

مجلة جامعة البعث

سلسلة العلوم الهندسية الميكانيكية
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 45 . العدد 18

1444 هـ - 2023 م

الأستاذ الدكتور عبد الباسط الخطيب
رئيس جامعة البعث
المدير المسؤول عن المجلة

رئيس هيئة التحرير	أ. د. محمود حديد
رئيس التحرير	أ. د. درغام سلوم

مديرة مكتب مجلة جامعة البعث
بشرى مصطفى

عضو هيئة التحرير	د. محمد هلال
عضو هيئة التحرير	د. فهد شريباتي
عضو هيئة التحرير	د. معن سلامة
عضو هيئة التحرير	د. جمال العلي
عضو هيئة التحرير	د. عباد كاسوحة
عضو هيئة التحرير	د. محمود عامر
عضو هيئة التحرير	د. أحمد الحسن
عضو هيئة التحرير	د. سونيا عطية
عضو هيئة التحرير	د. ريم ديب
عضو هيئة التحرير	د. حسن مشرقي
عضو هيئة التحرير	د. هيثم حسن
عضو هيئة التحرير	د. نزار عبشي

تهدف المجلة إلى نشر البحوث العلمية الأصيلة، ويمكن للراغبين في طلبها

الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة البعث

سورية . حمص . جامعة البعث . الإدارة المركزية . ص . ب (77)

. هاتف / فاكس : 2138071 31 963 ++

. موقع الإنترنت : www.albaath-univ.edu.sy

. البريد الإلكتروني : [magazine@ albaath-univ.edu.sy](mailto:magazine@albaath-univ.edu.sy)

ISSN: 1022-467X

شروط النشر في مجلة جامعة البعث

الأوراق المطلوبة:

- 2 نسخة ورقية من البحث بدون اسم الباحث / الكلية / الجامعة) + CD / word من البحث منسق حسب شروط المجلة.
 - طابع بحث علمي + طابع نقابة معلمين.
 - إذا كان الباحث طالب دراسات عليا:
يجب إرفاق قرار تسجيل الدكتوراه / ماجستير + كتاب من الدكتور المشرف بموافقة على النشر في المجلة.
 - إذا كان الباحث عضو هيئة تدريسية:
يجب إرفاق قرار المجلس المختص بإنجاز البحث أو قرار قسم بالموافقة على اعتماده حسب الحال.
 - إذا كان الباحث عضو هيئة تدريسية من خارج جامعة البعث :
يجب إحضار كتاب من عمادة كليته تثبت أنه عضو بالهيئة التدريسية و على رأس عمله حتى تاريخه.
 - إذا كان الباحث عضواً في الهيئة الفنية :
يجب إرفاق كتاب يحدد فيه مكان و زمان إجراء البحث ، وما يثبت صفته وأنه على رأس عمله.
 - يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (العلوم الطبية والهندسية والأساسية والتطبيقية):
عنوان البحث .. ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).
- 1- مقدمة
 - 2- هدف البحث
 - 3- مواد وطرق البحث
 - 4- النتائج ومناقشتها .
 - 5- الاستنتاجات والتوصيات .
 - 6- المراجع.

- يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (الآداب - الاقتصاد - التربية - الحقوق - السياحة - التربية الموسيقية وجميع العلوم الإنسانية):
- عنوان البحث .. ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).
- 1. مقدمة.
- 2. مشكلة البحث وأهميته والجديد فيه.
- 3. أهداف البحث و أسئلته.
- 4. فرضيات البحث و حدوده.
- 5. مصطلحات البحث و تعريفاته الإجرائية.
- 6. الإطار النظري و الدراسات السابقة.
- 7. منهج البحث و إجراءاته.
- 8. عرض البحث و المناقشة والتحليل
- 9. نتائج البحث.
- 10. مقترحات البحث إن وجدت.
- 11. قائمة المصادر والمراجع.
- 7- يجب اعتماد الإعدادات الآتية أثناء طباعة البحث على الكمبيوتر:
 - أ- قياس الورق 25×17.5 B5.
 - ب- هوامش الصفحة: أعلى 2.54- أسفل 2.54 - يمين 2.5- يسار 2.5 سم
 - ت- رأس الصفحة 1.6 / تذييل الصفحة 1.8
 - ث- نوع الخط وقياسه: العنوان . Monotype Koufi قياس 20
- . كتابة النص Simplified Arabic قياس 13 عادي . العناوين الفرعية Simplified Arabic قياس 13 عريض.
- ج . يجب مراعاة أن يكون قياس الصور والجداول المدرجة في البحث لا يتعدى 12سم.
- 8- في حال عدم إجراء البحث وفقاً لما ورد أعلاه من إشارات فإن البحث سيهمل ولا يرد البحث إلى صاحبه.
- 9- تقديم أي بحث للنشر في المجلة يدل ضمناً على عدم نشره في أي مكان آخر، وفي حال قبول البحث للنشر في مجلة جامعة البعث يجب عدم نشره في أي مجلة أخرى.
- 10- الناشر غير مسؤول عن محتوى ما ينشر من مادة الموضوعات التي تنشر في المجلة

11- تكتب المراجع ضمن النص على الشكل التالي: [1] ثم رقم الصفحة ويفضل استخدام التهميش الإلكتروني المعمول به في نظام وورد WORD حيث يشير الرقم إلى رقم المرجع الوارد في قائمة المراجع.

تكتب جميع المراجع باللغة الانكليزية (الأحرف الرومانية) وفق التالي:
آ . إذا كان المرجع أجنبياً:

الكنية بالأحرف الكبيرة . الحرف الأول من الاسم تتبعه فاصلة . سنة النشر . وتتبعها معترضة (-) عنوان الكتاب ويوضع تحته خط وتتبعه نقطة . دار النشر وتتبعها فاصلة . الطبعة (ثانية . ثالثة) . بلد النشر وتتبعها فاصلة . عدد صفحات الكتاب وتتبعها نقطة . وفيما يلي مثال على ذلك:

-MAVRODEANUS, R1986- Flame Spectroscopy. Willy, New York, 373p.

ب . إذا كان المرجع بحثاً منشوراً في مجلة باللغة الأجنبية:

. بعد الكنية والاسم وسنة النشر يضاف عنوان البحث وتتبعه فاصلة، اسم المجلد ويوضع تحته خط وتتبعه فاصلة . المجلد والعدد (كتابة مختزلة) وبعدها فاصلة . أرقام الصفحات الخاصة بالبحث ضمن المجلة .
مثال على ذلك:

BUSSE,E 1980 Organic Brain Diseases Clinical Psychiatry News ,
Vol. 4. 20 – 60

ج . إذا كان المرجع أو البحث منشوراً باللغة العربية فيجب تحويله إلى اللغة الإنكليزية و
التقيد

بالبنود (أ و ب) ويكتب في نهاية المراجع العربية: (المراجع In Arabic)

رسوم النشر في مجلة جامعة البعث

1. دفع رسم نشر (40000) ل.س أربعون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.
2. دفع رسم نشر (100000) ل.س مئة الف ليرة سورية عن كل بحث للباحثين من الجامعة الخاصة والافتراضية .
3. دفع رسم نشر (200) مئتا دولار أمريكي فقط للباحثين من خارج القطر العربي السوري .
4. دفع مبلغ (6000) ل.س ستة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.

المحتوى

الصفحة	اسم الباحث	اسم البحث
42-11	د. صفاء العلي د. عبد الكريم السالم	تحليل ومحاكاة لبنية حوسبة الحافة المتنقلة (MEC) في شبكات الجيل الخامس
64-43	د. أحمد سلامة د. حسان حامد م. طارق عياشي	دراسة عملية للتآكل الخارجي في أنابيب نقل النفط تبعاً لعوامل مختلفة
92-65	د. محمد صلاحو د. غزوان الخوري د. يحيى فريد محمد أسعد	التنبؤ بقيمة المقطع العرضي الراداري للغيوم الإلكترونية باستخدام الشبكات العصبونية
130-93	م. محمد ديب د. ناصر أبو صالح	أنظمة التوصية المعتمدة على الفلترية التعاونية المعززة أداؤها بنماذج تعلم الآلة
154-131	م. وعد خويص	بناء مجموعة معطيات لتطبيق المنهجية الكلامية اللحنية في تصحيح نطق الأصوات العربية المُفخّمة حاسوبياً

تحليل ومحاكاة لبنية حوسبة الحافة المتنقلة

(MEC) في شبكات الجيل الخامس

طالب الدراسات العليا: صفاء العلي كلية: الهمةك - جامعة: البعث
الدكتور المشرف: عبد الكريم السالم

الملخص:

تدعم شبكات الجيل الخامس مجموعة واسعة ومتنوعة من التطبيقات ذات الكلفة الحسابية العالية التي تحتاج للتخزين والحساب خلال زمن انتقال منخفض وعرض مجال ترددي كبير، وهذا ما توافق مع زيادة التحديات في شبكة الاتصال، وخاصةً تلك المتعلقة بزمن الاستجابة وزيادة حركة المرور وتكاليف الاتصال. مما استدعى الحاجة إلى وجود بنية حوسبة متطورة وموزعة تعمل بالقرب من أجهزة المستخدمين النهائية التي تنتج البيانات وتستهلكها لمواجهة هذه المتطلبات وخاصةً خفض زمن الانتقال واستخدام المجال الترددي إلى أقصى إمكانياته وحساب البيانات وتخزينها.

تحقق تقنية ((MEC)) - Mobile Edge Computing تلك الغاية وتعتبر تقنية دعم أساسية في تنفيذ شبكات 5G وإبراز خدماتها وتطبيقاتها. يهدف هذا البحث إلى دراسة بنية (MEC) المقترحة من قبل المعهد الأوروبي لمعايير الاتصالات ETSI، وتوضيح كيف يتم استخدام الخدمات الموجودة على مخدمات (MEC) من قبل التطبيقات الموجودة على الأجهزة UEs في شبكة 5G. تم استخدام محاكي Simu5G لنمذجة شبكة 5G و (MEC)، بغرض الحصول على بعض مؤشرات الأداء لهذه البنية المتكاملة، مثل تأخير الرزم، وزمن الإرسال و SNR، فضلاً عن حركة المرور على الوصلة الهابطة والوصلة الصاعدة. بينت نتائج المحاكاة أن زيادة حجم وعدد التطبيقات المرسله أدى إلى زيادة التأخير في الشبكة على كلا الوصلتين الصاعدة والهابطة.

الكلمات المفتاحية: شبكات الجيل الخامس، حوسبة الحافة، حوسبة السحابة، Simu5G

Analysis and simulation of mobile edge computing architecture in fifth generation networks

Abstract:

Fifth generation networks support a wide variety of applications with high computational costs that need storage and computation through low latency and large bandwidth, and this corresponds to increase in network challenges, especially those related to response time, increased traffic and communication costs. This called the need for an advanced and distributed computing architecture that operates near the end-user devices that produce and consume data to meet these requirements, especially reducing latency, using the bandwidth to its maximum potential, calculating and storing data.. (Mobile Edge Computing-(MEC)) technology achieves this goal and is considered a basic support technology in the implementation of 5G and highlighting its services and applications, so this research aims to theoretical study of the (MEC) structure proposed by the European Telecommunications Standards Institute(ETSI),and demonstrate how services on (MEC) hosts are used by applications on UEs in a 5G network. The Simu5G simulator was used to model the (MEC) and 5G network, in order to obtain some performance indicators of this integrated architecture, such as packet delay, transmission time and SNR, as well as downlink and uplink traffic. Simulation results showed that increasing the size and number of applications sent leads to an increase in network delays on both uplinks and downlinks.

Keywords: fifth generation networks, edge computing, cloud computing, Simu5G

1. مقدمة

أدت التطورات المتلاحقة في شبكات الجيل الرابع 4G إلى تحقيق سرعات وإمكانات جديدة في الشبكة، حيث تمكن المستخدمون من الوصول عبر أجهزتهم المحمولة إلى الخدمات التي لم تكن متاحة من قبل، مثل مشاهدة فيلم أو المشاركة في الألعاب عبر الإنترنت وغيرها. وبمرور الوقت ازداد استخدام الأجهزة المحمولة بشكل كبير، وأصبح من الممكن المشاركة في الشبكات الاجتماعية، والتواصل مع الناس حول العالم أثناء مشاهدة فيلم أو مقطع فيديو على Youtube مثلاً، وهو الأمر الذي كان من الصعب جداً تخيله قبل بضع سنوات. تستطيع شبكات 4G أن توفر هذا الاتصال، لكن شبكات الجيل الخامس 5G تأتي كتطور يتجاوز الاستخدام الشخصي للإنترنت عبر الأجهزة المحمولة. لأنها توفر معدل بيانات مرتفع للغاية وإمكانية اتصال كبيرة بأجهزة أخرى غير المحمولة، كونها أجهزة يمكن وضعها في البيئة المحلية وفي البيئة الصناعية. وبالتالي لا تهدف 5G للتواصل الشخصي فقط إنما لتوصيل أي جهاز لديه إمكانية الوصول إلى الإنترنت باعتباره مكوناً مهماً للغاية في تطوير الاتصالات من آلة إلى آلة **Machine-to-Machine (M2M)** وإنترنت الأشياء **Internet of things (IoT)** [1].

أصبح إنترنت الأشياء IoT المصطلح المستخدم لتعريف الاتصال بين الأشياء المستخدمة في الحياة اليومية كالأجهزة ووسائل النقل. والتي تهدف إلى تسهيل الحياة اليومية للمستخدمين وأتمتة العمليات التي يقوم بها الإنسان. لذلك سيكون هناك عدد كبير من الأجهزة المتصلة و سيزداد الطلب على تنفيذ المهام الحسابية للتطبيقات وبالتالي سيزداد حجم حركة المرور traffic في الشبكة. في ذات الوقت يمتلك معظم المستخدمين النهائيين سعة تخزين ومعالجة محدودة [2] ومع أخذ ذلك في الاعتبار فإن هذا البحث يهدف إلى تحليل أداء تقنية حوسبة الحافة المتنقلة (MEC) باعتبارها محركاً أساسياً في تطور شبكات 5G، كونها تجعل موارد الحوسبة السحابية أقرب إلى المستخدمين النهائيين.

تم تنظيم هذا البحث وفق مايلي: أولاً تم تقديم معلومات عن الانتقال من الحوسبة السحابية المتنقلة Mobile Cloud Computing (MCC) المركزية إلى حوسبة الحافة الموزعة (MEC)، ثم تم تقديم دراسة لبنية (MEC) مع ذكر أهم أدوات المحاكاة

المستخدمة لنمذجة هذه البنية، ومثال توضيحي لشرح كيفية تنفيذ الخدمات فيها، وبعد ذلك استُخدم محاكي Simu5G لنمذجة بنية (MEC) المدمجة في شبكات 5G، وأخيراً قُدِّمت الخلاصة والأعمال المستقبلية.

2. الدراسة المرجعية: يتضمن البحث محورين أساسين هما الحوسبة (Computing)

المتتملة بحوسبة الحافة وشبكة الاتصال (Communication) متمثلةً بشبكة 5G. الحوسبة: استعرض المرجع [3] القيود الأساسية بالنسبة لأجهزة المستخدمين والتي تتعلق بحجم الذاكرة وعمر البطارية والطاقة الحسائية المطلوبة لتنفيذ عمليات معقدة. زادت تلك القيود مع نمو الإنترنت وأصبح من الصعب إدارة موارد الشبكة بكفاءة، ومع أخذ ذلك في الاعتبار ظهرت الحوسبة السحابية النقالة (MCC) كبنية قادرة على دعم هذه العمليات المعقدة، ونقل تنفيذ المهام الحسائية من الأجهزة إلى السحابة، التي توفر الموارد مثل التخزين وقاعدة البيانات وسعة الحوسبة للخدمات والتطبيقات، ويمكن طلبها حسب الحاجة من خلال الإنترنت، وبالتالي يمكن أن تقلل من تكاليف التشغيل. لكن غالباً ما تكون مراكز البيانات بعيدة عن الأجهزة الطرفية مما قد يتسبب في تأخير وضعف جودة التجربة (Quality of Experience - QoE) للتطبيقات الحساسة لزمان الاستجابة. ونتيجة لذلك قد لا يكون (MCC) قادراً على تقديم الدعم الكامل لتنفيذ إطار عمل الجيل الخامس، مما يتطلب نموذجاً جديداً للحوسبة. في المرجع [4] قُدِّمت فكرة حوسبة الحافة أو الحوسبة المتطورة (MEC) كنموذج جديد قادر على أداء المهام الحسائية على حافة الشبكة، بحيث تكون أقرب إلى المستخدم مما يسمح باتخاذ القرار وتخزين البيانات الكبيرة وتحليل ومعالجة بيانات الزمن الفعلي التي يكون لها زمن انتقال محدد مسبقاً. لذلك تعد وسيلة إضافية لتحسين الأداء كونها تسمح باستخدام السحب الصغيرة الواقعة بالقرب من المستخدمين النهائيين [5-6].

في المرجع [7] اقترحت بنية (MEC) لأول مرة في عام 2014 من قبل (European Telecommunications Standards Institute - ETSI). في البداية كان اختصاراً لـ Mobile Edge Computing، الذي يشمل فقط شبكات المحمول في إشارة إلى استخدام المحطات الأساسية Base Station - BS لأداء مهام حسائية مكثفة من الأجهزة الطرفية المحمولة. وفي 2016 أعادت ETSI تسمية (MEC) باسم

(Multi-access Edge Computing)، سامحةً بتطبيق (MEC) على الشبكات غير المتجانسة مثل تقنية الوصول الثابت والواي فاي وغيرها. لذلك تعد الحوسبة ذات الوصول المتعدد تقنية أساسية لشبكة الجيل الخامس لأنها تدمج خدمات الاتصالات وتكنولوجيا المعلومات. وتسمح بتوزيع موارد الحوسبة السحابية المركزية على حافة شبكة الوصول الراديوي ((Radio Access Network-(RAN)، وبالتالي تقلل زمن الوصول للمستخدمين النهائيين. وتمكنهم من التغلب على التحديات من خلال توفير بيئات مفتوحة على منصات سحابية متوفرة على حافة شبكة [8].

شبكة الاتصال: أشار المرجع [9] إلى أنه من المتوقع بحلول عام 2027 سيكون أكثر من 62% من إجمالي حركة البيانات المتنقلة ناتجة عن شبكات الجيل الخامس، القدرة على توفير زمن انتقال منخفض ومعدلات إرسال بيانات عالية، مما يسمح للعديد من القطاعات مثل الصناعة والترفيه والزراعة والصحة بالاستفادة منها. لذلك من الواضح أن الجيل الجديد لن يقدم الفوائد للمستخدمين العاديين فحسب بل سيفتح أيضاً منظوراً جديداً للتقدم في الاتصالات بين الآلات. ولتحقيق ذلك حددت (3rd Generation Partnership Project-(3GPP) مجموعة من متطلبات الأداء لتكنولوجيا الجيل الخامس لم تكن موجودة في الشبكات السابقة [10-11] ، وهي مبينة في الجدول (1):

الجدول (1): متطلبات الأداء لشبكات الجيل الخامس المحددة من 3GPP

متطلبات الأداء	
20 Gbps	معدلات ذروة الوصلة الهابطة
10 Gbps	معدلات ذروة الوصلة الصاعدة
1-100 ms	زمن الانتقال
10-500 km/h	التنقل (mobility)

حددت المراجع [12-13-14] فئات الخدمة، والتي تكون وفقاً لاحتياجات المستخدمين سواء للاستخدام التقليدي أو لآلاف الأجهزة المتصلة بالشبكة، لذلك فإن سيناريو واحد فقط غير قادر على وصف المتطلبات المذكورة سابقاً. وبالتالي تم تحديد ثلاثة سيناريوهات لفئات الخدمة التي تجمع الخصائص المشتركة لـ 5G:

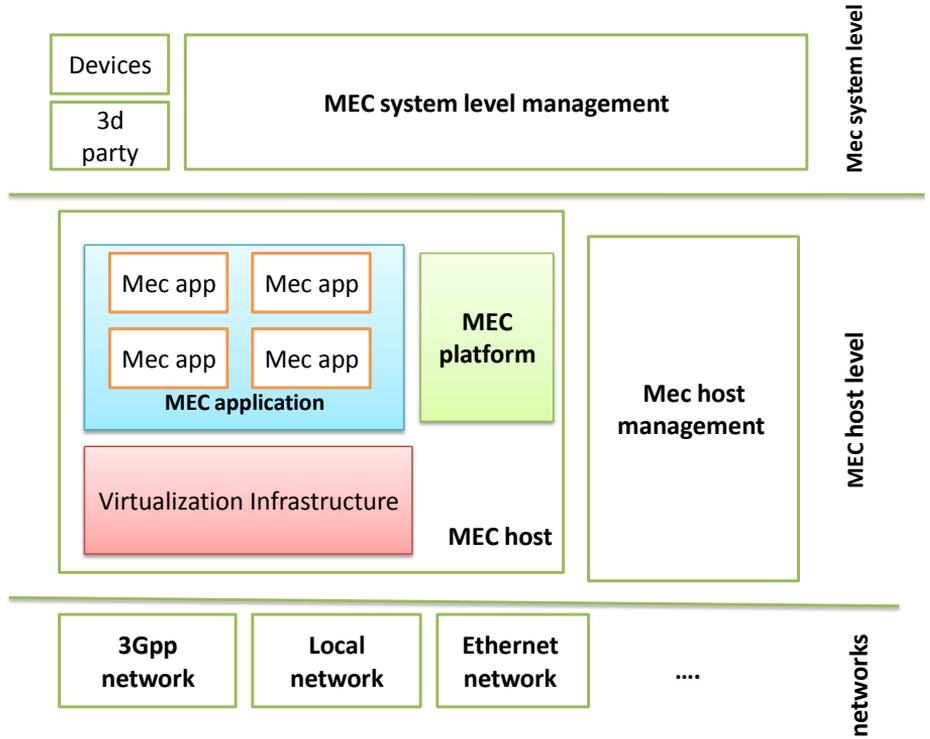
1. **(Enhanced Mobile Broadband-(eMBB))**: يضم الخدمات التي تتطلب اتصالات مستقرة ذات معدلات إرسال عالية قادرة على تغطية مساحة كبيرة، مثل تطبيقات الواقع المعزز.

2. **(ultra-reliable low latency communications - (URLLC))**: تتعلق بالخدمات التي تتطلب عمليات إرسال بزمن انتقال منخفض والموثوقية العالية مثل أنظمة القيادة الذاتية للسيارة.

3. **(Massive Machine-Type Communications-(mMTC))**: يتسم بدعم الاتصالات بعدد كبير من الأجهزة، مما يؤدي إلى إرسال حجم منخفض نسبياً من البيانات غير الحساسة للتأخير، مثل انترنت الأشياء [13]. إن التعايش بين المجموعات الثلاثة السابقة من الخدمات في نفس بنية شبكة 5G سيكون ممكناً بفضل ظهور تقنيات جديدة مثل افتراضية وظائف الشبكة (NFV) و الشبكات المعرفة بالبرمجيات (SDN) وتشريح الشبكة (slicing)، الذي يسمح بتخصيص موارد الحوسبة والتخزين والاتصالات بين الخدمات النشطة من أجل ضمان الفصل بين فئات الخدمة الثلاثة وتقديم أداء جيد.

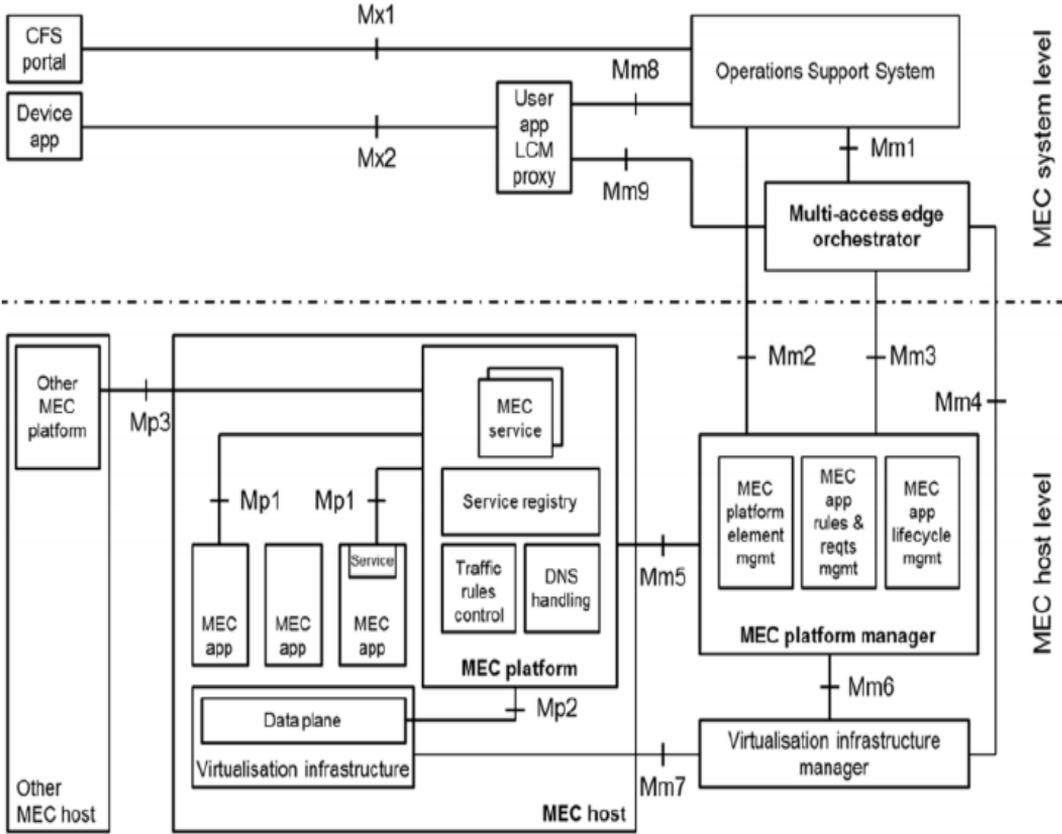
3. بنية نظام :

يعرض الشكل (1) النموذج المرجعي لبنية (MEC) الذي تم توحيد من قبل ETSI. تمكن هذه البنية التطبيقات من العمل بشكل مثالي في بيئة الشبكة بغض النظر عن نوع الوصول، وتتكون من ثلاث مستويات هي مستوى المضيف ومستوي النظام ومستوي الشبكة [16]. تربط بين هذه المستويات نقاط مرجعية، حيث توجد ثلاث مجموعات من النقاط المرجعية:



الشكل(1): بنية نظام MEC

1. النقاط المرجعية المتعلقة بوظيفة منصة (Mp): تربط بين مكونات المنصة، مثل Mp3 ، Mp2، Mp1
2. النقاط المرجعية للإدارة (Mm): عددها 9 نقاط تقوم بربط مكونات النظام المختلفة مع مستوى الإدارة
3. النقاط المرجعية التي تتصل بكيانات خارجية (Mx): مثل النقطتين Mx2، Mx1 اللتان ترتبطان مع الكيانات الخارجية [16]. توضح هذه النقاط في الشكل (2) الذي يصف البنية التفصيلية لل (MEC).



الشكل (2): البنية المفصلة لل (MEC)

3.1 مستوى مضيف (MEC) : يحوي على قسمين هما المضيف وإدارة المضيف.

1. مضيف (MEC) : يتضمن ثلاث كتل (المنصة والتطبيقات والبنية التحتية

الافتراضية)، يهدف المضيف إلى توفير موارد الحوسبة والتخزين والشبكات اللازمة لتشغيل التطبيقات. ففي المنصة تتوفر بيئة البنية التحتية حيث تكون التطبيقات قادرة على اكتشاف خدمات (MEC) واستهلاكها وتقديمها، أما التطبيقات فهي تعمل على بنية أساسية افتراضية والتي يمكن أن تكون آلات أو حاويات افتراضية، وفي البنية التحتية الافتراضية تتوفر الموارد الحاسوبية التي ستستخدمها تطبيقات (MEC) .

2. إدارة مضيف (MEC): يتكون مستوى الإدارة من مكونين أحدهما مدير المنصة الذي يدير دورة حياة التطبيقات والقواعد والمتطلبات، بما في ذلك الخدمات المصرح بها وقواعد المرور وتكوينات- Domain Name System (DNS)). والثاني مدير البنية التحتية الافتراضية (Virtualized Infrastructure Manager- (VIM)) المسؤول عن إدارة الموارد وإعداد البنية التحتية الافتراضية للعمل كصورة برمجية.

3.2 مستوى نظام (MEC): يتضمن.

1. المنسق (Orchestrator): يعتبر الوظيفة الرئيسية لنظام (MEC) لأن لديه نظرة عامة على النظام بناءً على مضيفي (MEC) المنفذين والموارد والخدمات المتاحة والطوبولوجيا. كما يقوم بالاختيار المناسب لمضيفي (MEC) لإنشاء مثل للتطبيقات وفقاً لقيود زمن الوصول والموارد والخدمات المتاحة. وأخيراً يقوم بإنشاء مثل وإنهاء التطبيقات وإعادة تخصيصها عند الضرورة.

2. نظام دعم المشغل ((Operating System Support- (OSS): يستلم الطلبات من تطبيقات الجهاز لإنشاء مثل لها وإنهائه عبر بوابة ((Customer Facing Service- (CFs (خدمة مواجهة العملاء)). ويتم إرسال الطلبات المقبولة إلى (Orchestrator) لمعالجتها.

3. (User Application LifeCycle Management Proxy- (UALCMP)) وكيل إدارة دورة حياة تطبيق المستخدم، يتفاعل هذا الوكيل مع (OSS) و (Orchestrator) من أجل معالجة طلبات التطبيق من الأجهزة أي يعمل كوسيط لإنشاء مثل لها أو إنهائها.

3.3 خدمات وتطبيقات (MEC): التطبيق والخدمة مفاهيم مرتبطة في سيناريو (MEC) هذا ولكنهما مختلفان، حيث يرتبط تطبيق (MEC) بما يحتاجه المستخدم من المتطلبات الضرورية، ثم يتم تنفيذه في تطبيق (MEC) خاص يقدمه الخادم للعميل. أما خدمات (MEC) فهي التي يقدمها مضيف (MEC) أي لن يؤدي المضيف إلا هذه الخدمة التي يطلبها تطبيق المستخدم إذا كانت هذه الخدمة متاحة على النظام الأساسي لل (MEC). يتم توفير هذه الخدمة واستهلاكها بواسطة المنصة أو بواسطة التطبيق. حيث

يتضح من الشكل (2) أنه عندما تتوفر خدمة من خلال (MEC app)، يتم تسجيلها على منصة Service registry عبر النقطة المرجعية Mp1 ثم يتم تقديمها إلى المستخدم. تلعب إدارة مستوى نظام (MEC) دورًا مهمًا فيما يتعلق بتطبيقات (MEC)، حيث يمكنها التحقق من القواعد والمتطلبات وتعيينها لتطبيقات (MEC). قد تكون هذه المتطلبات مرتبطة بالتطبيقات والموارد الضرورية لها مثل أقصى زمن انتقال والخدمات الضرورية، أو قد تكون غير مرتبطة [16-17].

3.4. أدوات المحاكاة:

تعتبر محاكاة نظام (MEC) في بيئة 5G عملية معقدة وخاصة إذا تم تناول جميع الجوانب ذات الصلة بالحوسبة والاتصال، فيما يلي شرح للخصائص الرئيسية التي يجب على المحاكاة دعمها:

1. دعم المحاكاة الافتراضية: تعتبر ضرورية لـ (MEC) لأنها تخلق بيئة معزولة للتطبيقات على الأجهزة المضيفة. وهذا يتيح لـ (MEC) أن تكون مرنة فيما يتعلق بإجراء التثبيتات دون التأثير على البرامج قيد التشغيل الأخرى.
2. دعم الشبكة: يمكن أن يؤثر إنشاء الشبكة على محاكاة (MEC). ففي حالات 1G و 2G ليست هناك حاجة لمنصة (MEC) لأن 1G و 2G يعتمدان على الراديو ولا يدعمان اتصالات ذات معدل بيانات مرتفع. بينما تحتاج 4G والأجيال اللاحقة إلى عقدة (MEC) تكفي للتحكم في التدفقات العالية للبيانات الضخمة. يجب أن تدعم شبكات 5G خوادم الحافة القريبة من الأبراج والتي تقلل زمن الوصول من 10-20 ms [20].
3. دعم (Orchestrator): المنسق هو وحدة تحكم لـ (MEC) كونه يتحمل مسؤولية العمل مع وحدة التحكم المحلية بناءً على حالة القياس الخاصة به مثل وحدة المعالجة المركزية وذاكرة الوصول العشوائي، ويقرر ما إذا كان (MEC) بحاجة إلى التفرغ [23].
4. وحدة التحكم المحلية: نظرًا لكون (MEC) أقرب إلى المستخدم لذلك يجب عليه قياس البيانات أو تجميعها قبل إرسالها إلى السحابة الرئيسية. ومن الأمثلة على ذلك البيانات الواردة من أجهزة إنترنت الأشياء التي يجب تحليلها

مسبقاً وإرسال التقارير فقط إلى السحابة أو اتخاذ قرارات محلية، ويتعين على (MEC) التحكم في مواردها بنفسها للإبلاغ عن استهلاك الطاقة واستخدام وحدة المعالجة المركزية وذاكرة الوصول العشوائي. ويجب أن تكون هذه المقاييس تحت السيطرة من خلال كل من (MEC) والمنسق بحيث لا يؤدي نشر ((Virtual Machine-VM) أو حاوية إلى حالة تحميل زائد في (MEC) للحفاظ على حالة الجهاز سليمة [20].

5. **التفريغ:** يهدف التفريغ في (MEC) إلى توفير طاقة الأجهزة التي تشغل تطبيقات ذات كلفة حسابية عالية وتقليل زمن حسابها، بحيث لا توجه الطلبات الواردة من جهاز المستخدم ((User equipment-UE) إلى الخادم البعيد. وبالتالي يُمكن التفريغ في (MEC) من دعم السحابة الرئيسية عندما لا تقوم بمعالجة جميع المهام .

6. **دعم التنقل:** هناك جانبان لدعم التنقل الأول هو محاكاة التنقل داخل أداة المحاكاة والثاني هو وحدة إدارة التنقل في (MEC) . هناك عملاء متحركون لذلك من المهم اختبار ما إذا كانت منصة (MEC) لديها حل كافٍ للمساهمة في خدمة العملاء المتحركين [20].

7. **دعم الترحيل:** ترحيل الخدمة هو عملية تحدث أثناء حدث التسليم وهي أحد الدعامات الرئيسية لإدارة التنقل. يتم نشر الخدمات في بيئة أو حاويات افتراضية وستقوم عملية الترحيل بنقل حالات الخدمات مثل نظام الملفات وذاكرة الوصول العشوائي ووحدة المعالجة المركزية إلى (MEC) التالي من (MEC) الحالي وإعادة توجيه حركة مرور الشبكة لإعادة الاتصال ب UE. يتطلب ترحيل الخدمة وحدة التنقل لإحداث عملية الترحيل.

8. **دعم الأمان:** قد تكون هناك مشكلات أمنية أثناء نقل جزء من الموارد والحساب من السحابة إلى الحافة. يمكن حماية الخادم السحابي بجدار حماية أو مضاد فيروسات، لكن لا يمكن قول الشيء نفسه عن (MEC). وبالتالي سيؤدي هذا إلى إنشاء نقطة نهاية ضعيفة لكل جهاز (MEC) . يمكن لوحدة المصادقة المثبتة على (MEC) حماية الشبكة السحابية وإيقاف الهجمات غير المصرح بها

على الشبكة [20]. حالياً يوجد العديد من أدوات المحاكاة ذات الإمكانيات المختلفة، والتي يمكن الاعتماد عليها في نمذجة (MEC)، مثل (EdgeCloudSim ، Simu5G ، IoTSim-Edge ، iFogSim ، Eclipsefogo5 ، Italtel-MEC) وغيرها، يوضح الجدول (2) مقارنة بين تلك الأدوات مع الأخذ في الاعتبار جميع الخصائص السابقة.

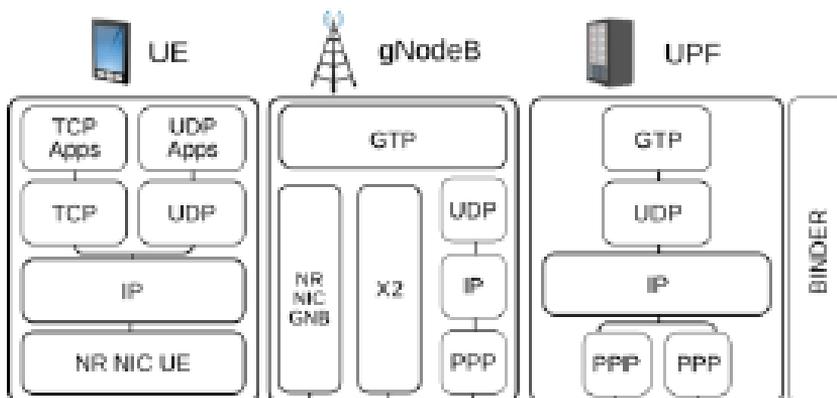
الجدول (2): مقارنة بين أدوات المحاكاة الحالية

Simulator	Virtualization	Network	Orechestr	RN I	Location	Traffic	Security
IoTSim-Edge	yes	common	No	yes	yes	No	No
EdgeCloudSim	yes	common	yes	yes	yes	yes	No
iFogSim	yes	Up to 4g	no	yes	yes	No	No
Eclipse Fogo5	yes	common	no	yes	yes	No	No
Italtel i-(MEC)	yes	Up to 5g	yes	yes	yes	yes	No
Simu5g	yes	Up to 5g	yes	yes	yes	yes	No

من هذا الجدول يتبين أن Simu5G يحقق جميع الخصائص السابقة لذلك اعتُمد في البحث.

