Lecture 28:
Problem Set 3 Presentations

Big Data and Machine Learning for Applied Economics
Econ 4676

Ignacio Sarmiento-Barbieri

Universidad de los Andes

November 19, 2020

Announcements

- Problem Set 4 is posted
- Course Perception Survey
- Final Exam Dates: 9/12 8am a 11/12 8am had 18 votes, now the question is can we move this to the weekend?
  - 10/12 2pm a 12/12 2pm
  - 10/12 8pm a 12/12 8pm
  - 11/12 8am a 13/12 8am

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"

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

- 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

	Group	FN	FP	Indp. Vars.	score1
1	gutierrez-cano-y-perea	0.063	0.041	82	0.105
2	cubillos-diaz-y-perez	0.077	0.049	53	0.126
3	mojica-noriega-y-porras	0.091	0.039	9	0.130
4	becerra-ochoa-y-roa	0.085	0.048	19	0.133
5	lopez-pardo-y-urdaneta	0.107	0.046	21	0.153
6	naranjo-pedraza-y-pulido	0.130	0.042	13	0.172
7	grandas-rios-y-rodriguez	0.185	0.026	4	0.212
8	gomez-gomez-y-palacio	0.041	0.220	71	0.261
9	aleatorio	0.100	0.402	10,000	0.502

Results: Best Overall Model

Table 2

	Group	FN	FP	Indp. Vars.	score1	score_final
1	mojica-noriega-y-porras	0.091	0.039	9	0.130	4.777
2	becerra-ochoa-y-roa	0.085	0.048	19	0.133	4.863
3	gutierrez-cano-y-perea	0.063	0.041	82	0.105	4.882
4	cubillos-diaz-y-perez	0.077	0.049	53	0.126	4.936
5	lopez-pardo-y-urdaneta	0.107	0.046	21	0.153	4.964
6	naranjo-pedraza-y-pulido	0.130	0.042	13	0.172	5.009
7	grandas-rios-y-rodriguez	0.185	0.026	4	0.212	5.081
8	gomez-gomez-y-palacio	0.041	0.220	71	0.261	5.622
9	aleatorio	0.100	0.402	10,000	0.502	7.529