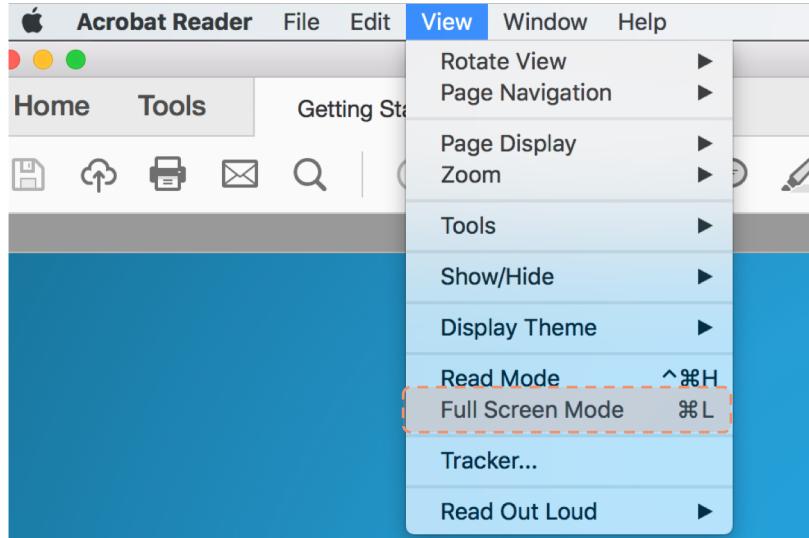


Getting Started with
Machine Learning

Jim Liang | Sep, 2018 Ver. 0.94

:: 友情提示

请在全屏状态下阅读此文档



原始文档是powerpoint制作并转成PDF文件
所以为了更好地阅读，请切换到全屏模式

如何在本文档里快速导航定位？

- Nearest Neighbor [↴](#)
- Support Vector Machines [↴](#)
- Linear Regression [↴](#) 点击此图标
- Logistic Regression [↴](#)
- Neural Network [↴](#)
- Gradient Descent [↴](#)
- Naïve Bayes [↴](#)
- K-means [↴](#)
- PCA [↴](#)
- Decision Trees [↴](#)
- AdaBoost [↴](#)
- Random Forest [↴](#)

Table of contents | ↪

:: Table of contents

Part 1: The Fundamental concept

- Overview [¶](#)
- Business Understanding [¶](#)
- Data Understanding [¶](#)
- Data Preparation [¶](#)
- Modelling [¶](#)
- Model Evaluation [¶](#)
- Model Deployment [¶](#)

Part 2: Well-known algorithms [¶](#)

- Nearest Neighbor [¶](#)
- Support Vector Machines [¶](#)
- Linear Regression [¶](#)
- Logistic Regression [¶](#)
- Neural Network -1 [¶](#)
- Gradient Descent [¶](#)
- Neural Network - 2 [¶](#)
- Naïve Bayes [¶](#)
- K-means [¶](#)
- PCA [¶](#)
- Decision Trees [¶](#)
- AdaBoost [¶](#)
- Gradient Boosting [¶](#)
- Random Forest [¶](#)

Part 3: Other topics [¶](#)

- Advice for designing & building machine learning system [¶](#)
- Large scale machine learning [¶](#)
- What to do when there is no enough data ? [¶](#)
- Popular machine learning frameworks [¶](#)

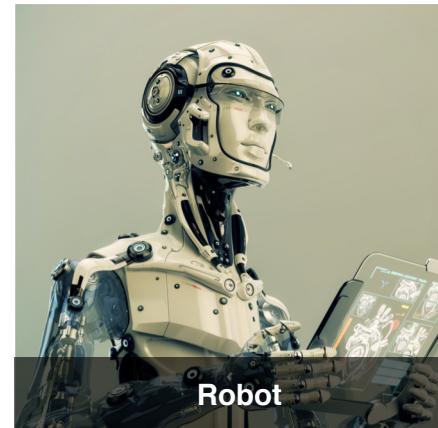
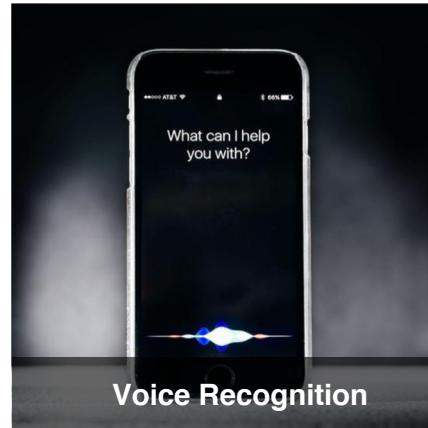
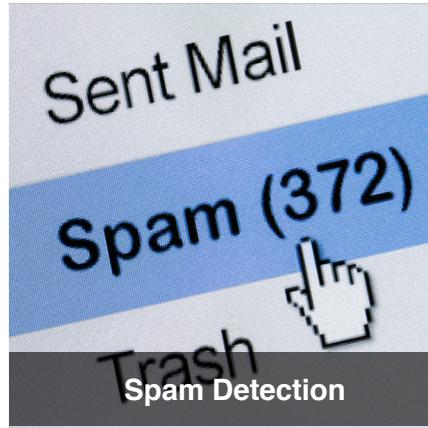
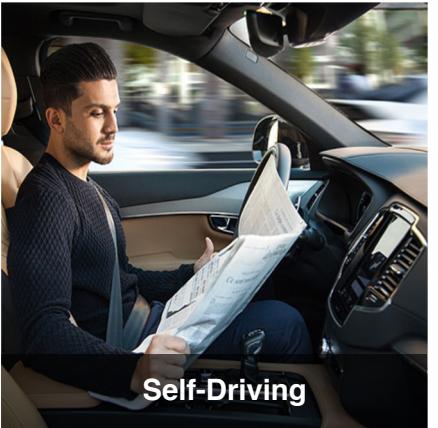


