The Developer Advocates present

SAP CodeJam 2025

Getting Started with Machine Learning using SAP HANA

Kolkata, India 🇮🇳
Mar 4, 2025

SAP TechEd

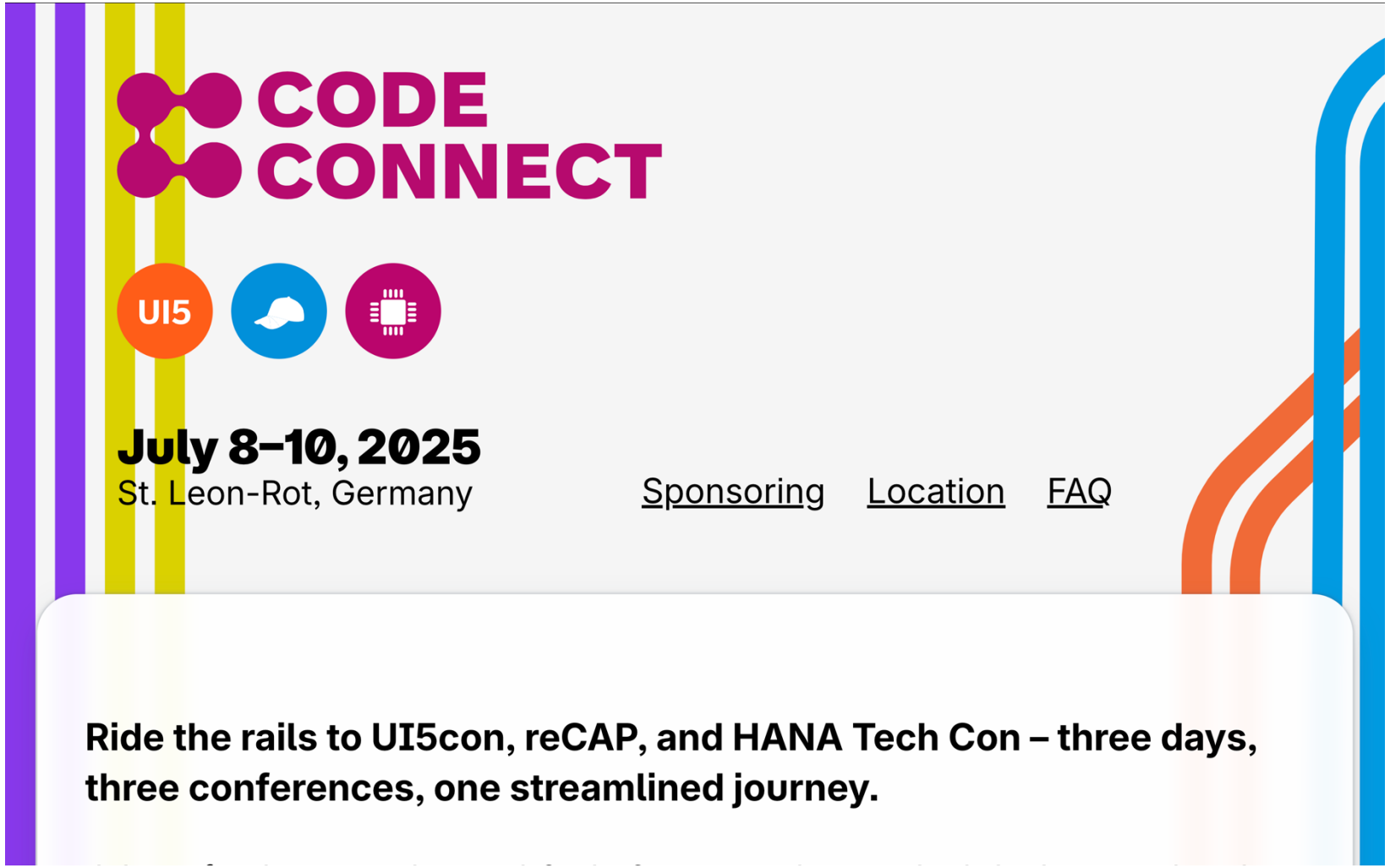
Witalij Rudnicki
aka @Sygyzmundovych
aka **Vitaliy Rudnytskiy**
aka **Віталій Рудницький**

- 10 years **tech consultant** in SAP BI/BW
- 10+ years as 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**

