

GIS Day, 2018, University of Buffalo

Machine Learning in ArcGIS

Mark Scott, Esri Local Government Team
NOT a GeoAI/ML Subject Matter Expert

Many Thanks to Lauren Bennett, Max Payson, and the
Spatial Statistics Team at Esri

mscott@esri.com



A word cloud featuring various terms in the field of Artificial Intelligence and Machine Learning. The words are arranged in a circular pattern around a central point, with their size and color varying. The largest word is 'Artificial Intelligence' in blue. Other prominent words include 'Machine Learning' in purple, 'TensorFlow' in green, and 'Neural Networks' in green. Smaller words include 'Caffe', 'Theano', 'CNTK', 'Object Tracking', 'Object Detection', 'scikit-learn', 'Computer Vision', 'Random Forest', 'Cognitive Computing', 'Natural Language Processing', 'T-SNE', 'GeoAI', 'Deep Learning', 'Keras', 'Support Vector Machines', and 'Dimensionality Reduction'.

Caffe

Object Tracking

CNTK

Object Detection

Artificial Intelligence

Theano

Computer Vision

scikit-learn

Random Forest

Machine Learning

Neural Networks

Cognitive Computing

Natural Language Processing

TensorFlow

T-SNE

GeoAI

Deep Learning

Keras

Support Vector Machines

Dimensionality Reduction

Artificial Intelligence

Machine Learning

Deep Learning

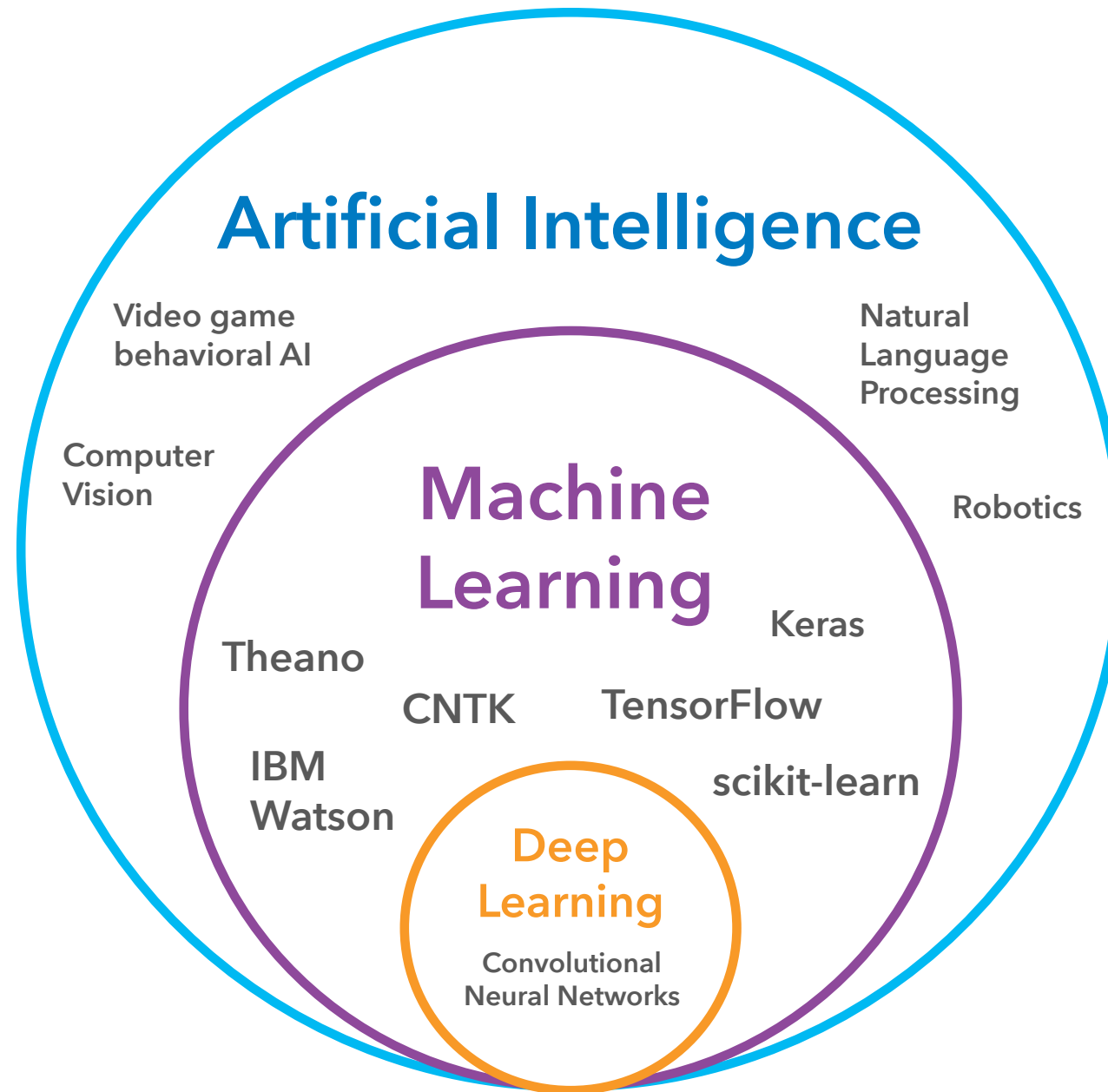


A Venn diagram consisting of three concentric circles. The outermost circle is light blue and contains the text 'Artificial Intelligence'. Inside it is a purple circle containing the text 'Machine Learning'. Inside the purple circle is a smaller orange circle containing the text 'Deep Learning'. This illustrates that Deep Learning is a subset of Machine Learning, which is a subset of Artificial Intelligence.