Part 1

The fundamentals
of machine learning

01 | overview

:: Machine Learning – all around us



:: Computer vision is surpassing human abilities

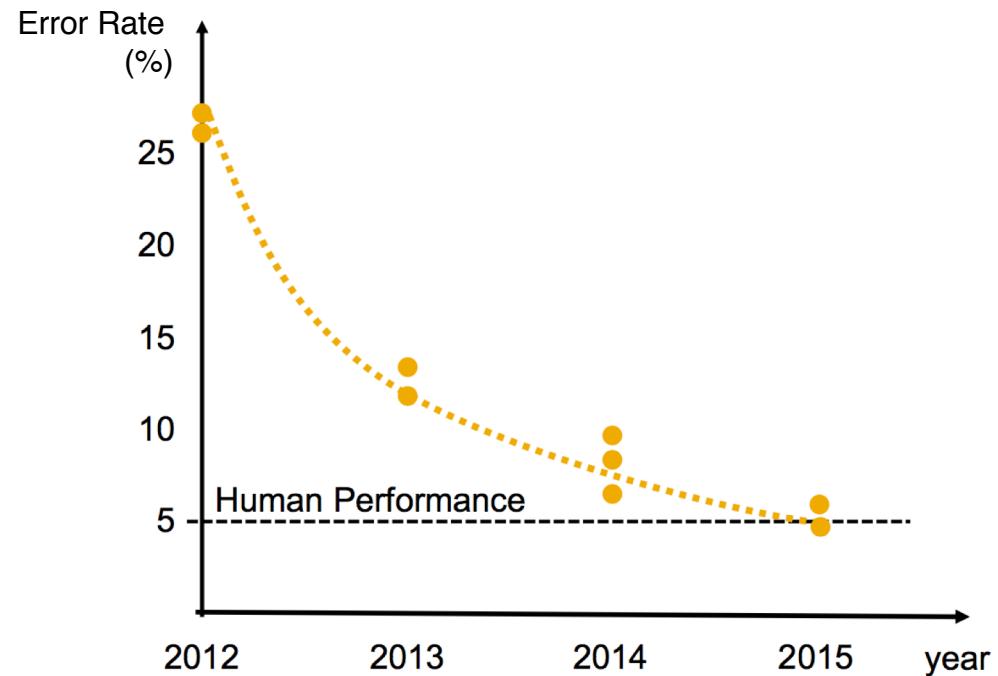


- Chair
- Dining Table
- Person



- Dog
- Person
- Leaves

ImageNet Performance

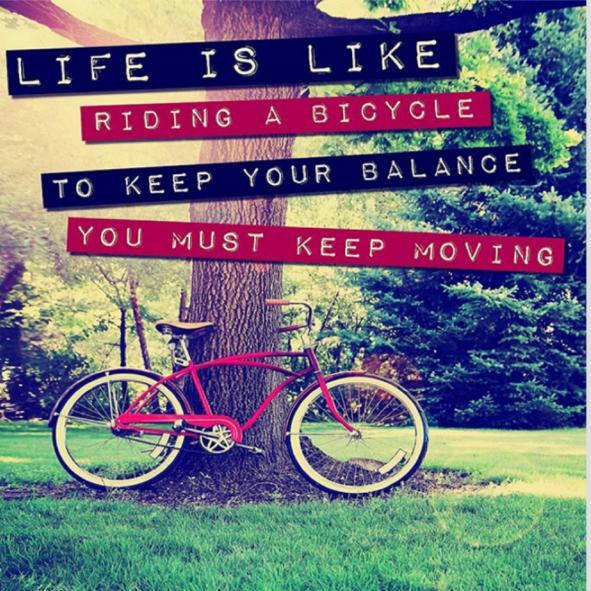


:: Computer vision - See it in action

Cool demos from Microsoft

There are couple of demos which show the capabilities of computer vision. For example, *Analyze an image*, *Read text in images*, *Read handwritten text from images*

Read text in images



The demo interface includes:

- A preview window showing the original image and the extracted text.
- Two tabs: "Preview" (selected) and "JSON".
- Input fields for "Image URL" and "Browse".
- A "Submit" button.

