

مدخل إحصائي لتحليل النماذج الوبائية

(دراسة تطبيقية)

Statistical Approach to Analyze Epidemiological Models (Applied Study)

شيماء مصطفى عبد الرحيم مصطفى سرحان

المعيدة بقسم الإحصاء التطبيقي والتأمين

كلية التجارة - جامعة المنصورة

الأستاذ الدكتور/ فاطمة علي عبد العاطي

أستاذ الإحصاء التطبيقي

كلية التجارة - جامعة المنصورة

الدكتور/ محمد مصطفى شهاب الدين

مدرس الإحصاء التطبيقي

كلية التجارة - جامعة المنصورة

المستخلص:

تؤثر الأمراض المعدية في البداية على منطقة واحدة فقط وسرعان ما تنتشر في جميع أنحاء العالم، وقد تصيح في النهاية جائحة أو وباء كما حدث في وباء ١٩ - COVID، وتعد المساهمة الأساسية لهذا البحث هي نمذجة الوباء في مصر، وكذلك القدرة على عمل تنبؤات بعدد حالات الإصابة والوفاة والشفاء في مصر على المدى طويل الأجل، من خلال استخدام أحد النماذج المتخصصة في النمذجة الوبائية وأحد نماذج التعلم العميق.

فلقد أجري البحث لدراسة ديناميكيات الوباء في مصر باستخدام خوارزميات التعلم العميق والنماذج الوبائية، ومن بين العديد من خوارزميات التعلم العميق، استخدمنا خوارزمية الشبكة العصبية متعددة الطبقات Multi - Layer Perceptron (MLP)، والنموذج الوبائي (SIR) - Susceptible - Infected - Removed وذلك لإظهار دور تلك النماذج في التنبؤ بتطور جائحة COVID-١٩ في مصر وتقدير الوقت اللازم لاختفائه، وأيضاً لبيان التوزيع الوبائي له في مصر.

استهدفت تلك الدراسة الوصول إلى نموذج مناسب للتنبؤ بوباء ١٩ - COVID في مصر، ولذلك تناولت الدراسة تطبيق نموذج SIR، ونموذج الشبكة العصبية متعددة الطبقات MLP، وتمت المقارنة بين تلك الأساليب باستخدام معايير قياس دقة التنبؤ، واتضح من خلالها أن نموذج MLP للتنبؤ بوباء ١٩ - COVID أكثر كفاءة من نموذج SIR.

الكلمات المفتاحية: النمذجة الوبائية، نموذج SIR، الذكاء الاصطناعي، الشبكات العصبية الاصطناعية، نموذج MLP، وباء ١٩ - COVID، رقم التكاثر الأساسي.

Abstract:

Infectious diseases initially affect only one specific region, but they quickly spread to all parts of the

world and may eventually become a pandemic or epidemic, as was the case with the COVID-19 pandemic. The primary contribution of this study is to model the epidemic in Egypt and to be able to make predictions about the number of infected, deaths, and recovered cases in Egypt over the long term. This was achieved by using a specialized epidemiological modeling model and a deep learning model.

This study aimed to find a suitable model for predicting the evolution of COVID-19 epidemic in Egypt. Therefore, the study applied the following models in forecasting: The SIR epidemic model, and The Multi-Layer Perceptron (MLP) neural network model.

Comparison between these models has been done to reach the best model that can be used in prediction. These methods were compared using prediction accuracy measures, and it was found that the MLP model for predicting the COVID-19 epidemic was more efficient than the SIR model.

Keywords:

Epidemiological Modeling, SIR model, Artificial Intelligence, Neural Networks, MLP Model, COVID-19 Epidemic, Basic Reproduction Number.

المقدمة ومشكلة الدراسة

منذ ظهور جائحة COVID – 19 في ديسمبر ٢٠١٩، أثرت على حياة الإنسان والاقتصاد العالمي، وتسببت في اضطرابات كبيرة. فلقد انتشر الوباء في كل قارات العالم، وتسبب في حوالي ما يقرب من ٧٠٥ مليون إصابة، و٧٠١ مليون حالة وفاة حتى ٣٠ أبريل ٢٠٢٣ في جميع أنحاء العالم. وتتساقط الدول لإبطاء انتشار المرض فلا يزال الوباء مستمرا، وتبذل الجهود العالمية لفهم طبيعة الفيروس وإدراك علم الأوبئة وكذلك محاولة إبطاء انتشاره من خلال تطبيق التدابير الاحترازية اللازمة.

ومن هنا جاءت أهمية النمذجة الوبائية Epidemiological Modeling، حيث تلعب النماذج الرياضية (النماذج الوبائية)، وكذلك الذكاء الاصطناعي Artificial Intelligence، دورا هاما في صحة العامة في العديد من الجوانب، مثل: التصور السريع للمعلومات الوبائية، وتقدير والتنبؤ بالوباء، ورصد سرعة انتشاره، وبالتالي المساعدة في اتخاذ القرارات بشأن الوقاية من الأوبئة ومكافحتها.

أجري البحث لدراسة ديناميكيات الوباء في مصر باستخدام خوارزميات التعلم العميق والنماذج الوبائية. ومن بين العديد من خوارزميات التعلم العميق استخدمنا خوارزمية لشبكة العصبية متعددة الطبقات Multi – Layer Perceptron (MLP)، والنموذج الوبائي Susceptible – Infected – Removed (SIR)، وذلك لإظهار دور نموذجي SIR، و MLP في التنبؤ بتطور جائحة COVID-19 في مصر وتقدير الوقت اللازم لاختفائه، وأيضا لبيان التوزيع الوبائي له في مصر.

أهداف الدراسة

يمكن بلورة أهداف الدراسة في النقاط التالية:

(١) الوصول إلى نموذج منسوب لفهم ديناميكيات انتشار الوباء وتحليل جزئ تفاصيل وخصائص الأمراض الوبائية.

٢) التنبؤ بعدد لسكان للصابين والمتعافين والوفيات في المجتمع مع استبعاد الزيادات المحتملة في فس الوقت في عدد الأفراد المعرضين للإصابة، ومن ثم التنبؤ بالاتجاه العام للوباء وتطوره.

٣) تقدير المعلمات الوبائية المختلفة مثل: معدل الانتشار (Transmission Coefficient or Rate (α))، ومعدل لشفاء (Recovery Coefficient (γ))، ومعدل الوفاة (Death Rate (μ))، والتي منها يتم تقدير المقاييس الوبائية المختلفة كتقدير عدد التكاثر الأساسي Basic Reproduction Number (R_0))، ومستوى المناعة لجماعية (Herd Immunity Limit (HIL))، ومستوى التطعيم اللازم للقضاء لقضاء على المرض V_c ، ومدة الوباء المتبقية لانتهائه Epidemic Time، ومن ثم يمكننا حساب والتنبؤ بحجم لسكان المعرضين للإصابة S ، وللصابين I ، والمتعافين من المرض R ، وبالتالي القدرة على على تتبع مسار الوباء والتنبؤ بالاتجاه العام له.

٤) عرض لخوارزميات الأساسية لنموذجي SIR، وMLP، واستخدام هذه لخوارزميات في مجال النمذجة الوبائية.

٥) عرض التحليل الرياضي بشكل أكثر وضوحا وبالتالي تحليل التفسيرات النظرية بشكل سليم.

٦) تحديد ما إذا كان الوباء يتبع التوزيع لطبيعي أم لا.

٧) عمل تنبؤات وتوقعات على المدى طويل الأجل.

٨) التعبير عن فائدة اللقاح وإجراءات الوقاية في تنمية المناعة لجماعية وزيادتها.

٩) حساب المؤشرات الإحصائية التي تسمح لمسؤولي لصحة معرفة ما إذا كانت لسياسات المتبعة فعالة أم لا، وإذا كان معدل انتشار الوباء في تزايد أو تنقص أو في حالة مستقرة.

أهمية الدراسة

تكن أهمية الدراسة فيما يلي :

أولاً: الأهمية العلمية

تتمثل الأهمية العلمية والإحصائية لهذا البحث في عرض الأساس الرياضي والإحصائي لأحد أساليب للسلاسل الزمنية وهو نموذج SIR، وأحد نماذج التكاثر الاصطناعي (التعلم العميق)، وهو أسلوب للشبكات العصبية الاصطناعية ANN والمقارنة بين النموذجين لتحديد الأفضل في التنبؤ بالاتجاه العام للوباء COVID-19 في جمهورية مصر العربية.

