



# Introduction to Machine Learning in Python (ML)

The beginning of it all

Jake Parbs [japarbs@42.student.org](mailto:japarbs@42.student.org)

*Summary: Intro to Linear and Non-Linear models, methods, environment setup and getting started.*



*This work is licensed under a [Creative Commons Attribution-Noncommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).*

# Contents

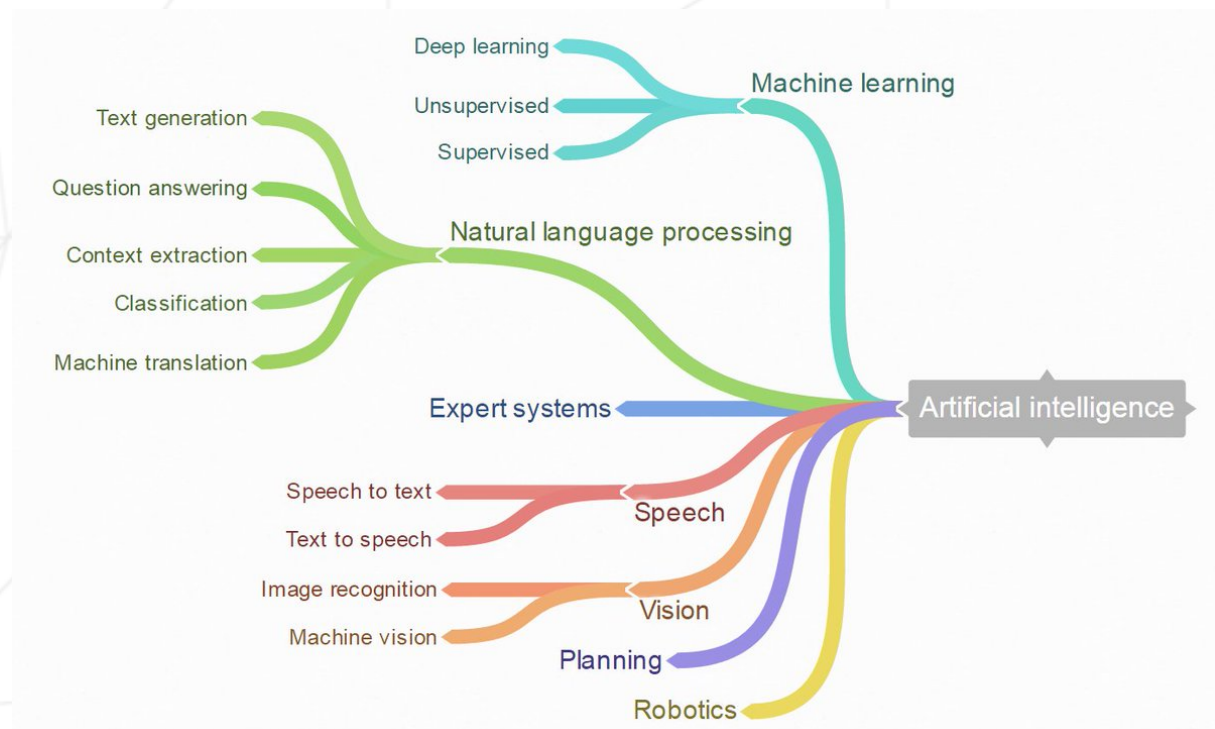
<b>I</b>	<b>Foreword</b>	<b>2</b>
<b>II</b>	<b>About the course</b>	<b>3</b>
<b>III</b>	<b>Coursera setup</b>	<b>4</b>
<b>IV</b>	<b>Ex00: Introduction to Machine Learning</b>	<b>5</b>
<b>V</b>	<b>Linear and Non-Linear regression</b>	<b>6</b>
V.1	Ex01: Linear Regression . . . . .	6
V.2	Ex02: Multiple Linear Regression . . . . .	6
V.3	Ex03: Non-Linear Regression . . . . .	6
<b>VI</b>	<b>Showcase Project Setup</b>	<b>7</b>
VI.1	Anaconda . . . . .	7
VI.2	PyCharm . . . . .	7
VI.3	Packages . . . . .	8
<b>VII</b>	<b>Ex04: Showcase Project</b>	<b>9</b>
VII.1	Demonstration . . . . .	10
<b>VIII</b>	<b>Ex05: Please Readme!</b>	<b>11</b>
VIII.1	Markdown language . . . . .	11
VIII.2	Demonstration . . . . .	12
<b>IX</b>	<b>Turn-in and peer-evaluation</b>	<b>13</b>

