Machine Learning

With Azure Cloud and Excel API

SRIKRISHNA KRISHNARAO SRINIVASAN
aka
KRISHNA

Masters in Business Analytics and Information Systems

UNIVERSITY OF SOUTH FLORIDA

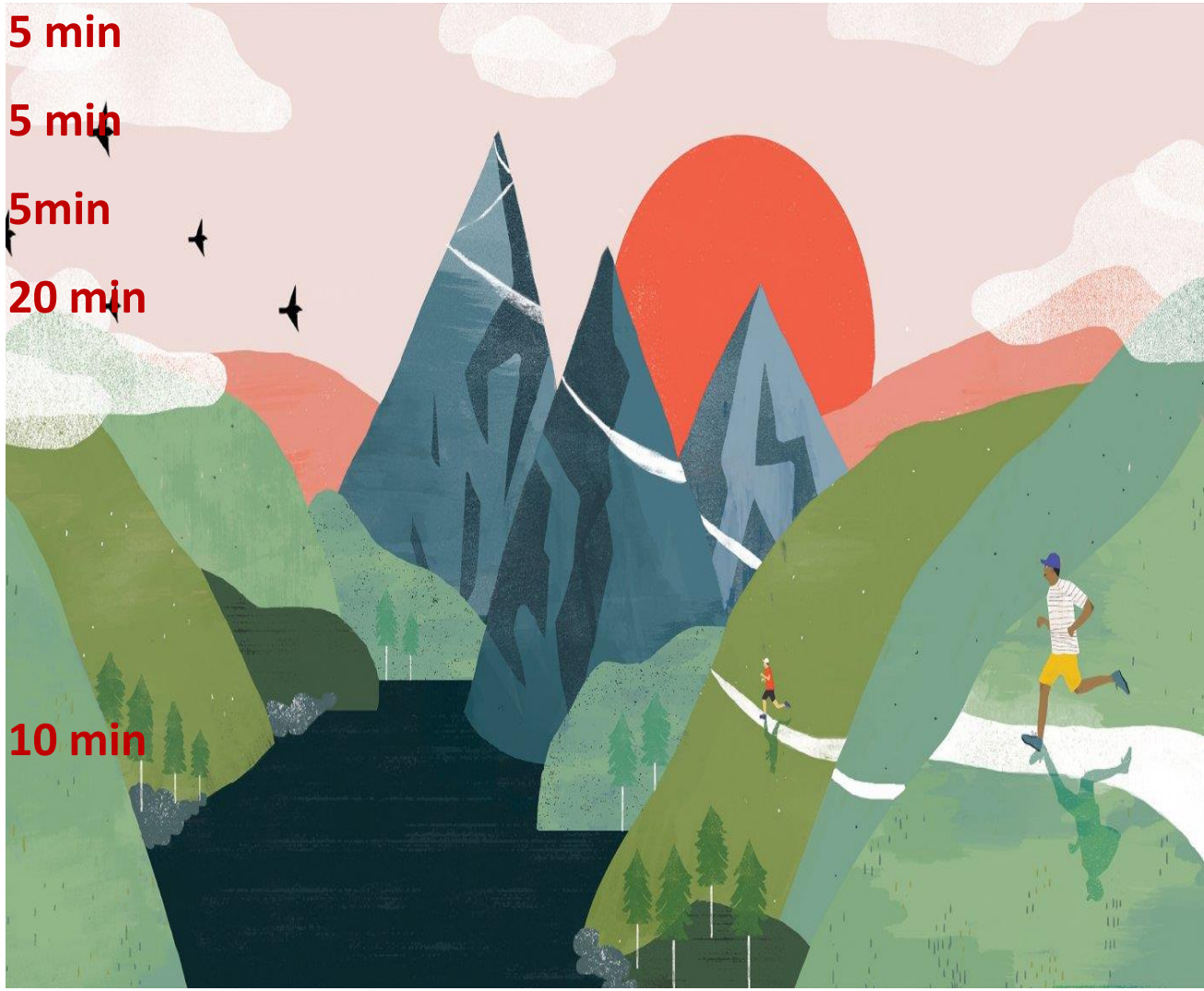
BaaN
HEXAWARE
ORACLE
Infosys
LANDMARK GROUP

I ♥ Microsoft Azure

PASS SQLSATURDAY
TAMPA | FEB 29 2020

Agenda

- Introduction & Agenda
- Business challenge
- Dataset description
- Machine Learning - AzureML
  - Load the dataset to Azure Cloud
  - Setup ML experiment
  - Evaluate predictive model
  - Create Webservice
- Microsoft Excel API
  - Access ML with Excel API
  - Predict for new dataset

