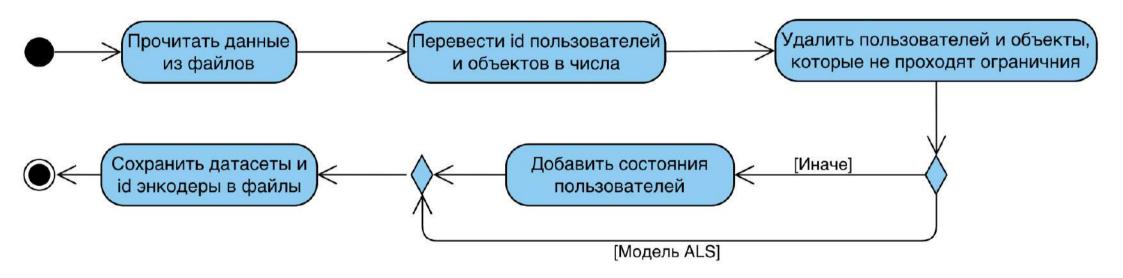
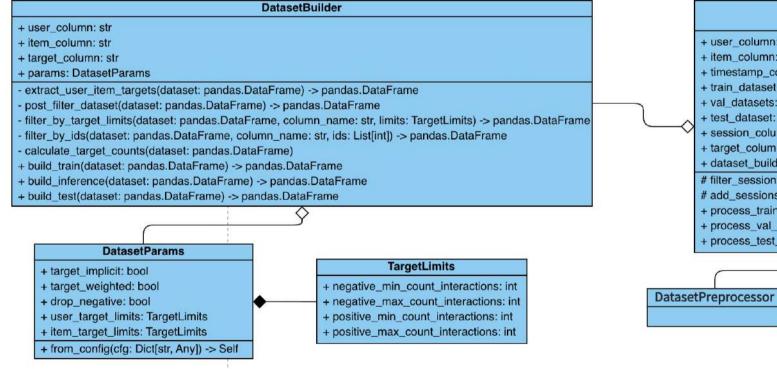


```
FROM docker-hosted.artifactory.tcsbank.ru/recsys-core/default-cpu-pack-python3.11:stable
      LABEL description="Image for train and inference torch models for Tinkoff Journal personal feed"
      USER root
      ENV TZ="Europe/Moscow"
      RUN mkdir -p /app \
          && ln -snf /usr/share/zoneinfo/$TZ /etc/localtime \
          && echo $TZ > /etc/timezone
      WORKDIR /app
      COPY pyproject.toml poetry.lock Makefile /app/
      RUN poetry config virtualenvs.in-project true && make install-models MODELS="torch" && rm -rf /root/.cache
      COPY configs/dssm /app/configs/dssm
      COPY configs/gru /app/configs/gru
      COPY configs/sasrec /app/configs/sasrec
      COPY src /app/src
      COPY models/dssm /app/models/dssm
      COPY models/gru /app/models/gru
      COPY models/sasrec /app/models/sasrec
      COPY models/base /app/models/base
      COPY cli/dssm /app/cli/dssm
      COPY cli/gru /app/cli/gru
      COPY cli/sasrec /app/cli/sasrec
      COPY cli/base /app/cli/base
34
```





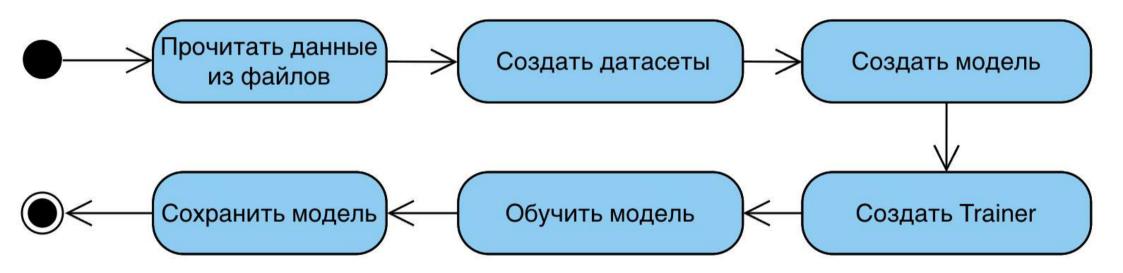
## <<abstract>> BaseDatasetPreprocessor + user column: str + item column: str + timestamp\_column: str + train dataset: pandas.DataFrame + val\_datasets: Dict[str, pandas.DataFrame] + test\_dataset: pandas.DataFrame + session columns: List[str] + target\_column: str + dataset\_builder: DatasetBuilder # filter sessions(dataset: pandas.DataFrame) -> pandas.DataFrame # add\_sessions(dataset: pandas.DataFrame) -> pandas.DataFrame + process train dataset(add sessions: bool) -> None + process\_val\_datasets(add\_sessions: List[bool]) -> None + process\_test\_dataset(add\_sessions: bool) -> None StateDatasetPreprocessor # filter\_without\_history(dataset: pandas.DataFrame)

