

The Developer Advocates present

# SAP CodeJam

2024

> Get Started with  
Machine Learning using SAP HANA



Poznań, Poland  
May 8, 2024



# Witalij Rudnicki

@Sygymundowych

aka Vitaliy Rudnytskiy,  
Віталій Рудницький

- 15 years **tech consultant** in SAP BI/BW
- SAP **Developer Advocate**  
in **SAP Community & Developer Relations**
- All things Data  
<http://bit.ly/SAPDevsData>
- Based in Wrocław, Poland
- Organizer of local SAP Community meetups and **SAP Inside Track**



# **Events not be missed**



## Devtoberfest:

<https://www.youtube.com/playlist?list=PL6RpkC85SLQDHz97qsNTNAE2jnUKj8X5d>

- ABAP & CAP
- Low-code/No-code
- Data, Analytics, & AI
- User Interface
- Integration Suite

# SAP TechEd

Where ideas get real

The SAP TechEd event brings together developers, practitioners, enterprise architects, and global IT leaders. Joined by SAP experts and partners, they'll unite to explore innovations in app development tools, generative AI, clean core for cloud ERP, and much more.



<https://news.sap.com/2024/01/sap-teched-in-2024-expanded-global-program>:

**SAP TechEd Virtual:** October 8-9, 2024

**SAP TechEd On Tour:**

- North America,
- Europe, the Middle East and Africa (EMEA),
- Latin America and the Caribbean (LAC), and
- Asian-Pacific and Japan (APJ)

**ASUG Tech Connect with SAP TechEd, West Palm Beach, Florida: November 12-14, 2024**

# May Developer Challenge: SAP AI Services

## May Developer Challenge - SAP AI Services



SAP noravonthenen

Developer Advocate



10 Kudos



Wednesday

- last edited Thursday

Welcome to week 1 of the May Developer Challenge on AI at SAP! The topic of this month's challenge are the SAP AI Services; Document Information Extraction and Data Attribute Recommendation. To participate in the challenge you just have to post a screenshot of your solution as a reply in this discussion of the corresponding week.

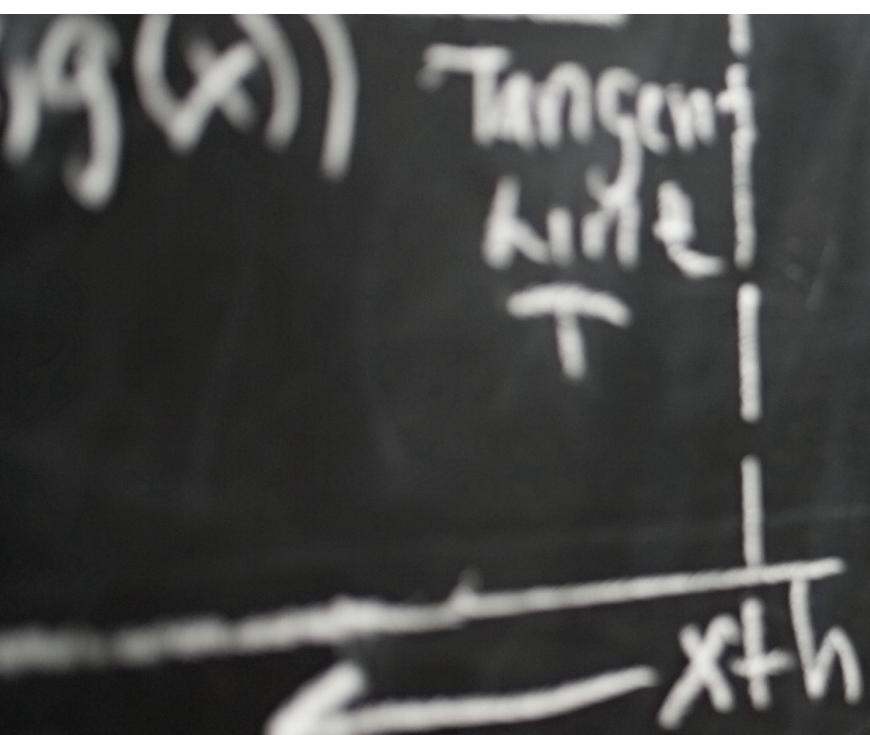
SAP AI Services help you implement custom use cases by providing powerful algorithms specifically tailored to business problems.

source: <https://community.sap.com/t5/application-development-discussions/may-developer-challenge-sap-ai-services/m-p/13692229#M2027917>

# SAP Developers channel: <https://www.youtube.com/@sapdevs>

The screenshot shows a YouTube video player interface. At the top, the SAP Developers channel logo is visible. The main video frame displays two hosts: Thomas Jung, a developer advocate, on the left, and DJ Adams on the right. Thomas Jung is shown against a blue background with a 'SAP' logo. DJ Adams is in an office setting with a window and blinds. A green box highlights the video frame. The bottom of the screen shows a control bar with play buttons, a timestamp (0:45 / 1:01:20), and other video controls. Below the video, a caption reads: "Let's test drive Joule's generative AI features in SAP Build Code together! 1 of 2 (EMEA / APJ)". The video has 34.4K subscribers. The SAP logo and "Powered by Zoom" are visible in the bottom right corner of the video frame. A sidebar on the right contains a "Top chat replay" section.

# A bit of theory before we start



$$f'(x) = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h}$$
$$= \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$$
$$= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h}$$
$$= \lim_{h \rightarrow 0} (2x + h)$$
$$= 2x$$
$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

# **How many of you are familiar with...**

...ChatGPT?

...Machine Learning / Artificial Intelligence?

...SAP HANA database?

...SAP HANA embedded ML (PAL)?

...SAP Business Technology Platform?

...SAP Business Application Studio? ...VS Code?

...Python?

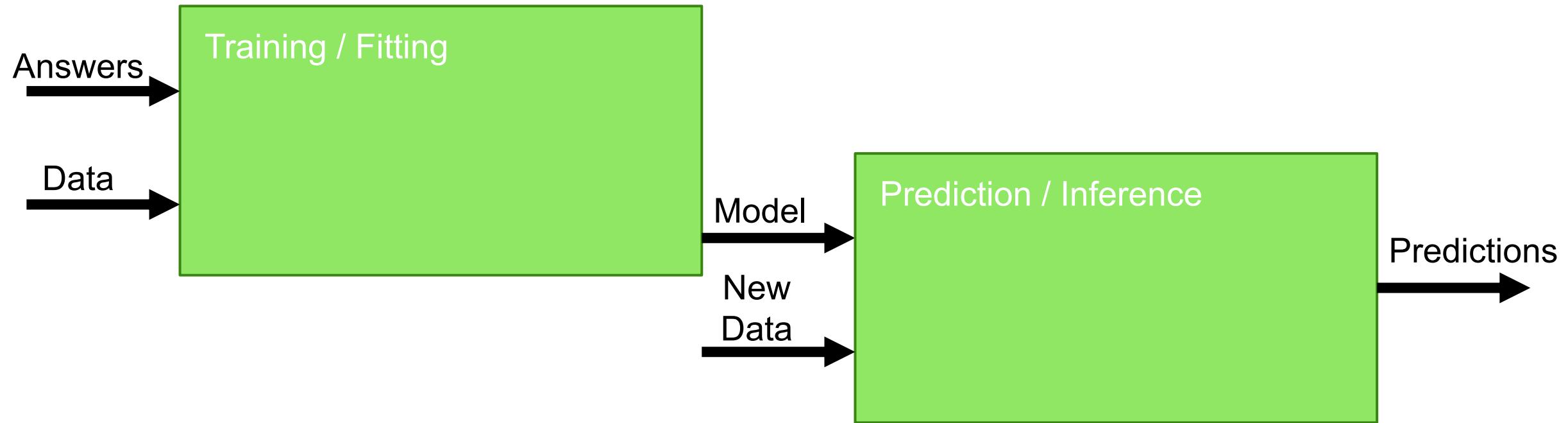
...Jupyter Notebooks?

...Python ML libraries, like `sklearn`?

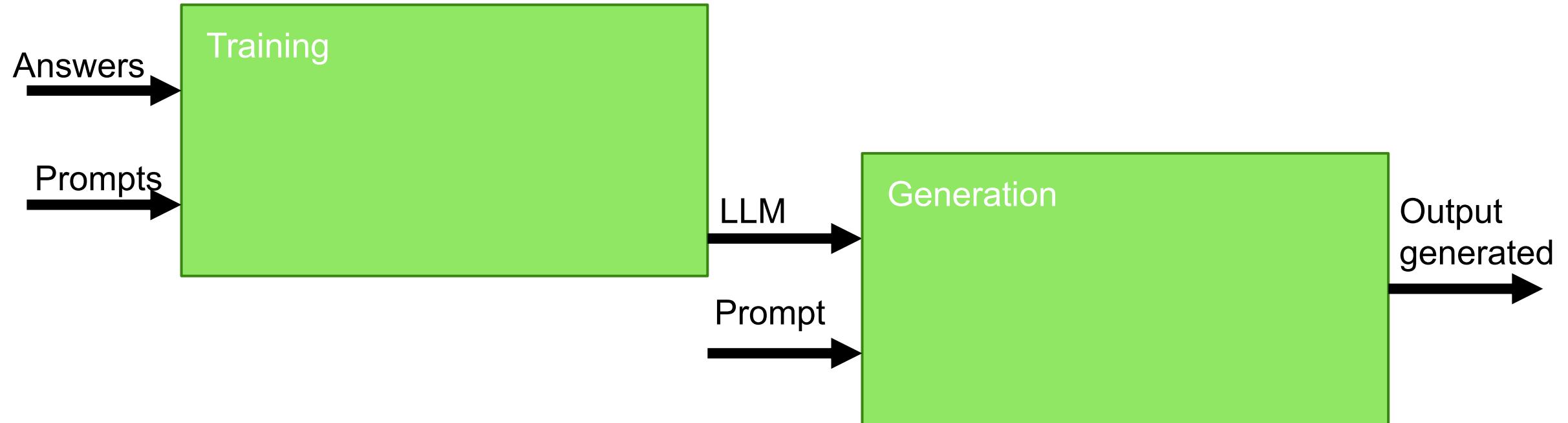
...GitHub?

...Generative AI?

