

الجمهورية العربية السورية
المعهد العالي للعلوم التطبيقية والتكنولوجيا
قسم المعلومات
العام الدراسي 2024/2025

مشروع تخرج

أعد لنيل درجة الإجازة في هندسة البرمجيات والذكاء الصناعي

منصة لخدمات الذكاء الصناعي موجهة لمواقع التجارة الإلكترونية

تقديم

آرام محمد

إشراف

د. رياض سنبل

5/8/2025

كلمة شكر

أتقدم بجزيل الشكر إلى مشرف المشروع الدكتور رياض سنبل، على دعمه المستمر وصبره الكريم وتوجيهاته السديدة خلال مراحل التنفيذ. لقد كان لحرصه وملاحظاته الدور الكبير في إثراء هذا العمل.

كما لا يفوتني أن أوجه شكري وتقديري العميق إلى جميع الدكاترة والأساتذة الذين كان لهم الفضل في تعليمنا وتوجيهنا طيلة السنوات الماضية. فقد غرسوا فينا قيم العلم والاجتهاد. فلکم منا كل الوفاء والاحترام.

الإهداء

إلى كل من يبحث عن اسمه بين سطور الإهداء...

اعلم أن مجرد بحثك يعني لي الكثير، ويكفيني دليلاً على حبك واهتمامك.

لك مني كل الشكر على وفائك، وامتناناً لا يُكتب بل يُحسّ.

وجودك في رحلتي كان نعمة، وإن لم يُذكر اسمك صراحة،

فمكانك محفوظ في القلب، حيث لا تُمحى الأسماء.

أهدي هذا العمل إلى كل من يهتمه الأمر

الخلاصة

في ظل تزايد أهمية الذكاء الصناعي في دعم القطاعات الحيوية كالـتجارة الإلكترونية، تبرز الحاجة إلى حلول ذكية تمكن أصحاب المتاجر من فهم سلوك العملاء واتخاذ قرارات تسويقية دون الحاجة إلى خبرات تقنية متقدمة. يهدف هذا العمل إلى تطوير منصة تعتمد على تقنيات الذكاء الصناعي لتحليل سلوك العملاء في بيئات التجارة الإلكترونية. يشمل العمل دراسة النماذج المستخدمة حالياً، وتحليل متطلبات أصحاب المتاجر، وتطوير منظومة برمجية مرنة قابلة للتوسع تدمج التحليل السلوكي للعملاء مع واجهات استخدام سهلة. توفر المنصة خدمات مثل تصنيف العملاء وفق شرائح سلوكية، ونموذج تنبؤي لاحتمالية انسحاب العملاء. وقد تم اختبار فعالية هذه النماذج باستخدام بيانات حقيقية، حيث أظهر النموذج التنبؤي أداءً واعداداً في تقديم تنبؤات بدقة إيجابية (Precision) بلغت 0.86، مع معدل إنذارات كاذبة منخفض، انعكس في استرجاع (Recall) بنسبة 0.81. تؤكد هذه النتائج قدرة المنصة على تقديم تحليل سلوكي موثوق يدعم التحول الذكي والمستدام في قطاع التجارة الإلكترونية.

Abstract

With the growing importance of artificial intelligence (AI) in supporting critical sectors such as e-commerce, there is an increasing need for intelligent solutions that enable store owners to understand customer behavior and make informed marketing decisions without requiring advanced technical expertise. This work aims to develop a platform based on AI techniques to analyze customer behavior in e-commerce environments. The study covers existing models, analyzes the specific needs of store owners, and develops a flexible, scalable software system that integrates customer behavior analytics with user-friendly interfaces. The platform provides services such as customer segmentation based on behavioral clusters and a predictive model for customer churn probability. The effectiveness of these models was tested using real-world data, where the predictive model demonstrated promising performance with a Precision of 0.86 and a low false alarm rate, reflected in a Recall of 0.81. These results highlight the platform's capability to deliver reliable behavioral analytics that support smart and sustainable transformation in the e-commerce sector.

فهرس المحتويات

13	الاختصارات
14	1 الفصل الأول: التعريف بالمشروع
14	1.1 مقدمة
14	2.1 الهدف من المشروع
15	3.1 المتطلبات الوظيفية
15	4.1 المتطلبات غير الوظيفية
16	2 الفصل الثاني: الدراسة النظرية
16	1.2 توقع انسحاب العملاء Churn Prediction
17	2.2 القيمة الدائمة لحياة العميل Customer Lifetime Value
17	3.2 الانتباه (Attention)
17	4.2 المحولات Transformers
17	5.2 تقسيم العملاء إلى شرائح
18	6.2 خوارزمية K-means
18	7.2 نموذج RFM
19	3 الفصل الثالث: الدراسة المرجعية
19	1.3 مقدمة
19	2.3 توقع انسحاب العملاء Churn Prediction
19	1.2.3 تطور تقنيات توقع انسحاب العملاء: من التقنيات التقليدية إلى التعلم العميق
20	2.2.3 استخدام المحولات في توقع انسحاب عملاء بناءً على سلوك الاستخدام اليومي

22	3.2.3 استخدام المحولات مع تمثيلات RFM الزمنية.
24	4.2.3 تنبؤ وتمثيل مبني على سلاسل نشاط العملاء في مواقع التجارة الالكترونية نموذج CASPR
26	3.3 استخدام خوارزمية K-Means في تقسيم عملاء متجر تجارة إلكترونية.
27	4.3 مجموعات البيانات في مجال التجارة الإلكترونية
27	1.4.3 مجموعة بيانات E-commerce Customer Data For Behavior Analysis
28	2.4.3 مجموعة بيانات Brazilian E-Commerce Public Dataset by OLIST
29	5.3 دراسة التطبيقات المشابهة
29	1.5.3 Baremetrics
29	2.5.3 Churnly
30	3.5.3 Pecan AI
31	4 الفصل الرابع: الدراسة التحليلية
31	1.4 مخططات حالات الاستخدام
33	2.4 السرد النصي لحالات الاستخدام
33	1.2.4 إنشاء حساب جديد
34	2.2.4 رفع بيانات المستخدم عن طريق ملف CSV
36	3.2.4 رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API
37	4.2.4 استعراض العملاء المتوقع انسحابها
38	5.2.4 استعراض شرائح العملاء
38	6.2.4 استعراض معلومات عن القيمة الدائمة للعملاء
40	5 الفصل الخامس: المنهجية المقترحة
40	1.5 مقدمة
40	2.5 النموذج المقترح لتوقع انسحاب العملاء

41	1.2.5	بنية النموذج المقترح لتوقع انسحاب العملاء
43	2.2.5	تحضير البيانات
43	3.2.5	منهجية التدريب
44	3.5	النموذج المقترح لتقسيم العملاء إلى شرائح
44	4.5	النموذج المقترح لحساب القيمة الدائمة للعميل
45	6	الفصل السادس: تصميم النظام
45	1.6	مقدمة
46	2.6	خدمة البيانات Data service
46	3.6	خدمة المصادقة Auth service
47	4.6	خدمة توقع انسحاب العملاء Churn service
47	5.6	خدمة تقسيم العملاء إلى شرائح Customer Segmentation Service
47	6.6	خدمة القيمة الدائمة للعملاء Customer Lifetime Value
48	7.6	خدمة البوابة API-Gateway
48	8.6	خدمة واجهة المستخدم Front-end
49	9.6	مخطط تصميم النظام
50	7	الفصل السابع: الأدوات المستخدمة
50	1.7	PostgreSQL
50	2.7	React
50	3.7	Python
50	4.7	FastAPI
50	5.7	Json Web Token (JWT)
51	6.7	Git

51 GitHub 7.7
51 Redis 8.7
51 PyTorch 9.7
51 Jupyter Notebook 10.7
52 8 الفصل الثامن: تنجيز النظام
52 1.8 مقدمة
53 2.8 بنية النظام ومكوناته
54 3.8 خدمة توقع انسحاب العملاء
54 1.3.8 تفاصيل التنجيز
56 4.8 خدمة القيمة الدائمة للعميل
56 1.4.8 تفاصيل التنجيز
58 5.8 خدمة تقسيم البيانات إلى شرائح
58 1.5.8 تفاصيل التنجيز
59 6.8 خدمة البيانات
59 1.6.8 تفاصيل التنجيز
60 2.6.8 جداول البيانات في قاعدة البيانات
62 7.8 خدمة المصادقة
62 1.7.8 تفاصيل التنجيز
63 8.8 خدمة البوابة
63 1.8.8 تفاصيل التنجيز
63 9.8 خدمة وجهات المستخدم
65 9 الفصل التاسع: اختبارات النظام ومناقشة النتائج

1.9	اختبار النموذج المقترح لتوقع انسحاب العملاء	65
1.1.9	منهجية التقييم المتبعة	65
2.9	نتائج نموذج تقسيم العملاء إلى شرائح	69
	الخاتمة والآفاق المستقبلية	71
	المراجع	72

فهرس الأشكال

الشكل 1	المحولات في توقع انسحاب العملاء بناءً على سلوك الاستخدام اليومي	20
الشكل 2	نموذج المحولات مع تمثيلات RFM الزمنية	22
الشكل 3	نموذج CASPR	24
الشكل 4	مجموعة بيانات OLIST	28
الشكل 5	بنية النموذج المقترح لتوقع الحساب العملاء	41
الشكل 6	مخطط تصميم النظام	49
الشكل 7	مجلدات النظام	53
الشكل 8	مجلدات خدمة توقع انسحاب العملاء	54
الشكل 9	مكونات مجلد utils في خدمة توقع انسحاب العملاء	54
الشكل 10	تواصل خدمة انسحاب العملاء بالخدمات الأخرى	55
الشكل 11	مجلدات خدمة القيمة الدائمة للعميل	56
الشكل 12	محتويات مجلد utils	56
الشكل 13	تواصل خدمة القيمة الدائمة للعميل بالخدمات الأخرى	57
الشكل 14	مجلدات خدمة تقسيم العملاء إلى شرائح	58
الشكل 15	تواصل خدمة تقسيم العملاء مع الخدمات الأخرى	58
الشكل 16	مجلدات خدمة البيانات	59
الشكل 17	أعمدة جدول upload_history	60
الشكل 18	أعمد جدول user_data	60
الشكل 19	أعمدة جدول user_predictions	61
الشكل 20	مجلدات خدمة المصادقة	62
الشكل 21	مجلدات خدمة واجهات المستخدم	63
الشكل 22	علاقة precision recall مع العتبة Threshold على مجموعة بيانات E-commerce	66
	Customer Data For Behavior Analysis	66

- الشكل 23 علاقة Precision و Recall مع العتبة Threshold على مجموعة بيانات OLIST 68
- الشكل 24 مؤشر السيولويت لتحديد عدد العناقيد الأمثلي 69
- الشكل 25 العلاقة بين الميزات بدلالة بعضها البعض 70
- الشكل 26 العلاقة بين عمر العميل وكمية الشراء والمبلغ الإجمالي 70

فهرس الجداول

جدول 1 نتائج دراسة استخدام المحولات في توقع انسحاب عملاء بناءً على سلوك الاستخدام اليومي.....	21
جدول 2 نتائج دراسة استخدام المحولات مع تمثيلات RFM الزمنية.....	23
جدول 3 نتائج دراسة تنبؤ وتمثيل مبني على سلاسل نشاط العملاء في مواقع التجارة الالكترونية.....	25
جدول 4 أعمدة مجموعة بيانات E-commerce Customer Data For Behavior Analysis ...	27
جدول 5 أهم أعمدة مجموعة بيانات OLIST.....	28
جدول 6: حالة استخدام إنشاء حساب جديد.....	33
جدول 7: السيناريو الناجح لحالة استخدام إنشاء حساب جديد.....	33
جدول 8: حالة استخدام رفع بيانات المستخدم عن طريق ملف CSV.....	34
جدول 9: السيناريو الناجح لحالة استخدام رفع بيانات المستخدم عن طريق ملف CSV.....	35
جدول 10: حالة استخدام رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API.....	36
جدول 11: السيناريو الناجح لحالة رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API.....	36
جدول 12: حالة استخدام استعراض العملاء المتوقع انسحابها.....	37
جدول 13: السيناريو الناجح لحالة استعراض العملاء المتوقع انسحابها.....	37
جدول 14: حالة استخدام استعراض شرائح العملاء.....	38
جدول 15: السيناريو الناجح لحالة استخدام استعراض شرائح العملاء.....	38
جدول 16: حالة استخدام استعراض معلومات عن القيمة الدائمة للعملاء.....	38
جدول 17: السيناريو الناجح لحالة استخدام استعراض معلومات عن القيمة الدائمة للعملاء.....	39
جدول 18 المعاملات الفوقية للنموذج المقترح.....	42
جدول 19 نتائج مجموعة بيانات E-commerce Customer Data For Behavior Analysis ..	65
جدول 20 النتائج وفق مجموعة بيانات OLIST.....	67

الاختصارات

RFM	Recency, Frequency, Monetary Value	التكرار والتردد والقيمة النقدية
ML	Machine Learning	التعلم الآلي
SVM	Support Vector Machine	آلة المتجهات الداعمة
RF	Random Forests	الغابات العشوائية
NN	Neural Networks	الشبكات العصبية
DL	Deep Learning	التعلم العميق
RNN	Recurrent Neural Networks	الشبكات العصبية المتكررة
LSTM	Long short-term memory	الذاكرة الطويلة قصيرة المدى
SPA	Single Page Application	تطبيقات الصفحة الواحدة
CLTV	Customer Lifetime Value	القيمة الدائمة لحياة العميل
B2B	Business-to-business	عمل لعمل
CRM	customer relationship management	إدارة علاقات العملاء
SaaS	Software as a service	تقديم البرمجية كخدمة
API	Application Programming Interface	واجهة برمجة التطبيقات

الفصل الأول

التعريف بالمشروع

يتضمن هذا الفصل التعريف بالمشروع ومتطلباته.

1.1 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 الصناعي كحلول استراتيجية تمكن 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 مؤتمنة وسهلة الاستخدام، لتتيح لمواقع التجارة الإلكترونية، بمختلف أحجامها، الاستفادة من الذكاء الصناعي في تعزيز استراتيجياتها التسويقية وتحقيق نمو مستدام.

2.1 الهدف من المشروع

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

3.1 المتطلبات الوظيفية

يجب أن يقدم النظام للمستخدم ما يلي:

- 1- السماح بإنشاء حساب جديد ضمن النظام
- 2- السماح بتسجيل الدخول من حساب مُنشأ سابقاً.
- 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 الإلكترونية الخاصة به الذين يتوقع انسحابهم.
- 6- السماح للمستخدم باستعراض معلومات القيمة الدائمة للعميل لدى موقع التجارة الإلكترونية الخاص به.
- 7- السماح للمستخدم باستعراض بيانات تقسيم العملاء إلى شرائح.

4.1 المتطلبات غير الوظيفية

- 1- يجب أن يكون النظام آمناً، حيث يسمح فقط للمستخدمين المسجلين 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 نسبياً مليون عملية شراء للعملاء.

الفصل الثاني

الدراسة النظرية

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

1.2 توقع انسحاب العملاء Churn Prediction.

يعد انسحاب العملاء (Customer Churn) إحدى أبرز التحديات التي تواجه الشركات في مختلف القطاعات، لاسيما في بيئات الأعمال التنافسية مثل التجارة الإلكترونية والاتصالات. يُعد فهم أسباب انسحاب العملاء والتنبؤ به عاملاً حاسماً في صياغة استراتيجيات فعالة للاحتفاظ بالعملاء، خاصة أن تكلفة اكتساب عميل جديد تُقدَّر بأنها أعلى بعشر مرات من تكلفة الاحتفاظ بعميل حالي [1]. وقد دفع هذا المبدأ الاقتصادي العديد من الشركات إلى الاستثمار في تطوير نماذج تحليلية وتنبؤية تعتمد على تقنيات الذكاء الصناعي والتعلم الآلي، بهدف اكتشاف الأنماط السلوكية التي تسبق الانسحاب [2].

وفي سياق التجارة الإلكترونية، يكتسب تحليل انسحاب العملاء أهمية متزايدة، نظراً لاعتماد هذه المنصات بشكل كبير على ولاء العملاء وتكرار عمليات الشراء. فقد يؤدي انسحاب عدد محدود من العملاء إلى تأثير كبير على الإيرادات، خاصة مع ارتفاع تكلفة الإعلانات والاستحواذ على عملاء جدد. ولهذا، تعتمد المتاجر الإلكترونية الناجحة على تتبع سلوك العملاء في الوقت الفعلي، وتحليل بياناتهم مثل تكرار الشراء، وزمن التفاعل، وسلال التسوق المهجورة، لتحديد المؤشرات المبكرة للانسحاب واتخاذ تدابير استباقية للحد منه. من خلال تقليل معدل الانسحاب، يمكن للشركات تحسين كفاءة الإنفاق التسويقي وزيادة العائد على الاستثمار، ما يجعل من الاحتفاظ بالعملاء أولوية استراتيجية لها تأثير مباشر على الربحية والاستدامة [2].

