ML

Machine Learning Cheat Sheet

1. What is Machine Learning?

- **Definition:** Field of study that gives computers the ability to learn without being explicitly programmed (Arthur Samuel, 1959).
- **Tom Mitchell's Definition:** A computer program learns from experience (E) with respect to some task (T) and some performance measure (P) if its performance on T, as measured by P, improves with experience E.
- **Examples:**
  - **T:** Playing checkers, **P:** Win percentage, **E:** Playing practice games.
  - **T:** Recognizing handwritten words, **P:** Classification accuracy, **E:** Labeled handwritten data.

2. Traditional vs. Machine Learning Approach

- **Traditional Approach:** Rule-based programming.
- **ML Approach:** Learns from data instead of being explicitly programmed.

3. When to Use Machine Learning?

- **No human expertise available** (e.g., navigating on Mars).
- **Humans cannot explain their expertise** (e.g., speech recognition).
- **Models must be customized** (e.g., personalized medicine).
- **Large-scale data is required** (e.g., genomics).

4. Types of Learning

- **Supervised Learning:** Labeled training data.
  - **Example:** Email spam classification.
- **Unsupervised Learning:** No labels, finds patterns.

    - **Example:** Customer segmentation.
- **Semi-supervised Learning:** Some labeled data, mostly unlabeled.
    - **Example:** Medical diagnosis.
- **Reinforcement Learning:** Learns by trial and error using rewards.
    - **Example:** Self-driving cars.

5. Learning Algorithm Tasks

- **Pattern Recognition:** Face recognition, handwriting recognition.
- **Pattern Generation:** Generating images, motion sequences.
- **Anomaly Detection:** Credit card fraud detection.
- **Prediction:** Stock price forecasting.

6. Sample Applications

- Web Search, Computational Biology, Finance, E-commerce, Robotics, Medical Imaging.

7. Autonomous Cars

- **Legalization:** First allowed in Nevada (2011), later in 4 US states (by 2013).
- **Key Technologies:** Path Planning, Adaptive Vision.

8. ML System Classification

- **Batch Learning:** Trained once, used without updates.
- **Online Learning:** Continuously learns and updates.
- **Instance-Based Learning:** Memorizes examples and compares new data.
- **Model-Based Learning:** Builds a model from training data.

9. Challenges in Machine Learning

- **Insufficient Data:** More data needed for effective learning.
- **Non-representative Data:** Bias in sampling.

- **Poor Data Quality:** Noisy or irrelevant features.
- **Overfitting:** Model learns noise, not patterns.
- **Underfitting:** Model too simple to capture trends.
- **Hyperparameter Tuning:** Finding the best configuration.

50-Question Quiz

1. What is the definition of machine learning according to Arthur Samuel?
2. How does Tom Mitchell define machine learning?
3. What is an example of a task (T) in machine learning?
4. How is performance (P) measured in ML?
5. What is experience (E) in ML?
6. How does traditional programming differ from ML?
7. When is ML useful?
8. What is supervised learning?
9. Give an example of supervised learning.
10. What is unsupervised learning?
11. Provide an example of unsupervised learning.
12. What is semi-supervised learning?
13. Give an example of semi-supervised learning.
14. What is reinforcement learning?
15. Give an example of reinforcement learning.
16. What are some ML tasks?
17. What is pattern recognition?
18. What is anomaly detection?
19. What is an example of prediction in ML?
20. List three ML applications.

21. Name a use case of ML in robotics.
22. What was the first state to legalize autonomous cars?
23. What are key technologies used in autonomous cars?
24. What is batch learning?
25. What is online learning?
26. What is instance-based learning?
27. What is model-based learning?
28. Why is insufficient data a problem in ML?
29. What is non-representative data?
30. Why is poor data quality an issue?
31. What is overfitting?
32. How does underfitting affect a model?
33. Why is hyperparameter tuning important?
34. What is the difference between batch and online learning?
35. How does reinforcement learning differ from supervised learning?
36. What is the role of data in ML?
37. How can ML help humans learn?
38. What is feature selection?
39. How does ML handle new data?
40. Why is model evaluation important?
41. What is cross-validation?
42. What are ML pipelines?
43. How is ML used in speech recognition?
44. What is deep learning?
45. How does ML impact medical imaging?
46. What is ChatGPT’s learning approach?

