GeoAI Use-cases

Machine Learning meets ArcGIS

Omar Maher

Practice Lead, Advanced Analytics

1. AI > ML > DL

Artificial Intelligence

Reasoning

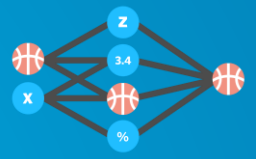
Knowledge Representation

Perception

NLP

Robotics

Machine Learning



Machine Learning

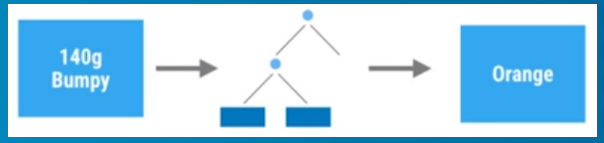
Supervised Learning

1. Training *features* *Labels*

Examples

Weight	Texture	Label
150g	Bumpy	Orange
170g	Bumpy	Orange
140g	Smooth	Apple
130g	Smooth	Apple
…	…	…

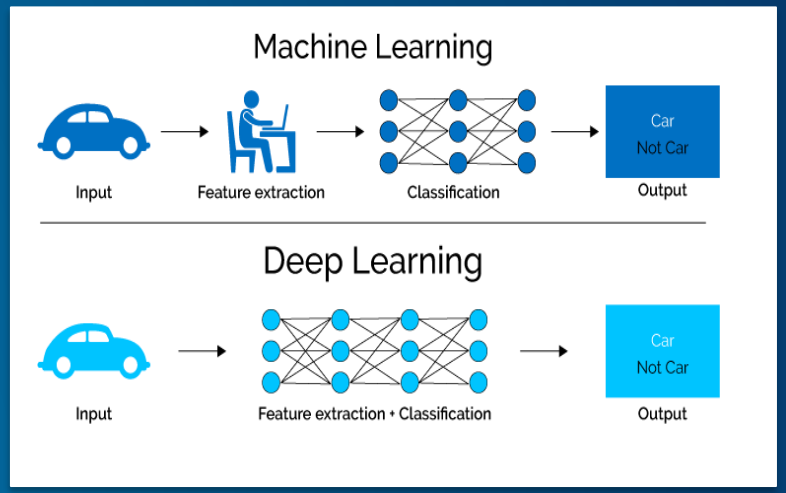
2. Predicting



Unsupervised Learning
Reinforcement Learning

Deep Learning

Deep Supervised Learning



Dog

2. ML Value

Prediction

Automation

Anomaly Detection

Object/Feature Detection from Imagery

Labelled Data Sets: Features + Output

Problem: Predicting Churn

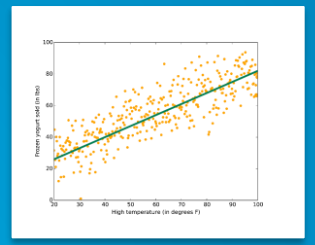
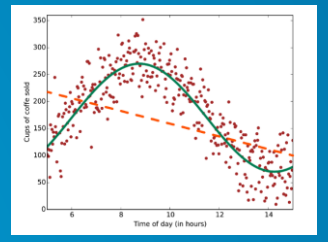
Features | **Output (label)**

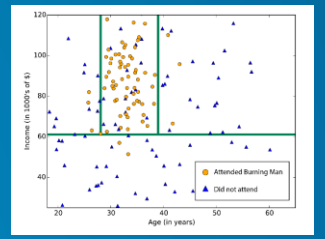
Training Data (Historical)

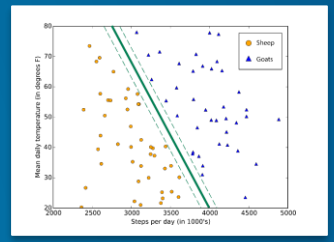
Call Drops	# Complains	Subscribed Package	Call Rate Decline	Churned?
4	5	ABC	20%	Yes
6	2	ABC	5%	No
9	4	XYZ	12%	Yes

For **Learning..**

Learning Models





Trained Model

New Data

7	5	KLM	8%

Yes (75%)

Prediction

Labelled Data Sets: Features + Output

Road Accidents Prediction

Segment Type	Proximity to Intersection	Time of Day	Weather	Accident
Highway	0.1 M	Morning	Raining	Injury
Tunnel	0.3 M	Evening	Sunny	Property
Inner	0.2 M	Noon	Foggy	Injury

Water Leakages Prediction

Pipe Age	Depth	Temperature	Pressure	Break
20	3 m	95 F	20 P	Yes
15	4 m	5 F	35 P	Yes
6	3.5 m	2 F	17 P	Yes

ISIS Location Prediction

IS Indirect Fire	IS Explosive Attack	IS Vehicle Attack	Coalition Air Ops	ISIS Camp
0.2 M	2 M	3 M	0.7 M	Yes
4 M	9 M	6 M	1 M	No
1.2 M	5 M	5 M	0.5 M	Yes

Retail Sales Prediction

# Females 35 - 50	F&B Sales	% Urban Chic	# Competitors	Sales
35,000	$400M	55%	15	$20M
14,000	$150M	15%	27	$5M
27,000	$210M	26%	9	$12M

Vector use-cases

Points, Lines, Polygons..

Retail, Real-Estate & Banking

Boost Annual Revenues by 5-10% through Predictive Geo-Targeted Marketing

Demographics

More than 2,000 variables on accurate current-year estimates and five-year projection of US demographics including population, households, income, age, and ethnicity. Data on education, labor force, journey to work, marital status, languages spoken, home value, and more.

Data Source: Esri, US Census Bureau, and ACS

Sample Report: Demographics and Income Profile → [PDF]

Consumer Spending

What products and services US consumers are buying where? Includes expenditures by households for 20+ categories including apparel, food and beverage, financial, and more.

Data Source: Esri, Consumer Expenditure Surveys, and Bureau of Labor Statistics

Sample Report: Retail Goods and Services Expenditure → [PDF]

Tapestry Segmentation

Detailed market segmentation of US residential neighborhoods divided into 67 segments based on socioeconomic and demographic characteristics. Provides insights into customer lifestyle and behaviors.

Data Source: Esri

Sample Report: Tapestry Segmentation Area Profile → [PDF]

Market Potential

Data on products and services consumers use and demand. Expected number of consumers and Market Potential Index data for goods, services, attitudes, and activities.

Data Source: Esri and GfK MRI

Sample Report: Retail Market Potential → [PDF]

Business Locations and Summary

Data on more than 13 million businesses including name and location, industrial classification code, number of employees, and sales.

Data Source: Esri and Infogroup

Sample Report: Business Locations → [PDF]

Major Shopping Centers

Statistics for nearly 7,000 major US shopping centers including name, total retail sales, gross leasable area, location, and more.

Data Source: Directory of Major Malls, Inc.

Sample Report: Major Shopping Center Locator → [PDF]

Traffic Counts

Extensive US traffic data for trade analysis, routing, and mathematical modeling. Data on the number of vehicles that cross a certain point or street location in the US.

Data Source: Market Planning Solutions Inc.

Sample Report: Traffic Count Map - Close Up → [PDF]

Retail MarketPlace

Compare retail sales and consumer spending to measure the gap between supply and demand. Learn where consumer spending might be leaking out of a market area and find opportunities to supply goods and services to new customers.

Data Source: Esri and Infogroup

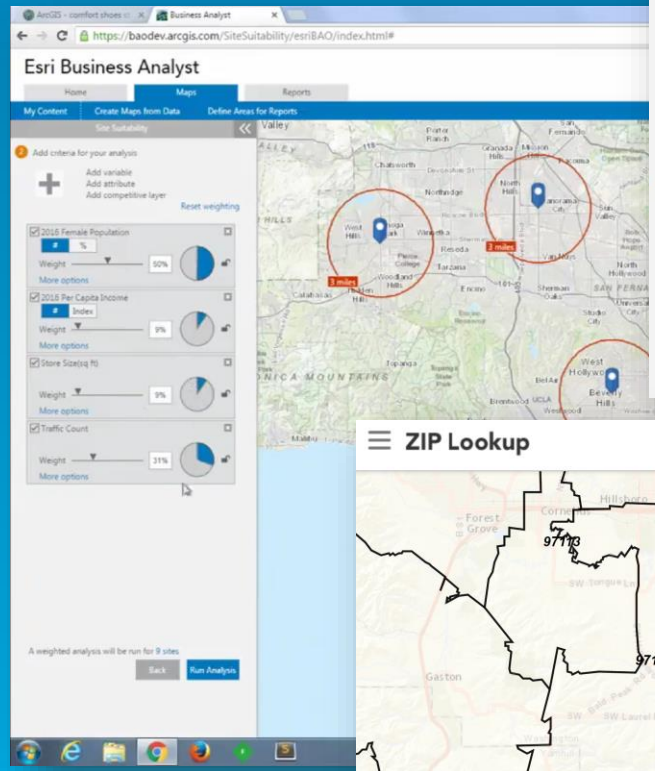
Sample Report: Retail MarketPlace Profile → [PDF]