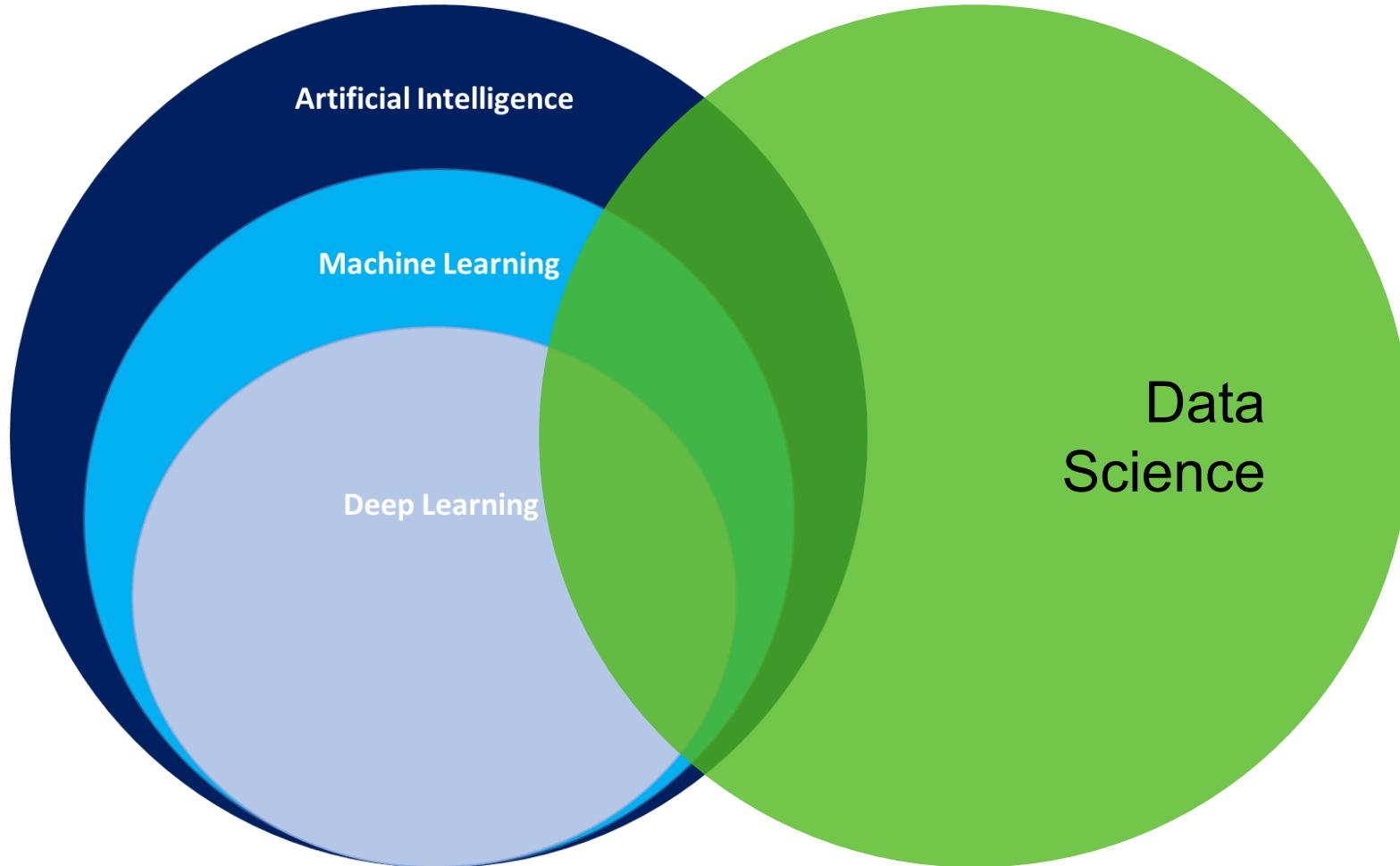
# Machine Learning



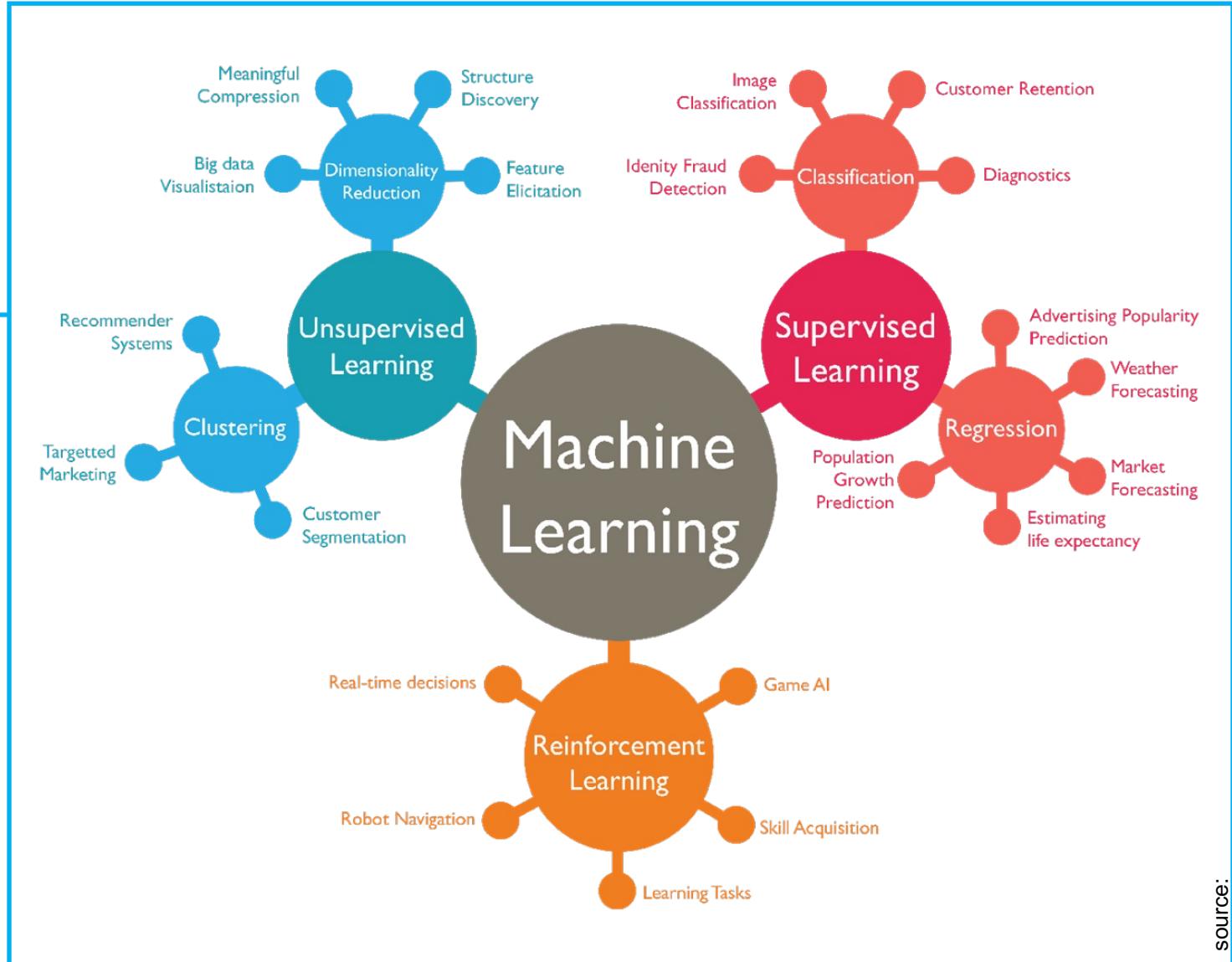
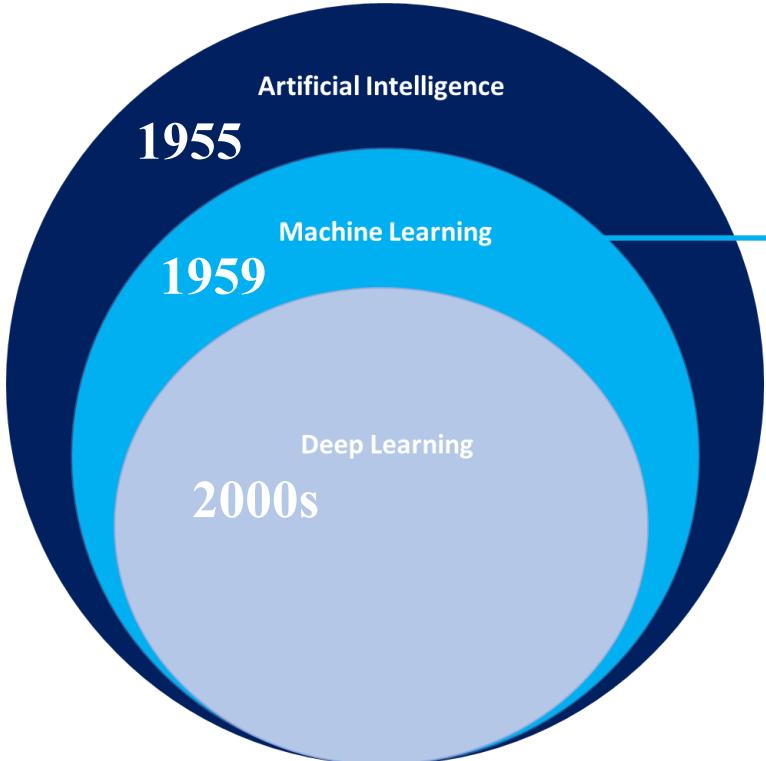
# Gen AI



