

Projekt DL – JDSZR6

Założenia:

1. Wybieramy temat. Może być to temat z poprzedniego projektu – o ile nie będzie realizować tego samego lub podobnego projektu trzeci raz.
2. Wyznaczamy zmienną celu.
3. Dokonujemy eksploracji zbioru. Uzupełniamy braki danych. Dokonujemy wstępnej selekcji zmiennych.
4. Dzielimy zbiór danych na treningowy testowy i w miarę możliwości na walidacyjny.
5. Tworzymy różnego rodzaju modele sieci neuronowych – sprawdzając różne możliwe parametry.
6. Porównujemy i wybieramy najlepszy model.
7. Dokonujemy ostatecznej oceny wybranego modelu na zbiorze walidacyjnym.
8. Porównujemy rozwiązanie z przynajmniej jednym, poznanym modelem w bloku ML.
9. **Wskazujemy uzysk dla firmy / odbiorcy (np. ile pieniędzy z tego będzie miała, jak wdrożenie wpłynie na rozpoznawalność przypadków, dla których firma wykona działania prewencyjne itd.) – to jest bardzo ważny punkt.**
10. Budujemy aplikację wykorzystującą model.
11. Na prezentacji powinna znaleźć się krótka analiza zmiennych, omówienie wyników wszystkich zastosowanych technik modelowania i ich porównanie, wybór modelu i przede wszystkim wartość dodana jaką niesie za sobą wdrożenie.

Tematy:

A. Tematy z poprzedniego projektu:

1. Prognoza prądu
 - Prognozowanie zapotrzebowania na energię elektryczną w ramach gospodarstwa domowego
 - Aplikacja szacująca opłaty za prąd
 - <https://www.kaggle.com/uciml/electric-power-consumption-data-set>
2. Szczęście Klienta
 - Prognozowanie czy Klient banku jest zadowolony z usług
 - Aplikacja dla pracowników informująca o zagrożeniu niezadowolenia Klienta
 - Uwaga: Dane pochodzą z zakończonego już konkursu na Kaggle, do pobrania danych jest wymagane zatwierdzenie warunków Banku, który te dane wystawił.
 - <https://www.kaggle.com/c/santander-customer-satisfaction/overview>
3. Odejdzie czy nie odejdzie? – prognozowanie zjawiska churn wśród Klientów telecom
 - Prognozowanie czy jest zagrożenie, że Klient zrezygnuje z usług firmy
 - Aplikacja dla pracowników informująca o zagrożeniu odejścia Klienta
 - <https://www.kaggle.com/blatchar/telco-customer-churn>
4. Czas choroby
 - Estymacja czasu pobytu w szpitalu
 - Aplikacja szacująca obciążenie szpitala
 - <https://www.kaggle.com/nehaprabhavalkar/av-healthcare-analytics-ii?>
5. Śmierć pacjenta

- Estymacja niewydolności serca
- Aplikacja pokazująca lekarzowi prawdopodobieństwo zdarzenia
- <https://www.kaggle.com/andrewmvd/heart-failure-clinical-data>

B. Nowe tematy:

6. Greenpeace

- Rozpoznawanie ptaków
- Szukanie gatunków zagrożonych wyginięciem
- typ zadania: klasyfikacja zdjęć
- <https://www.kaggle.com/gpiosenska/100-bird-species>

7. Sortownia owoców

- Klasyfikacja jabłek
- aplikacja na taśmie w fabryce dżemów
- typ zadania: klasyfikacja zdjęć
- wersja podstawowa: trzy klasy silnie odróżniające się od siebie
- wersja zaawansowana: wszystkie klasy
- <https://www.kaggle.com/chrisfilo/fruit-recognition>

8. W końcu inny mnist

- rozpoznawanie gestów
- aplikacja działająca real-time pozwalająca zrozumieć co chce przekazać niema osoba
- typ zadania: klasyfikacja zdjęć
- <https://www.kaggle.com/datamunge/sign-language-mnist>

9. Chłopiec czy dziewczynka?

- Rozpoznanie płci na podstawie zdjęcia
- <https://www.kaggle.com/nipunarora8/age-gender-and-ethnicity-face-data-csv>

10. Prognozowanie sprzedaży makro

- Model szeregów czasowych prognozujący sprzedaż na podstawie zmiennych makroekonomicznych.
- Stworzenie aplikacji prognozującej sprzedaż na kolejny rok lub aplikacji szacującej sprzedaż na zadanych przez użytkownika parametrach.
- <https://www.kaggle.com/datasets/yasserh/walmart-dataset>

11. Prognozowanie sprzedaży mikro

- Model szeregów czasowych prognozujący sprzedaż na podstawie zmiennych makroekonomicznych.
- Stworzenie aplikacji prognozującej sprzedaż na kolejny rok lub aplikacji szacującej zysk ze sprzedaży lub sugerującą liczbę promocji w celu uzyskania oczekiwanego zysku.
- <https://www.kaggle.com/competitions/favorita-grocery-sales-forecasting/data>

12. Czy ta firma ma sens?

- Model prognozujący bankructwo firmy
- Ocena sensowności inwestowania w firmę.
- <https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction>