**DRAFT Agenda**

**EU TA CopPhil**

**4-Day Advanced Online Training on AI/ML for Earth Observation**

**for Philippine EO Professionals**

**The Technical Assistance for the Philippines’ Copernicus Capacity Support Programme**

The Technical Assistance (TA) for the Philippines’ Copernicus Capacity Support Programme (CopPhil) is part of the broader European Union (EU)-Philippines cooperation programme and is a unique flagship initiative of the EU's Global Gateway strategy.

The Global Gateway strategy is building strong partnerships to boost smart, clean and secure digital links to strengthen health, education and research systems across the world.

Within this context, CopPhil EU-funded project that supports the Philippine Space Agency (PhilSA) and the Department of Science and Technology (DOST) and other national partners to improve the use of Earth Observation (EO) data for disaster risk reduction (DRR), climate change adaptation (CCA), and natural resource management (NRM), effectively positioning the Philippines as a pioneer in the EU's international cooperation on Copernicus.

**Day 1: EO Data, AI/ML Fundamentals, and Python for Geospatial Analysis**

**Introduction:**

* Course introduction
* Video message of EU Ambassador to the Philippines H.E. Massimo Santoro
* EU Global Gateway: Copernicus Programme in the Philippines and mention co-chairs PhilSA and DOST

**Session 1: Copernicus Sentinel Data Deep Dive & Philippine EO Ecosystem (2 hours)**

* + **Module: Copernicus Program Overview; Sentinel-1 & Sentinel-2:** This module will provide a comprehensive overview of the Copernicus Earth Observation program, focusing on the Sentinel-1 (SAR) and Sentinel-2 (Optical) missions. Key data characteristics such as spectral bands, spatial resolutions (10m, 20m, 60m for Sentinel-2; up to 5m x 20m for Sentinel-1 IW mode), and temporal resolutions (5 days for S2, 6-12 days for S1 with constellation) will be detailed. Standard data products (e.g., Level-1C, Level-2A for Sentinel-2; GRD for Sentinel-1) and methods for accessing these data, including Copernicus Hubs and Google Earth Engine, will be thoroughly explained.
  + **Module: The Philippine EO Landscape:** This segment will introduce participants to the key national agencies and initiatives relevant to Earth Observation in the Philippines. This includes the Philippine Space Agency (PhilSA) and its data platforms like the Space+ Data Dashboard, the National Mapping and Resource Information Authority (NAMRIA) and its Geoportal providing access to basemaps, hazard maps, and land cover data, the Department of Science and Technology - Advanced Science and Technology Institute (DOST-ASTI) with projects like DATOS (Remote Sensing and Data Science Help Desk), SkAI-Pinas (Philippine Sky Artificial Intelligence Program), DIMER (Democratized Intelligent Model Exchange Repository), and AIPI (AI Processing Interface), and the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) for climate and weather data. The discussion will highlight how locally available datasets can complement Sentinel data for richer AI/ML analysis.
  + **Activity: Introduction to CopPhil Mirror Site and Digital Space Campus:** A brief overview of the CopPhil data repository (Mirror Site) and the Digital Space Campus will be provided, setting the stage for how training materials will be accessible for future use and contribute to national capacity building.

**Session 2: Core Concepts of AI/ML for Earth Observation (2 hours)**

* + **Module: What is AI/ML? The AI/ML workflow in EO:** This module will demystify AI and ML, outlining a typical workflow for EO applications. This workflow encompasses problem definition, data acquisition, crucial pre-processing steps, feature engineering, model selection and training, rigorous validation, and eventual deployment for operational use.
  + **Module: Types of ML: Supervised (Classification, Regression), Unsupervised (Clustering) with EO examples:** The main paradigms of machine learning will be introduced. Supervised learning, where models learn from labeled data, will be exemplified by tasks like land cover classification (assigning pixels to classes like 'forest' or 'water') and regression (predicting continuous values like 'biomass' or 'soil moisture'). Unsupervised learning, which finds patterns in unlabeled data, will be briefly touched upon with examples like image clustering.
  + **Module: Introduction to Deep Learning: Neural Networks, basic concepts:** Participants will be introduced to the foundational concepts of deep learning, including the structure of artificial neural networks (neurons, layers), activation functions (which introduce non-linearity), loss functions (which measure model error), and optimizers (which adjust model parameters to minimize error).
  + **Module: Data-Centric AI in EO: Importance of data quality, quantity, diversity, and annotation:** This module will emphasize that the success of AI/ML in Earth Observation is profoundly dependent on the data itself. The principles of Data-Centric AI will be introduced, highlighting the critical importance of high-quality, well-annotated, diverse, and voluminous datasets for training robust and reliable models. Given that EO data often presents challenges like noise, atmospheric interference, and label scarcity, a data-centric mindset is crucial for practitioners to avoid common pitfalls where models underperform due to data issues rather than inherent model flaws.

