Al for official statistics

InKyung Choi

Statistical Management and Modernisation Unit

United Nations Economic Commission for Europe (UNECE)

@ITU WTIS24



Disclaimer: The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of UNECE, and it does not mean endorsement of any services or tools



ML application areas







AI and official statistics



AI can help statistical organizations to

- 1. Increase efficiency and productivity
- 2. Offer better and new products and services
- 3. Improve service delivery and user satisfaction





Buzz, must or burst?



Efficiency & productivity

Text classification

Survey response "I am a cook at a restaurant"

SOC "35-2014.00 Cooks, Restaurant"

SOC for occupation COICOP for product

ISIC for activity

Efficiency & productivity

Text classification

Statistics Poland (2020)

Efficiency & productivity

Coding and IT development

- Use case that has impacts on many working areas in NSOs
- Particularly important these days as many of them are transitioning to open-source such as R and python

Better & new products

Satellite imagery classification

Objective of this Imagery Pilot Project (Practical Application)

Expand the use of imagery data in the production of official statistics through the further development of knowledge and sharing of ML solutions and practices.

INEGI Mexico (2020)

1,975,719 (1km x 1km) grid cells

Better & new products

<u>Clustering for berth areas using Automatic Identification System</u> (AIS) data

Step1. Apply DBSCAN on high-resolution H3 hexagons (e.g., resolution level: 13)

ONS-UNECE Machine Learning Group AIS Theme Group (2022)

Step 2. Obtain a set of lower resolution H3 hexagons (parents of high-resolution H3 hexagons)

Step 3. Connect H3 polygons to create berth areas

Better & new products

Edit & imputation

Federal Statistical Office of Germany (2020)

Creating survey framework

Use case – Turkish Statistical Institute

 Project 3 – Create a framework for government R&D survey

Turkish Statistical Institute (2022)

Service delivery & user needs

Natural language-based dissemination

- Provision of statistics is fundamental role of NSOs
- Data platform is not easy to navigate
- Help users find data with natural language, improving user experience

Service delivery & user needs

Targeted communication

- Growing emphasis on tailored communication
- Image/audio/video generation
- Generation of communication materials in line with certain styles and tone

Resources on ML/AI for official statistics

use cases

JECE

modernstats

Thank you!

