انجمن انفورماتیک ایران

از سلسله وبینارهای

گروه تخصصی معماری سازمانی انجمن انفورماتیک ایران

{مقدمه ای بر MLOps و معماری سازمانی مبتنی بر هوش مصنوعی }

{Introduction to MLops and AI Based Enterprise architecture}

{رضا گنجی}

{مدیر تولید نرم افزار}

{عنوان وبینار}
{webinar title}

آنچه خواهیم گفت...

- معرفی
- مقدمه ای بر DevOps
- چرایی MLOps
- معرفی MLOps
- ساختار سازمانی مبتنی بر MLOps
- معرفی ابزار ها
- معرفی MlFlow

{عنوان وبینار}
{webinar title}

مراجع :

- Practical MLOps by Noah Gift, Alfredo Deza
- RedHat MLOps base model

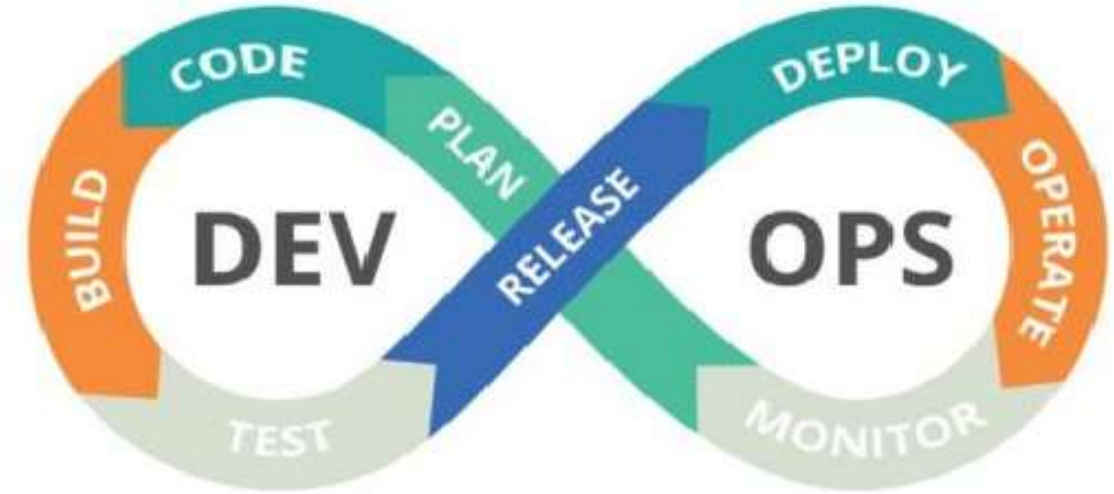
{عنوان وبينار}
{webinar title}

مشکلات موجود در ML

- عدم نگهداری تاریخچه الگوریتم ها و دلیل انتخاب مدل ها
- عدم وجود مانیتورینگ
- عدم وجود چرخه CI/CD
- عدم چابکی
- عدم نگهداری اجراهای مختلف
- عدم امکان مقایسه سریع الگوریتم های موجود

مقدمه ای بر DevOps

- تضمین کیفیت – عملکردی، غیرعملکردی و یکپارچگی داده‌ها
- امنیت و کنترل – انطباق با استانداردهای صنعت
- عملیات – انطباق در دسترس بودن بر اساس قراردادهای سطح خدمات
- چابکی



{عنوان وبینار}
{webinar title}

MlOps چیست؟

Wikipedia :

MLOps or ML Ops is a paradigm that aims to deploy and maintain machine learning models in production reliably and efficiently. The word is a compound of "machine learning" and the continuous development practice of DevOps in the software field. Machine learning models are tested and developed in isolated experimental systems. When an algorithm is ready to be launched, MLOps is practiced between Data Scientists, DevOps, and Machine Learning engineers to transition the algorithm to production systems.

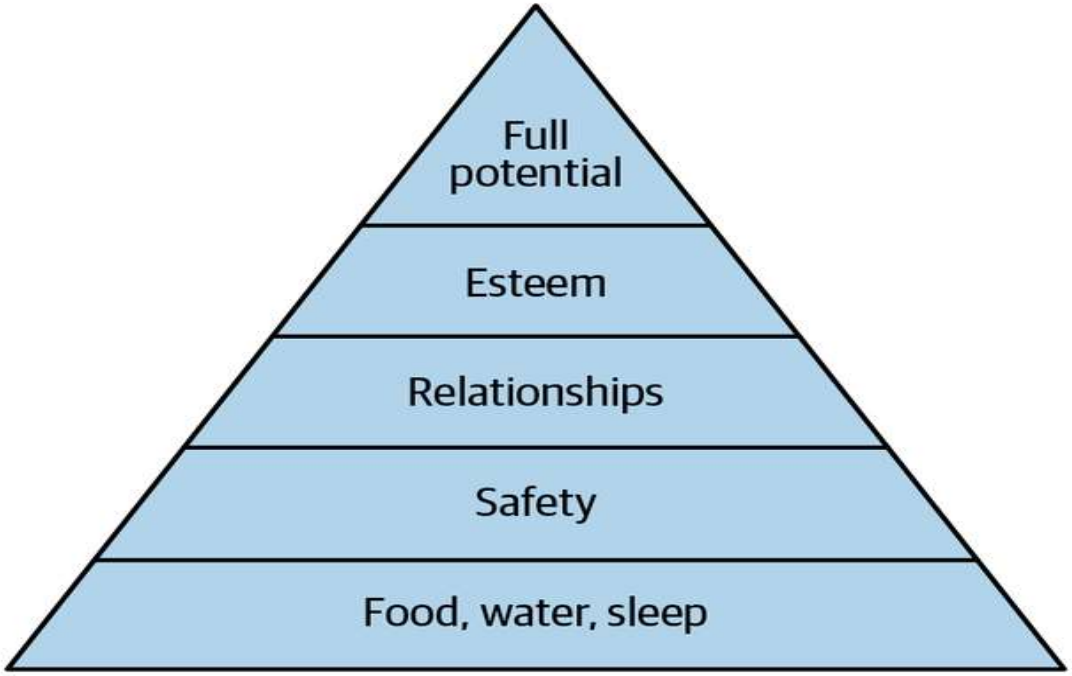
چرا MLOps ؟

- سرعت بیشتر ارائه محصولات (Faster Time To market)
- بهره وری بیشتر : استاندارد کردن و هزینه های عملیاتی کمتر
- تشخیص ریسک : نظارت بیشتر , پیش بینی ریسک , امنیت و ...

شباهت MLOps و DevOps

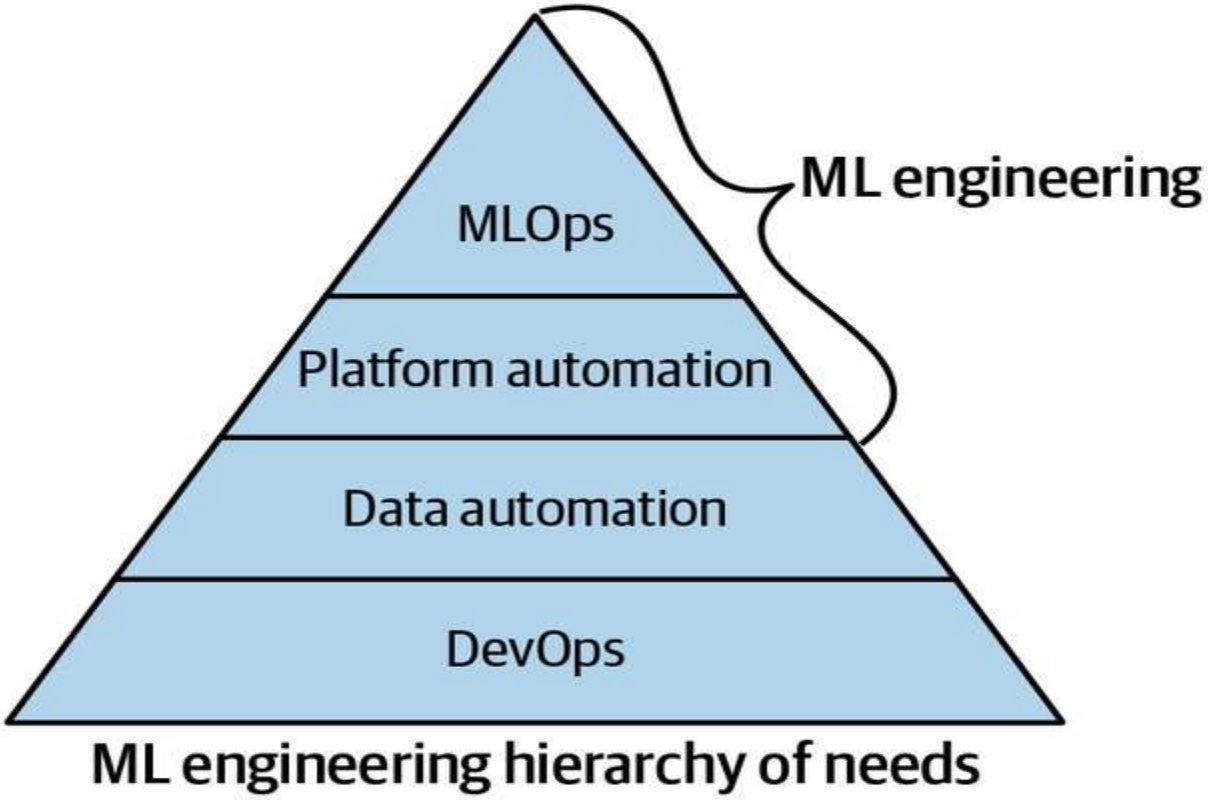
- Continues Integration
- Continues Delivery
- Microservices
- Infrastructure As a code
- Monitoring And Instrumentation
- Effective Technical Communication
- Effective Technical Project Management

