Microsoft Research
Summit 2022

Food Security Workshop

Nov 9 – Nov 10, 2022

Organized by: Ranveer Chandra, B. Ashok, Jiang Bian, Riyaz Pishori

Housekeeping

- This workshop is being recorded.
- The recording will be made available for on-demand viewing one week after the event.
- Please use the Q&A icon for engagement with the speakers to allow for extended dialogue while keeping the sessions on time.
- When posting a question, please note you have the option to post with your name or anonymously.

Agenda

Wednesday, November 9: Modern R&D

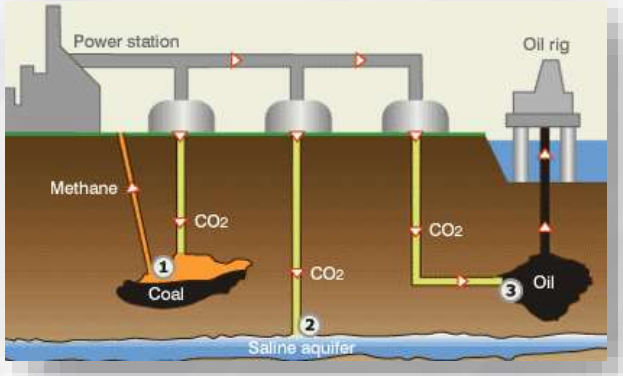
Time (Pacific Time)	Session
5:00 PM	**Opening** Jiang Bian, *Microsoft Research Asia*
5:05 PM	**Recent AI advances for accelerating R&D** Jiang Bian, *Microsoft Research Asia*
5:20 PM	**Confidential clean rooms for privacy preserving inference and model Evaluation** Satya Lokam, *Microsoft Research India*
5:35 PM	**Developing ML models to approximate and predict food protein digestibility** Sara Malvar, *Microsoft Research*
5:55 PM	**Applications of AI and ML in food process engineering: state of the art** Ronit Mandal, *Plant Protein Innovation Center (PPIC)*
6:20 PM	**Generative models and transformers for chemistry** Morris Sharp, *Microsoft Research Redmond*
6:35 PM	Break
6:45 PM	**Introduction to ag-genomics** Angels de Luis Balaguer, *Microsoft Research Redmond*
7:15 PM	**Harnessing the power of AI for animal phenotyping** Joao Dorea, *University of Wisconsin Madison*
7:55 PM	**Leveraging AI to improve crop production** Swati Sharma, *Microsoft Research Redmond*
8:25 PM	**Trustworthy supply chains using DNA tags** Yuan-Jyue Chen, *Microsoft*
8:40 PM	**Closing** Ranveer Chandra, *Microsoft Research Redmond*

Microsoft Research
Summit 2022

Recent AI Advances for Accelerating R&D

Jiang Bian (Microsoft Research Asia)

Research for Industry in MSR



Decarbonizing energy through affordable CCS

Data-driven agri-food systems

Multi-agent optimization for connected supply chains

Cloud-hosted financial exchanges

Real-time simulations for Grid 2.0

Scalable content creation & distribution for media & entertainment

R&D Innovation Pillars Across Industries

Last mile & Cloud/Edge compute

Data models & data sharing

Simulations & AI

Industry Sustainability

Key AI Techniques to Accelerate R&D

Supply Chain

- Resource Repositioning
- Future Warehouse
- Last mile delivery
- Route optimization
- Inventory Management
- Order fulfillment
- Demand Forecasting

CPG-Food

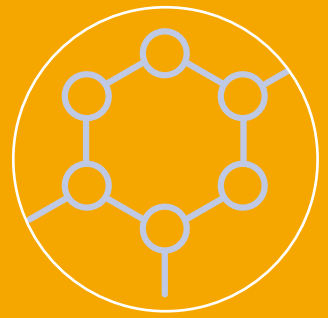
- Food Genomics
- Soil Carbon
- Food Transparency
- Food Production
- High Throughput Phenotyping

Energy

- Subsurface optimization
- Carbon sensing & estimation
- Grid/renewables optimization
- Virtual Battery
- Smart facilities

FSI

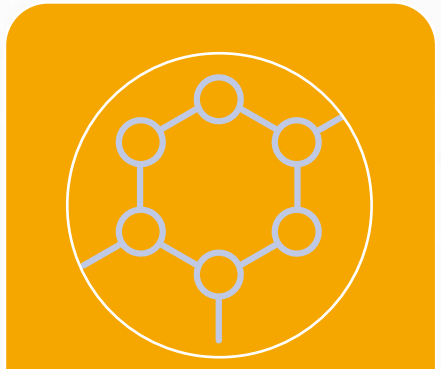
- AI based trading
- Cloud Hosted Exchange
- Investment asset rating
- Data sharing
- Intelligent RegTech
- Risk simulation and control



