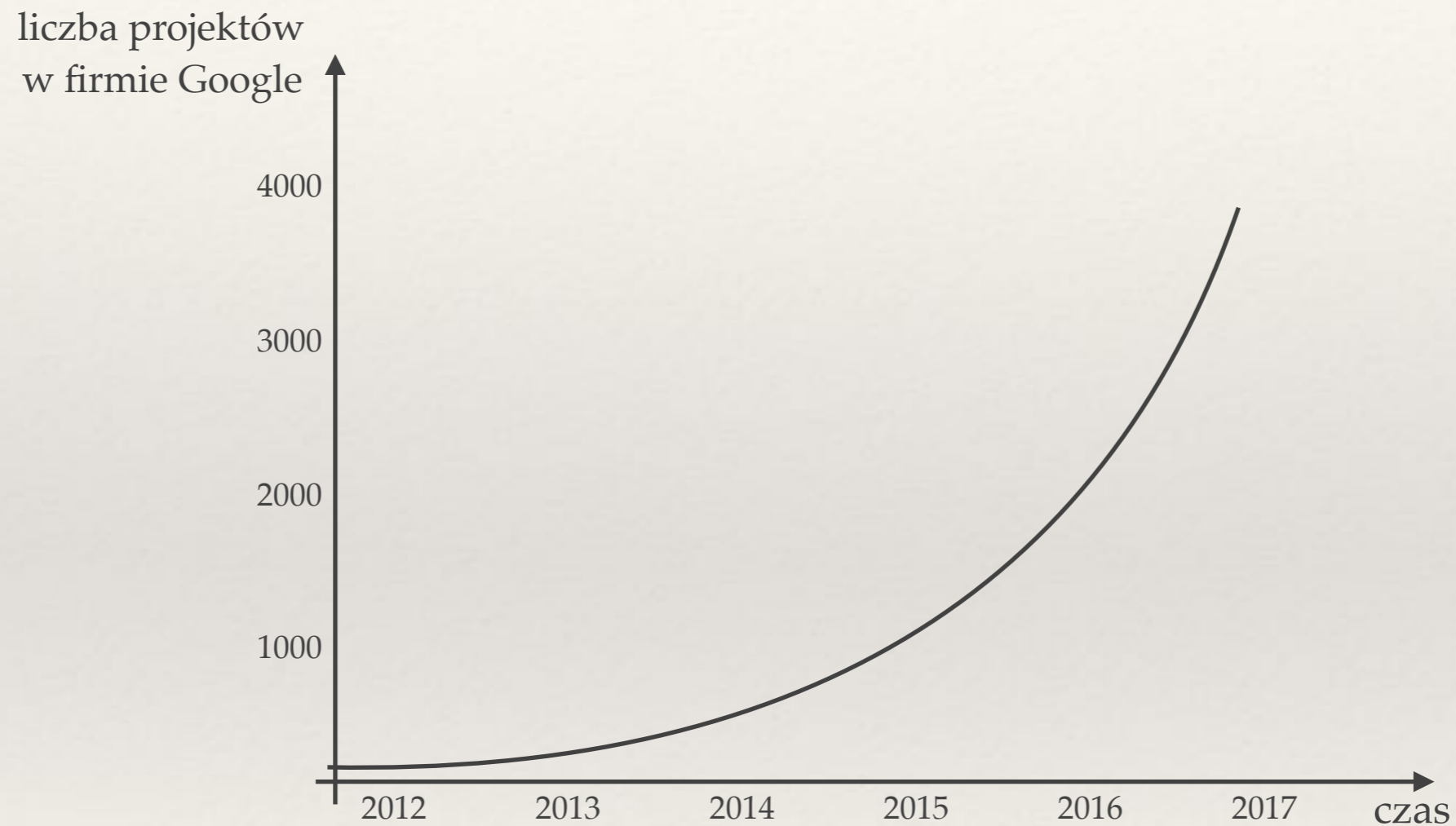

Wstęp do głębokich sieci neuronowych

Paweł Morawiecki
IPI PAN

Liczba projektów z głębokim uczeniem rośnie bardzo szybko



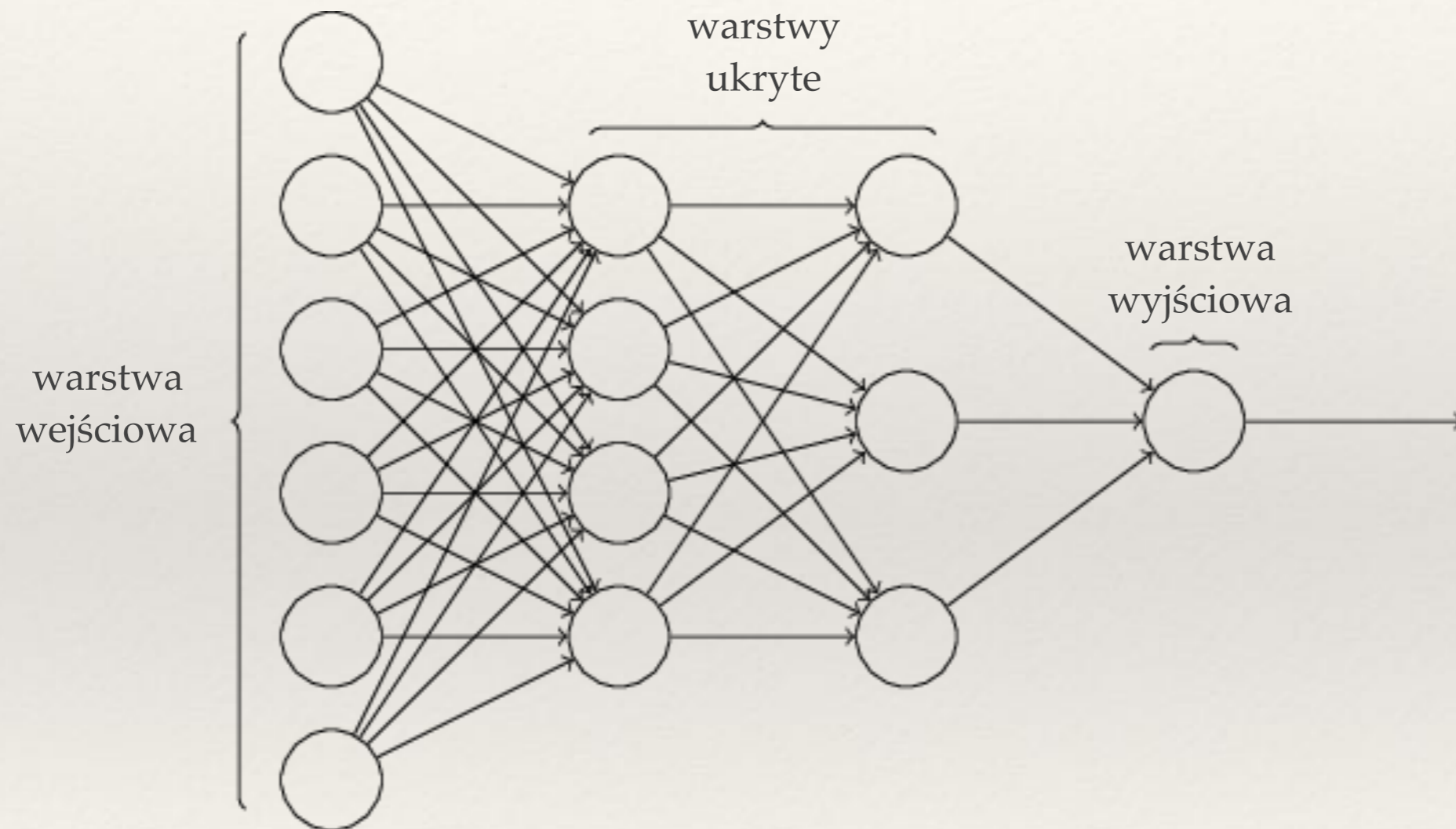
Głębokie sieci neuronowe podbijają NLP

Z Facebookowego profilu Stanford NLP Group (21 marca 2017):

