الجمهورية العربية السورية

المعهد العالي للعلوم التطبيقية والتكنولوجيا

نمذجة محاكاة ماذا–لو في ذكاء الأعمال

تطبيق على العروض في شركة سيريتل

What-If Analysis Simulation Model in Business Intelligence

Application on Offers' Selection in Syriatel Telecom Company

أطروحة أُعدت لنيل درجة الماجستر في النظم المعلوماتية ودعم القرار

إشراف

د.صلاح الدوه جي

د.كادان الجمعة

تقديم

رؤى عشي

كلمة شكر

لابدّ لي في البداية من تقديم الشكر والامتنان إلى الدكتور كادان الجمعة والدكتور صلاح الدوه جي اللذين قدما الأفكار والنصائح وأشرفا على هذا العمل المتواضع.

أيضاً أخص بالشكر شركة سيريتل التي أتاحت لي إمكانية اختبار النموذج المقترح على بيانات حقيقية والتأكد من صحة ودقة النتائج التي توصلنا إليها في هذا البحث.

إهداء

إلى ياسمين الشآم ...

إلى عائلتي الصغيرة التي علمتني
المعنى الحقيقي
للطموح والحب والعطاء والصبر
والإصرار لتحقيق الأحلام والأهداف
في هذه الحياة ...
والتي وقفت بجانبي ودعمتني
لإنجاز هذا العمل وإتمامه
على أحسن وجه ممكن
أضاءت طريقي ونورت دربي ...

أبي، أمي، الياس، يوسف

المنشورات

قمنا بنشر هذا العمل في مجلة محكّمة International Review on Computers and Software (I.RE.CO.S.)، هذه المجلة هي إيطالية معترف بها في Scopus & Thomson Reoters.

في الجدول التالي نجد معلومات تفصيلية عن المجلة والتي حصلنا عليها من الموقع الرسمي https://journalmetrics.scopus.com :

Year	Citescore	Cite source Precentile	Cite score rank	Citations 2015	Documents 2012-14	Cited %	SNIP	SJR
2015	0.31	22%	155/200	372	1,205	22%	0.339	0.299

ملحق

قاموس المصطلحات

Glossary

لاختصار	المفهوم الإنكليزي	المفهوم العربي
	What-if analysis	تحليل ماذا–لو
BI	Business intelligence	ذكاء الأعمال
PISWAS	Predict Impact & Suggest What-If Scenarios	اسم النظام المقترح في هذا البحث
MLR	Multiple Linear Regression	الانحدار الخطي المتعدد
SMS	Short Message Service, it is a type of text messaging in mobiles.	الرسائل النصية
GPRS	General Packet Radio Service, it is oriented mobile data services for mobile communications.	نقل البيانات
	Customer Churn: is the loss of clients or customers	خسارة الزبائن
CDRs	Call Data Records.	سجلات الاتصالات المخزنة
On net mobile	Phone calls which are made between users of the same service provider.	المكالمات الهاتفية بين الزبائن على نفس المخدّم
	Segmented offers	العروض الموجّهة لشرائح الزبائن
Syriatel Telecom company	Which has been leading the Syrian mobile telecommunication market since 2000. The company has successfully established its reputation by focusing on customer satisfaction and social responsibility.	شركة سيريتل الموقع الرسمي الخاص بها هو: http://www.syriatel.sy/
	Mixed Bundles	باقات العروض الموجّهة لكل أنواع الاتصالات SMS, GPRS, Mobile Calls
ARPU	Average Revenue per User	مقياس لمعرفة معدّل إيرادات الزبائن
MapReduce	A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner	-
	On the fly	على الطاير

الفهرس

قائمة الأشكال

قائمة الجداول

قائمة المعادلات

ملخص

انطلاقاً من أنّه على صنّاع القرار توخّي الحذر والدقة عند اتخاذ القرارات، ظهرت الأنظمة التي تعتمد على تحليل ماذا–لو للتنبؤ بأثر سيناريو معين ناتج عن اتخاذ قرار ما وبالتالي يكون القرار مبني على دراسة وتحليل أبعاده. لكن نجد أنّ هناك فجوة ناتجة عن هذه الأنظمة سابقاً وهي عدم تقديم النصيحة لصنّاع القرار بالسيناريوهات الأفضل، وإنّما الاكتفاء بالتنبؤ بتأثير سيناريو محدد. لإزالة تلك الفجوة توجّهنا في هذا البحث لبناء نموذج تحليل ماذا–لو يقوم بالتنبؤ بالبدائل الممكنة واقتراحها على المستخدمين بهدف الوصول إلى غاياتهم النهائية، مشيراً إلى أثر كل سيناريو على المؤسسة.

عادةً ما يكون هذا النظام موجّه فقط لتطبيق معين ضمن مجال عمل محدد، لذلك حاولنا قدر المستطاع أن نضع إطار عام للنموذج بحيث يتلاءم مع العديد من التخصصات، ولكن مع ذلك يبقى هناك نقاط تتناسب فقط مع مجال البحث الذي اخترناه.

باعتبار أنّ العروض في مجال الاتصالات هي أفضل طريقة لزيادة الإيرادات، فإنّه من الأفضل أن يتنبأ صنّاع القرار بالإيرادات المحتمل أن تتحقق فيما لو تمّ إطلاق عرض ما، وذلك باللجوء إلى تحليل ماذا–لو. اعتمدنا في إنجاز هذا التحسين في النموذج على خوارزمية k–means لتوزيع الزبائن في شرائح متجانسة من حيث الإيرادات، كما طبّقناMultiple Linear Regression للتنبؤ بالإيرادات التي من المحتمل أن يحققها العرض. إضافةً إلى أننا وضعنا خوارزمية للتنبؤ بالسيناريوهات الممكنة وطبقناها في هذا البحث.

للتحقق من نتائج النموذج الذي اقترحناه، قمنا ببناء تطبيق تفاعلي موجّه PISWAS واختبرنا معطيات حقيقية من شركة سيريتل Syriatel وقارنا بين النتائج الحقيقية للعروض والنتائج المقترحة من النموذج. يقدّم PISWAS المساعد لصنّاع القرار لاختيار وتقييم البدائل والسيناريوهات الممكنة بناءً على دراسة وتحليل البيانات المشابهة القديمة وذلك بما يتناسب مع غايات الشركة وتطلعاتها المستقبلية.

كلمات مفتاحية:

تحليل ماذا–لو، k-means، Multiple Linear Regression، الاتصالات، صناعة القرار، العروض.

Abstract

Decision making process have to be much more accurate and careful. Therefore, decision makers depend on what-if analysis systems to predict an impact of a specific scenario.

Usually, previous what-if analysis models in the literature have been directed just to predict an impact of a specific scenario. Therefore, our main goal in this approach is to enhance what-if analysis to suggest the best scenarios, in addition to predict their impacts.

Affordable offers are one of the best ways to increase the revenue in telecom companies. Decision makers can predict a potential revenue before launching an offer, depending on what-if analysis system.

This research depends on enhanced k-means algorithm to categorize customers into segments of the same behavior or usage. We applied also Multiple Linear Regression algorithm to predict revenues of offers' scenarios suggested by our system. To validate our model, we built a new interactive guided application, so named PISWAS system. This system was tested with real customers' data from a local important telecom company in our country which is Syriatel. Then we compared the obtained results with two of the selected launched offers in this company. As a result, PISWAS system was very helpful for decision makers to asses and select the perfect offers' scenarios relating to their company business need.

Keywords:

What-if analysis, K-means, Multiple Linear Regression, Telecom, Decision Making, Offers.

الفصل الأول

الإطار العام للبحث

يتم توضيح الإطار العام للبحث في هذا الفصل، مشاراً إلى الإشكالية الأساسية التي تمّ حلها من خلال هذا البحث.

"Not making a decision means forgoing on opportunity ". Auliq Ice.

1. مقدمة

يندرج هذا العمل ضمن إطار دعم القرار في الشركات والمؤسسات الكبيرة وذلك بهدف المساهمة في تحسين وتطوير هذه القرارات، وزيادة جودتها، إضافةً إلى تجنب المخاطر التي يمكن حدوثها على أثر اتخاذ قرار معين غير مدروس، وبالتالي هذا يؤدي إلى الارتقاء بالمؤسسة والمحافظة على نموها وتطورها بشكل سليم والسعي للعمل نحو الأفضل باستمرار.

يقضي المديرون معظم وقتهم في العمل بالتواصل مع الموظفين واتخاذ القرارات، ولكن ليس كل القرارات متشابهة ومن ممكن أن تتطلب الاعتماد على عمليات صنع قرار مختلفة. تختلف القرارات في أنواع، فمنها قرارات بغاية الأهمية وتأثيرها كبير على المؤسسة وفي هذه الحالة تكون بحاجة إلى تنظيم ومتابعة أكثر من غيرها من القرارات الأقل أهمية في المؤسسة والتي يكون تأثيرها محدود 1 [1].

هناك حاجة لمعرفة كيف يتخذ المديرون في كافة أنحاء العالم القرارات، وكيف تستطيع نظم المعلومات دعم عملية صنع القرار [22]. يتجلى الهدف الأساسي وراء هذا البحث في الابتعاد عن اتخاذ القرارات المبنية على خبرة الإداريين وتخصصاتهم في مجال معين، إلى جانب رؤيتهم وتصورهم وتقديرهم الشخصي وتحليلهم للمشاكل المطروحة دون دراسة وتحليل أبعاد تنفيذ أي قرار وأثره على كيان المؤسسة.

زادت شروط السوق من إجبار الشركات لتخفيض الضياع والفساد التدريجي وتحسين القرارات، وهذا التحسين ليس فقط شيء مهم وحساس وإنّما أيضاً يعتبر العامل الحيوي في المؤسسات، لتحقيق ذلك التحسين في سلوك المؤسسة، فإنّه يلجأ عادةً صنّاع القرار إلى الاعتماد على أنظمة تنبؤية موثوقة تدعم القرار في مجالات عمل معينة [3]، يتمحور جوهرها الأساسي حول تقييم جميع البدائل الممكنة بالإضافة إلى دراسة وتحليل تأثيرها قبل اتخاذ القرارات في مجال معين لتحقيق غاية محددة، وذلك بهدف التخطيط الأمثل والأدق للوصول إلى الأهداف المرجوة بأفضل سيناريو ممكن وأقل خسائر وأضرار من المحتمل أن تنتج عن تحقيق هذا الهدف وتؤثّر على استقرار كيان المؤسسة.

لهذا الغرض ظهر مفهوم تحليل ماذا–لو الذي يتمثّل بأنّه أداة محاكاة لبيانات مكثّفة، الهدف منه هو التنبؤ بسلوك المؤسسة بعد تطبيق فرضية معينة ناتجة عن اتخاذ قرار ما [3]، وانطلاقاً من هذه الأنظمة أصبح باستطاعة المؤسسة دراسة وتحليل أبعاد ونتائج أي قرار قبل اتخاذه، وبالتالي اختيار الحل والطريق الأمثل بشكل أدق ليعود عليها بالفائدة الأكبر، ولتجنب المخاطر التي يمكن حدوثها أثر هذا القرار المنفّذ.

2. لمحة عن تحليل ماذا–لو

بازدياد عدد المؤسسات وتضخمه، ظهرت الحاجة إلى الحصول على معلومات عن مستقبلها المهني من خلال التخطيط لاستراتيجيات مثالية لتحقيق أهداف هذه المؤسسات [40]، وبالتالي لتقييم أثر هذه الاستراتيجيات فإنّ صُنّاع القرار بحاجة إلى أنظمة موثوقة للاعتماد عليها.

إن مخازن المعطيات Data Warehouse تلعب الدور الأساسي في ذكاء الأعمال Business Intelligence، حيث أنّ الغرض منها هو تخزين البيانات القديمة وتحليلها ولكنها لا تستطيع التنبؤ والتوقع بالاتجاهات المستقبلية [3] [5]، وهذا ما دعا إلى ظهور تحليل ماذا–لو وتعزيز الهدف منه والحاجة له في دعم القرار والتنبؤ بالآثار المستقبلية لقرار ما. بالنهاية فإنّ تحليل ماذا–لو يعتمد على تقنيات قاعدة المعطيات Database ومخازن المعطيات Data Warehouse.

يتمحور مفهوم التحليل ماذا–لو حول قياس تأثير التغييرات في مجموعة من المتحولات المستقلة على مجموعة من المتحولات التابعة لها استناداً إلى نموذج محاكاة، يتكوّن هذا النموذج من سيناريو واحد أو أكثر وكل منها يعبّر عن أحد الطرق البديلة المحتملة للوصول إلى حاجة ورغبة صنّاع القرار [05]. حيث أنّ كل سيناريو يتحدد بما يلي:

- عدد من المتحولات العائدة لإدارة أعمال المؤسسة والتي سيتم التنبؤ بأثرها في حال تحقق السيناريو، بمعنى آخر إنّ هذه المتحولات تعبّر عن العوامل التي تهتم المؤسسة بدراستها ومعرفة مدى تأثرها بقرار معين، حيث تكون الدراسة كلها مبنية بهدف التنبؤ بقيم هذه المتحولات.
- عدد من المعاملات التي يتم تحديد قيمها من جهة المستخدم النهائي وذلك حسب حاجة المؤسسة، تساعدنا هذه المعاملات في دراسة السيناريو بشكل أفضل للتنبؤ بأثره بدقة.

وبالتالي يمكن تغيير قيم المتحولات والمعاملات السابقة بحسب رغبة أصحاب القرار حتى الوصول لقيم مرضية قريبة من تطلعاتهم المستقبلية في مجال معين.

يجب التنويه أيضاً إلى أنّه يختلف صلب وطريقة بناء النموذج العائد لتحليل ماذا–لو حسب مجال توجهه والغرض الأساسي من بنائه، بالإضافة إلى أنّ طبيعة المؤسسة وأهدافها تلعب دوراً أيضاً في تحديد بنية النموذج، فمن الممكن أن يعتمد النموذج إمّا على خوارزميات مبنية على معادلات وطرق حسابية، أو على مجموعة من خوارزميات تنقيب البيانات Data Mining وبالتالي تحليل النتائج من خلال الاعتماد على ذكاء الأعمال Business Intelligence، أو غيرها من التقنيات والأساليب والخوارزميات. ولكن مفهوم هذا النموذج الجوهري والهدف الأساسي منه يبقى ثابتاً مهما تغيرت أساليب تطبيقه ومجالاته والغايات الأساسية من بنائه وهيكلته .

قبل البدء بشرح إشكالية البحث المطروحة سوف نوضّح ونؤكّد على مجموعة من المصطلحات والمفاهيم المهمة في إطار هذا البحث، والتي تتمثّل فيما يلي:

- تحليل ماذا–لو: هو أداة محاكاة تدعم صنّاع القرار للتنبؤ والكشف عن أثر حدوث سيناريو معين على معاملات محددة في المؤسسة.
- القرارات في المؤسسة: تتمحور حول الحاجة لوضع مجموعة من الخيارات التي يجب على المؤسسة القيام بها أو الابتعاد عنها، وهذه الخيارات تقوم على متطلبات المؤسسة وحاجتها.

القرارات هي العملة القانونية في مجال الأعمال التجارية، فكل نجاح وفشل وفرصة تمّ ضبطها أو غابت عن المؤسسة تعود إلى قرار معين تمّ اتخاذه من قبل طرف ما أو فشل في تنفيذ هذا القرار، كما أنّ العديد من القرارات المتخذة بشكل روتيني داخل المؤسسة تؤذي وتؤثّر سلباً على الأداء الداخلي لها [6].

- القرارات الناجحة في المؤسسة: عادةً تكون مبنية على الفهم الدقيق لاحتياجات المؤسسة وأهدافها وغاياتها، إضافةً إلى تقييم البدائل وانتقاء الأفضل منها بعد تحليلها ودراسة أبعادها ومدى تأثيرها على المؤسسة، إلى جانب الابتعاد عن اتخاذ القرارات بناءً على خبرة الأطراف المعنية في الموضوع.

في الحقيقة إنّ صنع القرارات الجيدة والصحيحة وبشكل سريع هو أحد أسباب نجاح المنظمات والمؤسسات ذات الأداء العالي [6].

- المخاطر الناتجة عن عملية صنع القرار: هي أحداث غير متوقعة ممكن أن تؤثّر على أحد أهداف المؤسسة وغاياتها، إلى جانب أنّها تهدد أمن واستقرار المؤسسة. يعاني الإداريون عند وضع استراتيجية الخيارات والقرارات من اتخاذ الخيارات السيئة وفقدان الفرص وإضاعة الوقت وخسارته [7].

3. إشكالية البحث

بعد دراسة تحليل ماذا–لو في العديد من المقاربات السابقة والتعرف على النماذج المستخدمة والمجالات المدروسة، وجدنا أنّ جميع هذه الأنظمة تتنبأ فقط بمدى أثر تطبيق سيناريو معين أي أنّها تجيب على سؤال ماذا–لو الذي يتم طرحه من قبل صنّاع القرار. مما يطرح الإشكالية التالية وهي:

- ماهي السيناريوهات المقترحة لتحقيق هدف وغاية صنّاع القرار في تطوير مجال معين؟
- ما هو أثر كل سيناريو ضمن إطار محدد في المؤسسة؟

ينتج عن الإشكالية السابقة مجموعة من الاستفسارات والمقترحات التي تتجلى في مقاربتنا وقد تمّ أخذها بعين الاعتبار لتطوير نموذج ماذا–لو، وهي:

1.3. كيفية صياغة الاستعلام ماذا–لو من قبل صنّاع القرار

هناك عدد من الأسئلة التي يجب أخذها بعين الاعتبار لمعرفة كيفية إدخال الاستعلام، تتجلى فيما يلي:

- هل يمكن للمستخدم صياغة الاستعلام بلغة طبيعية أم لغة برمجية أم أنّ هناك لغة خاصة بالنظام على المستخدم الإدخال بواسطتها؟
- هل سيقوم المستخدم بكتابة الاستعلام أم أنّ النظام سوف يقوده لمعرفة واستنتاج الاستعلام الذي يريد الاستفسار عنه، وفي هذه الحالة المستخدم لا يقوم بإدخال الاستعلام وإنّما فقط يجيب على أسئلة يطرحها النظام، بالإضافة إلى متحولات يطلب منه أن يدخل قيمها؟

مقترحنا

الاعتماد على بناء نظام موجّه لا يطلب من المستخدم كتابة أي استعلام، وإنّما فقط يقوده للإجابة وتعبئة قيم للمتحولات المطلوبة والتي تساعد في عملية التحليل ودراسة كافة السيناريوهات الممكنة في مجال ما. وبالتالي الهدف من ذلك هو الابتعاد عن تقيد المستخدم بطريقة الإدخال واللغة بالإضافة إلى تسهيل استخدام النظام والوصول للأهداف المرجوّة بعيداً عن التعقيد.

