

Zadanie rekrutacyjne, DSC PJATK 2025

Witajcie w konkursie rekrutacyjnym organizowanym przez DSC PJATK – drugim etapie eliminacyjnym w procesie rekrutacji. W ramach tego przedsięwzięcia macie okazję zaprezentować swoje umiejętności analizy danych oraz modelowania predykcyjnego, pracując na rzeczywistych danych pochodzących z ofert sprzedaży samochodów.

Harmonogram

- **Rozpoczęcie zadania:** 12 marca 2025 r., godz. 08:00 na Kaggle
- **Zamknięcie zadania:** 26 marca 2025 r., godz. 23:59 (zgłoszenia na Kaggle oraz mailowo)
- **Ogłoszenie wyników:** Najpóźniej do 30 marca 2025 r.

Cel Konkursu

Celem konkursu jest opracowanie modelu predykcyjnego umożliwiającego precyzyjne przewidywanie cen pojazdów na podstawie szeregu zmiennych opisowych, takich jak marka, model, rok produkcji, przebieg, moc silnika czy emisja CO₂.

Wymagania

1. **Rozwiązanie Kaggle Competition**
 - Każdy zespół ma za zadanie opracować model predykcyjny przewidujący cenę pojazdu.
 - Szczegóły zasady oceniania oraz informacje o dataset'cie, znajdują się na stronie konkursu: [Kaggle Competition](#).
2. **Repozytorium GitHub**
 - Należy utworzyć repozytorium na **GitHub** zawierające cały kod, opis rozwiązania i prezentację.
 - Kod powinien być czytelny, dobrze udokumentowany i zgodny z dobrymi praktykami.
 - Mile widziane są notatniki Jupyter z analizą oraz komentarzami.
3. **Prezentacja wyników**
 - Na koniec zadania każda drużyna musi przygotować prezentację zawierającą:
 - i. **Minimum 2 wizualizacje** z danych, które opowiadają jakąś historię,
 - ii. **Opis procesu** analizy i modelowania,
 - iii. **Największe wyzwania** i jak sobie z nimi poradziliście.
 - Prezentacja może być przygotowana w formie klasycznej (np. w PowerPoint) lub jako notatnik Jupyter Notebook.
 - Jeśli prezentacja jest w formacie Jupyter Notebook, powinna być podzielona na dwa osobne pliki:
 - i. **EDA + wnioski** (Eksploracyjna analiza danych i kluczowe obserwacje)
 - ii. **Modelowanie + wnioski** (Opis użytych metod modelowania i ich efektywność)
4. **Zgłoszenie mailowe**
 - Każdy zespół musi przesłać do **26 marca 2025 do godz. 23:59** maila na dsc@pjwstk.edu.pl z następującymi informacjami:
 - i. **Link do repozytorium GitHub,**
 - ii. **Numer grupy i skład zespołu,**
 - iii. **Załączoną prezentację** (w przypadku prezentacji w formacie PDF lub PPTX)

Ocena zespołów

- **4 najlepsze zespoły** (zgodnie z wynikami na Kaggle) **dostają się do DSC PJATK automatycznie.**¹
- **Zespoły poniżej 4. miejsca** będą oceniane na podstawie:
 - Wyników modelu,
 - Jakości kodu i organizacji repozytorium,
 - Kreatywności w podejściu do problemu,
 - Użyteczności i jakości przetwarzania danych.

Zasady fair play

1. Możecie korzystać z dowolnych źródeł i narzędzi do pracy z danymi i modelem.
2. **Nie wolno kopiować kodu od innych zespołów** – jeśli jakimś cudem uzyskacie dostęp do ich kodu, nie możecie się nim “inspirować”.
3. Wszelkie podejrzenia o plagiat będą skutkowały dyskwalifikacją zespołu.
4. Każda drużyna ma ustalony **limit pięciu dziennych zgłoszeń**, każda drużyna musi się do niego stosować. Wszelkie próby obejścia limitu (np. tworzenie wielu kont) będą skutkować dyskwalifikacją.
5. Wszelkie próby sztucznego podbijania oceny, takie jak używanie bezpośrednio ukrytych wartości testowych, mogą skutkować obniżeniem oceny lub całkowitym pominięciem wyników zespołu.

Motywacja dla uczestników

To zadanie ma na celu nie tylko sprawdzenie Waszych umiejętności, ale także stworzenie wartościowego projektu, który może stać się częścią Waszego portfolio. Repozytorium GitHub, które przygotujecie, może być świetnym przykładem Waszej pracy, który będziecie mogli pokazać w przyszłych rekrutacjach lub aplikacjach na staże i praktyki.

Nie przejmujcie się, jeśli napotkacie trudności – każda analiza, nawet ta mniej skuteczna, przynosi nową wiedzę. Podejdźcie do zadania kreatywnie, eksperymentujcie i bawcie się danymi! 🎯

Powodzenia i do zobaczenia na leaderboardze! 🚀

¹ Jeżeli w danych pojawi się błąd, który wpłynie na ocenę zadania, organizatorzy zastrzegają sobie prawo do wycofania zasady automatycznego przyjęcia 4 najlepszych zespołów do koła DSC PJATK.