionální homogenní tyči a v dvoudimenzionální homogenní čtvercové desce neuronové sítě informované fyzikou. Řešili jsme rovnici vedení tepla. Neuronová síť se úspěšně naučila aproximovat teplotu v závislosti na poloze na tyči a čase a na desce s nízkou hodnotou ztrátové funkce. Naše simulace potvrdily, že lze simulovat fyziku pomocí neuronových sítí, a že není nutné používat existující data k trénování, pokud naučíme neuronovou síť používat fyzikální zákony na své aproximace.

Poděkování

Tímto bychom rádi poděkovali našemu garantovi Bc. Filipovi Bartošovi za odborné vedení našeho miniprojektu. Dále děkujeme Fakultě jaderné a fyzikálně inženýrské ČVUT za poskytnutí prostor a hardwarového vybavení.