

模式识别：通过计算机用数学技术方法来研究模式的自动处理和判读。

机器学习：能通过经验自动改进的计算机算法；用数据或以往的经验，以此优化计算机程序的性能标准。

手机拍照或视频直播中的自动美颜是直播中最常见的功能之一，其主要原理是磨皮和美白达到美颜的效果，磨皮是通过去噪实现的，指的是去除图像噪点或模糊化处理等，这一环节需要涉及到人脸和皮肤检测技术，这一环节采用了模式识别和机器学习的方法。

无人驾驶汽车是广义上的智能汽车，利用车载传感系统感知道路环境，自动规划行车路线并控制车辆到达预定目标。无人车根据感知所获得的道路、车辆位置和障碍物信息，控制车辆的转向和速度，从而使车辆能够安全、可靠地在道路上行驶。有的无人驾驶汽车使用了计算机视觉技术，识别道路边缘从而达到自动驾驶、自动规划路线的效果，道路识别使用到了模式识别和机器学习的方法。

DeepCode是一种代码助手，可以自动生成代码。DeepCode从开源代码库中学习，并已建立了知识库，以提出有关改进代码的建议。每次更改都会分析代码。DeepCode对程序员的代码有识别能力，这一点体现了模式识别和机器学习的方法的使用。

DeepFake是一种新兴的人工智能技术，视频伪造是Deepfake技术最为主要的代表，其核心原理是利用生成对抗网络或者卷积神经网络等算法将目标对象的面部“嫁接”到被模仿对象上。由于视频是连续的图片组成，因此只需要把每一张图片中的脸替换，就能得到变脸的新视频。具体而言，首先将模仿对象的视频逐帧转化成大量图片，然后将目标模仿对象面部替换成目标对象面部。最后，将替换完成的图片重新合成为假视频，而深度学习技术可以使这一过程实现自动化。随着深度学习技术的发展，自动编码器、生成对抗网络等技术逐渐被应用到深度伪造中。对于人脸的识别需要使用到模式识别和机器学习技术。

疫情期间商城等场所使用的自动摄像测温设备的原理是摄像机抓拍照片，使用模式识别和机器学习技术识别人脸图片，从而实现自动摄像测温。

其他感兴趣的系统：无人机。无人机的传感器采集信息，进行模式识别，利用机器学习技术准确识别出道路信息和物体信息。

2.1

$$(1) \text{ Tom患病的可能性 (使用A) : } P_{Tom}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 1$$

$$\text{Bob患病的可能性是 (使用A) : } P_{Bob}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 0.202\%$$

$$\text{John患病的可能性是 (使用B) : } P_{John}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 47.6\%$$

$$\text{Don患病的可能性是 (使用B) : } P_{Don}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 0.1\%$$

(2) 代入新的数据可算得：

$$\text{Tom患病的可能性 (使用A) : } P_{Tom}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 1\%$$

$$\text{Bob患病的可能性是 (使用A) : } P_{Bob}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 0.02\%$$

John患病的可能性是（使用B）：

$$P_{John}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 8.3\%$$

Don患病的可能性是（使用B）：

$$P_{Don}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 0.01\%$$

(3) 代入新的数据可算得：

$$\text{Tom患病的可能性（使用A）： } P_{Tom}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 1\%$$

Bob患病的可能性是（使用A）：

$$P_{Bob}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 16.7\%$$

John患病的可能性是（使用B）：

$$P_{John}(D^+ | T^+) = \frac{P(T^+|D^+)P(D^+)}{P(T^+)} = \frac{SnP(D^+)}{Sn \cdot P(D^+) + (1-Sp) \cdot (1-P(D^+))} \cong 98.9\%$$

Don患病的可能性是（使用B）：

$$P_{Don}(D^+ | T^-) = \frac{P(T^-|D^+)P(D^+)}{P(T^-)} = \frac{(1-Sn)P(D^+)}{(1-Sn)P(D^+) + Sp \cdot (1-P(D^+))} \cong 9.17\%$$

(4) 直观上的数据不一定能真正反映其本身的意义，因此需要用专业的方法分析数据的意义。对于人数比较少的地区，其数据分析可能不具有统计意义，因此研究者需要谨慎判断；通常在判断阳性概率时，应该用上述公式而不是简单的使用准确率。

2.2

- 数学形式

D 维空间的空间向量：

$$\vec{x} = (x_1, x_2, \dots, x_d)$$

线性分类面：

$$g(x) = \sum_{i=1}^d x_i w_i + w_0 = 0$$

- 证明

根据示意图，法向量可以表示为：

$$\vec{w} = (w_1, w_2, \dots, w_d)$$

设向量 x 在平面上的投影为 P ，有 $\overrightarrow{OP} = \vec{x}_p$ ，距离为 r

可得：

$$\vec{x} = \vec{x}_p + r \frac{\vec{w}}{\|\vec{w}\|}$$

两边同乘 \vec{w}

$$\sum_{i=1}^d x_i w_i = \sum_{i=1}^d x_{pi} w_i + r ||w||$$

又平面的性质为：

$$\sum_{i=1}^d x_{pi} w_i + w_0 = 0$$

可得距离为

$$|r| = \frac{\left| \sum_{i=1}^d x_i w_i + w_0 \right|}{||w||}$$

其中

$$g(x) = \sum_{i=1}^d x_i w_i + w_0$$

对于原点则有

$$|r_0| = \frac{w_0}{||w||}$$