| **WEEK 1** | **WHAT'S IN THERE** | **TASK 1** | **TASK 2** | **TASK 3** |
| --- | --- | --- | --- | --- |
| **Day 0** | Getting Started by setting up Anaconda environment. | [Installing anaconda for Windows](https://docs.anaconda.com/free/anaconda/install/windows/) | [Installing anaconda for Mac](https://docs.anaconda.com/free/anaconda/install/mac-os/) | [Installing anaconda for Linux](https://docs.anaconda.com/free/anaconda/install/linux/) |
| **Day 1** | Hey there, excited to start learning? First,  we will begin with learning the basics of Python. | [Beginner Tutorial for Python Programming.(videos 1-3)](https://www.youtube.com/playlist?list=PL-osiE80TeTskrapNbzXhwoFUiLCjGgY7) | [Moving Ahead (videos 4 - 6)](https://www.youtube.com/playlist?list=PL-osiE80TeTskrapNbzXhwoFUiLCjGgY7) | [Wrapping up on basics (videos 7 - 10)](https://www.youtube.com/playlist?list=PL-osiE80TeTskrapNbzXhwoFUiLCjGgY7) |
| **Day 2** | Gaining an overview of NumPy and Pandas. You will be using them extensively in your Data Science journey. | [Basics of Numpy (till 40 mins is sufficient)](https://www.youtube.com/watch?v=QUT1VHiLmmI)   [Numpy Notebook](https://www.kaggle.com/code/orhansertkaya/numpy-tutorial-for-beginners#Computation-on-NumPy-Arrays:-Universal-Functions) | [Intro to Pandas (videos 1 to 10)](https://www.youtube.com/playlist?list=PL5-da3qGB5ICCsgW1MxlZ0Hq8LL5U3u9y) | [Data Analysis with Pandas (videos 11 to 20)](https://www.youtube.com/playlist?list=PL5-da3qGB5ICCsgW1MxlZ0Hq8LL5U3u9y) |
| **Day 3** | Continuing with Pandas, let's use it for data cleaning and transformations | [Diving into Pandas (videos 25 to 33)](https://www.youtube.com/playlist?list=PL5-da3qGB5ICCsgW1MxlZ0Hq8LL5U3u9y) | [Pandas Notebook](https://github.com/justmarkham/pandas-videos/blob/master/pandas.ipynb) | [Kaggle Micro-course on Pandas (only exercise)](https://www.kaggle.com/learn/pandas) |
| **Day 4** | Data Visualization helps us in gaining insights from the data through visuals like graphs and maps. We would look into some common libraries which are Matplotlib, Seaborn, and Plotly. | [Intro to Matplotlib](https://www.youtube.com/watch?v=MbKrSmoMads)  [GFG article on Matplotlib (optional but helpful)](https://www.geeksforgeeks.org/matplotlib-tutorial/) | [Intro to Seaborn](https://www.youtube.com/playlist?list=PLBfyvFO_aKGRaJmdo501Hu_wXwgmjbR50)  [GFG Tutorial on Seaborn](https://www.geeksforgeeks.org/python-seaborn-tutorial/) | [Intro to Plotly (required till 1 hr)](https://www.youtube.com/watch?v=GGL6U0k8WYA) |  |
| **Day 5** | Dealing with large tables can at times become overwhelming. Thus we may want to summarize the content of tables using Descriptive Statistics. | [Data Types in Statistics.](https://towardsdatascience.com/data-types-in-statistics-347e152e8bee) | [Measurement of Central Tendency. (videos 3 - 7)](https://www.youtube.com/playlist?list=PL0KQuRyPJoe6KjlUM6iNYgt8d0DwI-IGR) | [Normal Distribution (videos 18 - 20)](https://www.youtube.com/playlist?list=PL0KQuRyPJoe6KjlUM6iNYgt8d0DwI-IGR)  [Quantile Plots](https://towardsdatascience.com/q-q-plots-explained-5aa8495426c0) |  |

| **WEEK 2** | **WHAT'S IN THERE** | **TASK 1** | **TASK 2** | **TASK 3** |
| --- | --- | --- | --- | --- |