Sustainability

- Emission tracking
- Climate Adaptation
- Carbon Sequestration

Key AI Techniques to Accelerate R&D



Advanced forecasting

FOST Forecasting Open Source Tool

https://github.com/microsoft/fost

Large-scale optimization

MARO

https://github.com/microsoft/maro

Deep & Offline RL
[NeurIPS'21, ICLR'22]

Causal analysis & inference

DoWhy library

https://github.com/microsoft/dowhy

EconML

https://github.com/microsoft/econml

Simulation & back-testing

Qlib

https://github.com/microsoft/qlib

CO_2 Sim

DeepMC

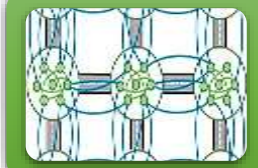
Key Challenges in Forecasting

Noisy, irregular, and missing data

→ **Data-driven imputation, alignment and normalization**

Rich spatial & temporal patterns

→ **A hybrid NN structure: GNN + RNN**



Complex signal with multi-granularity

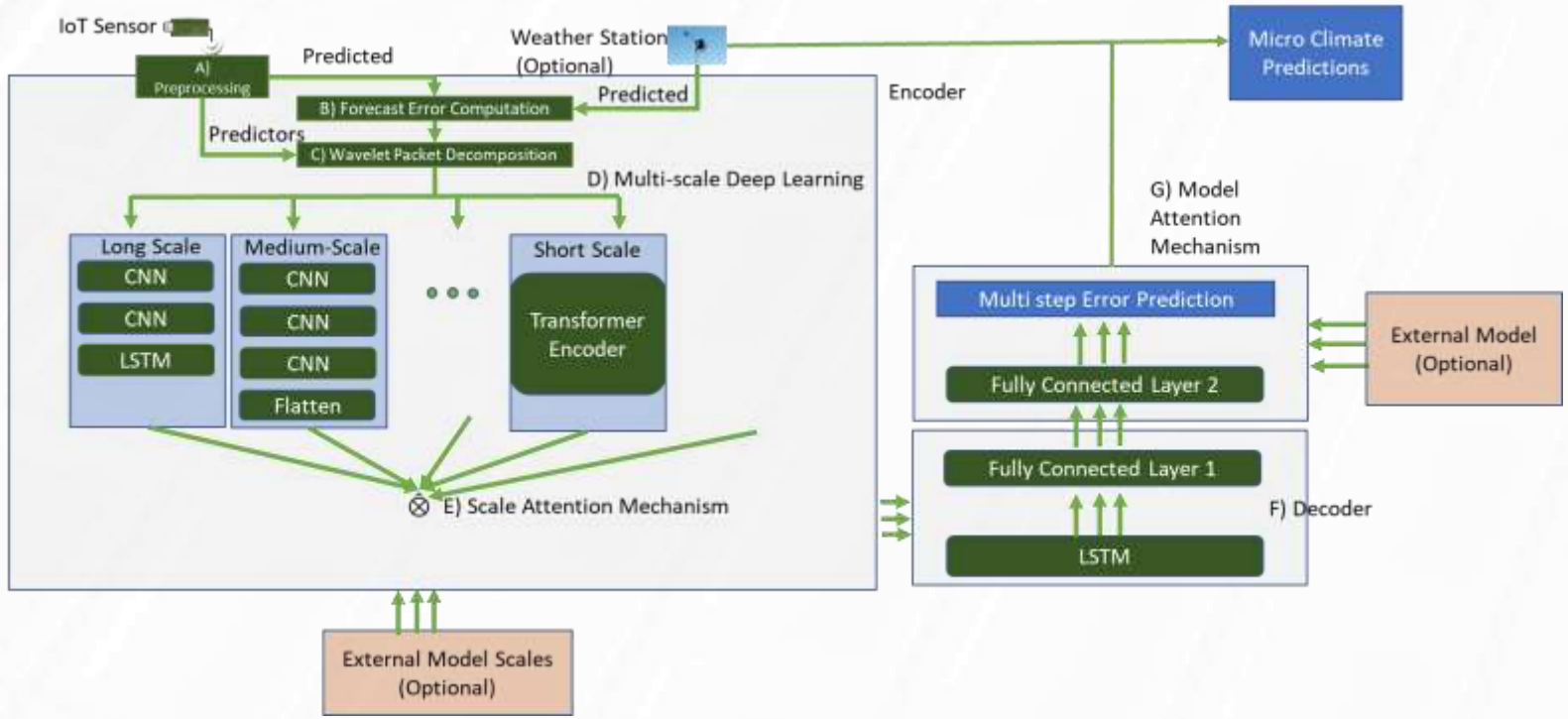
→ **Multi-scale learning**

Advanced Forecasting Framework - DeepMC

