**Roadmap for Machine Learning**

( [**Watch on Youtube for explanation**](https://www.google.com/url?q=https://www.youtube.com/watch?v%3D3hU5cxcOvVE&sa=D&source=editors&ust=1628566151585000&usg=AOvVaw1PrKiMABTLiR1U_TCxgiV1) )

Among many paths one can follow, below is one way describing minimum things that need to be done by a beginner to make a ML project.

**Must Know Theory**

Complete till week 5 of this course : [Machine Learning Coursera course](https://www.google.com/url?q=https://www.coursera.org/learn/machine-learning&sa=D&source=editors&ust=1628566151586000&usg=AOvVaw24728J2B3p7wylzeEhBDsC)

* You will cover the introduction of ML, Linear and Logistic Regression and the idea of neural networks.
* There are other algorithms like decision trees, ensemble algorithms like randomforest, catboost, xgboost etc. not covered above but you may use in your project directly from libraries.

*If you want to learn more :*

* *you can complete the course and related assignments.*
* *explore the maths behind ML :*[*Mathematics for Machine Learning*](https://www.google.com/url?q=https://mml-book.github.io/&sa=D&source=editors&ust=1628566151586000&usg=AOvVaw2EqYEBUXPMv3YFYUtgBDbg)

**Must Know Coding**

[**Python**](https://www.google.com/url?q=http://w3schools.com/python&sa=D&source=editors&ust=1628566151587000&usg=AOvVaw2bGfzbAhLwwQVar15Vyrj7) is the preferred programming language due to the following reasons:

* simple and easy to learn
* vast amount of available libraries and frameworks
* If you know any other language, you can learn the basics of python in a few hours.

If you are not familiar with python, you can learn through this free course:

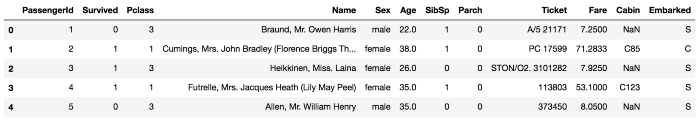
[Python Tutorial - Python for Beginners [Full Course]](https://www.google.com/url?q=https://www.youtube.com/watch?v%3D_uQrJ0TkZlc&sa=D&source=editors&ust=1628566151587000&usg=AOvVaw0KThMvs_72udImRa9IKRGI)

(at least up to Pypi and Pip)

1. Google Colaboratory is preferred for editing your code. Learn about it [here](https://www.google.com/url?q=https://medium.com/swlh/the-best-place-to-get-started-with-ai-google-colab-tutorial-for-beginners-715e64bb603b&sa=D&source=editors&ust=1628566151588000&usg=AOvVaw3D3p-fDdrrFErf6ujprJxl). It has many features. Both python codes and terminal commands can be run in the cells.
2. Must know Python Libraries - Numpy, Pandas and Matplotlib. Refer the following tutorials -

* [Python, Numpy, Matplotlib Tutorial](https://www.google.com/url?q=https://colab.research.google.com/github/cs231n/cs231n.github.io/blob/master/python-colab.ipynb&sa=D&source=editors&ust=1628566151589000&usg=AOvVaw2dDFzSdEsP249-f39FgcOK)
* [Another Numpy Tutorial](https://www.google.com/url?q=https://github.com/jamesdietle/fastaipart1v2/blob/master/Tutorials/NumpyTutorial.ipynb&sa=D&source=editors&ust=1628566151589000&usg=AOvVaw0v-Gyt4WFF82-3mc27xj1E)
* [Pandas Tutorial](https://www.google.com/url?q=https://mlcourse.ai/articles/topic1-exploratory-data-analysis-with-pandas/&sa=D&source=editors&ust=1628566151590000&usg=AOvVaw1zmlwwkwlBswNKBLoeZWLA)

**FOR PROJECTS BASED ON STRUCTURED / TABULAR DATA**



Here, many data points are given. A data point ( a row, also called a record ) is defined by its attributes ( columns or features).

One of the columns is the target variable.

Trivial parts involved in such projects are

1. **EDA (Exploratory Data Analysis)**

* analysis features and visualizing them.
* finding relations or trends between features.
* it’s not a must part but it's good to do it.

For visualizations refer [here](https://www.google.com/url?q=https://mlcourse.ai/articles/topic2-visual-data-analysis-in-python/&sa=D&source=editors&ust=1628566151591000&usg=AOvVaw1qoMM-cgzNmtlZxS_5m5Is).  No need to memorize the code. Just notice what can be done. Refer to the code whenever you need it.

1. **Feature Engineering and Data Cleaning**

* handling missing data
* making features useful for prediction, feature importance, adding new features, removing unnecessary features, etc

Refer [this.](https://www.google.com/url?q=https://medium.com/analytics-vidhya/data-cleaning-and-preprocessing-a4b751f4066f&sa=D&source=editors&ust=1628566151592000&usg=AOvVaw3E7axWvlpe_Otn0ha217v9)

1. **Training and Testing**

* Running different ML algorithms on data
* Testing model on test data
* Ensembling results, cross-validation, etc techniques can be used.

