# Inteligentne systemy przeciw atakom sieciowym

## Wykład 1

## Wprowadzenie

Joanna Kołodziejczyk

wrzesień/październik 2016

# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Wykrywanie ataków sieciowych — przeszłość a teraźniejszość

- Wykrywanie ręczne przez operatora sieci — początki Internetu, znaczna wiedza o anomaliach - ekspert.
- Automatyczne wykrywanie — potrzeba rozwijającej się szybkiej sieci.
  - prędkość sieci i szybkość znajdowania włamań i reakcji
  - dokładność
  - jak zautomatyzować -> sztuczna inteligencja, a w szczególności maszynowe uczenie się.

# Skąd wiadomo co nam zagraża?

Firmy od zabezpieczeń, skanerów tworzą roczne lub inne okresowe raporty. Ich treść to prawdziwy horror.

- Web Application Vulnerability Report Acunnetix 2016 — `https://d3eaqdewfg2crq.cloudfront.net/resources/acunetix-web-application-vulnerability-report-2016.pdf`
- Coroczny raport firmy Cisco

**Zadanie:**

Znaleźć inne raporty w tym raport Cisco i porównać wzrost podatności aplikacji sieciowych na ataki. Jaka jest zgodność? Co jest najczęstszą słabością? Ciekawostki?

Osoby piszące krótki raport zawierający spis analizowanych raportów dostaną punkty, które będą prowadzić do zwolnienia z obowiązku pisania testu zaliczeniowego.

# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Wykrywanie włamań

Źródła udostępniane w sieci są narażone na:

- zniszczenie
- ujawnienie
- modyfikację
- wyłączenie/uszkodzenie
- kradzież
- nieautoryzowany dostęp
- niewłaściwe wykorzystanie

**Wykrywanie włamań — intrusion detection**

jest to działanie polegające na wykryciu akcji, która zagraża poufności, integralności czy dostępności chronionych źródeł.

# Inteligencja—motywacja

## Słabości

Standardowe IDS wykrywają tylko rozpoznane i skatalogowane włamania. Nie potrafią rozpoznać nic nowego.

## Remedium

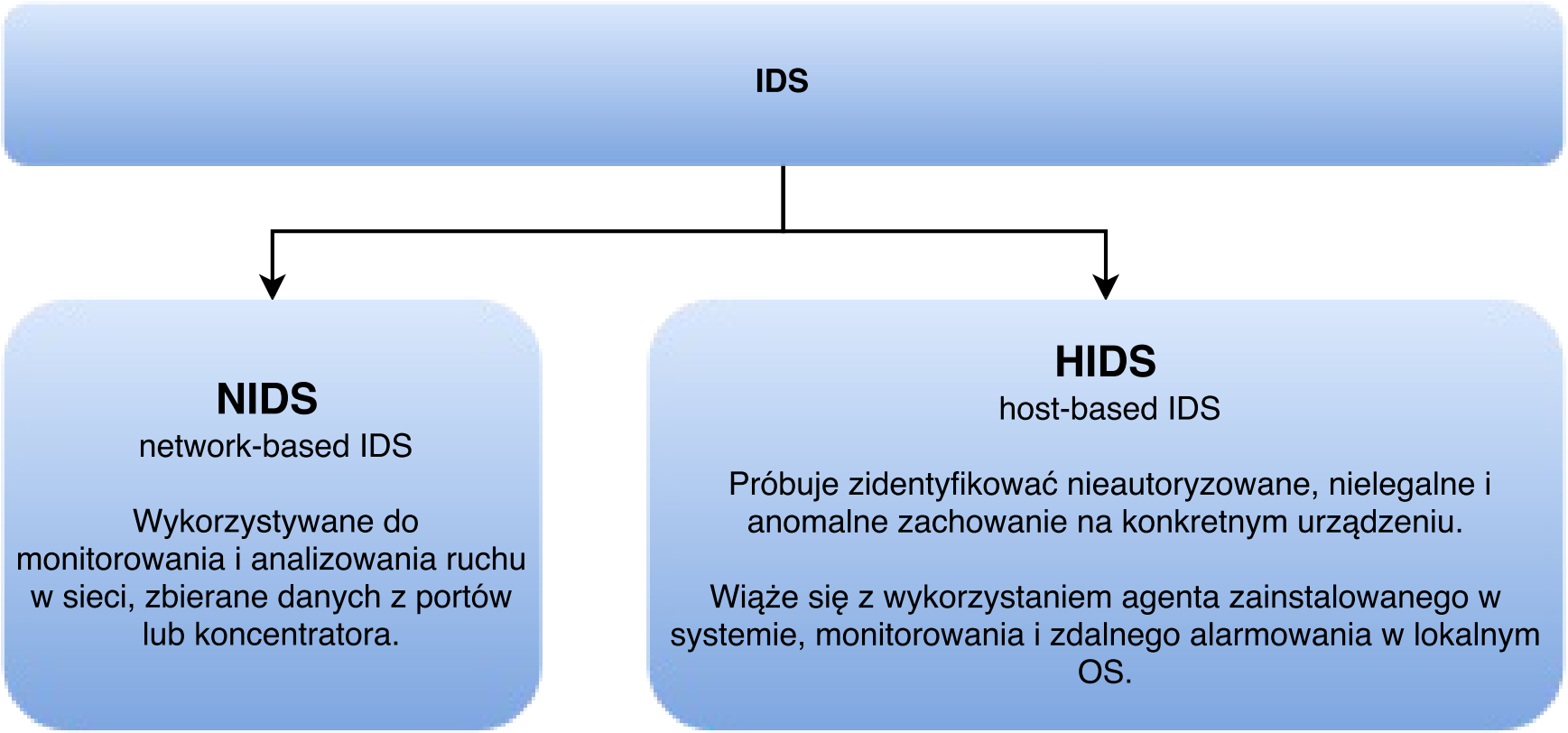
Metody ML (machine learning) potrafią się uczyć, często z przykładów, jak człowiek i dzięki temu są w stanie reagować na nowe sytuacje przez wytwozrzenie uogólnionego modelu włamań.

