Artificial intelligence in the 21st century: the treasure hunt for systematic mining of natural products

Janani Manochkumar and Siva Ramamoorthy*

School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore 632 014, India

Advancements in genome mining, high-throughput sequencing and experimental techniques have generated an enormous amount of data on natural products. This has led to the design and development of advanced machine learning (ML) and artificial intelligence (AI) algorithms which have simplified the search for novel natural products in the 21st century. These algorithms could effectively analyse the chemical structure of natural products and predict their biological function. They could also effectively analyse large sets of data in a sophisticated manner. In this context, this article reviews the various AI/ML algorithms employed in natural products-based drug discovery. Particular attention is paid to case studies employing AI tools in plant and microbial research. Challenges associated with the use of AI tools for natural products research have also been discussed.

Keywords: Artificial intelligence, dereplication, drug discovery, genome mining, machine learning, natural products.

ARTIFICIAL INTELLIGENCE (AI) utilizes [illegible] of NPs makes them a treasured source of novel bioactive compounds with various applications in the agricultural, biotechnological, food, cosmetics and pharmaceutical industries[5,6].

There are over 465,000 plant species existing on the Earth, of which 391,000 are vascular plants[7]. One of the enthralling facts about plants is their unique metabolic pathway which corresponds to the synthesis of highly complex bioactive metabolites[8]. The diversity of plant metabolites is estimated to exceed 1 million with each plant contributing to more than 4.7 structurally unique compounds[9]. The use of plant extracts as a commercial product in food and flavour, cosmetics, and pharma industries has been predicted to reach USD 59.4 billion by 2025 (ref. 10). Plants have also been used for the treatment of several diseases worldwide[11]. Based on this evidence, researchers are now focusing their studies on the potential of plants and microbes to render NPs with beneficial therapeutic effects[8]. Over the last few decades, AI has been utilized in the screening of plant extracts, chemical taxonomy, chemical fingerprinting, phylogenetic studies, predicting toxic properties and determining the structure of phytochemicals based