LOGIN / REGISTER (HTTPS://ID.ANALYTICSVIDHYA.COM/ACCOUNTS/LOGIN/?...VIDHYA.COM/BLOG/2019/08/11-IMPORTANT-MODEL-EVALUATION-ERROR-METRICS/)

Analytics Vidhya
Learn everything about analytics
(https://www.analyticsvidhya.com/blog/)

...LOG/CATEGORY/MACHINE-LEARNING/)
...GORY/PYTHON-2/)
...TEGORY/STATISTICS/)

11 Important Model Evaluation Metrics for Machine Learning Everyone should know

TAVISH SRIVASTAVA (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/AUTHOR/TAVISH1/), AUGUST 6, 2019 LOGIN TO BOOKMARK THIS A...

Overview

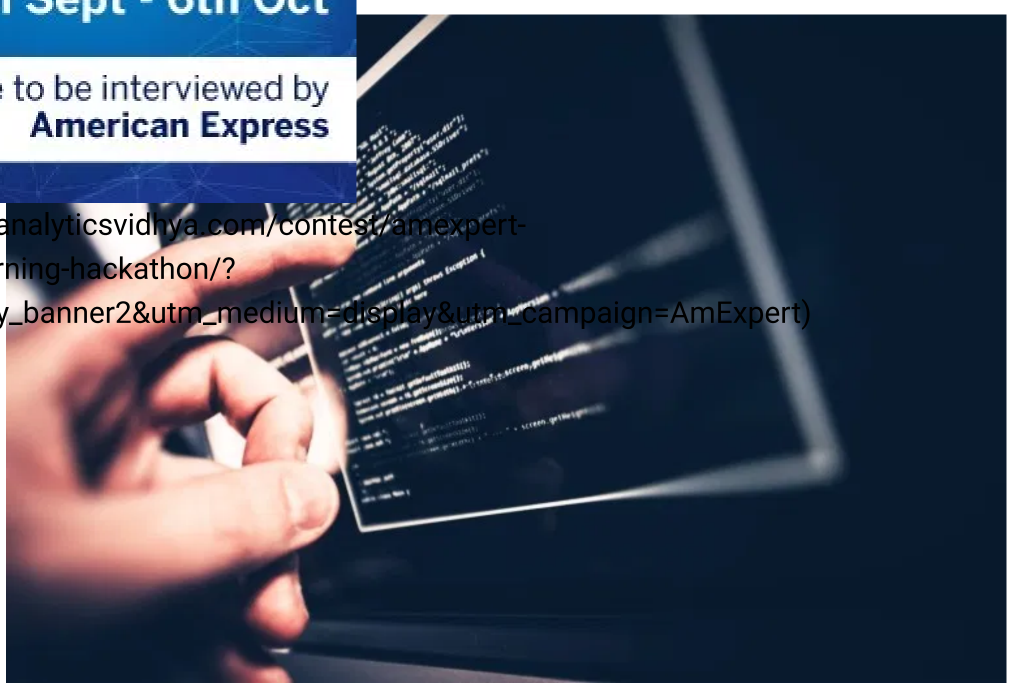
- Evaluating a model is a core part of building an effective machine learning model
- There are several evaluation metrics, like confusion matrix, cross-validation, AUC-ROC curve, etc.
- Different evaluation metrics are used for different kinds of problems

This article was originally published in February 2016 and updated in August 2019. with four new evaluation

...els (https://courses.analyticsvidhya.com/courses/applied-machine-
...e=blog&utm_medium=11-important-model-evaluation-error-metrics)
...You build a model, get feedback from metrics, make improvements
...ccuracy. Evaluation metrics explain the performance of a model. An
...r capability to discriminate among model results.

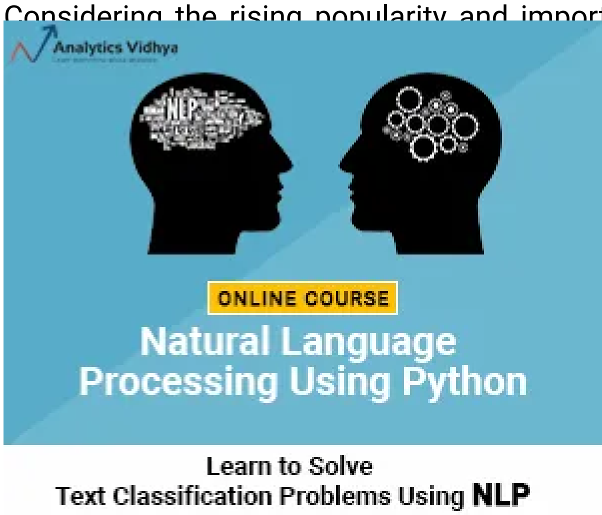
I have seen plenty of analysts and aspiring data scientists not even bothering to check how robust their model is.
...hurriedly map predicted values on unseen data. This is an incorrect

...r motive. It's about creating and selecting a model which gives high
...s crucial to check the accuracy of your model prior to computing



In our industry, we consider different kinds of metrics to evaluate our models. The choice of metric completely depends on the type of model and the implementation plan of the model.

After you are finished building your model, these 11 metrics will help you in evaluating your model's accuracy. Considering the rising popularity and importance of cross-validation, I've also mentioned its principles in this



ng journey, you should check out the comprehensive and popular *rses.analyticsvidhya.com/courses/applied-machine-learning-* *utm_medium=11-important-model-evaluation-error-metrics)e which* *h the various algorithms and components of machine learning.*

1. Confusion Matrix
2. F1 Score

Warming up: Types of Predictive models

When we talk about predictive models, we are talking either about a regression model (continuous output) or a classification model (nominal or binary output). The evaluation metrics used in each of these models are different.

In classification problems, we use two types of algorithms (dependent on the kind of output it creates):

1. **Class output**: Algorithms like SVM and KNN create a class output. For instance, in a binary classification problem, the outputs will be either 0 or 1. However, today we have algorithms which can convert these class outputs to probability. But these algorithms are not well accepted by the statistics community.

2. **Probability output**: Algorithms like Logistic Regression, Random Forest, Gradient Boosting, Adaboost etc. give probability outputs. Converting probability outputs to class output is just a matter of creating a threshold probability.

In regression problems, we do not have such inconsistencies in output. The output is always continuous in

ic discussion, I have used my predictions for the problem BCI blem is out of the scope of our discussion here. However the final ised for this article. The predictions made for this problem were d to class outputs assuming a threshold of 0.5.

