

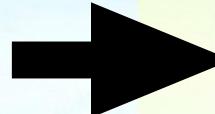


Business Process Automation with VBA and Python

Mr. Eddie Chow / 8 March 2025



Table Of Contents



Introduction to business process automation

Business Process automation with VBA

Business process automation with Python

Introduction to project management for business process automation

Development and implementation of business process automation

Final Group Presentation



Lecturer Biography

<https://www.linkedin.com/in/eddie-chow-a3860464/>



- Founder of InnoVi, Co-founder of 4 companies, covering from property, retail, education and ESG
- 19+ Years' programming experience, including 7 Years' IT Management Experience
- 7 years' teaching experience in tertiary education
- 10+ years' coaching or mentoring experience on large project development..
- Development experience on blockchain, computer vision, machine learning/deep learning
- Researcher in Cutting-edge AI Research, particular in area in Computer vision
- AI Tech Speaker, AI Mentor
- Personal Website : eddiecityu.github.io



Assessment

- **Continuous Assessment**

In-class discussions and computer assignments which include the usage of different applications & methodologies

- **Final Assessment**

20-mins (3 persons) group project presentation which integrate knowledge in this module, apply different AI methodologies and software to business applications of process automation



What is Process Automation, RPA, IPA?

Overview of the technological building blocks related to process automation

Contemporary tools for process automation

Challenges and opportunities of business process automation

Business implications of process automation

Intended Learning Outcomes

1. describe the contemporary trends of process automation and explain the opportunities and challenges of business process automation;
2. outline key steps in process automation project management and illustrate the significance of each step;
3. apply computational tools to implement process automation;
4. discuss the development of process automation and practical cases for business.

Assessment

Memory Test?

- NO Midterm / Final exams

Essays?

- NO research paper
- NO project report

Practical Work!

- 2 in-class coding exercises
- 1 Group Project



Taking
Exams

Building
Solutions

Lecture 1 - Business Process Automation with VBA and Python



Business Case Studies: Artificial Intelligence

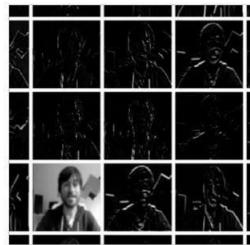
Artificial Intelligence

Common Use Cases:

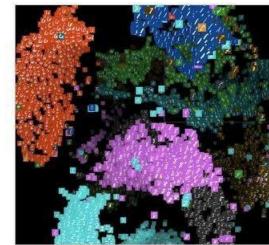
Object Detection
Speech Recognition
Natural Language Processing
Translation between languages
Creativity - Style Transfer
Art Restoration



TEACHABLE MACHINE



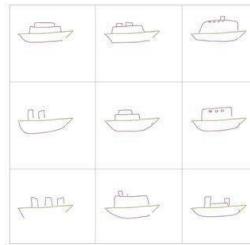
WHAT NEURAL NETWORKS SEE



VISUALIZING HIGH-DIMENSIONAL SPACE



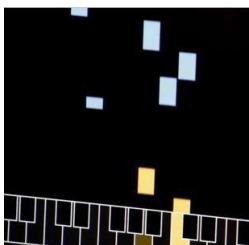
QUICK, DRAW!



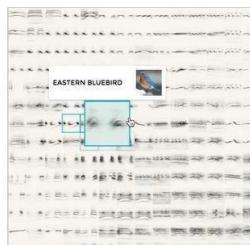
SKETCH-RNN DEMOS



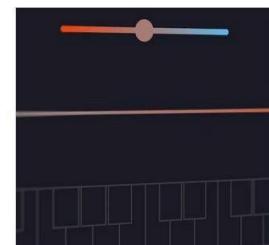
HANDWRITING WITH A NEURAL NET



AI DUET



BIRD SOUNDS



NSYNTH: SOUND MAKER

Machine Learning Business Use Case

Face Recognition Algorithm



Machine Learning Business Use Case

Hotel Recommendation Engine



A photograph showing the Boston Harbor Hotel at night, illuminated with warm lights. The hotel's distinctive arched entrance is visible, along with the surrounding city skyline across the water.

Heart icon Viewed

Boston Harbor Hotel

★★★★★

Downtown Boston

Come Experience our 5 Star Luxury Hotel.

Iconic Waterfront Hotel with a Convenient Location to Everything Boston Has to Offer. Every Room Boasts City or Harbor Views.

1-866-286-0843 • Expedia Rate

4.8/5 Exceptional!
(1,132 reviews)

\$693-\$581

nightly price
Sponsored

Get member price

People who looked at the Boston Harbor Hotel also viewed these:



Four Seasons Hotel Boston

★★★★★

\$645



The Ritz-Carlton, Boston

★★★★★

\$595

Shifting mindsets of customers: From Products To Relationships



Products
1970s

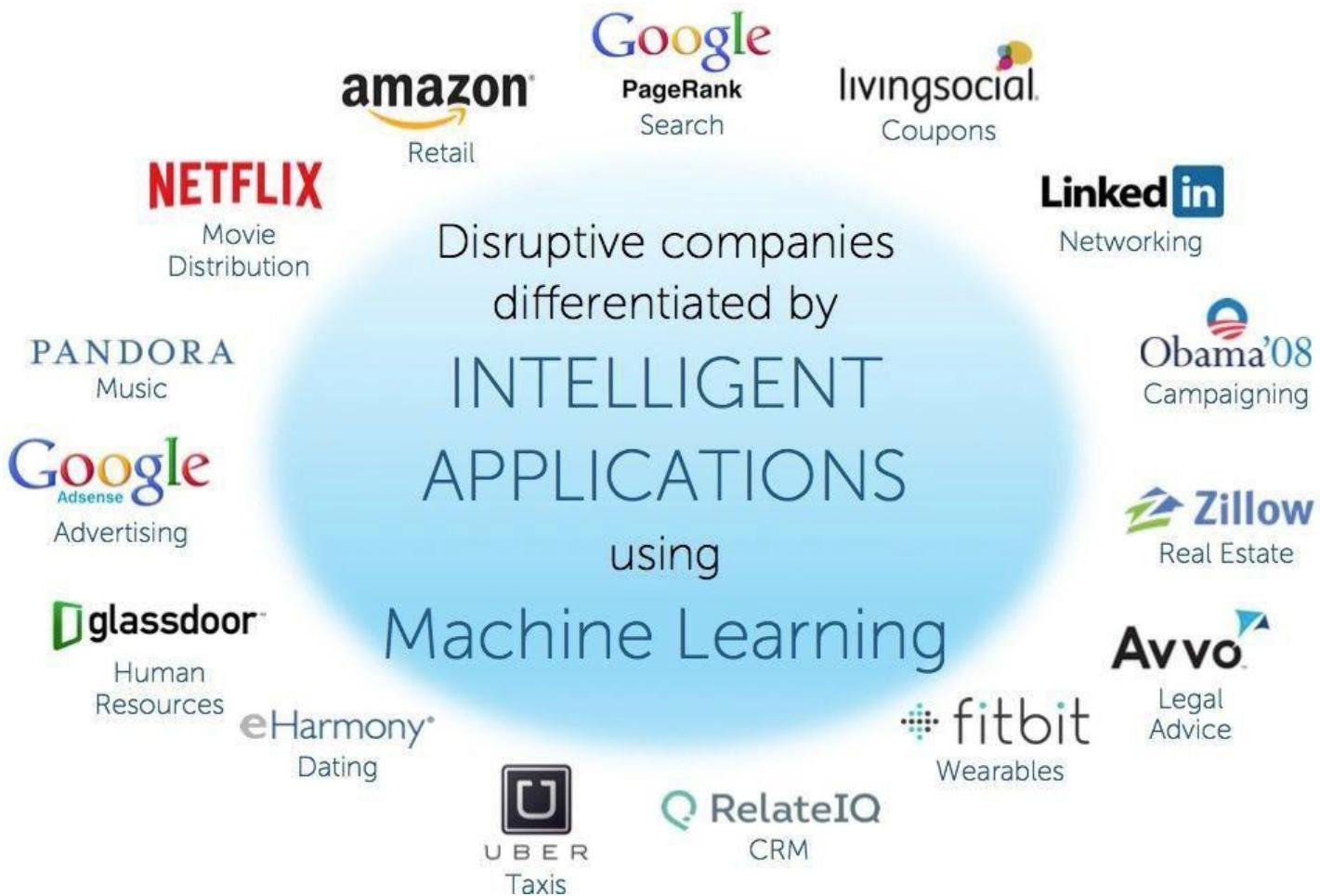


Customer
Centric



Relationship
Centric Today

The new disruptors



These companies applied analytics to stay 1 step ahead of the competition



AI Features in Retail

- Chatbots, Virtual Assistants
- Facial Recognition
- Demographic Segmentation
- Facial Personality Analytics
- AI-enabled Customer Analytics
- Real-time insights
- Personalization
- Inventory Management
- Visual Recognition
- Customer Purchase Prediction
- Trend Prediction
- Purchasing Recommendation
- Ideal Price Point Recommendation
- Predict the trends in Fashion, e-commerce, etc
- Predictive Analysis
- Predictive Personalization

NLP in real life

- Machine Translation
 - Google Translate translates language from one language to another.
- Text Simplification
 - Rewordify simplifies the meaning of sentences.
- Sentiment Analysis
 - Hater News gives us the sentiment of the user.
- Text Summarization
 - Smmry or Reddit's autotldr gives a summary of sentences.
- Auto-Predict
 - Google Search predicts user search results.
- Auto-Correct
 - Google Keyboard and Grammarly correct words otherwise spelled wrong.
- Speech Recognition
 - Google WebSpeech or Vocalware.
- Question Answering
 - IBM Watson's answers to a query.
- Natural Language Generation
 - Generation of text from image or video data

Cases : Facial recognition for customer analytics



Face Detection

Reliably detect human faces in a photo, along with their coordinates and sizes.



Emotions, Gender, Age

Recognize people's facial expressions and detects primary and secondary emotions. Also detects age and gender

<https://findface.pro/en/solutions/face-recognition-for-customer-analytics.html>

Introduction to Business Process Automation

What is Business Automation?

From ChatGPT...

JA

What is business automation? asking for a friend :)

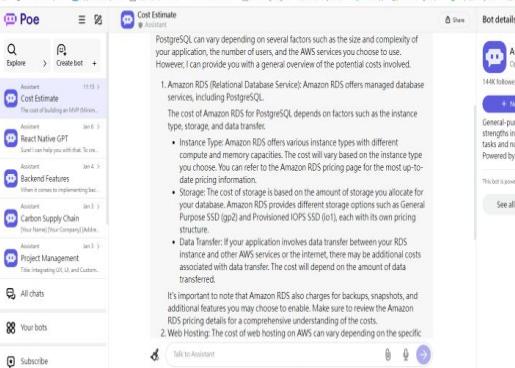


Business automation refers to the process of using technology to automate repetitive or routine tasks and processes within a business, such as data entry, invoicing, and inventory management. It can also involve the integration of various software applications and systems to streamline workflows and improve overall efficiency.

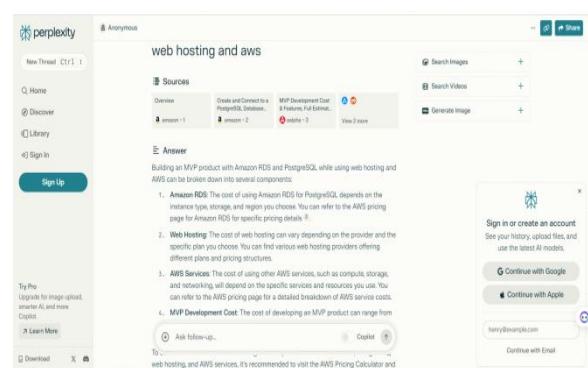
Step 1 : Tools Introduction – ChatGPT

- Poe (<https://poe.com/>) - Poe 讓您提出問題、立即獲得答案，並與人工智慧進行來回對話。允許存取 GPT-4、gpt-3.5-turbo、Anthropic 的 Claude 以及各種其他機器人
- Perplexity AI (<https://www.perplexity.ai/>) – 由人工智慧驅動的搜尋引擎和聊天機器人，利用自然語言處理(NLP)和機器學習等先進技術為用戶查詢提供準確、全面的答案。
- Chatgpt (<https://www.opera.com/features/browser-ai>) - Opera 瀏覽器內建最受歡迎的AI 聊天服務，可在瀏覽時一鍵存取。無需在網路上搜尋資訊即可立即獲得任何問題的答案，並在ChatGPT 的幫助下輕鬆生成內容。

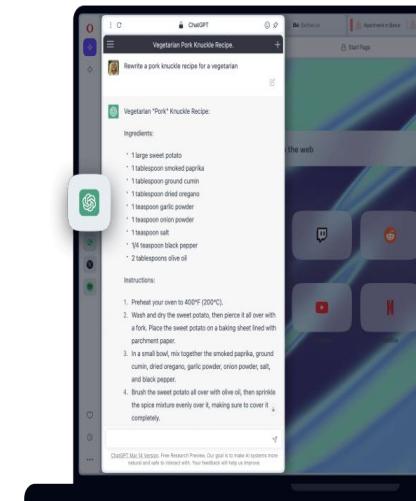
Poe



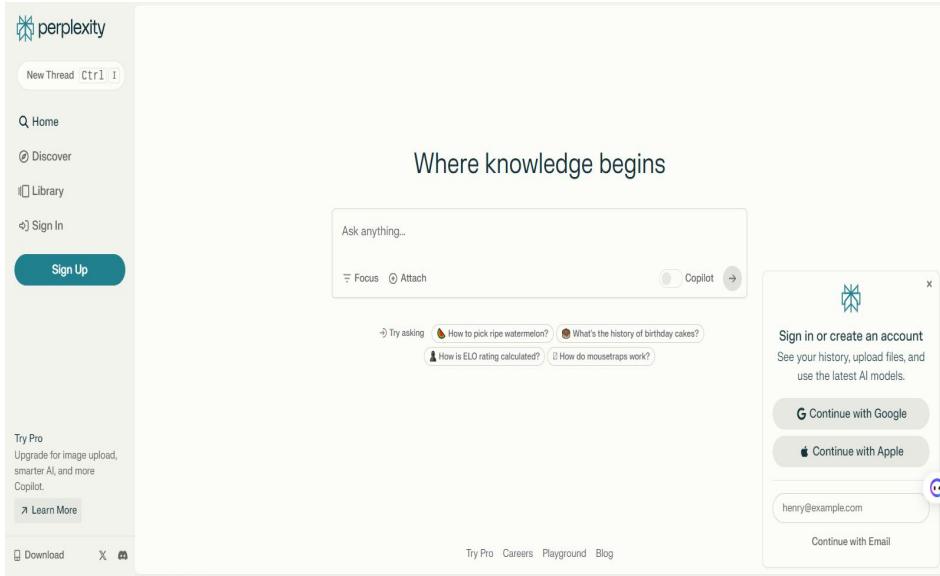
Perplexity



Chatgpt

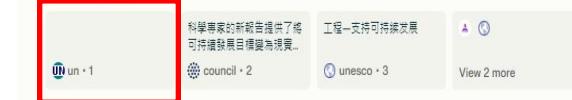


Step 1 : Tools Introduction – Perplexity



編寫生成可持續發展項目的技術提案

Sources



Answer

生成可持續發展項目的技術提案需要考慮以下幾個方面：

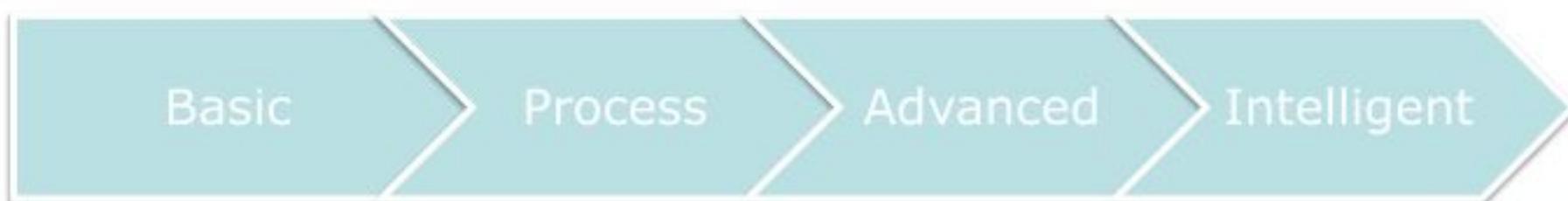
1. 目標設定：確定項目的目標，以便實現可持續發展。這可以包括減少碳排放、提高能源效率、減少廢棄物產生等。
2. 技術選擇：選擇適合項目目標的技術，例如新能源汽車、可再生能源、智能建築等。
3. 數據收集和分析：收集和分析相關數據，以評估項目的進展和效果。這可以包括碳排放量、能源消耗、廢棄物產生等。
4. 政策支持：了解政策環境，以便在項目實施過程中應對政策變化。這可以包括稅收政策、環保政策等。
5. 培訓和技能發展：提供培訓和技能發展機會，以便項目團隊能夠應對項目挑戰。這可以包括技術培訓、項目管理培訓等。
6. 環境可持續性：在決策中納入環境因素，以便實現可持續發展。這可以包括碳排放量、能

- 輸入提示“編寫生成可持續發展項目的技術提案”
- 或輸入提示“撰寫一份商業提案，用於利用人工智慧、區塊鏈和碳市場生成可持續發展技術項目。”

Introduction to Business Process Automation

Types of Business Automation

1. Basic Automation
2. Process Automation
3. Advanced Automation
4. Intelligent Automation



1. *The Basics of Business Automation.* (2021, May 12). The Basics of Business Automation | IBM.
<https://www.ibm.com/cloud/blog/basics-of-business-automation>

Introduction to Business Process Automation

Types of Business Automation – Basic Automation

- **Target →** Simple / Repetitive tasks
- **Why**
 - Eliminate human errors
 - Accelerate pace of transactional work
- **Example →** Robotic Process Automation (RPA)*



* To be covered in the next section

Introduction to Business Process Automation

Types of Business Automation – Process Automation

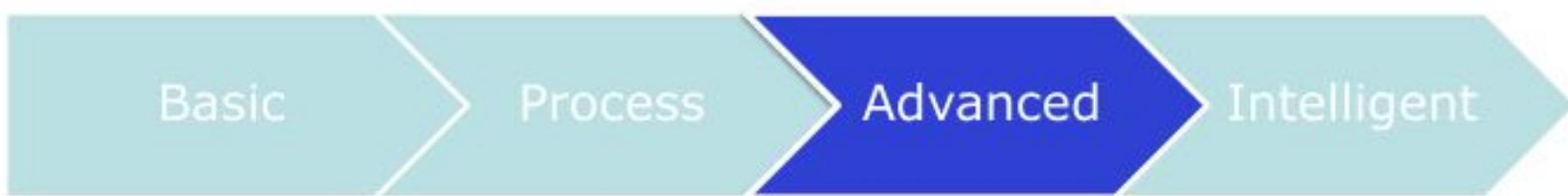
- **Target →** Process requiring dedicated solutions
- **Why**
 - Increase productivity and efficiency
- **Example →** Process Mining, workflow automation



Introduction to Business Process Automation

Types of Business Automation – Advanced Automation

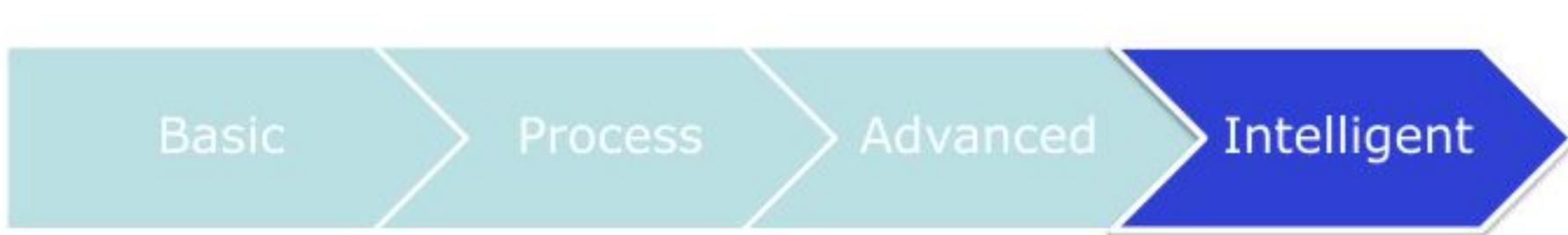
- **Target →** Complex process: Multiple systems + Human
- **Why**
 - Handles unstructured data / methods of access
- **Example →** Machine learning, natural language processing



Introduction to Business Process Automation

Types of Business Automation - Intelligent Automation

- **Target →** Processes requiring AI/customized decisions
- **Why**
 - Smarter interactions, personalization
- **Example →** Voice-based interfaces



Introduction to Business Process Automation

Types of Business Automation - Checkpoint

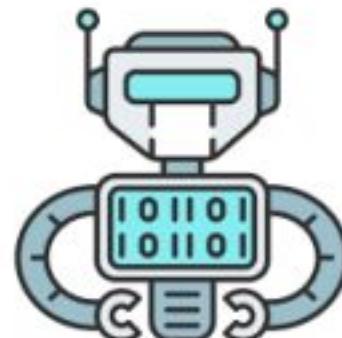
1. Daily email template to enter Market commentary
2. HR time off workflow
3. Online supermarket order processing
4. Voice-automated credit card hotline



Introduction to Business Process Automation

Robotic Process Automation (RPA) - Introduction

- Rule-based software
- Automating high-volume activities
- Free up human workers (for more meaningful tasks)

