

x

-

(http://play.google.com/store/apps/details?id=com.analyticsvidhya.android)



Analytics Vidhya
Learn everything about analytics
www.analyticsvidhya.com/blog/

(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=AVtopblogBanner)

(https://datahack.analyticsvidhya.com/contest/american-)



[/BLOG/CATEGORY/MACHINE-LEARNING/](#),
[CATEGORY/PYTHON-2/](#),
[Y/R/](#)

Naive Bayes Algorithm (with codes in Python and

[G/AUTHOR/SUNIL-RAY/](#)), SEPTEMBER 11, 2017

• **Note:** This article was originally published on Sep 13th, 2015 and updated on Sept 11th, 2017
(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?)

Introduction
utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

Here's a situation you've got into:

You are working on a classification problem and you have generated your set of hypothesis, created features and discussed the importance of variables. Within an hour, stakeholders want to see the first cut of the model.

What will you do? You have hundreds of thousands of data points and quite a few variables in your training data set. In such situation, if I were at your place, I would have used '**Naive Bayes**', which can be extremely fast relative to other classification algorithms. It works on Bayes theorem of probability to predict the class of unknown data set.

Subscribe!

In this article, I'll explain the basics of this algorithm, so that next time when you come across large data sets, you can bring this algorithm to action. In addition, if you are a [newbie in Python or R](https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/learning-path-data-science-python/) (<https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/learning-path-data-science-python/>), you should be overwhelmed by the presence of available codes in this article



Naive Bayes?

Naive Bayes

Naive Bayes Model in Python

Naive Bayes Model

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=CV101AVBlogBanner&utm_medium=StickyBanner2&utm_campaign=CV101Banner)



Bayes' Theorem (https://en.wikipedia.org/wiki/Bayes%27_theorem)

predictors. In simple terms, a Naive Bayes classifier assumes that a class is unrelated to the presence of any other feature. For example, if it is red, round, and about 3 inches in diameter. Even if we know the existence of the other features, all of these properties are independent of each other. This is why it is known as 'Naive'.

Naive Bayes is particularly useful for very large data sets. Along with simplicity,

Naive Bayes is known to outperform even highly sophisticated classification methods.

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=StickyBanner2&utm_campaign=CV101Banner)

Bayes theorem provides a way of calculating posterior probability $P(c|x)$ from $P(c)$, $P(x)$ and $P(x|c)$. Look at the equation below:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability

Posterior Probability
Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

(https://www.analyticsvidhya.com/wp-content/uploads/2015/09/Bayes_rule-300x172.png) Above,

Subscribe!

- $P(c|x)$ is the posterior probability of *class* (c , *target*) given *predictor* (x , *attributes*).
- $P(c)$ is the prior probability of *class*.
- $P(x|c)$ is the likelihood which is the probability of *predictor* given *class*.
- $P(x)$ is the prior probability of *predictor*.



I have a training data set of weather and corresponding target (playing/not playing). Now, we need to classify whether players will play or not using below steps to perform it.

table

probabilities like Overcast probability = 0.29 and probability of

[https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=AnalyticsVidhya&utm_medium=Banner2utm_campaign=CV101banner)



Table	
0	Yes
4	
3	2
2	3
5	9

Likelihood table			
Weather	No	Yes	
Overcast		4	=4/14 0.29
Rainy	3	2	=5/14 0.36
Sunny	2	3	=5/14 0.36
All	5	9	
	=5/14 0.36	=9/14 0.64	

[uploads/2015/08/Bayes_41.png](#)

[https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=AnalyticsVidhya&utm_medium=Banner2utm_campaign=CV101banner)

Step 1: Now, use Naive Bayes equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of predictor.

Problem: Players will play if weather is sunny. Is this statement is correct?

