**گزارش بخش objective**

**پارامترها:**

- batch\_size (اندازه دسته): تعداد نمونه‌هایی که در یک بار از آموزش به مدل داده می‌شود.

- learning\_rate (نرخ یادگیری): سرعت به‌روزرسانی وزن‌های مدل در طول آموزش.

**متریک:**

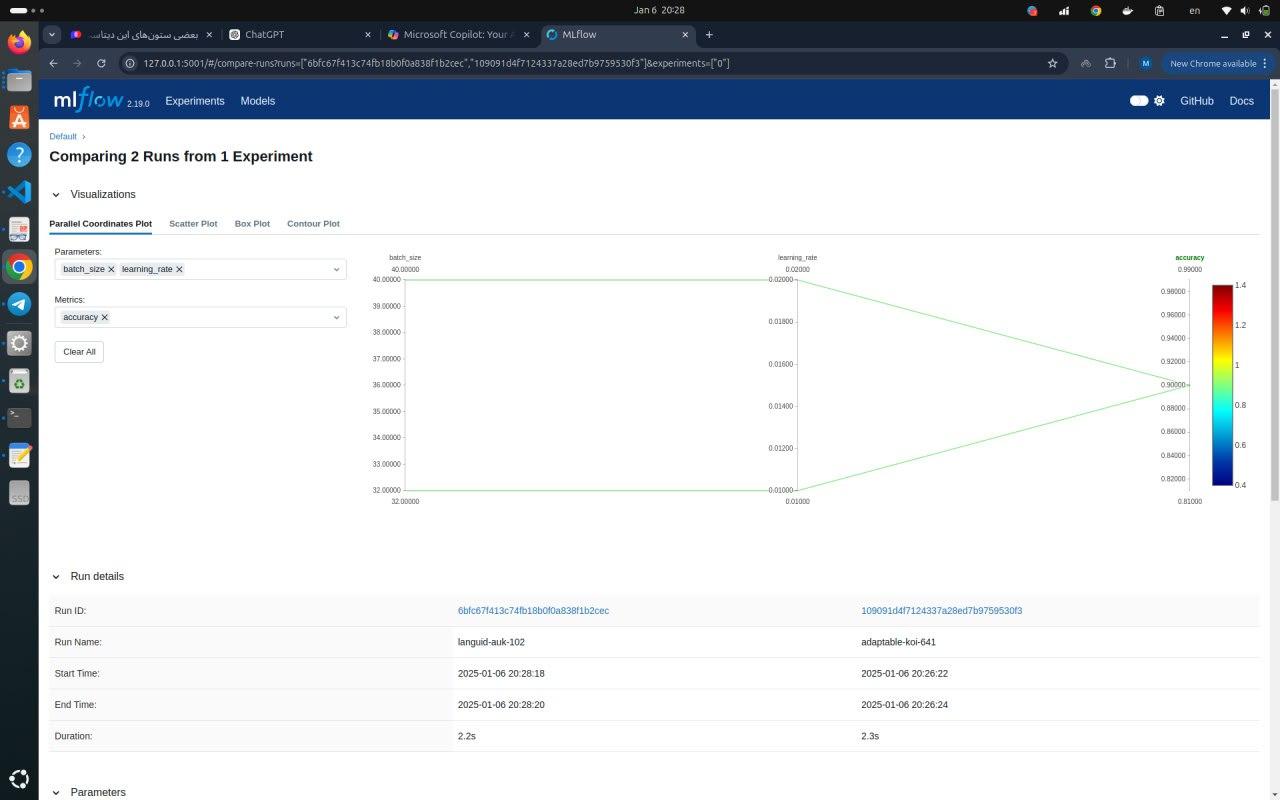
- دقت (accuracy): میزان صحت پیش‌بینی‌های مدل که به صورت درصدی از پیش‌بینی‌های درست محاسبه می‌شود.

**توضیح نمودار مختصات موازی:**

نمودار مختصات موازی نمایشی از ارتباط بین پارامترهای batch\_size و learning\_rate با متریک accuracy است. در این نمودار، مقیاس رنگی که در سمت راست قرار دارد، مقادیر دقت را از 0.81 تا 1.4 نشان می‌دهد. خطوط مختلف در نمودار نشان‌دهنده دو اجرای مختلف هستند که اثر تغییرات در مقادیر پارامترها روی دقت مدل را نمایش می‌دهند.

نتیجه‌گیری:

این نمودار می‌تواند برای درک بهتر چگونگی تأثیر تغییرات در پارامترهای batch\_size و learning\_rate بر دقت مدل مفید باشد. به‌ویژه با مقایسه میان دو اجرای مختلف، می‌توان مشاهده کرد که در کدام ترکیب از پارامترها، دقت مدل بهبود یافته است. این نوع بصری‌سازی می‌تواند به تسهیل فرآیند انتخاب بهترین تنظیمات پارامترها کمک کند تا عملکرد مدل بهینه شود.