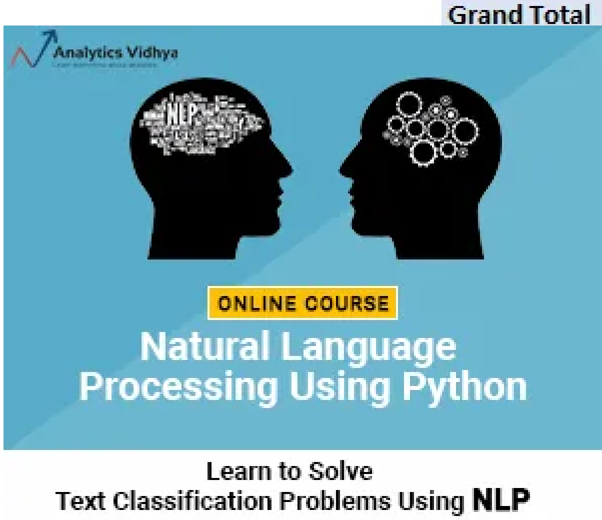
re N is the number of classes being predicted. For the problem in 2 matrix. Here are a few definitions, you need to remember for a

umber of predictions that were correct.
the proportion of positive cases that were correctly identified.
on of negative cases that were correctly identified.
actual positive cases which are correctly identified.
gative cases which are correctly identified.

Confusion Matrix		Target			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		

(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/Confusion_matrix.png)

Count of ID	Target			
Model	1	0	Grand Total	
1	3,834	639	4,473	85.7%
0	16	951	967	1.7%
Grand Total	3,850	1,590	5,440	
	99.6%	40.19%		88.0%



ontent/uploads/2015/01/Confusion_matrix1.png)

s out to be 88%. As you can see from the above two tables, the ive predictive value is quite low. Same holds for Sensitivity and shold value we have chosen. If we decrease our threshold value, the ne closer.

above defined metric. For instance, in a pharmaceutical company, rong positive diagnosis. Hence, they will be more concerned about high Specificity. On the other hand an attrition model will be more concerned with Sensitivity. Confusion matrix are generally used only with class output models.

d recall for classification problems and also highlighted the s our use case. What if for a use case, we are trying to get the best re is the harmonic mean of precision and recall values for a ore is as follows:

$$\ldots \frac{\ldots\text{precision}^{-1}}{2}\Bigg)^{-1} = 2\cdot\frac{\text{precision}\cdot\text{recall}}{\text{precision}+\text{recall}}.$$

Now, an obvious question that comes to mind is why are taking a harmonic mean and not an arithmetic mean. This is because HM punishes extreme values more. Let us understand this with an example. We have a binary classification model with the following results:

Precision: 0, Recall: 1

Here, if we take the arithmetic mean, we get 0.5. It is clear that the above result comes from a dumb classifier which just ignores the input and just predicts one of the classes as output. Now, if we were to take HM, we will get 0 which is accurate as this model is useless for all purposes.

This seems simple. There are situations however for which a data scientist would like to give a percentage more importance/weight to either precision or recall. Altering the above expression a bit such that we can include an adjustable parameter beta for this purpose, we get:

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}.$$

with respect to a user who attaches β times as much importance

...neck the rank ordering of the probabilities. Here are the steps to build a Lift/Gain chart:

...ation

...g order.

...almost 10% of the observations.

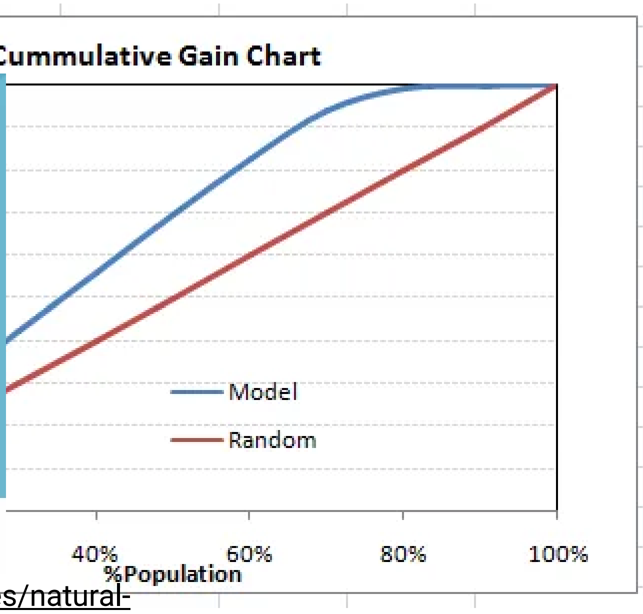
...eciles for Good (Responders) ,Bad (Non-responders) and total.

...eed to plot Gain/Lift charts:

Row Labels	0	1	Grand Total	%Rights	%Wrongs	%Population	Cum %Right	Cum %Pop	Lift @decile	Total Lift
				0%	0%	0%	0%	0%		
1		543	543	14%	0%	10%	14%	10%	141%	141%
2	2	542	544	14%	0%	10%	28%	20%	141%	141%
3	[illegible]	537	544	[illegible]	[illegible]	[illegible]	42%	30%	139%	141%
4	15	529	544	14%	1%	10%	56%	40%	137%	140%
5	20	524	544	14%	1%	10%	69%	50%	136%	139%
6	42	502	544	13%	3%	10%	83%	60%	130%	138%
7	104	440	544	11%	7%	10%	94%	70%	114%	134%
8	345	199	544	5%	22%	10%	99%	80%	52%	124%
9	515	29	544	1%	32%	10%	100%	90%	8%	111%
10	540	5	545	0%	34%	10%	100%	100%	1%	100%
Grand Total	1590	3850	5440							

(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/LiftnGain.png)

This is a very informative table. Cumulative Gain chart is the graph between Cumulative %Right and Cummulative %Population. For the case in hand here is the graph :