**Session 3: Hands-on: Python for Geospatial Data (2 hours)**

* + **Platform:** Google Colaboratory.
  + **Module: Setting up Colab, Google Drive integration, installing packages:** Practical guidance on navigating the Colab environment, mounting Google Drive for data access and storage, and installing necessary Python packages using pip.
  + **Module: Python basics recap:** A tailored recap of Python fundamentals (data types, control structures like loops, functions) will be provided, adjusted based on a pre-training assessment of participants' Python proficiency.
  + **Hands-on: Loading, exploring, and visualizing vector data with GeoPandas:** Participants will work with sample vector datasets, such as Philippine administrative boundaries or example Areas of Interest (AOIs), learning how to read, manipulate, and plot this data using the GeoPandas library.
  + **Hands-on: Loading, exploring, and visualizing raster data with Rasterio and Matplotlib:** This exercise will involve working with a small subset of a Sentinel-2 L2A tile. Participants will learn to open, inspect metadata, read pixel values, perform basic raster operations like cropping and resampling, and visualize the data using Rasterio and Matplotlib. A solid grasp of Python for handling both vector and raster data, as covered in this session, is a critical prerequisite. These skills form the bedrock upon which more complex AI/ML workflows will be built in subsequent days; any weakness here will likely impede understanding and progress in later modules.

**Session 4: Introduction to Google Earth Engine (GEE) for Data Access & Pre-processing (2 hours)**

* + **Platform:** GEE Code Editor and GEE Python API within Google Colab.
  + **Module: GEE Concepts: Image, ImageCollection, Feature, FeatureCollection, Filters, Reducers:** Core GEE data structures and operations will be explained, enabling participants to understand how GEE handles and processes vast amounts of geospatial data efficiently.
  + **Hands-on: Searching and accessing Sentinel-1 and Sentinel-2 collections in GEE:** Participants will learn to query the GEE data catalog for Sentinel imagery, applying spatial and temporal filters to find relevant scenes for specific AOIs and time periods.
  + **Hands-on: Basic pre-processing in GEE:** Practical exercises will demonstrate common pre-processing tasks directly within GEE, such as cloud masking for Sentinel-2 imagery using the QA bands, creating temporal composites (e.g., median or mean composites over a time range to reduce cloud effects and noise), and clipping imagery to an AOI.
  + **Module: Exporting data from GEE for use in external AI/ML workflows:** A crucial skill is moving data from GEE to platforms like Colab for custom model training. This module will cover methods for exporting processed images or feature collections (e.g., training samples) to Google Drive or Cloud Storage. The early introduction to Colab best practices and a clear understanding of GEE's capabilities alongside its limitations for certain types of advanced AI/ML (e.g., training complex custom deep learning models natively) will set realistic expectations. This empowers users to troubleshoot common issues and plan their workflows effectively, contributing to the sustainable application of these skills.

