×

-

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

7 Regression Techniques you should know!

SUNIL RAY (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/AUTHOR/SUNIL-RAY/), AUGUST 14, 2015

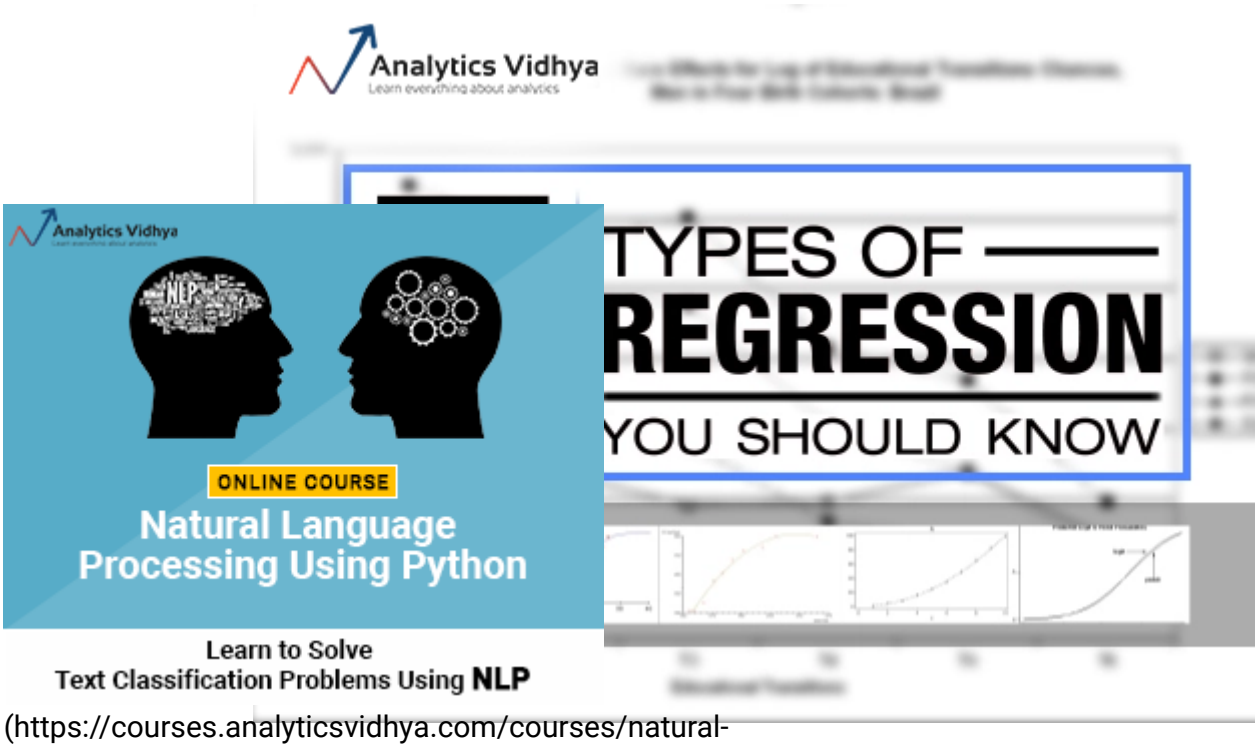
Overview

- Learn about the different regression types in machine learning (https://www.analyticsvidhya.com/machine-learning/?utm_source=blog&utm_medium=7-regression-techniques), including linear and logistic regression
- Each regression technique has its own regression equation and regression coefficients
- We cover 7 different regression types in this article

e first algorithms people learn in data science
s/introduction-to-data-science-2/?
typesarticle). Due to their popularity, a lot of analysts even end up
ions. The ones who are slightly more involved think that they are
ssion analysis.

There are many forms of regressions, which can be performed. Each form has its own importance and specific condition where they are best suited to apply. In this article, I have explained the most commonly used regressions (https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp?utm_source=blog&utm_medium=displayad) in data science
s/introduction-to-data-science-2/?
typesarticle) in a simple manner.

e develop an idea of the breadth of regressions, instead of just
ry machine learning (https://www.analyticsvidhya.com/machine-
regression-techniques) problem they come across and hoping that



(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp?utm_source=blog&utm_medium=typesarticle)

(https://s3-ap-south-1.amazonaws.com/av-blog-media/wp-content/uploads/2015/08/7-type-regression.png)

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master?utm_source=blog&utm_medium=Stickybanner2)

isplay&utm_campaign=NLPcourse)

for a place to start your journey, the 'data science

s/introduction-to-data-science-2/?

typesarticle)' course is as good a place as any to start! Covering the

e Modeling, it is the perfect way to take your first steps into data

1. What is Regression Analysis?
2. Why do we use Regression Analysis?
3. What are the types of Regressions?
    - Linear Regression
    - Logistic Regression
    - Polynomial Regression
    - Stepwise Regression
    - Ridge Regression
    - Lasso Regression
    - ElasticNet Regression
4. How to select the right Regression Model?