# IDS — Intrusion Detection System

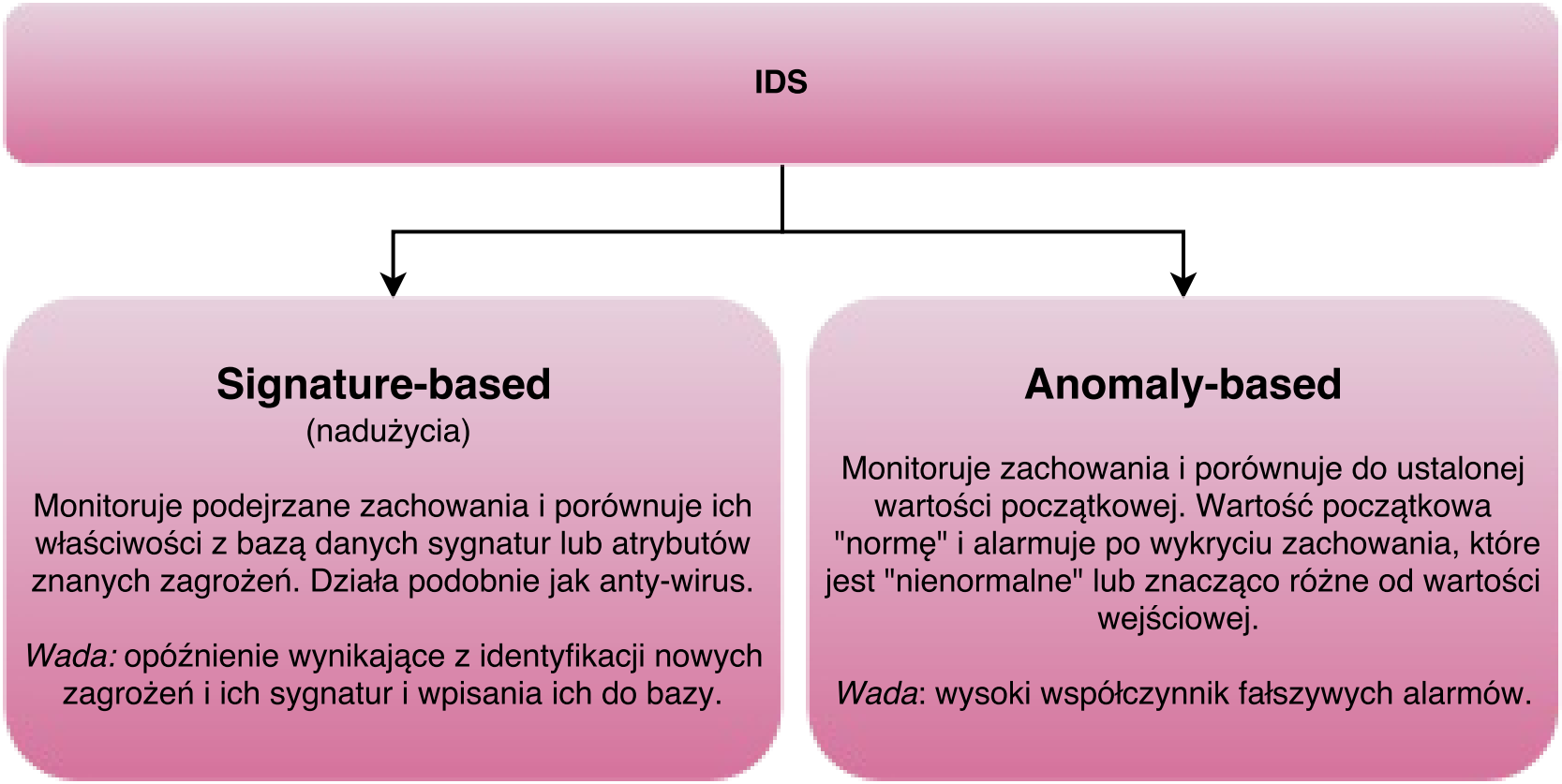
## Systemy wykrywania włamań

jest aplikacją lub urządzeniem, które monitoruje działania w sieci lub systemie, aby wykryć działania szkodliwe lub naruszenia i generuje raporty. IDS zbiera i analizuje informację z różnych źródeł w sieci by zidentyfikować prawdopodobne naruszenia zarówno intrusion (włamanie — pochodzące spoza organizacji), oraz misuse (nadużycia — pochodzące z wewnątrz organizacji).

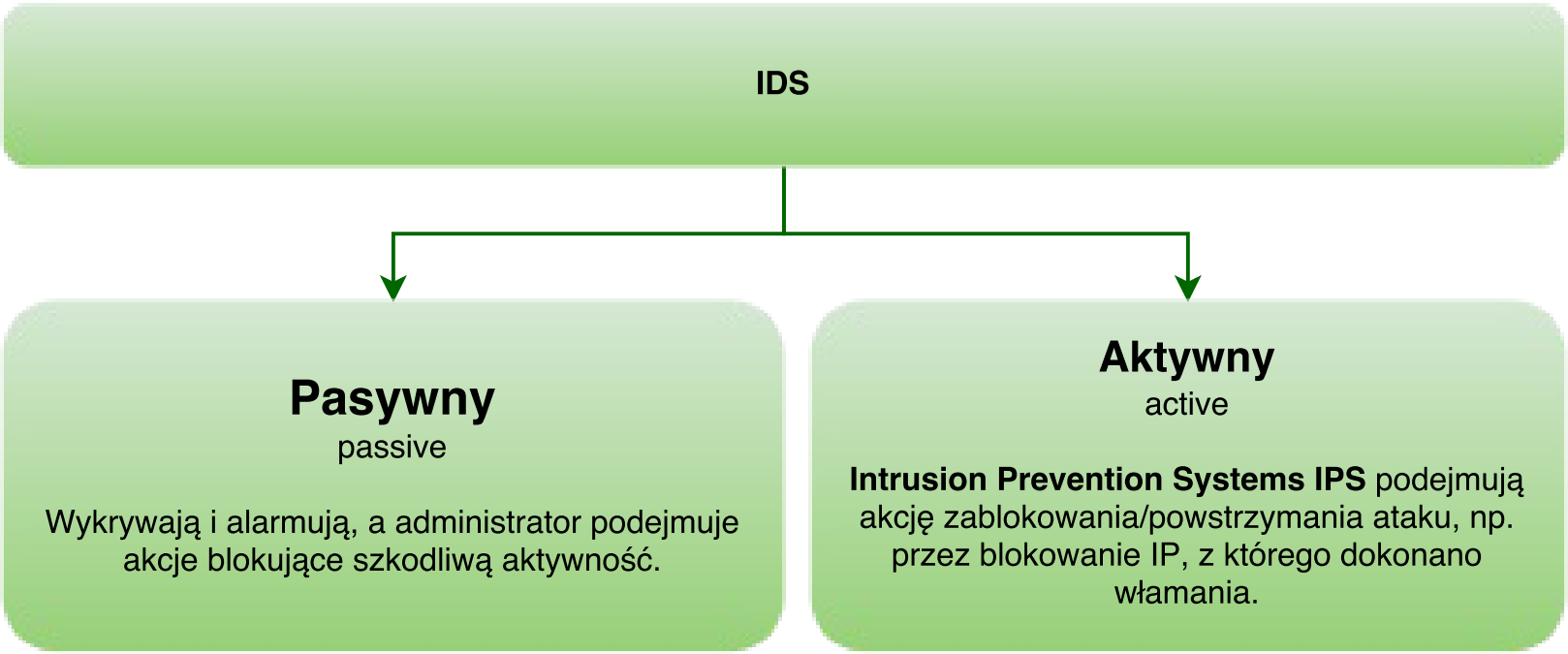
# IDS – podział ze względu na lokalizację ataku



