

Krótki wstęp do uczenia maszynowego

“Machine learning allows us to tackle tasks that are too difficult to solve with fixed programs written and designed by human beings.”

Goodfellow, I. et.al. (2016). *Deep Learning*. MIT Press

A computer program is said to learn from **experience E** with respect to some class of **tasks T** and performance **measure P**, if its performance at tasks in T, as measured by P, improves with experience E.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York.

The Task T

- dane składają się z **przykładów** (instancji): $\vec{x} \in \mathbb{R}^n$
- poszczególne wartości wektora $\vec{x} \in \mathbb{R}^n$ nazywamy **cechami**.

Rodzaje zadań:

- Klasyfikacja
- Regresja
- Transkrypcja
- Tłumaczenie
- Synteza i sampling
- ...

Task T: Klasyfikacja

Zadaniem programu jest “nauczyć” się funkcji:

$$\begin{array}{ccc} f: \mathbb{R}^n \rightarrow \{1, \dots, k\} \simeq \mathbb{Z}_k & \text{lub} & f: \mathbb{R}^n \rightarrow [0, 1] \times [0, 1] \times \dots \times [0, 1] \subset \mathbb{R}^k \\ \downarrow & & \downarrow \\ \text{klasy} & & \text{prawdopodobieństwo przynależności do klasy} \end{array}$$

Przykłady:

- klasyfikacja obrazów
- rozpoznawanie twarzy
- detekcja anomalii
- ...

Task T: Regresja

Zadaniem programu jest “nauczyć” się funkcji:

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

Przykłady:

- przewidywanie sprzedaży, cen
- oczekiwana liczba zgłoszeń w firmie ubezpieczeniowej
- ...

Task T: Transkrypcja

Zadaniem programu jest “nauczyć” się funkcji:

$$f: \mathbb{R}^n \rightarrow \mathbb{Z}_k^d$$

Przykłady:

- OCR (Optical Character Recognition): rozpoznawanie tekstu na obrazku
- Speech-to-text: transkrypcja mowy na tekst
- ...

Task T: Tłumaczenie maszynowe

Zadaniem programu jest “nauczyć” się funkcji:

$$f : \mathbb{Z}_k^d \rightarrow \mathbb{Z}_k^d$$

Przykłady:

- OCR (Optical Character Recognition): rozpoznawanie tekstu na obrazku
- Speech-to-text: transkrypcja mowy na tekst
- ...

Task T: Synteza i sampling

Zadaniem programu jest “nauczyć” się funkcji:

$$f : \mathbb{Z}_k^d \rightarrow \mathbb{R}^n$$

Przykłady:

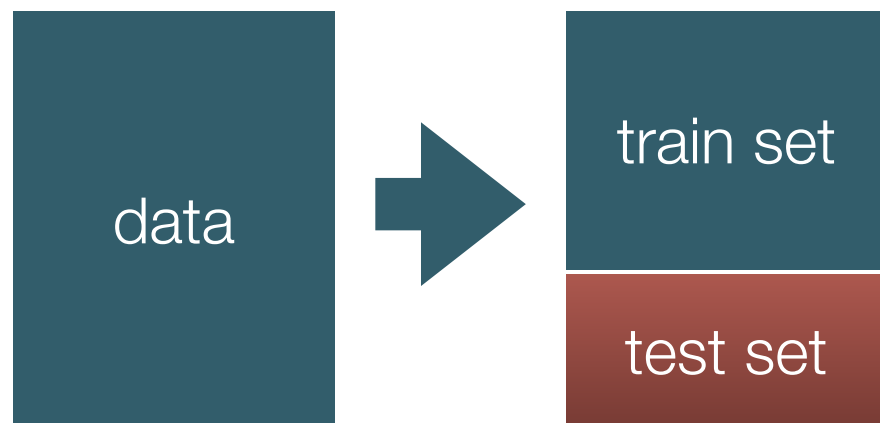
- Text-to-speech: synteza mowy
- generowanie danych dla innych zadań
- ...

