



# Kissipo Learning for Deep Learning

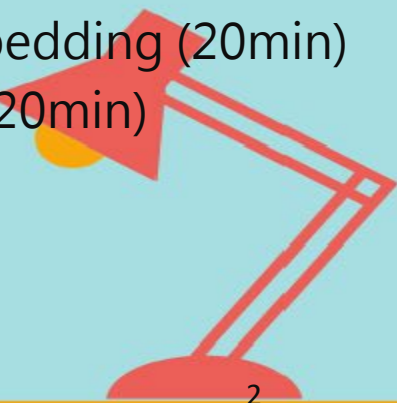
## Topic 6: Scikit-learn quick tutorial (15min)

Hsueh-Ting Chu

KLDL-W2-06

# Topics

- Topic 01: Introduction to Deep Learning (20min)
- Topic 02: KISS Learning for Deep Learning (20min)
- Topic 03: Python quick tutorial (20min)
- Topic 04: Numpy quick tutorial (15min)
- Topic 05: Pandas quick tutorial (15min)
- **Topic 06: Scikit-learn quick tutorial (15min)**
- Topic 07: OpenCV quick tutorial (15min)
- Topic 08: Image Processing basics (20min)
- Topic 09: Machine Learning basics (20min)
- Topic 10: Deep Learning basics (20min)
- Topic 11: TensorFlow overview (20min)
- Topic 12: CNN with TensorFlow (20min)
- Topic 13: RNN with TensorFlow (20min)
- Topic 14: PyTorch overview (20min)
- Topic 15: CNN with PyTorch (20min)
- Topic 16: RNN with Pytorch (20min)
- Topic 17: Introduction to AOI (20min)
- Topic 18: AOI simple Pipeline (A) (20min)
- Topic 19: AOI simple Pipeline (B) (20min)
- Topic 20: Introduction to Object detection (20min)
- Topic 21: YoloV5 Quick Tutorial (20min)
- Topic 22: Using YoloV5 for RSD (20min)
- Topic 23: Introduction to NLP (20min)
- Topic 24: Introduction to Word Embedding (20min)
- Topic 25: Name prediction project (20min)



# Content

- Topic 6: Scikit-learn quick tutorial
  - Introduction to Scikit-learn
  - Scikit-learn algorithms
  - Iris example



# scikit-learn

## *Machine Learning in Python*

[Install](#) [User Guide](#) [API](#) [Examples](#) [Community](#) [More](#)

### scikit-learn

*Machine Learning in Python*

[Getting Started](#) [Release Highlights for 1.1](#) [GitHub](#)

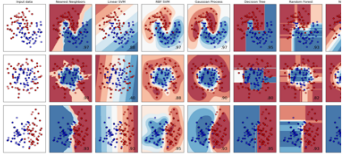
- Simple and efficient tools for predictive data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

#### Classification

Identifying which category an object belongs to.

**Applications:** Spam detection, image recognition.

**Algorithms:** SVM, nearest neighbors, random forest, and more...



Examples

#### Regression

Predicting a continuous-valued attribute associated with an object.

**Applications:** Drug response, Stock prices.

**Algorithms:** SVR, nearest neighbors, random forest, and more...



Examples

#### Clustering

Automatic grouping of similar objects into sets.

**Applications:** Customer segmentation, Grouping experiment outcomes

**Algorithms:** k-Means, spectral clustering, mean-shift, and more...



Examples

#### Dimensionality reduction

Reducing the number of random variables to consider.

**Applications:** Visualization, Increased efficiency

**Algorithms:** PCA, feature selection, non-negative matrix factorization, and more...



#### Model selection

Comparing, validating and choosing parameters and models.

**Applications:** Improved accuracy via parameter tuning

**Algorithms:** grid search, cross validation, metrics, and more...

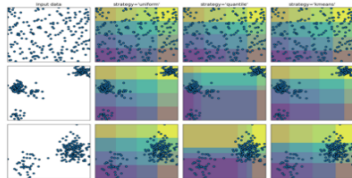


#### Preprocessing

Feature extraction and normalization.

