Lecture 25:
Problem Set 6 Presentations

Big Data and Machine Learning for Applied Economics
Econ 4676

Ignacio Sarmiento-Barbieri

Universidad de los Andes

November 5, 2021

Results

"measuring poverty is hard, time consuming, and expensive. By building better models, we can run surveys with fewer, more targeted questions that rapidly and cheaply measure the effectiveness of new policies and interventions. The more accurate our models, the more accurately we can target interventions and iterate on policies, maximizing the impact and cost-effectiveness of these strategies"

$$Score = f(False\ Positive\ Rate, False\ Negative\ Rate, log(Indp.\ Vars.)) \tag{1}$$

- False Positive Rate: Classified as Poor when it's not (over n)
- False Negative Rate: Classified as Not Poor when it is (over n)

Results: Best Predictions

Table 1

	name	FN	FP	parameters	score1	score_final
1	Salazar, Cortes, Rojas, y Pena	0.131	0.028	168	0.159	5.204
2	Rodriguez y Montero	0.108	0.059	47	0.167	5.110
3	Prieto, Segura, y Navarro	0.096	0.101	7	0.197	5.065
4	Cepeda, Angulo, Cifuentes y Mosquera	0.009	0.238	16	0.247	5.398
5	Gonzalez y Rengifo	0.057	0.197	4	0.254	5.293
6	Castro, Ramirez y Miranda	0.050	0.222	6	0.273	5.433
7	Acero, Pacheco, Saenz	-	-	-	-	-

Results: Best Overall Model

Table 2

	name	FN	FP	parameters	score1	score_final
1	Prieto, Segura, y Navarro	0.096	0.101	7	0.197	5.065
2	Rodriguez y Montero	0.108	0.059	47	0.167	5.110
3	Salazar, Cortes, Rojas, y Pena	0.131	0.028	168	0.159	5.204
4	Gonzalez y Rengifo	0.057	0.197	4	0.254	5.293
5	Cepeda, Angulo, Cifuentes y Mosquera	0.009	0.238	16	0.247	5.398
6	Castro, Ramirez y Miranda	0.050	0.222	6	0.273	5.433
7	Acero, Pacheco, Saenz	-	-	-	-	-