What is Regression Analysis?

modelling technique which investigates the relationship between a **(s)** (predictor). This technique is used for forecasting, time series ionship (https://www.analyticsvidhya.com/blog/2015/06/establish-r example, relationship between rash driving and number of road regression.


t/uploads/2015/08/Regression_Line.png)

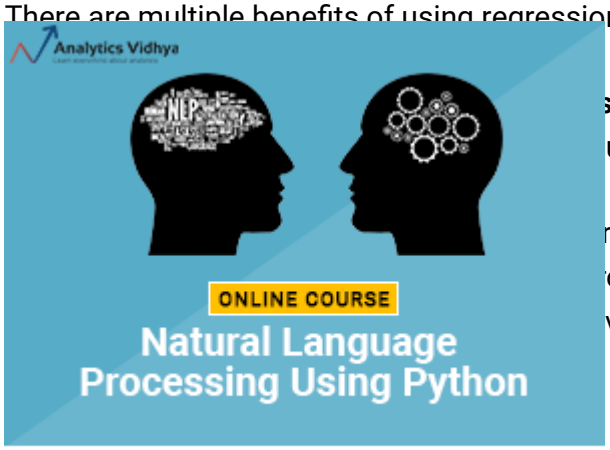
Regression analysis is an important tool for modelling and analyzing data. Here, we fit a curve / line to the data points, in such a manner that the differences between the distances of data points from the curve or line is minimized. I'll explain this in more details in coming sections.

Why do we use Regression Analysis?

As mentioned above, regression analysis estimates the relationship between two or more variables. Let's understand this with an easy example:

Let's say, you want to estimate growth in sales of a company based on current economic conditions. You have the recent company data which indicates that the growth in sales is around two and a half times the growth in the economy. Using this insight, we can predict future sales of the company based on current & past information.

There are multiple benefits of using regression analysis. They are as follows:



(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?utm_source=blog&utm_medium=Stickybanner1&utm_campaign=NLP_course)

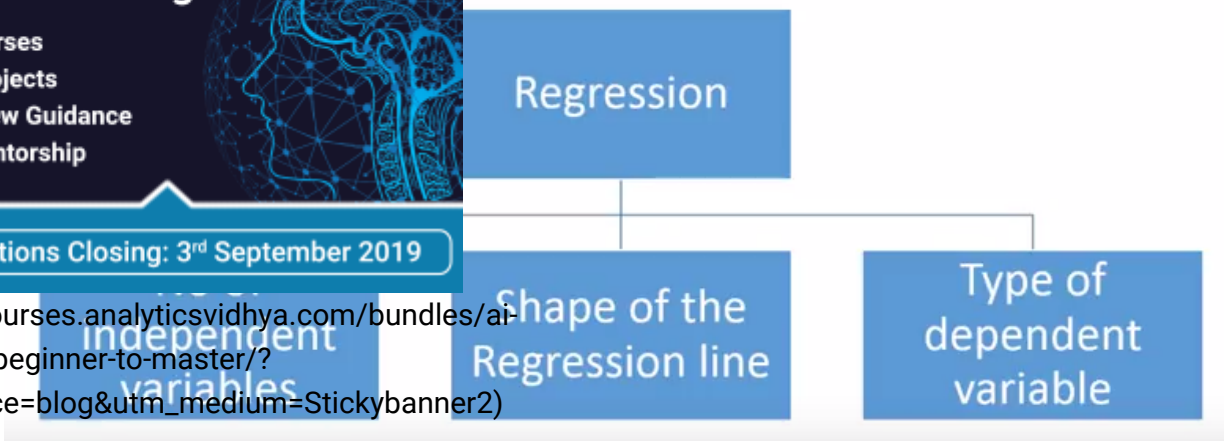
s between dependent variable and independent variable.
ultiple independent variables on a dependent variable.

re the effects of variables measured on different scales, such as the omotional activities. These benefits help market researchers / data valuate the best set of variables to be used for building predictive

How many types of regression techniques do we have?

ies available to make predictions. These techniques are mostly nt variables, type of dependent variables and shape of regression ing sections.

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)



(https://www.analyticsvidhya.com/wp-content/uploads/2015/08/Regression_Type.png)

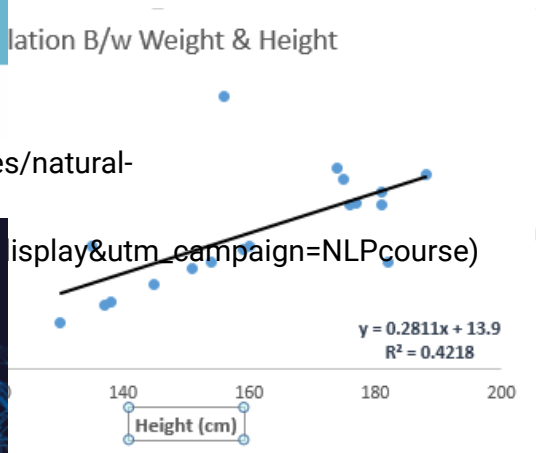
For the creative ones, you can even cook up new regressions, if you feel the need to use a combination of the parameters above, which people haven't used before. But before you start that, let us understand the most commonly used regressions:

1. Linear Regression

It is one of the most widely known modeling technique. Linear regression is usually among the first few topics which people pick while learning predictive modeling. In this technique, the dependent variable is continuous, independent variable(s) can be continuous or discrete d_discrete_variables), and nature of regression line is linear.

nship between **dependent variable (Y)** and one or more **ight line** (also known as regression line).

