Accelerated Data Science, AI and GeoAI for Sustainable Finance in Central Banking and Supervision

Int. Conference on Statistics for Sustainable Finance (14th & 15th Sept 2021)

Session 5: Leveraging Innovation

- Dr. Jochen Papenbrock, NVIDIA GmbH, Germany (jpapenbrock@nvidia.com)
- Dr. John Ashley, NVIDIA Corp., USA (jashley@nvidia.com)
- Professor Dr. Peter Schwendner, Zurich University of Applied Sciences, Switzerland (scwp@zhaw.ch)

LEVERAGING INNOVATION

Abstract

- Sustainability and Data Science / Artificial Intelligence (AI) are the two megatrends in FSI and supervision
- Use of massive amount of complex data
- Demand for computing platforms, HPC, and ‘AI’
- We see systems with data centric design, explainable/trustworthy AI models, GPU acceleration as key feature,
- Practical examples in ESG investing and data-driven policy design and investment/risk management

TRANSFORMED SUPERVISORY MODELS

Supervisory Technology (SupTech)

- synthetizes the vast quantities of structured and unstructured data
- improves actuality and timeliness (‘real-time’) of (emerging) risk identification, monitoring, early warning/intervention.
- supports both ‘full picture’ as near cast and granular, zooming level
- The same applies to Sustainable Finance Technology

NEED FOR COMPUTING PLATFORM

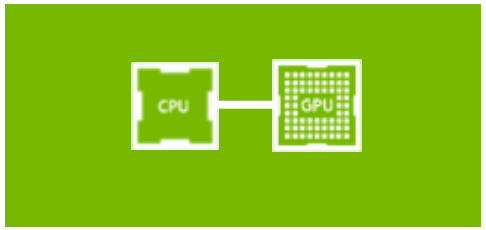
Establishment of computing platforms going hand in hand with the data collection and curation process. Several factors drive this need for a co-evolutionary approach:

1. Data, IT systems, operations and policies must be fully aligned and integrated:
   adaptive, auditable process to evolving data quantity and quality; validation, interpretability, narrative
2. filling data gaps and improve data quality (outliers)
3. HPC for simulating data and for complex optimization
4. large scale visualization, clustering and network analysis for data exploration
5. size of data and number of sources will increase, and so will the number, complexity, and frequency of updates of models

! Compliance with the proposed AI Act for high-risk AI might also require a computing platform !

RISE OF GPU COMPUTING

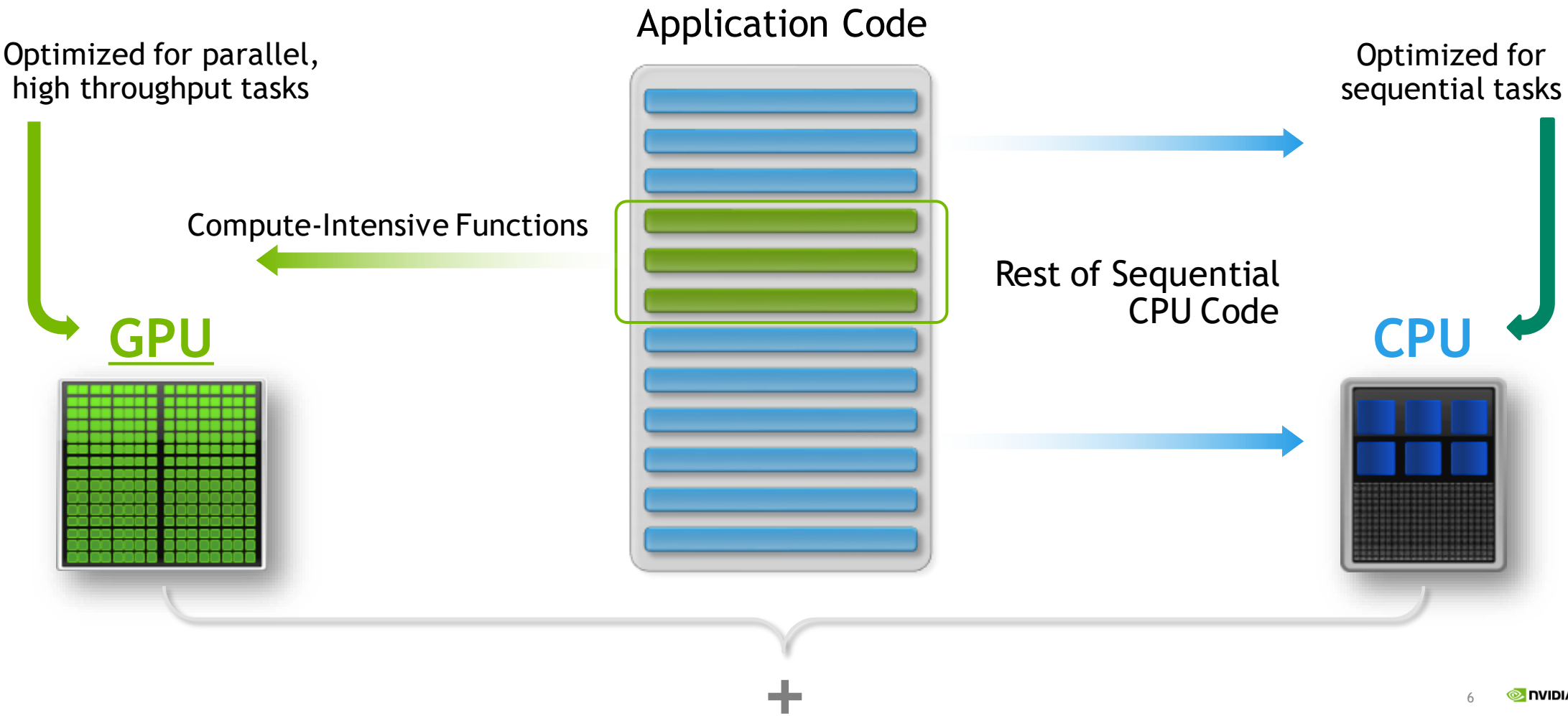
'Computing platforms for big data analytics and artificial intelligence.'
IFC Reports 11. Bank for International Settlements.
https://ideas.repec.org/p/bis/bisifr/11.html



