

مجلة جامعة البعث

سلسلة العلوم الهندسية الميكانيكية
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 44 . العدد 8

1443 هـ - 2022 م

الأستاذ الدكتور عبد الباسط الخطيب

رئيس جامعة البعث

المدير المسؤول عن المجلة

رئيس هيئة التحرير

أ. د. ناصر سعد الدين

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 : 963 31 2138071 ++

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

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

ISSN: 1022-467X

قيمة العدد الواحد : 100 ل.س داخل القطر العربي السوري

25 دولاراً أمريكياً خارج القطر العربي السوري

قيمة الاشتراك السنوي : 1000 ل.س للعموم

500 ل.س لأعضاء الهيئة التدريسية والطلاب

250 دولاراً أمريكياً خارج القطر العربي السوري

توجه الطلبات الخاصة بالاشتراك في المجلة إلى العنوان المبين أعلاه.

يرسل المبلغ المطلوب من خارج القطر بالدولارات الأمريكية بموجب شيكات

باسم جامعة البعث.

تضاف نسبة 50% إذا كان الاشتراك أكثر من نسخة.

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 (كتابية 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 (20000) ل.س عشرون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.
2. دفع رسم نشر (50000) ل.س خمسون ألف 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 (3000) ل.س ثلاثة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.

المحتوى

الصفحة	اسم الباحث	اسم البحث
28-11	علاء عواد أ.د.م. نصر ياسين	تحديد سماكة المادة المتحولة الطور خلال عملية التجمد حول اسطوانة دورانية تحليلياً وعددياً وتجريبياً
64-29	علي ياسين كمال السلوم وسيم رمضان	تحديد عتبة التصنيف المثلى ديناميكياً في أنظمة الكشف المبكر عن الشذوذ القائمة على التعلم العميق
92-65	منال العمر د. أكرم المرعي د. محسن عبود	تحسين أداء خوارزمية التقطير في انترنت الأشياء
128-93	نهله سعد الدين د. يسر الأتاسي	تطبيق التعلم الجماعي في صنع القرار الطبي

تحديد سماكة المادة المتحولة الطور خلال عملية التجمد حول اسطوانة دورانية تحليلياً وعددياً وتجريبياً

أ.د.م. نصر ياسين**

م. علاء عواد*

المخلص

إن نظام تخزين الطاقة الحرارية الباردة هو وسيلة مبتكرة لتخزين الطاقة ليلاً خارج اوقات الذروة لاستخدامها في اوقات الذروة نهاراً في العديد من المواقع. ومن المعلوم أن تكييف وتبريد الهواء يشكل الاستهلاك الأعظمي للطاقة الكهربائية في فصل الصيف حيث في بعض المناطق تمثل ما يصل إلى نصف معدل الاستهلاك الطاقوي خلال ساعات الذروة في منتصف اليوم، وبما أن المرافق لديها ليلاً قدرة توليد كهربائية، فإن الكهرباء المولدة خلال خارج اوقات الذروة تكون أقل تكلفة بكثير. الهدف الرئيس من هذا البحث هو إيجاد علاقة تربط بين سماكة الطور الصلب المتشكل (الجليد) عند استخدام الماء كمادة متغيرة الطور مع الزمن. تم استخدام طريقة المقاومات الحرارية لإيجاد الحل التحليلي. تم استخدام برنامج ANSYS15 لحل المعادلات الناظمة للمسألة المدروسة بطريقة الحجم المنتهية، وكانت النتائج بين الطريقتين السابقتين متقاربة بنسبة خطأ 9.5% في سماكة الطور الصلب المتشكل بعد 8 ساعات عمل. لمعرفة مدى تطابق الحلول النظرية مع الواقع، تم إجراء تجربة عملية باستخدام محلول غليكول إيثيلين كوسيط تبريد ثانوي (سائل ناقل الحرارة) وكان الخطأ النسبي في سماكة الطور الصلب المتشكل مع الزمن بعد 8 ساعات عمل بين الطريقتين التحليلية والتجريبية 13.5%، وبين الطريقتين الحاسوبية العددية والتجريبية 2.7%.

* طالب دكتوراه في قسم الميكانيك العام في كلية الهندسة الميكانيكية والكهربائية - جامعة دمشق.

** أستاذ في قسم الميكانيك العام في كلية الهندسة الميكانيكية والكهربائية - جامعة دمشق.

Estimate the Solidification Phase Thickness Around a Circular Cylinder Analytically, Numerically and Experimentally

Eng. Alaa Awad*

Dr. Nasr Yasin**

Abstract

CTES (Cold Thermal Energy Storage) is an innovative way of storing night-time off-peak energy for daytime peak use. In many locations, demand for electrical power peaks during summer. Air-conditioning and refrigerating is the main reason, in some areas accounting for as much as half of the power demand during the hot mid-day hours when electricity is most expensive. Since, at night, utilities have spare electrical generating capacity, electricity generated during this “off-peak” is much less expensive. The main objective of this research is to find a relationship between the thickness of the solid phase formed (ice) when using water as a phase change material with time. Thermal resistances method was used to find the analytical solution. ANSYS15 program was used to solve the equations of the problem by the finite volumes method, and the results between the two previous methods were close with an error of 9.5% in the thickness of the solid phase formed after 8 working hours. To find out the accuracy of the theoretical solutions, an experiment was done using an ethylene glycol solution as a secondary cooling medium (heat transfer fluid). The error in the thickness of the solid phase formed with time after 8 working hours between the analytical and experimental methods was 13.5%, and between the numerical and experimental methods was 2.7%.

Keywords: Thermal Ice Storage System (TISS), Thermal resistance, Phase Change Material (PCM).

* PHD student in Mechanical Department at Mechanical and Electrical Engineering Faculty - Damascus University

** Professor in Mechanical Department at the Mechanical and Electrical Engineering Faculty - Damascus University

1. Introduction

The world energy consumption is increasing rapidly, and the developed countries have become heavily reliant upon energy in all activities. It is well accepted now that electrical energy is the most versatile form of energy and its conservation will solve part of the world energy problems. Moreover, in most cases the “offer” and “demand” in thermal energy do not coincide in time. Hence, the obvious key to solve both problems is an efficient storage for low temperature thermal energy.

Bareiss et al. [1] studied melting enhancement above a heated horizontal cylinder by means of numerical and perturbation methods. Zhang et al. [2] studied analytically and experimentally melting of an unfixed solid PCM inside a horizontal tube. Murray et al. [3] developed a successful method for the numerical solution of freezing adjacent to a flat plate. Habeebullah [4] conducted an experimental study on ice formation around a horizontal long copper tube ($L = 12.3$ m with three returned bends of 180° and $d = 19.5$ mm) immersed in water, he found that there was a slope of the ice thickness, in which the axial distance depended on time but varied with coolant flow rate and Stanton and Biot numbers. The axial growth rate of ice was distinct for low values of the coolant Reynolds number and short freezing times. He also discovered an unexpected enlargement of ice thickness on the surface of the tube bends. Sait et al. [5,6] performed an experimental investigation on the freezing of water falling film on a vertical bank of horizontal cold tubes. The brine flows into the concentric tubes in parallel. The authors focused their work on ice formation characteristics and heat transfer for the three main modes of falling film: droplets, jets and sheets. These researchers determined that the formation of ice depends on falling film and coolant flow rates. In addition to that, he found that the thermal resistance of ice controlled the overall heat transfer coefficient. Hosseini et al. [7] experimentally and numerically investigated the melting process of the commercial paraffin RT50 inside a shell-and-tube heat exchanger with water flowing through the tube. It was found that natural convection inside the liquid PCM could enhance the heat storage process.

Yingxin et al. [8] developed a theoretical model to analyze the heat transfer during the melting process. A series of experiments on internal melting of unrestrained ice around a fixed horizontal tube were reported. The temporal geometric shape, melting rates under various experimental conditions were determined using a photograph technique. Sait et al. [9] investigated ice formation on cold vertical banks of horizontal tubes subjected to falling-film – jet mode– experimentally. In the charging process, a set of internally cooled vertical banks of horizontal tubes of brine is subjected to a falling film of water. The formed ice is periodically observed, photographed and measured in falling-film jet mode at specific internal coolant (ethylene–glycol solution) flow rates and temperatures. In the discharge process, the same solution is heated and used internally to release ice. Different thicknesses of the released ice were observed and measured. The maximum quantity of released ice is obtained and the optimum ice formation is determined. Tien et al. [10] were concerned with the transient location of a freezing front in the region outside an infinitely long circular cylinder. The surroundings were assumed to be initially at a uniform temperature above the freezing temperature T_f . The surface of the cylinder was suddenly reduced to, and/or maintained at, a constant temperature, below T_∞ . The frozen and unfrozen regions were assumed to have constant, but different, heat capacities and thermal conductivities. The density is assumed to be the same in both phases. Heat transfer is by conduction only, and the latent heat is all absorbed at T_f .

M. Yari et al. [11] used a numerical method for solving energy equation and describing solidification phenomenon around a circular pipe. Further the effect of pipe surface temperature and initial water temperature to increase freezing zone have been analyzed. Ismail et al. [12] studied a numerical study validated by experimental measurements on the solidification of PCM along a horizontal tube by using the boundary immobilization technique. Jiajia Liu et al. [13] performed a transient simulation of the charging processes based on a two-dimensional numerical model, and the melting processes of the storage units with staggered tube bundle structure and parallel tube bundle structure are compared

with that of flat plate structure. Sugawara et al. [14] investigated Freezing/melting of water/ice around a horizontal cylinder placed in a square cavity of the inner side length numerically.

in this research, we will present the relation between solidification phase thickness around a circular cylinder (δ) and time analytically, numerically and experimentally.

2. Analytical Solution

2-1. Mathematical Model.

Consider a circular cylinder of external radius (r_e) surrounded by PCM of a temperature equal to the phase change temperature (T_f) and heat transfer fluid (HTF) flow inside circular cylinder the temperature of (HTF) lower than the PCM temperature ($T_\infty < T_f$), as seen in Figure (1).

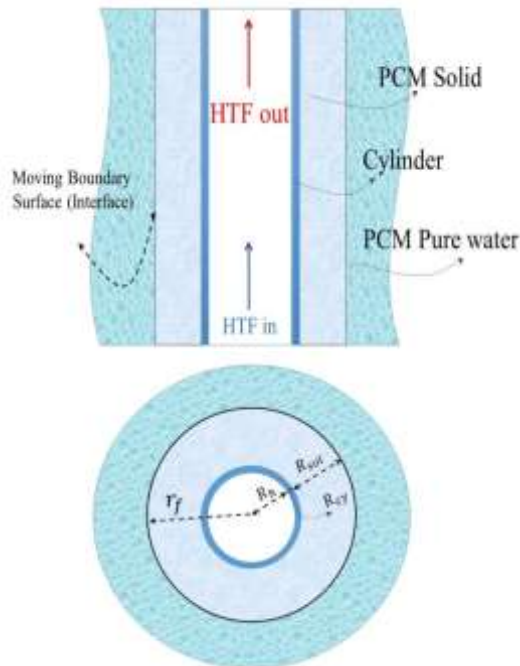


Fig. (1) Schematic view of a cylindrical thermal resistance network.

The heat exchange between the HTF and PCM is dominated by convection heat transfer. The PCM starts to solidify from the outer boundary of the container and the phase change front moves towards the surface of the tank. Three assumptions are taken into considerations:

- 1- The temperature of the free-stream of HTF (T_{∞}) and the heat transfer coefficient (h) are constants.
- 2- The heat transfer process inside the container is conduction dominant and in the radial direction.
- 3- The densities of solid and liquid phases of PCM are equal.

The thermal resistant network has three resistances to deliver the heat from the PCM to HTF, as per Fig (1):

1. Convection resistance at the surface of the cylinder's container (R_h).
2. Conductive resistance at the shell walls (R_{cy}).
3. Conductive resistance for the PCM solid phase (R_{sol}).

The heat delivered from the cylinder to the HTF (solidification) can be calculated by the thermal resistance network as follow:

$$Q_f = \frac{T_f - T_{\infty}}{R_{cy} + R_{sol} + R_h} \quad (1)$$

$$R_{cy} = \frac{\ln(r_{out}/r_{in})}{2\pi l k_{cy}} \quad (2)$$

$$R_{sol} = \frac{\ln(r_f/r_{out})}{2\pi l k_{sol}} \quad (3)$$

$$R_h = \frac{1}{2\pi r_{in} l h} \quad (4)$$

Also the energy balance around the cylinder is:

$$Q_f = \rho L \frac{dv}{dt} = 2\pi \rho l L r_f \frac{dr_f}{dt} \quad (5)$$

Substituting equation (1) into (5), we get:

$$2\pi\rho Lr_f \frac{dr_f}{dt} = \frac{T_f - T_\infty}{\frac{\ln(r_{out}/r_{in})}{2\pi l k_{cy}} + \frac{\ln(r_f/r_{out})}{2\pi l k_{sol}} + \frac{1}{2\pi r_{in} h}} \quad (6)$$

By using first order separation of variables for a homogeneous linear differential equation, the solution would be:

$$\frac{1}{2} \rho L r_f^2 \left[k_{cy} k_{sol} + r_{in} h k_{sol} \ln(r_{out}/r_{in}) + r_{in} h k_{cy} * \left(\ln(r_f/r_{out}) - \frac{1}{2} \right) \right] = r_{in} h k_{cy} k_{sol} (T_f - T_\infty) t + C \quad (7)$$

By using the following initial condition in order to find the constant C in equation (7):

$$r_f = r_{out} \quad at \quad t = 0 \quad (8)$$

The equation (7) becomes:

$$t = \frac{\rho L (r_f^2 - r_{out}^2)}{2(T_f - T_\infty)} \left(\frac{1}{r_{in} h} + \frac{\ln(r_{out}/r_{in})}{k_{cy}} - \frac{1}{2k_{sol}} \right) + \frac{\rho L r_f^2}{2k_{sol}(T_f - T_\infty)} \ln(r_f/r_{out}) \quad (9)$$

If we assume that

$$C_1 = \frac{\rho L}{2(T_f - T_\infty)} \left(\frac{1}{r_{in} h} + \frac{\ln(r_{out}/r_{in})}{k_{cy}} - \frac{1}{2k_{sol}} \right) \quad (10)$$

$$C_2 = \frac{\rho L}{2k_{sol}(T_f - T_\infty)} \quad (11)$$

The time equation (9) becomes:

$$t = C_1 (r_f^2 - r_{out}^2) + C_2 \ln(r_f/r_{out}) r_f^2 \quad (12)$$

2-2. Average convective heat transfer coefficient.

To solve equation (11), we need to obtain the heat transfer coefficient form a flow inside a cylinder, [10] recommends the following relations for the heat transfer coefficient:

if $Re_f < 2300$

$$Nu = 3.66 + \frac{(0.19RePr^{d/l})^{0.8}}{1 + 0.117(RePr^{d/l})^{0.467}} \quad (13)$$

if $2300 < Re_f < 10^4$, $5 < Pr_f < 500$

and $0.05 < \left(\frac{Pr_f}{Pr_w}\right) < 20$

$$Nu = 0.012(Re_f^{0.87} - 280)Pr_f^{0.4} * [1 + \left(\frac{d}{L}\right)^{\frac{2}{3}}] (Pr_f/Pr_w)^{0.11} \quad (14)$$

The fluid properties, in this case, are taken from ref [16] and evaluated at the free-stream temperature (T_∞).

2-3. Results

We will show the relationship between **Radial Interface Position (Solidification Phase Thickness)** and the time. The thermophysical properties shown in Table (1) are used in eq. (12).

Figure (2) shows the relation between solidification phase thickness (δ) and time. At ($t = 0$), the solidification starts at the cylinder outer surface ($\delta = 0$). The thermophysical properties of Ethylene Glycol (EG) which is the heat transfer fluid are taken from ASHRAE [16]. Figure (2) shows a solidification phase thickness after 8 hours equal to 4.2 (cm).

Table (1): Thermo-Physical Properties.

Water Solid Properties

Water initial temperature	273.15 (K) = 0 (°C)
Thermal conductivity of solid PCM	1.88 (W.m ⁻¹ .K ⁻¹)
Density of solid PCM	1000 (kg.m ⁻³)
Heat of fusion	334000 (J/kg)
HTF (EG) Properties	
Volume Concentration	40%
Velocity	1.0 (m.s ⁻¹)
Initial temperature	261.15 (K) = -12 (°C)
Cylinder Materials	
Radius	2 (cm)
Thermal conductivity	110 (W/m.K)
Wall Thickness	1 (mm)

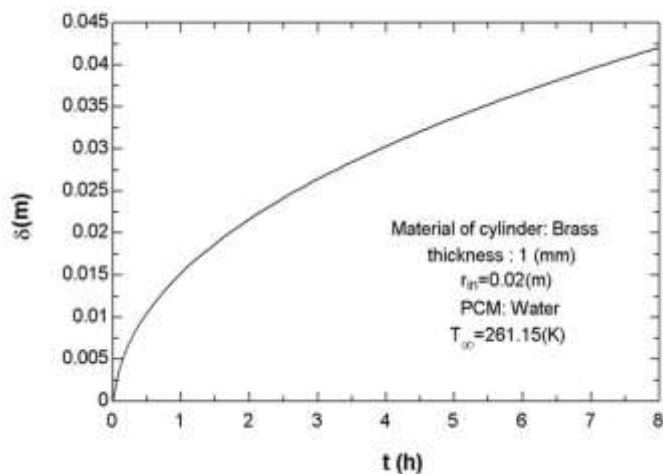


Fig. (2) The relationship between solidification phase thickness and time.

3. Numerical solution

The numerical solution was done by ANSYS program, we used The thermos-physical properties shown in Table (1) Figure (3) shows the relation between solidification phase thickness (δ) and time. At ($t = 0$), the solidification starts at the cylinder outer surface ($\delta = 0$). The thermos-physical properties of Ethylene Glycol (EG) which is the heat transfer fluid are taken from ASHRAE [16]. table (2) shows a relation between solidification phase thickness and time every hour. the solidification was after 8 hours equal to 3.8 (cm).

Table (2): Relation between solidification phase thickness and time.

solidification phase thickness (<i>cm</i>)	Time [<i>h</i>]
0.7	1
1.4	2
2	3
2.6	4
3	5
3.4	6
3.6	7
3.8	8

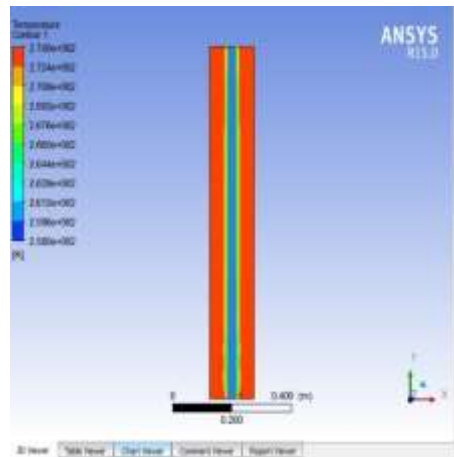


Fig. (4) Temperature distribution by the program after 2 hour.

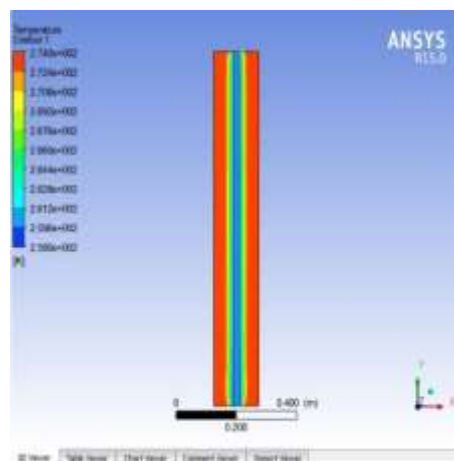


Fig. (5) Temperature distribution by the program after 4 hour.

4. Experiment

The experimental solution was done by used a glass box containing a brass tube of length 1 m. The outside tube diameter was 2.2cm and the inner diameter 2 cm. The glass box was $10 \times 10 \times 100$ cm, fig. (8).



Fig. (8) Experimental water tank.

During the ice formation process, the cooling fluid was a glycol solution of 40% wt concentration whose inlet temperature was -12 °C and a flow rate of $0.5 \text{ m}^3 \cdot \text{h}^{-1}$.

Table (3) shows the relation between solidification phase thickness and time. The solidification phase thickness after 8 hours was equal to 3.7 (cm).

Table (3): Relation between solidification phase thickness and time.

solidification phase thickness (<i>cm</i>)	Time [<i>h</i>]
0.6	1
1.3	2
1.8	3
2.5	4
2.9	5
3.1	6
3.4	7
3.7	8

5. Conclusion

The main objective of this research is to find a relationship between the thickness of the solid phase formed (ice) when using water as a phase change material with time. Thermal resistances method was used to find the analytical solution. ANSYS15 program was used to solve the equations of the problem by the finite volumes method, and the results between the two previous methods were close with an error of 9.5% in the thickness of the solid phase formed after 8 working hours. To find out the accuracy of the theoretical solutions, an experiment was done using an ethylene glycol solution as a secondary cooling medium (heat transfer fluid). The error in the thickness of the solid phase formed with time after 8 working hours between the analytical and experimental methods was 13.5%, and between the numerical and experimental methods was 2.7%.

Fig. (9) Shows the relation between thickness of ice (meter) and time (hours) for Analytical, Numerical and Experimental Methods.

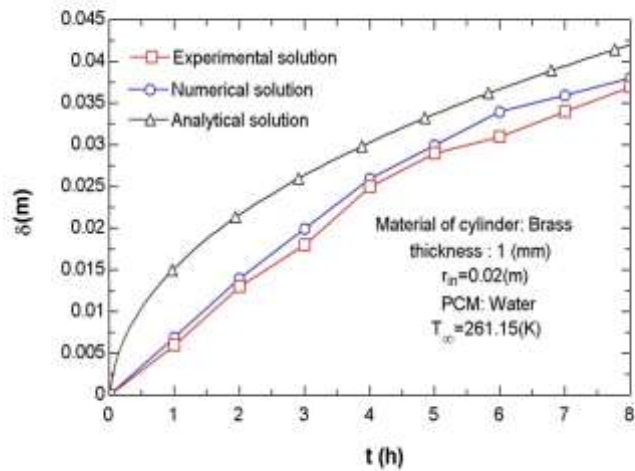


Fig. (9) Relation between thickness of ice and time for Analytical, Numerical and Experimental Methods.

Nomenclature

Q_f	heat flux (W)
L	latent solidification heat ($J.kg^{-1}$)
V	volume (m^3)
r_{in}	internal radius (m)
r_{out}	external radius (m)
r_f	interface radius (m)
δ	solidification phase thickness ($r_f - r_{out}$) (m)
R	thermal resistance ($K.W^{-1}$)
k	thermal conductivity ($W.m^{-1}.K^{-1}$)
h	convective heat transfer coefficient ($W.m^{-2}.K^{-1}$)
T	temperature (K)
T_f	phase change temperature (K)
T_∞	free-stream temperature (K)
t	time (s)
l	length of cylinder (m)
C_p	specific heat ($J.kg^{-1}.K^{-1}$)
g	gravity $m.s^{-2}$
v_{htf}	heat transfer velocity ($m.s^{-1}$)
EG	ethylene glycol

Greek symbols

ρ	density ($kg.m^{-3}$)
μ	dynamic viscosity ($Pa.s$)
ν	kinematic viscosity ($m^2.s^{-1}$)
α	thermal Diffusivity ($m^2.s^{-1}$)

Subscripts

sol	solid
h	free convection on external radius
cy	cylinder

Dimensionless numbers

Nu	Nusselt number	$Nu = \frac{2hR_e}{k}$
Ra	rayleigh number	$Ra = \frac{2VR_e}{\nu}$
Pr	Prandtl number	$r = \frac{\nu}{\alpha} = \frac{\mu C_p}{k}$
Re	Reynolds Number	$Re = \frac{2V_{HTF}R_e}{\nu}$

References

- [1] BAREISS M, BEER, H, 1984, An analytical solution of the heat transfer process during melting of an unfixed solid phase change material inside a horizontal tube. Int J Heat Mass Transfer, VOL 27. 739–746
- [2] ZHANG GP, WEINBAUM, S, 1985, Quasi-three-dimensional steady-state analytic solution for melting or freezing around a buried pipe in a semi-infinite medium. ASME J Heat Transf, VOL. 107. 245–247
- [3] MURRAY, W. D, LANDIS, F, J. Heat Transfer.
- [4] HABEEBULLAH, BA, 2007, An experimental study on ice formation around horizontal long tubes. Int J Refrigerating, VOL.30. 789–801.
- [5] SAIT, HH, HUSSAIN, A, SELIM, AM, 2012, Experimental investigation on freezing of water falling film on vertical bank of horizontal cold tubes. J Thermal Sci Eng Appl, VOL.4. 1-7.
- [6] SAIT, HH, 2012, Heat transfer analysis and effects of feeding tubes arrangement, falling film behavior and backsplash on ice formation around horizontal tubes bundles. Energy Convers Manage. VOL.73. 317–328.
- [7] HOSSEINI, M.J, 2012, A combined experimental and computational study on the melting behavior of a medium temperature phase change storage material inside shell and tube heat exchanger, Int. Commun. Heat Mass Transf. VOL. 39. 1416–1424.
- [8] 10.1016/j.icheatmasstransfer.2012.07.028.
- [9] SAIT, H2013, Heat transfer analysis and effects of feeding tubes arrangement , falling film behavior and backsplash on ice formation around horizontal tubes bundles. Energy Conversion and Management, VOL. 73. 317–328.
- [10] YINGXIN, H, 2001, Heat transfer processes during an unfixed solid phase change material melting outside a horizontal tube, Int. J. Therm. VOL. 1. 1–14.

- [11] SAIT, H, 2014, Charging and discharging characteristics of cool thermal energy storage system with horizontal pipes using water as phase change material.
- [12] YARI, M, Monsef , H, 2010, Numerical Analysis Of Water Solidification Around Horizontals Tubes Using For Ice Storage Systems, Journal of Economics and Engineering, ISSN: 2078-0346.
- [13] KAMAL A.R. ISMAIL, A.M 2014, Experimentally validated two dimensional numerical model for the solidification of PCM along a horizontal long tube. International Journal of Thermal Sciences, VOL. 75.
- [14] JIAJIA, L, 2017. Numerical investigation on the heat transfer enhancement of a latent heat thermal energy storage system with bundled tube structures. Applied Thermal Engineering, VOL. 112. 820–831.
- [15] SUGAWARA, Y, Komatsu, H, 2008. Melting and freezing around a horizontal cylinder placed in a square cavity. Heat Mass Transfer, VOL. 45. 83–92
- [16] ASHRAE Handbook—Fundamentals (2013). CH31. “PHYSICAL PROPERTIES OF SECONDARY COOLANTS (BRINES)”.

تحديد عتبة التصنيف المثلى ديناميكياً في أنظمة الكشف المبكر عن الشذوذ القائمة على التعلم العميق

المهندس علي ياسين¹ - الدكتور كمال السلوم² - الدكتور وسيم رمضان³

المخلص

لازال اكتشاف الشذوذ محط اهتمام الكثير من الباحثين لما له من دور في حل مشكلات العالم الحقيقية مثل الكشف عن الحالات غير الطبيعية ومنها عيوب الأنظمة الصناعية واكتشاف الاحتيال.. الخ.

يهدف البحث إلى بناء نظام عام للكشف المبكر عن حالات الشذوذ، يمكن استخدامه لتقليل من من الخسائر الكبيرة الناجمة عن وجود خلل في بيانات النظام. أهم المشاكل التي تواجه أنظمة كشف الشذوذ الحالية هي عدم قدرتها على (1) استخراج الميزات من بيانات متعددة عالية الأبعاد، (2) تحديد عتبة التصنيف الأمثل بشكل ديناميكي.

إعتمدَ البحث لحل المشاكل السابقة على شبكة الترميز الآلي ذات الذاكرة طويلة المدى (LSTM-Autoencoder) لاستخراج الميزات الأكثر أهمية، والتقاط التبعيات الزمنية لتسلسل الشذوذ ضمن سلاسل البيانات. بالإضافة لنظرية تشبيبيشيف (Chebyshev) لتحديد العتبة الأمثل ديناميكياً.

طُبِقَ النظام المُقترح على مشكلة كسر الورق في معامل صناعة الورق، حيث استطاع اكتشاف 53% من الحالات الشاذة قبل حدوثها بأربعة دقائق، ونسبة إشارات كاذبة لم تتجاوز 9%، والحد من تكاليف الخسارة بنسبة 54%. مما يؤكد قدرة النظام المُقترح على التكيف مع المسألة المطروحة في اكتشاف الشذوذ والتقليل من أثاره.

الكلمات المفتاحية: أنظمة كشف الشذوذ، استخراج الميزات، بيانات متعددة عالية الأبعاد، التبعيات الزمنية، عتبة التصنيف، شبكة الترميز الآلي ذات الذاكرة طويلة المدى، نظرية تشبيبيشيف.

¹ طالب دكتوراه - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سوريا

² أستاذ دكتور - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سوريا

³ دكتور مشارك - قسم الاقتصاد الزراعي - كلية الزراعة - جامعة البعث - حمص - سوريا

Selecting the Optimal Classification Threshold Dynamically in Early Anomaly Detection Systems Based on Deep Learning

Eng. Ali Yassin, Dr. Kamal Al-Salloum, Dr. Wassim Ramadan

Abstract

Anomaly detection is still an important subject for many researchers. It has a crucial role in solving real-world problems such as detecting abnormal cases, including defects in industrial systems, and fraud detection...etc.

The research aims to build a general system for early anomaly detection, which can be used to reduce the large losses caused by abnormality in the system data.

The main problems facing current anomaly detection systems are their inability to 1) extract features from high-dimensional, multiple data, 2) selecting the optimal classification threshold dynamically.

In order to solve the problems above, the research employed an LSTM-Autoencoder network to extract the most important features and capture the temporal dependencies of the occurrence of anomalies within the data series. In addition to Chebyshev's theory for selecting the optimal threshold dynamically.

The proposed system was applied to the problem of paper breaks in paper factories, where it was able to detect 53% of the anomalies four minutes before its occurrence, with a rate of false alarms that did not exceed 9%, and reduce the costs of loss by 54%. This confirms the ability of the proposed system to adapt to the issue at hand in detecting anomalies and minimizing its effects

Keywords: Anomaly Detection Systems, Feature Extraction, Chebyshev Theorem, High-Dimensional Multiple Data, Time Dependencies, Threshold Classification, LSTM-Autoencoder.

1- مقدمة

يعدّ اتخاذ القرارات الصائبة عاملاً رئيسياً في نجاح أي عمل تجاري أو صناعي مهما يكن، لذلك تعدّ مشكلة دعم اتخاذ القرارات المناسبة في إدارة سلسلة التوريد (Supply Chain Management) مصدر قلق كبير لعدد من الدراسات الحديثة [1][2][3]. تتحكم إدارة سلسلة التوريد (SCM) بجميع العمليات التي تُحوّل المواد الخام إلى منتجات نهائية.

