

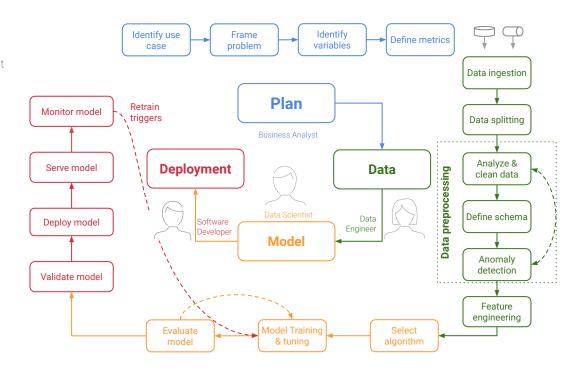
Data Science Lifecycle & TensorFlow Extended

Prof. Dr. Jan Kirenz HdM Stuttgart

Lifecycle

of an ML System

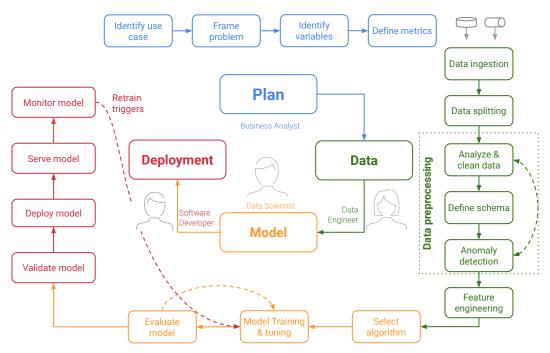
Plan | Data | Model | Deployment



Common issues

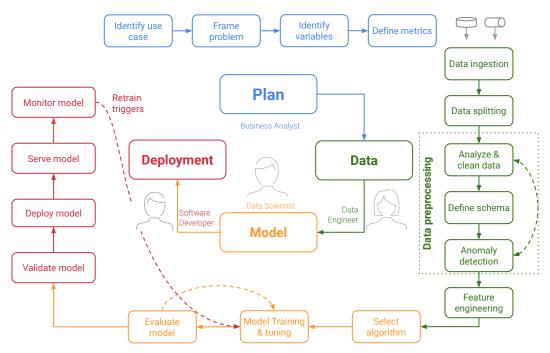
- Lack of reuse and duplication
- Inconsistency (data, code, models)
- Manual and slow transition from PoC to production

