Microsoft

Microsoft Ignite

Getting started on your health-tech journey

Lee Stott
Principal Program Manager- Microsoft

Agenda

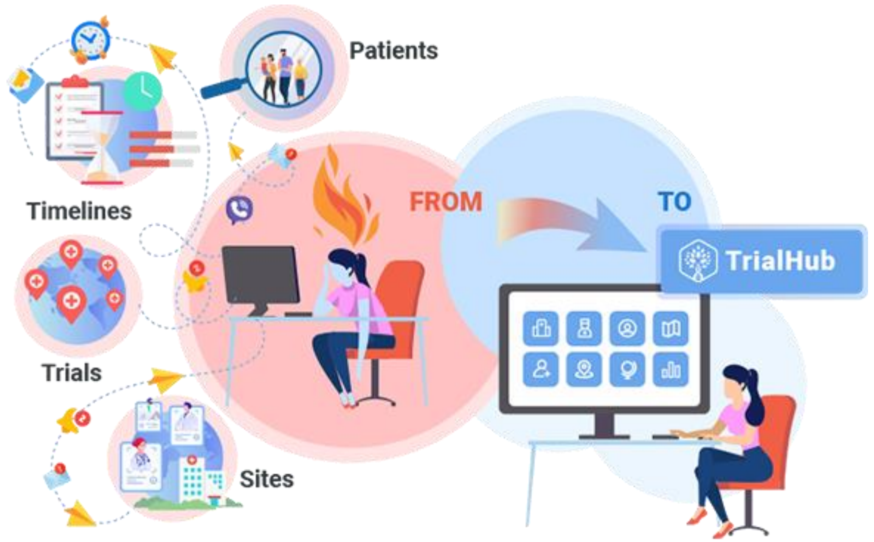
- Introduction to Health Tech - evolution through the decades [5min]
- Responsible AI Principles in Health [1min]
- Architecting your first Responsible AI Health Pipeline in AML using Cardiovascular risk [7min]
- Demo [7min]
- Q&A [10min]

NOVARTIS

Are reimagining medicine. Novartis focuses on innovative medicines as well as generics and biosimilars.
They are a leading global medicines company powered by data and digital.

Harnessing the power of science to push boundaries, develop breakthrough treatments and deliver them directly to the millions of people.

https://www.novartis.com/



FindMeCure:"Bringing Clinical Trials Closer to Patients"

https://www.findmecure.com/

Why go to the GP when you could send your health avatar instead?

Fri 26 Apr 2019 | Iain Buchan

https://techerati.com/features-hub/opinions/why-go-to-the-gp-when-you-could-send-your-health-avatar-instead/

4 Decades of Health-Tech and The Decade Ahead

1980s / 90s → 2000s : Clinical Audit & Governance

- Evidence Based Medicine
- 1-way: research → practice
- Hand-crafted models & scarce data

2010s : Learning Health Systems

- Electronic Health Records and Big Data
- 2-way: research <-> practice
- Models start to be learned from fuller data

2020s : Precision & Population Health Systems

- Personalization & Well Being
- Privacy & Bias
- Population Health & Smart Cities

Credit : University of Liverpool Study for NHSX

Digital Biomarkers

npj | Digital Medicine

NPJ Digit Med. 2019; 2: 108. PMCID: PMC6841669
Published online 2019 Nov 8. doi: 10.1038/s41746-019-0182-1 PMID: 31728415

GPS mobility as a digital biomarker of negative symptoms in schizophrenia: a case control study

Colin A. Depp,[1,2] Jesse Bashem,[2] Raeanne C. Moore,[1,2] Jason L. Holden,[1] Tanya Mikhael,[3] Joel Swendsen,[4] Philip D. Harvey,[5] and Eric L. Granholm[1,2]

▸ Author information ▸ Article notes ▸ Copyright and License information Disclaimer

Associated Data

▸ Supplementary Materials

▸ Data Availability Statement

Abstract

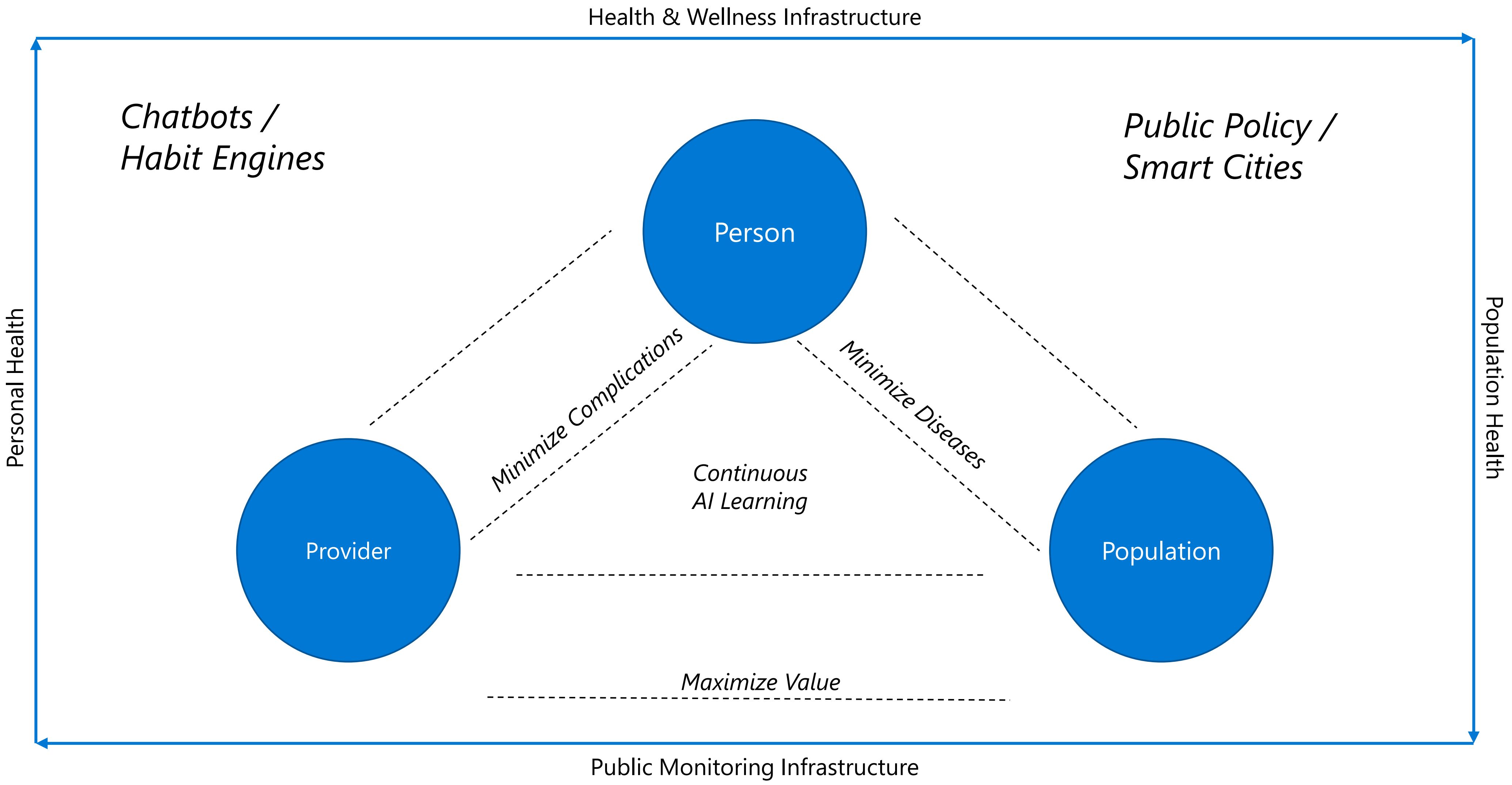
Go to:

Mobility is an important correlate of physical, cognitive, and mental health in chronic illness, and can be measured passively with mobile phone global positional satellite (GPS) sensors. To date, GPS data have been reported in a few studies of schizophrenia, yet it is unclear whether these data correlate with concurrent momentary reports of location, vary by people with schizophrenia and healthy comparison subjects, or associate with symptom clusters in schizophrenia. A total of 142 participants with schizophrenia ($n = 86$) or healthy comparison subjects ($n = 56$) completed 7 days of ecological momentary assessment (EMA) reports of location and behavior, and simultaneous GPS locations were tracked every five minutes. We found that GPS-derived indicators of average distance travelled overall and distance from home, as well as percent of GPS samples at home were highly correlated with EMA reports of location at the day- and week-averaged level. GPS-based mobility indicators were lower in schizophrenia with medium to large effect sizes. Less GPS mobility was related to greater negative symptom severity, particularly diminished motivation, whereas greater GPS mobility was weakly associated with more community functioning. Neurocognition, depression, and positive symptoms were not associated with mobility indicators. Therefore, passive GPS sensing could provide a low-burden proxy measure of important outcomes in schizophrenia, including negative symptoms and possibly of functioning. As such, passive GPS sensing could be used for monitoring and timely interventions for negative symptoms in young persons at high risk for schizophrenia.

Subject terms: Health care, Medical research

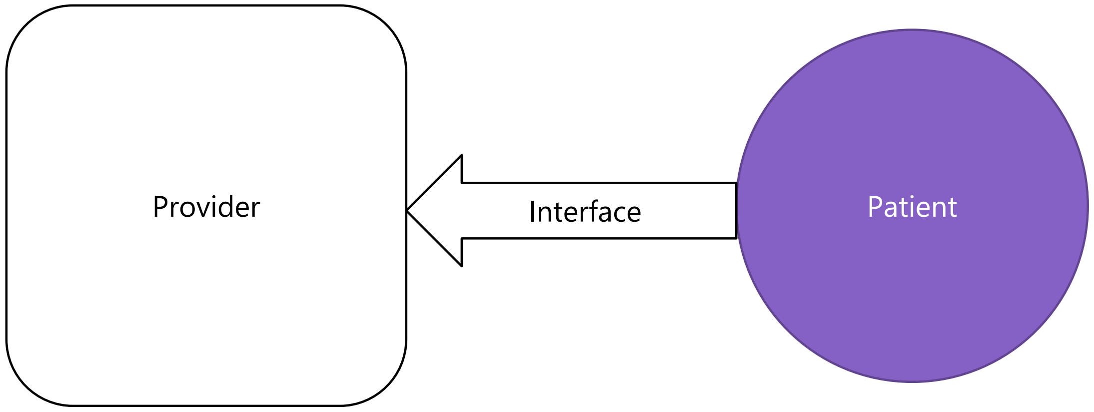
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6841669/

Health Stakeholder Relationships

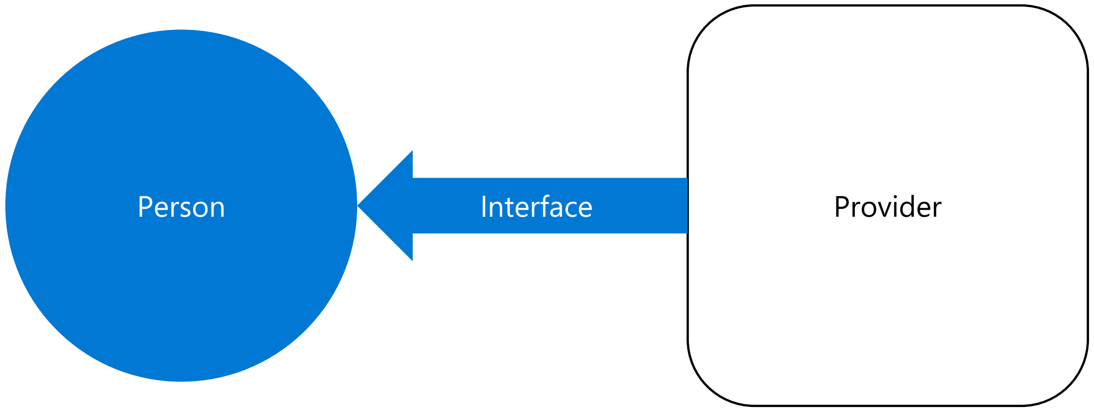


Prof. Iain Buchan, Chair in Public Health and Clinical Informatics, University of Liverpool

Crossing the Trust-chasm



Clinic-centred supply chain
Episodic Treatments



Person –centred supply chain
Preventive / Journey

Responsible AI Principles

- Fairness
  - AI systems should treat all people fairly
- Inclusiveness
  - AI systems should empower everyone and engage people
- Reliability & Safety
  - AI systems should perform reliably and safely
- Transparency
  - AI systems should be understandable
- Privacy & Security
  - AI systems should be secure and respect privacy
- Accountability
  - AI systems should have algorithmic accountability

