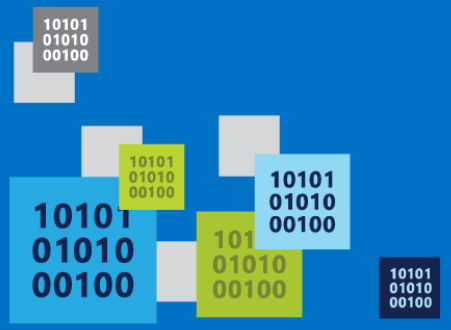
CloudBread Camp
Machine Learning

프로젝트 : http://aka.ms/cbp

프로젝트 Facebook : http://aka.ms/cbfb



Agenda

- Machine learning
- Azure Machine Learning 소개
- Machine learning 흐름 시연
- 비즈니스 시나리오 소개
- 요약

예측 - Prediction

과거의 데이터를 이용해 미래를 예측

Machine learning과 예측분석(predictive analytics)는 비즈니스 확장을 위한 핵심 역량


Churn analysis


Social network analysis

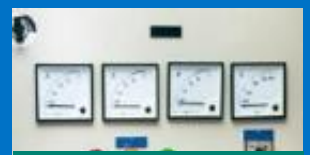

Recommenda-tion engines


Location-based tracking and services


Vision Analytics


Weather forecasting for business planning


Legal discovery and document archiving


Equipment monitoring


Advertising analysis


Pricing analysis


Fraud detection


Personalized Insurance

Machine learning - ?

Machine Learning

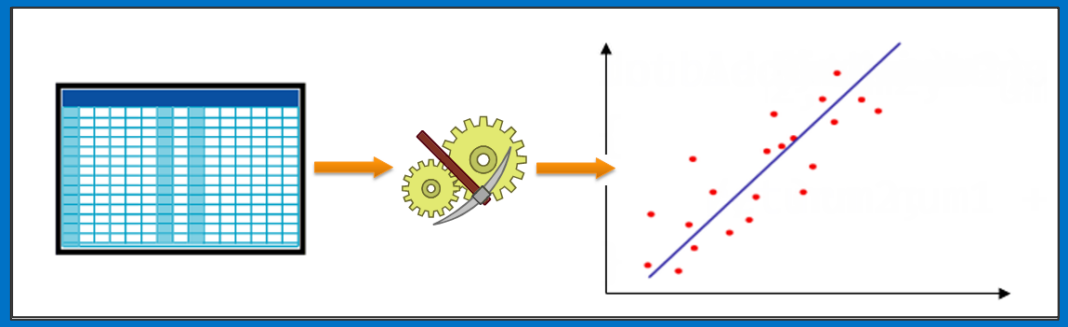
Subfield of computer science and statistics that deals with the construction and study of systems that can learn from data, rather than follow only explicitly programmed instructions

