The Convergence of Learning, Intelligence, and Emotion in Artificial Systems: A Literature Synthesis

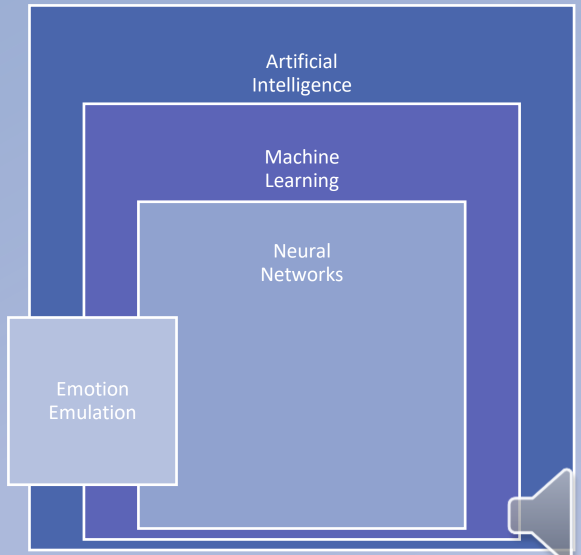
Anthony J. Coots

National University

May 7, 2025

Introduction to AI and Its Interconnected Aspects

- Artificial Intelligence (AI) has become more prevalent in recent years, much more than its mathematical concepts developed less than a century ago. AI can be defined as “intelligence demonstrated by machines” (National University, 2025), and it genuinely exhibits intelligent behavior in various forms.
- Important areas of AI are, but not limited to, neural networks, machine learning, and emotion emulation. It is important to understand that not all AI is based on machine learning, not all machine learning utilizes neural networks, and not all of these areas are intended to participate in emotion emulation.
- The purpose of this presentation is to demonstrate the core approaches to artificial intelligence and the following topics:
    - Understanding neural networks
    - Machine learning overview
        - Supervised, unsupervised, semi-supervised learning, and reinforcement learning
    - The role of emotions in AI and how they are emulated
    - Anthropomorphism in AI
    - Contrasting viewpoints from literature
    - Future research directions



Approaches to Artificial Intelligence

- According to Dietrich (2005), AI methodologies align with cognitive science perspectives and are categorized as:
  - Symbolic AI
  - Connectionist AI
  - Dynamic/Evolutionary AI
  - Embodied Robotic AI

And generally describe the approaches to artificial intelligence in a wide variety of applications. **Symbolic AI** uses explicit symbols and rules to represent knowledge and perform reasoning to make AI more understandable. Wajahat et al. (2024) highlight this concept when demonstrating accuracy and explainability in Diabetes prognosis by integrating knowledge through knowledge graphs. Medical data was transformed into graphical representations to present relationships, such as those between Body Mass Index and blood sugar levels. The graphs were used as embeddings in machine learning models to enhance prediction accuracy.

Connectionist AI is inspired by the human brain's current understanding of how it works. This AI, unlike symbolic AI, does not rely on explicit symbols or rules and is often seen in the form of neural networks. This approach is often seen in large-scale neural networks such as ChatGPT, Gemini, or Meta's LLaMA, is commonly referred to as that "black box" AI since the decision-making processes are opaque or unclear to the uninformed user.

Evolutionary AI uses principles from biological evolution. MirzagoltabarRoshan et al. (2022) use evolutionary AI by integrating a Particle Swarm Optimization algorithm to optimize a Multivariate Adaptive Regression Splines-based regression model. In layman's terms, the study utilized AI to fine-tune the parameters of a mathematical model, enabling it to predict the strength of eco-friendly concrete by simulating intelligent group behavior, thereby continually improving its prediction quality.

Lastly, **Embodied Robotic AI** is AI that takes a physical form, learning from not just abstract but also analog data and adapting to new environments. It can take the form of humanoids or even the small vacuums that float around homes, such as Roomba, utilizing real-time data from sensors and actuators.