ثانياً: الأهمية العملية

ترجع أهمية البحث العملية في بناء نموذج إحصائي للتنبؤ بالاتجاه العام للوباء COVID-19، وتحديد حدود التطور المستقبلي للوباء في جمهورية مصر العربية، مما له تأثير على الاقتصاد ونظم لصحة

ونظم لصحة والموارد المادية والبشرية، ومن ثم يمكن تقييم درجة تلعب الدولة لمواجهة هذا الوباء.
الوباء.

النماذج المستخدمة

هناك العديد من نماذج التنبؤ بالأمراض المعدية، وفي هذا البحث سنقوم باستخدام أحد النماذج المقصصة في النمذجة الوبائية، وأحد نماذج التعلم العميق، وهما نموذجي SIR، MLP.

أولاً: نموذج SIR

نموذج (Susceptible, Infected, Removed) SIR هو أحد أبسط النماذج الوبائية، يستخدم لتقدير عدد الأفراد للمصابين والمتعافين من مرض وبائي معين، وذلك في مجتمع كبير ومغلق، كما يعد يعد أيضاً أحد النماذج المستخدمة في السلاسل الزمنية، وبالتالي فهو نموذج ديناميكي dynamic model ولتكوين نموذج ديناميكي فإننا نحتاج لما يسمى بمتغيرات الحالة State Variables، ونحتاج ونحتاج أيضاً لمعرفة كيف تتغير هذه المتغيرات بمرور الوقت (معدل التغير أي الميل Slope). ومنه فإننا ومنه فإننا نستطيع التنبؤ بالاتجاه المستقبلي للوباء، والعدد المحتمل للمصابين في حالة وجود أو عدم وجود تطعيم، وتقدير السمات الرئيسية للمرض، وكذلك التأثير المحتمل لتدخلات لصحة العامة المختلفة. العامة المختلفة.

ولتحديد متغيرات الحالة الأساسية في هذا النموذج فإننا نقوم بتقسيم المجتمع بشكل أساسي إلى ثلاث فئات (Jakhar, et al., ٢٠٢٠):

١- مجموعة السكان المعرضين للإصابة susceptible: تمثل مجموعة السكان الذين ليس لديهم مناعة ضد العمل المعني، والمعرضة للإصابة بالمرض، ولكنها سليمة وتمثل هذه المجموعة سكان مجتمع الدراسة عند بداية الوباء.

٢- مجموعة السكان للصابة بالفعل Infected: مجموعة السكان للصابة بالمرض وتنقل الوباء لمن يتواصل معهم من أفراد المجموعة S.

٣- Removed تشمل مجموعة السكان المعافاة من الوباء والمتوفين أيضاً بسبب الوباء (Recovered, Deaths).

تتوقف جودة النماذج على دقة الافتراضات التي تستند إليها، إذا وجدنا أن توقعات النموذج لا تتماشى مع القيم الفعلية، يعني ذلك ضرورة تغيير الافتراضات الأولية ليصبح النموذج مفيداً (Wang, Jiale, et al., ٢٠٢٢)، (Paul, Amit, et al., ٢٠٢٢, p٤)

ويمكن عرض افتراضات النموذج في النقاط التالية:

- (١) مجتمع الدراسة كبير وثبت أي ثبات حجم السكان (Closed Population).
- (٢) تجاهل التغيرات الحيوية والسكانية للمجتمع (الوفيات والمواليد والهجرة)، عدا الوفيات بسبب المرض.
- (٣) لا يؤثر العمر والجنس والعرق والوضع الاجتماعي على احتمالية إصابة الشخص حيث إن لظاهرة تنشر بشكل عشوائي، وبالتالي تتساوى فرص الإصابة لجميع السكان.
- (٤) لا توجد مناعة موروثية (مشتقة من الأم) في السكان.
- (٥) يعتبر اختلاط الأفراد في جميع السكان يحدث بشكل متجانس.
- (٦) لشفاء من المرض ينتج مناعة دائمة أي لا توجد فرصة للعدوى مرة أخرى.
- (٧) غالباً ما يقترح أن يكون رقم التكاثر (RO) يتبع التوزيع الأسي (من إحصائيات بولتزمان الكلاسيكية (classical Boltzmann Gibbs (BG) statistics).
- (٨) لا توجد تحركات مباشرة لأفراد المجتمع من الفئة S إلى الفئة R. يعد نموذج SIR (أو SIRD) أحد النماذج التي تستخدم في نمذجة الأمراض المعدية وأبسطها، فيجب عدد للصائين بمرض محلي في مجتمع ثابت ومغلق بمرور الوقت. ويحتوي النموذج على أربع معادلات تفاضلية متعلقة بعدد الأشخاص المعرضين للإصابة (S)، وللصائين (I)، وكذلك (R) ويمثل المتعافين و (D) التي يمثل المتوفين (وهنا تم تقسيم المتغير Removed إلى المتغيرين Recovered, (Deaths).

تم صياغة النموذج على النحو التالي:

$$\frac{dS}{dt} = -\alpha \frac{S}{N} I \rightarrow (1)$$

$$\frac{dI}{dt} = +\alpha \frac{SI}{N} - \beta I_t \rightarrow (2)$$

$$\frac{dR}{dt} = +\gamma I_t \rightarrow (3)$$

$$\frac{dD}{dt} = +\mu I_t \rightarrow (4)$$

ويوضح لجدول التالي متغيرات الدراسة وكذلك المعلمات كما يلي:

| | |
|-----------|----------------------------------------------------|
| S | المعرضين للإصابة بالوباء |
| I | حالات الإصابة |
| R | يشمل حالات الشفاء |
| D | حالات الوفاة بسبب الوباء |
| N | حجم السكان (المجتمع محل الدراسة) |
| α | معدل انتشار الوباء (معدل الانتقال أو معدل الإصابة) |
| β | معدل الإزالة |
| γ | معدل الشفاء |
| μ | معدل الوفاة بسبب الوباء |
| λ | معدل النمو الأسي |
| R_0 | رقم التكاثر الأساسي |
| HIT | عتبة المناعة الجماعية |
| V_c | تجاوز الحد |

جدول (١) يوضح متغيرات الدراسة وكذلك المعلمات

ويمكن توضيح جز المقاييس الوبائية المستخدمة ذات صلة كما يلي:

(١) رقم التكاثر الأساسي **Basic Reproduction Number**: يعتبر مقياس أساسي لانتقال

العدوي ويمثل متوسط عدد حالات الإصابة الثانوية الناجمة عن حالة إصابة واحدة ويشار إليه بالرمز

R_0 (Katris, Christos, ٢٠٢١, p.٤).

بصيغة أخرى هو متوسط عدد الحالات الثانوية المكتسبة من الحالة الأولية في مجموعة سكانية معرضة

للإصابة في جميع مراحل العدوى.

أو متوسط حالات العدوى الجديدة المسببة بوسيلة كل حالة في مجموعة معرضة للعدوى بالكامل والتي هي متجلسة، أو مجموعة مختلطة جيداً، أي يمكن لكل شخص فيها الاتصال مع كل فرد آخر معرض للعدوى.

(٢) **المناعة الجماعية Herd Immunity**: المناعة الجماعية أو مناعة القطيع (HI) هي شكل من أشكال الحماية غير المباشرة من مرض معد، وتحدث عندما تكتسب نسبة كبيرة كافية من المجتمع مناعة لعدوى معينة، إما بسبب الإصابة بها سابقاً أو التطعيم، مما يوفر حماية للأفراد الذين ليس لديهم مناعة ضد المرض (Randolph, H. E., & Barreiro, L. B., ٢٠٢٠).

حين تكتسب نسبة حاسمة من السكان المناعة ضد الوباء، يطلق على هذه الحالة عتبة المناعة الجماعية (HIT) أو مستوى المناعة الجماعية (HIL) وهي النقطة التي تنخفض فيها نسبة الأفراد المعرضين للإصابة في مجتمع ما عن الحد الأدنى المطلوب لانتقال المرض. وحينها لا يستمر المرض في الانتشار بين السكان ويتوقف المرض المتوطن. يمكن قياس قيمة هذه العتبة بضرب R_0 في S وهو هنا نسبة السكان المعرضين للإصابة ومنها نجد:

$$\therefore HIT = 1 - \frac{1}{R_0} = \frac{R_0 - 1}{R_0}$$

(٣) **التطعيم Vaccination**: لتحقيق المناعة الجماعية HI من خلال التطعيم، يهدف مصنعو اللقاحات إلى إنتاج لقاحات ذات معدلات فشل منخفضة، ويهدف صانعو السياسات إلى تشجيع استخدامها. بعد الإدخال الناجح للقاح واستخدامه على نطاق واسع، يمكن ملاحظة انخفاض حاد في الإصابة بالأمراض التي بقي من الإصابة بها، مما يقلل من عدد حالات الإصابة والوفيات التي تسببها مثل هذه الأمراض.

