GAUCHER INSTITUTE
an educational platform developed and funded by Takeda

How might artificial intelligence and machine learning bring greater understanding to Gaucher disease?

What are the current challenges faced by patients with rare diseases?

Up to 40% of patients with a rare disease are **misdiagnosed** or receive no diagnosis at all.[1]

Patients with rare diseases (e.g., Gaucher disease) often consult **up to 8 different physicians** about their condition.[2]

6–8 years

It can take on average **6–8 years** before a person with a rare disease receives the correct diagnosis.[3]

Diagnostic delays in patients with rare diseases, such as Gaucher disease, may lead to irreversible complications associated with disease progression.[2]

Knowledge gaps exist in the pathophysiology underlying rare diseases, and the available treatment options for many rare diseases can be limited.[4]

Improving the diagnosis and treatment of rare diseases may be regarded as an important public health concern.[4]

Given the specific challenges associated with their diagnosis and treatment, patients with rare diseases may benefit from advances in information technology, such as in the fields of artificial intelligence (AI) and machine learning.[4]

What is artificial intelligence?

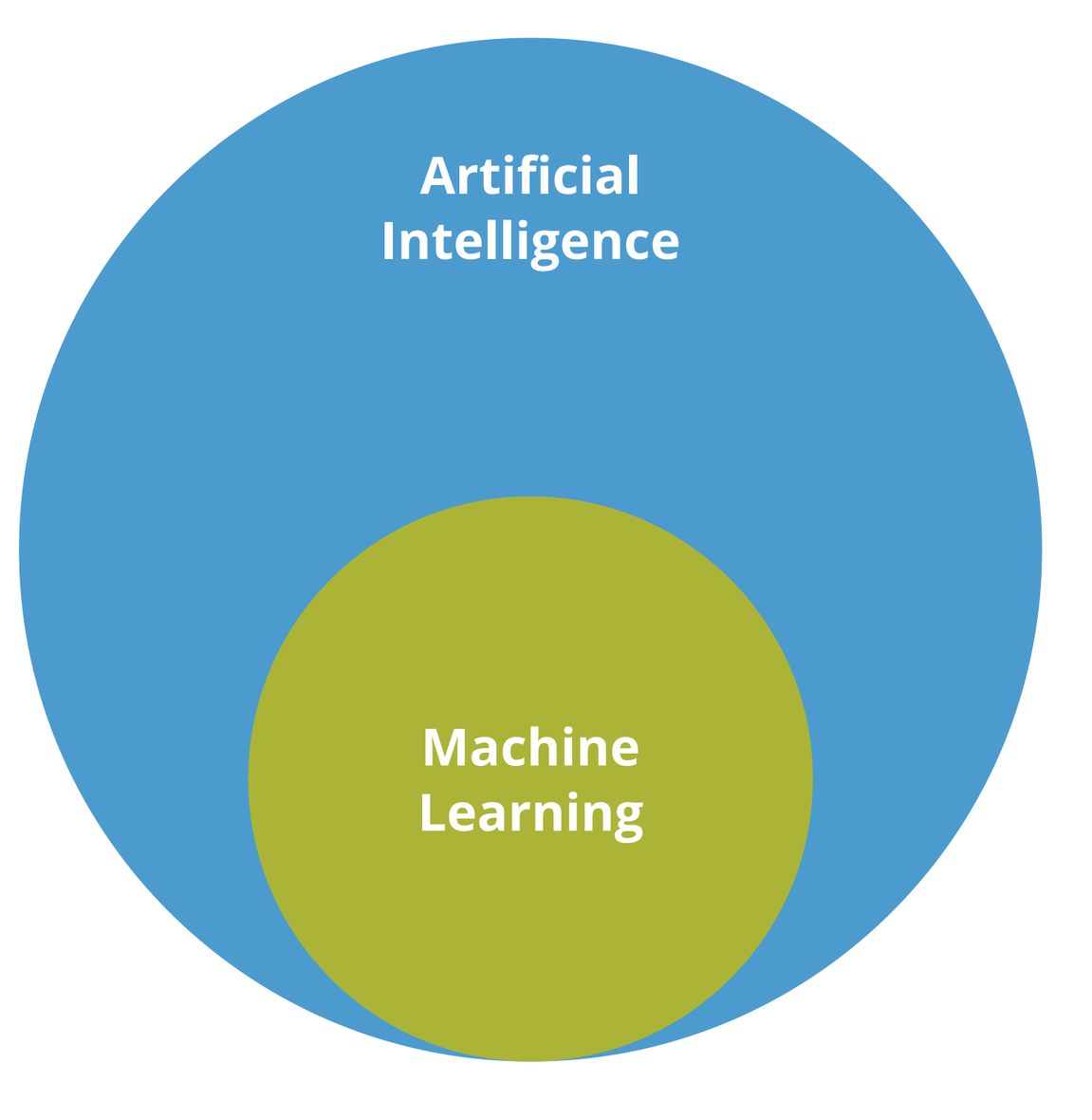
AI is defined as the 'science and engineering of making intelligent machines, especially intelligent computer programmes'.[5] AI relates to computer systems being able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.[6]