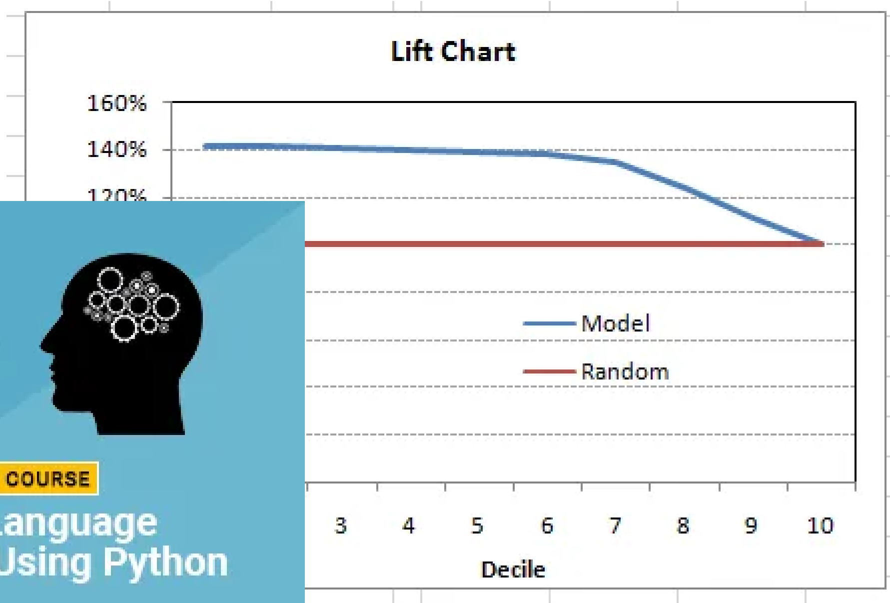


...ontent/uploads/2015/01/CumGain.png)

...segregating responders from non-responders. For example, the first ...s 14% of responders. This means we have a 140% lift at first decile.

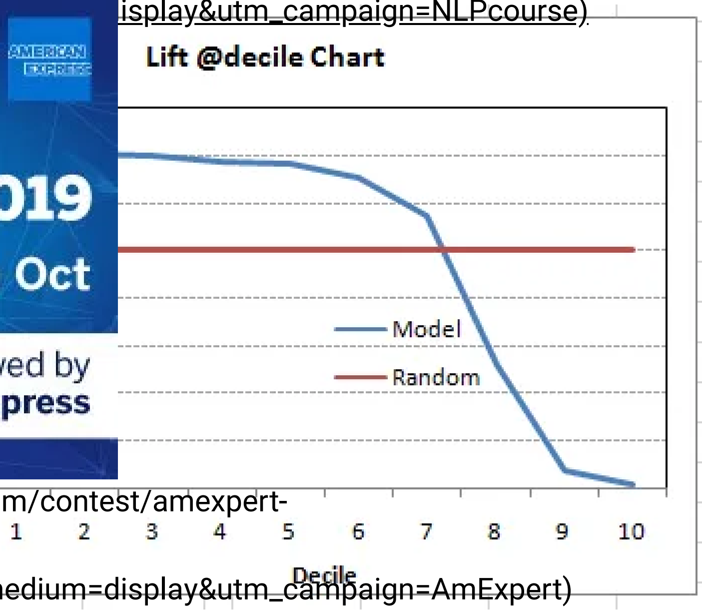
...ched in first decile? From the first table of this article, we know that ... Also the first decile will contains 543 observations. Hence, the ...543/3850 ~ 14.1%. Hence, we are quite close to perfection with this

...e plot between total lift and %population. Note that for a random ... the plot for the case in hand :



ontent/uploads/2015/01/Lift.png)

You can also plot decile wise lift with decile number :



(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/Liftdecile.png)

What does this graph tell you? It tells you that our model does well till the 7th decile. Post which every decile will be skewed towards non-responders. Any model with lift @ decile above 100% till minimum 3rd decile and maximum 7th decile is a good model. Else you might consider over sampling first.

Lift / Gain charts are widely used in campaign targeting problems. This tells us till which decile can we target customers for an specific campaign. Also, it tells you how much response do you expect from the new target base.

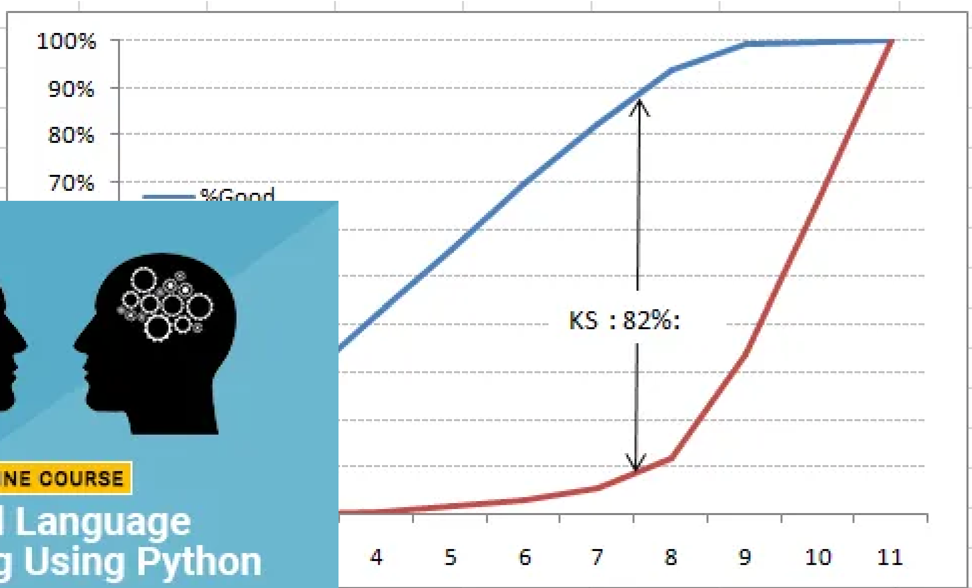
…performance of classification models. More accurately, K-S is a …the positive and negative distributions. The K-S is 100, if the scores …oups in which one group contains all the positives and the other all

…rentiate between positives and negatives, then it is as if the model … The K-S would be 0. In most classification models the K-S will fall …alue the better the model is at separating the positive from negative …cases.

						Cummulative			
				%Rights	%Wrongs	Cum %Rig	Cum %Wrong	K-S	
				0%	0%	0%	0%	0%	
				14%	0%	14%	0%	14%	
				14%	0%	28%	0%	28%	
				14%	0%	42%	1%	42%	
				14%	1%	56%	2%	54%	
				14%	1%	69%	3%	67%	
				13%	3%	83%	5%	77%	
				11%	7%	94%	12%	82%	K-S
8	[illegible]	[illegible]	[illegible]	[illegible]	22%	99%	34%	65%	
9	[illegible]	[illegible]	544	1%	32%	100%	66%	34%	
10	[illegible]	5	[illegible]	[illegible]	[illegible]	[illegible]	100%	0%	
Grand Tot	1590	3850	5440						

(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/KS.png)

We can also plot the %Cumulative Good and Bad to see the maximum separation. Following is a sample plot :



ontent/uploads/2015/01/KS_plot.png)

The metrics covered till here are mostly used in classification problems. Till here, we learnt about confusion ...irnov chart. Let's proceed and learn few more important metrics.