Artificial Intelligence

**Machine
Learning**

**Deep
Learning**



Artificial Intelligence

Video game
behavioral AI

Natural
Language
Processing

Computer
Vision

Robotics

ArcGIS

Machine Learning

Keras

Theano

CNTK

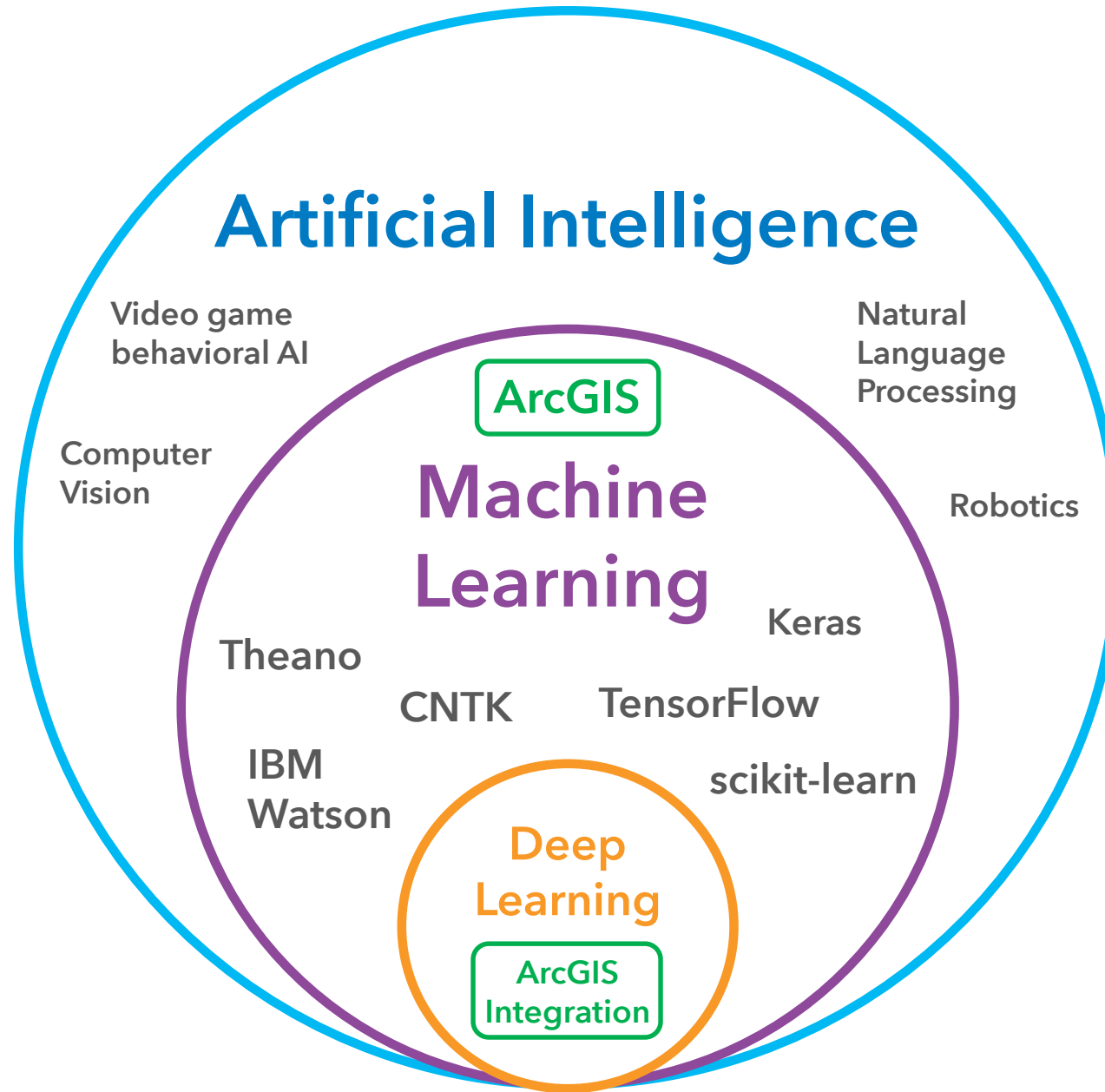
TensorFlow

IBM
Watson

scikit-learn

Deep Learning

ArcGIS
