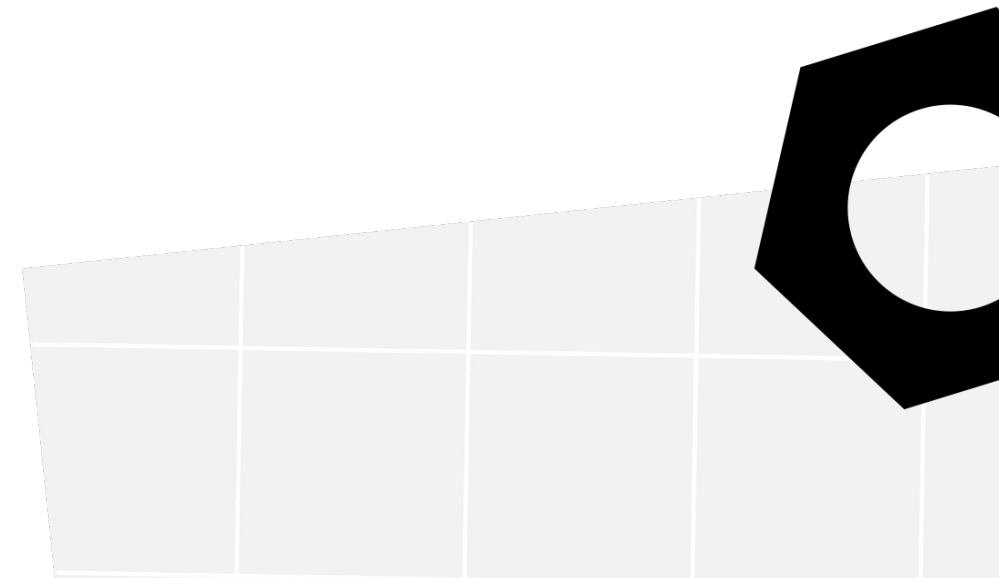




Projekt praktyczny (klasyfikacja)

Kurs Data Science





**Projekt praktyczny jest podsumowaniem
dotychczasowych zajęć poprzez udział w
konkursie! :)**



PLAN DNIA



- Podział na grupy projektowe (max. 2-3 osoby) – koniecznie nazwijcie swoją grupę :)
- Praca w grupach nad zadaniem konkursowym:
 - Zrozumienie i analiza problemu, który prezentują dane
 - Wstępne zapoznanie się z przedstawionymi danymi
 - Analiza, wizualizacja danych oraz przeprowadzenie EDA
 - Użycie i przetestowanie różnych algorytmów klasyfikacji
 - W trakcie dnia jest możliwość 3-krotnej próby na zbiorze konkursowym. Pod uwagę będzie brany najwyższy wynik f1-score. Wygrywa zespół z najwyższym wynikiem.
 - Czas na zadanie kończy się godzinę przed końcem zajęć



PLAN DNIA



- Ostatnia godzina to przygotowanie (30 min) i krótka prezentacja (30 min) użytych algorytmów i wyciągniętych wniosków, przed resztą grup.



Uwagi

- Prezentacja może przybrać dowolną formę.
- Kod powinien być dzielony pomiędzy członkami zespołu przy użyciu zdalnego repozytorium git.
- W przypadku problemów z dobraniem się w zespoły trener dokona podziału losowego.
- Niech wygra najlepszy! :)



Dane do pobrania



Podczas projektu będziemy korzystać z danych, które znajdziesz w dodatkowych plikach do bloku.

Możesz je pobrać z sekcji
"Dodatkowe materiały do bloku" [tutaj](#).



Treść zadania

Jakie czynniki warunkują chęć pozostania pracownika u obecnego pracodawcy?
Spróbuj przewidzieć, czy dany pracownik odejdzie z pracy, czy też nie.



Wyjaśnienie części danych:

Education

- 1 'Below College'
- 2 'College'
- 3 'Bachelor'
- 4 'Master'
- 5 'Doctor'

EnvironmentSatisfaction

- 1 'Low'
- 2 'Medium'
- 3 'High'
- 4 'Very High'

JobInvolvement

- 1 'Low'
- 2 'Medium'
- 3 'High'
- 4 'Very High'

JobSatisfaction

- 1 'Low'
- 2 'Medium'
- 3 'High'
- 4 'Very High'



Wyjaśnienie części danych:

PerformanceRating

- 1 'Low'
- 2 'Good'
- 3 'Excellent'
- 4 'Outstanding'

RelationshipSatisfaction

- 1 'Low'
- 2 'Medium'
- 3 'High'
- 4 'Very High'

WorkLifeBalance

- 1 'Bad'
- 2 'Good'
- 3 'Better'
- 4 'Best'



Wyjaśnienie części danych:

EmployeeCount

Ilość osób reprezentowanych przez dany wpis.

Monthly Rate a Monthly income

Monthly rate jest to koszt pracodawcy, jaki ponosi w związku z zatrudnieniem pracownika.

DailyRate analogicznie.

Attrition

Informacja czy dany pracownik odszedł z pracy, czy też nie.

STOCK OPTIONS LEVEL

Jaki procent akcji posiada/może posiadać pracownik.

PERCENT SALARY HIKE

Procentowy wzrost płacy w ciągu ostatnich 2 lat.