**گزارش درباره‌ی کد مرحله سوم جداسازی داده‌ها**

این کد به منظور پردازش و جداسازی داده‌های نقاط کلیدی استخراج‌شده از ویدیوهای حرکات ورزشی طراحی شده است. هدف این برنامه، سازمان‌دهی داده‌ها به تفکیک نوع حرکت ورزشی و ذخیره آن‌ها در قالب دیتاست‌های جداگانه برای هر حرکت است. در این فرآیند، داده‌ها از فایل‌های CSV خوانده می‌شوند و پس از بررسی صحت و یکپارچگی، داده‌ها به قالبی مناسب برای استفاده در مدل‌های یادگیری ماشین تبدیل می‌شوند.

**هدف برنامه:**

هدف این برنامه، جداسازی داده‌های نقاط کلیدی و سازمان‌دهی آن‌ها بر اساس نوع حرکت ورزشی است. این داده‌ها از فایل‌های CSV استخراج‌شده به‌صورت نقاط (x, y) ذخیره می‌شوند که نمایانگر موقعیت نقاط بدن انسان در هر فریم از ویدیوهای ورزشی هستند. این برنامه، این داده‌ها را بررسی کرده، از نظر صحت و یکپارچگی، فیلتر کرده و در دیتاست‌های جداگانه برای هر حرکت ذخیره می‌کند.

**مراحل اجرای برنامه:**

1. **بررسی وجود فایل‌ها و پوشه‌ها:**
   * در ابتدا، مسیر پوشه حاوی فایل‌های CSV (fixed\_keypoints\_output) بررسی می‌شود.
   * اگر این پوشه یا فایل‌ها موجود نباشند، یک خطا به اطلاع کاربر رسانده می‌شود.
2. **ایجاد دیکشنری برای ذخیره داده‌ها:**
   * یک دیکشنری dataset\_dict برای ذخیره داده‌ها به تفکیک نام حرکت ورزشی ایجاد می‌شود. هر کلید در این دیکشنری نام حرکت ورزشی است و مقدار آن لیستی از بردارهای ویژگی (نقاط کلیدی در قالب یک آرایه یک‌بعدی) برای هر فایل مربوط به آن حرکت می‌باشد.
3. **پردازش فایل‌های CSV:**
   * تمامی فایل‌های CSV موجود در پوشه بررسی و بارگذاری می‌شوند. در این مرحله:
     + نام حرکت ورزشی از نام فایل استخراج می‌شود (قبل از عبارت "*frame*").
     + داده‌های CSV خوانده شده و بررسی می‌شود که آیا ساختار داده‌ها مطابق با استاندارد (ستون‌های "x" و "y" برای هر نقطه) است یا خیر.
     + نقاط (x, y) به یک بردار ویژگی یک‌بعدی تبدیل می‌شود.
     + در صورتی که تعداد نقاط کلیدی برای هر حرکت ورزشی با تعداد استاندارد مطابقت نداشته باشد یا ساختار فایل نامعتبر باشد، آن فایل نادیده گرفته می‌شود.
4. **یکپارچگی داده‌ها:**
   * در حین پردازش، برنامه بررسی می‌کند که تعداد نقاط کلیدی برای هر حرکت ورزشی یکسان باشد. اگر تعداد نقاط کلیدی در هر فایل برای یک حرکت ورزشی متفاوت باشد، آن فایل نادیده گرفته می‌شود.
5. **ذخیره داده‌ها به تفکیک حرکت ورزشی:**
   * پس از پردازش و فیلتر کردن فایل‌ها، داده‌ها به تفکیک حرکت ورزشی در دیکشنری dataset\_dict ذخیره می‌شود.
   * این داده‌ها به قالب دیتافریم تبدیل شده و در فایل‌های CSV جداگانه برای هر حرکت ورزشی ذخیره می‌شوند.
6. **ایجاد پوشه‌های خروجی:**
   * یک پوشه جدید به نام final\_datasets\_by\_exercise\_1 ایجاد می‌شود که داده‌های نهایی به تفکیک حرکت ورزشی در آن ذخیره می‌شود.
7. **ذخیره داده‌ها در فایل CSV:**
   * برای هر حرکت ورزشی، یک دیتافریم ایجاد می‌شود که شامل ستون‌های "x" و "y" برای هر نقطه کلیدی است.
   * این دیتافریم‌ها در فایل‌های CSV جداگانه ذخیره می‌شوند و پیام تایید ذخیره هر فایل در کنسول نمایش داده می‌شود.

**جزئیات کد:**

* **خواندن و بررسی فایل‌های CSV:**
  + در این بخش، نام فایل‌ها بررسی می‌شود که با فرمت صحیحی ذخیره شده باشند و ساختار آن‌ها شامل ستون‌های "x" و "y" باشد. اگر ساختار داده‌ها نامعتبر باشد، فایل نادیده گرفته می‌شود.
* **دیکشنری برای ذخیره داده‌ها:**
  + داده‌ها به تفکیک نام حرکت ورزشی در دیکشنری dataset\_dict ذخیره می‌شوند. این دیکشنری شامل نام حرکت ورزشی به عنوان کلید و لیستی از نقاط کلیدی به عنوان مقدار است.
* **یکپارچگی داده‌ها:**
  + برنامه مطمئن می‌شود که برای هر حرکت ورزشی، تعداد نقاط کلیدی (بردار ویژگی‌ها) یکسان باشد. اگر تعداد نقاط در یک فایل با سایر فایل‌ها برای همان حرکت ورزشی متفاوت باشد، آن فایل نادیده گرفته می‌شود.
* **ذخیره فایل‌های نهایی:**
  + در نهایت، دیتافریم‌هایی برای هر حرکت ورزشی ایجاد می‌شود و در پوشه‌ای جدید ذخیره می‌شوند. این فایل‌ها شامل اطلاعات به‌صورت "x1", "x2", ..., "xn", "y1", "y2", ..., "yn" برای نقاط کلیدی هستند.

**نتیجه‌گیری:**

این کد یک فرایند کارآمد برای پردازش و جداسازی داده‌های نقاط کلیدی ورزشی است که به‌طور خودکار داده‌ها را از فایل‌های CSV فیلتر کرده و آن‌ها را به قالبی استاندارد برای هر حرکت ورزشی تبدیل می‌کند. پس از پردازش و اصلاح داده‌ها، آن‌ها در فایل‌های CSV جداگانه ذخیره می‌شوند که می‌توان از آن‌ها برای آموزش مدل‌های یادگیری ماشین یا تجزیه و تحلیل‌های دیگر استفاده کرد.

با استفاده از این کد، می‌توان به راحتی داده‌های نقاط کلیدی را از ویدیوهای حرکات ورزشی سازمان‌دهی کرده و آن‌ها را برای تحقیقات بعدی یا استفاده در مدل‌های یادگیری ماشین آماده کرد.