# Eggs diagram: AI vs ML vs DS



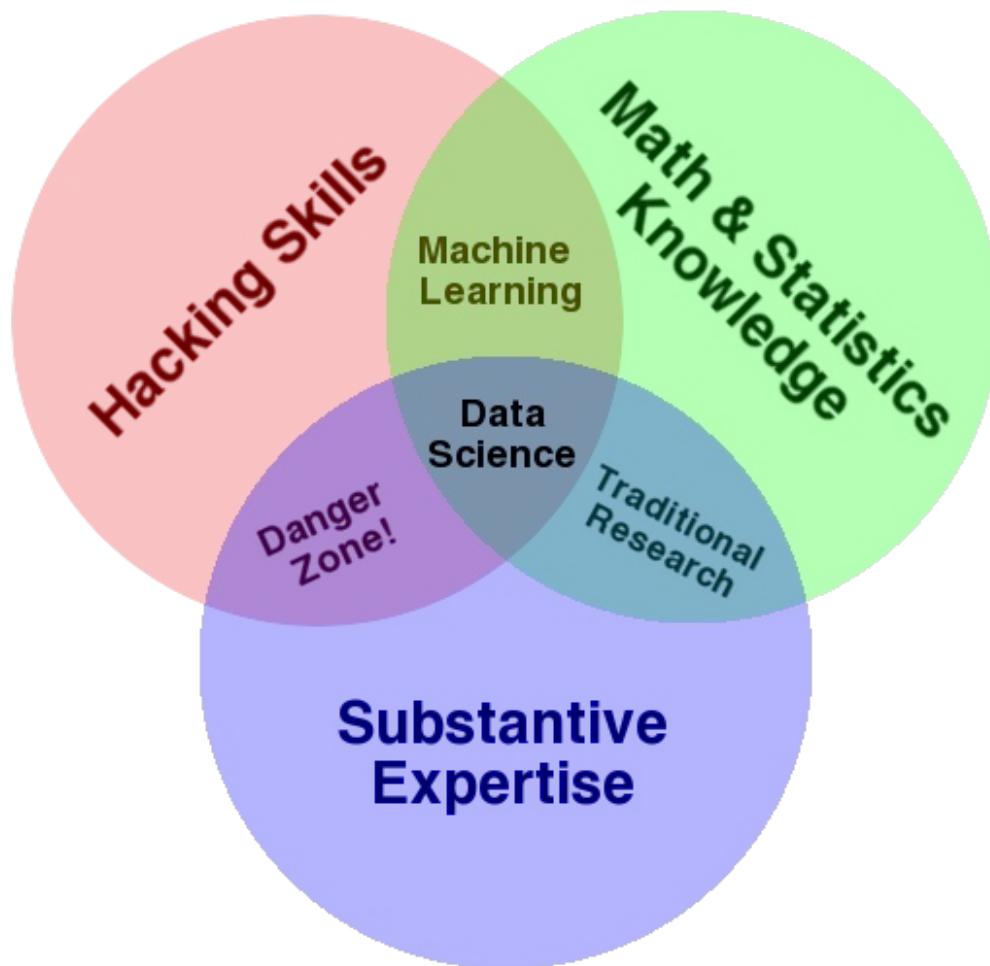
# Machine Learning Terminology



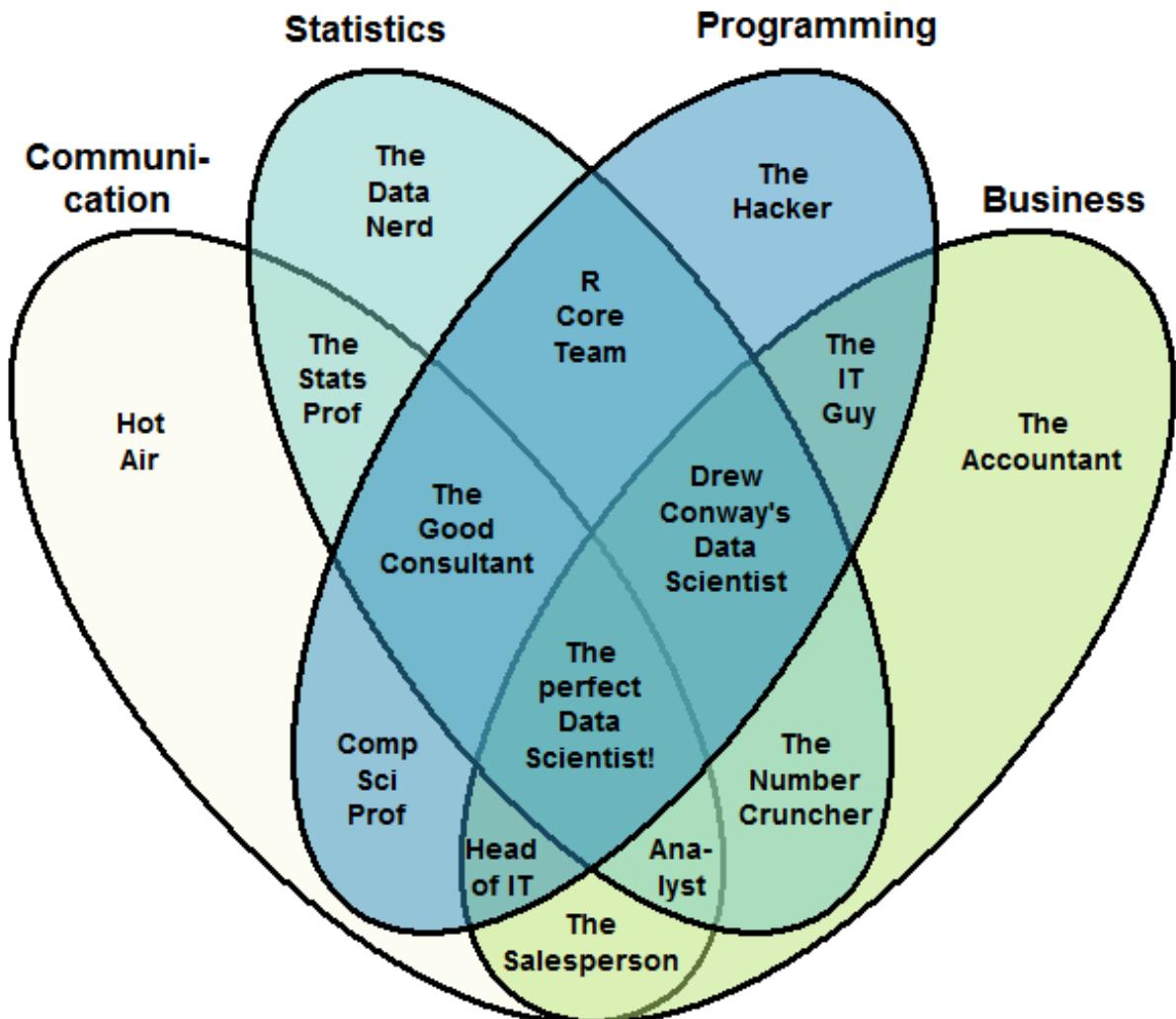
source:

[https://www.researchgate.net/publication/337958773 A Novel Approach for Improving Breast Cancer Risk Prediction using Machine Learning Algorithms A Survey](https://www.researchgate.net/publication/337958773_A_Novel_Approach_for_Improving_Breast_Cancer_Risk_Prediction_using_Machine_Learning_Algorithms_A_Survey)