يوجد مجموعة من الإجراءات التي تُساعد في دعم اتخاذ القرارات ضمن سلاسل التوريد. يُعتبر كل من التنبؤ (Forecasting) واكتشاف الشذوذ (Anomaly Detection) مهمتين أساسيتين في إدارة سلاسل التوريد (SCM). حيث تُساعد عملية التنبؤ في إدارة المواعيد المحددة وتخطيط الإنتاج والتسعير. لكن تؤدي كميات البيانات الهائلة التي يتم إنشاؤها في كل مرحلة من مراحل سلسلة التوريد، إلى زيادة تحميل البيانات وصعوبة تمييز الإشارات المفيدة التي تُتيح اتخاذ قرارات ذات مغزى. في المقابل تسمّح أساليب اكتشاف الشذوذ بتحديد تلك الإشارات الشاذة الخارجة عن نطاق البيانات الطبيعية.

تهدف عملية الكشف عن الحالات الشاذة [4] إلى تحديد كل نقاط البيانات التي تسلك سلوكاً مختلفاً عن باقي نقاط المجموعة. يُمكن أن تنتج الحالات الشاذة عن خطأ في البيانات؛ ولكنها تدل أيضاً على عمليات أساسية جديدة لم تكن معروفة مسبقاً وغالباً ما تكون حرجة في مجموعة واسعة من التطبيقات. يوجد إسقاط كبير لمفهوم اكتشاف الشذوذ في تطبيقات العالم الحقيقي كتطبيقات اكتشاف الاحتيال المالي، الاختراقات الشبكية، الأمراض النادرة، اكتشاف العيوب الصناعية وهو ما يتم التركيز عليه في الدراسة الحالية. توجّهت العديد من شركات القطاع الصناعي مؤخراً لاستخدام أساليب كشف الشذوذ لتطوير خطوط الإنتاج. تعدّ معامل صناعة الورق من أكثر القطاعات توجّهاً نحو ذلك، حيث يُنتج العالم في الوقت الحالي نحو 400 مليون طن من الورق، لتصنيع الاحتياجات الأساسية من النقود، علب الورق المقوى، الإيصالات، ورق الزبدة، الخ. بينما يبلغ استهلاك الفرد في الجمهورية العربية السورية 12 كيلو غرام من الورق في العام الواحد. يتطلّب إنتاج اللب والورق عمليات معقدة ومتكاملة للغاية بوسائل كيميائية أو ميكانيكية

[[5]]. كما تَتَطَلَّبُ مراقبة الأنظمة باستمرار حتى يَتِمَّكَّنَ المشغلون من إدارة العمليات واكتشاف أي حوادث محتملة، قد تَتَسَبَّبُ في انقطاع إنتاجي مفاجئ. لَتَحَقِّقَ ذلك يَتِمُّ نشر مجموعة كبيرة من أجهزة الاستشعار في أجزاء مختلفة من معدات التصنيع، لقياس متغيرات العملية الهامة ومراقبة حالة النظام. تولد هذه المستشعرات كميات كبيرة من بيانات السلاسل الزمنية متعددة المتغيرات (Multivariate Time Series). تحتوي السلسلة على أكثر من متغير تابع (Dependent Variable). لا يَتَعَمَدُ كُلُّ متغير فقط على قيمه السابقة، بل يَتَعَمَدُ أيضاً على بعض المتغيرات الأخرى ضمن السلسلة[[6]]. تُعَبَّرُ السلاسل الناتجة عن قراءات المقاييس على طول خط الإنتاج.

يجب معالجة هذه القراءات الأولية وتحليلها بشكل مناسب، للحصول على معلومات مفيدة بشأن الحالة الطبيعية للنظام، حيث يَتِمُّ بناء أنظمة قادرة على تحليل وتفسير بيانات المراقبة، لاكتشاف أي فشل مُمَكِّن في المستقبل. تُدْعَى هذه الأنظمة أنظمة كشف الشذوذ.

يُمكن تعريف أنظمة كشف الشذوذ على أنها أجهزة أو برامج تقوم بمراقبة البيانات المتعلقة بمجال مُعين لتحديد الحالات الشاذة والمختلفة عن السلوك الطبيعي باستخدام طرائق معينة مثل التعلُّم العميق (Deep Learning).

وعلى النقيض من ذلك، تَتَمَثَّلُ أحد أكبر القضايا التي تواجه الأنظمة الحالية لكشف العيوب (الشذوذ) الصناعية؛ أنه يجب توقع فشل الآلة قبل حدوثه. حيث كانت تُركِّز الأنظمة التقليدية، على اكتشاف الأسباب المؤدية للحالات غير الطبيعية للنظام [[7]] في أقرب وقت ممكن بعد حدوثها. يُساعد الكشف المبكر ولو قبل بضع ثوانٍ على تجنب أي انقطاع مُحتمَل في سير عملية الإنتاج. فعلى سبيل المثال تُسَبَّبُ مشكلة كسر الورق (Paper Breaks) توقف الآلات لأكثر من ساعة عن كُلِّ عملية كسر، وتُقدَّرُ تكلفة الخسارة الناتجة عن التوقف بحوالي 15 مليون دولار عن كُلِّ خط إنتاج [[8]].

يُساعد الاعتماد على السلاسل الزمنية متعددة المتغيرات، في نمذجة الترابطات المثيرة للاهتمام والتنبؤ بحالات الشذوذ المستقبلية. سَاعَدَتِ التطورات السريعة في الآونة الأخيرة للشبكات العصبونية العميقة [[9]] في توفير أداة قوية للتعامل مع هذا النوع من

البيانات، كاخترالها في بُعد أقل. يمكن اعتبار شبكات الذاكرة طويلة المدى LSTMs (Long Short-Term Memory) الحل الأكثر فعالية للتسلسلات الزمنية، حيث تَمَيَّزَ عن الشبكات العصبونية التقليدية بامتلاكها ذاكرة تساعد على تذكر الأنماط لفترات طويلة من الزمن.

لكن على الجانب الأخر، تَتَطَلَّبُ الشبكات العصبونية العميقة قدراً كبيراً من البيانات الشاذة، وهذا غير ممكن في العديد من تطبيقات العالم الحقيقي، كما هو الحال بالنسبة لمشكلة البحث الحالية. حيث تشتمل مجموعة البيانات فقط على 124 نقطة زمنية لانهيار الآلات (كسر الورق) ضِمَّنَ أكثر من 18000 نقطة زمنية، مما يجعل مجموعة البيانات غير متوازنة للغاية. يؤدي ذلك إلى صعوبة إنشاء نماذج ذات كفاءة عالية نظراً لعدم توفر بيانات شاذة كافية لعملية التدريب، وعلى الرغم من ذلك فإن التخفيض ولو بنسبة قليلة من أحداث كسر الورق سيعطي توفيراً كبيراً في تكلفة الإنتاج.

ومن جهة أخرى، تَتَطَلَّبُ أنظمة كشف الشذوذ القائمة على الشبكات العصبونية تحديد العتبة (Threshold) لفصل الحالات الشاذة والطبيعية ضِمَّنَ البيانات، تَعَمَّدُ معظم الدراسات على تحديد قيمة العتبة تجريبياً وبشكل ثابت (Static)؛ ولكن من خلال الاعتماد على فرضية أن متجهات البيانات (Data Vectors) تتبَّع توزيع غاوس متعدد المتغيرات، فيمكن استخدام تابع احتمالي مثل Likelihood لتقدير قيمة العتبة [10]. بالمقابل يبدو في العديد من المواقف العملية أنه من الصعب الاحتفاظ بهذا الافتراض لأن توزيع البيانات (Data Distribution) غالباً ما يكون غير معروف. يمكن تَجَنُّبُ وضع أي افتراضات حول توزيع البيانات من خلال تطبيق طريقة غير معلمية (Nonparametric Method)، وبالتالي يُتيح لهذه الأنظمة تَعْيِيرَ قيمة العتبة بشكل ديناميكي (Dynamic) في حال حدوث اختلاف بحجم وتوزيع البيانات. تُعتبر نظرية تشيبشيف (Chebyshev's Theory) طريقة غير معلمية، فهي تحدد الحد الأدنى لنسبة الملاحظات (Observations) التي تقع ضِمَّنَ عدد محدد من الانحرافات المعيارية عن المتوسط.

إِعْتَمَدَ البحث لحل المشاكل السابقة على استخدام تقنيات التعلُّم العميق وتحديداً شبكة الترميز الآلي ذات الذاكرة طويلة المدى (LSTM-Autoencoder) ونظرية

تشبيبيشيف؛ لبناء نظام ديناميكي للكشف عن الحالات الشاذة قبل حدوثها بوقت مناسب. حيث تمّ تجريبه على اكتشاف حالات فشل آلة الورق قبل حدوثها بأربعة دقائق. بالإضافة الى تَحْدِيدِ قِيمِ البارامترات الفائقة (Hyperparameters) للنظام المُقْتَرَحِ لتحقيق أعلى نسبة كشف لحالات كسر الورق.

2- أهمية وأهداف البحث

تساعد أنظمة كشف الشذوذ ضمن السلاسل الزمنية متعددة المتغيرات، على اكتشاف الحالات الشاذة قبل حدوثها. لكن بالمقابل يَنْطَلِبُ ذلك الاعتماد على تقنيات أكثر تعقيداً مثل بعض خوارزميات التَّعَلُّمِ العميق؛ لقدرتها على استخراج الميزات الهامة من جهة، ومن جهة أخرى لقدرتها على تذكر الأنماط المعقدة والخفية لفترات طويلة.

يَهْدَفُ البحث بشكل أساسي إلى بناء وتطوير نظام ديناميكي للكشف المبكر عن

الحالات الشاذة ضمن السلاسل الزمنية، وذلك من خلال تحقيق ما يلي:

- 1- توقع الحالات الشاذة ضمن مجموعة البيانات قبل حدوثها.
- 2- النقاط التبعيات الزمنية لتسلسل الشذوذ ضمن سلاسل البيانات.
- 3- إيجاد أفضل قيم البارامترات الفائقة للنظام المُقْتَرَحِ.
- 4- إيجاد عتبة التصنيف الأمثل ديناميكياً، لتحقيق أفضل دقة كشف ممكنة.

3- مواد وطرائق البحث

3-1- البيانات البحثية

تُمَثِّلُ البيانات المستخدمة مجموعة بيانات حقيقية تتعلق بمشكلة كسر الورق. قُدِّمَتْ مجموعة بيانات البحث ضمن مسابقة البيانات المقامة في معهد أنظمة المهندسين الصناعيين (Institute of Industrial and Systems Engineers) لعام 2019. تمّ جَمْعُ المشاهدات من أحد معامل صناعة اللب والورق [11] على مدار شهر واحد باستخدام مجموعة متنوعة من أجهزة الاستشعار. نُقِيَِسُ هذه المستشعرات المواد الخام

(مثل كمية الألياف الليبية والمواد الكيميائية وغيرها)، ومتغيرات العملية (مثل نوع الشفرة وسرعة الدوران). تتضمّن مجموعة البيانات قراءات 61 مستشعر عند 18398 نقطة زمنية ضمّن فترات منتظمة (كل دقيقتين)، مع تحديد حالة النظام (طبيعي أو شاذ) عند كل نقطة. على الرغم من هذا العدد الكبير من القياسات، إلا أن حالات الفشل تحدث فقط عند 124 نقطة زمنية (0.67% من إجمالي الملاحظات) أثناء التشغيل مما يجعل من الصعب التنبؤ بالفشل قبل حدوثه؛ لكن أي تقليل لحالات الفشل بواسطة الاكتشاف المبكر، يوفر قدراً كبيراً من تكاليف الانتاج. يُلخص الجدول 1 مجموعة بيانات البحث.

الجدول 1 وصف مجموعة البيانات

Element	No.	Remark
Continuous) المتغيرات المستمرة (variables)	59	X1 ~X27, X29 ~X60
(Categorical variables) المتغيرات الفئوية	2	X28 (8 categories), X61 (2 categories)
(Normal) الحالات الطبيعية	18274	2-minute intervals
(Abnormal) الحالات الشاذة	124	-

3-2- الحزم والمكتبات المستخدمة Used Packages and Libraries

تُبنى نظام الكشف عن الشذوذ في مجموعة البيانات السابقة باستخدام لغة بايثون Python وبالاعتماد على مجموعة من المكتبات البرمجية [12] وأهمها:

- **Keras**: توفر هذه الحزمة واجهات برمجية متناسقة وبسيطة من أجل التخاطب مع المستخدم النهائي وليس الآلة ، وتحتوي على مجموعة من النماذج (Models) مثل الشبكات العصبونية وأشجار القرار وتوابع التنشيط، كما تتميز بقابليتها للتوسع أي القدرة على إضافة نماذج جديدة. تكون المهمة الأساسية للمكتبة جعل التطبيق أكثر استجابة مع إمكانية إعطاء المستخدم المزيد من القدرة على التحكم به.

● **TensorFlow** : تُطبّق هذه الحزمة في العديد من المجالات كحساب المشتقات والمصفوفات الضخمة بالإضافة إلى استخدامها في توزيع العمليات الحاسوبية على وحدات المعالجة المركزية CPU، وكذلك على شبكة موزعة مكونة من مجموعة أجهزة بعيدة تتضمّن هذه المكتبة. يُستخدم TensorFlow بشكل أساسي في تعلم الآلة في الوقت الحالي.

3-3 - مقاييس الأداء Performance Metrics

يعتبر مقياس الدقة Accuracy المقياس الأكثر استخداماً والأكثر منطقية في تقييم أداء خوارزميات التصنيف وذلك عندما تكون مجموعات البيانات متوازنة. فهو يمثل النسبة بين عدد العينات المصنفة بشكل صحيح وعدد العينات الكلية. ولكن بالمقابل لا يمكن اعتباره مقياساً جيداً في البيانات غير المتوازنة (كما في حالة اكتشاف الشذوذ) لأنه يسبب ملائمة زائدة (Overfitting) لجهة صف الأغلبية.

يوجد مجموعة واسعة من المقاييس التي يمكن استخدامها لتقييم أداء أنظمة الكشف عن الشذوذ [13] ونذكر منها هنا فقط ما تم استخدامه لتقييم الأداء كونها الأكثر استخداماً في تقييم أنظمة اكتشاف الشذوذ:

1. مصفوفة الارتباك Confusion Matrix

يُتنبأ النموذج بصف عينات البيانات وينسب إلى كل عينة تسميتها (Label) المتوقعة (إيجابية أو سلبية)، لتقع كل عينة في نهاية المطاف ضمن أحد الحالات التالية:

(1) الإيجابيات الحقيقية (TP) True Positives تدل على العينات الإيجابية المتوقعة بشكل صحيح.

(2) السلبيات الحقيقية (TN) True Negatives تدل على العينات السلبية المتوقعة بشكل صحيح.

(3) الإيجابيات الخاطئة (FP) False Positives والسلبيات الخاطئة (FN) False Negatives عندما يتعارض الصف الحقيقي مع الصف المتوقع.

يُمكن تلخيص الحالات السابقة ضمن مصفوفة 2×2 تسمى مصفوفة الارتباك

$$.M = (TP \quad FN \quad FP \quad TN)$$

2. الدقة Precision والاستدكار Recall

يُمكن تمثيل العينات الإيجابية والسلبية في مصفوفة الارتباك على إنها عينات ذات صلة أو غير ذات صلة بالبيانات الشاذة على التوالي. بالتالي يُعبر مقياس الاستدكار (Recall) -مقياس الدقة في صف العينات الإيجابية، معدل الإيجابيات الحقيقية (TPR)- عن نسبة العينات ذات الصلة التي تم استرجاعها بشكل صحيح، بينما يُعرف مقياس الدقة (Precision) على أنه نسبة العينات المسترجعة ذات الصلة. يوفر كل من زوج دقة واستدكار رؤية مفيدة في سلوك المصنف، ويؤدي الجمع بينهما إلى إنشاء مقاييس تقييم ذات فعالية وعلى وجه التحديد مقياس F1 والذي يمثل المتوسط التوافقي بين كل من الدقة والاستدكار.

3. مقياس F1 Score

يُعدّ F1 Score في التحليل الإحصائي مقياساً لدقة الاختبار، فهو المتوسط التوافقي للدقة (Precision) والاستدكار (Recall). تحسب قيمة F1 بالاعتماد على مصفوفة الارتباك بالشكل التالي:

$$F1\ score = \frac{2 * TP}{2 * TP + FP + FN} = 2 * \frac{precision * recall}{precision + recall} \quad (1)$$

تمثل القيمة 1 أفضل نتيجة، بينما 0 أسوأها.

4. معدل الإيجابيات الخاطئة False Positive Rate

يُقَيِّس معدل الإيجابيات الخاطئة (FPR) النسبة بين عدد الحالات السلبية التي صنفت على أنها إيجابية (الإيجابيات الخاطئة FP) وإجمالي عدد الحالات السلبية بغض النظر عن تصنيفها.

$$FPR = \frac{FP}{FP + FN} \quad (2)$$

3-4 - نظرية تشيبشيف (Chebyshev's Theory)

إعتمدَ البحث لتحديد قيمة عتبة التصنيف ديناميكياً في نظام كشف الشذوذ المُقترح على اختبارات غير معلمية، تقادياً لوضع أي فرضية حول توزيع البيانات. ركزَ البحث الحالي على استخدام نظرية تشيبشيف كطريقة غير معلمية.

تتص النظرية [14] على أنه يوجد على الأقل نسبة $1 - \frac{1}{k^2}$ من البيانات تقع ضمن k انحراف معياري عن المتوسط $\mu \pm k\sigma$ ، حيث $k > 1$. وذلك من أجل مجموعة واسعة من التوزيعات الاحتمالية.

تطبق النظرية على متجهات الخطأ (e_t) للبيانات الطبيعية، لتحديد المجال الذي تقع ضمنه معظم هذه المتجهات. تمثل متجهة الخطأ الفرق بين السلسلة الزمنية المتوقعة والحقيقة في الزمن t .

$$e_t = ||\hat{x}_t - x_t|| \quad (3)$$

يختار النظام قيمة عتبة التصنيف ضمن المجال المُحدّد، حيث تم ضبط القيمة الافتراضية للعتبة على متوسط المجال، بينما القيمة العليا على الحد الأعلى للمجال.

4 - الدراسة المرجعية

قدّمت أدبيات الدراسات السابقة العديد من الطرائق والخوارزميات لمعالجة مسائل التصنيف ضمن السلاسل الزمنية أحادية المتغير (Univariate Time Series). حيث اعتمدت منهجية هجينة لتحسين دقة مصنّفات السلاسل الزمنية باستخدام نماذج ماركوف المخفية (Hidden Markov Models) [15]. كما تم مقارنة ثلاث منهجيات من خوارزميات أشجار القرار (Decision Tree) الشائعة، وهي الغابات العشوائية (Random Forest)، AdaBoost، Multi Boost [16] في اكتشاف حالات عدم انتظام ضربات القلب ضمن السلاسل البيولوجية الزمنية (Biomedical Time-Series)، وأشارت النتائج إلى تفوق AdaBoost على المصنّفات الأخرى. أُستخدِمت

مؤخراً الشبكات العصبونية الالتفافية (Convolutional Neural Networks) في مسائل تصنيف السلاسل الزمنية أحادية المتغير [17]، بهدف دمج كل من مهام التصنيف واستخراج الميزات (Feature Extraction) في إطار عمل واحد.

لُقَّت السلاسل الزمنية متعددة المتغيرات في الآونة الأخيرة اهتماماً كبيراً لدى العديد من الباحثين. حيث تمَّ معالجة التحديات التي تواجهها [18] ومن أهمها: (1) تدفقات البيانات عالية الأبعاد، (2) الحالات الشاذة الطارئة على عمل النظام، وذلك من خلال بناء نظام مراقبة تدفق بيانات حقيقية. إستطاع النظام المُقترح استخراج الميزات من خلال دراسة الارتباطات بين عناقد (Cluster) البيانات، واكتشاف الحالات الشاذة باستخدام تحليل المكونات الرئيسية (Principal Component Analysis)، لكن بالمقابل لم يتمَّ تحديد عتبة التصنيف بشكل ديناميكي، فضلاً عن عدم استخدام خوارزميات التعلُّم العميق.

حقَّقت خوارزميات التعلُّم العميق نتائج بارزة في مجال تصنيف السلاسل الزمنية متعددة المتغيرات. تَفَوَّقت الشبكات العصبونية العميقة على الأساليب القائمة على الميزات في تعلُّم ميزات السلاسل الزمنية [19]، تمَّ تطوير شبكة عصبونية التفافية متعددة القنوات (Multi-Channels Convolution Networks) لتصنيف السلاسل الزمنية ضمن مجموعتين من بيانات العالم الحقيقي، حيث تتعلَّم الشبكة الميزات من سلاسل زمنية أحادية المتغير في كل قناة، ومن ثم يتمَّ جمع المعلومات من جميع القنوات لتمثيل الميزات المهمة في الطبقة النهائية. أظهرت النتائج التجريبية قدرة التعلُّم العميق على تعلُّم الميزات الأكثر أهمية، مما يُساعد على تحسين أداء التصنيف. بالمقابل لم تستخدم الدراسة تقنيات ضبط البارامترات الفائقة، بالإضافة لتحديد عتبة التصنيف بشكل يدوي.

تمَّ مقارنة أربع نماذج مختلفة للتعلُّم العميق [20] وهي: الشبكات العصبونية الاصطناعية (Artificial Neural Networks)، والشبكات العصبونية التكرارية

(Recurrent Neural Networks)، والوحدة المتكررة ذات البوابات (Gated Recurrent Units)، والذاكرة طويلة المدى (LSTM) في مسائل كشف الاحتيال المالي القائم على التسلسل. أظهرت النتائج تفوق LSTM على النماذج الأخرى، وتوصلت الدراسة إلى تحديد الطوبولوجيا المناسبة لكل نموذج باستخدام ضبط البارامترات الفائقة. بالمقابل لم تتطرق الدراسة إلى موضوع تحديد عتبة التصنيف بشكل ديناميكي.

أستُخدمت الذاكرة طويلة المدى (LSTM) لبناء نموذج كشف الشذوذ ضمن طلبات المستهلك في إداة سلسلة التوريد [21]]. حَقَّقَت الشبكة تفوقاً على خوارزمية آلة شعاع الدعم (Support Vector Machine) في دقة الكشف، كما اغتمدت الدراسة على طريقة توقع الكميات (Quantiles Estimation) لتحديد قيمة العتبة. على الرغم من تفوق النموذج المقترح في كشف الحالات الشاذة، إلا إنه من المُمكِن تطويره من خلال التحكم بتدفق البيانات واختزالها قدر الإمكان، من خلال بناء نموذج هجين بين LSTM وأحد الشبكات العصبونية العميقة مثل Autoencoder و CNN.

تمَّ بناء عدة نماذج هجينة باستخدام شبكات عصبونية عميقة لتصنيف السلاسل الزمنية متعددة المتغيرات، ومن أهمها شبكة الذاكرة طويلة المدى (LSTM) والشبكة الالتفافية (CNN) وشبكة الترميز الآلي (Autoencoder). أستخدمت الشبكة (CNN-LSTM) لبناء نموذج اكتشاف الأنماط غير الطبيعية [22] في حركة بيانات السلاسل الزمنية عبر الويب. حيث تُستخدم CNN لتقليل أبعاد الميزات المكانية (Spatial Features)، بينما يكون الهدف من LSTM نمذجة معلومات الوقت. حَقَّقَ النموذج المقترح أداء غير مسبوق للكشف عن الحالات الشاذة حتى مع البيانات المتشابهة جداً. بالرغم من تفوق هذه الطريقة، إلا أن دقتها تميل للانخفاض عند اكتشاف القيم الشاذة التي يتم إنشاؤها عند حدود نوافذ البيانات، كما هناك تأخير زمني في كشف الحالات الشاذة.

إستطاع النموذج (LSTM-CNN) التفوق على جميع النماذج الأخرى [23] في التعامل مع السلاسل الزمنية. حيث لا يحتاج للكثير من عمليات المعالجة المسبقة للبيانات، وذو كفاءة عالية في وقت الاختبار. تمّ اختبار النموذج على أكثر من 30 مجموعة بيانات متنوعة ومن ضمنها بيانات صناعية. لكن على الجانب الأخر، لم يتمّ ضبط البارامترات الفائقة، بالإضافة لعدم تحديد العتبة للنموذج بشكل ديناميكي.

تمّ الاعتماد على خصائص كل من شبكتي LSTM و Autoencoder لبناء نموذج هجين لكشف الاحتيال القائم على التسلسل ضمن مجموعة بيانات حقيقية وصناعية [24]. أظهرت النتائج أن لدمج الشبكتين السابقتين تأثير كبير في زيادة دقة الكشف مقارنة مع استخدام كل شبكة على حدة. حيث يتميز النموذج (LSTM-Autoencoder) بقدرته على التعامل مع البيانات عالية الأبعاد من جهة، ومن جهة تذكر الأنماط المعقدة لفترات أطول. بالمقابل لم تعتمد الدراسة على أي منهجية لتحديد العتبة بشكل ديناميكي.

تمّ ملاحظة التفوق الكبير من خلال الدراسات السابقة للشبكات العصبونية العميقة في مجال اكتشاف الشذوذ ضمن السلاسل الزمنية متعددة المتغيرات على باقي الخوارزميات، كما أكدت الدراسات على ضرورة التوجه لبناء نماذج هجينة بين هذه الشبكات لتحقيق أعلى دقة كشف. لكن وبنفس الوقت تمّ ملاحظة القصر الحاصل فيما يتعلق ببناء هذه النماذج ضمن اتجاهين هُما: (1) ضبط البارامترات الفائقة، (2) تحديد عتبة التصنيف بشكل ديناميكي. استوجب كل ذلك المزيد من البحث لبناء نظام ديناميكي للكشف عن الحالات الشاذة قبل حدوثها بوقت مناسب. يكون قادر على اختزال الميزات عالية الأبعاد، والتقاط التبعيات الزمنية لتسلسل الشذوذ ضمن سلاسل البيانات. بالإضافة لتحديد عتبة التصنيف الأمثل بشكل ديناميكي. لتحقيق ذلك كان لابد من الاعتماد على النماذج الهجينة بين شبكات التعلم العميق، وإيجاد أفضل قيم البارامترات الفائقة لها.

5- الخوارزميات المقترحة للدراسة

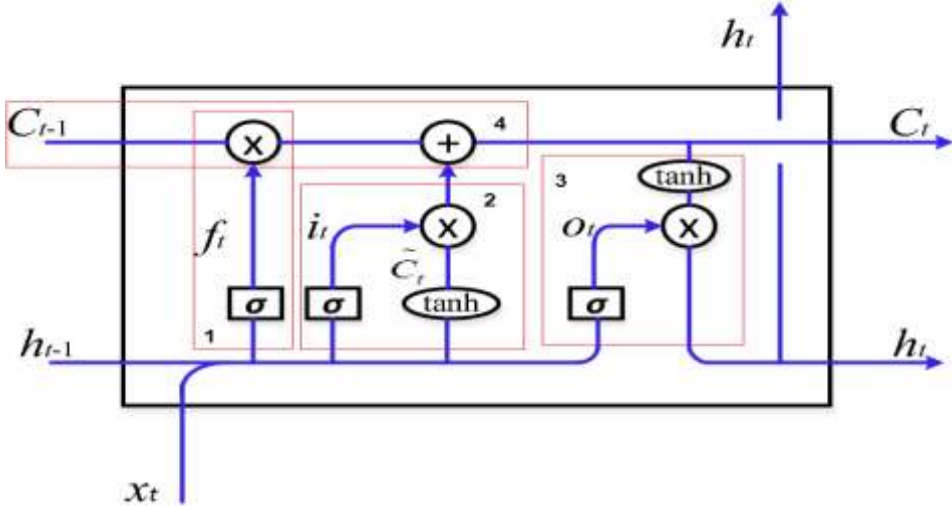
بناء على ما تمّ ملاحظته من الدراسات المرجعية السابقة، قد تمّ التركيز في هذا البحث على دراسة بعض خوارزميات التعلّم العميق التي أعطت أفضل أداء في اكتشاف الشذوذ.

5-1- شبكة الذاكرة طويلة المدى Long Short-Term Memory

تعدّ شبكة LSTM أحد أنواع الشبكات العصبونية المتكررة، والتي تسمح بالاحتفاظ بالتبعية طويلة المدى بين البيانات عند نقطة زمنية معينة، انطلاقاً من عدة خطوات زمنية سابقة. تحتوي LSTM على ما يسمى حالة الخلية (Cell State) وهي تمثل ذاكرة الشبكة، ويتمّ في كلّ خطوة الاحتفاظ بالمعلومات الضرورية والتخلص من المعلومات غير ذات الصلة [25] عن طريق بوابة النسيان (Forget Gate)، وبوابة الإدخال (Input Gate)، وبوابة الإخراج (Output Gate).

5-1-1- بنية وآلية عمل شبكة الذاكرة طويلة المدى

يوضح الشكل 1 بنية شبكة LSTM [26] والتي تتكون من ثلاث بوابات.



الشكل 1 بنية شبكة LSTM

1. بوابة النسيان Forget Gate

يُمثل المربع الأول (1) بوابة النسيان: يَتَمُّ في هذه البوابة الإجابة عن السؤال التالي: ماهي حالة الخلية السابقة التي يجب الاحتفاظ بها أو نسيانها؟

$$f_t = \sigma(w_f [h_{t-1}, x_t]) + b_f \quad (4)$$

حيث، f_t : مُؤَيَّر (Tensor) يستخدم كمرشح (Filter) للمعلومات الموجودة في حالة الخلية، σ : تابع Sigmoid، x_t : حالة الدخل الجديد في اللحظة t ، h_{t-1} : الخرج السابق لحالة الخلية في الزمن $t - 1$.

يكون خرج البوابة إما 0 أو 1، حيث تُقابل القيمة 0 نسيان الحالة الداخلية (Internal State) بشكل كامل، بينما تُقابل القيمة 1 تمرير الحالة بدون تغيير.

2. بوابة التحديث (الادخال) Update Gate

يُمثل المربع الثاني (2) بوابة الادخال: يَتَمُّ في هذه البوابة تحديد القيم المرشحة (Candidate) للدخل الحالي للاحتفاظ بها في حالة الخلية الجديدة.

$$c_t = \tanh(w_c [h_{t-1}, x_t]) + b_c \quad (5)$$

$$i_t = \sigma(w_i [h_{t-1}, x_t]) + b_i \quad (6)$$

يُستخدم التابع Tanh من أجل عملية تنظيم (Regulates) النموذج. حيث يُنتج عن التابع متجهة تحتوي على كل القيم (الحالات) التي يمكن إضافتها إلى حالة الخلية الجديدة، بينما يكون الهدف من استخدام تابع Sigmoid تحديد ماهي المعلومات الحالية التي يتم الاحتفاظ بها أو نسيانها.

3. بوابة الخرج Output Gate

يَتَمُّ في المربع الرابع (4) تحديث حالة الخلية السابقة c_{t-1} إلى حالة الخلية الجديدة c_t ، من خلال ضرب الخلية السابقة بخرج بوابة النسيان، ومن ثم إضافة القيم المرشحة الجديد الناتجة عن بوابة الإدخال.

$$c_t = f_t * c_{t-1} + i_t * c_t \quad (7)$$

حيث: تمثل $c_t * i_t$ قيم مرشحة (خرج بوابة الإدخال)، f_t خرج بوابة النسيان

يُمثل المربع الثالث (3) بوابة الإخراج: يَتَمُّ في هذه البوابة الإجابة عن السؤال

التالي: ماذا سوف يكون خرج حالة الخلية الجديدة h_t ؟

$$O_t = \sigma(w_o [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = O_t * \tanh(c_t) \quad (9)$$

5-2-2 - شبكة الترميز الآلي Autoencoder Network

شبكة عصبونية متعدّدة الطبقات، غير خاضعة للإشراف (Unsupervised)، ذات تغذية متقدمة (Feed-Forward). تتكون من طبقة إدخال، وطبقة إخراج، ومُرَمِّز (Encoder)، ومفكك ترميز (Decoder)، وفضاء كامن (Latent Space). يَتَمُّ تغذية دخل الشبكة بشعاع مكوّن من ميزات (Features) مجموعة بيانات الدخل، ومن ثمّ تَعْمَلُ الشبكة على إعادة إنتاج هذا الشعاع في طبقة الخرج. يَتَمُّ تدريب الشبكة بالاعتماد خوارزميات الانتشار الخلفي (Backpropagation)، وتابع خسارة (Loss Function) لحساب الخسارة الناتجة عن إعادة بناء شعاع الدخل في طبقة الخرج **Error!** [27]. [Reference source not found].

