MODULE 1: AI Development Techniques

Haitam EL-KHAMALI

22 March, 2025

1 Artificial Intelligence Overview

Artificial intelligence refers to computer programs that can complete cognitive tasks typically associated with human intelligence. There are two main techniques used to design AI programs :

- **Rule-based techniques** involve creating AI programs that strictly follow predefined rules to make decisions. For example, a spam filter using rule-based techniques might block emails that contain specific keywords using its predefined logic.
- **Machine learning techniques** involve creating AI programs that can analyze and learn from patterns in data to make independent decisions. For example, a spam filter using these techniques might flag potential spam for the recipient to review, preventing automatic blocking. If the recipient marks emails from trusted sources as safe, the spam filter learns and adapts its logic to include similar emails from that sender in the future.

AI tools can use either rule-based or ML techniques, or even a combination of both. In general, rule-based techniques are commonly used for tasks that require rigidity, such as blocking messages from untrusted senders that are obviously spam, like requests for bank transfers or private information. Conversely, ML techniques are better suited for tasks demanding flexibility and adaptability, like learning to recognize that messages from trusted senders containing typos are not spam.

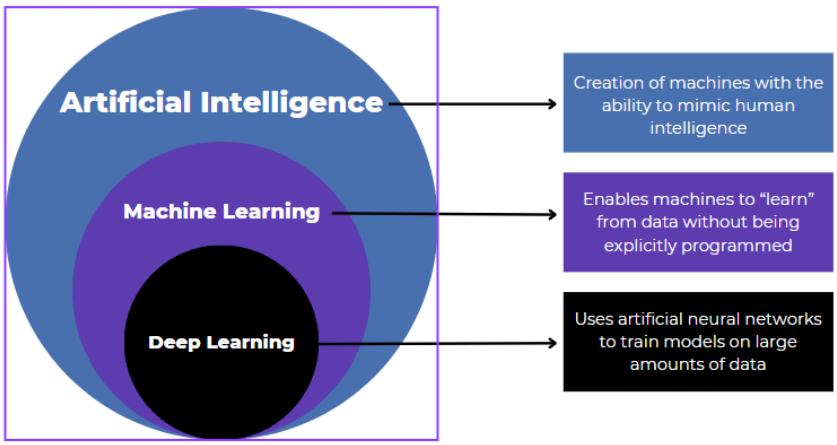
2 Approaches to Training ML Programs

2.1 Definition of Machine Learning

Machine learning is a subset of artificial intelligence that focuses on developing computer systems and algorithms capable of learning from data without being explicitly programmed. These systems improve their performance over time as they process more information. More specifically, machine learning :

- Uses statistical techniques to enable computers to "learn" patterns from data
- Allows systems to automatically identify relationships, make decisions, and predict outcomes
- Improves performance through experience without human intervention
- Adapts based on new examples it encounters
- Functions through various approaches including supervised learning (trained on labeled data), unsupervised learning (finding patterns in unlabeled data), and reinforcement learning (learning through trial and error with rewards)

Recall that machine learning is a subset of AI focused on developing computer programs that can analyze data to make decisions or predictions. AI designers often use ML in their AI programs because it doesn't have the limitations of rule-based techniques.



There are three common approaches to training ML programs :

- **Supervised learning.**
- **Unsupervised learning.**
- **Reinforcement learning.**

2.2 Supervised Learning

In this approach, the ML program learns from a labeled training set. A labeled training set includes data that is labeled or tagged, which provides context and meaning to the data. For instance, an email spam filter that's trained with supervised learning would use a training set of emails that are labeled as "spam" or "not spam." Supervised learning is often used when there's a specific output in mind.

2.3 Unsupervised Learning

In this approach, the ML program learns from an unlabeled training set. An unlabeled training set includes data that does not have labels or tags. For instance, ML might be used to analyze a dataset of unsorted email messages and find patterns in topics, keywords, or contacts. In other words, unsupervised learning is used to identify patterns in data without a specific output in mind.

2.4 Reinforcement Learning

In this approach, the ML program uses trial-and-error to learn which actions lead to the best outcome. The program learns to do this by getting rewarded for making good choices that lead to the desired results. Reinforcement learning is commonly used by conversational AI tools. As these tools receive feedback from users and AI designers, they learn to generate effective responses.

Each ML technique has its own strengths and weaknesses. Depending on the type of data that's available and what's needed to solve the particular problem, AI designers may use one, two, or all three of these techniques to produce an AI-powered solution.

3 Generative AI

Advancements in supervised, unsupervised, and reinforcement learning have helped pave the way for generative AI—AI that can generate new content, like text, images, or other media. When users provide input in the form of a prompt—text input that provides instructions to the AI model on how to generate output—the AI tool processes this input to generate new content.

For instance, all three approaches play distinct roles in conversational AI tools. Supervised learning equips conversational AI tools with foundational dialogue data, enabling them to respond to common conversational cues appropriately. Unsupervised learning enables them to interpret nuances in language, like colloquialisms, that occur naturally in conversation. Reinforcement learning further strengthens these tools by allowing them to improve their responses in real-time based on user feedback. This enables them to adapt to the conversational context and engage in natural conversations.

Generative AI's ability to create and innovate offers a range of benefits to all sorts of workplaces and professions, such as marketing, product development, engineering, education, manufacturing, and research and development. These benefits include :

- **Greater efficiency** : Generative AI can automate or augment routine tasks, allowing workers to focus on other work priorities.
- **Personalized experiences** : Generative AI can tailor its interactions to individual preferences and needs.
- **Better decisions** : Generative AI can quickly analyze vast amounts of data to uncover useful insights.

These are just some of the ways that generative AI can enhance your work.

4 Important Limitations

Despite its capabilities, AI has significant limitations that require human oversight :

- **No independent learning** : AI needs continual human updates to its training.
- **Training data issues** : Shortcomings in training data can reflect or amplify biases, leading to unfair outcomes.
- **Hallucinations** : AI can produce inaccurate outputs (hallucinations) ranging from minor errors to significant distortions.

4.1 Example of Hallucination Risk

A sales manager using AI to analyze quarterly data might receive a recommendation to remove a product due to declining sales. However, if seasonal factors affecting sales weren't accounted for in the AI's analysis, following this recommendation without human review could lead to a misguided business decision.

4.2 The Need for Human Oversight

Effective AI management in the workplace requires :

- Careful review of AI-generated outputs.
- Collaboration between technical and non-technical roles.
- Alignment of AI output with human values and ethics.

5 AI Augmentation

Definition : Using AI to improve work products by making them easier to create or higher in quality.
Application : AI can augment various job tasks—from simple duties like email responses to complex activities like collaborative brainstorming—helping you complete work more quickly and efficiently.

6 AI Automation

Definition : Using AI to accomplish tasks without any human action required.
Example : A customer support representative dealing with hundreds of daily emails could use AI to :

- Automatically sort incoming messages by priority.
- Draft responses to low-priority inquiries.
- Allow the representative to focus on complex issues requiring personal attention.

7 Finding the Right Balance

Achieving the optimal balance between augmentation and automation requires time, practice, and careful consideration. Successful businesses distinguish themselves by applying people-first approaches to AI across their products, services, and jobs.