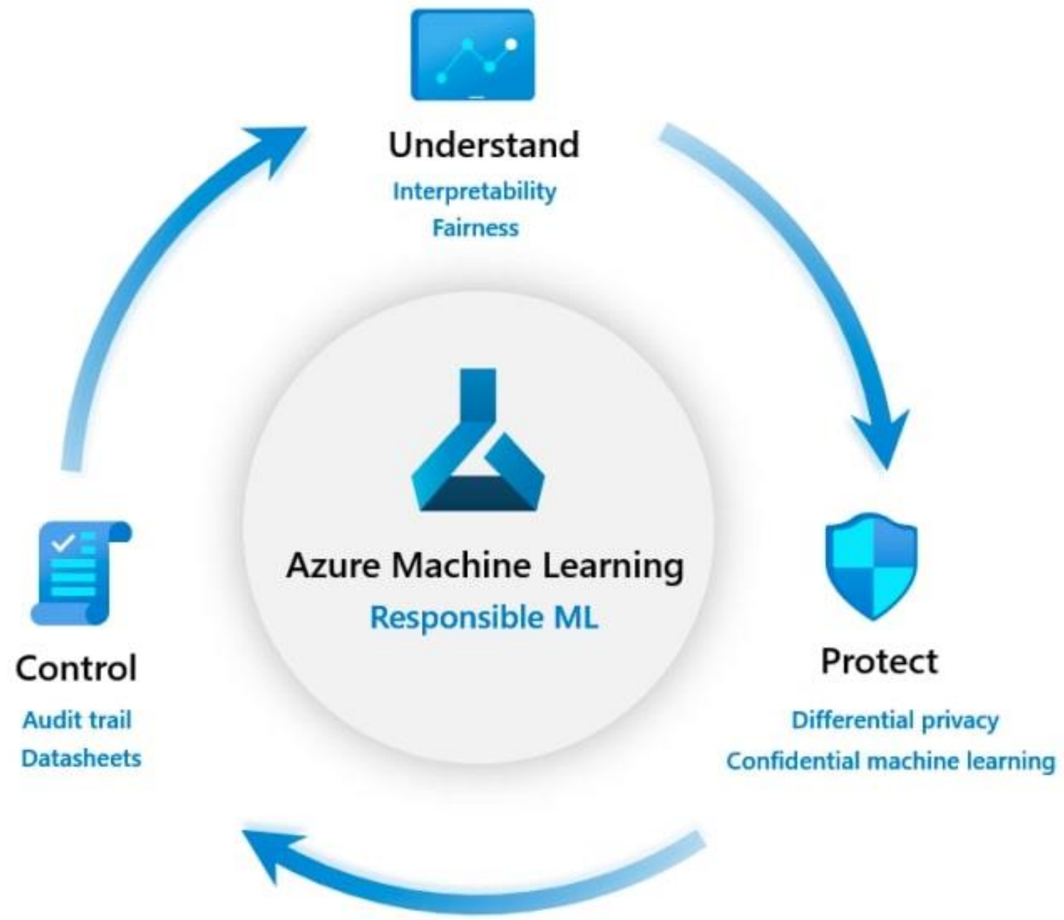
Read More: https://www.microsoft.com/en-us/ai/ai-business-school

Responsible ML

Responsible ML encompasses the following values and principles:

- Understand machine learning models
- Protect people and their data
- Control the end-to-end machine learning process

http://aka.ms/responsibleML



Responsible ML Healthcare

Goal: Detect if a person is suitable for receiving a treatment for heart disease.

Use Azure Machine Learning as a tool to cover all the Machine Learning and Responsible AI workflow

01

Data anonymization

Preprocessing Data.

02

Data protection

Apply privacy techniques to protect sensitive data and prevent leaks

03

Data understanding

Exploratory Analysis and Data preprocessing

04

Model training

Train model over the preprocessed dataset

05

Model understanding

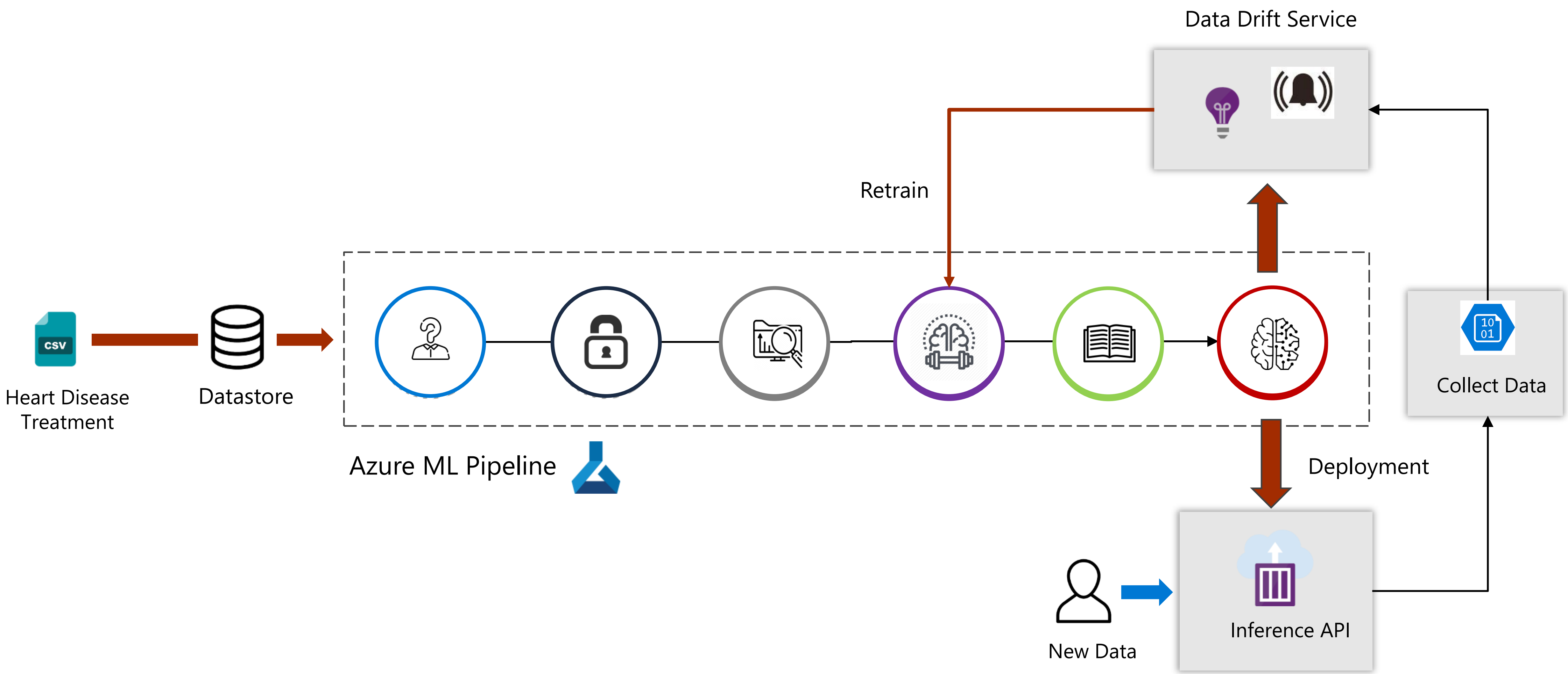
Explain model behavior and test it for fairness

06

Model serving

Register, deploy and monitor the best model with Data Drift

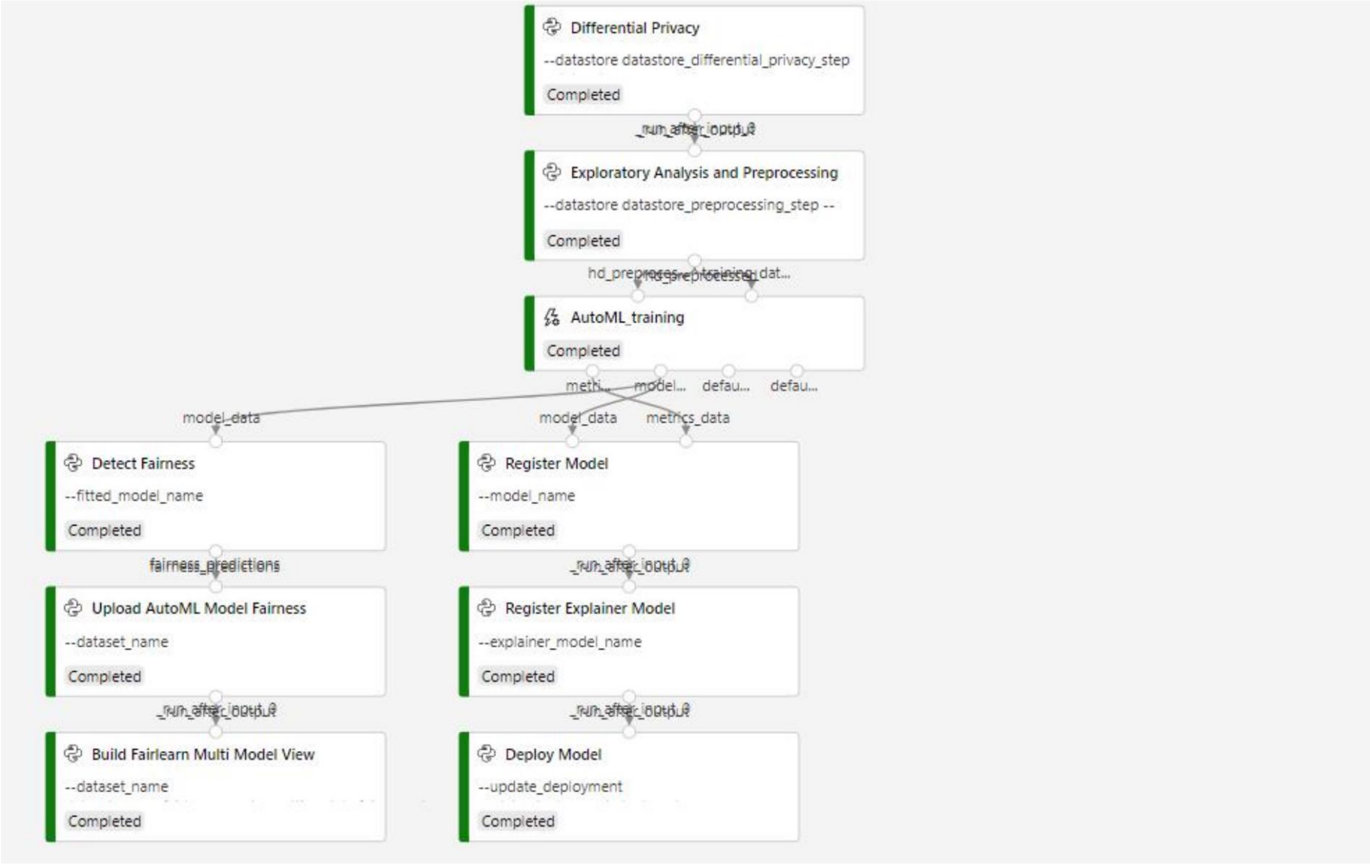
Process Diagram



http://github.com/leestott/ResponsibleAI

Azure ML Pipeline

- An Azure Machine Learning pipeline is an independently executable workflow of a complete machine learning task
- It performs a complete logical workflow with an ordered sequence of steps.
- We have developed a programatic pipeline creation based on the described “machine learning workflow”
    - Data Protection
    - Exploratory Analysis and Preprocessing
    - AutoML training
    - Fairness detection
    - Model explanation
    - Model deployment



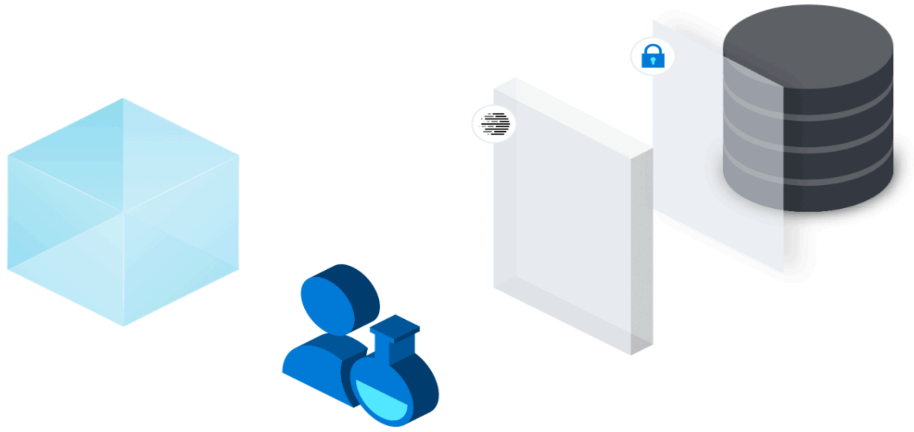
Data anonymization

- For data anonymization process we use **Presidio** to remove Personal Identificable Information from different Text columns of our dataset.
- We detect main entities such as names, locations and remove them from dataset.
- After this step the Data scientists could make the exploratory analysis without see sensible information about patients.

