Lift and Gain Charts

So far, we learnt about Precision, Recall, F1 score, ROC AUC and what not?

These metrics are good to evaluate the model and check its performance

• but they metric do not convey how a model would do business wise. How will the model impact the business?

These metrics are focusing on making sure making sure model are good at prediction

• but they don't help in translate it to business perspective

What's the need?

Business people need to know how our model's differences would make in the business term compared to random targeting.

For this purpose, Lift and Gain chart came into existence. To convey/To measure how much better our prediction model compared without the model to the stakeholders.

Where are they used?

Let's try understand the context

It is often used in target marketing.

Suppose zomato wants to run a marketing campaigns

But the customer response to these campaings are usually very low (1% in many cases)

But it cost money to run the campaign, right?

• The company would like to save money

Taking an inspiration from this, in ML

at the same time, it would like to achieve the desired response from the customers

• Lift and Gain chart help us graphically understand the benefit of using that model (in layman terms)

How to make Lift and Gain chart?

Suppose we built a classification model

• we have D_{cv} (cross validation data)

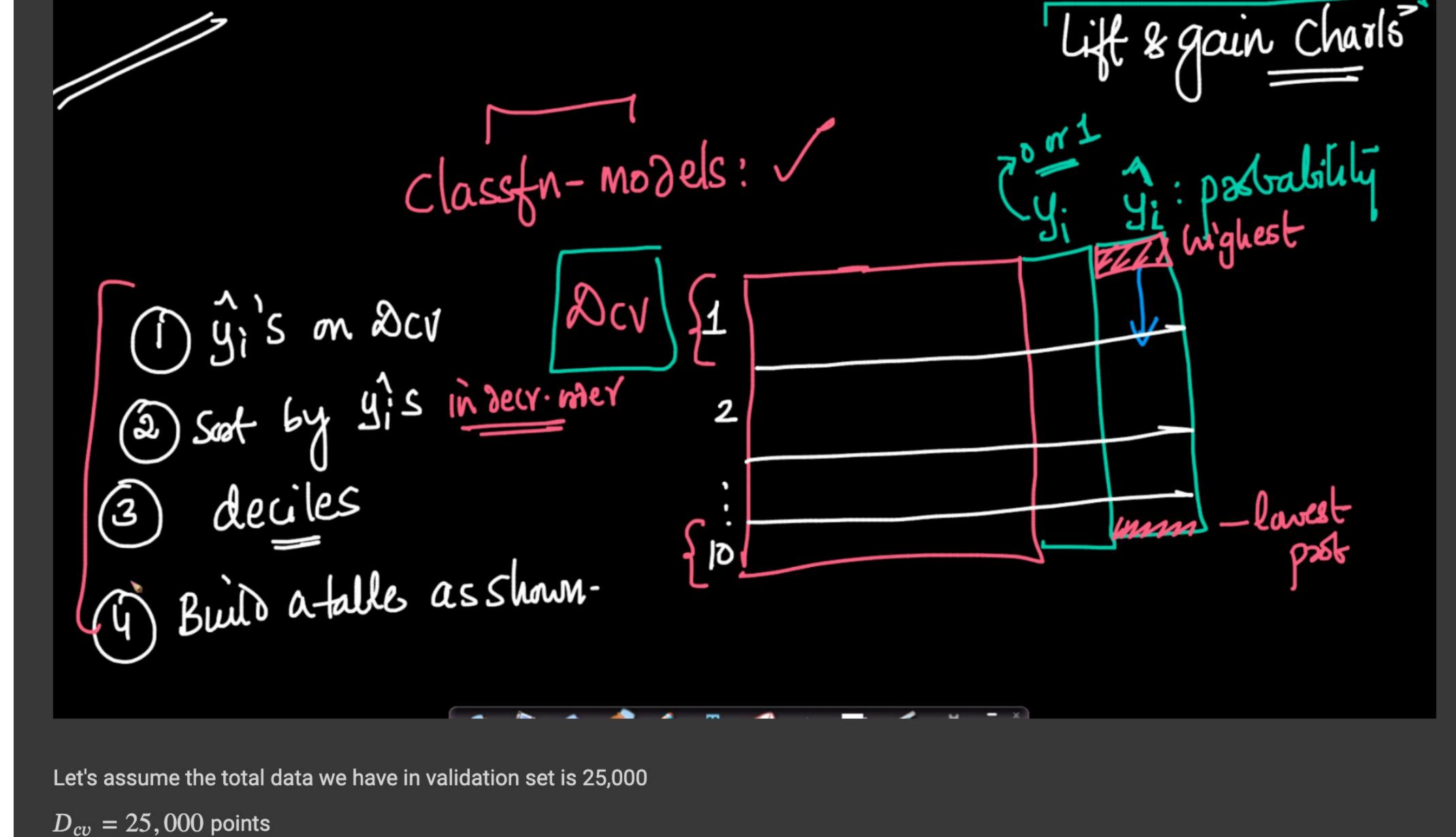
• Ground truth y_i (Assume 0 or 1)

• Predicted probability $\hat{y_i}$ (between 0 to 1) **Step 1**: Obtain the $\hat{y_i}$ on D_{cv}

Step 2: Sort the data predicted probability in desc order We will have highest probability at the top, lowest at the bottom.