2.2 القيمة الدائمة لحياة العميل Customer Lifetime Value

تُعد قيمة القيمة الدائمة لحياة العميل (Customer Lifetime Value - CLTV) من المفاهيم الأساسية في التسويق الحديث، وتُستخدم كأداة استراتيجية لتقدير العائد المالي المتوقع من العميل طوال فترة تعامله مع الشركة. تكتسب القيمة الدائمة للعميل أهمية متزايدة في البيئات الرقمية مثل التجارة الإلكترونية، حيث تسهم في توجيه القرارات المتعلقة بالاستحواذ على العملاء، وتخصيص الميزانيات التسويقية، وتصميم استراتيجيات الاحتفاظ [3].

3.2 الانتباه (Attention)

الانتباه هو آلية تعلم تمكن النماذج من التركيز على أجزاء معينة من بيانات الدخل عند توليد المخرجات، حيث تسمح للنموذج بوزن أهمية المدخلات المختلفة بشكل ديناميكي، مما يعزز قدرته على التقاط العلاقات والتبعيات داخل البيانات، بعض النظر عن المسافة بينها ضمن سلسلة الدخل [4].

4.2 المحولات Transformers

هي بنية شبكة عصبونية تستخدم في المهام التي تتضمن معالجة بيانات متسلسلة، مثل معالجة اللغة الطبيعية، وهي تستخدم آليات الانتباه بشكل أساسي، حيث تسمح بمعالجة بيانات الإدخال بالتوازي بدلاً من التسلسل. تُعد هذه البنية أكثر كفاءة من الشبكات التقليدية مثل RNN وLSTM، حيث تتفوق في فهم العلاقات الطويلة بين عناصر السلسلة وتُسهّل تدريب النماذج على نطاق واسع [4].

5.2 تقسيم العملاء إلى شرائح

يُعد تقسيم العملاء إلى شرائح أحد الأسس الجوهرية في علم التسويق الحديث، حيث يُعرّف بأنه عملية تصنيف قاعدة العملاء إلى مجموعات متجانسة تشترك في خصائص أو سلوكيات معينة، مثل العمر، الدخل، العادات الشرائية، أو التفاعل. يهدف هذا التقسيم إلى تمكين المؤسسات من تطوير استراتيجيات تسويقية أكثر تخصيصاً وفعالية، من خلال فهم احتياجات ودوافع كل شريحة على حدة بدلاً من التعامل مع العملاء ككتلة واحدة [5].

تتبع أهمية هذا النهج من أن العملاء لا يتصرفون بنفس الطريقة، فبعضهم أكثر ولاءً، وآخرون أكثر حساسية للسعر، وهناك من يركز على القيمة أو الخدمة. ولذلك، يتيح تقسيم العملاء للشركات إمكانية تصميم حملات موجهة، وتخصيص العروض والخدمات، وتحسين تجربة العميل، مما يؤدي إلى زيادة معدلات الاحتفاظ والربحية.

وقد تنوعت أساليب التقسيم ما بين التقسيم الديموغرافي (العمر، الجنس، الدخل)، والسلوكي (نمط الشراء، التكرار، الولاء)، والنفسي (الأسلوب الحياتي، القيم)، وصولاً إلى التقسيم باستخدام التحليل الكمي وعلوم البيانات، مثل خوارزميات التعلم الآلي، التي تسمح بتحليل قواعد بيانات ضخمة واكتشاف أنماط معقدة من العلاقات بين السمات والسلوك [5].

6.2 خوارزمية K-means

تُعد خوارزمية K-Means من أكثر خوارزميات التجميع (Clustering) استخداماً في تعلم الآلة، وخصوصاً في تحليل البيانات الاستكشافية وتقسيم العملاء في أنظمة دعم القرار. تهدف هذه الخوارزمية إلى تقسيم مجموعة من البيانات إلى K عنقود غير متداخلين، بحيث تكون البيانات ضمن كل عنقود متقاربة قدر الإمكان، وتكون المسافة إلى مركز العنقود (Centroid) أقل ما يمكن.

تبدأ الخوارزمية بتحديد عدد العناقيد K واختيار مراكز أولية عشوائية لها، ثم تتكرر خطوتان رئيسيتان حتى يحدث الاستقرار:

- إسناد كل نقطة بيانات إلى أقرب مركز عنقود باستخدام مسافة إقليدية.
- تحديث مراكز العناقيد كمتوسط للنقاط المنتمية لها.

تستمر العملية إلى أن تتوقف مراكز العناقيد عن التغير أو يتم الوصول إلى عدد محدد من التكرارات.

تُستخدم K-Means في تطبيقات متنوعة تشمل: تحليل شرائح العملاء، تجزئة الصور، تصنيف المستندات، واكتشاف الأنماط السلوكية. وتُعد مناسبة بشكل خاص للبيانات العددية القابلة للقياس [5].

7.2 نموذج RFM

نموذج RFM (Recency, Frequency, Monetary) هو أداة تحليلية تُستخدم لفهم سلوك العملاء في التسويق، حيث يتم تقسيم العملاء استناداً إلى ثلاثة معايير رئيسية: الحداثة (كم من الوقت مضى منذ آخر شراء)، التكرار (عدد مرات الشراء)، والقيمة النقدية (إجمالي المبالغ المنفقة). يساعد هذا النموذج الشركات على استهداف العملاء بشكل أكثر فاعلية، وتحسين الحملات التسويقية، وزيادة العوائد على الاستثمار عن طريق تخصيص العروض بناءً على سلوك العملاء. رغم فعاليته، يتطلب النموذج تحديثاً دورياً للبيانات لضمان دقة التصنيفات [6].

الفصل الثالث

الدراسة المرجعية

يعرض هذا الفصل الأبحاث والبيانات المرتبطة بالعمل المقدم.

1.3 مقدمة

إن مجال تحليل بيانات عملاء مواقع التجارة الإلكترونية هو مجال واسع ويحتوي على العديد من التحليل 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 التحليل والقياسات التي استخرجت من الأبحاث ضمن هذا المجال.

2.3 توقع انسحاب العملاء Churn Prediction.

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

في توقع انسحاب العملاء، غالبا ما تشمل الطرق التقليدية النماذج الإحصائية والخوارزميات مثل نموذج 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 على الطرق التقليدية إلا أنها غالبا ما تكون غير قادرة على نمذجة الأنماط الزمنية أو نمذجة الأنماط التتابعية في بيانات العملاء [7].

ومع التطورات في مجال التعلم العميق Deep Learning أظهرت الشبكات العصبية التكرارية (RNNs) Recurrent Neural Networks والمحولات 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 [8].

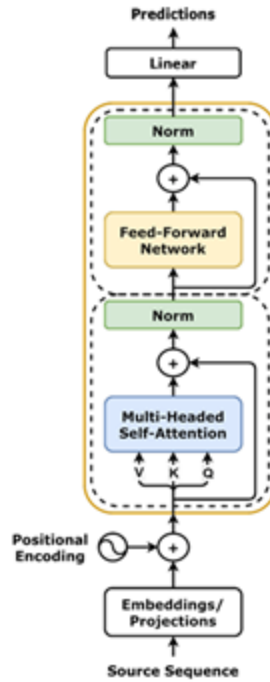
2.2.3 استخدام المحولات في توقع انسحاب عملاء بناءً على سلوك الاستخدام اليومي.

جرى استخدام المحول في قطاع B2B business-to-business، 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 [9].

1.2.2.3 طبيعة البيانات المستخدمة.

هي بيانات يومية على مستوى الاشتراك مثل: طول فترة التعاقد، عدد المكالمات، استهلاك البيانات، وعدد الباقات [9].

2.2.2.3 بنية النموذج.



الشكل 1 المحولات في توقع انسحاب العملاء بناءً على سلوك الاستخدام اليومي [9]

يعتمد النموذج الموضح في الشكل 1 على بنية المحول (Transformer) لمعالجة البيانات اليومية للمشاركين مثل مدة التعاقد، عدد المكالمات، استهلاك البيانات، وعدد الباقات. تبدأ العملية بتحويل هذه البيانات إلى تمثيلات عددية، ثم تُضاف ترميزات موضعية لتحديد ترتيب الأيام. بعد ذلك، تمر البيانات عبر طبقة الانتباه الذاتي متعدد الرؤوس التي تُمكن النموذج من فهم العلاقات بين الأيام المختلفة، تليها طبقات Normalization وشبكة تغذية أمامية لمعالجة المعلومات بشكل أعمق. في النهاية، تُستخدم طبقة خطية لإنتاج التنبؤات، مثل احتمال إنهاء الاشتراك أو توقع الاستخدام المستقبلي [9].

3.2.2.3 النتائج

تم تدريب النموذج لمدة 15 دورة (epochs) باستخدام معدل تعلّم قدره 0.001، واستغرق التدريب أقل من ساعة. اعتمدت مجموعة التدريب على توازن الصفوف من خلال تقليل عدد العينات السالبة (Under Sampling)، بينما تُركت مجموعة التحقق دون توازن لتمثل سيناريو واقعي، حيث يشكل العملاء المنسحبون حوالي 1.3% فقط. أظهرت النتائج تبايناً كبيراً في الأداء بين التدريب والتحقق (validation)، حيث انخفضت دقة F1 من 0.935 في التدريب إلى 0.286 في التحقق. كما انخفضت الدقة (Precision) من 0.923 إلى 0.217، في حين بلغت قيمة AUROC على مجموعة التحقق 0.766، مما يشير إلى تحديات واضحة عند التعامل مع بيانات غير متوازنة [9].

التحقق Validation	التدريب Training	مقياس التقييم
0.286	0.935	F1-score
0.217	0.923	Precision
0.419	0.964	Recall
0.766	0.982	AUROC

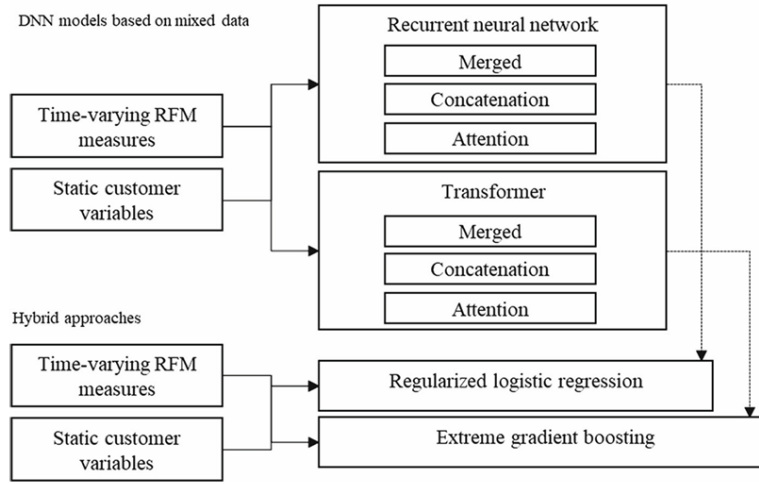
جدول 1 نتائج دراسة استخدام المحولات في توقع انسحاب عملاء بناءً على سلوك الاستخدام اليومي [9]

3.2.3 استخدام المحولات مع تمثيلات RFM الزمنية.

1.3.2.3 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 [8].

2.3.2.3 بنية النموذج.



الشكل 2 نموذج المحولات مع تمثيلات RFM الزمنية [8]

تم بتطبيق نماذج Transformers و RNN على مقاييس RFM المتغيرة عبر الزمن، بهدف تحسين التنبؤ بالانسحاب العملاء في قطاع الخدمات المالية. يُعدّ هذا النهج مهماً لأن مقاييس RFM تعتبر من المؤشرات الأساسية لتقييم سلوك العملاء وقدرةهم على التنبؤ بالانسحاب [8].

3.3.2.3 النتائج

في هذه الدراسة، تم تقييم فعالية نماذج الشبكات العصبية العميقة (DNN) في التنبؤ بالانسحاب (Churn Prediction) باستخدام مقاييس RFM المتغيرة زمنياً. أظهرت النتائج أن نماذج RNN تتفوق على نماذج Transformer عند التعامل مع هذه البيانات، وذلك من خلال مؤشرات رئيسية مثل: مساحة منحنى AUC، حيث

بلغ متوسط AUC لنموذج RNN مع دمج المتغيرات الثابتة 0.8146، مقارنة بـ 0.7761 لنموذج Transformer المماثل. كما دعمت الاختبارات الإحصائية هذه النتائج، حيث أظهرت فروقاً معنوية بين النموذجين على جميع المقاييس. كذلك، تم اختبار عدة أشكال من دمج البيانات (كالدمج المباشر، الانتباه، والانتباه متعدد الرؤوس)، وأظهرت النماذج الهجينة (مثل XGBoost مع خصائص RNN) أفضل أداءً إجماليًا، حيث حققت أعلى AUC (0.8428). توضح هذه النتائج أن اختيار المعمارية المناسبة للنموذج يلعب دوراً حاسماً في دقة التنبؤ في سياق انسحاب العملاء [8].

AUC	نوع البنية	الخوارزمية	مجموعة الميزات
0.7671	–	RNN	Time-varying RFM
0.7130	–	Transformer	
0.8013	Merged	RNN	Time-varying RFM and static variables
0.8146	Concatenation		
0.8084	Attention		
0.7567	Concatenation	Transformers	
0.7761	Merged		
0.7839	Attention		
0.7854	Multi-head Attention		

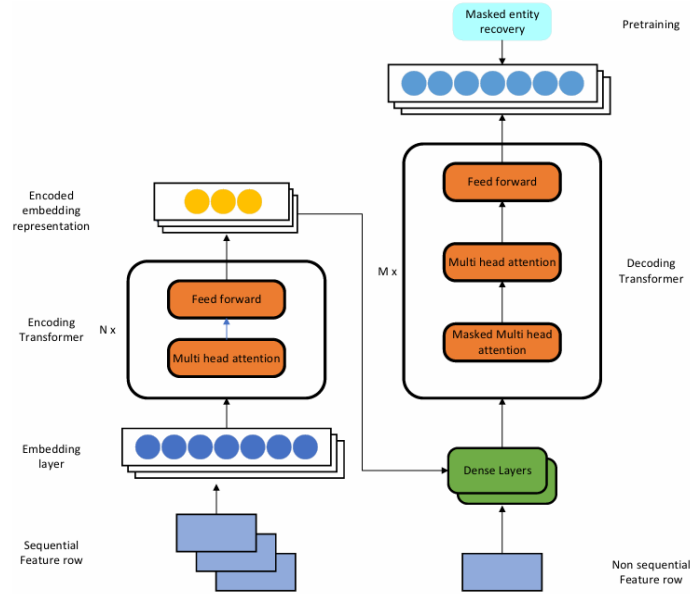
جدول 2 نتائج دراسة استخدام المحولات مع تمثيلات RFM الزمنية [8]

4.2.3 تنبؤ وتمثيل مبني على سلاسل نشاط العملاء في مواقع التجارة الالكترونية نموذج CASPR.

1.4.2.3 طبيعة البيانات المستخدمة.

بيانات سجلات المشتريات أو الجلسات من مواقع مثل Microsoft Retail Stores أو Instacart [1].

2.4.2.3 بنية النموذج



الشكل 3 نموذج CASPR [1]

تم إعادة توظيف بنية المحول (Transformer) المستخدمة في مهام معالجة اللغة الطبيعية لتعمل على بيانات نشاط العملاء المتسلسلة. يتمثل مدخل النموذج في سلسلة من المتجهات تمثل كل نشاط للعميل عند كل نقطة زمنية، حيث يتكون من دمج السمات الرقمية مع تمثيلات كثيفة للسمات التصنيفية الناتجة عن طبقة تضمين. يستخدم المحول آلية الانتباه الذاتي متعدد الرؤوس (Multi-Head Attention) مع شبكة عصبية أمامية مكونة من طبقتين لكل رأس. بعد الترميز، يتم دمج المخرجات مع سمات غير زمنية (مثل العمر والحالة وفترة الاشتراك) قبل تمريرها إلى طبقات كثيفة إضافية. يتضمن الجزء الخاص بالمفسر (Decoder) آلية انتباه مقنّع (Masked Attention) لضمان خاصية التوليد التتابعي، ما يُحسن فعالية النموذج في مهام مثل توصية المنتجات [1].

