**Explainable AI (XAI)**

While exploring the explainable AI topic in data science, I came across an intriguing paper titled ***"Interpretable Machine Learning: A Brief Overview".*** The paper has beautifully explained the concept of Interpretable ML models, and it has mainly emphasized on the definition of **“Interpretability”**. While going through the paper, I came across a very an exciting topic named LIME (Local Interpretable Model-Agnostics Explanation).

To simplify the concept of LIME beyond its acronym, let me offer an analogy:

LIME (Local Interpretable Model Agnostics Explanation) helps us understand the working of Black Box by playing a role of detective.

Allow me to elaborate on how this process unfolds:

**Step 1 – Selecting a prediction:**

Imagine opting for a scenario such as a loan application or an image of a cat, both of which are predictions generated by a black box algorithm.

**Step 2 – Creation of a Mysterious Case:**

Having chosen a prediction, the next step involves crafting a mysterious case. This is achieved by making a minor alteration to the chosen prediction. For instance, if we consider a loan application, a slight tweak could be applied to the income or credit score. Similarly, in the case of a cat picture, a slight adjustment to the color might suffice.

**Step 3 – Ask the Back Box:**

Feed these mystery cases (modified outputs) back to the Back Box and check what it predicts. This was LIME is simply playing a role of detective to understand what parameters the Black Box is using to make predictions.

**Step 4 – Repeat and Learn:**

Repeat feeding such mystery cases to the Black Box by making small changes. Our detective (LIME) will study how black box reacts to these cases.

**Step 5 – Build a Story:**

LIME now studies all the cases and predictions Black Box just made. It then creates a small story – an explanation to why black box did the prediction it did for the original case that we were interested in.