وإذا افترضنا أن اللقاح فعال بنسبة ١٠٠٪، فإن المعادلة المستخدمة لحساب عتبة مناعة القطيع يمكن استخدامها لحساب مستوى التطعيم اللازم لفضاء على المرض، مكتوباً بصيغة V_c ، ولكن عادة ما تكون فعالية اللقاحات غير كاملة، لذلك يجب مراعاة فعالية اللقاح E عند حساب V_c :

$$V_c = \frac{HIT}{E} = \frac{1 - \frac{1}{R_0}}{E}$$

(٤) **تجاوز الحد Overshoot**: تجاوز الحد هو نتيجة غير مرغوب فيها، قد تحدث بشكل كبير حتى بعد تحقق المناعة الجماعية، مما يؤدي إلى عدد كبير من الإصابات والوفيات قبل أن ينخفض الانتشار. الانتشار. وهو الفرق بين النسبة التراكمية للأفراد الصائين والعتبة النظرية للمناعة الجماعية): (Nguyen, M. M., ٢٠٢٤)

Overshoot=Proportion of Total Infections – HIT

٥) الحالات النشطة **Active Cases**: قصد بالحالات الشطة لحالات التي لا تزال لا نتيجة لها

(سواء بالشفاء أو الوفاة)، ويتم حساب لحالات الشطة اليومية من المعادلة التالية:

$$\text{Daily Active Cases} = \text{Total Infected} - [\text{Total Recovered} + \text{Total Deaths}]$$

طريقة تقدير النموذج

إن إيجاد قيم معاملات النموذج التي تعطي أقرب تطابق بين تنبؤات النموذج والبيانات الفعلية هي المشكلة الأساسية للنموذج، ومن أهم طرق المستخدمة في تقدير معاملات النموذج هي خوارزمية التحسين Optimization Algorithm لمعرفة أفضل ما يفعله النموذج من الناحية العملية ولفهم الآليات ولها

العديد من الوسائل أهمها:

(١) طريقة المربعات لصغى العادية Ordinary Least Squares

(٢) طريقة الإمكان الأعظم Maximum Likelihood Method

(٣) طريقة إيلر Euler Method

وسوف يقصر التنبؤ في هذا البحث بتطبيق طريقة المربعات لصغى العادية OLS، وحيث إن γ ,

إن $\mu, \beta, \alpha, \gamma$ هي معدلات يومية، سيتم تحسين قيمها يوميا بتطبيق OLS وبالاعتماد على البيانات

التاريخية اليومية للمتغيرات (S, I, R, D) كما يلي (Sebbagh, A., & Kechida, S. (٢٠٢٢))

:P٢)

$$\alpha = \frac{\Delta S}{\frac{S_{n-1}}{S_0} \cdot I_{n-1}} = \frac{S_{n-1} - S_n}{\frac{S_{n-1}}{S_0} \cdot I_{n-1}}$$

$$\beta = \frac{\Delta R(\text{Removed})}{I_n} = \frac{R_n - R_{n-1}}{I_n}$$

$$\gamma = \frac{\Delta R(\text{Recovered})}{I_n} = \frac{R_n - R_{n-1}}{I_n}$$

$$\mu = \frac{\Delta D}{I_n} = \frac{D_n - D_{n-1}}{I_n}$$

$$\lambda (\text{Exponential Growth Rate}) = \alpha - \beta$$

$$\text{Epidemic Time} = \ln(\gamma) / \lambda$$

ومن خلال التقديرات الخاصة بمعلمات النموذج نقوم بتقدير بعض المقاييس الوبائية والتنبؤ

بالمتغيرات ($S_{n+1}, I_{n+1}, R_{n+1}, D_{n+1}$) من خلال المعادلات التالية:

$$R_0 = \frac{\alpha}{\beta}$$

$$\text{HIT} = \frac{R_0 - 1}{R_0} = 1 - \frac{1}{R_0}$$

$$D(\text{duration of infection}) = 1 / \gamma$$

$$\text{OV} = \text{Proportion of Total Infections} - \text{HIT}$$

$$V_c = \frac{1 - \frac{1}{R_0}}{E} = \frac{HIT}{E} = \frac{P_c}{E}$$

$$S_{n+1} = S_n + \hat{\alpha} \cdot \frac{S_{n-1}}{S_0} \cdot I_{n-1}$$

$$I_{n+1} = I_n + \hat{\alpha} \cdot \frac{I_{n-1}}{S_0} \cdot I_{n-1} - \beta I_{n-1}$$

$$(Removed)R_{n+1} = R_n + \beta I_{n-1}$$

$$(Recovered)R_{n+1} = R_n + \gamma I_{n-1}$$

$$D_{n+1} = D_n + \mu I_{n-1}$$

ثانيا: الشبكات العصبية الاصطناعية ونموذج MLP

لشبكات العصبية الاصطناعية (ANNs) هي نماذج حسابية مستوحاة من عمل الدماغ البشري. تستخدم ANNs في العديد من التطبيقات بما في ذلك التنبؤ وتحليل البيانات فليها قدرة عالية علي تحديد وحل العلاقات غير لخطية Non – Linear Relations. وفي هذا البحث، سنناقش استخدام نموذج لشبكات العصبية متعددة لطبقات (MLP) Multilayer Perceptron في التنبؤ بوباء كورونا في مصر، مع التركيز على حالات الإصابة، التعافي، والوفاة.

الشبكات العصبية متعددة الطبقات Multilayer Perceptron

MLP هو نوع من لشبكات العصبية الاصطناعية تم اقتراحه بولسطة Minsky و Papert في الخمسينيات من القرن الماضي ١٩٥٠s. ويمثل Perceptron اول نموذج بسيط للتعلم الآلي، حيث إن Perceptron هو الغر الأصغر في الشبكة ويقوم بإجراء حسابات معينة لتحليل خصص المدخلات، وتضع عملية تدريب شبكة MLP الي التدريب الإشرافي والتي تعمل على تعديل الأوزان تدريجيا بناء على الفرق بين المخرجات المتوقعة والمخرجات الفعلية. تتكرر هذه العملية حتى يتم تقليل خطأ إلى حد الأدنى (Azeem, Muhammad, et al., ٢٠٢٣, p٥) ولشبكة MLP ذاكرة ديناميكية dynamic memory وهي مناسبة جدا في مجال التنبؤ بالسلسلة الزمنية، وفي الوقت لحالي، تستخدم على نطاق واسع في مخف المجالات وقد حقت نتائج ملحوظة وناجحة في مجال التنبؤ.

تتكون لشبكة من ثلاث طبقات رئيسية:

(١) طبقة المدخلات (Input Layer): تستقبل البيانات الخام وتمثل المتغيرات المستقلة التي يتم تقديمها إلى النموذج وكل متغير يتم تمثيله بوزن محدد. والأوزان هي قيم عددية مرتبطة بكل متغير يتم تعديلها خلال عملية التدريب لتحقيق أفضل أداء.

٢) طبقة المخفية (Hidden Layers): تقوم بعمليات حسابية ومعالجة البيانات. يمكن أن تكون هناك واحدة أو أكثر من هذه الطبقات.

٣) طبقة المخرجات (Output Layer): تمثل القيم المتوقعة للمتغيرات التابعة بناء على قيم المدخلات.

وللتنبؤ بتطور حالات كورونا في مصر، نحتاج إلى بيانات تاريخية تشمل:

- حالات الإصابة اليومية.

- حالات الشفاء اليومية.

- حالات الوفاة اليومية.

وحيث إن الشبكة العصبية لا تبرمج، بل إنها تقوم بالتعلم والتدريب، ويقصد بتدريب لشبكة هو إجراء يتم فيه استخدام قواعد التدريب والتعلم لضبط الأوزان. وطرق التدريب الأساسية هي: التدريب الإشرافي supervised training (التدريب في وجود معلم)، التدريب غير الإشرافي unsupervised training (التدريب في عدم وجود معلم)، التدريب بإعادة التعزيز reinforcement training or hybrid training، التدريب المتأخر delay learning.