هرم مازلو



Maslow's hierarchy of needs

هرم مازلو در MLOps



ML engineering hierarchy of needs

{عنوان وبینار}
{webinar title}

Data Automation یا Data Ops

Data Ops پایه اصلی MlOps هست و هدف ان ادغام DevOps و متدولوژی های چابک و تولید ناب هست برای رسیدن به دادهایی در راستای اهداف سازمان

Data Ops

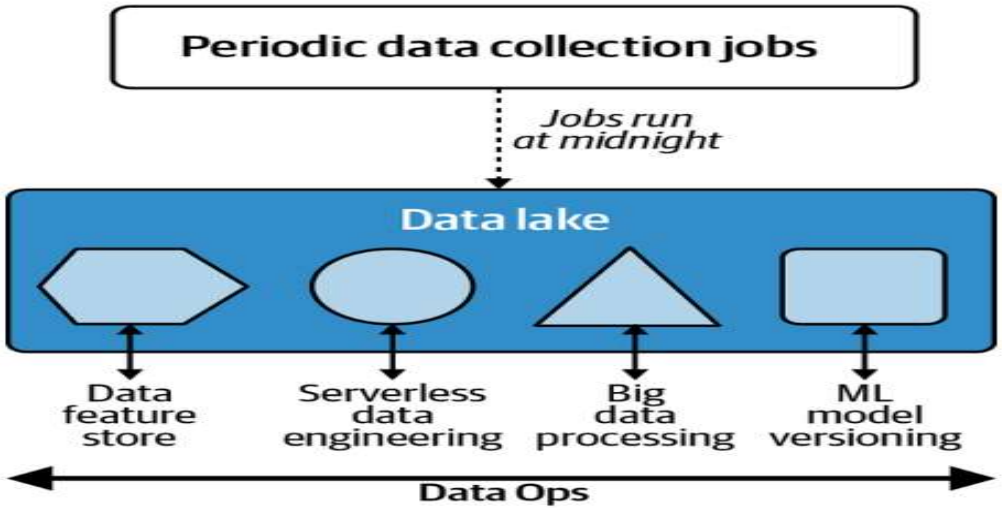
- همسو کردن داده ها با کسب و کار
- فرآیندهای چابک برای مدیریت داده ها و توسعه تجزیه و تحلیل فرآیندهای
- DevOpsبرای بهینه سازی کد، ساخت محصول و تحویل

پارادایم های اصلی Data Ops

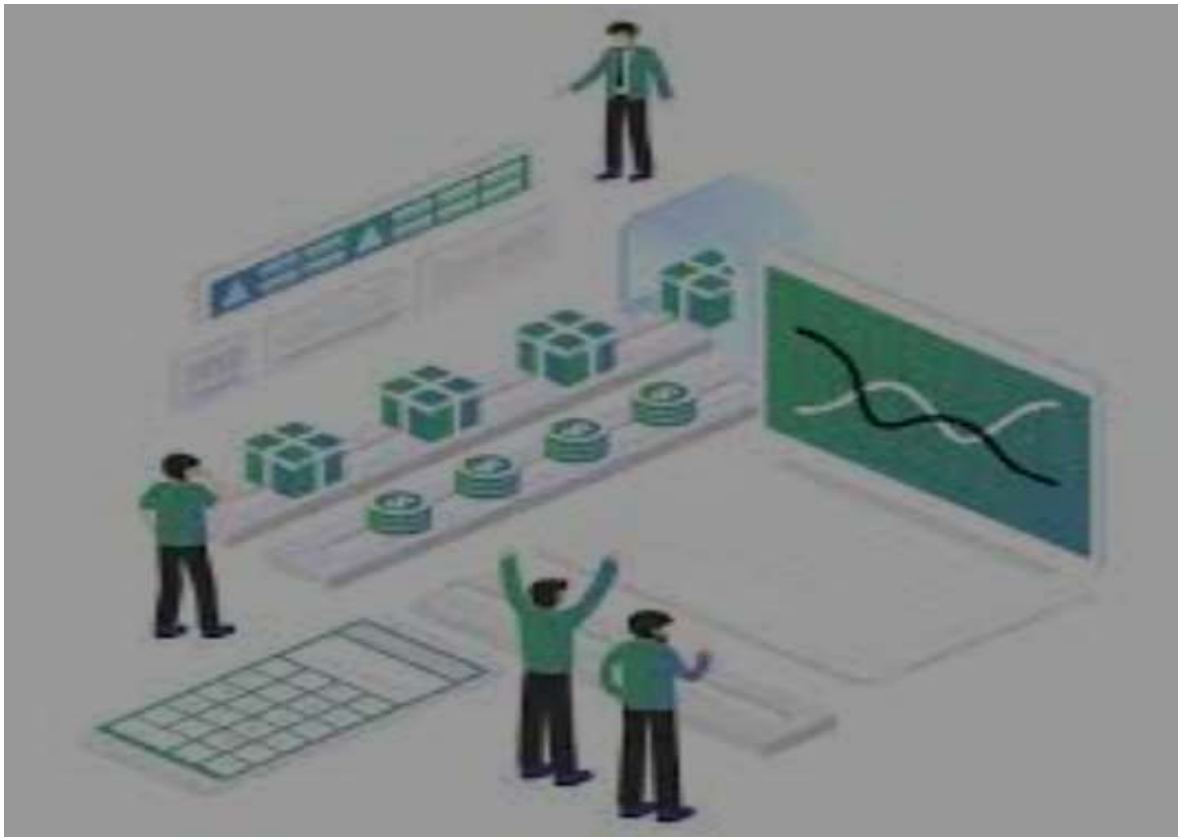
- Cross-functional communication
- Agile mindset
- Integrated data pipeline
- Data-driven culture
- Continuous feedback

وظایف اصلی Data Ops

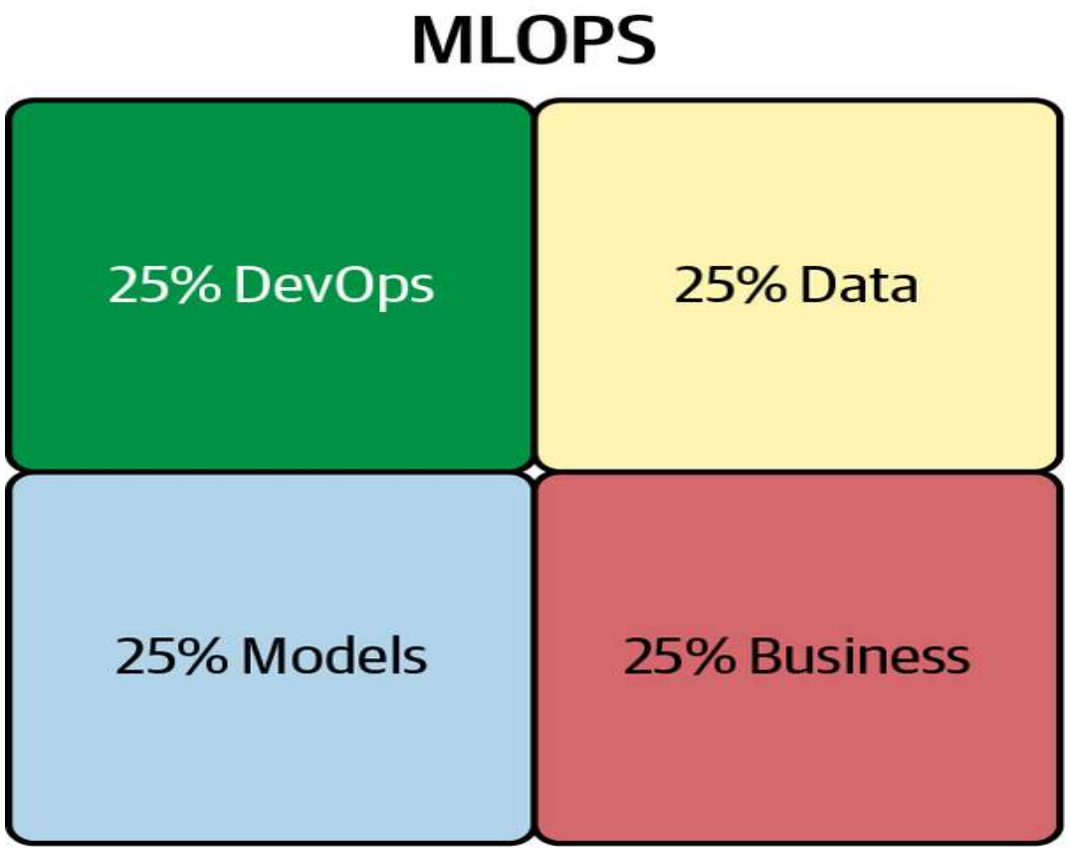
- Periodic collection of data and running of jobs
- Processing streaming data
- Serverless and event-driven data
- Big data jobs
- Data and model versioning for ML engineering tasks