# Who is the Data Scientist?

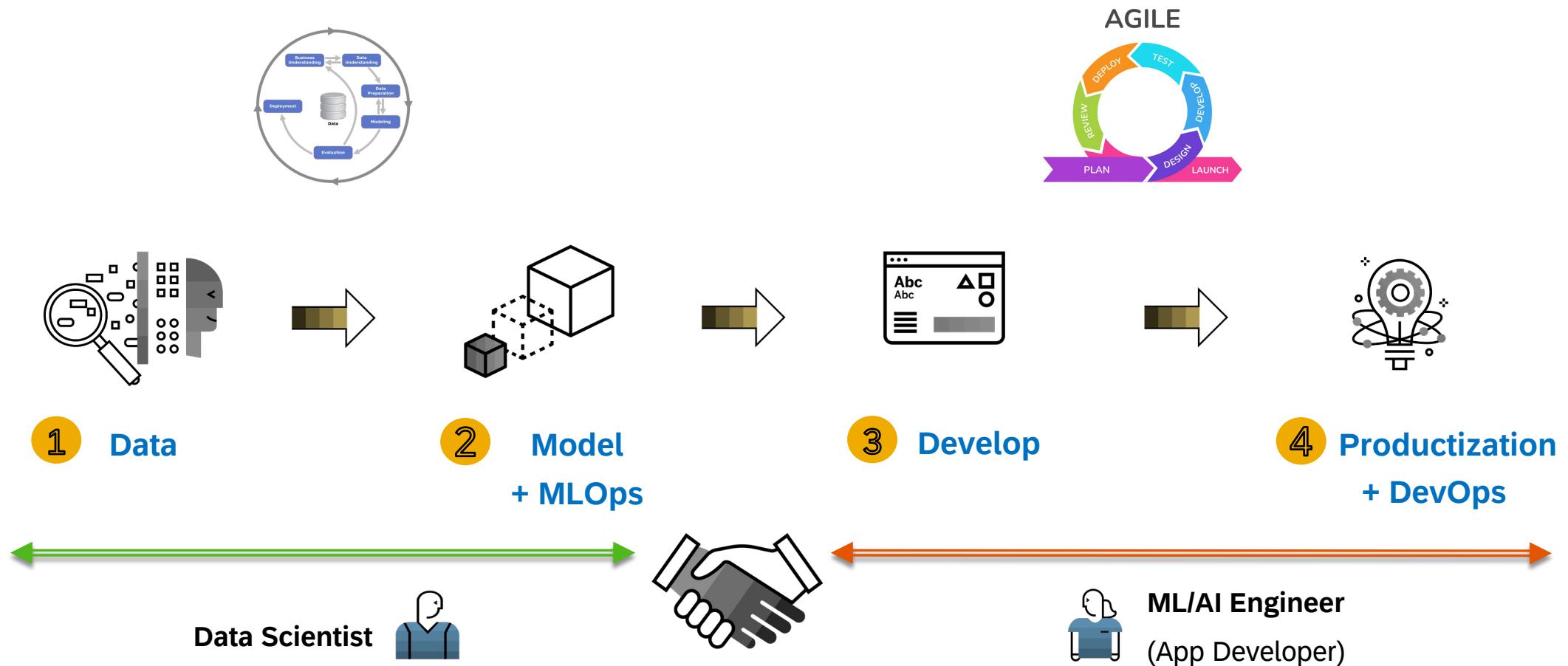


Drew Conway's Data Scientist Venn Diagram

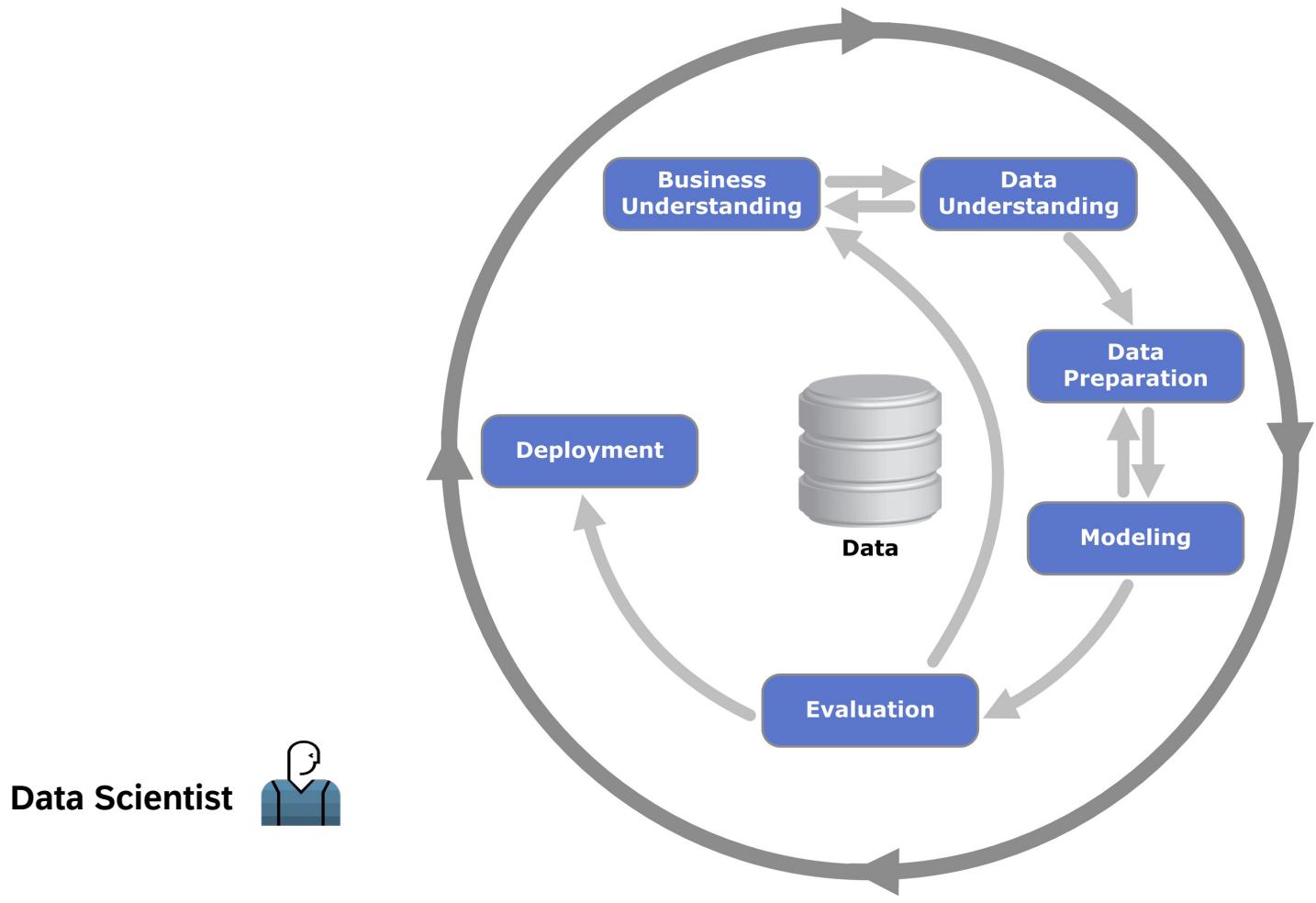


Stephan Kolassa's Data Scientist Venn Diagram

# Development Approach | Building an Intelligent Data Application



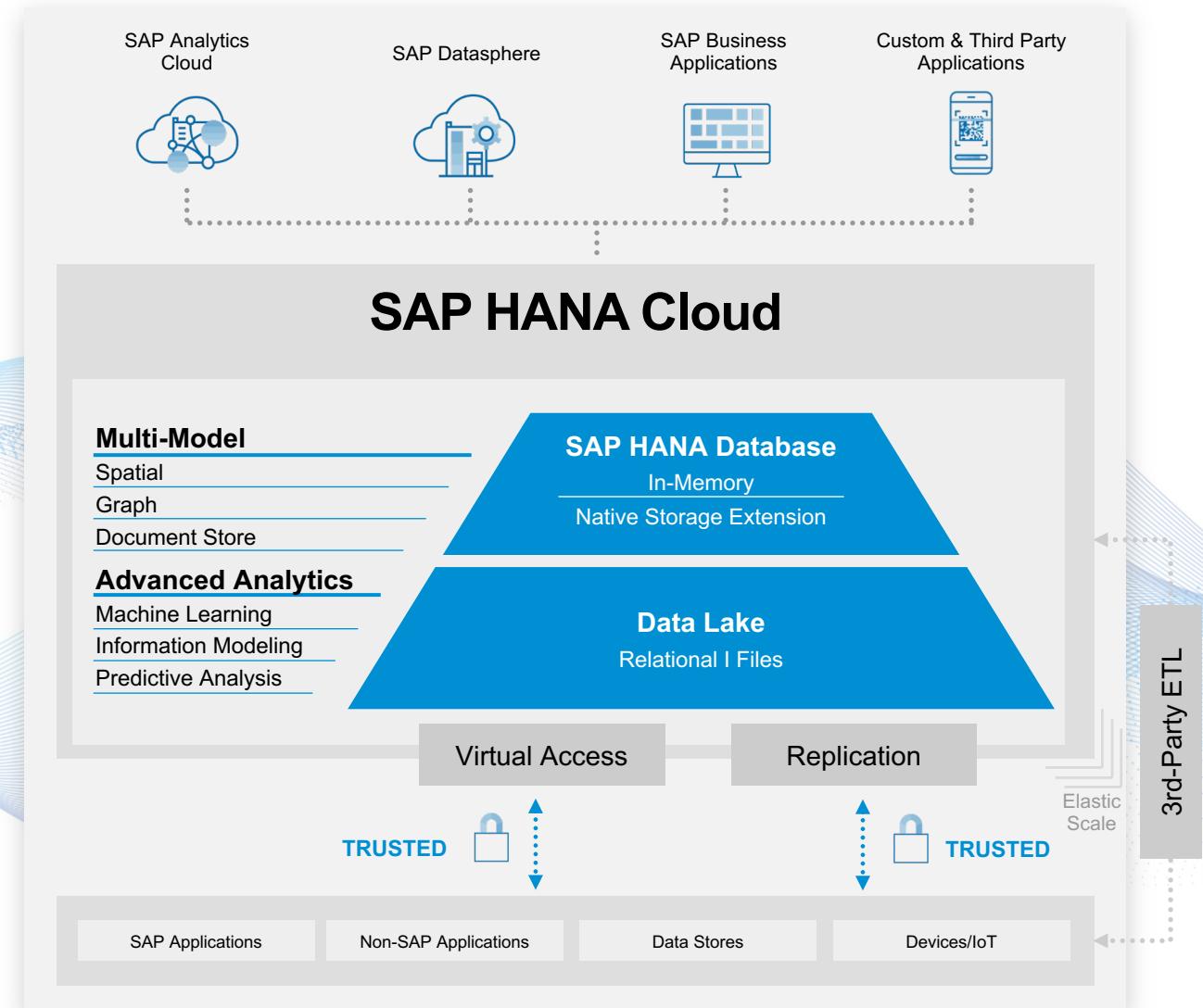
# Development Approach | Training a Model



# SAP HANA Cloud | Foundation for Intelligent Data Applications on SAP BTP

## Power Intelligent Data Applications with SAP HANA Cloud

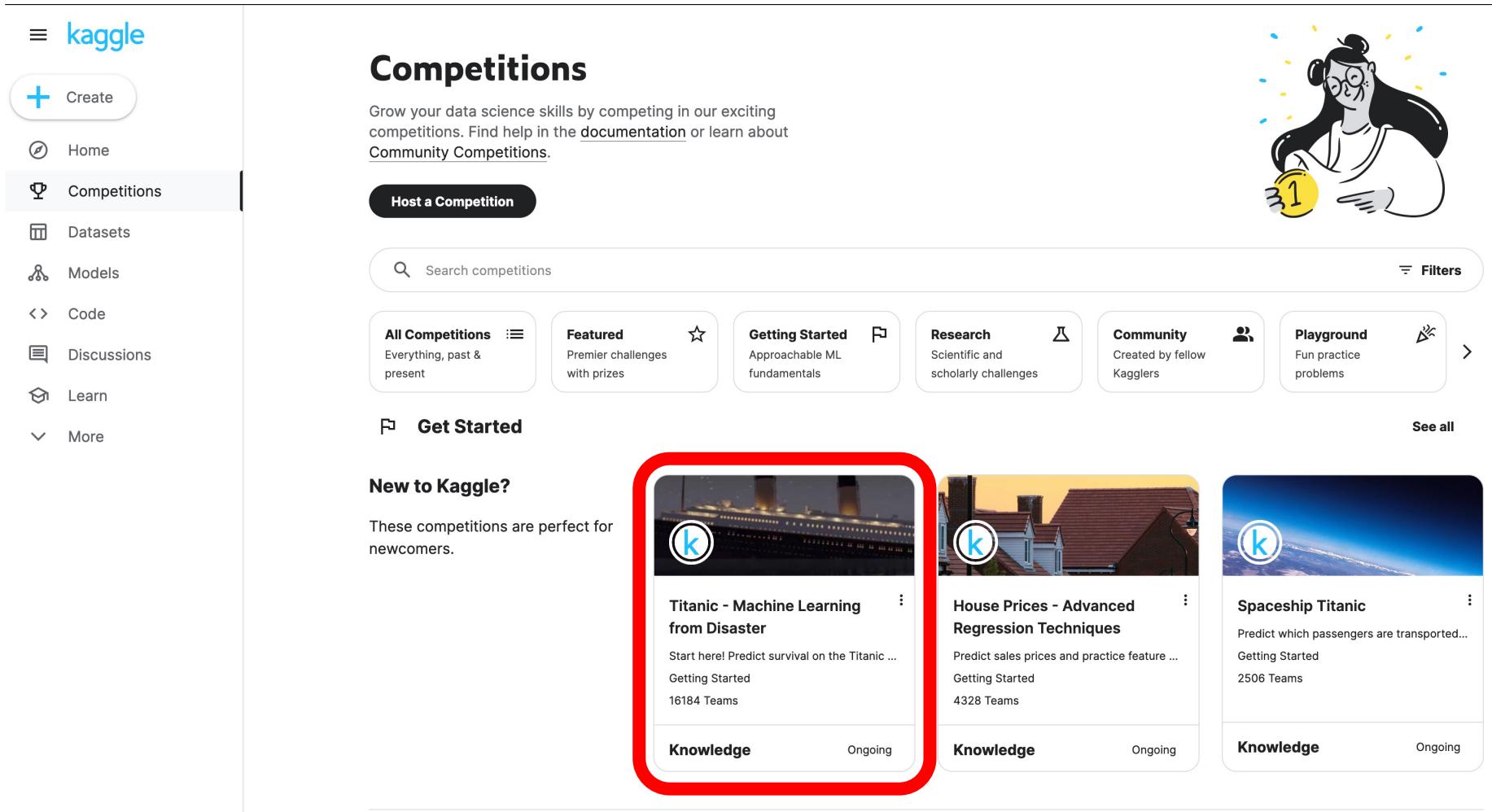
- Running mission-critical applications and analytics with **one solution**
- Seamlessly blend **multi-model data** to enhance business processes
- Federate data **across hybrid systems** and multiple clouds
- Provide **proven in-memory performance** for all data needs



# Let's get busy



# Our ML Challenge today:



The screenshot shows the Kaggle Competitions page. On the left is a sidebar with navigation links: kaggle, Create, Home, Competitions (which is selected), Datasets, Models, Code, Discussions, Learn, and More. The main content area has a title "Competitions" and a sub-section "Host a Competition". Below this is a search bar with "Search competitions" and a "Filters" button. A row of category cards includes "All Competitions" (Everything, past & present), "Featured" (Premier challenges with prizes), "Getting Started" (Approachable ML fundamentals), "Research" (Scientific and scholarly challenges), "Community" (Created by fellow Kagglers), and "Playground" (Fun practice problems). A "Get Started" section follows, featuring three competitions: "Titanic - Machine Learning from Disaster" (highlighted with a red box), "House Prices - Advanced Regression Techniques", and "Spaceship Titanic". Each card shows the competition name, a thumbnail image, a brief description, the number of teams, and the status (Knowledge, Ongoing).

**Competitions**

Grow your data science skills by competing in our exciting competitions. Find help in the [documentation](#) or learn about [Community Competitions](#).

[Host a Competition](#)

Search competitions Filters

**All Competitions** Everything, past & present **Featured** Premier challenges with prizes **Getting Started** Approachable ML fundamentals **Research** Scientific and scholarly challenges **Community** Created by fellow Kagglers **Playground** Fun practice problems

[Get Started](#) See all

**New to Kaggle?**

These competitions are perfect for newcomers.

**Titanic - Machine Learning from Disaster**  
Start here! Predict survival on the Titanic ...  
Getting Started  
16184 Teams  
**Knowledge** Ongoing

**House Prices - Advanced Regression Techniques**  
Predict sales prices and practice feature ...  
Getting Started  
4328 Teams  
**Knowledge** Ongoing

**Spaceship Titanic**  
Predict which passengers are transported...  
Getting Started  
2506 Teams  
**Knowledge** Ongoing

# Exercises



**Repo:** <https://github.com/SAP-samples/hana-ml-py-codejam/>

(→ [https://bit.ly/CJ\\_HANAML](https://bit.ly/CJ_HANAML))

## **1. Pre-requisites:**

<https://github.com/SAP-samples/hana-ml-py-codejam/blob/main/prerequisites.md>

## **2. Exercises:**

<https://github.com/SAP-samples/hana-ml-py-codejam#the-exercises>

## **3. Worth watching:**

<https://github.com/SAP-samples/hana-ml-py-codejam#additional-learning-material>

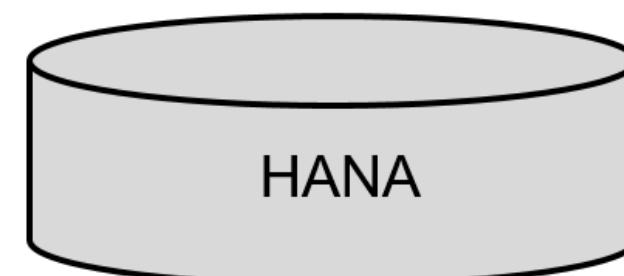
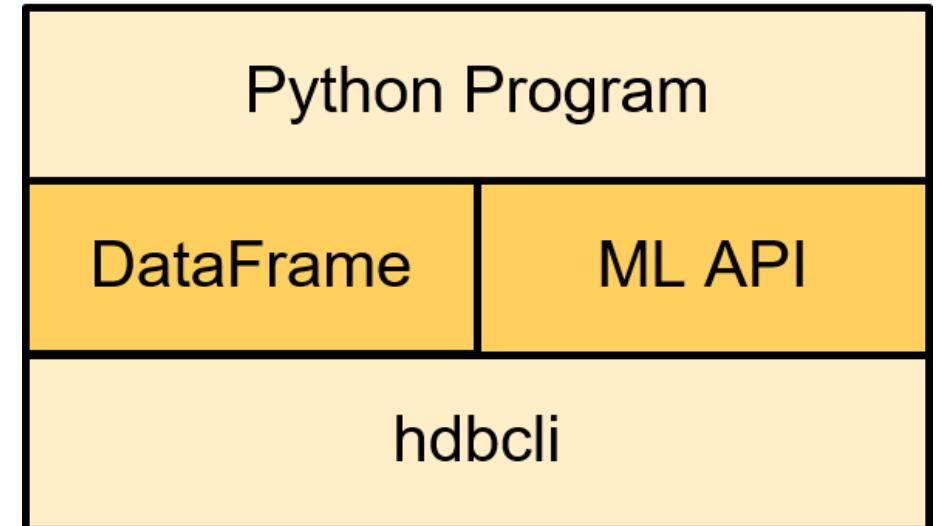
## **4. Try at home in your SAP BTP Trial:**

