

A Concise Introduction to Machine Learning and Artificial Intelligence in Psychological Research

Matteo Orsoni
University of Bologna

This chapter introduces some of the basic principles of Machine Learning and Artificial Intelligence for readers who may not have a strong quantitative background but already know a little about statistics and probability. The goal is to make complex concepts more accessible, with the broader intent of democratizing the comprehension and application of these methods within various fields of psychology. Because this is such a broad and fast-moving area, the chapter cannot go into every detail. Rather, the goal is to provide a strong starting point that helps readers continue learning on their own, while also sparking interest, creativity, and innovative applications in psychology.

Introduction

The landscape of Artificial Intelligence (AI) reached the global public in November 2022 with the release of a conversational chatbot (OpenAI, 2022) based on a well-established deep learning architecture known as transformers (Vaswani et al., 2017). This architecture is grounded in the attention mechanism first proposed in 2014 by Bahdanau, Cho, and Bengio (2014), developed to address the limitations of Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997), which themselves had been introduced to overcome the shortcomings of Recurrent Neural Networks (RNNs) (Rumelhart et al., 1988), as part of the ongoing evolution of neural network architectures.

This brief introduction serves to highlight that modern Artificial Intelligence has roots going much further back than many might assume. Decades of research and development have paved the way for today's widespread access: anyone with an internet connection can now use algorithmic and computational innovations that once required years of specialized expertise. With this democratization, however, comes the responsibility to understand not only how these systems function

and what benefits they bring, but also their boundaries and limitations. It is crucial to recognize what AI and Machine Learning can realistically accomplish, and, just as importantly, what they cannot, so that their application in psychology and other fields remains both effective and responsible. With this in mind, we aim to provide readers with the foundations needed to better understand ML/AI, while also encouraging curiosity and creativity in applying these methods to their own psychological research.

The chapter will be organized into several sections.

In Section 18, we will define Machine Learning, Artificial Intelligence, and their main subfields.

Section 19 will explain why ML/AI methods are increasingly applied in psychology, examining their benefits and implications.

Section 20 will address the fundamental principles and methodological foundations needed to develop reliable, valid, consistent, and interpretable ML systems using R and Python. This section will also introduce some of the most relevant classical ML algorithms.

Finally, Section 21 will focus on the most important libraries and packages, as well as key manuals and scientific works that can support independent study and deeper exploration.

What

In recent years, the terms Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have often been used interchangeably. However, it is important to recognize the hierarchical relationship among these concepts. The goal of this section is to clarify this structure, beginning with individual definitions.

Artificial Intelligence (AI) is closely linked to the idea of human intelligence. As described by Russell and Norvig (2021), AI refers to the development of artificial agents or systems capable of carrying out tasks that are typically associated with humans. Such tasks include learning, perception, reasoning, and decision-making, as well as more concrete activities like playing chess, driving cars, writing poetry, or making medical diagnoses.

Machine Learning (ML), in contrast, focuses on algorithms that can identify patterns in data without explicit human programming. ML refers to methods and techniques designed to learn from data, either by representing it in new ways or by extracting meaningful information. ML methods can be grouped into several categories, the most common being *Supervised Learning*, *Unsupervised Learning*, and *Reinforcement Learning*. Other approaches, such as Semi-Supervised and Self-Supervised Learning, also exist but will not be addressed here. Widely used ML algorithms found in the literature include Classification and Regression Trees (CART) (Breiman, 1984), Random Forests (Breiman, 2001), Support Vector Machines (Cortes & Vapnik, 1995), Bayesian Networks (Koller &

Friedman, 2009), Artificial Neural Networks (ANNs), K-means clustering (MacQueen, 1967), Fuzzy C-means (Nayak et al., 2015b), and Gaussian Mixture Models (Reynolds., 2009), among others.

At the next level, **Deep Learning (DL)** can be seen as a subfield of Machine Learning that relies on Artificial Neural Networks as its core components. DL methods are designed to learn at multiple levels of abstraction, with each layer capturing increasingly complex patterns (Chen, 2015). Inspired by the structure of the human brain, deep learning has been successfully applied in areas such as computer vision, robotics, automation, and artificial intelligence more broadly (Polson & Sokolov, 2018). Whether supervised or unsupervised, DL techniques excel at discovering hidden structures in data and have achieved state-of-the-art performance in tasks like dimensionality reduction, object recognition, speech processing, and natural language understanding (Firdaus & Dixit, 2018; Orsoni, Giovagnoli, Garofalo, et al., 2025).

Supervised Machine Learning

Among ML methods, the most widely used are Supervised Learning algorithms (SML) (Lecun et al., 2015). In this setting, training data provide both inputs and their corresponding output labels, guiding the model to learn associations between them. Through this process, the algorithm uncovers patterns and decision rules, enabling it to generalize to new, unseen cases (Alloghani et al., 2020).

For example, consider building a model to distinguish between students with and without dyslexia. Cognitive data serve as input, while diagnoses by psychologists act as labels. The algorithm iteratively adjusts its internal parameters to minimize the difference between its predictions and the true diagnoses, improving its ability to classify students accurately.

The ultimate goal of SML is to predict or classify outcomes as reliably as possible. A central challenge is the **bias–variance tradeoff**: overly simple models (high bias) underfit the data, while overly complex ones (high variance) overfit. Effective models strike a balance, ensuring robust generalization.

Developing SML solutions begins with understanding the problem and data. The **preprocessing** phase is critical: exploratory data analysis (EDA) helps identify distributions, outliers, and missing values, which must be addressed to avoid bias. Features are often standardized or normalized, depending on the data type, and techniques such as **feature selection** (choosing the most informative variables) and **feature engineering** (creating new variables) improve model performance (Orrù et al., 2020).

Datasets are then split into **training, validation, and test sets** (commonly 70/10/20 in % over the total sample), though proportions may vary. Training data fit the model, validation data tune it, and test data evaluate its generalization. During training, careful model selection and monitoring for underfitting or

overfitting are crucial. Tools such as **cross-validation**, **regularization**, and **hyperparameter tuning** (via grid search, random search or bayesian optimization) help optimize performance. The final stage is **model evaluation** and **interpretability**. Performance metrics depend on whether the task is classification or regression type. Beyond accuracy, understanding how the model makes decisions is fundamental. Methods such as feature importance analysis, SHAP (Lundberg & Lee, 2017), and LIME (Ribeiro et al., 2016) help explain predictions, making results more transparent and actionable.

This overview outlines the key stages involved in implementing SML. In section 20 we will revisit these concepts and provide concise code snippets for each step. Given the inherent complexity, we also present a roadmap that summarizes the essential procedures for developing reliable SML systems independent to the context applications.

Unsupervised Machine Learning

Unsupervised learning models (UML), unlike supervised models, are trained on unlabeled examples and seek to uncover structure by grouping similar observations (Orrù et al., 2020). Two primary techniques dominate this setting: clustering and dimensionality reduction. Although they may appear similar at first glance, they serve distinct purposes. **Clustering** groups similar data points to reveal patterns or latent structure in the data, whereas **dimensionality reduction** projects the data into a lower-dimensional feature space to simplify analysis and visualization and, in many cases, to improve subsequent model performance (Parvathi et al., 2017). In some contexts, dimensionality reduction can also be viewed as a form of feature engineering (Orsoni, Giovagnoli, Garofalo, et al., 2025).

Classic techniques for linear and nonlinear dimensionality reduction include Principal Component Analysis (PCA) (Kurita, 2020) and t-Distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten & Hinton, 2008), which transform high-dimensional data into a lower-dimensional representation. Popular clustering algorithms include k-means (MacQueen, 1967), Fuzzy C-means (Nayak et al., 2015a), and Gaussian Mixture Models (Reynolds., 2009). Section 19 present these topics further by combining dimensionality reduction with clustering to discover underlying data structures.

Reinforcement Learning

This paragraph introduces the fundamental concepts of Reinforcement Learning. It's important to note, however, that implementation details won't be covered here or in later sections.

In Reinforcement Learning (RL), an agent learns to make optimal decisions by interacting with an environment over discrete time steps. At each step t , the agent

observes the environment's state S_t and, following a policy π , selects an action A_t . The environment then transitions to a new state S_{t+1} and provides a scalar reward R_{t+1} that reflects the immediate outcome of the action.

This agent–environment loop is formalized as a Markov Decision Process (MDP), which assumes the Markov property: the future is independent of the past given the present. In other words, the next state and reward depend only on the current state and action, not on the full history.

