|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr. No.** | **Date** | **Module No.** | Module Name | **Topic** | **Link** |
| 1 | 18/05/2020 |  | Data Preprocessing in Python | Python libraries - Numpy  Pandas Dataframe Reading data  Matplotlib  Seaborn  sklearn  Categorical data  Feature Scaling | [**https://data-flair.training/blogs/python-ml-data-preprocessing/**](https://data-flair.training/blogs/python-ml-data-preprocessing/)  [**https://towardsdatascience.com/data-preprocessing-3cd01eefd438**](https://towardsdatascience.com/data-preprocessing-3cd01eefd438)  [**https://www.kdnuggets.com/2019/11/data-cleaning-preprocessing-beginners.html**](https://www.kdnuggets.com/2019/11/data-cleaning-preprocessing-beginners.html)  [**https://www.geeksforgeeks.org/numpy-in-python-set-1-introduction**](https://www.geeksforgeeks.org/numpy-in-python-set-1-introduction)  [**https://www.tutorialspoint.com/python\_pandas/index.htm**](https://www.tutorialspoint.com/python_pandas/index.htm)  .drop(), .iloc(), isnan(), .values, .describe(), .shape(), reshape(), .head(10), .to\_csv(), .read\_csv(), .ravel()  [**https://www.tutorialspoint.com/matplotlib/index.htm**](https://www.tutorialspoint.com/matplotlib/index.htm) (scatter plot, histogram, box plot)  vs  [**https://seaborn.pydata.org/generated/seaborn.heatmap.html**](https://seaborn.pydata.org/generated/seaborn.heatmap.html) (heatmap)  [**https://scikit-learn.org/stable/tutorial/basic/tutorial.html**](https://scikit-learn.org/stable/tutorial/basic/tutorial.html) (datasets, models)  fit\_transform(), transform(), random\_state  **sklearn.preperocessing** **Normaliser**, **OneHotEncoder**, **LabelEncoder**, **OrdinalEncoder**  scaler – **sklearn.preprocessing StandardScaler, MinMaxScaler, RobustScaler** (Outliers)  **sklearn.model\_selection.train\_test\_split** – 70:30 evaluation metrics  [**https://www.kaggle.com/pmarcelino/data-analysis-and-feature-extraction-with-python**](https://www.kaggle.com/pmarcelino/data-analysis-and-feature-extraction-with-python)Titanic  [**https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/**](https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/)PIMA Indian diabetes  [**https://www.analyticsvidhya.com/blog/2016/07/practical-guide-data-preprocessing-python-scikit-learn/**](https://www.analyticsvidhya.com/blog/2016/07/practical-guide-data-preprocessing-python-scikit-learn/)loan prediction |
| 2 | 19/05/2020 |  | Data Preprocessing in Python  Dimensionality Reduction | Missing data  Data imbalance, Multiclass  Feature selection – Filter, Wrapper, Embedded, Hybrid | missing value [**https://scikit-learn.org/stable/modules/impute.html**](https://scikit-learn.org/stable/modules/impute.html)  data imbalance [**https://www.analyticsvidhya.com/blog/2017/03/imbalanced-data-classification/**](https://www.analyticsvidhya.com/blog/2017/03/imbalanced-data-classification/)  [**https://www.datacamp.com/community/tutorials/feature-selection-python**](https://www.datacamp.com/community/tutorials/feature-selection-python)  [**https://medium.com/@shiwanigupta3005/feature-selection-techniques-for-classification-models-832ebfc6564d**](https://medium.com/@shiwanigupta3005/feature-selection-techniques-for-classification-models-832ebfc6564d)  [**http://ijcttjournal.org/archives/ijctt-v67i6p109**](http://ijcttjournal.org/archives/ijctt-v67i6p109)  [**https://medium.com/analytics-vidhya/feature-selection-methods-with-code-examples-a78439477cd4**](https://medium.com/analytics-vidhya/feature-selection-methods-with-code-examples-a78439477cd4)  [**https://www.kaggle.com/pmarcelino/data-analysis-and-feature-extraction-with-python**](https://www.kaggle.com/pmarcelino/data-analysis-and-feature-extraction-with-python) |