ESRI Business Data

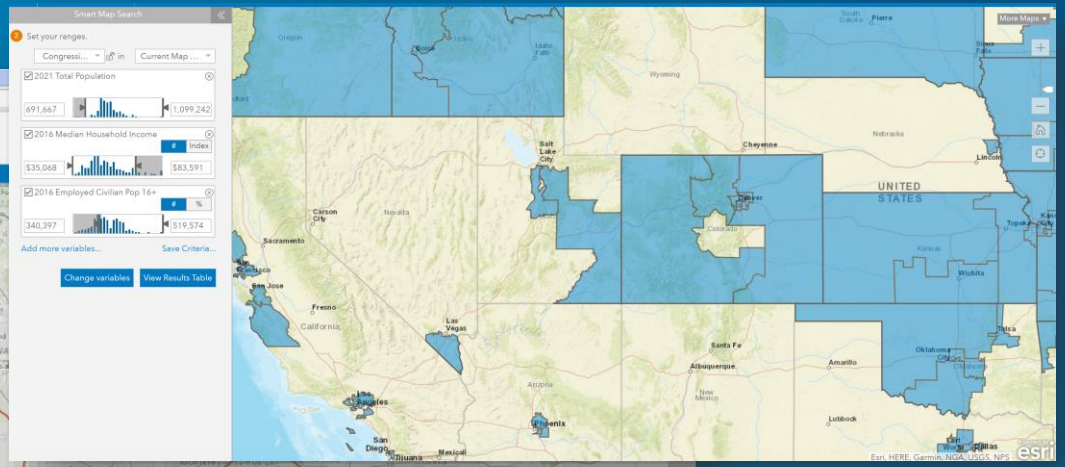
10,000 Features per Block Group (217K across US)

ESRI Business Data: Current Utilization

Manual Assignment of 10K+ Features w/ Weights *(Best Locations, Site Selection)*



Manual Filtering for 217K+ areas based Feature Values *(Targeted Promotion)*



Exploring Tapestry Segmentation per Area *(Market Expansion)*

Challenges

- Which Features are mostly affecting my Sales/Marketing? (10K+)
- What Weights should I assign per Feature?
- Do the Features differ per Area? To which extent?

Outcome

Significant Time spent on Manual Analysis with limited value realized

Solution

Machine Learning over Esri's data + Customer's Historical Sales data to **Predict Highest Sales Opportunities per 217K US Block Groups with Justification**

Esri's Data + Customer's Data = **Predictions for 5-10% Boost in Annual Revenues**

Esri's Data

- Demographics
- Consumer Spending
- Tapestry Segmentation
- Market Potential
- Business Locations
- Shopping Centers
- Traffic Counts
- Retail MarketPlace

Customer's Data

Mandatory

- Sales per Product, Shop , Time, & (Customer Segment)

Optional

- Marketing ROI (SMS, Email, Coupons..)
- Customers Location
- Competitors Location
- Customer Segments

Esri's AI Engine

Machine Learning Models over 15K Features per Block Group + Customer's Historical Data (1+ year) to Predict Highest Sales Opportunities

Output

Do 1, 2, & 3 to **increase Annual Revenues by 5-10%**

Geo-Targeted Marketing

Product per Shop per Segment per Channel

Geo-Cross & Upselling

Recommended Products Bundling per Shop

High Potential Areas

Top Areas for New Store, or Promotion per Product

Incident Management

Predict Incident Locations, Classify by Severity, Optimize Asset Allocation

Road Data

Road Alignment

Road Type

Road Width

Segment Type

Num of Lanes

Road Junction

Accidents' Data

Accident Type

Severity

Day of Week

Rush Hour

Time of Day

Weekend

Metadata (Optional)

Weather

Events

Road Works

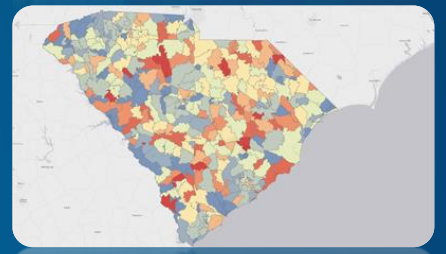
Road Block

Speed Limit

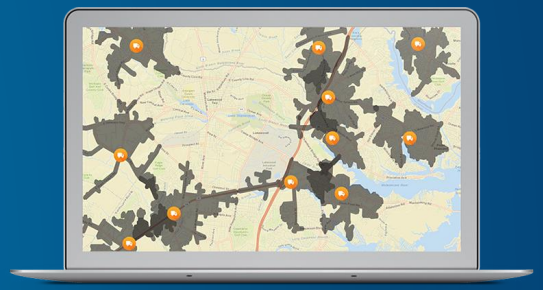
Output

Predicted Accident Probability

Categorize by Severity



Optimize Ambulance Placement



Predict the Location, Type and Time of Accidents before they occur

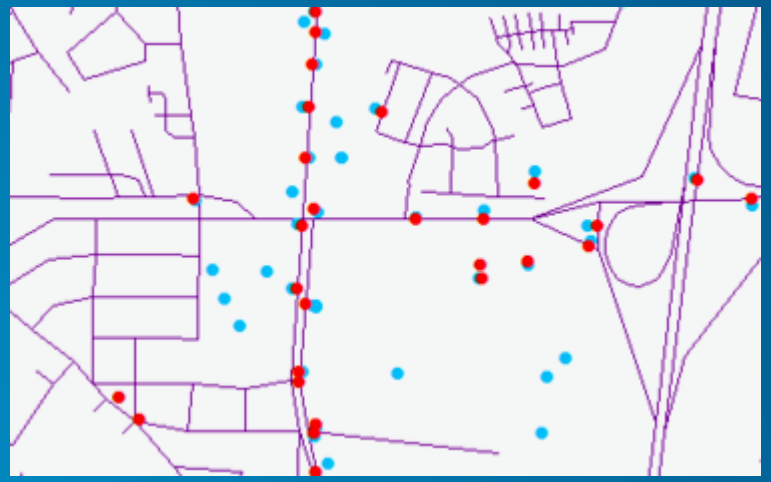
- Cut Incident Rates
- Optimize Patrols Allocation
- Explore Accidents Root-Causes
- Improve Response Time

Machine Learning Models over 1+ year of Accidents Data

1. Spatial Data Exploration

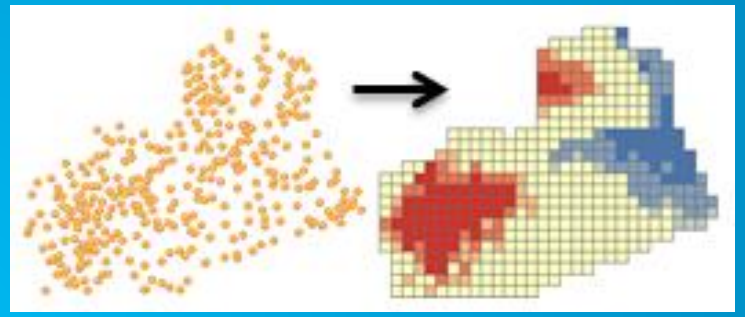
Add the Road Network Layer to derive the Road Geo-Features (curvature, speed..)

Data not Accurate: Crashes location slightly way from Road Links

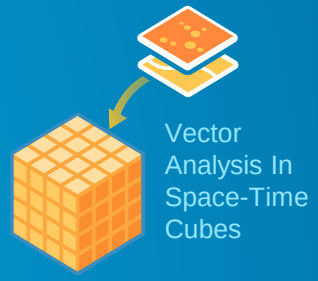


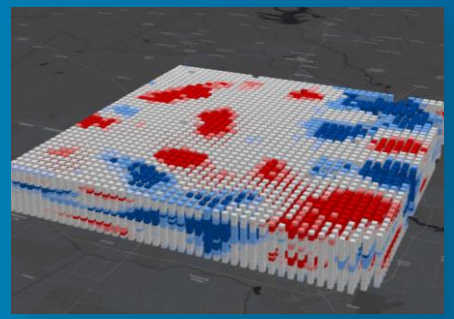
Patterns: Crashes mostly on Crossroads. Add Proximity to Crossroads as a Feature

Random VS Non-Random Events?
Run Hotspot Analysis

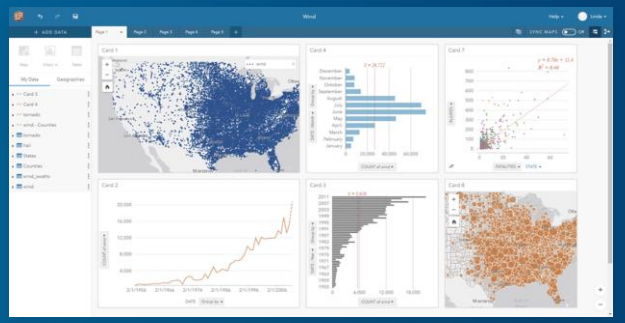


Emerging VS Fading Hotspots?
SpaceTime Cube





Explore Spatiotemporal Patterns..

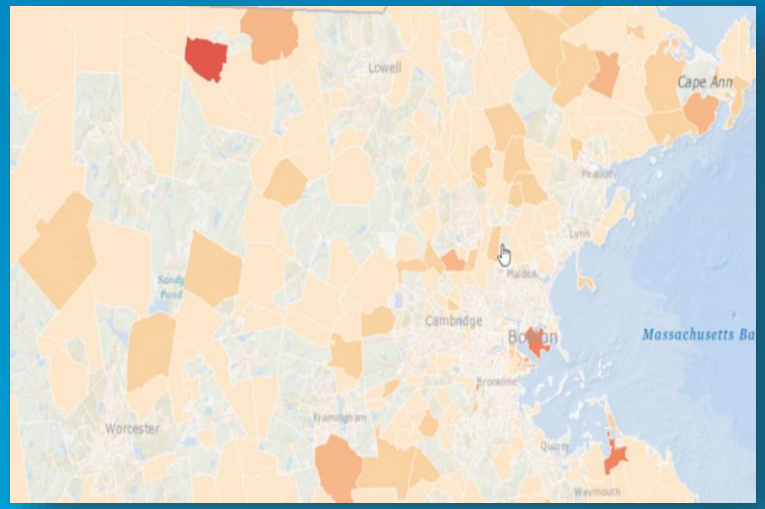


Predict Sales, Marketing ROI, Cross-Selling, and Potential per Location

With high Accuracy, Speed & Scale

Where should I Promote for Product XYZ?