Text extracted from the image:

LIFE IS LIKE
RIDING A BICYCLE
TO KEEP YOUR BALANCE
YOU MUST KEEP MOVING

You can try these demos by yourself. Just upload your image and see what will happen. It's nice!

<https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/>

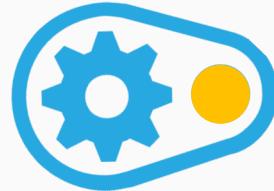
:: Future of knowledge work

Another Industrial revolution

\$ 4bn

Enterprise machine learning projected to grow at 56% to \$4bn in 2020

Automate knowledge work



Machine Learning will enhance or automate repetitive knowledge work

Do the impossible



New application and business models

:: Use predictive analytics to solve a variety of business challenges



SALES & MARKETING

- Churn Reduction
- Customer Acquisition
- Lead Scoring
- Product Recommendation
- Campaign Optimization
- Customer Segmentation
- Next Best Offer/Action
- Predictive Maintenance
- Load Forecasting
- Inventory/Demand Optimization
- Product Recommendation
- Price Optimization
- Manufacturing Process Optimization
- Quality Management
- Yield Management



OPERATIONS



FRAUD & RISK

- Fraud and Abuse Detection
- Claims Analysis
- Collection and Delinquency
- Credit Scoring
- Operation Risk Modelling
- Crime Threat
- Revenue and Loss Analysis



FINANACE & HR

- Cash Flow and Forecasting
- Budgeting Simulation
- Profitability and Margin Analysis
- Financial Risk Modeling
- Employee Retention Modeling
- Succession Planning
- Life Sciences
- Healthcare
- Media
- Higher Education
- Public Sector/Social Sciences
- Construction and Mining
- Travel and Hospitality
- Big Data and IoT

