Being Smarter with Azure Machine Learning and R

Leila Etaati

Who Am I ?

RADACAD

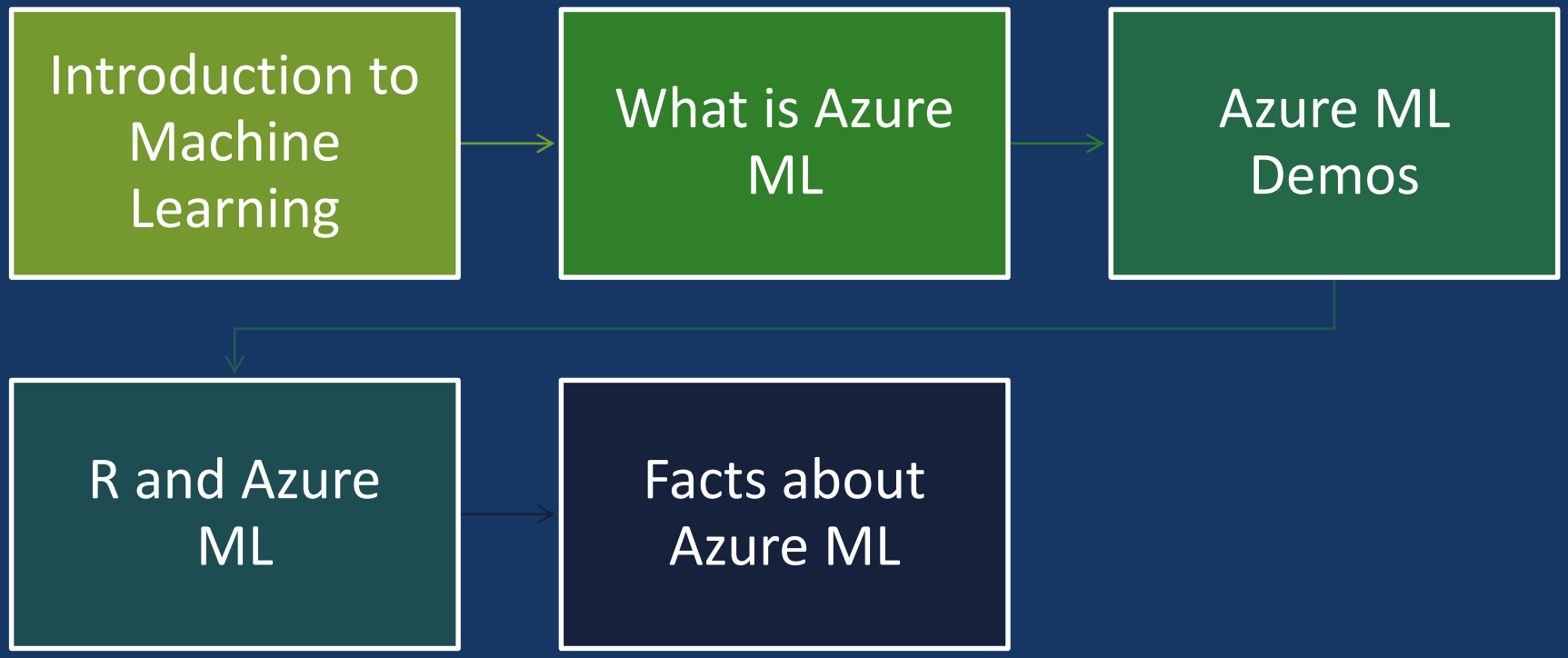
Leila Etaati

- PhD, Data Mining and BI Consultant
- Speakers in at least 20 Microsoft SQL Server Conferences (NZ ignite, SQL Rally, Code camp and SQL Saturday)
- 10 Years experiences in SQL Server
- Co-Lead New Zealand Business Intelligence User Group (Meet up)
- Lecturer and Tutor of BI and database

@leila_etaati

Leila@radacad.com

Agenda



@leila_etaati

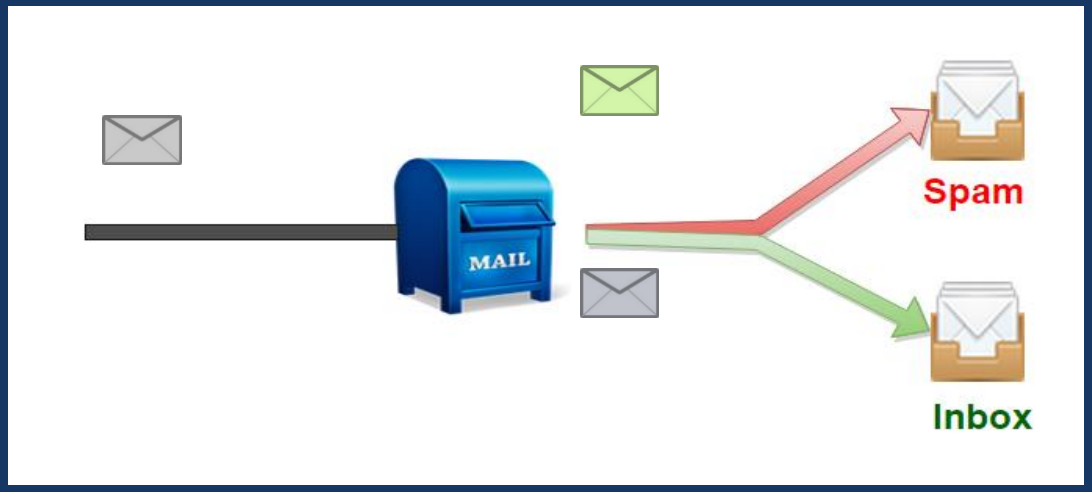
What is Machin Learning/Data mining?

"The goal of machine learning is to build computer systems that can adapt and learn from their experience."