"Central banks' experience shows that HPC platforms are primarily developed to ensure that computing resources are used in the most efficient way, so that analytical processes can be completed as rapidly as possible. […]

A processor core (or "core") is a single processing unit. Today's computers – or CPUs (central processing units) – have multiple processing units, with each of these cores able to focus on a different task. Depending on the analytical or statistical problem at hand, clusters of GPUs (graphics processing units, which have a highly parallel structure and were initially designed for efficient image processing) might also be embedded in computers, for instance, to support mass calculations"

HOW GPU ACCELERATION WORKS



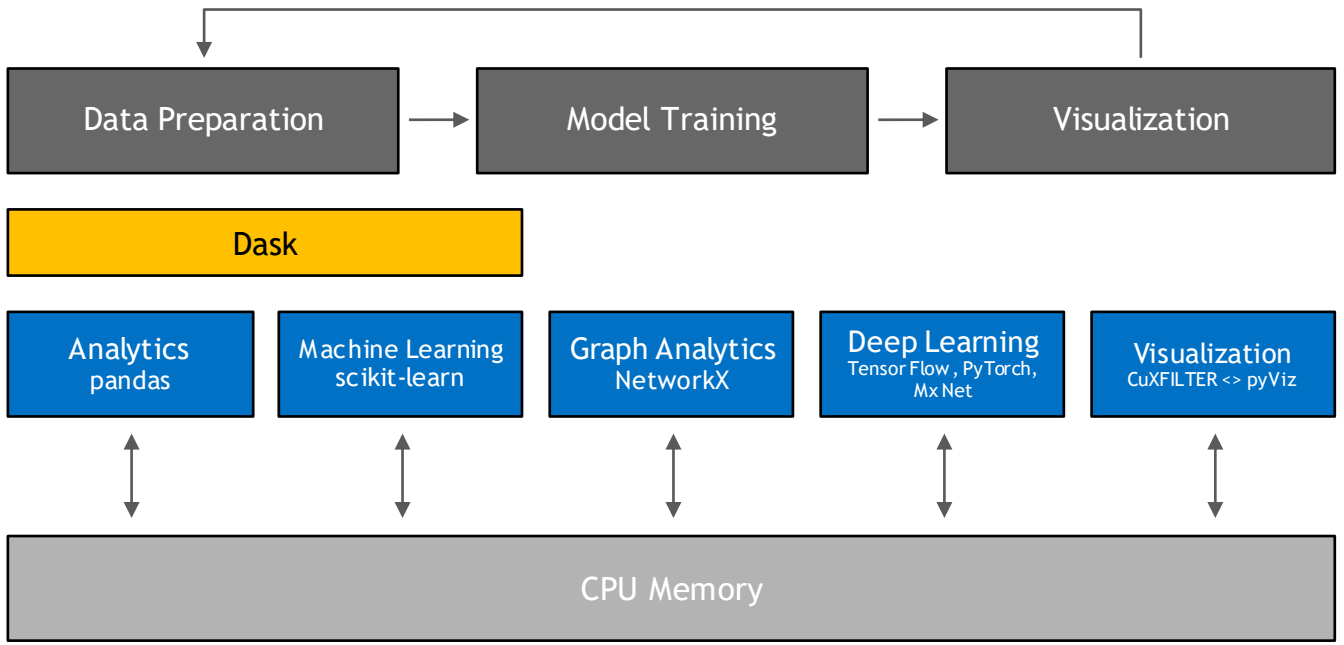
PYTHON TOOLS HAVE DEMOCRATIZED DATA SCIENCE

