Journal of Basic Sciences (JBS)
مجلة العلوم الأساسية
Volume 38, Issue 1, January – March 2025
Page no.: 92- 104
Email: jbs@journals.asmarya.edu.ly Website: journals.asmarya.edu.ly/jbs

استخدام نموذج BERT لحل مشكلة البداية الباردة في أنظمة التوصية

إيمان خليفة علي خليفة *، آمال عبدالسلام ارحومة الجائر، أسماء علي أبودينة

قسم علوم الحاسوب ، كلية تقنية المعلومات ، الجامعة الأسمرية الإسلامية ، زليتن ، ليبيا

*** لمراسلة المؤلف:** em.khalifa@asmarya.edu.ly

استلمت: 20 فبراير 2025	قبلت: 03 مارس 2025	نشرت: 24 مارس 2025

الملخص:

تعد أنظمة التوصية **(Recommendation Systems)** من التقنيات المهمة في العديد من التطبيقات الحديثة مثل التجارة الإلكترونية، شبكات التواصل الاجتماعي، ومنصات البث الرقمية مثل نتفليكس **(Netflix)** ويوتيوب (YouTupe)، حيث تقوم بتقديم اقتراحات مخصصة للمستخدمين بناءً على بيانات تفاعلهم السابقة أو خصائص المحتوى الذي يتفاعلون معه.ومن أبرز التحديات والمشاكل التي تواجه هذه الأنظمة هي مشكلة البداية الباردة (Cold Start Problem)، وهي الحالة التي يصعب فيها تقديم توصيات دقيقة عندما تكون بيانات تفاعل المستخدمين أو العناصر جديدة أو محدودة. هذه المشكلة تصبح أكثر تعقيدًا في الحالات التي يتعامل فيها النظام مع المستخدمين الجدد أو العناصر غير المعروفة.

لحل هذه المشكلة تم استخدام نموذج التعلم العميق BERT لتحسين قدرة النظام على تقديم توصيات دقيقة للمستخدمين بناء على خصائصهم النصية بالإضافة إلى الاستفادة من بيانات المستخدم الديموغرافية كالعمر ، الجنس ، لتحقيق حلول فعّالة في تحسين أداء هذه الأنظمة، وقد تم الوصول إلى دقة قدرها 96% في تصنيف النصوص، مما يبرز قدرة النموذج على التعميم بشكل جيد على بيانات جديدة. كما اثبت نموذج التعلم العميق bert كفاءته في هذه الدراسة مقارنة بنموذج الشبكة العصبية Neural Collaborative Filtering (NCF)

الكلمات المفتاحية: نموذج BERT، مشكلة البداية الباردة (Cold Start Problem)، الفلترة التعاونية العصبية .(NCF) .

The Cold Start Problem Solution Using BERT model in Recommendation Systems

Ayiman Khalleefah Ali Khalleefah * , Amal Abdalslam Arhoma Aljaer and Asma Ali Abodina

Computer Science Department, Faculty of Information Technology,
Alasmarya Islamic University, Zliten, Libya

*** Corresponding author: em.khalifa@asmarya.edu.ly**

Received: Feb. 20, 2025	Accepted: Mar. 03, 2025	Published: Mar. 24, 2025

Abstract:
Recommendation systems are crucial technologies in many modern applications such as e-commerce, social networks, and digital streaming platforms like Netflix and YouTube. These systems provide personalized suggestions to users based on their past interaction data or the characteristics of the content they engage with. One of the most significant challenges facing these systems is the Cold Start Problem, which occurs when it is difficult to provide accurate recommendations due to limited or new user interaction data or item information. This issue becomes more complex when dealing with new users or unknown items.

To address this problem, the BERT deep learning model was utilized to enhance the system's ability to provide accurate recommendations based on textual features of users, in addition to leveraging demographic data such as age and gender. This approach has proven effective in improving the performance of these systems, achieving a text classification accuracy of up to 96%, demonstrating the model's strong generalization capabilities on new data. Furthermore, the BERT model showed superior efficiency in this study compared to the Neural Collaborative Filtering (NCF) model.

Keywords: BERT Model, Cold Start Problem, Neural Collaborative Filtering (NCF).