3.4.1 التعريف ب Simu5G

يعتمد Simu5G على إطار عمل OMNeT++، الذي يحاكي عملياً كافة أنواع الشبكات سواءً السلكية واللاسلكية [18]. يعد استخدامه مناسباً للنماذج الأولية للتطبيقات التي تحتاج إلى اختبار في بيئة يمكن التحكم فيها لقياس مؤشرات الأداء. يتم تمثيل مكونات النظام بوحدات نمطية بسيطة تتواصل مع بعضها البعض عبر الاتصالات بين منافذها (بواباتها)، والتي تعمل كواجهات. في كل تبادل للرسائل بين الوحدات النمطية يتم استدعاء حدث بواسطة نواة المحاكاة. يرمز سلوك الوحدات باستخدام لغة C++. ويمثل تكوين الاتصالات والمنافذ والبارامترات والطبولوجيا بواسطة ((Network - (NED) definition) لغة وصف الشبكة، وتكتب قيم البارامترات في ملفات -INI Initialization. يسمح هذا البرنامج بنمذجة العناصر التالية الموضحة في الشكل (3) وهي كالتالي:



الشكل(3): العناصر التي يتيح Simu5G نمذجتها

1. **نمذجة عناصر الشبكة:** توفر Simu5G نمذجة كل من ((User Plane) (UE) و Function-(UPF) و (gNodeB). ونظراً لأن هذا الإطار قادر على نمذجة السيناريو الذي يتضمن فقط مستوى المستخدم لبنية شبكة G5 الأساسية، يحتوي Simu5G على وحدة نمطية خاصة تمثل مستوى التحكم بالكامل يسمى الموثق. بالإضافة إلى ذلك هناك وحدة تجميع الحوامل (carrierAggregation)، حيث يتمكن كل من هذا والموثق من الحفاظ على نظرة عامة على معلومات الشبكة التي يمكن طلبها بواسطة الوحدات النمطية الأخرى [18].
2. **نمذجة (MEC):** يتيح Simu5G إمكانية نمذجة مكونات مستوى نظام (MEC) جميعها مثل المنسق و(UALCMP) و(OSS). ويتيح أيضاً نمذجة مكونات مستوى مضيف(MEC) مثل مدير البنية التحتية الافتراضية (VIM) ومدير المنصة ((MECPM-(MEC platform manager)، والوحدات النمطية اللازمة لتشغيل تطبيقات (MEC)، والوحدات النمطية نفسها لمنصة(MEC) والبنية التحتية الافتراضية.
3. **نمذجة الجهاز(UE):** يتم تفاعل (UE) مع نظام(MEC) من خلال كيانين هما تطبيق (UE) وتطبيق الجهاز[17]. الأول هو نقطة نهاية الاتصال في مستوى البيانات بين UE وتطبيق (MEC)، بينما يكون تطبيق الجهاز مسؤولاً عن التفاعل

مع نظام (MEC)، ويطلب إنشاء مثيل أو إنهاء تطبيق (MEC) [17]. يمكن الاستنتاج أن تطبيق UE هو الجزء الرسومي الذي يراه مستخدم UE ويطلب الخدمة وأن تطبيق الجهاز هو الجزء الذي لا يراه مستخدم UE أي كل جزء من الكود والواجهة التي تطلب الخدمة لنظام (MEC). أخيرًا يبدأ العميل فقط في استخدام الخدمة المطلوبة بعد أن يؤكد تطبيق الجهاز إنشاء تطبيق (MEC). يمكن لتطبيق UE أن يكون داخليًا أو خارجيًا لـ Simu5G [17].

4. **نمذجة التطبيقات (نقاط النهاية):** كل تطبيق (MEC) له قواعده ومتطلباته. يتم نمذجة البارامترات من خلال ملف JavaScript Object Notation-JSON ليسمى Application Descriptor، والذي يتكون من الحقول اللازمة لتكامل حزمة التطبيق في نظام (MEC). تم وصفها في الجدول (3) [17].

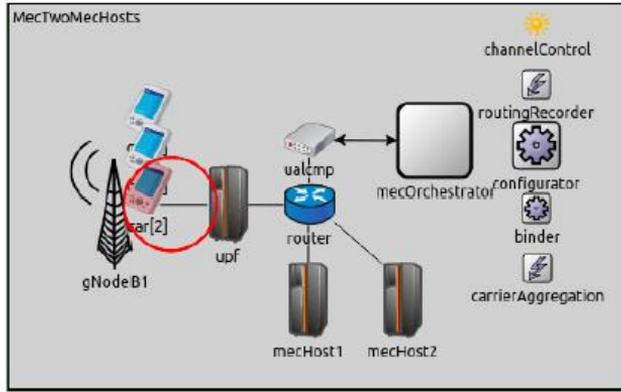
الجدول (3): البارامترات التي تصف التطبيق والموجودة في ملف JSON

AppId	معرف واصف التطبيق
App Name	اسم تطبيق (MEC)
AppProvider	اسم الوحدة المطلوبة لتشغيل التطبيق
AppServiceRequired	الخدمات اللازمة لتنفيذ التطبيق
VirtualComputeDescriptor	موارد الحوسبة (الذاكرة والقرص ووحدة المعالجة)

5. **نمذجة خدمات (MEC):** في Simu5G يتم توفير وحدة تسمى (MEC) ServiceBase، والتي تستخدم للنماذج الأولية السريعة لهذه الخدمات. تنفذ هذه الوحدة النمطية خدمتين هما خدمة الموقع و خدمة معلومات الشبكة الراديوية، (Radio Network Information Services-RNIS) فيما يلي مثال توضيحي حول كيفية تنفيذ الخدمات في مضيف (MEC) يعمل على شبكة (5G) باستخدام محاكي Simu5g.

3.5 مثال توضيحي باستخدام Simu5G:

في هذا المثال سيتم شرح سيناريو multMECHost كونه يمكن من استهلاك خدمات (MEC). وبالأخص خدمة تحديد الموقع. يوضّح الشكل(4) بنية الشبكة المقابلة لـ (MEC) و 5G المتصلة مع أجهزة مستخدمين UEs الممثلة بسيارات مجهزة بواجهات (New Radio-NR). لدينا ثلاث سيارات، سيارة [0] وسيارة [1] وسيارة [2]. كل سيارة لها تطبيق (UE) الخاص بها والموصوف بواسطة ملف (JSON) الذي يعطي الحد الأدنى من الموارد المطلوبة، وهي في هذه الحالة virtualDisk=10، virtualCPU=1500، (virtual memory=10) والموضحة في الشكل(5). تتحرك السيارات في مساحة معينة داخل شبكة (5G) وبسرعة 10m/s. وتريد تلك السيارات أن يتم إنذارها عندما تدخل منطقة الخطر، يتألف نظام (MEC) من مضيفين (MECHost1) و (MECHost2) و (ualcmp) والمنسق. يدير MECHost2 خدمة تحديد الموقع هنا، ولكن ليس لديه الموارد الحسابية الكافية. يوضّح الجدول(4) الموارد المتاحة على عند كل مضيف (MEC).



الشكل(4): بنية الشبكة المقابلة لـ (MEC) و 5G

الجدول (4): الموارد الموجودة عند كل مضيف

الموارد	MEC) Host1)	MEC) Host2)
RAM	32	32
Disk	100TB	100TB
MIPS Million Instruction per Seconds	40000	500

يتبين من هذا الجدول أن موارد المضيف الثاني (500 MIPS) غير كافية لتنفيذ التطبيق الحالي الذي يتطلب (1500 MIPS)، على الرغم من أن هذا المضيف يمتلك الخدمة المطلوبة.

```

1 {
2   "appId" : "WAMECAPP",
3   "appName" : "MECWarningAlertApp",
4   "appProvider" : "simu5g.apps.mec.WarningAlert.MEC",
5   "appInfoName" : "appInfoName_",
6   "appDescription" : "appDescription_",
7   "virtualComputeDescriptor" : {
8     "virtualDisk": 10,
9     "virtualCpu" : 1500,
10    "virtualMemory":10
11  },
12  "appServiceRequired": [
13    {
14      "ServiceDependency" :{
15        "serName" : "LocationService",
16        "version" : "v2",
17        "serCategory": "Location"
18      }
19    }
20 ],
21
22  "omnetppServiceRequired": "MEWarningAlertService"
23 }

```

الشكل (5): الحد الأدنى من الموارد المطلوبة للتطبيق

الهدف الأساسي من هذا المثال هو توضيح كيف يتم استخدام الخدمات الموجودة على مخدمات (MEC) من قبل التطبيقات الموجودة على الأجهزة UEs المتصلة مع gNodeB والتي تتحرك نحو منطقة خطر [17]. عند الدخول إلى منطقة الخطر أو مغادرتها يجب إنذار السيارات من خلال التطبيق الموجود على أجهزتهم والمسمى (UEWarningAlertApp). وبالتالي فإن الخدمة الأساسية هي خدمة الموقع، حيث

((MAC-control)) عبر القناة اللاسلكية بين gNodeBs و UEs، وتتألف من أربع عمليات رئيسية ((FEC-Forward error correction)) (التصحيح الأمامي للأخطاء) يليها اكتشاف الأخطاء ثم الجمع وإعادة الإرسال [25]. يمكن أن يشير (HARQ) إلى رسالة إقرار ACK للإقرار بإرسال الإطار الصحيح، أو (NotAcknowledgement-NACK) معرباً عن الفشل. يظهر في الشكل (7) جزء من السجل الذي يحتوي على HARQFeedback، حيث تكون النتيجة عبارة ACK. لم يتم العثور على NACKs في السجلات أثناء تنفيذ المحاكاة.

```
INFO: pdu_has been sent and received correctly
INFO: H-ARQ TX: feedback received for process 0 codeword 0 result is ACK
```

الشكل(7): نتيجة تبادل حزمة harqFeedback

3. ترسل السيارة [2] حزمة Airframe إلى gNodeB بحجم يساوي 53B تحتوي على معلومات عن القناة التي تم إنشاؤها للسيارة مثل (SNR) و- (Block Error Rate (BLER)) (معدل خطأ الكتلة) ومؤشر جودة القناة (CQI). يعرض الشكل (8) مقطعاً من حزمة Airframe الذي تم إرساله.

```
INFO: **** Background Cell Interference
INFO: RealisticChannelModel::getSINR - distance from my gNb=117.71 . DIR=UL
INFO: bgCell - ext[0] - multi[0] - recvpwr[ 3.48336e-09] - sinr[٤٤.٩٢]
INFO: bler computation: [ itxMode=1] . [ cqj-1=14] . [ snr=44]
INFO: RealisticChannelModel::error direction UL node 2051 remote unit MACRO Band 3 SNR 44 CQI 15 BLER 0 success probability 1
INFO: RealisticChannelModel:: error direction UL node 2051 total ERROR probability 0 per with H-ARQ error reduction 0 - CQI[١٥]
INFO: this is your lucky day (0.157752 > 0) -> Receive AirFrame
INFO: Handled Airframe with ID 697 with result RECEIVED
```

الشكل (8): إرسال مقطع من حزمة Airframe

- إنشاء اتصال بين العناصر الأخرى: بعد 9ms من بدء المحاكاة تنتهي الرسائل بين UEs و gNodeB، ويبدأ إنشاء الاتصال بين العناصر الأخرى وصولاً إلى (MEC). يمكن هنا معرفة البروتوكولات وحجم الرزم. حيث يستخدم بروتوكول النقل ((User Datagram Protocol-(UDP)) عند الاتصال بين gNodeB وUPF لأنه سريع وبسيط. وفي الوقت نفسه يستخدم بروتوكول Transmission ((Control Protocol-(TCP)) في حال الاتصالات الأخرى بين العناصر لتسليم الحزم بشكلٍ مضمون وموثوق. بهذه الطريقة من الممكن مراقبة تدفق الرسائل التي

تبدأ باتصال gNodeB مع (UPF) ثم تستمر باتصال UPF مع جهاز التوجيه الذي بدوره ينشئ الاتصال مع مضيفي (MEC) و ualcmp من خلال المنسق. وكنتيجة للمرحلة الأولى يتبين أن الاتصال يبدأ من UEs وينتهي عند أحد مضيفي (MEC) باستخدام التسلسل التالي (حزمة ملاحظات-حزمة منح-حزمة Airframe-حزمة UDP-حزمة TCP).

3.5.2 المرحلة الثانية: اختيار مضيف (MEC) بواسطة Orchestrator: بعد العملية الأولية التي أنشئت اتصالاً بين gNodeB و UE تم التحقق من تبادل الرسائل بين مكونات نظام (MEC)، حيث يرسل ualcmp إلى منسق (MEC) الطلب الذي يتطلبه تطبيق الجهاز. بعد ذلك يتحقق المنسق من أنسب مضيف (MEC) لإنشاء مثل للتطبيق ويتأكد من وجود الخدمة والموارد اللازمة لتنفيذها. يوضح الشكل (9) نتيجة عملية اختيار مضيف (MEC) من بين المضيفين المتاحين، حيث يبحث المنسق عن أنسب مضيف يلبي المتطلبات اللازمة لإنشاء مثل التطبيق، ثم يبحث عن وجود الخدمة والتي في هذه الحالة خدمة تحديد موقع UEs داخل السيناريو. وبعد أن يتبين للمنسق أن موارد (MECHost2) غير كافية لمثل هذا التطبيق يختار (MECHost1) لأنه ذو موارد كافية. وبالنتيجة يتم إنشاء تطبيق (MEC) على MECHost1 كما هو موضح في السطر 23 من الشكل (9) كونه المضيف الأفضل من حيث الموارد. على الرغم من أن الخدمة التي يحتاجها المستخدم موجودة على منصة (MECHost2)، إلا أن الاختيار يتم وفقاً للحد الأدنى من الموارد التي يحددها التطبيق. وبالتالي في هذا السيناريو يجب أن يستهلك التطبيق الذي تم إنشاؤها على (MECHost1) خدمة موجودة على مضيف مختلف وتحديداً (MECHost2).

20 Initializing module MecTwoMechosts.mechost1.MECWarningAlertApp[?] , stage 13

21 INFO:VirtualisationInfrastructureManager :: instantiateMEApp - MECWarningAlertAPP[?] instanced

22 INFO:VirtualisationInfrastructureManager :: instantiateMEApp - currentMEApps: 3 / 100

23 INFO: MecOrchestrator :: startMECApp - new MEC application with name: **MECWarningAlertApp instantiated on MEC host[] (omnetpp ::cModule)mechost1 id=13 at 192.168.5.2:4003**

الشكل (9): عملية اختيار مضيف (MEC)

3.5.3 سلوك التطبيق: في هذه الحالة هناك تبادل لحزمتين:

1. WarningAlertPacketStart: يستخدمها تطبيق UE بعد إنشاء مثيل للتطبيق يطلب فيها من تطبيق (MEC) مراقبة منطقة معينة، فيقوم تطبيق (MEC) باستخدام خدمة الموقع الخاصة بـMECHost2 من خلال رسائل ((Hypertext Transfer Protocol)-(HTTP)، هنا يتم توجيه هذه الحزمة من gNodeB إلى (MECHost1) مع معلومات المنطقة، حيث يحدد تطبيق UE الإحداثيات التي يجب مراقبتها. في البداية تكون UEs خارج منطقة الخطر الدائرية مما يعني أن التوقيع الأول يتم في حالة الدخول إلى منطقة الخطر. وبمجرد إندار تطبيق (MEC) بدخول السيارة إلى المنطقة يقوم بتعديل التوقيع، ولا يتلقى تطبيق UE سوى أحداث الإنذار إلى أن يتم إنشاء المورد الجديد بنجاح. يستخدم بروتوكول (HTTP POST) هنا لإنشاء التوقيع .

2. WarningAlertPacketInfo: تحتوي على حقل في الحزمة مشار إليه بـ " risk = true". فأتثناء تحرك السيارة ودخولها منطقة الخطر يتم إنشاء إشعار الدخول وتوجيهه من المصدر (MECHost1) إلى وجهة gNodeB ثم إرساله إلى تطبيق UE عبر Airframe. ونظرًا لأن (MECHost2) هو مالك خدمة الموقع فإنه يبدأ بعملية التنبيه والتي ستم معالجتها بواسطة (MECHost1). بعد دخول السيارة إلى منطقة الخطر يستخدم بروتوكول (HTTP PUT) لتحديث توقيع الإنذار. عندما تغادر السيارة هذه المنطقة الدائرية يتم إنشاء إشعار مغادرة و إرساله إلى تطبيق UE. تتبع تفاعلات الإرسال الخاصة بهذه الحزمة نفس التسلسل حزمة (WarningAlertPacketStart) مع الاختلاف الذي تُعلمه الحزمة الآن وهو " risk = false". ثم يتم حذف التوقيع باستخدام بروتوكول (HTTP DELETE) بعد الخروج من منطقة الخطر. و**كنتيجة** تم استخدام كل من (MECHost1) و (MECHost2) في هذا السيناريو حيث يستخدم الأول لمعالجة التطبيق والثاني لخدمة الموقع المحدد. أي أن (MECHost1) يستهلك خدمة الموقع الخاصة بـ (MECHost2) ويعالج هذه المعلومات من خلال مقارنتها بالإحداثيات المحددة للمراقبة من أجل إندار المستخدم بالخطر الوشيك.

3.5.4 تعديل بارامترات موارد التطبيق: ذكرنا سابقاً أن التطبيق يحتاج معدل معالجة MIPS 1500. ومن خلال الجدول (4) يتبين أن (MECHost2) لا يلبي هذه الحاجة. لذلك من أجل التحقق من كيفية اختيار المنسق إذا كانت بارامترات الموارد المتاحة لمضيفي (MEC) هي نفسها، سنقوم بتغيير قيمة الموارد عند (MECHost2) إلى نفس قيمة (MECHost1). وكنتيجة سيختار المنسق (MECHost2) باعتباره الأفضل لإنشاء مثيل للتطبيق، يوضح الشكل (10) جزء السجل الذي يعرض هذه النتيجة.

```
20 Initializing module MecTwoMechHosts.mechHost2.MECWarningAlertApp[?] , stage 13
21 INFO:VirtualisationInfrastructureManager :: instantiateMEApp - MECWarningAlertAPP[?] instanced
22 INFO:VirtualisationInfrastructureManager :: instantiateMEApp - currentMEApps: 3 / 100
23 INFO: MecOrchestrator :: startMECApp - new MEC application with name: MECWarningAlertApp[?] instantiated on MEC host[ ]
(omnetpp ::Module)mechHost2 id=14 at 192.168.6.2 : 4003
```

الشكل (10): الأحداث الناتجة بعد زيادة موارد (MECHost2)

3.5.5 الاستنتاج:

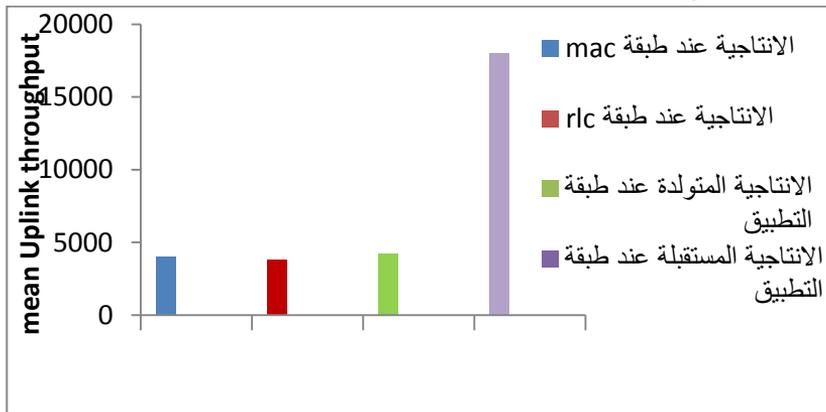
1. إذا كان المضيفين يمتلكان نفس القدر من الموارد، سيحدث التعريف من خلال معيار وجود الخدمة في منصة (MEC) طالما أن مضيف (MEC) لديه الحد الأدنى الذي يشير إليه التطبيق. وللتأكد أُجري اختبار إضافي بجعل (MECHost1) يمتلك موارد أكبر من (MECHost2) وقيمت النتيجة نفسها كما هو موضح في الشكل (10).
2. هناك عدد أقل من القفزات منذ أن تم إرسال رسالة (WarningAlertPacketInfo) بواسطة (MECHost2) حتى وصول Airframe إلى السيارة. لذلك لا يوجد تفاعل حتى مع (MECHost1) مما يجعل التوجيه أكثر ديناميكية.
3. بالنسبة للتطبيقات الهامة التي تتطلب زمن انتقال منخفض، فإن السيناريو المناسب هو أن يتم تقديم التطبيقات والخدمات من قبل مضيف (MEC) نفسه.

4. المحاكاة:

4.1 وصف السيناريو: اختير في هذه السيناريو جهاز مستخدم واحد UE متصل مع gNB التي تتصل بدورها مع UPF الذي يعتبر أساسى كونه يفصل بين مستوى التحكم ومستوى المستخدم في شبكة 5G. ترسل وحدة UPF البيانات إلى مضيف (MEC) واحد من خلال جهاز توجيه router، فيعالج المضيف البيانات ويرسلها إلى جهاز المستخدم. يهدف هذا السيناريو إلى جمع المعلومات المحدثة حول حالة الشبكة ووحدات UE المتصلة بالمحطة الأساسية المرتبطة بمضيف (MEC). تم دراسة الأداء من خلال مقاييس الأداء الأكثر أهمية مثل تأخير الرزم، وزمن الإرسال و SNR، فضلاً عن حركة المرور على الوصلة الهابطة والوصلة الصاعدة.

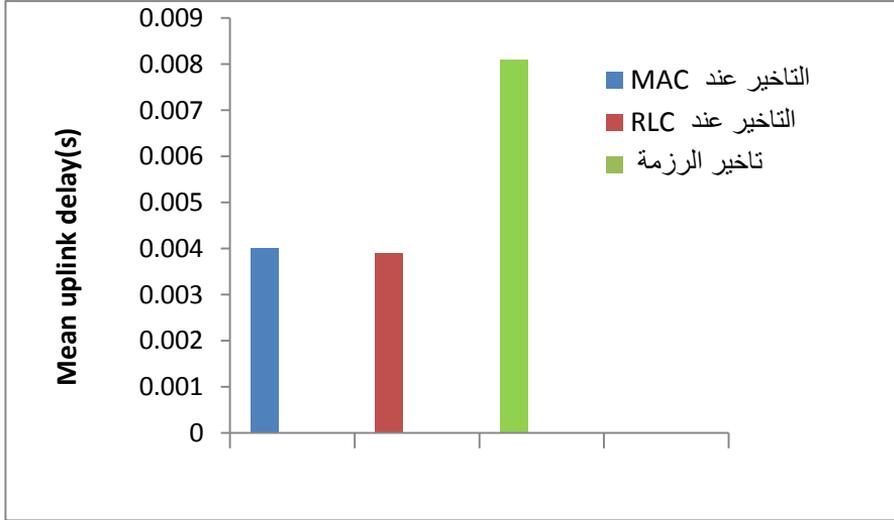
4.2 سير المحاكاة: يستخدم UE تطبيقاً واحداً ليكن VoIP بحجم رزمة يبلغ (50 B). تُشغل المحاكاة لرؤية نتائج الوصلة الهابطة والوصلة الصاعدة في هذه الحالة. بعد أن يتم ضبط طاقة gNB إلى 40w وطاقة UE إلى 26w.

تحليل الإرسال على الوصلة الصاعدة: في الوصلة الصاعدة يرسل تطبيق VoIP-UL البيانات إلى gNB ثم إلى مضيف (MEC). يسمى التطبيق الذي يقوم بتشغيل المضيف VoIPReceiver، و من جانب UE تضبط بارمترات حجم الرزمة (50 B)، وعنوان الإرسال destAddress، ومنفذ الاتصال للتطبيق VoIPSender الذي يرسل رزم الصوت. يوضح الشكل (11) حركة مرور البيانات على الوصلة الصاعدة.



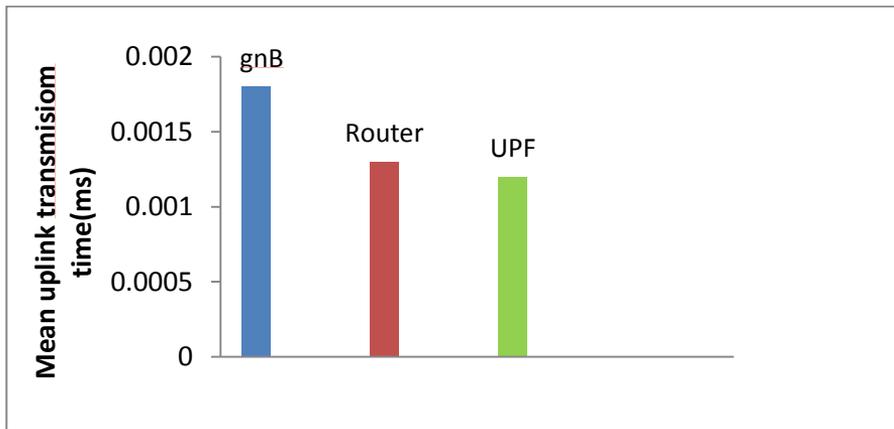
الشكل (12): حركة المرور المتولدة على الوصلة الصاعدة

يُلاحظ من الشكل(12) أن معظم حركة المرور تمر عبر طبقة الفيزيائية حيث يتم تبادل البيانات ومعالجتها بين الخادم و UE. يوضح الشكل (13) التأخير الحاصل على الوصلة الصاعدة عند أجزاء مختلفة من الشبكة مثل MAC و Radio link control- RLC والرزمة المرسله.



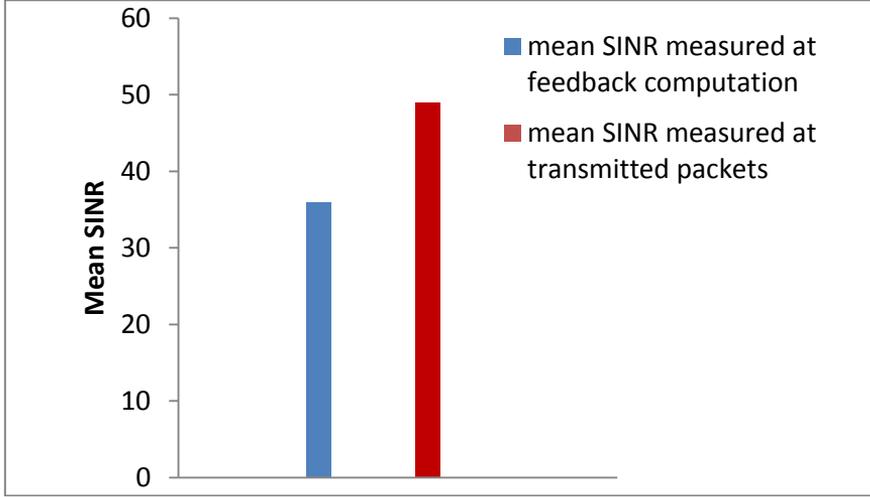
الشكل(12): التأخير الوسطي على الوصلة الصاعدة

تنتقل الرزم في هذا السيناريو من جهاز المستخدم إلى الخادم، لذلك سيحصل التأخير الأكبر عند إرسال رزمة VoIP الإجمالية. يعرض الشكل(13) زمن الإرسال عند كل من UPF و Router و gnB.



الشكل(13): متوسط زمن الإرسال

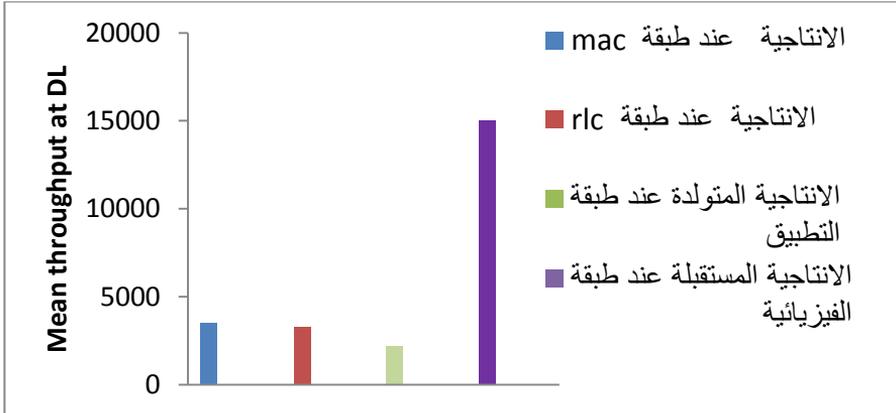
من الملاحظ أن أكبر تأخير موجود في gnb لأنه يعالج جميع الرزم التي ترسلها UE إلى الشبكة المركزية. يبين الشكل (14) نسبة الإشارة إلى الضجيج SNR بعد بدء الإرسال على Uplink.



الشكل (14): نسبة الإشارة إلى الضجيج

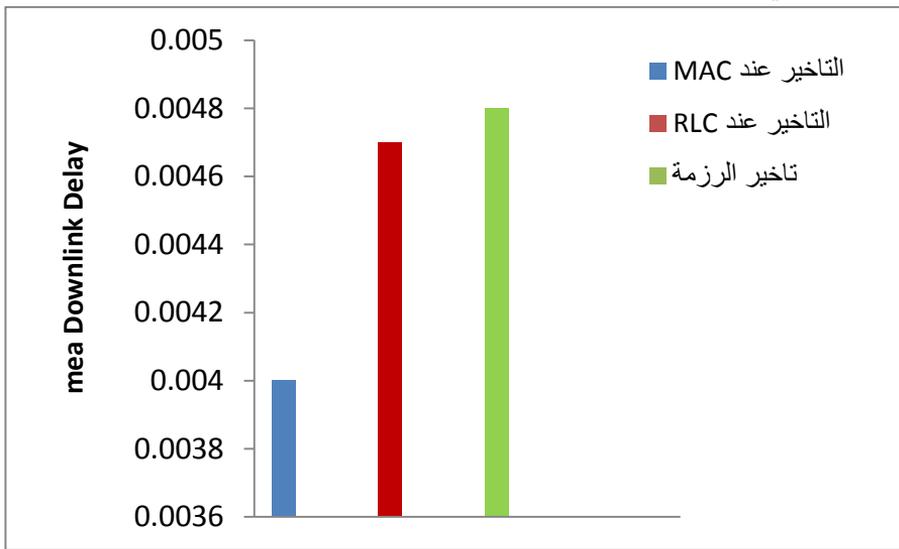
من الملاحظ أن قيمة SNR الناتجة بعد إرسال الرزم تكون أكبر من feedback SNR لأن الأخيرة متعلقة بالمعلومات الراجعة feedback من المستقبل، حيث تكون تلك المعلومات متأثرة بظروف القناة والتداخل لذلك تكون قيمتها أقل من SNR المرسل.

تحليل الإرسال على الوصلة الهابطة: في الوصلة الهابطة يرسل تطبيق VoIP-DL البيانات من الخادم إلى المستخدم. لذلك من جانب UE يتم تشغيل تطبيق (VoIPReceiver) لتلقي بيانات VoIP ويحدد منفذ الاتصال على UE. ثم من جانب المضيف تحدد الوجهة ($destAddress = ue(0)$) وحجم الرزمة وتطبيق (VoIPSender) الذي يرسل البيانات إلى UE. تم إيجاد كل من حركة المرور والتأخير و زمن الإرسال و SNR. يبين الشكل (15) حركة مرور البيانات على الوصلة الهابطة.



الشكل(15): حركة المرور المتولدة في حال إرسال الوصلة الهابطة

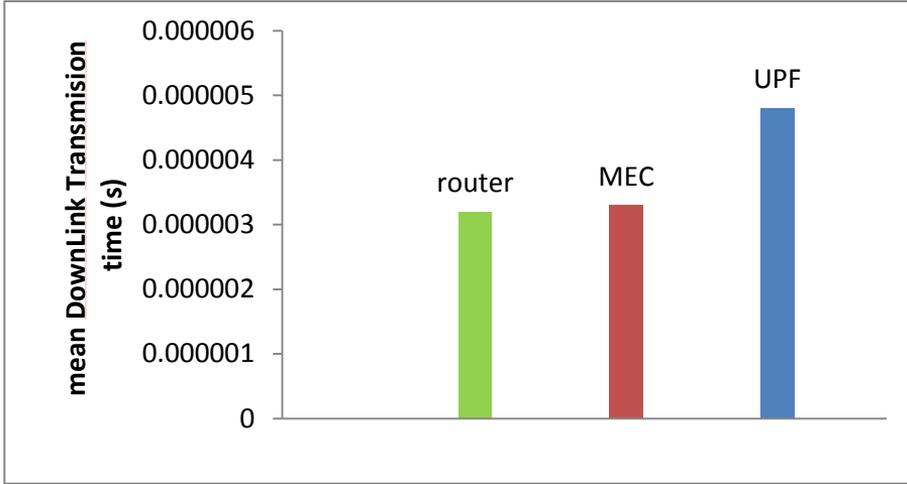
يُلاحظ من الشكل (15) أن معظم حركة مرور البيانات يتم توليدها بواسطة طبقة الفيزيائية حيث يتم تبادل البيانات ومعالجتها بين الخادم و UE. يوضح الشكل(16) التالي التأخير الوسطي عند كل من MAC و RLC والرزمة المرسله.



الشكل(17):التأخير

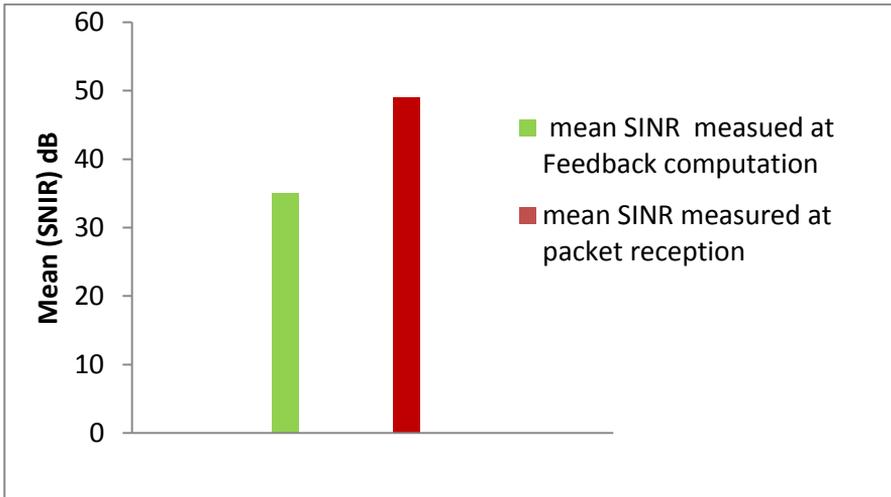
من الملاحظ أن التأخير يحصل في أجزاء مختلفة من الشبكة مثل طبقة MAC و RLC ويتراوح بين(4-5ms)، ويحصل التأخير الأكبر على الوصلة الهابطة عند رزمة VoIP الإجمالية. ومن الملاحظ أيضاً أن تأخير الإرسال في هذه الحالة أقل من التأخير

السابق على Uplink لأن الرزم ترسل من gNB إلى UE. يوضح الشكل (17) زمن الإرسال المستغرق على الوصلة الهابطة عند كل من Router و UPF و (MEC).



الشكل (17): زمن الإرسال في الوصلة الهابطة

من الملاحظ أن UPF لديه أكبر زمن إرسال نظرًا لأنه مسؤول عن عمليات المستخدم و يربط RAN بالشبكة الأساسية. يعرض الشكل (18) نسبة الإشارة إلى الضجيج المتولدة عند استقبال الرزم.

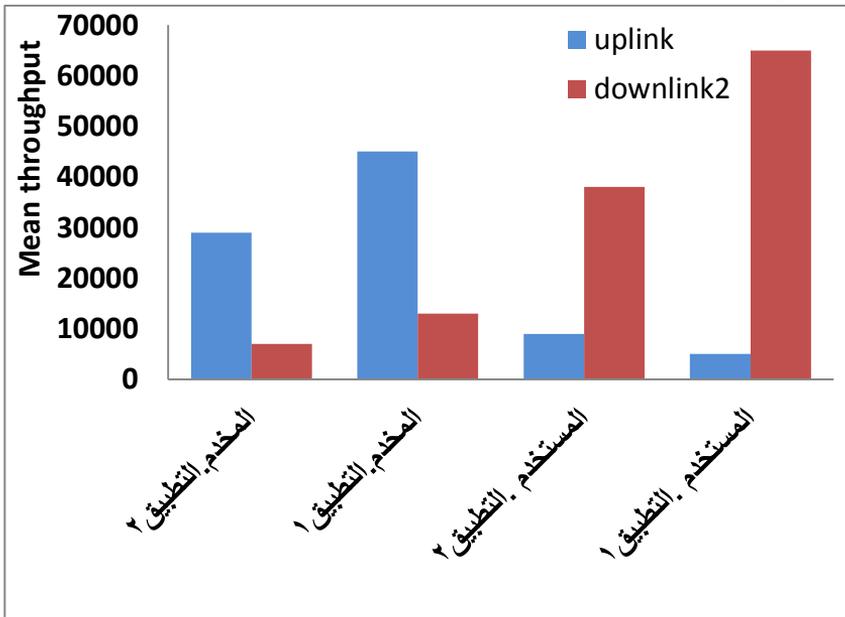


الشكل (18) نسبة الإشارة إلى الضجيج

من الملاحظ أن قيمة (SNR) الناتجة بعد استقبال الرزم تكون أكبر من FeedbackSNR لان قيمة FeedbackSNR تعتمد على الرزم التي يعاد إرسالها إلى المرسل وبالتالي تكون متعرضة لظروف القناة والتداخل.

حالة إرسال تطبيقين:

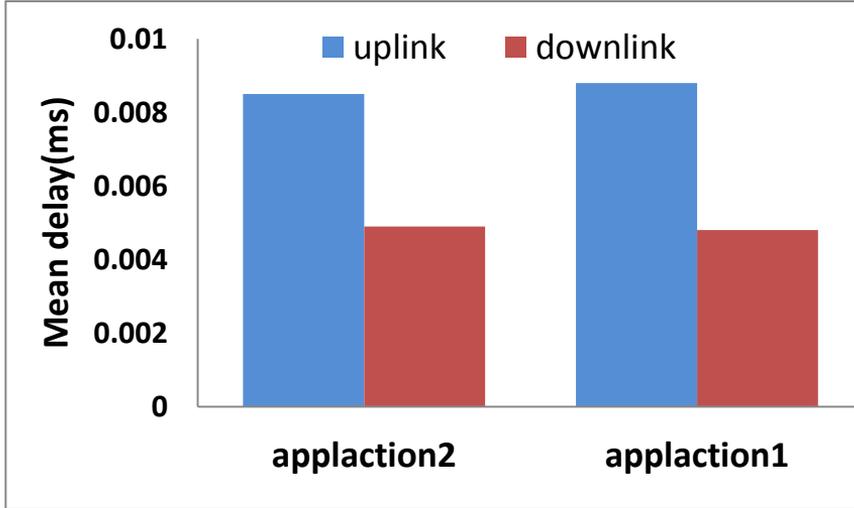
تم في هذا السيناريو قياس حركة المرور البيانات المتولدة عن تشغيل تطبيقين على UE. لذلك سنجري التغييرات على الكود بإضافة منفذ اتصال آخر عند UE ومخدم (MEC)، للتواصل على منفذين و نغير رقم التطبيقات إلى 2 ونغير حجم حزمة التطبيقات إلى 200. لندرس حركة المرور المتولدة في الشبكة عن التطبيقين على كل من الوصلة الهابطة والصاعدة عند كل من UE ومخدم (MEC). يوضح الشكل (19) النتائج الوسطية التي تم الحصول عليها من سجل الأحداث .



الشكل(19): حركة المرور البيانات traffic المتولدة عن تشغيل تطبيقين

نلاحظ من الشكل (19) أن حركة مرور كبيرة تصل إلى 70000 تم إنشاؤه في UE واحد على الوصلة الهابطة وهذا أمر واضح لأن تطبيقين يعملان على UE واحد. وفي الوصلة الصاعدة يتم توليد حركة مرور أقل لأن البيانات تذهب إلى الخادم وتتم معالجتها هناك.

أدى هذا التغيير في حجم وعدد التطبيقات المرسل إلى زيادة التأخير في الشبكة كما يوضح الشكل (20).