-Wikipedia

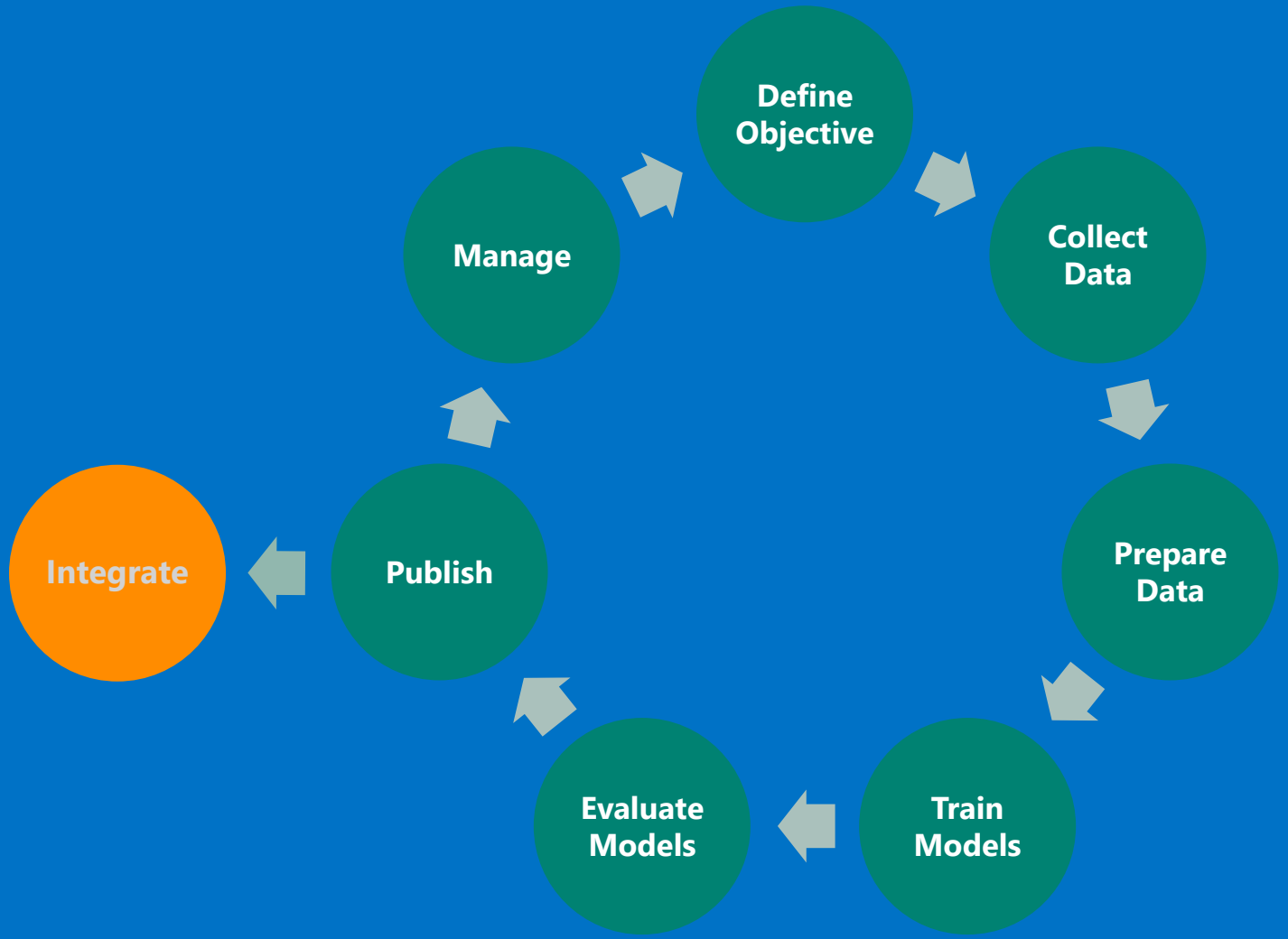
고객사의 수익 예측이 필요해

$f()$

Age, Marital Status, Gender, Yearly Income, Total Children, Education, Occupation, Home Owner, Commute Distance



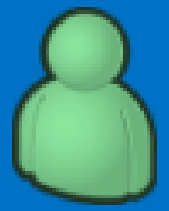
Machine learning 워크플로우



Machine learning 역할

Data scientist

A highly educated and skilled person who can solve complex data problems by employing deep expertise in scientific disciplines (mathematics, statistics or computer science)



