NTU Journal for Administrative and Human Sciences (2026) 6(2):171-198
DOI: https://doi.org/10.56286/VOL.6NO.2

IRAQI Academic Scientific Journals

JAHS NTU

P-ISSN: 2789-1240 E-ISSN:2789-1259
NTU Journal for Administrative and Human Sciences
Available online at: https://journals.ntu.edu.iq/index.php/NTU-JMS/index

The role of data mining technology in customer churn rates in telecommunications companies: An analytical study at NBTeL in Nineveh.

Hiba Abd AL-Salam Saeed Alhilo
Northern Technical University
Administrative Technical College
hiba_abdalsalam@ntu.edu.iq

Article Informations

Received: 18-03- 2026,
Accepted: 30-03-2026
Published online: 27-06-2026

Corresponding author:
Name: Hiba Abd AL-Salam Saeed Alhilo
Affiliation : Northern Technical University/Administrative Technical College
Email:hiba_abdalsalam@ntu.edu.iq

Key Words:
Customer Churn,
Data Mining.

A B S T R A C T

This research aims to study the role of data mining techniques in predicting customer churn rates in telecommunications companies, thereby contributing to supporting customer retention decisions and improving customer relationship management strategies at NBTel in Mosul, Nineveh Governorate. The research relied on analyzing customer data through a set of variables related to demographic characteristics, usage behavior, subscription details, contracts, complaint records, technical support, marketing campaign interactions, and financial data.
The research population consisted of NBTel's telecommunications customers, while the research sample comprised (25) customers whose data were selected for analysis. To achieve the research objectives, a set of statistical tools and methods were used, and data was analyzed using SPSS software. The methods employed included descriptive statistics, analysis of variance (ANOVA), chi-squared test, correlation analysis, logistic regression, decision trees, discriminant analysis, and K-means cluster analysis.

The research reached a number of conclusions indicating the importance of employing data mining techniques and predictive models in analyzing customer behavior and identifying the factors influencing their decision to leave the service. This contributes to enhancing the ability of telecommunications companies to understand customer behavior patterns and make data-driven decisions. In light of this, the research recommended the necessity of adopting predictive models within customer relationship management systems, developing data analysis mechanisms on a regular basis, and strengthening customer retention strategies by improving service quality and developing loyalty programs and targeted offers. This will contribute to reducing customer churn rates and achieving a sustainable competitive advantage for telecommunications companies.

©2026 NTU JOURNAL FOR ADMINISTRATIVE AND HUMAN SCIENCES, NORTHERN TECHNICAL UNIVERSITY.
THIS IS AN OPEN ACCESS ARTICLE UNDER THE CC BY LICENSE: https://creativecommons.org/licenses/by/4.0/

دور تقنية التنقيب عن البيانات بمعدل فقدان الزبائن في شركات الاتصالات/ دراسة تحليلية في شركة NBTeL في مدينة نينوى

م.م. هبه عبدالسلام سعيد الحلو

الجامعة التقنية الشمالية

الكلية التقنية الادارية

hiba_abdalsalam@ntu.edu.iq

المستخلص

يهدف هذا البحث إلى دراسة دور تقنيات التنقيب عن البيانات في التنبؤ بمعدلات فقدان الزبائن في شركات الاتصالات، وبما يسهم في دعم قرارات الاحتفاظ بالزبائن وتحسين استراتيجيات إدارة علاقات الزبائن في شركة NBTel بمحافظة نينوى/ مدينة الموصل. اعتمد البحث على تحليل بيانات الزبائن من خلال مجموعة من المتغيرات المرتبطة بالخصائص الديموغرافية وسلوك الاستخدام وتفاصيل الاشتراك والعقود وسجلات الشكاوى والدعم الفني وتفاعلات الحملات التسويقية إضافة إلى البيانات المالية.

تكوّن مجتمع البحث من زبائن شركة NBTel للاتصالات، في حين تمثلت عينة البحث في (25) زبوناً تم اختيار بياناتهم لغرض التحليل. ولتحقيق أهداف البحث تم استخدام مجموعة من الأدوات والأساليب الإحصائية وتحليل البيانات بالاعتماد على برنامجSPSS ، حيث شملت الأساليب المستخدمة الإحصاءات الوصفية، وتحليل التباين(ANOVA) ، واختبار مربع كاي(Chi-Square) ، وتحليل الارتباط، والانحدار اللوجستي، وأشجار القرار، والتحليل التمييزي، فضلاً عن التحليل العنقودي (K-Means).

وتوصل البحث إلى مجموعة من الاستنتاجات التي تشير إلى أهمية توظيف تقنيات التنقيب عن البيانات والنماذج التنبؤية في تحليل سلوك الزبائن وتحديد العوامل المؤثرة في قرار مغادرتهم للخدمة، الأمر الذي يسهم في تعزيز قدرة شركات الاتصالات على فهم أنماط سلوك الزبائن واتخاذ قرارات مبنية على البيانات. وفي ضوء ذلك، أوصى البحث بضرورة تبني النماذج التنبؤية ضمن أنظمة إدارة علاقات الزبائن، وتطوير آليات تحليل البيانات بصورة دورية، وتعزيز استراتيجيات الاحتفاظ بالزبائن من خلال تحسين جودة الخدمة وتطوير برامج الولاء والعروض الموجهة بما يسهم في تقليل معدلات فقدان الزبائن وتحقيق ميزة تنافسية مستدامة لشركات الاتصالات.

الكلمات المفتاحية: فقدان العملاء، التنقيب عن البيانات

المقدمة: يشهد قطاع الاتصالات في العالم تطوراً متسارعاً نتيجة التقدم التكنولوجي واتساع استخدام الخدمات الرقمية، الأمر الذي أدى إلى زيادة حدة المنافسة بين الشركات المقدمة لهذه الخدمات. وفي ظل هذه البيئة التنافسية، أصبح الاحتفاظ بالزبائن أحد أهم

التحديات التي تواجه شركات الاتصالات، إذ إن فقدان الزبائن (Customer Churn) يؤدي إلى خسائر مالية مباشرة ويؤثر في الحصة السوقية والاستقرار التشغيلي للشركات. كما تشير العديد من الدراسات إلى أن تكلفة استقطاب زبون جديد تفوق بكثير تكلفة الاحتفاظ بالزبائن الحاليين، الأمر الذي يجعل من إدارة علاقات الزبائن والحد من مغادرتهم أولوية استراتيجية لدى المؤسسات الخدمية، ولاسيما في قطاع الاتصالات.

ومع التزايد الكبير في حجم البيانات التي تنتجها شركات الاتصالات، مثل سجلات المكالمات واستخدام الإنترنت وسجلات الفواتير والشكاوى والتفاعلات التسويقية، برزت الحاجة إلى استخدام تقنيات تحليل متقدمة قادرة على استخراج المعرفة من هذه البيانات وتحويلها إلى معلومات تدعم عملية اتخاذ القرار. ومن بين أبرز هذه التقنيات تقنية التنقيب عن البيانات(Data Mining) ، التي تُعد من الأدوات التحليلية الحديثة القادرة على اكتشاف الأنماط والعلاقات الخفية داخل قواعد البيانات الكبيرة، بما يسهم في التنبؤ بسلوك الزبائن وتحديد العوامل المؤثرة في قراراتهم.