# Chapter I

## Foreword

Machine Learning is such a massive field. It's everywhere and you use it everyday. From spam filters, price matching, and path finding on a GPS to robots, theoretical research and more.

The simple Machine Learning models that you're going to learn in this course are the building blocks to the cutting edge technology we have today, and this is only scratching the surface of the power of ML technology. Without ML so many technologies wouldn't be possible due to their scale and the flexibility some systems need. You might not think much of it at first but every single social media site: Facebook, Twitter, Youtube, Instagram are all powered by Machine Learning in some kind of way. They just wouldn't function the same way today due to their scale and would require much more brain power to program and maintain.

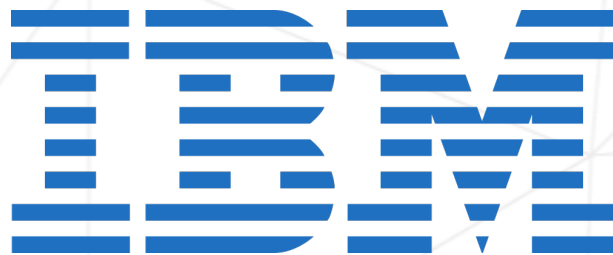


# Chapter II

## About the course

You will be using "Machine Learning with Python" by IBM for the learning material via Coursera. Practice and learn code in the environments that Coursera provides via Jupyter Notebooks that contain working practice code.

Next, you will then setup a local environment that you will use to build your final Showcase Project.



# Chapter III

## Coursera setup




You will need to create a account in order to use Coursera. It's the same as any other web sign up and is free. Once you've done that, enroll into the course. Once in the course you should be brought to the overview page. Now we can get into the swing of things!

We will not be using the Quizzes that are built into the curriculum on Coursera. However do feel free to use the in-video questions to make sure you understand the topic.

# Chapter IV

## Ex00: Introduction to Machine Learning

	Introduction to Machine Learning
Topics to study : <a href="#">Introduction to Machine Learning on Coursera</a>	
Files to turn in : None	
Notes : This will tell you what Machine Learning is, and some use cases in machine learning. Watch all of week 1 of Coursera.	




There's a lot of information here. Take notes! You'll thank yourself later.


# Chapter V

## Linear and Non-Linear regression


### V.1 Ex01: Linear Regression

	Linear Regression
Topics to study : <a href="#">Linear Regression on Coursera</a>	
Files to turn in : None	
Notes : Within week 2 learn about Linear Regression	

### V.2 Ex02: Multiple Linear Regression

	Multiple Linear Regression
Topics to study : <a href="#">Multiple Linear Regression on Coursera</a>	
Files to turn in : None	
Notes : Within week 2 learn about Multiple Linear Regression	

### V.3 Ex03: Non-Linear Regression

	Non-Linear Regression
Topics to study : <a href="#">Non-Linear Regression on Coursera</a>	
Files to turn in : None	
Notes : Within week 2 learn about Non-Linear Regression	

# Chapter VI

## Showcase Project Setup

This Showcase Project is for you to practice what you just learned about Linear Regression and learn how to setup your own environment. You've been using the Jupyter notebooks for practicing code. Now it is time for you to create your own environment to do this final project!