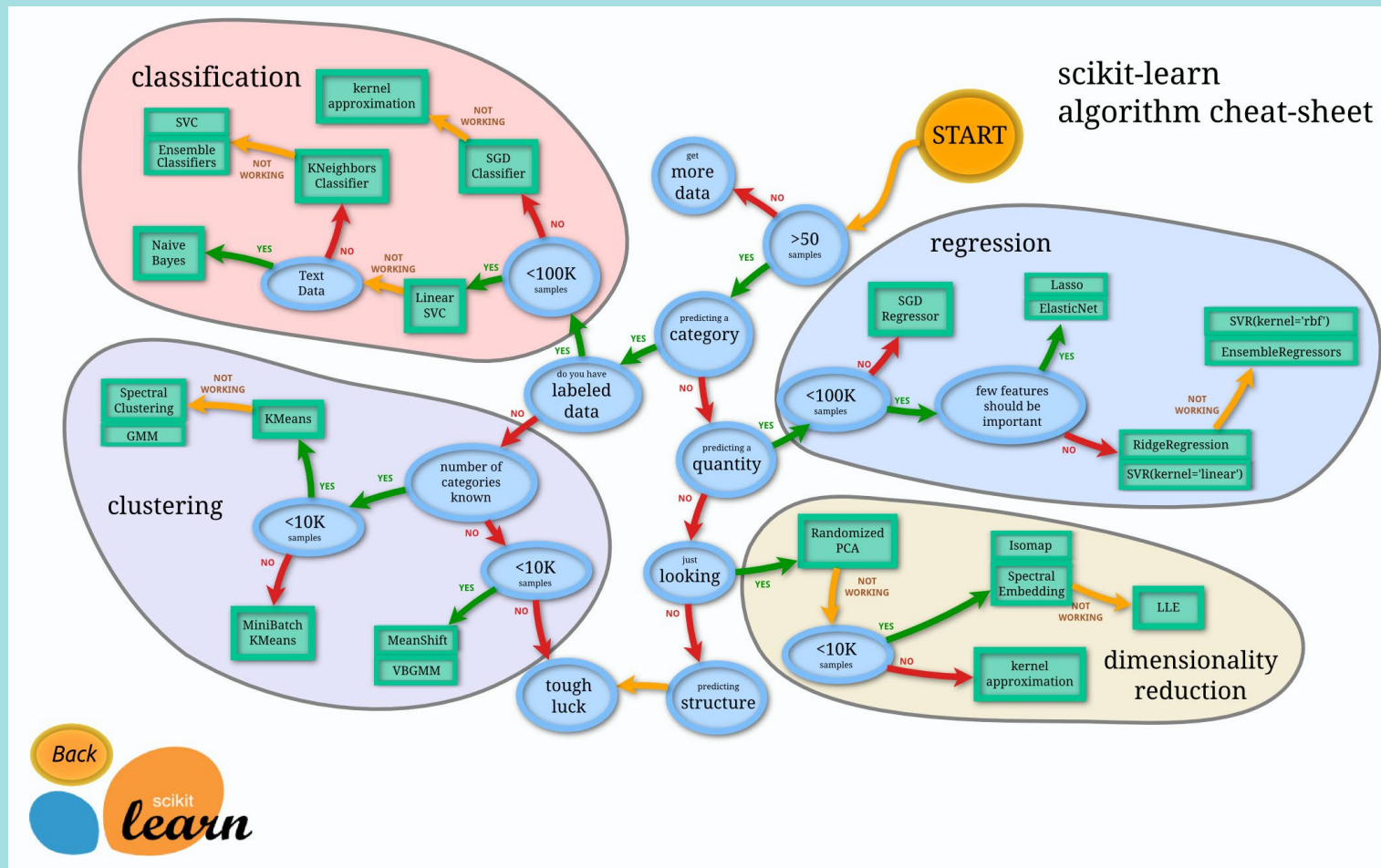
**Applications:** Transforming input data such as text for use with machine learning algorithms.

**Algorithms:** preprocessing, feature extraction, and more...





# ML algorithm cheat sheet



# *Iris dataset*

- The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor).
- Four features were measured from each sample: the length and the width of the sepals and petals
- Examples:
  - 5.1,3.5,1.4,0.2,Iris-setosa
  - 4.9,3.0,1.4,0.2,Iris-setosa
  - 6.7,3.1,4.4,1.4,Iris-versicolor
  - 5.6,3.0,4.5,1.5,Iris-versicolor
  - 6.1,2.6,5.6,1.4,Iris-virginica
  - 7.7,3.0,6.1,2.3,Iris-virginica





# *Iris dataset*

- *Iris setosa*
- Iris versicolor
- Iris virginica



Thanks!

Q&A

