4.11 ML ops and monitoring plan

و مانیتورینگ MLOps برنامه - ProDecks
نسخه: 1.0
تاریخ: 1403/11/17
وضعیت: MVP - پیاده‌سازی شده

MLOps فصل ۱: چشم‌انداز

۱.۱. مأموریت

"ایجاد و نگهداری یک سیستم MLOps پایدار، مقیاس‌پذیر و قابل اطمینان که تضمین کند مدل‌های یادگیری ماشین ProDecks همیشه بهینه، به‌روز و ایمن عمل می‌کنند."

۱.۲. اصول MLOps

• اتوماسیون: خودکارسازی تا حد امکان
• reproduciblity: قابلیت تکرارپذیری کامل
• مانیتورینگ: نظارت مستمر
• collaboration: همکاری بین تیم‌ها

۱.۳. محدوده

• کلیه مدل‌های تولیدی
• خط لوله داده
• زیرساخت آموزش و استقرار
• سیستم‌های مانیتورینگ

MLOps فصل ۲: چرخه حیات

Data Management → Model Development → Model Deployment → Monitoring & Maintenance → Model Retirement

۲.۱. مرحله ۱: مدیریت داده

• Data Versioning: با DVC (Data Version Control)
• Data Quality Monitoring: با Great Expectations
• Feature Store: با Feast
• Pipeline Orchestration: با Apache Airflow

مرحله ۲: توسعه مدل ۲.۲.

• Experiment Tracking: با MLflow
• Model Registry: با MLflow Model Registry
• Hyperparameter Tuning: با Optuna
• Code Versioning: با Git

مرحله ۳: استقرار مدل ۲.۳.

• Containerization: با Docker
• Orchestration: با Kubernetes
• Serving: با TensorFlow Serving / Seldon Core
• A/B Testing: با Istio

مرحله ۴: مانیتورینگ و نگهداری ۲.۴.

• Performance Monitoring: با Prometheus
• Drift Detection: با Evidently AI
• Logging: با ELK Stack
• Alerting: با PagerDuty

مرحله ۵: بازنشستگی مدل ۲.۵.

• Archival: ذخیره مدل‌های قدیمی
• Documentation: مستندات بازنشستگی
• Data Retention: نگهداری داده‌های مرتبط

فصل ۳: خط لوله آموزش مدل

Pipeline معماری ۳.۱.

[Trigger] → [Data Fetch] → [Preprocessing] → [Training] →
[Evaluation] → [Model Registry] → [Deployment]

Triggerها ۳.۲.

۱. زمان‌بندی شده:

• هفتگی برای مدل‌های اصلی
• ماهانه برای مدل‌های کم‌اهمیت

۲. event-driven:

• Drift detection
• Performance degradation
• Data threshold reached

۳. دستی:

• جدید feature انتشار
• business requirements تغییر

۳.۳. مراحل Pipeline

مرحله ۱: دریافت داده
• منبع: پایگاه داده production
• فیلتر: داده‌های ۶ ماه گذشته
• حجم: حداکثر ۱M نمونه

مرحله ۲: پیش‌پردازش
• پاکسازی داده
• feature engineering
• train/test split

مرحله ۳: آموزش
• اجرای experiment جدید
• hyperparameter tuning
• cross-validation

مرحله ۴: ارزیابی
• تست بر روی داده hold-out
• مقایسه با baseline
• بررسی metrics کسب‌وکار

مرحله ۵: ثبت
• ذخیره در Model Registry
• versioning
• metadata storage

فصل ۴: استراتژی استقرار

۴.۱. الگوهای استقرار

۱. Canary Deployment:
• ترافیک به نسخه جدید ۱۰٪
• مانیتورینگ دقیق
• در صورت مشکل rollback

۲. Blue-Green Deployment:
• دو محیط موازی
• ترافیک switch
• zero-downtime

۳. Shadow Deployment:
• اجرای موازی

• مقایسه خروجی‌ها
• بدون تأثیر بر کاربران

۴.۲. محیط‌های استقرار

• Development: برای آزمایش‌های داخلی
• Staging: شبیه‌سازی production
• Production: کاربران واقعی
• Canary: درصد کمی از کاربران

۴.۳. رویکرد استقرار

• مدل‌های غیرحساس: استقرار خودکار
• مدل‌های حساس: تأیید دستی
• مدل‌های بحرانی: تأیید کمیته

فصل ۵: مانیتورینگ مدل‌ها

۵.۱. مانیتورینگ عملکرد (Performance Monitoring)

• metrics فنی:
- Latency: P50, P95, P99
- Throughput: درخواست/ثانیه
- Error Rate: درصد خطاها
- Resource Usage: CPU, Memory, GPU