The agent's objective is to learn an *optimal policy* π^* that maximizes the cumulative discounted reward, or *expected return*. Future rewards are discounted by a factor $\gamma \in [0, 1]$, so that a reward received k steps in the future is worth γ^{k-1} times an immediate reward. A value of γ close to 1 makes the agent far-sighted, while γ close to 0 makes it short-sighted (Sutton & Barto, 2018).

The expected return is then strictly related to another fundamental concept in RL, that is the concept of value functions. Value functions estimate how good it is to be in a particular state or to take a particular action from a state under a given policy. This notion of “how good” represents the anticipation of future rewards that the agent can expect, according to the actions. Many RL algorithms employ a value function to assign utility to states and actions. There are two main value functions.

State-value function. The expected return when starting in state s and following policy π thereafter:

$$v_\pi(s) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s \right]. \quad (1)$$

Action-value function. The expected return when starting in state s , taking action a once, and then following policy π :

$$q_\pi(s, a) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s, A_t = a \right]. \quad (2)$$

The value functions satisfy recursive relationships that decompose value into immediate reward plus discounted future value. To derive the optimal policy, it is essential to articulate the connection between the value of a state and the values of its successor states, considering the expected cumulative reward. The Bellman Expectation Equation for the state-value and action-value functions compute this relationship

Bellman expectation (for q_π).

$$q_\pi(s, a) = R(s, a) + \gamma \sum_{s'} P(s' \mid s, a) \sum_{a'} \pi(a' \mid s') q_\pi(s', a'). \quad (3)$$

Bellman optimality (for q^*).

$$q^*(s, a) = R(s, a) + \gamma \sum_{s'} P(s' | s, a) \max_{a'} q^*(s', a'). \quad (4)$$

Many reinforcement learning algorithms, such as value iteration, seek the optimal action-value function q^* by repeatedly applying the Bellman optimality update. Once q^* is known, an optimal policy is obtained by acting greedily:

$$\pi^*(s) \in \arg \max_a q^*(s, a). \quad (5)$$

Over the years, a diverse landscape of RL algorithms has emerged. These algorithms can be categorized along several key dimensions, each reflecting a different approach to the problem of learning optimal behavior.

Model-Based vs. Model-Free A primary distinction lies in whether the agent learns a model of the environment's dynamics.

Model-Based RL. These algorithms first learn a model of the environment, which includes the transition function $P(s' | s, a)$ and the reward function $R(s, a)$. The agent then uses this learned model for planning, often by simulating potential future trajectories to find the best course of action. Monte Carlo Tree Search (MCTS) (Świechowski et al., 2023) is a prominent example, where a tree of possible futures is built and evaluated to guide decisions.

Model-Free RL. These algorithms learn a policy or value function directly from experience without creating an explicit model of the environment. They learn *what* to do, not necessarily *how* the environment works. This is a more direct, trial-and-error approach. Q-learning (Watkins & Dayan, 1992) is a classic model-free algorithm that learns the optimal action-value function q^* from interactions, allowing it to choose the best action in a given state without needing to know the transition probabilities.

Value-Based vs. Policy-Based vs. Actor-Critic Methods Algorithms also differ in what they learn and optimize.

Value-Based Methods. These algorithms focus on learning an optimal value function, such as the action-value function $q^*(s, a)$. The policy is then derived implicitly by always selecting the action with the highest value in a given state (a greedy policy). Q-learning is a canonical value-based method.

Policy-Based Methods. These algorithms directly parameterize and optimize the policy $\pi(a | s)$ itself, without necessarily learning a value function. They adjust the policy's parameters to maximize the expected return. This approach is particularly effective in continuous or high-dimensional action spaces. Examples include REINFORCE and Proximal Policy Optimization (PPO) (Sutton & Barto, 2018).

Actor-Critic Methods. These hybrid methods combine the strengths of both value-based and policy-based approaches. They consist of two components: an *actor*, which is a policy that decides which action to take, and a *critic*, which is a value function that evaluates the actions taken by the actor. The critic's feedback is then used to “teach” the actor how to improve its policy. Prominent examples include Asynchronous Advantage Actor-Critic (A3C) and Deep Deterministic Policy Gradient (DDPG) (Dong et al., 2020).

On-Policy vs. Off-Policy Learning Another key difference is how the agent uses the data it collects.

On-Policy Learning. The agent learns from actions taken according to its current policy. It improves the same policy that it uses to make decisions. This can be thought of as “learning on the job”. SARSA is a classic on-policy algorithm (Sutton & Barto, 2018).

Off-Policy Learning. The agent can learn from data generated by a different policy. For instance, it can learn the optimal policy while executing a more exploratory (e.g., random) policy. This allows for greater data efficiency, as experiences can be stored in a replay buffer and reused. Q-learning is a well-known off-policy algorithm.

Online vs. Offline RL Finally, algorithms can be differentiated by when they learn.

Online RL. The agent learns incrementally and updates its knowledge after each interaction with the environment. This allows it to adapt in real-time to dynamic or changing environments.

Batch RL (or Offline RL). The agent learns from a fixed, finite dataset of previously collected experiences, without any further interaction with the environment. The goal is to learn the best possible policy from this static batch of data.

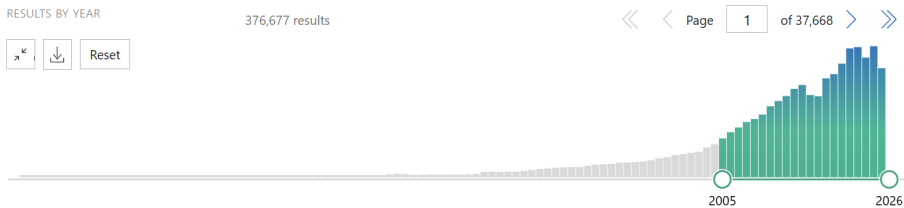
These classifications provide a framework for understanding the vast array of RL algorithms. The field has been further revolutionized by Deep Reinforcement Learning (DRL), which integrates deep neural networks as powerful function approximators for representing complex policies and value functions, enabling RL to solve problems of previously intractable scale and complexity (Sutton & Barto, 2018).

Why

In this section, we will examine why Artificial Intelligence (AI) and Machine Learning (ML) are becoming increasingly important in psychology. Some readers may already have considered potential research questions or application areas where

these tools could be applied, for example, to gain deeper insights into psychological processes or to open new avenues of investigation.

The figure 7 illustrates this trend by showing the number of PubMed articles indexed with the keywords “Artificial Intelligence,” “Machine Learning,” and “Psychology.” The rise has been nearly exponential: in the past two decades, publications in this area have grown dramatically. This rapid expansion underscores the need for students and researchers to build a solid understanding of AI and ML methods, as they are now a common feature in the literature and play an increasingly central role in the design, analysis, and interpretation of psychological research.



7. Publications between 2005 and 2026 on Artificial Intelligence, Machine Learning, and Psychology, retrieved from PubMed using the search query: (((Artificial Intelligence) OR (Machine Learning)) AND (Psychology)) OR (Psychological Research)).

Social Psychology

Machine learning and in recent years generative artificial intelligence (GenAI), are increasingly transforming research practices in social psychology by enabling scalable, theory-driven analyses of human interaction and behavior (Bail, 2024; Salah et al., 2023). Advances in large language models (LLMs) allow researchers to simulate social exchanges, ranging from everyday conversations to complex group deliberations, thereby offering controlled environments for examining social influence, opinion change, persuasion, and conformity over time (Bail, 2024). When embedded within agent-based models, these systems can empower synthetic agents with personalities, memories, and adaptive behaviors, making it possible to reproduce phenomena such as social movement dynamics, organizational decision-making, and competitive interactions (Bail, 2024). GenAI also can enhance traditional research methodologies by supporting the creation of synthetic survey respondents, the production of ecologically valid experimental stimuli, and even the replication of classic experiments in behavioral science (Hoey et al., 2018). In parallel, machine learning techniques in natural language processing (NLP) can automate large-scale content analysis, enabling rapid and accurate classification of sentiment, emotional expression, topic framing, and ideological positioning across vast corpora of social media or political communication (Hoey et al., 2018). These capabilities address long-standing challenges in studying group processes, particularly the

difficulty of analyzing high-volume interaction data (“the scaling problem”) and modeling complex, nonlinear social dynamics (“the dynamics problem”) (Hoey et al., 2018). Emerging frameworks such as Bayesian Affect Control Theory and its BayesAct implementation integrate social psychological principles into probabilistic AI agents that reason about identity, emotion, and uncertainty (Catellani et al., 2021). Moreover, reinforcement learning and dynamic Bayesian models are being leveraged to design adaptive, personalized strategies for behavior change, such as promoting physical activity through tailored interaction policies (Catellani et al., 2021).

