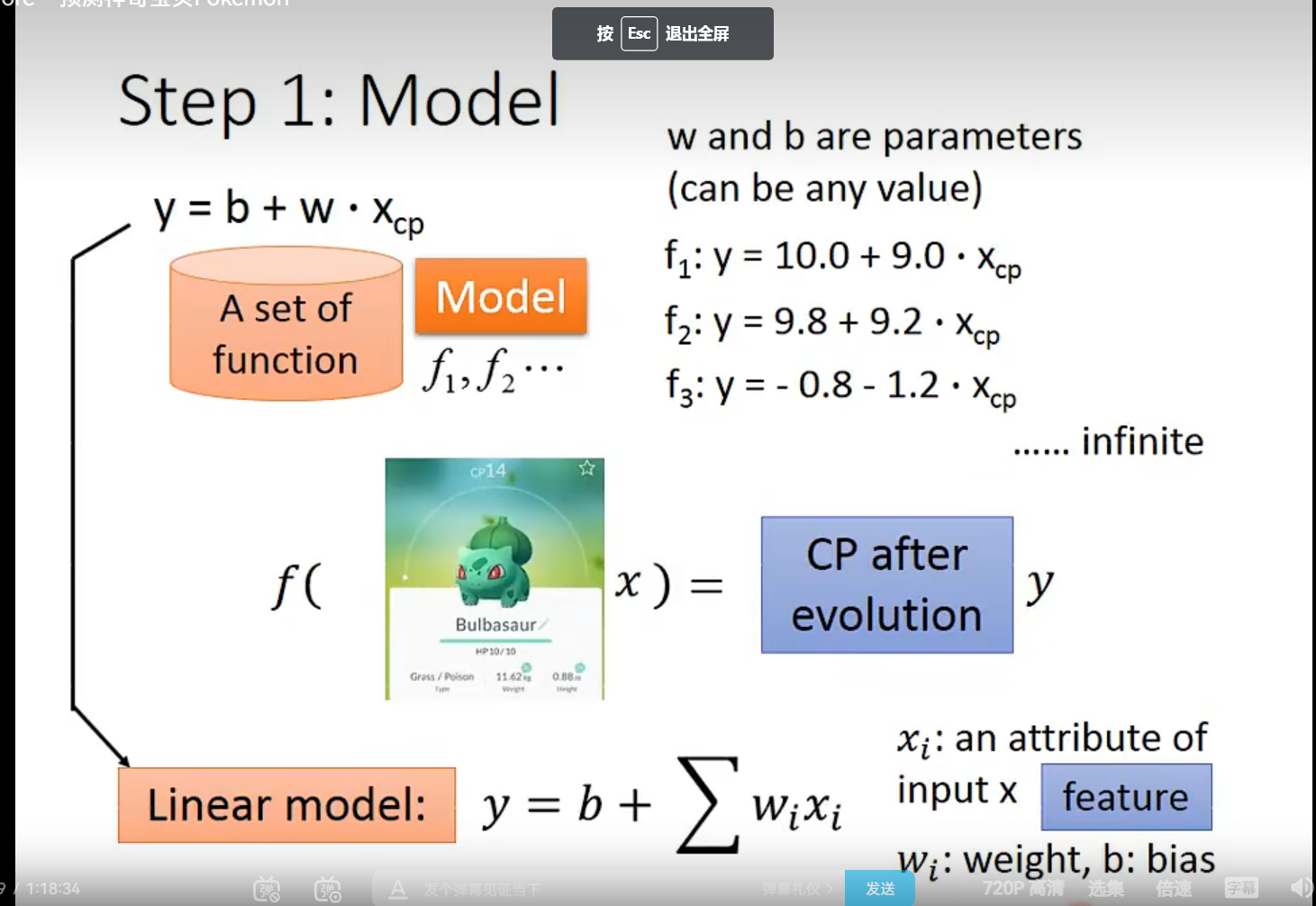
