

MITOCW | MIT15_071S17_Session_4.3.07_300k

Let us introduce the error measures we used in building the analytics models.

We of course used R squared, but we also used other measures.

Next measure, the so-called "penalty error," is motivated by the fact that if you classify a very high-risk patient as a low-risk patient, this is more costly than the reverse, namely classifying a low-risk patient as a very high-risk patient.

Motivated by this, we developed a penalty error.

And the idea is to use asymmetric penalties.

The graph here-- so it's a matrix-- where this is the outcome and this is the forecast.

For example, whenever we classify a low-risk patient as high-risk, we pay a penalty of 2, which is a difference of 3 minus 1, the difference in the error.

But inversely, when you classify a bracket 3 patient as bracket 1 patient, this is double.

The cost-- the penalty-- is double the amount.

So you observe that the off diagonal penalties are double the corresponding penalties in the lower diagonal.

To judge the quality of the analytics models we developed, we compare it with a baseline.

And the baseline is to simply predict that the cost in the next "period" will be the cost in the current period.

We have observed that as far as identification of brackets is concerned, the accuracy was 75%.

So namely, whenever we predict that the risk is bracket 3-- indeed it is bracket 3-- this happens 75% of the time, and the penalty error-- the average penalty error of the baseline-- was 0.56.