CH5650: Molecular Data Science and Informatics
End Term Examination (Maximum Marks – 30)

Start Time: 8 AM, May 10, 2023; End Time: 6 PM, May 10, 2023

Question 1 (Max. Marks 10)

. The radius of gyration of a polymer is written as

$$R_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i - r_{CM})^2}.$$

Here r_i is the position of the i-th monomer, r_{CM} is the center of mass of the polymer, and N is the total number of monomers in the polymer. Develop a model for predicting the radius of gyration of an AB-polymer as a function of the sequence A and B moieties in the molecule. Use the **DataSet1** to build and test the model performance.

Question 2 (Max. Marks 10)

Develop a model to predict the glass transition temperature of a polymer based on its monomer's physicochemical descriptors. You can use the RDKit to extract physicochemical descriptors of a monomer. Use **DataSet2** for building and testing the performance of the model.

Question 3 (Max. Marks 10)

Develop a model to predict the formation energy, enthalpy and specific heat of a molecule based on its structural fingerprints. Use **DataSet3** for building and testing the performance of the mode.

Note: You can use any algorithm for building the models. You should generate parity plots for test and train data sets for all the cases. In addition to codes, please submit a write-up (PDF file) wherein you should include parity plots, loss functions, discuss how you build the models, and any other relevant information.