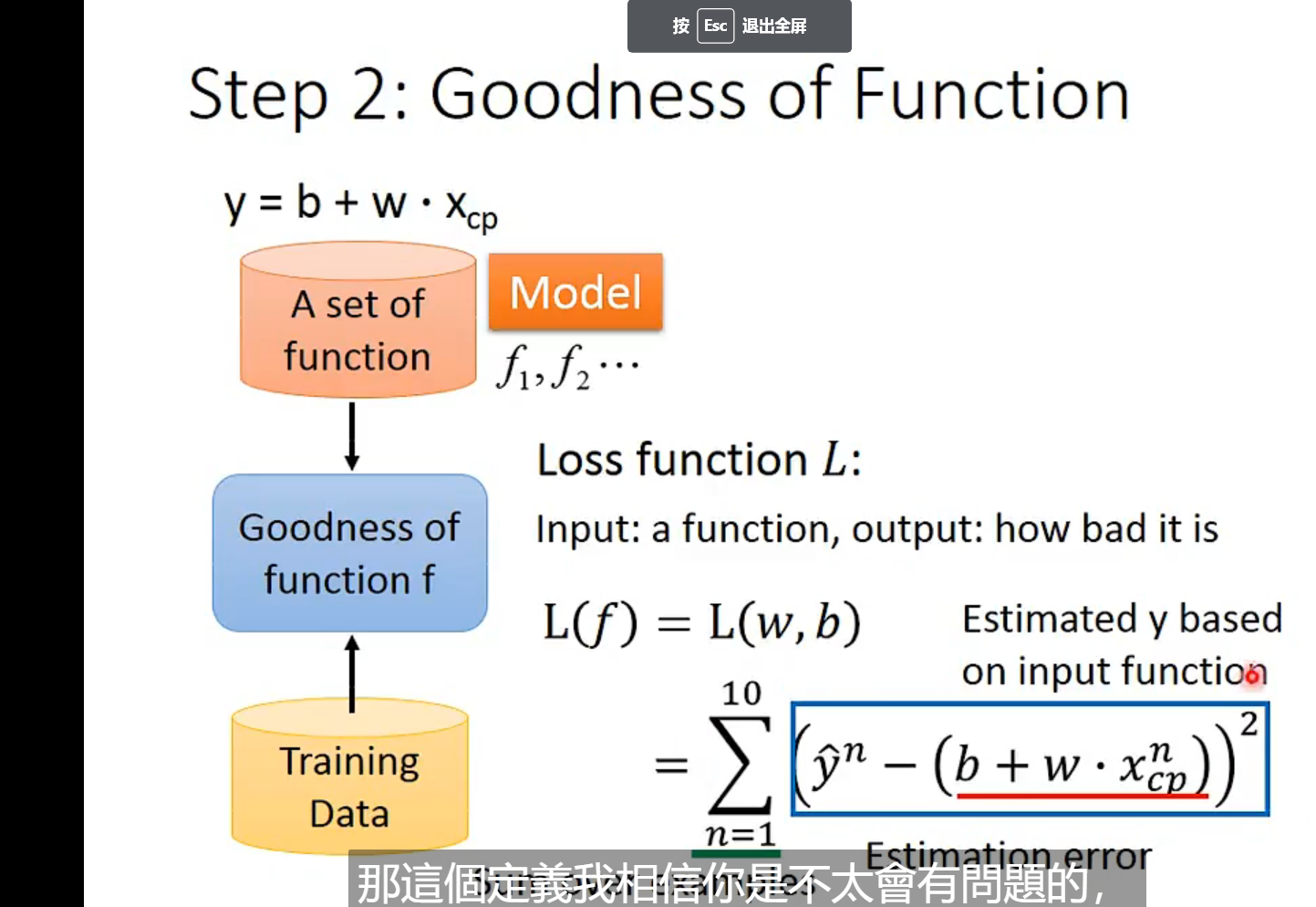
# Regression

