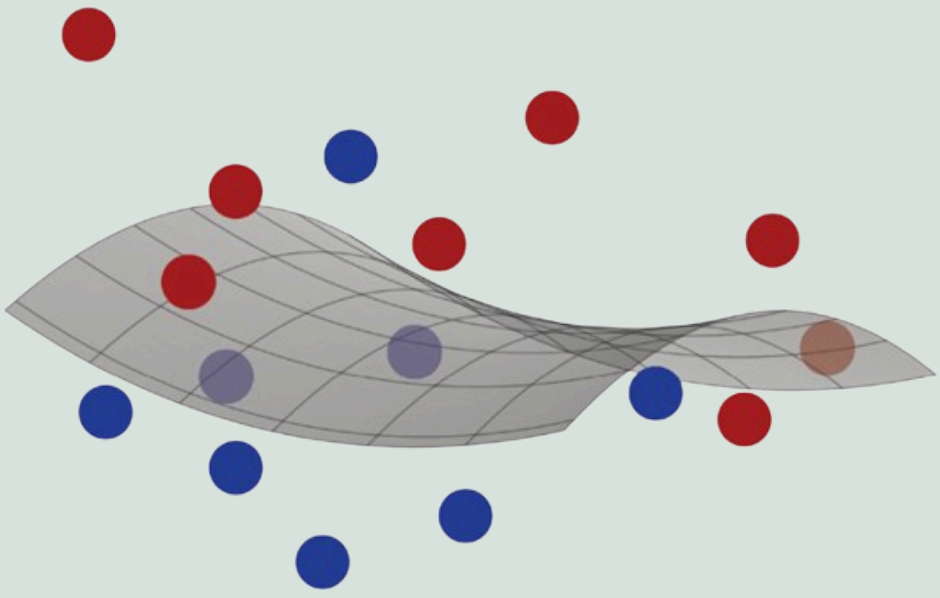
Foundations of Machine Learning

DAY - 3

Some Standard Learning Tasks



Some Standard Learning Tasks

Created By: **Birva Dave**
medium.com/@birva1809
github.com/Birva1809
linkedin.com/in/birva-dave
birvadave1809@gmail.com

Machine learning deals with various types of tasks — some of which are extensively studied and widely used in both research and real-world applications. These are standard problems ML researchers and practitioners often work with:

- 🗂️ **Classification**
  - Classification is about assigning a label or category to each item.
  - For example:
    - In document classification, the task is to assign a topic like “politics,” “sports,” or “business” to each document.
    - In image classification, the goal might be to label an image as a “car,” “train,” or “airplane.”
  - Usually, the number of categories is limited to a few hundred. But in more complex tasks like Optical Character Recognition (OCR), speech recognition, or text classification, the number of possible categories can be very large — or even unlimited.
  - The key idea: the model learns from labeled examples and predicts the correct category for new, unseen items.
- 📈 **Regression**
  - Regression focuses on predicting a real-valued number rather than a category.
  - Examples include:
    - Predicting the price of a stock tomorrow
    - Estimating economic growth rates or house prices
  - In regression, the penalty for a wrong prediction isn’t just “right or wrong” — it depends on how far off the prediction is from the actual value. This differs from classification, where the goal is simply to match the correct label.
  - It’s ideal for any task where the output is continuous, not categorical.

Some Standard Learning Tasks

Created By: **Birva Dave**
medium.com/@birva1809
github.com/Birva1809
linkedin.com/in/birva-dave
birvadave1809@gmail.com

- 🧭 **Ranking**
  - Ranking is about ordering items based on a specific criterion.
  - The most familiar example is web search — when you Google something, it ranks the most relevant pages at the top.
  - Ranking is also important in recommendation systems and NLP, where outputs are prioritized based on relevance, confidence, or user behavior.
  - The goal isn't just to predict something, but to learn the correct order of relevance or importance.
- 🧱 **Clustering**
  - Clustering is about grouping similar items together — even when no labels are provided.
  - It's an unsupervised learning task, meaning the model must discover patterns without being told what the groups should look like.
  - For example, in social network analysis, clustering can identify communities — people who interact more frequently with each other.
  - It's widely used in exploratory data analysis, market segmentation, and organizing large datasets.
- 🔻 **Dimensionality Reduction / Manifold Learning**
  - This reduces the number of features while preserving key structure.
  - Like shrinking a high-res image without losing important details.
  - Common in computer vision, it helps:
    - Speed up computation
    - Remove noise
    - Improve performance
    - Make data easier to visualize
  - It's ideal when working with high-dimensional datasets.

Practical Objectives of Machine Learning

Created By: **Birva Dave**
medium.com/@birva1809
github.com/Birva1809
linkedin.com/in/birva-dave
birvadave1809@gmail.com

The two main goals of machine learning are:

- Making accurate predictions on new, unseen data
- Designing algorithms that are efficient, robust, and scalable, even when the dataset is huge

To reach these goals, ML brings up a bunch of theoretical and practical questions like:

- Which types of problems (or concept families) can actually be learned using machines?
- Under what conditions is learning even possible?
- How effectively (in terms of time, memory, and data) can we learn those concepts?

These questions form the foundation of ML theory, and finding answers helps us design better algorithms for real-world problems.