– ROC)

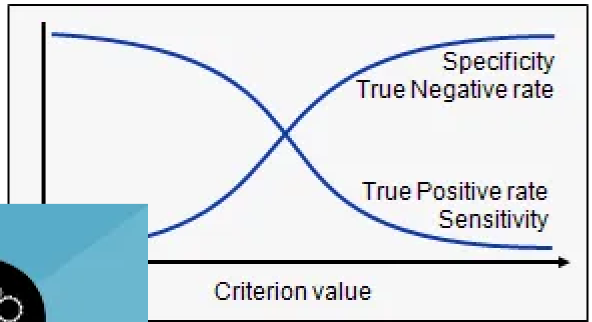
...in the industry. The biggest advantage of using ROC curve is that it ...n of responders. This statement will get clearer in the following

...eceiver operating characteristic) curve. If we look at the confusion ...tic model, we get different value for each metric.

		Target			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		

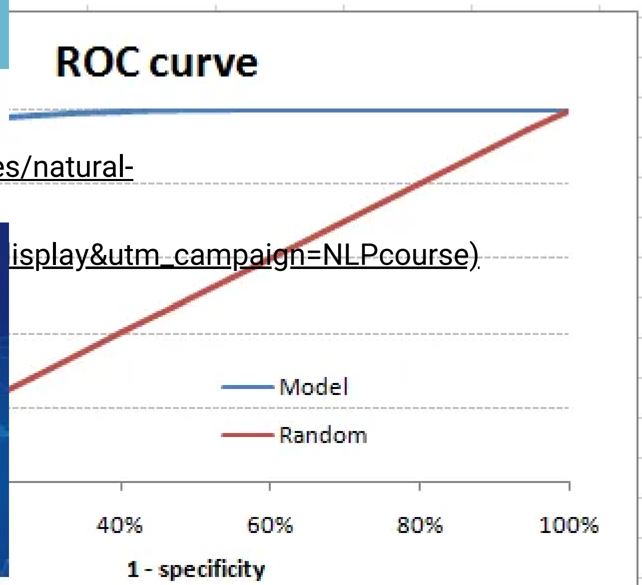
(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/Confusion_matrix.png)

Hence, for each sensitivity, we get a different specificity.The two vary as follows:



ontent/uploads/2015/01/curves.png)

and (1- specificity). (1- specificity) is also known as false positive
itive rate. Following is the ROC curve for the case in hand.



ontent/uploads/2015/01/ROC.png)

Let's take an example of threshold = 0.5 (refer to confusion matrix). Here is the confusion matrix :

Target		
1	0	Grand
3,834	639	
16	951	
3,850	1,590	
99.6%	40.19%	

(https://www.analyticsvidhya.com/blog/wp-content/uploads/2015/01/Confusion_matrix2.png)

Enroll Now

As you can see, the sensitivity at this threshold is 99.6% and the (1-specificity) is ~60%. This coordinate becomes on point in our ROC curve. To bring this curve down to a single number, we find the area under this curve (AUC).

Note that the area of entire square is 1*1 = 1. Hence AUC itself is the ratio under the curve and the total area. For the case in hand, we get AUC ROC as 96.4%. Following are a few thumb rules:

for the current model. But this might simply be over-fitting. In such and out-of-time validations.

1. For a model which gives class as output, will be represented as a single point in ROC plot.

h other as the judgement needs to be taken on a single metric and el with parameters (0.2,0.8) and model with parameter (0.8,0.2) can e metrics should not be directly compared.

tunate enough to get a single number which was AUC-ROC. But still, conclusive decisions. It is also possible that one model performs ter in other.

Advantages of using ROC

Why should you use ROC and not metrics like lift curve?

Lift is dependent on total response rate of the population. Hence, if the response rate of the population changes, the same model will give a different lift chart. A solution to this concern can be true lift chart (finding the ratio of lift and perfect model lift at each decile). But such ratio rarely makes sense for the business.

ROC curve on the other hand is almost independent of the response rate. This is because it has the two axis coming out from columnar calculations of confusion matrix. The numerator and denominator of both x and y axis will change on similar scale in case of response rate shift.

6. Log Loss

AUC ROC considers the predicted probabilities for determining our model's performance. However, there is an ... t the order of probabilities and hence it does not take into account ... bility for samples more likely to be positive. In that case, we could ... average of the log of corrected predicted probabilities for each

$$...og(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

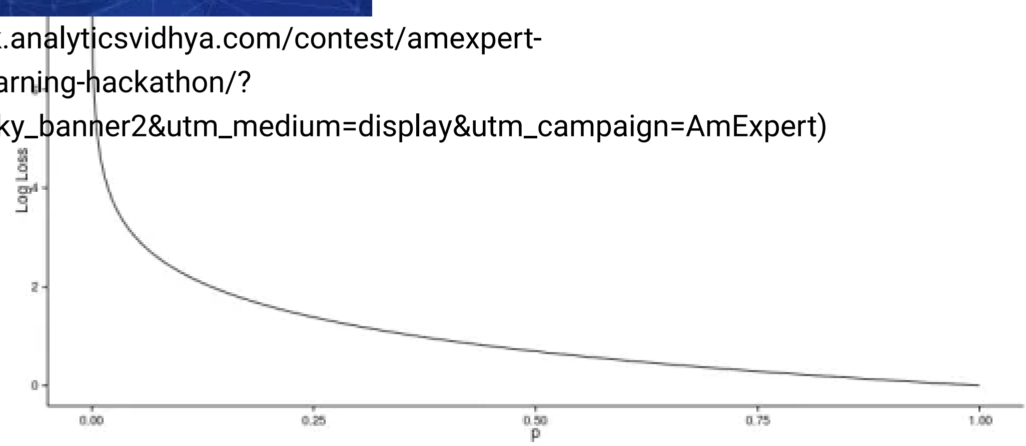
... class

... ve class

- y_i = 1 for positive class and 0 for negative class (actual values)

... ues to get the gist of the above mathematical function:

... as follows:



It’s apparent from the gentle downward slope towards the right that the Log Loss gradually declines as the predicted probability improves. Moving in the opposite direction though, the Log Loss ramps up very rapidly as the predicted probability approaches 0.

So, lower the log loss, better the model. However, there is no absolute measure on a good log loss and it is use-

...o binary classification with a varying decision threshold, log loss ...o account.