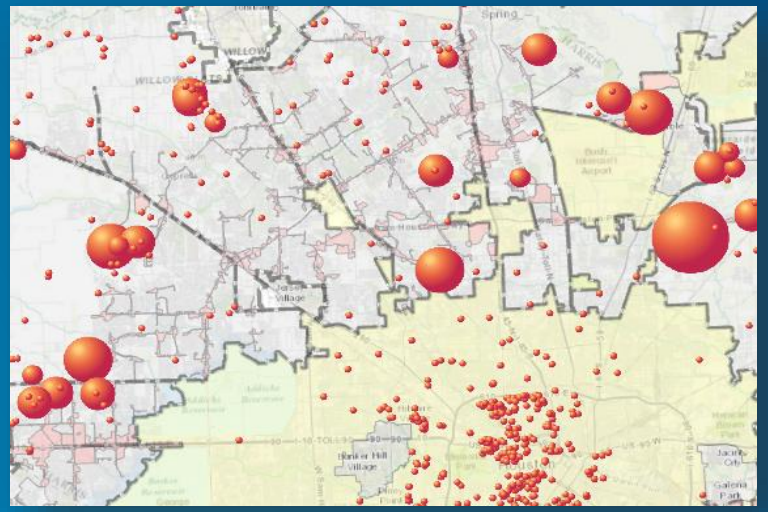
What's the expected Return? Why?



- ***Areas:*** *1,2 & 3*
- ***Predicted Sales:*** *$12.7M*
- ***Reasons****: 55K Hispanic Male Teenagers 12-17, 75% Comfortable Empty Nesters (Tap), 10+ mins Away from 70% of Competitors*

Where should I open a new store?

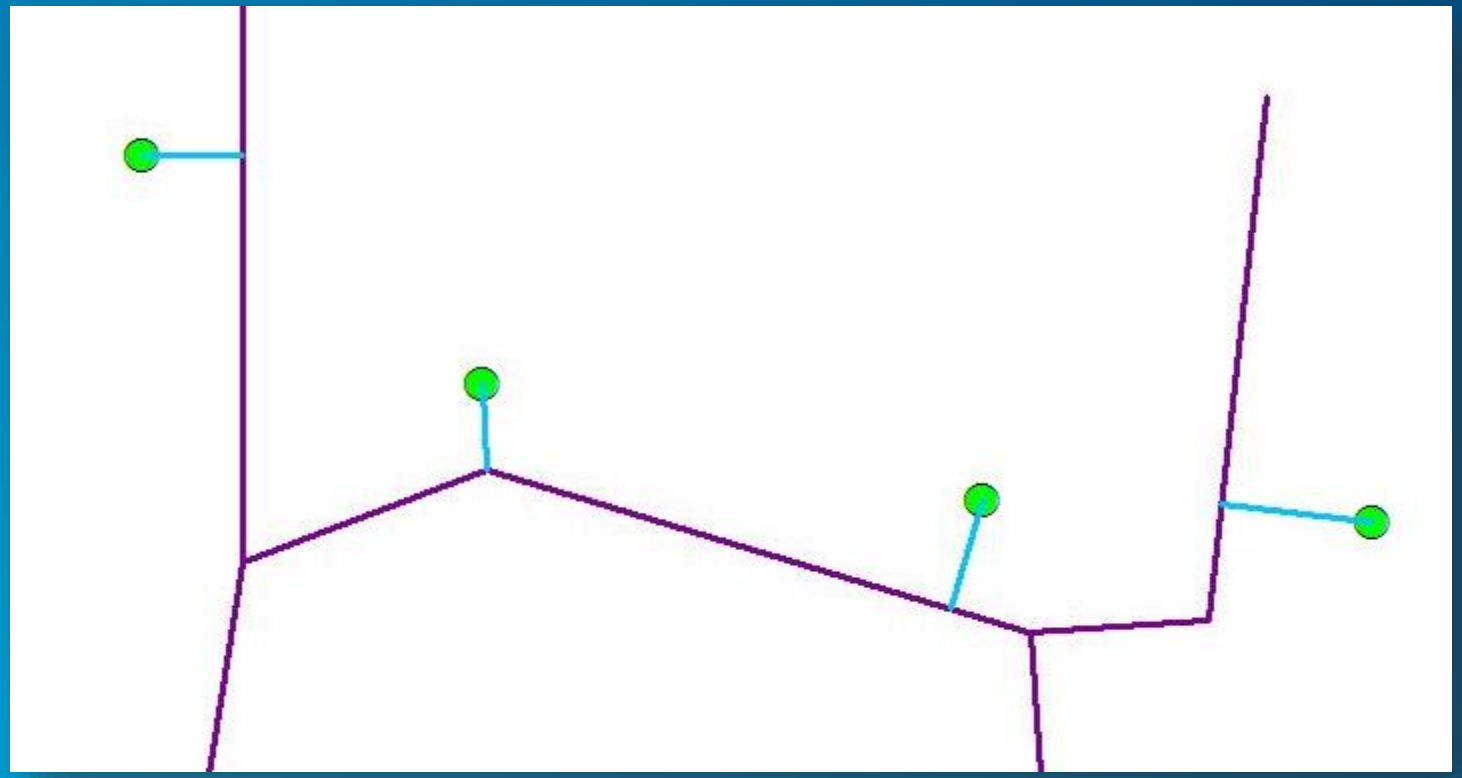
What's the expected Sales? Why?



- ***Block G:*** *31B*
- ***Predicted Sales:*** *$2.1M*
- ***Reasons****: 125K Employed Female 25-28s whose income > $80K, 2016 "Food Away from Home - Dinner at Full Service Restaurant" Index 125, 67% Diners & Miners*

2. Spatial Data Preparation

Example: Snapping Crash Points to the Nearest Line (Road Link), GeoEnrichment w/ 10K Features



3. Spatial Feature Extraction

Road Alignment
Straight / Curved

Road Type
Double / Single Carriageway

Num of Lanes
1,2,3..

Road Width
20-30 M

Segment Type
Highway, Bridge, Tunnel..

Road Junction
Crossroads, Roundabout..

Day of the Week
Sun, Mon, Fri..

Time of the Day
12:45, 23:00

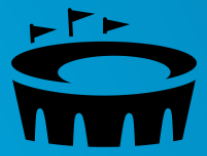
Rush Hour
Yes/No

Weekend
Yes/No

Raining
Yes/No

Fog
Yes/No

Temperature
Sun, Mon, Fri..



Nearby Event
Sports, Festivals..

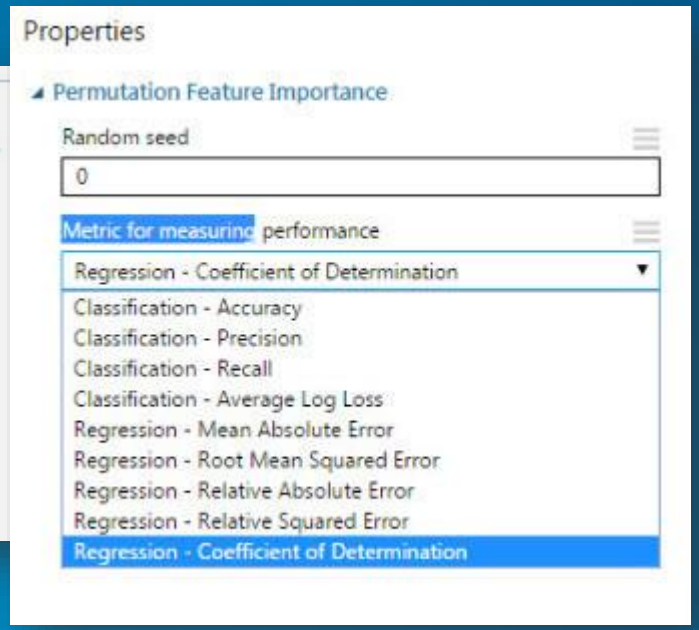
Nearby block
From related roads

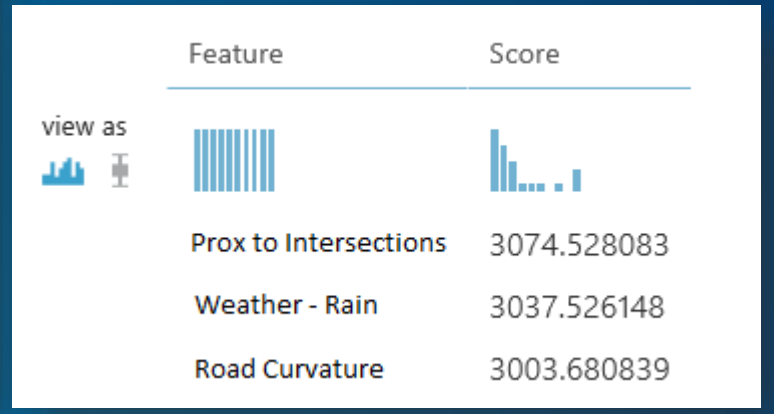
Speed Limit
120 km/h

Road Works
Yes/No

4. Feature Selection: listing the values of all independent features that were associated with historical accident incidents – and running Feature Selection techniques to select the most relevant features