5-2-1 - بنية وآلية عمل شبكة الترميز الآلي

نُفرض بدايةً عند تصميم بنية شبكة Autoencoder وجود مشكلة عنق الزجاجة (Bottleneck) في المنتصف مما يفرض تمثيل مضغوط للمدخلات الأصلية؛ لذلك تُعتبر الشبكة مناسبة جداً في مجموعات البيانات ذات الأبعاد الكبيرة حيث تَعْمَلُ على اختزال البيانات في بُعد أصغر مع الاحتفاظ بالمعلومات الرئيسية لهيكلية البيانات.

كما ذكرنا سابقاً أن الشبكة تتكون من ثلاث مكونات رئيسية **Error!** **Reference source not found.** هي: المُرَمِّز (Encoder)، الترميز (Code)، مفكك الترميز (Decoder). كما تحتاج إلى ثلاث أشياء: طريقة الترميز، طريقة فك الترميز، وتابع الخسارة.

يُعتبر المُرَمِّز ومفكك الترميز شبكة عصبونية ذات تغذية متقدمة، ومتصلة بالكامل (Fully-Connected) فيما بينها. يقوم المُرَمِّز بضغط البيانات الأصلية إلى بُعد أصغر، بينما يَعْمَلُ مفكك الترميز بفك ترميز البيانات المضغوطة وإعادة بناءها إلى بعدها الأصلي. يُمثل الترميز طبقة واحدة في الشبكة بحجم اختياري يَتَمُّ تحديده قبل عملية

التدريب (يسمى أيضاً عنق الزجاجاة). حيث يتحكم حجم هذه الطبقة ببعدها المدخلات المضغوطة التي يتم تغذيتها إلى وحدة فك الترميز. تُمرر بيانات الإدخال عبر المُرمِّز لضغط البيانات كمرحلة أولى، ومن ثم إلى الترميز لضغطها إلى بُعد أصغر، بينما يقوم مفكك الترميز الذي له بنية مماثلة للمُرمِّز بإعادة إنتاج الدخل في طبقة الخرج بالاعتماد على البيانات المضغوطة ضمن طبقة الترميز. يجب التنويه أن يكون لطبقتي الدخل والخرج نفس عدد العصبونات، بينما يمكن تعديل أحجام الطبقات في الوسط كما نريد.

5-2-2- اكتشاف الشذوذ باستخدام شبكة الترميز الآلي

يتم الكشف عن الشذوذ باستخدام شبكة الترميز الآلي، من خلال تعلم الأنماط والعلاقات المعقدة ضمن البيانات الطبيعية. تُصَبَّح الشبكة بعد عملية التدريب قادرة على إعادة بناء البيانات الطبيعية بشكل جيد للغاية، بينما تقشل في القيام بذلك مع البيانات الشاذة. حيث تُعتبر نقاط البيانات التي لها خطأ إعادة بناء (Reconstruction Error) أعلى من حد معين نقاط شاذة.

5-3- شبكة الترميز الآلي ذات الذاكرة طويلة المدى LSTM Autoencoder

يُمكن تنظيم شبكة LSTM Autoencoder بالاعتماد على بنية مشابهة لشبكة Autoencoder، تُسمى LSTM Encoder-Decoder [28] والتي يكون فيها كل من المُرمِّز ومفكك الترميز هو شبكة LSTM. يكون الهدف من استخدام LSTM ضمن المُرمِّز ومفكك الترميز هو التقاط التبعية الزمنية (Temporal Dependencies) ضمن سلاسل البيانات.

تُساعد شبكة LSTM Autoencoder على تقليل أبعاد البيانات واختزالها في بُعد أصغر مع الاحتفاظ بالمعلومات الرئيسية لهيكلية البيانات باستخدام خصائص شبكة الترميز الآلي، وتعلم الأنماط المعقدة داخل الترتيب الزمني لتسلسل الإدخال باستخدام شبكة الذاكرة طويلة المدى. كما تساعد في المحافظة على التبعية بين مكونات متجهة الخرج المتوقع.

5-3-1- اكتشاف الشذوذ باستخدام شبكة LSTM AUTOENCODER

تُشاهد الشبكة التسلسلات الطبيعية فقط أثناء عملية التدريب وتتعلم إعادة بنائها، مما يجعلها قادرة على اكتشاف الشذوذ في السلاسل الزمنية متعددة المتغيرات. فعندما يتم تغذيتها بتسلسل غير طبيعي (شاذ) لا تكون قادرة على إعادة بنائها بشكل جيد. أن تدريب الشبكة بهذه الطريقة له معنى عملي لأن البيانات الشاذة ليست متاحة دائماً، كما أنه من المستحيل تغطية جميع انماط هذه البيانات.

6- التجارب والنتائج Experiments and Results

يُستعرض في هذا القسم بدايةً وصفاً تجريبياً لتصميم التجارب التي أجراها البحث لبناء أنظمة الكشف المبكر عن الشذوذ (اكتشاف كسر الورق) باستخدام خوارزميات التعلم العميق المقترحة، وبالاعتماد على أساليب ضبط البارامترات الفائقة لتحديد أفضل قيم البارامترات، ونظرية تشبيشيف لتحديد قيمة عتبة التصنيف بشكل ديناميكي. يتبع التوصيف استعراض النتائج وتقييم الأداء.

6-1- بناء النموذج Model Creation

تتضمن مراحل بناء النموذج مجموعة من الإجراءات وهي كما يلي:

1. تحضير البيانات Data Preprocessing

يتم معالجة وتحضير البيانات قبل عملية تطبيق النموذج المقترح. يجب بدايةً تقسيم البيانات بأكملها إلى مجموعات تدريب واختبار (Training and Test)، ونظراً لأن الشبكات المقترحة ضمن البحث، تحتاج للبيانات الطبيعية فقط أثناء عملية التدريب، لذا فأولاً يتم وضع جميع العينات الشاذة ضمن بيانات الاختبار، وبعد ذلك تقسم البيانات الطبيعية إلى 80% للتدريب و20% للاختبار (مجموعة الاختبار: 20% من البيانات طبيعية + كل البيانات الشاذة).

وبعد عملية التَّقْسِيم، يُتَمَّ تَقْيِيس (Standardization) بيانات التدريب، ثم تَقْيِيس بيانات الاختبار باستخدام المتوسط والانحراف المعياري لبيانات التدريب بعد تَقْيِيسها، مما يُبْقِي بيانات الاختبار غير معروفة أثناء النمذجة.

يُهْدَفُ البحث إلى اكتشاف الشذوذ قبل حدوثه، تَتَمَثَّلُ أحد الأساليب لتحقيق ذلك بإزاحة تصنيف النقطة الزمنية المقابلة لحالة الشذوذ إلى للأعلى بمقدار s ، بحيث تكون المصنفات قادرة على التنبؤ بحالة النظام الحالية (c_t) قبل s وحدة زمنية.

$$c_t \leftarrow c_{t+s}, \text{ where } s = 1, 2, \dots \quad (10)$$

تُتَمَّ تَعْيِين $s = 2$ ضِمَّنَ البحث الحالي، مما يعني أن النموذج قادر على توقع الشذوذ قبل حدوثه بمقدار 4 دقائق (يفصل بين كل قراءة دقيقتين). تُطَبَّقُ عملية الإزاحة على كُلِّ من بيانات التدريب والاختبار.

2. النافذة الزمنية Time Window

يحتاج النموذج حتى يَتَمَكَّنَ من النقاط التبعيات الزمنية خلال فترة معينة، أن يَعْتَمِدَ على جميع النقاط الزمنية ضِمَّنَ تلك الفترة. لذا يَتَمَّ الاحتفاظ بمجموعة من النقاط الزمنية المتتالية باستخدام مفهوم النوافذ الزمنية (Time Windows). تَضِمَّنَ النافذة بحجم m ، النقاط الزمنية المقابلة لأخر m قراءة وصولاً للوقت الحالي t ، يَتَمَّ تصنيف مثل النافذة (Window Instance)، بالاعتماد على تصنيف آخر نقطة زمنية في النافذة (c_t).

يُسْتَخْدَمُ مثل النافذة في تدريب الشبكات العصبونية، بالإضافة لمعالجة مشكلة البيانات غير المتوازنة (Imbalanced Data)، من خلال إنشاء مثل النافذة عند جميع النقاط الزمنية المقابلة لحالة الشذوذ ($c_t = 1$)، وبعض النقاط الزمنية المقابلة للحالة الطبيعية ($c_t = 0$)، مما يُقَلِّصُ الفرق بين صفوف البيانات.

تُتَمَّ تَطْبِيقُ حجوم مختلفة للنافذة على بيانات التدريب والاختبار ضِمَّنَ البحث وهي $m = 3, 5, 7, 9$ ، وبالتالي فإن الشبكة تكون قادرة على التقاط التبعيات الزمنية ضِمَّنَ 6، 10، 14، 18 دقيقة على الترتيب.

3. ضبط البارامترات الفائقة Tuning Hyperparameters

تَعْنِي عملية ضبط (أو تَحْسِين) البارامترات الفائقة إيجاد مجموعة قِيم البارامترات التي تُحَقِّق أفضل أداء للنموذج. رَكَزَ هذا البحث على استخدام طريقة البحث العشوائي لضبط البارامترات الفائقة، حيث أن البحث العشوائي (Random Search) [29] Error! Reference source not found. أكثر كفاءة في تَحْسِين البارامترات من البحث الشبكي (Grid Search) من الناحية التجريبية والنظرية.

يَعْمَل البحث العشوائي بشكل أفضل عندما يكون فضاء البحث عالي الأبعاد، ويحتوي على عدد كبير من التركيبات المختلفة للبارامترات الفائقة. حيث تَبَحَّث تقنية البحث العشوائي ضِمْنَ مجموعة عشوائية من هذه التركيبات لاختيار أفضل القيم للبارامترات الفائقة الخاصة بالنموذج المُقْتَرَح، وبالتالي فإن الوقت المُستَعْرَق للعثور على المجموعة الصحيحة يكون أقل مع عدد أقل من التكرارات. فيما يلي وصفاً مختصراً لجميع البارامترات الفائقة التي أَعْتَمَدَهَا البحث من أجل تطوير عمل النموذج المُقْتَرَح.

• تابع التنشيط (Activation Function)

يُحَدَد آلية تحويل الأوزان الموزونة (Weights Weighted) للمدخلات إلى قيمة محددة تمثل خرج العُقد. يوجد مجموعة واسعة من توابع التنشيط أهمها:
(1) Logistic (Sigmoid): يحدد قِيمَةَ الخرج ضمن المجال $[0, 1]$ ، والصيغة الرياضية للتابع هي بالشكل التالي:

$$1.0/(1.0 + e^{-x}) \quad (11)$$

(2) Hyperbolic Tangent (Tanh): يحدد قِيمَةَ الخرج ضمن المجال $[-1, 1]$ ، والصيغة الرياضية للتابع هي بالشكل التالي:

$$(e^x - e^{-x})/(e^x + e^{-x}) \quad (12)$$

(3) Rectified Linear Activation (ReLU): يحافظ على قِيمَ الإدخال الموجبة، بينما يكون الخرج مساوياً للصفر من أجل قِيمَ الإدخال السالبة. الصيغة الرياضية للتابع هي بالشكل التالي:

$$\max(0.0, x) \quad (13)$$

• **معدل التعلّم (Learning Rate)**

يُشير إلى مقدار خطوة الانتشار الخلفي عند تحديث الأوزان. يتم ضبطه بقيمة صغيرة جداً.

• **حجم الدفعة (Batch Size)**

يُمثل عدد العينات التي يتم تغذية الشبكة بها في كل مرة. أن الحجم النموذجي هو 32، ويكون الحجم من مضاعفات العدد 2.

• **معدل الحذف (Dropout Rate)**

يُفيد في إزالة بعض العقد ضمن الطبقة والتي تكون زائدة وغير مجدية.

• **مرحلة (Epoch)**

يُمثل مقدار الوقت الذي تحتاجه الشبكة من أجل أن تتدرب على مجموعة البيانات بأكملها.

• **العقد (Nodes)**

تُمثل حجم العصبونات في طبقات الشبكة. يتم ضبط الحجم ضمن النموذج المُقترح إلى النصف في كل طبقة تالية في المُرّمز، ومن ثم مضاعفته رجوعاً في مفكك الترميز. فعلى سبيل المثال، من أجل حجم 32 وشبكة مكونة من ثلاث طبقات ضمن المُرّمز ومفكك الترميز، تكون العقد بالشكل التالي [32,16,8,8,16,32]

4. **بناء النموذج**

كما تمّ ذكره يتم بناء نموذج الكشف المبكر عن الحالات الشاذة، باستخدام شبكة الترميز الآلي ذات الذاكرة طويلة المدى (LSTM–Autoencoder). من ثمّ مقارنته مع شبكة الترميز الآلي التقليدية (Autoencoder).

تمّت تسمية النماذج المُقترحة للكشف الديناميكي المبكر عن الشذوذ، كما هو موضح في الجدول التالي.

الجدول 2 أسماء النماذج المُقترحة ضمن البحث

Name	Description
LSTM–AEDT	LSTM–Autoencoder with a Dynamic Threshold

5. تحديد العتبة Threshold Selection

تُستخدم عتبة التصنيف (T) للفصل بين الحالات الشاذة والطبيعية. إن اختيار قيمة عتبة مناسبة دوراً مهماً في عمَل أنظمة الكشف، يتم تحديد العتبة ديناميكياً ضمن النظام المقترح. تُصنف حالة النظام الحالية c_t ذو متجهة الخطأ e_t بالاعتماد على قيمة العتبة المحددة بالشكل التالي:

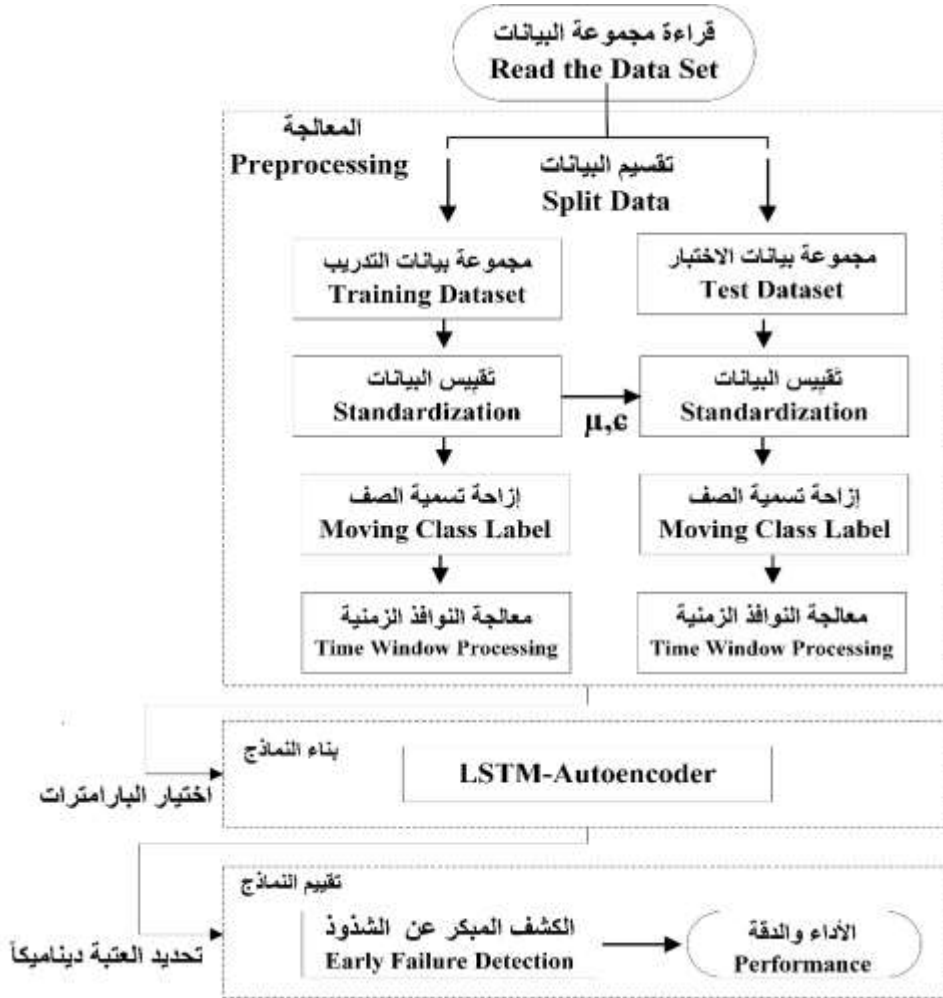
$$c_t \rightarrow \text{normal if } e_t \leq T$$

$$c_t \rightarrow \text{anomaly if } e_t > T$$

6. اختبار النموذج

المرحلة الأخيرة وسيتم مناقشتها في القسم التالي.

يُلخّص المخطط التالي مراحل العمل لبناء نظام كشف شذوذ مبكر (كشف فواصل الورق كمثال تطبيقي على النموذج).



الشكل 2 النظام المقترح لاكتشاف الشذوذ

2-6-2- تقييم الأداء Performance Evaluation

2-6-1- نموذج LSTM-AEDT

رَكَزَ البحث على دراسة أداء النموذج المقترح باستخدام حجوم مختلفة للنافذة الزمنية، وذلك لمعرفة مدى تأثيرها على كُـلِّ من نسبة الكشف وزمن التنفيذ. حيث يتم بدايةً إيجاد قيم البارامترات الفائقة اللازمة لبناء النموذج، ومن ثم مناقشة النتائج وتفاضل العتبة بالاعتماد على المقاييس المقترحة.

1 ضبط البارامترات الفائقة

تُظهر النتائج المُبيّنة في الجدول 3 أفضل تكوينات البارامترات الفائقة لنموذج LSTM-AEDT، باستخدام ضبط البارامترات الفائقة، ومن أجل حجوم النافذة الزمنية المختلفة.

الجدول 3 البارامترات الفائقة لنموذج LSTM-AEDT

Hyperparameters Configurations						Window Size
Neurons	#Epochs	#Batch	Learn Rate	Dropout	Activation	
32	200	256	0.001	0.0	Tanh	3
32	50	128	0.01	0.0	Relu	5
32	50	32	0.01	0.0	Tanh	7
32	50	32	0.001	0.2	Tanh	9

2 تفاضل العتبة (Threshold Tradeoff)

يتم التفاضل بين قيم العتبة حسب طبيعة المسألة، فإذا كان الهدف من نظام الكشف إسترجاع أكبر قدر ممكن من الحالات الشاذة، دون التركيز على دقة التصنيف، يتم اختيار العتبة التي تُحقق أفضل قيمة لمقياس Recall. أما عندما يكون دقة التصنيف عاملاً مهماً، فإنه يتم اختيار العتبة التي تُحقق أفضل قيمة لمقياس Precision. يمكن جمع الحالتين السابقتين (إسترجاع أكبر قدر من الحالات الشاذة، وأفضل دقة تصنيف) من خلال اختيار العتبة التي تُحقق أفضل قيمة لمقياس F1Score.

بالرُجوع إلى نظام كشف كسر الورق المُقترح ضمنَ البحث، يكون الهدف غالباً هو إسترجاع أكبر قدر ممكن من حالات الكسر، تقادياً لتوقف الآلات لساعات طويلة، حيث يبقى زمن فحص الحالة والتأكد من صحتها، أقل تكلفة مادية وزمنية من حدوث الكسر، لكن بالمقابل يجب ألا تتجاوز نسبة الإيجابيات الخاطئة (FPR) حدّ معين موصى به.

• من أجل حجم نافذة $m = 3$

تُظهر النتائج المُبيّنة في الجدول 4 أداء نموذج LSTM-AEDT في اكتشاف الشذوذ (كسر الورق). نلاحظ من خلال الجدول أن النموذج يقوم بتحديد عتبة التصنيف بشكل ديناميكي ضمنَ المجال $[0.2, 1.7]$ ، حيث أن القيمة الافتراضية للعتبة و1.7 هي القيمة العظمى.

الجدول 4 نتائج نموذج LSTM-AEDT من أجل $m = 3$

Threshold	F1Score	Recall	Precision	FPR	Run Time
Default:0.2	0.54	0.59	0.54	0.74	2.5 min
0.5	0.56	0.54	0.58	0.27	
0.8	0.53	0.52	0.56	0.18	
1.1	0.54	0.52	0.60	0.13	
1.4	0.54	0.52	0.59	0.11	
Max:1.7	0.53	0.51	0.60	0.11	

تصل أفضل قيمة لمقياس Recall إلى 59% عند العتبة ($T = 0.2$)، وقيمة Precision إلى 60% عند العتبة ($T = 1.1, 1.7$)، أما بالنسبة لمقياس F1Score فإن أفضل قيمة تصل إلى 56% عند العتبة ($T = 0.5$)، بينما كانت أقل نسبة للإيجابيات الخاطئة FPR هي 11% عند العتبة ($T = 1.7$).

استطاع نموذج LSTM-AEDT المُقترح من أجل حجم نافذة زمنية $m = 3$ ، كشف 51% من حالات الكسر قبل 4 دقائق من حدوثها ($\text{Recall} = 0.51$) وبنسبة إنذارات خاطئة 11% ($\text{FPR} = 0.11$) عند العتبة ($T = 1.7$)، علماً أن زمن التنفيذ 2.5 دقيقة.

• من أجل حجم نافذة $m = 5$

تُظهِر النتائج المُبيّنة في الجدول 5 أداء نموذج LSTM-AEDT من أجل حجم نافذة $m = 5$. حيث نلاحظ أن النموذج يقوم بتحديد عتبة التصنيف بشكل ديناميكي ضمن المجال $[0.2, 1.4]$ ، علماً أن القيمة الافتراضية للعتبة و1.4 هي القيمة العظمى.

الجدول 5 نتائج نموذج LSTM-AEDT من أجل $m = 5$

Threshold	F1Score	Recall	Precision	FPR	Run Time
Default:0.2	0.53	0.61	0.54	0.84	
0.5	0.56	0.55	0.57	0.41	
0.8	0.56	0.53	0.58	0.22	3.2 min
1.1	0.55	0.52	0.60	0.15	
Max:1.4	0.54	0.52	0.60	0.10	

تصل أفضل قيمة لمقياس Recall إلى 61% عند العتبة ($T = 0.2$)، وقيمة Precision إلى 60% عند العتبة ($T = 1.1, 1.4$)، أما بالنسبة لمقياس F1Score فإن أفضل قيمة تصل إلى 56% عند العتبة ($T = 0.5, 0.8$)، بينما كانت أقل نسبة للإيجابيات الخاطئة FPR هي 10% عند العتبة ($T = 1.4$).

حقّق نموذج LSTM-AEDT المُقترح من أجل حجم نافذة زمنية $m = 5$ كشف 52% من حالات الكسر قبل 4 دقائق من حدوثها ($\text{Recall} = 0.52$) وبنسبة إنذارات خاطئة 10% ($\text{FPR} = 0.1$) عند العتبة ($T = 1.4$)، بالمقابل ارتفع زمن تنفيذ النموذج إلى 3.2 دقيقة.

• من أجل حجم نافذة $m = 7$

تُظهِر النتائج المُبيّنة في الجدول 6 أداء نموذج LSTM-AEDT من أجل حجم نافذة $m = 7$. حيث نلاحظ أن النموذج يقوم بتحديد عتبة التصنيف بشكل ديناميكي ضمن المجال $[0.2, 1.4]$ ، علماً أن القيمة الافتراضية للعتبة و1.4 هي القيمة العظمى.

الجدول 6 نتائج نموذج LSTM-AEDT من أجل $m = 7$

Threshold	F1Score	Recall	Precision	FPR	Run Time
Default:0.2	0.54	0.62	0.54	0.82	
0.5	0.55	0.54	0.55	0.41	
0.8	0.55	0.53	0.57	0.28	5.3 min
1.1	0.56	0.53	0.59	0.21	
Max:1.4	0.54	0.53	0.58	0.09	

تصل أفضل قيمة لمقياس Recall إلى 62% عند العتبة $(T = 0.2)$ ، وقيمة Precision إلى 59% عند العتبة $(T = 1.1)$ ، أما بالنسبة لمقياس F1Score فإن أفضل قيمة تصل إلى 56% عند العتبة $(T = 1.1)$ ، بينما كانت أقل نسبة للإيجابيات الخاطئة FPR هي 9% عند العتبة $(T = 1.4)$.

حقّق نموذج LSTM-AEDT من أجل حجم نافذة زمنية $m = 7$ ، كشف 53% من حالات الكسر قبل 4 دقائق من حدوثها ($\text{Recall} = 0.53$) وبنسبة إنذارات خاطئة 9% ($\text{FPR} = 0.09$) عند العتبة $(T = 1.4)$ ، بينما ارتفع زمن تنفيذ النموذج إلى 5.3 دقيقة.

• من أجل حجم نافذة $m = 9$

تُظهر النتائج المُبيّنة في الجدول 7 أداء نموذج LSTM-AEDT من أجل حجم نافذة $m = 9$. حيث نلاحظ أن النموذج يقوم بتحديد عتبة التصنيف بشكل ديناميكي ضمن المجال $[0.2, 2.0]$ ، علماً أن القيمة الافتراضية للعتبة و 2.0 هي القيمة العظمى.

الجدول 7 نتائج نموذج LSTM-AEDT من أجل $m = 9$

Threshold	F1Score	Recall	Precision	FPR	Run Time
Default:0.2	0.53	0.6	0.54	0.81	
0.5	0.57	0.55	0.57	0.4	
0.8	0.56	0.53	0.59	0.24	
1.1	0.57	0.53	0.6	0.19	9.7 min
1.4	0.56	0.53	0.61	0.16	
1.7	0.55	0.52	0.6	0.13	
Max:2.0	0.52	0.50	0.57	0.12	

تصل أفضل قيمة لقياس Recall إلى 60% عند العتبة $(T = 0.2)$ ، وقيمة Precision إلى 61% عند العتبة $(T = 1.4)$ ، أما بالنسبة لقياس F1Score فإن أفضل قيمة تصل إلى 57% عند العتبة $(T = 0.5, 1.1)$ ، بينما كانت أقل نسبة للإيجابيات الخاطئة FPR هي 12% عند العتبة $(T = 2.0)$.

حقّق نموذج LSTM-AEDT من أجل حجم نافذة زمنية $m = 9$ كشف 50% من حالات الكسر قبل 4 دقائق من حدوثها ($\text{Recall} = 0.5$) وبنسبة إنذارات خاطئة 12% ($\text{FPR} = 0.12$) عند العتبة $(T = 2.0)$ ، بينما ارتفع زمن تنفيذ النموذج إلى 9.7 دقيقة.

أكدت التجارب أن لحجم النافذة دوراً مهماً في تحسين أداء النموذج، لكن من جهة أخرى، يجب المحافظة على حدّ معين لحجم النافذة، حيث أن النموذج يسلك سلوكاً

عكسياً بعد حجم مُعيّن. حَقَّقَ LSTM-AEDT أفضل أداء له من أجل حجم نافذة $m = 7$ ، من حيث نسبة الكشف (Recall = 53%)، والإنذارات الخاطئة (FPR = 9%). بينما سلك النموذج بعد حجم نافذة $m = 7$ سلوكاً معاكساً، فمن أجل حجم نافذة $m = 9$ ، إنخَفَصَت نسبة الكشف إلى 50%، وازْتَمَعَت الإنذارات الخاطئة إلى 12%.

(3) تحليل التكاليف والفوائد (Cost Benefit Analysis)

تمَّ تَحْلِيل مساهمة النموذج المُقْتَرَح في الحَدِّ من الخسارة المالية ضِمَّنَ الصناعة (مشكلة كسر الورق). حيث تمَّ إحتِسَاب الربح على أساس الاستنكار، والخسارة الناتجة على أساس الإيجابيات الخاطئة، علماً أنه قُدرت تكلفة كُلِّ عملية كسر بحوالي 10,000 دولار أمريكي، بينما تكلفة كل إنذار خاطئ بحوالي 100 دولار أمريكي فقط [11]. يُبيِّن الجدول 8 مقدار الخسارة الذي يوفره النموذج المُقْتَرَح من أجل نافذة $m = 7$ في كل عام.

الجدول 8 تحليل التكلفة والفائدة لنموذج LSTM-AEDT

Item	Gain (by TPR)	Loss (by FPR)	Remark
تكلفة	10,000\$	100\$	-
عدد الحالات	124x12month =1488	(2 minx30x24 hr.x365 d)- 1488=524112	1 year
معدل الحالات	Recall=53%	FPR=9%	Test Result
تكلفة/ سنة	7,886,400\$	-47,170\$	
الإجمالي	7,839,230\$		

نُلاحظ أن نموذج LSTM-AEDT المُقترح إستطاع مع عدد ليس بالكبير من عمليات الكشف المبكر، أن يوفر قدراً كبيراً من تكاليف الخسارة وصلت لحوالي 8 مليون دولار سنوياً عن كل خط إنتاج، وبمعدل 54% من الخسارة الإجمالية.

6-2-2- نموذج AEDT

يهدف البحث من خلال استخدام شبكة الترميز الآلي التقليدية في بناء نموذج الكشف، إلى تقديم مقارنة بين الأنظمة القائمة على المنطق الزمني مثل شبكة الترميز الآلي ذات الذاكرة طويلة المدى، والأنظمة التي تعتمد على دراسة النقطة الزمنية الحالية فقط مثل شبكة الترميز الآلي.

1) ضبط البارامترات الفائقة

تُظهر النتائج المُبيّنة في الجدول 9 أفضل تكوينات البارامترات الفائقة لنموذج AEDT، باستخدام ضبط البارامترات الفائقة.

الجدول 9 البارامترات الفائقة لنموذج AEDT

Hyperparameters Configurations						
Neurons	#Epochs	#Batch	Learn Rate	Dropout	Hidden Activation	Out Layer Activation
32	200	64	0.001	0.4	Tanh	Linear

2) تفاضل العتبة (Threshold Tradeoff)

تُظهر النتائج المُبيّنة في الجدول 10 أداء نموذج AEDT في اكتشاف الشذوذ لمشكلة كسر الورق. نلاحظ من خلال الجدول أن النموذج يقوم بتحديد عتبة التصنيف بشكل ديناميكي ضمن المجال [0.2,1.4]، حيث أن 0.2 هي القيمة الافتراضية و1.4 هي القيمة العظمى.