...cation problems. Gini coefficient can be straigh away derived from ...io between area between the ROC curve and the diagnol line & the area of the above triangle. Following is the formulae used :

... in hand we get Gini as 92.7%.

...c for any classification predictions problem. To understand this let’s ...ikelihood to pass this year. Following are our predictions :

A – 0.9

B – 0.5

C – 0.3

Now picture this. if we were to fetch pairs of two from these three student, how many pairs will we have? We will have 3 pairs : AB , BC, CA. Now, after the year ends we saw that A and C passed this year while B failed. No, we choose all the pairs where we will find one responder and other non-responder. How many such pairs do we have?

We have two pairs AB and BC. Now for each of the 2 pairs, the concordant pair is where the probability of responder was higher than non-responder. Whereas discordant pair is where the vice-versa holds true. In case both the probabilities were equal, we say its a tie. Let's see what happens in our case :

AB – Concordant

BC – Discordant

his example. Concordant ratio of more than 60% is considered to be
sed when deciding how many customer to target etc. It is primarily
. For decisions like how many to target are again taken by KS / Lift

)

RMSE is the most popular evaluation metric used in regression problems. It follows an assumption that error are
re are the key points to consider on RMSE:

his metric to show large number deviations.
s to deliver more robust results which prevents cancelling the
her words, this metric aptly displays the plausible magnitude of

s which is highly undesirable in mathematical calculations.
ucting the error distribution using RMSE is considered to be more

s. Hence, make sure you've removed outliers from your data set

6. As compared to mean absolute error, RMSE gives higher weightage and punishes large errors.

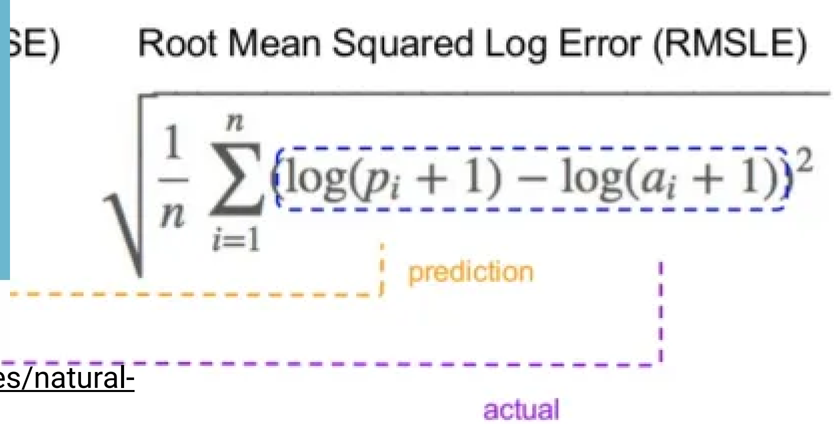
RMSE metric is given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

where, N is Total Number of Observations.

10. Root Mean Squared Logarithmic Error

In case of Root mean squared logarithmic error, we take the log of the predictions and actual values. So basically, ...easuring. RMSLE is usually used when we don't want to penalize ...ual values when both predicted and true values are huge numbers.



$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(p_i+1)-\log(a_i+1))^2}$$

...small: RMSE and RMSLE are same.

...big: RMSE > RMSLE

...big: RMSE > RMSLE (RMSLE becomes almost negligible)

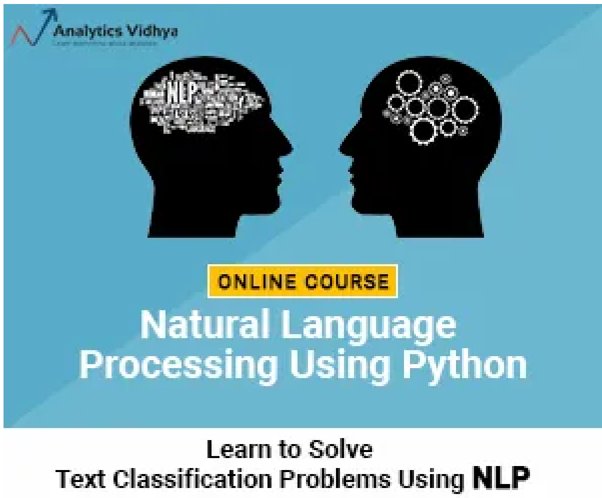
...e model's performance will improve. But these values alone are not

In the case of a classification problem, if the model has an accuracy of 0.8, we could gauge how good our model is against a random model, which has an accuracy of 0.5. So the random model can be treated as a benchmark. But when we talk about the RMSE metrics, we do not *have* a benchmark to compare.

This is where we can use R-Squared metric. The formula for R-Squared is as follows:

$$R^2 = 1 - \frac{\mathrm{MSE(model)}}{\mathrm{MSE(baseline)}}$$

$$\frac{\text{MSE(model)}}{\text{MSE(baseline)}} \qquad \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(\overline{y}_i - \hat{y}_i)^2}$$



...ictions against the actual values

...prediction against the actual values

...el as compared to a very simple model that just predicts the mean ...ns.

A model performing equal to baseline would give R-Squared as 0. Better the model, higher the r2 value. The best model with all correct predictions would give R-Squared as 1. However, on adding new features to the model, the ...he same. R-Squared does not penalize for adding features that add ...n over the R-Squared is the adjusted R-Squared. The formula for

$$1-\left(1-R^2\right)\left[\frac{n-1}{n-(k+1)}\right]$$

As you can see, this metric takes the number of features into account. When we add more features, the term in the denominator n-(k+1) decreases, so the whole expression increases.

If R-Squared does not increase, that means the feature added isn't valuable for our model. So overall we subtract a greater value from 1 and adjusted r2, in turn, would decrease.

Beyond these 11 metrics, there is another method to check the model performance. These 7 methods are statistically prominent in data science. But, with arrival of machine learning, we are now blessed with more robust methods of model selection. **Yes! I'm talking about Cross Validation**.

