بسم الله الرحمن الرحیم

علوم شناختی

جلسه ۲۲ (الف)

سیستم‌های خبره و یادگیری ماشینی

Expert systems and Machine Learning

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/cogsci

PART 3:
APPLICATIONS

CAMBRIDGE

Chapter 12:
Machine Learning: From Expert Systems to Deep Learning

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE

Chapter 12.1: Expert systems and machine learning

CAMBRIDGE UNIVERSITY PRESS

Expert systems

Designed to reproduce the performance of human experts

- E.g., MYCIN
  - Collected medical information on patients and produced a diagnosis
  - 69% accurate
  - More accurate than human experts

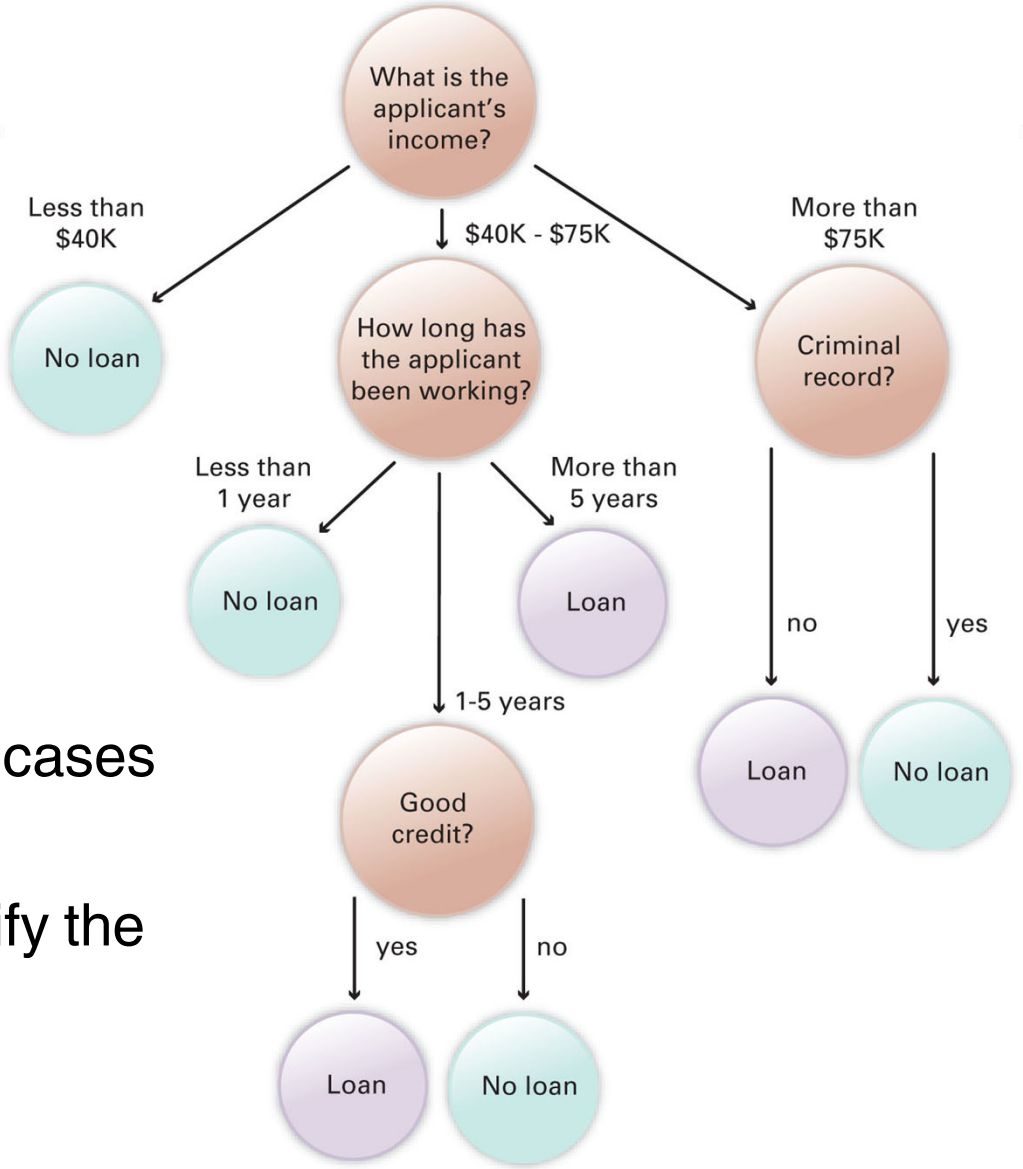
Decision trees

- Many expert systems use decision trees
  - Fixed decision procedure
  - Operate according to basic IF… THEN… rules

