IoT期末(10-11)

Lecture 9 47

5 competitive forces竞争力量驱动力

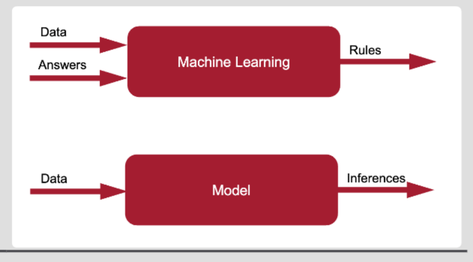
- The bargaining power of buyers — giving buyers a better understanding of true product performance.
- Rivalry among Competitors — higher fixed costs and lower variable costs
- The threat of new entrants
  - face significant new obstacles
  - connected products can also increase buyer loyalty and switching costs
- The threat of substitute products or services替代品
  - product-as-a-service
  - the shared-usage model — eg. shared bike systems
- The bargaining power of suppliers供应商的议价能力 — eg. the Open Automotive Alliance

Machine Learning (Definition And Applications)

- ML：Machine Learning is a subfield of Artificial Intelligence focused on developing algorithms that learn to solve problems by analysing data for patterns.
- DL：Deep Learning is a type of Machine Learning that leverages Neural Networks and Big Data

Lecture 10

Machine Learning Paradigm（机器学习范式）



Basic Concepts of Machine Learning（以Linear Regression为例子）

1. 随机初始化：初始猜测
2. 损失函数：衡量准确性（accuracy）
3. 优化：通过梯度下降最小化损失（minimize the loss by gradient descent） — 学习率（Learning rate）：学习时的步长（step size）
4. Epochs：重复次数

强化学习是一种机器学习方法，通过智能体与环境互动、尝试不同动作，并根据反馈(feedback)奖励(reward)或惩罚(penality)来优化(optimize)行为策略(behavior)，以最大化累积奖励。它广泛应用于自动驾驶、游戏AI和机器人控制等领域。

Reinforcement learning is a type of machine learning where an agent interacts with an environment, tries different actions, and optimizes its behavior based on feedback in the form of rewards or penalties. The goal is to maximize cumulative rewards, and it's widely used in applications like autonomous driving, game AI, and robotic control.

Ethical AI & Machine Learning

- Human agency & oversight(人力代理和监督) — AI系统应该赋予人类权利，让人类作出明智的决定促进人类的基本权利
- Robustness & Safety（技术稳健型和安全性）
- Privacy（隐私的数据治理）
- Transparency（透明度） — 数据、系统和AI商业模式应该是透明的
- Diversity & Fairness（多元化和公平） — AI系统应该向所有人开放，无论是否有残疾，并让利益相关者参与整个生命周期
- Societal & environmental wellbeing（社会和环境福祉） — 可持续的且环境友好的，利于人类和生物的
- Accountability（责任性） — 建立机制确保AI系统及结果的问责 Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes.

Lecture 11

Tiny Machine Learning

定义：是一个快速发展的机器学习技术和应用领域，包括算法、硬件和软件，能够以极低的功率（extremely low power）执行设备上(on-device)传感器数据分析(sensor data analytics)

Process

1. Step 1: Audio input from microphone (sensor)
2. Step 2: Process Input translation, then execute command
3. Step 3: Generate Output Play response through embedded speaker

为什么TinyML很重要

1. Tiny computers are ubiquitous
2. MCU(Microcontroller) demand forecast
3. MCUs are cheap
4. MCUs are Ultra-low power system
5. No good data left behind（没有留下任何好的数据） — 5 quintillion（5千万亿）Bytes data are produced every day by IoT, but less than 1% is used.

TinyML的挑战

- 硬件
  - Microprocessor和Microcontroller的区别
    - Microprocessor
      - 计算机系统的Heart
      - 内存和外存是external
      - 设计时很flexible
      - System size is big
    - Microcontroller
      - 嵌入式系统的Heart
      - 内存和外存是internal
      - 设计时不太flexible
      - System size is tiny
- 软件
  - 操作系统 — Windows, MacOS, Linux, iOS, Android
- 机器学习
  - Model Compression Techniques（模型压缩技术）
    - Pruning 修剪 — 消除贡献不大的权重或神经元来降低神经网络的复杂性
    - Quantization 量化 — 具体的量
    - Knowledge Distillation 知识蒸馏 — 将大型模型学到的知识转移到较小的模型中，从而实现更快的推理并减少内存需求，同时保持可比较的性能
    - 模型压缩工具
      - TensorFlow Lite
      - Edge Impulse
      - arm

模型评估：性能指标

Confusion Matrix Example

	Predicted Fall	Predicted No Fall
Actual Fall	True Positives (TP) = 50	False Negatives (FN) = 5
Actual No Fall	False Positives (FP) = 10	True Negatives (TN) = 35

Metric Calculations

1. **Precision**:
$$\text{Precision} = \frac{TP}{TP + FP} = \frac{50}{50 + 10} = \frac{50}{60} \approx 0.833 \text{ (or } 83.3\%)$$
2. **Recall (Sensitivity)**:
$$\text{Recall} = \frac{TP}{TP + FN} = \frac{50}{50 + 5} = \frac{50}{55} \approx 0.909 \text{ (or } 90.9\%)$$
3. **Accuracy**:
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{50 + 35}{50 + 35 + 10 + 5} = \frac{85}{100} = 0.85 \text{ (or } 85\%)$$
4. **F1 Score**:
$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.833 \times 0.909}{0.833 + 0.909} \approx 0.869 \text{ (or } 86.9\%)$$
5. **Sensitivity** (Same as Recall):
$$\text{Sensitivity} = \frac{TP}{TP + FN} = 0.909 \text{ (or } 90.9\%)$$
6. **Specificity**:
$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{35}{35 + 10} = \frac{35}{45} \approx 0.778 \text{ (or } 77.8\%)$$

Confusion Matrix

Terms

- Precision：ratio between true positives versus all positives
- Accuracy: The probability of a correct decision — **Accuracy = (TP+TN)/(TP+TN+FP+FN)**
- Recall：the measure of how accurate the model is inidentifying true positives
- Sensitivity (True Positive Rate)：is the probability of a positive test result, conditioned on the individual truly being positive — = Recall = TPR
- Specificity (True Negative Rate)：is the probability of a negative test result, conditioned on the individual truly being negative. — = FPR
- F1 Score：是精确率（Precision）和召回率（Recall）的调和平均数（harmonic mean），值介于 0 和 1 之间，值越高表示模型的性能越好 — $F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

Balanced Accuracy

Balanced Accuracy = $\frac{Sensitivity+Specificity}{2}$

在**不平衡数据集**上表现优于普通准确率，因为普通准确率在数据集严重不平衡时容易产生偏差

The optimal classifier is one with sensitivity (Probability of True Positive) as close to 100% as possible, and at the same time with specificity (Probability of True Negative) as close to 100% as possible

Receiver Operator Characteristics Curve(ROC)