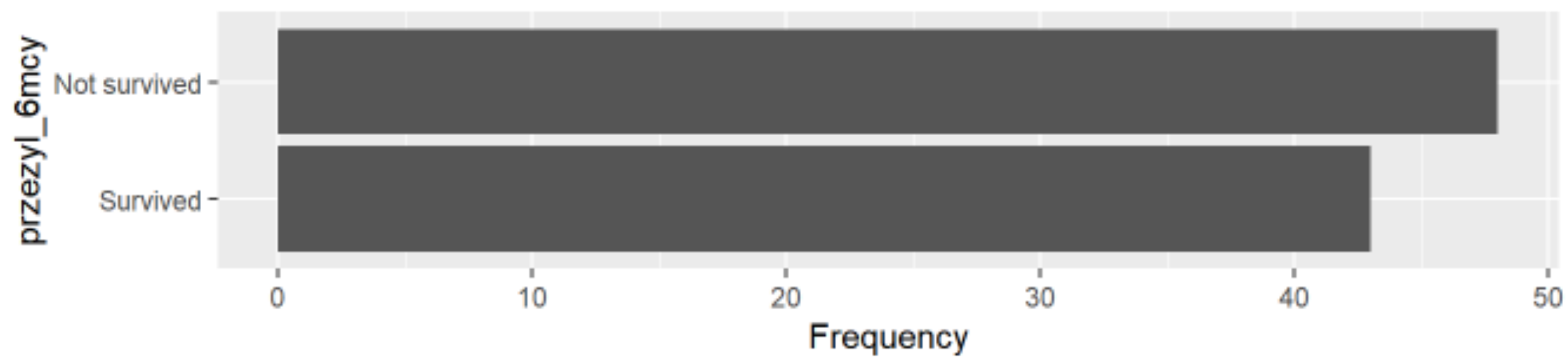
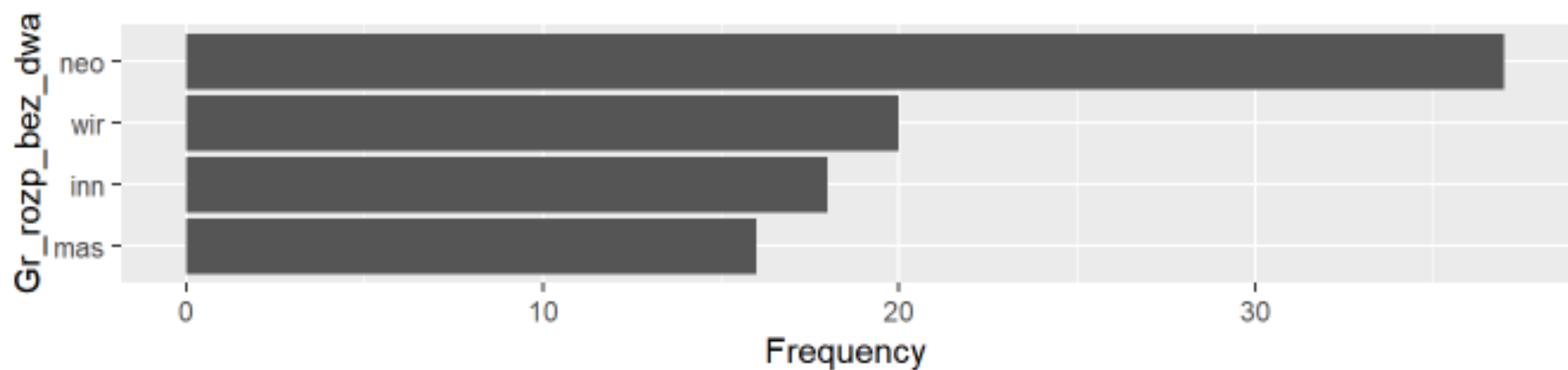
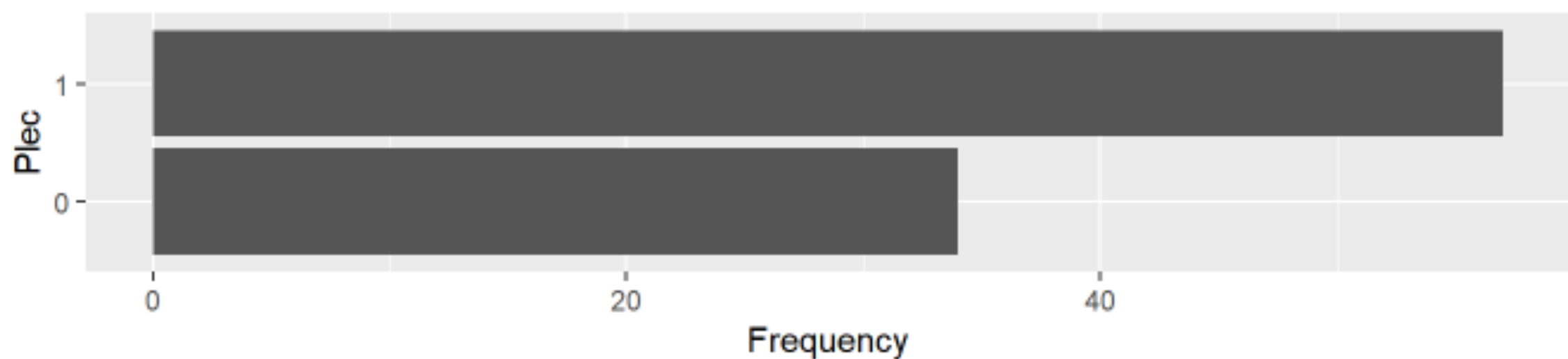


