**Machine Learning**

The Variational AutoEncoder is designed to capture and learn intricate patterns from the vast chemical space of silanes by compressing and expanding the embedded information. The model acquires fluency in the language of molecules, for the purpose of being leveraged as a tool to discover optimal chemistries.

**Model Architecture**

At the forefront of our model lies the encoder, deftly encoding intricate chemistries onto a precise mathematical space, capturing essential features and relationships. Complementing the encoder, the **property predictor** artfully constrains high-strength candidates within a defined region, ensuring precision in material characteristics and guiding the generation of silanes with desired properties. The decoder, our third cornerstone, exemplifies finesse as it learns to effortlessly reconstruct encoded molecules.

**Latent & Property Predictor**

The latent dimension serves as a compressed representation of the high-dimensional chemical space of silanes, effectively reducing the data's dimensionality while retaining essential information, capturing meaningful features in a more concise form.

**Marking Optimum Regions**

With this strategic combination, we efficiently explore the latent space, targeting regions that yield high-strength silanes and revolutionizing material design.

**Acknowledgements**

- Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-22-2-0014. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.
- I would like to acknowledge the guidance received from Timothy W. Sirk and Chris Rinderspacher (U.S. DEVCOM Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005, USA).