

Microsoft Research

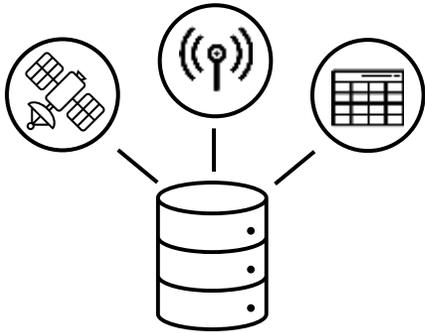
Summit 2022

Model training in FarmVibes.AI

Rafael Padilha

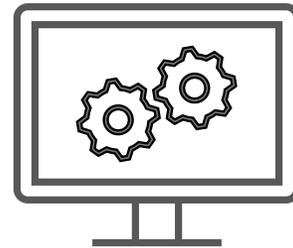
Microsoft Research

Empowering ML pipelines



Data Ingestion

Sentinel
Landsat
USDA
USGS 3DEP



Data Processing

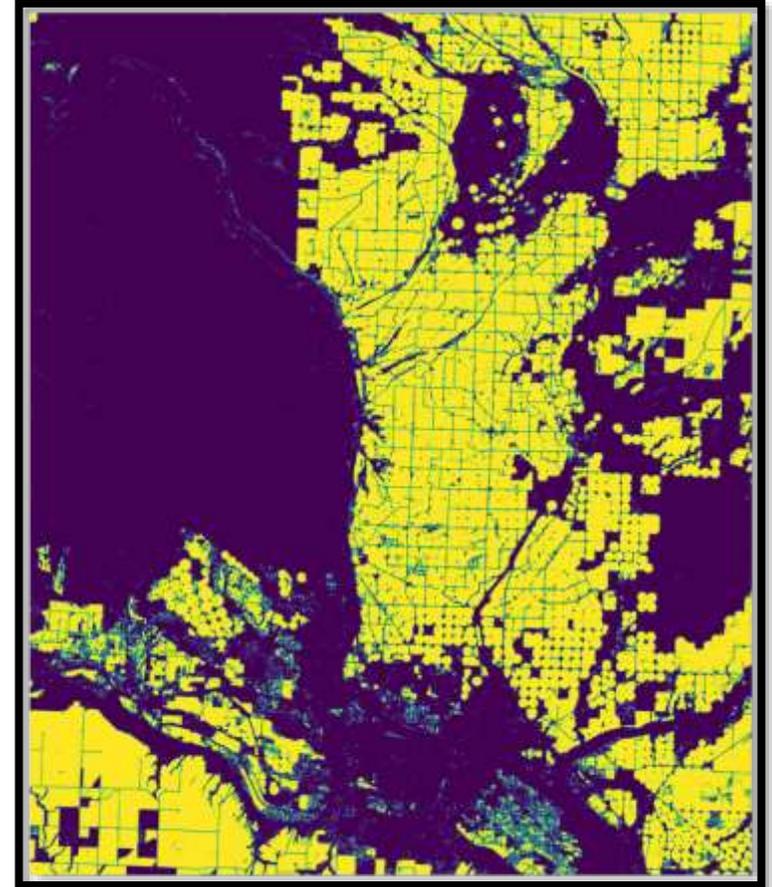
Index Computation
Linear Trends
Aggregation



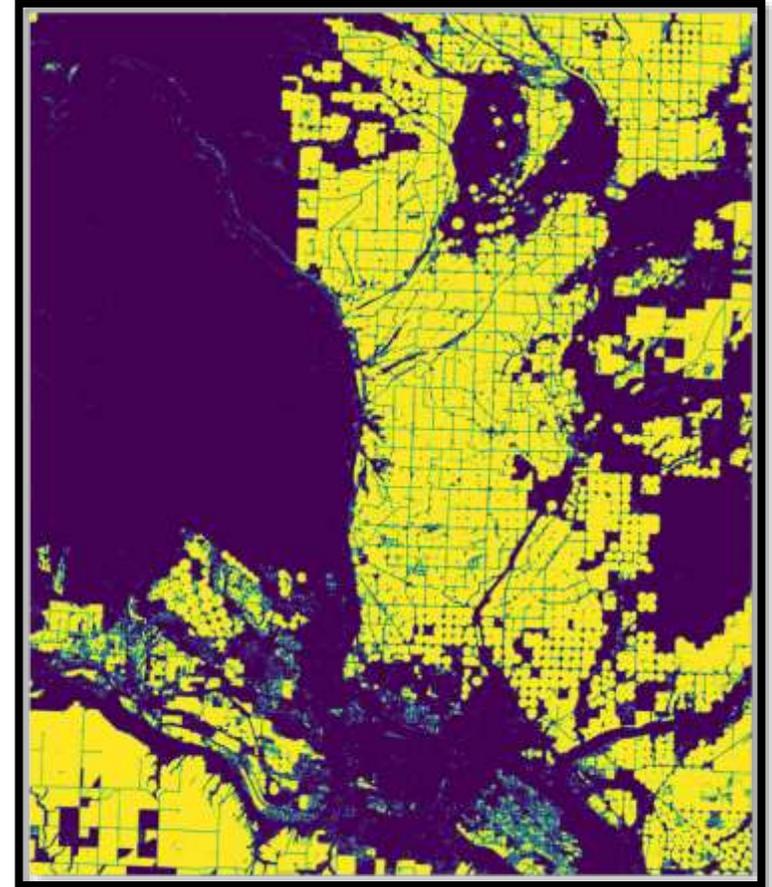
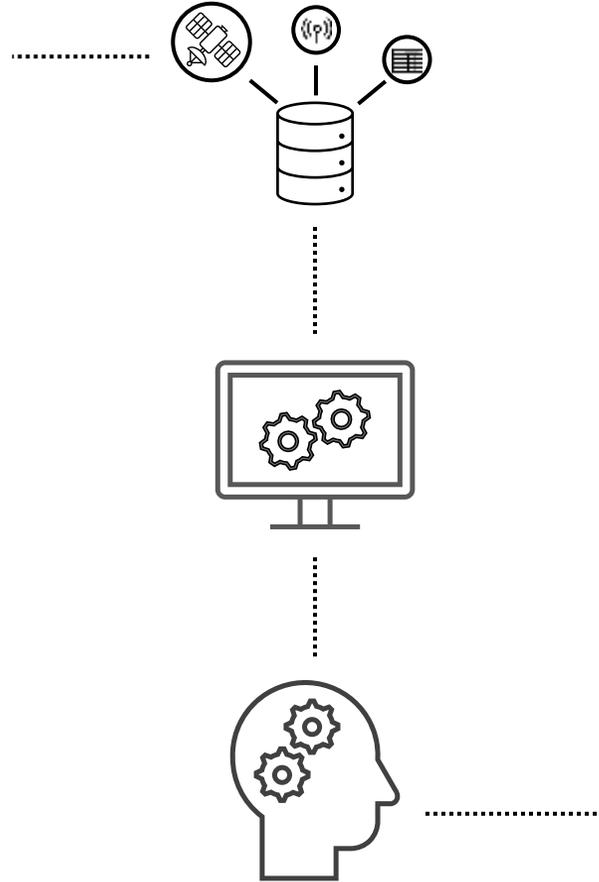
Model Lifecycle

