تقدير عدد مطالبات تأمين السيارات التكميلي باستخدام النماذج التجميعية المعممة Generalized Additive Models وأساليب التسوية LASSO Regularization بالتطبيق على إحدى شركات التأمين العاملة في سوق التأمين المصري

إعداد

د. زهرة صلاح الدين جمعة حسن

مدرس بقسم التأمين والعلوم الأكتوارية – كلية التجارة - جامعة القاهرة

أ.د/ محمد عبد المنعم جودة حُزين

أستاذ بقسم التأمين والعلوم الأكتوارية – كلية التجارة - جامعة القاهرة

د. أسماء محمد سعد حسين

مدرس بقسم التأمين والعلوم الأكتوارية – كلية التجارة - جامعة القاهرة

المستخلص:

يتمثل هدف البحث في تقدير عدد المطالبات في فرع تأمين السيارات التكميلي، باستخدام النماذج التجميعية المعممة Generalized Additive Models، واقترح الباحثون دمج أساليب التسوية LASSO Regularization لزباده قدرة النموذج على التنبؤ. واعتمد الباحثون على بيانات تأمين السيارات التكميلي لإحدى شركات التأمين العاملة في سوق التأمين المصري. وتم إدخال تأثير عوامل خطر إضافية لم تؤخذ في الحسبان من قبل ومنها المسافات المقطوعة والتقييم الائتماني I-SCORES. وقد تم استخدام النماذج التجميعية المعممة GAM في برنامج *R package*, 'mgcv'، والتي تعتمد في التنبؤ بالمتغيرات غير الخطية على الدوال الممهدة smooth Functions، وهذا ما يعطيه الدقة العالية في التنبؤ أكثر من غيره من النماذج.

وتم التوصل إلى أن هناك العديد من المتغيرات التي تؤثر في التنبؤ بعدد المطالبات، وهي: عمر السيارة، وعمر قائد السيارة، والمسافات المقطوعة بالأميال، والأداء الائتماني. وقد أثبت النموذج قدرته على التنبؤ، حيث أن النموذج التجميعي المعمم مع استخدام معيار التسوية L1 regularization، يعمل على تقليل معاملات المتغيرات βj ومن الممكن أن تصبح صفر، مما يؤدي إلى ملائمة النموذج التجميعي المعمم ذو معيار التسوية للدوال الممهدة.

الكلمات الدالة: (تأمين السيارات التكميلي – النماذج التجميعية المعممة– عدد المطالبات – الدوال الممهدة -أسلوب التسويةLASSO Regularization).

Estimating The Frequency of Claims in The Private Motor Insurance using Generalized Additive with Regularization Lasso with Application to one of The Insurance Companies in The Egyptian Insurance Market.

Abstract

This research aims to estimate the frequency of claims in the private Motor Insurance industry in Egypt by using Generalized Additive Models (GAMs), which can capture common non-linear complex relationships by combining multiple smooth functions of predictor variables additive models (AMs).

The researchers proposed a model for pricing Motor insurance that incorporates **LASSO** regularization methods into GAM models, to increase the predictive performance of the model. Additionally, this Research introduced new risk factors that had not been taken into consideration to estimate the number of claims in pricing Motor insurance in the Egyptian Market, **Distances and I-SCORES credit rating.** The researchers used Motor insurance Line of Business data from one of the insurance companies in the Egyptian insurance market.

To achieve this objective, The statistical-analytical and Machine learning methods with **"mgcv"** R package are used.

The findings reveal that the proposed model of GAM model with LASSO regularization L1 Techniques (ML) helps in improving its interpretability of the results and increasing predictive performance, selecting the most predictive risk factors on Pricing Motor insurance.

مقدمة:

يعد تحديد تكلفة خدمة الحماية التأمينية من أهم القرارات التي تتخذها شركات التأمين، ويرجع استمرار شركات التأمين في العمل إلى التحديد الدقيق لأسعار الخدمات التأمينية المقدمة؛ وأن يحقق هذا السعر العدالة لكلاً من شركة التأمين والمؤمن له، والكفاية للمطالبات المتوقعة الحدوث وما تتطلبه من مصاريف تسوية، والمعقولية في تحقيق هامش الربح. وبالتالي فإن الدور الأساسي للتأمين هو توفير الحماية المالية، وتقديم وسيلة لنقل الخطر مقابل دفع قسط التأمين. (درويش، ٢٠١٣)

ويتكون قسط التأمين من عدة عناصر؛ وهي: المبلغ المطلوب لسداد المطالبات معبراً عنه بنسبة مئوية ويعرف بمعدل المطالبات، ومخصص لكل من التكاليف الثابتة والمتغيرة، وتكاليف التشغيل للشركة، وأيضا تكاليف الدعاية. وبما أن جميع الأخطار ليست متساوية، فمن الطبيعي أن يدفع كل مؤمن له قسطاً يتناسب مع درجة الخطر. (المعداوي وآخرون، ٢٠١٧) (عبد الرحمن وآخرون، ٢٠١٤)

وتتمثل مشكلة حساب سعر تأمين السيارات التكميلي في السوق المصري بصفة خاصة في جانبين: (David, etal، ٢٠١٥) (عبدالعليم، ٢٠١٦)

- الأول: أن تسعير تأمين السيارات التكميلي يتم على أساس القيمة السوقية للسيارة ونوعها سواء كان ملاكي أم غير ملاكي لتحديد سعر التأمين مع إغفال باقي العوامل.
- الثاني: أن التسعير يعتمد على خبرة الأسواق الأخرى وخاصة السوق الإنجليزي مع إضافة بعض التعديلات التي تتم على أساس التقدير الشخصي للمكتتب.

ويتم إغفال الأسس الرياضية والإحصائية المفترض استخدامها لتقدير الأسعار المناسبة، وهذا قد يعرض شركات التأمين لهزات مالية لا تسمح لها بالاستمرار في مزاولة أنشطتها التأمينية وقد يؤثر بالتبعية على مقدرة شركات التأمين على الوفاء بالتزاماتها المستقبلية تجاه حملة الوثائق.

مشكلة البحث:

يحظى فرع تأمين السيارات التكميلي باهتمام متزايد، لأنه يعد أهم وأكبر فروع التأمينات العامة بعد التأمين الطبي من حيث حجم الأقساط وحجم التعويضات. حيث بلغ حجم أقساط تأمين السيارات التكميلي نسبة ٢١,٢% من إجمالي أقساط محفظة التأمينات العامة في السوق المصري، وبلغ حجم التعويضات نسبة ٢٧,٧% من إجمالي التعويضات المسددة عن العمليات المباشرة في عام ٢٠٢٢.

ويشير قسط تأمين السيارات التكميلي إلى القيمة المتوقعة للخسائر المستقبلية، والذي يعكس نجاح شركة التأمين في التنبؤ بمبالغ المطالبات المتوقعة، وبالتالي يتم النظر إلى كل من: الأقساط والتعويضات ومعدل نمو كل منهما ومعدل الخسائر لفرع تأمين السيارات التكميلي في سوق التأمين المصري خلال الفترة من عام ٢٠١٣/٢٠١٢ إلى عام ٢٠٢٢/٢٠٢١ والتي يعرضها الجدول رقم (١) التالي:

جدول رقم (١)
معدل نمو الأقساط والتعويضات ومعدل الخسائر لفرع تأمين السيارات التكميلي
خلال الفترة من عام ٢٠١٣/٢٠١٢ إلى عام ٢٠٢٢/٢٠٢١

القيم بالألف جنيه

العام	الأقساط		التعويضات		معدل الخسائر %
	قيمة الأقساط	معدل نمو الأقساط %	قيمة التعويضات	معدل نمو التعويضات %	
٢٠١٣/٢٠١٢	١٣٢٥١٥٧	-	٩٣٠٤٤٦	-	٦٧,٦
٢٠١٤/٢٠١٣	١٥٧١٩٧٦	١٨,٦	٩١٩٩١١	٠,٩٨	٥٥,٦
٢٠١٥/٢٠١٤	١٧٩٦١٦٧	٣٥,٥	١٠٩٧٨٠١	١٧,٩	٦٧,٦
٢٠١٦/٢٠١٥	٢٠٤٢٢٩٨	٥٤,١	١١٦٥٧٤٢	٢٥,٣	٥٧,٦
٢٠١٧/٢٠١٦	٢٦٩٦٠٢٤	١٠٣,٤	١٣٥٩٦٩٦	٤٦,١	٥٦,٦
٢٠١٨/٢٠١٧	٣٢٢٥٣٨٤	١٤٣,٤	١٧٢٧٠٠٠	٨٥,٦	٦١,٩
٢٠١٩/٢٠١٨	٣٧٧١٤٩٩	١٨٤,٦	١٩٩٥٤٨٢	١١٤,٥	٥٧,٩
٢٠٢٠/٢٠١٩	٣٩٥٦٥١٢	١٩٨,٦	٢١٤٥٩٣٥	١٣٠,٦	٥٧,٧
٢٠٢١/٢٠٢٠	٤٥٢٧٧٥٣	٢٤١,٧	٢٣٧٤١٣٨	١٥٥,٢	٥٧,١
٢٠٢٢/٢٠٢١	٥٤٣١١٨٣	٣٠٩,٩	٢٨١٣٢٥٥	٢٠٢,٤	٦٢,٢

المصدر: الكتاب الإحصائي السنوي– هيئة الرقابة المالية– أعداد مختلفة من٢٠١٣/٢٠١٢حتى ٢٠٢٢/٢٠٢١

ويتضح من الجدول السابق أن:

- قيمة الأقساط في تزايد مستمر خلال مدة الدراسة فقد كانت ١٣٢٥١٥٧ ألف جنيه في عام ٢٠١٣/٢٠١٢ ثم أخذت في الزيادة إلى أن وصلت إلى ٥٤٣١١٨٣ ألف جنيه في عام ٢٠٢٢/٢٠٢١. أما قيمة التعويضات فقد كانت ٩٣٠٤٤٦ ألف جنيه في عام ٢٠١٣/٢٠١٢ ثم قلت بنسبة طفيفة لا تتعدى ٢% في عام

٢٠١٤/٢٠١٣ إلى ٩١٩٩١١ ألف جنيه، ثم بعد ذلك أخذت في الزيادة إلى أن وصلت إلى ٢٨١٣٢٥٥ ألف جنيه في عام ٢٠٢٢/٢٠٢١.