Data Anonymization
Presido
https://github.com/Microsoft/presidio

Differential Privacy

- In this step, we apply differential privacy techniques, which injects precise levels of statistical noise in data to limit the disclosure of sensitive information
- Identify data leaks and intelligently limit repeat queries to manage exposure risk.
- We have integrate the results of add noise to our data in Azure Machine Learning Workspace.



Differential Privacy

Add noise to our data in Azure Machine Learning Workspace.

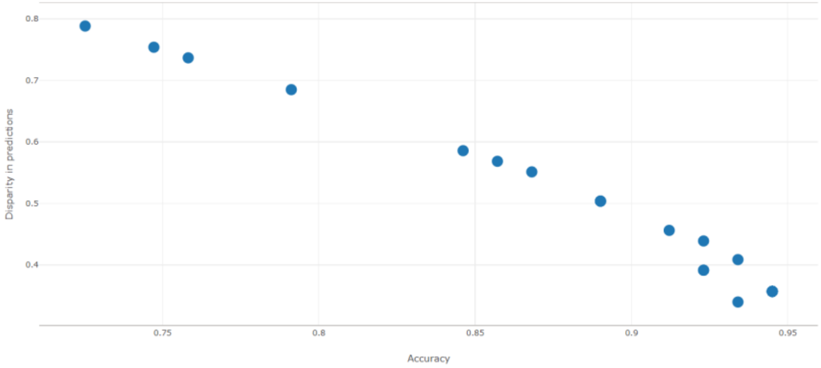
Fairness detection

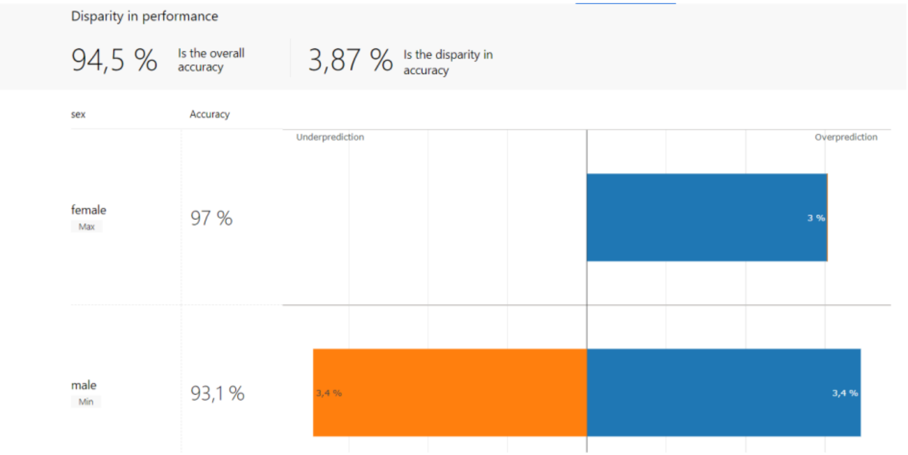
- In this Azure ML Step we use FairLearn to assess the system’s fairness and mitigate any observed unfairness issues.

- We use the fitted automl from the previous step to detect any unfairness over the sensitive features like: gender, pregnant, smoker, diabetic that could be important to decide if a Patient recieve a treatment or not.

- Enable the mitigation of unfairness using the Fairlearn GridSearch algorithm with Demographic parity for gender.

Fairness Detection
Fairlearn
https://fairlearn.org





Model explainability

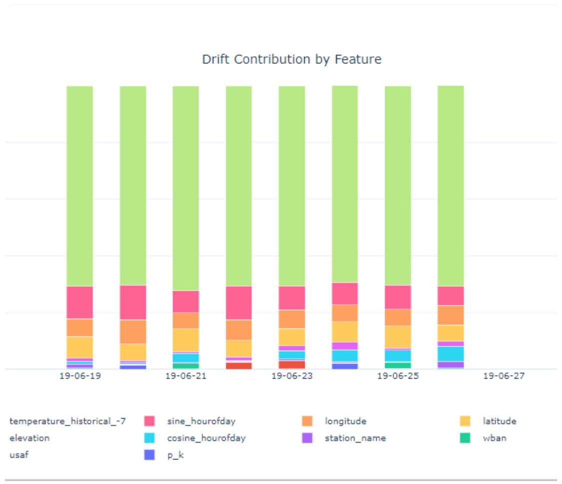
- In this step, we explain our model using Interpret ML. Enabling the capability of explain our AutoML model is important during two main phases.
- **Training**: We activate the AutoML explanation option to verify hypotheses and build trust stakeholders.
- **Inference**: We extract our model explainability for local explanations in order to have transparency around deployed models to know how the model is working and how its decisions are treating and impacting patients

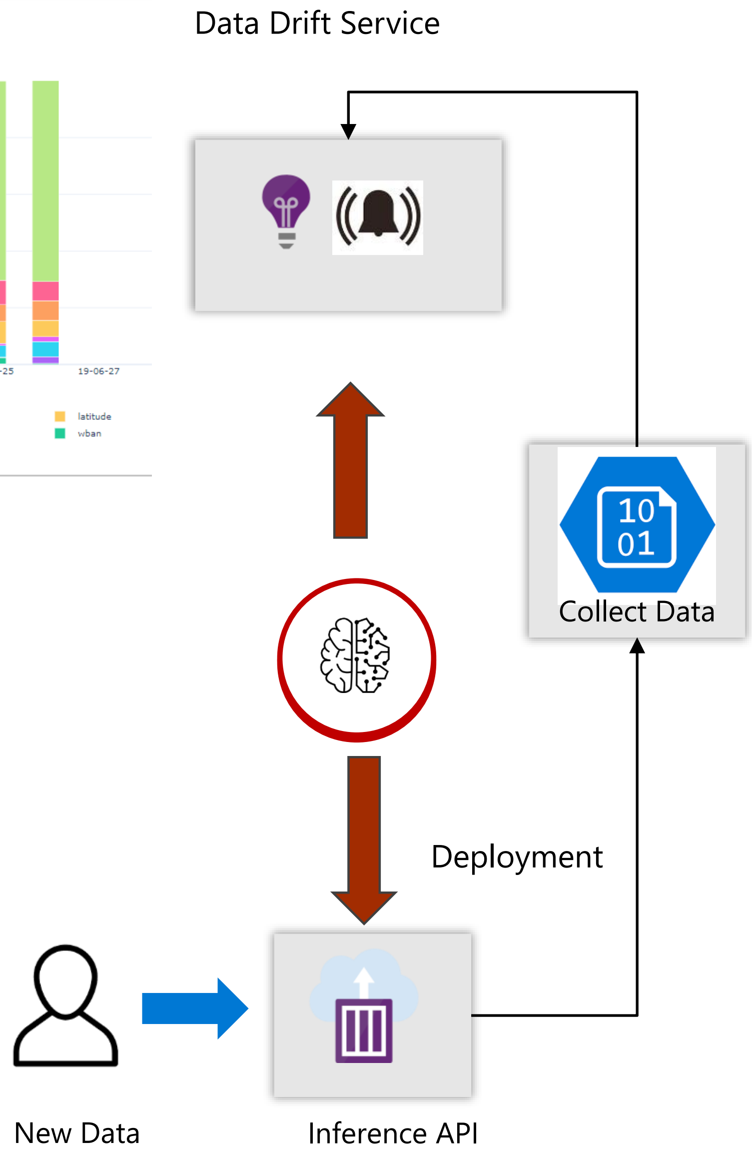
Interpret ML
https://interpret.ml/

Model deployment

During this step, we register a new version of our model and we prepare it for deployment, doing the following tasks:

- Deploy the last version of our model with his explanation model.
- Create a Docker image with all the logic to obtain predictions with the local explainability
- Deploy our Docker image in an ACI
- Start collecting data and logs in application insight.
- Activate Data Drift service to monitor the inputs of the deployed model and compare this data to the training dataset in order to measure the magnitude of data drift.



