Machine Learning to Accelerate Screening for Marcus Reorganization Energies

Omri Abarbanel and Geoffrey Hutchison • Department of Chemistry, University of Pittsburgh

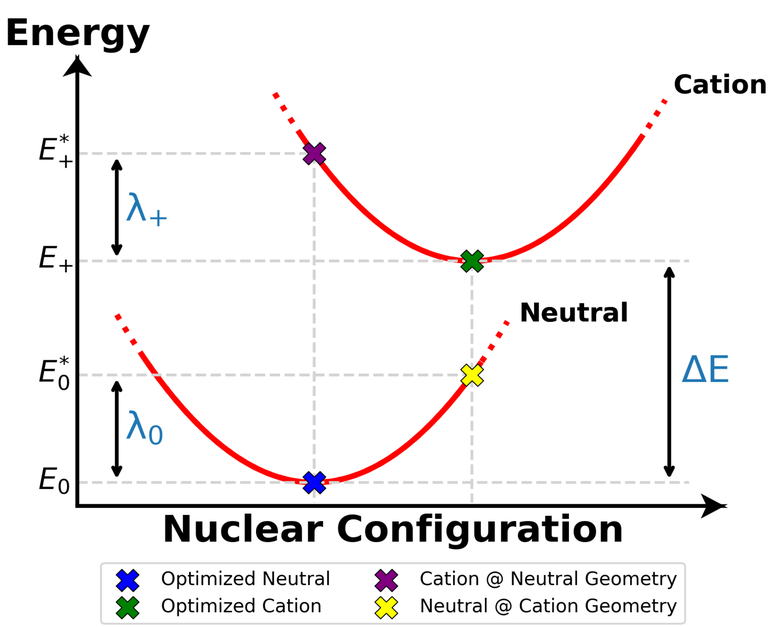
Reorganization Energy

Reorganization energy (λ) acts as an activation barrier during charge transfer between two molecules:

$$M + M^+ \rightarrow M^+ + M$$

During charge transfer the conformation of the molecule changes to minimize its energy. Computationally, the reorganization energy of a molecule can be calculated by calculating 4 points - the energy of the neutral species, the energy of the cationic species, the energy of the neutral species at the cation conformation, and the energy of the cation at the neutral conformations:

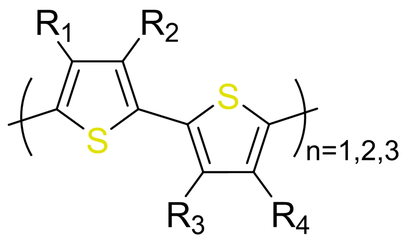
$$\lambda = \lambda_0 + \lambda_+ = (E_0^* - E_0) + (E_+^* - E_+)$$

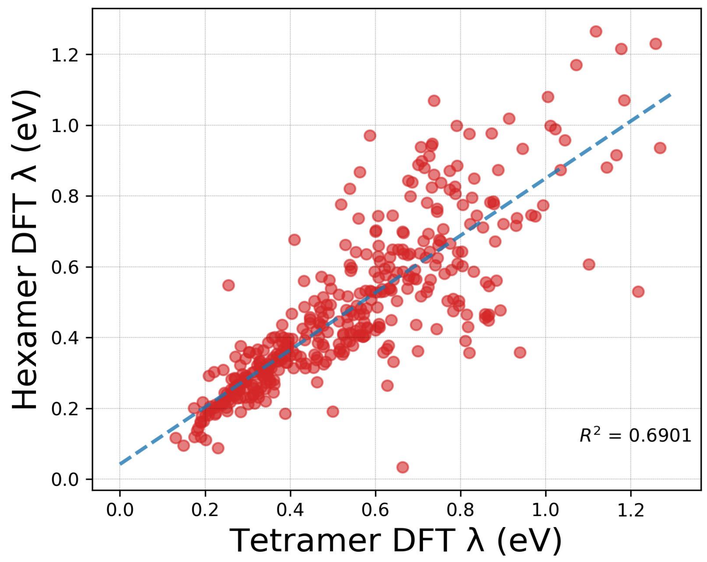


Calculating those energies using DFT methods can be long and resource intensive, which is why we employed a Machine Learning (ML) model for predicting the reorganization energy of π-conjugated oligomers. This ML model acts as a screening tool to find candidates with low reorganization energy, as low reorganization energy means a lower energy barrier for charge transfer, and a polymer with better conductivity.

Data Set

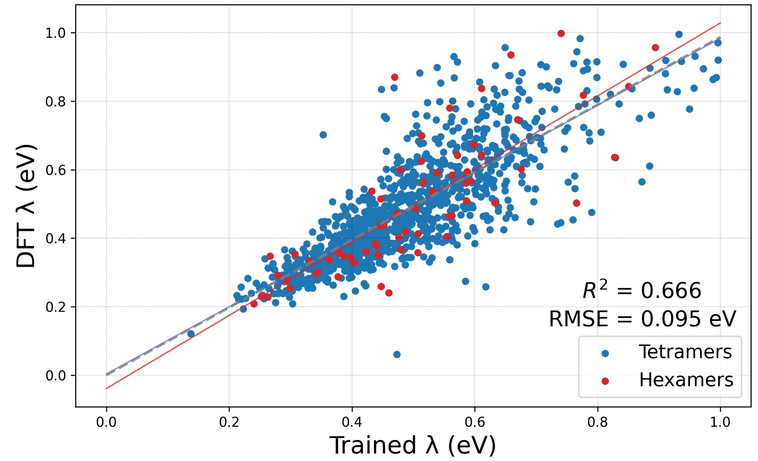
We used a list of 253 thiophene-based monomers previously used in our group, which gave us a set of ~32K possible copolymers and homopolymers. For each possible polymer we created an oligomer with a length of 2, 4, and 6 monomers, dubbed dimers, tetramers, and hexamers. Since our goal was to minimize the number of DFT calculations needed, we found that the reorganization energies of the tetramers correlate with these of the hexamers. Therefore we calculated the reorganization energy of ~7.2K tetramers and ~400 hexamers to be used as for training and testing the ML model.

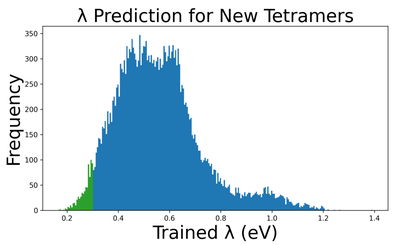


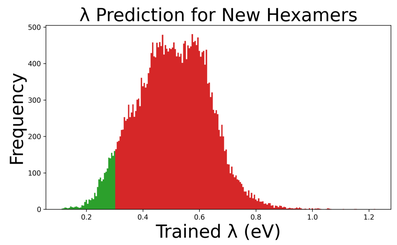


Machine Learning Model

We have found that a Random Forrest regression model gives better results and is more consistent than a Neural Network model. Similar to the oligomer length correlation plot between the tetramers and hexamers, the predictions show heteroscedasticity, however we are interested only in oligomers with low reorganization energy, where the model gives a good prediction. Using this model we predicted the reorganization energy of the rest of the tetramers and hexamers and found candidates with λ < 0.3 eV, as well as the best performing monomers.







Conclusion and Future Work

We have found that ML can perform as a fast screening tool with good accuracy - a first step in the road to find high-performance conductive polymers. We are currently working on running DFT calculations to find the reorganization energy of the best performing oligomers and comparing the results to our model.