

Raport walidacyjny

Jan Cwalina
Bartłomiej Borycki

Kwiecień 2024

1 Wstęp

Naszym zadaniem jako walidatorów tematu Bank Marketing Classification Task wykonywanego przez Bartosza Jezierskiego i Michała Iwaniuka, było finalne przetestowanie zaproponowanych przez nich modeli oraz kontrolowanie ich pracy w trakcie całego projektu.

2 Proces walidacyjny

Po ukończeniu każdego z etapów przez grupę modelarzy dostawaliśmy do wglądu ich postęp w pracach i zwracaliśmy im naszą ocenę i ewentualne sugestie co należałoby poprawić.

2.1 Dane walidacyjne

Na samym początku prac otrzymaliśmy walidacyjny zbiór danych składający się z 30% oryginalnego datasetu. Grupa modelarzy nie miała dostępu do tych danych podczas tworzenia modelu, natomiast naszym zadaniem było wykorzystanie tych danych do ostatecznej oceny modelu.

2.2 EDA

Podczas tego etapu grupa modelarzy świetnie sobie poradziła i nie mieliśmy do niej żadnych zastrzeżeń. Spojrzeli na dane z każdej możliwej strony, wykonali wszystkie wymagane kroki na tym etapie projektu oraz wyciągnęli poprawne i logiczne wnioski.

2.3 Feature engineering

W tym etapie również wykonano wszystkie potrzebne kroki. Mieliśmy jednak kilka sugestii, które zostały wcielone w życie przez grupę modelarzy. Poradziliśmy zastosowanie crosswalidacji już na tym etapie oraz kilka sposobów transformacji zmiennych takich jak np na okrąg.

2.4 Hiperparametry

W tym etapie również wszystko co dostaliśmy było poprawne oraz logiczne. Dodatkowo zaproponowano bardzo przejrzysty sposób porównywania modeli. Po otrzymaniu wstępnej wersji kodu zaproponowaliśmy modelarzom oraz wypróbowanie modeli typu stacked i soft.

3 Wyniki modelu

Po otrzymaniu wytrenowanego na danych projektowych najlepszego modelu, którym był model typu soft, przeprowadziliśmy jego ostateczny test na zbiorze walidacyjnym. Kod odpalił się bezproblemowo. Testowaliśmy wyniki modelu tylko w metrykach które zostały wybrane przez grupę modelarzy.

Wyniki modelu na zbiorze walidacyjnym:

- precision: 0.490
- gmean: 0.758
- fbeta: 0.512
- balanced accuracy: 0.771
- recall: 0.628

Wyniki te są zgodne z deklaracjami grupy modelującej i mieszczą się w przekazanych nam zakresach.

4 Podsumowanie

Podsumowując, rozwiązanie zaproponowane przez grupę modelarzy działa poprawnie na danych które nie był wykorzystywane do jego stworzenia. Uważamy, że z danych które mieli do dyspozycji, wyciągnęli wszystko co się dało.