Machine Learning and applied Artificial Intelligence in cognitive science and psychology: a tutorial.

Caroline König[1] and Alfredo Vellido[1] *

1- Computer Science Department and Intelligent Data Science
and Artificial Intelligence (IDEAI-UPC) Research Center
Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract. Artificial Intelligence (AI) both in general and in its current predominant version, mostly based on connectionist tenets, lives in the paradox of aiming to reproduce and simulate the workings of an immensely complex system, the biological brain, which are still to a large extent unknown. This gives us latitude for some interesting domain interplay: concepts from the cognitive sciences can be used to improve AI models, while AI can be used in data science mode to analyze cognitive processes in neuroscience, as well as brain pathologies from a medical standpoint.

1 Introduction

Artificial Intelligence (AI) both in general and in its current predominant version, which is mostly based on connectionist tenets (that is, Machine Learning (ML) and, more in particular Deep Learning (DL) in the form of large artificial neural networks (ANN)) lives in the paradox of aiming to reproduce and simulate the workings of an immensely complex system, the biological brain, which are still to a large extent unknown. The ongoing groundbreaking advances in large language models (LLMs) with their mastery of human-like dialogue and other Transformer-based models (such as the AlphaFold variants that led to the recent concession of the Nobel Prize in Chemistry to Demis Hassabis and John Jumper [1]) have reignited the debate about the feasibility of Artificial General Intelligence (AGI) as an AI with human-level cognitive abilities.

The faith in the eventual advent of AGI should at the very least be tempered by our limited knowledge of biological intelligence. The contrast is stark when assessing the outcomes of recently finished neuroscience flagship efforts such as the European Human Brain Project (HBP) [2],[3] and the US BRAIN Initiative [4][5] and other similar efforts around the world [6]. The decade-long HBP began in 2013 with the stated general objective of understanding the human brain through computer simulation, from the molecular level all the way to the operational systems level. It quickly run into managerial problems requiring a complete overhaul [7]. Its overall goal was not achieved, although partial impressive results included a full and accessible map of the brain: the Julich Atlas [8].

*This research was partially funded by grant ALECTORS-2023 from the Polytechnic University of Catalonia (UPC) and the Spanish PID2022-143299OB-I00/AEI/10.13039/501100011033/ FEDER, UE, research project.

This paradoxical situation gives us latitude for some interesting scientific domain interplay: on the one hand, specific and limited-on-the-scope concepts from neuroscience and the cognitive sciences can be used as inspiration for the design and improvement of different AI models, even if we do not mean to claim their biological plausibility as *in silico* models of intelligence. On the other hand, AI can be used in data science mode (and mostly using ML-based methods at the core of analytical pipelines) to investigate cognitive processes in neuroscience, as well as brain pathologies from a medical standpoint.

This brief tutorial paper reflects on all these issues, without attempting completeness, and collects the contributions to the ESANN 2025 special session on *Machine Learning and applied Artificial Intelligence in cognitive sciences and psychology.*

2 Cognitive science concepts as foundations of AI systems

The interplay between the developments in AI and cognitive neuroscience, often mediated by the concept of computational neuroscience, is hardly ever fluid, as exemplified by DL. As a connectionist apotheosis of sorts, you might expect it to draw heavily on ideas of bio-plausibility as, for instance, Self-Organizing Maps did from the eighties [9] (see some comments on this below). Instead, the literature on this issue is rather thin on the ground. Elements of the integration of DL and Neuroscience were provided by Marblestone and co-workers [10], who work with three hypotheses, namely that the biological brain operates by optimizing diverse and dynamic cost functions; that cost functions associated to this optimization vary across brain regions; and that the brain utilizes pre-structured architectures optimized for specific behavioral and computational tasks. Authors claim that standard ML algorithms for optimization like backpropagation may have, at least partially, some correspondences in biological brains and they stress the importance of determining "whether and how brains implement these algorithms". Indeed, the bio-plausibility of backpropation has been a historically contentious matter, but one that has of late been successfully defended [11]. The use of DL as a "framework of neuroscience" has recently been advocated in [12], arguing that the same building blocks investigated for DL design: objective functions, learning rules and model architectures, would benefit Systems Neuroscience Research.

