

# نظرة عامة على مشروع تحليل والتنبؤ بالتحويلات المالية في اليمن

## المقدمة

يهدف هذا المشروع إلى تحليل شامل للتحويلات المالية السنوية التي تلقاها اليمن من عام 1990 إلى عام 2023، بالإضافة إلى بناء نموذج تعلم آلي للتنبؤ بالتحويلات المستقبلية. تعتمد الدراسة على بيانات البنك الدولي المفتوحة، وتهدف إلى تقديم رؤى حول الاتجاهات التاريخية للتحويلات، وتقييم أداء نموذج التنبؤ، وتقديم توصيات بناءً على النتائج.

تُعد التحويلات المالية من المغتربين عنصرًا حيويًا للاقتصاد اليمني، حيث تساهم بشكل كبير في دعم الأسر، وتخفيف حدة الفقر، وتوفير العملة الصعبة اللازمة لاستيراد السلع الأساسية، خاصة في ظل الظروف الاقتصادية الصعبة التي تمر بها البلاد.

## مراحل المشروع

تم تنفيذ هذا المشروع على عدة مراحل رئيسية، كل مرحلة تركز على جانب معين من تحليل البيانات وبناء النموذج:

### المرحلة 1: جمع البيانات وتنظيفها

**الهدف:** الحصول على بيانات التحويلات المالية لليمن من المصادر المتاحة وتنظيفها لتكون جاهزة للتحليل.

### الخطوات المنفذة:

- 1. تحديد المصادر:** تم تحديد بيانات التحويلات المالية من موقع البنك الدولي للبيانات المفتوحة (World Bank Open Data) ضمن مؤشرات التنمية العالمية (World Development Indicators).
- 2. تنزيل البيانات:** تم تنزيل الملفات بصيغة CSV التي تحتوي على بيانات التحويلات المالية.
- 3. قراءة البيانات الأولية:** تم قراءة ملف `API_BX.TRF.PWKR.CD.DT_DS2_en_csv_v2_38454.csv` باستخدام مكتبة `pandas` في `Python`. لوحظ أن الملف يحتوي على صفوف علوية تحتوي على بيانات وصفية (metadata)، يجب تخطيها.
- 4. تنظيف البيانات:**
  - تم تصفية البيانات لتشمل اليمن فقط (`'Country Code == 'YEM'`).
  - تم حذف الأعمدة غير الضرورية مثل `Country Name, Country Code, Indicator`، بالإضافة إلى أي أعمدة فارغة غير مسماة قد تظهر في نهاية الملف.
  - تم تحويل البيانات من التنسيق العريض (wide format) إلى التنسيق الطويل (long format) باستخدام `melt`، حيث أصبحت السنوات عمودًا واحدًا وقيم التحويلات في عمود آخر.
  - تم تحويل أعمدة `Year` و `Remittances` إلى أنواع بيانات رقمية، مع التعامل مع القيم غير الصالحة كـ `NaN`.
  - تم إزالة الصفوف التي تحتوي على قيم `NaN` في أي من الأعمدة الرئيسية.
  - تم تصفية البيانات لتشمل الفترة الزمنية المطلوبة فقط (من 1990 إلى 2023).
- 5. حفظ البيانات النظيفة:** تم حفظ البيانات النظيفة في ملف CSV جديد باسم `yemen_remittances_cleaned.csv`.

## المرحلة 2: التحليل الإحصائي والاستكشافي

**الهدف:** فهم الاتجاهات والأنماط الرئيسية في بيانات التحويلات المالية من خلال التحليل الإحصائي والاستكشافي.

### الخطوات المنفذة:

1. **الإحصائيات الوصفية:** تم حساب وعرض الإحصائيات الوصفية للبيانات (المتوسط، الانحراف المعياري، الحد الأدنى، الحد الأقصى، الأرباع) لفهم التوزيع العام للتحويلات.
2. **التحويلات عبر الزمن:** تم عرض بيانات التحويلات مرتبة حسب السنة لمراقبة التغيرات السنوية.
3. **التغير السنوي:** تم حساب التغير السنوي في التحويلات (الفرق بين قيمة التحويلات في سنة معينة وقيمتها في السنة السابقة) لإبراز فترات النمو والانكماش.
4. **حفظ البيانات المحللة:** تم حفظ البيانات التي تحتوي على عمود التغير السنوي في ملف CSV جديد باسم `yemen_remittances_analyzed.csv`.

## المرحلة 3: التصور البياني للبيانات

**الهدف:** إنشاء رسوم بيانية واضحة ومفهومة لتصوير الاتجاهات والأنماط المكتشفة في بيانات التحويلات.

### الخطوات المنفذة:

1. **إعداد بيئة الرسم:** تم استخدام مكتبة `matplotlib` في Python لإنشاء الرسوم البيانية. تم التأكد من دعم الخطوط العربية (خط Amiri) لضمان عرض صحيح للعناوين والتسميات باللغة العربية.
2. **رسم بياني للتحويلات عبر الزمن:** تم إنشاء رسم بياني خطي يوضح تطور التحويلات المالية السنوية في اليمن من 1990 إلى 2023. تم حفظ هذا الرسم البياني باسم `yemen_remittances_time_series.png`.
3. **رسم بياني للتغير السنوي:** تم إنشاء رسم بياني شريطي يوضح التغير السنوي في التحويلات المالية. تم حفظ هذا الرسم البياني باسم `yemen_remittances_yoy_change.png`.