Work and Organizational Psychology

The use of ML and AI is reshaping work and organisational psychology. These are transforming both core research methodologies and the full spectrum of human resource management functions (Pereira et al., 2023). In organizational science, AI-driven analytic techniques enhance predictive accuracy for critical outcomes such as turnover, safety incidents, career trajectories, and unethical behavior, surpassing the limits of traditional explanatory models (Woo et al., 2024). Advances in natural language processing, deep learning, and multimodal analytics are enabling the extraction of psychological constructs from previously inaccessible data sources, including video recordings, linguistic traces, and passive sensor data, thereby refining measurement of personality, attitudes, emotions, and dynamic behavioral processes (Woo et al., 2024). Continuous data streams from wearable technologies and mobile sensing further allow researchers to examine temporal constructs such as affective fluctuations, performance rhythms, and human–machine interactions with unprecedented granularity (Woo et al., 2024). These methodological innovations also support inductive theorizing and computational grounded theory, opening pathways for discovering emergent patterns and nonlinear relationships that conventional statistical approaches often overlook (Oswald et al., 2020). Within organizational practice, AI and ML increasingly drive key Human Resource Management (HRM) functions, from recruitment and selection to training, performance management, and workforce planning (Pereira et al., 2023). In talent acquisition, NLP-based screening tools automate resume analysis, identify suitable candidates across large information ecosystems, and evaluate applicant behavior through video analytics that assess vocal and facial cues. These systems can reduce human bias by adhering to standardized decision criteria (Ganatra & Pandya, 2023). In employee development, machine learning enables tailored learning pathways, VR- and AR-enabled simulations, and intelligent knowledge management systems that adapt training content to evolving skill profiles (Baki et al., 2023). Performance management is similarly augmented through real-time monitoring systems that generate continuous, data-driven assessments, highlight performance risks, and support evidence-based appraisal decisions (Pereira et al., 2023). AI-

enhanced communication tools such as chatbots improve employee engagement by facilitating HR support and enabling large-scale sentiment analysis to detect dissatisfaction or emerging relational issues (Ganatra & Pandya, 2023). Workforce optimization techniques, including genetic algorithms, data-driven team formation, and automated scheduling, help organizations align talent with dynamic operational demands (Ganatra & Pandya, 2023). In parallel, machine learning contributes to occupational health and safety through predictive modeling of hazards, optimization of job rotation schedules, and the design of adaptive work environments that support employee well-being (Pereira et al., 2023).

Cognitive Psychology, Neuropsychology, and Neuroscience

ML and AI, particularly DL, have become central to contemporary research in cognitive psychology, neuropsychology, and neuroscience, where they support both foundational scientific investigation and highly specialized clinical applications (Burgos et al., 2021). In neuroscience, ML models increasingly function as theoretical tools, offering computational accounts of learning, representation, and decision-making that inspire hypotheses about neural mechanisms (Hinton, 2011). For example, reinforcement learning models of reward prediction error have shaped theories of dopaminergic signaling (Vu et al., 2018). The conceptual parallels between deep learning systems and neural information processing have further encouraged the development of unified explanatory frameworks, including explainable AI models designed to integrate prediction, interpretation, and biological plausibility (Vu et al., 2018). On the technical front, AI enables the automation of complex neuroimaging workflows: deep neural networks reconstruct high-resolution MRI and PET images from undersampled data, perform denoising or cross-modality synthesis, and support robust registration and segmentation of anatomical and pathological structures through architectures such as U-Net (Burgos et al., 2021; Vu et al., 2018; Zhang et al., 2020). These methods also facilitate biomarker discovery, whether through extracting structural or functional metrics, estimating “brain age” as a general marker of neuropathology, or using neural network models to detect genetic variants associated with conditions such as autism (Burgos et al., 2021; Noor et al., 2020).

In cognitive and clinical neuropsychology, AI is widely adopted for disease classification, prognosis, and refinement of assessment practices (Orsoni, Giovagnoli, Garofalo, et al., 2023, 2025). Deep learning methods applied to neuroimaging and physiological data reliably distinguish individuals with Alzheimer’s disease, Parkinson’s disease, or schizophrenia from healthy controls, while more sophisticated architectures, such as generative models, graph convolutional networks, and recurrent networks, enable subtype differentiation and early detection of mild cognitive impairment or dopamine-deficit disorders (Burgos et al., 2021; Noor et al., 2020; Zhang et al., 2020). Predictive models are also used to forecast disease pro-

gression, identify high-risk patients, and anticipate treatment response, thereby supporting personalized intervention strategies (Burgos et al., 2021; Noor et al., 2020). Beyond diagnosis, ML enhances cognitive assessment by classifying neuropsychological functioning, reducing testing burdens, and improving discrimination among clinically overlapping conditions (Orsoni, Giovagnoli, Garofalo, et al., 2025; Ying et al., 2024). Importantly, AI techniques also uncover latent heterogeneity within clinical populations by clustering patients into biologically or functionally distinct subtypes, such as alternative progression pathways in Parkinson's disease or connectivity-based profiles of depression (Burgos et al., 2021; Orsoni, Giovagnoli, Garofalo, et al., 2025; Vu et al., 2018).

Psychometrics

ML and AI are reshaping psychometric methodology through innovations in different branches of researches, for example in extracting and measuring psychological dimensions (Ahmad et al., 2020), in the individual classification and diagnostic assessment (Orsoni, Giovagnoli, Garofalo, et al., 2025), in enhancing and automating psychometric procedures (Liu et al., 2024; Orsoni, Benassi, & Scutari, 2025), and in assessing and profiling AI systems (AI Psychometrics) (Benosman, 2025). Regarding the computerized adaptive testing (CAT), reinforcement learning, meta-learning, and active learning are progressively used to optimize item selection, increase measurement precision, and reduce test burden (Liu et al., 2024). Deep cognitive diagnosis models and ML-based question bank analytics further strengthen test fairness, robustness, and interpretability (Pellert et al., 2024). Related techniques treat scale reduction as a feature-selection challenge, using genetic algorithms or decision-tree structures to generate efficient short-form assessments that preserve predictive validity (Orsoni, Benassi, & Scutari, 2025). Finally, psychometric tools are increasingly applied to AI systems themselves: large language models are evaluated using adapted psychological inventories and bias assessments to quantify implicit associations, moral orientations, or personality-like response patterns (Benosman, 2025). Such "AI psychometrics" not only reveal the latent biases embedded in AI models but also provide a principled framework for monitoring and shaping their psychological profiles (Benosman, 2025).

Educational Psychology

In the broad field of educational psychology, ML and AI play an increasingly central role by advancing assessment, personalization, cognitive skill development, and data-driven analysis of learning processes (Benvenuti et al., 2023). A major contribution lies in evaluating and predicting student performance (Alruwais & Zakariah, 2023). ML models can classify learners by knowledge level, forecast academic outcomes, detect at-risk students, and identify psychological and behavioral variables, such as motivation, self-control, or personality, that shape learning

trajectories (Alruwais & Zakariah, 2023; Benvenuti et al., 2023; de Chiusole et al., 2024). AI also automates complex scoring tasks, from essay evaluation to conceptual problem solving, with convolutional neural networks enabling intelligent tutoring systems to assess open-ended responses (Caspari-Sadeghi, 2022). These capabilities support the development of adaptive instructional systems, where AI-driven tutors model learners' psychological states and tailor instruction accordingly (Caspari-Sadeghi, 2022; Pardos & Heffernan, 2010). Knowledge tracing frameworks such as Bayesian Knowledge Tracing (BKT) and Deep Knowledge Tracing (DKT) use probabilistic models and recurrent neural networks to track students' mastery of skills over time, while individual-specific parameters enhance precision in estimating learning progress (Pögelt et al., 2024). Reinforcement learning further drives dynamic difficulty adjustment and personalized recommendation systems by optimizing task sequences based on students' goals, performance, and estimated proficiency, sometimes in combination with more standard psychometric systems such as the item response theory (IRT) (Orsoni, Pögelt, et al., 2023).

