**گزارش مفصل سیستم تحلیل تراکنش‌های بانکی**

**1. مقدمه**

این کد با استفاده از زبان پایتون و کتابخانه‌های Streamlit، OpenCV، EasyOCR، Tesseract و Pandas یک سیستم جامع برای تحلیل تراکنش‌های بانکی ارائه می‌دهد. هدف اصلی این سیستم استخراج نام مشتری از تصویر فیش بانکی، جستجو در فایل اکسل مشتریان و ارائه آمارهای کلیدی همراه با نمودارهای تعاملی و تحلیل‌های پیشرفته مانند روند ماهانه، هفتگی، ساعتی و محاسبه میانگین و انحراف معیار تراکنش‌ها است.

**2. ساختار کلی کد**

کد به چند بخش اصلی تقسیم می‌شود:

* **تنظیمات اولیه:** پیکربندی مسیرهای Tesseract و راه‌اندازی EasyOCR با پشتیبانی از زبان‌های فارسی و انگلیسی.
* **تعریف توابع اصلی:** شامل توابع پردازش تصویر، تبدیل اعداد انگلیسی به فارسی، اجرای OCR با استفاده از دو موتور Tesseract و EasyOCR، استخراج نام مشتری و جستجو در سوابق اکسل.
* **رابط کاربری با استفاده از Streamlit:** بخش‌هایی برای آپلود فایل‌های اکسل و تصویر فیش بانکی، نمایش نتایج استخراج نام، آمارهای اولیه و تحلیلی، نمودارها و گزارش‌های آماری.
* **تحلیل‌های اضافی:** بخش‌هایی برای بررسی فراوانی بیشترین مبالغ تراکنش‌ها (واریز و برداشت) در یک روز خاص، روند ماهانه، هفتگی و ساعتی تراکنش‌ها و محاسبه آماری میانگین و انحراف معیار مبالغ.

**3. بررسی اجزای کلیدی کد**

**3.1. OCR و پردازش تصویر**

* **enhance\_image** این تابع تصویر را به حالت خاکستری تبدیل کرده، نویز تصویر را کاهش داده و از تکنیک CLAHE برای بهبود کنتراست استفاده می‌کند.
* **ocr\_ensemble** با استفاده از دو موتور OCR (Tesseract و EasyOCR) متن استخراج شده از تصویر را ترکیب می‌کند تا دقت استخراج افزایش یابد.
* **extract\_customer\_name** این تابع با استفاده از الگوهای منظم (Regex) نام مشتری را از متن استخراج می‌کند و اعداد انگلیسی موجود را به فارسی تبدیل می‌کند.

**3.2. پردازش و جستجوی داده‌های اکسل**

* **find\_customer\_records** سوابق مشتری در فایل اکسل بر اساس نام استخراج شده جستجو می‌شود. این تابع از چندین نام محتمل برای ستون نام مشتری استفاده می‌کند تا در صورت تغییر نام ستون نیز کارکرد داشته باشد.

**3.3. تحلیل‌های آماری و نمودارها**

* **تحلیل‌های اولیه:** شامل محاسبه مجموع واریزها، برداشت‌ها و موجودی خالص مشتری.
* **تحلیل‌های اضافی:**
  + بررسی فراوانی بیشترین مبلغ واریز و برداشت (با توجه به ثبت اعداد منفی برای برداشت).
  + تحلیل تراکنش‌ها به تفکیک روزانه، ماهانه، هفتگی و حتی ساعتی.
  + استفاده از نمودارهای خطی و میله‌ای جهت نمایش روند تراکنش‌ها.
  + محاسبه میانگین و انحراف معیار مبالغ تراکنش‌ها برای ارائه دیدی کلی از پراکندگی داده‌ها.

**3.4. رابط کاربری با Streamlit**

* استفاده از Streamlit باعث می‌شود تا رابط کاربری بسیار تعاملی و کاربرپسند باشد.
* امکان آپلود دو فایل (اکسل و تصویر) به کاربر داده شده و در صورت عدم آپلود هر دو فایل، پیام مناسب نمایش داده می‌شود.
* استفاده از expander ها برای نمایش کامل سوابق و متن استخراج شده از فیش، امکان مشاهده جزئیات بیشتر را به کاربر می‌دهد.

**4. مزایا و نکات کلیدی**

* **دقت بالای استخراج متن:** ترکیب دو موتور OCR باعث افزایش دقت استخراج اطلاعات از فیش بانکی می‌شود.
* **انعطاف‌پذیری در پردازش داده‌ها:** استفاده از چندین الگو برای استخراج نام مشتری و جستجو در اکسل، کد را نسبت به تغییرات ساختار داده مقاوم می‌سازد.
* **تحلیل‌های جامع و بصری:** ارائه آمارهای کلیدی به همراه نمودارهای تعاملی و تحلیل‌های زمان‌بندی (روزانه، هفتگی، ماهانه و ساعتی) به مدیران بانکی یا کاربر نهایی دید دقیقی از تراکنش‌ها می‌دهد.
* **رابط کاربری ساده و کاربرپسند:** استفاده از Streamlit باعث می‌شود تا کاربران با حداقل دانش فنی بتوانند از سیستم بهره‌مند شوند.

**5. نتیجه‌گیری**

این کد نمونه‌ای جامع از چگونگی استفاده از فناوری‌های مختلف در پردازش تصویر، استخراج متن و تحلیل داده‌های بانکی است. با ترکیب تکنیک‌های پردازش تصویر، OCR و تحلیل داده‌های زمانی و آماری، سیستم قابلیت ارائه گزارش‌های دقیق و مفصل از تراکنش‌های بانکی را دارد. این سیستم می‌تواند به عنوان یک ابزار قدرتمند در حوزه بانکداری و مالی جهت بهبود فرآیندهای نظارتی و مدیریتی به کار رود.