Libraries such as sklearn are usually used for the ML algorithms. Training and testing can be done in 2-3 lines with the libraries.

Refer [this](https://www.google.com/url?q=https://medium.com/fintechexplained/lets-talk-about-machine-learning-ensemble-learning-in-python-382747e5fba8&sa=D&source=editors&ust=1628566151592000&usg=AOvVaw1xEKnIMCdQq09xDu05cejN).

*Examples :*

* A Simple Example [Titanic Data Analysis - Beginner Friendly](https://www.google.com/url?q=https://www.kaggle.com/jicksy/titanic-data-analysis-beginner-friendly&sa=D&source=editors&ust=1628566151593000&usg=AOvVaw0wcJhUShWd294RXiKg36Pq)
* [EDA To Prediction(DieTanic)](https://www.google.com/url?q=https://www.kaggle.com/ash316/eda-to-prediction-dietanic&sa=D&source=editors&ust=1628566151593000&usg=AOvVaw2OX_cPlSOie0LloqAX9IcA). This is a very detailed example to give you an overall idea.

*If you want to learn more :*

* *complete*[*mlcourse.ai*](https://www.google.com/url?q=https://mlcourse.ai/&sa=D&source=editors&ust=1628566151594000&usg=AOvVaw1Xzj-U6M1tZdE06XfWYLgZ)*machine learning course.*

**FOR IMAGE-RELATED PROJECT / DEEP LEARNING**

* Complete the course : [Neural Networks and Deep Learning Coursera course](https://www.google.com/url?q=https://www.coursera.org/learn/neural-networks-deep-learning?specialization%3Ddeep-learning&sa=D&source=editors&ust=1628566151595000&usg=AOvVaw1x8bz8B9C8Z9odt2zl-I4Q)
* Complete till week 2 of the course : [Convolutional Neural Networks (coursera.org)](https://www.google.com/url?q=https://www.coursera.org/learn/convolutional-neural-networks?specialization%3Ddeep-learning&sa=D&source=editors&ust=1628566151595000&usg=AOvVaw29bb1IOavsksIKfQj1BN14)
* Now there is a need to learn **any one** of many deep learning frameworks

For **Keras**, refer these links:

[Introduction to Deep Learning with Keras | by Gilbert Tanner](https://www.google.com/url?q=https://towardsdatascience.com/introduction-to-deep-learning-with-keras-17c09e4f0eb2&sa=D&source=editors&ust=1628566151596000&usg=AOvVaw2iabIdfhJFqusH6_RJrIJB)

[Building A Deep Learning Model using Keras | by Eijaz Allibhai](https://www.google.com/url?q=https://towardsdatascience.com/building-a-deep-learning-model-using-keras-1548ca149d37&sa=D&source=editors&ust=1628566151596000&usg=AOvVaw1rHOUTDrs2tST_pO0J1Z7x)

[Building Neural Network using Keras for Classification | by Renu Khandelwal](https://www.google.com/url?q=https://medium.datadriveninvestor.com/building-neural-network-using-keras-for-classification-3a3656c726c1&sa=D&source=editors&ust=1628566151596000&usg=AOvVaw3G8R2WJAVvQ717NXF2-BCK)

keras pretrained model : [https://keras.io/api/applications/](https://www.google.com/url?q=https://keras.io/api/applications/&sa=D&source=editors&ust=1628566151597000&usg=AOvVaw3T0WS2R21wAl0HH9W5Bd02)

For **Pytorch**, refer these links:

[jcjohnson/pytorch-examples: Simple examples to introduce PyTorch](https://www.google.com/url?q=https://github.com/jcjohnson/pytorch-examples&sa=D&source=editors&ust=1628566151597000&usg=AOvVaw0L0Ngq1U6dXPyafDRKXQsT)

[A Gentle Introduction to PyTorch 1.2 | by elvis | dair.ai](https://www.google.com/url?q=https://medium.com/dair-ai/pytorch-1-2-introduction-guide-f6fa9bb7597c&sa=D&source=editors&ust=1628566151597000&usg=AOvVaw3GTmi4d2tTWBIV5TkqfSRD)

[PyTorch For Deep Learning — Binary Classification](https://www.google.com/url?q=https://medium.com/analytics-vidhya/pytorch-for-deep-learning-binary-classification-logistic-regression-382abd97fb43&sa=D&source=editors&ust=1628566151598000&usg=AOvVaw2M9cbuvf7Qb02lvhuwxr9M)

pytorch pretrained models : [https://pytorch.org/vision/stable/models.html](https://www.google.com/url?q=https://pytorch.org/vision/stable/models.html&sa=D&source=editors&ust=1628566151598000&usg=AOvVaw2vuiFV-B9MD6B2gjkm3ayF)

