

# Bringing Machine Learning to Earth Observation Training

HAYLEY EVERS-KING<sup>1</sup>, NEIL FLETCHER<sup>1</sup>, RAVI KAPUR<sup>2</sup>, BEN LOVEDAY<sup>3</sup>, JULIA WAGEMANN<sup>4</sup>, JOANA MIGUENS<sup>1</sup>, FEDERICO FIERLI<sup>1</sup>, CHRIS STEWART<sup>5</sup>, FABRICE MESSAL<sup>6</sup>, MARK HIGGINS<sup>1</sup>

<sup>1</sup> *EUMETSAT, Germany*

<sup>2</sup> *Imperative Space, UK*

<sup>3</sup> *Innoflair, Germany*

<sup>4</sup> *Julia Wagemann Consulting*

<sup>5</sup> *ECMWF, UK*

<sup>6</sup> *Mercator Ocean International, France*

In 2021 EUMETSAT, with partners from Mercator Ocean International, ECMWF, and the European Environment agency, developed a Massive Open Online Course (MOOC) on the growing role that machine learning and related techniques are playing in the exploitation of Earth Observation data.

The aim of the MOOC was to show data users how machine learning techniques can be used in various parts of the Earth Observation value chain, and to help equip users with the skills and knowledge needed to leverage the potential machine learning offers to the sector. The MOOC focused largely on data available through the European Commission's Copernicus Programme, allowing for different thematic applications to be covered over the 6 week course duration. A specific week was dedicated to marine applications, drawing on data available from the Sentinel-1, 2, and 3 satellite missions, as well as from the Copernicus Marine Service. Case studies in the marine week showcased the role of machine learning in a number of different examples including identifying ships, merging data from different satellite sensors, retrieving water quality and subsurface parameters, and monitoring marine life.

A series of Jupyter Notebooks were developed in collaboration with expert users from the scientific community to showcase examples of machine learning in action. A strong emphasis was placed on participants running practical examples in the Copernicus WEkEO cloud computing environment. This approach required careful management of resources, and code design to allow for large numbers of users to undertake relatively computationally intensive tasks. However, this level of interactivity was essential to facilitate user uptake and independent application of knowledge gained.

Over 10,000 participants registered for the MOOC. Feedback from participants described the course as well structured and high quality, with the interactive elements provided by the Jupyter Notebooks especially beneficial. Limitations in the MOOC format have provided useful feedback for the partner organisation to design further future training interventions using machine learning techniques.

This poster will highlight the notebook resources made available through the MOOC, focussing on their utility and describing plans for future expansion of the catalogue

in the marine domain. In addition, it will highlight the lessons learned from the MOOC, including some of the challenges faced when working with large data sets in an educational context, and suggest ways to mitigate them.