We can solve it using above discussed method of posterior probability.

$$P(\text{Yes} | \text{Sunny}) = P(\text{Sunny} | \text{Yes}) * P(\text{Yes}) / P(\text{Sunny})$$

Here we have $P(\text{Sunny} | \text{Yes}) = 3/9 = 0.33$, $P(\text{Sunny}) = 5/14 = 0.36$, $P(\text{Yes}) = 9/14 = 0.64$

Now, $P(\text{Yes} | \text{Sunny}) = 0.33 * 0.64 / 0.36 = 0.60$, which has higher probability.

Subscribe!

Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification and with problems having multiple classes.



Naive Bayes?

It is a simple probabilistic classifier for data set. It also perform well in multi class prediction

holds, a Naive Bayes classifier performs better compare to and you need less training data.

It handles both categorical and numerical variables compared to numerical variable(s). For numerical variables, it assumes a Gaussian distribution (bell curve, which is a strong assumption).

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+CV101+CV102/about>)



If a categorical variable is observed in test data set, which was not observed in training data set, the probability will be zero and it will be unable to make a prediction. This is often called as zero probability problem. To solve this, we can use the smoothing technique. One of the simplest technique is Laplace smoothing.

Naive Bayes is also known as a bad estimator, so the probability outputs from Naive Bayes should be taken seriously.

Naive Bayes makes the assumption of independent predictors. In real life, it is almost impossible to find predictors which are completely independent.

Algorithms

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+CV101+CV102/about>)

• **Real time Prediction:** Naive Bayes is an eager learning classifier and it is sure fast. Thus, it could be used for making predictions in real time.

utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

- **Multi class Prediction:** This algorithm is also well known for multi class prediction feature. Here we can predict the probability of multiple classes of target variable.
- **Text classification/ Spam Filtering/ Sentiment Analysis:** Naive Bayes classifiers mostly used in text classification (due to better result in multi class problems and independence rule) have higher success rate as compared to other algorithms. As a result, it is widely used in Spam filtering (identify spam e-mail) and Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments)
- **Recommendation System:** Naive Bayes Classifier and Collaborative Filtering (https://en.wikipedia.org/wiki/Collaborative_filtering) together builds a Recommendation System that uses machine learning and data mining techniques to filter unseen information and predict whether a user would like a given resource or not

Subscribe!

How to build a basic model using Naive Bayes in Python?

Again, scikit learn (python library) will help here to build a Naive Bayes model in Python. There are three models in this library:



MultinomialNB (http://scikit-learn.org/stable/modules/naive_bayes.html) It is used in classification and discrete distribution.

BernoulliNB (http://scikit-learn.org/stable/modules/naive_bayes.html): It is used for discrete distribution. Here we can consider bernoulli instead of “word occurring in the document”, we have “count how many times word occurs in the document”. You can think of it as “number of times outcome number x_i is observed in n trials”.

BinomialNB (http://scikit-learn.org/stable/modules/naive_bayes.html): The binomial model is useful if you have a text classification problem with only two classes (zeros and ones). One application would be text classification with two classes. The features are “word occurs in the document” and “word does not occur in the document”.



Below is the example of Gaussian Naive Bayes model.

```
import numpy as np
from sklearn.naive_bayes import GaussianNB

# Create a Gaussian Naive Bayes model
model = GaussianNB()

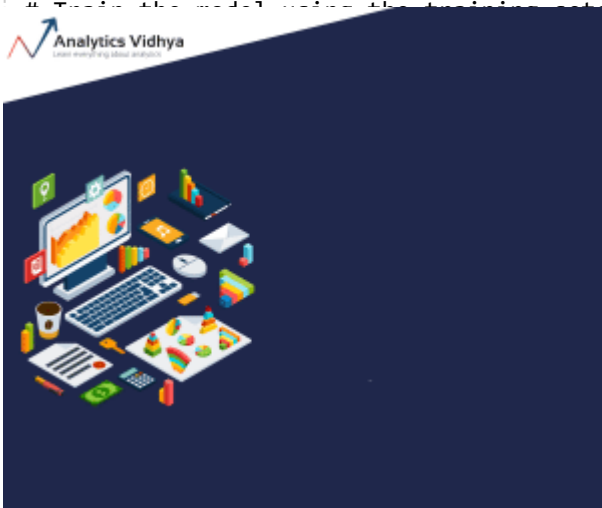
# Train the model with training data
model.fit(X_train, Y_train)

# Predict the class for new data points
Y_pred = model.predict(X_test)
```