, where a is intercept, b is slope of the line and e is error term. This arget variable based on given predictor variable(s).



t/uploads/2015/08/Linear_Regression1.png)

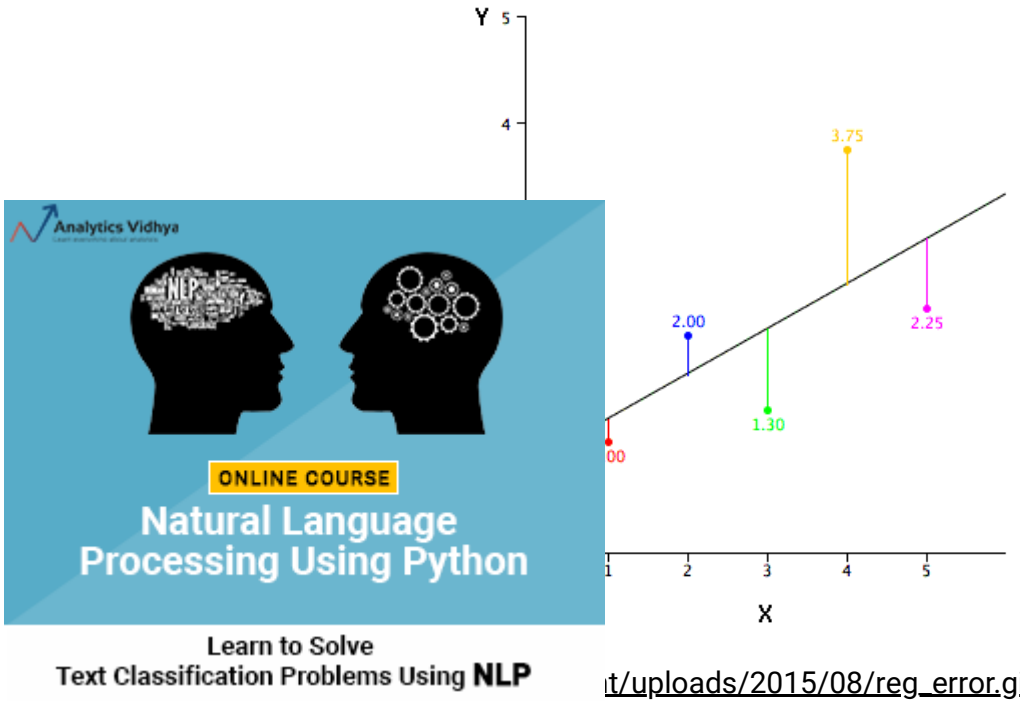
ion and multiple linear regression is that, multiple linear regression le linear regression has only 1 independent variable. Now, the

How to obtain best fit line (Value of a and b)?

This task can be easily accomplished by Least Square Method. It is the most common method used for fitting a regression line. It calculates the best-fit line for the observed data by minimizing the sum of the squares of the vertical deviations from each data point to the line. Because the deviations are first squared, when added, there is no cancelling out between positive and negative values.

$$\min_{w} ||Xw - y||_2^2$$

(https://www.analyticsvidhya.com/wp-content/uploads/2015/08/Least_Square.png)



t/uploads/2015/08/reg_error.gif)

We can evaluate the model performance using the metric **R-square**. To know more details about these metrics, you can read: [Model Performance metrics Part 1, Part 2](https://www.analyticsvidhya.com/blog/2015/01/model-performance/) (https://www.analyticsvidhya.com/blog/2015/01/model-perform-2/)

een independent and dependent variables

ollinearity, autocorrelation, heteroskedasticity**.**

Outliers. It can terribly affect the regression line and eventually the

iance of the coefficient estimates and make the estimates very sensitive to minor changes in the model. The result is that the coefficient estimates are unstable

- In case of multiple independent variables, we can go with **forward selection**, **backward elimination** and **step wise approach** for selection of most significant independent variables.

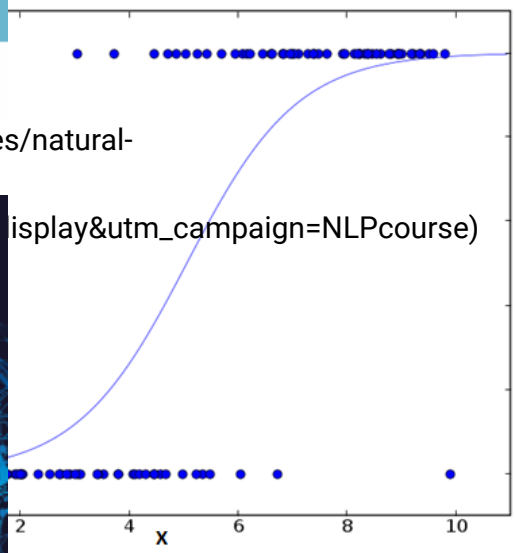
2. Logistic Regression

Logistic regression is used to find the probability of event=Success and event=Failure. We should use logistic regression when the dependent variable is binary (0/ 1, True/ False, Yes/ No) in nature. Here the value of Y ranges from 0 to 1 and it can represented by following equation.

```
odds= p/ (1-p) = probability of event occurrence / probability of not event occurrence
ln(odds) = ln(p/(1-p))
logit(p) = ln(p/(1-p)) = b0+b1X1+b2X2+b3X3....+bkXk
```


he characteristic of interest. A question that you should ask here

istribution (dependent variable), we need to choose a link function And, it is **logit (https://en.wikipedia.org/wiki/Logistic_function)** ers are chosen to maximize the likelihood of observing the sample ared errors (like in ordinary regression).



t/uploads/2015/08/Logistic_Regression.png)