- قام الباحثون بحساب معدلات النمو باتخاذ سنة ٢٠١٣/٢٠١٢ هي سنة الأساس، وكان معدل النمو للأقساط في تزايد مستمر وبمعدل أكبر بكثير من معدل النمو في التعويضات، حيث وصل معدل النمو في الأقساط إلى ٣٠٩,٩% في عام ٢٠٢٢/٢٠٢١، بينما وصل معدل النمو في التعويضات المدفوعة إلى ٢٠٢,٤% في عام ٢٠٢٢/٢٠٢١.
- يتراوح معدل الخسارة خلال سنوات الدراسة ما بين ٥٥,٦% و٦٧,٦%. مما يدل على وجود فجوة بين الأقساط التي يتم تحصيلها والتعويضات المدفوعة.

وبناءً على ما سبق، فإن مشكلة الدراسة تكمن في دقة التنبؤ بالمطالبات المتوقعة في فرع تأمين السيارات التكميلي، لأن الدقة في التنبؤ بالمطالبات المتوقعة يساعد شركات التأمين في تسعير الأخطار، ويكون هناك تناسبا بين القسط والقيمة المتوقعة لتكلفة المطالبات بالنسبة لدرجة الخطر. ويتم ذلك بالاعتماد على الأساليب الرياضية والاحصائية، وعدم الاعتماد على الأساليب التقليدية في التنبؤ لما تنتاب هذه الطرق التقليدية من مشاكل في التطبيق.

أي أن هناك حاجة إلى إعادة النظر في نماذج التسعير في شركات التأمين.

الدراسات السابقة:

يتم عرض أهم الدراسات التي تناولت تأمين السيارات التكميلي وتسعيره والعوامل التي تؤثر في تحديد المطالبات، وأيضا الدراسات التي طبقت النماذج التجميعية المعممة في مجال التأمين، ثم التعليق عليها وتحليلها، وذلك بغرض الاستفادة مما توصلت إليه تلك الدراسات من نتائج وتوصيات في موضوع البحث الحالي، وأخيراً الوصول إلى الفجوة البحثية والتي يبنى عليها هذا البحث. وقد تم تقسيم تلك الدراسات إلى ثلاث مجموعات، وهم:

أ. المجموعة الأولى: دراسات باللغة العربية:

هدفت دراسة (الديب، ١٩٩٢) إلى وضع نموذج لتسعير التأمين التكميلي للسيارات الملاكي معتمداً على ثلاثة من عوامل الخطر، وهم: عوامل تتعلق بقائد السيارة مثل النوع والسن- وعوامل تتعلق بالسيارة نفسها مثل سنة الصنع وبلد المنشأ- والمنطقة الجغرافية. وقد تم دراسة العلاقة بين كل عامل من عوامل الخطر مع معدل تكرار الحوادث والتعويضات.

وتوصلت الدراسة إلى أن الأسعار المطبقة غير كافية وغير عادلة لعدم اهتمامها بالعوامل المسببة للخطر.

هدفت دراسة (الفقي، ١٩٩٣) إلى صياغة نموذج رياضي يعتمد على أسلوب حاصل الضرب Multiplicative Model للمتغيرات التي تؤثر في عدد مطالبات السيارات مثل السن ونوع السيارة، وتقدير المعالم غير المعروفة ومن ثم تقدير عدد المطالبات، وذلك باستخدام التقريب المتتاليIterative Technique.

وقد توصلت الدراسة إلى شكل نموذج رياضي أثبت الباحث صحة اشتقاقه نظرياً - ولم يتم تطبيقه على بيانات لتقدير المطالبات في تأمين السيارات -.

هدفت دراسة (الفقي، ١٩٩٣) إلى صياغة نموذج رياضي يعتمد على الأسلوب التجميعي Additive model للمتغيرات التي تؤثر في عدد مطالبات السيارات، والتي تم تحديدها في السن ونوع السيارة. وذلك لتقدير المعالم غير المعروفة والتي تفيد في تقدير عدد مطالبات السيارات عن طريق minimum Chi- Square.

وقد توصلت الدراسة إلى نموذج رياضي أثبت الباحث صحة اشتقاقه نظرياً - ولم يتم اختياره وتطبيقه على بيانات عملية لتقدير المطالبات في تأمين السيارات -.

هدفت دراسة (علي، ٢٠٠١) إلى التنبؤ بقيم المطالبات في تأمين السيارات التكميلي باستخدام نظرية الفئات الفازية Fuzzy Set Theory.

وقد توصلت الدراسة إلى أن هذه النظرية أفضل في اتخاذ القرار من الأساليب التقليدية المتبعة، فهي تقوم على المزج بين الأساليب الاحصائية والتقدير الشخصي بطريقة منطقية ومنظمة.

هدفت دراسة (عثمان، ٢٠٠٦) إلى استخدام نظم إدارة قواعد البيانات العلاقية Relational Data Base Systems، وذلك من خلال بناء نظام توزيعات احتمالية لتسعير تأمين السيارات التكميلي بالتطبيق على سيارات الأجرة الميكروباص.

وقد توصلت الدراسة إلى أنه باتخاذ زمن التعرض للحادث في الاعتبار فإن النظم المستخدمة أدت إلى اختلاف في عدد المطالبات عن العدد الناتج من جدول توزيع قيمة المطالبة. وأن استخدام نظم قواعد البيانات العلاقية قد يكون هو الحل الأمثل لعملية إعداد التوزيع التجريبي لعدد وقيم المطالبات، ومن ثم التوصل إلى سعر تأمين صافي وسعر تأمين إجمالي عادل وهو أعلى من الأسعار المطبقة في السوق.

هدفت دراسة (مهدي و آخرون، ٢٠١٠) إلى استخدام تقدير بيز في الوصول إلى معادلة المصداقية التي يمكن استخدامها في تسعير تأمين السيارات التكميلي. وقد تم استخدام ثلاث أساليب للتسعير والمقارنة بينهم وهم: تقديرات بيز لنظرية المصداقية لنموذج بواسون وجاما، ونموذج المصداقية الأول، ونموذج المصداقية الثاني.

وتوصلت الدراسة إلى أن شركات التأمين المصرية تحدد سعر التأمين التكميلي للسيارات بناءً على نوع السيارة سواء كانت ملاكي أم غير ملاكي وعلى القيمة السوقية من أجل تحديد القسط وتغفل باقي العوامل، وأن استخدام النموذج المتعدد المتغيرات في توفيق البيانات الخاصة بمطالبات السيارات يعد استخدام مناسب.

هدفت دراسة (درويش، ٢٠١٣) إلى ترشيد قرارات الاكتتاب عن طريق استخدام نظام الشبكات العصبية الفازية Neuro-Fuzzy system للمساعدة في اتخاذ قرارات اكتتاب أكثر كفاءة وفاعلية عن طريق: التحديد السليم لدرجة الخطورة الخاصة بوحدة الخطر موضوع التأمين، وتحديد الفئة التي تنتمي إليها وحدة الخطر، واتخاذ قرار بقبول أو رفض طلب التأمين، وتحديد السعر العادل في حالة اتخاذ قرار بقبول طلب التأمين.

وقد توصلت الدراسة إلى قواعد تربط بين فئات المدخلات (عوامل التقييم) وبين فئات المخرجات (درجة الخطر المتوقعة)، واقتراح قواعد للقبول والرفض بناءً على فئات الخطر المتوقعة، وتم اقتراح إطار عام لتسعير تأمين السيارات التكميلي.

هدفت دراسة (نوار، ٢٠١٦) إلى تحديد المتغيرات الأكثر تأثيراً على صافي أقساط السيارات سواء كانت السيارات الملاكي أو التجاري، ومن ثم التنبؤ بقيمة صافي الأقساط باستخدام النماذج الإحصائية، وذلك من خلال تطبيق نموذجي الانحدار التدريجي والشبكات العصبية الاصطناعية والمقارنة بينهما واختيار أفضلهما والذي يعطى أقل قيمة لخطأ التنبؤ.

وقد توصلت الدراسة إلى أن نموذج الشبكات العصبية الاصطناعية أفضل نموذج للتنبؤ بقيمة أقساط تأمين السيارات التكميلي ويعطى أقل قيمة لخطأ التنبؤ.

هدفت دراسة (المعداوي وآخرون، ٢٠١٧) إلى استخدام الشبكات العصبية في دراسة العوامل المؤثرة على المطالبات في فرع تأمين السيارات التكميلي.

وقد توصلت الدراسة إلى أن أهم العوامل المؤثرة في فرع تأمين السيارات التكميلي تتمثل في: مبلغ التأمين، وعدد حوادث المؤمن له، وعدد حوادث الورش. وتوصلت الدراسة أيضاً إلى أن نموذج الشبكات العصبية متعدد الطبقات Multiplayer Neural Network (MLP) مناسب للتنبؤ بمبالغ المطالبات في حالة وجود ثلاث متغيرات مستقلة.

هدفت دراسة (الحصري، ٢٠١٨) إلى تسعير تأمين السيارات التكميلي وفقا لعوامل خطر مختلفة، وهي: الجنس، سن المؤمن له، قوة محرك السيارة، عمر السيارة وقت الحادث، وذلك باستخدام النماذج الخطية المعممة Generalized Linear Model (GLM) والنماذج الخطية المختلطة المعممة Generalized Linear Mixed Model (GLMM). حيث تم تحديد أهم عوامل الخطر واختبار مدى تأثير هذه العوامل في عدد المطالبات وقيمة التعويضات، ومن ثم استخدام النماذج الخطية المعممة والنماذج الخطية المختلطة المعممة في توفيق نموذج لعدد المطالبات وتوفيق نموذج لقيمة التعويضات؛ وبالتالي تسعير تأمين السيارات التكميلي.