1. المقدمة Introduction

تلعب أنظمة التوصية دورًا أساسيًا في العديد من التطبيقات عبر الإنترنت والخدمات التجارية الإلكترونية، مثل شبكات التواصل الاجتماعي وتوصيات المنتجات كالأفلام والموسيقى والمقالات [1]. وقد قامت شركات كبرى، مثل Amazon وeBay وNetflix، بدمج تقنيات التوصية في أنظمتها بهدف توقع تفضيلات العملاء المحتملة وتقديم توصيات للمنتجات أو العناصر التي تتوافق مع اهتماماتهم. ويؤثر أداء أنظمة التوصية بشكل كبير على النجاح التجاري لهذه الشركات من حيث زيادة الإيرادات ورفع مستوى رضا المستخدم.

يمكن تصنيف أساليب التوصية وفقًا لنوع البيانات التي يتم جمعها وطرق استخدامها في أنظمة التوصية إلى تصفية قائمة على المحتوى Content-Based Filtering (CB) وتصفية تعاونية Collaborative Filtering(CF) ونماذج هجينة Hybrid Models [2].

تُستخدم التصفية القائمة على المحتوى (CB) على نطاق واسع في تصميم أنظمة التوصية، حيث تعتمد على خصائص ومميزات محتوى العناصر لمطابقتها مع ملفات تعريف المستخدمين. يتم مقارنة العناصر مع تلك التي نالت إعجاب المستخدمين سابقًا، ومن ثم يُوصى بالعناصر التي تتطابق

بشكل أفضل. إحدى المشكلات الرئيسية في هذا النهج هي أن نظام التوصية يحتاج إلى تعلم تفضيلات المستخدمين لبعض أنواع العناصر وتطبيق هذه التفضيلات على أنواع أخرى.

يُعد نهج التصفية التعاونية (CF) الأكثر شيوعًا في تصميم أنظمة التوصية، حيث يستفيد من كميات كبيرة من البيانات المجمعة من سلوك المستخدمين في الماضي لتوقع العناصر التي قد تعجبهم. ولا يحتاج هذا النهج إلى تحليل محتوى العناصر، بل يعتمد على العلاقة بين المستخدمين والعناصر، والتي يتم تمثيلها عادةً في مصفوفة تقييم، حيث يمثل كل عنصر في المصفوفة تقييمًا معينًا لمستخدم محدد على عنصر معين.

النهج الهجين هو نهج يجمع بين التصفية القائمة على المحتوى (CB) والتصفية التعاونية (CF) بهدف التغلب على أوجه القصور في كل منهما وتقديم نتائج أكثر كفاءة.

ومع أن هذه الأنظمة قد حققت نجاحًا كبيرًا في تخصيص التجربة للمستخدمين، إلا أن هناك العديد من التحديات التي تواجهها. من بين هذه التحديات مشكلة البداية الباردة (Cold Start Problem)، وهي الحالة التي يعجز فيها النظام عن تقديم توصيات دقيقة عندما تكون البيانات المتاحة للمستخدمين أو العناصر الجديدة محدودة أو غائبة. في هذه الحالة، يعجز النظام عن الاستفادة من تفاعلات المستخدمين مع العناصر السابقة لأن البيانات غير كافية، مما يؤدي إلى صعوبة كبيرة في تقديم توصيات موثوقة. هذه المشكلة تصبح أكثر تعقيدًا عندما يتعلق الأمر بالمستخدمين الجدد الذين لم يكن لديهم تفاعلات كافية، أو العناصر الجديدة التي لم يحصل على تفاعل كافٍ حولها.

لحل هذه المشكلة تم استخدام التعلم العميق باستخدام نموذج تصنيف نصي BERT في تحسين دقة التوصيات، تم فيه دمج بيانات المستخدمين والأفلام والتقييمات في وصف نصي لكل تقييم، ثم تُصنّف التقييمات إلى إيجابية (4 أو أكثر) أو سلبية (أقل من 4) .بعد ذلك، تُقسّم البيانات إلى تدريب واختبار، وتُجهّز باستخدام مكتبة ktrain بتنسيق يدعم BERT لتحليل النصوص وبناء نموذج تصنيف فعّال.

