

Session D2:

Frafald og fastholdelse

Paper 21

Estimeret sandsynlighed for at bestå 1. studieår til tiden

Programtekst

Der er udviklet en model, som på grundlag af de baggrundsoplysninger, der allerede kendes om de studerende på optagelsestidspunktet, estimerer den studerendes sandsynlighed for at bestå til tiden.

Abstract

Projektets tema

De sidste år har der været en stigende opmærksomhed på mulighederne for brug af big data og statistisk metode til at identificere de studerende, som har stor risiko for at få studiemæssige problemer, som fører til frafald eller studietidsforsinkelse. Dette projekt sigter på, kun ud fra de baggrundsdata som allerede kendes i ansøgningsfasen, at estimere den enkelte studerendes sandsynlighed for at gennemføre 1. studieår til tiden.

Metode

Projektet er gennemført ved indsamling af baggrunds- og gennemførelsesdata for 1. studieår på HA almen, idet "gennemførelse til tiden" estimeres ved logistisk regression med baggrundsoplysningerne som forklarende variable.

Resultater

Først etableredes en model, der forklarede 2015-årgangens gennemførelse. Denne model blev herefter anvendt til, på baggrund af optagelsesdata at forudsige 2016-årgangens gennemførelse. Opdelt i 5 lige store grupper, sorteret efter den estimerede sandsynlighed for at bestå, observeres en udmærket overensstemmelse mellem det forudsagte og den faktiske gennemførelsesfrekvens:

2016-årgang:		Forudsagt sandsynlighed for at bestå 1. år til tiden					
		< 63%	63 - < 71%	71 - < 76%	76 - < 81%	81 - 100%	Total
Forudsigelse	N	112	120	114	110	111	567
	Gns	.541	.671	.737	.787	.852	.717
Faktisk bestået 1. år til tiden	%	.616	.650	.754	.818	.883	.743

Af de 20%, som modellen gav mindst chance for at bestå til tiden, bestod 38,4% ikke til tiden. Det gjaldt kun 11,7% i den femtedel, som modellen gav størst chance. Det har signifikant betydning, at de studerende kommer fra Norge (17%), om de optages i kvote 2 frem for kvote 1 (17%), hvor høj deres studentereksamen er (11% pr. point), at de holder netop ét sabbatår (10%) og hvor gamle de var, da de blev student (-3% pr. år). Procenterne angiver, hvor mange %-point sandsynligheden for at bestå til tiden påvirkes.

Implikationer

Resultaterne kan bruges i overvejelser om kvote 2 optagets størrelse og til holdsammensætning. Evt. til forskellige former for nudging. Men hvor er grænsen for det etisk forsvarlige?

Forfatter

Michael Møller Nielsen, CBS