1. Main page: http://cortanaanalytics.com
2. Before you begin this Module, you should have:
   1. Basic Math and Stats skills
   2. Business and Domain Awareness
   3. Familiarity with Azure ML
   4. An understanding of interpreting a problem Statement for a Customer Scenario

Learning objectives

Using Azure ML:
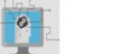
1. Split data into testing and training sets
2. Build a model for predicting customer churn
3. Compare two models
4. Share an experiment
5. Deploy a model as an API

When you have completed this Module, using Azure ML you will be able to:

1. Split data into testing and training sets

2. Build a model for predicting customer churn

3. Compare two models

4. Share an experiment

5. Deploy a model as an API

Cortana Analytics

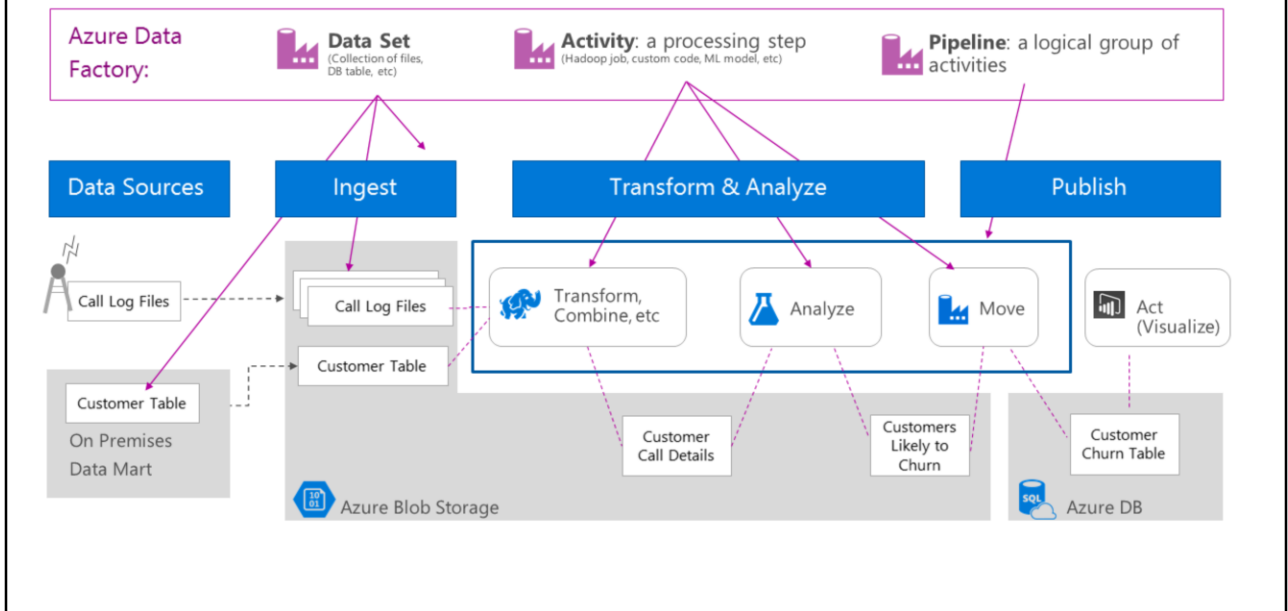|	
	Cortana
	Power BI
	Azure Stream Analytics
	Azure HDInsight
	Azure Machine Learning
	Azure SQL DB, Data Warehouse, DocumentDB
	Azure Data Lake
	Azure Event Hubs
	Azure Data Catalog
	Azure Data Factory
	Microsoft Azure