3.4.2.3 النتائج

تُظهر الدراسة أن نموذج CASPR يحقق تحسينات ملحوظة في الأداء التنبؤي عبر مجموعة متنوعة من المهام مثل توقع انسحاب العملاء (churn)، تقدير القيمة الدائمة للعميل (CLTV)، اكتشاف الحسابات الاحتيالية، وترتيب توصيات المنتجات. في مهام التصنيف، تم استخدام AUROC كمؤشر رئيسي للمقارنة، إضافةً إلى F1-score في حالة توقع الانسحاب. أظهرت نتائج التجارب أن CASPR حسّن أداء نموذج الأساس في توقع انسحاب العملاء بنسبة 2-3 نقاط في AUROC، مع تحسّن أكبر في F1-score خصوصاً على مجموعة بيانات KKBox (من 0.27 إلى 0.44). أما على مجموعة Microsoft Retail، فقد ارتفعت قيمة AUROC من 0.761 إلى 0.794 عند استخدام خصائص RFM. في المقابل، لم يحقق CASPR تحسّناً ملحوظاً على بيانات Google Online Stores بسبب قصر سجل النشاطات (متوسط 1.5 تفاعل لكل عميل)، ما حدّد من قدرته على تعلم الأنماط السلوكية. تُبرز هذه النتائج قدرة CASPR على التعلّم التتابعي وتحسين الأداء في السيناريوهات التي تتوفر فيها بيانات سلوكية كافية [1].

F1-Score	AUROC	مجموعة البيانات
0.44	0.91	KKBox
0.96	0.903	GoogleOnlineStores
0.837	0.777	MicrosoftRetailStores

جدول 3: نتائج دراسة تنبؤ وتمثيل مبني على سلاسل نشاط العملاء في مواقع التجارة الإلكترونية [1]

3.3 استخدام خوارزمية K-Means في تقسيم عملاء متجر تجارة إلكترونية.

يعد تقسيم العملاء إلى شرائح (Customer Segmentation) من أهم التطبيقات العملية لخوارزمية K-Means في بيئة التجارة الإلكترونية. تهدف هذه العملية إلى تقسيم قاعدة العملاء إلى شرائح متجانسة بناءً على سلوكهم الشرائي وخصائصهم، مما يساعد المتاجر الإلكترونية على اتخاذ قرارات تسويقية مخصصة، وتحسين تجربة المستخدم، وزيادة ولاء العملاء. تعتمد خوارزمية K-Means على تجميع العملاء وفقاً لسمات محددة مثل: عدد الطلبات، إجمالي الإنفاق، معدل التكرار، والمدة الزمنية منذ آخر عملية شراء. يتم تمثيل كل عميل كنقطة في فضاء متعدد الأبعاد، وتُقسم هذه النقاط إلى عناقيد باستخدام مبدأ تقليل المسافة بين العملاء ومراكز العناقيد. تساعد هذه الطريقة في تحديد أنماط مثل "عملاء مميزون"، "عملاء جدد"، أو "عملاء مهددون بالانسحاب"، وهو ما يمكّن الشركات من تخصيص استراتيجيات التسويق والعروض بشكل أكثر دقة [5].

وقد أظهرت فعالية K-Means نظراً لبساطتها وسرعتها في التعامل مع مجموعات بيانات كبيرة، مما يجعلها مناسبة للتطبيقات العملية في المتاجر الإلكترونية. كما يمكن دمج نتائج التقسيم مع تقنيات أخرى، مثل التوصية التنبؤية أو توقع الانسحاب، لتحقيق أداء أفضل في إدارة العلاقة مع العملاء [5].

4.3 مجموعات البيانات في مجال التجارة الإلكترونية

مجموعات البيانات في الأوراق البحثية السابقة كانت جميعها غير متوفرة للعامة. حيث أنها كانت مجموعات بيانات مقدمة للباحثين وفق اتفاقيات مع الشركات محمية من النشر تحت بند حماية بيانات المستخدمين. ولذلك لجأت الحاجة إلى البحث عن مجموعات بيانات في مجال التجارة الإلكترونية وكانت أهم النتائج هي:

1.4.3 مجموعة بيانات E-commerce Customer Data For Behavior Analysis

مجموعة البيانات هذه تتكون من 250000 عملية شراء للعملاء تخدم أغراض التحليل الإحصائي والنمذجة التنبؤية في مجال التجارة الإلكترونية، تحتوي البيانات على مجموعة عمليات شراء لعدة مستخدمين، وهي تناسب تطبيقات متعددة مثل [10]:

- التنبؤ بانسحاب العملاء (Churn Prediction).
- تحليل سلال التسوق (Market Basket Analysis).
- أنظمة التوصية (Recommendation Systems).

حيث أن كل سطر في قاعدة البيانات هذه هو عبارة عن عملية شراء قام بها أحد المستخدمين تضمن المعلومات عملية الشراء.

1.1.4.3 أعمدة قاعدة البيانات.

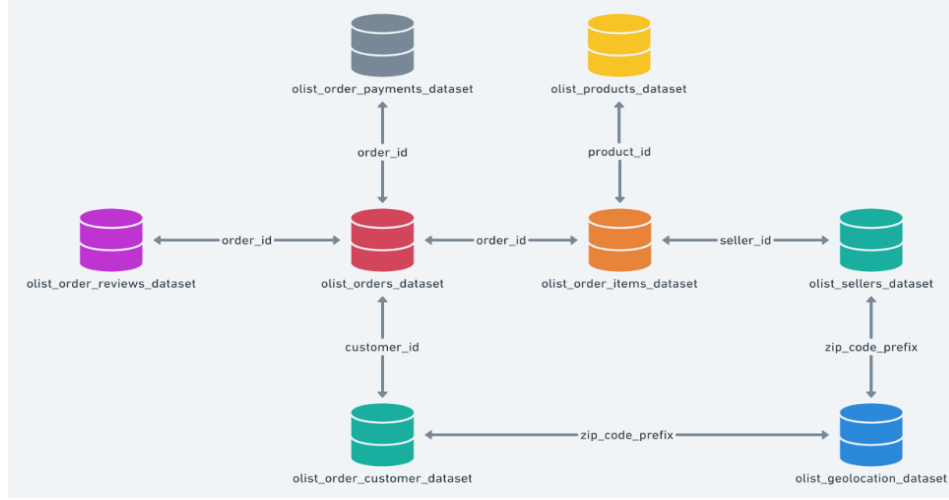
الاسم	وصف	نوع البيانات
Customer ID	رقم تعريفى لكل عميل	Integer
Customer Name	اسم العميل	String
Customer Age	عمر العميل	String
Gender	جنس العميل	String (Male,Female)
Purchase Date	تاريخ الشراء	Time stamp
Product Category	تصنيف المنتج الذي تم شراؤه	String
Product Price	السعر الفردي للمنتج	Integer
Quantity	عدد الوحدات المشتراة من المنتج	Integer
Total Purchase Amount	إجمالي مبلغ الشراء الذي تم دفعه في العملية	Integer
Payment Method	وسيلة الدفع المستخدمة	String
Returns	مؤشر يدل على إذا ما تم ارجاع المنتج أم لا	Boolean
Churn	مؤشر يدل على ما إذا كان العميل قد انسحب أم لا	Boolean

جدول 4 أعمدة مجموعة بيانات E-commerce Customer Data For Behavior Analysis [10]

2.4.3 مجموعة بيانات Brazilian E-Commerce Public Dataset by OLIST

مجموعة البيانات هذه هي قاعدة بيانات واقعية تتضمن معلومات تفصيلية حول أكثر من 100 ألف طلب شراء تم تنفيذها عبر متجر OLIST الإلكتروني والتي تعد من أكبر المتاجر ضمن الأسواق البرازيلية [11].

قسمت البيانات إلى عدة مجموعات لتسهيل الفهم والتنظيم حيث كانت البنية كما يلي:



الشكل 4 مجموعة بيانات OLIST [11]

1.2.4.3 أعمدة مجموعة البيانات

قاعدة البيانات تحتوي على عدة أعمدة في مختلف المجموعات سنذكر أهمها:

نوع البيانات	وصف	الاسم
Integer	رقم تعريفى لكل عميل	customer_unique_id
Time stamp	تاريخ الشراء	order_purchase_timestamp
String	تصنيف المنتج الذي تم شراؤه	product_category_name_english
Integer	السعر الفردي للمنتج	Price
Integer	الرسوم اللوجستية (سعر الشحن)	freight_value
String	وسيلة الدفع المستخدمة	payment_type

جدول 5 أهم أعمدة مجموعة بيانات OLIST [11]

5.3 دراسة التطبيقات المشابهة

Baremetrics 1.5.3

تُعد منصة Baremetrics من الأدوات المتقدمة في مجال تحليل بيانات الاشتراكات لمواقع ومنصات التجارة الإلكترونية، حيث تقدم لوحات تحكم تفاعلية تعرض مؤشرات مالية وتشغيلية مهمة مثل الإيرادات الشهرية المتكررة، معدل انسحاب العملاء، والقيمة الدائمة للعميل. توفر المنصة إمكانيات واسعة لتقسيم العملاء إلى شرائح بناءً على معايير متعددة مثل تاريخ الاشتراك، حجم الاستخدام، أو نوع الباقة، مما يساعد في فهم سلوك العملاء وتوجيه الحملات التسويقية بشكل أكثر دقة. كما تقوم منصة Baremetrics بحساب القيمة الدائمة للعميل باستخدام بيانات الأداء المالي للمشاركين، وتُصدر تنبيهات تلقائية عند حدوث تغيرات مفاجئة في معدلات الانسحاب أو الإيرادات، مما يمكن أصحاب المنصات من اتخاذ قرارات سريعة ومبنية على بيانات لحظية. تعتمد المنصة على التكامل مع بوابات الدفع مثل Stripe وShopify، وتتيح سحب البيانات وتحليلها عبر واجهات برمجة التطبيقات (APIs) بسهولة، ما يجعلها أداة فعالة لمراقبة وتحسين أداء الأعمال القائمة على الاشتراكات.

Churnly 2.5.3

تُعد منصة Churnly من الحلول المتقدمة المدعومة بالذكاء الاصطناعي، المصممة خصيصاً للشركات العاملة بنظام SaaS، وتركز على التنبؤ بانسحاب العملاء عبر تحليل بياناتهم وسلوكهم على طول رحلة الاستخدام. تعتمد المنصة على خوارزميات تعلم آلي متقدمة لتقييم عوامل متعددة مثل التفاعل مع المنتج، تاريخ الفواتير، تذاكر الدعم، ومعدلات التجديد، ومن ثمّ تعيين درجة خطورة الانسحاب لكل عميل، بدقة توقع تتراوح بين 92% و 96%. توفر Churnly لوحة تحكم تفاعلية تساعد فرق نجاح العملاء في تحديد العائد المعرض للخسارة، وفهم أسباب الانسحاب المحتملة، إضافة إلى إمكانيات متقدمة لتقسيم العملاء حسب معايير مثل مستوى المخاطرة، مدة الاستخدام، وتفاعل العملاء. كما تربط المنصة أنظمة إدارة علاقات المستخدمين CRM وفوترة ودعم مثل Zendesk وHubSpot وZoho لاستيراد البيانات وإنشاء صورة متكاملة لسلوك العملاء. توفر Churnly أداء في الوقت الفعلي ينبه عند تغير مؤشرات النجاح أو الانسحاب، مما يمكن اتخاذ قرارات استباقية لتعزيز القيمة الدائمة للعميل (CLTV) وتحسين معدلات الاحتفاظ.

Pecan AI 3.5.3

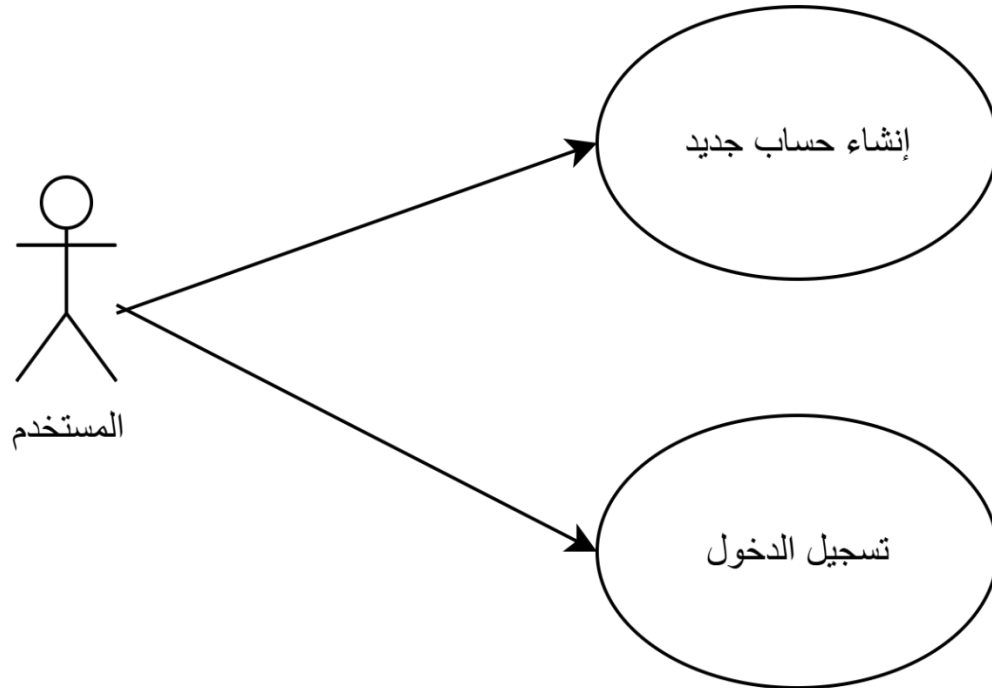
تُعدّ منصة **Pecan AI** منصة تحليل تنبؤية مدعومة بالذكاء الصناعي تتيح لمستخدمي الأعمال (مثل التسويق، التحليلات، وفرق العمليات) بناء نماذج تعلم آلي دون الحاجة إلى خبرة برمجية متقدمة. تعتمد المنصة على واجهة بديهية منخفضة الحاجة إلى رماز، وتقوم بعمليات تجهيز البيانات، وهندسة السمات، واختيار النموذج وتحسينه، ومن ثم توليد التنبؤات في وقت قياسي (أيام أو أسابيع). تستخدم Pecan بيانات العملاء التاريخية (مثل التفاعل مع المنتج، سجل الفواتير، الاستخدام، وتذاكر الدعم) لبناء نماذج تعلم تقيّم احتمالية انسحاب كل عميل. يتم تصنيف العملاء ضمن شرائح "مرتفع/متوسط/منخفض المخاطرة" بدقة تنبؤية عالية، مع القدرة أيضاً على تفسير العوامل التي تؤثر على الانسحاب كما توفر المنصة إمكانيات متقدمة في تقسيم العملاء بناءً على التنبؤ بسلوكهم المستقبلي مثل قيمة الدائمة للعميل وأنماط التفاعل، مما يسهل توجيه حملات تسويقية واستباقية مخصصة لكل شريحة. كما تدعم خاصية الـ Explainable AI التي تشرح للمستخدمين العوامل المؤثرة في التنبؤات، مما يعزز الثقة في النتائج ويُسهم في اتخاذ قرارات غير عشوائية.