الشكل (20): التأخير في حال إرسال تطبيقين

يتبين من الشكل (20) أن زيادة إنتاجية الشبكة في حالة إرسال تطبيقين سببت زيادة التأخير على كلا الوصلتين الصاعدة والهابطة. وبمقارنة نتائج هذه الدراسة مع الأبحاث [10-11] اللذين يحددان متطلبات الأداء الرئيسية لشبكات 5G، نرى بأن قيم التأخير التي حصلنا عليها مقبولة وتقع ضمن المجال الذي حددته 3GPP لخدمات الجيل الخامس.

7. الخلاصة والأعمال المستقبلية

تمت دراسة بنية (MEC) في شبكة 5G، من خلال سيناريو موجود على Simu5G، وتمكننا من:

1. مراقبة كيفية عمل عناصر نظام (MEC) في شبكة 5G، ومعرفة كيفية تدفق الرسائل في تطبيق WarningAlertApp.
2. لوحظ أن هنالك عدد قليل من قفزات حزم السيارات إلى مضيفي (MEC)، مما يؤكد أن بنية (MEC) قادرة على معالجة المعلومات بشكل أكثر ديناميكية وكفاءة من MCC التي تستغرق أمانة وصول كبيرة وتسبب ازدحام على الشبكة الأساسية، وخاصةً عند التعامل مع التطبيقات ذات حجم كبير.

3. عند نمذجة تطبيق VOIP تم الحصول على أزمنة تأخير من رتبة ميلي ثانية وهي متطابقة مع معايير الأداء لشبكات الجيل الخامس والتي أشرنا إليها في الدراسة المرجعية.
4. تبين أن التغيير في حجم وعدد التطبيقات المرسله أدى إلى زيادة حركة المرور و زيادة التأخير في الشبكة.
5. تم نمذجة بنية شبكة تحتوي على UE واحد وgnb واحدة، سنقوم في الأبحاث القادمة بدراسة شبكة واقعية تحتوي على عدد من الأجهزة ومحطات القاعدة، والعمل على تحسين مؤشرات الأداء في هذه الشبكة.

References	
[1]	Zeqiri, R., Idrizi, F. & Halimi, H. Comparison of Algorithms and Technologies 2G, 3G, 4G and 5G. 2019 3rd International Symposium On Multidisciplinary Studies And Innovative Technologies (ISMSIT). pp. 1-4 (2019)
[2]	Pham, Q., Fang, F., Ha, V., Piran, M., Le, M., Le, L., Hwang, W. & Ding, Z. A Survey of Multi-Access Edge Computing in 5G and Beyond: Fundamentals, Technology Integration, and State-of-the-Art. IEEE Access. 8 pp. 116974-117017 (2020)
[3]	Taleb, T., Samdanis, K., Mada, B., Flinck, H., Dutta, S. & Sabella, D. On Multi- Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration. IEEE Communications Surveys Tutorials. 19, 1657-1681 (2017)
[4]	Cao, K., Liu, Y., Meng, G. & Sun, Q. An Overview on Edge Computing Research. IEEE Access. 8 pp. 85714-85728 (2020)
[5]	Hassan, N., Yau, K. & Wu, C. Edge computing in 5G: A Review. IEEE Access. 7 pp.2019
[6]	Liu, Y., Peng, M., Shou, G., Chen, Y. & Chen, S. Toward Edge Intelligence :Multiaccess Edge Computing for 5G and Internet of Things. IEEE Internet Of Things Journal. 7, 6722-6747 (2020)127276-127289 (2019)
[7]	Morris, I. ETSI drops “mobile” from (MEC) , 2016. Light Reading, New York, NY, USA, Sep. (2016)
[8]	Hu, Y., Patel, M., Sabella, D., Sprecher, N. & Young, V. Mobile edge computing—A key technology towards 5G. ETSI White Paper. 11, 1-16 (2015)
[9]	Anatel Perguntas e Respostas - 5G. (https://www.gov.br/anatel/ptbr/assuntos/5G/perguntas-e-respostas), Acessado em: Fevereiro/2022
[10]	ITU Draft new Report ITU-R M.[IMT-2020.TECH PERF REQ] - Minimum requirements related to technical performance for IMT-2020 radio interface(s).

	(https://www.itu.int/md/R15-SG05-C-0040/en) ، Acessado em: Mar, co/2022
[11]	U, E. ITU-R Workshop on IMT-2020 terrestrial radio interfaces. (https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020/Documents/S01-1_Requirements%20for%20IMT-2020_Rev.pdf)، Acessadoem: Mar, co/2022
[12]	OFCOM Enabling 5G in the UK. (https://www.ofcom.org.uk/spectrum/information/innovationlicensing/enabling-5g-uk)، Acessado em Fevereiro/2022
[13]	Li, Z., Uusitalo, M., Shariatmadari, H. & Singh, B. 5G URLLC: Design challenges and system concepts. 2018 15th International Symposium On Wireless Communication Systems (ISWCS). pp. 1-6 (2018)
[14]	Popovski, P., Trillingsgaard, K., Simeone, O. & Durisi, G. 5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View. IEEE Access. 6 pp. 55765-55779 (2018)
[15]	Filali, A., Abouaomar, A., Cherkaoui, S., Kobbane, A. & Guizani, M. Multi-Access Edge Computing: A Survey. IEEE Access. 8 pp. 197017-197046 (2020)
[16]	Antipolis, S. ETSI GS (MEC) 003 V2.2.1 (2020-12) - Multi-access Edge Computing ((MEC)); Framework and Reference Architecture. ETSI. 2.2.1 (2020)
[17]	Noferi, A., Nardini, G., Stea, G. & Viridis, A. Rapid prototyping and performance evaluation of (MEC) -based applications. ArXiv Preprint ArXiv:2203.13511. (2022)
[18]	Nardini, G., Sabella, D., Stea, G., Thakkar, P. & Viridis, A. Simu5G—An OMNeT++ Library for End-to-End Performance Evaluation of 5G Networks. IEEE Access. 8 pp. 181176-181191 (2020)

[19]	Nardini, G., Stea, G. & Viridis, A. Simu5G - 5G NR and LTE/LTE-A userplane simulation model for OMNeT++ and INET. (Disponível na Internet via: https://github.com/Unipisa/Simu5G/tree/master/simulations/NR/(MEC)/multi(MEC) Host) ,Acessado em: Junho/2022
[20]	Thanh Van Le et al. "Mobile edge cloud computing simulation tool survey and literature reviews". In: (). URL: https://www.researchgate.net/profile/Van-Thanh-Le/publication/344404205_Mobile_edge_cloud_computing_simulation_tool_survey_and_literature_reviews/links/5f71e6f1a6fdcc008643914a/Mobile-edge-cloud-computing-simulationtool-survey-and-literature-reviews.pdf
[21]	Ahmed, A., Al-Dweik, A., Iraqi, Y., Mukhtar, H., Naeem, M. & Hossain, E. Hybrid Automatic Repeat Request (HARQ) in Wireless Communications Systems and Standards: A Contemporary Survey. IEEE Communications Surveys And Tutorials. 23,2711-2752 (2021)

دراسة عملية للتآكل الخارجي في أنابيب نقل النفط تبعاً لعوامل مختلفة

م. طارق عياشي * د.م. أحمد سلامة ** د.م. حسان حامد ***

المخلص:

يهدف البحث إلى دراسة الحالة التآكلية لعينات فولاذية من أنبوب نقل النفط الواصل بين مصبي بانياس وطرطوس النفطيين، في عدة مواقع من الخط، بطريقة عملية اعتماداً على طريقة الوزن المفقود، مع الأخذ بعين الاعتبار أهم العوامل المؤثرة على التآكل الخارجي من طرق موصلات ذات اتجاه واحد وسكك حديدية وأنهار وخطوط توتر عالي ومتوسط، بالإضافة إلى الخصائص الفيزيائية والكيميائية للتربة.

بينت الدراسة أن للعوامل الخارجية المذكورة تأثيراً كبيراً على التآكل، فكلما كانت النقاط أقرب إلى تلك العوامل، زادت نسبة التآكل، وكلما ازدادت نسبة الكلوريد والكربونات في التربة زادت نسبة التآكل، كما لوحظ في الدراسة وجود موقع يبدي خطورة تآكلية كبيرة، يجب مراقبته بشكل دوري، واتخاذ الإجراءات المناسبة للحماية من التآكل، من عزل بمواد خاصة تعطي حماية أكثر فعالية، أو رفع شدة الحماية المهبطية.

كلمات مفتاحية: التآكل، أنابيب النفط، طريقة الوزن المفقود.

* طالب دكتوراه - كلية الهندسة الميكانيكية والكهربائية - جامعة تشرين - اللاذقية - سوريا.

** أستاذ مساعد - كلية الهندسة الميكانيكية والكهربائية - جامعة تشرين - اللاذقية - سوريا.

*** أستاذ مساعد - كلية الهندسة الميكانيكية - جامعة البعث - حمص - سوريا.

Practical study of external corrosion in oil transport pipelines depending on different factors

Tarek Aiashee*

D. Ahmad Salamah**

D. Hassan Hamed ***

ABSTRACT

The research aims to study the corrosion state of steel samples from the oil transport pipeline connecting the Baniyas and Tartous oil estuaries, in several locations of the line, in a practical way depending on the lost weight method, taking into account the most important factors affecting external corrosion such as one-way transportation roads and railways. Iron, rivers, high and medium voltage lines, in addition to the physical and chemical properties of the soil.

The study showed that the aforementioned external factors have a significant impact on erosion. The closer the points are to those factors, the greater the erosion rate, and the higher the chloride and carbonate ratio in the soil, the higher the erosion rate. It was also noted in the study that there is a site that shows great erosive risk, and must be monitored periodically. And taking appropriate measures to protect against corrosion, by isolating with special materials that give more effective protection, or raising the intensity of cathodic protection.

Keywords: Corrosion, pipelines, lost weight method.

* **Doctoral student; Faculty of Mechanical & Electrical Engineering; Tishreen University; Lattakia: Syria.**

** **Professor Assistant; Faculty of Mechanical & Electrical Engineering; Tishreen University; Lattakia: Syria.**

*** **Professor Assistant; Faculty of Mechanical & Electrical Engineering; Al-Baath University; Homs: Syria.**

مقدمة:

إن التكاليف الباهظة جداً الناتجة عن التآكل، تعد من أهم الأسباب التي تحث على دراسة التآكل، ومحاولة التحكم فيه، والحد من آثاره الضارة، فقد قدرت الخسائر السنوية الناتجة عن مشكلة التآكل بحوالي 82 بليون دولار في الولايات المتحدة الأمريكية وحدها عام 1975، بسبب ما يسببه التآكل من تلف المعادن وسبائكها، وقد أجريت عدة أبحاث في الدول المتقدمة لتقدير تكلفة التآكل على الاقتصاد القومي، وكانت أول دراسة في بريطانيا عام 1970م، حيث قدرت نسبة الخسارة والتكلفة الناتجة عن التآكل بنحو 2.4-4% من الدخل القومي، كما قدرت كمية الحديد التي أتلها التآكل في بلجيكا بربع أو ثلث الإنتاج السنوي، كما أجريت عدة دراسات في مصر وتركيا وكانت النتائج مشابهة للدراسات التي سبقتها، حيث قدرت التكلفة الناتجة عن التآكل بنحو 4-5% من الدخل القومي، وكذلك تصرف الدول الكثير من الأموال للحماية من التآكل وإجراء عمليات الصيانة الوقائية للمنشآت والأجهزة المعدنية، فعلى سبيل المثال تصرف شركة نفط الكويت 300 ألف دينار كويتي سنوياً، لصيانة الأجهزة فقط من آثار التآكل الضارة على الصناعات المختلفة [1،2].

يعرف التآكل في أنابيب النفط على أنه عملية التلف أو النخر أو الاهتراء الذي يصيب الأنبوب من الداخل أو الخارج بفعل الوسط المحيط، ويؤدي إلى تخريبه وتقليل سماكته والإساءة إلى خواصه الكيميائية والميكانيكية، وإلى إحداث ثقب فيه تخرجه عن التشغيل والخدمة في وقت مبكر [1].

العوامل المؤثرة على تآكل أنابيب النفط:

هناك الكثير من العوامل التي تزيد من التآكل في أنابيب النفط نذكر منها:

1- المسيلات المائية: إن وجود المسيلات المائية بالقرب من أنابيب النفط، يؤدي إلى زيادة الرطوبة في التربة وزيادة البكتريا واختلاف درجة الحموضة (PH) في الوسط المحيط بالأنبوب، مما يؤدي إلى زيادة التآكل في أنبوب النفط، وفيما يلي نعرض تأثير كل منها:

➤ **قيمة pH الوسط:** إن تركيز شوارد الهيدروجين الذي تعبر عنه درجة الحموضة pH هو أحد العناصر الرئيسية لنشاط الوسط المحيط وتأثيره على العملية الكهروكيميائية لتآكل الفولاذ، حيث أن درجة الحموضة تؤثر على التفاعلات المهيطة (الكاثودية) في حين يقل تأثيرها على التفاعلات المصعدية (الأنودية)، كما تزداد سرعة الأكسدة بانخفاض قيمة pH الوسط المحيط إلى مادون (4) ($PH < 4$)، وعندما تتراوح قيمتها بين (4-9)، فإن سرعة الأكسدة تبقى ثابتة، إلا أنه في وسط قلوي $pH=14$ يلاحظ زيادة في نشاط الأكسدة بسبب انحلال نواتج الأكسدة في المحاليل الكهروكيميائية [2,3].

➤ **التآكل بسبب البكتريا:** هو التآكل الناتج عن نشاط حيوي، فمن المعروف أن هناك كثير من الأحياء الدقيقة من طحالب وأشنيات وبكتريا توجد في الهواء والماء والتربة، وإن لهذه الأحياء نشاطات حيوية، فهي تتغذى على المواد العضوية واللاعضوية وتنتج أحماضاً أو قواعد تكون أوساط أكالة للمعدن، حيث تسبب تآكل الطبقة العازلة (الطلاء) على الأنابيب، فتتحول إلى مصعد (أنود) على حساب الطبقة العازلة المجاورة (كاثود) [2,4].

➤ **تأثير الرطوبة وفرق تركيز الأملاح التربة:** إن محتويات التربة من أملاح ورطوبة ومعادن ومكونات مختلفة، بما في ذلك الأحياء الدقيقة المتواجدة تحت الأرض، مثل البكتريا اللاهوائية، تؤدي إلى تآكل الأنابيب والمعدات والتراكيب المدفونة في التربة، وتسبب ما يسمى بالتآكل تحت الأرضي (الجوفي)، حيث يمكن اعتبار الأرض الرطبة كمحلول إلكتروليتي موصل للإلكترونات والأيونات، ومعظم التآكلات الأرضية هي من نوع التآكلات الكهروكيميائية، هذا يعني بأن التآكل يحصل بسبب وجود أنود وكاثود، حيث أن المناطق الأنودية والكاثودية على سطح المعدن لا تحصل بسبب الإجهادات المسلطة على المعدن، أو بسبب اختلاف معدنين في محيط واحد فقط،

ولكن قد تحصل بسبب المحيط الخارجي الآكل واختلاف مواصفاته من منطقة لأخرى، مثل اختلاف نوعيّة الأيونات أو تركيزها في منطقتين، مسبباً ما يسمّى باستقطاب التّركيز Concentration Polarization، أو بسبب اختلاف التوصيل الكهربائي للأرض من منطقة إلى أخرى مسبباً استقطاب المقاومة Resistance Polarization.

وهذا يعني لو مدّ أنبوب مغلف مدفون في أرض، جزء منها جاف رملي، والآخر رطب رملي، فإنّ التآكل لا يحصل في الجزء الجاف بل في الجزء الرطب، وذلك لأنّ الأرض الرطبة موصلة للكهرباء، أمّا الجافة فهي غير موصلة بشكل جيد، أو عند مرور الأنبوب في أرض رسوبيّة وأرض رملية جافة، فإنّ الموصليّة الكهربائيّة، وبالتالي التآكل في الأرض الرسوبيّة الطينيّة، أعلى من الأرض الجافة الهشّة، وعليه قد تكون الموصليّة الكهربائيّة للتربة، أهمّ مؤشر على حجم التآكل الحاصل للأنابيب المدفونة [5].

2- تأثير التيارات الكهربائيّة الخارجيّة: يحصل هذا التآكل بسبب تيارات التحريض

الكهربائيّة المتولدة من حركة الآليات (سيارات وقطارات) على سطح الأرض باتجاه واحد، لأنها تؤثر على استقطابيّة المعدن داخل الأرض، بسبب التحريض الكهربائي المتولّد من الحركة النسبية لجسم خارجي على المعدن المدفون، فعند حركة جسم معدني (سيارات أو قطارات) باتجاه واحد فقط على طريق يحوي أنبوب معدني مدفون تحته، فإن الكثافة الإلكترونيّة في نهاية الخط، ستكون أعلى منها في بداية خط الأنبوب بسبب التحريض الذي يولده الجسم المتحرك (بشكل مشابه لإنتاج المغناطيس بالتحريض) وبذلك تتصرف نهاية الخط كمهبط وبدايته كمصدر [6].

3- التآكل بسبب الحقول الكهرطيسية الناتجة عن خطوط التوتر العالي: لقد أثبتت

الدراسات أن الحقول الكهرطيسية الناتجة عن خطوط الكهرباء ذات التوتر العالي تؤدي إلى تسريع حركة الشوارد في الأنبوب، مما يؤدي إلى تشكل مناطق مصعدية

ومناطق أخرى مهبطية، مما يؤدي إلى تآكل المعدن. وعند تصميم أنابيب النفط يجب الابتعاد عن خطوط الكهرباء مسافة 1cm لكل فولط كهرباء، أي بالنسبة لخطوط التوتر العالي يجب الابتعاد عنها حوالي 500m [7].

4- تآكل الوصلات اللحامية: يحدث فيها تآكل يدعى تآكل موضعي، وذلك بسبب وجود نقاط تحمل الصفة الأنودية بالنسبة لبقية السطح، وهي من الأنواع الخطرة، لأنها تعطي فكرة مغلوطة عما تبقى من سمك المعدن [8].

الدراسة المرجعية:

1- وفي عام 2006، درست المهندسة الصياح وآخرون تأثير خطوط التوتر العالي على تآكل الفولاذ الكربوني في التربة، حيث قاموا بزراعة عينات في منطقة يوجد فيها أبراج توتر عالي بتوتر 230 كيلو فولط، ومن ثم قياس مقدار التآكل بطريقة الفقد بالوزن، فتبين أنه يوجد علاقة خطية بين فرق الوزن والزمن، حيث يزداد فرق الوزن بازدياد الزمن، وهذا يعني حدوث تآكل مستمر للعينات بمعدل ثابت مع مرور الزمن، وبالتالي يزداد التآكل نتيجة التيارات الناتجة عن خطوط التوتر العالي [9].

2- في عام 2009، اقترح Velázquez وآخرون نموذج تنبؤي للتآكل النقطي في أنابيب النفط والغاز المدفونة في التربة، حيث أخذوا بعين الاعتبار الخصائص الفيزيائية والكيميائية للتربة والأنبوب للتنبؤ بعمق ومعدل نقطة التآكل [10].

3- وفي عام 2013، قامت شركة Petro Min Pipeliner بنشر بحث عن تقنيات تفتيش الأنابيب المستخدمة لكشف وتمييز الضرر والتلف والتي تسمى عموماً pigs، والتي تستخدم لفحص الأنابيب من الداخل حيث توضع داخل الأنبوب ويتم دفعها من خلال تدفق المنتج، وقد تبين أن هذه التقنيات لها دور كبير في كشف التآكل والعيوب في الأنابيب بالإضافة إلى إزالة العوالق والأكاسيد داخل الأنابيب [11].

4- وفي عام 2014، أجرى Hui Wang وآخرون عملية نمذجة باستخدام النظرية الافتراضية (Bayesian Modeling) للتآكل الخارجي في خطوط الأنابيب المدفونة تحت الأرض على أساس تكامل تقنيات مونتي كارلو، سلسلة ماركوف وبيانات التنقيش المجمع، فتوصل الباحثون إلى منهجية جديدة لتصنيف وتحليل عيوب التآكل الخارجية باستخدام نتائج التنقيش الفني والخصائص الفيزيائية والكيميائية للتربة [12].

5- وفي عام 2015، قام الدكتور حامد والمهندسة رجوب بدراسة الحماية الكاثودية لأنابيب نقل النفط الخام، حيث وضعت عينات محمية وعينات غير محمية في وسط التآكل المتمثل بموقعين من تربة محافظة حمص (تربة مدينة حمص- تربة منطقة الفرقلس)، ولقد أظهرت النتائج أن النماذج غير المحمية أعطت أقل مقاومة للتآكل مقارنة بالنماذج المحمية، كما لوحظ زيادة معدل التآكل في الموقع الذي يحوي نسبة عالية من الكلوريد في التربة [13].

6- في عام 2016، قام Bolaji و Nosa بتصميم نموذج لحساب معدل التآكل في الأنابيب المدفونة في التربة، باستخدام الشبكات العصبونية الاصطناعية المعدلة ومحاكاة مونتي كارلو. المؤشرات المقاسة هي PH التربة، درجة حرارة التربة، درجة حرارة الجو لترب مختلفة مأخوذة من موقع إنتاج النفط في ولاية دلتا في نيجيريا. تم حساب معدل التآكل لكل حالة باستخدام طريقة الفقد بالوزن [14].

7- في عام 2017، قام Oladipo وآخرون بتقييم التآكل لبعض الأنابيب المعدنية المدفونة في التربة باستخدام خوارزمية الشبكات العصبونية، وتم أخذ منطقة النفط والغاز في نيجيريا كحالة دراسية، كانت بيانات الإدخال درجة حرارة التربة، ودرجة حرارة المحيط المباشر، ودرجة حموضة التربة (pH)، في حين أن بيانات الخرج كان الفقد بوزن المعدن، فتوصلوا إلى أن درجات حرارة التربة والظروف المحيطة ودرجة

الحرارة المحيطة مباشرة لها تأثير أقوى على معدل انتشار التآكل، في حين أن درجة الحرارة المحيطة المباشرة وقيمة الـ (pH) مجتمعاً كان تأثيرها أقل على معدل انتشار التآكل [15].

8- في عام 2017، أوجد الباحثان Yang و Oterkus طريقة لكشف التآكل في خطوط الأنابيب، بناءً على قياسات الترددات الطبيعية، حيث قدما طريقة جديدة لتحديد موقع ونسبة التآكل وذلك عن طريق قياس الترددات الطبيعية الأولى والثانية لخطوط الأنابيب المتضررة، من ثم تحققاً من صحة النتائج باستخدام طريقة العناصر المنتهية [16].

9- في عام 2018، أجرى Tian و Xie مراجعة حول إدارة تكامل خطوط الأنابيب باستخدام بيانات التفتيش المباشر، حيث تم استعراض تقنيات التفتيش الرئيسي ILLI مع أدائها وتطبيقاتها، كما تم تقديم طرق ونماذج الفحص والتفتيش المستخدمة لكشف المخاطر [17].

10- كما أجرى Sumayh وآخرون دراسة بعنوان نموذج لكشف الشذوذ في خطوط أنابيب النفط والغاز باستخدام تقنيات التعلم الآلي في المملكة العربية السعودية، في عام 2022، حيث قاموا بإجراء دراسة لخمس نماذج للتعلم الآلي، تستخدم في كشف أماكن الضعف والتسرب في أنابيب النفط والغاز، وقاموا بتجربة اثنين منها على أرض الواقع، فأعطت هذه النماذج دقة عالية وصلت إلى 97.4%، في تحديد أماكن الضعف والتسرب في أنابيب النفط والغاز [18].

نستنتج من الدراسة المرجعية السابقة أن هناك كثير من الدراسات حول دراسة الحالة التآكلية لأنابيب النفط والغاز، فهناك كثير من الباحثين اهتموا بكشف أماكن التآكل الداخلي في الأنابيب، واستخدموا لذلك العديد من التقنيات الحديثة، مستفيدين من بيانات

التفتيش الفني، وبعضهم اهتم بدراسة التآكل الخارجي، آخذين خصائص التربة فقط كعوامل مسببة للتآكل، ولم يأخذوا باقي العوامل في عين الاعتبار.

ومن هنا كان الاهتمام بالبحث الحالي بدراسة الحالة التآكلية لأنابيب النفط باستخدام طريقة الفقد بالوزن، مع الأخذ بعين الاعتبار أكبر عدد ممكن من العوامل المؤثرة على التآكل الخارجي.

أهمية وهدف البحث:

- 1- أهمية البحث: تأتي أهمية هذا البحث، من أن الدراسة العملية للتآكل الخارجي في أنابيب النفط، سيؤدي إلى تحقيق الإدارة السليمة لتآكل أنابيب النفط، من خلال إجراء الصيانة الوقائية، مما يقلل من التكاليف الناتجة عن التلف والاهتراء، وما يتبعها من عمليات حفر وتوقف الانتاج، وبالتالي التقليل من الهدر في الأموال والجهد والوقت.
- 2- هدف البحث: يهدف هذا البحث إلى تحقيق دراسة عملية للتآكل الخارجي في أنابيب النفط في تربة مختلفة في خصائصها الفيزيائية والكيميائية، ودراسة مدى تأثير العوامل الخارجية المسببة للتآكل الخارجي عليها باستخدام طريقة الفقد بالوزن.

مواد البحث:

- ✚ عينات فولاذية من أنبوب النفط الواصل بين مصبي بانياس وطرطوس النفطيين، وهي من نوع X52.
- ✚ ميزان الكتروني حساس دقته 0.001g.
- ✚ برنامج الاكسل Excel.

النتائج والمناقشة:

حالة دراسية: أنبوب نقل النفط الواصل بين مصبي بانياس وطرطوس

النفطيين

يصل الانبوب المدروس بين مصبي بانياس وطرطوس النفطيين بطول 35 كم، وهو مدفون في التربة على عمق 1م، وقطره 24 إنش، وهو مصنوع من مجموعة من الأجزاء طول كل جزء منها حوالي 12 متر. الأنبوب مصمم حسب مواصفات ASTM، وهو من الفولاذ الكربوني تركيبه الكيميائي (من ASTM) موضح في الجدول (1).

الجدول (1): التركيب الكيميائي للعينات المدروسة

الكبريت (S)	الفوسفور (P)	المنغنيز (Mn)	الكربون (C)	طريقة صنع الأنابيب	ASTM Spec
0.03	0.03	1.25	0.28	لحام على البارد	5L X52

تم دفن عينات من أنبوب النفط المدروس في التربة، في أربع مواقع من منطقة الدراسة هي: القصور - الباصية - الرمال الذهبية - المحطة الحرارية، حيث وصلت على علب الحماية المهبطية الموجودة على طول الخط، لتخضع لنفس أسلوب الحماية المستخدم على الأنبوب، ودفنت العينات في التربة المختلفة لهذه المواقع، في نفس الظروف التي يتعرض لها الأنبوب المدروس، وذلك لمدة عامين، وتم حساب التآكل بطريقة الفقد بالوزن (الطريقة الوزنية)، عن طريق حساب الفرق بين وزن العينات قبل وبعد الاختبار، وتبين الفقرات التالية خطوات العمل:

• الحصول على العينات وتجهيزها:

تم الحصول على قطع من الأنبوب المدروس، من الشركة السورية لنقل النفط في بانياس، وقصها وتقطيعها إلى قطع صغيرة على شكل متوازي مستطيلات، ومن ثم جلقها وتنظيف سطوحها، كما وضع ثقب في كل منها، كما تم تجهيز قاعدة من معدن الأنبوب نفسه بطول متر تقريباً وعرض 12 cm، وقد نفر منها 8 قضبان من نفس المادة، تم

لحامها على القاعدة، لوضع العينات، بحيث يتم أخذ 8 قراءات من كل موقع تلافياً لأي خطأ، كما وصل كبل نحاسي معزول بطول متر إلى القاعدة، وذلك لوصل العينات مع علب الحماية المهبطية الموجودة على طول الخط المدروس، وقسمت العينات إلى مجموعات، يتم تمييزها بأحرف انكليزية A و B و C و G ، ولكل حرف 8 أدلة، مثلاً A1 و A2 و A3 و و A8، ويظهر الشكل (1) العينات بعد أن تم قصت وجهزت، ويظهر الشكل (2) القاعدة مع الكبل النحاسي المعزول والعينات.



الشكل (1): العينات بعد أن تم قصها وتجهيزها



الشكل (2): القاعدة الحديدية مع الكبل النحاسي المعزول والعينات

كما قيست أبعاد كل العينات وحساب مساحة سطوحها، ويظهر الجدول (2) أبعاد العينة A ومساحة سطحها.

الجدول (2): أبعاد العينة A ومساحة سطحها

العينة	العيونة طولcm	العرضcm	الاسماعةcm	السطح مساحة للتآكل المعرض (cm2)
A1	5.9	3.87	0.792	61.14168
A2	5.6	3.26	0.792	50.54624
A3	5.9	3.84	0.792	60.74016
A4	5.7	3.32	0.792	52.13568
A5	5.6	3.2	0.792	49.7792
A6	5.5	3.13	0.792	48.09992
A7	5.7	3.54	0.792	54.99216
A8	5.6	3.33	0.792	51.44112

- **وزن العينات قبل الاختبار:** بعد تحضير العينات وترقيمتها، تم وزن العينات باستخدام ميزان دقيق يعطي أوزان من رتبة أجزاء من الألف من الغرام (ثلاثة أرقام بعد الفاصلة)، وهو موجود في الشركة السورية لنقل النفط في بانياس، وهو من نوع BEL، ويظهر الشكل (3) طريقة ونتائج وزن بعض العينات.



الشكل (3): نتائج أوزان بعض العينات

- **تحليل التربة:** أخذت عدة مقاطع من التربة على طول منطقة الدراسة، حيث تم الحفر في مواقع مختلفة من منطقة الدراسة على عمق 70 cm تقريباً، وأخذت عينات من التربة من هذه المواقع، وتم تسميتها بأحرف مختلفة: A و B و C و G، لتمييزها عن بعضها أثناء تحليلها ومعرفة خصائصها.

وتم تحليل مقاطع التربة في مخبر مركز البحوث العلمية الزراعية في محافظة اللاذقية للحصول على خصائصها الفيزيائية والكيميائية، وهي: درجة الحموضة PH- الناقلية الكهربائية للتربة (Electrical Conductivity (EC) (ds/m) - درجة الملوحة (Salinity) - تركيز الكربونات (CO_3 (mm/100g) - تركيز الكلوريد (CL (mm/100g)) - نسبة الرطوبة في التربة (Humidity)، وجاءت النتائج كما هو مبين في الجدول (3).

الجدول (3): نتائج تحليل التربة للعينات المأخوذة من مناطق مختلفة

موقع أخذ العينة	البيانات	1:5معلق		م.م/100 غرام تربة		%
		PH	EC ds/m	CO_3^{-2}	Cl ⁻	
القصور	A	7.63	0.94	25	0.04	10.7
		قاعدية	قليل الملوحة			
الباصية	B	7.57	0.18	20	0.05	10.9
		قاعدية	غير مالحة			
الرمال الذهبية	C	7.6	0.22	16	0.03	9.8
		قاعدية	غير مالحة			
المحطة الحرارية	G	7.61	0.56	23	0.05	9.3
		قاعدية	غير مالحة			

• زراعة العينات:

تم زرع العينات بالتعاون مع الفنيين في الشركة السورية لنقل النفط، حيث تم اختيار أربعة مواقع بالقرب من العوامل المدروسة المؤثرة على التآكل الخارجي للأنبوب، وهي: طرق المواصلات ذات الاتجاه الواحد، السكك الحديدية، الأنهار، خطوط التوتر المتوسط، خطوط التوتر العالي، بالإضافة إلى محتويات التربة.

بحيث تشمل هذه المواقع جميع العوامل المدروسة، ويظهر الجدول (4) بيان للنقاط والعوامل المدروسة في كل نقطة.

الجدول (4): بيان للنقاط والعوامل المؤثرة على التآكل في كل نقطة

موقع النقطة	رمز النقطة	العوامل المسببة للتآكل	العينات المحمية
القصور	A	خط توتر متوسط - طريق	-A5 -A4 -A3-A2 -A1 A8 -A7 - A6
الباصية	B	البحر- توتر عالي- طريق - سكة حديد	-B5 -B4 -B3-B2 -B1 B8 -B7 - B6
الرمال الذهبية	C	نهر	-F5 -F4 -F3-F2 -F1 F8 -F7- F6
جانب المحطة الحرارية	G	توتر عالي - توتر متوسط - طريق - سكة حديد	-L5 -L4 -L3-L2 -1L L8 -L7 - L6

حيث تم الحفر في المناطق المذكورة سابقاً، لعمق 1 متر تقريباً، وكانت جميع المواقع بالقرب من علب الحماية المهبطية، الموجودة على الخط المدروس، وذلك كي يتم ربط العينات المزروعة بمنظومة الحماية المهبطية التي تحمي أنبوب النفط المدروس من التآكل، ويظهر الشكل (5) العينات بعد أن وصلت مع علب الحماية المهبطية.



الشكل (5): العينات بعد أن وصلت مع علب الحماية المهبطية

• إزالة العينات ووزنها:

تم إزالة العينات على أربع مراحل خلال عامين من الزراعة، بعد ستة أشهر، وبعد عام، وبعد عام ونصف، وأخيراً بعد مضي عامين، وفي كل مرحلة كان يتم إزالة عينتين، وذلك لضمان نجاح التجربة، والاطلاع على أن التجربة تسير في الاتجاه الصحيح. وفي كل مرحلة تم فيها إزالة العينات من التربة، كان يتم تنظيفها من التربة والعوالق ونواتج التآكل، وذلك كما يلي [19]:

1. غسل العينات بالماء لإزالة التربة عنها.
 2. تجفيف العينات بواسطة أشعة الشمس.
 3. غسل العينات بحمض كلور الماء الممدد بنسبة (1:3) باستخدام فرشاة بلاستيكية.
 4. غسل العينات بالماء لإزالة آثار الحمض.
 5. تجفيف العينات مرة أخرى بشكل جيد بوضعها في الشمس.
- وبعد عملية التنظيف يتم وزن العينات بواسطة الميزان الإلكتروني الذي دقته 0.001gr، ثم يتم حساب نسب التآكل.

• نتائج تآكل العينات بعد عامين من الاختبار:

تم حساب فرق الوزن بالعلاقة:
 $\Delta w = w_1 - w_0$ حيث w_0 هو وزن العينات قبل الاختبار و w_1 هو وزن العينات بعد عامين من الاختبار.

كما حسب معدل التآكل مقاساً بـ ملغ/سم² خلال عامين، وذلك بالعلاقة:

$$\text{معدل التآكل} = \frac{1000 \times \Delta w}{S}$$

حيث: Δw فرق الوزن والذي يدل على الكمية المتأكلة، S مساحة السطح المعرض للتآكل.

ثم حسب معدل سماكة (عمق تغلغل) التآكل مقاساً بالسنتيمتر خلال عامين، وذلك بالعلاقة:

$$\text{معدل سماكة التآكل} = \frac{\text{معدل التآكل}}{1000 \times 7.85}$$

حيث 7.85 تمثل كثافة معدن الأنابيب.

ثم حسب معدل سماكة التآكل مقاساً بالميليمتر خلال عامين، وذلك بضرب العلاقة بالعدد 10.

ثم حساب النسبة المئوية للتآكل بالعلاقة:

$$\text{النسبة المئوية للتآكل} = \frac{\text{معدل سماكة التآكل}}{\text{سماكة الانبوب}} \times 100$$

ثم تم حساب المتوسط الحسابي للنسبة المئوية لتآكل العينات الموجودة في نفس الموقع، لتمثل النسبة المئوية النهائية للتآكل في ذلك الموقع. ويظهر الجدول (5) نتائج تآكل العينات المزروعة بعد عامين.

الجدول (5): نتائج تآكل العينات المدفونة في التربة بعد عامين

المتوسط الحسابي لنسبة التآكل (%)	نسبة التآكل (%)	سماكة معدل التآكل (مم/ سنتين)	سماكة معدل التآكل (سم/ سنتين)	معدل التآكل (ملغ /سم ²)	الوزن فرق (g)	مساحة السطح (cm ²)	بعد الوزن من عامين (g)	قبل الوزن (g)	العينات	الموقع
32.00826	31.97403	2.53234279	0.25323428	32.25914	1.774	54.99216	123.547	125.321	A7	القصور
	32.04249	2.53776551	0.25377655	32.32822	1.663	51.44112	114.435	116.098	A8	
55.30704	55.64583	4.40714955	0.44071495	56.14203	3.102	55.25272	122.675	125.777	B7	البياضية
	54.96826	4.35348585	0.43534858	55.45842	2.54	45.80008	98.558	101.098	B8	
27.50159	27.08348	2.14501133	0.21450113	27.32499	1.62	59.2864	117.216	118.836	C7	الذهبية الزمال
	27.9197	2.21123985	0.22112399	28.16866	1.571	55.7712	109.387	110.958	C8	
66.29578	66.32051	5.25258423	0.52525842	66.9119	3.061	45.74672	98.204	101.265	L7	الحرارية المحطة
	66.27105	5.24866707	0.52486671	66.862	3.726	55.72672	123.708	127.434	L8	

ويتضح من الجدول (5) أنه بعد دفن العينات لمدة عامين في التربة، في عدة مواقع، وربطها بالحماية المهبطية الموجودة على خطوط أنابيب النفط، كانت نسب التآكل كما في الجدول (6).

الجدول (6): نسب التآكل بعد دفن العينات في التربة لمدة عامين

المحطة الحرارية	الرمال الذهبية	الباصية	القصور	الموقع	البعد عن (m)
G	C	B	A	النقطة	
103.681	586.850	37.6647	210.273	الطرق	
269.320	757.780	632.325	1040.357	السكك الحديدية	
68.9907	1895.68	1199.86	554.103	خطوط التوتر المتوسط	
85.5596	4504.27	1467.581	1227.061	خطوط التوتر العالي	
723.398	110.155	19.39635	585.7498	الانهار	
7.61	7.6	7.57	7.63	pH	
0.56	0.22	0.18	0.94	الناقلية الكهربائية (ds/m)	
قليل الملوحة	غير مالحة	غير مالحة	قليل الملوحة	الملوحة	
23	16	20	25	تركيز الكربونات (mm/100g)	
0.05	0.03	0.05	0.04	تركيز الكلوريد (mm/100g)	
9.3	9.8	10.9	10.7	نسبة الرطوبة %	
66.295	27.501	55.307	32.008	نسبة التآكل %	

مناقشة النتائج:

بعد عملية زراعة العينات في أربع مواقع مختلفة من منطقة الدراسة وحساب نسب التآكل في كل منها، كانت النتائج على النحو الآتي:

- ❖ **منطقة المحطة الحرارية:** نجد فيها أعلى نسبة تآكل (66.295%)، أي أن قسم كبير من سماكة العينات قد تعرض للنخر والاهتراء، وذلك بسبب قرب المنطقة من عدد كبير من العوامل التي تسبب زيادة في التآكل الخارجي، وهي الطرق (103.68 m) والسكك الحديدية (269.32 m) وخطوط التوتر العالي (85.5 m) والمتوسط (68.9 m)، بالإضافة إلى خصائص التربة الأكاله، حيث امتازت التربة في هذه المنطقة بناقلية كهربائية متوسطة نسبياً (0.56 ds/m) ، مما يزيد من سرعة التآكل، كما أن نسبة الكلوريد والكربونات عالية نسبياً، مما يسبب زيادة في سرعة التآكل.
- ❖ **في منطقة القصور:** كانت نسبة التآكل 32%، ويمكن تفسير هذه النسبة بسبب القيمة العالية للناقلية الكهربائية للتربة بالإضافة إلى نسبة الرطوبة العالية فيها، كما أن لزيادة نسبة الكربونات تأثير أيضاً، بالإضافة إلى وجود طرق موصلات باتجاه واحد بالقرب من العينات A، حيث يبعد الطريق عنها حوالي 210 متر، مما يسبب زيادة في حدوث التآكل.
- ❖ **في منطقة الباصية:** كما نجد أن اجتماع طرق الموصلات ذات الاتجاه الواحد مع الأنهار يؤدي إلى زيادة نسبة التآكل في منطقة الباصية بشكل كبير، بالإضافة إلى نسبة الرطوبة العالية في التربة، فكانت نسبة التآكل في العينة B المزروعة في منطقة الباصية (55.3%).
- ❖ **في منطقة الرمال الذهبية:** تناقصت نسبة التآكل في العينات C المزروعة في منطقة الرمال الذهبية إلى (27.501%)، بسبب تناقص نسبة الكلوريد والكربونات في التربة، ويعزى حدوث التآكل في العينة إلى القرب من مياه الأنهار، التي تؤدي إلى زيادة التآكل.