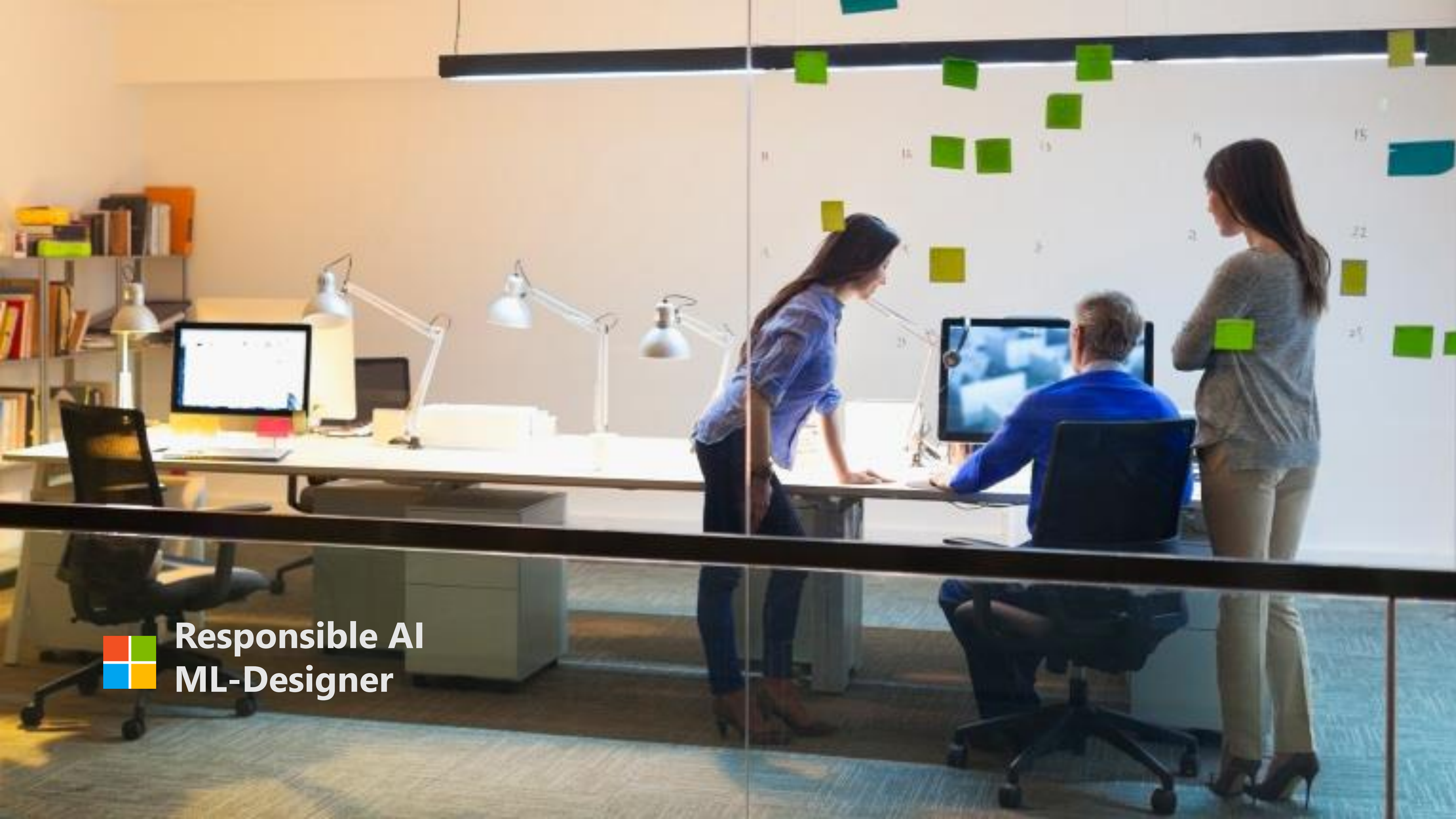


Demo

Lee Stott

Q&A

Lee Stott

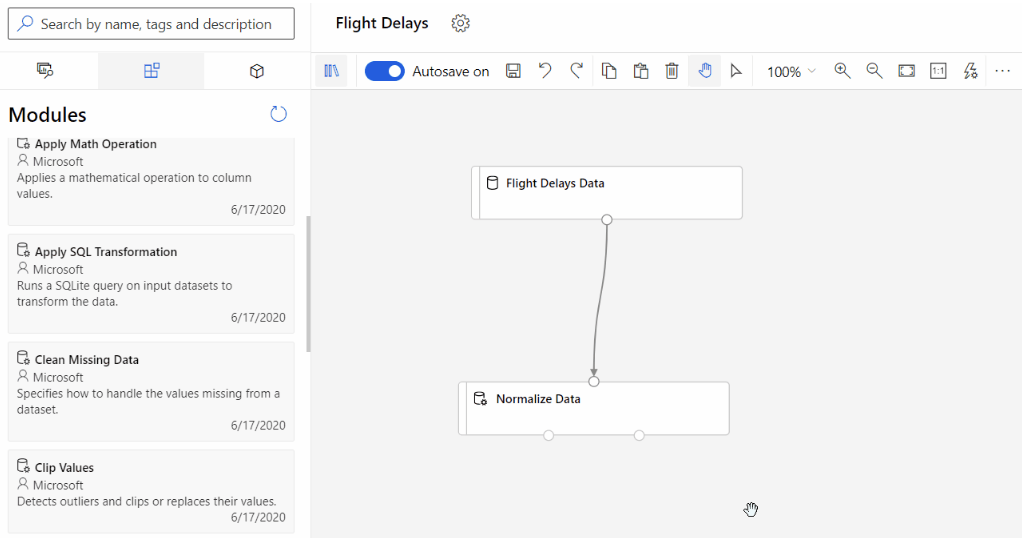


Responsible AI
ML-Designer

ML Designer

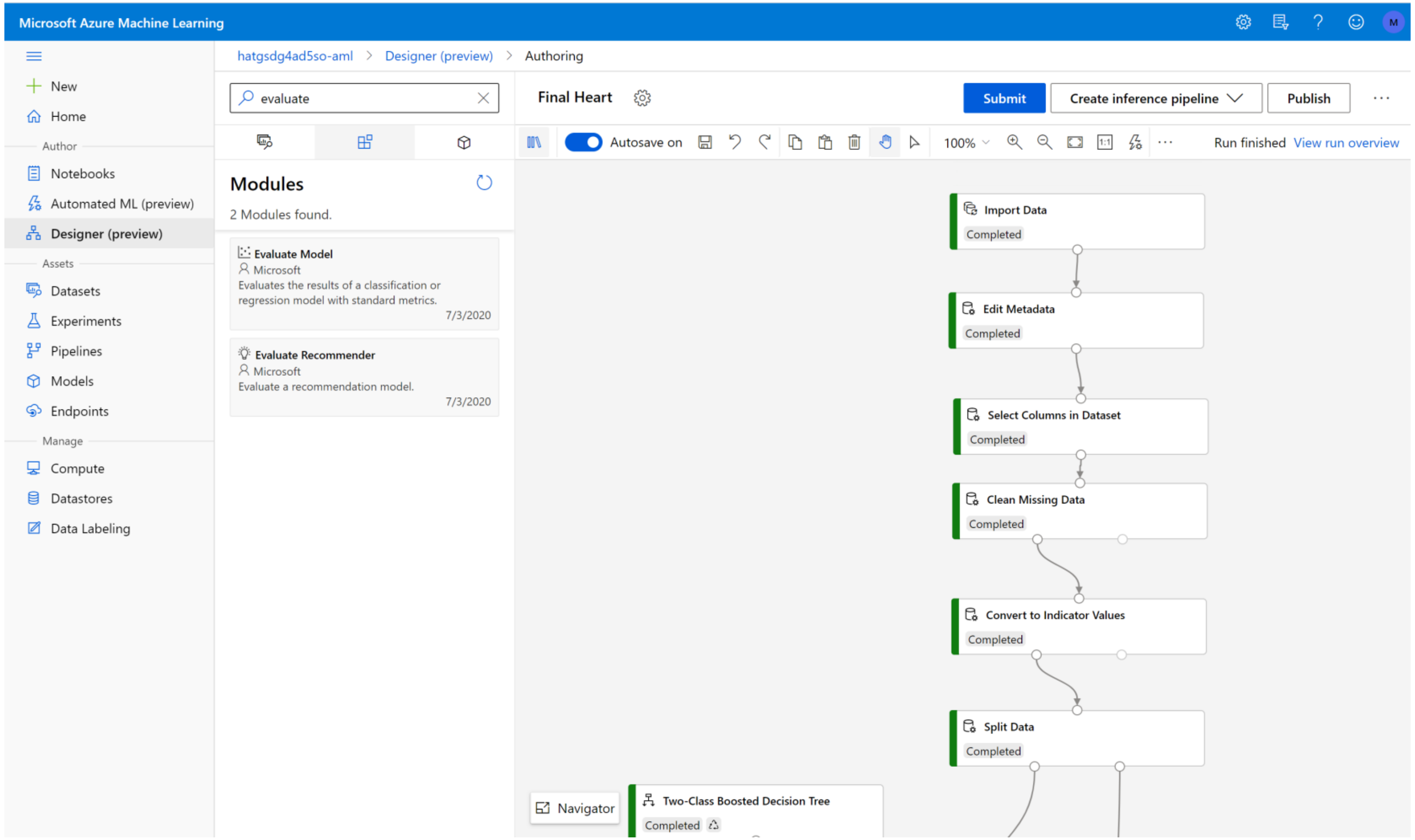
Azure ML Designer lets you visually conect datasets and modules on an interactive canvas to define machine learning workflows. With the designer we can:

- Drag-and-drop datasets and modules onto the canvas
- Connect the modules to create an azure machine learning pipeline using the visual editor
- Submit a pipeline run using the compute resources in your Azure Machine Learning workspace.
- Deploy a real-time inference pipeline to a real-time endpoint to make predictions on new data in real-time.



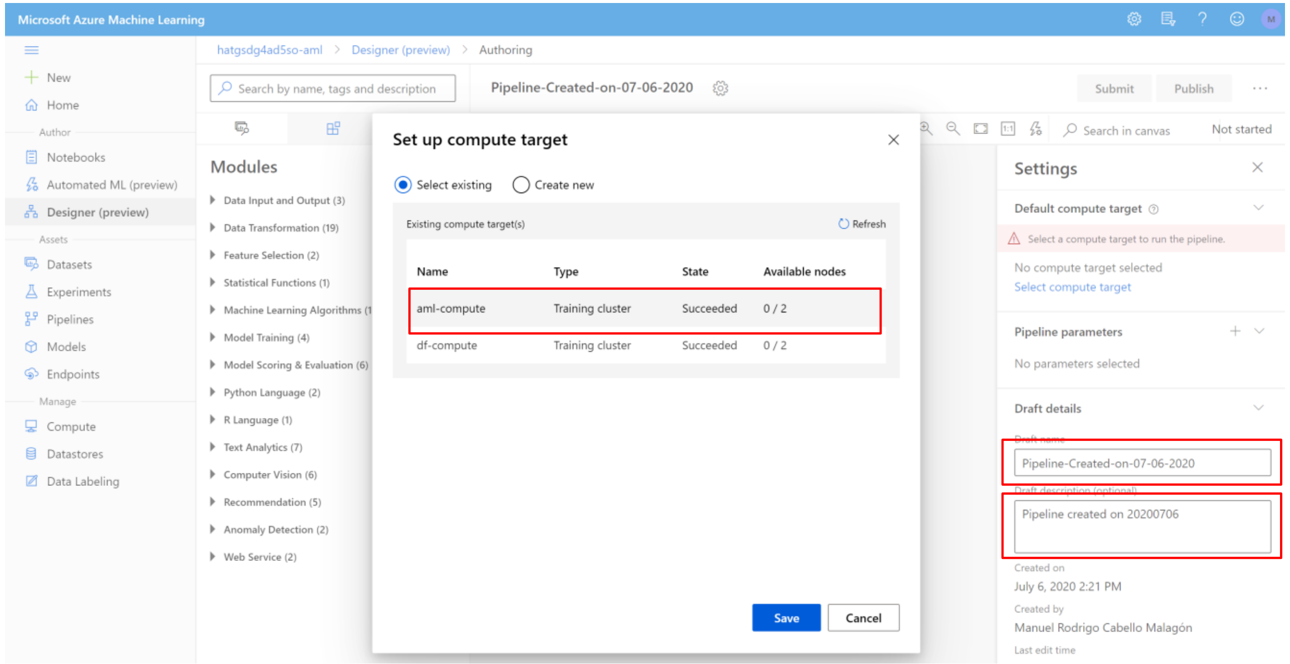
ML Designer

- Before training our model, we need to prepare our dataset, making some tranformations



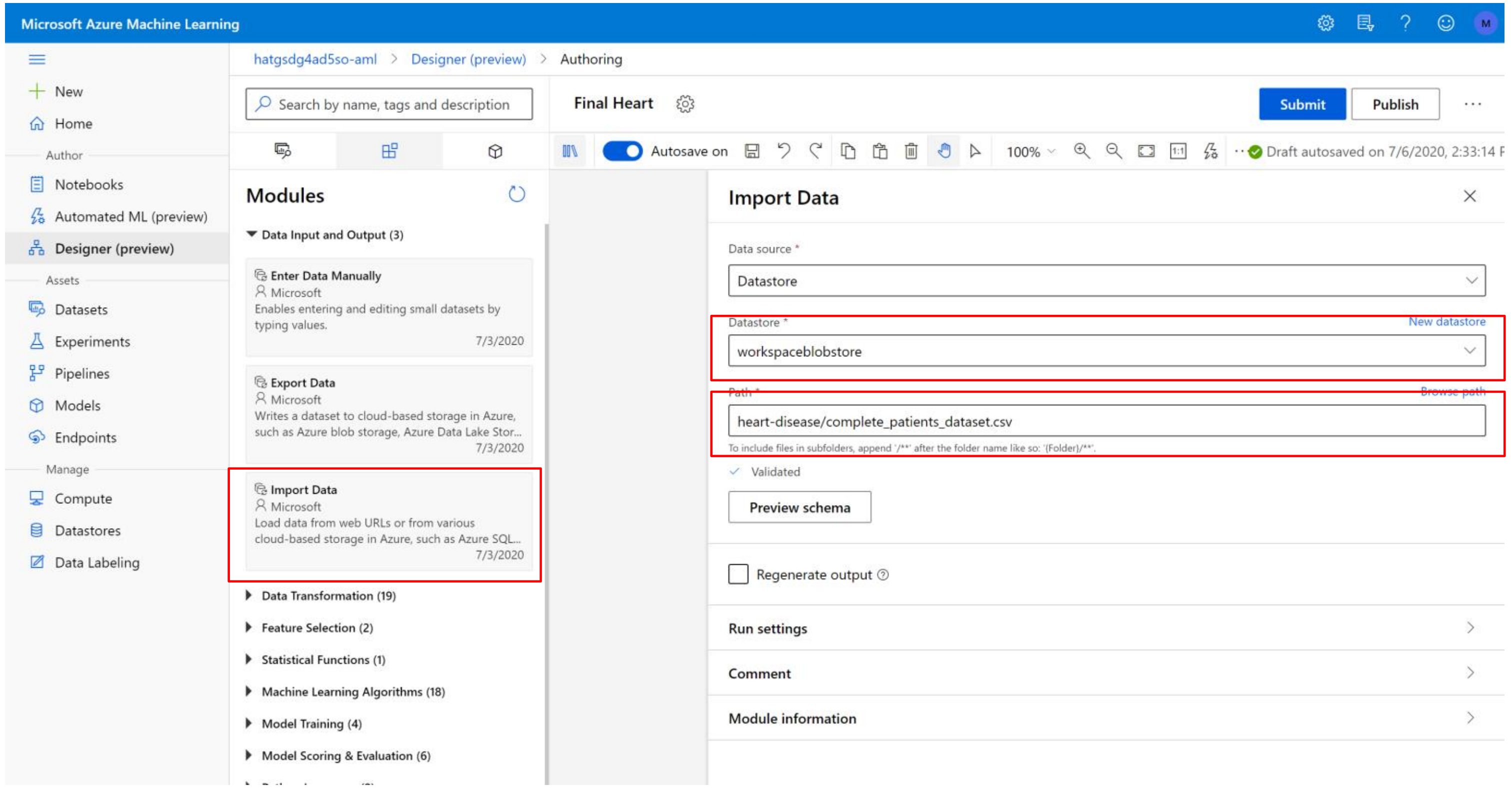
ML Designer

- In this case we will use the ML-Designer to build an end2end machine learning workflow to predict if a patient will be receive a treatment for heart disease or not.
- First of all we need to set up a compute target to execute our pipeline.
- (optional) Introduce the pipeline name and description.