Beyond assessment and personalization, AI fosters the development of higher-order cognitive and psychological skills (Benvenuti et al., 2023). Intelligent cognitive tutors provide metacognitive feedback that strengthens planning and problem solving, and AI systems, including robotic tutors, support creativity, curiosity, and motivation by adapting interactions to students' developmental needs (Benvenuti et al., 2023). In areas such as computational creativity, AI tools can even augment creative output, particularly when integrated with immersive technologies like virtual reality (Benvenuti et al., 2023). Finally, educational data mining and learning analytics leverage ML to handle the vast, multimodal datasets produced in digital learning environments (Jiao et al., 2021). These methods enable fine-grained analysis of process data, including log files, response times, action sequences, and eye-tracking patterns to uncover the strategies students use during problem solving and to refine estimates of ability or cognitive profiles. Coupled with affective computing approaches that employ sensors and wearables to infer engagement, emotion, and motivation, these analytic tools provide a comprehensive, data-rich framework for understanding and optimizing learning (Alruwais & Zakariah, 2023; Caspari-Sadeghi, 2022).

Clical Psychology and Psychotherapy

ML and AI are reshaping clinical psychology, and psychotherapy by enhancing diagnostic precision, personalizing treatment, optimizing therapeutic processes, and enabling continuous monitoring (Althoff et al., 2016; Olawade et al., 2024). In diagnostic contexts, AI systems analyze large-scale clinical datasets, including electronic health records, clinician notes, and unstructured text, to classify mental health conditions and predict risks such as suicide or treatment nonresponse (Althoff et al., 2016; Chekroud et al., 2021; Olawade et al., 2024; Zhou et al., 2022).

Natural language processing supports the detection of symptoms in clinical documentation and the analysis of text-based counseling interactions, while social media monitoring provides additional insights into emotional states and mood fluctuations (Chekroud et al., 2021; Graham et al., 2019). Beyond linguistic data, deep learning methods applied to neuroimaging and electrophysiology identified neural markers of disorders such as depression, bipolar disorder, and schizophrenia, and sensor-based technologies use smartphone, video, and wearable data to infer behavioral patterns and physiological indicators relevant to mental health (Althoff et al., 2016; Graham et al., 2019). These diagnostic and predictive capabilities also inform personalized treatment planning, where machine learning models identify which individuals are most likely to respond to specific psychotherapies or medications, derive symptom-based subtypes, and integrate genetic information to guide pharmacological decisions (Spytska, 2025).

AI-driven therapeutic tools further expand access to and delivery of psychological care (Althoff et al., 2016; Olawade et al., 2024; Spytska, 2025). Digital mental health interventions such as chatbots and virtual therapists provide scalable platforms for cognitive-behavioral therapy, mindfulness training, and crisis support, while advanced systems employing computer vision and speech analysis can detect subtle markers of distress during virtual consultations (Spytska, 2025). Virtual reality-based treatments enhance exposure therapy and skills training, and AI-supported robotics have shown promise in assisting social development in populations such as children with autism. In traditional psychotherapy settings, AI augments the work of human clinicians by analyzing therapeutic dialogue, identifying evidence-based conversational strategies, and modeling session progression (Zhou et al., 2022). Tools offering real-time emotional analytics, automated administrative support, and data-informed suggestions for interventions can improve both clinician efficiency and therapeutic quality, and routine outcome monitoring powered by ML provides individualized feedback that enhances treatment efficacy (Olawade et al., 2024).

How

In this paragraph we are going to present an application of classical SML to a possible psychological case study.

The dataset has been retrieved from a recent open-access repository, namely: **NeuroDataSets** accessible through the following Github page: <https://github.com/cracan/NeuroDataSets> (Caceres Rossi, 2025).

As a didactic scope of this application, we will follow the 20.1, however for limnited space is not possible to present an application of UML; yet the interested reader can foud a roadmap for this kind of investigation here: 20.1. For a didactive

perspective, the code of the analysis can be found at the following OSF repository: https://osf.io/9y5ew/overview?view_only=8ef744f212404bb1bbf3cf854cbc39a7. Indeed, for brevity, we are not going to attach in the main text all the script, but only the most relevant parts.

Roadmaps for SML and UML

A Roadmap of Supervised Machine Learning

1. Problem Understanding

- Define the research question.
- Identify input features (X) and target variable (Y).

2. Data Preprocessing

- Explore data: distributions, outliers, missing values.
- Clean data: handle missing values, correct errors.
- Standardize/normalize features for consistency.
- Feature selection: keep the most informative variables.
- Feature engineering: create new variables if useful.
- Split dataset: Training (70%) – Validation (10%) – Test (20%).

3. Model Training

- Choose an algorithm (e.g., Decision Trees, SVM, Neural Networks).
- Train model on training set.
- Tune complexity to balance bias vs. variance.
- Apply cross-validation and regularization to prevent overfitting.
- Adjust hyperparameters (grid search, random search, bayesian optimization).

4. Model Evaluation

- Use validation/test sets to check generalization.
- Choose metrics based on the task (e.g., F1, accuracy, precision, RMSE).

5. Interpretability

- Go beyond accuracy: understand *why* the model makes its predictions.
- Use tools like:
 - Feature importance
 - SHAP (SHapley values)
 - LIME (Local explanations)

Goal: Build a model that is accurate, generalizable, and interpretable.

Roadmap of Unsupervised Machine Learning

1. Problem Understanding

- Define the task (clustering, dimensionality reduction).
 - Set success criteria: internal validity (e.g., silhouette), external validity, stability across resamples.
2. **Data Preprocessing**
 - Explore and clean data; encode categoricals; scale features.
 - Use resampling/hold-out splits for unbiased validation and stability checks.
 3. **Algorithm and Metric Choice**
 - Select appropriate methods (e.g., k-means/GMM/hierarchical/DBSCAN; PCA/UMAP/t-SNE).
 - Choose a similarity/distance metric consistent with the data (e.g., Euclidean, cosine).
 4. **Training and Tuning**
 - Tune key hyperparameters (e.g., number of clusters, density thresholds, embedding dimension).
 - Use internal indices (Silhouette, Calinski–Harabasz, Davies–Bouldin) and stability-based selection across seeds/resamples.
 5. **Evaluation and Utility**
 - Assess separation/compactness and solution stability; if labels exist, use ARI (Adjusted Rank Index)/NMI (Normalized Mutual Information) as external checks.
 - Verify usefulness (e.g., improved simple supervised models or domain KPIs).
 6. **Interpretability and Reporting**
 - Summarize clusters (centroids/medoids, distinguishing features) and visualize (embeddings, heatmaps, dendrograms).
 - Document assumptions, limitations, and key findings.
 7. **Operationalization and Monitoring**
 - Build a reproducible pipeline.
 - Monitor data/cluster drift and anomaly rates, address fairness/privacy.

Goal: Discover structure that is stable, valid, interpretable, and useful.

A real case scenario

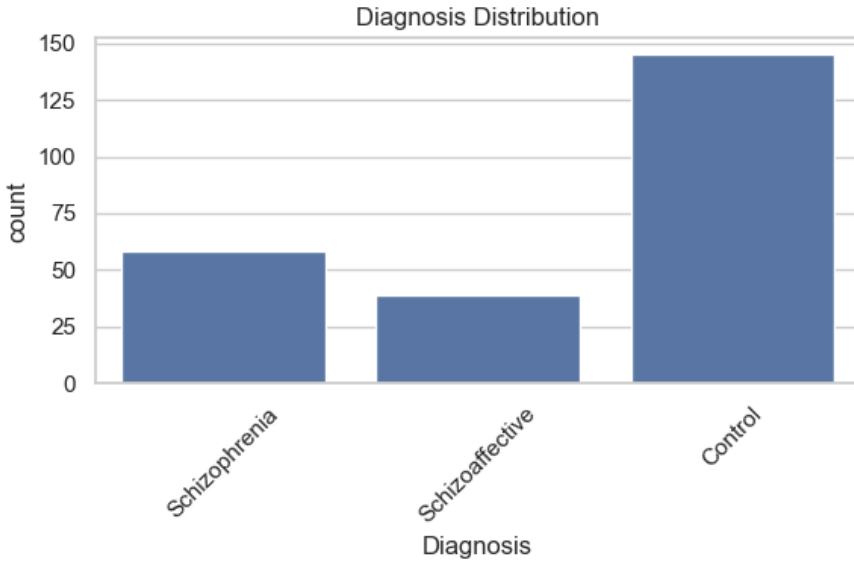
In this example, we applied a SML workflow to a clinical-cognitive dataset including individuals with schizophrenia, schizoaffective disorder, and healthy controls. Each participant was described by eight cognitive/behavioral scores (e.g., processing speed, attention, memory, verbal and visual abilities, problem solving, social cognition, age) plus sex. In figure 8 is depicted the distribution of the outcome variable (“Diagnosis”), and in figure 9 the feature values according to the diagnosis type.

```

1
2 df = pd.read_csv("neurodataset_book_chapter.csv", header=0, sep
   =";")
3
4 print(df.head())
5 print(df.info())
6 print(df["Diagnosis"].value_counts())

