**Towards Interpretability and Explainability in Intelligent Systems**

**Abstract:**

An intelligent system is a software or hardware device that uses artificial intelligence (AI) and machine learning (ML) algorithms to make decisions based on data analysis. They are the answer to the accelerated technological growth of recent years and the needs of people and organizations in an increasingly interconnected world. However, many advanced machine learning techniques (such as ResNet, VGG, etc.) are often perceived as black-boxes, making their decision less understandable to humans and prohibiting their usage in safety-critical applications. In this talk, we will discuss some of the basic ideas in explainable AI and how we can incorporate them into intelligent systems. We will study the use of a prototypical part network and show how we can use it in the detection of respiratory diseases from chest x-ray images. We will also examine the application of explainable AI techniques to state-of-the-art pattern detection algorithms such as You Only Look Once (YOLO) to detect faulty insulators in transmission lines.