Pattern Recognition & Machine Learning

Preface

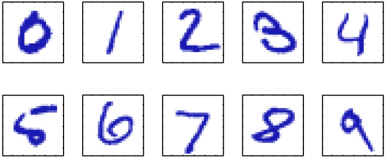
Prof. Reza Azadeh

University of Massachusetts Lowell

Pattern recognition: the automatic discovery of regularities in data through the use of computer algorithms. Also, with the use of these regularities, taking actions such as classifying the data into different categories.

Johannes Kepler discovered the empirical laws of planetary motion, based on the astronomical observations from the 16th century (Image generated by DALL.E).

Example: Recognizing handwritten digits



Each digit is a 28×28 pixel image, represented by a vector $\mathbf{x} \in \mathbb{R}^{784}$.

Goal: build a machine that takes a vector $\mathbf{x}$ and input and outputs the identity of the digit as $0, \dots, 9$.

Example: Recognizing handwritten digits

Machine learning: use a *training set*, a large set of N digits $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, together with a *target vector*, $\mathbf{t}$, for each image, including the identity of the samples.
Our model:

$$\mathbf{y}(\mathbf{x})$$

We then use a *test set* to check the precision and *generalization* of the learned model.

Pre-processing

Pre-processing: transforming the raw data into some new space of variables where, hopefully, the pattern recognition problem is easier to solve. This is also called *feature extraction*.

For instance, each image is translated and scaled so that each digit is contained within a box of a fixed size.

Supervised Learning

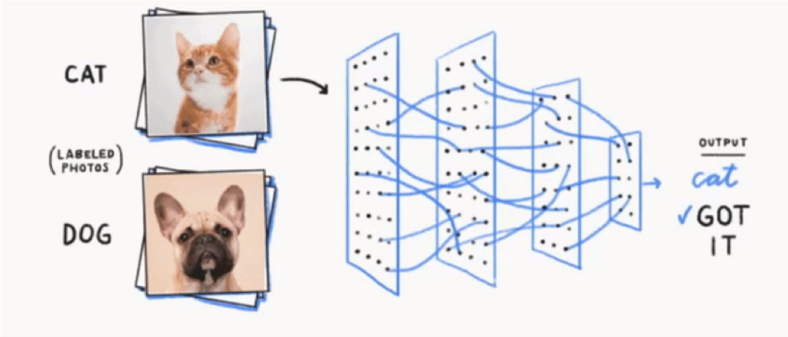
Applications in which the training data comprises examples of the input vectors along with their corresponding target vectors are known as *supervised learning* problems.

- **Classification:** assigning each input vector to one of the finite number of discrete categories.
- **Regression:** when the output consists of one or more continuous variables.

Supervised Learning - Example

Sample of cats & dogs images from Kaggle Dataset

Supervised Learning - Example

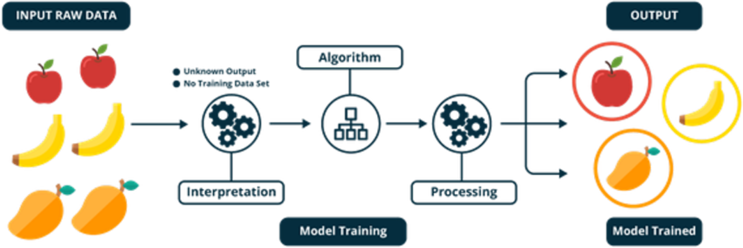


Unsupervised Learning

When the training data consists of a set of input vectors $\mathbf{x}$ without any corresponding target values, the problem is called *unsupervised learning*.

- **Clustering:** discovering groups of similar examples within data.
- **Density estimation:** determine the distribution of data within the input space.

Unsupervised Learning - Example



Machine Learning Applications - Healthcare

- Disease diagnosis (e.g., cancer detection)
- Personalized medicine
- Drug discovery and development
- Medical image analysis
- Predictive analytics for patient care
- Health monitoring through wearables
- Virtual health assistants

Machine Learning Applications - Finance

- Fraud detection
- Algorithmic trading
- Risk assessment
- Credit scoring
- Portfolio management
- Customer segmentation
- Personalized financial advice

Machine Learning Applications - Retail and E-commerce

- Product recommendation systems
- Inventory management
- Dynamic pricing
- Customer sentiment analysis
- Supply chain optimization
- Virtual shopping assistants
- Targeted marketing and advertising

Machine Learning Applications - Transportation

- Autonomous vehicles (self-driving cars)
- Traffic prediction and optimization
- Ride-sharing services (e.g., Uber, Lyft)
- Route planning and navigation
- Predictive maintenance for vehicles
- Airline pricing optimization

Machine Learning Applications - Manufacturing

- Predictive maintenance for machinery
- Quality control and defect detection
- Supply chain optimization
- Production planning and scheduling
- Robotics and automation
- Energy efficiency optimization