Several cognitive science concepts have inspired the development of AI systems. Arguably, one of the most successful DL models available today in different variations is the Transformer [13], and one of its key components is a mechanism of attention, namely self-attention. Attention mechanisms (AM) in AI have been investigated for over forty years and draw inspiration from biological AM [14], instantiated, for instance, by visual attention mechanisms [15] in sensory-motor loops, such as eye movements used to focus on the relevant part of an image while disregarding irrelevant information. AM leverages the concept of relevance by enabling the model to selectively concentrate on specific parts of the input, thereby improving the effectiveness of task performance. Over the years, various types of AM inspired by the human visual system have been developed and integrated into neural networks.

The self-attention mechanism has yielded impressive advances in applications such as computer vision [16],[17] and natural language processing [18], [19], among other areas of application. Two prominent self-attention-based models, namely BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) [20] implement different cognitive-like approaches to solve NLP tasks. BERT follows a bidirectional design to capture the context of a word in the sequence. BERT is trained as a masked language modeling task and next-section prediction, so that they are optimal for tasks related to a deep understanding of the text, such as text classification and question answering [21]. Meanwhile, the GPT architecture is trained with a one-directional design (left-to-right) and as a language modeling task to predict the next word in a sequence, therefore being optimal for text generation. Transformer models based on self-attention are widely used for other generative tasks such as images [22] and music [23], just to name a couple of areas.

Another important concept of cognitive science that has inspired computational models, such as associative memory, are the topographically structured maps [24] in the cortex responsible for transmitting visual signals from the retina to the brain. Inspired by the observation that associative areas are created through self-organization from ongoing learning, much research has aimed to emulate the mechanism of self-organization in neural networks through Hebbian learning [25]. Some examples include early attempts corresponding to the "Theory of cerebellar cortex" [26], and the "Non holographic Associative Memory" [27].

Visual processing in the human occipital lobe has been a constant source of inspiration for connectionist models in ML. Half a century ago, Malsburg and Willshaw emulated self-organization in the retina-cortex mapping through self-organizing maps [28],[29]. Arguably, though, Kohonen's Self-organizing Maps (SOM) [9] are, by far, the most popular self-organizing ML architecture. They incorporated a very useful feature, namely discrete topographically ordered projections of feature mappings (in ideas that can be traced back to Sun Ichi Amari's proposal of topographic organization of nerve fields [30]) learned through self-organization. In the same way the visual cortex uses feature maps to capture the information of orientation and spatial frequency from visual stimuli [31], SOMs learn feature maps that preserve the topological characteristics of the original input [32], resulting in a useful form of low-dimensional data representation in tasks of pattern recognition and data exploration for high-dimensional datasets in unsupervised learning [33]. This low-dimensional representation is a heuristic approach akin to that of probabilistic latent models such as Generative Topographic Mapping (GTM) [34] with discrete constrained latent space representations that keep the model within the connectionist framework. SOM-based large-scale simulations of the visual cortex such as the different variants of LISSOM at the core of the *Topographica* software were created by Bednar, Mikkulainen and co-workers from the 1990s [35].

In the field of behavioral neuroscience, the human decision making process has been described through the paradigm of Pavlovian conditioned behavior [36]. From a computational point of view, conditioned behavior was formally defined through the reinforcement learning (RL) normative framework [37]. Individual's decisions are oriented in a long term to maximize reward and minimize punishments simultaneously [38]. RL is based on temporal difference (TD) learning, where the objective is the prediction of the future reward and a temporal difference error, which is the difference between the predicted reward and the actual reward. This would be the key reinforcement signal to guide learning, which would have a parallel in the brain, the so-called *temporal difference reward prediction error*. This links RL to the function of dopaminergic neurons in the mammalian midbrain. Phasic firing of dopaminergic neurons may reflect a reward prediction error [39]. According to [40], dopamine acts as a signal for reward prediction errors and it can be utilized for forecasting and for action learning in dopaminergic targets. Furthermore, physiological evidence for dopamine-dependent (or even dopamine-gated) plasticity in the synapses between the cortex and the striatum has been found [41].