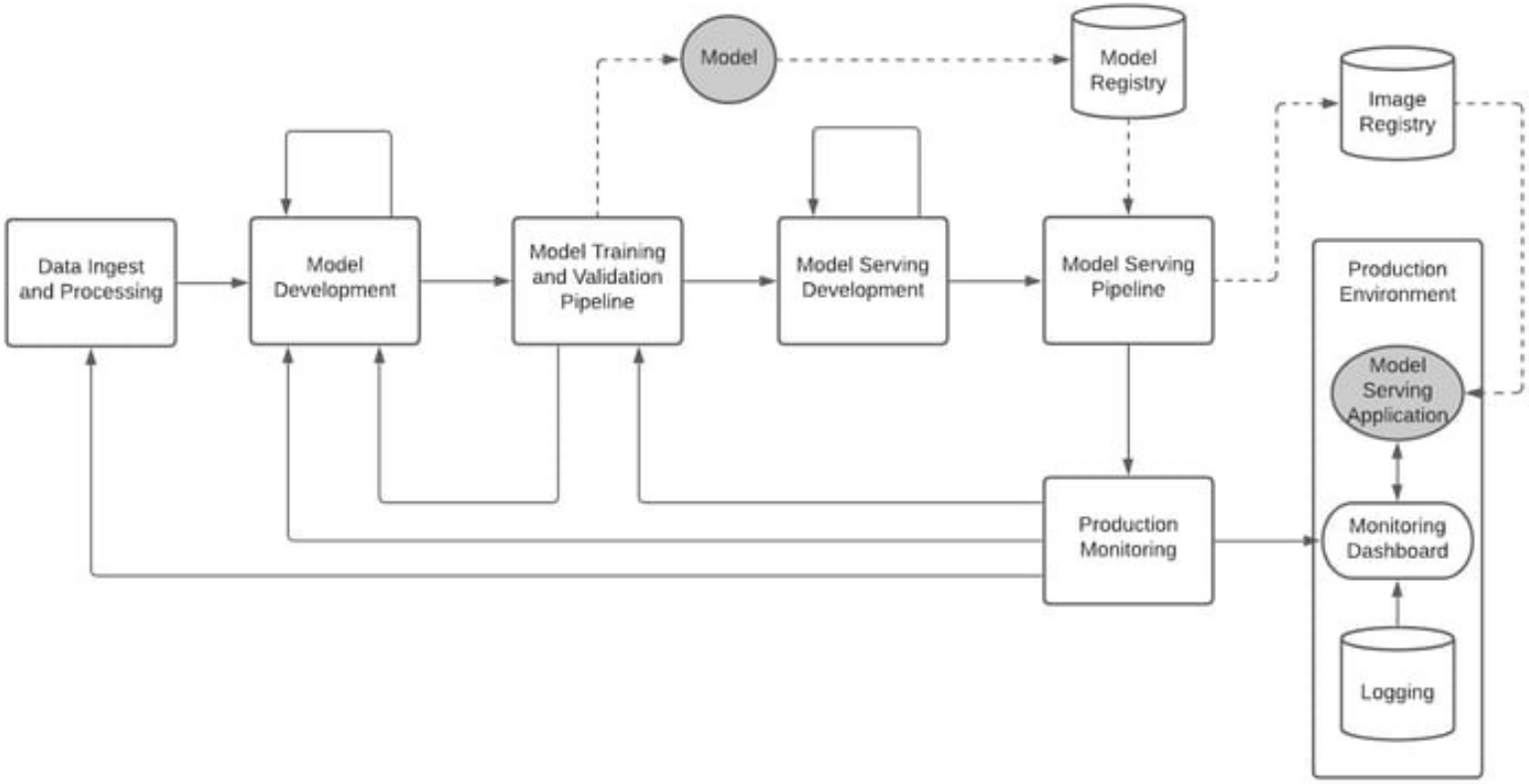
Platform Automation



توزیع حوزه ها در MLOps



معماری مرجع MLOps

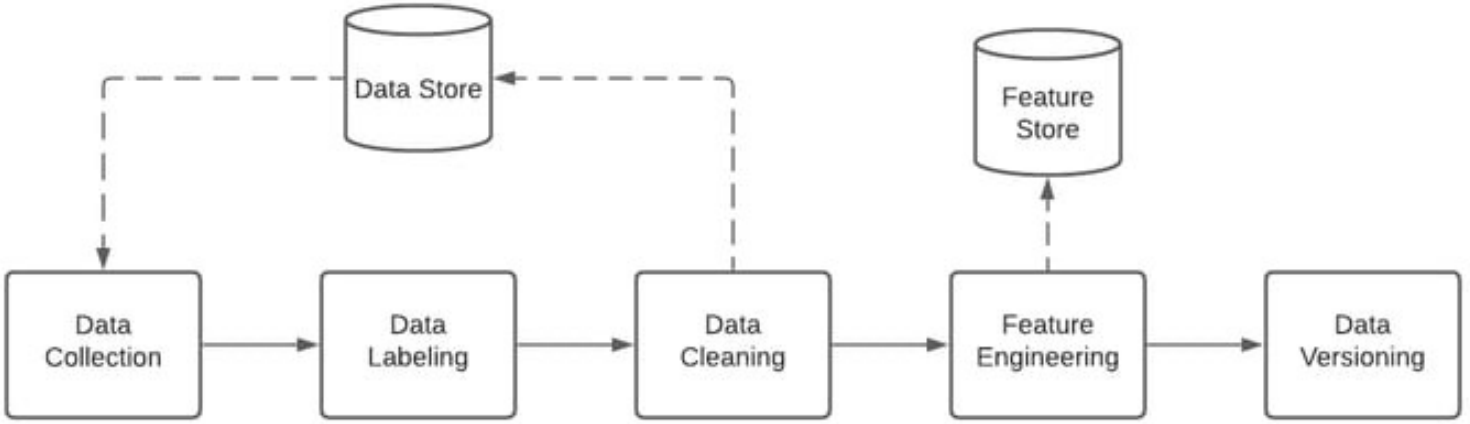


{عنوان وبینار}
{webinar title}

جمع اوری و پردازش داده ها (Data Ingest and Processing)

مهندس داده ها :

- جمع اوری داده ها از دیتابیس یا data lake (SQL database or HDFS/S3) یا از چند منبع مختلف
- برچسب گزاری داده ها
- پاکسازی داده ها (Normalization or Deduplication)
- مهندسی بردار ویژگی ها



