**SEMINAR SYNOPSIS**

**Title**: LIME(Local Interpretable Model-agnostic Explanation)

**Speaker**: Soundarya G

**Abstract:** In this 21 century, Artificial Intelligence and Machine Learning have emerged as powerful technology for handling big data. The complex structured ML models can provide insights into business problems, but their complexity makes it difficult for the business user to understand the logic behind their predictions. For instance, the random forest model gives high accuracy, but it is not easy to interpret and trust its predictions, while the linear regression model with less accuracy is easy to interpret. Here raises the question, "How do we trust in Machine Learning models? "

However, Interpretable machine learning has made it possible to identify significant logic behind the ML models that humans can understand and trust. This talk will discuss various theoretical and practical aspects of interpretable machine learning using LIME framework, and both specific directions feature attribution and example-based learning.

The first segment of the talk will cover the ML Interpretability approach's specific important terminologies and taxonomies for better understanding. Specifically, I will talk about LIME, the acronym for Local Interpretable Model-agnostic Explanations. It is a technique that explains the individual predictions of black-box machine learning models for both regression and classification. Moreover, I will discuss the working, advantages, and disadvantages of LIME.

In the second segment of the talk, we will discuss the model-agnostic methods in precise LIME for different datasets using Python. We will be learning the usage of three main modules, i.e., lime\_tabular, lime\_text and lime\_image, with various datasets and ML models.