

End-to-End Machine Learning with High Performance and Cloud Computing

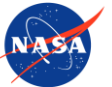
Lecture 2: Introduction and Motivations

Manil Maskey, Gabriele Cavallaro, Iksha Gurung, Muthukumaran Ramasubramanian, Rocco Sedona

NASA Interagency Implementation and Advanced Concept Team
Jülich Supercomputing Centre

High-Performance and Disruptive Computing in Remote Sensing
IEEE Geoscience and Remote Sensing Society
Earth Science Informatics Technical Committee

Tutorial, IGARSS 2022
July 17, 2022



Motivations

- Provide technical guidance on performing end-to-end machine learning use case for remote sensing
- Utilize hybrid: HPC+cloud computing to optimize cost and efficiency
- Promote open science via collaboration
- Develop machine learning expertise within remote sensing community
- Provide a venue for sharing experiences and lessons learned
- Promote collaboration amongst machine learning experts, domain experts, and software developers

Expected outcomes

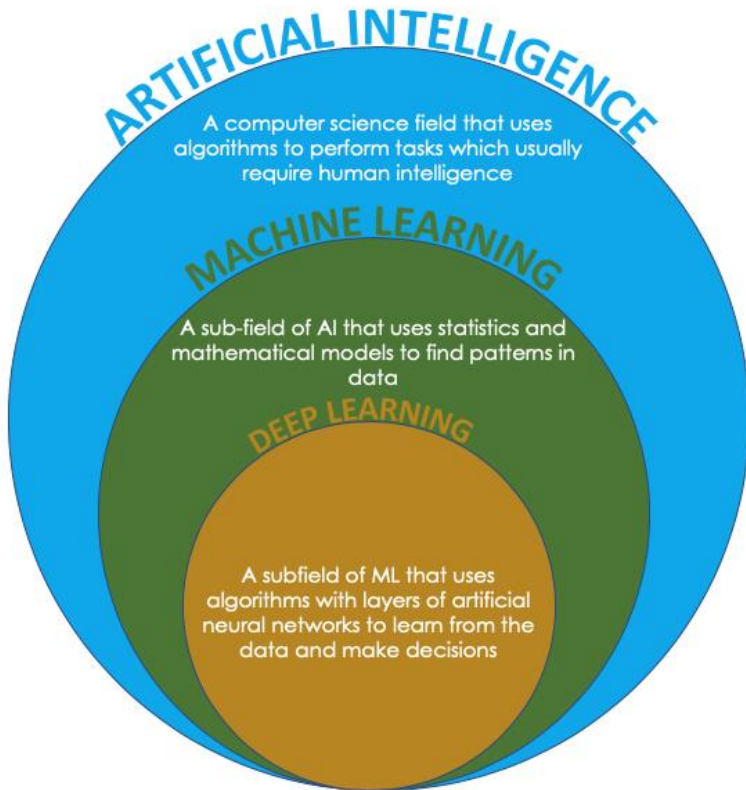
Participants are expected to:

- Learn the fundamentals of end to end machine learning life cycle
- Design, implement, deep learning models in HPC and deploy in cloud

Everyone is expected to:

- Exchange ideas
- Foster collaboration

Machine Learning



Rapid adoption of ML due to:

- Large data volumes
- Advanced algorithms
- Networks
- Cloud computing
- Hardware

Cloud computing

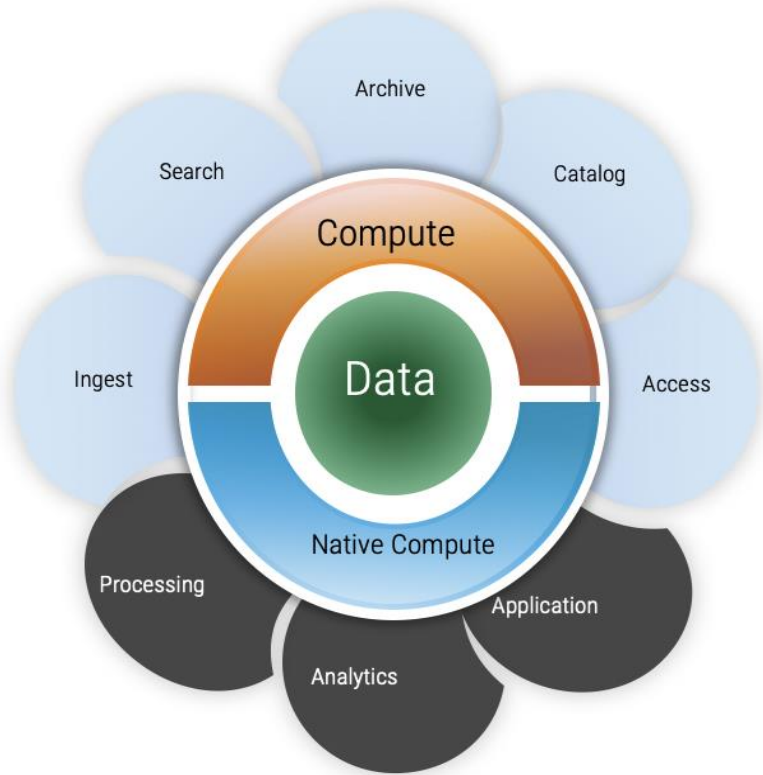
On-demand computing services: storage, compute, software,... as a service

