

## إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام

### نماذج التعلم العميق المعتمدة على الانتباه لتحسين

#### التنبؤ بالتقييمات

م. محمد ديب<sup>1</sup>، أ.د. ناصر أبو صالح<sup>2</sup>

كلية الهندسة المعلوماتية - جامعة البعث

#### المخلص

تستخدم أنظمة التوصية طرقاً معتمدة على ميزات المعطيات المتعلقة بكل من المستخدمين والمنتجات والتفاعلات السابقة فيما بينهم بهدف تقديم أفضل التوصيات من المنتجات للمستخدم. نظام التوصية التحليل العاملي للمصفوفة العصبوني (NeuMF) المعتمد على الفلتر التعاونية يستخدم فقط معرفات المستخدمين والمنتجات في حساب تقييمات المنتجات من أجل تقديم التوصيات المناسبة للمستخدمين الهدف، ولا يستخدم كامل ميزات المعطيات المتعلقة بالمستخدم والمنتج والتفاعلات فيما بينهم. في هذه الدراسة، نقوم بتحسين NeuMF بتسخير جميع ميزات المعطيات باستخدام التعلم العميق للاستفادة من العلاقات غير الخطية بين ميزات المعطيات والاستفادة من الميزات النصية، ونستخدم أيضاً نموذج XGBoost لتجميع خرج طبقات الشبكات العصبونية العميقة المستخدمة. تم تطبيق التجارب والمقارنات على مجموعات المعطيات الثلاث التالية: Book Crossing، Amazon Gift Cards، Shopify App Store. يتفوق النموذج المقترح المحسن على أنظمة التوصية الأساس قبل استخدام الميزات النصية بنسبة تخفيض خطأ تصل 20%-، ويعد استخدام الميزات النصية NeuMF-XT بنسبة تصل 60%- بالنسبة لمقاييس الخطأ MAE و RMSE.

**الكلمات المفتاحية:** التعلم العميق، الفلتر التعاونية، تضمين الكلمات، الانتباه، المعالجة المسبقة.

<sup>1</sup> طالب دكتوراه - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة البعث.

<sup>2</sup> أستاذ دكتور - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة البعث.

## Enriching NeuMF System with Full Data Features Using Attention-Based Deep Learning Models to Improve Ratings Prediction

Eng. Mohammed Deeb<sup>3</sup>, Dr. Nasser Abu Saleh<sup>4</sup>

Faculty of Informatics Engineering, Al-Baath University

### Abstract

Recommendation systems use methods based on data features related to users, products, and previous interactions between them to provide the best products recommendations to the user. The collaborative filtering-based recommendation system, Neural Matrix Factorization (NeuMF) only uses users and products identifiers in calculating products ratings to provide appropriate recommendations to the target users, and does not use the full data features related to the user, product, and interactions between them. In this study, we improve NeuMF by harnessing all data features using deep learning to take advantage of the non-linear relationships between the data features and taking advantage of the textual features. We also use the XGBoost model to aggregate the output of the deep neural networks' layers. Experiments and comparisons were applied to the following three datasets: Book Crossing, Shopify App Store, and Amazon Gift Cards. The proposed improved model outperforms the baseline recommendation systems before using the textual features with an error reduction of up to -20%, and after using textual features NeuMF-XT by up to -60% for MAE and RMSE error metrics.

**Keywords**— Deep Learning || Collaborative Filtering || Word Embedding || Attention || Preprocessing.

<sup>3</sup> PhD Student - Department of Software Engineering and Information Systems, Faculty of Informatics Engineering.

<sup>4</sup> Professor Dr. - Department of Software Engineering and Information Systems, Faculty of Informatics Engineering.

## 1- المقدمة

تلعب أنظمة التوصية دوراً مهماً في تسهيل عمليات البحث واختيار المنتجات والخدمات المناسبة للمستخدمين لتوفر عليهم الوقت والجهد في البحث ضمن أعداد ضخمة من تلك المنتجات المتوفرة في مختلف أنواع المواقع الخدمية، وهذا الدور مهم باعتبار أن أكثر من نصف سكان العالم (تقريباً 67.9%) يستخدمون الإنترنت في معظم أنشطتهم اليومية<sup>5</sup>.

تستخدم أنظمة التوصية معطيات المستخدمين والمنتجات لإجراء العمليات عليها بهدف إيجاد وتقديم أفضل التوصيات بالمنتجات للمستخدمين [24]، وتقسّم إلى 3 أصناف: الفلتر التعاونية التي تعتمد على تفاعلات مستخدم-عنصر، والفلتر المعتمدة على المحتوى التي تعتمد على معطيات المستخدم أو العنصر، ومزيج هجين من التصنيفين السابقين [38]. التحليل العاملي للمصفوفة MF [24] هو نظام توصية معتمد على الفلتر التعاونية يقوم بتمثيل كل مستخدم وعنصر بمتجه من الميزات الكامنة ويسقطها على مصفوفة من المتجهات الكامنة، ينتج من جدائها داخلياً مصفوفة تفاعل عنصر-مستخدم تمثل قيمها تفضيلات المستخدم لذلك العنصر (تقييمات). إنّ MF تعتمد فقط على العلاقات الخطية بين تقييمات المستخدم والعناصر مما يجعل أداؤها التنبؤي غير جيد كفاية. طُبقت العديد من الدراسات لتطوير أنظمة MF للاستفادة من العلاقات غير الخطية بين ميزات المعطيات، من ضمنها المتعلقة بالتعلم العميق مثل نظام التحليل العاملي للمصفوفة العصبوني NeuMF [14]. يُستخدم نظام NeuMF بنيتين عصبونيتين وهما: GMF للتعلم من العلاقات الخطية وMLP للتعلم من العلاقات غير الخطية في تفاعلات مستخدم-عنصر [14]. مع ذلك تعاني NeuMF من نفس مشاكل أنظمة الفلتر التعاونية CF، إذ أنّ عدد التقييمات المقدمة من قبل المستخدمين قليل جداً بالنسبة للأعداد الهائلة المتوفرة من المنتجات وهذا ما يطرح مشاكل مثل تبعثر

<sup>5</sup> احصائيات الإنترنت العالمي، في 12 آذار 2024: <http://www.internetworldstats.com/stats.htm>

المعطيات والاقلاع البارد التي تؤثر سلباً على الأداء التنبؤي لأنظمة CF [33]. أُجريت العديد من الأبحاث من أجل حل تلك المشاكل من خلال استخدام نماذج تعلم الآلة وشبكات التعلم العميق العصبونية وذلك باستخدام كامل ميزات المعطيات المتعلقة بالمستخدم والعنصر [13، 27، 33]. إنّ ميزات المعطيات تملك معلومات مهمة تمثل خصائص المستخدم والعنصر، ويمكن الاستفادة منها في زيادة الدقة التنبؤية لأنظمة CF. تحمل الميزات النصية العديد من المعلومات المهمة التي توصف المستخدم والعنصر وأهمها التي تعبر عن مشاعر المستخدم تجاه العنصر ومدى تفضيله لها [7، 45]. يوجد العديد من الطرق التقليدية لتمثيل النصوص مثل BOW، TF-IDF، ولكنها لا تمثل مشاعر وأهمية الكلمات في النصوص ومراجعات المستخدمين [21، 45]. طُوّرت العديد من الطرق والنماذج لتحليل المشاعر وتصنيفها، إذ تم استخدام نماذج الشبكات العصبونية العميقة في تحليل المشاعر والآراء وقدمت تنبؤات بالتقييمات للمراجعات النصية بأداء عالي عند استخدامها نماذج مدربة مسبقاً لتضمين الكلمات [7، 21، 45]. بالإضافة لمشكلتي التبعرث والإقلاع البارد لدى نموذج NeuMF فإنّ دخله متبعرث عبارة عن أصفار تمثل مواضع معرفات المستخدمين أو العناصر (one-hot-encodings)، وتعتمد فقط على ميزة معرفات المستخدم والعنصر وهي لا تمثل سوى أنّ المستخدم قام بتفضيل العنصر [14]، وهذا غير كافي لتقديم أداء تنبؤي جيد.

في هذه الدراسة، سنقوم بالاستفادة من كامل ميزات المعطيات وخاصة الميزات النصية وذلك بتمكين النموذج من استخدام جميع ميزات المعطيات في طبقة الدخل، وتجنب التلاوم الزائد (Overfitting) والحصول على أداء جيد باستخدام نموذج تعلم الآلة XGBoost [4] كمكدس لنموذجي GMF، MLP ودمجها مع نموذج عصبوني نصي. تم تدريب واختبار 20 نموذج نصي من البنى العصبونية MLP [1، 14]، CNN [22، 25]، BiLSTM [12، 40، 42]، ونموذج هجين منها مع آلية الانتباه [26، 29، 41، 44]، بالإضافة إلى استخدام عدة نماذج مدربة مسبقاً لتضمين الكلمات Word2Vec [31]، FastText [20]، Glove [34]. تقدم الدراسة نموذجاً مقترحاً

NeuMF-XT يتفوق على جميع النماذج المقترحة، النموذج الأصلي NeuMF، وأنظمة الأساس بتطبيق التجارب عليها وإجراء المقارنات بينها بالنسبة لمعايير تقييم الدقة التنبؤية، مقياسي الخطأ MAE و RMSE وعلى 3 مجموعات للمعطيات Book Crossing [47]، Amazon Gift Cards [32]، Shopify App Store [36].

باقي البحث منسق كالتالي: الفقرة 2 تتحدث عن هدف البحث، الفقرة 3 تتحدث عن الأعمال السابقة المتعلقة بموضوع الدراسة، الفقرة 4 تشرح الخوارزميات والطرق المستخدمة في البحث، الفقرة 5 تعرض منهجية البحث وخطوات العمل المتبعة في معالجة المعطيات وبناء النماذج المقترحة، الفقرة 6 تعرض التجارب، أدوات البحث، مجموعات المعطيات، معايير التقييم، أنظمة الأساس، والتجارب والمقارنات للنماذج المقترحة لكل مرحلة، وأخيراً النتائج ومناقشتها في الفقرة 7 والتي تعرض ما تم التوصل إليه في هذه الدراسة.

## 2- هدف البحث

تعاني أنظمة CF من مشكلتي تبعثر المعطيات والاقلاع البارد بسبب قلة التقييمات التي تعتمد عليها في تفاعلات مستخدم-عنصر، ودخل NeuMF متبعثر جداً إذ يعتمد على طريقة One-hot-encoding لبناء متجهات معرفات المستخدمين بحجم العدد الكلي لهم ولا يحمل أي معلومات مفيدة سوى معرف المستخدم والعنصر. ونموذج MLP لا يحوي أي نوع من التنظيم في طبقاته والتي ستحسن من جودة النظام في التنبؤ. تقوم هذه الدراسة على تحسين وتمكين نظام NeuMF من استخدام كامل ميزات المعطيات بما فيها النصية لتحسين الدقة التنبؤية بالتقييمات من خلال تطوير بنية طبقات النموذج للحصول على أفضل النتائج، واستخدام نموذج تعلم الآلة XGBoost كطبقة لربط خرج طبقات النموذج المقترح، واستخدام بنى الشبكات العصبونية العميقة للاستفادة من الميزات النصية، وتمثيلاتها من خلال استخدام نماذج لتضمين الكلمات مدربة مسبقاً، والاستفادة من آلية الانتباه لزيادة الدقة التنبؤية للنموذج المقترح النهائي.