الجدول 10 نتائج نموذج AEDT

Threshold	F1Score	Recall	Precision	FPR	Run Time
Default:0.2	0.49	0.56	0.50	0.84	
0.5	0.52	0.51	0.53	0.42	
0.8	0.52	0.49	0.55	0.25	1.3 min
1.1	0.51	0.49	0.56	0.15	
Max:1.4	0.51	0.49	0.57	0.14	

تُصل أفضل قيمة لقياس Recall إلى 56% عند العتبة ($T = 0.2$)، وقيمة Precision إلى 57% عند العتبة ($T = 1.4$)، أما بالنسبة لقياس F1Score فإن أفضل قيمة تصل إلى 52% عند العتبة ($T = 0.5, 0.8$)، بينما كانت أقل نسبة للإيجابيات الخاطئة FPR هي 14% عند العتبة ($T = 1.4$).

استطاع نموذج AEDT كشف 49% من حالات الكسر قبل 4 دقائق من حدوثها (Recall = 0.49) بنسبة إنذارات خاطئة 14% ($FPR = 0.14$) عند العتبة ($T = 1.4$)، بينما وصل زمن تنفيذ النموذج إلى 1.3 دقيقة.

(3) تحليل التكاليف والفوائد (Cost Benefit Analysis)

يُبين الجدول 11 مقدار الخسارة الذي يوفره النموذج AEDT في كل عام.

الجدول 11 تحليل التكلفة والفائدة لنموذج AEDT

Item	Gain (by TPR)	Loss (by FPR)	Remark
تكلفة	10,000\$	100\$	-
عدد الحالات	124x12month =1488	(2 minx30x24 hr.x365 d)- 1488=524112	1 year
معدل الحالات	Recall=49%	FPR=14%	Test Result

تكلفة/ سنة	7,291,200\$	-73,375\$
الإجمالي	7,217,825\$	

إستطاع النموذج، أن يوفر قدراً جيداً من تكاليف الخسارة وصلت لحوالي 7 مليون دولار سنوياً عن كل خط إنتاج، وبمعدل 46% من الخسارة الإجمالية. أكّدت نتائج التجارب العملية لهذا البحث، قدرة النظام المُقترح على اكتشاف الحالات الشاذة قبل حدوثها بوقت مناسب، والتقاط التبعيات الزمنية لتسلسل الشذوذ ضمن سلاسل البيانات، وتحديد عتبة التصنيف ديناميكياً باستخدام أساليب غير معلميه. إستطاع نموذج LSTM-AEDT عند تجريبه على اكتشاف حالات كسر الورق، كشف 53% من الحالات قبل حدوثها بأربعة دقائق، وبنسبة إنذارات خاطئة لم تتجاوز 9%، كما حدّ من تكاليف الخسارة بنسبة وصلت إلى 54%، ومتقوفاً بذلك على نموذج AEDT الذي إستطاع كشف 49% من الحالات الشاذة، والحدّ من تكاليف الخسارة بنسبة 46%.

7- الخاتمة والاستنتاجات والتوصيات

ركّز البحث على معالجة مشكلة هامة تواجه أنظمة كشف الشذوذ القائمة على التعلّم العميق، وهي تحديّد عتبة التصنيف ديناميكياً. إستطاع البحث أن يتوصل إلى بناء نظام ديناميكي للكشف عن الحالات الشاذة قبل حدوثها بوقت مناسب. تمّت تسمية هذا النظام بـ: LSTM-AEDT، ويتمتع النظام بقدرته على دراسة الارتباطات بين البيانات عبر الزمن، واستخراج الميزات الأكثر أهمية (اختزل أبعاد البيانات)، بالإضافة إلى تحديّد عتبة التصنيف ديناميكياً. اعتمدَ نظام LSTM-AEDT المُقترح على شبكة الترميز الآلي ذات الذاكرة طويلة المدى، لالتقاط التبعيات الزمنية لتسلسل الشذوذ ضمن سلاسل البيانات باستخدام مفهوم النوافذ الزمنية. بالإضافة لنظرية تشبيثيف من أجل تحديّد المجال الذي يقع ضمنه معظم متجهات الخطأ للبيانات الطبيعية. بالتالي يستطيع النظام اختيار العتبة

الأفضل ضمنَ المجال المُحدّد ديناميكياً، دون وضع فرضيات حول توزيع البيانات التي يقوم بمعالجتها. كما إعتَمَدَ على تقنيات ضبط البارامترات الفائقة لتحقيق أفضل أداء له. تَوَصَّلَت نتائج البحث إلى أن أفضل أداء لنموذج LSTM-AEDT يكون عند اختيار حجم نافذة $m = 7$ ، وضبط البارامترات الفائقة بالشكل التالي: عدد العصبونات: 32، عدد الدورات: 50، حجم الدفعة: 32، معدل التعلم: 0.01، معدل الحذف: 0.0، تابع التنشيط: Tanh، وعدد طبقات مخفية: 6 طبقات، حيث بَلَغَت نسبة الكشف 53% من الحالات قبل حدوثها، ونسبة إنذارات خاطئة لم تتجاوز 9% عند العتبة ($T = 1.4$). كما استَطَاعَ الحد من الخسارة الإجمالية لِكُلِّ خط إنتاج، بحوالي 8 مليون دولار سنوياً وبمعدل 54%.

كما تَقَوَّقَ نموذج LSTM-AEDT على نموذج AEDT، الذي حَقَّقَ أفضل أداء له بنسبة كشف 49%، ونسبة إنذارات الخاطئة 14% عند العتبة ($T = 1.4$)، بينما بَلَغَت نسبة الحد من تكاليف الخسار 46%.

تُؤكِّد نتائج البحث على أهمية استخدام خوارزميات التعلُّم العميق القائم على الزمن عند التعامل مع بيانات السلاسل الزمنية، حيث حَقَّقَت أداء أفضل من الخوارزميات التي تَعْتَمِدُ فقط على النقطة الزمنية الحالية. كما تُؤكِّد على ضرورة اتباع أساليب غير معلمية لإيجاد قيم عتبة التصنيف ديناميكياً. وبناء على النتائج التي تم التوصل إليها يمكن أن نقدم بعض المُقترحات:

- محاولة التوسع في الدراسة لتشمل مجالات أخرى في كشف الشذوذ.
- استخدام شبكات عصبونية عميقة أخرى بهدف المقارنة.

المراجع

- [1] Dolgui, A., Ivanov, D., Potryasaev, S., Sokolov, B., Ivanova, M., & Werner, F. (2020). Blockchain-oriented dynamic modelling of smart contract design and execution in the supply chain. *International Journal of Production Research*, 58(7), 2184-2199.
- [2] Ivanov, D., & Dolgui, A. (2020). A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control*, 1-14.
- [3] What is Supply Chain Management? [Online] Available: <https://www.oracle.com> [Accessed 29 Dec 2021]
- [4] Zenati, H., Romain, M., Foo, C. S., Lecouat, B., & Chandrasekhar, V. (2018, November). Adversarially learned anomaly detection. In 2018 IEEE International conference on data mining (ICDM) (pp. 727-736).
- [5] Bajpai, P. (2015). Basic overview of pulp and paper manufacturing process. In *Green chemistry and sustainability in pulp and paper industry* (pp. 11-39). Springer, Cham.
- [6] Nguyen, H. D., Tran, K. P., Thomassey, S., & Hamad, M. (2021). Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, 57, 102282
- [7] Montgomery, D. C. (2020). *Introduction to statistical quality control*. John Wiley & Sons.
- [8] Pathak, P., & Sharma, C. (2021). Processes and problems of pulp and paper industry: an overview. *Physical Sciences Reviews*.
- [9] Karim, F., Majumdar, S., Darabi, H., & Harford, S. (2019). Multivariate LSTM-FCNs for time series classification. *Neural Networks*, 116, 237-245.

- [10] Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. arXiv preprint arXiv:1607.00148.
- [11] Ranjan, C., Reddy, M., Mustonen, M., Paynabar, K., & Pourak, K. (2018). Dataset: rare event classification in multivariate time series. arXiv preprint arXiv:1809.10717.
- [12] Python 3.10.0 documentation [Online] Available: <https://docs.python.org/3/> [Accessed 1 May. 2021]
- [13] Chicco, D. (2017). Ten quick tips for machine learning in computational biology. *BioData mining*, 10(1), 1-17.
- [14] Alsmeyer, G. "Chebyshev's Inequality." (2011): 239-240.
- [15] Esmael, B., Arnaout, A., Fruhwirth, R. K., & Thonhauser, G. (2012, December). Improving time series classification using Hidden Markov Models. In 2012 12th International Conference on Hybrid Intelligent Systems (HIS) (pp. 502-507). IEEE.
- [16] Jović, A., Brkić, K., & Bogunović, N. (2012, August). Decision tree ensembles in biomedical time-series classification. In Joint DAGM (German Association for Pattern Recognition) and OAGM Symposium (pp. 408-417). Springer, Berlin, Heidelberg.
- [17] Cui, Z., Chen, W., & Chen, Y. (2016). Multi-scale convolutional neural networks for time series classification. arXiv preprint arXiv:1603.06995.
- [18] Zhang, C., Yan, H., Lee, S., & Shi, J. (2018). Multiple profiles sensor-based monitoring and anomaly detection. *Journal of Quality Technology*, 50(4), 344-362.
- [19] Zheng, Y., Liu, Q., Chen, E., Ge, Y., & Zhao, J. L. (2014, June). Time series classification using multi-channels deep convolutional neural networks. In International conference on web-age information management (pp. 298-310). Springer, Cham.
- [20] Roy, A., Sun, J., Mahoney, R., Alonzi, L., Adams, S., & Beling, P. (2018, April). Deep learning detecting fraud in credit card transactions. In 2018 Systems and Information Engineering Design Symposium (SIEDS) (pp. 129-134). IEEE.
- [21] Tran, K. P., Du Nguyen, H., & Thomassey, S. (2019). Anomaly detection using long short term memory networks and its applications in supply chain management. *IFAC-PapersOnLine*, 52(13), 2408-2412.
- [22] Kim, T. Y., & Cho, S. B. (2018). Web traffic anomaly detection using C-LSTM neural networks. *Expert Systems with Applications*, 106, 66-76.
- [23] Karim, F., Majumdar, S., Darabi, H., & Harford, S. (2019). Multivariate LSTM-FCNs for time series classification. *Neural Networks*, 116, 237-245.
- [24] Nordling, C. (2020). Anomaly Detection in Credit Card Transactions using Autoencoders.
- [25] Nguyen, T. T., Tahir, H., Abdelrazek, M., & Babar, A. (2020). Deep Learning Methods for Credit Card Fraud Detection. arXiv preprint arXiv:2012.03754.

- [26] Marie-Sainte, S. L., Alamir, M. B., Alsaleh, D., Albakri, G., & Zouhair, J. (2020, July). Enhancing Credit Card Fraud Detection Using Deep Neural Network. In Science and Information Conference (pp. 301-313). Springer, Cham
- [27] Yu, S. and Principe, J.C., 2018. Understanding autoencoders with information theoretic concepts. arXiv preprint arXiv:1804.00057.
- [28] Ghrib, Z., Jaziri, R., & Romdhane, R. (2020, July). Hybrid approach for anomaly detection in time series data. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.
- [29] Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of machine learning research, 13(2).

تحسين أداء خوارزمية التقطير

في انترنت الأشياء

طالبة الدكتوراه: منال سعد الله العمر

كلية الهندسة المعلوماتية - جامعة البعث

إشراف الدكتور: أكرم المرعي + د. محسن عبود

الملخص

اكتسبت خوارزمية Trickle التي تم تقديمها وتوحيدها في RFC 6206 شعبية كبيرة حيث يمكنها أن تضمن حل تناقض البيانات في البيئات الموزعة بدقة وسرعة مع انخفاض تكلفة الصيانة ودعم جيد للتوسع وزيادة كثافة الشبكة. بالنسبة لهذه الميزات الجذابة تشكل Trickle أساس العديد من معايير الإنترنت ويتم نشرها في العديد من التطبيقات مثل البث الموثوق واكتشاف الخدمة والموارد الموزعة وكذلك بروتوكولات التوجيه ومنها بروتوكول RPL المستخدم في شبكات انترنت الأشياء. قمنا في هذا البحث بتقديم خوارزمية تقطير جديدة (New Sys-Trickle) قادرة على تقليل كل من زمن التقارب و عدد حزم التحكم المرسل في الشبكة و مقدار استهلاك الطاقة ومقدار استهلاك المعالج في آن واحد، وهذا ما يميزها عن الخوارزميات السابقة التي كانت تحسن معيار على حساب معيار آخر. حيث قمنا بدمج ثلاث خوارزميات تقطير وهي Trickle-F و Adaptive-k و Trickle-plus مع إضافة تحسيناتنا على كل منها.

الكلمات المفتاحية: خوارزمية التقطير، انترنت الأشياء، بروتوكول RPL ، نظم

موزعة.

Improving the performance of IoT Trickle algorithm

Abstract

The Trickle algorithm that was introduced and standardized in RFC 6206 is gaining popularity as it can ensure fast and accurate resolution of data inconsistency in distributed environments with low maintenance cost, good support for scaling and increased network density. It is for these attractive features that Trickle forms the basis of many Internet standards and is deployed in many applications such as reliable broadcasting, service discovery, distributed resources, as well as routing protocols, including RPL, which is used in IoT networks. In this paper, we presented a new Trickle algorithm (New Sys-Trickle) capable of reducing both the convergence time, the number of control packets sent in the network, the amount of power consumption and the amount of processor operating time at the same time, and this is what distinguishes it from the previous algorithms that were improving a standard at the expense of a standard Else. We have combined three Trickle algorithms Trickle-F, Adaptive-k and Trickle-plus with our improvements to each of them.

KEYWORDS: Trickle algorithm, Internet of Things, RPL protocol, distributed system.

1. المقدمة:

تتطلب شبكات (Low-Power and Lossy Networks) LLN بناء طوبولوجيا الشبكة بأسلوب يتسم بالكفاءة والسرعة، وبالتالي من الضروري وجود خطة فعالة لاختيار اتجاه البيانات والتوجيه بشكل سريع. يعتمد RPL (IPv6 Routing Protocol for LLN) وهو بروتوكول التوجيه في شبكات LLN من أجل الحفاظ على الطاقة والحفاظ على طوبولوجيا الشبكة مع أقل تكلفة للتوجيه خوارزمية Trickle التي تضمن الانتشار السريع لمعلومات التوجيه وتكلفة صيانة منخفضة لطوبولوجيا الشبكة. تستخدم هذه الخوارزمية أسلوب يسمى بالثرثرة المهذبة "polite gossip" وتعني أن المشارك لن ينشر ثرثرة إذا قام شخص آخر بذلك بالفعل وبالتالي بواسطة خوارزمية Trickle يتم التنظيم الذاتي لتبادل الحزم مع العقد المجاورة [2] .

2. مشكلة البحث:

على الرغم من الجهود البحثية الموجودة لتحسين خوارزمية Trickle إلا أنه لا توجد خوارزمية تقطير قادرة على تقليل كل من زمن التقارب وعدد حزم التحكم واستهلاك المعالج والطاقة معاً، لذلك عملنا على إيجاد خوارزمية تقطير قادرة على تحسين هذه المعايير الأربعة معاً.

3. الهدف من البحث:

يهدف هذا البحث إلى تقديم خوارزمية تقطير جديدة ذات أداء أفضل من أداء خوارزميات التقطير في الدراسات السابقة وذلك باتباع الخطوات التالية :

- الهدف 1: الهدف الأول هو اكتساب معرفة متعمقة وإتقان أحدث ما توصلت إليه مفاهيم إنترنت الأشياء مع التركيز بشكل أكبر على بروتوكول التوجيه RPL وخوارزمية Trickle.

- الهدف 2: فهم الكود البرمجي الخاص ببروتوكول RPL داخل نظام التشغيل contiki وذلك بهدف التعديل عليه .
- الهدف 3: دراسة وتحقيق ومقارنة التحسينات على خوارزمية Trickle في بروتوكول التوجيه RPL.
- الهدف 4: تصميم خوارزمية Trickle جديدة بهدف تحسين أداء RPL من حيث تخفيف استهلاك الطاقة وتخفيض وقت تقارب الشبكة وعدد حزم التحكم ومقدار استهلاك المعالج. حيث سيتم التعديل على الملفات التالية-rpl , rpl.h , rpl.c , dag.c , rpl-icmp6.c , rpl-timers.c .

4. أهمية البحث:

نظرًا لحقيقة أن بروتوكول التوجيه هو أحد الأعمدة الرئيسية لهندسة الشبكات، ويتوقع بثقة ضرورة وجود شبكات LLN ، أصبح RPL سريعًا بروتوكول التوجيه الفعلي لـ IoT (Internet Of Thing) ، علاوة على ذلك لأن RPL هو بروتوكول التوجيه الموحد الوحيد لشبكات LLN حتى الآن ، فإن عدد الأعمال المنشورة حول RPL يزداد كل عام بشكل واضح ومن أهمية خوارزمية Trickle في ضبط انتشار بيانات توجيه بروتوكول RPL تتبع أهمية بحثنا هذا الذي يقدم دراسة شاملة لأهم التحسينات على هذه الخوارزمية وتقييمها ومقارنتها ومن ثم تقديم خوارزمية تقطير جديدة ومحسنة قادرة على تقديم أفضل أداء ممكن.

5. انترنت الأشياء:

تتمثل الرؤية الرئيسية لإنترنت الأشياء في خلق عالم ذكي يوفر مزيدًا من الذكاء للطاقة والصحة والنقل والمدن والصناعة والمباني والعديد من المجالات الأخرى، سيتم تحقيق ذلك من خلال جعل الكائنات التي نتعامل معها يوميًا مزودة بأجهزة استشعار ومحدد

هوية ومحدد مواقع ولها عنوان IP لتصبح كائنات ذكية وقادرة على التواصل ليس فقط مع الكائنات الذكية الأخرى ولكن أيضًا مع البشر [3].

6. بروتوكول توجيه IPv6 للشبكات منخفضة الطاقة والخسارة (RPL) [2]:

هو بروتوكول توجيه IPv6 لشبكات انترنت الأشياء حيث تم اقتراحه للشبكات ذات الوصلات المفقودة التي تكشف عن معدل عالي في أخطاء الحزم وبالإضافة إلى فصل وانقطاع الوصلات. يعتبر بروتوكول RPL من النوع Source Routing (path addressing) حيث يسمح لمُرسل الحزمة بتحديد المسار الذي تتبعه الحزمة عبر الشبكة جزئيًا أو كليًا ويتيح للعقدة اكتشاف جميع الطرق الممكنة للهدف كما يوفر RPL دعم التوجيه لنقل البيانات إما multicast أو unicast. تنظم RPL طوبولوجيا الشبكة على شكل DAG (Directed Acyclic Graph) الذي يتم تقسيمه إلى واحد أو أكثر من DODAG (Destination Oriented DAGs) حيث يتم تحديد المسارات الافتراضية بين العقد في LLN بواسطة هذه البنية الشجرية. يمكن أن تنتمي العقدة الموجودة في DAG إلى أكثر من والد مقارنة بالشجرة التقليدية. ترتبط العقد في بنية DODAG ليس فقط بوالدها ولكن أيضًا مع العقد الأخوية [8]. يكون جذر هذه الشجرة في عقدة المصرف ويعمل هذا الجذر كنقطة عبور إلى شبكات IPv6 ويتميز بكونه عقدة لا تحتوي على نفس القيود مثل عقدة المستشعر ويسمى جهاز التوجيه الحدودي. علاوة على ذلك يتم التحكم في إرسال التحديثات إلى العقد الأخرى بواسطة جهاز التوجيه الحدودي.

6-1. رسائل التحكم في RPL [8]:

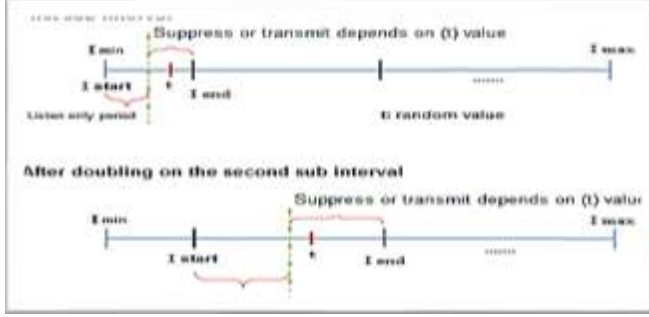
DODAG Information Solicitation (DIS) (a): يتم إرسال هذه الرسالة من خلال عقد جديدة للانضمام إلى DODAG تلتمس من خلالها كائن المعلومات (DIO) من عقدة RPL.

- (b) **DODAG Information Object (DIO)**: رسالة بث مرسله بواسطة عقدة الجذر لإنشاء DODAG يحتوي DIO على معلومات عامة مثل RPLInstanceID أو الترتيب أو DODAGID ، ، إلخ.
- (c) **Destination Advertisement Object (DAO)**: إنها رسالة تُستخدم لإنشاء مسارات توجيهه للأسفل من جذر DODAG إلى العقد الأخرى.
- (d) **Destination Advertisement Object (DAO-ACK)**: ترسل هذه الرسالة من قبل مستلم رسالة DAO (جذر الـ DODAG) كرد على رسالة DAO أحادية الإرسال المستلمة. تحتوي على البيانات المتعلقة بالحالة، تسلسل DAO ، ومعرف حالة الـ RPL .

2-6 . خوارزمية Trickle [9]:

يتكون RPL من مجموعة من الخوارزميات، لكل خوارزمية مهام محددة. الخوارزمية الرئيسية في RPL هي خوارزمية Trickle. يتمثل الهدف الرئيسي لخوارزمية Trickle في إدارة عملية الإرسال في الشبكة، حيث تقوم Trickle بتقليل نشر الرسائل التي لا تحتاج إلى إعادة الإرسال مثل الرسائل المتكررة في الشبكة. يتم ذلك باستخدام آليتين الأولى تحدث عندما تظهر حالة عدم تناسق في الشبكة، تزيد الخوارزمية من معدل التحكم في الإشارة لتعود إلى الوضع التوافقي في الشبكة، والثانية تحدث عند تكرار نفس الرسالة في الشبكة والعقد لم تعد بحاجة إليها لأنها متصلة بجيرانها تعمل عندها الخوارزمية على كبح إرسال الرسالة وهذا يساعد على تقليل الرسائل المنتشرة على الشبكة وتوفير الطاقة. تقوم خوارزمية Trickle بتعيين الفاصل الزمني الرئيسي لكل عقدة في الشبكة، ويبدأ هذا الفاصل الزمني من l_{min} وينتهي عند l_{max} حيث كل من l_{max} و l_{min} متغيرات. تقسم العقدة الفاصل الزمني الرئيسي الخاص بها إلى مجموعة من المناطق الفرعية، كل فاصل زمني يبدأ من l_{start} وينتهي في l_{end} وكل من l_{start} و l_{end} هي متغيرات. يبدأ التنفيذ من الفترة الفرعية الأولى في العقدة، وعند

انتهاء الفترة الفرعية الأولى يبدأ الفترة الفرعية التالية، وما إلى ذلك حتى ينتهي من جميع المناطق الفرعية عندما يصل المؤقت إلى قيمة I_{max} .

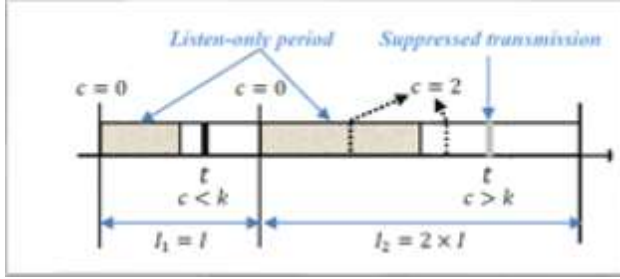


الشكل (1) خوارزمية Trickle القياسية لكل عقدة [9].

تحتوي خوارزمية Trickle القياسية على ثلاثة بارامترات أساسية:

- I_{max} : الحد الأقصى لحجم الفاصل الزمني يوصف I_{max} على أنه عدد من مضاعفات I_{min} .
 - I_{min} : الحد الأدنى لحجم الفاصل.
 - K : القيمة الحدية وتسمى أيضاً بثابت التكرار k .
- محلياً، تحافظ كل عقدة في الشبكة على مؤقت وثلاثة متغيرات:
- a : حجم الفاصل الحالي.
 - C : عدد رسائل Trickle المستلمة أثناء الفاصل الحالي ويسمى عداد الاتساق.
 - t : وقت الإرسال وهو عشوائي محدد داخل الفاصل الحالي (a و $1/2$).
- يمكن التعبير عن خوارزمية Trickle من خلال القواعد أو الخطوات الستة أدناه [1]:
1. **الخطوة 1:** عندما يبدأ Trickle التنفيذ فإنه يختار (a) بشكل عشوائي وموحد من $[I_{min}, I_{min} * 2]$ ويبدأ الفاصل الأول.
 2. **الخطوة 2:** عند بدء فاصل زمني (a) تعيد Trickle إعادة تعيين c إلى 0 وتختار t بشكل عشوائي من النطاق $(1, 1/2]$.
 3. **الخطوة 3:** كلما سمعت عقدة انتقال متناسق مع بياناتها تزيد Trickle عداد الاتساق c .

4. **الخطوة 4:** في الوقت t يرسل Trickle إذا كان $(c < k)$ فقط وخلاف ذلك يتم قمع الإرسال.
5. **الخطوة 5:** عند انتهاء الفاصل الزمني I يضاعف Trickle طول الفاصل حتى الوقت المحدد بواسطة I_{max} ثم يبدأ فاصل زمني جديد كما في الخطوة 2.
6. **الخطوة 6:** إذا سمع Trickle انتقال غير متناسق بينما I أكبر من I_{min} فإنه يصفر المؤقت. للقيام بذلك يقوم Trickle بجعل I يساوي I_{min} ويبدأ فاصلاً جديداً كما في الخطوة 2، وإلا مثلاً I تساوي I_{min} عند اكتشاف التناقض ولن يفعل Trickle شيئاً. يضمن اختيار t من النصف الثاني من الفاصل الزمني في الخطوة 2 فترة استماع تعادل فقط نصف الفاصل ويتم ذلك استجابة لمشكلة الاستماع القصير.



- الشكل (2) خوارزمية Trickle على فترتين مع $k = 1$. الخط الأسود عبارة عن عملية إرسال ونقل، والرمادي هو إرسال تم منعه والخطوط المنقطه هي عمليات استقبال رسائل [1].
7. **الدراسات السابقة:** سنشرح في هذا القسم أهم الدراسات التي حسنت خوارزمية التقطير لنقوم في الأقسام التالية بتحقيقها ومقارنتها ثم إضافة تحسيننا عليها.
- **Trickle-F:** اقترح الباحث في [11] نسخة معدلة من خوارزمية Trickle وهي Trickle-F والتي تهدف إلى حل مشكلة موازنة التحميل من خلال ضمان آلية عادلة لكبح البث على المدى القصير بين العقد في المنطقة لتسهيل الاكتشاف السريع لجميع المسارات المتاحة. يتمثل الأساس المنطقي وراء Trickle-F في تحديد أولوية كل عقدة اعتماداً على عدد عمليات القمع المتتالية بمعنى أنه كلما طال الوقت الذي تقضيه العقدة دون الإرسال زادت أولويتها للإرسال في الجولة التالية.

- **I-Trickle**: الخوارزمية المقترحة في [12] والتي تدعى I-Trickle تضع قيمة عداد التكرار redundancy counter عند قيمة الصفر ليس في بداية الفاصل الزمني الجديد ولكن في وقت منع أو إرسال رسالة كائن معلومات DODAG في خوارزمية Trickle من أجل حل مشكلة موازنة الحمل وتقليل استهلاك الطاقة من خلال مراعاة الرسائل التي يتم سماعها من وقت t إلى نهاية الفاصل الزمني.
- **Optimized Trickle**: تم الاقتراح في [1] تحسين بسيط وقوي يقلل من التأخير دون تكبد أي نفقات عامة إضافية. في هذه الخوارزمية المحسنة عندما يتم اكتشاف عدم تناسق، يعيد تعيين I إلى Imin الذي يتم أيضاً في خوارزمية Trickle الأصلية. لكنه يقول أنه في هذا الوقت سيختار الوقت العشوائي t في [0, Imin] بدلاً من [I/2, I] كما هو الحال في خوارزمية Trickle الأصلية.
- **E-Trickle**: في الورقة [14] تم اقتراح نسخة محسنة من Trickle وهي E-Trickle حيث تقدم هذه الخوارزمية ثلاث تعديلات لخوارزمية Trickle. أولاً بدلاً من تحديد الوقت العشوائي t من النطاق [I/2, I] يتم تحديد قيمة t من النطاق [0, I]. ثانياً بدلاً من تعيين عداد التكرار c على قيمة 0 في بداية كل فاصل زمني يتم تعيين c إلى قيمة 0 فقط في بداية الفاصل الزمني الأول Imin وأيضاً في الوقت الذي تم اختياره عشوائياً t وذلك من أجل القضاء على التأثير التراكمي لمشكلة الاستماع القصير. كما قد يلاحظ المرء، فإن ضبط عداد التكرار c إلى 0 في الوقت المحدد عشوائياً سيؤدي إلى فواصل زمنية غير متساوية بين العقد من حيث الطول وبالتالي، فإن العقد ذات طول الفاصل الأقل سيكون لها فرص أكبر للإرسال. للتغلب على هذا الموقف تم إضافة خطوة ثالثة لتعديل قيمة عامل التكرار k حيث تصبح قيمته $(newk = ((2 * \ln z - 1) / I) * k)$ بشكل متكيف لتعكس التمدد الذي يحدث في حجم الفاصل.
- **Adaptive-k**: يقترح الباحث في [15] امتداداً لخوارزمية Trickle التي تسمى adaptive-k والتي تسمح للعقد بتعيين ثابت التكرار الخاص بها وفقاً للمعلومات المحلية حول كثافة الشبكة. حيث تعتمد كل عقدة على العداد c لتخمين عدد جيرانها وتسمح للعقد بتعيين قيمتها k بشكل مستقل.

- **Trickle-Plus**: يوجد في [10] نسخة موسعة من خوارزمية Trickle. في هذه النسخة بدلاً من زيادة حجم نافذة الإرسال أضعافاً مضاعفة، يتم وضع حجم الفاصل الزمني الجديد بالقيمة $(I * SF * 2)$ حيث SF هو عامل الاراحة ويعني عدد مرات مضاعفة الفاصل الزمني التي يمكن تخطيها، وبالتالي تتقارب الشبكة بسرعة مع استهلاك أقل للطاقة.
- **RANDOMIZED DYNAMIC TRICKLE**: اقترح الباحث في [7] تحسيناً لخوارزمية Trickle القياسية مرتبطاً بقيمة t. في الخوارزمية المقترحة يتم تحديد وقت لفترة الاستماع وفترة الإرسال بناءً على قيمة t ضمن فترة زمنية فرعية. المتغير t هو وقت عشوائي يستخدم كمحدد للاستماع والإرسال. لذلك يركز هذا العمل بشكل أساسي على اختيار قيمة t. حيث يتم اختيار قيمة t وفقاً لكثافة الشبكة، لتحقيق ذلك يتم اقتراح أربع حالات لاختيار قيمة t أفضل في الفترة الفرعية. في كل مرة يتم التحقق من عدد الجيران (C) لتحديد النطاق الذي يجب أن يتم اختيار فيه وقتاً عشوائياً (t).
- **FL-Trickle**: قدم الباحثون في [6] تحسيناً جديداً لخوارزمية Trickle القياسية وهي خوارزمية Flexible Trickle (FL-Trickle) تقلل FL-Trickle من التأخير في إرسال رسائل التحكم عن طريق تحديد وقت الإرسال T عند $1/2$ بدلاً من اختياره بشكل عشوائي في $[1, 1/2]$. تقلل FL-Trickle من معدل النقل من خلال العمل بقيمة عالية من الحد الأدنى للفاصل الزمني (I_{min}) من أجل الحصول على عبء منخفض مع عدم وجود زيادة في استهلاك الطاقة.