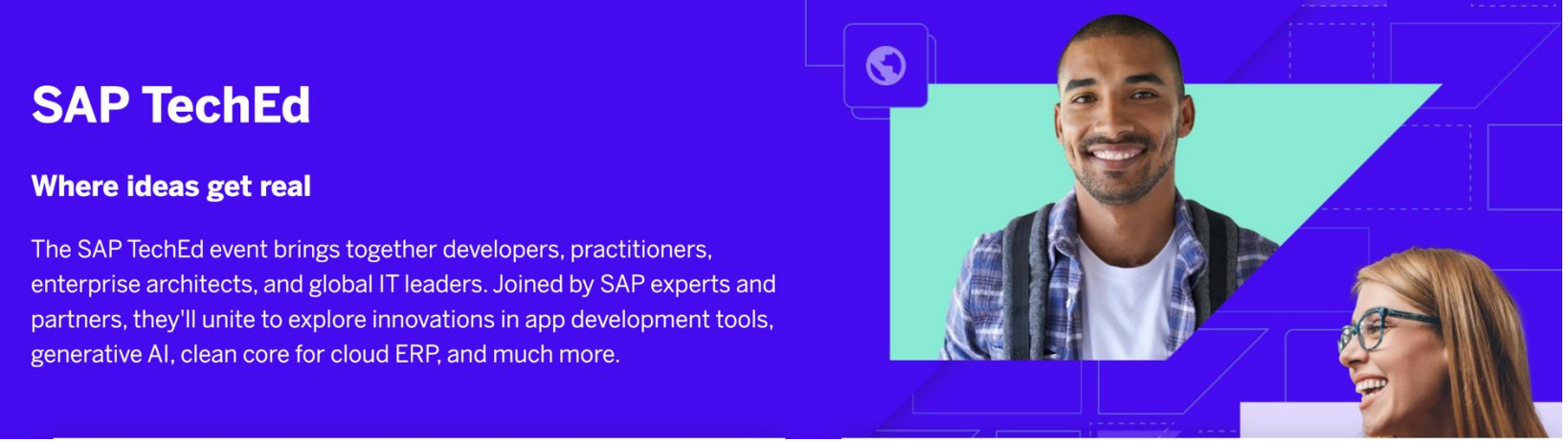
Events not be missed

Code Connect:
https://code-connect.dev/



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

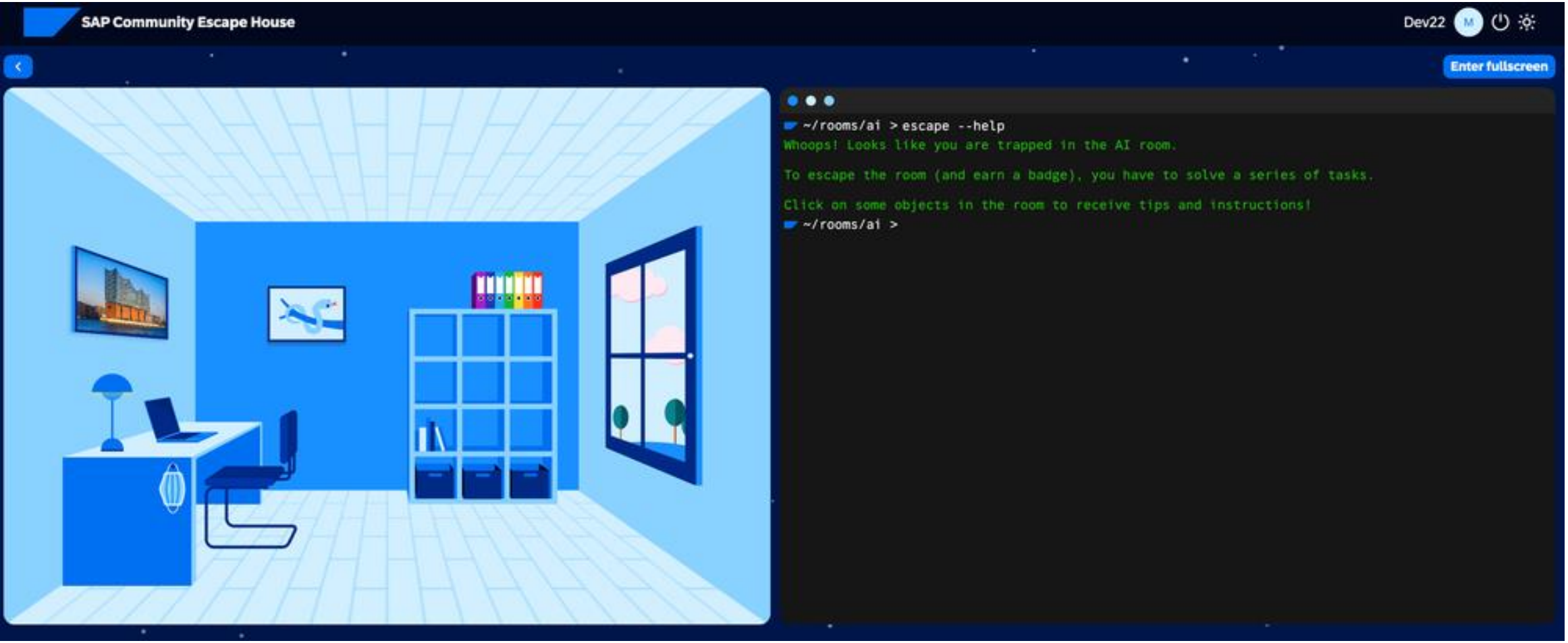
1. ABAP and SAP Cloud Application Programming Model
2. Tooling
3. Integration
4. MAD (Machine Learning, AI, and Data)
5. Frontend



https://www.sap.com/events/teched/virtual.html:

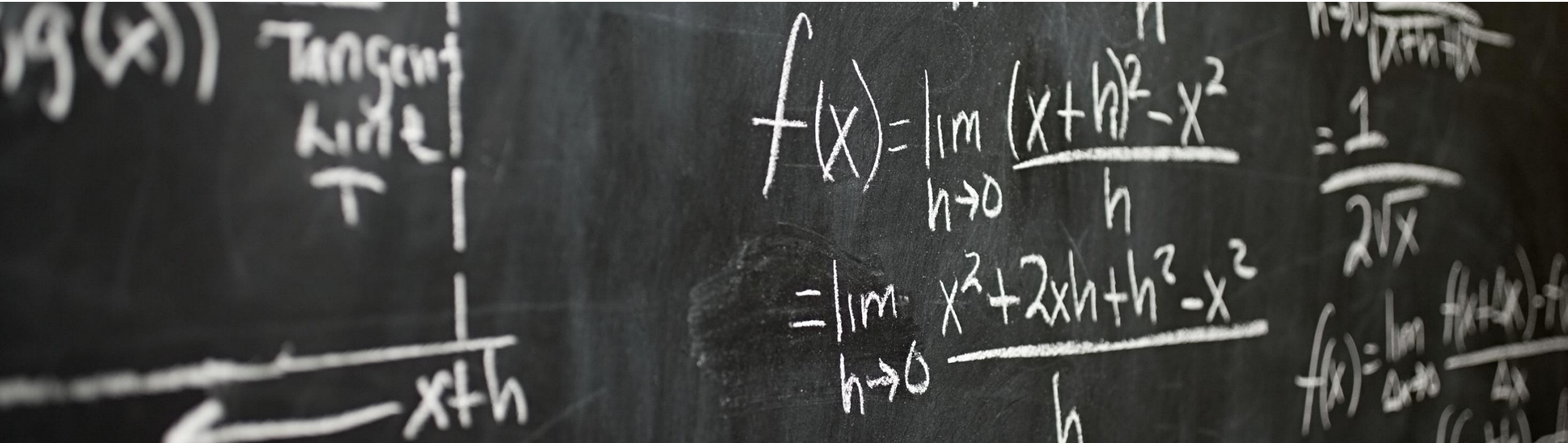
Berlin and Virtual: **November 4-6, 2025**

SAP Developer Challenge March - Community Escape House - AI

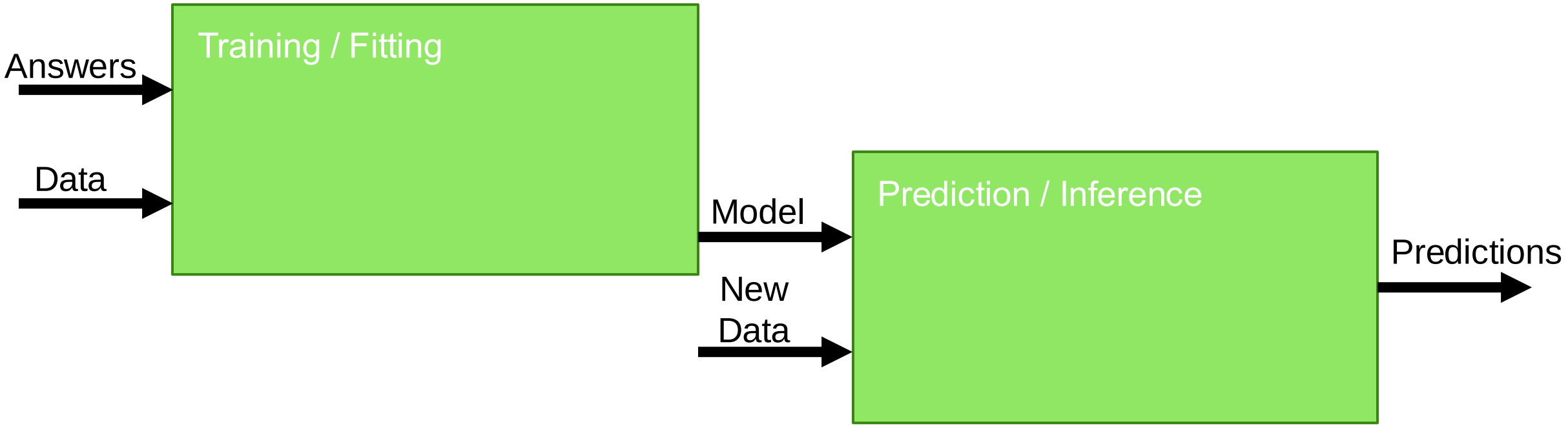


source: https://community.sap.com/t5/artificial-intelligence-and-machine-learning-blogs/sap-developer-challenge-march-sap-community-escape-house-ai/ba-p/14030669