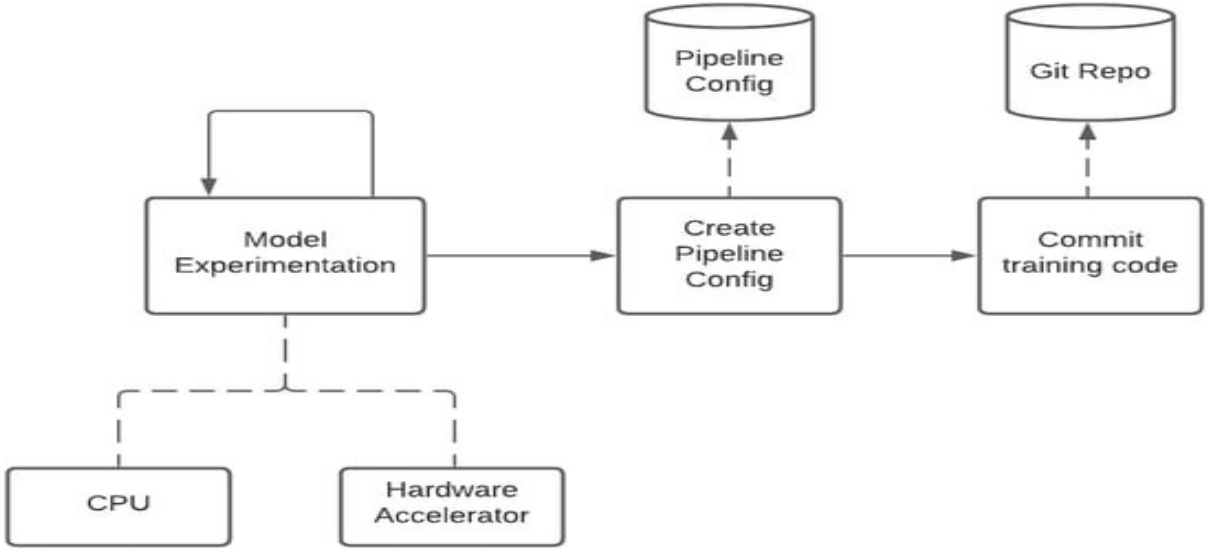
{عنوان وبینار}
{webinar title}

توسعه مدل(Model Development)

Data scientist:

- نوشتن کد برای ساختن مدل یادگیری ماشین

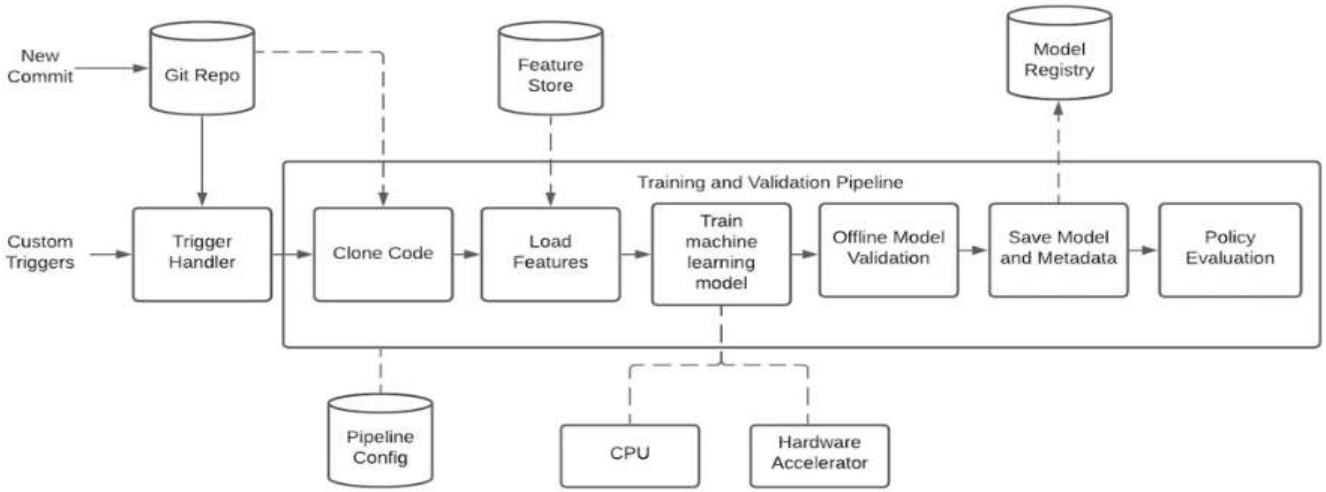
- تولید مدل های مختلف با استفاده از داده ها
- مقایسه مدل ها با ابزار یا ماتریس سردرگمی (Confusing matrix)
- ارزیابی توان سخت افزار و پردازنده برای الگوریتم ها
- تنظیمات خط تولید



{عنوان وبینار}
{webinar title}

یادگیری و اعتبار سنجی (Model Training and Validation)

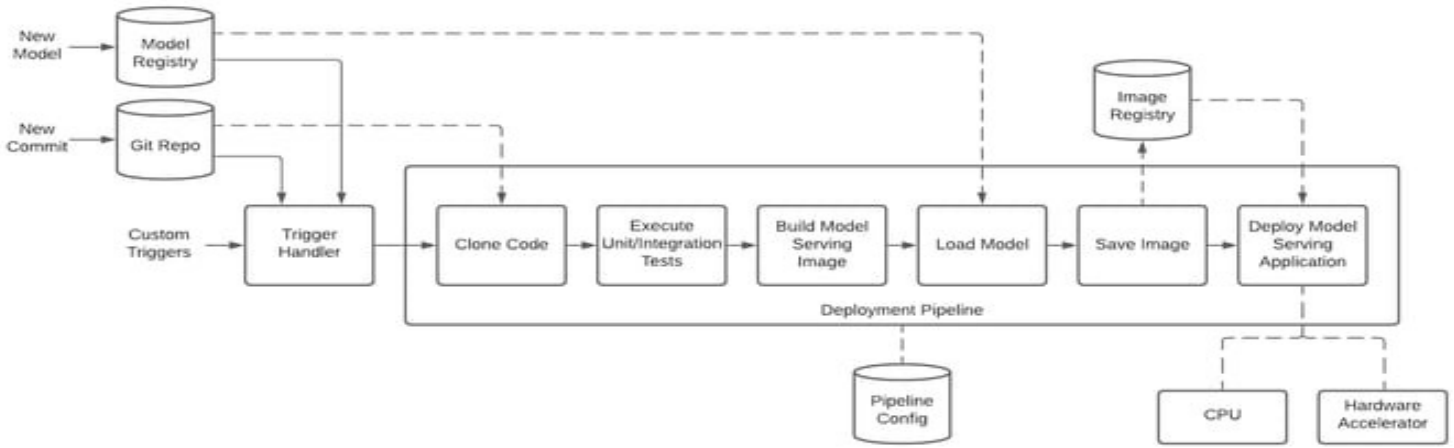
- فرایند اتوماتیک و تکرار شونده
- فعال شدن در زمان تغییر کد ها , رفع باگ , تغییر کتابخانه های نرم افزاری



{عنوان وبینار}
{webinar title}

Model Serving

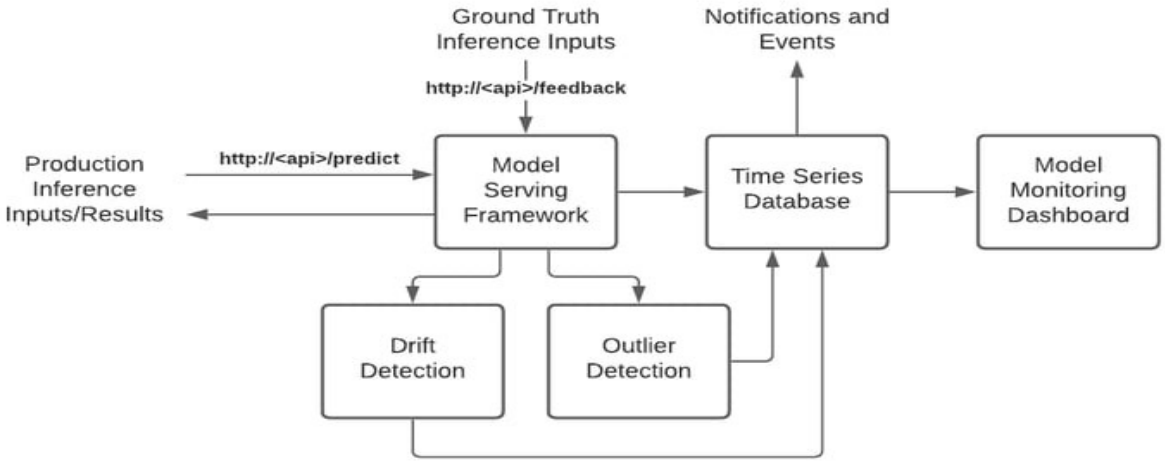
- معماری مبتنی بر میکروسرویس
- خروجی مبتنی بر داکر
- انتخاب بهترین مدل



{عنوان وبینار}
{webinar title}

مانیتورینگ محصول(Production Monitoring)