• metrics مدل:
- Accuracy, Precision, Recall
- AUC-ROC, F1-Score
- Business metrics (تأثیر بر retention, conversion)

۵.۲. مانیتورینگ داده (Data Monitoring)

۱. Data Drift:
• تغییر توزیع داده ورودی
• روش: Statistical Distance (Wasserstein, KL-divergence)
• threshold: ۰.۱

۲. Concept Drift:
• تغییر رابطه ویژگی‌ها و هدف
• روش: Model Performance Degradation
• threshold: ۱۰٪ کاهش performance

۳. Data Quality:
• missing values
• outliers

• schema changes

۵.۳. مانیتورینگ bias

• زیرگروه‌ها: جنسیت، تجربه، منطقه
• معیارها:
- Demographic Parity
- Equal Opportunity
- Predictive Parity
• threshold: Δ < ۵٪

۵.۴. مانیتورینگ امنیت

• Adversarial Attacks Detection
• Model Extraction Attempts
• Unauthorized Access

فصل ۶: سیستم هشدار (Alerting System)

۶.۱. سطوح هشدار

• Critical (P0):
- Model completely down
- Performance degradation > ۵۰٪
- Security breach
- Response: Immediate (within ۱۵ minutes)

• High (P1):
- Performance degradation > ۲۰٪
- Data drift > threshold
- Bias increase > ۱۰٪
- Response: Within ۱ hour

• Medium (P2):
- Performance degradation > ۱۰٪
- Resource usage > ۸۰٪
- Warning signs
- Response: Within ۴ hours

• Low (P3):
- Informational alerts
- Trends monitoring
- Response: Within ۱ day

۶.۲. کانال‌های هشدار

• P0: Phone Call + SMS + PagerDuty
• P1: SMS + Email + Slack
• P2: Email + Slack
• P3: Slack only

۶.۳. Escalation Policy

• Level ۱: ML Engineer on-call
• Level ۲: Senior ML Engineer
• Level ۳: Head of AI
• Level ۴: CTO

فصل ۷: بازآموزی و به‌روزرسانی

استراتژی بازآموزی ۷.۱.

• زمان‌بندی شده:
- مدل‌های اصلی: هفتگی
- مدل‌های ثانویه: ماهانه

• event-driven:
- Data drift detected
- Performance degradation
- New data threshold reached

• دستی:
- انتشار feature جدید
- تغییر business logic

رویکرد بازآموزی ۷.۲.

۱. Full Retraining:
• آموزش کامل از ابتدا
• زمان‌بر اما دقیق
• برای تغییرات اساسی

۲. Incremental Learning:
• آموزش با داده‌های جدید
• سریع‌تر
• برای تغییرات تدریجی

۳. Transfer Learning:
• fine-tuning مدل موجود
• برای adaptation به دامنه جدید

تست قبل از استقرار ۷.۳.

• تست کد: Unit Tests
• تست pipeline: Integration Tests

• Performance Tests: تست metrics

• A/B Tests: تست impact کسب‌وکار

فصل ۸: Rollback و بازیابی

۸.۱. استراتژی Rollback

• Automatic Rollback:
- Performance degradation > ۳۰٪
- Error rate > ۱۰٪
- Latency increase > ۲x

• Manual Rollback:
- Business logic issues
- Customer complaints
- Regulatory concerns

۸.۲. نسخه‌های پشتیبان

• Model Registry: تمام نسخه‌های مدل

• Data Snapshots: داده‌های آموزشی

• Configuration: تنظیمات pipeline

۸.۳. زمان بازیابی

• RTO (Recovery Time Objective): ۱ ساعت

• RPO (Recovery Point Objective): ۱ روز

• Rollback Procedure: مستند شده و tested

فصل ۹: ابزارها و فناوری‌ها

۹.۱. Stack فنی

• Version Control: Git, DVC
• Experiment Tracking: MLflow
• Orchestration: Apache Airflow
• Serving: TensorFlow Serving, FastAPI
• Monitoring: Prometheus, Grafana, Evidently AI
• Infrastructure: Kubernetes, Docker
• Cloud: AWS (S3, SageMaker, EKS)

۹.۲. انتخاب ابزارها

معیارهای انتخاب:

• Open Source بودن

• جامعه فعال

• مستندات کامل

• قابلیت integration

• مقیاس‌پذیری

۹.۳. هزینه ابزارها

• ابزارهای Open Source: رایگان

• خدمات cloud: ۲ میلیون تومان ماهانه

• توسعه و نگهداری: ۳ نفر full-time

فصل ۱۰: مسئولیت‌ها و نقش‌ها

۱۰.۱. تیم MLOps

• MLOps Engineer (2 نفر):

