GROWNET — MLOps & Monitoring Plan

(Trustworthy, Observable, Rollback-Ready ML)

---

1. Purpose & Scope

هدف این سند:

تضمین اینکه سیستم ML **قابل اعتماد، قابل مانیتور، قابل بازگشت (rollback)** و **قابل نگهداری در رشد** است.

In scope

- Ranking models
- Recommendation
- Fraud / spam detection (در صورت فعال شدن)

Out of scope (فعلاً)

- Real-time deep learning
- AutoMLپیچیده بدون کنترل

تمرکز **ML**: **پایدار، نهflashy**

---

2. ML System Overview

ML Use-Cases in GROWNET

Area	Model Role
Content ranking	Score & order
Reputation score	Trust signals
Recommendation	Similar content

Decision Boundary

- MLتصمیم نهایی نیست
- MLپیشنهاد می‌دهد، rule-basedکنترل می‌کند

کاهش ریسک رفتار غیرقابل توضیح

---

3. Model Lifecycle (End-to-End)

Data → Feature → Train → Validate → Deploy → Monitor → Retrain

Ownership

Stage	Owner
Data quality	Data Eng
Training	ML Eng
Deployment	Platform
Monitoring	ML + SRE

---

4. Model Versioning Strategy

Versioning Layers

Layer	Version
Dataset	data_vX
Features	feature_vX
Model	model_vX
Pipeline	pipeline_vX

Rule

No model runs in prod بدون:

- نسخه مشخص
- hash artifact
- training metadata

---

5. Pre-Deployment Testing

Mandatory Checks

Test	Purpose
Offline metrics	Accuracy / NDCG
Bias check	Distribution shift
Backtest	Compare to baseline
Canary run	Limited exposure

ML بدون تست feature = خطرناک

---

6. Deployment Strategy

Deployment Types

- Shadow mode
- Canary release
- Gradual rollout

Rollback Rule

Trigger	Action
KPI drop	Auto rollback
Drift alert	Freeze model
Incident	Switch to rules

rollback < 5 min

---

7. Monitoring Dimensions

Data Monitoring

Metric	Why
Missing values	Pipeline break
Distribution shift	Drift
Feature ranges	Input sanity

Model Performance

Metric	Target
Ranking CTR	Stable ±5%
Precision@k	≥ baseline
False positives	bounded

Concept Drift Detection

- KS test
- Population stability index
- Rolling window comparison

فوریdrift ≠ retrain
drift + KPI drop = retrain

8. Alerting & Incident Response

Alert Levels

Level	Trigger
Warning	Feature drift
Critical	KPI drop
Emergency	Data corruption

Response Playbook

1. Freeze deployment
2. Switch fallback
3. Root cause
4. Fix + postmortem

9. Retraining Strategy

Retraining Triggers

- Scheduled (monthly)
- Event-based (drift)
- Business-driven (new segment)

Retraining Controls

- Same pipeline
- Same validation
- Human approval gate

No silent retraining

10. Tooling Stack

Area	Tool
Experiment tracking	MLflow

Area	Tool
Model registry	MLflow
Monitoring	Evidently / custom
CI/CD ML	GitHub Actions
Data quality	Great Expectations

11. Security & Access Control

- Model artifacts encrypted
- Limited prod access
- Training data masked
- Audit logs enabled

ML = data risk surface

12. Documentation & Knowledge Transfer

Required Docs

- Model cards
- Feature definitions
- Known failure modes

Bus Factor Mitigation

- No single-owner model
- Shared reviews
- Recorded walkthroughs

13. KPIs for ML Health

KPI	Target
Time to detect drift	<24h
Time to rollback	<5 min

KPI	Target
Model incidents	<1/q
Unexplained drops	0

---

14. Known Risks & Mitigations

Risk	Mitigation
Silent drift	Automated alerts
Overfitting	Holdout sets
Bias	Periodic audits
Tool lock-in	Portable pipelines

---

15. Executive Takeaway (Investor Lens)

ما مدل را deploy نمی‌کنیم و رها نمی‌کنیم؛
ما **کنترل، مشاهده و بازگشت سریع** داریم.

یعنی:

- MLدارایی است
- ریسک فنی کنترل‌شده
- هزینه آینده قابل پیش‌بینی