A computer program is said to learn from **experience E** with respect to some class of **tasks T** and performance **measure P**, if its performance at tasks in T, as measured by P, improves with experience E.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York.

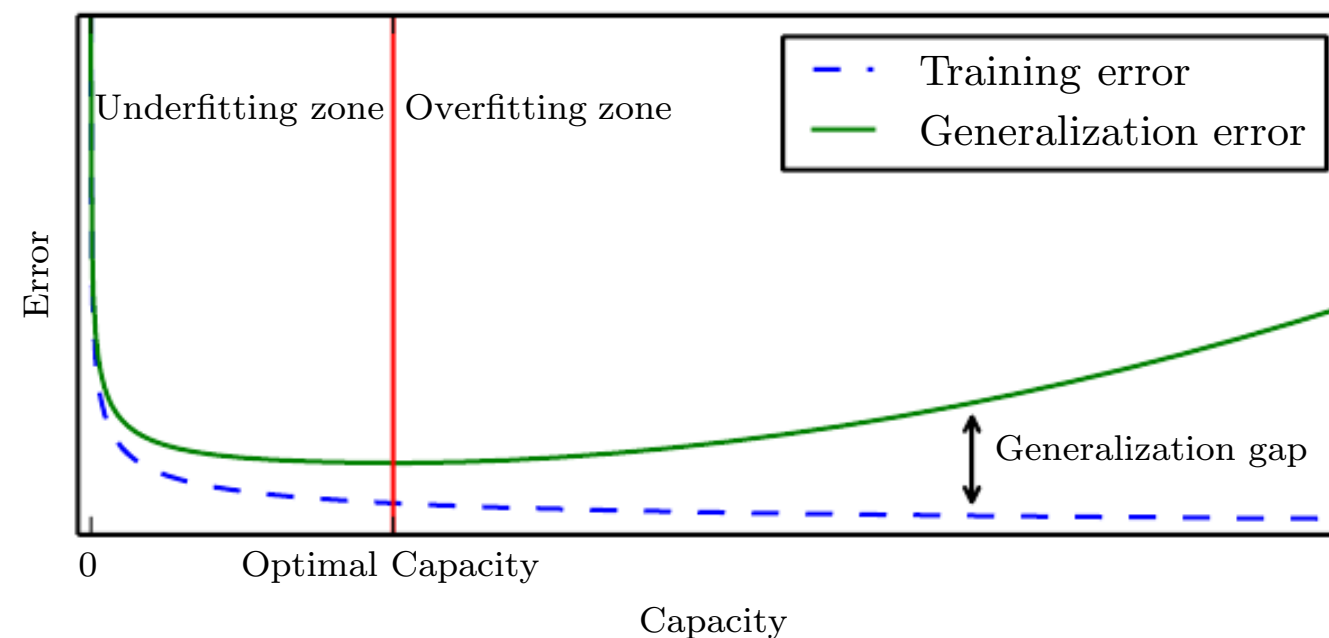
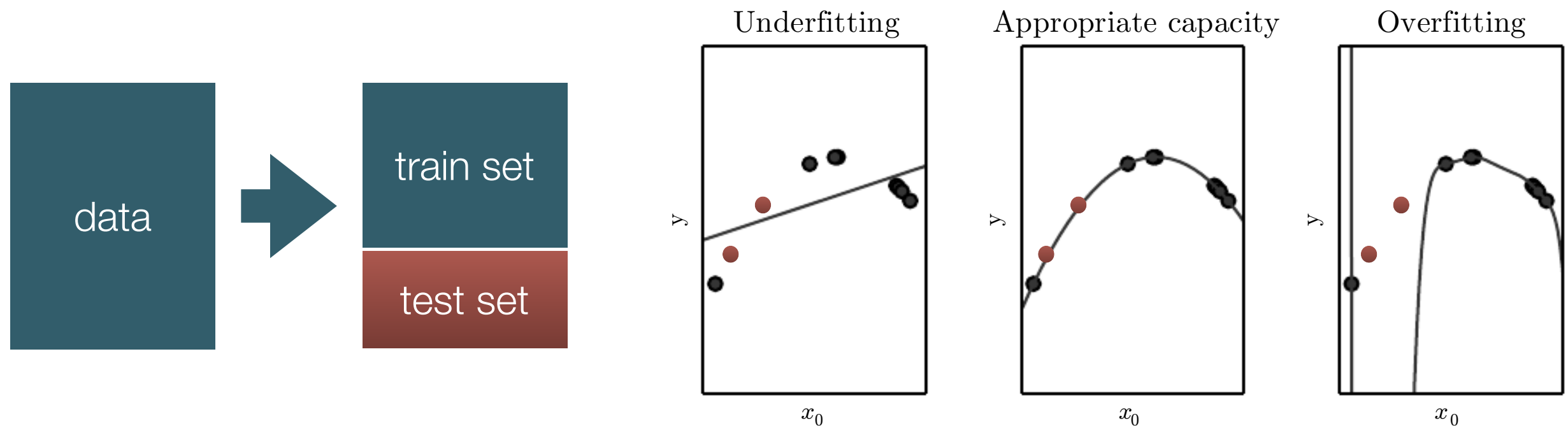
The Performance measure P