```

Listing 1: Dataset loading and info



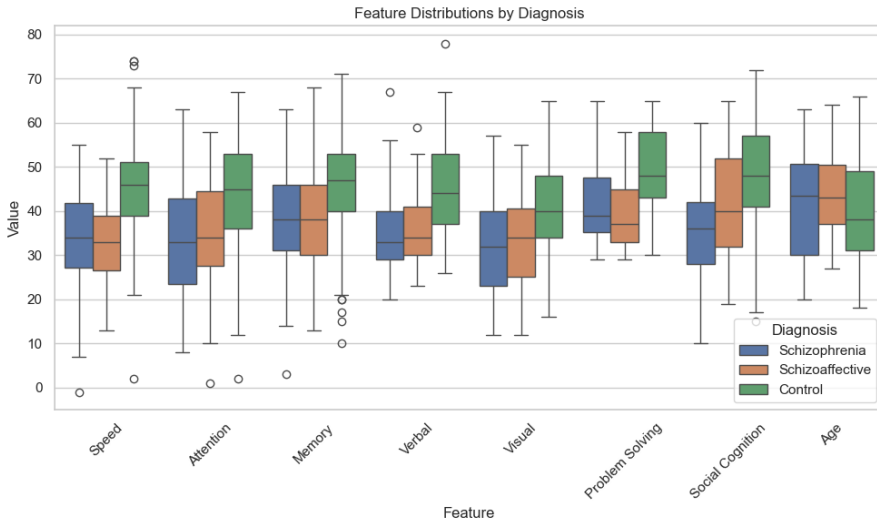
8. Distribution of the outcome variable (Diagnosis)

After basic cleaning (treating an out-of-range value as missing and imputing numerical variables with the mean) and one-hot encoding of sex (categorical variable), we standardized all predictors and split the data into training, validation, and test sets (approximately 70%, 10%, and 20% of the sample, while preserving class proportions).

```

1
2 imputer_num = SimpleImputer(strategy="mean")
3
4 ohe = OneHotEncoder(drop="first", sparse_output=False) # drop
   ='first' to avoid collinearity
5
6 X = X_full.copy()
7
8 # First split: train_val (80%) e test (20%)
9 X_train_val, X_test, y_train_val, y_test = train_test_split(
10 X, y, test_size=0.20, stratify=y, random_state=42

```



9. Boxplots of the feature distributions according to the diagnosis

```

11 )
12
13 # Second split: train (70%) e val (10%) su train_val (80% del
14 # 70/80 = 0.875 train_val will be the train set
15 X_train, X_val, y_train, y_val = train_test_split(
16     X_train_val, y_train_val,
17     test_size=0.125,           # 0.125 * 0.8 = 0.10 del totale
18     stratify=y_train_val,
19     random_state=42
20 )
21
22 print("Dimensions:")
23 print("Train:", X_train.shape, "Val:", X_val.shape, "Test:", X_
24     test.shape)
25
26 scaler = StandardScaler()
27 X_train_scaled = scaler.fit_transform(X_train)
28 X_val_scaled   = scaler.transform(X_val)
29 X_test_scaled  = scaler.transform(X_test)

```

Listing 2: Imputation, One-Hot encoding, Train-Validation-Test splitting, and Standardization

We then trained a Random Forest (RF) classifier and used cross-validated grid search to tune its hyperparameters, optimizing macro-averaged F1 to account for the imbalance between the majority Control group and the smaller schizophrenia and schizoaffective groups.

RF is a non parametric, *ensemble learning*, which combines the predictions of multiple individual models to achieve a more accurate and stable result than any single model could provide. The “forest” is composed of a multitude of Decision Trees. The robustness of the final model comes from two key mechanisms that ensure these trees are diverse and uncorrelated:

- Bootstrap Aggregation (Bagging): For each new tree, a random sample of the training data is drawn with replacement. Some original data points may appear multiple times in one tree’s training set, while others are left out (known as Out-of-Bag (OOB) samples, which can be used for internal validation).
- Random Feature Selection

The final prediction from the Random Forest is determined by combining the outputs of all its individual trees. In classification tasks, the model usually use the majority voting approach, that is the most frequent aggregated output derived from the individual trees. In regression tesks, the model takes the average of the continuous predictions made by all the individual trees.

```

1
2 # Base Model
3 rf = RandomForestClassifier(
4     n_estimators=200,
5     max_depth=None,
6     random_state=42,
7     class_weight="balanced" # for small imbalance
8 )
9
10 rf.fit(X_train_scaled, y_train)
11
12
13 # Cross-Validation and Hyperparameter Tuning
14
15 param_grid = {
16     "n_estimators": [200, 400, 600],
17     "max_depth": [None, 5, 10, 15],
18     "min_samples_split": [2, 6, 10, 20],
19     "min_samples_leaf": [1, 2, 4, 6]
20 }
21
22 cv = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)
23
24 rf_cv = RandomForestClassifier(random_state=42, class_weight="
    balanced")
25
26 grid_search = GridSearchCV(
27     estimator=rf_cv,
28     param_grid=param_grid,

```

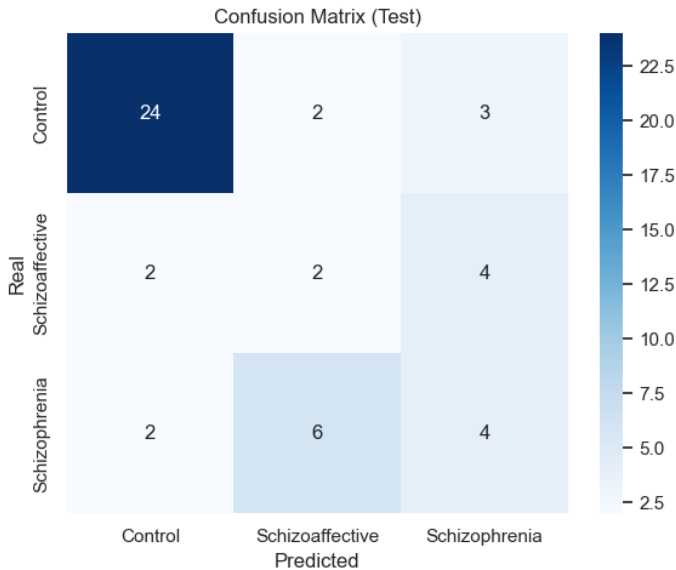
```

29     cv=cv,
30     scoring="f1_macro", # For multi-class classification
31     n_jobs=-1,
32     verbose=1
33 )
34
35 grid_search.fit(X_train_scaled, y_train)

```

Listing 3: Random Forest implementation and hyperparameter search

The final model achieved an accuracy of about 0.61 on the held-out test set, with a macro F1 score of approximately 0.47. Performance was clearly asymmetric across classes: the model classified Control participants relatively well (precision ≈ 0.86 , recall ≈ 0.83 , F1 ≈ 0.84), but performance dropped markedly for Schizophrenia (precision ≈ 0.36 , recall ≈ 0.33 , F1 ≈ 0.35) and especially for Schizoaffective disorder (precision ≈ 0.20 , recall ≈ 0.25 , F1 ≈ 0.22). The confusion matrix (figure 10) provide an overview of the classification accuracy of the winner model.