| 3 | 20/05/2020 |  | Dimensionality Reduction  Supervised Learning - Regression | Feature Extraction – Principal Component Analysis, Singular Value Decomposition  scipy (linear algebra, stats)  Linear Regression  Regularization  Performance Metric | [**https://towardsdatascience.com/dimensionality-reduction-toolbox-in-python-9a18995927cd**](https://towardsdatascience.com/dimensionality-reduction-toolbox-in-python-9a18995927cd)  [**https://www.edureka.co/blog/principal-component-analysis/**](https://www.edureka.co/blog/principal-component-analysis/)  [**https://www.analyticsvidhya.com/blog/2019/08/5-applications-singular-value-decomposition-svd-data-science/**](https://www.analyticsvidhya.com/blog/2019/08/5-applications-singular-value-decomposition-svd-data-science/)  [**https://www.guru99.com/scipy-tutorial.html**](https://www.guru99.com/scipy-tutorial.html)  <https://towardsdatascience.com/linear-regression-detailed-view-ea73175f6e86>  <https://towardsdatascience.com/regularization-in-machine-learning-76441ddcf99a>  <https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/> using function  <https://www.dataquest.io/blog/understanding-regression-error-metrics/>  <https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/beginners-guide-regression-analysis-plot-interpretations/tutorial/> |
| 4 | 21/05/2020 |  | Supervised Learning - Classification | Logistic Regression  Support Vector Machine  Neural Network  Decision Tree  Ensembles  k Nearest Neighbor  Deep Learning  Evaluation Metric | [**https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8**](https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8)  [**https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python**](https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python)  [**https://heartbeat.fritz.ai/building-a-neural-network-from-scratch-using-python-part-1-6d399df8d432**](https://heartbeat.fritz.ai/building-a-neural-network-from-scratch-using-python-part-1-6d399df8d432)  [**https://www.python-course.eu/neural\_networks\_with\_scikit.php**](https://www.python-course.eu/neural_networks_with_scikit.php)  [**https://www.datacamp.com/community/tutorials/decision-tree-classification-python**](https://www.datacamp.com/community/tutorials/decision-tree-classification-python)  [**https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/**](https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/)  [**https://towardsdatascience.com/k-nearest-neighbor-python-2fccc47d2a55**](https://towardsdatascience.com/k-nearest-neighbor-python-2fccc47d2a55)  [**https://pythonprogramming.net/introduction-deep-learning-python-tensorflow-keras/**](https://pythonprogramming.net/introduction-deep-learning-python-tensorflow-keras/)  [**https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/**](https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/) |
| 5 | 22/05/2020 |  | Optimization  Use Cases | Genetic Algorithm  Gradient Descent  Underfitting, Overfitting, Bias, Variance, model\_selection learning\_curve, validation\_curve cross\_val\_score  Grid Search random search  Cross Validation  scipy(optimisation)  Virtual Personal Assistant, Traffic Predictions, Price Prediction, Video Surveillance, Facebook friend suggestions, Pinterest, Email Spam and Malware Filter, Online customer support, Search engine result refining, Product recommendation, Online fraud detection | [**https://blog.alookanalytics.com/2017/05/25/automate-your-machine-learning/**](https://blog.alookanalytics.com/2017/05/25/automate-your-machine-learning/) AML  [**https://towardsdatascience.com/gradient-descent-training-with-logistic-regression-c5516f5344f7**](https://towardsdatascience.com/gradient-descent-training-with-logistic-regression-c5516f5344f7)  [**https://www.steveklosterman.com/over-under**](https://www.steveklosterman.com/over-under)  **<https://www.kaggle.com/pmarcelino/data-analysis-and-feature-extraction-with-python>**  [**https://machinelearningmastery.com/how-to-tune-algorithm-parameters-with-scikit-learn/**](https://machinelearningmastery.com/how-to-tune-algorithm-parameters-with-scikit-learn/)  [**https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/**](https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/)  [**https://www.guru99.com/scipy-tutorial.html**](https://www.guru99.com/scipy-tutorial.html)  [**https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0**](https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0)  [**https://www.dezyre.com/article/top-10-machine-learning-projects-for-beginners/397**](https://www.dezyre.com/article/top-10-machine-learning-projects-for-beginners/397) |