- musimy **zmierzyć** w jakim stopniu model poprawnie rozwiązuje zadanie
- miara zależy od rodzaju zadania, np.:
 - **accuracy/error rate** w przypadku klasyfikacji, transkrypcji, itp.
 - **MSE (mean square error)** w przypadku regresji
- nazywamy ją funkcją kosztu lub błędu (**loss function**)
- kluczowym jest ocenianie modelu na **zbiorze testowym**



The Performance measure P

- kluczowym jest ocenianie modelu na **zbiorze testowym**

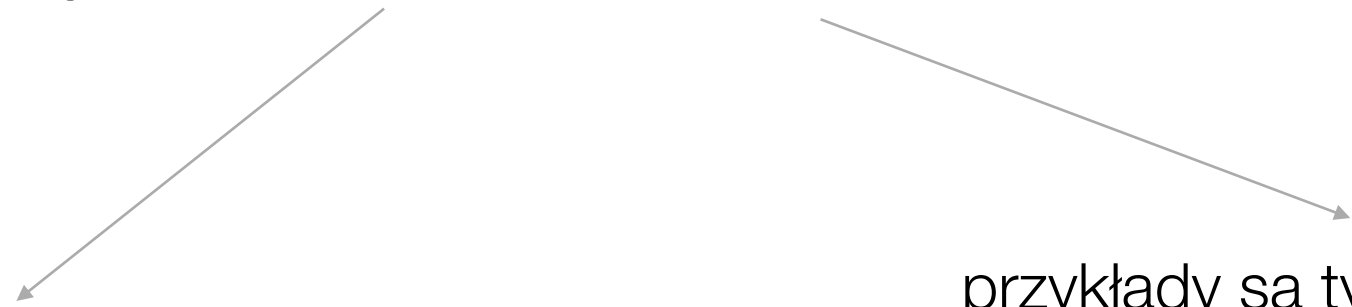


A computer program is said to learn from **experience E** with respect to some class of **tasks T** and performance **measure P**, if its performance at tasks in T, as measured by P, improves with experience E.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York.