- مسئول pipeline‌ها و زیرساخت

- مانیتورینگ و هشدار

- استقرار و rollback

• Data Engineer (1 نفر):

- مدیریت داده‌ها

- feature store

- data quality

• ML Engineer (2 نفر):

- توسعه مدل‌ها

- آزمایش‌ها

- ارزیابی

۱۰.۲. مسئولیت‌های کلیدی

• MLOps Engineer: مدل‌های تولیدی

• Data Engineer: داده‌های آموزشی

• ML Engineer: مدل performance

• امنیت: کل تیم

۱۰.۳. فرآیندهای تصمیم‌گیری

• تغییرات کوچک: تأیید تیم

• تغییرات متوسط: تأیید سرپرست

• تغییرات بزرگ: تأیید کمیته AI

فصل ۱۱: مستندات و گزارش‌گیری

۱۱.۱. مستندات الزامی

• Model Cards: برای هر مدل
• Pipeline Documentation: pipeline برای هر
• Deployment Guides: راهنماهای استقرار
• Runbooks: incident response برای

۱۱.۲. گزارش‌های منظم

• روزانه: سلامت سیستم
• هفتگی: performance مدل‌ها
• ماهانه: trends و بهبودها
• فصلی: review کامل

۱۱.۳. dashboards

• Real-time Dashboard: وضعیت لحظه‌ای
• Performance Dashboard: metrics مدل‌ها
• Business Impact Dashboard: تأثیر بر کسب‌وکار
• Cost Dashboard: هزینه‌های MLOps

فصل ۱۲: برنامه بهبود MLOps

۱۲.۱. فاز ۱ (ماه‌های ۱-۳): پایه‌ها

• راه‌اندازی pipeline‌های پایه
• مانیتورینگ اولیه
• مستندات اولیه

۱۲.۲. فاز ۲ (ماه‌های ۴-۶): پیشرفته‌سازی

• automated retraining
• advanced monitoring
• A/B testing framework

۱۲.۳. فاز ۳ (ماه‌های ۷-۹): optimization

• cost optimization
• performance optimization
• automation کامل

۱۲.۴. فاز ۴ (ماه‌های ۱۰-۱۲): enterprise

• multi-model management
• federated learning
• advanced security

فصل ۱۳: معیارهای موفقیت

معیارهای فنی ۱۳.۱.

• Uptime: ۹۹.۹٪
• Pipeline Success Rate: ۹۵٪
• Model Retraining Frequency: مطابق برنامه
• Incident Response Time: مطابق SLA

معیارهای کسب‌وکار ۱۳.۲.

• Model Impact: بهبود metrics کسب‌وکار
• Cost Efficiency: کاهش هزینه‌های عملیاتی
• Team Productivity: افزایش سرعت توسعه

معیارهای کیفیت ۱۳.۳.

• Documentation Completeness: ۱۰۰٪
• Test Coverage: ۸۰٪
• Incident Frequency: کاهش ماهانه

فصل ۱۴: ریسک‌ها و mitigations

ریسک‌های فنی ۱۴.۱.

• ریسک: Pipeline failures
• Mitigation: Monitoring, alerting, automatic recovery

• ریسک: Model degradation
• Mitigation: Regular retraining, drift detection

• ریسک: Infrastructure issues
• Mitigation: Redundancy, backup, disaster recovery

ریسک‌های عملیاتی ۱۴.۲.

• ریسک: Knowledge silos
• Mitigation: Documentation, cross-training

• ریسک: Resource constraints
• Mitigation: Capacity planning, optimization

• ریسک: Security breaches
• Mitigation: Security protocols, regular audits

ریسک‌های کسب‌وکار ۱۴.۳.

• ریسک: Business impact from model failures
• Mitigation: Rollback procedures, SLAs

• ریسک: Compliance issues
• Mitigation: Regular audits, documentation

• ریسک: Cost overruns
• Mitigation: Budget monitoring, optimization

فصل ۱۵: نتیجه‌گیری

یک چارچوب جامع و پایدار برای مدیریت چرخه حیات کامل MLOps ProDecks سیستم
مدل‌های یادگیری ماشین ارائه می‌دهد. با تمرکز بر اتوماسیون، مانیتورینگ و
مسئولیت‌پذیری، این سیستم تضمین می‌کند که مدل‌ها همیشه بهینه عمل می‌کنند
.و ارزش پایدار برای کسب‌وکار ایجاد می‌کنند

ضمیمه‌ها

ضمیمه ۱: نمودارهای کامل Pipeline
ضمیمه ۲: Configuration Files
ضمیمه ۳: Runbooks برای Incident Response
ضمیمه ۴: گزارش‌های مانیتورینگ نمونه

تهیه‌کنندگان:
تیم MLOps ProDecks
حامد کوهی (سرپرست MLOps)