الاستنتاجات والتوصيات:

- 1- طريقة الوزن المفقود في حساب نسب التآكل، طريقة فعالة ومجدية وتساعد في معرفة النقاط الخطرة من أنبوب النفط.
- 2- منطقة المحطة الحرارية في المنطقة المدروسة، منطقة تآكلية عالية الخطورة، بسبب اجتماع عوامل كثيرة مسببة للتآكل فيها، يجب مراقبتها بشكل دائم، واتباع أكثر من طريقة حماية فيها، من عزل أو رفع شدة تيار الحماية المهبطية أو غير ذلك.
- 3- تفيد نتائج دراسة تأثير العوامل المختلفة التي تزيد من خطر تآكل أنابيب نقل النفط والغاز، في تحديد المسارات الأفضل لمد هذه الخطوط مستقبلاً بعيداً عن العوامل المختلفة المسببة للتآكل.
- 4- كلما زادت نسبة الكلوريد والكربونات في التربة كانت التربة أكالة بشدة، لذلك يجب زيادة كمية العزل لأنابيب التي تمر من الأراضي ذات التربة الغنية بهذه العناصر.

المراجع : References

- 1- BADMOS, A, AJIMOTOKAN, H, EMMANUEL, E 2009 **Corrosion In Petroleum Pipelines**. New York Science Journal, ISSN 1554-0200, Vol.2, NO.5, 36-40.
- 2- ALAHMAD, N, ALHAZOURY, A, AMMAR, S 2011 **Metal Technology**. Equipment and Machinery Engineering, Technical Engineering, Tishreen University Publications, Syria. (In Arabic)
- 3- PROTOPOPOFF, E, MARCUS, P 2003 Potential versus pH (Pourbaix) Diagrams, **Corrosion: Fundamentals, Testing, and Protection**, Volume 13A of the ASM Handbook. 50-52.
- 4- DEXTER, C 2003 Microbiologically Influenced Corrosion, **Corrosion: Fundamentals, Testing, and Protection**. Volume 13A of the ASM Handbook. 1022-1051.
- 5- BAKER, M, FESSLER, R, CONSULTING, B 2008 **Pipeline Corrosion, FINAL REPORT**. U.S. Department of Transportation Pipeline and Hazardous Materials Safety Administration Office of Pipeline Safety, 14-25.
- 6- KOCH, H, BRONGERS, M, THOMPSON, N 2003 **Direct Costs of Corrosion in the United States**. Corrosion: Fundamentals, Testing, and Protection, Volume 13A of the ASM Handbook. 2442-2463.
- 7- MA, J, DAWALIBI, F; SOUTHEY, R 1997 **Computation and Measurement of Electrical Interference Effects in Aqueducts Due to a Nearby Parallel Transmission Line**, International Symposium on Electromagnetic Compatibility, Beijing, 1997.
- 8- FRANKEL, G 2003 **Corrosion Science in The 21st Century**. Journal of Corrosion Science and Engineering, Vol. 6, pp. 1-15.
- 9- ALOMAR, L, ALSAYAH, R, HAMED, H, KARFOUL, M 2006 **A study of the effect of high tension lines on corrosion of carbon steel in the soil**. Unpublished master's thesis,

Department of Chemical Engineering, College of Chemical and Petroleum Engineering, Al-Baath University, Syria.

- 10- VELÁZQUEZ, J, CALEYO, F, VALOR, A, HALLEN, J 2009 **Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines.** NACE International, Corrosion Engineering Section, ISSN 0010-9312.
- 11- 2013 **Pipeline Inspection Techniques.** PetroMin Pipeliner. Vol.09 (1). 48-52.
- 12- WANG, H, YAJIMA, A, LIANG, R, CASTANEDA, H 2014 **Bayesian Modeling of External Corrosion in Underground Pipelines Based on the Integration of Markov Chain Monte Carlo Techniques and . Clustered Inspection Data.** Computer-Aided Civil and Infrastructure Engineering.1-17.
- 13- HAMED, H, RAJJOOB, B 2015 **Study of the cathodic protection of crude oil transport pipelines.** Al-Baath University Journal, Vol. 7(6), 193-217. (In Arabic)
- 14- NOSA, I, BOLAJI, A 2016 **Modelling the Corrosion Rate of Buried Pipes Using Modified Artificial Neural Network (MANN) Coupled with Monte Carlo Simulation.** Archives of Current Research International, Vol.4(2). NO. ACRI.25564.1-8.
- 15- OLADIPO, B, AJIDE, O, MONYEI, C 2017 **Corrosion Assessment of some Buried Metal Pipes using Neural Network Algorithm.** I.J. Engineering and Manufacturing, 27-42.
- 16- YANG, Z, OTERKUS, E 2017 **Corrosion detection in pipelines based on measurement of natural frequencies.** Open Journal of Oceanography.
- 17- XIE, M, TIAN, Z 2018 **A review on pipeline integrity management utilizing in-line inspection data.** Engineering Failure Analysis, Vol. (92), 222-239.

- 18- SUMAYH, A, DORIEH, A, SHATHA, A, FATIMAH, K, ALJAWHARAH, A, FATIMAH, A, RAZAN, A 2022 **An Anomaly Detection Model for Oil and Gas Pipelines Using Machine Learning**. Computation 2022,10, 138.
- 19- HASSAN, U, SHAMSUDDEEN, A, KAULAHA, A 2022 **Effects of Soil Properties on Corrosion Rate of Buried Steel Pipes - A Case Study of Bauchi State Nigeria**. International Journal of Trend in Research and Development, Volume 6(1), ISSN: 2394-9333.

النبؤ بقيمة المقطع العرضي الراداري للغيوم الإلكترونية باستخدام الشبكات العصبونية

محمد أسعد، محمد نجيب صلاحو*، غزوان الخوري**، يحيى فريد***
طالب دكتوراه قسم هندسة الاتصالات، كلية الهندسة الكهربائية والإلكترونية، جامعة حلب
* أستاذ في قسم هندسة الاتصالات، كلية الهندسة الكهربائية والإلكترونية، جامعة حلب
** أستاذ مساعد في فرع هندسة اللاسلكي والرادار، أكاديمية الأسد للهندسة العسكرية
*** أستاذ في قسم هندسة الاتصالات، كلية الهندسة الكهربائية والإلكترونية، جامعة حلب

المخلص

يعتبر التنبؤ بقيمة المقطع الراداري RCS (Radar Cross Section) لغيمة إلكترونية قبل اطلاقها من العوامل الحاسمة في التطبيقات التكتيكية عموماً والكهرطيسية خصوصاً، فمن الصعب حساب هذه القيمة لغيمة تتألف من ملايين العناصر متعددة الأطوال والمبعثرة بشكل عشوائي والخاضعة لظروف البيئة المحيطة من سرعة الرياح والضغط الجوي ودرجة الحرارة وحجم العناصر وعددها الكلي مع المادة المصنوعة منها والمسافة عن الرادار وغيرها الكثير.. ناهيك عن استحالة تقدير الانعكاس الكهرطيسي عن الغيمة نظراً لتعقيد الاقتران الكهرطيسي ما بين عناصرها وبالتالي تم اقتراح استخدام تقنيات الذكاء الصناعي ونخص منها الشبكات العصبونية للتنبؤ بهذه القيمة، وقد تم اختيار بنية الشبكة العصبونية المناسبة.

إن عملية اختيار الشبكة العصبونية المناسبة وتدريبها هي مسألة معقدة ولا توجد ضوابط واضحة لها حيث تختلف الشبكة وبنيتها باختلاف التطبيق ومدخلاته وفي هذا البحث تم اختيار الشبكة المناسبة وتدريبها واختبارها لنصل إلى تنبؤ صحيح بالاعتماد على قيم برنامج newFasant المتخصص بمحاكاة الغيوم الإلكترونية ونتائج الأبحاث العملية في هذا

الكلمات المفتاحية: الغيمة الإلكترونية، الشبكات العصبونية، RCS، التشويش والتمويه الراداري.

RCS Prediction of Chaff Clouds Using Neural Networks

**Mohammad Asaad, M.Najib Salaho*
, Ghazwan Alkhoury**, Yahia Fareed*****

Postgraduate Student (PhD) Telecommunication Engineering Dep., Faculty of Electrical & Electronic Engineering, University of Aleppo

*Professor in Telecommunication Engineering Dep., Faculty of Electrical & Electronic Engineering, University of Aleppo

**Assistant Professor in Wireless and Radar Engineering Branch, Al-Assad Academy of Military Engineering

***Professor in Telecommunication Engineering Dep., Faculty of Electrical & Electronic Engineering, University of Aleppo

Abstract

Prediction the value of chaff cloud's RCS (Radar Cross Section) before launching it is considered one of the decisive factors in tactical applications in general and electromagnetics in particular. It is impossible to calculate RCS for a cloud consisting of millions of elements with different lengths scattered randomly and subjected to surrounding conditions of the environment that affect the chaff cloud shape and behavior; such as wind speed, atmospheric pressure, temperature, size of the elements, total dipoles number, material they are made of, distance from the radar, and many others... Not to mention the impossibility of estimating the electromagnetic reflection from the cloud, due to the complexity of the electromagnetic coupling between its elements and describing its nature. This research proposes use artificial intelligence techniques, especially neural networks, to predict RCS value, by choosing the appropriate neural network structure.

The process of designing and training the appropriate (NN) neural network is a complex issue because there are no clear rules for it. NN structure differs according to the application and its inputs. In this research, the appropriate network was designed, trained and tested to reach a correct RCS prediction based on the values of the newFasant program specialized in simulating chaff clouds beside to practical researches in this area.

Keywords: Chaff cloud, Neural Networks, Radar Cross Section, Passive jamming

1- مقدمة:

بالتزامن مع تطور أنظمة الاتصالات وخاصة في السنوات العشرين الماضية تطورت تقنيات تعلم الآلة (Machine Learning) بشكل كبير لتلبي طيف واسع من التطبيقات في مجالات التحكم ومعالجة الصورة وغيرها .. ومن التطبيقات التي دخلها الذكاء الصناعي بقوة هي التطبيقات العسكرية على اختلاف أنواعها كتحسين دقة الملاحقة وتوجيه الصواريخ والتنبؤ بمسارات الأهداف، ولا تزال الغيوم الإلكترونية موضع الدراسة لأهميتها وفعاليتها في المعارك الجوية والبحرية والاختباء من الرادارات والتمويه على الأهداف ولكن يجد الطيار صعوبة واضحة في تحديد عدد عناصر الغيمة الكافي مع أطوالها المناسبة وفقاً لترددات رادارات العدو وموقعها والطبيعة الجغرافية للمكان، ولهذا تم بناء شبكة عصبونية قادرة على التنبؤ بقيمة المقطع الراداري RCS عند رادار الهدف وفقاً لمتغيرات ثابتة سنتحدث عنها بالتفصيل.

تُعرف الغيمة على أنها مجموعة من الرقائق أو القطع المعدنية التي تنتثر بكميات كبيرة على هيئة حزم، والتي تعكس اشارات الرادار التي تصطدم بها لتحاكي هدف ما، وعادة ما تكون الرقائق من الألمنيوم أو الألياف الزجاجية المطلية بمعدن ما، وتحوي كل حزمة على آلاف العواكس التي يحدد طولها وفقاً لطول موجة الرادار، كما أنّ لها عدد من الأسماء وفقاً للدولة حيث كانت تسمى أساساً بـ "النافذة" (window) في بريطانيا وسُميت بـ Düppel في ألمانيا أثناء الحرب العالمية الثانية، وتستخدم الغيوم الإلكترونية عموماً في ثلاثة تطبيقات أساسية: [1]

1- يمكن أن تحاكي الغيمة الثابتة هدف حقيقي وبالتالي تصبح بمثابة فخ للرادارات التي لا تمتلك القدرة على ترشيح أثر دوبلر.

2- يمكن أن تخفي هدف حقيقي ضمنها أو خلفها كطائرة مثلاً.

3- يمكن نشر عدد كبير من الغيوم الإلكترونية على مساحة جغرافية كبيرة، وبالتالي تؤمن التشويش على الرادارات قبل البدء بغارة أو أن تكون مجرد عملية إرباك للعدو.

تكون العواكس عادة من رتبة طول موجة الرادار المراد التشويش عليه ومن الممكن استخدام عدد من الأطوال المختلفة وبالتالي تغطية مجال ترددي أكبر والحصول على

قيمة كبيرة لسطح المقطع الراداري (RCS (Radar Cross Section). يتم إطلاق هذه الغيوم عبر الألعاب النارية أو الصواريخ أو قذائف الهاون وعادة يتم النشر بطريقة ميكانيكية أو نارية، حيث يستخدم الطرد الميكانيكي رقائق صغيرة من الورق المقوى التي تفتح أثناء الاطلاق. ويبين الشكل (1) أحد أشكال الغيوم الإلكترونية، كما يمكن استخدام القذف الناري وذلك من خلال الغازات الساخنة المتولدة في حجرة الانفجار التي تدفع مكبساً بلاستيكيّاً صغيراً لأسفل أنبوب مملوء بالعواكس الذي يكون بطول 8 بوصات وبمقطع عرضي 1 بوصة مربعة، مما يؤدي إلى إخراج الغطاء البلاستيكي لتتبعه ألياف الغيمة. [2]

تعتبر هذه الطريقة في التمويه من طرق التشويش السلبي المصطنع ويهدف التشويش عموماً إلى تقليل فعالية الوسائط الإلكترونية ويمكن تنفيذه بعدة أشكال منها:

- استخدام مشعات خاصة تؤثر على المستقبلات.
- استخدام أهداف كاذبة.
- استخدام إرساليات خادعة.



الشكل (1) أحد أشكال الغيوم الإلكترونية ومكان تركيبها على جدار الطائرة [3]

نستطيع اعتبار التنبؤ بسلوك الغيمة الإلكترونية باستخدام تقنيات الذكاء الصناعي كطريقة متممة لطرق الحسابات الكهرومغناطيسية (CEM) المعتمدة في مختلف برامج تصميم البنى الكهرومغناطيسية مثل HFSS و Comsol وغيرها، حيث استخدم تعلم الآلة العديد من الخوارزميات في التطبيقات المتنوعة منها الجينية ومنها المنطق الضبابي ونماذج

ماركوف ولكن أهم تقنية هي الشبكات العصبونية بأنواعها والتي أعطت فعالية كبيرة في تطبيقات تنبؤ المتنوعة. [3]

وحتى تكون الدراسة أقرب ما يمكن للواقع يجب اختيار برنامج المحاكاة والتحليل المناسب لهذا النوع من البنى حيث تتعدد في وسط الأبحاث الكهربية البرمجيات المعنية ومنها ما هو الأكاديمي ومنها التجاري ولكن ليس كلها يقدم نتائج موثوقة ما بعد المحاكاة، ناهيك عن مدى سهولة التعامل مع واجهتها ومدى حجم الدعم الذي توفره للأدوات والعناصر المختلفة، وفي هذا الإطار تم اختيار برنامج تصميم المقدم من newFasant كونه يوفر البيئة المثالية لبحثنا هذا، أما في مجال تصميم الشبكات العصبونية واختبارها فتم اختيار برنامج MATLAB لتنوع مكتباته وتخصيص مكتبة كاملة في الذكاء الصناعي.

تم في هذا البحث تقديم شبكة عصبونية تستطيع التنبؤ بقيمة RCS للغمية الإلكترونية لدى رادار الهدف بالاعتماد على البارامترات المؤثرة التالية:

- 1- عدد العواكس
- 2- اطوال العواكس
- 3- تردد الرادارات المعادية
- 4- المسافة عن الرادار
- 5- وزن الغيمة الكلي
- 6- حجم الحيز الذي تشغله الغيمة

تعامل هذه البارامترات كمتغيرات دخل للشبكة العصبونية ليكون خرجها قيمة RCS ويتم تدريب الشبكة العصبونية بناء على معطيات تم استصدارها من بيانات وكالة ناسا وبرنامج المحاكاة المتخصص بالغيوم الإلكترونية newFasant.

2- الدراسة المرجعية:

إن الهدف الأساسي للبحث والتطوير يتمحور حول توفير حلول تقنية ذات أداء عالٍ وسعرٍ منخفض وأكثر تقدماً وأقل تعقيداً عن سابقتها، فكان لزاماً علينا أن نقوم بجمع الدراسات السابقة في المجال ذاته، ولحساسية الموضوع سنصطدم بندرة المراجع

العلمية، حيث يوجد العديد من الدراسات التي حاولت حساب قيمة RCS باستخدام تقنيات الذكاء الصناعي لبنى هندسية نظامية (كطائرة معينة أو سفينة) أو تلك التي قامت بنمذجة الغيمة الإلكترونية وإيجاد المعادلات الرياضية التي تصفها ضمن بيئات عمل مختلفة للوصول بالحاكاة الى أقرب سلوك يحاكي سلوك الغيمة الحقيقي في الجو وتفاعلها مع البيئة من جهة والإشارة الكهروضوئية من جهة أخرى، حيث قدم البحث رقم [4] دراسة تحليلية لانتشار الأمواج الكهروضوئية خلال الغيمة الإلكترونية وتقدير قيمة RCS لها بشكل ديناميكي باستخدام طريقة الناقلية المكافئة العامة (GEC) ومن ثم تقديم موديل رياضي لحساب تلك القيمة أثناء الاستقرار وفي المرحلة السابقة لاستقرار الغيمة، كما قام البحث [5] بالتنبؤ بقيمة RCS للغيمة بشكل ديناميكي بالاعتماد على تأثير الفرز (Screening effect) من خلال تطوير بيئة عمل برمجية باستخدام MATLAB على أشكال مختلفة للغيمة الكروية والاسطوانية والمكعبة، وقام ببناء مجسمات مصغرة تحاكي الموجة وإجراء الاختبارات للمقارنة مع المحاكاة في بيئة عمل مخبرية ضمن غرفة ماعة للصدى ضمن المجال الترددي 8 - 18 GHz ولكن لم يتم اختبار القيم الناتجة على أهداف حقيقية. وقد درس البحث [6] إمكانية تحديد الغيمة الإلكترونية من خلال تسخير بارامترات المسافة وتردد دوبلر والطاقة ضمن صورة مدى الدوبلر والتي تعتمد على ديناميكية الغيمة المدروسة، بينما قدم البحث [7] نموذجاً رياضياً للغيمة الإلكترونية المنتشرة في الجو بالاعتماد على استخدام معادلة الانتشار - رد الفعل العكسي ومن ثم تطبيق نظرية نقل الشعاع (VRT Vector Transport Theory) والحصول على بارامترات الانعكاس للغيمة الإلكترونية، وقد قدم البحث [8] برنامجاً يقدم نموذجاً رياضياً لحساب RCS الغيمة الإلكترونية، أما من وجهة نظر أخرى فقد قام الباحثون [9] بتحديد الهدف الراداري التلقائي باستخدام الشبكات العصبونية التعاونية RNN وذلك بالاعتماد على قاعدة بيانات مؤلفة من قيم RCS بزوايا مختلفة تحاكي حركة الهدف وحصل على دقة تصنيف بقيمة % 98.

مما سبق يُلاحظ بأن الأبحاث السابقة لم تناقش استخدام تقنيات الذكاء الصناعي لتقدير قيمة RCS في الغيوم الإلكترونية، وهو ما تم القيام به في هذا البحث لفتح باب جديد من استخدام الشبكات العصبونية في التطبيقات العسكرية.

3- أهمية البحث وأهدافه:

تتجلى أهمية البحث بتضمين خوارزميات الذكاء الصناعي في تطبيقات الحروب الكهرطيسية وبالتالي بناء الشبكة العصبونية المناسبة للتعويض بقيمة المقطع الراداري وتشكيل الغيمة الإلكترونية المناسبة وذلك بالزمن الحقيقي، لنصل إلى الأهداف التالية:

1- دراسة علاقة قيمة المقطع الراداري مع المتغيرات آنفة الذكر المتنوعة (عدد العواكس - طول العواكس - تردد الرادار - المسافة عن الرادار - وزن الغيمة - حجمها) والتي تم استخلاصها من الدراسات المرجعية السابقة وبرامج المحاكاة وبناء جداول خاصة لتدريب الشبكة العصبونية.

2- دراسة أنواع الشبكات العصبونية المختلفة واختيار النوع المناسب لهذا التطبيق واعتماد البنية المناسبة للشبكة من عدد طبقات مخفية وتوابع التفعيل والقيم الابتدائية وطريقة التدريب.

3- إن التعامل مع هذا القدر الكبير من البيانات واتخاذ القرار المناسب من قبل الطيار أمر يصعب تطبيقه في الزمن الحقيقي وبالتالي القرار الخاطئ من الممكن أن تكون نتائجه كارثية، وهنا أتت فكرة استخدام وتدريب شبكة عصبونية مناسبة التي تتخذ القرار مباشرة وتسهل عملية إطلاق الغيمة الإلكترونية المناسبة.

4- سرعة التنفيذ في الزمن الحقيقي

5- تجنب التعقيد الكهرطيسي والمعادلات الرياضية لنمذجة عمل الغيمة الإلكترونية.

4- بنية الغيمة الإلكترونية:

غالباً ما تستخدم القصاصات المعدنية (الديبولات) العاكسة المولدة للتشويش السلبي المصطنع، ويتم اختيار طول وسماكة الديبولات بشكل يؤمن انعكاس كافة الأمواج تقريباً عن هذه الديبولات مما يؤدي إلى ظهور بقعة مضيئة للتشويش السلبي على شاشات جميع المحطات الرادارية بمختلف أطوال أمواجها، وكقاعدة يجب أن يختار طول الديبول مساوياً لعدد صحيح من أنصاف طول الموجة للمحطة الرادارية المراد التشويش عليها بهدف زيادة استطاعة التشويش السلبي.

يتم تمويه إشارة الأهداف الحقيقية والتأثير على عمل محطة الرادار باستخدام التشويش السلبي بطريقتين:

1- رمي القصاصات المعدنية.

2- رمي العواكس الديبولية.

بالنسبة للقصاصات المعدنية والتي هي شرائح معدنية مصنوعة من معدن خفيف مثلاً الألمنيوم (ممكن أن تكون كذلك أوراق قصدير) تُرمى من الطائرات بأعداد كبيرة على شكل حزم تحوي كل منها آلاف الشرائح بأطوال موجية قريبة، وأحياناً مساوية تماماً طول الموجة العاملة للرادار وتقوم بتمويه الأهداف ضمن منطقة الكشف الراداري ويمكن أن تظهر في بعض الأحيان كأهداف كاذبة على الشاشة، ويسمى التشويش السلبي باستخدام القصاصات ضمن مجال الترددات (VHF، HF) بالتشويش المقيد rope chaff ويرمى هذا النوع من التشويش بثلاثة أشكال هي:

أولاً- مركز بحيث تقلد الهدف الحقيقي.

ثانياً- يُشكل ممر خلف الطائرات المشاركة بالهجوم لإخفائها وتضليل صواريخ الدفاع الجوي عنها.

ثالثاً- يُشكل غيمة منتشرة على مجال واسع تخفي الممر الحقيقي للهجوم والأهداف المهاجمة.

أما الديبولات العاكسة فهي عناصر مصنوعة من معدن خفيف مثل الألمنيوم أو من الزجاج المكسو بالألمنيوم ترمى في الجو لتعكس طاقة الرادار مقلدةً بذلك إشارة منعكسة عن هدف حقيقي ذي سطح عاكس فعال كبير وغالباً ما يتم رميها برزم تحوي ديبولات مختلفة الأطوال لتؤثر على مجموعة من الرادارات بأن واحد. تتمتع العواكس الديبولية بسطح عاكس فعال كبير نسبياً، تفوق شدة إشارات التشويش السلبي أحياناً مستوى الضجيج الداخلي للمستقبل بمقدار $(30 \div 80)dB$ مما يؤدي إلى إشباع مستقبل المحطة الرادارية وضياع الإشارات المفيدة كذلك يمكن أن تصل الإشارات المفيدة مع إشارات التشويش السلبي في آن واحد وهنا يصعب تمييز إشارة الهدف الحقيقي عن التشويش وبالتالي عدم الكشف.

إن إشارات الأهداف الثابتة هي إشارات الأصداء الثابتة المنعكسة عن أجسام واقعة ضمن حجم نبضي لمخطط الإشعاع الهوائي.

إن جميع أنواع الأصداء الثابتة أو المتحركة بشكل بطيء مثل التضاريس الأرضية أو الغيوم أو مجموعات كبيرة من الطيور أو حتى رقائق التشويش السلبي تعطي إشارات انعكاس غير مرغوبة تدعى ضوضاء (*Cluter*) والتي تملأ شاشة الرادار. يُشكل التشويش السلبي خطراً كبيراً على الوسائط اللاسلكية وخاصة الرادارية منها وذلك لأنه موجود في زمن الحرب والسلام، وأهم الاختلافات الأساسية بين الأهداف ومنابع التشويش السلبي هي:

1- إن الطائرات والصواريخ والأهداف الأخرى كقاعدة تعتبر أهداف نقطية، حيث تعتبر أبعاد الهدف الحقيقي أصغر من الحجم النبضي، بينما يكون التشويش السلبي أهدافاً منتشرة.

2- في أغلب الحالات تكون سرعة الأهداف الحقيقية أكبر بكثير من سرعة منابع التشويش. فمثلاً تكون سرعة الغيوم الديبولية أو ظرف الطقس (غيوم، تلوج، أمطار) مساوية لسرعة الريح وتصل إلى عشرات الكيلومترات في الساعة، وتكون سرعة الأجسام الثابتة مساوية للصفر، أما سرعة الطائرات والصواريخ فتصل إلى مئات وآلاف الكيلومترات في الساعة، إن الاختلاف بالسرعة يؤدي إلى اختلاف بتردد الإشارة المنعكسة عن الهدف عن تردد إشارة التشويش السلبي.

3- إن منابع التشويش السلبي الناتجة عن ظروف الطقس تكون ذات شكل أقرب إلى الكروي، بينما تكون الأهداف الحقيقية على الغالب ذات شكل يفتقر إلى التناظر المركزي، وهذا يؤدي إلى وجود اختلافات في استقطاب الإشارات المنعكسة عن الأهداف والإشارات المنعكسة عن ظروف الطقس. [2].

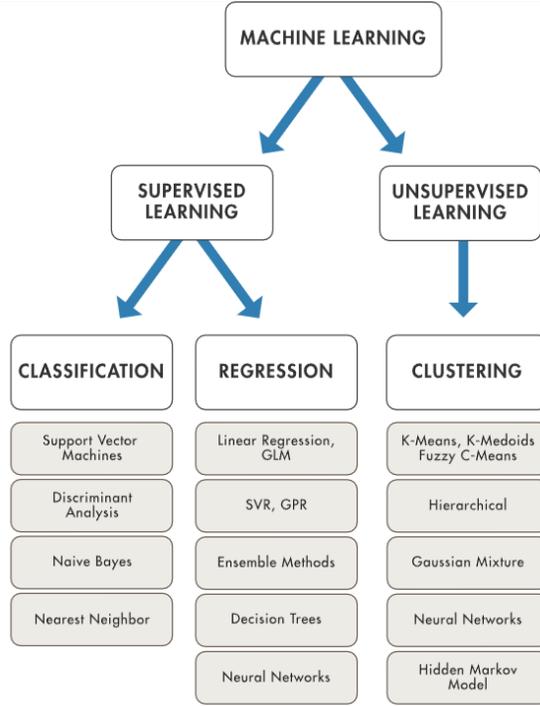
5- الشبكات العصبونية وتعلم الآلة:

علم الذكاء الاصطناعي هو فرع من علم الحاسوب، يختص بتصميم أنظمة حاسوبية ذكية تحاكي السلوك الذكي للإنسان مثل فهم اللغات الطبيعية، والتعلم، والتفكير، والتخطيط، وحل المشكلات والاستدلال المنطقي، وغيرها، ويهدف علم الذكاء

الاصطناعي إلى فهم طبيعة الذكاء الإنساني عن طريق عمل برامج للحاسب الآلي قادرة على محاكاة السلوك الإنساني المتمم بالذكاء. وتعني قدرة برنامج الحاسب على حل مسألة ما، قد تكون فهم نص لغوي منطوق أو مكتوب، أو حل لغز، أو مسألة رياضية. ويبدأ الباحث عمله في علم الذكاء الاصطناعي باختبار أحد الأنشطة المتفق على أنها "ذكية"، ثم يضع بعض الفروض عما يستخدمه الإنسان لدى قيامه بهذا النشاط من معلومات واستدلالات، ثم يدخل هذه المعلومات في برنامج للحاسب الآلي، وأخيراً يراقب سلوك هذا البرنامج. وقد تؤدي ملاحظة البرنامج إلى اكتشاف أوجه القصور فيه مما يفضي إلى إدخال تعديلات وتطوير في أسسه النظرية، ويوضح الشكل (2) تقنيات تعلم الآلة المختلفة.

ومن أهم تقنيات الذكاء الصناعي هي الشبكات العصبونية التي تُعرف على أنها محاولة رياضية برمجية لمحاكاة طريقة عمل المخ البشري. حيث أن العلماء قد اكتشفوا تقريباً طريقة عمل المخ البشري من حيث قابلية التعلم وقابلية التذكر والقدرة على تمييز الأشياء والقدرة على اتخاذ القرارات [12].

إنّ الشبكات العصبونية هي عبارة عن معالج ضخم ذو توزيع تفرعي massively parallel distributed processor مؤلفة من وحدات معالجة قادرة على تصنيف المعطيات التجريبية و جعلها متوفرة لتأدية الغرض المطلوب، تمتلك هذه الشبكات القدرة على نمذجة الأنظمة الخطية و غير الخطية دون الحاجة إلى لوضع افتراضات ضمنية كما في معظم التقريبات الاحصائية الاعتيادية وهذا ما جعلها نقطة استقطاب مستخدمة في العديد من أنواع العلوم الهندسية.



الشكل (2): تقنيات وتطبيقات تعلم الآلة

وتتألف الشبكة العصبونية من:

- مجموعة من وحدات المعالجة (العصبونات).
- مجموعة من توابع التنفيع توضع على خرج كل عصبون (Activation Function) وهو تابع غير خطي، وهناك العديد من توابع التنفيع نذكر منها: [12]

تابع النبضة الواحدية:

$$\text{step}_t = \begin{cases} 1 & \text{if } in_i \geq t \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

تابع الإشارة:

$$\text{sign} = \begin{cases} +1 & \text{if } in_i \geq 0 \\ -1 & \text{if } in_i < 0 \end{cases} \quad (2)$$

تابع السيكمويند:

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (3)$$

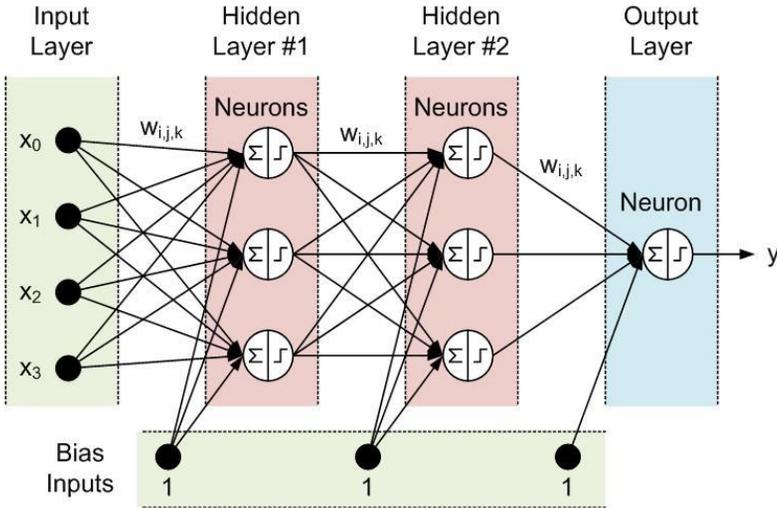
- مجموعة من الوصلات ذات الأوزان لربط العصبونات فيما بينها وربط العصبونات بالوسط الخارجي، ونستخدم القاعدة التالية لحساب وزن العصبون:

$$y_k = f(S_k) \quad (4)$$

حيث أن S_k هو الدخل الفعال للعصبون وهو تابع للأوزان والمداخل السابقة ويسمى الخرج الخطي أو الدخل الكلي:

$$S_k = f(w_{1k}, w_{2k}, \dots, x_1, x_2) \quad (5)$$

- لكل عصبون عامل انحياز نسميه b_k
- وسط توضع فيه الشبكة العصبونية لتتعلم منه وتؤثر فيه.
- خوارزمية أو طريقة لتعليم الشبكة العصبونية بحيث تستطيع جمع المعلومات من الوسط المحيط. والشكل (3) يبين بنية مكونات الشبكة العصبونية.



الشكل (3): بنية الشبكة العصبونية

وللعصبونات المسميات التالية حسب الموقع:

1. عصبون دخل (Input Neuron): العصبونات التي ترتبط مع الوسط الخارجي وليس لها أوزان وليس لها أي وظيفة إلا الربط مع الوسط الخارجي.
2. عصبون مخفي (Hidden Neuron): العصبونات التي تتلقى إشارات الدخل من عصبونات أخرى وتعطي إشارة خرج إلى عصبونات أخرى.
3. عصبون الخرج (Output Neuron): العصبونات التي تعطي الخرج الكلي للشبكة.

وللشبكات العصبونية المزايا التالية:

1. تمتلك الشبكات العصبونية خاصة التعلّم والتدريب من خلال مجموعة من الأمثلة التي قامت بمعالجتها سابقاً، أي يكفي في بعض الأحيان أن تعالج معلومة ما مرة واحدة حتى تستطيع إنتاج ذات الخرج في حالة إدخال دخل مشابه في مرة أخرى.
 2. يمكنها التعبير عن تابع معقد من خلال مجموع حدود جبرية بسيطة.
 3. لا تطلب فهماً عميقاً للمسألة المدروسة، تتم النمذجة بطريقة الصندوق الأسود والتي تتعامل مع النظام من خلال داخله ومخارجه.
 4. ذات بنية تفرعية سهلة التطبيق والتنفيذ.
 5. معدلات حساب تفرعية عالية السرعة.
- ومن مساوئها:

1. تحتاج لزمان تدريب يتناسب طردياً مع تعقيد النظام المدروس.
 2. لا تعطي أية فكرة عن علاقة المتحولات ضمن النظام المدروس.
- ويُعرف الاجراء الذي يتم من خلاله ضبط الأوزان W والانزياح b بتدريب الشبكة العصبونية من أجل قيام الشبكة بإيجاد الخرج المناسب بالنسبة لدخل محدد، ويوجد نوعان من التدريب [12]:

- التدريب بمعلم Supervised.
- التدريب بدون معلم Unsupervised.

5-1- التدريب بمعلم:

في هذه الطريقة يدرّب الحاسوب باستخدام مجموعة تدريب دخل معروفة الخرج مسبقاً، حيث يُعطى النظام الخاضع للتدريب مجموعة بيانات الدخل، بالإضافة للخرج المتوقع، وأثناء التدريب إذا كان الخرج خاطئاً (غير موافق للخرج المرغوب فيه، يقوم النظام بإعادة دورة الحساب لكن بعد إجراء تعديلات مناسبة على الحسابات، تمتدّ هذه العملية على عدّة تكرارات على مجموعة التدريب، حتى يتمكن النظام من إعطاء نتائج متوافقة مع الخرج المرغوب فيه. من الأمثلة على هذه الطريقة في التعلّم:

- مسائل التنبؤ Prediction: مثل التنبؤ بالطقس (Weather - predicting) بالاعتماد على مجموعة تدريب تحتوي على بيانات دخل تخص مثلاً الضغط، الرطوبة، وسرعة الرياح، مع خرج متوقع هو درجة الحرارة.
- مسائل التصنيف Classification: على سبيل المثال تصنيف الكتب (علمية، أو تاريخية، أو سيرة ذاتية، أو روايات...) بالاعتماد على مجموعات تدريب تتضمن أسماء كتب وتصنيفاتها، أو تصنيف رسائل البريد الإلكتروني إلى مهمة أو غير مهمة.

5-2- التدريب بدون معلم:

في هذه الطريقة يدرّب الحاسوب باستخدام مجموعة تدريب، لكن على عكس الطريقة السابقة، حيث الدخل هنا بدون معرفة الخرج مسبقاً. يتمثل التدريب بإجراء تصنيفات منطقية على البيانات، واستنتاج علاقات ربط فيما بينها، أي عنقدة (تجميع) بيانات الدخل ضمن مجموعات غير معروفة مسبقاً، ومن الأمثلة على هذه الطريقة في التعلّم هي التنبؤ كالتنبؤ بالسلوك على مواقع التجارة الإلكترونية، لكنه ليس كالتنبؤ بإشراف معلم، حيث يتم هنا تجميع المستخدمين في مجموعات تبعاً لسلوكيات التصفح، والتقييمات، وعمليات الشراء، ومن ثمّ التنبؤ بالأخبار والإعلانات التي تهمهم، واقتراحها عليهم.

هنالك العديد من تقنيات تدريب الشبكات العصبونية لا نريد أن نسهب في شرحها في هذا البحث فهي متوفرة في العديد من المراجع المتخصصة، وقد استخدمنا في بحثنا هذا خوارزمية الانتشار العكسي Back Propagation.

