

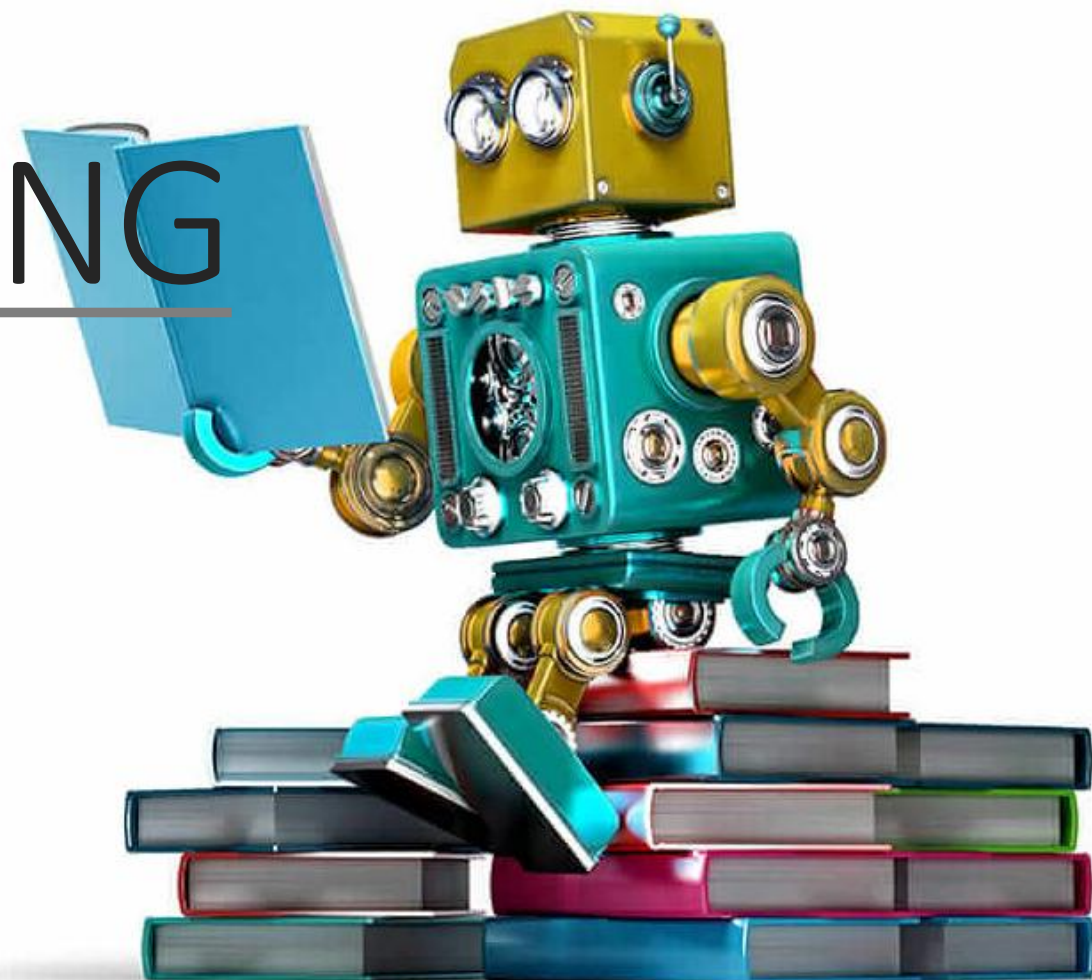
# MACHINE LEARNING

---

## LAB1

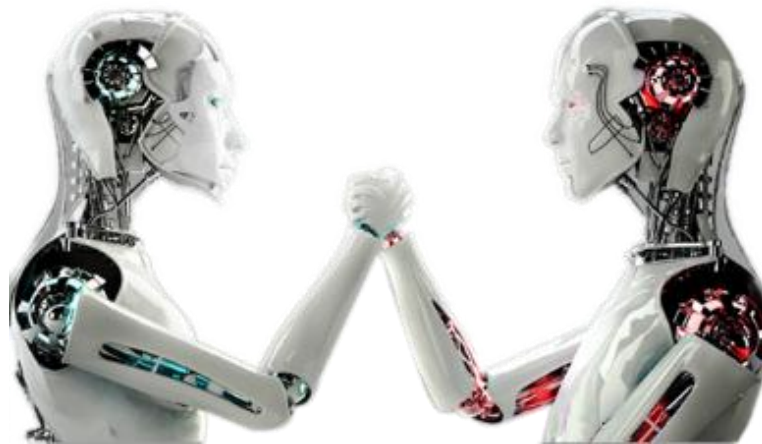
**贾艳红** Jana

Email: [jiayh@mail.sustech.edu.cn](mailto:jiayh@mail.sustech.edu.cn)



PART ONE

# Lab Introduction





# Class Schedule



■ Lab: Wednesday 16:20-18:10 Teaching Building III Room 504

## ■ Grading policy

Final Exam (in-class):	20%	Midterm Exam(take-home): 10%
Assignments (8~12 times):	20%	Quizzes(<=10 times): 10%
<b>LAB Task:</b>	<b>15%</b>	<b>Final Projects (4 per group): 20%</b>
<b>LAB attendance:</b>	<b>5%</b>	

90~93: A-	94~97: A	98~100: A+
80~82: B-	83~86: B	87~89: B+
70~72: C-	73~76: C	77~79: C+
60~62: D-	63~66: D	67~69: D+



# Attendance



- Mandatory
- Be on Time!
- Don't expect labs to end early.
- Let instructor know if you anticipate an absence.





# Textbook and Lab Notes



## Textbooks:

- 1 Pattern Recognition and Machine Learning, by Christopher M. Bishop, 2006 Springer
- 2 Machine Learning in Action, by Peter Harrington, 2012, Manning
- 3 Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow

## Other books:

- 1 机器学习, 周志华
- 2 Dive in Deep Learning, by Aston zhang, Zachary C.Lipton, Mu Li, and Alexander J.Smola
- 3 Reinforcement Learning: An Introduction, by Richard S. Sutton
- 4 the Elements of Statistical Learning Data Mining, Inference, and Prediction

## Paper reading:

- 1 Ghahramani Z. Probabilistic machine learning and artificial intelligence, Nature, 2015
- 2 Lecun Y, Bengio Y, Hinton G. Deep learning, Nature, 2015
- 3 Littman M L. Reinforcement learning improves behavior from evaluative feedback, Nature, 2015

BB:[https://bb.sustech.edu.cn/webapps/blackboard/execute/announcement?method=search&context=course&course\\_id=4109\\_1&handle=cp\\_announcements&mode=cpview](https://bb.sustech.edu.cn/webapps/blackboard/execute/announcement?method=search&context=course&course_id=4109_1&handle=cp_announcements&mode=cpview)



# Other considerations



- TA:

LAB	TA
Wednesday 16:20-18:10 Teaching Building III Room 504	骆家睿、张睿豪、张俊翔、 黄珂邈

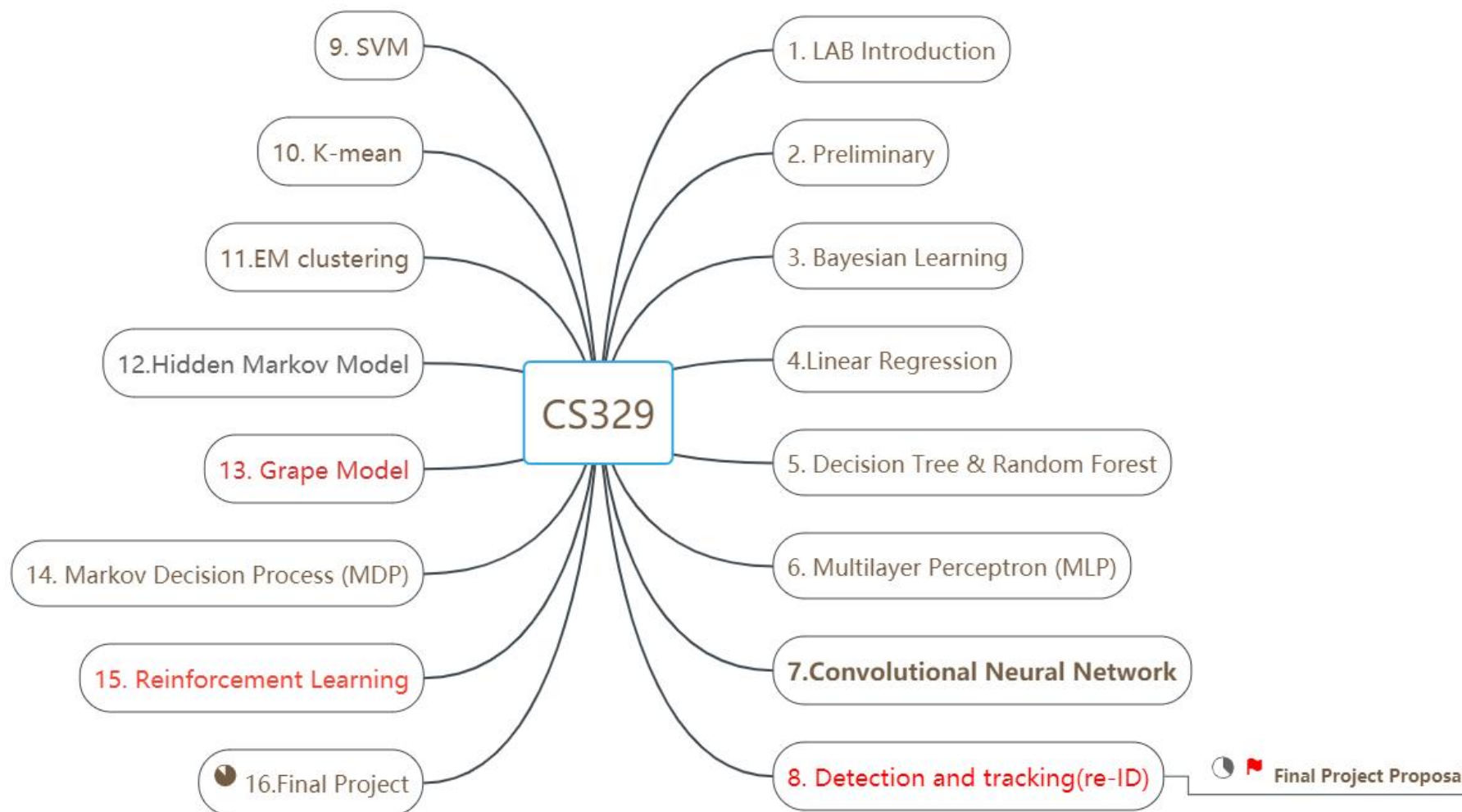
- Communication platform



群名称:2022机器学习H  
群 号:820664075



# Lab Schedule





# Final Projects



- 1 Reinforcement learning based planning using a self-driving car simulator
- 2 Generation of annotated self-driving datasets with the CARLA simulator
- 3 Segmentation of 2D/3D measurements for self-driving applications
- 4 Detection and recognition of traffic signs for self-driving applications
- 5 Detection and tracking of 2D/3D objects for self-driving applications





# Lab Report & Grading Guidelines



- Lab report
  - ✓ Follow the lab instructions
  - ✓ Testing the algorithms and discuss your observations [25%]
  - ✓ Develop new algorithms (bonus)
  - ✓ Describe your methods (bonus)
  - ✓ Validating the algorithms and discuss your observations[75%]

**we will deduct scores according to this rule, but we will not strictly follow it all the time. In some cases, we may decide to detract more or less points based on the perceived gravity of the mistakes.**



# Final Project & Grading Guidelines



- Proposal (with Presentation) (2~4 people) [20%]
- Midterm report [10%]
- Final Report(with Presentation and Demos [70%]
- Abstract (motivation, method, novelties, validation) [10%]
- Introduction (Background, Motivation, Challenges, Rationale, Contributions) [15%]
- Existing Methods of Related Work, Novelties of This Work [15%]
- System Setup, Problem Statement [10%]
- Proposed Method [15%]
- Results (Experiment platform, experiment setup, results) [10%]
- Discussion (Comparison and analysis, speed, accuracy, robustness, strengths and limits) [15%]
- Conclusion (Main method, major results, future work) [05%]
- References [05%]

**we will deduct scores according to this rule, but we will not strictly follow it all the time. In some cases, we may decide to detract more or less points based on the perceived gravity of the mistakes.**



CS329 Project team r...