Lifecycleof an ML System Plan | Data | Model | Deployment



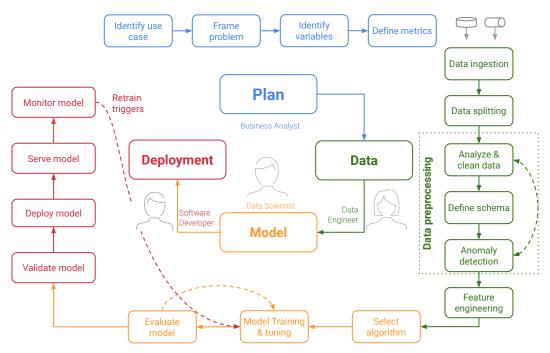


Lifecycleof an ML System Plan | Data | Model | Deployment



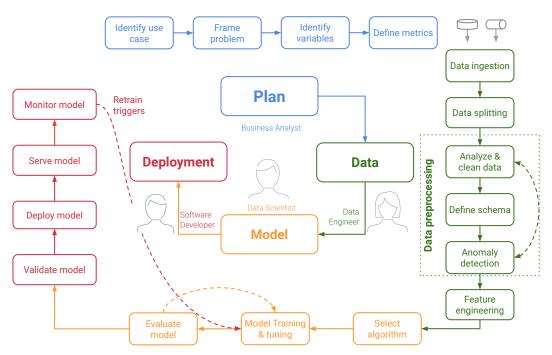


LifecycleOf an ML System Plan | Data | Model | Deployment



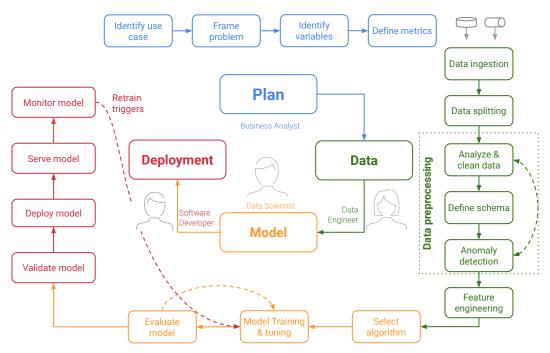


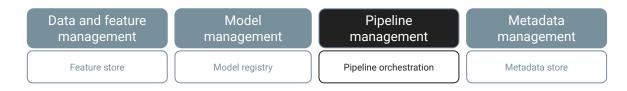
Lifecycleof an ML System Plan | Data | Model | Deployment



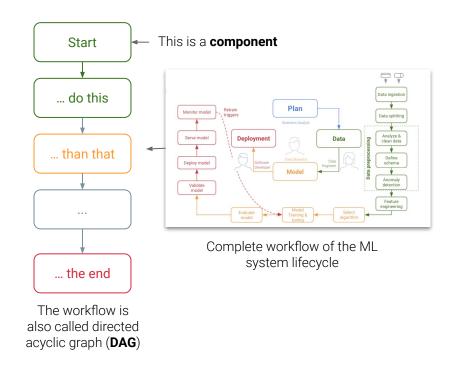


Lifecycleof an ML System Plan | Data | Model | Deployment

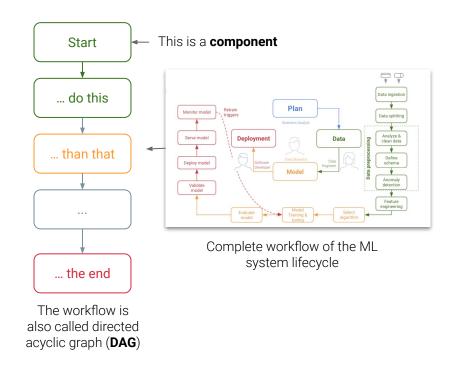




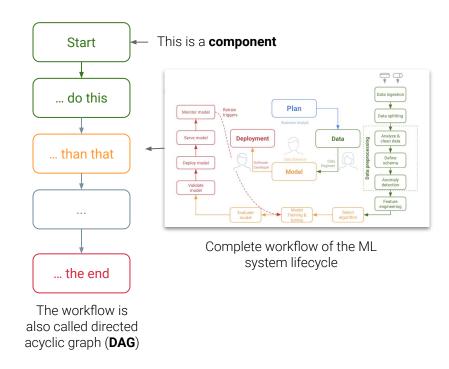
- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)



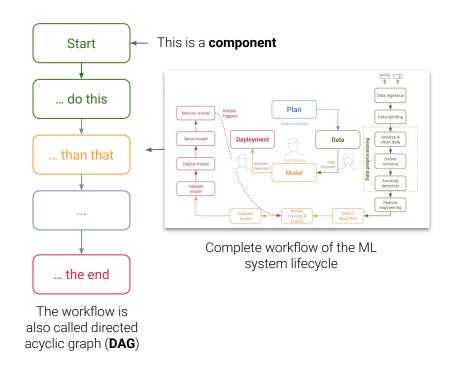
- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)



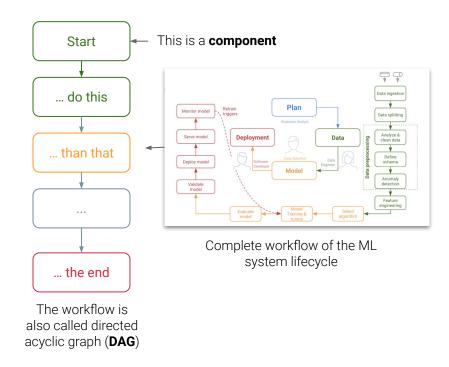
- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)