3 Advances in the application of ML to cognitive science and brain pathologies

3.1 ML and cognitive science

Brain modeling is a thriving field of research at the crossroads of several disciplines, such as neuroscience, image and signal processing, and data science, with the overall aim of creating a theoretical framework to understand the structure and functioning of the brain. This tutorial does, by no means, attempt to cover the many instances of ML use in cognitive science; instead, it aims to provide an illustrative example of its potential. The Human Connectome Project (HCP) [42], is a large-scale endeavor in cognitive science: a joint transnational effort to accelerate advances in human neuroimaging toward the creation of a comprehensive map of the human brain. The connectome is a graph-based representation of the organization of the brain, where its specific regions, at different levels, are treated as nodes and their multiple interactions as edges [43]. Different modalities of non-invasive neuroimaging techniques, including structural Magnetic Resonance Imaging (MRI), resting-state functional MRI (rfMRI), task-evoked functional MRI (tfMRI), and diffusion imaging (dMRI) allow acquiring the valuable information about structural and functional features of the brain [44] that is required to build the graphs that are the foundations of such connectomes, either at the anatomical or functional level [45].

Graph theory provides a framework to investigate the topological organization of these large-scale brain networks [46], in terms of their organization in small-world networks [47], modular structure [48], or the presence of highly connected nodes relevant for general intelligence [49].

Different types of manifold learning can also be used to efficiently analyze the high-dimensional connectivity matrices of brain models [50]. Neural manifold learning exploits the advantages of modeling the spatio-temporal patterns of neural activity in lower-dimensional subspaces of the original high-dimensional space of the brain model, using dimensionality reduction techniques [51]. This method has been extensively used in neuroscience research to gain insight into the neural mechanisms involved in decision making [52], spatial navigation [53], and movement execution [54].

3.2 ML for the analysis of brain pathologies

The data-based analysis of brain pathologies has long been an area of interest for Computational Intelligence and ML. Its reach has been thoroughly reviewed elsewhere [55]. For this reason, again, we aim here only to provide a few examples of its relevance.

An area of obvious clinical interest is neuro-oncology: the analysis of cancers of the brain, often from data acquired with imaging techniques [56]. Brain cancers are, in any case, only a sub-family of the many diseases and disorders of the brain that have been analyzed using AI-based approaches, which also include, among others, Alzheimer's disease, mild cognitive impairment, schizophrenia, depressive disorders, Parkinson's disease (for a thorough review on ML for diagnosis and prognosis of Parkinson's disease using brain imaging, see [57]), attention-deficit hyperactivity disorder, autism spectrum disease, epilepsy, multiple sclerosis, stroke, and traumatic brain injury [58]. A perspective on brain diseases analysis centered on DL can be found in [59]. Going back to the analysis of the human connectome discussed in the previous subsection, many studies have investigated brain connectivity patterns related to mental disorders with healthy and disease model states, such as for depression [60], schizophrenia [61], or autism spectrum disorder [62], to name a few.

In the investigation of brain diseases from an ML-based modelling perspective, guidelines are required to bridge the gap between data and their preprocessing and clinical routine [63]. For that to happen, a key issue is data harmonization in multi-center studies [64].

4 Contributions to the ESANN 2025 special session

The 33^{th} European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025) hosted a special session on "Machine learning and applied Artificial Intelligence in cognitive sciences and psychology" that included a total of eight studies. This section summarizes their contributions in light of the previous sections.

