ML workshop 2020 - outline

# Zielsetzung

* Überblick AI, ML, Big Data, Data Mining, ... Begriffsabgrenzung: Was gehört wohin?
* Methoden, Anwendungsfelder und relevante Use Cases
* Daten: Voraussetzungen, Preprocessing Schritte und Einschätzung von Ressourcenbedarf
* Explorativen / experimentellen Character von Data Analytics ML Projekten
* Hands-On: Realistischer Praxis-Mini-Use-Case in Teamarbeit

# Rahmenbedingungen

* Zeitraum Mitte März 2020
* Dauer 2x 1 Tag ML-Grundlagen/Training plus ein ½ Tag Ideation/Konsolidierung
* Teilnehmer: ca. 2x 8-10
* Datenbasis: Offene public verfügbare Datensets (Kaggle) mit fachlichem Bezug
* Software: Python, Jupyter Notebooks gehostet in der Cloud (VMs via Google oder CoCalc)
* Business Laptops der Teilnehmer mit Standard Webbrowser (Chrome)
* Vorbereitete Jupyter Notebooks, die die Teilnehmer öffnen und bearbeiten

# **Themenüberblick**

**A Grundlagen Know-How**

1. Einführung & Überblick Themenfelder
2. Daten & Preprocessing
   * Überblick Workflow
   * Trainingsmatrix und Labelvektor
   * Datenaufbereitung (Normalisierung, kategorielle Daten, ...)
3. Unsupervised Learning
   * K-means Clustering
4. Supervised Learning
   * Lernen von Fehlerfunktionen
   * Datenaufbereitung (Missings, Normalisierung, kategorielle Daten, ...)
   * Lineare Regression
   * Classification: Naive Bayes, Decision Trees / Random Forest, MLP
   * Text Mining
   * Modellauswahl: Confusion Matrix / Cost Matrix
   * Accuracy / Precision / Recall
   * Bias-Variance Dilemma (Overfitting & Generalisierung)

# **Ablaufplan tag 1 | 8.00 – 17.30 Uhr**

**8.00 - 12.00 (15 Min. integrierte Kaffeepause)**

* Intro / Vorstellung / Erwartungen
* Grundlagen Know How Machine Learning (siehe Workshopinhalte oben)

**12.00 - 13.00 Mittagspause**

**13.00 - 14.30 Gemeinsame Bearbeitung - Anwendungsbeispiel**

* Restwertschätzung Gebrauchtwagen (Regressionsmodell)
* Einführung Jupyiter Notebooks, Daten laden & explorieren
* Regressionsmodell trainieren und validieren

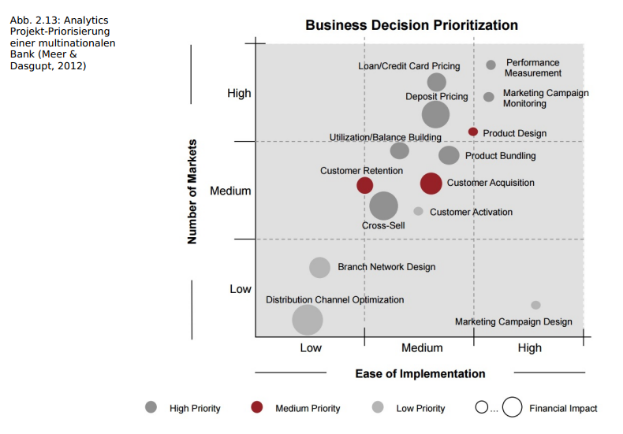
**14.30 - 17.00 “Hands-On”: Mini Use Case in 3 Teams (Kaffeepause individuell)**

* Vorstellung Case u. Aufgabenstellung / Teameinteilung
* Teamarbeit auf Mini-Use Case
* Kurze Vorstellung & Diskussion Teamergebnisse (jeweils 10 Minuten) ab 16.30 Uhr

**17.00 - 17.30 Closing Session Tag 1**

* Eindrücke Workshoptag 1: Überlegungen, Fragen, Ideen
* Next Steps / Vorbereitung Ideation Workshop am Tag 3

# **Ablaufplan tag 2 – 27.3.2020 | 8.00 – 12.00 Uhr**

1. **- 8.30 Kick-Off Tag 2**

* Recap Learnings Tag I / Fragen

**8.30 – 11.00 Vorstellung Bereichsspezifische  
Ideen und Ansätze**

**11.00 - 12.00 Erstellung Ideation Matrix o.ä.**

**12.00- 12.30 Wrap-Up / Closing**