1389135600000	07.01.2014.	10.31	9.83	1	92.51	4.185
1389222000000	08.01.2014.	10.31	9.83	1	92.26	4.02
1389308400000	09.01.2014.	10.31	9.83	1	92.85	4.076
1389567600000	12.01.2014.	10.31	9.83	1	91.52	4.3
1389654000000	13.01.2014.	10.31	9.83	1	92.62	4.342
1389740400000	14.01.2014.	10.31	9.83	1	94.29	4.328
1389826800000	15.01.2014.	10.31	9.83	1	94.13	4.367
1389913200000	16.01.2014.	10.31	9.83	1	94.26	4.295
1390172400000	19.01.2014.	10.31	9.83	1	93.84	4.262



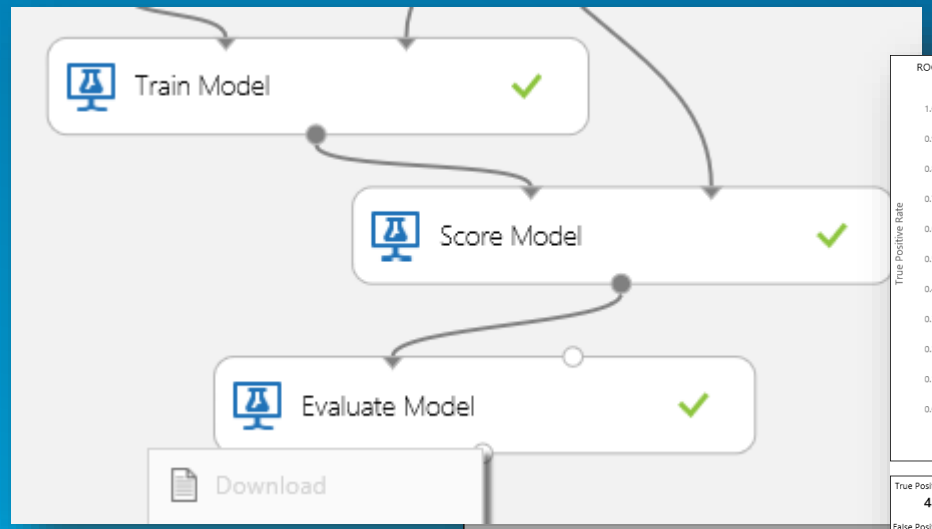


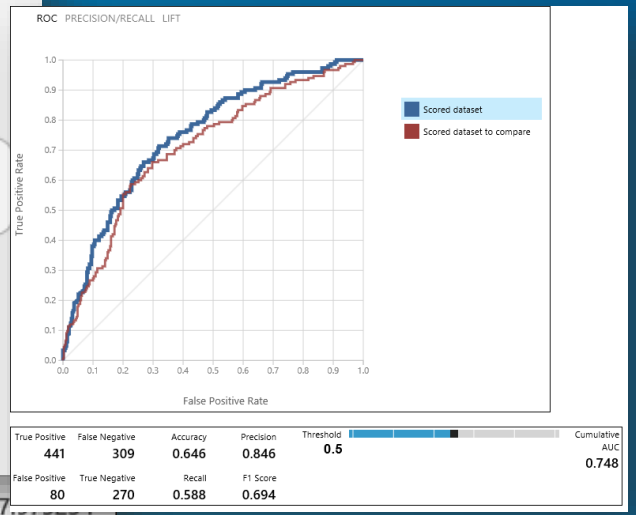
Independent Features

Dependent Feature

Time	Segment	Curved	Lanes	Type	Width	Rain	Wind	Temp	Prox Inters	Weekend	…	Accident
17:05	A123	Yes	2	Highway	30	Yes	65 km/h	28	0.3	No	…	Two Cars
23:50	B742	No	3	Tunnel	45	No	23 km/h	18	0.1	Yes	…	Road Deviation

5. Model Development: Experimenting with different models to see which yield highest Matric of interest (e.g. RMSE)





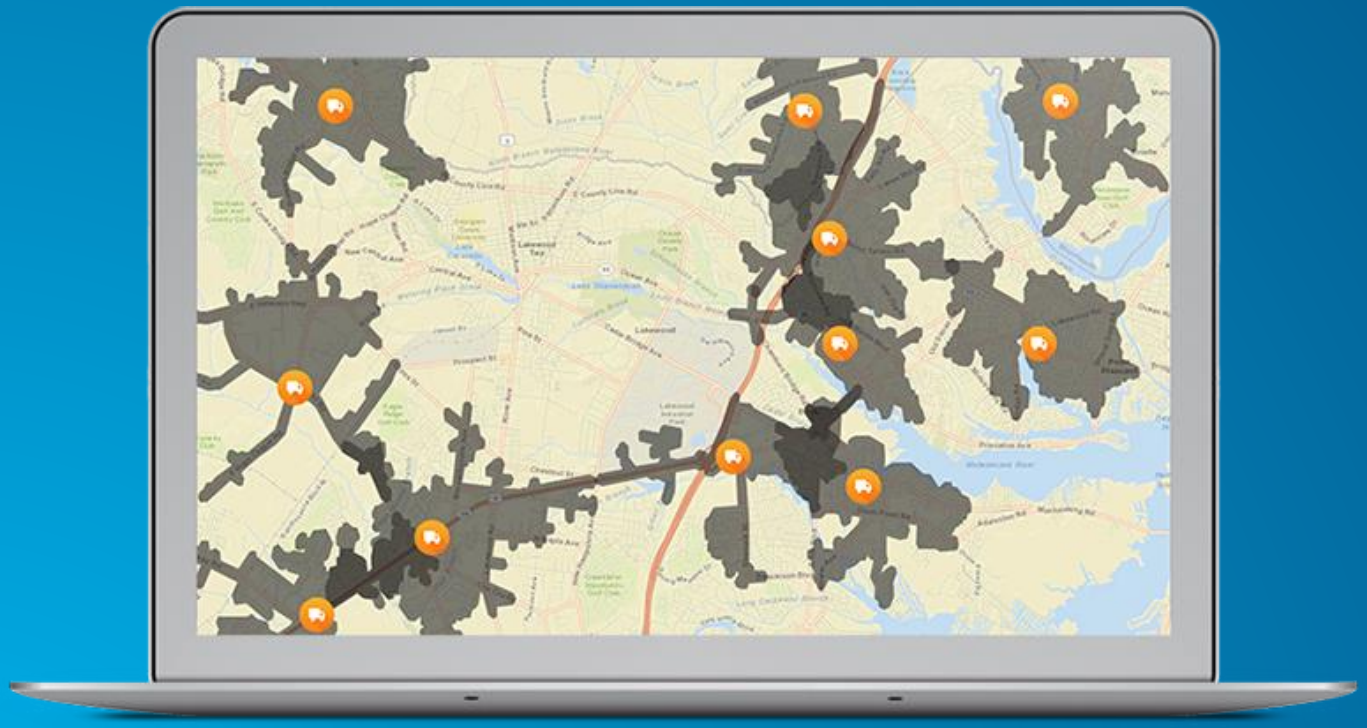
Mean Absolute Error	747.[illegible]
Root Mean Squared Error	955.587783
Relative Absolute Error	0.163528
Relative Squared Error	0.026598
Coefficient of Determination	0.973402

Fold Number	Number of examples in fold	Model	Accuracy	Precision	Recall	F-Score	AUC	Average Log Loss	Training Log Loss
0	3256	Logistic Regression	0.850399	0.737342	0.621333	0.674385	0.905809	0.325788	41.989626
1	3256	Logistic Regression	0.842245	0.707358	0.585062	0.640424	0.901871	0.32527	40.990733
2	3256	Logistic Regression	0.852583	0.767007	0.591864	0.668148	0.905203	0.32936	41.51439

Accidents Probability

Time	Segment	Curved	Lanes	Type	Width	Rain	Wind	Temp	Prox Inters	Weekend	Accident Probability
17:05	A123	Yes	2	Highway	30	Yes	65 km/h	28	No	No	**82%**
23:50	B742	No	3	Tunnel	45	No	23 km/h	18	No	Yes	**25%**

6. Post Prediction Actions: Optimizing Asset Allocation: based on Accident Predictions



- Optimize the allocation of Police Patrols & Ambulance Cars based on the Predicted Accident Locations
- Take Accident type into consideration (Critical, Medium..)
- Inputs: Predictions, Asset types, numbers, allocation constraints, shifts.. Etc
- Root-cause Analysis