10. Confusion Matrix of the best CV RF model.

This pattern shows two important points for applied work: first, overall accuracy can be misleading when classes are imbalanced as we can see in figure 8; second, even reasonably sophisticated models may struggle to discriminate between closely related diagnostic categories when their feature distributions overlap substantially, as we can see in figure 9.

```

1
2 best_rf = grid_search.best_estimator_
3
4 X_train_val_scaled = scaler.fit_transform(X_train_val) # fit
   over the full train_val sample
5 X_test_scaled_final = scaler.transform(X_test)
6 best_rf.fit(X_train_val_scaled, y_train_val)
7
8 y_test_pred = best_rf.predict(X_test_scaled_final)
9
10 print("=== PERFORMANCE OVER TEST SET ===")
11 print(classification_report(y_test, y_test_pred))
12
13 acc = accuracy_score(y_test, y_test_pred)
14 f1m = f1_score(y_test, y_test_pred, average="macro")
15 prec = precision_score(y_test, y_test_pred, average="macro")
16 rec = recall_score(y_test, y_test_pred, average="macro")
17
18 print(f"Accuracy: {acc:.3f}")
19 print(f"F1-macro: {f1m:.3f}")
20 print(f"Precision-macro: {prec:.3f}")
21 print(f"Recall-macro: {rec:.3f}")

```

The last point regards the interpretability of the results. Beyond global performance, we examined which features the model relied on using both traditional feature importance and SHAP (Lundberg & Lee, 2017). RF feature importance (figure 11), suggested that processing speed, problem solving, age, visual abilities, and social cognition were the most informative predictors, whereas memory and sex contributed relatively little to the classification.

```

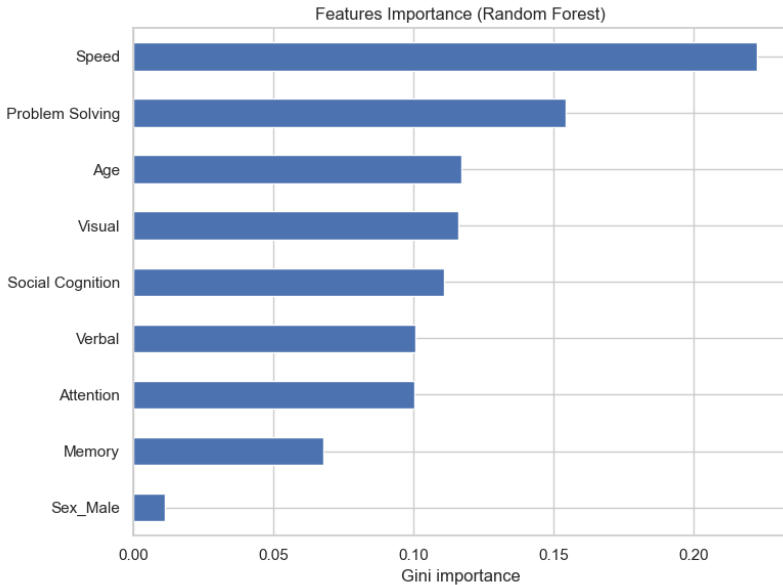
1
2 importances = best_rf.feature_importances_
3 feat_names = X_train_val.columns # X_train_val_scaled column
   names
4
5 feat_importances = pd.Series(importances, index=feat_names).
   sort_values(ascending=False)
6 print(feat_importances)

```

```

1
2 X_train_val_sample = X_train_val.sample(n=n_background, random_
   state=42)
3 X_train_val_sample_scaled = scaler.transform(X_train_val_sample
   )
4 X_shap = pd.DataFrame(X_train_val_sample_scaled, columns=feat_
   names)
5 print("Shape X_shap:", X_shap.shape)
6
7 # SHAP for RandomForest: TreeExplainer
8 explainer = shap.TreeExplainer(best_rf)

```



11. Features importance derived by the best RF model.

```

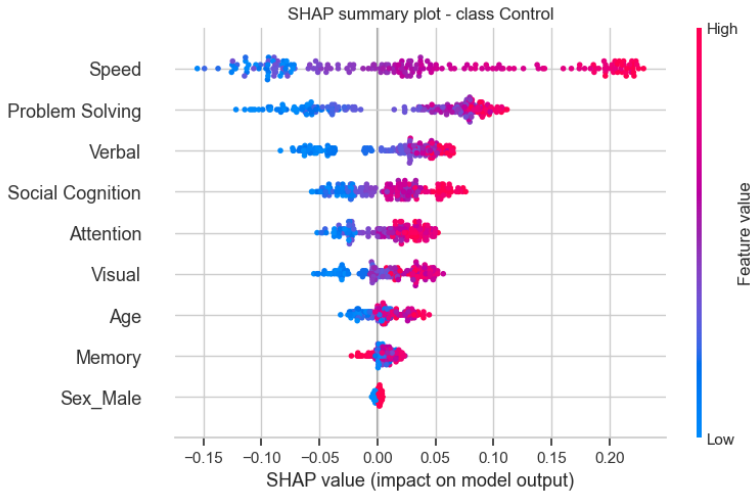
9
10 shap_values = explainer.shap_values(X_shap)
11 print("Type of shap_values:", type(shap_values))

```

SHAP summary plots (figures 12, 13, 14) provided a more nuanced, class-specific picture: for the control class, high values across several cognitive domains, especially high speed scores and high problem solving abilities, pushed predictions towards this group, consistent with the idea that globally better performance is more typical of healthy participants. For schizophrenia and schizoaffective disorder, different combinations of features became influential. In the schizophrenia sample, greater social cognition scores are associated with a decreased likelihood of the model classifying individuals into this diagnostic category. This inverse relationship also holds true for the cognitive functions of processing speed, verbal reasoning, problem-solving, and attention, although the magnitude of their influence on classification varies. For the schizoaffective disorder sample, the pattern is largely comparable to the schizophrenia group. However, processing speed and problem-solving abilities emerge as the most influential variables in determining the model's classification output for this group.

Taken together, this small case study shows an important contribution of the full supervised machine learning pipeline for psychological science; from the data preprocessing steps, to the model training and tuning, un-

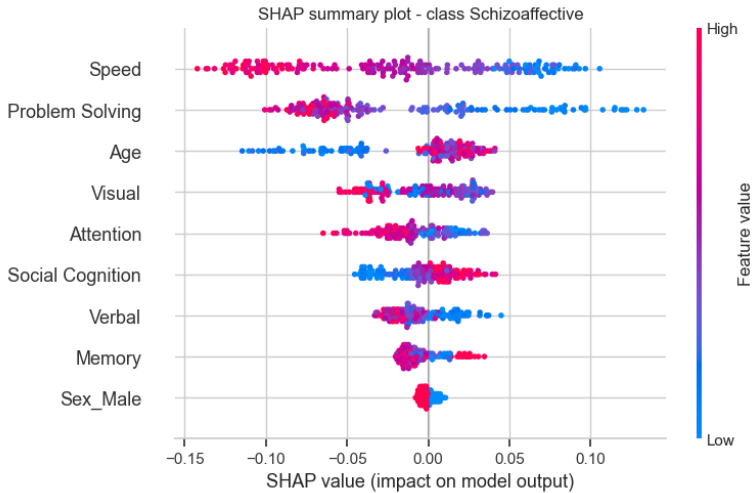
til the rigorous evaluation, and post-hoc interpretability. Beyond merely creating predictive models (e.g., classifying diagnostic categories), this systematic approach could provide a powerful tool to generate psychologically meaningful and testable hypotheses, and allowing to describe complex, and non-linear phenomena.



12. SHAP values of the control sample.



13. SHAP values of the schizophrenic sample



14. SHAP values of the schizoaffective sample.

Where

By the fact that ML applications are divided into classical ML, and Deep Learning (DL) algorithms, the following libraries are the well-documented standards in their respective domains. Python is the industry and academic standard for developing and deploying complex DL architectures, which build fundamentally upon data structures provided by libraries like NumPy and Pandas. Python's `Scikit-learn` (Pedregosa et al., 2011) library <https://scikit-learn.org/stable/index.html> library remains the essential toolkit for robust classical ML tasks (SML and UML) such as predicting clinical outcomes or clustering participant responses. Conversely, R maintains its dominance in traditional statistical modeling, offering powerful packages like the `caret` (Kuhn, 2008), <https://topepo.github.io/caret/index.html> and `tidymodels` (Kuhn & Wickham, 2020), <https://www.tidymodels.org/> ecosystem. These R tools are highly valued by statisticians for their emphasis on rigorous model fit, interpretability, and robust analysis workflows in classical ML settings. However, recognizing the growth of DL, recent years have seen the development of mature R interfaces for Deep Learning, allowing researchers to leverage powerful frameworks like Keras without leaving the R environment. Tables 24, and 25 summarize some of the most used, and useful libraries and packages for ML and DL implementations in R and python.