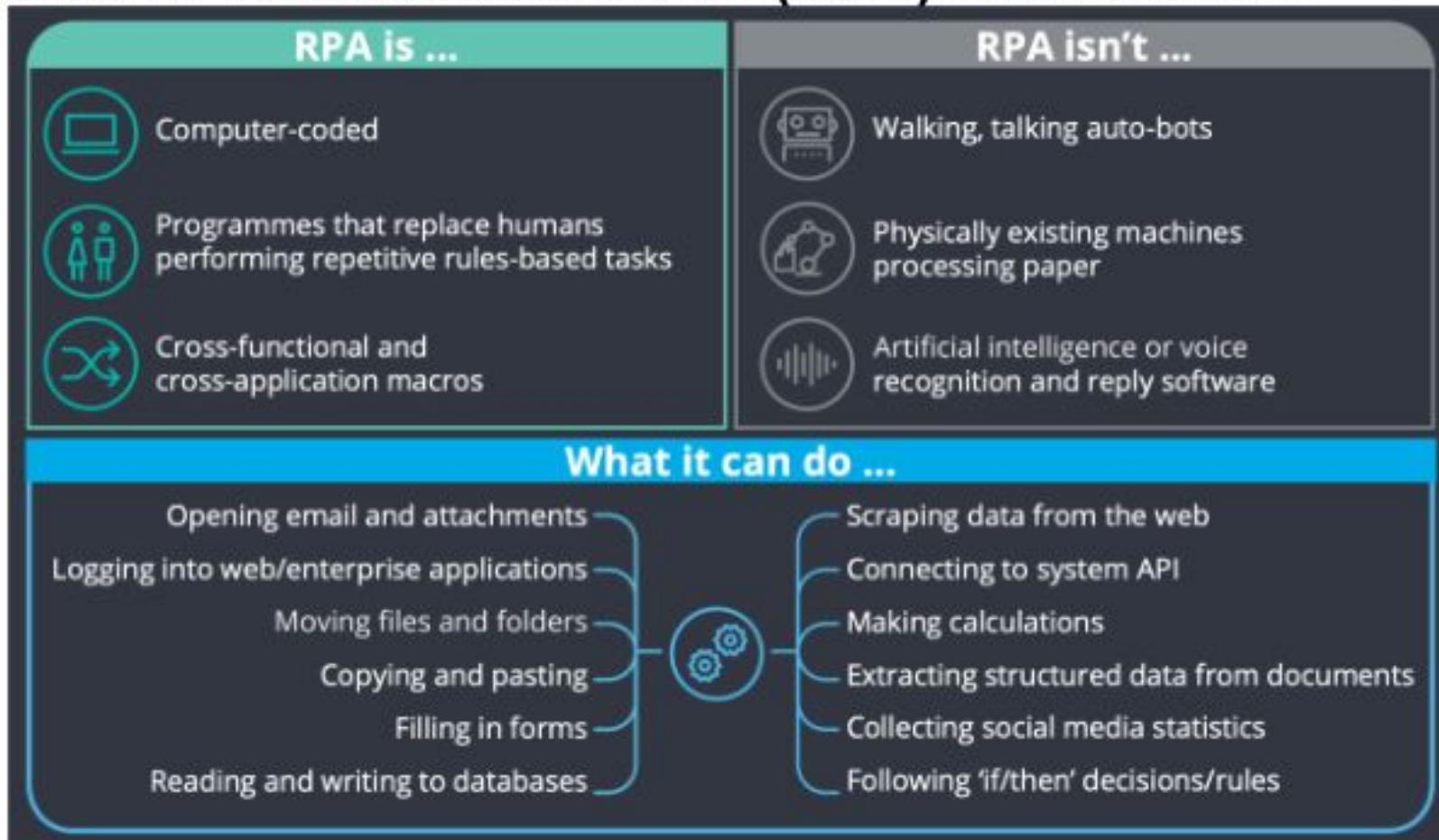


RPA bot

1. *What is Robotic Process Automation (RPA)?* | IBM. What Is Robotic Process Automation (RPA)? | IBM. <https://www.ibm.com/topics/rpa>

Introduction to Business Process Automation

Robotic Process Automation (RPA) - Introduction



1. Deloitte. (2018). *The robots are ready. Are you?*

<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

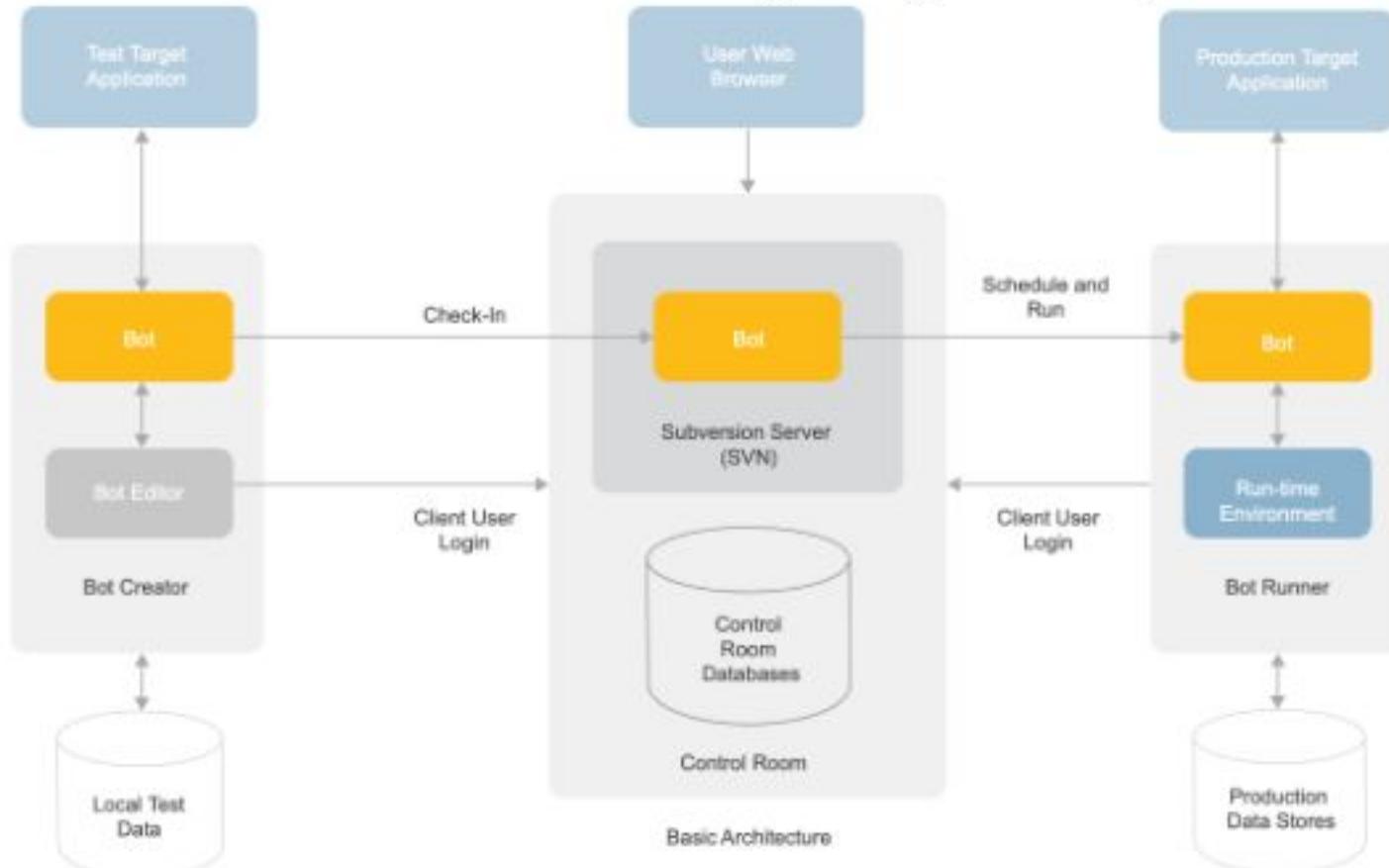
Robotic Process Automation (RPA) - Components

- **Bot modeling**
 - Bot Creator → development/ testing
 - Bot Runner → execution of tested codes in production
- **Control Room**
 - Deployment / Scheduler
 - Audit trails
 - Performance Analytics
 - Role-based access control



Introduction to Business Process Automation

Robotic Process Automation (RPA) - Components



1. Automation Anywhere. *Enterprise-class Security for Robotic Process Automation.*
https://www.automationanywhere.com/sites/default/files/internal-assets/uberflip/security-whitepaper_en.pdf

Introduction to Business Process Automation

Intelligent Process Automation (IPA)

- Using Artificial Intelligence (AI), e.g.,
 - Machine Learning (ML)
 - Natural Language Processing (NLP)
- Enhancing cognitive ability of the automation



Introduction to Business Process Automation

Intelligent Process Automation (IPA)

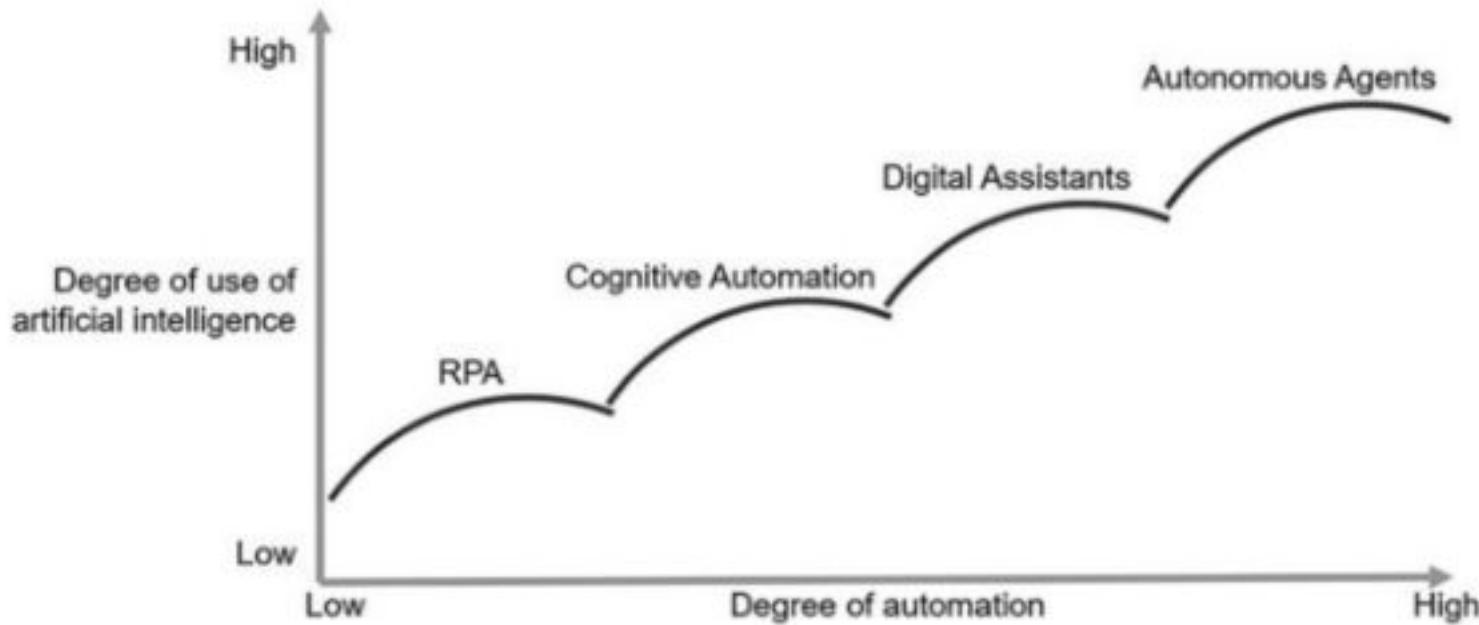
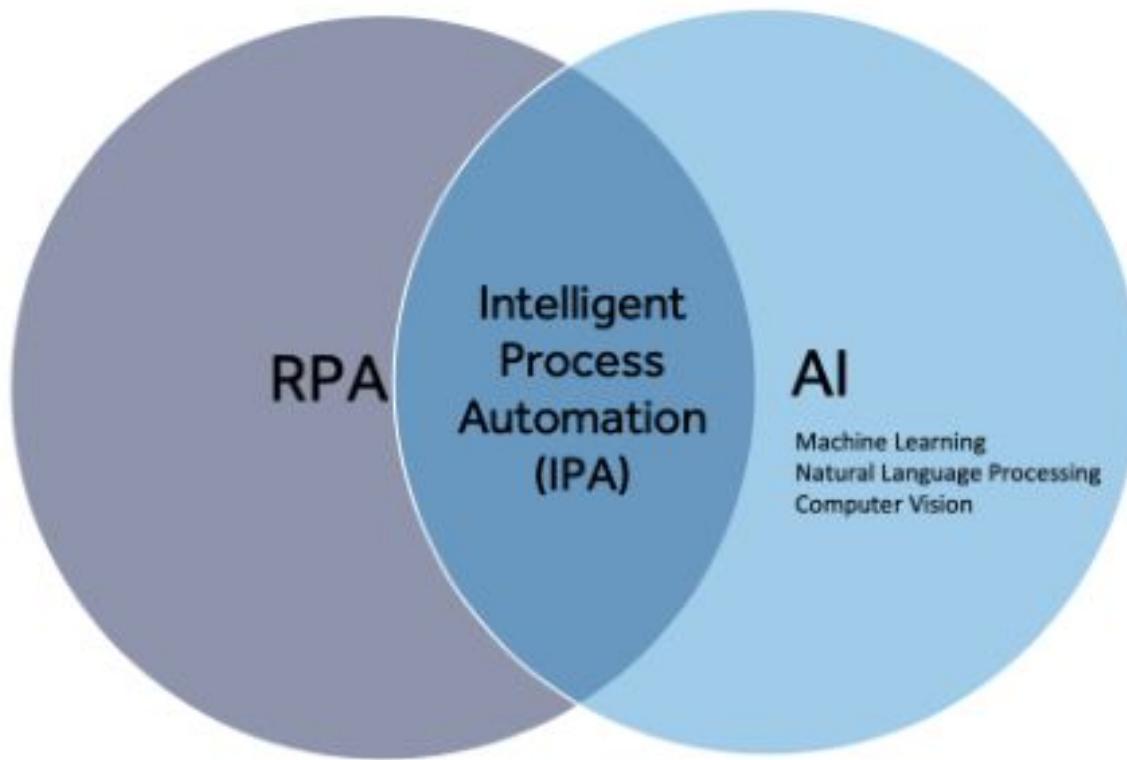


Fig. 2.2 Classification of RPA according to the degree of automation and the use of artificial intelligence (based on Ostrowicz 2018, p. 4)

1. Smeets, M., Erhard, R., & Kaußler, T. (2021, July 30). *Robotic Process Automation (RPA) in the Financial Sector: Technology - Implementation - Success for Decision Makers and Users*. Springer. <https://doi.org/10.1007/978-3-658-32974-7>

Introduction to Business Process Automation

Intelligent Process Automation (IPA)



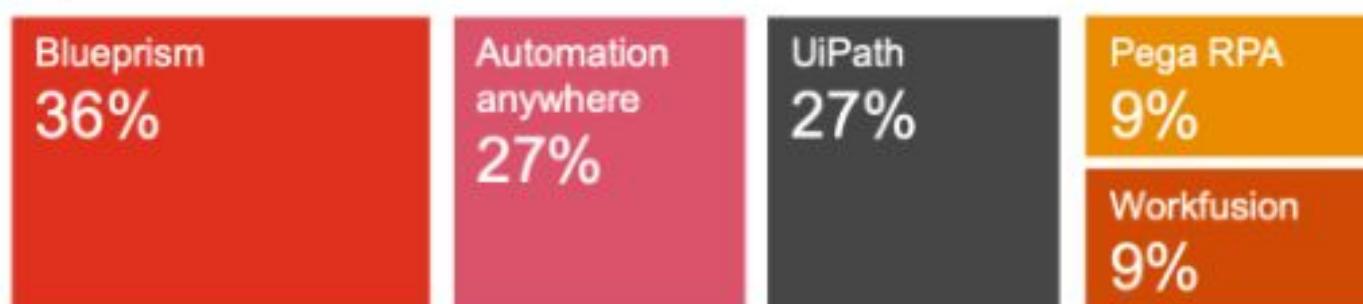
1. Kosmopoulos, C. (2021, May 11). *What is Intelligent Process Automation (IPA)?* | Blueprint. What Is Intelligent Process Automation (IPA)? | Blueprint. <https://www.blueprintsys.com/blog/rpa/what-is-intelligent-process-automation-ipa>

Introduction to Business Process Automation

Tools for Automation - Overview

Which tool is being used ?

As some companies use more than one tool, the percentages in this graphic total to more than 100%

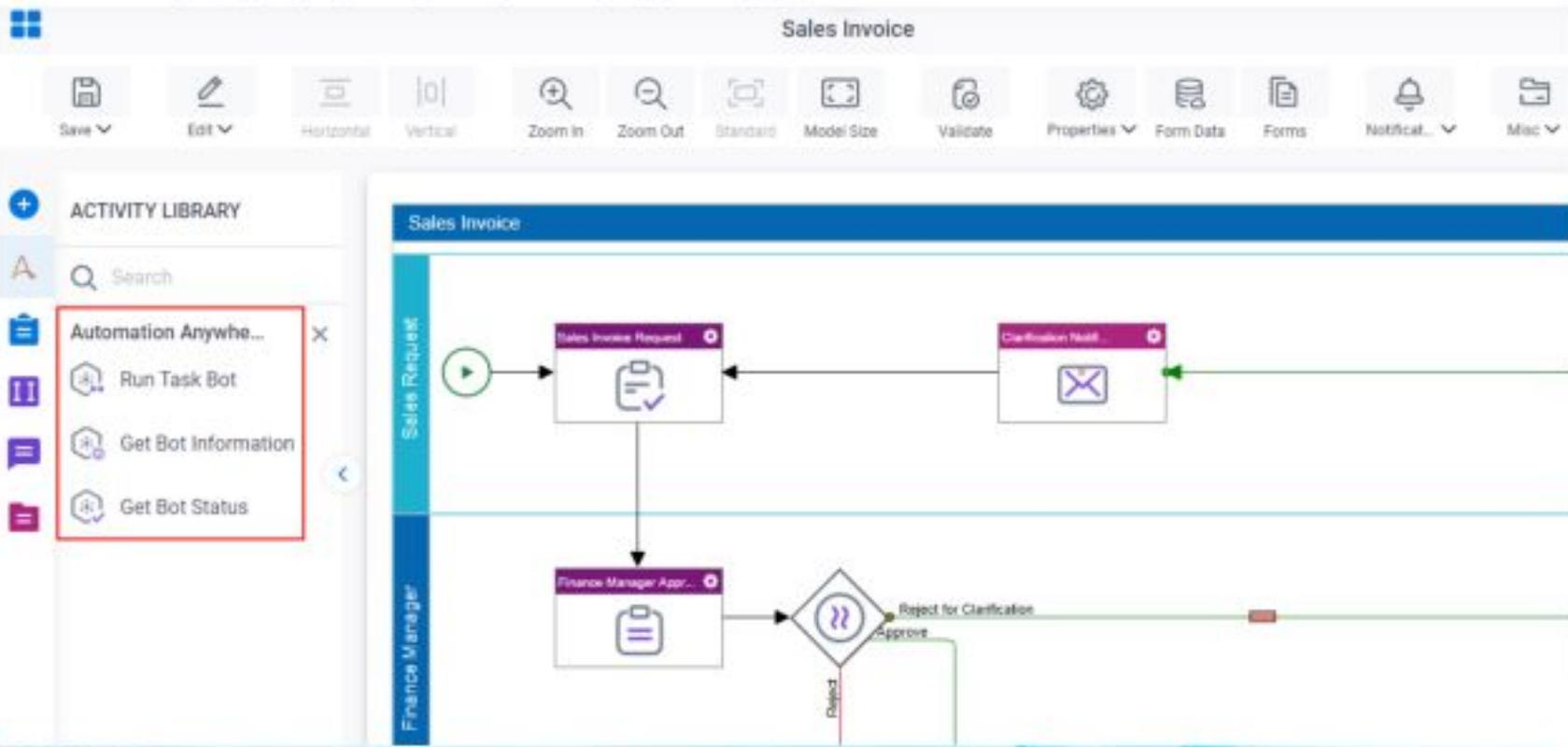


1. 2019 actuarial robotic process automation (RPA) survey report. PwC.
<https://www.pwc.com/gx/en/industries/financial-services/publications/pwc-2019-actuarial-robotic-process-automation-survey-report.html>

Introduction to Business Process Automation

Tools for Automation - Types

- **Low-code platforms**
 - Automation Anywhere, XCEPTOR, Microsoft Power Automate



Introduction to Business Process Automation

Tools for Automation - Types

- In-house Solutions

- Excel VBA, Java, Python, Web

Microsoft®
Excel VBA



Most in-demand functional areas

Compliance
Operational Risk
Fund Accounting
Corporate Accounting
Actuarial
Finance Operations
IT Audit

In-demand technical skills and experience

IFRS 17 Reporting and Accounting Policy
RBC and Capital Reporting
IT skills (Python, C+ BASIC, C, C++, COBOL, Java, R, Tableau, SQL, VBA, and Macro)
Operational Risk, Information and IT Risk, Risk Transformation, Risk Analytics, Internal Control Framework



1. Robert Half. 2024 Salary Guide HK.
<https://content.roberthalfonline.com/SG24/SG24-PDF/2024-Salary-Guide-HK.pdf>

Introduction to Business Process Automation

Tools for Automation – Getting a taste

- **Visualping** (<https://visualping.io/>)
 - Websites monitoring



<https://hkuspace.hku.hk/prog/cert-for-module-business-process-automation-with-vba>

Certificate for Module (Business Process Automation with VBA ...)

(2) Mr Jackie Liu. Mr. Jackie Liu is veteran quantitative strategist currently working in one of the top global investment banks. With over 10 years of ...

hku.hk
<https://hkuspace.hku.hk/prog/cert-for-module-business-process-automation-with-vba>

Certificate for Module (Business Process Automation with VBA ...)

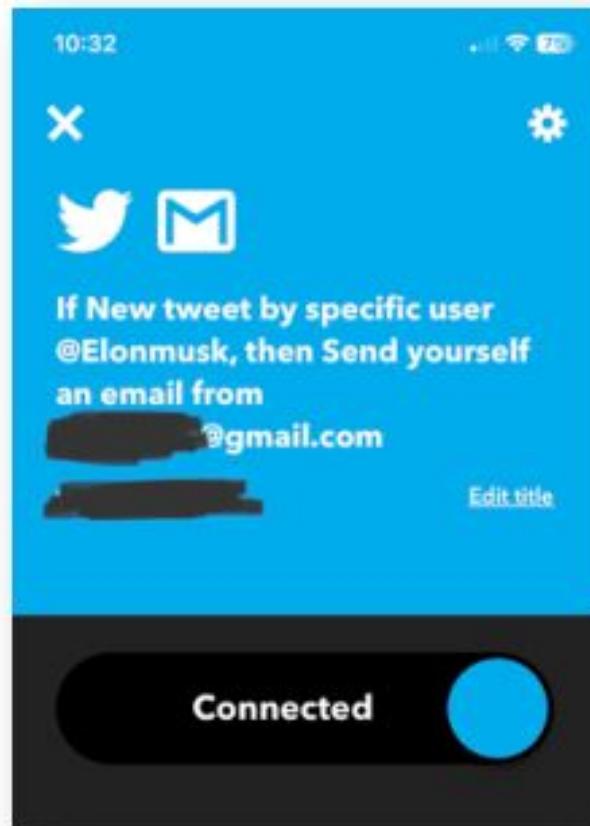
(2) Mr Jackie Liu. Mr. Jackie Liu is veteran quantitative strategist currently working in one of the top global investment banks. With over 10 years of ...