1. Platform: Microsoft Azure – http://microsoftazure.com
2. Storage: https://azure.microsoft.com/en-us/documentation/services/storage/ (Blobs)
3. Azure Data Lake: http://azure.microsoft.com/en-us/campaigns/data-lake/ (from Bing)
4. Azure Data Factory: http://azure.microsoft.com/en-us/services/data-factory/ (Orchestration and Movement)
5. Azure Event Hubs: http://azure.microsoft.com/en-us/services/event-hubs/ (Ingest to streaming (IoT))
6. Azure Data Catalog: http://azure.microsoft.com/en-us/services/data-catalog (Social networking for data)
7. Azure DocumentDB: https://azure.microsoft.com/en-us/services/documentdb/?WT.srch=1&WT.mc_ID=SEM_JQ3fO8dU
8. Azure SQL Data Warehouse: http://azure.microsoft.com/en-us/services/sql-data-warehouse/
9. Azure Stream Analytics: http://azure.microsoft.com/en-us/services/stream-analytics/ (Consume from EH, Output to Alerting (Telco Example))
10. Azure HDInsight: http://azure.microsoft.com/en-us/services/hdinsight/ (Platform on a Platform)
11. Azure Machine Learning: http://azure.microsoft.com/en-us/services/machine-learning/ (Prediction and Prescription - to API)
12. Revolution R Enterprise: http://www.revolutionanalytics.com/revolution-r-enterprise (ML)
13. Power BI: https://powerbi.microsoft.com/
14. Cortana: http://blogs.windows.com/buildingapps/2014/09/23/cortana-integration-and-speech-recognition-new-code-samples/ and https://blogs.windows.com/buildingapps/2015/08/25/using-cortana-to-interact-with-your-customers-10-by-10/
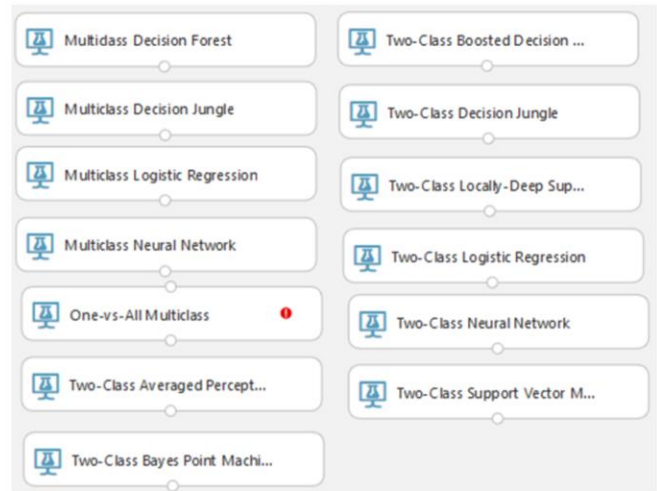
1. Full explanation of this example: https://azure.microsoft.com/en-us/blog/getting-started-with-azure-data-factory-and-azure-machine-learning-4/

1. Head link for this topic – note that it has a section for choosing the right algorithm: https://msdn.microsoft.com/en-us/library/azure/dn905812.aspx
2. Example using flight times: http://blogs.msdn.com/b/cdndevs/archive/2016/02/18/step-by-step-how-to-predict-the-future-with-machine-learning.aspx

Azure Machine Learning
Initialize Model – Regression

- Algorithms that learn to predict the value of a real function for a single instance of data

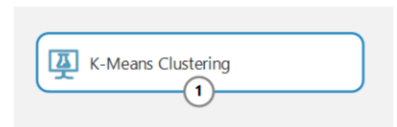|	
Bayesian Linear Regression	Neural Network Regression
Boosted Decision Tree Regre...	Ordinal Regression
Decision Forest Regression	Poisson Regression
Fast Forest Quantile Regress...	Linear Regression

1. Creating a model with a SQL Server Database:
   http://blogs.msdn.com/b/business_intelligence1/archive/2014/09/14/how-to-setup-machine-learning-and-run-a-linear-regression-model-on-the-azure-cloud.aspx

Azure Machine Learning
Initialize Model – Clustering

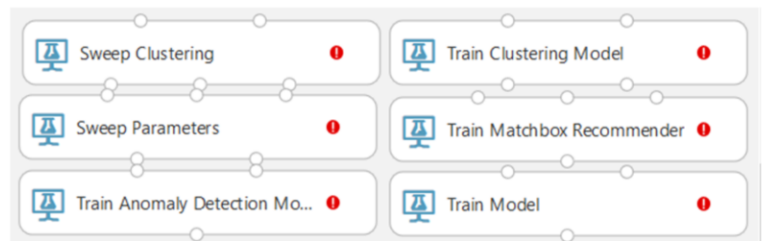- Create an untrained K-means clustering model