-Tom Dietterich

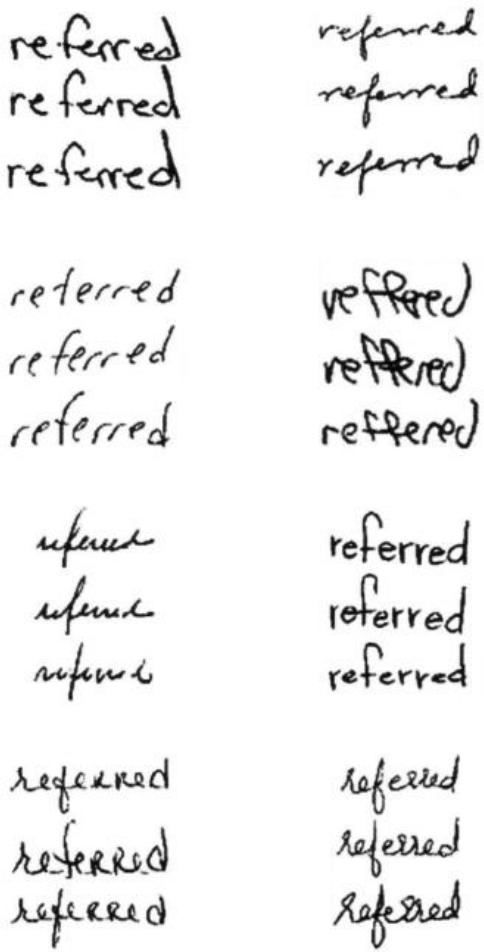
What is Machin Learning/Data mining?

Example of Using Machine Learning



The United States Postal Service processed over 150 billion pieces of mail in 2013—far too much for efficient human sorting.

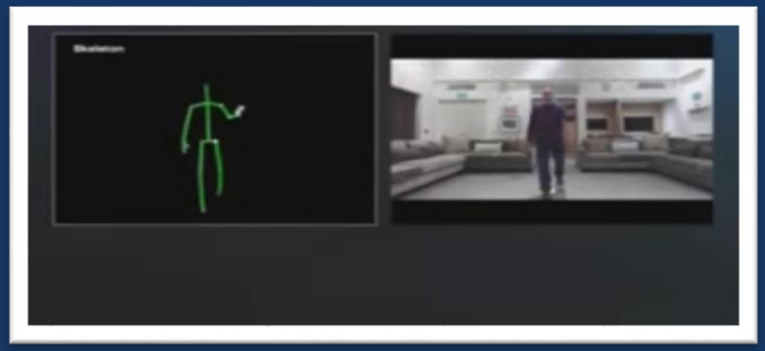
But as recently as 1997, only 10% of hand-addressed mail was successfully sorted automatically.



The challenge in automation is enabling computers to interpret endless variation in handwriting.

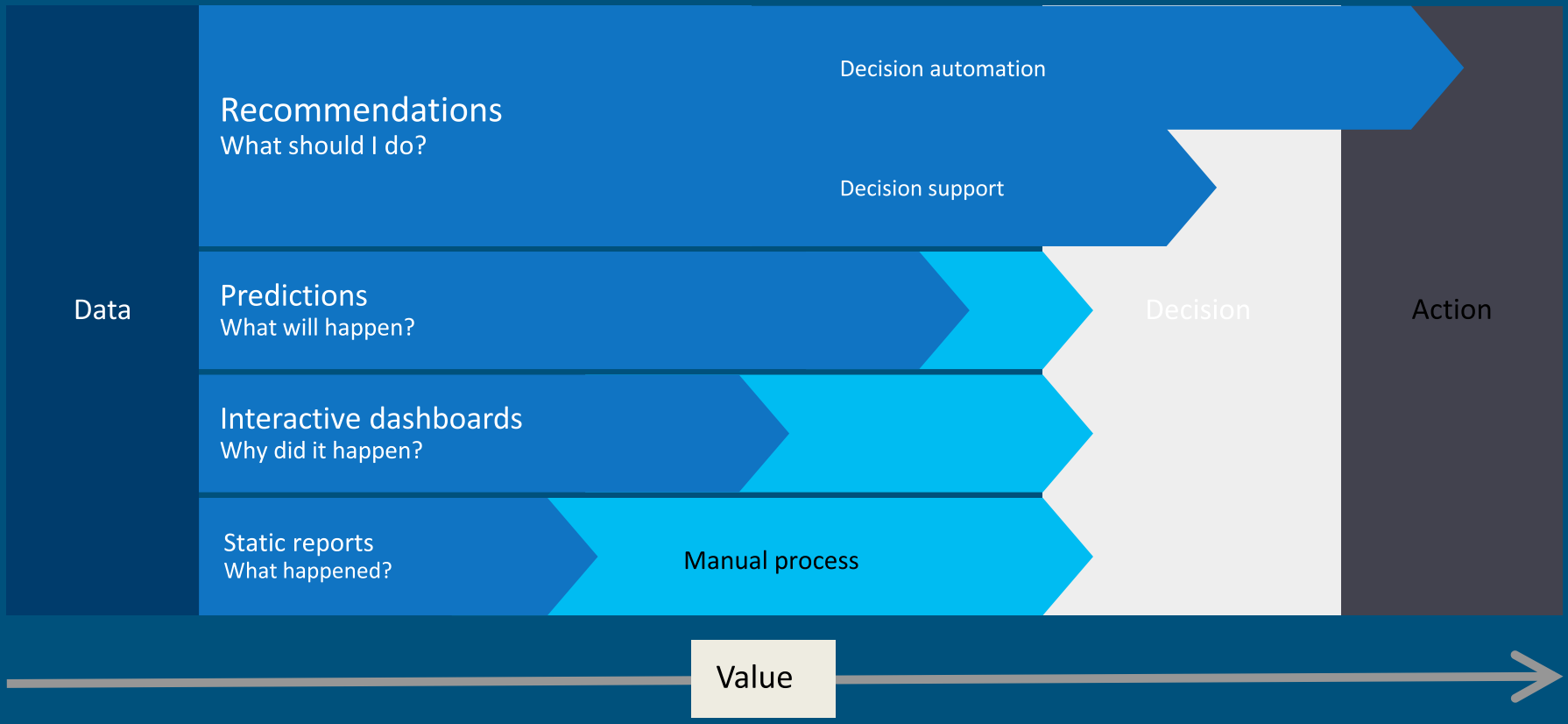
What is Machin Learning/Data mining?