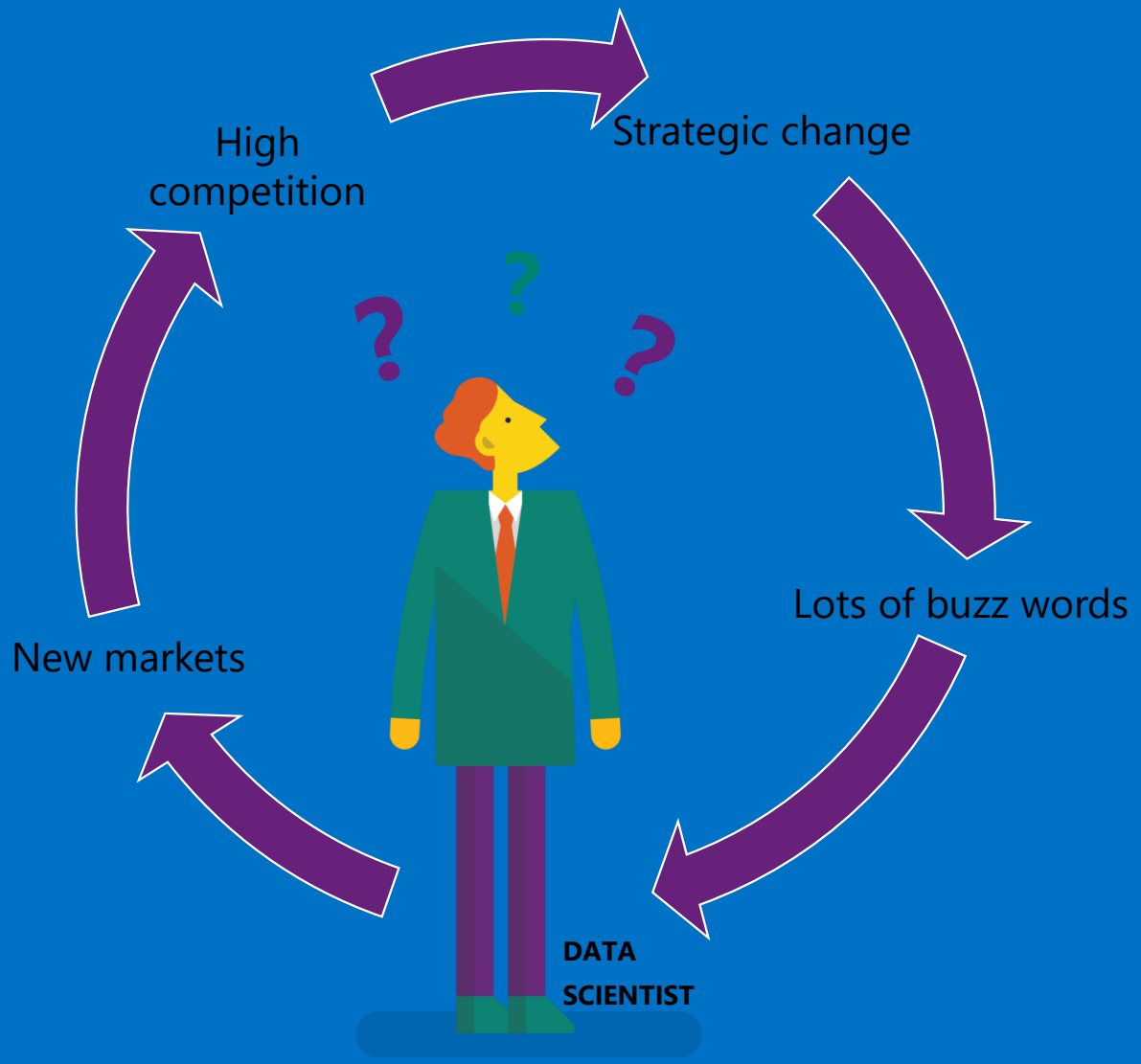
Data professional

A skilled person who creates or maintains data systems, data solutions, or implements predictive modelling
Roles: Database Administrator, Database Developer, or BI Developer

Software developer

A skilled person who designs and develops programming logic, and can apply machine learning to integrate predictive functionality into applications

Machine learning 난제



Expensive

Isolated data

Tool chaos

Complexity

Traditional approach

- Guessing
- Rules of thumb
- Trial and error

Consequences

- Lost opportunities
- Expensive operative mistakes

Azure Machine Learning 소개

Azure Machine Learning

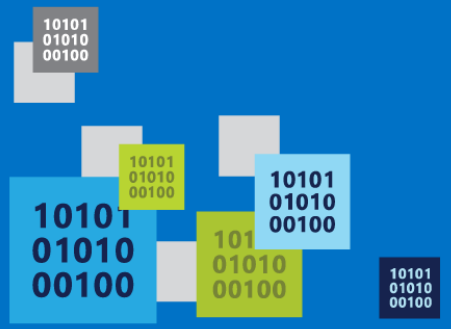
강력한 클라우드 기반 예측 분석 가능
Advanced Analytics 솔루션을 손쉽게 빌드, 배포, 공유 가능
브라우저 기반 빠른 개발
다양한 Azure의 데이터 서비스들과 연계 가능