تهدف هذه الورقة البحثية إلى دراسة كيف يمكن للتعلم العميق أن يساعد أنظمة التوصية في التغلب على مشكلة البداية الباردة.

2. الدراسات السابقة (Literature Review)

هناك العديد من الأبحاث التي تناولت مشكلة التوصيات في سياق البداية الباردة (Cold-start Recommendations) وأساليب تحسين أداء النماذج في هذا السياق. نستعرض فيما يلي بعض الدراسات:

قدم Luo وزملاؤه [3] خوارزمية مبتكرة تُدعى Popularity-Aware Meta-Learning (PAM). تركز هذه الخوارزمية على تقسيم المهام بناءً على مستويات معدل تفاعل العناصر باستخدام إطار التعلم فوق المعرفي (Meta-Learning). تهدف إلى تحسين توصيات العناصر ذات المعدل الأقل تفاعلا في بيئات البيانات المتدفقة مع الحفاظ على كفاءة الحوسبة. وأظهرت الدراسة نتائج متميزة على عدة مجموعات بيانات عامة وسيناريوهات حقيقية، مما يُبرز فعالية الخوارزمية مقارنة بالطرق التقليدية. ركز الباحثين في هذه الدراسة على بيانات التفاعلات بين المستخدمين والعناصر

وفي دراسة أخرى اقترح He واخرون [4] نموذج يُدعى GAE (Graph Attention neural networks with meta-learning) الذي يجمع بين شبكات الانتباه الرسومية (Graph Attention Neural Networks) والتعلم الفوقي (Meta-Learning) لتوليد توصيات للمستخدمين الجدد.

يتميز هذا النموذج باستخدامه لشبكات الانتباه الرسومية (GAT) لتجميع معلومات السمات من العقد المجاورة للمستخدمين الجدد، ثم استخدام مولد التضمين (Embedding Generator) لدمج هذه المعلومات مع سمات المستخدم نفسه. بعد ذلك، يتم استخدام التعلم الفوقي لتوليد قائمة بالتوصيات للمستخدمين الجدد.

نموذج GAE يجمع بين تقنيات متقدمة مثل شبكات الانتباه الرسومية (GAT) والتعلم الفوقي (Meta-Learning)، مما يجعله معقدًا من حيث التطبيق والتنفيذ.

كما انه قد يؤدي إلى زيادة في تكاليف الحوسبة والوقت المطلوب للتدريب، خاصة في الأنظمة الكبيرة التي تحتوي على عدد كبير من المستخدمين والعناصر.

كما تناول Wu واخرون [5] مشكلة البداية الباردة في أنظمة التوصيات من خلال تقديم نموذج PromptRec الذي يحول مهمة التوصيات إلى مهمة تحليل مشاعر(Sentiment Analysis) باستخدام قوالب لغوية (Prompt Templates)، مع تعزيز النماذج الصغيرة (التي تحتوي على

معلمات اقل) عبر منهجيات (Data-Centric Approaches) والتي تركز على تحسين جودة البيانات وتمثيلها بدلًا من تحسين النموذج نفسه.

أظهرت نتائج الدراسة أن PromptRecيمكن أن يحقق أداءً تنافسيًا مقارنة بالنماذج الكبيرة، مع تقليل كبير في تكلفة الحسابات (Computational Cost)وزمن الاستجابة .(Response Time) هذا يجعل النماذج الصغيرة خيارًا عمليًا وفعّالًا للتطبيقات الواقعية، خاصة في البيئات التي تتطلب سرعة في الاستجابة وكفاءة في استخدام الموارد. ومع ذلك، لا يزال هناك تأخير زمني (Latency) مرتبط باستخدام النماذج الكبيرة للغة(Large Language Models - LLMs) ، مما يشكل تحديًا في التطبيقات التي تتطلب معالجة في الوقت الفعلي.(Real-Time Processing)