### 3- الأعمال ذات الصلة

استخدمت العديد من الدراسات نظام NeuMF كقالب في بناء أنظمة توصية جديدة تحل مشاكل النماذج والأنظمة مثل تبعثر المعطيات والإقلاع البارد بتطبيق مجموعة من خوارزميات تعلّم الآلة والتعلّم العميق مما يحسن الدقة التنبؤية لها. قدّم الباحثون في [13] نموذجاً جديداً يعتمد على نموذج NeuMF بتبديل البنية العصبونية MLP ببنية عصبونية مكررة LSTM ويدعى النموذج المقترح بـ G-LSTM. أجريت المقارنات مع نماذج عصبونية LSTM، RNN، BPR وذلك بالنسبة لمعايير التقييم HR و NDGC على مجموعتي معطيات من موقع تعليم صيني "أكاديمية جوني"، وقد تفوّق النموذج المقترح عليها من ناحية الدقة والفعالية. قام الباحثون في [27] باستبدال بنية النموذج MLP في النظام NeuMF ببنية عصبونية CNN لنمذجة تفاعلات مستخدم-عصر بشكل مباشر، حيث استُخدمت GMF لنمذجة الميزات الضمنية باستخدام الطريقة الخطية و CNN لنمذجة التفاعلات بالطريقة غير الخطية. تم استخدام مجموعة المعطيات MovieLens ومقارنة النتائج مع الأنظمة الأساس NeuMF، BPR، ItemKNN، GMF، eALS، إذ تفوّق النموذج المقترح على تلك الأنظمة بقيم 0.7068 و 0.4259 لمعايير التقييم HR و NDCG. قدمت الدراسة [7] نموذجاً جديداً InterSentiment للتنبؤ بتقييمات المراجعات من خلال ربط نموذج لتفاعلات مستخدم-عصر مع نموذج لتحليل المشاعر للمراجعات النصية. يقوم النموذج باستخدام طبقة GMF لدمج بنية MLP مع بنية CNN بحيث يكون دخل MLP هو المعارف الفريدة للمستخدم والعنصر أما دخل بنية CNN فهو الميزة النصية للمراجعات. تم استخدام 3 مجموعات معطيات: IMDB، Yelp-2013، Yelp-2014، واستخدام مقياس الخطأ RMSE كمعيار لمقارنة الدقة التنبؤية مع كل من النماذج MF، SVM، TFM، إذ أظهرت المقارنات تفوّق InterSentiment على تلك النماذج. من أجل تقليل مشكلتي التبعثر والإقلاع البارد التي تعاني منها أنظمة الفلترة التعاونية CF وبالأخص NeuMF، قدّمت الدراسة [33] نموذجاً هجيناً NeuMF++ يستخدم شبكات الترميز

التقائمي لإزالة الضوضاء Denoising Autoencoders للحصول على تمثيلات كامنة لميزات المستخدمين والعناصر لُستخدم كدخل لنظام NeuMF للتعلم من تلك الميزات. تمت المقارنات مع NeuMF، GMF، MLP، على مجموعة معطيات MovieLens وتفق النمؤج بقيمة خطأ  $RMSE=0.8681$ .

إن طرق تصنيف وتحليل النصوص التقليدية بحاجة لتطوير بسبب طبيعة النصوص المعقدة والشعورية، مما دعت الحاجة إلى استخدام طرق أفضل للتنبؤ وتحليل المشاعر من الميزات النصية والمراجعات. قَدّم الباحثون في الدراسة [45] نمؤج الشبكة التلافيفية المكررة ثنائية الاتجاه المعتمدة على الانتباه (BRCAN) وهو نمؤج ناتج عن دمج بنية BiLSTM وبنية CNN مع آلية الانتباه. تستخدم طريقة تضمين الكلمات Word2Vec للحصول على متجهات الكلمات ونمؤج BiLSTM لاستخراج معلومات السياق والاعتماديات الطويلة للجملة. أما نمؤج CNN يُستخدم لتحديد الكلمات التي لها دور أساسي في التصنيف وآلية الانتباه تعطي أوزان كبيرة للميزات المفتاحية من النص. أظهرت التجارب تفوق النمؤج المقترح في دقة التصنيف على 4 مجموعات للمعطيات وهي Yahoo Answers، Sogou News، Yelp Reviews، Douba، بالمقارنة مع أنظمة شبكية وحيدة الطبقة، هجينة، معتمدة على الانتباه. أيضاً اقترح الباحثون في الدراسة [21] نمؤجاً جديداً يستخدم بنية CNN لاستخراج واكتشاف الميزات السياقية وتقليل أبعاد الميزات، ويطبق الانتباه على خرج CNN للحصول على مستوى الانتباه لكل كلمة، وبنية BiLSTM للحفاظ على الاعتماديات الطويلة. بالنسبة لتهيئة أوزان تضمين الكلمات تم استخدام TF-IDF و GloVe للحصول على تمثيلات معبرة للنصوص. تمت مقارنة النمؤج المقترح مع نماذج تعلم آلة تقليدية تستخدم أوزان TF-IDF ونماذج عصبونية عميقة تستخدم أوزان TF-IDF و GloVe، ومعيار التقييم هو

دقة التصنيف على 4 مجموعات معطيات وهي US-airline، Sentiment140، SA4A، Sentiment140-MV.

في عمليات تصنيف النصوص وتحليل المشاعر، يتم استخدام الطبقات الكثيفة لربط نتائج الطبقات الخفية وإجراء عملية التصنيف النهائية للنموذج. هذه النماذج عرضة للتلاؤم الزائد بسبب تعلّم أوزان تلك الطبقة لكامل تعقيد مشكلة التصنيفات والتنبؤ. اقترح الباحثون في الدراسة [19] نموذجاً هجيناً يستخدم التعلّم العميق وتعلّم الآلة، إذ تتضمن بنية LSTM لتعلّم الاعتماديات الطويلة للسلاسل النصية وبنية RNN لاستخراج الميزات من تلك السلاسل. يتم معالجة خرج LSTM و RNN من خلال نموذج تعلّم الآلة XGBoost عوضاً عن الطبقات الكثيفة MLP من أجل عمليات التنبؤ النهائية. تفوّق النموذج على نماذج أساس CNN، RNN، CNN-LSTM، RNN-LSTM بمعايير تصنيف الدقة، الإصابة Precision، الاسترجاع Recall، قيمة F1 وذلك على مجموعة معطيات Twitter Airline Sentiment.

في هذه الدراسة، سنقوم بتطوير وتحسين نظام التوصية NeuMF باستخدام نموذج XGBoost كبديل لطبقة الربط بين GMF و MLP، وتمكين MLP من التعلّم من كامل ميزات المعطيات، بالإضافة إلى استخدام بنى عصبونية للتعلّم من الميزات النصية وتجربة أوزان تضمين للكلمات مختلفة ودمج تلك البنى مع NeuMF المُحسّن ومقارنة النتائج.

#### 4- الخوارزميات والطرق المستخدمة

يتكون نظام التوصية NeuMF من ربط بنيتين شبكيتين وهما GMF و MLP [14]، ومن أجل تحسين بنية نظام التوصية استخدمت هذه الدراسة مجموعة من الطرق والنماذج النصية المدربة المسبقة وإضافة بنى شبكية عميقة بهدف تسخير جميع ميزات المعطيات المتوفرة وخاصة النصية من أجل تحسين الدقة التنبؤية لنظام NeuMF، وذلك



وفق منهجيات مقترحة تم تطبيق نتائجها على مجموعات معطيات مختلفة لإجراء التجارب ومقارنة النتائج النهائية مع النظام الأساس NeuMF وعدد من أنظمة التوصية.

#### 4-1-1 إطار عمل نظام توصية عصبوني معتمد على الفترة التعاونية (NCF)

إطار عمل يضم 3 أنظمة توصية CF، تستخدم المعارف الفريدة للمستخدمين والعناصر مع التفاعلات فيما بينهم (التقييمات) في إيجاد تنبؤات بالتقييمات، وتلك الأنظمة هي التحليل العاملي للمصفوفة المعمّم GMF، البيرسيترون متعدد الطبقات MLP، والتحليل العاملي للمصفوفة العصبوني NeuMF المكوّن من ضم البنيتين السابقتين [14].

#### 4-1-1-1 التحليل العاملي للمصفوفة المعمّم (GMF)

GMF هو نظام توصية يُعتبر تعميم وامتداد لنظام التحليلي العاملي للمصفوفة Matrix Factorization (MF) [14]، GMF مكوّن من عدة طبقات شبكية عصبونية تتمثل بطبقة دخل لمتجهات المعارف الفريدة للمستخدمين والعناصر، الشكل (1)، يليها طبقة تضمين Embeddings تمثل أوزانها المتجهات الكامنة الخاصة بالمستخدمين  $p_u$  والعناصر  $q_i$ . بعد ذلك تأتي طبقة مسطحة Flatten تقوم بتسطيح مصفوفتي التضمين ثنائية الأبعاد إلى متجهة وحيدة البعد، بذلك نحصل على متجهتين تمثلان كل من المستخدم والعنصر. يأتي بعدها طبقة جداء عنصر بعنصر  $\odot$ ، المعادلة (1). ثم يتم إدخال المتجهة الناتجة من عملية الجداء إلى طبقة خرج متمثلة بطبقة كثيفة Dense ذات تابع تنشيط غير خطي sigmoid [5] مهينة أوزانها بطريقة التهيئة LeCun uniform، لكن تم استخدام ReLu كتابع تنشيط في الدراسة باعتبار أنّ مشكلة البحث انحدار (هدف ذو قيم مستمرة - التقييمات Ratings). وتابع التحسين المستخدم هو هبوط التدرج العشوائي SGD. توضح المعادلة (2) خرج شبكة GMF (التقييم المتوقع):

$$\phi^{GMF} = p_u^{GMF} \odot q_i^{GMF}, \quad (1)$$

$$\hat{y}_{ui} = a_{out}(h^T(\phi^{GMF})), \quad (2)$$

حيث  $h$  : متجهة أوزان طبقة الخرج،  $a_{out}$  : تابع التنشيط. يمكن الحصول على نظام MF من GMF باستخدام تابع تنشيط خطي Identity و  $h$  متجهة موحدة بقيم 1.

#### 4-1-2- MLP- البيرسبيترون متعدد الطبقات (MLP)

MLP هي نوع من الشبكات العصبونية ذات التغذية الأمامية تتكون من طبقات متعددة كثيفة Dense متصلة بالكامل، تتمثل بطبقة الدخل، وأكثر من طبقة مخفية واحدة (تعلّم عميق)، وطبقة الخرج، الشكل (2). يتم حساب مجموع أوزان كل خلية عصبونية من الطبقات المذكورة باستخدام توابع تنشيط غير خطية. إنّ ضم الميزات الكامنة للمستخدم  $p_u$  والعنصر  $q_i$  غير كافي لنمذجة تأثير الفترة التعاونية ولذلك يقترح الباحثون في الدراسة [14] استخدام MLP لتعلّم التفاعلات والروابط غير الخطية بين تلك الميزات الكامنة. تابع التنشيط المستخدم هو ReLU [10] لأن sigmoid و tanh يعانيان من مشكلة تلاشي التدرج Vanishing Gradient التي تحدث بسبب التدرجات الصغيرة جداً الناتجة أثناء عملية الانتشار الخلفي لتحديث أوزان خلايا طبقات الشبكة والتي ستتلاشى بسبب عمق الشبكة ولن يتم تحديثها. استخدم الباحثون بنية هرمية تضيق نحو الخرج بنسبة 50% من عدد خلايا الطبقة السابقة. توضح المعادلة (3) خرج الشبكة MLP:

$$\phi^{MLP} = a_L(W_L^T(a_{L-1}(\dots a_2(W_2^T \begin{bmatrix} p_u^{MLP} \\ q_i^{MLP} \end{bmatrix} + b_1) \dots)) + b_L), \quad (3)$$

حيث  $W_x, b_x, a_x$ : مصفوفة الأوزان، الانحياز، وتابع تنشيط لخلية الطبقة  $x$ .