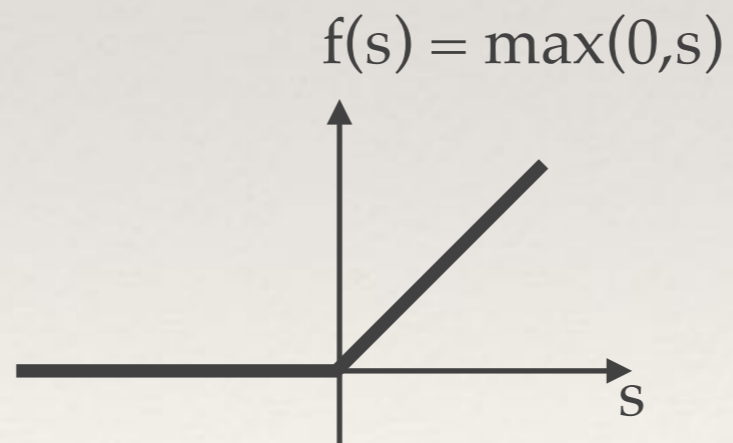
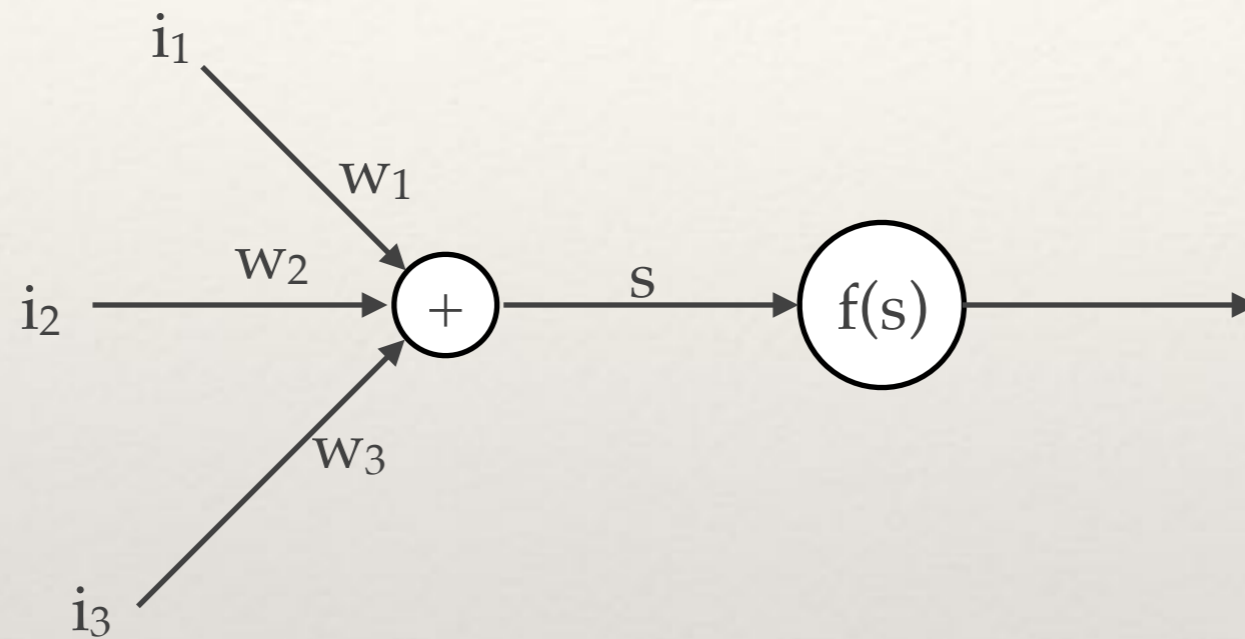
The 2017 Stanford CS224N poster session will showcase projects in Natural Language Processing with Deep Learning that students have worked on over the past quarter. This year, over 650 students will be presenting over 300 projects, organized over two sessions.

!

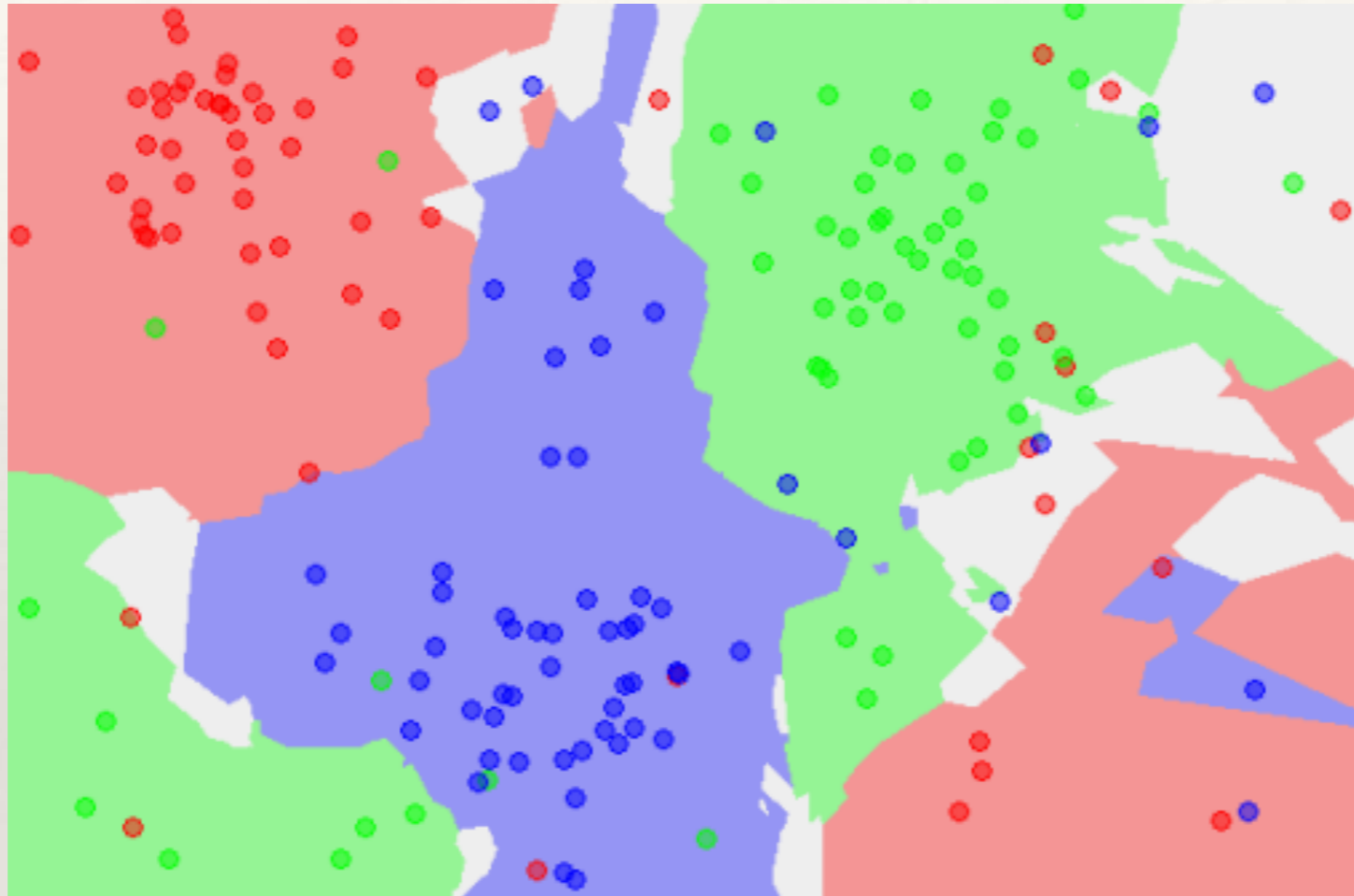
Sieć neuronowa jest (prostym) modelem obliczeniowym