Most of the contributions to the session addressed, directly or in a related manner, the use of ML methods for data analysis as a way to investigate human cognitive pathologies. They are, therefore, related to the contents of Section 3.

Ben Yahia *et al.*'s work [65] investigates the broad and complex Autism Spectrum Disorder, a neurodevelopmental pathology with a difficult diagnosis due to significant individual variability and the absence of clearly described biomarkers.

Despite the fact that authors address diagnosis as a simplified binary classification problem (occurrence or absence of the disorder), they introduce several interesting elements in the analysis: First, multi-modal data are used, combining fMRI information and phenotypic data, and including a feature selection process. Second, and acknowledging the fact that relying on DL black-box methods for classification would dramatically limit the interpretability of the results and, consequently, their practical medical application, the study resorts to explainable AI (xAI) to enhance model transparency in the form of SHAP [66] to provide the type of insights into the model decision-making process that may assist a medical expert.

Dislexia, yet another complex and not fully understood disorder related to learning processes, is investigated in [67]. Acknowledging such complexity, this work focuses on studying the relative impact of key attributes on its diagnosis. In particular, it focuses on the demographic variables utilized in a computer-based linguistic game for dislexia screening. The analysis "highlights the heterogeneity present in these variables" and provides insights for the development of future ML-based approaches to dislexia screening.

Dementia can be characterized as a syndrome that is often associated with neurodegenerative diseases (often, for instance, with Alzheimer's disease). It is characterized by a global decline in cognitive abilities. Kumpik and colleagues [68] undertake the complex task of analyzing unstructured conversations for the early detection of dementia, which is required for palliative medical treatment. This work explores "whether cross-domain (from semi-structured to unstructured) transfer learning improves dementia classification from conversational speech". A BERT-family model is fine-tuned using semi-structured narratives for which contextual information was available, with further fine-tuning on naturalistic conversations. Authors find that direct transfer from BERT to conversations is more effective in improving generalization.

Although not directly an investigation about a cognitive pathology, the characterization of sleep stages can be a door to the study of sleep disorders. Moctezuma *et al.* [69] investigate sleep staging using Gradient Boosting. They do so without the requirement of extensive computational resources or high density electroencephalograms (EEG). A key to this study is the use of data transformations using the Discrete Wavelet Transform and Power Spectral Density. This method is shown to achieve competitive performance when compared to more complex DL-based methods, even with fewer subjects in the training set and the use of low-density EEG combined with electrooculograms.

Two of the contributions to the session make use of human cognition processes as inspiration for ML developments and they are, consequently, aligned with the scope of Section 2 of this tutorial.

The cognitive process, in this case a "functional emotion", put forward by Schöfer *et al.* [70] might seem unlikely at first. Authors propose a regulatory role for *boredom* in AI models. Being an emotion increasingly recognized to play a role in regulating human behavior and learning processes, this study transfers the regulatory function of boredom into a control loop modeled with spiking neural networks. Preliminary evidence of the ability of the model to replicate the regulatory mechanism of boredom is provided for synthetic data, thus setting the bases for the future development of self-regulating systems based on spiking neural networks "capable of entering a state of boredom".

The work of Guiducci *et al.* [71] comes almost from the other side of the spectrum of cognitive processes: authors use *intrinsic motivation*, related to the psychological concept of curiosity, as a generator of agent-driven rewards for exploratory behavior, thus aiming to enhance the biological plausibility of the AI model. In this study, intrinsic motivation is added to the Elastic Decision Transformer (EDT) framework for offline RL through an auxiliary intrinsic loss, to "enhance representation learning without altering fixed reward signals". The validity of this *curiosity-driven* approach is illustrated using several locomotion tasks from an open-source benchmark for offline RL.

