

## ONS-UNECE ML 2022 Group

### TERMS OF REFERENCE

#### Introduction

Many National Statistical Offices (NSOs) are exploring how machine learning (ML) can be used to increase the relevance and quality of official statistics in an environment of growing demands for trusted information, rapidly developing and accessible technologies, and numerous competitors.

While specific business environment may vary depending on country, NSOs face similar types of challenges which can benefit from sharing knowledge, experiences and collaborating on developing common solutions within the broad official statistical community.

To address this need, the [UNECE High-Level Group for the Modernisation of Official Statistics \(HLG-MOS\)](#) launched a Machine Learning Project (2019-20). The project aimed to demonstrate the added value of ML by, for instance, investigating whether it enables production of more relevant, timely, accurate and trusted data in an efficient manner.

Building on the work of the Project, the [UK Office for National Statistics Data Science Campus](#) in partnership with the [UNECE HLG-MOS](#), launched an international initiative, the Machine Learning Group, to continue supporting the international collaboration in building the capability of NSOs to use ML as well as identifying and addressing some common challenges encountered when incorporating ML in organisations and their production processes.

#### ML 2022 Group

#### Objectives

The general objectives of ML 2022 Group are to:

1. Facilitate the creation, development and implementation of research projects and skills building activities that meet the global statistical community needs.
2. Build and engage a strong machine learning community by sharing resources, good practice, exchanging ideas and experiences, and keeping abreast of new developments in the field.
3. Offer open, shareable and easily accessible resources to the statistical community.
4. Facilitate machine learning capacity building for official statistics.

#### Membership

The criteria for membership is:

1. You currently work in a National Statistical Office or other organisation in the national and international statistical system (e.g. government department, international institution, central bank) that produces official statistics.
2. You have a strong interest in how machine learning can help improve the quality of official statistics and statistical output, and how it can be integrated into production systems.
3. You are ready to contribute to the group's activities, for example by joining one of our Theme Groups or activities (see "Themes and modes of work" section), proposing content for our all group meetings and newsletter, or presenting at one of our meetings or external engagements.

We consider applications from others who are not directly working in official statistics where they would bring expertise of benefit to our purpose of using ML to improve the production of official statistics. We are open to all levels of ML and data science expertise.

As a platform for collaboration and exchange, we depend upon our members to provide content and shape the group's direction. All members can play an active role in the group. Those with more time can collaborate regularly in one of our Theme Groups (see below). Those with time to contribute on more of a monthly or bi-monthly basis can be a member. Those with more limited time can keep in touch with key developments by joining our mailing list.

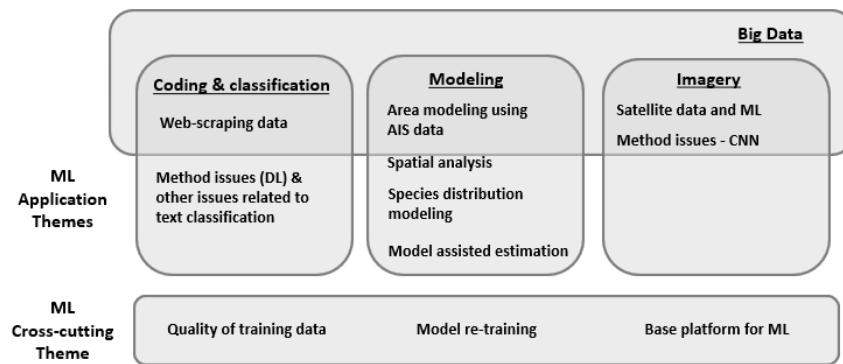
Diagram showing different membership levels



### Themes and modes of work

Following initial consultation with current and prospective members of the ML 2022 Group via survey polls and a call for activity proposals and leads, activities for 2022 were grouped into Themes as shown below:

## ML 2022 Programme



1. Text Classification
2. Modelling
3. Imagery
4. Web scraping data
5. Quality of Training Data
6. Model Retraining
7. IT infrastructure

The ML 2022 Group will publish additional information, such as objectives and ways of working, for each Theme via the [Members' website](#) in due course.

It is suggested that activity leads within a specific theme group agree ways of working. For instance, one or more colleagues can agree to lead the theme group jointly or on rotation. The role of the lead is essential to the successful functioning of the Group and will entail:

- Agreeing planned activity with group members contributors and overseeing its completion, including progress reports to be published on Wiki in July and November, along with codes and any other relevant resources
- Coordinating communication between their contributors to agree on, for instance, meetings frequency and platform (Zoom, Teams, WebEx, etc.)
- Updating the ML 2022 Group members on monthly meetings when needed
- Representing the Theme Group at higher-level meetings, if agreed

## ML Group Coordination Team

The ML 2022 Group is co-ordinated by ONS Data Science Campus and UNECE and their responsibilities entail:

- Setting up the Group and drafting its terms of reference
- Communicating with other UN agencies and exploring related initiatives
- Presenting at the HLG-MOS Executive Board through its bimonthly updates for advice and suggestions
- Co-ordinating activity proposals, setting up workstreams and contributors' membership
- Co-ordinating and hosting monthly meetings via Teams, inviting guest speakers, supporting members and workstreams leads on general queries
- Managing a dedicated email account and spaces on UNECE wiki:
  - Public wiki: <https://statswiki.unece.org/display/ML/Machine+Learning+Group+2022>
  - Members wiki:  
<https://statswiki.unece.org/display/MLP/Machine+Learning+Group+2022>
  - Dedicated inbox: [ML2022@ons.gov.uk](mailto:ML2022@ons.gov.uk)
- Developing a series of one-off tutorials in a technical aspect of machine learning
- Compiling and publishing a monthly newsletter on latest developments and opportunities in the ML for official statistics space.
- Creating social media posts (twitter, UNECE stats account) and publishing blogs (e.g. Data Science Campus webpage) to promote the Group's activities
- Setting up and managing an online platform for collaboration between Members
- Moderating the [ML group](#) under the UN Global Network of Data Officers and Statisticians:

## Potential Outputs and Timeline

1. Organise and run monthly meetings
2. Activity reports, workstreams progress reports, codes and other resources made publicly available on Wiki in July and November 2022
3. Final report published in December 2022
4. Webinar delivered in November 2022
5. Establish the [ML group](#) under the UN Global Network of Data Officers and Statisticians as an active community channel using high-impact and relevant applications of ML in Official Statistics from different NSOs
6. Facilitate delivery of engaging activities, such as short courses, masterclasses, hackathons, coffee and coding, ML for leaders talks, and webinars with support of NSOs and guest speakers
7. Publish and promote blogs and articles across key channels