8. القسم العملي :

8-1 . نظام التشغيل Contiki ومحاكي الشبكات Cooja [16]:

تم اختيار Contiki لأنه تم تصميمه خصيصاً لأجهزة IoT منخفضة الطاقة والخسارة ولديه تطبيق أساسي لخوارزمية Trickle في مكتبة ContikiRPL التي سيتم استخدامها كأساس لبحثنا . كما تم اختيار المحاكي Cooja الذي يعمل على نظام تشغيل

Contiki (الإصدار 3.0). يعتمد اختيار هذا المحاكى على اعتبارين رئيسيين: الأول هو محاكي مفتوح المصدر مصمم لتطبيقات إنترنت الأشياء. ثانيًا يجعل تعديل وتحسين خوارزمية Trickle سهلًا جدًا بسبب تنفيذه على RPL الأساسي. كما يوفر محاكي Cooja أدوات لإخراج البيانات من كل اختبار بسهولة شديدة بتنسيق قابل للقراءة. هذا ما يسهل علينا استخلاص البيانات منها وإنتاج الرسوم البيانية من المخرجات باستخدام نصوص Perl لتصفية البيانات الناتجة إلى مقاييس الأداء المرغوب فيها، بالإضافة إلى ذلك يوفر هذا المحاكى أداة تدعى collect view التي يمكن للمبرمج دمجها بالكود الخاص به وتعرض معلومات ومخططات مفيدة في عملية قياس الأداء.

8-2. المقاييس المستخدمة لتقييم الأداء ومقارنته:

- **The Convergence Time (setup time):** وقت تقارب الشبكة مقدار الوقت الذي تحتاجه جميع العقد القابلة للوصول (من حيث الراديو) في الشبكة للانضمام إلى DAG.

$$\text{Convergence Time} = \text{Last DIO joined DAG} - \text{First DIO sent}$$

- **Energy Consumption:** مقدار استهلاك الطاقة بالميلي واط من قبل كل من المعالج والاستماع والارسال ونمط توفير الطاقة LPM.
- **Control Traffic Overhead:** وهذا يشمل رسائل DIO و DIS و DAO التي تم إنشاؤها بواسطة كل عقدة.

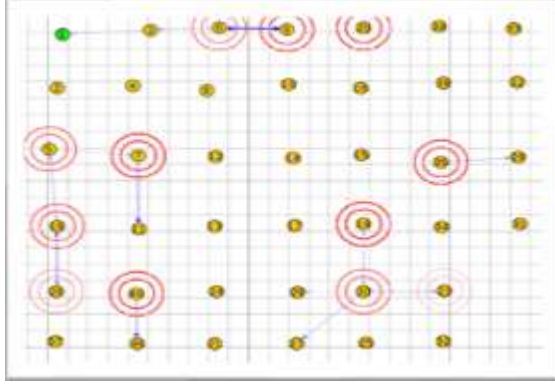
$$\text{Control Traffic Overhead} = \sum_{k=1}^n \text{DIO}(k) + \sum_{k=1}^m \text{DIS}(k) + \sum_{k=1}^o \text{DAO}(k)$$

- **Total CPU consumption (ticks):** مقدار الوقت الذي تم تشغيل المعالج به وذلك لجميع العقد.

لم يتم استخدام معدل تسليم الرزم والتأخير من طرف لطرف في المقارنة وذلك لأنَّ خوارزمية التقطير تقلل عدد حزم التحكم وليس لها تأثير مهم على هذين المعاملين [4].

3-8. معاملات المحاكاة [5]:

نظرًا لأن الاختبارات التي تم إجراؤها لمدة 7 دقائق أسفرت عن نتائج مشابهة لتلك التي أجريت لمدة 20 دقيقة، فقد اخترنا 7 دقائق كوقت المحاكاة لجميع عمليات محاكاة Cooja. تحوي طوبولوجيا الشبكة على 20 ثم 40 عقدة عميل في منطقة 100 م × 100 م. تمثل العقدة 1 في الزاوية اليسرى العليا جهاز توجيه الحدودي. يتم وضع جهاز توجيه الحدود في الزاوية بحيث تكون هناك عُقد ذات مسافات إرسال متعددة الفقرات من جهاز توجيه الحدود. يوضح الشكل التالي طوبولوجيا الشبكة المستخدمة في المحاكاة:



الشكل (3) طوبولوجيا الشبكة المستخدمة في المحاكاة تحوي 40 عقدة بمصرف واحد.

4-8. مستوى التطبيق Application Level:

لتشغيل الاختبارات قمنا باستخدام نموذج تطبيق UDP Contiki يسمى "Hello World!". يرسل هذا التطبيق البسيط رسالة "مرحبًا" في فاصل زمني محدد. بحيث ترسل كل عقدة عميل رسالة "مرحبًا" إلى جهاز التوجيه الحدودي. يستخدم موجه الحدود الملف `udpserver.c` وجميع عُقد المستشعر تستخدم الملف `udp-client.c`. لقد قمنا بتعديل شيفرة المصدر لـ RPL في Contiki لكتابة خوارزميات Trickle حيث تم التعديل على الملف `rpl-timers.c`.

سنستخدم ملحق Cooja المسمى Contiki Test Editor لقياس وقت المحاكاة وإيقاف المحاكاة بعد الوقت المحدد. ينشئ هذا البرنامج المساعد أيضًا ملف سجل

(COOJA.testlog) لجميع مخرجات المحاكاة التي سنقوم بتحليلها في نهاية المحاكاة باستخدام برنامج نصي مكتوب بلغة Perl. من أجل إدخال الضياع في الوسط اللاسلكي نستخدم Cooja Unit Disk Graph Medium، الذي يقدم الضياع فيما يتعلق بالمسافات النسبية للعقد في وسيط الراديو. كما هو موضح في الجدول التالي، فإن تأخير البدء start delay هو تأخير البدء الأولي للتطبيق لبدء إرسال الرسائل إلى عقدة المصرف. وقت البدء الأولي هذا هو الوقت التقريبي الكافي لتقارب الشبكة الأولي. وهذا يضمن أيضاً عدم فقدان الحزمة المرسلة إلى الخادم بسبب نقص اتصال الشبكة. لذلك يمكن إجراء تقييم صحيح لعدد الحزم المرسلة. يتم تعيين DIO و DIO Min و Doublings على القيم الافتراضية في ContikiRPL. تمثل نسبة الاستقبال (RX) مدى ضياع الوسيط الراديوي ويتم تعيينه بالنسب المئوية أثناء التكرار المتتالي للمحاكاة. يتم تعيين نسبة الإرسال (TX) على 100% (بدون خسارة) لأننا لا نهدف إلى إحداث خسائر عند طرف الإرسال ولكن فقط عند طرف الاستقبال. تم ضبط نطاق TX على 50 م ونطاق التداخل على 55 م .

الجدول (1) اعدادات المحاكاة.

Parameters	Value
Start Delay	65 s
RPL MOP	NO_DOWNWARD_ROUTE
OF	ETX
DIO Min	8
DIO Doublings	12
RDC Chanel Check Rate	16
Send Interval	4 s
RX Ratio	30-100%
TX Ratio	100%
TX Range	50m
Interference Range	55m
Simulation Time	7 min
Client Nodes	20,40

9. التجارب والنتائج:

قمنا ومن أجل كل خوارزمية بإعادة التنفيذ عشر مرات وأخذ متوسط القيم الناتجة عن التجارب وذلك بعد استثناء القيم الشاذة والمتطرفة لبعض التجارب بتطبيق قانون المدى الربيعي عن القيم العشر الناتجة.

- مثال لنتيجة تحليل الملف analysis.pl باستخدام COOJA.testlog ملف التحليل analysis.pl الذي قمنا بكتابته بلغة PERL وذلك لخوارزمية Trickle الأصلية:

```
user@instant-contiki:~/contiki-3.9/perl$ perl analysis.pl COOJA.testlog
NETWORK SETUP TIME
=====
First DID      Last DID  Joined DAG  Setup Time(ms)
1338934.888    1388888.888    18466154.888

ENERGY CONSUMPTION
=====
Nodes  Total Transmit ticks  Total Listen ticks  Total Consumption(ticks)  Total cpu  Total Ipm  Total Time  Radio ON Time
20     2810612                2732615              5549227                   7388828   233499414  242799234   2.286

NETWORK TRAFFIC
=====
DID      DTS      DAG
772      10       71
```

الشكل (4) نتيجة تحليل ملف الخرج COOJA.testlog حالة خوارزمية Trickle وشبكة تحتوي 20 عقدة.

```
user@instant-contiki:~/contiki-3.9/perl$ perl analysis.pl COOJA.testlog
NETWORK SETUP TIME
=====
First DID      Last DID  Joined DAG  Setup Time(ms)
1405564.888    28295269.888    20886645.888

ENERGY CONSUMPTION
=====
Nodes  Total Transmit ticks  Total Listen ticks  Total Consumption(ticks)  Total cpu  Total Ipm  Total Time  Radio ON Time
20     9239884                7457961              16727865                   28139452   478285440  488379972   3.354

NETWORK TRAFFIC
=====
DID      DTS      DAG
588      37       148
```

الشكل (5) نتيجة تحليل ملف الخرج COOJA.testlog حالة خوارزمية Trickle وشبكة تحتوي 40 عقدة.

9-1. ملخص النتائج قبل إضافة التحسين:

بناءً على ما سبق يمكننا رسم الجدول التالي الذي يحوي ملخص النتائج حيث ترمز الإشارة (+) إلى أن الخوارزمية المقابلة قدمت أداء أفضل من الخوارزمية الأصلية

بالنسبة للمعيار المقابل وترمز الإشارة (-) إلى أن الخوارزمية المقابلة قد قدمت أداء أسوأ من الخوارزمية الأصلية بالنسبة للمعيار المقابل بينما ترمز الإشارة (*) إلى أن الخوارزمية المقابلة قدمت أفضل أداء من بين جميع الخوارزميات الباقية أما الفراغ يعني أن لها نفس أداء الخوارزمية الأصلية:

الجدول (2) تلخيص النتائج قبل إضافة التحسين.

Total CPU	Energy Cosum 40 20	Traffic Overhead	Setup Time	Metric Algorithm
-			+	Trickle-F
-	-	-	*	I-trickle
-	-	-	+	Opt-trickle
-	-	-	+	E-trickle
+	*	+		Adaptive-k
*	+	+	-	Trickle-plus
-		-	+	RD-trickle
+	-	*	-	FL-trickle

نلاحظ من الجدول السابق أن الخوارزمية التي تقدم أفضل (أقل) زمن تقارب هي خوارزمية I-trickle ولكن ذلك كان على حساب عدد حزم التحكم المرسل في الشبكة وكمية استهلاك الطاقة للذين ازدادا عما كانا عليه في Trickle الأصلية. تقدم Trickle-plus أقل عدد حزم تحكم مرسل في الشبكة لكنها بالمقابل تزيد من زمن التقارب وكذلك FL-trickle التي تقلل من عدد حزم التحكم على حساب زيادة زمن التقارب. تقدم Adaptive-k أقل كمية استهلاك طاقة في الشبكة .

9-2. مناقشة النتائج:

- تحسن Trickle-F من زمن التقارب نعل ذلك بسبب جعل فترة الاستماع فقط للعقد المحرومة من الإرسال لفترات أطول من غيرها أصغر ما يمكن بحيث تتناسب طرذاً مع عدد فترات الحرمان من الإرسال مما يعطي فرصة أكبر للعقد العائمة

الغير منضمة إلى DODAG من تلقي رسالة DIO أسرع ما يمكن . أما بالنسبة لعدد حزم التحكم الذي بقي بنفس القيمة تقريباً كما في خوارزمية التقطير الأصلية فيعود ذلك إلى كون إبقاء الخوارزمية أطوال الفواصل الزمنية كما هي في الخوارزمية الأصلية وعند دعمها لعقد للإرسال كانت بالمقابل تحرم عقد أخرى من الإرسال لذلك لم تشكل زيادة في عدد حزم التحكم وتقديم الخوارزمية فترة استماع فقط كافية بالنسبة لوضع كل عقدة كما نعلم أن مقدار استهلاك الطاقة يعود إلى أربع أشياء تستهلك الطاقة وهي $low\ power\ mode -cpu\ listen\ power -transmit\ power$ و $power-$ بمعنى أن استهلاك الطاقة يتأثر بعدد حزم التحكم المرسل والمستقبل وبما أنها بقيت كما في الخوارزمية الأصلية لذلك فإن استهلاك الطاقة بقي كما في الخوارزمية الأصلية . أما بالنسبة لعدد أجزاء الوقت التي تم استخدام المعالج فيها فقد ازداد بشكل بسيط وذلك بسبب عملية حساب طول فترة الاستماع فقط .

- تحسن trickle-من زمن التقارب وذلك لأنها تتبع نفس أسلوب trickle-F ولكنها تجعل فترة اختيار t من 0 إلى $(1/2^s)$ أي أن فترة الاستماع فقط قد تكون معدومة في بعض الفواصل وبالتالي جعلت فترة الاستماع فقط أقل من سابقتها مما يتيح المجال إلى إرسال الرسائل التحكم بشكل أسرع وانضمام أسرع للبيان الموجه الخاص بالشبكة. تظهر جلياً في هذه الخوارزمية مشكلة الاستماع القصير وعدم التزامن حيث بدون التزامن يمكن أن تعاني Trickle من مشكلة الاستماع القصير. أي أن بعض المجموعات الفرعية من العقد تستمع وقت قصير من بداية الفاصل الزمني وقبل قدرة أي عقدة أخرى لديها فرصة للإرسال على الإرسال. إذا تمت مزامنة جميع الفواصل الزمنية عندها فستمنع عملية إرسال العقدة الأولى أي عقدة أخرى من الإرسال، ومع ذلك إذا لم تتم مزامنتها فربما تبدأ العقدة الفاصل الزمني بعد البث مباشرة وقد تختار أيضاً فترة استماع قصيرة مما ينتج عنه إرسال متكررة. إن الاستماع القصير أدى إلى زيادة عدد حزم التحكم واستهلاك الطاقة واستخدام المعالج زاد بسبب الحسابات.

- بالنسبة لخوارزمية optimal trickle و E-trickle فإن اختيار t ضمن المجال $[0, I_{min}]$ يقلل من زمن تقارب الشبكة ولكن ممكن أن يؤدي إلى ظهور فترة استماع فقط شبه معدومة وتظهر مشكلة الاستماع القصير التي تعني ان العقدة قد تستمع لوقت قصير وتكرر ارسالها لمعلومات مكررة وبالتالي زيادة حزم التحكم في الشبكة وتحسن زمن التقارب في الوقت ذاته .
- خوارزمية Adaptive-k تعتمد على فكرة ضبط قيمة k للفاصل القادم اعتمادا على قيمة c السابق إن هذه العملية والحصول على قيمة k تمثل الجوار يقضي على الارسال الزائد عن الحاجة الذي قد يظهر عندما تكون قيمة k منخفضة والشبكة كثيفة ويقضي على عملية قمع العقدة من الارسال عند ضبط قيمة k على قيمة عالية والشبكة كانت منخفضة الكثافة . لا تحسن هذه الخوارزمية زمن التقارب الذي يبقى كما هو في الخوارزمية الأصلية بينما بسبب تلافي الارسالات المتكررة بشكل فعال تؤدي الى تقليل عدد حزم التحكم واستهلاك الطاقة والمعالج.
- تقوم فكرة خوارزمية trickle plus على عملية تلافي أطوال زمنية وتجاوزها والانتقال إلى الاطوال الأعلى بشكل أسرع مما يجعل عدد رسائل التحكم أقل من حالة trickle الأصلية. تعتمد trickle plus على فترة استماع فقط تساوي نصف الفاصل الزمني كما وأنه من أجل بداية عملية الازاحة بالقرب من I_{min} أدى ذلك إلى ازدياد زمن تقارب الشبكة لأن الخوارزمية تبدأ بمضاعفة الفاصل الزمني قبل انضمام جميع العقد للشبكة وتلقيها رسائل DIO. إن تخفيض عدد رسائل التحكم يخفض بدوره مقدار استهلاك الطاقة اللازمة لإرسال الرسائل كما أن تجاوز اطوال زمنية وانتقال الى اطوال اعلى يجعل عدد عمليات زيادة c ومقارنتها ب k أقل مما يجعل هذه الخوارزمية الأقل استخداماً للمعالج.
- أما خوارزمية RD-trickle نلاحظ انخفاض زمن التقارب بسبب جعل t وفترة الاستماع فقط تتناسب مع عدد الجيران وبالتالي في العقد الحدودية تصبح هذه الفترة

شبه معدومة مما يولد رسائل تحكم زائدة عن الحاجة مما يؤدي الى زيادة استهلاك الطاقة والمعالج.

- تتشابه FL-trickle مع trickle plus حيث حسنت عدد حزم التحكم ولكن ساء زمن التقارب بالمقابل .

10. خوارزمية التقطير الجديدة المتكيفة والمتزامنة New sys-trickle:

انطلاقاً من كون وجود ثلاث خوارزميات Trickle كل منها قادرة على تحسين معامل أداء واحد قررنا في هذا البحث دمج الخوارزميات الثلاثة مع بعضها البعض مع إضافة تحسيننا على كل منها .

الخوارزمية الأولى هي خوارزمية I-trickle التي هي بالأصل تحسين على Trickle-F وانطلاقاً من أن مزامنة الفواصل الزمنية أمر هام جداً وذلك لأنه عند عدم وجود تزامن قد تتزامن فترة الاستماع فقط مع فترة الارسال لعقدة بشكل مستمر مما يؤدي إلى تكرار الرسائل من العقدة الأولى وقمع الرسائل من العقدة الثانية لذلك قدمنا تعديل هام على خوارزمية Trickle-F يجعل من فواصلها متزامنة كما في الشكل التالي :

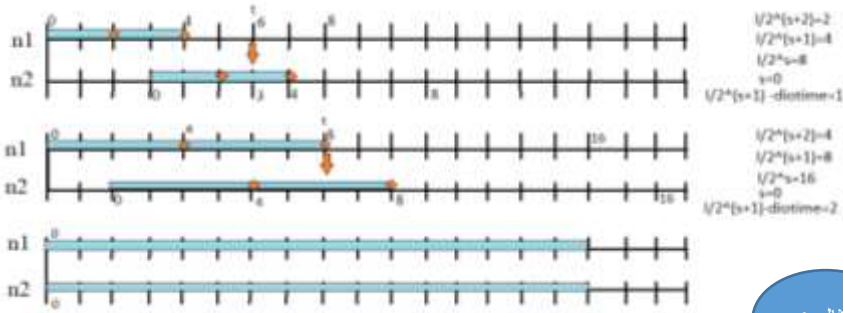
في Trickle-F هناك فترة استماع فقط تنتمي إلى المجال $[0, 1/2^{s+1}]$ بالتالي فإن الفترة المخصصة لاختيار t ضمنها هي $[1/2^{s+1}, 1/2^s]$ نلاحظ أن العقدة تكون غير متزامنة عندما تستلم في النصف الثاني من فترة الاستماع فقط رسائل DIO لذلك نقوم بحساب الاختلاف الزمني بين العقدتين من خلال طرح زمن وصول أول رسالة DIO ضمن الفاصل الزمني $[1/2^{s+1}, 1/2^s]$ من $1/2^{s+1}$.

عند طرح قيمة اللحظة التي استلمت فيها العقدة رسالة DIO (diotime) من $1/2^{s+1}$ سنحصل هنا على زمن tsys الذي سنقوم ببدء الفاصل الزمني التالي على أساسه حيث أن $I=1-tsys$ أي سنجعل الفاصل الزمني التالي يبدأ بشكل أكبر بمقدار يساوي tsys .

إن جعل الفاصل الزمني للعقد يبدأ بشكل أكبر يؤدي إلى جعل الشبكة تتقارب بشكل أسرع كما يؤدي إلى تلافي منع وقمع عقد قد تمتلك معلومات توجيه جديدة بشكل متكرر

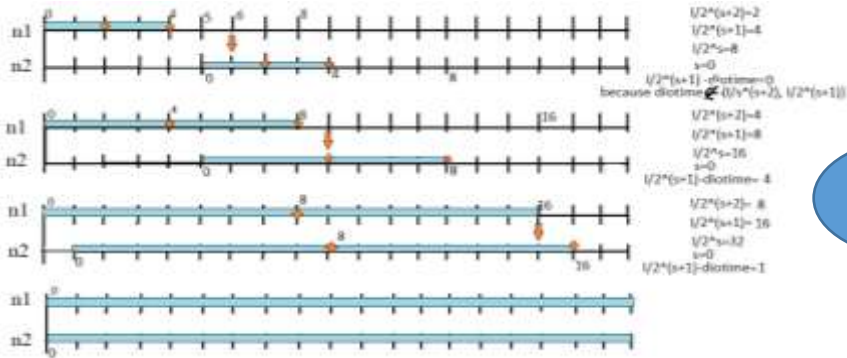
من الارسال وتلافي منح فرصة الارسال بشكل متكرر لعقد أخرى الذي يؤدي إلى نقل معلومات مكررة ضمن الشبكة , وبالتالي تحقيق موازنة جيدة للحمل.

سنقوم بهذه الخطوة فقط في بداية المحاكاة وأثناء عملية الانضمام إلى الشبكة أي في أول أربع فواصل زمنية وذلك لأنه في هذه الفترة تكون قيمة S التي تمثل عدد الفواصل التي تمنع فيها العقدة من الارسال صفر , وبالتالي تكون فترات الاستماع للأب والابن متساوية لأنه عندما تكون هذه الفترات غير متساوية سوف تتلقى العقدة من جيرانها رسائل DIO في الجزء الثاني من فترة الاستماع فقط على الرغم من تزامن الفواصل الزمنية للعقد الذي قمنا بتحقيقه مسبقاً.



مثال 1

الشكل (6) إضافة تزامن للعقد خلال فاصلين زمنين.



مثال 2

الشكل (7) إضافة تزامن للعقد خلال ثلاث فواصل زمنية.

الفكرة الثانية هي في خوارزمية FL-trickle التي هي عبارة عن تحسين على خوارزمية trickle-plus تفرض هذه الخوارزمية أن الازاحة ستكون عشوائية وفي بعض التجارب قد تبدأ الازاحة قبل انضمام جميع العقد وبالتالي زيادة زمن التقارب. التحسين هو أنه سنعمل على جميع الفواصل الزمنية ونجعل iss تساوي Imin و ise ستكون Imax ثم سنعرف متحول يدعى is1 الذي يعبر عن النقطة الفاصلة بين مضاعفة الفاصل بشكل معتاد كما في الخوارزمية الأصلية و الانتقال مباشرة إلى طول الفاصل الزمني الاعظمي , أي أنه عند المرور بمجموعة من الفواصل الزمنية التي تعتبر كافية لتقارب الشبكة سننتقل مباشرة إلى طول الفاصل الأعظمي Imax وبالتالي سنكون ضمناً أن الإزاحة وتلافي الفواصل لن يؤثر على زمن التقارب كما أن الانتقال مباشرة إلى الطول الأعظمي يقلل من عدد رسائل التحكم بشكل واضح.

الفكرة الثالثة هي في خوارزمية adaptive-k تقوم هذه الخوارزمية بضبط قيمة k حسب عدد الجيران وبالتالي مع كل فاصل زمني جديد قد تظهر قيمة جديدة لـ k ولكن يمكن في بعض الفواصل استقبال رسائل من نصف عدد الجيران أو ربعهم أو من لا احد وبالتالي تغيير قيمة k غير فعال. لذلك اقترحنا إضافة متحول نضع فيه قيمة k من الفاصل السابق إذا كانت قيمة k الجديدة أكبر من القديمة عندها لن نغير قيمة المتحول وستبقى في k القيمة الأكبر التي حصلنا عليها من الفواصل السابقة.

```

user@instant-cootiki:~/cootiki-3.0/perls/perl5 perl analysis.pl COOJA.testlog
NETWORK SETUP TIME
-----
First DIO      Last DIO joined DAG      Setup Time(ms)
1185748.000    4746038.000              3642894.000

ENERGY CONSUMPTION
-----
Nodes  Total Transmit ticks  Total Listen ticks  Total Consumption(ticks)  Total cpu  Total Ipm  Total Time  Wradio ON Time
29     1949825                2350943             4300768                 5598758   237281088  242798836  1.771

NETWORK TRAFFIC
-----
DIO      805      840
29       14       38
    
```

الشكل (8) نتيجة تحليل ملف الخرج COOJA.testlog حالة خوارزمية New Sys-

Trickle وشبكة تحتوي 20 عقدة.

```

user@minotaur:~/cooja-3.9/per13/per13_analysis.pl: COOJA: testing
NETWORK SETUP TIME
-----
First DIO      Last DIO  Joined DAG      Setup Time(us)
1813850.988    18132700.988    18819641.988

ENERGY CONSUMPTION
-----
Nodes  Total Transmit ticks  Total Listen ticks  Total Consumption(ticks)  Total cpu  Total Ipe  Total Time  Max(CPU Time)
49     8382113                7388073              15771186                  18488127  479812471  498378508  3.109

NETWORK TRAFFIC
-----
I/O    S/S    O/O
1/3    11     111
    
```

الشكل (9) نتيجة تحليل ملف الخرج COOJA.testlog حالة خوارزمية New Sys- Trickle وشبكة تحتوي 40 عقدة.

```

Algorithm new sys-Trickle
function INITIALIZATION()
    I ← Imin sys = 0 kmid = 0 lss = lmin
    s ← 0 diotime = 0 lsl = 5
function INTERVALBEGINS()
    c ← 0
    sys = 0
    t ← random(1/(2s+1), 1/2s)
function CONSISTENTTRANSMISSIONRECEIVED()
    c ← c + 1
    diotime = clock_time()
    if diotime > I / 2s+2 && diotime < I / 2s+1
        sys = I / 2s+1 - diotime
function TIMEREXPIRES()
    if k ≥ c then
        Transmit DIO
        s ← 0
    else
        s ← s + 1
    end if
    if sys > 0 && l < 5 l = l - sys
function INTERVALENDS()
    c ← 0
    newk = 0.5 * c
    if newk < kmin k = kmin
    else if newk > kmax k = kmax
    else if newk > kmin && newk < kmax k = floor(newk)
    if kmid > k k = kmid
    else kmid = k
    if I > lss && I < lsl
        I = I * 2
    else
        I = Imax
if InconsistentTransmissionReceived then
    I ← Imin
    s ← 0
    
```

الشكل (10) كود الخوارزمية.

11. النتائج النهائية:

❖ زمن التقارب (us) The Convergence Time :

- الحالة الأولى شبكة تحتوي 20 عقدة موزعة بشكل منتظم:

الجدول (3) متوسط زمن التقارب لجميع الخوارزميات حالة شبكة تحوي 20 عقدة.

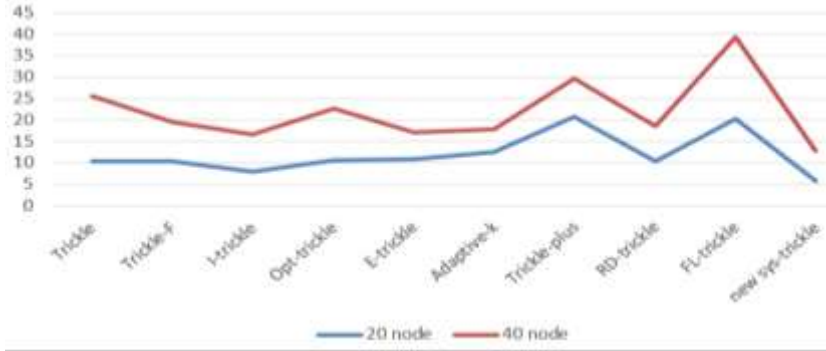
تحسين أداء خوارزمية التقطير في انترنت الأشياء

Setup Time (us)										
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle	
10466754	17288865	7274411	9833959	7351746	8788846	53738188	11814344	81726613	3642894	
16389861	10955983	5442737		9194282	7088198	13879747	9841382	13879746	7521462	
9533161	10660428	8788513	13238877	13918589	8542192	28662655	9917268	28662654	7397528	
8782718	12763878	5948187	10208975	11839528	15819565	1197414	8883927	1197413	7781419	
6498822	9681874	9733645	12492798	9883283	22856277	57299255	7733258	4733248	5163264	
8562928	8128167	11899811	8749572	13923888	15943888	13872648	13876755	13876759	3884738	
9688831	6314797	5348189	9562648	15618232	18842928		7846437	61543851	4243321	
7375552	7251512	8318819	9818738	9874613	11146887	14364333	18837183	14364352	4489128	
14188758	13888881	11312994	15626188	11888882	11252843	4773977	11981822	4773976	8332814	
9159267	5989226	5543832	8883483	6818887	8283888	4886785	12758895	4886784	7955311	
18253528.1	18253224.1	7988769.4	18128266	18765838	12466264	28787427.13	18248885.5	28334828.8	5981188.1	
18.2535	18.2532	7.988769	18.128	18.76183	12.466	28.787	18.248	28.334	5.98	

• الحالة الثانية شبكة تحتوي 40 عقدة موزعة بشكل منتظم:

الجدول (4) متوسط زمن التقارب لجميع الخوارزميات حالة شبكة تحوي 40 عقدة.

Setup Time (us)										
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle	
48855148	17416753	14648888	18898322	15897859	12158532	31715529	18773931	52238323	18818641	
17349523	22843382	28183758	28178289	28842879	14385461	18246483	15118874	21378268	17691435	
28357426	24887818	14287715	2881817	23283283	21856557	18944838	16844838	61855813	14972835	
58418853	29788828	21178419	11488959	15849781	18898253	61974434	13575888	39973721	14288358	
17475568	16936437	15724871	22288839	17988915	18892852	21185586	22688188	31824718	8423179	
26888845		11198886	18718811	19386832	11714759	13953223	22931517	14363945	15811296	
11881187	16488855	13857995	38895884	18813311	14613986	18371649	24178348	18371648	18545884	
18524198	18365188	8638825	28842332	24883226	18886878	25333581	22413838	25885222	18888388	
18112137	18613518	2738813	15888888	28818759	11488285	28888885	13377894	78888878	13872883	
21518818	17989378	18878917	18895362	15489556	17813884	61288884	15523752	71885217	8493647	
25438228.9	19588882.6	16695811.7	22651888.5	171218812.8	17733812.5	29678886.1	18624595	99177817.3	12768557.4	
25.438	19.58	16.69	22.65	17.12	17.73	29.67	18.624	28.177	12.786	



الشكل (11) المخطط الخطي لزمن التقارب.