الفصل الرابع

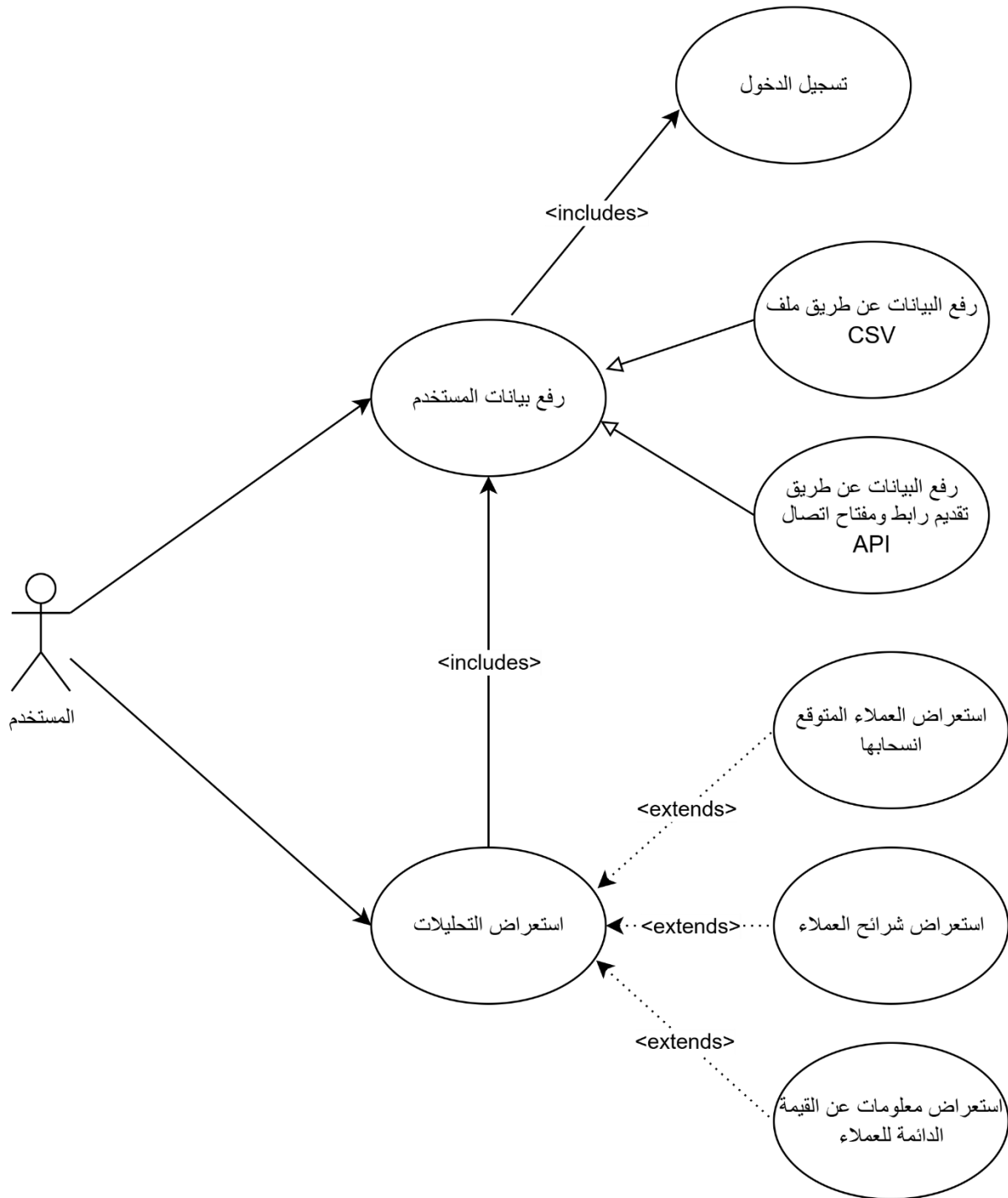
الدراسة التحليلية

يوضح هذا الفصل عملية تحليل النظام ودراسة متطلباته.

1.4 مخططات حالات الاستخدام



الشكل 1 مخطط حالات استخدام إدارة حسابات المستخدمين



الشكل 2 مخطط استخدام إدارة عمليات النظام

2.4 السرد النصي لحالات الاستخدام

1.2.4 إنشاء حساب جديد.

الوصف Description	اسم الحالة: إنشاء حساب جديد
يقوم المستخدم بإنشاء حساب جديد خاص به.	
الفاعلين Actors	المستخدم
الشروط السابقة Preconditions	لا يوجد.
الشروط اللاحقة Postconditions	تم إنشاء حساب جديد خاص بالمستخدم ضمن النظام.

جدول 6: حالة استخدام إنشاء حساب جديد.

سير الأحداث

السيناريو الأساسي الناجح:

النظام	المستخدم
1. يطلب إنشاء مشروع جديد	
2. يطلب النظام المعلومات التالية:	
• اسم المستخدم	
• البريد الإلكتروني الخاص بالمستخدم	
• كلمة المرور	
3. يحدد المستخدم المعلومات المطلوبة ويطلب تأكيد العملية	
4. يتحقق النظام من صحة المعلومات ويعيد للمستخدم رسالة توضح انتهاء العملية.	

جدول 7: السيناريو الناجح لحالة استخدام إنشاء حساب جديد.

المسارات البديلة:

لا يوجد

مسارات الأخطاء:

E1: في المرحلة رقم 2 في حال كان:

- اسم المستخدم غير صالح
- البريد الإلكتروني غير صالح
- كلمة المرور ضعيفة

في حال تحققت إحدى الحالات السابقة، يتم استبدال الخطوة الثانية في السيناريو الأساسي بالخطوة التالية:

2. يعيد النظام رسالة توضح سبب الخطأ ويطلب تحديد المعلومات من جديد.

E2: في المرحلة رقم 4 في حال كان:

- اسم المستخدم مسجل سابقاً ضمن النظام.
- البريد الإلكتروني مسجل سابقاً ضمن النظام.

في حال تحققت إحدى الحالات السابقة، يتم استبدال الخطوة الثانية في السيناريو الأساسي بالخطوة التالية:

4. يعيد النظام رسالة توضح سبب الخطأ ويطلب تحديد المعلومات من جديد.

2.2.4 رفع بيانات المستخدم عن طريق ملف csv.

اسم الحالة: رفع بيانات المستخدم عن طريق ملف csv	
الوصف Description	يقوم المستخدم برفع البيانات الخاصة به عن طريق إرفاق ملف بصيغة CSV.
الفاعلين Actors	المستخدم
الشروط السابقة Preconditions	المستخدم مسجل في النظام سابقاً.
الشروط اللاحقة Postconditions	تم رفع بيانات المستخدم وحفظها ضمن النظام.

جدول 8: حالة استخدام رفع بيانات المستخدم عن طريق ملف CSV

سير الأحداث

السيناريو الأساسي الناجح:

النظام	المستخدم
1. يطلب رفع بياناته خاصة.	
2. يطلب النظام مسار الملف المراد رفعه.	
3. يحدد المعلومات المطلوبة ويطلب تأكيد العملية.	
4. يتحقق النظام من صحة المعلومات ويعيد للمستخدم رسالة توضح رفع الملف بنجاح.	
5. يطلب النظام ربط الأعمدة الخاصة ببيانات المستخدم بالأعمدة المعرفة ضمن النظام.	
6. يحدد ربط الأعمدة ويطلب تأكيد العملية.	
7. يتحقق النظام من صحة ربط الأعمدة ويعيد للمستخدم رسالة توضح صحة ربط الأعمدة.	
8. يطلب النظام من المستخدم تأكيد رفع البيانات.	
9. يطلب المستخدم تأكيد العملية.	
10. يتم رفع بيانات المستخدم ويعيد رسالة توضح انتهاء العملية.	

جدول 9: السيناريو الناجح لحالة استخدام رفع بيانات المستخدم عن طريق ملف CSV.

المسارات البديلة:

لا يوجد

مسارات الأخطاء:

E1: في المرحلة رقم 2 في حال كان الملف غير موجود في المسار المشار إليه او الملف غير صالح.

في حال تحققت إحدى الحالات السابقة، يتم استبدال الخطوة الثانية في السيناريو الأساسي بالخطوة التالية:

4. يعيد النظام رسالة توضح سبب الخطأ ويطلب تحديد المسار من جديد.

E2: في المرحلة رقم 6 في حال لم يحدد المستخدم أحد الأعمدة المطلوب ربطها. يتم استبدال الخطوة السادسة في السيناريو الأساسي بالخطوة التالية:

5. يعد النظام رسالة توضح سبب الخطأ ويطلب تحديد ربط الأعمدة من جديد.

3.2.4 رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API.

الوصف Description	اسم الحالة: رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API
يقوم المستخدم برفع البيانات الخاصة به عن طريق تقديم رابط إلى نقطة اتصال API endpoint في نظامه ومفتاح اتصال.	
الفاعلين Actors	المستخدم
الشروط السابقة Preconditions	المستخدم مسجل في النظام سابقاً.
الشروط اللاحقة Postconditions	تم رفع بيانات المستخدم وحفظها ضمن النظام.

جدول 10: حالة استخدام رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API.

سير الأحداث

السيناريو الأساسي الناجح:

النظام	المستخدم
	1. يطلب رفع بياناته الخاصة.
2. يطلب النظام المعلومات التالية:	
• رابط نقطة الاتصال الخاصة بالمستخدم	
• مفتاح الاتصال الخاص بنقطة الاتصال	
	3. يحدد المستخدم المعلومات المطلوبة ويطلب تأكيد العملية
4. يتحقق النظام من صحة المعلومات ويقوم بطلب البيانات من نقطة الاتصال وتخزينها ضمن النظام ويعيد للمستخدم رسالة توضح انتهاء العملية.	

جدول 11: السيناريو الناجح لحالة رفع البيانات عن طريق تقديم رابط ومفتاح اتصال API.

المسارات البديلة:

لا يوجد

مسارات الأخطاء:

E1: في المرحلة رقم 4 في حال كان:

- رابط نقطة الاتصال غير صالح.
- مفتاح الاتصال غير صحيح.

يتم استبدال الخطوة الرابعة في السيناريو الأساسي بالخطوة التالية:

4. يعيد النظام رسالة توضّح سبب الخطأ ويطلب تحديد المعلومات من جديد.

4.2.4 استعراض العملاء المتوقع انسحابها.

الوصف Description	اسم الحالة: استعراض العملاء المتوقع انسحابها
يقوم المستخدم باستعراض العملاء المتوقع انسحابها ومعلومات تحليلية عنها.	
المستخدم	الفاعلين Actors
المستخدم مسجل في النظام سابقاً، وقد رفع بياناته سابقاً إلى النظام.	الشروط السابقة Preconditions
تم عرض النتائج الخاصة بالعملاء المتوقع انسحابها.	الشروط اللاحقة Postconditions

جدول 12: حالة استخدام استعراض العملاء المتوقع انسحابها

سير الأحداث

السيناريو الأساسي الناجح:

النظام	المستخدم
1. يطلب استعراض العملاء المتوقع انسحابها.	
2. يقوم النظام بتحليل بيانات العملاء وعرض العملاء المتوقع انسحابها.	

جدول 13: السيناريو الناجح لحالة استعراض العملاء المتوقع انسحابها.

المسارات البديلة:

لا يوجد.

مسارات الأخطاء:

لا يوجد.

5.2.4 استعراض شرائح العملاء.

اسم الحالة: استعراض شرائح العملاء	
الوصف Description	يقوم المستخدم باستعراض شرائح العملاء.
الفاعلين Actors	المستخدم
الشروط السابقة Preconditions	المستخدم مسجل في النظام سابقاً، وقد رفع بياناته سابقاً إلى النظام.
الشروط اللاحقة Postconditions	تم عرض النتائج الخاصة بشرائح العملاء.

جدول 14: حالة استخدام استعراض شرائح العملاء

سير الأحداث

السيناريو الأساسي الناجح:

النظام	المستخدم
	3. يطلب استعراض شرائح العملاء.
	4. يقوم النظام بتحليل بيانات العملاء وتقسيمهم إلى شرائح وعرضها.

جدول 15: السيناريو الناجح لحالة استخدام استعراض شرائح العملاء.

المسارات البديلة:

لا يوجد.

مسارات الأخطاء:

لا يوجد.

6.2.4 استعراض معلومات عن القيمة الدائمة للعملاء.

اسم الحالة: استعراض معلومات عن القيمة الدائمة للعملاء.	
الوصف Description	يقوم المستخدم باستعراض معلومات عن القيمة الدائمة للعملاء.
الفاعلين Actors	المستخدم
الشروط السابقة Preconditions	المستخدم مسجل في النظام سابقاً، وقد رفع بياناته سابقاً إلى النظام.
الشروط اللاحقة Postconditions	تم عرض النتائج الخاصة بالقيمة الدائمة للعملاء.

جدول 16: حالة استخدام استعراض معلومات عن القيمة الدائمة للعملاء

سير الأحداث

السيناريو الأساسي الناجح:

المستخدم	النظام
5. يطلب استعراض معلومات عن القيمة الدائمة للعملاء.	
	6. يقوم النظام بتحليل بيانات العملاء وعرضها.

جدول 17: السيناريو الناجح لحالة استخدام استعراض معلومات عن القيمة الدائمة للعملاء.

المسارات البديلة:

لا يوجد.

مسارات الأخطاء:

لا يوجد.

الفصل الخامس

المنهجية المقترحة

يعرض هذا الفصل المنهجية المقترحة ضمن العمل المقدم.

1.5 مقدمة

بعد دراسة التطبيقات المشابهة تم استخلاص الميزات الأساسية التي سيتم تقديمها في التطبيق وكانت:

- توقع انسحاب العملاء.
- تقسيم العملاء إلى شرائح.
- حساب القيمة الدائمة للعملاء.

سنذكر في الفقرات القادمة المنهجية المقترحة لكل خدمة.

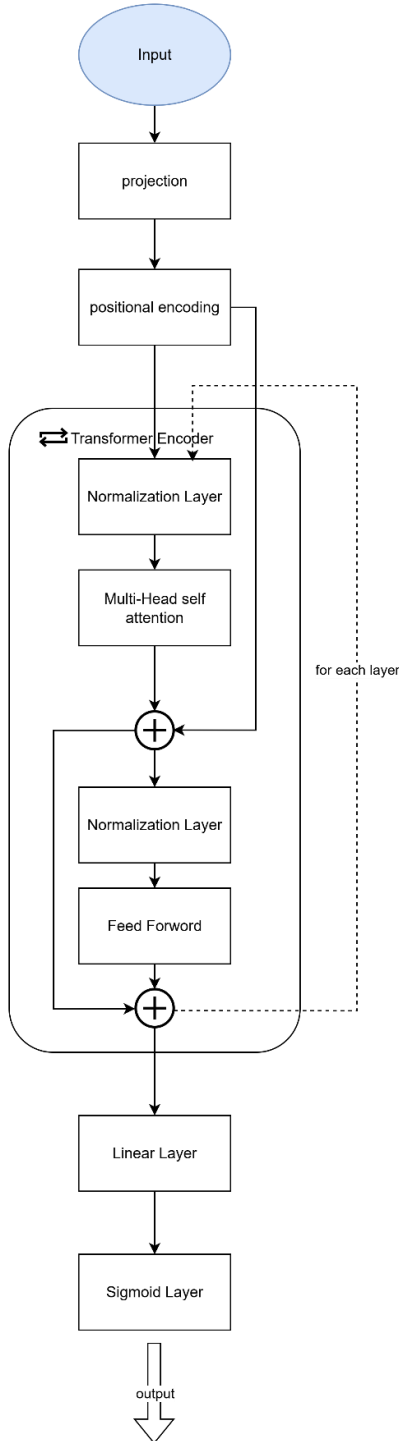
2.5 النموذج المقترح لتوقع انسحاب العملاء

بعد النظر إلى الأبحاث في مجال توقع انسحاب العملاء جرى اتباع منهجية بناء نموذج مبني على المحولات (Transformers) وتدريبه على عمليات شراء العملاء للتوقع بانسحاب العملاء نظرا لقدرة المحولات على فهم سلسلة عمليات وتمثيلها.

بدايةً نعرف دخل النموذج، حيث يتكون الدخل من مجموعة عمليات شراء للعملاء تحتوي على معلومات عن كل عملية شراء ومعلومات عن العميل حيث كانت:

- معلومات عن كل عملية شراء مثل: تاريخ عملية الشراء، الكمية، طريقة الدفع، المبلغ الإجمالي المدفوع، نوع المنتج ضمن عملية الشراء، سعر المنتج الفردي، هل تم إعادة المنتج.
- معلومات عن العميل مثل: عمر العميل وجنس العميل.

1.2.5 بنية النموذج المقترح لتوقع انسحاب العملاء



الشكل 5 بنية النموذج المقترح لتوقع انسحاب العملاء

النموذج المقترح يكون من طبقتين من المحولات وكل طبقة تحتوي على 8 رؤوس من الانتباه الذاتي (Self-attention).

حيث كانت المعاملات الفوقية (Hyper-Parameters) هي:

المعامل	القيمة
الطبقة العرض projection	128
عدد الرؤوس Number of heads	8
عدد طبقات المحول	2
Number of Transformers level	
طول سلسلة العمليات Sequence Length	10 * 14
معامل ترك التعلم Dropout	0.1

جدول 18 المعاملات الفوقية للنموذج المقترح

يتكون النموذج المقترح لتوقع انسحاب العملاء من سلسلة من المكونات المترابطة التي تعالج سلوك المستخدم كسياق زمني متسلسل.

