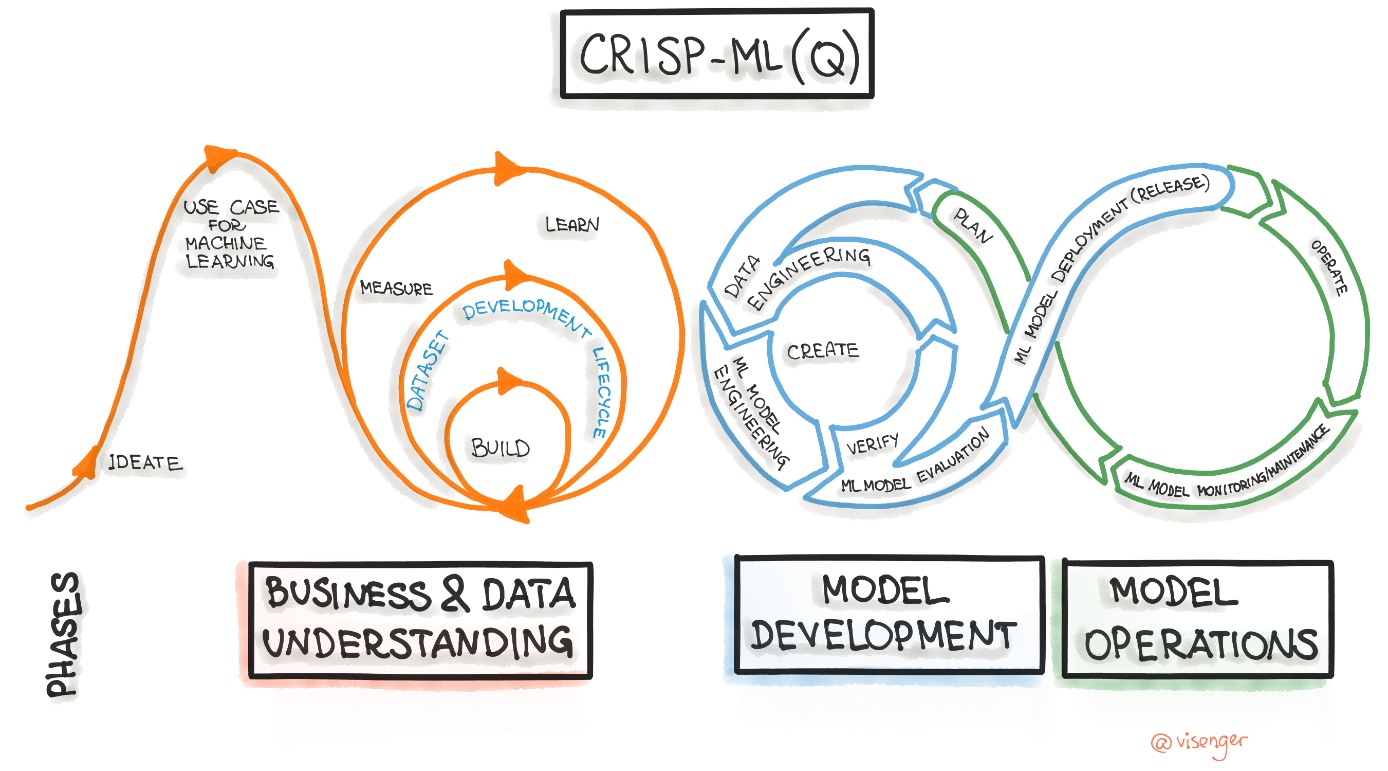
1. **基于机器学习的软件开发**应该遵照一定的**开发标准或规范**，标准制定的目的是保证通用和质量；



注意：上图显示，机器学习开发流程中大多步骤是**迭代和探索性**的；

2. 开发流程中的每个阶段都需要进行质量保证，质量保证方法的实施需要定义**需求和约束** (如：性能要求、数据质量要求、模型稳健性要求等)，然后在实例化特定任务 (如：ML算法选择)的前提下，查明影响ML性能和质量的因素 (如：偏差、过拟合、缺乏可复现性)，提供风险因素出现时的处理方案 (如：交叉验证、过程文档化)。

3. 基于机器学习的软件开发过程可分为6个阶段：

**[1]** **任务和数据理解**：目标是确保项目可行性。a. 明确应用程序的范围、成功标准和质量验证方案；b. 明确任务的可行性；c. 了解数据质量和数据生成过程；d. 将任务目标转化为机器学习目标；

**[2]** **数据工程**：目标是为后续建模准备数据。a. 数据选择. 即选择对任务有价值并且质量合格的数据、过采样/欠采样解决类不平衡问题；b. 数据清洗. 即降噪、数据插补、错误数据检测与处理； c. 特征工程与数据增强. 即连续数据离散化、类别变量编码、数据变换；d. 数据标准化. 即去量纲化，特别对于采用梯度下降算法的模型而言；

**[3]** **机器学习模型工程**：目标是指定要构建的模型。a. 确定模型质量评估指标. 即常见的**预测性能(含可泛化性)**、**健壮性**、**公平性**fairness、**可扩展性**scalability、**可解释性**、**可复现性**reproducibility、**模型复杂度**和**模型资源需求**； b. 建模，包括模型选择、模型限定(model specialization)和模型训练. 即根据领域知识将模型实例化；c. 记录模型训练元数据 (如：算法、数据集划分、超参数、运行环境、随机种子). 即为保证模型可复现性；

**[4]** **机器学习应用程序的质量保证**：目标是检验和保证模型的质量。a. 检验模型质量. 即在测试集上验证模型性能、使用带噪声数据或错误数据评估模型稳健性，等；b. 评估模型是否满足[1]中指定的成功标准；

**[5]** **模型部署**：目标是将ML模型集成到现有软件系统。a. 封装模型. 如：插件、组件等；b. 生产环境中的模型评估. 如在线测试；c. 确保用户接受度和可用性. 如提供用户验收和可用性测试；

**[6]** **质量监控与模型维护**：目标是监控模型性能并对其维护。a.如果需要，重新训练模型 (迭代前述流程)；b. 收集新数据并标注。

参考文献：

[1] 本文翻译自：CRISP-ML(Q). The ML Lifecycle Process. <https://ml-ops.org/content/crisp-ml>.

[2] ML模型评价指标：Zhang, J., et al., 2019. Machine Learning Testing: Survey, Landscapes and Horizons. <https://arxiv.org/pdf/1906.10742.pdf>.