وتتيح تقنيات التنقيب عن البيانات إمكانية بناء نماذج تنبؤية تساعد شركات الاتصالات على التعرف المبكر إلى الزبائن المعرضين لخطر مغادرة الخدمة، وذلك من خلال تحليل مجموعة من المتغيرات المرتبطة بخصائص الزبائن وأنماط استخدامهم ومستوى رضاهم عن الخدمة. ويسهم هذا النهج الاستباقي في تمكين الشركات من اتخاذ إجراءات وقائية مثل تقديم عروض مخصصة أو تحسين مستوى الخدمة أو تعزيز برامج الولاء، بما يساعد على تقليل معدلات فقدان الزبائن وتعزيز استمرارية العلاقة معهم.

المحور الأول: منهجية البحث

اولا: مشكلة البحث : في ظل التنافس المتزايد في قطاع الاتصالات والتطور المستمر في الخدمات الرقمية، تواجه شركات الاتصالات تحدياً مهماً يتمثل في تزايد معدلات فقدان الزبائن وانتقالهم إلى شركات منافسة. وتؤثر هذه الظاهرة بشكل مباشر في إيرادات الشركات واستقرارها السوقي، الأمر الذي يستدعي تبني أساليب تحليلية متقدمة تساعد على فهم سلوك الزبائن والتنبؤ باحتمالية مغادرتهم للخدمة قبل حدوثها فعلياً. وفي هذا السياق برزت تقنيات التنقيب عن البيانات كأحد الأدوات التحليلية الحديثة التي يمكن توظيفها لاكتشاف الأنماط والعلاقات الكامنة في البيانات الضخمة للزبائن، بما يسهم في

بناء نماذج تنبؤية تدعم اتخاذ القرار وتساعد في وضع استراتيجيات فعالة للحد من فقدان الزبائن في شركات الاتصالات.

وانطلاقاً من ذلك يمكن صياغة مشكلة البحث في التساؤل الرئيس الآتي: كيف يمكن توظيف تقنيات التنقيب عن البيانات والأساليب الإحصائية في التنبؤ بفقدان الزبائن في شركات الاتصالات، بما يسهم في دعم استراتيجيات الاحتفاظ بالزبائن في شركة NBTel بمحافظة نينوى/ مدينة الموصل؟

1. ما طبيعة العلاقة بين بعض الخصائص الديموغرافية والسلوكية والمالية للزبائن واحتمالية فقدانهم للخدمة؟

2. ما أهم العوامل المؤثرة في قرار الزبائن بمغادرة خدمات شركة الاتصالات؟

3. إلى أي مدى يمكن لتقنيات التنقيب عن البيانات والنماذج التنبؤية أن تسهم في تحديد الزبائن الأكثر عرضة لفقدان الخدمة؟

4. ما مدى فاعلية استخدام الأساليب الإحصائية وتقنيات التنقيب عن البيانات في دعم قرارات إدارة الشركة المتعلقة بالاحتفاظ بالزبائن؟

5. كيف يمكن توظيف نتائج النماذج التنبؤية في تطوير استراتيجيات فعالة للحد من فقدان الزبائن في قطاع الاتصالات؟

ثانيا: أهمية البحث:

1- الأهمية العلمية

- يساهم البحث في إثراء الأدبيات العلمية المتعلقة بتطبيق تقنيات التنقيب عن البيانات في مجال إدارة علاقات الزبائن والتنبؤ بفقدان الزبائن في قطاع الاتصالات.
- يسلط الضوء على دور الأساليب الإحصائية وتقنيات التنقيب عن البيانات في تحليل سلوك الزبائن واكتشاف الأنماط والعلاقات المؤثرة في قرارات مغادرة الخدمة.

- يقدم إطاراً علمياً يوضح كيفية توظيف النماذج التنبؤية في دراسة ظاهرة فقدان الزبائن، مما يدعم الدراسات المستقبلية في مجالات تحليل البيانات وذكاء الأعمال.
- يربط بين المفاهيم النظرية للتنقيب عن البيانات والتطبيقات العملية في قطاع الاتصالات، الأمر الذي يسهم في توسيع مجالات البحث في هذا المجال.

2- الأهمية التطبيقية

- يساعد شركات الاتصالات على فهم العوامل المؤثرة في فقدان الزبائن من خلال تحليل البيانات المتاحة لديها.
- يساهم في دعم متخذي القرار في شركات الاتصالات من خلال توفير أدوات تحليلية تساعد على التنبؤ بسلوك الزبائن واتخاذ قرارات مبنية على البيانات.
- يوفر نموذجاً تحليلياً يمكن الاستفادة منه في تحديد الزبائن الأكثر عرضة لمغادرة الخدمة ووضع استراتيجيات مناسبة للاحتفاظ بهم.
- يسهم في تحسين كفاءة إدارة علاقات الزبائن وتطوير برامج ولاء وعروض تسويقية موجهة تقلل من معدلات فقدان الزبائن.
- يدعم توجه المؤسسات نحو استخدام التحليلات المتقدمة والتنقيب عن البيانات في تحسين الأداء وتحقيق ميزة تنافسية مستدامة في سوق الاتصالات

ثالثا: أهداف البحث: يسعى هذا البحث إلى دراسة ظاهرة مغادرة العملاء في شركة NBTel من خلال توظيف الأساليب الإحصائية وتقنيات التعلم الآلي، بهدف بناء نموذج علمي يساعد في التنبؤ بسلوك العملاء ودعم اتخاذ القرار الإداري في مجال الاحتفاظ بهم. ولتحقيق ذلك، تم تحديد الأهداف الآتية:

1. تطبيق تقنيات التحليل الإحصائي المتقدم لاستكشاف العلاقات بين متغيرات البحث
2. تطوير نموذج تنبؤي باستخدام الانحدار اللوجستي وأشجار القرار
3. تحديد أهم العوامل المؤثرة في قرار مغادرة العملاء
4. تقييم دقة النماذج التنبؤية باستخدام مقاييس الأداء المعيارية
5. تقديم توصيات عملية لاستراتيجيات الاحتفاظ بالعملاء.

رابعا: منهج البحث: اعتمد البحث على المنهج الوصفي التحليلي لملاءمته لطبيعة الدراسة التي تهدف إلى تحليل بيانات الزبائن واكتشاف العلاقات بين المتغيرات المختلفة

المرتبطة بسلوك الزبائن واحتمالية مغادرتهم للخدمة. ويقوم هذا المنهج على وصف الظاهرة محل الدراسة وتحليلها باستخدام الأساليب الإحصائية وتقنيات التنقيب عن البيانات من أجل التوصل إلى نتائج يمكن الاستفادة منها في دعم عملية اتخاذ القرار في شركات الاتصالات.

خامسا: أساليب جمع البيانات والمعلومات: اعتمد البحث على مجموعة من الأساليب لجمع البيانات والمعلومات، تمثلت في الآتي:

- البيانات الثانوية :تم الحصول عليها من قواعد البيانات المتوفرة لدى شركة NBTel، والتي تضمنت بيانات الزبائن المرتبطة بالخصائص الديموغرافية وأنماط الاستخدام وتفاصيل الاشتراك والعقود وسجلات الشكاوى والدعم الفني والبيانات المالية.
- المصادر العلمية :تم الاعتماد على الكتب والمراجع العلمية والدوريات الأكاديمية والدراسات السابقة ذات العلاقة بموضوع التنقيب عن البيانات والتنبؤ بفقدان الزبائن في قطاع الاتصالات.
- المصادر الإلكترونية :تم الاستفادة من قواعد البيانات العلمية والمواقع الأكاديمية المتخصصة للحصول على معلومات حديثة تدعم الإطار النظري للبحث.