1. تبدأ العملية بمرحلة المدخلات (Input) التي تمثل سلسلة عمليات الشراء للعميل. حيث تعبر بيانات الدخل عن مصفوفة بطول 10×14 حيث 10 هو طول سلسلة عمليات الشراء و14 هو عدد الميزات المرتبطة بكل عملية شراء.
2. تليها طبقة الإسقاط الخطي (Projection layer) والتي تقوم بتحويل كل عملية إلى تمثيل شعاعي ثابت البعد ليتوافق مع متطلبات دخل طبقة المحول.
3. بعدها يتم دمج تمثيل الموضع (Positional Encoding) مع التمثيلات الشعاعية لإعطاء النموذج القدرة على التمييز بين ترتيب الأحداث، وهي خطوة مهمة لأن بنية المحول لا تتعامل تلقائياً مع الترتيب الزمني.
4. ثم تُمرر البيانات إلى المحول الذي يتكرر على طبقتين. كل طبقة تتكون من:
 - a. طبقة (Normalization) لتثبيت التوزع الإحصائي.
 - b. آلية الانتباه الذاتي متعدد الرؤوس وهي العنصر الأساسي الذي يسمح للنموذج بفهم العلاقات بين جميع عمليات الشراء ضمن سلسلة الدخل، بما في ذلك العلاقات بعيدة المدى. وهو يتكون من ثماني رؤوس انتباه.
 - c. وصلات التجاوز (residual connections) تتجاوز الطبقات لتجنب تلاشي التغيرات.
 - d. طبقة التغذية الأمامية وهي شبكة عصبية بسيطة الهدف منها هو استخراج ميزات أعلى من التمثيلات الناتجة عن آلية الانتباه.
5. بعد خروج من جميع طبقات المحول تمر البيانات عبر طبقة خطية تقلص التمثيل النهائي لكل عميل إلى قيمة عددية.
6. ثم عبر طبقة Sigmoid لتحويل القيمة إلى احتمال يتراوح بين 0 و1، يمثل احتمالية انسحاب العميل.

2.2.5 تحضير البيانات

في مرحلة التدريب تم استخدام مجموعة بيانات E-commerce Customer Data for Behavior Analysis.

جرى تحضير هذه البيانات قبل تدريب النموذج المقترح لاختيار أفضل الميزات التي يمكن أن يستخلص النموذج منها المعلومات وتقديمها للنموذج بأفضل شكل. حيث قمنا بفصل التاريخ إلى ثلاث ميزات (السنة، الشهر، اليوم) بدلا من ميزة تحتوي على المعلومات كلها. وأيضاً تم اتباع منهجية One-Hot encoding لتحويل القيم الفئوية (Categorical values) إلى قيم عددية 0 أو 1 بحيث يمكن للنموذج استخلاص معلومات منها. وبالإضافة إلى ذلك تمت معالجة البيانات الفارغة Null وتعبئتها (في مجموعة البيانات المستخدمة كانت لا تحتاج إلى معالجة كبيرة للبيانات الفارغة).

بعد تحضير البيانات تم تجميع عمليات شراء كل عميل على شكل سلسلة أو أكثر تتكون من 10 (الرقم يعتبر متحول يجب ملائمته للحصول على أفضل نتائج) عمليات شراء وبعدها تم تقييس البيانات (Scaling).

في بعض الحالات وخاصةً في مسائل توقع انسحاب العملاء تكون البيانات غير متوازنة من حيث عدد العملاء المنسحبين لذلك جرى اتباع منهجية إغناء بيانات التدريب oversampling لمنع النموذج عند التدريب من الانحياز إلى تصنيف العملاء على أنهم عملاء غير منسحبين أو العكس.

3.2.5 منهجية التدريب

بعد تحضير البيانات، تم تقسيمها إلى ثلاث أقسام وهي:

- مجموعة التدريب Training set وكانت نسبتها 70% من البيانات الكلية.
- مجموعة التحقق Validation set وكانت نسبتها 15% من البيانات الكلية.
- مجموعة الاختبار Testing set وكانت نسبتها 15% من البيانات الكلية.

تم تدريب النموذج على فترات (Epochs) على بيانات التدريب، بعد كل فترة (epoch) يتم تقييم النموذج والتحقق من التدريب على مجموعة التحقق وعند الانتهاء من كل الفترات يتم اختبار النموذج على مجموعة الاختبار. تم اعتماد منهجية التوقف المبكر (Early Stopping) وترك التعلم (Dropout) للحد من انحياز النموذج خلال عملية التدريب على بيانات التدريب. حيث جرى استخدام BCELoss (Binary Cross-Entropy Loss) وهي دالة خسارة مناسبة لمهام التصنيف الثنائي حيث يتم حساب الخطأ، حيث تكون قيمة الخسارة كبيرة كلما كانت التنبؤات بعيدة عن القيمة الحقيقية. لأن النموذج المقترح يخرج قيمة احتمالية بعد مروره عبر الطبقة الأخيرة sigmoid.

كما أنه تم تهيئة الأوزان قبل بدء التدريب باستخدام (Xavier uniform) من أجل تحسين بدء التدريب وتجنب مشاكل بطء التعلم والتدرجات غير المستمرة.

3.5 النموذج المقترح لتقسيم العملاء إلى شرائح

تم اعتماد خوارزمية K-Means لتقسيم العملاء إلى شرائح بناء على خصائصهم وسلوكهم الشرائي، نظراً لقدرتها الفعالة على اكتشاف الأنماط ضمن مجموعات كبيرة من البيانات بشكل غير إشرافي (Unsupervised). تتميز K-Means بالبساطة وسرعة الأداء، كما أنها مناسبة للمسائل التي يكون فيها الهدف هو تصنيف العملاء إلى مجموعات متجانسة اعتماداً على معايير مثل العمر، الإنفاق الكلي، وعدد المشتريات.

لضمان دقة التجميع، جرى تقييس البيانات (Scaling). ثم جرى استخدام طريقة "مخطط المرفق (Elbow Method)" لتحديد العدد الأمثل للشرائح (K). تعتمد هذه الطريقة على حساب مجموع التباين داخل المجموعات (Within-Cluster Sum of Squares) لكل قيمة ممكنة لـ K، ثم اختيار العدد الذي يحدث عنده انحناء، مما يشير إلى أن إضافة شرائح أخرى لن تحسن من جودة التصنيف بشكل كبير. بالإضافة إلى ذلك، تم الاستعانة بمؤشر Silhouette Score كأداة داعمة لقياس مدى تمايز الشرائح الناتجة، حيث تعكس القيم الأعلى جودة التقسيم.

4.5 النموذج المقترح لحساب القيمة الدائمة للعميل

يعتمد النموذج المقترح لحساب القيمة الدائمة للعميل على تحليل البيانات التاريخية لسلوك العملاء بهدف تقدير الإيرادات المستقبلية المتوقعة من كل عميل على حدة. يقوم النموذج بحساب متوسط قيمة الطلب (AOV) من خلال قسمة إجمالي المشتريات على عدد الطلبات، ثم تحديد تكرار الشراء (Purchase Frequency PF) لكل عميل خلال فترة زمنية محددة. كما يتم تقدير فترة حياة العميل (Customer Lifespan CL) من خلال معدل الانسحاب (Churn Rate)، حيث يُفترض أن العمر المتوقع للعميل يعادل مقلوب معدل الانسحاب.

ويُدرج هامش الربح الإجمالي (Gross Margin GM) لتقديم تقدير دقيق للقيمة الفعلية المكتسبة من العميل. تُدمج هذه العناصر في المعادلة التالية:

$$CLTV = AOV \times PF \times GM \times CL$$

الفصل السادس

تصميم النظام

يعرض هذا الفصل القرارات التصميمية التي بني من خلالها النظام

1.6 مقدمة

جرى اعتماد بنية الخدمات المصغرة (micro-services) لبناء التطبيق، إذ يتكون النظام من مجموعة من الخدمات التي تعمل مع بعضها بشكل متكامل لخدمة المستخدم النهائي. تقدم هذه البنية مجموعة من الفوائد أهمها:

- **التوسع المستقل:** يمكن توسيع كل خدمة مصغرة بشكل مستقل بناءً على الطلب، حيث تسمح هذه المرونة باستخدام الموارد بكفاءة، إذ يتم تسريع أجزاء التطبيق التي تحتاج إلى المزيد من الموارد فقط، بدا من النظام بأكمله.
- **المرونة في استخدام التقنيات:** يمكن استخدام لغات برمجة أطر عمل وقاعد بيانات وتقنيات أخرى مختلفة للخدمات المصغرة المختلفة، وتحسين كل خدمة بناءً على متطلباتها المحددة، إذ يمكن أن يؤدي هذا إلى أداء أفضل وحلول أكثر تخصيصاً لكل خدمة.
- **النشر المستقل:** يمكن نشر الخدمات المصغرة بشكل مستقل، مما يسمح بالتكامل المستمر والتسليم المستمر (CI/CD) مما يعطي إمكانية إصدار التحديثات وإصلاح الأخطاء بسرعة دون انتظار دورة إصدار النظام الكاملة.

نعرض في الفقرات التالية شرحاً لتصميم كل خدمة.

2.6 خدمة البيانات Data service

الهدف الأساسي من هذه الخدمة هو تخزين الموارد الخاصة بعملاء المستخدمين. لتزويد جميع الخدمات الأخرى بالمعلومات المطلوبة دون الحاجة إلى تكرار تنجز الوصول إلى هذه البيانات في كل خدمة من الخدمات الأخرى. تقوم هذه الخدمة بجلب البيانات وتخزينها في قاعدة البيانات وتوفر واجهات تخاطب لطلب استرجاع البيانات أو تخزين بيانات جديدة من الخدمات الأخرى.

لعدم تعقيد التنجيز تم استخدام منهجية 3-tier التي تعتمد على وجود API يطلب خدمة منطق العمل Business logic التي تتخاطب بدورها مع قاعدة المعطيات لتخزين واسترجاع الموارد.

من خلال فصل هذه الخدمة في خدمة مستقلة نحقق قابلية التوسع من حيث عدد المستخدمين من جهة وقابلية التوسع في أنماط البيانات القابلة للتخزين وكيفية تخزينها من جهة أخرى. حيث معالجة أي نوع جديد من الموارد أو أي تعديل في بنية تخزين البيانات سيتم ضمن هذه الخدمة بمعزل عن بقية الخدمات، مما يزيد من قابلية الصيانة ويسرع عملية التطوير.

3.6 خدمة المصادقة Auth service

الهدف الأساسي من هذه الخدمة هو المصادقة على طلبات المستخدمين، حيث يتم تزويد هذه الخدمة برمز الوصول (Access Token) الخاص بالمستخدم الذي طلب هذا الطلب وتقوم هذه الخدمة بالمصادقة عليه، كما أن هذه الخدمة مسؤولة عن عمليات تسجيل دخول وإنشاء حسابات المستخدمين.

من خلال فصل هذه الخدمة على حدة، يمكن لكل الخدمات الأخرى التحقق من أي رمز وصول (Access Token) من أجل مصادقة المستخدمين دون الحاجة إلى تجنيز فهم العمل ضمنها.

4.6 خدمة توقع انسحاب العملاء Churn service

الهدف الأساسي من هذه الخدمة هو استخدام نموذج ذكاء صناعي قادر على التنبؤ بالعملاء الذين يتوقع انسحابهم من منصة التجارة الإلكترونية الخاصة بالمستخدم. حيث تقوم هذه الخدمة بالتواصل مع خدمة البيانات للحصول على بيانات عملاء المستخدم وتقوم بتحليلها والتنبؤ بالعملاء المتوقع انسحابهم وعرضهم على المستخدم.

تم استخدام منهجية 3-tier في هذه الخدمة لتوفير خدمات التوقع بانسحاب العملاء. من خلال فصل هذه الخدمة لتحقيق قابلية التوسع وتزويد سرعة الاستجابة لأنه من الممكن ان تكون هذه الخدمة هي عنق الزجاجة في النظام بسبب الزمن لازم لتوقع انسحاب العملاء وبذلك يمكن التوسع على عدد مخدمات أكبر لتحقيق أفضل استجابة للمستخدمين دون تأثير أحدهم على الآخر.

5.6 خدمة تقسيم العملاء إلى شرائح Customer Segmentation Service

الهدف من هذه الخدمة هو تحديد الشريحة التي ينتمي إليها كل عميل وعرض معلومات عن شرائح العملاء للمستخدم. حيث تتواصل هذه الخدمة مع خدمة البيانات للحصول على بيانات المستخدمين.

تم استخدام منهجية 3-tier في هذه الخدمة أيضاً، ومن خلال فصل هذه الخدمة يمكن أيضاً توسيعها ونشرها على أكثر من مخدم حسب الطلب لتحقيق قابلية التوسع.

6.6 خدمة القيمة الدائمة للعملاء Customer Lifetime Value

الهدف من هذه الخدمة هو حساب القيمة الدائمة للعملاء الخاصة بالمستخدم، حيث تتواصل هذه الخدمة مع خدمة البيانات لجلب بيانات المستخدمين وتتواصل أيضاً مع خدمة توقع انسحاب العملاء للحصول على بعض القيم والمقاييس التي تفيد في حساب أفضل للقيمة الدائمة للعملاء.

تم استخدام منهجية 3-tiers في هذه الخدمة أيضاً، ومن خلال فصل هذه الخدمة يمكن توسيعها ونشرها على أكثر من مخدم وفي حال قمنا بتغيير المنهجية المتبعة في حساب القيمة الدائمة للعملاء أو إضافة ميزات جديدة يسهل اضافتها وتشغيلها دون الحاجة إلى التحقق من تأثير هذه الإضافات على النظام ككل.

7.6 خدمة البوابة API-Gateway

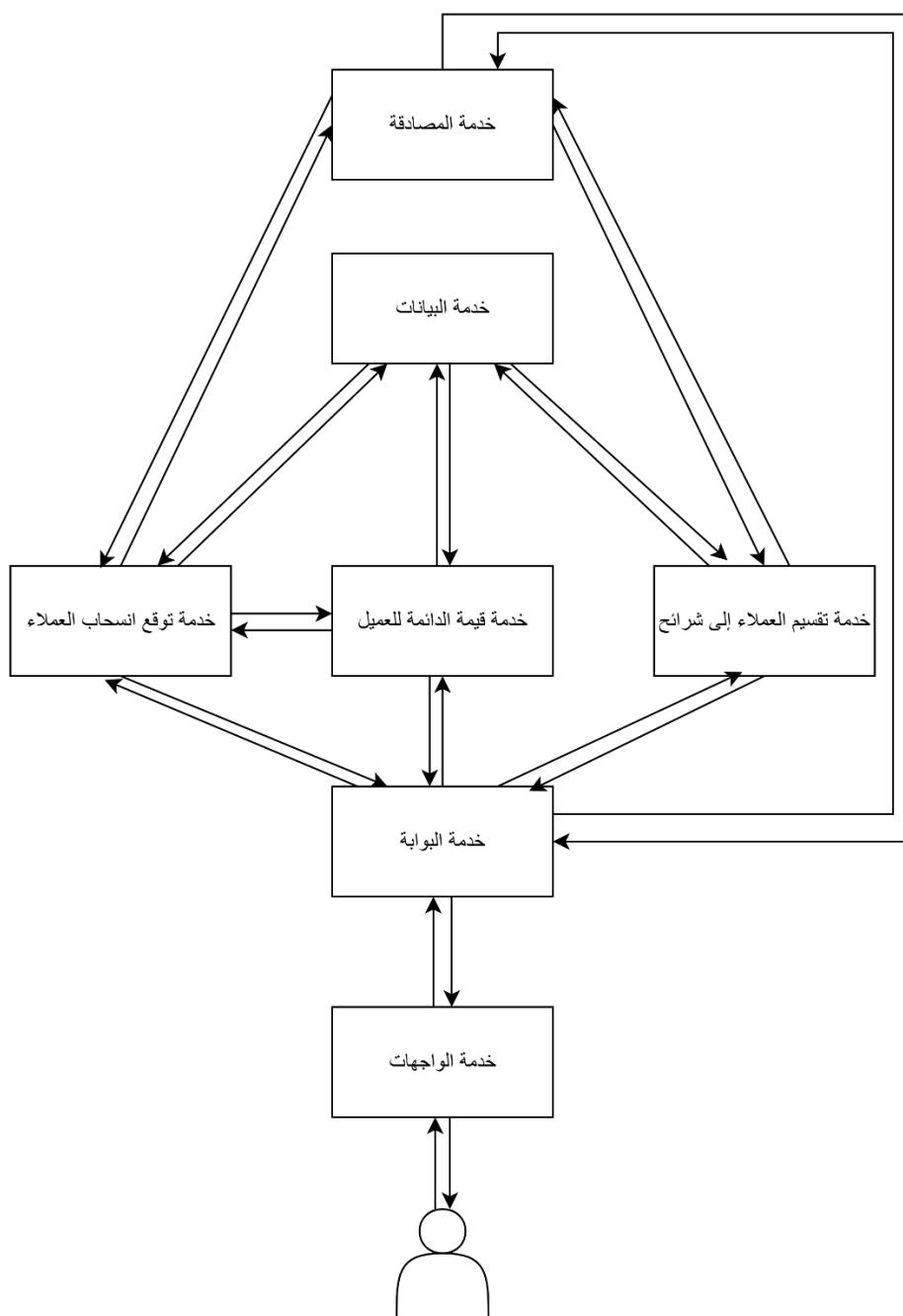
بعد النظر إلى الخدمات التي تكوّن النظام الكلي، نجد أن تواصل المستخدم مع كل خدمة على حدة يشكل مجموعة من التحديات مما يؤثر على سهولة صيانة وتشغيل النظام لذلك ظهرت الحاجة لوجود مخدم يعمل كواجهة تخاطب للنظام بحيث تقوم بتوجيه الطلبات على الخدمات الأخرى حسب الطلب.