**Day 2: Machine Learning for Image Classification & Introduction to Deep Learning**

**Session 1: Supervised Classification with Random Forest for EO (1.5 hours)**

* + **Module: Theory of Decision Trees and Random Forest (RF) algorithm:** This module will explain the principles behind decision trees and how the Random Forest (RF) algorithm builds an ensemble of such trees to improve classification accuracy and robustness. Key advantages in EO, such as its ability to handle high-dimensional data (many spectral bands or features) and its non-parametric nature (making no assumptions about data distribution), will be discussed. Its sensitivity to the quality and design of training samples will also be highlighted as a critical consideration.
  + **Module: Feature selection and importance in RF:** RF provides measures of variable importance (e.g., Gini importance), which can be used to understand which input features (e.g., spectral bands, vegetation indices) are most influential in the classification process and for feature selection to potentially improve model efficiency and accuracy.
  + **Module: Training data preparation:** Best practices for preparing training data for supervised classification will be covered, including strategies for sampling (e.g., stratified random sampling), defining clear and representative land cover classes, and ensuring an adequate number of high-quality samples.
  + **Module: Model training, prediction, and accuracy assessment:** The process of training an RF classifier, using it to predict classes for an entire image, and evaluating its performance using standard metrics like the confusion matrix, overall accuracy, producer's and user's accuracy, and the Kappa coefficient will be detailed.

**Session 2: Hands-on: Land Cover Classification (NRM Focus) using Sentinel-2 & Random Forest (2.5 hours)**

* + **Case Study:** Land Cover Classification in Palawan. This case study directly aligns with the Natural Resource Management (NRM) thematic area specified in the TOR. Palawan, with its rich biodiversity and pressures from development, serves as an excellent example for NRM applications.
  + **Platform:** Google Earth Engine (for data preparation and RF classification) and/or Python with Scikit-learn in Google Colab (for a more detailed local analysis or comparison).
  + **Data:** Sentinel-2 multispectral imagery for a selected area in Palawan. Optionally, Shuttle Radar Topography Mission (SRTM) DEM data can be incorporated as an additional feature to aid classification, as terrain can influence land cover types.
  + **Workflow:**
    1. Define an Area of Interest (AOI) in Palawan.
    2. Access and pre-process Sentinel-2 imagery within GEE (e.g., cloud masking, creating a median composite for a specific period).
    3. Collect training samples: Participants will learn to digitize polygons representing different land cover classes (e.g., forest, mangrove, agriculture, urban, water) directly in GEE, or use pre-existing shapefiles if available.
    4. Extract predictor variables for training points: This includes spectral band values from Sentinel-2 and derived indices like NDVI. If DEM is used, elevation and slope can also be added as features.
    5. Train a Random Forest classifier in GEE using ee.Classifier.smileRandomForest() or a similar function.
    6. Classify the Sentinel-2 image composite for the AOI.
    7. Perform accuracy assessment using a reserved set of validation samples.
    8. (Optional) Export the classified map and discuss how QGIS can be used for further cartographic refinement, map layout, and post-processing tasks.
  + **Practical Tips/Pitfalls to Highlight:** The importance of creating high-quality, representative training samples cannot be overstated. Issues such as mixed pixels (pixels containing multiple land cover types), the impact of atmospheric conditions (even with L2A data), and the careful selection of input features (bands and indices) will be discussed as common challenges.

**Session 3: Introduction to Deep Learning: Neural Networks & CNNs (1.5 hours)**

* + **Module: Recap Neural Networks. Introduction to Deep Learning concepts:** Building upon basic neural network ideas, this module will formally introduce Deep Learning as a subfield of ML characterized by networks with multiple layers (deep architectures).
  + **Module: Convolutional Neural Networks (CNNs):** The core architecture of CNNs will be explained, including the role and function of convolutional layers (applying filters to extract features like edges, textures), pooling layers (reducing dimensionality and creating invariance to small translations), and fully connected layers (for final classification or regression). The concept of learnable filters and hierarchical feature extraction (from simple to complex features) will be central.Common activation functions like ReLU will be introduced.
  + **Module: How CNNs learn features for image analysis. Applications in EO:** The power of CNNs in automatically learning relevant features from raw pixel data will be emphasized, contrasting with traditional methods that require manual feature engineering. Diverse applications in EO, such as image classification (assigning a single label to an image patch), object detection (locating objects), and semantic segmentation (pixel-wise classification), will be showcased.
  + **Module: Introduction to TensorFlow and PyTorch for building CNNs:** A brief overview of these two leading deep learning frameworks, highlighting their main components for defining and training CNN models, will be provided. The transition from Random Forest, a more intuitively understandable ensemble method, to CNNs allows participants to appreciate the significant leap in automated feature extraction offered by deep learning. The hands-on RF case study provides a concrete example of a successful EO application before delving into the more abstract workings of CNNs.