预测宝可梦的cp值：

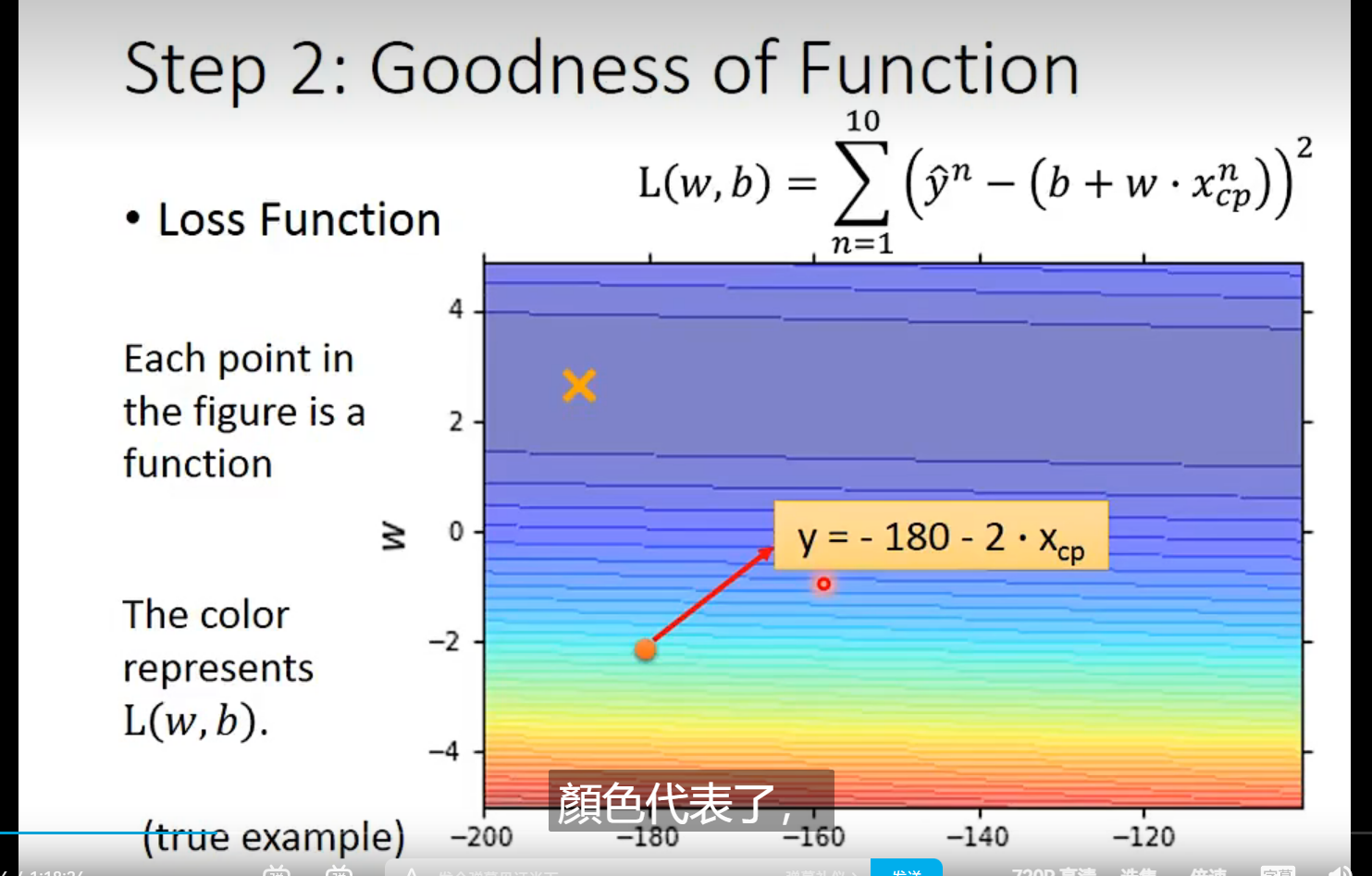
Step 1：Model



估测误差（用方差）

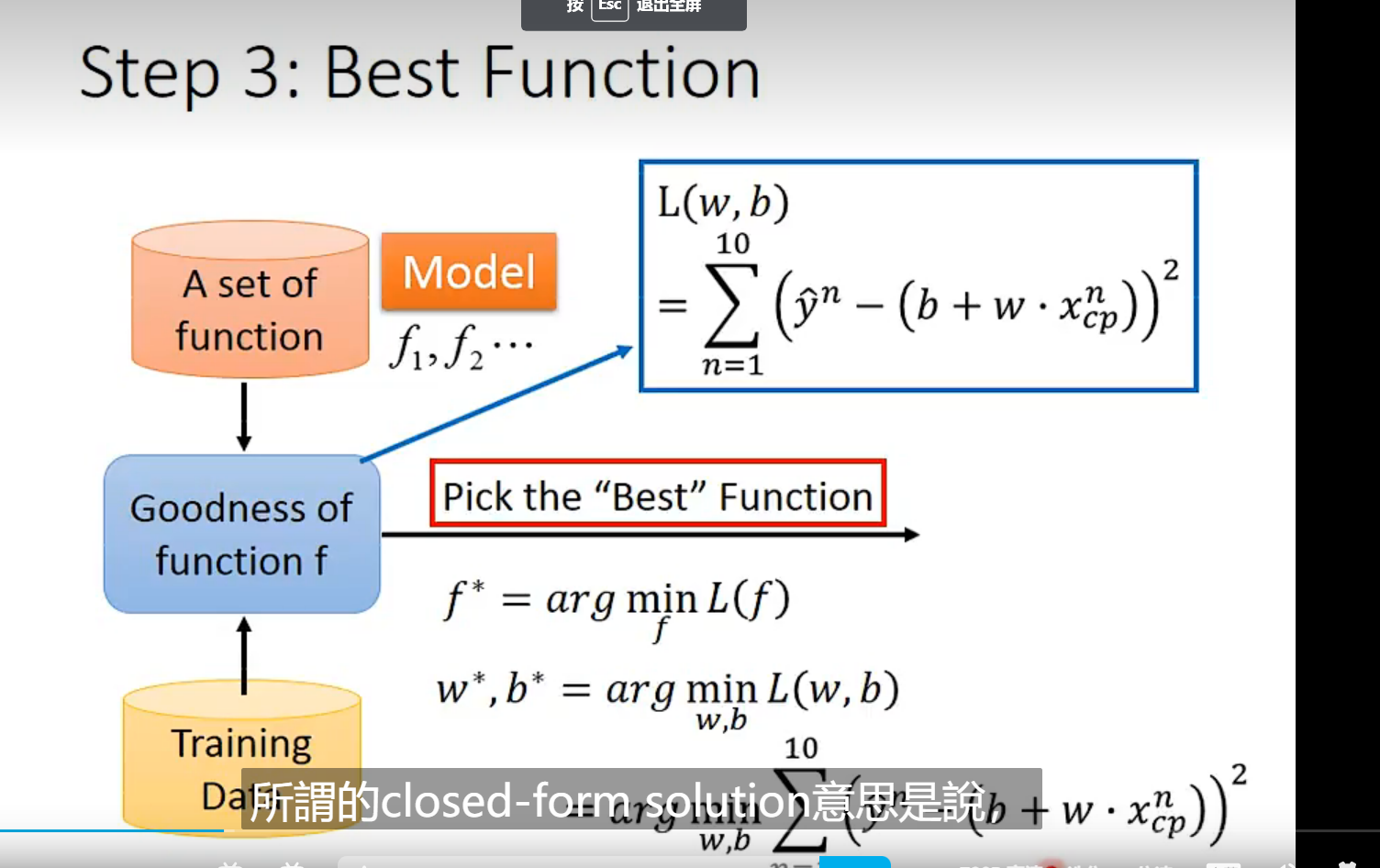


