American Mineralogist, Volume 111, pages 345-359, 2026

Uncovering the mineral assemblages of six major gold deposit types: A machine learning approach to big data analytics of a global mineralogical database

Bin Wang[1], Renguang Zuo[1,*,†], and Oliver P. Kreuzer[2]

[1]State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China
[2]Economic Geology Research Centre (EGRU), College of Science and Engineering, James Cook University, Townsville, Queensland 4811, Australia

Abstract

Identifying mineral assemblages is crucial for developing a better understanding of ore genesis and improving mineral exploration efficiency. Traditional geological methods have provided significant insights into the classification of the many different gold deposit types and their genesis, but they typically occur at the deposit to thin section scale and focus on small, local data sets. This study developed a data-driven approach, leveraging machine learning and big data to determine the characteristic mineral assemblages of six globally significant gold deposit types: orogenic, epithermal, porphyry, Carlin, iron oxide-copper-gold (IOCG), and volcanogenic massive sulfide (VMS). We utilized association rule mining (ARM) with an improved Apriori algorithm, which constrains rules consequent to deposit types, on a global database of 454 gold deposits, aiming to unravel the characteristic mineral assemblages of six of the world's most economically significant gold deposit types. Visualization of the rule set through bipartite and unipartite networks revealed distinct mineral-to-gold deposit relationships. This study also showed that a machine learning approach to big data analytics of a global mineralogical database can detect both known and as-of-yet-unrecognized mineral associations. As such, our approach, which links geology and big data, offers new opportunities for mineral exploration targeting and gold deposit research.

Keywords: Big mineralogical data, machine learning, gold deposit types, mineral assemblages, Apriori algorithm; Mineral Informatics: Revolutionizing Mineralogy, Petrology, and Geochemistry