Dataset Generation
Model Training
Inference

Crop Land Segmentation



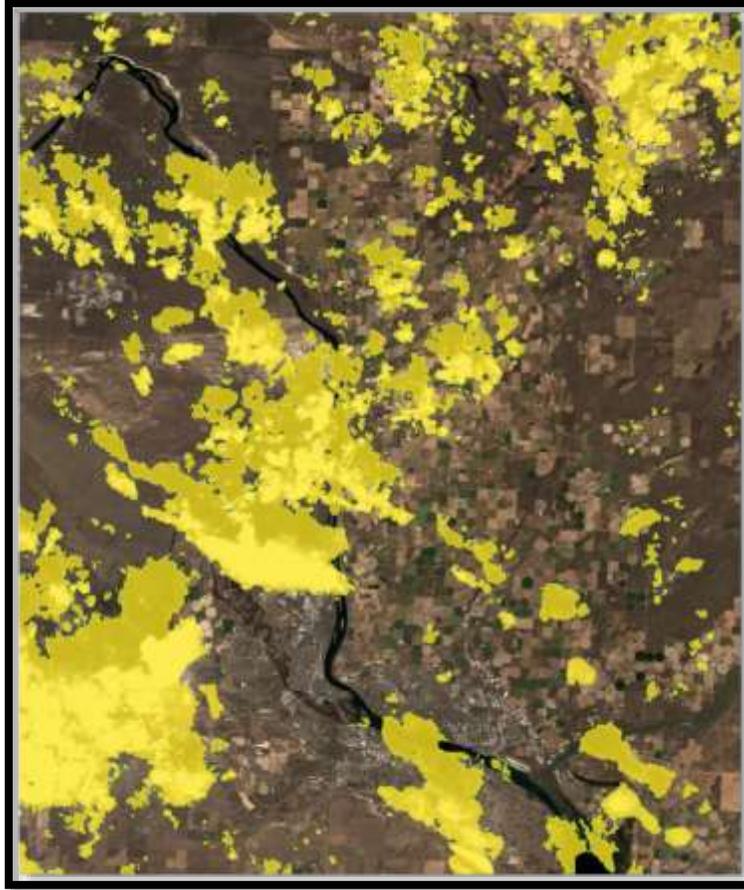
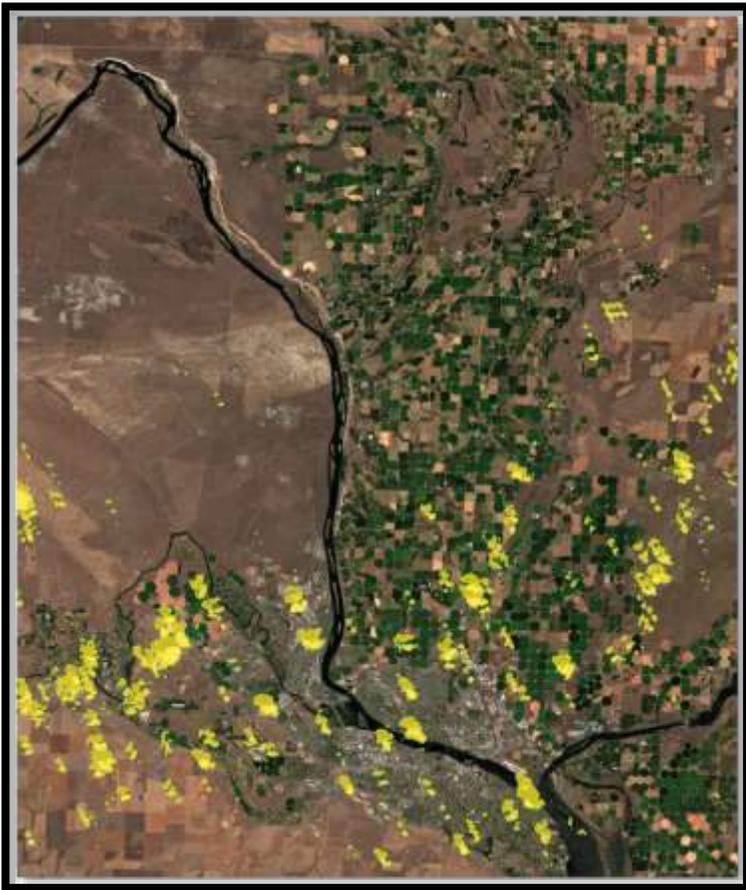
Notebook Overview



