



# HLG-MOS Machine Learning Project

Presented to 2020 Workshop on the Modernisation of  
Official Statistics

Claude Julien – UNECE Project Manager



The great work from a large group of professional, engaged, talented, perseverant and very pleasant people, and their numerous collaborators and supporters

**Very happily** presented to 2020 Workshop on the  
Modernisation of Official Statistics

Claude Julien – UNECE Project Manager

# Is ML just a buzz, a bust or a must?

- ML is definitely NOT just a BUZZ
- ML is a MUST... where it adds value
- ML is still up against some BUSTers (challengers and resistors)

Note: the ML group icon



in this talk points at ML project outputs.

# ML is definitely NOT a just a buzz!

- The ML group has continued to grow (40 participants and 80 other collaborators and followers)
  - Organisational changes within several NSOs
  - Some NSOs have set-up internal and external fora for data science and ML
  - Many talks, organized sessions and entire on ML
- 325+ registrants from 46 countries and 95 organisations at ML webinar



# Background

- Project objective: Advance the sound and efficient use of ML in the production of official statistics
- WP1: Pilot Studies (demonstrate value added)
  - Coding & Classification; Edit & Imputation; Imagery
- WP2: Quality
- WP3: Integration challenges

# ML is a MUST... where it adds value

- 21 study reports and 3 summary theme reports shared on public wiki last week
- On C&C, ML does deliver better quality at same/lower cost
- On E&I, ML can add value in some situations
- On Imagery, ML is essential and must be as accessible as possible
- Generic Pipeline for Production of Official Statistics Using Satellite Data and Machine Learning



# Some takeaways

- Clear business needs and/or corporate-wide strategy
- Mutual support/trust between SM/operations and ML
  - Focus on how ML can improve process results, not on “taking over” the process
- Sufficient supply of good quality data for training and monitoring
  - Good quality = accurate + representative
  - More at the start, but some throughout its use
- Sound and efficient proof of concepts
- Focus on medium/long-term efficiencies
- Combination of multiple disciplines

# ML is still up against some BUSTers (challengers and resistors)

- More ML methods have made it to production processes, but this “acceptance” is still a challenge
- Quality Framework for Statistical Algorithms (QF4SA)
- Integration challenges and practices





# Quality Framework for Statistical Algorithms (QF4SA)

- Guide official statisticians on the choice of algorithms to be used in the production process
- Statistical algorithms cover both *traditional* and *modern* statistical methods
- Accuracy, Timeliness, Efficiency, Explainability and Reproducibility
- The QF4SA is not a replacement for existing quality frameworks but is a supplement to them

# Integration of ML

- Identify and discuss challenges and practices
  - Organisation, skills, acceptance, align ML solutions to the right needs
- Multidisciplinary nature of ML
- Skill acquisition
  - Broad agreement that internal and external training courses valuable
  - Lots of great free resources on the web
  - Broad agreement that knowledge sharing, through internal communities of practice and international organizations valuable

# Conclusions

- The ML project did advance the use of ML in participating organisations
- Its shared outputs should further advance it in more organisations
- Everybody should have ML in mind as a potential solution to their business needs
  - Everybody = all functions + all levels
- Have access to a centre of expertise to assess and support the development of ML
- Have clear ownership and accountabilities in the implementation and maintenance of ML

# Recommendations

- Let the ML group continue to share and collaborate
  - ONS DSC has come forward to “lead” the group
  - Need continued support from the Heads of NSOs (HLG-MOS)
  - Stay connected to the HLG-MOS Exec
- Heads of NSO data science units should create a formal network to share management practices
- The ML group should continue to communicate and share with other groups looking into other aspects or other applications of ML within their respective broader agenda

# I thank the ML group for all their great work

- Thank you
- Merci
- Tak
- Dank u
- Kiitos
- Danke
- Takk
- Grazie
- 감사합니다
- Takk
- Dziękuję
- Obrigado
- Хвала / Hvala
- Gracias
- Tack
- Mauruuru

I thank the HLG-MOS and the Exec Board for giving me this very pleasant and enriching opportunity