Integration



What is Machine Learning?

Data-driven algorithms and techniques that automate **prediction**, **classification** and **clustering** of data

Traditional Machine Learning

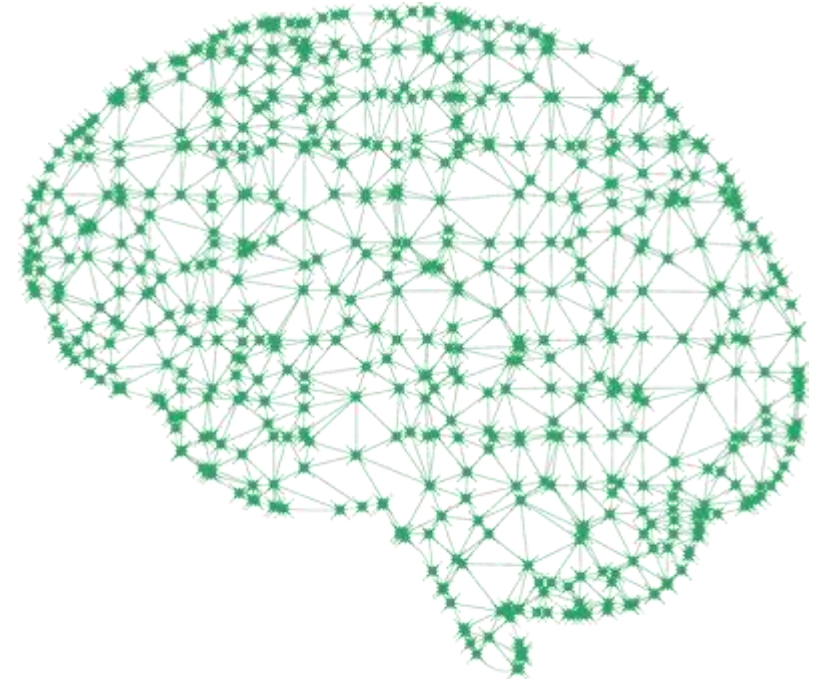
- Useful to solve a wide range of spatial problems
- Geography often acts as the 'key' for disparate data