وقد توصلت الدراسة إلى أن النموذج المقترح باستخدام النماذج الخطية المختلطة المعممة الذي تم توفيقه باستخدام توزيع بواسون أفضل النماذج التي تم توفيقها لعدد المطالبات، والنموذج المقترح باستخدام النماذج الخطية المختلطة المعممة الذي تم توفيقه باستخدام توزيع جاما أفضل النماذج التي تم توفيقها لقيمة التعويضات.

هدفت دراسة (هاشم، ٢٠١٩) إلى الوصول إلى التوزيع الاحتمالي المناسب للمطالبات في تأمين السيارات التكميلي ومحاولة استخدام هذا التوزيع في التنبؤ بالمطالبات.

وقد توصلت الدراسة إلى أن التوزيع الاحتمالي لعدد المطالبات يخضع إلى توزيع بواسون، وأن التوزيع الاحتمالي لقيم المطالبات يخضع إلى توزيعات بيرسون. كما تم التوصل إلى المعادلة التي يمكن تطبيقها للتنبؤ بالمطالبات في فرع التأمين التكميلي على السيارات.

هدفت دراسة (عجوة، ٢٠١٩) إلى تحسين جودة توفيق البيانات التي تحتوي على أصفار زائدة باستخدام توزيع بواسون ذي الأصفار الزائدة وتوزيع هاردل بواسون. وقد تم اشتقاق صيغ الوسط الحسابي والتباين لتوزيع هاردل بواسون. وتم نمذجة البيانات، واختبار جودة التوفيق باستخدام اختبار كاي تربيع، وتم الاختيار بين التوزيعات المختلفة المستخدمة في الدراسة باستخدام معياري(AIC) ، (BIC).

وتوصلت الدراسة إلى أن توزيع هاردل بواسون هو الأفضل لتمثيل البيانات.

هدفت دراسة (عجوة، ٢٠٢٠) إلى تقديم منهجية لتسعير تأمين السيارات التكميلي باستخدام نموذج مقلوب جاوس ونموذج بواسون ذي التشتت الزائد، وتم استخدام نموذج انحدار بواسون ذي التشتت الزائد للتنبؤ بمتوسط عدد الحوادث باستخدام العوامل المؤثرة في متوسط عدد الحوادث وهي: جنس المؤمن له، وعمر السيارة، ونسبة خصم عدم المطالبة. وقد تم استخدام نموذج مقلوب جاوس للتنبؤ بمتوسط حجم المطالبة الواحدة بالوثيقة باستخدام العوامل المؤثرة في متوسط حجم المطالبة الواحدة وهي: نوع المؤمن له، وعمر السيارة، ومبلغ التأمين.

وقد توصلت الدراسة إلى أن متوسط عدد الحوادث ومتوسط قيمة المطالبة الخاصة بالإناث أعلى من الذكور. وهناك علاقة طردية بين كل من عمر السيارة ومتوسط عدد الحوادث للوثيقة - قيمة مبلغ التأمين ومتوسط قيمة المطالبة الواحدة،

ويوجد علاقة عكسية بين كل من: عمر السيارة ومتوسط قيمة المطالبة - نسبة خصم التحمل ونسبة خصم عدم المطالبة.

هدفت دراسة (عجوة، ٢٠٢٣) إلى حساب سعر تأمين السيارات التكميلي باستخدام بعض التوزيعات الاحتمالية المتضخمة الأصفار؛ وذلك لحل مشكلة عدم توافر بيانات عن عدد المطالبات في محفظة تأمين السيارات التكميلي خاصة عندما تكون نسبة المطالبات الصفرية كبيرة.

وقد تم استخدام ثلاث توزيعات متضخمة الأصفار، وهي: التوزيع الأسى المتضخم الأصفار، وتوزيع ويبل المتضخم الأصفار، وتوزيع باريتو المتضخم الأصفار. وتم اشتقاق صيغ المتوسط والتباين للتوزيعات محل الدراسة، وتم تقدير المعلمات المجهولة باستخدام طريقة المئويات.

وتوصلت الدراسة إلى أن توزيع ويبل المتضخم الأصفار وتوزيع باريتو المتضخم الأصفار أفضل من التوزيع الأسى المتضخم الأصفار.

هدفت دراسة (هاشم، ٢٠٢٣) إلى تقييم تسعير تأمين السيارات التكميلي باستخدام نظام (تحفيز– عقوبة) Bonus-Malus system (BMS) للحصول على القسط الذي يختلف من شخص لآخر نتيجة اختلاف الخصائص الفردية للمؤمن لهم؛ وذلك من خلال أخذ العوامل المؤثرة في درجة الخطر في الاعتبار عند تطبيق النموذج المقترح، وتحليل محفظة بيانات تأمين السيارات التكميلي على السيارات الخاصة باستخدام النماذج الخطية المعممة (GLM)، وتقدير تكرار المطالبات السنوية.

وقد توصلت الدراسة إلى أن نموذج (BMS) المقترح يستند على عدد الحوادث الماضية ووطأتها في عملية تقدير تسعير التأمين، وأن استخدام هذا النموذج في تسعير تأمين السيارات التكميلي يساعد على وضع أساس واضح ومحدد لتحديد الأخطار الجيدة والأخطار الرديئة مما يؤدي إلى المساعدة في اتخاذ قرارات الاكتتاب المناسبة، ويضمن للمؤمن له دفع القسط الذي يتناسب مع درجة تعرضه للخطر.

ب. المجموعة الثانية: دراسات باللغة الإنجليزية:

هدفت دراسة (Fahrmeir, 2007) إلى تقييم تأثير المتغير الجغرافي على تكرار وقيم مطالبات تأمين السيارات باستخدام نماذج إحصائية مناسبة، من خلال نمذجة وتقدير تأثير المتغير الجغرافي في ذات الوقت مع عوامل الخطر الخطية وغير الخطية مثل عمر المؤمن له وعمر السيارة.

وقد توصلت الدراسة إلى أن النماذج الجيوتجميعية Geoadditive Models هي الأفضل، حيث يتم نمذجة المتغير الجغرافي بناءً على Gaussian random fields، المستخدمة في الإحصاء الجغرافي.

هدفت دراسة (Mazviona etal, 2013) إلى الوصول إلى التوزيع الاحتمالي المناسب للمطالبات في تأمين السيارات التكميلي من بين أربع توزيعات وهي جاما وباريتو والأسي واللوغاريتمي الطبيعي. ومحاولة استخدام هذا التوزيع في التنبؤ بالمطالبات.

وقد توصلت الدراسة إلى أنه من بين التوزيعات الأربعة، فإن البيانات تخضع بشكل كبير إلى توزيعي باريتو واللوغاريتمي الطبيعي. حيث أن التوزيع اللوغاريتمي الطبيعي أنتج أفضل ملاءمة للمطالبات المنخفضة بينما أنتج باريتو أفضل ملاءمة للمطالبات الكبيرة.

هدفت دراسة (Kafková, etal, 2014) إلى التنبؤ بعدد المطالبات السنوية في تأمين السيارات بالاعتماد على خمسة من عوامل الخطر وهم: نوع السيارة، وعمر السيارة، والمنطقة الجغرافية، وجنس وعمر المؤمن له، وذلك عن طريق استخدام النماذج الخطية المعممة GLMs، لشرح العلاقة بين تكرار المطالبات السنوية وعوامل الخطر المحددة.

وقد توصلت الدراسة إلى وجود تأثير لثلاثة عوامل فقط من الخمسة وهم: سن المؤمن له، وعمر السيارة، والمنطقة الجغرافية، بينما لا يوجد أي تأثير ذو دلالة إحصائية لكلاً من جنس المؤمن له أو نوع السيارة.

هدفت دراسة (David, 2015) إلى حساب القسط الصافي لتأمين السيارات التكميلي باستخدام النماذج الخطية المعممة (GLMs)، وذلك مع الأخذ في الاعتبار الخصائص الفردية للمؤمن لهم من سن ومهنة، وذلك للحصول على قسط تأمين عادل يختلف من شخص لآخر نتيجة اختلاف الخصائص الفردية. وقد تم استخدام نموذج انحدار بواسون للتنبؤ بتكرار الحوادث، ونموذج جاما للتنبؤ بقيمة المطالبة.

وتوصلت الدراسة إلى تحديد قسط تأمين عادل باستخدام النماذج المعممة يأخذ في الاعتبار عدد من عوامل الخطر، مثل: عمر ومهنة المؤمن عليه، والغرض من استخدام السيارة، ومدة وثيقة التأمين.

هدفت دراسة (David, etal, 2015) إلى دراسة تأثير عوامل الخطر والمتمثلة في: الفئة العمرية للمؤمن له ومهنته، ونوع السيارة، ونظام تحديد المواقع (GPS)، ومدة وثيقة التأمين على تكرار المطالبات بتطبيق كلاً من انحدار بواسون ونماذج ذات الحدين السالبة على محفظة تأمين سيارات فرنسية.

وتوصلت الدراسة إلى أن أهم عاملين مؤثرين في تكرار المطالبات هما مهنة المؤمن له ونظام تحديد المواقع (GPS)، بينما العمر لم يكن له تأثير كبير في مجتمع الدراسة.

هدفت دراسة (Masese, 2020) إلى استخدام النماذج الخطية المعممة في تقييم تأثير عوامل الخطر، والمتمثلة في: مسافة وسرعة ووقت القيادة، والتي تعتبر عوامل أساسية ويجب أخذ تأثيرها في الحسبان عند تسعير التأمين وتحديد القسط على كلاً من تكرار المطالبة وشدتها. وقد تم استخدام نماذج جاما وبواسون المعممة في هذه الدراسة، حيث تم استخدام نموذج جاما للتنبؤ بشدة المطالبة ونموذج بواسون لتكرار المطالبة.

وقد توصلت الدراسة إلى أن نموذج تسعير تأمين السيارات يحدد أقساط التأمين بناءً على عوامل القيادة في الوقت الفعلي، حيث تم تحليل محفظة التأمين والتوصل إلى أن متغيرات السرعة والمسافة ذات دلالة إحصائية معنوية بينما متغير الزمن لم يكن معنوياً في كلا النموذجين، وأن الشدة والتكرار يزدادان مع المسافة والسرعة، وعليه يزداد القسط الصافي.