Moreover, the emergence of Transformer architectures (Vaswani et al., 2017) and Large Language Models (LLMs) represents a paradigm shift from

| Library | Primary Focus | SML/UML/DL Applications |
|-------------------------------|--------------------------|---|
| Scikit-learn (sklearn) | Classical ML (SML & UML) | The all-in-one tool for regression, classification (SML), and clustering, feature selection, and dimensionality reduction (UML). Excellent for baseline models and statistical ML. |
| TensorFlow / Keras | Deep Learning | Core frameworks for developing complex neural networks (DL). Keras provides a high-level API over TensorFlow, simplifying the construction of Convolutional, and Recurrent Neural Networks. |
| PyTorch | Deep Learning / Research | Known for its dynamic computational graph, preferred for advanced DL research, computer vision, and Natural Language Processing (NLP). Offers greater flexibility in model design. |
| NumPy & Pandas | Data Handling | Essential for array manipulation (NumPy) and efficient tabular data cleaning and preprocessing (Pandas)—necessary groundwork for all ML algorithms. |

24. Essential Python Libraries for Machine Learning and Deep Learning

building task-specific models to fine-tuning massive, pre-trained models. The Hugging Face Transformers (Wolf et al., 2020) <https://huggingface.co/docs/transformers/installation> library, in Python is the central ecosystem for accessing and deploying state-of-the-art models (like BERT, GPT, and Llama) for tasks such as sentiment analysis, automatic coding of thematic content, and even simulating behavioral or cognitive biases in model-based agents.

| Package | Primary Focus | SML/UML/DL Applications |
|-------------------|--|--|
| caret | Classification And REgression Training | Provides a unified interface to train, tune, and evaluate nearly 200 different ML algorithms, making comparative modeling straightforward. |
| tidymodels | ML with Tidyverse Principles | A modern collection of packages (e.g., <i>parsnip</i> , <i>recipes</i>) that structures the ML workflow in a consistent, interpretable way, highly favored for reproducible analysis. |
| xgboost | Gradient Boosting | An implementation of Extreme Gradient Boosting, offering one of the most powerful and fastest algorithms for highly accurate classification and regression on structured data. |
| e1071 | Statistical/ML Algorithms | Implements Support Vector Machines (SVM), Naive Bayes, and includes tools for K-Means clustering, often used for exploratory or classical multivariate modeling. |
| keras | Deep Learning | Provides a complete R interface to the Python TensorFlow/Keras library, allowing users to build complex neural networks without leaving the R environment. |

25. Essential R Packages for Machine Learning and Deep Learning

References

- Ahmad, F., Abbasi, A., Li, J., Dobolyi, D., Netemeyer, R., Clifford, G., & Chen, H. (2020). A deep learning architecture for psychometric natural language processing. *ACM Transactions on Information Systems (TOIS)*, 38(2), 1–29. <https://doi.org/10.1145/3365211>
- Alruwais, N., & Zakariah, M. (2023). Evaluating student knowledge assessment using machine learning techniques. *Sustainability*, 15, 6229. <https://doi.org/10.3390/su15076229>
- Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics*, 4, 463–476. https://doi.org/10.1162/tacl_a_00111
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. <https://arxiv.org/abs/1409.0473>

- Bail, C. A. (2024). Can generative ai improve social science? *Proceedings of the National Academy of Sciences of the United States of America*, 121(21), e2314021121. <https://doi.org/10.1073/pnas.2314021121>
- Baki, N., Mohd Rasdi, R., Krauss, S., & Omar, M. (2023). Integrating artificial intelligence in human resource functions: Challenges and opportunities. *International Journal of Academic Research in Business and Social Sciences*, 13(8), 18071. <https://doi.org/10.6007/IJARBS/v13-i8/18071>
- Benosman, M. (2025). Designing psychometric bias measures for chatbots: An application to racial bias measurement. <https://doi.org/10.48550/arXiv.2509.13324>
- Benvenuti, M., Cangelosi, A., Weinberger, A., Mazzoni, E., Benassi, M., Barbaresi, M., & Orsoni, M. (2023). Artificial intelligence and human behavioral development: A perspective on new skills and competences acquisition for the educational context. *Computers in Human Behavior*, 148, 107903. <https://doi.org/10.1016/j.chb.2023.107903>
- Breiman, L. (1984). *Classification and regression trees* (1st ed.). Routledge. <https://doi.org/10.1201/9781315139470>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Burgos, N., Bottani, S., Faouzi, J., Thibeau-Sutre, E., & Colliot, O. (2021). Deep learning for brain disorders: From data processing to disease treatment. *Briefings in Bioinformatics*, 22(2), 1560–1576. <https://doi.org/10.1093/bib/bbaa310>
- Caceres Rossi, R. (2025). *Neurodatasets: A comprehensive collection of neuroscience and brain-related datasets* [R package version 0.3.0]. <https://github.com/lightbluetitan/neurodatasets>
- Caspari-Sadeghi, S. (2022). Artificial intelligence in technology-enhanced assessment: A survey of machine learning. *Journal of Educational Technology Systems*, 004723952211387. <https://doi.org/10.1177/00472395221138791>
- Catellani, P., Carfora, V., & Piastra, M. (2021). Connecting social psychology and deep reinforcement learning: A probabilistic predictor on the intention to do home-based physical activity after message exposure. *Frontiers in Psychology*, 12, 696770. <https://doi.org/10.3389/fpsyg.2021.696770>
- Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, 20(2), 154–170. <https://doi.org/10.1002/wps.20882>

- Chen, C. L. P. (2015). Deep learning for pattern learning and recognition. *2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics*, 17. <https://api.semanticscholar.org/CorpusID:6265469>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- de Chiusole, D., Spinoso, M., Anselmi, P., Bacherini, A., Balboni, G., Mazzoni, N., Brancaccio, A., Epifania, O. M., Orsoni, M., Giovagnoli, S., Garofalo, S., Benassi, M., Robusto, E., Stefanutti, L., & Pierluigi, I. (2024). Psycassist: A web-based artificial intelligence system designed for adaptive neuropsychological assessment and training. *Brain Sciences*, 14(2), 122. <https://doi.org/10.3390/brainsci14020122>
- Dong, H., Ding, Z., & Zhang, S. (Eds.). (2020). *Deep reinforcement learning: Fundamentals, research and applications* (1st ed.). Springer. <https://doi.org/10.1007/978-981-15-4095-0>
- Firdaus, N., & Dixit, M. (2018). *Deep learning technique*. <https://api.semanticscholar.org/CorpusID:212495631>
- Ganatra, N., & Pandya, J. (2023). The transformative impact of artificial intelligence on hr practices and employee experience: A review. *Journal of Management Research and Analysis*, 10, 106–111. <https://doi.org/10.18231/j.jmra.2023.018>
- Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports*, 21(11), 116. <https://doi.org/10.1007/s11920-019-1094-0>
- Hinton, G. E. (2011). Machine learning for neuroscience. *Neural Systems & Circuits*, 1(1), 12. <https://doi.org/10.1186/2042-1001-1-12>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hoey, J., Schröder, T., Morgan, J., Rogers, K. B., Rishi, D., & Nagappan, M. (2018). Artificial intelligence and social simulation: Studying group dynamics on a massive scale. *Small Group Research*, 49(6), 647–683. <https://doi.org/10.1177/1046496418802362>
- Jiao, H., He, Q., & Veldkamp, B. P. (2021). Editorial: Process data in educational and psychological measurement. *Frontiers in Psychology*, 12, 793399. <https://doi.org/10.3389/fpsyg.2021.793399>
- Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. MIT Press. <https://books.google.it/books?id=7dzpHCHzNQ4C>

