

LAMBDA Lab Presents
Online Course on
Explainable Artificial Intelligence
with Python

Duration: 8 Weeks (1 Session Per week)

Contact Hours: 16 Hours

Start Date: 25th Oct 2021

Mode of Communication: English

**Specially Designed for Students and Researchers who are
interested to apply XAI in their own Research Area**

Platform: Zoom

Course Instructor

Dr. Parteek Bhatia

Visiting Professor

LAMBDA Lab, Tel Aviv University, Israel

www.parteekbhatia.com

Explainable Artificial Intelligence with Python

Course overview: This course will provide a broad introduction to latest developments in the field of Explainable Artificial Intelligence and its application in various research domains. As our reliance on AI models is increasing day by day, it's also becoming equally important to explain how and why AI is making a particular decision. Recent laws have also caused the urgency about explaining and defending the decisions made by AI systems.

In this course, you will learn about tools and techniques using Python to visualize, explain, and build trustworthy AI systems. We will discuss all-important XAI techniques like LIME, SHAP, DiCE, LRP, Contrastive and Counterfactual Explanations Method in this course. You will be introduced to several open-source explainable AI tools for Python that can be used throughout the machine learning project lifecycle.

Who is the course for?

This seminar is designed for the learners who are:

- Beginner Python programmers who already have some foundational knowledge with machine learning libraries.
- Researchers who already use Python for building AI models and can benefit from learning the latest explainable AI open-source toolkits and techniques.
- Data analysts and data scientists that want an introduction to explainable AI tools and techniques using Python for machine learning models.

Interested researchers and students

Please register by 22nd Oct 2021

For Registration, please fill following form

<https://bit.ly/3A6KfNv>

Or visit www.paratekbhatia.com/XAI

Scan QR Code to Register



Tentative Detailed Contents and Plan

Meeting-1	Introduction to Explaining Artificial Intelligence <ul style="list-style-type: none"> • Need and Applications of Explaining Artificial Intelligence • Categorization of XAI • Various Case Studies XAI • Overview of various techniques for XAI
Meeting-2	XAI through Local Interpretable Model-Agnostic Explanations (LIME) <ul style="list-style-type: none"> • Working Principle of LIME • Understanding Mathematical representation of LIME • Applications of LIME in various Case studies
Meeting-3 (One hour of Demonstration and one hour of Self Practice)	Implementing LIME over various Datasets Case Study-1: Applying LIME over Stroke/No Stroke Health Dataset Case Study-2: Applying LIME over Newsgroup Dataset Through these case studies following topics will be covered: <ul style="list-style-type: none"> • Getting started with LIME • An experimental AutoML module • Interpreting the scores • Training the model and making predictions • Creating the LIME explainer • Interpreting LIME explanations
Meeting-4	XAI through SHapley Additive exPlanations (SHAP) <ul style="list-style-type: none"> • Working Principle of SHAP • Key SHAP Principles Symmetry, Null Player and Additivity • Understanding Mathematical representation of SHAP • Applications of SHAP in various Case studies

<p>Meeting-5</p>	<p>Research Directions for XAI By Prof. Irad E. Ben-Gal Prof. and Head, Laboratory of AI Machine Learning Business & Data Analytics (LAMBDA)</p> <p>Implementing SHAP over various Datasets—Parteek Bhatia</p> <p>Case Study-1: Applying SHAP over Stroke/No Stroke Health Dataset</p> <p>Case Study-2: Applying SHAP for understanding results of Sentiment Analyser</p> <p>Through these case studies following topics will be covered:</p> <ul style="list-style-type: none"> • Installing SHAP • Intercepting the dataset • Vectorizing the datasets • Creating, training, and visualizing the output of a linear model • Agnostic model explaining with SHAP • Explaining the original IMDb reviews with SHAP
<p>Assignment-1</p> <p>Objective: To apply LIME and SHAP on various classifiers for XAI</p>	<ul style="list-style-type: none"> • Applying LIME over mushroom dataset to explain prediction that is edible or poisonous. • Applying LIME and SHAP over MNIST Digit classification and to perform comparative analysis of both techniques for XAI. <p>References</p> <p>https://github.com/marcotcr/lime</p> <p>https://github.com/slundberg/shap</p>
<p>Meeting-6</p>	<p>AI Fairness with What-If Tool (WIT) and Counterfactual Explanations Method for XAI</p> <ul style="list-style-type: none"> • Understanding Fairness in AI

	<ul style="list-style-type: none"> • Demonstration of What-If Tool (WIT) for COMPAS dataset • Working Principle of Counterfactual Explanations Method • Understanding Mathematical representation of Counterfactual • Diverse Counterfactual Explanations (DiCE) • Concept of Belief, Truth, Justification and Sensitivity • Understanding various distance functions <p>Case Study-1 Applying Counterfactual Explanations over CelebA dataset for identification of Smile</p> <p>Through these case studies following topics will be covered:</p> <ul style="list-style-type: none"> • Installing DICE • WIT datapoint explorer and editor • Visualizing counterfactual distances in WIT <p>Exploring various data point distances</p>
Meeting-7	<p>XAI for Neural Networks with Layer wise Relevance Propagation (LRP) and Its Implementing</p> <ul style="list-style-type: none"> • Working Principle of LRP • Understanding Mathematical representation of LRP <p>Case Study Applying LRP over Brain MRI dataset</p> <p>Through this case study following topics will be covered:</p> <ul style="list-style-type: none"> • Loading the dataset • Pre-processing the dataset • Building VGG16 Model for identification of tumour • Layerwise relevance propagation for VGG16 • Calculating relevance for images
Assignment-2 Objective:	<p><u>Compare income classification on UCI census data</u> <u>binary classification</u> <u>model comparison</u></p> <p>DATA SOURCE</p>

<p>To demonstrate the use of What-If Tool for XAI.</p>	<p><u>UCI Census Income Dataset</u> Compare two binary classification models that predict whether a person earns more than \$50k a year, based on their census information. Examine how different features affect each models' prediction, in relation to each other.</p> <p><u>Text toxicity classifiers</u> binary classification model comparison keras model custom distance</p> <p>DATA SOURCE <u>Wikipedia Comments Dataset</u> Use the What-If Tool to compare two pre-trained models from ConversationAI that determine sentence toxicity, one of which was trained on a more balanced dataset. Examine their performance side-by-side on the Wikipedia Comments dataset. These are keras models which do not use TensorFlow examples as an input format.</p>
<p>Meeting-8</p>	<p>Concept of Contrastive XAI, Cognitive XAI and Future Research Directions</p> <ul style="list-style-type: none"> • Understand the working principle of Contrastive Explanations Method (CEM) • Concept of Cognitive XAI Explanations • Discussion of Future Directions for XAI

Interested researchers and students

Please register by 22nd Oct 2021

For Registration, please fill following form

<https://bit.ly/3A6KfNv>

Or Visit

www.parteekbhatia.com/XAI



Scan QR Code to Register