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

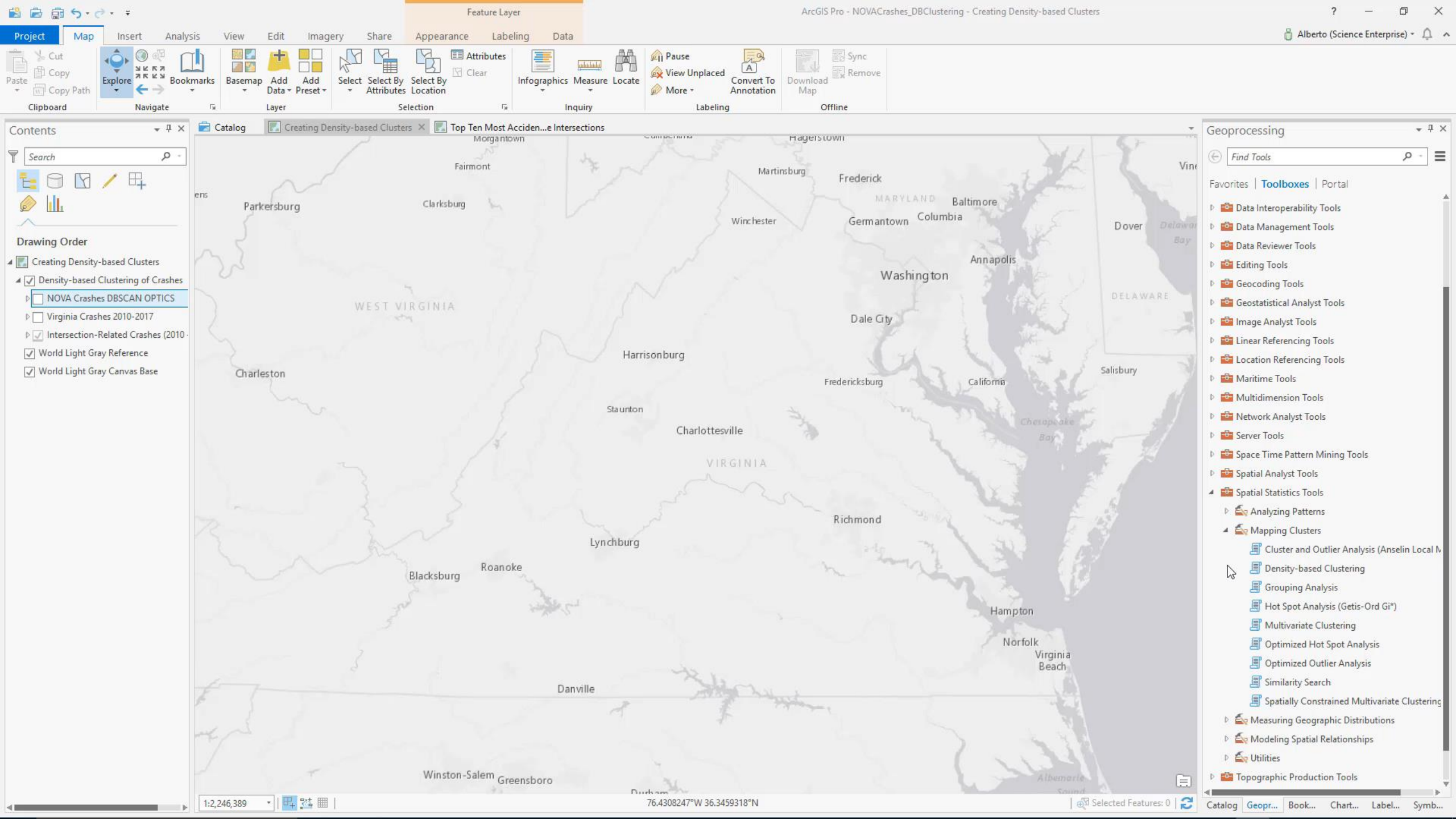
Exploring Input Feature Correlation with the Output Feature. Feature Selection Techniques

Model Development

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

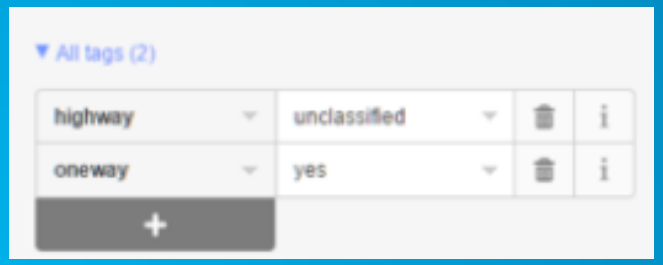
Spatial Action Facilitation

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

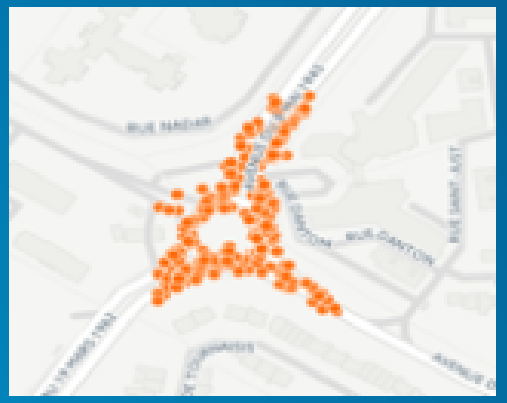

Feature Layer
ArcGIS Pro - NOVACrashes_DBClustering - Creating Density-based Clusters
Project
Map
Insert
Analysis
View
Edit
Imagery
Share
Appearance
Labeling
Data
Alberto (Science Enterprise)
Cut
Copy
Copy Path
Paste
Clipboard
Explore
Bookmarks
Navigate
Basemap
Add Data
Add Preset
Layer
Select
Select By Attributes
Select By Location
Attributes
Clear
Selection
Infographics
Measure
Locate
Inquiry
Pause
View Unplaced
More
Convert To Annotation
Labeling
Download Map
Sync
Remove
Offline
Contents
Search
Drawing Order
Creating Density-based Clusters
Density-based Clustering of Crashes
NOVA Crashes DBSCAN OPTICS
Virginia Crashes 2010-2017
Intersection-Related Crashes (2010
World Light Gray Reference
World Light Gray Canvas Base
Catalog
Creating Density-based Clusters
Top Ten Most Acciden...e Intersections
Morgantown
Fairmont
Parkersburg
Clarksburg
Martinsburg
Hagerstown
Frederick
Winchester
Germantown
Columbia
Baltimore
Annapolis
Washington
Dover
Dale City
Harrisonburg
Charleston
Fredericksburg
California
Salisbury
Staunton
Charlottesville
Richmond
Lynchburg
Roanoke
Blacksburg
Hampton
Norfolk
Virginia Beach
Danville
Winston-Salem
Greensboro
WEST VIRGINIA
VIRGINIA
MARYLAND
DELAWARE
Geoprocessing
Find Tools
Favorites
Toolboxes
Portal
Data Interoperability Tools
Data Management Tools
Data Reviewer Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Image Analyst Tools
Linear Referencing Tools
Location Referencing Tools
Maritime Tools
Multidimension Tools
Network Analyst Tools
Server Tools
Space Time Pattern Mining Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Cluster and Outlier Analysis (Anselin Local M
Density-based Clustering
Grouping Analysis
Hot Spot Analysis (Getis-Ord Gi*)
Multivariate Clustering
Optimized Hot Spot Analysis
Optimized Outlier Analysis
Similarity Search
Spatially Constrained Multivariate Clustering
Measuring Geographic Distributions
Modeling Spatial Relationships
Utilities
Topographic Production Tools
1:2,246,389
76.4308247°W 36.3459318°N
Selected Features: 0
Catalog
Geopr...
Book...
Chart...
Label...
Symb...

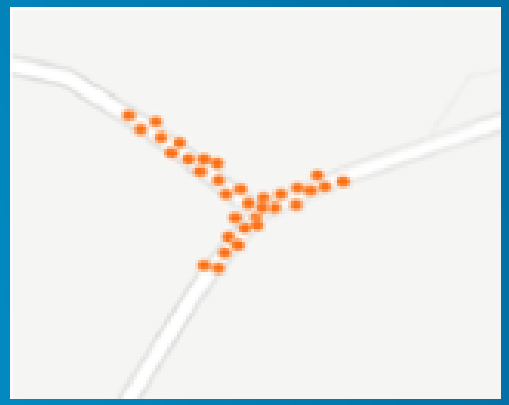
Predicting Roundabouts from GPS Probe Data

OSM Maps has many undetected roundabouts

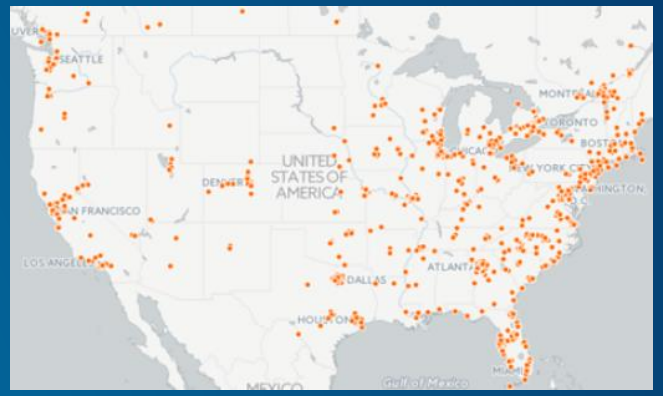


ML trained to detect circular vehicle movements with a whole





9,000 roundabouts were detected across US & Europe



Transportation

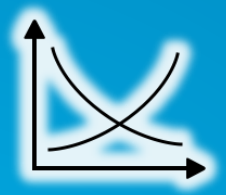
Forecast Demand, Reduce Congestion, Optimize Pricing, & Optimize Urban Planning

Data collected from different sources, processed, and integrated

Predictive Analytics Engine captures the data and runs it against Advanced Statistical Models

Insights are shown in real-time over ArcGIS Apps with rich Spatial Analysis tools for even deeper understanding

Demand Forecasting



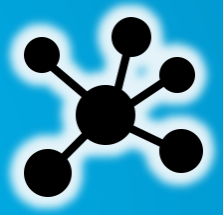
Traffic State Prediction

Trip Duration Prediction

Congestion Root-Cause Analysis

Price Optimization & Demand Control

Network Effect Prediction



Network Effect Prediction

Predictive Urban Planning

Predictive Parking

Predictive Policing

Predict Crime Possibility, Classify by Severity, Optimize Asset Allocation

Predictive Policing

Analyzing historical data from multiple sources to highlight areas of high Crime Likelihood, and the optimum placement of police assets accordingly