Spatial Machine Learning

- Incorporate geography in their computation
- Shape, density, contiguity, spatial distribution, or proximity

Computationally Intensive

- Esri's continued advancements in storage and both parallel and distributed computing make solving problems at the intersection of ML and GIS increasingly possible



ArcGIS has Machine Learning Tools



GeoAI Project Lifecycle

Spatial Data Exploration



How is the Data distributed Spatially?
Any Spatial Patterns of interest?

Spatial Data Preparation



Example:
Snapping Car Crashes to Road Links, Geo-enrichment

Spatial Feature Extraction



Example:
Road Curvature, Number of Lanes, Proximity to Crossroads

Feature Engineering



Exploring Input Feature Correlation with the Output Feature.
Feature Selection Techniques

Model Development



Iterating with different Models. Choosing Models per Matrix of choice (e.g. RMSE)

Spatial Action Facilitation

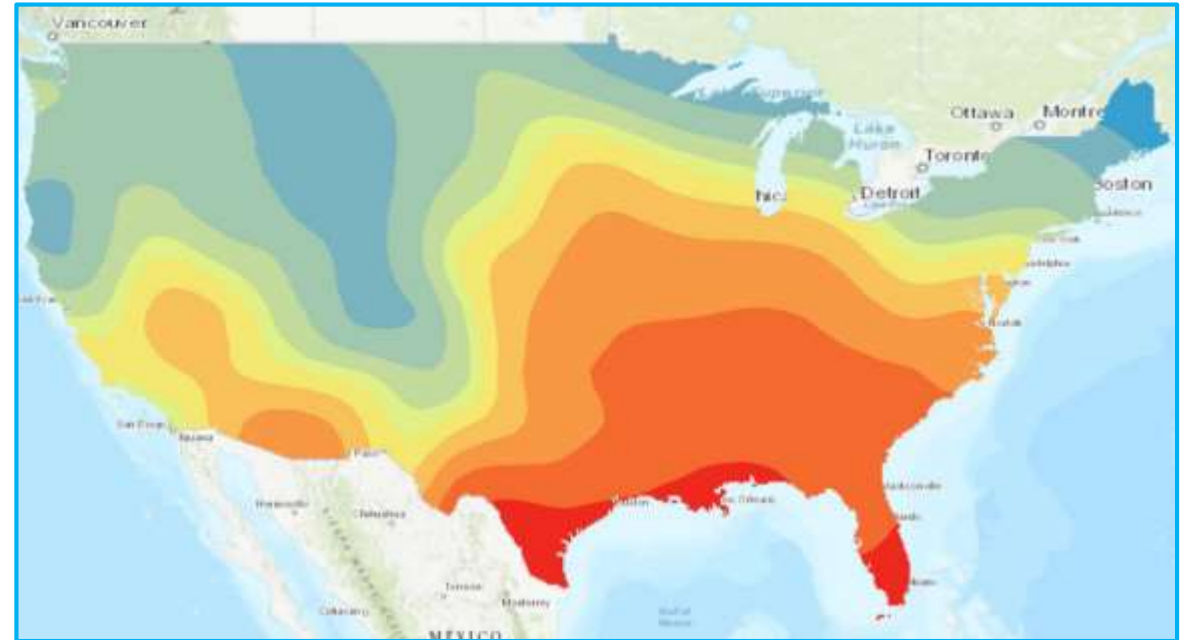
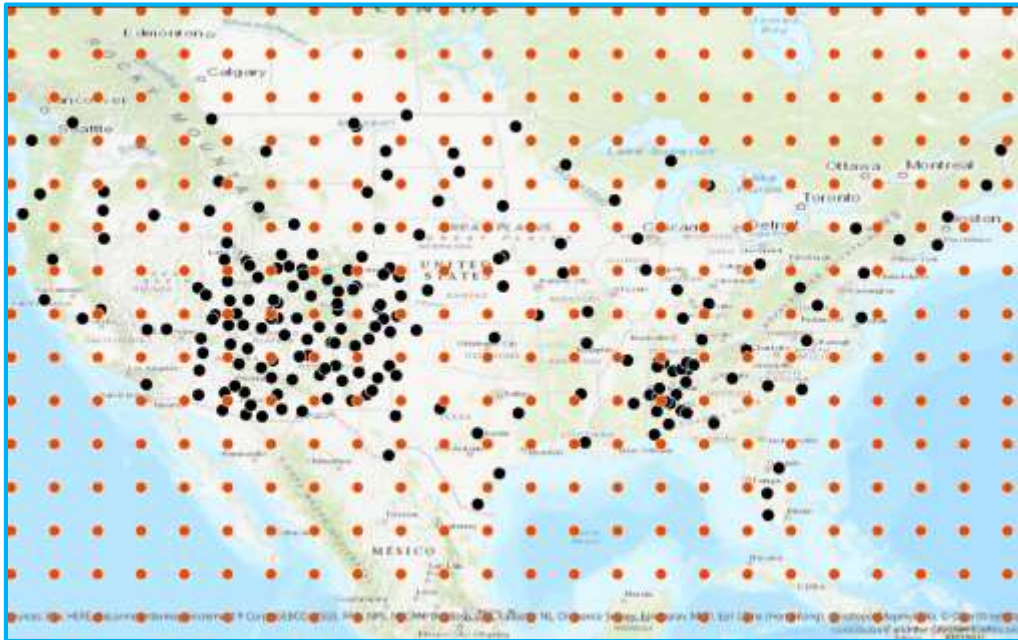