### VI.1 Anaconda



Anaconda is a platform that has tons and tons of tools for Machine Learning and Python. It's great for Python versioning and package control. For now however we just want to use it to have our own development environment.

[Download Anaconda for Python 3.7 here.](#)

Anaconda will take a while to download and install. Lets install PyCharm as we wait.

### VI.2 PyCharm



You'll also need an IDE to develop in. I recommend PyCharm because it has built-in support for Anaconda environments. Optionally, you can use your preferred IDE. [You can download PyCharm Community here.](#)



## VI.3 Packages

With Anaconda installed we can now create our environment and install all needed packages in a Unix terminal. The following command will install an Anaconda environment named "mlenv" in Python 3.7

To use Anaconda within Unix we call its command, Conda.

```
conda create -n mlenv python=3.7
```

This command will then activate the Conda environment and allow us to configure it.

```
conda activate mlenv
```

Now that the environment is installed and selected, run the following commands to install our packages. This is handled just like a normal Python environment since our environment is active.

```
pip install numpy  
pip install pandas  
pip install sklearn  
pip install matplotlib
```

Now that we have all packages installed and everything setup. We're ready to start developing. Open PyCharm and create a new project and select the Conda environment you've just created.

# Chapter VII


## Ex04: Showcase Project

Run the following command within your working directory to download the dataset.

```
wget -O FuelConsumption.csv https://s3-api.us-gso.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/ML0101ENv3/labs/FuelConsumptionCo2.csv
```

If you don't have wget then you can install it via Homebrew (This will take a bit).

```
brew install wget
```

	Showcase Project
Topics to study : Create your Showcase Project	
Files to turn in : showcase.py	
Notes : This will be what you create for your main grade.	

Create a Showcase Project with this dataset showing the linear relation of Engine Size and Fuel Consumption(Highway) with matplotlib with a line of best fit. And a Linear Regression model to predict outside of dataset cases.

There are more details on what exactly to be done below. Everything here has already been covered in the course.



Use the Jupyter Notebooks if you're lost!

## VII.1 Demonstration

The Showcase Project should have the following.

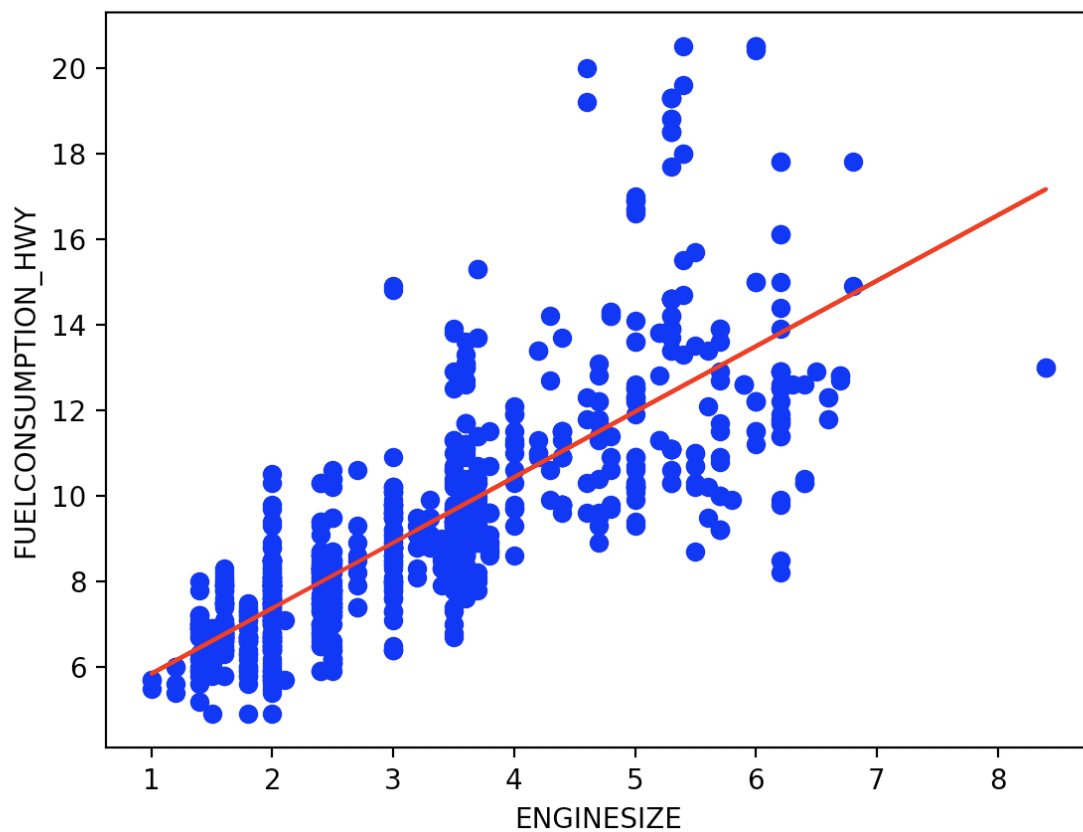
### VII.1.1 A read of the filtered data frame in command line.

```
#      ENGINESIZE  FUELCONSUMPTION_HWY
0         2.0         6.7
1         2.4         7.7
2         1.5         5.8
3         3.5         9.1
4         3.5         8.7
...      ...      ...
1062      3.0         9.8
1063      3.2         9.5
1064      3.0         9.8
1065      3.2         9.3
1066      3.2        10.2
```