- نگهداری رفتار سیستم
- Drift detection
- Outlier detection
- سیستم هشدار



{عنوان وبینار}
{webinar title}

پروفایل ریسک در استقرار

نوع تغییر	شرح	سطح ریسک
Model Serving Microservice - Bug Fixes	تغییرات کد کم در سطح رفع باگ	کم
Model Serving Microservice - Library updates	تغییرات کتابخانه های نرم افرازی	کم
Model Serving Microservice - Base Image Update	تغییرات base image در کانتایرنر	کم
Model Serving Microservice - Feature Enhancements	اضافه کردن ویژگی های جدید مثل متریک های جدید	متوسط
Model Serving Framework - API Change	تغییرات API مورد استفاده سایر برنامه های کاربردی	متوسط
Model - Minor Training Dataset Change	تغییرات جزیی در داده ها	کم
Model - Major Training Dataset Change	تغییرات مهم در داده ها مانند نوسازی داده ها , تغییر برچسب ها و یا تغییر بخش مهمی از داده ها	متوسط
Model - Hyperparameter Tuning	تغییرات در هایپر پارامتر ها مثل نرخ یادگیری , دقت مدل و over/under fit	متوسط
Model - Architecture or Model Type Change	تغییرات مدل ها مثل سویچ از الگوریتم random forest به الگوریتم Gradient boosting ویا تغییر لایه های نرم افزاری	زیاد