Facilitating Post-Prediction Actions
E.g.: Optimizing Ambulance Allocation based on Crash Prediction

Prediction

Using the known to estimate the unknown

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

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

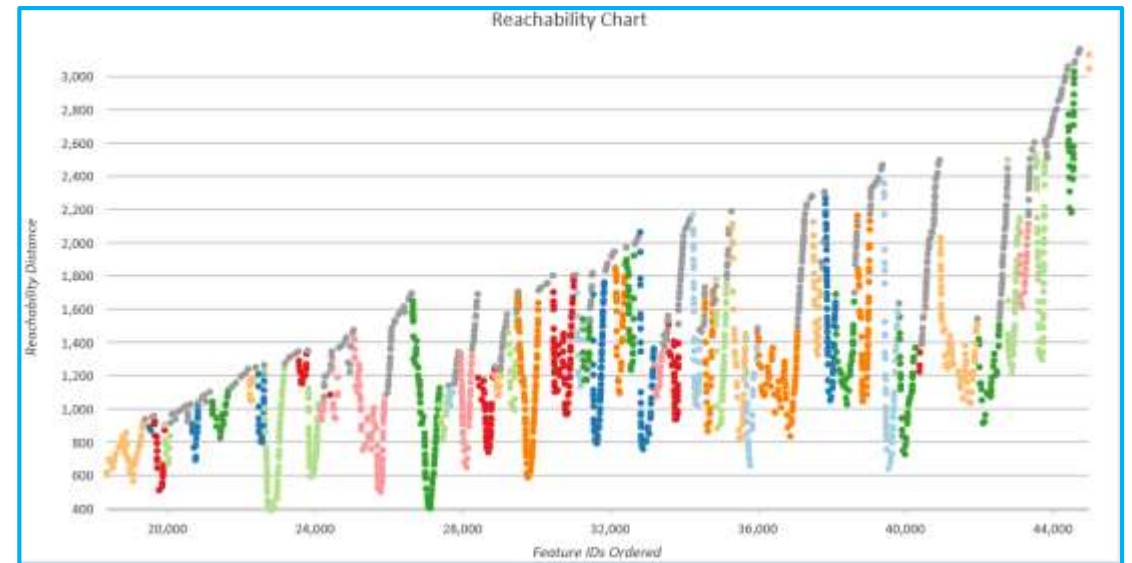


In ArcGIS: Maximum Likelihood Classification, Random Trees, Support Vector Machine

Clustering

The grouping of observations based on similarities of values or locations

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?