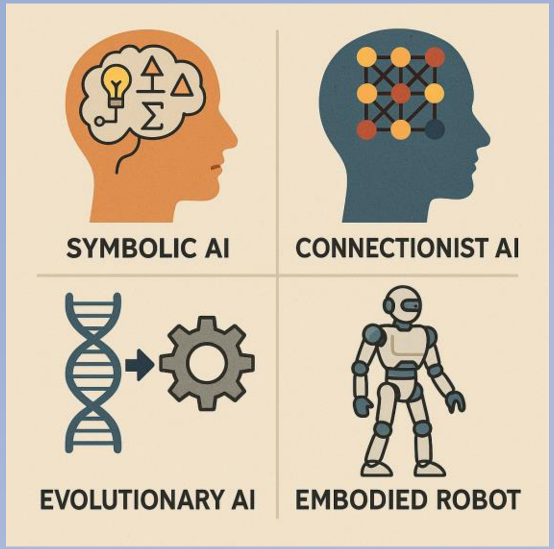


Image Courtesy of Connectionist AI.

Understanding Neural Networks

- A Neural Network is a model of computation that is inspired by the architecture and activity of the human brain. It consists of interconnected nodes, or “layers,” that work together to produce an output based on the weighted inputs from each node (Neural network, 2016). In other words, it is a simplified version of how the human brain works. Consider the following:
  - Nodes are a team of decision makers who all have to talk to each other.
  - Each node listens to some information, makes a small judgment, and passes it to the next layer.
  - These decisions are combined to determine an answer, such as recognizing a feature in an image.
- Neural networks are widely used in connectionist AI for tasks that require pattern recognition and decision making across large datasets. Input data is processed through multiple layers of interconnected nodes, where each node adjusts its weights based on the data until a final output is produced. For example, Torabzadeh et al. (2022) applied a neural network model to analyze survey data assessing trust and patient satisfaction in a home healthcare (HHC) setting. Unlike symbolic AI, this approach does not rely on explicit rules; instead, it interprets sentiment based on contextual patterns in language. Words like “great” or “terrible” may typically signal positive or negative sentiment, but their meaning is subject to the context. For instance, in the sentence “The day was great, the service was not,” the sentiment is mixed, despite the word “great” appearing. Neural networks can learn to navigate these nuances by identifying deeper patterns across this entire dataset, which offers a flexibility that rule-based systems lack.

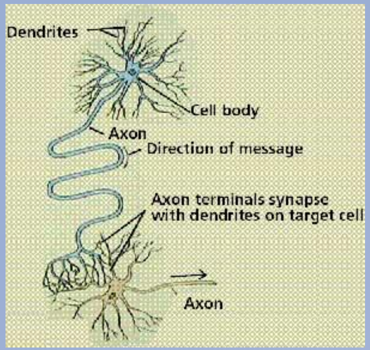


Figure 1
Note. From *Neural Networks – Biology* by Stanford University students, 2003, Stanford University. Retrieved from https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Biology/index.html. Used under fair use.

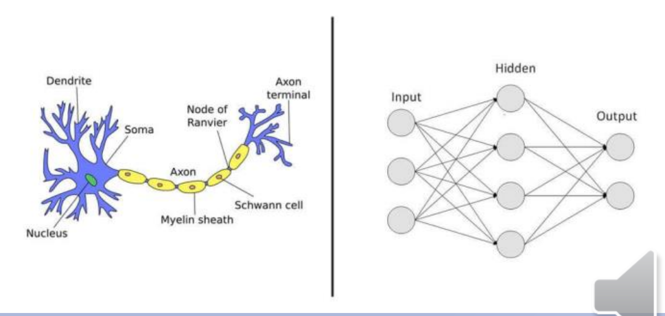


Figure 2
Note. From A gentle introduction to neural networks by J. Joseph, 2025, CleverTap. Retrieved from https://clevertap.com/blog/neural-networks/. Used under fair use.