Introduction to Business Process Automation

Tools for Automation – Getting a taste

- **IFTTT (<https://ifttt.com/>)**

- Web + App interfaces
- Service-based triggers and actions



Introduction to Business Process Automation

Tools for Automation – Considerations

- Cost

- One-time cost

- Initial effort of adaptation
 - Implementation and rollout

- Running cost

- License, e.g., Power Automate: \$Millions/year¹
 - Infrastructure, e.g., Cloud storage
 - Training, e.g., Automation Anywhere, 10+ hours²
 - Support, e.g., bot re-configuration

Cost		
Production	Gold	Maint.
60 🎁	240 🎁	1 🎁

1. RPA Pricing: Comparison of Leading RPA Vendors' Fees in 2023. <https://research.aimultiple.com/rpa-pricing/>

2. RPA Course Learning Trails | Automation Anywhere University. <https://university.automationanywhere.com/training/rpa-learning-trails/>

Tools for Automation – Considerations

- **Flexibility**
 - Specialized vs Generalized, e.g., XCEPTOR → Tax
- **Scalability**
 - Desktop vs Enterprise, e.g., Excel VBA
 - Version Control, e.g., Use of repository
- **Time to delivery**
 - “Perfect” solution does not exist



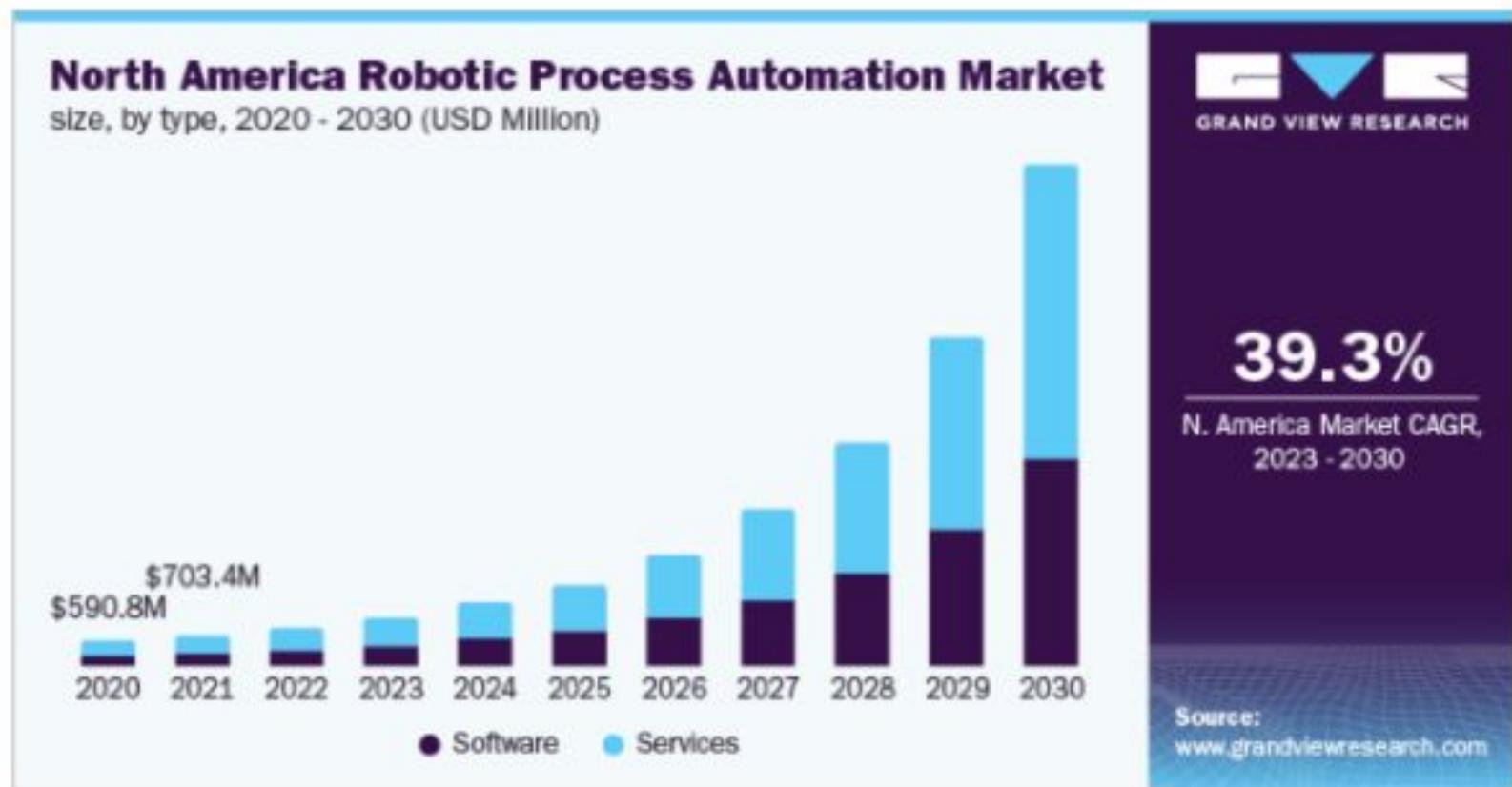
GitHub



1. Smeets, M., Erhard, R., & Kaußler, T. (2021, July 30). *Robotic Process Automation (RPA) in the Financial Sector: Technology - Implementation - Success for Decision Makers and Users*. Springer. <https://doi.org/10.1007/978-3-658-32974-7>

Introduction to Business Process Automation

Opportunities of BPA – The Money



1. *Robotic Process Automation Market Size & Share Report 2030.* Robotic Process Automation Market Size & Share Report 2030.
<https://www.grandviewresearch.com/industry-analysis/robotic-process-automation-rpa-market>

Introduction to Business Process Automation

Opportunities of BPA – Sectors

Industries	% of RPA Solutions
Banking, Financial Services & Insurance	51 %
Business Process Outsourcing (BPO)	14 %
Manufacturing – Consumer Packaged Goods (CPG)	7 %
Professional, Legal & Accountancy Services	7 %
Retail Trade	7 %
Technology (IT, Internet, SAAS)	7 %
Utilities	7 %

1. Top 15 RPA Use Cases & Examples in Banking in 2023. <https://research.aimultiple.com/banking-rpa/>

Opportunities of BPA – Sectors

- **Healthcare**
 - Monitoring, drug dispensing
- **Retail**
 - Auto-checkout, smart shelf, inventory management
- **Marketing**
 - Lead generation
 - social media management
- **Finance**
 - Fraud detection, Credit management, Regtech



Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Follow the money

 CNN

Wells Fargo ordered to pay \$3.7 billion for 'illegal activity' including unjust foreclosures and vehicle repossession

The CFPB said the more than \$2 billion in customer refunds Wells Fargo has been ordered to pay includes more than \$1.3 billion to consumers hurt...

20 Dec 2022



 Spiceworks

SEC Penalizes Major Wall Street Firms \$1.97B For Using Unauthorized Messaging Apps

If there are allegations of wrongdoing or misconduct, we must be ... Morgan Stanley recently agreed to a \$35 million fine by the SEC for...

28 Sept 2022



Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Follow the money



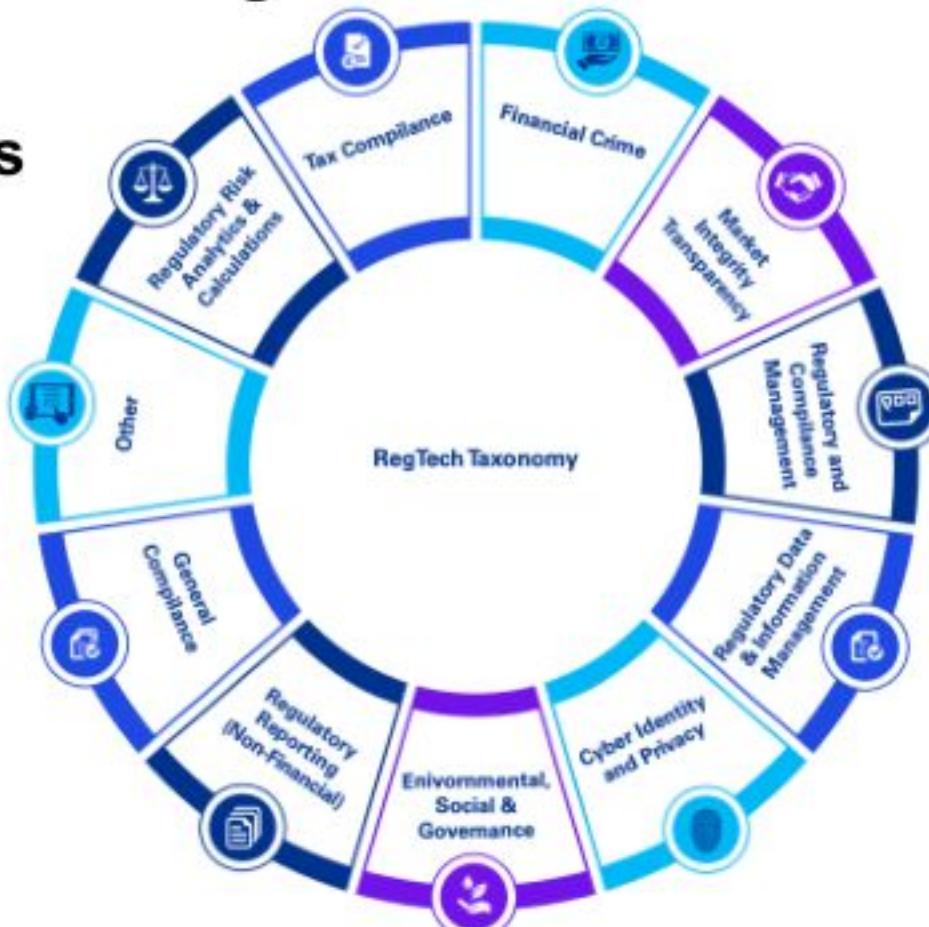
1. KPMG. (2019, June). *There's a revolution coming.*

<https://assets.kpmg.com/content/dam/kpmg/cn/pdf/en/2019/06/embracing-the-challenge-of-the-new-regtech-era.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Common themes



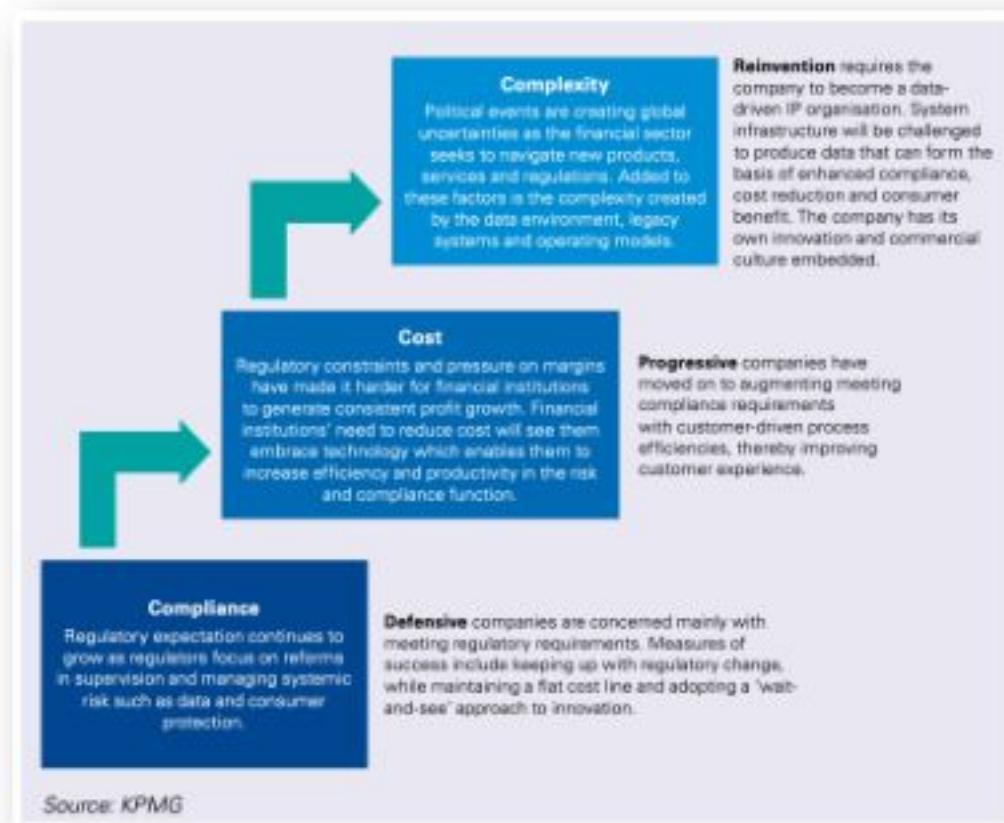
1. KPMG. (2022, November). <https://assets.kpmg.com/content/dam/kpmg/uk/pdf/2022/11/innovate-finance-regtech-industry-and-adoption.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

Adoption phases:

1. Defensive
2. Progressive
3. Reinvention



1. KPMG. (2019, June). *There's a revolution coming.* <https://assets.kpmg.com/content/dam/kpmg/cn/pdf/en/2019/06/embracing-the-challenge-of-the-new-regtech-era.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- HKMA 3-Year Roadmap
 - RPA: automating rule-based processes
 - Network Analysis: Company holdings and exposures

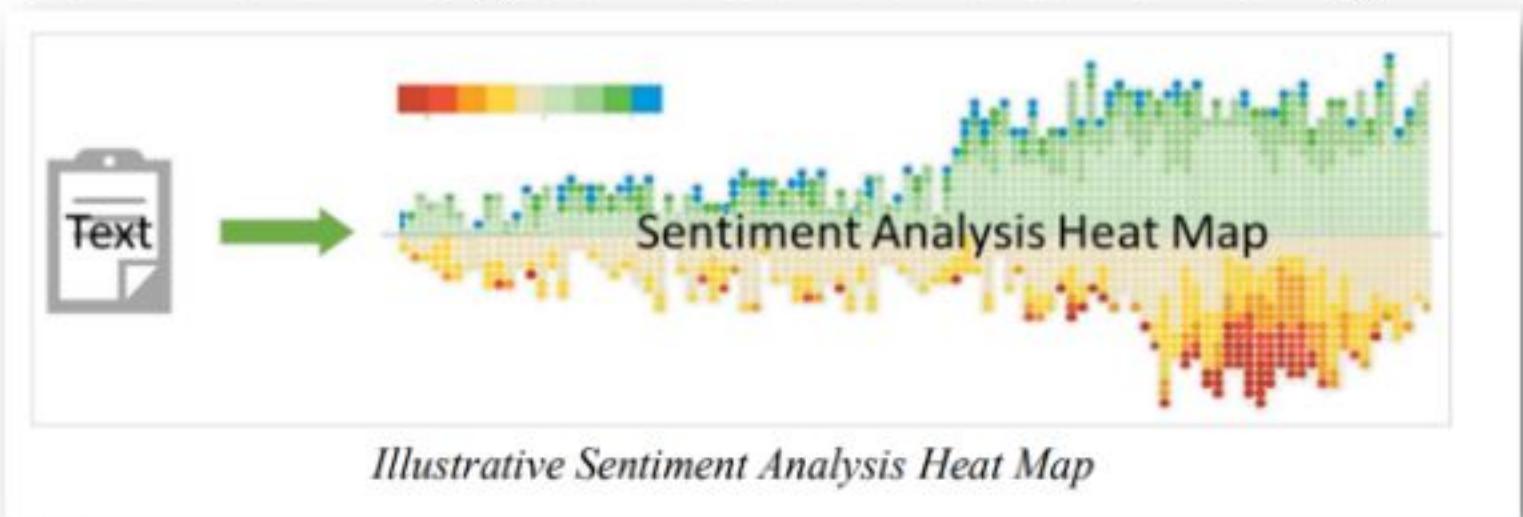


1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- **HKMA 3-Year Roadmap**
 - **Speech-to-text (STT)**: Audio to time-stamped text
 - **Sentiment Analysis**: Realtime news monitoring

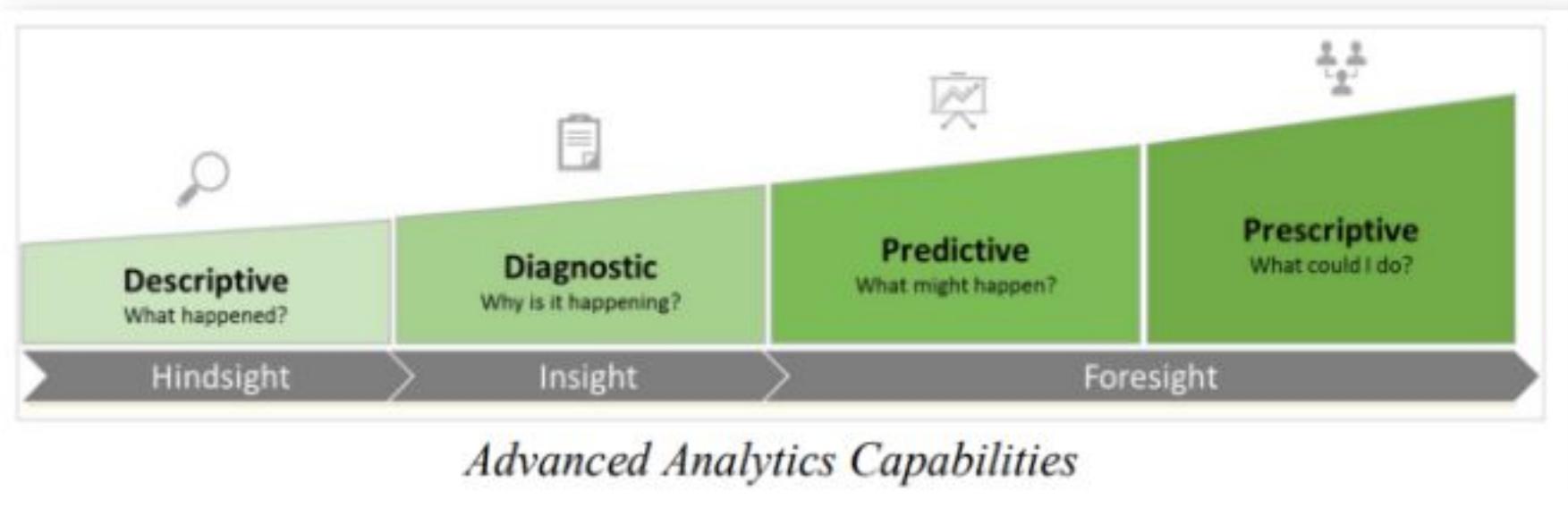


1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Analytics enabled by BPA



1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Challenges of BPA – Insight from the banks

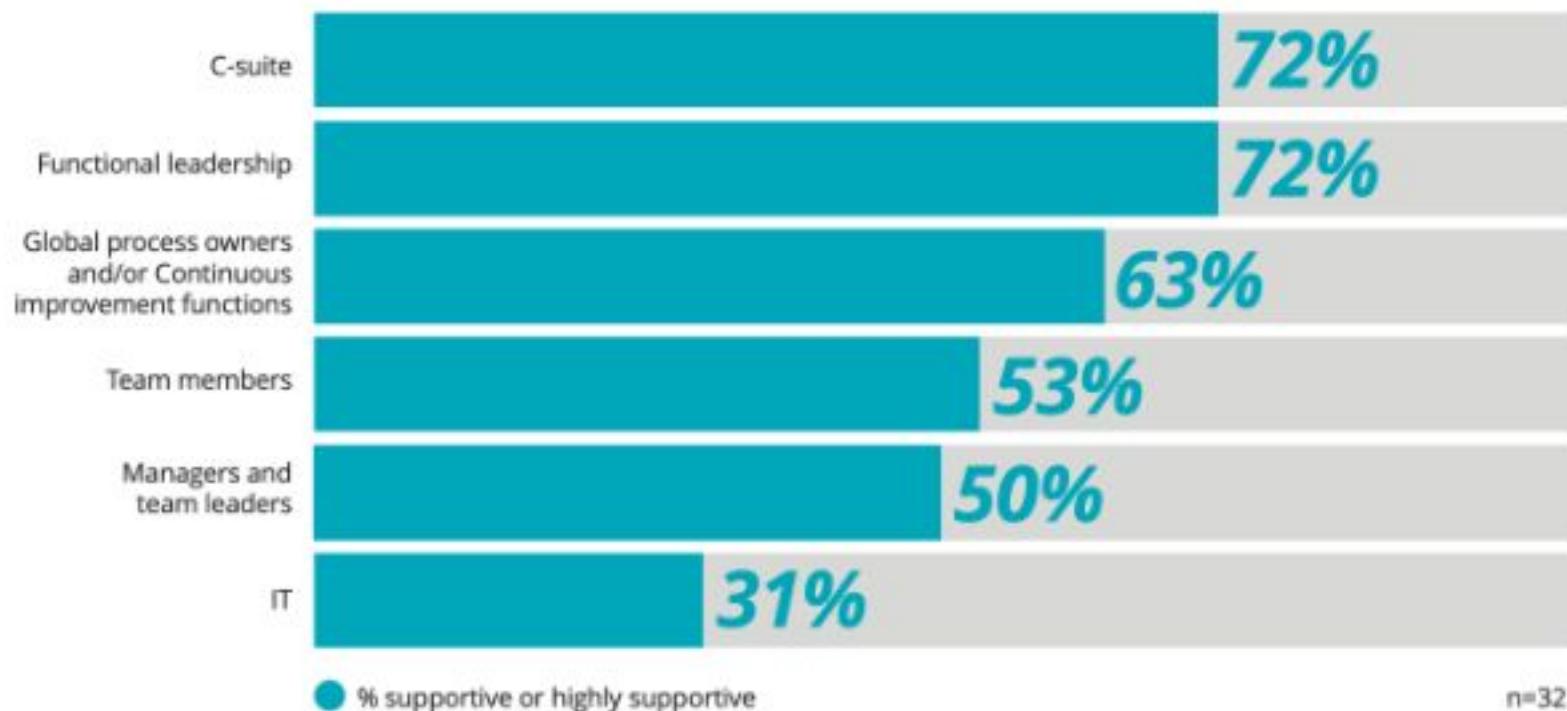


1. KPMG & HKMA. (2020). *Transforming Risk Management and Compliance: Harnessing the Power of Regtech.* <https://www.hkma.gov.hk/media/chi/doc/key-information/press-release/2020/20201102c3a1.pdf>

Introduction to Business Process Automation

Challenges of BPA – Worries from the workers?

Figure 5: How supportive of the RPA implementation were your stakeholder groups?