كما أقترح Lin واخرون في دراسة [6] ، نموذج يسمى (Temporal Dynamic Risk Optimization)TDROيعمل على تحسين الأداء في أسوأ الحالات من خلال التركيز على المجموعات(مستخدمين جدد – عناصر جديدة) الأقل أداءً، بالإضافة إلى التقاط الاتجاهات الزمنية لتحديد المجموعات التي قد تصبح شائعة في المستقبل. ومع ذلك، تواجه TDRO بعض القيود، مثل الاعتماد على تقسيم المجموعات بشكل مسبق، مما قد لا يعكس دائمًا التغيرات الفعلية في تفضيلات المستخدمين. بالإضافة إلى ذلك، يمكن أن يكون تحسين الأداء في أسوأ الحالات على حساب أداء المجموعات الأخرى. لتحسين هذه الطريقة، أقترح دمج تقنيات التعلم الذاتي (Self-Supervised Learning)لتعزيز تمثيلات العناصر الباردة دون الاعتماد على بيانات تفاعل إضافية، واستخدام نماذج اللغة الكبيرة (LLMs) لتحسين دقة التوصيات الباردة.

كما تناول Li واخرون [7] مشكلة البداية الباردة في أنظمة التوصية من خلال دمج نوعين من المعرفة المبدئية: المعرفة الخاصة بالتوصية المستخلصة باستخدام تقنيات التعلم المعدني-(Meta Learning)والمعرفة العامة المستخلصة من نماذج اللغة الكبيرة (LLMs) ، تم ترميز المعلومات النصية حول المستخدمين والعناصر باستخدام نماذج LLMs لتعزيز قدرة النظام على فهم التفضيلات والتوصية بالعناصر للمستخدمين الجدد. أظهرت النتائج أن دمج هذين النوعين من المعرفة يحسن أداء النظام في مواقف البداية الباردة، مما يعزز دقة التوصيات للمستخدمين أو العناصر الذين يفتقرون إلى تاريخ تفاعل كافٍ وقد أظهرت النتائج أن النماذج الكبيرة قادرة على تحقيق دقة عالية، ولكن بتكلفة حسابية كبيرة وزمن استجابة طويل.

وفي دراسة أخرى اقترح الباحث Zheng واخرون [8] واخرون طريقة نهج التعيين التباييني (VM-Rec: Variational Mapping for Cold-Start User Recommendation)، الذي

يستند إلى فكرة أن المستخدمين الحاليين ذوي التفاعلات الكافية كانوا في الأصل مستخدمين جددًا، وبذلك يمكن محاكاة عملية انتقالهم من مرحلة قلة البيانات إلى مرحلة امتلاك تمثيلات غنية. يعتمد هذا النهج على إنشاء دالة تعيين مخصصة للمستخدمين الجدد، يتم توليد معالمها باستخدام توزيع تبايني احتمالي (Variational Distribution) لضمان تعبيرية عالية حتى مع البيانات القليلة المتاحة. ومن أجل تحسين الفعالية الحسابية والتفسيرية، تم تصميم دالة التعيين على شكل نموذج خطي متفرق(Sparse Linear Model) ، بحيث تمثل كل معلمة فيه ارتباط المستخدم الجديد بمستخدمين آخرين لديهم تفاعلات أكثر. بعد ذلك، يتم استخدام هذه الدالة لنقل تمثيلات المستخدمين الحاليين إلى فضاء المستخدم الجديد، مما يُنتج تمثيلًا ذا تعبيرية مشابهة لهم.

تم تقييم النموذج على مجموعتي بيانات LastFM :و Ali Display Ad Click، باستخدام ثلاثة نماذج توصية أساسية.(BPR, NGCF, LightGCN) كأنظمة أساسية، حيث تفوق نهج VM-Rec على أربعة طرق شائعة لحل مشكلة البداية الباردة في مجموعتي بيانات مختلفتين .أظهرت النتائج أن النهج المقترح قادر على تحسين دقة التوصية للمستخدمين الجدد مقارنةً بالأساليب التقليدية، مما يجعله حلاً واعدًا لمشكلة البداية الباردة في أنظمة التوصية.