Historical Crime Incidents

Demographics per Area

Repeat Victimization

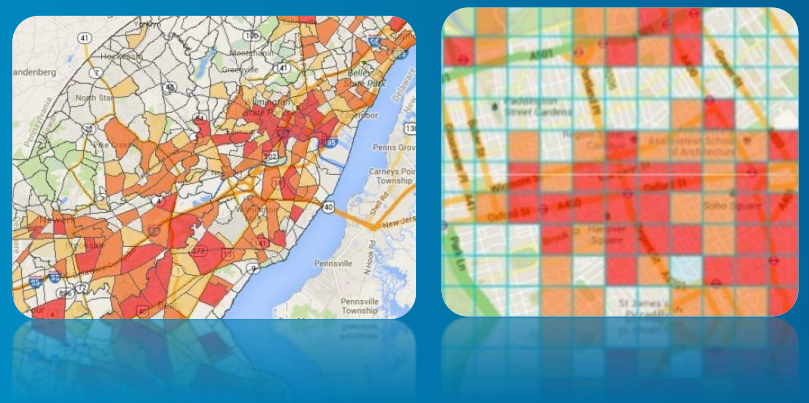
Risk Terrain Modelling

Taskforce Placement

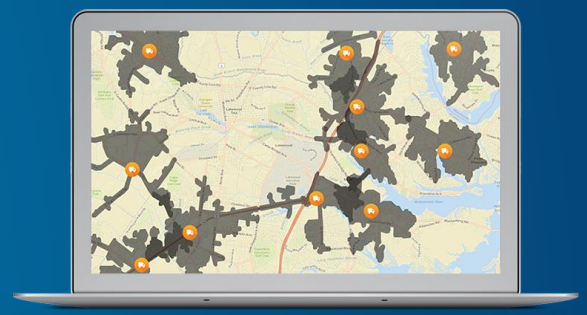
Events Location

Predictive Model

Crime Probability per Area



Recommended Asset Locations

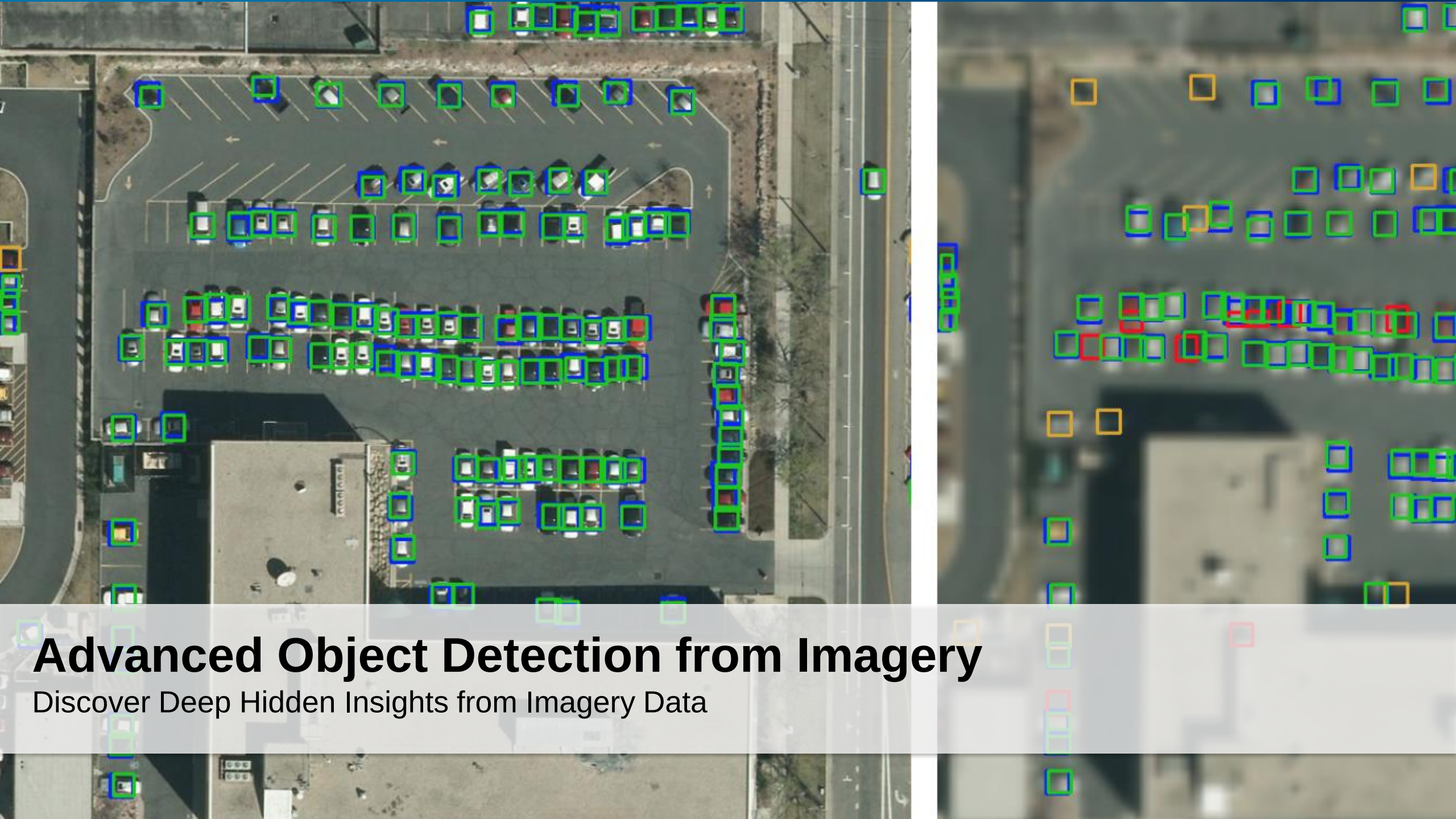


Machine learning for Smart Cities

- Accidents Prediction
- Crime Prediction
- Predicting Building most likely to catch fire
- Predict Restaurant Inspections Priority
- Predict Ridership Patterns (volume, public transportation utilization
- Correlate Land-use to Ridership/Congestion and Predict the effect
- Imagery
  - Automating Detection: of phenomena of interest: street cracks/packing violations/garbage/violence/traffic jams/crashes from satellite imagery, drone feeds, street cameras...etc.
  - Labelled Counting: trees and their condition, displaced garbage cans (instead of dispatching human workers to count them)
- Census: predicting the amounts of people with a specific phenomenon/trait (e.g. blind)
- Mapping the unmapped
- Predicting Parking slots utilization (and dynamically direct people to other nearby areas)
- Understanding people’s movement & transportation patterns (Telecom Data)

Imagery use-cases

Points, Lines, Polygons..

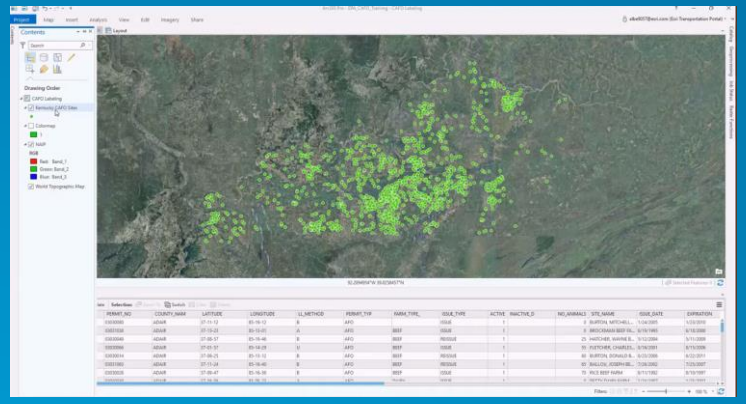


Advanced Object Detection from Imagery

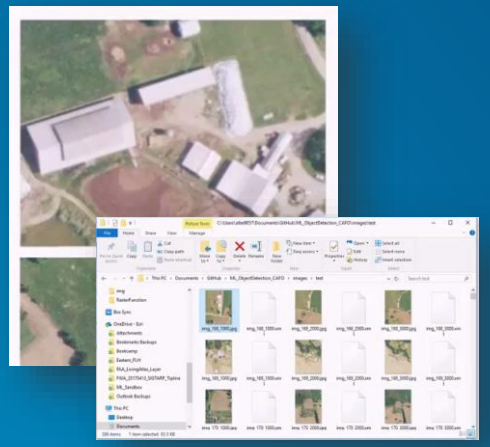
Discover Deep Hidden Insights from Imagery Data

Object Detection using Deep Learning + ArcGIS Pro

1. Sample Training Data



2. Extracting Images

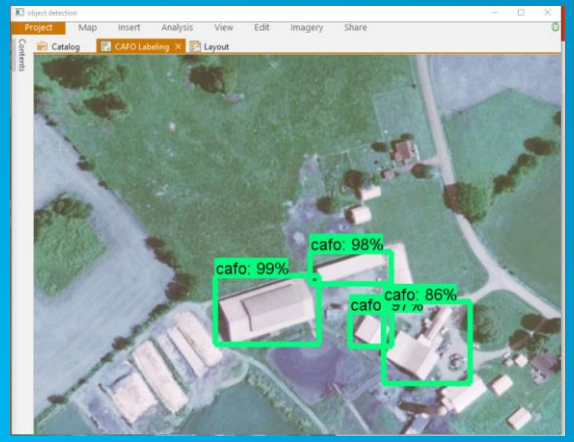


3. Bounding Boxes

4. Train TF CNN

TensorFlow

5. Detect Objects



6. Call the model directly from Pro..

..Via Python Raster Function

```
134 lines (112 sloc)  5.15 KB
import numpy as np
import tensorflow as tf

class Detect():

    def __init__(self):
        self.name = "Detect Object Function"
        self.description = ""

        self.modelPath = None
        self.clsLblPath = None
        self.threshold = 0.8
        # self.num_classes = 1
        #self.batch = 128

    def getParameterInfo(self):
        return [
```