**Session 4: Hands-on: Basic CNN for Image Classification with TensorFlow/Keras or PyTorch (2.5 hours)**

* + **Platform:** Google Colab with GPU acceleration enabled.
  + **Dataset:** A small, pre-prepared EO image patch dataset will be used. Options include a subset of the EuroSAT dataset (derived from Sentinel-2, as used in several tutorials) or image patches extracted from the Palawan LULC case study data. The dataset will consist of image patches labeled with specific classes (e.g., 'forest', 'water', 'urban').
  + **Workflow:**
    1. Load and pre-process image patches: This includes normalizing pixel values (e.g., to a 0-1 range) and resizing patches to a uniform input size for the CNN.
    2. Define a simple CNN architecture: Participants will build a basic CNN, for example, with 2-3 convolutional layers, interspersed with pooling layers, and followed by one or more dense (fully connected) layers for classification.
    3. Compile the model: This involves specifying a loss function suitable for multi-class classification (e.g., categorical cross-entropy) and an optimizer (e.g., Adam).
    4. Train the model: The CNN will be trained on the training set of image patches.
    5. Evaluate the model: Performance will be assessed on a separate validation or test set of patches using metrics like accuracy.
    6. Visualize some predictions: A few example predictions on test images will be shown to qualitatively assess performance.
  + **Focus:** The primary goal of this session is to provide a foundational understanding of the code structure and workflow for building, training, and evaluating a CNN. The emphasis is on the process rather than achieving state-of-the-art accuracy. Reference tutorials like those found for EuroSAT classification can serve as a basis.
  + **Tooling Choice Consideration:** The choice of TensorFlow or PyTorch for this initial CNN exercise will influence subsequent deep learning sessions. While PyTorch's "Pythonic" nature is often favored in research for its ease of experimentation, TensorFlow with the high-level Keras API is also very accessible for beginners. Maintaining consistency with one framework for the more advanced U-Net and LSTM models later in the training could reduce the cognitive load on participants. Providing well-commented Colab notebooks for this exercise is a direct contribution to the materials for the Digital Space Campus 1, facilitating self-paced learning.

**Day 3: Advanced Deep Learning: Semantic Segmentation & Object Detection**

**Session 1: Semantic Segmentation with U-Net for EO (1.5 hours)**

* + **Module: Concept of semantic segmentation:** This module will introduce semantic segmentation as a pixel-wise classification task, where each pixel in an image is assigned a class label (e.g., water, building, forest). This contrasts with image classification (one label per image) and object detection (bounding boxes around objects).
  + **Module: U-Net Architecture:** The U-Net architecture, highly popular for semantic segmentation in EO and medical imaging, will be detailed. This includes its characteristic encoder (contracting path) for feature extraction and context aggregation, decoder (expansive path) for precise localization and upsampling, and the crucial skip connections that fuse high-resolution features from the encoder with upsampled features in the decoder to preserve spatial detail. The role of the bottleneck layer will also be explained.
  + **Module: Applications in EO:** Various applications of U-Net in Earth Observation will be showcased, such as flood mapping from SAR/optical data, detailed land cover mapping, road network extraction, and building footprint delineation.
  + **Module: Loss functions for segmentation:** Common loss functions used for training segmentation models, such as pixel-wise cross-entropy, Dice loss, and Jaccard index (Intersection over Union - IoU), will be introduced, highlighting their suitability for handling class imbalance and focusing on boundary accuracy.