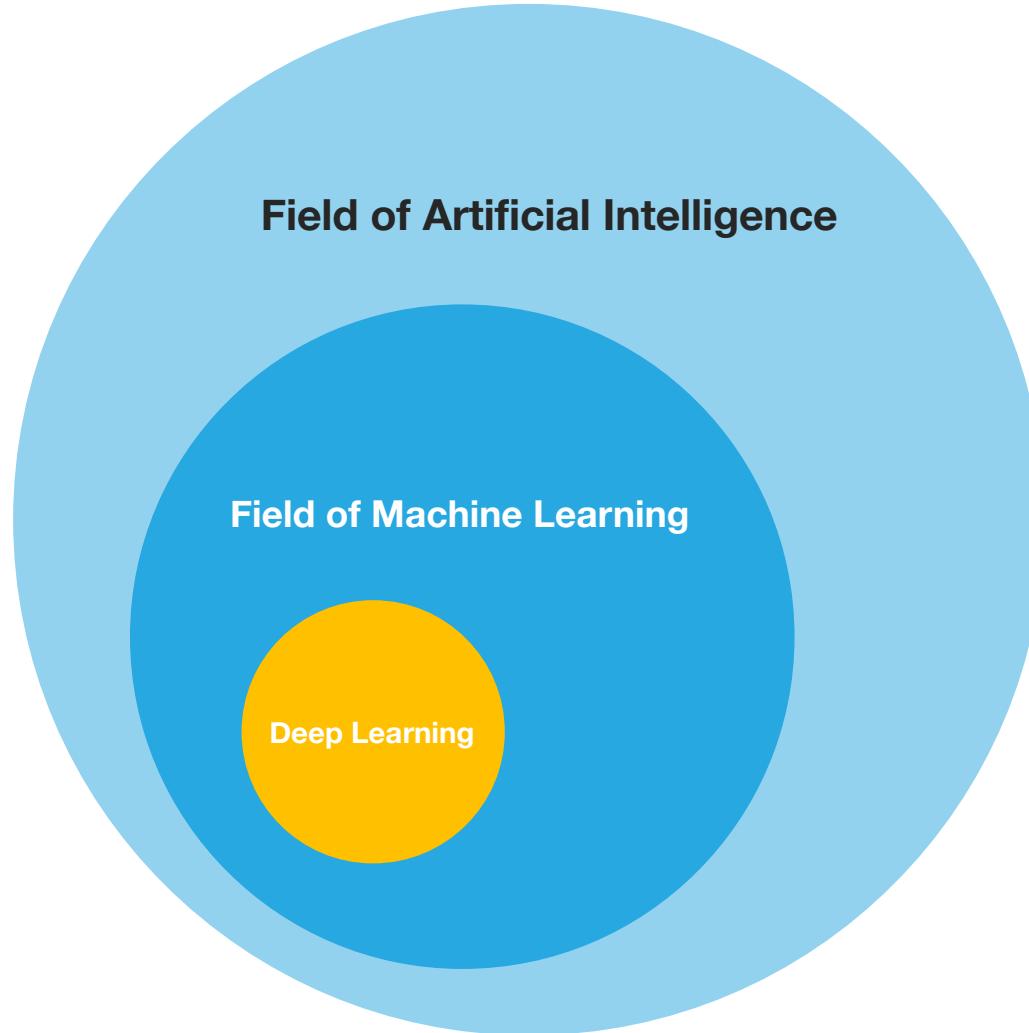


OTHER SECTORS

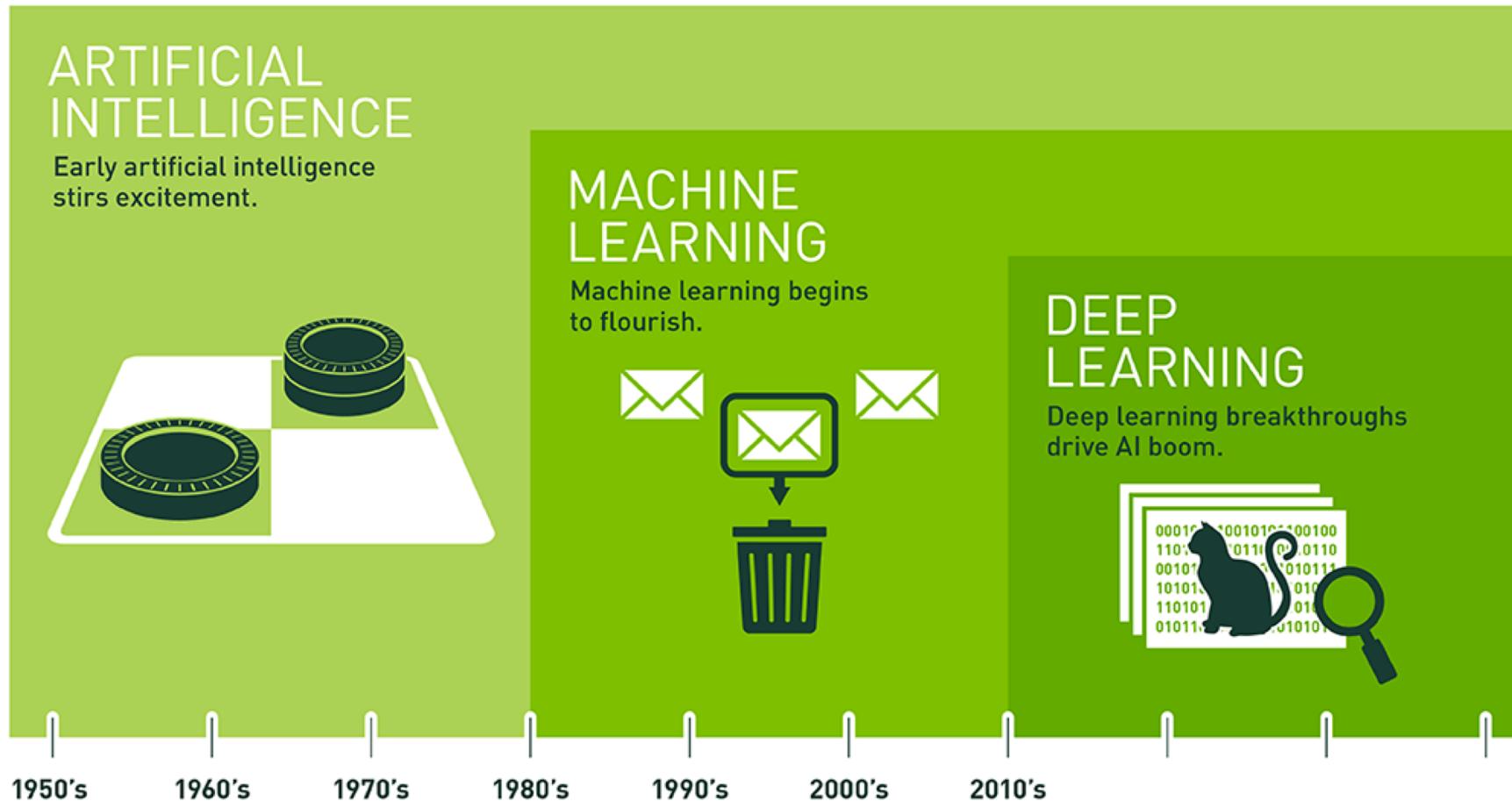
:: Big Pool of Machine Learning Use Cases