Sentinel-2



Sentinel-2



Cloud and Shadow Masks

SpaceEye Cloud Removal

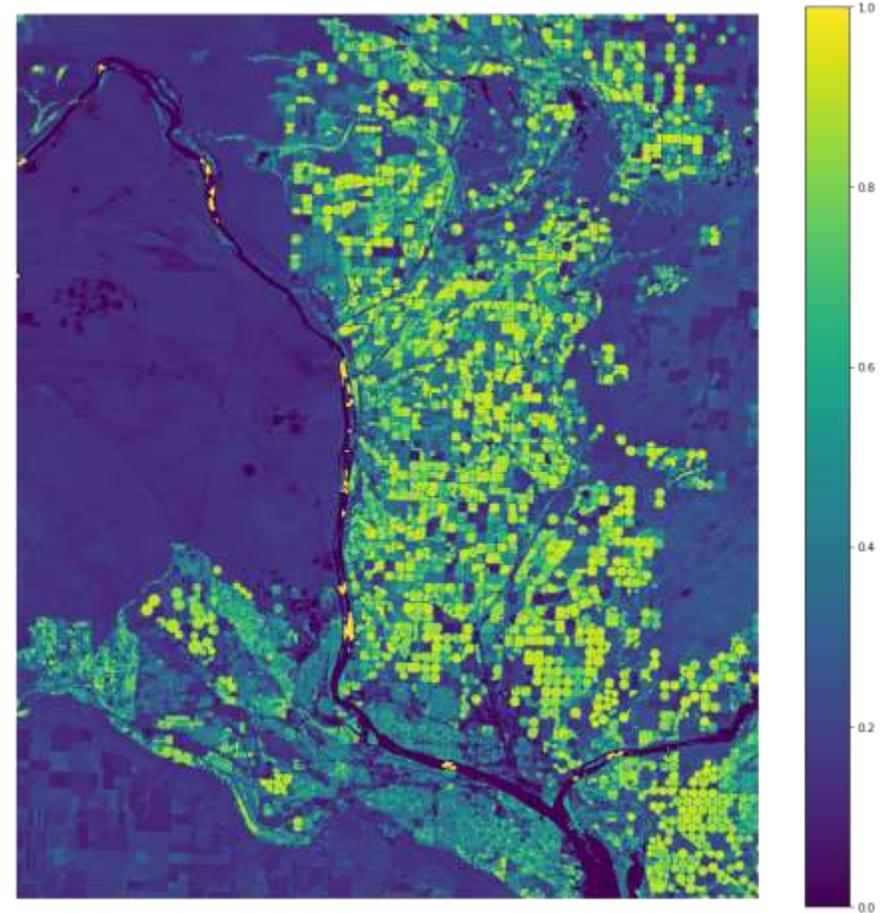


Cloud-free Image

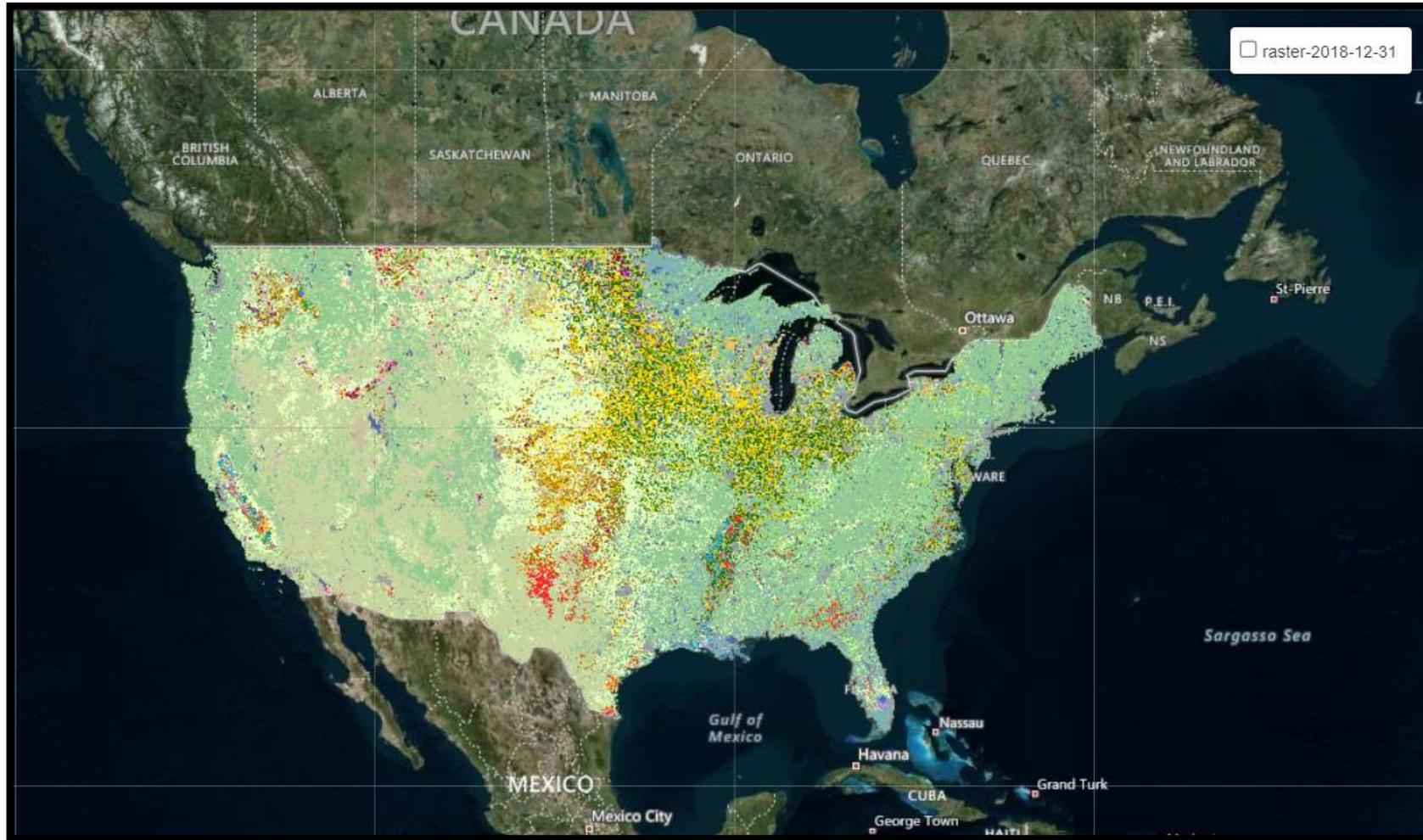
Normalized Difference Vegetation Index