سادسا: مجتمع البحث وعينته: يتكون مجتمع البحث من زبائن شركة NBTel للاتصالات في محافظة نينوى/ مدينة الموصل. أما عينة البحث فقد تمثلت في بيانات (25) زبوناً تم اختيارهم من قاعدة بيانات الشركة لغرض التحليل الإحصائي وتطبيق تقنيات التنقيب عن البيانات.

سابعا: الأدوات والأساليب الإحصائية المستخدمة: اعتمد البحث على برنامج SPSS لتحليل البيانات وتطبيق مجموعة من الأساليب الإحصائية وتقنيات التنقيب عن البيانات، والتي شملت:

- الإحصاءات الوصفية.
- تحليل التباين.(ANOVA)
- اختبار مربع كاي.(Chi-Square)
- تحليل الارتباط.(Correlation Analysis)

- الانحدار اللوجستي.(Logistic Regression)
- أشجار القرار.(Decision Trees)
- التحليل التمييزي.(Discriminant Analysis)
- التحليل العنقودي.(K-Means Clustering)

ثامنا: مخطط البحث الفرضي: يعتمد البحث على مخطط فرضي يوضح العلاقة بين المتغيرات المستقلة والمتغير التابع. حيث تتمثل المتغيرات المستقلة في مجموعة من العوامل المرتبطة بخصائص الزبائن وسلوكهم، مثل:

- مستوى نشاط الزبون.
- نوع العقد (شهري/سنوي).
- عدد الشكاوى المقدمة.
- تأخر الدفع.
- مستوى الرضا عن خدمات الدعم الفني.

أما المتغير التابع فيتمثل في:

- فقدان الزبائن (Customer Churn) أي احتمال مغادرة الزبون لخدمات الشركة.

ويفترض المخطط الفرضي أن المتغيرات المستقلة تؤثر بدرجات متفاوتة في المتغير التابع، حيث يتم استخدام تقنيات التنقيب عن البيانات والأساليب الإحصائية لاختبار هذه العلاقات وتحليل تأثيرها في التنبؤ بفقدان الزبائن.

تاسعا: حدود البحث : تتحدد هذه الدراسة بمجموعة من الحدود التي ينبغي أخذها في الاعتبار عند تفسير النتائج وتعميمها، وتتمثل في الآتي:

1. **الحدود المكانية:** اقتصرت الدراسة على شركة NBTel للاتصالات العاملة في محافظة نينوى، مما قد يحد من إمكانية تعميم النتائج على شركات اتصالات أخرى أو مناطق جغرافية مختلفة.

2. **الحدود الزمنية:** اعتمدت الدراسة على بيانات متاحة لفترة زمنية محددة من 2025/10/1 الى 2026/1/5 ، وبالتالي فإن النتائج تعكس واقع تلك الفترة دون فترات زمنية أخرى.
3. **الحدود البشرية (العينة):** شملت الدراسة عينة مكونة من 25 عميلاً فقط.

عاشرا: فرضيات البحث : انطلاقاً من مشكلة البحث وأهدافه، تم صياغة الفرضيات الآتية لاختبار العلاقات بين بعض المتغيرات المستقلة واحتمال مغادرة العميل:

1- **الفرضية الاولى:** توجد علاقة ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) بين مستوى نشاط العميل واحتمال مغادرته للخدمة.

2- **الفرضية الثانية** : توجد علاقة ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) بين نوع العقد (شهري/سنوي) واحتمال مغادرة العميل.

احدى عشر: الدراسات السابقة :لقد حظيت ظاهرة مغادرة العملاء (Customer Churn) في قطاع الاتصالات باهتمام كبير في الأدبيات الأكاديمية خلال العقدين الأخيرين، نظرًا لتأثيرها المباشر على استقرار الإيرادات واستراتيجيات الاحتفاظ بالعملاء. ركزت هذه الدراسات على تحليل سلوك العملاء باستخدام تقنيات إحصائية وتقنيات تعلم آلي متقدمة مثل الانحدار اللوجستي، أشجار القرار، الشبكات العصبية، وآلات الدعم المتجهية (SVM)، بهدف التنبؤ بالعملاء الأكثر عرضة للمغادرة قبل حدوث ذلك. كما اهتمت بعض الدراسات بدمج البيانات الديموغرافية، السلوكية، المالية والتسويقية لتعزيز دقة النماذج التنبؤية.

يأتي الجدول أدناه ليعرض أبرز الدراسات السابقة مع أهدافها ونتائجها، لتوفير إطار علمي يساعد على فهم مجالات البحث الحالية والفجوات التي يسعى هذا البحث إلى معالجتها.

جدول (1) الدراسات السابقة ذات العلاقة

العنصر	المحتوى
عنوان الدراسة	(V. Kavitha, G. Hemanth Kumar, S. V Mohan Kumar &M. Harish, 2020) – التنبؤ بانقطاع العملاء في قطاع الاتصالات باستخدام تقنيات التعلم الآلي
مشكلة الدراسة	تواجه شركات الاتصالات صعوبة في التنبؤ المبكر بالزبائن الذين قد يتركون الخدمة، مما يؤدي إلى خسائر مالية وزيادة تكاليف استقطاب زبائن جدد.
أهم الأهداف	تطوير نماذج تعلم آلي قادرة على التنبؤ بفقدان الزبائن وتحليل العوامل المؤثرة في قرار مغادرتهم للخدمة.
مجتمع الدراسة وعينتها	بيانات زبائن في قطاع الاتصالات تم تحليلها باستخدام خوارزميات تعلم آلي مختلفة.
أهم الاستنتاجات	أظهرت النتائج أن نماذج **الانحدار اللوجستي وأشجار القرار** حققت أداءً جيداً في التنبؤ بفقدان الزبائن.
أهم المقترحات أو التوصيات	توصي الدراسة باستخدام تقنيات التعلم الآلي بشكل مستمر في تحليل بيانات الزبائن لدعم قرارات الاحتفاظ بهم.

جدول (2) الدراسات السابقة ذات العلاقة

العنصر	المحتوى
عنوان الدراسة	دراسة مقارنة لتقنيات استخراج البيانات للتنبؤ بانقطاع العملاء في قطاع الاتصالات – (Ahmad, A. K., Jafar, A., & Aljoumaa, K. 2019)
مشكلة الدراسة	عدم وضوح أي من تقنيات التنقيب عن البيانات أكثر كفاءة في التنبؤ بفقدان الزبائن في قطاع الاتصالات.
أهم الأهداف	مقارنة مجموعة من تقنيات التنقيب عن البيانات والتعلم الآلي لتحديد أكثرها دقة في التنبؤ بمغادرة الزبائن للخدمة.
مجتمع الدراسة وعينتها	بيانات زبائن من شركات الاتصالات تم تحليلها باستخدام عدة خوارزميات تعلم آلي.
أهم الاستنتاجات	أظهرت النتائج أن خوارزميات **Random Forest** و**Decision Tree** حققت أفضل أداء في التنبؤ بفقدان الزبائن.
أهم المقترحات أو التوصيات	توصي الدراسة بدمج عدة نماذج تنبؤية لتحسين دقة التنبؤ ودعم قرارات إدارة علاقات الزبائن.

