Using Data Mining Algorithms in Solid State Physics

TROY LYONS, Central Michigan University, NICHOLAS MECHOLSKY, The Catholic University of America, STEFANO CURTAROLO, Duke University, MARCO BUONGIORNO NARDELLI, University of North Texas, MARCO FORNARI, Central Michigan University — We processed large materials databases with data mining methods such as clustering and classification in order to answer specific questions in the field of thermoelectric materials and transparent conductors. Our goal is to extract meaningful information from band structures repositories such as AFLOWLIB. Our implementation is validated using a toy database that mimics the complexity of AFLOWLIB, which has also been solved analytically. We found that even when the analytical solution is known, proper data analysis can help to understand physical phenomena.