Machine Learning



@leila_etaati

From data to decisions and actions



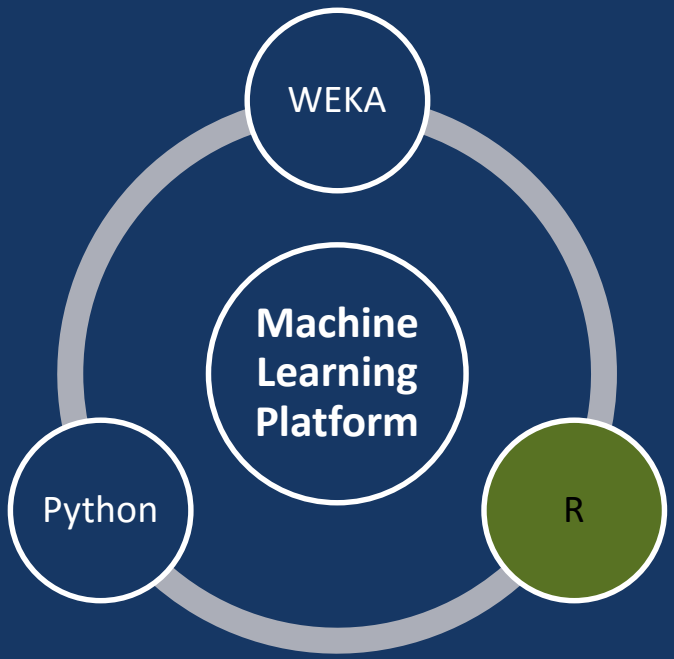
Should I used Machine Learning

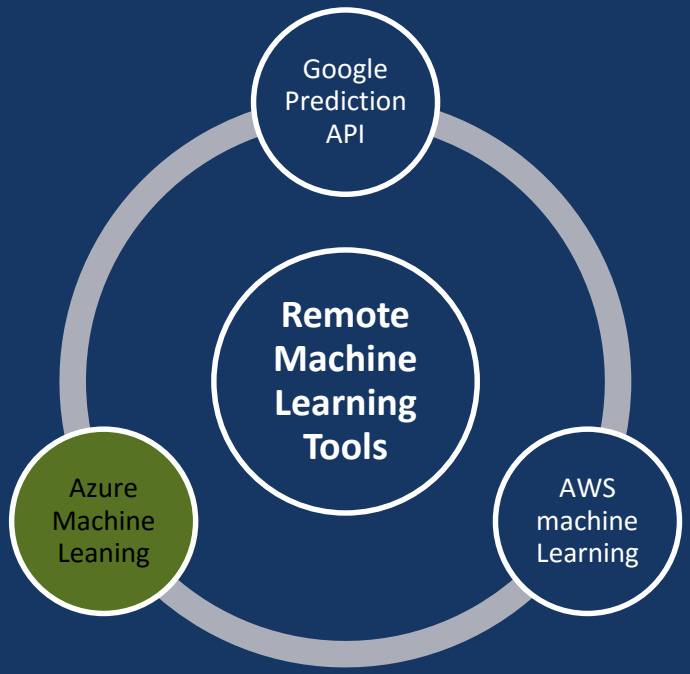
- Predication is **Small** part of experiences
- No Past data
- Many Rules govern Experience

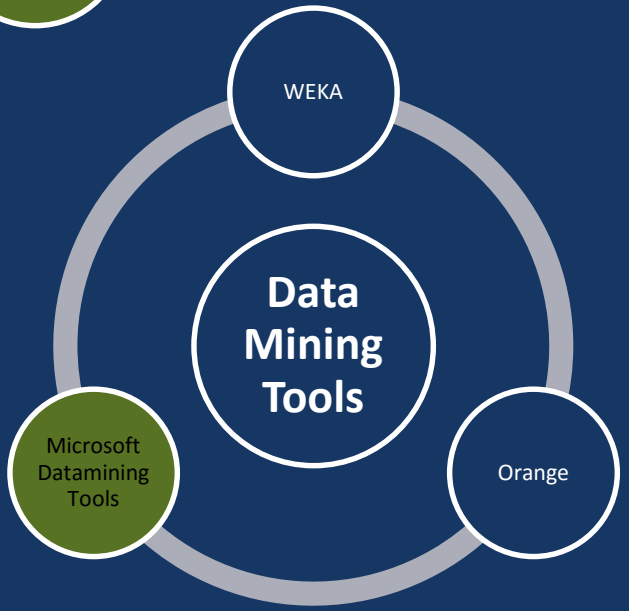
- Automated Predication is **Core**
- Lots of History
- Magic numbers in current prediction system