6- تصميم الشبكة العصبونية

تتراءى أمام المصمم العديد من الطرق والأساليب في بناء وتصميم الشبكة العصبونية، ولكن اتفق العديد من المصممين على مجموعة من الأساسيات أو المسلمات في التصميم جاءت بعد التجربة والخبرة ومن أبرزها: [13]

- يعتمد المبدأ الأول: التطوير التدريجي لنموذج الشبكة العصبونية حيث يتم العمل بداية على بنية بسيطة والتي من الممكن أن تكون كافية ومناسبة ومن ثم يتم تطوير البنية تدريجياً، فعلى سبيل المثال، إذا كنا لا نعرف شيئاً عن شكل الدالة، فيجب أن نفترض مبدئياً أن المشكلة خطية وأن نتعامل معها وفقاً لذلك، وعند فشل هذا النهج، نتحرك نحو الدوال الأكثر تعقيداً، وذلك لأن الجزء الأكثر تكلفة من الناحية الحسابية لتطوير الشبكة العصبية يتكون من تدريب معلّماتها. إذا تمكنا من إيجاد نموذج خطي لحل مشكلة معينة، فسيوفر لنا هذا وقتاً حسابياً كبيراً وموارد مالية، أما إذا لم يكن هناك استطاعة، يجب المحاولة مع طبقة أو طبقتين مخفيتين.
- ينص المبدأ الثاني: على زيادة عدد الطبقات المخفية تدريجياً والبدء بعدد مناسب وقليل من الطبقات التي من الممكن أن تؤدي الغرض والتركيز في حال كانت الشبكة العصبية مع عدد معين من الطبقات المخفية غير قادرة على اتخاذ القرار الصحيح فإنه يفضل تجنب زيادة عدد الطبقات حتى إذا فشل التدريب، وبدلاً من ذلك، يجب توسيع الشبكة بإضافة المزيد من الخلايا العصبية المخفية، حيث تعد مضاعفة حجم الطبقة المخفية أقل تكلفة من الناحية الحسابية، من مضاعفة عدد الطبقات المخفية.
- المبدأ الثالث: ويُطبق دائماً عندما يتم العمل مع بيانات جديدة ويفشل التدريب الأمر الذي يشير إلى أنه ربما تتطلب البيانات التي يتم استخدامها خطوات معالجة إضافية، وهذا بدوره يعني أن المشكلة في التدريب لا تتعلق بعدد الطبقات المخفية في حد ذاتها، بل تتعلق بتحسين معايير الطبقات الموجودة، قد تعني معالجة البيانات بشكل أفضل أشياء مختلفة وفقاً للطبيعة المحددة للمشكلة، فعلى سبيل المثال، ربما هناك حاجة إلى إجراء تقليل في الأبعاد لاستخراج ميزات مستقلة صحيحة أو ربما يجب القيام بتوحيد أو تطبيع المدخلات، لتخفيف صعوبة التدريب، أو ربما من الممكن إضافة طبقة تسرب، خاصة إذا كان النموذج يتناسب مع الدُفعات الأولى من البيانات.

وجدنا سابقاً بأن دخل الخلايا العصبية يعبر عن عدد الميزات التي يتم استخدامها في الشبكة العصبية لعمل التنبؤات واتخاذ القرار، ويحتاج متجه الدخل إلى خلية دخل واحدة لكل ميزة، وبالنسبة إلى البيانات المجدولة هو عدد الميزات ذات الصلة في مجموعة البيانات الخاصة ويجب تحديد هذه الميزات بعناية وإزالة أي منها ممن يحتوي على أنماط لا يمكن تعميمها خارج مجموعة التدريب (وتتسبب في زيادة التخصيص)، فبالنسبة للصور، مثلاً أبعاد الصورة $(28 * 28 = 784)$. وبالتالي يوجد 784 مدخل.

بينما نعبر عن خلايا الخرج بأنها عدد التوقعات المراد القيام بها، ومن جهة أخرى يعتمد عدد الطبقات المخفية بشكل كبير على المشكلة وبنية الشبكة العصبية، وبشكل عام يتم البدء بعدد من 1 إلى 5 طبقات مخفية في عدد كبير من التطبيقات البسيطة والتي تخدم جيداً لحل معظم المشكلات. عند العمل باستخدام بيانات الصورة أو الكلام أو دراسة سلوك قناة نقل لاسلكية، قد تحتوي الشبكة على عشرات أو المئات من الطبقات، والتي قد لا تكون جميعها متصلة بالكامل، وإنما يتم تجزئة الشبكة إلى عدد من الشبكات الفرعية والتي من الممكن استخدام أجزاء أو نماذج مدربة مسبقاً لعملية معينة مثل (VGG, ResNet, YOLO)، وبشكل عام يكفي استخدام نفس العدد من الخلايا العصبية لجميع الطبقات المخفية. بالنسبة لبعض مجموعات البيانات سيؤدي وجود طبقة أولى كبيرة ومتابعتها بطبقات أصغر إلى أداء أفضل حيث يمكن للطبقة الأولى أن تتعلم الكثير من ميزات المستوى الأدنى التي يمكن أن تغذي بعض الميزات ذات الترتيب الأعلى في الطبقات اللاحقة.

6-1- بناء الشبكة العصبونية:

تم بناء الشبكة العصبونية في هذا البحث باستخدام برنامج MATLAB كبيئة برمجية (تم اعتماد المرجع [10, 11] كدليل برمجي)، حيث تم برمجة النماذج الرياضية وإدخال المعطيات وتدريب الشبكة واستخلاص الشبكة النهائية مع أوزانها، وذلك لسهولة استخدام واجهة الـ MATLAB وتوفر مكتبات خاصة بالذكاء الصناعي، ويمكن تقسيم العمل إلى الخطوات التالية:

1. اختيار نوع (معمارية) الشبكة العصبية.
2. اختيار الأوزان (التدريب) للشبكة العصبية.

الخطوة الأولى هي اختيار عدد الخلايا العصبية المراد استخدامها (عدد المدخلات، وظائف النقل)، وكيفية الربط ببعضهم البعض، وتبدو هذه المهمة هائلة، لكن لا يتعين ابتكار شبكة عصبية من البداية، فهناك العشرات من هياكل الشبكات العصبية المختلفة، وقد تم إثبات فعالية العديد منها رياضياً، وقد قمنا باختيار عدد المداخل بناء على البارامترات المؤثرة على تشكيل الغيمة الإلكترونية واستجابتها للرادار، وبالتالي يتوفر لدينا ستة مداخل (عدد العواكس - طول العواكس - وزنها - حجم الغيمة - تردد الرادار - المسافة عن الرادار) بخرج واحد هو قيمة RCS المتنبأ بها.

والخطوة الثانية هي تدريب الشبكة العصبية المختارة، أي تحديد قيم الأوزان بحيث تعمل بالطريقة المرغوبة. يمكن أن يصل عدد الأوزان إلى عدة عشرات الآلاف، لذا فإن التدريب عملية صعبة. تم تطوير خوارزميات تعليمية خاصة تسمح لك بضبط أوزان الشبكة العصبية بطريقة معينة، وأكثر هذه الخوارزميات شيوعاً هي طريقة (الانتشار العكسي).

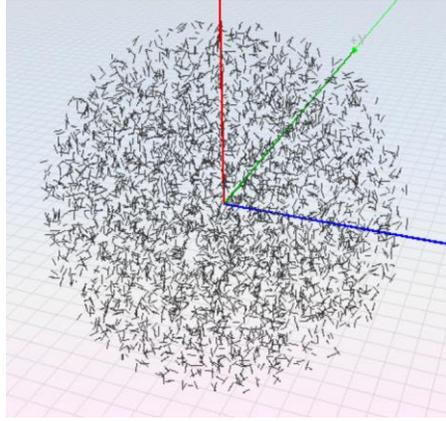
ويمكن تلخيص مراحل بناء الشبكة العصبية بالنقاط التالية: [13]

1. جمع البيانات للتدريب.
2. إعداد البيانات وتطبيعها.
3. اختيار طوبولوجيا الشبكة.
4. الاختيار التجريبي لخصائص الشبكة.
5. الاختيار التجريبي لمعايير التدريب.
6. التدريب الفعلي والتحقق من كفاية التدريب.

6-1-1 جمع البيانات للتدريب:

يعد اختيار البيانات لتدريب الشبكة ومعالجتها أصعب مرحلة في حل المشكلة، حيث يجب أن تستوفي مجموعة بيانات التدريب عدة معايير منها، التمثيل الصحيح (توضيح الحالة الحقيقية للمشكلة المعالجة من خلال البيانات)، والانتظام حيث ستؤدي البيانات غير المنتظمة في مجموعة التدريب إلى رداءة جودة تدريب الشبكة، ويتم تحويل البيانات الأصلية إلى نموذج يمكن تغذيته بمدخلات الشبكة، حيث يتم استدعاء زوج

التعليم أو ناقلات التدريب لكل إدخال في ملف البيانات، وقد تم جمع البيانات الخاصة بتدريب الشبكة العصبونية من خلال بيانات تقنية تم الحصول عليها من المركز الوطني للملاحة الجوية وإدارة الفضاء NASA [14, 15]، وبرنامج newFasant المتخصص بمحاكاة الحقول الكهرومغناطيسية على البنى المختلفة وخاصة الغيوم الإلكترونية حيث قمنا ببناء عدد من الغيوم الإلكترونية باستخدام البرنامج وإجراء المحاكاة عليها، بالإضافة إلى العودة لمجموعة من المعادلات الرياضية التي تقوم بنمذجة عمل الغيمة الإلكترونية. تم إجراء المحاكاة ضمن بيئة newFasant من أجل غيمة الكترونية بنوع واحد من الديبولات العاكسة بطول نصف طول موجة $15mm$ وبنصف قطر $0.3mm$ من أجل رادار يعمل على التردد $8GHz$ من النوع *mono*، وللغيمة قطر $10m$ ، حيث تم تغيير عدد العاكسات في الغيمة من $(5000 \div 11000)$ قطعة ليتم الحصول على قيم RCS في الحقل البعيد، كما هو موضح في الشكل (4).



الشكل (4): محاكاة بنية غيمة كروية

وتم تم مقارنة نتائج المحاكاة مع العلاقات الرياضية المنمذجة لعمل الغيمة الإلكترونية فعادة ما يتم تحديد السطح العاكس الفعال لديبول بتوضع واحد عشوائي وفق العلاقة (6): [1]

$$\sigma_{\pi} = 0.85 \cdot \lambda^2 \cdot \cos^4 \theta \quad (6)$$

θ : الزاوية بين محور العاكس وشعاع الحقل الكهربائي \vec{E} .

بفرض أن التوضع يخضع لقانون التوزيع الطبيعي فتحدد القيمة المتوسطة للسطح العاكس الفعال لديبول واحد وفق العلاقة (7): [1]

$$\bar{\sigma} = 0.17 \cdot \lambda^2 \quad (7)$$

تُعطى قيمة السطح العاكس الفعال لحزمة ديبولية بشكل عام بالعلاقة (8) [7]:

$$\sigma_{chaff} \approx 0.18 \cdot \lambda^2 \cdot N \quad (8)$$

حيث: λ طول الموجة بالمتر، N العدد الكلي للديبولات ضمن الحجم.

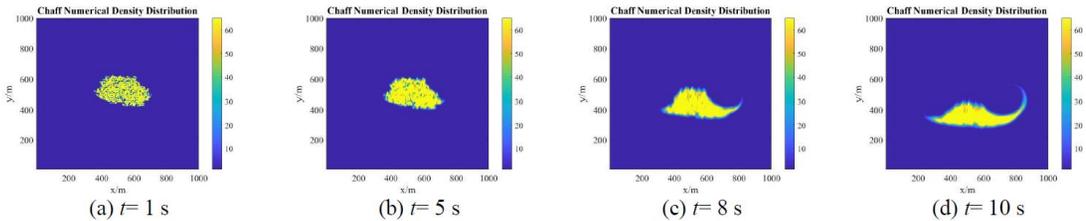
كما يتعلق السطح العاكس الفعال الكلي للغيمة بكتلة حزمة الديبولات العاكسة ويعطى بالنسبة للنماذج الأكثر استخداماً بالعلاقة (9) [7]:

$$\sigma_T \approx 22000 \lambda \cdot W_C \quad (9)$$

حيث: W_C كتلة حزمة الديبولات مقدرةً بـ kg .

حيث يُلاحظ من العلاقة بأنه مع زيادة طول الموجة تتناقص كتلة الغيمة الكلية من أجل مقطع راداري معين، وذلك بسبب الانخفاض في عدد العواكس

عند دراسة تأثير سرعة الغيمة في الجو والتي تتشكل خلال أقل من 0.5 s من الاطلاق وتحت تأثير السرعة الكبيرة للمركبة الحاملة لها يُلاحظ أن رقائق الديبولات تنتشر في طبقات الجو أفقياً وتتساقط ببطء متذبذبةً تبعاً لاضطرابات حركة الريح مما يؤدي لتضخم معامل انعكاسها (η) وازدياد نسبة الاستقطاب اللاخطي لها وفقاً لكثافتها وبالتالي ينخفض مدى الترابط (ρ) فيما بينها، وتستمر دورة حياة الغيمة من بضع ثوان إلى عدة ساعات ويبين الشكل (5) مراحل تشكل الغيمة.



الشكل (5): مراحل تشكل الغيمة الإلكترونية مع الزمن

يوضح الجدول (1) جزء من البيانات التي تم بناؤها (7 عينات فقط)، والمؤلفة من زوج التعليم (المداخل الست - خرج واحد).

الجدول (1): جزء من بيانات تدريب الشبكة العصبونية

العينة	1	2	3	4	5	6	7
عدد العناصر*1000	5	6	7	8	9	10	11
التردد [GHz]	6	6	6	6	6	6	6
طول العاكس [cm]	2.5	2.5	2.5	2.5	2.5	2.5	2.5
نصف قطر [m]	5	5	5	5	5	5	5
المسافة [km]	22	22	22	22	22	22	22
الوزن [kg]	0.278	0.334	0.39	0.445	0.501	0.557	0.612
rsc [dBsm]	0.48960283	1.281415	1.950883	2.530803	3.042328	3.499903	3.91383

نلاحظ من الجدول (1) بأنه قد تم اختيار الوحدات القياسية بما يتلاءم مع إمكانية تدريب الشبكة والحصول على نتائج أفضل بزمان تدريب أقل، فعلى سبيل المثال لا يجوز أن يشمل أحد مداخل الشبكة وليكن التردد على رقم يتسع خانات وطول العاكس عبارة عن رقم بخانتين، ولذلك تم اختيار الوحدات بحيث تكون عدد خانات الأرقام متقاربة لسلامة عملية التدريب.

6-1-2 إعداد البيانات وتطبيعتها:

يتم اعداد البيانات وتطبيعتها لكافة المداخل والمخارج بأرقام ما بين [0-1] حصراً حتى لا يحدث اشباع سريع أثناء عملية التدريب وبالتالي عدم الحصول على أوزان صحيحة لعناصر الشبكة العصبونية.

6-1-3 اختبار طوبولوجيا الشبكة:

يجب أن يعتمد اختيار نوع الشبكة على الفهم الدقيق للمشكلة وكذلك معرفة البيانات المتاحة للتدريب، بالإضافة إلى وجود خرج واضح لكل عنصر من العينات عند إجراء عملية التدريب بمعلم، كون المشكلة الحالية هي التنبؤ، وتجدر الإشارة إلى عدم توفر صيغة واضحة لبناء الشبكة العصبونية وإنما تعتمد على خبرة الدارس ونوعية المداخل وعددها وطبيعتها المشكلة، فهناك الملايين من الاحتمالات المتاحة، ولكن عموماً يمكننا أن نقول بأن عدد العصبونات في طبقة المخفية ما قبل طبقة الخرج يجب أن يكون 3 - 5 أضعاف عدد عصبونات الخرج وبالتالي يجب أن يكون لدينا 3 عصبونات خرج، وبالنسبة للمداخل يمكن أن تزداد عدد العصبونات في الطبقات المخفية من الدخل إلى الخرج بشكل تدريجي ثم تنخفض من المنتصف بشكل تدريجي بحيث تبدأ ب 3

أضعاف عدد طبقات الدخل وبالتالي 18 عصبون مخفي، ناهيك عن اختيار توابع التفعيل المناسبة للمشكلة كما يمكن اختيار عدة أنواع توابع تفعيل وفق موقعها في الشبكة، وهناك عدة نماذج سنقدمها لاحقاً،

6-1-4 الاختيار التجريبي لمعايير التدريب

يجب تحديد بارامترات تدريب الشبكة العصبونية وذلك بعد اختيار طوبولوجيا معينة، هذه الخطوة مهمة بشكل خاص للشبكات الخاضعة للإشراف، ولا يعتمد الاختيار الصحيح لقيم بارامترات التدريب فقط على مدى سرعة تقارب الخطأ فمثلاً سيؤدي اختيار معدل تعلم منخفض إلى زيادة وقت التقارب، ولكنه في بعض الأحيان يتجنب شلل الشبكة وفي أوقات أخرى يوقع الشبكة في منطقة انخفاض محلي وبالتالي عدم الوصول إلى الحل الصحيح.

6-1-5 التدريب الفعلي للشبكة

تقوم الشبكة أثناء التدريب بمسح عينة التدريب بترتيب معين، يمكن أن يكون ترتيب المسح متسلسلاً أو عشوائياً. مثلاً بعض الشبكات غير الخاضعة للإشراف (شبكات Hopfield) تحصح العينة مرة واحدة فقط، بينما شبكات Kohonen والشبكات الخاضعة للإشراف تقوم بمسح العينة عدة مرات بتمريرة كاملة واحدة عبر العينة تسمى عصر التعلّم (Epoch).

عند التعليم مع معلّم يتم تقسيم مجموعة البيانات الأولية إلى ثلاثة أجزاء (عينة التدريب نفسها وبيانات الاختبار وبيانات التقييم)، حيث يتم تغذية بيانات التدريب إلى الشبكة من أجل التدريب، ويتم استخدام بيانات التحقق من الصحة لحساب خطأ الشبكة (لا يتم استخدام بيانات التحقق مطلقاً لتدريب الشبكة)، وبالتالي إذا تم انقاص الخطأ في بيانات الاختبار فإن الشبكة تتعلم، أما إذا استمر الخطأ في بيانات التدريب في الانخفاض، وزاد الخطأ في بيانات الاختبار فقد توقفت الشبكة عن التعلّم وتتذكر بيانات التدريب، تسمى هذه الظاهرة باحتواء الشبكة أو فرط التخصيص، في مثل هذه الحالات عادة ما يتم إيقاف التدريب.

6-2- تهيئة بارامترات تدريب الشبكات العصبونية

هنالك العشرات من البارامترات الهامة والتي تؤثر مباشرة على عملية تدريب الشبكة العصبونية نذكر منها:

1- **تابع الخسارة:** متوسط الخطأ التربيعي هو دالة الخسارة الأكثر شيوعاً التي يجب تحسينها ما لم يكن هناك عدد كبير من القيم المتطرفة.

2- **عدد عصور التدريب (دورات التدريب):** يوصى بالبدء بعدد كبير من مرات التدريب واستخدام التوقف المبكر لإيقاف التدريب عندما يتوقف الأداء عن التحسن.

3- **معدل التعلّم:** يعد اختيار معدل التعلّم أمراً مهماً للغاية، للعثور على أفضل معدل تعلم، نبدأ بقيمة منخفضة جداً (10^{-6}) ونضربها ببطء في ثابت حتى تصل إلى قيمة أعلى. عادةً ما يكون أفضل معدل تعلم هو نصف معدل التعلّم الذي يتسبب في تشعب النموذج. كما يوصى باستخدام طريقة مكتشف معدل التعلّم التي اقترحتها ليزلي سميث، إنها طريقة ممتازة للعثور على معدل تعليمي جيد لمعظم مُحسّن التدرج (SGD) وتعمل مع معظم بنى الشبكات.

4- **الرّخم:** يأخذ التدرج خطوات صغيرة ومتسقة نحو الحدود الدنيا المحلية وعندما تكون التدرجات صغيرة، يمكن أن يستغرق الكثير من الوقت من أجل التقارب. ومن ناحية أخرى، يأخذ الرّخم في الاعتبار التدرجات السابقة ويسرع التقارب عن طريق الارتفاع نحو القاع أو الوديان بشكل أسرع وتجنب الحدود الدنيا المحلية. تعد القيمة 0.9 مكاناً جيداً للبدء بمجموعات البيانات الصغيرة.

5- **التلاشي - التدرجات المتفجرة:** جميع طبقات الشبكة العصبونية لا تتعلم بنفس السرعة، لذلك عندما تنتشر خوارزمية back-propagation من طبقة المخرجات إلى الطبقات الأولى، تصبح التدرجات أصغر وأصغر حتى تصبح ضئيلة تقريباً عندما تصل إلى الطبقات الأولى. هذا يعني أن أوزان الطبقات الأولى لم يتم تحديثها بشكل ملحوظ في كل خطوة، هذه هي مشكلة تلاشي التدرجات (تحدث مشكلة مماثلة لانفجار التدرجات عندما تصبح التدرجات لطبقات معينة أكبر بشكل تدريجي، مما يؤدي إلى تحديثات ضخمة للوزن لبعض الطبقات على عكس الطبقات الأخرى)،

وهنا يتم تعديل أنواع توابع التفعيل ما بين الطبقات القريبة من الخرج والطبقات القريبة من الدخل.

6- توابع التفعيل: بشكل عام يتحسن الأداء الناتج عن استخدام وظائف التنشيط المختلفة بهذا الترتيب من الأدنى ← الأعلى أداءً:

logistic → tanh → ReLU → Leaky ReLU → ELU → SELU

وتعد ReLU هي وظيفة التنشيط الأكثر شيوعاً. وبالنسبة لطبقة الخرج فتكون توابع التفعيل مثلاً وفق نطاق المخرجات فيمكننا استخدام القيم $[-1, 1]$ ، أما من أجل عملية التصنيف فيمكن استخدام تابع التفعيل sigmoid للتصنيف الثنائي للتأكد من أن الناتج بين 0 و 1 أو استخدم softmax للتصنيف متعدد الفئات لضمان أن احتمالات الإخراج تضيف ما يصل إلى 1.

7- طريقة تهيئة الأوزان: يمكن أن تؤدي طريقة تهيئة الوزن الصحيحة إلى تسريع وقت التقارب إلى حد كبير. يعتمد اختيار طريقة التهيئة على تابع التفعيل وطريقة التدريب، من الممكن الاعتماد على الطرق التالية في التهيئة:

1. عند استخدام ReLU أو Leaky RELU، يتم استخدام تهيئة He

2. عند استخدام SELU أو ELU، يستخدم تهيئة LeCun

3. عند استخدام softmax أو logistic أو tanh، يستخدم تهيئة Glorot

تأتي معظم طرق التهيئة بطبيعة توزيع منتظمة وطبيعية.

8- التوقف المبكر للتدريب: يتيح الإيقاف المبكر من خلال تدريب نموذج يحتوي على طبقات مخفية أكثر، وخلايا عصبية مخفية واعدد مرات تدريب أكثر مما تحتاج إليه، وإيقاف التدريب فقط عندما يتوقف الأداء عن التحسن على التوالي لعدد من الفترات.

9- التخطي: التخطي هو أسلوب تنظيم رائع يمنح أداءً هائلاً (حوالي 2٪ للأساليب الحديثة) بالرغم من بساطة هذه التقنية، والتي تتلخص بإيقاف تشغيل نسبة مئوية من الخلايا العصبية في كل طبقة بشكل عشوائي في كل خطوة تدريب وهذا يجعل الشبكة أكثر قوة لأنها لا تعتمد على أي مجموعة معينة من الخلايا العصبية المدخلة لعمل التنبؤات وبالتالي توزيع المعرفة بين الشبكة بأكملها. يتراوح معدل التخطي

الجيد بين 0.1 إلى 0.5 فعادة يكون $RNN \downarrow 0.3$ و $CNN \downarrow 0.5$. ويتم استخدام معدلات أكبر للطبقات الأكبر.

7- النتائج والمناقشة:

تدرج المسألة التي نعالجها في هذا البحث ضمن إطار مسائل التنبؤ والتي تُصنف بدورها في التدريب بمعلم، وبالتالي تتطلب عملية اعداد الشبكة العصبونية المناسبة الكثير من الوقت والجهد والخبرة، وقد تم تصميم عدة شبكات عصبونية ومقارنة أدائها مع بعضها البعض للحصول على الشبكة ذات الطوبولوجية الأفضل بمصفوفة الأوزان المناسبة، والتي تقوم بالزمن الحقيقي بعملية التنبؤ بقيمة RCS للغيمة المراد اطلاقها بعد عدة خطوات أساسية:

- اجراء عمليات معالجة الإشارة الخاصة بالتجهيزات
- استخلاص المعلومات عن الرادار (تردد - مسافة)
- تحديد عدد عناصر الغيمة
- التنبؤ بقيمة RCS

وتتشابه بنى كافة الشبكات العصبونية المقترحة بالنقاط الأساسية التالية:

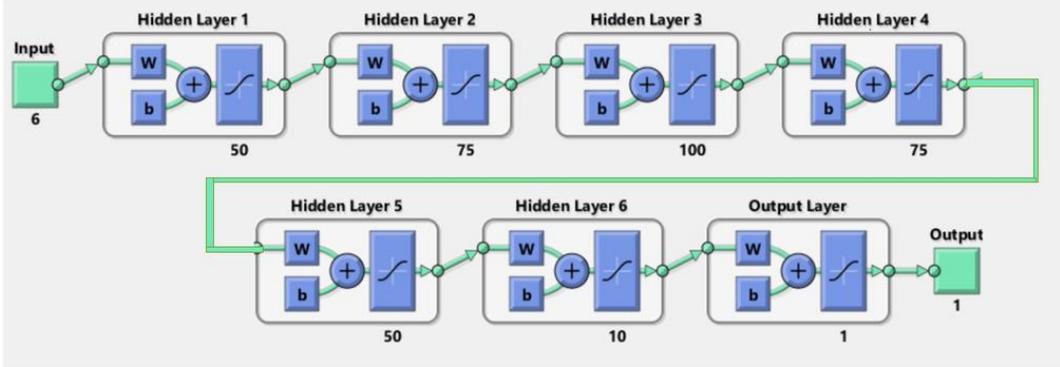
- 1- 6 ستة عناصر دخل (المطبعة) هي بيانات خصائص الرادار وبنية الغيمة الإلكترونية المقترحة.
- 2- عنصر خرج وحيد قيمة RCS الممتبأ بها وذلك ضمن إطار التدريب بمعلم
- 3- طريقة التدريب هي الانتشار العكسي.
- 4- التهيئة الأولية للأوزان باستخدام HE.
- 5- تقسيم بيانات التدريب إلى 85% تدريب و 10 % تقييم و 5 % اختبار، وتجدر الإشارة إلى أن بيانات الاختبار لا تدخل في عملية التدريب.

وتختلف فيما بينها:

- 1- عدد الطبقات المخفية
- 2- عدد عصبونات كل طبقة
- 3- توابع التفعيل وتوزيعها بين طبقات الدخل والخرج
- 4- خوارزمية التحسين

5- عدد عصور التدريب

وكانت بنية الشبكة العصبونية ذات دقة النتائج الأفضل تلك الموضحة في الشكل (6) والتي لها البارامترات الموضحة في الجدول (2).

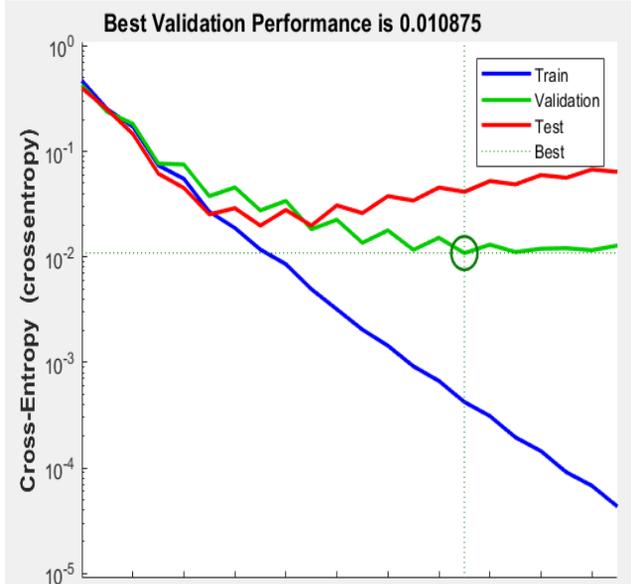


الشكل (6): بنية الشبكة العصبونية المقترحة
الجدول (2): بارامترات الشبكة العصبونية المقترحة

المعاملات	الشبكة العصبونية الخاصة بالنموذج
عدد سمات الدخل	6
عناصر الخرج	1
الطبقات المخفية	7
عدد العصبونات	بنية هرمية بخطوة 25 عصبون 360 عصبون
تابع التفعيل	سيجمويد
خوارزمية التدريب	الانتشار العكسي
تقسيم البيانات	5-10-85
معدل التعلم	0.01
MSE goal	e-2
عدد عصور التدريب	65

والتي لها بنية شجرية مؤلفة من سبع طبقات تتزايد تدريجياً من الدخل بحيث تحقق عصبونات الطبقة الأولى عدد أكبر من خمسة أضعاف عدد سمات الدخل وبالتالي تم اختيار 50 عصبون دخل ليتزايد تدريجياً حتى 100 عصبون في الطبقة المخفية الثالثة ليعود ويتناقص ليحقق 10 عصبونات في طبقة الخرج النهائية، ولها توابع تفعيل من نوع سيغمويد.

حيث يتم تقسيم البيانات ما بين عملية التدريب والتقييم والاختبار إلى نسبة مئوية (85 - 10 - 5) بالمئة على الترتيب، ويبيّن الشكل (7) أداء الشبكة العصبونية التي تم اختيارها حيث أعطت نتائج مناسبة ودقيقة بنسبة خطأ لا تتعدى 0.01.



الشكل (7): أداء الشبكة العصبونية المقترحة

وقد تم اختبار عدد من الشبكات العصبونية للوصول إلى نتائج صحيحة والتي كما تم ذكره سابقاً تعتمد على الخبرة والتجريب وعلى نوع سمات الدخل وطبيعة الشبكة والموصوفة وفق الجدول (3).

الجدول (3): مقارنة ما بين عدد من الشبكات العصبونية المقترحة

رقم الشبكة	بنية الشبكة	طريقة التدريب	عدد الطبقات المخفية	عدد العصبونات	تابع التفعيل	الدقة
1	هرمية	الانتشار الخلفي	7	360	سيجمويد	98
2	هرمية	الانتشار الخلفي	15	900	مختلط	65
3	متزايدة	الانتشار الخلفي	5	200	سيجمويد	80
4	متزايدة	الانتشار الخلفي	9	630	مختلط	73
5	متناقصة	الانتشار الخلفي	5	200	سيجمويد	83

حيث تم اعتماد نوعين من طريقة توابع التفعيل، فإما تم استخدام توابع تفعيل (سيجمويد) خالصة أو استخدام توابع تفعيل مختلطة (سيجمويد - RELU طبقة خرج) حيث تعمل طبقة RELU على زيادة كفاءة التدريب بالنسبة للشبكات ذات عدد الطبقات المخفية الكبير، ومنه تم اعتماد الشبكة العصبونية في الجدول (2).

8- الخلاصة

مما تقدم نخلص إلى أن الشبكة العصبونية المقترحة هي طريقة فعالة ومناسبة لنمذجة الغيمة الإلكترونية وحساب قيمة المقطع الراداري في الزمن الحقيقي بناء على المعطيات آنفة الذكر وبالتالي مساعدة الطيار في الظروف الخطرة لاتخاذ القرار المناسب، وتمتاز بأنها سهلة التطبيق مع انخفاض في كلفة التصنيع والبرمجة ولكن تكمن المشكلة في عملية انتقاء البيانات المناسبة لتدريب الشبكة.

المراجع

- [1] ZAK J.; VACH M.; DVORACEK F., 2016 – **Advanced Chaff Usage in Modern EW.** *IEEE*, 4.
- [2] KNOTTE E.; SHAEFFER J.; TUELY M., 2004 - **Radar Cross Section.** SciTech Publishing, second edition, USA, 634.
- [3] HARRISON A., 2020 - **Introduction to Radar Using Python and MATLAB.** ARTECH HOUSE, London, 491.
- [4] D. SEO; H. NAM; O. KWON; N. MYUNG, 2011 - **Dynamic RCS Estimation of Chaff Clouds.** *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS*, 48(3).
- [5] V. Kumar; A. Singh; P. Vasistha; R. Kumar, 2018 - **Dynamic RCS Prediction of Rapidly Blooming Chaff Cloud and its Validation using Measurement on Scaled-down.** *International Journal of Advances in Microwave Technology (IJAMT)*, 3(4).
- [6] H. Wang; B. Chen; D. Zhu; F. Huang; X. Yu2; Q. Ye; X. Cheng; S. Peng; J. Jing, 2022 - **Chaff identification method based on Range Doppler imaging feature.** *IET Radar Sonar Navig*, 1861 -1871.
- [7] LIXINEN G.; YANCHUAN Z., 2020 – **Investigation on the Electromagnetic Scattering from the Chaff Cloud in Airflow with VRT.** *IEEE*, 4.
- [8] PINCHOT L.; BECHU O.; POULIGUEN P., 2005– **A chaff cloud modelisation.** *ANTEM*, 4.
- [9] B. Sehgal; H. Shekhawat; S. Jana, 2019 - **Automatic Radar Target Identification Using Radar Cross Section Fluctuations and Recurrent Neural Networks.** *IEEE Region 10 Conference*, 6.
- [10] MATLAB, 2015 - **Creating Graphical User Interfaces.** The Mathworks Inc., USA, 502.
- [11] KIM P., 2017 – **MATLAB Deep Learning.** APRESS, USA, 162.
- [12] A. Malla; M. S. Modabbes, 2022 – **Design and Implementation of Self-Tunable Smart Array Antenna.** *Aleppo University*, 250.
- [13] A. ANAND ; AMURU S., 2020- **How to choose a neural network architecture ? - A modulation classification example,** *IEEE*, 5.
- [14] W. Joss; H. Prinsen; C. Easterbrook, 1971 – **An Experimental Investigation and Evaluation of a Doppler RADAR Technique for the Measurement of the flow Generated by Wing Tip Vortices.** *NASA*, 56.
- [15] H. U. Widdel; U. Zahn, 1990 – **A Comparison of Results Obtained from Foil Chaff Cloud at 69 ° Northern Latitude During Winter, Summer and Autumn.** *NASA*, 56.

أنظمة التوصية المعتمدة على الفلتر التعاونية المعززة أداؤها بنماذج تعلم الآلة

م. محمد ديب¹، أ.د. ناصر أبو صالح²

كلية الهندسة المعلوماتية - جامعة البعث

المخلص

تتزايد الأبحاث والدراسات في أنظمة التوصية باستمرار لأهميتها في تقديم التوصيات في مختلف المجالات العلمية، التعليمية، الترفيهية وبالأخص التجارية. يوجد أنواع عديدة لأنظمة التوصية أشهرها وأكثرها استخداماً أنظمة الفلتر التعاونية (Collaborative Filtering CF). تقوم أنظمة التوصية CF بتقديم توصيات للمستخدمين الهدف من خلال إيجاد العلاقات الخطية بين التفاعلات السابقة (التقييمات) للمستخدمين مع العناصر، ولكنها تعاني من مشاكل عديدة مثل الإقلاع البارد وتبعثر المعطيات الناتج من قلة التقييمات الخاصة بالعنصر من قبل كل مستخدم. يمكن حل تلك المشاكل بتسخير معطيات جديدة غير التقييمات التي تعتمد عليها أنظمة التوصية CF فقط وذلك من خلال نماذج تعلم الآلة (Machine Learning ML) التي يمكن من خلالها الاستفادة من جميع ميزات العناصر والمستخدمين المتوفرة في مجموعة المعطيات، إذ أنّ معظم نماذج ML قادرة على التعلم من العلاقات الخطية وغير الخطية بين ميزات المعطيات والهدف للحصول على نتائج تنبؤية جيدة للتقييمات. في هذا البحث، سنقوم باستخدام نماذج ML لتمكين أنظمة التوصية CF من الحصول على نتائج تنبؤية أفضل للتقييمات. أظهرت نتائج

¹ طالب دكتوراه - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة البعث.

² أستاذ دكتور - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة البعث.

الدراسة تفوق المنهجية المقترحة على كل من نتائج نماذج ML وأنظمة CF على حداء، وحقت النماذج المدمجة ذات طرق التكديس أعلى نسب تحسين، وخاصة بالنسبة للنموذج المقترح المطبق الناتج من دمج نموذج التكديس Stacking XGBoost مع نظام التحليل العاملي للمصفوفة (Matrix Factorization ALS).
قمنا باستخدام مجموعة معطيات Book Crossing التي تحتوي على ميزات الكتب والمستخدمين وتقييمات الكتب التي تُعبر عن تفاعل المستخدم مع الكتب، ومقياس الأداء التنبؤي للنماذج والأنظمة المستخدمة في الدراسة هو مقياس دقة الخطأ جذر متوسط مربع الخطأ (Root Mean Square Error RMSE).

الكلمات المفتاحية: تعلم الآلة، التصنيفية التعاونية، التكديس، التجميع، تعزيز التدرج.

Collaborative Filtering Recommendation Systems Performance Boosted with Machine Learning Models

Eng. Mohammed Deeb³, Dr. Nasser Abu Saleh⁴

Faculty of Informatics Engineering, Al-Baath University

Abstract

Research and studies on recommendation systems are constantly increasing due to their importance in providing recommendations in various scientific, educational, entertainment, and especially commercial fields. There are many types of recommendation systems, but the most famous and most widely used are Collaborative Filtering (CF) systems. CF recommendation systems make recommendations to target users by finding linear relationships between the users' past interactions (ratings) with items. However, they suffer from several problems such as Cold Start and Data Sparsity resulting from the lack of ratings for the item by each user. These problems can be solved by harnessing new data other than ratings on which CF recommendation systems only depend, through machine learning (ML) models that can take advantage of all the features of the items and users available in the data set, as most ML models are able to learn from linear relationships and nonlinearity between dataset features to obtain good predictive results of ratings. In this paper, we will use ML models to enable CF systems to obtain better predictive results for ratings. The results of the study showed the superiority of the proposed methodology over each of the results of ML models and

³ Master's student - Department of Software Engineering and Information Systems, Faculty of Informatics Engineering.

⁴ Professor Dr. - Department of Software Engineering and Information Systems, Faculty of Informatics Engineering.

CF systems separately, and the merged models with stacking methods achieved the highest improvement rates, especially for the applied proposed model resulting from merging the Stacking XGBoost model with the Matrix Factorization ALS system. We used the Book Crossing dataset that contains books and users' features, and the ratings that express user interaction with the books, and the predictive performance measure of the models and systems used in the study is Root Mean Square Error RMSE.

Keywords— Machine Learning || Collaborative Filtering || Stacking || Ensemble || Gradient Boosting.

