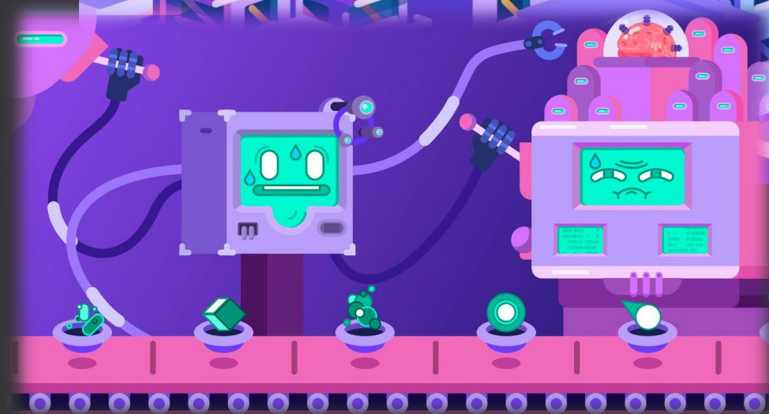




Metody i narzędzia *Big Data*

Organizacja zajęć



Materiał kursu

potoki → uczenia → maszynowego
dla masowej ilości danych

Potoki uczenia maszynowego z wykorzystaniem prawdziwych danych: CRISP-DM

- Inżynieria cech:
 - redukcja wymiarowości
 - selekcja cech
- Potoki przetwarzania tekstu

Obsługa dużych wolumenów danych

- model 3V
- modele obliczeń:
 - *MapReduce, MPI, barierowy*
- ekosystemy *Hadoop* i *Spark*

Egzamin

- I termin – 3 lipiec 2025, godz. 17:05-18:55, bud. D1 s. 30
- II termin – 10 lipiec 2025 (samoocena: anuluję ocenę uzyskaną w I terminie),
godz. 17:05-18:55, bud. D1 s. 30
- Forma pisemna: **zrozumienie pojęć i zasad działania** metod
 - 1 otwarte pytanie dotyczące
stosowalności metod uczenia maszynowego w dużej skali

| CZERWIEC | | | | | LIPIEC | |
|----------|-----------|------------|------------|----|--------|----|
| 2 | 9 Cz/P | 16 | 23 | 30 | 7 | 14 |
| 3 | 10 | 17 | 24 Pt/P | 1 | 8 | 15 |
| 4 | 11 | 18 Pt/N | 25 Cz/N | 2 | 9 | |
| 5 | 12 | 19 | 26 | 3 | 10 | |
| 6 | 13 | 20 | 27 | 4 | 11 | |
| 7 | 14 | 21 | 28 | 5 | 12 | |
| 8 | 15 | 22 | 29 | 6 | 13 | |
| N | P | N | P | N | P | N |

Literatura