## المرحلة 4: بناء وتدريب نموذج التعلم الآلي

**الهدف:** بناء نموذج تعلم آلي يمكنه التنبؤ بقيمة التحويلات المالية المستقبلية بناءً على البيانات التاريخية.

### الخطوات المنفذة:

1. **اختيار النموذج:** تم اختيار نموذج الانحدار الخطي البسيط (Linear Regression) نظرًا لطبيعة البيانات (التنبؤ بقيمة عددية مستمرة) والبساطة الأولية للنموذج.
2. **تقسيم البيانات:** تم تقسيم البيانات إلى مجموعتي تدريب واختبار (80% للتدريب و 20% للاختبار) لتقييم أداء النموذج بشكل موضوعي.
3. **تدريب النموذج:** تم تدريب نموذج الانحدار الخطي باستخدام بيانات التدريب، حيث كانت السنة هي المتغير المستقل والتحويلات هي المتغير التابع.
4. **التنبؤ بالتحويلات المستقبلية:** تم استخدام النموذج المدرب للتنبؤ بقيمة التحويلات للسنوات من 2024 إلى 2030.
5. **حفظ التنبؤات:** تم حفظ التنبؤات المستقبلية في ملف CSV باسم `yemen_remittances_predictions.csv`.

## المرحلة 5: تقييم أداء النموذج

الهدف: قياس مدى دقة وفعالية النموذج التنبؤي.

### الخطوات المنفذة:

1. **مقاييس التقييم:** تم استخدام مقياسين رئيسيين لتقييم أداء النموذج:
  - **متوسط الخطأ التربيعي (Mean Squared Error - MSE):** يقيس متوسط مربعات الأخطاء بين القيم المتوقعة والفعلية.
  - **معامل التحديد (R-squared - R2):** يوضح نسبة التباين في المتغير التابع التي يفسرها النموذج.
2. **عرض النتائج:** تم عرض قيم MSE و R2 المحسوبة لتقييم مدى ملائمة النموذج للبيانات.

## المرحلة 6: إعداد التقرير النهائي

الهدف: تجميع جميع النتائج والتحليلات والتوصيات في تقرير شامل ومفهوم.

### الخطوات المنفذة:

1. **هيكلية التقرير:** تم إنشاء ملف Markdown باسم `yemen_remittances_report.md` يحتوي على الأقسام التالية:
  - مقدمة المشروع وأهدافه.
  - نظرة عامة على البيانات المستخدمة.
  - التحليل الإحصائي والاستكشافي للبيانات.
  - التصورات البيانية مع تضمين الصور التي تم إنشاؤها.
  - تفاصيل بناء وتقييم نموذج التعلم الآلي، بما في ذلك مقاييس الأداء والتنبؤات المستقبلية.
  - الخلاصة والتوصيات المستنبطة من التحليل.
2. **تضمين النتائج:** تم دمج جميع النتائج العددية والرسوم البيانية والتنبؤات في التقرير بشكل منظم.

## الخلاصة والتوصيات

أظهر التحليل أن التحويلات المالية إلى اليمن شهدت نموًا ملحوظًا، مع قفزة كبيرة في عام 2012 واستقرار عند مستويات عالية في السنوات اللاحقة. على الرغم من أن نموذج الانحدار الخطي البسيط قدم تنبؤات معقولة، إلا أن قيمة MSE المرتفعة تشير إلى أن هناك مجالًا لتحسين دقة التنبؤات، ربما من خلال استخدام نماذج أكثر تعقيدًا أو دمج متغيرات إضافية.

### التوصيات الرئيسية:

- **استكشاف نماذج متقدمة:** النظر في استخدام نماذج تعلم آلي أكثر تعقيدًا (مثل نماذج السلاسل الزمنية المتقدمة كـ ARIMA أو Prophet) لتحسين دقة التنبؤات.
- **دمج عوامل إضافية:** البحث عن بيانات إضافية قد تؤثر على التحويلات (مثل أسعار النفط، الأوضاع الاقتصادية في الدول المضيفة للمغتربين، الأحداث السياسية في اليمن) ودمجها في النموذج.
- **تحسين جودة البيانات:** العمل على تحسين دقة واكتمال بيانات التحويلات، خاصة في السنوات الأخيرة، لضمان تنبؤات أكثر موثوقية.

● **تسهيل التحويلات:** الاستمرار في دعم وتسهيل قنوات التحويل الرسمية لضمان استمرارية تدفق هذه الأموال الحيوية.

يهدف هذا المشروع إلى توفير أساس لفهم أعمق للتحويلات المالية في اليمن، ويمكن توسيعه في المستقبل ليشمل تحليلات أكثر تعقيداً وتنبؤات أكثر دقة.