و في دراسة أخرى قدمها Hao وأخرون [9] ، تم فيها تطوير نموذج MPT (Multi-strategy based Pre-Training) لتحسين توصيات البداية الباردة. يعتمد النموذج على دمج GNN للعلاقات قصيرة المدى و Transformer للعلاقات طويلة المدى، مع إضافة مهمة التعلم التباينيلتعزيز التمثيلات. تم اختبار النموذج على مجموعات بيانات مثل MovieLens-1M و MOOCs وGowalla، حيث أظهر تفوقًا في استنتاج التضمينات وأداء التوصية. ومع ذلك، يواجه النموذج تحديات مثل التعقيد الحسابي واعتماده على التفاعلات الأولية. بالإضافة إلى ذلك، لا يستخدم النموذج المعلومات الجانبية مثل الرسوم البيانية المعرفية، مما قد يحد من فعاليته في حالات ندرة البيانات. يحتاج النموذج أيضًا إلى ضبط دقيق للمعاملات، وقد لا يكون فعالًا في حالات التفاعلات الشديدة الندرة. هذه القيود تشير إلى الحاجة لمزيد من التحسينات في الأبحاث المستقبلية.

قدمت Volkovs وأخرون [10] نموذج DropoutNet، وهو نموذج كامن يعتمد على الشبكات العصبية العميقة لحل مشكلة البداية الباردة في أنظمة التوصية. على عكس الأساليب التقليدية التي تعتمد على تضمين ميزات إضافية، يركز النموذج على تحسين عملية التدريب باستخدام الإسقاط العشوائي (Dropout) لتعزيز التعميم. يمكن تطبيقه على أي نموذج كامن (Latent

Representation) موجود مسبقًا، مما يجعله حلاً مرنًا. أظهرت التجارب على بيانات عامة تفوق النموذج على الأساليب التقليدية، مما يؤكد فعالية التعلم العميق في هذا السياق.

لكنه يأتي بتكلفة حسابية أعلى وحاجة لضبط دقيق للمعلمات لتحقيق أقصى استفادة.

وفي دراسة اجراها Leung [11] واخرون حيث قدموا نموذج تفضيل يعتمد على علاقات المستخدم-العنصر والعنصر-العنصر، وقدمت خوارزمية CLAREلتوليد توصيات البداية الباردة باستخدام خصائص العناصر. عندما لا تكون بيانات التقييمات كافية لتوليد توصيات، تعتمد CLAREعلى خصائص العناصر لتقديم توصيات فعالة. أظهرت النتائج التجريبية أن CLARE قادرة على تحسين جودة التوصيات وزيادة عدد العناصر القابلة للتوصية بشكل كبير. بالإضافة إلى ذلك، تم اختبار الخوارزمية باستخدام أنواع مختلفة من الخصائص، مما أدى إلى تحسين أداء التوصيات وتوسيع نطاق التغطية. ركزت الدراسة على التحقق من قدرة CLARE على التعامل مع مشكلة البداية الباردة، مع إجراء تحليل أداء شامل، وسيتم تقديم نتائج أكثر تفصيلاً ومقارنات مع أعمال ذات صلة في أبحاث لاحقة.

3. المنهجية (Methodology)

تهدف هذه الدراسة إلى توقع تفضيلات المستخدمين في حالة مشكلة البداية الباردة (Cold-Start Problem)باستخدام نموذج BERT (Bidirectional Encoder Representations from Transformers). يتم ذلك عبر الاستفادة من قدرات BERT في فهم النصوص وتحليل أوصاف الأفلام وتفضيلات المستخدمين بشكل دقيق.

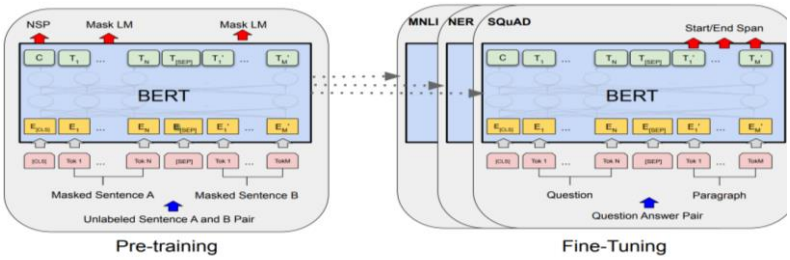
نموذج BERT من اشهر نماذج اللغة تم تطويره بواسطة Google من قبل الباحثين جاكوب ديفلين Jacob Devlin ، مينغ-وي تشانغ، كينتون لي Ming-Wei Chang ، وكريستينا توتانونا Kristina Toutanova ، يعتمد على بنية (Transformer) وجاءت هذه الفكرة من خلال الورقة البحثية التي قدمت قي سنة 2017 بعنوان (Attention Is All You Need)هذه البنية تعتمد على خاصية Attention وهي السماح للنموذج بالتركيز على أجزاء معينة من النص لفهم السياق بشكل افضل [12].