جدول (3) الدراسات السابقة ذات العلاقة

العنصر	المحتوى
عنوان الدراسة	تحليل الشبكات الاجتماعية للتنبؤ بانقطاع العملاء في قطاع الاتصالات – (Verbeke, Martens & Baesens, 2014)
مشكلة الدراسة	محدودية النماذج التقليدية في التنبؤ بفقدان الزبائن دون الأخذ في الاعتبار تأثير العلاقات الاجتماعية بين الزبائن.
أهم الأهداف	دراسة دور تحليل الشبكات الاجتماعية في تحسين نماذج التنبؤ بفقدان الزبائن في شركات الاتصالات.
مجتمع الدراسة وعينتها	بيانات زبائن في قطاع الاتصالات مع تحليل العلاقات والتفاعلات بينهم.
أهم الاستنتاجات	دمج تحليل الشبكات الاجتماعية مع نماذج التعلم الآلي أدى إلى تحسين دقة التنبؤ بفقدان الزبائن.
أهم المقترحات أو التوصيات	توصي الدراسة بدمج البيانات السلوكية والاجتماعية مع تقنيات التنقيب عن البيانات للحصول على نماذج تنبؤية أكثر دقة.

جدول (4) الدراسات السابقة ذات العلاقة

العنصر	المحتوى
عنوان الدراسة	التنبؤ بانقطاع العملاء في خدمات الاشتراك: تطبيق لآلات المتجهات الداعمة – (Coussement & Van den Poel, 2008)
مشكلة الدراسة	الحاجة إلى تطوير نماذج تنبؤية أكثر دقة لتحديد الزبائن المعرضين لمغادرة خدمات الاشتراك.
أهم الأهداف	اختبار كفاءة خوارزمية **آلات الدعم المتجهية (SVM)** في التنبؤ بفقدان الزبائن في خدمات الاشتراك.
مجتمع الدراسة وعينتها	بيانات زبائن في خدمات الاشتراك تم تحليلها باستخدام خوارزميات التعلم الآلي.
أهم الاستنتاجات	أثبت نموذج **SVM** قدرة عالية على التنبؤ بفقدان الزبائن مقارنة ببعض الأساليب الإحصائية التقليدية.
أهم المقترحات أو التوصيات	توصي الدراسة بدمج البيانات الديموغرافية وسلوك الاستخدام في النماذج التنبؤية لتحسين دقة النتائج.

المصدر :من اعداد الباحثة

المحور الثاني: الإطار النظري

اولا: مفهوم التنقيب عن البيانات

تنقيب البيانات (Data Mining) هو عملية استخراج المعرفة المفيدة من مجموعات البيانات الضخمة باستخدام أساليب حسابية متقدمة. وفقاً لـ (Fayyad & other،

1996) يُعرّف تنقيب البيانات بأنه "عملية غير تافهة لتحديد الأنماط الصحيحة والجديدة والمفيدة والمفهومة في البيانات" (Knowledge Discovery in Databases).

وترى الباحثة أن تنقيب البيانات يمثل أداة تحليلية متقدمة تمكّن المؤسسات من تحويل الكم الهائل من البيانات المتراكمة إلى معرفة قابلة للتطبيق تدعم اتخاذ القرار الاستراتيجي. ومع التزايد المستمر في حجم البيانات الناتجة عن الأنظمة الرقمية، أصبحت تقنيات تنقيب البيانات ضرورة أساسية لاكتشاف الأنماط الخفية والعلاقات غير الظاهرة التي يصعب اكتشافها بالطرق التقليدية للتحليل.

كما ترى الباحثة أن أهمية تنقيب البيانات لا تكمن فقط في استخراج الأنماط من البيانات، بل في دمج نتائجه ضمن إطار أوسع لاكتشاف المعرفة ((KDD، حيث يتم تفسير النتائج وربطها بسياق المشكلة المدروسة، الأمر الذي يسهم في تحويل البيانات الخام إلى معرفة مفيدة يمكن توظيفها في تحسين الأداء التنظيمي وتطوير الاستراتيجيات المستقبلية.(Sadat, Qiam & Kohistani 2025) .

أن عملية اكتشاف المعرفة في قواعد البيانات (Knowledge Discovery in Databases – KDD)تعد إطاراً منهجياً متكاملاً لتحويل البيانات الخام إلى معرفة قابلة للاستخدام، حيث تبدأ بجمع البيانات وإعدادها، ثم تطبيق تقنيات تنقيب البيانات لاستخراج الأنماط، يلي ذلك تفسير النتائج وتقديمها بطريقة تدعم اتخاذ القرار. وبذلك فإن تنقيب البيانات يمثل المرحلة التحليلية الأساسية أو القلب التقني داخل عملية اكتشاف المعرفة الأوسع.

كما ترى الباحثة أن القيمة الحقيقية لتنقيب البيانات لا تتحقق إلا عند دمجه ضمن إطار اكتشاف المعرفة ((KDD، حيث يتم تفسير الأنماط المكتشفة وربطها بسياق المشكلة البحثية أو التطبيق العملي. ومن هذا المنطلق، فإن نجاح مشاريع تحليل البيانات لا يعتمد فقط على استخدام تقنيات التنقيب المتقدمة، بل يتطلب أيضاً إدارة متكاملة لمراحل اكتشاف المعرفة لضمان تحويل البيانات الخام إلى معلومات ذات قيمة تسهم في دعم اتخاذ القرار وتحسين الأداء المؤسسي.

هذا التعريف يشدد على أربعة معايير أساسية للمعرفة المستخرجة:

1. الصحة: (Validity) يجب أن تكون الأنماط المكتشفة صحيحة على بيانات جديدة بدرجة معينة من الثقة

2. الجدة: (Novelty) يجب أن تكون الأنماط جديدة وغير معروفة مسبقاً

3.الفائدة : (Usefulness) يجب أن تؤدي الأنماط إلى إجراءات مفيدة حسب معايير المستخدم

4.المفهومية: (Understandability) يجب أن تكون الأنماط قابلة للفهم والتفسير بشكل واضح

من جهة أخرى، يعرّف (Han & Kamber,2011) تنقيب البيانات بأنه "عملية اكتشاف الأنماط المثيرة للاهتمام والمعرفة من كميات كبيرة من البيانات." يؤكد هذا التعريف على الطبيعة متعددة التخصصات لتنقيب البيانات، حيث يجمع بين علوم الحاسوب، الإحصاء، التعلم الآلي، وأنظمة قواعد البيانات

ثانيا: تنقيب البيانات مقابل اكتشاف المعرفة

من المهم التمييز بين مصطلحي تنقيب البيانات (Data Mining) واكتشاف المعرفة في قواعد البيانات(Knowledge Discovery in Databases – KDD). وفقاً لـ (Fayyad & other، 1996) فإن KDD هو العملية الشاملة التي تتضمن عدة مراحل، بينما تنقيب البيانات هو مرحلة محددة ضمن هذه العملية. يمكن وصف العلاقة كالتالي:

KDD = اختيار البيانات + المعالجة المسبقة + التحويل + تنقيب البيانات + التفسير/التقييم

بمعنى آخر، تنقيب البيانات هو القلب التقني لعملية KDD الأوسع، حيث يتم تطبيق الخوارزميات الفعلية لاستخراج الأنماط من البيانات المُعدة.