- Azure HDInsight (Big Data)
- Azure SQL Database, and
- Virtual Machines

Azure Machine Learning
동작방식

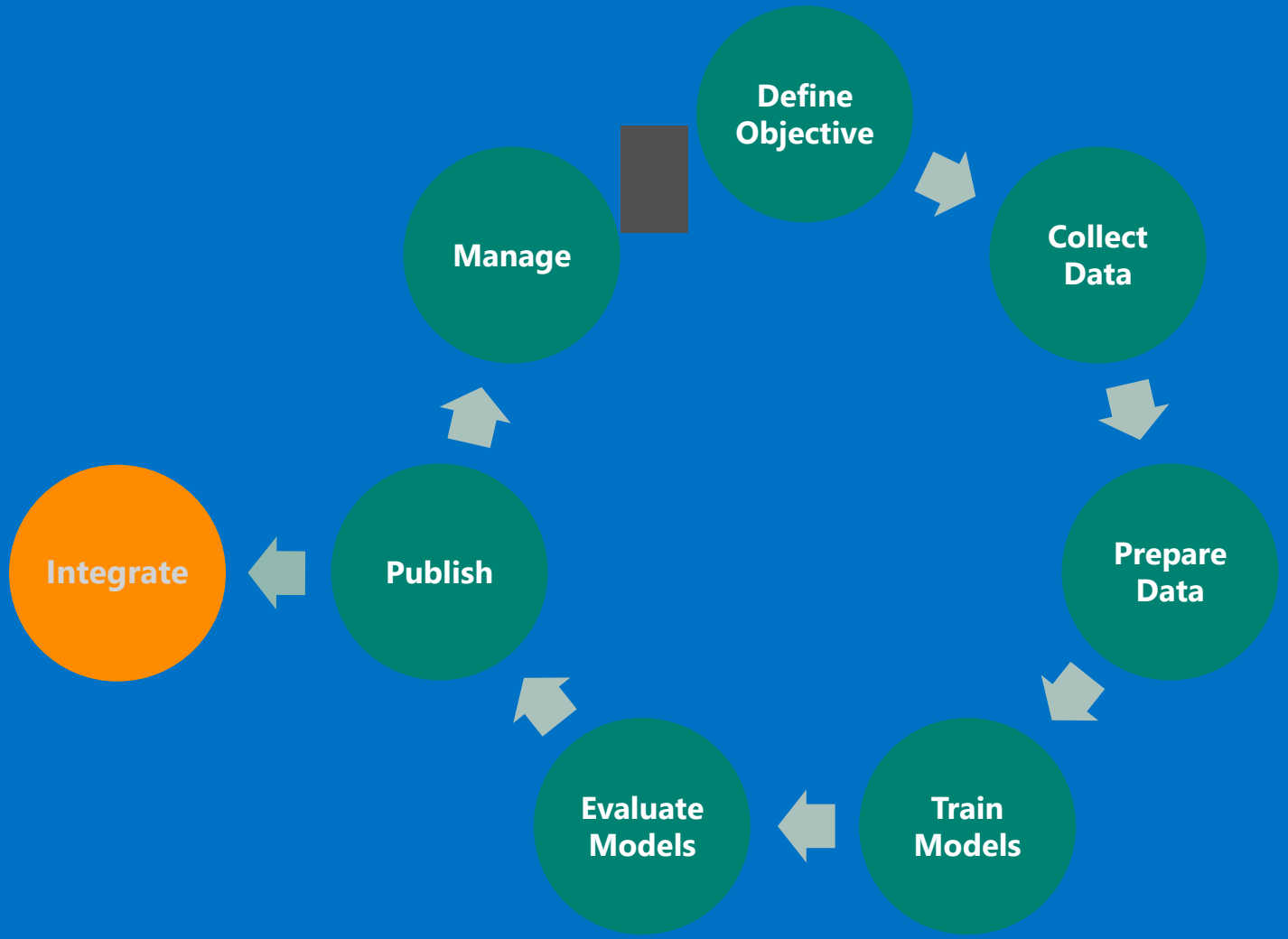
Azure Portal		Azure Ops team
ML Studio	+	Data professionals & Data scientists
ML API service		Software developers