❖ Traffic Overhead (عدد حزم التحكم):

• الحالة الأولى شبكة تحتوي 20 عقدة موزعة بشكل منتظم:

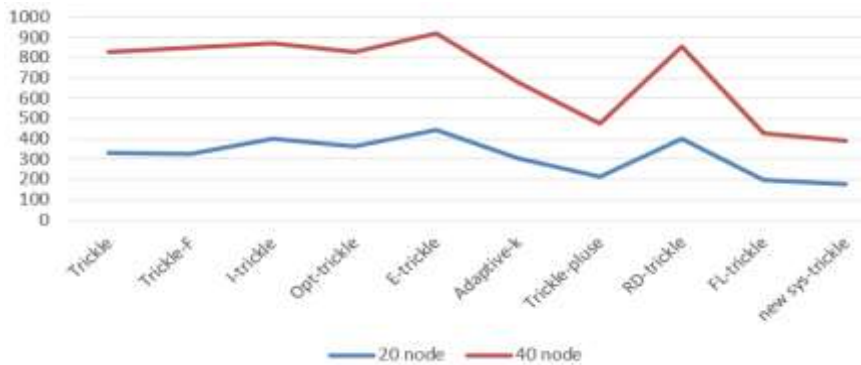
الجدول (4) متوسط عدد حزم التحكم لجميع الخوارزميات حالة شبكة تحوي 20 عقدة.

Traffic Overhead										
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle	
306	308	364	359	388	283	232	364	199	178	
332	324	478		456	311	211	374	177	182	
291	306	350	336	505	285	254	448	264	169	
312	314	391	374	421	324	222	354	183	161	
298	358	430	418	457	299	190	392	196	182	
352	344	372	354	429	329	218	379	186	184	
321	295	375	352	429	286		521	203	187	
302	317	375	364	438	323	282	389	181	165	
342	306	418	366	436	323	287	358	192	194	
383	333	438	337	512	288	233	441	177	179	
330.9	325.7	399.9	361.3	445.4	305.1	235.2	402	195.8	178.1	

• الحالة الثانية شبكة تحتوي 40 عقدة موزعة بشكل منتظم:

الجدول (5) متوسط عدد حزم التحكم لجميع الخوارزميات حالة شبكة تحوي 40 عقدة.

Traffic Overhead										
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle	
933	778	958	767	1198	842	451	763	378	372	
755	777	871	766		798	476	946	389	412	
971	877	915	898	1088	835	543	833	458	395	
728	819	829	832	873	866	537	918	447	466	
737		787	855	856	897	448	853	415	329	
931	932	921	1058	828	827	489	795	488	395	
795	1025	815	842	843	791	487	823	458	338	
888	728	898	878	874	862	493	816	487	378	
788	829	883	738	818	824	456	959	486	366	
831.4	847.8	873	838.2	917.8	879.1	478.5	857.1	427.8	389.2	



الشكل (12)المخطط الخطي لعدد حزم التحكم.

❖ Total CPU (ticks):

• الحالة الأولى شبكة تحتوي 20 عقدة موزعة بشكل منتظم:

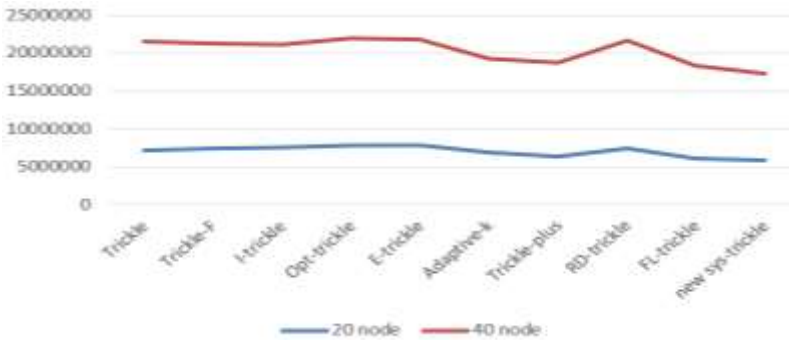
الجدول (6) متوسط زمن تشغيل المعالج لجميع الخوارزميات حالة شبكة تحوي 20 عقدة.

Total cpu (ticks)									
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle
7345945	6952532	7825506	7238631	7429884	4958385	5504658	7838214	5468691	5588756
7971606	7713851	8713367	8292419	4366148	8841336	6841336	7602217	1746159	5132746
6988951	7406868	6521418	6712185	7747148	4918734	8010889	8821897	7531047	5163049
8913738	7875322	8153134	8078936	8401013	7524329	8951406	6731958	8657423	6977117
8877934	7809271	8872173	7881737	8847714	4479787	8182338	7277488	1988478	8885133
7542731	7579624	7588855	7897483	7518727	7888959	5838413	6827438	1798388	6574861
7526724	6962872	6416846	6653428	7875683	6871429		7882186	8866996	5771887
7188442	7191783	7727335	7982395	8854088	6758744	5896255	7871861	1022298	5747467
7848588	7191168	8138371	7277621	8348414	6889558	6429553	6895798	6218423	6198614
7898345	7486841	7589467	4387198	8841788	4737118	8926698	7364755	1818777	5696281
7148188.4	7379184.8	7583487.2	7853312.8	7983139.8	6828311.3	8311282.7	7361818.5	6889664.8	5877511.1

- الحالة الثانية شبكة تحتوي 40 عقدة موزعة بشكل منتظم:

الجدول (7) متوسط زمن تشغيل المعالج لجميع الخوارزميات حالة شبكة تحوي 40 عقدة.

Total cpu (ticks)									
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new syn-trickle
23517493	19578845	22875640	20883128	22553158	28359388	17933422	22272861	17172979	18468127
23849982	21892411	21589461	19363458	22522998	17825284	17252774	28417211	16499228	16524545
21888834	19989682	22649196	23872393	18625888	17797888	23980321	23980321	17648846	17639429
23874484	21841546	21883526	20529876	22888541	18486481	28116742	21617888	21228578	15599488
19719138	28699182	28499888	28943665	28983418	18278298	18833584	22725412	17544774	18788141
28138432	19697381	22348824	21770701	23339714	19455967	28988387	28988387	18421735	16588887
23538947	24764771	19643342	27428898	19799888	18693492	17728887	28354181	17863871	16739733
21465372	23188667	21776572	23147446	22684888	19819319	19788474	21628886	28271881	15857885
19418695	18943855	28394449	23192598	22849893	18237213	18682172	28957117	19312415	28954444
19683391	21642255	28279529	18988918	23146864	19388834	18379439	21132886	17624753	16278869
21458878	21187621.5	21125999.2	21895819	21884881	19381492.5	18788822.9	21693338.8	18378128.2	17245386



الشكل (13) المخطط الخطي لزمن تشغيل المعالج.

❖ Energy Consumption

لحساب استهلاك الطاقة نستخدم أداة Powertrace المتوفرة في Contiki التي تقوم بتحديد ملامح الطاقة على مستوى الشبكة للشبكات اللاسلكية منخفضة الطاقة، حيث تقدر استهلاك الطاقة من قبل وحدة المعالجة المركزية، ونقل الحزم والاستماع . يتم ارسال هذه المعلومات إلى الأداة collect view التي تعرضه على الشكل التالي:

Node Control		Sensor Map		Network Graph		Sensors		Network		Power		Node Info		Serial Console	
Node	Recv.	Dups	Lost	Hops	Reverts	RTT	Churn	Beacon Interval	Repacks	FU Power	LPA Power	Listen Power	Transmit Power	Power	
1.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
2.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
3.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
4.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
5.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
6.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
7.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
8.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
9.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
10.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
11.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
12.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
13.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
14.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
15.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
16.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
17.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
18.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
19.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
20.0	0	0	0	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
Avg	6.000	0.000	0.000	0.000	1091.368	52.211	0.000	5 min, 30 sec	0.000	0.152	0.152	0.660	0.580	1.561	

الشكل (14) استهلاك الطاقة بالميلي واط لجميع العقد في الشبكة التي تحتوي 20 عقدة حالة خوارزمية Trickle الأصلية.

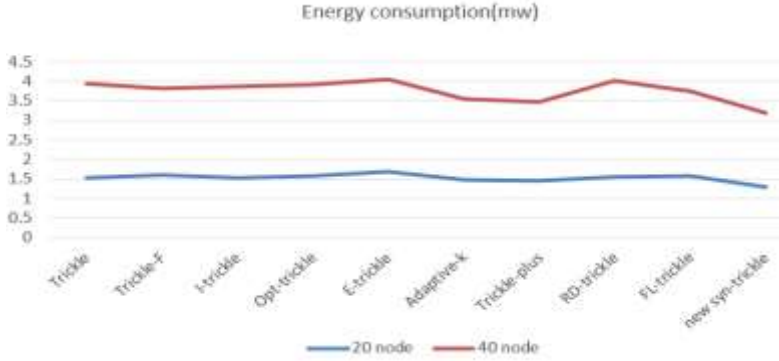
كما في الشكل السابق فإنه ومن أجل كل خوارزمية ومن أجل شبكة بعدد عقد 20 وشبكة بعدد عقد 40 يوجد قيمة للطاقة Power سنزيتها في الجدول التالي :

الجدول (10) متوسط استهلاك الطاقة لجميع الخوارزميات حالة شبكة تحوي 20 عقدة.

power consumption (mW)										
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new sys-trickle	
1.561	1.476	1.506	1.406	1.583	1.569	1.337	1.686	1.625	1.366	
1.439	1.658	1.811	1.532	1.679	1.342	1.336	1.685	1.485	1.165	
1.661	1.628	1.346	1.413	1.666	1.464	1.884	1.654	1.425	1.249	
1.462	1.693	1.627	1.675	1.786	1.587	1.335	1.385	1.832	1.318	
1.492		1.586	1.743	1.773	1.431	1.407	1.451	1.372	1.383	
								1.747	1.351	
1.523	1.61	1.527	1.569	1.697	1.478	1.443	1.556	1.58	1.292	

الجدول (11) متوسط استهلاك الطاقة لجميع الخوارزميات حالة شبكة تحوي 40 عقدة.

power consumption(mw)									
Trickle	Trickle-F	I-trickle	Opt-trickle	E-trickle	Adaptive-k	Trickle-plus	RD-trickle	FL-TRICKLE	new sys-trickle
2.546	2.106	2.398	2.308	2.292	2.091	2.819	2.610	2.456	1.926
2.474	2.296	2.248	2.248	2.549	2.129	1.916	2.443	1.888	1.961
2.295	2.311	2.298	2.196	2.402	1.979	2.328	2.258	1.992	1.775
2.261	2.163	2.487	2.529	2.409	1.913	1.989	2.591	2.458	1.856
2.474	2.276	2.284	2.462	2.166	2.219	1.968	2.357	1.997	1.968
2.41	2.22	2.337	2.347	2.344	2.066	2.04	2.451	2.158	1.897



الشكل (15)المخطط الخطي لاستهلاك الطاقة.

12. التوصيات:

قدمنا في هذا البحث خوارزمية تقطير جديدة متزامنة قادرة على إعطاء أقل زمن تقارب ممكن وذلك بسبب تزامن الفواصل الزمنية للعقد واختيارها فترة استماع فقط منطقية متغيرة مع حالة القمع لكل عقدة على حدى. بالإضافة على قدرة هذه الخوارزمية تخفيض العبء اللازم للتحكم بالتوجيه ببروتوكول RPL من خلال تلافي فواصل زمنية قد لا يكون للمرور فيها داعي أثناء التنفيذ والانتقال مباشرة إلى الطول الأعظمي للفواصل الزمني بعد ضمان تقارب الشبكة . كما أثبتت هذه الخوارزمية قدرتها تخفيض استهلاك الطاقة مقارنة بنظرائها من الخوارزميات لاستخدامها قيمة k تتناسب مع الجوار وبالتالي تقليل تكرار رسائل التحكم في الشبكة وتعتبر هذه الخوارزمية غير مستهلكة للمعالج بما يتناسب مع طبيعة العقد في شبكات الطاقة المنخفضة والخسارة التي تتميز بمواردها المحدودة .

13. المراجع:

- [1] Djamaa, B., & Richardson, M. (2015). Optimizing the Trickle Algorithm, 13(9), 10–13.
- [2] Zhao, M., Kumar, A., Han, P., Chong, J., & Lu, R. (2016). A comprehensive study of RPL and P2P-RPL routing protocols : Implementation , challenges and opportunities. Peer-to-Peer Networking and Applications, (October 2017).
- [3] Bagula, A., & Pietrosemoli, E. (n.d.). Internet of Things IN 5 DAYS.
- [4] Goyal, S., Chand, T., & Member, S. (2018). Improved Trickle Algorithm for Routing Protocol for Low Power and Lossy Networks, 18(5), 2178–2183.
- [5] Benson, D. J. (2016). A Performance Study of RPL with Trickle Algorithm Variants.
- [6] Lamaazi, H., Benamar, N., Kahili, N. E. L., & Taleb, T. (2019). FL-Trickle : New Enhancement of Trickle Algorithm for Low Power and Lossy Networks.
- [7] Yassein, M. B., Alnadi, A., & Bataineh, A. (2018). RANDOMIZED DYNAMIC TRICKLE TIMER ALGORITHM FOR INTERNET OF THINGS, 187–197.
- [8] Idrees, A. K., & Witwit, A. J. H. (2018). A Comprehensive Review for {RPL} Routing Protocol in Low Power and Lossy Networks.
- [9] Yassein, M. B., Hmeidi, I., Shehadeh, H., & Yaseen, W. B. (2015). Performance Evaluation of “ Dynamic Double Trickle Timer Algorithm ” in RPL for Internet of Things (IoT).
- [10] Ghaleb, B., Al-dubai, A., Ekonomou, E., Paechter, B., & Qasem, M. (2016). Trickle-Plus : Elastic Trickle Algorithm for Low- Power Networks and Internet of Things, (MEIoT).

- [11] Vallati, C., & Mingozzi, E. (2013). Trickle-F : fair broadcast suppression to improve energy-efficient route formation with the RPL routing protocol.
- [12] Yassein, M. B., Mohammad, R., Masadeh, T., World, T., & Science, I. (2017). A New Dynamic Trickle Algorithm for Low Power and Lossy Networks, (September 2016).
- [13] Ph . D , Performance Evaluation and Improvement of the RPL Protocol. (n.d.).
- [14] Ghaleb, B., Al-dubai, A., & Ekonomou, E. (2015). E-Trickle: Enhanced Trickle Algorithm for Low-Power and Lossy Networks,
- [15] Meyfroyt, T. M. M., Stolikj, M., & Lukkien, J. J. (2015). Adaptive Broadcast Suppression for Trickle-Based Protocols, 0–8.
- [16] F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt, “Cross-level sensor network simulation with cooja,” in Local computer networks, proceedings 2006 31st IEEE conference on, 2006, pp. 641–648.

تطبيق التعلم الجماعي في صنع القرار الطبي

طالب الدكتوراه م. نهله سعيد سعد الدين

كلية الهندسة المعلوماتية - جامعة البعث

إشراف الدكتورة يسر السيد سليمان الأتاسي

الملخص:

لوحظ أنه مع توفر المزيد من البيانات، فإن الزيادة في استخدام خوارزميات التعلم الآلي ستسمح للمؤسسات بالنظر في نطاق أوسع بكثير من مجموعات البيانات أو المدخلات مما كان ممكناً في السابق. وهذا ما يوفر فرصة لاتخاذ قرارات أفضل - من خلال الجمع بين ذكاء الإنسان والآلة بطريقة ذكية.

تجدر الإشارة إلى أن مدى الاستخدام المستقبلي للخوارزميات في صنع القرار سيختلف حسب مجال تطبيقها، إلا أن المشكلة الأساسية هي الافتراض أن الخوارزميات قريبة من الكمال، أو ستحل محل البشر تماماً في جميع عمليات صنع القرار.

تزداد أهمية الخوارزميات في اتخاذ القرار، وقد تعزز أيضاً القرار البشري وتدعمه.

نعرض في هذه المقالة دراسة حول تطبيق صنع القرار باستخدام التعلم الجماعي (Ensemble Learning)، كما نحاول تسليط الضوء على صنع القرار في المجال الطبي بهدف مساعدة الأطباء عموماً في صنع القرار المناسب للحالة المدروسة.

الكلمات المفتاحية: التعلم الجماعي - خوارزميات - صنع قرار - تعلم الآلة - شبكات

InceptionV3- ResNet - تصنيف الصور.

Apply Ensemble Learning in Medical Decision-Making

Abstract:

It has been observed that as more data becomes available, the increase in the use of machine learning algorithms will allow organizations to consider a much wider range of data sets or inputs than was previously possible. This provides an opportunity to make better decisions – by combining human and machine intelligence in an intelligent way.

It should be noted that the extent of future use of algorithms in decision-making will vary according to their field of application, but the main problem is the assumption that algorithms are close to perfect, or will completely replace humans in all decision-making processes.

Algorithms are becoming increasingly important in decision-making, and they may also enhance and support human decision-making.

In this article, we present a study on the application of decision-making using group learning, and we try to shed light on decision-making in the medical field with the aim of assisting physicians in general in making the appropriate decision for the case studied.

Keywords: Ensemble Learning – Algorithms – Decision Making – Machine Learning – ResNet Networks – InceptionV3 – Image Classification.

1 - مقدمة:

صنع القرارات وظيفية ضرورية للغاية لأي شخص ولأي مجموعة ولأي مجتمع بشكل عام. فبدون القرارات، لا يحدث شيء فعلياً، كما تتحكم القرارات في مصائرنا كأفراد مستقلين أو كأفراد داخل مجموعات (مؤسسة أو شركة أو المجتمع ككل).

نحن نتخذ قرارات كل يوم. هذا جزء من عملنا وحياتنا، وفي معظم الأوقات تسيير الأمور في نصابها الصحيح مع تلك القرارات. فإذا كنا بحاجة إلى اتخاذ قرارات كبيرة وطويلة الأمد سيكون لها تأثير على ما نفعله في المستقبل، فسيكون من الجيد أن يكون لدينا نوع من نموذج صنع القرار الذي يساعدنا ويوفر أساساً ثابتاً للقرار الذي يجب أن نتخذه.

القرارات مهمة، لكن نلاحظ أنه نادراً ما يتم الاهتمام بكيفية صنع القرار. فبدلاً من اختيار نموذج اتخاذ القرار الصحيح للموقف المطروح، إما أن نتراجع عن القواعد أو نقرر ببساطة بأنفسنا بمعزل عن الآخرين. إلا رد الفعل السريع يمكن أن يكون له عواقب سلبية اقتصادية وثقافية وشخصياً وصحياً داخل المجتمع أو المؤسسات.

قرارات قرارات، هذا هو واقع الحياة، هل أنا أتخذ القرار الصحيح؟، هل هذا هو الخيار الأفضل؟

توفر خوارزميات لكل أنواع المشكلات التي من الممكن أن تحدث في العالم يتيح لنا القول إن هناك خوارزمية لذلك (لصنع القرار) أيضاً. وفعلياً عندما نتخذ أي قرار في حياتنا، فإننا نريد أفضل نتيجة بأقل قدر من الأسف.

كما نعلم فإن الخوارزميات تستخدم منذ سنوات كثيرة. إلا أن التوافر المتزايد للبيانات، وإمكانية الوصول إلى أجهزة الكمبيوتر ذات قوة المعالجة الكبيرة، أثر في تطوير خوارزميات جديدة للمساعدة في التنبؤ بالسلوك وأتمتة القرارات. وأدى ذلك إلى تمكين اتخاذ قرارات أكثر كفاءة ومستنيرة في الحكومات والشركات والهيئات والمؤسسات عموماً.

عند اتخاذ قرار ما من الممكن أن يكون لدينا خيارات متعددة للاختيار بينها، فأياً منها يجب أن نختاره حتى يكون لدينا أكبر فرصة لاختيار الخيار الأفضل؟

نواجه هذه الحالة في عدة حالات، مثلاً في عملية التوظيف لاختيار المرشح المناسب والأفضل بين المرشحين المتقدمين، أو اختيار منزل بين عدة خيارات متاحة أو.... الخ،

هذه العملية مقسمة إلى مرحلتين، مرحلة الاستكشاف ومرحلة الاختيار. أولاً يجب استكشاف عدد محدد من الخيارات لإنشاء خط الأساس الخاص بنا الذي سنعتمده كمعيار تقاضل بين الخيارات، ثم اختيار الخيار الأول الذي يتجاوز خط الأساس الذي تم تحديده. كما أنه في السنوات الأخيرة، تركزت الدراسات حول فرص استخدام خوارزميات التعلم الآلي في صنع القرار. وكما نعلم التعلم الآلي هو نوع من الذكاء الصناعي الذي يزود أجهزة الحاسوب بالقدرة على الوصول للنتائج تلقائياً عند تلقيها بيانات جديدة، دون أن تكون مبرمجة بشكل صريح على تلك البيانات. وهناك العديد من الأمثلة على الاستخدام الحالي والمستقبلي لخوارزميات التعلم الآلي في صنع القرار، نذكر في هذا السياق:

- في المجال المالي: يتم استخدام التعلم الآلي لأتمتة قرارات التداول واكتشاف فرص الاستثمار للعملاء بشكل سريع.
- في المجال التعليمي: يتم استخدام التعلم الآلي لتقييم أداء التدريس في المدارس. كما يمكن استخدام التعلم الآلي لدعم كفاءة التعلم عن طريق اختيار تقييمات وموارد تعلم أخرى للطلاب.
- في المجال الطبي: اكتسب التشخيص المعتمد على التعلم الآلي اهتماماً في السنوات الأخيرة، ويرجع ذلك إلى نتائج استدلالها الأسرع وقدرتها على أداء المهام المعقدة، والتي تتطلب خبرة وتجربة متخصصة. يتم استخدام التعلم الآلي لتحسين دقة تشخيص الأمراض المختلفة من خلال اكتشاف الأنماط، كما يمكن استخدام التعلم الآلي للتنبؤ بالاستجابة لمسارات علاج معينة، ما يتيح اتخاذ قرارات أكثر فائدة حول خيارات العلاج المخصصة (وهو المجال الذي تم العمل عليه في هذه الدراسة).

2- هدف البحث:

نحاول من خلال الدراسة دعم المجموعات في صنع القرار من خلال تطبيق التعلم الجماعي في عملية صنع القرار الطبي (التشخيص كما تم ذكره في القسم السابق) أصبح الذكاء الصناعي منتشراً في معظم جوانب صنع القرار في العالم. كما يتزايد استخدام الخوارزميات بسرعة بسبب توفر كميات كبيرة من البيانات وتحليلها من قبل

الحكومات والشركات والمؤسسات عموماً للاستفادة منها في عمليات صنع القرار العام. نقدم في هذه الدراسة تطبيقاً لتجميع التعلم لعدد من خوارزميات تعلم الآلة المستخدمة حديثاً في جانب صنع القرار.

يمكن لعملية صنع القرار باستخدام الخوارزميات القضاء على التحيز أو التمييز في عملية صنع القرار، وذلك يعتمد على جودة وتمثيل البيانات الأساسية التي تستخدمها الخوارزمية، بحيث تتأثر جودة البيانات بعدد من العوامل التي تشمل دقة التوقيت والتفصيل وجودة البيانات الوصفية ... الخ.

نحاول في هذه الدراسة أيضاً توضيح ما إذا كان نظام الذكاء الصناعي يمكن أن يعمل كصانع قرار أساسي، أو كمساعد وداعم مهم لصانع القرار البشري، بحيث إن الجمع بين أساليب الذكاء الصناعي في دعم القرار يمكن من استخدام كميات كبيرة من البيانات لإنتاج معلومات صنع القرار ذات الصلة بالمجال المدروس. كما نعمل في هذا البحث على توضيح كيف يمكن دمج منهجيات الذكاء الصناعي المختلفة في أنظمة صنع القرار، بحيث يتم تقديم إطار عمل لوصف تطبيقات الذكاء الصناعي في أنظمة صنع ودعم القرار لاحقاً.

أحد تطبيقات التعلم الآلي الأكثر شهرة هو في مجال صنع القرار، بحيث يتم تنفيذ أنظمة صنع القرار الحديثة على مجموعات بيانات ضخمة، وتعمل بشكل خاص على التعلم العميق وهو شكل من أشكال التعلم الآلي، ولكن ليس في قالب الأنظمة الخبيرة التقليدية التي تتم برمجتها يدوياً لإيجاد حلول لمشاكل محددة.

في نظم دعم القرار الأكثر نموذجية، يمكن القول إن برامج التشخيص والتنبؤ بالحالة تقترب من المعيار الأفضل من المعيار البشري، بحيث توجد في الوقت الحالي أنظمة ذكاء صناعي يمكنها اكتشاف مرض الزهايمر المبكر بدقة تزيد عن 80% حتى عقد من الزمن قبل ظهور الأعراض لأول مرة، وهو إنجاز يتفوق بشكل كبير على أفضل أخصائي علم الأمراض البشري الذي يحاول أي شيء مشابه Amoroso et al. (2017).

تقول الدراسة [1] أنه لا ينبغي استخدام أدوات القرار الخوارزمية في المسائل الكبيرة أو القرارات الحرجة فيما يتعلق بالسلامة ما لم تكن الأنظمة المعنية أفضل من الإنسان في المجال ذي الصلة أو المجال الفرعي لصنع القرار. أوصت الدراسة بثلاث استراتيجيات لمعالجة مشكلة التحكم، وأكثرها واعدة تتضمن تعاوناً بين أدوات خوارزمية عالية الكفاءة وأشخاص بشريين. كما حددت الدراسة ستة مبادئ رئيسية يجب أن تعكسها جميع أنظمة الإنسان والآلة في تصميمها.

3- عملية صنع القرار:

يشير مصطلح صنع القرار إلى اتخاذ الخيارات بين مسارات العمل البديلة - والتي قد تشمل أيضاً عدم اتخاذ أي إجراء. يستخدم الأفراد في جميع المؤسسات المعلومات التي يجمعونها لاتخاذ مجموعة واسعة من القرارات، وقد تؤثر هذه القرارات على حياة الآخرين وتغير مسار العمل. مثلاً، أدت القرارات التي اتخذها التنفيذيون والشركات الاستشارية لشركة Enron في النهاية إلى خسارة 60 مليار دولار للمستثمرين، وآلاف الموظفين بدون وظائف، وخسارة جميع صناديق التقاعد للموظفين.

أنواع القرارات:

قرارات مبرمجة: هناك قرارات بسيطة واعتيادية مثل ماذا ترتدي، وماذا تأكل، والطريق الذي تسلكه وأنت تذهب من وإلى المنزل. ربما لا تقضي الكثير من الوقت في هذه القرارات ويسمى هذا النوع من القرارات المباشرة بالقرارات المبرمجة، وهي قرارات تحدث بشكل متكرر بما يكفي لتطوير استجابة آلية لها. تسمى الاستجابة الآلية التي نستخدمها لاتخاذ هذه القرارات بقاعدة القرار.

قرارات غير مبرمجة: تتطلب القرارات الفريدة والمهمة التفكير الواعي وجمع المعلومات والنظر بعناية في البدائل، وتسمى هذه القرارات قرارات غير مبرمجة.

يعد اتخاذ القرارات الاستراتيجية والتكتيكية والتشغيلية جزءاً من وظيفة التخطيط في نموذج P-O-L-C (التخطيط / Planning - التنظيم / Organizing - القيادة / Leading - التحكم / Controlling).

نتخذ مئات القرارات كل يوم. قد يبدو اتخاذ القرار وكأنه مهمة بسيطة، إلا أن هناك عوامل رئيسية نحتاج إلى أخذها في عين الاعتبار للتأكد من أن القرار النهائي يؤدي إلى النتيجة التي نريدها.

عملية صنع القرار بالشكل الأبسط هي:

- الاعتراف بضرورة اتخاذ القرار: فهم لماذا نتخذ القرار، هل نفهم فعلياً المشكلة المطروحة أم هناك حاجة للحصول على مزيد من المعلومات؟

- تحديد جميع البدائل المختلفة المتاحة: محاولة العثور على مجموعة من الحلول للمشكلة حتى يتم مراجعتها واختيار الأفضل منها. وفقاً للطبيعة البشرية، فإنه يتم القفز إلى الحل الأول الذي نفكر فيه فوراً، ولكن هذا ليس دائماً المسار الأكثر إنتاجية للعمل.

- التعرف على البدائل وتقييم النتائج واختيار أفضل خيار: مراجعة كل بديل تم اكتشافه وتقييم ما إذا كان تنفيذ هذه الأفكار سيكون له تأثير إيجابي أو سلبي على الأهداف النهائية.

- تنفيذ القرار وتقييم النتيجة: نتأكد من تقييم النتيجة للتأكد من أن القرار الذي تم اتخاذه كان الاختيار الصحيح. لا تتردد في إجراء تغييرات إذا رأيت مجالاً للتحسين.

القرارات التي نتخذها لا تتطوي دائماً على مشاكل. فقد تأتي بعض مواقف اتخاذ القرار كنتيجة للفرصة، حيث تدرك أن لديك فرصة لإجراء تغيير ويجب عليك تحديد أفضل طريقة للاستفادة منها.

4- التعلم الجماعي:

ظهر مؤخراً التعلم الجماعي أو التعلم القائم على المجموعات أو تعلم أنظمة التصنيف المتعددة وبتعريف هذا التعلم فهو مفهوم استخدام نماذج عصبية أو نماذج تدريب متعددة معاً لإنشاء خوارزمية أكثر دقة من الخوارزمية الأصلية التي تعمل بمفردها. ويعتبر أحد الحلول الملائمة لبناء نماذج تصنيف قوية ودقيقة [2]. فبدلاً من استخدام خوارزمية واحدة لبناء نموذج، يتم استخدام طرق المجموعات بشكل واسع لبناء مجموعة من المصنفات والجمع بينها، سواء كانت هذه المصنفات ضعيفة أو قوية، وتسمى هذه المجموعة

مجموعة المتعلمين الأساسيين. تتعلق النقاط الأساسية لفعالية طريقة التجميع بالاختيار الدقيق لكل من المتعلمين الأساسيين [3] وطريقة الدمج لإنتاج النتيجة النهائية. من الطرق الجماعية الشهيرة وشائعة الاستخدام، نذكر:

- حساب المتوسط (البسيط أو الموزون)
- التصويت (الأغلبية، الإجماع، التعددية، أو حتى الأصوات المرجحة)

في [4]، تم اقتراح مخطط تدريب مشترك قائم على مجموعة التصويت فيما يتعلق بمشاكل التصنيف الثنائي. الأساس في الخوارزمية المقترحة هو بناء متعلم مجموعة يختار من بين خمس مصنفات مختلفة، بناءً على وظيفة موضوعية جديدة تقيس فعالية أي زوج من المصنفات لكل مجموعة بيانات، تعمل بموجب نظام التصويت. يتم التحقق من فعالية الآلية المقترحة لاختيار المشاركين في المجموعة من خلال العديد من التجارب - من خلال مخطط التدريب الذاتي المعروف - ومخطط التدريب المشترك، حيث تم تطبيق الأزواج العشرة المختلفة من الخوارزميات في نفس متعلم التصويت والمصنفات الفردية الخمسة، على عدد كبير من مجموعات البيانات المعيارية عبر خمس قيم نسبية معنونة منفصلة. تم تأكيد النتائج التي تم الحصول عليها فيما يتعلق بمقياسين تصنيفين معروفين إحصائياً.

تم ملاحظة أن عدداً كبيراً من القرارات التي تم تصنيفها بشكل خاطئ في البداية بناءً على التنبؤات الفردية تم إرجاعها نحو التسمية الصحيحة. بينما في الوقت نفسه، لم تتأثر العديد من الحالات التي يختلف فيها المشتركان حول التسمية الثنائية. يحدث هذا لأن قيمة ثقة كبيرة وصحيحة مقترنة بقيمة أصغر غير صحيحة تظل كما هي في ظل نظام التصويت المقترح.