ثالثا: منهجية :CRISP-DM المعيار الصناعي لتنقيب البيانات

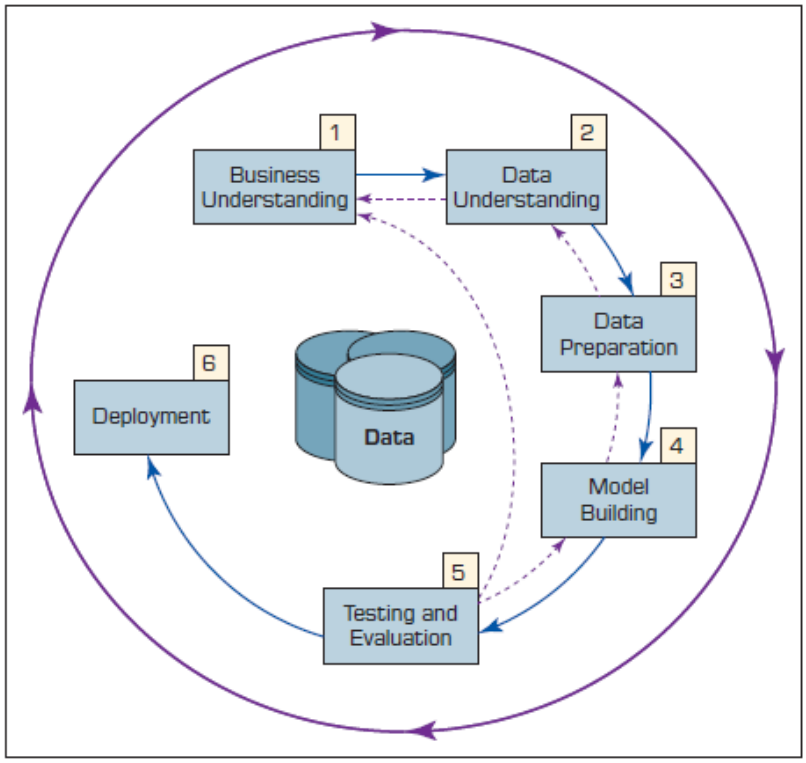
منهجية (CRISP-DM (Cross-Industry Standard Process for Data Mining هي إطار عمل معياري يصف المراحل الرئيسية لمشاريع تنقيب البيانات. تم تطويرها عام 1996 من قبل اتحاد من الشركات الرائدة يضم DaimlerChrysler،

SPSS، NCR Corporation ، وOHRA، بتمويل من المفوضية الأوروبية ضمن برنامج ESPRIT (Wirth & Hipp, 2000).

نُشرت النسخة الأولى من المنهجية رسمياً في عام 1999، وسرعان ما أصبحت المعيار الصناعي الفعلي (de facto standard) لمشاريع تنقيب البيانات. وفقاً لمراجعة أجراها (Martínez-Plumed & other, 2021)، تظل CRISP-DM المنهجية الأكثر استخداماً في الصناعة، حيث تُطبق في أكثر من 60% من مشاريع تنقيب البيانات عالمياً.

المراحل الست لمنهجية CRISP-DM المستخدمة في الحالة الدراسية

تتكون منهجية CRISP-DM من ست مراحل رئيسية مترابطة، كما هو موضح في الشكل (1) نموذج المراحل ليست خطية بالضرورة، بل تتميز بطبيعة دورية تسمح بالتكرار والرجوع بين المراحل حسب الحاجة:



الشكل (1) :عملية CRISP-DM للتنقيب عن البيانات

المرحلة الأولى: فهم الأعمال (Business Understanding)

تركز هذه المرحلة على فهم أهداف المشروع ومتطلباته من منظور الأعمال، ثم ترجمة هذه المعرفة إلى تعريف لمشكلة تنقيب البيانات وخطة أولية لتحقيق الأهداف. المهام الرئيسية تشمل:

الجدول (2):مرحلة فهم الاعمال حسب منهجية CRISP-DM

العنصر	التطبيق في بحث NBTel
هدف العمل	تقليل معدلات مغادرة العملاء في شركة NBTel
المشكلة	خسارة 40% من العملاء (فئة خطر عالي)
الهدف التقني	بناء نموذج تنبؤي بدقة ≥75%
الموارد	بيانات 25 عميل، 34 متغير، SPSS

المصدر :من اعداد الباحثة

المرحلة الثانية: فهم البيانات ((Data Understanding

تبدأ هذه المرحلة بجمع البيانات الأولية وتتقدم نحو الألفة بالبيانات، وتحديد مشاكل جودة البيانات، واكتشاف الرؤى الأولية. المهام تشمل:

الجدول (3):مرحلة فهم البيانات حسب منهجية CRISP-DM

المهمة	التنفيذ
جمع البيانات	بيانات منNBTel 25:سجل عميل
وصف البيانات	34 متغير (ديموغرافي، سلوكي، مالي)
استكشاف البيانات	الإحصاءات الوصفية
جودة البيانات	التحقق من القيم المفقودة والشاذة

المصدر: من اعداد الباحثة

المرحلة الثالثة: تحضير البيانات ((Data Preparation

تغطي هذه المرحلة جميع الأنشطة اللازمة لبناء مجموعة البيانات النهائية من البيانات الأولية. غالباً ما تكون المرحلة الأكثر استهلاكاً للوقت، حيث قد تستغرق 50–70% من إجمالي وقت المشروع (Pyle, 1999) المهام تشمل:

الجدول (4):مرحلة تحضير البيانات حسب منهجية CRISP–DM

المهمة	التنفيذ
جمع البيانات	بيانات منNBTel 25سجل عميل
وصف البيانات	34 متغير (ديموغرافي، سلوكي، مالي)
استكشاف البيانات	الإحصاءات الوصفية
جودة البيانات	التحقق من القيم المفقودة والشاذة

المصدر: من اعداد الباحثة

المرحلة الرابعة: النمذجة ((Modeling

في هذه المرحلة، يتم اختيار وتطبيق تقنيات النمذجة المختلفة، وضبط معاملاتها للحصول على القيم المثلى. عادةً ما توجد عدة تقنيات لنفس نوع مشكلة تنقيب البيانات، وبعضها يتطلب صيغاً محددة للبيانات. المهام تشمل:

الجدول (5):مرحلة النمذجة حسب منهجية CRISP–DM

العنصر	التنفيذ
تقنيات النمذجة	1. الانحدار اللوجستي (رئيسي) 2. أشجار القرار 3. التحليل التمييزي
تصميم الاختبار	Train-Test Split + Cross-Validation
بناء النموذج	SPSS Logistic Regression
المتغيرات المدخلة	8 متغيرات مستقلة
المتغير التابع	WillChurn (ثنائي)

المصدر:من اعداد الباحثة

المرحلة الخامسة: التقييم ((Evaluation

في هذه المرحلة، يتم تقييم النموذج من منظور الأعمال بدلاً من المنظور التقني فقط. الهدف هو التأكد من أن النموذج يحقق أهداف العمل بشكل كافٍ. المهام تشمل:

الجدول (6):مرحلة التقييم حسب منهجية CRISP-DM

المقياس	النتيجة	التفسير
Accuracy	88.0%	ممتاز (> 75% المطلوب)
AUC	0.910	ممتاز (0.90-1.00)
Precision	93.3%	دقة عالية جداً
Recall	87.5%	تغطية ممتازة

