SORBONNE UNIVERSITY ABU DHABI

Machine learning: Trends, perspectives, and prospects

M. I. Jordan* and T. M. Mitchell*

Presented By: Abrar Al-Sharei

E-mail: abrark_19@outlook.com

Despite practical challenges, we are hopeful that informed discussions among policy-makers and the public about data and the capabilities of machine learning, will lead to insightful designs of programs and policies that can balance the goals of protecting privacy and ensuring fairness with those of reaping the benefits to scientific research and to individual and public health. Our commitments to privacy and fairness are evergreen, but our policy choices must adapt to advance them, and support new techniques for deepening our knowledge.

REFERENCES AND NOTES

1. M. De Choudhury, S. Counts, E. Horvitz, A. Hoff, in *Proceedings of International Conference on Weblogs and Social Media* [Association for the Advancement of Artificial Intelligence (AAAI), Palo Alto, CA, 2014].
2. J. S. Brownstein, C. C. Freifeld, L. C. Madoff, *N. Engl. J. Med.* **360**, 2153–2155 (2009).
3. G. Eysenbach, *J. Med. Internet Res.* **11**, e11 (2009).
4. D. A. Broniatowski, M. J. Paul, M. Dredze, *PLOS ONE* **8**, e83672 (2013).
5. A. Sadilek, H. Kautz, V. Silenzio, in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence* (AAAI, Palo Alto, CA, 2012).
6. M. De Choudhury, S. Counts, E. Horvitz, in *Proceedings of the*

REVIEW

Machine learning: Trends, perspectives, and prospects

M. I. Jordan[1]* and T. M. Mitchell[2]*

Machine learning addresses the question of how to build computers that improve automatically through experience. It is one of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science. Recent progress in machine learning has been driven both by the development of new learning algorithms and theory and by the ongoing explosion in the availability of online data and low-cost computation. The adoption of data-intensive machine-learning methods can be found throughout science, technology and commerce, leading to more evidence-based decision-making across many walks of life, including health care, manufacturing, education, financial modeling, policing, and marketing.

Machine learning is a discipline focused on two interrelated questions: How can one construct computer systems that auto- ance when executing some task, through some type of training experience. For example, in learning to detect credit-card fraud, the task is to as-

About the authors:

1. Michael I. Jordan (M.I. Jordan):
    - Michael I. Jordan is a renowned figure in the field of machine learning and artificial intelligence.
    - He is a Professor at the University of California, Berkeley, with appointments in the Department of Electrical Engineering and Computer Sciences and the Department of Statistics.
    - Jordan's research spans a wide range of topics in machine learning, including probabilistic graphical models, neural networks, and Bayesian statistics.
    - He has made significant contributions to the development of machine learning algorithms and is known for his work in areas such as deep learning, reinforcement learning, and Bayesian networks.
    - Jordan's research has had a substantial impact on both academia and industry, and he has received numerous awards and honors for his contributions to the field.

2. Tom M. Mitchell (T.M. Mitchell):

- Tom M. Mitchell is another influential figure in the field of machine learning and artificial intelligence.
- He is a Professor at Carnegie Mellon University, where he holds the E. Fredkin University Professorship in the School of Computer Science.
- Mitchell's research focuses on a range of topics in machine learning and natural language processing. He has contributed to areas like text mining, knowledge representation, and computational cognitive modeling.
- He is perhaps best known for his work on the development of machine learning techniques for natural language processing tasks, including text classification and information retrieval.
- Mitchell's research has made significant contributions to the intersection of artificial intelligence and cognitive science, and he has received recognition for his achievements in the field.