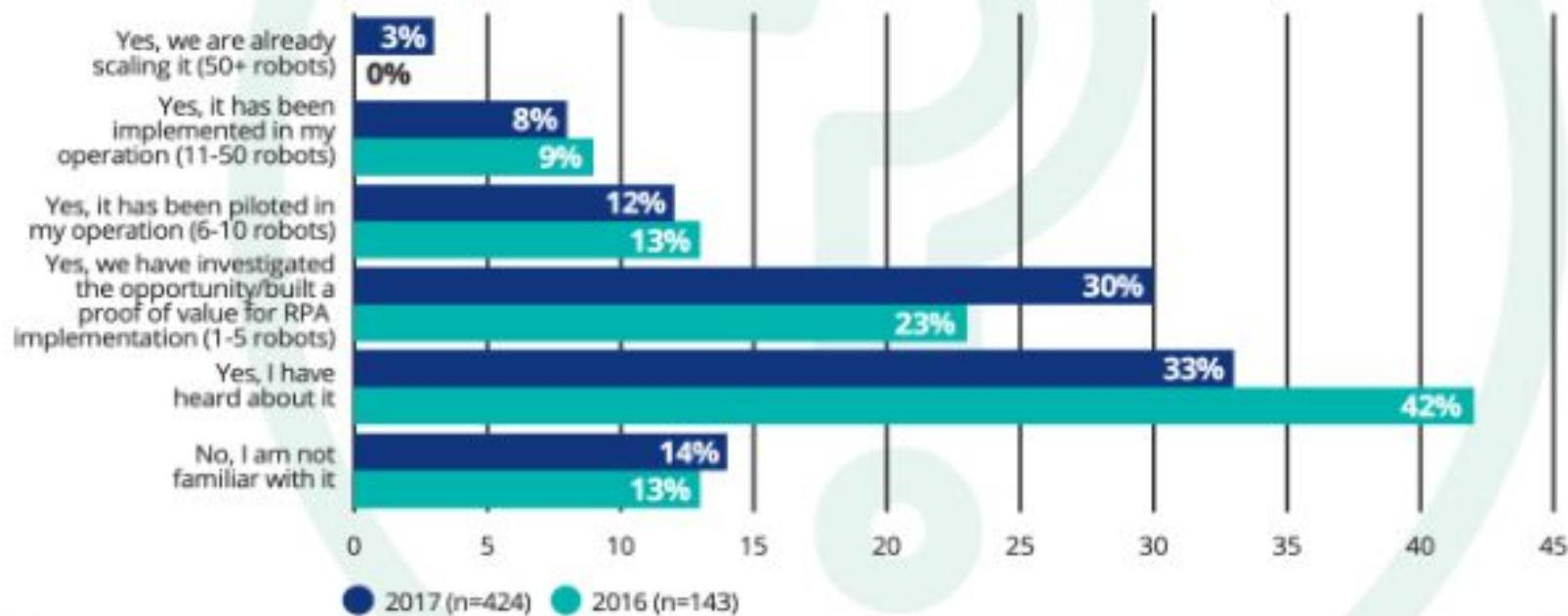


1. Deloitte. (2018). *The robots are ready. Are you?*
<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

Introduction to Business Process Automation

Challenges of BPA – Hesitant to scale up

Figure 2: Are you familiar with RPA?



1. Deloitte. (2018). *The robots are ready. Are you?*
<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

Challenges of BPA

- **Shortage**
 - Budget / Resource/ Talent
- **Concerns**
 - Risk of rushing / failing
 - Moral / Structural unemployment

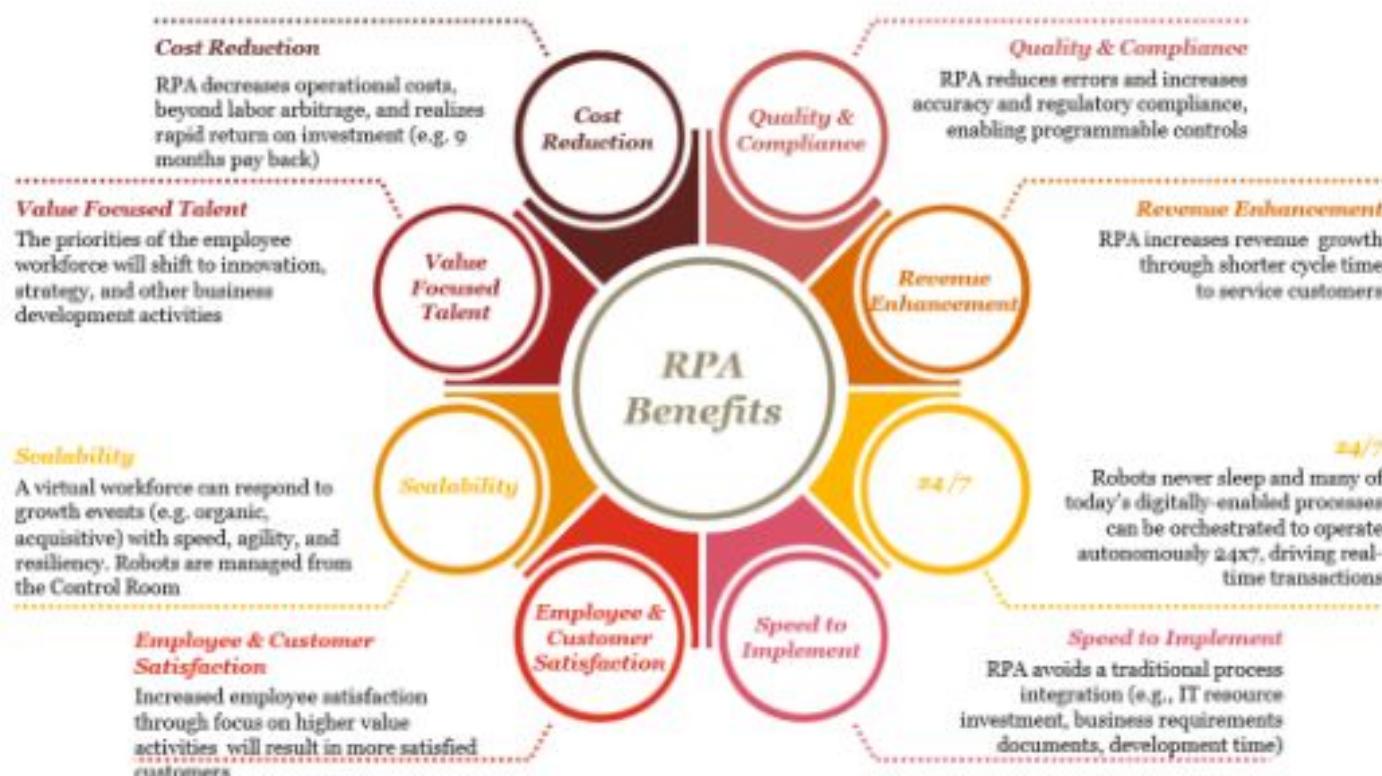


* Success depends on planning
(More on project management later)

Introduction to Business Process Automation

Implications of BPA – The Good

- Cost, Efficiency, Customer Experience



1. PwC Robotics Process Automation Solutions. PwC.

<https://www.pwchk.com/en/services/entrepreneurial-and-private-business/new-technology-digitalisation-and-transformation/process-automation-solutions.html>

Introduction to Business Process Automation

Implications of BPA – The Bad

- Automation complacency / dependency
- Risk of Technical Issues / Downtime

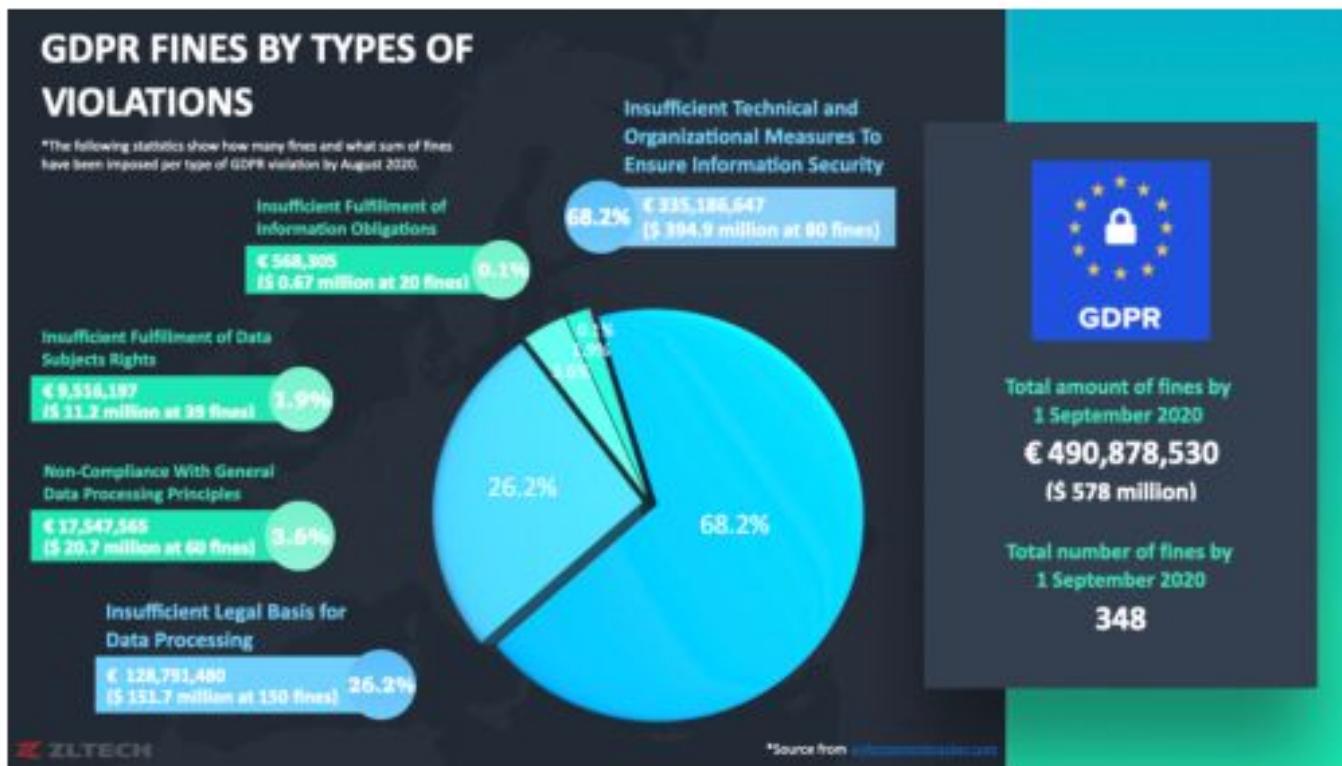


Image: Civil Aviation Safety Authority

Introduction to Business Process Automation

Implications of BPA – The Ugly

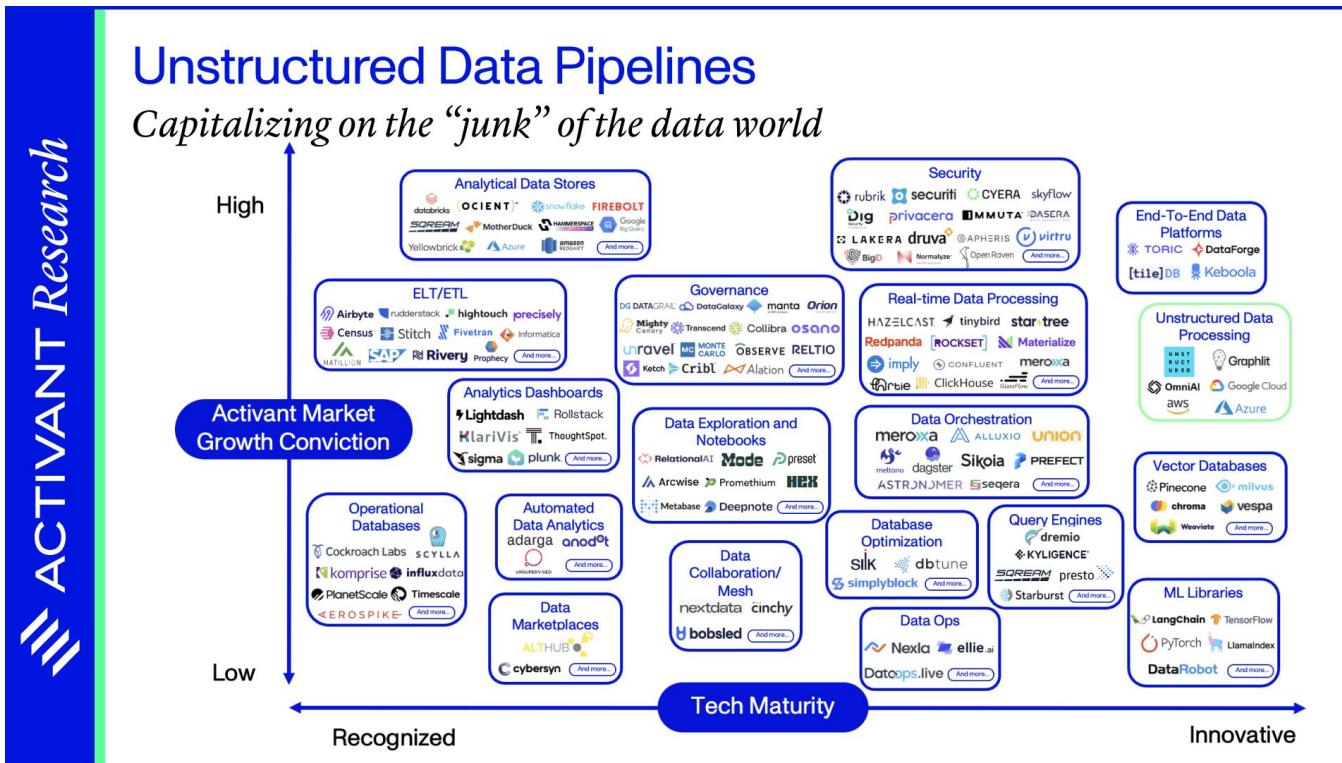
- Data Collection and Analysis¹
- Vulnerability to Cyberattacks



1. Chen, B. (2020, September 9). GDPR Recap: 28 Nations, 348 Fines, Half a Billion Euros. ZL Tech.
<https://www.zlti.com/blog/gdpr-recap-28-nations-348-fines-half-a-billion-euros/>

IPA – Data Science

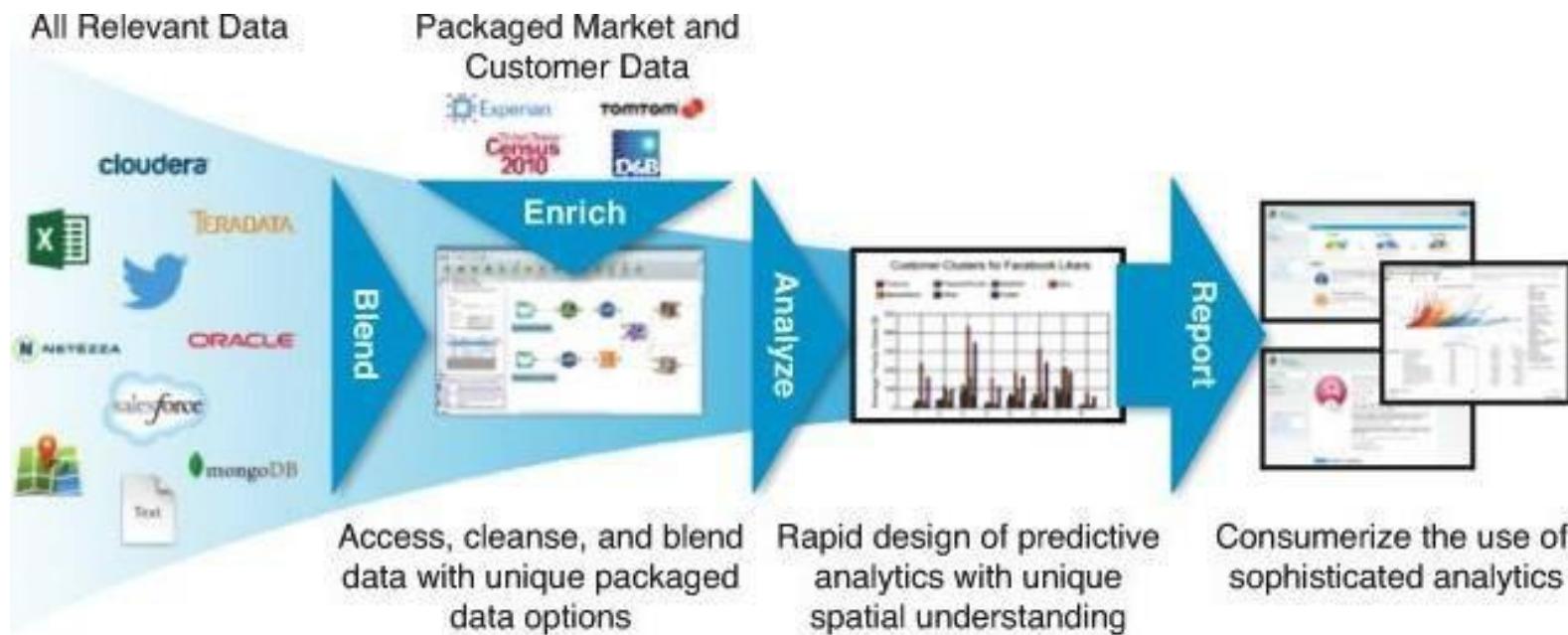
IPA's integration with data science is particularly impactful in several areas:
Unstructured Data Processing: IPA can convert unstructured data into structured formats suitable for analysis, enabling businesses to leverage insights from previously untapped data source.



Modern BI Process - Blend, Enrich, Analyze, and Report

Predictive Analytics: By applying machine learning algorithms to historical project data, organizations can forecast outcomes more accurately and mitigate risks in project delivery

Business Intelligence: Automating the preparation of data for analytics tools allows organizations to gain actionable insights more rapidly, enhancing their decision-making capabilities.



What Is Data Science?

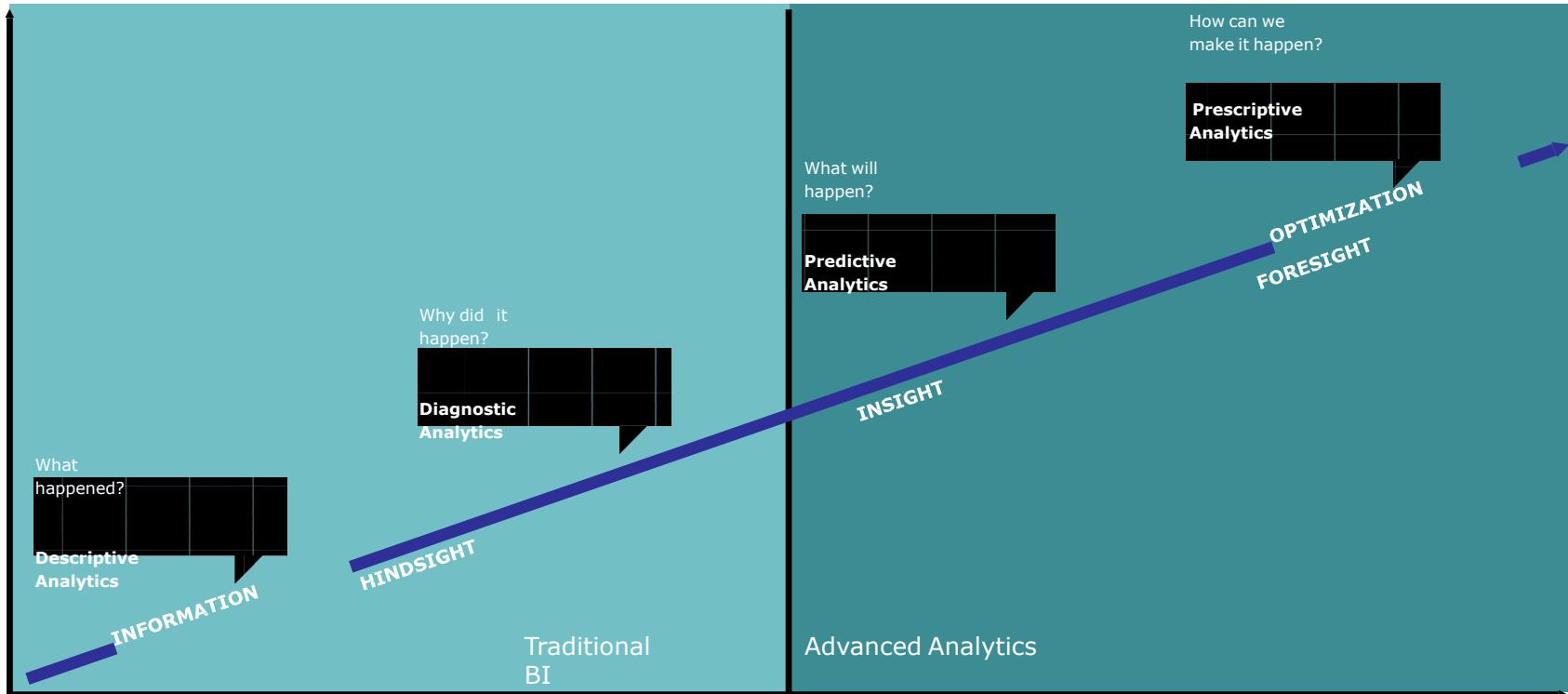
From Wikipedia:

Data science, also known as data-driven science, is an interdisciplinary field about scientific processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, machine learning, data mining ,and predictive analytics similar to Knowledge Discovery in Databases (KDD).