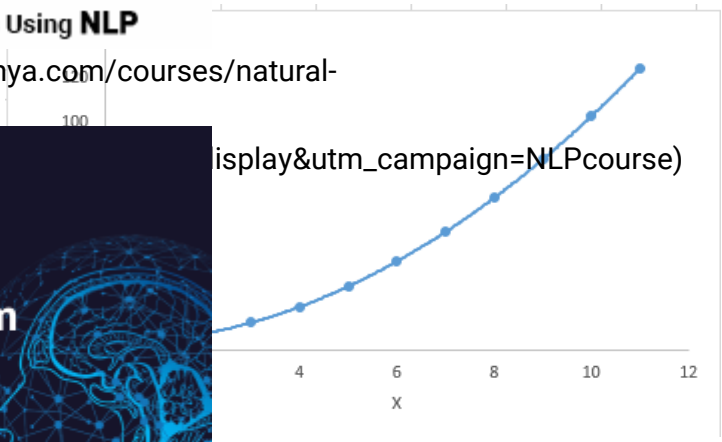
- It is widely used for **classification problems**
- Logistic regression doesn't require linear relationship between dependent and independent variables. It can handle various types of relationships because it applies a non-linear log transformation to the predicted odds ratio
- To avoid over fitting and under fitting, we should include all significant variables. A good approach to ensure this practice is to use a step wise method to estimate the logistic regression
- It requires **large sample sizes** because maximum likelihood estimates are less powerful at low sample sizes than ordinary least square
- The independent variables should not be correlated with each other i.e. **no multi collinearity**. However, we have the options to include interaction effects of categorical variables in the analysis and in the model.

- If the values of dependent variable is ordinal, then it is called as **Ordinal logistic regression**
- If dependent variable is multi class then it is known as **Multinomial Logistic regression**.

sion equation if the power of independent variable is more than 1. equation:

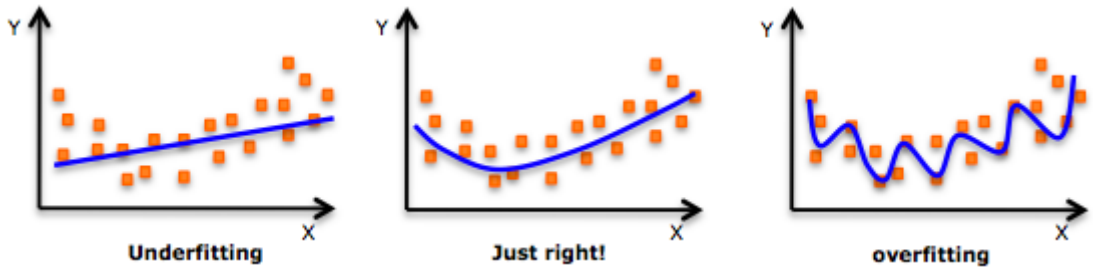
```
y=a+b*x^2
```

s not a straight line. It is rather a curve that fits into the data points.



t/uploads/2015/08/Polynomial.png)

- While there might be a temptation to fit a higher degree polynomial to get lower error, this can result in over-fitting. Always plot the relationships to see the fit and focus on making sure that the curve fits the nature of the problem. Here is an example of how plotting can help:



(https://www.analyticsvidhya.com/wp-content/uploads/2015/02/underfitting-overfitting.png)

- Especially look out for curve towards the ends and see whether those shapes and trends make sense. Higher polynomials can end up producing wierd results on extrapolation.

with multiple independent variables. In this technique, the selection of an automatic process, which involves *no* human intervention.

l values like R-square, t-stats and AIC metric to discern significant the regression model by adding/dropping co-variates one at a time ost commonly used Stepwise regression methods are listed below:

things. It adds and removes predictors as needed for each step.

nificant predictor in the model and adds variable for each step.

- Backward elimination starts with all predictors in the model and removes the least significant variable for each step.

isplay&utm_campaign=NLPcourse)

aximize the prediction power with minimum number of predictor the method to handle higher dimensionality /07/dimension-reduction-methods/) of data set.

the data suffers from multicollinearity (independent variables are though the least squares estimates (OLS) are unbiased, their variances are large which deviates the observed value far from the true value. By adding a degree of bias to the regression estimates, ridge regression reduces the standard errors.

Above, we saw the equation for linear regression. Remember? It can be represented as:

y=a+ b*x

This equation also has an error term. The complete equation becomes:

```
y=a+b*x+e (error term),  [error term is the value needed to correct for a prediction error between th
e observed and predicted value]
```

```
=> y=a+y= a+ b1x1+ b2x2+....+e, for multiple independent variables.
```

In a linear equation, prediction errors can be decomposed into two sub components. First is due to the **biased** and second is due to the **variance**. Prediction error can occur due to any one of these two or both components. e to variance.

ulticollinearity problem through shrinkage parameter mator) λ (lambda). Look at the equation below.

$$\underset{\beta \in \mathbb{R}^p}{\text{gmin}} \underbrace{\|y - X\beta\|_2^2}_{\text{Loss}} + \lambda \underbrace{\|\beta\|_2^2}_{\text{Penalty}}$$

(https://www.analyticsvidhya.com/wp-content/uploads/2015/08/Ridge2.png)

First one is least square term and other one is lambda of the he coefficient. This is added to least square term in order to shrink

ame as least squared regression except normality is not to

oesn't reaches zero, which suggests no feature selection feature s l2 regularization zation_(mathematics)).