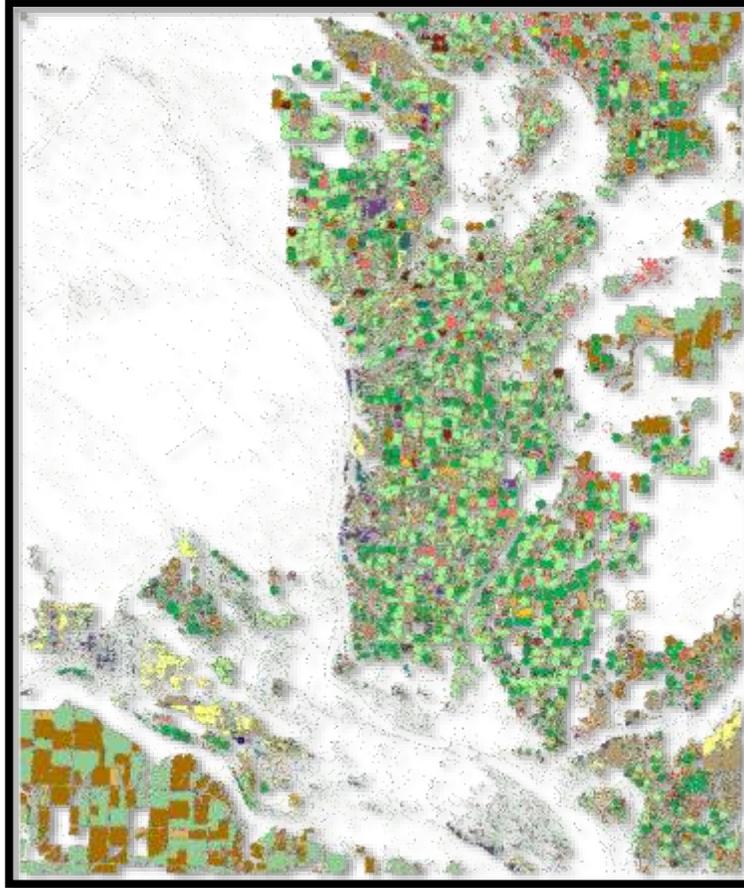
$$\text{NDVI} = \left(\frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} \right)$$



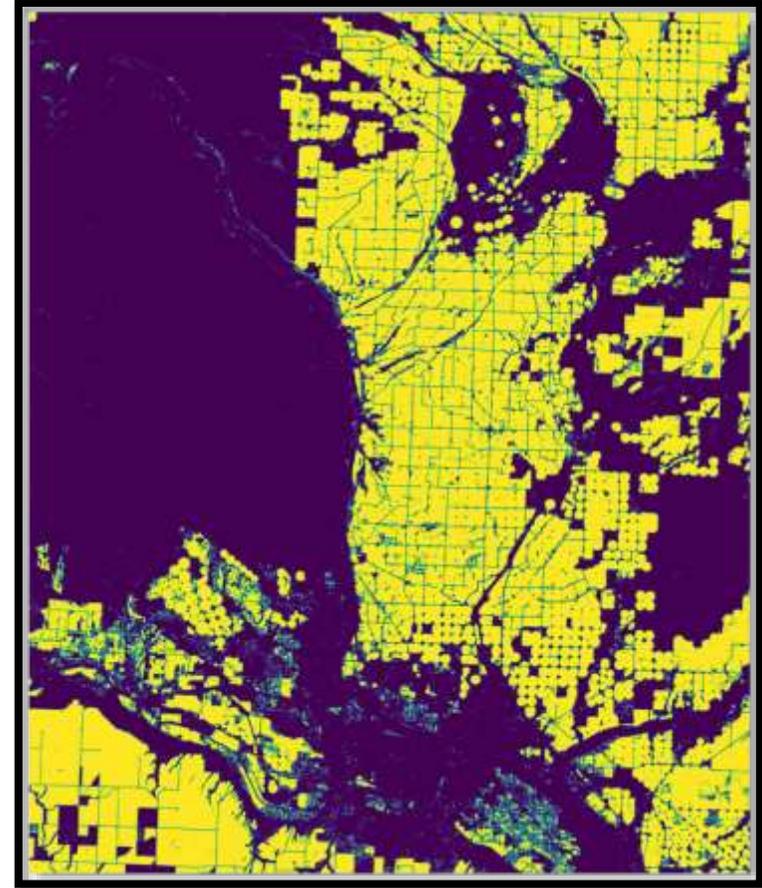
USDA Crop Data Layer (CDL)



USDA Crop Data Layer (CDL)