المصدر :من اعداد الباحثة

المرحلة السادسة: النشر ((Deployment

بناء النموذج ليس نهاية المشروع. حتى إذا كان الهدف هو زيادة المعرفة، فإن المعرفة المكتسبة يجب تنظيمها وعرضها بطريقة يمكن للعميل استخدامها. المهام تشمل:

- تخطيط النشر: تطوير استراتيجية شاملة لنشر النموذج
- التخطيط للمراقبة والصيانة: وضع خطة لمراقبة أداء النموذج بمرور الوقت
- إنتاج التقرير النهائي: توثيق كامل للمشروع ونتائجه
- مراجعة المشروع: تقييم ما حدث بشكل صحيح وما يمكن تحسينه

رابعا: مزايا وتحديات CRISP-DM: مثل منهجية CRISP-DM إطاراً معيارياً شائعاً لتحليل البيانات واستخراج المعرفة، حيث تقدم مزايا واضحة في التنظيم والتوثيق والمرونة، لكنها تواجه تحديات مرتبطة بالتعامل مع البيانات الضخمة وسرعة التنفيذ ونقص التغطية لمرحلة النشر.

المزايا:

1. عامة ومستقلة عن الصناعة: يمكن تطبيقها في أي قطاع
2. موثقة جيداً: دليل مستخدم شامل بأكثر من 200 صفحة
3. منطقية ومنظمة: تتبع تسلسلاً منطقياً للمراحل
4. مرنة وتكرارية: تسمح بالرجوع والتكرار بين المراحل
5. مدعومة على نطاق واسع: أدوات وموارد عديدة متاحة

التحديات:

1. غير محدثة للبيانات الضخمة: تم تطويرها قبل عصر Big Data
2. ليست منهجية إدارة مشروع كاملة: تركز على الجوانب التقنية أكثر
3. قد تكون بطيئة: الطبيعة الشاملة قد تبطئ المشاريع السريعة
4. نقص في مرحلة النشر: معظم الدراسات لا تغطي النشر الكامل

خامسا: تقنيات التنقيب عن البيانات المستخدمة في البحث

تنقيب البيانات يشمل مجموعة واسعة من التقنيات والخوارزميات، كل منها مناسب لأنواع معينة من المشاكل والبيانات.

التقنية الاولى: التحليل الوصفي (Descriptive Analysis)

التحليل الوصفي هو عملية إحصائية تهدف إلى تلخيص وتنظيم وعرض البيانات بطريقة ذات معنى، لتوفير فهم أفضل لخصائص مجموعة البيانات. يشمل حساب المقاييس الإحصائية الأساسية والتوزيعات التكرارية. (Hair et al., 2019)

ترى الباحثة أن المراحل الست لمنهجية CRISP-DMتمثل إطارًا منهجيًا متكاملاً لإدارة مشاريع تنقيب البيانات وتحويل البيانات الخام إلى معرفة قابلة للتطبيق. إذ تبدأ العملية بمرحلة فهم الأعمال التي تُعد الأساس في تحديد المشكلة البحثية وربطها بالأهداف التنظيمية، وهو ما يضمن توجيه الجهود التحليلية نحو تحقيق قيمة عملية للمؤسسة. تلي ذلك مرحلة فهم البيانات التي تسهم في استكشاف خصائص البيانات المتاحة وتحديد المشكلات المرتبطة بجودتها أو اكتمالها، مما يساعد في بناء تصور واضح لطبيعة البيانات قبل البدء بالتحليل.

كما تؤكد الباحثة أن مرحلة تحضير البيانات تعد من أكثر المراحل أهمية واستهلاكًا للوقت في مشاريع تحليل البيانات، نظرًا لما تتطلبه من عمليات تنظيف ودمج وتحويل للبيانات بهدف إعدادها بالشكل المناسب لعمليات النمذجة. أما مرحلة النمذجة فتمثل الجانب التحليلي الأساسي في عملية التنقيب عن البيانات، حيث يتم تطبيق الخوارزميات الإحصائية وتقنيات التعلم الآلي لاكتشاف الأنماط والعلاقات الكامنة في البيانات.

وفيما يتعلق بمرحلة التقييم، ترى الباحثة أنها مرحلة حاسمة لضمان أن النماذج المطورة تحقق الأهداف المطلوبة من منظور تقني ومنظور الأعمال، إذ يتم فيها اختبار دقة

النماذج ومدى قدرتها على التنبؤ واتخاذ القرار. وأخيرًا تأتي مرحلة النشر التي تمثل التطبيق العملي لنتائج التحليل، حيث يتم تحويل المعرفة المستخرجة إلى أدوات أو تقارير أو أنظمة دعم قرار يمكن الاستفادة منها في تحسين الأداء المؤسسي.

وبناءً على ذلك، ترى الباحثة أن منهجية CRISP–DMتوفر إطارًا عمليًا مرنًا ومترابطًا يساهم في تنظيم مراحل تحليل البيانات بشكل منهجي، كما أنها تساعد الباحثين والمؤسسات على الاستفادة من البيانات المتاحة وتحويلها إلى معرفة تدعم اتخاذ القرارات وتحسين الكفاءة التشغيلية.

المحور الثالث: الاطار العملي

اولا: المقاييس المستخدمة في البحث: اعتمدت الدراسة على مجموعة من المقاييس الإحصائية لوصف وتحليل البيانات، وتشمل:

أ. مقاييس النزعة المركزية:

- **المتوسط الحسابي: (Mean) متوسط جميع القيم**
- **الوسيط : (Median) القيمة الوسطى عند ترتيب البيانات**
- **المنوال: (Mode) القيمة الأكثر تكراراً**

ب. مقاييس التشتت:

- **الانحراف المعياري : (SD) مقياس تشتت البيانات حول المتوسط**
- **المدى : (Range) الفرق بين أكبر وأصغر قيمة**
- **التباين: (Variance) مربع الانحراف المعياري**

ج. التوزيعات التكرارية:

- **التكرارات : (Frequencies) عدد مرات ظهور كل فئة**

- **النسب المئوية: (Percentages) نسبة كل فئة من المجموع**

ثانياً: تقنيات التحليل الإحصائي والتنبؤ المستخدمة في البحث

1. الانحدار اللوجستي**Logistic Regression)) : هو تقنية إحصائية للتنبؤ بمتغير تابع ثنائي** (binary) **بناءً على متغير أو أكثر من المتغيرات المستقلة. يستخدم دالة اللوجيت** (logit function) **لتحويل الاحتمالات الخطية إلى نطاق بين (0 و 1 .(** (Hosmer, Lemeshow & Sturdivant, 2013)

الصيغة الرياضية:

$$\ln(p / (1-p)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$$

نسب الأرجحية (Odds Ratios – OR):