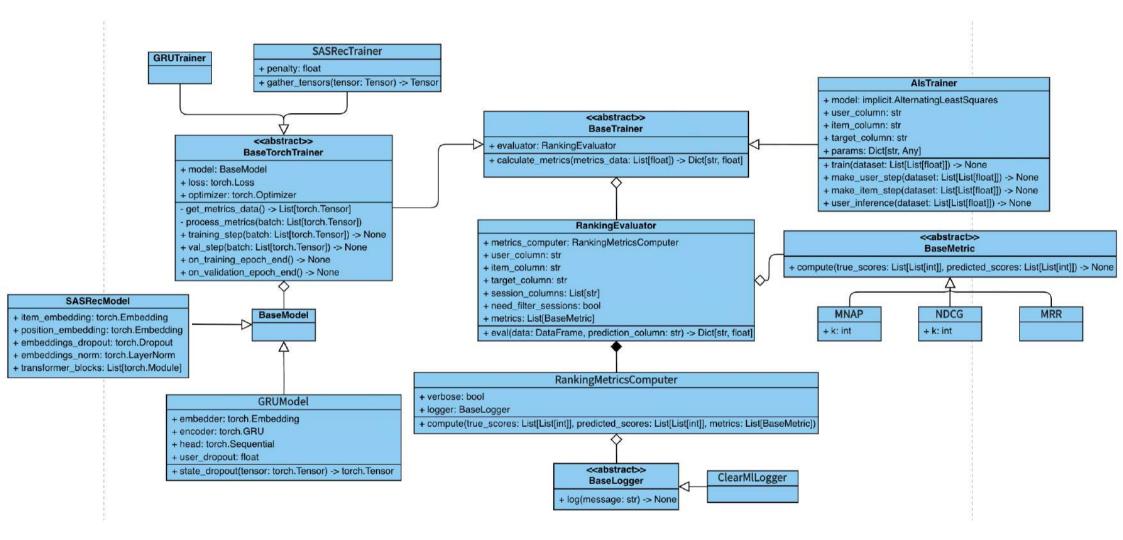
### HistoryBasedStateBuilder

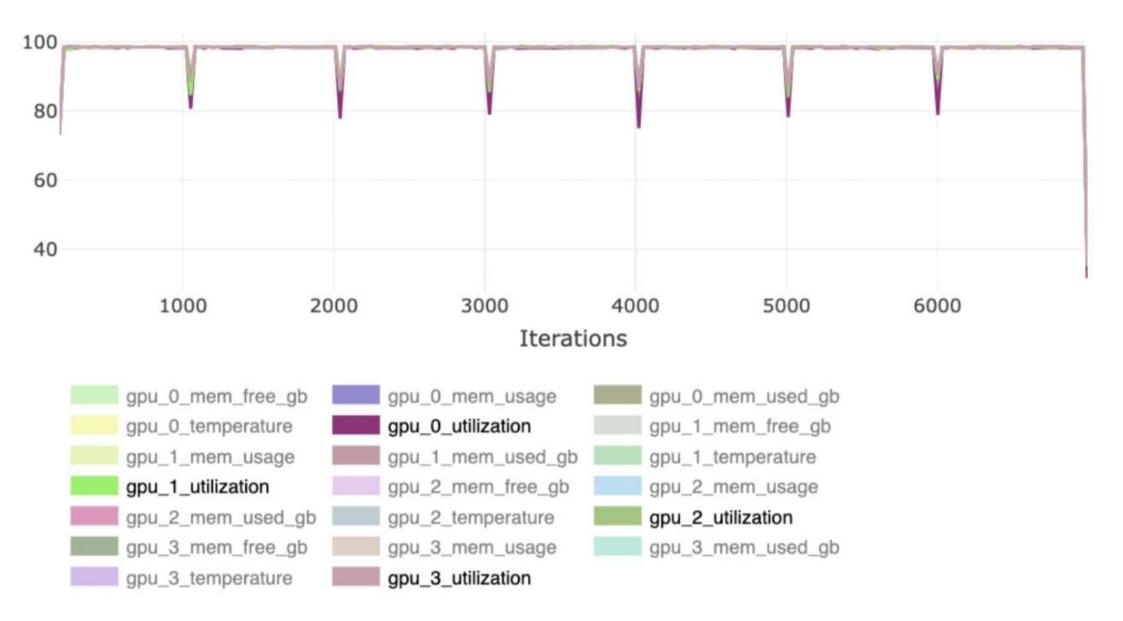
- + mapper: SortedDict[pandas.Timedelta, Dict]
- + user\_column: str
- + item\_column: str
- + target column: str
- + timestamp\_column: str
- + state\_column: str
- + last n: int
- dict\_merge(new\_interactions: Dict[str, List[int]], history: Dict[str, List[int]]) -> Dict[str, List[int]]
- filter\_by\_id(interactions: pandas.DataFrame, column\_name: str, ids: Set[int]) -> pandas.DataFrame
- choose\_date(timestamp: pandas.Timestamp) -> pandas.Timestamp | None
- find state(timestamp: pandas.Timestamp, user: int) -> list[int]
- + fit(interactions: pandas.DataFrame,step\_day: int, users: Set[int] | None = None, items: Set[int] | None = None
- + transform(interactions: pandas.DataFrame) -> pandas.DataFrame

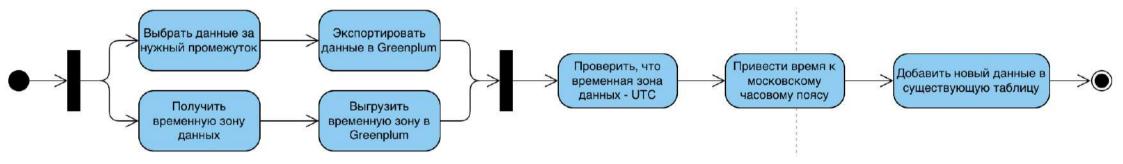
### **IDEncoder**

- + padding: int
- + encoder: sklearn.LabelEncoder
- + fit(data: List[Any]) -> None
- + transform(data: pandas.DataFrame) -> pandas.DataFrame
- + fit\_transform(data: pandas.DataFrame) -> pandas.DataFrame
- + inverse\_transform(data: pandas.DataFrame) -> pandas.DataFrame
- + from\_pretrained(path: str) -> Self
- + save(path: str) -> None