Machine Learning Applications - Entertainment and Media

- Content recommendation (e.g., Netflix, YouTube)
- Automated content generation
- Video and audio recognition
- Sentiment analysis for media reviews
- Personalized marketing
- Fake news detection

Machine Learning Applications - Education

- Personalized learning systems
- Automated grading and feedback
- Language translation and learning tools
- Intelligent tutoring systems
- Dropout prediction
- Administrative process automation

Machine Learning Applications - Energy

- Energy demand forecasting
- Renewable energy management
- Smart grid optimization
- Predictive maintenance of energy systems
- Energy efficiency monitoring

Machine Learning Applications - Agriculture

- Crop monitoring and yield prediction
- Precision farming
- Pest and disease detection
- Soil quality analysis
- Weather prediction for farming
- Autonomous agricultural equipment

Machine Learning Applications - Security and Surveillance

- Facial recognition systems
- Intrusion detection
- Cybersecurity threat analysis
- Biometric authentication
- Fraud detection
- Video analytics for surveillance

Machine Learning Applications - Natural Language Processing (NLP)

- Chatbots and virtual assistants
- Language translation
- Sentiment analysis
- Text summarization
- Document classification
- Speech recognition

Machine Learning Applications - Robotics

- Autonomous navigation
- Human-robot interaction
- Robotic process automation (RPA)
- Industrial robotics
- Robotic surgery

Machine Learning Applications - Environmental Science

- Climate modeling and prediction
- Wildlife conservation
- Disaster prediction and management
- Waste management
- Renewable energy optimization

Machine Learning Applications - Gaming

- AI opponents in games
- Real-time decision-making in gameplay
- Procedural content generation
- Player behavior analysis

Machine Learning Applications - Social Media

- Content moderation
- Trend prediction
- Fake profile detection
- Engagement prediction
- Sentiment analysis of posts

Next

In the first chapter of the textbook, we

- introduce some of these problems using simple examples,
- give an introduction to three important tools that will be used throughout the course, namely
  - probability theory,
  - decision theory, and
  - information theory.

Mathematical Notation (1)

- Vectors are denoted by lower case bold Roman letter such as $\mathbf{x}$, and all vectors assumed to be column vectors.
- A superscript $\top$ denotes the transpose of a matrix or vector, so that $\mathbf{x}^\top$ will be a row vector. Note that the textbook uses T instead.
- Upper case bold Roman letters, such as $\mathbf{M}$, denote matrices.
- The notation $(w_1, \ldots, w_M)$ denotes a row vector with M elements, while the corresponding column vector us written as $\mathbf{w} = (w_1, \ldots, w_M)^\top$.
- the notation $[a, b]$ is used to denote the closed interval from a to b, that is the interval including the values a and b themselves, while (a, b) denotes the corresponding open interval.

Mathematical Notation (2)

- The $M \times M$ identity matrix is denoted $\mathbf{I}_M$, which will be abbreviated to $\mathbf{I}$. It has elements I_{ij} that equal 1 if $i = j$ and 0 otherwise.
- a functional is denoted $f[y]$ where $y(x)$ is some function.
- The notation $g(x) = O(f(x))$ denotes that $|f(x)/g(x)|$ is bounded as $x \to \infty$. For instance, if $g(x) = 3x^2 + 2$, then $g(x) = O(x^2)$.
- The expectation of a function $f(x, y)$ with respect to a random variable x is denoted by $\mathbb{E}_x[f(x, y)]$. In situations where there is no ambiguity as to which variable is being averaged over, this will be simplified by omitting the suffix.

Mathematical Notation (3)

- The variance is denoted $var[f(x)]$, and for vector variables the covariance is written $cov[\mathbf{x}, \mathbf{y}]$. We shall use $cov[\mathbf{x}]$ as a shorthand notation for $cov[\mathbf{x}, \mathbf{x}]$.
- if we have N values $\mathbf{x}_1, \ldots, \mathbf{x}_N$ of a D-dimensional vector $\mathbf{x} = (x_1, \ldots, x_D)^\top$, we can combine the observations into a data matrix $\mathbf{X}$ in which the n^{th} row of $\mathbf{X}$ corresponds to the row vector $\mathbf{x}_n^\top$. Thus the n, i element of $\mathbf{X}$ corresponds to the i^{th} element of the n^{th} observation $\mathbf{x}_n$.

$$\mathbf{X} = \begin{pmatrix} | & \ldots & | \\ \mathbf{x}_1 & \ldots & \mathbf{x}_N \\ | & \ldots & | \end{pmatrix} \in \mathbb{R}^{D \times N}$$

For next session

1. Read pp. 1-12 from the textbook.
2. Read the mathematical notation (page xi of the book, or from this lecture notes).
3. Complete checkpoint 1 - instructions in Blackboard.