Demo: Azure Machine Learning workspace 배포



ML 워크플로우

Machine learning 워크플로우



Machine learning 워크플로우

Define Objective

고객의 수익 예측이 필요해

...to deliver targeted display advertising on the company eCommerce web site, to:

- Present relevant product suggestions
- Increase sales and profitability

Machine learning 워크플로우

Garbage in ► Garbage out ☹

- Internal sources, i.e. operational systems, data warehouse, etc.
- External sources
- Different formats, i.e. relational, multidimensional, text, map-reduce

E.g., integrate internal data to external data like weather, or market intelligence data

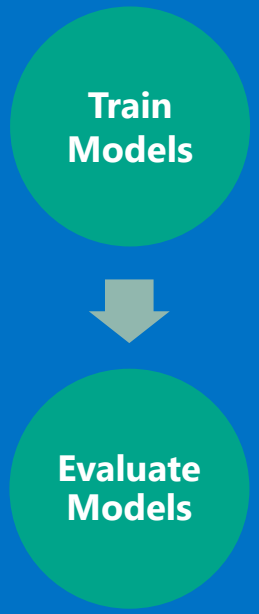
Machine learning 워크플로우

Prepare Data

- Transform to cleanse, reduce or reformat
- Isolate and flag abnormal data
- Appropriately substitute missing values
- Categorize continuous values into ranges
- Normalize continuous values between 0 and 1

When designing systems, give consideration to attributes that may be required as inputs for future modeling, e.g. demographic data: Birth date, gender, etc.

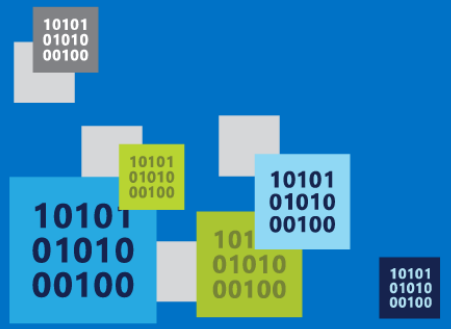
Machine learning 워크플로우



- Selecting a machine learning algorithm
- Defining inputs and outputs
- Optimizing by configuring algorithm parameters

Accuracy, Reliability, Usefulness

Demo: ML Experiment 실행

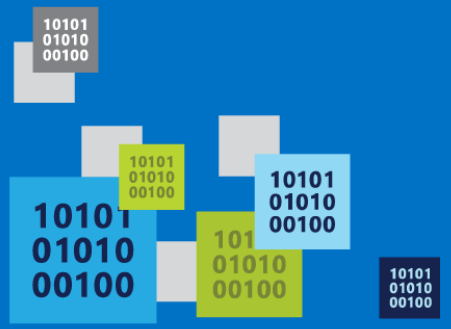


Machine learning 워크플로우

- Transformational logic is replaced with a re-usable transformation resource
- Training logic is replaced with a trained model
- Web service inputs and outputs are added
- Module properties can be parameterized

Learn from others by discovering experiments
Contribute and showcase your experiments