وفي هذا البحث سنقوم بتطبيق شبكة MLP، التي تضع قسط للتدريب الإشرافي ويقصد به التدريب في وجود معلم، حيث تعد هذه الطريقة هي الأكثر انتشاراً في تدريب لشبكة العصبية، وفيها تجمع عينات من البيانات تحتوي قيماً للمتغيرات المدخلة وقيماً للمتغيرات المخرجة، وتقوم لشبكة بمقارنة النتائج التي تقدرها للمتغيرات الخارجة لكل عينة مدخلة بالقيم الفعلية لهذه المتغيرات، وبناء على ذلك تقوم لشبكة بإجراء التعديلات على أوزان الاصل بهدف تقليل الأخطاء في النتائج، ثم تعاد عملية التدريب عدة مرات إلى أن يتم الوصول إلى نتائج مقبولة ومقاربة للقيم الفعلية.

النهج المستخدم في بناء وتطبيق الشبكة العصبية متعددة الطبقات (MLP) لتحليل البيانات، وكذلك الأهداف التي يسعى كل نهج لتحقيقها.

لإجراء التحليل على برنامج SPSS، فإننا بعد جمع البيانات نحتاج إلى إعدادها كي نتأكد من خلوها من القيم المفقودة، ثم يتم تقسيم المتغيرات إلى مستقلة وأخرى تابعة (مستهدفة)، وفي تحليلنا هذا نقوم بتحويل المتغيرات (I, R, D) الي Lagged Values والتي يطلق عليها القيم السابقة أو التاريخية أو المتأخرة حيث تمثل $[I(n-1), R(n-1), D(n-1)]$ المتغيرات المستقلة وتستخدم في التنبؤ بالمتغير التابع.

وهناك ثلاث طرق لإجراء تحليل لشبكة العصبية متعددة الطبقات (MLP) وهم:

(١) شبكة واحدة لكل متغير Separate Network for Each Variable : يتم وضع متغير واحد
قطر كمتغير تابع، واستخدام المتغير المتأخر لهذا المتغير كمتغير مستقل.

(٢) شبكة مفصلة لكل متغير تابع على حدة (تحليل مفصل لكل متغير) Separate Network for
Each Dependent Variable: يتم بناء شبكة عصبية مفصلة لكل متغير تابع (I, R, D)،
يكون لكل شبكة عصبية متغير واحد تابع التي تحاول الشبكة التنبؤ به أو تصنيفه، بالإضافة الي
الي مجموعة من المتغيرات المستقلة (ويتم استخدام المتغيرات المتأخرة كمتغيرات مستقلة للشبكة).

(٣) شبكة واحدة لجميع المتغيرات (المستقلة والتابعة معا) Single Network for All
Variables: هنا نقوم بالتعامل مع جميع المتغيرات ككل وبناء شبكة عصبية واحدة تشمل جميع
جميع المتغيرات (Infected, Recovered, Deaths) كمتغيرات تابعة، وجميع المتغيرات
المتأخرة (Lagged Infected, Lagged Recovered, Lagged Deaths) كمتغيرات
مستقلة. وبالتالي تدخل جميع المتغيرات الي الشبكة، ويتم معالجة العلاقة بينها داخل فس الهيكل
الهيكل العصبي.

الخلاصة: نختار النهج الأول عندما يكون لدينا بيانات سلسلة زمنية، ويكون المتغير نفسه هو
العمل الرئيسي التي يؤثر على التوقعات المستقبلية، وفي الحالات التي تفقر إلى بيانات خارجية ذات
صلة ، أما النهج الثاني فهو منسب عندما يكون لكل متغير خصص مختلفة جدا لتحسين الدقة والموثوقية
لكل متغير على حدة وتحليل خصص لكل نتيجة. ونختاره عندما تكون التفاعلات بين المتغيرات أقل أهمية
أو يمكن معالجتها بطرق أخرى، أما النهج الثالث فهو مناسب للوصول على رؤية شاملة ومراعاة العلاقات
المتبادلة بين المتغيرات.

وباتباع أي نهج منهم، يمكننا تحقيق فهم أفضل وتنبؤات أكثر دقة باستخدام تحليل لشبكات العصبية
متعددة لطبقات (MLP). ويعتمد الاختيار على الأهداف البحثية ومتطلبات التحليل. هنا سوف نختار
النهج الثالث فهو منسب أكثر عند مقارنته بنموذج SIR فهو نموذج ديناميكي يعتمد التغير في كل
متغير على التغيرات في المتغيرات الأخرى؛ للوصول على رؤية شاملة ومراعاة العلاقات المتبادلة بين
المتغيرات.

مصدر البيانات

يتمثل مصدر البيانات في البيانات اليومية المتعلقة بعدد حالات الإصابة، والشفاء، والوفيات جراء
جاء وباء ١٩ - COVID في الفترة من ١٤ فبراير ٢٠٢٠ الي ٣٠ ابريل ٢٠٢٣، والمنشورة في

موقع منظمة لصحة العالمية (WHO)، وموقع Worldometer، وموقع Our World in Data، وكذلك وزارة لصحة للصربية MoHP، والجهاز المركزي للتعبئة العامة والإحصاء بجمهورية مصر العربية. مصر العربية.

الدراسة التطبيقية

أولاً: التحليل الوصفي للبيانات الفعلية لوباء ١٩ - COVID في مصر:

يقدم هذا القسم التحليل الوصفي للبيانات اليومية الفعلية عن حالات الإصابة والوفاة والتعافي لوباء ١٩ - COVID منذ بداية ظهوره في مصف فبراير ٢٠٢٠ حتى ٣٠ ابريل ٢٠٢٣. وكما هو هو موضح في الجدول التالي الإحصاءات الوصفية خلال الموجات الأربع ولأضاً خلال فترة الوباء بالكامل:

| Measures | Wave ١ | | | Wave ٢ | | | Wave ٣ | | |
|----------|---------|---------|-------|---------|--------|---------|--------|---------|--------|
| | I | R | D | I | R | D | I | R | D |
| Mean | ٣٩٩.٥٦ | ٣٦٠.٩١ | ٢٣.١٧ | ٦٨٨.١١٥ | ٤٢١ | ٤٢ | ٥٥٦.٠٦ | ٥٣٣ | ٢٩.١٦ |
| Median | ١٧٠.٥ | ١٤٧ | ١٤ | ٦٢٧ | ٤٠٠.٨٥ | ٤٦ | ٥٦٩ | ٤٩٥.٩٩ | ٣١ |
| Max | ١٧٧٤ | ١٧١٦ | ٩٧ | ١٤١٨ | ٨١٣ | ٦٤ | ١٢٠.٣ | ١٠٠.٥ | ٦٨ |
| Min | ٠ | ٠ | ٠ | ٣٥١ | ٦٧ | ١١ | ٠ | ٢٣ | ٠ |
| Sd | ٤٧٢.٧٨٩ | ٣٩٥.٠٧٨ | ٢٤.١٦ | ٢٥٦.٥٢ | ١٧٠.٩٤ | ١٣.٧٦٦٦ | ٣٨٩.٥٢ | ٢٧٤.٠٩٦ | ١٩.٥٢١ |
| Sum | ١١٢٦٧٦ | ١٠١٧٧٧ | ٦٣٥٣ | ٨٩٤٥٥ | ٥٢١١١ | ٥٤٦٠ | ١٠١٧٥٩ | ٩.٧٦٧ | ٥٣٣٦ |

| Measures | Wave ٤ | | | Total | | |
|----------|--------|--------|---------|--------|--------|-------|
| | I | R | D | I | R | D |
| Mean | ٣٦٨.٤٧ | ٣٤٢.٨٤ | ١٣.٠٢ | ٤٤٠.٧٦ | ٣٩٩ | ٢١.٢ |
| Median | ٠ | ٣٤٢ | ٠ | ١٦٨ | ٣٧٧.٥٦ | ١٣ |
| Max | ٥٥١٦ | ١٨١٢ | ١٣٢ | ٥٥١٦ | ١٨١٢ | ١٣٢ |
| Min | ٠ | ٠ | ٠ | ٠ | ٠ | ٠ |
| Sd | ٦٧٣.٧٣ | ٢٩٤.٦٩ | ٢١.٧٩٩٩ | ٥٦٥.٢٣ | ٣١٣.٠٣ | ٢٣.٣٦ |
| Sum | ٢١٢٢٤٠ | ١٩٧٤٧٣ | ٧٤٩٩ | ٥١٦.٢٣ | ٤٤٢١٢٨ | ٢٤٨٣٠ |