1- المقدمة:

أنظمة التوصية أصبحت جزء لا يتجزأ في معظم المواقع الالكترونية لدى العديد من الشركات العالمية في مختلف المجالات العلمية، التعليمية، الترفيهية، والتجارية. وبسبب الازدياد والنمو الهائل للمعلومات والمنتجات المقدمة من تلك المواقع أصبح من الصعب على المستخدمين استعراض جميع تلك المنتجات وإيجاد ما يناسب احتياجاتهم وتفضيلاتهم، بالإضافة إلى الوقت الطويل والجهد المبذول في البحث ضمن تلك المعطيات والمنتجات العديدة للوصول إلى ما يحتاج إليه ويرغب به المستخدمين، وهنا كان لأنظمة التوصية دور في عملية إيجاد أفضل المنتجات التي تناسب تفضيلات واحتياجات المستخدمين [6]. أنظمة التوصية مقسمة إلى عدة تصنيفات [14] وهي: الفترة التعاونية، الفترة المعتمدة على المحتوى، ومزيج هجين من التصنيفين السابقين. من منهجيات أنظمة التوصية الأكثر استخداماً ونجاحاً المعتمدة على الفترة التعاونية CF [22]، إذ أنها تقوم بتقديم قائمة من المنتجات المقترحة عن طريق التنبؤ بتقييمات المنتجات من خلال إيجاد التشابه بين تفاعلات المستخدمين السابقة مع المنتجات (تقييماتهم السابقة للكتب). يتم إيجاد التشابه بين المستخدمين أو الكتب من خلال إيجاد علاقات ارتباط خطية بين تفاعلاتهم السابقة مع الكتب (تقييمات الكتب). التقييمات لها نوعين وهما التقييمات الصريحة والضمنية [20]، التقييمات الصريحة هي التي يقوم المستخدم بتقديمها للمنتجات بشكل مباشر عن طريق مقاييس متعددة تتعلق بطبيعة ما تقدمه تلك المواقع من منتجات وعناصر، مثل الأحادية (يحب)، الثنائية (يحب أو يكره)، مجال عددي محدد (3 من 5)، أو مقياس استينياني (مجال بين موافق بشدة وغير موافق بشدة). أما بالنسبة للتقييمات الضمنية فهي التي يقوم الموقع الالكتروني بالحصول عليها من خلال تفاعلات المستخدم مع المنتجات المطروحة، وتتمثل بشراء منتج ما، البحث عنه بكتابة كلمات مفتاحية، استعراض المنتج، وكل ذلك بدون أن يقوم المستخدم بإعطاء تقييم مباشر لتلك المنتجات. إن التقييمات التي سيتم تطبيق الدراسة عليها هي من النوع الصريح المباشر. أنظمة الفترة التعاونية تستخدم مصفوفة تقييمات $R_{u,i}$ للمستخدمين u والمنتجات i ، تلك المصفوفة تحتوي على عدد قليل من التقييمات إذ لا

يمكن للمستخدم الواحد تقييم عشرات الآلاف من المنتجات والعناصر الموجودة لدى موقع ما، وهذا ما يُطلق عليه مشكلة التبعثر (ضالة المعطيات) Data Sparsity، حيث يقدم نظام التوصية مقترحات ضعيفة غير جيدة لقلة التقييمات المتوفرة [10]. بالإضافة إلى أنّ المستخدمين أو الكتب الجديدة لا تقييمات لهم أو لديهم عدد تقييمات قليل لتساهم في عمل أنظمة التوصية بشكل جيد في تقديم التوصيات للمستخدمين الهدف، وهذه المشكلة تدعى بالإقلاع البارد Cold Start [10]. بالإضافة إلى ذلك، فإنّ أنظمة التوصية تقوم فقط باستخدام العلاقات الخطية بين تقييمات المستخدمين والكتب مع تقييمات المستخدم الهدف عند إيجاد الأكثر تشابهاً معها، ولا تستخدم ميزات أخرى يمكن أن تساهم في زيادة أداء التنبؤ لدى تلك الأنظمة سوى ميزة التقييمات [2]. يوجد العديد من الطرق والتقنيات الأخرى التي تستخدم جميع الميزات الخاصة بالمستخدمين والمنتجات من مجموعة المعطيات في إيجاد تنبؤات بالتقييمات والتي تساهم في تقديم توصيات أفضل، وتعرف تلك التقنيات بنماذج تعلم الآلة [11]. نماذج تعلم الآلة الخاضعة للإشراف التنبؤية تعمل على التنبؤ بتقييمات المنتجات (الكتب) من خلال التعلم من العلاقات الخطية وغير الخطية بين ميزات المعطيات المتوفرة فيما بينها والهدف (التقييمات). وبالتالي سنستخدم نماذج تعلم الآلة الخاضعة للإشراف ذات الطبيعة التنبؤية في زيادة دقة الأداء التنبؤي لأنظمة التوصية ذات الفترة التعاونية CF، والتخلص من مشاكل أنظمة التوصية السابقة من خلال استخدام نماذج تعلم الآلة في التنبؤ بالتقييمات غير المتوفرة لدى أنظمة التوصية CF. قمنا بتطبيق الدراسة والتجارب على عدة طرق ونماذج من خوارزميات تعلم الآلة وأنظمة التوصية المعتمدة على الفترة التعاونية بهدف الحصول على نتائج شاملة للدراسة وإيجاد أفضل مزيج بينها من خلال متابعة دقة أدائها التنبؤي باستخدام مقياس دقة الخطأ التنبؤي RMSE وإيجاد أفضل نموذج من النماذج المقترحة والذي يقدم تنبؤات بأقل خطأ ممكن. نماذج تعلم الآلة المستخدمة في الدراسة هي كالتالي: الجيران الأقرب KNN [19، 26]، متنبئ متجهة الدعم SVR [4، 21]، الغابة العشوائية RF [5، 17، 25]، تعزيز التدرج الخفيف LightGBM [16، 31]، تعزيز التدرج الشديد XGBoost [7، 29]. تم تجميع النماذج الخمسة السابقة باستخدام

4 طرق تكديس [12] وهي: المتوسط Mean، الانحدار الخطي LR، الغابة العشوائية RF، تعزيز التدرج الشديد XGBoost. أما أنظمة التوصية ذات الفترة التعاونية أيضاً تقسم إلى المعتمدة على الذاكرة والمعتمدة على النموذج [10]، سنستخدم في الدراسة أنظمة التوصية المعتمدة على المستخدم User-based CF والمعتمدة على العنصر Item-based CF، ونظامي توصية معتمدين على النموذج وهما: التحليل العملي للمصفوفة MF [18، 28، 30] بطريقتي تحسين هبوط التدرج العشوائي SGD والمربعات الصغرى المتناوية ALS.

في هذا البحث سنقوم بما يلي:

- تقديم 36 نموذجاً مقترحاً ناتجاً عن دمج كل من نماذج ML التسعة مع كل من أنظمة CF الأربعة بهدف حل مشاكل أنظمة التوصية والحصول على أداء تنبؤي أفضل من كل من تلك النماذج والأنظمة على حدا.
- مجموعة المعطيات المستخدمة هي Book Crossing [32] المستخدمة في أبحاث ودراسات عديدة. تحوي ميزات لكل من الكتب والمستخدمين وتقييمات المستخدمين لتلك الكتب. تعاني من القيم المفقودة وكثرة الأخطاء والقيم الشاذة بها، لذلك سنقوم بإجراء العديد من العمليات والطرق للحصول على مجموعة معطيات نظيفة متكاملة صالحة لإجراء التجارب وتقديم أفضل النتائج منها.
- بناء وتطوير تطبيق "Amazon Books Scraper" باستخدام لغة Python وأداة Qt Designer، وذلك بهدف جلب المزيد من ميزات الكتب من أمازون للتخلص من الأخطاء والقيم الشاذة الموجودة في مجموعة المعطيات Book Crossing المستخدمة في الدراسة. ويمكن أن يُستخدم التطبيق من قبل الباحثين لجلب معلومات وميزات عن أي كتب متوفرة لدى موقع أمازون.
- معيار مقياس الأداء المستخدم هو معيار دقة الخطأ التنبؤي RMSE وهو الجذر التربيعي لمربع الخطأ. قمنا باستخدام RMSE في اختيار النماذج ذات أفضل جودة أثناء معايرة بارامتراتها، ومقارنة النماذج فيما بينها لإيجاد الأفضل من حيث الأداء

- التنبؤي في التوقع بتقييمات الكتب، إذ أنه كلما كانت قيمة RMSE أقل، كان الأداء التنبؤي للنموذج أفضل.
- الحصول على نتائج ومقارنات تُقدم أفضل التوقعات لتقييمات الكتب، وذلك بالمقارنة بين نماذج ML وأنظمة CF والنماذج المقترحة.

باقي البحث منسق كالتالي: الفقرة 2 تتحدث عن الأعمال والمنهجيات المتبعة في تسخير نماذج ML في تعزيز الأداء التنبؤي لأنظمة CF، الفقرة 3 تتحدث عن نماذج ML وأنظمة CF والمنهجية المتبعة في هذا البحث، الفقرة 4 تعرض مخطط منهجية البحث وخطوات العمل المتبعة في مرحلة معالجة وبناء مجموع معطيات التجربة، الفقرة 5 تعرض الأدوات ومجموعة المعطيات والتجارب من بناء ومعايرة للنماذج والأنظمة وإجراء المقارنات بينها، وأخيراً النتائج والمناقشة في الفقرة 6 والتي تعرض ما تم التوصل إليه من إجراء هذه الدراسة.

2- الأعمال ذات الصلة:

العديد من الأبحاث والدراسات ناقشت وقامت بتحسين أنظمة التوصية باستخدام منهجيات مختلفة معتمدة على نماذج تعلم الآلة.

ناقش البحث في [3] دراسة تأثير تطبيق طرق التجميع على أداء أنظمة الفترة التعاونية CF، حيث قدمت الدراسة تطبيقاً للعديد من تقنيات التجميع من تعلم الآلة في مجال الفترة التعاونية، ومنها التجميع، التعزيز، الدمج والحقن العشوائي، وتم خلالها تقييم المنهج المقترح على أنواع عديدة من أنظمة CF، مثل: الجيران الأقرب KNN، التحليل العائلي للمصفوفة MF والتحليل العائلي للمصفوفة الجيران Neighborhood MF. أظهرت المقارنات التجريبية تحسناً في الأداء التنبؤي لأنظمة CF الأساس، إذ أظهر الباحثون أن الأداء التنبؤي لتجميع عدة أنظمة CF ضعيفة مثل KNN هو منافس لنموذج CF قوي مثل التحليل العائلي للمصفوفة. تم استخدام مجموعة المعطيات MovieLens، ومقياس دقة الخطأ التنبؤي المستخدم هو RMSE. تم استخدام تقسيم

80% من مجموعة المعطيات للتدريب و20% للاختبار، إذ كانت نسبة التحسين بين 0.87% و2.76%. تقترح الدراسة إضافة المزيد من النماذج الأساس إلى نموذج التجميع للحصول على دقة أفضل.

قام الباحثون "ستار وآخرون" [23] باقتراح إطار عمل هجين يجمع الفترة المعتمدة على المحتوى مع CF لحل مشكلتي الإقلاع البارد وتبعثر المعطيات. أظهرت النتائج أنّ المنهج المقترح أفضل أو قريب من أداء الأنظمة المعتمدة على المحتوى والطرق المهجنة على حدٍ. قام الباحثون باستخدام نماذج تعلّم الآلة ML لحساب التقييمات عوضاً عن مرحلة التنبؤ الخاصة بأنظمة التوصية، حيث بعد عملية إيجاد المستخدمين أو العناصر الجيران المشابهة للمستخدم الهدف يتم استخدام ميزات المعطيات المتعلقة بهم وبناء مجموعة تدريب من أجل نماذج ML لتتعلّم منها وتُقدّم تنبؤات بالتقييمات للعناصر المطلوبة. تم استخدام نماذج ML التالية على حدٍ: آلة متجهة الدعم، أشجار القرار، والمصنف البايزي. أما بالنسبة لمجموعة المعطيات فقد تم استخدام كل من MovieLens و FilmTrust، بالإضافة إلى جلب ميزات نصية للأفلام من موقع قاعدة معطيات الويب للأفلام IMDB. تم استخدام مقياس دقة الخطأ التنبؤي MAE، إذ أنّ أفضل أداء في مجموعة المعطيات FilmTrust كان لنموذج KNNHybrid (NBIB) بقيمة MAE=1.38، وأفضل أداء في مجموعة MovieLens كان لنموذج KNNHybrid (SVMIB) بقيمة MAE=0.74. أظهرت الدراسة أنّ المناهج الهجينة المقترحة تُقدّم نتائج جيدة من أجل المعطيات غير المتوازنة وحالات الإقلاع البارد.

في [27] اقترح الباحثون منهجاً يدمج تنبؤات كل من نظامي الفترة التعاونية CF المعتمد على المستخدم وعلى العنصر باستخدام نموذجي تعلّم الآلة متنبئ متجه الدعم SVR والانحدار الخطي المتعدد MLR. أما بالنسبة لمجموعة المعطيات تم استخدام كل من MovieLens و Hetrec2011، وكان مقياس دقة الخطأ التنبؤي المستخدم MAPE. أظهرت الدراسة أنّ الطرق الهجينة المقترحة باستخدام MLR و SVM قدمت أفضل قيمة

MAPE 1.66 و 1.161، وبنسبة تحسين 9% إلى 11% لكل من نظامي CF المعتمد على المستخدم وعلى العنصر.

في [24] قدم الباحثون نظام فلترة تعاونية CF بمنهجية مقترحة تعتمد على دمج نموذج XGBoost مع تقنية Word2Vec لاكتشاف المنتجات التي تم شراؤها اعتماداً على ميزة نقرات المستخدم. تم جلب ميزات مجموعة المعطيات من موقع الويب للتسوق Jeju، واعتماد نسبة 80% للتدريب و20% لاختبار النموذج المقترح. بالإضافة إلى إجراء عمليات تنظيف وضمان تكاملية المعطيات واستخراج واكتشاف ميزات جديدة منها. و MAE, RMSE كمقاييس لدقة الخطأ التنبؤي. من خلال المقارنات، أظهر النموذج المقترح تفوق واضح بالنسبة لنماذج ML الأخرى مثل: الغابة العشوائية، متنبئ متجه الدعم، والانحدار الخطي.

تقدم الدراسات السابقة نتائج مهمة ومفيدة في استخدام طرق تعلم الآلة بتقديم أنظمة توصية ذات أداء تنبؤي أفضل. سنقوم في هذه الدراسة بتحسين أنظمة CF باستخدام نماذج تعلم الآلة التقليدية المعتمدة على المسافة والشجرية وطرق التكديس ومقارنتها وإيجاد أفضل نموذج مقترح ناتج من دمج نماذج تعلم الآلة مع أنظمة التوصية CF.

3- الخوارزميات والتقنيات المستخدمة في البحث:

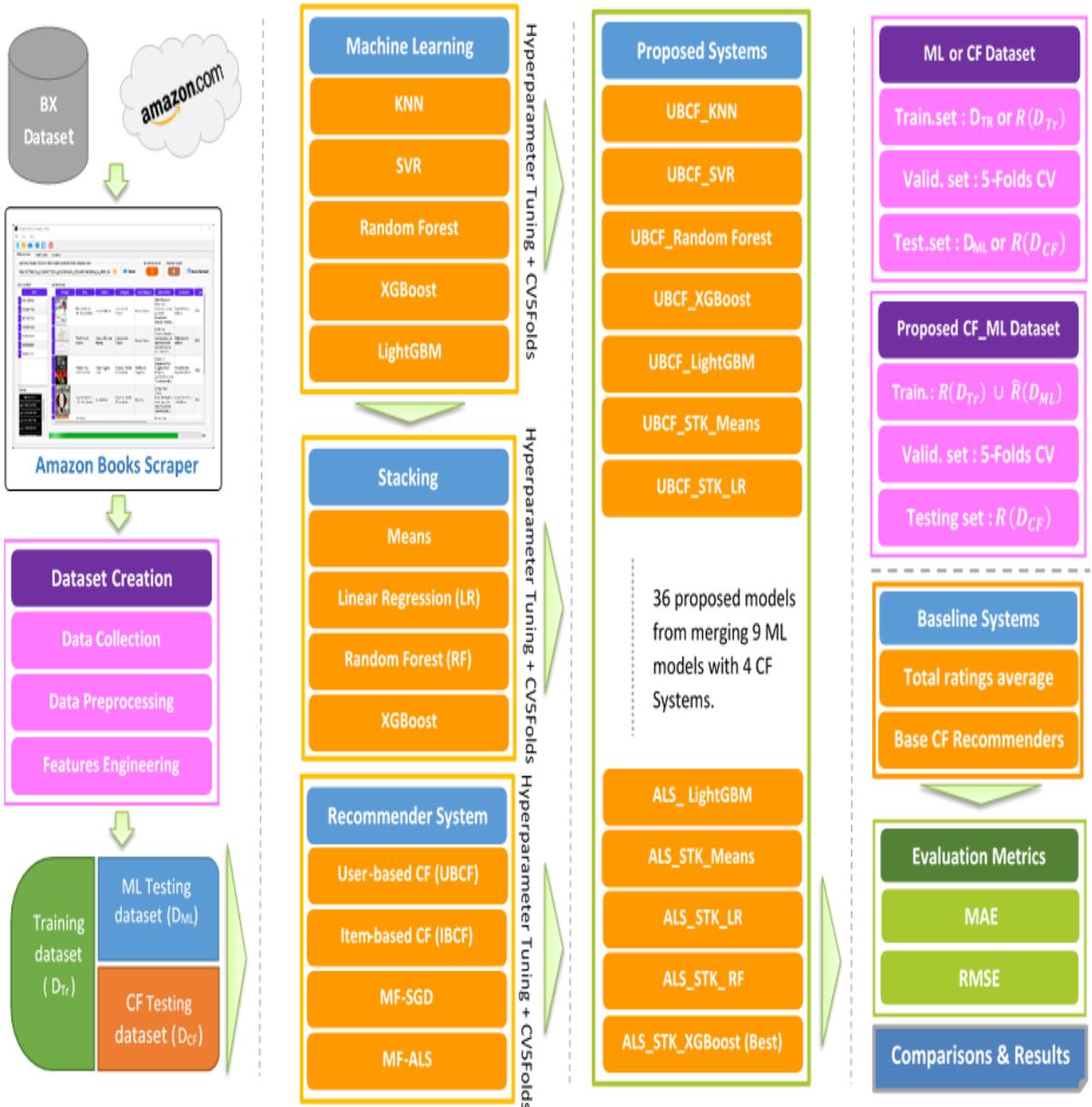
هذا البحث يستخدم 5 نماذج من تعلم الآلة التقليدية (المعتمد على المسافة [1] والمعتمد على الأشجار)، بالإضافة إلى تكديسهم باستخدام 4 نماذج [9]، وذلك بهدف التنبؤ بتقييمات الكتب ليتم إغناء مصفوفة تقييمات أنظمة التوصية CF وبالتالي تحسين نتائج هذه الأنظمة لتقدم بدورها توصيات بدقة أفضل مما كانت تقدمه بدون إغناءها.

نماذج ML المستخدمة في الدراسة هي كالتالي: الجيران الأقرب KNN، متنبئ متجهة الدعم SVR، الغابة العشوائية RF، تعزيز التدرج الخفيف LightGBM، تعزيز التدرج الشديد XGBoost. تم تجميع النماذج الخمسة السابقة باستخدام 4 طرق تكديس وهي:

المتوسط Mean، الانحدار الخطي LR، الغابة العشوائية RF، تعزيز التدرج الشديد XGBoost. أما أنظمة CF سنستخدم في الدراسة المعتمدة على المستخدم User-based CF والمعتمدة على العنصر Item-based CF، ونظامي توصية معتمدين على النموذج وهما: التحليل العاملي للمصفوفة MF بطريقتي تحسين هبوط التدرج العشوائي SGD والمربعات الصغرى المتناوبة ALS. تم شرح عمل ما سبق بالتفصيل في [9].

3-1- النماذج الناتجة من تطبيق المنهجية المقترحة

تعتمد الدراسة على استخدام نماذج تعلم الآلة التقليدية كمغذي لمصفوفة التقييمات التي تستخدمها أنظمة التوصية CF بهدف تسخير العلاقات غير الخطية بين ميزات المعطيات المتعلقة بالكتب والمستخدمين، وحل مشاكل الإقلاع البارد وتبعثر المعطيات التي تعاني منها أنظمة CF التي تعتمد فقط على التقييمات المتوفرة لديها والعلاقات الخطية فيما بينها الخاصة بكل من الكتب والمستخدمين [2]. سيتم في فقرة "منهجية البحث" وفقرة "التجارب" شرح النماذج والمنهجية المقترحة خطوة بخطوة مع المخطط والمعادلات وتقسيمات مجموعة المعطيات الناتجة وإجراء المقارنات ومناقشة النتائج.



المخطط (1) منهجية البحث المتبعة وخطوات العمل مع ذكر نماذج ML وأنظمة CF المستخدمة وتقسيمات مجموعة

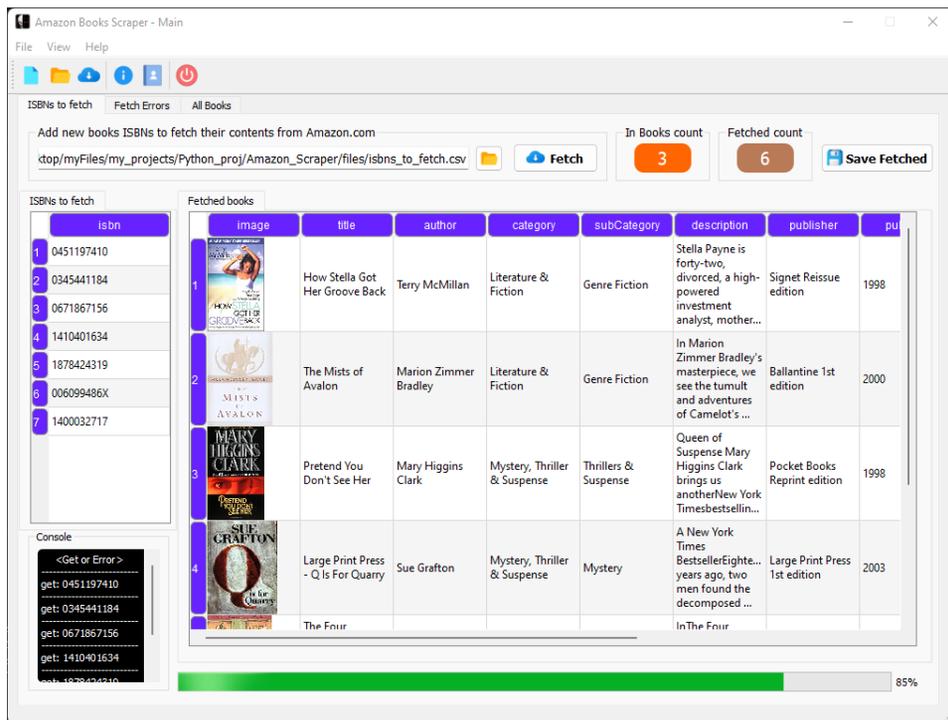
4- منهجية البحث:

المخطط السابق يوضح خطوات العمل المتبعة والمنهجية المقترحة في دمج نماذج تعلم الآلة مع أنظمة التوصية CF وآلية جمع ومعالجة ميزات مجموعة المعطيات المستخدمة في البحث مع ذكر للطرق التي تم استخدامها وتقسيمات مجموعة المعطيات، والمقارنات التي تم إجراؤها أثناء إجراء الدراسة، المخطط (1).

4-1- مرحلة معالجة وبناء مجموعة معطيات التجربة

قمنا باختيار مجموعة معطيات Book Crossing [32] بسبب ضعف النتائج التي تقدمها الأبحاث الأخرى [13، 15] وذلك يعود إلى طبيعة معطياتها الحاوية على أخطاء كثيرة، إذ أنّ ميزات الكتب تحوي العديد من الأخطاء والمعطيات المفقودة التي تحتاج إلى معالجة وتصحيح بهدف ضمان تكاملية المعطيات والحصول على مجموعة معطيات صالحة للاستخدام في الأبحاث والتجارب العلمية وتقديم نتائج صالحة موثوقة جيدة.

يوضح الشكل (1) تطبيقاً برمجياً قمنا ببنائه وتطويره، يقوم ب جلب ميزات الكتب من موقع أمازون للكتب، إذ استخدمنا لغة Python والمكاتب التالية: Qt Designer PyQt5، urllib، Matplotlib، Pandas، BeautifulSoup، requests، re ب جلب ميزات الكتب من موقع أمازون باستخدام رقم الكتاب المعياري العالمي ISBN، ويتم تخزين الميزات الجديدة في قاعدة معطيات محلية SQLite. يسمح التطبيق باستعراض محتويات الكتب التي تم جلبها والوصول إلى صفحة الويب الخاصة بالكتاب في أمازون، وتخزينها إلى ملف CSV لاستخدامها في الدراسة والأبحاث لاحقاً. يدعم استكمال عملية جلب الميزات، ويخزن الأخطاء الحاصلة أثناء عملية جلب الميزات وإمكانية البحث عنها في غوغل، ويسمح باستعراض جميع الكتب الموجودة في قاعدة المعطيات الرئيسية والحصول على مخطط إحصائي يظهر عدد الكتب المنشورة خلال جميع السنوات أثناء عملية جلب الميزات، ويمكن مراقبة تقدم العملية ومعرفة عدد الكتب التي تم جلبها وعدد الأخطاء التي تحدث أثناء تلك المرحلة.



الشكل (1) التطبيق البرمجي Amazon Books Scraper.

قمنا بتنظيف المعطيات وتصحيحها وجلب المفقودة منها من خلال التطبيق البرمجي السابق، بالإضافة إلى تطبيق عمليات عديدة على المعطيات النصية بهدف تصحيحها وضمان تكاملتها، واستخراج ميزات جديدة ممثلة لمزايا مجموعة المعطيات، وإضافة ميزات إحصائية تتعلق بتقييمات الكتب والمستخدمين [9]. وهكذا نكون حصلنا على مجموعة معطيات متكاملة نظيفة قابلة لإجراء الدراسة والتجارب عليها. تم تطبيق عملية التسوية على الميزات العددية لتسهيل عملية إيجاد العلاقات فيما بينها والحصول على تنبؤات أفضل.

قمنا باستخدام اختيار العينات باستخدام التقسيم الطبقي (Stratified Splitting) لتقسيم المعطيات إلى مجموعتي اختبار وتدريب بنسب 20%، 80% بالترتيب، وذلك للحصول على نسب متوازنة للخروج والحصول على مجموعتي تدريب واختبار متوازنتين، وبالتالي الحصول على نتائج تصنيف جيدة موثوقة لتقييمات الكتب.

استخدمنا مقياسي دقة الخطأ RMSE و MAE كمعيار لجودة النماذج والأنظمة لاختيار أفضل البارامترات في مرحلة المعايرة Hyperparameters Tuning، والبحث الشبكي Grid Search الذي يختبر كافة الاحتمالات الممكنة من قيم البارامترات الخاصة لكل من نماذج ML وأنظمة CF لضمان الحصول على أفضل البارامترات في بناء أفضل النماذج والأنظمة لاستخدامها في عمليات الدمج وتطبيق المنهج المقترح والحصول على نتائج منطقية جيدة. تمت معايرة البارامترات الخاصة بخوارزميات التعلم باستخدام طريقة التحقق المتقاطع على 5 تقسيمات لمجموعة التدريب (5-folds Cross Validation) [8].

أما بالنسبة لعملية اختبار النتائج والمقارنات فقد قمنا أيضاً باستخدام مقياس الأداء RMSE ومقارنة النتائج من ناحية نسب التحسين على أنظمة CF الأربعة كأنظمة أساس Baseline لأنماطها المهجنة الناتجة من المنهج المقترح في دمج خرج نماذج ML كدخل لها.

5- التجارب:

5-1- أدوات البحث

- حاسب محمول (Windows 10 Pro 64bit, Core i3 @ 2.5GHz, RAM 4GB).
- بيئة عمل سحابية Google Collaboratory (Xeon Processors @ 2.3Ghz,) (RAM 12.72GB) وذلك لتدريب نماذج Word2Vec لاستخراج الميزات.
- حاسب محمول (Windows 11 Home 64bit, Core i5 @ 2.6GHz, RAM 8GB).
استخدمنا لغة بايثون ومكتباتها NumPy, Pandas, Matplotlib, Seaborn, Sklearn لكتابة وبناء جميع الخوارزميات والنماذج وأنظمة التوصية، وحتى أننا قمنا بكتابة توابع لمعالجة البيانات وبناء تطبيق ذو واجهة رسومية لجلب ميزات معطيات جديدة من أمازون، ومعالجة النتائج قبل عرضها للحصول على مخططات رسومية معيرة. واستخدمنا بيئة عمل Anaconda Jupyter Notebook لكتابة برامج Python & R

مع القدرة على كتابة الملاحظات والمعادلات الرياضية وعرض الرسوميات بشكل مباشر وتوثيق الأكواد البرمجية باستخدام Markdown.

5-2- مجموعة المعطيات

مجموعة المعطيات المستخدمة هي قاعدة البيانات Book Crossing [32]، حيث قمنا بتطبيق طرق لتنظيف البيانات وضمان تكاملتها وحذف التكرارات والتقييمات الضمنية، وجلبنا ميزات إضافية مثل (تصنيف الكتاب الرئيسي والفرعي، عدد الصفحات، ملخص عن الكتاب، اللغة)، مع ذلك فإن بعض الكتب ما زالت تفتقد للبيانات في ميزات لذلك تم حذفها. قمنا باختيار فقط الكتب والمستخدمين الذين لهم ما لا يقل عن 20 تقييماً، وقمنا بتسوية ميزات مجموعة المعطيات Standardization لتسهيل إيجاد العلاقات بينها باستخدام النماذج المعتمدة على المسافة ومقاييس التشابه، وهكذا يكون لدينا مجموعة المعطيات الأفضل للدراسة [9]، جدول (1)، والميزات المستخدمة في البحث كما في الجدول (2).

جدول (1) حجم مجموعة البيانات قبل وبعد معالجتها للدراسة

BX Book Crossing		مجموعة المعطيات
بعد المعالجة	قبل المعالجة	الخصائص
14,239	1,149,780	عدد العينات
1,091	278,858	عدد المستخدمين
603	271,379	عدد الكتب
97.8356 %	99.9985 %	معدل الضالة
	العمر: 26.94 %	
لا يوجد	المؤلف: 3e-4 %	نسب القيم المفقودة
	الناشر: 2e-4 %	
[1, 10]	[0, 10]	مجال التقييمات
69	14	عدد الميزات
[9]	[32]	المرجع

جدول (2) الميزات المستخدمة والمستخرجة في الدراسة

الميزات المستخدمة والمستخرجة	التوصيف
user_Age	عمر المستخدم
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 _ephocs_PCA_1	الميزة الأولى لتحليل المكون الأساسي ذو 100 بعد لعناوين الكتب المطبق عليها عملية lemmatization
w2v_lemmatized_title_100 _ephocs_PCA_2	الميزة الثانية لتحليل المكون الأساسي ذو 100 بعد لعناوين 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
rating_1	عدد تقييمات الكتاب ذات القيمة 1
rating_2	عدد تقييمات الكتاب ذات القيمة 2
rating_3	عدد تقييمات الكتاب ذات القيمة 3
rating_4	عدد تقييمات الكتاب ذات القيمة 4
rating_5	عدد تقييمات الكتاب ذات القيمة 5
rating_6	عدد تقييمات الكتاب ذات القيمة 6
rating_7	عدد تقييمات الكتاب ذات القيمة 7

rating_8	عدد تقييمات الكتاب ذات القيمة 8
rating_9	عدد تقييمات الكتاب ذات القيمة 9
rating_10	عدد تقييمات الكتاب ذات القيمة 10
min_Ri	أدنى تقييم للكتاب
max_Ri	أعلى تقييم للكتاب
median_Ri	الوسيط لتقييمات لكل كتاب
count_Ri	عدد التقييمات لكل كتاب
min_Ru	أدنى تقييم للمستخدم
max_Ru	أعلى تقييم للمستخدم
median_Ru	الوسيط لتقييمات لكل مستخدم
count_Ru	عدد التقييمات لكل مستخدم

بالنسبة لمجال التقييمات قبل المعالجة [0-10]، فقد تم حذف جميع التقييمات الصفرية لأنها تمثل التقييمات الضمنية implicit الغير مباشرة من المستخدمين، إذ أنهم تفاعلوا مع الكتاب ولم يقوموا بتقديم تقييمات عددية مباشرة explicit ضمن المجال [1-10] والذي يعبر عن مدى إعجابهم للكتاب.

5-3- معايير تقييم الأداء التنبؤي

استخدمنا في الدراسة كمقياس للأداء بين نماذج ML وأنظمة CF مقياس دقة الخطأ (جذر متوسط مربع الخطأ RMSE)، أما بالنسبة لاختبار جودة نماذج ML فاستخدمنا RMSE، ومقياس دقة الخطأ (متوسط الخطأ المطلق MAE) لاختبار جودة أنظمة CF، حيث كلما صغرت تلك القيم كان الأداء أفضل:

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

5-4- مراحل بناء ومعايرة نماذج ML وأنظمة CF والمنهجيات المقترحة

5-4-1- تدريب ومعايرة بارامترات نماذج تعلم الآلة التقليدية

في هذه الخطوة نقوم بمعايرة وتدريب 5 نماذج من تعلم الآلة التقليدية على عينات مجموعة المعطيات المستخدمة في الدراسة، إذ أن لكل نموذج مجموعة من البارامترات الفائقة Hyperparameters التي تحسن من دقة تنبؤ تلك النماذج عند معايرتها وتساعد على تجنب مشكلة التناسب المفرط Overfitting والمنخفض Underfitting. جدول (3):

جدول (3) قيم البارامترات المستخدمة في معايرة نماذج ML

ML Model	Parameter	Range	Best
KNN	n_neighbors	5,10, 20, 50, 75, 100, 150	20
SVR	C	5,10, 20, 50, 75, 100, 150	50
	Epsilon (ϵ)	0.01, 0.1, 1, 2,3	1
RF	n_estimators	200, 700	200
	max_depth	1, 3, 5, 10, 20, 30	5
	max_features	'auto', 'sqrt', 'log2'	'auto'
LightGBM	n_estimators	[1, 200]	90
	max_depth	1, 4, 5, 8	4
	learning_rate	0.01, 0.05, 0.1, 0.5	0.05
XGBoost	n_estimators	XGBoost uses 'fit' method	111
	max_depth	to tune its parameters and	4
	learning_rate	returns the best ones.	0.05
	tree_method	fast and memory efficient	'hist'

5-4-2- تكديس نماذج تعلم الآلة باستخدام أربع نماذج مكدسة

في هذه الخطوة نقوم باستخدام 4 نماذج تعلم آلة مكدسة وتدريبها على نتائج وتنبؤات الخطوة السابقة للحصول على نتائج أدق وأكثر شمولية. تستخدم الدراسة كل من نموذج المتوسط العام للتقييمات Mean ratings، الانحدار الخطي Linear Regression، الغابة العشوائية Random Forest، تعزيز التدرج الشديد XGBoost.

5-4-3- تدريب ومعايرة بارامترات أنظمة التوصية CF

في هذه الخطوة نقوم بتدريب ومعايرة بارامترات أنظمة الفترة التعاونية الأربعة (المعتمدة على العنصر، المعتمدة على المستخدم، التحليل العاملي للمصفوفة ALS، SGD)، جدول (4). وذلك بهدف الحصول على أفضل أنظمة منها ليتم اعتمادها كأساس مقارنة مع النماذج الهجينة الناتجة من تطبيق المنهجية المقترحة.

جدول (4) قيم البارامترات المستخدمة في معايرة أنظمة CF

CF System	Parameter	Range	Best
User-based	k (neighbors)	1, 3, 5, 10, 20, 30, 50	10
Item-Based	k (neighbors)	1, 3, 5, 10, 20, 30, 50	5
SGD	k (latent features)	1, 5, 10, 20, 50	20
	epochs	1, 5, 10, 15, 20, 25, 30	20
ALS	k (latent features)	1, 2, 5, 10, 20, 50, 100	20
	epochs	1, 10, 20, 30, 40, 50	30

5-4-4- بناء النماذج الهجينة وفق المنهجية المقترحة

تقوم المنهجية المقترحة على فكرة دمج وتغذية أنظمة التوصية بالتقييمات الناتجة عن نماذج تعلم الآلة بهدف التخلص من مشكلة ضالة المعطيات، الإقلاع البارد، وتحسين

الأداء التنبؤي بالتقييمات من خلال تسخير العلاقات الخطية وغير الخطية بين كامل ميزات مجموعة المعطيات المستخدمة.

تم دمج كل من نماذج تعلم الآلة الخمسة والنماذج المكدسة الأربعة مع كل من أنظمة الفلترة التعاونية الأربعة لنحصل على 36 نموذج هجين مقترح تقدم نتائج وتنبؤات أفضل مما تقدمه أنظمة الفلترة التعاونية الأساسية قبل تطبيق المنهج المقترح.

على سبيل المثال، سنقدم مجموعة من المعادلات المعبرة عن مراحل المنهجية المقترحة لأحد النماذج المقترحة، وهنا سنختار النموذج المدمج ALS_STK_XGBoost الناتج عن دمج نموذج التكديس Stacking XGBoost مع نظام MF-ALS.

سنبدأ أولاً بمجموعة المعطيات الكلية D وتقسيمها إلى مجموعة تدريب ومجموعة اختبار لنماذج ML ومجموعة اختبار لأنظمة CF، كما في المعادلة (3).

$$D = D_{Tr} \cup D_{ML} \cup D_{CF} \quad (3)$$

بعد تحديد مجموعات التدريب والاختبار سنبدأ بتدريب ومعايرة بارامترات نماذج ML، توضح المعادلة التالية استخدام نموذج KNN لمجموعة D_{Tr} للتدريب وللاختبار D_{ML} من اجل التنبؤ بتقييمات المستخدمين للكتب \hat{R}_{KNN} .

$$\hat{R}_{KNN} = KNN(D_{ML}|D_{Tr}) \quad (4)$$

بعد الحصول على تنبؤات تقييمات المستخدمين للكتب من نماذج ML الخمسة نقوم بتكديسها باستخدام 4 نماذج تعلم آلة، في المعادلة (5) مجموعة معطيات \hat{R}_{ML} المكونة من تقييمات نماذج ML الخمسة المستخدمة من قبل المكدسات الأربعة. وتوضح المعادلة (6) حالة استخدام نموذج XGBoost كنموذج لتكديس تنبؤات نماذج ML الخمسة للحصول على تنبؤات بالتقييمات \hat{R}_{stk_xgb} .

$$\hat{R}_{ML} = (\hat{R}_{KNN}, \hat{R}_{SVR}, \hat{R}_{RF}, \hat{R}_{LGBN}, \hat{R}_{XGB}) \quad (5)$$

$$\hat{R}_{stk_xgb} = stk_xgb(\hat{R}_{ML}) \quad (6)$$

أما بالنسبة لأنظمة التوصية CF سنستخدم نفس مجموعة التدريب D_{Tr} وللاختبار D_{CF} ، إذ نستخدم فقط ميزة التقييمات R لتشكيل مصفوفات التقييمات لأنظمة CF. توضح المعادلة التالية استخدام نظام ALS لتقييمات D_{Tr} كمجموعة تدريب وللاختبار تقييمات D_{CF} للحصول على تنبؤات بالتقييمات \hat{R}_{ALS} .

$$\hat{R}_{ALS} = ALS(R(D_{CF})|R(D_{Tr})) \quad (7)$$

بعد الحصول على تنبؤات نماذج ML التقليدية الخمسة ونماذج التكديس الأربعة $\hat{R}(D_{ML})$ (تنبؤات أحد نماذج ML) نقوم بدمجها مع تقييمات مجموعة التدريب $R(D_{Tr})$ لتصبح مجموعة التدريب الخاصة بأنظمة التوصية المدمجة المقترحة كما في المعادلة التالية:

Proposed

$$\text{Training dataset: } R_{ProposedTraining} = \hat{R}(D_{ML}) \cup R(D_{Tr}) \quad (8)$$

dataset:

سنستخدم مصفوفة التقييمات في المعادلة (8) لتدريب الأنظمة المقترحة المدمجة ومصفوفة تقييمات $R(D_{CF})$ للاختبار للحصول على التنبؤات النهائية بالتقييمات لمقارنتها مع تنبؤات أنظمة التوصية CF.