# IDS – podział ze względu logikę detekcji



# IDS – podział ze względu reakcję



# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

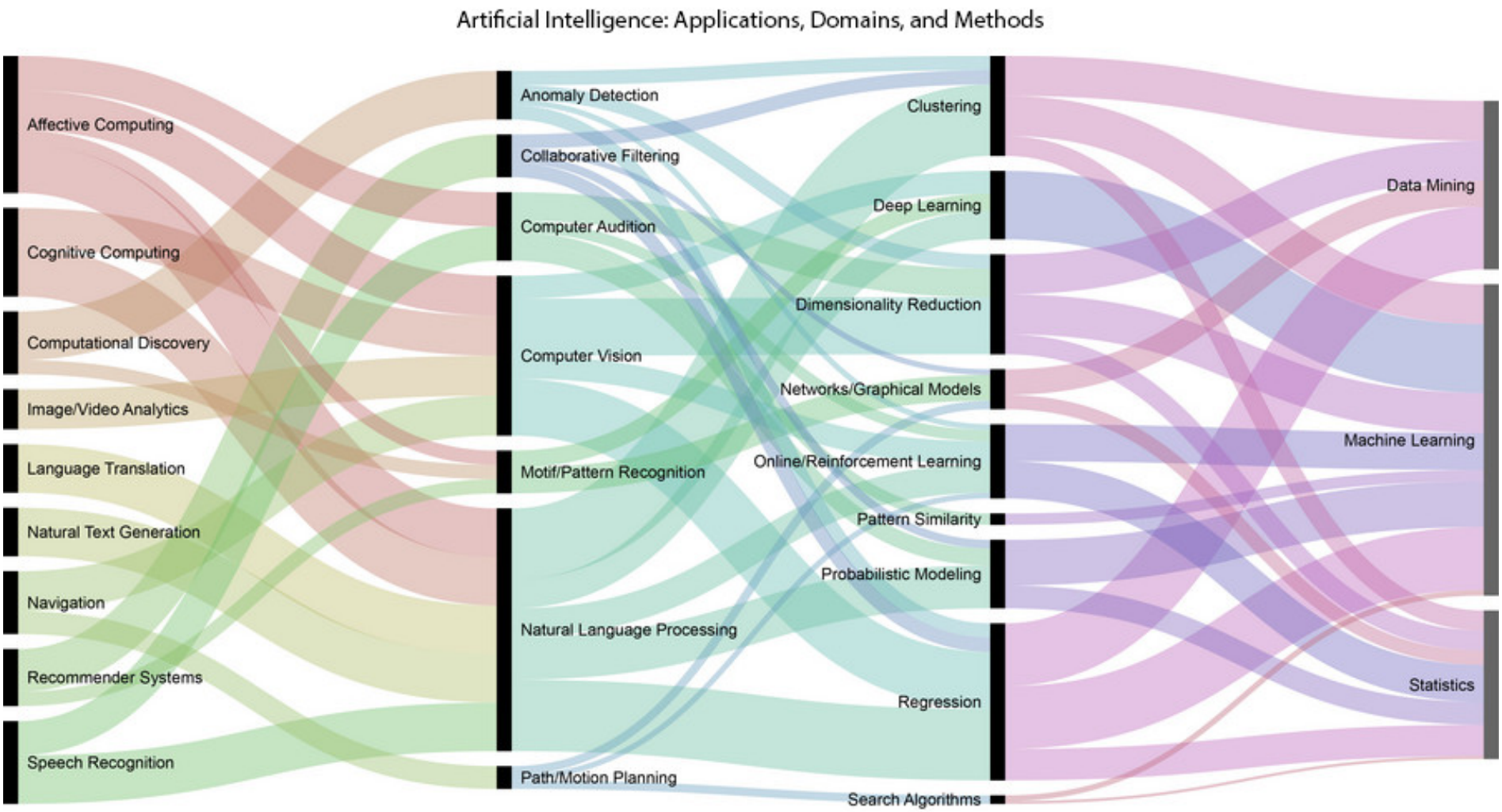
# Definicja

**Sztuczna Inteligencja — Artificial Intelligence (AI)**

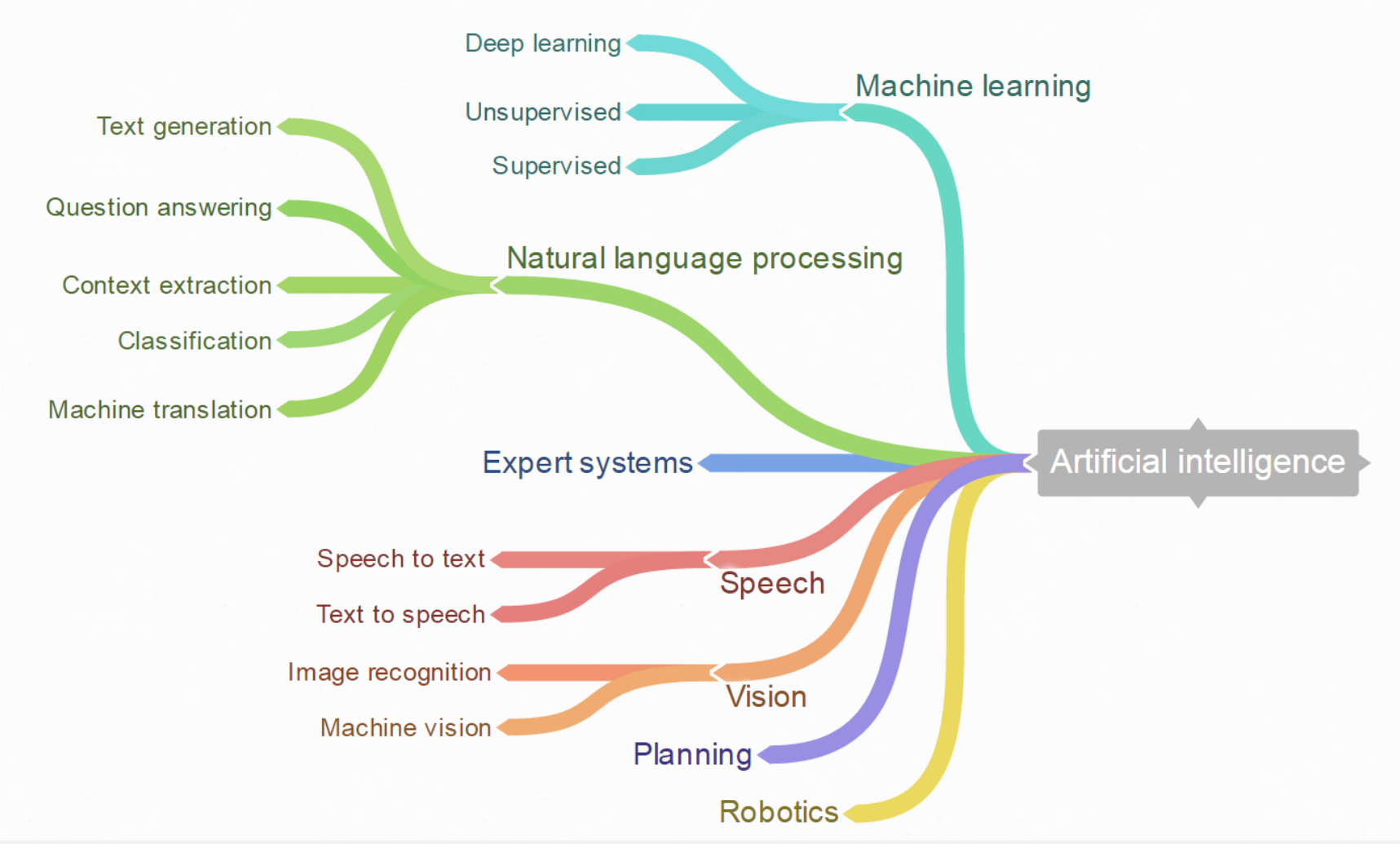
Dział informatyki. Zajmuje się tworzeniem inteligentnych artefaktów, czyli programów lub urządzeń/maszyn.