2.3. كيفية اقتراح السيناريوهات الممكنة

سيتم توضيح عملية الاقتراح من خلال أخذ الأسئلة التالية بعين الاعتبار:

- ماهي الخوارزميات المستخدمة لمعرفة السيناريوهات الممكنة واقتراحها على المستخدمين النهائيين؟
- هل هذه الخوارزميات ثابتة وموحّدة بغض النظر عن ماهية مجال التطبيق وطبيعة المؤسسة؟ أم أنّها تتغير وتختلف حسب مجال الدراسة؟
- هل سيتم الاعتماد على الحالات المشابهة السابقة؟ أم أنّه من الممكن استخلاص السيناريوهات من الاستعلام الحالي فقط؟
- ما هو مدى دقة ومصداقية هذه النتائج التي تم الوصول إليها؟

مقترحنا

تطبيق خوارزمية تهدف للتنبؤ بالسيناريوهات المتوقعة، وهذه الخوارزمية تتناسب مع طبيعة التطبيق ومجال وهدف المؤسسة، والاستفادة من السيناريوهات المقترحة في الحالات المشابهة.

3.3. كيفية التنبؤ بأثر كل سيناريو

ماهي الخوارزميات المطبقة للحصول على أثر كل سيناريو؟

مقترحنا

اعتماداً على خوارزميات التنقيب، يمكننا التنبؤ بأثر السيناريو على المؤسسة في حال تحققه من خلال تحليل ودراسة الحالات المشابهة السابقة للوصول إلى تنبؤ أدق وأقرب إلى الواقع، فكلّما كانت النتائج السابقة دقيقة ومدروسة كلما كانت نتائج بحثنا أدق وأفضل.

4. لمحة عن الحل المقترح

يبين الشكل (1) بنية الحل المقترح للإشكالية التي ذكرناها في الفقرة السابقة، حيث يتضمن هذا الحل مسارين أساسيين بالإضافة إلى مجموعة من المراحل الجزئية في كل منهما، بدايةً على المستخدم اختيار وجهة الحل والهدف من الاستعانة بهذا النظام، حيث يتجلّى ذلك فيما يلي:

1.4. اقتراح السيناريوهات الأفضل

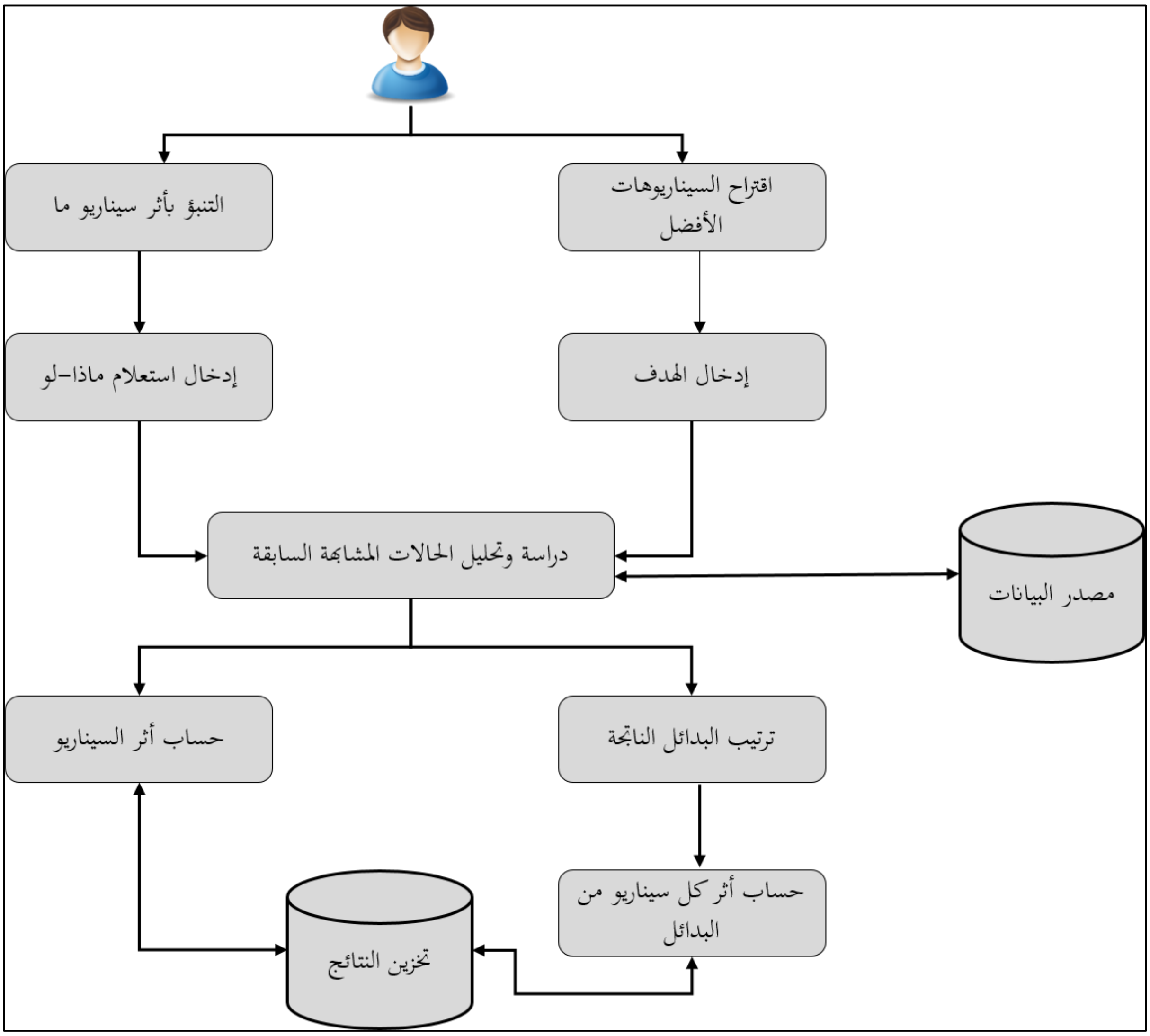
يتضمن هذا المسار مجموعة من المهام الجزئية، وهي:

- إدخال الغاية النهائية من قبل المستخدم التي يسعى في الوصول إليها ضمن إطار العمل والمؤسسة ومجال التطبيق، وذلك بهدف التطوير والتحسين اعتماداً على رغبة المستخدم.
- تحليل ودراسة الحالات السابقة المشابهة واستخراج معلومات مفيدة عن السلوك الخاص بها.
- ترتيب الاحتمالات البديلة الناتجة عن تطبيق خوارزميات معينة.
- حساب أثر كل سيناريو من السيناريوهات السابقة والتنبؤ به.

2.4. التنبؤ بأثر سيناريو معين

يتجلى هذا المسار بالخطوات التالية:

- إدخال استعلام ماذا-لو من قبل المستخدم، مشيراً إلى المتحولات التي ستدخل في الدراسة والتحليل، وإلى المتغيرات الخاصة بهذا الاستعلام التي سوف ينعكس من خلالها أثر تطبيق السيناريو.
- تحليل ودراسة الحالات السابقة المشابهة، بهدف الوصول إلى نتائج أدق قائمة على أمثلة واقعية حقيقية قد حصلت في مجال ما يطابق مجال الدراسة الحالية التي يستفسر عنها المستخدم.
- حساب أثر هذا السيناريو اعتماداً على خوارزميات معينة.



الشكل 1– لمحة عامة عن الحل المقترح

5. مخطط البحث

نظِّم هذا العمل على عدة فصول وهي:

- الفصل الثاني : يحتوي على دراسة مرجعية لمجموعة من المقاربات المختارة في مجال تحليل ماذا–لو وتطبيق نموذج المحاكاة التابع له.
- الفصل الثالث: يتضمن شرح مفصّل عن النموذج المقترح للحل، وتوضيح الإطار العام له الذي يطبّق في جميع المجالات والأبحاث بغض النظر عن مجال الدراسة وطبيعته.
- الفصل الرابع: يشمل البنية التفصيلية للنموذج العائد إلى التطبيق الذي قمنا بدراسته في هذا البحث.

- الفصل الخامس: يشير إلى حالة الاستخدام التي طبقّناها للتأكد من النتائج التي توصلنا إليها.
- الفصل السادس: يتجلى فيه الخاتمة والآفاق المستقبلية.

الفصل الثاني

الدراسة المرجعية

سنعرض في هذا الفصل مجموعة من الدراسات السابقة في تاريخ تحليل ماذا–لو مع مقارنتها وتأطيرها ضمن إطار خاص للمقارنة ومن ثمّ تحديد موقع بحثنا هذا.

"Making good decisions is a crucial skill at every level." Peter Drucker.

1. مقدمة

يوجد العديد من المقاربات التي تعود لدعم القرار من خلال الاعتماد على تحليل ماذا–لو، تتمحور في مجالات مختلفة منها التقني ومنها الإداري. سنشير في هذا الفصل إلى مجموعة مختارة من تلك المقاربات التي تهتم بتحليل ماذا–لو، كما قمنا بدراستها وتحليلها واستخلاص قيمها الإيجابية والسلبية من حيث النماذج المستخدمة في التحليل ولغات الاستعلام وآلية معالجته وطريقة البحث في مصادر البيانات المتاحة.

نهدف من خلال عرض المقاربات إلى تأطيرها ومقارنتها مع بعضها البعض عبر استخدام طريقة للمقارنة مستوحاة من الإطار المسمّى "إطار وجهات النظر الأربعة" [8]. يوفّر هذا الإطار توصيف هيكلي لأربع منظورات تحدد المقاربة المدروسة، بحيث يسهل علينا مقارنة المقاربات مع بعضها من خلال توضيح وتحديد نقاط القوة والضعف لكل مقاربة للتمكن من فهمها بشكل صحيح. كما أنّه يتوافق كل منظور مع بُعد وجانب أساسي من جوانب المقاربة المدروسة.

من خلال الاعتماد على هذا الإطار، فقد استطعنا معرفة وتحديد موضع بحثنا بسهولة وبدقة أكبر مشيرين إلى القيم المضافة التي حققناها، إلى جانب السمات الأساسية التي يتمتع بها البحث والتي قمنا باستخلاصها من المقاربات السابقة، حيث تمّ تحديد السلبيات والإيجابيات في كل مقاربة وبالتالي معرفة أفضل نهج على مستوى كل منظور في الإطار واعتماده أو إضافة قيم جديدة لتحسين الأداء وتحقيق التمييز في مقاربتنا، وهذا يعود بشكل أساسي إلى الهدف وراء البحث.

نستعرض في الفقرات التالية إطار المقارنة المقترح بشكل تفصيلي من حيث تعريف المنظورات الأربع ومركباتها، ومن ثمّ بتطبيق إطار المقارنة المستخرج على المقاربات المدروسة بما فيها مقاربنا المقترحة.

2. إطار المقارنة "وجهات النظر الأربعة"

1.2. تعريف

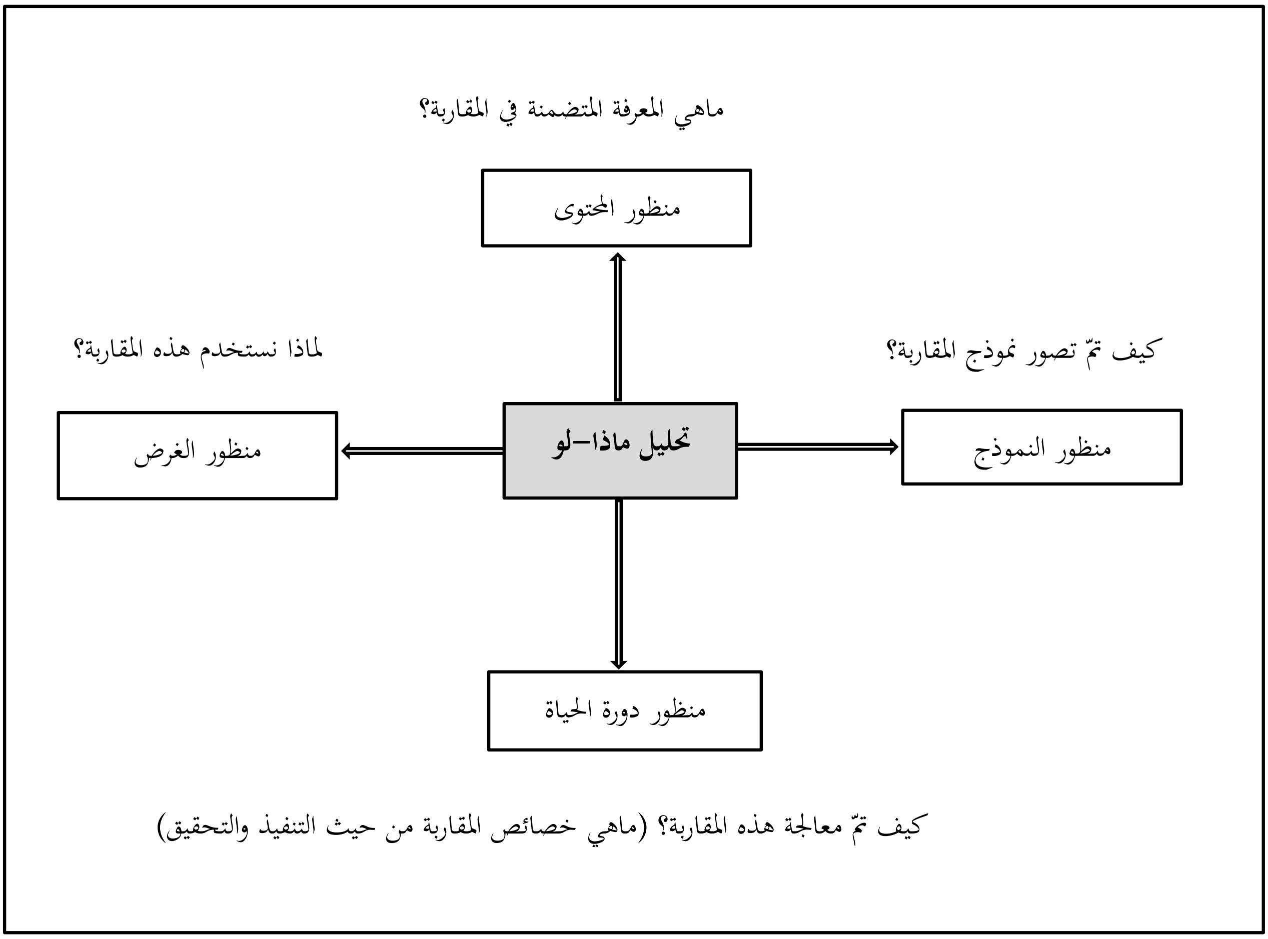
يضع إطار المقارنة "وجهات النظر الأربعة" بين أيدينا أداة علمية وفكرية ومفاهيمية لتحليل المقاربات المدروسة وتسليط الضوء على النقاط الأساسية التي تتمتع بها كل مقاربة، فقد أثبت هذا الإطار فعاليته في تحسين الفهم للعديد من الاختصاصات الهندسية مثل هندسة المتطلبات [9]، وهندسة النظم المعلوماتية [10]، وهندسة الإجراءات [11].

2.2. الهدف من إطار المقارنة

- المساعدة في فهم وتوضيح المقاربة.
- تحديد موقع الناحية العملية الصناعية من المقاربة (التقنيات والأدوات والأساليب المستخدمة في كل مقاربة).
- مساعدة الباحثين لتطوير مقاربة أكثر ابتكاراً.

3.2. بنية إطار المقارنة

يوضّح الشكل (2) هذه البنية، حيث يتكوّن إطار المقارنة من أربعة أبعاد وهي: النموذج، المحتوى، الغرض، ودورة حياة المقارنة. كل بعد من هذه الأبعاد له مجموعة من الأوجه التي تعود إلى مجال البحث في المقارنة، وكل وجه منها يتحدد بمجموعة من القيم الممكنة.



الشكل 2 – بنية إطار المقارنة

1.3.2. منظور المحتوى (الموضوع)

يجيب هذا المنظور على السؤال التالي:

ماهي المعرفة المتضمنة في المقارنة؟ وما هو الغرض الأساسي من تحليل ماذا–لو؟

فمثلاً يمكن أن يركّز الموضوع على توصيف وظائف النظام الذي تدرسه المقارنة أو الاكتفاء بوصف نظرة واسعة وشاملة عن الوظائف دون الإشارة إلى التفاصيل، ويتضمن الأوجه التالية:

- **مجال الرؤية**: يحدد درجة التفصيل في مشكلة البحث ومراحل الخوارزميات المستخدمة في تحليل الاستعلام واسترجاع النتائج، له نوعين هما:
  - صندوق أسود: يكون مستوى التفصيل محدود في نموذج تحليل ماذا–لو، حيث لا يتم الدخول بالتفاصيل وإنّما فقط الاكتفاء بنظرة سطحية عامة عن موضوع البحث.
  - صندوق أبيض: حيث يتمتع البحث بوضوح وتفصيل الخطوات في الخوارزميات المستخدمة وطريقة نمذجة تحليل ماذا–لو.

- **مستوى العمل**: يكون على مستوى أحد الخيارين التاليين:
  - مستوى معطيات الدخل وفي هذه الحالة تتم مساعدة أصحاب القرار في اتخاذ القرارات المناسبة المتعلقة بمجال معين.
  - تعديل في البنية التحتية المستخدمة ومعرفة مدى تأثير هذا التعديل على مكونات وأداء البنية المدروسة، وبالتالي في هذه الحالة ليس هناك دخل لقيم ومعطيات وإنّما فقط تهتم بالبنية.