1. Case study: HLH
2. Praca domowa 1
3. Książka z projektami i bookdown



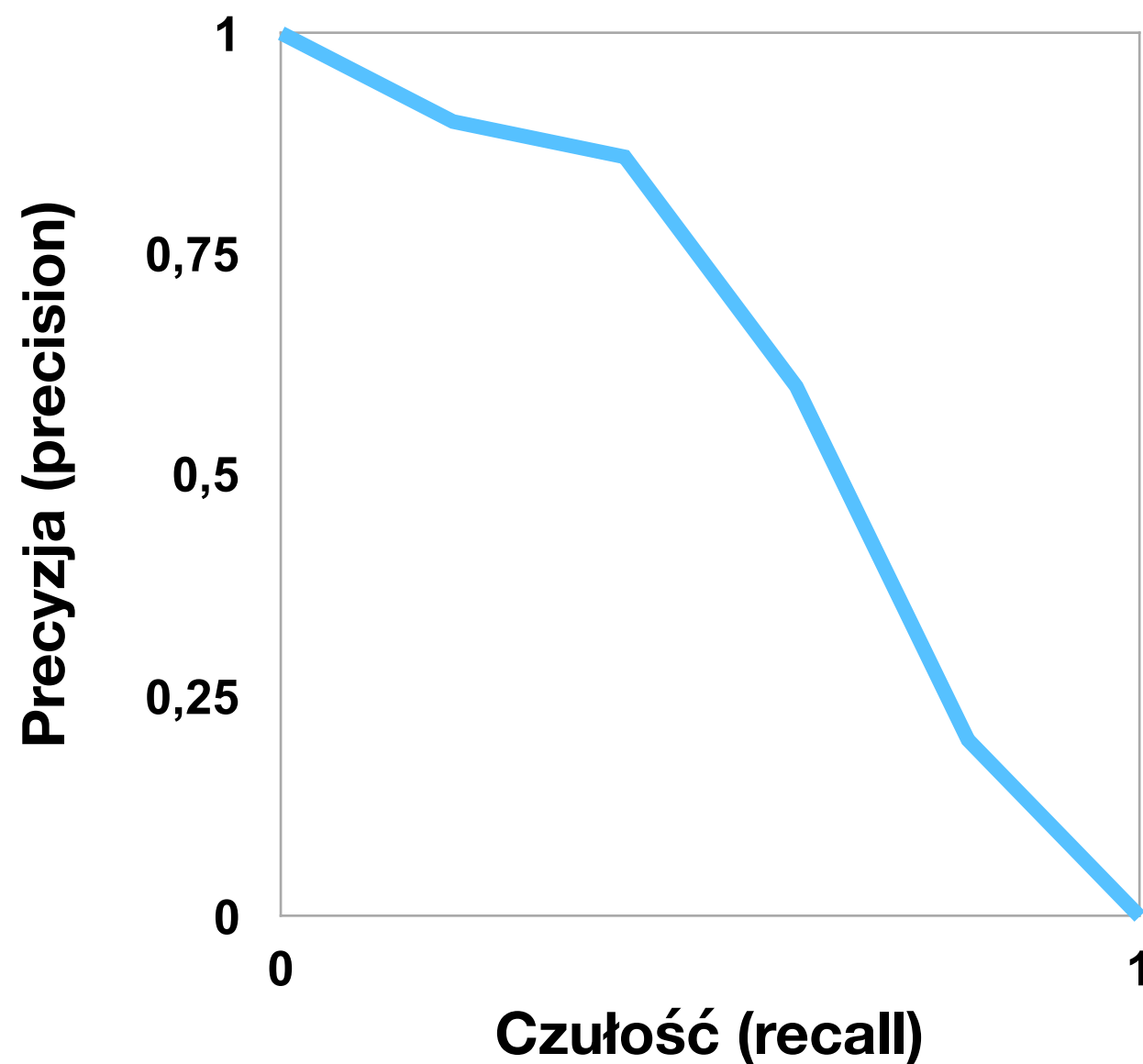
Problem niezbalansowanych klas

- Zbierz więcej danych
- Zmień metryki
- Dolosuj z klasy mniejszościowej
- Dobierz odpowiedni algorytm

Problem niezbalansowanych klas

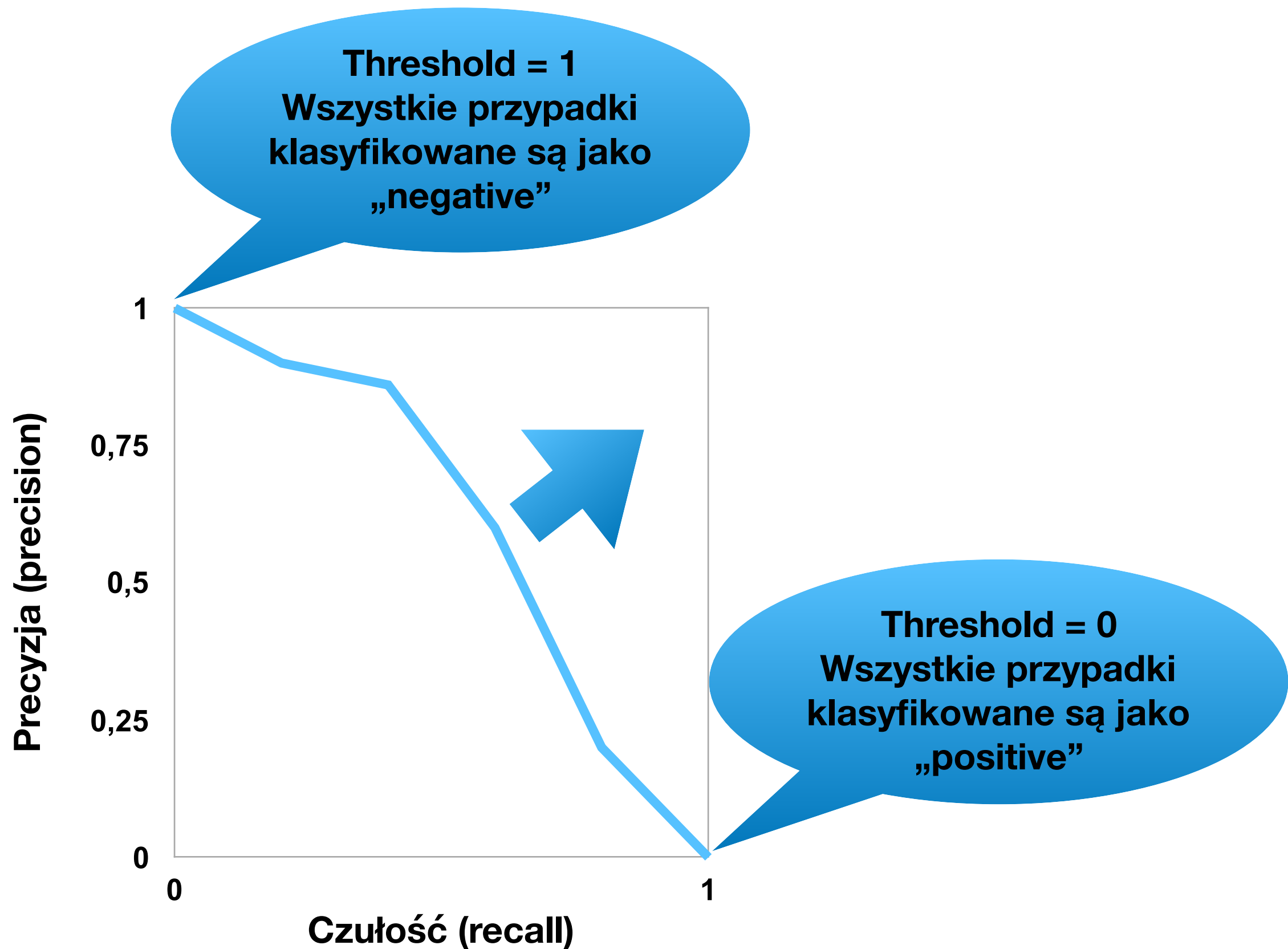
- Zbierz więcej danych
- Zmień metryki
- Dołosuj z klasy mniejszościowej
- Dobierz odpowiedni algorytm

