SECTION 4
USING OF MODERN ARTIFICIAL INTELLIGENCE MODELS IN INFORMATION CONTROL SYSTEMS

Derbentsev V.D., Candidate of Economic Sciences, Professor, Kyiv National Economic University named after Vadym Hetman,
Bezkorovainyi V.S., Candidate of Economic Sciences, Associate Professor, Kyiv National Economic University named after Vadym Hetman

4.1. Basic concepts of Artificial Intelligence

Artificial Intelligence (AI) has become one of the most important and discussed fields in modern science and technology. The concept of AI arouses admiration and indignation, creates hopes and fears, because it involves the creation of systems capable of self-education, self-learning and self-improvement.

AI is a branch of computer science that seeks to create intelligent agents that can reason, learn, and act autonomously. These agents can be software systems, robots, or other types of machines that possess a certain level of intelligence.

There are many different concepts that underlie AI. Some of the most important of these include [1–4]:

– *Machine Learning* (ML) – explores methods by which computer systems can learn from data without explicit programming;

– *Deep Learning* (DL) is a subfield of ML that uses ANNs with many layers to automatically perform learning tasks;

– *Artificial Neural Networks* (ANNs) – this type of models imitates the work of the human brain and is used in many applications of AI, in particular in Machine Learning;

– *Natural Language Processing* (NLP) – studies the interaction between computers and humans through natural language, it includes language recognition, language understanding, language generation and other aspects;

– *Computer Vision* (CV) – explores methods by which computers can interpret visual information;

– *Expert Systems* – software systems that use a knowledge base, rules and information to solve problems in a specific fields, providing recommendations or solutions based on input data;

– *Autonomous Systems* – systems that can act independently of external influence, making their own decisions and interacting with the environment;

– *Distributed Systems* – systems that combine various AI components to work together and achieve common goals;

– *Search* – focuses on the development of algorithms that can find optimal solutions for the different complex tasks.

Let's look at the main differences between the main components of AI as a scientific and applied field (Fig. 4.1).

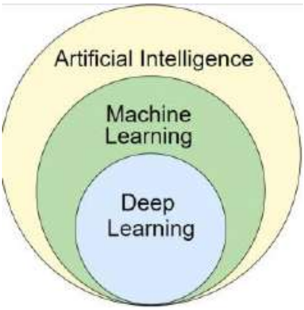

Figure 4.1. Relations between AI, ML and DL

Source: designed by the authors based on [1–4]

Artificial Intelligence: AI is a broad field that aims to create intelligent machines or systems that can perform tasks that typically require human intelligence, such as perception, reasoning, learning, problem-solving, and decision-making. AI encompasses various techniques and approaches, including machine learning, deep learning, expert systems, natural language processing, and robotics.

The main concept behind AI is to develop algorithms and computational models that can mimic human-like intelligence and behaviour. AI systems can be rule-based, where they follow a set of predefined rules and logic, or they can be based on machine learning and deep learning techniques, where they can learn from data and adapt their behaviours accordingly.

ML is a subset of AI that focuses on developing algorithms and statistical models that enable systems to learn from data and make predictions or decisions without being explicitly programmed. The main concept behind ML is to identify patterns and relationships within data and use these insights to make accurate predictions or decisions on new, unseen data.

ML algorithms can be broadly categorized into Supervised Learning (where the data is labelled, and the algorithm learns to map inputs to outputs), Unsupervised Learning (where the data is unlabelled, and the algorithm tries to find inherent patterns and structures), and Reinforcement Learning (where an agent learns to make decisions by interacting with an environment and receiving rewards or penalties).

DL is a subset of ML that is inspired by the structure and function of the human brain. It involves the use ANNs with multiple layers, allowing the model to learn and represent complex patterns and features from data automatically.

The main concept behind DL is to create hierarchical representations of data by stacking multiple layers of non-linear transformations. Each layer in a deep neural network learns to extract increasingly abstract and complex features from the input data. These learned features are then used to make predictions or decisions.

DL algorithms, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders and Transformers have been particularly successful in areas like CV, NLP, speech recognition, and time series forecasting, where they can effectively capture and model complex patterns and relationships within the data.

Distinguishing between AI, ML, and DL:

– AI is the broadest field, encompassing various techniques and approaches for creating intelligent systems, including ML and DL.

– ML is a subset of AI that focuses on developing algorithms and models that can learn from data and make predictions or decisions without being explicitly programmed.

– DL is a subset of ML that uses deep neural networks with multiple layers to automatically learn hierarchical representations and complex patterns from data.

So, AI is the overarching field, ML is a specific approach within AI that enables systems to learn from data, and DL is a powerful technique within ML that uses deep neural networks to automatically learn hierarchical representations and patterns from data.