Finally, this session includes a paper that does not quite fit neither the remit of Section 2, nor that of Section 3. Marquise *et al.* [72] investigate human behavior in a spatial navigation task involving locomotion and gaze dynamics of human subjects that was reproduced in a virtual environment. The data generated in such navigation task is used to train a CNN to reproduce human decisions, and the CNN inner workings are revealed using several xAI methods. This leads to the discovery of an specific oculomotor marker leading the behavioral strategy used by human participants in the navigation task.

References

[1] J. Jumper, R. Evans, A. Pritzel, et al. Highly accurate protein structure prediction with Alphafold. *Nature*, 596(7873):583–589, 2021.

[2] H. Markram, K. Meier, T. Lippert, et al. Introducing the human brain project. *Procedia Computer Science*, 7:39–42, 2011.

[3] I.E. Bjerke, M. Øvsthus, E.A. Papp, et al. Data integration through brain atlasing: Human brain project tools and strategies. *European Psychiatry*, 50:70–76, 2018.

[4] J.R. Ecker, D.H. Geschwind, A.R. Kriegstein, et al. The brain initiative cell census consortium: lessons learned toward generating a comprehensive brain cell atlas. *Neuron*, 96(3):542–557, 2017.

[5] L.A. Jorgenson, W.T. Newsome, D.J. Anderson, et al. The brain initiative: developing technology to catalyse neuroscience discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668):20140164, 2015.

[6] R. Yuste and C. Bargmann. Toward a global brain initiative. *Cell*, 168(6):956–959, 2017.

[7] N. Siva. What happened to the human brain project? *The Lancet*, 402(10411):1408–1409, 2023.

[8] K. Amunts, H. Mohlberg, S. Bludau, and K. Zilles. Julich-brain: A 3d probabilistic atlas of the human brain's cytoarchitecture. *Science*, 369(6506):988–992, 2020.

[9] T. Kohonen. Self-organizing maps, ser. *Information Sciences. Berlin: Springer*, 30, 2001.

[10] A.H. Marblestone, G. Wayne, and K.P. Kording. Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, 10:215943, 2016.

[11] T.P. Lillicrap, A. Santoro, L. Marris, C.J. Akerman, and G. Hinton. Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6):335–346, 2020.

[12] B.A. Richards, T.P. Lillicrap, P. Beaudoin, et al. A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11):1761–1770, 2019.

[13] A Vaswani. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017.

[14] M.I. Posner, S.E. Petersen, et al. The attention system of the human brain. *Annual Review of Neuroscience*, 13(1):25–42, 1990.

[15] S. Ahmad. VISIT: A neural model of covert visual attention. *Advances in Neural Information Processing Systems*, 4, 1991.

[16] P. Ramachandran, N. Parmar, A. Vaswani, et al. Stand-alone self-attention in vision models. *Advances in Neural Information Processing Systems*, 32, 2019.

[17] A. Dosovitskiy. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[18] A. Galassi, M. Lippi, and P. Torroni. Attention in natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(10):4291–4308, 2020.

[19] A. Vyas, A. Katharopoulos, and F. Fleuret. Fast transformers with clustered attention. *Advances in Neural Information Processing Systems*, 33:21665–21674, 2020.

[20] T. Brown, B. Mann, N. Ryder, et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901, 2020.

[21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, volume 1. Minneapolis, Minnesota, 2019.

[22] N. Parmar, A. Vaswani, J. Uszkoreit, et al. Image transformer. In *International Conference on Machine Learning*, pages 4055–4064. PMLR, 2018.

[23] B. Yu, P. Lu, R. Wang, W. Hu, X. Tan, W. Ye, S. Zhang, T. Qin, and T.-Y. Liu. Museformer: Transformer with fine-and coarse-grained attention for music generation. *Advances in Neural Information Processing Systems*, 35:1376–1388, 2022.

[24] D.H Hubel and T.N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106, 1962.

[25] D.O. Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.

[26] D. Marr and W.T. Thach. A theory of cerebellar cortex. *From the retina to the neocortex: selected papers of David Marr*, pages 11–50, 1991.