@leila_etaati

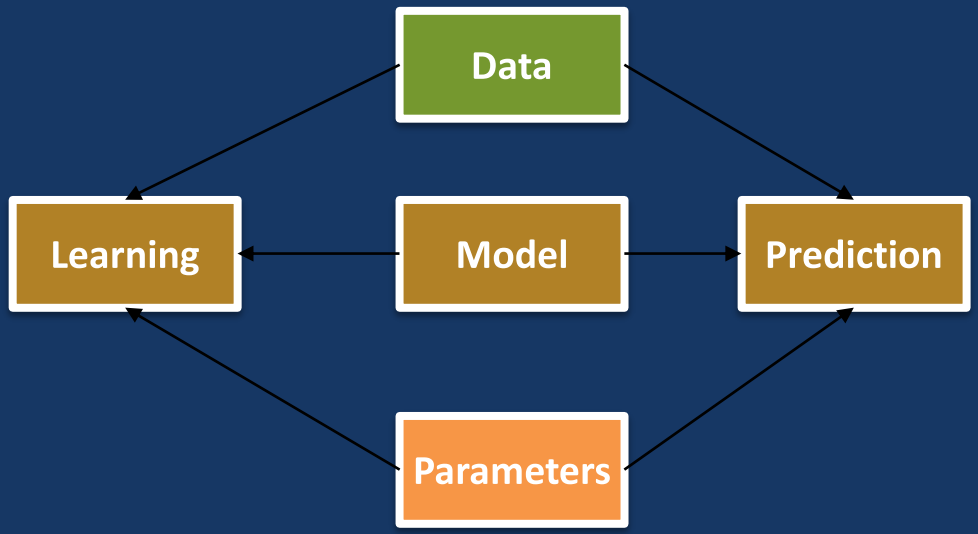
Various Tools for Machin Learning





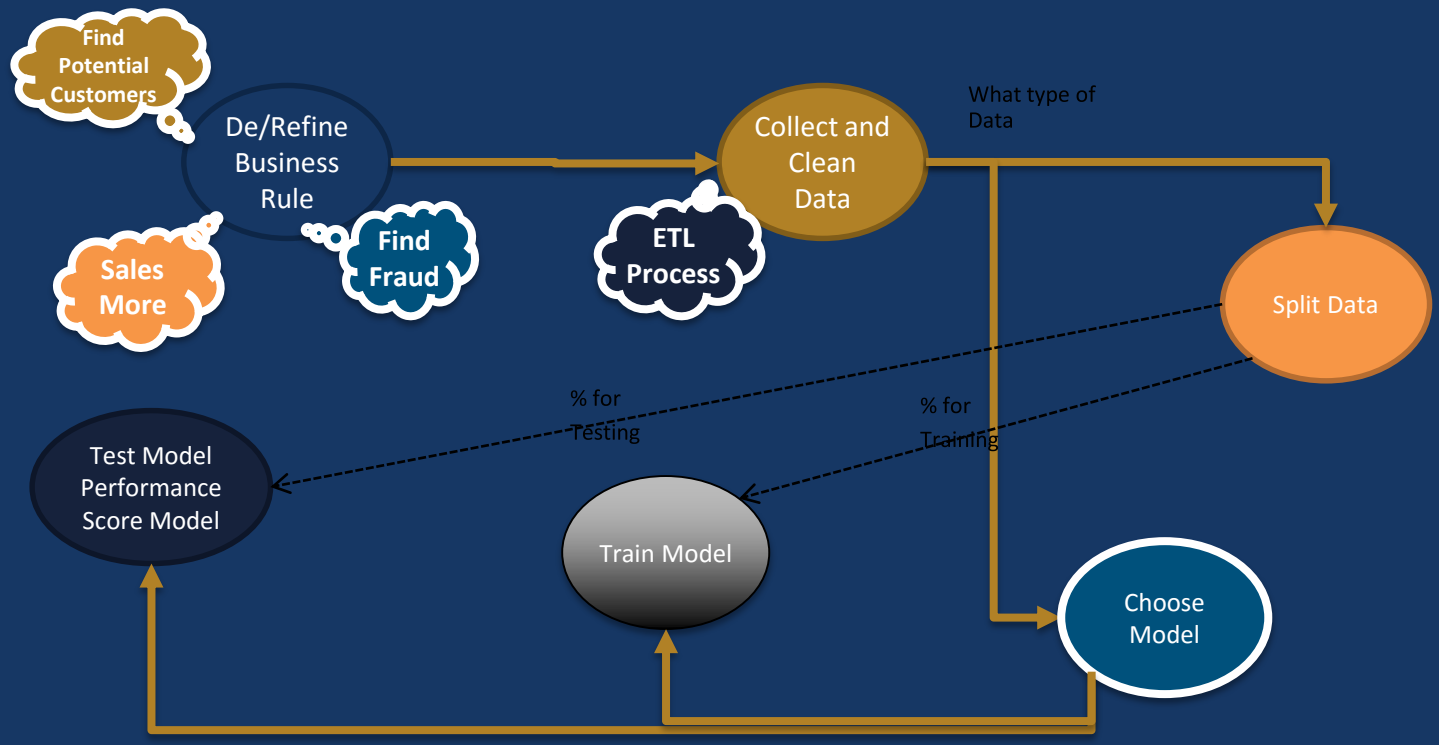


Machine Learning Concepts



@leila_etaati

Steps to Build a Machin Learning



@leila_etaati

Cortana Analytics Suite

Transform data into intelligent action

Microsoft & Machine Learning

15 years of realizing innovation

1999 — Computers work on users behalf, filtering junk email

2004 — Microsoft search engine built with machine learning

2005 — SQL Server enables data mining

2008 — Bing Maps ships with ML traffic-prediction service