**گزارش بخش model registry**

**ثبت مدل**

- مدل‌ها پس از دستیابی به عملکرد مناسب در داده‌های آزمایشی در MLflow Model Registry ثبت می‌شوند.

- فرایند ثبت از طریق API mlflow.register\_model() انجام می‌شود.

**معیارهای انتقال مدل**

- انتقال از Staging به Production: مدل زمانی به مرحله "Production" منتقل می‌شود که معیارهای زیر را داشته باشد:

- دقت بیشتر از 80٪ در داده‌های آزمایشی.

- عدم وجود مشکلات عمده در آزمایش‌های عملکرد.

**فرایند نسخه‌بندی**

- هر بار که یک اجرای جدید آموزش که بهبود در عملکرد یا ویژگی‌های مدل ایجاد می‌کند، انجام شود، یک نسخه جدید ثبت می‌شود.

- نسخه‌های مدل بر اساس معیارهایی مانند دقت مقایسه می‌شوند و آخرین نسخه پایدار به مرحله بالاتری منتقل می‌شود.

گزارش مقایسه مدل ها:

### **۱. بررسی RandomForestClassifier**

* **مدل استفاده‌شده**: طبقه‌بند جنگل تصادفی (پیاده‌سازی شده در Scikit-learn)
* **هابرپارامترها**:
  + تعداد تخمین‌گرها (n\_estimators): ۱۰۰
  + وضعیت تصادفی: ۴۲
* **داده‌ها**: مجموعه داده طبقه‌بندی دودویی مصنوعی تولیدشده با استفاده از make\_classification.
  + ویژگی‌ها: ۲۰
  + نمونه‌ها: ۱۰۰۰
  + کلاس‌ها: ۲
* **معیارهای عملکرد**:
  + **دقت**: 0.9

گزارش lifecycle مدلها:  
**1. توسعه و تست مدل**

* **هدف**: توسعه و آموزش یک مدل با استفاده از RandomForestClassifier برای دسته‌بندی.
* **مراحل**:
  + تولید داده‌های مصنوعی با استفاده از make\_classification شامل 20 ویژگی و 1000 نمونه.
  + تقسیم داده‌ها به مجموعه‌های آموزشی و آزمایشی با استفاده از train\_test\_split.
  + آموزش مدل بر روی داده‌های آموزشی.
  + ارزیابی مدل بر روی داده‌های آزمایشی با استفاده از معیار **دقت (accuracy)**.
* **ثبت در MLflow**:
  + **پارامترها**:
    - learning\_rate، batch\_size، cpu، ram.
  + **متریک‌ها**:
    - **دقت** بر روی مجموعه آزمایشی.
  + **آرتیفکت‌های مدل**:
    - مدل آموزش‌دیده با استفاده از mlflow.sklearn.log\_model ذخیره می‌شود.

#### **2. ثبت مدل**

* **هدف**: ثبت مدل آموزش‌دیده در **MLflow Model Registry** برای ردیابی و نسخه‌بندی.
* **مراحل**:
  1. ثبت مدل آموزش‌دیده با یک **URI** یکتا (runs:/[run\_id]/model).
  2. ثبت مدل با یک نام یکتا (RandomForestModel).

#### **3. اعتبارسنجی مدل**

* **هدف**: اعتبارسنجی مدل برای آماده‌سازی جهت تولید بر اساس معیارهای تعریف‌شده.
* **معیارها**:
  1. **آستانه دقت**: دقت مدل باید بیشتر از **0.8** باشد.
  2. متریک‌های اختیاری اضافی (در کد پیاده‌سازی نشده اما پیشنهاد می‌شوند):
     + دقت (precision)، یادآوری (recall)، نمره F1.
     + متریک‌های تأخیر (latency) و استفاده از منابع.
* **فرآیند**:
  1. بازیابی **آخرین نسخه ثبت‌شده** از مدل با استفاده از MlflowClient.
  2. دریافت متریک accuracy برای مدل با استفاده از run\_id.
  3. مقایسه دقت با آستانه تعیین‌شده.

#### **4. انتقال به تولید**

* **هدف**: ارتقا مدل به مرحله "تولید" اگر معیارهای اعتبارسنجی را برآورده کند.
* **مراحل**:
  1. اگر دقت از آستانه بیشتر باشد:
     + مرحله مدل به **Production** منتقل می‌شود با استفاده از transition\_model\_version\_stage.
     + به ذی‌نفعان اطلاع داده می‌شود.
  2. اگر مدل در اعتبارسنجی رد شود:
     + مدل در مرحله پیش‌فرض باقی می‌ماند (مثل "Staging").
     + به ذی‌نفعان درباره دلایل رد شدن و مراحل بعدی برای بهبود اطلاع داده می‌شود.