K-Means Clustering

1

---

1. A good comparison of clustering versus classification:
   http://www.scientificcomputing.com/blogs/2015/09/machine-learning-techniques-help-find-patterns-big-data-sets

Azure Machine Learning
Train

Variety of methods for training models, which can also incorporate parameter sweeping and cross-validation

⊞ Sweep Clustering ❶	⊞ Train Clustering Model ❶
⊞ Sweep Parameters ❶	⊞ Train Matchbox Recommender ❶
⊞ Train Anomaly Detection Mo... ❶	⊞ Train Model ❶

1. https://msdn.microsoft.com/library/azure/038d91b6-c2f2-42a1-9215-1f2c20ed1b40
2. More on tuning parameters (aka hyperparameter optimization):
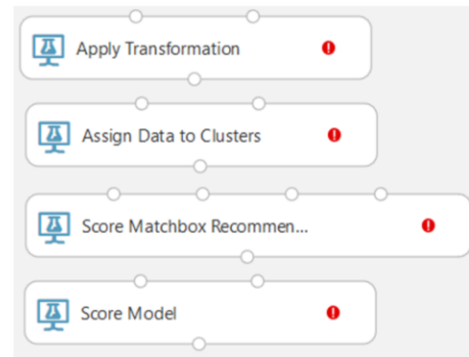   http://www.johnmyleswhite.com/notebook/2012/07/21/automatic-hyperparameter-tuning-methods/

Azure Machine Learning
Score

Apply a trained model to:

- A list of recommended items
- Forecasts for time series models
- Estimates of projected demand, volume, or other numeric quantity, for regression models
- Cluster assignments
- A predicted class or outcome, for classification models
- Probability scores associated with these outputs

Apply Transformation		❶
Assign Data to Clusters		❶
Score Matchbox Recommen...		❶
Score Model		❶

1. An example of training and scoring a model: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-walkthrough-4-train-and-evaluate-models/
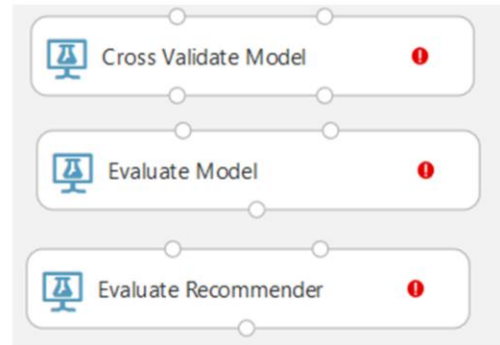
Azure Machine Learning
Evaluate
Metrics for Classification Models
- Accuracy
- Recall
- Precision
- F-Score
- AUC
- Average Log Loss
- Training Log Loss

Metrics for Regression Models
- Mean absolute error (MAE)
- Root mean squared error (RMSE)
- Relative absolute error (RAE)
- Relative squared error (RSE)
- Coefficient of determination

Cross Validate Model

Evaluate Model

Evaluate Recommender

1. A discussion on using AUC eefectively: http://academic.research.microsoft.com/Paper/428291.aspx

Azure Machine Learning
Deploy

- Access through the web service API
- Manage through the Azure management portal
- Update if the model changes

1. A Video of the connection between the app and the API – we'll cover this again in the next module:
https://channel9.msdn.com/Blogs/Windows-Azure/Introduction-to-Azure-ML-API-Service

1. Follow the subsequent slides for the lab

Start Azure ML Studio

- Go to: http://studio.azureml.net/
- Click **Get started →**
- If you have an Azure account, sign in; otherwise click Guest Access
- In the bottom left corner, click **+ NEW**
- Click "Blank Experiment"

Our Dataset

- Telecom data for 4,667 fictitious customers
- Feature selection already complete, leaving us with:
  - Day_Mins: daytime call usage
  - Eve_Mins: evening call usage
  - Intl_Mins: international call usage
  - VMail_Message: Number of voicemails received
  - CustServ_Calls: number of customer service calls
  - Churn_: indicating whether the customer churned