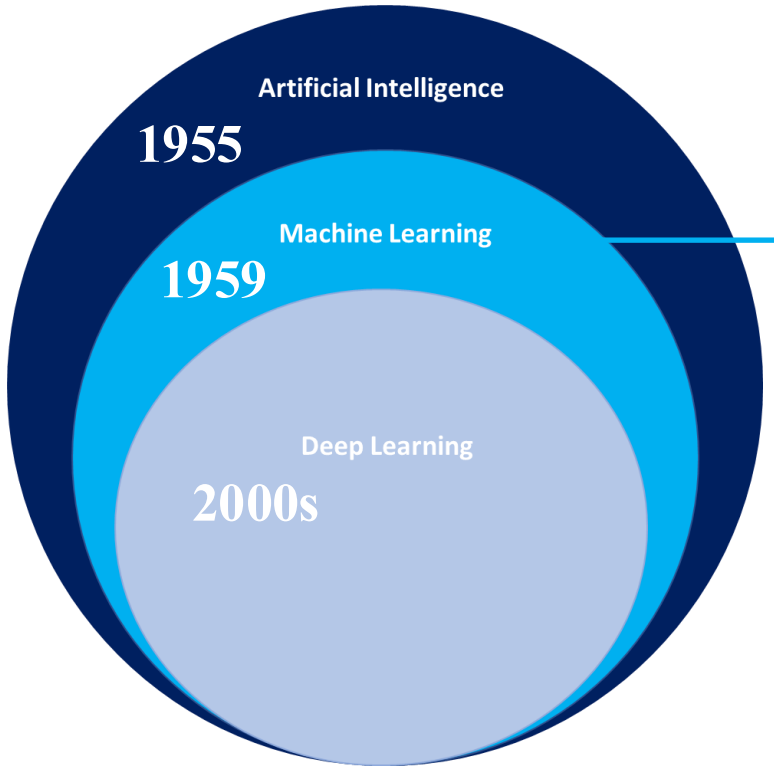
A bit of theory
before we start

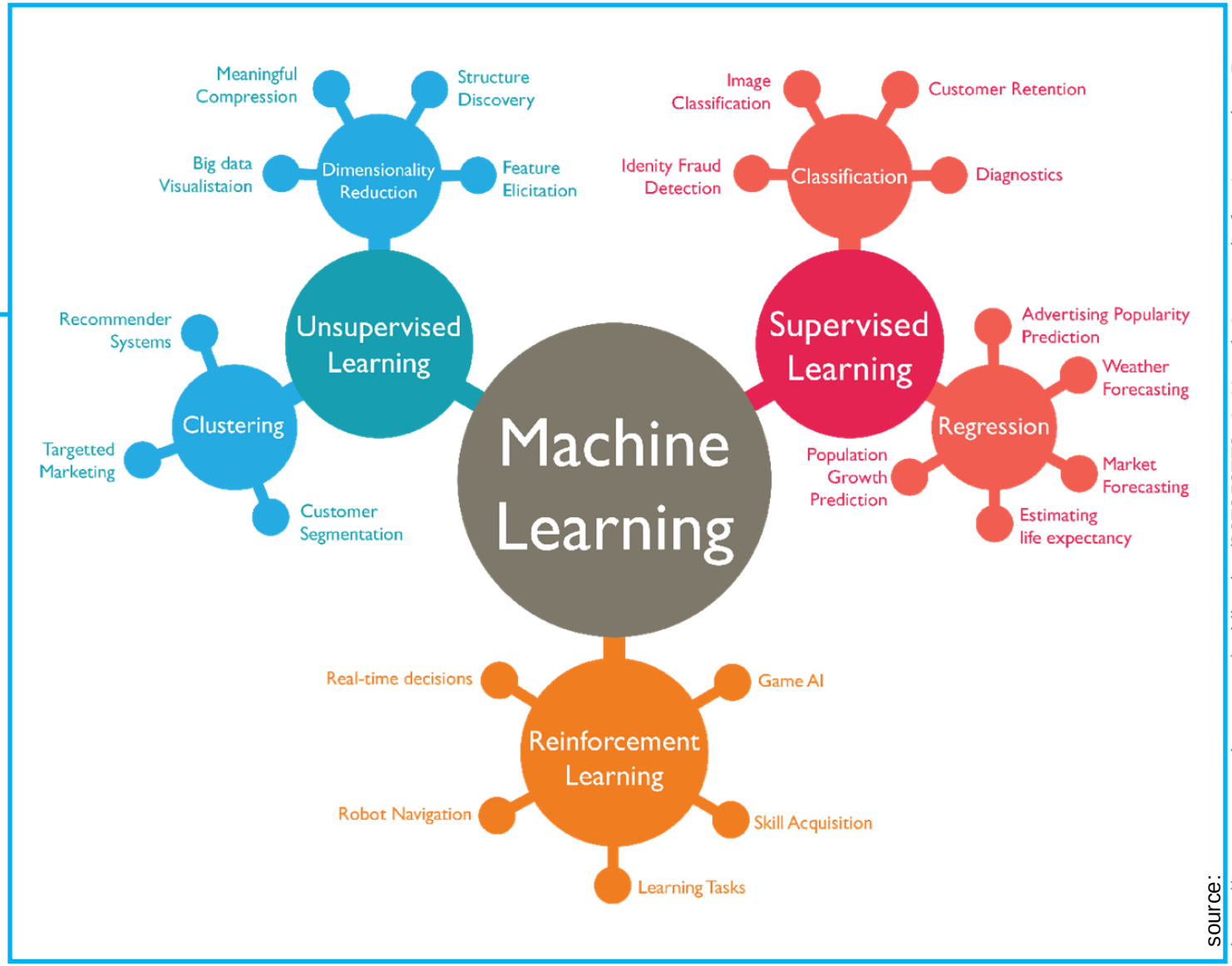


Machine Learning



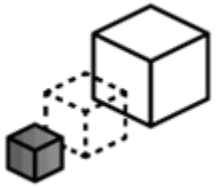
Machine Learning Terminology

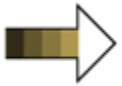


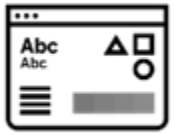


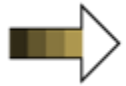
source: https://www.researchgate.net/publication/337958773_A_Novel_Approach_for_Improving_Breast_Cancer_Risk_Prediction_using_Machine_Learning_Algorithms_A_Survey

Development Approach | Building an Intelligent Data Application









1 Data

2 Model + MLOps

3 Develop

4 Productization + DevOps

Data Scientist

ML/AI Engineer

(App Developer)

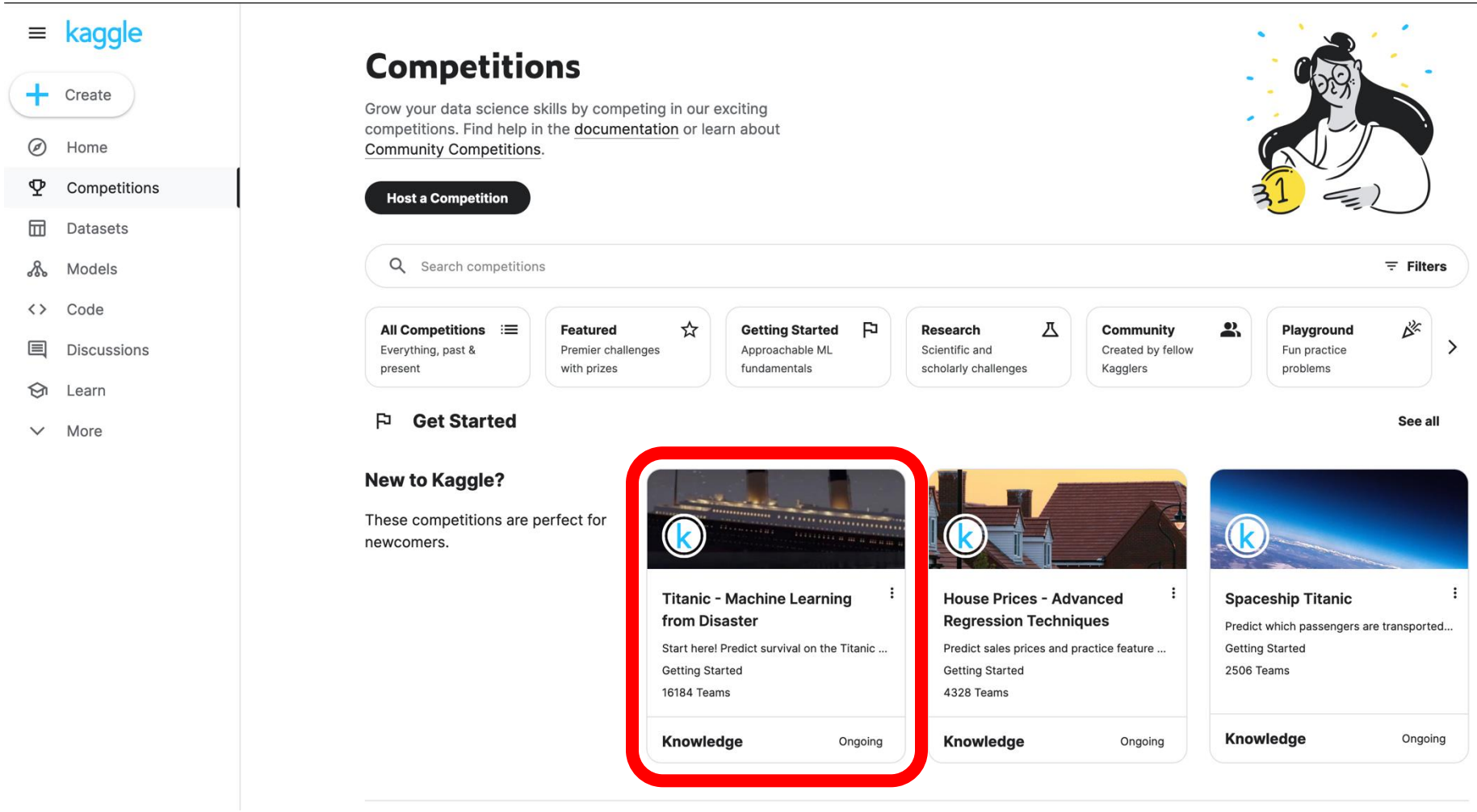
Development Approach | Training a Model

Data Scientist 

source: https://en.wikipedia.org/wiki/Cross-industry_standard_process_for_data_mining