<https://github.com/SAP-samples/sap-community-developer-challenge-eda-hana>

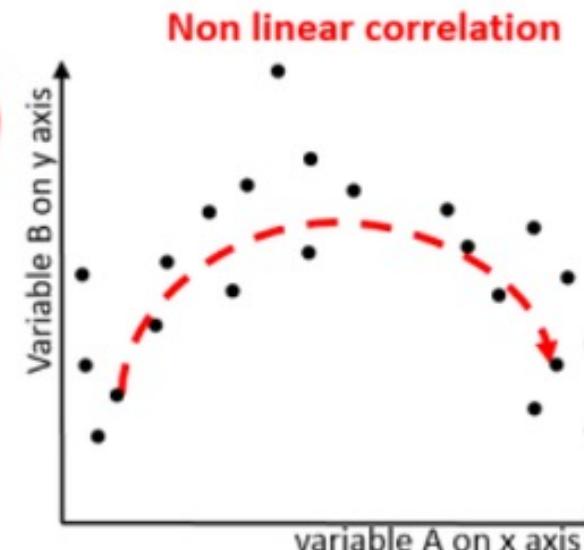
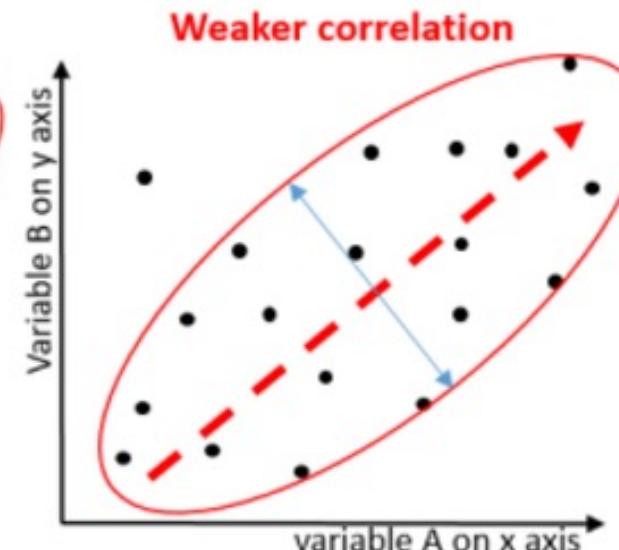
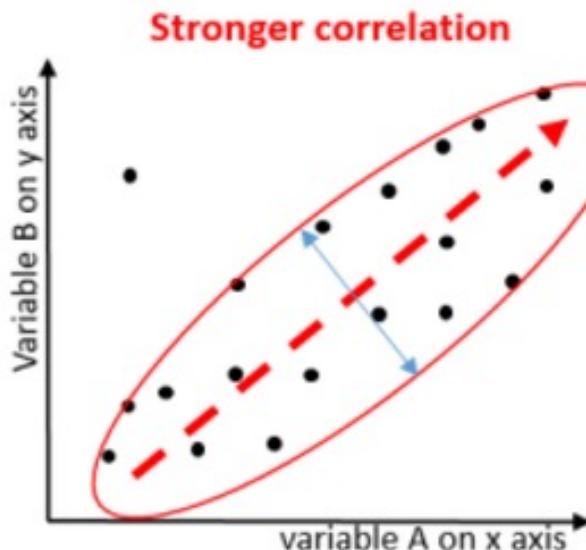
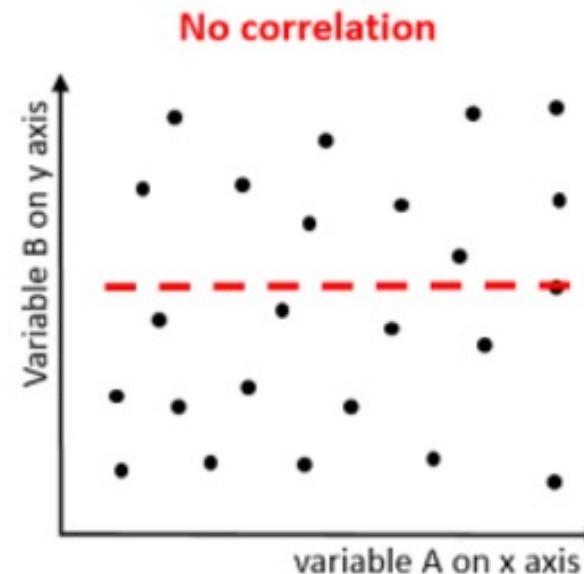
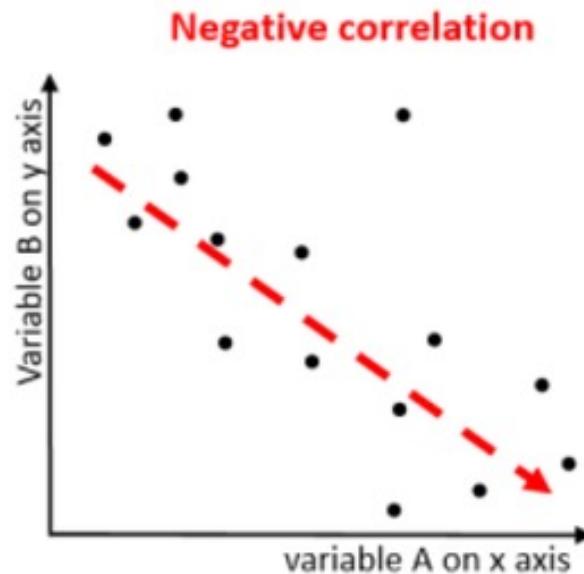
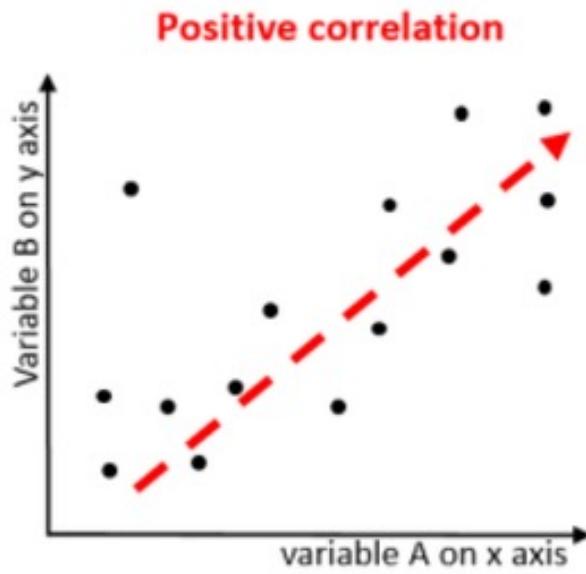
# Understand DataFrame(s)