6. Lasso Regression

$$= \underset{\beta \in \mathbb{R}^p}{\text{argmin}} \underbrace{\|y - X\beta\|_2^2}_{\text{Loss}} + \lambda \underbrace{\|\beta\|_1}_{\text{Penalty}}$$

(https://www.analyticsvidhya.com/wp-content/uploads/2015/08/Lasso.png)

Similar to Ridge Regression, Lasso (Least Absolute Shrinkage and Selection Operator) also penalizes the absolute size of the regression coefficients. In addition, it is capable of reducing the variability and improving the accuracy of linear regression models. Look at the equation below: Lasso regression differs from ridge regression in a way that it uses absolute values in the penalty function, instead of squares. This leads to penalizing (or equivalently constraining the sum of the absolute values of the estimates) values which to turn out exactly zero. Larger the penalty applied, further the This results to variable selection out of given n variables.

ame as least squared regression except normality is not to

ero), which certainly helps in feature selection
s l1 regularization
ation_(mathematics))
ted, lasso picks only one of them and shrinks the others to zero

Regression techniques. It is trained with L1 and L2 prior as re multiple features which are correlated. Lasso is likely to pick one o pick both.

$$\hat{\beta} \equiv \operatorname{argmin}(\|y - X\beta\|^2 + \lambda_2\|\beta\|^2 + \lambda_1\|\beta\|_1).$$

t/uploads/2015/08/Elastic_Net.png)

A practical advantage of trading-off between Lasso and Ridge is that, it allows Elastic-Net to inherit some of Ridge's stability under rotation.

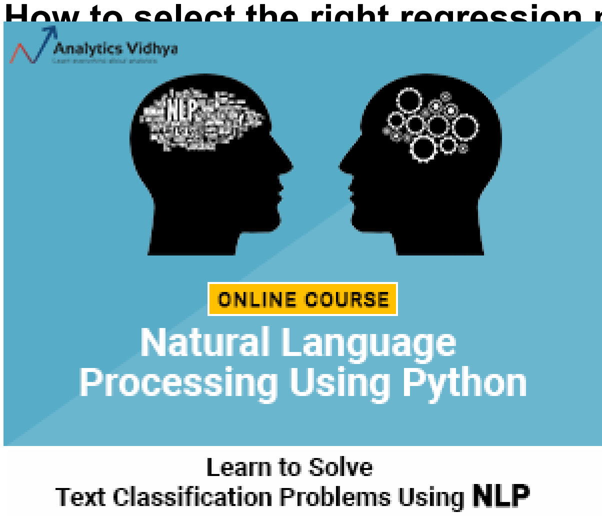
Important Points:

- It encourages group effect in case of highly correlated variables
- There are no limitations on the number of selected variables
- It can suffer with double shrinkage

Beyond these 7 most commonly used regression techniques, you can also look at other models like Bayesian (https://en.wikipedia.org/wiki/Bayesian_linear_regression), Ecological (https://en.wikipedia.org/wiki/Ecological_regression) and Robust regression

(https://en.wikipedia.org/wiki/Robust_regression).

How to select the right regression model?



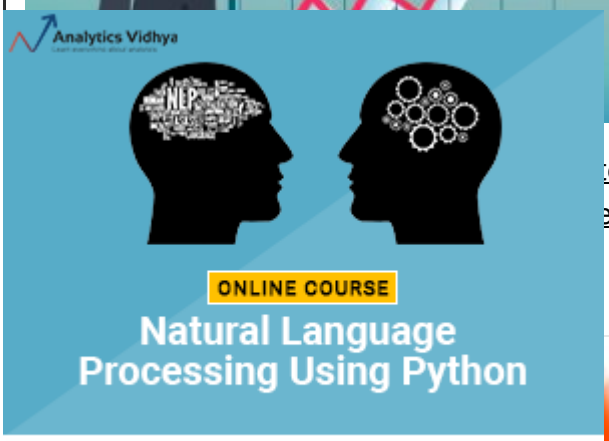
…ne or two techniques. One of the training institutes I know of tells …s – apply linear regression. If it is binary – use logistic regression! …able at our disposal, more difficult it becomes to choose the right …models.

…t is important to choose the best suited technique based on type of …sionality in the data and other essential characteristics of the data. …ctice to select the right regression model:

1. … building predictive model. It should be you first step before selecting the right model like identify the relationship and impact of variables
2. To compare the goodness of fit for different models, we can analyse different metrics like statistical …djusted r-square, AIC, BIC and error term. Another one is the …m/en-us/minitab/17/topic-library/modeling-statistics/regression-…cs/what-is-mallows-cp/) criterion. This essentially checks for …ing the model with all possible submodels (or a careful selection of
3. …uate models used for prediction. Here you divide your data set into … mean squared difference between the observed and predicted …ction accuracy.
4. …g variables, you should not choose automatic model selection …t these in a model at the same time.
5. …n occur that a less powerful model is easy to implement as compared to a highly statistically significant model.
6. Regression regularization methods(Lasso, Ridge and ElasticNet) works well in case of high dimensionality and multicollinearity among the variables in the data set.

Projects

Now, its time to take the plunge and actually play with some other real datasets. Try the techniques learnt in this post on the datasets provided in the following practice problems and let us know in the comment section how it worked out for you!

  	Practice Problem: Food Demand Forecasting Challeng (https://datahack.analyticsvidhya.com/contest/genpa machine-learning-hackathon-1/?utm_source=complet tutorial-learn-data-science-python-scratch-2&utm_medium=blog)
test/genpact- e=complete-	Practice Problem: Predict Number of Upvotes (https://datahack.analyticsvidhya.com/contest/enigm codefest-machine-learning-1/?utm_source=complete- tutorial-learn-data-science-python-scratch-2&utm_medium=blog)

(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?utm_source=blog&utm_medium=displayBanner2&utm_campaign=NLPcourse)

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master?utm_source=blog&utm_medium=Stickybanner2)

test/enigma- complete-

view of regression. These regression techniques should be applied considering the conditions of data. One of the best trick to find out which technique to use, is by checking the family of variables i.e. discrete or continuous.

