

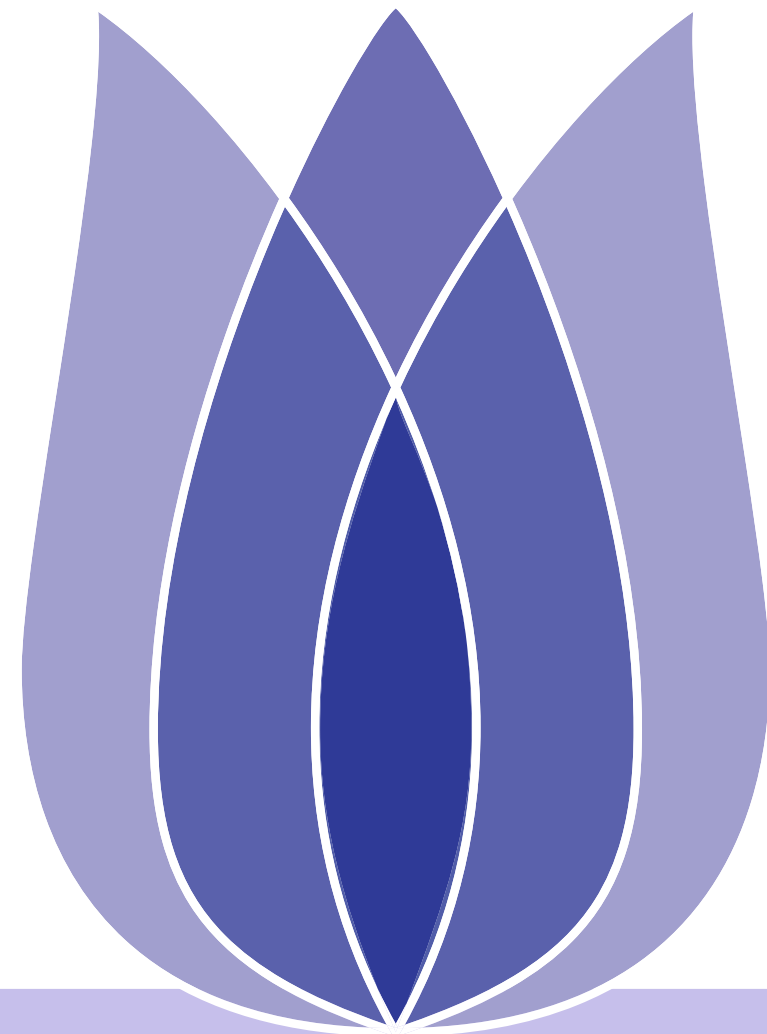
The First Report

Wang Mingxi

Jilin University

College of Computer Science and Technology

(None)





Overview

[Learning content](#)

[Hands on practice](#)

[Plan for the following week](#)

Learning content

Decision tree

Ensemble Learning

Neural Network

Hands on practice

Generative Adversarial Network GAN

Plan for the following week

Plan for the following week



Learning content

- Decision tree
- Ensemble Learning
- Neural Network

Hands on practice

Plan for the following week

Learning content



Decision tree: A decision tree is a prediction model used to predict the category of samples. In the structure of these trees, leaf nodes give categories and inner nodes represent attributes.