توضح المعادلة (9) حساب التنبؤات للنظام المقترح الناتج من دمج نموذج Stacking XGBoost مع نظام التوصية ALS، إذ أنّ مجموعة التدريب هي اجتماع لتقييمات نموذج التكديس مع مجموعة التدريب الأساسية ومجموعة الاختبار هي نفسها الخاصة

بأنظمة التوصية. عملية الاجتماع هذه تقلل من مشكلة تبعثر المعطيات (ضالة التقييمات).

$$\text{Proposed } \hat{R}_{ALS_stk_xgb} = ALS(R(D_{CF}) | R(D_{Tr}) \cup \hat{R}_{stk_xgb}) \quad (9)$$

يستخدم نظام ALS متوسط التقييمات \bar{R} كعامل أساس لتقديم تنبؤ للمستخدمين والعناصر الجديدة مما يحل مشكلة الإفلاق البارد، ولكن بعد تطبيق المنهجية المقترحة بعملية الدمج يصبح عامل الأساس عبارة عن متوسط مجموع كل من تقييمات التدريب بالإضافة إلى التقييمات الناتجة عن نماذج ML مثل نموذج التكديس XGBoost، كما في العلاقة التالية:

$$\bar{R}_{baseline} = \frac{1}{N} \left(\sum_{i,j} R((i,j) | D_{Tr}) + \sum_{i,j} \hat{R}_{stk_xgb}(i,j) \right) \quad (10)$$

بعد الحصول على التنبؤات النهائية نقوم بتسوية التنبؤات الخارجة عن مجال التقييمات، إذ في حالة تقييمات الكتب فهي ضمن المجال [1, 10] يتم تسوية التقييمات كما يلي:

$$\hat{R}_{Proposed} = \begin{cases} \hat{R}_{Proposed} & , \text{ If } r \in [1, 10] \\ \min(r \in R) & , \text{ If } r < 1 \\ \max(r \in R) & , \text{ If } r > 10 \end{cases} \quad (11)$$

بالنهاية نكون قد قمنا ببناء الأنظمة المقترحة وللحصول على تنبؤات بالتقييمات من أجل الحصول على نتائج لمقارنة أدائها التنبؤي مع نماذج ML وأنظمة CF الأساس. تمثل المعادلة التالية طريقة حساب نسب التحسين من قيم دقة الخطأ RMSE للنماذج والأنظمة بالنسبة للنظام الأساس.

$$\text{Improv. Percent} = \frac{\text{BaselineRMSE} - [\text{CF, ML, Proposed}] \text{RMSE}}{\text{BaselineRMSE}} * 100 \quad (12)$$

5-5- مقارنة الأداء التنبؤي ونسب التحسين

سنقارن بين نماذج تعلم الآلة وتكديسها وأنظمة التوصية والمنهج المقترح في دمجها مع تلك النماذج.

تم إجراء مقارنة بين أنظمة التوصية والنماذج الناجمة عن دمجها مع أنظمة تعلم الآلة، وذلك باستخدام معيار دقة الخطأ التنبؤي RMSE، وقياس نسب تحسينها على نظام الأساس Baseline المتوسط العام لتقييمات الكتب في مجموعة التدريب والمقدر بدقة خطأ تساوي 1.69، وأيضاً نسب التحسين بالنسبة لأنظمة التوصية CF كأساس للنماذج الهجينة الناتجة من تطبيق المنهج المقترح في دمج نماذج تعلم آلة مع أنظمة التوصية، جدول (5) يوضح الاختصارات المستخدمة في الشكل (2)، الشكل (3) والشكل (4)، علماً أنّ السطر والعمود المظلل يمثل قيم النماذج ML وأنظمة CF بدون دمج.

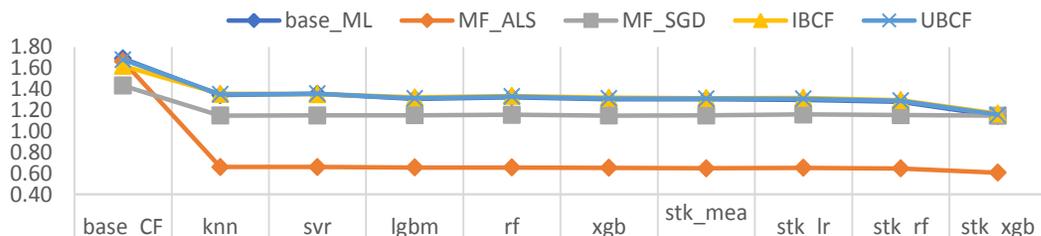
نلاحظ من الشكل (2) والشكل (4) أنّ أفضل النتائج بين نماذج تعلم الآلة كان للنماذج المعتمدة على الأشجار rf، xgb، lgbm بالمقارنة مع النماذج المعتمدة على المسافة، إذ حققت قيم لدقة الخطأ RMSE تساوي 1.327 و 1.302 و 1.307 بالترتيب بنسب تحسين تساوي 21.79% و 22.99% و 22.7% بالترتيب بالنسبة للنظام الأساس متوسط التقييمات y_mean. ونلاحظ أيضاً أنّ طرق التكديس قدمت نتائج أفضل من نماذج ML بدون تكديسها، حيث أفضل نتيجة كانت لنموذج stk_xgb بقيمة RMSE تساوي 1.145 وبنسبة تحسين 32.31%. ونلاحظ أيضاً أنّ نظام التوصية MF_SGD قدم أفضل أداء تنبؤي بالنسبة لباقي أنظمة CF بقيم تصل إلى RMSE = 1.43 ونسبة تحسين بالنسبة للنموذج الأساس بنسبة 15.21% بالترتيب.

بعد تطبيق المنهجية المقترحة في دمج نماذج تعلم الآلة بأنظمة التوصية الأربعة نلاحظ من الشكل (3) والشكل (4) أنّ جميع النماذج المدمجة قد حققت نسب تحسين تفوق النظام الأساس وتفوق نظام التوصية الأساس المطبق عليه عملية الدمج أيضاً. حيث نجد أنّ النماذج المقترحة المطبقة بدمج نماذج تعلم الآلة مع كل من أنظمة التوصية الثلاث IBCF, UBCF, MF_SGD حققت نسب تحسين تتراوح بين 15% إلى 30% بالنسبة لكل من نظام CF الأساس والنظام الأساس y_mean. بينما نلاحظ أنّ النماذج المقترحة المطبقة على نظام التوصية MF_ALS قد حققت نسب تحسين عالية تفوق 60%، حيث أنّ أفضل نموذج مقترح كان بدمج نموذج تكديس stk_xgb مع نظام التوصية MF_ALS والذي قدم أعلى نسبة تحسين تصل إلى 63.65% بقيمة أداء تنبؤي RMSE=0.605 وذلك بالنسبة لنظام أساس MF_ALS، الشكل (2) والشكل (3)، ونسبة تحسين 64.22% بالنسبة للنظام الأساس متوسط التقييمات y_mean، الشكل (4).

جدول (5) اختصارات لأسماء نماذج تعلم الآلة وأنظمة التوصية CF المستخدمة

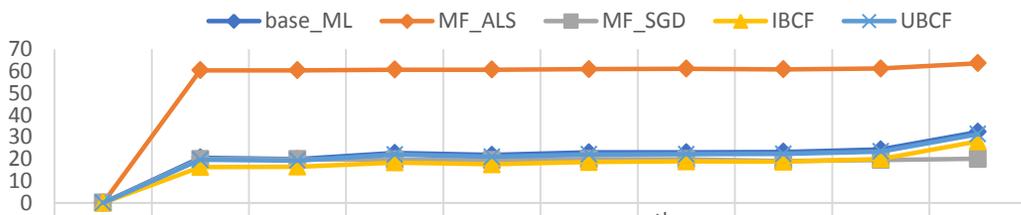
الاختصار	الاسم الكامل
base_CF	Collaborative Filtering Recommenders (Baselines)
knn	K Nearest Neighbors
svr	Support Vector Regressor
lgbm	Light Gradient Boosting Machine
rf	Random Forest
xgb	eXtreme Gradient Boosting
stk_means	Stacking Means
stk_lr	Stacking Linear Regression
stk_rf	Stacking Random Forest
stk_xgb	Stacking eXtreme Gradient Boosting
base_ML	Machine Learning Models (Baselines)
MF_ALS	Matrix Factorization Alternating Least Squares
MF_SGD	Matrix Factorization Stochastic Gradient Descent
IBCF	Item-Based Collaborative Filtering
UBCF	User-Based Collaborative Filtering

ROOT MEAN SQUARE ERROR (RMSE)



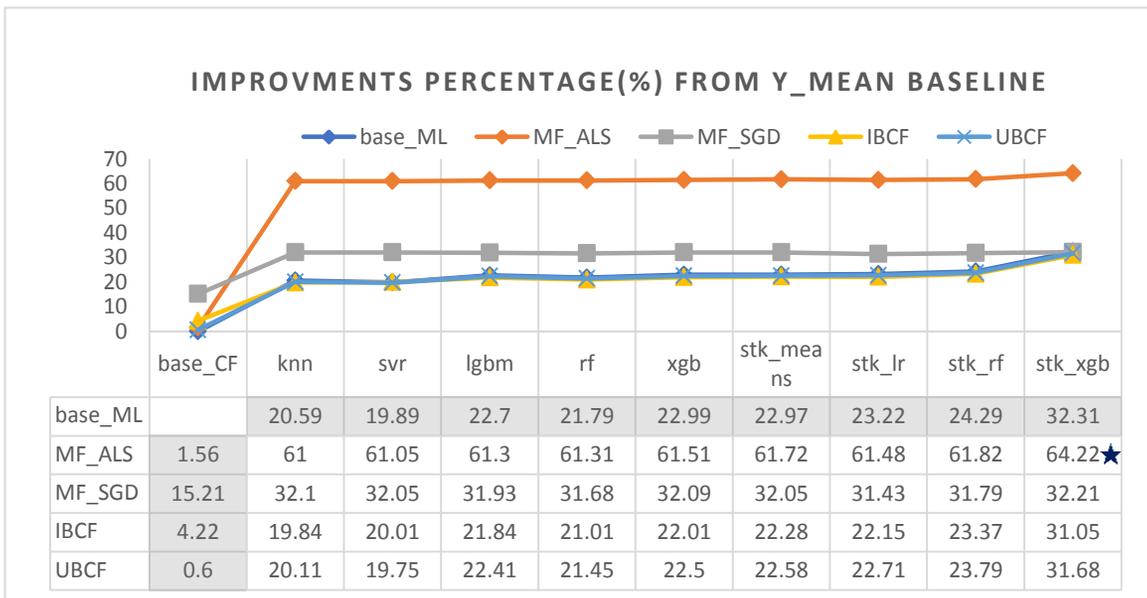
	base_CF	knn	svr	lgbm	rf	xgb	stk_means	stk_lr	stk_rf	stk_xgb
base_ML		1.3430	1.3549	1.3072	1.3227	1.3024	1.3027	1.2984	1.2804	1.1447
MF_ALS	1.6647	0.6596	0.6587	0.6544	0.6543	0.6509	0.6474	0.6514	0.6458	0.6051★
MF_SGD	1.4339	1.1482	1.1491	1.1512	1.1555	1.1485	1.1491	1.1597	1.1536	1.1465
IBCF	1.6198	1.3556	1.3528	1.3218	1.3358	1.3189	1.3144	1.3165	1.2960	1.1660
UBCF	1.6810	1.3510	1.3572	1.3122	1.3284	1.3106	1.3093	1.3071	1.2889	1.1554

IMPROVEMENTS PERCENTAGE(%) FROM CF BASELINES



	base_CF	knn	svr	lgbm	rf	xgb	stk_means	stk_lr	stk_rf	stk_xgb
base_ML		20.59	19.89	22.7	21.79	22.99	22.97	23.22	24.29	32.31
MF_ALS	0	60.38	60.43	60.69	60.7	60.9	61.11	60.87	61.21	63.65★
MF_SGD	0	19.92	19.86	19.72	19.42	19.9	19.86	19.12	19.54	20.04
IBCF	0	16.31	16.48	18.4	17.53	18.58	18.85	18.72	19.99	28.02
UBCF	0	19.63	19.26	21.94	20.97	22.03	22.11	22.24	23.33	31.27

الشكل (3) نسب التحسين على RMSE بالنسبة لأنظمة CF كأساس.



الشكل (4) نسب التحسين على RMSE بالنسبة لمتوسط التقييمات Ratings كأساس.

6- النتائج والمناقشة:

من خلال مقارنات الأداء التنبؤي (قيم RMSE أصغر، أداء تنبؤي أعلى) ونسب التحسين لنماذج ML وأنظمة CF نلاحظ أنّ نماذج ML تتفوق على أنظمة CF وذلك بسبب أنّ تلك النماذج تستخدم جميع ميزات مجموعة المعطيات والعلاقات الخطية وغير الخطية فيما بينها بالمقارنة مع أنظمة التوصية التي تستخدم فقط الميزة الهدف وهي تقييمات المستخدمين للكتب Ratings التي لا تقدم معلومات كافية لتقديم توصيات تتفوق على تلك التي تقدمها نماذج ML، وقد كان للتطبيق البرمجي "Amazon Books Scraper" دور هام في جلب ميزات جديدة لمجموعة المعطيات المستخدمة وتصحيح الأخطاء فيها وضمان تكاملية معطياتها. إنّ أفضل أداء ونسب تحسين كانت للنماذج المعتمدة على الأشجار وبالأخص المعتمد على تعزيز التدرج وذلك لسرعتها وقدرتها على التعامل مع البيانات الكبيرة وتقليل الانحياز والتباين في المعطيات وأدائها

التنبؤي العالي بسبب استخدامها لطرق التحسين هبوط التدرج Gradient Decent .optimization.

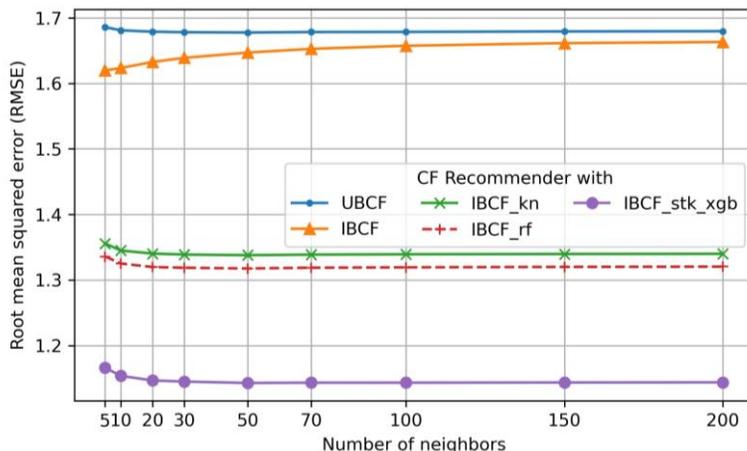
قدمت طرق التكدس الأربعة أفضل النتائج بنسب تحسين +22% عن النظام الأساس المعتمد وذلك من خلال تجميعها لتنبؤات نماذج ML الخمسة، حيث أن النموذج المكس stk_xgb قدم أفضل أداء ونسبة تحسين من تلك الطرق بمقدار +30% حيث أنه سريع، ذو أداء تنبؤي عالي وقدرة على التعلم من العلاقات الخطية وغير الخطية بين ميزات المعطيات وهي هنا التقييمات الناتجة من نماذج ML الخمسة.

بالنسبة للأداء التنبؤي لأنظمة التوصية CF في التنبؤ بالتقييمات فقد كان أقل من أداء نماذج ML إذ أن أعلى نسبة تحسين كانت 15.21%، إذ تستخدم تلك الأنظمة تقييمات المستخدمين فقط ولا يوجد أي استفادة من العلاقات الخطية وغير الخطية بين ميزات المستخدمين والكتب التي تقدم معلومات مهمة لتقديم أداء تنبؤي ونسب جيدة أفضل من تلك التي تقدمها ML.

من خلال تطبيق المنهجية المقترحة في دمج نماذج ML مع أنظمة CF فقد حصلنا على أنظمة مقترحة لأنظمة التوصية CF تفوق من ناحية الأداء التنبؤي ونسب التحسين الأنظمة الأساس CF الأربعة قبل تطبيق الدمج، إذ أنّ إضافة ودمج تنبؤات نماذج ML بمصفوفة الدخل لتقييمات المستخدمين للكتب المستخدمة من قبل أنظمة CF قد كان له دور أساسي في زيادة أداء تلك الأنظمة ويحل مشكلتي الإقلاع البارد وتبعثر المعطيات التي تعاني منها أنظمة CF بسبب قلة التقييمات المتوفرة والتي قمنا بحلها بزيادة عدد التقييمات من خلال التنبؤ بها باستخدام نماذج ML.

من الشكل (5)، نلاحظ أنّ خطأ التنبؤ لنظام التوصية المعتمد على العنصر IBCF يتناسب بشكل منخفض (Underfitting) مع زيادة عدد الجيران بسبب أنّه كلما أضفنا جيراناً أقل تشابهاً مع الهدف حصلنا على عناصر غير متعلقة بالهدف مما يقلل من دقة التنبؤ، بينما الأنظمة المقترحة المدمجة تحسن من دقة التنبؤ لإضافتها المزيد من التقييمات المتنبئ بها من قبل نماذج ML ونلاحظ استقرار الأداء مع زيادة عدد الجيران

بدون حدوث انخفاض في الأداء التنبؤي للأنظمة المقترحة، إذ أن تقييمات الجيران أصبحت أكثر تشابهاً مع الهدف لحصولها على تقييمات ML المتنبئ بها.

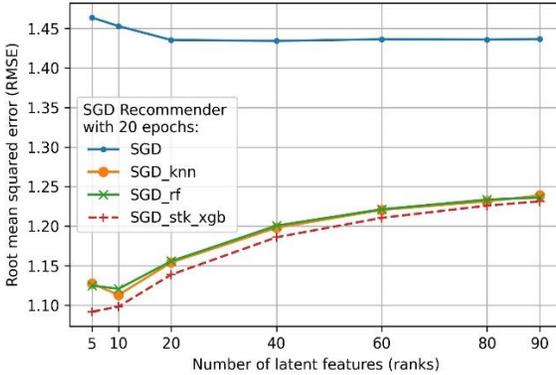


الشكل (5) مقارنة مقياس الخطأ RMSE لكل من أنظمة IBCF الأساس والمقترحة.

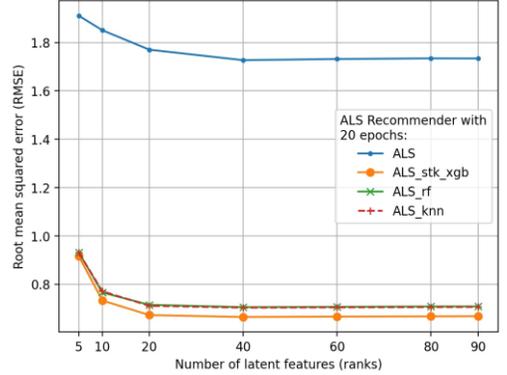
من خلال المقارنات نجد أن أفضل أداء للأنظمة المقترحة كان لأنظمة CF المعتمدة على التحليل العاملي للمصفوفة MF، وكان للأنظمة المقترحة الناتجة من دمج نظام التوصية MF-ALS مع نماذج ML أداء عالي ونسب تحسين عالية جداً تفوق +60% بالنسبة للأساس إذ أنها أقل تأثراً بأخطاء التنبؤ الناتجة من نماذج ML بالمقارنة مع الأنظمة المقترحة لأنظمة CF الأخرى. إن نظام MF-ALS يعتمد على مصفوفتي ميزات كامنة P و Q في تحديث مصفوفة التقييمات للوصول إلى الهدف بأقل خطأ ممكن. قمنا بتهيئة المصفوفتين بقيم عشوائية وتهيئة المصفوفة Q بمتوسط التقييمات قبل عملية إضافة التقييمات الناتجة عن نماذج ML وبالتالي لا يتأثر ALS بالأخطاء كما في SGD، و ALS يقوم بتحديث أوزان كل من المصفوفتين بشكل منفصل مستقل على عكس SGD والذي يقوم بتحديث مستمر متتالي على كل من مصفوفتي الميزات الكامنة، وهذا ما يجعل ALS أقل تأثراً بالأخطاء. سنلاحظ تلك التأثيرات في كل من الشكلين (6، 7). يبين الكود التالي عملية تهيئة المصفوفة Q لأول سطر فيها لتكون منطلقاً لتحديث أوزان باقي قيم المصفوفة:

```
self.P = 3 * np.random.rand(self.k, n_users)
self.Q = 3 * np.random.rand(self.k, n_items)
# Avg. rating for each movie
self.Q[0, :] = self.R[self.R != 0].mean(axis=0)
```

نلاحظ من الشكلين (6، 7) أن أنظمة CF المعتمدة على MF بطريقتي ALS و SGD المدمجة مع نماذج ML المقترحة قد حققت تحسينات ملحوظة على الأداء التنبؤي بتقييمات المستخدمين للكتب بالنسبة للأنظمة CF الأساس، ولكن نلاحظ اختلاف الأداء التنبؤي بين الطريقتين عند 20 دورة تدريب epochs، إذ أنه في الشكل (6 ب) عند زيادة عدد الميزات الكامنة ranks في طريقة SGD نلاحظ أن مقياس الخطأ RMSE يزداد مما يدل على مشكلة التناسب المنخفض Underfitting، وهذه المشكلة تحدث عند ازدياد مقياس خطأ التدريب والاختبار وعدم قدرة النموذج على التعلم والتنبؤ على الرغم من تدريبها على المزيد من العينات. أما طريقة ALS فهي أكثر استقراراً، إذ لا يتأثر أداؤها التنبؤي بعد 20 ميزة كامنة على 20 دورة تدريب، الشكل (6 أ).



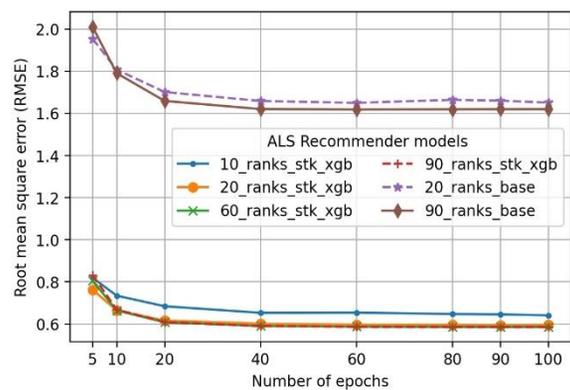
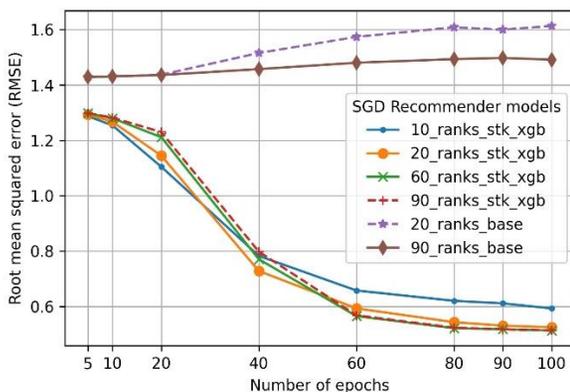
(ب) أنظمة التوصية SGD المدمجة مع نماذج ML KNN, RF, Stacking XGBoost



(أ) أنظمة التوصية ALS المدمجة مع نماذج ML KNN, RF, Stacking XGBoost

الشكل (6) العلاقة بين مقياس الخطأ RMSE وعدد الميزات الكامنة ranks في أنظمة CF التحليل العاملي للمصفوفة MF.

نلاحظ من الشكل (7) أنه بالنسبة لطريقة SGD، يتفوق النظام المقترح المدمج مع Stacking XGBoost على نظام التوصية الأساس SGD ولكن نلاحظ أنه كلما زاد عدد دورات التدريب كلما قل مقياس الخطأ بالمقارنة مع أنظمة ALS المدمجة والتي تصل بسرعة إلى أفضل نتيجة بأقل عدد دورات وتستقر. من الشكل (7 ب) إن زيادة عدد الدورات في SGD للوصول إلى أداء تنبؤي عالي تدل على مشكلة التناسب المفرط Overfitting، إذ أن النظام أصبح يتذكر جميع عينات التدريب ولا يدرك العينات الجديدة. أما طريقة ALS فهي أكثر استقراراً، إذ لا يتأثر أداؤها التنبؤي بعد 20 دورة تدريب حتى لو ازداد عدد الميزات الكامنة، الشكل (7 أ).



(ب) أنظمة التوصية SGD المدمجة مع نموذج

Stacking XGBoost

(أ) أنظمة التوصية ALS المدمجة مع نموذج

Stacking XGBoost

الشكل (7) العلاقة بين مقياس الخطأ RMSE وعدد دورات التدريب epochs

في أنظمة CF التحليل العاملي للمصفوفة MF.

نلاحظ مما سبق أن زيادة العوامل الكامنة ranks لا تؤثر على دقة تنبؤ الأنظمة المدمجة المقترحة في ALS بفضل التقييمات الناتجة من نماذج ML التي أغنت مصفوفة تقييمات أنظمة CF، أما بالنسبة لطريقة SGD فنلاحظ أن زيادة عدد العوامل الكامنة ranks يتطلب زيادة في عدد دورات التدريب وبالتالي استهلاك طاقة معالجة

وذاكرة أكبر وإمكانية الوصول إلى مشكلة التناسب المفرط. تم اختيار عدد الميزات الكامنة بعد الاختبار على مجموعة قيم واختيار الأفضل منها كما في الجدول (4)، إذ 20 ميزة كامنة كانت الأفضل.

بهذا نقترح أن النظام المدمج ALS مع Stacking XGBoost (ALS_stk_xgb) هو أفضل نموذج بين جميع النماذج بأداء تنبؤي عالي بأصغر قيمة خطأ $RMSE=0.6051$ وبتحسين عالي 64.22% على نموذج المقارنة الأساس، وهو أكثر استقراراً، ولا يحتاج إلى دورات تدريب عديدة للوصول إلى أعلى أداء تنبؤي، ولا يحتاج إلى طاقة معالجة أو ذاكرة أكبر بسبب إمكانية معالجتها بشكل موزع على التفرع. بالإضافة إلى ذلك، فإن المنهجية المقترحة تقوم باستخدام تقييم أساس (Baseline rating) كقيمة تنبؤية للمستخدمين والكتب الجدد (حل لمشكلة Cold Start) وهذا التنبؤ هو عبارة عن متوسط التقييمات المتعلقة بمجموعة التدريب الأساسية مضاف إليها تقييمات نماذج ML، كما هو مبين في العلاقة (10)، إذ أن زيادة التقييمات الناتجة عن ML في مصفوفة التقييمات في أنظمة CF تزيد من الدقة التنبؤية للتقييمات الأساس الخاصة بالكتب أو المستخدم، وبالتالي فإن المستخدمين الجدد يحصلون على توصيات بأفضل الكتب على العموم. وبإضافة تقييمات ML إلى مجموعة التدريب الأساسية لأنظمة CF فإننا قللنا من تأثير مشكلة التبعر (Sparsity)، كما هو مبين في العلاقة (9)، إذ أننا قمنا بالتنبؤ بالتقييمات المفقودة في مصفوفة التقييمات في أنظمة CF وتقليل نسبة التبعر. أيضاً قمنا بتقويم جميع التنبؤات خارج مجال التقييمات (1، 10) إذ أن أي قيمة أقل من الحد الأدنى تصبح 1 وأكثر من الحد الأعلى تصبح 10، مما ساهم في زيادة بسطة أيضاً في دقة التنبؤ بالتقييمات.

7- الخاتمة:

في هذا البحث، قدمنا تطبيق يقوم بجلب ميزات ومعلومات للكتب من موقع أمازون ويمكن للباحثين الاستفادة منه في جلب معلومات لأي كتاب متوفر لدى موقع أمازون.

قدمنا أيضاً منهجية مقترحة تقوم على تسخير نماذج تعلم الآلة ML في استخدام العلاقات الخطية وغير الخطية لميزات المستخدمين والكتب بشكل كامل وإغناء مصفوفة تقييمات أنظمة التوصية CF بتقييمات المستخدمين للكتب الناتجة من نماذج ML.

قدمنا في هذا البحث مجموعة من المقارنات والنتائج بين النماذج الهجينة المقترحة ونماذج تعلم الآلة ML وأنظمة التوصية CF المستخدمة في البحث، إذ حصلنا على نتائج تظهر تفوق المنهج المقترح على أنظمة CF الأساس الأربعة.

وجدنا أيضاً أنّ أفضل النتائج من ناحية الأداء التنبؤي ونسبة التحسين كانت لنظام ALS المدمج مع نموذج التعلم Stacking XGBoost، إذ أنّ طريقة التكديس باستخدام XGBoost قد قدمت نتائج جيدة جداً بالمقارنة مع بقية النماذج، واستقرار ALS واستخدام التقييمات من مرحلة ML قد قدم نسبة تحسين عالية بالمقارنة مع باقي نسب التحسين بالنسبة للنظام الأساس تصل إلى 63.65%، علماً النظام الأساس المستخدم في المقارنات والنتائج هو المتوسط العام لتقييمات المستخدمين للكتب في مجموعة التدريب (y_mean)، وحقق نسبة تحسين 64.22% بالنسبة لنظام ALS الأساس.

8- الأعمال المستقبلية:

لقد استخدمنا في هذه الدراسة طبقة واحد من نماذج التكديس للحصول على نتائج جيدة، يمكن استخدام طبقات متعددة للتكديس والتجميع وإجراء المقارنات لنتائجها.

يمكن تطبيق المنهج المقترح على مجموعات معطيات أخرى وعلى مهام مختلفة تبين نقاط قوة المنهج المقترح على نماذج وأنظمة أخرى. بالنسبة للتطبيق البرمجي الذي يقوم بجلب الميزات والمعلومات من أمازون للكتب فإنه من الممكن تطويره ليشمل جميع المنتجات المختلفة في موقع أمازون مما يسمح للباحثين بالحصول على كم وفير من المعلومات وبناء مجموعات معطيات جديدة لمنتجات وعناصر مختلفة تفيد في إجراء الأبحاث عليها.

أحد الميزات التي يمكن استخدامها في دراسات لاحقة هي النصوص، مثل: ملخص عن محتوى الكتاب وردود المستخدمين عن تلك الكتب، وأيضاً الصور، إذ يمكن استخدام شبكات التعلم العميق لاستخراج المعرفة من تلك النصوص والصور والحصول على أنظمة ذات أداء تنبؤي عالي أفضل.

المراجع References

- [1] Abanda A, Mori U, and Lozano JA, 2019. "A review on distance based time series classification", *Data Mining and Knowledge Discovery*, vol. 33, no. 2, pp. 378-412.
- [2] Al-Ghuribi SM and Mohd Noah SA, 2019. "Multi-Criteria Review-Based Recommender System–The State of the Art", *IEEE Access*, vol. 7, pp. 169446-169468.
- [3] Bar A, Rokach L, Shani G, Shapira B, and Schlar A, 2013. "Improving Simple Collaborative Filtering Models Using Ensemble Methods", in *Multiple Classifier Systems*, vol. 7872, Z.-H. Zhou, F. Roli, and J. Kittler, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1-12.
- [4] Basak D, Pal S, and Patranabis DC, 2007. "Support vector regression", *Neural Information Processing-Letters and Reviews*, vol. 11, no. 10, pp. 203–224.
- [5] Breiman L, 2001. "Random forests", *Machine learning*, vol. 45, no. 1, pp. 5–32.
- [6] Burke R, Felfernig A, and Göker MH, 2011. "Recommender Systems: An Overview", *AI Magazine*, vol. 32, no. 3, pp. 13-18.
- [7] 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.
- [8] Choi IY, Oh MG, Kim JK, and Ryu YU, 2016. "Collaborative filtering with facial expressions for online video recommendation", *International Journal of Information Management*, vol. 36, no. 3, pp. 397-402.
- [9] Deeb M and Saleh N, 2019. "Performance Evaluation of Learning Models in Comparison with Recommendation Systems by Studying Data Features", *Journal of Al-Baath University*, vol. 41, pp. 11-52.

- [10] Ekstrand MD, 2011. "**Collaborative Filtering Recommender Systems**", *Foundations and Trends® in Human-Computer Interaction*, vol. 4, no. 2, pp. 81-173.
- [11] Fernández-Delgado M, Sirsat MS, Cernadas E, Alawadi S, Barro S, and Febrero-Bande M, 2019. "**An extensive experimental survey of regression methods**", *Neural Networks*, vol. 111, pp. 11-34.
- [12] Goyal A and Kaur R, 2016. "**A Survey on Ensemble Model for Loan Prediction**", *International Journal of Advanced research and Innovative Ideas in Education (IJARIIE)*, vol. 2, no. 1, pp. 623-628.
- [13] Huang X, Yang Y, and Bao X, 2017. "**Grid-based Gaussian Processes Factorization Machine for Recommender Systems**", in *Proceedings of the 9th International Conference on Machine Learning and Computing*, pp. 92-97, Singapore Singapore: ACM.
- [14] Jalili M, Ahmadian S, Izadi M, Moradi P, and Salehi M, 2018. "**Evaluating Collaborative Filtering Recommender Algorithms: A Survey**", *IEEE Access*, vol. 6, pp. 74003-74024.
- [15] Kannan R, Ishteva M, Drake B, and Park H, 2016. "**Bounded Matrix Low Rank Approximation**", in *Non-negative Matrix Factorization Techniques*, G. R. Naik, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 89-118.
- [16] Ke G *et al.*, 2017. "**Lightgbm: A highly efficient gradient boosting decision tree**", in *Advances in Neural Information Processing Systems*, pp. 3146-3154.
- [17] Kern C, Klausch T, and Kreuter F, 2019. "**Tree-based Machine Learning Methods for Survey Research**", in *Survey Research Methods*, vol. 13, pp. 73-93.
- [18] Koren Y, Bell R, and Volinsky C, 2009. "**Matrix Factorization Techniques for Recommender Systems**", *Computer*, vol. 42, no. 8, pp. 30-37.

- [19] Kotsiantis SB, Zaharakis I, and Pintelas P, 2007. "**Supervised machine learning: A review of classification techniques**", *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3–24.
- [20] Li G and Chen Q, 2016. "**Exploiting Explicit and Implicit Feedback for Personalized Ranking**", *Mathematical Problems in Engineering*, vol. 2016, pp. 1-11.
- [21] Saberioon M, Cisař P, Labbé L, Souček P, Pelissier P, and Kerneis T, 2018. "**Comparative Performance Analysis of Support Vector Machine, Random Forest, Logistic Regression and k-Nearest Neighbours in Rainbow Trout (Oncorhynchus Mykiss) Classification Using Image-Based Features**", *Sensors*, vol. 18, no. 4, p. 1027.
- [22] Sarwar B, Karypis G, Konstan J, and Reidl J, 2001. "**Item-based collaborative filtering recommendation algorithms**", in *Proceedings of the tenth international conference on World Wide Web - WWW '01*, pp. 285-295, Hong Kong, Hong Kong: ACM Press.
- [23] Sattar A, Ghazanfar MA, and Iqbal M, 2017. "**Building Accurate and Practical Recommender System Algorithms Using Machine Learning Classifier and Collaborative Filtering**", *Arabian Journal for Science and Engineering*, vol. 42, no. 8, pp. 3229-3247.
- [24] Shahbazi Z, Hazra D, Park S, and Byun YC, 2020. "**Toward Improving the Prediction Accuracy of Product Recommendation System Using Extreme Gradient Boosting and Encoding Approaches**", *Symmetry*, vol. 12, no. 9, p. 1566.
- [25] Shreyas R, Akshata DM, Mahanand BS, Shagun B, and Abhishek CM, 2016. "**Predicting popularity of online articles using Random Forest regression**", in *2016 Second International Conference on Cognitive Computing and Information Processing (CCIP)*, pp. 1-5, Mysuru, India: IEEE.

- [26] Song Y, Liang J, Lu J, and Zhao X, 2017. "**An efficient instance selection algorithm for k nearest neighbor regression**", *Neurocomputing*, vol. 251, pp. 26-34.
- [27] Thakkar P, Varma K, Ukani V, Mankad S, and Tanwar S, 2019. "**Combining User-Based and Item-Based Collaborative Filtering Using Machine Learning**", in *Information and Communication Technology for Intelligent Systems*, vol. 107, S. C. Satapathy and A. Joshi, Eds. Singapore: Springer Singapore, pp. 173-180.
- [28] Tu X, Liu S, and Li R, 2017. "**Improving matrix factorization recommendations for problems in big data**", in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*(pp. 193-197, Beijing, China: IEEE.
- [29] Xu AL, Liu BJ, and Gu CY, 2018. "**A Recommendation System Based on Extreme Gradient Boosting Classifier**", in *2018 10th International Conference on Modelling, Identification and Control (ICMIC)*, pp. 1-5, Guiyang: IEEE.
- [30] Yu H-F, Hsieh C-J, Si S, and Dhillon IS, 2014. "**Parallel matrix factorization for recommender systems**", *Knowledge and Information Systems*, vol. 41, no. 3, pp. 793-819.
- [31] Zhang H, Si S, and Hsieh C-J, 2017. "**GPU-acceleration for Large-scale Tree Boosting**", *CoRR*, vol. abs/1706.08359.
- [32] Ziegler C-N, McNeer 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.

بناء مجموعة معطيات لتطبيق المنهجية الكلامية اللحنية في تصحيح نطق الأصوات العربية المُفخَّمة حاسوبياً

الباحثة: م. وعد خويص

كلية الهك - جامعة دمشق

الملخص

يجد الناطقون بغير العربية صعوبةً في تعلم بعضٍ من حروفها التي تتفرد بها، ويصعب عليهم نطقها. ويهدف مساعدتهم في تعلم النطق السليم قمنا بدراسة جدوى تطبيق المنهجية الكلامية اللحنية في تصحيح نطق الأحرف المفخمة وهي القاف /ق/ والطاء /ط/ والظاء /ظ/ والصاد /ص/ والضاد /ض/. ولهذا الغرض لزم بناء مجموعة معطيات من التسجيلات الصوتية لكلمات متوافقة مع المنهجية الكلامية اللحنية وتستخدم تقنياتها في التصحيح بما يراعي اللغة العربية وخصوصيتها. ونظراً لعدم توفر مجموعة المعطيات هذه، كان لا بدّ لنا من إعدادها واختبارها على متعلمين أجانب لانتقاء الأفضل منها والذي يساعد على إظهار خاصية التفخيم. وقمنا برسم المنحنيات البيانية للنتائج الإحصائية والمنحنيات الطيفية للإشارات الكلامية باستخدام برنامج ماتلاب. وبذلك تمكنا من تحديد المواضع الأنسب لهذه الحروف في الكلمة وحصلنا على مجموعة المعطيات للكلمات المناسبة لتصحيح النطق والتي يمكن الاستعانة بها في التصحيح الآلي لاحقاً.