[27] D.J. Willshaw, O.P. Buneman, and H.C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222(5197):960–962, 1969.

[28] C. von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:413–445, 1976.

[29] D.J. Willshaw and C. von der Malsburg. How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 194(1117):431–445, 1976.

[30] S.I. Amari. Topographic organization of nerve fields. *Bulletin of Mathematical Biology*, 42:339–364, 1980.

[31] R. Miikkulainen, J.A Bednar, Y. Choe, and J. Sirosh. *Computational maps in the visual cortex*. Springer Science & Business Media, 2006.

[32] T. Kohonen. Essentials of the self-organizing map. *Neural Networks*, 37:52–65, 2013.

[33] H. Yin. The self-organizing maps: background, theories, extensions and applications. In *Computational Intelligence: A Compendium*, pages 715–762. Springer, 2008.

[34] C.M. Bishop, M. Svensén, and C.K. Williams. GTM: The generative topographic mapping. *Neural Computation*, 10:215–234, 1998.

[35] J.A. Bednar, A. Kelkar, and R. Miikkulainen. Scaling self-organizing maps to model large cortical networks. *Neuroinformatics*, 2:275–301, 2004.

[36] P.I. Pavlov. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, 17(3):136, 2010.

[37] R.S. Sutton and A.G. Barto. *Reinforcement learning: an introduction.* MIT Press, Cambridge, MA, 1998.

[38] A.G. Barto. Reinforcement learning control. *Current Opinion in Neurobiology*, 4(6):888–893, 1994.

[39] P.R. Montague, P. Dayan, and T.J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, 16(5):1936–1947, 1996.

[40] Y. Niv and P.R. Montague. Theoretical and empirical studies of learning. In *Neuroeconomics*, pages 331–351. Elsevier, 2009.

[41] G.W. Arbuthnott, C.A. Ingham, and J.R. Wickens. Dopamine and synaptic plasticity in the neostriatum. *The Journal of Anatomy*, 196(4):587–596, 2000.

[42] J.S. Elam, M.F. Glasser, M.P. Harms, et al. The human connectome project: a retrospective. *NeuroImage*, 244:118543, 2021.

[43] E.T. Bullmore and D.S. Bassett. Brain graphs: graphical models of the human brain connectome. *Annual Review of Clinical Psychology*, 7(1):113–140, 2011.

[44] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[45] O. Sporns, G. Tononi, and R. Kötter. The human connectome: a structural description of the human brain. *PLoS Computational Biology*, 1(4):e42, 2005.

[46] F.V. Farahani, W. Karwowski, and N.R. Lighthall. Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review. *Frontiers in Neuroscience*, 13:585, 2019.

[47] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442, 1998.

[48] Y. He, J. Wang, L. Wang, et al. Uncovering intrinsic modular organization of spontaneous brain activity in humans. *PloS One*, 4(4):e5226, 2009.

[49] K. Hilger, M. Ekman, C.J. Fiebach, and U. Basten. Efficient hubs in the intelligent brain: Nodal efficiency of hub regions in the salience network is associated with general intelligence. *Intelligence*, 60:10–25, 2017.

[50] S. Gao, G. Mishne, and D. Scheinost. Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics. *Human Brain Mapping*, 42(14):4510–4524, 2021.

[51] R. Mitchell-Heggs, S. Prado, G.P. Gava, M.A. Go, and S.R. Schultz. Neural manifold analysis of brain circuit dynamics in health and disease. *Journal of Computational Neuroscience*, 51(1):1–21, 2023.

[52] V. Mante, D. Sussillo, K.V. Shenoy, and W.T. Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84, 2013.

[53] R.J. Gardner, E. Hermansen, M. Pachitariu, Y. Burak, N.A. Baas, B.A. Dunn, M.-B. Moser, and E.I. Moser. Toroidal topology of population activity in grid cells. *Nature*, 602(7895):123–128, 2022.