Neuron „odpala” przy odpowiednim pobudzeniu

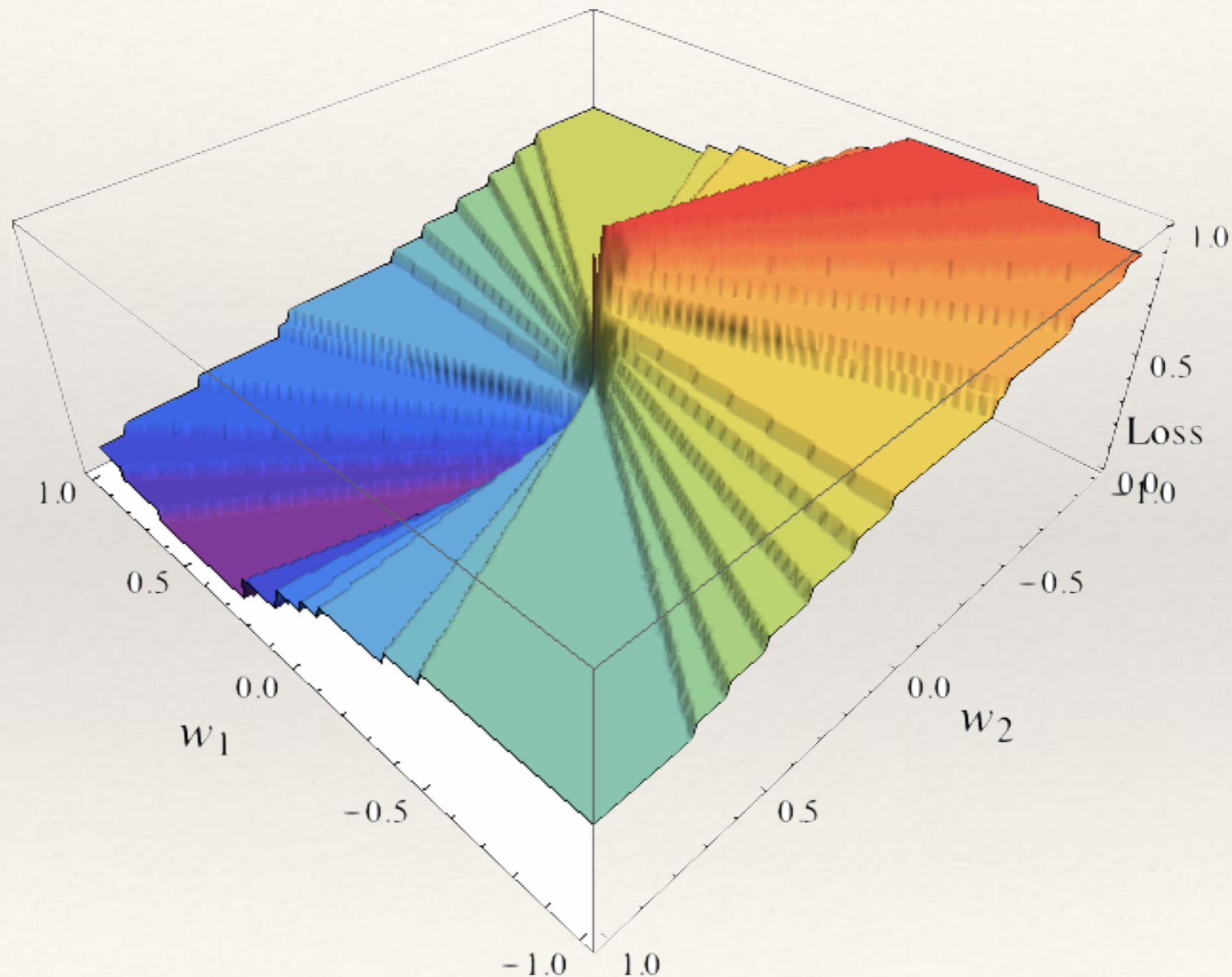


Uczenie sieci to w istocie zadanie minimalizacji

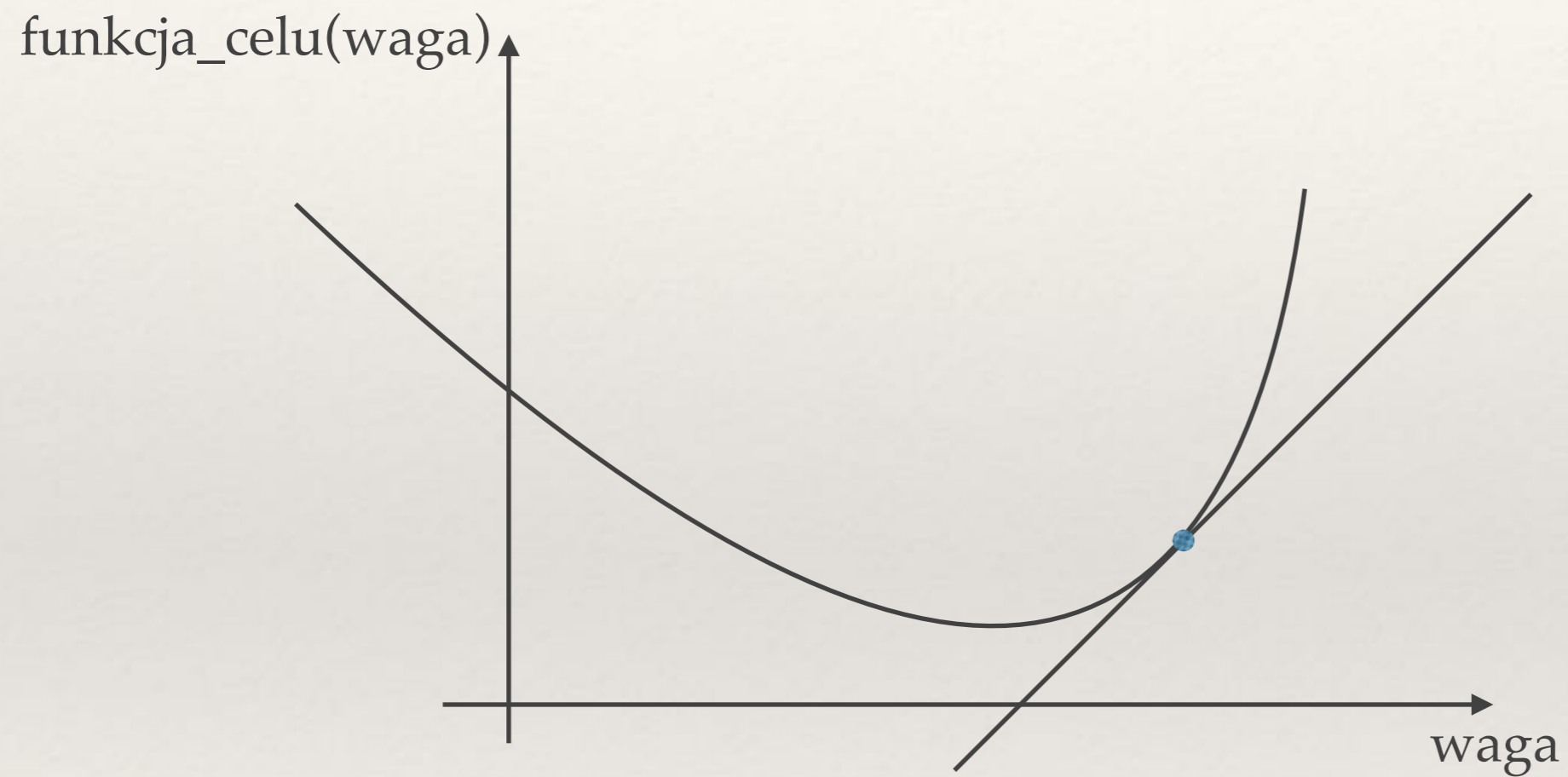


Tak próbujemy dostroić sieć by liczba błędów była jak najmniejsza
(wartość funkcji celu jak najmniejsza)