# Data Scientist using Python

Pandas DataFrame <> HANA DataFrame

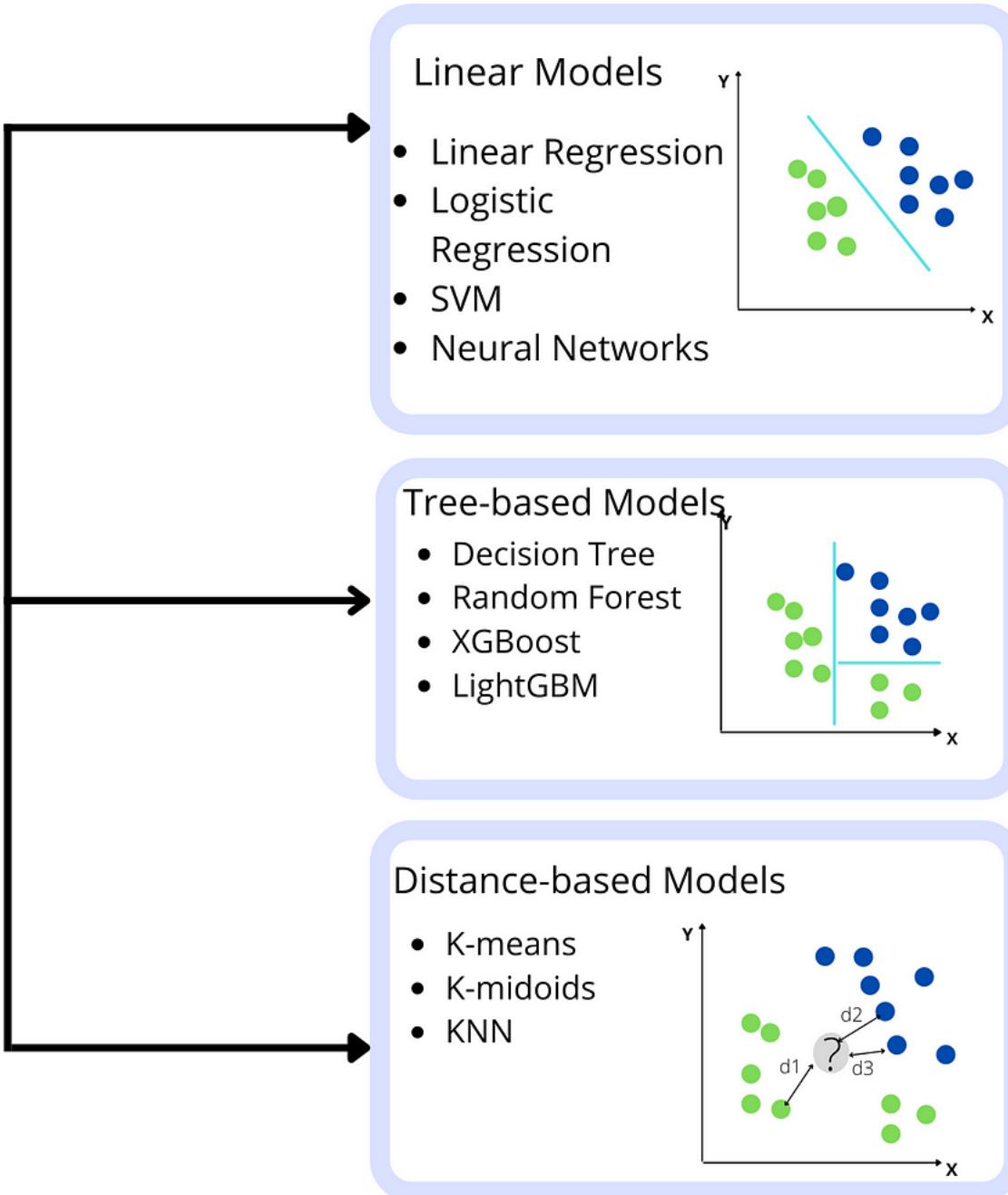


# Correlation based on direction, form, and dispersion strength

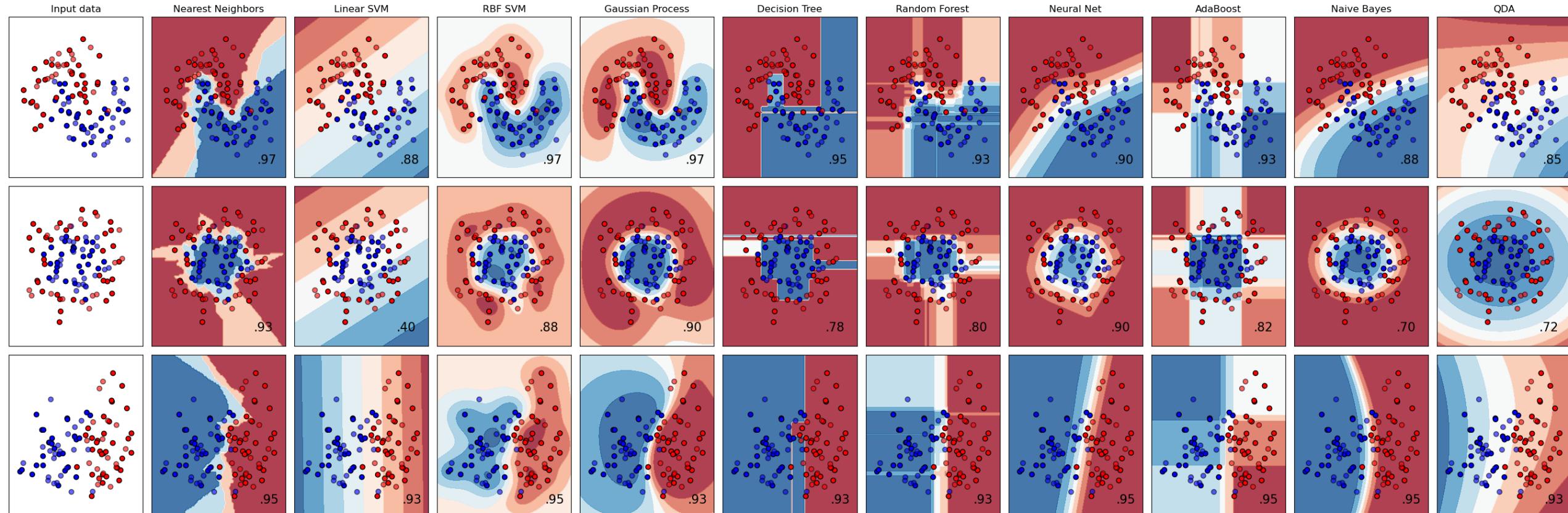


# Decision Boundary

Decision  
Boundary

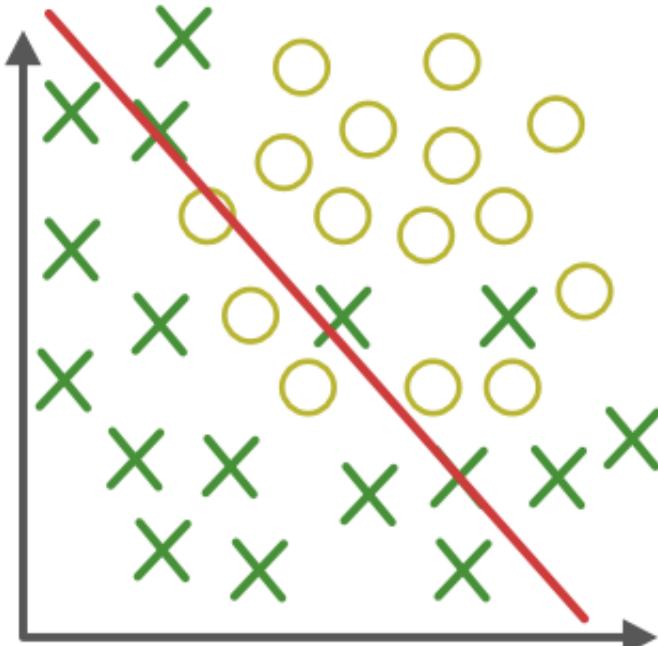


# Decision Boundary (different classification algorithms)

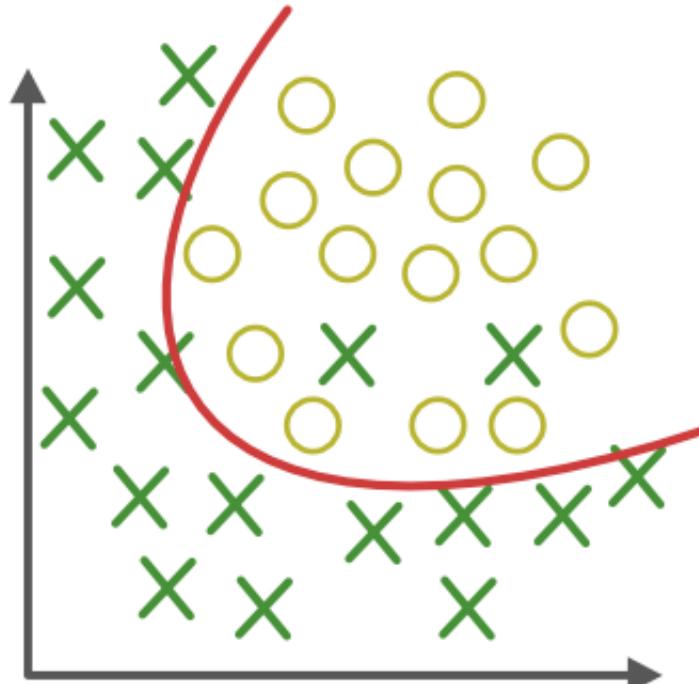


source: [https://scikit-learn.org/stable/auto\\_examples/classification/plot\\_classifier\\_comparison.html](https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html)

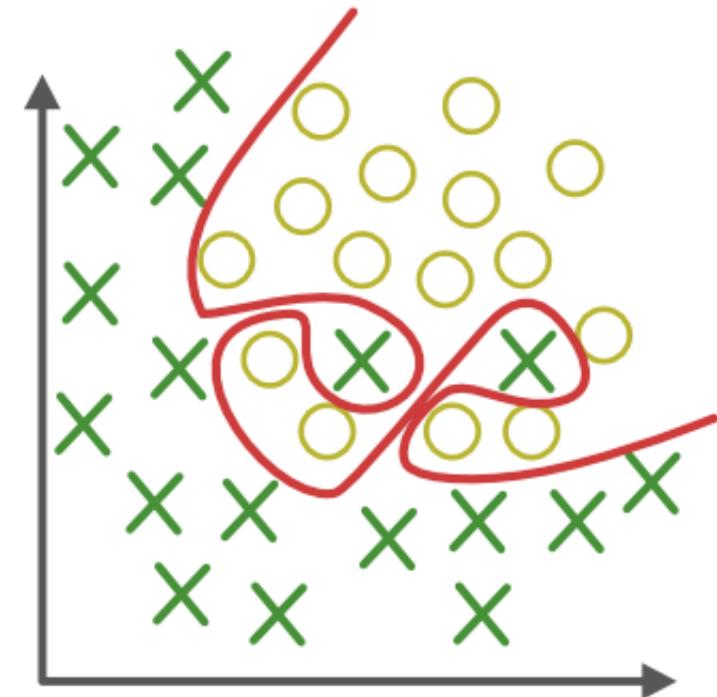
# Underfitting and Overfitting



**Under-fitting**  
(too simple to  
explain the variance)



**Appropriate-fitting**



**Over-fitting**  
(forcefitting--too  
good to be true)

DG

## Confusion matrix

In predictive analytics, a **table of confusion** (sometimes also called a **confusion matrix**) is a table that reports the number of true positives, false negatives, false positives, and true negatives.

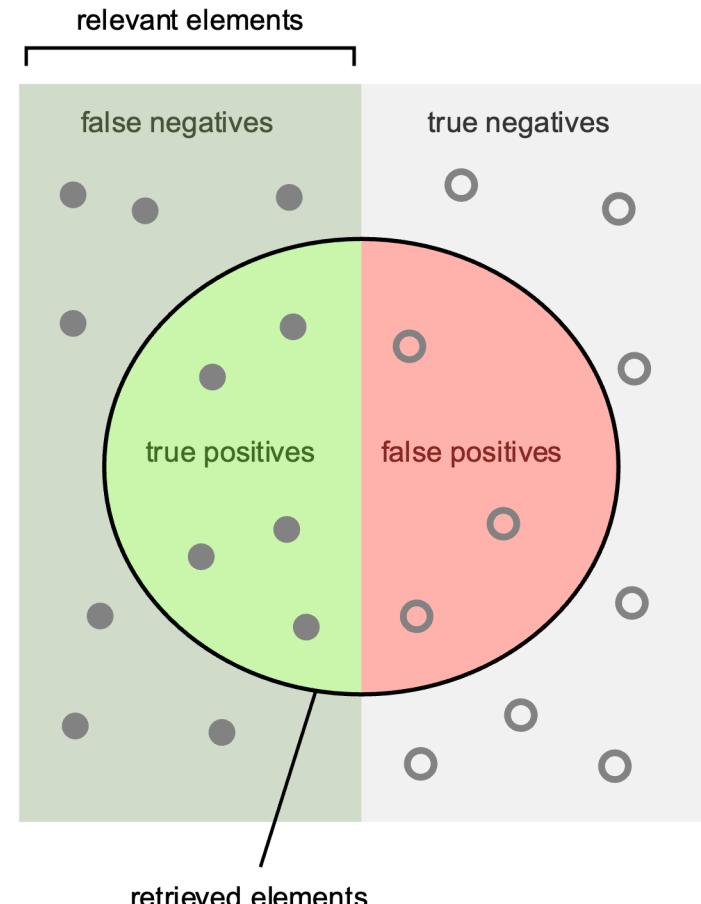
		Predicted class
		$P$
		$N$
Actual class	$P$	True positives (TP)
	$N$	False negatives (FN)
$N$	$P$	False positives (FP)
	$N$	True negatives (TN)

source: <https://subscription.packtpub.com/book/data/9781787125933/6/ch06lvl1sec41/looking-at-different-performance-evaluation-metrics>

# Performance metrics

**Precision** (also called **positive predictive value**) is the fraction of relevant instances among the retrieved instances.

**Recall** (also known as **sensitivity**) is the fraction of relevant instances that were retrieved.



$$\text{Recall} = \frac{\text{Relevant retrieved instances}}{\text{All relevant instances}}$$

How many relevant items are retrieved?

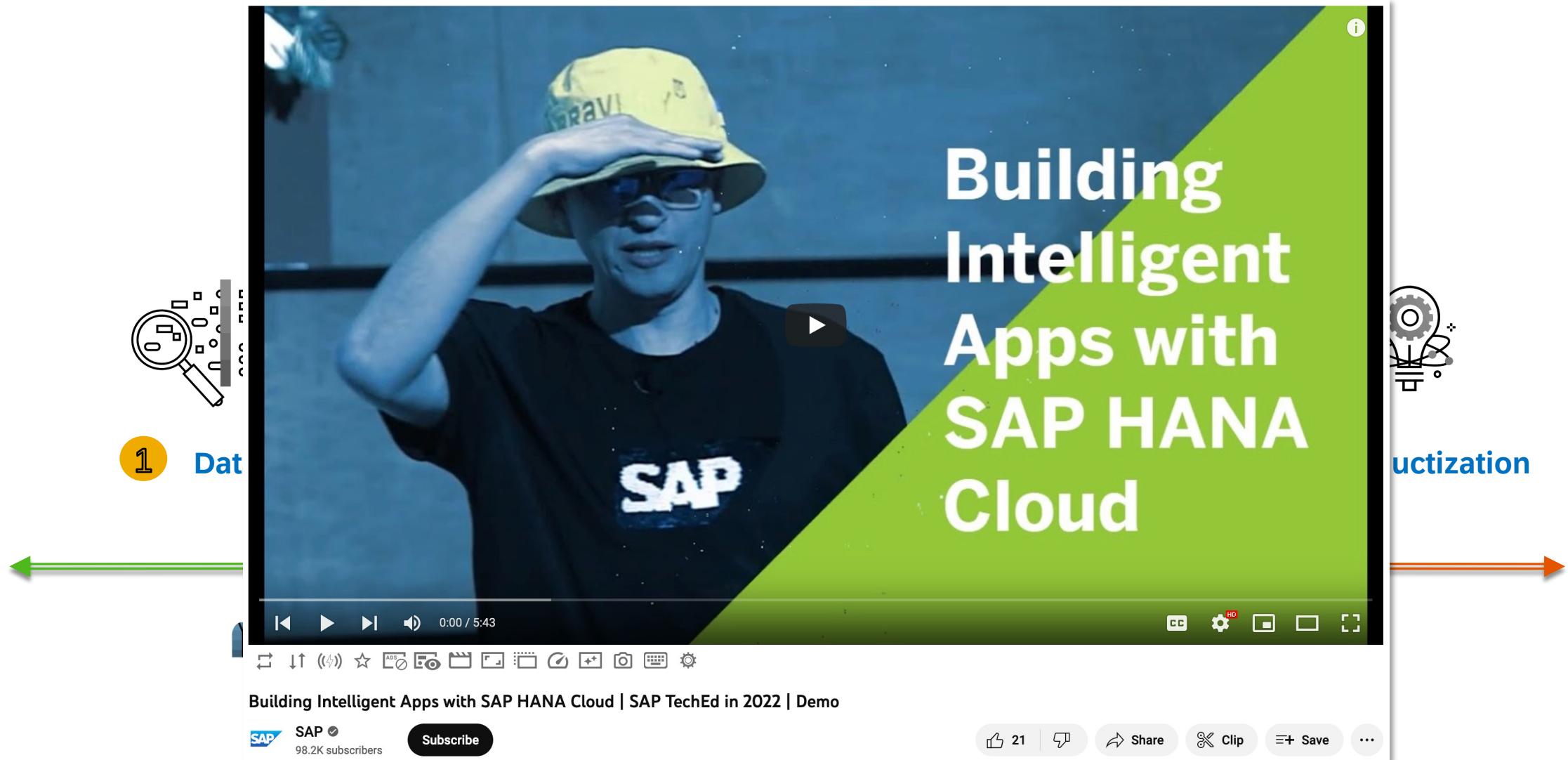
$$\text{Precision} = \frac{\text{True positives}}{\text{All retrieved instances}}$$