## 4-1-3- التحليل العائلي للمصفوفة العصبوني (NeuMF)

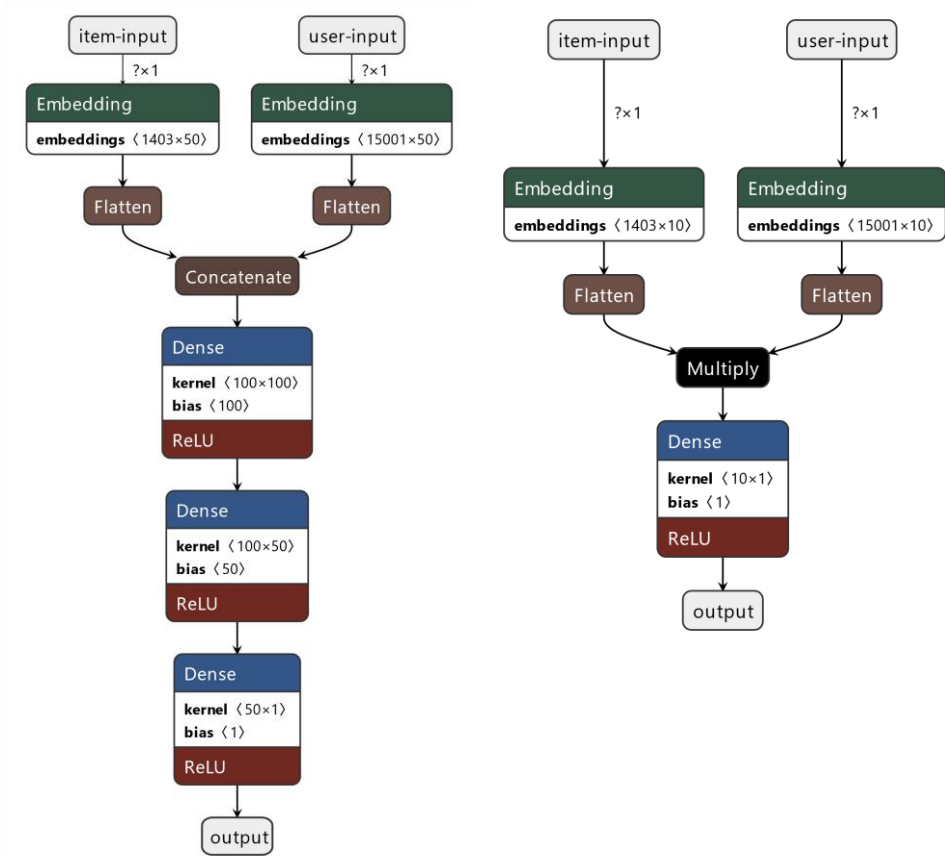
NeuMF هو نظام توصية CF مكون من دمج الطبقة المخفية الأخيرة لكل من GMF و MLP بطبقة ربط Concatenation، الشكل (3)، وبالتالي الاستفادة من العملية الخطية من نمذجة روابط الميزات الكامنة للمستخدمين والعناصر في نموذج GMF، والعملية غير الخطية من روابط الميزات من المعطيات في نموذج MLP [14].  
توضح المعادلة (4) التقييم المتوقع للعنصر  $i$  من أجل المستخدم  $u$ :

$$\hat{y}_{ui} = \sigma \left( h^T \begin{bmatrix} \phi^{GMF} \\ \phi^{MLP} \end{bmatrix} \right), \quad (4)$$

حيث أن  $h$ : متجهة أوزان طبقة الربط،  $\sigma$ : تمثل تابع التنشيط.

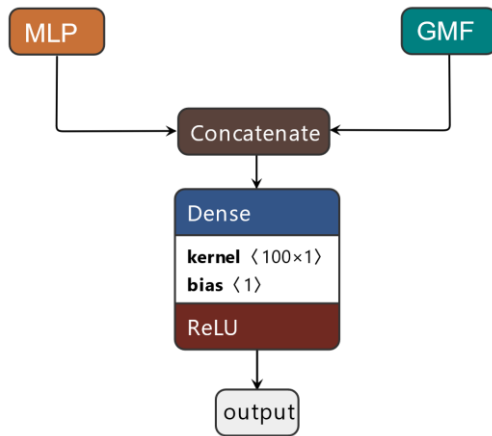
تابع التنشيط المستخدم هو ReLU وطريقة تهيئة الأوزان He\_uniform. بالنسبة لتوابع التحسين Optimizers، كل من GMF و MLP تستخدم تابع التحسين Adam [23] لتحسين مقاييس الدقة والخطأ، أما NeuMF يستخدم تابع التحسين SGD، ومعدل التعلم يساوي 0.01. باعتبار أن النظام يقوم بالتنبؤ بالتقييمات التي تقع ضمن مجال بأعداد حقيقية مستمرة فإن المشكلة هي انحدار Regression وبالتالي سنستخدم مقاييس خطأ MAE و RMSE في تدريب واختبار النماذج.

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه لتحسين التنبؤ بالتقييمات



الشكل (2): نموذج MLP.

الشكل (1): نموذج GMF.



الشكل (3): نموذج التحليل العائلي للمصفوفة العصبوني NeuMF.

## 4-2- الشبكات العصبونية العميقة للميزات النصية

تستخدم الشبكات العصبونية العميقة مع مختلف أنواع المعطيات الرقمية، النصية، الصور والفيديو في مختلف المهام والمسائل، مثل: التصنيف والتنبؤ بالتقييمات، تحليل المشاعر، ترجمة النصوص، معالجة الصور والفيديو [1، 11]. تتناول الدراسة 3 أنواع للشبكات العصبونية بهدف الاستفادة من الميزات النصية في مجموعات المعطيات المستخدمة، وهي البيرسبيترون متعدد الطبقات MLP [1، 14]، الشبكة العصبونية التلافيفية CNN [22، 25، 35، 43]، الشبكة المتكررة ذات الذاكرة الطويلة قصيرة المدى LSTM [8، 9، 15، 40]، بالإضافة إلى نموذج شبكي هجين مقترح يستخدم آلية الانتباه Attention في طبقاته.

بالنسبة للشبكات العصبونية المذكورة، إنّ طبقة الدخل عبارة عن سلاسل عددية تمثل الكلمات في الميزات النصية من مجموعات المعطيات المستخدمة. يليها طبقة تضمين Embedding لأوزان الكلمات  $E$  ذات حجم يساوي معجم الكلمات  $V$  و 300 بُعد لأوزان الطبقة  $d$ ، حيث  $E \in \mathbb{R}^{d \times V}$ . في البداية يتم تدريب الشبكات بدون أوزان تضمين للكلمات ومن ثم يتم تهيئتها بأوزان تضمين باستخدام أوزان ل 3 نماذج من نماذج تضمين الكلمات المدربة مسبقاً، وهي Word2Vec، FastText، GloVe. بالنسبة لطبقات الشبكة، تم استخدام تابع التنشيط ReLU، ومقاييس خطأ MAE و RMSE لتدريب واختبار الأداء التنبؤي للشبكات، وتابع التحسين هو Adam.

## 4-2-1- البيرسبيترون متعدد الطبقات (MLP)

MLP [11] هي نوع من الشبكات العصبونية ذات التغذية الأمامية، تتكون من طبقة مخفية أو أكثر من نوع الطبقة الكثيفة Dense المتصلة بالكامل ذات تابع تنشيط غير خطي. في هذه الدراسة، تم بناء الشبكة باستخدام 4 طبقات كثيفة Dense بعدد خلايا

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه لتحسين التنبؤ بالتقييمات

(64، 32، 5، 1). بالإضافة إلى استخدام الطبقة المسطحة Flatten، وطبقة الإسقاط Dropout [39] بمعدل 20% لخلايا كل طبقة مما يقلل من التعرض لمشكلة التلاؤم الزائد، الشكل (4).

#### 4-2-2- الشبكات العصبونية التلافيفية (CNN)

CNN هي نوع من الشبكات العصبونية ذات التغذية الأمامية طوّرت من أجل مهام الرؤية الحاسوبية [25] وتُستخدم في تصنيف النصوص [22] وتحليل المشاعر [43]. تتكون CNN من 3 مراحل: مرحلة الدخل (صورة ثنائية البعد، نصوص أحادية البعد)، مرحلة استخراج الميزات (على الأقل طبقة تلافيفية convolution وطبقة تجميع pooling)، ومرحلة التصنيف (على الأقل طبقة كثيفة Dense مثل شبكة MLP).

في هذه الدراسة، تم بناء شبكة CNN باستخدام اثنتين من الطبقات التلافيفية أحادية البعد Conv1D بعدد فلاتر 5 و 10، بعدها طبقة تجميع MaxPooling يليها طبقة تسوية الدفعة BatchNormalization [18] التي تحسن استقرار الشبكة بتسوية أوزان دخل كل طبقة على مستوى الميزة لكل العينات في الدفعة الصغرى mini-batch. يلي ذلك طبقة الإسقاط Dropout بمعدل 20% لخلايا الطبقة، يليها طبقتين Dense بحجم 5 و 1 من أجل عملية التنبؤ بالتقييمات وتحليل النصوص. تابع التنشيط المستخدم هو ReLU في كل من طبقتي Conv1D و Dense، الشكل (4).

#### 4-2-3- الذاكرة الطويلة قصيرة المدى ثنائية الاتجاه (BiLSTM)

LSTM هي نوع من الشبكات العصبونية المتكررة طوّرت لحل مشكلة تلاشي التدرج (Vanishing Gradient) التي تعاني منها الشبكات العصبونية المتكررة RNN، بالإضافة إلى قدرتها على الاحتفاظ بالاعتماديات طويلة المدى [9، 15، 40]. تُستخدم

LSTM في المهام المتعلقة بالتسلسلات الزمنية من نصوص، صوت، فيديو، وتحليل المشاعر منها وتصنيفها، ترجمة اللغات، وأنظمة التوصية.