تم في النهاية اقتراح تطبيق اعتماد مراحل معالجة مسبقة أكثر تخصصاً موجهة نحو مشاكل أكثر تحديداً لتعزيز أداء الاستراتيجية المقترحة في هذه الورقة، ومع ذلك فقد سجلت الخوارزمية المقترحة أداءً قوياً ودقيقاً. كما أوصت الدراسة بإمكانية استخدام الشبكات العصبية العميقة (DNNs) [5] لتعزيز أداء التنبؤ، مع تغذية إما بالبيانات

الأولية أو أنواع عامة أخرى من مجموعات البيانات. فقد أثبتت شبكات الذاكرة طويلة المدى (LSTM) فعاليتها بدرجة كافية عند دمجها مع طرق SSL لبناء أنظمة دعم القرار الطبي [3].

في البحث [6] [7] ، تم اقتراح خوارزمية هجينة للتعلم الآلي تتكون من أشجار القرار والشبكات العصبية التي يمكنها تصنيف البيانات بكفاءة وفعالية كبيرة. كان هيكل الخوارزمية الهجينة معتمداً على شجرة قرار ثنائية بحيث كل عقدة في الشجرة عبارة عن شبكة عصبية مدربة على تصنيف فئة معينة من المخرجات باستخدام التصنيف الثنائي. البيانات التي تم استخدامها في هذه الدراسة لتدريب واختبار قدرة الخوارزمية كانت مجموعة بيانات لصيانة طائرة بوينج 737 التابعة لإدارة الطيران الفيدرالية (FAA) والتي تكونت من 137236 سجلاً مميزاً يتكون كل منها من 72 بارامتر. تم العمل بذلك عن طريق تصنيف التناقض أو سبب الحادث فيما إذا كان الحادث قد وقع أثناء عمليات الصيانة المجدولة أم لا ثم تصنيف التفاصيل المحددة المتعلقة بالحادث. أشارت النتائج إلى أن الخوارزمية الهجينة قادرة على تصنيف الحوادث بشكل فعال بدقة عالية. أشارت الدراسة أيضاً إلى قابلية تطبيق الخوارزمية في سيناريوهات العالم الحقيقي بحيث عرضت فوائد الجمع بين أشجار القرار والشبكات العصبية بدلاً من استخدامها بشكل فردي كونها أعطت تنبؤ أكثر دقة عند مقارنتها باستخدام نموذج عصبي واحد فقط للتنبؤات.

كما تم اقتراح نموذج هجين أيضاً في الدراسة [8] بحيث دمج نموذج الانتروبي الأقصى، وآلة نقل الدعم (Support Vector Machine)، وناظرة بايز لتلخيص نص متعدد المستندات لتحسين دقة التصنيف.

اقترح الباحثون في الدراسة [9] أيضاً نموذجاً هجيناً للتعلم الآلي للمشكلات متعددة التصنيف. تتكون الطريقة من مصنف شجرة القرار C4.5 ونهج واحد مقابل الكل. تم عرض فعالية الطريقة الهجينة على مجموعات البيانات مفتوحة المصدر لتجزئة الصور والأمراض الجلدية والتصوير للمفاوي.

في الدراسة [10] جمعت شبكة عصبية مع خوارزمية جينية لزيادة أداء الشبكات العصبية وتم تطبيق الطريقة الهجينة على مجموعة بيانات أمراض القلب.

5- دراسة حول الاستخدام الحالي للخوارزميات والتعلم الآلي في صنع القرار:

في طريقة المجموعة، تم دمج الشبكة العصبية (MLP) مع AdaBoost و Bagging و Dagging و MultiBoost و Rotation Forest و Random SubSpace. كما تم تطوير نموذج تنبؤي هجين آخر بواسطة Chen et al [11]. وفي هذه الطريقة، تم دمج مجموعات K-mean مع شجرة قرار J48 لتشخيص مرض السكري من النوع 2.

كما استخدمت دراسة أخرى مبنية على التنبؤ أسلوب التعلم الجماعي. اقترح المؤلفون نهج التعلم الجماعي المكون من طبقتين (TLEL) بناءً على أشجار القرار [12]. تستخدم الطبقة الخارجية نماذج Random Forest مختلفة للتدريب بينما الطبقة الداخلية عبارة عن تكامل لشجرة القرار والتعبئة لبناء نموذج Random Forest.

استخدمت دراسة أخرى [13] التعلم الجماعي من أجل أداء تنبؤي أفضل للحياة المتبقية (RUL) لمحركات الطائرات واستخدمت عدة طرق مثل المتعلمين الأساسيين المتعددين، بما في ذلك الغابات العشوائية (RFs) والشبكات العصبية التكرارية (RNN)، نظام الاستدلال الضبابي القائم على الشبكة التكيفية (ANFIS) كما استخدمت الطريقة أيضاً تحسين سرب الجسيمات (PSO) والتحسين التريبيعي المتسلسل (SQP) لتحقيق أفضل مزيج من الأوزان لاستخدامها في المتعلمين الأساسيين.

كما كان هناك امتداد للدراسة السابقة بواسطة Li et al [14]. واستخدم في الدراسة البيان غير الدوري الموجه (DAG) المهجن بذاكرة طويلة المدى (LSTM) وشبكة عصبية تلافيفية (CNN) للتنبؤ بـ RUL. قام باختبار الطريقة باستخدام مجموعة بيانات توفرها وكالة ناسا.

6- خوارزميات تعلم المجموعات شائعة الاستخدام

/Bagging /التعبئة:

هي واحدة من أقدم الخوارزميات المبنية على المجموعات وأكثرها سهولة وربما أبسطها، مع أداء جيد بشكل مدهش (Breiman 1996)[15]. يتم الحصول على تنوع المصنفات في التعبئة باستخدام نسخ متماثلة لبيانات التدريب. بحيث يتم الحصول على مجموعات فرعية مختلفة من بيانات التدريب بشكل عشوائي (مع الاستبدال) من مجموعة بيانات التدريب بأكملها. يتم استخدام كل مجموعة فرعية من بيانات التدريب لتدريب مصنف مختلف من نفس النوع. يتم بعد ذلك دمج المصنفات الفردية عن طريق التصويت بالأغلبية البسيطة لقراراتهم. بالنسبة لأي حالة، فإن التصنيف الذي يختاره أكبر عدد من المصنفات هو قرار المجموعة.

:Boosting

يتم هنا إنشاء مجموعة من المصنفات من خلال إعادة تجميع البيانات، والتي يتم دمجها بعد ذلك عن طريق التصويت بالأغلبية. بحيث يتم توجيه عملية إعادة التشكيل بشكل استراتيجي لتوفير بيانات التدريب الأكثر إفادة لكل مصنف متتالي. كل تكرار للتعزيز يقوم بإنشاء ثلاث مصنفات ضعيفة: المصنف الأول C1 يتم تدريبه بمجموعة فرعية عشوائية من بيانات التدريب المتاحة. يتم اختيار المجموعة الفرعية لبيانات التدريب للمصنف الثاني C2 على أنها المجموعة الفرعية الأكثر فائدة، بالاعتماد على C1. بحيث يتم تدريب C2 على بيانات التدريب التي تم تصنيف نصفها فقط بشكل صحيح بواسطة C1، والنصف الآخر مصنف بشكل خاطئ. يتم تدريب المصنف الثالث C3 مع الحالات التي يختلف عليها C1 وC2. وفي النهاية يتم الجمع بين المصنفات الثلاثة من خلال تصويت الأغلبية الثلاثية.

:AdaBoost

يمكن القول إن AdaBoost (Adaptive Boosting) هي من أكثر الخوارزميات القائمة على المجموعات شهرةً ولها العديد من الأنواع، مثل AdaBoost.M1 لمشاكل التصنيف، AdaBoost.M2 للمصنفات الضعيفة، AdaBoost.R لمشاكل regression.

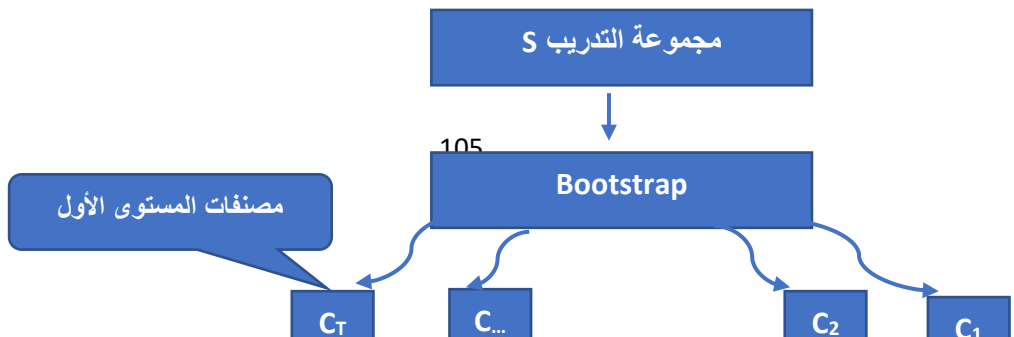
التعميم المكس / Stacked Generalization

في التعميم المكس λ Wolpert (أو التكديس)، يتم تدريب مجموعة من المصنفات أولاً باستخدام عينات من بيانات التدريب، وإنشاء مصنفات المستوى الأول، والتي تُستخدم مخرجاتها بعد ذلك لتدريب مصنفات الطبقة الثانية (Wolpert 1992) [16]. الفكرة الأساسية هي معرفة ما إذا تم تعلم بيانات التدريب بشكل صحيح. على سبيل المثال، إذا تعلم مصنف معين بشكل غير صحيح منطقة معينة من مساحة الميزة، وبالتالي أخطأ باستمرار في تصنيف الحالات القادمة من تلك المنطقة، فقد يكون مصنف المستوى الثاني قادراً على تعلم هذا السلوك بالتعاون مع السلوكيات المكتسبة للمصنفات الأخرى، وبالتالي يمكنه تصحيح مثل هذا التدريب غير المناسب. يتم تقسيم مجموعة بيانات التدريب بأكملها إلى T كتلة، ويتم تدريب كل مصنف من المستوى الأول بدايةً على مجموعة مختلفة من $T-1$ كتلة من بيانات التدريب. ثم يتم تقييم كل مصنف على الكتلة T (الاختبار الزائف)، ولا يتم رؤيته أثناء التدريب. تشكل مخرجات هذه المصنفات على كتل التدريب الزائفة مع الـ Labels الصحيحة الفعلية لتلك الكتل مجموعة بيانات التدريب لمصنف المستوى الثاني.

يمكننا تخيل إجراء التكديس على أنه يحتوي على مستويين: المستوى الأول والمستوى الثاني:

- المستوى الأول: بيانات هذا المستوى هي مدخلات مجموعة بيانات التدريب وتتعلم نماذج هذا المستوى القيام بتنبؤات من هذه البيانات.
- المستوى الثاني: بيانات المستوى الثاني تأخذ مخرجات نماذج المستوى الأول كمدخلات، ويقوم النموذج الوحيد في هذا المستوى (أو المتعلم الفوقي / meta-learner) بتعلم عمل تنبؤات من هذه البيانات.

نوضح في الشكل الآتي التعميم المكس:



الشكل (1) توضيح التعميم المكّس

مجموعة التعميم المكّس (Stacked Generalization Ensemble):

المجمّع هو نموذج يحسب المتوسط للتنبؤات من عدة نماذج مدربة، ويتمثل أحد قيود هذا النهج هو أن كل نموذج يساهم بنفس المقدار في تنبؤ المجموعة، بغض النظر عن مدى جودة أداء النموذج.

أحد أشكال هذا النهج يُطلق عليه اسم مجموعة المتوسط الموزون (Weighted Average Ensemble) يقوم بوزن مساهمة كل عضو في المجموعة من خلال الثقة أو الأداء المتوقع للنموذج على عينة بيانات محددة. وهذا يسمح للنماذج ذات الأداء الجيد بالمساهمة أكثر من النماذج ذات الأداء الأقل. يوفر نموذج مجموعة المتوسط الموزون تحسناً على نموذج مجموعة المتوسط العادي.

التحسين الآخر الذي يمكن تقديمه لهذا النهج هو استبدال نموذج المجموع الموزون الخطي (مثل الانحدار الخطي) المستخدم لدمج تنبؤات النماذج الفرعية بأي خوارزمية تعلم. وهذا النهج يسمى التعميم المكّس (التكديس).

في التكديس، تأخذ خوارزمية التعلم مخرجات النماذج الفرعية كمدخلات وتحاول تعلم أفضل طريقة تجمع بين تنبؤات الإدخال للحصول على تنبؤ خرج أفضل ما يمكن.

7- أهمية نموذج التعلم الجماعي:

يكون عادةً نموذج التعلم الجماعي أكثر تعقيداً من استخدام نموذج واحد للتعلم الآلي، وقد يتطلب هذا أسلوباً أكثر تعقيداً لإعداد النموذج، وموارد حسابية أكثر لتدريب النموذج، ومع ذلك، يمكن القول إنه لدينا سببان لاستخدام خوارزميات تعلم المجموعات، وهما:

متانة أفضل: تقدم العديد من خوارزميات التعلم الآلي تنبؤات مختلفة في كل مرة يتم فيها تدريب النموذج على نفس البيانات أو حتى على بيانات مختلفة قليلاً. يشار إلى هذا باسم التباين في التنبؤات أو استقرار النموذج. يمكن أن تعطي خوارزميات التعلم الجماعي تنبؤات أكثر استقراراً من نموذج واحد.

تنبؤات أفضل: هناك بعض المجالات ومهام التنبؤ يكون الأداء التنبؤي هو الأكثر أهمية من أي شيء آخر. وهذا هو الحال بشكل خاص في التعلم الآلي التنافسي حيث تكون المهارة التنبؤية هي النتيجة الوحيدة التي يتم أخذها في الاعتبار وهذا ما توفره خوارزميات التعلم الجماعي فهي توفر مهارة تنبؤية أفضل من استخدام نموذج واحد.

تُستخدم المجموعات بشكل شائع لتثبيت التنبؤات التي قدمتها نماذج التعلم العميق بالنظر إلى كل من الطبيعة العشوائية لبنية النموذج التي تمت تهيئتها باستخدام أوزان عشوائية.

يشير تنوع المجموعة إلى الاختلافات في التنبؤات أو أخطاء التنبؤ التي تحدثها النماذج التي تساهم في المجموعة. ومن الناحية المثالية، سيكون كل عضو في المجموعة مستقلاً، ما يعني أنه سيقدم تنبؤات بطريقة مختلفة تماماً عن جميع أعضاء المجموعة الآخرين إلا أن هذا غير ممكن لأن لدينا مجموعة بيانات تدريب واحدة فقط.

تشير طرق التجميع إلى الطريقة المحددة التي يتم بها استخدام الإخراج من كل عضو في المجموعة لإجراء التنبؤ النهائي. يمكن أن يتضمن ذلك استخدام طريقة إحصائية بسيطة كوسيلة للتصويت، على الرغم من أنه قد يتضمن وزن التنبؤات بمهارة النموذج المساهم أو حتى استخدام نموذج آخر لمعرفة أفضل طريقة للجمع بين التنبؤات.

8- الدراسة العملية:

تتكون المهمة المختارة من تصنيف صور شعاعية طبية لسبعة أنواع مفاصل وعظام في الجسم إلى فئتين. يرجع اختيار هذه المهمة بشكل أساسي إلى توفر هذه المجموعة من البيانات، بالإضافة إلى مجال المشكلة، وهو عام بما يكفي ليكون مناسباً للتطبيق المطلوب.

حسب الاحصائيات الأخيرة المتوفرة في مجال الحالات المتعلقة بمجال بحثنا (صور شعاعية) فإن أمراض العضلات والعظام تؤثر على أكثر من 1.7 مليار شخص في جميع أنحاء العالم، وهي السبب الأكثر شيوعاً للألم والإعاقة الشديدة وطويلة الأجل، بحيث تزداد 30 مليون زيارة لقسم الطوارئ سنوياً وتتزايد بسبب هذا النوع من الحالات [17][18].

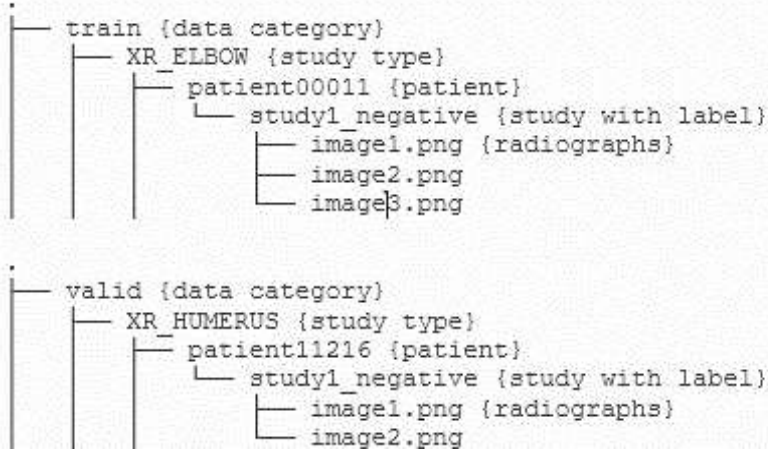
مهمة الكشف عن المرض، أو بمعنى آخر تحديد ما إذا كانت دراسة التصوير الشعاعي طبيعية أو غير طبيعية، هي عملية إشعاعية مهمة بحيث تعتبر دراسة تفسر على أنها تستبعد المرض بشكل طبيعي ويمكن أن تلغي حاجة المرضى للخضوع لمزيد من الإجراءات التشخيصية أو التدخلات.

9- عينة البيانات المدروسة:

نعمل من خلال بحثنا على MURA Dataset، وهي مجموعة بيانات كبيرة للكشف عن التشوهات في التصوير الشعاعي العضلي الهيكلي تتألف من 14863 دراسة (وتحتوي كل دراسة على واحد أو أكثر من المشاهدات (الصور)) من 12173 مريضاً، بإجمالي 40,561 صورة إشعاعية متعددة الرؤية. تحتوي مجموعة البيانات على 9,045 دراسة طبيعية و 5818 دراسات إشعاعية عضلية هيكلية غير طبيعية [19].

ينتمي كل منها إلى واحد من سبعة أنواع من الدراسات التصويرية للأطراف العلوية: الكوع، الإصبع، الساعد، العضد، الكتف والمعصم واليد. تم تصنيف كل دراسة يدوياً على أنها طبيعية أو غير طبيعية من قبل أطباء الأشعة المعتمدين من مستشفى ستانفورد في وقت التفسير الشعاعي السريري في بيئة الأشعة التشخيصية بين 2001 و 2012.

أغلب الدراسات التي تمت على نفس عينة البيانات تقسم الصور إلى 36808 صورة تدريبية و3197 صورة للتحقق منها. بحيث تأتي عينة البيانات كمجموعة بيانات منقسمة إلى مجموعتين: مجموعة التدريب ومجموعة التحقق في مجلدات منفصلة. كل مجموعة من مجموعات البيانات (التدريب والتحقق) لها مجلدات فرعية.



الشكل (2): طريقة تقسيم عينة البيانات المدروسة

كما نرى من الشكل أنه يتم تقسيم البيانات ضمن المجموعتين المذكورتين سابقاً (في مجلد منفصل لكل مجموعة منها)، كما أنه يوجد مجلد خاص بكل فئة من الفئات السبعة والمذكورة أيضاً، ضمن كل فئة يوجد مجلد خاص لكل مريض (حالة دراسة مستقلة) ويتم تحديد حالة الدراسة مع وضع الـ label أو التصنيف لها، بحيث أن التصنيفين المحددين هما positive وnegative، وضمن المجلد الخاص بحالة الدراسة هناك مجموعة من المشاهدات (الصور).

تقسيم البيانات وزيادتها:

كما ذكرنا سابقاً عن البيانات المدروسة وتصنيفها إلى فئتين وفقاً للتسميات المقدمة من اختصاصي الأشعة. قمنا بتقسيم 80% بشكل عشوائي في مجموعة التدريب (Train Set) من مجموعة البيانات بالكامل، بينما بقي الجزء الباقي 20% في مجموعة الاختبار (Test Set).

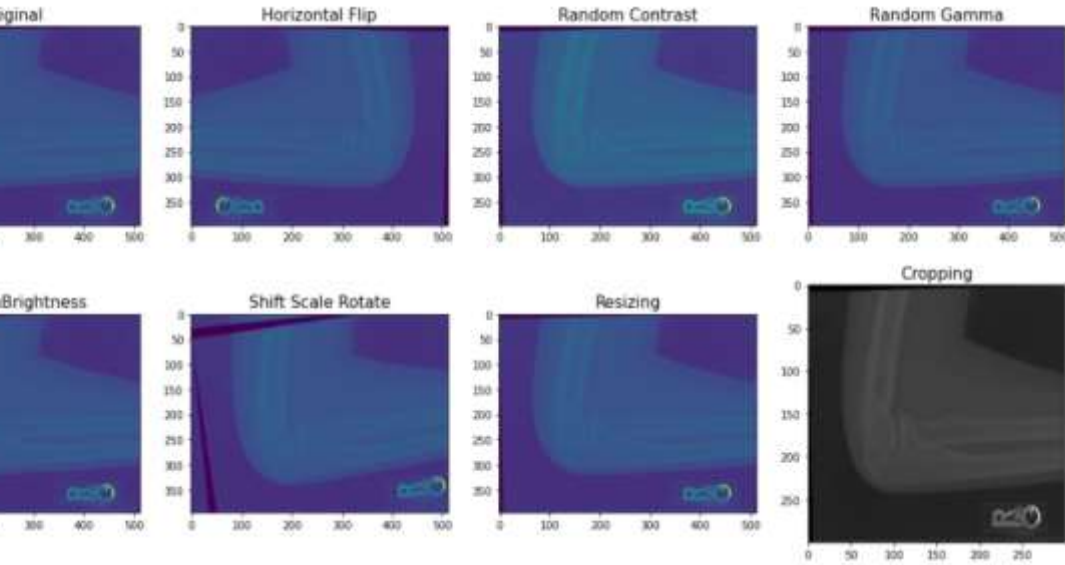
تم تقييم أداء المصنف (النموذج) وفقاً لمتوسط دقة التصنيف الناتج، كما تم الاعتماد على مقياس كوهين كبا كمقياس جديد في عملية التصنيف (هناك شرح له في قسم النتائج والتقييم) وذلك فترة تشغيل تبلغ 43 تكرار.

المعالجة المسبقة للبيانات:

قمنا بمعالجة كل صورة بحيث يكون لها نفس المتوسط الانحراف المعياري للصورة في مجموعة التدريب ImageNet ، تم استخدام صور متغيرة الحجم وتحويلها لحجم ثابت $224 \times 224 \times 3$. كما قمنا بزيادة البيانات أثناء التدريب عن طريق تطبيق بعض التحويلات باستخدام ImageDataGenerator, وما قمنا بتطبيقه فعلياً هو (Horizontal Flip, Random Contrast, Random Gamma, Random Brightness, Shift Scale Rotate, Resizing, Cropping)

كما اعتمدنا على Batch Size = 32

يوضح الشكل الآتي مثال على نتائج التحويلات المطبقة على إحدى الصور:



الشكل (3): مثال على نتيجة تطبيق تحويلات المعالجة المسبقة للبيانات على صورة من عينة البيانات (Data Augmentation)

مسابقة MURA

نشرت مجموعة التعلم الآلي في جامعة ستانفورد ورقة بحثية [19] تتعلق بمشكلة تصنيف الصور الشعاعية الطبية وقدمت واحدة من أكبر مجموعة بيانات صور شعاعية عامة في العالم تسمى MURA. MURA هي اختصار للتصوير الشعاعي العضلي الهيكلي (Musculoskeletal Radiographs). استخدمت مجموعة ML بجامعة ستانفورد خوارزمية DenseNet169 لتدريب شبكة عصبية عميقة يمكنها الكشف عن التشوهات في الصور الشعاعية بدقة أقرب إلى كبار أطباء الأشعة. وكان الهدف من مسابقة MURA هو تصنيف كل دراسة إلى طبيعية أو غير طبيعية (تنبؤات ثنائية)، وليس كل صورة. وكان أفضل أداء في جامعة ستانفورد. في مجموعة البيانات هذه، قام الفريق الذي نشرها بتدريب شبكة تلافيفية متصلة بكثافة DenseNet169 مكونة من 169 طبقة للكشف عن التشوهات وتحديد مكانها.

10- مواد وطرق البحث:

قمنا بتطبيق ثلاثة من نماذج شبكة Resnet المتوفرة حتى تاريخ إعداد هذا البحث من خلال تدريب نموذج لكل منها ثم قمنا بمقارنة أداء هذه النماذج. نماذج الشبكات التي قمنا بتجريبها هي ثلاثة نماذج، وهي ResNet152V2, InceptionV3, InceptionResNetV2. نود أن يحقق نموذجنا أداءً مشابهاً لأداء أطباء الأشعة. بالنسبة للدراسات التي أجريت على اليد والرسغ والمعصم، فإن نتائج النموذج الخاصة بنا أعلى قليلاً من ناحية الدقة من تلك النتائج الخاصة بأطباء الأشعة؛ وفي دراسات الكوع والساعد والعضد والكتف، كانت الدقة الخاصة بنموذجنا أقل قليلاً من درجات أخصائي الأشعة، ولكنها ليست ذات دلالة إحصائية، مما يشير ويؤكد إلى أن مجموعة البيانات تمثل مشكلة تحدي جيدة للبحث في المستقبل.

:Residual Networks (ResNet)

بعد أول بنية قائمة على CNN (AlexNet) والتي فازت بمسابقة ImageNet 2012، تستخدم كل بنية فائزة لاحقة المزيد من الطبقات في شبكة عصبية عميقة لتقليل معدل الخطأ، ويعمل هذا مع عدد أقل من الطبقات، ولكن عندما نزيد عدد الطبقات، توجد مشكلة شائعة في التعلم العميق مرتبطة بما يسمى Vanishing/Exploding gradient. يؤدي هذا إلى أن يصبح التدرج / 0 gradient أو كبيراً جداً. وبالتالي عندما نزيد عدد الطبقات، يزداد أيضاً معدل خطأ التدريب والاختبار.

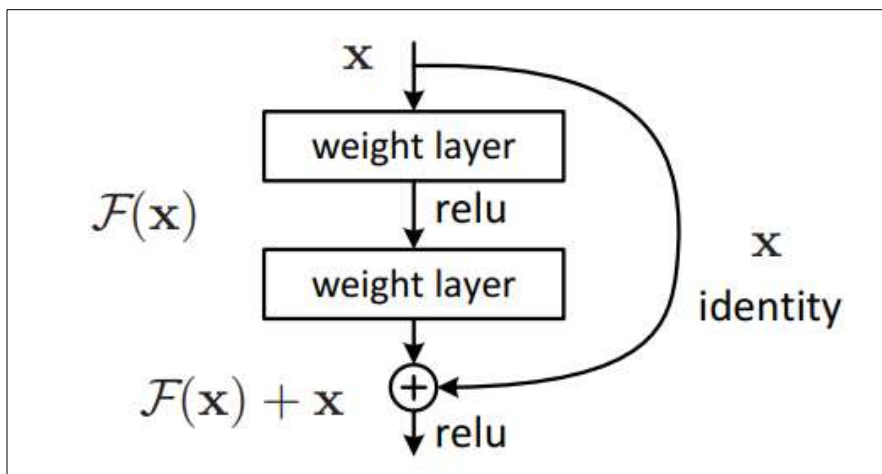
تم تقديم ResNet والتي اقترحها باحثون في Microsoft Research في عام 2015، كبنية جديدة تسمى الشبكة المتبقية (Residual Network).

الكتلة المتبقية (Residual Block):

من أجل حل مشكلة التدرج المتلاشي / المتفجر (vanishing/exploding gradient)، قدمت هذه البنية مفهوماً جديداً يسمى الشبكة المتبقية. في هذه الشبكة نستخدم تقنية تسمى تخطي الاتصالات. يتخطى الاتصال من خلال تخطي التدريب من عدة طبقات ويتصل مباشرة بطبقة الخرج.

النهج المستخدم في هذه الشبكة هو أنه بدلاً من أن تتعلم الطبقات التعيين الأساسي (mapping)، نسمح للشبكة بتناسب التعيين المتبقي.

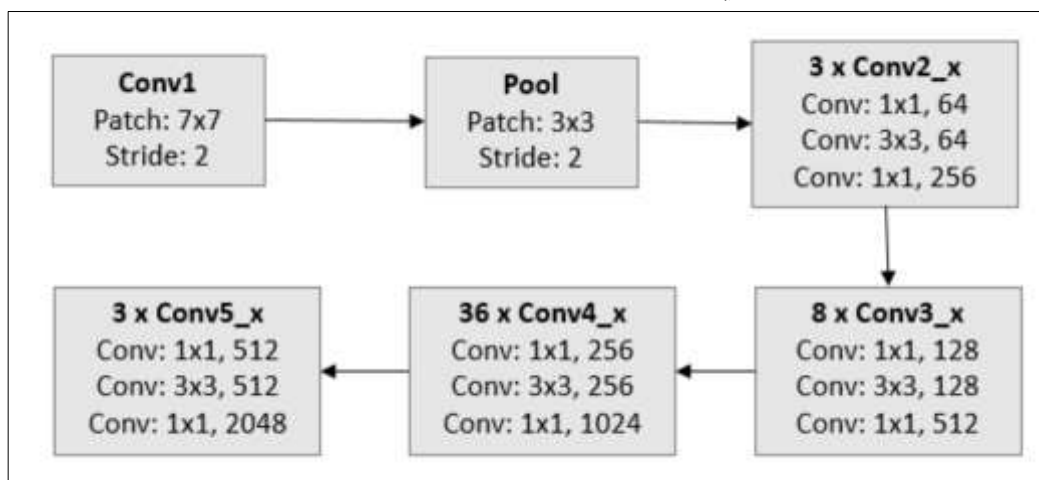
لذلك، بدلاً من قول $H(x)$ ، التعيين الأولي، تقوم الشبكة بما يلي $F(x) = H(x) - x$ والذي يعطي $H(x) = F(x) + x$. والشكل الآتي يوضح بنية الكتلة المتبقية (Residual Block)



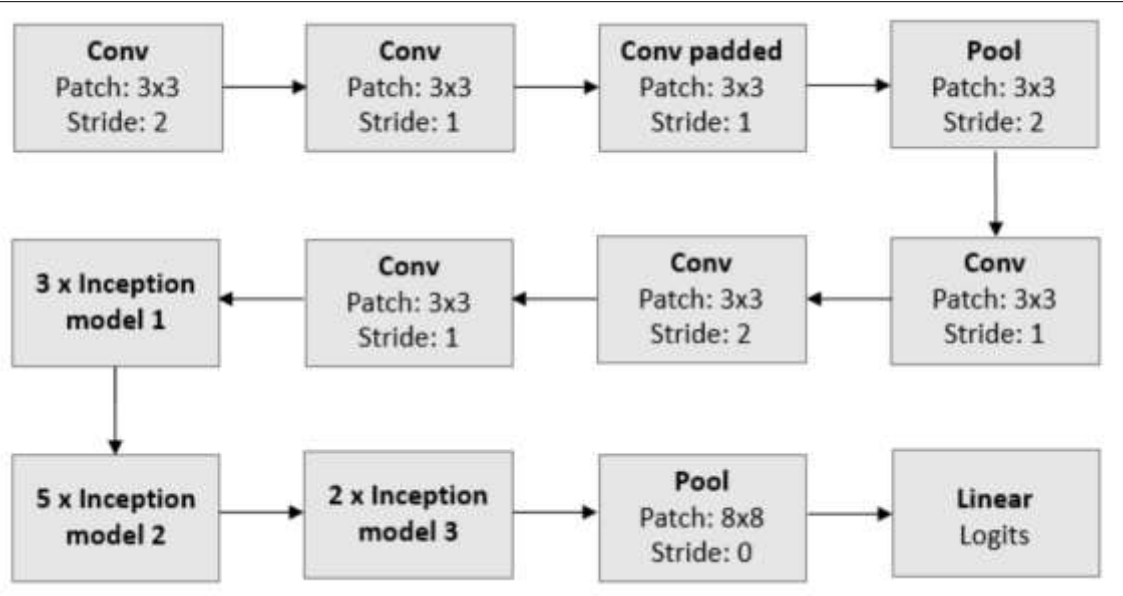
الشكل (4): بنية الكتلة المتبقية (Residual Block)

بنية الشبكة:

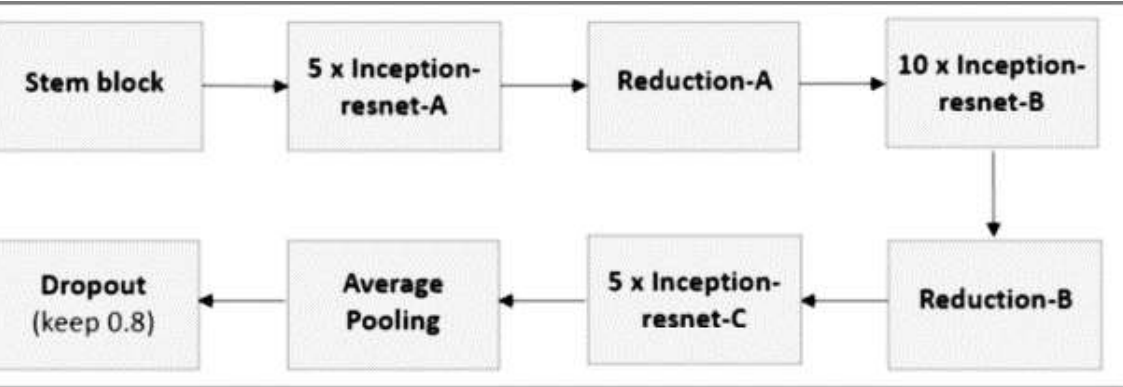
تستخدم هذه الشبكة بنية شبكة عادية مكونة من عدد من الطبقات مستوحاة من VGG-19 حيث يتم إضافة اتصال الاختصار (shortcut connection) بحيث تقوم اتصالات الاختصار هذه بتحويل البنية إلى شبكة متبقية.