47. What is classification in ML?
48. Give an example of a classification algorithm.
49. What is clustering?
50. Name a challenge in ML and how to address it.

50-Question Quiz with Answers

1. **What is the definition of machine learning according to Arthur Samuel?**
    - The ability of computers to learn without being explicitly programmed.
2. **How does Tom Mitchell define machine learning?**
    - Learning occurs if a system improves performance on a task with experience.
3. **What is an example of a task (T) in machine learning?**
    - Recognizing handwritten digits.
4. **How is performance (P) measured in ML?**
    - Accuracy, error rate, or other evaluation metrics.
5. **What is experience (E) in ML?**
    - The dataset used for training the model.
6. **How does traditional programming differ from ML?**
    - Traditional programming follows fixed rules, ML learns from data.
7. **When is ML useful?**
    - When human expertise is unavailable or complex problems arise.
8. **What is supervised learning?**
    - Learning from labeled data.
9. **Give an example of supervised learning.**
    - Spam email classification.

10. **What is unsupervised learning?**
    - Learning from unlabeled data to find patterns.
11. **Provide an example of unsupervised learning.**
    - Customer segmentation.
12. **What is semi-supervised learning?**
    - Learning from a mix of labeled and unlabeled data.
13. **Give an example of semi-supervised learning.**
    - Medical diagnosis with limited labeled data.
14. **What is reinforcement learning?**
    - Learning from actions through rewards and penalties.
15. **Give an example of reinforcement learning.**
    - Training a self-driving car.
16. **What are some ML tasks?**
    - Classification, regression, clustering, anomaly detection.
17. **What is pattern recognition?**
    - Identifying recurring data patterns.
18. **What is anomaly detection?**
    - Detecting unusual patterns, e.g., fraud detection.
19. **What is an example of prediction in ML?**
    - Stock price forecasting.
20. **List three ML applications.**
    - Web search, robotics, finance.
21. **Name a use case of ML in robotics.**
    - Autonomous navigation.
22. **What was the first state to legalize autonomous cars?**
    - Nevada (2011).

23. **What are key technologies used in autonomous cars?**
    - Path Planning, Adaptive Vision.
24. **What is batch learning?**
    - Training a model once and using it without updates.
25. **What is online learning?**
    - Continuous model training with incoming data.
26. **What is instance-based learning?**
    - Memorizing past cases and comparing new ones.
27. **What is model-based learning?**
    - Creating a predictive model from training data.
28. **Why is insufficient data a problem in ML?**
    - It limits the model's ability to learn patterns.
29. **What is non-representative data?**
    - Data that does not reflect real-world scenarios.
30. **Why is poor data quality an issue?**
    - Noisy or irrelevant features reduce accuracy.
31. **What is overfitting?**
    - The model learns noise instead of patterns.
32. **How does underfitting affect a model?**
    - The model is too simple and fails to learn trends.
33. **Why is hyperparameter tuning important?**
    - It optimizes model performance.
34. **What is the difference between batch and online learning?**
    - Batch learning is static, online learning updates continuously.
35. **How does reinforcement learning differ from supervised learning?**
    - Reinforcement learning learns through trial and error.

36. **What is the role of data in ML?**
    - It is the foundation for training models.
37. **How can ML help humans learn?**
    - By identifying patterns and automating complex tasks.
38. **What is feature selection?**
    - Choosing relevant input variables for a model.
39. **How does ML handle new data?**
    - Through continuous learning and adaptation.
40. **Why is model evaluation important?**
    - To assess performance and improve accuracy.