ACCESSIBLE, EASY TO USE TOOLS ABSTRACT COMPLEXITY

Python is the most-used language in Data Science today. Libraries like NumPy, Scikit-Learn, and Pandas have changed how we think about accessibility in Data Science and Machine Learning.

While great for experimentation, PyData tools lack the power necessary for enterprise-scale workloads. This leads to substantial refactoring to handle the size of modern problems, increasing cycle time, overhead, and time to insight.

These pain points are further compounded by computational bottlenecks of CPU-based processing.



Code refactors and inter-team handoffs decrease data-driven ROI

RAPIDS

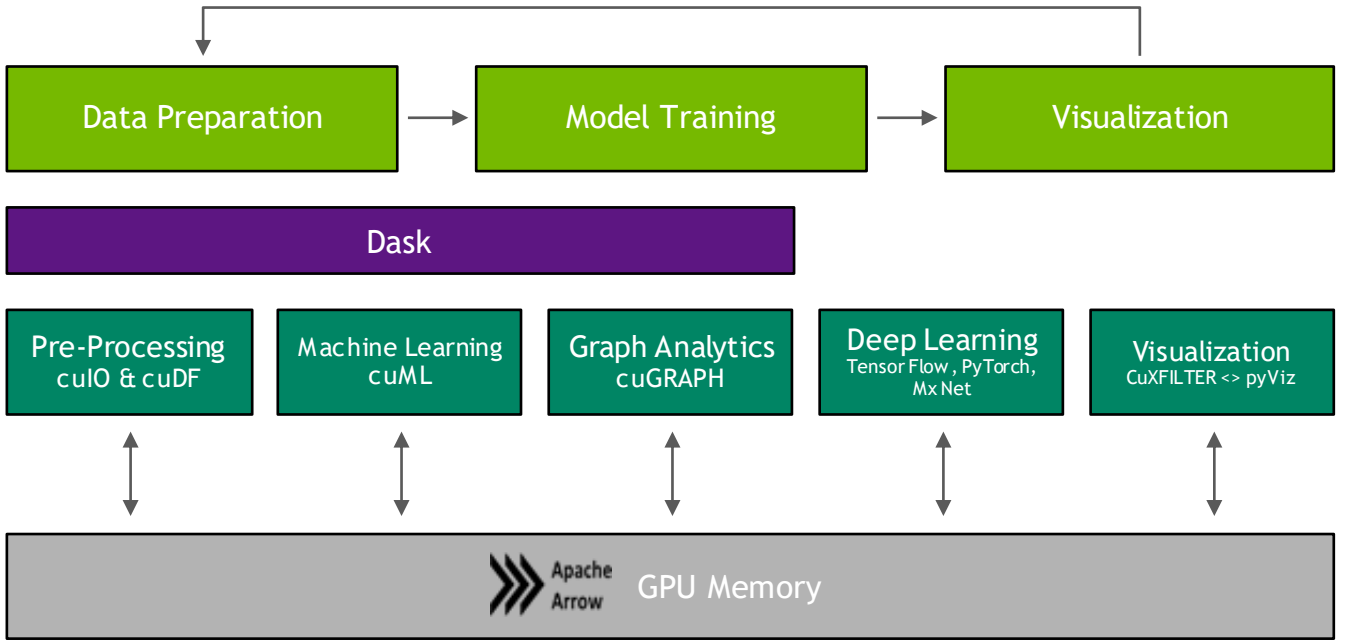
RAPIDS ACCELERATES POPULAR DATA SCIENCE TOOLS

DELIVERING ENTERPRISE-GRADE DATA SCIENCE SOLUTIONS IN PURE PYTHON

The RAPIDS suite of open source software libraries gives you the freedom to execute end-to-end data science and analytics pipelines entirely on GPUs.