# Czym się zajmuje AI?

# Czym się zajmuje AI?



# Czym się zajmuje AI?




https://www.nuix.com/blog/will-artificial-intelligence-outlive-hype-cybersecurity

# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Co to jest?

**Znaczenie:**

Jest motorem dziedziny nazywanej: **data science**. Jest związana z dziedziną **data mining**, czyli drążeniem danych.
Każda metoda ML (algorytm) pobiera dane, przetwarza i generuje odpowiedź. Najistotniejszą częścią algorytmów ML jest uczenie się, które wykonują często matematyczne obliczenia.

# Uczenie nadzorowane w ML

## Uczenie nadzorowane:

Zadanie: predykcja
Cecha: w zbiorze danych istnieją wartości klasy/wyścia

Maszyna uczy się z danych wejście/wyjście potem stara się odgadnąć jakie jest wyjście dla zadanego wejścia.

**Regresja**

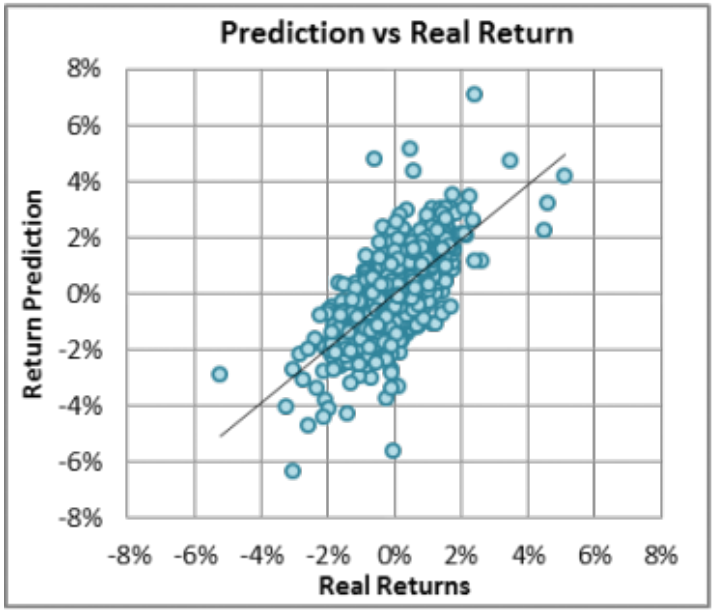
Szacuje wartości ciągłe
Wyjście: liczba rzeczywista

**Klasyfikacja**:

Identyfikuje przynależność do klasy
Wyjście: wartości dyskretne: binarne, kategorie

# Uczenie nadzorowane w ML - różnice

Regression vs Classification




# Klasyfikacja i Regresja

## Klasyfikacja

Dane układa się w ustalonych grupach (klasach). Dane składają się z atrybutów obiektów i przypisanych im etykiet klas. Jeżeli pojawi się nowy obiekt o znanych atrybutach a nieznanej klasie zostaje on dopasowany do jednej z klas. Modelowanie zależności wejście-wyjście nazywane jest uczeniem nadzorowanym.

## Regresja

Metoda pozwalająca na badanie związku pomiędzy wielkościami danych i przewidywanie na tej podstawie nieznanych wartości jednych wielkości na podstawie znanych wartości innych.

# Klasyfikacja i Regresja

Zarówno klasyfikacja jak i regresja używane są w dwóch fazach:

1. konstruowanie modelu – budowa tzw. modelu regresyjnego lub klasyfikacyjnego, czyli funkcji opisującej, jak zależy wartość oczekiwana zmiennej objaśnianej lub klasy od zmiennych objaśniających. Funkcja ta może być zadana nie tylko czystym wzorem matematycznym, ale także całym algorytmem, np. w postaci drzewa regresyjnego, sieci neuronowej, itp.. Model konstruuje się tak, aby jak najlepiej pasował do danych z próby, zawierającej zarówno zmienne objaśniające, jak i objaśniane (tzw. zbiór uczący).
2. stosowanie modelu – użycie wyliczonego modelu do danych, w których znamy tylko zmienne objaśniające, w celu wyznaczenia wartości oczekiwanej zmiennej objaśnianej lub klasy.

# Uczenie nadzorowane w ML - Regresja przykłady

**Pytanie:Jak dużo? Ile?**

Zwykle algorytmy regresji dają w odpowiedzi wartości rzeczywiste, często dużej precyzji z wieloma miejscami po przecinku, a nawet wartości ujemne. Dla niektórych pytań, zwłaszcza pytań rozpoczynających się od "Ile ...", negatywne odpowiedzi mogą być zaokrąglane do zera a wartości ułamkowe do najbliższej liczby całkowitej.

- Jaka będzie temperatura powietrza w następny wtorek?
- Jaka będzie sprzedaż w Portugalii w czwartym kwartale?
- Jakie będzie zapotrzebowanie na energię w kW z farmy wiatrowej za 30 minut?
- Ile nowych "followers" będę mieć w przyszłym tygodniu?
- Spośród tysiąca sztuk, ile z tego modelu łożysk przetrwa 10.000 godzin pracy?

# Uczenie nadzorowane w ML - Klasyfikacja binarna przykłady

## Pytanie: Czy A czy B?

Jest to dowolny problem, na który uzyskujemy tylko dwojaką odpowiedź: tak lub nie, włączony lub wyłączony, chory lub zdrowy. Wiele pytań w eksploracji danych brzmi właśnie tak, lub można je przeformułować na postać binarną. Jest to najprostsze i najczęściej zadawane pytanie w analizie danych.

- Czy ten klient odnowi subskrypcję?
- Czy to obraz kota lub psa?
- Czy ten klient kliknie na górny link?
- Czy opona się zniszczy w przeciągu kolejnego 1000 km?
- Czy kupon na 5$ czy na 25% rabatu przyciągnie ponownie więcej klientów?

# Uczenie nadzorowane w ML - Klasyfikacja do wielu klas- przykłady

**Pytanie: Czy to A czy B czy C czy D....?**

Odpowiedzi zatem jest wiele, czyli kilka lub kilkanaście, np. jaki to smak, która osoba, jaka część. Większość algorytmów klasyfikacji do wielu klas jest tylko rozszerzeniem algorytmów klasyfikacyjnych binarnej.