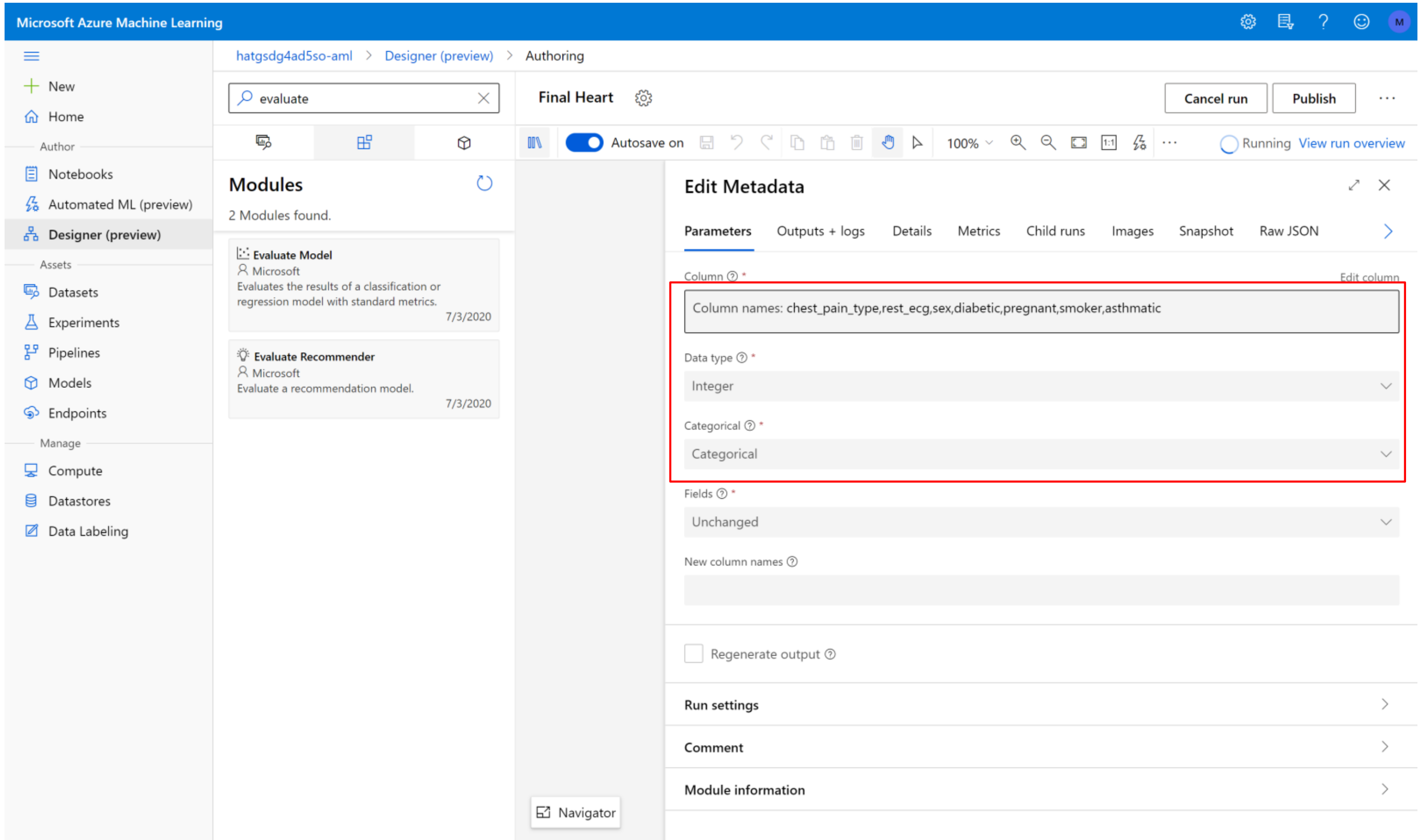
ML Designer

- Add an import module data and select the Dataset “**complete_patients_dataset**” from the Datastore



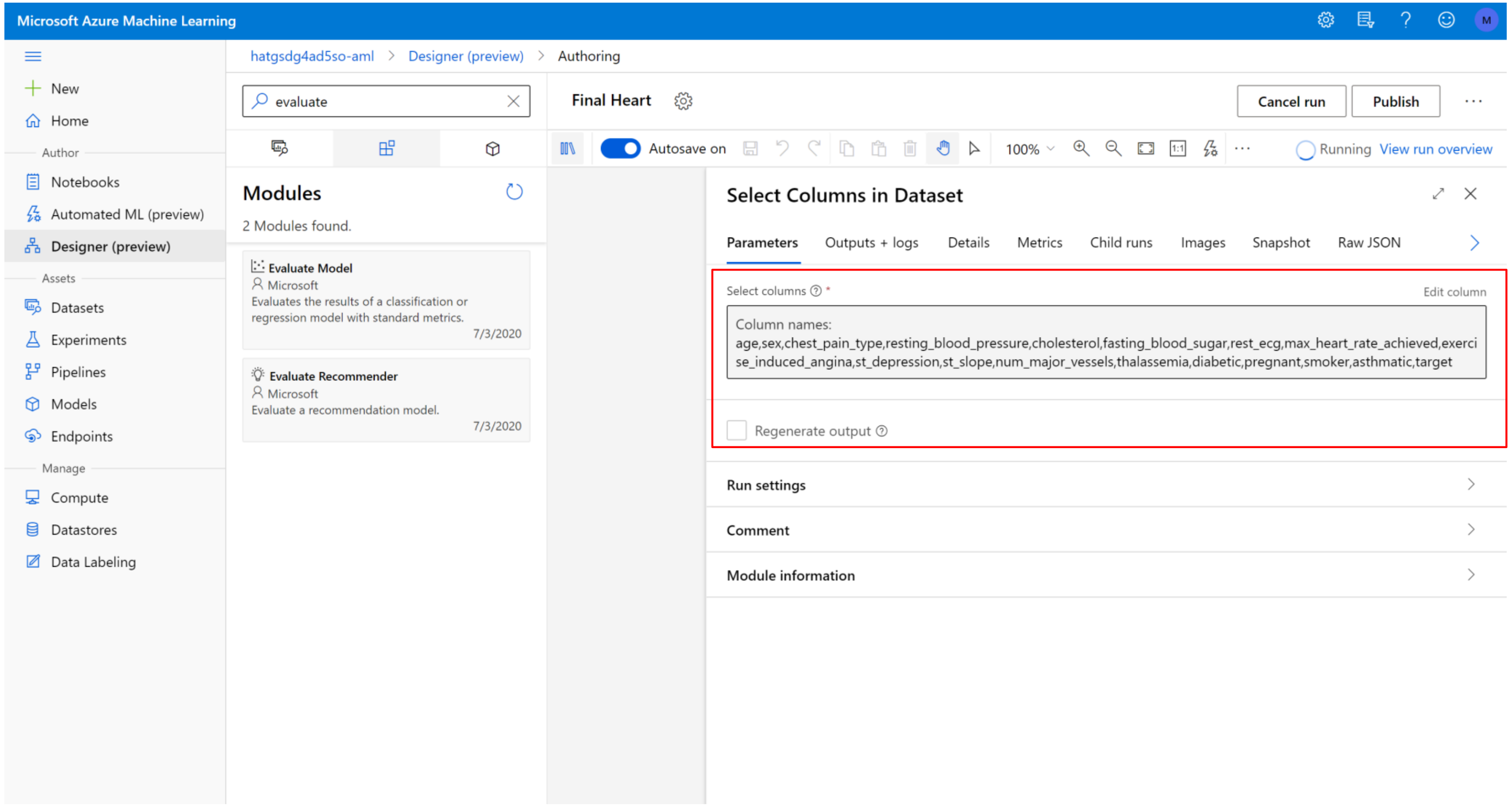
ML Designer

- Convert some integer columns to categorical values using the Edit Metadata module



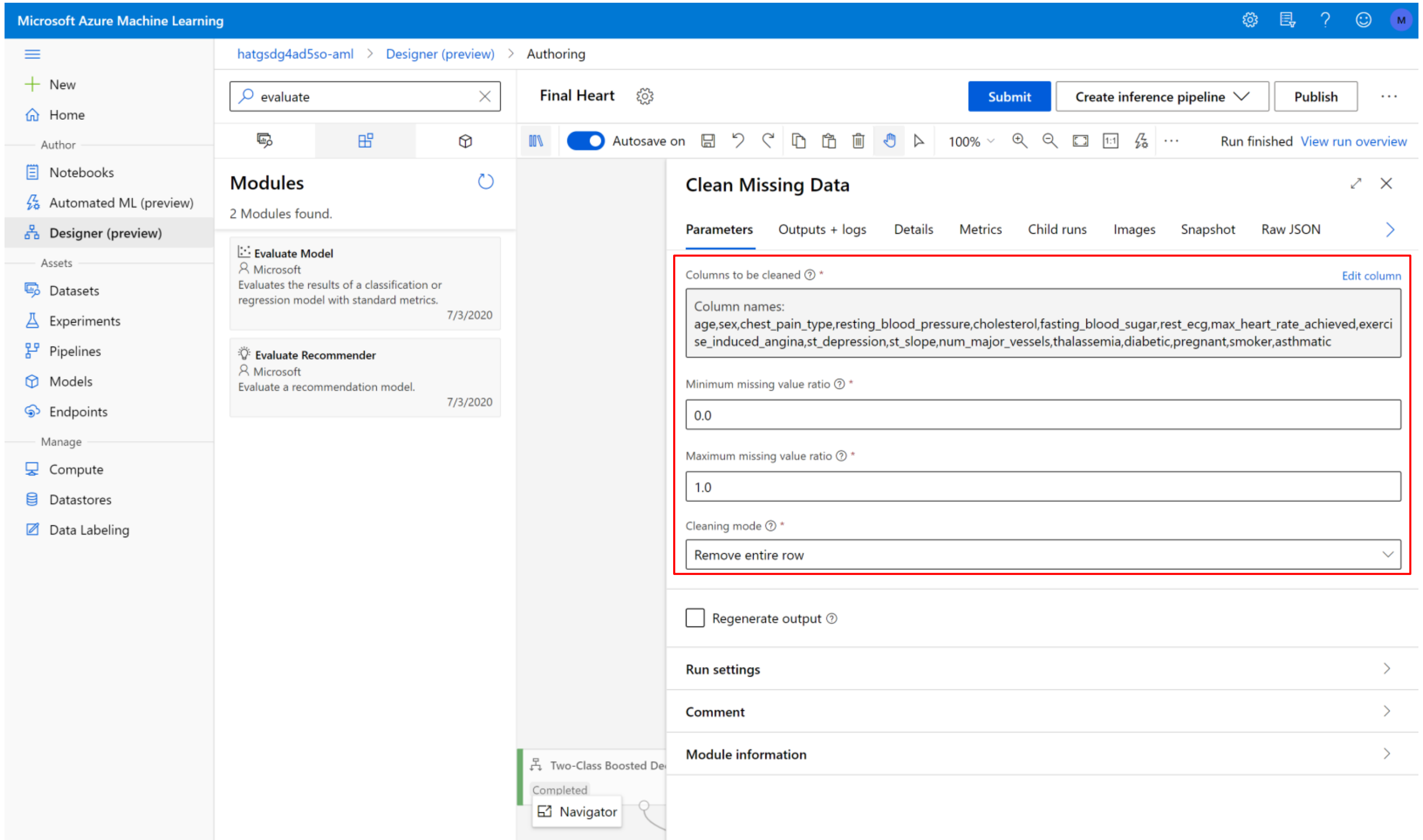
ML Designer

- Select the final columns that we will consider in our dataset to train our model. We will use the **Select Columns in Dataset module**