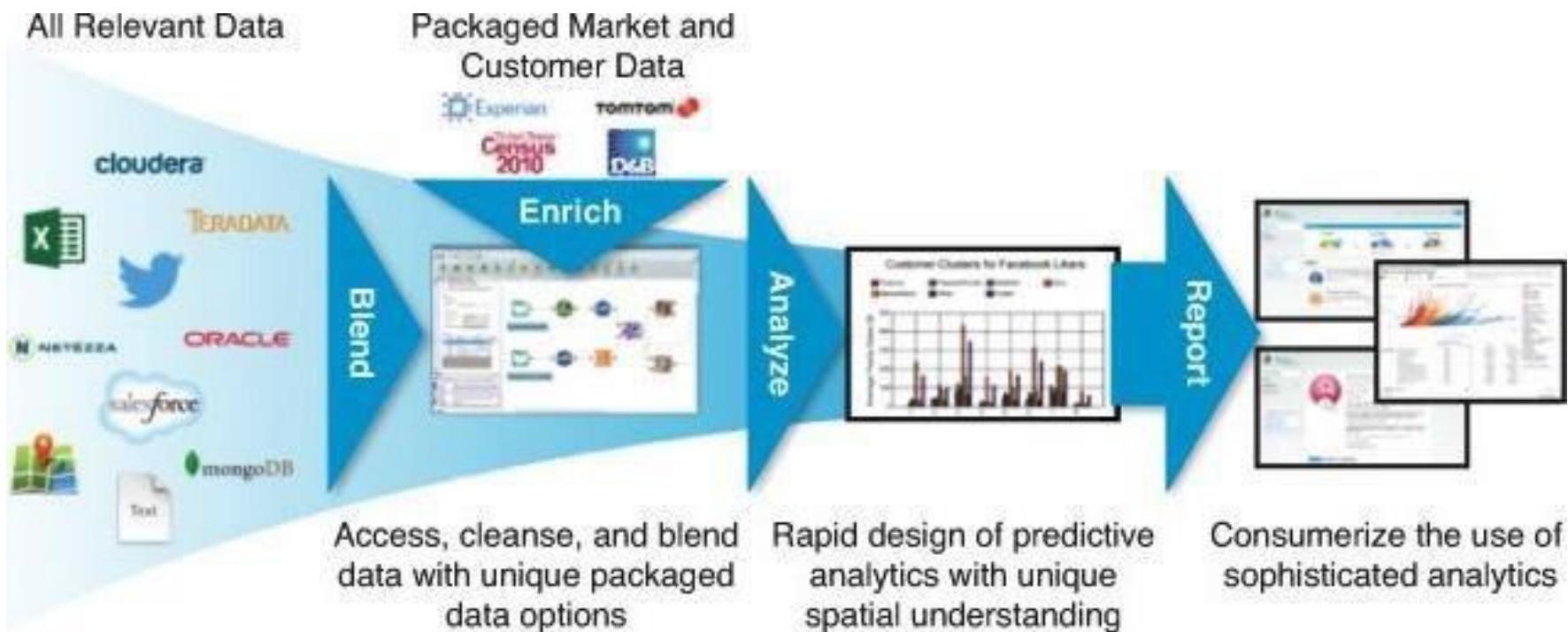


What is Process Automation, RPA, IPA?
Overview of the technological building blocks related to process automation
Contemporary tools for process automation
Challenges and opportunities of business process automation
Business implications of process automation

How Is Business Intelligence Different From Data Science?



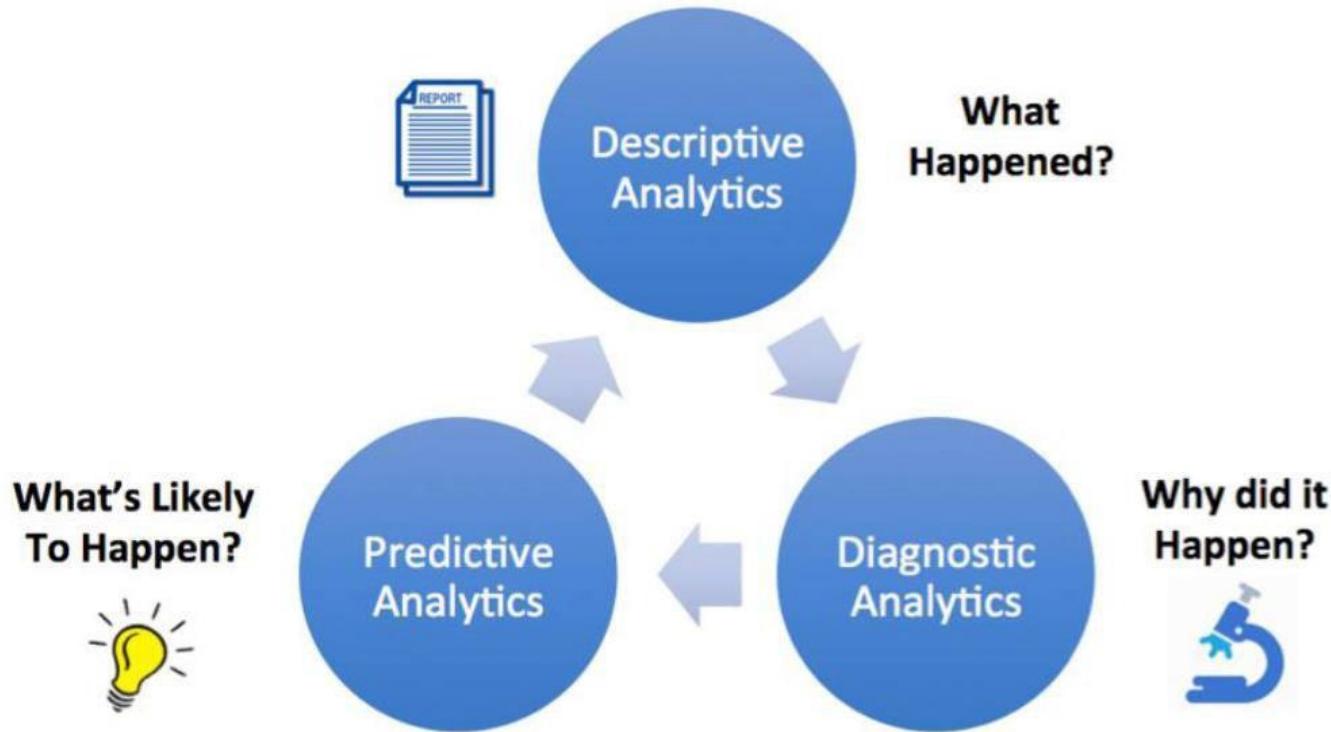
Modern BI Process - Blend, Enrich, Analyze, and Report



Source: Advanced Analytics Methodologies: Driving Business Value with Analytics by Michele Chambers and Thomas W. Dinsmore (0133498603) Copyright © 2015 Pearson Education, Inc. All rights reserved.

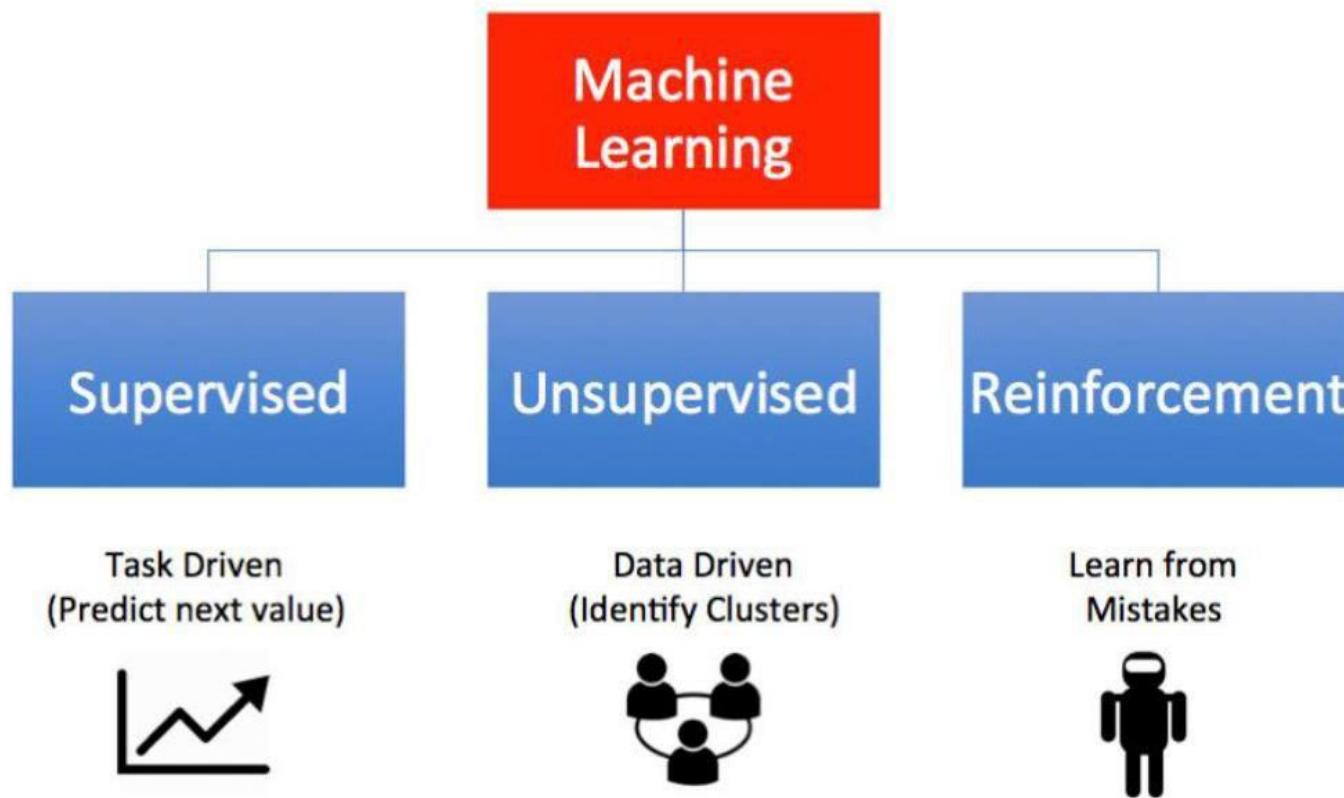
What is Business Analytics & Machine Learning?

Types of Analytics

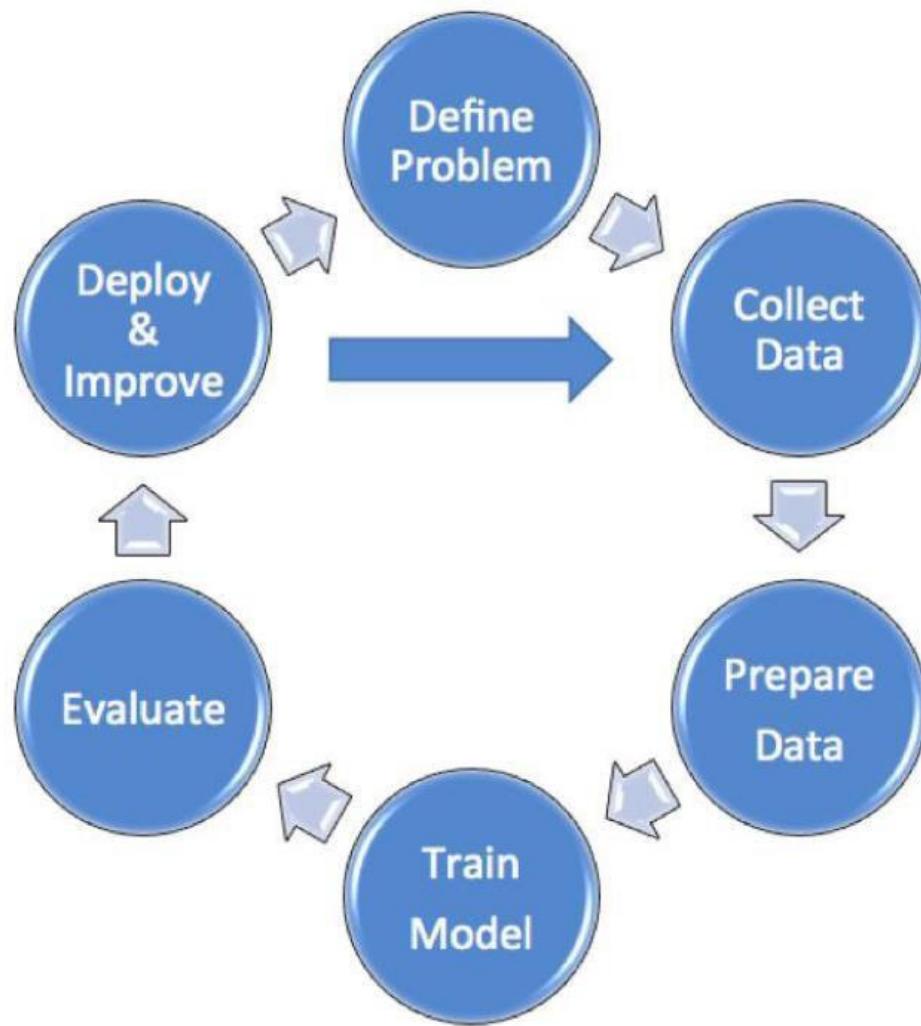


What is Business Analytics & Machine Learning?

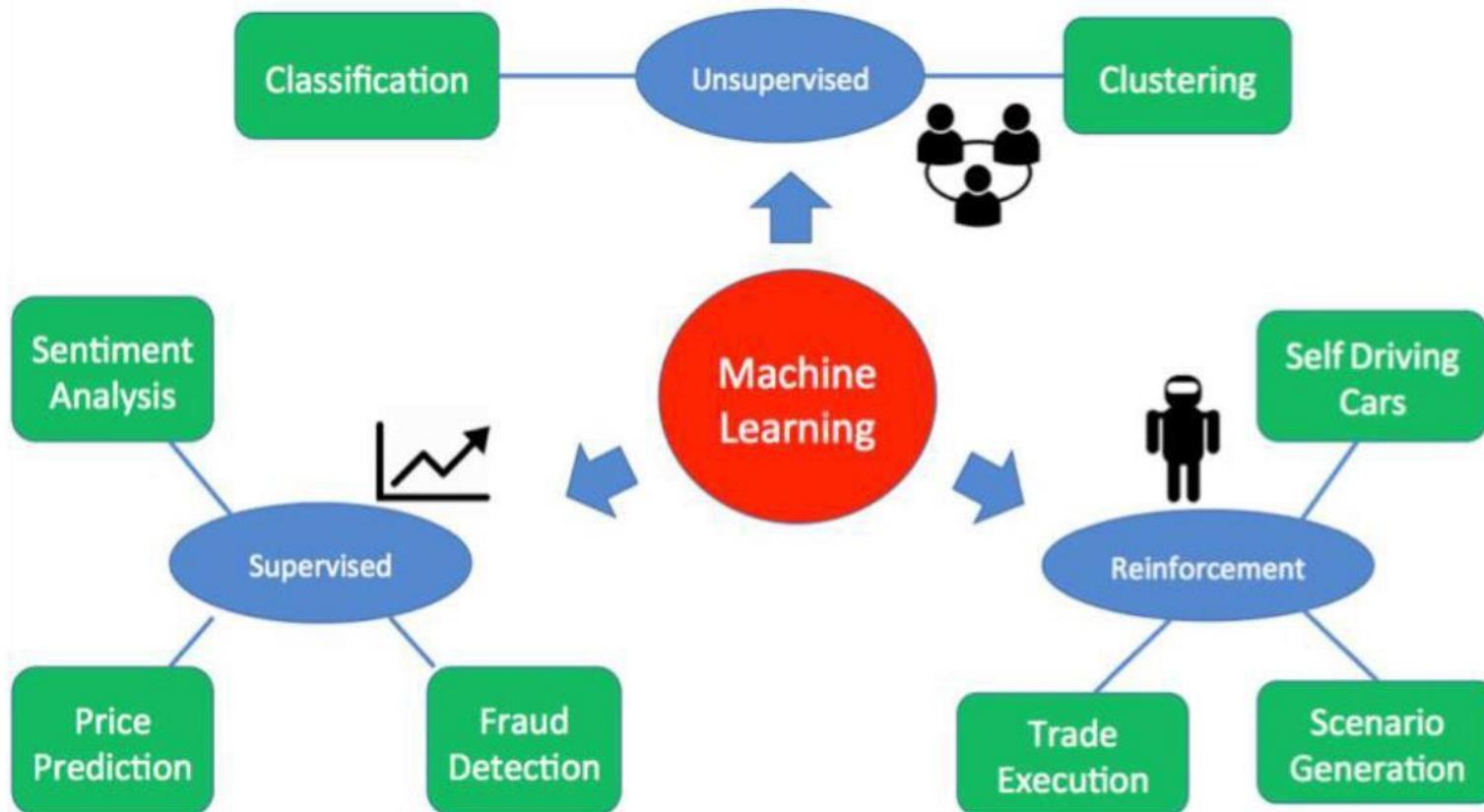
Types of Machine Learning



What is Business Analytics & Machine Learning?



What is Business Analytics & Machine Learning?



How To Do Data Mining: Machine Learning Techniques

Essential Concepts of Machine Learning

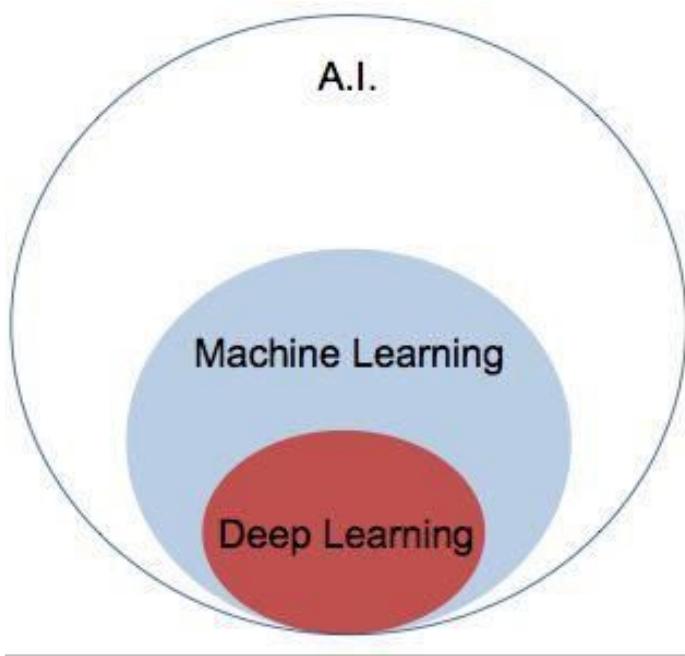
Artificial Intelligence

Machine Learning

Deep Learning

Source: <https://www.becomingadatascientist.com/2017/07/17/introductory-machine-learning-terminology-with-food/>

Clearing The Confusion About Data Science And Big Data



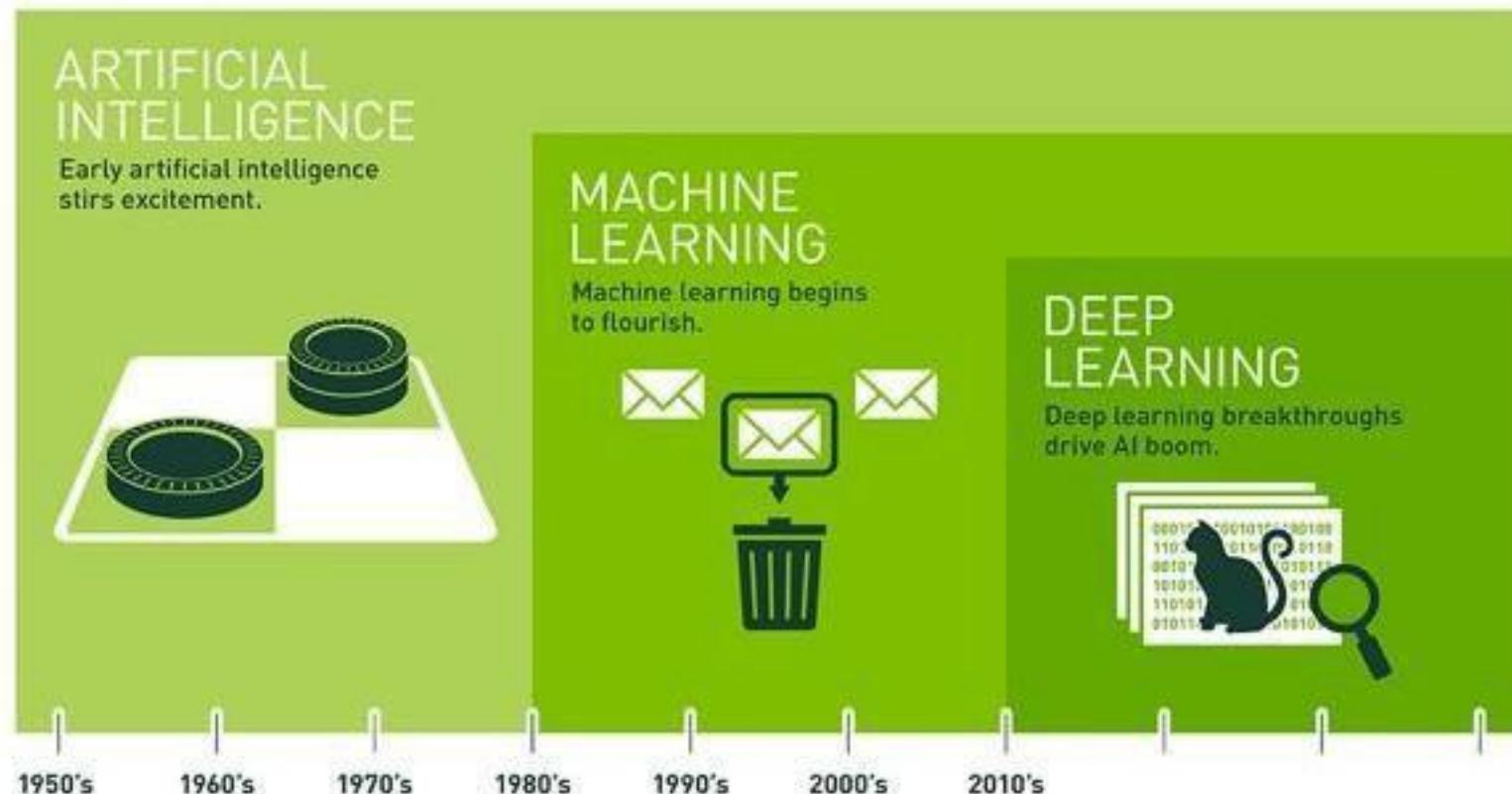
Definition of AI: a branch of computer science trying to get computers to exhibit intelligence.

- Machine Learning is a range of techniques for computers to perform cognitive functions
- Deep Learning (a type of Machine Learning that uses layers of neural networks) has been the most accurate and productive technique in AI research

Frank Chen of a 16z

Big Data Tools and Algorithm Exploration

Difference Between AI, Machine Learning & Deep Learning



Since an early flush of 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.

How To Do Data Mining: Machine Learning Techniques

Artificial Intelligence

Artificial Intelligence (AI) is the science of making things smart. Can be defined as:

“Human intelligence exhibited by machines”

A broad term for getting computers to perform human tasks. The scope of AI is disputed and constantly changing over time. Let's go deeper...

How To Do Data Mining: Machine Learning Techniques

Machine Learning

An approach to achieve artificial intelligence through ‘software’ systems that can **learn** from **experience** to find **patterns** in a **set of data**.