نموذج BERT يتميز بقدرته على فهم النصوص بشكل ثنائي الاتجاه (Bidirectional) من خلال السياق الكامل حيث يتم معالجة النص في الاتجاهين (قبل وبعد الكلمة المستهدفة)، مما يجعله قويًا في فهم المعاني الدقيقة للنصوص مقارنة بالنماذج أحادية الاتجاه،تم تدريب BERTعلى مجموعة

بيانات ضخمة تتضمن ويكيبيديا وBookCorpus، وهي قاعدة بيانات تحتوي على أكثر من 10,000كتاب من مختلف الأنواع الأدبية.

الهيكلية البنية العامة لنموذج BERT:

الشكل [1] يوضح البنية العامة لتدريب وضبط نموذج BERT، وفقًا لما قدمه Devlin واخرون (2019)



الشكل 1تمثيل معماري لنموذج BERTيوضح مرحلتي التدريب المسبق والضبط الدقيق.
المصدر.[13]

من الشكل 1 نوضح عمل المرحلتين الآتيتين:

مرحلة (Pre-training) وهي مرحلة التدريب المسبق، يتم تدريب النموذج باستخدام مهمتين رئيسيتين:

1. Masked Language Model (MLM) يتم استبدال بعض الكلمات بـ [MASK]، ويطلب من النموذج توقعها.

2. Next Sentence Prediction (NSP) يحدد النموذج ما إذا كان الجملتان المدخلتان متتابعتين أم لا.

أما في مرحلة (Fine-Tuning) تبين كيف يمكن تكييف BERT لمهام مختلفة مثل تصنيف النصوص (MNLI, NER) أو إجابة الأسئلة (SQuAD) عبر ضبط الطبقات العليا على بيانات جديدة.

طبقات نموذج BERT:

1- طبقة التضمين Embedding Layer

وهي الطبقة السفلى وهي تمثل عملية تحويل الكلمات المدخلة إلى تمثيلات رقمية (Embeddings)، حيث يتم تحويل الكلمات إلى متجهات.(Vectors) و يتم استخدام طبقات خاصة مثل Positional Embedding لإضافة معلومات عن موقع الكلمات في النص.

2- طبقة Transformer

تتكون من عدة طبقات من المحولات (Transformers) التي تعتمد على آلية الاهتمام الذاتي (Self-Attention). الهدف منها هو تحليل العلاقات بين الكلمات في السياق الكامل للنص، بحيث يمكن للنموذج فهم الكلمة بناءً على الكلمات التي قبلها وبعدها.

3- طبقة [MASK] Token

يُستخدم الرمز "[MASK]" كإدخال خاص في مرحلة تدريب النموذج ،حيث تُزال بعض الكلمات من النص وتُستبدل بالرمز"[MASK]" ، ويُطلب من النموذج التنبؤ بالكلمة الأصلية وإعادتها إلى مكانها. تهدف هذه الطريقة إلى تحسين قدرة النموذج على تعلم التمثيلات السياقية للكلمات.

4- طبقة Output Layer

تُعتبر الطبقة العليا (Classification Layer) المسؤولة عن معالجة المخرجات الصادرة من طبقة المحول .(Transformer Layer) تعتمد هذه الطبقة على شبكة عصبية متصلة بالكامل (Fully-Connected Layer)لإعطاء النتائج النهائية. وتشمل هذه النتائج مجموعة من المهام المختلفة، مثل تصنيف النصوص، التنبؤ بالكلمات المفقودة، أو تحليل وفهم العبارات.