Wartość funkcji celu zależy od wag



Gradient podpowiada jak zmieniać wagi



Zmieniamy wagi by sieć działała coraz lepiej

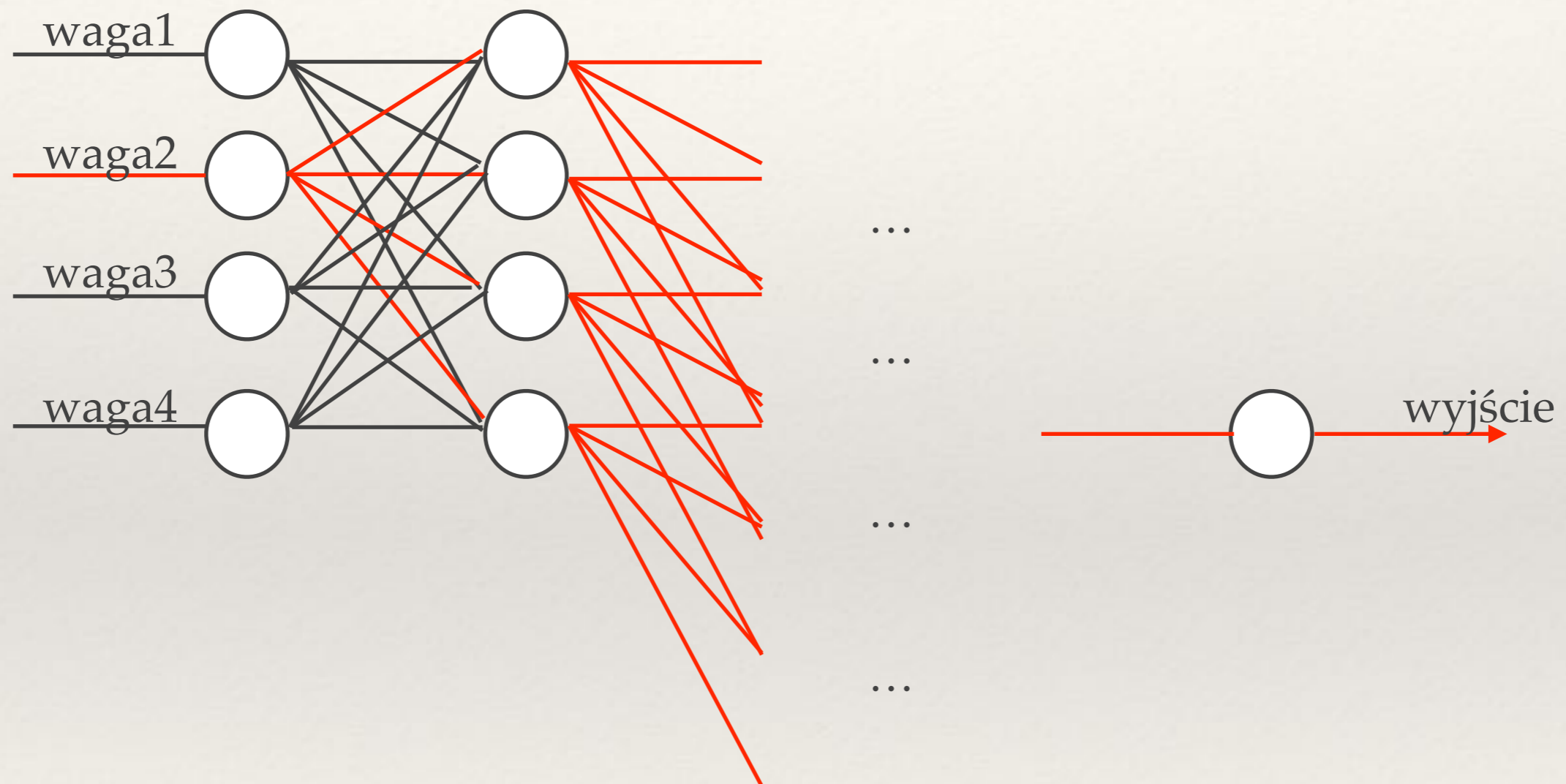
$$w := w - \eta \nabla$$

gdzie w to wektor wag,

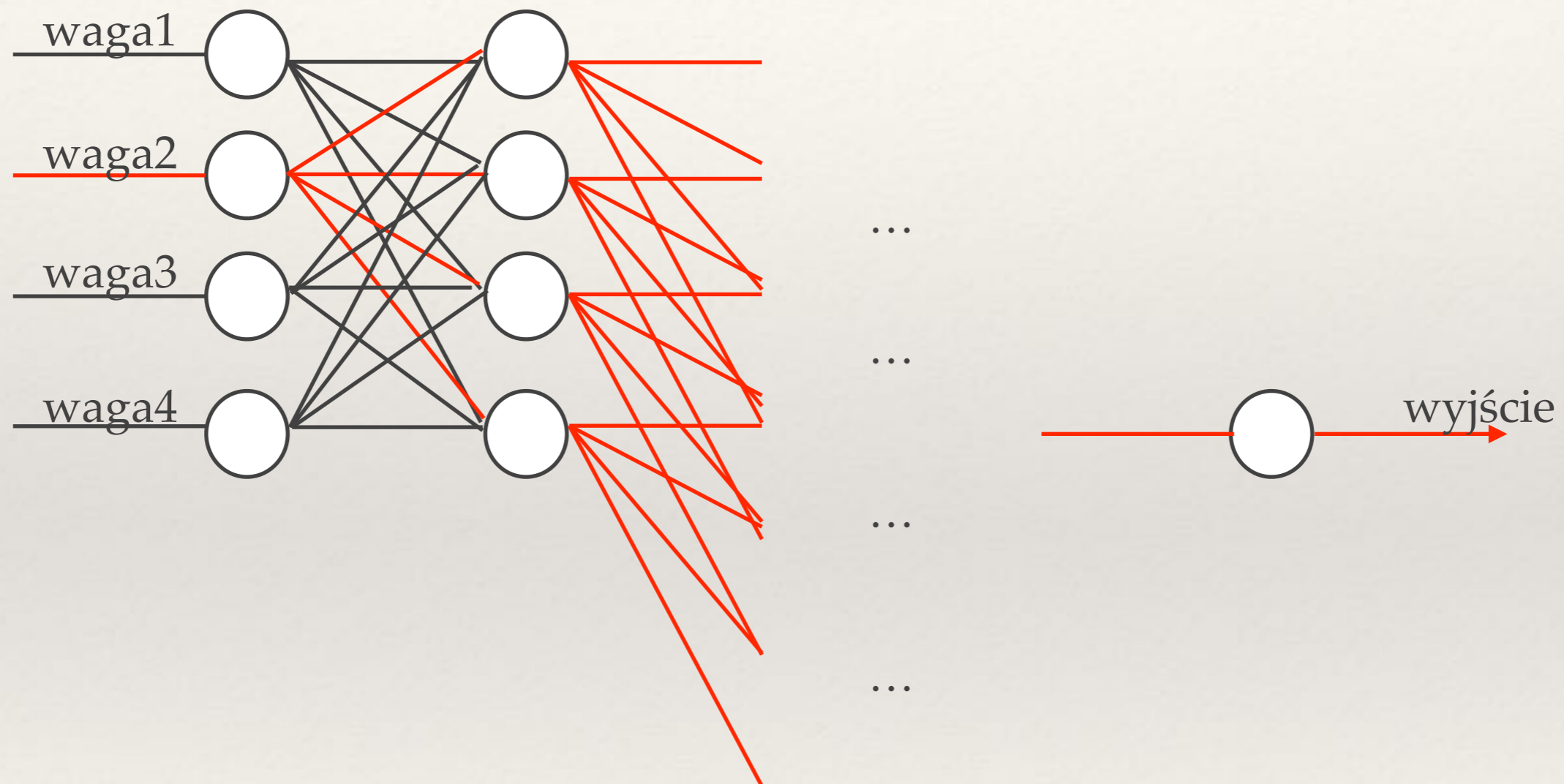
η to współczynnik uczenia,

∇ to gradient (wektor pochodnych cząstkowych)

Naiwne liczenie gradientu działa za wolno

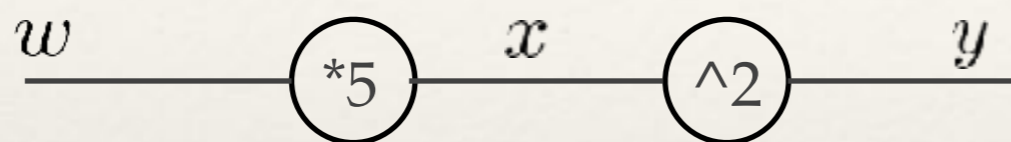


Naiwne liczenie gradientu działa za wolno



Wydajne rozwiązanie: algorytm wstecznej propagacji!

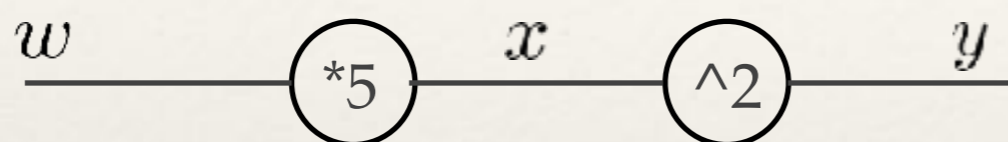
Algorytm wstecznej propagacji bazuje na regule łańcuchowej