Machine Learning Overview

- Reilly et al. (2003), in the Encyclopedia of Computer Science, define machine learning as "the study of methods for constructing and improving software systems by analyzing examples of their desired behavior rather than by direct programming". Instead of explicitly instructing a computer on how to perform specific tasks, machine learning involves showing the computer what a "correct" outcome looks like, allowing it to learn and achieve that outcome independently when presented with data and thus is a subset of artificial intelligence, "intelligence demonstrated by machines" (National University, 2025).
- There are three roles in machine learning: data, training, and algorithms.
- Lin et al. (2022) demonstrate these roles in their study on predicting human decision-making using machine learning in the following ways to support a recurrent neural network.
    - Data: The researchers utilized behavioral data from various studies that involved two concepts: the Iterated Prisoner's Dilemma and the Iowa Gambling Task.
    - Training: The data is fed into a Long Short-Term Memory model to learn patterns and examine what someone did before, aiming to predict their next action, improving with time.
    - Algorithms: The model used is capable of handling complex thinking patterns, such as learning or changing strategies, and performed better than simpler models at predicting the next move when prompted.
- Ultimately, there are four different types of machine learning: supervised, unsupervised, semi-supervised, and reinforcement learning.

Supervised Learning

- Supervised learning is a subset of machine learning where a model is trained on a "labeled dataset."
- A "labeled dataset" includes the input features and the target output.
- The goal of supervised learning is to train a model to learn the relationships between input variable(s) and an output variable in order to predict a correct output when presented with new data the model had not yet seen when trained.
- Some examples of supervised learning are, but not limited to:

Scenario	Input	Output
A realtor agency wishes to predict the value of a future residential development based on several known factors about existing homes in a similar, nearby neighborhood.	The number of bedrooms, square footage, location, etc.	Price of the house based on the relationship to the input features, to predict pricing in the new development.
A hospital aims to utilize machine learning for diagnostic assistance with diseases that may otherwise go undetected in routine operations.	The symptoms, test results, and other data such as age of a patient.	Identify potential medical diagnoses that might otherwise not be immediately apparent to an Emergency Room doctor.
A tech company aims to detect spam emails before an employee falls victim to a phishing attack.	The text, email address, and subject.	Based on the data it was trained on, determine if an email is spam or not spam before it officially waits in an employee's inbox.

Unsupervised Learning

- Unsupervised learning is a subset of machine learning where a model is trained using an unlabeled dataset, the opposite of supervised learning.
- The involved algorithm *tries* to identify patterns or groups in the data with no guidance, such as a target variable. For example, market basket analysis utilizing Apriori, which follows a "if a buys x and y, then they are likely to buy z" rule set.
- It is typically used for clustering, dimensionality reduction (i.e., when more than one column is identical in nature), and anomaly detection.
- Some examples of unsupervised learning are, but not limited to:

Scenario	Input	Output
An e-commerce company aims to create customer groups for targeted marketing.	Purchase history, browsing pattern, and demographics.	Group customers into **clusters** with similar behavior.
A genetics lab is analyzing DNA samples to identify groupings that may reveal insights into traits, evolutionary links, or health risks among different populations.	Sequences found in the DNA and other biological data.	Discover clusters of genetic similarity.
A tech company aims to detect unusual behavior to identify potential security threats.	User metadata, such as login times, access locations, and typical usage patterns on the machine.	Identify **anomalies** in user behavior, typically indicating a potential security threat.

Reinforcement Learning

- Reinforcement learning is a subset of machine learning where an “agent” learns to make decisions by performing in an environment to maximize a hypothetical “reward.”
- The “agent” learns through trial and error and receives feedback from its actions in the form of “rewards” or “penalties.”
- Different from supervised learning, there is no fixed dataset of labeled input-output pairs; the model (agent) learns from its experiences.
- Used in robotics, gaming, and real-time decision making systems.
- Some examples of reinforcement learning are, but not limited to:

Scenario	Input	Output
A game development team aims to create an AI that learns to play a video game.	The game environment state including position, score, enemy, and location.	Learns to take “optimal” actions such as moving around, jumping, and shooting to get a maximized score over time.
A robotics team develops a walking robot that may adapt to a new terrain.	Sensor readings involving leg position, stability, tilt, and grip.	Determine the control actions necessary to result in stable movement and fall less over time.
A logistics company wants to optimize delivery routes *using reinforcement learning*.	Real-time traffic data, delivery locations, and vehicle status.	The best delivery path is chosen to minimize time and fuel consumption, thereby maximizing efficiency.