| **Day 1** | Hey, excited for Week 2? Often the data we deal with can have various issues like **missing values**, **categorical values** and **outliers**. Today we will learn about basic techniques to deal with such issues! | [Outlier Analysis](https://www.youtube.com/watch?v=rzR_cKnkD18) | [Handling Missing Values](https://www.youtube.com/watch?v=P_iMSYQnqac)  [Practical Handling Missing Values](https://www.kaggle.com/code/debarshichanda/handling-missing-values/notebook) | [Exploratory Data Analysis](https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15)   [Handling Categorical Variables](https://medium.com/analytics-vidhya/how-to-handle-categorical-features-ab65c3cf498e#:~:text=1)%20Using%20the%20categorical%20variable,category%20with%20a%20probability%20ratio.) |
| **Day 2** | Today we aren't going to be too technical, let us just motivate ourselves about Machine Learning, get to know its application, and have a rudimentary understanding of what **Machine Learning** is. | [What is ML , common problems solved by ML](https://www.youtube.com/watch?v=_1QtMPuYIVw&t=582s) | [Supervised and unsupervised learning](https://www.youtube.com/watch?v=xtOg44r6dsE) | [(optional)[Linear Algebra, Refresher required for those who don't have mathematical base]](https://cs229.stanford.edu/summer2019/cs229-linalg.pdf) |
| **Day 3** | Starting at the grassroots level, we study in depth the simplest ML model **Linear Regression**, alongwith **Cost function** and **Gradient Descent**. Don't worry if it sounds too hard, trust us it isn't. | [Linear Regression Blog](https://towardsdatascience.com/everything-you-need-to-know-about-linear-regression-b791e8f4bd7a) | [Linear Regression with One Variable  (Videos 9 - 14)](https://www.youtube.com/watch?v=dLc-lfEEYss&list=PLkDaE6sCZn6FNC6YRfRQc_FbeQrF8BwGI&index=9) | [Linear Regression with One Variable (Videos 15 - 20)](https://www.youtube.com/watch?v=WtlvKq_zxPI&list=PLkDaE6sCZn6FNC6YRfRQc_FbeQrF8BwGI&index=15) |
| **Day 4** | Let us spice things up a bit, we study Linear regression again but this time with **Multiple features**. | [Linear Regression with Multiple Variables (Videos 21 - 24)](https://www.youtube.com/watch?v=WtlvKq_zxPI&list=PLkDaE6sCZn6FNC6YRfRQc_FbeQrF8BwGI&index=21) | [Linear Regression without Scikit-learn](https://inria.github.io/scikit-learn-mooc/python_scripts/linear_regression_without_sklearn.html) | [Linear Regression with Scikit-learn](https://inria.github.io/scikit-learn-mooc/python_scripts/linear_regression_in_sklearn.html)   [Linear Regression with 1 variable from scratch](https://www.kaggle.com/code/navinpatwari09/linear-regression-with-1-feature?scriptVersionId=131306785) |
| **Day 5** | Today we will be introduced to our first ever classification model, **Logistic Regression**. Let's get to it. | [Logistic Regression Videos 31 - 36](https://www.youtube.com/watch?v=p-ltr1C7u2o&list=PLkDaE6sCZn6FNC6YRfRQc_FbeQrF8BwGI&index=31) | [Logistic Regression Blog](https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148) | [Logistic Regression with Scikit-learn](https://www.youtube.com/watch?v=71iXeuKFcQM&t=5s)   [Logistic Regression from scratch](https://www.kaggle.com/code/arush207/basic-implementation-of-logistic-regression) |

| **WEEK 3** | **WHAT'S IN THERE** | **TASK 1** | **TASK 2** | **TASK 3** |
| --- | --- | --- | --- | --- |
| **Day 1** | Since we have covered 2 basic ML models, let us take a break and learn about **Overfitting, Underfitting and the Bias-Variance Tradeoff**. These can help in telling you the complexity of your model - how well your model has used your data. This will be followed by Regularization. | [Bias-Variance Video](https://www.youtube.com/watch?v=EuBBz3bI-aA)  [Blog](https://medium.com/@itbodhi/bias-and-variance-trade-off-542b57ac7ff4) | [Overfitting and Regularisation (37 - 41)](https://www.youtube.com/watch?v=8upNQi-40Q8&list=PLkDaE6sCZn6FNC6YRfRQc_FbeQrF8BwGI&index=37) | [L1 L2 Regularization](https://towardsdatascience.com/intuitions-on-l1-and-l2-regularisation-235f2db4c261)  [Lasso and Ridge Regression](https://www.youtube.com/watch?v=VqKq78PVO9g) |