جدول (٢) يوضح الإحصاءات الوصفية للبيانات الفعلية

يوضح الجدول (٢) أن المتوسط اليومي للحالات للصابة والوفيات والتعافي أكبر من قيم الوسيط للمتغيرات الثلاث خلال الأربع موجات ولأضاً خلال فترة الوباء بالكامل. ونستنتج من تلك أن البيانات لا تتبع التوزيع لطبيعي كما أنها ملتوية تجاه اليمين. وهذا يشير الي وجود عدد قليل من الأيام ذات عدد حالات مرتفع بشكل غير طبيعي نتيجة لنفشي مفاجئ أو موجات جديدة، هذه القيم العالية سحب المتوسط الي أعلي، مما يجعله أكبر من الوسيط، مما يعني أن أعداد الحالات عند نهاية كل موجة صحوية بالتواء تجاه اليمين (أي في تزايد، وبالتالي أعداد الحالات عند نهاية كل موجة تميل الي القيم المرتفعة).

كما يلاحظ ارتفاع قيم الانحراف المعياري، وهذا يشير الي وجود فترات من الزيادات والانخفاضات والانخفاضات الكبيرة (ذروة وانكاس) وهو ما يكن تقبلاً ملحوظاً في أعداد حالات الإصابة،

والوفيات، ولشفاء، وهذا يشير إلى عدم استقرار الوضع الوبائي مما يتطلب اتخاذ إجراءات احترازية احترازية سريعة.

ويمكن أيضا من خلال الإحصاءات الوصفية توضيح العلاقة بين المتغيرات؛ على سبيل المثال، قد نلاحظ أنه مع ارتفاع حالات الإصابة، يمكن أن ترتفع معدلات الوفيات والتعافي، ولكن بشكل غير متساو. يمكن أن يساعدنا هذا في تقديم تفسير أعمق حول فعالية الاستجابة لصحية وتأثير الفيروس على المجتمع.

إلى جانب ذلك، يمكن للإحصاءات الوصفية أن تساعد في التنبؤ باتجاه الوباء على المدى القصير. فمثلا، إذا أظهر تحليل البيانات أن هناك ارتفاعا متزايدا في عدد الحالات للصابة، يمكن استخدام هذه المعلومات للتخمين من موجات جديدة، وبالتالي معرفة فط التغيير في البيانات بمرور الوقت، مثل ارتفاع أو انخفاض الوفيات، قد يساعد في تخصيص الموارد لصحية بشكل أفضل.

ثانيا: بالنسبة لنموذج SIR

تم استخدام لغة البرمجة R إصدار (4.3.1) لتحليل البيانات الخاصة بوباء 19 - COVID، حيث تم الاعتماد على عدد من لحزم الإحصائية Packages لاجراء التحليل واستخراج النتائج، وهي: (EpiModel, reshape2, ggplot2, Covid - 19. analytics, desolve).

ووفقاً للمنهجية التي تم عرضها سابقاً سيتم تقسيم الدراسة التطبيقية الى مرحلتين، الأولى توضح الإحصاءات الوصفية للبيانات من خلال نموذج SIR، والثانية توضح الإحصاء التحليلي للنموذج من خلال تقدير معاملات المتغيرات، وحساب بعض المقاييس الوبائية ذات صلة، وكذلك التنبؤ بالمتغيرات، واستخراج الأخطاء.

باستخدام طريقة OLS لتقدير المعلمات يوميا خلال الأربع موجات وأيضاً خلال فترة الوباء كاملة وبالتطبيق على برنامج R صلنا على النتائج التالية:

| Waves | α | γ | μ | β | λ | Epidemic Time | R_0 | P_c and V_c | OV |
|-------|----------|----------|-------|---------|-----------|---------------|-------|-----------------|----|
| A | 0.105 | 0.078 | 0.005 | 0.083 | 0.023 | 30 | 1.27 | 21.3% | 0 |
| B | 0.185 | 0.117 | 0.057 | 0.069 | 0.116 | 6 | 2.68 | 62.7% | 0 |
| C | 0.258 | 0.107 | 0.058 | 0.165 | 0.093 | 8 | 1.56 | 35.9% | 0 |
| D | 0.077 | 0.055 | 0.05 | 0.105 | -0.03 | 0 | 0.73 | 0 | 0 |
| Total | 0.225 | 0.252 | 0.17 | 0.422 | 0.203 | 3 | 1.48 | 32.5% | 0 |

جدول (3) يوضح قيم المعلمات المقدرة لنموذج SIR

بالنسبة للموجة الأولى A: نجد أن معدل انتشار الوباء $\alpha=0.105$ وكذلك معدل الإزالة $\beta=0.083$ ومنها نجد أن $R_0 = 1.27$ (أكبر من الواحد لصحيح)؛ مما يعني أن الوباء انتشر بسرعه بسرعه ويجب اتخاذ الإجراءات اللازمة؛ كي نتمكن من السيطرة عليه ووقف وإبطاء انتشاره؛ لذا نجد أننا نجد أننا نحتاج أن ناصد تطعيم حوالي 2% من إجمالي حجم السكان.

بالنسبة للموجة الثانية B: نجد أن معدل انتشار الوباء $\alpha=0.185$ وكذلك الإزالة $\beta=0.069$ نلاحظ زيادة معدل الانتشار خلال الموجه الثانية عن الموجه الأولى وصحوبا أيضا بانخفاض معدل الإزالة؛ لذا فإننا نجد $R_0 = 2.68$ وذلك يعني أن كل شخص مصاب يعي في المتوسط ثلاث أشخاص تقريبا أي أننا نحتاج لتطعيم 62.7% من إجمالي حجم لسكان. مما يعني أنه وفق تقديرات النموذج أن الموجه الثانية أكثر حدة من الموجه الأولى مما يستدعي سرعة تصنيع على اللقاحات.

أما الموجه الثالثة C: فإننا نجد معدل انتشار الوباء $\alpha=0.258$ وكذلك الإزالة $\beta=0.165$ وهنا نلاحظ ارتفاع معدل الانتشار، ولكنه صحوبا بارتفاع معدل الإزالة التي لى إلى تخفاض قيمة $R_0 = 1.06$ عن الموجه الثانية وذلك بسبب ظهور اللقاحات حيث قت أول جرعة لقاح في 24/1/2021. 24/1/2021.

وارتفاع معدل الانتشار هنا يعني زيادة مستوي التواصل بين الأفراد وذلك صحيحا لأن قرارات الدولة للصربية في تلك الفترة أشارت الي رجوع الحياة الي طبيعتها وتخفيف الإجراءات لصارمة المتخذة من قبل، وبالتالي عادت لحياة كما كان لحال قبل ظهور لجئحة وازداد التواصل بين الأفراد وبالتالي زيادة في معدل الانتشار، ولكن وجود اللقاحات في تلك الفترة لى إلى ارتفاع في معدل الشفاء وبالتالي انخفاض في متوسط أعداد الإصابة الناتجة عن التواصل بين الأفراد.

أما الموجه الرابعة D: فإننا نجد تخفاض في معدل الانتشار، وصحوبا بارتفاع في معدل الإزالة الإزالة مما يعني فعالية اللقاحات المستخدمة. وبالتالي تخفاض في متوسط أعداد الإصابة الناتجة عن التواصل بين الأفراد حتى وصل إلى $R_0 = 0.73$ (أي أن كل شخص مصاب يعي في المتوسط أقل من المتوسط أقل من شخص واحد فقط وهذه هي المرحلة المراد الوصول إليها) وذلك بسبب وجود مناعة جماعية مناعة جماعية لي الناس ووفقا لهذه التقديرات فإننا نجد أننا لا نحتاج إلى مزيد من التطعيم $HIT = 0\%$.

ومن المعلوم أنه إذا كان $(\lambda = \alpha - \beta) > 0$ فان نسبة الأعداد للصابة تنمو بشكل أسي. ومن خلال خلال قيمة هذا المعدل يمكننا معرفة الوقت المتبقي لانتهاى الوباء Epidemic Time من العلاقة $\ln(\lambda) / \lambda$.