1. https://aka.ms/AzureMLChurnTutorial

Read the Data into Azure ML

Data URL:
https://telecomchurndemo.blob.core.windows.net/churnfeatures/churnfeatures.csv

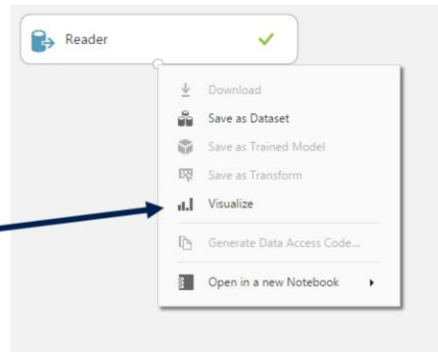1. Paste from here: https://telecomchurndemo.blob.core.windows.net/churnfeatures/churnfeatures.csv

Examine the Imported Data

Hover over the output port

Right-click to see options

Click "Visualize"

Examine the Imported Data

Split the Data

- 70% training (1st port)
- 30% testing (2nd port)
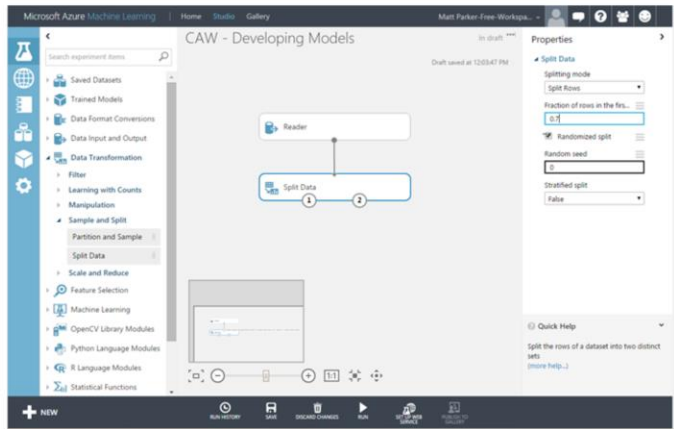
1. Splitting helps us avoid overfitting: http://www.kdnuggets.com/2015/01/clever-methods-overfitting-avoid.html
2. An alternative to the train-test split: Crossvalidation https://msdn.microsoft.com/en-us/library/azure/dn905852.aspx

Select the Model

1. Search for "Logistic" in the top left box
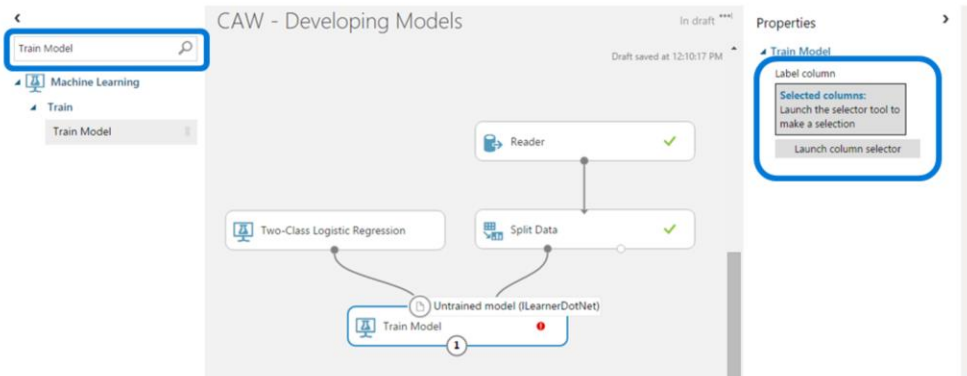2. Drag "Two-Class Logistic Regression" into the experiment