| **Day 2** | Today we will give you an introduction to Feature Transforations and how these are used for different types of data | [All Feature Transformation](https://www.youtube.com/watch?v=3gfhbXt9TcQ) | [Scaling methods](https://www.analyticsvidhya.com/blog/2020/07/types-of-feature-transformation-and-scaling/) | [Categorical Encoding](https://pbpython.com/categorical-encoding.html)  [Bag of Words](https://www.youtube.com/watch?v=IKgBLTeQQL8) |
| **Day 3** | Today, we'll take a closer look at what the AUC-ROC Score is and various other Evaluation Metrics to evaluate our machine leaning algorithms. | [AUC - ROC curve](https://www.youtube.com/watch?v=4jRBRDbJemM)  [AUC - ROC curve Blog](https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/?fbclid=IwAR3NiyvLoVEQxRCerb5A3YVU8Qtuf9fpnG5ERWGLBQsfKbpvfuccI-7DI7U) | [Confusion Matrix](https://www.youtube.com/watch?v=Kdsp6soqA7o)  [Confusion Matrix Blog](https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/) | [Evaluation Metrics](https://medium.com/ml-cheat-sheet/machine-learning-evaluation-metrics-b89b8832e275)  [Evaluation Metrics (Optional Topics included)](https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html) |
| **Day 4** | Today, we look into Naive and Gaussian Naive Baye's Algorithms. Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. | [Multinomial Naive Bias Classifier](https://www.youtube.com/watch?v=jS1CKhALUBQ) | [Gaussian Naive Bias](https://www.youtube.com/watch?v=H3EjCKtlVog) | [Naive Bias implementation using Scikit-learn](https://www.youtube.com/watch?v=nHIUYwN-5rM) |
| **Day 5** | Let's have a look at Support Vector Machine (SVM) | [SVM 1](https://www.youtube.com/watch?v=efR1C6CvhmE) | [SVM 2](https://www.youtube.com/watch?v=Toet3EiSFcM) | [SVM 3](https://www.youtube.com/watch?v=Qc5IyLW_hns)  [SVM Blog with code implementation](https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/) |

| **WEEK 4** | **WHAT'S IN THERE** | **TASK 1** | **TASK 2** | **TASK 3** |
| --- | --- | --- | --- | --- |
| **Day 1** | A model's performance can be greatly increased by tuning its hyperparameters and at the same time it is also important to look for how accurate our model is. For this, we can use Grid search methods and Cross-Validation. | [Cross Validation](https://www.youtube.com/watch?v=fSytzGwwBVw)  [Code Implementation](https://www.youtube.com/watch?v=gJo0uNL-5Qw) | [What is hyperparameter tuning ?](https://www.youtube.com/watch?v=DTcfH5W6o08) | [Implementing Random Search method.](https://www.youtube.com/watch?v=aijB8qbEOQ4) |
| **Day 2** | Today, we shall learn about Decision Trees and Random Forest which will create the foundation for many advanced Machine Learning Algorithms. | [DecisionTrees (videos 46 -49)](https://www.youtube.com/watch?v=_L39rN6gz7Y&list=PLblh5JKOoLUICTaGLRoHQDuF_7q2GfuJF&index=46) | [Random Forest (videos 52 - 53)](https://www.youtube.com/watch?v=J4Wdy0Wc_xQ&list=PLblh5JKOoLUICTaGLRoHQDuF_7q2GfuJF&index=52) | [Decision Trees Notebook](https://www.kaggle.com/code/faressayah/decision-trees-random-forest-for-beginners)  [Random Forest implementation](https://www.youtube.com/watch?v=ok2s1vV9XW0) |
| **Day 3** | Let's explore what boosting is and some of its variations | [Gradient Boosting (videos 59 - 61)](https://www.youtube.com/watch?v=3CC4N4z3GJc&list=PLblh5JKOoLUICTaGLRoHQDuF_7q2GfuJF&index=59)  [XGBoost](https://www.youtube.com/watch?v=OtD8wVaFm6E) | [XGBoost continued](https://www.youtube.com/watch?v=8b1JEDvenQU) | [Catboost 1](https://www.youtube.com/watch?v=KXOTSkPL2X4)  [Catboost 2](https://www.youtube.com/watch?v=3Bg2XRFOTzg) |
