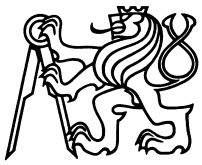
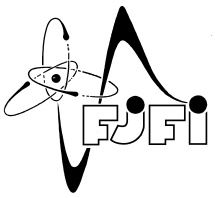


České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská



Robustní strojové učení a adversariální vzorky

Robust machine learning and adversarial examples

Bakalářská práce

Autor: **Pavel Jakš**

Vedoucí práce: **Mgr. Lukáš Adam, Ph.D.**

Akademický rok: 2021/2022

- Zadání práce -

- Zadání práce (zadní strana) -

Poděkování:
Chtěl bych zde poděkovat především svému školiteli - panu doktoru Adamovi - za pečlivost, ochotu, vstřícnost a odborné i lidské zázemí při vedení mé bakalářské práce.

Čestné prohlášení:
Prohlašuji, že jsem tuto práci vypracoval samostatně a uvedl jsem všechnu použitou literaturu.

V Praze dne 7. července 2022 Pavel Jakš

Název práce:

Robustní strojové učení a adversariální vzorky

Autor: Pavel Jakš

Obor: Matematická informatika

Druh práce: Bakalářská práce

Vedoucí práce: Mgr. Lukáš Adam, Ph.D., Katedra počítačů, Fakulta elektrotechnická, České vysoké učení technické v Praze, Karlovo náměstí 13, 121 35, Praha 2

Abstrakt: Abstrakt max. na 10 řádků.

Klíčová slova: klíčová slova (nebo výrazy) seřazená podle abecedy a oddělená čárkou

Title:

Robust machine learning and adversarial examples

Author: Pavel Jakš

Abstract: Max. 10 lines of English abstract text.

Key words: keywords in alphabetical order separated by commas

Obsah

Úvod

Text úvodu....

Kapitola 1

Neuronové sítě

1.1 Hluboká dopředná neuronová síť

1.2 Konvoluční neuronové sítě

1.2.1 Konvoluce

1.2.2 Pooling

Kapitola 2

Učení neuronové sítě

2.1 Účelové funkce

2.1.1 Střední kvadratická chyba

2.1.2 Ztráta křížové entropie

2.2 Algoritmus zpětného šíření chyby

2.3 Algoritmy učení

2.3.1 Gradientní sestup

2.3.2 Stochastický gradientní sestup

2.3.3 Adam

Kapitola 3

Adversariální vzorky

3.1 Metody generování adversariálních vzorků

3.1.1 FGSM

3.1.2 Iterativní FGSM

Kapitola 4

Robustní učení neuronové sítě

Závěr

Text závěru....

Literatura

[1] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*. MIT Press, 2016.

[2] I. Goodfellow, J. Shlens, C. Szegedy, *Explaining and Harnessing Adversarial Examples*. In 'International Conference on Learning Representations', ICLR 2015.

[3] J. Nocedal, S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.