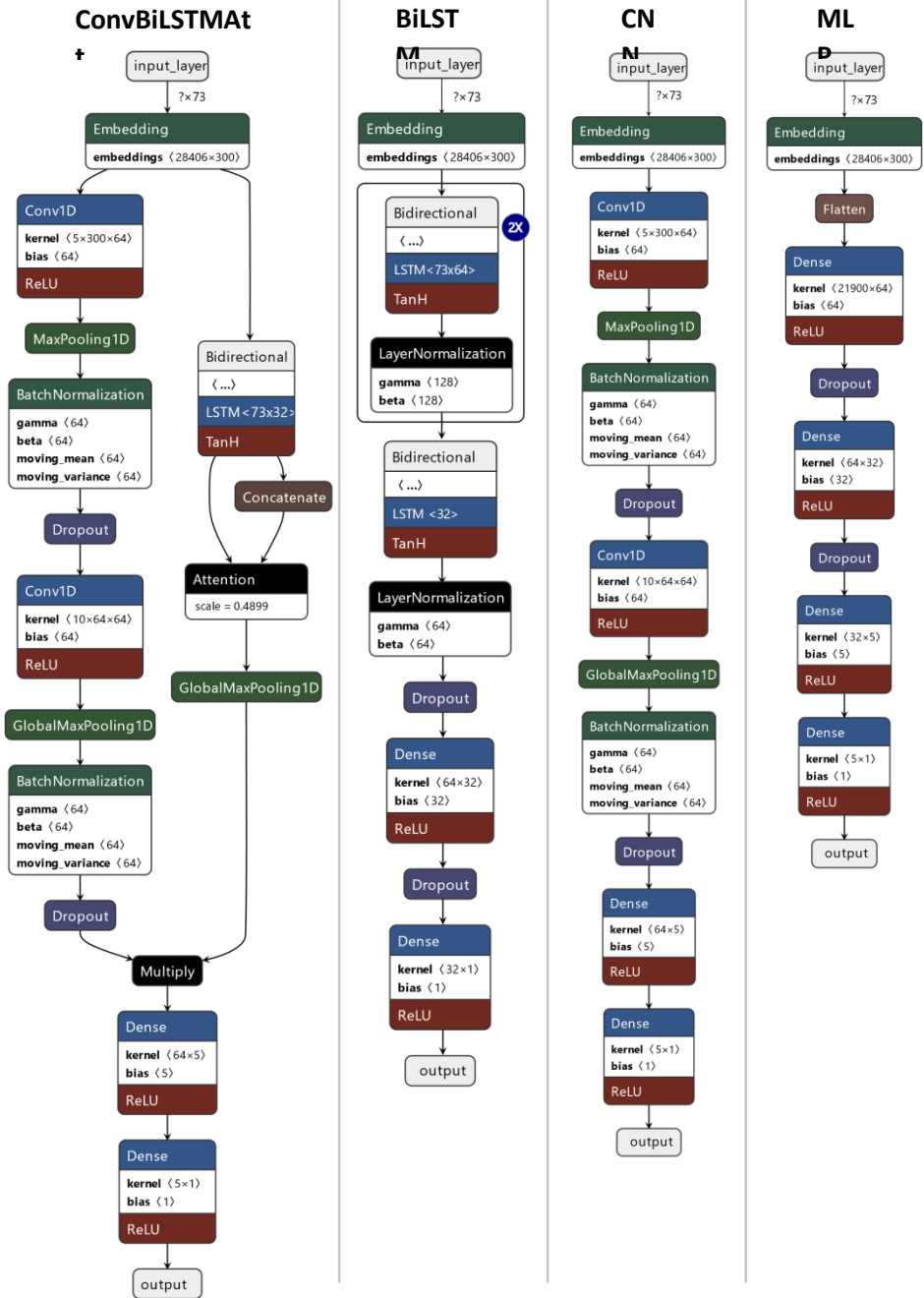
في هذه الدراسة، تم بناء شبكة BiLSTM باستخدام 3 طبقات BiLSTM بحجم 64، 64، 32، يلي كل منها طبقة تسوية الطبقة Layer Normalization [2] التي تحسن استقرار LSTM بتسوية أوزان دخل خلايا الطبقة بأكملها. يلي ذلك طبقة الإسقاط Dropout بمعدل 20% لخلايا الطبقة، يليها طبقتين Dense بحجم 32 و 1 من أجل عملية التنبؤ بالتقييمات وتحليل النصوص. تابع التنشيط المستخدم هو ReLU في طبقات Dense، tanh و sigmoid في طبقات LSTM، الشكل (4).

#### 4-2-4- آليّة الانتباه Attention

آلية الانتباه هي بنية تُستخدم في الشبكات العصبونية لنمذجة التفاعلات طويلة المدى، وهي تُقدم متجهة سياق تساعد الشبكة في التركيز على الأجزاء الأكثر صلة في سلسلة الدخل (مثل الكلمات والعبارات المفتاحية الممثلة لمشاعر النص) [3، 16، 26، 29، 41، 46-44].

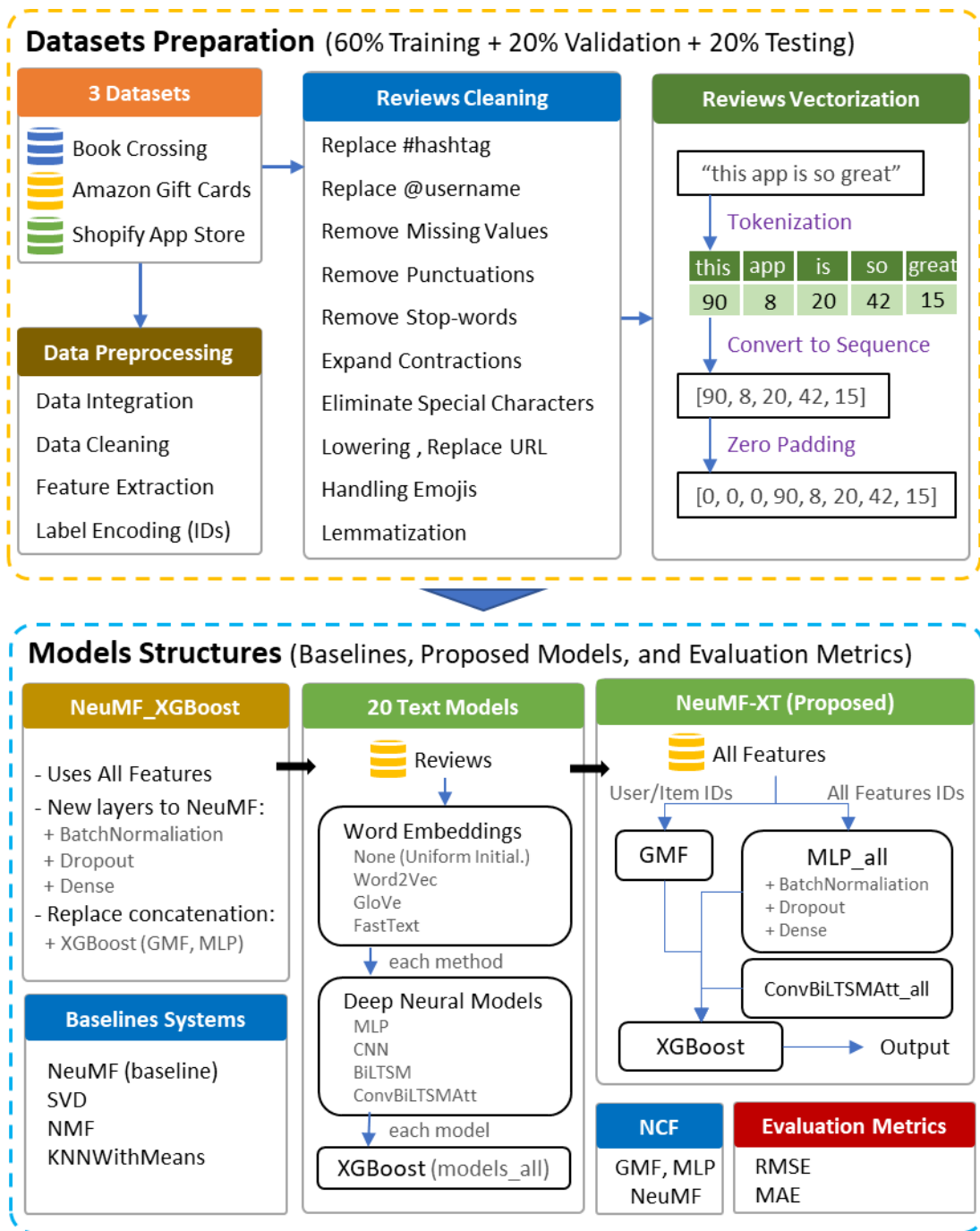
في هذه الدراسة، تم بناء نموذج عصبي هجين مكون من فرعين، الأول طبقات تلافيفية، والثاني طبقة مكررة ثنائية الاتجاه مع الانتباه. بنية النموذج عبارة عن طبقتين تلافيفيتين بعدد فلاتر 64 وحجم 5، 10 وتابع التنشيط ReLU يتخللها طبقة تجميع وتسوية بالدفعة وإسقاط بمعدل 20%. طبقة BiLSTM بعدد خلايا 32 وتابع التنشيط tanh. طبقة انتباه تأخذ الحالة الأخيرة للطبقة المخفية من BiLSTM وإضافة الانتباه إلى خرج الطبقة الكلي. ثم يتم دمج الفرعين بطبقة جداء عنصر بعنصر Multiply، يليها طبقتين كثيفتين Dense بحجم 5 و 1 بتابع تنشيط ReLU، الشكل (4).

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه لتحسين التنبؤ بالتقييمات



الشكل (4): الشبكات العصبونية العميقة المبنية للتعامل مع الميزات النصية.





الشكل (5): منهجية البحث والطرق المستخدمة.

## 5- منهجية البحث

تتضمن منهجية البحث مرحلتين أساسيتين في إجراء هذه الدراسة لتطوير وتحسين نظام NeuMF، الشكل (5):

- مرحلة معالجة وإعداد مجموعات المعطيات التي يتم من خلالها إجراء التجارب والمقارنات بهدف اختيار النماذج ذات أفضل أداء تنبؤي (بأقل خطأ)، وتشمل عمليات المعالجة المسبقة لميزات المعطيات وعمليات تنظيف وتحويل للميزات النصية.
- مرحلة بناء النماذج المقترحة والمقارنات فيما بينها مع النظام الأساس NeuMF وأنظمة توصية CF أخرى، حيث تمت المقارنات على 3 مجموعات معطيات لشمولية النتائج واستخدام مقياسي الخطأ MAE و RMSE لمقارنة دقة التنبؤ.

### 5-1-1- مرحلة معالجة وإعداد مجموعات المعطيات

في هذه الدراسة قمنا باختيار مجموعات المعطيات الثلاث التالية: Book Crossing [47]، Amazon Gift Cards [32]، Shopify App Store [36]، وتطبيق مجموعة من عمليات المعالجة المسبقة عليها، ومعالجة الميزات النصية من خلال تنظيفها وتحويلها إلى سلاسل عددية.

### 5-1-1-1- المعالجة المسبقة للمعطيات Data Preprocessing

تم التحقق من كمالية وصحة المعطيات من خلال تفقد القيم المفقودة والخاطئة في ميزات مجموعات المعطيات وتصحيحها بإعطائها قيم تمثل العينة المنتمية لها وتشابه قيم العينات المشابهة لها. بعدها تم تنظيف المعطيات من العينات المكررة والميزات التي

تحتوي قيمة وحيدة فقط ولا تقدم أي أهمية في عملية التنبؤ، مثل: "Amazon Gift" "Cards(marketplace='us')، إذ جميع قيم الميزة هي 'us' ولذلك تم حذف جميع الميزات المشابهة لها. تم حذف الميزات غير المهمة في إجراء الدراسة مثل عناوين الويب والصور للعناصر. قمنا باستخراج ميزات جديدة من خلال إيجاد معرفات فريدة لقيم كل ميزة من مجموعات المعطيات المستخدمة من أجل النماذج الشبكية العصبونية والتي تتطلب قيم عديدة كدخل لها.

بالنسبة للمجموعة Shopify App Store، فقد قمنا بدمج مجموعتي معطيات Apps.csv و reviews.csv للحصول عليها وإجراء عمليات التنظيف وضمان تكاملية المعطيات، بالإضافة إلى إعادة ترميز معرفات المستخدمين، المنتجات، والمراجعات لاحتوائها على قيم ذات طبيعة مختلفة (عددية ونصية) تُخل بإجراء العمليات عليها. من أجل الحصول على عينات معبرة وذات أهمية للتجارب اخترنا المستخدمين والمنتجات الذين لديهم أكثر من 3 مراجعات، إذ إنَّ اختيار قيمة أعلى ستقصي عدداً كبيراً من عينات التجربة. بالنسبة للمجموعة Book Crossing، قمنا باستخدام ميزاتنا المستخرجة كما هي [6]، وقمنا باستخراج معرفات عديدة فريدة لها كدخل للنماذج العصبونية العميقة.

يوضح كل من الجدول (1)، الجدول (2)، الجدول (3) الميزات المستخدمة والمستخرجة في استخراج معرفات فريدة لميزاتها (باستثناء rating) تمثيلها كدخل في نماذج الدراسة.