Semi-Supervised Learning

- Semi-supervised learning is a subset of machine learning that combines a small amount of labeled data with a large amount of unlabeled data during training.
- It makes a connection between supervised and unsupervised learning by using the labeled data to guide the learning process, while leveraging the unlabeled data to improve a model's accuracy and ability to generalize.
- It is useful when labeled data is expensive or time-consuming to collect or obtain, but unlabeled data is not.
- Applications include image recognition, content classification, and speech analysis
- Some examples of semi-supervised learning are, but not limited to:

Scenario	Input	Output
A photo studio wants to create a system trained with a set of labeled images and a larger set of images to identify.	A small number of images with known labels and more unlabeled images.	The ability to accurately label new, unseen images from both labeled and unlabeled examples.
A tech company wants to improve their moderation bot or system using a small set of abusive content and more unlabeled content.	A few posts or comments that are deemed abusive and more unlabeled comments.	Classified future comments as abusive or not abusive with improved confidence.
A research team has approved access to diseased labeled patient records and more unlabeled records.	Demographic, clinical, and test result data with limited labels that indicate disease presence.	Predicting disease risk in new patients more reliably than labeled data alone.

Emotions in AI

- While different approaches try to theoretically frame how we can get computers to “feel,” computers do not feel. Instead, in AI, emotions are modeled based on computational frameworks of what it is like to “feel.”
- These models can simulate human-like emotional responses to stimuli (things that provoke action). Still, they require input data such as text, facial expressions, and tone to even begin to comprehend how to demonstrate such intelligence.
- Emotions in AI could enable better human-to-agent interaction. For example, emotional modeling contributes to *self-regulated learning* environments (Channa et al., 2021), where AI adapts based on user affect, such as frustration or confusion, to maintain an engaging learning environment.
  - This demonstrates the intelligent ability to identify relationships between facial expressions and potential learning outcomes as a way to interpret emotions. It is important to note that not everyone needs to smile broadly to be content; in fact, a person may appear discontented yet still be actively learning, and that may be demonstrated in their learning outcomes.
- For working fields such as healthcare, particularly home healthcare, neural networks (a subset of machine learning, which is a subset of artificial intelligence) have been utilized to interpret emotional input and assess patient trust and satisfaction (Torabzadeh et al., 2022).
  - This demonstrates the intelligent ability to *interpret* emotion in order to perceive trust and satisfaction in health services.

Emulating Emotions in AI

- Emulating emotions through artificial intelligence can take different forms, as there are generally four approaches to artificial intelligence: **symbolic**, **connectionist**, **evolutionary**, and **embodied robot**.
- Each approach, of course, has its strategy:

Artificial Intelligence Approach	Strategy	Literature Examples
Symbolic	Explicit rule and symbol-based logic systems; emotional responses are predefined through structured logic.	Bizzarri et al. (2024)
Connectionist	Neural network-based; learns emotional behavioral from patterns in data.	Channa et al. (2021), Torabzadeh et al. (2022)
Evolutionary	Adaptive models that evolve emotional responses through optimization over time.	MirzagoltabarRoshan et al. (2022)
Embodied Robotic	Emotions are simulated from an interaction between sensors and internal states that are based from real-world behavior.	Canamero (2021)

- The approaches listed above demonstrate that there are philosophical differences in how intelligence elicits “emotion,” ranging from explicit logic to more adaptive systems.
- Some perspectives argue that genuine emotion requires consciousness, which is often defined in a manner that deliberately excludes artificial systems. This raises ongoing philosophical debates about whether AI can genuinely “feel” or only simulate. Conversely, others express concerns that emotionally responsive, anthropomorphic AI could render man and machine interchangeable, which contributes to the fear of AI autonomy and dominance.