Write a computer program
with **explicit rules** to follow

```
if email contains V!agra  
  then mark is-spam;  
if email contains ...  
if email contains ...
```

Write a computer program
to **learn from examples**

```
try to classify some emails;  
change self to reduce errors;  
repeat;
```

Traditional Programming

Machine Learning Programs

How To Do Data Mining: Machine Learning Techniques

Gathering Data



Inspiring Your Future

上環/中環/金鐘 寶泰大廈 [室內相]
售 \$530 萬 月供 \$13,711 按揭計算機»

單位詳情

成交紀錄

屋苑資訊

附近學校



樓盤	上環/中環/金鐘 寶泰大廈
樓盤編號	LCX458-34676
地址	水坑口街9號 <small>中原地圖»</small>
價格	售 \$530 萬
實用面積	357 呎 (\$14,845/呎)
建築面積	519 呎 (\$10,211/呎)
實用率	69%
樓層	低層 (B室)
樓齡	33年
單位類型	分層單位
輟學特色	<small>小學: 11» 中學: 中西區»</small>

Business Education @ HKUSPACE

Source: <http://www.become-a-scientist.com/2017/07/17/introductory-machine-learning-terminology-with-food/>

How To Do Data Mining: Machine Learning Techniques

Features / Attributes

Are used to train a Machine Learning algorithm

Facts and Features

	Type Single Family		Year Built 1989		Heating Other
	Cooling No Data		Parking 2 spaces		Lot 6,300 sqft

INTERIOR FEATURES

Bedrooms	Flooring
Beds: 6	Floor size: 3,138 sqft



Using a Home Valuation Prediction Model
as an example

How To Do Data Mining: Machine Learning Techniques

Features / Attributes

Are used to train a Machine Learning algorithm

Sales Price

Saleable

Area

Number of Bedrooms

Year built

(Refurbished) Schools
in the area

Transportation (MTR, buses)

Address: 270-280 QUEEN'S ROAD
WEST
Price per sqft.: 17,841
Qty transaction: 2
Year: 2017



Using a Home Valuation Prediction Model as an example

How To Do Data Mining: Machine Learning Techniques

Machine Learning

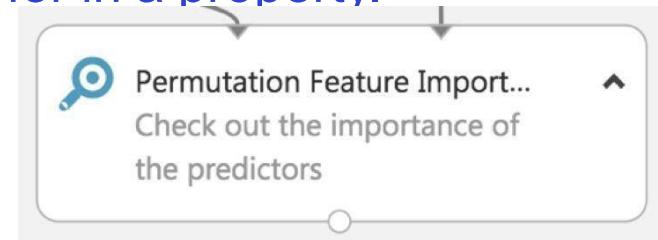
Feature	Score
	1.
SaleableArea	0.858545
Eff_Build_Ages	0.166877
Bus_Route	0.076519
Parks	0.019011

Choosing a ‘useful’ or **significant** feature (predictors) **makes** or **breaks** a data science project.

Domain experts, property agents will tell you what property buyers value in a home.

Data scientists would **NOT** have a clue what home buyers look for in a property.

But not all is lost!



Data Processing & Analysis - What Is Machine Learning?

Supervised Learning

Used as an advanced form of *predictive modeling*

Each observation must be labeled with a "correct answer"

Only then can you build a predictive model because you must tell the algorithm what's "correct" while training it (hence, "supervising" it)

Regression is the task for modeling continuous target variables

Classification is the task for modeling categorical (a.k.a. "class") target variables



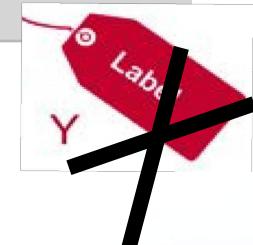
Unsupervised Learning

Used either as a form of automated data analysis or automated signal extraction

Unlabeled data has no predetermined "correct answer"

You'll allow the algorithm to directly learn patterns from the data (without "supervision")

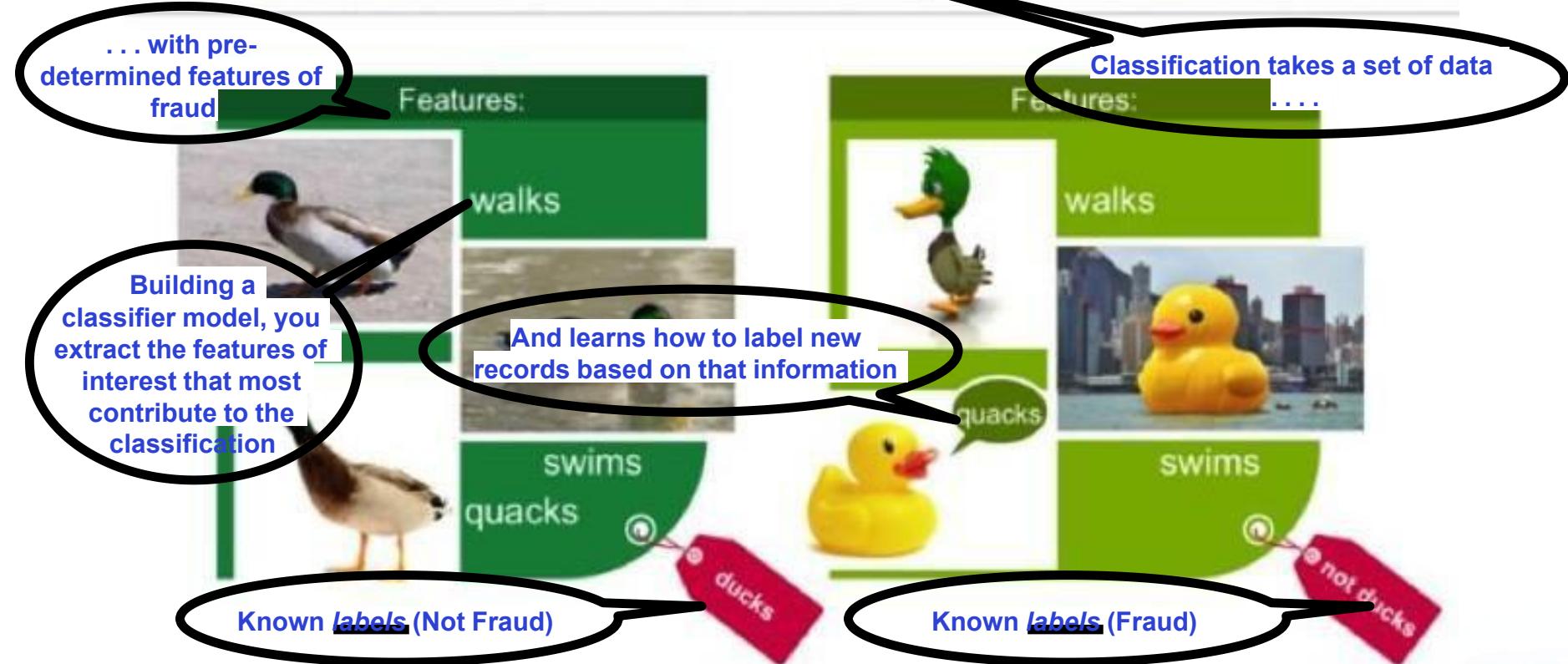
Clustering is the most common unsupervised learning task, and it's for *finding groups within your data*



Data Processing & Analysis

Supervised Learning: Classification For Debit Card Fraud

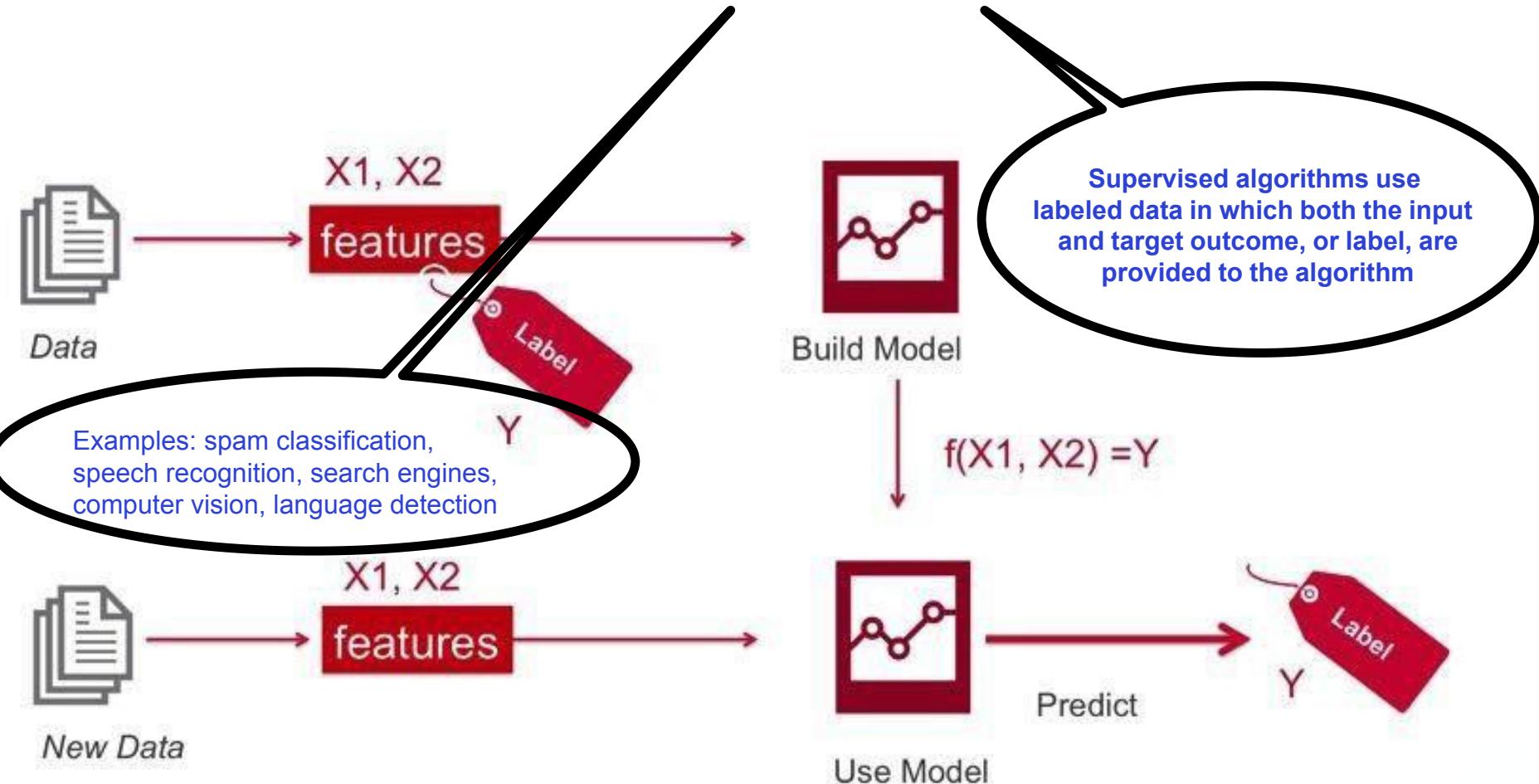
If it Walks/Swims/Quacks Like a Duck . . . Then It Must Be a Duck



Source:
<https://dzone.com/articles/demystifying-ai-machine-learning-and-deep-learning>

Data Processing & Analysis

Machine Learning: Supervised Learning



Source:
<https://dzone.com/articles/demystifying-ai-machine-learning-and-deep-learning>

Data Processing & Analysis

Machine Learning: Decision Trees

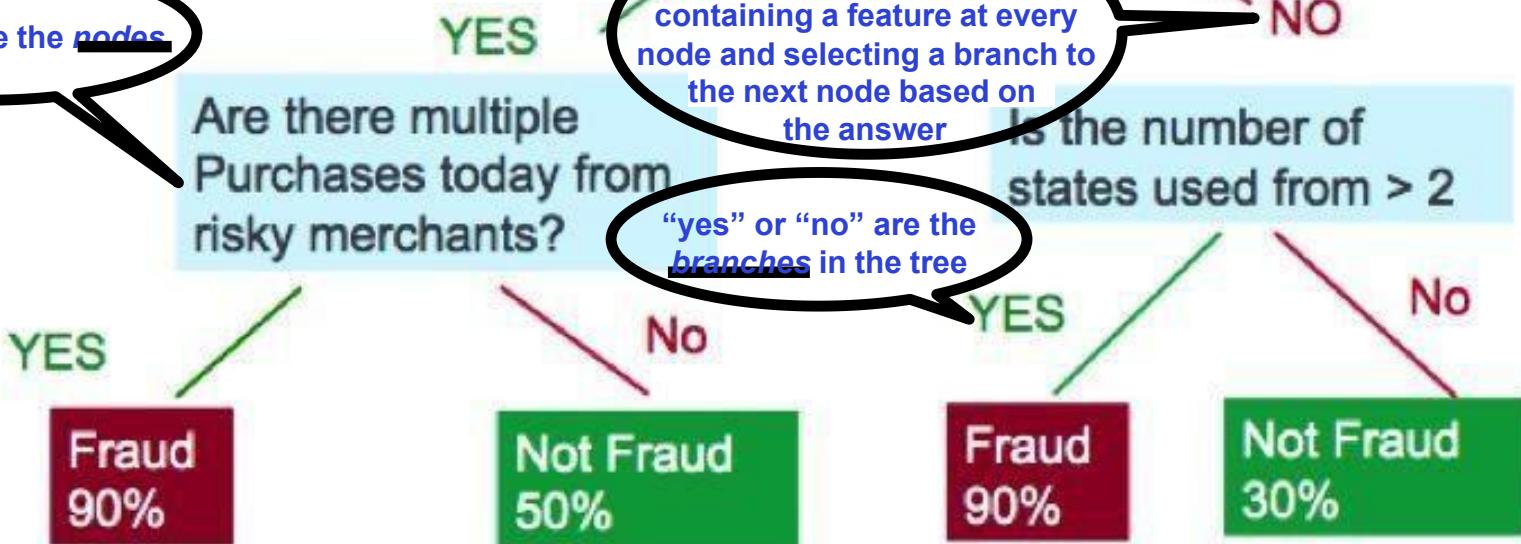
decision tree for predicting debit card fraud

Decision trees create a model that predicts the class or label based on several input features

questions are the nodes

Is the amount spent in 24 hours > average

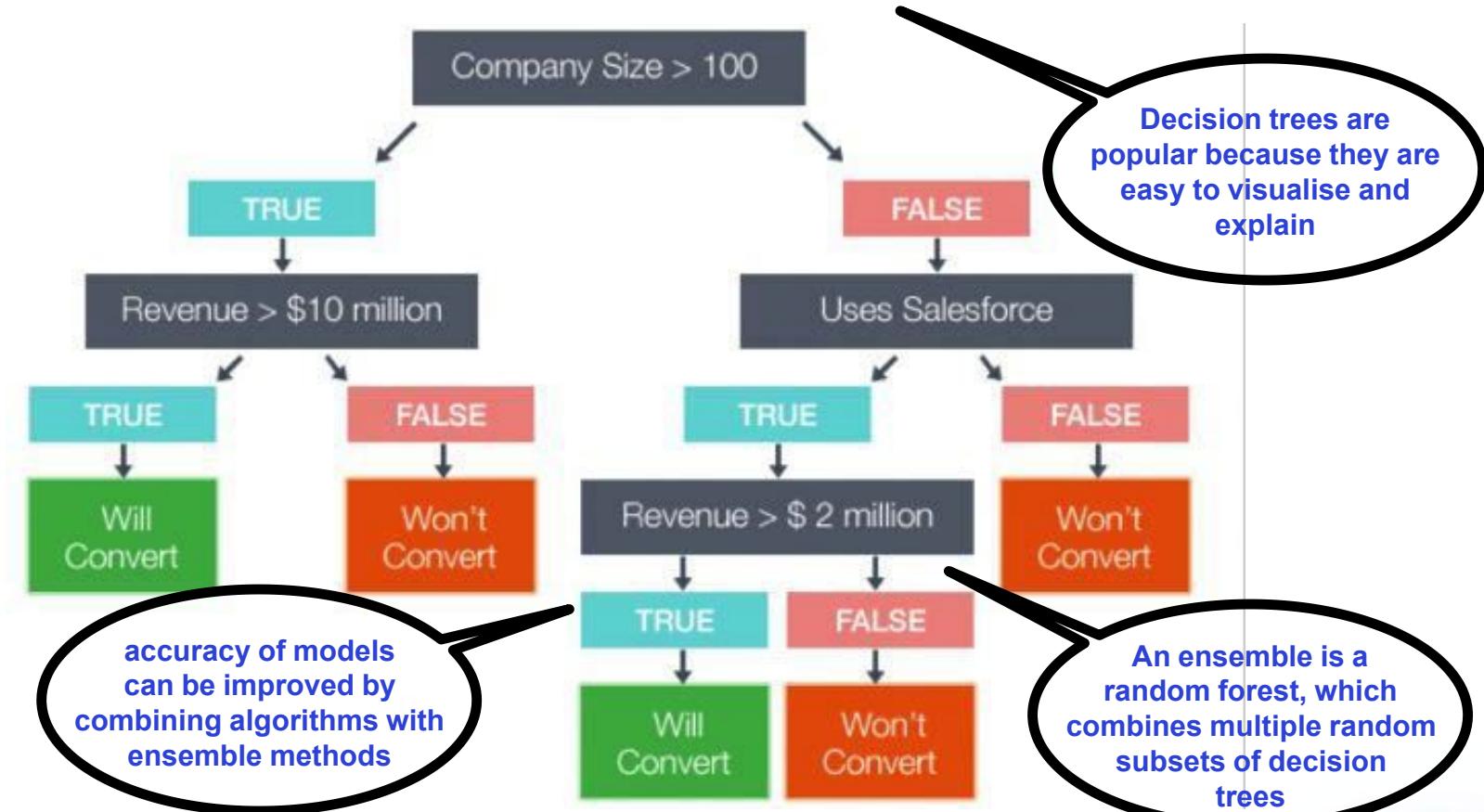
work by evaluating a question containing a feature at every node and selecting a branch to the next node based on the answer



"yes" or "no" are the branches in the tree

Data Processing & Analysis

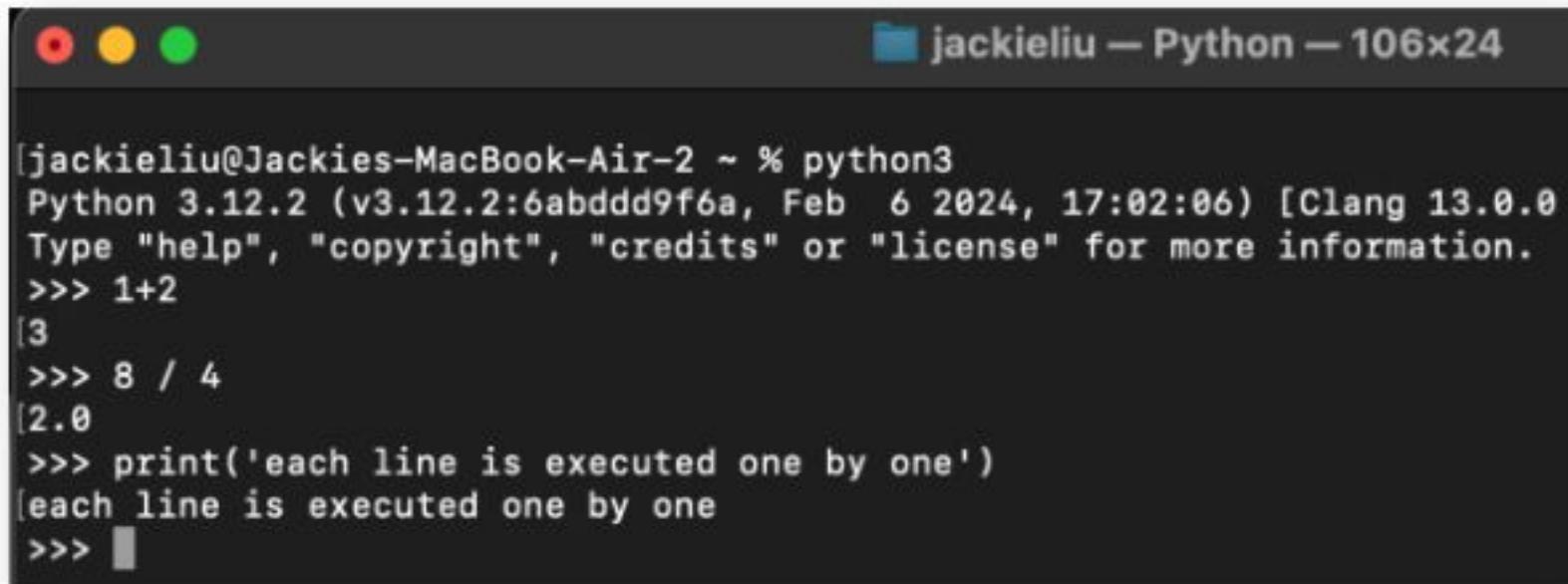
Decision Trees: Ensemble Random Forests



Source: <http://dataeconomy.com/2015/05/predictive-machine-learning-behind-the-scenes-at-flipkart-and-predictions-for-the-future-of-martech/>

How to run Python – In the old days

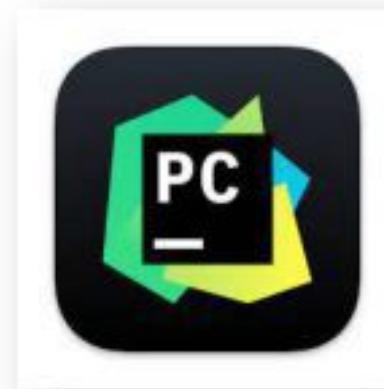
- Python interpreter
 - This is the core of Python
 - Input a line of codes → Press Enter → Run



```
[jackieliu@Jackies-MacBook-Air-2 ~ % python3
Python 3.12.2 (v3.12.2:6abddd9f6a, Feb  6 2024, 17:02:06) [Clang 13.0.0
Type "help", "copyright", "credits" or "license" for more information.
>>> 1+2
[3
>>> 8 / 4
[2.0
>>> print('each line is executed one by one')
each line is executed one by one
>>> ]
```

How to run Python – IDE

- Integrated Development Environment (IDE)
 - A smart editor tailored for writing codes
 - e.g., Syntax highlighting, variable tracking, debugging mode
 - May support useful extensions/plugins (e.g., GitHub Copilot)
 - Common IDEs
 - e.g., Spyder, PyCharm, Visual Studio Code



How to run Python – Jupyter Notebook

- Google Colaboratory (Colab)
 - Jupyter Notebook powered by Google Cloud ¹
 - <https://colab.research.google.com/>
 - Benefits ²
 - “Zero configuration required”
 - Easy sharing
 - Free access: CPU & GPU

Google Colaboratory



1. Google Colab FAQ. <https://research.google.com/colaboratory/faq.html>

2. Google Colaboratory. <https://colab.research.google.com>

Software Introduction

Colaboratory, or Colab for short, is a Google Research product, which allows developers to write and execute Python code through their browser. Google Colab is an excellent tool for deep learning tasks. It is a hosted Jupyter notebook that requires no setup and has an excellent free version, which gives free access to Google computing resources such as GPUs and TPUs.