Learning content

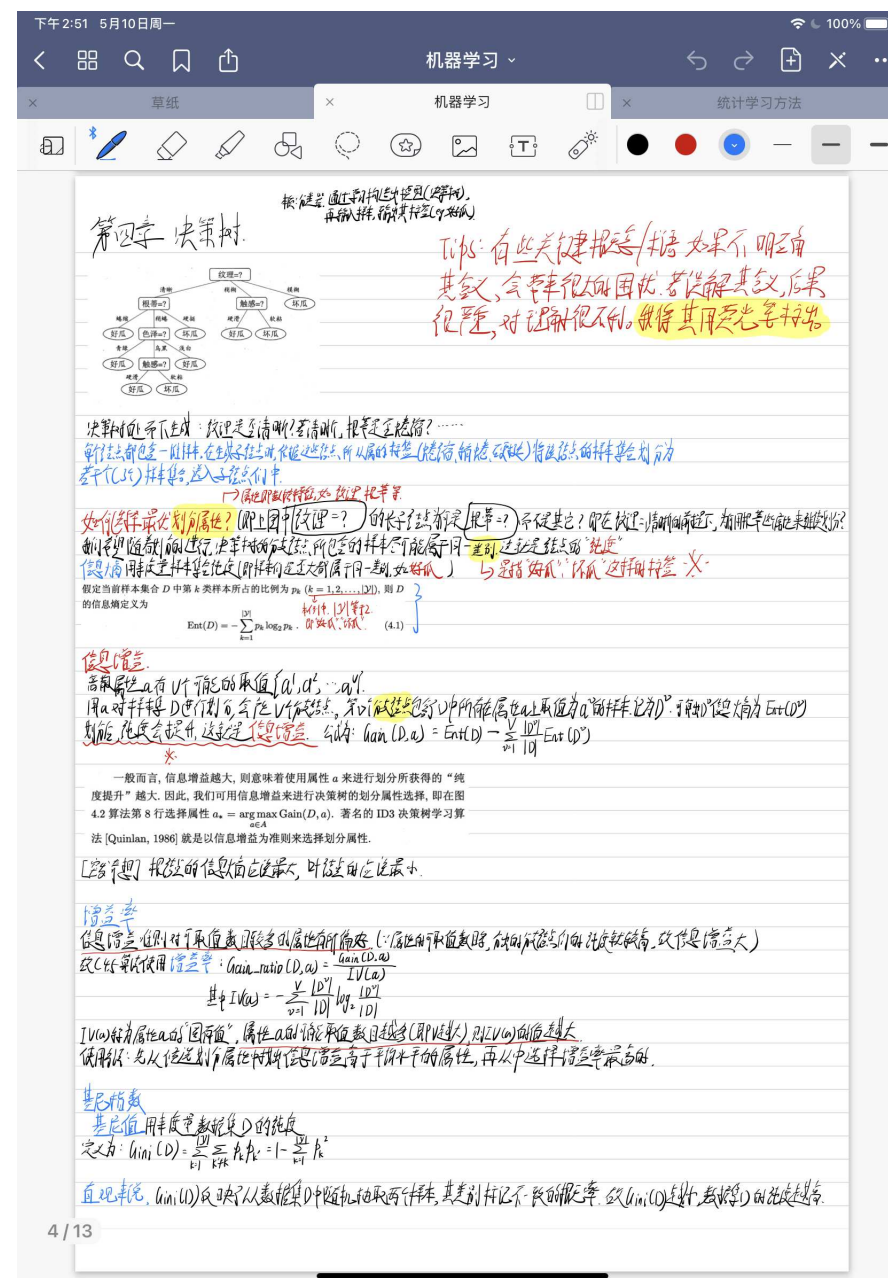
Decision tree

Ensemble Learning

Neural Network

Hands on practice

Plan for the following week



A close-up photograph of several bright yellow tulips. The petals are in various stages of opening, with some showing the characteristic six-petaled shape. The background is dark and out of focus, making the vibrant yellow flowers stand out.

Plan for the following week

下午 2:51 5月10日周一

机器学习

草纸
机器学习
统计学习方法

第8章 集成学习

Boosting 的框架：AdaBoost

(一种弱学习)

Boosting 是一种可解释学习提升为强学习的算法。这算法的工作机制是：先从一个弱分类器得到一个初步学习，然后根据学习到的结果对训练样本分布进行调整，使得那些被弱分类器分错的样本分布权重增加多次迭代，然后基于调整后的样本分布来训练下一个基学习器，如此重复进行，直至基学习器的性能达到预先设定的值，最后将若干个基学习器进行加权组合。

$$\text{集成学习模型} : G(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \dots + \alpha_m f_m(x)$$