Decision trees

Decision tree for loan applicants

- Needs to apply to novel cases
- Uses heuristics to simplify the problem space



Heuristic search

- Problems are solved by generating and modifying symbol structures until a solution structure is reached
- Problem-spaces are generally too large to be searched exhaustively (brute force algorithms)
- Search must be selective ⇒ heuristic search rules

ID3 and PSSH

- ID3 is an illustration of the heuristic search hypothesis
- ID3 transforms one highly complex symbol structure (database) into another (a decision tree)
- The decision tree itself works by transforming symbol structures according to rules
    - Rules are the IF. . . THEN . . . rules built into the decision tree

ID3 as heuristic search

- ID3 is a tool for navigating through the search space of decision trees

- The algorithm that it uses sorts through possible decision trees using measures of information gain and entropy reduction

Basic concepts

- **Entropy** – level of uncertainty (relative to a target attribute)

- **Information gain** – reduction of entropy (relative to the target attribute)

The algorithm compares different classifying attributes in order to determine which has the highest information gain

An example

- We want to decide whether or not the weather is suitable for playing tennis.
- We collect information for two weeks. For each day we log the principal meteorological data and note whether or not we decide to play tennis on that day.
- We want to use this information to construct a decision tree that we can use in the future.

Attributes

- The target attribute: *Play Tennis?*

- The other attributes with the values they can take:

Outlook?	{sunny, overcast, rain}
Temperature?	{hot, mild, cool}
Humidity?	{high, low, normal}
Wind?	{weak, strong}

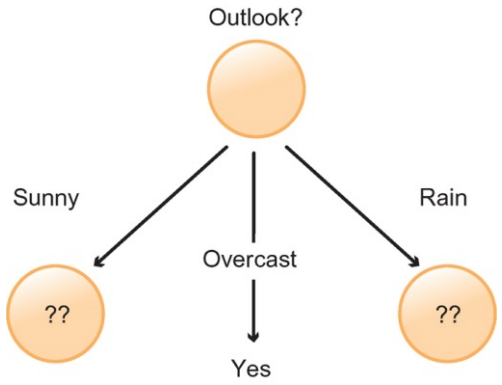
Database

Day	Outlook?	Temperature?	Humidity?	Wind?	Play Tennis?
D1	Sunny	Hot	High	Weak	**No**
D2	Sunny	Hot	High	Strong	**No**
D3	Overcast	Hot	High	Weak	**Yes**
D4	Rain	Mild	High	Weak	**Yes**
D5	Rain	Cool	Normal	Weak	**Yes**
D6	Rain	Cool	Normal	Strong	**No**
D7	Overcast	Cool	Normal	Strong	**Yes**
D8	Sunny	Mild	High	Weak	**No**
D9	Sunny	Cool	Normal	Weak	**Yes**
D10	Rain	Mild	Normal	Weak	**Yes**
D11	Sunny	Mild	Normal	Strong	**Yes**
D12	Overcast	Mild	High	Strong	**Yes**
D13	Overcast	Hot	Normal	Weak	**Yes**
D14	**Rain**	**Mild**	**High**	**Strong**	**No**