8.6 خدمة واجهة المستخدم Front-end

تم بناء واجهة المستخدم باستخدام بنية قائمة على المكونات مما يوفر العديد من الفوائد أهمها:

- النسقية (Modularity): تسمح بتقسيم موقع الويب إلى مكونات أصغر مستقلة. يغلف كل مكون وظيفة محددة، مما يجعله سهل الإدارة والفهم.
- قابلية إعادة الاستخدام: يمكن إعادة استخدام كل مكون من مكونات موقع الويب في أجزاء مختلفة ضمن الموقع أو حتى في مشاريع مختلفة لاحقة. مما يؤدي إلى تسريع عملية التطوير.
- سهولة تصحيح الأخطاء: بسبب عزل المكونات عن بعضها البعض تكون عملية تصحيح الأخطاء وتتبعها عملية بسيطة لا تتطلب الوقت والجهد في غالب الأحيان.

9.6 مخطط تصميم النظام



الشكل 6 مخطط تصميم النظام

الفصل السابع

الأدوات المستخدمة

يعرض هذا الفصل الأدوات المستخدمة لتنفيذ العمل.

PostgreSQL 1.7

PostgreSQL هو نظام إدارة قواعد بيانات علائقية مفتوح المصدر وقوي يدعم الاستعلامات المعقدة والمعاملات. يُستخدم على نطاق واسع في التطبيقات التي تتطلب استقراراً وأماناً عالياً.

React 2.7

React مكتبة JavaScript تُستخدم لبناء واجهات المستخدم التفاعلية، تتيح إنشاء مكونات قابلة لإعادة الاستخدام وتحديث واجهة المستخدم بشكل فعال عند تغيير البيانات.

طورتها Meta وتُعد أساساً لتطبيقات (Single Page Application SPA).

Python 3.7

Python هي لغة برمجة عالية المستوى، تتميز بالبساطة والوضوح وتُستخدم في تطوير الويب والذكاء الاصطناعي وتحليل البيانات.

FastAPI 4.7

FastAPI إطار عمل Python حديث لإنشاء واجهات REST API عالية الأداء. يتميز بالسرعة ودعم توثيق API تلقائي باستخدام Swagger.

Json Web Token (JWT) 5.7

هو معيار مفتوح يُستخدم لنقل المعلومات بين طرفين بشكل آمن كرمز مشفر (Token)، يُستخدم عادةً للتحقق من الهوية في تطبيقات الويب.

Git 6.7

نظام تحكم في الإصدارات يُستخدم لتتبع التغييرات في الملفات البرمجية وتنسيق العمل الجماعي بين المطورين. ويدعم العمل الموزع.

GitHub 7.7

GitHub منصة استضافة لمشاريع Git تسمح بالتعاون في تطوير البرمجيات، تشمل ميزات مثل طلبات السحب، تتبع المشاكل، وأتمتة CI/CD.

Redis 8.7

Redis قاعدة بيانات NoSQL في الذاكرة تدعم هياكل بيانات مثل السلاسل والقوائم والمجموعات. تُستخدم في التخزين المؤقت (caching).

PyTorch 9.7

مكتبة تعلم عميق طُورت بواسطة Meta، تُستخدم في بناء وتدريب الشبكات العصبية باستخدام واجهة Python مرنة وسهلة التجريب.

Jupyter Notebook 10.7

Jupyter Notebook بيئة تفاعلية تُستخدم لكتابة وتشغيل كود Python وتحليل البيانات وتوثيق النتائج، خاصة في مجالات التعليم والبحث العلمي.

الفصل الثامن

تنجيز النظام

يعرض هذا الفصل كيفية تنجيز النظام مع تفصيل كل جزء من أجزائه.

1.8 مقدمة

خلال عملية تطوير الخدمات يتم رفع الرماز المصدري الخاص بكل خدمة باستخدام أداة Git إلى مستودعات (Repositories) ضمن موقع GitHub حيث يتم تجميع الرماز المصدري ضمن فرع (Branch) أساسي اسمه main وعند إضافة ميزات جديدة يتم العمل على فروع فرعية بحيث يتم التطوير عليها وإجراء الاختبار دون التأثير على الفرع الأساسي الذي يحتوي على رماز التطبيق النهائي، عند الانتهاء من تطوير الميزة الجديدة يتم دمج الفرعين بحيث يتحدث الفرع الأساسي ويكون يحتوي على الميزة الجديدة.

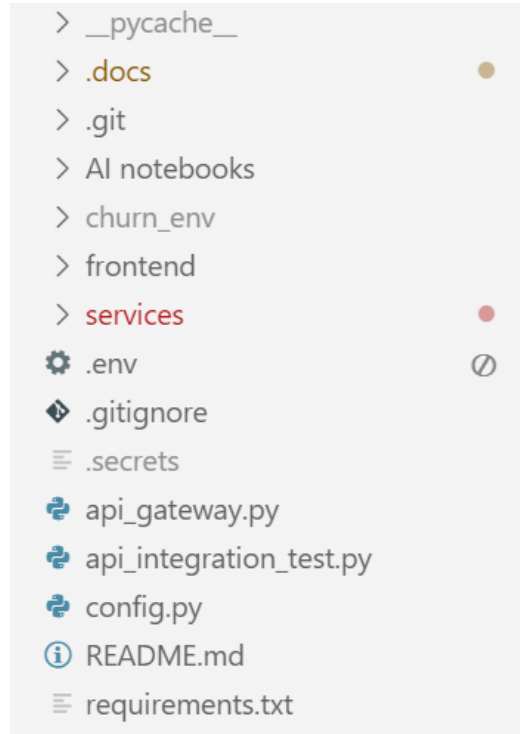
يساعد هذا الأسلوب في حفظ النسخ (Version Control) مما يمنع فقدان الرماز المصدري بشكل كامل ويسهل عملية التطوير ومراقبة التعديلات. وفي حال كان هنالك فريق من المطورين يعمل هذا تطوير التطبيق يساعد هذا الأسلوب في تنسيق عمل الفريق وعدم تداخل العمل بين المطورين مما يقلل الأخطاء ويسرع عملية التطوير.

تم بناء هذه الخدمات باستخدام إطار عمل FastAPI بلغة Python، تم اختيار Python عن غيرها من اللغات بسبب الدعم الكبير لها في مجال الذكاء الصناعي حيث تحتوي على العديد من المكتبات التي تساعد في بناء التطبيقات ونماذج الذكاء الصناعي مثل TensorFlow, PyTorch, Sklearn مما يسهل ويسرع عملية التطوير.

تم الاعتماد على إطار العمل FastAPI لكونه إطار العمل الأكثر دعماً من حيث المجتمعات والمكتبات ولتوافق مع البنية التي طورت عليها نماذج الذكاء الصناعي مما يوفر سرعة في عملية التطوير.

نعرض في الفقرات التالية شرحاً تفصيلياً عن كل خدمة.

2.8 بنية النظام ومكوناته



الشكل 7 مجلدات النظام

تم تجميع النظام بكل تطبيقاته وخدماته ضمن بيئة واحدة لسهولة الوصول والتشغيل، حيث كانت المكونات النظام هي:

ReadMe: ملف توثيقي بصيغة Markdown يحتوي على شرح عام للمشروع، طريقة التثبيت، التشغيل، واستخدامه.

.env: ملف لتخزين متغيرات البيئة مثل روابط تشغيل الخدمات وهي بيانات متغيرة لا تُضمّن في الرماز مباشرةً.

.secrets: يُستخدم لتخزين معلومات سرّية مثل مفاتيح التشفير بشكل آمن وهو مستثنى من التتبع في Git.

Churn_env: يُستخدم كبيئة افتراضية خاصة بالمشروع تحتوي على الحزم المثبتة ضمن المشروع.

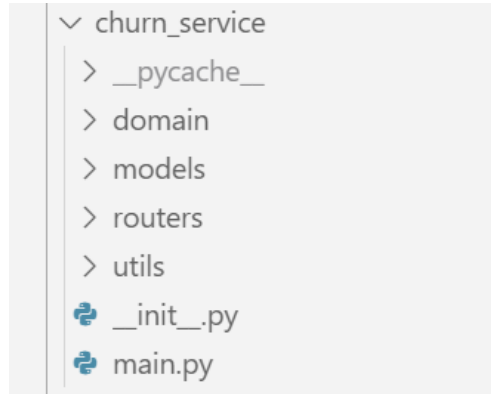
frontend: المجلد يحتوي على خدمة واجهات المستخدم.

services: مجلد يحتوي على الخدمات المقدمة في النظام.

Requirements.txt: يحتوي الملف على قائمة بجميع الحزم والمكتبات المطلوبة لتشغيل المشروع مع تحديد إصداراتها، يستخدم هذا الملف لتثبيت الاعتماديات عند البدء بالمشروع من أجل التوافقية.

3.8 خدمة توقع انسحاب العملاء

1.3.8 تفاصيل التنفيذ



الشكل 8 مجلدات خدمة توقع انسحاب العملاء

تتكون هيكلية مجلدات خدمة انسحاب العملاء من المكونات التالية:

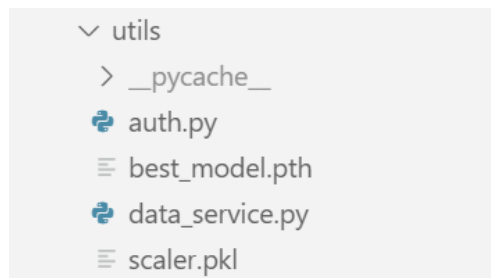
Domain: هو الجزء الذي يحتوي على منطق العمل Business Logic وتجزئ التوابع المستخدمة في هذه الخدمة.

Models: يحتوي على الأنماط والصفوف التي ستتفاعل معها الخدمة لتحقيق منطق العمل Business Logic.

Routers: تشمل تعريف نقاط النهاية (Endpoints) التي تتعامل مع الطلبات (Requests) والاستجابات (Responses) ضمن FastAPI.

Utils: تحتوي على وظائف أو أدوات مساعدة (Helper Functions) مثل الاتصال مع الخدمات الأخرى وتحتوي على نموذج الذكاء الصناعي وأداة التقييم.

ضمن مجلد utils يوجد نموذج مدرب للذكاء الصناعي model ومقيس scaler مقاسة عليه البيانات سابقاً عند التدريب.

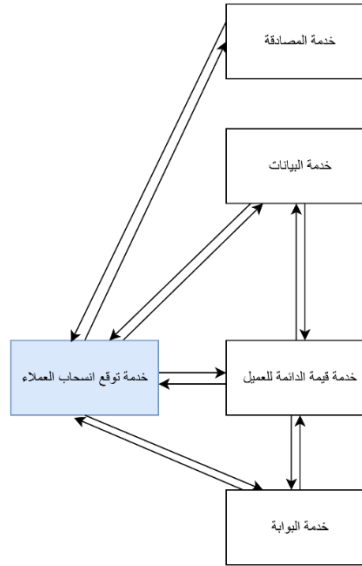


الشكل 9 مكونات مجلد utils في خدمة توقع انسحاب العملاء

سيناريو العمل الأساسي للخدمة: يتم تحميل الـ Model والـ Scaler عند بدء التطبيق، عند وصول طلب يتم جلب البيانات الخاصة بالمستخدم من خدمة البيانات ومعالجتها باستخدام التوابيع الموجودة ضمن مجلد Domain الذي بدوره يقيس البيانات باستخدام scaler وتقرير هذه البيانات على الـ model للحصول على التوقع لكل عميل.

جرى إضافة تخزين مؤقت باستخدام (Redis) حيث انه عندما يبدأ توقع انسحاب العملاء يتم انشاء مهمة Task خاصة بالطلب ويتم تخزين التقدم progress لمهمة التوقع مع نتائج التي قد تمت معالجتها ويتم طلب تحديث بالتقدم كل فترة زمنية لإعلام المستخدم النهائي بمقدار التقدم الحاصل في توقع انسحاب العملاء.

تواصل خدمة توقع انسحاب العملاء مع الخدمات الأخرى



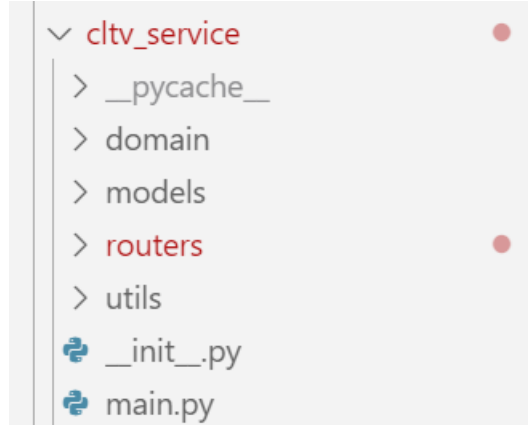
الشكل 10 تواصل خدمة انسحاب العملاء بالخدمات الأخرى

تستقبل خدمة توقع انسحاب العملاء من خدمة البوابة طلبات المستخدمين للحصول على العملاء المتوقع انسحابهم وتعيد لخدمة البوابة التوقعات بعد معالجتها وتحليل البيانات. عند وصول الطلب، يتم مصادقة المستخدم من خلال ارسال رمز الوصول إلى خدمة المصادقة وتعيد تأكيد بهوية المستخدم، في حال التأكد، تطلب خدمة توقع انسحاب العملاء بيانات المستخدم المصادق عليه من خدمة البيانات وتقوم بمعالجتها وهيئتها بحيث يتم توقع الانسحاب باستخدام نموذج الذكاء الصناعي وتعيد الخدمة النتائج لخدمة البوابة.

أيضا تتواصل خدمة توقع انسحاب العملاء مع خدمة القيمة الدائمة للعميل حيث تطلب خدمة القيمة الدائمة للعميل حساب نسبة العملاء المنسحبين.

4.8 خدمة القيمة الدائمة للعميل

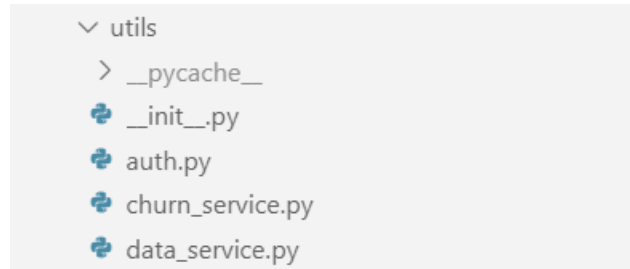
1.4.8 تفاصيل التنجيز



الشكل 11 مجلدات خدمة القيمة الدائمة للعميل

هيكلية مجلدات خدمة القيمة الدائمة للعميل هي نفسها هيكلية ملفات خدمة انسحاب العملاء المذكورة سابقا ولكن يختلف التنجيز ضمنها لكي تتوافق مع الهدف الأساسي من الخدمة.

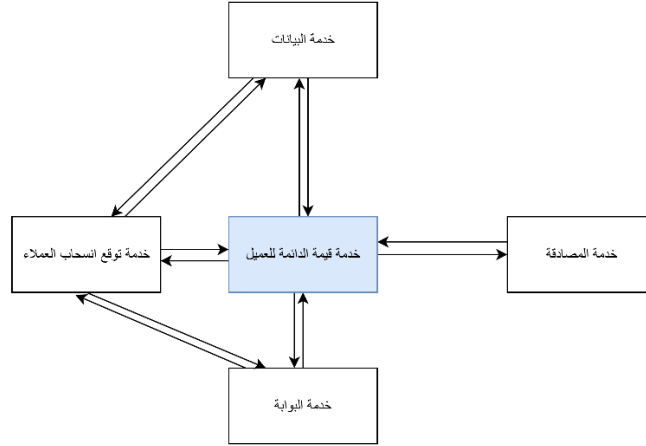
ضمن مجلد `utils` يوجد طرق التخاطب مع الخدمات الأخرى (خدمة البيانات، توقع انسحاب العملاء، المصدقة).