### جدول (1): الميزات المستخدمة لمجموعة معطيات Book Crossing

الميزات المستخدمة	التوصيف
user_Age	عمر المستخدم
rating	تقييم المستخدم للكتاب
book_year_of_publication	تاريخ نشر الكتاب
book_num_pages	عدد الصفحات
user_ratings_mean	معدل تقييمات المستخدم
book_title_length	طول عنوان الكتاب
book_word_counts	عدد الكلمات
w2v_lemmatized_title_100_e	الميزة الأولى لتحليل المكون الأساسي ذو 100 بعد لعناوين
phocs_PCA_1	الكتب المطبق عليها عملية lemmatization
w2v_lemmatized_title_100_e	الميزة الثانية لتحليل المكون الأساسي ذو 100 بعد لعناوين
phocs_PCA_2	الكتب المطبق عليها عملية lemmatization
book_isbn_mean_ratings	معدل تقييمات الكتب
book_isbn_counts_ratings	عدد التقييمات الكتب
book_author_mean_rating	معدل تقييمات لكل مؤلف
book_author_count_rating	عدد تقييمات لكل مؤلف

### جدول (2): الميزات المستخدمة والمستخرجة لمجموعة معطيات Amazon Gift Cards

الميزات المستخدمة والمستخرجة	التوصيف
customerID	رقم تعريف الزبون
reviewID	رقم تعريف المراجعة
productID	رقم تعريف المنتج
product_parent	صنف المنتج
product_title	اسم المنتج
rating	تقييم الزبون للمنتج
helpful_votes	عدد الزبائن الذين أعجبته المراجعة
total_votes	العدد الكلي للزبائن الذين قِيموا المراجعة
verified_purchase	تأكيد الشراء للمنتج من قبل الزبون
review_headline	عنوان المراجعة
review_body	جسم المراجعة النصية
review_date	تاريخ المراجعة

## جدول (3): الميزات المستخدمة والمستخرجة لمجموعة معطيات Shopify App Store

التوصيف	الميزات المستخدمة والمستخرجة
رقم تعريف التطبيق	appId
رقم تعريف المراجعة	reviewId
اسم التطبيق	title
الشركة المطورة	developer
تقييم الموقع	site_rating
المستخدم الذي كتب المراجعة	author
جسم المراجعة النصية	body
تقييم المستخدم للتطبيق	rating

## 5-1-2- تنظيف الميزات النصية Reviews Cleaning

بالنسبة للميزات النصية في مجموعات المعطيات المستخدمة، فقد قمنا بتطبيق مجموعة من عمليات تنظيف النصوص، مثل: حذف العينات ذات القيم المفقودة، حذف علامات التعجب والاستفهام وتبديلها بنقطة، حذف كلمات التوقف والمحارف الخاصة، تبديل الهاشتاج واسم المستخدم وعناوين URLs بقيمها فقط (حذف #، @)، توسيع الاختصارات (“she should have” -> “She should’ve”), تصغير حالة الأحرف، تحويل الرموز التعبيرية emojis إلى نص معبر عنها (😊): grinning face with smiling eyes).

تم تطبيق تلك العمليات على الميزة النصية مراجعات المستخدمين (reviews) لكل من مجموعتي المعطيات Amazon Gift Cards، Shopify App Store، بالإضافة إلى إجراء عملية Lemmatization لإعادة الكلمات لمصادرها لتوحيد المعنى الخاص بها في مجموعة معطيات Book Crossing للميزات النصية الممثلة للكتاب (عنوان الكتاب، تصنيفه، تصنيفه الفرعي، ملخص الكتاب).

### 5-1-3- تحويل المراجعات إلى سلاسل عددية Reviews Vectorization

بعد إجراء عمليات تنظيف النصوص على الميزات النصية يجب تحويلها إلى قيم عددية تمثلها لتكون مناسبة كدخل للشبكات العصبونية. في البداية، نقوم بتحويل نص المراجعات إلى متجهات من كلمات النص وتمثيلها بمعرفات فريدة ترمز لها في معجم الكلمات الخاص باستخدام عملية الترميز Tokenization لنحصل على سلسلة عددية تمثل قيمها المعرفات الخاصة بكل كلمة، وفي النهاية نقوم بحشو تلك السلاسل الناتجة بأصفار لتصبح بحجم متساوي ثابت لتكون دخلاً لنماذج الشبكات العصبونية المستخدمة في الدراسة.

قمنا باستخدام متوسط طول كلمات المراجعات مضافاً إليها الانحراف المعياري لتوزيعها الطبيعي للحصول على الطول الأعظمي الثابت للسلاسل العددية المستخدمة كدخل للنماذج، جدول (4).

جدول (4): حجم معجم الكلمات والطول الأعظمي لسلسلة الدخل مقاسة بعدد الكلمات

Shopify App Store	Amazon Gift Cards	Book Crossing	مجموعة المعطيات
73	54	264	الطول الأعظمي للدخل
28,405	31,516	12,755	حجم معجم الكلمات

بالنسبة لتقسيمات مجموعات المعطيات الثلاث، فقد قمنا بتقسيمها إلى 60% كمجموعة تدريب، 20% كمجموعة تحقق، 20% كمجموعة اختبار، أما حجم الدفعة=32 [30]. واستخدمنا مقاييس الخطأ RMSE و MAE لقياس دقة التنبؤ بالتقييمات لنماذج الدراسة.

### 5-2- مرحلة بناء النماذج المقترحة

يشرح هذا القسم التعديلات على نظام NeuMF لتسخير كامل ميزات المعطيات، والنماذج العصبونية العميقة للتعامل مع الميزات النصية وتحليل المشاعر منها. بالإضافة إلى شرح النموذج المقترح النهائي الناتج من جمع العمليات السابقة، الشكل (5).

**5-2-1- نموذج التحليل العاملي للمصفوفة كامل الميزات NeuMF\_XGBoost**

النموذج الأساس NeuMF يتكون من ربط نموذجي GMF و MLP بطبقة ربط Concatenation. من أجل تمكين استخدام كامل ميزات المعطيات، قمنا بتعديل طبقة الدخل لتشمل كامل الميزات (معرفاتها IDs). ومن أجل تنظيم خرج نموذج MLP استخدمنا طبقة BatchNormalization لتسوية أوزان دخل طبقات النموذج لتسهيل تهيئة الأوزان وتسريع عملية التدريب واستقرار الشبكة للحصول على دقة أفضل. من أجل تقليل مشكلة التلاؤم الزائد استخدمنا طبقة الإسقاط Dropout بمعدل 20% من خلايا الطبقة.

إنّ نموذج الأساس NeuMF يستخدم طريقة Lecun uniform لتهيئة الأوزان مع تابع التنشيط ReLU ولكن في دراسة لاحقة [5] وجد الباحثون أنّ استخدام طريقة تهيئة الأوزان He\_uniform تقدم أداءً أفضل للنماذج ولذلك استخدمنا تلك الطريقة في تهيئة أوزان النموذج لنحصل في النهاية على نموذج MLP\_all. قمنا بتبديل طبقة الربط Concatenation بنموذج تعلم الآلة المكدمس XGBoost لتكديس خرج طبقة GMF و MLP\_all لنحصل على النموذج المقترح الأولي NeuMF\_XGBoost. إنّ استخدام نموذج XGBoost سيساعد في إضافة نماذج أخرى مما يساهم في توسيع وظيفة النموذج.

**5-2-2- النماذج العصبونية العميقة للتعامل مع الميزات النصية**

بالنسبة للميزات النصية، فقد استخدمنا مجموعة من الشبكات العصبونية ذات طبقة التضمين Embedding لتحويل سلاسل النصوص إلى أوزان قابلة للتدريب لتعطي كل كلمة متجهات أوزان تُمثل مشاعر ومعنى كل كلمة في سياق النصوص (مراجعات المستخدمين، عناوين وتوصيفات الكتب). أولاً، قمنا ببناء 3 نماذج شبكية MLP، CNN، BiLSTM، الشكل (4)، وتهيئة أوزان طبقة التضمين بقيم عشوائية بدون استخدام أوزان نموذج تضمين كلمات مدربة مسبقاً مثل GloVe. ثانياً، بعد تدريب

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه لتحسين التنبؤ بالتقييمات

النماذج الثلاث بدون أوزان نماذج تضمين قمنا باستخدام 3 نماذج لتضمين الكلمات مدربة مسبقاً لهيئة أوزان طبقات تضمين النماذج الشبكية، وهي Word2Vec، GloVe، FastText، جدول (5). هذه النماذج تقدم متجهات ذات 300 بُعد للكلمات تُمثل المعنى الدلالي وتموضعها في سياق النصوص. ثم تم تدريب النماذج العصبونية عليها في التنبؤ بالتقييمات Ratings.

جدول (5): نماذج تضمين الكلمات المدربة مسبقاً المستخدمة في الدراسة

الباحثون	عدد عينات التدريب	مجموعة المعطيات المدربة عليها	النماذج المدربة مسبقاً المستخدمة	المطورون	الخوارزميات والطرق في تمثيل الكلمات
(Mikolov et al., 2013) [31]	3 million words	Google News dataset (100 billion words)	GoogleNews-vectors-negative300.bin	Google	Word2Vec Word 2 Vector
(Joulin et al., 2017) [20]	1 million words	Wikipedia 2017, UMBC corpus and statmt.org news dataset (16B tokens)	wiki-news-300d-1M.vec	Facebook AI Research (FAIR) Lab	FastText
(Pennington et al., 2014) [34]	6 billion words	Wikipedia 2014 and Gigaword 5	glove.6B.300d.txt	Stanford University	GloVe Global Vectors

ثالثاً، في مرحلة لاحقة خلال إجراء الدراسة وجدنا أنه يمكن الاستفادة من فوائد شبكات CNN، BiLSTM، وتسخير آلية الانتباه لتحسين دقة تنبؤ النموذج. قمنا ببناء نموذج عصبوني هجين مكون من فرعين، الأول يعتمد على CNN، والثاني يعتمد على كل من BiLSTM وآلية الانتباه، الشكل (4). لنحصل بذلك على نموذج ConvBiLSTMAtt يستخدم قدرة الطبقات التلافيفية لـ CNN على استخراج واكتشاف الميزات الدلالية المحلية في كلمات ميزة الدخل النصية، وقدرة BiLSTM على الاحتفاظ



بالاعتماديات طويلة المدى بين ميزات الدخل النصية والتعلم منها. بالإضافة إلى طبقة الانتباه Attention التي تقدم أوزان الانتباه إلى خرج BiLSTM للتركيز على أهم الميزات التي تعبر عن مشاعر السياق المتمثلة بالتقييمات. رابعاً، قمنا باستخدام نموذج تعلم الآلة XGBoost لتكديس خرج كل من النماذج الشبكية على حدا مع مقابلاتها من نماذج تضمين الكلمات المدربة مسبقاً بحيث نحصل على نموذج يكسب خرج 4 نماذج شبكية لنفس النوع مستخدمةً 4 طرق لتهيئة أوزان طبقة التضمين None، Word2Vec، Glove، FastText. مثلاً، نموذج CNN\_all هو نموذج نصي يكسب 4 نتائج معتمدة على كل من None، Word2Vec، Glove، FastText، و"all" ترمز إلى استخدام جميع طرق تهيئة الأوزان.

بالنسبة لبارامترات النماذج، فقد قمنا بمعايرة معدل التعلم للمُحسن Adam لكل النماذج وعدد الخلايا لطبقات MLP و BiLSTM وعدد وحجم فلاتر CNN لاختيار أفضل القيم لها، جدول (6).

يبين الجدول (7) قيم البارامترات الخاصة لكل نموذج عصبوني نصي، والجدول (8) يبين قيم البارامترات العامة المشتركة بين النماذج.