نسبة الأرجحية تقيس التغير في احتمالية وقوع الحدث عند تغيير المتغير المستقل بوحدة واحدة:

$$OR = e^{\beta}$$

2. **أشجار القرار ((Decision Trees:** هي نموذج تنبؤي يستخدم هيكلاً شجرياً لتمثيل القرارات ونتائجها المحتملة. كل عقدة داخلية تمثل اختباراً على سمة، كل فرع يمثل نتيجة الاختبار، وكل عقدة ورقية تمثل تصنيفاً أو قراراً نهائياً. (Breiman et al., 1984)
3. **التحليل التمييزي ((Discriminant Analysis:**هو تقنية إحصائية تستخدم لتصنيف الحالات إلى مجموعتين أو أكثر بناءً على مجموعة من المتغيرات المستقلة. يبحث عن الدالة الخطية للمتغيرات التي تفصل بشكل أفضل بين المجموعات. ((Fisher, 1936; Tabachnick & Fidell, 2013
4. **التجميع العنقودي K–Means Clustering** هي خوارزمية تجميع غير إشرافية (unsupervised) تقسم مجموعة البيانات إلى K عناقيد (clusters) بناءً على التشابه. الهدف هو تجميع نقاط البيانات المتشابهة معاً وفصل النقاط غير المتشابهة.((MacQueen, 1967; Hartigan & Wong, 1979

5. **تحليل الارتباط ((Correlation Analysis:** هو تقنية إحصائية تقيس قوة واتجاه العلاقة الخطية بين متغيرين كميين. النتيجة هي معامل الارتباط الذي يتراوح بين −1 و + 1 . ((Pearson, 1896

معادلة بيرسون:

$$r = \Sigma[(X_i - \bar{X})(Y_i - \bar{Y})] / \sqrt{[\Sigma(X_i - \bar{X})^2 \times \Sigma(Y_i - \bar{Y})^2]}$$

حيث:

X_i, Y_i =قيم المتغيرات

$\bar{X}, \bar{Y}$ = متوسطات المتغيرات

ثانيا: اختبارات الفرضيات الإحصائية

اختبار الفرضيات هو إجراء إحصائي لاتخاذ قرارات حول معلمات المجتمع بناءً على بيانات العينة. يتضمن صياغة فرضيتين (صفرية وبديلة) واستخدام البيانات لتحديد أيهما أكثر ترجيحاً.

- **الاختبارات المستخدمة في البحث:**

أ. اختبار مربع كاي (Chi-Square Test of Independence):

الاستخدام: اختبار العلاقة بين متغيرين فئويين

المعادلة:

$$\chi^2 = \Sigma [(O_i - E_i)^2 / E_i]$$

حيث:

O_i = التكرار المُشاهد (Observed)

E_i = التكرار المتوقع (Expected)

ب. تحليل التباين الأحادي (One-Way ANOVA):

الاستخدام: مقارنة متوسطات 3 مجموعات أو أكثر

يقارن:

- التباين بين المجموعات (Between-group variance)
- التباين داخل المجموعات (Within-group variance)

F إحصائية:

F = (التباين بين المجموعات) / (التباين داخل المجموعات)

إذا كان F كبيراً ⟵ فروق معنوية بين المجموعات

الجدول (7) تقنيات التنقيب في البيانات المستخدمة في البحث

التقنية	النوع	الهدف	المتغير التابع	المخرجات الرئيسية
الاحصاءات الوصفية	وصفي	فهم البيانات	–	متوسطات, انحرافات,تكرارات
الانحدار اللوجستي	تنبؤي	التصنيف	ثنائي(0/1)	احتمالات, OR, Accuracy=88%
اشجار القرار	تنبؤي	التصنيف	ثنائي/متعدد	قواعد،Accuracy=80%
التحليل التمييزي	تنبؤي	التصنيف	ثنائي/متعدد	دالةتمييزية,Accuracy=78%
K-Means	وصفي	التجميع	–	3مراكز , عناقيد
الارتباط	وصفي	العلاقات	–	معاملاتr (−1 إلى +1)
Chi-Square	استدلالي	العلاقات	فئوي	χ^2, p-value
ANOVA	استدلالي	المقارنات	رقمي	F, p-value

المصدر: من اعداد الباحثة

رابعا: تقييم نماذج التنقيب في البيانات

تقييم النماذج خطوة حاسمة لضمان أن النموذج المطور يحقق الأهداف المطلوبة ويؤدي بشكل جيد على بيانات جديدة غير مرئية.

1. استراتيجيات التحقق

- Holdout Validation: تقسيم البيانات إلى مجموعتي تدريب واختبار (مثلاً 30/70 أو 20/80)

- K–Fold Cross–Validation : تقسيم البيانات إلى k أجزاء، استخدام k−1 للتدريب و1 للاختبار، وتكرار العملية k مرات

2.مقاييس التقييم لمشاكل التصنيف

تعتمد مقاييس التقييم على مصفوفة الارتباك (Confusion Matrix) التي تلخص أداء النموذج:

- True Positives (TP) : التنبؤات الصحيحة للفئة الإيجابية
- True Negatives (TN) : التنبؤات الصحيحة للفئة السلبية
- False Positives (FP) : خطأ النوع الأول – تنبؤ خاطئ بالإيجاب
- False Negatives (FN):خطأ النوع الثاني – تنبؤ خاطئ بالسلب

3. منحنى ROC ومساحة AUC

منحنى ROC (Receiver Operating Characteristic) يرسم معدل الإيجابية الحقيقية (TPR/Sensitivity) مقابل معدل الإيجابية الخاطئة (FPR = 1 − Specificity) عند عتبات قرار مختلفة. مساحة تحت المنحنى (AUC – Area Under Curve) توفر مقياساً واحداً لأداء النموذج:

- AUC =1.0 :نموذج مثالي
- AUC ≤ 0.9 : < 1.0 ممتاز
- AUC ≤ 0.8 : < 0.9 جيد جداً
- AUC ≤ 0.7: < 0.8 جيد
- AUC ≤ 0.6: < 0.7 مقبول
- AUC : < 0.6 ضعيف
- AUC =0.5 : مساوٍ للتخمين العشوائي

المحور الثالث: الاستنتاجات والتوصيات

اولا: الاستنتاجات

في ضوء التحليل الإحصائي وتطبيق تقنيات تنقيب البيانات للتنبؤ بمغادرة العملاء في شركة NBTel، توصلت الدراسة إلى مجموعة من الاستنتاجات المهمة التي تعكس طبيعة العلاقة بين متغيرات الدراسة وسلوك العملاء. فقد أظهرت النتائج أن استخدام تقنيات التحليل التنبئي، ولاسيما نماذج الانحدار اللوجستي وأشجار القرار، يوفر قدرة

فعالة على تحليل بيانات العملاء والتنبؤ باحتمالية مغادرتهم للخدمة، مما يؤكد أهمية توظيف تقنيات تنقيب البيانات في دعم القرارات الإدارية في قطاع الاتصالات.

كما بينت الدراسة أن سلوك العملاء وخصائص تعاملهم مع الخدمة تعد من العوامل الحاسمة في تفسير ظاهرة مغادرة العملاء، إذ يتأثر قرار المغادرة بعدة متغيرات تتعلق بمستوى نشاط العميل وطبيعة العقد ودرجة رضاه عن الخدمات المقدمة، فضلاً عن تكرار الشكاوى ومستوى التفاعل مع الشركة. وتشير هذه النتائج إلى أن تحليل بيانات العملاء يمكن أن يوفر مؤشرات مبكرة تساعد الشركات على التعرف على العملاء المعرضين لخطر المغادرة قبل حدوثها فعلياً.