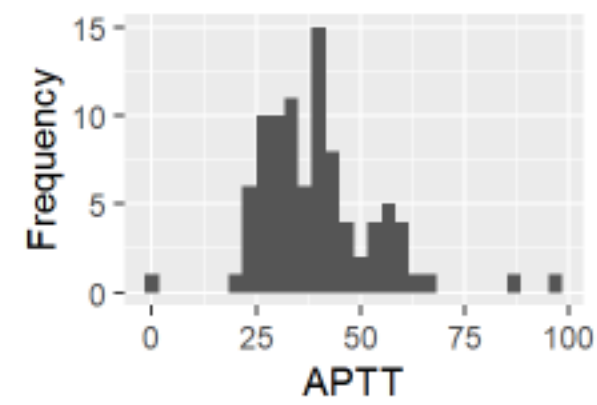
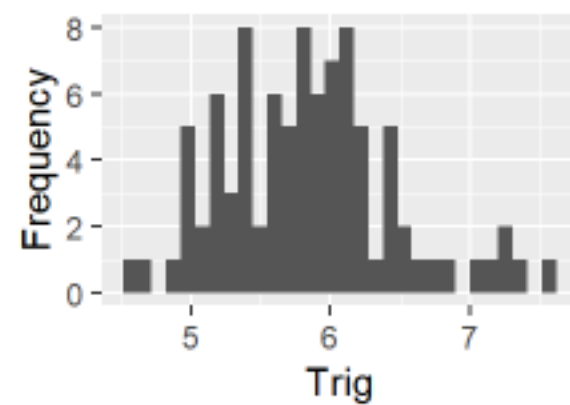
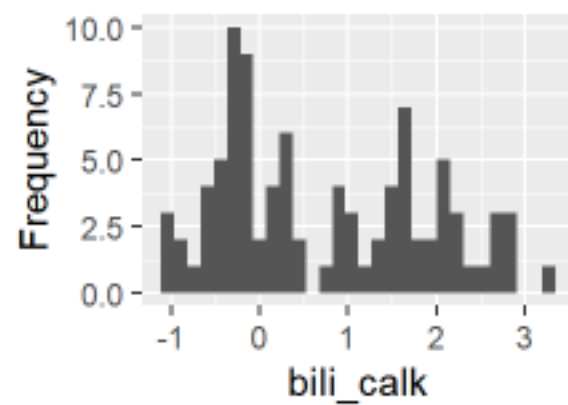
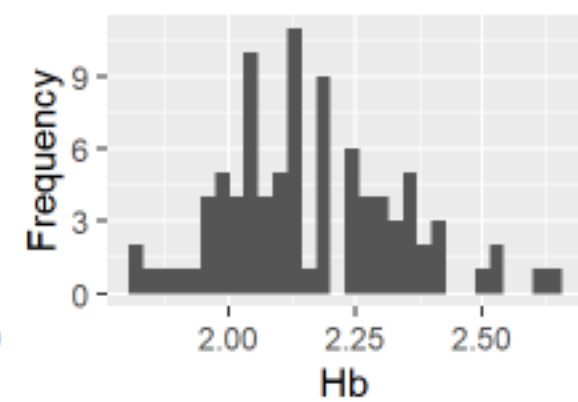
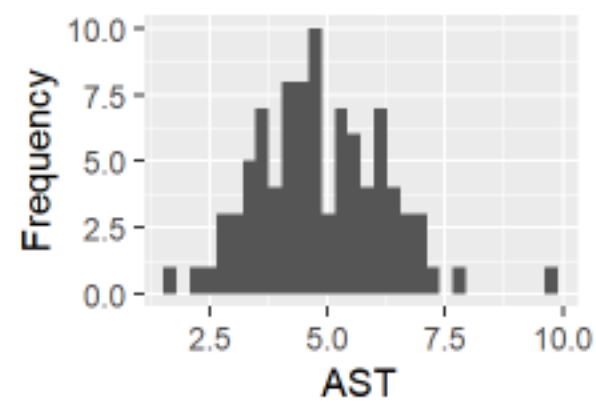
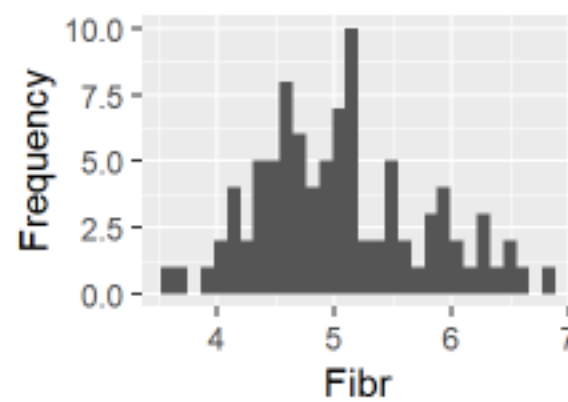
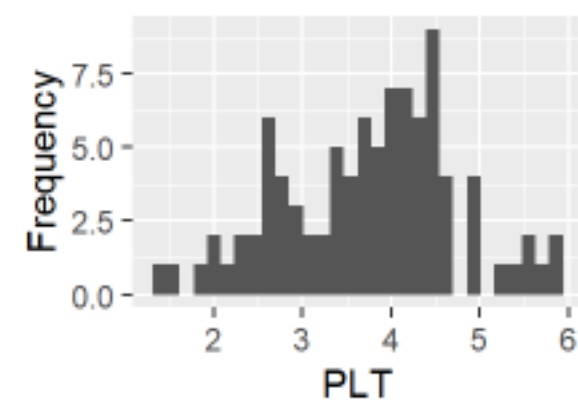
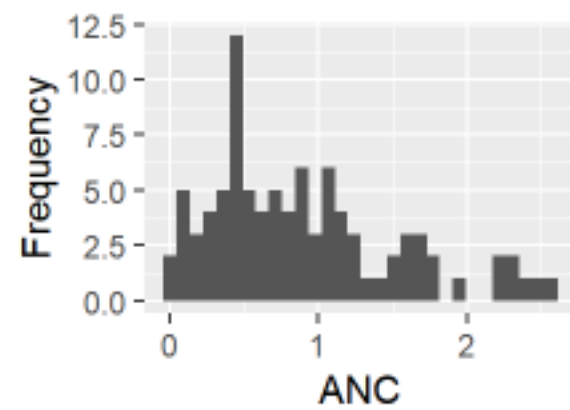
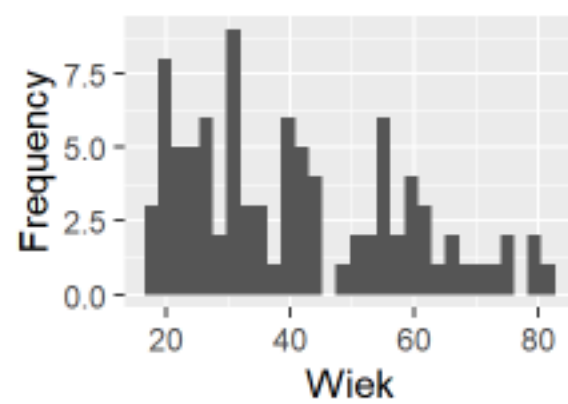
AUPRC