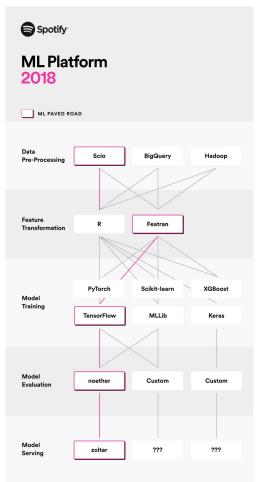


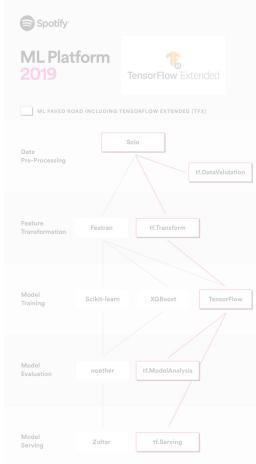
- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)

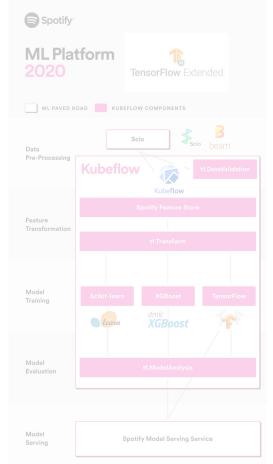


- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)



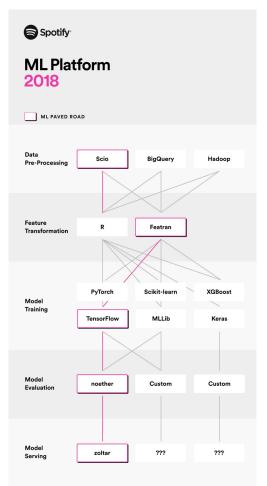


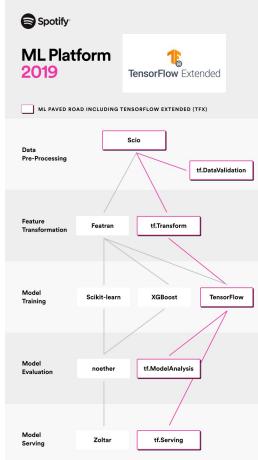


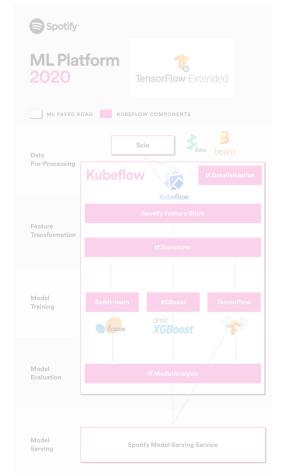




Source: Baer & Ngahane (2019) Prof. Dr. Jan Kirenz

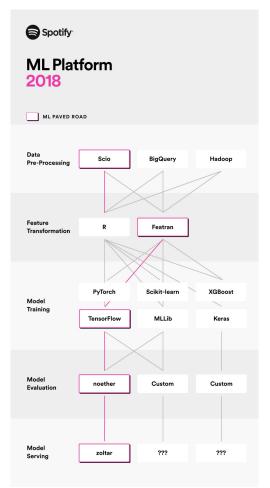


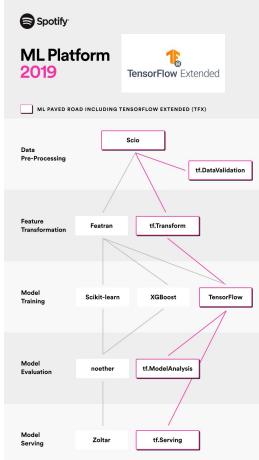


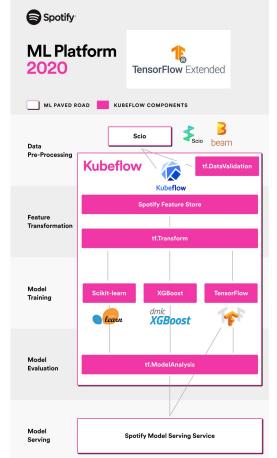




Source: Baer & Ngahane (2019) Prof. Dr. Jan Kirenz









Source: Baer & Ngahane (2019)

- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines









- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure, AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines









- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure, AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines









- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure, AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines









- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure, AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines





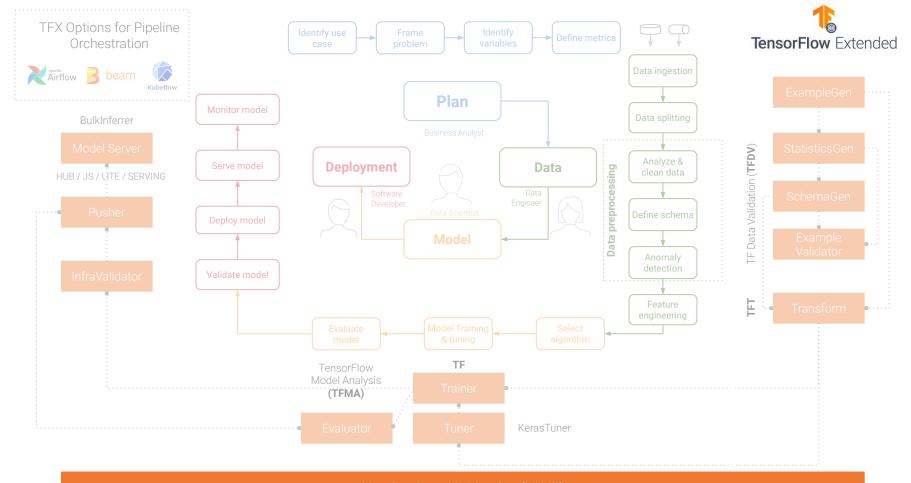


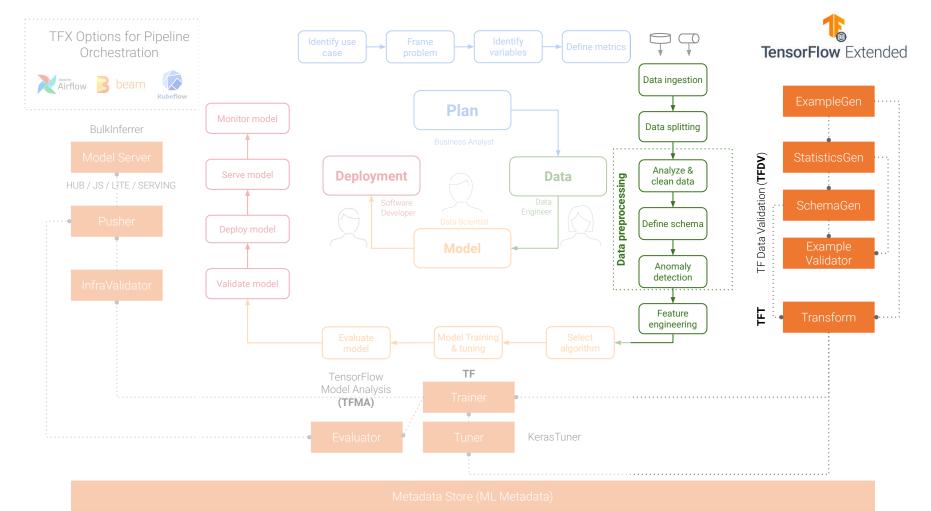


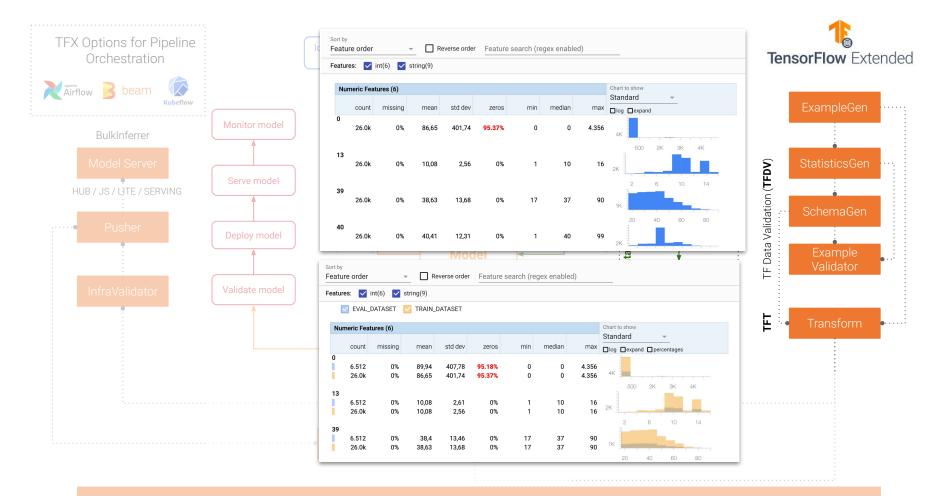
TFX 1.0 (19.05.21)

