GIS & Spatial Machine Learning: Transforming Our Planet's Pulse to Action

Orhun Aydin, PhD
Senior Researcher

oaydin@esri.com
@orhun_sow
Geographer’s Guide to Galaxy (and this Keynote)
The Way We Think & Create

Need for Interoperable GIS
Advanced Analysis, Scientific Methods, Tools, Algorithms

Description | Prediction | Diagnostics | Prescription
Think Spatially
For Spatial Problems

Find a Framework

Develop/Use Tools
Outline for Today

Data Science in GIS
Spatial Data Science
Machine Learning
Deep Learning
Artificial Intelligence

Our Planet’s Memory
Citizen Science
Living Atlas

Integration with Platforms
Integration for Machine Learning
Integration for Data Science
Spatial (and Temporal) Data Wrangling
Projections
Mapping
Collecting Data
Empirical Models
Niche Frameworks
Space-implicit models
Traditional Spatial Statistics
Geostatistics
GIS

Object Detection
Global Models
Non-Spatial Classification
Regression
Clustering

Space-time explicit ML/DL

Empirical Models
Niche Frameworks

Machine Learning

Spatial Statistics
Artificial Intelligence

Machine Learning

Deep Learning
Artificial Intelligence

Machine Learning

Deep Learning
Artificial Intelligence: Pre-Machine Learning

- Programs with common sense (McCarthy, 1960) → Set of **predefined** logical operators
  * is (GIS Day, Nov 13\(^{th}\)), in (Nov 13\(^{th}\), today) → *in* (GIS, today!)

- Complex and extensive representations of human knowledge - Knowledgebases
  * CYC\(^2\)
  * SenticNet (1, 2, 3, 4, 5)\(^3\)

Artificial Intelligence: Machine Learning

- Learn rules and patterns from data
- Data is represented explicitly, knowledge is NOT
  - Data-driven
- Explicit rules do NOT exist, instead inferred from data

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance … improves with experience E.” ¹

Machine Learning in ArcGIS
Data-driven algorithms and techniques that automate tasks such as **prediction**, **classification** and **clustering**

**Traditional Machine Learning**
- Useful to solve a wide range of spatial problems
- Geography often acts as the ‘key’ for disparate data

**Spatial Machine Learning**
- Incorporate geography in their computation
- Shape, density, contiguity, spatial distribution, or proximity
Machine Learning Tools in ArcGIS

Classification
- Maximum Likelihood Classification
- Random Trees
- Support Vector Machine

Clustering
- Spatially Constrained Multivariate Clustering
- Multivariate Clustering
- Density-based Clustering
- Image Segmentation
- Hot Spot Analysis
- Cluster and Outlier Analysis
- Space Time Pattern Mining
- Time Series Clustering

Prediction
- Empirical Bayesian Kriging
- Areal Interpolation
- EBK Regression Prediction
- Ordinary Least Squares Regression and Exploratory Regression
- Geographically Weighted Regression
- Forest Based Prediction
Artificial Intelligence

Machine Learning

Deep Learning

ArcGIS Integration

ArcGIS

Video game behavioral AI
Computer Vision
IBM Watson

Natural Language Processing
Robotics

Theano
CNTK
TensorFlow
Keras

scikit-learn
Solving Spatial Problems

ArcGIS the Scientific Workbench
**Prediction**

Exploring Relationships Between Variables via Regression

**Use Case:** Understanding house price drivers in King County, WA from a rich dataset on house condition.

**In ArcGIS:** Forest Based Classification and Regression, Geographically Weighted Regression, Ordinary Least Squares Regression
Use Case: Accurately predict impacts of climate change on local temperature using global climate model data

In ArcGIS: Empirical Bayesian Kriging, Areal Interpolation, EBK Regression Prediction, Ordinary Least Squares Regression and Exploratory Regression, Geographically Weighted Regression
Use Case: Given the nearly 50,000 reports of traffic between 5pm and 6pm in Los Angeles (from Traffic Alerts by Waze), where are traffic zones that can be used to elicit feedback from current drivers in the area?

In ArcGIS: Spatially Constrained Multivariate Clustering, Multivariate Clustering, Density-based Clustering, Image Segmentation, Hot Spot Analysis, Cluster and Outlier Analysis, Space Time Pattern Mining
**Classification**

The process of deciding to which category an object should be assigned based on a training dataset

**Use Case:** Classify impervious surfaces to help effectively prepare for storm and flood events based on the latest high-resolution imagery

**In ArcGIS:** Maximum Likelihood Classification, Random Trees, Support Vector Machine
Interoperability with External Frameworks
Working in ArcGIS and Further
Open and Interoperable Platform for Science

Open Standards and Formats

- XLSForm
- GML
- SQL
- SLD
- SOAP
- WMTS
- KML
- LAS
- INSPIRE
- Shapefiles
- IMDF
- WCS
- IFC
- Web Scene (I3S)
- LERC
- CSW
- WPS
- REST
- OneGeology
- WFS
- WMS
- OPeNDAP
- JSON
- WaterML
- netCDF
- GeoPackage
- CityGML

Direct Product Integration

- MS Office
- SQL Server
- SharePoint
- Azure
- Power BI
- Netezza
- SAP HANA
- Adobe Creative Cloud
- Jupyter Notebook
- Teradata
- R
- AWS
- Altibase
- Python
- Oracle
- AutoCAD
- Revit
- PostgreSQL
- Dameng
- SQLite
- Spark
- Hadoop