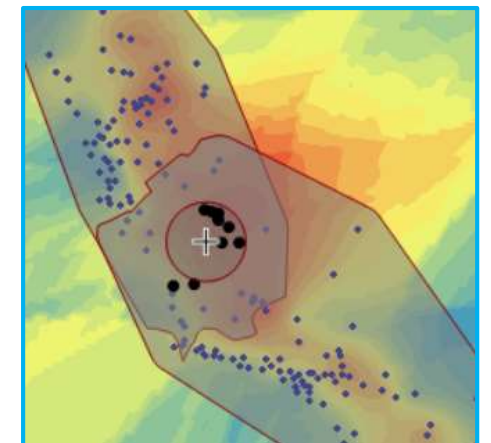
In ArcGIS: Spatially Constrained Multivariate Clustering, Multivariate Clustering, Density-based Clustering, Image Segmentation, Hot Spot Analysis, Cluster and Outlier Analysis, Space Time Pattern Mining

Behind the scenes...

Machine learning is also used throughout the platform as a means of choosing smart, data-driven defaults, automating workflows, and optimizing results

Examples:

- EBK Regression Prediction uses principal component analysis (PCA) as a means of dimension reduction to improve predictions
- The OPTICS method within Density-based clustering uses ML techniques to choose a cluster tolerance based on a given reachability plot
- The Spatially Constrained Multivariate Clustering tool uses an approach called evidence accumulation to provide the user with probabilities related to clustering results



Integration with External Frameworks



Integration

ArcGIS is an open, interoperable platform that allows for the integration of complementary methods and techniques

- ArcGIS API for Python
- ArcPy
- R-ArcGIS Bridge

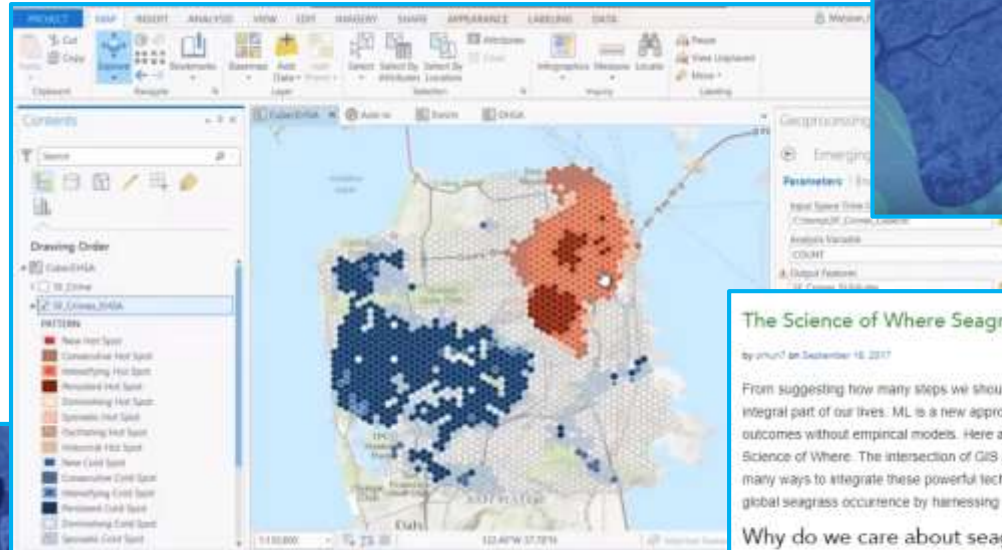
This integration empowers ArcGIS users to solve complex problems by combining powerful built-in tools with any machine learning package they need, *from scikit-learn and TensorFlow in Python to caret in R to IBM Watson and Microsoft AI* – all while benefiting from the spatial validation, geoenrichment, and visualization of results in ArcGIS.