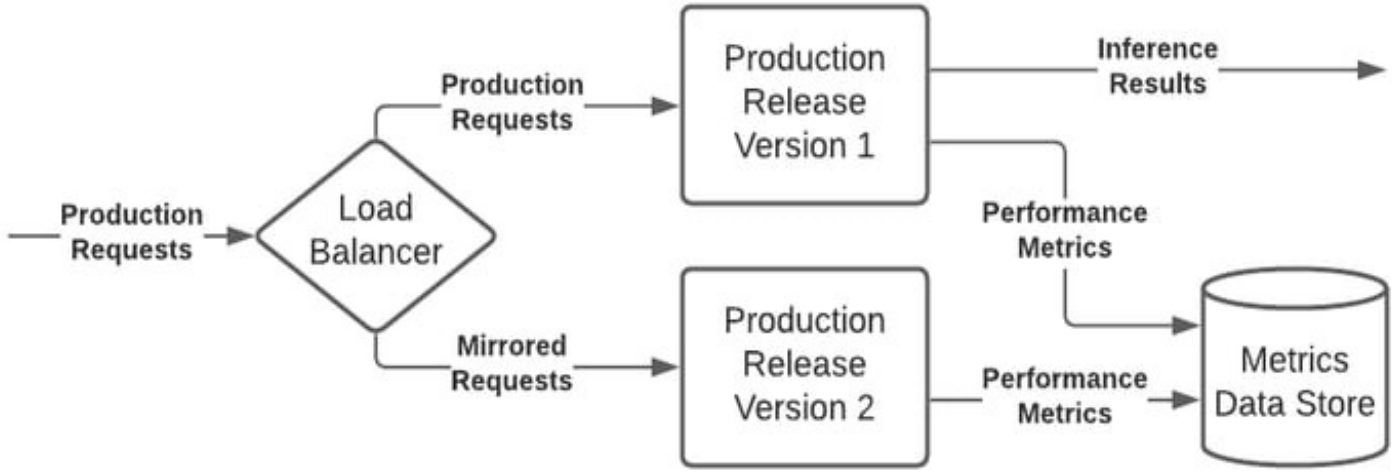
{عنوان وبینار}
{webinar title}

روش های تست و استقرار

- Shadow Testing
- A/B Test Pattern
- Canary Test Pattern

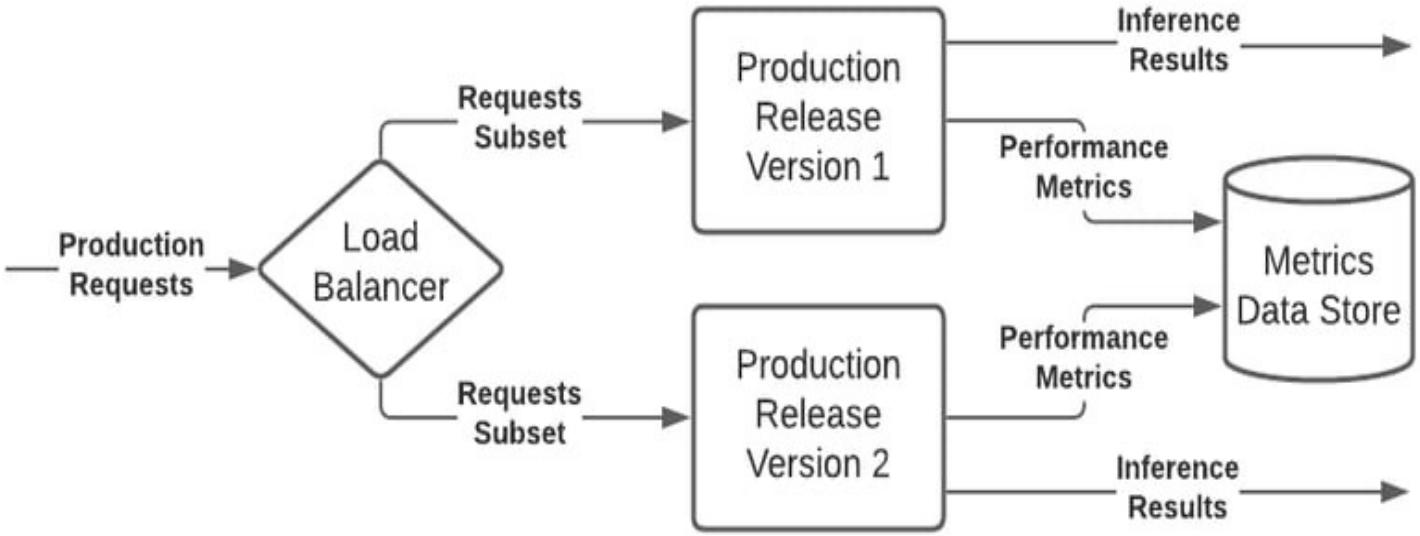
Shadow Testing

- ریسک گریزترین حالت تست
- تاخیر در اجرا



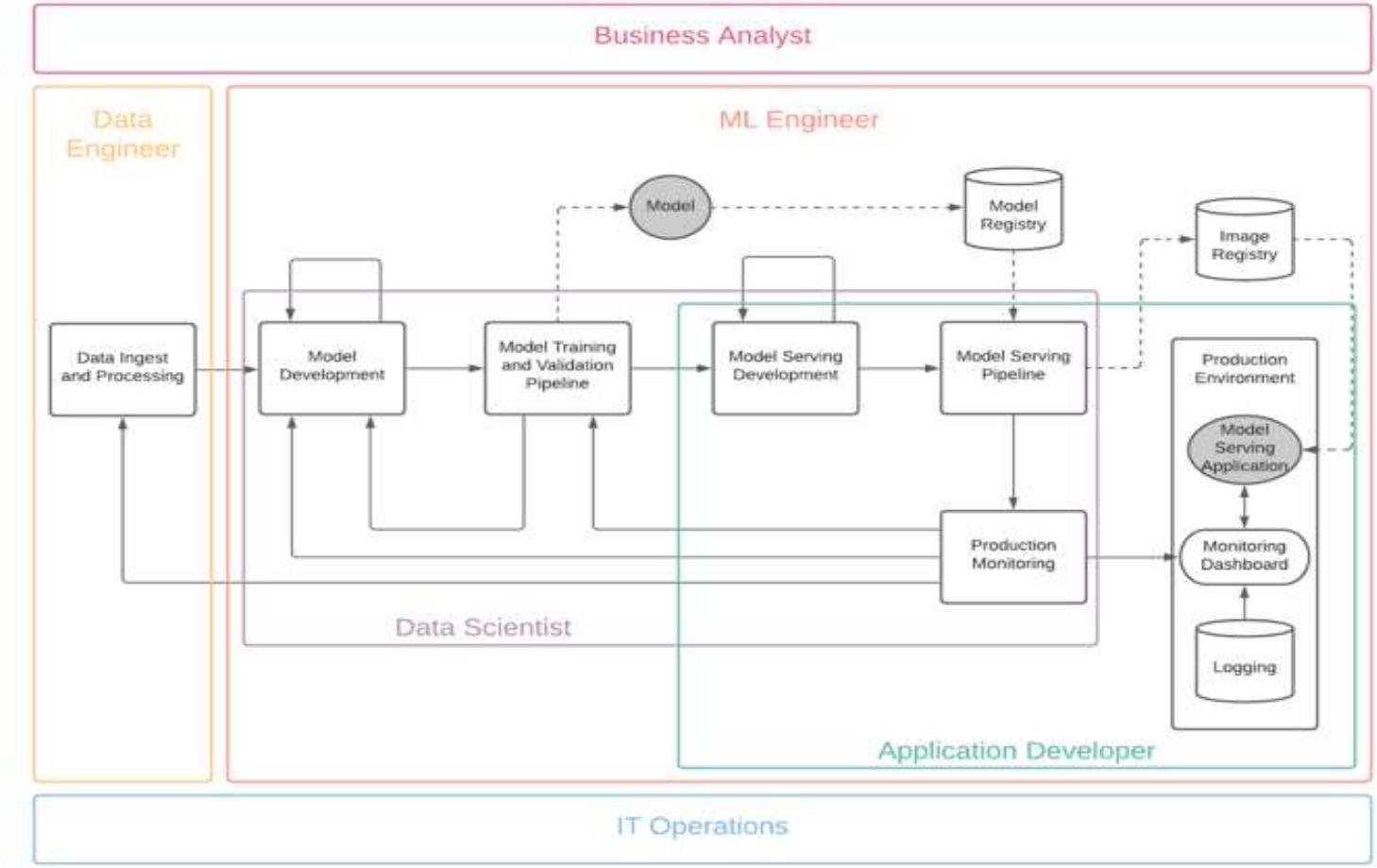
A/B Testing

- انتخاب سرور ها به صورت تصادفی
- تاخیر در اجرا



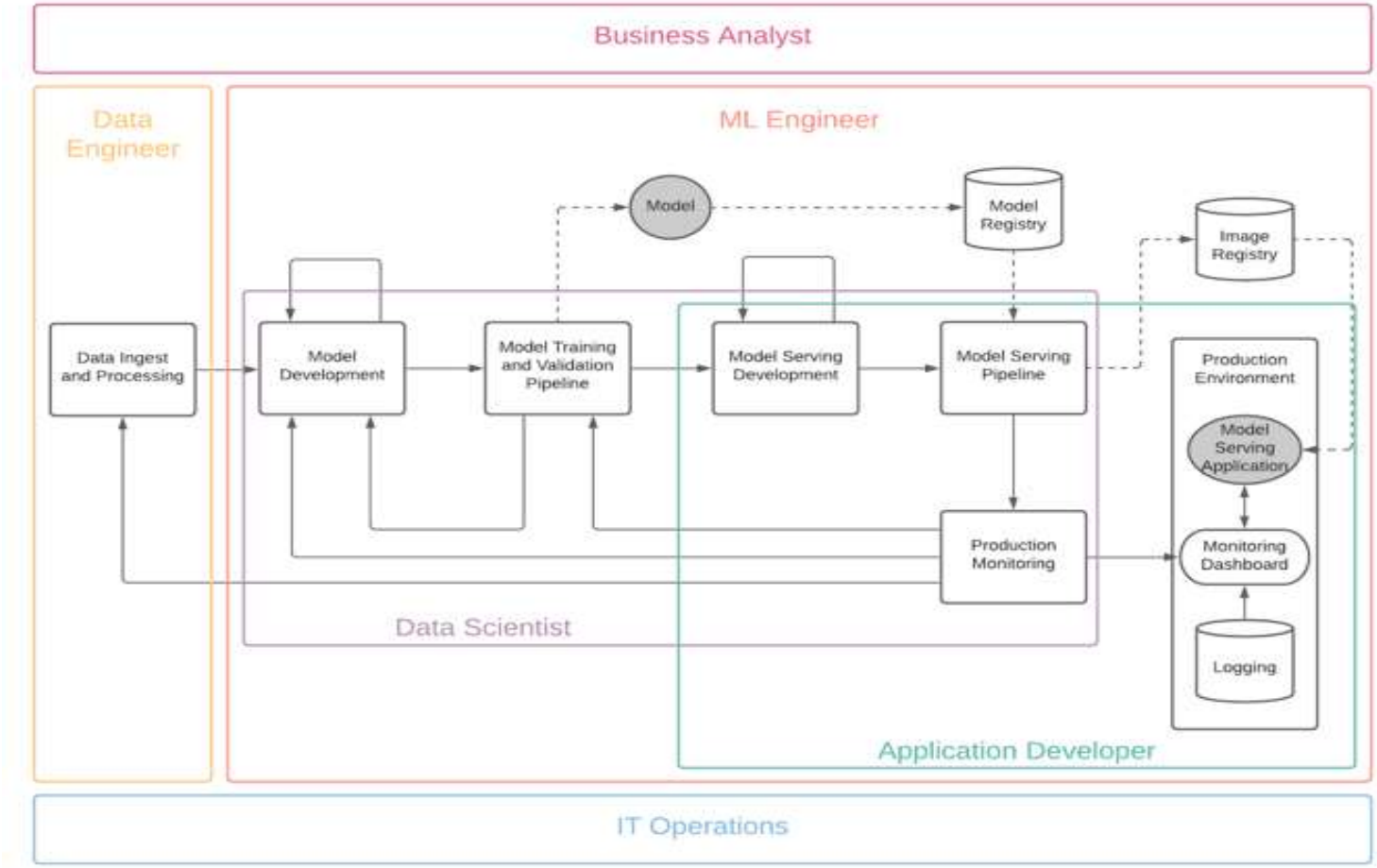
Canary Testing

ساختار سازمانی مبتنی بر ML/Ops



{عنوان وبینار}
{webinar title}

ساختار سازمانی مبتنی بر ML/Ops



{عنوان وبینار}
{webinar title}

Business Analyst

- تعریف یوزکیس ها و نیازمندی های بیزینسی
- پیگیری ROI
- اطمینان از پوشش نیازمندی بعد از پیاده سازی

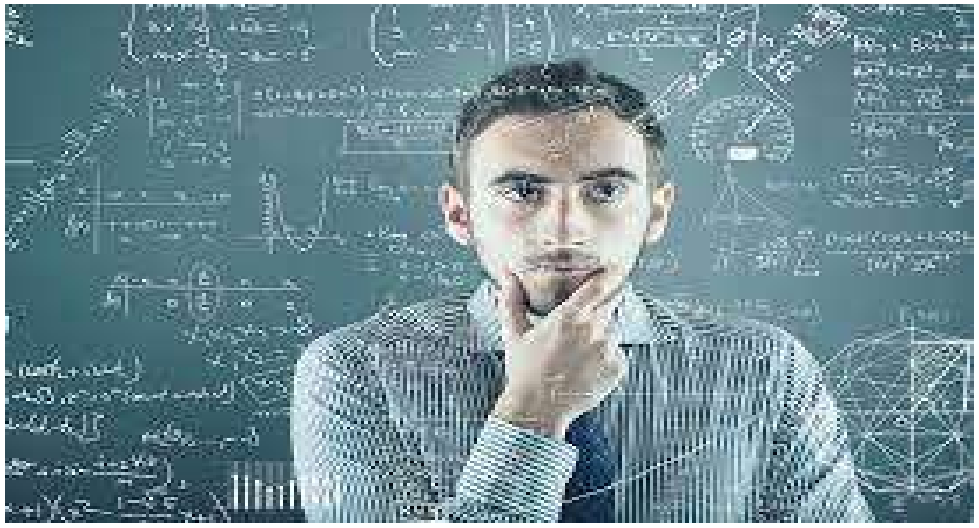
{عنوان وبینار}
{webinar title}

Data Engineer

- دریافت داده خام و اماده سازی (clean,prepare,organize)
- مجوز دریافت و در صورت نیاز تغییر داده های حساس و قابل فهم
- ارائه داده در فرمت استاندارد سازمان
- ایجاد ساختار های توسعه و معماری و ابزاری برای داده ها

Data Scientist

- فهم معنایی داده ها
- فرایند تکرار پذیر برای انتخاب الگوریتم مناسب
- ایجاد مدل نهایی یادگیری ماشین