الشكل (5): بنية شبكة Resnet152



الشكل (6): بنية شبكة InceptionV3



الشكل (6): بنية شبكة Inception-Resnet-v2

التدريب والاختبار العملي:

بيئة التكوين وإعداد البارامترات:

تم تطوير جميع التجارب وتنفيذها على منصة Google Colaboratory السحابية باستخدام Keras مع TensorFlow, كما تم التعامل مع منصة Kaggle لما تقدمه من ميزات للتعامل مع نماذج الشبكات العصبية أيضاً وبسبب القيود التي بدأت

تقرضها منصة Google Colab من حيث محدودية الموارد المتاحة وتحديد موارد الجلسة الواحدة وتقييد مدتها.

Colaboratory أو Colab ، هو أحد منتجات Google Research ، والذي يسمح للمطورين بكتابة وتنفيذ كود Python من خلال متصفحهم. يعد Google Colab أداة ممتازة لمهام التعلم العميق. كما يعتبر فعلياً كإصدار متخصص من Jupyter Notebook، والذي يعمل على السحابة ويوفر موارد حوسبة مجانية مثل GPUs وTPUs.

هناك عدة أسباب لاختيار استخدام Google Colab بدلاً من Jupyter Notebook العادي، وأهمها:

- وجود مكتبات مثبتة مسبقاً (Keras, TensorFlow, and PyTorch مثلاً)
 - يتم الحفظ على السحابة (يتم حفظ جميع دفاتر Google Colab ضمن حساب Google Drive، تماماً مثل ملفات Google Docs و Google Sheets)
 - التعاون (وهي ميزة أخرى يقدمها Colab فإذا كنت تعمل مع مطورين متعددين في مشروع ما، فمن الأفضل استخدام Colab. تماماً مثل التعاون في مستند محرر مستندات Google، كما يمكن البرمجة المشتركة مع مطورين متعددين باستخدام Colab notebooks)
 - استخدام GPU وTPU مجاناً (تتيح Google Research استخدام وحدات معالجة الرسومات (GPU) و (TPU) المخصصة لمشاريع التعلم الآلي كون تسريع GPU وTPU يحدثان فرقاً كبيراً حتى في المشاريع الصغيرة)
- أما بالنسبة لمنصة Kaggle فهي تابعة لشركة Google LLC، وهي مجتمع عبر الإنترنت لعلماء البيانات وممارسي التعلم الآلي. يسمح Kaggle للمستخدمين بالعثور على مجموعات البيانات ونشرها، واستكشاف النماذج وإنشائها في بيئة علم البيانات المستندة إلى الويب، والعمل مع علماء البيانات الآخرين ومهندسي التعلم الآلي، والدخول في مسابقات لحل تحديات علوم البيانات.

وقد كانت بدايتها في عام 2010 من خلال تقديم مسابقات التعلم الآلي وتقدم الآن أيضاً منصة بيانات عامة وبيئة عمل قائمة على السحابة (cloud-based workbench) لعلوم البيانات وتعليم الذكاء الاصطناعي.

خدمات Kaggle (إضافةً إلى الخدمات التي يقدمها Colab):

- مسابقات التعلم الآلي: كانت أول منتج لشركة Kaggle, بحيث تشر الشركات المشكلات ويتنافس المتسابقون لبناء أفضل خوارزمية، عادةً تكون بجوائز نقدية قيّمة.
- Kaggle Kernels: وهي بيئة عمل قائمة على السحابة لعلوم البيانات والتعلم الآلي, تسمح لعلماء البيانات بمشاركة الكود والتحليل في Python و R و R Markdown, وقد تمت مشاركة أكثر من 150 ألف نواة على Kaggle تغطي عدة مشاريع ومواضيع هامة.
- منصة لمجموعات البيانات العامة: يشارك أعضاء مجتمع Kaggle مجموعات البيانات مع بعضهم البعض. ونجد على هذه المنصة مجموعات بيانات عن كل شيء.
- Kaggle Learn: وهي منصة لتعليم الذكاء الاصطناعي.

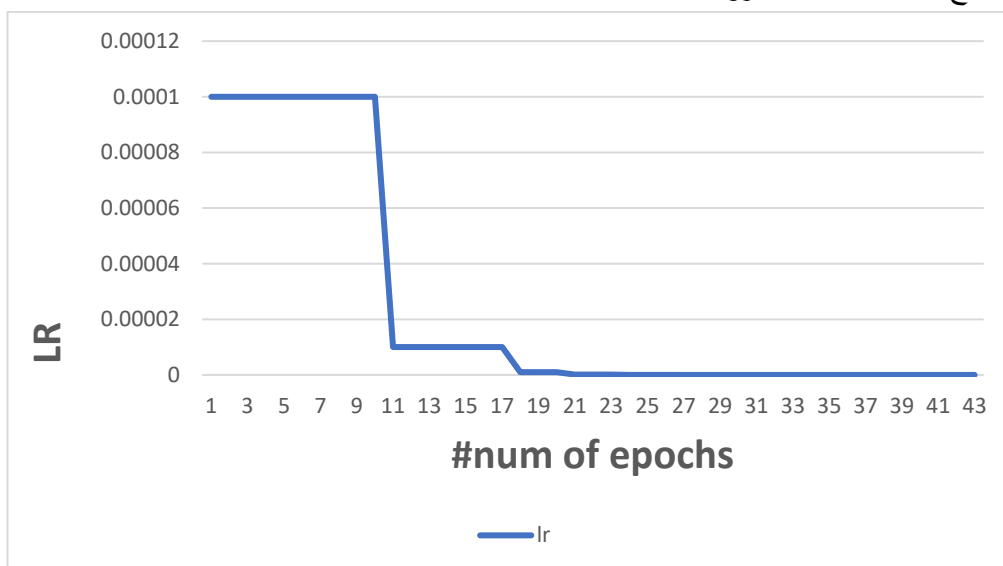
لمقارنة أداء الطرق المختلفة، قمنا بتعيين نفس البارامترات لكل الطرق المستخدمة، البارامترات هي: معدل التعلم الأولي 0.0001 ويتناقص تدريجياً ليصل إلى قيمة 0.000000001، حجم الـ Batch هو 32، وعدد الفترات القصوى هو 43 فترة (Epochs=43).

11-دراسة النتائج العملية

- مرحلة التدريب:

تمت تهيئة الشبكة بأوزان من نماذج الشبكات المختارة والتي تم تدريبها مسبقاً على عينة البيانات imagenet وتم تدريب الشبكة باستخدام Adam Optimizer مع

البارامترات الافتراضية. استخدمنا حجم batch = 32 ومعدل التعلم 0.0001 يتناقص تدريجياً مع مراحل التكرار (الشكل (7))، وقمنا بتدريب الشبكات لكل أنواع الدراسة بشكل منفصل بحيث قمنا بتطبيق نموذج الشبكة المحددة على عينة البيانات المدروسة على كل نوع بشكل مستقل (أي التطبيق على فئة المعصم بشكل مستقل، على فئة الأصابع بشكل مستقل... الخ) وبعدها مرات 43 لكل الأنواع، وقمنا بتسجيل النتائج لكل الحالات المدروسة.



الشكل (7): توضيح لقيم معدل التعلم (Learning Rate) خلال العمل على كل النماذج المدروسة

- تقييم الأداء:

قمنا بتقييم أداء النماذج المدروسة على الاختبار الذي تم تعيينه لكل أنواع الدراسة ومقارنتها معاً كما تمت مقارنتها بنموذج المجموعة (القسم التالي). تم إنشاء نموذج المجموعة باستخدام النماذج الثلاثة لنوع الدراسة (/ humerus العضد). تمت مقارنة الأداء باستخدام مقياس الدقة ومقياس كايا. لاحظنا أن أداء النماذج الفردية المدربة على نوع الدراسة المحددة أفضل بكثير مقارنةً بنموذج المجموعة.

- مقياس كوهين كابا (Cohen's kappa statistic)

المقياس المستخدم في مسابقة MURA ليس هو الدقة الكلاسيكية، ولكن درجة kappa أو مقياس كابا. وهو مقياس أكثر قوة من الدقة البسيطة، لأنه يأخذ في الاعتبار إمكانية حدوث الاتفاقية عن طريق الصدفة بطرحها من الاتفاقية الملاحظة.

ما هو كوهين كابا؟

يقيس كوهين كابا الاتفاق بين اثنين من المقيمين يصنف كل منهما عناصر N إلى فئات حصرية متبادلة. هناك طريقة بسيطة للتفكير في ذلك وهي أنه مقياس كمي للاعتمادية لاثنتين من المقيمين الذين يصنفون نفس الشيء، ويتم تصحيحه وفقاً لعدد المرات التي قد يتفق فيها المقيّمون عن طريق الصدفة.

تقييم كوهين كابا:

يمكن أن تكون قيمة kappa أقل من 0 (سالبة). تعني الدرجة 0 أن هناك اتفاقاً عشوائياً بين المقيمين، في حين أن الدرجة 1 تعني أن هناك اتفاقاً تاماً بين المقيمين. لذلك، فإن الدرجة الأقل من 0 تعني أن هناك اتفاقاً أقل من فرصة عشوائية.

		Rater 2	
		Correct	Incorrect
Rater 1	Correct	A	B
	Incorrect	C	D

الشكل (8): شبكة $N \times N$ المستخدمة لتفسير نتائج المقيمين

في الشكل تقسم الشبكة إلى الأجزاء الآتية:

A: إجمالي عدد الحالات التي قال كلا المقيمين إنها صحيحة، أي أن المقيمين متفقون.

B: العدد الإجمالي للحالات التي قال Rater 2 إنها غير صحيحة، لكن Rater 1 قال إنها صحيحة. وهذه تمثل حالة خلاف.

C: العدد الإجمالي للحالات التي قال Rater 1 إنها غير صحيحة ، لكن Rater 2 قال إنها صحيحة. وهذه تمثل أيضاً حالة خلاف.

D: إجمالي عدد الحالات التي قال كلا المقيمين إنها غير صحيحة، أي أن المقيمين متفقون.

من أجل إيجاد قيمة kappa، نحتاج أولاً إلى معرفة احتمالية الاتفاق (وهذا يفسر سبب تسليط الضوء على قطر الاتفاقية). بحيث يتم اشتقاق هذه الصيغة عن طريق إضافة عدد الاختبارات التي يتفق فيها المقيمون ثم تقسيمها على العدد الإجمالي للاختبارات. وفقاً للشكل، فهذا يعني $(A + D) / (A + B + C + D)$.

$$P_o = \text{Number in Agreement} / \text{Total}$$

والخطوة التالية هي إيجاد احتمالية الاتفاق العشوائي وهي إجمالي عدد المرات التي قال Rater 1 أنها صحيحة مقسومة على العدد الإجمالي للحالات، مضروبة في إجمالي عدد المرات التي قال Rater 2 أنها صحيحة مقسومة على العدد الإجمالي للحالات، المضافة إلى إجمالي عدد المرات التي قال Rater 1 أنها غير صحيحة مضروباً في إجمالي عدد المرات التي قال Rater 2 إنها غير صحيحة. هناك الكثير من المعلومات التي يجب أخذها هناك، ونختصرها بالمعادلة الآتية:

$$P_{(\text{correct})} = (A + B / A + B + C + D) * (A + C / A + B + C + D)$$

$$P_{(\text{incorrect})} = (C + D / A + B + C + D) * (B + D / A + B + C + D)$$

$$P_e = P_{(\text{correct})} + P_{(\text{incorrect})}$$

معادلة احتمالية الاتفاق العشوائي

وفي النهاية فإن صيغة Cohen's Kappa هي احتمال الاتفاق الذي يزيد احتمال الاتفاق العشوائي مقسوماً على 1 ناقصاً احتمال الاتفاق العشوائي.

$$K = P_o - P_e / 1 - P_e$$

صيغة معامل كبا كوهين

ملاحظات حول النتائج:

نوضح في الجداول أدناه نتائج المقارنة (من ناحية دقة التصنيف، مقياس كابا)، التي قمنا بإجرائها بين الشبكات الثلاث التي عملنا عليها وتم ذكرها سابقاً لتصنيف صور عينة البيانات المدروسة.

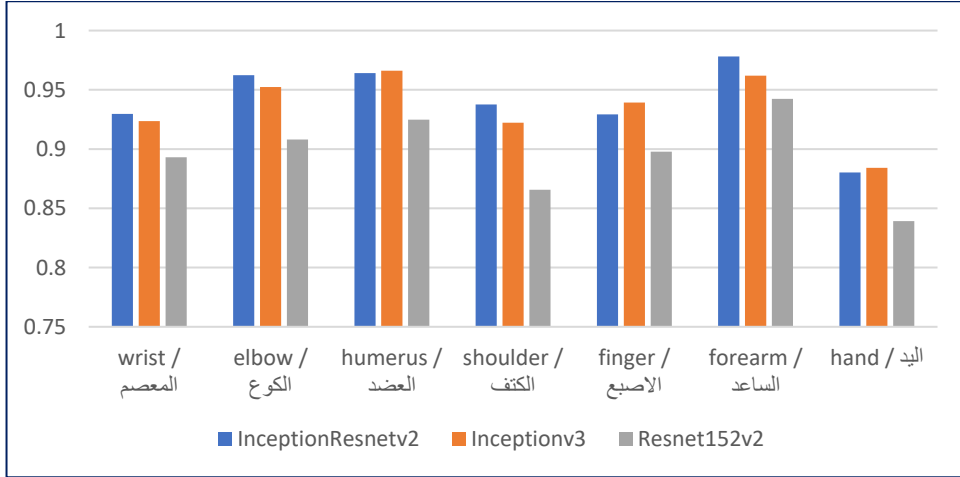
نلاحظ من الجدول الأول أننا حصلنا على أفضل دقة لأغلب (الكل ماعدا نوع المعصم) حالات الدراسة ضمن عينة البيانات من خلال نموذج الشبكة **Resnet152v2** بحيث وصلنا إلى دقة تصنيف تقريبا 88% في نوع الكوع والعضد وهي كانت الأفضل بين الأنواع الأخرى فقد وصلت دقة تصنيف نوع الكتف في هذا النموذج حوالي 78.5% ووصلت 79% في نوع الأصابع واقترب من 80% في تصنيف نوع الساعد وتجاوز هذه النسبة قليلاً في نوع اليد.

أما بالنسبة لنوع المعصم فقد تفوق أداء نموذج **Inceptionv3** على الأنواع الأخرى بدقة اقتربت من 86%.

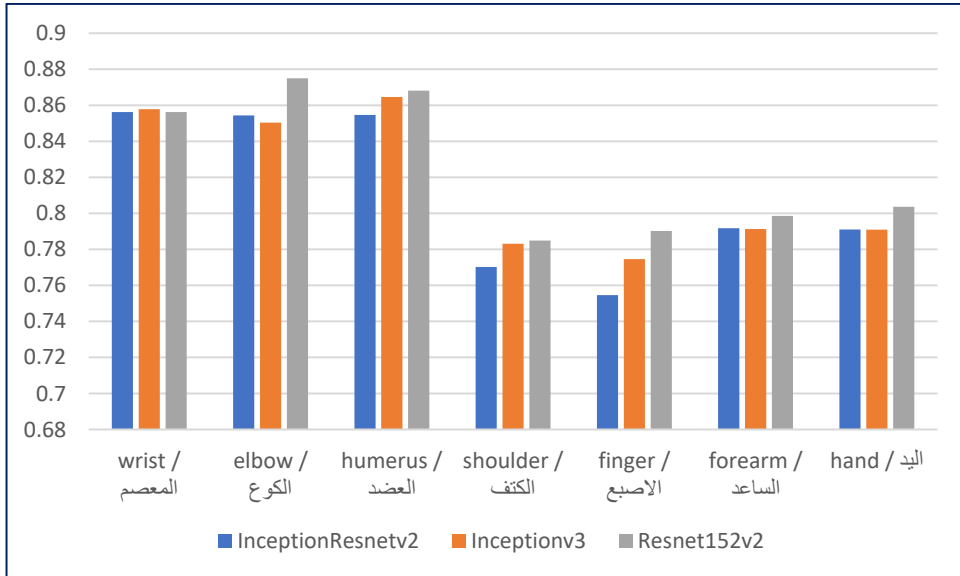
الجدول (1) نتائج مقارنة أداء الشبكات الثلاث المدروسة من ناحية دقة التصنيف

دقة التصنيف						نوع الدراسة
InceptionResnetv2		Inceptionv3		Resnet152v2		الشبكة المختارة
Train	test	train	Test	train	Test	
0.9296	0.8562	0.9236	0.8578	0.8931	0.8562	/ wrist المعصم
0.9624	0.8544	0.9524	0.8504	0.908	0.875	/ elbow الكوع
0.9641	0.8546	0.9661	0.8646	0.9248	0.8681	humerus / العضد
0.9376	0.7702	0.9222	0.7831	0.8656	0.7849	shoulder / الكتف
0.9292	0.7546	0.9392	0.7746	0.8978	0.7902	/ finger الاصبع
0.9782	0.7917	0.962	0.7913	0.9424	0.7986	/ forearm الساعد
0.8802	0.791	0.8842	0.7909	0.8392	0.8036	hand / اليد

كما نعرض في المخططين الآتيين توضيح لنتائج المقارنة من خلال المخططين التوضيحين لأداء النماذج الثلاثة على كل أنواع الدراسة لكل من التدريب والاختبار.



المخطط (1) مقارنة نتائج التدريب لأنواع الشبكات الثلاثة على كل أنواع الدراسة



المخطط (2) مقارنة نتائج الاختبار لأنواع الشبكات الثلاثة على كل أنواع الدراسة

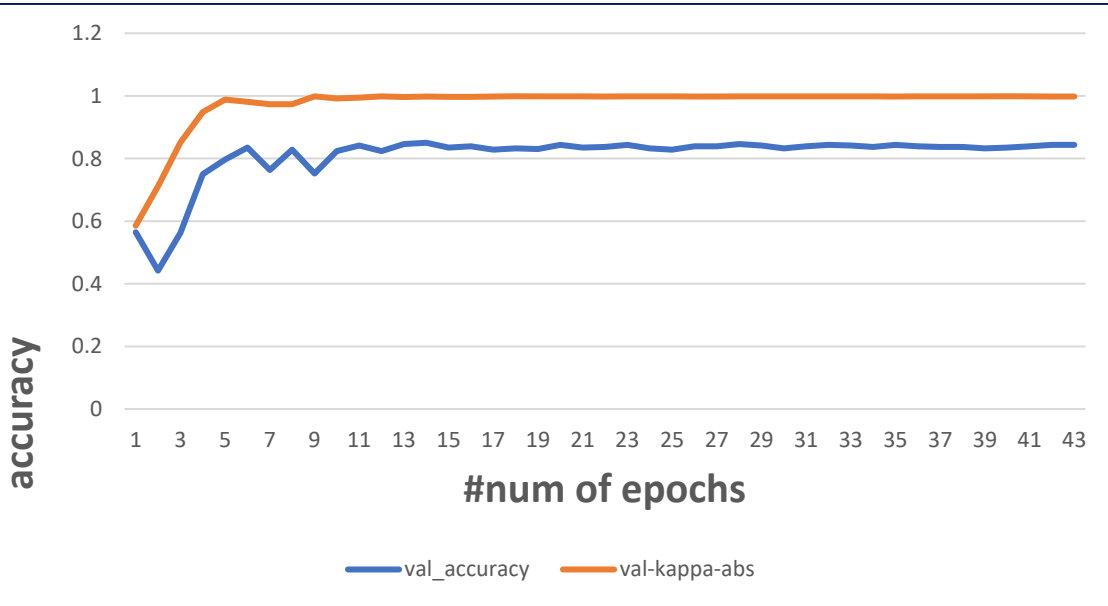
نعرض في الجدول الثاني نتائج مقياس كابا لحالات الدراسة ضمن عينة البيانات من خلال نماذج الشبكات الثلاثة.

Resnet152v2 بحيث حصلنا على قيم جيدة تتناسب مع دقة التصنيف المعروضة في الجدول الأول.

نجد من الجدول أنه بالنسبة لنوع المعصم والكوع فقد أعطت شبكة InceptionResnetv2 أعلى قيمة لمقياس كابا من ناحية الاختبار، بالنسبة للعضد كانت القيمة الأكبر لمعامل كابا لنموذج Inceptionv3 وتفوقت Resnet152v2 في باقي أنواع الدراسة بحصولها على أعلى قيم لهذا المقياس.

الجدول (2) نتائج مقارنة أداء الشبكات الثلاث المدروسة من ناحية مقياس كابا					
kappa static					
InceptionResnetv2		Inceptionv3		Resnet152v2	
Train	test	train	Test	train	Test
0.9545	0.9577	0.9373	0.9207	0.896	0.8497
0.9541	0.9488	0.9418	0.9099	0.9384	0.895
0.99578	0.9299	0.99621	0.948	0.997	0.942
0.9997	0.85	0.9993	0.889	0.9863	0.8627
0.9675	0.6649	0.9572	0.9598	0.9535	0.9649
0.9379	0.9138	0.9204	0.8894	0.9337	0.9194
0.841	0.8191	0.8649	0.8224	0.8965	0.8826

ونعرض في المخطط الآتي قيم الدقة وقيم مقياس كبا لإحدى حالات الدراسة ويظهر من خلاله التناسب الذي نتج في المقياسين (مقياس الدقة ومقياس كبا للتوافق).



المخطط (3) مخطط توضيحي لنتائج مقياسي الدقة وكبا لحالة دراسة الكوع لشبكة InceptionResnetv2

تطبيق طريقة التعلم الجماعي (Ensemble Method):

الخطوة الأخيرة في هذه الدراسة هي تطبيق تجميع (تم شرح التعلم الجماعي وآلياته سابقاً) للنماذج الثلاثة على واحدة فقط من حالات الدراسة وهي نوع العضد من خلال أكثر آلية تجميع شهيرة بين الآليات التي ذكرناها

الجدول (3) نتائج تطبيق التعلم الجماعي على نوع العضد (من ناحية دقة التصنيف)

دقة التصنيف				وع راسة
Avg	InceptionResnetv2	Inceptionv3	Resnet152v2	شبكة ختارة

single		ensemble		single		ensemble		single		ensemble		single		ensemble		
Test	train	Test	train	Test	train	Test	train	Test	train	test	train	Test	train	test		
0.85997																
0.98253																
0.8713																
0.9661																
0.8646																
0.9976																
0.8846																
0.9169																
0.8472																
0.979																
0.8572																
0.9248																
0.8681																
0.971																
0.8721																
																/ humerus

نلاحظ من الجدول الأخير التحسن في دقة نتائج التصنيف للنماذج الثلاثة للنوع المدروس (نوع العضد HUMERUS) فقد ارتفع أداء نموذج شبكة Resnet152v2 من ناحية دقة الاختبار من 0.8681 إلى 0.8721 وكذلك بالنسبة لنماذج شبكتي Inceptionv3 و InceptionResnetv2 فقد حصلنا أيضاً على زيادة في دقة التصنيف بحوالي 2% لكل من النموذجين من ناحية دقة الاختبار، في العمود الأخير للجدول قمنا بحساب متوسط الدقة الفردية للشبكات الثلاثة ومتوسط دقتها عند جمعها من خلال التكديس في نموذج جماعي لتوضيح التحسين النهائي الذي تم الحصول عليه.

12-النتائج والتوصيات:

- هناك العديد من التأثيرات البشرية في اتخاذ القرارات، بما في ذلك تحديد خيارات المعايير ووظائف التحسين. وبالتالي سنحتاج للخوارزميات للتفكير في طرق تدعم آليات التغذية الراجعة والتسجيل. كما تحتاج الحكومات والشركات والمؤسسات النظر في استخدام الخوارزميات في صنع القرار، والتشاور على نطاق واسع، والتأكد من وجود آليات لاكتشاف ومعالجة أي أخطاء أو عواقب غير مقصودة للقرارات المتخذة.

• تحديد نوع البيانات وكيفية استخدامها يمكن أن يساعد عملية صنع القرار في تقليل أي عدم يقين أو شكوك بشأن كيفية اتخاذ القرارات.

تسلط توصياتنا الضوء على الحالات التي يعمل فيها نموذج صنع قرار معين بشكل جيد (الحالة الخاصة التي تمت دراستها)، ولكن هذا النموذج قد لا يعمل مع جميع الحالات وفي جميع الظروف.

من الناحية البشرية فإنه مع نمو الفريق، يصبح اتخاذ القرارات معاً أكثر صعوبة. فعندما نضيف المزيد من الأشخاص، غالباً ما يصبح الوصول إلى الإجماع بطيئاً للغاية، وتصبح التنازلات التي يتعين القيام بها مؤلمة للغاية. رداً على ذلك، تضع بعض الفرق جميع مسؤوليات اتخاذ القرار على عاتق قائد الفريق، مما يثقل كاهلهم ويثبط عزيمة أي شخص آخر. ولكن هناك العديد من الطرق لاتخاذ القرار معاً، أكثر بكثير من مجرد إعطاء رأي لكل فرد أو إعطاء سلطة لشخص واحد فقط.

من الناحية الآلية (العمل مع الآلة) فإن نمو الفريق يمكن أن يدعم اتخاذ القرار ولا يكون هناك معاناة كما في العامل البشري والآراء المختلفة.

الأعمال المستقبلية:

يمكن توسيع هذا البحث ليشمل تطبيق التعلم الجماعي من خلال العمل على عدة شبكات أخرى وتكديس عدة نماذج من نماذج الشبكات المدروسة على عينة البيانات ومقارنة نتائجها مع نتائج البحث الحالي (في حال العمل على تقنية الجمع بالتكديس).

كما يمكن توسيع البحث بطريقة أخرى وهي تطبيق طرائق التعلم الجماعي المختلفة على عينة البيانات المحددة ودراسة نتائج كل طريقة وتحديد من الأفضل بينها حسب الحالة المدروسة.

13-المراجع:

- 1- John Z., Alistair K., James M., Colin G., 2019 – Algorithmic Decision-Making and the Control Problem, Springer, China.

- 2- Schmidt, J., Marques, M.R.G., Botti, S. *et al.* 2019 - Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater.
- 3- Panagiotis P., Ioannis E. L., (2020) – Ensemble Algorithms and Their Applications, Department of Mathematics, University of Patras, Greece.
- 4- Sarker I.H., (2021) – Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI.
- 5- Kong Y., Yu T., (2018) - A Deep Neural Network Model using Random Forest to Extract Feature Representation for Gene Expression Data Classification. Sci Rep.
- 6- Jarrod C., Kane H., Rituparna D., Aviv S., (2019) - Failing & Falling (F&! F): Learning to Classify Accidents and Incidents in Aircraft Data, IEEE.
- 7- Kyoungok Kim, (2016) - A Hybrid classification algorithm by subspace partitioning through semi-supervised decision tree, Springer.
- 8- Congbo M.A., Wei Emma Z., Mingyu G., Hu W., Quan Z. Sh., (2020) - Multi-document Summarization via Deep Learning Techniques: A Survey, arXiv:2011.04843v2 [cs.CL].
- 9- Yanyang Gu., (2019) - Melanoma Detection Based on Color and Hyperspectral Imaging, School of Info & Comm Tech.
- 10- Rohit B., Aditya Kh., Mohammad Sh., Gaurav D., Parneet S., (2021) - Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning, Computational Intelligence and Neuroscience.
- 11- Wenqian Ch., Shuyu Ch., Hancui Z., Tianshu W., (2017) - A hybrid prediction model for type 2 diabetes using K-means and decision tree, 8th IEEE International Conference on Software Engineering and Service Science (ICSESS).
- 12- Xinli Y., David L., Xin X., Jianling S., (2017) - TLEL: A two-layer ensemble learning approach for just-in-time defect prediction, Information and Software Technology.

- 13- Peyman Sh. M., Sławomir N., Sepideh P., (2020) - Stacked Ensemble of Recurrent Neural Networks for Predicting Turbocharger Remaining Useful Life, Center for Applied Intelligent Systems Research (CAISR).
- 14- Li J., Li X., He D., (2019) - A Directed Acyclic Graph Network Combined with CNN and LSTM for Remaining Useful Life Prediction, IEEE.
- 15- Breiman L., (1996) - Bagging Predictors, Springer.
- 16- D H. Wolpert, (1992) - Stacked Generalization, SiteSeer.
- 17- www.usbji.org
- 18- [Musculoskeletal conditions \(who.int\)](http://www.who.int)
- 19- Pranav R., Jeremy I. , Aarti B. , Daisy D., Tony D., Hershel M., Brandon Y., Kaylie Z., Dillon L., Robyn L. B., Curtis L., Katie S., Matthew P. L., Andrew Y. Ng, 2018 - MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs, 1st Conference on Medical Imaging with Deep Learning.