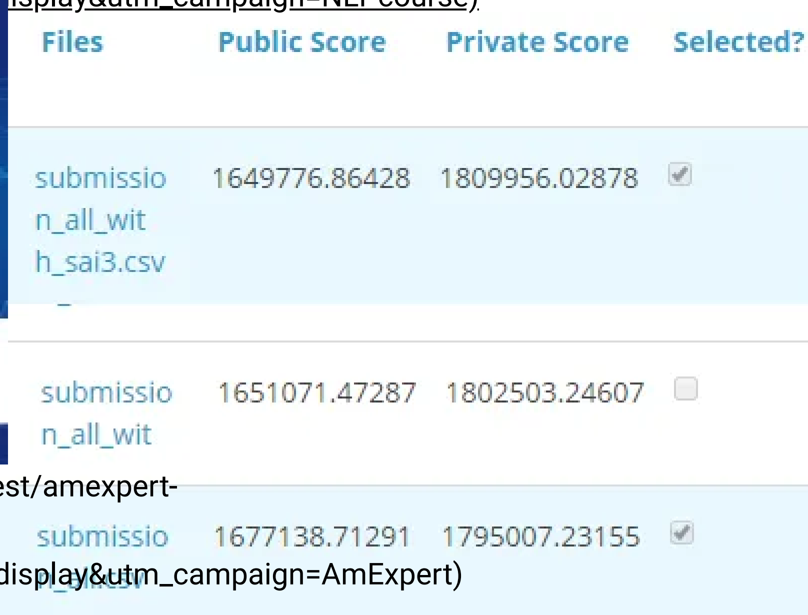
Though, cross validation isn't a really an evaluation metric which is used openly to communicate model accuracy. But, the result of cross validation provides good enough intuitive result to generalize the performance of a model.

Let's now understand cross validation in detail.

...ss validation. Due to busy schedules, these days I don't get much ...ions. Long time back, I participated in TFI Competition on Kaggle. ...ance, I would like to show you the dissimilarity between my public

...gle!

For TFI competition, following were three of my solution and scores (Lesser the better) :



Files	Public Score	Private Score	Selected?
submission_all_with_sai3.csv	1649776.86428	1809956.02878	✓
submission_all_wit...	1651071.47287	1802503.24607	
submissio...	1677138.71291	1795007.23155	✓

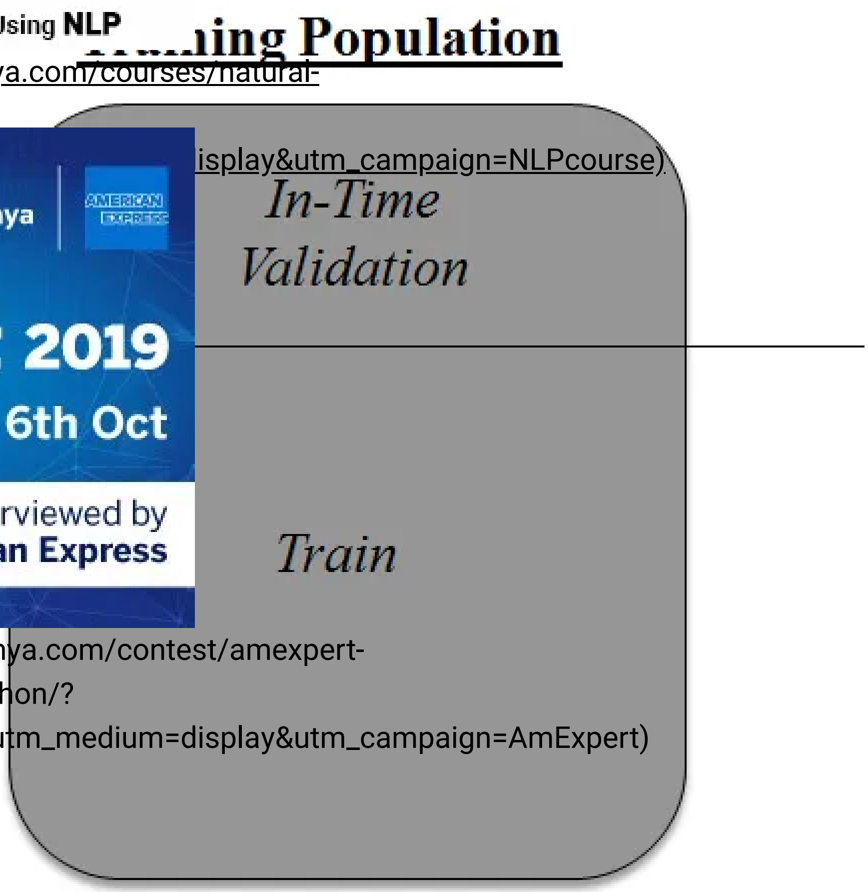
(https://www.analyticsvidhya.com/wp-content/uploads/2015/05/kagglescores.png)

You will notice that the third entry which has the worst Public score turned to be the best model on Private ranking. There were more than 20 models above the "submission_all.csv", but I still chose "submission_all.csv" as my final entry (which really worked out well). What caused this phenomenon ? The dissimilarity in my public and private leaderboard is caused by over-fitting.

Over-fitting is nothing but when you model become highly complex that it starts capturing noise also. This 'noise' adds no value to model, but only inaccuracy.

In the following section, I will discuss how you can know if a solution is an over-fit or not before we actually know the test results.

t concepts in any type of data modelling. It simply says, try to leave l and test the model on this sample before finalizing the model.



(https://www.analyticsvidhya.com/wp-content/uploads/2015/05/validation.png)

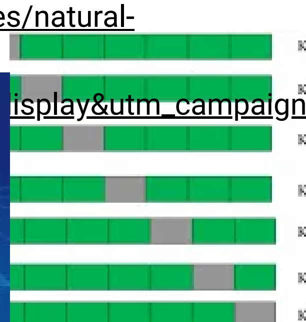
Above diagram shows how to validate model with in-time sample. We simply divide the population into 2 samples, and build model on one sample. Rest of the population is used for in-time validation.

Could there be a negative side of the above approach?

I believe, a negative side of this approach is that we loose a good amount of data from training the model. Hence, the model is very high bias. And this won't give best estimate for the coefficients. So what's the next best option?

pulation and the train on first 50 and validate on rest 50. Then, we ay we train the model on the entire population, however on 50% in le selection to some extent but gives a smaller sample to train the ross validation.

om 2-fold cross validation. Now, we will try to visualize how does a



t/uploads/2015/05/kfolds.png)

divide the entire population into 7 equal samples. Now we train models on 6 samples (Green boxes) and validate on 1 sample (grey box). Then, at the second iteration we train the model with a different sample held as validation. In 7 iterations, we have basically built model on each sample and held each of them as validation. This is a way to reduce the selection bias and reduce the variance in prediction power. Once we have all the 7 models, we take average of the error terms to find which of the models is best.

How does this help to find best (non over-fit) model?

k-fold cross validation is widely used to check whether a model is an overfit or not. If the performance metrics at each of the k times modelling are close to each other and the mean of metric is highest. In a Kaggle competition, you might rely more on the cross validation score and not on the Kaggle public score. This way you will be sure that the Public score is not just by chance.

model?