Let’s get busy 

Your ML Challenge today:



https://www.kaggle.com/competitions

Exercises

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

(→ https://bit.ly/CJ_HANAML)

1. Pre-requisites:
https://github.com/SAP-samples/hana-ml-py-codejam/blob/main/prerequisites.md

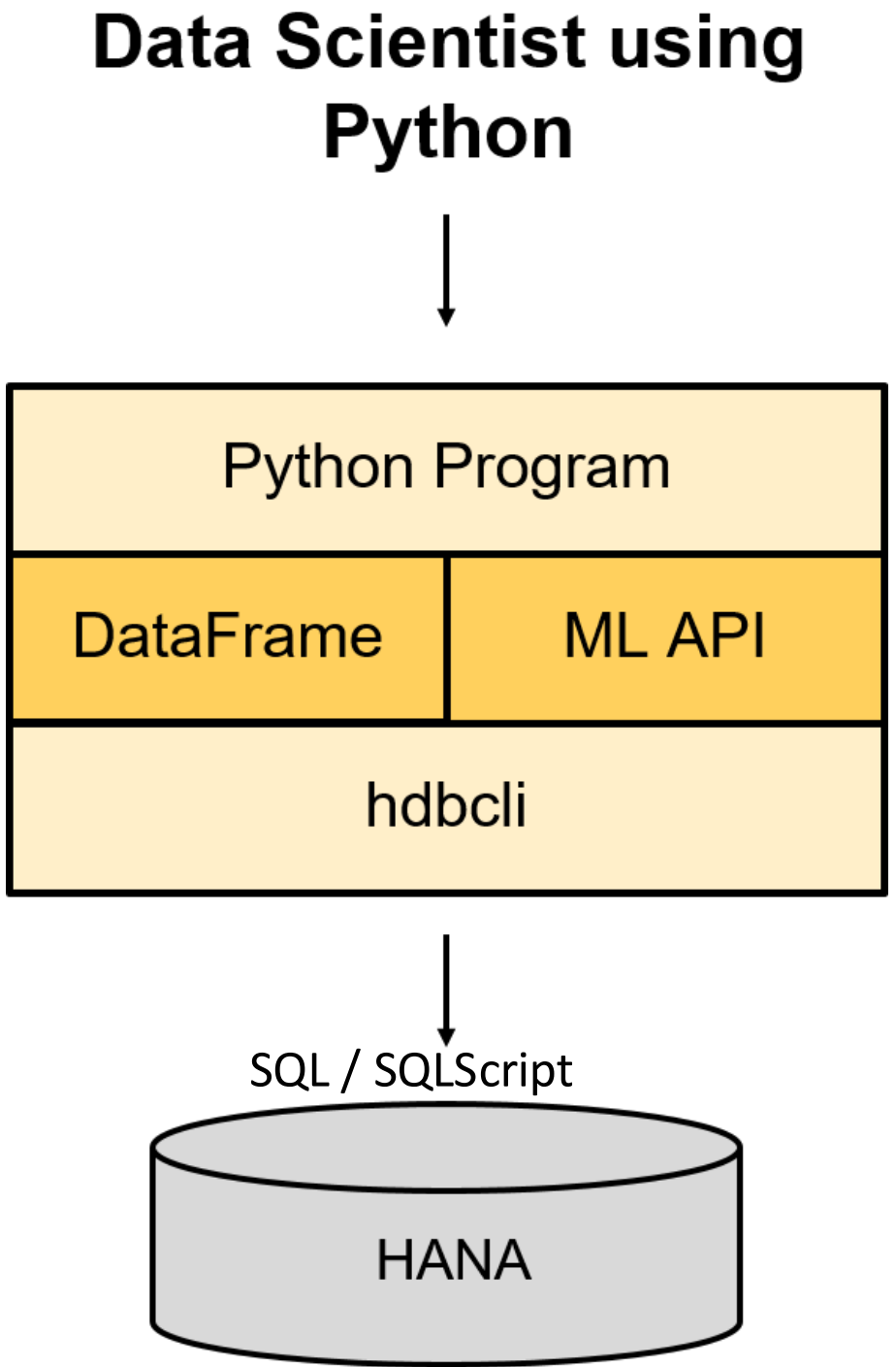
2. Exercises:
https://github.com/SAP-samples/hana-ml-py-codejam#the-exercises

3. Worth watching:
https://github.com/SAP-samples/hana-ml-py-codejam#additional-learning-material

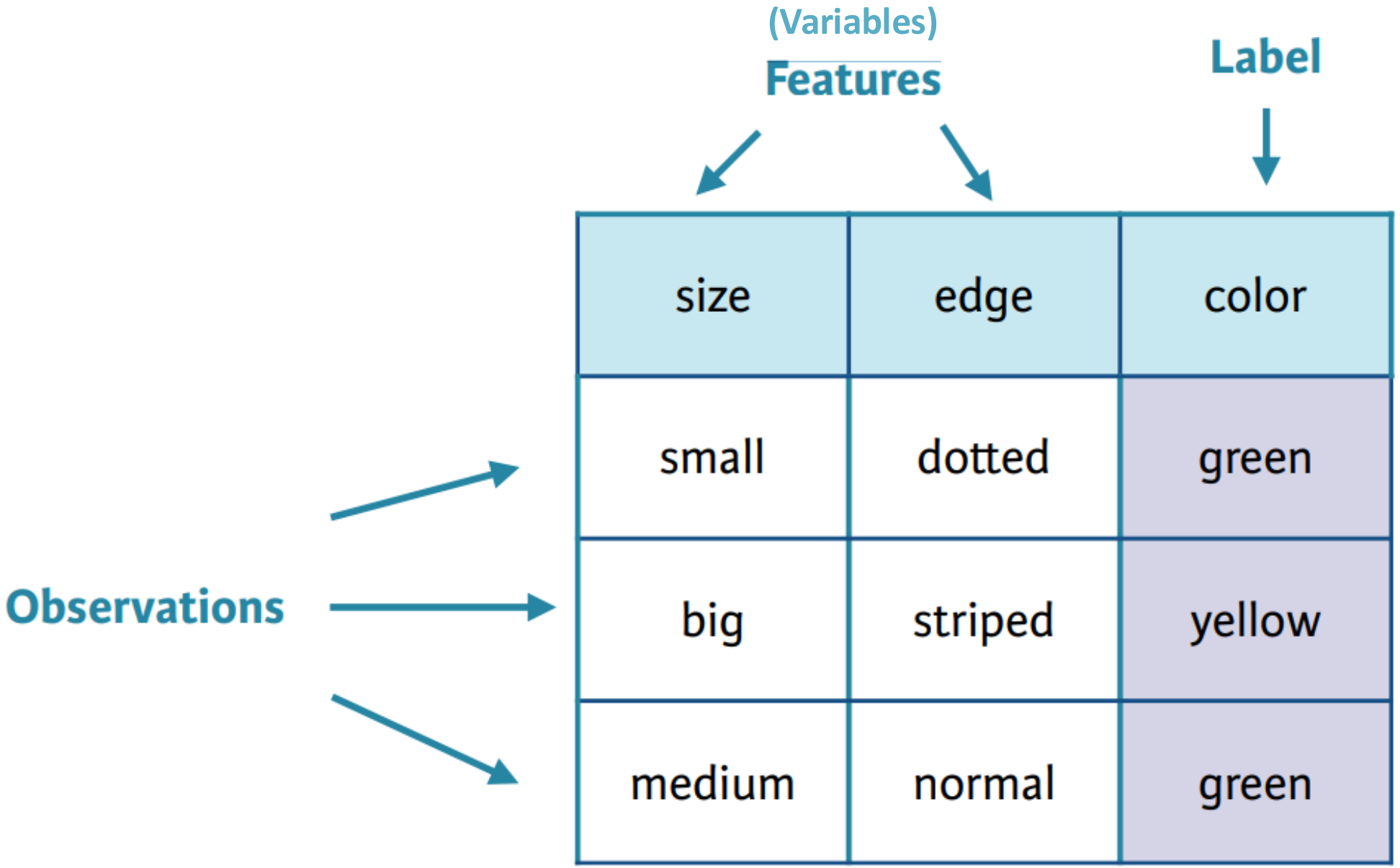
4. Try at home in your SAP BTP Trial:
https://github.com/SAP-samples/sap-community-developer-challenge-eda-hana