Step 3: Break the sorted cross valdiation data into 10 groups (or deciles). We get the data for each decile

Step 4: Using these deciles, we built a table as follows



Now, after sorting these datapoints based on predicted probability, we split these into 10 deciles How many points will each decile or group will have ? => 2500

As we sorted the data in desc order of predicted probability, • 1st decile will have datapoints with highest pred. probability

Terminology

• 10th decile will have datapoints with lowest pred. probability

Number of cases means number of datapoints in each decile - As we have 10 decile and 25,000 pts - Each decile will have 2500 points

Number of Responses

Number of cases

Number of responses means the total number of points with class label 1 or ground truth value of 1 (positive class) • In first decile, out of 2500, we have 2179 points with class label 1

• It makes sense as data is sorted based on predicted probability

Notice few observations here

Out of 4874, approx half of them (2179) belong to 1st Decile

We have total of 4874 points belonging to class 1

• As we go down to 10th, there are lesser number of datapoints belonging to class 1

Cummulative Responses

Gain

• For 1st decile, Cummulative Responses will be 2179 o For 2nd decile, it'll be sum of 2179 and 1753 i.e. 3932 (sum of responses for 1st and 2nd decile)

Using the Number of responses, we make Cummulative Responses

o Similarly, we calculate cummulative responses for rest of the decile • The CR (Cummulative Responses) for last decile will be equal to total number of responses.

• Cummulative Responses, for a decile, is calcualted by summing up the number of positive (Responses) upto that decile

• % of events of a decile tells us "of all the points belonging to positive class, how many of them belong to that decile" Total number of positive points are 4874

 Positive points in 1st decile is 2179 \circ % of events for 1st decile will be (2179/4874) × 100 = 44.71

Now, we calculate Gain Gain for a decile is calculated by

 dividing the cumulative responses by total number of responses • i.e. cummulative number of postives by total number of positives.

For 2nd decile, gain will be (3932/4874) × 100 = 80.67 ... and so on **What does gain mean (in maths term)?** Gain for ith decile tells us "what percentage of positive points are in i^{th} or smaller decile" For example: For 8th decile, gain is 97.87 %. It means till 8th decile, we cover 97% of positive datapoints.

Now, Cummulative lift for a decile is calculated by

For 1st decile, gain will be (2179/4874) × 100 = 44.71

Cummulative Lift Imagine we have a random model How many number of response will there be for random model in 1st decile?

number of response for 1st decile will be roughly 10% of 4874 For a smart model, we have 44% of positive points (gain) in 1st decile

Since it is random model, each decile will have 10% of total positive points or number of response

• in other terms, it is ratio of gain of our model to a gain of random model For example: cummulative lift for 1st decile will be

• 44.71/ 10 i.e. gain for model for 1st decile /gain for random model for 1st decile

Cummulative lift for 2nd decile will be • 80.67/20 i.e. gain for model (2nd decile) / gain for random model in 2nd decile What does lift mean (in maths terms)? It means Cummulative percentage of positive points till ith decile divided by cummulative percentage of

positive points by random model It is intuitively telling how much better is model compared to random model.

