

Zespoły projektowe:

Projekty wykonywane są w zespołach **3 – 4 osobowych**. Dopuszczane są zespoły 2 lub 5 osobowe, ale niechętnie, po konsultacji ze mną. Projekty indywidualne, 1 – osobowe są także możliwe, ale tylko w szczególnych przypadkach i z odpowiednim, przekonującym umotywowaniem. Nie przewiduję tzw. szybkiej ścieżki.

Tematy projektów:

Tematów projektów można szukać na:

<https://www.drivendata.org/competitions/>

<https://www.data-mining-cup.com/>

<https://www.kaggle.com/competitions>

<https://datahack.analyticsvidhya.com/contest/all/>

Nie muszą to być aktywne konkursy, może być coś z przeszłości. Nie można korzystać z gotowych rozwiązań, wszystko należy opracować i wykonać samodzielnie.

Kilka uwag:

- Należy przygotować instrukcję użytkownika opisującą, krok po kroku, szczegółowo, co należy zrobić, aby uruchomić program i wszystkie jego podprogramy.
- Trzeba przygotować dokumentację dotyczącą wdrożonych metod. Dokumentacja musi obejmować:
 - podstawową wiedzę wykorzystaną do opracowania tych metod z odniesieniami do wykorzystanych źródeł,
 - szczegółowy opis algorytmów, główne funkcje przetwarzania danych powinny posiadać pseudokod,
 - studium przypadków, w których należy sprawdzić, jak działają metody dla różnych przykładów.

Ostatnia część jest bardzo ważna. Wymaga dobrze przemyślanego pomysłu. Pomyśl, co chcesz pokazać i udowodnić. Celem jest zbadanie wdrożonych metod i przedstawienie ich działania. W zależności od rodzaju projektu zakres tego raportu będzie różny.

Niektóre standardowe elementy, które należy wziąć pod uwagę, to:

- opis wykorzystanych zbiorów danych,
- ocena efektywności za pomocą miar, takich jak dokładność, błąd, krzywe ROC itp.,
- hipotezy badawcze (wyrażone ustnie, a następnie potwierdzone za pomocą eksperymentów z danymi) lub
- pytania badawcze (wyrażane ustnie, a następnie odpowiadane za pomocą eksperymentów),
- analiza porównawcza wyników dla różnych parametrów/podprogramów.

Projekt musi zawierać zaimplementowane scenariusze testowe (scenariusze opisane w dokumentacji). Scenariusze testowe powinny być wystarczająco elastyczne, aby umożliwić testowanie kodu z innymi zestawami danych.

Harmonogram projektów

NR ZAJĘĆ	DATA	TEMAT
2.	03.03.2022 04.03.2022	Przedstawienie tematów projektów Powinniście być podzieleni na zespoły (3-4 osoby). Każdy zespół wybiera temat projektu i prezentuje temat ze wstępnymi założeniami, hipotezami... Ogólna dyskusja na temat projektów.
3.	10.03.2022 11.03.2022	Ostateczny wybór projektu Zespoły przygotowują dokument, w którym przedstawiony będzie temat, krótki opis, wymienione zostaną: podział zadań na członków zespołu (ogólnie), wybrany język programowania, wstępna propozycja doboru narzędzi do przetwarzania danych – może ona ulec później zmianie. Od tego momentu rozpoczyna się etap bardziej teoretyczny – macie się zaznajomić z literaturą i wybrać konkretne narzędzia i ścieżki przetwarzania danych. (około 10 % oceny)
5.	24.03.2022 25.03.2022	Omówienie projektów Dyskusja o postępie w pracach nad projektami. Poszczególne zespoły omawiają swoje projekty. Powinien być tu przygotowany już szczegółowy plan zadań dla każdego członka zespołu. Nie jest wymagany duży postęp w programowaniu rozwiązania, wyłącznie w zakresie wczytania i zaznajomienia się z danymi. Od tego momentu rozpoczyna się etap bardziej praktyczny. (około 10 % oceny)
8.	14.04.2022 22.04.2022	Omówienie postępów Dyskusja o postępie w pracach nad projektami. Poszczególne zespoły omawiają swoje projekty. W tym terminie oczekuje się intensywnie prowadzonych prac implementacyjnych i eksperymentów. (około 10 % oceny)
11.	05.05.2022 12.05.2022(czw)	Omówienie postępów Dyskusja o postępie w pracach nad projektami. Poszczególne zespoły omawiają swoje projekty. (około 10 % oceny)
13.	26.05.2022 27.05.2022	Prezentacje projektów Prezentacje projektów. Ma na celu ustalenie zakresu końcowych poprawek i zmian.
14.	02.06.2022 03.06.2022	Termin oddania projektów Standardowy termin oddania projektów, max ocena 5.0.
15.	10.06.2022	Termin ostateczny oddania projektów Spóźnienie o 1 tydzień: maksymalna ocena to 4.0, spóźnienie o 2 tygodnie: maksymalna ocena to 3.0.