2010 — Microsoft Kinect can watch users gestures

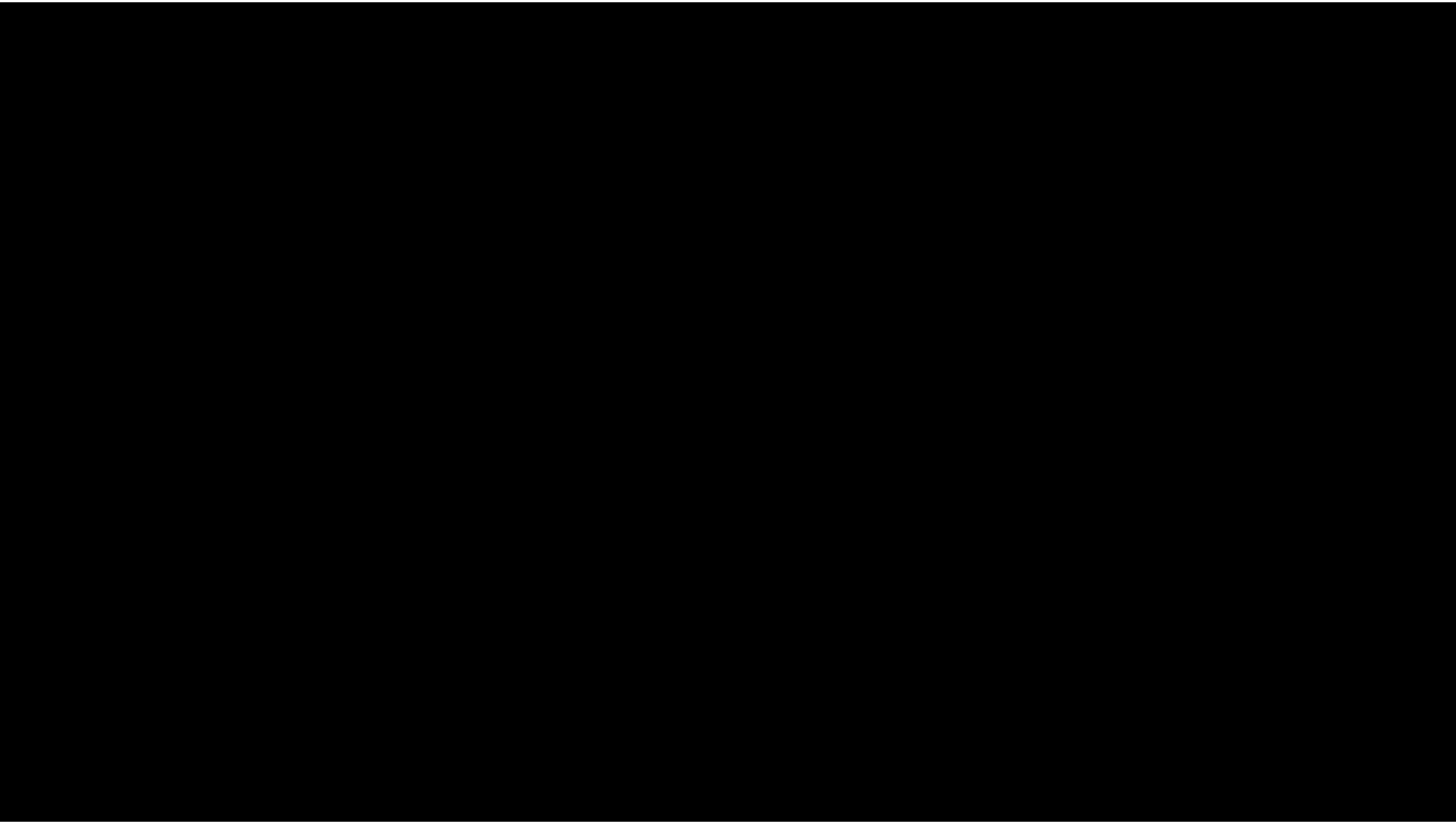
2012 — Successful, real-time, speech-to-speech translation

2014 — Microsoft launches Azure Machine Learning

John Platt,
Distinguished scientist at Microsoft Research

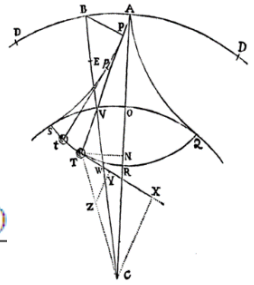
"Machine learning is pervasive throughout Microsoft products."

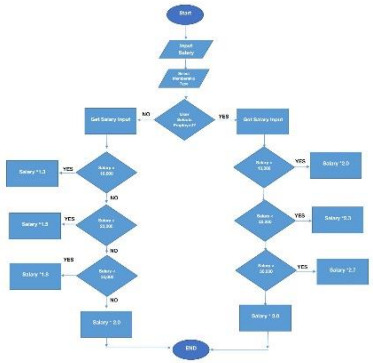
See how Microsoft Azure and HDInsight enables **AccuWeather** to use big data to examine traffic patterns and other user information, helping move their business forward.