1. How to choose a model: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/
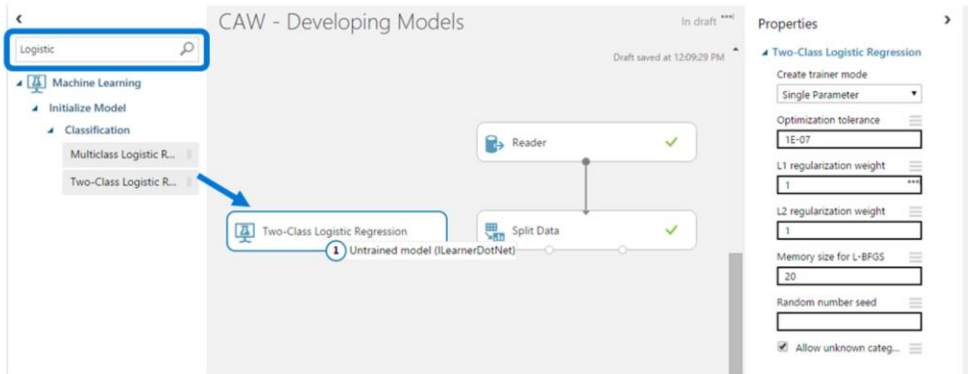
Train the Model

1. Search for "Train Model" and drag into experiment
2. Link *untrained* model to left port, *training* data to right port

Score the Model (aka Prediction)

1. Search for "Score Model" and drag into experiment
2. Link *trained* model to left port, *testing* data to right port
3. Output: the testing dataset, *with predictions*

CAW - Developing Models

Reader ✓

Two-Class Logistic Regression ✓   Split Data ✓

Train Model ✓

Score Model

... but are they good predictions?
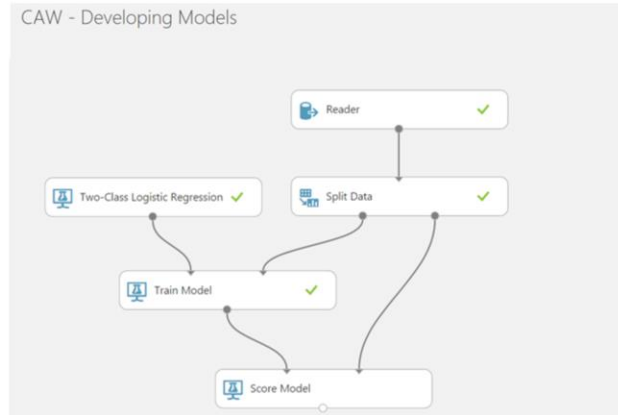
Evaluate the Model

1. Search for "Evaluate Model" and link to "Score"
2. Run experiment
3. Click the "Evaluate" out port and "Visualize"

1. Overview of evaluation statistics: https://msdn.microsoft.com/library/azure/927d65ac-3b50-4694-9903-20f6c1672089/#bkmk_classification
2. Compared with Crossvalidation: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-evaluate-model-performance/#evaluating-a-binary-classification-model
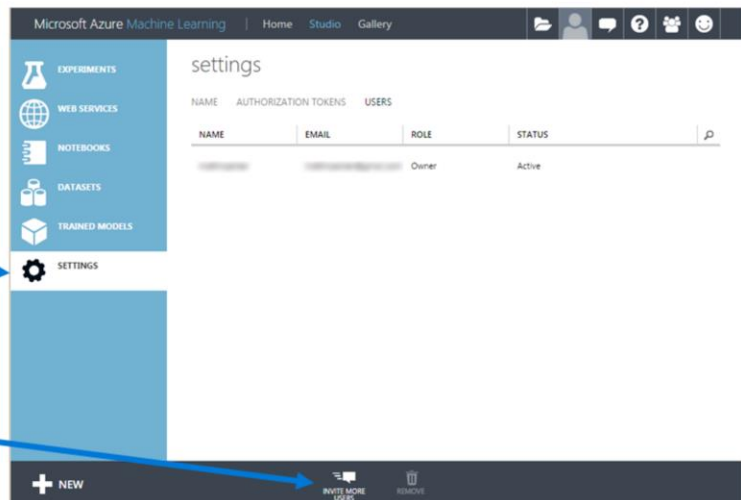
Sharing Experiments