ج. المجموعة الثالثة: الدراسات السابقة المتعلقة بالنماذج التجميعية المعممة Generalized Additive Models في تسعير التأمين:

هدفت دراسة (Kaivanipour, 2015) إلى دراسة أسلوبين مختلفتين وهما أسلوب التصحيح المتقاطع (cross validation) وأسلوب المنحنى L (L-curve) للوصول إلى معامل التجانس الأمثل لنماذج Generalized Additive Models (GAM)، حتى يتم التسعير بدقة أكبر.

وقد توصلت الدراسة إلى أن أداء أسلوب التصحيح المتقاطع كان جيد في معظم المواقف، ولكنه واجه بعض المشاكل في كثرة وصعوبة العمليات الحسابية، وهذا عكس أسلوب المنحنى L فهو أسرع وأسهل، ولكن يعاني من التجانس الشديد، وبالتالي تم رفض طريقة المنحنى L من كونها طريقة مناسبة لتقدير معامل التجانس الأمثل.

هدفت دراسة (Tingting, 2018) إلى التوصل إلى نموذج إحصائي مناسب لتقدير قسط التأمين الصافي، وذلك من خلال حل مشكلة التقدير الدقيق للقيمة المتوقعة والتباين في إجمالي المطالبات لكل حامل وثيقة. وقد تم توسيع إطار النماذج الخطية المعممة (GLMs) للمطالبات الإجمالية ليشمل هيكل النماذج التجميعية المعممة (GAMs)، حيث أن النماذج التجميعية المعممة تسمح بدخول معلمات غير خطية وبها مرونة تكرار المطالبات، وقد تم التطبيق على بيانات التأمين على السيارات.

وتوصلت الدراسة إلى أن النماذج التجميعية المعممة تعطي نتائج أفضل من النماذج الخطية المعممة.

هدفت دراسة (Jo˜ao, 2019) إلى اقتراح نماذج إحصائية دقيقة لتكرار المطالبات باستخدام المنهجية الحديثة في تطبيقات التأمين على السيارات التكميلي، وذلك عن طريق تحديد بيانات عوامل الخطر وتحليل كيفية تأثر تكرار المطالبات بعوامل الخطر، ومن ثم التوصل إلى أقساط تأمين عادلة ومناسبة لحاملي الوثائق، وذلك بالاعتماد على النماذج التجميعية المعممة (GAMs) باستخدام تقدير بيز من خلال Markov chain Monte Carlo (MCMC).

وتوصلت الدراسة إلى أن دمج النماذج التي تم اختيارها كان الأنسب للوصول إلى أقساط عادلة، حيث كان الناتج نموذج ملائم للبيانات التي تحتوي على أصفار زائدة (تضخم صفري)، ويعمل على تقليل التشتت. كما أنه ملائم للمتغيرات غير الخطية أيضاً، وكان جنس وعمر المؤمن له أكثر عاملين من عوامل الخطر تأثيراً على قسط التأمين.

هدفت دراسة (Díaz etal, 2023) إلى المقارنة بين أسلوبي النماذج الخطية المعممة (GLMs) والنماذج التجميعية المعممة (GAMs) وذلك باستخدام عينة مكونة من عدد كبير من وثائق تأمين السيارات لتقييم الاختلافات بين النموذجين من حيث جودة التنبؤات، وتعقيد الاستخدام، ووقت التنفيذ.

وقد توصلت الدراسة إلى أن النماذج التجميعية المعممة GAMs هي البديل القوي للنماذج الخطية المعممة GLMs، وخاصة عند استخدام البيانات الضخمة.

هدفت دراسة (Xie, etal, 2023) إلى بناء نماذج للتنبؤ بتكرار المطالبات وشدتها، وذلك بالاعتماد على مجموعة من عوامل الخطر على أساس الاستخدام الاصطناعي Usage Based Insurance (UBI)، حيث تقوم بسد الفجوة بين التسعير على المستوى الفردي وتنظيم الأسعار باستخدام قاعدة بيانات UBI. ويقوم النموذج بالاعتماد على عدة عوامل خطر، وهي: دخل المؤمن له، وعمره، وموقعه الجغرافي، وعمر السيارة، ودرجات الائتمان، والمسافات المقطوعة سنوياً، وسنوات عدم وجود مطالبات.

وقد توصلت الدراسة إلى أن النماذج التجميعية المعممة GAM كشفت وجود علاقة غير خطية مهمة بين عوامل الخطر الرئيسية واحتمالية المطالبة أو مبلغها. وأن أخطار المنطقة الجغرافية تساهم بشكل كبير في التسعير.

ونخلص من الدراسات السابقة - رغم اختلاف الهدف لكل منها – إلى مجموعة من النقاط، وهي:

- عدم عدالة الأسعار في تأمين السيارات التكميلي في جمهورية مصر العربية نظراً لجمودها فترة طويلة وعدم إتباع الأسس العلمية والرياضية في تحديد عدد وقيم المطالبات وحساب الأسعار أو مقدار الزيادة فيها.
- أن هناك مجموعة من العوامل التي تؤثر في درجة الخطورة بالنسبة لحوادث السيارات، ومنها: سن المؤمن له – نوع المؤمن له – مهنة المؤمن له – عمر السيارة – قيمة السيارة – استخدام السيارة – نوع السيارة – المنطقة الجغرافية.
- اعتمدت الدراسات على بعض النماذج في التنبؤ بمطالبات تأمين السيارات، ولكن لم تتطرق أي من الدراسات السابقة إلى استخدام النماذج التجميعية المعممة في تقدير المطالبات الخاصة بتأمين السيارات التكميلي في جمهورية مصر العربية.
- لم تتطرق أي من الدراسات السابقة إلى دمج أسلوب التسوية LASSO Regularization في تقدير المطالبات الخاصة بتأمين السيارات التكميلي في جمهورية مصر العربية، حيث تسوية نموذج التنبؤ من النتوءات وتقليل تعقد النموذج تعمل على تقليل نسب الخطأ وزيادة قدرة النموذج على التنبؤ.

وبناءً على ما سبق يهتم البحث الحالي باستخدام النماذج التجميعية المعممة في تقدير المطالبات وتسعير تأمين السيارات التكميلي، وذلك عن طريق دراسة العوامل المؤثرة في تأمين السيارات التكميلي ثم التنبؤ بعدد المطالبات التي يمكن أن تتحملها مستقبلا في هذا الفرع حتى لا تتعرض شركات التأمين إلى مشاكل مالية.

وهذا ما دفع الباحثون إلى القيام بهذه الدراسة، ومحاولة الاستفادة بما تم تطبيقه في الدراسات السابقة.

هدف البحث:

يتمثل هدف البحث الرئيسي في تقدير عدد المطالبات في فرع تأمين السيارات التكميلي في جمهورية مصر العربية، وذلك من خلال الأهداف الفرعية التالية:

- تحليل بيانات محفظة تأمين السيارات التكميلي باستخدام النماذج التجميعية المعممة Generalized Additive Models (GAMs).
- دراسة وتحديد عوامل الخطر المؤثرة على تسعير السيارات.
- إدخال عوامل خطر لم تؤخذ في الحسبان من قبل عند تسعير تأمين السيارات التكميلي في جمهورية مصر العربية، وهي المسافات المقطوعة والتقييم الائتماني I-SCORES.
- تطبيق طريقة لاسو LASSO Regularization من طرق الـ Machine Learning لزباده قدرة النموذج على التنبؤ.

حدود البحث:

سوف تقتصر الدراسة على:

- فرع السيارات التكميلي لإحدى شركات التأمين العاملة في سوق التأمين المصري.
- بيانات التقييم الائتماني من الشركة المصرية للاستعلام الائتماني
- الفترة الزمنية من عام ٢٠١٣/٢٠١٢ إلى عام ٢٠٢٢/٢٠٢١.

أهمية البحث:

وتكمن أهمية البحث في:

- تحديد عوامل الخطر التي يتوقف عليها تسعير تأمين السيارات التكميلي.
- مساعدة شركات التأمين التجاري في تقدير عدد المطالبات في ظل ازدياد عوامل الخطر.
- الاعتماد على دمج الأساليب الرياضية والاحصائية في تسعير تأمين السيارات التكميلي.

محتويات البحث:

أولاً: النماذج التجميعية المعممة
ثانياً: التطبيق العملي للنموذج المقترح
ثالثاً: النتائج والتوصيات

أولاً: النماذج التجميعية المعممة Generalized Additive Models):

تعد النماذج التجميعية المعممة GAMs أحد مداخل الانحدار اللامعلمي، وهي امتداد للنماذج الخطية المعممة Generalized Linear Model (GLMs). وتجمع بين خصائص النماذج الخطية المعممة GLMs والنماذج التجميعية (AMs) Additive Models. (Guisan, etal، ٢٠٠٢)(Wood, 2017)

وقد أخذت النماذج التجميعية المعممة من النماذج التجميعية (AMs) ميزة الحفاظ على الشكل الجمعي المتعدد فقط، حيث يعاب على النموذج التجميعي أنه قد يقدم نتائج زائفة عند التطبيق على عينة تتكون من قيم غير كافية، حيث يمكن أن ينتج قيماً سالبة في حساب الأقساط أو المطالبات. (فرحان، وآخرون، ٢٠٢٠)

بينما يعاب على النماذج الخطية المعممة GLMs أنها تتعامل مع جميع القيم التي تنتمي لفئة معينة بأنها متطابقة. وهذا ما تم مراعاته في النماذج التجميعية المعممة، حيث تم تجنب جزء التصنيف من النماذج الخطية المعممة، واستبدال الحدود البسيطة في المعادلة الخطية GLMs بالدوال؛ وهي دوال لامعلمية للمتغيرات، أي أن النماذج التجميعية المعممة تقدر دوال لامعلمية غير محددة لكل متغير مستقل بدلاً من المعامل، وذلك للتوصل لأفضل تنبؤ لقيم المتغير التابع. (Irwin, 2005)

وتعد النماذج التجميعية المعممة أكثر تقييداً من النماذج التجميعية والنماذج الخطية المعممة لما تقدمه من تجانس مكونات النموذج، حيث تكمن قوة النماذج التجميعية المعممة في: (Guisan, etal، ٢٠٠٢) (Kaivanipour، ٢٠١٥) (Baayen, 2020)

- أنها أداة فعالة لتحليل البيانات وتوفير نتائج قابلة للتفسير، وهذا ما يميزها عن كثير من النماذج مثل نماذج الشبكات العصبية التابعة إلى الـ Machine Learning والتي يعاب عليها صعوبة التفسير.