What is machine learning?

Machine learning is a subset of AI that provides computer systems or algorithms with the ability to automatically learn and improve from experience without being explicitly programmed.[7–9] For example, software algorithms can be designed and trained to learn from and act on data without human intervention or assistance.[7,10]

There are three main categories of machine learning:
i) supervised learning;
ii) unsupervised learning; and
iii) reinforcement learning.[11]

Machine learning is a subset of artificial intelligence[8]

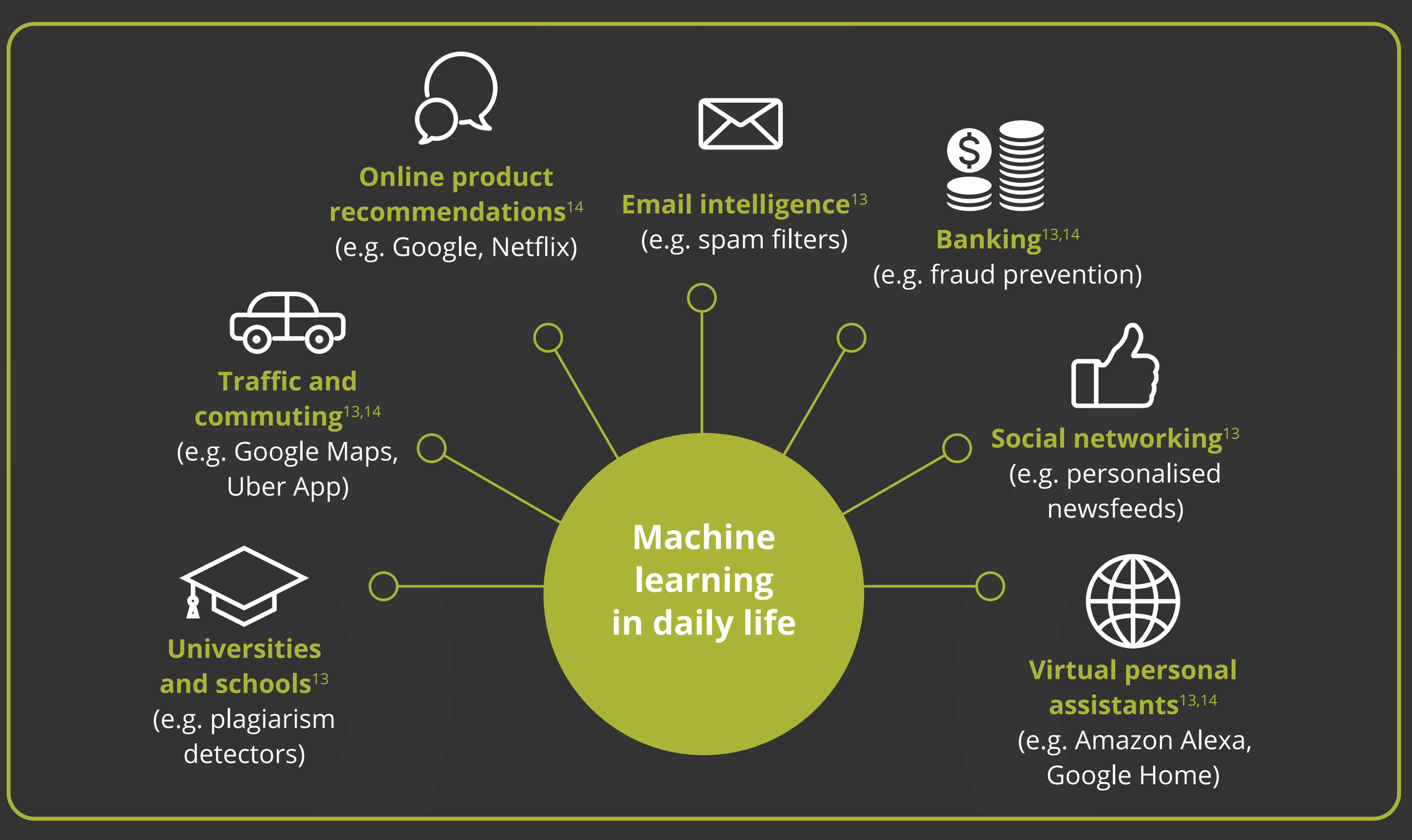


Machine learning enables algorithms to:

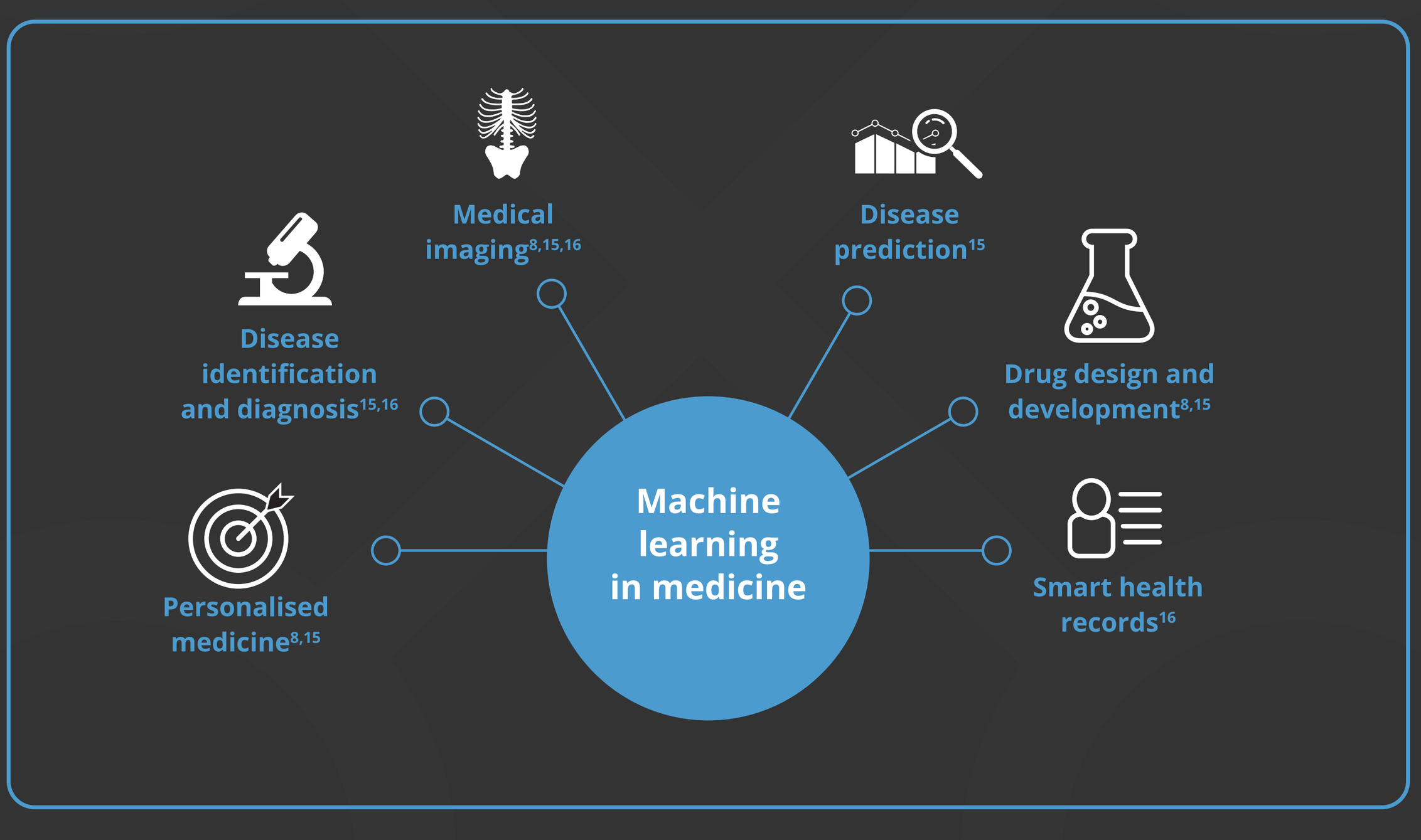
- Absorb input data and predict output values[9]
- Identify patterns and trends within data[9]
- Learn from previous experience[9]