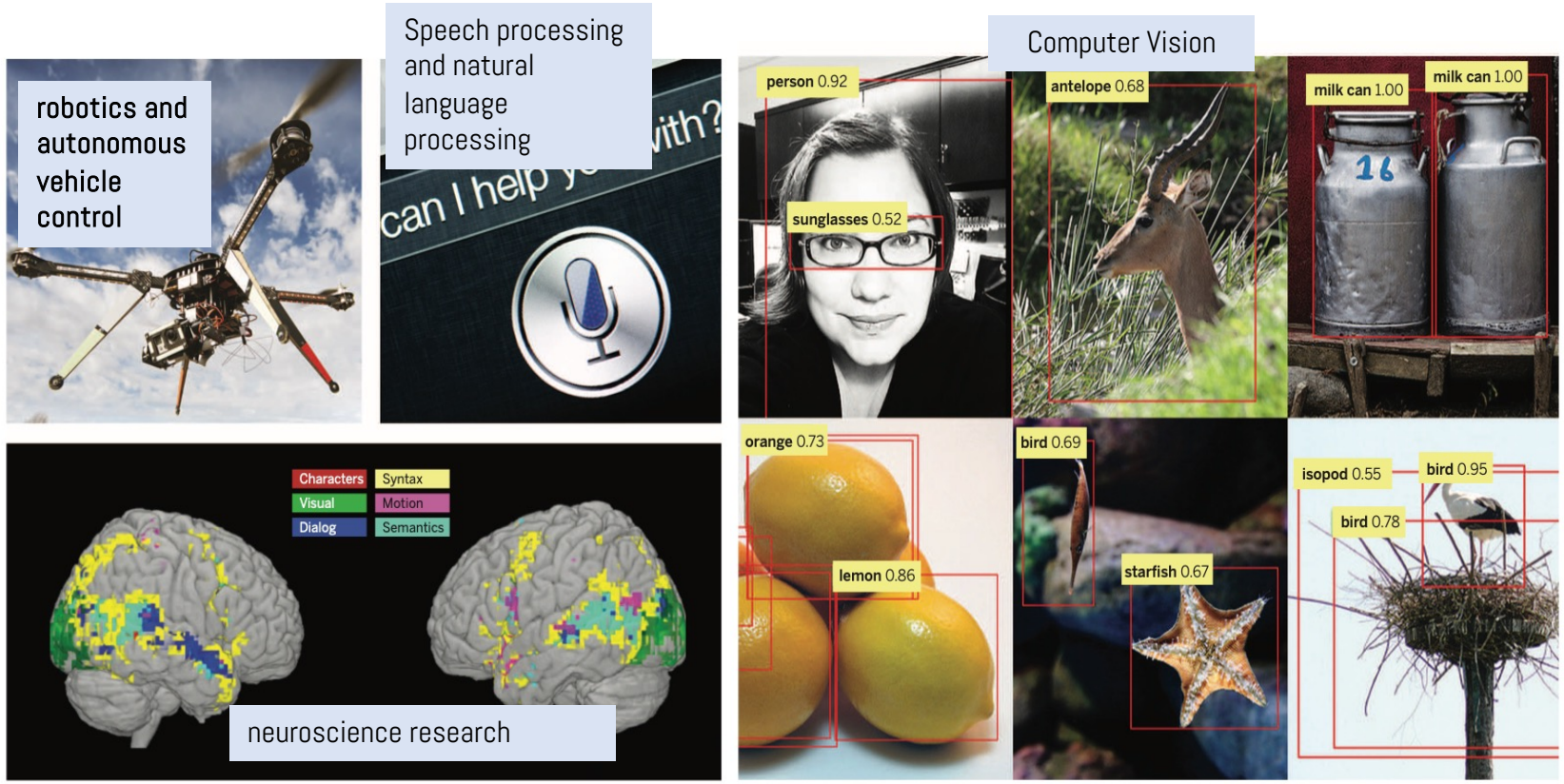
Outline

1. Overview
2. Drivers of machine-learning progress
3. Core methods and recent progress
4. Emerging trends
5. Opportunities & Challenges

Machine Learning

Q1) How can computer systems improve automatically through experience?

Q2) What are the fundamental laws that govern learning systems, including computers, humans, and organizations?



Example

"For example, in learning to detect credit-card fraud, the task is to assign a label of **"fraud" or "not fraud"** to any given credit-card transaction. The performance metric to be improved might be the accuracy of this fraud classifier, and the training experience might consist of a collection of **historical credit-card transactions**, each labeled in retrospect as fraudulent or not."

Three main components of a machine learning system:

1. What is the Input ?
2. What is the Data?
3. What is the desired Output?

Optimization Of Machine Learning:

- When you are dealing with the machine learning problem you need to **optimize 3 main properties**:

1. **Sample Complexity** - how much data are required to learn accurately?
2. **Computational Complexity** - how much computation is required? / how much improvement complexity is there?
3. **Feature Complexity** - Which will be the important one where you put together all your features?

2. Drivers of machine-learning progress

What was the main reason that the machine learning algorithms could make such a difference in the world?

•**Big Data & Personalization**: The growth of "Big Data" and personalized data collection drives machine learning's need for insights from vast, granular datasets.

•**Connected Services**: The trend of connected personalized services leverages collective data while customizing experiences, fostering innovation.

•**Data-Driven Decision-Making**: Across industries, data-intensive, evidence-based decision-making is rising, making machine learning pivotal in enhancing decision processes.