- اللاخطية (Non-Linearity): وهذه النماذج تحدد طبيعة العلاقة بين المتغيرات ويمكن نمذجة العلاقات المعقدة وغير الخطية، بدلاً من افتراض شكل لتمثيل العلاقة مثل GLMs والحاجة إلى تدخلات أو تحويل (Transformation).
- المرونة: حيث تمتلك القدرة على التعامل مع العلاقات غير الخطية بشكل كبير بين المتغير التابع وعدد من المتغيرات المستقلة، سواء كانت متغيرات كمية أو وصفية.
- التسوية (Regularization): حيث أن النماذج التجميعية المعممة تندمج بها أساليب تسوية لتخليص النموذج من النتوءات(Shrinkage) ، وذلك من خلال اختيار عوامل الخطر وتقليل قدرة النموذج على استخدام جميع خصائص المتغيرات وعلاقتها بمتغير الاستجابة (response Variable) مما يسبب أخطاء في تقدير النموذج (overfitting)، والتحكم فيه.

وبناءً على ذلك تحظى النماذج التجميعية المعممة باهتمام متزايد من الباحثين والاكتواريين لاستخدامها في تسعير التأمين، والفرض الذي تقوم عليه عملية التقدير لتحقيق المرونة وقابلية التفسير وتحديد المتغيرات وعوامل الخطر هو أن دوال التقدير تكون تجميعية ومتغيراتها ممهدة.

وفي هذه الدراسة سوف يقوم الباحثون باقتراح دمج أساليب التسوية LASSO Regularization وهو من أساليب الـ Machine Learning إلى النماذج التجميعية المعممة، ويتم عرض مكونات النموذج المقترح من LASSO Regularization، والنماذج التجميعية المعممة، بحيث يكون الإطار النظري للنماذج التجميعية المعممة (GAM) كما يلي:

يتم تحديد العلاقة بين المتغيرات مع متغير الاستجابة، وقياس مدى التجانس بين مكونات النموذج من خلال التوصيل بين نقطتين بدالة خطية من الدرجة الأولى. كما يعمل النموذج أيضا بشكل مميز في العلاقات اللامعلمية Non-Parametricوذلك عندما تكون العلاقة بين المتغيرات ومتغير الاستجابة غير خطية، ولكنها ممهدة، حيث تعتمد النماذج التجميعية المجمعة في التنبؤ بالمتغيرات غير الخطية على الدوال الممهدة smooth Functions، ويوجد دالة للربط بين متوسط متغير الاستجابة والدوال الممهدة للمتغير التابع وهذا ما يعطيه الدقة العالية في التنبؤ أكثر من غيره من النماذج. ويعتمد النموذج التجميعي المعمم على كلاً من: (Goldburd, etal, 2016)

$$y_i \sim Exponential\ (M_i, \Phi) \qquad (1)$$

$$F\ [\theta i)\ ] = exp[y_i\ (\theta)_i - b_i\ (\theta)_i] \qquad (2)$$

حيث أن:

Y: توزيع متغير الاستجابة، حيث أن المتغير التابع لابد أن يتبع أحد التوزيعات الأسية، ويمثل هذا المركب العنصر العشوائي(Random Component) .

θ- : المعلمة الطبيعية للتوزيع.

ويمر النموذج بالمراحل التالية:

١. تحديد الدوال الممهدة Smooth Functions

وتكون الصيغة الرياضية للنموذج التجميعي المعمم لعدد **n من** المتغيرات هو: (Xie, etal، ٢٠٢٣)

$$g(\mu i) = ß_0 + f_1(xi_1) + f_2(xi_2) + \ldots\ldots\ldots f_n(xi_n) \qquad (3)$$

حيث أن:

$f_n(xi_n)$: المركب التجميعي للنموذج لعددn من المتغيرات.

f_i: دالة خطية للمتغيرات x_i .

ويتم تقليل دالة الخسارة المربعة التالية إلى الحد الأدنى وتكون:

$$L(ß_{0,\ldots\ldots,} ß_n) = (Y- ß_0 -\sum_{i=1}^{n} f_n\, x_n)^2 \qquad (4)$$

ويتم تمديد دالة الخسارة للتحكم في تمهيد دقة منحني التقدير، كما يتم الاعتماد على استخدام دوال تحيز للتجانس بين الشرائح، وتكون هذه الدوال عبارة عن $\{b_{ni}(x_n)\}_{i=1,2,\ldots\ldots Ln}$ لكل عدد **n** من المتغيرات لنمذجة المركب التجميعي $\mathbf{f_n\ x_n}$ في هذا النموذج.

٢. <u>تحديد الانحراف بتسوية النموذج Deviance by Regularization</u>

يتم التحكم في تحيز دوال الشرائح باستخدام معامل التسوية penalty term لدالة الخسارة لتصبح:

$$L(ß_{0,\ldots\ldots,} ß_n) = (Y- ß_0 -\sum_{i=1}^{n} f_n\, x_n)^2 + \sum_{i=1}^{n}\sum_{l=1}^{Ln} \lambda_n\, (\ ß_{n,I+1,\ldots\ldots}\ ß_{n1}\ )^2 \qquad (5)$$

حيث أن:

λ_n : المعاملات الممهدة التي تعمل علي تسوية مربع الفرق بين معاملات دوال التحيز لكل متغير.

يتم اختيار المعاملات الممهدة λ_n بإيجاد الدالة التي تعمل على تقليل انحراف جودة التوفيق للبيانات الفعلية وتقليل أخطاء تقديرات النموذج، كما يلي:

$$D(Y,exp\{f(x)\}+\lambda_n R(f(x) \qquad (6)$$

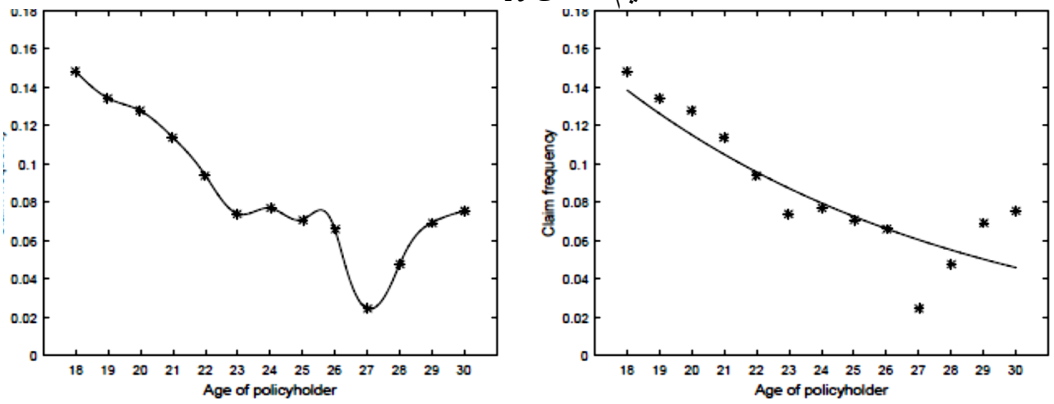
حيث أن:

المقدار الأول (D(Y,exp{f(x)}) يعبر عن مقدار الانحراف، بينما يعبر المقدار الثاني القيمة المثلى $\boldsymbol{\lambda}$ ، حيث أن اختيارها له دور هام في تقليل دالة الخسارة والتحكم في دوال الشرائح عن طريق تعظيم دالة التسوية، وذلك لتقليل نسبة الخطأ عند التوقع.

ويوضح الشكل التالي رقم (١) استخدام قيم كبيرة أو صغيرة لمعامل $\boldsymbol{\lambda}$:

شكل رقم (١)

قيم معامل $\boldsymbol{\lambda}$



المصدر: (Kaivanipour، ٢٠١٥)

ويتضح من الشكل السابق أن تقليل قيم معامل $\boldsymbol{\lambda}$ يؤدي إلي زيادة البيانات الفعلية المستخدمة، وبالتالي تقليل معامل دالة التسوية وزيادة الانحراف المعياري للبيانات في النموذج المقترح كما في الشكل السابق جهة اليسار، والعكس صحيح كما في الشكل

جهة اليمين. لذا فإن اختيار قيم مثلى لمعامل λ ووضعها في معادلة التسوية (معادلة رقم ٦)، يعمل على تعظيم المعاملات الممهدة للنموذج التجميعي المعمم.

٣. تحديد المعاملات الممهدة λ_n باستخدام معيار التصحيح المتقاطع المعمم Generalized Cross Validation Criteria:

يعتبر معيار التصحيح المتقاطع Cross Validation وسيلة لقياس أداء النموذج، ويتم عن طريقه إجراء الاختبارات للبيانات ثم تشغيلها، حيث يتم تقسيم البيانات إلى عدد من القطاعات العشوائية k-folds وإجراء هذا الاختبار لعدد k من المرات. ويتم تقدير معامل الخطأ عند كل λ ينتج عنه أفضل تنبؤ للنموذج، حيث يعمل معدل الخطأ C(λ) على تقليل جميع الشرائح $S_1{}^{\lambda}$,…., $S_m{}^{\lambda}$ بالنموذج التجميعي المعمم .