الشكل 12 محتويات مجلد `utils`

سيناريو العمل الأساسي للخدمة: يتم طلب هذه الخدمة من خدمة البوابة للحصول على معلومات القيمة الدائمة للعميل تقوم الخدمة بطلب بيانات المستخدم ومعدل انسحاب العملاء من خدمة توقع انسحاب العملاء وتقوم بإجراء العمليات الحسابية وتعيد نتائج القيمة الدائمة للعملاء.

تواصل خدمة القيمة الدائمة للعميل مع الخدمات الأخرى

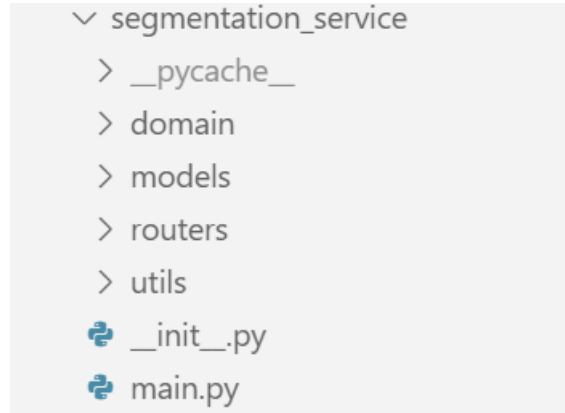


الشكل 1.3 تواصل خدمة القيمة الدائمة للعميل بالخدمات الأخرى

تستقبل خدمة القيمة الدائمة للعميل من خدمة البوابة طلبات المستخدمين للحصول على القيمة الدائمة للعملاء وتعيد إلى خدمة البوابة النتائج بعد معالجتها وتحليلها للبيانات. عند وصول طلب المستخدم تتم مصادقته مع خدمة المصادقة حيث يتم ارسال رمز الوصول إلى الخدمة وتعيد تأكيد بهوية المستخدم، في حال التأكيد يتم طلب بيانات المستخدم المصادق عليه من خدمة البيانات وتتم معالجتها وهيئتها بحيث يتم حساب القيمة الدائمة للعملاء وتعيد الخدمة النتائج إلى خدمة البوابة.

5.8 خدمة تقسيم البيانات إلى شرائح

1.5.8 تفاصيل التنفيذ

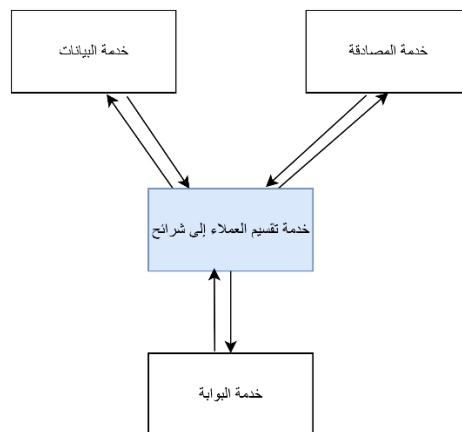


الشكل 14 مجلدات خدمة تقسيم العملاء إلى شرائح

هيكلية مجلدات خدمة تقسيم العملاء إلى شرائح هي نفسها هيكلية ملفات خدمة انسحاب العملاء المذكورة سابقا ولكن يختلف التنفيذ ضمنها لكي تتوافق مع الهدف الأساسي من الخدمة.

سيناريو العمل الأساسي للخدمة: يتم طلب هذه الخدمة من خدمة البوابة للحصول على شرائح العملاء ومعلومات عنها تقوم الخدمة بطلب بيانات المستخدم وتقوم بإجراء العمليات الحسابية وتعيد نتائج شرائح العملاء.

تواصل خدمة تقسيم العملاء إلى شرائح مع الخدمات الأخرى

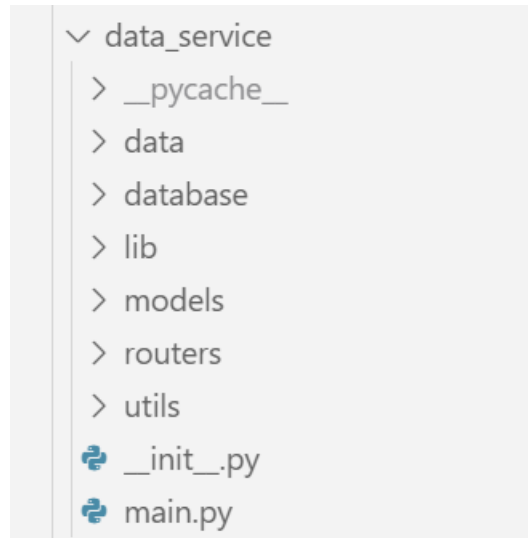


الشكل 15 تواصل خدمة تقسيم العملاء مع الخدمات الأخرى

تستقبل خدمة تقسيم العملاء إلى شرائح من خدمة البوابة طلبات المستخدمين للحصول على شرائح العملاء وتعيد الخدمة البوابة التوقعات بعد معالجتها وتحليلها للبيانات. عند وصول طلب المستخدم تتم مصادقته مع خدمة المصادقة حيث يتم ارسال رمز الوصول إلى الخدمة وتعيد تأكيد بهوية المستخدم، في حال التأكيد تطلب الخدمة بيانات المستخدم المصادق عليه من خدمة البيانات وتقوم بمعالجتها وتهيئتها بحيث يتم تقسيم العملاء إلى شرائح وتعيد الخدمة النتائج لخدمة البوابة.

6.8 خدمة البيانات

1.6.8 تفاصيل التنفيذ



الشكل 16 مجلدات خدمة البيانات

هيكلية مجلدات خدمة البيانات هي نفسها هيكلية ملفات خدمة انسحاب العملاء المذكورة سابقا ولكن يختلف التنفيذ ضمنها لكي تتوافق مع الهدف الأساسي من الخدمة مع إضافة مجلد database الذي يحتوي على ملفات إعدادات الاتصال بقاعدة البيانات وتعريف الجداول وتنفيذ توابع لعمليات تخزين واسترجاع البيانات.

الهدف الأساسي من هذه الخدمة: هو ان يتم عزل تنفيذ استرجاع البيانات من قاعدة البيانات عن الخدمات الأخرى بحيث لا يتم تكرار التنفيذ ضمن الخدمات، إضافة إلى تحقيق معايير الأمن، حيث انه لا يمكن للمستخدمين الوصول إلى بيانات عملاء مستخدمين آخرين حيث ان تنفيذ هذا المفهوم موجود ضمن هذه الخدمة.

2.6.8 جداول البيانات في قاعدة البيانات

تحتوي قاعدة الجداول التالية:

- `upload_history`: هو جدول يتم فيه تخزين معلومات عن عملية رفع بيانات المستخدمين.

Name	upload_time
id	status
user_id	file_size
filename	records_count
table_name	error_message

الشكل 17 أعمدة جدول `upload_history`

- `user_data`: هو جدول يتم فيه تخزين بيانات المستخدمين.

Name	Churn
Customer ID	Year
Purchase Date	Month
Product Price	Day
Quantity	Gender_Male
Total Purchase Amount	Payment Method_Credit Card
Customer Age	Payment Method_PayPal
Returns	Product Category_Clothing
Customer Name	Product Category_Electronics
Age	Product Category_Home

الشكل 18 أعمدة جدول `user_data`

- `user_predictions`: هو جدول يحتوي على بيانات الناتجة عن عملية توقع الحساب العملاء بعد معالجتها.

id
customer_id
churn_probability
confidence
name
email
totalSpent
last_purchase_date

الشكل 19 أعمدة جدول `user_predictions`

7.8 خدمة المصادقة

1.7.8 تفاصيل التنفيذ

```
▼ auth_service
  > __pycache__
  > models
  > routers
  📄 __init__.py
  📄 auth_utils.py
  📄 config.py
  📄 database.py
  📄 main.py
```

الشكل 20 مجلدات خدمة المصادقة

هيكلية مجلدات خدمة المصادقة هي نفسها هيكلية ملفات خدمة انسحاب العملاء المذكورة سابقا ولكن يختلف التنفيذ ضمنها لكي تتوافق مع الهدف الأساسي من الخدمة مع إضافة مجلد database الذي يحتوي على ملفات إعدادات الاتصال بقاعدة البيانات وتعريف الجداول وتنفيذ توابع لعمليات تخزين واسترجاع بيانات المستخدمين.

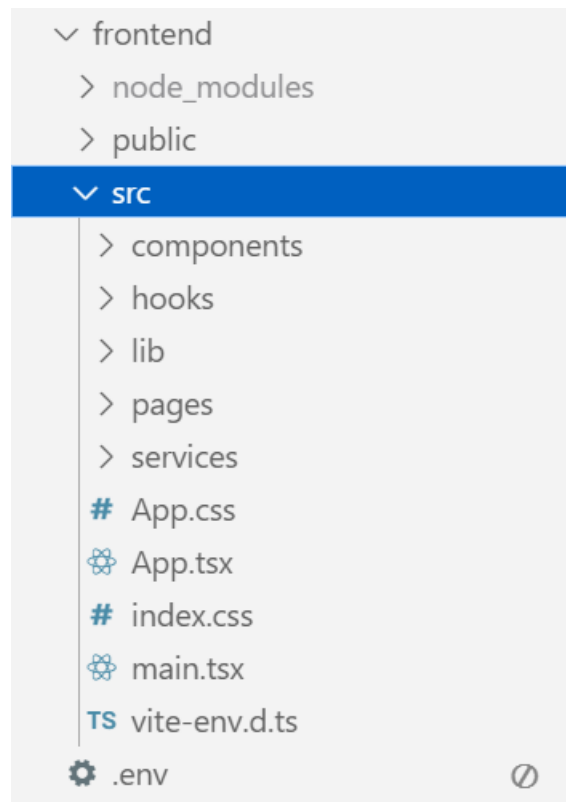
الهدف الأساسي من هذه الخدمة: هو إدارة المستخدمين وبهذا يتم عزل تنفيذ مصادقة المستخدمين عن الخدمات الأخرى بحيث لا يتم تكرار التنفيذ ضمن الخدمات.

8.8 خدمة البوابة

1.8.8 تفاصيل التنجيز

تحتوي خدمة البوابة على تابع proxy يقوم بتنجيز مفهوم العمل الخاص بهذه الخدمة Business Logic، خدمة البوابة هي خدمة مسؤولة عن تحويل طلبات المستخدمين إلى باقي خدمات النظام بحيث توجه الطلبات حسب نوع الطلب والخدمة. تم تنجيز هذه الخدمة باستخدام إطار عمل FastAPI عوضا عن استخدام أدوات جاهزة وذلك لسهولة التحكم بمنطق العمل وسهولة التوسعة عند الحاجة ولأنها تتكامل بشكل مباشر مع البنية الخلفية حيث يمكن إضافة سجل عمليات النظام ضمنها logging.

9.8 خدمة وجهات المستخدم



الشكل 21 مجلدات خدمة واجهات المستخدم

تتكون هيكلية مجلدات خدمة واجهات المستخدم من المكونات التالية:

Components: يحتوي على مكونات واجهة المستخدم القابلة لإعادة الاستخدام.

Pages: يتضمن مكونات تمثل الصفحات الكاملة (views) التي تُعرض في مسارات (Routes) معينة.

Services: يضم ملفات لتنفيذ طلبات API والتفاعل مع الخادم (backend) مثل جلب البيانات، إرسال نماذج، أو تنفيذ مصادقة.

Lib: يحتوي على أدوات أو وظائف عامة

Hooks: يحتوي على Hooks مخصصة (Custom React Hooks) لتعزيز وإعادة استخدام المنطق البرمجي المرتبط بالحالة أو التأثيرات مثل Toast.

الهدف الأساسي من هذه الخدمة: هو تقديم واجهات للمستخدم يستطيع من خلالها التفاعل مع النظام.

تتواصل هذه الخدمة مع خدمة البوابة فقط حيث أن خدمة البوابة تقوم بتوجيه الطلبات الى الخدمات المطلوبة وتعيد الإجابة لواجهات المستخدم.

الفصل التاسع

اختبارات النظام ومناقشة النتائج

يوضح هذا الفصل الاختبارات التي تم تطبيقها ونتائج هذا العمل.

1.9 اختبار النموذج المقترح لتوقع انسحاب العملاء

1.1.9 منهجية التقييم المتبعة

تم تقييم النموذج المقترح باستخدام مجموعتي بيانات المتوفرين في الدراسة المرجعية وفق المنهجية المتبعة في البحث [1]. اعتمدت الورقة البحثية على تقييم أداء نموذج CASPR في التنبؤ بانسحاب العملاء عبر استخدام مقياسي AUROC و-F1 score، وذلك بمقارنته مع نماذج تقليدية تعتمد على ميزات هندسية مثل RFM. تم تطبيق المنهجية على بيانات واقعية مثل KKBox وMicrosoft Retail، حيث تم اعتبار العميل منسحباً إذا لم يسجل أي تفاعل خلال فترة زمنية محددة. أظهرت النتائج أن تمثيلات CASPR حسّنت الأداء بشكل ملحوظ، خصوصاً في F1-score، ما يدل على قدرتها في التقاط أنماط سلوكية دقيقة دون الحاجة لهندسة ميزات معقدة [1].

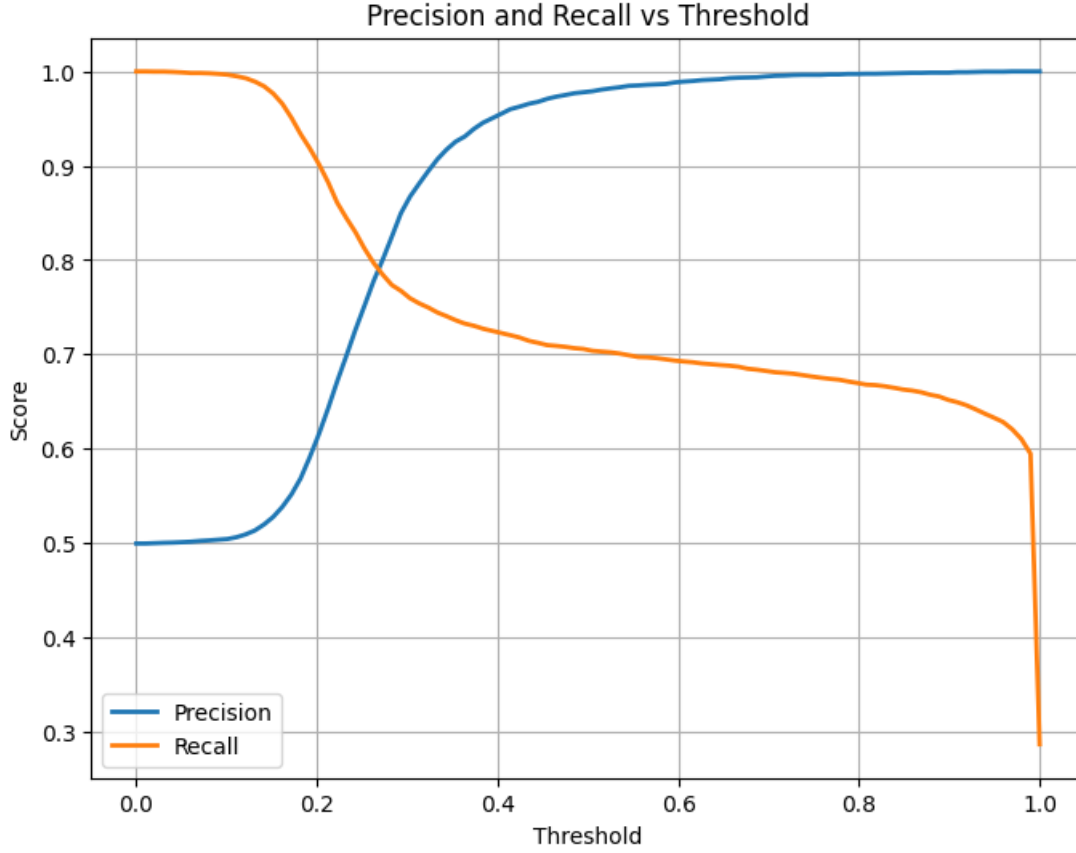
1.1.1.9 تقييم النموذج المقترح على مجموعة بيانات E-commerce Customer Data For Behavior Analysis

بعد تنفيذ النموذج المقترح وتقييمه على جزء من مجموعة البيانات التي تم التدريب عليها ولكن دون أن يتدرب النموذج على جزء الاختبار، كانت النتائج عند عتبة هي 0.5 على الشكل التالي:

0.84	Accuracy
0.96	Precision
0.71	Recall
0.82	F1 Score
0.87	AUROC

جدول 19 النتائج وفق مجموعة بيانات E-commerce Customer Data For Behavior Analysis

يعرض الشكل التالي علاقة Precision و Recall مع العتبة Threshold:



الشكل 22 علاقة precision recall مع العتبة Threshold على مجموعة بيانات E-commerce Customer Data For Behavior Analysis

تظهر النتائج التوازن بين الدقة Precision والاسترجاع Recall حيث كان F1 Score جيد يبلغ 0.82، بينما يشير AUROC البالغ 0.87 إلى قدرة تمييز عالية للنموذج عبر مختلف العتبات.