Learning-based imputation and **time-gated alignment** to address missing and irregular time series data.

Both **temporal (LSTM, Transformer) & spatial (GNN) modules** to capture rich temporal patterns & spatial dependency.

Multi-scale (micro-meso) learning for extracting complex signals of multi-granularity.



Cases of DeepMC

Hornsdale : Wind Forecasting

Solar Farm: Micro-radiation prediction

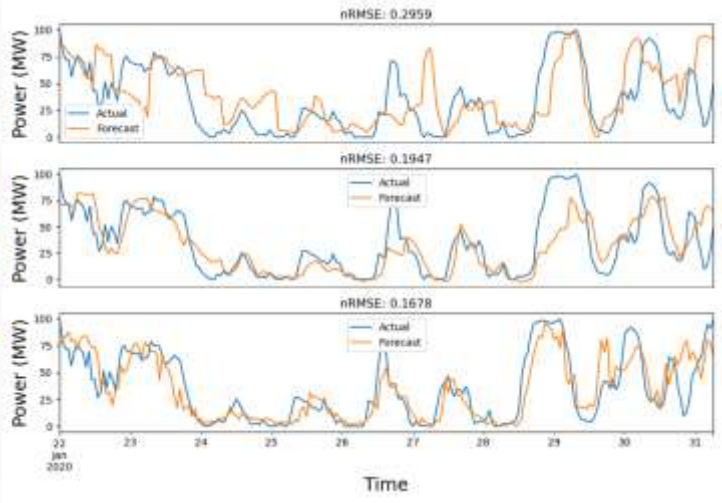
TFT w/o External Forecast

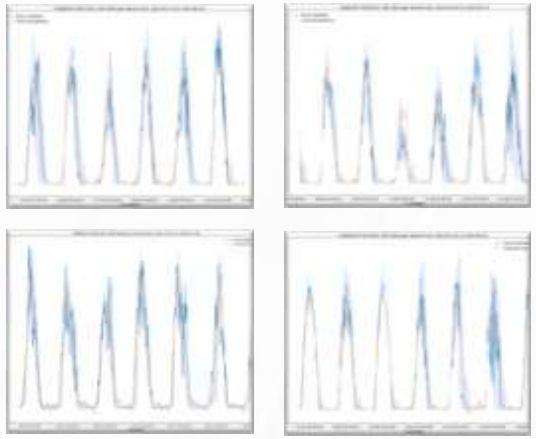
34% relative improvement

TFT Fusion w External Forecast

14% relative improvement

DeepMC Fusion w External Forecast





	DeepMC	CNN	LSTM	CNN-LSTM	ARIMA
RMSE	124.5	167.4	192.3	155.6	530.6
MAE	68.15	111.77	130.99	90.02	397.45
MASE	1.95	3.2	3.75	2.89	11.39

Key Challenges in Large-Scale Optimization

Delayed Feedback & multiple objectives → **RL with reward shaping**

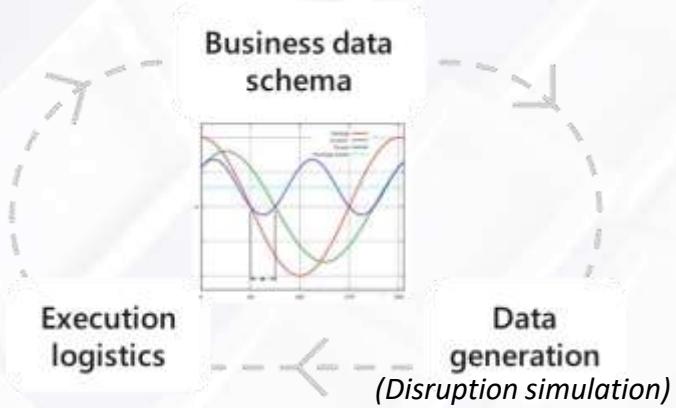
Huge cost for online interaction → **Continual offline RL**