In this article, I discussed about 7 types of regression and some key facts associated with each technique. As somebody who's new in this industry, I'd advise you to learn these techniques and later implement them in your models.

Did you find this article useful ? Share your opinions / views in the comments section below.

Note – The discussions of this article are going on at AV's Discuss portal. Join here (https://discuss.analyticsvidhya.com/t/discussions-for-article-7-types-of-regression-techniques-you-should-know/65230?u=jalfaizy)!

If you like what you just read & want to continue your analytics learning, subscribe to our emails (http://feedburner.google.com/fb/a/mailverify?uri=analyticsvidhya), follow us on twitter (http://twitter.com/analyticsvidhya) or like our facebook page (http://facebook.com/analyticsvidhya).

lhya's Android APP GET IT ON Google Play

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

mprehensive-guide-regression/?share=linkedin&nb=1)

mprehensive-guide-regression/?share=facebook&nb=1)

(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?

com/blog/2015/08/comprehensive-guide-regression/?share=twitter&nb=1)

isplay&utm_campaign=NLPcourse)

mprehensive-guide-regression/?share=pocket&nb=1)

mprehensive-guide-regression/?share=reddit&nb=1)

YTICSVIDHYA.COM/BLOG/TAG/ELASTICNET-REGRESSION/), LASSO REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/LASSO-REGRESSION/), LINEAR REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/LINEAR-REGRESSION/), LOGISTIC REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/LOGISTIC-REGRESSION/), MULTINOMIAL LOGISTIC RGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/MULTINOMIAL-LOGISTIC-RGRESSION/), MULTIPLE REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/MULTIPLE-REGRESSION/), ORDINAL LOGISTIC REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/ORDINAL-LOGISTIC-REGRESSION/), POLYNOMIAL REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/POLYNOMIAL-REGRESSION/), REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/REGRESSION/), REGRESSION ANALYSIS (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/REGRESSION-ANALYSIS/), REGRESSION MODELS (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/REGRESSION-MODELS/), REGULARIZATION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/REGULARIZATION/), RIDGE REGRESSION

(HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/RIDGE-REGRESSION/), STEPWISE REGRESSION (HTTPS://WWW.ANALYTICSVIDHYA.COM/BLOG/TAG/STEPWISE-REGRESSION/)

NEXT ARTICLE

should know as a Data Scientist

g/2015/08/ready-data-science-resources-common-questions-answered/)

•••

PREVIOUS ARTICLE

nches Analytics Program in Bangalore, India

/blog/2015/08/great-lakes-analytics-program-bangalore/)

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

alyticsvidhya.Com/Blog/Author/Sunil-Ray/)

nce professional with deep experience in the Indian Insurance industry. I have worked for various multi-national Insurance companies in last 7 years.

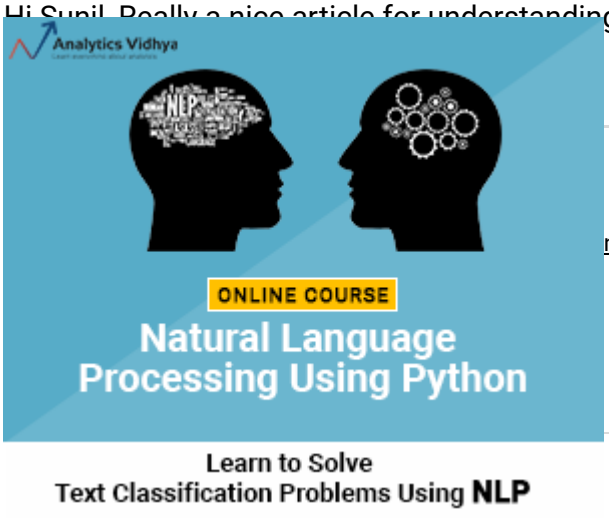
This article is quite old and you might not get a prompt response from the author. We request you to post this comment on Analytics Vidhya's **Discussion portal** (https://discuss.analyticsvidhya.com/) to get your queries resolved

34 COMMENTS

PRAKASHPCS Reply

August 14, 2015 at 6:02 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92695)

Hi Sunil, Really a nice article for understanding the regression models. Especially for novice like me who are



Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?utm_source=blog&utm_medium=display&utm_campaign=NLPcourse)

Reply

August 14, 2015 at 6:24 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

ebsite) where I can understand concept underlying in such

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

You can read book “The Elements of Statistical Learning”, it has detailed explanation of these regression models.

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)

Regards,
Sunil

JULIUS MKUMBO Reply

August 14, 2015 at 9:23 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92724)

I agree with you Sunil, but before reading “The Elements of Statistical Learning”, I would recommend reading An Introduction to Statistical Learning: with application in R, which is more practical because you have to practise with R codes, or you may take Statistical Learning course which is offered by authors of these books, in addition they are inventors of some of these model as well (e.g. Lasso by Tibshirani).