[54] J.A. Gallego, M.G. Perich, L.E. Miller, and S.A. Solla. Neural manifolds for the control of movement. *Neuron*, 94(5):978–984, 2017.

[55] A. Segato, A. Marzullo, F. Calimeri, and E. De Momi. Artificial intelligence for brain diseases: A systematic review. *APL Bioengineering*, 4(4), 2020.

[56] C. Pitarch, V. Ribas, and A. Vellido. AI-based glioma grading for a trustworthy diagnosis: An analytical pipeline for improved reliability. *Cancers*, 15:3369, 2023.

[57] B. García Santa Cruz, A. Husch, and F. Hertel. Machine learning models for diagnosis and prognosis of parkinson's disease using brain imaging: general overview, main challenges, and future directions. *Frontiers in Aging Neuroscience*, 15:1216163, 2023.

[58] K. Sakai and K. Yamada. Machine learning studies on major brain diseases: 5-year trends of 2014–2018. *Japanese Journal of Radiology*, 3:34–72, 2019.

[59] L. Zhang, M. Wang, M. Liu, and D. Zhang. A survey on deep learning for neuroimaging-based brain disorder analysis. *Frontiers in Neuroscience*, 14:779, 2020.

[60] S.-X. Xu, W.-F. Deng, Y.-Y. Qu, et al. The integrated understanding of structural and functional connectomes in depression: A multimodal meta-analysis of graph metrics. *Journal of Affective Disorders*, 295:759–770, 2021.

[61] K.L. Narr and A.M. Leaver. Connectome and schizophrenia. *Current Opinion in Psychiatry*, 28(3):229–235, 2015.

[62] A. Irimia, X. Lei, C.M. Torgerson, Z.J. Jacokes, S. Abe, and J.D. Van Horn. Support vector machines, multidimensional scaling and magnetic resonance imaging reveal structural brain abnormalities associated with the interaction between autism spectrum disorder and sex. *Frontiers in Computational Neuroscience*, 12:93, 2018.

[63] N. Burgos, S. Bottani, J. Faouzi, E. Thibeau-Sutre, and O. Colliot. Deep learning for brain disorders: from data processing to disease treatment. *Briefings in Bioinformatics*, 22:1560–1576, 2021.

[64] S. Lövdal, R. van Veen, G. Carli, et al. IRMA: Machine learning-based harmonization of ^{18}F-FDG PET brain scans in multi-center studies. *European Journal of Nuclear Medicine and Molecular Imaging*, 2025.

[65] M. Ben Yahia, M. Garouani, and J. Aligon. Multimodal explainable automated diagnosis of autistic spectrum disorder. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[66] S.M. Lundberg. A unified approach to interpreting model predictions. In *arXiv preprint*, page arXiv:1705.07874, 2017.

[67] E. Romero and L. Rello. Screening dyslexia for english: Impact of heterogeneity in demographic variables. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[68] D.P. Kumpik, Y. Ben-Shlomo, E. Coulthard, A. Hepburn, and R. Santos-Rodríguez. Direct versus intermediate multi-task transfer learning for dementia detection from unstructured conversations. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[69] L.A. Moctezuma, Y. Suzuki, J. Furuki, M. Molinas, and T. Abe. Sleep staging with gradient boosting and dwt-psd features from eeg/eog signals. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[70] P. Schöfer, J. Danckert, P.F. Stadler, and M. Bogdan. The regulatory character of boredom in ai - towards a self-regulating system based on spiking neural networks. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[71] L. Guiducci, G.M. Dimitri, G. Palma, and A. Rizzo. Introducing intrinsic motivation in elastic decision transformers. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.

[72] Y. Marquise, B. Benziane, H. Bani, D. Lambert, Y. Cherifi, and D. Sheynikhovich. Explainable deep learning reveals a behavioral strategy underlying human decisions in a spatial navigation task. In *Procs. of the 33th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2025)*, 2025.