Understand DataFrame(s)

Pandas DataFrame <-> HANA DataFrame



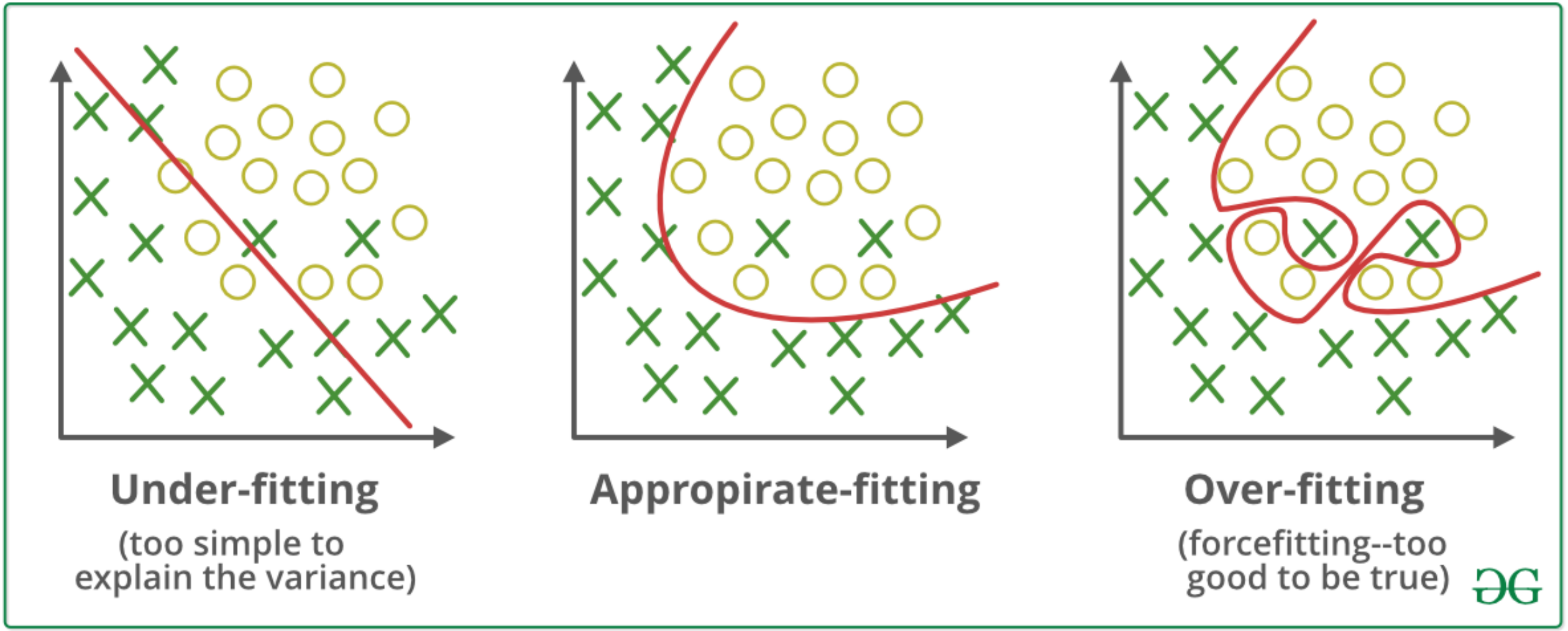
Train datasets in Supervised Machine Learning



source: https://zinayouhan33.medium.com/machine-learning-can-be-divided-into-3-categorizations-supervised-unsupervised-and-reinforcement-9a1b47460f5d

Public

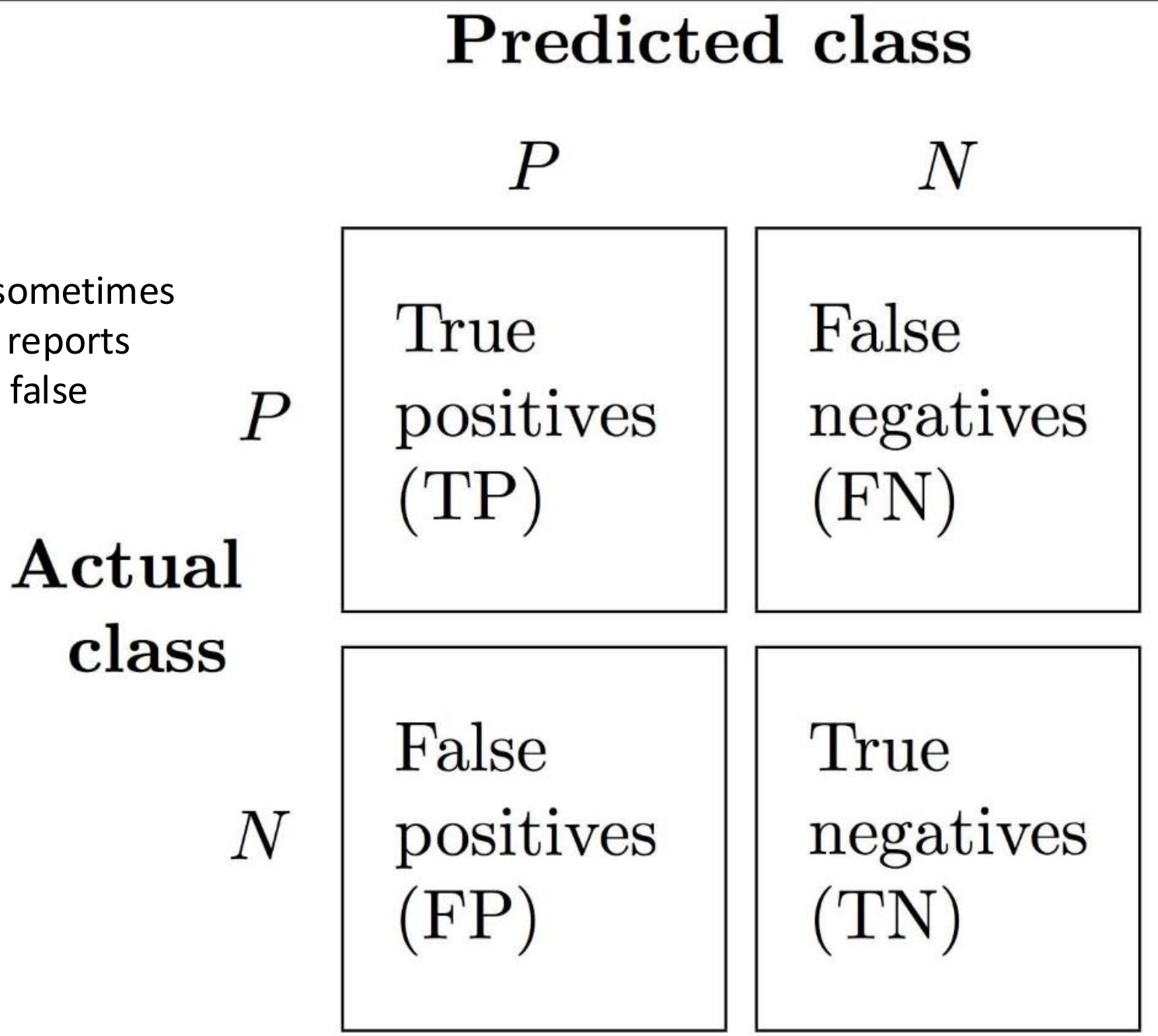
Underfitting and Overfitting



source: https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/

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.

