

ML- «HS- + → »  
 , - , ( « », « - -1»).

, :

;

- ;

.

1. , hs\_code (6- ), country\_origin, country\_destination, consignee\_name, description<sup>1</sup> risk\_score [0,1], risk\_level {low, medium, high}, violations ( / )  
 Recall 0.95 High ( ) Latency 200 95-  
 20 req/s - - ,  
 . 1) inline- , 2) , 3)  
 REST-API ERP. <sup>1</sup> description = « » ( ENG/RUS).
2. 2.1 OFAC SDN/NS- CSV, EU FSF (XML), UN SC (XML) CSV/XML/JSON ./ . HS- + Harmonized-System CSV . UN Comtrade API ( ) CSV / ( NDA) Parquet

2.2 pgsql data/ raw/ sanctions\_ofac.csv sanctions\_eu.xml sanctions\_un.xml hs\_full.csv comtrade\_2024.csv processed/ train.parquet # sanctions\_flat.parquet 2.3 normal-ize\_sanctions.py → : entity, country\_iso, program, list\_id, last\_update.

HS- Euristics v0 — + «dual-use», «explosives» . . v1 — TF-IDF + , v2 — fine-tuned MiniLM (Siamese).  
 → train.parquet (hs\_code, features\_json, label, source).

3. v0 TF-IDF(eng+rus)+LogReg baseline  
 1 ROC-AUC 0.80 v1 MiniLM-S + (CatBoost) multi-input, mix embeddings + one-hot ROC-AUC 0.90 v2 Cross-encoder (E5-large) + LoRA- heavy, - F1-macro 0.92  
 — v1 ( / ).
4. API- (FastAPI) python POST /v1/check Request: { "hs\_code": "860900", "country\_origin": "CN", "country\_destination": "RU", "consignee\_name": "OOO Logistika", "description": "Electric locomotives control module" }  
 Response 200: { "risk\_score": 0.87, "risk\_level": "high", "violations": ["OFAC\_SDN", "EU\_CFSP"], "model\_version": "2025.07.14\_v1" } :  
 - /status ( ) /metrics (Prom- ).

5. API- FastAPI + Uvicorn - 200 MB  
 Docker cron job , .  
 MLflow Artifacts (S3- minio) . CI/CD  
 GitHub Actions → Docker Hub → Fly.io → pytest → docker build  
 → deploy.
6. - Recall@High 0.95 F1-macro 0.90  
 P95 latency 200 Throughput 20 req/s Pop-recall-7d  
 -5 pp avg. / ↓ 50 %  
 Prometheus + Grafana; Slack / Telegram.
7. D + 0 doc v1 D + 3  
 train.parquet v0 80 k D + 5 Baseline v0 (notebook + pickle)  
 ROC-AUC 0.80 D + 8 API /health + /v1/check (mock) Docker image  
 D + 10 v1 (MiniLM+CatBoost), unit- 80 % tag v1 D + 12 - ,  
 . latency < 200 JMeter D + 14 - : , ,  
 KPI PDF
- ( — « 13 – 14» .)
8. : - pytest 80 % , CI .  
 ML: hold-out (2024Q4) — Recall@High 0.95.  
 Prod: Fly.io, 1000 wrk — P95 200 , 0.  
 : README , ml\_spec\_v0.1.md, Mermaid, JMeter.
9. cron- , retrain ↑ HS Recall  
 top-false-negatives Latency > 200 v1 inline- fallback:  
 + Faiss
10. ML- A. Bernikov , Back-end N. Dev FastAPI,  
 CI/CD DevOps P. Ops Docker, Fly.io, PM / E. Manager ,
- « » ( ). , « - » ( )  
 KPI ,