$$x = 5 * w$$

$$y = x^2$$

Algorytm wstecznej propagacji bazuje na regule łańcuchowej



$$x = 5 * w$$

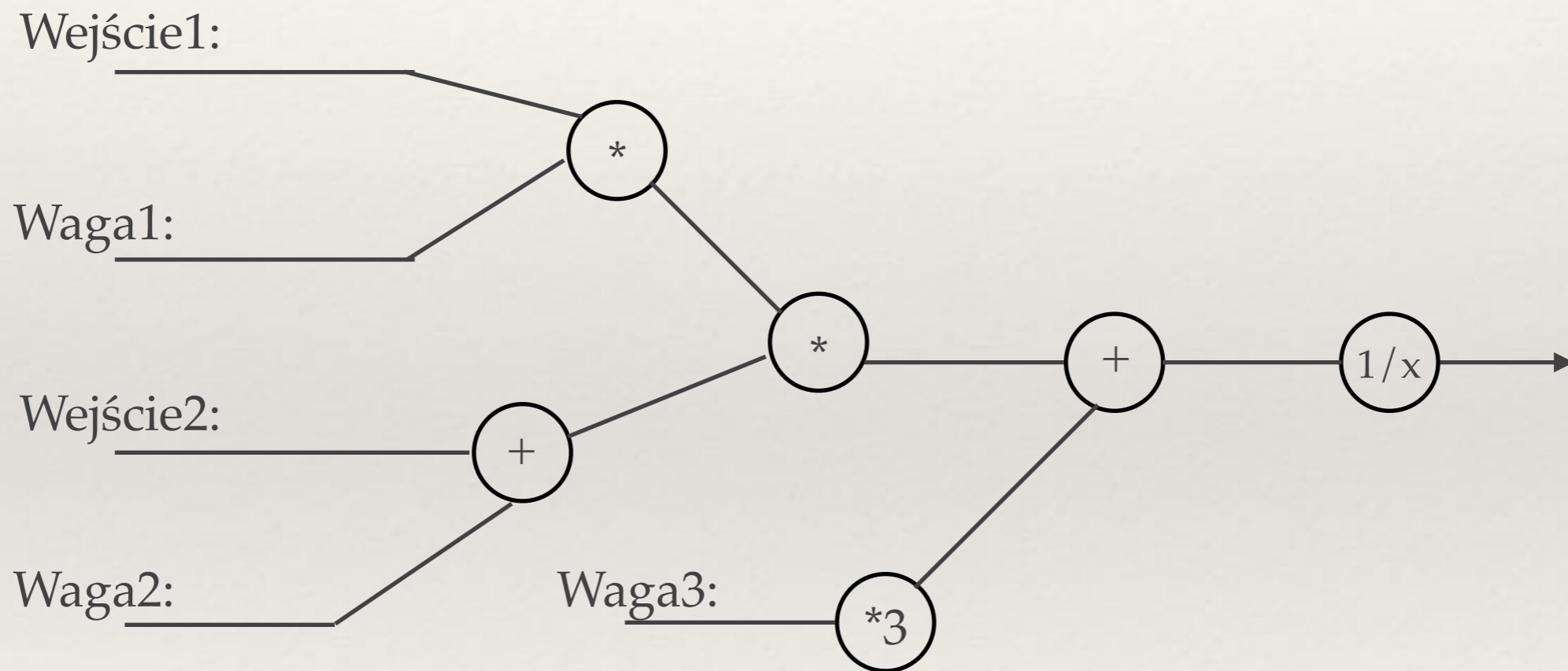
$$y = x^2$$

$$\frac{dx}{dw} = 5$$

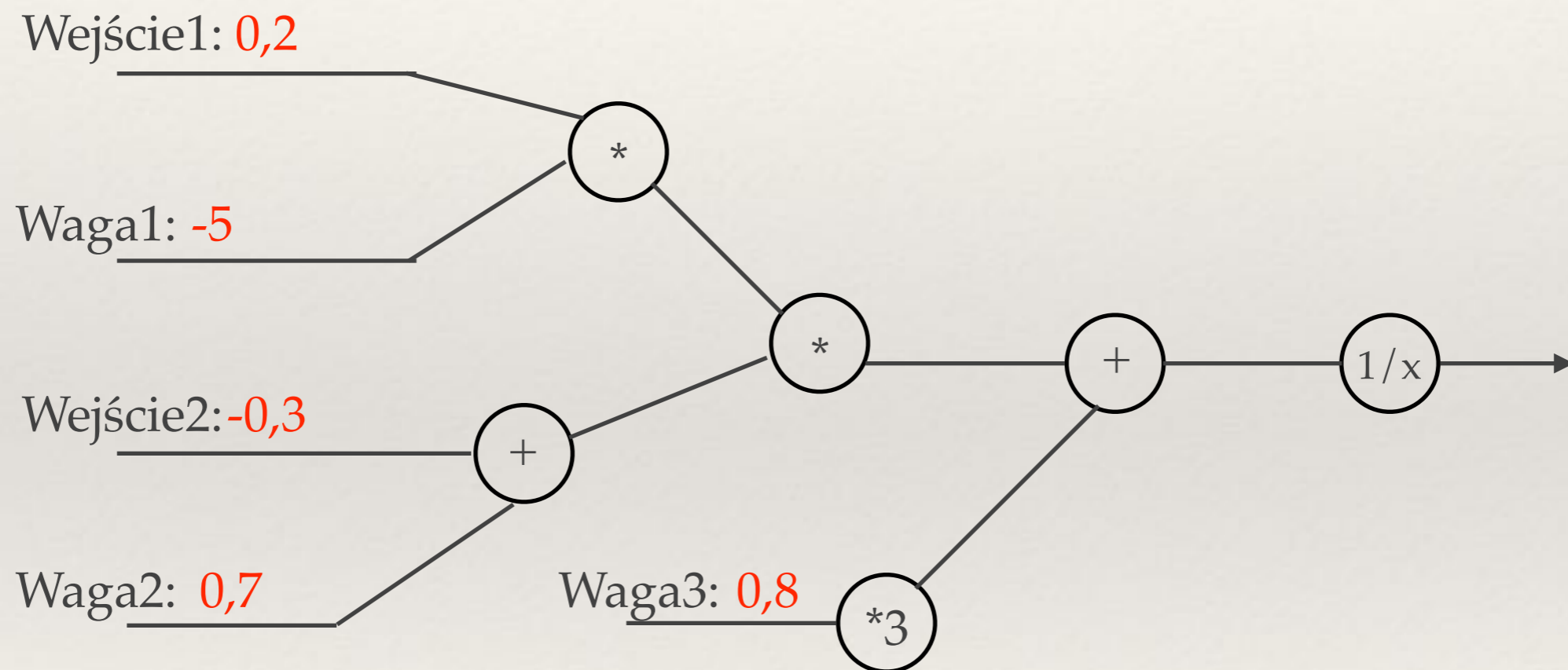
$$\frac{dy}{dw} = \frac{dy}{dx} \frac{dx}{dw} = 2x * 5 = 50w$$

$$\frac{dy}{dx} = 2x$$

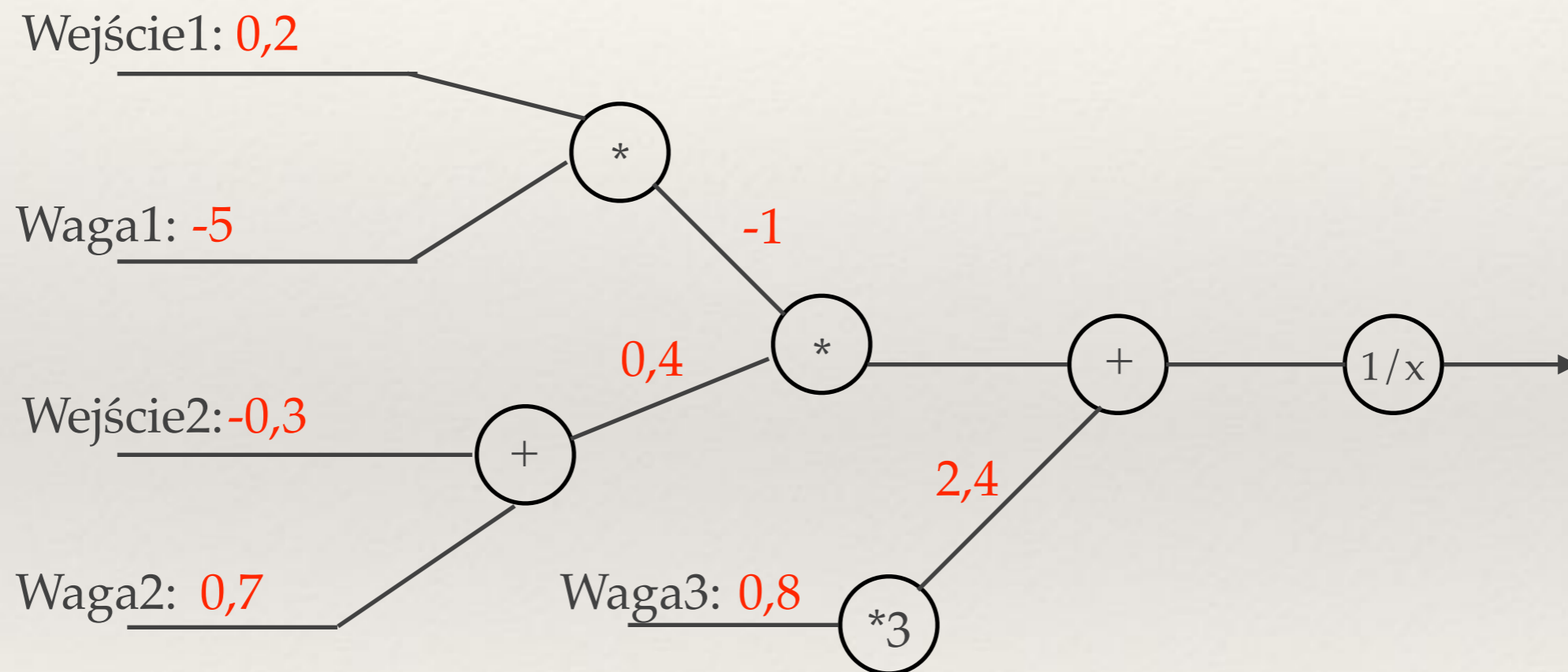
Algorytm wstecznej propagacji jest wydajny



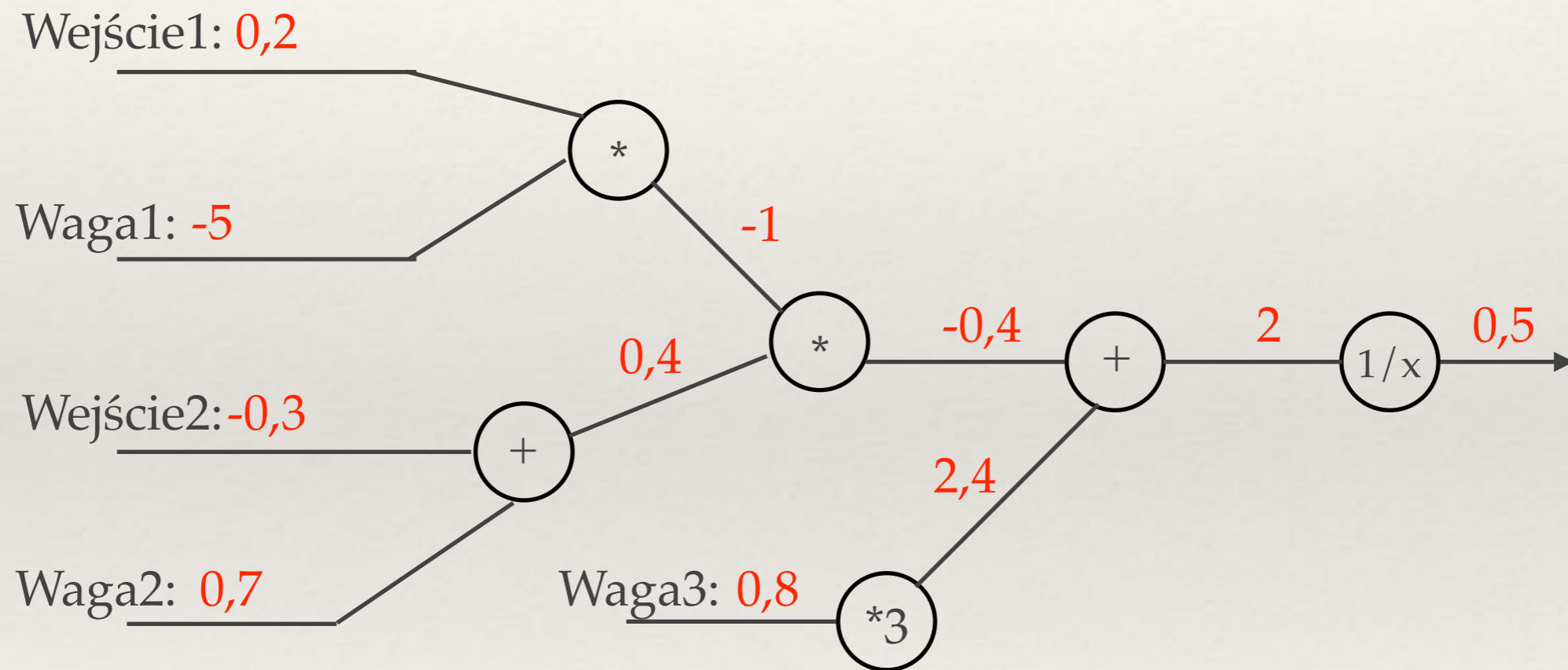
Algorytm wstecznej propagacji jest wydajny