$\alpha_j f_j(x)$ 称为该子模型的贡献。 $\alpha > 0$ 且满足： $\sum_{j=1}^m \alpha_j = 1$

$G(x)$ 中每个子模型 f_j 表示该个体学习器的重要性！

(带权投票)

一个弱学习器标：

$$\text{sign}(f_i(x)) = \begin{cases} 1 & x < 2.5 \\ -1 & 2.5 < x < 5.5 \\ 1 & 5.5 < x < 8.5 \\ 1 & x > 8.5 \end{cases}$$

x	1	2	3	4	5	6	7	8	9
输出 sign(f _i (x))	-1	-1	-1	-1	1	1	1	1	1

注意：sign(f_i(x)) 返回的是数据上属于哪个类别（Sampling）！

(结合上面有3个假设)

Bagging (一种弱学习)

Bagging [Breiman, 1996a] 是并行式集成学习方法最著名代表。从名字即可看出，它直接源于我们在 2.2.3 节介绍的自助采样法 (bootstrap sampling)。它包含 m 个样本的数据集，我们先随机取出一个样本放入训练集中，再把这个样本放回数据集做替换，使得下次采样时该样本有可能被采到。这样，经过多次随机采样操作，我们将得到 m 个样本的采集集；若训练集中的所有样本在采样集里多次出现，有的则从未出现。(由式(2.1)可知，初始训练集中的某些 x_i 的副本可能出现多次)

利用这些，我们就可以得到 m 个弱基分类器的采样集，然后将每个弱基分类器训练成一个基学习器，再将这些基学习器进行综合，这就是 Bagging 的基本思想。

与之相对的，adaboost 是用一个（这里重点）数据做训练，每次训练出一个基学习器，与之前训练的 adaboost 是一个（这里重点）数据做训练，每次训练出一个基学习器。

Bagging 的算法描述如图 8.5 所示。

```

输入：训练集 D = {(x1, y1), (x2, y2), ..., (xm, ym)};
      基学习算法 L;
      弱函数 T;
过程：
1. for i = 1, 2, ..., T do
2.   bi ← L(D, yb);
3. end for
输出：H(x) = argmax Σt=1T ht(x) = pL → 这可能就是良算法，那个 y (即类别) 得到的基分类器 y 的票数总和最多，那个 y 作为集成分类器的输出。
    
```