All machine learning techniques are AI techniques, but not all AI techniques are machine learning techniques.[12]

Everyday uses of machine learning have become increasingly commonplace.[8,13,14]



Machine learning is also rapidly becoming an important approach used across medicine and healthcare.[15,16]



How might the use of machine learning benefit patients with a rare disease, such as Gaucher disease?

Almost 85% of rare diseases affect less than one patient in a million, so very experienced physicians may never be visited by a patient with a rare disease throughout their entire career.[4,17]

Therefore, physicians can lack an awareness of the early clinical signs and symptoms of disease.[2,18] Diagnosis of Gaucher disease can be particularly challenging for non-Gaucher disease specialists due to the wide variability in age, severity and type of clinical manifestations at presentation.[18]

Advanced machine learning algorithms can be 'trained' to memorise vast quantities of digital information about thousands of rare diseases.[4] These algorithms may help to detect patients with rare diseases, independent of a physician's ability to recognise the signs and symptoms of disease.[19]

Consequently, advanced machine learning algorithms for the early diagnosis of patients with rare diseases, such as Gaucher disease, have the potential to:

- **reduce diagnostic delays[19]**
- **facilitate prompt initiation of appropriate therapy[19]**
- **enable earlier decision-making[19]**
- **avoid potentially irreversible morbidities[19]**

What examples are there of applying machine learning in Gaucher disease?

A study is underway to develop an advanced machine learning algorithm for the **early diagnosis** of Gaucher disease using real-world data.[19] A large sample size of 2.2 million detailed medical records and state of the art machine learning techniques are expected to result in a predictive model with a high degree of diagnostic accuracy.[19]

A machine learning approach has been used to **analyse diagnostic data** of patients in the Spanish Gaucher Disease Registry in order to **determine any potential correlation** of these data with **long-term complications of Gaucher disease** such as bone crises, development of cancer or Parkinson's disease.[20] The study found the following correlations:

- Previous splenectomy with the development of skeletal complications and severe bone disease.[20]
- Serum levels of IgA and delayed therapy with severe bone disease at diagnosis and with the incidence of bone crisis during therapy.[20]
- High IgG levels and age over 60 years at diagnosis with the development of cancer.[20]

Deep learning, a subset of machine learning, has been used to **define the mechanism of glucocerebrosidase enzyme dysfunction** that underlies the pathophysiology of Gaucher disease.[21] Two of the common pathogenic *GBA1* gene variants, *N370S* and *L444P*, were shown to cause destabilisation of the complex usually formed between glucocerebrosidase and its facilitator protein saposin C, leading to reduced glucocerebrosidase activation.[21]

A study has used an advanced machine learning approach to **predict the responsiveness of Gaucher disease variants to pharmacological chaperone treatment.**[22] The *N370S* variant was predicted to be responsive, whereas the *L444P* variant was predicted to be non-responsive.[22]

What is potentially the biggest challenge associated with the use of machine learning in rare diseases?