كذلك أظهرت نتائج التحليل أن دمج البيانات الديموغرافية والسلوكية والمالية في نموذج تحليلي واحد يسهم في تحسين دقة التنبؤ بسلوك العملاء، الأمر الذي يعزز من قدرة المؤسسة على تطوير استراتيجيات احتفاظ قائمة على تحليل البيانات بدلاً من الاعتماد على الأساليب التقليدية. كما تؤكد النتائج أن تطبيق النماذج التنبئية في بيئة العمل الفعلية يمكن أن يدعم التخطيط الاستراتيجي ويساعد في توجيه الموارد نحو الفئات الأكثر عرضة للمغادرة.

وبناءً على ذلك، يمكن الاستنتاج أن اعتماد تقنيات تنقيب البيانات في تحليل بيانات العملاء يمثل أداة فعالة لتعزيز إدارة علاقات العملاء، حيث يتيح للشركات إمكانية التنبؤ بالسلوك المستقبلي للعملاء وتصميم برامج احتفاظ أكثر كفاءة. كما يسهم هذا النهج التحليلي في تقليل معدلات فقدان العملاء وتحسين القدرة التنافسية للمؤسسات في بيئات الأعمال التي تعتمد بشكل متزايد على تحليل البيانات الضخمة.

ثانيا: **التوصيات** : تهدف التوصيات إلى تعزيز قدرة الشركة على الاحتفاظ بالعملاء وتقليل معدل المغادرة من خلال خطة تنفيذية متعددة المدى. تعتمد هذه التوصيات على نتائج التحليل التنبؤي وتقسيم العملاء حسب مستوى الخطر، لتوجيه الموارد والجهود بشكل فعال. تشمل التوصيات إجراءات قصيرة المدى لتعزيز الاحتفاظ الفوري، وخططاً متوسطة المدى لتطوير الأنظمة والتقنيات، واستراتيجيات طويلة المدى لبناء ثقافة تنظيمية مستدامة وتحسين البنية التحتية والمنتجات بما يدعم النمو المستدام للشركة.

أ- التوصيات الزمنية

- **قصيرة المدى (1-3 أشهر):**

- تطبيق النموذج على كامل قاعدة العملاء

- إنشاء فريق احتفاظ متخصص
- تطوير باقة عروض مستهدفة
- تحسين نظام الدعم الفني
- حملة للتحويل للعقود السنوية

- **متوسطة المدى (3–12 شهر):**

- تطوير نظام CRM متكامل
- توسيع العينة اولية وإعادة تدريب النماذج
- استكشاف تقنيات متقدمة)الشبكات العصبية(
- تطوير تطبيق جوال محسّن
- إنشاء برنامج ولاء شامل

- **طويلة المدى (1–3 سنوات):**

- بناء ثقافة تنظيمية تركز على الاحتفاظ
- الاستثمار في البنية التحتية
- تطوير منتجات جديدة بناءً على الاحتياجات
- تطبيق المنهجية على شركات أخرى
- دراسات دورية لتحديث النماذج

ب–**استراتيجيات الاحتفاظ المقترحة** :تسعى استراتيجيات الاحتفاظ المقترحة إلى تقليل معدل مغادرة العملاء من خلال تقديم تدخلات موجهة حسب مستوى الخطر لكل عميل. تعتمد هذه الاستراتيجيات على تحليل البيانات التنبؤية لتصنيف العملاء إلى مجموعات ذات خطورة عالية، متوسطة، ومخلصة، بحيث يتم تصميم برامج وعروض مخصصة لكل مجموعة لتعزيز رضاهم وزيادة ارتباطهم بالشركة،

وبالتالي تحسين استدامة الإيرادات وتعظيم القيمة الاقتصادية لكل عميل.

- **للعملاء ذوي الخطر العالي: (40%)**

✓ اتصال شخصي فوري خلال 48 ساعة

✓ خصم 30–40% لمدة 3 أشهر عند التحويل لعقد سنوي

✓ ترقية مجانية للباقة الأعلى لشهرين

✓ إلغاء الغرامات والرسوم المتأخرة

✓ أولوية في الدعم الفني (< 1 ساعة)

✓ زيارة فنية مجانية لحل المشاكل

- **للعملاء ذوي الخطر المتوسط (24%):**

- رسائل شهرية توضح قيمة الخدمات
- مكالمة ربع سنوية لقياس الرضا
- خصم 15–20% عند التجديد
- عروض موسمية حصرية
- مكافآت على الإحالات

- **للعملاء المخلصين (36%):**

❖ برنامج ولاء متدرج مع مكافآت

❖ أولوية في الخدمات الجديدة

❖ خصومات حسب سنوات الاشتراك

❖ دعوات لفعاليات خاصة

❖ إشراكهم كسفراء للعلامة

❖ تطوير نظام تحليل بيانات العملاء ضمن البنية التحتية الرقمية للشركة.

- ❖ دمج النماذج التنبؤية ضمن نظام إدارة علاقات العملاء (CRM) لدعم اتخاذ القرار.
- ❖ استخدام نتائج النماذج التنبؤية لتصنيف العملاء حسب مستوى خطر المغادرة وتوجيه العروض التسويقية وفق هذا التصنيف.

المراجع

المراجع العربية:

1. العبيدي، محمد حسن (2021). تطبيقات الذكاء الاصطناعي في قطاع الاتصالات العراقي. مجلة العلوم الإدارية والاقتصادية، جامعة بغداد.
2. الموصلي، أحمد سعيد (2020). استراتيجيات الاحتفاظ بالعملاء في الشركات الخدمية. دار الكتب العلمية، بيروت.
3. النعيمي، فاطمة علي (2022). تنقيب البيانات وتطبيقاته في إدارة علاقات العملاء. مجلة الإدارة والاقتصاد، جامعة الموصل.

المراجع الأجنبية:

1. Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In Proceedings of the 20th International Conference on Very Large Data Bases (VLDB), 487-499.
2. Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning. Journal of Big Data, 6(1), 1-24.
3. Britannica, The Editors of Encyclopaedia (2009). Data mining. Encyclopedia Britannica. Retrieved from https://www.britannica.com/technology/data-mining
4. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI Magazine, 17(3), 37-54.
5. Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann Publishers.

6. Keramati, A., et al. (2014). Improved churn prediction in telecommunication industry. Applied Soft Computing, 24, 994-1012.
7. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
8. Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., ... & Flach, P. (2021). CRISP-DM twenty years later: From data mining processes to data science trajectories. IEEE Transactions on Knowledge and Data Engineering, 33(8), 3048-3061.
9. Neslin, S. A., et al. (2006). Defection detection: Measuring and understanding customer churn models. Journal of Marketing Research, 43(2), 204-211.
10. Provost, F., & Fawcett, T. (2013). Data Science for Business. O'Reilly Media.
11. Pyle, D. (1999). Data Preparation for Data Mining. Morgan Kaufmann Publishers.
12. Vafeiadis, T., et al. (2015). A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 55, 1-9.
13. Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 29-39.
14. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann.
15. V. Kavitha, G. Hemanth Kumar, S. V Mohan Kumar &M. Harish: Churn Prediction of Customer in Telecom Industry using Machine Learning Algorithms. May 2020International Journal of Engineering Research and V9(05).
16. Sadat, S., Qiam, M., & Kohistani, A.Data Mining Techniques for Knowledge Discovery in Large-Scale Information Systems, October 2025Journal of Artificial Intelligence in Education 1(2):80-88.
17. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate Data Analysis (8th ed.).
18. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression (3rd ed.).