RAPIDS exposes GPU parallelism and high-bandwidth memory speed through user-friendly Python interfaces like PyData.

With Dask, RAPIDS can scale out to multi-node, multi-GPU cluster to power through big data processes.

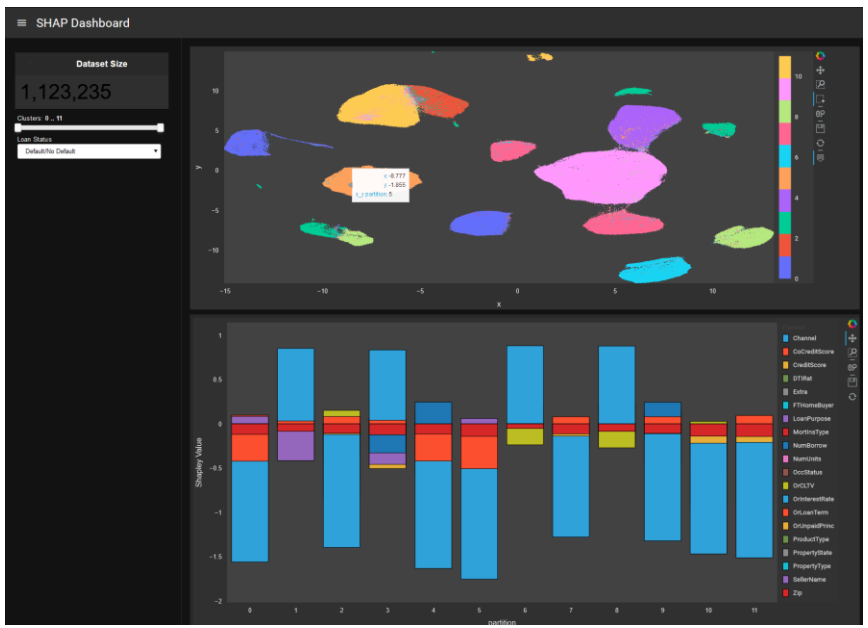


RAPIDS enables the PyData stack with the power of GPUs

XAI USE CASE ON ENTERPRISE PRODUCTION LEVEL

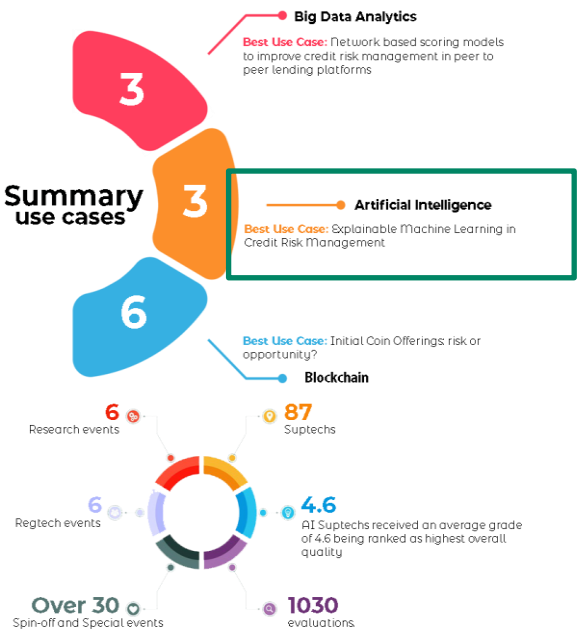
NVIDIA acceleration of XAI use case accelerated with RAPIDS

- Use case on credit portfolio risk with realistic data (11.2 million dataset from Fannie Mae)
- General speedups of up to 19x for SHAP values, and 340x for SHAP interaction values
- use case based on best AI use case in EU Horizon2020 project FIN-TECH *)