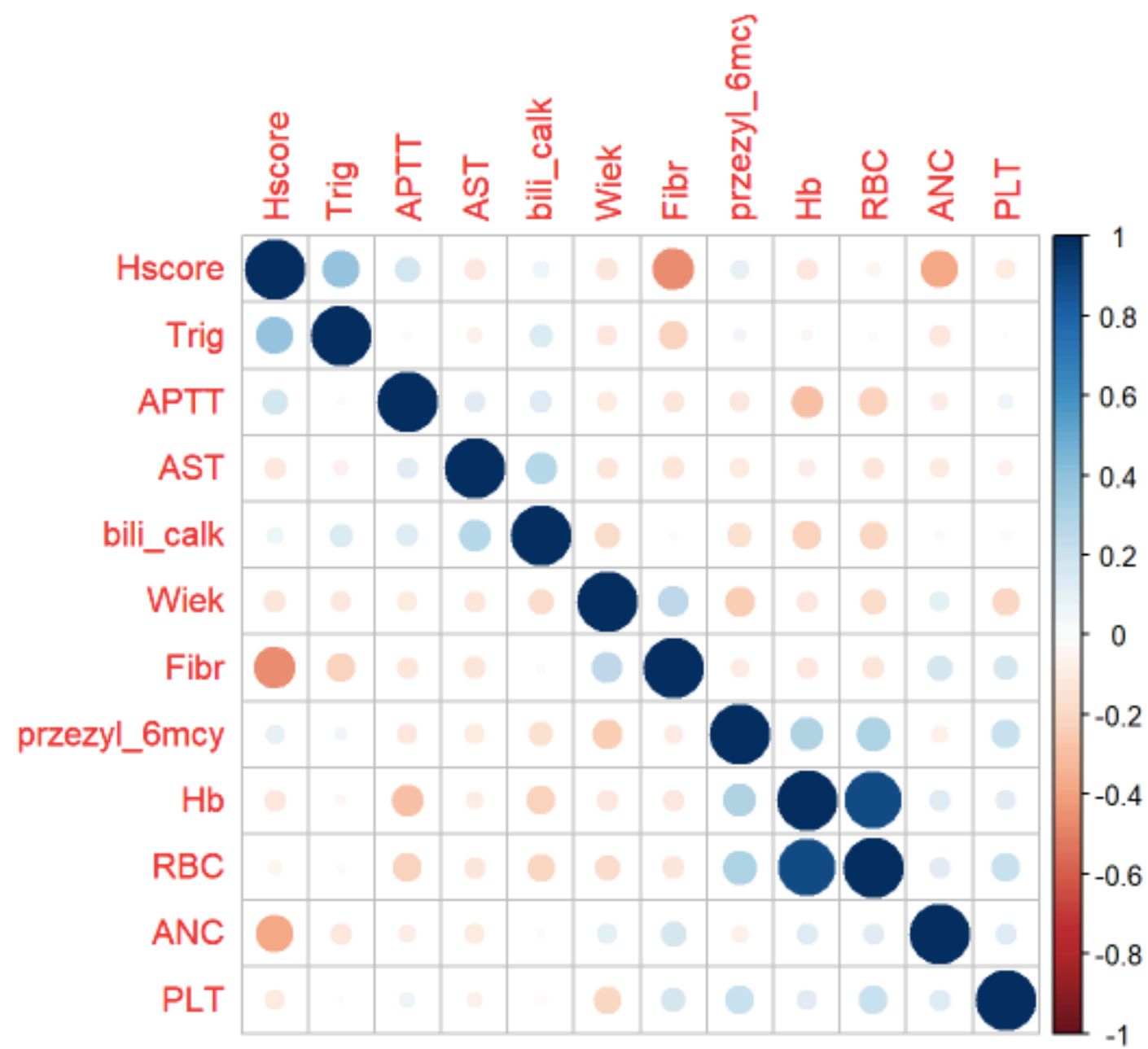


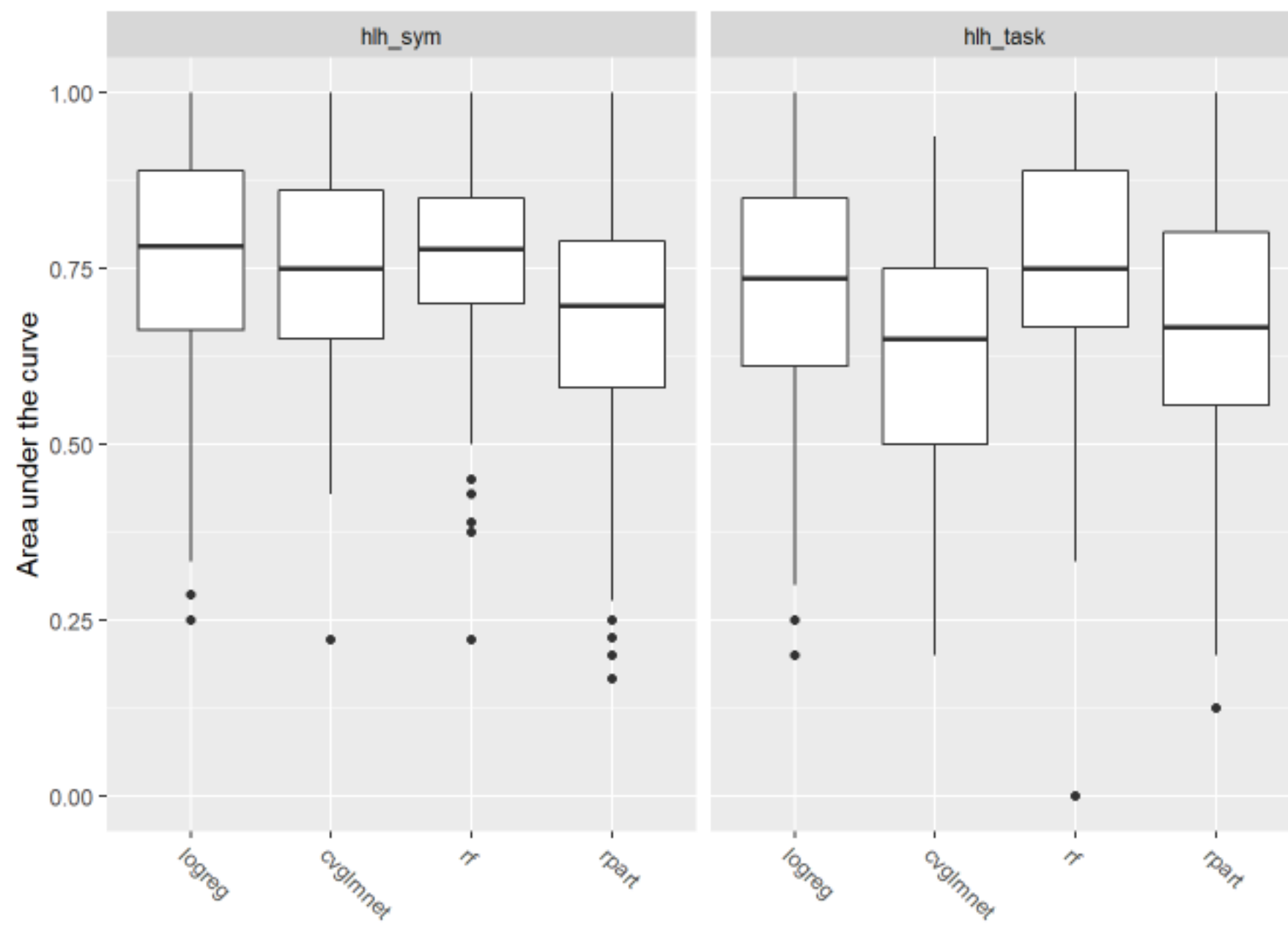
Precyzja = $TP/(TP+FP)$

Czułość = $TP/(TP+FN)$