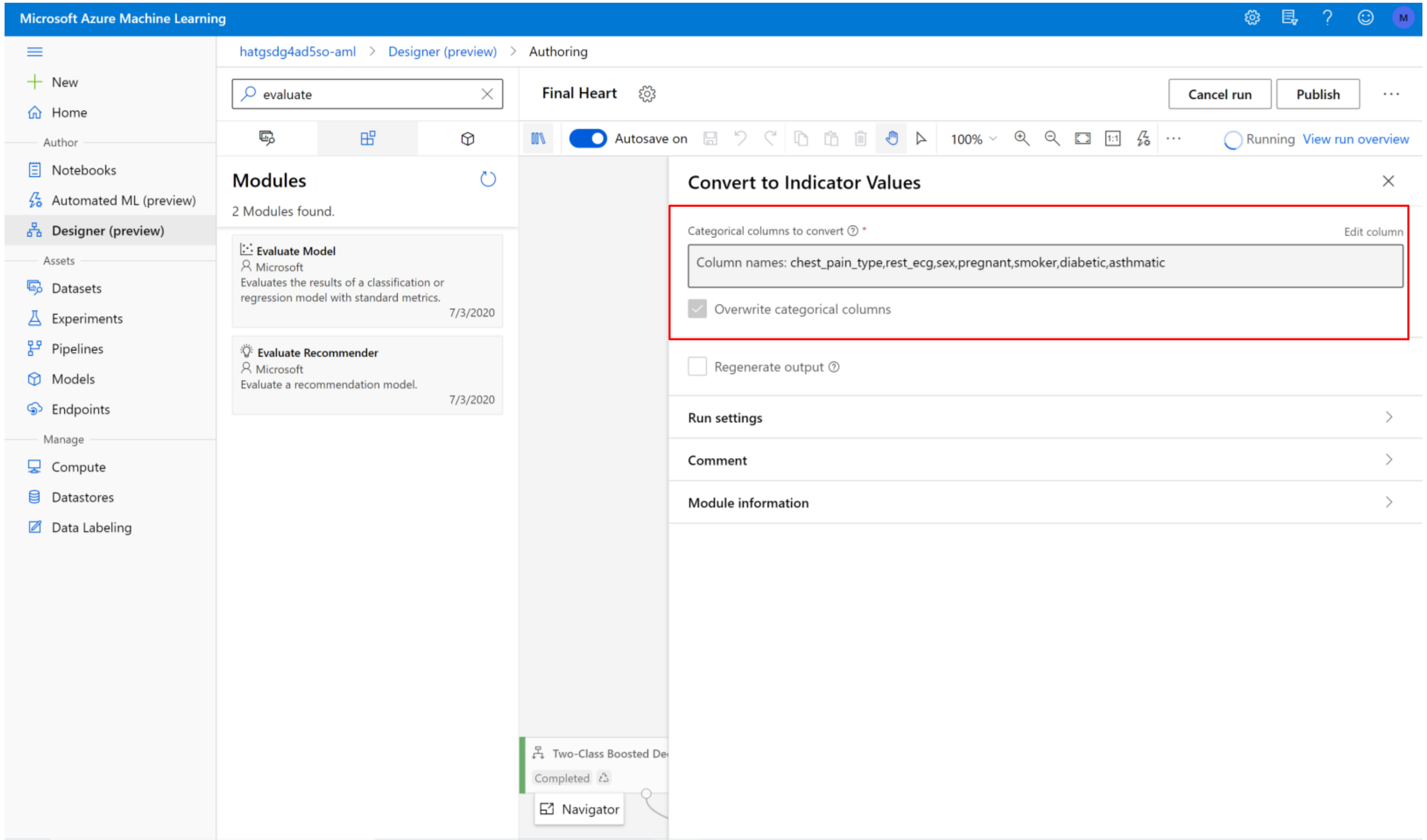
ML Designer

- In this step we clean missing values from dataset using the **Clean Missing Data module**. In this case, we will remove the entire row if some value of the columns is missed.



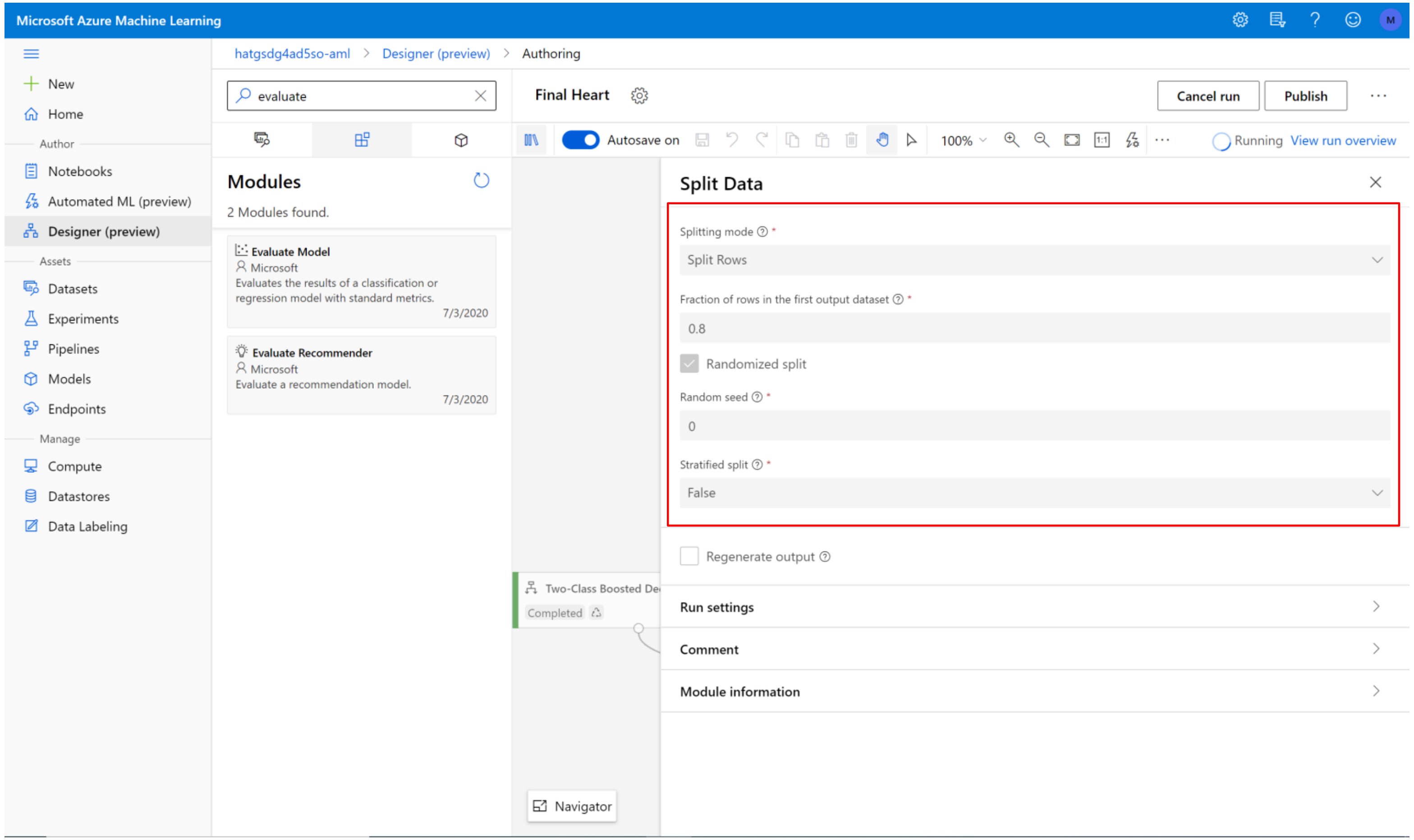
ML Designer

- We transform our categorical columns using one hot encoding, in this case, we will use **the convert to indicator module**



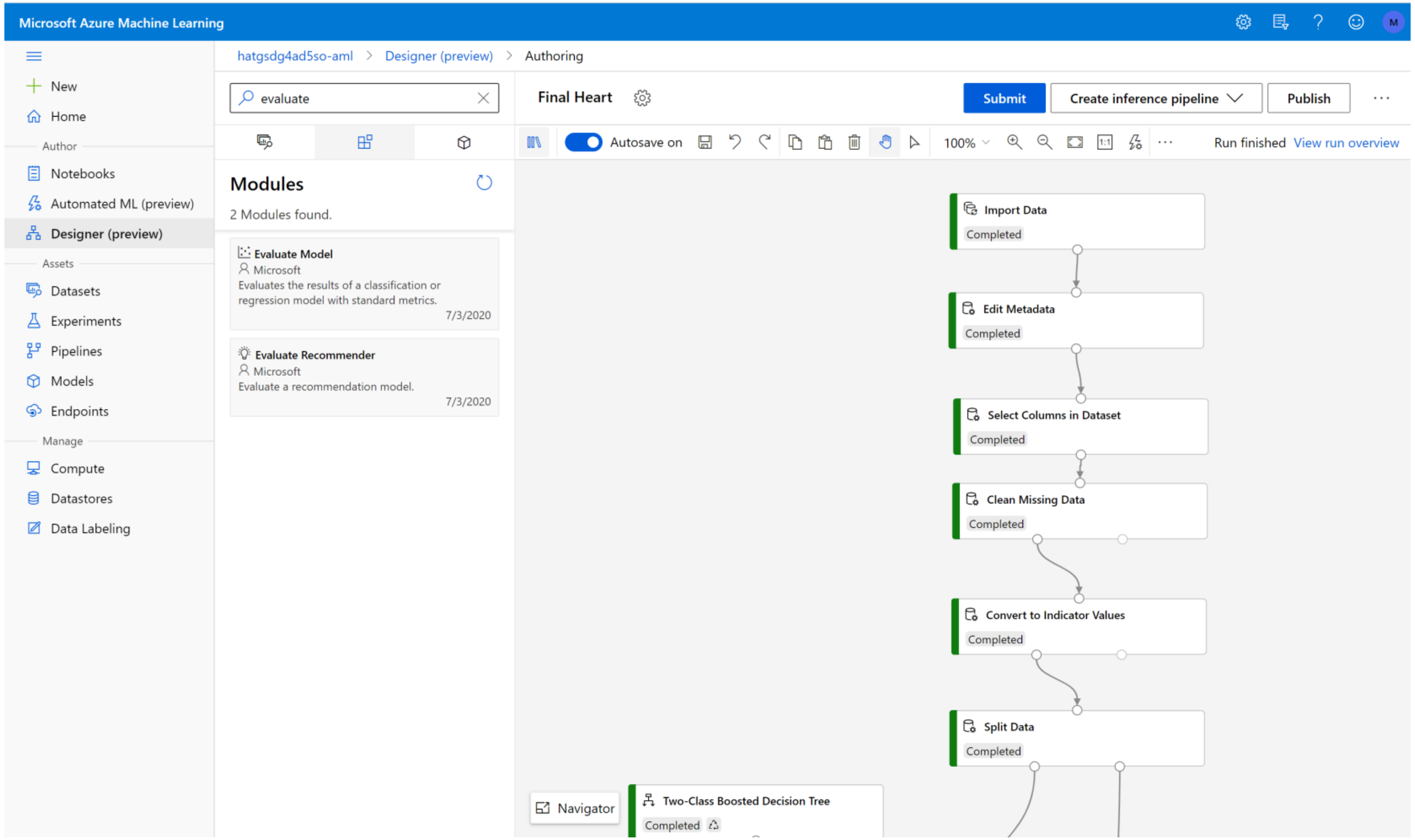
ML Designer

- We split our dataset into training and test, taking 80% for training and 20% for testing,



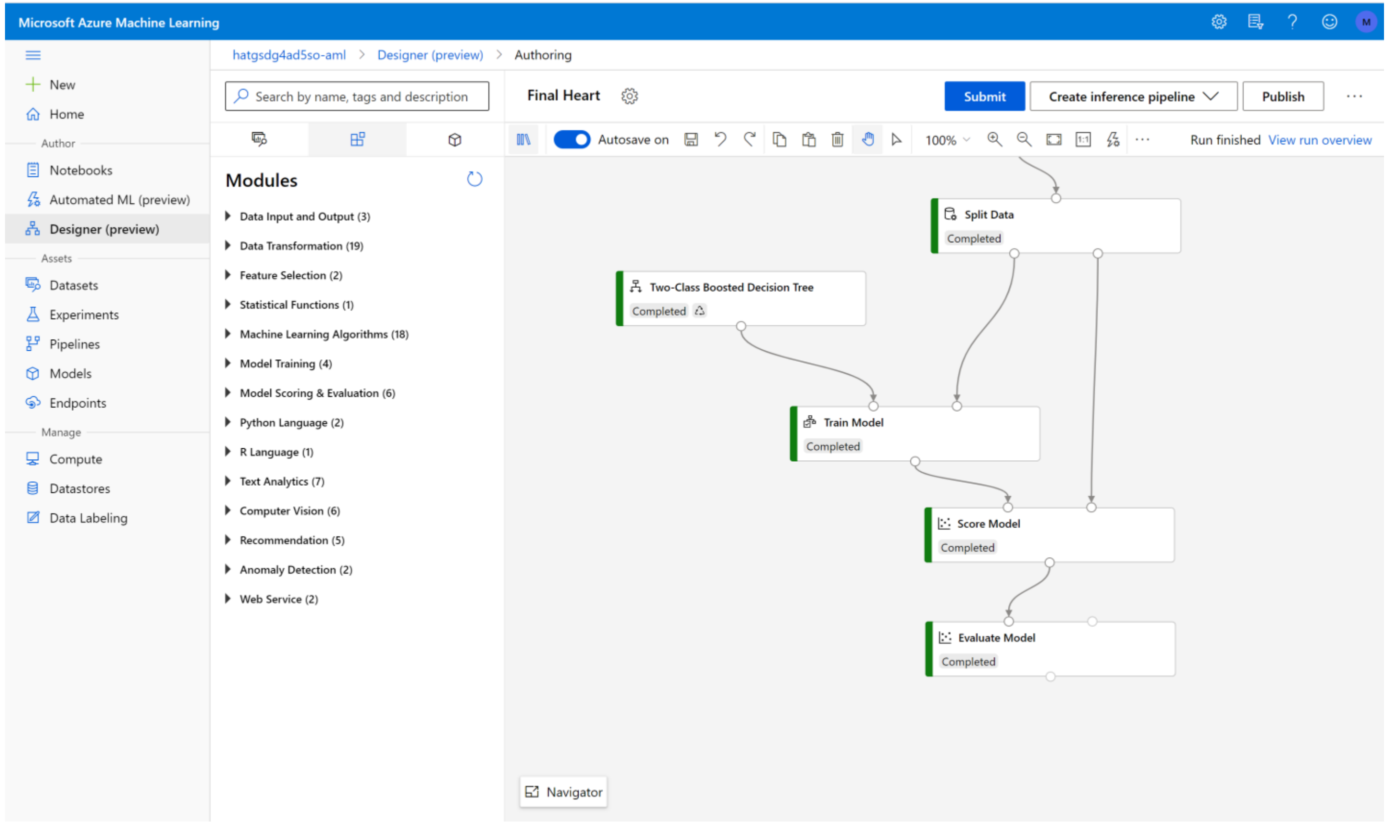
ML Designer

- Before training our model, we need to prepare our dataset, making some tranformations