- Jakie zwierzę jest na tym obrazie?
- Który samolot wysyła ten sygnał radarowy?
- Co jest tematem artykułu prasowego?
- Jaki jest nastrój tego tweeta?
- Kto przemawia na nagraniu?

# Zmiana zadania klasyfikacji na regresję

Czasami pytania o wielowartościową lub binarną klasyfikację właściwie lepiej przekształcić na zadanie regresji. Przykład:

- Pytanie: „Która książka jest najbardziej interesujące dla czytelnika?" oznacza przypisanie pojedynczej kategorii, a można zapytać: „Jak bardzo interesująca jest ta książka dla czytelnika?" co oznacza przypisanie wartości liczbowej do odpowiedzi.
- Pytanie: „Które 5% moich klientów przeniesie się do konkurencji w przyszłym roku?" można zmienić na: „Jak prawdopodobne jest, że ten klient przeniesie się do konkurencji w przyszłym roku?".
- Pytanie: „Czy użytkownik kliknie w reklamę?" zamienić na „Jak prawdopodobne jest, że użytkownik kliknie w reklamę?"
- Pytanie: „Czy ten samolot doleci bez opóźnienia?" zmienić na „Jaki odsetek samolotów dolec bez opóźnienia?"

# Uczenie nadzorowane w ML - wykrywanie anomalii - przykłady

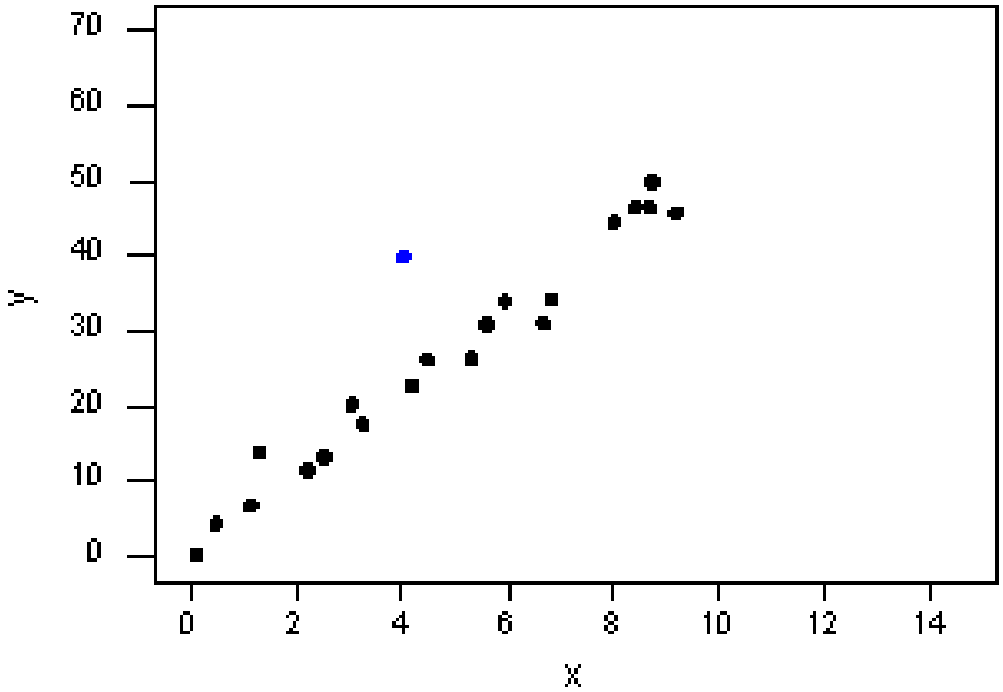
**Pytanie: Czy to nie dziwne?**

Znajdują punkty w danych, które są inne. Wykrywanie anomalii nie jest klasyfikacją binarną, bo nie istnieje w danych etykieta 'dziwności'. Celem jest zauważenie tak wyjątkowych danych, sytuacji, których nawet się nie etykietuje, bo ich nie zauważamy.

- Czy ten odczyt ciśnienie jest niezwykły?
- Czy jest to typowe zapytanie do serwera WWW?
- Czy ta płatność odbiega od standardowej procedury wykonywanej przez użytkownika?
- Czy takie wartości napięcia są typowe dla tej pory roku?

# Anomalie - odstające rekordy (outliers)

Stosując ML można łatwo zidentyfikować odstające rekordy i wskazać przyczynę tego stanu rzeczy.



# Uczenie nienadzorowane w ML

**Pytanie:**

Podstawowym pytaniem zadawanym przez uczenie nienadzorowane jest organizacja danych.

## Uczenie nienadzorowane:

Zadanie: szukaj struktur w danych
Cecha: w zbiorze danych nie istnieją wartości klasy/wyścia

Maszyna znajduje użyteczne informacje ukryte w danych.

**Analiza skupień**

Układa dane z zbiory.

**Szacowanie gęstości**

Przybliżony rozkład

**Redukcja wymiarowości**:

Wybiera istotne atrybuty.

# Uczenie nienadzorowane w ML - analiza skupień

- Istnieje wiele różnych technik, które starają się określić strukturę danych.
- Cel: podzielić cały zbiór danych na grupki (podzbiory).
- Uczenie nadzorowane można porównać do wybierania planet spośród gwiazd, to uczenie nienadzorowane jest tworzeniem konstelacji.
- Klastry, to grupki danych, które analityk może opisać i łatwiej przedstawić, czy interpretować, gdyż są logicznie powiązane.
- Klasteryzacja zawsze opiera się na pomiarze odległości tzw. metryce. Tą metryką może być dowolna mierzalna cecha np. różnica w IQ, liczba wspólnych genów, odległość w km.
- Klastry są mniej więcej jednolitymi grupami.

# Uczenie nienadzorowane w ML - Analiza skupień - przykłady

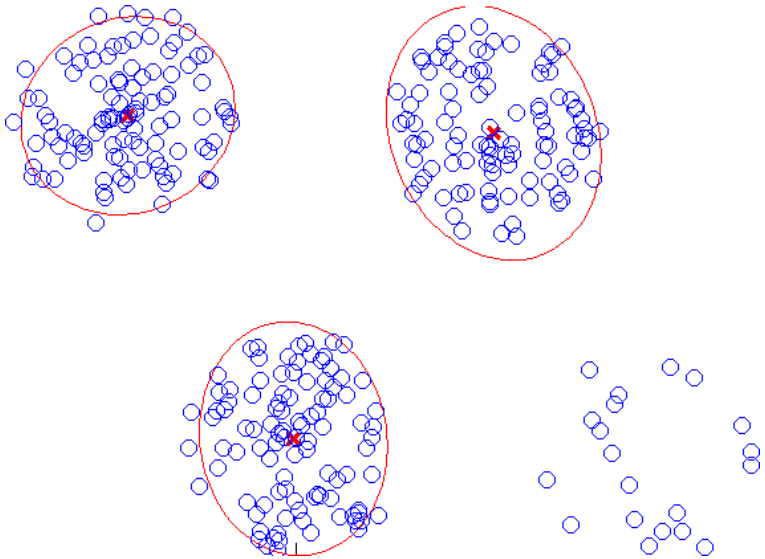
**Pytanie:**

Wszystkie pytania dotyczące grupowania starają się dzielić dane na jednolite grupy.