- **الكيان**: يمثّل شريحة الأشخاص الذين يتفاعلون مع النظام ومجال عملهم وخبرتهم الموجه لها النظام، وجدنا أنّ هناك وجهتي نظر وهما:
  - تقنية: بحيث أنّه بإمكان مدراء النظام IT Administrators فقط إدخال استعلامات ماذا–لو.
  - عملية: بإمكان المستخدمين النهائيين الإداريين إدخال الاستعلامات.

- **التفاعلية مع المستخدم**: تعبّر عن مدى تفاعل المستخدم النهائي مع النظام، فمثلاً:
  - تعديل أو حذف الاستعلامات المدخلة.
  - تغير قيم المتحولات المدخلة.
  - تصدير النتائج التي تمّ الوصول إليها، وسحب تقارير خاصة بالمعطيات والبيانات المدروسة، وغيرها من المهام التفاعلية مع النظام.

2.3.2. منظور الغرض

يجيب هذا المنظور على السؤال التالي:

لماذا تستخدم هذه المقاربة؟

ويتضمن الأوجه التالية:

- **الأهداف الجزئية**: هناك مجموعة من الوظائف التي يقوم نظام تحليل ماذا–لو بها أو ببعضها حسب طبيعة بناء هذا النظام، وهي: النشر، البحث، الاستعلام، والتحليل بالإضافة إلى التحويل من لغة الإدخال إلى لغة أخرى تقنية يفهمها الحاسوب.
- **الغاية النهائية**: وجدنا في جميع المقاربات المدروسة أنّ الغاية النهائية من اللجوء إلى تحليل ماذا–لو هي تقييم السيناريوهات المدخلة والتنبؤ بنتائجها الممكنة، أمّا في بحثنا هذا فقد قمنا بإضافة غاية نهائية أخرى إلى جانب الغاية السابقة، تتمثل باقتراح السيناريوهات المتوقعة وترتيبها حسب الأفضل وذلك لمساعدة المستخدمين النهائيين في اتخاذ القرار الأمثل.
- **مرحلة العملية الهندسية**: إمّا أن يتم بناء النموذج on the fly أي في مرحلة التنفيذ أو أن يتم ذلك بشكل مسبق أي في مرحلة التصميم.

3.3.2. منظور النموذج

يجيب على السؤال التالي:

كيف تمّ تصور نموذج المقاربة؟

بعض المقاربات لم تقتصر على شرح نموذج ماذا–لو فقط وإنما توسّعت أيضاً بشرح نموذج بنية المعطيات التي تمّت من خلاله عملية البحث عن نتائج الاستعلامات:

- **نموذج البنية**: يشار في هذا النموذج إلى النقاط التي تتعلق بطبيعة البيانات المخزنة المدروسة والتي ستتم عليها عمليات التحليل والاستنتاج والتنبؤ:
  - هيكلية بنية المعطيات وكيفية تخزينها، تأخذ إحدى القيمتين: مخزنة على شكل بيان حلقي موجّه أو جداول تخزين في قاعدة معطيات.
  - تحديد آلية النمذجة والخوارزميات المستخدمة في ذلك والمطبّقة على هذه البيانات المخزّنة للوصول إلى النتائج النهائية.
- **نموذج ماذا–لو**: يتحدد في هذا النموذج كيفية معالجة البيانات لنشر استعلام ماذا–لو ومعالجته للحصول على النتائج الممكنة، ويتمتع بالخصائص التالية:

- **صياغة الاستعلام**: أي لغة الاستعلام وهي إما أن تكون لغة طبيعية وبالتالي يمكن إدخال الاستعلام من قبل المستخدمين النهائيين بشكل مباشر، أو أن تكون رمّاز برمجي وفي هذه الحالة لا يمكن إدخالها إلّا من قبل التقنيين.
- **آلية البحث**: التي تستخدم لتحليل الاستعلام ونشره ومعرفة نتائجه الممكنة، مثل خوارزميات التنقيب أو OLAP.
- **طريقة التخزين**: ممكن أن تكون عقد في بيان أو جداول تخزين في قاعدة معطيات.
- **إجرائية الاستعلام**: إمّا أن تكون بدائية بسيطة حيث يقوم المستخدم بإدخال الاستعلام بطريقة معينة يفرضها النظام عليه دون أن يتدخل في اختياراته، أو أن تكون موجّهة فيقوم النظام بتوجيه المستخدم ومساعدته لطلب الاستعلام.
- **طريقة المعالجة**: ممكن أن تكون عمليات حسابية، أو خوارزميات بحث، أو التحليل باستخدام ذكاء الأعمال Business Intelligence.
- **الاستفادة من الاستعلامات السابقة**: ممكن أن تكون عن طريق استخدام ذاكرة مخبأة Cache لتخزين الاستعلامات والاستفادة منها لاحقاً أو من خلال Parameterized View وذلك باستخدام متحولات عند صياغة الاستعلام واستبدالها بالقيم المناسبة في استعلامات أخرى.

4.3.2. منظور دورة الحياة

يجيب هذا المنظور على الأسئلة التالية:

- كيف تمّ معالجة وتصوّر هذه المقاربة؟
- ماهي خصائص المقاربة من حيث التنفيذ والتحقيق؟

ويتضمن الأوجه التالية:

- **الأدوات المستخدمة** في تنفيذ النموذج.
- **وسائل التحقق** من النتائج التي تم الوصول إليها في المقاربة، إمّا أن تكون تطبيقات أو استفتاءات أو إحصائيات.
- **التحقق من النتائج**: فمثلاً يمكن المقارنة بين النتائج المتوقعة من المقاربة والنتائج الواقعية الحقيقية.

يلخّص الجدول (1) المناظير الأربعة والأوجه المختلفة في كل منظور والقيم الممكنة لكل وجه.

جدول 1 إطار التنصيف العام

المنظور		الوجه	القيم الممكنة
المحتوى		مجال الرؤية	صندوق أسود، صندوق أبيض
		مستوى العمل	بنية تحتية، معطيات الدخل
		الكيان	تقني، أعمال
		التفاعلية مع المستخدم	تفاعلي، غير تفاعلي
الغرض		الغاية النهائية	التنبؤ بنتائج السيناريو المدخل، اقتراح السيناريوهات الأفضل
		الوظائف	النشر، البحث، الاستعلام، التحليل، التحويل
		مرحلة العملية الهندسية	مرحلة التنفيذ، مرحلة التصميم
النموذج	**نموذج البنية**	هيكلية البنية	بيان حلقي موجّه، قاعدة معطيات
		آلية النمذجة	linear function (System-level Models), partition and sub partition tables in DW
	نموذج ماذا–لو	صياغة الاستعلام	لغة برمجية خاصة ، لغة طبيعية
		آلية البحث	خوارزمية البحث بالعمق أولاً، OLAP، Data Mining
		نموذج التخزين	عقد في بيان، قاعدة معطيات
		إجرائية الاستعلام	بدائية بسيطة، موجهة
		طريقة المعالجة	عمليات حسابية، BI، خوارزميات بحث
		الاستفادة من الاستعلامات السابقة	Parameterized view ، Cash
دورة الحياة		الأدوات المستخدمة للتنفيذ	Hadoop ، My sql ،R Language ،Python

	وسائل التحقق	تطبيقات، استفتاءات، إحصائيات
	التحقق	مقارنة

3. الأعمال ذات الصلة

هناك العديد من المقاربات الموجودة في مجال التحليل ماذا–لو، كل واحدة منها صُمّمت لتحسين عملية صنع القرار في مجال معين إمّا تقني أو إداري.

تمّ انتقاء خمسة أعمال مرتبطة بمجال البحث، نقدّم فيما يلي لمحة عامّة وموجزة عن موضوع بحث كل من هذه المقاربات ومن ثمّ تحديد موقع هذه المقاربة بعد إسقاطها على إطار التصنيف السابق:

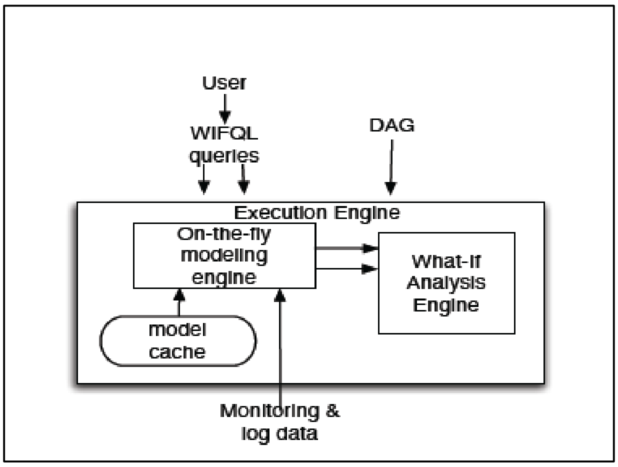
1.3. مقاربة Analytical Modeling for What–If Analysis In Complex Cloud Computing Applications – 2013

تتمحور هذه المقاربة حول تقديم نظام تحليل ماذا–لو للتنبؤ بتأثير التغييرات في حجم العمل على سلوك تطبيقات الحوسبة السحابية [12]. فقد طبّقت إطار نظري لنمذجة التطبيقات الموزعة الكبيرة، وهذا النموذج عبارة عن شبكة من الطوابير التي تجسّد الاعتمادية بين حجم العمل الواقع على كل مركبة من مركبات الحوسبة السحابية المعقدة.

تمّ بناء لغة برمجية خاصة لإدخال استعلام ماذا–لو من قبل مستخدمين من خلفية معلوماتية، وباعتبار أنّ هذا النموذج يستخدم لتطبيقات كبيرة مع مئات المكونات التي تتفاعل مع بعضها فنجد أنّ هناك عدة تحسينات فيه مثل:

- يستخدم نموذج on the fly.
- يعمل على التخزين في ذاكرة مؤقتة cache لتحسين الدقة والسرعة.

الخوارزمية المطبقة لإجراء الاستعلام الذكي هي نشر التغيير الجديد "Novel change propagation"، حيث تقوم هذه الخوارزمية أولاً برسم البيان المتأثر أي تحدد المكونات التي تأثّرت في الاستعلام المدخل، وبعد ذلك تستخدم تقنية نشر التغيير لنشر تغيير حجم العمل المحدد من خلال كل مركبة في الرسم البياني المتأثر، وهذا ما تم توضيحه في الشكل (3).



الشكل 3 – بنية النظام

جرى تقييم هذه المقاربة بناءً على تطبيقات صناعية في العالم الحقيقي، أكدّت نتائج هذا التقييم على دقة النموذج لتحليل استعلام ماذا–لو.

الإيجابيات في هذه المقاربة

- شرح تفصيلي لنموذجي بنية المعطيات وماذا–لو، إضافةً إلى شرح تفصيلي عن طريقة نشر الاستعلام على المعطيات المتاحة للوصول إلى النتائج المطلوبة.
- التخزين في ذاكرة مؤقتة cache لتحقيق الدقة في النتائج والسرعة في الأداء.

السلبيات في هذه المقاربة

- لغة الاستعلام هي لغة برمجية خاصة وليست لغة طبيعية.

يبين الجدول (2) تموضع المقاربة المدروسة في إطار المقارنة المعرّف في الفقرة الثانية من هذا الفصل.

*جدول 2 مقاربة **Singh et al.**, 2013 حسب إطار المقارنة*

المنظور	الوجه	القيم الممكنة
المحتوى	مجال الرؤية	صندوق أبيض
	مستوى العمل	بنية تحتية
	الكيان	تقني، الاعمال
	التفاعلية مع المستخدم	N/A
الغرض	الغاية النهائية	التنبؤ بنتائج السيناريو المدخل
	الوظائف	النشر، البحث، الاستعلام، التحليل

		مرحلة العملية الهندسية	مرحلة التنفيذ
النموذج	نموذج البنية	هيكلية البنية	بيان حلقي موجّه
		آلية النمذجة	linear function (System-level Models)
	نموذج ماذا–لو	صياغة الاستعلام	لغة برمجية خاصة
		آلية البحث	البحث بالعمق أولاً
		نموذج التخزين	عقد في بيان
		إجرائية الاستعلام	بدائية بسيطة
		طريقة المعالجة	خوارزميات بحث
		الاستفادة من الاستعلامات السابقة	Cache
دورة الحياة		الأدوات المستخدمة للتنفيذ	R + (lex and yacc parsing tools) Python language
		وسائل التحقق	بناء نموذج أولي لتطبيقين كبيرين في مجال المالية، تطبيق في مجال التجارة الالكترونية، جافا
		التحقق	مقارنة

3.3. مقاربة “What–if” Analysis in Investment Decision –Making 2012

نظراً لأهمية تقييم التدفقات النقدية، فإنّ هذه المقاربة تتضمن أسئلة ماذا–لو [13]، وتناقش بعض الطرق المنظّمة في تحليل ماذا–لو، وهدفها في ذلك هو تقييم درجة المخاطرة في التنبؤ وتحديد العناصر الأكثر أهمية في نجاح أو فشل الاستثمار ولكن بالمقابل هي لا تعطي الحلول التي يجب فعلها حيال الأخطاء المحتملة.

هناك ثلاثة أنواع من تحليل ماذا–لو المعتمد في هذه المقاربة، وهي:

1- تحليل السيناريو: النموذج الأساسي بحيث تتغير كل المتحولات ولكن تأخذ أصغر عدد ممكن من القيم.

2- تحليل الحساسية: هناك متحول واحد فقط يتغير ولكن يأخذ أكبر عدد ممكن من القيم.

3- تحليل المحاكاة : الدمج بين التحليلين السابقين.

الإيجابيات في هذه المقاربة

- شرح تفصيلي واضح للمعادلات المالية المعتمد عليها للوصول لنتائج تحليل ماذا-لو.
- لغة الاستعلام هي لغة طبيعية سهل إدخالها من قبل المستخدم.

السلبيات في هذه المقاربة

- عدم وجود نموذج لتحليل ماذا-لو.

يبين الجدول (3) تموضع المقاربة المدروسة في إطار المقارنة المعرّف في الفقرة الثانية من هذا الفصل.

جدول 3 – مقاربة ***Belanová, 2012*** حسب إطار المقارنة

المنظور		**الوجه**	**القيم الممكنة**
المحتوى		مجال الرؤية	صندوق أبيض
		مستوى العمل	معطيات الدخل
		الكيان	أعمال
		التفاعلية مع المستخدم	N/A
الغرض		الغاية النهائية	التنبؤ بنتائج السيناريو المدخل
		الوظائف	الاستعلام، التحليل
		مرحلة العملية الهندسية	N/A
النموذج	**نموذج البنية**	هيكلية البنية	N/A
		آلية النمذجة	N/A
	نموذج ماذا-لو	صياغة الاستعلام	لغة طبيعية
		آلية البحث	N/A
		نموذج التخزين	N/A
		إجرائية الاستعلام	بدائية بسيطة
		طريقة المعالجة	عمليات حسابية (معادلات)
		الاستفادة من الاستعلامات السابقة	N/A

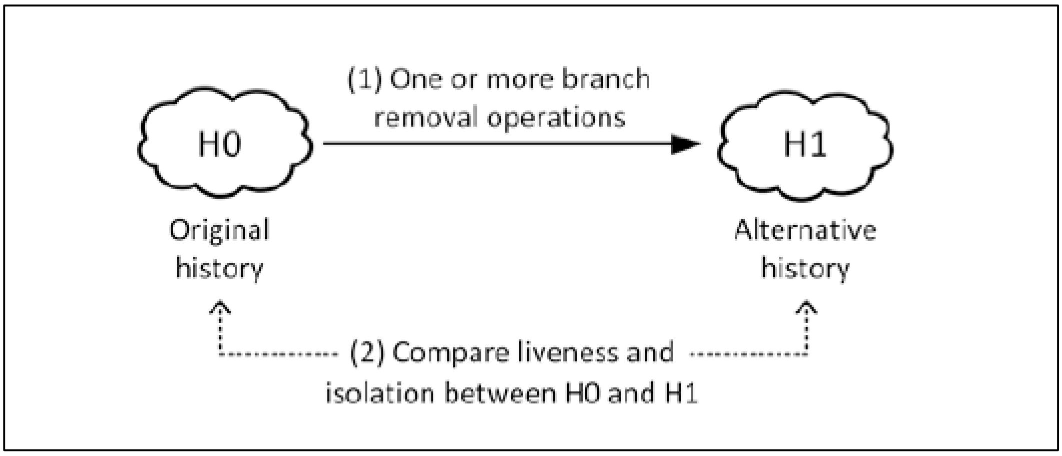
دورة الحياة	الأدوات المستخدمة للتنفيذ	N/A
	وسائل التحقق	تطبيق مثال عملي لمنتج جديد
	التحقق	مقارنة

3.4. مقارنة ASSESSING THE VALUE OF BRANCHES WITH WHAT-IF ANALYSIS–2012

إنّ الفروع ضمن نظم إدارة مصدر الرمّاز تسمح للمشروع البرمجي بتقسيم العمل بين فِرق من أجل التطوير المتزامن من خلال عزل التغييرات، ولكن هذه المنفعة لها العديد من التكاليف مثل زيادة الوقت المطلوب للتغيرات من أجل التحرك من خلال النظام، والجهد، بالإضافة إلى احتمال الخطأ عند دمج التغييرات عبر الفروع.

تقدّم المقاربة [14] نتائج استبيان في شركة مايكروسوفت عن كيفية استخدام المطورين للفروع في مشروع صناعي كبير والمشاكل الشائعة التي تواجههم. أحد هذه المشاكل الرئيسية هو التأخير الطويل الذي تستغرقه التغييرات للانتقال من فرع إلى آخر والتي غالباً ما تظهر عند وجود عدد كبير من الفروع.

حسب الشكل (4)، نلاحظ أنّ تحليل ماذا–لو يقوم بمراقبة صحة ودقة الفرع وذلك بتقييم بنية الفروع البديلة فيما يتعلق بخاصيتين هما العزل والحيوية، علماً أنّ الجزء الهام في تحليل ماذا–لو في هذه المقاربة هو حذف الفرع الوحيد.



الشكل 4. العمليات التي يقوم بها تحليل ماذا–لو

تمّ تنفيذ تحليل ماذا–لو لتقييم الفروع اعتماداً على عملية التطوير الخاصة Windows 7 وإزالة جميع الفروع الضعيفة التي تسبب تأخيراً كبيراً، وأكدّت النتائج التي حصلوا عليها أنّ تحليل ماذا–لو هو تحليل موثوق.