Correlations between the target attribute and the other attributes

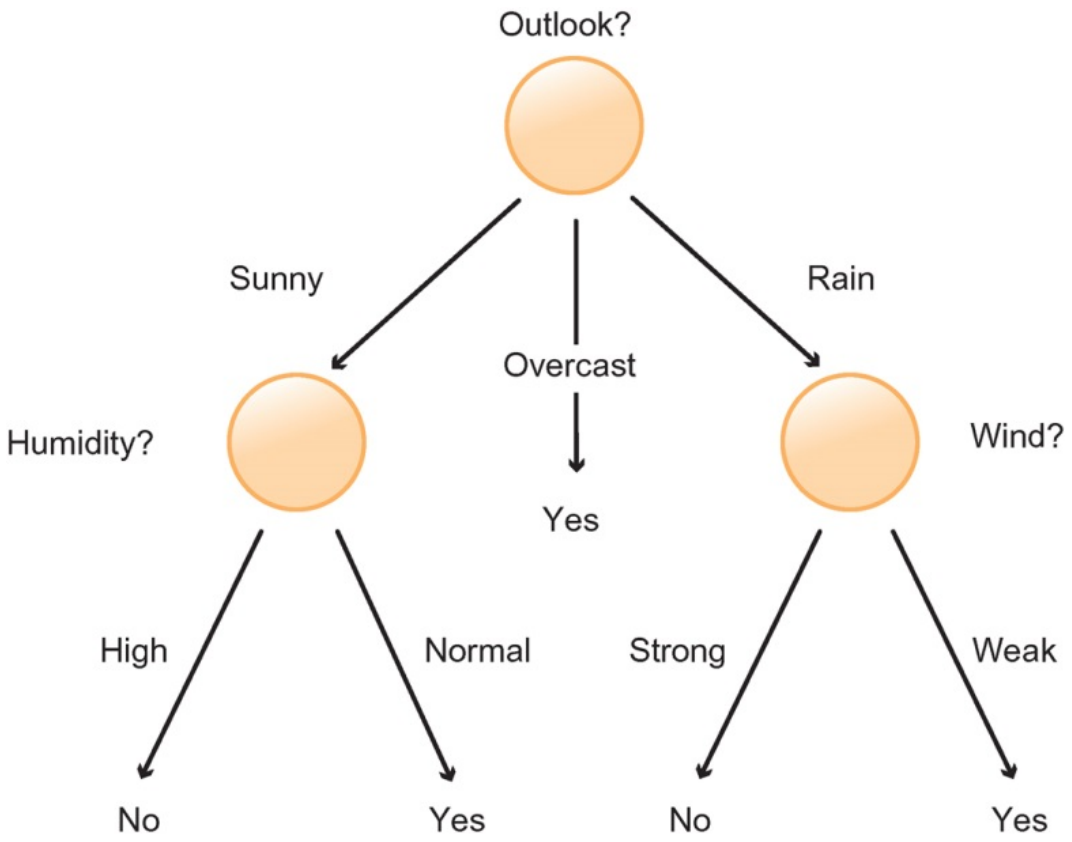
- **The first step:** find the attribute with the highest information gain.

Gain (*S, Outlook?)*	=	0.246
Gain (*S, Temperature?)*	=	0.029
Gain (*S, Humidity?)*	=	0.151
Gain (*S, Wind?)*	=	0.048

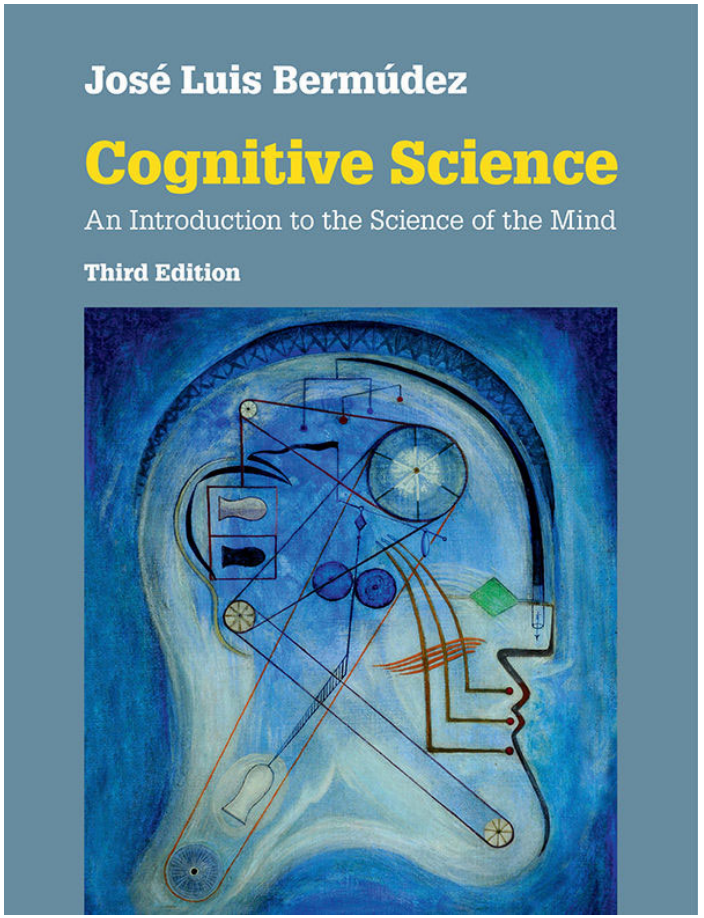


- The information gain is highest for *Outlook?.* That is the attribute it assigns to the first node in the decision tree.