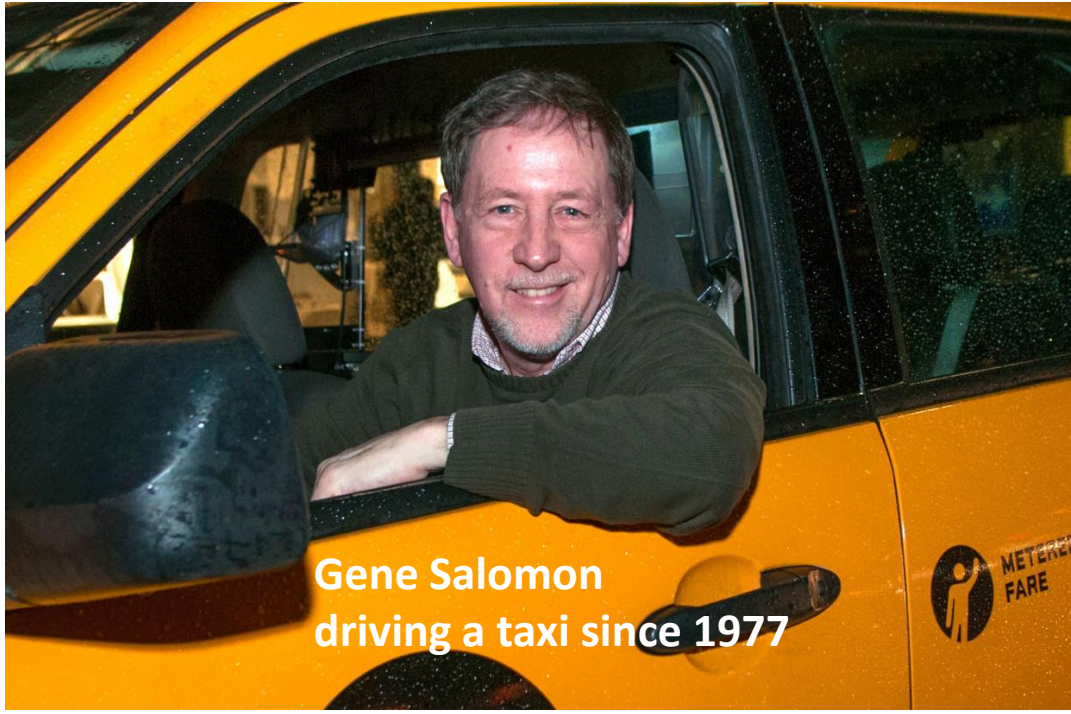


Business Challenge

"Was it the girl who rushed out of the cab seven times to puke?

Or was it the guy who, without the slightest provocation, would just start screaming?

Or the basket case who got out of the cab in the middle of the 59th Street Bridge?"



The NYC Taxi and Limousine Commission's official stance is that "it is impossible to pre-calculate a fare, because the meter rate depends on

traffic,
construction,
weather,
and route to the destination."

About the Dataset

	Column name	Description
1	vendor_id	The ID of the taxi vendor is a feature 1= Creative Mobile Technologies, LLC; 2= VeriFone Inc
2	rate_code	The rate type of the taxi trip is a feature 1=Standard rate 2=JFK 3=Newark 4=Nassau or Westchester 5=Negotiated fare 6=Group ride
3	passenger_count	The number of passengers on the trip is a feature
4	trip_time_in_secs	The amount of time the trip took. You want to predict the fare of the trip before the trip is completed. At that moment, you don't know how long the trip would take. Thus, the trip time is not a feature and you'll exclude this column from the model
5	trip_distance	The distance of the trip is a feature
6	payment_type	The payment method (cash or credit card) is a feature 1=Credit card 2=Cash 3=No charge 4=dispute 5=Unknown 6=voided trip
7	**fare_amount**	The total taxi fare paid is the label **(column to predict)**
	Github	https://github.com/sks4world/NYC_Taxi_Fare_Prediction
	Original link	https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

Regression machine learning task to predict the price based on factors in dataset