**Session 2: Hands-on: Flood Mapping (DRR Focus) using Sentinel-1 SAR & U-Net (2.5 hours)**

* + **Case Study:** Flood Mapping in Central Luzon (Pampanga River Basin), focusing on a specific past typhoon event like Ulysses (2020) or Karding (2022).
  + **Platform:** Google Colab with GPU acceleration. The deep learning framework (PyTorch or TensorFlow) will be consistent with that used on Day 2.
  + **Data:** Pre-processed Sentinel-1 SAR image patches (e.g., 128x128 or 256x256 pixels) containing both VV and VH polarizations will be provided. Corresponding binary flood masks (1 for flood, 0 for non-flood) for these patches, derived from a chosen typhoon event, will also be supplied.
    - *Data Preparation Pitfall to Highlight:* The significant effort required for preparing analysis-ready SAR data (speckle filtering, radiometric calibration, geometric terrain correction) and generating accurate ground truth flood masks will be emphasized. For this hands-on session, providing pre-processed patches is essential to allow participants to focus on understanding and implementing the U-Net model itself.
  + **Workflow:**
    1. Load SAR image patches and their corresponding binary flood masks.
    2. Implement data augmentation techniques suitable for SAR data if time permits and deemed beneficial (e.g., rotation, flip).
    3. Define the U-Net model architecture using the chosen framework.
    4. Compile the model, selecting an appropriate loss function (e.g., Dice loss or a combination) and optimizer.
    5. Train the U-Net model on the training set of patches, monitoring performance on a validation set.
    6. Evaluate the trained model's performance on a test set using metrics like Intersection over Union (IoU), F1-score, precision, and recall.
    7. Visualize some of the model's predictions on test patches, overlaying the predicted flood extent on the SAR imagery.
  + *Conceptual Hurdle for EO Users:* A key challenge for participants might be grasping how the U-Net architecture, particularly its convolutional layers, processes SAR data (which is inherently different from optical data in terms of image characteristics) and how the skip connections contribute to the precise delineation of flood boundaries. Explaining the flow of information and the role of multi-resolution feature fusion will be important. The use of a highly relevant DRR case study like flood mapping in a major Philippine river basin, employing advanced deep learning, makes this session particularly impactful. The complexity is managed by providing pre-processed data, allowing focus on the AI/ML aspects.

**Session 3: Object Detection Techniques for EO Imagery (1.5 hours)**

* + **Module: Concept of object detection:** This module will clearly define object detection as the task of not only classifying objects within an image but also localizing them, typically by drawing bounding boxes around each detected instance.
  + **Module: Overview of popular architectures:** A high-level overview of common object detection architectures will be provided:
    - *Two-stage detectors:* These first propose regions of interest and then classify those regions (e.g., R-CNN, Fast R-CNN, Faster R-CNN).
    - *Single-stage detectors:* These perform localization and classification in a single pass, often being faster (e.g., YOLO - You Only Look Once, SSD - Single Shot MultiBox Detector).
    - *Transformer-based detectors (e.g., DETR):* Briefly introduced as an emerging and powerful approach, if appropriate for an advanced audience.
  + **Module: Applications in EO:** Numerous applications of object detection in Earth Observation will be discussed, such as detecting and counting ships, vehicles, aircraft, buildings, oil tanks, and other infrastructure elements from satellite or aerial imagery.
  + **Module: Challenges in EO object detection:** Specific challenges pertinent to EO will be highlighted, including the detection of small objects relative to image size, variations in object scale and orientation, complex backgrounds (e.g., urban clutter), atmospheric effects, and often, the limited availability of large, accurately labeled EO datasets for training object detectors.

**Session 4: Hands-on: Feature/Object Detection from Sentinel Imagery (Urban Monitoring Focus) (2.5 hours)**

* + **Case Study:** Informal Settlement Growth/Building Detection in Metro Manila (e.g., focusing on areas like Quezon City or along the Pasig River corridor). This aligns with urban monitoring aspects of DRR and NRM as per the TOR.
  + **Platform:** Google Colab with GPU. The deep learning framework will be consistent with previous sessions.
  + **Data:** Pre-prepared Sentinel-2 optical image patches of an urban area in Metro Manila. These patches will come with annotations, specifically bounding boxes delineating buildings or informal settlement clusters. While Sentinel-1 can also be used for detecting built-up change, Sentinel-2 offers richer spectral information for visual object characteristics, which might be more intuitive for an initial object detection exercise.
  + **Workflow (simplified to be feasible within the time, focusing on practical application):**
    1. Load image patches and their corresponding bounding box annotations (coordinates and class labels).
    2. **Option A (Simpler, Recommended):** Utilize a pre-trained object detection model available through TensorFlow Hub or PyTorch Hub (e.g., a lightweight version of SSD or YOLO). Fine-tune this pre-trained model on the provided settlement/building dataset. This approach leverages transfer learning and is more practical for a short training session.
    3. **Option B (More Involved, if time and audience proficiency allow):** Implement a very simplified version of a single-stage detector like YOLO or SSD. This would involve understanding the grid cell approach, anchor boxes (if used), and the prediction of bounding box coordinates, objectness score, and class probabilities.
    4. Train or fine-tune the model using the prepared dataset.
    5. Evaluate the model's performance using appropriate object detection metrics (e.g., mean Average Precision - mAP).
    6. Visualize the detected bounding boxes overlaid on test images to qualitatively assess performance.
  + *Conceptual Hurdle for EO Users:* A common point of confusion can be the distinction between image classification (one label per image/patch), semantic segmentation (pixel-wise labels), and object detection (bounding boxes around instances). Clearly explaining these differences with visual examples is crucial. For object detection specifically, concepts like anchor boxes (in YOLO/SSD) and non-max suppression (NMS) for refining predictions can be challenging and will be explained intuitively. The urban monitoring case study directly addresses issues relevant to the Philippines, such as unplanned urban growth and disaster vulnerability in dense settlements. By focusing on using pre-trained models or simplified architectures, the hands-on session remains achievable while introducing core object detection concepts.