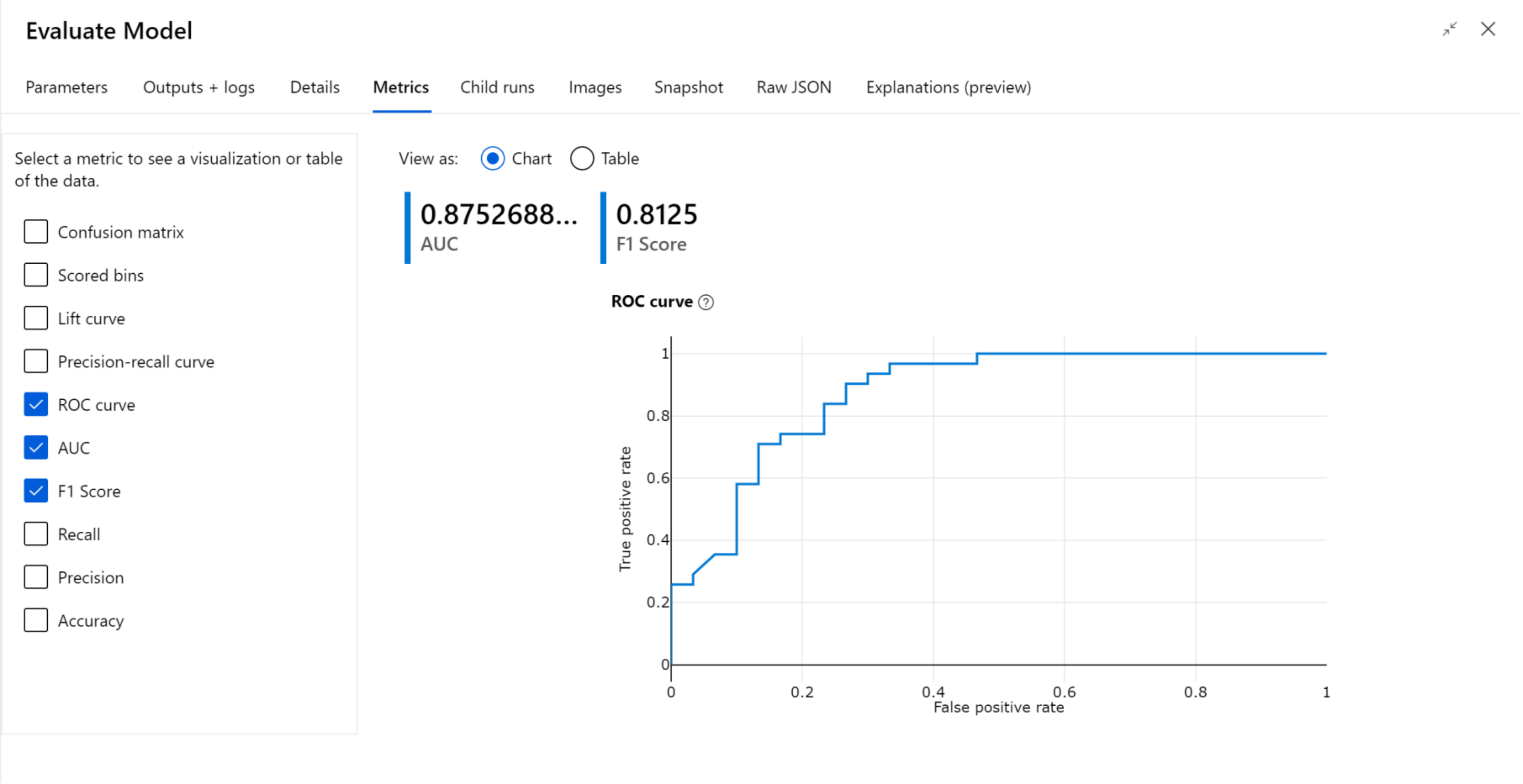
ML Designer

- Once we have splitted our dataset into training and testing, we must select a Machine Learning algorithm to start the training process
- In this case we have selected a classification algorithm: "**Two-class boosted decisión tree"** because we want to predict if a patient is going to receive a treatment or not.
- In the train module we must to select the "target" column
- In the score model module we make predictions over our trained model using the testing dataset.



ML Designer

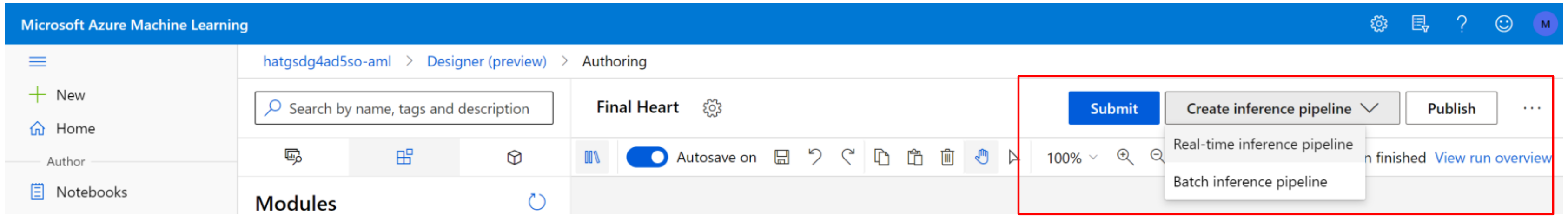
- After finish our ML workflow, we add the evaluate module to see the performance of our model.
- In the metrics tab appears the most common metrics of a classification problem: AUC, F1 Score, Precision, recal...

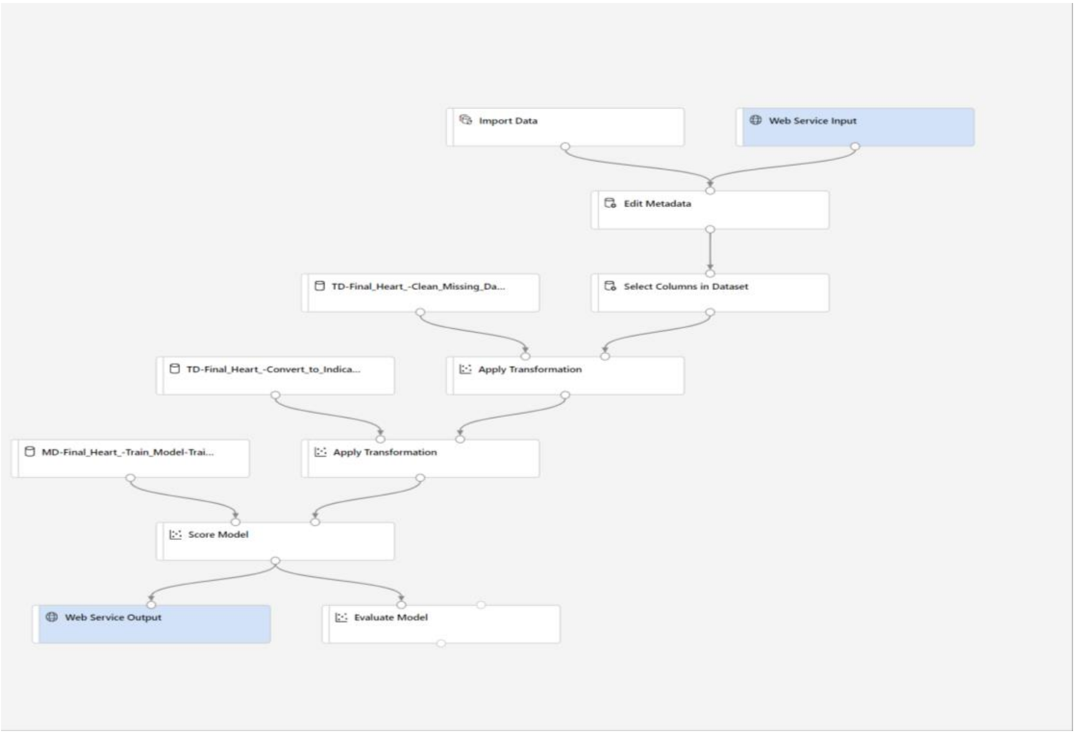


ML Designer (Inference)

- Once we hace finished our training pipeline, we can create a real-time inference pipeline

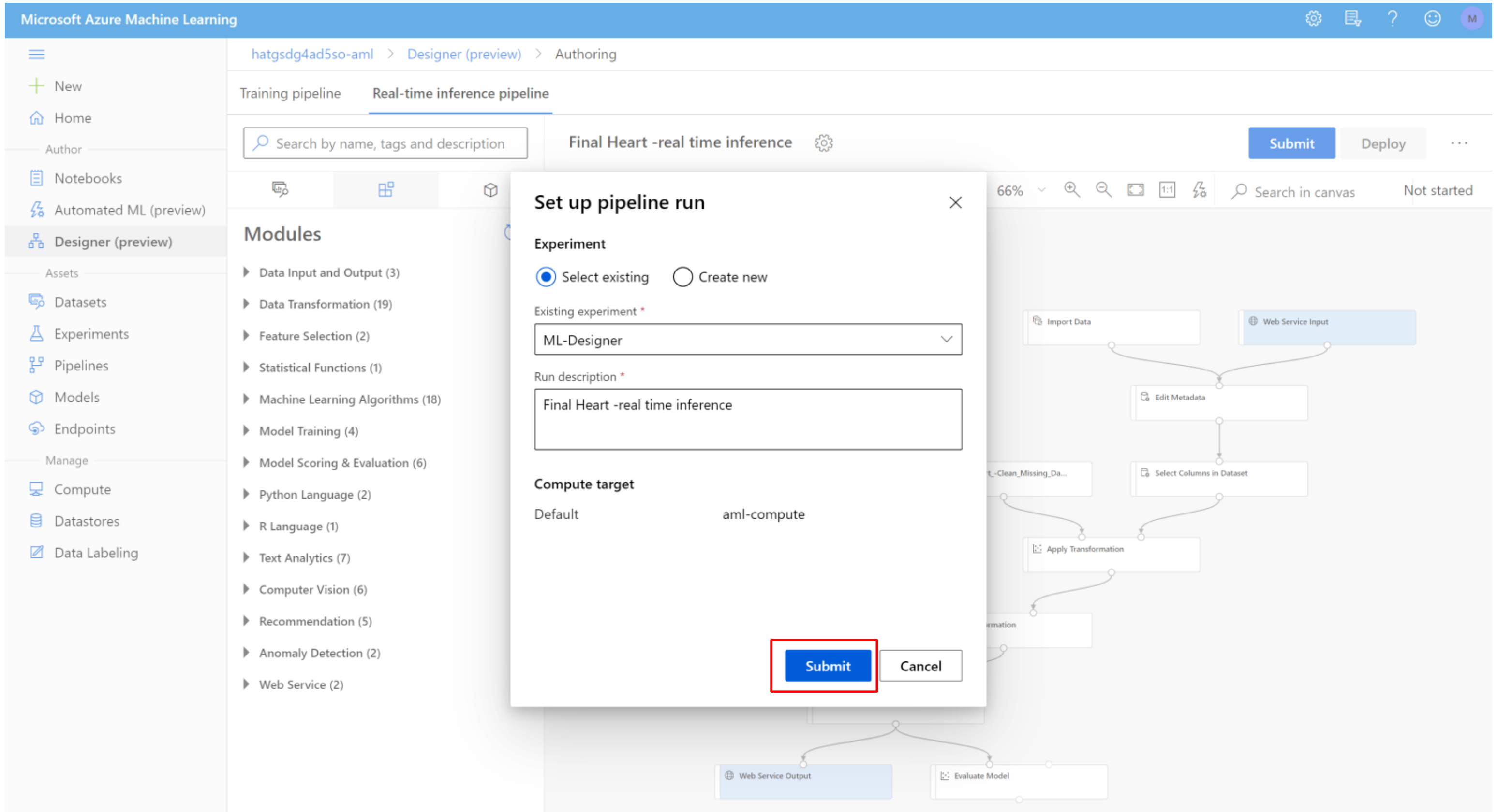
- Azure Machine Learning Studio mades some transformation to convert our train pipeline into an inference pipeline