الإيجابيات في هذه المقاربة

- القيام باستبيان مسبق لبناء النموذج.
- لغة الاستعلام هي لغة طبيعية سهل إدخالها من قبل المستخدم.

السلبيات في هذه المقاربة

- عدم شرح نموذج ماذا–لو بشكل تفصيلي.

يبين الجدول (4) تموضع المقاربة المدروسة في إطار المقارنة المعرّف في الفقرة الثانية من هذا الفصل.

جدول 4مقارنة ***Bird et al., 2012*** *حسب إطار المقارنة*

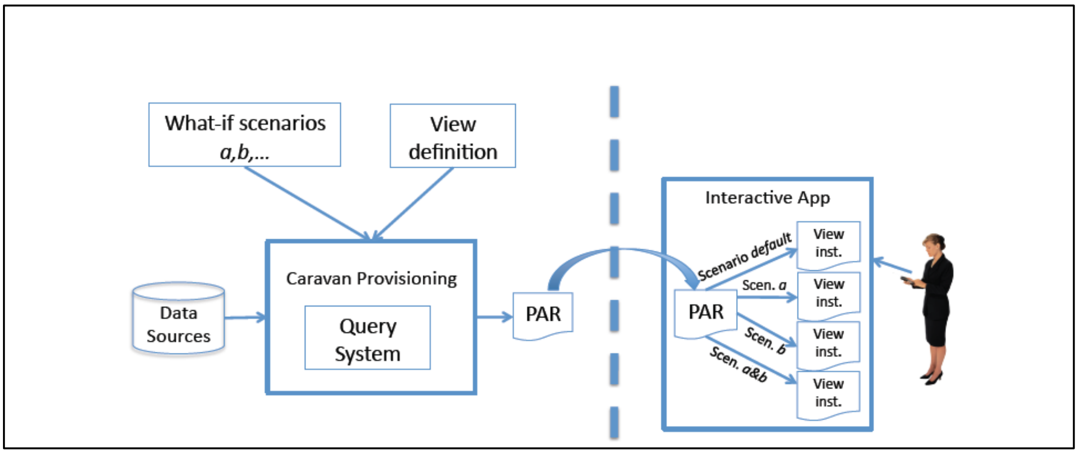
المنظور		الوجه	القيم الممكنة
المحتوى		مجال الرؤية	صندوق أسود
		مستوى العمل	بنية تحتية
		الكيان	الاعمال
		التفاعلية مع المستخدم	N/A
الغرض		الغاية النهائية	التنبؤ بنتائج السيناريو المدخل
		الوظائف	الاستعلام، التحليل
		مرحلة العملية الهندسية	N/A
النموذج	**نموذج البنية**	هيكلية البنية	N/A
		آلية النمذجة	N/A
	نموذج ماذا–لو	صياغة الاستعلام (لغة)	لغة طبيعية
		آلية البحث	N/A
		نموذج التخزين	N/A
		إجرائية الاستعلام	بدائية بسيطة
		طريقة المعالجة	N/A
		الاستفادة من الاستعلامات السابقة	N/A

دورة الحياة	الأدوات المستخدمة للتنفيذ	N/A
	وسائل التحقق	Microsoft – Windows 7 development
	التحقق	مقارنة

4.3. مقارنة Caravan: Provisioning for What-If Analysis – 2013

تقوم المقارنة [1515] بخلق تمثيلات مخصصة مستقلة مشروطة، ويتيح النظام المستخدم في هذه المقارنة تحليل ماذا-لو أسرع، أخف، وتفاعلي. يفضل المستخدمون النهائيون للنظام عادةً التفاعل مع النظام عند إدخال الاستعلامات والتعديل على الفرضيات المدخلة وغيرها من المهام التفاعلية.

يوضّح الشكل (5) أنّ هناك مجموعة من السيناريوهات، كل سيناريو منها مرتبط بمتحول بولياني يعبّر عنه، يأخذ القيمة TRUE أو FALSE، وعند التقييم في حال كان هذا المتحول هو TRUE نحتفظ به ويكون هذا السيناريو مختار من قبل المستخدم وإلا يتم حذفه في حال كانت قيمة المتحول هي FALSE، علماً أنّ حجم النظام المستخدم متحكم به من قبل المتحولات الخاصة بتحليل ماذا-لو وأيضاً هو متناسب مع حجم الإجابة الأولية على الاستعلام بدلاً من حجم مصدر المعطيات.



الشكل 5. بنية النظام

هناك عدة أطوار للتنفيذ موضحة في الشكل (5)، وهي:

1- عرض الهدف: أخذ تعريف رؤية (هدف) المستخدمين النهائيين ومجموعة من سيناريوهات ماذا–لو، ومن ثمّ إعادة صياغة هذا الهدف وفقاً لقواعد معينة يتم تجميعها من خلال النظام بناءً على كل سيناريو من السيناريوهات باستخدام لغة مخصصة محددة (sql).

2- ترميز التمثيل في علاقات: بما أنّ لغة sql القياسية لا تدعم الترميز المباشر من التعليقات الموضحة والتعبيرات ضمن الجداول، لذلك يتم تحويل النتيجة التي حصلنا عليها من (1) إلى ترميز مجسم من خلال materialized view .

3- حساب السيناريوهات المشابهة: خلق parameterized view والتي يمكن استخدامها لتقييم السيناريوهات المشابهة.

الإيجابيات في هذه المقاربة

- النموذج الناتج عن هذه المقاربة هو أسرع ،أخف، وتفاعلي.
- شرح تفصيلي واضح للنموذج.
- لغة الاستعلام هي لغة طبيعية سهل إدخالها من قبل المستخدم.
- الاستفادة من الاستعلامات السابقة لتحقيق السرعة في تقييم السيناريوهات من دون الحاجة إلى الوصول للبيانات الأصلية أو تنفيذ الاستعلامات المعقدة.

السلبيات في هذه المقاربة

- لغة الترميز بدائية ومن الممكن تطويرها.

يبين الجدول (5) تموضع المقاربة المدروسة في إطار المقارنة المعرّف في الفقرة الثانية من هذا الفصل.

*جدول 5 - مقاربة **Deutch etal., 2013** حسب إطار المقارنة*

المنظور	الوجه	القيم الممكنة
المحتوى	مجال الرؤية	صندوق أبيض
	مستوى العمل	معطيات الدخل
	الكيان	أعمال
	التفاعلية مع المستخدم	تفاعلي
الغرض	الغاية النهائية	التنبؤ بنتائج السيناريو المدخل
	الوظائف	البحث، الاستعلام، التحليل، التحويل
	مرحلة العملية الهندسية	مرحلة التنفيذ
	هيكلية البنية	N/A

النموذج	نموذج	آلية النمذجة	N/A
	نموذج ماذا–لو	صياغة الاستعلام (لغة)	لغة طبيعية
		آلية البحث	N/A
		نموذج التخزين	OLAP queries: tables/views (materialized, parameterized)
		إجرائية الاستعلام	بدائية، بسيطة
		طريقة المعالجة	BI
		الاستفادة من الاستعلامات السابقة	Parameterized view
التنفيذ والتحقيق		الأدوات المستخدمة للتنفيذ	Mysql
		وسائل التحقق	بناء تطبيق مع مولّد بيانات بسيط
		التحقق	مقارنة

5.3. مقاربة PROFILING, WHAT–IF ANALYSIS, AND COST–BASED OPTIMIZATION OF MAPREDUCE PROGRAMS – 2011

تقدّم المقاربة [16] ما يلي:

- محسّن على أساس التكلفة: من أجل إيجاد إعدادات تهيئة جيدة خاصة بوظائف MapReduce الاعتباطية بشكل تلقائي.
- محلّل: يحرر برامج MapReduce الغير معدّلة بشكل ديناميكي لتوليد ملخصات إحصائية موجزة بتنفيذ وظيفة MapReduce.
- محرّك ماذا–لو: للتفكير حول تأثير إعدادات التهيئة والبيانات وخصائص مورد العنقود على أداء عملMapReduce.

يقوم محرّك ماذا–لو بالخطوتين التاليتين للإجابة على الأسئلة ماذا–لو:

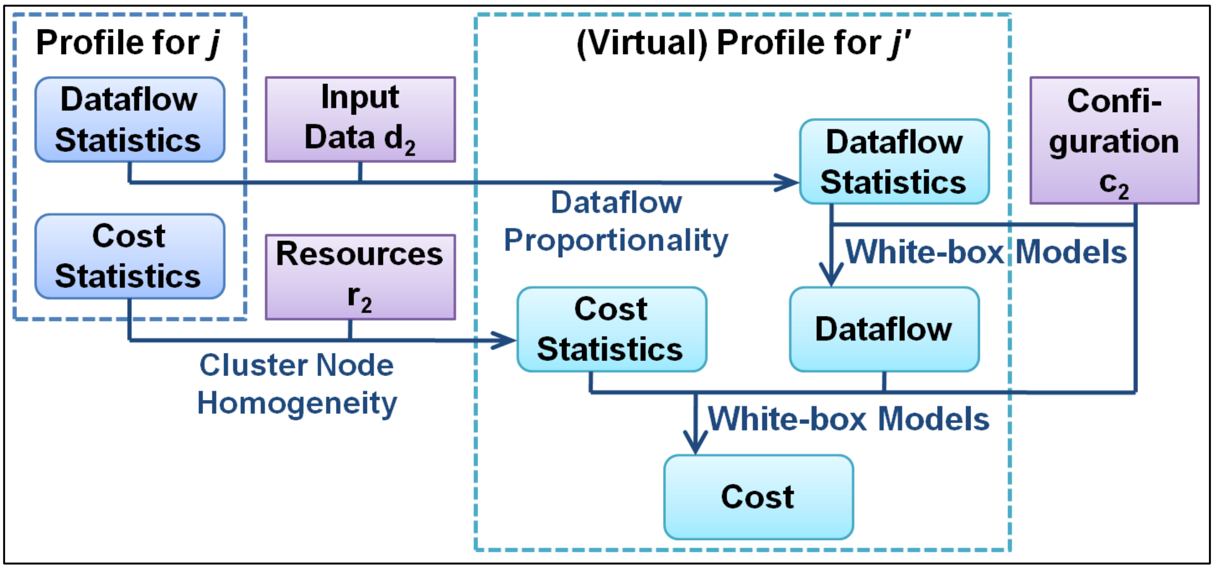
1– تقدير ملامح الوظيفة الواقعية للوظيفة الافتراضية، كما هو موضّح في الشكل (6)، من أجل الوظيفة التي يتم السؤال عنها في تحليل ماذا–لو، تنقسم عملية التقدير هذه إلى عدة خطوات، وهي:

1.1. تقدير مجالات التكلفة وتدفق البيانات: يجب تحديد هذه المجالات كبداية للنموذج.

2.1. تقدير مجالات إحصائيات تدفق البيانات: بشكل افتراضي تكون هذه المجالات في الوظيفة الافتراضية هي نفسها في الوظيفة الأولية الأصلية كإدخال، وفي حال توافر معلومات إضافية يسمح محرك ماذا–لو بتجاوز الحالة الافتراضية من خلال توفير المجالات لإحصائيات تدفق البيانات في التشكيل الافتراضي مباشرة كدخل.

3.1. تقدير مجالات إحصائيات التكلفة: تتضمن CPU , I/O هي مطابقة لحالة الوظيفة الأصلية.

2– استخدام الوظيفة الافتراضية لمعرفة كيف سيتم تنفيذ الوظيفة التي يتم السؤال عنها.



الشكل 6. عملية تقدير الوظيفة الافتراضية

الإيجابيات في هذه المقاربة

- تحسين أداء برنامج MapReduce.

السلبيات في هذه المقاربة

- عدم شرح النموذج بشكل واضح.
- لغة الاستعلام هي لغة طبيعية يقوم بإدخالها مستخدم من خلفية معلوماتية فقط.

يبين الجدول (6) تموضع المقاربة المدروسة في إطار المقارنة المعرّف في الفقرة الثانية من هذا الفصل.

*جدول 6– مقارنة **Herodotou et al., 2011** حسب إطار المقارنة*

المنظور	الوجه	القيم الممكنة
المحتوى	مجال الرؤية	صندوق أسود
	مستوى العمل	بنية تحتية
	الكيان	تقني، أعمال

		التفاعلية مع المستخدم	N/A
الغرض		الغاية النهائية	التنبؤ بنتائج السيناريو المدخل
		الوظائف	البحث، الاستعلام، التحليل
		مرحلة العملية الهندسية	N/A
النموذج	نموذج البنية	هيكلية البنية	N/A
		آلية النمذجة	N/A
	نموذج ماذا–لو	صياغة الاستعلام (لغة)	لغة طبيعية
		آلية البحث	N/A
		نموذج التخزين	N/A
		إجرائية الاستعلام	بدائية بسيطة
		طريقة المعالجة	N/A
		الاستفادة الاستعلامات السابقة	N/A
دورة الحياة		الأدوات المستخدمة للتنفيذ	Hadoop
		وسائل التحقق	Hadoop cluster
		التحقق	مقارنة

4. مقارنة بين المقاربات المدروسة

بعد دراسة المقاربات السابقة قمنا بوضع جدول اعتماداً على إطار المقاربة، يتضمن الخمس المقاربات السابقة بالإضافة إلى مقاربتنا الجديدة التي هي محور هذه الدراسة، حيث قمنا بالمقارنة بين قيم كل وجه من الأوجه واختيار الأفضل فيما بينها والأنسب لبحثنا أو قمنا بإضافة قيمة جديدة للوجه وطبقناها في هذا العمل، تتجلى هذه المقارنة في الجدول (7).

جدول7- تلخيص مقارنة المقاربات المدروسة

المنظور	الوجه	المقاربة 1	المقاربة 2	المقاربة 3	المقاربة 4	المقاربة 5	مقاربتنا
المحتوى	مجال الرؤية	صندوق أبيض	صندوق أبيض	صندوق أسود	صندوق أبيض	صندوق	صندوق أبيض
	مستوى العمل	بنية تحتية	معطيات الدخل	بنية تحتية	معطيات الدخل	بنية تحتية	معطيات الدخل
	الكيان	تقني، أعمال	أعمال	أعمال	أعمال	تقني، أعمال	أعمال
	التفاعلية مع المستخدم	N/A	N/A	N/A	تفاعلي	N/A	تفاعلي
الغرض	الغاية النهائية	التنبؤ بنتائج السيناريو المدخل	التنبؤ بنتائج السيناريو المدخل	التنبؤ بنتائج السيناريو المدخل	التنبؤ بنتائج السيناريو المدخل	التنبؤ بنتائج السيناريو المدخل	التنبؤ بنتائج السيناريو المدخل، اقتراح السيناريوهات الأفضل
	الوظائف	النشر، البحث، الاستعلام، التحليل	الاستعلام، التحليل	الاستعلام، التحليل	البحث، الاستعلام، التحليل، التحويل	البحث، الاستعلام، التحليل	البحث، الاستعلام، التحليل

		مرحلة العمل الهندسية	مرحلة التنفيذ	N/A	N/A	مرحلة التنفيذ	N/A	مرحلة التنفيذ
النموذج	نموذج البنية	هيكلية البنية	بيان حلقي موجّه	N/A	N/A	N/A	N/A	Data warehouse
		آلية النمذجة	linear function (System-level Models)	N/A	N/A	N/A	N/A	Partition and sub partition tables
	نموذج ماذا–لو	صياغة الاستعلام	لغة برمجية خاصة	لغة طبيعية	لغة طبيعية	لغة طبيعية	لغة طبيعية	لغة طبيعية
		آلية البحث	البحث بالعمق أولاً	N/A	N/A	N/A	N/A	Data Mining Algorithms

		نموذج التخزين	عقد في بيان	N/A	N/A	OLAP queries: tables/views	N/A	OLAP
		إجرائية الاستعلام	بدائية بسيطة	بدائية بسيطة	بدائية بسيطة	بدائية بسيطة	بدائية بسيطة	موجهة
		طريقة المعالجة	خوارزميات بحث	عمليات حسابية (معادلات)	N/A	BI	N/A	BI
		الاستفادة من الاستعلامات السابقة	Cache	N/A	N/A	Parameterized view	N/A	cache
دورة الحياة		الأدوات المستخدمة للتنفيذ	Python lex and) yacc parsing + (tools R	N/A	N/A	Mysql	Hadoop	SQL Server , ASP.net

	وسائل التحقق	بناء نموذج أولي لتطبيقين كبيرين في مجال المالية، تطبيق في مجال التجارة الالكترونية، جافا	تطبيق مثال عملي لمنتج جديد	Windows 7 development Microsoft –	بناء تطبيق مع مولّد بيانات بسيط	Hadoop cluster	بناء تطبيق ماذا–لو لمساعدة الإدارة في اختيار العروض الأفضل لشركة سيريتل.
	التحقق	مقارنة	مقارنة	مقارنة	مقارنة	مقارنة	مقارنة

5. الخلاصة

عرضنا في هذا الفصل مجموعة من الدراسات والأعمال السابقة في مجال البحث العائد للتحليل ماذا-لو والتي تجلّت في المجالات العملية والتقنية، وأيضاً أشرنا إلى نقاط الضعف والقوة في كل مقاربة حتى نتمكن من معرفة الإيجابيات والسلبيات في كل منها، بعدها تمّت المقارنة فيما بينهم من خلال الإطار العام للمقارنة الذي تعرفنا عليه وعلى مكوناته بشكل تفصيلي.

بعد هذه المقارنة التوضيحية، تمّ إسقاط كل مقاربة سابقة إضافةً إلى مقاربتنا على إطار المقارنة، وبهذا تمّ تحديد موقع عملنا هذا وأبعاده وخصائصه وإيجابياته إلى جانب نقاط الضعف. استفدنا من النقاط الإيجابية التي كانت متبلورة في بعض المقاربات، وقمنا بتحسين وتجنب النقاط السلبية والضعيفة.

وبالتالي تم تحديد القيم المضافة في هذا البحث والتي جعلته يتميز عن باقي الأبحاث في نفس المجال، وهذا من خلال الفهم الصحيح والدقيق لطبيعة بحثنا وطبيعة الأبحاث السابقة.

الفصل الثالث

مقاربة PISWAS للتنبؤ واقتراح

السيناريو "ماذا-لو"

سنعرض في هذا الفصل النموذج المصمم بإطاره العام الذي من الممكن إسقاطه على أي تطبيق ومجال عملي.