Microsoft Azure

Share Your Experiment

Specifically, share your *workspace:*

1. Click the "Settings" gear on the left

2. Click "Invite More Users" at the bottom

---

1. More on managing your workspace: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-manage-workspace/

Share Your Experiment

1. Enter emails of colleagues

2. Set permissions

1. More on managing your workspace: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-manage-workspace/
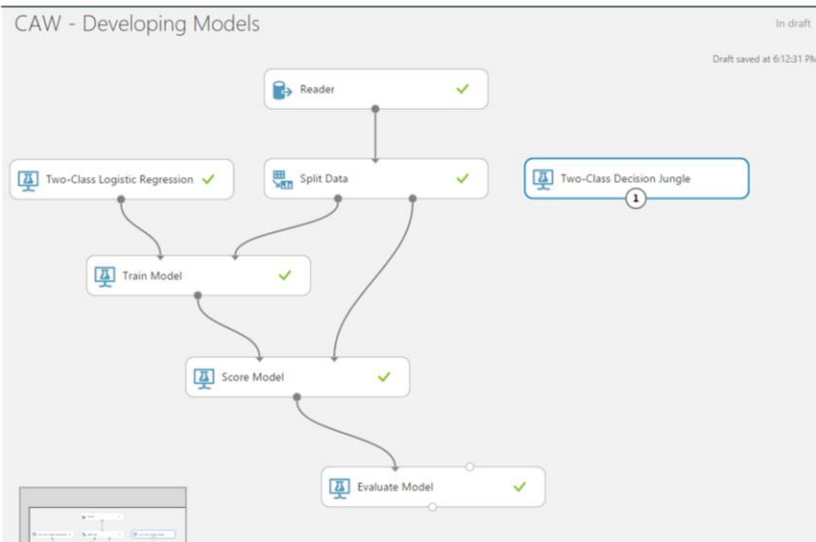
Add a Fancy Model

CAW - Developing Models

In draft

Draft saved at 6:12:31 PM

- Reader ✓
- Two-Class Logistic Regression ✓
- Split Data ✓
- Train Model ✓
- Score Model ✓
- Evaluate Model ✓