**Day 4: Time Series Analysis, Emerging Trends, and Sustainable Learning**

**Session 1: AI for Time Series Analysis in EO: LSTMs (1.5 hours)**

* + **Module: Introduction to time series data in EO:** This module will cover the nature and importance of time series data derived from Earth Observation satellites, such as sequences of Normalized Difference Vegetation Index (NDVI) values for monitoring vegetation phenology and health, or SAR backscatter time series for tracking land surface changes over time.
  + **Module: Recurrent Neural Networks (RNNs) basics and the vanishing/exploding gradient problem:** A brief introduction to RNNs as networks designed to process sequential data will be provided, along with a discussion of their limitations in learning long-range dependencies due to issues like vanishing or exploding gradients.
  + **Module: Long Short-Term Memory (LSTM) Networks:** The LSTM architecture will be introduced as a specialized type of RNN designed to overcome these limitations. Key components like the memory cell and the gating mechanisms (input gate, forget gate, output gate) that allow LSTMs to selectively remember or forget information over long sequences will be explained conceptually. Analogies to human memory can be helpful here.
  + **Module: Applications in EO:** Practical applications of LSTMs in Earth Observation will be discussed, including drought monitoring and forecasting using vegetation indices, crop yield prediction, land cover change detection and phenological analysis.
* **Session 2: Hands-on: Drought Monitoring (CCA Focus) using Sentinel-2 NDVI & LSTMs (2.5 hours)**
  + **Case Study:** Drought Monitoring in Mindanao Agricultural Zones (e.g., Bukidnon or South Cotabato).This case study directly addresses the Climate Change Adaptation (CCA) thematic area from the TOR and is of high relevance to Philippine agriculture.
  + **Platform:** Google Colab with GPU acceleration. The deep learning framework (TensorFlow/Keras or PyTorch) will be consistent with previous sessions.
  + **Data:** Pre-prepared time series data will be provided. This could consist of:
    - Monthly or bi-monthly mean NDVI values derived from Sentinel-2 imagery for selected agricultural plots in Mindanao over several years.
    - Optionally, corresponding historical drought indices (e.g., Standardized Precipitation Evapotranspiration Index - SPEI, if available from PAGASA) or rainfall data (e.g., from CHIRPS) could serve as target variables for prediction or as correlative data to interpret NDVI trends.
    - *Data Preparation Pitfall to Highlight:* Creating a consistent, cloud-free NDVI time series from Sentinel-2 imagery requires meticulous pre-processing, including accurate cloud masking, atmospheric correction, and potentially interpolation of missing values. For the training session, providing curated time series data is essential to allow participants to focus on the LSTM modeling aspects. Challenges in acquiring consistent climate data in some regions will also be noted.
  + **Workflow:**
    1. Load and prepare the time series data (e.g., sequences of NDVI values as input features, and a corresponding drought index or vegetation stress level as the output to be predicted/classified).
    2. Normalize the data (scaling values to a suitable range for the LSTM).
    3. Create input sequences for the LSTM (e.g., using a sliding window approach to generate input sequences of past NDVI values and corresponding target values).
    4. Define the LSTM model architecture (e.g., one or more LSTM layers followed by dense layers for output).
    5. Compile and train the LSTM model.
    6. Evaluate the model's performance (e.g., using Root Mean Squared Error - RMSE for regression tasks like predicting a drought index, or accuracy for classification tasks like categorizing drought severity).
    7. Plot actual vs. predicted time series to visualize the model's forecasting capability.
  + *Conceptual Hurdle for EO Users:* Understanding how LSTMs process sequential data (e.g., a series of NDVI values over several months or years) and how they "remember" past information to make future predictions or classifications can be challenging. Visualizing the input-output structure (sequence-to-sequence or sequence-to-value) and explaining the flow of information through the LSTM gates will be key.

