

Nature Reviews Materials: Machine learning for alloys

主題: High entropy alloys

心得:

選擇高熵合金這個主題的原因，是我在大二的材料熱力學課堂中，曾學習過使用 Pandat 這個軟體來繪製相圖，當初老師給我們的作業是繪製一個二元合金的相圖，但光是二元合金，就必須參考好幾篇論文實驗出的熱力學公式才能繪製而成，若是要研究多元的高熵合金，需要花費的時間成本可想而知是相當可觀，這也使我好奇目前的機器學習究竟能對高熵合金的研究帶來怎樣的幫助，是否能夠建立一個有效率的研究流程幫助這個領域的研究者，減少 trial and error 所消耗的時間與金錢成本。

高熵合金是由至少四種主元素形成的單相固溶體，與直覺相反的是，混入如此多種的元素並沒有讓高熵合金的相變化變得複雜，反而藉由每種原子隨機散布的特性，抑制了脆性化合物的生成，使材料獲得更好的韌性。然而在原子排列的方面，當各種不同大小的元素要排列成單一晶格時，就會出現晶格扭曲和應變現象，使高熵合金的性質難以預測，也讓不同元素比例的拿捏變得困難。要憑藉人腦在如此複雜的特性中找出規律顯然是不切實際，而機器學習在這樣的任務中便有了最好的發揮空間。

從這篇文章當中可以知道，在相變化的領域，已經有大量的研究利用常見

的機器學習模型如神經網路、K 鄰近法、SVM、主成份分析法、隨機森林等，找出了對於相變化來說重要與不重要的特徵。另一方面，科學家們也透過已成熟的 cluster expansions、ab initio method 等理論來產生訓練資料，並成功訓練出能達到與理論計算相同效果的模型。上述的這些研究成果，無疑有為高熵合金的發展帶來更多進步，但我認為對於機器來說都不是太困難的挑戰，畢竟機器學習最擅長的就是篩選出特徵，與擬合某些公式。

真正讓我覺得對高熵合金的開發有很大幫助的，是在 Thermo-mechanical properties 這個段落提到的《Machine learning assisted design of high entropy alloys with desired property》這篇論文，在這篇論文當中，作者想要探索 Al-Co-Cr-Cu-Fe-Ni 合金系統的成份空間，並找出高硬度的組成成分，為此，他們設計了一個實驗循環：首先，利用已知的性質的訓練資料加上人為篩選的特徵池，訓練出用來預測性質的 Surrogate Model，並用這個 model 開始在成份空間中找出高硬度的成份組合，接著用 utility function 評估進行開發或搜索，在開發完成後，將新獲得的資料加入訓練資料集中，並從一直重複循環下去直到開發出理想硬度的材料。透過這個巧妙的循環設計，即使研究者一開始手上資料不多，也可以在整個實驗循環重複進行的過程當中，逐漸提升 Surrogate Model 的準確性，藉此讓每次開發新組合時都變得有意義，省去盲目探索六元合金成份空間的過程。

這份研究讓我對材料科學與機器學習領域之間的關係有了新的想法，過去

我會認為要用機器學習進行材料領域的研究，就必須要先有資料量足夠的資料集，但其實先使用少量的資料集訓練出準確度不那麼高的模型，就已經能對實驗產生幫助，並且隨著實驗的進行，不僅能夠使用較少的時間和資源得到需要的材料，過程中累積下來的資料集還有可能再以更複雜的機器學習模型做出更高準確率，或是更全面的性質預測，我認為機器學習將會成為加速高熵合金研究的一大助力。