图 8.5 Bagging 算法

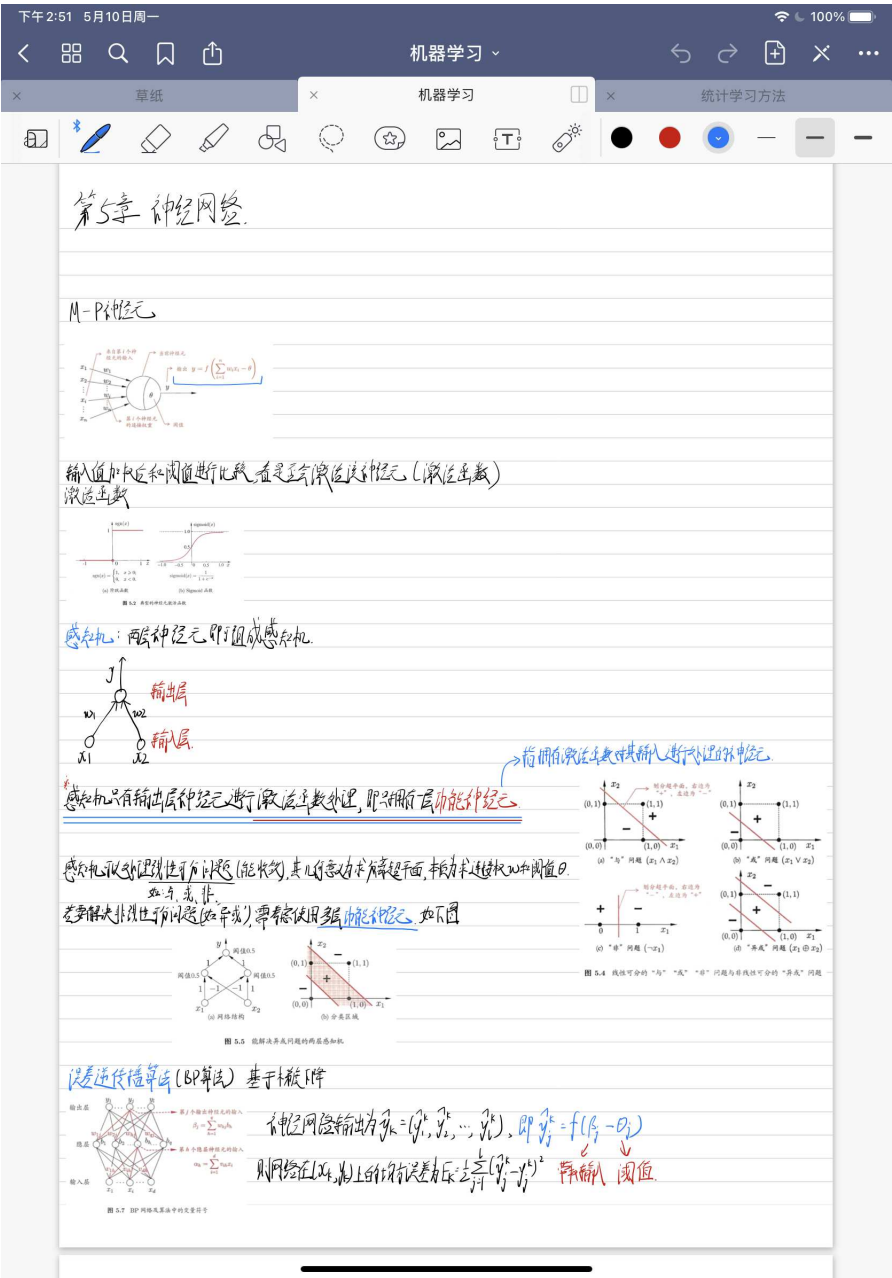
AdaBoost	只适用二分类任务	单一弱分类
Bagging	适用于多类任务	多个弱分类器下个数据集



Neural Network

Learning content
Decision tree
Ensemble Learning
Neural Network
Hands on practice
Plan for the following week

Neural Network:A mathematical or computational model that imitates the structure and function of a biological neural network and is used for estimating or approximating functions.





[Learning content](#)

Hands on practice

Generative Adversarial Network GAN

[Plan for the following week](#)

Hands on practice

Generative Adversarial Network GAN

Learning content

Hands on practice

Generative Adversarial Network GAN

Plan for the following week

- Generating Adversarial networks (GAN) is a method of unsupervised learning in which two neural networks play each other against each other.

