الاختصارات

|  |  |  |
| --- | --- | --- |
| التكرار والتردد والقيمة النقدية | Recency, Frequency, Monetary Value | RFM |
| التعلم الآلي | Machine Learning | ML |
| آلة المتجهات الداعمة | Support Vector Machine | SVM |
| الغابات العشوائية | Random Forests | RF |
| الشبكات العصبية | Neural Networks | NN |
| التعلم العميق | Deep Learning | DL |
| الشبكات العصبية المتكررة | Recurrent Neural Networks | RNN |
|  |  |  |

# الفصل الأول

التعريف بالمشروع

يتضمن هذا الفصل التعريف بالمشروع ومتطلباته.

## مقدمة

في العصر الحديث للتجارة الرقمية، أصبحت القدرة على فهم سلوك العملاء واتخاذ قرارات مدعومة بالبيانات عنصراً جوهرياً لنجاح مواقع التجارة الإلكترونية. وفي هذا السياق، تبرز منصات الذكاء الصنعي كحلول استراتيجية تمكّن هذه المواقع من تحسين أدائها وتعزيز ولاء العملاء. توفر هذه المنصات مجموعة من الخدمات الذكية، مثل التنبؤ بانقطاع العملاء، وتصنيفهم إلى شرائح ذات أنماط سلوكية متقاربة، وتقدير القيمة المستقبلية لكل عميل والعائد المتوقع منه. ومع ذلك، فإن الاستفادة من هذه القدرات غالباً ما تتطلب خبرات متقدمة في علوم البيانات وتحليلات الأعمال، مما قد يشكل حاجزاً أمام العديد من المواقع. ومن هنا تبرز الحاجة إلى منصة توفر هذه الخدمات بطريقة مؤتمتة وسهلة الاستخدام، لتتيح لمواقع التجارة الإلكترونية، بمختلف أحجامها، الاستفادة من الذكاء الاصطناعي في تعزيز استراتيجياتها التسويقية وتحقيق نمو مستدام.

## الهدف من المشروع

يندرج عملنا في هذا المشروع ضمن سياقين: السياق الأول نتطرق فيه للمسألة من منظور الذكاء الصنعي، حيث نحاول الاستفادة من آخر التطورات في مجال تعلم الآلة والتعلم العميق للوصول إلى نماذج ذكاء صنعي توفر خدمات تعزز استراتيجياتها وتزيد من أرباحها. والسياق الثاني نتطرق فيه إلى بناء تطبيق برمجي يستفيد من نماذج الذكاء الصنعي التي توصلنا إليها في السياق الأول للوصول إلى تطبيق مفيد عملياً، مع مراعاة أسس ومبادئ هندسة البرمجيات.

## المتطلبات الوظيفية

يجب أن يقدم النظام للمستخدم ما يلي:

1. السماح بإنشاء حساب جديد ضمن النظام
2. السماح بتسجيل الدخول من حساب مُنشأ سابقا.
3. السماح للمستخدم برفع بيانات عملاء موقع التجارة الإلكترونية الخاصة به بطرق متعددة.
4. السماح للمستخدم باستعراض بيانات تحليلية عامة حول العملاء الخاصة به.
5. السماح للمستخدم باستعراض عملاء موقع التجارة الالكترونية الخاصة به الذين يتوقع انسحابهم.
6. السماح للمستخدم باستعراض معلومات القيمة الدائمة للعميل لدى موقع التجارة الالكترونية الخاص به.
7. السماح للمستخدم باستعراض بيانات تقسيم العملاء إلى شرائح.

## المتطلبات غير الوظيفية

1. يجب أن يكون النظام آمناً، حيث يسمح فقط للمستخدمين المسجّلين باستخدامه.
2. يجب أن يوفر النظام واجهات سهلة الاستخدام وجيدة المظهر.
3. يجب أن يكون الكود البرمجي قابلاً للتعديل والصيانة.
4. يجب أن يكون النظام قابلاً للتوسع.
5. يجب أن يكون النظام قادراً على التعامل مع حجم بيانات كبير نسبيا مليون عملية شراء للعملاء.

# الفصل الثاني

الدراسة النظرية

يوضّح هذا الفصل مجموعة من المفاهيم النظرية المستخدمة ضمن العمل المقدم.

## توقع انسحاب العملاء Churn Prediction.

مقدمة احكي انو بتوفر مصاري للشركات

## المحولات

# الفصل الثالث

الدراسة المرجعية

يعرض هذا الفصل الأبحاث والبيانات المرتبطة بالعمل المقدم.