- Amazon web services, Azure, Google cloud

Common characteristics:

- Elasticity: the ability to scale resources both up and down as needed
- Reliability: implies that the service is available and works as intended
- Pay as you go: only users pay for what they use
- Resource pooling: allows a cloud provider to serve its users in a multitenant model
- Minimal management effort: users can use and procure cloud services without much difficulty

Cloud computing



- **Big data close to compute**
- **Data storage**
- **Scalable compute**
- **Cloud native**

Cloud computing services

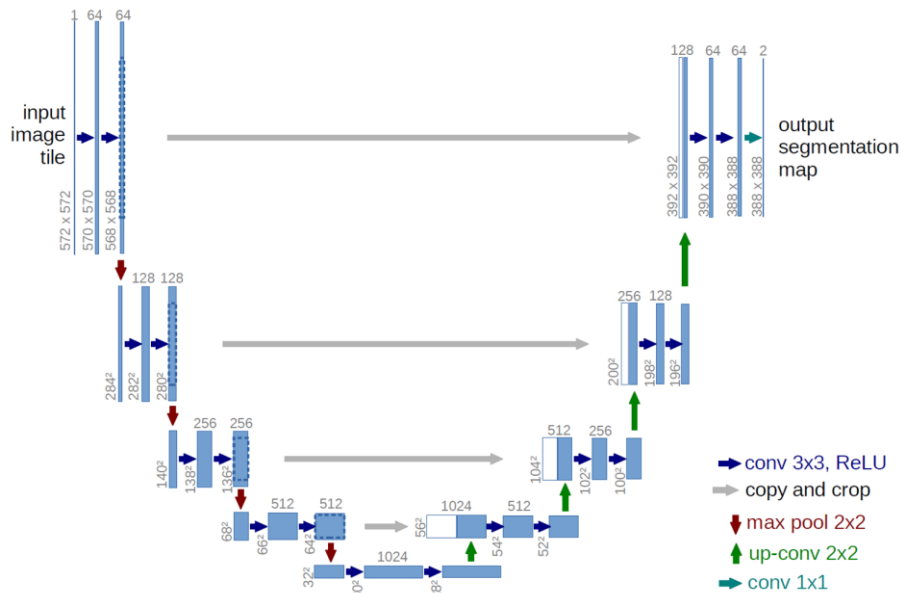
SaaS	Browser	Cloud Applications User interface, Reporting, Content management
PaaS	Development Environment	Cloud Platform Programming languages, Editors, Frameworks
IaaS	Console	Cloud Infrastructure Servers, Storage, Load balancers

Practical use case

Smoke detection



Smoke detection



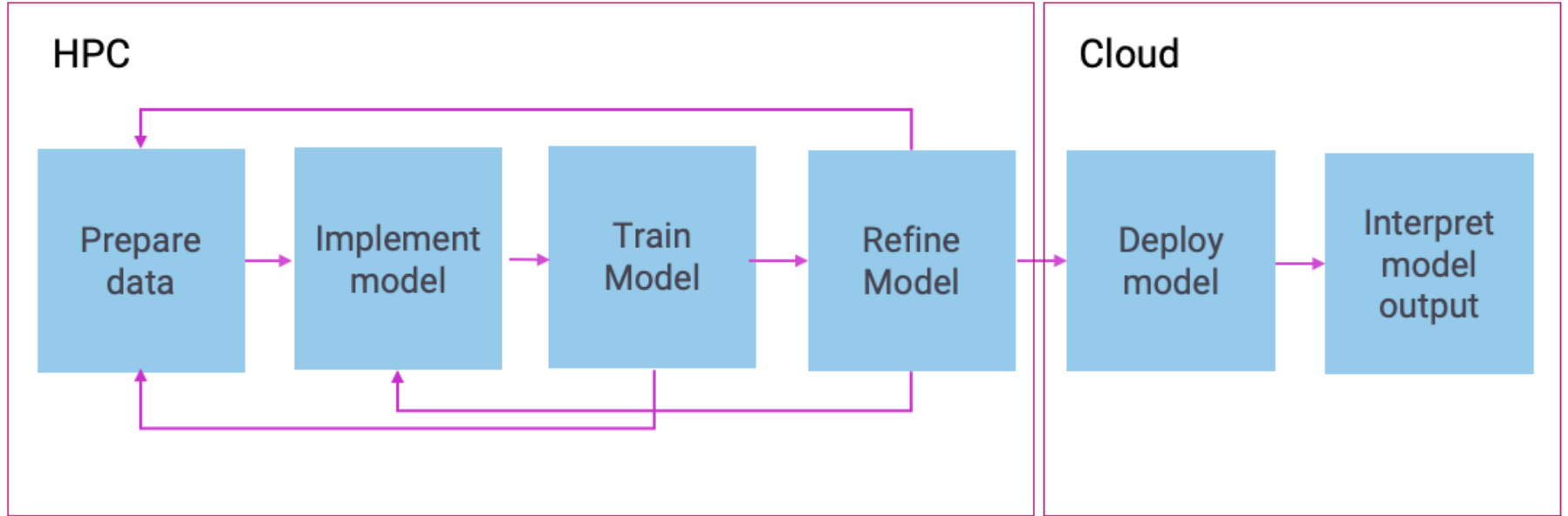
Model:

- U-Net model for semantic segmentation

Data:

- GOES
- Geostationary
- 15 min updates
- High-quality labeled

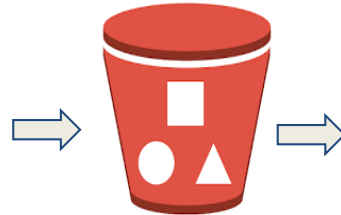
Machine Learning Lifecycle



Tutorial: Machine Learning Workflow



Train using Nodes in Jülich Supercomputing Centre

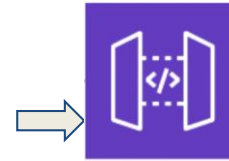


Upload model to AWS S3



Amazon SageMaker Endpoint

Load and host model in AWS SageMaker



Amazon API Gateway

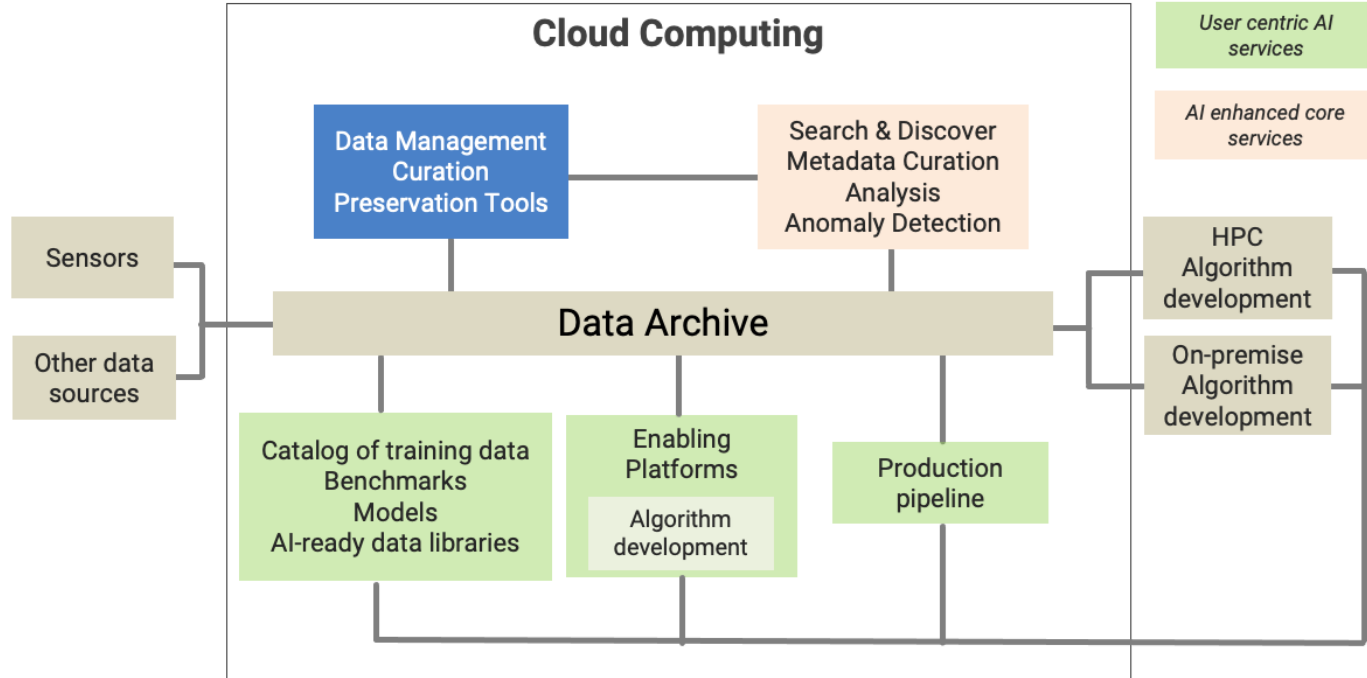
Serve Inference via API

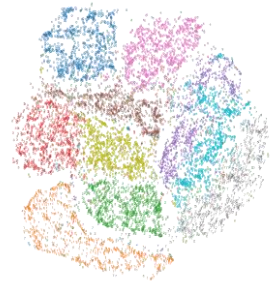


Lambda function

<https://github.com/NASA-IMPACT/pixel-detector>

Machine Learning Architecture





Thank you.