آلية العمل المقترح وبيئة العمل المستخدمة:

وفي هذا الورقة تم استخدام مجموعة من الخصائص لتدريب النموذج ومعالجة مشكلة البداية الباردة في أنظمة التوصية، حيث شملت هذه الخصائص معلومات عن المستخدمين (العمر، الجنس). تقييمات الأفلام من قِبل المستخدمين. وأوصاف الأفلام (العناوين، الأنواع، التقييمات) .وبدمج هذه الخصائص يساعد النموذج على فهم العلاقات بين المستخدمين والعناصر، كما أن هذه الخصائص تُستخدم كمدخلات لنموذج BERT، الذي يتيح استغلال قدراته في التمثيل السياقي للنصوص لتحسين دقة التوصيات في مشكلة البداية الباردة (Cold Start Problem)،

وتم استخدام بناء نموذج باستخدام Neural Collaborative Filtering (NCF) لنفس البيانات لمقارنته مع نموذج BERT

نموذج Neural Collaborative Filtering (NCF) تم فيه استخدام الشبكات العصبية العميقة بدلًا من تفكيك المصفوفات التقليدي، حيث تم تدريب النموذج لاستخراج الميزات الكامنة تلقائيًا بدلاً من فرض قيود رياضية مثل التعامد و تمرير الميزات إلى طبقات غير خطية، مما يسمح للنموذج بالتقاط العلاقات المعقدة بين المستخدمين والأفلام.

4. إعداد البيانات في نموذج BERT (Data Preprocessing)

في هذا العمل تم استخدام برنامج Python على السحابة google colab ، وقد تم استخدام قاعدة البيانات "MovieLens" والتي تحتوي على 1000209 سجل ، حيث تحتوي على ملف بيانات المستخدمين، الأفلام، والتقييمات، أخذ منها عينة عشوائية عددها 20000 سجل ،تم حذف عمودي "Zip-code" و "Timestamp"من بيانات المستخدمين والتقييمات، ثم تم دمج هذه البيانات في إطار بيانات واحد باستخدام دالة merge بناءً على العمود المشترك "UserID" و "MovieID".بذلك يتم تكوين إطار بيانات يحتوي على معلومات المستخدمين، الأفلام، والتقييمات.

وتمت تنقية النصوص من الرموز غير الضرورية و تحويلها إلى (Tokenization) باستخدام Tokenizerالخاص بـ BERT، وذلك من خلال وظيفة texts_from_array من مكتبة ktrain.

وبعد ذلك تم توحيد النصوص عن طريق (Padding & Truncation) إلى 128رمزًا.

ثم تم انشاء النموذج لتصنيف النصوص باستخدام تقنية BERT وتدريبه على البيانات التدريبية (x_train, y_train) مع معالجة مسبقة محددة بـ preproc.

تم اخذ بيانات تدريب 70% ، واختبار 30 %.

تم إنشاء كائن learner لإدارة عملية تدريب النموذج باستخدام مكتبة ktrain، حيث يتم تمرير بيانات التدريب والاختبار مع تحديد حجم الدفعة (batch_size=6). بعد ذلك، يتم إعداد كائن Callbackباستخدام ModelCheckpoint لحفظ أوزان النموذج بعد كل دورة تدريبية (Epoch)في المسار المحدد. أخيرًا، يبدأ تدريب النموذج باستخدام الطريقة fit_onecycleمع تحديد معدل التعلم وعدد الدورات التدريبية المطلوبة.

ثم تم استخدام النموذج لتوقع المشاعر لمستخدمين جدد بحيث عندما يسجل مستخدم جديد تقييمًا لفيلم معين، يمكن للنموذج التنبؤ بموقفه بناءً على بياناته الشخصية ومحتوى الفيلم.

5. النتائج والمناقشة (Results & Discussion)

بعد انجاز هذه الدراسة أثبتت النتائج التي تم الحصول عليها أن دقة النتائج في حالة استخدام نموذج BERT أفضل من النتائج المتحصل عليها باستخدام نموذج **NCF** وذلك كما موضح بالجدول التالي

TABLE I. THIS IS A SIMPLE EXAMPLE

accuarcy	RMSE	النموذج
0.96	0.006	BERT
0.74	0.8616	NCF

وتم ادخال بيانات وصفية للنموذج مثل :

predictor.predict(["User 102 (F, Age 25) rated movie 'Titanic' with 5 stars. Genre: Romance|Drama"])

predictor.predict(["User 501 (M, Age 35) rated movie 'Cats' with 2 stars. Genre: Musical|Fantasy"])

وقد تنبأ بالاتي :

positive

negative

هذه النتائج تشير إلى قدرة نموذج BERT على تحديد التصنيفات الصحيحة بناءً على محتوى النصوص، مما يعزز استخدامه في تحسين التوصيات في أنظمة البداية الباردة، حيث يمكن استغلال هذه التنبؤات لتحليل تفضيلات المستخدمين الجدد".

