

Unbiased kernel shap

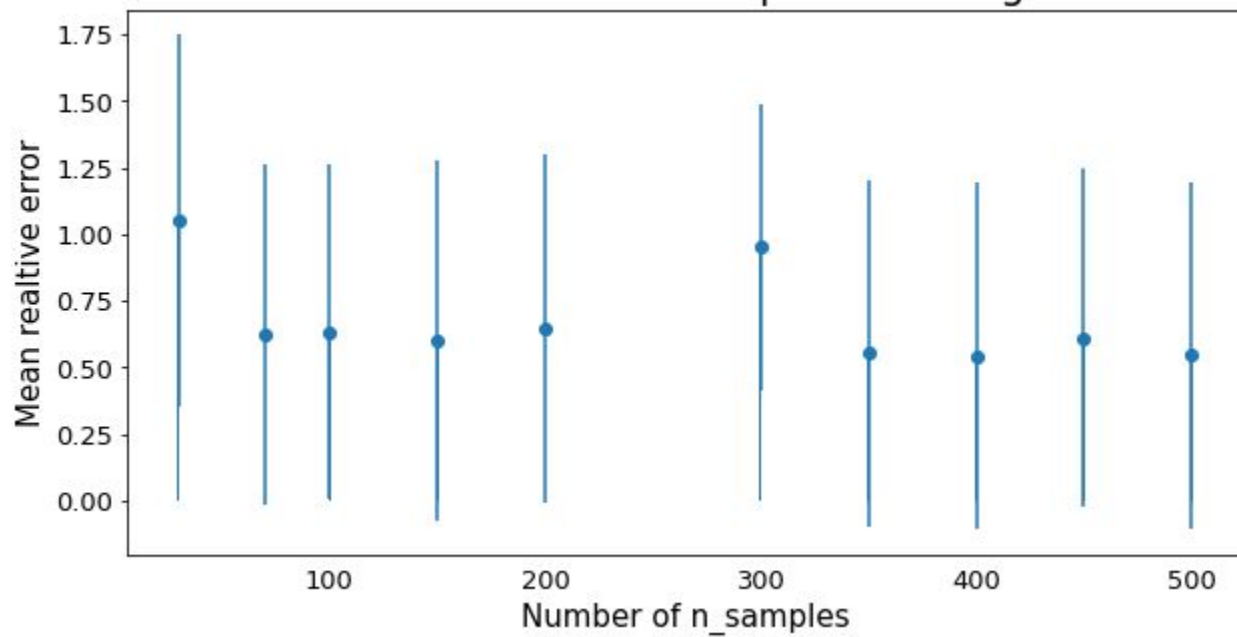
Michał Kucharczyk i Edward Sucharda

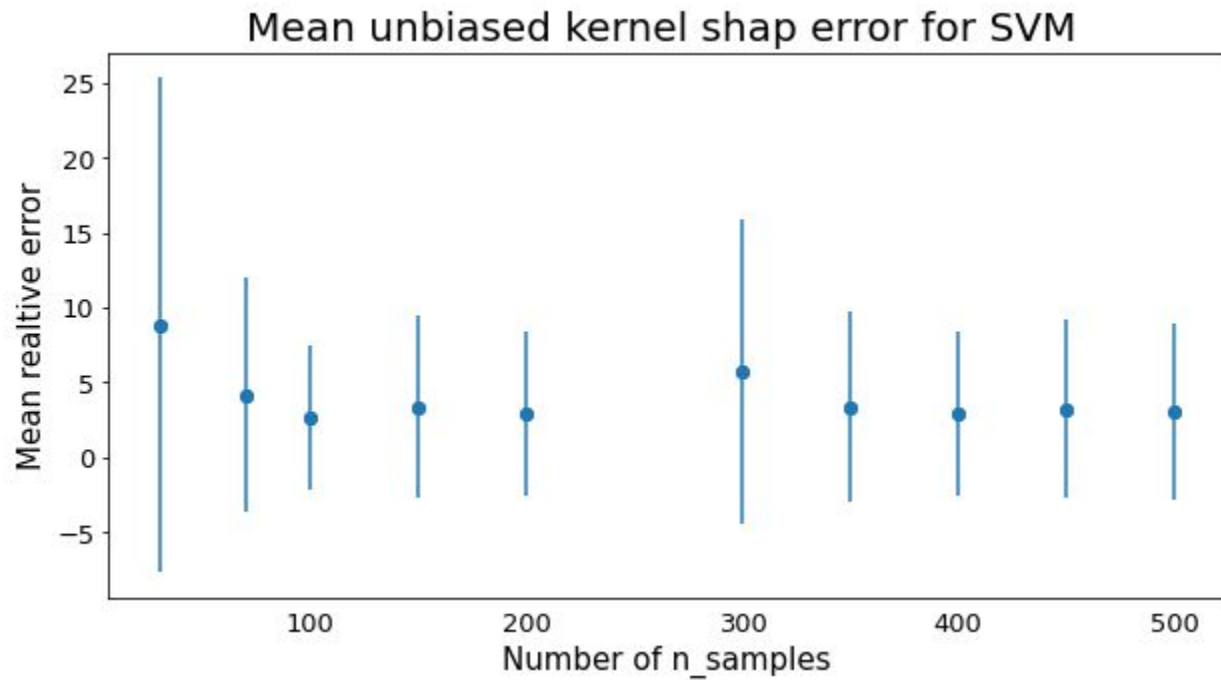
Test poprawności i nieobciążenia modelu

- Zbiór danych: 9 cech mieszkania określających jego cenę
- Modele: SVM, xgboost
- 500 pomiarów
- 10 różnych wartości `n_samples`
- 5 próbek badanych