Subscribe!

```
#Create a Gaussian Classifier
```

```
model = GaussianNB()
```



([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)



([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)
[table\(pred\)](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)
[table\(pred\)](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)
[data = Train\)](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

Above, we looked at the basic Naive Bayes model, you can improve the power of this basic model by tuning parameters and handle assumption intelligently. Let's look at the methods to improve the performance of Naive Bayes Model. I'd recommend you to go through [this document](http://www.inf.ed.ac.uk/teaching/courses/inf2b/learnnotes/inf2b-learn-note07-2up.pdf) (<http://www.inf.ed.ac.uk/teaching/courses/inf2b/learnnotes/inf2b-learn-note07-2up.pdf>) for more details on Text classification using Naive Bayes.

Tips to improve the power of Naive Bayes Model

Here are some tips for improving power of Naive Bayes Model:

Subscribe!

- If continuous features do not have normal distribution, we should use transformation or different methods to convert it in normal distribution.
- If test data set has zero frequency issue, apply smoothing techniques “Laplace Correction” to predict the class of test data set.
- Remove correlated features, as the highly correlated features are voted twice in the model and it can



(<https://trainings.analyticsvidhya.com/courses/course->

<https://trainings.analyticsvidhya.com/courses/course->



Learn (<https://www.analyticsvidhya.com/blog/>), **engage**,

(<https://discuss.analyticsvidhya.com/>), **compete**,

(<https://datahack.analyticsvidhya.com/>) and **get hired**

(<https://www.analyticsvidhya.com/jobs/#/user/>)!

You can also read this article on Analytics Vidhya's Android APP



([https://play.google.com/store/apps/details?](https://play.google.com/store/apps/details?id=com.analyticsvidhya.android&utm_source=blog_article&utm_campaign=blog&pcampaignid=MKT-Other-global-all-co-prtnr-py-PartBadge-Mar2515-1)

[id=com.analyticsvidhya.android&utm_source=blog_article&utm_campaign=blog&pcampaignid=MKT-Other-global-all-co-prtnr-py-PartBadge-Mar2515-1](https://play.google.com/store/apps/details?id=com.analyticsvidhya.android&utm_source=blog_article&utm_campaign=blog&pcampaignid=MKT-Other-global-all-co-prtnr-py-PartBadge-Mar2515-1))

Share this:

Subscribe!

 (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=linkedin&nb=1>)

 (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=facebook&nb=1>)

395

 (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=google-plus-1&nb=1>)



[naive-bayes-explained/?share=twitter&nb=1](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=twitter&nb=1))

[naive-bayes-explained/?share=pocket&nb=1](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=pocket&nb=1))

[naive-bayes-explained/?share=reddit&nb=1](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/?share=reddit&nb=1))

([https://trainings.analyticsvidhya.com/courses/course-](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=Essential+Tools+any+Data+Scientists+can+use+to+improve+their+productivity)

TAGS : MACHINE LEARNING ([HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/MACHINE-LEARNING/](https://www.analyticsvidhya.com/blog/tag/machine-learning/)), NAIVE BAYES

VI:AnalyticsVidhya+DS101+201612/about?



[/E-BAYES/](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)), [NAIVE BAYES IN R](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)

[/E-BAYES-IN-R/](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)), [PROBABILITY](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)

[BABILITY/](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)), [PYTHON](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)

[HON/](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)

NEXT ARTICLE

SAS) – which tool should I learn?

[com/blog/2017/09/sas-vs-vs-python-tool-learn/](https://www.analyticsvidhya.com/blog/2017/09/sas-vs-vs-python-tool-learn/))

...