# clickstream\_logs

hit\_dttm: TIMESTAMP NOT NULL

wuid: UUID NOT NULL

social\_id: UUID

event\_name: str NOT NULL

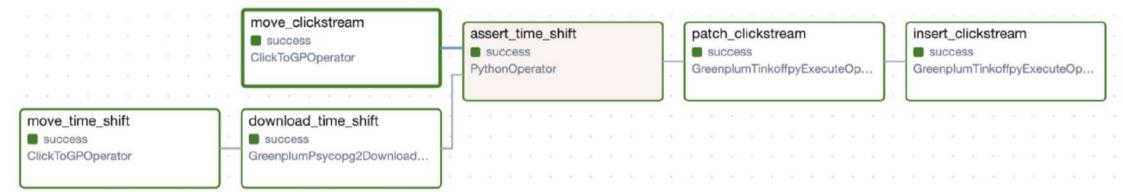
event\_param\_1: str

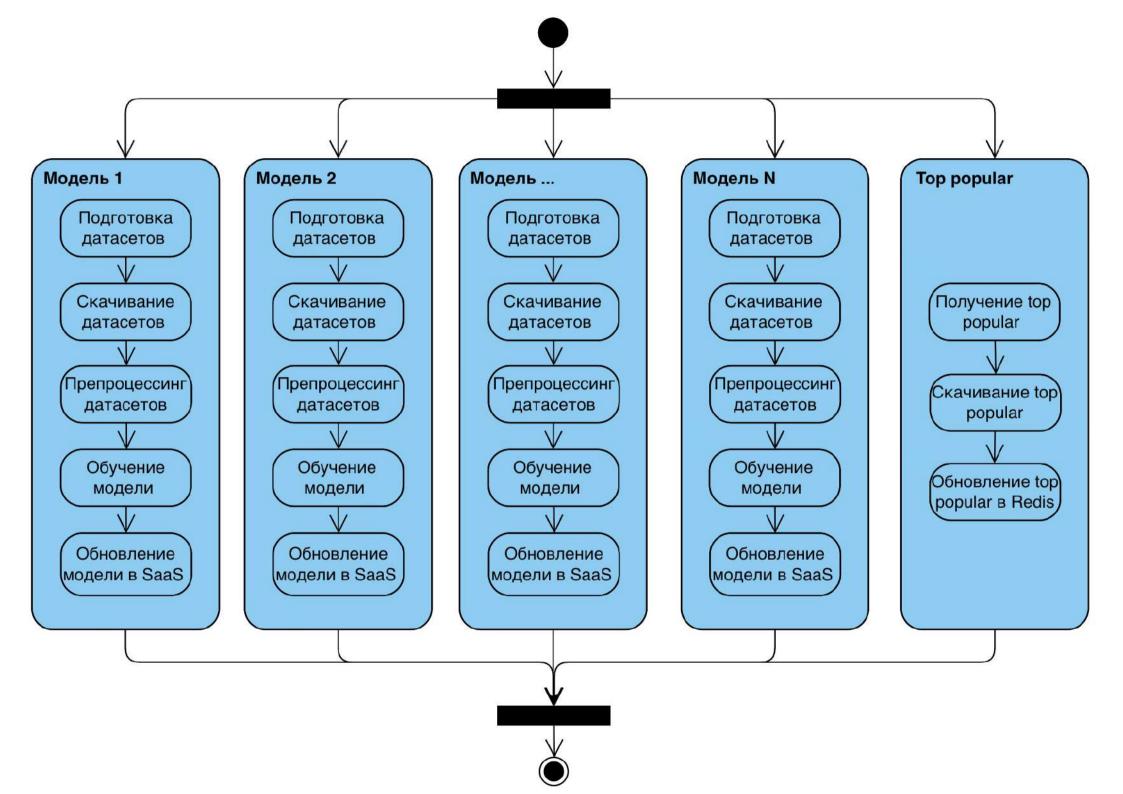
event\_param\_17: str

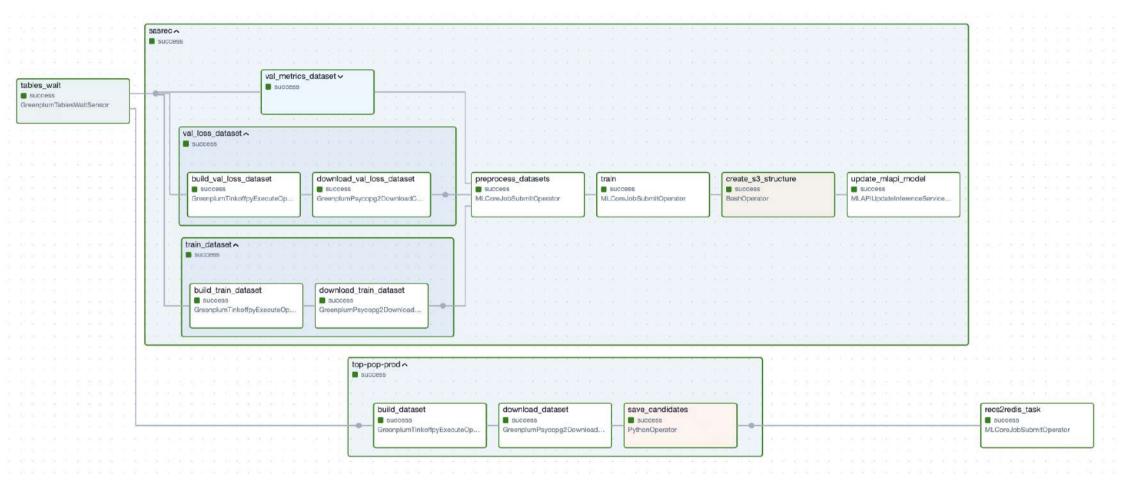
event\_param\_18: str

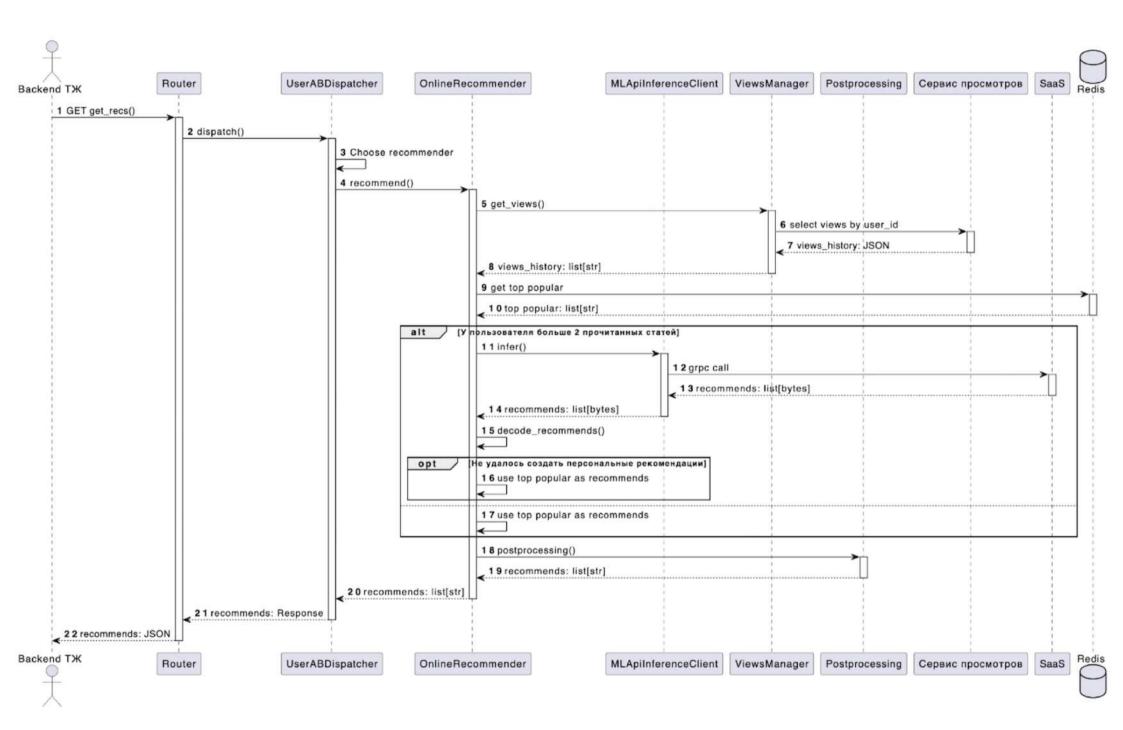