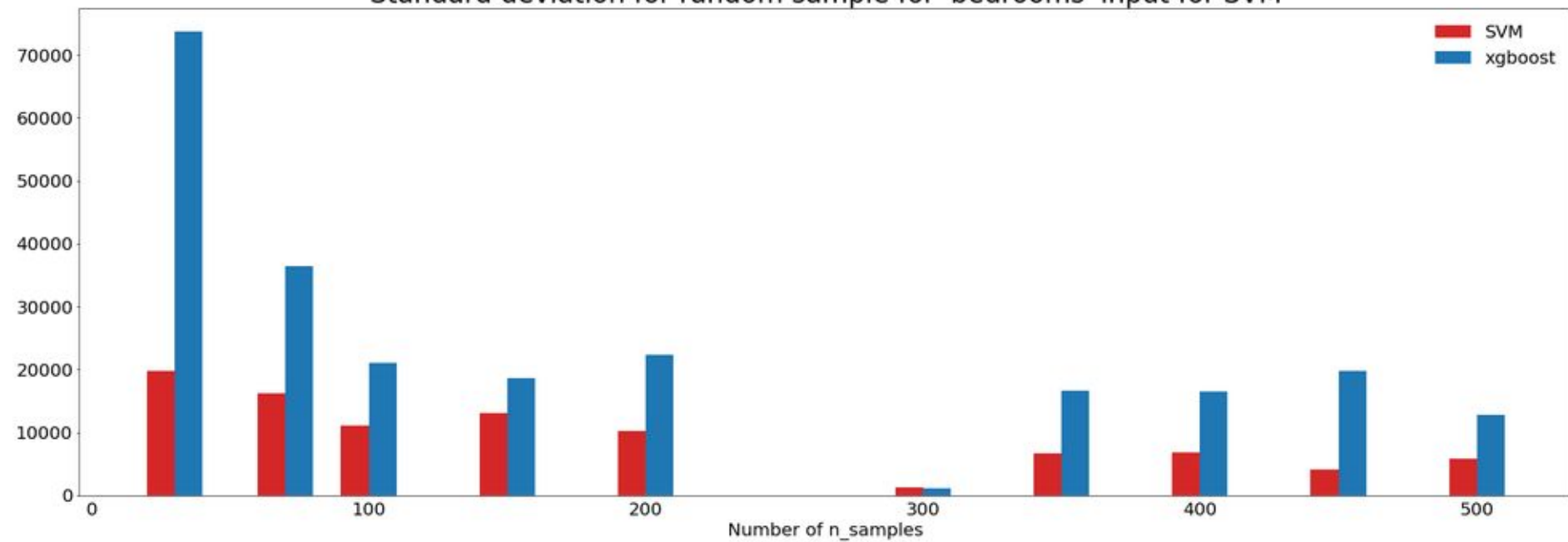


Mean unbiased kernel shap error for xgboost





Standard deviation for random sample for 'bedrooms' input for SVM

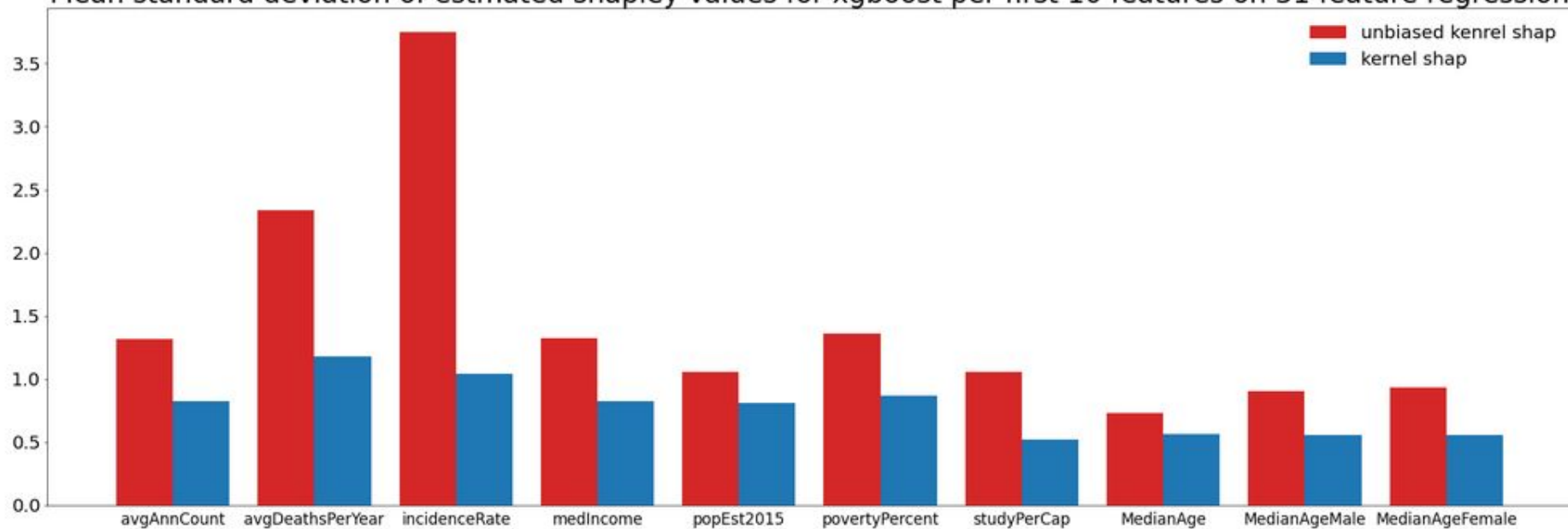


Porównanie jakości z Kernel Shap

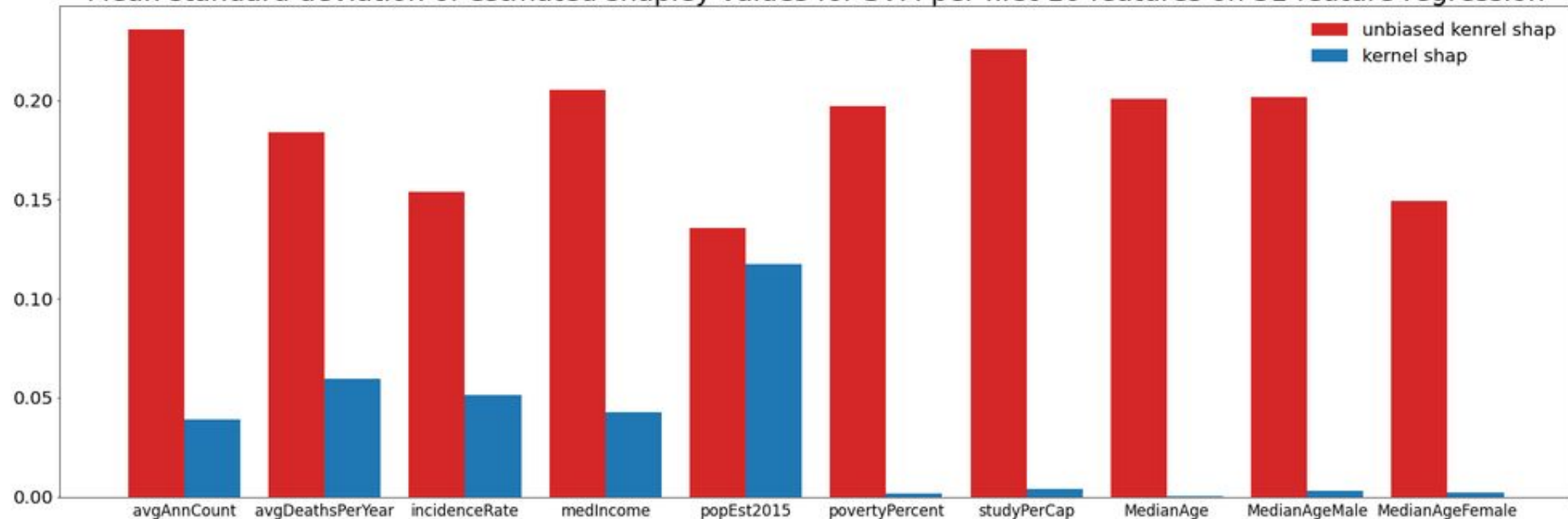
- Dwa zbiory danych:
 - 9 cech mieszkania określających jego cenę
 - 31 cech określających współczynnik śmiertelności
- Dobór `n_samples` i liczby danych uczących tak by obie metody miały podobny czas wyjaśniania pojedynczej próbki
- 500 pomiarów próbki
- 20 próbek