数据分析建立误差（红色代表误差越大，蓝色代表误差越小）



通过检验无数个w，b形成无数条liner model来看Loss

通过找到Loss最小，来找到the best function

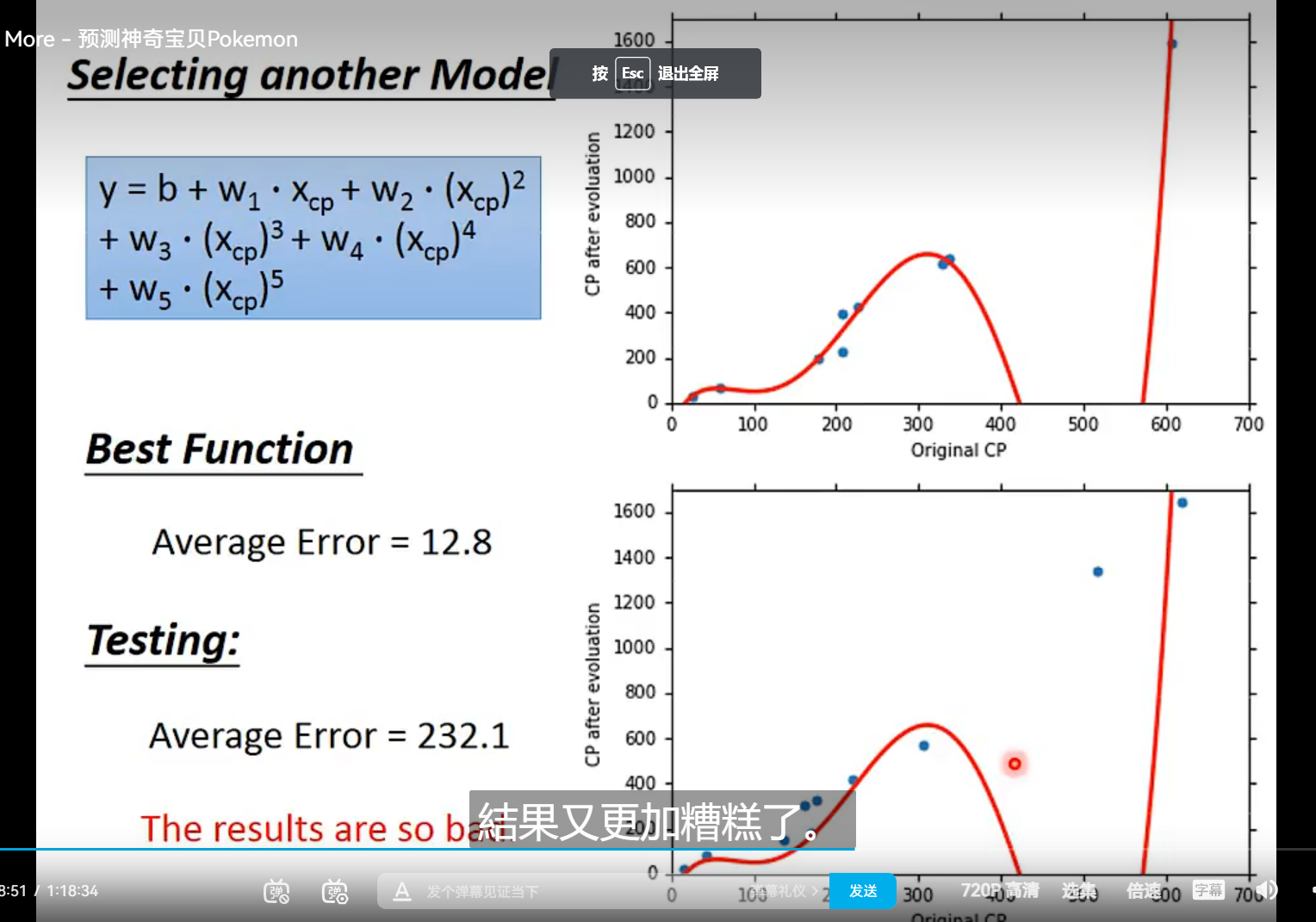


Step 3：也可以通过Gradient Descent（微分判断，找极值），这个在第一课有讲。

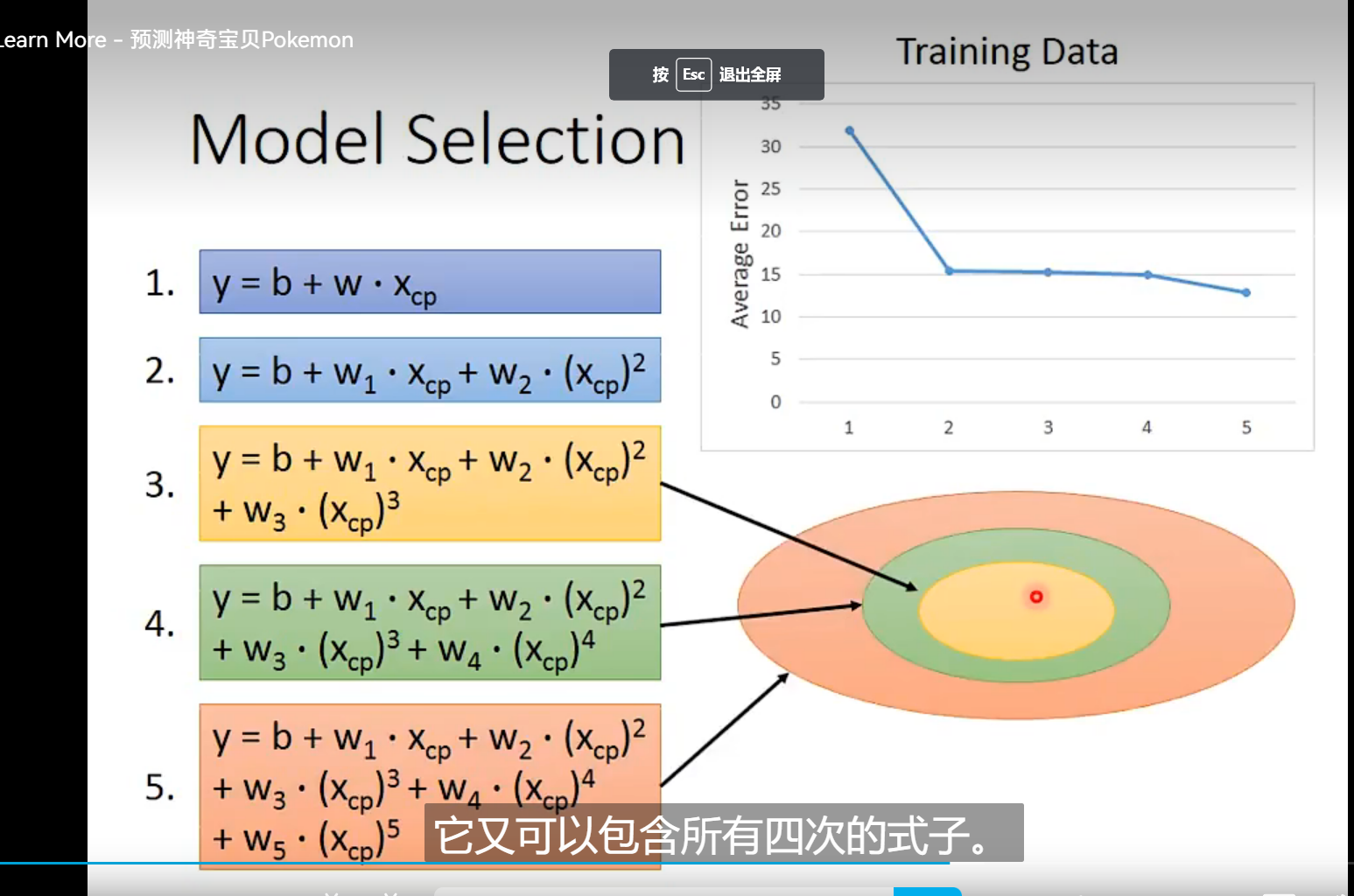
Training date算出来的Loss值是要比Testing date的Loss值要大，因为你在training date算出来的Loss的Liner Model去的Loss最小，最贴切。