الكلمات المفتاحية: المنهجية الكلامية اللحنية - تعليم العربية للناطقين بغيرها - الأصوات المُفخَّمة - حوسبة اللغة العربية - مجموعة معطيات لتصحيح النطق.

Constructing a Dataset to Apply Verbotonal Method for Automated Pronunciation Correction of Emphatic Arabic Sounds

Abstract

It is difficult for non-Arabic speaker to learn some Arabic letters, especially those which do not exist in other languages. In order to help foreigners to learn Arabic pronunciation properly, we studied the feasibility of applying verbotonal method as a new pronunciation correction method, taking emphatic letters as a case study. For this purpose, a dataset of audio files of proper words is necessary. We had to construct this dataset as it is not available for Arabic. Verbotonal method has been applied in practice to foreign learners using well prepared training words and appropriate techniques of correcting the target voices. The curves of statistical results and spectral verbal signals were drawn using special computer software (MATLAB). As a result, the most appropriate positions of emphatic letters in words have been determined to clearly show their characteristics. Finally, a dataset of proper words for pronunciation correction has been constructed to be used, later, in automated correction.

Key words: Verbotonal method – teaching Arabic for foreigners – emphatic sounds – Arabic Language automation – dataset for pronunciation correction.

1- المقدمة

لكل لغة مظهرها الصوتي الذي يتضمن مخارج الأصوات وكيفية نطقها وتفاعل الأصوات المتجاورة بعضها مع بعض، كما يتضمن النبر والتنغيم الخاص باللغة. وبما أن المظهر الصوتي من أهم ما يجب اكتسابه لتعلم اللغة وأول ما يسترعي السمع عند تعلم تلك اللغة، وآخر ما نستطيع تقليده، كان لا بدّ من الاهتمام بتصحيح نطق أصوات اللغة المراد تعلمها [1]. ويجري تصحيح النطق باستخدام منهجيات مختلفة.

ونظراً لغنى اللغة العربية وتنوعها اللفظي وتفردتها بحروف غير موجودة في لغات أخرى، يشكّل تعلمها تحدياً لدى الناطقين بغيرها. الأمر الذي دعانا لدراسة جدوى تطبيق المنهجية الكلامية اللحنية لتصحيح نطق الحروف العربية واخترنا الحروف المفخمة نموذجاً. ولزم لذلك توفر مجموعة معطيات dataset من الكلمات المناسبة للمنهجية الكلامية اللحنية والتي تراعي خصوصية اللغة العربية وتناسبها. ولعدم توفرها، كان لا بدّ لنا من بنائها.

تجدر الإشارة إلى أننا قمنا بنشر ورقة بحثية سابقة تتحدث عن تطبيق المنهجية الكلامية اللحنية في تصحيح النطق آلياً باستخدام نماذج ماركوف المخفية¹ والنتائج التي توصلنا إليها، وقد قمنا بتطوير نظام تصحيح نطق آلي يعتمد على مجموعة معطيات الكلمات التي نستعرض في هذه الورقة مراحل بنائها.

تتضمن هذه الورقة ثلاثة أقسام. يقدم القسم الأول شرحاً تعريفياً بالمنهجية الكلامية اللحنية (Verbotonal Method)²، ثم استعراضاً لكيفية تصحيح النطق باستخدامها. أما القسم

(1) وتشير الدراسة [14] إلى أنّ نماذج ماركوف المخفية هي من أكثر التقنيات نجاحاً في التعرف الآلي على الصوت والتي تعطي نتائج بدقة عالية.

(2) وقد اصطلحنا تسمية هذه المنهجية بالكلامية اللحنية لأننا لم نجد ترجمة مناسبة لها ووجدنا أن هذا التعبير يظهر خصائصها وطريقة عملها.

الثاني، فيتناول بناء مجموعة المعطيات المطلوبة والتجارب التي قمنا بها لذلك. ويستعرض القسم الثالث النتائج التي حصلنا عليها والخاتمة.

2- هدف البحث

2-1- المنهجية الكلامية اللحنية

يمكن تصحيح النطق باستخدام منهجيات مختلفة، منها: المنهجية اللفظية المفصلية (Articulatory Method) [8]، [13]، [14]، ومنهجية الاستماع والإعادة (Listen and Repeat Method) [8]، والمنهجية الكلامية اللحنية. وتركز هذه الأخيرة على تشخيص كل خطأ في اللفظ، وعلى تصحيح الإيقاع والنبر أولاً (أي قبل أي مظهر آخر من مظاهر اللغة المستهدفة)، وتأخذ بالاعتبار أهمية تعريف المتعلم بموسيقا اللغة المستهدفة وإيقاعها ثم التركيز على لفظ الأصوات بدقة [4]، [7].

وبما أن موسيقا اللغة من أهم ما يميّزها فلا بدّ من التركيز على إتقانها لضمان تعلّم اللغة كما ينطقها أهلها [1]، حيث أنّ اللفظ الصحيح أيضاً يسهم في الفهم الصحيح للمعنى المقصود من الكلام [15]، لذلك اخترنا المنهجية الكلامية اللحنية لتطبيقها في هذا البحث.

أسّس هذه المنهجية البروفيسور بيتر غوبيرينا (Petar Guberina)، يوغسلافي الجنسية، وهو أستاذ في علم الصوتيات بجامعة زغرب، عام 1954، ولم تنتشر حتى عام 1970. وتعتبر من أكثر منهجيات تصحيح النطق انتشاراً بسبب النتائج الباهرة التي حققتها، وهي تهدف إلى تنمية مهارات الاستماع والنطق والتحدث [4]، [7].

توصل البروفيسور غوبيرينا إلى فكرة هامة، مفادها أنّ أذن الفرد الذي يتدرب على لغة ثانية غير لغته الأم هي أذن ضعيف السمع نفسها من حيث استقبال الأصوات وإدراكها، فلا بد من أن يُنقى الصوت تنقيّة تامّة لكي يُسمَع على نحوٍ واضحٍ يُساعد الفرد على النطق السليم بدقة [4]، [7].

2-1-1- أسس المنهجية الكلامية اللحنية

تركز هذه المنهجية على العلاقة الأساسية بين إدراك الصوت وإنتاجه، وتتطلق من الخطأ الذي يرتكبه المتعلم. وتولي أهمية بالغة إلى التدريب السمعي والعناصر الفوققطعية (Suprasegmental)⁽³⁾، وتعتمد مبدأ « الإدراك الصحيح من أجل النطق الصحيح ».

ولفهم استراتيجيات المنهجية الكلامية اللحنية في تقويم نطق الأصوات اللغوية، لا بدّ من معرفة مخارج الأصوات اللغوية وصفاتها من ناحية، ومعرفة تصنيف هذه الأصوات إلى حادة وخشنة (أو غليظة) ومحايدة من ناحية ثانية [4]. كما تعمل هذه المنهجية على وضع الأصوات في سياق يتدرج من السهولة إلى الصعوبة للوصول إلى اللفظ الصحيح، وتعتمد الشرح باستخدام حركات الجسد وإيماءاته (كاستخدام اليدين والرأس) لتوضيح خصائص الصوت المراد تعليمه كالشدة والرخاوة والقوة والضعف والغلو والانخفاض وغيرها [7]، [9].

2-1-2- تصحيح النطق باستخدام المنهجية الكلامية اللحنية

تعتمد عملية التصحيح باستخدام المنهجية الكلامية اللحنية على الخصائص التالية للأصوات: الشدة أو المطال (Amplitude)، والزمن أو سرعة الكلام (Tempo)، والتوقف (Pause)، والنغمة (Tone) هل هي حادة (Sharp أو Clear) أو غليظة (Dim أو Dark)، والشد (Tension) هل الصوت مشدود أو مرتخ، وخاصية الجهر والهمس هل الصوت مجهور (Voiced) أو مهموس (Voiceless)، والنبر (Stress) هل الموضع مشدد أو غير مشدد، والمنحني اللحن هل هو صاعد أو هابط، وموضع الصوت من المقطع هل هو أولي أو نهائي [7]، [8]. أما أسلوبا التصحيح الأساسيان في

(3) العناصر الفوققطعية: أي التي تتجاوز الكلمة الواحدة والمقطع الواحد لتشمل علاقة الكلمات والمقاطع المتجاورة بعضها مع بعض وتأثيراتها المتبادلة، وتضم التنغيم، والنبر، وسرعة الكلام [6].

المنهجية الكلامية اللحنية فهما: ضبط الإيقاع والتنغيم (العروض - Prosody)، والاستعانة بالأصوات المرافقة (Combinatory Phonetics) [7].

2-1-3- تطبيق المنهجية الكلامية اللحنية في تصحيح أخطاء النطق بالعربية

سيجري الحديث فيما يلي عن الآلية المقترحة لتحديد نوع الخطأ ومقدار استجابة المنهجية الكلامية اللحنية في حالة اللغة العربية بهدف تصحيح نطق الأصوات المُفخّمة، وتبقى المراحل التالية (إدراج الصوت في جمل والاستعانة بالإيماءات الجسدية) لأبحاث أخرى مستقبلية.

يُعالج النظامُ المقترح تصحيح نطق الأصوات المُفخّمة التالية في اللغة العربية: الطاء /ط/، الضاد /ض/، الصاد /ص/، الظاء /ظ/، القاف /ق/. يبين الجدول 1 خصائص هذه الأصوات المُفخّمة ومثيلاتها المرفقة [2]، [3]، [13].

الطاء	الظاء	الصاد	السين	القاف	الكاف	الضاد	الذال	التاء	التنوين	الاحتكاكية	الرخوة	التشديد	الطعم	الجر	الزنج	الزنج	
X												X	X				الطاء - لثوي - أسناني
	X											X	X				التاء - لثوي - أسناني
		X										X		X			الضاد - لثوي - أسناني
			X									X		X			الذال - لثوي - أسناني
	X	X	X										X				الظاء - أسناني
X		X												X			الذال - أسناني
	X	X	X										X				الصاد - لثوي
X		X	X										X				السين - لثوي
	X											X	X				القاف - لهوي
X												X	X				الكاف - طبقي

الجدول 1. خصائص الأصوات المُفخّمة المعالجة في بحثنا ومثيلاتها المرفقة.

وبناءً على خصائص هذه الأصوات نستخدم الأسلوبين التاليين: ضبط الإيقاع والتنغيم المخصص للصوامت الشديدة والرخوة، والأصوات المرافقة المساعدة على نطق هذه الصوامت.

2-1-3-1- ضبط الإيقاع والتنغيم (العروض)

يجري ضبط الإيقاع والتنغيم للأصوات الشديدة، أي الطاء /ط/ والضاد /ض/ والقاف /ق/، كما يلي:

نزيد الشد بوضع الصامت في الموضع الأولي من المقطع الصوتي المشدد من الكلمة أو المقطع أو المجموعة الصوتية أو الجملة، ونستعمل المنحني اللحني الصاعد، ونزيد سرعة اللفظ، ونستخدم اللفظ المعدل وذلك باستعمال ألوفون⁽⁴⁾ أكثر حدة، مثل " طا " .

كما نعمل على عزل الصوت المطلوب على نحو تقديمي بجعل الصوت المطلوب في بداية المقطع، أو على نحو وسطي بجعل الصوت المطلوب وسط المقطع، مع مراعاة الحصول على مقطع مغلق (ينتهي بساكن) يأخذ فيه الصامت دفعة من الصامت السابق له، مثل " يَطْ " .

أما بالنسبة للأصوات الرخوة، أي الصاد /ص/ والظاء /ظ/، فنعمل على زيادة الارتخاء باستعمال المواضع بين الصوتيات (أي بجعل الصوت الرخو بين صائتين) أو في الموضع النهائي للكلمة أو المقطع، واستعمال المنحني اللحني الهابط، واستخدام اللفظ المعدل بزيادة الارتخاء، فمثلاً عندما تلفظ الزاي سيئاً، والجيم شيئاً؛ نلفظ الزاي أو الجيم بطريقة مسترخية أكثر ما يمكن، كما نعمل على اختيار الألوفون الأكثر ارتخاءً (أي تطويل لفظ الصوت) مثل، " بوظ " .

(4) الألوفون (المتحول الصوتي): يمكن للفونيم (الصوتيم) أو أصغر وحدة صوتية في كلام البشر [15] أن يظهر في أشكال مختلفة حسب الأصوات التي تجاوره وعندها يصبح ألوفوناً [6].

2-3-1-2- الاستعانة بالأصوات المرافقة

لمعرفة الأصوات المساعدة على نطق الأصوات المُفخّمة التي نعالجها في بحثنا، قمنا بتجربة عملية لتعليم نطق اللغة العربية للناطقين بغيرها، وذلك باستخدام مجموعة من الكلمات والمقاطع الصوتية التي تشمل كل الحالات الممكنة لموقع الصوت المراد تعليمه في الكلمة ومجاورته لجميع حروف الهجاء العربية بموضع سابق أو لاحق، واستندنا إلى الطريقة الإحصائية لاختيار الأصوات الأكثر مساعدة وبالتالي اعتمادها في مجموعة معطيات البرنامج. وسنوضح تفاصيل التجربة العملية هذه في الفقرة التالية.

3- بناء مجموعة معطيات كلمات التدريب

يُعدّ بناء مجموعة معطيات التدريب الخطوة الأولى لتطوير نظام التعرف على الكلام [14] وفق المنهجية المتبّعة والتي تلائم خصائصها وتسهم في إنجاحها. ولبناء مجموعة معطيات كلمات التدريب الملائمة للمنهجية الكلامية اللحنية، قسّمنا الأصوات العربية في ثلاث مجموعات مبدئية اعتماداً على ترددها⁽⁵⁾: أصوات حادة وأصوات غليظة وأصوات متوسطة وفق الجدول 2. وذلك للاعتماد عليها في اختيار كلمات التدريب.

غليظ	متوسط	حاد
و	ا	ي
/u/	/a/	/i/
ط	ج	ت
ظ	ث	ز
ص	ش	س
ض	ر	د

⁽⁵⁾ جرت عملية التقسيم اعتماداً على دراسة للمنهجية الكلامية اللحنية مُطبّقة على اللغة الفرنسية [9]، فقمنا بمحاكاة تلك التجربة وإسقاطها على اللغة العربية بشيء من الاجتهاد والابتكار.

حاد	متوسط	غليظ
ن	ل	ق
ف	ك	خ
ح	ذ	غ
—		ب
—		م
ء		ع
—		هـ
ي ساكنة		و ساكنة

الجدول 2. تصنيف الأصوات العربية وفق التردد.

ولدى اختيار الكلمات، بدأنا بمقاطع صوتية لا تعبر عن كلمات حقيقية في اللغة العربية بالضرورة، وذلك بوضع صوت قبل الصوت المفخم أو بعده، ثم استخدمنا كلمات فعلية يتوضع فيها الصوت المفخم في مواضع مختلفة من الكلمة وبحركات إعرابية مختلفة. وتم ترتيب الكلمات من الأسهل للأصعب بالاعتماد على التقسيم الذي اتبعناه في الجدول 2، والذي يعتمد على المجموعة التي ينتمي إليها الصوت المجاور للصوت المفخم، فإذا كان هذا الصوت المجاور ينتمي إلى المجموعة نفسها للصوت المفخم أي غليظاً يكون السياق سهلاً ويأتي أولاً. أما إذا كان من المجموعة المتوسطة فيعتبر أصعب ويأتي ثانياً من حيث الصعوبة. وأخيراً إذا كان من المجموعة الثالثة (يعني حاداً) فيكون الأكثر صعوبة.

وبذلك حصلنا على نحو 150 كلمة وسطياً لكل حرف.

باستخدام كلمات التدريب هذه جرت عملية التدريب للأجانب، ثم قمنا بتحليل نتائجها حاسوبياً.

3-1- تجربة عملية لتعليم الأجنبي

قمنا بتدريب ثمانية طلاب أجنبي من مختلف الجنسيات والأعمار والمستويات في دراستهم للغة العربية وبيّن الجدول 3 معلومات عن الطلاب المدربين.

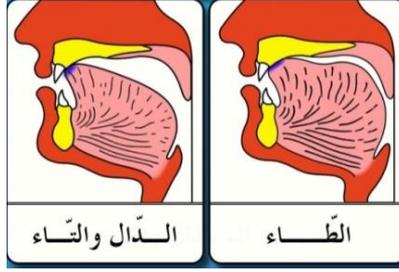
الجدول 3. معلومات عن الطلاب المتدربين.

وجرت عملية التدريب وفق ما يلي:

الاسم	اللغة الأم	لغات أخرى	مستوى اللغة العربية
فكتوريا 1	روسية	إنكليزية	أربعة دروس في تعلم العربية
فكتوريا 2	روسية	إنكليزية	ستة دروس في تعلم العربية
خوسيه	إسبانية	إنكليزية	يعرف بعض الكلمات
أوغو	إسبانية كوبية	إنكليزية - فرنسية	اثنا عشر درساً في تعلم العربية
يوجي 1	يابانية	إنكليزية	تعرف بعض الكلمات
يوجي 2	يابانية	إنكليزية	تعرف بعض الكلمات
مارينيه	روسية	بولونية	تتحدث العربية بطلاقة ولكن لفظها سيء
آجي	هندية	إنكليزية	تعلم ذاتي ويعرف القراءة والكتابة وبعض الكلمات لكن لفظه سيء

1- توضيح مخرج الصوت المفخَّم ومقارنته مع مقابله المرقق والتركيز على موضع

اللسان وشكل جوف الفم بالاستعانة بالصور، كما في الشكل 1.



الشكل 1. مخارج الحروف المُفخّمة ومثيلاتها المرققة [5].

2. نطق كلمة من الكلمات المعدة للتدريب، والتي أخذت بعين الاعتبار وضع الصوت المُفخّم مع حروف الهجاء كافة وفي مواضع مختلفة ومع حركات مختلفة، وطلب تكرارها من المتعلّم.

3. عند نطق المتعلّم للصوت الهدف على نحو خاطئ (كنطق الطاء تاءً)، جرى الانتقال التدريجي من خطأ الطالب (التاء مثلاً ذي التردد الحاد) تدريجياً إلى الصواب (الطاء في هذه الحالة ذي التردد الغليظ) وذلك بوضع الصوت الهدف بجوار صوت أقل حدة لجذبه من التردد الحاد إلى التردد الغليظ (مثل صوت الباء واستخدام كلمة "طَب").

4. إظهار ضرورة التمييز بين الصوتين (المُفخّم ومقابلته المرقق) فونولوجياً، وذلك عن طريق نطق ثنائيات لفظية نحو "سار/صار".

3-2- تحليل النتائج باستخدام الحاسوب

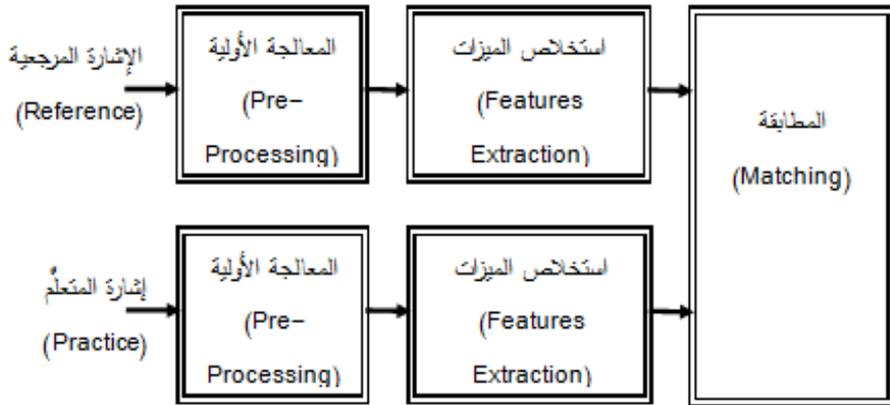
لتحليل النتائج حاسوبياً، سجلنا كلمات التدريب المستخدمة في تجربة تدريب المتعلّمين من الناطقين بغير العربية، وتمت عملية التدريب آنفة الذكر. وسجلنا كذلك كلمات المتعلّمين لاستخلاص معاملات السبيسترال والمنحنيات الطيفية الخاصة بها، وقد استخدمنا برنامج ماتلاب لهذه الغاية، وذلك عن طريق بناء برمجية خاصة لتحليل النتائج.

يتألف برنامج التحليل من مراحل المعالجة التالية [11]:

مرحلة المعالجة الأولية لإشارة الكلام: وتتألف بدورها من مراحل عدة: مرحلة تحصيل الصوت، ويجري فيها التقاط الصوت باستخدام الميكروفون وأخذ عينات من إشارة الصوت، ثم مرحلة ترشيح، تستخدم مرشحاً لتميرير المجال المحصور بين (400Hz – 4KHz). ثم تأتي مرحلة التأطير التي تُستخدم فيها نافذة هامينغ.

مرحلة استخلاص الميزات: وتستخدم تقنية معاملات السبيسترال بمقياس ميل (Mel Frequency Cepstral Coefficients – MFCC) [10].

مرحلة المطابقة: وتأخذ مرحلة المطابقة دخلاً لها معاملات السبيسترال بمقياس ميل الخاصة بالإشارة الكلامية المنطوقة من قبل المتعلم، وتقوم بمقارنة هذه الكلمة مع الكلمة المرجعية بالاستعانة بالمنحنيات الطيفية لكل منهما. يبين الشكل 2 المخطط الصندوقي لبرنامج التحليل.

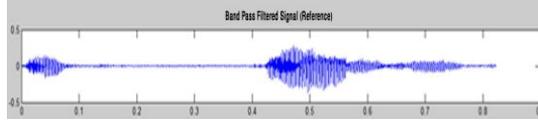


الشكل 2. برنامج التحليل باستخدام ماتلاب.

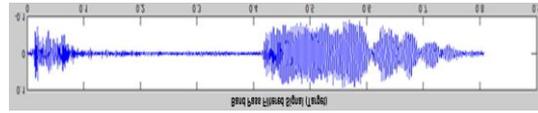
4- نتائج التحليل

باستخدام برنامج التحليل آنف الذكر، استخلصنا معاملات السبيسترال لكلمات التدريب المرجعية والمنحنيات الطيفية الخاصة بها، كما حصلنا على معاملات السبيسترال لكلمات

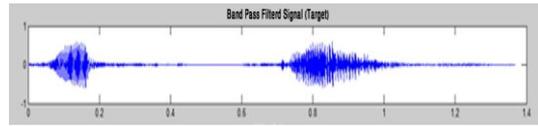
المتعلّمين والمنحنيات الطيفية الخاصة بها، وقمنا بمقارنة المنحنيات الطيفية لكل كلمة من كلمات كل متدرّب مع منحنيات مثيلتها المرجعية لمعرفة مدى التطابق بينهما ودقة اللفظ. يبيّن الشكل 3 الإشارات الكلامية لكلمة " أفُقُّ " بعد تمريرها بمرشح للمجال الترددي (4KHz - 400Hz)، حيث يبين الشكل (أ) الإشارة المرجعية، وتبين الأشكال (ب) و(ت) إشارات كلام المتعلّمين.



(أ) الإشارة المرجعية المرشحة.



(ب) إشارة متعلّم صحيحة مرشحة.

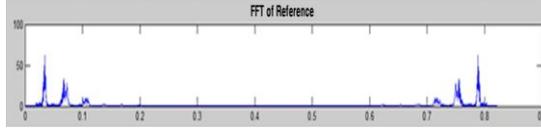


(ت) إشارة متعلّم خاطئة مرشحة.

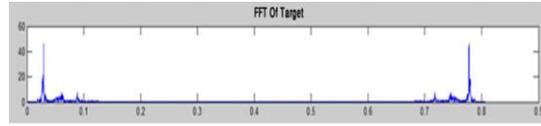
الشكل 3. الإشارات الكلامية لكلمة " أفُقُّ " بعد تمريرها بمرشح تمرير مجال.

ويبين الشكل 4 المنحنيات الطيفية للإشارات السابقة، حيث يظهر الشكل (أ) المنحني الطيفي للإشارة المرجعية، بينما تظهر بقية الأشكال المنحنيات الطيفية لإشارات كلام المتعلمين. بمقارنة الشكلين 4 (ب) و 4 (أ) نلاحظ أنّ المنحنيين الطيفيين متقاربان كثيراً والاختلاف هو فقط في شدة الصوت، حيث كانت شدة صوت المتعلّم هنا أخفض من شدة صوت الإشارة المرجعية. وبالمقارنة بين الشكلين 4 (ت) و 4 (أ)، نلاحظ أنّ المنحنيين الطيفيين مختلفان كثيراً من حيث الشكل إضافة إلى الاختلاف في شدة الصوت،

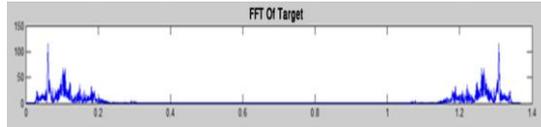
ونتيجة عمليات المقارنة هذه تبين: أن الإشارة الموضحة في الشكل 4 (ب) صحيحة وأن الإشارة الموضحة في الشكل 4 (ت) خاطئة.



(أ) المنحني الطيفي للإشارة المرجعية.



(ب) المنحني الطيفي لإشارة متعلّم صحيحة.



(ت) المنحني الطيفي لإشارة متعلّم خاطئة

الشكل 4. المنحنيات الطيفية لكلمة "أُفُق".

4-1- المواضع الأمثل للأصوات الهدف

تبيّن لنا من خلال نتائج التجربة العملية المذكورة أعلاه، وبالاعتماد على النتائج الإحصائية التي سجلناها بعد تدريب المتعلّمين الأجانب الثمانية، المواضع الأمثل للأصوات الهدف، والتي تُسهّل لفظها، والموضحة في الجدول 4.

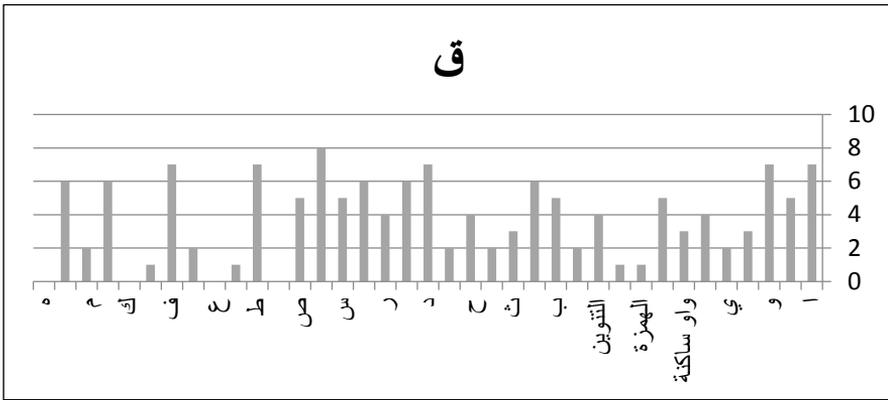
الصوت	الموضع الأفضل للصوت
الطاء /ط/	أول الكلمة مثل " طلب ". فاء الفعل في صيغة على وزن يَفْعَل مثل " يَطْمَح ". عين الفعل في صيغة على وزن يَفْعَل مثل " يَسْطُر ".
القاف /ق/	أول الكلمة مثل " قال ". في أول مقطع وسط الكلمة بعد ساكن نحو " أْفُق ". فاء الفعل في صيغة على وزن يَفْعَل مثل " يَقْلِب ".
الضاد /ض/	أول الكلمة نحو " ضرب ". فاء الفعل في صيغة على وزن يَفْعَل مثل " يَضْحَك ".
الصاد /ص/	آخر الكلمة بعد ساكن وتحديداً الواو نحو " غَوْص ". فاء الفعل في صيغة على وزن يَفْعَل مثل " يَصْرَف ". قبل الصائت نحو " صاب ". بعد الألف نحو " راصف ".
الظاء /ظ/	آخر الكلمة نحو " باظ ". أول الكلمة قبل الألف نحو " ظاق ". عين الفعل في صيغة على وزن يَفْعَل مثل " يَنْظُر ".

الجدول 4. المواضع الأمثل للأصوات الهدف.

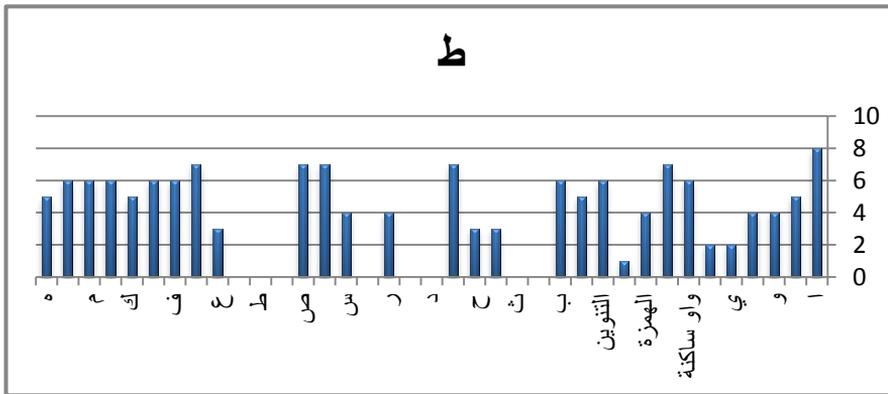
4-2- المنحنيات البيانية

لقد حوّلنا النتائج الإحصائية لتجربة تعليم الأجنبيات المنفّذة لمعرفة المواضع الأمثل للأصوات المُفحّمة إلى منحنيات بيانية تبين علاقة كل صوت مُفحّم مع جميع الحروف

الهجائية (أي مدى سهولة لفظ هذا الصوت المُفخّم عندما يجاوره حرف هجاء ما) المتوضعة على المحور X بينما يمثل المحور Y عدد المتعلّمين. ويبين الشكل 5 المنحني البياني المخصص لحرف القاف. كما يبين الشكل 6 المنحني البياني المخصص لحرف الطاء، والشكل 7 المنحني البياني المخصص لحرف الظاء، والشكل 8 المنحني البياني المخصص لحرف الصاد، والشكل 9 المنحني البياني المخصص لحرف الضاد.



الشكل 5. المنحني البياني المخصص لحرف القاف.

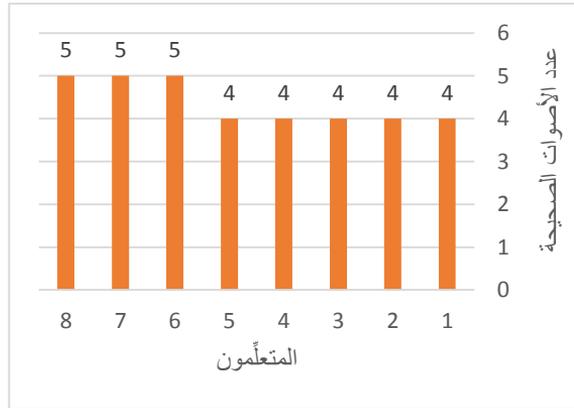


الشكل 6. المنحني البياني المخصص لحرف الطاء.

الصوت	الكلمة	المقطع الهدف	درجة الصعوبة
ص	يَصْرَف	يَصْ	3
ظ	ظَا	ظَا	1
ظ	وَضَّ	وَضَّ	2
ظ	يَنْظُرُ	ظُرُ	3

الجدول 5. مجموعة معطيات الكلمات المستخرجة (نماذج مختارة).

ولا بدّ من الإشارة هنا إلى سرعة تعلّم المتدربين على نطق الأصوات المُفخّمة باستخدام المنهجية الكلامية اللحنية. فقد استغرقت جلسة التدريب الواحدة مدة ساعة من الزمن تم فيها التدريب على نطق الأصوات الخمسة المُفخّمة الموضحة في الجدول 1. ففي بداية الجلسة لم يكن المتدرب قادراً على نطق الأصوات المُفخّمة، بل كانت تخرج معه كمثيلاتها المرققة، وفي نهاية الجلسة الأولى تمكّن معظمهم (5 من 8) من نطق أربعة أصوات من خمسة (أي 80% من الأصوات) على نحو صحيح، وتمكّن الباقي (3 من 8) من نطق الأصوات الخمسة على نحو صحيح، كما يوضح الشكل 10.



الشكل 10. عدد الأصوات الصحيحة لكل متعلّم.

- تسجيل كلمات في مجموعة المعطيات

قمنا بتسجيل الإشارة الصوتية وفقاً لما يلي: التردد 16 KHz والترميز 8 أو 16 بتاً وطريقة التسجيل Mono [12]. كما قمنا بتسجيل كل كلمة من كلمات التدريب للصوت الواحد 20 مرة على الأقل وكل كلمة اختبار 5 مرات. علماً أن كلمات الاختبار كانت مختلفة عن كلمات التدريب. وقد قمنا بتسمية كل ملف صوتي بالكلمة المسجلة فيه. ولهذا الغرض قمنا بترميز الحروف العربية بحروف إنكليزية لتتمكّن من استخدامها حاسوبياً. يوضح الجدول 6 الترميز المستخدم.

وعليه، كان عدد العينات ما يقارب 100 كلمة \times 20 تكرار = 2000 عينة لكل حرف في مجموعة التدريب، و 50 كلمة \times 5 تكرارات = 250 عينة لكل حرف في مجموعة الاختبار.

الأصوات	الرمز الصوتي	الصوت
الصائتة	A	ـَ
	Aa	ا
	U	ـُ
	Uu	و
	I	ـِ
	li	ي
الصامتة	Ax	ء
	B	ب
	T	ت
	Th	ث
	J	ج
	Hh	ح
	Kh	خ
	D	د

الأصوات	الرمز الصوتي	الصوت
	Dh	ذ
	R	ر
	Z	ز
	S	س
	Sh	ش
	Ss	ص
	Dd	ض
	Tt	ط
	Zz	ظ
	C	ع
	Gh	غ
	F	ف
	Q	ق
	K	ك
	L	ل
	M	م
	N	ن
	H	هـ
	W	و
	Y	ي

الجدول 6. الترميز المستخدم لكتابة كلمات مجموعة المعطيات المستنتجة بحروف إنكليزية.

وعليه، تكتب كلمة " طا " بالشكل " TTA "، وكلمة " طَلَبَ " بالشكل " TTALABA ".

5- الخاتمة

تمت في هذا البحث دراسة مراحل بناء مجموعة المعطيات الملائمة لتطبيق المنهجية الكلامية اللحنية على اللغة العربية لتصحيح نطق الأصوات المُفخّمة التي تتميز بها اللغة العربية عن غيرها من اللغات، وطَبّقنا ذلك تجريبياً مع متعلّمين أجانب، وأثبتت التجربة نجاح هذه المنهجية في تصحيح نطق الأصوات التالية (الطاء /ط/)، الضاد /ض/)، الصاد /ص/)، الظاء /ظ/)، القاف /ق/). وبالنتيجة توصلنا إلى تحديد الأصوات المجاورة التي يمكنها المساعدة على إظهار صفة التفخيم وتمكين المتعلّم من إدراكها بصورة أفضل، وذلك لكل حرف من الحروف المُفخّمة. وقد تم إظهار نتائج الدراسة ورسم المنحنيات الطيفية للإشارات الكلامية والمنحنيات البيانية للنتائج الإحصائية بكتابة برمجيات خاصة ضمن نظام ماتلاب. ثم قمنا بتسجيل الكلمات في ملفات صوتية مع تكرارات تتناسب عملية التصحيح المحوسب للأصوات وبذلك أصبح لدينا مجموعة معطيات جاهزة لتطبيق المنهجية الكلامية اللحنية لتصحيح نطق الأصوات المفخمة في اللغة العربية وهي على قسمين مجموعة تدريب من نحو 2000 عينة لكل حرف، ومجموعة اختبار مما يقارب 250 عينة لكل حرف.

6- المراجع

6-1- المراجع العربية

- [1] أنيس، إ، 2010 - الأصوات اللغوية. مطبعة نهضة مصر، مصر 203 صفحة.
- [2] بشر، ك، 2000 - علم الأصوات. دار غريب، القاهرة، 640 صفحة.
- [3] حسان، ت، 1994 - اللغة العربية - معنى ومبنى، دار الثقافة، الدار البيضاء، 373 صفحة.
- [4] أديب، أ، 2010/06/28 - اللفظ المنعم. تمت زيارتها 2023/03/11.
<http://bafree.net/alhisn/showthread.php?t=111257&page=1>
- [5] سويد، أيمن. صور لفظ الأصوات العربية. دار الوثقائي للدراسات القرآنية، دمشق، سورية. تمت زيارتها 2023/05/20.
<https://i.pinimg.com/originals/e1/b4/ec/e1b4ec1e8c57e9aff3d58acd34df228f.jpg>
- [6] الغامدي، م، 2001 - الصوتيات العربية. مكتبة التوبة، الرياض، 196 صفحة.

6-2- المراجع الأجنبية

- [7] INTRAVAIA P, 2003 – Formation des Professeurs de Langue en Phonétique Corrective. Le Système Verbo-tonal.
- [8] LLISTERRI J – El Método Verbo-tonal Corrección de Elementos Suprasegmentales y de Elementos Segmentales. Universitat Autònoma de Barcelona. Retrieved 10/03/2012
http://liceu.uab.es/~joaquim/teaching/Applied_linguistics/L2_Phonetics/Verbo_to_nal/Estrat_correc_VT.pdf
- [9] MAJBOUR A, BILLIERES M – 2008. La Correction Phonétique En Français Pour Les Apprenants Arabophones, Centre de Documentation Pédagogique. (CD-ROM)
- [10] PLANNERER B, 2005 – An Introduction to Speech Recognition. Munich, Germany.
- [11] WANG H, 2002 – Application of Speech Technology to the Assistance of Speech and Auditory Training. National Tsing Hua University, Hsinchu, ISCSLP2002, Taipei.
- [12] JURAFSKY D, MARTIN J, 2007 – Speech and Language Processing. Pearson Education, inc, London. 1037p.
- [13] ALGABRI M, MATHKOUR H, ALSULAIMAN M, BENCHERIF MA, 2022 – Mispronunciation Detection and Diagnosis with Articulatory-Level Feedback Generation for Non-Native Arabic Speech. Mathematics, 10(15), p.2727. MDPI.
- [14] ALQADHEEB F, ASIF A, AHMAD HF. 2021 – Correct Pronunciation Detection for Classical Arabic Phonemes Using Deep Learning. International Conference of Women in Data Science at Taif University (WiDSTaif), pp. 1-6. IEEE.
- [15] ASIF A, MUKHTAR H, ALQADHEEB F, AHMAD HF, ALHUMAM A. 2021 – An approach for pronunciation classification of classical Arabic phonemes using deep learning. Applied Sciences, 12(1):238. MDPI.