"Nothing is more difficult, and therefore more precious, than to be able to decide".
Napoean Bonaparte

1. مقدمة

تعاني جميع المقاربات التي درسناها في الفصل السابق من أنّها لا تقدّم النصيحة لأصحاب القرار بالحل الأمثل والقرار الأفضل الذي يعود على المؤسسة بالنفع، وإنّما كانت موجّهة للتنبؤ بأثر سيناريو ما مطبق نتيجة حاجة أدّت إلى اتخاذ قرار معين، وهذا التنبؤ يكون عائد على متغير معين يحدده صنّاع القرار بناءً على حاجتهم، ولكن دون اقتراح الحلول الأفضل على صانعي القرار أو نصحهم بأبعاد ومخاطر تنفيذ قرار محدد على كيان المؤسسة.

قمنا في هذا البحث بمحاولة إزالة الفجوة الناتجة عن الدراسات السابقة، وذلك بتحسين آلية التحليل لنصح ودعم أصحاب القرار وبالتالي المساهمة بتطوير المؤسسة بشكل فعّال أكثر.

قمنا ببناء النظام PISWAS [1] الذي يقوم بتحليل ودراسة مجموعة من البدائل الممكنة الأقرب من هدف معين بحيث تسعى المؤسسة في الوصول إليه، وبعدها يقوم باقتراح السيناريوهات الأفضل مع التنبؤ بأثر كل سيناريو بديل، مما حسّن عملية اتخاذ القرار في المؤسسات وعزز الفهم الدقيق والأفضل للمعطيات والقرارات وأبعاد تنفيذها ودرجة المخاطرة المحتملة من أثرها.

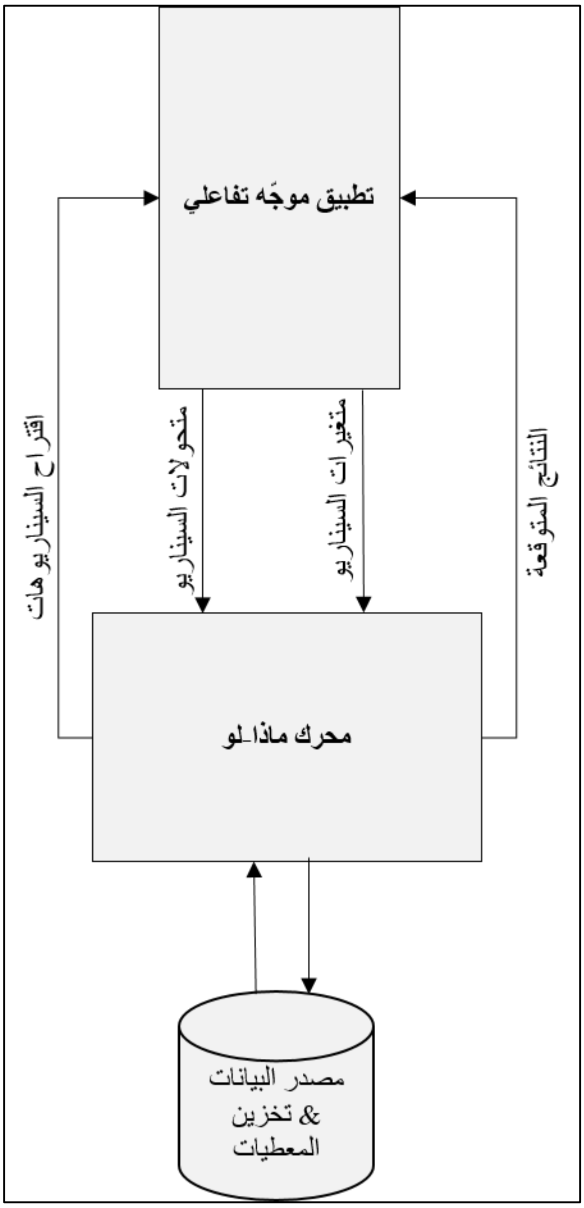
سنستعرض في الفقرات التالية نموذج الحل المقترح متضمناً شرح بنية النموذج:

2. بنية الحل

يوضّح الشكل (7) بنية وهيكلية النموذج المقترح بإطاره العام الذي سيعالج المعطيات المدخلة لمحرّك ماذا-لو، ويرد النتائج المتوقعة لأصحاب القرار الذين يتفاعلون مع النظام.

---

[1] PISWAS: is an abbreviation to "Predict Impact & Suggest What-If Scenarios".



الشكل 7 – بنية النموذج المقترح

حيث أنّ متحولات الدخل هي:

- **مصدر البيانات**: تتضمن المعطيات Data التي سوف نقوم بدراستها والعائدة لمجال عمل معين في المؤسسة، الذي يتوافق مع الهدف الذي دفع الإداريين لاستخدام النظام PISWAS.
- **متحولات السيناريو**: هذه المتحولات يحدد قيمها أصحاب القرار، وهي تستخدم في عملية التنبؤ بالسيناريوهات الممكنة وأثر كل منها، من خلال تنفيذ النموذج.
- **متغيرات السيناريو**: تعود للمجال الإداري للمؤسسة والتي سوف نتنبأ بأثر السيناريو عليها.

ومن جهة أخرى فإنّ معطيات الخرج، هي:

- **النتائج المتوقعة**: وهي تأثير سيناريو ما بعد تطبيق خوارزميات محددة للتنبؤ بذلك.

- **السيناريوهات المقترحة**: بعد تحليل ودراسة المعطيات الموجودة والمخزنة في مصدر معين واعتماداً على محرّك ماذا–لو، يقوم النظام بعرض مجموعة من البدائل المقترحة الممكنة، التي تكون أقرب ما يمكن لتحقيق الهدف المدخل والمحدد من قبل أصحاب القرار.
- **البيانات المخبأة Cached Data**: بعد الحصول على النتائج السابقة، يقوم النظام بتخزين هذه السيناريوهات العائدة لمعطيات دخل محددة، وذلك لتسريع وتحسين عملية استرجاع المعلومات بحال تمّ طلب نفس الفرضيات المدخلة من قبل المستخدمين مرة أخرى.

وبالتالي قمنا بتحديد الإطار العام لتحسين دعم القرار في التحليل ماذا–لو، فجميع الدراسات السابقة التي ذكرناها في الفصل الثاني كانت مصممة ومبينة لتخدم مجال بحث معين وتطبيق واحد، بينما في هذا البحث حاولنا قدر المستطاع تعميم النموذج المقترح المحسّن حتى يتناسب مع مجموعة كبيرة من التطبيقات الممكنة، ولكن يبقى الاختلاف بالتحديد في بنية محرك ماذا–لو الذي يعود لطبيعة المجال المدروس والمعطيات المستخدمة للتحليل إلى جانب توجّه المؤسسة وتخصصها.

3. بنية المحرّك ماذا–لو

أشرنا مسبقاً أنّ بنية المحرّك ماذا–لو تختلف باختلاف طبيعة مجال التطبيق ولكن هناك هيكلية عامة تشمل جميع المجالات، وهي ممثلة في الشكل (8)، حيث يشير هذا المخطط إلى طريقين مختلفين بناءً على رغبة أصحاب القرار، وهما:

1.3. التنبؤ بأثر السيناريو فقط

في هذا المنحى يهتم أصحاب القرار فقط بمعرفة أثر سيناريو معين كما كان في المقاربات السابقة لتحليل ماذا–لو، وبالتالي يحدد أصحاب القرار السيناريو الذي يريدون دراسته والتنبؤ بأثره، وأيضاً يحددون العامل الذي سندرس أثر السيناريو عليه، بالإضافة لمجموعة من المتحولات التي تساعد النظام في عملية التحليل.

اعتمدنا على خوارزمية التنقيب Multiple Linear Regression (MLR) للتنبؤ بالأثر، إنّ هذه الخوارزمية عريقة وقديمة منذ 1908 وكانت موظفة للتنبؤ بالتباين في التبعية المنفصلة، وذلك اعتماداً على التركيبات الخطية من الفاصلة Intervals أو ثنائية الفرع Dichotomous أو المتغيرات المستقلة الوهمية Dummy Independent Variables [17].

الهدف الأساسي من هذه الخوارزمية هو فهم وتعلّم المزيد عن العلاقة بين العديد من المتغيرات المستقلة أو المتغيرات المتوقعة والمتغير التابع أو متغير المعيار، يفترض هذا النموذج أنّ العلاقة تكون خطية بين المتغير التابع Y والشعاع P من الانحدارات Regressions [17],[18].

تتجلى معادلة MLR كما يلي:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \varepsilon \qquad (1)$$

حيث أنّ:

Y: هي القيمة المتوقعة التي تمثّل أثر السيناريو على متغيرات الدخل التي حددها أصحاب القرار.

X_i : هي مجموعة المتحولات المستقلة العائدة إلى المؤسسة والتي تؤثر على النتيجة المتوقعة لتطبيق السيناريو، وهذه المتحولات إمّا يمكن الوصول لها بشكل مباشر، أو من خلال معادلات معينة تعتمد على متحولات أخرى.

β0: Intercept، يتم حسابها وفقاً للمعادلة التالية:

$$\beta_0 = \overline{Y} - \sum_{j=1}^{k} \beta_j X_j \qquad (2)$$

β_i : Slopes، تحسب وفقاً للمعادلة التالية واعتماداً على المعطيات السابقة:

$$\beta_i = \frac{\sum X_i Y - \frac{\sum X_i \sum Y}{count}}{\sum X_i^2 - \frac{\sum X_i \sum X_i}{count}} \qquad (3)$$

حيث أنّ:

X_i تعبّر عن مصفوفة من قيم سابقة حقيقية للمتحول X.

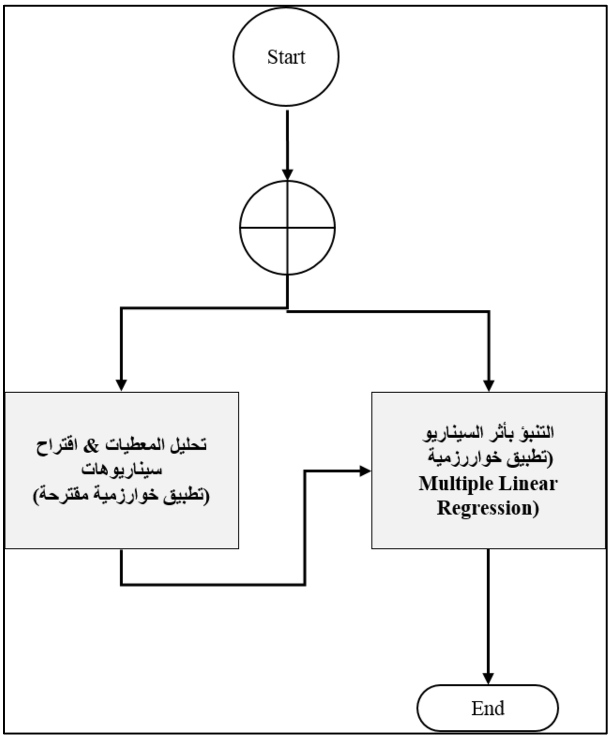
Count: هي عدد الحالات المشابهة للدراسة الحالية.

بالعودة إلى المعادلة الأولى، فإنّه عادةً يمكن تفسير القيمة R-squared (R) ليس فقط على أنّها الانخفاض النسبي في إجمالي مجموع المربعات بسبب إدراج الانحدارات، ولكن أيضاً هي جزء صغير من إجمالي مجموع المربعات الموضّحة في نموذج الانحدار [19] .

تحسب قيمة Adjusted R-squared error € كما يلي:

$$\begin{aligned}
&\in = 1 - R \\
&R = \sum_{i=1}^{k} \left(\frac{T_1 - T_2}{\sqrt{T_3 \times T_4}}\right)^2 \\
&T_1 = (count \times \sum X_i \times Y) \\
&T_2 = (\sum X_i \times \sum Y) \qquad\qquad (4) \\
&T_3 = (count \times \sum X_i^2 - (\sum X_i \times \sum X_i)) \\
&T_4 = (count \times \sum Y^2 - (\sum Y \times \sum Y))
\end{aligned}$$

بعد التعويض في المعادلة (1) نكون قد حصلنا على النتيجة المطلوبة بها.



الشكل 8- بنية محرّك البحث ماذا-لو

2.3. اقتراح السيناريوهات الممكنة والتنبؤ بأثر كل منها

في هذا المسار يتم اقتراح السيناريوهات المحتملة التي تحقق غاية وهدف أصحاب القرار في المؤسسة، حيث يتم دراسة الحالات المشابهة السابقة وتحليلها بناءً على خوارزمية محددة تختلف بحسب طبيعة المعطيات ومجال الدراسة، وهذا ما سنقوم بشرحه بشكل تفصيلي في الفصل القادم.

بعد عملية التحليل يتم ترتيب البدائل الممكنة مع التنبؤ بأثر كل منها، فبذلك يتمكن المستخدمون النهائيون من تقييم هذه الاقتراحات ودراستها وتحليلها للتعرف على أثر اتخاذ القرار بتنفيذ كل بديل مقترح.

4. الخلاصة

في هذا الفصل قمنا بشرح الآلية العامة لطريقة عمل المحرّك ماذا-لو، والتي تتجلى بمسارين مختلفين حسب رغبة مستخدمي النظام، كما تعرفنا على المنهجية المطبقة في المسار الأول التي تتجلى بالتنبؤ بأثر السيناريو المدخل فقط، وأشرنا إلى الخوارزمية المطبقة للوصول إلى هذا الغرض.

أمّا عن المسار الثاني الذي يتجلى باقتراح السيناريوهات الأفضل على المستخدمين النهائيين إلى جانب التنبؤ بأثر كل من هذه السيناريوهات، فهو يرتبط بطبيعة مجال البحث ولذلك فسوف نقوم بتفصيله في الفصل القادم.

الفصل الرابع

طريقة اقتراح السيناريوهات في مجال

الاتصالات

سنعرض في هذا الفصل النموذج المطبّق الذي

يتناسب مع بيئة العمل في قطّاع الاتصلات.

الهدف من هذا النموذج هو المساعدة في دعم

القرار باختيار العروض الأفضل للزبائن.

"If you don't drive your business, you will be driven out of business." Founder of Forbes Magazine.

1. مقدمة

نجد حالياً في مجال الاتصالات أن استراتيجيات التسويق تركّز على استقطاب وجذب زبائن جدد باستمرار إلى جانب الحفاظ على الزبائن الحاليين وزيادة ولائهم وذلك بتقديم عروض خاصة لهم تهدف لتخفيض الأسعار [20]. كما أنّه أصبح هناك تحدّي في زيادة الأرباح بشكل مستمر ويعود ذلك إلى المنافسة الشديدة، حيث أنّ جميع الشركات المنافسة تقدّم نفس المنفعة غالباً ولكن بأسعار تنافسية معقولة [21].

مما لاشك فيه أنّه على الشركة فهم وإدراك حاجات الزبائن وتفضيلاتهم وإرضائهم قبل أن تقوم بإصدار أي عرض، وذلك يعود إلى وجود علاقة إيجابية بين رضا الزبون وإخلاصه للشركة، بحيث أنّ إرضاء الزبائن هي النتيجة من عملية المقارنة بين التوقعات وأداء الخدمة في الواقع [22]. إدراك ورضا الزبائن هي جوانب لها قيمة ما بين أهداف وغايات التسويق [23].

لذلك يجب أن يكونوا دقيقين وحذرين عند اختيار خصائص العروض بعد تحليل سلوك الزبائن واستهلاكهم والأرباح التي حققوها للشركة في كل أنواع الاتصالات: مكالمات هاتفية Mobile Calls، رسائل نصية SMS [2]،ونقل البيانات GPRS[3]. وذلك بهدف تجنب العروض التي تخفض من الأرباح أو تزيد من خسارة الزبائن Customer Churn[4]، لذلك فإنّ إدارة الزبائن في شركات الاتصالات تلعب دوراً هاماً لمحاربة خسارة الزبائن والحد منه والحفاظ عليهم [23].

فبدلاً من أن تطلق الشركة العرض لجميع الزبائن بغض النظر عن استهلاكهم، تستطيع أن توزّعهم على شرائح متجانسة من حيث السلوك في الاستهلاك، وبالتالي تستطيع أن تصدر العروض المناسبة لكل شريحة بعد دراسة سلوكها. ومن هنا وجدنا أنّ هذه المؤسسات بحاجة لنظام PISWAS الذي سوف يساعد أصحاب القرار على معرفة جميع سيناريوهات العروض الممكنة لكل شريحة من الزبائن في الشركة.

سنركّز في هذا الفصل علي الخوارزمية المستخدمة لاقتراح السيناريوهات الأفضل التي ترتبط بطبيعة المؤسسة ومجالاتها إضافةً إلى الحاجة لهذا النظام التي تتمحور حول عنصر مهم في المؤسسة. وفيما يلي سنستعرض ماهية نموذج التطبيق وكيفية استخدام النموذج المجرد الذي أشرنا إليه في الفصل السابق.

2. بنية نموذج التطبيق

سنشير إلى بنية النموذج بالاعتماد على الإطار العام الذي شرحناه في الفصل السابق:

---

2 Short Message Service, it is a type of text messaging in mobiles.

3 General Packet Radio Service, it is an oriented mobile data services for mobile communications.

4 The percentage of subscribers to a service who discontinue their subscriptions to that service within a given time period.

1.2. طبيعة المعطيات المدخلة

بالعودة إلى الشكل (7)، سنوضح فيما يلي أمثلة عن معطيات الدخل والتي سوف نتعامل معها في بحثنا هذا:

- **مصدر المعلومات:** تتضمن بيانات عن تفاصيل الاتصالات بين الزبائن CDRs[5]، ومجموعة من العروض الحقيقية التي أصدرت مسبقاً وكانت موجّهة لشرائح الزبائن متضمنة تفاصيل عن هذه العروض والأرباح التي حققتها، بالإضافة إلى بيانات عن الاشتراكات في العروض.
- **متحولات السيناريو:** مثل عدد الشرائح التي سوف نوزع الزبائن من خلالها، أو فترة الدراسة للبيانات المخزنة، وغيرها.
- **متغيرات السيناريو:** مثل الأرباح التي تهدف المؤسسة إلى تحقيقها.