Anthropomorphism in AI

- Anthropomorphism in artificial intelligence refers to the “intelligence demonstrated by machines” (National University, 2025) in a form that ascribes human characteristics; these machines appear, sound, or behave like humans.
- AI anthropomorphism includes physical traits such as voice, face, personality, and emotional characteristics (Alabed et al., 2022).
- Reasons researchers aim to implement anthropomorphism in this field include:
  - Enhance trust, emotional connection, and engagement.
  - Promote self-congruence so that users relate more to the AI that is perceived as similar to themselves.
  - May ultimately contribute to “self-AI integration,” where users psychologically incorporate the AI as part of their identity.
- There are perceived cautionary implications for this research:
  - Raises questions of digital dependency, privacy, or psychological well-being, Alabed et al. (2022) cite ‘digital dementia,’ a term that refers to cognitive decline from dependence on this AI.
    - I would like to add for fairness: similar concerns were raised when calculators first became available for purchase. People predicted that their impact on society would be negative, but those fears ultimately proved otherwise.
  - There is a risk of identity threats with self-AI integration, as humans have developed a sense of self prior to AI.
- Further research must involve the ethical design, boundaries, and long-term impacts that anthropomorphic AI may have on individuals and society as a whole.

Contrasting Viewpoints from Literature

- Artificial intelligence is a significant global concern that has created extensive debate among educational professionals and settings (K-12, college), governments, and others regarding its transparency, anthropomorphic implementation, emotion emulation, and ethical implications.

- Artificial intelligence is being researched and implemented at a much higher rate than its early theoretical foundations from nearly a century ago, despite concerns. The following is four different matters with both support and cautious or skeptical viewpoints:

Matter	Supportive Viewpoint	Cautious Viewpoint
Transparency/Trust	Channa et al. (2021) demonstrate an intellectual tutoring system that works with learner affect, utilizing emotional data.	Schultz et al. (2024) caution that automated decision making may lack transparency and can result in performance degradation in real-world AI systems.
Anthropomorphism	Alabed et al. (2022) found that Human-like AI can build trust and engagement through self-AI integration, highlighting positive outcomes such as emotional support.	Walan (2024) states that primary school students express both excitement and anxiety, raising concerns about AI's autonomy.
Emotion Emulation	Channa et al. (2021) again demonstrate that emotional modeling increases adaptive support in an intelligent tutoring system.	Alabed et al. (2022) note that over-personalized AI may cause identity confusion or dependency.
Ethics	Walan (2024) argues that educating young people about AI helps them develop informed perspectives.	Schultz et al. (2024) note that the lack of explainability and control in these systems poses a threat to access and long-term decision integrity.

Future Research Directions

- Researchers should always consider the perspectives on artificial intelligence (AI). While AI can be beneficial depending on its application, there is a subset of the population that fears AI, particularly as it becomes more autonomous. This raises significant ethical concerns, making it crucial to establish design standards (Alabed et al., 2022; Walan, 2024). Furthermore, AI may be perceived as a “black box” (Schultz et al., 2024) by those who are not considered experts in AI.
- Although neural networks are undeniably important, as demonstrated by their functionality, attempting to replicate something as complex as the human brain is a challenge, as the human brain is rather complex. This task should involve research collaboration with neuroscience experts to advance both the development of neural networks and our understanding of them.
- Schultz et al. (2024) highlight an important issue for the future of machine learning: small errors in input data can significantly reduce the performance of algorithms. Therefore, future research should concentrate not only on optimizing these algorithms but also on ensuring transparency in their decision-making processes.
- Emotion emulation is a significant concern in artificial intelligence. Future research should focus on the psychological and social consequences of emotionally responsive AI agents (Alabed et al., 2022; Walan, 2024). Additionally, it is essential to create carefully designed ethical guidelines regarding the use of emotion in AI.

Conclusion & Synthesis

- Artificial intelligence is multifaceted and can be defined as “intelligence demonstrated by machines” (National University, 2025). It can take various forms, including:
    - Symbolic AI, which follows explicit rules and symbols.
    - Connectionist AI, which makes connections in a manner like the human brain, often in less transparent ways.
    - Evolutionary AI, which evolves and adapts over time.
    - Embodied AI, which exists in the form of physical robots. These different approaches highlight the diverse nature of artificial intelligence.
- Neural networks and machine learning provide tools for intelligent behavior, but encounter challenges in areas where transparency and data sensitivity are priorities, such as higher education and FERPA, as well as medical institutions under HIPAA.
- Emotion emulation in AI is progressing in different contexts but comes with challenges and risks.
- Anthropomorphism in AI design may lead to better user engagement and trust, but also raises a flag about dependency and individuality.
- There are contrasting viewpoints in the literature demonstrating the tension between different matters involving AI and its use.
- Future research should prioritize how to make AI more explainable or transparent, along with the development of ethical standards.