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

TOM

August 31, 2015 at 8:49 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

Reply

n a technique used so often but underutilised when looking at the rested in doing something similar for classification

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

92726)

Thanks Tom , you can refer article on most common machine learning algorithms http://www.analyticsvidhya.com/blog/2015/08/common-machine-learning-algorithms/ (http://www.analyticsvidhya.com/blog/2015/08/common-machine-learning-algorithms/). Here I have discussed various types of classification algorithms like decision tree, random forest, KNN, Naive Bayes…

Regards,
Sunil

R RAJ KUMAR

Reply

August 14, 2015 at 9:14 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92722)

Dear sir,

sults of a college for the coming academic year.

R.COM/PRATZJOSHI) Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

n and multiple regression needs correction. When there is just one called “simple linear regression” not just linear regression.

(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?

isplay&utm_campaign=NLPcourse) Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)

ROHAN POTA Reply

August 14, 2015 at 11:47 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92735)

Hey,quite nice article.It did help me broaden my perspective regarding the regression techniques(specially ElasticNet),but still it would be nice to elucidate upon the differences between l1 and l2 regularization techniques.For this,http://www.quora.com/What-is-the-difference-between-L1-and-L2-regularization (http://www.quora.com/What-is-the-difference-between-L1-and-L2-regularization) will be very helpful.Though it could be incorporated into a new article I think.

PAUL **Reply**

August 14, 2015 at 1:17 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92743)

I'm sorry but I am going to complain again

o to save it or print it, it is a mess!

my network, all the ads and hypertext links cover the article text;

se a button to convert the article to a PDF, which can them be

d in once, then stopped. Please start again!

Reply

August 14, 2015 at 6:27 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

u have a comprehensive data set upon which we can apply all/few

h regression behaves…thanks again

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

You can look at scikit-learn example data sets.

Regards
Sunil

LALIT SACHAN **Reply**

August 17, 2015 at 5:33 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-92929)

Hi Sunil,

Nice compilation. Suggesting a correction , elastic net penalty has another parameter too and is written as lambda * summation ([alpha * L2 penalty + (1-alpha)* L1 Penalty])….

Also quoting book by trevor & hastie “The elastic-net selects variables like the lasso, and shrinks together the coefficients of correlated predictors like ridge.”

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

Can you please explain this point mentioned under the logistic regression – multi collinearity part “However, we have the options to include interaction effects of categorical variables in the analysis and in the model.”

isplay&utm_campaign=NLPcourse)

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

nteraction in detail.

g-interactions-in-regression/

ng-interactions-in-regression/)

Regards,
Sunil

GAURAV (HTTP://WWW.ANALYTICSVIDHYA.COM/BLOG/2015/08/COMPREHENSIVE-GUIDE-REGRESSION/) **Reply**

September 23, 2015 at 9:34 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-95801)

Hi sunil,
The article seems very interesting. Please can you let me know how can we implement Forward stepwise Regression in python as we dont have any inbuilt lib for it.

Reply

analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

students.please provide this and further articles in pdf.thank you.

Reply

w.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?utm_source=blog&utm_medium=display&utm_campaign=NLPcourse)

Reply

analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

ou really think that “stepwise regression” is a type of regression? In thods for selecting relevant variables. And it is performed by echnics (linear, polynomial, ridge or lasso…).

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-mastery?utm_source=blog&utm_medium=Stickybanner2)

MUKUND Reply
January 13, 2016 at 9:30 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-103147)

Very useful article. Are there any specific types of regression techniques which can be used for a time series stationary data?

BHANUSHREE Reply
February 3, 2016 at 1:41 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-105116)

Hi,

Very nice article, crisp n neat! Thank you ☺
In 2. Logistic regression
We use log because while calculating MLE (Maximum Likelihood Estimate) it is easy to differentiate and equate

st (a * b)

Reply

nalyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

(just love this article ⓘ very very useful ⓘ

isplay&utm_campaign=NLPcourse)

ETECHGUYS.COM) **Reply**

ticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-110274)

what I learnt 35 years ago. Excellent. Professional practicing today
no option to choose Data Analytics as a popular profession in 80's.
es. Compliment to you for such a vast subject so lucidly worded

a tutor teaching students in an institute – if outcome is continuous
t's simplistic for a reader to appreciate the importance of

In case of multiple independent variables, we can go with forward selection, backward elimination and step wise approach for selection of most significant independent variables.

Please let me know where to get little details on these?
Compliments once again. All is well.
Asesh Datta

ROHIT **Reply**

July 3, 2016 at 1:37 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-113012)

Thanks Sunil ,
Useful article . Why is poisson regression not mentioned here.

FAZAL AZEEM **Reply**

November 20, 2018 at 4:16 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

d interesting.

Reply

November 20, 2018 at 4:16 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

narrow concept and gives food for thought. Thank you.

Reply

RAWIYAH ALRADDADI

January 29, 2018 at 4:00 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-151077)

Thank you very much!

RYAN ROSARIO **Reply**

March 22, 2018 at 1:49 pm (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-152080)

This is an awesome article. I just want to point out something important for people that may be new to modeling as many data science students I've mentored get overwhelmed and confused about what the different types of regression truly are, and what the true definition of linear regression is.

These aren't really different types of regression models per se. This is a mix of different techniques with different
near regression, logistic regression or any other kind of generalized

ase models covered.

e variables, but the model is still linear in the beta parameters. Thus
like y = exp(a + bx) is a generalized linear model if we use the log-
a + bx. This is a concept that bewilders a lot of people.

by adding and removing variables based on the F statistic.