جدول (6): معايرة البارامترات المستخدمة في النماذج العصبونية النصية

النماذج	البارامترات	مجال القيم	الأفضل
MLP	units	(64, 128, 256)	64
	learning_rate	(0.001, 0.01, 0.1)	0.01
CNN	filters	(32, 64, 128)	64
	kernel_size	(3, 5, 10)	5
	learning_rate	(0.001, 0.01, 0.1)	0.01
BiLSTM	units	(16, 32, 64)	64
	learning_rate	(0.001, 0.01, 0.1)	0.001

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه  
لتحسين التنبؤ بالتقييمات

جدول (7): قيم البارامترات الخاصة للنماذج العميقة النصية المستخدمة في الدراسة

البارامترات المستخدمة	قيم البارامترات	النماذج
Dense عدد الخلايا في طبقات معدل الإسقاط تابع التنشيط في Dense	units=(64, 32, 5, 1) rate=0.2 activation='relu'	<b>MLP</b>
Conv1D عدد الخلايا في طبقات Dense عدد الخلايا في طبقات معدل الإسقاط تابع التنشيط في Conv1D تابع التنشيط في Dense نوع الحشو	filters=(64, 64) units=(5, 1) rate=0.2 activation='relu' activation='relu' padding='same'	<b>CNN</b>
عدد الخلايا في طبقات BiLSTM Dense عدد الخلايا في طبقات معدل الإسقاط تابع التنشيط في Dense	units=(64, 64, 32) units=(32, 1) rate=0.2 activation='relu'	<b>BiLSTM</b>
عدد الخلايا في طبقات Conv1D حجم الفلاتر معدل الإسقاط تابع التنشيط في Conv1D تابع التنشيط في Dense نوع الحشو عدد الخلايا في طبقات BiLSTM نمط التسجيل في طبقة Attention استخدام مُقيس في طبقة Attention عدد الخلايا في طبقات Dense	filters=(64, 64) kernel_size=(5, 10) rate=0.2 activation='relu' activation='relu' padding='same' units=32 score_mode='dot' use_scale=True units=5, 1	<b>ConvBiLSTMAtt</b>

## جدول (8): قيم البارامترات العامة المشتركة للنماذج العميقة النصية المستخدمة

البارامترات العامة في النماذج	قيم البارامترات
طريقة التحسين	optimizer='adam'
تابع الخسارة	loss='mse'
معيار الخطأ	metrics=['mae']
تقسيمات مجموعة المعطيات	60% training, 20% validation, 20% testing
حجم الدفعة	batch_size=32
عدد دورات التدريب	epochs=10
حجم طبقة التضمين	embed_dim=300
أوزان طبقة التضمين	None, Word2Vec, FastText, GloVe
الطول الأعظمي لدخل طبقة التضمين	متوسط أطوال المراجعات + انحرافها المعياري

## 5-2-3- NeuMF-XT النموذج المقترح

إنّ ناتج المرحلة السابقة هو 20 نموذج عصبوني نصي (4 نماذج شبكية  $\times$  4 نماذج تضمين + 4 نماذج شبكية مكدسة). قمنا بدمج كل نموذج نصي على حدا مع نموذج NeuMF\_XGBoost لتمكينه من الاستفادة من كامل ميزات المعطيات العددية والفئوية والنصية. قمنا بإجراء عملية الدمج باستخدام نموذج XGBoost الخاص بنموذج NeuMF\_XGBoost نفسه، لنحصل على نموذج ذو دقة تنبؤية أعلى. من خلال إجراء المقارنات واعتماد مقاييس الخطأ RMSE، MAE لمقارنة دقة أداء التنبؤ بالتقييمات لكل من النماذج الناتجة وجدنا أنّ النموذج الذي يربط NeuMF\_XGBoost مع النموذج النصي ConvBiLSTMAAtt\_all قدّم أفضل النتائج، ويدعى NeuMF-XT، الشكل (5).

## 6- التجارب

### 6-1- أدوات البحث

تم تطبيق تجارب الدراسة على بيئتي عمل محلية وسحابية، وهما:

- حاسب محمول بنظام Windows 11 Home 64bit ومعالج Intel Core i5 بسرعة 2.6GHz وذاكرة RAM 8GB، بطاقة الفيديو المستخدمة NVIDIA GeForce RTX 3050 Ti 4GB وCUDA 11.2.

- بيئة عمل سحابية Google Collaboratory بنظام Ubuntu 22.04.3 LTS ومعالج Xeon بسرعة 2.3GHz وذاكرة RAM 12.7GB، بطاقة الفيديو المستخدمة NVIDIA Tesla T4 16GB وCUDA 12.2.

استخدمنا لغة بايثون 3.10.9 ومكتباتها NumPy, Pandas, Matplotlib, Sklearn في إجراء عمليات المعالجة على مجموعات المعطيات وعرض الأشكال لمقارنة النتائج، ومكتبات re، contractions، emoji، nltk لمعالجة النصوص. واستخدمنا مكتبة Tensorflow 2.10.1 لبناء وتدريب النماذج المقترحة واختبارها ومقارنة النتائج. استخدمنا بيئة عمل Anaconda Jupyter Notebook لكتابة وتوثيق الأكواد البرمجية.

### 6-2- مجموعات المعطيات

قمنا باستخدام مجموعات المعطيات الثلاثة التالية في الدراسة: Book Crossing، Amazon Gift Cards، Shopify App Store. قمنا باستخدام Book Crossing [6، 47] لاحتوائها على ميزات عديدة ومفيدة في عمليات التنبؤ بالتقييمات، مثل: الميزات الإحصائية، التمثيلات النصية، ميزات توصف الكتاب بأكثر قدر ممكن. واستخدمنا مجموعتي معطيات Amazon Gift Cards [32]، Shopify App Store [36] لاحتوائهما على الميزة النصية مراجعات المستخدمين للمنتجات والتي تعبر عن

مدى تفضيل المستخدم للمنتج، ومجموعة الكتب لا تحوي ميزة نصية معبرة عن تفضيلات المستخدم، إنما تحوي ميزات نصية تُمثل الكتاب فقط (عنوانه، تصنيفه، تصنيفه الرئيسي، ملخص شامل). أجرينا مجموعة من عمليات المعالجة المسبقة والتنظيف على كامل الميزات واستخلاص معرفات فريدة لقيم الميزات ما عدا ميزة التقييمات Rating وهي هدف التنبؤ.

تم تقسيم المجموعات إلى 60% للتدريب، 20% للتحقق، و20% اختبار. يوضح الجدول (9) خصائص مجموعات المعطيات الإحصائية ومجال التقييمات قبل وبعد معالجتها.

جدول (9): خصائص مجموعات المعطيات قبل وبعد معالجتها للدراسة

Shopify App Store		Amazon Gift Cards		Book Crossing		مجموعة المعطيات
بعد	قبل	بعد	قبل	بعد	قبل	
<b>الخصائص</b>						
77,538	287,467	-	148,310	14,239	1,149,780	عدد العينات
1,402	179,656	-	142,467	1,091	278,858	عدد المستخدمين
15,000	2,169	-	1,776	603	271,379	عدد العناصر
76,155	283,864	-	148,310	603	271,379	عدد النصوص
لا يوجد	المؤلف: 151 الجسم: 186	لا يوجد	عنوان: 3 جسم: 2 تاريخ: 1	لا يوجد	العمر: 110,762 المؤلف: 1 الناشر: 2	القيم المفقودة
0.369%	0.074%	-	0.059%	2.164%	0.002%	كثافة التقييمات
-	(1, 5)	-	(1, 5)	-	(1, 10)	مجال التقييمات
19	22	23	15	69	14	عدد الميزات

### 6-3- معايير تقييم دقة التنبؤ

من أجل اختبار جودة النماذج ومقارنة الدقة التنبؤية لكل منها مع النظام الأساس NeuMF استخدمنا مقاييس الخطأ MAE متوسط الخطأ المطلق و RMSE جذر متوسط مربع الخطأ [37]، إذ كلما صغرت قيمها كانت دقة التنبؤ أفضل.

$$MAE = \frac{\sum_{j=1}^N |\hat{r}_j - r_j|}{N} \quad (6) \quad \left| \quad RMSE = \sqrt{\frac{\sum_{j=1}^N (\hat{r}_j - r_j)^2}{N}} \quad (5)$$

ونسبة التخفيض في الخطأ RMSE أو MAE فتحسب من المعادلة (7):

$$\text{Reduction\%} = \frac{[RMSE|MAE]_{\text{model}} - [RMSE|MAE]_{\text{baseline}}}{[RMSE|MAE]_{\text{baseline}}} * 100, \quad (7)$$

### 6-4- أنظمة الأساس Baselines

من أجل مقارنة النموذج المقترح مع أنظمة CF مختلفة عن الأساس NeuMF استخدمنا مكتبة بايثون Surprise [17] والتي تقدم مجموعة من الأنظمة التي يمكن اختبار المعطيات عليها ومقارنة نتائجها مع النموذج المقترح. استخدمنا الأنظمة التالية: KNNWithMeans، [28] Non-negative Matrix Factorization (NMF)، [38] Singular Value Decomposition (SVD)، [24].

### 6-5- التجارب والمقارنات خلال مراحل منهجية الدراسة

يُقدّم هذا القسم مجموعة من المقارنات وفقاً لمراحل بناء النماذج المقترحة، وذلك بالنسبة لمعايير دقة التنبؤ، مقاييس الخطأ RMSE و MAE ومقارنتها مع NeuMF الأساس.

في الجدول (10)، قمنا بمقارنة النموذج المحسن NeuMF\_XGBoost مع النظام الأساس NeuMF وذلك على مجموعتي معطيات Book Crossing و Amazon

Gift Cards بالنسبة لقيمتي الخطأ RMSE و MAE، إذ حقق النموذج المحسن قيم  $MAE = 0.4049$  و  $RMSE = 0.8715$  بنسبة تخفيض  $66.24\%$  و  $46.87\%$  بالترتيب من أجل Book Crossing بالمقارنة مع قيم الأساس NeuMF. بالمثل من أجل Amazon Gift Cards، حقق النموذج المحسن قيم  $MAE = 0.3189$  و  $RMSE = 0.6659$  بنسبة تخفيض  $9.15\%$  و  $22.88\%$  بالترتيب. بعدها قمنا بتوسيع النموذج المحسن بنموذج شبكي نصي CNN بدون أوزان تضمين كلمات مدربة مسبقاً لنحصل نموذج NeuMF\_XGBoost\_CNN والذي حقق نتائج أفضل من الأساس والنموذج المحسن بقيم  $MAE = 0.3133$  و  $RMSE = 0.8283$  بنسبة تخفيض  $73.88\%$  و  $49.50\%$  بالترتيب لـ Book Crossing، وقيم  $MAE = 0.1750$  و  $RMSE = 0.4240$  بنسبة تخفيض  $50.14\%$  و  $50.90\%$  بالترتيب لـ Amazon Gift Cards. قمنا بعدها باستخدام مجموعة معطيات Shopify App Store والحصول على نفس النتائج.

جدول (10): مقارنة دقة تنبؤ النماذج المقترحة ونسبة تخفيض الخطأ مع الأساس NeuMF

Shopify App Store		Amazon Gift Cards		Book Crossing		مجموعة المعطيات
RMSE	MAE	RMSE	MAE	RMSE	MAE	
						النماذج
0.8104	0.4111	0.8635	0.3510	1.6402	1.1993	NeuMF (الأساس)
0.5960	0.2670	0.6659	0.3189	0.8715	0.4049	NeuMF_XGBoost
(-26.46 %)	(-35.05 %)	(-22.88 %)	(-9.15 %)	(-46.87 %)	(-66.24 %)	
0.3333	0.1374	0.4240	0.1750	0.8283	0.3133	NeuMF_XGBoost_CNN
(-58.87 %)	(-66.58 %)	(-50.90 %)	(-50.14 %)	(-49.50 %)	(-73.88 %)	

من المقارنة السابقة وجدنا أن استخدام نموذج شبكي CNN للتعامل مع النصوص قدم أداءً تنبؤياً أفضل من النموذج الأساس والنموذج المحسن من NeuMF، لذلك قمنا

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه لتحسين التنبؤ بالتقييمات

بناء مجموعة من النماذج العصبونية واستخدمنا عدة طرق لتهيئة أوزان تضمين الكلمات ودمجها مع النموذج المحسن NeuMF\_XGBoost، ومن ثم قمنا بإجراء المقارنات عليها بالنسبة لقيم RMSE و MAE على كل من مجموعتي المعطيات Amazon Gift Cards و Shopify App Store، جدول (11) وجدول (12). من الجدولين وجدنا أنّ النموذج المقترح NeuMF\_XGBoost+ConvBiLSTMAtt حقق أفضل دقة تنبؤ على مقاييس الخطأ MAE و RMSE بالنسبة لمعظم طرق تضمين الكلمات وكان أفضل نموذج هو النموذج NeuMF\_XGBoost+ConvBiLSTMAtt\_all والذي يكس نتائج النماذج النصية الأربعة المعتمدة على كل طريقة تضمين على النموذج ConvBiLSTMAtt، وكاختصار للنموذج المقترح أشرنا إليه بـ NeuMF\_XT وهو أفضل نموذج في الدراسة، إذ حقق قيم MAE=0.1080 و RMSE=0.3297 لمجموعة معطيات Amazon Gift Cards، وقيم MAE=0.0585 و RMSE=0.2824 لمجموعة معطيات Shopify App Store.

