

Author (Year)	Cohort Size	Retained (#)	Retained (%)	Measure of Accuracy	Coeffes Used?	Techniques Used
Spady (1971)	683	615	90.04%	R^2 of .3132 for men and .3879 for women	Yes	Multiple regression
Bean (1980)	906	769	84.88%	R^2 of .22 for women and 0.09 for men	Yes	Multiple regression
Terenzini (1980)	study 1	379	60	R^2 of .1580% R^2 of .246	Yes	discriminate analyses
	study 3	518	428	R^2 of .82.63% R^2 of .256	Yes	Multiple regression
	study 5	763	673	R^2 of .88.20% R^2 of .309	Yes	discriminate analyses
	study 6	763	673	R^2 of .88.20% R^2 of .476 for men and .553 for women	Yes	discriminate analyses
Stage (1989)	323		91.00%	Multiple R	Yes	Logistic regression
Dey & Astin (1993)	947		16.00%	0.354, and 0.323	Yes	logit, probit, and regression
Waugh et al. (1994)	8667		60%	estimated ret	Yes	Survival Analysis/ Hazard regression
Murtaugh et al. (1999)	3535		88.30%	R^2 of 0.022	Yes	Logistic regression
Bresciani & Carson (2002)	any dropout; not only first-year;	3244	49.08%	overall accuracy of 83%	Yes	Logistic regression
Glynn et al. (2003)	accuracies based on the training data					
Herzog (2005)	5261		76.30%	77.4% accuracy	Yes	Logistic regression
	4298		77.10%		Yes	
	4671		83.50%	85.4% accuracy	Yes	
	2,444		79.50%	81.6% accuracy on training;	on	Logistic regression
Sujitparapitaya (2006)	2,445		79.50%	80.7% on validation		
Herzog (2006)	8,018		75.29%	83.9% accuracy on training; 82.1% on validation		Neural Network
Atwell et al. (2006)	2,445		79.50%	85.5% on training; 84.4% on validation		C4.5
DeLong et al. (2007)	3,829	3149	82.24%	accuracy close to 75%		Neural Networks; CHAID, C4.5, CR&T; Logistic regression
Pittman (2008)	5,990	4,881	81.49%			decision trees (entropy, chi-sq, gini) and logistic regression
			50%	precision varied from 57% to 60%		AdaBoost M1 with Decision Stumps
	21136	17139	81.10%	overall accuracy of 78-81%; not-retained precision from 44-63%		Logistic regression, neural network, bayes, J48

Table 1: Techniques and Accuracies Reported in Literature