扫一扫二维码打开或分享给好友



- 腾讯文档 -

可多人实时在线编辑，权限安全可控

***Please open the Shared Excel, and complete the grouping. DDL is week8.***



- **Lab Tutotial web sites**

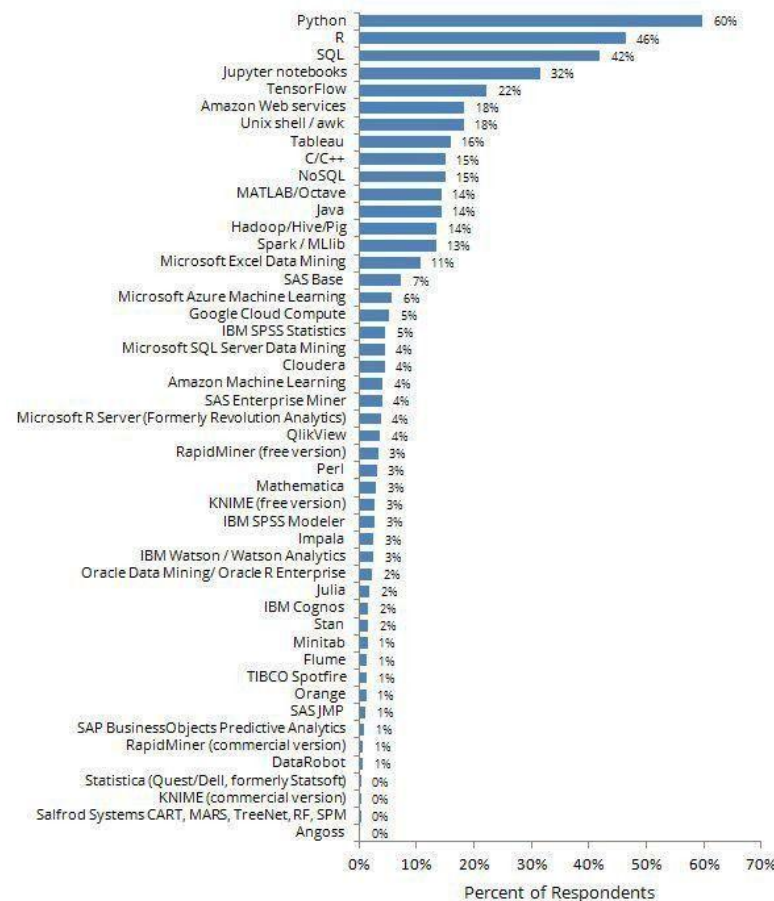
- <http://scikit-learn.org/stable/tutorial/basic/tutorial.html>
- <https://pythonprogramming.net/python-fundamental-tutorials/>
- <https://pythonprogramming.net/data-analysis-tutorials/>
- <https://pythonprogramming.net/robotics-tutorials/>
- <http://prml.github.io/>
- [https://scikit-learn.org/dev/user\\_guide.html](https://scikit-learn.org/dev/user_guide.html)