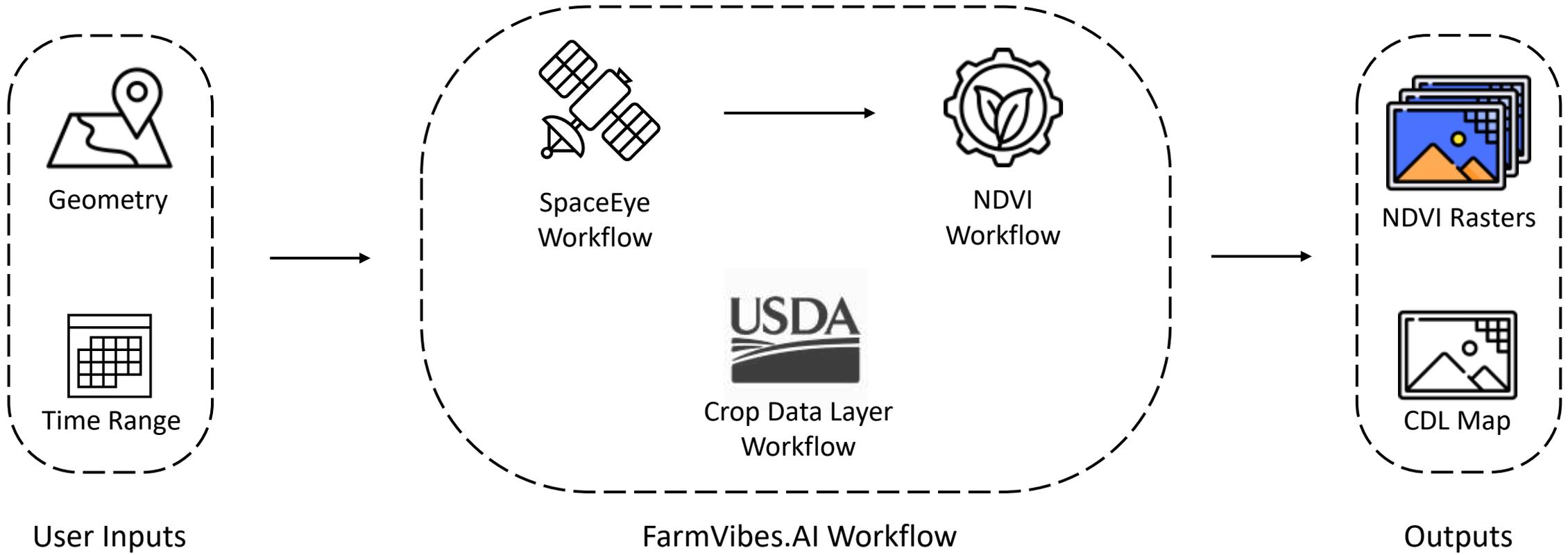


CDL Map

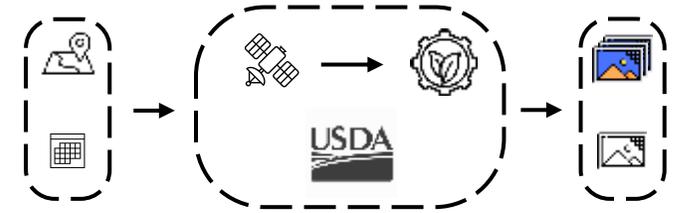


Crop Land Segmentation Map

FarmVibes.AI Dataset Generation Workflow



The Anatomy of a Workflow

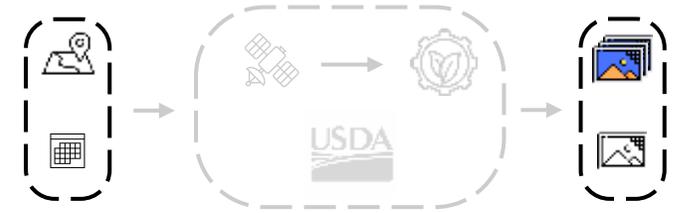


```

File Edit Selection View Go Run Terminal Help datagen_crop_segmentation.yaml - TerraVibes [SSH: AzureVM] - Visual Studio Code
1  name: datagen_crop_segmentation
2  sources:
3
4
5
6  sinks:
7
8
9  parameters:
10
11 tasks:
12
13
14
15
16
17
18
19
20
21
22 edges:
23
24
25

```

The Anatomy of a Workflow

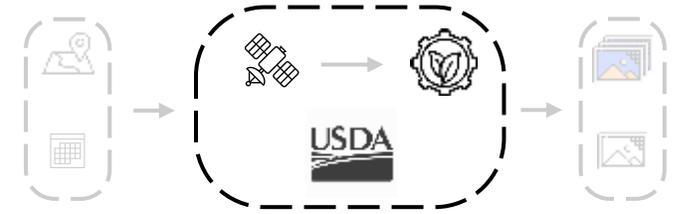


```

1  name: datagen_crop_segmentation
2  sources:
3    user_input:
4
5
6  sinks:
7    ndvi:
8    cdl:
9  parameters:
10
11 tasks:
12
13
14
15
16
17
18
19
20
21
22 edges:
23
24
25

```

The Anatomy of a Workflow

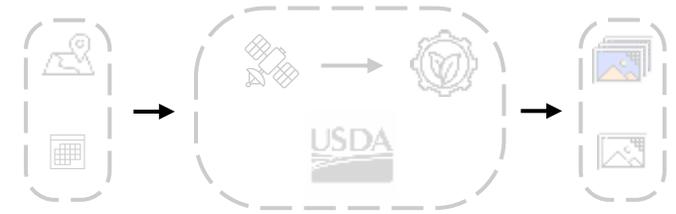


```

1  name: datagen_crop_segmentation
2  sources:
3    user_input:
4
5
6  sinks:
7    ndvi:
8    cdl:
9  parameters:
10
11 tasks:
12   spaceeye:
13
14
15
16   ndvi:
17
18
19
20   cdl:
21
22 edges:
23
24
25

```

The Anatomy of a Workflow

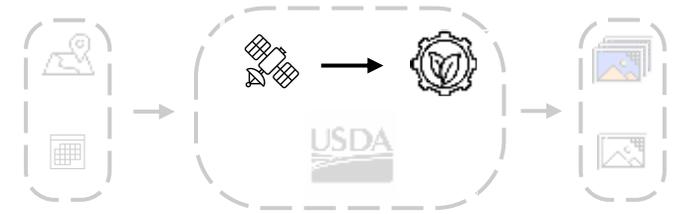


```

1  name: datagen_crop_segmentation
2  sources:
3    user_input:
4      - spaceeye.user_input
5      - cdl.user_input
6  sinks:
7    ndvi: ndvi.index_raster
8    cdl: cdl.raster
9  parameters:
10
11 tasks:
12   spaceeye:
13
14
15
16   ndvi:
17
18
19
20   cdl:
21
22 edges:
23
24
25

```

The Anatomy of a Workflow

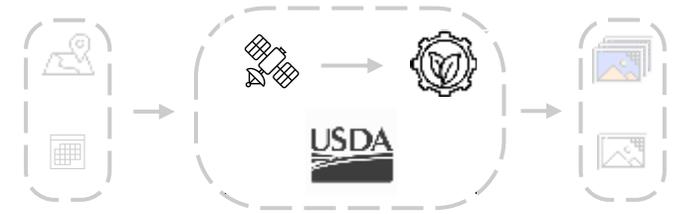


```

1  name: datagen_crop_segmentation
2  sources:
3    user_input:
4      - spaceeye.user_input
5      - cdl.user_input
6  sinks:
7    ndvi: ndvi.index_raster
8    cdl: cdl.raster
9  parameters:
10
11 tasks:
12   spaceeye:
13
14
15
16   ndvi:
17
18
19
20   cdl:
21
22 edges:
23   - origin: spaceeye.split.rasters
24     destination:
25       - ndvi.raster