([https://trainings.analyticsvidhya.com/courses/course-](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=Essential+Tools+any+Data+Scientists+can+use+to+improve+their+productivity)

v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about? PREVIOUS ARTICLE

utm_source=Essential+Tools+any+Data+Scientists+can+use+to+improve+their+productivity

(<https://www.analyticsvidhya.com/blog/2017/09/essential-tools-data-scientist-improve-productivity/>)



(<https://www.analyticsvidhya.com/blog/author/sunil->

Subscribe!

ray/).

Sunil Ray (https://www.analyticsvidhya.com/blog/author/sunil-ray/)

I am a Business Analytics and Intelligence professional with deep experience in the Indian Insurance industry. I have worked for various multi-national Insurance companies in last 7 years.



(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)



(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

(https://www.analyticsvidhya.com/blog/2017/06/introductory-guide-to-deal-imbalanced-classification-problems/)

Introductory guide to Generative Adversarial Networks (GANs) and their promise!

(https://www.analyticsvidhya.com/blog/2017/06/introductory-guide-to-deal-imbalanced-classification-problems/)

This article is quite old and you might not get a prompt response from the author. We request you to post this comment on Analytics Vidhya's [Discussion portal \(https://discuss.analyticsvidhya.com/\)](https://discuss.analyticsvidhya.com/) to get your queries resolved

BLOG (https://www.analyticsvidhya.com/blog/2017/06/introductory-guide-to-deal-imbalanced-classification-problems/)



(https://www.analyticsvidhya.com/blog/2017/06/introductory-guide-to-deal-imbalanced-classification-problems/)

Complete tutorial on Text Classification using Conditional Random Fields Model (in Python)

(https://www.analyticsvidhya.com/blog/2016/03/practical-guide-to-deal-imbalanced-classification-problems/)



(https://www.analyticsvidhya.com/blog/2016/04/deep-learning-computer-vision-introduction-convolution-neural-networks/)

Deep Learning for Computer Vision – Introduction to Convolution Neural Networks

(https://www.analyticsvidhya.com/blog/2016/04/deep-learning-computer-vision-introduction-convolution-neural-networks/)

ANALYTICS VIDHYA CONTENT TEAM (https://www.analyticsvidhya.com/blog/2016/03/practical-guide-to-deal-imbalanced-classification-problems/)



(https://www.analyticsvidhya.com/blog/2016/03/practical-guide-to-deal-imbalanced-classification-problems/)

Practical Guide to deal with Imbalanced Classification Problems in R

(https://www.analyticsvidhya.com/blog/2016/03/practical-guide-to-deal-imbalanced-classification-problems/)



(https://www.analyticsvidhya.com/blog/2016/01/xgboost-algorithm-easy-steps/)

How to use XGBoost algorithm in R in easy steps

(https://www.analyticsvidhya.com/blog/2016/01/xgboost-algorithm-easy-steps/)

Subscribe!

41 COMMENTS



ARUN CR

[Reply](#)

September 14, 2015 at 5:36 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-94915>)



...e [1] we know that

...ve this?

...ng conditional probability. can you suggest such examples.

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)



[Reply](#)

[analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-104496](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-104496)),

...e of conditional probability is the Monty Hall problem.
...lem (https://en.wikipedia.org/wiki/Monty_Hall_problem)

...articular problem because the solution depends on Bayes'

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)



ERDEM KARAKOĞLU

[Reply](#)

March 17, 2016 at 2:26 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-107604>)

It's a trivial example for illustration. The "Likelihood table" (a confusing misnomer, I think) is in fact a probability table that has the JOINT weather and play outcome probabilities in the center, and the MARGINAL probabilities of one variable (from integrating out the other variable from the joint) on the side and bottom.

Say, weather type = w and play outcome = p .