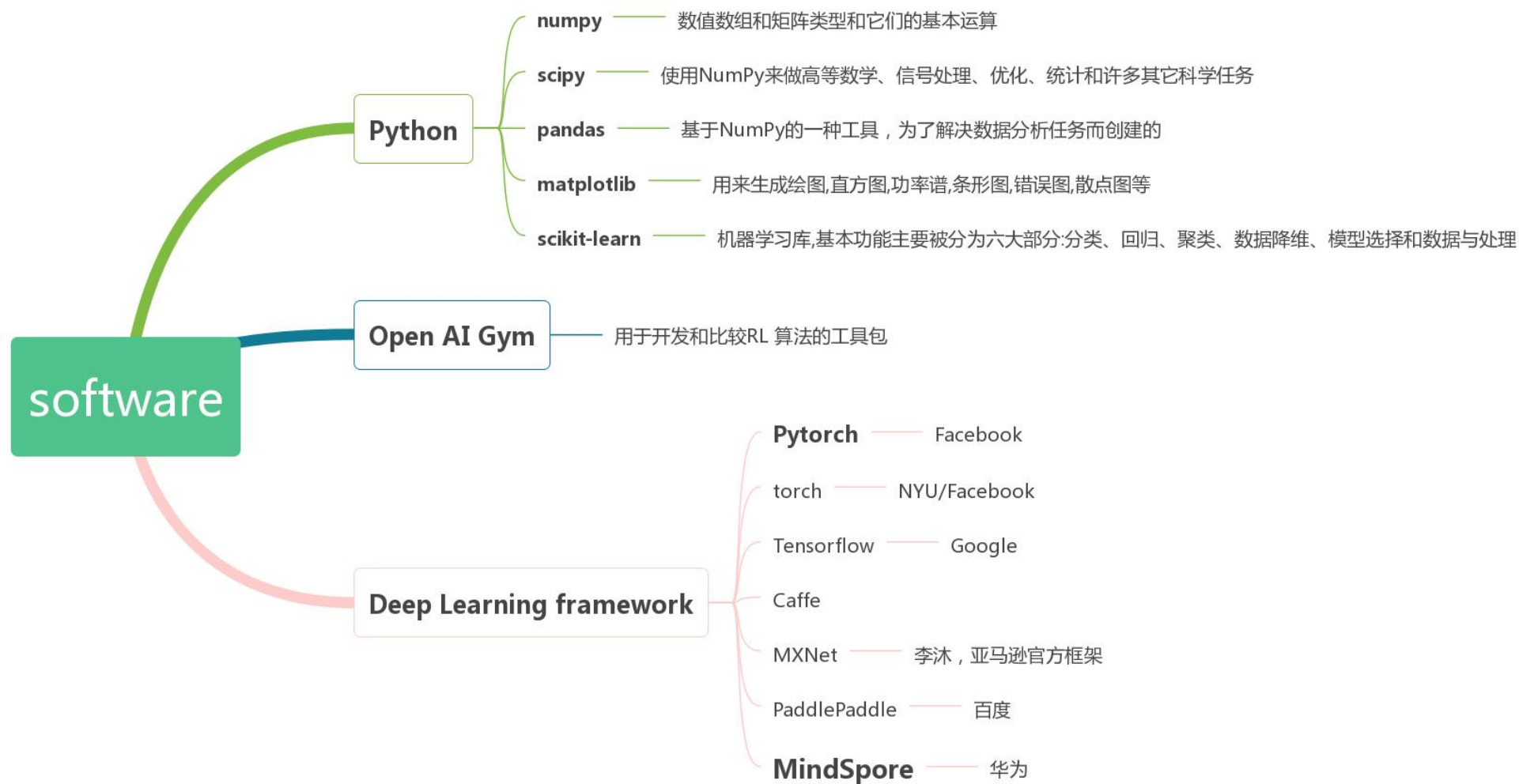
# Which Programming language is popular for ML?