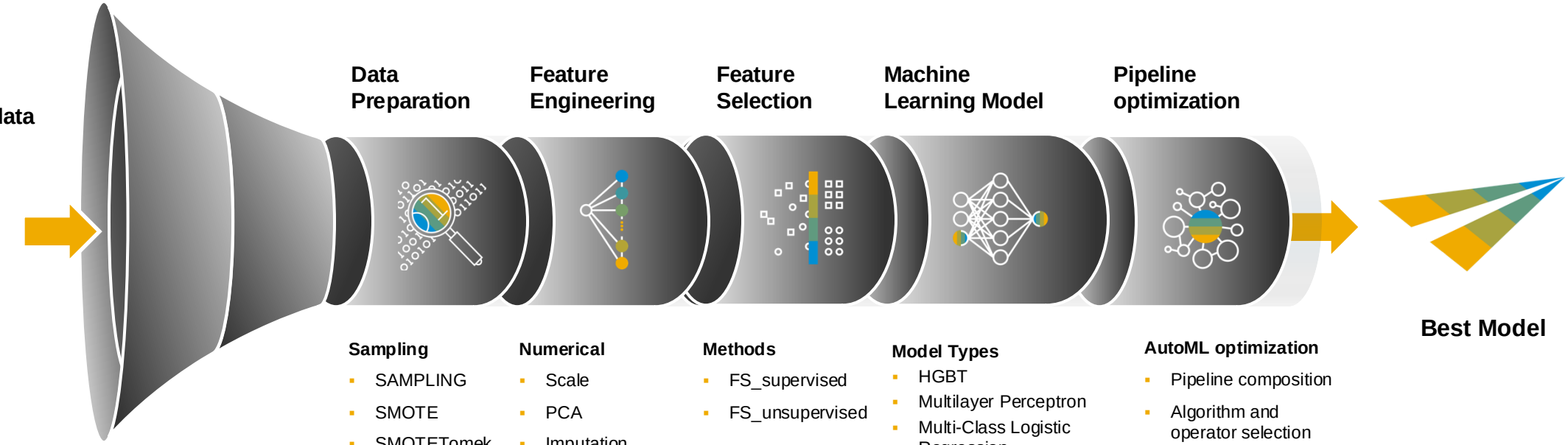


source: https://subscription.packtpub.com/book/data/9781787125933/6/ch06lvl1sec41/looking-at-different-performance-evaluation-metrics

Public

Predictive Analysis Library | Automated Machine Learning – Supported Operators

Structured data
- Numerical
- Categorical



Data Preparation

Sampling
- SAMPLING
- SMOTE
- SMOTETomek
- TomekLinks

Time Series
- Outlier
- Timepoint / Timestamp data transformation

Feature Engineering

Numerical
- Scale
- PCA
- Imputation
- PolynomialFeatures Generation

Categorical
- Categorical Principal Component Analysis
- Label Encoder
- One-Hot Encoder
- Imputation

Time Series
- Imputation

Feature Selection

Methods
- FS_supervised
- FS_unsupervised

See PAL Feature Selection

Machine Learning Model

Model Types
- HGBT
- Multilayer Perceptron
- Multi-Class Logistic Regression
- Naive Bayes
- Random Decision Tree
- Support Vector Machine
- Decision Trees
- Multiple Linear Regression
- Bi-Variate Natural Logarithmic Regression
- Exponential Regression
- Bi-Variate Geom. Regr.
- GLM
- ARIMA, Exponential Smoothing, BSTS, AdditiveModel Analysis

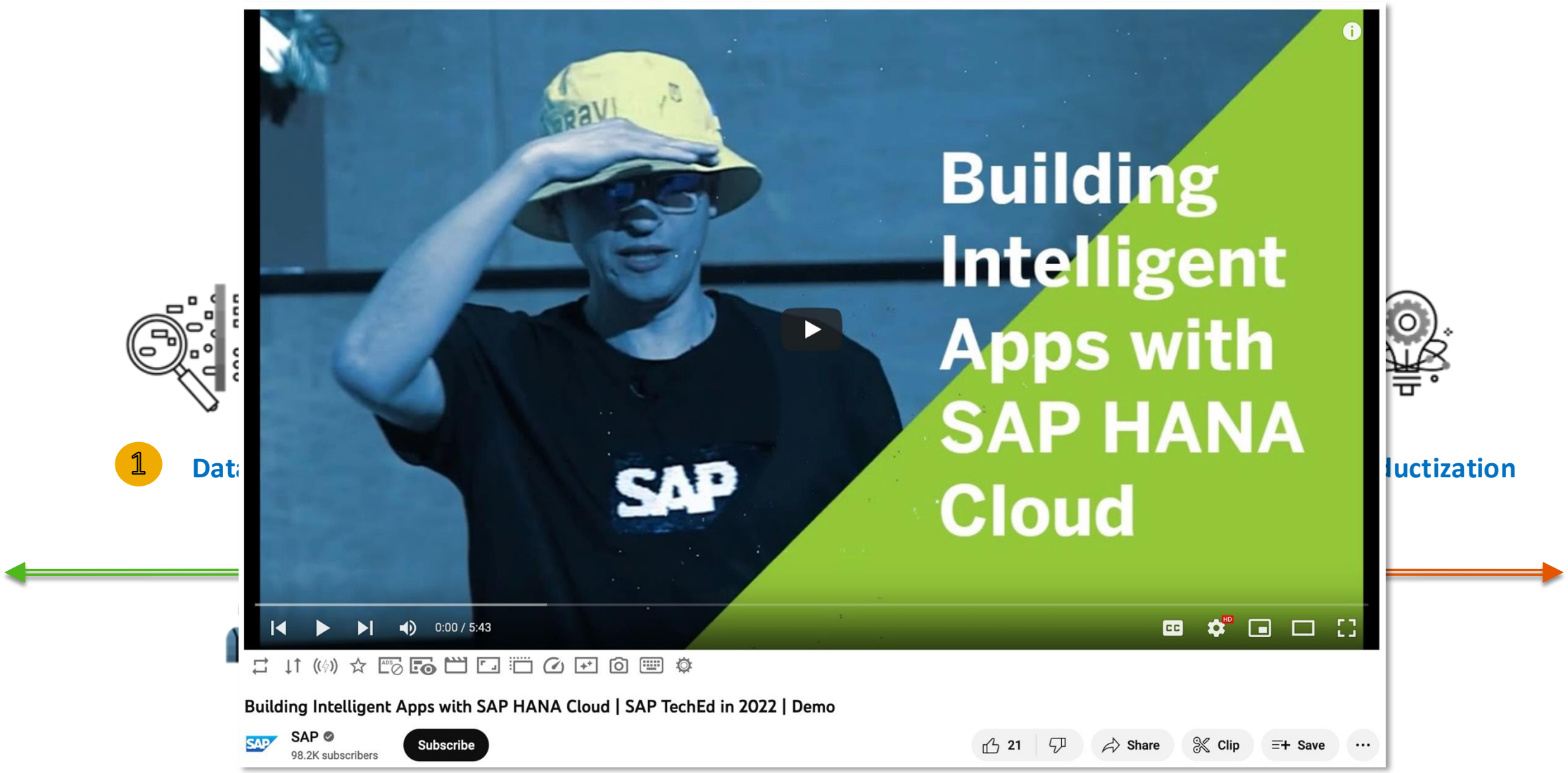
Pipeline optimization

AutoML optimization
- Pipeline composition
- Algorithm and operator selection
- (Hyper-) parameter selection and tuning
- Time series rolling and blocked cross validation
- Genetic algorithm based pipeline optimization (TPOT-like)
- Successive Halving
- Connectivity Optimization