2.2. بنية النموذج المستخدم في التطبيق

يوضّح الشكل (9) نفس بنية محرك ماذا-لو الذي أشرنا لها في الفصل السابق، ولكن القيمة المضافة هنا والتي كانت أساسية حتى تناسب مجال الاتصالات هي توزيع الزبائن في شرائح متجانسة من حيث السلوك في الإيرادات التي تحققها، حيث أنّ عملية التوزيع هذه تمثّل خطوة أولية ومهمة لابدّ منها قبل البدء بالمهام الأساسية للنظام PISWAS.

3.2. توزيع الزبائن في شرائح CUSTOMER SEGMENTATION

تساعد عملية توزيع الزبائن في تحليل سلوكهم لإصدار العروض المناسبة لكل شريحة بشكل أدق. للقيام بهذه العملية اعتمدنا على خوارزمية التنقيب K-means التي أثبتت أنّها واحدة من أشهر خوارزميات التصنيف إلى جانب أنّها واحدة من أسهل وأبسط خوارزميات التعلم غير المراقبة التي تستطيع أن تحل مشاكل التصنيف. الاستراتيجية المتبعة لتصنيف البيانات المعطاة في عدد محدد من العناقيد هي استراتيجية بسيطة جداً [24].

الخوارزمية التقليدية الأصلية من K-means تعمل من خلال تحديد مراكز العناقيد[25]. في الواقع إنّ الخوارزمية الأصلية من K-means عند تجريبها عدة مرات وعلى نفس العينة من المعطيات نجد أنها لا ترد دائماً نفس النتيجة في التوزيع وذلك يعود إلى الطريقة العشوائية في اختيار المراكز البدائية للعناقيد Clusters، لذلك فقد اعتمدنا على خوارزمية مطورة منها [26] التي حسّنت الدقة والفعالية في النتائج مقارنةً مع الخوارزمية الأصلية.

[5] Call Data Records.

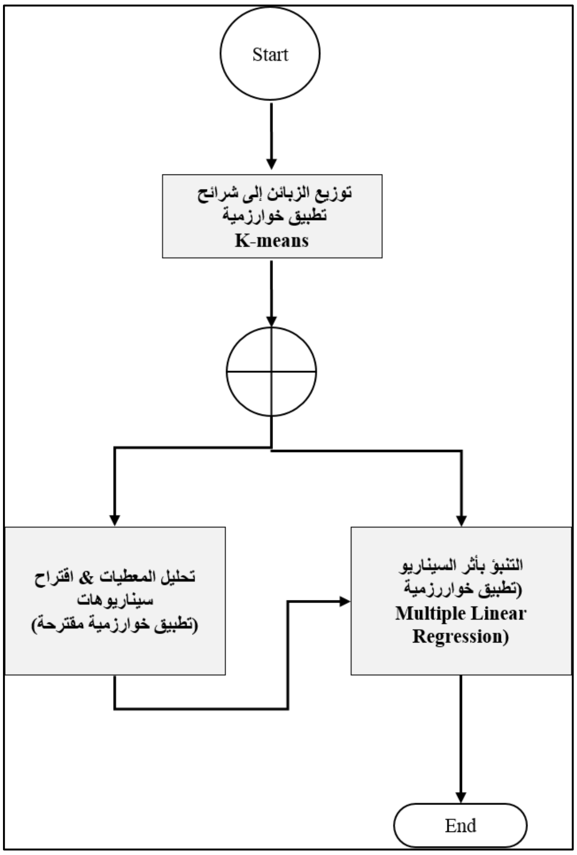
كلمحة موجزة عن آلية عمل هذه الخوارزمية المحسّنة فنجد أنّها غيّرت الطريقة العشوائية في تحديد المراكز البدائية للعناقيد لتصبح تشكّل العناقيد البدائية بناءً على المسافة الإقليدية بين كل نقطة بيانات Data Point في العنقود عن المراكز البدائية التي تتولد من خلال النظام وليس بطريقة عشوائية.

تطبّق هذه الخوارزمية من خلال مرحلتين، هما:

- المرحلة الأولى: هي تحديد المراكز البدائية وذلك اعتماداً على خوارزمية محددة هي "Finding the initial centroids"، وهي قائمة على تشكيل عناقيد من نقاط البيانات التي تعود لأقرب مسافة إقليدية.
- المرحلة الثانية: تستخدم هذه المرحلة الخوارزمية "Assigning data-points to clusters" التي تدعم عملية تكرارية لوضع كل نقطة بيانات في العنقود الأقرب لها العائد إلى أصغر مسافة إقليدية بين كل نقطة وجميع المراكز البدائية.

باعتبار أنّ مجال البحث هو الاتصالات فإننا سنقوم بتوزيع الزبائن وفقاً لكل العوامل التالية أو بعضها على حسب رغبة صنّاع القرار:

- On net mobile revenue: أي المكالمات الهاتفية التي تجرى بين الزبائن العائدين إلى نفس المخدّم.
- SMS revenue
- GPRS revenue



الشكل 9- بنية النموذج المستخدم في التطبيق

4.2. البحث عن نتائج الاستعلام

بعد الانتهاء من عملية توزيع الزبائن، فإنّ النظام أمام خيارين ليكمل البحث، يتجليان فيما يلي:

1.4.2. التنبؤ بالأثر فقط

لقد قمنا بشرح هذه الخطوة في الفصل السابق ولكن هنا فقط سوف نشير إلى قيم المتحولات التي سنستخدمها في الأربع المعادلات السابقة:

Y: نسبة زيادة إيرادات الشركة خلال فترة العرض.

X_1 : سعر العرض.

X_2 : نسبة الزبائن المشتركين بالعرض، ويتم حسابها اعتماداً على الخوارزمية Multiple Linear Regression وفقاً للمعادلة التالية:

$$X_2 = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \qquad (5)$$

حيث أنّ:

X_3 : متوسط الاستهلاك اليومي للزبائن.

X_4 : نسبة التوفير التي يقدّمها هذا العرض، ويتم حسابها حسب المعادلة التالية:

$$X_4 = 1 - \frac{X_1}{U \times P} \qquad (6)$$

حيث أنّ:

U: تأخذ إحدى القيم التالية:

- دقائق بحال كان العرض مطبق على الاتصالات الهاتفية.
- Megabytes عندما يكون العرض موجّه لمستخدمي GPRS.
- عدد الرسائل في حالة عروض الرسائل.

P: سعر الواحدة من القيم السابقة U.

2.4.2. اقتراح السيناريوهات الأفضل والتنبؤ بالأثر

قبل البدء بشرح الخوارزمية المستخدمة في هذا المجال، سوف نذكر أحد الإشكاليات التي واجهتنا وكيفية حلها:

- الإشكالية: لا يوجد أي طريقة معتمدة لمعرفة الشرائح المتشابهة وذلك بسبب عدم وجود عدد ثابت منها، كما أنّ الحد الأدنى والأعلى من الإيرادات لكل شريحة غير ثابت.
- اقتراحنا: تقسيم الشرائح بغض النظر عن عددها إل ثلاثة تصنيفات، هي:
  - منخفض Low.
  - متوسط Medium.
  - مرتفع High.

لتحقيق هذه العملية فإنّ النظام PISWAS يطبّق الخوارزمية التالية:

الدخل:

S1: مجموعة من الشرائح الناتجة عن الخوارزمية K-means.

I: نسبة زيادة الأرباح لكل شريحة من S، هذه القيم يحددها أصحاب القرار حسب حاجة الشركة.

الخرج:

S2: السيناريوهات المقترحة الإيراد المتوقع لكل منها.

الخطوات:

1- من أجل كل شريحة segment في S1، قم بالخطوات التالية:
   a) احضر كل العروض المشابهة والتي تعود لنفس تصنيف الشريحة.
   b) احسب وسطي الإيرادات median revenue للزبائن في الشريحة ونرمز له ب M.
   c) احسب قيمة المتحول X_3 متوسط الاستهلاك اليومي للزبائن في هذه الشريحة.
   d) احسب قيمة المتحول X_1 حسب المعادلة التالية:

$$X_1 = M \times I + M$$

2- من أجل كل X_4 يأخذ قيمة بين المتحولين P_1 و P_2 والتي تتحدد قيمها من جهة المستخدمين النهائيين، قم بالخطوات التالية:
   a) احسب X_2 اعتماداً على المعادلة (5).
   b) احسب الإيراد المتوقع بناءً على تطبيق المعادلة الأولى.

3. الخلاصة

قدّمنا في هذا الفصل بنية النموذج المقترح، حيث شرحنا الإطار العام للبحث في مجال الاتصالات والحاجة الأساسية وراء عملية اختيار العروض بطريقة مدروسة، وأشرنا أيضاً إلى أهمية عملية توزيع الزبائن في شرائح متجانسة في إصدار العروض بما يتناسب مع طبيعة إيرادات واستهلاك الزبائن في الشريحة نفسها.

كما تعرّفنا على دلالة المتحولات في المعادلات التي أشرنا إليها في الفصل السابق بعد إسقاطها على مجال التطبيق في قطّاع الاتصالات. وأشرنا أيضاً بشكل أساسي إلى الخوارزمية المستخدمة في هذا البحث والتي تعود إلى مجال الاتصالات تحديداً، حيث أنّها تساهم في دعم صنّاع القرار في إصدار العروض المناسبة لكل شريحة من الزبائن المتجانسة بالسلوك من حيث الاستهلاك والإيرادات المادية على الشركة.

وبالنتيجة تدرس هذه الخوارزمية العروض السابقة المشابهة للعرض الذي يراد إصداره وتتم دراسته، وبعدها ترد مجموعة من السيناريوهات التي يحللها الإداريون في مجال التسويق ويختارون الأفضل منها بما يتناسب مع استراتيجية وعمل الشركة.

الفصل الخامس

حالة الاستخدام

نشرح في هذا الفصل كيفية التحقق من النتائج التي توصلنا لها في النموذج، وذلك بالاستعانه ببيانات زبائن حقيقية من شركة سيريتل.

"Whenever you see a successful business, someone once made a courageous decision." Peter Drucker.

1. مقدمة

لإثبات صحة النموذج الذي شرحناه في الفصلين السابقين ومعرفة النتائج المتوقعة من تطبيقه على بيانات حقيقية، قمنا ببناء تطبيق تفاعلي موجّه جديد لدعم قرار الإداريين في التسويق لاختيار العروض الأفضل والأنسب لكل شريحة من الزبائن حسب طبيعة سلوكها.

لإثبات فعالية هذا النموذج بشكل أدق وأعلى فقد حصلنا على ببيانات حقيقية من شركة سيريتل Syriatel Telecom Company[6]، تتمثل هذه البيانات فيما يلي:

- خصائص أربعة عروض تمّ إصدارها سابقاً في مجال تقسيم الزبائن إلى شرائح Segmented Offers.
- نسبة الإيرادات التي زادت خلال فترة العروض.
- معلومات عن تفاصيل الاتصالات بكل أنواعها التي قام بها الزبائن أي ما يسمى Call Details Records (CDRs) التي تعود ل 7000 رقم GSM كعينة للدراسة.
- تعود CDRs للاتصالات خلال فترة العرض، وقبل وبعد شهر من بداية العرض.
- ثلاثة عروض منها صدرت من أجل المكالمات الهاتفية المحلية فقط Mobile Calls، والعرض الرابع كان موجّه لجميع أنواع الاتصالات وهو ما يسمى Mixed Bundles .
- بعد بناء هذا التطبيق، قمنا بمقارنة نتائج العروض الحقيقية مع العروض التي تمّ اقتراحها من خلال النظام PISWAS.

2. المقارنة بين النتائج المدروسة والواقعية

اخترنا اثنين من العروض الحقيقية التي حصلنا عليها من شركة سيريتل.

يوضّح الجدولان (8) و(9) كيفية توزيع الشرائح في هذين العرضين الصادرين مسبقاً:

جدول 8 – الشرائح الخاصة بالعرض الأول

الشريحة	الإيرادات من	الإيرادات إلى	تصنيف الشريحة
1	25	40	Low
2	41	70	Medium
3	70	-	High

[6] Syriatel Telecom Company which has been leading the Syrian mobile telecommunication market since 2000. The company has successfully established its reputation by focusing on customer satisfaction and social responsibility.

جدول 9 – توزيع الشرائح في العرض الحقيقي الثاني

الشريحة	الإيرادات من	الإيرادات إلى	تصنيف الشريحة
1	0	16	Low
2	17	33	Medium
3	34	71	Medium
4	72	137	High

كما يوضّح الجدولين التاليين (10) و (11) الباقات الموافقة للعرضين السابقين والتي تخض كل شريحة من الشرائح:

جدول 10 – خصائص باقة العرض الأول

الشريحة	الدقائق	سعر الباقة	نسبة التوفير	نسبة المشاركة اليومية	نسبة زيادة الإيراد اليومية
1	60	30	94.4	1.9	8.76
2	60	40	92.5	1.8	9.1
3	60	75	86.1	1.6	2.5

جدول 11– خصائص باقة العرض الثاني

الشريحة	الدقائق	سعر الباقة	نسبة التوفير	نسبة المشاركة اليومية	نسبة زيادة الإيراد اليومية
1	50	75	78.5	1.1	4.88
2	100	100	78.5	1.2	3.5
3	250	300	82.8	1.4	8.5
4	500	600	82.8	1.2	3.6

أدخلنا إلى النظام PISWAS المدخلات التالية، والتي هي نفس خصائص العرضين الحقيقين (اختبرنا كل عرض على حدى):

- زيادة الإيرادات لكل شريحة حسب الجدول (12):

جدول 12- نسبة زيادة الإيرادات

الشريحة	العرض الأول (%)	العرض الثاني (%)
1	90	80
2	70	70
3	50	20
4		10

قيم المتحولات P1,P2 هي 90%، 50% على الترتيب.

- تحديد فترات الدراسة لكل عرض والتي هي نفسها الفترات التي تمت الدراسة عليها في العروض الحقيقية.

بعدها طلبنا من النظام اقتراح السيناريوهات المتوقعة للوصول إلى الأهداف السابقة، يوضّح الجدولين (13) و(14) توزيع الزبائن على شرائح متجانسة لكل من العرضين السابقين كخطوة بدائية قبل البدء بالتحليل واقتراح السيناريوهات الأفضل:

جدول 13- توزيع الشرائح في العرض الأول

الشريحة	الإيرادات من	الإيرادات إلى	تصنيف الشريحة
1	0	14.9	Low
2	14.10	43.6	Medium
3	43.8	150.2	High

جدول 14- توزيع الشرائح في العرض الثاني

الشريحة	الإيرادات من	الإيرادات إلى	تصنيف الشريحة
1	0	27.3	Low
2	27.4	57.7	Low
3	59.3	144.1	Medium
4	191.7	308.9	High

بعد ذلك تمّ اقتراح السيناريوهات من قبل نظام PISWAS كما هي مذكورة في الجدولين التاليين (15) و(16):

جدول 15- السيناريوهات المقترحة للعرض الأول

الشريحة	الدقائق	سعر الباقة	نسبة التوفير	نسبة المشاركة اليومية	نسبة زيادة الإيراد اليومية
1	3	17	50	2.87	12.31
1	4	17	60	2.37	9.89
1	5	17	70	1.87	7.46
1	7	17	80	1.37	5.04
1	13	17	90	0.87	2.61
2	4	28	50	0.35	0.03
2	5	28	60	0.68	2.36
2	7	28	70	1.02	4.77
2	11	28	80	1.36	7.17
2	22	28	90	1.69	9.51
3	30	198	50	0.07	0.4
3	38	198	60	0.27	0.6
3	51	198	70	0.6	1.56
3	76	198	80	0.94	2.55
3	152	198	90	1.27	3.51

جدول 16- السيناريوهات المقترحة للعرض الثاني

الشريحة	الدقائق	سعر الباقة	نسبة التوفير	نسبة المشاركة اليومية	نسبة زيادة الإيراد اليومية
1	4	25	50	3.01	13.68
1	5	25	60	2.51	11.26
1	6	25	70	2.01	8.83
1	10	25	80	1.51	6.41
1	19	25	90	1.01	3.98
2	8	52	5	2.69	14.46
2	10	52	60	2.19	12.03
2	13	52	70	1.69	9.61

2	20	52	80	1.19	7.18
2	40	52	90	0.69	4.76
3	8	51	50	0.29	0.4
3	10	51	60	0.63	2.01
3	13	51	70	0.96	4.34
3	20	51	80	1.3	6.75
3	39	51	90	1.64	9.15
4	10	64	50	0.03	0.05
4	12	64	60	0.3	0.92
4	16	64	70	0.64	1.91
4	25	64	80	0.97	2.87
4	49	64	90	1.31	3.86

1.2. مناقشة النتائج السابقة

- الزبائن العائدة للشرائح ذات التصنيف "المنخفض" تميل وتفضل الاشتراك في العروض التي تقدّم باقات من الدقائق القريبة إلى استهلاكها، فحتى لو زاد عدد الدقائق مع زيادة في نسبة التوفير فهي لن تفضل الاشتراك في هذه العروض لأنّ استهلاكها محدود.
- الزبائن العائدة للشرائح ذات التصنيف "المتوسط" و"المرتفع" تفضّل العروض التي تقدّم لها أكبر نسبة توفير على عكس سلوك الزبائن في الشرائح ذات التصنيف المنخفض.
- **الإيرادات**: نجد أنّ السيناريوهات التي تمّ اقتراحها من قبل النظام PISWAS تحقق إيرادات أكبر من السيناريوهات الحقيقية المطبّقة في العروض المدروسة، وكمثال على ذلك: الشريحة الثالثة في الجدول (9) زادت إيراد الشركة خلال فترة العرض بنسبة 2.5%، بينما بحال لو طبقت الشركة السيناريو الخامس المقترح للشريحة الثالثة حسب الجدول (14) فكان قد زادت الإيرادات بنسبة 3.51%.
- بحال كان المستخدمون النهائيون يرغبون بزيادة الإيرادات بشكل أكبر، فإنّهم يستطيعون التفاعل مع التطبيق وإدخال نسب أعلى من زيادة الإيرادات على مستوى كل شريحة وبالتالي ينعكس ذلك على الإيراد الكلي للشركة.
- **الدقة**: أثبت النظام PISWAS دقة نتائجه أكثر من النتائج التي تعود للعروض الحقيقية، وذلك لأنّه يعتمد على دراسة وتحليل الحالات المشابهة الحقيقية السابقة وليس على خبرة الإداريين أصحاب القرار في إصدار العروض وتقسيم الزبائن في شرائح غير مدروسة وغير معتمدة على خوارزميات تنقيب.