Demo: ML 웹서비스



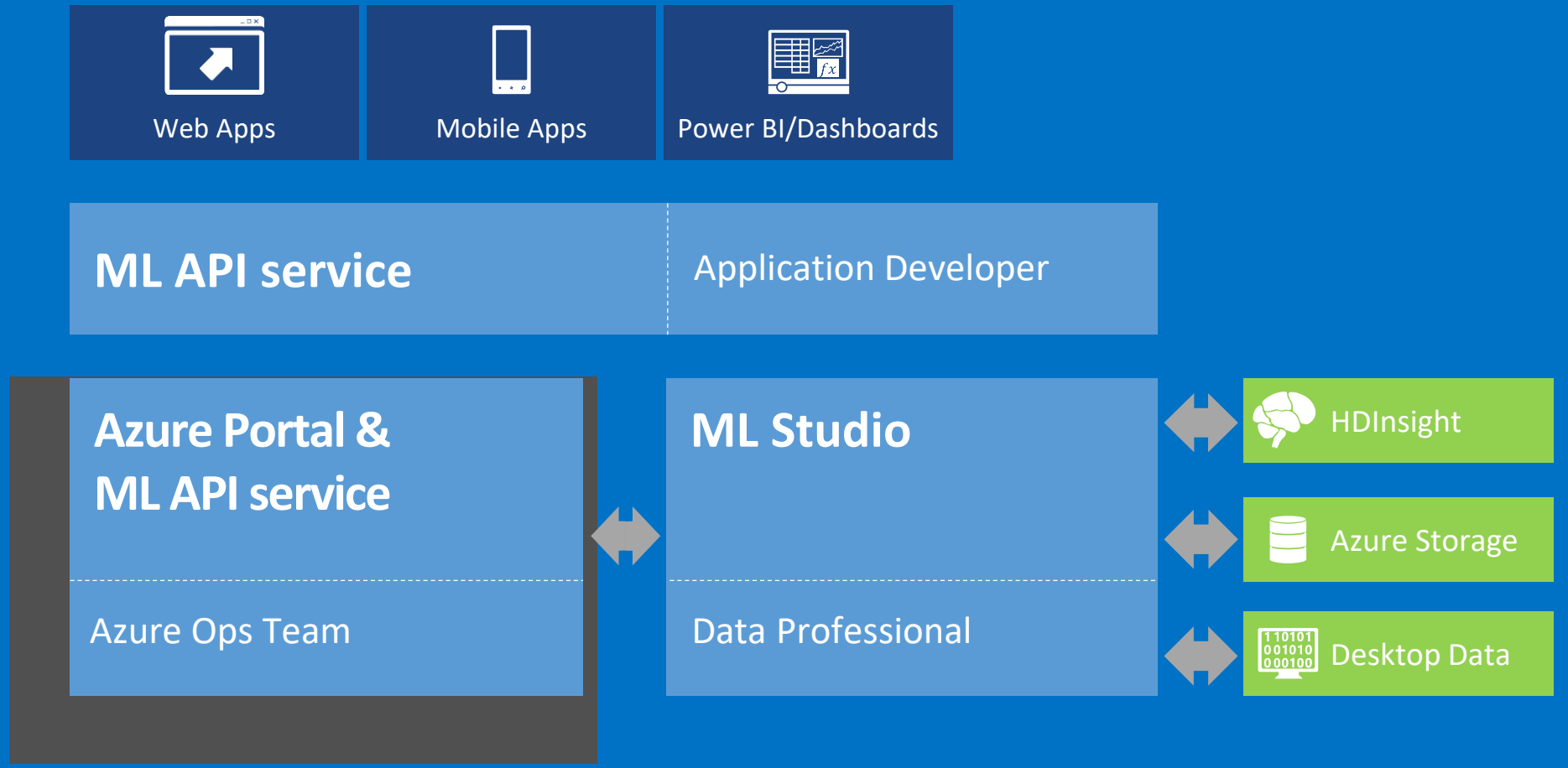
Machine learning 워크플로우

Each web service offers two methods:

- Request/Response Service (RRS) ► Low latency, highly scalable web service
- Batch Execution Service (BES) ► High volume, asynchronous scoring of many records