$P(w,p)$ is the joint probabilities and $P(p)$ and $P(w)$ are the marginals. Bayes rule described above by Sunil stems from:

$$P(w,p) = P(w|p) * P(p) = P(p|w) * P(w).$$

From the center cells we have $P(w,p)$ and from the side/bottom we get $P(p)$ and $P(w)$.

[Subscribe!](#)

Depending on what you need to calculate, it follows that:

(1): $P(w|p) = P(w,p) / P(p)$ and

(2): $P(p|w) = P(w,p) / P(w)$, which is what you did with $P(\text{sunny}, \text{yes}) = 3/14$ and $P(w) = 5/14$, yielding $(3/14) / (5/14)$, with the 14's cancelling out.



of the two quantities above, $P(w|p)$ or $P(p|w)$ is much harder to calculate. But you'll come to see this as one of two mathematical miracles made possible by the flexibility of Markov Chain Monte Carlo in circumventing some of these difficulties. I digress.

[Reply](#)

<https://trainings.analyticsvidhya.com/courses/course-113751>

<https://trainings.analyticsvidhya.com/courses/course-113751>

What is the answer? Thanks!



[Reply](#)

www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-94918



[Reply](#)

www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-94918

<https://trainings.analyticsvidhya.com/courses/course-113751>

Amazing content and useful information

utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)



LEENA

[Reply](#)

February 8, 2016 at 5:59 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-105356>)

I'm new to machine learning and Python. Could you please help to read data from CSV and to separate the same data set to training and test data



ISMAEL EZEQUIEL ([HTTPS://GITHUB.COM/ISMAELEZEQUIEL](https://github.com/ISMAELEZEQUIEL))

[Reply](#)

October 15, 2016 at 12:12 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-117171>)

import pandas as pd

[Subscribe!](#)

```

person = pd.read_csv('example.csv')
mask = np.random.rand(len(sales)) < 0.8
train = sales[mask]
test = sales[~mask]

```

[Reply.](#)<https://analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-106807>[Reply.](#)<https://analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-107493>

please give me that code in JAVA

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?>

[Reply.](#)<https://analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-108955>

data mining tool??

[Reply.](#)<https://analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-109291>

(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?

utm_source=CVDL101Banner&utm_medium=Sidebar&utm_campaign=CV101banner)

[Reply.](#)

DEVENIR RICHE AU MAROC (HTTP://WWW.COACHOUTLET-FACTORYSTORE3.COM/2015/10/25/JEUX-D'ARGENT-VS-TRADING/)

April 14, 2016 at 5:13 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-109474>).

I am really impressed together with your writing skills as wwell as with the format to your weblog.

Is that this a paid theme or did you customiz it yourself?

Anyway stay up the excellent quality writing, it's uncommon tto see a great weblog like this one these days..



รับซื้อ PATEK PHILIPPE (HTTPS://WWW.FACEBOOK.COM/SHINBRANDNAME)

[Reply.](#)

April 22, 2016 at 3:52 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-109828>).

Subscribe!

You should be a part of a contest for one of the most useful websites online.
I will highly recommend this blog!


[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-109857>).

f models, that 's really precious experience.

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-110810>)

e require a CA OCF-1 , We used a sample document here

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=EV-1201AV-Edg&utm_medium=Stack&utm_campaign=Clarifai+Banner)

[Reply](#)


<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-110816>)

tribution, we should use transformation or different methods to

techniques?

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=EV-1201AV-Edg&utm_medium=Stack&utm_campaign=Clarifai+Banner)

[Reply](#)

<http://www.bangladeshilivetelevision.com/>)
utm_source=EV-1201AV-Edg&utm_medium=Stack&utm_campaign=Clarifai+Banner

Can I simply just say what a comfort too find someone
who actually knows what they're talking about over the
internet. You certainly know how to bring an issue to light and make it important.

A lot more people ought to read this and understand this side of your story.
I can't believe you aren't more popular because you certainly have the gift.