- Kuhn, M. (2008). Building predictive models in r using the caret package. *Journal of Statistical Software*, 28(5), 1–26. <https://doi.org/10.18637/jss.v028.i05>
- Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles*. Retrieved December 30, 2025, from <https://www.tidymodels.org>
- Kurita, T. (2020). Principal component analysis (pca). In *Computer vision*. Springer. https://doi.org/10.1007/978-3-030-03243-2_649-1
- Liu, Q., Zhuang, Y., Bi, H., Huang, Z., Huang, W., Li, J., Yu, J., Liu, Z., Hu, Z., Hong, Y., Pardos, Z. A., Ma, H., Zhu, M., Wang, S., & Chen, E. (2024). Survey of computerized adaptive testing: A machine learning perspective. <https://arxiv.org/abs/2404.00712>
- Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. <https://arxiv.org/abs/1705.07874>
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th berkeley symposium on mathematical statistics and probability* (pp. 281–297, Vol. 1). University of California Press. https://dигicoll.lib.berkeley.edu/record/113015/files/math_s5_v1_article-17.pdf
- Nayak, J., Naik, B., & Behera, H. S. (2015a). Fuzzy c-means (fcm) clustering algorithm: A decade review from 2000 to 2014. In L. Jain, H. Behera, J. Mandal, & D. Mohapatra (Eds.), *Computational intelligence in data mining – volume 2* (Vol. 32). Springer. https://doi.org/10.1007/978-81-322-2208-8_14
- Nayak, J., Naik, B., & Behera, H. S. (2015b). Fuzzy c-means (fcm) clustering algorithm: A decade review from 2000 to 2014. In L. C. Jain, H. S. Behera, J. K. Mandal, & D. P. Mohapatra (Eds.), *Computational intelligence in data mining - volume 2* (pp. 133–149). Springer India.
- Noor, M. B. T., Zenia, N. Z., Kaiser, M. S., Al Mamun, S. A., & Mahmud, M. (2020). Application of deep learning in detecting neurological disorders from magnetic resonance images: A survey on the detection of alzheimer’s disease, parkinson’s disease and schizophrenia. *Brain Informatics*, 7(1), 11. <https://doi.org/10.1186/s40708-020-00112-2>
- Olawade, D. B., Wada, O. Z., Odetayo, A., David-Olawade, A. C., Asaolu, F., & Eberhardt, J. (2024). Enhancing mental health with artificial intelligence: Current trends and future prospects. *Journal of Medicine, Surgery, and Public Health*, 3, 100099. <https://doi.org/10.1016/j.glmedi.2024.100099>
- OpenAI. (2022). *Chatgpt: Optimizing language models for dialogue*. Retrieved December 30, 2025, from <https://openai.com/blog/>

- Orrù, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2020). Machine learning in psychometrics and psychological research. *Frontiers in Psychology, 10*, 2970. <https://doi.org/10.3389/fpsyg.2019.02970>
- Orsoni, M., Giovagnoli, S., Garofalo, S., Magri, S., Benvenuti, M., Mazzoni, E., & Benassi, M. (2023). Preliminary evidence on machine learning approaches for clusterizing students' cognitive profile. *Heliyon, 9*(3), e14506. <https://doi.org/10.1016/j.heliyon.2023.e14506>
- Orsoni, M., Giovagnoli, S., Garofalo, S., Mazzoni, N., Spinoso, M., & Benassi, M. (2025). Comparing factor mixture modeling and conditional gaussian mixture variational autoencoders for cognitive profile clustering. *Frontiers in Psychology, 16*, 1474292. <https://doi.org/10.3389/fpsyg.2025.1474292>
- Orsoni, M., Pögel, A., Duong-Trung, N., Benassi, M., Kravcik, M., & Grützmüller, M. (2023). Recommending mathematical tasks based on reinforcement learning and item response theory. In C. Frasson, P. Mylonas, & C. Troussas (Eds.), *Augmented intelligence and intelligent tutoring systems* (Vol. 13891). Springer. https://doi.org/10.1007/978-3-031-32883-1_2
- Orsoni, M., Benassi, M., & Scutari, M. (2025). Information theory, machine learning, and bayesian networks in the analysis of dichotomous and likert responses for questionnaire psychometric validation [Advance online publication]. *Psychological Methods*. <https://doi.org/10.1037/met0000713>
- Oswald, F. L., Behrend, T. S., Putka, D. J., & Sinar, E. (2020). Big data in industrial-organizational psychology and human resource management: Forward progress for organizational research and practice. *Annual Review of Organizational Psychology and Organizational Behavior, 7*, 505–533. <https://doi.org/10.1146/annurev-orgpsych-032117-104553>
- Pardos, Z. A., & Heffernan, N. T. (2010). Modeling individualization in a bayesian networks implementation of knowledge tracing. In P. De Bra, A. Kobsa, & D. Chin (Eds.), *User modeling, adaptation, and personalization* (pp. 255–266, Vol. 6075). Springer. https://doi.org/10.1007/978-3-642-13470-8_24
- Parvathi, R., Pavithra, M. S., & Parvathi, R. M. S. (2017). *Cluster ensemble approach for semi supervised clustering*. <https://api.semanticscholar.org/CorpusID:249602997>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research, 12*, 2825–2830.
- Pellert, M., Lechner, C. M., Wagner, C., Rammstedt, B., & Strohmaier, M. (2024). Ai psychometrics: Assessing the psychological profiles of large language models

- through psychometric inventories. *Perspectives on Psychological Science*, 19(5), 808–826. <https://doi.org/10.1177/17456916231214460>
- Pereira, V., Hadjilias, E., Christofi, M., & Vrontis, D. (2023). A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective. *Human Resource Management Review*, 33(1), 100857. <https://doi.org/10.1016/j.hrmr.2021.100857>
- Pögelt, A., Ihsberner, K., Pengel, N., Kravcik, M., Grützmüller, M., & Hardt, W. (2024). Individualised mathematical task recommendations through intended learning outcomes and reinforcement learning. In A. Sifaleras & F. Lin (Eds.), *Generative intelligence and intelligent tutoring systems* (Vol. 14798). Springer. https://doi.org/10.1007/978-3-031-63028-6_10
- Polson, N. G., & Sokolov, V. O. (2018). Deep learning. <https://arxiv.org/abs/1807.00000>
- Reynolds. (2009). Gaussian mixture models. In S. Z. Li & A. Jain (Eds.), *Encyclopedia of biometrics* (pp. 659–663). Springer. https://doi.org/10.1007/978-0-387-73003-5_196
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “why should i trust you?”: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning internal representations by error propagation. In A. Collins & E. E. Smith (Eds.), *Readings in cognitive science* (pp. 399–421). Morgan Kaufmann. <https://doi.org/10.1016/B978-1-4832-1446-7.50035-2>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Salah, M., Al Halbusi, H., & Abdelfattah, F. (2023). May the force of text data analysis be with you: Unleashing the power of generative ai for social psychology research. *Computers in Human Behavior: Artificial Humans*, 1(2), 100006. <https://doi.org/10.1016/j.chbah.2023.100006>
- Spytska, L. (2025). The use of artificial intelligence in psychotherapy: Development of intelligent therapeutic systems. *BMC Psychology*, 13(1), 175. <https://doi.org/10.1186/s40359-025-02491-9>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.
- Świechowski, M., Godlewski, K., Sawicki, B., et al. (2023). Monte carlo tree search: A review of recent modifications and applications. *Artificial Intelligence Review*, 56, 2497–2562. <https://doi.org/10.1007/s10462-022-10228-y>

- van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-sne [Nov]. *Journal of Machine Learning Research*, *9*, 2579–2605.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. <https://doi.org/10.48550/arXiv.1706.03762>
- Vu, M.-A. T., Adalı, T., Ba, D., Buzsáki, G., Carlson, D., Heller, K., Liston, C., Rudin, C., Sohal, V. S., Widge, A. S., Mayberg, H. S., Sapiro, G., & Dzirasa, K. (2018). A shared vision for machine learning in neuroscience. *Journal of Neuroscience*, *38*(7), 1601–1607. <https://doi.org/10.1523/JNEUROSCI.0508-17.2018>
- Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*, 279–292. <https://doi.org/10.1007/BF00992698>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., ... Rush, A. M. (2020). Huggingface's transformers: State-of-the-art natural language processing. <http://doi.org/10.48550/arXiv.1910.03771>
- Woo, S. E., Tay, L., & Oswald, F. L. (2024). Artificial intelligence, machine learning, and big data: Improvements to the science of people at work and applications to practice. *Personnel Psychology*, *77*(4), 1387–1402. <https://doi.org/10.1111/peps.12643>
- Ying, H., Pranolo, A., Nuryana, Z., & Syafitri, A. I. (2024). Emerging trends in the evolution of neuropsychology and artificial intelligence: A comprehensive analysis. *Telematics and Informatics Reports*, *16*, 100171. <https://doi.org/10.1016/j.teler.2024.100171>
- Zhang, L., Wang, M., Liu, M., & Zhang, D. (2020). A survey on deep learning for neuroimaging-based brain disorder analysis. *Frontiers in Neuroscience*, *14*, 779. <https://doi.org/10.3389/fnins.2020.00779>
- Zhou, S., Zhao, J., & Zhang, L. (2022). Application of artificial intelligence on psychological interventions and diagnosis: An overview. *Frontiers in Psychiatry*, *13*, 811665. <https://doi.org/10.3389/fpsyt.2022.811665>