### VII.1.2 R2 Score


```
R2-score: 0.28
```

### VII.1.3 Matplotlib graph with line of best fit from the model



# Chapter VIII

## Ex05: Please Readme!

	Ex05: Readme!
Topics to study : Create a Readme file that describes the nature of the project.	
Files to turn in : README.md	
Notes : There are more details on what exactly to be done below.	

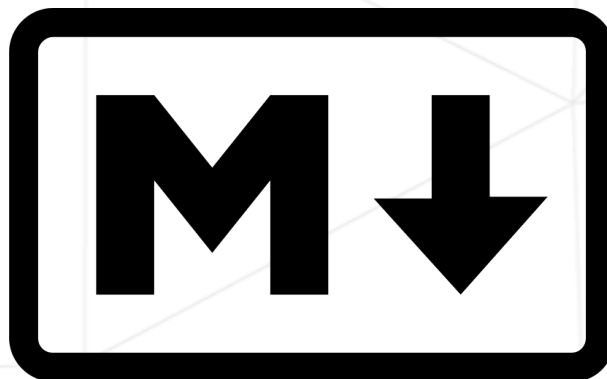
### VIII.1 Markdown language

A Readme is an essential document on any project. It can be used for things such as describing functions in a custom library, recording known bugs, a detailed description about what the project is and what it's purpose is, ect.

Readme files on Github are made in Markdown. A simple formatting language that's designed to be quick and powerful to write.

[Here is a Github page explaining the ins and outs of Markdown.](#)

[dillinger.io](#) is a great online Markdown editor if you don't feel like using Github's editor.



## VIII.2 Demonstration

The Readme should have the following things. It is encouraged to add more and polish it as much as you feel necessary.

### VIII.2.1 Summary of the project

What it does, what does it predict, why you made it, ect.

### VIII.2.2 What is Machine Learning?

In your own words what is machine learning and what can it do?

### VIII.2.3 Challenges

What were things that challenged you? How did you overcome them?

### VIII.2.4 What you've learned

Give an explanation of what you've learned from doing this project. Has it made you a better developer?

Here's an example.

## I'm a Header! ML Something Example

### What is the Thing?

things are here on what did do. It predicted the thing.

### What is the MLs?

It brain but not. Right?

### What did learn?

I learned this example is silly.

# Chapter IX

## Turn-in and peer-evaluation

Turn your work in using your `Git` repository, as usual. Only work present on your repository will be graded in defense.