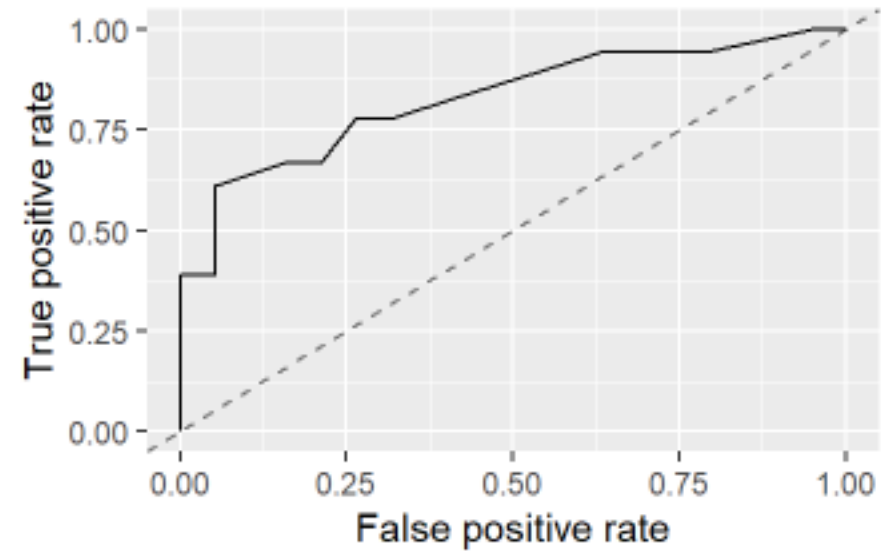




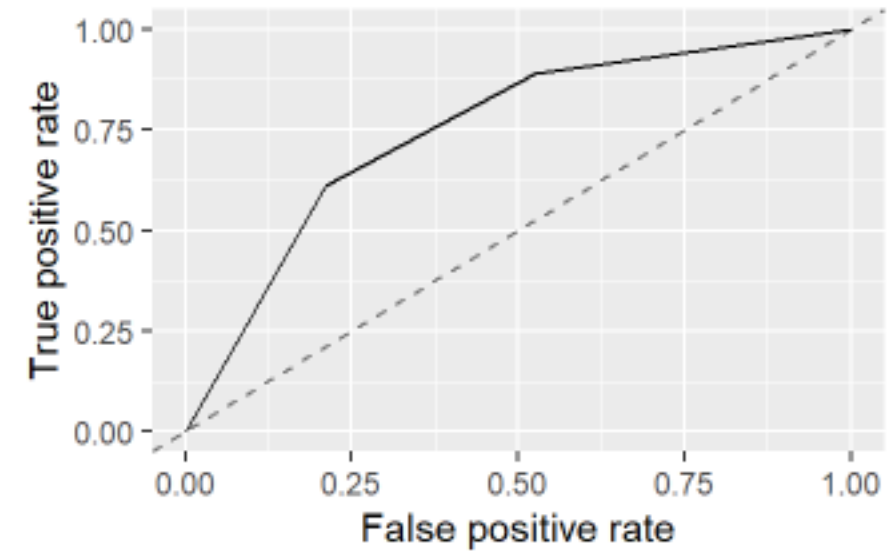




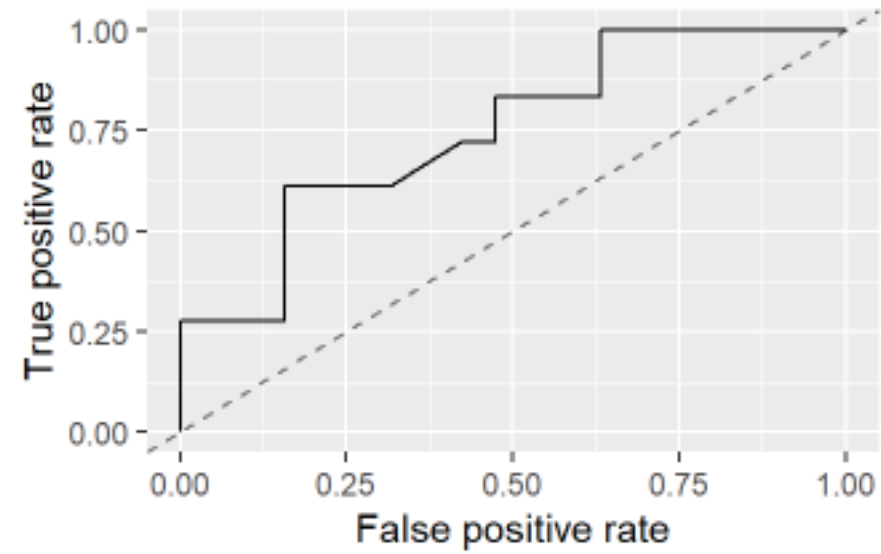
Lasso