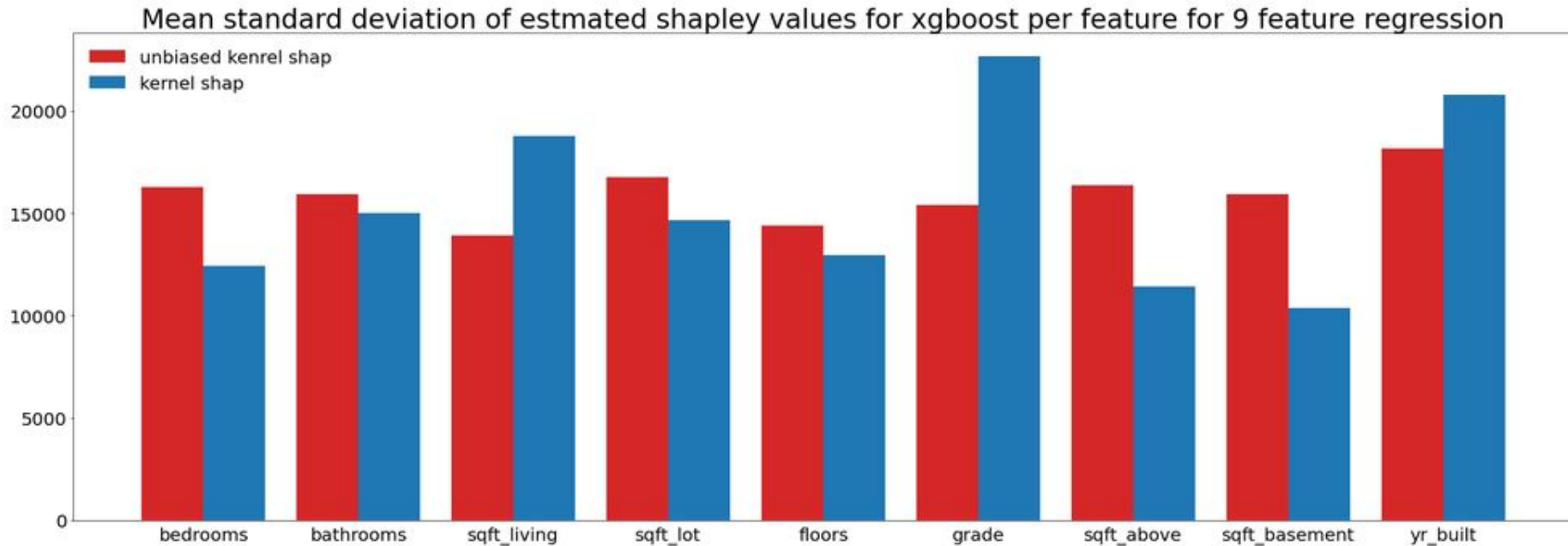


Mean standard deviation of estimated shapley values for xgboost per first 10 features on 31 feature regression



Mean standard deviation of estimated shapley values for SVM per first 10 features on 31 feature regression





*dla SVM kernel shap był dużo lepszy i wykresy nie oddawały skali różnicy