2.2. حساب ال Average Revenue Per User (ARPU)

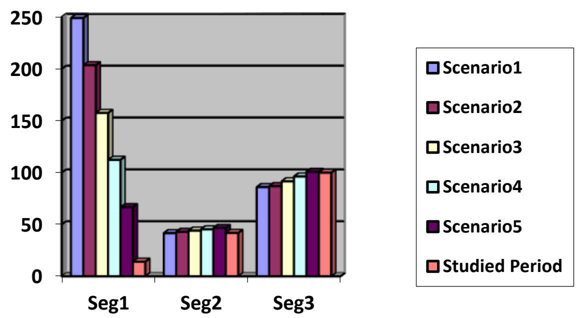
يعرّف هذا المصطلح على أنّه مقياس الإيراد على مستوى الزبون، وبشكل عام يقاس بتقسيم المجموع الكلي للإيرادات على عدد الزبائن[27] ، فهو يعطينا بشكل أدق مقدار الزيادة في الإيرادات على مستوى الزبون.

قمنا بحساب قيم ARPU التالية:

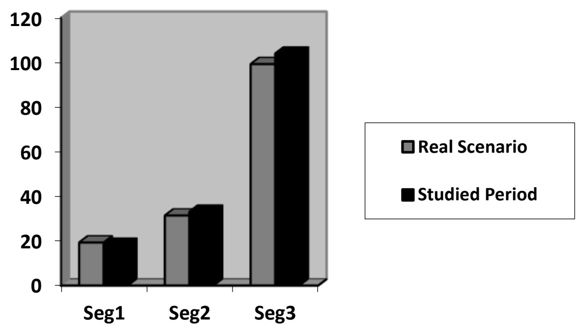
- خلال فترة الدراسة.
- خلال فترة العروض الحقيقة والعروض المقترحة من النظام PISWAS.

وذلك لاكتشاف فيما إذاكان العرض رابحاً أم لا بناءً على مقارنة ال ARPU قبل إصدار العرض وبعده على مستوى كل شريحة.

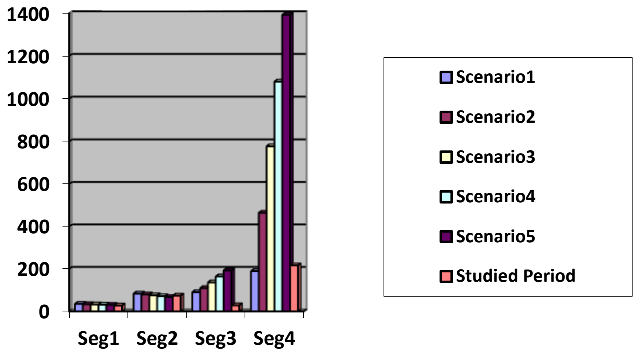
وكانت النتيجة في المخططات التالية:



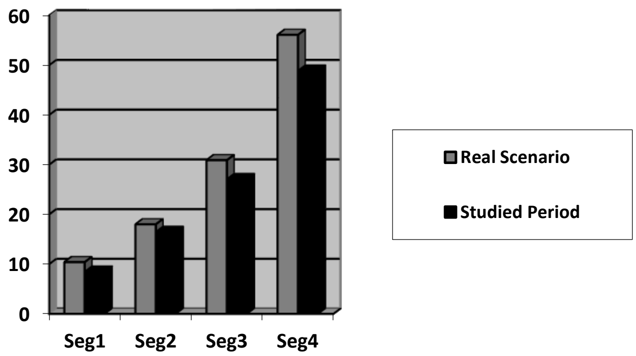
الشكل 10- ARPU من أجل السيناريوهات المقترحة العرض الأول



الشكل 11- ARPU من أجل العرض الأول



الشكل 12- ARPU من أجل السيناريوهات المقترحة العرض الثاني



الشكل 13- ARPU من أجل العرض الثاني

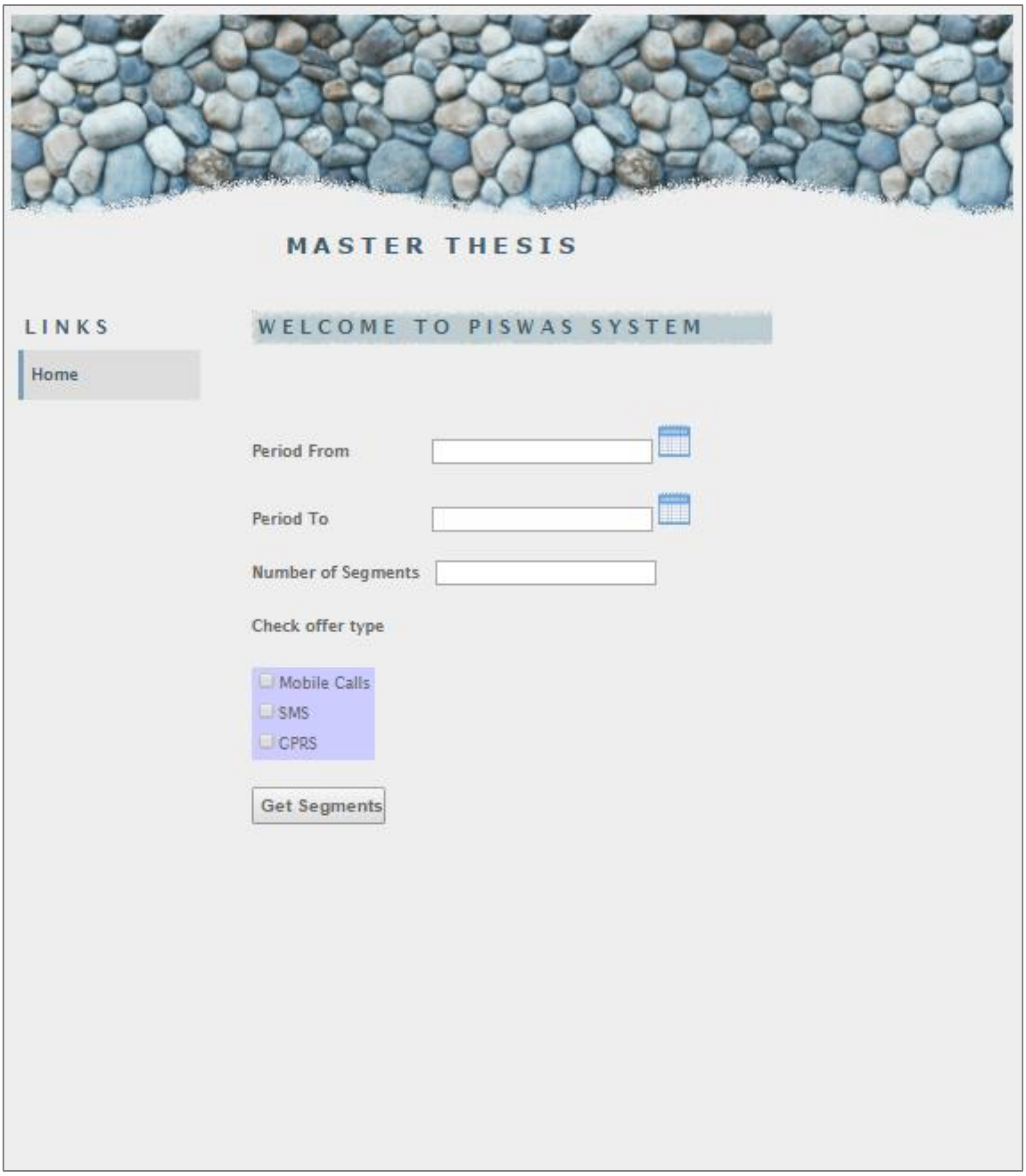
3. واجهة التطبيق

سنعرض في هذه الفقرة واجهات التعامل مع النظام PISWAS، يوضح الشكل (14) الصفحة الرئيسية التي تعطي لمحة شاملة عن النظام وآلية عمله.

يختار المستخدم أحد الرابطين Predict Impacts أو Suggest Scenarios وذلك حسب الرغبة وحاجة المؤسسة.

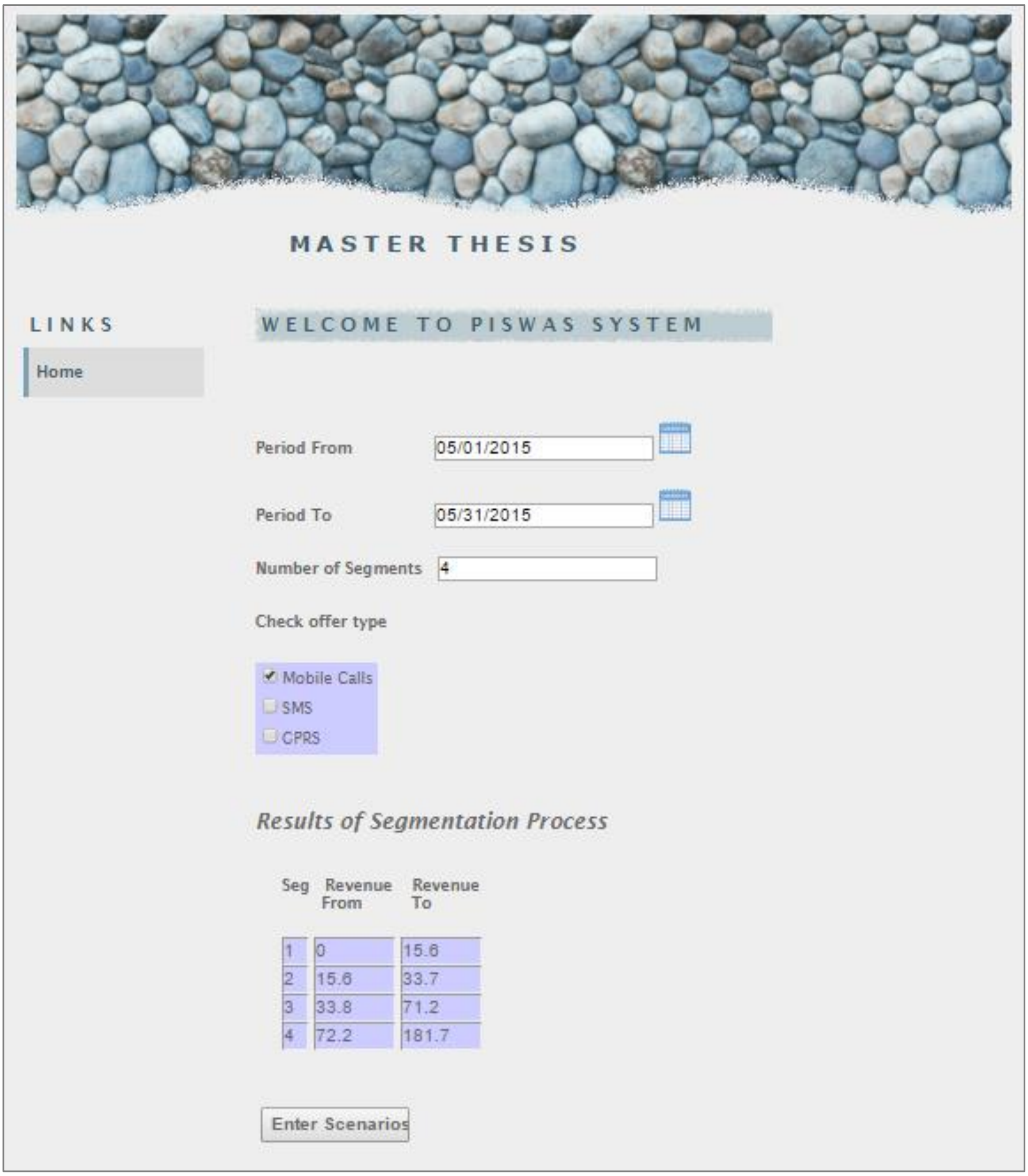
الشكل 14- الصفحة الرئيسية

في حال تم الضغط على الرابط الأول، ستظهر الصفحة التالية حيث سيتم إدخال البيانات المطلوبة (فترة الدراسة، عدد الشرائح، نوع العرض المدروس) ومن ثمّ الضغط على الزر "Get Segments".



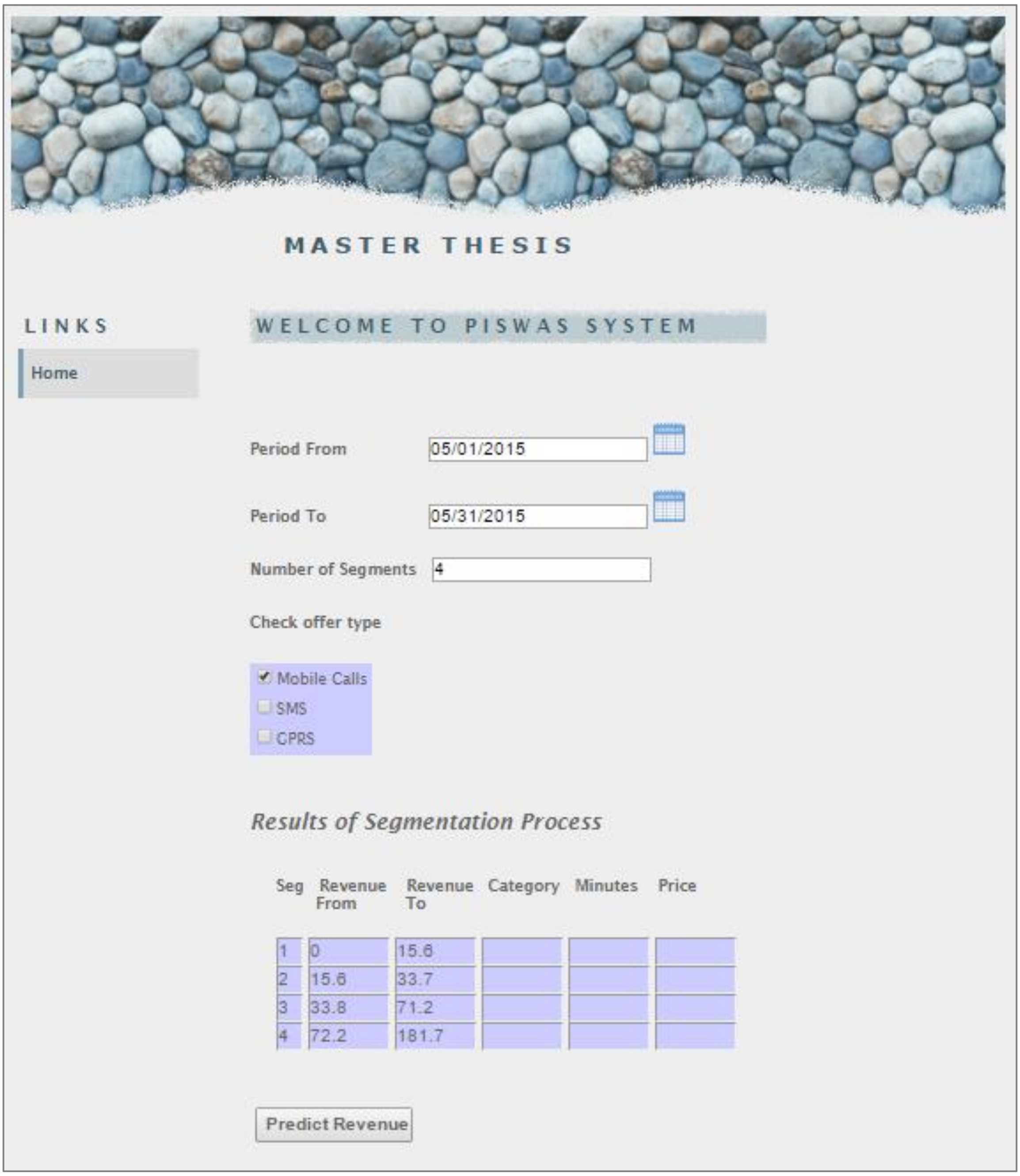
الشكل 15- صفحة ال Predict

بعدها تظهر نتائج توزيع الزبائن في شرائح متجانسة كمايلي:



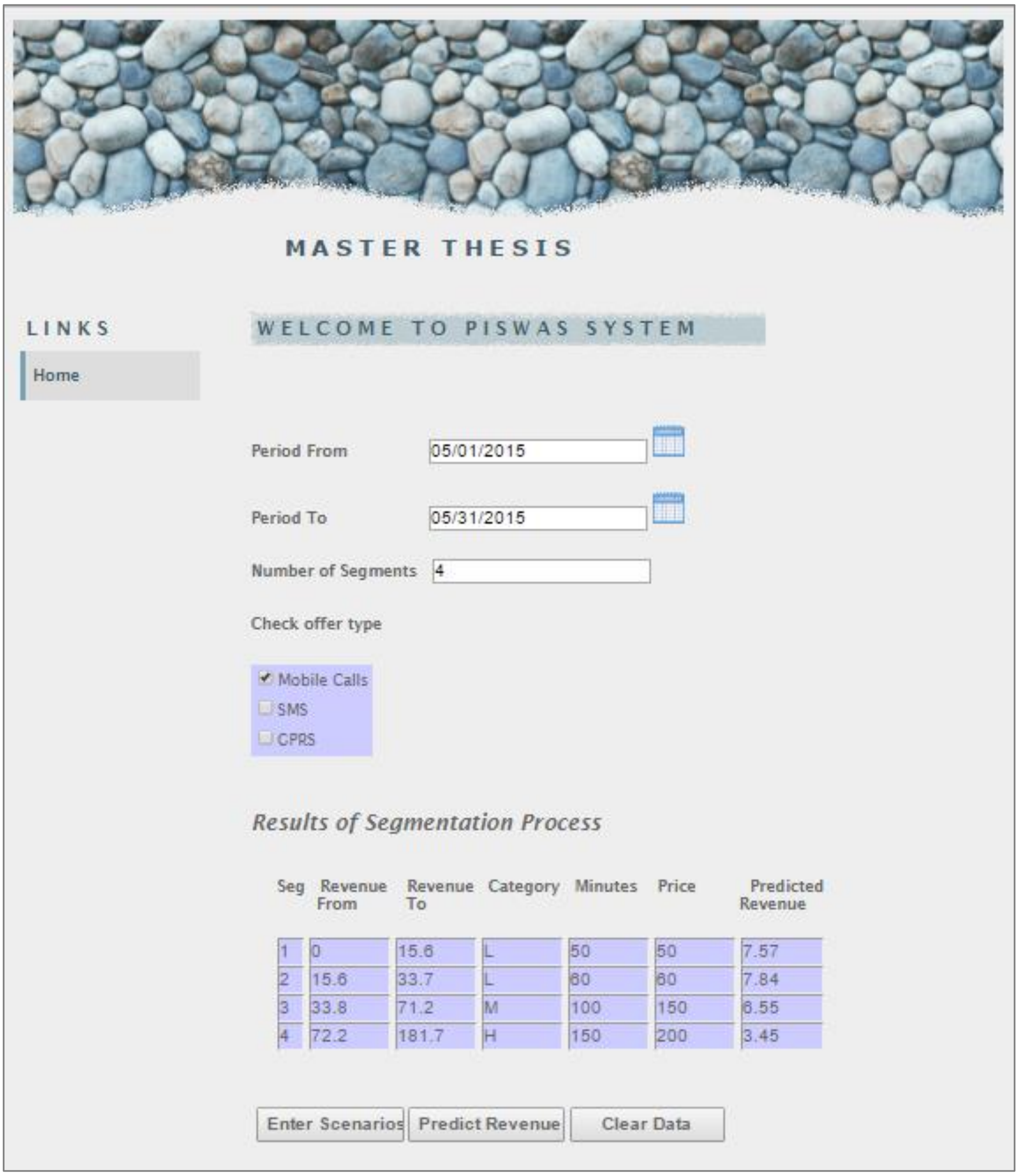
الشكل 16- عرض نتيجة التوزيع إلى شرائح

بعدها نضغط على "Enter Scenarios" ونحدد المعلومات التي تظهر في الشكل التالي:



الشكل 17- إدخال السيناريوهات المرادة

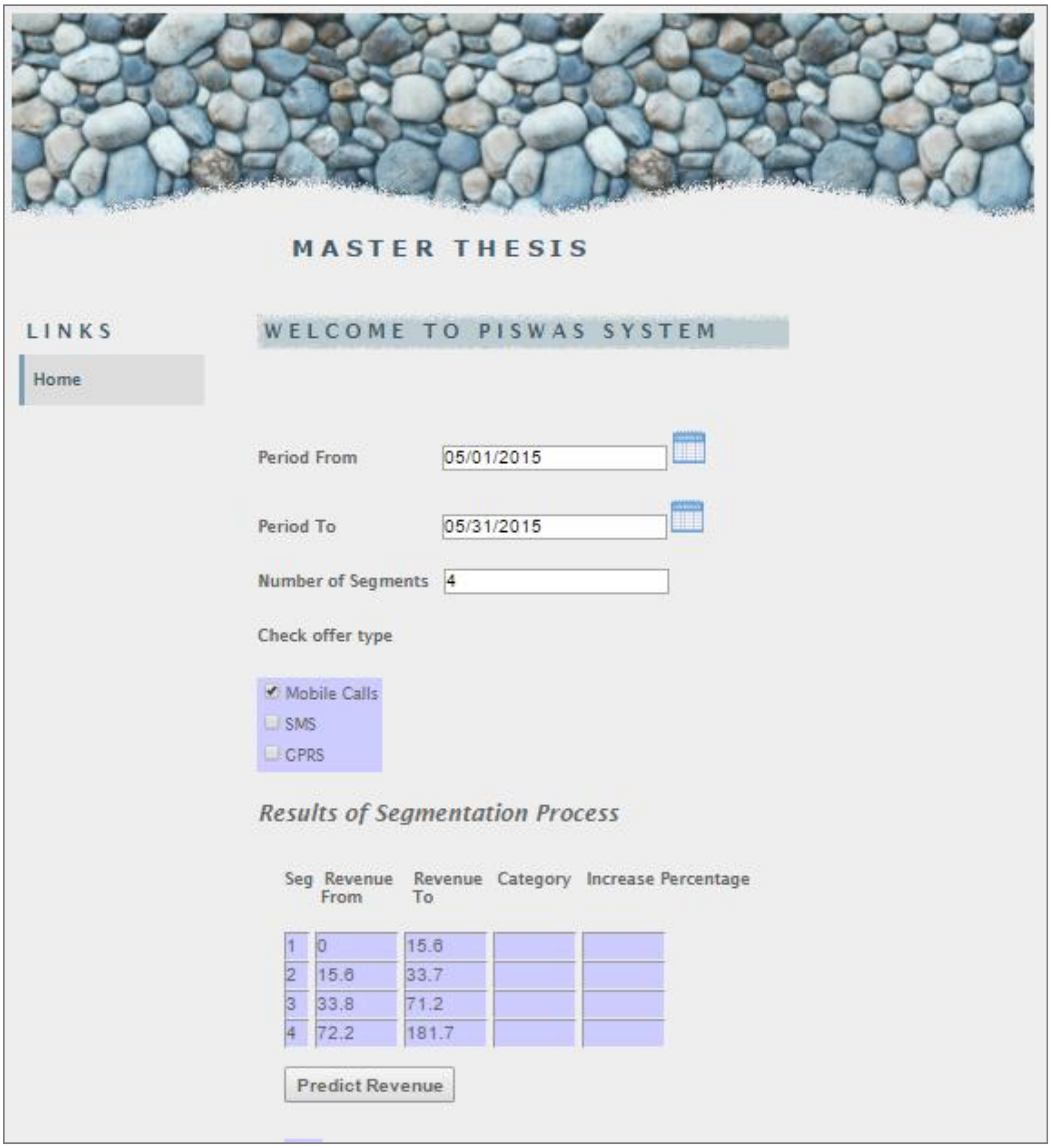
وبعد الضغط على "Predict Revenue" تظهر النتائج المتوقعة لكل سيناريو وأيضاً ممكن أن يتفاعل المستخدم النهائي مع النظام ويعيد إدخال سيناريوهات أخرى ويحسب أثرها، كمايلي:



الشكل 18- الإيرادات المتوقعة لكل سيناريو

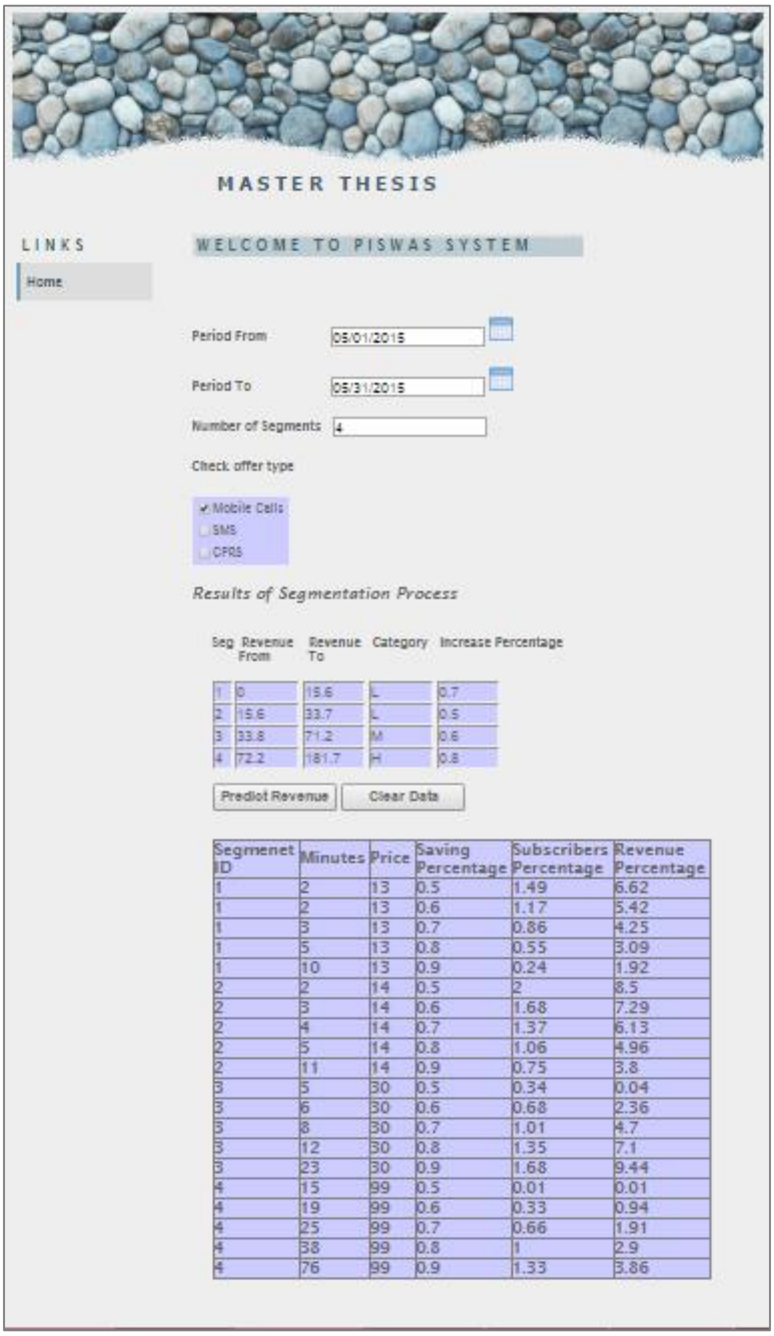
أمّا في حال ضغط المستخدم على الرابط "Suggest Scenarios" فسيظهر له بدايةً كما ظهر في الشكل (15) حيث ستتم عملية التوزيع إلى شرائح.

وبعد ظهور نتائج التوزيع، نضغط على "More details " وندخل المعطيات المطلوبة:



الشكل 19– معلومات إضافية للمساعدة في التنبؤ

ومن ثمّ يتم تحديد تصنيف الشرائح ونسبة زيادة الإيرادات المطلوبة لكل شريحة، وبعدها يقترح النظام مجموعة من السيناريوهات المتوقعة:



الشكل 20– السيناريوهات المقترحة

4. الخلاصة

تحدثنا بشكل مفصّل عن كيفية التحقق من النموذج الذي اقترحناه، وذلك ببناء تطبيق جديد يعتمد على بيانات حقيقية من شركة سيريتل. أدخلنا إلى النظام PISWAS نفس خصائص عرضين صادرين من بين الأربعة عروض التي حصلنا عليها من الشركة، وقد كانا موجّهين للمكالمات الهاتفية المحلية فقط.

تحققنا من صحة النتائج التي توصلنا إليها بمقارنتها مع نتائج العروض الحقيقية من خلال حساب ARPU، وكانت كل النتائج تشير إلى أنّها دقيقة وتزيد الأرباح أكثر من العروض الحقيقية. وفي النهاية قمنا بعرض واجهات التطبيق الذي قمنا ببنائه حيث أنّه باستطاعة المستخدمين النهائيين التفاعل معه للحصول على نتائج مرضية بالنسبة لهم ولشركتهم.

الفصل السادس

الخاتمة والآفاق المستقبلة

"Public decision – making does not lend itself to certitude. " Jim Leach.

قدمنا في هذا العمل طريقة لدعم القرار في المؤسسات الكبيرة والابتعاد عن الاعتماد على خبرة الإداريين في اتخاذ القرارات، وذلك من خلال الاعتماد على تحليل ماذا–لو. اقترحنا نظام جديد يعتمد على تحليل البيانات المخزّنة ودراستها بناءً على خوارزميات تنقيب وبعدها يقوم باقتراح السيناريوهات الممكنة على أصحاب القرار.

تحليل ماذا–لو عادةً ما يكون موجّهاً لتطوير دعم القرار لغرض معين في مجال ما إمّا تقني أو مهني، لذلك حاولنا في هذا البحث أن نعمّم النموذج المقترح قدر الإمكان ضمن إطار عام شامل إلى حدٍ ما، ولكن على الرغم من ذلك تبقى هناك أمور تتناسب فقط مع طبيعة مجال العمل في التطبيق والغاية من الاعتماد على التحليل ماذا–لو، فمثلاً عملية توزيع الزبائن إلى شرائح متجانسة هي عملية تخص فقط العمل في شركات الاتصالات ولا يمكن تطبيقها في قطاعات أخرى لأننا اعتمدنا على إيرادات الزبائن في المكالمات الهاتفية والرسائل النصية ونقل البيانات.

أيضاً يجب أن ننوه إلى أنّ النتائج التي توصلنا إليها هي نتائج أولية وبالتأكيد سوف تصبح أدق مع استخدام النظام باستمرار، وذلك لأنّ النظام PISWAS هو نظام تعلّمي لاعتماده على خوارزميات تنقيب تعلمية.

كإضافات وتطلعات مستقبلية، فإننّا سنعمل على تطوير النظام للتنبؤ بالأثر على أكثر من متحول واحد، وأيضاً العمل على اقتراح عدد الشرائح الأمثل في عملية توزيع الزبائن أي تحديد عدد العناقيد K عند تطبيق خوارزمية K-means .

المراجع

1. Manon G. Guillemette, Maxime Laroche, Jean Cadieux, **Defining decision making process performance: Conceptualization and validation of an index**, (Elsevier) 618–626, 2014.
2. Maris G. Martinsons, Robert M. Davison, **Strategic decision making and support systems: Comparing American, Japanese and Chinese management**, (Elsevier) 284–300, 2007.
3. Matteo Golefarelli, Stefano Rizzi, **What-if simulation modeling in business intellegance, international journal of data warehousing and mining**, X(X), X-X, Oct-Dec 2008 1, 2008.
4. Matteo Golfarelli, Stefano Rizzi, Andrea Proli, **Designing What-if Analysis: Towards a Methodology**, 2006.
5. Steeano Rizzi - Encyclopedia of Database Systems, **What-If Analysis**, 2009.
6. Paul Rogers, Marcia Blenko, **what you don't know about making decisions**, Harvard business review, 2006.
7. Bob Frisch, **What Makes a Decisive Leadership Team 3rd Edition, Harvard business review**, 2006.
8. C. Rolland, C. Ben Achour, C. Cauvet, J. Ralyté, A. Sutcliffe, N-A. M. Maiden, M. Jarke, P. Haumer, K. Pohl, E. Dubois and P. Heymans, **A Proposal for a Scenario Classification Framework, Requirements** Engineering Journal (REJ), Vol. 3, N°1, pp. 23-47,1998.
9. M. Jarke, K. Pohl**, Requirements Engineering: An Integrated View of Representation, Process and Domain**, In Proc. of the 4th European Software Conference, Springer Verlag, 1993.
10. M. Jarke, J. Mylopoulos, J.M. Schmidt, Y. Vassiliou, **DAIDA - An Environment for Evolving Information Systems, ACM Trans., In Information Systems**, vol. 10, n° 1, 1992.
11. C. Rolland, **A primer for method engineering, proceeding of INFORSID'97**, Toulouse, France, 1997.
12. Rahul Singh, Prashant Shenoy, Maitreya Natu, Vaishali Sadaphal and Harrick Vin, **Analytical Modeling for What-if Analysis in Complex Cloud Computing Applications**, 2013.
13. Katarína Belanová, **"What-if" Analysis In Investment Decision Making** 2012.
14. Christian Bird, Thomas Zimmermann, **Assessing the value of Branches with What-if Analysis**, 2012.
15. Daniel Deutch, Zachary Ives, Tova Milo, Val Tannen, **Caravan: Provisioning for What-If Analysis,** 2013**.**
16. Herodotos Herodotou, Shivnath Babu, **Profiling, What-if analysis, and Cost-based Optimization of MapReduce Programs**, 2011.
17. M.Esmaeili, M.Osanloo, F.Rashidinejad, A. A.Bazzazi, M.Taji, **Multiple regression, ANN and ANFIS models for prediction of backbreak in the open pit blasting. Engineering with computers**, Vol. 30 n. 4, pp. 549-558, 2014.
18. Jan Adamowski, Hiu Fung Chan, Shiv O. Prasher, Bogdan Ozga-Zielinski, Anna Sliusarieva, **Comparision of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network methods for urban water demand forecasting in Montreal, Canada**, 10.1029/2010WR009945, 2012.

19. Colin Cameron, Frank A.G. Windmeijer, **An R-squared measure of goodness of fit for some common nonlinear regression models**, 1995.
20. Mohammad Ridwan Ismail1, Mohd Khalid Awang2, M Nordin A Rahman2 and Mokhairi Makhtar2, **A Multi-Layer Perceptron Approach for Customer Churn Prediction**, Vol.10, No.7, pp.213-222, 2015
21. Dr. Mrinal Kanti Das, **Developing Payout-Loyalty Matrix for Telecom Retailers – A Study on Murshidabad and Nadia districts, Social Science Electronic Publishing** (SSRN), 2016.
22. Dr. Hurmaz D. Patel, Dr. Mehul P.Desai, **A Study on Relationship between Customer Satisfaction, Customer Loyalty and Intention to switch from one bank to another bank surat ciy**, (ISSN 0974-7028) - Vol. : 9 Issue : 1, 2016
23. Nazeeh A. Ghatasheh, Anas R. Najdawi, Mua'ad M. Abu-Faraj, Hossam Faris, **Corporate e-learning environment using concept maps: a case study**, (IRECOS) Vol. 8, N. 11 ISSN 1828-6003, 2013.
24. D. Selvaraj, R.Dhanasekran, **AReview on current MRI brain tissue segmentation, feature extraction and classification techniques**, (IJECIERD) Vol. 3, Issue 5, 2013, Dec 2013, 11-30, 2013.
25. J.E. Judith, J.Jayakumari, **Enhanced Distributed text document clustering based on semantics**, (IRECOS) Vol. 8,N. 10 ISSN 1828-6003, 2013.
26. K. A. Abdul Nazeer, M. P. Sebastian **Improving the Accuracy and Efficiency of the k-means Clustering Algorithm**, ISBN: 978-988-17012-5-1, 2009
27. Soorena Merat, P.Eng, Dr. Wahab Almuhtadi, P.Eng, **STANDARD ARPU CALCULATION IMPROVEMENT USING ARTIFICIAL INTELLIGENT TECHNIQUES, INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS** VOL. 8, NO. 4, DECEMBER 2015 1917, 2015.