Machine Learning using AzureML

Sign in to Azure cloud https://studio.azureml.net

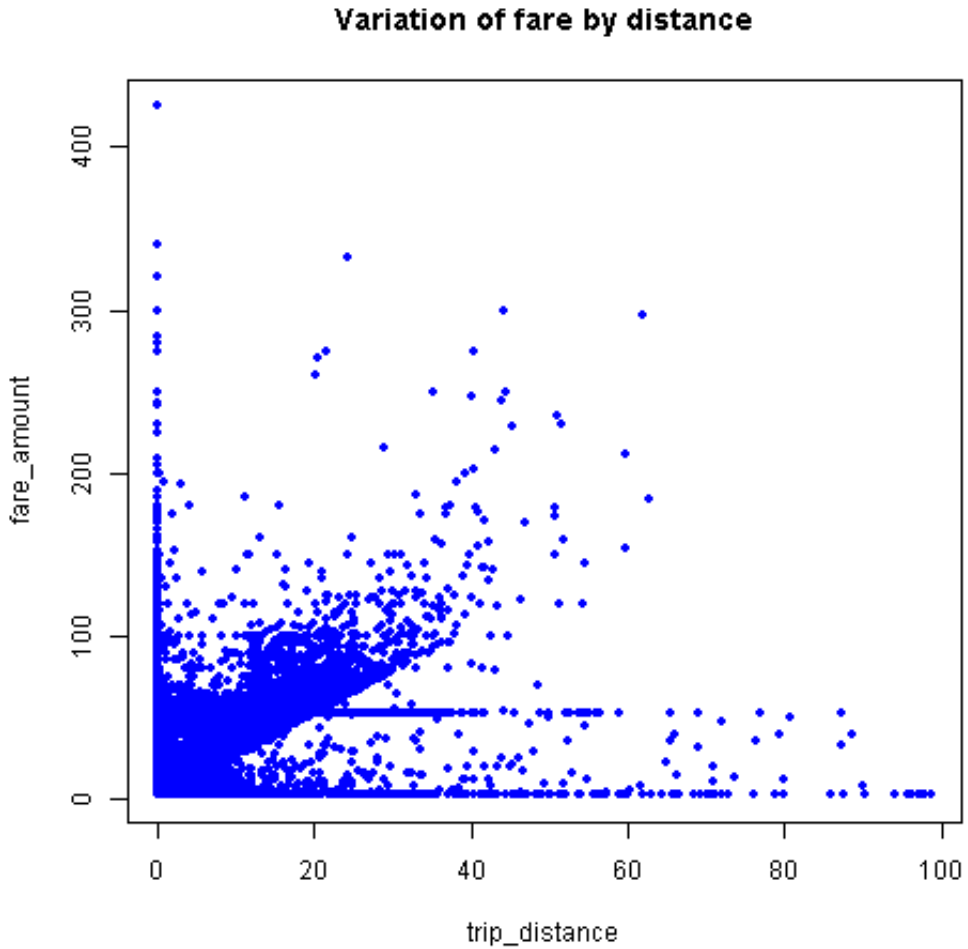
Sign in → Load Dataset → Pre-process data → Setup ML experiment → Evaluate Model → Create Predictive Webservice

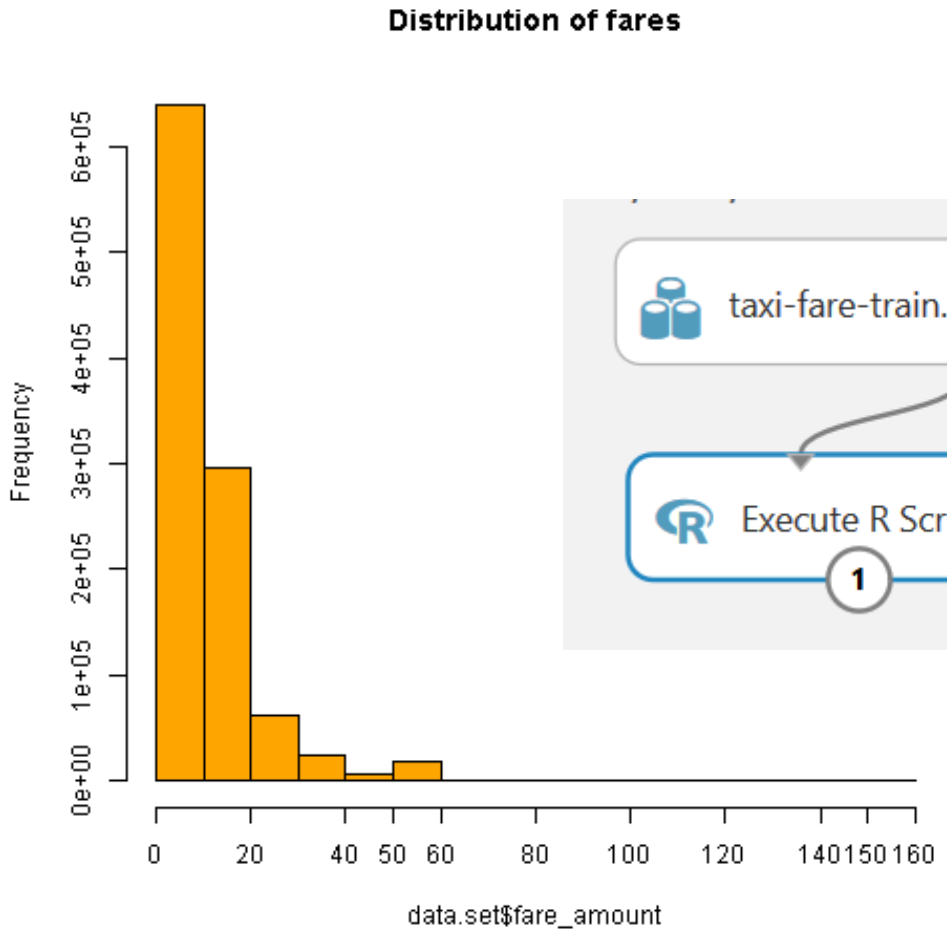
Quick Demo of Azure ML predictive model using Excel API

