Prediction of percentage aggregate in a sample from DLS data using machine learning algorithm

Anuj Shrivastava, S Mandal, Sudip K Pattanayek, Anurag S. Rathore*
Department of Chemical Engineering, Indian Institute of Technology Delhi, New Delhi, India
**Corresponding author Email: asrathore@biotechcmz.com*

Abstract: Dynamic Light scattering have been a well-established technique to estimate the average size of the aggregate or to evaluate the stability of the sample. In this study we have discussed one more prominent solution that can be achieved by DLS, that is calculating composition of a mixture. Mostly researchers use size exclusion chromatography to determine the composition of components in a mixture which is time consuming. We have demonstrated the technique based upon machine learning algorithm and the implication of regression model to predict the oligomers in the monoclonal antibodies. The two machine learning models based on neural network algorithm and support vectors machine algorithm models were used to determine the oligomer composition of a sample with the help of its DLS data. We were able to predict the percentage of monomer, dimer etc with root mean square error of less than 11% using both the algorithms. Also found neural network algorithm is better fitting the data in comparison with support vector machine algorithm. This proposed technique could be economical and an alternative way for rapid estimation of oligomers in the sample.