References

Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI anthropomorphism and its effect on users' self-congruence and self–AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change, 182*, 121786. https://doi.org/10.1016/j.techfore.2022.121786

Bizzarri, A., Jalaian, B., Riguzzi, F., & Bastian, N. D. (2024). A Neuro-Symbolic Artificial Intelligence Network Intrusion Detection System. *2024 33rd International Conference on Computer Communications and Networks (ICCCN)*, 1–9. https://doi.org/10.1109/ICCCN61486.2024.10637618

Canamero, L. (2021). Embodied robot models for interdisciplinary emotion research. *IEEE Transactions on Affective Computing, 12*(2), 340–351. https://doi.org/10.1109/TAFFC.2019.2908162

Channa, F. R., Sarhandi, P. S. A., Bugti, F., & Brohi, I. A. (2021). Supporting self-regulated learning by affect detection and responding in AI-driven learning systems. *Elementary Education Online, 20*(5), 3205–3211. http://ilkogretim-online.org

Dietrich, E. (2005). Artificial intelligence, philosophy of. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (1st ed.). Wiley. https://search.credoreference.com/articles/Qm9va0FydGljbGU6Mjg2Mjkw?aid=102577

References

Joseph, J. (2025, January 30). A gentle introduction to neural networks. *CleverTap*. https://clevertap.com/blog/neural-networks/

MirzagoltabarRoshan, M., AlizadeElizei, M., & Esmaeilabadi, R. (2022). Evolutionary artificial intelligence model to formulate compressive strength of eco-friendly concrete containing recycled polyethylene terephthalate. *Arabian Journal for Science & Engineering, 47*(10), 13229–13247. https://doi.org/10.1007/s13369-021-06432-7

National University. (2025). Artificial intelligence, neural networks, learning, and emotions. https://ncuone.ncu.edu/d2l/le/content/493832/viewContent/3619746/View

Neural network. (2016). In A. Burdett & D. Bowen (Eds.), *BCS Glossary of Computing and ICT* (14th ed.). BCS, The Chartered Institute for IT. https://search.credoreference.com/articles/Qm9va0FydGljbGU6NDU2MDU3Nw==?aid=102577

Reilly, E. D., Ralston, A., & Hemmendinger, D. (Eds.). (2003). Machine learning. In *Encyclopedia of Computer Science* (4th ed.). Wiley. https://search.credoreference.com/articles/Qm9va0FydGljbGU6MTY2NTU0Nw==?aid=102577

References

Schultz, C. D., Koch, C., & Olbrich, R. (2024). Dark sides of artificial intelligence: The dangers of automated decision-making in search engine advertising. *Journal of the Association for Information Science and Technology, 75*(5), 550–566. https://doi.org/10.1002/asi.24798

Stanford University students. (2003). Neural networks – Biology. *Stanford University*. https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Biology/index.html

Torabzadeh, S. A., Tavakkoli-Moghaddam, R., Samieinasab, M., & Hamid, M. (2022). An intelligent algorithm to evaluate and improve the performance of a home healthcare center considering trust indicators. *Computers in Biology and Medicine, 146*, 105656. https://doi.org/10.1016/j.compbiomed.2022.105656

Wajahat, I., Singh, A., Keshtkar, F., & Bukhari, S. A. C. (2024). A symbolic AI framework for enhanced diabetes prognosis accuracy and explainability. *2024 IEEE International Conference on Future Machine Learning and Data Science (FMLDS)*, 15–20. https://doi.org/10.1109/FMLDS63805.2024.00013

Walan, S. (2025). Primary school students' perceptions of artificial intelligence – for good or bad. *International Journal of Technology and Design Education, 35*(1), 25–40. https://doi.org/10.1007/s10798-024-09898-2