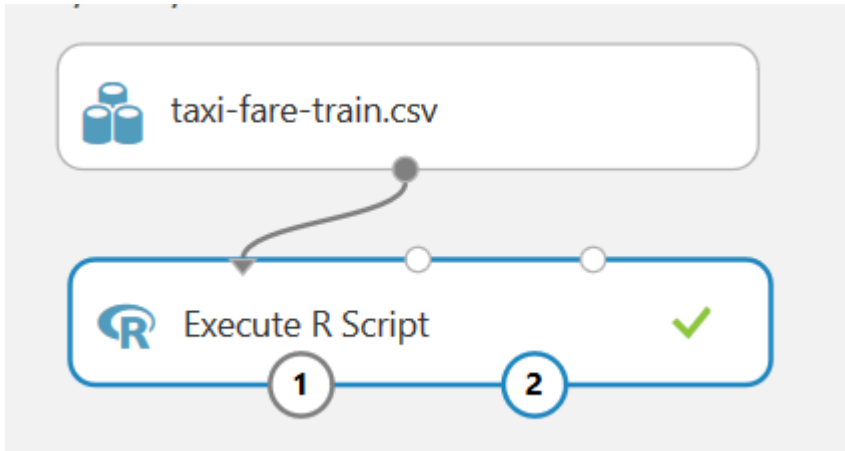
Model Evaluation

FARE AMOUNT	SCORED LABELS	ERROR	ABS.ERROR	SQ.ERROR
13	12.15	0.84	0.84	0.706
5.5	5.47	0.02	0.02	0.0007
7	6.84	0.15	0.15	0.024
9.5	9.53	-0.03	0.03	0.001
4.5	5.18	-0.68	0.68	0.468
4	4.27	-0.27	0.27	0.076
6	6.22	-0.22	0.22	0.049
5.5	6.39	-0.89	0.89	0.797
7	6.83	0.16	0.16	0.026
10.5	10.46	0.03	0.03	0.001
	MEAN BIAS ERROR	-0.88		
		MEAN ABS.ERROR	0.33	
			MEAN SQ.ERROR	0.21
			ROOT MSE	0.46