event\_param\_20: str

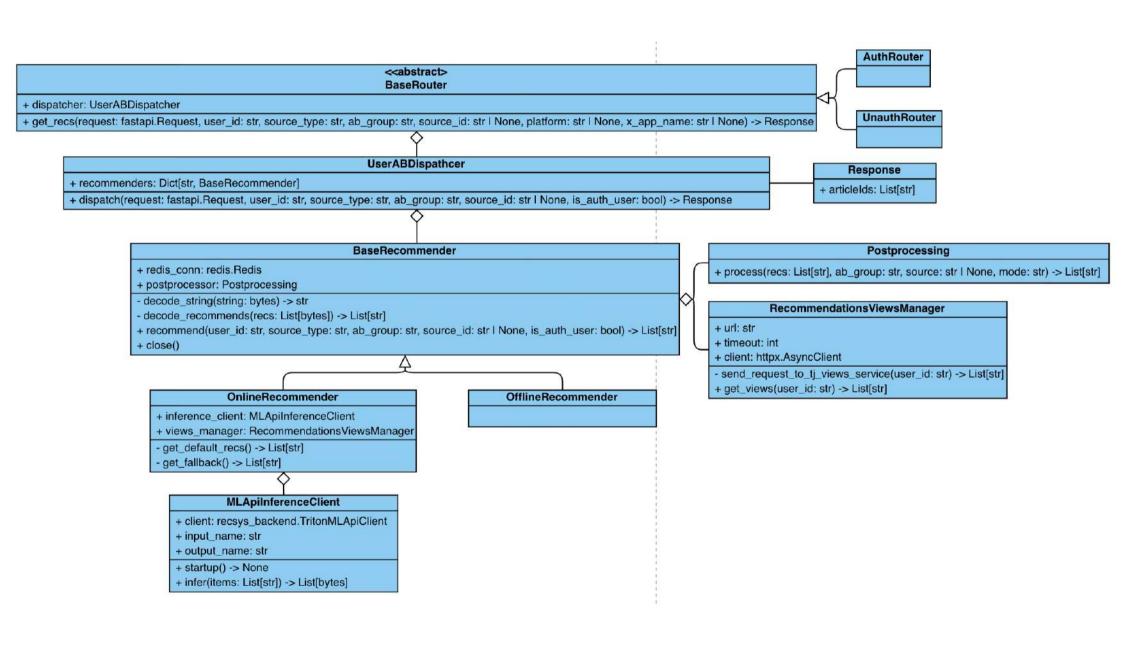
event\_param\_27: str

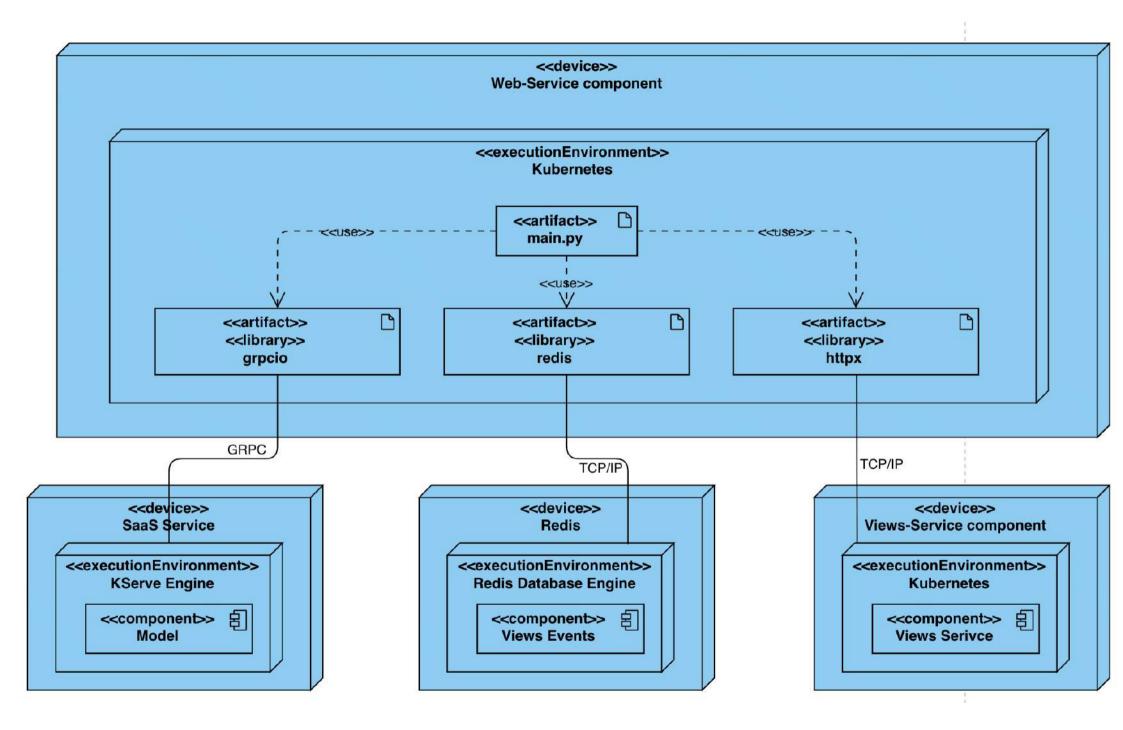


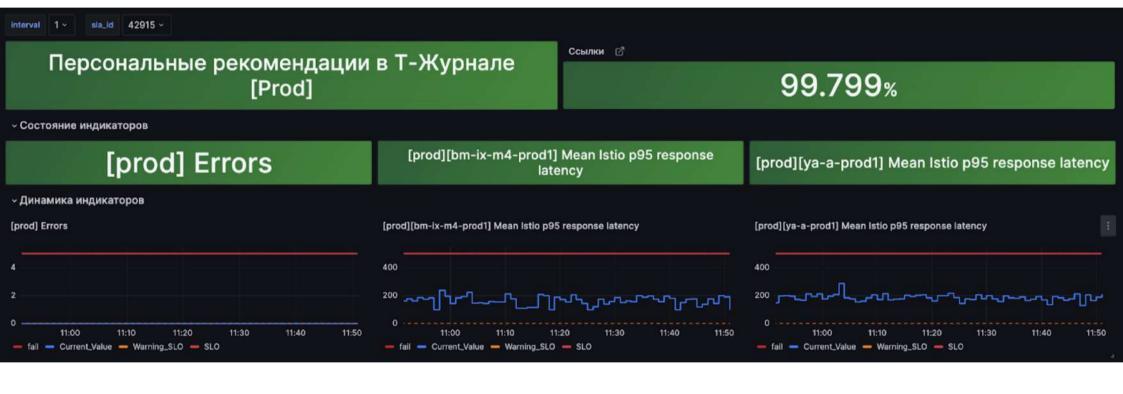






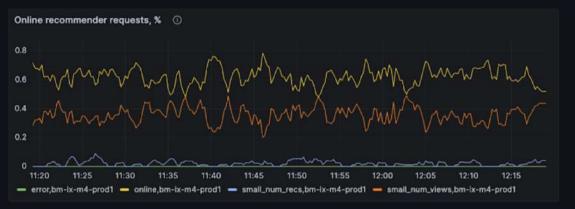




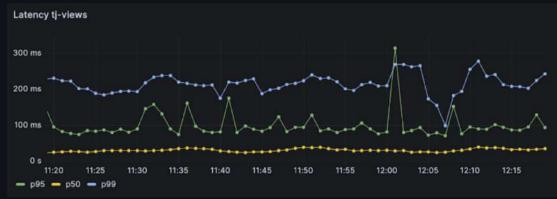


- tech\_name: istio-latency-bm-ix-m4-prod1-p95

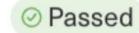
#### - online recommender



### ~ tj-filter-views







Overall code **②** 84.70 > 80.00





Контрольная группа	Тестовая группа	Метрика	Изменение метрики	p value
Отсутствие рекомендаций	iALS	Конверсия в клик	+28.7%	0
		Retention 1d	+0.6%	0.3748
iALS	GRU4Rec	Конверсия в клик	+8.35%	0.0007
		Retention 1d	+13.55%	0
GRU4Rec	SASRec	Конверсия в клик	+24.5%	0.006
		Retention 1d	+6.4%	0.17

# Подобрано специально для вас



В регионах России все чаще отключают мобильный интернет: что нужно знать и как подготовиться



Сколько должен длиться идеальный отпуск с точки зрения науки

© 21 Q 51 Q 30



В 10 мегаполисах РФ цена жилья с учетом ипотеки превысила миллион рублей за квадратный метр