## مقدمة

إن مجال تحليل بيانات مستخدمين مواقع التجارة الالكترونية هو مجال واسع ويحتوي على العديد من التحاليل والمخططات التي تفيد مدراء هذه المواقع في فهم تعامل العملاء مع الموقع وتوقع الأرباح والقيام بالإجراءات الوقائية لتحقيق أكبر ربح، سنستعرض في الفقرات التالية بعض التحاليل والقياسات التي استخرجت من الأبحاث ضمن هذا المجال.

## توقع انسحاب العملاء Churn Prediction.

### تطور تقنيات توقع انسحاب العملاء: من التقنيات التقليدية إلى التعلم العميق.

في توقع انسحاب العملاء، غالبا ما تشمل الطرق التقليدية النماذج الإحصائية والخوارزميات مثل خوارزمية التكرار والتردد والقيمة النقدية (RFM) أو سلاسل ماركوف ولكن غالبا ما تكون قدرتها التنبؤية محدودة، نتيجة لذلك تم تطبيق تقنيات التعلم الآلي Machine Learning في هذا المجال وساهمت بشكل كبير في تطويره، حيث أظهرت الدراسات فعالية خوارزميات مثل Support Vector Machine (SVM)، والغابات العشوائية Random Forests (RF)، والشبكات العصبية Neural Networks (NNs) وبالرغم من تفوقها إلا أنها غالبا ما تكون غير قادرة على نمذجة الأنماط الزمنية أو نمذجة الأنماط التتابعية في بيانات العملاء [1].

ومع التطورات في مجال التعلم العميق Deep Learning أظهرت الشبكات العصبية التكرارية (RNNs) Recurrent Neural Networks ومحولات البيانات Transformers إمكانيات واعدة في نمذجة بيانات سلوك العملاء المتغيرة مع الزمن، وتم استخدامها لالتقاط الأنماط التتابعية في بيانات العملاء، مما أدى إلى تفوقها على الطرق التقليدية وطرق التعلم الآلي في العديد من الحالات [2] [3].

### استخدام المحولات في توقع انسحاب العملاء.

في توقع انسحاب العملاء، ظهرت بنى المحول Transformer كأداة قوية لتحليل بيانات سلوك العملاء المتسلسلة. ويأتي تفوق هذه البنى في التقاط الأنماط طويلة المدى ضمن بيانات السلاسل الزمنية دون الحاجة الى هندسة ميزات معقدة. وخصوصا نماذج المحولات التي تستخدم آليات الانتباه متعددة الرؤوس Multi-Head Attention، فهي فعالة عند التعامل مع بيانات مختلطة، تجمع بين مقاييس RFM المتغيرة زمنياً والمتغيرات الثابتة للعملاء. ان المعالجة المباشرة لبيانات السلاسل الزمنية بواسطة المحولات تقلل من الحاجة إلى هندسة ميزات معقدة. كما نماذج المحول توفر تقديرات لعدم اليقين، مما يساعد في حملات مخصصة للعملاء أكثر استهدافا، للحد من انسحاب العملاء وتقدير خطر انسحاب كل عميل [3].

### استخدام المحولات في توقع انسحاب عملاء بناءً على سلوك الاستخدام اليومي.

جرى استخدام المحول في قطاع B2B، تحديداً في مجال الاتصالات، للتنبؤ بانسحاب العملاء بناءً على بيانات سلوك الاستخدام اليومية. يتميز هذا النموذج بقدرته على التعامل مع بيانات سلاسل زمنية متعددة المتغيرات، وهو ما يسمح بتحليل سلوك العملاء المتغير بشكل دقيق عبر الزمن [3].