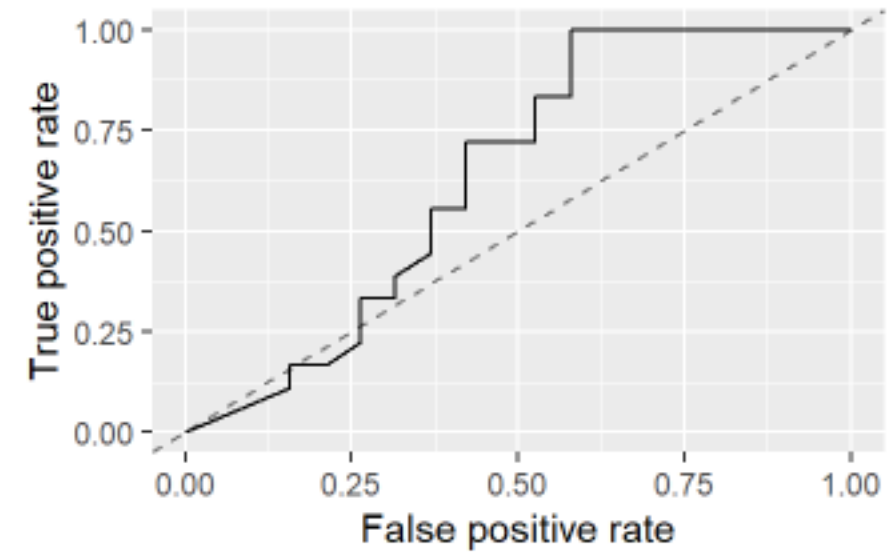
Decision tree



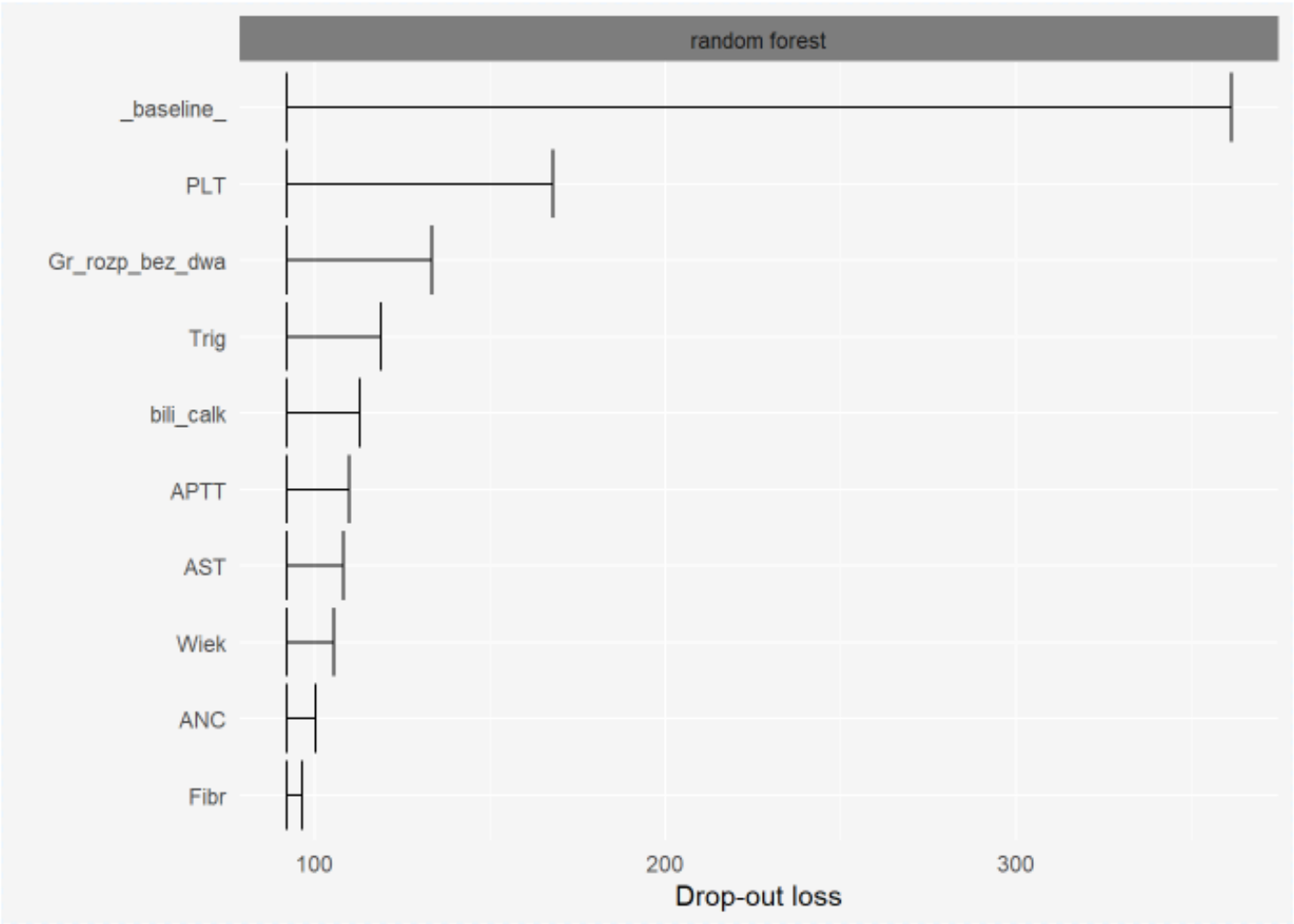
Random forest

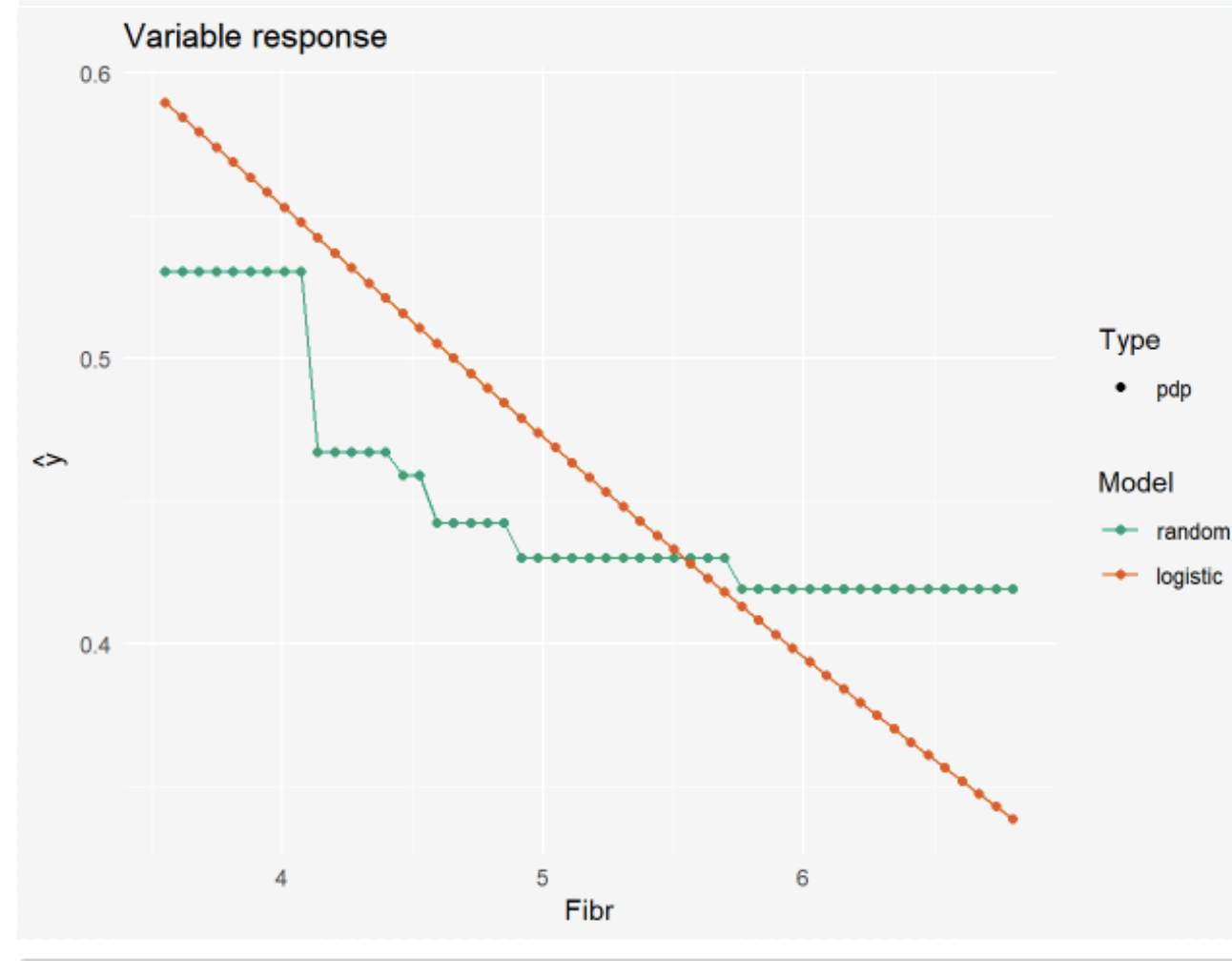
