

# Praktyczne problemy w Data Science

## oraz

# Dylematy etyczne i konflikty interesów

Ziemowit Dworakowski  
AGH University in Krakow,  
Department of Robotics and Mechatronics

1

---

---

---

---

---

---

---

---

- ZDm*
- ❖ Nieprecyzyjne pojęcia bywają różnie rozumiane przez DS i przemysł
  - ❖ Większość osób współpracujących z „ekspertami od ML” **nie ma pojęcia** które zadania są łatwe a które są trudne
  - ❖ Większość osób postrzega metody ML jako magiczną czarną skrzynkę która może zrobić wszystko, jeśli tylko ekspert jest wystarczająco dobry
  - ❖ Większość osób współpracujących z „ekspertami od ML” **błędnie ocenia jakość dostępnych danych**
  - ❖ Większość osób współpracujących z „ekspertami od ML” **nie rozumie ryzyka przeuczenia i wycieku danych**
  - ❖ Większość osób współpracujących z „ekspertami od ML” **nie docenia istotności wstępnego przetwarzania danych**
  - ❖ Większość osób współpracujących z „ekspertami DS” **przypisuje zbyt duże znaczenie rodzajowi użytego klasyfikatora**
  - ❖ W przemyśle powszechna jest presja na „efekt marketingowy” nawet kosztem jakości merytorycznej rozwiązania.

2

---

---

---

---

---

---

---

---

- ZDm*
- Większość ekspertów od ML nie rozumie jak trudno w przemyśle zbiera się dane
  - Większość ekspertów od ML ćwiczyła wyłącznie na danych wysokiej jakości o dobrze reprezentowanych klasach
  - Eksperti od ML lubią „iść na łatwiznę” wrzucając wektory danych do jakiegoś systemu decyzyjnego i jeśli nie wychodzi, żądać „więcej danych!” zamiast spróbować zbudować **lepsze cechy** i wykorzystać **wiedzę kontekstową**
  - Większość „ekspertów od ML” nie docenia znaczenia wiedzy kontekstowej i przecenia możliwości użytego klasyfikatora
  - Większość ekspertów od ML błędnie ocenia istotność fałszywie pozytywnych wskazań i źle interpretuje skuteczność metod
  - Większość ekspertów od ML błędnie ocenia istotność fałszywie pozytywnych wskazań i źle interpretuje skuteczność metod
  - Większość ekspertów od ML błędnie identyfikuje znaczenie czynników wpływających na ostateczny wynik

3

---

---

---

---

---

---

---

---

## Wymagane nastawienie



**Bądź gotów na pracę z „brzydkimi” zbiorami danych**  
(Niepełne opisy, outliery, braki itp.)



**Bądź gotów na kompromisy i niestandardowe rozwiązania**  
(Niektórych problemów nie da się rozwiązać bez nagięcia części zasad)



**Potrzebujesz kredytu zaufania**  
(na podstawie doświadczenia i umiejętności komunikacyjnych)



**Chroń marketingowców przed nimi samymi**  
(„Hej, chcesz krótką kontrolę merytoryczną Waszej ulotki?”)



**Pamiętaj o rzeczywistym celu swojej pracy**  
(Rzadko kiedy celem jest „wykonanie klasyfikacji”)

4

---

---

---

---

---

---

---

---

---

---

## Wymagane umiejętności



**Podstawowa wiedza inżynierska**  
(By zrozumieć kontekst danych i cel swojej pracy)



Dobre zrozumienie problemów wycieku danych, przeczenia i interpretacji wyników  
(To one w praktyce limitują możliwości systemów AI w inżynierii)



**Umiejętność uczenia się nowych rzeczy i znajdowania źródeł**  
(Często praca w DS wymaga szybkiego przeskakowania między dziedzinami)



**Podstawowa wiedza o AI**  
(Głębokie zrozumienie szczegółów metod nie jest na początku aż tak istotne...)



**Doświadczenie z przeszłych projektów z nastawieniem na projektowanie i optymalizację własnych rozwiązań**



**Python**  
(razem z bibliotekami do AI)

5

---

---

---

---

---

---

---

---

---

---

## Dylematy etyczne i konflikty interesów

6

---

---

---

---

---

---

---

---

---

---



### Studium przypadku IV

**Józef** jest specjalistą od *data science* w międzynarodowej firmie. Pełni również rolę merytorycznego eksperta w procesie rekrutacji – zatrudniając innych inżynierów o tej specjalności. Firma chce zatrudnić jednego nowego pracownika. Do firmy zgłosiło się dwóch kandydatów, zdanie Józefa będzie decydujące:



**Rashid** – świetne CV. Spełnia wszystkie wymagania zawiązką, Józef go nie zna.



**Maks** – kandydat który spełnia prawie wszystkie wymagania. Prywatnie znajomy Józefa ze studiów, Józef wie, że Hans się umie szybko uczyć i w razie czego doczy się brakujących kwestii.

10

---

---

---

---

---

---

---

---

---

---

### Proste rozwiązania do rozważenia w sytuacji problemów etycznych:



**1. Konflikt interesów, nacechowanie emocjonalne:**

Wycofaj się, przekaż problem osobie obiektywnej



**2. Konflikty wartości:** Rozważ możliwość fałszywej dychotomii. Być może rozwiązanie kompromisowe jest możliwe?



**3. Jeśli nie możesz rozwiązać problemu sam, **eskaluj!****  
(to znaczy: przekaż problem na wyższy szczebel hierarchii)



**4. Znaj swoje wartości i wiedz jak się efektywnie komunikować**



**5. Korzystaj z różnych strategii decyzyjnych, staraj się spojrzeć na problem w całości**  
(Wiele konfliktów jest podtrzymywanych tylko dlatego, bo osoby zaangażowane nie zauważają możliwych rozwiązań)

11

---

---

---

---

---

---

---

---

---

---

### Informacja zwrotna...!

*Skorzystajcie proszę z ankiet studenckich na koniec semestru. Zróbcie to dla jak największej ilości przedmiotów i nauczycieli. One mają dużo większe znaczenie niż myślicie :)*

12

---

---

---

---

---

---

---

---

---

---