```

The Anatomy of a Workflow

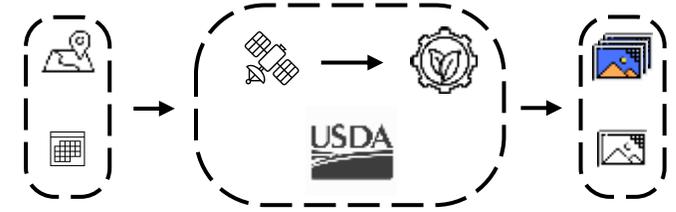


```

1  name: datagen_crop_segmentation
2  sources:
3    user_input:
4      - spaceeye.user_input
5      - cdl.user_input
6  sinks:
7    ndvi: ndvi.index_raster
8    cdl: cdl.raster
9  parameters:
10
11 tasks:
12   spaceeye:
13     workflow: data_ingestion/spaceeye/spaceeye_interpolation
14
15
16   ndvi:
17     workflow: data_processing/index/index
18
19
20   cdl:
21     workflow: data_ingestion/cdl/download_cdl
22 edges:
23   - origin: spaceeye.split.rasters
24     destination:
25       - ndvi.raster

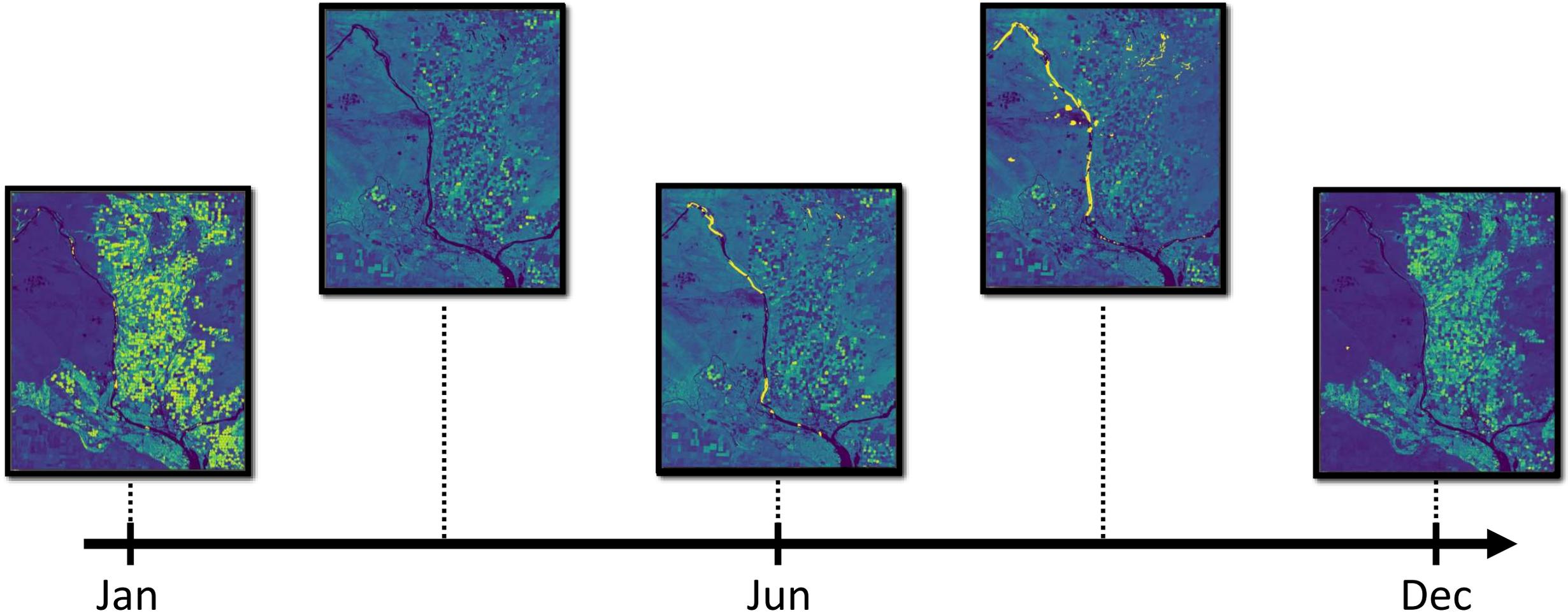
```

The Anatomy of a Workflow

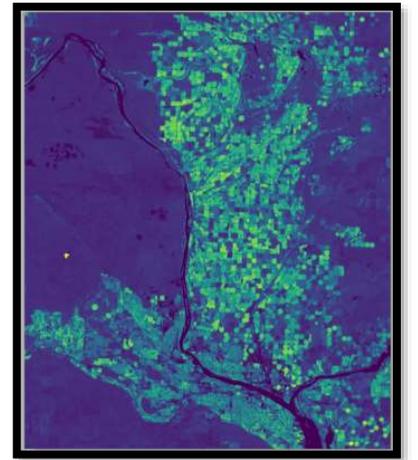
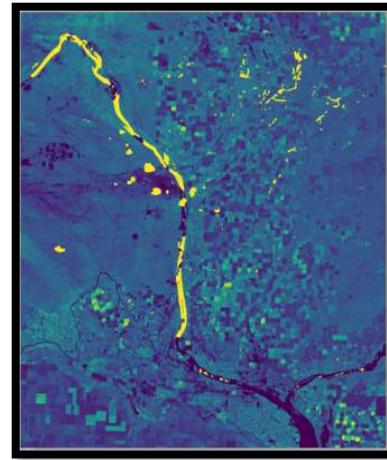
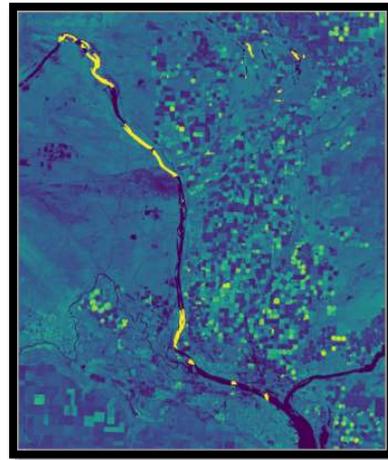
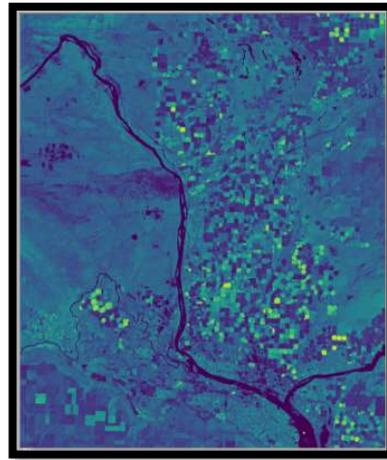
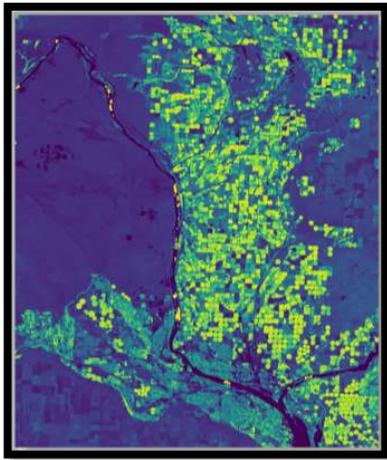


```
File Edit Selection View Go Run Terminal Help datagen_crop_segmentation.yaml - TerraVibes [SSH: AzureVM] - Visual Studio Code
1 name: datagen_crop_segmentation
2 sources:
3   user_input:
4     - spaceeye.user_input
5     - cdl.user_input
6 sinks:
7   ndvi: ndvi.index_raster
8   cdl: cdl.raster
9 parameters:
10  pc_key:
11 tasks:
12   spaceeye:
13     workflow: data_ingestion/spaceeye/spaceeye_interpolation
14     parameters:
15       pc_key: "@from(pc_key)"
16   ndvi:
17     workflow: data_processing/index/index
18     parameters:
19       index: ndvi
20   cdl:
21     workflow: data_ingestion/cdl/download_cdl
22 edges:
23   - origin: spaceeye.split.rasters
24     destination:
25       - ndvi.raster
```