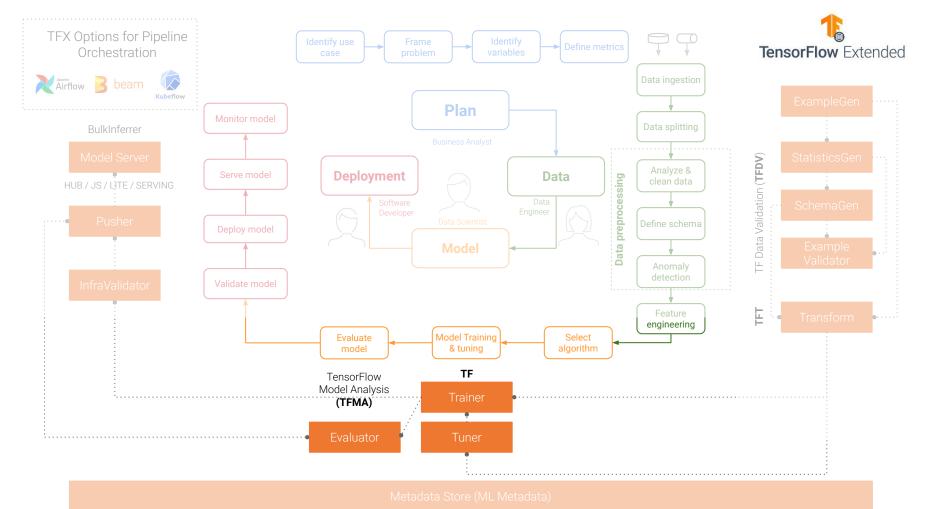
- Enterprise-grade support
- Security patches and select bug fixes for up to three years
- Guaranteed API & Artifact backward compatibility