. Here is how you code a k-fold in Python :

```
el = RandomForestClassifier(n_estimators=100) #Simple K-Fold
lder scikit-learn versions the "n_folds" argument is named
ain), n_folds=5, indices=False) results = [] # "model" can b
_function" can be replaced by the error function of your anal
ysis for traincv, testcv in cv:        probas = model.fit(train[traincv], target[traincv]).predi
lts.append( Error_function ) #print out the mean of the cross
r( np.array(results).mean() )
```


choose k.

but low variance in the performances.

For a large k, we have a small selection bias but high variance in the performances.

Think of extreme cases :

k = 2 : We have only 2 samples similar to our 50-50 example. Here we build model only on 50% of the population each time. But as the validation is a significant population, the variance of validation performance is minimal.

k = number of observations (n) : This is also known as "Leave one out". We have n samples and modelling repeated n number of times leaving only one observation out for cross validation. Hence, the selection bias is minimal but the variance of validation performance is very large.

Generally a value of k = 10 is recommended for most purpose.

End Notes

Measuring the performance on training sample is point less. And leaving a in-time validation batch aside is a

very singe datapoint which can reduce this selection bias to a good ed with any modelling technique.

are some of the most used metrics of evaluation in a classification

fication and regression problem ? Have you used k-fold cross d you see any significant benefits against using a batch validation? e in the comments section below.

You want to apply your analytical skills and test your potential? Then participate in our (…vidhya.com/contest/all) **and compete with Top Data**

hya's Android APP GET IT ON Google Play

=blog_article&utm_campaign=blog&pcampaignid=MKT-Other-

...OG/TAG/AUC/), **CONCORDANT RATIO**
...ICORDANT-RATIO/), **CONFUSION MATRIX**
...IFUSION-MATRIX/), **CROSS-VALIDATION**
...SS-VALIDATION/), **DISCORDANT RATIO**
...CORDANT-RATIO/), **ERROR METRICS**
...OR-METRICS/), **GAIN AND LIFT CHARTS**
...N-AND-LIFT-CHARTS/), **GINI COEFFICIENT**
...-COEFFICIENT/), **K FOLD VALIDATION**
...OLD-VALIDATION/), **KOLMOGOROV SMIRNOV CHARTS**
...MOGOROV-SMIRNOV-CHARTS/), **PREDICTIVE MODELING**
...DICTIVE-MODELING/), **ROC**
(HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/ROC/)

NEXT ARTICLE

...me Object Detection using the Powerful SlimYOLOv3 Framework

.../2019/08/introduction-slimyolov3-real-time-object-detection/)

•••

PREVIOUS ARTICLE

...with these 5 Must-Know Applications of Singular Value Decomposition (SVD) in Data Science

(https://www.analyticsvidhya.com/blog/2019/08/5-applications-singular-value-decomposition-svd-data-science/)

(https://www.analyticsvidhya.com/blog/author/tavish1/)

[avish Srivastava csvidhya.Com/Blog/Author/Tavish1/)

rategy Officer of Analytics Vidhya, is an IIT Madras graduate and with 8+ years of diverse experience in markets including the US, ling Digital Acquisitions, Customer Servicing and Customer ail Banking, Credit Cards and Insurance. He is fascinated by the human intelligence and enjoys every discussion, theory or even movie related to this idea.

analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

seful. It is a good matrix to identify better model in case of multi

VENUGOPAL

February 19, 2016 at 6:27 am (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-105911)

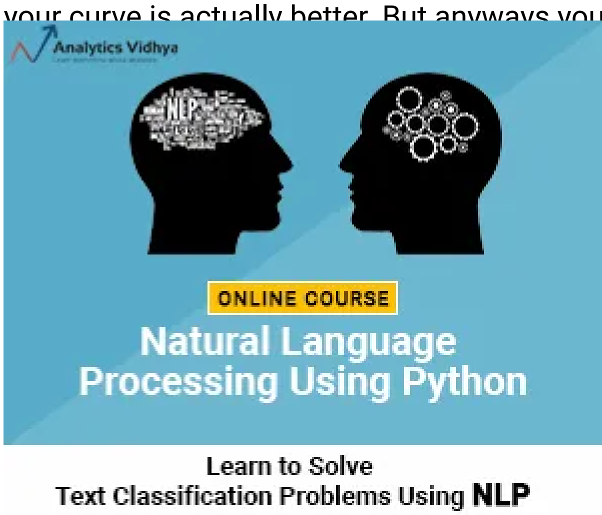
Very useful

MARIO

February 20, 2016 at 5:00 pm (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-105976)

Hi,

Great post thanks. Just in number one confusion matrix you miscalculated the negative predicted value.
It is not 1.7% but 98.3%. Same for specificity (59.81% instead of 40.19%). Since you reuse this example for ROC, your curve is actually better. But anyways your argument still holds. Nicely presented.



...ANDULA

...w.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

...analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-106243)

...analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

...ive predicrive value is 951/967 or not? Is there an error in confusion

AMOGH (HTTP://ANALYTICS%20VIDHYA)

February 25, 2016 at 12:32 pm (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-106228)

Yes I agree with Tamara the negative predictive value should be 98.3454%. Can you please confirm Tavish

RISHABH0709

March 2, 2016 at 8:03 am (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-106471)

Hi Tavish,
Thanks again , for this valuable article.
It would be great , if along with this very informative explanation , you can also provide how to code it , preferably in R.
Thanks

...alyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

...constructive feedback principle. You build a model. Get feedback

July 29, 2016 at 2:10 pm (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-114256)

...yticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

SUDHINDRA
August 11, 2016 at 12:31 am (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-114691)

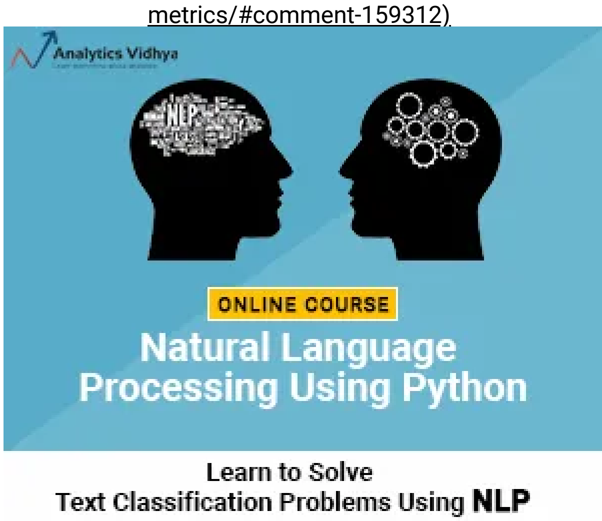
Can we all the list of Statistical Models and the application scenarios please for a novice person like me?