**Session 3: Emerging AI in EO: Foundation Models, Self-Supervised Learning, Explainable AI (XAI) (2 hours)**

* + **Module: Introduction to Foundation Models for EO (GeoFMs):** This module will introduce the concept of Foundation Models – large-scale AI models pre-trained on vast amounts of diverse, often unlabeled data, which can then be adapted (fine-tuned) for various downstream tasks with minimal task-specific labeled data. The emergence of Geospatial Foundation Models (GeoFMs) specifically pre-trained on EO data (e.g., Prithvi, Clay, SatMAE, DOFA) will be discussed, highlighting their potential to revolutionize EO analysis by providing powerful, general-purpose representations.
  + **Module: Self-Supervised Learning (SSL) in EO:** SSL techniques, which enable models to learn meaningful representations from unlabeled data by defining pretext tasks (e.g., masked autoencoding, contrastive learning), will be introduced. SSL is particularly relevant for EO due to the abundance of unlabeled satellite imagery and the high cost of acquiring labels. This approach helps address data scarcity and improves model transferability and efficiency.
  + **Module: Explainable AI (XAI) for EO:** The importance of understanding the decision-making process of complex "black box" AI models like deep neural networks will be emphasized. XAI techniques aim to make these models more transparent and interpretable. Methods such as SHAP (SHapley Additive exPlanations) for feature importance, LIME (Local Interpretable Model-agnostic Explanations) for local explanations, and Grad-CAM (Gradient-weighted Class Activation Mapping) for visualizing which parts of an image a CNN focuses on, will be introduced. Discussing XAI is crucial for building trust in AI-driven EO solutions and for debugging and improving models.
  + **Activity: Brief demo of an XAI technique:** If feasible within the Colab environment and time constraints, a short demonstration of an XAI method applied to one of the models trained earlier in the workshop (e.g., showing feature importance for the Random Forest land cover model using SHAP, or visualizing activation maps for the CNN/U-Net using Grad-CAM) will be conducted.

**Session 4: Synthesis, Q&A, and Pathway to Continued Learning (2 hours)**

* + **Module: Recap of key AI/ML techniques and their applications in the Philippine DRR, CCA, NRM context:** A synthesis of the AI/ML methods covered (RF, CNNs, U-Net, LSTMs, Object Detection) and a reiteration of their relevance and application to the Philippine-specific case studies (Flood Mapping, Drought Monitoring, Land Cover Classification, Urban Monitoring).
  + **Module: Best practices for AI/ML model training, validation, and deployment in EO:** A summary of best practices, including data-centric approaches, robust validation strategies (beyond simple accuracy metrics), and considerations for deploying models operationally.
  + **Module: Introduction to the CopPhil Digital Space Campus:** A more detailed look at how the training materials (presentations, Colab notebooks, datasets, guides) will be made available on the CopPhil Digital Space Campus for self-paced learning, wider access, and continued skill development by the participants and their colleagues.
  + **Module: Fostering a Community of Practice:** Discussion on the importance of building a community among EO and AI/ML practitioners in the Philippines. Participants will be informed about relevant national initiatives like SkAI-Pinas (including DIMER and AIPI) and other PhilSA and DOST programs, creating avenues for collaboration and knowledge sharing.
  + **Open Q&A and Troubleshooting:** Dedicated time to address any remaining questions from the four days of training, discuss specific challenges participants anticipate in applying these techniques in their work, and troubleshoot any lingering technical issues.
  + **Feedback session on the training:** Collection of participant feedback to improve future iterations of the training. The final session is crucial not just for Q&A, but for empowering participants with concrete ideas on how to apply their newly acquired skills within their institutions and how to access further support and resources.