



UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE CONFERENCE OF EUROPEAN STATISTICIANS

Machine Learning for Official Statistics Workshop

5-7 June 2023, Geneva, Switzerland Palais des Nations (Room VII)

VERSION: 5 June 2023

Provisional Timetable

DAY 1 - Monday 5 June

8:30-9:30	Registration		
9:30-9:40	Welcome – InKyung Choi (UNECE)		
	Session 1: Machine Learning Applications		
Session Ch	Session Chairs: Michael Reusens (Statistics Flanders, Belgium) and Joni Karanka (ONS, UK)		
9:40-9:45	Introduction to Session – Session Chairs		
9:45-10:00	Classifying companies in France using machine learning - Thomas Faria and Tom Seimandi (Insee, France)		
10:00-10:05	Q&A		
10:05-10:20	Using Webdata to derive the Economic Activity of Enterprises - Manveer Mangat (Statistics Austria)		
10:20-10:25	Q&A		
10:25-10:40	Clothing Price Index using Web-Scraped Data – Laura Christen and Ahmet Aydin (ONS, UK)		
10:40-10:45	Q&A		
10:45-11:05	Break		
11:05-11:20	Imputation of occupation in the Occupational Register - Jens Malmros (Statistics Sweden)		
11:20-11:25	Q&A		
11:25-11:40	Too good to be true? Machine learning in the editing process - Eva Charlotte Berner and Solveig Bjørkholt (Statistics Norway)		
11:40-11:45	Q&A		
11:45-12:00	Data-driven earthquake multi-impact modelling: A comparison of approches - Hamish Patten (University of Oxford)		





	·
12:00-12:05	Q&A
12:05-12:20	Progression patterns in the Swiss social security system based on Machine Learning: methods for evaluating quality and model drift - Athanassia Chalimourda (Swiss Federal Statistics Office)
12:20-12:25	Q&A
12:25-13:40	Lunch
13:40-14:10	Soapbox session
14:10-14:25	ML Poverty: Using Machine Learning to estimate poverty rates in Switzerland at the canton level - Yara Abu Awad (Swiss Federal Statistics Office)
14:25-14:30	Q&A
14:30-14:45	Creating a modern business index: Machine learning record linkage at scale - Isabela Breton and Joanne Sheppard (ONS, UK)
14:45-14:50	Q&A
14:50-15:10	Break
15:10-15:25	Time Series Outlier Detection using Metadata and Data Machine Learning in Statistical Production - Olivier Sirello (BIS)
15:25-15:30	Q&A
15:30-15:45	Timeliness and Accuracy with Machine Learning Algorithms: Early Estimates of the Industrial Turnover Index - David Salgado (Statistics Spain)
15:45-15:50	Q&A
15:50-16:05	Nowcasting TiVA indicators: improving timeliness of trade data - Polina Knutsson (OECD)
16:05-16:10	Q&A
16:10-17:00	Small group discussion and reporting back Read descriptions of topics and select topic you are interested – scan below QR or use this https://tinyurl.com/ywxm3444
17:00-17:05	Closing Day 1





DAY 2 - Tuesday 6 June

	Session 2: Quality Aspects of Machine Learning in Official Statistics		
Session Chairs: Florian Dumpert (Federal Statistical Office of Germany) and Ralf Becker (UNSD)			
9:30-9:35	Introduction to Session – Session Chairs		
9:35-9:45	Quality Framework for Statistical Algorithms – InKyung Choi (UNECE)		
9:45-10:00	A Quality Concept for the Use of Machine Learning in Official Statistics - Florian Dumpert (Federal Statistical Office of Germany)		
10:00-10:20	Exploring quality dimensions in trustworthy Machine Learning in the context of official statistics: model explainability, accuracy and uncertainty - Saeid Molladavoudi (Statistics Canada) [remote]		
10:20-10:30	Q&A		
10:30-10:45	Understanding model quality in the context of trustworthiness and value - Emily Barrington (Office for Statistics Regulation, UK) [remote]		
10:45-10:50	Q&A		
10:50-11:05	Break		
11:05-11:20	Lessons learned when applying Machine Learning in Official Statistics: Why it helps to be a survey statistician and a data scientist! - Piet Daas (Statistics Netherlands)		
11:20-11:35	Changing Data Sources in the Age of Data Science for Official Statistics - Cedric De Boom (Statistics Flanders, Belgium)		
11:35-11:45	Q&A		
11:45-12:25	Small group discussion and reporting back Read descriptions of topics and select topic you are interested – scan below QR or use this https://tinyurl.com/ywxm3444		
12:25-13:30	Lunch		





Session 3: Toward system-wide transformation of statistical production		
Session Chairs: Riitta Piela (Statistics Finland) and Dominika Nowak (Statistics Poland)		
13:30-13:35	Introduction to Session – Session Chair	
13:35-14:05	Keynote presentation: How to Manage Artificial Intelligence and Data Science as a Process of Continuous Improvement in the Context of Official Statistics? – Prof. Diego Kuonen (Statoo Consulting & University of Geneva, Switzerland)	
14:05-14:10	Q&A	
14:10-14:25	Facilitators and Blockers of ML Adoption in Official Statistics - Joni Karanka (ONS, UK)	
14:25-14:30	Q&A	
14:30-14:45	A Machine Learning Capability Uplift Strategy - Claire Clarke (ABS, Australia) [remote]	
14:45-15:00	ML training: Who? What? How? and What for? - Christophe Bontemps (UN SIAP, ESCAP) [remote]	
15:00-15:10	Q&A	
15:10-15:30	Break	
15:30-15:45	Balsam: A Collaborative Platform to Support ML and ML-Ops initiatives - Jakob Engdahl (Statistics Sweden)	
15:45-16:00	An open-source data science platform to foster innovative and production-ready machine learning systems - Romain Avouac (Insee, France)	
16:00-16:10	Q&A	
16:10-17:00	Small group discussion and reporting back Slido poll for Session 3 - scan below QR or use this https://tinyurl.com/4u8v9yf2	
17:00-17:05	Closing Day 2	





DAY 3 – Wednesday 7 June

Session 3: Toward system-wide transformation of statistical production (continue) Session Chairs: Riitta Piela (Statistics Finland) and Dominika Nowak (Statistics Poland)		
9:30-9:35	Opening	
9:35-11:05	Hands-on Lab: An introduction to MLOps with Mlflow (Part I) - Tom Seimandi, Romain Avouac and Thomas Faria (Insee, France)	
11:05-11:25	Break	
11:25-12:55	Hands-on Lab: An introduction to MLOps with Mlflow (Part II) - Tom Seimandi, Romain Avouac, Thomas Faria (Insee, France)	
12:55-13:05	End of workshop	