جدول (11): مقارنة دقة تنبؤ النماذج المقترحة بحسب النماذج النصية لكل من أوزان تضمين الكلمات المدربة مسبقاً

#### لمجموعة المعطيات Amazon Gift Cards

All		Word2Vec		Glove		FastText		No Weights		أوزان تضمين الكلمات
RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	
0.3663	0.1202	0.4459	0.1846	0.4521	0.1874	0.4519	0.1847	0.4488	0.1864	MLP
0.3407	0.1088	0.4192	0.1697	0.4266	0.1759	0.4279	0.1732	0.4338	0.1773	LSTM
0.3396	0.1101	0.4231	0.1740	0.4210	0.1754	0.4163	0.1722	<b>0.4240</b>	<b>0.1750</b>	CNN
<b>0.3297</b>	<b>0.1080</b>	<b>0.4094</b>	<b>0.1659</b>	<b>0.4103</b>	<b>0.1673</b>	<b>0.4159</b>	<b>0.1707</b>	0.4244	0.1751	ConvBiLSTMAtt
-	-	-	-	-	-	-	-	0.8635	0.3510	NeuMF (الأساس)



جدول (12): مقارنة دقة تنبؤ النماذج المقترحة بحسب النماذج النصية لكل من أوزان تضمين الكلمات المدربة مسبقاً

لمجموعة المعطيات Shopify App Store

All		Word2Vec		Glove		FastText		No Weights		أوزان تضمين الكلمات
RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	
0.3170	0.0708	0.3514	0.1411	0.3525	0.1419	0.3493	0.1417	0.3595	0.1521	النماذج MLP
0.2971	0.0635	0.3281	0.1310	0.3322	0.1321	0.3281	0.1294	0.3314	0.1310	LSTM
0.2841	0.0595	0.3243	<b>0.1284</b>	<b>0.3231</b>	0.1307	0.3314	0.1286	0.3333	0.1374	CNN
<b>0.2824</b>	<b>0.0585</b>	<b>0.3189</b>	0.1286	0.3293	<b>0.1304</b>	<b>0.3108</b>	<b>0.1255</b>	<b>0.3199</b>	<b>0.1263</b>	ConvBiLSTMAtt
-	-	-	-	-	-	-	-	0.8104	0.4111	NeuMF (الأساس)

بعد الانتهاء من إجراء التجارب والمقارنات والحصول على أفضل نموذج وهو NeuMF\_XT (والذي يتألف من دمج النموذج المحسن NeuMF\_XGBoost مع نموذج الشبكة العصبونية النصية ConvBiLSTMAtt\_all والذي بدوره يدمج 4 نماذج لـ ConvBiLSTMAtt معتمدة على كل طريقة تضمين) قمنا بإجراء مقارنات بين النموذج المقترح وعدة أنظمة CF، ولاحظنا تفوق NeuMF\_XT على تلك الأنظمة، جدول (13) حيث قدم نسب تخفيض تصل إلى 60% لقيم RMSE و MAE على المجموعات الثلاث.

إغناء نظام NeuMF بكامل ميزات المعطيات باستخدام نماذج التعلّم العميق المعتمدة على الانتباه  
لتحسين التنبؤ بالتقييمات

جدول (13): مقارنة دقة تنبؤ النموذج المقترح ونسبة تخفيض الخطأ مع أنظمة التوصية CF

Shopify App Store		Amazon Gift Cards		Book Crossing		مجموعة المعطيات
RMSE	MAE	RMSE	MAE	RMSE	MAE	
النماذج						
0.8104	0.4111	0.8635	0.3510	1.6402	1.1993	NeuMF (الأساس)
0.9762 (+20.46 %)	0.7103 (+72.79 %)	0.8517 (-1.36 %)	0.4948 (+40.98 %)	2.6152 (+59.45 %)	2.2970 (+91.53 %)	NMF
0.8287 (+2.25 %)	0.3798 (-7.61 %)	0.8289 (-4 %)	0.4638 (+32.15 %)	1.6399 (-0.02 %)	1.2004 (+0.09 %)	KNNWithMeans
0.7522 (-7.18 %)	0.3852 (-6.31 %)	0.8224 (-4.76 %)	0.4543 (+29.43 %)	1.4868 (-9.35 %)	1.1333 (-5.50 %)	SVD
<b>0.2824</b> (-65.15 %)	<b>0.0585</b> (-85.77 %)	<b>0.3297</b> (-61.82 %)	<b>0.1080</b> (-69.23 %)	<b>0.9451</b> (-42.38 %)	<b>0.3297</b> (-72.51 %)	NeuMF_XT (النموذج المقترح)

### 7- النتائج ومناقشتها

من نتائج المقارنات السابقة، بملاحظة نسب التخفيض في خطأ MAE و RMSE للنموذج الأساس في الجدول (10)، فقد وجدنا أنّ النموذج المحسن NeuFM\_XGBoost تفوق على الأساس بنسبة تخفيض جيدة لكل من قيمتي الخطأ MAE و RMSE على مجموعات المعطيات الثلاث، وذلك بسبب إضافة طبقات التسوية بالدفعات BatchNormaliztion، الإسقاط Dropout إلى بنية MLP في NeuMF والتي تمنع مشكلة التلاؤم الزائد وتزيد من استقرار النموذج، وتمكيننا للنموذج من إدخال كامل معرفات ميزات مجموعة معطيات التجربة، بالإضافة إلى استخدام مكس XGBoost الذي يدمج خرج كل من MLP\_all و GMF معاً للحصول على نتائج أفضل من الأساس NeuMF. وجدنا أيضاً أنّ استخدام النموذج NeuFM\_XGBoost\_CNN (النموذج النصي CNN + النموذج المحسن) يتفوق

على كل من النموذج الأساس والمحسن وينسب تخفيض جيدة جداً تفوق 50% لكل من MAE و RMSE على مجموعات المعطيات الثلاث. وهذا ما شجعنا إلى متابعة بناء نماذج شبكية نصية مختلفة لمقارنة أدائها في تحسين الأداء التنبؤي للأساس NeuMF. وأيضاً لاحظنا أن النموذج المحسن في مجموعة Book Crossing قدم نسب تخفيض أعلى بالمقارنة مع نسب التخفيض بالنسبة لمجموعتي المعطيات الثانية والثالثة وذلك لأن المجموعة الأولى Book Crossing تحوي على ميزات ممثلة أكثر للعناصر (الكتب) والمستخدمين، وبالمثل نلاحظ أن النموذج المقترح NeuFM\_XGBoost\_CNN الذي يُمكن الأساس من الاستفادة من الميزات النصية قد حقق نسب تخفيض لمقاييس الخطأ MAE و RMSE لكل من مجموعتي المعطيات Amazon Gift Cards و Shopify و App Store أفضل من نسبة التخفيض لمجموعة المعطيات Book Crossing بالمقارنة مع النموذج المحسن وذلك لأن Book Crossing يحوي ميزات نصية لا تعبر عن مشاعر وآراء المستخدم حول المنتج، بينما Amazon Gift Cards و Shopify و App Store يحويان الميزة النصية (مراجعات المستخدم) والتي تعبر عن آراء المستخدم وتُقدّم معلومات مفيدة لتحسين الدقة التنبؤية.

بالنسبة للنماذج المقترحة المعتمدة على النماذج العصبونية النصية، وجدنا من نتائج المقارنات السابقة في الجدول (11) والجدول (12) أن جميع النماذج حققت قيم خطأ MAE و RMSE أقل من النموذج الأساس، وأفضلها كان النموذج المعتمد على ConvBiLSTMAtt والذي حقق أدنى خطأ في معظم التجارب لكل من نماذج تضمين الكلمات، إذ أن ConvBiLSTMAtt يعتمد على قدرة الطبقات التلافيفية في استخراج واكتشاف الميزات الدلالية المحلية (المعبّرة عن المشاعر) في كلمات السياق للدخل، وقدرة BiLSTM على الاحتفاظ بالاعتماديات الطويلة بين ميزات الدخل النصية، بالإضافة إلى قدرة الانتباه Attention في تقديم أوزان الانتباه إلى خرج BiLSTM

للتركز على أهم الميزات المفتاحية التي تعبر عن مشاعر السياق المتمثلة بالتقييمات. بالنسبة لمجموعتي المعطيات Amazon Gift Cards و Shopify App Store وجدنا أنّ أفضل نموذج مقترح من هذه المقارنات هو النموذج NeuMF-XT الذي يضم النموذج المحسن NeuFM\_XGBoost مع النموذج العصبوني النصي المكس ConvBiLSTMAtt\_all. وقد حقق النموذج قيمة NSE [37] جيدة جداً تساوي 84% لـ Amazon Gift Cards و 88% لـ Shopify App Store، إذ أنّ القيمة تدل على مقدار تشابه التنبؤات مع التقييمات الحقيقية وقيمها بين  $-\infty$  و 100%، القيمة 0% أي النموذج يشابه متوسط قيم الهدف أما 100% أي النموذج يقدم تنبؤات مماثلة تماماً للهدف.

من المقارنات في ال جدول (13) وجدنا أنّ النموذج المقترح NeuMF-XT تفوّق على كل من أنظمة التوصية CF: NeuMF، NMF، KNNWithMeans، SVD بالنسبة لقيم الخطأ MAE و RMSE، وحقق أفضل نسب تخفيض للخطأ بالنسبة للنموذج الأساس على مجموعات المعطيات الثلاث. بما أنّ نسب التخفيض الخاصة بالنموذج المقترح جيدة جداً وهي أفضل بكثير من أنظمة التوصية CF فإننا نقترح استخدام هذه المنهجية التي تقوم على استخدام نموذج مكس XGBoost لتحسين نتائج NeuMF وتوسيع مهام النظام الأساس ليتمكن إضافة نماذج عصبونية نصية تمكن النظام من الاستفادة من الميزات النصية المعبرة عن آراء المستخدمين مثل نموذج ConvBiLSTMAtt\_all الذي يعتمد على خصائص كل من CNN و BiLSTM والانتباه لزيادة الدقة التنبؤية بتقييمات العناصر (كتاب، بطاقة هدية، تطبيق).

