Contents

* [Python](#_fkkru5l18mb5)
* [Pandas](#_frqv5epq3vvf)
* Matplotlib
* [Module 3](#_5w323tbs2zte)
  + [18. Real world problem: Predict rating given product reviews on Amazon](#_8q4qk83lojeq)
* [Document Notes](#_iaggr9bk5tpb)

# Python

* Single-line if with lambda operation:- lambda x: True if x % 2 == 0 else False
* Nested if-else: lambda x: x\*10 if x<2 else (x\*\*2 if x<4 else x+10)
* Random.choices
* argmax

# Pandas

# Linear Algebra

* Normal to the plane
* Distance between points
* Distance from origin

# Statistics and probability

* Stratified sampling (test\_train\_split)
* Regularization / normalisation / standardisation

# Module 3

## 18. Real world problem: Predict rating given product reviews on Amazon

Topics - Word vectors, bag of words, word2vec, tf-idf

* ~~Sparse matrix~~
* ~~Word vectors~~
* ~~Bag of words~~
* Word2vec
* TF-IDF
* Norm and how is it represented
* Explain text being geometrically closer
* Problem with a bag of words and how to fix it
* Binary BoW / Boolean BoW
* BoW v/s Binary BoW
* Featurization in NLP
* Stop word removal, tokenization, lemmatization
* Tokenization v/s lemmatization

# Logistic Regression

## Document Notes:

* Use Heading1 for topics and Heading 2 for subtopics
* Heading1 and Heading2 should be arranged based on the notes.
* Add the video no., if you include a topic of interest from video (including time if necessary)