3. Core methods and recent progress

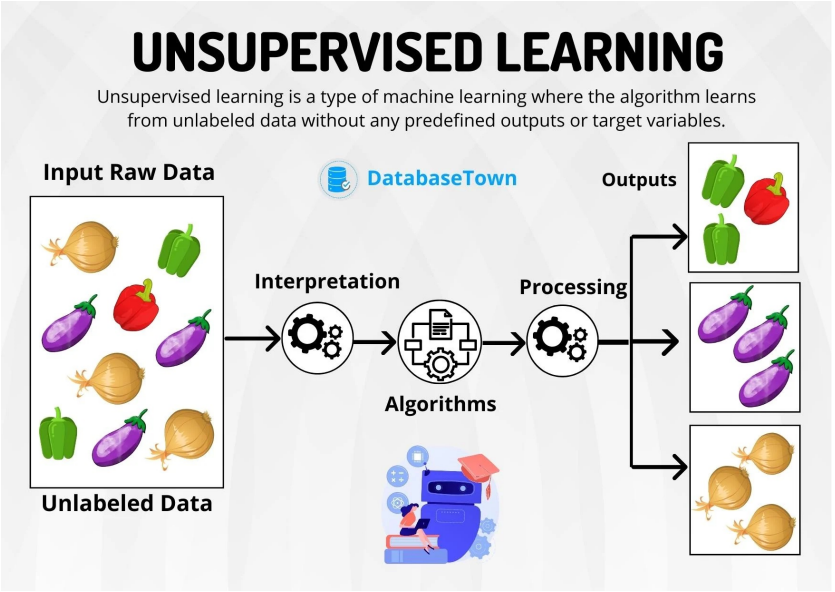
- **Supervised Learning:**
  - Training data: in the form of (x,y) pairs
  - Goal: Generate y* (prediction) in response to x*(query)

Input	Output	
Doccument	Binary	A form that has one of two values. For Example: spam or not spam.
Image	Multiclass	A form that has one value among k labels.
Graph		

3. Core methods and recent progress

Unsupervised Learning:

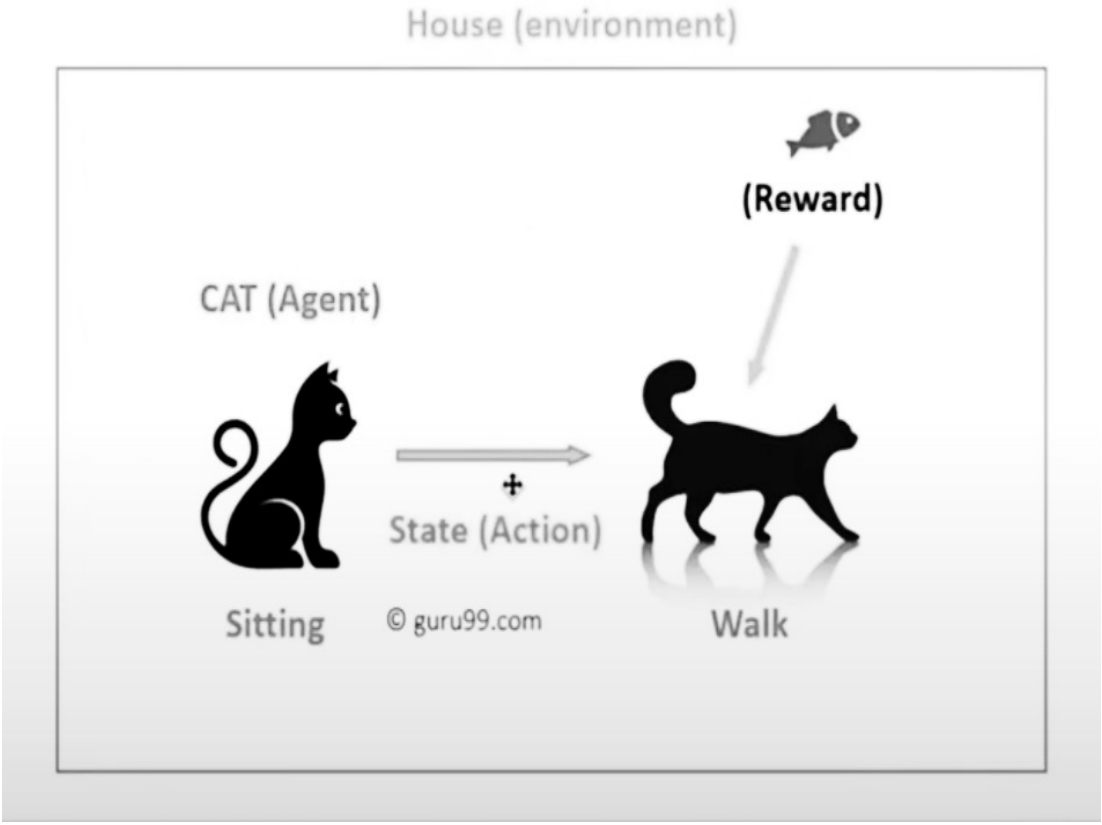
- **Unsupervised Learning**: Involves analyzing unlabeled data with assumptions about structural data properties (e.g., algebraic, combinatorial, or probabilistic).
- **Dimension Reduction**: Techniques like principal components analysis, manifold learning, and autoencoders aim to reveal low-dimensional patterns in data, each making specific assumptions about the underlying data structure.
- **Clustering**: Focuses on partitioning data without explicit labels, with various algorithms developed based on assumptions about what defines a "cluster."
- **Computational Complexity**: Managing computational complexity is vital as these techniques are designed to handle large, unlabeled datasets, eliminating the need for supervised labels.