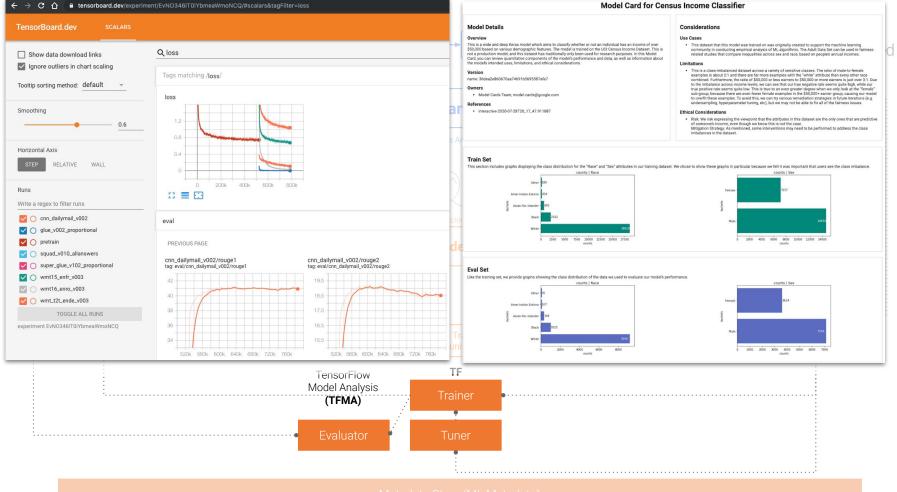


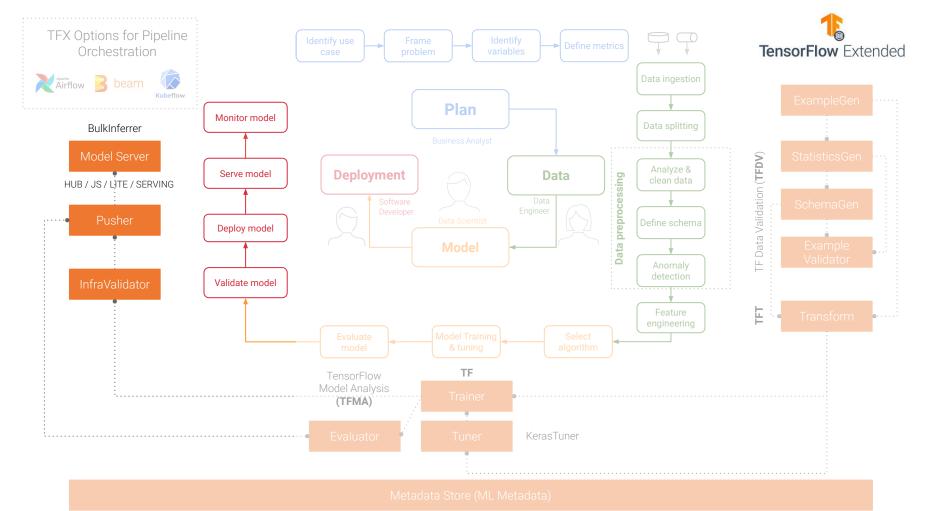


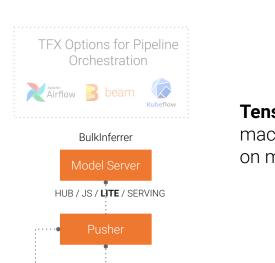








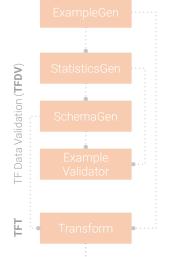


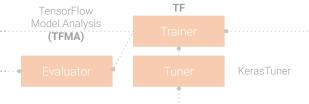


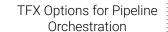


TensorFlow Lite is a set of tools that enables on-device machine learning by helping developers run their models on mobile, embedded, and IoT devices.









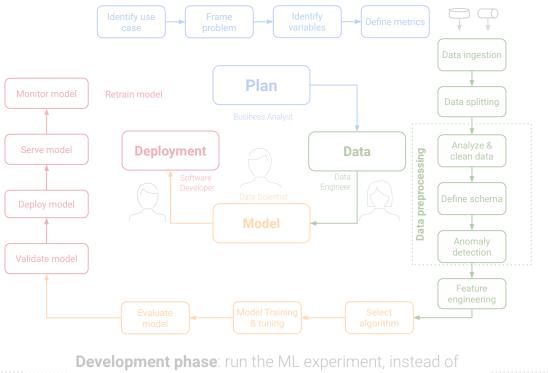








of the ML pipeline based triggering conditions.



TensorFlow Extended

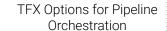
Data preparation

automatically

ingest, validate and transform

data and provide features to models

phase:



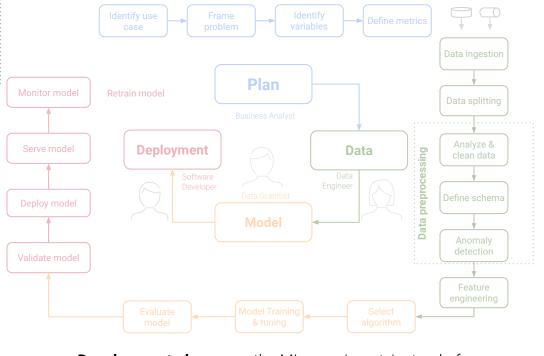








of the ML pipeline based triggering conditions.

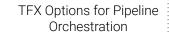




Data preparation phase:

ingest, validate features to models

Development phase: run the ML experiment, instead of manually executing each step.

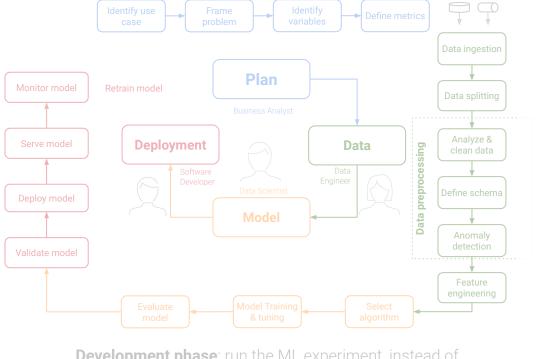








automate the execution of the ML pipeline based on a schedule or certain triggering conditions.