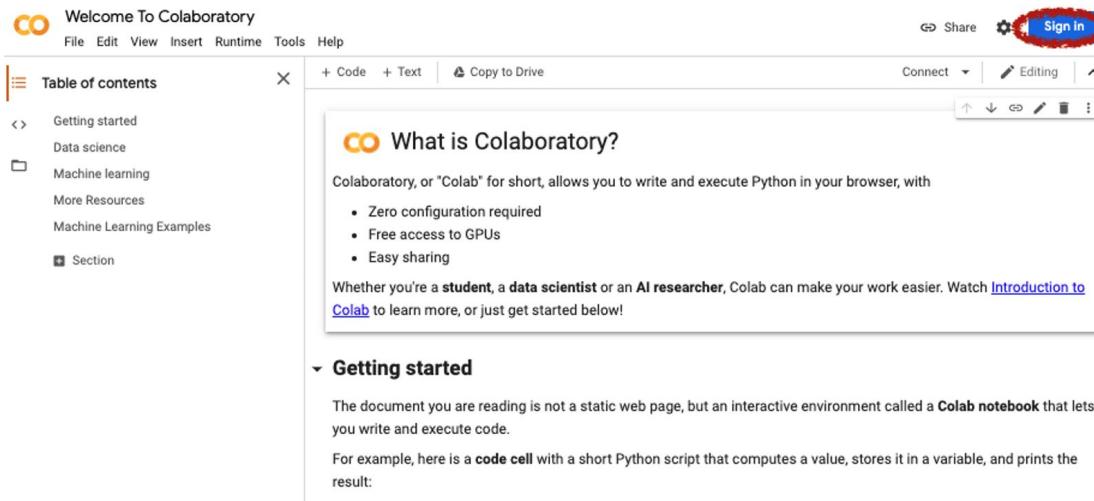
Why Should I Use Google Colab?

There are several reasons to opt to use Google Colab instead of a plain Jupyter Notebook instance:

- Pre-Installed Libraries
- Saved on the Cloud
- Collaboration
- Free GPU and TPU Use

Google Colab Setup

1. Visit the [Google Colab](http://colab.research.google.com/) (<http://colab.research.google.com/>) page, which will direct you to the [Google Colaboratory Welcome Page](#).
2. Click the **Sign in** button on the right top.



Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

Table of contents

- Getting started
- Data science
- Machine learning
- More Resources
- Machine Learning Examples
- Section

+ Code + Text Copy to Drive Connect Editing

What is Colaboratory?

Colaboratory, or "Colab" for short, allows you to write and execute Python in your browser, with

- Zero configuration required
- Free access to GPUs
- Easy sharing

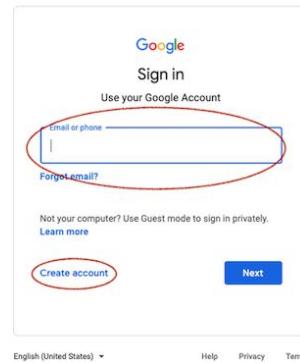
Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or just get started below!

Getting started

The document you are reading is not a static web page, but an interactive environment called a **Colab notebook** that lets you write and execute code.

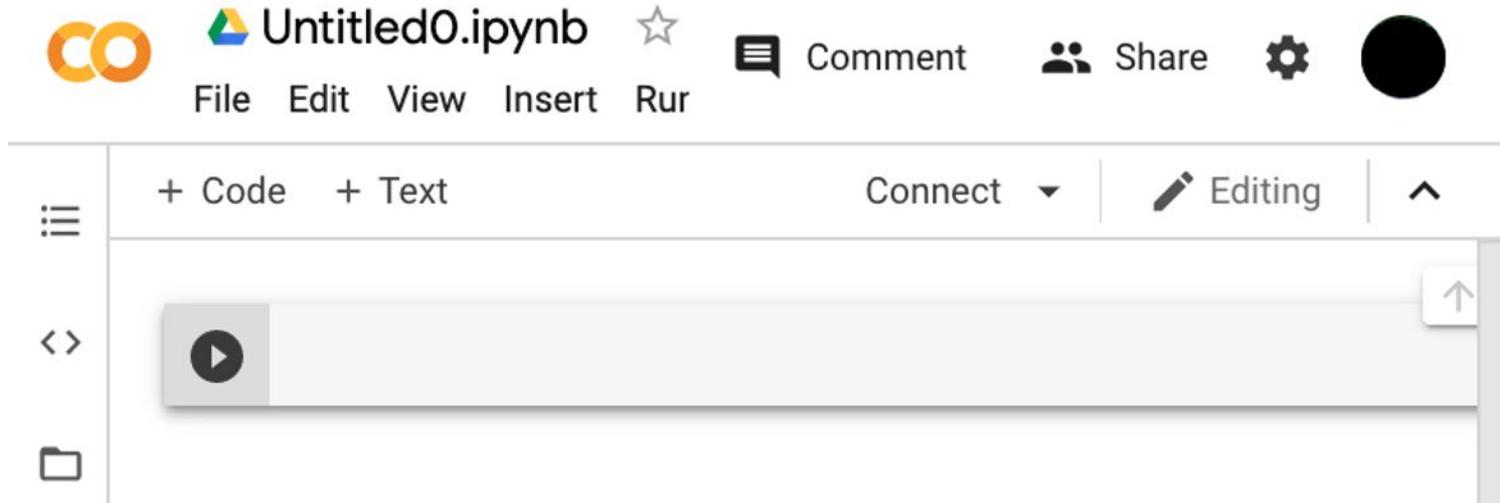
For example, here is a **code cell** with a short Python script that computes a value, stores it in a variable, and prints the result:

3. Sign in with your Gmail account.



Google Colab Setup

4. As soon as you complete the sign-in process, you are ready to use Google Colab.
5. You may easily create a new Colab notebook on this page by clicking *File> New notebook*.



Business Process Automation with Python

Practice 1 – Market Capitalization

- Compute market capitalization of 2 companies
- Show the difference between the two
- Format the result up to 2 decimal places



- Starting notebook:
<https://drive.google.com/file/d/1cVUe4Wcum3jMjVtVi-z2fG5kTHd1dsY/view?usp=sharing>

Data Processing & Analysis

Decision Trees: Ensemble Random Forests

Practice 2 – Run machine learning model

Colab Lab Materials :

https://colab.research.google.com/drive/1UQUaZhRj5ThSCXN50zW_SXUk3wqLPDIQ?usp=sharing

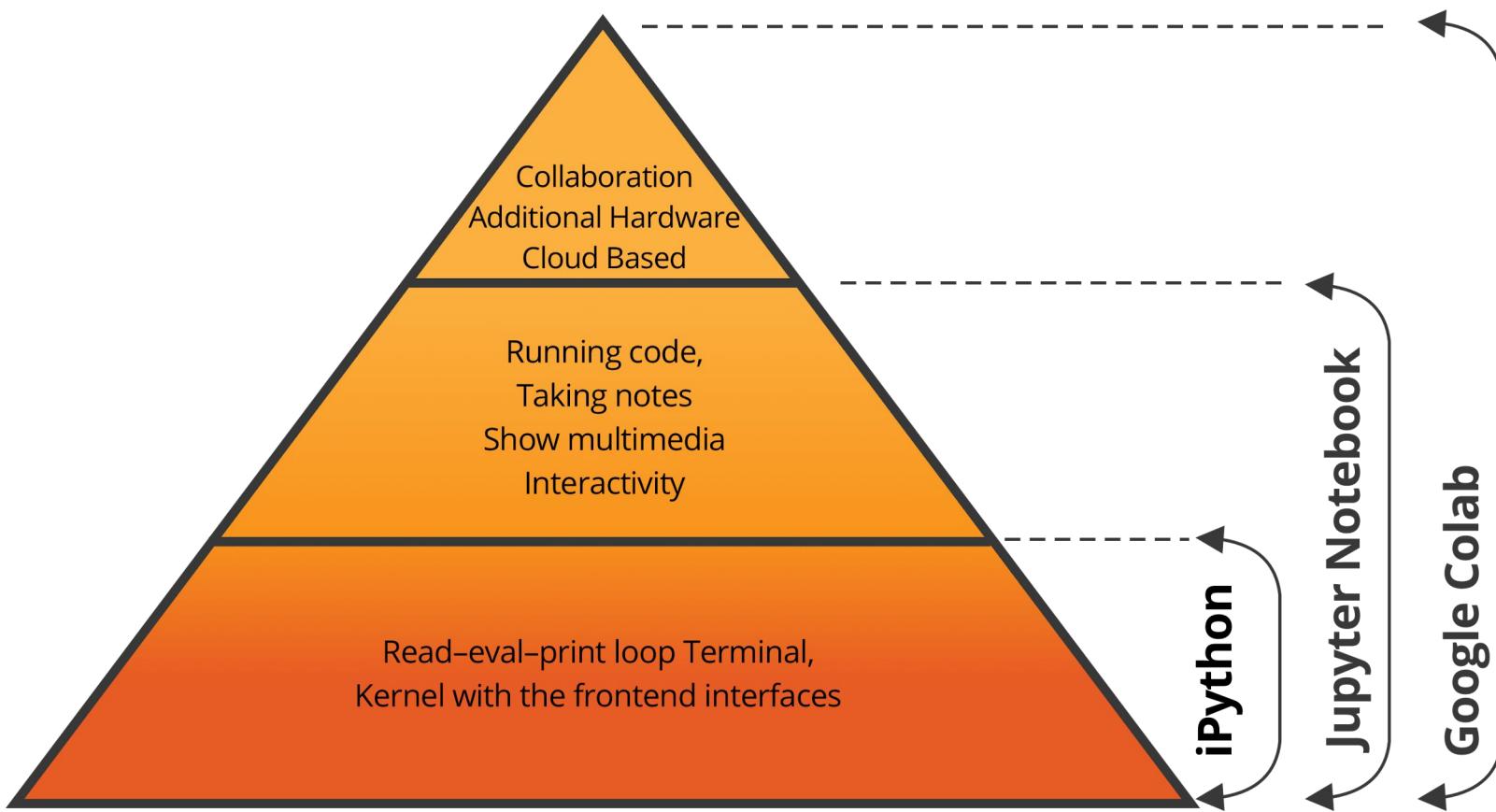
Data :

C:\Users\Eddie Chow\Downloads\archive\data.csv

How to Fit a Decision Tree Model using Scikit-Learn

<https://www.kdnuggets.com/2020/04/visualizing-decision-trees-python.html>

The relationship between iPython, Jupyter Notebook, and Google Colab



Data Processing & Analysis

Quick Test

Which Data Type is Structured/Unstructured,
Discrete/Continuous?

E-mails

Digital Images

Stock Market Logs

Historical Gold Prices

Credit Approval Records

Social Media Friend Relationships

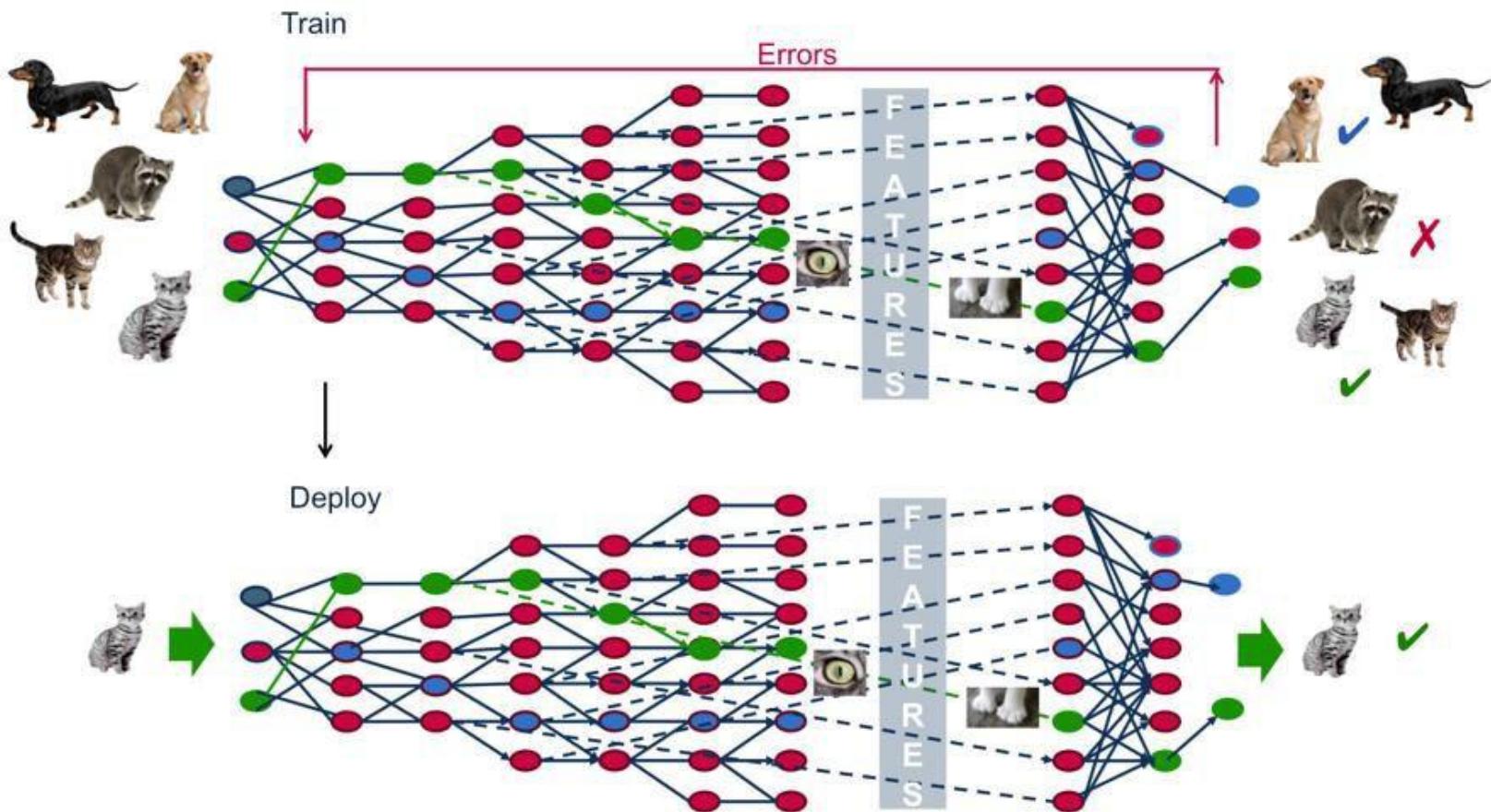
Tweets & Trending Topics

Sales Records

Source: Practical Data Analysis - Second Edition Hector Cuesta, Dr. Sampath Kumar OACKT Publishing BIRMINGHAM - MUMBAI

Data Protection & Information Governance

Deep Learning (multilayered neural networks)

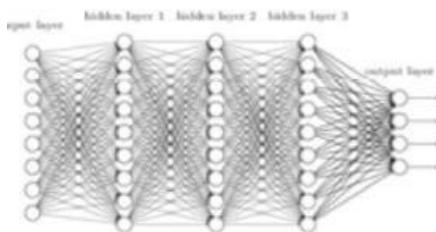


Data Protection & Information Governance

Deep Learning Algorithms

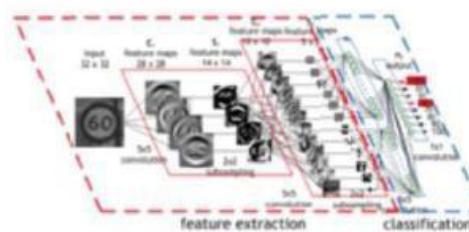
providing lift for classification and forecasting models

Deep Neural Networks



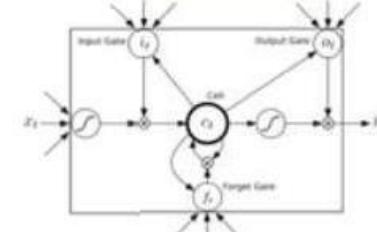
feature extraction and classification of images

Convolutional Neural Networks



for sequence of events, language models, time series, etc.

Recurrent Neural Networks



Data Protection & Information Governance

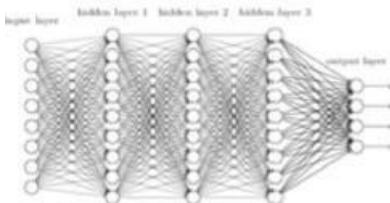
Deep Learning Algorithms

providing lift for
classification and
forecasting models

Deep Neural Networks

Deep neural networks for improved traditional algorithms.

- **Finance:** Enhanced fraud detection through identification of more complex patterns.
- **Manufacturing:** Enhanced identification of defects based on deeper anomaly detection.

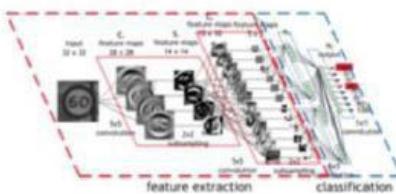


Data Protection & Information Governance

Deep Learning Algorithms

feature extraction
and classification of
images

Convolutional Neural Networks



Convolutional neural networks for images.

- **Retail:** In-store activity analysis of video to measure traffic.
- **Satellite images:** Labeling terrain and classifying objects.
- **Automotive:** Recognition of roadways and obstacles.
- **Healthcare:** Diagnostic opportunities from x-rays, scans, etc.
- **Insurance:** Estimating claim severity based on photographs.

Big Data Tools and Algorithm Exploration

The Problem Space

Machines are '*unable*' to classify images by recognising patterns, generalise from prior knowledge, nor adapt to different image environments



What We See

Humans recognise these dogs naturally, we've instinctively done this since we were a baby

Image classification is the task of taking an input image and outputting a class or a probability of classes that best describes the image (a cat, dog)

```
08 02 22 97 38 15 00 40 00 75 04 05 07 78 52 12 50 77 91 08  
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00  
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 36 65  
52 70 95 23 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91  
22 31 16 71 51 67 63 89 41 92 36 54 22 40 40 28 66 33 13 80  
24 47 32 60 99 03 45 02 44 75 33 53 78 36 84 20 35 17 12 50  
32 98 81 28 64 23 67 10 26 38 40 67 59 54 70 66 18 38 64 70  
67 26 20 68 02 62 12 20 95 63 94 39 63 08 40 91 66 49 94 21  
24 55 58 05 66 73 99 26 97 17 78 78 96 83 14 88 34 89 63 72  
21 36 23 09 75 00 76 44 20 45 35 14 00 61 33 97 34 31 33 95  
78 17 53 28 22 75 31 67 15 94 03 80 04 62 16 14 09 53 56 92  
16 39 05 42 96 35 31 47 55 58 88 24 00 17 54 24 36 29 85 57  
86 56 00 48 35 71 89 07 05 44 44 37 44 60 21 58 51 54 17 58  
19 60 61 68 05 94 47 69 26 73 92 13 86 52 17 77 04 89 55 40  
04 52 08 83 97 35 99 16 07 97 57 32 16 26 26 79 33 27 98 66  
88 36 65 87 57 62 20 72 03 46 33 67 46 55 12 32 63 93 53 69  
04 42 16 73 38 25 39 11 24 94 72 18 08 46 29 32 40 62 76 36  
20 69 36 41 72 30 23 88 34 62 99 69 82 67 59 85 74 04 36 16  
20 73 35 29 78 31 90 01 74 31 49 71 40 86 81 16 23 57 05 54  
01 70 54 71 83 51 54 69 16 92 33 48 61 43 52 01 89 19 67 48
```

What Computers See

source /J. Seng GPU Teaching Kit is licensed by NVIDIA and New York University

Big Data Tools and Algorithm Exploration

Confusion Matrix: Measuring Success for Classification

Prediction:



Image:



True
Positive

True
Negative

False
Negative

False
Positive

True
Negative

True
Positive

- **True Positive:** Correctly identified as relevant
- **True Negative:** Correctly identified as not relevant
- **False Positive:** Incorrectly labeled as relevant
- **False Negative:** Incorrectly labeled as not relevant

+ 'plus' sign is an image that is predicted to be a cat

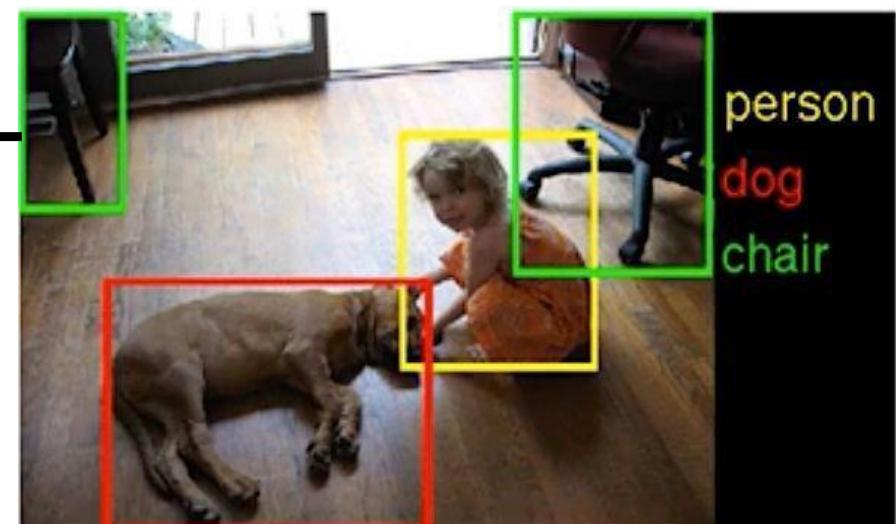
– 'negative' sign is an image that is predicted to be NOT a cat

source /J. Seng GPU Teaching Kit is licensed by NVIDIA and New York University

Big Data Tools and Algorithm Exploration

Convolutional Neural Nets

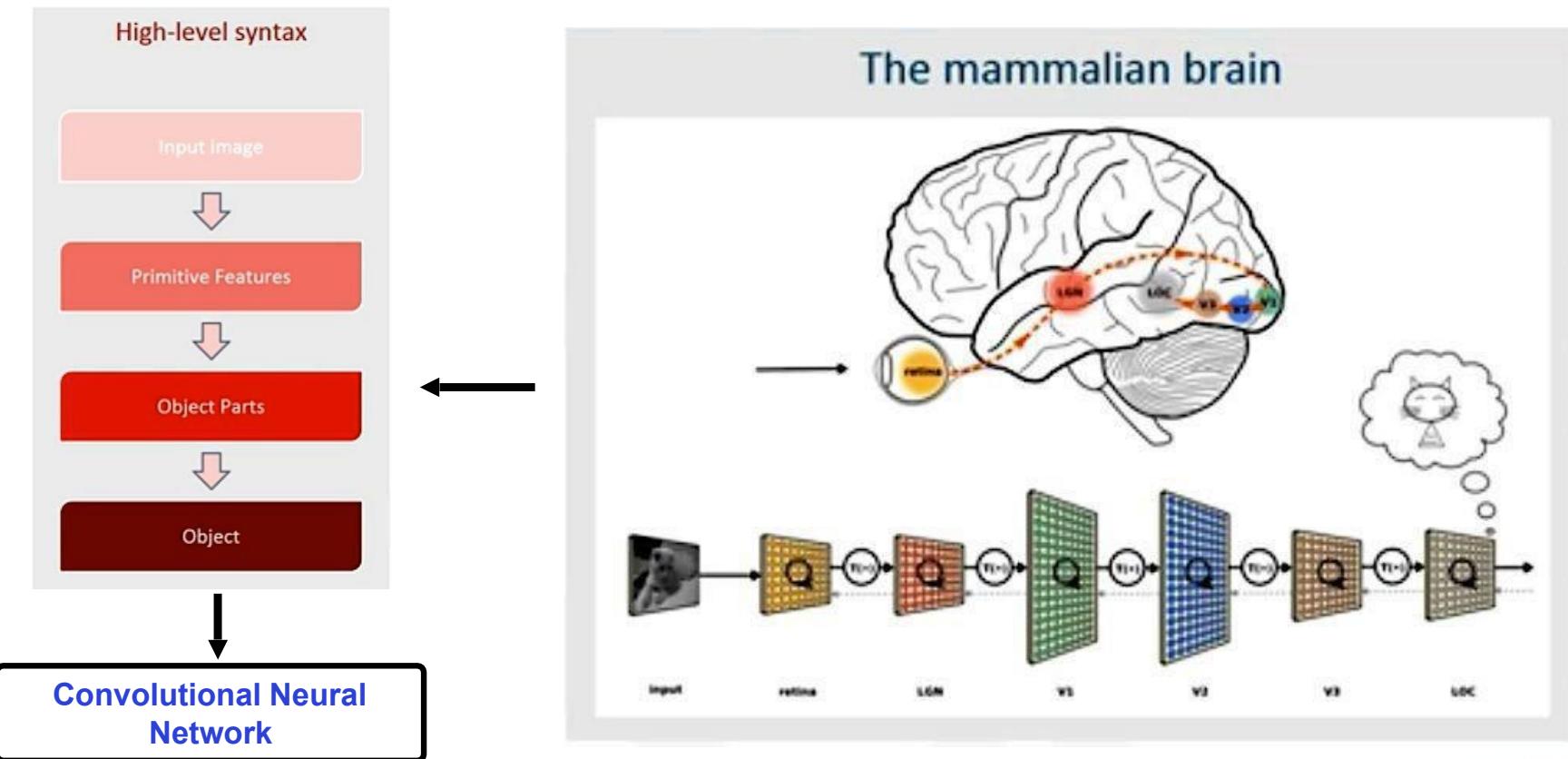
- Signal and image processing
- Handwritten text/digits recognition
- Natural object classification (photos and videos)
- Segmentation
- Face detection
- Recommender systems
- Speech recognition
- Natural Language Processing