في الشكل 17، نلاحظ العلاقة العكسية بين الدقة والاسترجاع عند تغيير قيمة العتبة (Threshold): عند العتبات المنخفضة، يكون الاسترجاع مرتفعاً لكن الدقة منخفضة، بينما يحدث العكس عند رفع العتبة. يتقاطع المنحنيان تقريباً عند عتبة 0.25، وهي نقطة مناسبة للموازنة بين الدقة والاسترجاع، وتدعم اختيار عتبة تحقق توازناً جيداً حسب الهدف التجاري (مثلاً: تقليل خسارة العملاء أو تقليل الإنذارات الكاذبة).

2.1.1.9 Brazilian E-Commerce Public Dataset by OLIST

تحضير البيانات بالنسبة لمجموعة بيانات OLIST.

مجموعة البيانات هذه كانت البيانات من غير طبيعة البيانات السابقة، ولم يتدرب عليها النموذج حيث أنها كانت لا تحتوي على بعض الميزات التي دُرِب عليها النموذج سابقاً مثل ميزة عمر العميل وجنسه وكمية العناصر المشتراة في عملية الشراء وما إذا كان العميل أعاد الطلب (لم يقبل بعملية الشراء) وأيضاً لم يكن هنالك ميزة ما إذا كان العميل منسحباً أم لا.

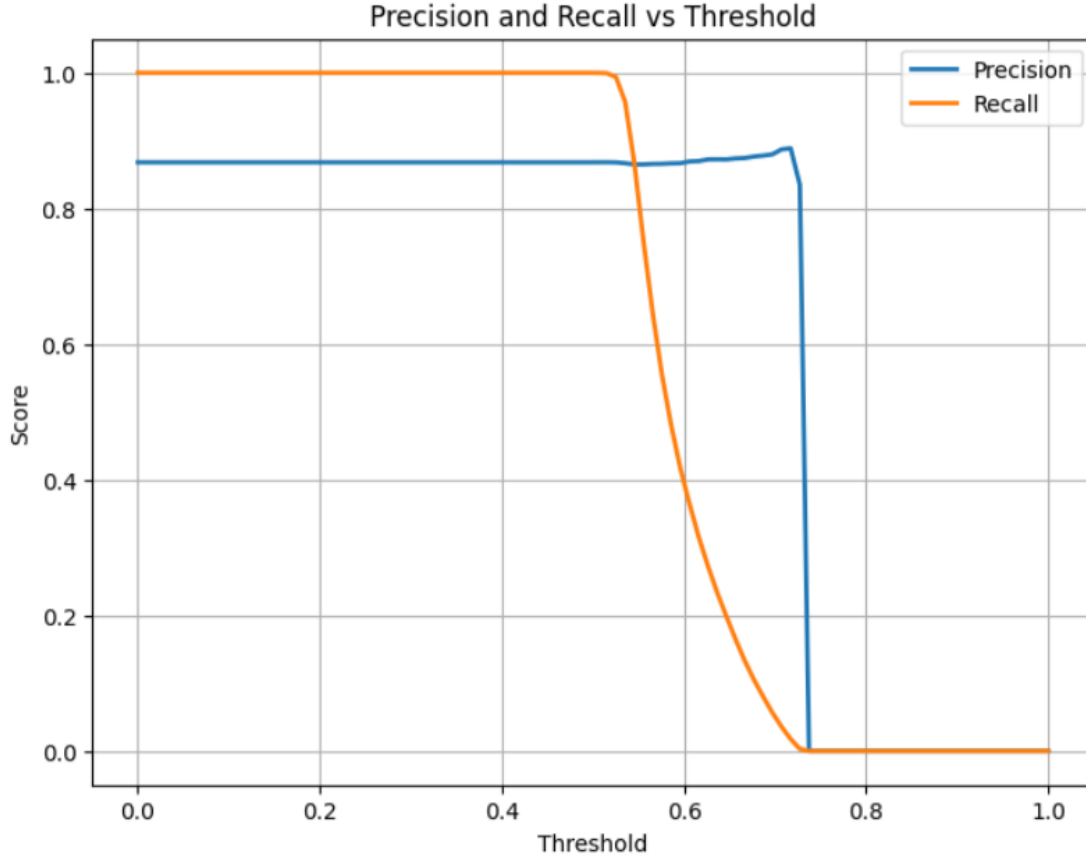
لمعالجة فقدان المعلومات تم وضع القيمة الوسطية في مجموعة البيانات السابقة للميزات المفقودة أما بالنسبة لتحديد ما إذا كان العميل منسحباً أم لا حيث تم اعتبار العميل منسحباً إذا لم يسجل أي تفاعل خلال فترة زمنية محددة.

بعد تنفيذ النموذج المقترح وتقييمه على مجموعة البيانات كانت النتائج عند العتبة 0.2 على الشكل التالي

0.72	Accuracy
0.86	Precision
0.81	Recall
0.83	F1 Score
0.49	AUROC

جدول 20 النتائج وفق مجموعة بيانات OLIST

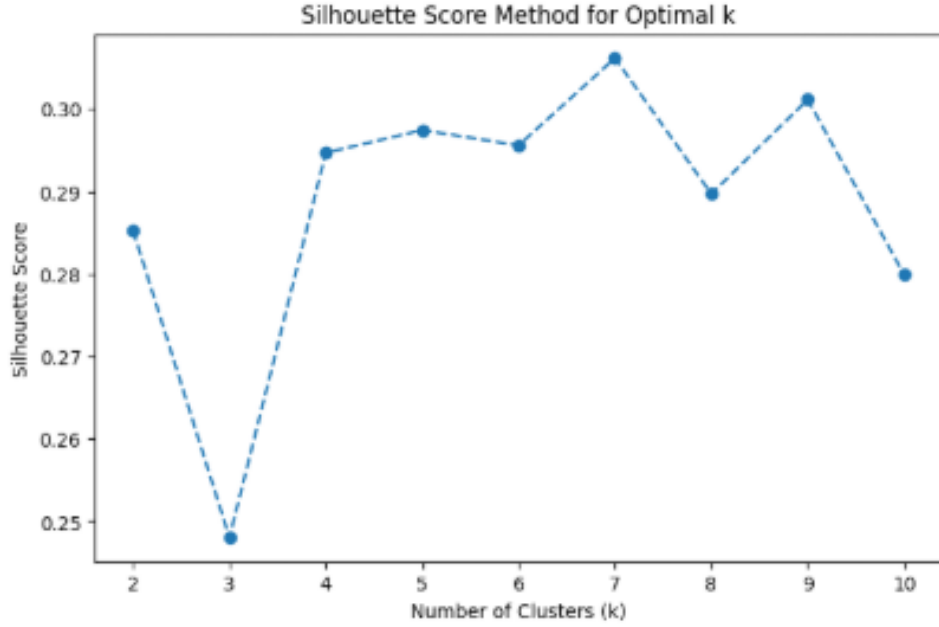
يعرض الشكل التالي علاقة Precision و Recall مع العتبة Threshold:



الشكل 23 علاقة Precision و Recall مع العتبة Threshold على مجموعة بيانات OLIST

تعكس نتائج النموذج على مجموعة بيانات OLIST أداءً مشجعاً في عدة جوانب مهمة. فقد حقق النموذج دقة إيجابية (Precision) مرتفعة بلغت 0.86، مما يدل على أن غالبية التنبؤات الإيجابية كانت صحيحة، وهو أمر بالغ الأهمية في تطبيقات تتطلب تقليل الإنذارات الكاذبة. كما أن الاسترجاع (Recall) الجيد عند 0.81 يشير إلى قدرة النموذج على اكتشاف عدد كبير من الحالات الإيجابية الحقيقية، مما يعزز موثوقيته في الكشف عن الأنماط الفعلية. بالإضافة إلى ذلك، فإن درجة F1 البالغة 0.83 تعكس توازناً ممتازاً بين الدقة والاسترجاع، وهي مؤشر واضح على فعالية النموذج في معالجة التحديات التنبؤية. ورغم أن قيمة AUROC كانت منخفضة، فإن النتائج الأخرى تثبت أن النموذج يقدم أداءً عملياً جيداً عند اختيار عتبة مناسبة، ويمكن تحسينه بسهولة بمزيد من الضبط أو التوسيع في البيانات.

2.9 نتائج نموذج تقسيم العملاء إلى شرائح



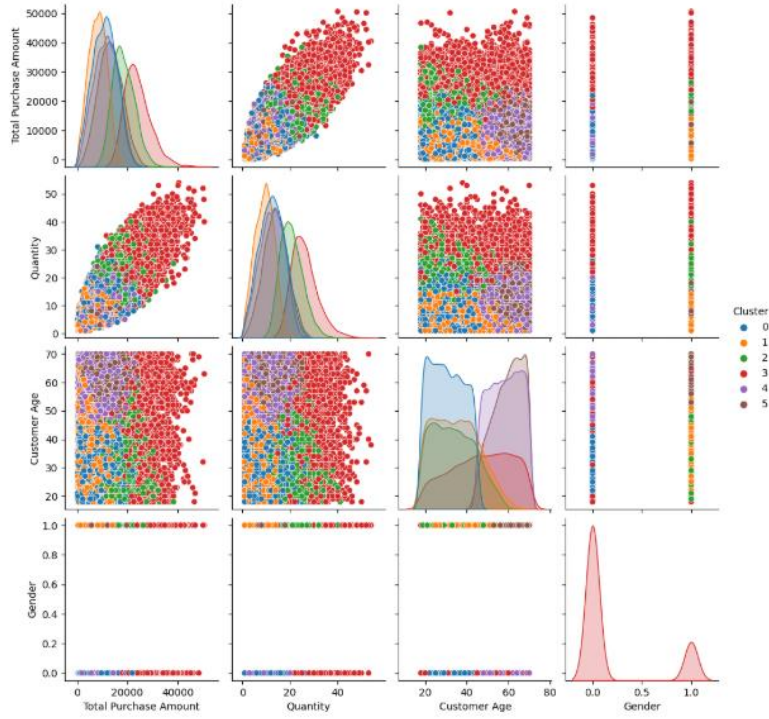
الشكل 24 مؤشر السيلويت لتحديد عدد العناقيد الأمثل

بناءً على نتائج مؤشر السيلويت الموضحة في الشكل، نجد أن أفضل عدد للعناقيد هو سبعة عناقيد ($k=7$) حيث يحقق أعلى قيمة للمؤشر، مما يدل على أن تقسيم العملاء إلى سبع مجموعات يعطي أفضل تمايز بين الشرائح. ويشير ذلك إلى أن كل شريحة من العملاء لها خصائص مميزة بشكل كافٍ عن الشرائح الأخرى في هذا التقسيم. مع ذلك، يمكن ملاحظة أن قيم المؤشر ليست عالية جداً مما يعني أن هناك بعض التداخل بين الشرائح، وهذا أمر شائع في بيانات العملاء.

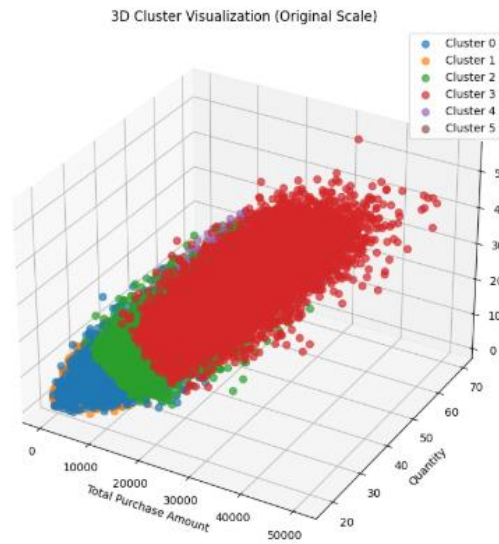
تم تقسيم العملاء إلى 6 شرائح في التطبيق النهائي وهي كانت:

- صغار العمر الذكور ذوي الصرف الكبير وكميات شراء متوسطة.
- صغار العمر الإناث ذوي الصرف القليل وكميات شراء قليلة.
- صغار العمر الإناث ذوي الصرف الكبير وكميات شراء كبيرة.
- متوسطي العمر من الذكور والإناث ذوي الصرف الكبير وكميات شراء كبيرة.
- كبار العمر الذكور ذوي الصرف الكبير وكميات شراء كبيرة.
- كبار العمر الإناث ذوي الصرف المتوسط وكميات شراء قليلة.

وكانت العلاقة بين الشرائح وفق القيم كالتالي:



الشكل 25 العلاقة بين الميزات بدلالة بعضها البعض



الشكل 26 العلاقة بين عمر العميل وكمية الشراء والمبلغ الإجمالي

الخاتمة والآفاق المستقبلية

في هذا العمل، قدّمنا منصة متكاملة تعتمد على تقنيات الذكاء الصناعي لتحليل سلوك العملاء في مواقع التجارة الإلكترونية، بهدف دعم اتخاذ القرار وتخصيص الخدمات التسويقية بشكل أكثر فعالية. تم استخدام نماذج مثل K-Means لتقسيم العملاء إلى شرائح سلوكية اعتماداً على بياناتهم، بالإضافة إلى إنشاء نموذج تنبؤي لتقدير احتمالية انسحاب العملاء. تتيح هذه المنصة لملاك المتاجر الإلكترونية فهم شرائح عملائهم بشكل أعمق، مما يساهم في تقليل معدل الانسحاب وزيادة العائد من الحملات التسويقية. حيث تم تغليف النماذج بطريقة قابلة للتوسع والاختبار بحيث يمكن مستقبلاً إضافة خدمات جديدة. أمّا بالنسبة للآفاق المستقبلية، بعد تقييم نتائج المنهجية المقترحة يمكننا استخلاص عدّة نقاط من شأنها تحسن الأداء أهمّها:

- تدريب النموذج على بيانات أكبر وذات صفات مخصصة.
- تخصيص النموذج بناءً على بيانات المستخدم من خلال ملائمة الأوزان.
- ومن الممكن إضافة ميزات جديدة تؤدي إلى تقديم تجربة أفضل للمستخدم، أهمّها:
- تقديم تحليل للأسباب المؤدية إلى انسحاب العملاء.
- تقديم العروض والتوصيات مخصصة لشرائح العملاء.

المراجع

- [1] "Chen, P. J., Bhatnagar, S., & Kowalczyk, D. K. (2022, November). CASPR: Customer Activity Sequence-based Prediction and Representation. In NeurIPS 2022 First Table Representation Workshop."
- [2] "Zhang, T., Moro, S., & Ramos, R. F. (2022). A data-driven approach to improve customer churn prediction based on telecom customer segmentation. Future Internet."
- [3] "Kumar, V., & Reinartz, W. (2016). Creating enduring customer value. Journal of marketing."
- [4] "Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30."
- [5] "Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern recognition letters."
- [6] "Kotler, P., & Keller, K. (2011). Marketing management 14th edition. prentice Hall."
- [7] "Bauer, J., & Jannach, D. (2021). Improved customer lifetime value prediction with sequence-to-sequence learning and feature-based models. ACM Transactions on Knowledge Discovery from Data (TKDD)."
- [8] "Mena, G., Coussement, K., De Bock, K. W., De Caigny, A., & Lessmann, S. (2024). Exploiting time-varying RFM measures for customer churn prediction with deep neural networks. Annals of Operations Research."
- [9] "Ahlstrand, J., Borg, A., Grahm, H., & Boldt, M. (2025). Using Transformers for B2B Contractual Churn Prediction Based on Customer Behavior Data. In International Conference on Enterprise Information Systems (ICEIS) 2025, Apr 4-6."
- [10] "E-commerce Customer Data For Behavior Analysis," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/shriyashjagtap/e-commerce-customer-for-behavior-analysis>.
- [11] "Brazilian E-Commerce Public Dataset by Olist," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce>.