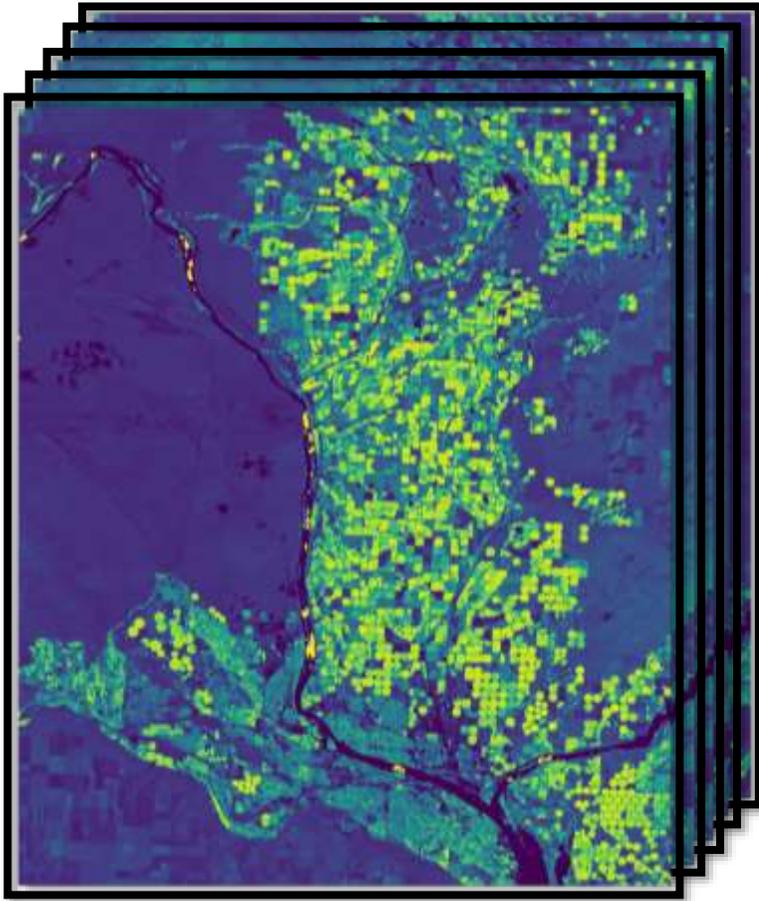
NDVI Temporal Sampling



NDVI Temporal Sampling



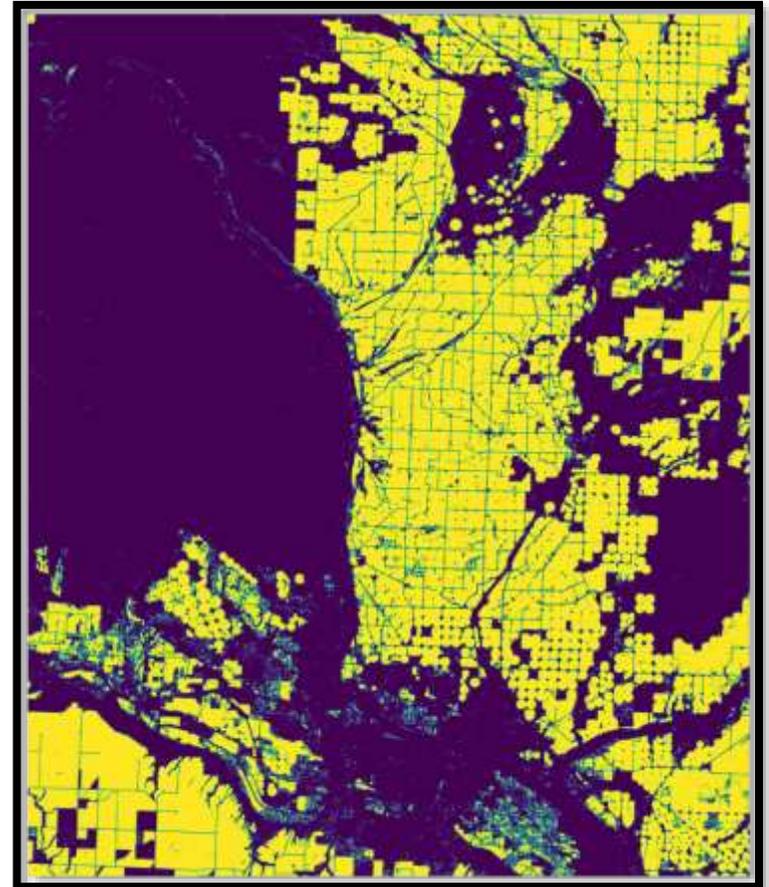
Model Training



NDVI Stack



Local Training
Azure Machine Learning



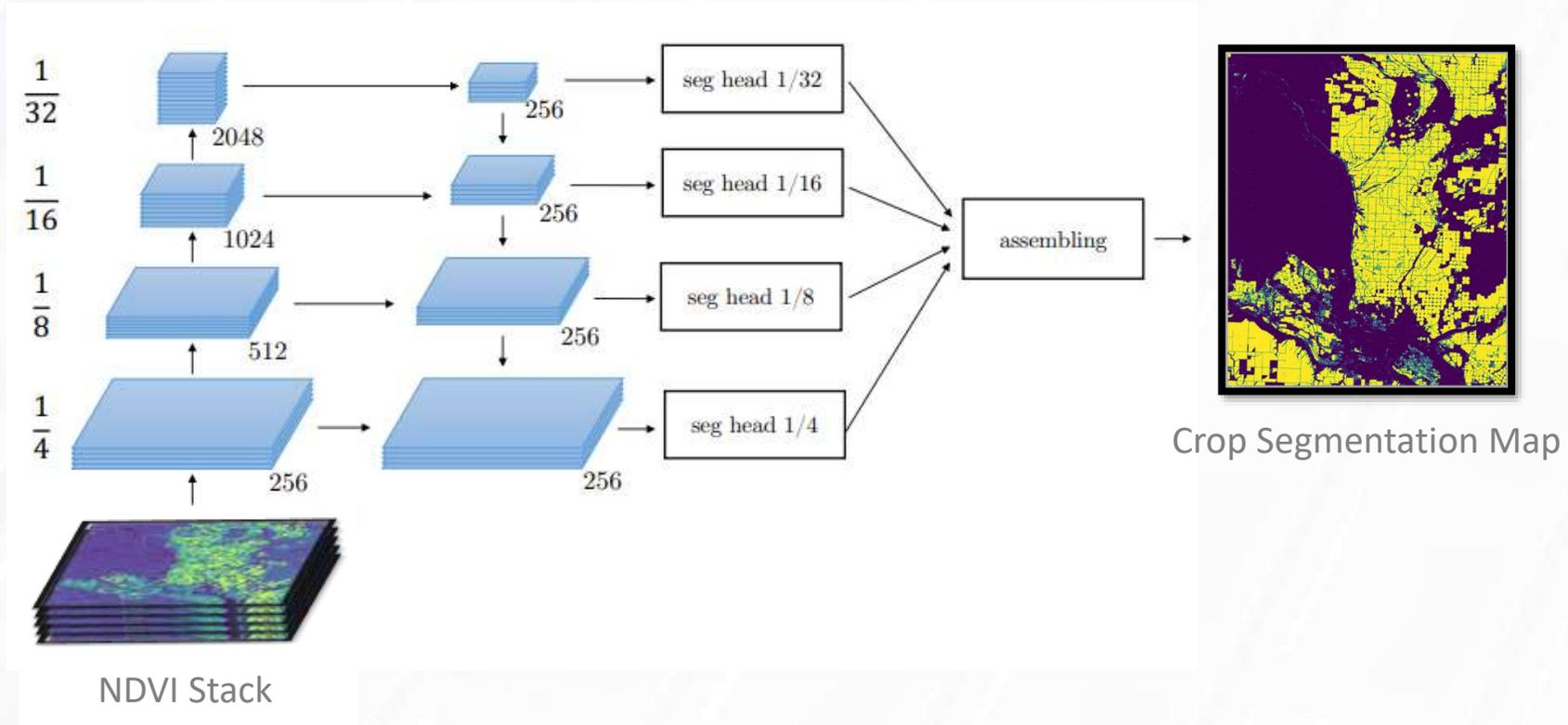
Crop Land Segmentation Map

Model Training

Feature Pyramid Network (FPN)
ResNet34 encoder



Segmentation
Models 



Inference within FarmVibes.AI

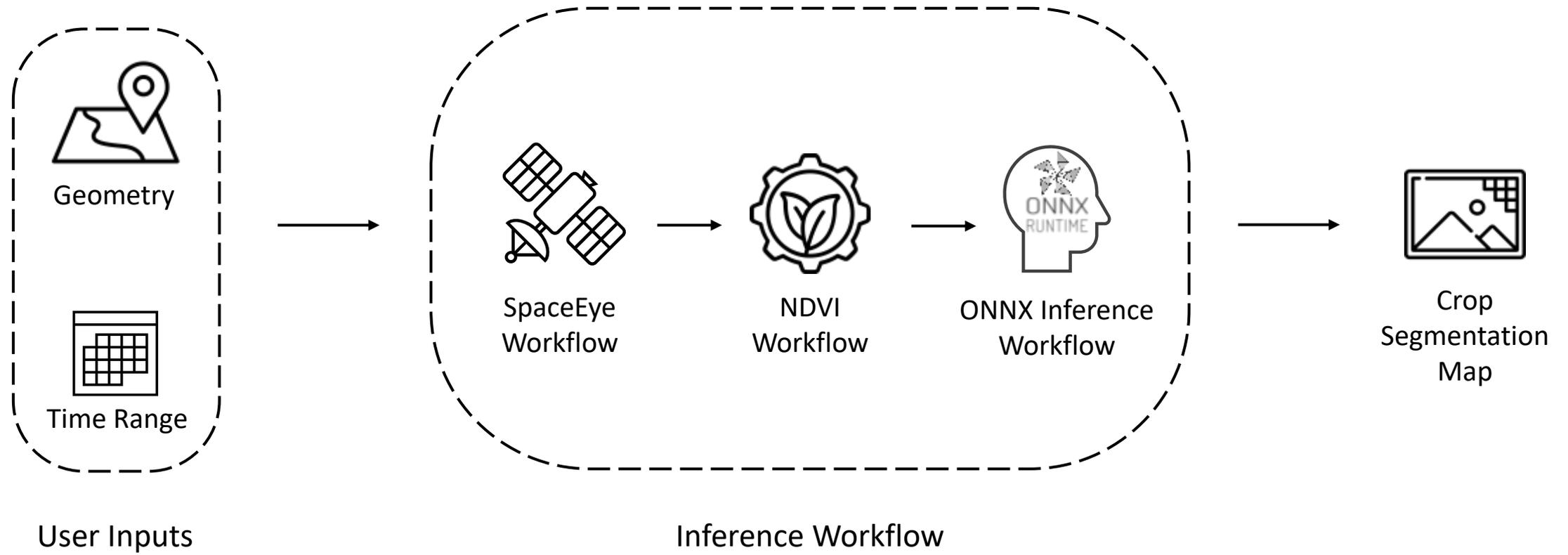


Cross-platform training and inference accelerator

Supports most deep learning and classical machine learning frameworks (e.g., PyTorch, scikit-learn)

Compatible with different hardware, drivers and OS, while providing several optimizations

FarmVibes.AI Inference Workflow





Notebook Organization

Notebooks

- `01_dataset_generation` run FarmVibes.AI data generation workflow
- `02_visualize_dataset` explore workflow intermediate outputs
- `03_local_training` train segmentation model locally
- `03_aml_training` integrate training with Azure Machine Learning
- `04_inference` predict segmentation map with exported model

Supporting code

- `datasets.py` loads and preprocesses FarmVibes.AI outputs
- `modules.py` defines the PyTorch Lightning data module
- `models.py` defines model architecture and training configuration
- `constants.py` defines crop indices constants
- `utils.py` implements utility methods



To the notebook...

Sources and image attributions

- Some icons were made with assets from www.flaticon.com