وتعمل النماذج التجميعية المعممة على التوازن بين درجة التوافق والقيمة المثلى λ، وذلك لتعظيم معدل التسوية، وتقليلoverfitting، كما أن إيجاد القيم المثلى للدوال الممهدة تعطي للنموذج التجميعي المعمم قدرة عالية على الدقة وتفسير النتائج، كما يتضح من المعادلة التالية:

$$logl(\beta, \beta_j) - \lambda_j . S_j.\beta_j \qquad (7)$$

حيث أن:

λ_j .: المعامل الممهد الذي يوازن بين درجة التوفيق وتمهيد الدالة.

β_j: *معامل يتم تقديره* في الدالة.

٤. تحديد المتغيرات (Feature Selection) أو عوامل الخطر لزيادة دقة تنبؤ النموذج التجميعي عند تسعير السيارات:

تزداد نسبة توقع الخطأ في النموذج التجميعي المعمم عند زيادة العينة والمتغيرات، لذلك يتم استخدام أساليب التسوية Regularization، حيث تسوية نموذج التنبؤ المقترح من التنوءات (Shrinkage) وتقليل تعقد النموذج تعمل على تقليل خطأ overfitting وزيادة قدرة النموذج على التنبؤ، وتتمثل الطرق الشائعة للتسوية في: (Fujita, 2020) Ridge (L1) , Lasso (L2)

$$Ridge\ (L2\ regularization)\ R(\beta; \lambda) = \lambda \sum_{I=1}^{P} |\beta j|^2 \qquad (8)$$

$$Lasso\ (L1\ regularization)\ R(\beta; \lambda) = \lambda p \sum_{I=1}^{P} |\beta j|^2 \qquad (9)$$

حيث أن:

λ ، β : معاملات المتغيرات في النموذج الخطي .

٥. اختيار النموذج: MODEL SELCTION

يتم الاختيار بين النماذج عن طريق اختيار مصفوفة التنبؤ المناسبة للبيانات التي تؤدي إلى أفضل درجة تنبؤ للنموذج التجميعي المعمم.

وقد تم اختيار التوزيع الأسي لأنه يناسب بيانات عدد مطالبات السيارات، ويتم الاختيار عن طريق معيارAkaike information criterion (AIC) وانحراف بواسون Poisson Deviance، حيث أن:

ا. **Akaike information criterion (AIC):** تعتمد دالة المعاملات على دالة التعظيم المستوية والتي تستخدم λ لعدد من المعلمات، ويمكن التعبير عنها رياضيا كما يلي: (Jo˜ao، ٢٠١٩)

$$AIC(\lambda) = -2log\{L(\lambda)\} + 2 \times df_{fit}(\lambda) \qquad (10)$$

حيث أن:

$df_{f\,it}$: هي درجات حرية المعاملات عند تقدير النموذج.

II. انحراف بواسون Poisson Deviance:

يتم الاختيار بين النماذج التجميعية المعممة من خلال تحليل فروق الانحراف، ويمكن تفسيرها على أنها المجموع المرجح لمتوسط المسافات المقدرة للبيانات μ^{Λ} ، ويمكن تعريفها من خلال المعادلة التالية:

$$D(Y,exp\{f(x)\}= \ ٢\sum_1^n w_i(Y_i \ _{log_i}- Y_i\ f(x_i)-Y_i + e^{f(x_i)}) \qquad \textbf{(11)}$$

حيث أن:

Y : متغير الاستجابة للمطالبات

x_i : *المتغيرات وعوامل الخطر لتقدير عدد مطالبات تأمين السيارات التكميلي*

ثانياً: التطبيق العملي للنماذج التجميعية المعممة على بيانات تأمين السيارات التكميلي لإحدى شركات التأمين العاملة في سوق التأمين المصري:

يعتمد الباحثون في هذه الدراسة على بيانات تأمين السيارات التكميلي لإحدى شركات التأمين العاملة في سوق التأمين المصري، وذلك لنمذجة عدد المطالبات، والتي يرمز لها بالرمز Y، وتم تقسيم الوثائق إلى مجموعات وفقا لعوامل الخطر، حيث تم إضافة عوامل خطر لم تؤخذ في الحسبان من قبل عند تسعير السيارات التكميلي في جمهورية مصر العربية، ومنها المسافات المقطوعة بالأميال للسيارات، والأداء الائتماني I-SCORES والذي يتراوح من ٤٠٠ إلي ٨٥٠ وتم الحصول على بيانات التقييم الائتماني من الشركة المصرية للاستعلام الائتماني.

ويعرض الجدول رقم (٢) التالي تقسيم المتغيرات وعوامل الخطر إلى بيانات وصفية أو كمية، وعددهم ١٣ متغير، ويتم استخدام الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج 'mgcv'. *R package,* Version1.9-0 **:** (Wood, 2023)

جدول رقم (٢)
متغيرات عدد مطالبات السيارات التكميلي

المتغيرات	الوصف	النوع
ID	*رقم الوثيقة*	*Integer*
Exposure	*عدد الوحدات المعرضة للخطر*	*Integer*
ClaimNb	*عدد المطالبات أثناء فترة سريان الوثيقة*	*Integer*
ClaimAMT	*قيمة المطالبات بالجنيه المصري*	*Integer*
Vehvalue	*قيمة السيارة بالجنيه المصري*	*Integer*
VehAge	*عمر السيارة بالسنوات*	*Integer*
VehBrand	*موديل السيارة*	*Facto*
Insured.age	*عمر السائق*	*Integer*
Insured.Sex	*جنس السائق*	*Factor*
Marital Status	*الحالة الاجتماعية*	*Factor*
AREA	*المنطقة وتنقسم إلى مركز: حضر F وريف A*	*Factor*
Annual Miles Drives	*المسافة المقطوعة*	*Integer*

Integer	*CREDIT* الأداء الائتماني *SCORE*	Credit I-Score

المصدر: الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R package,* Version1.9-0 .'mgcv'

ويتم استخدام دالة للربط بين المتغيرات مكونة من التوزيع الأسي البواسوني والدالة اللوغاريتمية Log Link function؛ حيث توضح أثر استخدام الدوال الممهدة للمتغيرات الكمية المتصلة في تقدير عدد مطالبات تأمين السيارات التكميلي بعد دمج عوامل الخطر الوصفية والكمية، وذلك باستخدام النموذج التجميعي المعمم البواسوني، وقد تم صياغة النموذج كما يلي:

$$Y^{\sim} = \beta 0 + \beta 1_{Insured.sex} + \beta 2_{Marital} + \beta 3_{VehBrand} + \beta 5_{Car.use} + \beta 5_{AREA} + s(VehAge\ i) + s(CreditScorei) + s(Insured.age) + s(Annual.miles.drive) + \epsilon \quad (12)$$

حيث أن:

Y: تعبر عن عدد المطالبات.

S(.): الدوال الممهدة smooth functions

وبعد تطبيق النموذج المقترح تم التوصل إلى نتائج الإحصاءات الوصفية، حيث يعرض الجدول رقم (٣) التالي الإحصاءات الوصفية الناتجة عن استخدام النموذج التجميعي المعمم وأثر الدوال الممهدة وعلاقتها غير الخطية بعدد المطالبات:

جدول رقم (٣) الإحصاء الوصفية لعدد مطالبات السيارات التكميلي

	Min.Value	Mean	Max.Valu	Median
ClaimNb	0	8	5	0

Parametric coefficients:				
	Estimate Std.	Error	z value	Pr(>|z|)
(Intercept)	- 0.143008	0.086218	-1.659	0.09718
Insured.sexMale	0.051118	0.022342	2.288	0.02214 *
MaritalSingle	-0.048288	0.024942	-1.936	0.05287

المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الجدول السابق أن عدد المطالبات تعاني من مشكلة تضخم الأصفار لذا يجب أن يؤخذ في الاعتبار التوزيع الاحتمالي المناسب لمشكلة تضخم الأصفار أو عدم وجود بيانات لمحفظة السيارات التكميلي عند التسعير.

كما تم التوصل إلى تأثير عوامل الخطر على تقدير عدد المطالبات، ويتم عرض مخرجات النموذج المطبق في الجدول رقم (٤) التالي:

جدول (٤) مخرجات النموذج التجميعي المعمم

Car.use Commute	-0.008531	0.074616	-0.114	0.0٣214 *
Car.usePrivate	0.005982	0.076192	0.079	0.93742
RegionUrban	0.050313	0.027561	1.826	0.06792
VehBrandPickup	-0.099772	0.044370	-2.249	0.02453*
VehBrandSedan	-0.315972	0.042912	-7.363	1.8e-13 ***
VehBrandSUV	-0.122646	0.041400	-2.962	0.00305 **

المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الجدول السابق أن جنس السائق، وسبب استخدام السيارة، وموديل السيارة من عوامل الخطر المعنوية ذات الدلالة الإحصائية على تقدير عدد المطالبات في السيارات التكميلي.

ويعتبر المكون التجميعي Additive Term من أهم خصائص النموذج التجميعي المعمم؛ والذي يوضح تأثير معنوية الدوال الممهدة Smooth Functions على المتغيرات ذات العلاقات غير الخطية، لذلك يتم في الجدول رقم (٥) التالي عرض تحليل ANOVAللدلالة المعنوية للدوال الممهدة:

جدول رقم (٥) تحليل ANOVA للدوال الممهدة للنموذج التجميعي المعمم

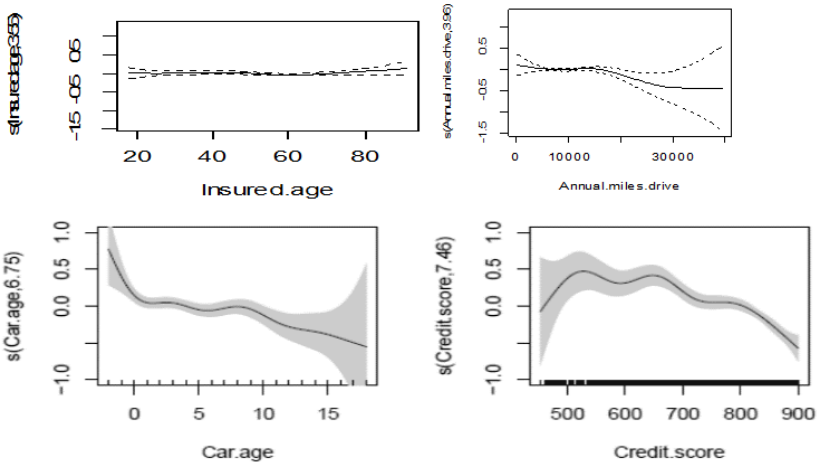
ANOVA OF Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(*VehAge*)	8.753	8.971	28.11	0.00014 ***
s(*DrivAge*)	3.554	4.444	5.678	0.002861 **
s(Credit i-Score)	2.274	2.874	9.5	0.0001***
s(*Distance Bands*)	3.959	4.924	8.677	0.044*

المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الجدول السابق أن درجات الحرية لكل من عمر السيارة، وعمر قائد السيارة، والمسافات المقطوعة بالأميال، والأداء الائتماني أكبر من ١، وهذا يدل على وجود علاقة معنوية ذات دلالة إحصائية بين هذه المتغيرات وعدد المطالبات، كما يؤكد على دقة وقدرة النموذج على التنبؤ.