3. Core methods and recent progress

Reinforcement Learning:

- **Reinforcement Learning**: A major machine learning paradigm that lies between supervised and unsupervised learning.
- **Training Data and Reward Signals**: Unlike supervised learning, it uses training data indicating whether actions are correct. Reward signals apply to entire action sequences.
- **Control Strategy Learning**: Focuses on learning control strategies (policies) in unknown dynamic environments to maximize expected rewards over time.
- **Ties to Control Theory and Psychology**: Draws from control theory concepts and has connections to psychology and neuroscience, including predicting responses of neurons associated with reward learning.



4. Emerging trends

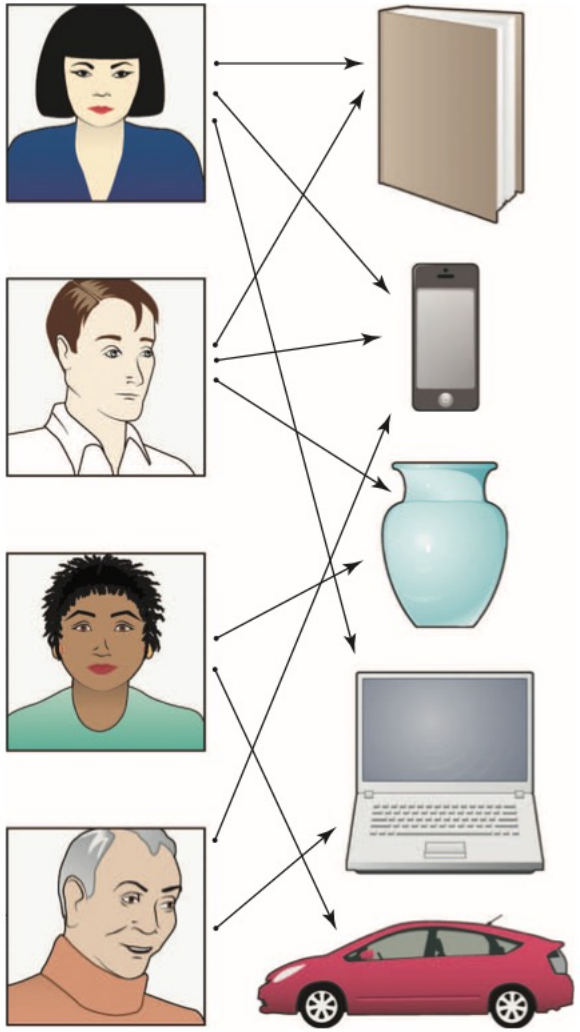


Fig. 4. Recommendation systems

A recommendation system is a machine-learning system that is based on data that indicate links between a set of a users (e.g., people) and a set of items (e.g., products). A link between a user and a product means that the user has indicated an interest in the product in some fashion (perhaps by purchasing that item in the past). The machine-learning problem is to suggest other items to a given user that he or she may also be interested in, based on the data across all users.

5. Opportunities & Challenges

How can we construct some networks that can not only learn once, but it also learn continuously?

Opportunities

- **Lifelong Learning:** Exploring the development of computer systems that continuously learn multiple skills and functions over years, building on previous knowledge.
- **Collaborative Learning:** Investigating the potential for machine learning systems to work collaboratively with humans, combining machine's data analysis capabilities with human background knowledge.

Challenges

- **Privacy Concerns:** The collection of personal data for economic purposes raises significant privacy issues, leading to questions about data security and ethical use.
- **Data Ownership:** Deciding who owns and has access to online data and who benefits from it raises ethical and societal questions. While data is collected by corporations for profit, there is potential for public good if more data sharing were to occur.