The main tasks of ML can be defined as follows [1-5]:

– *Classification*: Classification is a supervised learning task where the goal is to assign a class label or category to a given input instance based on its features. Examples include image classification (recognizing objects in images), spam detection (classifying emails as spam or not spam), and sentiment analysis (classifying text as positive, negative, or neutral).

– *Clustering* (Clusterization): Clustering is an unsupervised learning task that aims to group similar instances or data points together based on their features or characteristics. The goal is to identify inherent patterns or structures within the data without any predefined labels. Examples include customer segmentation (grouping customers based on their behaviours or preferences) and image segmentation (identifying distinct regions or objects in an image).

– *Regression*: Regression is a supervised learning task that involves predicting a continuous numerical value based on input features. It is commonly used for tasks like stock price prediction, sales forecasting, and estimating real estate prices based on various factors.

– *Anomaly Detection*: Anomaly detection, also known as outlier detection, is the task of identifying rare or unusual instances or patterns within a dataset that deviate significantly from the norm. It has applications in fraud detection, intrusion detection systems, and predictive maintenance, where identifying anomalies can help prevent failures or detect malicious activities.

– *Dimensionality Reduction:* Dimensionality reduction is the process of reducing the number of features or variables in a dataset while retaining most of the relevant information. It is useful for visualizing high-dimensional data, reducing computational complexity, and improving the performance of machine learning models by removing redundant or irrelevant features.

– *Reinforcement Learning*: Reinforcement learning is a type of machine learning where an agent learns to make decisions by interacting with an environment and receiving rewards or penalties for its actions. It is used in applications like game playing (e.g., chess, Go), robotics, and autonomous systems (e.g., self-driving cars).

– *Time Series Forecasting*: Time series forecasting is the task of predicting future values of a sequence of data points based on historical observations. It is used in various domains, such as finance (stock market prediction), weather forecasting, sales forecasting, and energy demand prediction.

To solve these problems, a large number of effective models and algorithms have been developed within the ML paradigm:

Linear Regression: Linear regression is a supervised learning algorithm used for regression tasks. It models the relationship between a dependent variable and one or more independent variables using a linear equation. It is commonly used for predicting continuous values, such as house prices or stock prices.

Logistic Regression: Logistic regression is a supervised learning algorithm used for classification tasks. It models the probability of an instance belonging to a binary class (0 or 1) based on its features. It is widely used in areas like spam detection, credit scoring, and medical diagnosis.

Decision Trees: Decision trees are a type of supervised learning algorithm that can be used for both classification and regression tasks. They work by recursively splitting the data based on the feature values, creating a tree-like structure of decisions. Decision trees are easy to interpret and can handle both numerical and categorical data.

Random Forests: Random Forests are an ensemble learning method that combines multiple decision trees. Each tree is trained on a random subset of the data and features, introducing randomness and reducing overfitting. Random Forests are highly accurate and can handle high-dimensional data and complex relationships.

Support Vector Machines (SVMs): SVMs are supervised learning algorithms used for both classification and regression tasks. They work by finding the optimal hyperplane that maximizes the margin between classes or data points. SVMs are effective for high-dimensional data and can handle non-linear problems using kernel tricks.

Naive Bayes: Naive Bayes is a family of probabilistic classifiers based on Bayes' theorem. It assumes independence between features and calculates the probability of an instance belonging to each class given its feature values. Naive Bayes classifiers are simple and computationally efficient, making them suitable for large datasets and real-time applications.

K-Nearest Neighbors (KNN): KNN is a non-parametric algorithm used for both classification and regression tasks. It works by finding the k nearest neighbors of a new instance in the feature space and predicting the class or value based on the majority vote or average of the neighbors.

K-Means Clustering: K-Means is a popular unsupervised learning algorithm used for clustering tasks. It partitions the data into k clusters by iteratively minimizing the sum of squared distances between data points and their assigned cluster centroids.

Principal Component Analysis (PCA): PCA is a dimensionality reduction technique used for feature extraction and data visualization. It transforms the high-dimensional data into a lower-dimensional space while preserving the maximum variance in the data.

ANNs: ANNs are a class of algorithms inspired by the structure and function of the human brain. They consist of interconnected nodes (neurons) that can learn complex patterns and relationships from data. ANNs are widely used for tasks like image recognition, natural language processing, and time series forecasting.

Stages of an AI models design and development

The development of an AI model is a complex process that involves multiple stages and requires careful consideration of various factors (Table 4.1).