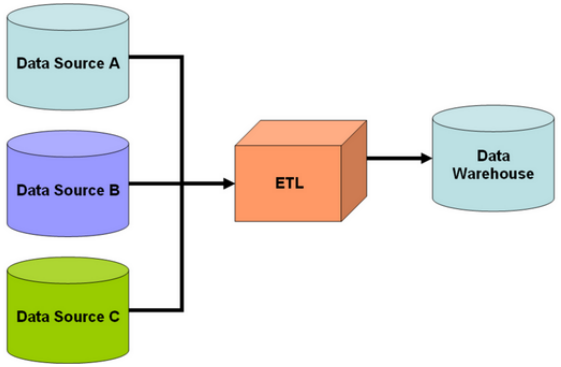


Data scientist

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$





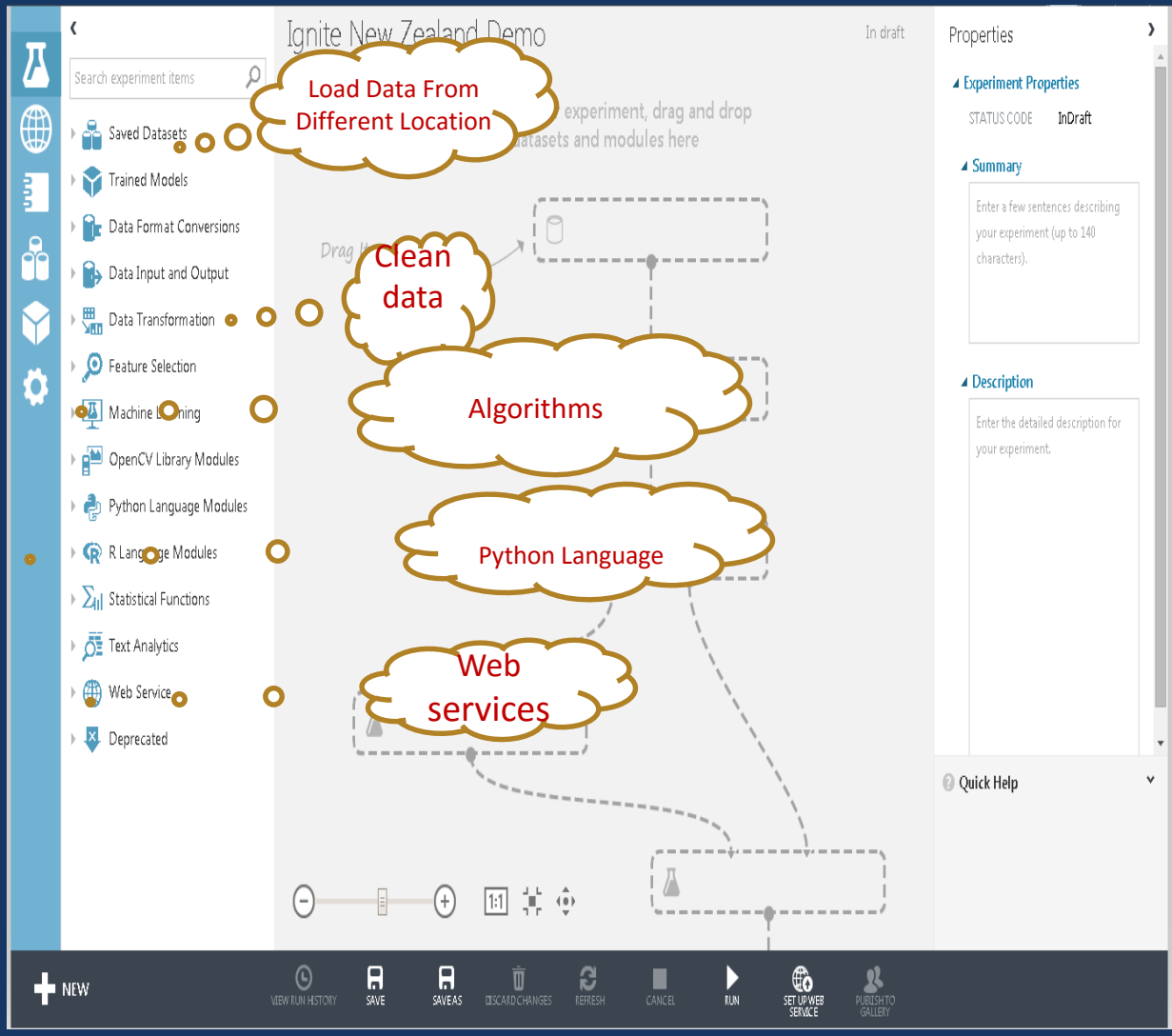


CoffeeScript
Microsoft .NET
python
php
Java
RAILS
Caml

@leila_etaati

Microsoft Azure Machine Learning

- Reduced Complexity
- Access Through Web Browser, no need to install any thing
- Collaborate work with anyone
- Visual composition, easy to use, No Coding
- Good storage of Algorithm (Use in Bing search, Xbox..)
- Have good support for R studio, Python and Jupyter notebook

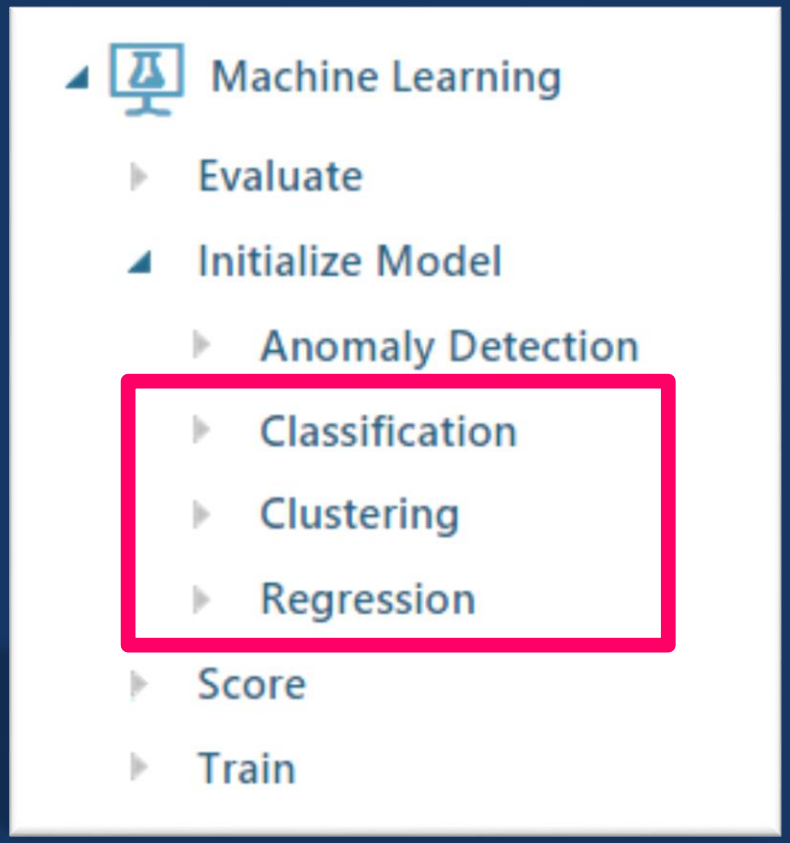


Using past data to predict the future

Imagine what machine learning could do for your business.

Churn analysis	Ad targeting	Image detection & classification
Equipment monitoring	Recommendations	Forecasting
Spam filtering	Fraud detection	Anomaly detection

Azure ML Algorithm

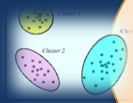


@leila_etaati

Descriptive and Predictive Analysis

1-Which other customers have similar preferences to this one?

2-What are the most common patterns in gasoline price changes?



Clustering Algorithm

Descriptive Analysis

1-When will this customer make another purchase?

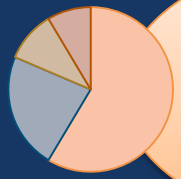
2-How many new followers will I get next week?

Regression Algorithm

Predictive Analysis

1-Will this customer click on the top link?

2-Which offer should this customer receive?



