INF721 - Deep Learning (2024/2)
L2: Machine Learning

Prof. Lucas N. Ferreira
Universidade Federal de Viçosa

1 Introduction

This lecture provides an introduction to machine learning, which is a subfield of artificial intelligence (AI). Machine learning is a key component of deep learning and forms the foundation for understanding neural networks.

1.1 Definitions

- **Artificial Intelligence (AI):** The field of computer science focused on creating systems that can perform tasks requiring human intelligence.
- **Machine Learning:** A subset of AI that focuses on developing algorithms and models that allow computers to learn from and make predictions or decisions based on data.
- **Deep Learning:** A subset of machine learning based on artificial neural networks with multiple layers.

1.2 Brief History of Machine Learning and AI

- 1940s: McCulloch and Pitts design the first artificial neurons (not learned)
- 1950s: Rosenblatt develops the perceptron, capable of learning linear problems
- 1960s: Minsky shows limitations of single-layer perceptrons (e.g., XOR problem)
- 1980s: Multi-layer perceptrons and backpropagation algorithm developed
- 2000s: Support Vector Machines (SVMs) gain popularity
- 2010s onwards: Deep learning breakthroughs and rapid advancements

2 Types of Machine Learning

Machine learning can be broadly categorized into three main types based on the nature of the learning process and the data available:

2.1 Supervised Learning

In supervised learning, the algorithm learns from labeled data. The dataset contains both input features and corresponding target labels.

Formal definition: Given a dataset $D = \{(x^{(i)}, y^{(i)})\}_{i=1}^{m}$ where:

- $x^{(i)} \in \mathbb{R}^d$ is the feature vector of the i-th example
- $y^{(i)}$ is the label or target value of the i-th example
- m is the number of examples in the dataset
- d is the dimensionality of the feature vector

The goal is to learn a function $h : \mathbb{R}^d \to C$ that maps input features to output labels, where C is the set of possible labels.

2.1.1 Types of Supervised Learning Problems

1. **Classification:** The target variable is categorical (discrete classes).
   - Binary Classification: Two possible classes (e.g., spam detection)
   - Multi-class Classification: More than two classes (e.g., handwritten digit recognition)
2. **Regression:** The target variable is continuous (e.g., house price prediction).

Examples:

1. Spam Detection (Binary Classification):
   - Input: Email content (text)
   - Output: Spam (1) or Not Spam (0)
   - Feature representation: Word frequency counts
2. Handwritten Digit Recognition (Multi-class Classification):
   - Input: Image of a handwritten digit
   - Output: Digit class (0-9)
   - Feature representation: Pixel values
3. House Price Prediction (Regression):
   - Input: House features (size, location, etc.)
   - Output: Predicted price (continuous value)
   - Feature representation: Tabular data

2.2 Unsupervised Learning

In unsupervised learning, the algorithm learns from unlabeled data. The dataset contains only input features without corresponding target labels.

Formal definition: Given a dataset $D = \{x^{(i)}\}_{i=1}^{m}$ where:

- $x^{(i)} \in \mathbb{R}^d$ is the feature vector of the i-th example
- m is the number of examples in the dataset
- d is the dimensionality of the feature vector

The goal is to find patterns, structures, or relationships in the data without explicit labels.

2.2.1 Types of Unsupervised Learning Problems

1. **Clustering:** Group similar data points together.
2. **Dimensionality Reduction:** Reduce the number of features while preserving important information.
3. **Generative AI:** Learn to generate new data similar to the training data.

Examples:

1. Customer Segmentation (Clustering):
   - Input: Customer behavior data
   - Output: Groups of similar customers
2. Feature Compression (Dimensionality Reduction):
   - Input: High-dimensional data
   - Output: Lower-dimensional representation
3. Image Generation (Generative AI):
   - Input: Large dataset of images
   - Output: New, synthetic images

2.3 Reinforcement Learning

Reinforcement learning involves an agent learning to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties.

Formal definition:

- Agent: The learning algorithm
- Environment: The world in which the agent operates

- State (s): The current situation of the agent in the environment
- Action (a): A decision made by the agent
- Reward (r): Feedback from the environment
- Policy (π): A function that maps states to actions

The goal is to learn a policy $\pi(s)$ that maximizes the expected sum of rewards over time.

Example:

- Game-playing AI (e.g., AlphaGo)
- Robotic control systems

3 Data Types in Machine Learning

Understanding the types of data used in machine learning is crucial for selecting appropriate algorithms and preprocessing techniques.

3.1 Structured Data

Structured data is organized in a predefined format, typically in tables with rows and columns. Each column represents a specific attribute or feature.

Examples:

- Relational databases
- Spreadsheets
- CSV files

3.2 Unstructured Data

Unstructured data lacks a predefined format or structure. It requires more complex preprocessing and feature extraction techniques.

Examples:

- Text documents
- Images
- Audio files
- Video files

4 Hypothesis Space and Loss Functions

4.1 Hypothesis Space

The hypothesis space $\mathcal{H}$ is the set of all possible functions that a learning algorithm can consider as potential solutions.

Examples of hypothesis spaces:

1. Linear functions: $\mathcal{H} = \{h(x) = w_1 x + w_0 | w_1, w_0 \in \mathbb{R}\}$
2. Sinusoidal functions: $\mathcal{H} = \{h(x) = A \sin(Bx + C) | A, B, C \in \mathbb{R}\}$
3. Polynomial functions: $\mathcal{H} = \{h(x) = \sum_{i=0}^{n} a_i x^i | a_i \in \mathbb{R}, n \in \mathbb{N}\}$

4.2 Loss Functions

A loss function $L(h, D)$ measures how well a hypothesis h performs on a dataset D. It quantifies the difference between predicted and actual values.

Properties of loss functions:

- Measures the discrepancy between predictions and true labels
- Always non-negative
- Lower values indicate better performance

Common loss functions:

1. 0-1 Loss (for classification):

$$L_{0-1}(h, D) = \frac{1}{m} \sum_{i=1}^{m} \mathbb{1}[h(x^{(i)}) \neq y^{(i)}]$$

   where $\mathbb{1}[\cdot]$ is the indicator function.

2. Mean Squared Error (for regression):

$$L_{MSE}(h, D) = \frac{1}{m} \sum_{i=1}^{m} (h(x^{(i)}) - y^{(i)})^2$$

3. Mean Absolute Error (for regression):

$$L_{MAE}(h, D) = \frac{1}{m} \sum_{i=1}^{m} |h(x^{(i)}) - y^{(i)}|$$

5 Evaluating Model Performance

Proper evaluation of machine learning models is crucial to assess their generalization ability and avoid overfitting.

5.1 Train-Validation-Test Split

To evaluate model performance, the dataset is typically divided into three subsets:

- Training set (D_{train}): Used to train the model
- Validation set (D_{val}): Used to tune hyperparameters and select the best model
- Test set (D_{test}): Used to evaluate the final model performance

These subsets should be mutually exclusive and come from the same distribution as the original dataset.

5.2 Overfitting and Underfitting

- **Underfitting:** The model has high error on both training and validation sets.
- **Overfitting:** The model has low error on the training set but high error on the validation set.
- **Good fit:** The model has low error on both training and validation sets, and generalizes well to the test set.

6 Conclusion

This introduction to machine learning covers the fundamental concepts, types of learning problems, and evaluation techniques. Understanding these basics is crucial for delving deeper into deep learning and neural networks. In the following lectures, we will explore specific algorithms, starting with linear regression, and gradually build up to more complex neural network architectures.