# 以宝可梦与数码宝贝分类器为例，解析机器学习中过拟合的原理

## 背景与目标：

因之前图片格式问题分类器失败，本次重新处理图片，旨在从理论解释 “参数越多越易过拟合” 的原因，预设听众具备机器学习基础。

## 分类依据与函数设计：

宝可梦与数码宝贝虽相似，但画风有别（数码宝贝线条复杂、多人形；宝可梦线条简单、圆润）。据此设计分类函数，通过边缘检测后白色像素点数量判断线条复杂程度，以参数 h 为阈值分类（超过 h 为数码宝贝，反之则为宝可梦），h 的所有可能取值集合为大 H，其大小代表模型复杂度。

## 损失函数：

基于数据集（含 x 即样本、y 即类别标注）计算，定义为错误率（预测错误记 1，正确记 0，取平均值）。

## 理想与现实训练：

理想是用全宇宙数据（Dio）找到最优参数 h0（使 Dio 上损失最小）；现实只能用训练集（Dtrain，从 Dio 抽样，满足独立同分布）找到 htrain，需关注 htrain 与 h0 在 Dio 上的损失差距。

## 训练集影响与理论推导：

好的训练集能让 htrain 在 Dio 上表现接近 h0，差的则差距大。理论上，若训练集对所有 h 在 Dtrain 和 Dio 上的损失差距小（有代表性），则理想与现实接近。

## 坏训练集概率：

抽样到坏训练集（导致理想与现实差距大）的概率上限，与模型复杂度（大 H 大小）成正相关，与训练资料数（n）成负相关（n 越大、模型越简单，概率越低）。

## 连续参数与模型选择：

连续参数模型可通过计算精度或 VC dimension 评估复杂度；模型选择存在两难 —— 需多训练数据和简单模型，但数据量常不可控，简单模型可能性能不足。