Large-scale problem with high complexity of global optimization → **Distributed multi-agent RL**

Multi-agent Optimization Framework

MARO

https://github.com/microsoft/maro



Continual offline RL

Optimistic policy π_{optim}

Pessimistic policy π_{pessim}

$$\pi_{mix} = (1-\alpha)\pi_{pessim} + \alpha\pi_{optim}$$

Distributed multi-agent RL framework to solve large-scale global optimization

Contextualized Training

Scale up MARL under the coopetitive environment

Balanced Task Dispatching

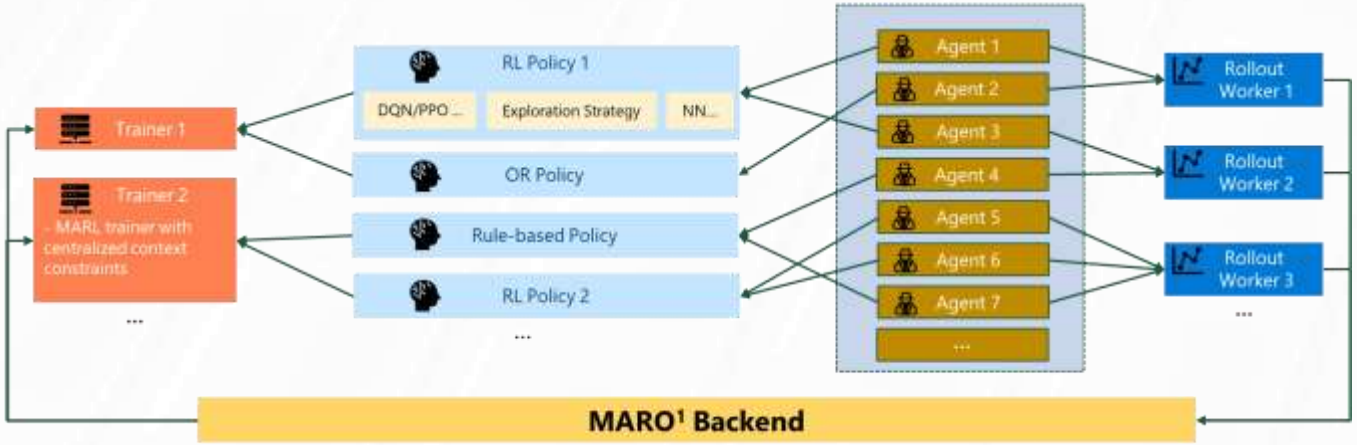
Balanced training workload in the distributed learning system

Flexible Policy Mapping

Tradeoff between learning efficiency and generalization

Hybrid Experience Management

Efficient management of massive offline experience data



Cases of MARO

Food products flow from producer to customers and consumers

Application Domain	Production & Processing Scheduling	Resource Repositioning	Inventory Management	Routing & Order Fulfillment
Optimization Task	Scheduling orders to factories to minimize total production cost	Reposition resources efficiently to accommodate imbalanced and dynamic supply and demand.	Optimize replenishing decisions to balance supply and demand.	Optimize on-call stops routing and fulfillment strategy to minimize total cost.
Results		*Used by **50%** of ports operated by OOCL* ***1000 times** speed-up of decision* ***10%** operational cost saving.*	*Scale up to **1000+** SKUs.* ***150 times** speed-up of decision efficiency* ***10%** profit increment in a POC experiment*	*Experiments a dispatch center in San Diego:* ***10%** increase of fulfillment ratio* ***Instant** decision making*

OOCL Walmart+ FedEx

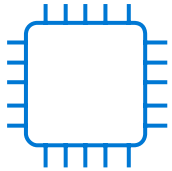
Key Challenges in Causal Analysis & Inference