1. How to choose a model: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/

Add a Fancy Model

CAW - Developing Models

In draft

Draft saved at 6:12:31 PM

- Reader ✓
- Two-Class Logistic Regression ✓
- Split Data ✓
- Two-Class Decision Jungle ①
- Train Model ✓
- Score Model ✓
- Evaluate Model ✓
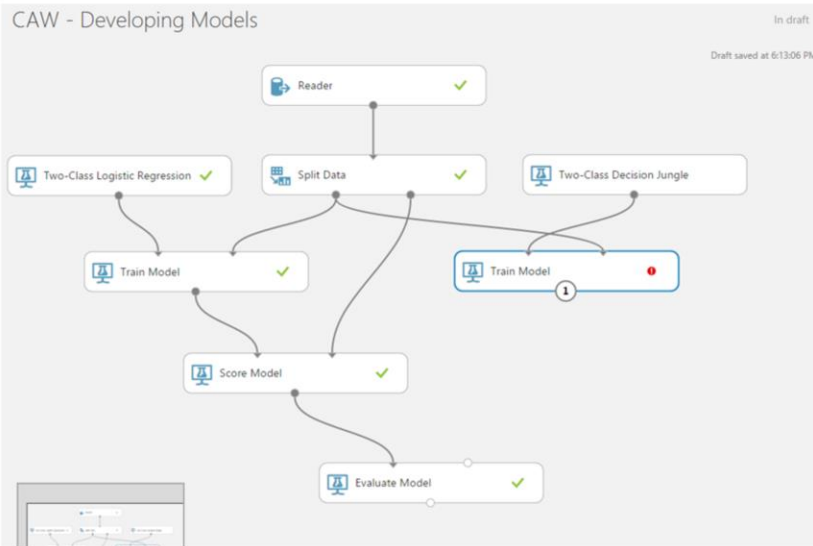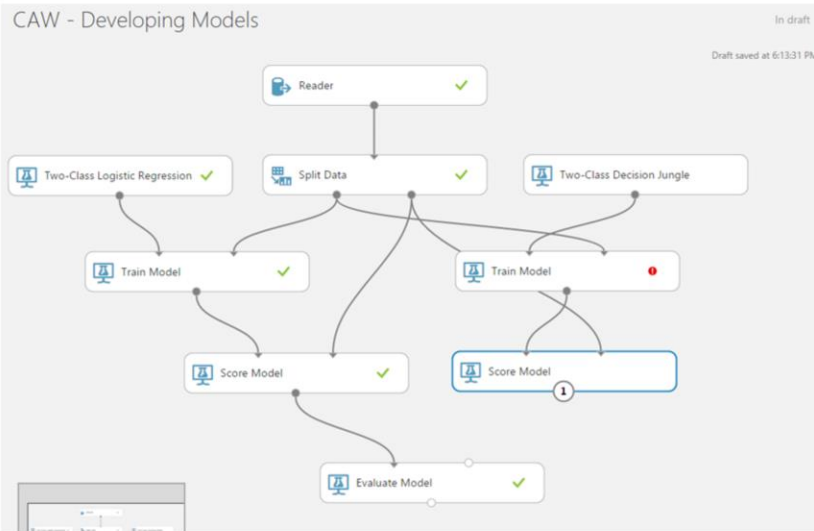
1. How to choose a model: https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/
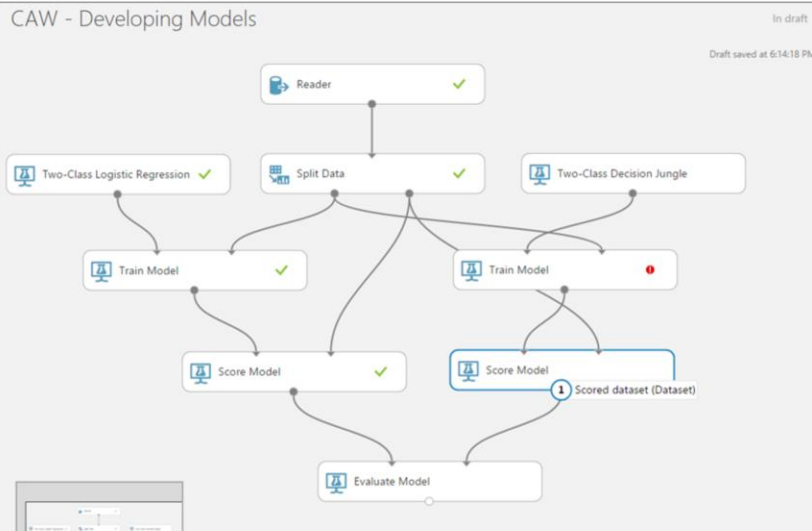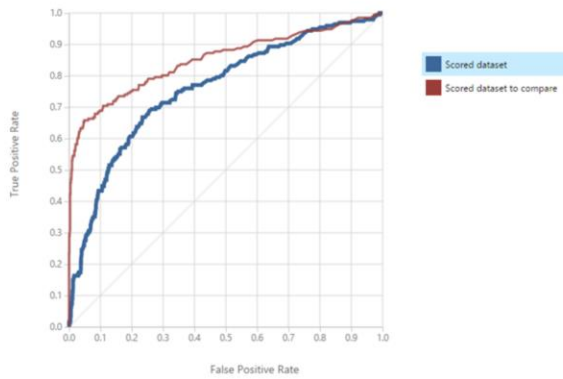
Add a Fancy Model

Add a Fancy Model

Add a Fancy Model

Compare Models

CAW - Developing Models › Evaluate Model › Evaluation results

ROC PRECISION/RECALL LIFT

Scored dataset
Scored dataset to compare

True Positive	False Negative	Accuracy	Precision	Threshold		AUC
14	182	0.863	0.583	0.5		0.760
False Positive	True Negative	Recall	F1 Score			
10	1194	0.071	0.127			

1. A more complex model comparison example: https://gallery.cortanaanalytics.com/Experiment/Compare-Binary-Classifiers-2?share=1

Using Azure ML:

1. Split data into testing and training sets
2. Build a model for predicting customer churn
3. Compare two models
4. Share an experiment
5. Deploy a model as an API

1. Use this for Q/A time