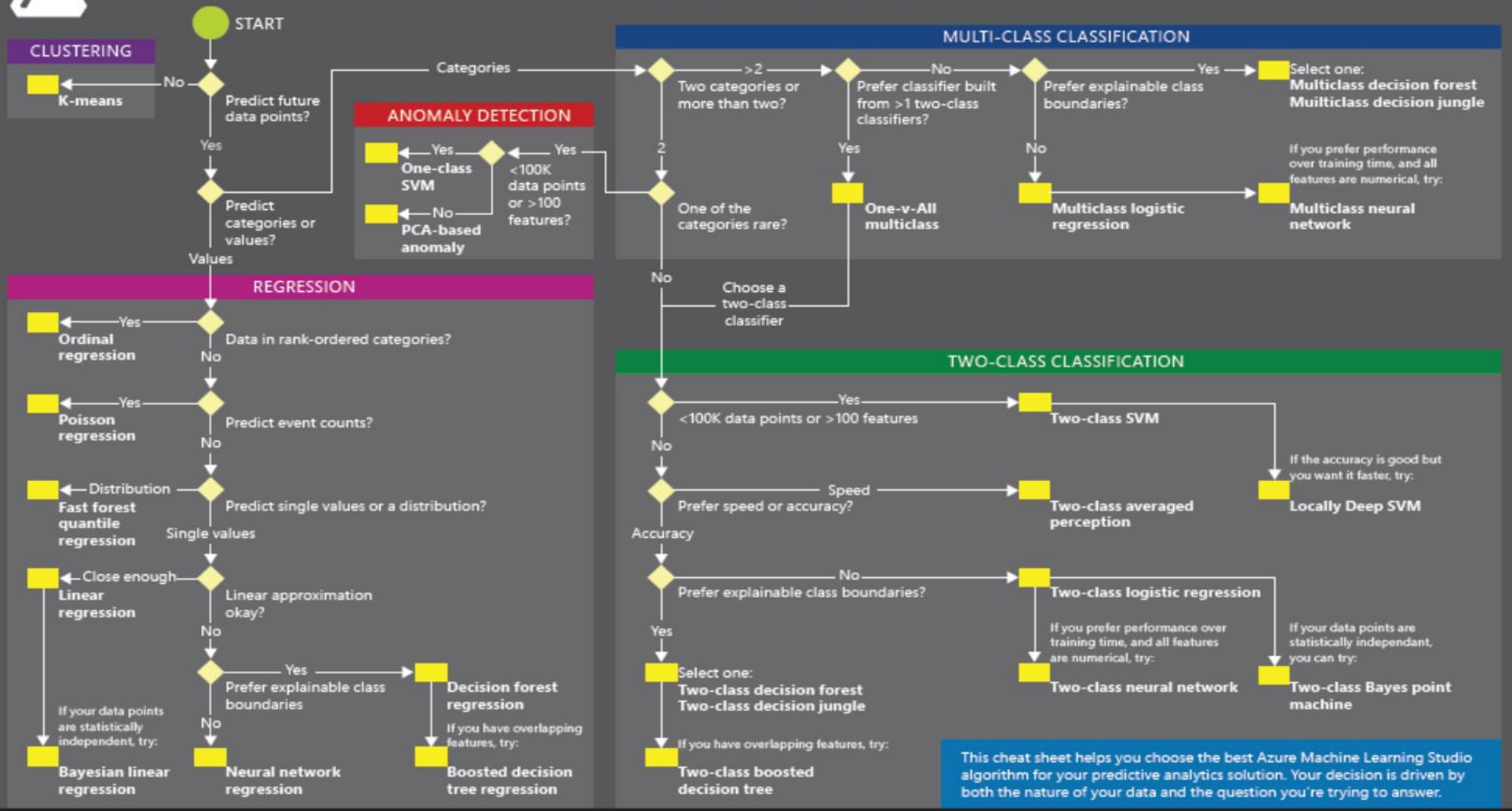
Classification Algorithm

Predictive Analysis

@leila_etaati

Predictive Maintenance	Two-class classification	Multi-class classification	Regression	Unsupervised learning
Should I replace this part now?	✓			
When will this part fail?		✓	✓	
Will this tire fail in the next thousand miles?	✓			
What is the remaining useful life of this aircraft engine?		✓	✓	
Which vehicle needs servicing most urgently?		✓		
Which groups of sensors in this jet engine tend to vary with (and against) each other?				✓

Microsoft Azure Machine Learning: Algorithm Cheat Sheet



This cheat sheet helps you choose the best Azure Machine Learning Studio algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you're trying to answer.

© 2015 Microsoft Corporation. All rights reserved.
Created by the Azure Machine Learning Team
Email: AzurePoster@microsoft.com
Download this poster: http://aka.ms/MLCheatSheet

Microsoft

Demo On Azure ML

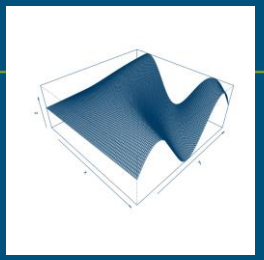
Titanic –Classification Algorithm

Survived or Not

Titanic sank on15 April 1912 after colliding with an iceberg

@leila_etaati

What is ?



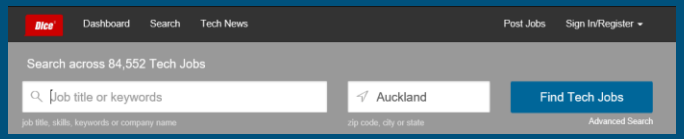
- The R statistical programming language is a free open source package based on the S language developed by Bell Labs.
- The language is very powerful for writing programs.
- Many statistical functions are already built in.
- Contributed packages expand the functionality to cutting edge research.
- Since it is a programming language, generating computer code to complete tasks is required.
- http://www.r-project.org

R and Data Mining : Examples and Case Studies by Yanchang Zhao
https://cran.r-project.org/doc/contrib/Zhao_R_and_data_mining.pdf

Growth In Demand ?