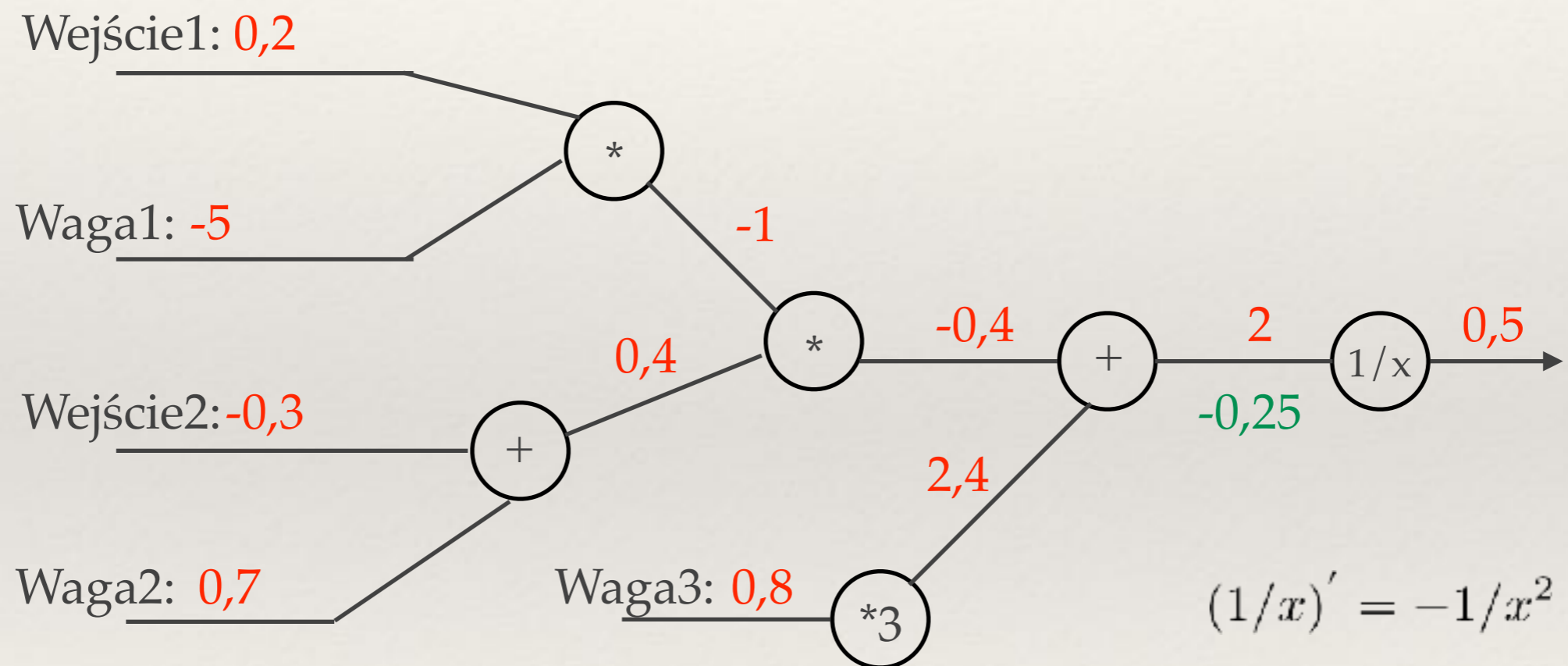
Algorytm wstecznej propagacji jest wydajny



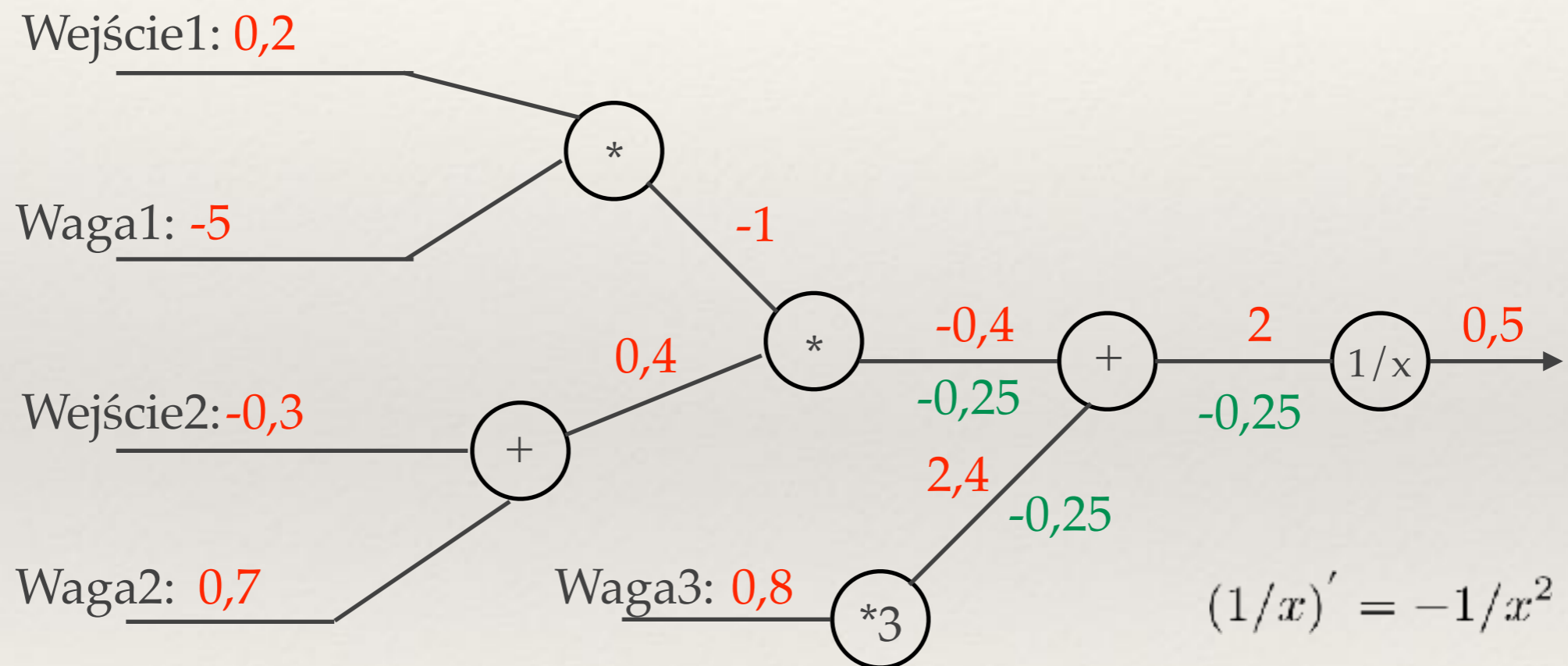
Algorytm wstecznej propagacji jest wydajny



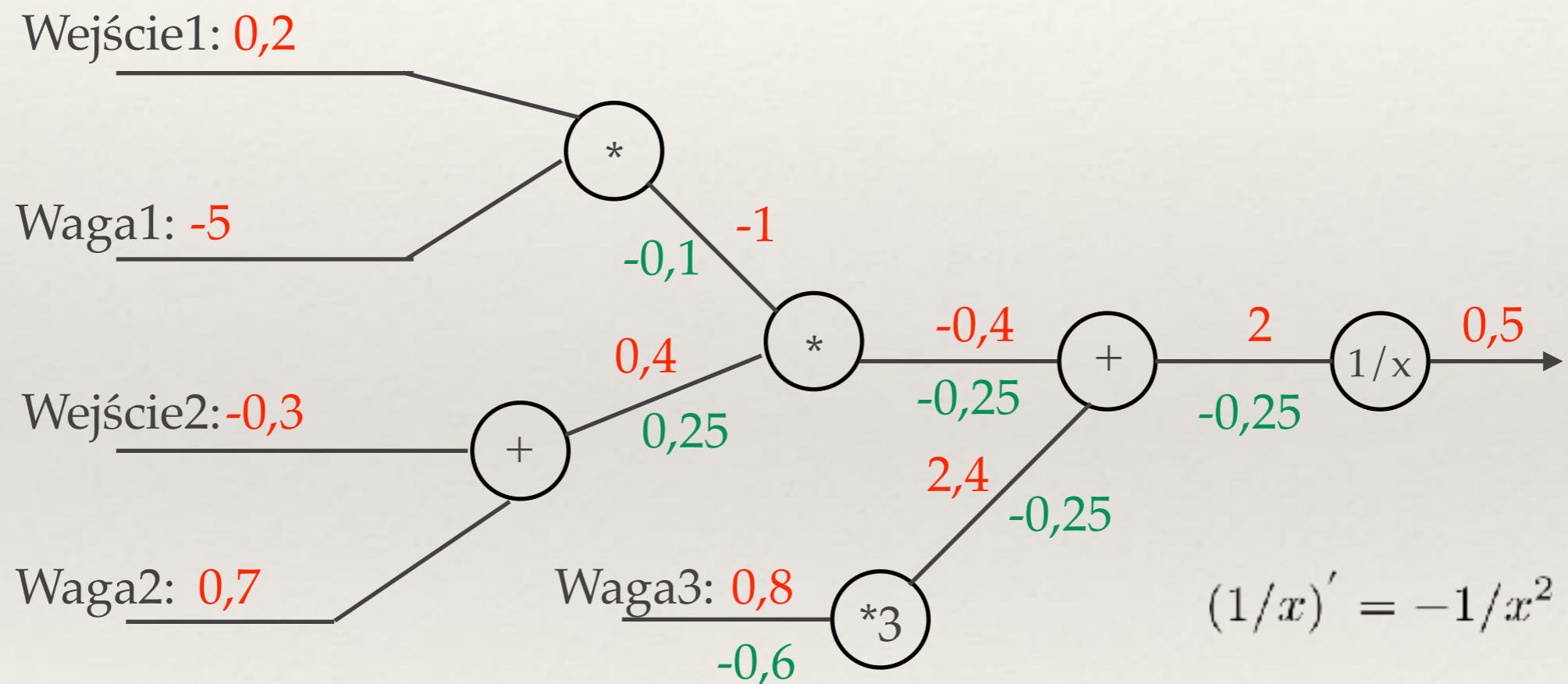
Algorytm wstecznej propagacji jest wydajny