#### بنية النموذج:

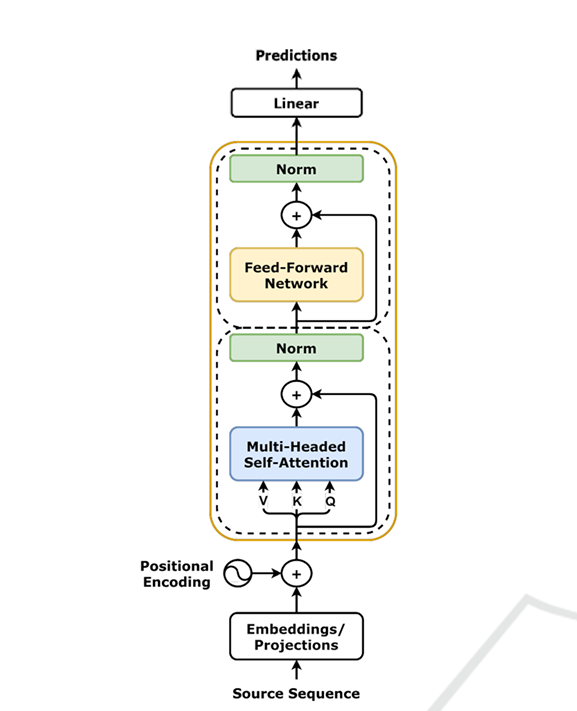
1. 

Figure 1 استخدام المحولات في توقع انسحاب العملاء بناءً على سلوك الاستخدام اليومي [3]

* تم استبدال طبقة التضمين (Embedding) التقليدية بـ 1D Depth wise Convolution.
* تم استخدام 8 طبقات Transformer مع رؤوس الانتباه المتعدد (Multi-Head Attention).
* يتم استخدام Global Average Pooling لتقليل حجم البيانات الناتجة وتحويلها إلى تمثيل نهائي.
* تم تطبيق وظيفة تفعيل Sigmoid في الطبقة النهائية لإنتاج احتمالية انسحاب العميل [3].

#### طبيعة البيانات المستخدمة:

كانت هي بيانات يومية على مستوى الاشتراك مثل: طول فترة التعاقد، عدد المكالمات، استهلاك البيانات، وعدد الباقات [3].

### استخدام المحولات مع تمثيلات RFM الزمنية.

تم بتطبيق نماذج Transformers على مقاييس RFM المتغيرة عبر الزمن، بهدف تحسين التنبؤ بانسحاب العملاء في قطاع الخدمات المالية. يُعدّ هذا النهج مهمًا لأن مقاييس RFM تعتبر من المؤشرات الأساسية لتقييم سلوك العملاء وقدرتهم على التنبؤ بالانسحاب [2].

#### بنية النموذج.

* دمج البيانات الثابتة والمتغيرة: تم دمج مقاييس RFM مع البيانات الثابتة باستخدام آلية Concatenation أو Merging.
* Attention Mechanism: تم استخدام آلية Multi-Head Attention لتعزيز قدرة النموذج على التركيز على أهم السمات الزمنية [2].

#### طبيعة البيانات المستخدمة.

مجموعة بيانات هي بيانات أحد كبار مزودي الخدمات المالية في فرنسا، وتحتوي على سجلات شهرية للعملاء. تتكون العينة المقدمة من عملاء تعتبر الشركة موفرهم الأساسي للخدمات المالية، وذلك لضمان جودة عالية في بيانات السلوك. تحتوي قاعدة البيانات على متغيرات تستخدم كثيراً في التنبؤ بانسحاب العميل، مثل الخصائص الديموغرافية والمعلومات السلوكية لحساب متغيرات RFM المتتابعة [2].

### تنبؤ وتمثيل مبني على سلاسل نشاط العملاء في مواقع التجارة الالكترونية.

جرى تحسين التنبؤ بسلوك العملاء من خلال معالجة بيانات النشاطات اليومية للعملاء كسلسلة زمنية. تم تكييف بنية المحولات لتحليل البيانات المجدولة أو السلوكية للعملاء. حيث توفر هذه الطريقة تمثيلاً عميقاً للبيانات الصعبة والمعقدة، دون الحاجة إلى هندسة خصائص مكثفة. اعتمدت هذه الطريقة في نموذج Customer Activity Sequence-based Prediction and Representation.

#### بنية النموذج.

* تحويل البيانات إلى سلاسل: تتحول البيانات المجدولة إلى سلسلة من الأحداث لكل عميل.
* تدريب النموذج باستخدام Transformer: تستخدم بنية Transformer لتحليل هذه السلاسل وتوليد تمثيلات متجهية (Embeddings) للعملاء.
* استخدام التمثيلات في المهام المختلفة: يمكن استخدام هذه التمثيلات في مهام متعددة مثل تنبؤ الانسحاب أو توصيات المنتجات.

#### طبيعة البيانات المستخدمة.

بيانات سجلات المشتريات أو الجلسات من مواقع مثل Microsoft Retail Stores او Instacart.