Access all GeoAI Demos from: https://github.com/ArcGIS/geo-ai

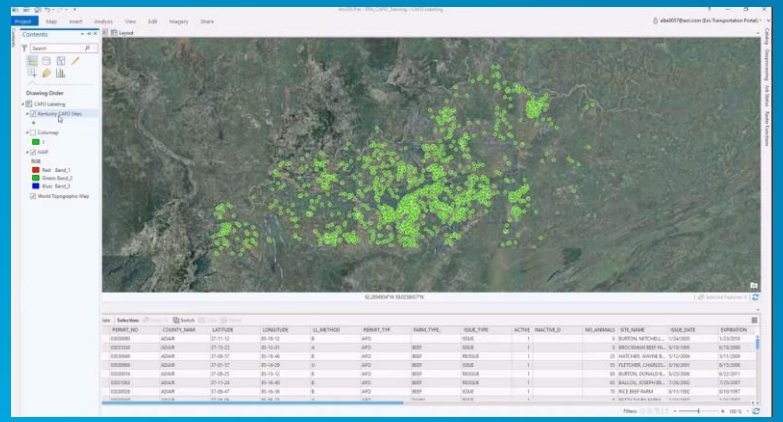
Request Access by sending an email to github_admin@esri.com

Imagery Tools for Deep Learning

Export Training Data For Deep Learning

Available with Spatial Analyst / Image Analyst license

GIS Feature Class



Imagery Data

Bounding Boxes

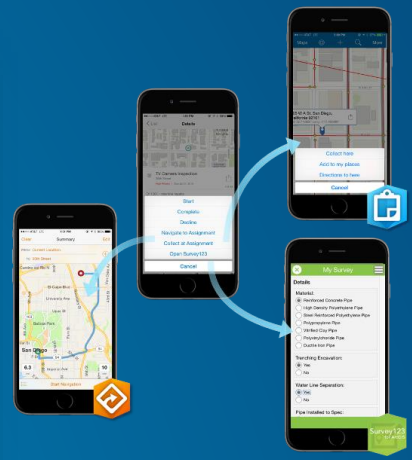
Labelled Pixels

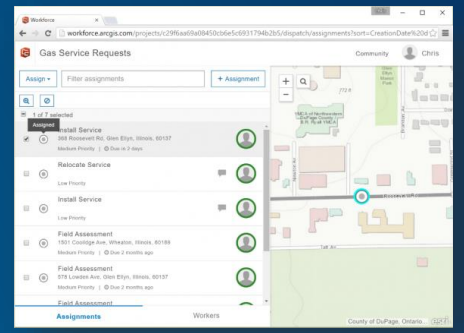
Link for tool here:
http://prodev.arcgis.com/en/pro-app/tool-reference/image-analyst/export-training-data-for-deep-learning.htm

Advanced Object Detection

Detecting complex Objects from Satellite Imagery using a trained deep learning CNN TensorFlow Model. The model is called on-the-fly via a Raster Function from Pro. Detected Sites are then converted to vector points to apply further spatial analytics

1. **ArcGIS *"Export Training Data For Deep Learning"*** GP tool used to Prepare the a labelled training data set from feature class





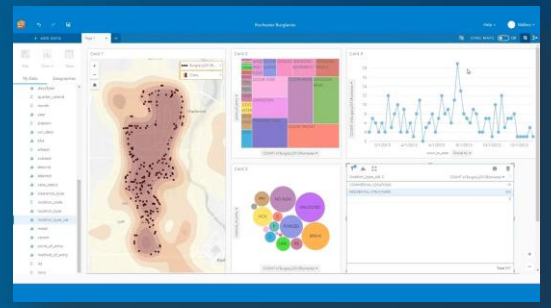
Microsoft CNTK

TensorFlow

2. CNTK or TensorFlow used to train a CNN to detect objects of interest using the labelled training data set

3. ArcGIS Imagery tools used for imagery management and analytics. A Raster Function is used to call the trained CNN and generate the results directly at Pro, allowing for further vector and raster analytics

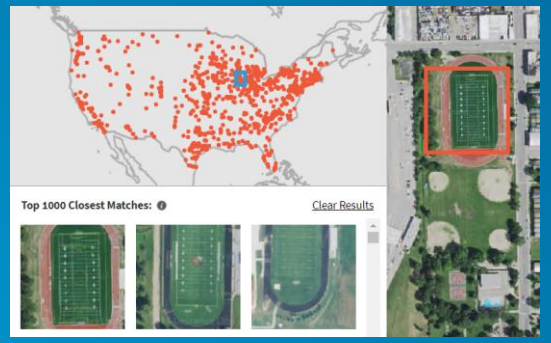
4. ArcGIS Field Apps like Workforce + Collector and Survery123 could be used to plan inspections for detected sites. Operations Dashboard could be used to monitor execution of assigned tasks in Real-Time Inspections Results could then be analyzed in ArcGIS



Deep Learning with Imagery

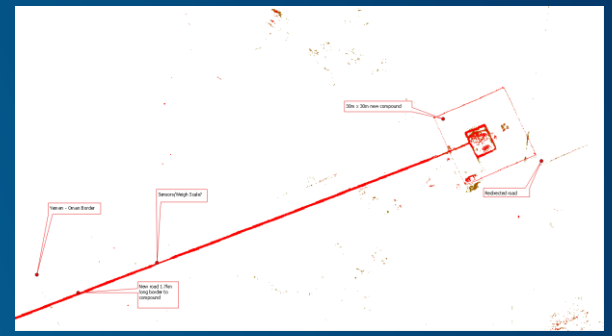
Quantify Oil reserve from the shadow of oil rigs

Visual Map Search (Similar Places)

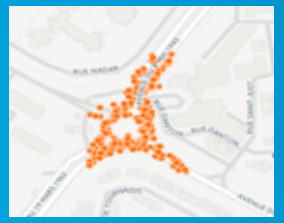


Abnormal Objects Pattern (Patriot Missiles that are near RUH)

Smuggling Attempts Detection near Oman-Yemen border

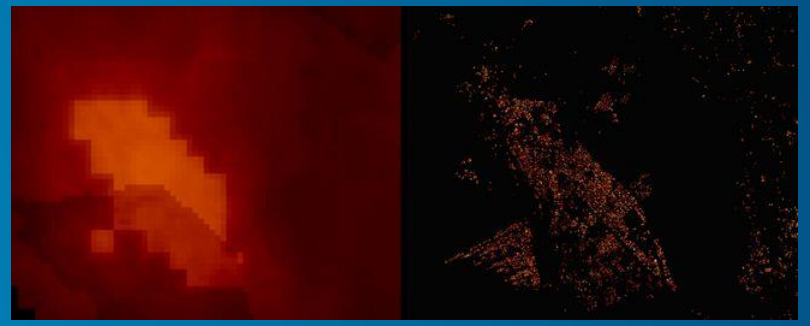


Mapping the unmapped (HD Maps) Roundabouts from GPS data



Auto-Registration of New Roads in un-developed areas (OSM data + Imagery)

More accurate Population Densities (8K labelled images > detecting buildings)



Visual Map Search (Similar Places)

Mapping the unmapped

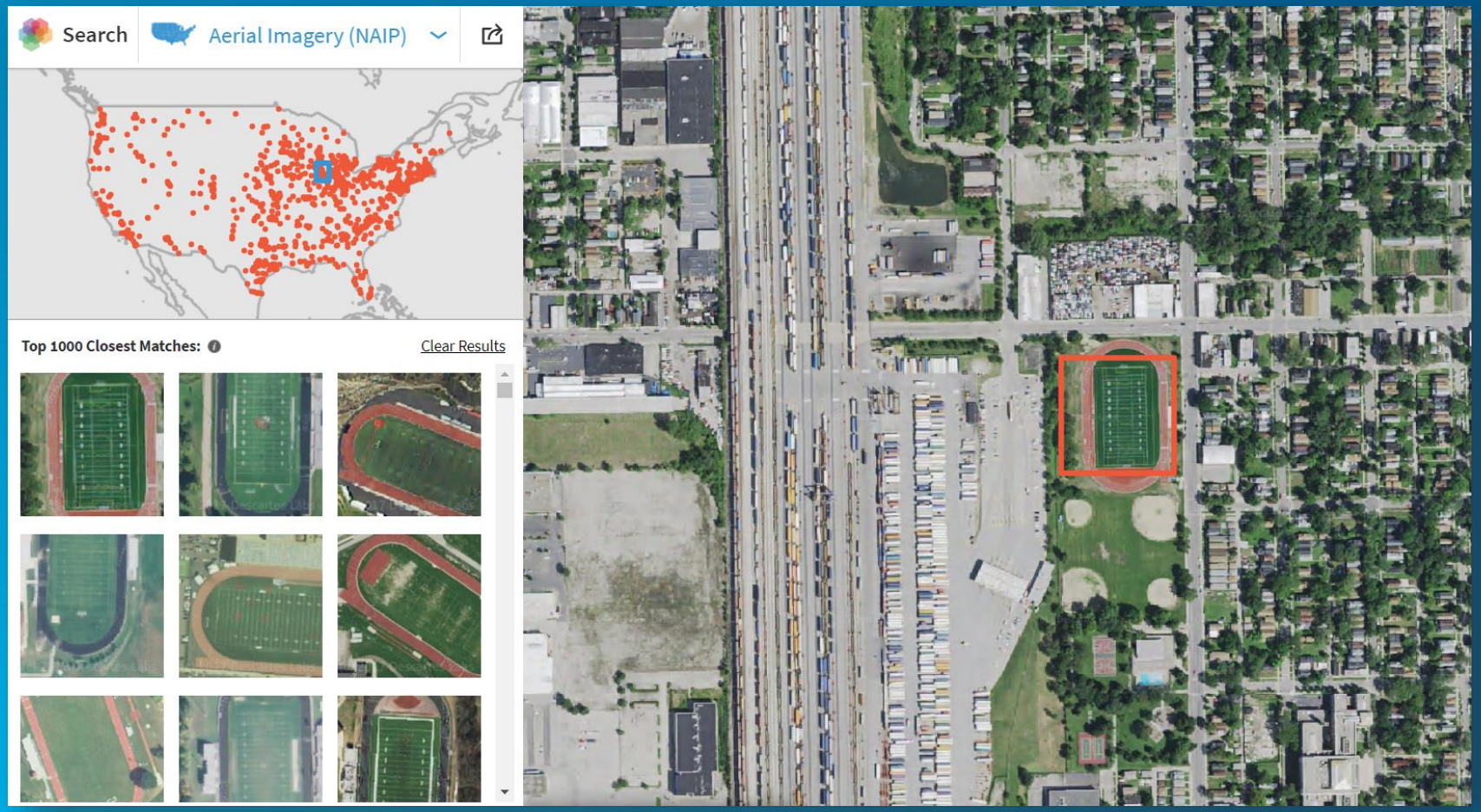
Auto-Registration of new Map Features from Satellite Imagery