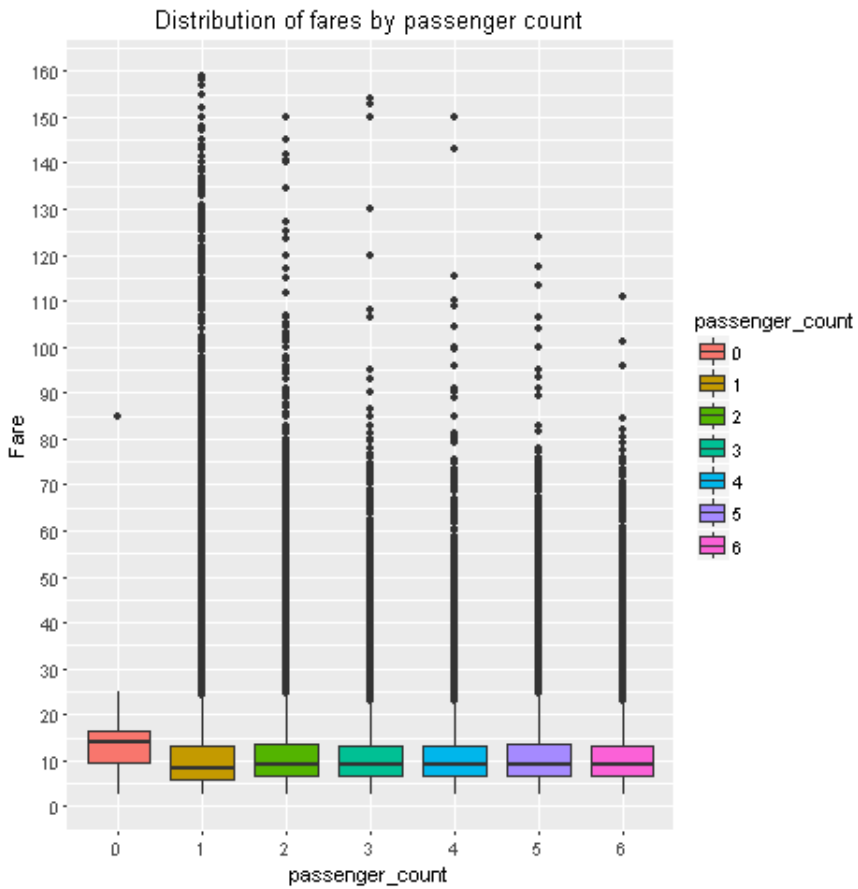
Load, visualize and Pre-process 1 of 2

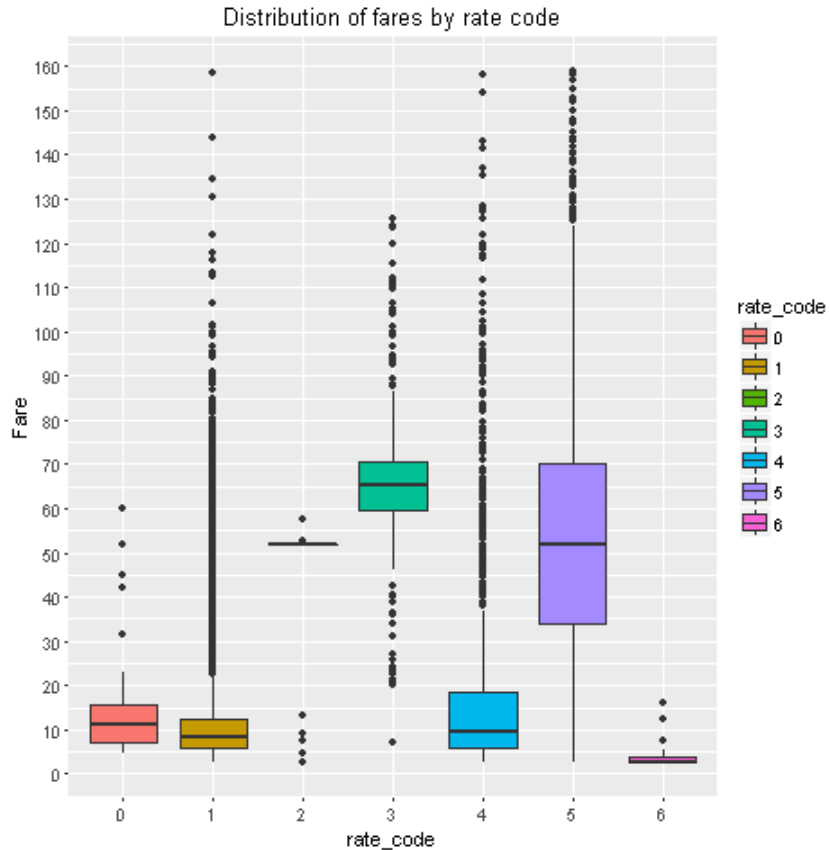




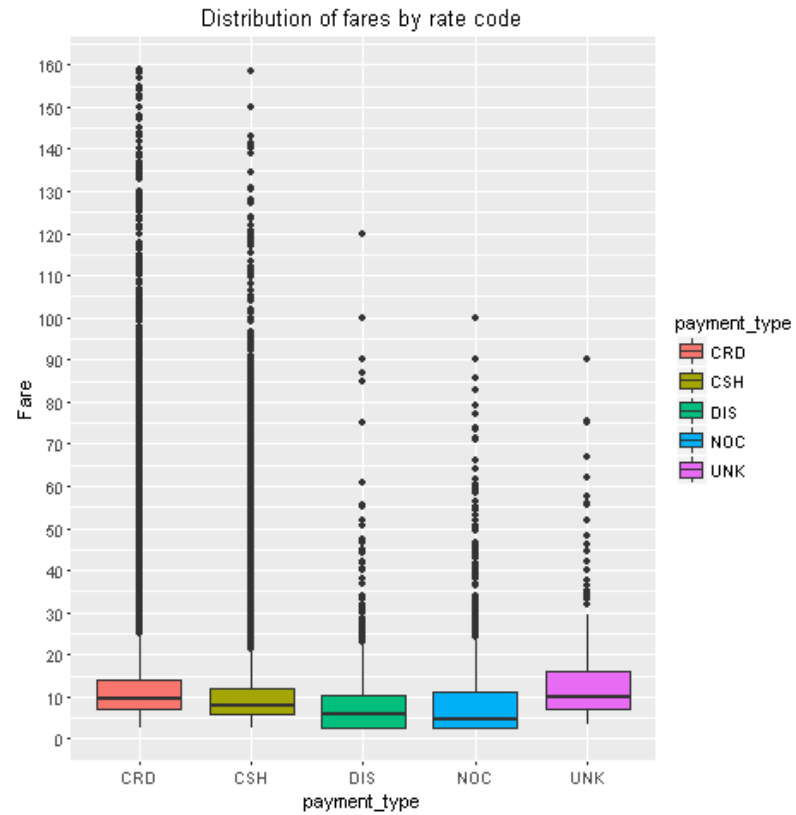


Load, visualize and Pre-process 2 of 2

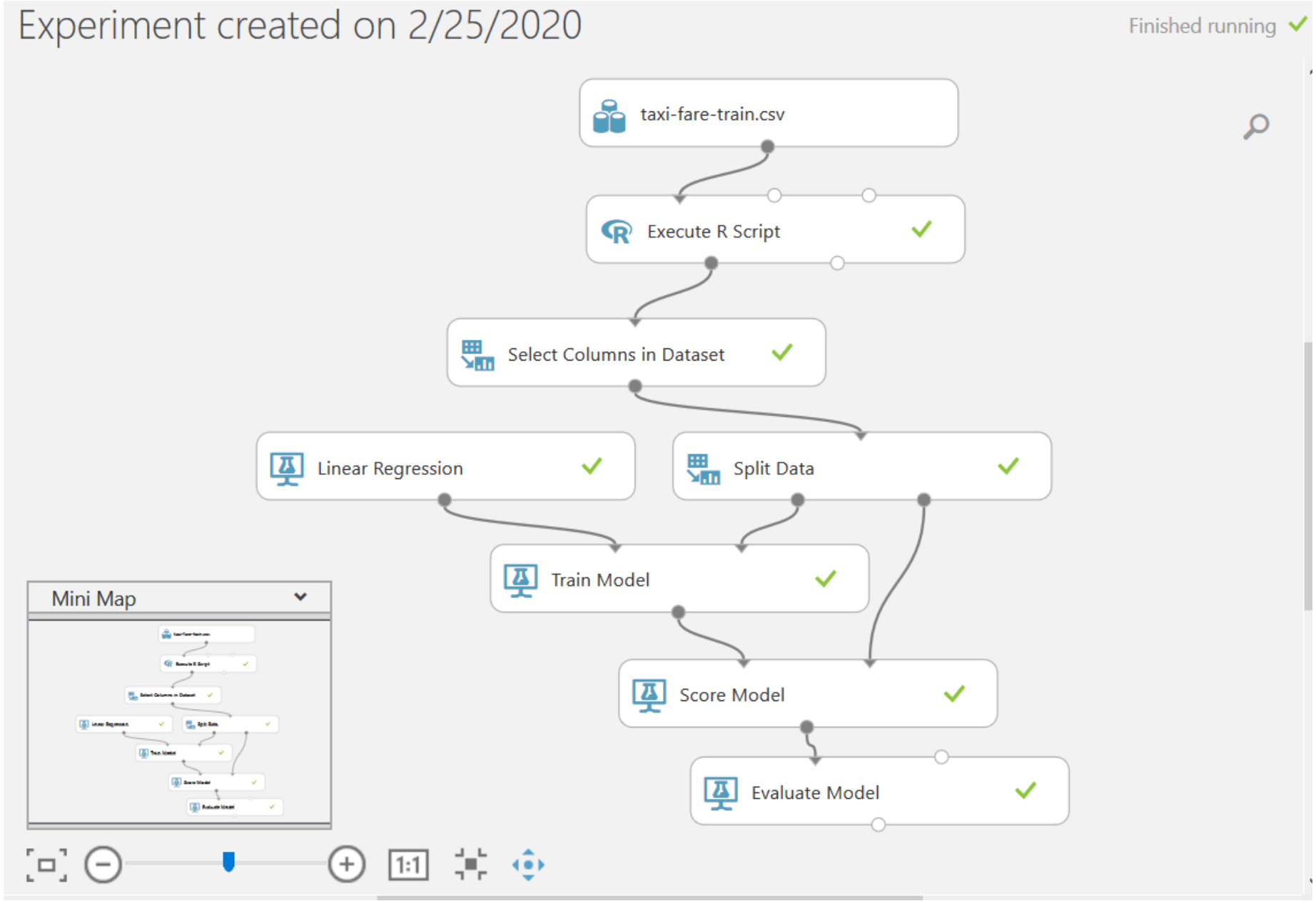




Convert to Categorical values

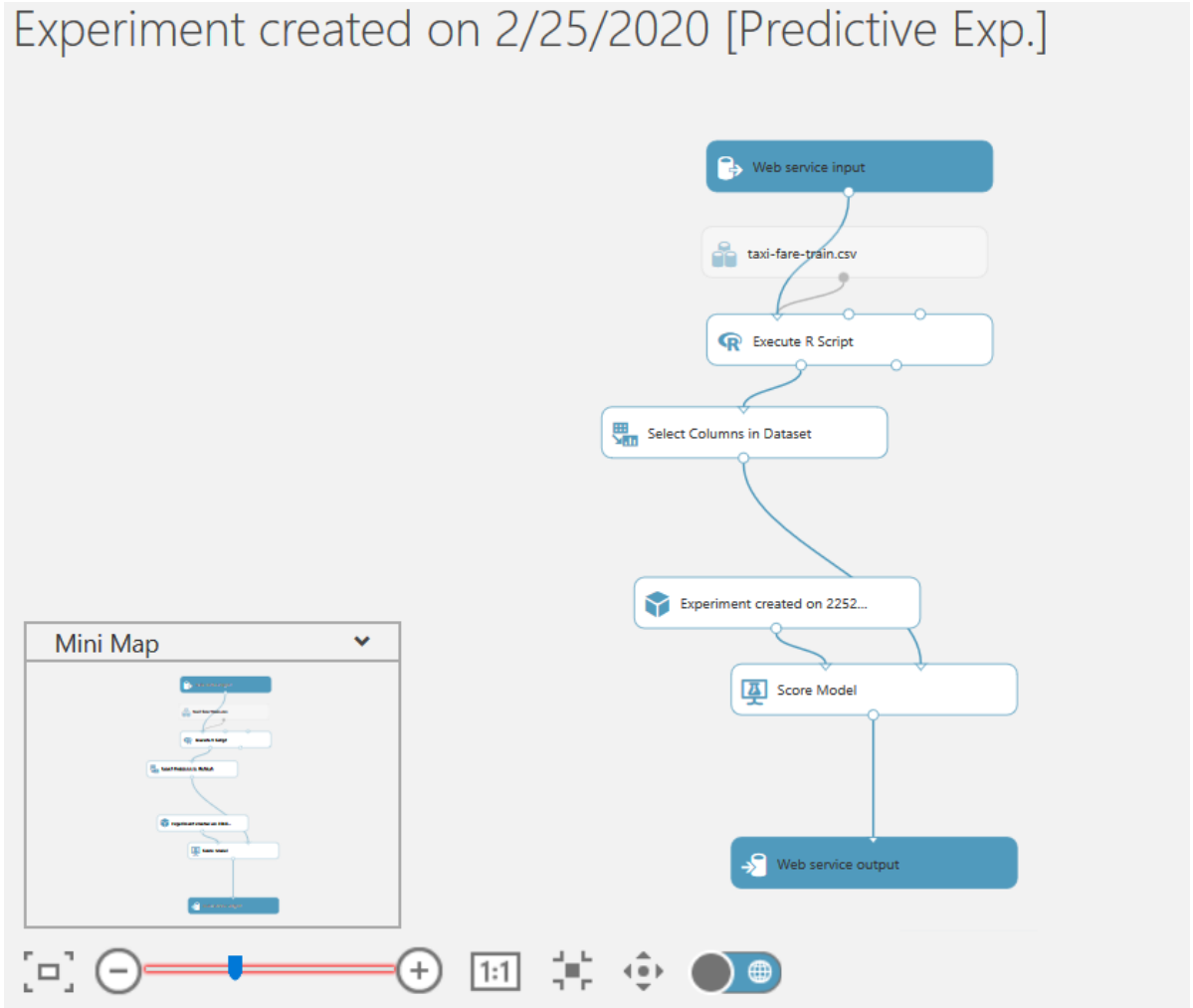


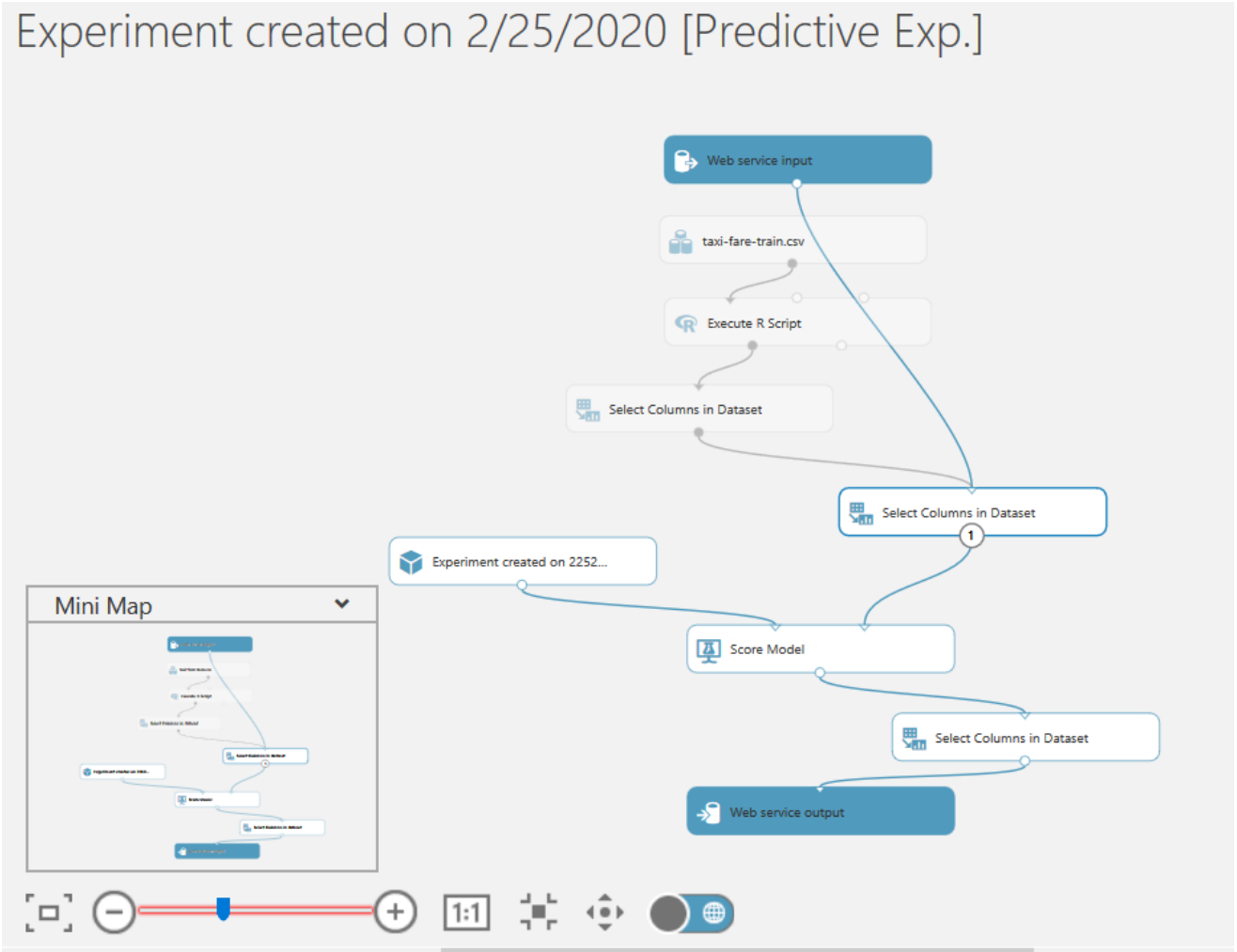
Setup ML Experiment and Evaluate



Mean Absolute Error	1.66
Root Mean Squared Error	3.65

Predictive Webservice





Microsoft Excel API