ويعرض الشكل التالي رقم (٢) مدى جودة ملائمة النموذج التجميعي المعمم للدوال الممهدة:

شكل (٢) جودة ملائمة النموذج المقترح للدوال الممهدة



المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الشكل السابق أن النموذج التجميعي المعمم أثبت جودة ملائمته للدوال الممهدة على المتغيرات غير الخطية (عمر السيارة، وعمر المؤمن، والمسافات المقطوعة والأداء الائتماني).

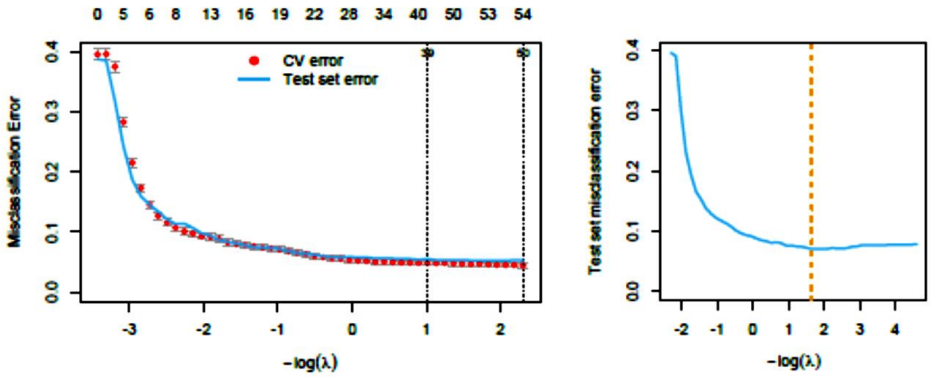
ومن ثم تم تطبيق المعادلتين (١١)، (١٢) لاختبار قدرة النموذج على التنبؤ بعدد مطالبات السيارات التكميلي، وقد تم الاختبار باستخدام معياري التنبؤ AIC, Poisson Deviance، وقد نتج عن ذلك أن قيمة معيار AIC كانت ٢٦٢٨٩٨٧، وقيمة معيار Poisson Deviance كانت ٠,٣١١١٩٢٠، مما يدل على أن النموذج التجميعي المعمم له قدرة عالية في التعامل مع العلاقات غير الخطية الممهدة.

ثم تم تحديد المعاملات الممهدة λ_n باستخدام معيار التصحيح المتقاطع المعمم Generalized Cross Validation Criteria

ويوضح الشكل التالي رقم (٣) إجراء معيار التصحيح المتقاطع cross Validation، وتحديد أقل قيمة خطأ للمعلمةλ^{1se} :

شكل (٣)

استخدام معيار التصحيح المتقاطعcross Validation



المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الشكل السابق أنه تم تقدير معامل الخطأ عند كل λ ينتج عنه أفضل تنبؤ للنموذج، حيث يعمل معدل الخطأ $C(\lambda)$ على تقليل جميع الشرائح، ,....

S_1^{λ} S_m^{λ} بالنموذج التجميعي المعمم، حيث أن قيمة λ^{1se} تعطي أقل قيمة خطأ وهي ٥,٥ %.

وقد تم تقسيم البيانات إلى عدد من القطاعات العشوائية ١٠-folds وإجراء هذا الاختبار لعدد ١٠ مرات. وتم تقدير معامل الخطأ عند كل λ وينتج أفضل تنبؤ للنموذج (tunning Parameter) من خلال أقل مربع خطأ Mean Square error وتسمى $\lambda\hat{}^{min}$. وتقدر قيمة $\lambda\hat{}^{min}$ بما يعادل ٠,١٤١٥٤١٦.

وبعد تحديد قيمة $\lambda\hat{}^{min}$ قام الباحثون في هذه الدراسة باستخدام أسلوب Lasso من أساليب التسوية Regularization. وبالرجوع للمعادلة رقم ٩، واستخدام حزمة glmnet ببرنامج R، واستخدام المعاملات hyper parameter ($\lambda\hat{}^{min}$) على النموذج التجميعي المعمم، وكانت نتائج تحليل ANOVA للدلالة المعنوية كما موضحة في الجدول رقم (٦) التالي:

جدول رقم (٦)

تحليل ANOVA للدوال الممهدة للنموذج التجميعي المعمم بعد تطبيق معيار التسوية *L1 regularization*

ANOVA OF Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(*VehAge*)	5.753	8.971	28.11	0.004 **
s(*DrivAge*)	2.554	4.444	5.678	0.002861 **
s(Credit i-Score)	1.274	2.874	9.5	0.0002***
s(*Distance Bands*)	2.959	4.924	8.677	0.0005***

المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

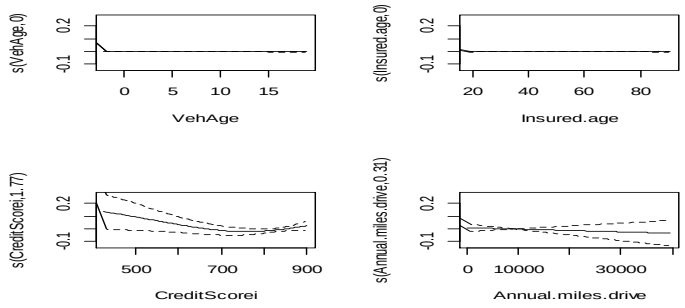
ويتضح من الجدول السابق أن درجات الحرية لكل من عمر السيارة، وعمر قائد السيارة، والمسافات المقطوعة بالأميال، والأداء الائتماني بعد تطبيق معيار التسوية Lasso قد تحسن عن قبله، وهذا يدل على دقة أكبر للنموذج بعد التسوية.

وقد أصبح معيار AIC للنموذج التجميعي المعمم٢٣١٥٦بعد استخدام معيار التسوية *L1 regularization* لتقليل معدل النتوءات وoverfitting عند تقدير عدد مطالبات السيارات التكميلي، وهذا يدل على زيادة قدرة وجودة النموذج على التنبؤ.

ويتم في الشكل رقم (٤) التالي توضيح مدى ملائمة النماذج التجميعية المعممة ذات معيار التسوية *Lasso* للدوال الممهدة على المتغيرات غير الخطية لعوامل الخطر: عمر السيارة، وعمر المؤمن، والمسافات المقطوعة والأداء الائتماني:

شكل (٤)

مدى ملائمة النماذج التجميعية المعممة ذات معيار التسوية *Lasso* للدوال الممهدة على المتغيرات غير الخطية



المصدر: مخرجات الحزمة الخاصة بالنماذج التجميعية المعممة GAM في برنامج *R*

ويتضح من الشكل السابق أن النموذج التجميعي المعمم مع استخدام معيار التسوية L1 regularization، يعمل على تقليل معاملات المتغيرات βj ومن الممكن أن تصبح صفر، مما يؤدي إلى ملائمة النموذج التجميعي المعمم ذو معيار التسوية للدوال الممهدة، ويساعد النموذج على اختيار أهم عوامل الخطر ذات العلاقات اللامعلمية وغير الخطية، (وهما: الأداء الائتماني والمسافات المقطوعة) على متغير الاستجابة، مما يؤدي إلى تأثير معنوي ذو دلالة إحصائية عند تسعير السيارات التكميلي.

ثالثاً: النتائج والتوصيات:

أ. نتائج البحث:

١. جنس السائق، وسبب استخدام السيارة، وموديل السيارة من عوامل الخطر المعنوية ذات الدلالة الإحصائية على تقدير عدد المطالبات في تأمين السيارات التكميلي.
٢. درجات الحرية لكل من عمر السيارة، وعمر قائد السيارة، والمسافات المقطوعة بالأميال، والأداء الائتماني أكبر من ١، وهذا يدل على وجود علاقة معنوية ذات دلالة إحصائية بين هذه المتغيرات وعدد المطالبات، كما يؤكد على دقة وقدرة النموذج على التنبؤ.
٣. أثبت النموذج التجميعي المعمم جودة ملائمته للدوال الممهدة على المتغيرات غير الخطية (عمر السيارة، وعمر المؤمن، والمسافات المقطوعة والأداء الائتماني).
٤. تم تقدير أفضل تنبؤ للنموذج عند كل λ باستخدام معيار التصحيح المتقاطع المعمم Generalized Cross Validation Criteria من خلال أقل مربع خطأ Mean Square error وتقدر قيمة λ^{min} بما يعادل ٠,١٤١٥٤١٦.
٥. أصبح معيار AIC للنموذج التجميعي المعمم باستخدام معيار التسوية *L1 regularization* ٢٣١٥٦ مقابل ٢٦٢٨٩ قبل استخدام معيار التسوية، وهذا يدل على زيادة قدرة وجودة النموذج على التنبؤ مع دخول أساليب التسوية.
٦. النموذج التجميعي المعمم مع استخدام معيار التسوية L1 regularization، يعمل على تقليل معاملات المتغيرات βj ومن الممكن أن تصبح صفر، مما يساعد النموذج على اختيار أهم عوامل الخطر ذات العلاقات اللامعلمية وغير الخطية على متغير الاستجابة وهو عدد المطالبات بالنموذج المقترح.

