

7

Dość proste sieci neuronowe

Kiedy słyszymy o postępach w dziedzinie sztucznej inteligencji, zazwyczaj dotyczą one konkretnej specjalizacji nazywanej *uczeniem maszynowym* (zdobywaniem przez komputery wiedzy, która nie została im bezpośrednio przekazana). Te postępy wynikają zwykle z zastosowania techniki uczenia maszynowego nazywanej *sieciami neuronowymi*. Choć sieci neuronowe zostały odkryte dziesiątki lat temu, przechodzą swego rodzaju renesans, ponieważ lepszy sprzęt i odkryte w wyniku badań nowe techniki programistyczne otworzyły drogę dla nowego modelu o nazwie *Deep learning* (*głębokie uczenie maszynowe*).

Techniki Deep learning znalazły mnóstwo zastosowań. Przydają się w wielu różnych dziedzinach, począwszy od handlu algorytmicznego po bioinformatykę. Dwa zastosowania modelu Deep learning, które są znane konsumentom, to rozpoznawanie obrazów oraz rozpoznawanie mowy. Jeśli kiedykolwiek spytałeś wirtualnego asystenta, jaka jest pogoda, lub użyłeś programu rozpoznającego twarze na zdjęciach, najprawdopodobniej w grę wchodziły jakieś techniki Deep learning.

Techniki Deep learning wykorzystują te same komponenty co prostsze sieci neuronowe. W tym rozdziale przeanalizujemy te komponenty, budując prostą sieć neuronową. Nie będzie to arcydzieło, ale da podstawę do zrozumienia modelu Deep learning (bazującego na bardziej złożonych sieciach neuronowych niż ta, którą zbudujemy). Większość praktyków uczenia maszynowego zamiast budować sieci neuronowe od podstaw, korzysta z popularnych, wysoce zoptymalizowanych, gotowych do użycia bibliotek, które realizują podstawowe zadania. Choć w tym rozdziale nie tłumaczę, jak korzystać z określonej biblioteki, a budowana przez nas sieć nie będzie przydatna w praktyce, pomoże w zrozumieniu podstaw działania tych bibliotek.

7.1. Podłoże biologiczne?

Ludzki mózg jest najbardziej fascynującym urządzeniem obliczeniowym. Nie potrafi przetwarzać liczb tak szybko jak mikroprocesor, ale potrafi przystosowywać się do nowych sytuacji, zdobywać nowe umiejętności i kreatywnością przewyższa wszystkie znane maszyny. Od czasu pojawienia się komputerów naukowcy badali możliwość zamodelowania maszynierii mózgu. Komórki nerwowe w mózgu są nazywane *neuronami*. Neurony w mózgu są za pomocą *synaps* połączone w sieć. Sygnały elektryczne przechodzą przez synapsy i zasilają te sieci neuronów, zwane również *sieciami neuronowymi*.

UWAGA Przedstawiony opis biologicznych neuronów jest ogromnym uproszczeniem służącym do celów porównawczych. W rzeczywistości biologiczne neurony składają się z takich części jak aksony, dendryty i jądro neuronu, jak być może pamiętasz z lekcji biologii. Natomiast synapsy są tak naprawdę lukami między neuronami, a sygnał elektryczny może być przekazywany dzięki wydzielanym neuroprzekaźnikom.

Chociaż naukowcy zidentyfikowali części i funkcje neuronów, nadal nie do końca rozumiemy, w jaki sposób biologiczne sieci neuronowe formułują złożone myśli. W jaki sposób przetwarzają informacje? W jaki sposób rodzą się nowe idee? Większość naszej wiedzy o sposobie działania mózgu wynika z obserwowania go w skali makro. Skany mózgu, otrzymywane przez funkcjonalne obrazowanie metodą rezonansu magnetycznego (ang. *functional magnetic resonance imaging*, fMRI) pokazują, gdzie płynie krew, gdy badana osoba wykonuje określone czynności lub myśli o czymś konkretnym (jak na rys. 7.1). Na podstawie tej i innych technik makro wnioskujemy, w jaki sposób są ze sobą połączone różne części mózgu. Jednak nie zdradzają one, w jaki sposób pojedyncze neurony pomagają w rozwijaniu nowych myśli.



Rysunek 7.1. Naukowiec analizuje obrazy fMRI mózgu. Obrazy fMRI niewiele mówią o działaniu pojedynczych neuronów lub organizacji sieci neuronowych