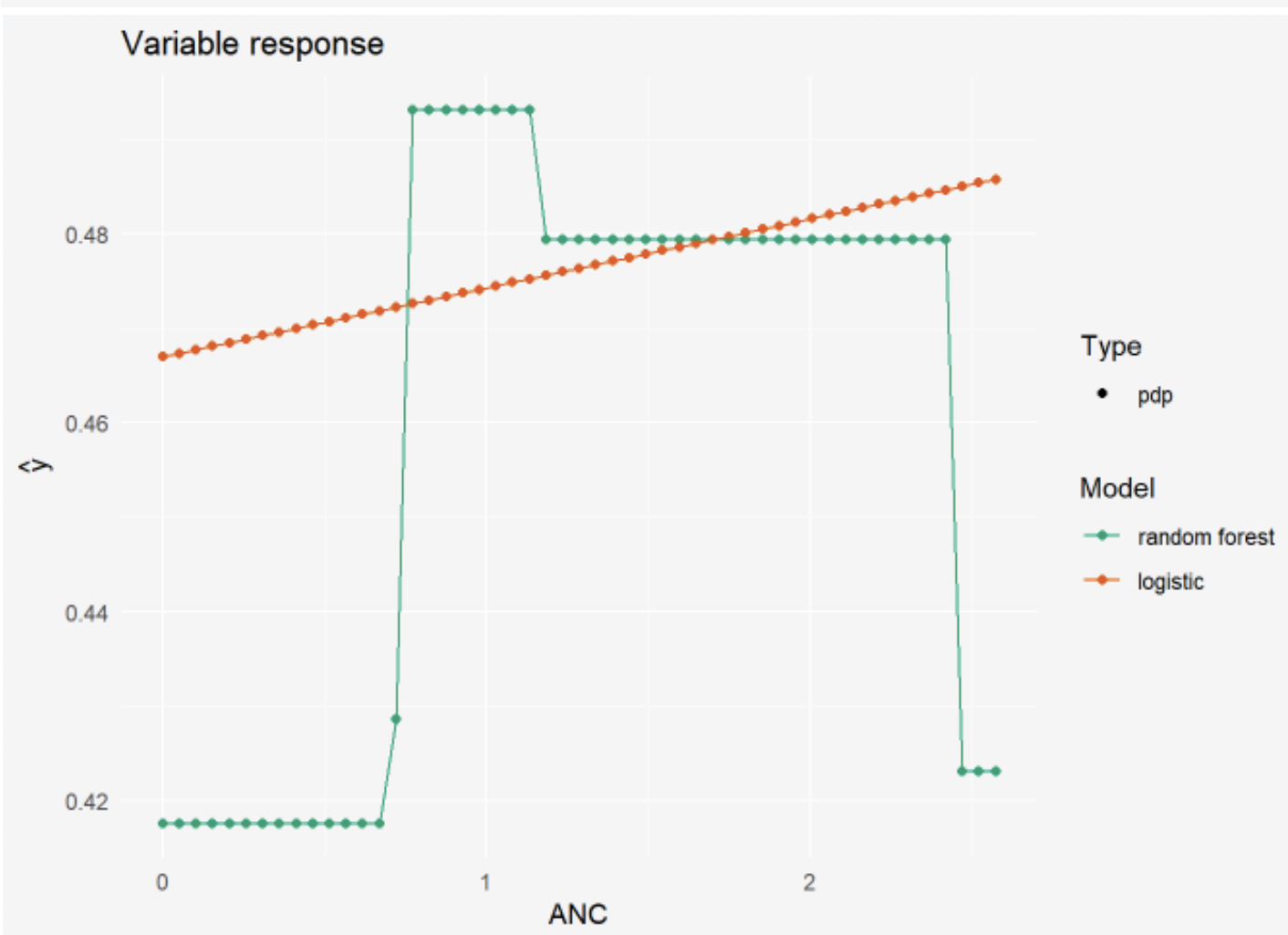
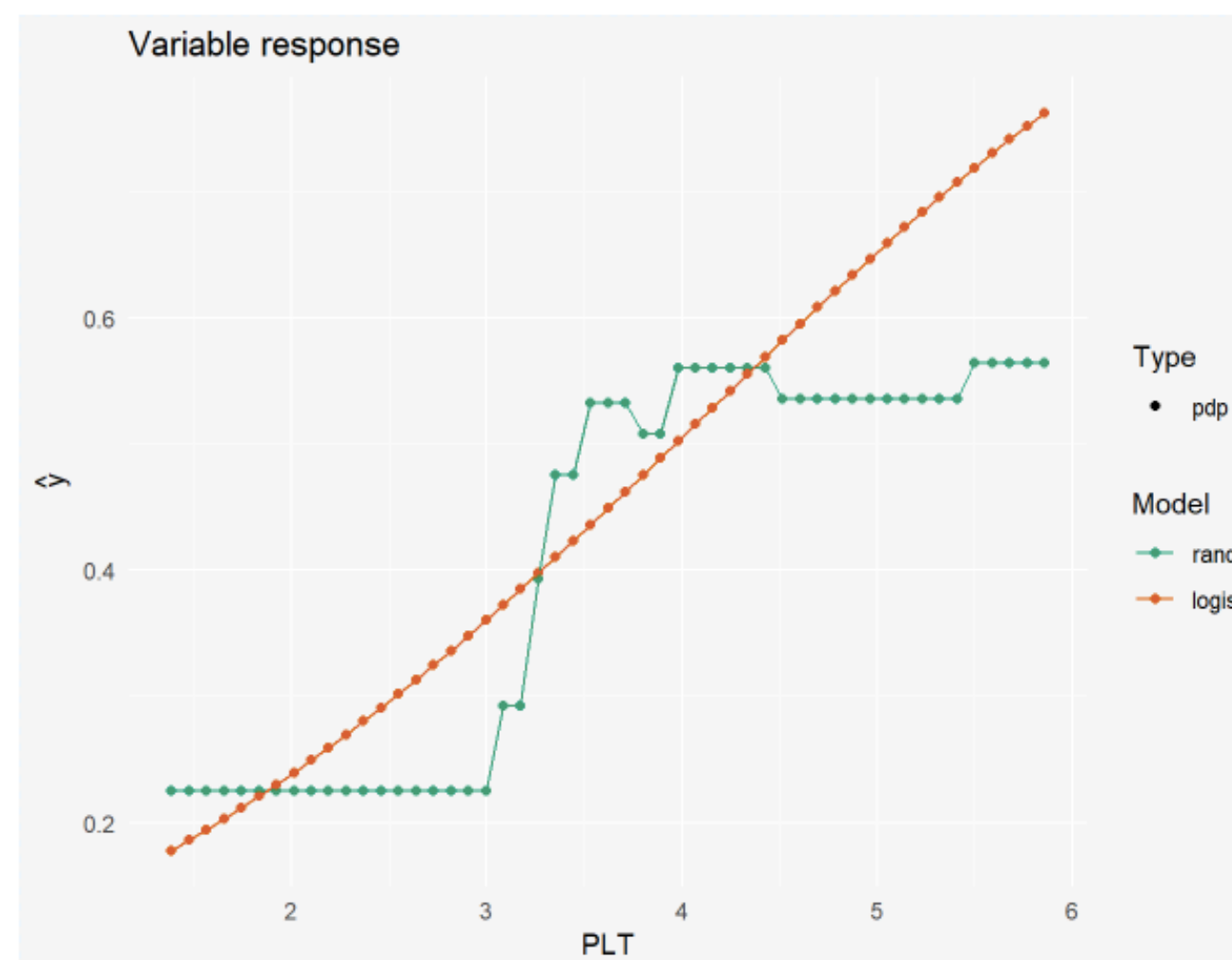
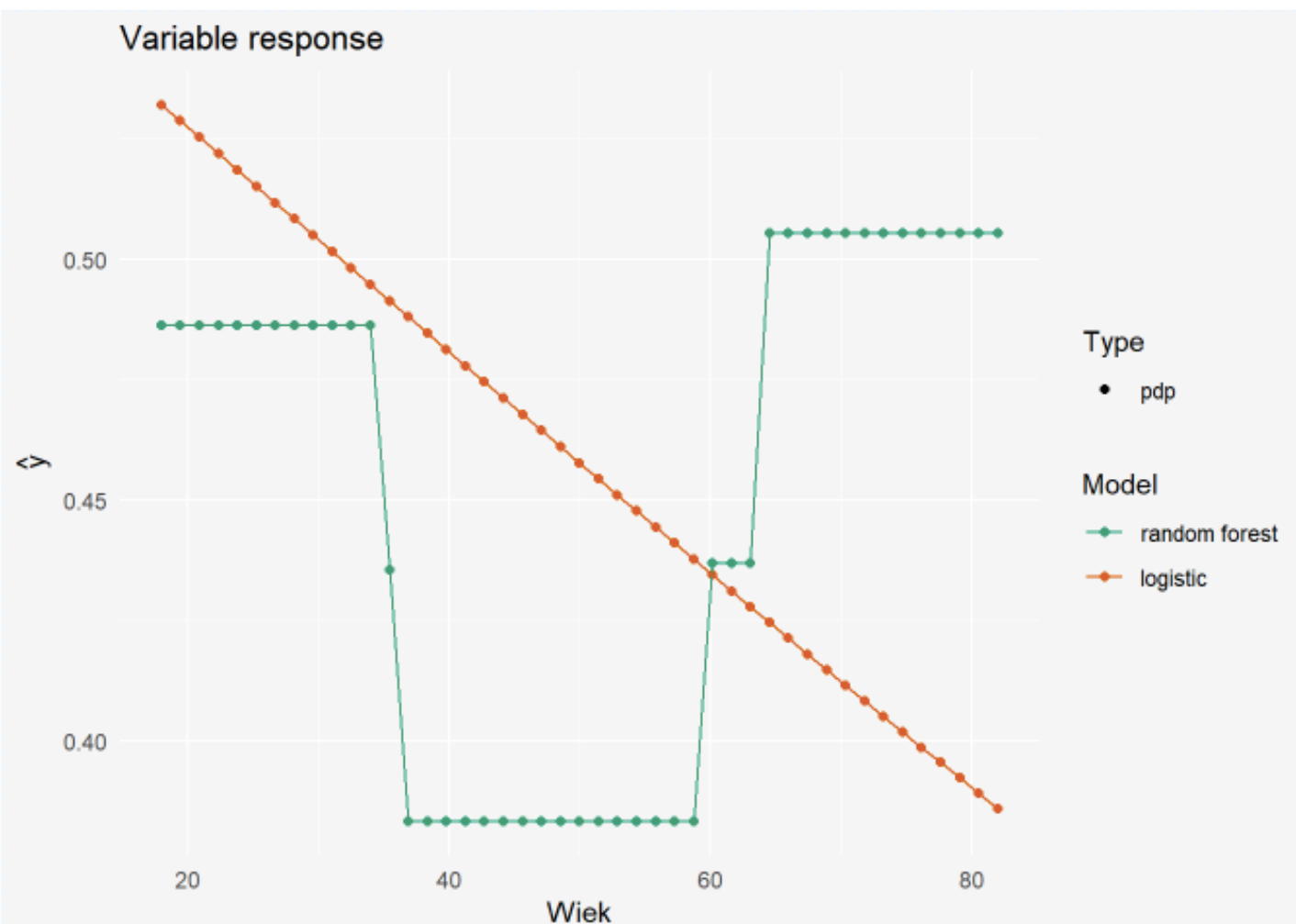