	A	B	C	D	E	F	G	H	I	J
1	**The downloadable Excel workbook in this page contains your Microsoft Azure Machine Learning Web Service API Key.**									
2	**WEB SERVICE URL**	https://ussouthcentral.services.azureml.net/workspaces/4b349e891d3e48759078da20c5b3b2cc/services								
3	**ACCESS KEY**	ufCM8i2ExPnVbfEbMXqvKZLiY28IPpKr1KoWb2GrJLuOm8yPeMPBmLtWOc10eQCwcjueINpmxvzMYygEo								
4	**SCHEMA**	https://ussouthcentral.services.azureml.net/odata/workspaces/4b349e891d3e48759078da20c5b3b2cc/s								

INSTRUCTIONS

Once you have enabled macros and the table has been generated, please enter your input values in the **PARAMETERS** section. (

If the web service you consume is hosted in a **Free Workspace** you may experience delay due to throttling. Upgrade to **a Stand**

PARAMETERS						EDICTED VAL
vendor_id	rate_code	passenger_count	trip_distance	payment_type	fare_amount	ScoredLabels
VTS	1	1	0.97	CSH	0	6.95527602

A	B	C	D	E	F	G	H	I	J	K	L
vendor_id	rate_code	passenger_	trip_distance	payment_type	fare_amount		Scored Labels				fare_amount
VTS	1	1	3.75	CRD			13.78474575				15.5
VTS	1	1	2.72	CRD			11.40379677				10
VTS	1	1	7.8	CSH			22.74351048				26.5
VTS	1	1	4.73	CSH			15.64689558				14.5
VTS	1	1	2.18	CRD			10.15553226				9.5
VTS	1	1	10.33	CSH			28.59186088				29.5
VTS	1	1	2.01	CSH			9.359341001				9
VTS	1	1	1.5	CRD			8.583643614				7.5
VTS	1	1	2.49	CSH			10.46890946				10.5

Thank you, to connect with me scan QR code

Interested to discuss about quick wins in Machine Learning for your team?