## Data Science / Analytics Tools, Technologies and Languages Used in Past Year

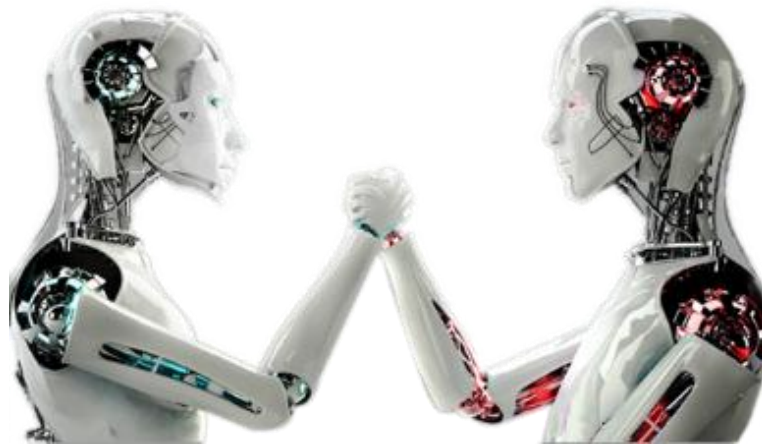


Data are from the Kaggle 2017 The State of Data Science and Machine Learning study. You can learn more about the study and download the data here: <https://www.kaggle.com/surveys/2017>. Respondents were asked to indicate for work, which data science/analytics tools, technologies, and languages they used in the past year. A total of 10153 respondents answered the question.



## PART TWO

# Prepare the development environmentIntroduction





# Prepare the development env intro



## 1. Install python + vscode

- [Getting Started with Python in VS Codev2](#)
- [Install Miniconda on Win10](#)

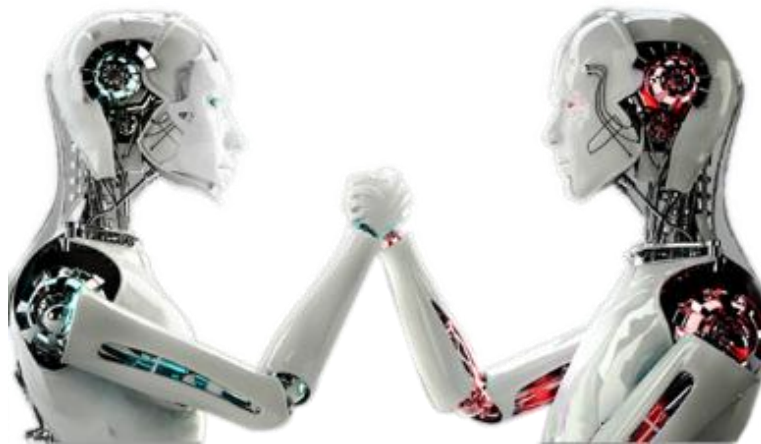
## 2. Install essential libraries for Machine Learning in Python

- **Numpy**, a package with support for mathematical operations on multidimensional data—was the most imported package, used in nearly three-quarters of machine learning and data science projects. (<https://numpy.org/>)
- **scipy**, a package for scientific computation. (<https://www.scipy.org/>)
- **pandas**, a very popular library, it can easily fetch data from different sources like SQL databases, CSV, Excel, JSON files and manipulate the data to perform operations on it. (<https://pandas.pydata.org/>)
- **matplotlib**, a visualization library. (<https://matplotlib.org/>)
- **scikit-learn**, Scikit-learn is one the most popular ML libraries. It supports many supervised and unsupervised learning algorithms. Examples include linear and logistic regressions, decision trees, clustering, k-means and so on. For a novice in ML, Scikit-learn is a more-than-sufficient tool to work with, until you start implementing more complex algorithms. (<https://scikit-learn.org/stable/index.html>)



## PART FOUR

# Lab Task





1. Install and configure your development environment on your own computer.

2. Implement more than one methods use the Scikit-learn Toolbox.

- Parameter regularization method: Lasso(L1 norm)/Ridge(L2 norm) (<https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.Normalizer.html#sklearn.preprocessing.Normalizer>)
- Linear Models: LDA ( Linear Discriminant Analysis ), SVM, Perceptron ([https://scikit-learn.org/stable/modules/linear\\_model.html](https://scikit-learn.org/stable/modules/linear_model.html))
- Nonlinear Models: KNN( Nearest Neighbors ), Decision Trees, Ensemble methods, Kernel SVM, Multi-layer Perceptron

# Thanks

贾艳红 Jana

Email: [jiayh@mail.sustech.edu.cn](mailto:jiayh@mail.sustech.edu.cn)