JORGE
September 21, 2016 at 9:49 pm (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-116313)

Are there metrics for non supervised models as k-means for example? Thanks!

PRANAV DAR

August 16, 2019 at 4:24 pm (https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/#comment-159312)



…metrics for unsupervised learning as well.

…w.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-

…ent on total response rate of the population"? Is this only applicable when you are able to correctly predict 100% of the 1st (or more) deciles?

…ECOSYSTEM

…dhya

…lytics Vidhya

POPULAR POSTS

Here are 7 Data Science Projects on GitHub to Showcase your Machine Learning Skills! (https://www.analyticsvidhya.com/blog/2019/09/7-data-science-projects-github-showcase-your-skills/)

Commonly used Machine Learning Algorithms (with Python and R Codes) (https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/)

24 Ultimate Data Science Projects To Boost Your Knowledge and Skills (& can be accessed freely) (https://www.analyticsvidhya.com/blog/2018/05/24-ultimate-data-science-projects-to-boost-your-knowledge-and-skills/)

A Complete Python Tutorial to Learn Data Science from Scratch

016/01/complete-tutorial-learn-data-science-python-scratch-2/)

! (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-

oduction to Interpretable Machine Learning Models in Python

019/08/decoding-black-box-step-by-step-guide-interpretable-

uld Learn (in Python, R, Tableau and D3.js)

019/08/11-data-visualizations-python-r-tableau-d3js/)

Features from Image Data using Python

019/08/3-techniques-extract-features-from-image-data-machine-

... How it Works from Scratch
...19/09/introduction-to-pytorch-from-scratch/)

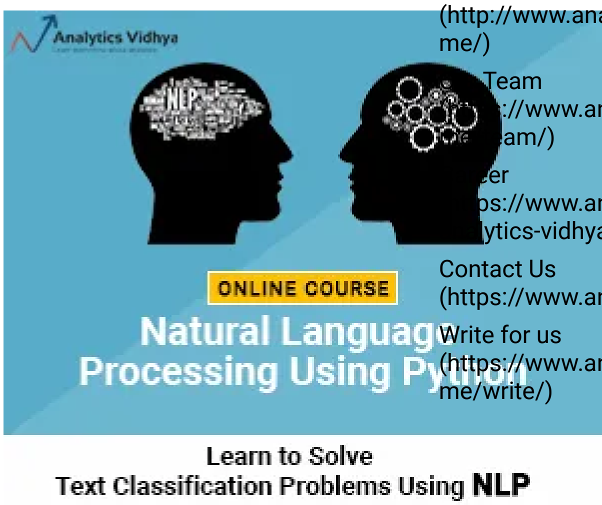
...h Image Data using skimage in Python
...19/09/9-powerful-tricks-for-working-image-data-skimage-

...very Data Scientist and Leader Should Know
...19/09/4-key-aspects-data-science-project/)

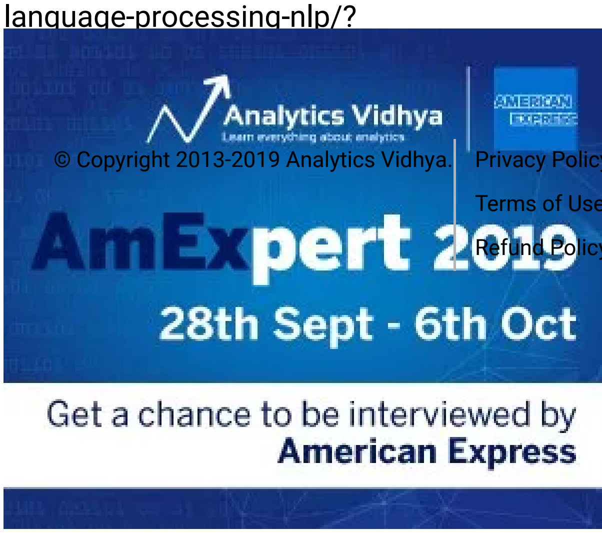
SEPTEMBER 12, 2019

A Data Scientist's Guide to 8 Types of Sampling Techniques
(https://www.analyticsvidhya.com/blog/2019/09/data-scientists-guide-8-types-of-sampling-techniques/)

SEPTEMBER 11, 2019



(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?...isplay&utm_campaign=NLPcourse)



(https://datahack.analyticsvidhya.com/contest/amexpert-2019-machine-learning-hackathon/?utm_source=Sticky_banner2&utm_medium=display&utm_campaign=AmExpert)

ANALYTICS VIDHYA

About Us (http://www.analyticsvidhya.com/about-me/)

Our Team (https://www.analyticsvidhya.com/about-me/team/)

Career (https://www.analyticsvidhya.com/career-analytics-vidhya/)

Contact Us (https://www.analyticsvidhya.com/contact/)

Write for us (https://www.analyticsvidhya.com/about-me/write/)

DATA SCIENTISTS

Blog (https://www.analyticsvidhya.com/blog/)

Hackathon (https://datahack.analyticsvidhya.com/)

Discussions (https://discuss.analyticsvidhya.com/)

Apply Jobs (https://www.analyticsvidhya.com/jobs/)

Leaderboard (https://datahack.analyticsvidhya.com/users/)

COMPANIES

Post Jobs (https://www.analyticsvidhya.com/corporate/)

Trainings (https://trainings.analyticsvidhya.com)

Hiring Hackathons (https://datahack.analyticsvidhya.com/)

Advertising (https://www.analyticsvidhya.com/contact/)

Reach Us (https://www.analyticsvidhya.com/contact/)

JOIN OUR COMMUNITY :

(https://www.facebook.com/AnalyticsVidhya/) 46336 Followers

(https://www.twitter.com/analyticsvidhya) 22180 Followers

(https://plus.google.com/+Analyticsvidhya) 7513 Followers

(https://in.linkedin.com/company/analytics-vidhya) Followers

Subscribe to emailer >

Privacy Policy (https://www.analyticsvidhya.com/privacy-policy/)

Terms of Use (https://www.analyticsvidhya.com/terms/)

Refund Policy (https://www.analyticsvidhya.com/refund-policy/)

Don't have an account? Sign up (https

Enroll Now