Selecting another Moder：

找多次函数使得Loss更加偏小



也不一定你选的函数越复杂你的准确度就越高。

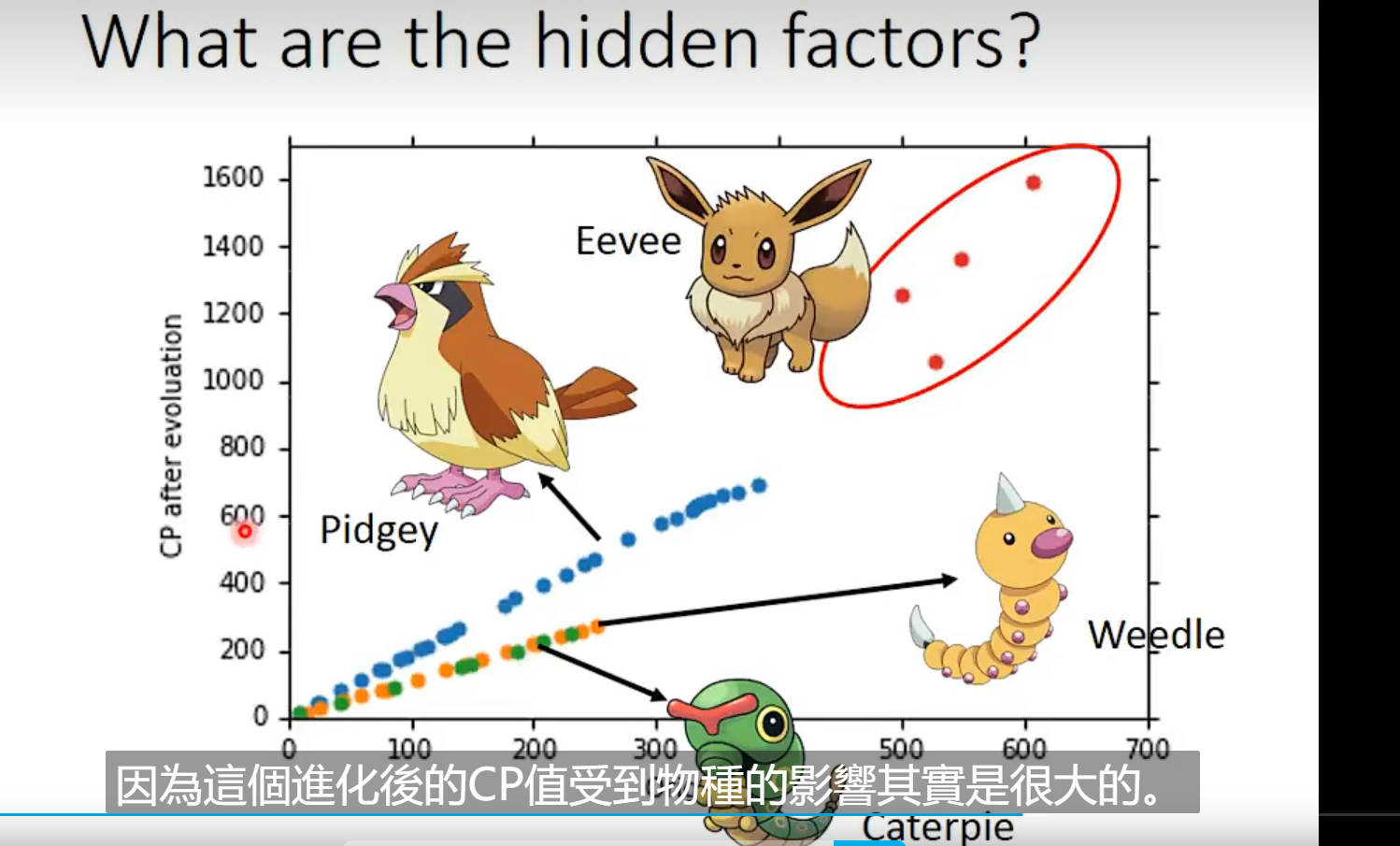


Over fitting：（过拟合）

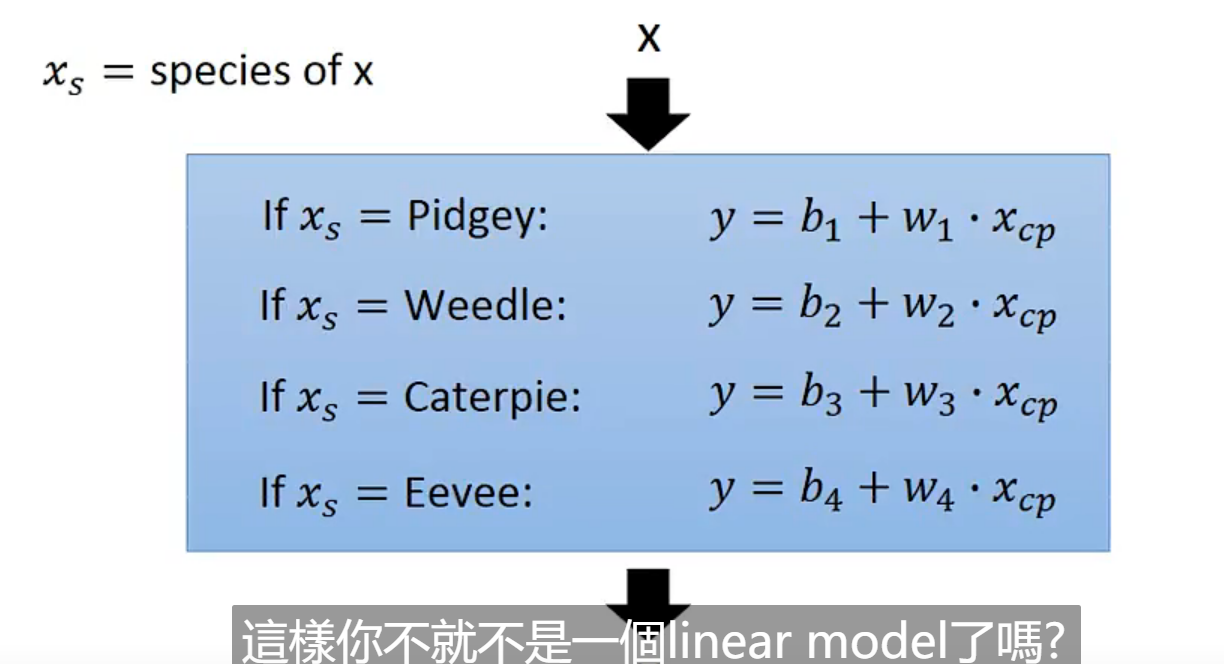
越复杂的model是能确保你在training date上的Loss越小，但不一定能保证你的testing date的Loss越小

4.