Computes the error derivative with respect to a weight for the output layer

Big Data Tools and Algorithm Exploration

Convolutional Neural Nets

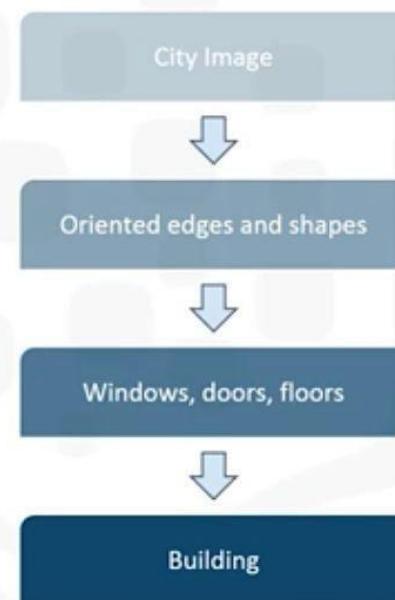


Big Data Tools and Algorithm Exploration

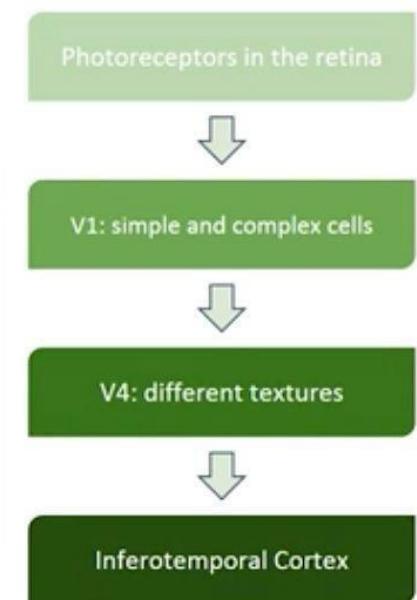
Convolutional Neural Nets



How the scene is organized

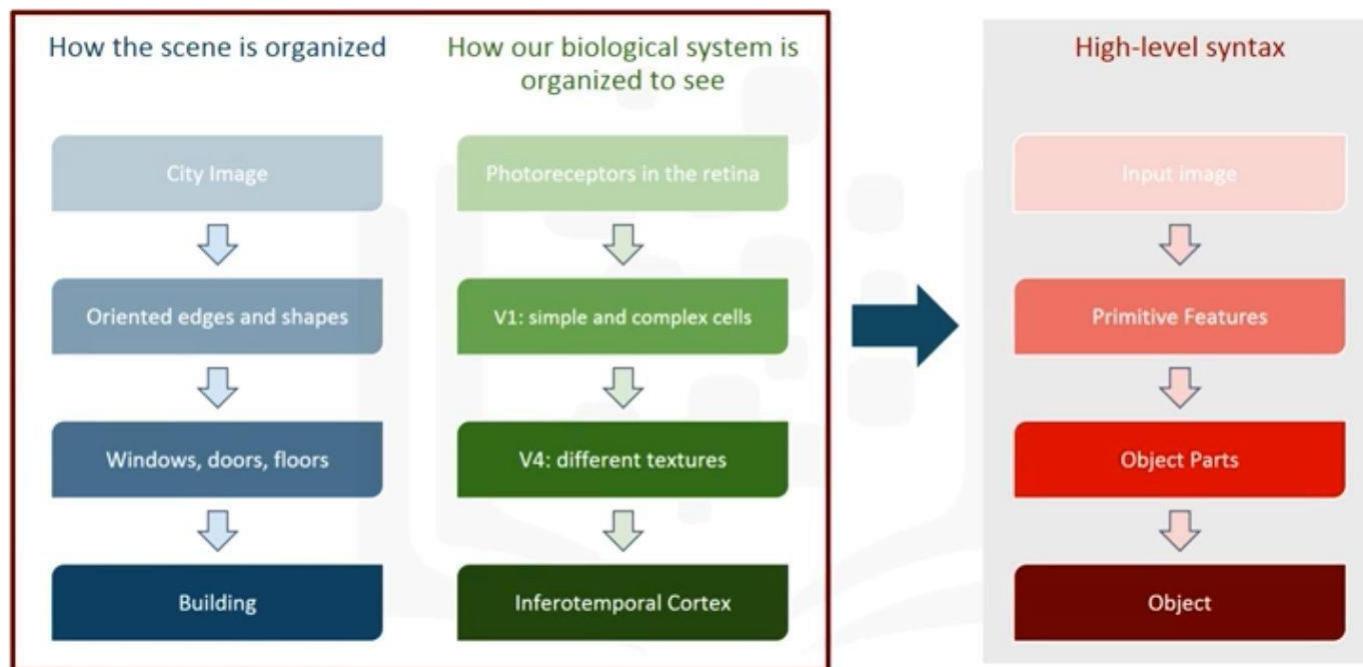


How our biological system is organized to see



Big Data Tools and Algorithm Exploration

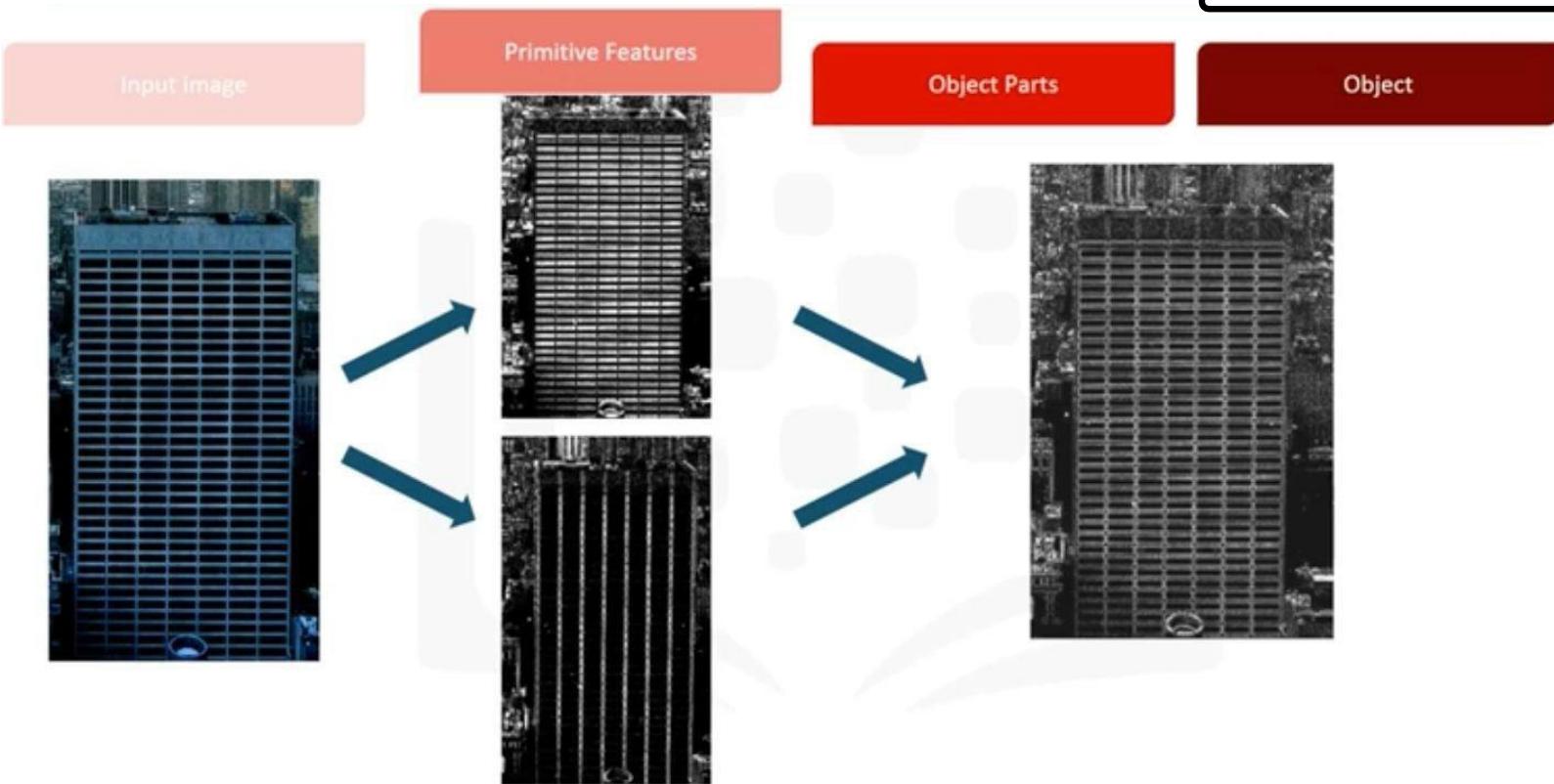
Convolutional Neural Nets



Big Data Tools and Algorithm Exploration

Convolutional Neural Nets

A Visual Explanation



AI Applications

- You could be next Steve Jobs
but please check what could be reused first!
- You can revisit this after you mastered Python 😊
- <https://www.kaggle.com/models>



Introduction to Business Process Automation

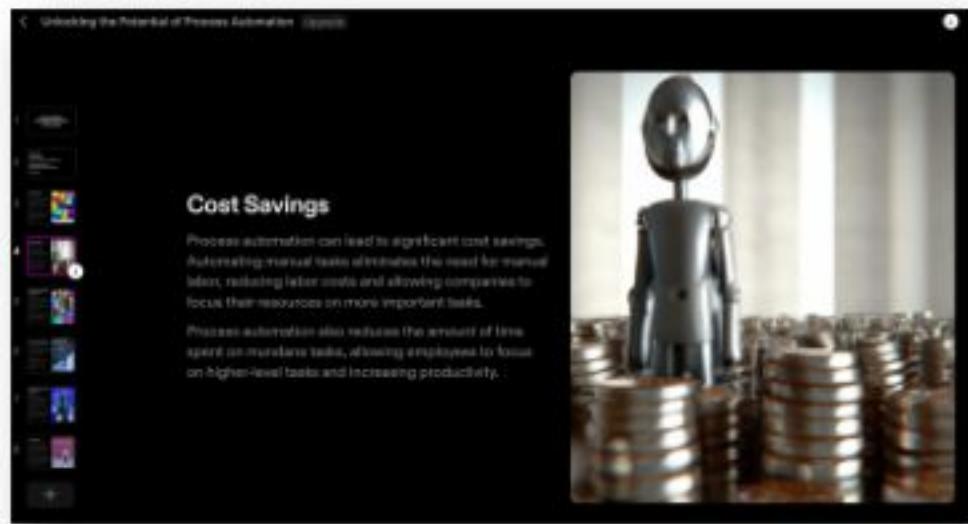
AI Applications – Tome

- Input → Presentation topic
- Output → Slides with outlines and artworks
- <https://tome.app/>

Unlocking the Potential of Process Automation

Introduction
Cost Savings
Improved Customer Experience
Increased Efficiency
Data Collection and Analysis
Conclusion

S - Unlocking the Potential of Process Automation (2020)



Cost Savings

Process automation can lead to significant cost savings. Automating manual tasks eliminates the need for manual labor, reducing labor costs and allowing companies to focus their resources on more important tasks.

Process automation also reduces the amount of time spent on mundane tasks, allowing employees to focus on higher-level tasks and increasing productivity.

PPT AI Generator – SlideGPT

1. Open Browser and enter <https://slidesgpt.com>
2. Enter Prompt “我想創建一個包含商業 ESG 的套牌”

The screenshot shows a web browser displaying a generated PPT slide from SlidesGPT. The slide has a dark blue header with the SlidesGPT logo and navigation links: Pricing, Enterprise plan, Suggest a feature, Contact, and Follow @SlidesGPT. The main content area is titled "Agenda" and contains a bulleted list of 15 items related to Business ESG. At the bottom, there are "Download" and "Share" buttons, along with a link to the slide's URL: <https://slidesgpt.com/l/YeBS>.

- Introduction: What is Business ESG?
- Environmental Considerations
- Social Considerations
- Governance Considerations
- Key Metrics: ESG Reporting
- ESG Integration: Benefits and Challenges
- ESG Investing: Trends and Opportunities
- Regulations and Standards
- ESG Performance: Case Studies
- ESG Communication and Transparency
- Driving Change: Role of Leadership
- Conclusion: Embracing ESG for Future Success

<https://slidesgpt.com/l/YeBS>

PPT AI Generator – SlideGPT

What is Business ESG?

ESG, Environmental, Social, Governance, Definition

- **ESG Framework:** ESG refers to Environmental, Social, and Governance factors that measure the sustainability and ethical impact of an investment in a business or industry.
- **Environmental Considerations:** The environmental aspect of ESG encompasses the impact of a company's operations on the environment, including resource usage, emissions, and waste management.
- **Social Considerations:** The social aspect of ESG evaluates a company's relationships with its employees, stakeholders, and the communities in which it operates, focusing on diversity, labor practices, and human rights.
- **Governance Considerations:** The governance aspect of ESG assesses the internal systems and controls of a company, including leadership, executive pay, shareholder rights, and ethical business practices.



Photo by Omar Quezzi on Unsplash

Environmental Considerations

Environmental Impact, Climate Change, Carbon Footprint, Renewable Energy

- **Environmental Impact:** Assessing and mitigating the impacts of business activities on the natural environment, ecosystems, and resource depletion.
- **Climate Change:** Understanding and addressing the contribution of business activities to climate change through the reduction of greenhouse gas emissions and adaptation strategies.
- **Carbon Footprint:** Measuring and managing the total amount of greenhouse gases emitted directly or indirectly by a company, including Scope 1, 2, and 3 emissions.
- **Renewable Energy:** Exploring and investing in sustainable and renewable energy sources to reduce reliance on non-renewable resources and minimize environmental impact.



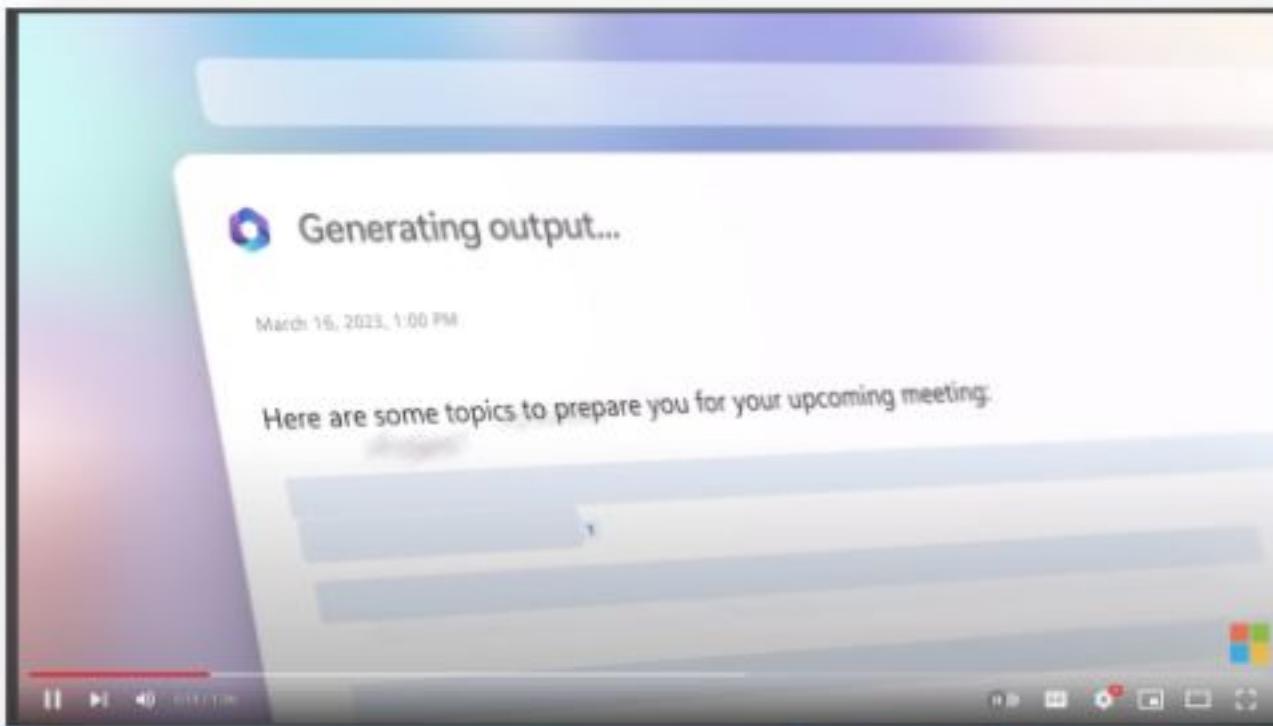
Photo by Tomasz Bazydłowski on Unsplash

<https://slidesgpt.com/I/YeBS>

Introduction to Business Process Automation

AI Applications – Microsoft 365 Copilot

- Virtual Assistant
 - Word / Excel / PowerPoint / Outlook / Teams
- Ad: <https://youtu.be/S7xTBa93TX8>

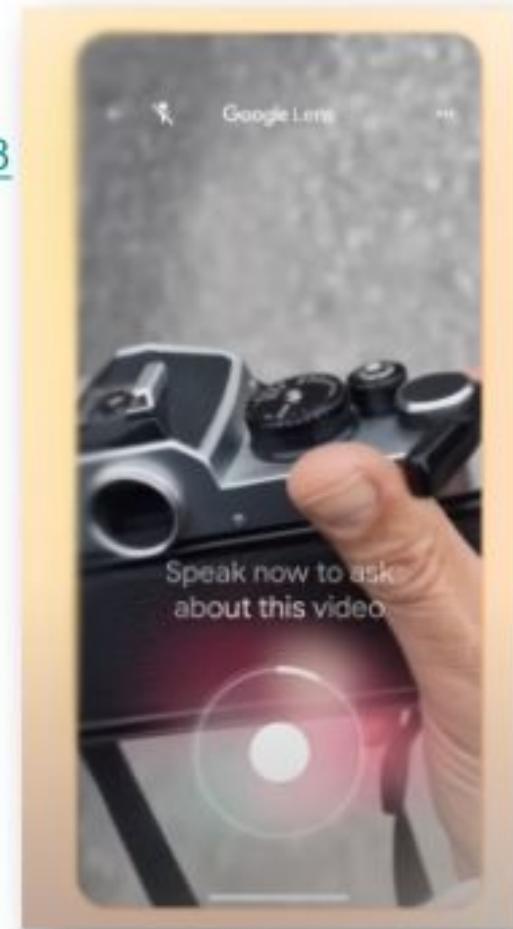
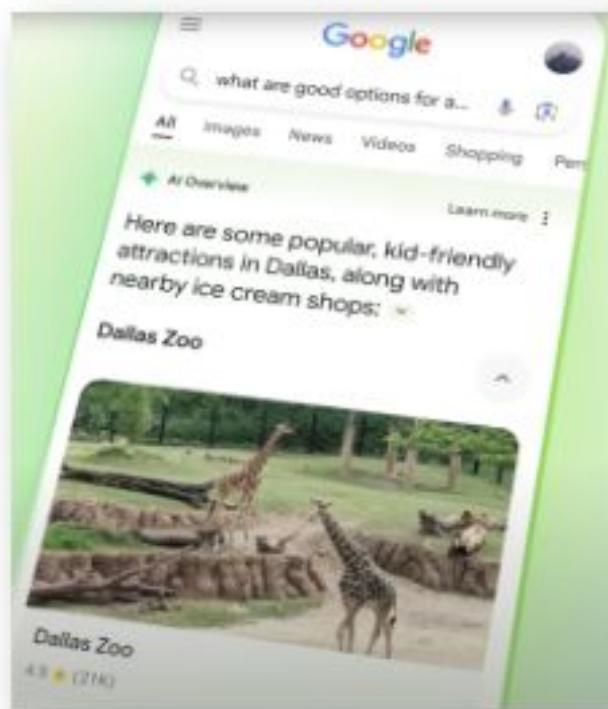


Introduction to Business Process Automation

AI Applications – Gemini

- Geolocation + Image/Video as Inputs
- Ad:

<https://youtu.be/s4InWsd-J6g?si=rwCwgMyHwUb40GAY&t=58>





HKU SPACE
香港大學專業進修學院
HKU School of Professional and Continuing Education

THANK YOU