- Rexer Data Mining survey

- R is the highest paid IT skill Dice.come, Jan 2014



- R most used-data science language after SQL-O'Reilly, Jan 2014
- R is used by 70% of data miners. Rexer, Sept 2013

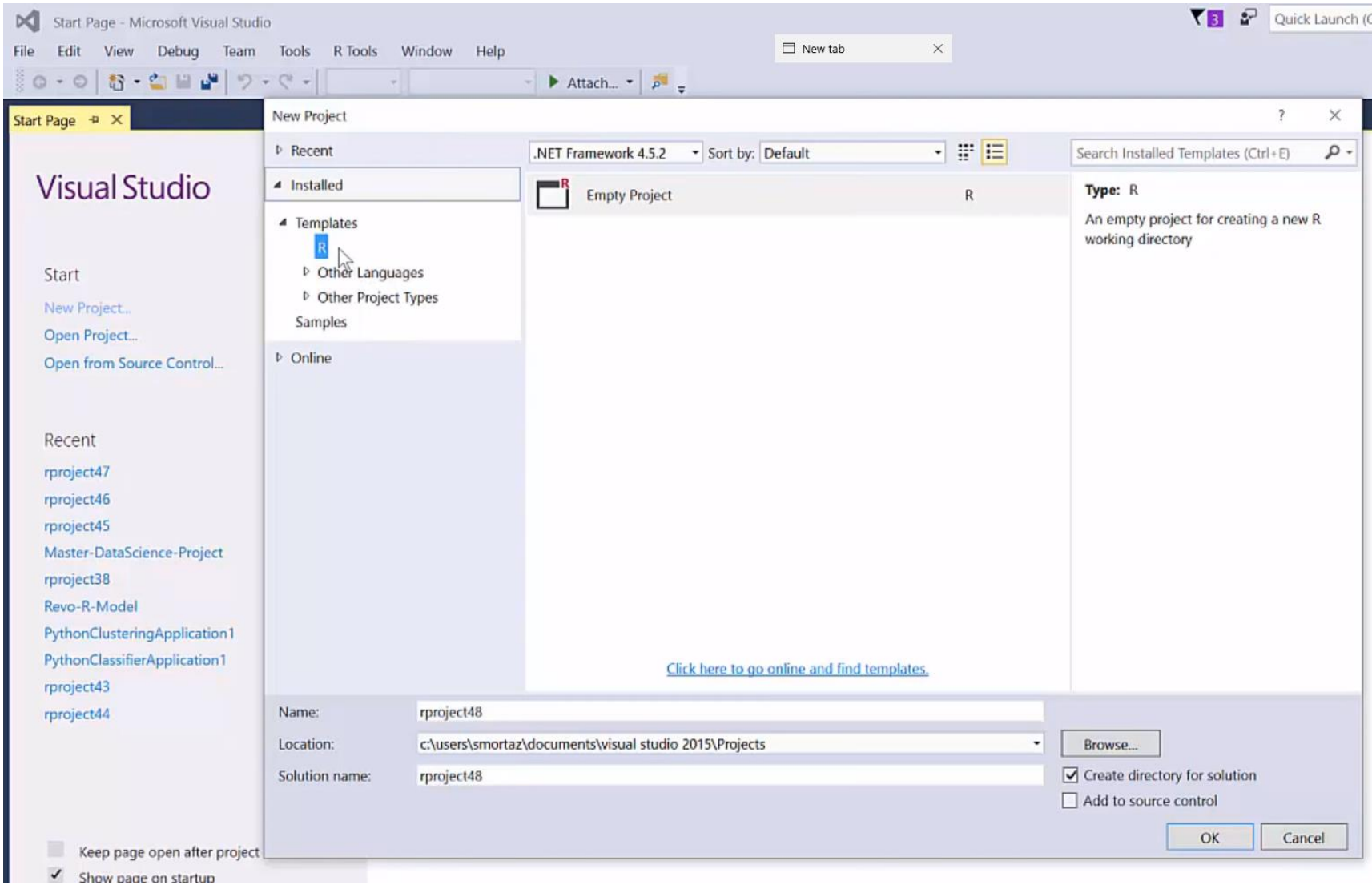
- R is #15 all programming languages. REdMonk, Jan 2014

- R growing faster than any other data science language. KDNuggs
- R is in-memory and limited in size of data that you can process.

Demo on R

Cortana Analytics Suite successful experience with Azure ML

https://www.youtube.com/watch?v=YxmAEMmwXYU


Start Page - Microsoft Visual Studio
File Edit View Debug Team Tools R Tools Window Help
New tab
Quick Launch (Ctrl+Q)
Attach...
Start Page
Visual Studio
Start
New Project...
Open Project...
Open from Source Control...
Recent
rproject47
rproject46
rproject45
Master-DataScience-Project
rproject38
Revo-R-Model
PythonClusteringApplication1
PythonClassifierApplication1
rproject43
rproject44
Keep page open after project
Show page on startup
New Project
Recent
Installed
Templates
R
Other Languages
Other Project Types
Samples
Online
.NET Framework 4.5.2
Sort by: Default
Search Installed Templates (Ctrl+E)
Empty Project
R
Type: R
An empty project for creating a new R working directory
Click here to go online and find templates.
Name: rproject48
Location: c:\users\smortaz\documents\visual studio 2015\Projects
Solution name: rproject48
Browse...
Create directory for solution
Add to source control
OK
Cancel

SSAS Vs Azure ML

Features	Usability	Cost	Support
• End to end Product • Canned algorithm • Not possible to change algorithm • DMX Code	• More Visual • Excel Add on • It's not easy to start • All users can use	If you purchase SQL Sever: Free	Few books and small online community
• Current and up to date algorithm • Integration with R and Python • Cloud base • REST format	• Hard to interpret • Drag and Drop UI • Customize the Algorithm • Excel Add on	Free version, limited options	More online Community

Azure ML Pricing

Pricing Details

Machine Learning is offered in two tiers: Free and Standard.

Features by tier are compared in the table below:

	FREE	STANDARD
Authentication	Microsoft account (does not require an Azure subscription or a credit card)	Requires Azure subscription
Max Number of Modules per Experiment	100	Unlimited
Max Experiment Duration	1 hour	None
Max Storage Space	10 GB	Unlimited - BYO
Execution / Performance	Single node	Multiple nodes
Staging Web API	Yes (Throttled performance)	Yes (Selectable performance)
Production Web API	No	Yes
SLA	No	Yes

Read More

Machine Learning blog

http://blogs.technet.com/b/machinelearning

Videos: PASS Data Science Virtual Chapter

https://www.youtube.com/channel/UCqB3xWdwjA9soFV6EOu7qfg

Microsoft Ignite Conference

https://channel9.msdn.com/Events/Ignite/Microsoft-Ignite-New-Zealand-2015/M234

- Leila.etaati@gmail.com
- leila@radacad.com