ومن لجدول السابق نلاحظ أن قيمة معدل النمو الأسي خلال أول ثلاث موجات موجبة، مما يوضح أن للحالات للصابة تنمو بشكل أسي ومنه فانه يوجد وقت متوقع لانتهاى الوباء، بينما أصبحت قيمته سالبة خلال الموجه الرابعة وهذا يعني أن الحالات للصابة خلال هذه الموجه لا تتزايد ولا تنمو بشكل أسي، بل تقل تدريجيا، وعليه فإن وقت الوباء المتوقع هو صفر أي ينتهي الوباء بنهاية هذه الموجه. وبالنظر إلى قيمته خلال فترة الوباء كاملة نجدها 0.203، وكذلك الوقت المقدر لانتهاى الوباء هو 3 أيام فقط أي ينتهي الوباء في 23/05/2023.

ومن الملاحظ أيضا أن قيمة تجاوز الحد Overshoot خلال الموجات الأربعة وكذلك خلال فترة الوباء كاملة مساوية لصفر، مما يشير الي أن النسبة التراكمية للأفراد للصابة لم تتجاوز الحد الأدنى (النسبة اللازمة) المطلوب لتحقيق المناعة لجماعية، وذلك يوضح لنا أنه وفق تقديرات النموذج تم السيطرة على الوباء بفعالية قبل أن ينتشر بشكل واسع بين السكان وأن التدابير الوقائية كلفت فعالة للغاية في كبح انتشار الوباء.

ملحوظة: تم افتراض أن فعالية اللقاح $E = 100\%$ ، لذا فإنه لا يوجد فرق بين تقديرات مستوي المناعة الجماعية HIT، ومستوي التطعيم المطلوب V_c فكلاهما واحد.

كما تم تقدير المتغيرات (S, I, R, D) بشكل يومي خلال الموجات الأربعة، ومنه استنتجنا

الإحصاءات الوصفية المقدرة للمتغيرات بولسطة نموذج SIRD:

| SIR | Wave ١ | | | Wave ٢ | | | Wave ٣ | | |
|--------|-----------|-----------|---------|----------|--------|---------|-----------|-----------|---------|
| | I | R | D | I | R | D | I | R | D |
| Mean | ٤٠٢.٨٨ | ٣٧٠.٤١ | ٢٤.٢٤ | ٧٣١.٦٦ | ٤٥٧.٠٤ | ٤٤.١٩ | ٥٥٦.٠٥ | ٥٦٨.٢٣ | ٣٠.٨٦ |
| Median | ١٧٥.٩٩٧ | ١٥٧.٤٢ | ١٤.٩٥ | ٦٤١.٣٧ | ٤٤٠.١٧ | ٤٨.٦٣ | ٦٠٩.٧٤ | ٥٦٥.١٦ | ٣٢.٧٧ |
| Max | ١٧٦٩.٢٥ | ١٨٠٦.٠٣ | ١٠٠.٥٠ | ١٨٤١.٧٦ | ١٢٣٦.٤ | ٦٧.٥١ | ١٣٠٠.٩٥ | ٢٤٣٤.٤٨ | ٧١.٩٨ |
| Min | ٠ | ٠ | ٠ | ٢٣٠.٦٨ | ٩٨ | ١١.٦٣ | ٠ | ٣٤.٣٦ | ٠ |
| Sd | ٤٧٨.٩٢ | ٤٠٤.٦٢ | ٢٥.١٨ | ٢٨٨.٠٣ | ١٥٨.٣٠ | ١٤.٧٥ | ٣٨٥.٧٧ | ٣٠٢.٩٧ | ٢٠.٦٧ |
| Sum | ١١٣٦١٢.٨٢ | ١٠٤٤٥٥.٧٦ | ٦٨٣٦.٦٠ | ٩٥١١٥.٢٥ | ٥٩٤١٥ | ٥٧٤٤.٤٩ | ١٠١٧٥٧.٤٢ | ١٠٣٤٢٤.٠٤ | ٥٦٤٨.٢٤ |

| SIR | Wave ٤ | | | Total | | |
|--------|-----------|------------|----------|-----------|----------|----------|
| | I | R | D | I | R | D |
| Mean | ٣٦٨.٤٢ | ٣٦٤.٥٧ | ١٣.٨٠ | ٤٤٥.٩٤ | ٤٤٠.٦٣ | ٢٢.٣٦ |
| Median | ٠ | ٣٥٨ | ٠ | ١٦٩.٠١ | ٤١٢.١٣ | ١٣.٦٩٦ |
| Max | ٥٥١٦ | ٣٩٧٥.٨٩ | ١٣٩.٥١ | ٥٥١٦ | ٣٩٧٥.٨٩ | ١٣٩.٥١ |
| Min | ٠ | ٠ | ٠ | ٠ | ٠ | ٠ |
| Sd | ٦٦٤.٨٧ | ٣٤٣.١٧ | ٢٣.٠٢ | ٥٦٥.٢٤ | ٣٤٥.٨٨ | ٢٤.٥٨ |
| Sum | ٢١٢٢٠٩.٥٤ | ٢٠٩٩٩٤.٦٩٦ | ٧٩٤٩.١٩٦ | ٥٢٢١٩٢.١٦ | ٤٧٧٢٨٩.٥ | ٢٦١٧٨.٥٤ |

جدول (٤) يوضح الإحصاءات الوصفية المقدرة للمتغيرات بواسطة نموذج SIRD

ووفقا لتقديرات النموذج، يشير الجدول لسبق الي أن البيانات لا تتبع التوزيع لطبيعي كما أنها ملتوية لليمين، حيث نجد أن قيمة المتوسط أكبر من قيمة الوسيط بالنسبة للمتغير I، وهذا يشير الي أن هناك أياما قليلة تشهد ارتفاعا كبيرا في عدد الإصابات، مما رفع قيمة المتوسط عن الوسيط. كما نلاحظ ارتفاع قيمة المتوسط عن قيمة الوسيط بالنسبة للمتغير R، وهذا يعني أن توزيع حالات الشفاء ملتوتجاه اليمين، كما يشير إلى وجود أيام استثنائية قليلة شهدت ارتفاعا كبيرا في حالات لشفاء مقارنة بباقي الأيام، بينما تبقي أغلب الأيام عند مستويات تعاف أقل نسبيا، مما جعل المتوسط أكبر من الوسيط.

كما نجد ارتفاع قيمة الوسط عن الوسيط بالنسبة للمتغير D خلال الموجتين الأولى والرابعة وكذلك خلال فترة الوباء كاملة، أي يوجد التواء لليمين والعكس بالنسبة للموجتين الثانية والثالثة، يوجد يوجد التواء سلب.

ثانياً: بالنسبة لنموذج MLP

من خلال نتائج مجموعة التدريب تم تقدير المتغيرات (I, R, D) بشكل يومي من الفترة ١٤/٠٢/٢٠٢٠ حتى ٣٠/٠٤/٢٠٢٣. ومنه استنتجنا الإحصاءات الوصفية المقدرة للمتغيرات بولسطة نموذج MLP كما يلي:

| MLP | Wave ١ | | | Wave ٢ | | | Wave ٣ | | |
|--------|--------|--------|--------|---------|--------|-------|---------|---------|--------|
| | I | R | D | I | R | D | I | R | D |
| Mean | ٤٣٦.٧١ | ٣٨٧.٠٥ | ٢٢.٣٠ | ٨٠.٨.٥ | ٤٥٩ | ٤٢.٥ | ٦٢٣ | ٥٢٤ | ٣١ |
| Median | ٢٤٤.٥ | ٢٣٧ | ١٥ | ٨٠.٥.٢٧ | ٤٤٢.٨٦ | ٣٩.٤٢ | ٥٩٣.٩٢٢ | ٥٠٤.١٦٢ | ٢٨.٥٦٢ |
| Max | ١٦٦٩ | ١٢٦٩ | ٨١ | ١٣٠.٩ | ٧٦٩ | ٥٩ | ١٢٣٨ | ٨٩٧ | ٦١ |
| Min | ٥ | ٩٠ | ١ | ٣٦٠ | ١٧١ | ١٣ | ٠ | ١٣١ | ٢ |
| Sd | ٤٥٣.٢٦ | ٢٩١.٣١ | ٢٠.٦٩٨ | ٢٣١.٢١ | ١٣٩.٧١ | ١١.٩٢ | ٣٩٣.٨٤ | ٢٠٩.٠١ | ١٧.٦٢ |
| Sum | ١٢٣١٥٣ | ١٠٩١٤٩ | ٦٢٨٩ | ١٠٤٦٨٥ | ٥٧٥٧١ | ٥١٢٤ | ١٠٨٦٨٨ | ٩٢٢٦٢ | ٥٢٢٧ |