- Którzy klienci mają podobne gusta/ wybierają podobne produkty?
- Którzy widzowie lubią ten sam gatunek filmów?
- Które modele drukarek psują się w taki sam sposób?
- W jakie dni tygodnia pojawiają się takie same wymagania dotyczące zapotrzebowania na energię elektryczną?
- Jaki jest naturalny sposób podziału artykułów na 5 grup tematycznych?

# Analiza skupień

W danych nie ma predykcji. Grupowanie polega na obserwacji rozkładu danych w przestrzeni wejść i nadawanie tej samej etykiety dla bliskich wg zadanej metryki rekordów i separowalnych od innych.



# Różne wyniki analizy skupień/grupowania

Według przychodu

| ID | Name | Prediction | Age | Balance | Income | Eyes | Gender |
|---|---|---|---|---|---|---|---|
| 3 | Betty | No | 47 | $16,543 | High | Brown | F |
| 5 | Carla | Yes | 21 | $2,300 | High | Blue | F |
| 6 | Carl | No | 27 | $5,400 | High | Brown | M |
| 8 | Don | Yes | 46 | $0 | High | Blue | M |
| 1 | Amy | No | 62 | $0 | Medium | Brown | F |
| 2 | Al | No | 53 | $1,800 | Medium | Green | M |
| 4 | Bob | Yes | 32 | $45 | Medium | Green | M |
| 7 | Donna | Yes | 50 | $165 | Low | Blue | F |
| 9 | Edna | Yes | 27 | $500 | Low | Blue | F |
| 10 | Ed | No | 68 | $1,200 | Low | Blue | M |

[1]

[1] źródło: http://www.thearling.com/text/dmtechniques/dmtechniques.htm

# Różne wyniki analizy skupień/grupowania

Według wieku i koloru oczu

| ID | Name | Prediction | Age | Balance | Income | Eyes | Gender |
|---|---|---|---|---|---|---|---|
| 5 | Carla | Yes | 21 | $2,300 | High | Blue | F |
| 9 | Edna | Yes | 27 | $500 | Low | Blue | F |
| 6 | Carl | No | 27 | $5,400 | High | Brown | M |
| 4 | Bob | Yes | 32 | $45 | Medium | Green | M |
| 8 | Don | Yes | 46 | $0 | High | Blue | M |
| 7 | Donna | Yes | 50 | $165 | Low | Blue | F |
| 10 | Ed | No | 68 | $1,200 | Low | Blue | M |
| 3 | Betty | No | 47 | $16,543 | High | Brown | F |
| 2 | Al | No | 53 | $1,800 | Medium | Green | M |
| 1 | Amy | No | 62 | $0 | Medium | Brown | F |

2

[2] źródło: http://www.thearling.com/text/dmtechniques/dmtechniques.htm

# Problemy z grupowaniem

- Który rekord do którego klastra? Algorytm podziału na klastry powinien mieć określone zasady, jaka cecha ma większy priorytet i jaki atrybut jest ważniejszy.
- Jak uzyskać kompromis liczności klastrów i jednorodności. Chcąc uzyskać najbardziej jednorodne klastry będziemy mieli tendencję do zwiększania liczby klastrów (aż do liczby rekordów). Natomiast chcąc uzyskać generalizację trzeba dla danego problemu próbować budować jak najmniej klastrów.

# Uczenie nienadzorowane w ML - redukcja wymiarowości

- To sposób na uproszczenie danych, co pozwala na łatwiejszą ich analizę, czy przechowywanie.
- Proces redukowania liczby zmiennych, poprzez uzyskanie zestawu zmiennych podstawowych.
- W wyniku redukcji wymiarowości metody regresji czy klasyfikacji mogą być efektywniejsze niż na pełnym zbiorze zmiennych.
- Stosowane metody, to np.: analizy głównych składowych (PCA), liniowa analiza dyskryminacyjna (LDA), kanoniczna Analiza korelacyjna (CCA).

# Uczenie nienadzorowane w ML - redukcja wymiarowości - przykłady

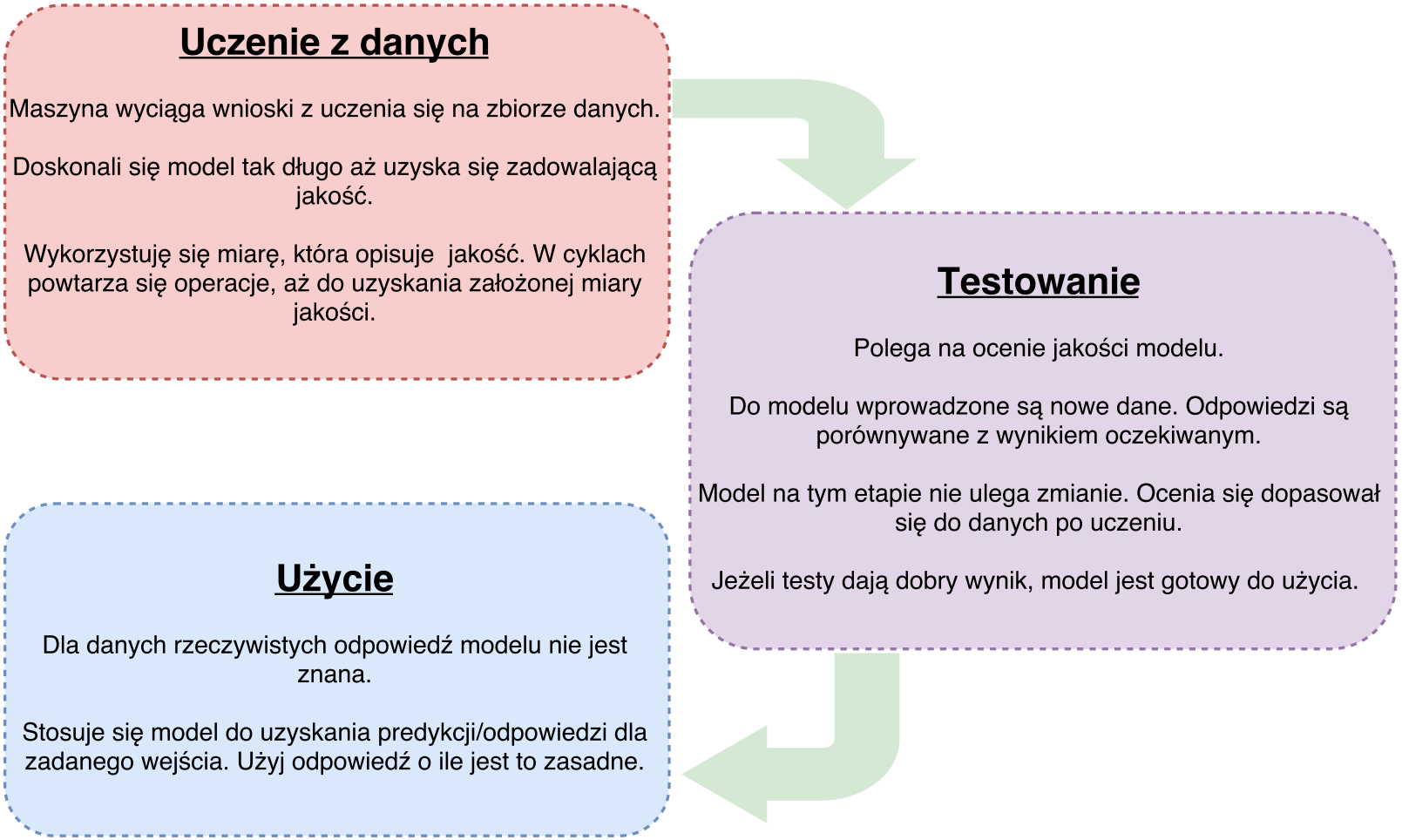
**Pytanie:**

Wszystkie pytania dotyczące grupowania starają się określić jakie dane są istotne, częstsze.