NIR

[Reply](#)

July 13, 2016 at 6:57 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-113416>)

Subscribe!

Great article! Thanks. Are there any similar articles for other classification algorithms specially target towards textual features and mix of textual/numeric features?



[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-114946>

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-115295>

t. Thank you!

([https://training.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://training.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=AnalyticsVidhya&utm_medium=StickyBanner&utm_campaign=CV101banner)

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-115354>



[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-115357>

ned! Loved this article.

[Reply](#)

([https://training.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?](https://training.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=AnalyticsVidhya&utm_medium=StickyBanner&utm_campaign=CV101banner)
The 'y' should be capitalized in your code – great article though.
`utm_source=CV101AVBlogBanner&utm_medium=StickyBanner2utm_campaign=CV101banner`)



AKASH SWAMY

[Reply](#)

September 6, 2016 at 7:24 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-115661>)

This is the best explanation of NB so far simple and short 😊



JOHN (HTTPS://FLOOFYDUGONG.GITHUB.IO/)

[Reply](#)

September 9, 2016 at 5:20 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-115863>)

Great article! Really enjoyed it. Just wanted to point out a small error in the Python code.

[Subscribe!](#)

Should be a capital "Y" in the predict like so : `model.fit(x, Y)`

Thanks!



[Reply](#)

www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-116232)

used as a newbie. Can you please guide?

[Reply](#)

[analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-117314](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-117314))

concept

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+2018T2/about?>

[Reply](#)



[analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-117495](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-117495))

was handy to me in understanding naive bayes especially the

[Reply](#)

www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-118208)

Thanks to you I can totally understand NB classifier.
(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?

`utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner`)

[Reply](#)



T B

July 3, 2017 at 12:17 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-131449>)

Really nice article, very use-full for concept building.



AKSHAY

[Reply](#)

July 4, 2017 at 8:40 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-131533>)

I didn't understand the 3rd step. Highest probability out of which probability values?

[Subscribe!](#)

>> Now, $P(\text{Yes} \mid \text{Sunny}) = 0.33 * 0.64 / 0.36 = 0.60$, which has higher probability.
Higher than what?

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-133229>),

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-135803>)

of this concept

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?utm_source=AnalyticsVidhya&utm_medium=StickyBanner&utm_campaign=DS101Banner)

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-144881>



documents classification using naive base algorithm.

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-144881>

the simplicity.Thanks for the effort.

([https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVL101+CVDL101_T1/about?](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVL101+CVDL101_T1/about?utm_source=AnalyticsVidhya&utm_medium=StickyBanner&utm_campaign=CVL101Banner)

[Reply](#)

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-144881>)

Good start point for beginners



ABDUL SAMAD

[Reply](#)

April 12, 2018 at 10:19 am (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-152525>)

Weldone sanil

I have a question regarding naive bayes,currently i am working on a project that is detect depression through naive bayes algorithm so plz suggest few links regarding my projects.i shall be gratefull to you.

Thanku so much

Subscribe!



AISHWARYA SINGH

[Reply](#)

April 16, 2018 at 4:37 pm (<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-152600>).



ed=0ahUKEwix5veO077aAhWCrY8KHeWeBj8QFghkMAQ&url=
w0RSIYviBCcjh9efgoKhKMz) link.

[Reply](#)

icsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-153048).

les. Can someone help me

([https://trainings.analyticsvidhya.com/courses/course-](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?)
v1:AnalyticsVidhya+DS101+2018T2/about?

AISHWARYA SINGH

[Reply](#)



icsvidhya.com/blog/2017/09/naive-bayes-explained/#comment-153068).

d y is used for target variable (which is to be predicted).

([https://trainings.analyticsvidhya.com/courses/course-](https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?)
v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?