Social Media Signal Discovery	Intelligent Fraud Detection & Management	Retention Risk Analysis	Career Path Recommender	Price Optimization	HCP Predictive Services
Solution Recommender	Deal Scoring	Text Analytics		Dynamic Pricing	
Intelligent Financing for Ariba Pay	SAP Clea for Cash Application	Customer Retention Insights	Business Forecasting	Automated Product Safety Classifications	Workforce Planning Recommender
Brand Monitoring	Receivables Intelligence	Predictive Forecasting	Invoice to Record	SAP Clea for Résumé Matching	Imaging Intelligence for Retail Execution (Smart Store)
Social Media Customer Service	Product Classification Suggestions	Guided Discounting	Self-Service Conversational Interface	Analyze User Interaction	Software Security Analysis
Predictive Analytics Integrator	Payment Risk	Learning Recommendations		Streaming Machine Learning	Marketing Efficiency
From Service to Sales	Predictive Analytics	Intercompany Reconciliation	Business Optimizations	Automatic Creation of a Semantic Hierarchy	Non-Expert Machine Learning
Best Contact Time	SAP Clea for Brand Intelligence	Lead Recommender	Job Matching	Payables Intelligence	IoT Machine Learning Services
Predictive Lead and Opportunity Scoring	Sales Assistant App	Spatial & Graph Machine Learning	SAP Clea for Service Ticket Intelligence		Intelligent Financing for Ariba Pay
Intelligent Self Service Bot	Machine Learning for Intelligent Services	Product Master Data Matching	Predictive Machine Maintenance		Predictive Modelling
	Machine Learning for Vehicle Insights	Job Posting Sentiment Analysis			

:: About AI , Machine Learning and Deep Learning

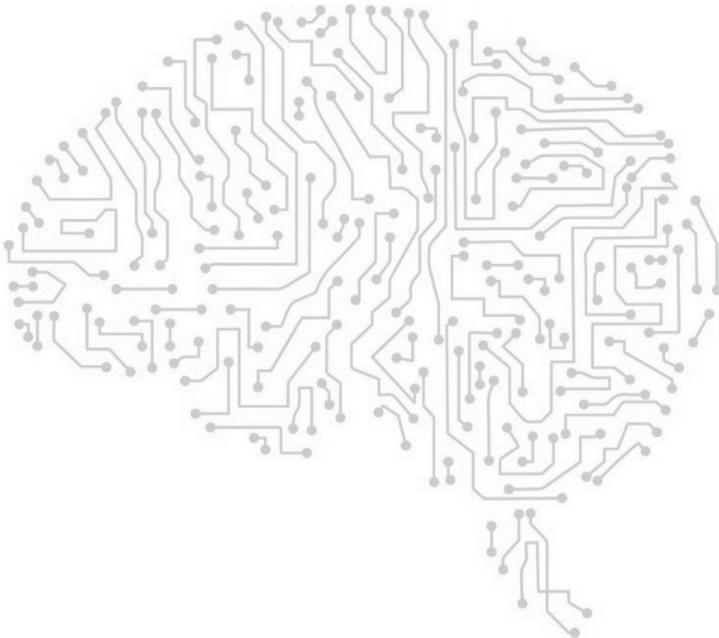


:: Artificial Intelligence has been evolving



Since an early flush optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions

:: Machines can now do things that were not possible before



Why now ?

- **Big Data**
- **Massive Computer power**
(e.g. GPU, multicore)
- **Better algorithms**
(e.g. Deep Learning algorithms)

:: What's machine learning ?

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T , as measured by P , improves with experience E .

- Mitchell, T. (1997). *Machine Learning*, McGraw Hill



:: More friendly definition of Machine Learning

*Machine Learning is the science of getting
computers to act without being
explicitly programmed*

- Andrew Ng (吴恩达)

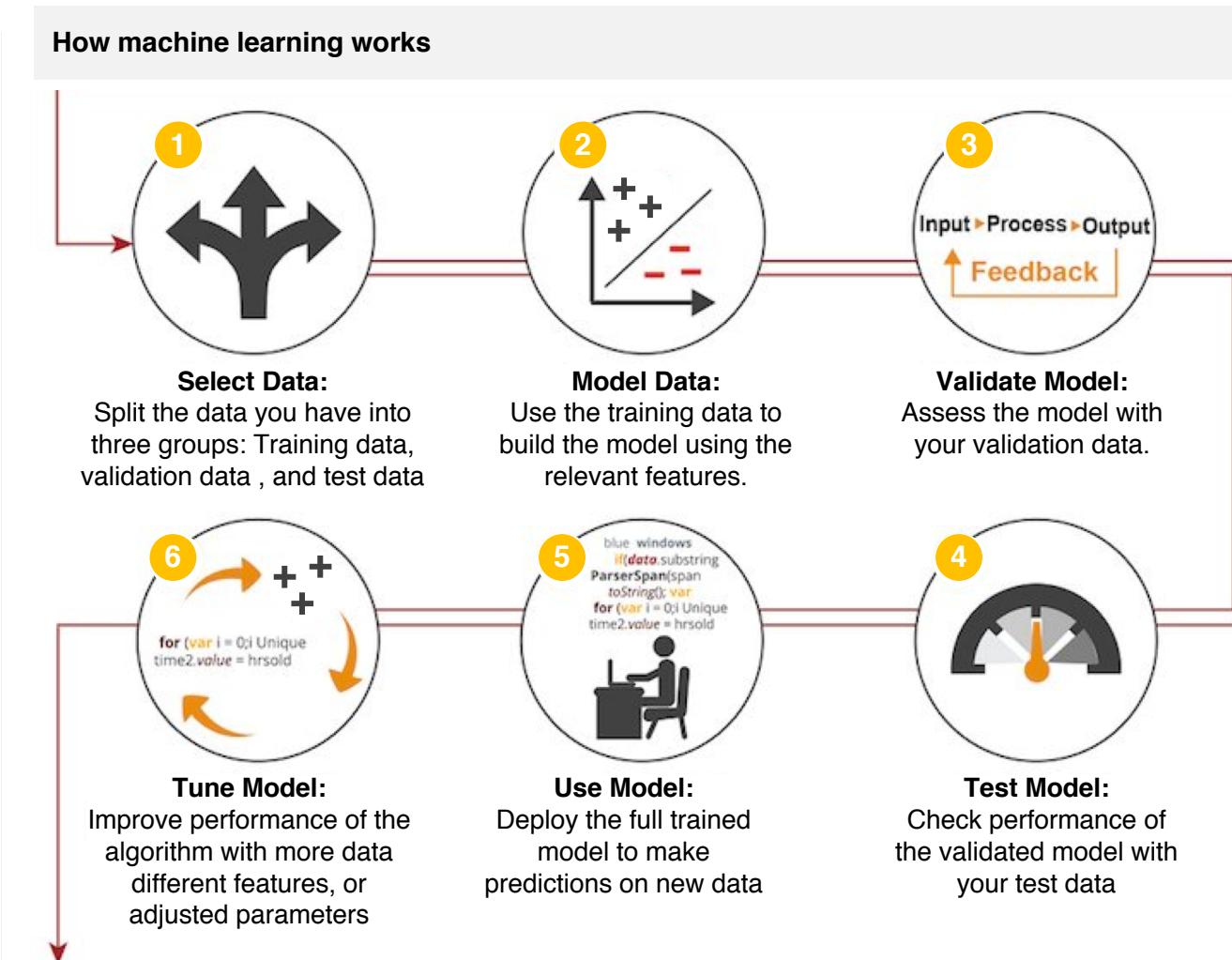
*Machine learning is a technique of data science that
helps computers **learn from existing data** in order to
forecast future behaviors, outcomes, and trends*