Answer "What-if" questions from **observational data at scale.**

~~High Cost~~
Low Cost

~~Small Scale~~
Large Scale

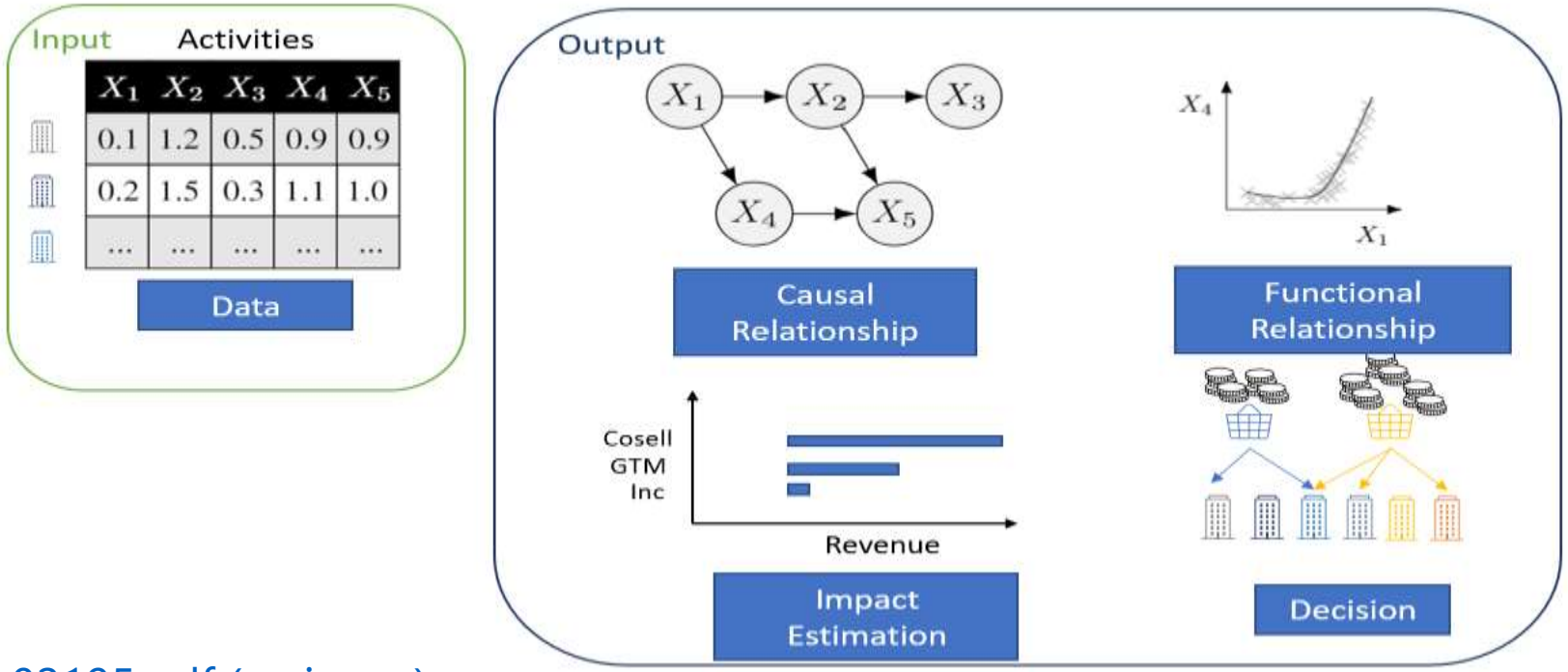
~~Long Time~~
Real Time



~~Low Resolution~~
High Resolution

New Tech for Causal Inference

Causal Discovery + Causal Inference + Deep learning → Real world impact



DoWhy library

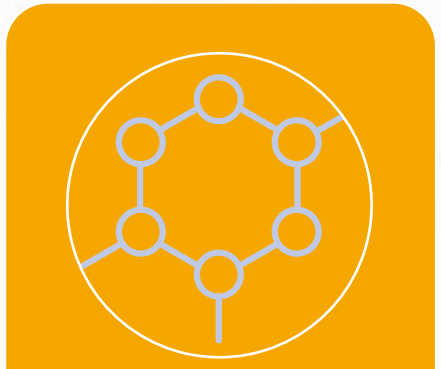
https://github.com/microsoft/dowhy

EconML

https://github.com/microsoft/econml

2202.02195.pdf (arxiv.org)

Key AI Techniques to Accelerate R&D



Advanced forecasting

FOST Forecasting Open Source Tool

https://github.com/microsoft/fost

Large-scale optimization

MARO

https://github.com/microsoft/maro

Deep & Offline RL
[NeurIPS'21, ICLR'22]

Causal analysis & inference

DoWhy library

https://github.com/microsoft/dowhy

EconML

https://github.com/microsoft/econml

Simulation & back-testing

Qlib

https://github.com/microsoft/qlib

CO_2 Sim

DeepMC

Microsoft Research
Summit 2022

Thank you

Stay in touch:

linkedin.com/in/jbian

jiang.bian@microsoft.com