只考虑几个因素往往是不可取的

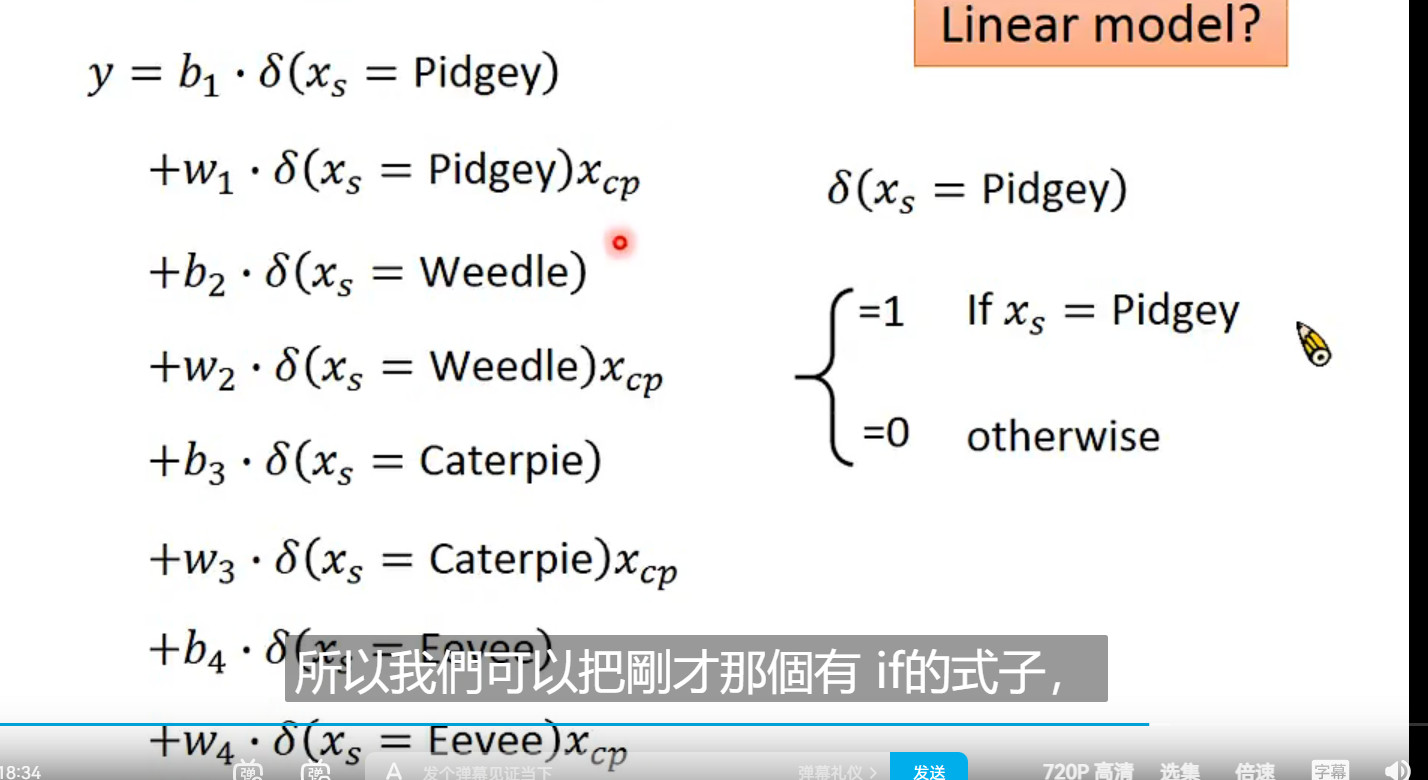


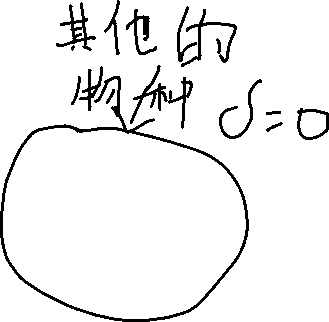
对于不同的物种我们需要来选择不同的Liner Model



但是我们可以把这些式子整合成一个式子







考虑所有的因素