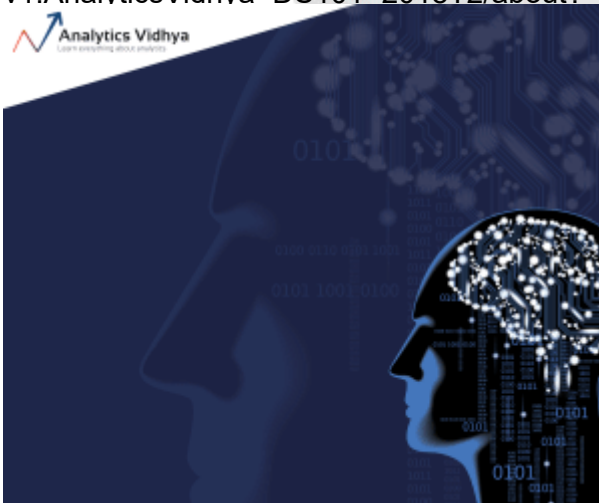
utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

Subscribe!



(<https://www.analyticsvidhya.com/datahack-summit-2018/>?)

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?>



(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

POPULAR POSTS

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 Tutorial to Learn Data Science with Python from Scratch
(<https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-learn-data-science-python-scratch-2/>)

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

Subscribe!

Understanding Support Vector Machine algorithm from examples (along with code)

(<https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>)

7 Types of Regression Techniques you should know!

(<https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/>)



thm (with codes in Python and R)

([2017/09/naive-bayes-explained/](https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/))

te a Time Series Forecast (with Codes in Python)

([2016/02/time-series-forecasting-codes-python/](https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/))

ified (with implementation in Python)

([2018/03/introduction-k-neighbours-algorithm-clustering/](https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/))

(<https://trainings.analyticsvidhya.com/courses/course->

W1:AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about? a Comprehensive List of Resources to get Started

([2018/11/data-engineer-comprehensive-list-resources-get-](https://www.analyticsvidhya.com/blog/2018/11/data-engineer-comprehensive-list-resources-get-)

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

-CNN Algorithm for Object Detection (Part 2 – with Python

([blog/2018/11/implementation-faster-r-cnn-python-object-](https://www.analyticsvidhya.com/blog/2018/11/implementation-faster-r-cnn-python-object-)

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

ies & Reddit Discussions (October 2018)

([2018/11/best-machine-learning-github-repositories-reddit-](https://www.analyticsvidhya.com/blog/2018/11/best-machine-learning-github-repositories-reddit-)

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

AnalyticsVidhya+DS101+2018T2/about?

(<https://www.analyticsvidhya.com/courses/course->

W1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

AnalyticsVidhya+CVDL101+CVDL101_T1/about?

An Introduction to Text Summarization using the TextRank Algorithm (with Python implementation)

(<https://www.analyticsvidhya.com/blog/2018/11/introduction-text-summarization-textrank-python/>)

NOVEMBER 1, 2018

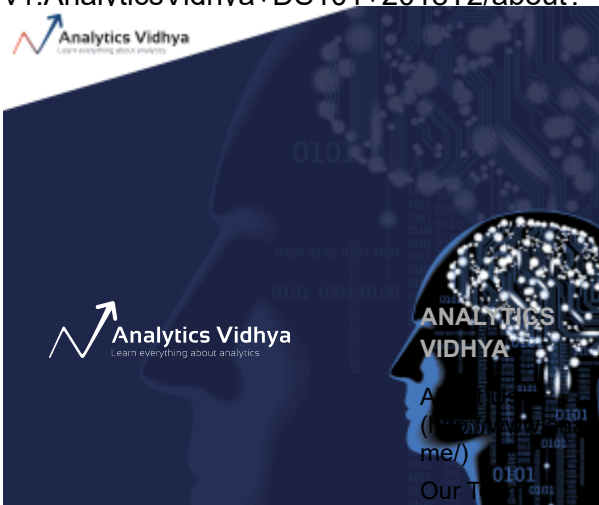
Subscribe!



(<http://www.edvancer.in/certified-data-scientist-with-python->

[s&utm_campaign=AVadsnonfc&utm_content=pythonavad\)](#)

(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?>



(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?