- Microsoft

:: Machines Learning uses historical data to make predictions



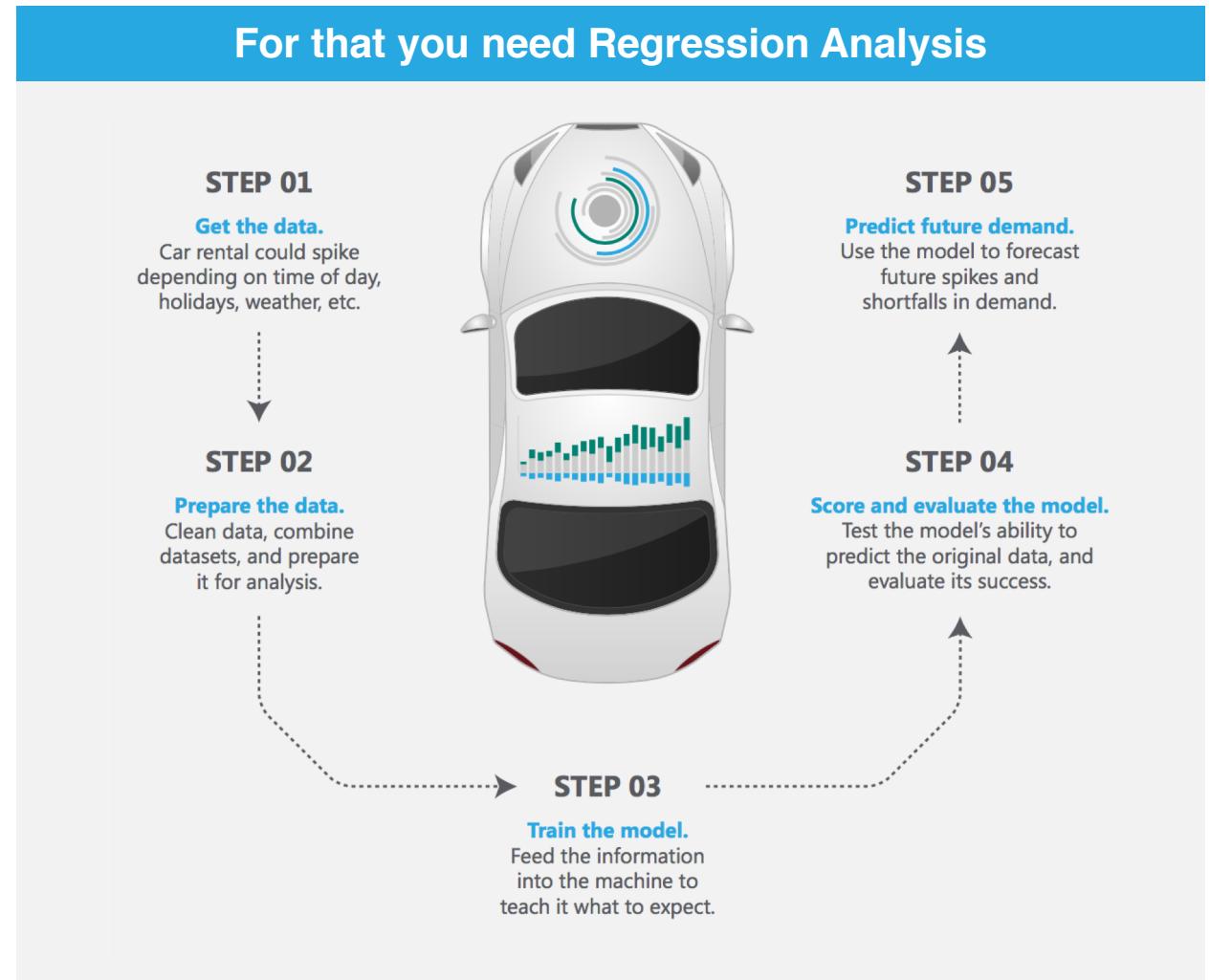
:: How machine learning works