- Które grupy czujników w tym silniku odrzutowym zwykle różnią się od siebie?
- Jakie zachowania są wspólne dla dobrych kierowników?
- Jakie są najczęstsze wzorce zmian cen paliwa?
- Jakie grupy słów występują często razem w danym zbiorze dokumentów?

# Technika uczeni-testowanie-użycie

Techniki ML stosują etapy uczenia i testowania przed użyciem modelu.



# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Definicja

## Eksploracja danych ED (Data mining)

Metody wydobywania ukrytych informacji z dużych baz danych.

## Cel

Do prognozowania przyszłych trendów i zachowań, które pozwolą przedsiębiorstwom na podejmowanie opartych na wiedzy decyzji.

## Zalety

- Zautomatyzowana prospektywna analiza danych wykracza poza zwykłe narzędzia wspomagania decyzji.
- ED udziela odpowiedzi na pytania, które nie znajdowały odpowiedzi ze względu na złożoność obliczeniową.
- Poszukują w bazach danych ukrytych wzorców, informacji, które ekspert może pominąć, gdyż znajdują się poza jego oczekiwaniami.

# Technologie pozwalające na rzeczywiste wykorzystanie ED

Zasoby zapewniające wykorzystanie ED:

- olbrzymie i prawie wszechobecne zbiory danych
- zwiększająca się moc obliczeniowa komputerów
- algorytmy eksploracji danych.

Technologie eksploracji danych wywodzą się z obszarów badań:

- statystyka
- sztuczna inteligencja
- maszynowe uczenie się.

# Zakres zastosowania eksploracji danych

## Automatyczne przewidywanie trendów i zachowań

Automatyzuje się proces wyszukiwania informacji i można szybko udzielać odpowiedzi na pytania dotyczące danych.
Przykłady:

- Ukierunkowany marketing: wykorzystanie np. danych z przeszłych korespondencji promocyjnych do określenia klientów maksymalizujących szansę ponownych inwestycji.
- Prognozowanie upadłości: identyfikacja segmentów biznesu, które mogą reagować podobnie na pewną sekwencję zdarzeń.

# Zakres zastosowania eksploracji danych

## Automatyczne wykrywanie nieznanych wcześniej wzorców

Narzędzia eksplorują bazy danych i identyfikują ukryte wzorce.
Przykłady odkrywania wzorców

- Analiza danych o sprzedaży detalicznej do identyfikacji pozornie niepowiązanych produktów, które często są nabywane razem.
- Wykrywanie wzorca fałszywych transakcji z użyciem kart kredytowych.
- Identyfikacja anomalii w danych.

# Plan wykładu

1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Zadania wykonywane w ramach eksploracji danych

- Klasyfikacja
- Grupowanie
- Asocjacje
- Wzory sekwencyjne

# Asocjacje

**Definicja**

Dane służą do identyfikacji związków pomiędzy atrybutami. W dużych bazach danych poszukuje się reguł, które określają silne (według przyjętego kryterium) powiązania pomiędzy cechami obiektu.

Przykłady:

- Reguła asocjacyjna: piwo-> pieluchy, cebula, ziemniaki -> mięso. Wykorzystywane do rozmieszczanie produktów i akcji promocyjnych.
- Wykrywanie włamań komputerowych. Wykrywanie reguł, które łączy się z atakiem. Wykonuje się to poprzez analizy tysięcy linii logów i poszukiwaniu anomalii.

# Wzory sekwencyjne

**Definicja**

Dane wykorzystuje się do przewidywania zachowań i trendów. Dane pojawiają się sekwencyjnie i przechowywane są w sposób wskazujący na kolejność ich pojawiania się. Poszukuje się w nich statystycznie istotnych wzorców.

Przykłady:

- Sprzedawcy sprzętu mogą przewidzieć prawdopodobieństwo nabycia ubezpieczenia w pewnym czasie po zakupie komórki na podstawie zakupu konkretnego typu telefonu komórkowego. W pewnym sensie jest to wykrycie reguły, w której ważne jest następstwo czasowej
- Wykrywanie brakującego fragmentu DNA lub ciągu znaków.

# Przykładowe techniki eksploracji danych

- **sztuczne sieci neuronowe**
- **drzewa decyzyjne**: struktura drzewiasta, które zawiera zestawy decyzji. Decyzje te generują zasad klasyfikacji zbioru danych. Metody wykorzystujące drzewa decyzyjne to drzewa klasyfikacyjne i regresyjne.
- **algorytmy genetyczne** do wykrywania reguł w danych
- **metoda najbliższego sąsiedztwa**: technika, która grupuje rekordy w zbiorze danych łącząc ze sobą $k$ rekordów najbliższych (podobnych do niego) dla pewnego wybranego rekordu.
- **indukcja reguł**: wydobycie reguł (jeśli-to) w oparciu o istotność statystyczną.

# Plan wykładu

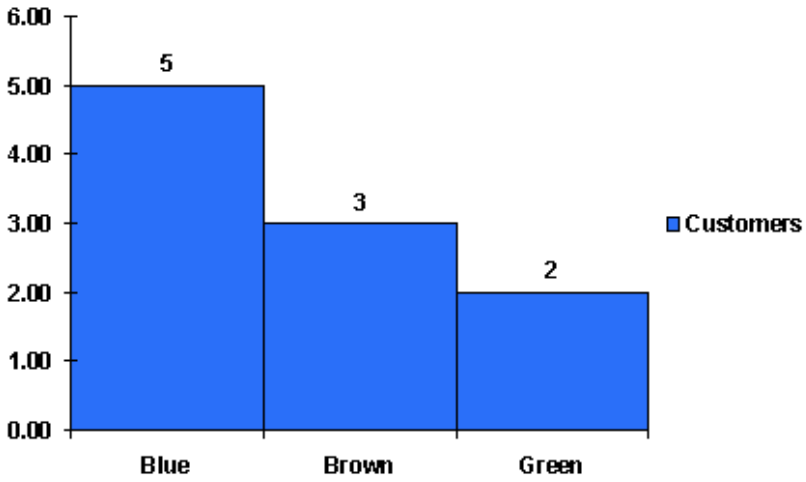
1. Wprowadzenie
2. IDS — definicje i klasyfikacja
3. Krótka powtórka z AI
4. Maszynowe uczenie się
   - Uczenie nadzorowane
   - Uczenie nienadzorowane
5. Eksploracja danych
6. Jak działa eksploracja danych?
7. Statystyka w eksploracji danych