[utm_source=CV101AVBlogBanner&utm_medium=BlogBanner&utm_campaign=CV101AVBlogBanner">utm_source=CV101AVBlogBanner&utm_medium=BlogBanner&utm_campaign=CV101AVBlogBanner](#)

<p>ANALYTICS VIDHYA</p> <p>Learn everything about analytics</p> <p>DATA SCIENTISTS</p> <p>Blog (https://www.analyticsvidhya.com/blog/)</p> <p>Hackathon (https://datahack.analyticsvidhya.com/)</p> <p>Our Team (https://www.analyticsvidhya.com/about-me/team/)</p> <p>Career (https://www.analyticsvidhya.com/career/)</p> <p>Apply Jobs (https://www.analyticsvidhya.com/apply-jobs/)</p> <p>Leaderboard (https://datahack.analyticsvidhya.com/leaderboard/)</p> <p>Contact Us (https://www.analyticsvidhya.com/contact/)</p> <p>Write for us (https://www.analyticsvidhya.com/about-me/write/)</p>	<p>COMPANIES</p> <p>Post Jobs (https://www.analyticsvidhya.com/corporate/)</p> <p>Blog (https://www.analyticsvidhya.com/blog/)</p> <p>Hackathon (https://trainings.analyticsvidhya.com/)</p> <p>Our Team (https://www.analyticsvidhya.com/about-me/team/)</p> <p>Career (https://www.analyticsvidhya.com/career/)</p> <p>Apply Jobs (https://www.analyticsvidhya.com/apply-jobs/)</p> <p>Leaderboard (https://datahack.analyticsvidhya.com/leaderboard/)</p> <p>Contact Us (https://www.analyticsvidhya.com/contact/)</p> <p>Write for us (https://www.analyticsvidhya.com/about-me/write/)</p>	<p>JOIN OUR COMMUNITY :</p> <p>f (https://www.facebook.com/analyticsvidhya/) 18556 followers</p> <p>t (https://twitter.com/analyticsvidhya) 18556 followers</p> <p>G+ (https://plus.google.com/+AnalyticsVidhya) 3016 followers</p> <p>in (https://www.linkedin.com/company/analyticsvidhya/) 7513 followers</p> <p>Subscribe to emailer <input type="button" value="Subscribe!"/></p>
--	--	--

© Copyright 2013-2018 Analytics Vidhya.

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

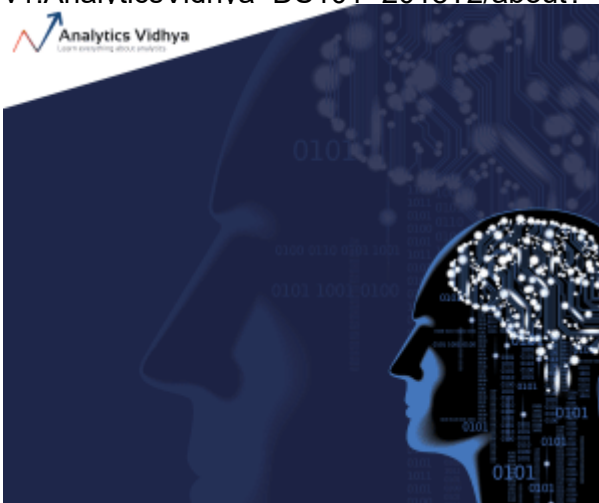
Don't have an account? Sign up (h

Terms of Use (<https://www.analyticsvidhya.com/terms/>)Refund Policy (<https://www.analyticsvidhya.com/refund-policy/>)

u



(<https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+DS101+2018T2/about?>



(https://trainings.analyticsvidhya.com/courses/course-v1:AnalyticsVidhya+CVDL101+CVDL101_T1/about?utm_source=CV101AVBlogBanner&utm_medium=Stickybanner2utm_campaign=CV101banner)

[Subscribe!](#)