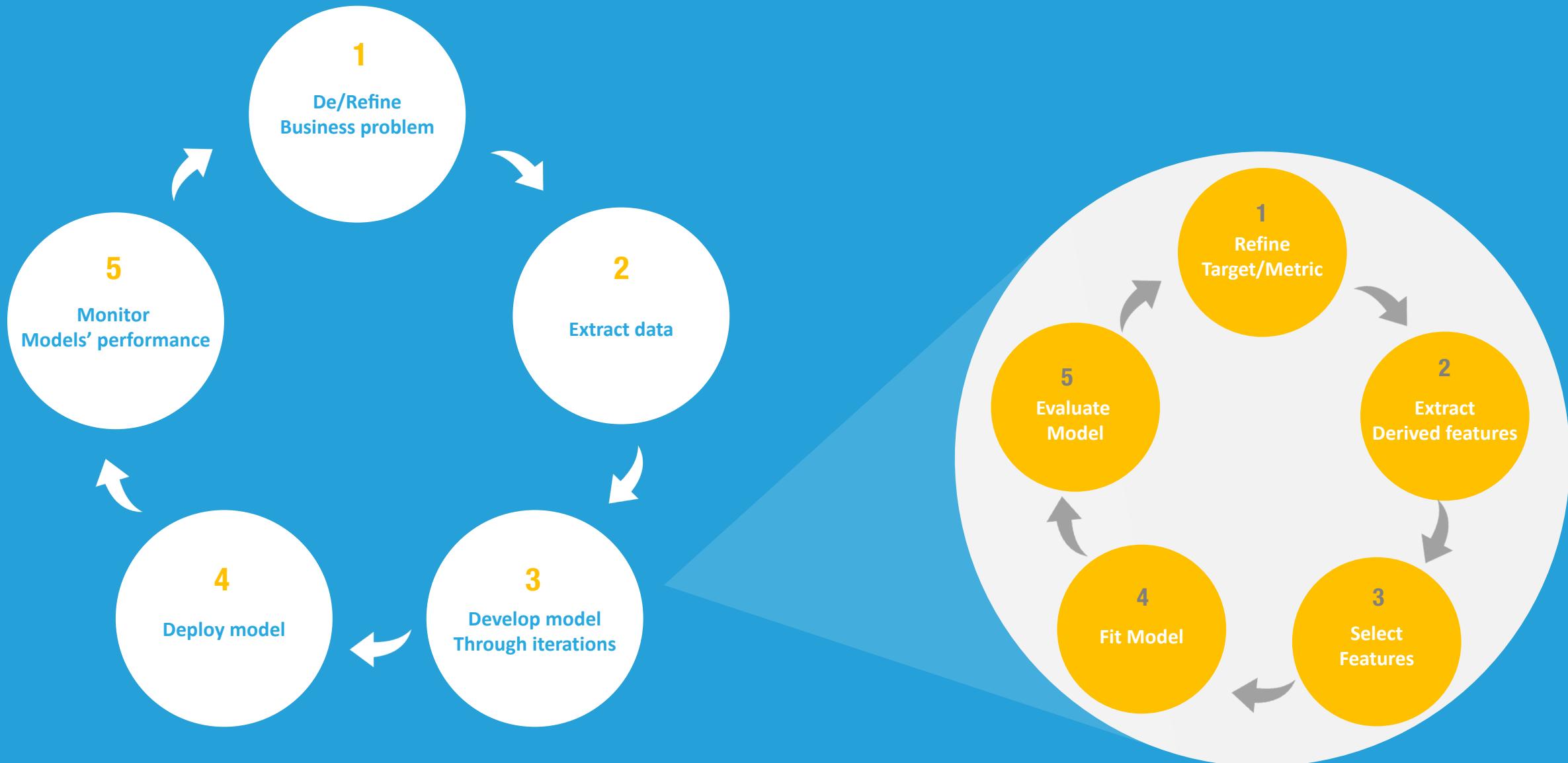
:: Example of machine learning task

Let's say you rent cars

How can you accurately predict demand for your product ?