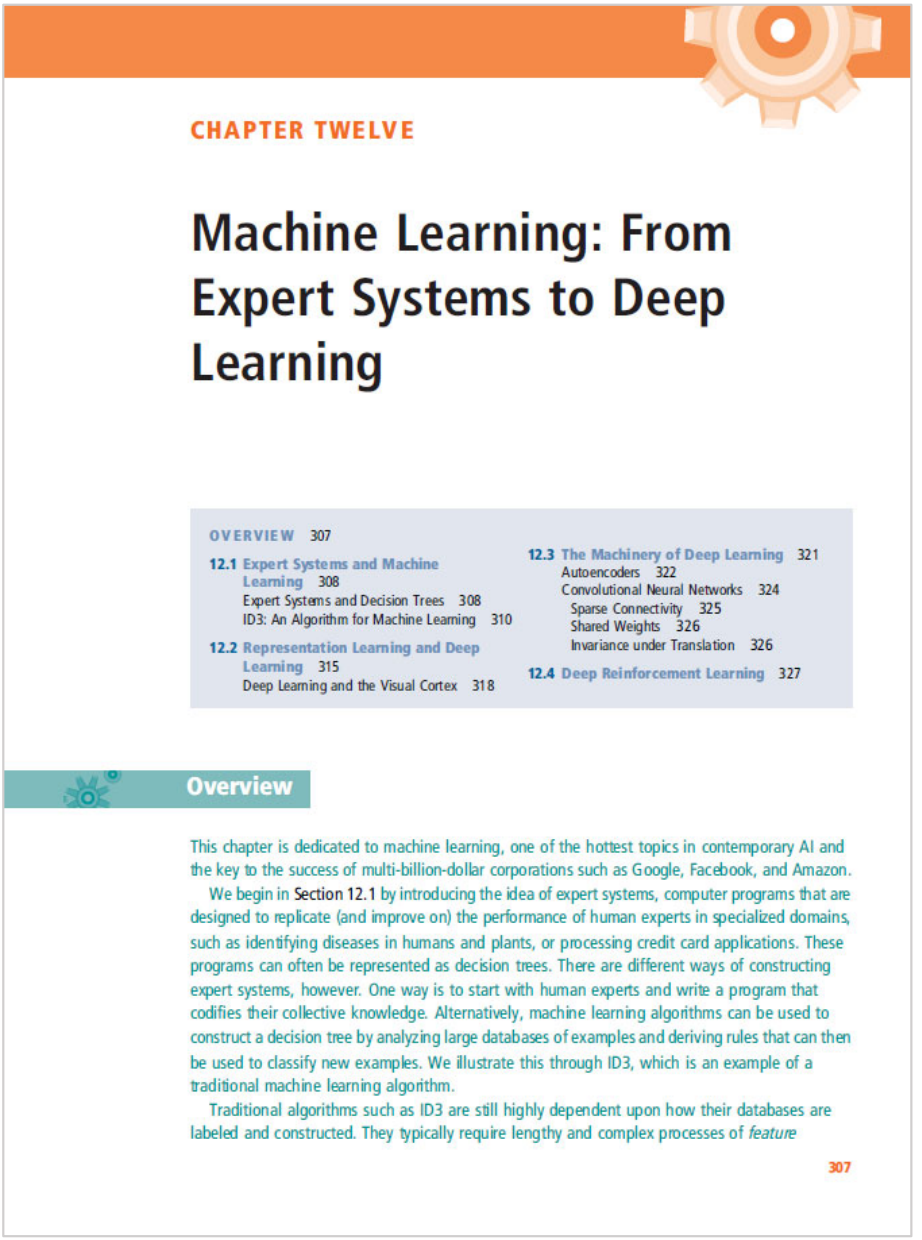
The decision tree



منبع اصلی



José Luis Bermúdez,
Cognitive Science:
An Introduction to the Science of the Mind,
3rd ed., Cambridge University Press, 2020.
Chapter 12 (Section 12.1)



CHAPTER TWELVE

Machine Learning: From Expert Systems to Deep Learning

Overview

This chapter is dedicated to machine learning, one of the hottest topics in contemporary AI and the key to the success of multi-billion-dollar corporations such as Google, Facebook, and Amazon.

We begin in Section 12.1 by introducing the idea of expert systems, computer programs that are designed to replicate (and improve on) the performance of human experts in specialized domains, such as identifying diseases in humans and plants, or processing credit card applications. These programs can often be represented as decision trees. There are different ways of constructing expert systems, however. One way is to start with human experts and write a program that codifies their collective knowledge. Alternatively, machine learning algorithms can be used to construct a decision tree by analyzing large databases of examples and deriving rules that can then be used to classify new examples. We illustrate this through ID3, which is an example of a traditional machine learning algorithm.

Traditional algorithms such as ID3 are still highly dependent upon how their databases are labeled and constructed. They typically require lengthy and complex processes of *feature*

307