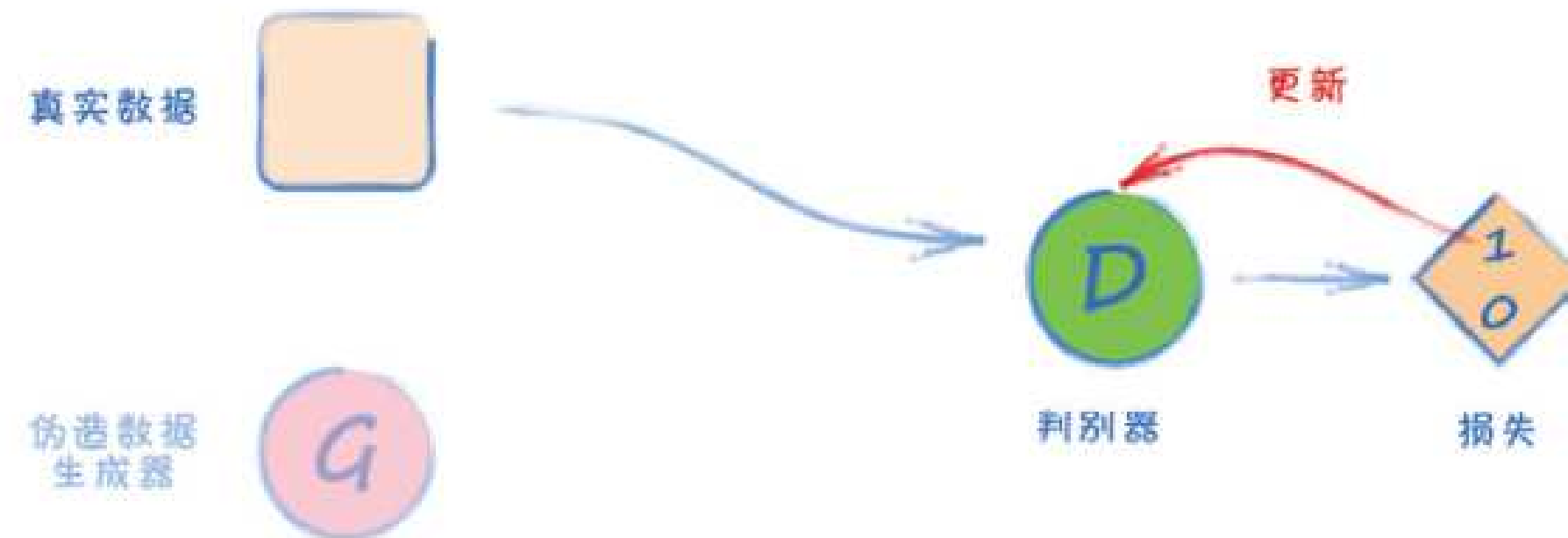


Figure 1



TULIP

Team for Universal Learning and Intelligent Processing

Generative Adversarial Network GAN

- Learning content
- Hands on practice
- Generative Adversarial Network GAN
- Plan for the following week

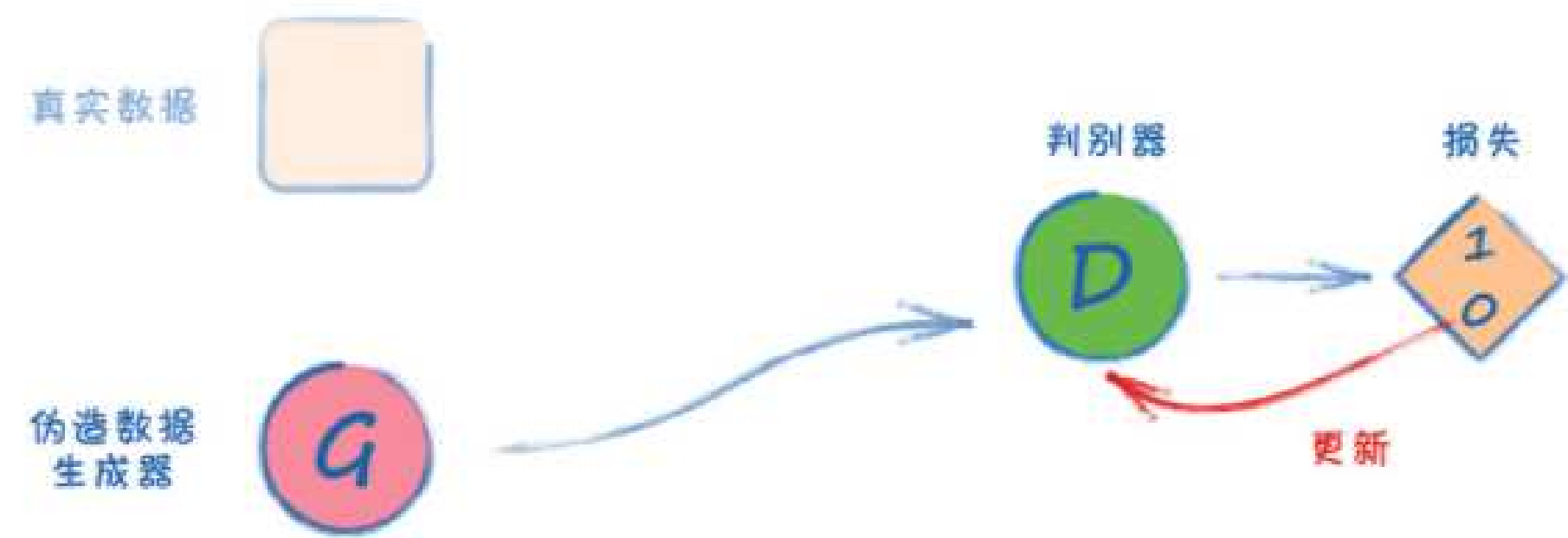


Figure 2: preview

Generative Adversarial Network GAN

- Learning content
- Hands on practice
- Generative Adversarial Network GAN
- Plan for the following week