$$\text{Precision} = \frac{\text{Relevant retrieved instances}}{\text{All retrieved instances}}$$

$$\text{Recall} = \frac{\text{True positives}}{\text{All relevant instances}}$$

# **Additional content**

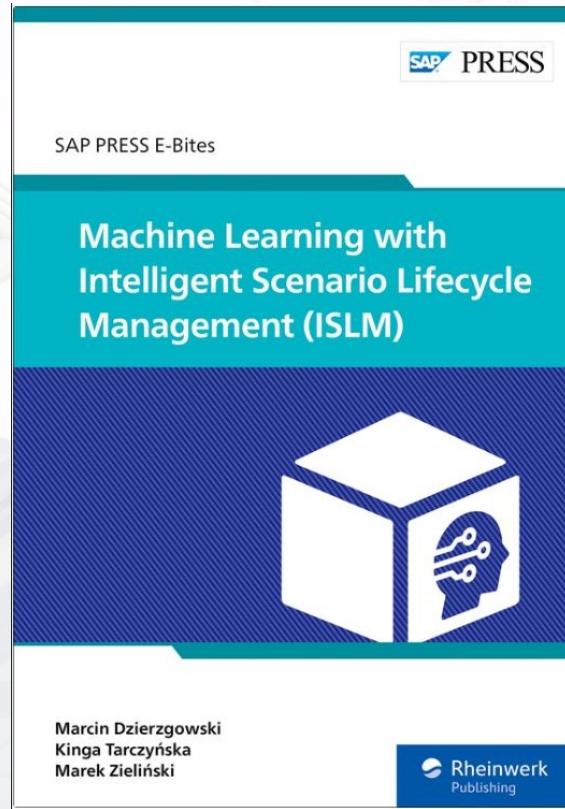
# Development Approach | Building an Intelligent Data Application (demo)



# Code samples: <https://github.com/SAP-samples/hana-ml-samples>

Screenshot of a GitHub repository page for "hana-ml-samples / Python-API / usecase-examples /". The page shows a list of code samples with their names, last commit messages, and dates.

Name	Last commit message	Last commit date
..		
diabetes-classification	Create OpenSAP-SAPHANA-HANA Machine Learning Demo (2...	2 years ago
estimate-car-price	update estimate car price	3 years ago
fairml-examples	FairML - Fair Recruiting Model	3 weeks ago
fraud-detection	fraud-detection use case	2 years ago
melbourne-housing-price	Add files via upload	4 years ago
ml-anonymized-data	Tutorial example - ML with HANA-ML Python ML client on HAN...	4 years ago
multimodel-analysis-airroutes	Update README.md	4 years ago
sapcommunity-automl-examples	update community call examples	2 years ago
sapcommunity-hanaml-challenge	Create SAP HANA Cloud Machine Learning Demo - Employee C...	7 months ago



## Machine Learning with Intelligent Scenario Lifecycle Management (ISLM)

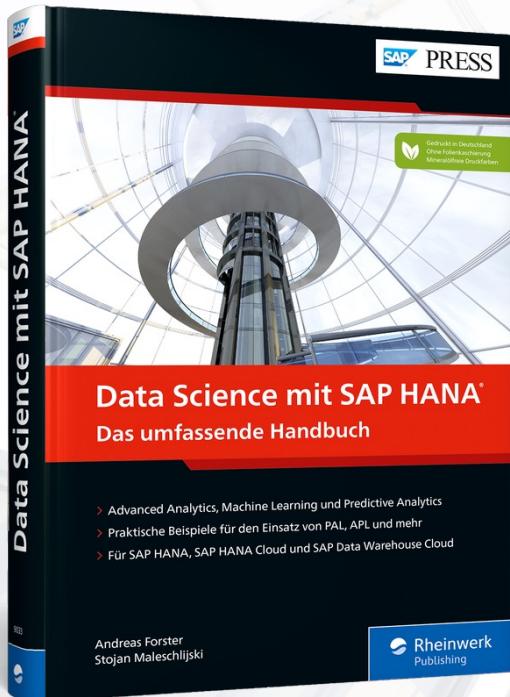
115 pages, 2023, E-Book  
ISBN 978-1-4932-2395-4

[www.sap-press.de/5668](http://www.sap-press.de/5668)

## Update your machine learning skills with Intelligent Scenario Lifecycle Management (ISLM)!

In this E-Bite, you'll develop a complete machine learning application for SAP S/4HANA using SAP HANA PAL, from data preparation and model building to training and prediction generation. You'll learn to use the ISLM framework to simplify machine learning implementation with standard apps for managing intelligent scenarios. Learn the ins and outs of machine learning with ISLM in this how-to guide!

- Learn to use the ISLM framework in SAP S/4HANA
- Develop an end-to-end machine learning scenario with SAP HANA PAL
- Prepare data, train models, and implement predictions with ISLM



## Mit SAP HANA, SAP HANA Cloud und SAP Data Warehouse Cloud ist viel mehr möglich als das Speichern großer Datenmengen.

In diesem Buch erfahren Sie, wie Sie die Automated Predictive Library (APL) und die Predictive Analysis Library (PAL) einsetzen können, um komplexe Auswertungen vorzunehmen und Vorhersagen zu treffen. Praktische Beispiele zu Klassifizierung, Textanalyse, Clustering, Regression u.v.m. zeigen Ihnen die vielfältigen Möglichkeiten auf und lassen sich direkt auf Ihre Anwendungsfälle übertragen.

- Advanced Analytics, Machine Learning und Predictive Analytics
- Praktische Beispiele für den Einsatz von PAL, APL und mehr
- Für SAP HANA, SAP HANA Cloud und SAP Data Warehouse Cloud

### Data Science mit SAP HANA

Das umfassende Handbuch

von [Andreas Forster, Stojan Maleschlijski](#)

<https://www.rheinwerk-verlag.de/data-science-mit-sap-hana/>

# SAP HANA Cloud Multi-model Further Learning

## SAP HANA Cloud Basic Trial

- 30-days own-schema shared-instance
- <https://www.sap.com/products/technology-platform/hana/guided-experience.html>

## SAP HANA Cloud Trial

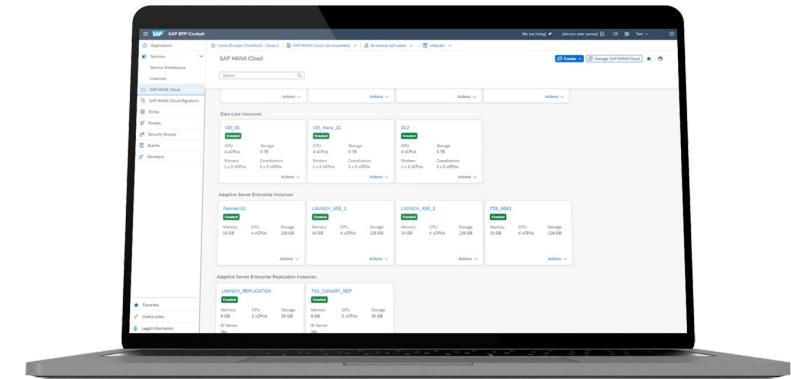
- 3x30-days own-instance in SAP BTP Trial
- 16GB RAM, 1 vCPU
- <https://developers.sap.com/tutorials/hana-trial-advanced-analytics.html>

## SAP HANA Cloud Free Tier

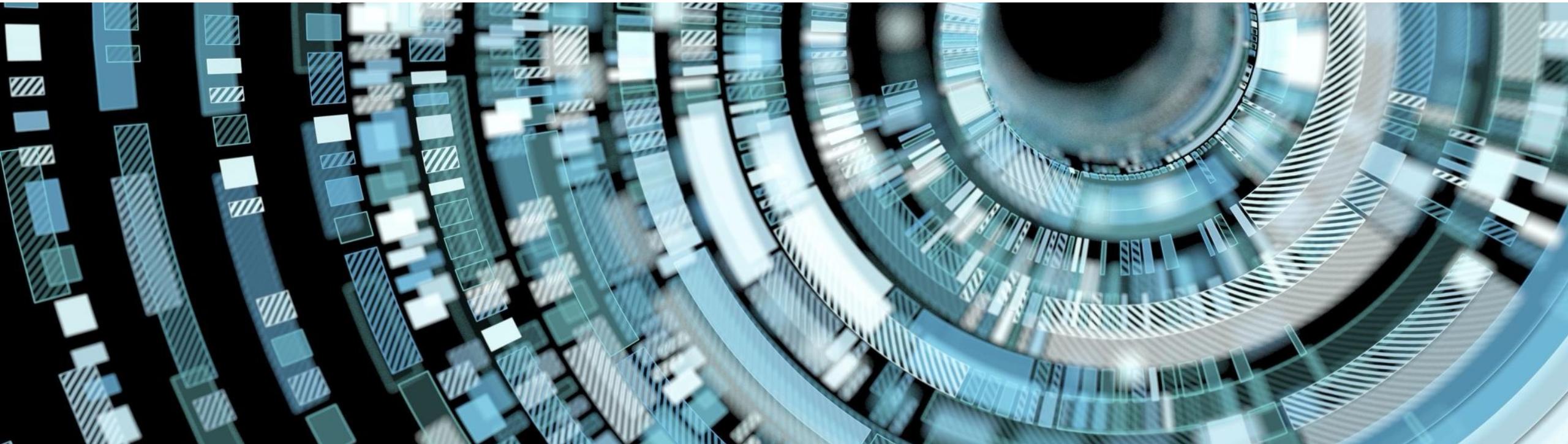
- 30 GB RAM, 2 vCPUs
- <https://developers.sap.com/mission.hana-cloud-database-get-started.html>

## SAP TechEd 2023 exercises

- DAT285v - Building Intelligent Data Applications with SAP HANA Cloud: <https://github.com/SAP-samples/teched2023-DA285v>
- DA263 - Build Innovative Business Applications with Database Services: <https://github.com/SAP-samples/teched2023-DA263>



# SAP Business AI portfolio



# SAP Business AI approach

Relevant. Reliable. Responsible.