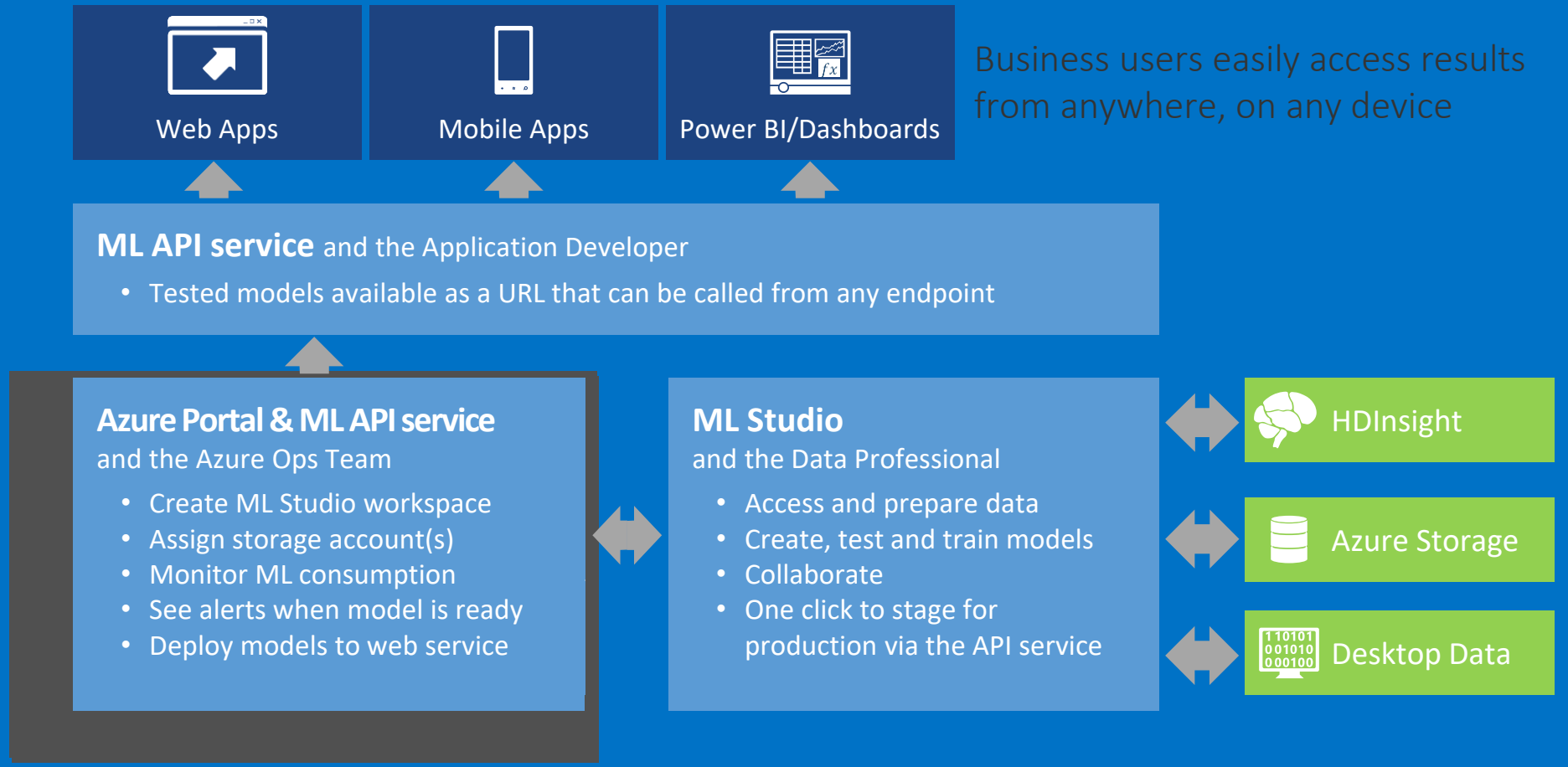
Azure Machine Learning

One solution for machine learning



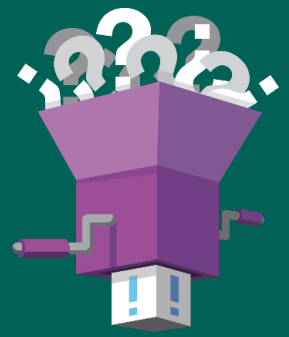
Azure Machine Learning

One solution for machine learning



Cloud-based machine learning today

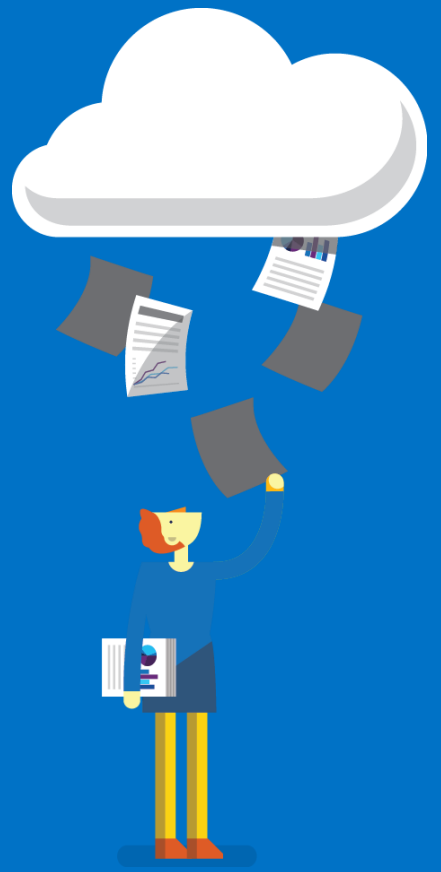
Faster towards solutions



Mashup of powerful algorithms

Elastic, pay-as-you-go model with low operative costs

Global scaling of solutions via cloud API



Quick and easy extensibility with cloud functions such as Power BI, Hadoop (Azure HDInsight) and cloud storage

Summary

Machine Learning is a subfield of computer science and statistics that deals with the construction and study of systems that can learn from data

Azure Machine Learning key attributes:

Fully managed ► No hardware or software to buy

Integrated ► Drag, drop, connect and configure

Best-in-class algorithms ► Proven solutions from Xbox and Bing