1. Training: left Satallite Imagery, right OSM Data

2. Predicting: left: input image, © Mapbox Satellite; middle: OpenStreetMap data; right: Deep Learning model

Visual Map Search

Using Deep Learning to find Similar Locations all over US. Possible use-cases include: locations with specific type of missile launchers.



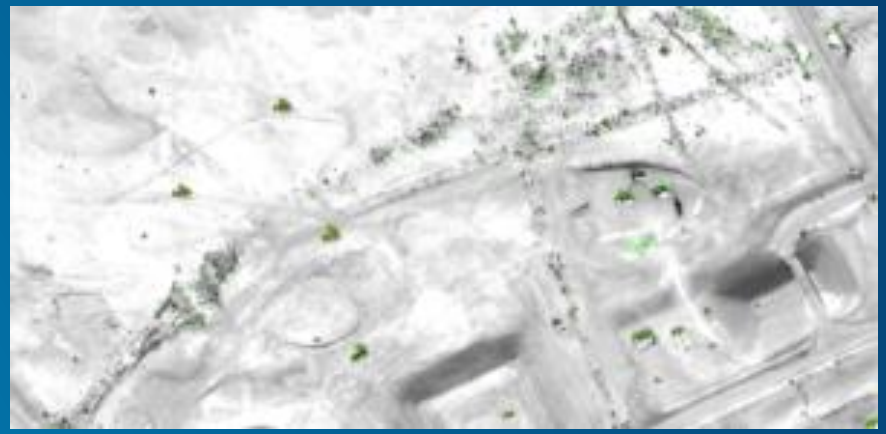
Possible Use-cases

- **Defense**: spotting places with missile launchers and specific geospatial features (e.g. proximity to borders)
- **Real-Estate**: quickly spotting places with specific kind of amenities (e.g. marketplace, big pools, green areas..etc)
- **Agriculture**: highlighting locations with potential crop diseases

Discovering Patriot Missiles near RUH

Default Images, nothing abnormal

Anomalous objects with specific pattern that just appeared recently

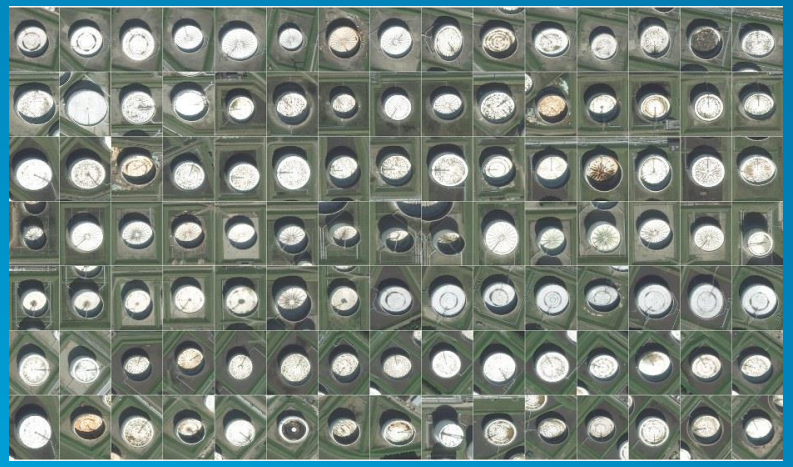


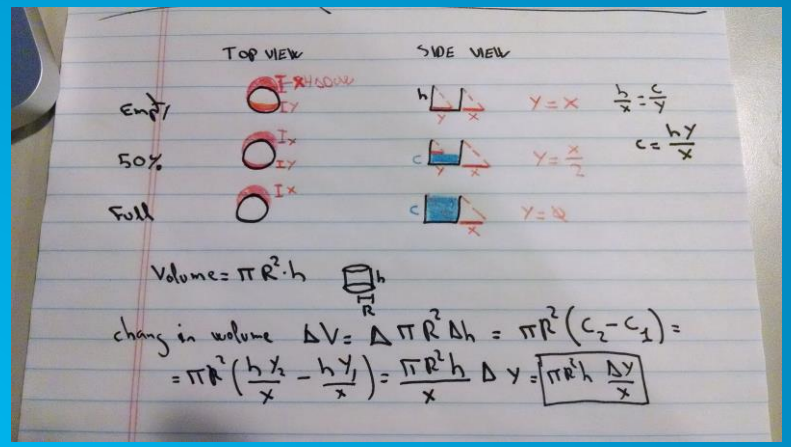
Not all anomalies are captured, but once that exceed specific threshold and meet specific pattern(s)

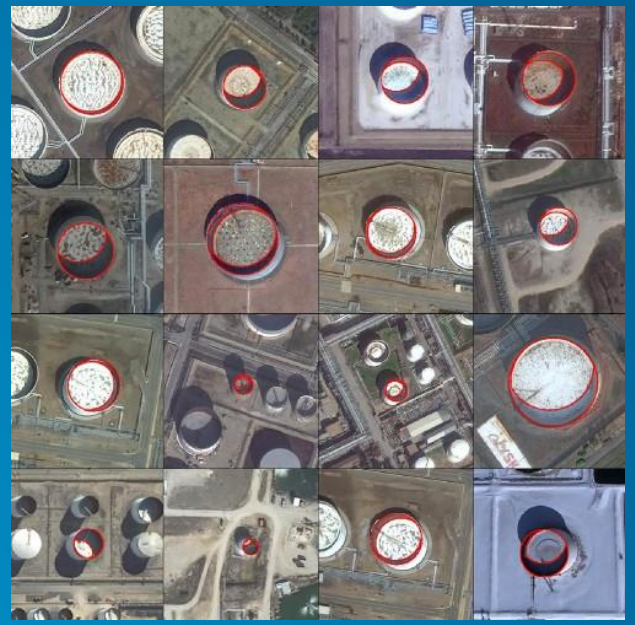
Patriot Missiles that are near RUH (Riyadh International Airport) Identified

Quantify Oil reserve from the shadow of oil rigs

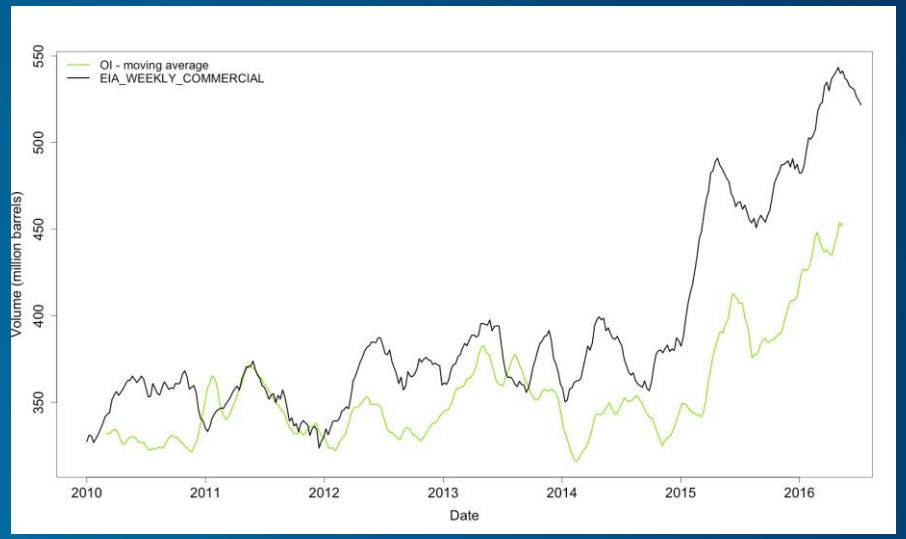
Shadow-detection algorithms continuously monitors global oil storage and produces daily-updated estimates of individual tank volumes







AI estimate of total US oil volume




esri
THE
SCIENCE
OF
WHERE