Finance | Supply Chain | HR | Procurement | Marketing and Commerce | Sales and Services | IT



A copilot that truly understands your business

## Embedded AI capabilities

Cloud ERP

Human capital management

Spend management  
and business network

Customer relationship  
management

Business Technology  
Platform

## AI Foundation

on Business Technology Platform

## AI ecosystem partnerships and investments



ANTHROPIC



cohere

databricks

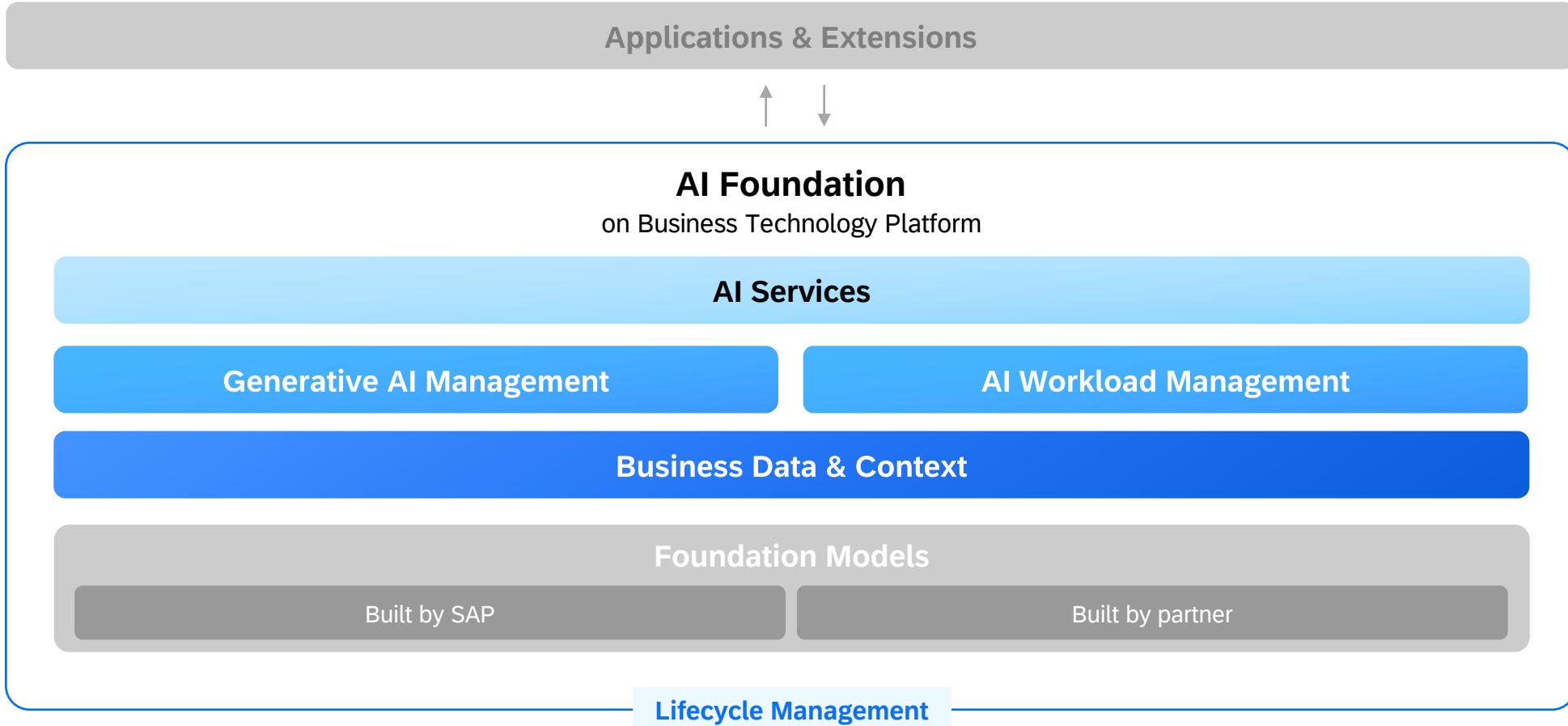
DataRobot

Google Cloud

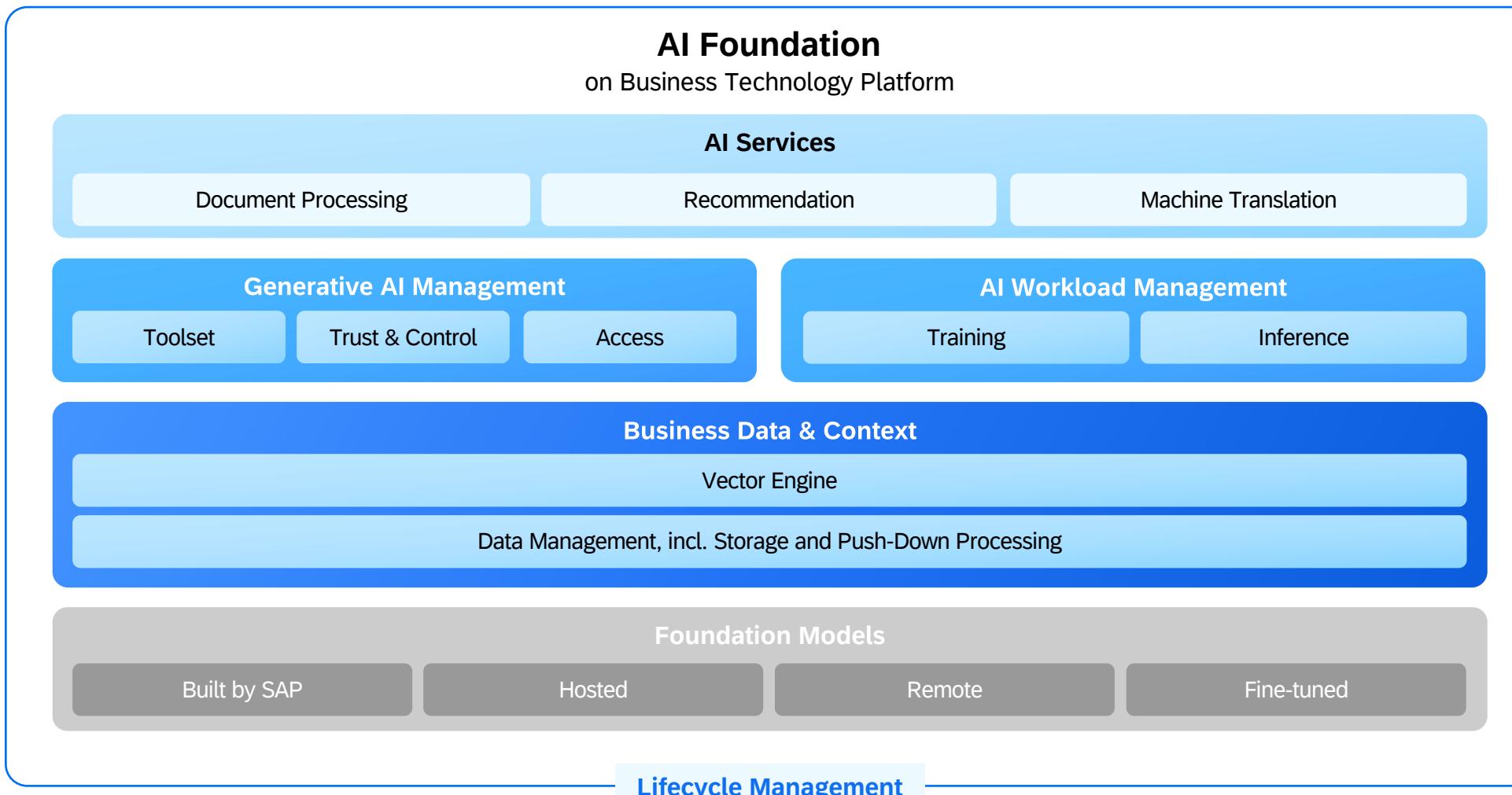
IBM

Microsoft

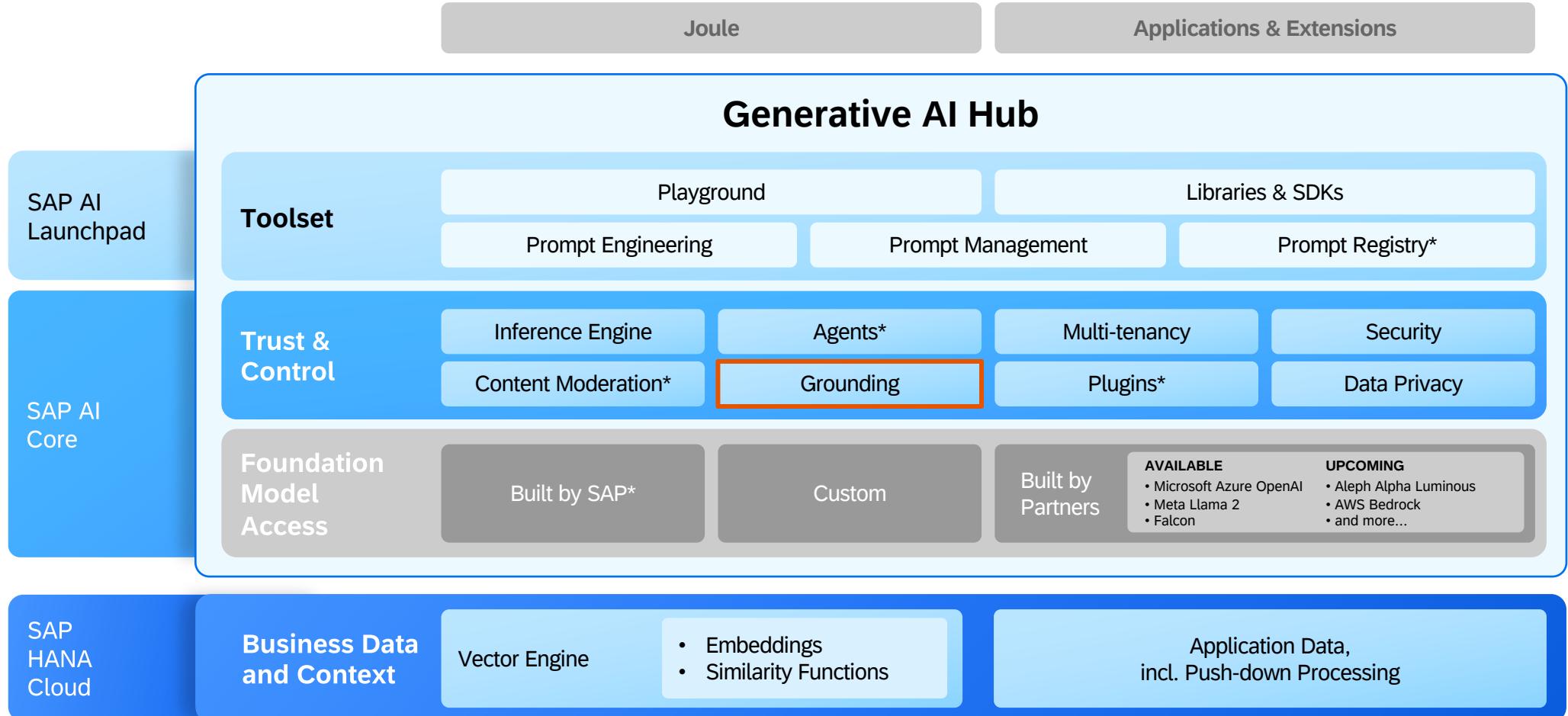
# A complete set of services for AI developers on SAP BTP



# A complete set of services for AI developers on SAP BTP



# Generative AI Hub: integrate AI into applications, seamlessly



Please, support Ukrainian business 💙💛 eg. Ugears Mechanical Models

<https://ugearsmodels.com/>

Free shipping on all orders of €60 or more!

The website features a green header bar with the text "Free shipping on all orders of €60 or more!". Below the header is the Ugears logo and a search bar. The main navigation menu includes links for CATALOG, BEST DEALS, CUSTOMER SERVICE, UGEARS WORLDWIDE, GIFT IDEAS, and language selection (English). The main content area features a large blue banner with the text "NEW RELEASE" in white. Below the banner are four product images: "SERENITY'S DREAM YACHT", "NASA SPACE SHUTTLE DISCOVERY", "RESCUE HOVERCRAFT", and "STEGOSAURUS". A "LEARN MORE" button is located next to the Hovercraft image. At the bottom of the page, there are four categories with icons: "SELF ASSEMBLY" (thumb up), "MECHANICAL" (cogwheel), "EDUCATIONAL" (brain), and a speech bubble icon.

<https://ugearsmodels.com/catalogue/preorder5/>



SCAN ME

**Thank you // Дякую // Dziękuję!**

Contact information:

Witalij Rudnicki