٧. إن استخدام النماذج التجميعية المعممة في تحديد عدد المطالبات في تأمين السيارات التكميلي يساعد على وضع أساس واضح ومحدد للتسعير.

ب. توصيات البحث:

١. يجب الاعتماد على الأساليب العلمية والطرق الاكتوارية في تسعير التأمين وعدم الاعتماد على خبرة الماضي، حتى يمكن للشركات المصرية المنافسة.
٢. دراسة أهم المتغيرات التي تؤثر في عدد وقيم المطالبات مع محاولة تقدير ودمج نماذج أخرى متعددة المتغيرات للتوصيل إلى أفضل تسعير للسيارات التكميلي.
٣. محاولة إدراج متغيرات جديدة لنماذج التسعير
٤. الاعتماد على النماذج التجميعية المعممة وضرورة إدخال أساليب تسوية مثل معيار التسوية *Lasso* للدوال الممهدة على المتغيرات غير الخطية لعوامل الخطر لدقة التنبؤ في تسعير تأمين السيارات التكميلي.
٥. ضرورة إجراء المزيد من الأبحاث التي يمكن استخدامها في مجال تسعير فروع التأمينات العامة المختلفة، وخصوصا استخدام النماذج التجميعية المعممة.

المراجع

أ. المراجع العربية:

- الحصري، محمد حسن. (٢٠١٨). استخدام النماذج المعممة Generalized Models في تسعير تأمينات الممتلكات: دراسة تطبيقية على سوق التأمين المصري. رسالة دكتوراة، كلية التجارة، جامعة القاهرة.
- الديب، علي السيد عبده. (١٩٩٢). تسعير التأمين التكميلي للسيارات الخاصة في ج.م.ع وفقاً للعوامل المؤثرة في درجة الخطر. رسالة دكتوراة، كلية التجارة، جامعة القاهرة.
- الفقي، السباعي محمد. (١٩٩٣). استخدام نموذج حاصل الضرب ذات المتغيرات المتعددة في تقدير عدد المطالبات لتأمين السيارات، المجلة المصرية للدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد ١٧، العدد ٣، يوليو ١٩٩٣، ص ص ٣٤٣-٣٥٧.
- الفقي، السباعي محمد. (١٩٩٣). استخدام النموذج التجميعي ذات المتغيرات المتعددة في الوصول لأحسن تقدير لمطالبات وثائق تأمين السيارات. المجلة المصرية للدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد ١٧، العدد ٣، يوليو ١٩٩٣، ص ص ٣٥٩-٣٧١.٢
- المعداوي، جيهان مسعد، عبد الله، مصطفى يوسف. (٢٠١٧). استخدام الشبكات العصبية في التنبؤ بمطالبات تأمين السيارات التكميلي. المجلة المصرية للدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد ٤١، العدد ٤، أكتوبر ٢٠١٧، ص ص ١٢٩-١٥٣.
- درويش، مروة سعيد. (٢٠١٣). استخدام الشبكات غير العصبية الفازية (Nero-fuzzy system) في ترشيد قرارات الاكتتاب في تأمينات الممتلكات والمسئوليات في السوق المصرية: بالتطبيق على فرع تأمين السيارات التكميلي. رسالة دكتوراه، كلية التجارة – جامعة القاهرة.
- عبد العليم، إيمان عماد. (٢٠١٦). تأمين السيارات التكميلي المشاكل والتحديات – دراسة تحليلية. مجلة البحوث التجارية، كلية التجارة – جامعة الزقازيق، المجلد ٣٨، العدد ٢، يوليو ٢٠١٦، ص ص ٤٣ - ٧٧.
- عثمان، شريف محمد محسن. (٢٠٠٦). تسعير تأمين السيارات التكميلي بالتطبيق على سيارات الميكروباص. رسالة ماجستير، كلية التجارة – جامعة المنوفية.
- علي، مها محمد زكى. (٢٠٠١). استخدام نظرية الفازى في تقدير حجم المطالبات في التأمين التكميلي على السيارات. رسالة ماجستير، كلية التجارة – جامعة المنصورة.

- عجوة، أماني محمد. (٢٠١٩). استخدام توزيع بواسون ذي الاصفار الزائدة وتوزيع هاردل بواسون في نمذجة تكرار المطالبات في تأمين السيارات. المجلة المصرية للدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد ٤٣، العدد ٤، أكتوبر ٢٠١٩، ص ص ١٧١-٢٠٥.
- عجوة، أماني محمد. (٢٠٢٠). نموذج مقترح لتسعير وثائق تأمين السيارات التكميلي. مجلة الدراسات التجارية المعاصرة، كلية التجارة – جامعة كفر الشيخ، المجلد ٦، العدد ٩، يناير ٢٠٢٠، ص ص ٥٤٢-٥٦١.
- عجوة، أماني محمد. (٢٠٢٣). استخدام بعض التوزيعات الاحتمالية المتضخمة الأصفار في حساب السعر الصافي لتأمين السيارات التكميلي. المجلة العلمية للدراسات والبحوث المالية والتجارية، كلية التجارة، جامعة دمياط، المجلد ٤، العدد ١، يناير ٢٠٢٣، ص ص ٤٩٧-٥٢٠.
- فرحلن، أحمد محمد، عبد الحميد، عبد الحميد مصطفى. (٢٠٢٠). نموذج كمي لتقدير القسط الصافي لتأمين السيارات بالتطبيق على شركة التعاونية للتأمين التعاوني بالمملكة العربية السعودية. المجلة العلمية للبحوث التجارية، كلية التجارة – جامعة المنوفية، المجلد ٣٦، العدد ١، يناير ٢٠٢٠، ص ص ١٠٧-١٤٩.
- مهدي، إبراهيم علي محمد، المعداوي، محمد مسعد، الحسيني، الإمام عبد العزيز. (٢٠١٠). نماذج بديلة لتسعير تأمين السيارات التكميلي دراسة – تطبيقية. المجلة المصرية الدراسات التجارية، كلية التجارة – جامعة المنصورة، المجلد ٣٤، العدد ٢، يناير ٢٠١٠، ص ص ٥٧٣-٥٩٤.
- هاشم، محمد محمود. (٢٠١٩). استخدام التوزيعات الاحتمالية في التنبؤ بالمطالبات بالتطبيق على تأمين السيارات التكميلي بشركات التأمين المصرية. مجلة الدراسات المالية والتجارية، المجلد ٢٩، العدد الثاني، كلية التجارة، جامعة بني سويف، يوليو٢٠١٩، ص ص ١-٦٣.
- هاشم، محمد محمود. (٢٠٢٣). تطبيق نظام Bonus-Malus System (BMS) في تسعير التأمين التكميلي على السيارات: دراسة تطبيقية على شركات التأمين المصرية. المجلة العلمية للدراسات والبحوث المالية والتجارية، المجلد الرابع، العدد الأول – الجزء الرابع- كلية التجارة، جامعة دمياط، يناير٢٠٢٣، ص ص ٦١٢-٦٤٣.

ب. المراجع الأجنبية:

- Baayen, R. H., & Linke, M. (2020). An introduction to the generalized additive model. *A practical handbook of corpus linguistics. New York: Springer*, 563-591.
- David, M. (2015) Auto insurance premium calculation using generalized linear models. *Procedia Economics and Finance*, *20*, 147-156.
- David, M., & Jemna, D. V. (2015). Modeling the frequency of auto insurance claims by means of poisson and negative binomial models. *Analele stiintifice ale Universitatii "Al. I. Cuza" din Iasi. Stiinte economice/Scientific Annals of the" Al. I. Cuza"*.
- Díaz Martínez, Z., Fernández Menéndez, J., & García Villalba, L. J. (2023). Tariff Analysis in Automobile Insurance: Is It Time to Switch from Generalized Linear Models to Generalized Additive Models? *Mathematics*, *11*(18), 3906.
- Fahrmeir, L., Sagerer, F., & Sussmann, G. (2007). Geoadditive regression for analyzing small-scale geographical variability in car insurance. *Blätter der DGVFM*, *28*(1), 47-65.
- Fujita, S., Tanaka, T., Kondo, K., & Iwasawa, H. (2020). AGLM: A Hybrid Modeling Method of GLM and Data Science Techniques. In *Actuarial Colloquium Paris 2020*.
- Goldburd, M., Khare, A., Tevet, D., & Guller, D. (2016). Generalized linear models for insurance rating. *Casualty Actuarial Society, CAS Monographs Series*, *5*.

- Guisan, A., Edwards Jr, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, *157*(2-3), 89-100.
- Irwin, M. E. (2005). Generalized Linear Models Introduction.". *Harvard University Statistics Department, študijný material*.
- Jo˜ao Pedro Rodrigues G´ ois. (2019). Bayesian Generalized Additive Models for Car Insurance Data. Thesis to obtain the Master of Science Degree, Tecnico Lisboa.
- Kafková, S., & Křivánková, L. (2014). Generalized linear models in vehicle insurance. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *62*(2), 383-388.
- Kaivanipour, K. (2015). Non-Life Insurance Pricing Using the Generalized Additive Model, Smoothing Splines and L-Curves. KTH ROYAL INSTITUTE OF TECHNOLOGY, SCI SCHOOL OF ENGINEERING SCIENCES.
- Masese, V. (2020). *Application of generalized linear Models in Pricing Usage-Based Insurance* (Doctoral dissertation, University of Nairobi).
- Mazviona, B. W., & Chiduza, T. (2013). The use of statistical distributions to model claims in motor insurance. *Journal of Business, Economics and Law*, *3*, 44-57.
- Tingting, C. (2018). *Generalized Additive Models for Dependent Frequency and Severity of Insurance Claims* (Doctoral dissertation, University of Guelph).
- Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.
- Wood, Simon (2023). Package ‘mgcv’. *R package*, Version1.9-0
- Xie, S., & Shi, K. (2023). Generalized Additive Modelling of Auto Insurance Data with Territory Design: A Rate Regulation Perspective. *Mathematics*, *11*(2), 334.