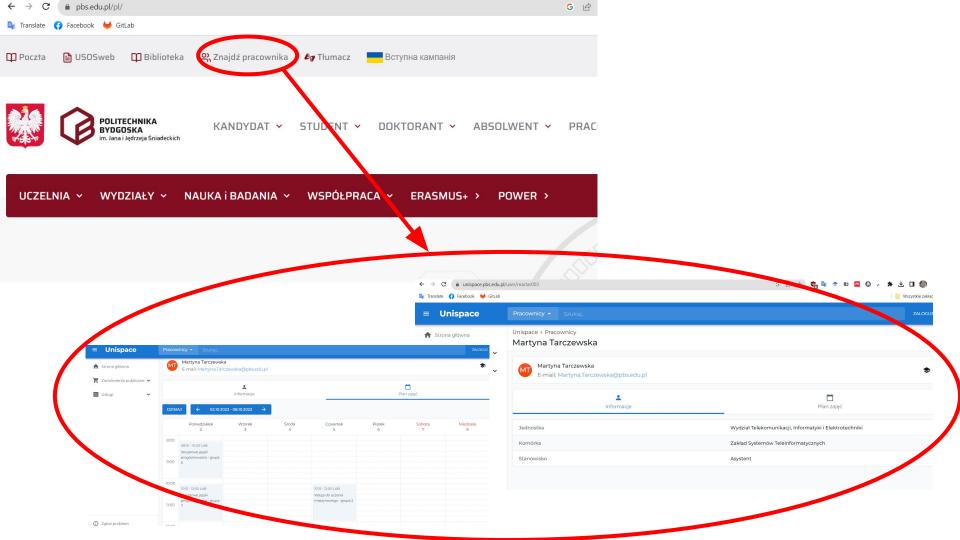
# Wdrażanie technik uczenia maszynowego - projekt

## Agenda

- 1. Wstęp
- 2. BHP
- 3. Tematyka zajęć
- 4. Zasady zaliczenia

#### Kontakt

- → mailowy martyna.tarczewska@pbs.edu.pl tylko z maila uczelnianego (pbs.edu.pl)
- → osobisty konsultacje (czwartki 8:00-10:00 , B 312 / 3 PIĘTRO), zdalnie poprzez Teams: hoj6q4f



#### Na zajęciach:

- → BHP podpisy na liście
- → punktualność (kwadrans akademicki)
- → obecność obowiązkowa (max. 3 nieobecności) lista na każdych zajęciach
- → usprawiedliwienie nieobecności od lekarza lub PBŚ
- → telefony, przerwy, wyjścia...
- → picie i jedzenie
- → komputery własne lub lab.

#### Zajęcia:

- → stacjonarnie + Teams
  - e3zihdw utworzyć folder IMIE\_NAZWISKO\_NR-ALBUMU
- → co poniedziałek (jeśli wypadną -> @)
- → starosta?
- → godzina rozpoczęcia? (15 min później start, przerwa czy 15 min wcześniej koniec)

## **BHP**

### Tematyka zajęć

- 1. Uczenie maszynowe, sztuczna inteligencja
- 2. Docker
- 3. CI/CD git
- 4. Rozwiązania chmurowe
- 5. Front w gradio

#### Adekwatne oprogramowanie:

- → środowisko programistyczne Python 3.X (Visual Studio Code, Jupyter Notebook, Google Colab, pyCharm, Tonny, Notepad) <a href="https://code.visualstudio.com/docs/languages/python">https://colab.research.google.com/</a> (lokalnie)
- → środowisko uczelniane <a href="https://jupyter.io.pbs.edu.pl">https://jupyter.io.pbs.edu.pl</a>

## Zasady zaliczenia

- 1. **Projekt** wykonany w parach
- 2. >3 nieobecności nieusprawiedliwione  $\rightarrow$  n.p.

