PRODUCT BRIEF

Fast, Accurate and Real-Time Predictions with PipelineAI

The Fastest, Safest, and Most-Flexible Machine Learning Application Platform on the Market

At a Glance

PipelineAI is a real-time model-prediction platform that optimizes, scales, and explains the predictions of any cognitive decision-making system at the speed of thought. Our platform clears the path from local-laptop research to global-scale production. PipelineAI is a zero-touch continuous model optimization platform that automatically generates, deploys, and compares hundreds of model variants - from just a single model. Scientists and Engineers can instantly launch any experiment using any language and any ML/AI framework.

KEY BENEFITS

- One-click model deployment directly from your laptop or staging server
- Continuous model optimizations throughout the life of the model
- Online model training to improve prediction relevancy after initial deploy
- Instant prediction explanations for each and every decision

Business Challenges

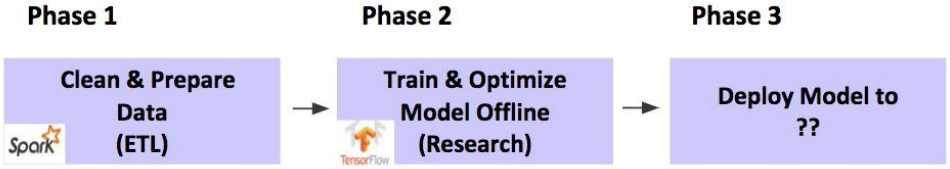
Today's production-focused machine learning requires a platform that continuously adapts to real-world conditions such as fraud detection, real-time recommendations, and autonomous driving. It is relatively easy to get started on a laptop and train a simple model with offline, historical data. However, moving from a single laptop to a large-scale, production-grade, multi-cloud cluster is a completely different challenge.

The Enterprise with traditional Machine Learning is struggling to take models from research to production due to widening skills gap, high maintenance cost, and poor runtime performance. Current ML solutions do not provide the trust, accuracy and speed that is needed to provide the financial benefits demanded by business leaders.

Solution Overview

PipelineAI platform reduces deployment time down to minutes by seamlessly optimizing and deploying all of your machine learning models as highly-available microservices regardless of the language or model framework. The PipelineAI Platform uses a serverless, cloud-agnostic architecture that allows you to 1) automate the infrastructure, 2) introduce new functionality as needed, and 3) production-ize models at the speed of thought. PipelineAI uses a unique combination of offline training metrics and online prediction performance to adaptively route user traffic to the best-performing models 24 hours a day. Additionally, multiple models can be dynamically combined to produce a more-accurate prediction. This is the PipelineAI Application Runtime Platform.

A typical ML pipeline consists of the following 3 major phases:



Phase 1 - Cleanse and Prepare Dataset

Data Scientist modifies the data to prepare for model training

Phase 2 - Develop and Train Model

Data Scientist and Data Engineer train the model offline in a research lab

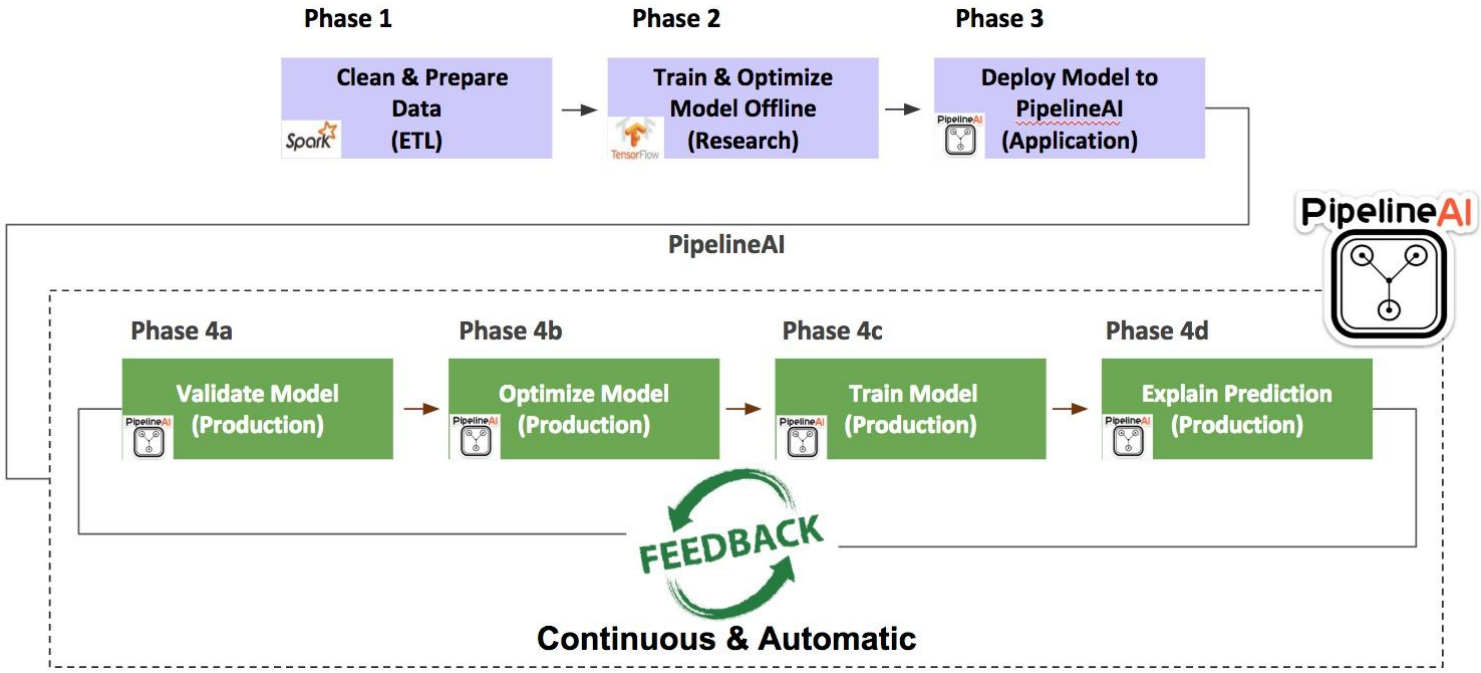
Phase 3 - Deploy the Model to Production Applications

Data Scientist hands-off the model to an ML Engineer or App Developer to re-implement for application use

For more information, please visit **pipeline.ai**

PipelineAI provides a high-performance implementation of Phase 3 - as well as a unique, adaptive, and continuous 4th phase that we call"VOTE". VOTE stands for "**V**alidation, **O**ptimization, **T**raining, and **E**xplanation."

Phase 4 - PipelineAI's Unique Continuous VOTE Phase



The PipelineAI automatically coordinates the validation, optimization, training and explainability of the model. This stage involves many automated cycles of model training and testing using different sets of online hyperparameters including system configuration, network speed, and real-time user data.

Validate Models Immediately After Deployment - and Through the Life of the Model

By default, PipelineAI generates many optimized model runtimes from a single uploaded model. These model runtimes are safely deployed and compared directly in live production using our PipelineAI "shadow" mode.

Optimize Predictions using Real-Time Conditions

Each model is continuously optimized based on predictive and system metrics collected from the PipelineAI Runtime

Train Your Models as New Data Arrives

The trained and optimized model and its metadata (e.g., accuracy) are exported to the PipelineAI ModelStore.

Explain Your Predictions for Maximum User Transparency

Explain & trace every single prediction for audit purposes (ie. GDPR)

For more information, please visit **pipeline.ai**