| MLP | Wave ٤ | | | Total | | |
|--------|--------|--------|-------|--------|--------|--------|
| | I | R | D | I | R | D |
| Mean | ٣٣٠.٠٣ | ٣٨٥ | ١٤.٣٤ | ٤٤٩.٦٢ | ٥٣٨.٧٦ | ٢٧.٥٣٥ |
| Median | ٥ | ٣٧٤.٢٤ | ٢ | ٢٣٩.٥ | ٤١٧ | ١٤ |
| Max | ٢٢٥٠ | ١٣٨٨ | ٩٥ | ٢٢٥٠ | ١٣٨٨ | ٩٥ |
| Min | ٠ | ٨٩ | ١ | ٠ | ٨٩ | ١ |
| Sd | ٥١٧.٨١ | ٢٢٤.٠٥ | ٢٠.١٦ | ٤٨٥.٣٥ | ٢٣٧.١٤ | ٢٠.٨٧ |
| Sum | ١٩٠٤٢٩ | ٢١٥٩٣٤ | ٨٢٧٢ | ٥٢٦٩٥٥ | ٤٧٤٩١٦ | ٢٤٩١٢ |

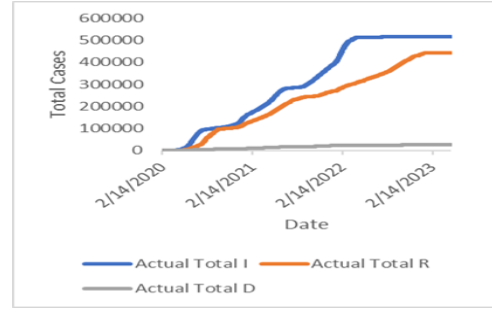
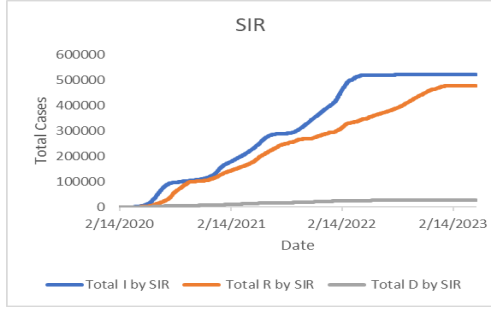
جدول (٥) يوضح الإحصاءات الوصفية المقدرة للمتغيرات بواسطة نموذج MLP

وفقاً لتقديرات النموذج يتضح أن البيانات لا تتبع التوزيع لطبيعي كما أنها ملتوية لليمين، وبمقارنة نتائج الإحصاءات الوصفية في هذا الجدول بنظيرها في جدول (٤) الخاص بنموذج SIR، وكذلك جدول (٢) الخاص بالقيم الفعلية، نجد أن نتائج نموذج MLP لا تخف كثيراً عن جدول (٢) وهذا يتوافق كثيراً مع الإحصاءات الوصفية للبيانات الفعلية، وعلى الرغم من عدم الاختلاف يبقى المعيار الأساسي في المقارنة بين النموذجين SIR، MLP هو مقياس دقة التنبؤ.

المقارنة بين النماذج

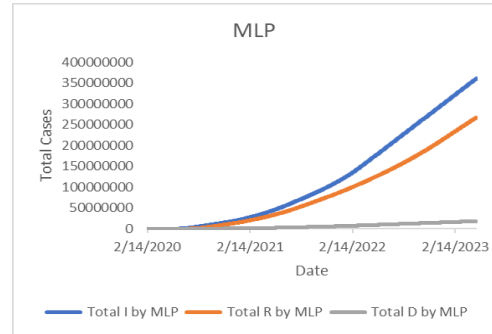
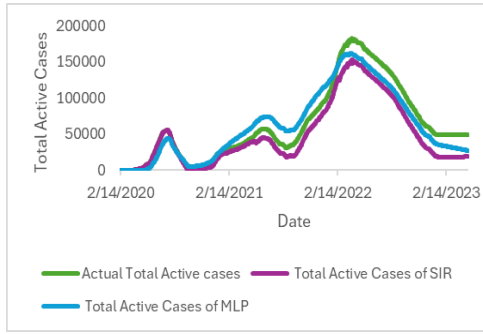
يعرض الشكل (١)، والشكل (٢)، والشكل (٣) الرسم البياني للقيم الفعلية والمقدرة والمتنبئ بها للمتغيرات الثلاثة (I, R, D) بولسطة نموذج SIR، ونموذج MLP. ومنه تم التنبؤ بعدد الحالات

الشطة كما هو موضح في الجدول (٦)، وكذلك يوضح الشكل (٤) الفرق بين لحالات الشطة الفعلية الفعلية والمقدرة بواسطة النماذج المستخدمة.



شكل (٢) إجمالي حالات الإصابة، والشفاء، والوفاة المقدرة بواسطة نموذج SIR لوباء COVID - ١٩ في مصر

شكل (١) إجمالي حالات الإصابة، والشفاء، والوفاة الفعلية لوباء COVID - ١٩ في مصر



شكل (٤) إجمالي الحالات النشطة الفعلية والمقدرة بواسطة نموذج SIR، ونموذج MLP لوباء COVID - ١٩ في مصر

شكل (٣) إجمالي حالات الإصابة، والشفاء، والوفاة المقدرة بواسطة نموذج MLP لوباء COVID - ١٩ في مصر

| Waves | Actual | | SIR | | MLP | |
|--------------|-------------------|------------------------|------------------------------|-----------------------------------|------------------------------|-----------------------------------|
| | Peak Active Cases | Peak Active Cases Date | Peak Active Cases Estimation | Peak Active Cases Date Estimation | Peak Active Cases Estimation | Peak Active Cases Date Estimation |
| A | ١٢٩٥ | ٦/١٩/٢٠٢٠ | ١٢٨٢.٤٤ | ٦/٢٠/٢٠٢٠ | ١٠٧١ | ٦/١٦/٢٠٢٠ |
| B | ٩٧٣ | ١٢/٢٥/٢٠٢٠ | ١٠٤٤.٥ | ١٢/٣١/٢٠٢٠ | ٨١٨ | ١٢/٢٦/٢٠٢٠ |
| C | ٨٥٤ | ٥/٢٥/٢٠٢١ | ١١٠١.٩٣ | ٩/٣٠/٢٠٢١ | ٦٧٢ | ٥/٩/٢٠٢١ |
| D | ٤٨٠.٣ | ٣/١٩/٢٠٢٢ | ٤٨٠.٣ | ٣/٢٠/٢٠٢٢ | ١٤٨٣ | ٣/٢٧/٢٠٢٢ |
| Total | ٤٩٠.٦٥ | | ١٨٧٢٤.١٢ | | ٢٧١٢٧ | |

جدول (٦) الحالات النشطة الفعلية والمتنبئ بها بواسطة نموذج MLP ونموذج SIR

يمكن حساب معايير قياس دقة التنبؤ من خلال القيم الفعلية والقيم المتنبئ بها للمتغيرات الثلاث وفقا لنموذج SIR ونموذج MLP. وبالنسبة لنموذج MLP يتم قياس دقة النموذج باستخدام معايير قياس دقة الأداء على مجموعة الاختبار، هذه البيانات التي لم يستخدمها النموذج أثناء التدريب، وبالتالي تعطى مؤشرا حقيقيا على أداء النموذج في التنبؤ ببيانات جديدة. أما جودة توفيق البيانات يتم قياسها من خلال مجموعة التدريب. كما هو موضح في الجدول التالي:

| Measures | SIR | | | MLP | | |
|----------|--------|----------|--------|--------|--------|--------|
| | I | R | D | I | R | D |
| R^2 | ٠.٧٥ | ٠.٥٧ | ٠.٨٤ | ٠.٩٨ | ٠.٩٧ | ٠.٩٤ |
| R | ٠.٨٦ | ٠.٧٥ | ٠.٩١ | ٠.٩٨٨ | ٠.٩٨٦ | ٠.٩٦٩٩ |
| MSE | ١٠٩.٠٥ | ١٣٨٤٣.٩٦ | ٢.٥١ | ٣٧.٧٩٦ | ٠.٠٠٠٧ | ٠.٢٢ |
| RMSE | ١٠.٤٤ | ١١٧.٦٦ | ١.٥٨ | ٦.١٥ | ٠.٠٢٦ | ٠.٤٧ |
| MAE | ١٤.١٧ | ٢٤.٤٩ | ١٨٤.٦٩ | ٣.٧٢ | ١٣.٣٣ | ٥.٥٦ |
| MAPE | ٧.٣٨ | ١٠١.٩٣ | ١.٤١ | ١.٣٧ | ٠.٠٠٣ | ٠.١١ |