6. الخلاصة (Conclusion)

تُظهر نتائج التجربة التي أُجريت في هذه الدراسة تفوق نموذج BERT على الطرق التقليدية في حل مشكلة البداية الباردة في أنظمة التوصية. فقد حقق دقة تصل إلى 96 %في تصنيف النصوص بعد فترة تدريب واحدة فقط، مما يعكس قدرته العالية على التعميم والتكيف مع بيانات جديدة .تؤكد هذه

النتائج أهمية BERT في تحسين دقة التوصيات، خاصة في الأنظمة التي تواجه تحديات نقص بيانات المستخدمين أو العناصر الجديدة، مما يجعله خيارًا واعدًا لتطوير أنظمة توصية أكثر كفاءة.

نوصي باستكشاف تقنيات دمج BERT مع نماذج توليدية مثل GPT لتعزيز فهم السياق في التوصيات .بالإضافة إلى ذلك، نقترح اختبار نموذج DistilBERT ومقارنته بـ BERTمن حيث الدقة، سرعة الاستدلال، وكفاءة استخدام الموارد، مما قد يسهم في تطوير حلول أكثر كفاءة وملاءمة للتطبيقات العملية.

المراجع (References)

[1] Y. Shi, M. Larson, and A. Hanjalic, "Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges," ACM Comput. Surv., vol. 47, no. 1, pp. 1–45, Jan. 2014.

[2] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," Computer, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[3] Y. Luo, Y. Jiang, Y. Jiang, G. Chen, J. Wang, K. Bian, ... and Q. Zhang, "Online Item Cold-Start Recommendation with Popularity-Aware Meta-Learning," arXiv preprint arXiv:2411.11225, 2024.

[4] B. He, J. Zhou, R. Zhao, L. Wang, L. Li, and Y. Yang, "Cold-Start Product Recommendation Method Based on GAE," Journal of Computing and Electronic Information Management, vol. 13, no. 3, pp. 8–15, 2024.

[5] X. Wu, H. Zhou, Y. Shi, W. Yao, X. Huang, and N. Liu, "Could Small Language Models Serve as Recommenders? Towards Data-centric Cold-start Recommendations," arXiv preprint arXiv:2306.17256, 2024.

[6] X. Lin, W. Wang, J. Zhao, Y. Li, F. Feng, and T.-S. Chua, "Temporally and Distributionally Robust Optimization for Cold-Start Recommendation," in Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI-24), 2024.

[7] Y. Li, Y. Liu, and T. Furukawa, "Integrating Prior Knowledge from Meta-Learning and Large Language Models for Cold-Start Recommendation," in Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES), 2023, pp. 327-333, doi: 10.5109/7157997.

[8] L. Zheng, J. Chen, P. Liu, G. Zhang, and J. Fang, "VM-Rec: A Variational Mapping Approach for Cold-start User Recommendation," arXiv preprint arXiv:2307.04709, 2023.

[9] B. Hao, H. Yin, J. Zhang, C. Li, and H. Chen, "A Multi-Strategy based Pre-Training Method for Cold-Start Recommendation," Proc. ACM Meas. Anal. Comput. Syst., vol. 37, no. 4, pp. 1–24, May 2022, doi: 10.1145/1122445.1122456.

[10] M. Volkovs, G. Yu, and T. Poutanen, "DropoutNet: Addressing Cold Start in Recommender Systems," Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 2017, pp. 4003–4013.

[11] C.W. Leung, S.C. Chan and K.K.F. Chung. Applying Cross-Level Association Rule Mining to Cold-Start Recommendations, Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence and International Conference on Intelligent Agent Technology, 2007, Silicon Valley, CA, USA, pag: 133-136

[12] A. Vaswani, “Attention is all you need,” in Advances in Neural Information Processing Systems, 2017.

[13] J. D. M. W. C. Kenton and L. K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proceedings of NAACL-HLT, vol. 1, no. 2, Jun. 2019.