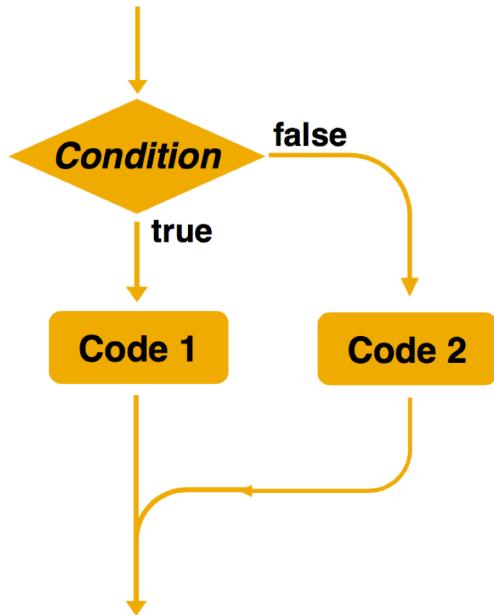


:: Steps to build a machine learning solution



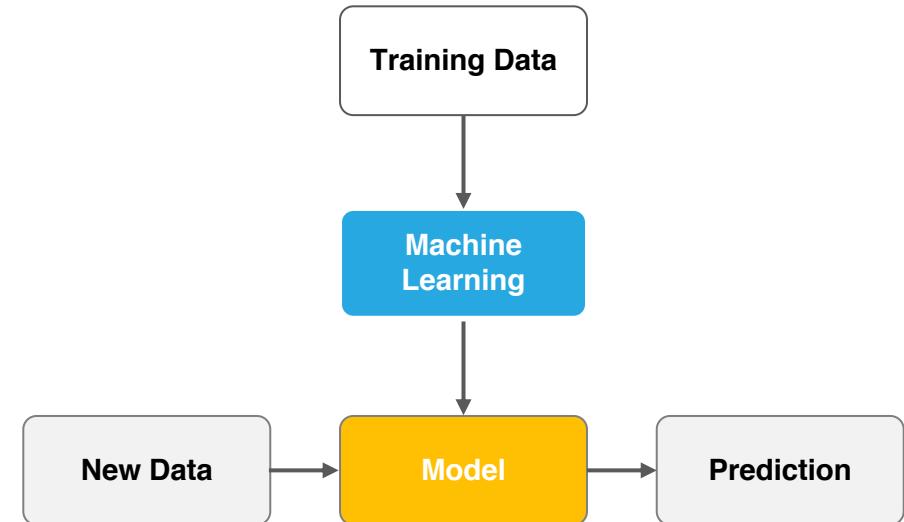
:: The difference between traditional approach and Machine Learning

Rule-based approach



- Explicitly programmed to solve problem
- Decision rules are clearly defined by humans

Machine learning



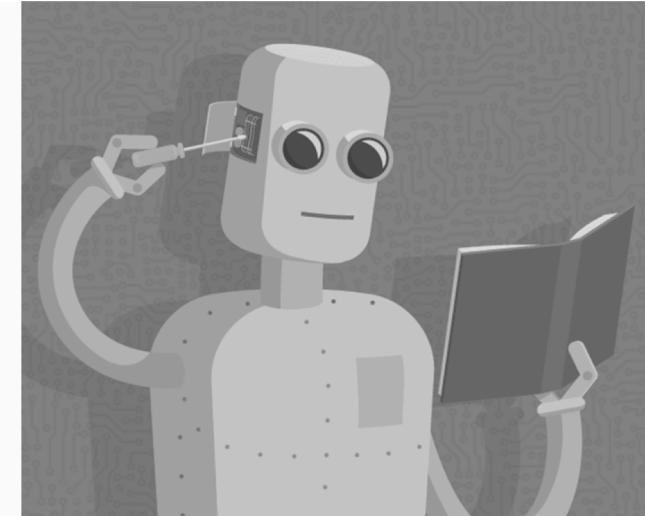
- Trained from examples
- Decision rules complex or fuzzy
- Rules are not defined by humans but learned by the machine from data

:: Wrap up

Machines Learning uses historical data to make predictions

Machine Learning is also similar to **Data Mining**, but whereas data mining is the science of discovering unknown patterns and relationships in data, ML applies previously inferred knowledge to **new data** to make decisions in real-life applications.

- Computers approximate complex functions from historical data
- Rules are not explicitly programmed but learned from data



:: When should you use Machine Learning ?

From business problem to Machine learning problem: a Recipe

Step-by-step “recipe” for qualifying a business problem as a machine learning problem

- 1. Do you need machine learning?**
- 2. Can you formulate your problem clearly?**
- 3. Do you have sufficient examples?**
- 4. Does your problem have a regular pattern?**
- 5. Can you find meaningful representations of your data?**
- 6. How do you define success?**



:: From business problem to Machine learning problem: a Recipe

When to use machine learning ?

1

Do you need machine learning?

- Do you need to automate the task?
- High volume tasks with complex rules and unstructured data are good candidates

Example: sentiment analysis

- High volume of reviews on the Web
- Unstructured text
- Human language is complex and ambiguous