Integration

Additional Resources

R-ArcGIS Bridge



Integration with Microsoft AI



Integration with sci-kit learn

What's Next

There are many key initiatives within Esri to advance both machine learning and deep learning methods, as well as integration approaches, across the platform

Methods

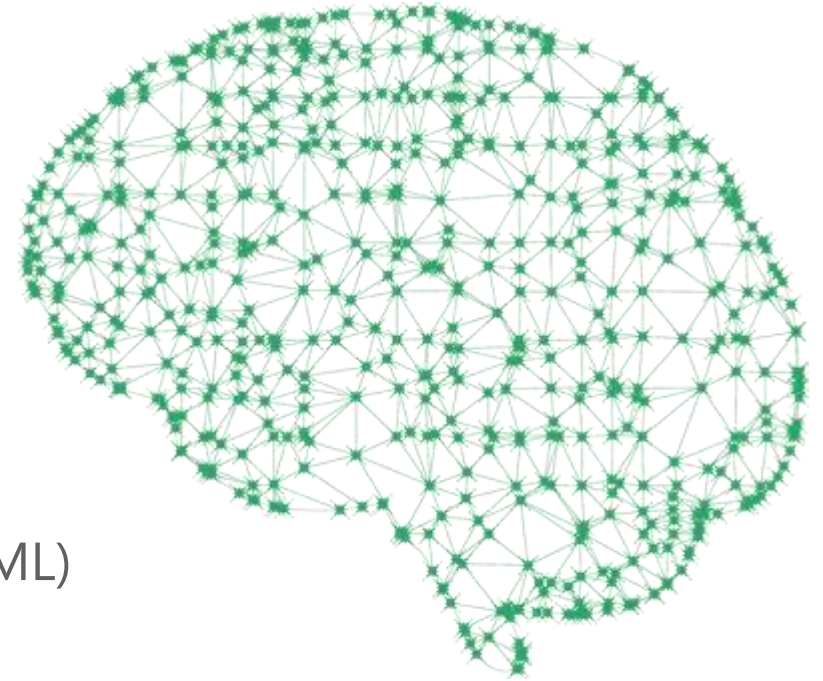
- Random forests
- Neural Networks
- Logistic Regression
- Time-series forecasting
- Space-time interpolation

Platform

- Major focus on big data and distributed computing (Spark ML)
- Simplified integration with popular packages

Research

- Continued innovation in spatial ML



Resources to learn more

Documentation

- [Geostatistical Analyst](#) (prediction)
- [Image Classification](#) (classification)
- [Spatial Statistics](#) (clustering, prediction)
- [Empirical Bayesian Kriging](#)
- [EBK Regression](#)

Other useful links

- <http://esriurl.com/analysis> (Case Studies)
- <http://esriurl.com/spatialstats> (Spatial Statistics resources page)

Resources to learn more

- Blog posts
 - <https://medium.com/geoai>
 - <https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/machine-learning-in-arcgis/>
 - <https://www.esri.com/arcgis-blog/products/product/3d-gis/restoring-3d-buildings-from-aerial-lidar-with-help-of-ai/>
 - <https://community.esri.com/community/gis/applications/arcgis-pro/blog/2018/03/27/predicting-travel-times-with-artificial-neural-network-and-historical-routes>
- **Google “ArcGIS Machine Learning” for a wide variety of videos from recent events**
 - https://www.youtube.com/watch?v=duZ7jor_YrU
 - <https://www.youtube.com/watch?v=aKq50YM8a8w>
 - <https://www.youtube.com/watch?v=KCkGif6wSMo>

Summary

- ArcGIS has AI/ML!
- You can use it to Predict, Cluster, Classify and more (e.g. Ecological NFA)
- You can interface with ML/L Frameworks
- Actions
 - Connect with a REAL ML/AI Subject Matter Expert

Questions?

Slide deck will be made available

mscott@esri.com