Jeśli jest problem ze zrozumieniem zadania /

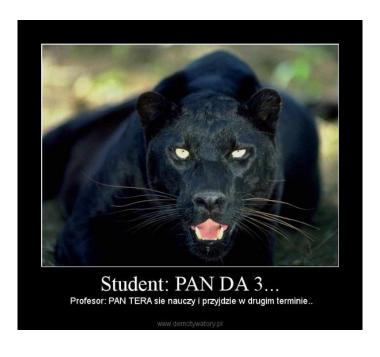
techniczny / zdrowotny / jakikolwiek inny → zgłaszać od razu!

#### Projekt:

- → czas **do 03.06.2024** godz. 23:59
- → wykonany w parach
- → kod (notebooki, skrypty) na GitHubie/GitLabie, sprawozdanie na Teamsie
- → **sprawozdanie** (jedno na parę) zawierające wypełnioną tabelkę (wzór dostarczę), link do kodu, wykresy, wyniki, uzasadnienie podjętych decyzji i **wnioski**
- projekt ma dotyczyć zagadnienia klasyfikacji jednego z zadań tzw. uczenia z nadzorem
- → każda para musi znaleźć zbiór danych (do akceptacji dane obrazkowe) i wymyślić temat projektu - <a href="https://www.kaggle.com/datasets">https://www.kaggle.com/datasets</a> (najlepiej jakiś z wysokimi ocenami)

#### W ramach projektu należy:

- znaleźć interesujący zbiór danych, który może zostać użyty w zadaniu klasyfikacji można wykorzystać kod z projektu WDUM
- → dokonać **eksploracyjnej** (ale zwięzłej) analizy danych (może być w notebooku)
- → wytrenować i odpowiednio zwalidować min. 2 algorytm (transfer learning albo dowolny inny)
- → zmierzyć i porównać **jakość** modeli wykorzystując do tego odpowiednie **miary**
- → strojenie hiperparametrów
- → Przygotować prosty front z użyciem biblioteki **Gradio / Dash** załadowanie przykładowego obrazka, odpowiedź modelu
- → Utworzyć Dockerfile, obraz dockerowy, odpalić program przez Docker,
- → Zaimplementować CI/CD na gicie
- → Wrzucić obraz dockerowy na dowolną chmurę wykonać trening modelu.



Ocena końcowa to ocena z wykonanego projektu oraz z aktywności na zajęciach. Dodatkowe plusy za obecność na pyData.

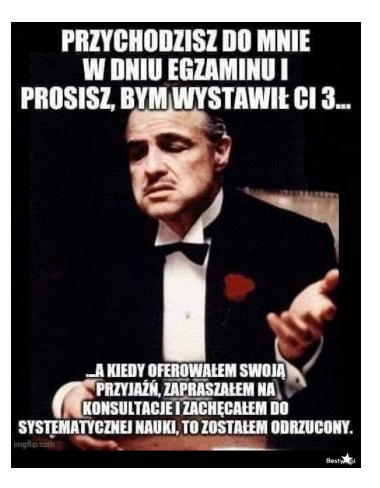


- plagiat jawny polegający na podpisaniu własnymi danymi osobowymi (imieniem i nazwiskiem, numerem indeksu, numerem Pesel etc.) pracy, której treść w 100% stanowi całość lub fragment cudzej pracy (czyli czynność, którą określamy pojęciem "kopiuj-wklej")
- plagiat ukryty polegający na zamieszczeniu w swojej pracy fragmentów pochodzących z pracy innego autora/autorów, bez podania informacji o źródle pochodzenia.

#### Co grozi w przypadku popełnienia plagiatu?

Studentowi, który zostanie obwiniony o popełnienie plagiatu grozi nie tylko postępowanie dyscyplinarne realizowane na uczelni, ale również postępowanie karne prowadzone przez prokuraturę, a nawet odpowiedzialność cywilna.

# Pytania?



# Gra integracyjna

(Kahoot albo skribbl.io)