Table 4.1

STAGES OF AN AI MODELS DEVELOPMENT

Stage	Description
Task Definition	• Formulate the problem the AI model should solve • Define success/performance metrics and expected results • Collect data for training and testing models
Algorithm Selection	• Select an AI algorithm based on the task type and available data • Algorithms can include ANNs, Decision Trees, SVMs, Genetic Algorithms, etc.
Data Preparation	• Clean and format data for the AI model to understand • Divide data into training, validation, and test datasets
Model Training	• Train the AI model on the training dataset Optimize model parameters to minimize error on the validation dataset
Model Testing	• Test the AI model on the test dataset to evaluate generalization ability identify strengths and weaknesses of the model
Model Improvement	• Improve the AI model based on test results Modify model parameters, algorithm, or data as needed
Model Deployment	• Deploy the improved AI model in a real environment • Use the model for decision-making, prediction, or task automation
Monitoring and Maintenance	• Continuously monitor the AI model's effectiveness • Update or retrain the model as new data or conditions emerge

Source: designed by the authors

The process of developing an AI model begins with a clear understanding of the problem or task at hand. This involves defining the objectives, success criteria, and performance metrics that the model should achieve. Additionally, it is crucial to determine the type and quality of data required for training and evaluating the model.

Once the task is well-defined, the next step is to select the appropriate AI algorithm or model architecture. This choice depends on the nature of the problem, the available data, and the desired performance characteristics. There are numerous algorithms and architectures to choose from, such as deep neural networks, decision trees, support vector machines, and ensemble methods, each with its own strengths and weaknesses.

Data preparation is a critical stage in the AI model development process. It involves collecting, cleaning, and formatting the data to ensure that it is suitable for training the chosen model. This may involve handling missing values, removing outliers, and transforming the data into a format that the model can understand. Additionally, the data is typically divided into training, validation, and testing sets to evaluate the model's performance and generalization ability.

The model training stage is where the AI algorithm learns from the prepared training data. This process involves iteratively adjusting the model's parameters to minimize the error or loss function between the model's predictions and the true labels or values in the training data. Techniques like regularization, early stopping, and learning rate scheduling are often employed to prevent overfitting and improve the model's generalization performance.

Once the training process is complete, the model is evaluated on the *validation and testing* datasets to assess its performance and generalization ability. This evaluation may reveal areas for improvement, such as the need for additional data, adjustments to the model architecture, or the selection of a different algorithm altogether.

The development process is iterative, and the model may go through multiple cycles of refinement and improvement before it is ready for deployment. This may involve exploring different model architectures, tuning hyperparameters, or incorporating additional data sources.

Finally, once the model meets the desired performance criteria, it can be *deployed* in a production environment for real-world use. However, even after deployment, the model may require on-going monitoring and maintenance to ensure that it continues to perform well as new data becomes available or as the underlying conditions change over time.

Overall, the development of an AI model is a complex and iterative process that requires careful planning, domain expertise, and a deep understanding of the underlying algorithms and techniques.

It is important to note that AI model development is an iterative process. At each stage, problems may arise that need to be solved. With careful planning, the right choice of algorithms and careful testing, you can create AI models that will help solve complex problems.

The 20th century witnessed the rapid development of AI. Starting with theoretical considerations about the possibility of creating artificial people, AI has turned into a powerful scientific discipline with many practical applications.

It should be noted that the development of AI has not been linear. There have been periods of rapid progress and periods of stagnation. However, the general trend is that AI is becoming more powerful and sophisticated, and has many potential applications to improve our lives.

Despite significant progress, AI still faces a number of challenges:

– Ethical issues – AI can have negative consequences for society, such as job losses, bias and abuse;

– Black box problem – some AI algorithms are unclear, making it difficult to understand their solutions;

– Security issues – AI can be used to create autonomous weapons or other dangerous systems.

4.2. The role of time series forecasting in the development and use of AI models for ICS

AI has a wide range of applications in many industries, such as:

– Medicine – for diagnosis of diseases, development of new drugs and personalization of treatment;

– Production – for optimization of production processes, demand forecasting and supply chain management;

– Transport – for the development of self-driving cars, optimization of routes and management of transport networks;

– Finance – to detect fraud, predict market trends and automate trade.

It should be noted that time series forecasting is an important task in various fields and industries due to its ability to predict future values or patterns based on historical data. Accurate forecasting can provide valuable insights and support decision-making processes, leading to improved planning, resource allocation, and risk management. Here are some key points highlighting the importance of time series forecasting and its applied fields:

1. Business and finance:

– Sales and revenue forecasting: Predicting future sales figures and revenue streams is crucial for budgeting, inventory management, and resource planning.

– Stock market forecasting: Forecasting stock prices, indices, and trends can aid in investment decisions and portfolio management.