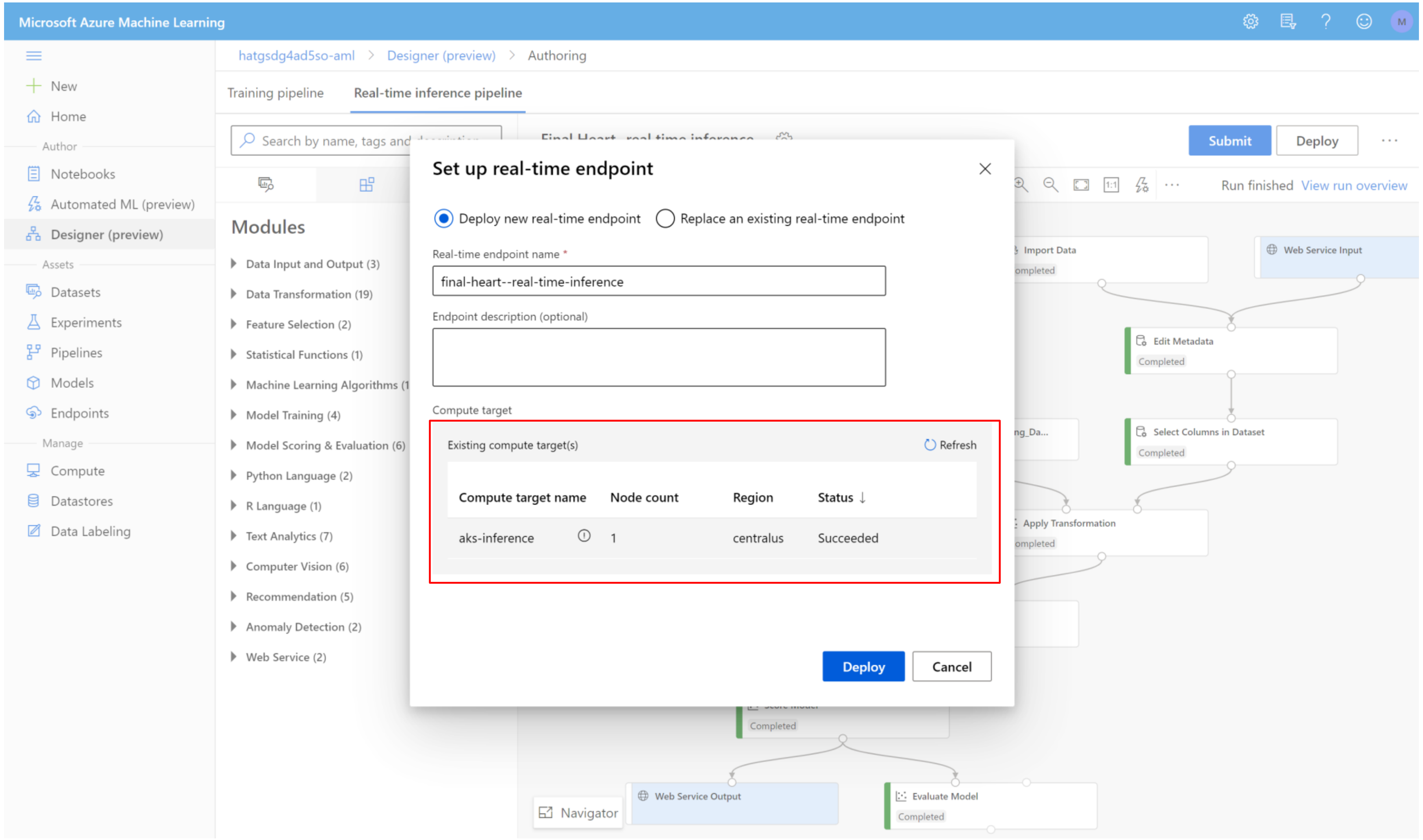
ML Designer (Inference)

- Before deploying our real-time service we need to submit the inference pipeline.



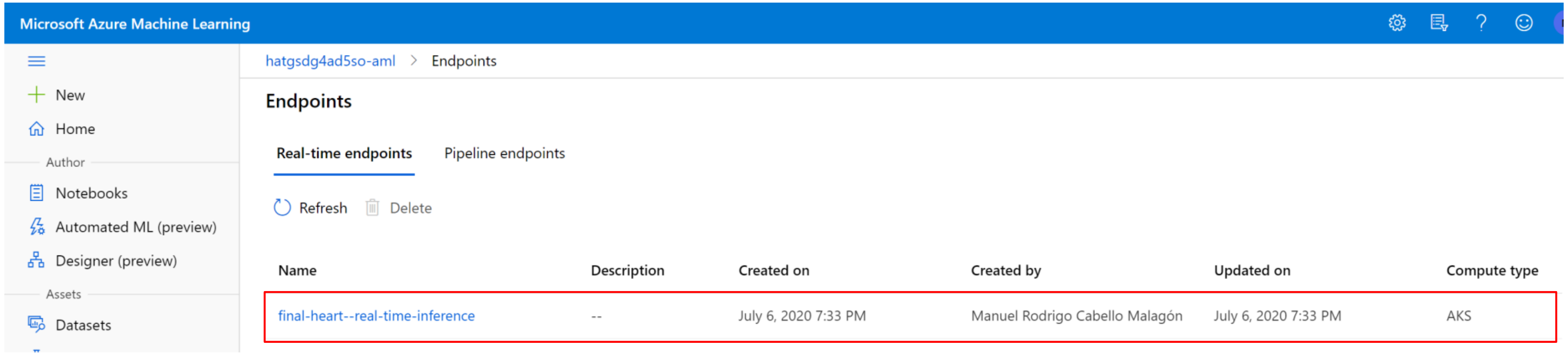
ML Designer (Deployment)

- We can convert our inference pipeline into a real-time endpoint.
- Before deploying our service we must create a inference clustering (AKS).
- Once we have created our inference pipeline, we can select it to deploy the web service.



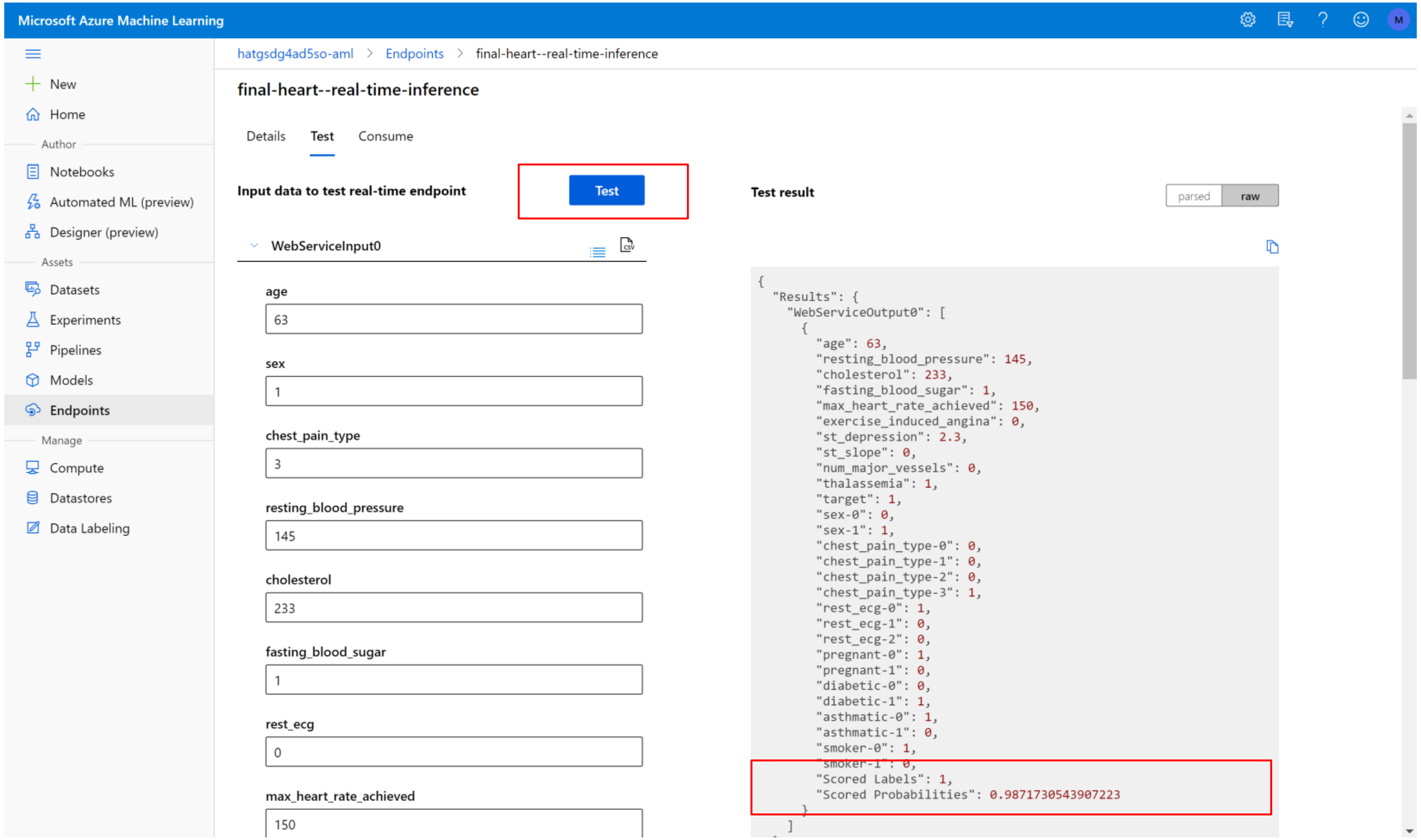
ML Designer (Deployment)

- In the endpoint sections, our new service will appear, click on it to get information about how to consume the web service.



ML Designer (Deployment)

- Into the details endpoint we can obtain information about how to consume our module through API.
- In the test tab, we can test our model introducing the values for inference. In the test results sections the predicted value appears with the scored labels and the scored probabilites



Useful Web Links & Resources

Responsible ML
http://aka.ms/responsibleML

Azure Machine Learning
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-monitor-datasets

Videos Demos
AI Show
https://www.youtube.com/watch?time_continue=240&v=Ts6tB2p97ek&feature=emb_logo
Dev Relations Webinar
https://www.youtube.com/watch?time_continue=9&v=vR7N4aUXmaQ&feature=emb_logo
InterpretML AI Show (video)
Fairlearn AI Show (video)

Fairlearn
http://Fairlean.org
http://github.com/fairlearn
https://www.microsoft.com/en-us/research/uploads/prod/2020/05/Fairlearn_whitepaper.pdf
•Fairlearn open source
•Fairlearn whitepaper
•Fairness Assessment in AzureML, concept
•Fairness Assessment in AzureML, how-to
•Fairlearn case study

Data Drift
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-monitor-datasets

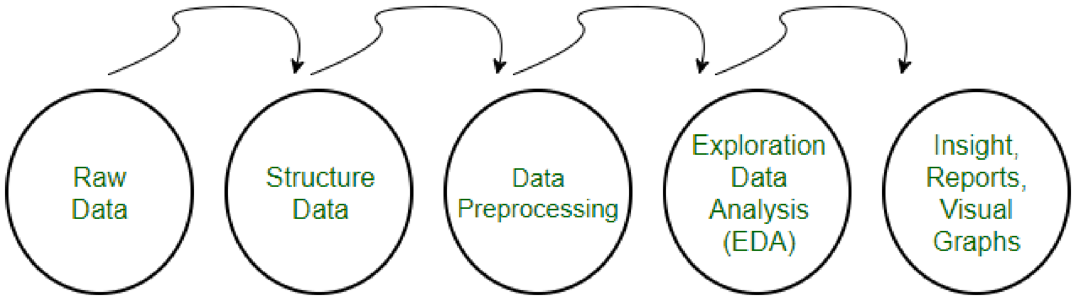
Github Repo
Health demo
https://github.com/leestott/ResponsibleAI
Responsible AI Airflit
https://github.com/microsoft/ResponsibleAI-Airlift

Data Annonymization
https://github.com/Microsoft/presidio
Presidio Features
Presidio Input and Output
API Spec
The Technology Stack
Architecture

Interpret ML https://interpret.ml/
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability-aml
http://interpret.ml
https://github.com/interpretml/interpret
•InterpretML open source
•Interpretability in AzureML, concept
•Interpretability in AzureML, how-to
•InterpretML case study

Exploratory analysis and preprocessing

- This step launches the following tasks:
    - Perform initial investigations on data to discover patterns, spot anomalies and check assumptions with the help o summary statistics and graphical representations.
    - Preprocess all the dataset:
        - Handle null values
        - Encode categorical and text values.
        - Normalize data.
    - Feature engineering



Explainable AI

Interpret ML

https://interpret.ml/