Best Model

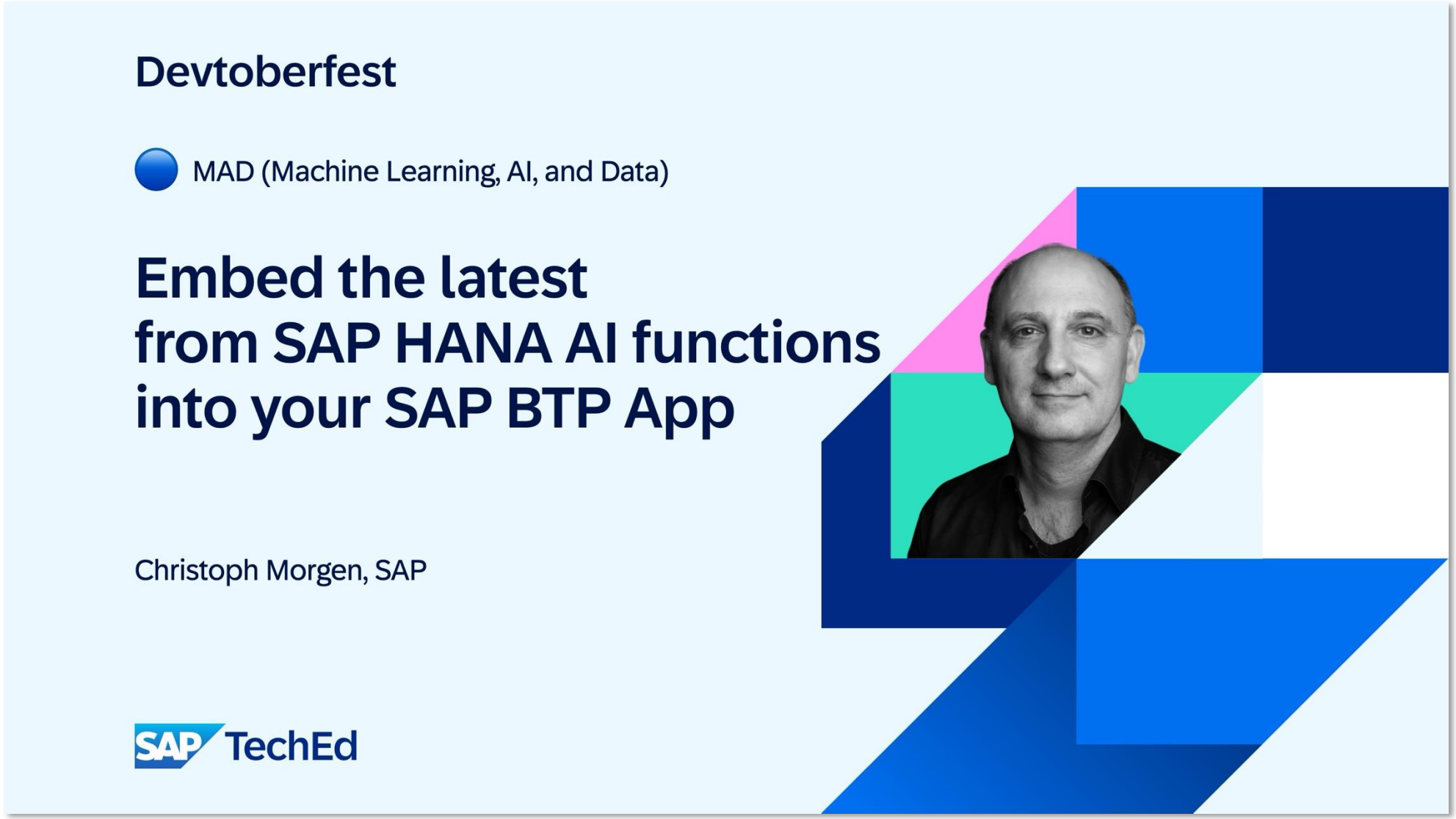
Additional content

Development Approach | Building an Intelligent Data Application (demo)



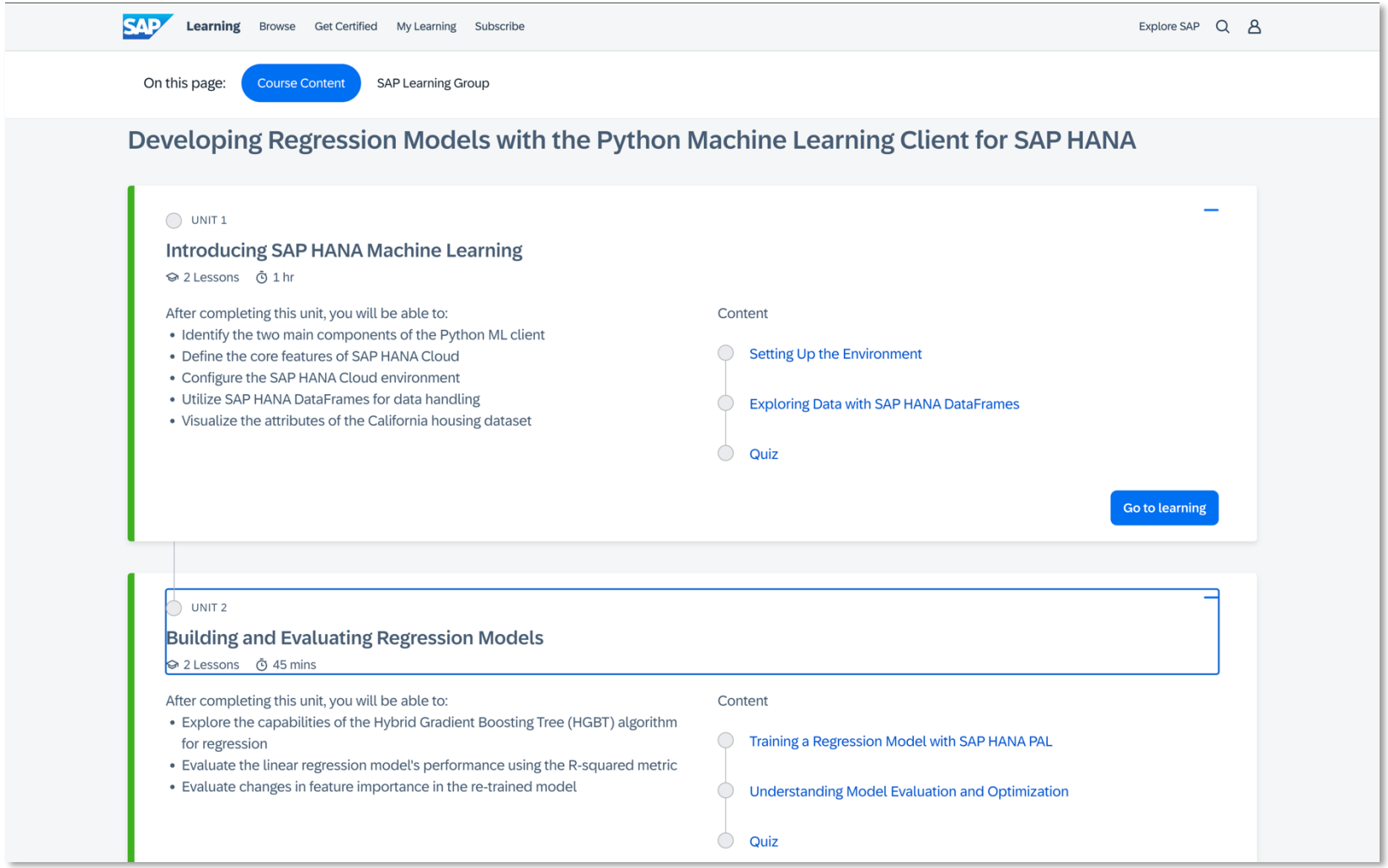
https://www.youtube.com/watch?v=yzA3IOIG5IQ&list=PLSXNnd21oW416dQIZu5-XGx9K0806AUKN

