

Metody walidacji modeli statystycznych

Dariusz Majerek

28.01.2023

Table of contents

Wprowadzenie	3
1 Introduction	6
2 Summary	7
References	8

Wprowadzenie

Niniejsza książka została przygotowana z myślą o studentach kierunków *Inżynierii i Analizy Danych* oraz *Matematyka*. Ma ona posłużyć w nauce szeroko rozumianej walidacji modeli statystycznych. Pojęcie zostanie dosyć szeroko potraktowane, ponieważ w książce nawiązuję do metod próbkowania (ang. *resampling*), stosowanych do oceny jakości modelu oraz w procesie tuningowaniu modelu, jak również do miar stosowanych do oceny dopasowania modeli. Ponieważ książka ta ma stanowić przegląd metod używanych w ocenie jak dobrze model przewiduje nieznane wartości, to zostanie również uzupełniona o analizę wrażliwości pozwalającą na określenie jakie czynniki (predyktory) i w jaki sposób wpływają na taką a nie inną predykcję.

W statystyce walidacja modelu jest zadaniem polegającym na ocenie, czy wybrany model statystyczny jest odpowiedni, czy nie. Często we wnioskowaniu statystycznym wnioski z modeli, które wydają się pasować do danych, mogą być błędne, co powoduje, że badacze nie rozumieją rzeczywistej istotności swojego modelu. Aby temu zapobiec, stosuje się walidację modelu, aby sprawdzić, czy model statystyczny posiada tzw zdolność generalizacji. Owa zdolność polega na możliwości poprawnego przewidywania wartości wyjściowej nie tylko dla obserwacji ze zbioru uczącego, ale również z testowego, bez znaczącej straty jakości wspomnianej predykcji. Tego tematu nie należy mylić z blisko związanym zadaniem wyboru modelu, czyli procesem rozróżniania wielu modeli kandydujących: walidacja modelu nie dotyczy tak bardzo konceptualnej konstrukcji modeli, lecz testuje jedynie spójność pomiędzy wybranym modelem a jego deklarowanymi wynikami.

Walidacja modelu występuje w wielu odmianach, a konkretna metoda walidacji modelu stosowana przez badacza jest często uwarunkowana jego projektem badawczym. Oznacza to, że nie

A GOOD EXAMPLE OF



OVERFITTING

ma jednej uniwersalnej metody walidacji modelu. Na przykład, jeśli badacz pracuje z bardzo ograniczonym zestawem danych, ale ma silne założenia wstępne dotyczące danych, może rozważyć walidację dopasowania swojego modelu poprzez zastosowanie metod bayesowskich i testowanie dopasowania modelu przy użyciu różnych rozkładów wstępnych. Jeśli jednak badacz ma dużo danych i testuje wiele zagnieżdżonych modeli, warunki te mogą sprzyjać walidacji krzyżowej i ewentualnie testowi “leave one out”. Są to dwa abstrakcyjne przykłady i każda rzeczywista walidacja modelu będzie musiała rozważyć znacznie więcej zawiłości niż tu opisano, ale przykład ten ilustruje, że metody walidacji modelu zawsze będą miały charakter poszlakowy. Ogólnie rzecz biorąc, modele mogą być walidowane przy użyciu istniejących danych lub przy użyciu nowych danych, a obie metody są omówione w kolejnych rozdziałach.



1 Introduction

This is a book created from markdown and executable code.

See (**knuth84?**) for additional discussion of literate programming.

2 Summary

In summary, this book has no content whatsoever.

References