xgBoost



term	estimate	std.error	statistic	p.value
(Intercept)	-0.945	1.035	-0.913	0.364
Wiek	-0.004	0.003	-1.276	0.206
Plec1	-0.042	0.099	-0.429	0.669
Gr_rozp_bez_dwamas	0.635	0.150	4.223	0.000
Gr_rozp_bez_dwaneo	0.301	0.126	2.387	0.019
Gr_rozp_bez_dwawir	0.265	0.143	1.857	0.067
ANC	-0.083	0.075	-1.107	0.272
PLT	0.159	0.049	3.235	0.002
Fibr	0.004	0.075	0.057	0.955
AST	0.008	0.040	0.192	0.849
Hb	0.284	0.287	0.991	0.325
bili_calk	-0.070	0.043	-1.625	0.108
Trig	0.062	0.057	1.089	0.281
APTT	-0.005	0.050	-0.100	0.920





Praca domowa

- Pracę domową należy wykonać pojedynczo. Na podstawie zbioru danych „sick” dostępnych w zbiorze OpenML należy wykonać analizę eksploracyjną oraz zbudować interpretowalny model klasyfikacyjny przewidujący czy pacjent jest chory czy zdrowy. Powinna zostać użyta 5-krotna krosvalidacja i wyliczone dwie miary: AUC i AUPRC.
- Praca w formie raportu .pdf i .Rmd w języku angielskim powinna być zamieszczona w folderze <https://github.com/mini-pw/2020L-WarsztatyBadawcze-InzynieriaCech/tree/master/PracaDomowa1/ImieNazwisko> do 17.04 do godz. 10 oraz zaprezentowana (max. 10 minut) podczas zajęć 17.04.

Ocena: max. 10pkt:

- 0-5pkt: ranking miary AUPRC

- 0-5pkt: jakość modelu, kodu, wykresów, prezentacja

Bookdown

- Mamy już repozytorium z ebookiem: <https://github.com/mini-pw/2020L-WB-Book>, w którym będziemy umieszczać artykuły. Należy założyć podrozdział odpowiadający artykułowi, który powstanie w ramach porojektu. Podrozdział zawierać ma autorów oraz tytuł.
- Książka będzie służyć do artykułów wszystkich grup WBI i WBII. Podzielona jest na 3 rozdziały.
- Pliki zaczynające się od 1-0, 2-0, 3-0 to wstępy do rozdziałów
- Pliki zaczynające się od 1-1, 1-2, 1-3 to podrozdziały, które powinny odpowiadać grupom.
- Istotne jest, że najwyższy stopień nagłówek w plikach z artykułami to ##, dzięki temu będą one podrozdziałami dla pliku 3-0-interpretability.Rmd
- Robiąc pull request zaczniście jego nazwę od INTER. Powinien zostać dodany tylko jeden plik .Rmd, bez renderowania książki na nowo.