Embed the latest from SAP HANA's AI functions into your SAP BTP Application



https://youtu.be/J_dTMTHFC9M

Learning journey: Developing Regression Models with the Python Machine Learning Client for SAP HANA



https://learning.sap.com/learning-journeys/developing-regression-models-with-the-python-machine-learning-client-for-sap-hana

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

main

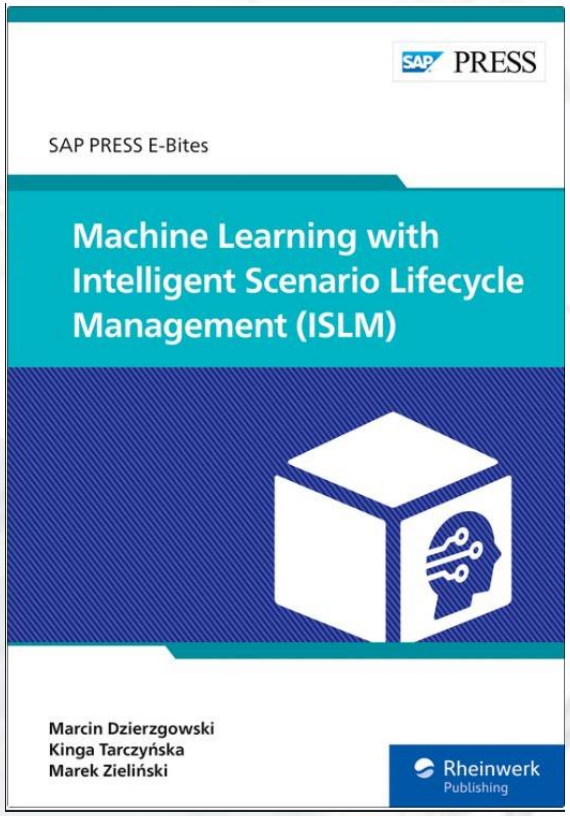
hana-ml-samples / Python-API / usecase-examples /

Go to file t

Add file

cmog FairML - Fair Recruiting Model f009115 · 3 weeks ago History

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

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

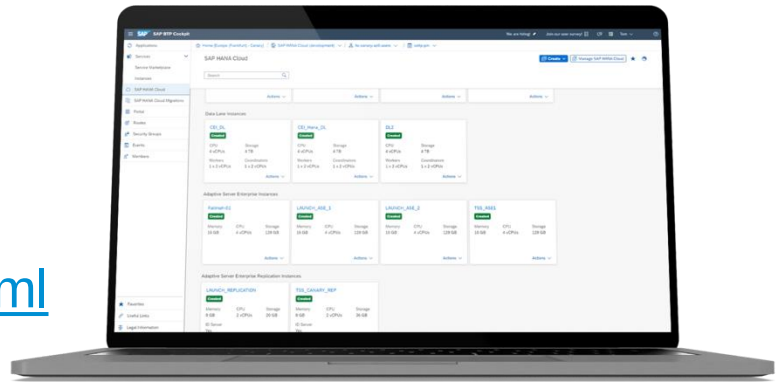
SAP HANA Cloud – Free offering

SAP HANA Cloud Basic Trial

- 30-days own-schema shared-instance
- https://www.sap.com/products/technology-platform/hana/trial.html

SAP HANA Cloud (Advanced) Trial

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

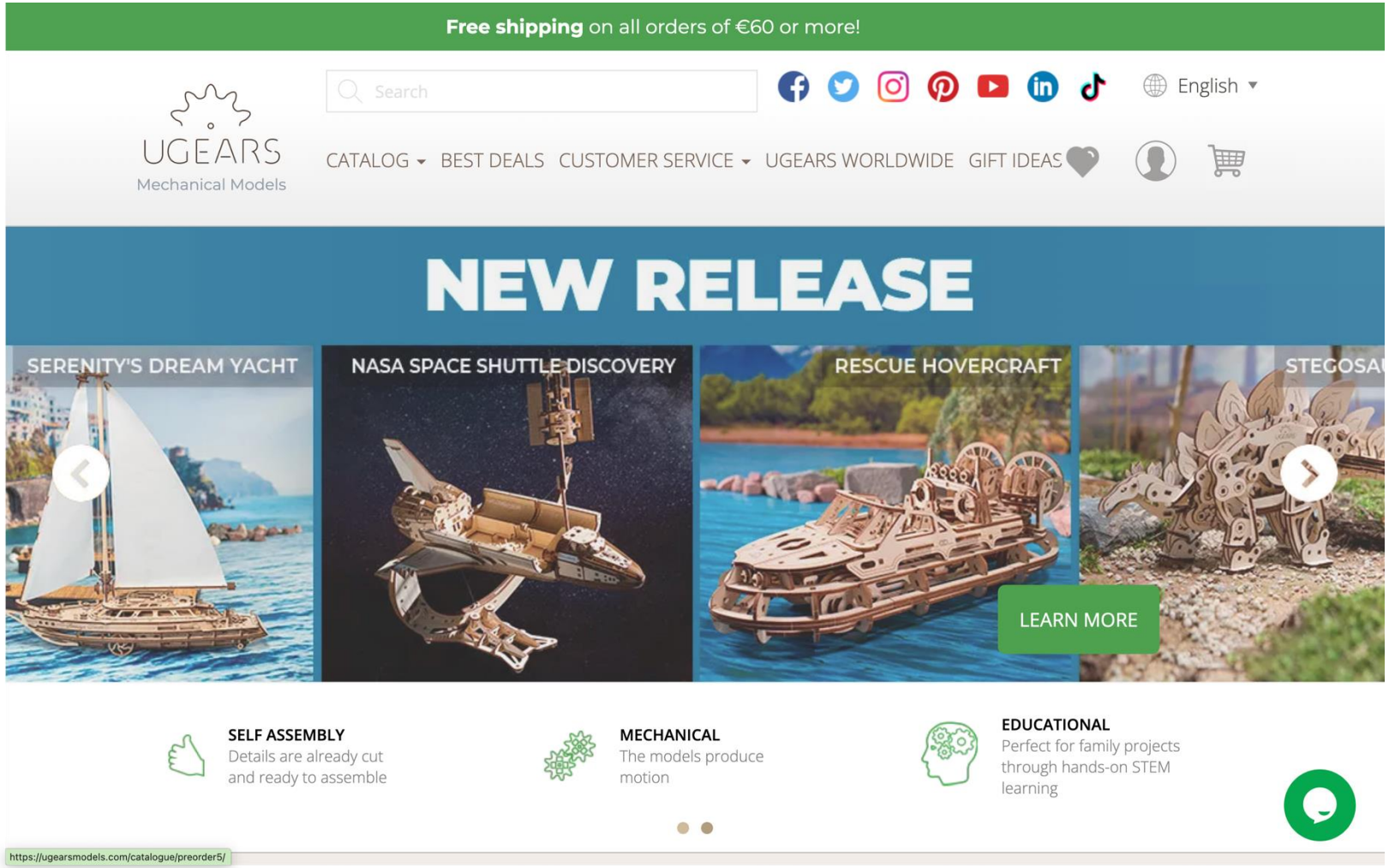


SAP HANA Cloud Free Tier

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

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

https://ugearsmodels.com/



SCAN ME

Thank you // Дякую! // ধন্যবাদ

Contact information:

Witalij Rudnicki, **SAP Developer Advocacy**

© 2023 SAP SE or an SAP affiliate company. All rights reserved. See Legal Notice on www.sap.com/legal-notice for use terms, disclaimers, disclosures, or restrictions related to SAP Materials for general audiences.