Ridge, Lasso and ElasticNet aren't really regressions, they're penalties (regularization) on the loss function (OLS), or log-likelihood function (logistic and GLM). Hence, they are useful for other models that are distinct from regression with a regularization term/penalty."

would be plain linear, logistic, multinomial, poisson, gamma, Cox,

---

LORE Reply

tics vidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-

Wonderfull put down information. Thanks for sharing this in detail.

JAGADISH Reply

May 25, 2018 at 11:29 am (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/#comment-153540)

Hi Sunil ,

Really a good gist on regression techniques. Thanks for sharing the article.

JOIN THE NEXTGEN DATA SCIENCE ECOSYSTEM

Get access to free courses on Analytics Vidhya

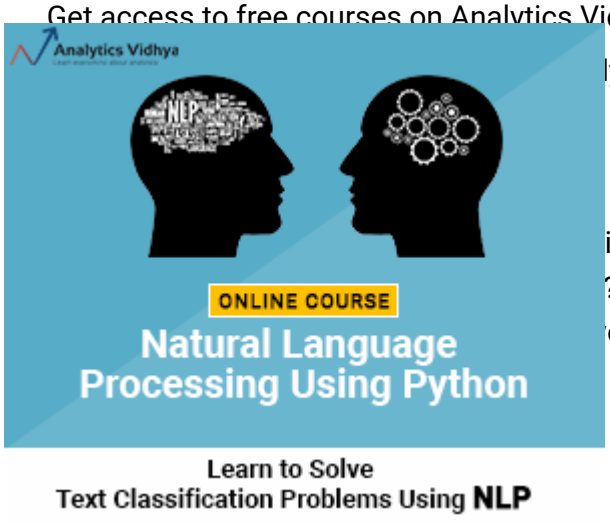
lytics Vidhya

in/?

?

eb)

Join Now



(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?

isplay&utm_campaign=NLPcourse)

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)

POPULAR POSTS

7 Regression Techniques you should know! (https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/)

7 Innovative Machine Learning GitHub Projects you Should Try Out in Python (https://www.analyticsvidhya.com/blog/2019/08/7-innovative-machine-learning-github-projects-in-python/)

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/)

11 Innovative Data Visualizations you Should Learn (in Python, R, Tableau and D3.js) (https://www.analyticsvidhya.com/blog/2019/08/11-data-visualizations-python-r-tableau-d3js/)

RECENT 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/)

SEPTEMBER 2, 2019

3 Beginner-Friendly Techniques to Extract Features from Image Data using Python (https://www.analyticsvidhya.com/blog/2019/08/3-techniques-extract-features-from-image-data-machine-learning-python/)

AUGUST 29, 2019

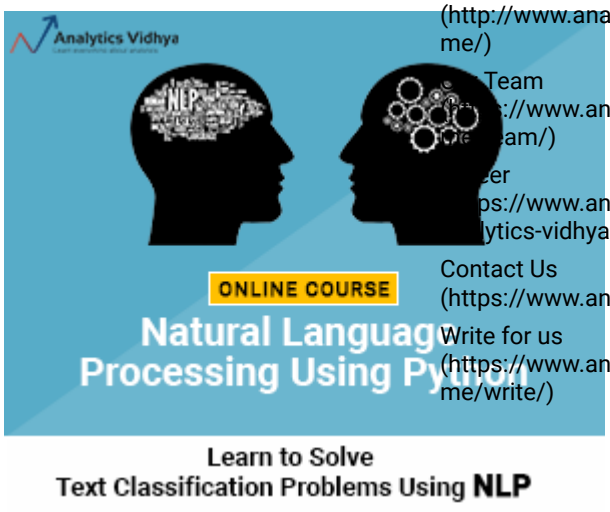
A Complete List of Important Natural Language Processing Frameworks you should Know (NLP Infographic) (https://www.analyticsvidhya.com/blog/2019/08/complete-list-important-frameworks-nlp/)

AUGUST 28, 2019

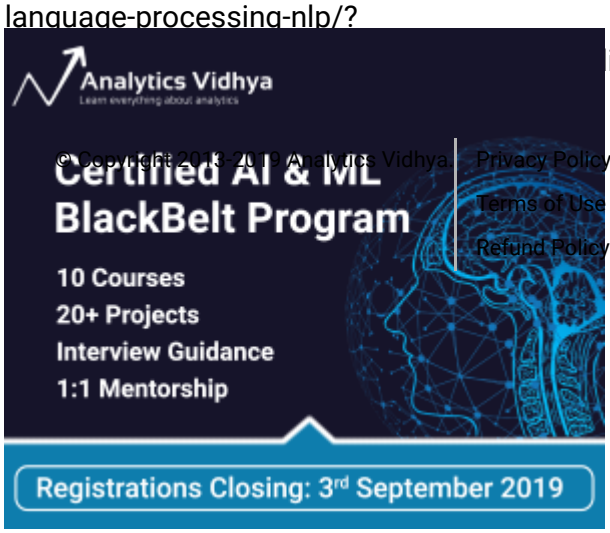
11 Innovative Data Visualizations you Should Learn (in Python, R, Tableau and D3.js) (https://www.analyticsvidhya.com/blog/2019/08/11-data-visualizations-python-r-tableau-d3js/)

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

(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)



(https://courses.analyticsvidhya.com/courses/natural-language-processing-nlp/?utm_source=blog&utm_medium=display&utm_campaign=NLPcourse)



(https://courses.analyticsvidhya.com/bundles/ai-blackbelt-beginner-to-master/?utm_source=blog&utm_medium=Stickybanner2)