The effectiveness of a machine-learning algorithm is strongly related to the amount of data used in the 'training' process – the larger the available dataset, the better the algorithm can learn and improve from experience.[4,8] Given this, the performance of machine learning algorithms in rare diseases may be hindered by the scarcity of available patient data.[23]

References
1. Eurordis. Survey of the delay in diagnosis for 8 rare diseases in Europe ('Eurordiscare 2'). Available at: https://www.eurordis.org/sites/default/files/publications/Fact_Sheet_Eurordiscare2.pdf. Accessed 04 November 2021.
2. Mistry PK, Sadan S, Yang R, et al. Consequences of diagnostic delays in type 1 Gaucher disease: the need for greater awareness among hematologists-oncologists and an opportunity for early diagnosis and intervention. Am J Hematol 2007; 82: 697-701.
3. Global Genes. Rare disease facts. Available at: https://globalgenes.org/rare-disease-facts/. Accessed 08 November 2021.
4. Schaefer J, Lehne M, Schepers J, et al. The use of machine learning in rare diseases: a scoping review. Orphanet J Rare Dis 2020; 15: 145.
5. McCarthy J. What Is artificial intelligence? Available at: http://jmc.stanford.edu/articles/whatisai/whatisai.pdf. Accessed 03 November 2021.
6. Oxford English Dictionary. Available at: https://www.lexico.com/definition/artificial_intelligence. Accessed 05 November 2021.
7. Expert.ai. What is machine learning? A definition. Available at: https://www.expert.ai/blog/machine-learning-definition/. Accessed 05 November 2021.
8. Koteluk O, Wartecki A, Mazurek S, et al. How do machines learn? Artificial intelligence as a new era in medicine. J Pers Med 2021; 11: 32.
9. Handelman GS, Kok HK, Chandra RV, et al. eDoctor: machine learning and the future of medicine. J Intern Med 2018; 284: 603-619.
10. US Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device. Available at: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device#whatis. Accessed 05 November 2021.
11. Towards data science. What are the types of machine learning? Available at: https://towardsdatascience.com/what-are-the-types-of-machine-learning-e2b9e5d1756f. Accessed 04 November 2021.
12. Toh TS, Dondelinger F, Wang D. Looking beyond the hype: applied AI and machine learning in translational medicine. EBioMedicine 2019; 47: 607-615.
13. Analytic steps. 7 popular applications of machine learning in daily life. Available at: https://www.analyticssteps.com/blogs/7-popular-applications-machine-learning-daily-life. Accessed 08 November 2021.
14. edureka! Top 10 applications of machine learning: machine learning applications in daily life. Available at: https://www.edureka.co/blog/machine-learning-applications/. Accessed 08 November 2021.
15. Jayatilake SMDAC, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. J Healthc Eng 2021; 2021: 6679512.
16. Built in beta. 15 Examples of machine learning in healthcare that are revolutionizing medicine. Available at: https://builtin.com/artificial-intelligence/machine-learning-healthcare. Accessed 08 November 2021.
17. Wakap SN, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet 2020; 28: 165-173.
18. Mehta A, Belmatoug N, Bembi B, et al. Exploring the patient journey to diagnosis of Gaucher disease from the perspective of 212 patients with Gaucher disease and 16 Gaucher expert physicians. Mol Genet Metab 2017; 122: 122-129.
19. Revel-Vilk S, Chodick G, Shalev V, et al. Study design: development of an advanced machine learning algorithm for the early diagnosis of Gaucher disease using real-world data. Blood 2020; 136: 13-14.
20. Andrade-Campos MM, de Frutos LL, Cebolla JJ, et al. Identification of risk features for complication in Gaucher's disease patients: a machine learning analysis of the Spanish registry of Gaucher disease. Orphanet J Rare Dis 2020; 15: 256.
21. Romero R, Ramanathan A, Yuen T, et al. Mechanism of glucocerebrosidase activation and dysfunction in Gaucher disease unraveled by molecular dynamics and deep learning. Proc Natl Acad Sci U S A 2019; 116: 5086-5095.
22. Woodard J, Zheng W, Zhang Y. Protein structural features predict responsiveness to pharmacological chaperone treatment for three lysosomal storage disorders. PLoS Comput Biol 2021; 17: e1009370.
23. Decherchi S, Pedrini E, Mordenti M, et al. Opportunities and challenges for machine learning in rare disease. Front Med (Lausanne) 2021; 8: 747612.

Takeda

C-ANPROM/INT/GAUD/0128
Date of preparation: January 2022