| **Day 4** | Let's look at some more variations of boosting algorithms and how they can be used for specialized tasks. | [Adaboost](https://www.youtube.com/watch?v=LsK-xG1cLYA) | [Kaggle Intermediate microcourse](https://www.kaggle.com/learn/intermediate-machine-learning) | [LightGBM](https://www.kaggle.com/code/prashant111/lightgbm-classifier-in-python#2.-LightGBM-intuition-)  [AdaBoost implementation](https://www.kaggle.com/code/prashant111/adaboost-classifier-tutorial/notebook) |
| **Day 5** | Today we learn about KNN. K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique. | [KNN video](https://www.youtube.com/watch?v=HVXime0nQeI) | [KNN Blog](https://towardsdatascience.com/k-nearest-neighbors-knn-explained-cbc31849a7e3) | [KNN Implementation](https://www.analyticsvidhya.com/blog/2021/01/a-quick-introduction-to-k-nearest-neighbor-knn-classification-using-python/) |

| **WEEK 5** | **WHAT'S IN THERE** | **TASK 1** | **TASK 2** | **TASK 3** |
| --- | --- | --- | --- | --- |
| **Day 1** | Welcome to the final week of Learning, as next week is Capstone Project This week we will discuss **Neural Networks** and **Unsupervised Learning**. Today Let us go into the foundations of **Neural Networks** | [Neural Networks **(1 - 23)**](https://www.youtube.com/watch?v=CS4cs9xVecg&list=PLkDaE6sCZn6Ec-XTbcX1uRg2_u4xOEky0&index=1) |  |  |
| **Day 2** | Today we will first get an intuitive understanding of how **Neural Networks** work and then implement a small Neural Network from scratch using Python | [Understanding Neural Networks](https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi) | [Neural Networks with Python](https://youtu.be/ubgwaNhU5v8) |  |
| **Day 3** | Today we will learn how to use the **Keras** library to implement Neural Networks. **Keras** is a popular Deep learning library which makes using Neural Networks very simple for us | [Regression with Keras](https://www.pluralsight.com/guides/regression-keras/) | [Classification with Keras](https://www.pluralsight.com/guides/classification-keras) |  |
| **Day 4** | Let us have a look at **unsupervised learning**, its uses and types. We will also look at one particular algorithm the **K-means method** | [Unsupervised Learning](https://towardsdatascience.com/unsupervised-learning-and-data-clustering-eeecb78b422a) | [K-means Clustering](https://youtu.be/4b5d3muPQmA) |  |
| **Day 5** | Today we will discuss **PCA** and its application through **scikit-learn** | [Tools and techniques for Deep Learning.  **24 - 43**](https://www.youtube.com/watch?v=k_S5fnKjO-4&list=PLkDaE6sCZn6Ec-XTbcX1uRg2_u4xOEky0&index=24) | [PCA using sklearn](https://youtu.be/QdBy02ExhGI) |  |

| **LINK** | **DEADLINE** | **INSTRUCTIONS** |
| --- | --- | --- |
| [**Final hackathon**](https://aiplanet.com/challenges/347/healthy-life-style-hackathon-by-iit-guwahati-d7ac6059) | **July 16, 2023** | NA |
| [**Quiz 5**](https://forms.gle/9k9M5WJkLFaG6ayX9) | **July 05, 2023** |  |
| Assignment 5   [**GDrive Link**](https://drive.google.com/drive/folders/1GHLdL9g3RBRcM7VmJMWX_uxIC4TXYKc2)  [**Datalore Link**](https://datalore.jetbrains.com/notebook/0xka1Qy8HKyNSjsVl2tMFy/gTVYVPxBztEoFrisdVZekP) | **July 05, 2023** | [You can share the link to you notebook or upload it directly here.](https://forms.gle/ojnTbGVCr8ecLVKp6)  We would recommend cloning datalore notebook as it won't require you to download dataset on your PC also there are many professional tools to easy your task. |