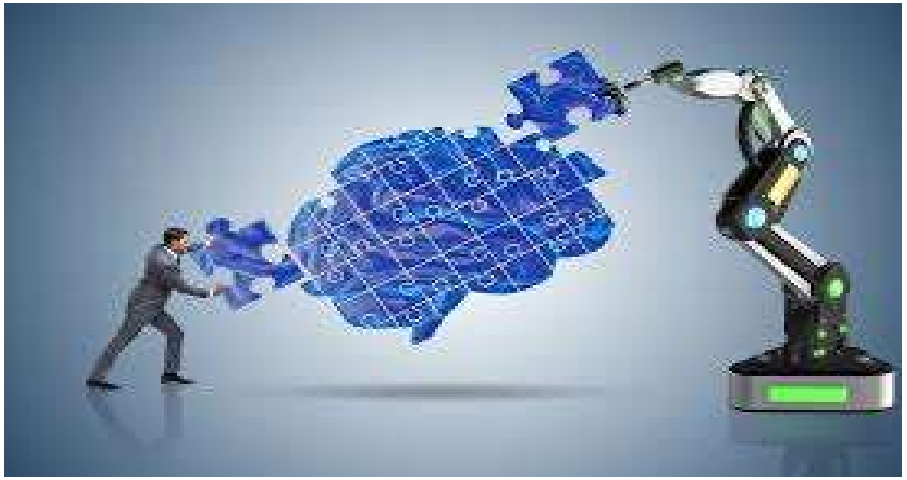


Application Developer

- ایجاد میکروسرویس برای مدل یادگیری
- تمرکز روی تحویل نهایی محصول

ML Engineer

- نظارت و کنترل و ساختن خط تولید مبتنی بر ML



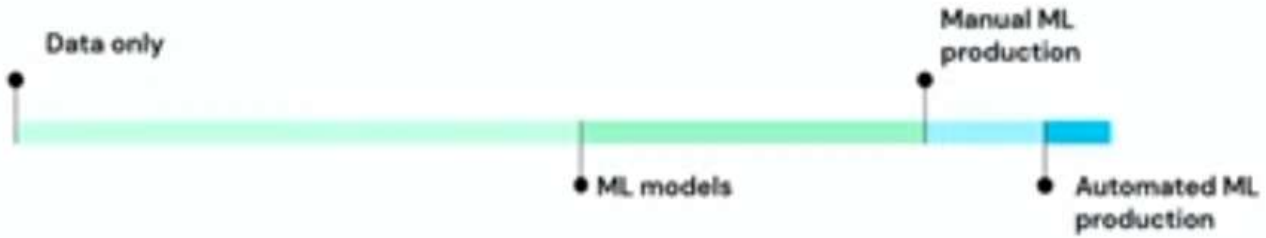
IT Operations

- نصب سیستم ها
- مانیتورینگ سیستم ها
- نگهداری از سیستم ها
- تصمیم برای توپولوژی استقرار

Data+AI

ML maturity levels

Almost every company wants to be a Data+AI company, but the majority are new to machine learning (ML).



ابزارهای مطرح این حوزه

MlOps tools	strengths	weakness
Kubeflow	Seamless integration with Kubernetes	Requires Kubernetes expertise
MLflow	Easy setup and usage	May lack some advanced features
TensorBoard	Powerful visualization for TensorFlow models	Focused on visualization; may lack other MLOps features
DVC	Strong data versioning and pipeline orchestration	Limited MLOps-specific features
Apache Airflow	Excellent for complex workflow orchestration	Additional components needed for full MLOps support
TFX	Scalable model training and serving with TensorFlow	Tightly integrated with TensorFlow
Metaflow	Easy to use and collaborate on data science projects	May lack extensive enterprise-grade features

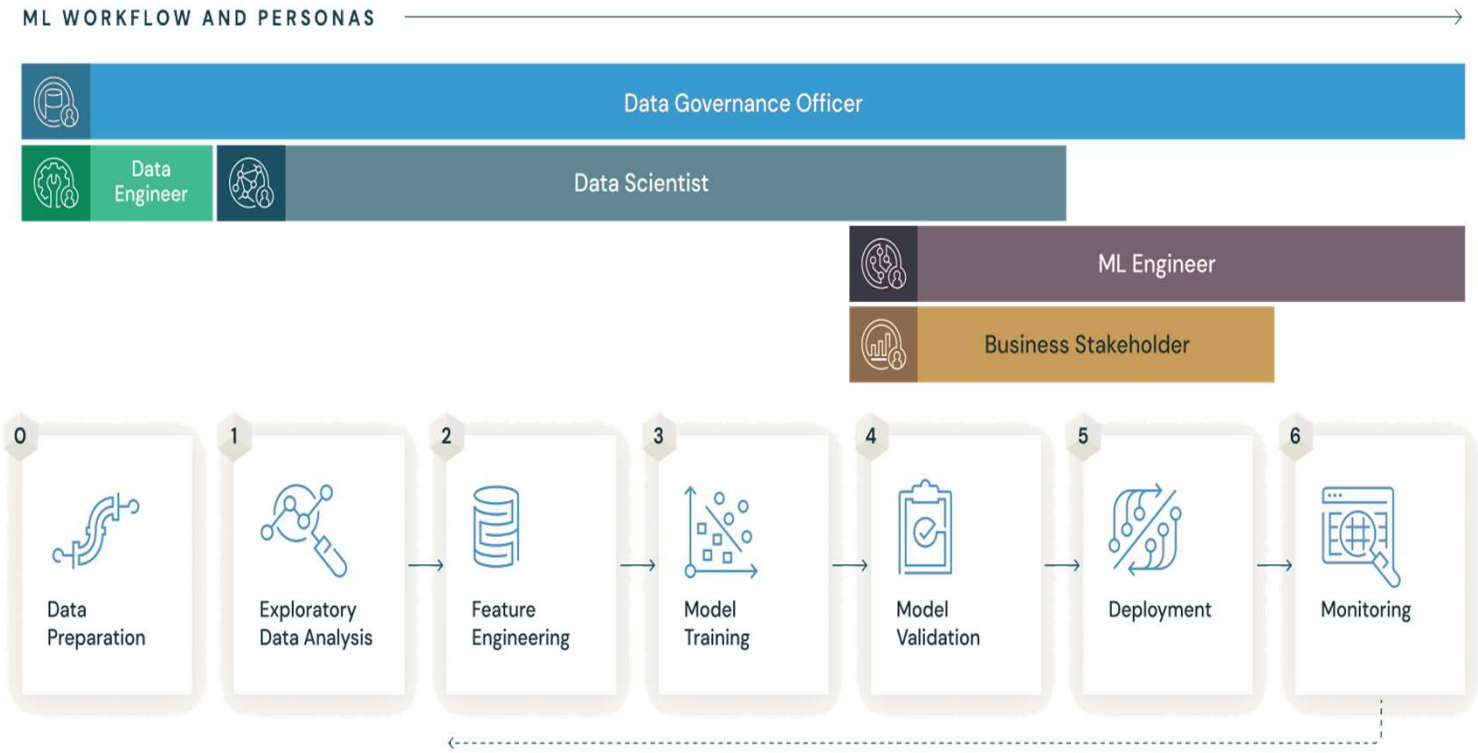
ابزارهای مطرح این حوزه

MlOps tools	strengths	weakness
Kubeflow	Seamless integration with Kubernetes	Requires Kubernetes expertise
MLflow	Easy setup and usage	May lack some advanced features
TensorBoard	Powerful visualization for TensorFlow models	Focused on visualization; may lack other MLOps features
DVC	Strong data versioning and pipeline orchestration	Limited MLOps-specific features
Apache Airflow	Excellent for complex workflow orchestration	Additional components needed for full MLOps support
TFX	Scalable model training and serving with TensorFlow	Tightly integrated with TensorFlow
Metaflow	Easy to use and collaborate on data science projects	May lack extensive enterprise-grade features

MLflow

- پروژه متن باز
- نگهداری تجارب(experiments) الگوریتم های مختلف روز داده ها به منظور مقایسه ان ها
- نگهداری کد های ML به منظوربسته بندی و ارسال نسخه و همینطور استفاده مجدد
- مدیریت و دیپلوی مدل های مختلف ML
- مدیریت چرخه تولید و عملیات در یادگیری ماشین