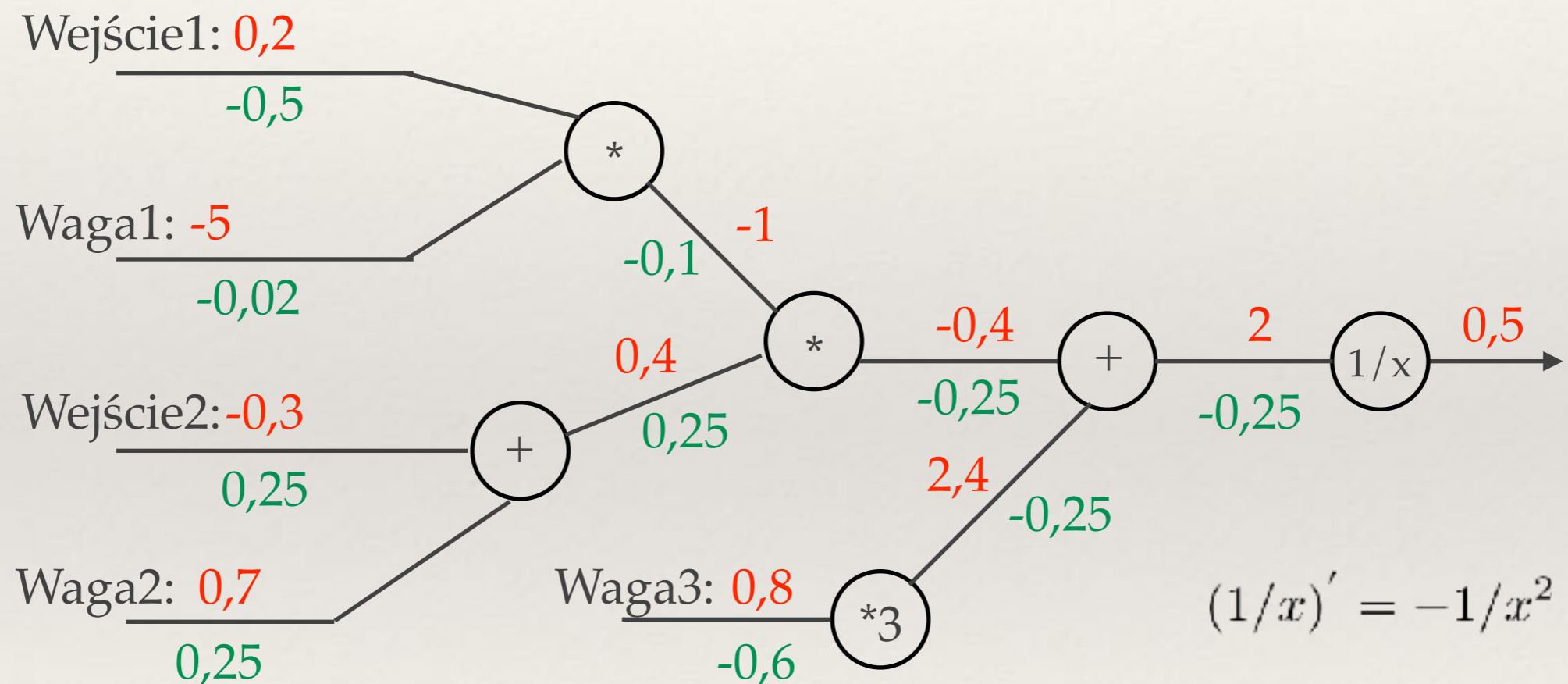
Algorytm wstecznej propagacji jest wydajny



Algorytm wstecznej propagacji jest wydajny



Algorytm wstecznej propagacji jest wydajny



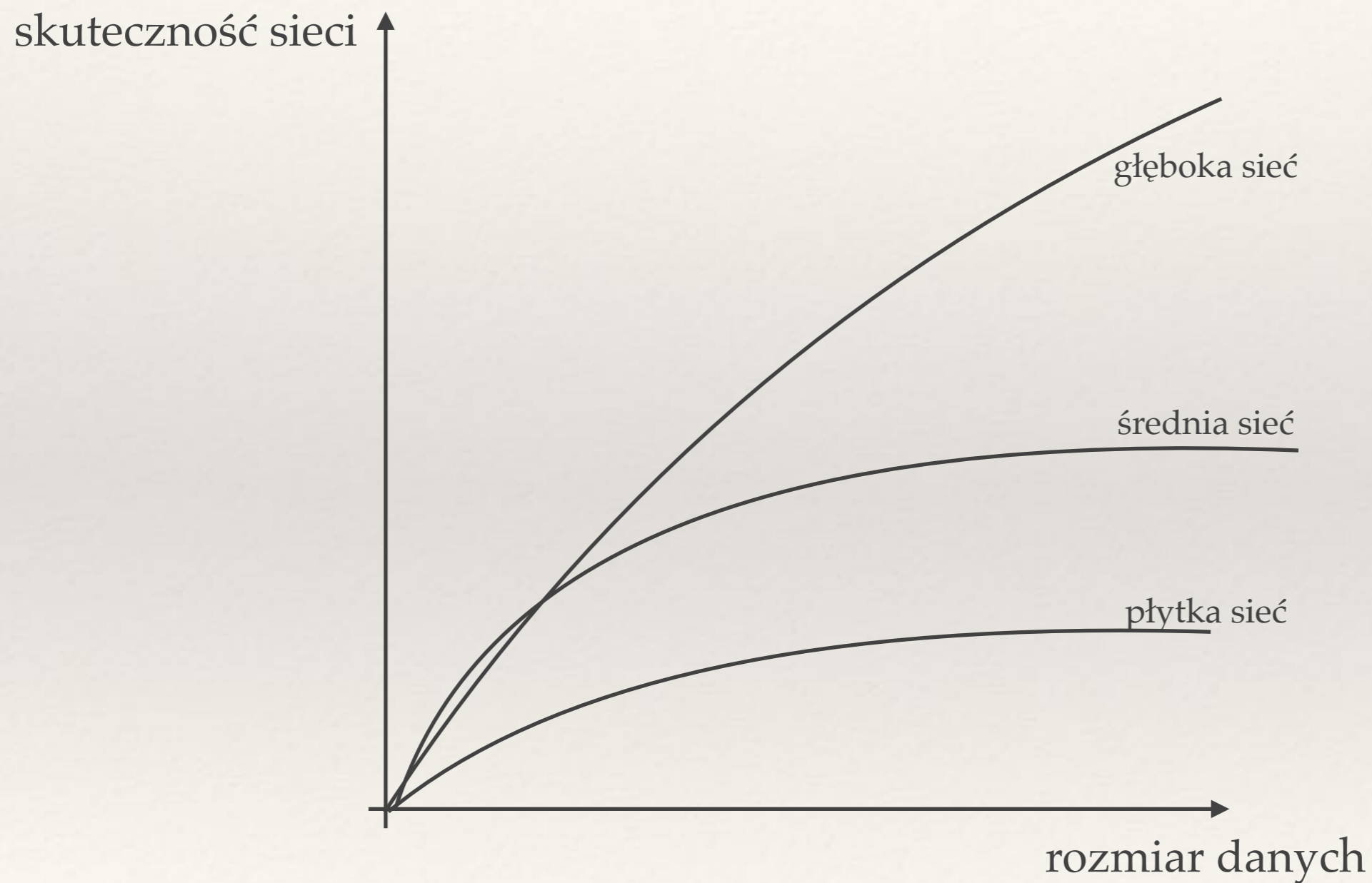
Uczenie się sieci to zmiany wag w dobrym kierunku

dopóki czas* pozwoli {

dla każdego** przykładu ze zbioru treningowego {oblicz gradient dla wag}
policz średni gradient i zaktualizuj wagi