For **Fastai**, refer these links:

[Deep Learning Image Classification with Fastai | by Blake Samaha](https://www.google.com/url?q=https://towardsdatascience.com/deep-learning-image-classification-with-fast-ai-fc4dc9052106&sa=D&source=editors&ust=1628566151599000&usg=AOvVaw3lC5lEnjI48xCSKgxfZ_Ya)

[Building an Object Classification Model with Fast.AI | by Ronak Bhatia |](https://www.google.com/url?q=https://ronak-k-bhatia.medium.com/building-an-object-detection-model-with-fast-ai-c1694c7a521e&sa=D&source=editors&ust=1628566151599000&usg=AOvVaw0Qmu6nxAnFuS203zrwSR9K)

[Image Classification using FastAI and Transfer Learning | by Parasrawatjr](https://www.google.com/url?q=https://medium.com/data-science-community-srm/image-classification-using-fastai-and-transfer-learning-463b47285a4f&sa=D&source=editors&ust=1628566151599000&usg=AOvVaw2LB4TlHwCwTr1mF8X0o3WH)

*If you want to learn more :*

* *you can complete the course and related assignments.*
* *complete*[*CS231n*](https://www.google.com/url?q=http://cs231n.stanford.edu/2019/&sa=D&source=editors&ust=1628566151600000&usg=AOvVaw0ddCoRK3D7KiG8408VkRhh), its assignment and lecture youtube videos.

**DEPLOY YOUR MODEL BY SETTING UP UI :**

1. If you or your friend knows *basic web development* (HTML, CSS), you can integrate ML models in the backend using Flask or FastAPI. These would be helpful for that:

[Learn Flask for Python - Full Tutorial](https://www.google.com/url?q=https://www.youtube.com/watch?v%3DZ1RJmh_OqeA&sa=D&source=editors&ust=1628566151601000&usg=AOvVaw1rWpWGMhmoPywiQIxJaQQc)

[Machine Learning Web App with Python](https://www.google.com/url?q=https://towardsdatascience.com/build-a-web-application-for-predicting-apple-leaf-diseases-using-pytorch-and-flask-413f9fa9276a&sa=D&source=editors&ust=1628566151601000&usg=AOvVaw1eD76svp8Prm4tPRSU1xbc)

1. Or you may use a Python library like Streamlit to build WebApp in just a few lines of python code:

[Model Deployment Using Streamlit | Deploy ML Models using Streamlit](https://www.google.com/url?q=https://www.analyticsvidhya.com/blog/2020/12/deploying-machine-learning-models-using-streamlit-an-introductory-guide-to-model-deployment/&sa=D&source=editors&ust=1628566151601000&usg=AOvVaw3VYZaQKXBzjdoRnCIGVc0p)

1. You can also make WebApp directly from Colab Notebook:

[Turning a Google Colab notebook into a web app](https://www.google.com/url?q=https://anvil.works/learn/tutorials/google-colab-to-web-app&sa=D&source=editors&ust=1628566151602000&usg=AOvVaw1Sjh_fcwJGIxPKWmQsmFie)

**UPLOAD PROJECT ON GITHUB :**

You can organize your code in classes and upload on github with proper Readme.

Refer [this](https://www.google.com/url?q=https://medium.com/analytics-vidhya/how-to-add-a-machine-learning-project-to-github-4d8857c853a0&sa=D&source=editors&ust=1628566151603000&usg=AOvVaw1JwvrgbYL-Bu0n5-2_GbpG) for adding your project on Github.

*Example :*[*identify-the-animal*](https://www.google.com/url?q=https://github.com/naineshhulke/identify-the-animal&sa=D&source=editors&ust=1628566151603000&usg=AOvVaw0bEeUNE7A52dx9jecwzRBI)*( Keras, Streamlit )*

**Note :**

1. For the **coursera courses** which are paid for, you can apply for **financial aid** to get the course for **free**. You can find this option to the right of the ‘Enroll Now’ button. Don’t forget to mention you are a student from India. Till the time the financial aid is accepted, you can audit the course or continue using the course without availing certificates.
2. **Google and Stackoverflow** are your best friends. At any point for any error, just google it. Use proper search terms. If you want to do a task like ‘use kaggle dataset in google colab’, search for it.
3. You can look up how to use a particular method of **library** from its **documentation**. Google for it.
4. There’s no need to learn all the different methods and functions that a library has. If you want to do anything, just search on google if a library has such a function. Also, **before writing a large code for some task, look up on google** if any library has a function that will accomplish the task in a single line. Many times there is.
5. Explore machine learning competition platforms like [Kaggle](https://www.google.com/url?q=https://kaggle.com/&sa=D&source=editors&ust=1628566151604000&usg=AOvVaw3HmywKDt39enN-DEg43qYN). You will find lots of dataset and related notebooks. Also, you get to learn a lot from Kaggle competitions.