# Statystyka

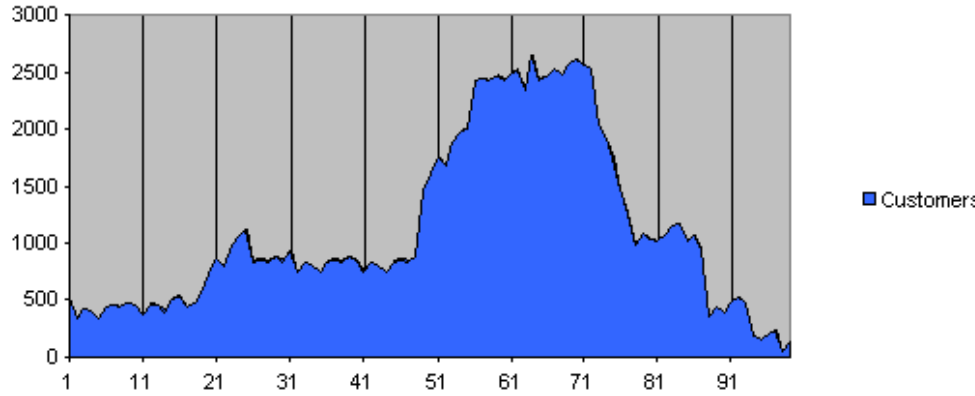
Używając narzędzi ze statystyki można udzielać odpowiedzi na pytania:

- Jakie wzorce są ukryte w bazie danych?
- Jaka jest szansa, że nastąpi pewne zdarzenie?
- Jakie wzorce są istotne?
- Co wynika z „podsumowania” (np. średnia) danych? Zyskuje się pewne wyobrażenie o tym, co jest zawarte w bazie danych.

# Histogramy




kolor oczu




wiek

[3] źródło: http://www.crisp-dm.org/Process/index.htm

# Użyteczne miary

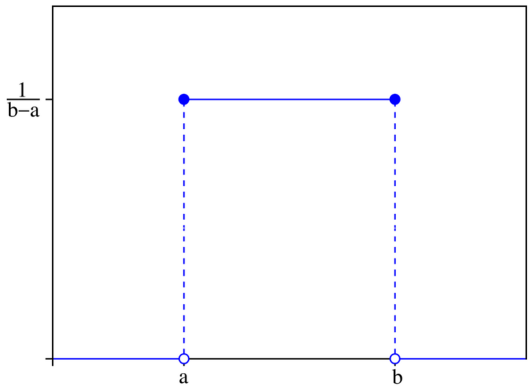
- **Max** - maksymalna wartość z danych.
- **Min** - minimalna wartość z danych.
- **Średnia** - średnia wartość w próbie.
- **Mediana** - wartość w bazie, powyżej i poniżej której znajduje się jednakowa liczba rekordów (dzieli bazę na połówki o równej liczbie rekordów).
- **Dominanta** - wartość najczęściej występująca (o największym prawdopodobieństwie wystąpienia).
- **Wariancja** - miara zmienności, tego, jak rozkładają się wartości od wartości średniej.

# Rozkłady

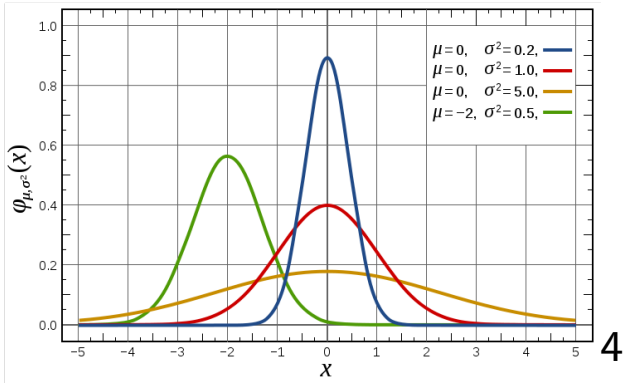
Czasami zamiast histogramu chce się opisać rozkład danych równaniem. W klasycznej statystyce zakłada się, że istnieje pewien „prawdziwy", podstawowy kształt rozkładu, który powstaje wtedy, gdy zostaną zebrane wszystkie możliwe dane.

Zadaniem statystyka jest określenie prawdopodobnego rozkładu z ograniczonej liczby danych .

Wiele rozkładów opisanych jest tylko przez średnią i wariancję.




jednostajny



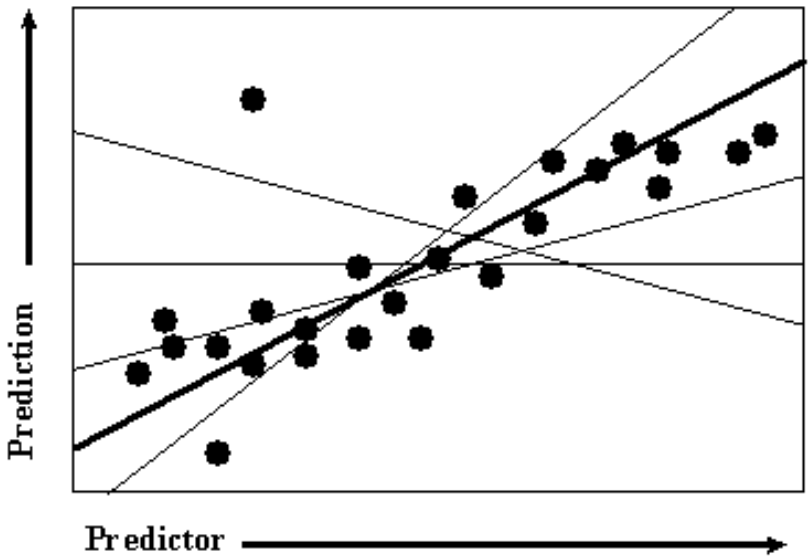

normalny

[4] źródło: wikipedia

# Regresja liniowa

Podstawowa zasada regresji jest taka, że z mapy wartości jest tworzony taki model, by uzyskać najniższy błąd (zazwyczaj średniokwadratowy).

$$Prediction = a + b \cdot Predictor$$




[5]

[5] źródło: http://www.thearling.com/text/dmtechniques/dmtechniques.htm

# Bardziej złożone modele niż liniowe

Złożoność modelu może wynikać z:

- zwiększenia liczby wejść (predictors) (zwiększenie wymiarowości)

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5$$

- regresja nieliniowa — zastosowania przekształcenie dla wejścia (podnoszenie do potęgi)

$$Y = a + b_1 X_1 + b_2 X_1^2$$

- wymnażania przez siebie wejść
- modyfikacji by odpowiedź modelu była binarna (regresja logistyczna)