End to End Machine Learning with Tensorflow Extended (TFX)

Introduction

Jaganadh Gopinadhan

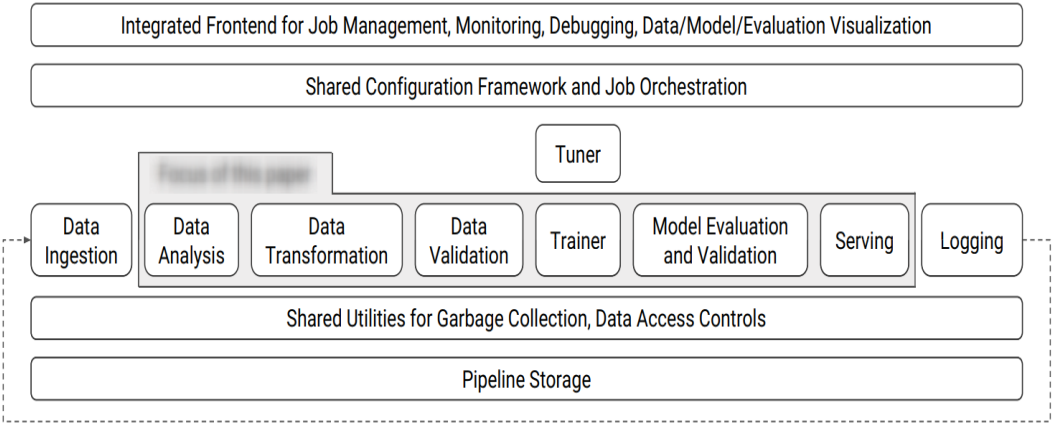
August 18, 2019

http://jaganadhg.github.io/blog

Why TFX

- Machine Learning is an integral part of any imaginable IT system today.
- Development, deployment and maintenance of a Machine Learning platfrom required orchestration of various components.
- Most of the times it is a team work of Machine Learning Engineers, DevOps (AIOps?), Data Engineers and IT support professionals.
- Glue codes ad custom scripts combinations used in Machine Learning applications production deployment is a major drawback.
- Monitoring prediction trends and drift in input is observed, analyzed and fixed by fire fighting.

Why TFX


Integrated Frontend for Job Management, Monitoring, Debugging, Data/Model/Evaluation Visualization
Shared Configuration Framework and Job Orchestration
Tuner
Data Ingestion
Data Analysis
Data Transformation
Data Validation
Trainer
Model Evaluation and Validation
Serving
Logging
Shared Utilities for Garbage Collection, Data Access Controls
Pipeline Storage

Key ML Challenges addressed

- One platform many learning tasks.
- Continuous training and serving
- Human-in-the-loop
- Production reliability and scalability

What is TFX

Tensorflow Extended a.k.a TFX is a Goolge Open Source project based on Tensrflow. TFX is a production scale Machine Learning Platform. The capabilities offered by TFX includes libraries to integrate common components required to define, launch and monitor Machine Learning systems in scale.
TFX provides integration to various modern tools such as Kubernets, Apache Beam and Apache Airflow.

Let’s Explore TFX

Necessary code and supporting materials for TFX tutorial is hosted at https://github.com/jaganadhg/tfx-tutorial

Thank You

Stay Tuned

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of my employer.