R built in ► Use over 400 R packages, or bring your own R or Python code

Deploy in minutes ► Operationalize with a click

Machine Learning is now approachable to developers

Resources

http://azure.microsoft.com/en-us/services/machine-learning

http://azure.microsoft.com/en-us/documentation/services/machine-learning

http://azure.microsoft.com/en-us/documentation/articles/machine-learning-faq

http://azure.microsoft.com/en-us/pricing/details/machine-learning/

Note: The Free tier does not require an Azure subscription or a credit card

Resources

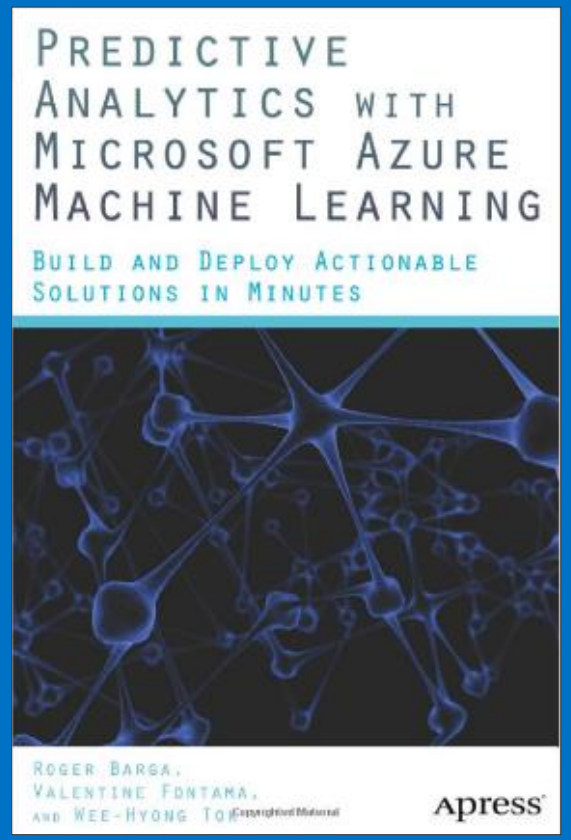
https://gallery.azureml.net

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

http://www.revolutionanalytics.com

Resources

Publisher: Apress
Authors: Roger Barga, Valentine Fontama, and Wee Hyong Tok



http://en.wikipedia.org/wiki/Paul_the_Octopus

Q&A