}

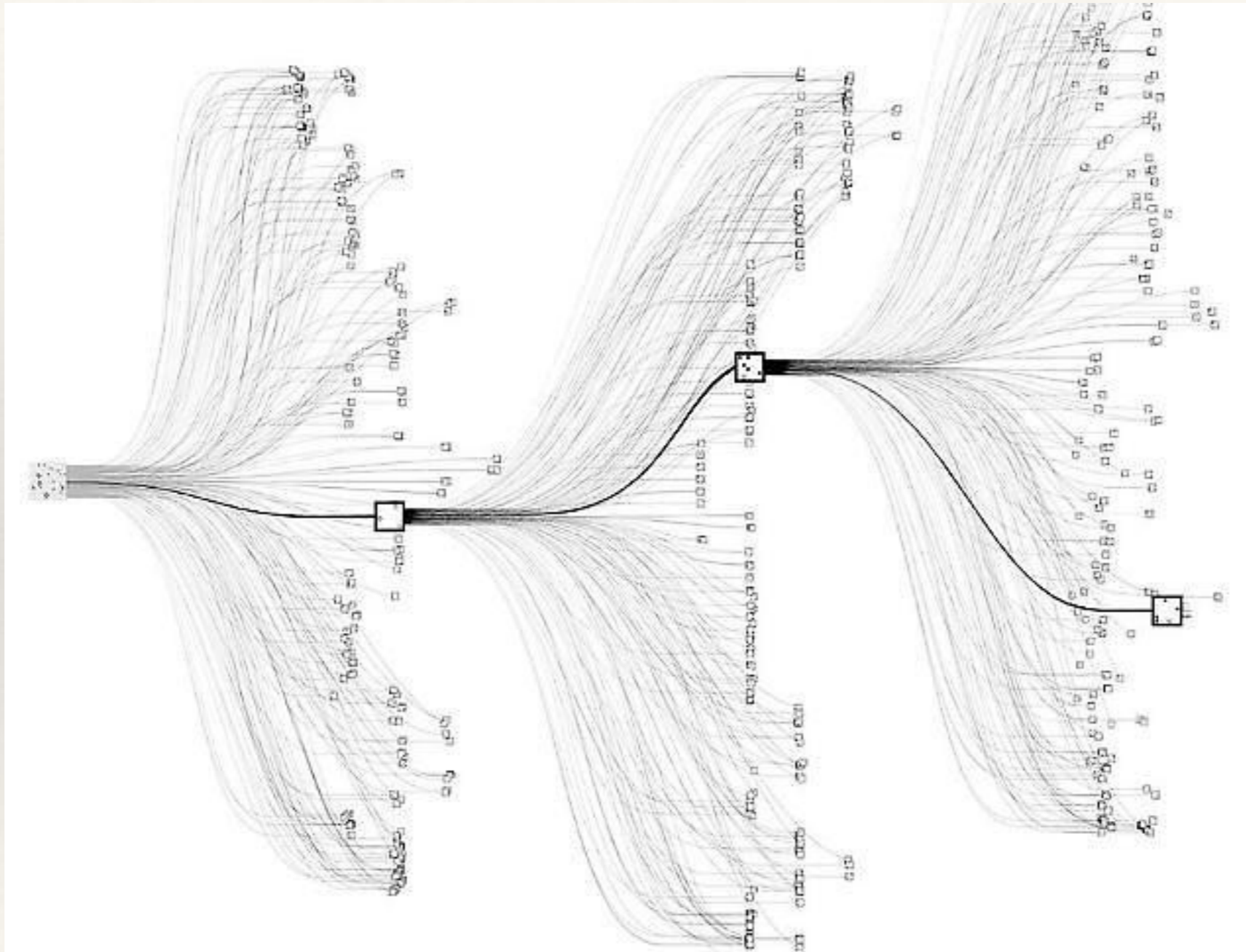
Skuteczność głębokich sieci rośnie wraz z rozmiarem danych



Program AlphaGo pokonał zawodowego mistrza gry Go



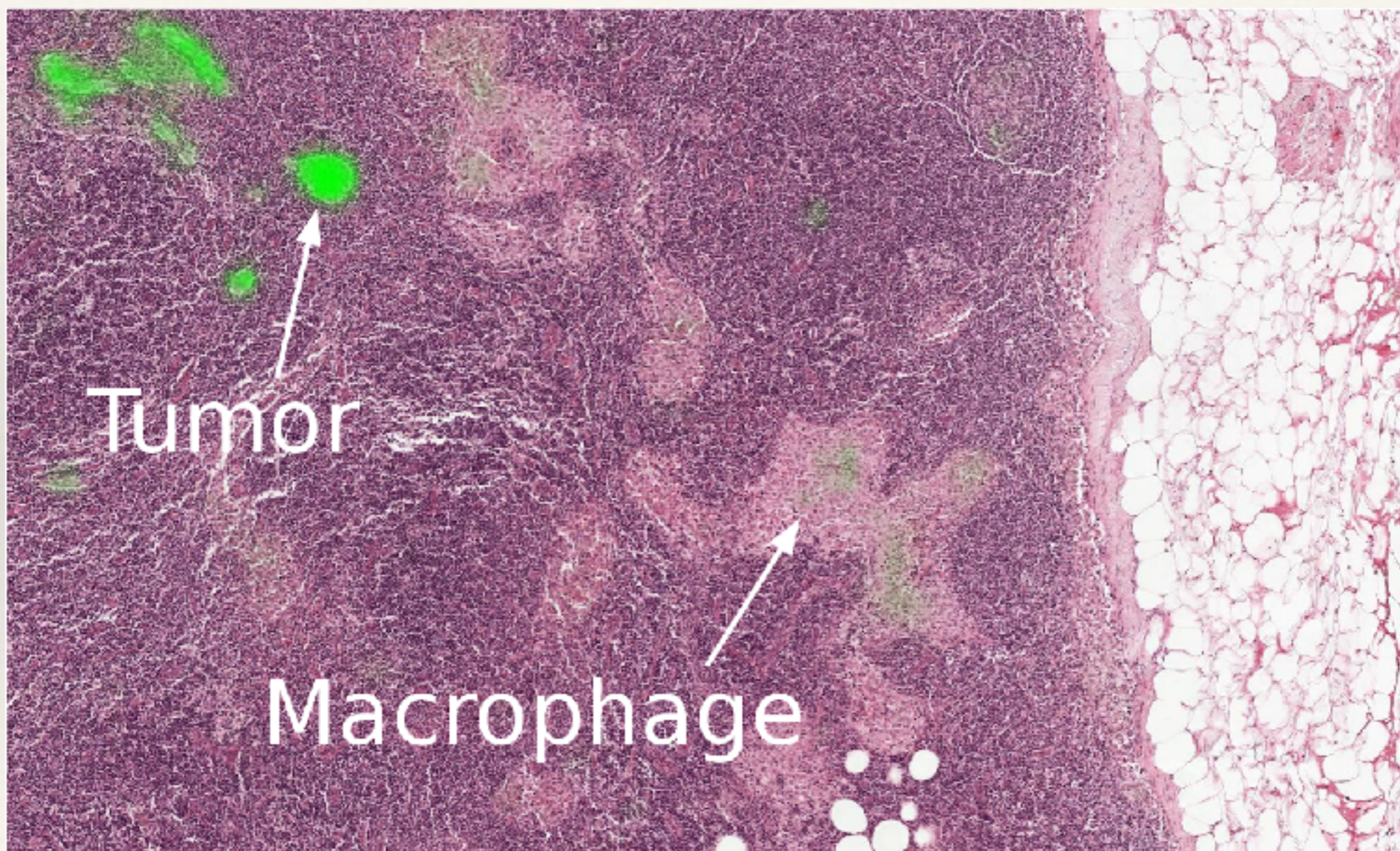
Głębokie sieci neuronowe przycięły drzewo gry



Źródło: Google Deepmind presentation on AlphaGo

Głębokie sieci potrafią diagnozować zmiany nowotworowe na poziomie doświadczonego eksperta (lub lepiej)

Wycinek tkanki z węzła chłonnego



Źródło: Detecting Cancer Metastases on Gigapixel Pathology Images, 2016
<https://arxiv.org/pdf/1703.02442.pdf>

Głębokie sieci neuronowe znakomicie służą NLP (np. pomagają wykrywać koreferencje)

„ - Znam planetę na której mieszka **pan** o czerwonej twarzy. **On** nigdy nie wachał kwiatów. Nigdy nie patrzył na gwiazdy. Nigdy nikogo nie kochał. Niczego w życiu nie robił poza rachunkami. I cały dzień powtarza tak jak ty: "Jestem **człowiekiem** poważnym, jestem **człowiekiem** poważnym. Nadyma się dumą.”

Vincent van Gogh maluje prof. Koronackiego



+



Można się samemu uczyć

- ❖ course.fast.ai (Practical Deep Learning for Coders)
- ❖ neuralnetworksanddeeplearning.com
- ❖ cs231n.github.io (Convolutional Neural Networks for Visual Recognition)

Coraz lepsze i przystępniejsze narzędzia ułatwiają start

- ❖ Tensorflow (www.tensorflow.org) i Keras (keras.io)
Biblioteka i API wspierane przez Google
- ❖ PyTorch (pytorch.org)
Używany i wspierany przez Facebook i Twitter

Amazon oferuje serwery z GPU

- ❖ Amazon Web Services (aws.amazon.com)
- ❖ dedykowane serwery z preinstalowanymi bibliotekami
- ❖ koszt około 1\$ za godzinę pracy na serwerze