Open Software Architecture

- Open Data Access
- Open APIs & SDKs
- Open-Source Integration
- Open-Source Contributions
- Extensible Architecture
- Embeddable

... Successfully Integrated Into Thousands of Systems
Machine Learning Integration with External Frameworks

ArcGIS

Desktop

Apps

APIs

TensorFlow

aws

mxnet

IBM Watson

scikit learn
ArcGIS Notebooks

- Performing analysis
  - Spatial analysis
  - Spatial data science
    - Machine learning
    - Deep learning

- Automating analytic and administrative workflows related to GIS

- Operationalizing data science models
Deep Learning Integration
What is the #1 Challenge?
Getting Everyone to SEAMLESSLY work together

Analyst
Access Imagery, Fix Data, Prepare Training Data, Formulate Ask
Consume Models for Analysis

AI Expert
Build and Optimize Models
Request Data

End User
Information Products
Analysis & Decision Making
Imagery AI Capabilities in ArcGIS

**Prepare Training Data**
*Export Training Data for Deep Learning*

- Feature Layer
- Imagery
- Chips

**Consume DL Models**
*Detect Objects & Classify Pixels*

- DL Model
- Imagery
- Detections

**Train DL Models**
*Arcgis.Learn*

**Deploy Models on Portal**
*As dlpk items*
Examples for other Imagery AI Workflows

Object Detection, Instance Segmentation, Land Cover, Change Detection...

- Damaged Structures
- Building Footprints
- Land Cover
- Pipeline Encroachment
- Roads
- Oil Pads
- Road Cracks
- Cars
- Swimming Pools
- Palm Trees
End to End GeoAI Life Cycle
with Imagery
1. Imagery Access & Preparation
Imagery Processing & Building Image Mosaic
2. Prepare Training Data
This is what we need to train a Deep Learning Model: Image Chips (Training Data)

ArcGIS has tools that can help with generating Training Data
Training Sample Manager
For easily labelling objects of interest
Export Training Data for Deep Learning

GIS Feature Class + Imagery Data → Geoprocessing

Parameters:
- Input Raster
- Output Folder
- Input Feature Class Or Classified Raster

→ Bounding Boxes
3. Train Model
Model Training using ArcGIS.Learn

**Train SingleShotDetector Model**

```python
from arcgis.learn import SingleShotDetector

ssd = SingleShotDetector(data, grids=[9], zooms=[1.0], ratios=[[1.0]],...)
```

The `arcgis.learn` module in ArcGIS API for Python enable GIS analysts and data scientists to easily adopt and apply deep learning in their workflows. It enables training state-of-the-art deep learning models with a simple, intuitive API.
4. Add Model to Portal
Upload Model to Portal as DLPK item
So it’s accessible to different people in the organization
5. Consume Model
Detect Objects Using Deep Learning

Imagery

Model (emd file)

Detected Objects GP Tool

Geoprocessing

Parameters

* Input Raster
* Output Detected Objects
* Model Definition

Arguments

Name | Value
--- | ---

Detected Objects
Model Definition File

- .emd extension with json formatted content
- describes the deep learning model to ArcGIS
  - deep learning framework
  - model
  - model type
  - model configuration
  - model inputs
  - model outputs
  - classes
  - if not using out-of-the-box model
    - optional Python Raster Function path
    - optional model parameters
- Image Scientist must understand and edit
6. Analysis
Selecting only NEW Buildings within the Pipelines’ Buffer
7. Feedback Loop
Image Visit App could be used as a QC tool to check model’s output
8. Action: Information Products
Workforce App to assign Inspection Tasks
Field Apps for Inspection and Data Collection
Operations Dashboard to view Inspection Results in Real-Time
End to End GeoAI Life Cycle with Imagery

- Imagery Access
- Imagery Prep
- Data Labelling
- Deploy Models to Production
- Train & Consume Models
- Run Inference at SCALE
- Feedback Loop
- Take Action
Introducing the R-ArcGIS Bridge

The R-ArcGIS Bridge

- The R-ArcGIS bridge allows you to connect ArcGIS to R and enables the seamless transfer of data back and forth, along with the ability to integrate R and ArcGIS functionality.
Patterns of Working with R

Local Data

Remote Data

R
Remote Data with R-ArcGIS Bridge

- Server
- Online Content & Services

Portal

R-ArcGIS Bridge

Spatial R Data frames
R Raster Types
R Mosaics
...
Living Data – Living Atlas
Spatial Memory of Our Planet
Living Atlas of the World

Open Data, Maps, and Apps from 1000s of Contributors + Esri

livingatlas.arcgis.com

Millions of Maps and Layers
Shared by Users

The Foremost Collection of Global Geographic Information

. . .
. . . A Living Atlas for the Planet
Citizen Science
esriurl.com/citizens

ArcGIS & Citizen Science

Citizen Science Introduction

Overview

The ArcGIS platform supports citizen science and crowdsourcing initiatives from data collection, analysis, monitoring, visualization and communication.

*Photo: Citizen scientist sampling in Rocky Mountain National Park. NPS Photo*
Spatial Storytelling
The Power of Storytelling

Science Communication is a growing movement among researchers, educators.
Get Involved
GeoAI Virtual Machine

Learn more at esriurl.com/geoai2018
Esri Applied Data Science Initiative

- Oregon State University
- EPA AirNow
- Conservation Intl
- USDOT Natl Highway Traffic Safety Admin
- Virginia Commonwealth U
- UC-Riverside
- DOE Natl Energy Technology Lab