FINTECH RISK MANAGEMENT
www.fintech-ho2020.eu

European Commission

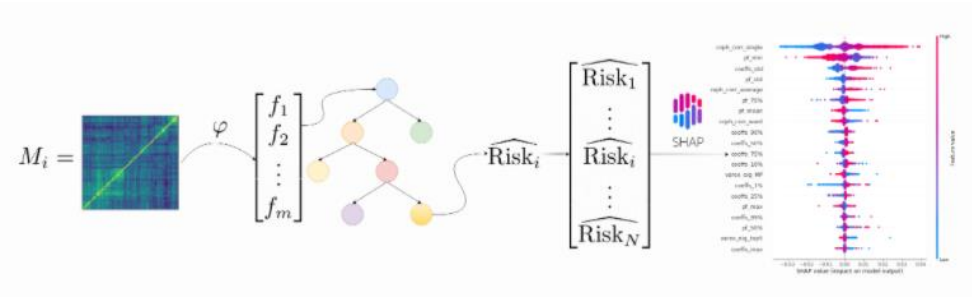


This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 825215 (topic ICT-35-2018, Type of action: CSA)The content reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

NVIDIA

XAI FOR PORTFOLIO CONSTRUCTION

- Collaboration with Munich Re
- 3 Publications in top Journal of Financial Data Science
- Implementing a new workflow and combine XAI with synthetic market data generation to enhance the explainability
- Additional paper with cryptos is in the pipeline:
  “Can adaptive seriational risk parity tame crypto portfolios?”



SUSTAINABLE FINANCE AND ESG INVESTING

Problem

- assessing the appropriate level of protection against sustainability risks is a challenge for central banks
- existing sustainability ratings from different providers diverge significantly for the same investable asset (Berg et al. 2019). This implies risk for greenwashing.
- also conflicting findings of ESG alpha (Bruno et al. 2021 vs Giese et al. 2021);

➡ calling for more scientific rigour in ESG evaluation studies (Edmans 2021)

Solution

- consider full set of data and information from multiple source
- process them with explainable AI/NLP models and aggregate them in a large-scale, transparent portfolio construction and optimization process
- This is where the two megatrends, Sustainable Finance and AI, that will shape the financial sector in the coming years overlap; computing platforms will play a crucial role

EMPOWERING REGIONS REGARDING THEIR ENVIRONMENTAL EXPOSURES (1/3)

Use Case

Purpose: empower regions to jointly tackle environmental challenges; identify regions with similar environmental exposures (climate, deforestation, land use, population, etc.)

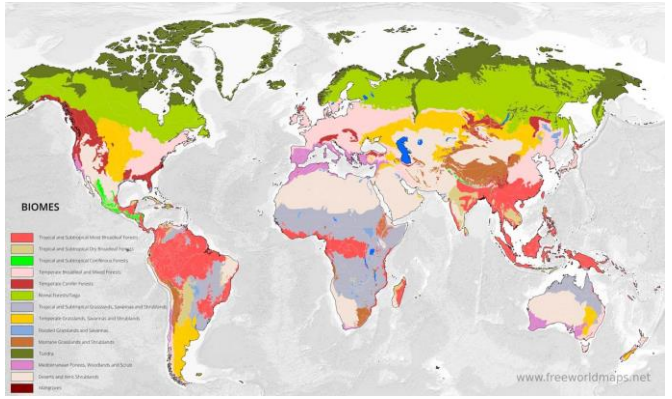
- Why?
  - improve and benchmark ESG data
  - exchange experiences
  - improve their negotiation leverage
  - support each other in joint projects or ad-hoc emergencies like floods or wildfires
  - decrease insurance and funding cost for public or private purposes
  - monitor transition and physical risk

EMPOWERING REGIONS REGARDING THEIR ENVIRONMENTAL EXPOSURES (2/3)

Use Case

Data

- Geospatial environmental data on land cover, land use, climate change measurements, etc.
- data layers derived from satellite measurements have a raster format and need to be connected and enriched with relevant country and sector data (e.g. also with economic and financial data)
- potential sources: https://globalforestwatch.org, https://earthdata.nasa.gov/, https://www.restor.eco/, etc.
- the data are of varying global coverage, spatial resolution, spatial granularity and time resolution across datasets
- Non-trivial task