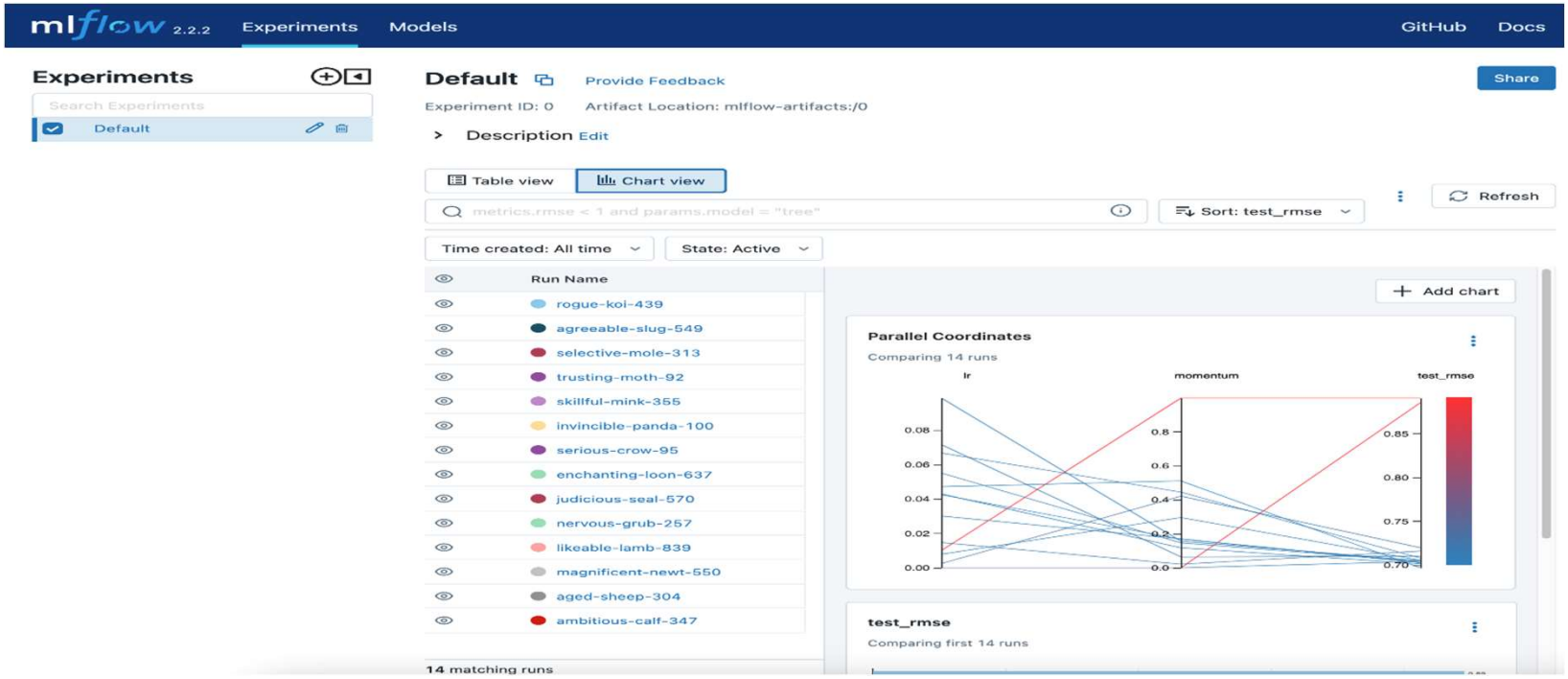
معماری سازمانی در MLflow



{عنوان وبینار}
{webinar title}

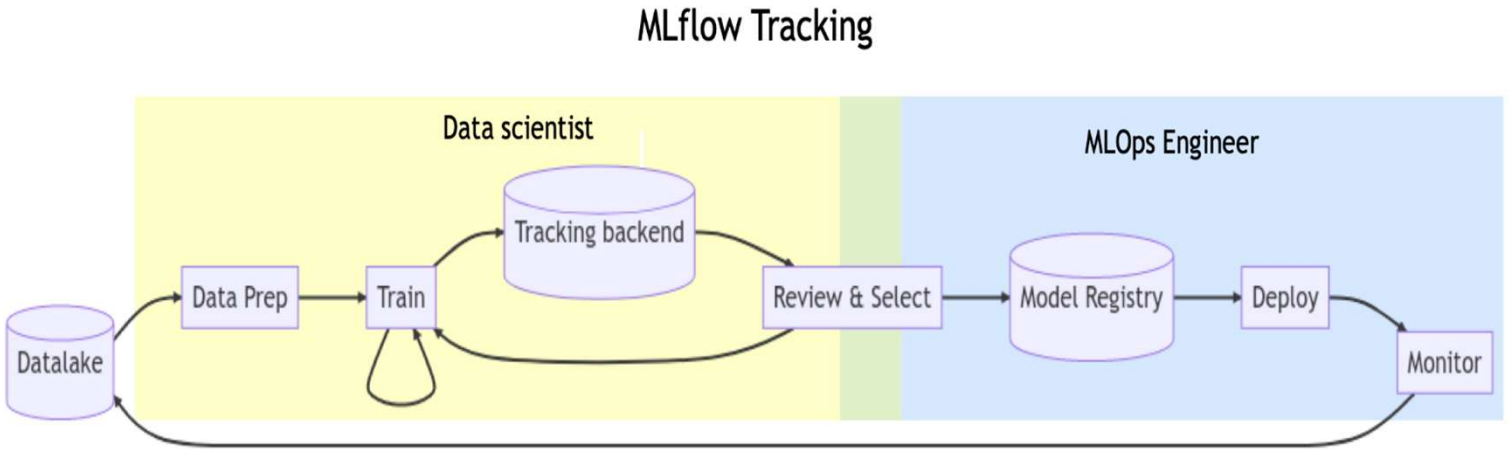
MIFlow Tracking

- لاگ و ثبت پارمتر ها
- نگهداری ورژن کد ها
- ثبت متریک ها
- داشبرد مدیریتی برای مشاهده وقایع



{عنوان وبینار}
{webinar title}

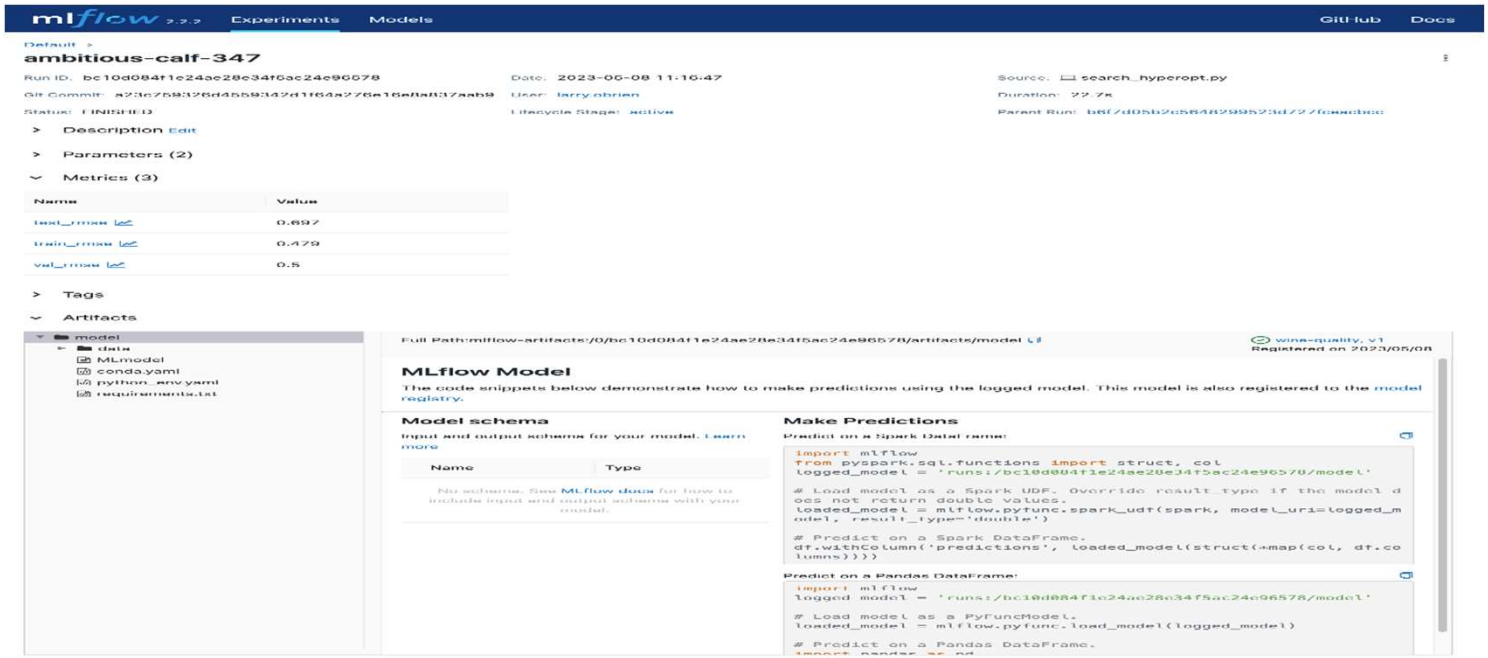
MIflow tracking



{عنوان وبینار}
{webinar title}

MLFlow Models

- مجموعه ای از ابزار ها برای یادگیری و استقرار مدل
- استفاده از ابزار هایی مثل داکر برای استقرار مدل
- مقایسه پارامتر های مدل برای انتخاب مدل بهینه



به گروه تخصصی معماری سازمانی انجمن انفورماتیک ایران بپیوندید...

www.isi-ea.ir