Chemical Hide and Seek - Using AI to efficiently find high quality leads

Edward O. Pyzer-Knapp

IBM Research, IBM United Kingdom Ltd.,
PO Box 41, North Harbour, Portsmouth, Hampshire PO6 3AU
EPyzerK3@uk.ibm.com

Library generation is increasingly an automated process with combinatorial methods vying with newer AI driven generative models. The order in which these libraries are investigated (I.e which molecules are tested when) is, perhaps surprisingly, an area where large efficiency gains can be made. Classic design of experiments is a static method, and can be maddening inefficient. I will show how Bayesian optimisation is a highly effective way to achieve effective yet dynamic design of experiments which are robust to challenging problems. I will also show a new method for parallel Bayesian optimisation which proves extremely effective on a complex lead discovery problem.