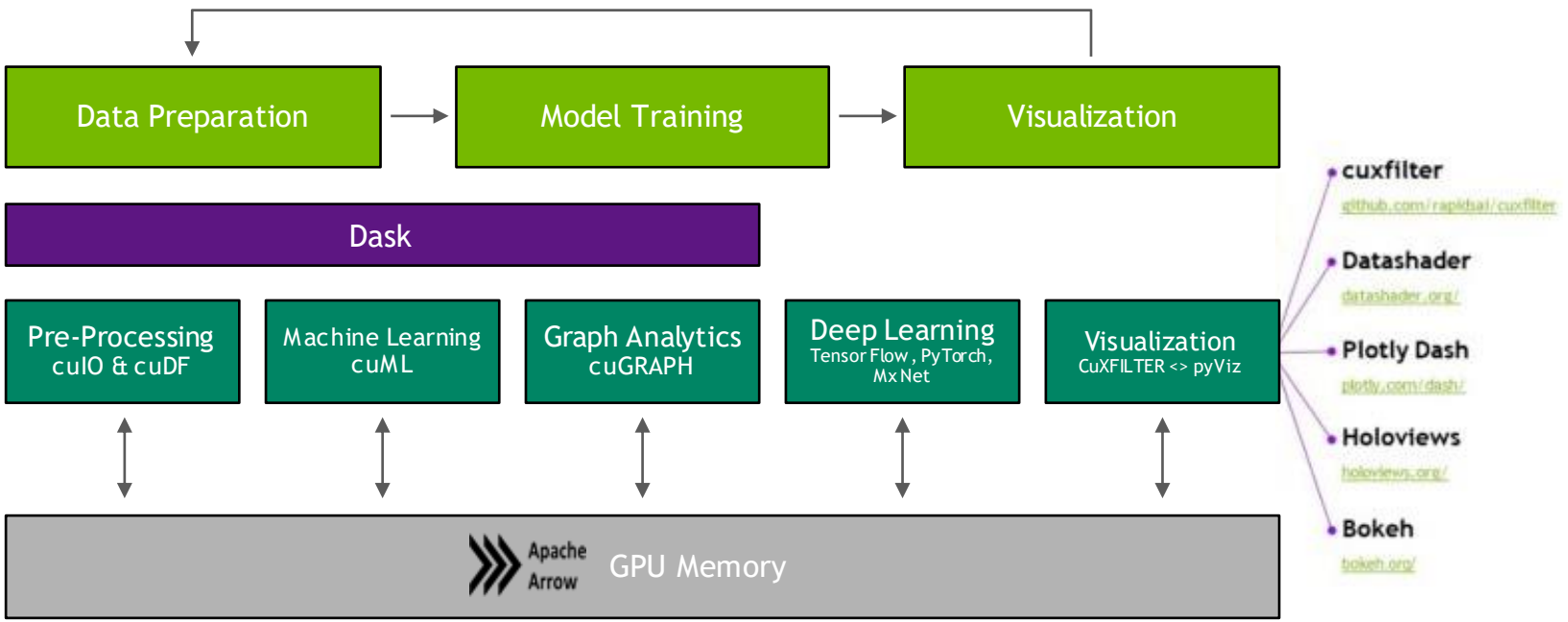
Method

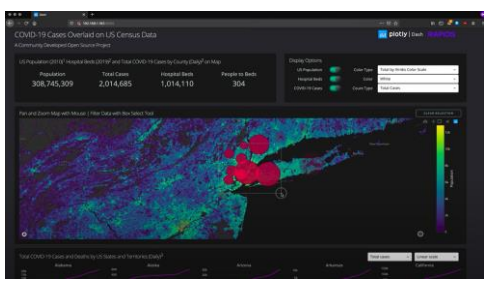
- Similarity Clustering, Network Analysis, Visualization and real-time cross-filtering, typical RAPIDS use case

EMPOWERING REGIONS REGARDING THEIR ENVIRONMENTAL EXPOSURES (3/3)

Use Case

Potential implementation with RAPIDS





NVIDIA

OUR ENGAGEMENT IN XAI IN FINANCIAL SERVICES

Webinar series on XAI

FINTECH RISK MANAGEMENT
www.fintech-h2020.eu
European Commission

Extension of a XAI use case of EU Horizon2020 project

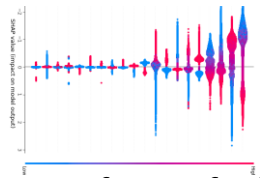
gaia-x

Initiated a project on XAI

frontiers
in Artificial Intelligence | Artificial Intelligence in Finance

Editorial and publishing activity

dmlc
XGBoost



Accelerating open-source software for XAI

NVIDIA