3932

4328

4439

1/53

111

% of events

44.71

35.97

8.12

2.28

1st-deale 2179 pts have y== Agrentisement

Gain

44.71 V

80.67

88.80

91.08

Cumulative

4.47

4.03

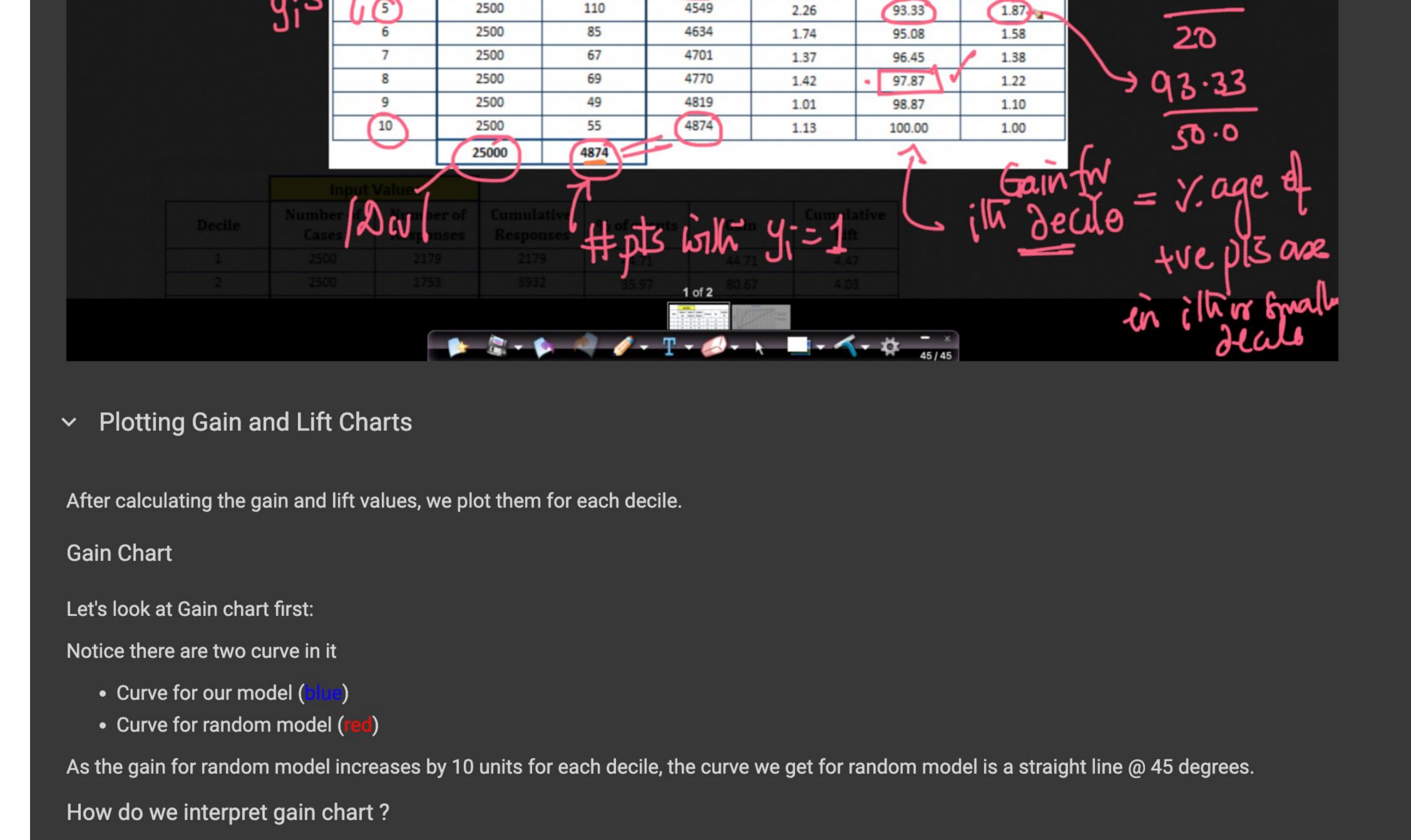
2.96

2.28

• dividing total % of positive point we have upto that decile (gain) for a smart model by total % of positive points if we had a random model

Input Values Number of Number of Cumulative Decile Responses Respor ses 2500 2179 2179

2500



It means that approx 80% of targets can be covered by top 20% of data. Suppose this was a churn prediction model. In that case, we could say, we can identify and target approx 80% of customer who are likely to churn just by targeting top 20% of customer.

Lift Chart

We have gain of approx 80 for second decible.

As you keep you keep coming to last decile, the lift will decrease eventually both model and random model will have lift of 1 for 10th decile How do we interpret lift chart?

What does it mean for business? Using fewer resources, we can avoid 80% of churn event

For calculating lift of decile, we simply divide gain of model by gain of random model

Note: The Greater the area between the Lift / Gain and Baseline, the Better the model.

A random model will have a lift of 1 for each decile. But, a good model will have lift as high as possible.

The lift for 2nd decile is 4. It means when selecting 20% of data based on model, • we could find target (actual positives) 4 times more than the randomly selected 20% of the data without a model.

Gain Chart

20 % of data sets Lift Chart 4.5 4.0 3.5 → Lift (Model) 3.0 ₹ 2.5 -Lift (Random) 2.0 1.5 1.0 0.5 40 50 % of data sets 2 of 2

ADVERTISEMENT