## 8- الاستنتاجات والتوصيات

في هذه الدراسة قمنا بتطوير بنية نظام NeuMF للحصول على نموذج ذو أداء تنبؤي أفضل من النظام الأساس NeuMF بأقل قيم ممكنة للخطأ MAE و RMSE في التنبؤ بالتقييمات. اقترحنا في هذه الدراسة نموذج NeuMF-XT والذي قمنا فيه بتعديل بنية نظام NeuMF بإضافة طبقات التسوية بالدفعات والإسقاط واستخدام نموذج تعلم الآلة XGBoost لتكديس كل من MLP\_all و GMF للحصول على نموذج محسن NeuMF\_XGBoost والذي قمنا أيضاً بربطه مع نموذج عصبوني نصي ConvBiLSTMatt\_all والذي يستفيد من قدرة الطبقات التلافيفية في استخراج واكتشاف الميزات الدلالية المحلية وقدرة BiLSTM على الاحتفاظ بالاعتماديات الطويلة بين ميزات الدخل النصية، وقدرة الانتباه في تقديم أوزان الانتباه إلى خرج BiLSTM للتركيز على أهم الميزات المفتاحية التي تعبر عن مشاعر السياق المتمثلة بالتقييمات. تُمكن المنهجية المقترحة النظام NeuMF من استخدام كامل ميزات المعطيات وخاصة النصية منها للحصول على نتائج جيدة بأفضل أداء تنبؤي بالتقييمات. أظهرت النتائج تفوق النماذج المقترحة بالنسبة لمجموعات المعطيات الثلاث Book Crossing، Amazon Gift Cards، Shopify App Store، بنسب تخفيض للخطأ تصل بين 60% و 70%.

بالنسبة للأعمال المستقبلية، النموذج المقترح قابل للتوسيع ليستفيد من ميزات غير النصية مثل: ميزات السلاسل الزمنية، ميزات الصور، الفيديو، الصوت وذلك بحسب المهام المطلوبة. يمكن استخدام نماذج لتضمين اللغات مثل BERT والشبكات العصبونية الصورية GNN وشبكات الخصومة التوليدية GAN مع الانتباه واختبار تأثيرها على الأداء التنبؤي للنموذج. استخدام النموذج المقترح في بناء أنظمة توصية تعتمد على طرق الترتيب Ranking لتقديم أفضل التوصيات. يمكن تطبيق النموذج المقترح على أنواع مختلفة من المهام ومجموعات المعطيات وإجراء التجارب والمقارنات عليها.

جدول (14): اختصارات لأسماء النماذج والتوابع المستخدمة في البحث

الاختصار	المعنى الكامل
NCF	Neural Collaborative Filtering
GMF	Generalized Matrix Factorization
MLP	Multi-Layer Perceptron
MLP_all	MLP تستخدم جميع ميزات المعطيات .all features
NeuMF	Neural Matrix Factorization
XGBoost	eXtreme Gradient Boosting
ReLU	Rectified Linear Unit
Adam	Adaptive Moments
CNN	Convolutional Neural Network
BiLSTM	Bidirectional
RNN	Recurrent Neural Network
ConvBiLSTMAtt	CNN+BiLSTM+Attention
(TextModelName)_all	نموذج نصي يستخدم XGBoost لتكديس 4 نماذج من أجل كل طريقة تضمين كلمات مدربة مسبقاً بالنسبة لكل نموذج عصبوني نصي، مثال: CNN_all هو نموذج نصي يكس 4 نتائج معتمدة على كل من None, Word2vec, FastText, Glove.
NeuMF-XT	نموذج مقترح يتألف من نموذج NeuMF معدّل بطبقات شبكية ونموذج مكس XGBoost لربط نتائج GMF و MLP، بالإضافة لنموذج عميق نصي Text model، والأفضل كان نموذج ConvBiLSTMAtt_all للتعامل مع الميزات النصية.
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
NSE	Nash-Sutcliffe Model Efficiency

### المراجع References

- [1] Aggarwal CC, 2023 - "**Neural Networks and Deep Learning: A Textbook**". Cham: Springer International Publishingp.
- [2] Ba JL, Kiros JR, and Hinton GE, 2016. "**Layer Normalization**".
- [3] Bahdanau D, Cho K, and Bengio Y, 2014. "**Neural Machine Translation by Jointly Learning to Align and Translate**".
- [4] Chen T and Guestrin C, 2016. "**XGBoost: A Scalable Tree Boosting System**", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, pp. 785-794.
- [5] Datta L, 2020. "**A Survey on Activation Functions and their relation with Xavier and He Normal Initialization**", *ArXiv*, vol. abs/2004.06632.
- [6] Deeb M and Abu Saleh N, 2019. "**Performance Evaluation of Learning Models in Comparison with Recommendation Systems by Studying Data Features**", *Journal of Al-Baath University*, vol. 41, no. 51, pp. 11-52.
- [7] Feng S, Song K, Wang D, Gao W, and Zhang Y, 2021. "**InterSentiment: combining deep neural models on interaction and sentiment for review rating prediction**", *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 2, pp. 477-488.
- [8] Gers FA, Schmidhuber J, and Cummins F, 2000. "**Learning to forget: continual prediction with LSTM**", *Neural Comput*, vol. 12, no. 10, pp. 2451-71.
- [9] Gers FA, Schraudolph NN, and Schmidhuber J, 2003. "**Learning precise timing with lstm recurrent networks**", *J. Mach. Learn. Res.*, vol. 3, no. null, pp. 115-143.
- [10] Glorot X, Bordes A, and Bengio Y, 2011. "**Deep Sparse Rectifier Neural Networks**".

- [11] Goodfellow I, Bengio Y, and Courville A, 2016 - "***Deep learning***". Cambridge, Massachusetts: The MIT Press.
- [12] Graves A and Schmidhuber J, 2005. "**Frame-wise phoneme classification with bidirectional LSTM and other neural network architectures**", *Neural Networks*, vol. 18, no. 5, pp. 602-610.
- [13] Guo T, Wen Y, Wang F, and Hou J, 2019. "**Learning Resource Recommendation Based on Generalized Matrix Factorization and Long Short-Term Memory Model**", in *2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, pp. 217-222, Sydney, Australia: IEEE.
- [14] He X, Liao L, Zhang H, Nie L, Hu X, and Chua T-S, 2017. "**Neural Collaborative Filtering**", in *Proceedings of the 26th International Conference on World Wide Web*, pp. 173-182, Perth Australia: International World Wide Web Conferences Steering Committee.
- [15] Hochreiter S and Schmidhuber J, 1997. "**Long short-term memory**", *Neural Comput*, vol. 9, no. 8, pp. 1735-80.
- [16] Huang R, McIntyre S, Song M, E H, and Ou Z, 2018. "**An Attention-Based Recommender System to Predict Contextual Intent Based on Choice Histories across and within Sessions**", *Applied Sciences*, vol. 8, no. 12, p. 2426.
- [17] Hug N, 2020. "**Surprise: A Python library for recommender systems**", *Journal of Open Source Software*, vol. 5, no. 52.
- [18] Ioffe S and Szegedy C, 2015. "**Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift**".
- [19] James V and Alagarsamy S, 2022. "**RNN LSTM-based Deep Hybrid Learning Model for Text Classification using Machine Learning Variant XGBoost**", *International*



*Journal of Performability Engineering*, vol. 18, no. 8, pp. 545-551.

- [20] Joulin A, Grave E, Bojanowski P, and Mikolov T, 2017 - "**Bag of Tricks for Efficient Text Classification**". Valencia, Spain: Association for Computational Linguistics, 427-431p.
- [21] Kamyab M, Liu G, and Adjeisah M, 2021. "**Attention-Based CNN and Bi-LSTM Model Based on TF-IDF and GloVe Word Embedding for Sentiment Analysis**", *Applied Sciences*, vol. 11, no. 23, p. 11255.
- [22] Kim Y, 2014. "**Convolutional Neural Networks for Sentence Classification**", in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746-1751, Doha, Qatar: Association for Computational Linguistics.
- [23] Kingma DP and Ba J, 2014. "**Adam: A Method for Stochastic Optimization**".
- [24] Koren Y, Bell R, and Volinsky C, 2009. "**Matrix Factorization Techniques for Recommender Systems**", *Computer*, vol. 42, no. 8, pp. 30-37.
- [25] LeCun Y *et al.*, 1989. "**Backpropagation Applied to Handwritten Zip Code Recognition**", *Neural Computation*, vol. 1, no. 4, pp. 541-551.
- [26] Liu G and Guo J, 2019. "**Bidirectional LSTM with attention mechanism and convolutional layer for text classification**", *Neurocomputing*, vol. 337, pp. 325-338.
- [27] Low YH, Yap W-S, and Tee YK, 2019. "**Convolutional Neural Network-Based Collaborative Filtering for Recommendation Systems**", Singapore, pp. 117-131: Springer Singapore.
- [28] Luo X, Zhou M, Xia Y, and Zhu Q, 2014. "**An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems**",

*IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1273-1284.

- [29] Luong M-T, Pham H, and Manning CD, 2015. "**Effective Approaches to Attention-based Neural Machine Translation**".
- [30] Masters D and Luschi C, 2018. "**Revisiting Small Batch Training for Deep Neural Networks**".
- [31] Mikolov T, Sutskever I, Chen K, Corrado G, and Dean J, 2013. "**Distributed Representations of Words and Phrases and their Compositionality**".
- [32] Ni J, Li J, and McAuley J, 2019. "**Justifying Recommendations using Distantly-Labeled Reviews and Fine-Grained Aspects**", in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 188-197, Hong Kong, China: Association for Computational Linguistics.
- [33] Ong K, Ng K-W, and Haw S-C, 2021. "**Neural matrix factorization++ based recommendation system**", *FI1000Research*, vol. 10, p. 1079.
- [34] Pennington J, Socher R, and Manning C, 2014. "**Glove: Global Vectors for Word Representation**", in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532-1543, Doha, Qatar: Association for Computational Linguistics.
- [35] Roy PK, 2020. "**Multilayer Convolutional Neural Network to Filter Low Quality Content from Quora**", *Neural Processing Letters*, vol. 52, no. 1, pp. 805-821.
- [36] Rustam F, Mehmood A, Ahmad M, Ullah S, Khan DM, and Choi GS, 2020. "**Classification of Shopify App User Reviews Using Novel Multi Text Features**", *IEEE Access*, vol. 8, pp. 30234-30244.

- [37] Shan S, Ni H, Wang D, Chen G, and Sang X, "A Deep Learning Model with Attention-BiLSTM Networks Combining XGBoost Residual Correction for Short-Term Water Demand Forecast," In Review, preprint 2022, Accessed on: 2024-03-15 14:00:58.
- [38] Sharma R, Gopalani D, and Meena Y, 2017. "**Collaborative filtering-based recommender system: Approaches and research challenges**", in *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICCT)*, pp. 1-6, Ghaziabad, India: IEEE.
- [39] Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, and Salakhutdinov R, 2014. "**Dropout: a simple way to prevent neural networks from overfitting**", *J. Mach. Learn. Res.*, vol. 15, pp. 1929-1958.
- [40] Van Houdt G, Mosquera C, and Nápoles G, 2020. "**A review on the long short-term memory model**", *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929-5955.
- [41] Vaswani A *et al.*, 2017. "**Attention Is All You Need**".
- [42] Wang H *et al.*, 2021. "**A Stock Closing Price Prediction Model Based on CNN-BiSLSTM**", *Complexity*, vol. 2021, pp. 1-12.
- [43] Xiaoyan L, Raga RC, and Xuemei S, 2022. "**GloVe-CNN-BiLSTM Model for Sentiment Analysis on Text Reviews**", *Journal of Sensors*, vol. 2022, pp. 1-12.
- [44] Yao X, 2018. "**Attention-based BiLSTM Neural Networks for Sentiment Classification of Short Texts**", in *Proceedings of Information Science and Cloud Computing — PoS(ISCC 2017)*, p. 014, Guangzhou, China: Sissa Medialab.
- [45] Zheng J and Zheng L, 2019. "**A Hybrid Bidirectional Recurrent Convolutional Neural Network Attention-Based Model for Text Classification**", *IEEE Access*, vol. 7, pp. 106673-106685.

- [46] Zhou P *et al.*, 2016. "**Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification**", in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 207-212, Berlin, Germany: Association for Computational Linguistics.
- [47] Ziegler C-N, McNee SM, Konstan JA, and Lausen G, 2005. "**Improving recommendation lists through topic diversification**", in *Proceedings of the 14th international conference on World Wide Web - WWW '05*, p. 22, Chiba, Japan: ACM Press.