The Experience E

- uczenie dzielimy na **z nadzorem** i **bez nadzoru**



każdy przykład ma określoną etykietkę (**label**, **target**)

$$p(y|\vec{x})$$

przykłady są tylko zbiorami cech;
algorytm “uczy się” rozkładu
prawdopodobieństwa, który wygenerował
zbiór

$$p(\vec{x})$$

$$p(\vec{x}) = \prod_{i=1}^n p(x_i | x_1, \dots, x_{i-1})$$



$$p(y|\vec{x}) = \frac{p(y, \vec{x})}{\sum_{y'} p(y', \vec{x})}$$

No Free Lunch Theorem

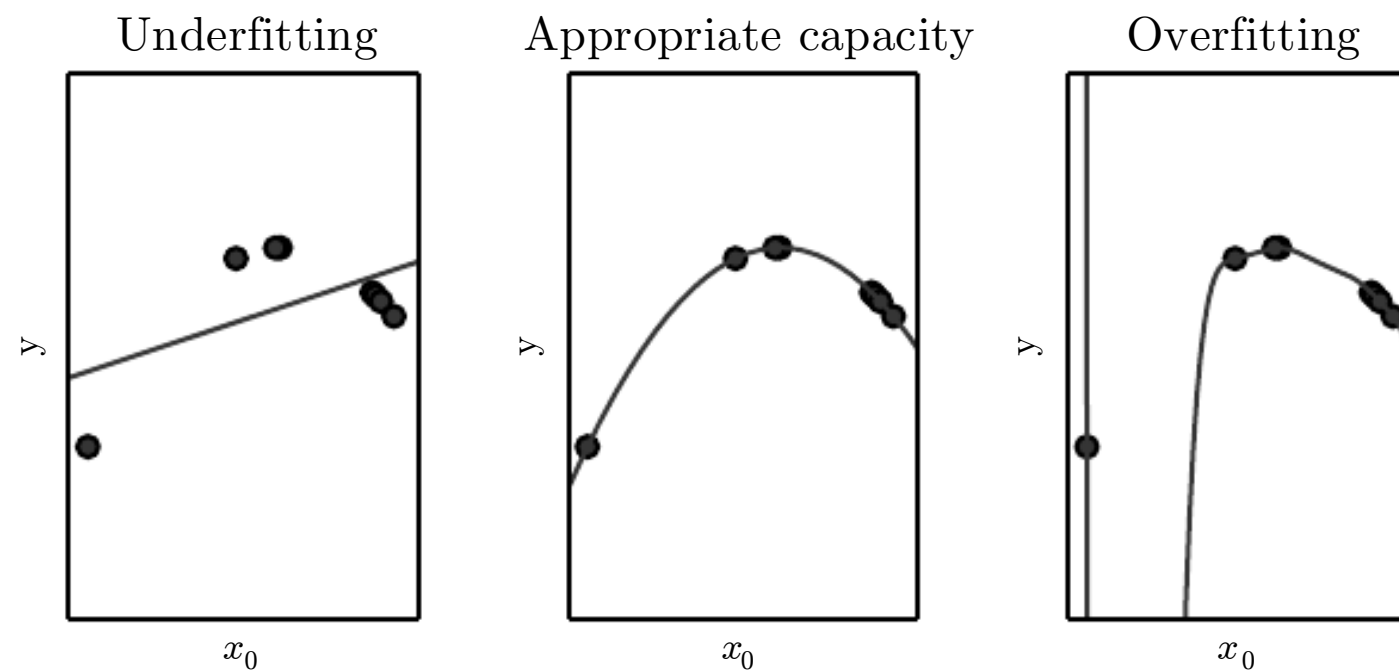
Every classification algorithm, when averaged over all possible data generating distributions, has the same error rate when classifying previously unobserved data.

Goodfellow, I. et.al. (2016). *Deep Learning*. MIT Press

- w szczególności, algorytm przypisujący zawsze tę samą klasę
- nie ma uniwersalnych algorytmów ML
- każdy algorytm ML musi być dostosowany do zadania, które rozwiązuje

Hyperparameters

czyli parametry określające “rodzinę” modeli



Na przykład:

- stopień wielomianu
- funkcja błędu
- szybkość uczenia się
- ...