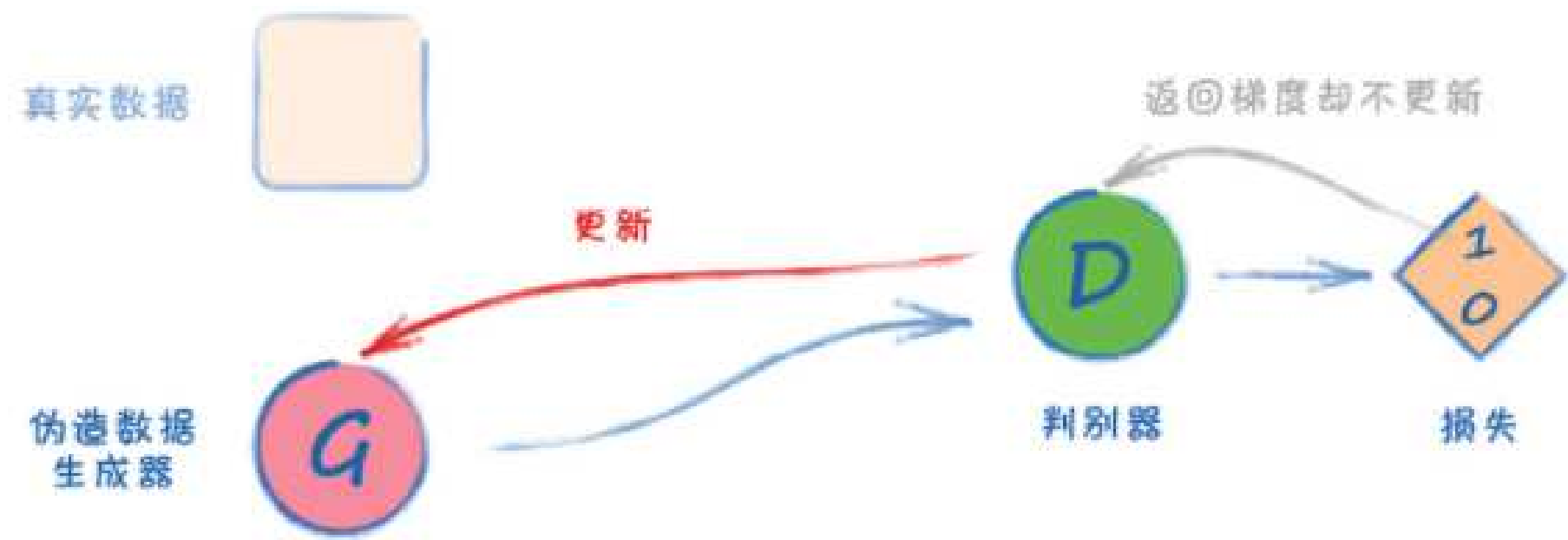


Figure 3: preview



[Learning content](#)

[Hands on practice](#)

[Plan for the following week](#)

[Plan for the following week](#)

Plan for the following week



Plan for the following week

- [Learning content](#)
- [Hands on practice](#)
- [Plan for the following week](#)
- [Plan for the following week](#)

- I’ll continue to learn the basics next week.
- I’m going to go ahead and program this stuff out.



Contact Information

Wang Mingxi
College of Computer Science and Technology
Jilin University, China

-  MXWANG@TULIP.ACADEMY
-  TEAM FOR UNIVERSAL LEARNING AND INTELLIGENT PROCESSING

