Ideas for capstone:

Use pre-trained protein language model (e.g. from meta: <https://www.pnas.org/doi/epdf/10.1073/pnas.2016239118>) to re-train against protein property data.