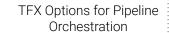




Data preparation phase:

ingest, validate features to models

Development phase: run the ML experiment, instead of

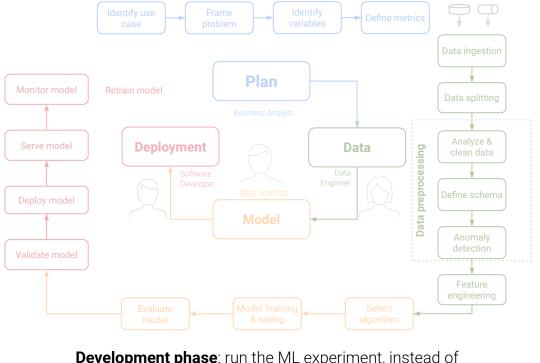








automate the execution of the ML pipeline based on a schedule or certain triggering conditions.





Data preparation phase:

automatically ingest, validate and transform data and provide features to models

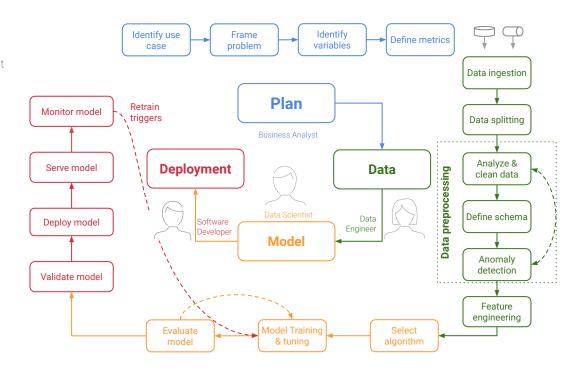
Development phase: run the ML experiment, instead of manually executing each step.

Summary

Lifecycle

of an ML System

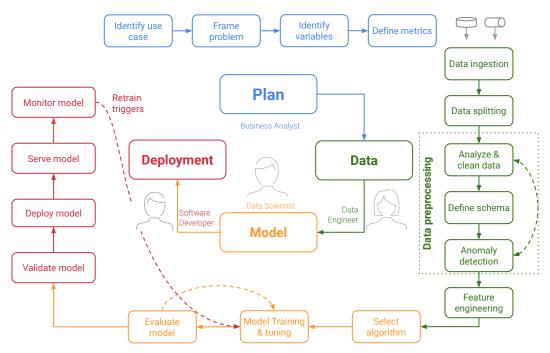
Plan | Data | Model | Deployment

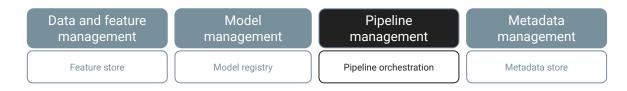


Common issues

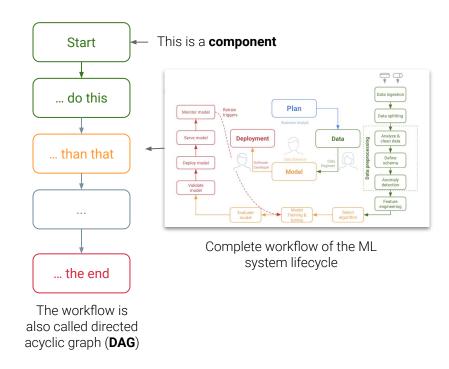
- Lack of reuse and duplication
- Inconsistency (data, code, models)
- Manual and slow transition from PoC to production

Lifecycleof an ML System Plan | Data | Model | Deployment





- Description of an ML workflow
- A pipeline component is a self-contained set of user code that performs one step in the pipeline
- Includes the definition of the configuration and inputs required to run the pipeline (e.g. model hyperparameters)



- Google-production-scale machine learning (ML) platform based on TensorFlow
- Portable to multiple environments (Azure, AWS, Google Cloud, IBM, ...)
- Python based toolkit; can be used with notebooks
- Helps you orchestrate your ML process:
 Apache Airflow, Apache Beam or Kubeflow pipelines









