

## به نام خدا

### پروژه: سامانه پایش و شناسایی ناهنجاری در داده‌های بازار بورس

#### چکیده اجرایی :

این پروژه یک سامانه تحلیلی برای شناسایی خودکار الگوهای غیرعادی (Anomalies) در داده‌های قیمت و حجم معاملات بورس بر اساس داده‌های شبیه‌سازی شده ارائه می‌دهد. هدف اصلی، کمک به تحلیلگران و نهادهای ناظر برای کشف سریع رخداد‌های مشکوک، دستکاری‌های بازار، یا نوسانات غیرمنتظره است. این سامانه با استفاده از هر دو روش آماری (IQR: interquartile range) و الگوریتم یادگیری ماشین (Isolation Forest) به تحلیل داده می‌پردازد و یک «امتیاز ریسک» یکپارچه برای هر روز معاملاتی محاسبه می‌کند.

#### اهداف پروژه:

کشف تقلب و دستکاری: شناسایی علائم هشداردهنده اولیه برای معاملات غیرمعارف.  
مدیریت ریسک: ارائه یک معیار کمی (امتیاز ریسک ۰ تا ۱۰) برای کمک به اولویت‌بندی بررسی‌های روزانه.  
اتوماسیون فرآیند: خودکارسازی نظارت بر بازار که به صورت دستی زمانبر است.

#### روش‌شناسی و معماری پروژه :

این پروژه در بستر برنامه پایتون نوشته و اجرا می‌گردد و داده‌ها بصورت شبیه‌سازی شده با یکی از کتابخانه‌های قدرتمند پایتون ایجاد شده‌اند. در این برنامه داده‌های شبیه‌سازی شده ایجاد و تحلیل می‌گردد و بر اساس متدولوژی IQR ناهنجاری‌های داده بر روی دو پارامتر حجم معاملات و قیمت پایانی در بازه‌های زمانی دلخواه (در این پروژه ۷ روزه) تشخیص داده شده و نتایج آن بصورت جدول و نمودار قابل دسترسی می‌باشد.  
در قسمت یادگیری ماشین از روش Isolation Forest Algorithm از نوع یادگیری نظارت نشده استفاده شده است که می‌توان نتایج این روش را با روش تحلیل IQR مقایسه نمود.

برخی کتابخانه‌های مورد استفاده در پروژه شامل Pandas, Numpy, Matplotlib.pyplot, Seaborn, و scikit-learn می‌باشند.

## نتیجه‌گیری و گام‌های آتی

این پروژه به عنوان یک فیلترینگ اولیه برای شناسایی نشانه‌های آنومالی داده‌ها می‌باشد و نشان می‌دهد چگونه می‌توان با ابزارهای ساده داده‌کاوی، بینش ارزشمندی از داده‌های مالی استخراج کرد.

### گام‌های بعدی برای توسعه و بهبود پروژه:

- استفاده از داده‌های واقعی و اتصال به API های مالی.
- مانا سازی (Stationary) داده‌های سهام خصوصا هنگام استفاده از مدل IQR و نیز برای مدل سازی داده‌های سری زمانی.
- الگوریتم‌های پیشرفته‌تر: استفاده از مدل‌های LSTM برای شناسایی ناهنجاری در داده‌های سری زمانی.
- در نظر گیری ویژگی‌های تکنیکال: افزودن اندیکاتورهایی مانند RSI، MACD، و Bollinger Bands به عنوان features مدل.
- ایجاد داشبورد تعاملی: ساخت یک وب اپلیکیشن یا داشبورد با استفاده از Plotly Dash، Streamlit یا Power BI برای تعامل کاربر.
- در نظر گیری روند (Trend) و فصلیت (Seasonality)، انجام تفکیک سری زمانی (Decomposition) یا استفاده از مدل‌هایی مانند ARIMA برای حذف روند قبل از شناسایی ناهنجاری.