جدول (٧) مقارنة معايير قياس دقة التنبؤ لنموذج SIR، ونموذج MLP

ونلاحظ من الجدول (٧) انخفاض معايير المقارنة (MAPE، RMSE، MSE، MEA) الخاصة بالخاصة بقياس خطأ ودقة التنبؤ لنموذج MLP مقارنة بنموذج SIR للثلاث متغيرات على الترتيب (I, R, D) مما يؤدي أفضلية نموذج MLP. وبالنسبة لقيمة معامل التحديد R^2 وهي خاصة بجودة توفيق البيانات، فجد أن نموذج MLP هو الأفضل من حيث جودة التوفيق لبيانات المتغيرات الثلاث وقيمته على الترتيب (٠.٩٨، ٠.٩٧، ٠.٩٤) (٠.٩٤) وهي الأقرب من ١٠٠٪ مقارنة بنموذج SIR حيث يضح من الجدول أن نموذج MLP هو هو الأفضل والأكثر ملائمة للبيانات.

ملحوظة: لشبكة العصبية نفسها لا يمكنها تقدير معدلات الإصابة والشفاء والوفاة، ولكنها تقوم بالتنبؤ بالمتغيرات (عدد الحالات)، وبالتالي لا يمكنها تقدير المعلمات الأخرى كرقم التكاثر الأساسي، وعتبة المناعة الجماعية، وتجاوز الحد.. الخ. بل تقوم بعملية تعلم غير خطي بين المدخلات والمخرجات من خلال عملية التدريب للبيانات التاريخية؛ هذا يعني أنها تقوم بنمذجة هذه البيانات بناء على الأنماط التي تجدها فيها دون استخدام معادلات صريحه، كما في النماذج الرياضية التقليدية، مما يجعل من الصعب ربط المخرجات بمعادلات معينة مثل معدل الإصابة والشفاء والوفاة، ولكن يمكن حساب هذه المعدلات يدويا بالاعتماد على نتائج لشبكة المستخرجة.

ومما سبق نستنتج أن الشبكات العصبية الاصطناعية قادرة على التنبؤ بالقيم التي يمكن استخدامها استخدامها لتقدير المعادلات يدويا، لكنها لا توفر معادلات رياضية تقليدية مثل تلك المستخدمة في

النماذج الرياضية (الإحصائية)، فبدلاً من تلك تقوم لشبكات العصبية بالتعلم من البيانات وتقديم تنبؤات تنبؤات بناء على الأنماط المعقدة التي تكتشفها، وللتغلب على هذا التحدي، يمكن دمج لشبكات العصبية مع العصبية مع الأساليب الإحصائية التقليدية لحساب المعدلات والمعلومات بشكل دقيق وتفسير النتائج بطريقة بطريقة أكثر دقة ووضوحاً.

أهم نتائج الدراسة:

يتم تقسيم نتائج الدراسة بناء على لطريقة المستخدمة حيث تناولت الدراسة أسلوبين للتنبؤ، هما أسلوب SIR، وأسلوب MLP وذلك بغرض التعرف على تأثير كل طريقة على نتائج الدراسة، وهدف التحليل هو مقارنة التقديرات وكذلك دقة وقدرة النموذجين على التنبؤ بعدد حالات الإصابة، والوفيات، والشفاء في فترات زمنية مختلفة بناء على البيانات المتاحة، وتوصلت الدراسة الي النتائج التالية:

(١) نموذج SIR هو نموذج رياضي يستخدم على نطاق واسع في تحليل انتشار الأمراض. يعتمد على المعادلات التفاضلية لتقسيم المجتمع إلى ثلاث فئات رئيسية هي: المعرضين للإصابة، وللصائين، والمتعافين.

(٢) على الرغم من قدرة نموذج SIR على تقديم صورة عامة عن انتشار الوباء، إلا أن نتائجه كانت أقل دقة في التنبؤ بالتقلبات اليومية في عدد لحالات. ومن الجدير بالذكر فهو يعتمد على افتراضات مبسطة حول معدلات الانتقال والشفاء، مما يجعله أقل مرونة في مواجهة تغييرات سريعة في طبيعة انتشار المرض.

(٣) عند مقارنة التوقعات بالبيانات الفعلية، أظهر نموذج SIR فروقات كبيرة بين القيم المتوقعة والقيم الفعلية، خاصة في الأيام التي شهدت زيادة حادة أو تراجعاً في عدد الإصابات .

(٤) تم استخدام نموذج MLP كأحد النماذج المعتمدة على لشبكات العصبية الاصطناعية، والتي يتعلم من البيانات التاريخية لتوقع انتشار الوباء.

(٥) يعد استخدام نماذج النكاء الاصطناعي مثل MLP خطوة مهمة نحو تحسين التنبؤات لصحية واتخاذ القرارات المستنيرة خلال الأوبئة.

(٦) أظهر نموذج MLP أداءً أفضل مقارنة بنموذج SIR في التنبؤ بأعداد حالات الإصابة، والوفيات، والشفاء. وتمكن نموذج MLP من محاكاة التقلبات اليومية بدقة أعلى، فضل قدرته على التعامل مع البيانات المتغيرة والأنماط غير الخطية في الانتشار.

(٧) علاوة على ذلك، أظهر نموذج MLP قدرة جيدة على التكيف مع الزيادات المفاجئة في عدد لحالات، مما يشير إلى أن النموذج يمكنه التعلم والتكيف مع أنماط انتشار الوباء المتغيرة.

٨) عند مقارنة أداء النموذجين باستخدام مؤشرات الدقة (MAPE، RMSE ، MSE ، MEA) الخاصة بقياس خطأ ودقة التنبؤ، تبين أن:

- نموذج MLP تفوق بشكل واضح على نموذج SIR في توقع لحالات الوباء اليومية.
- نموذج MLP أكثر قدرة على التقاط التعقيدات والتقلبات غير الخطية في البيانات، بينما اعتمد SIR على معادلات ثابتة لم تستطع التعامل مع جميع التغيرات المفاجئة في مسار انتشار الوباء.
- باستخدام نموذج MLP وجدنا أن قيم أخطاء التنبؤ الخاصة به أقل من قيمها لنموذج SIR، مما يعكس تحسنا كبيرا في دقة التوقعات وقدرته على التعامل مع البيانات المتغيرة والتقلبات اليومية في الإصابات الوفيات والشفاء.
- انخفاض قيم مؤشرات الدقة لنموذج SIR خلال فترة التحليل، مما يشير إلى محدودية قدرة النموذج على التنبؤ الدقيق في ظل التغيرات الديناميكية في أعداد الحالات.

توصيات ومقترحات الدراسة بناء على النتائج:

- بناء على ما تقدم من نتائج التحليل لوباء ١٩ - COVID نتقدم بالتوصيات والمقترحات التالية:
- (١) استخدام تقنيات الذكاء الاصطناعي في التقدير والتنبؤ ليعطي نتائج أفضل وتُصغّر حجم الأخطاء الى أقل قيمة لها.
 - (٢) نوصي باستخدام أسلوب لشبكات العصبية الاصطناعية وجميع أدوات الذكاء الاصطناعي لعمل تنبؤات طويلة الأجل، كبديل للنماذج الرياضية التي لا تصلح إلا للتنبؤات قصيرة الأجل وكذلك لتجنب الافتراضات المعقدة لتلك النماذج كثبات حجم المجتمع وتجاهل التغيرات الحيوية وتسليوي فرص الإصابة لجميع الأفراد وغيرها من الافتراضات غير المتوافقة مع الواقع، مع العلم إننا لا ننكر أهمية التنبؤات قصيرة الأجل.
 - (٣) بناء على نتائج التحليل، نوصي باستخدام نموذج MLP كأداة أكثر فعالية للتنبؤ بتطور الوباء في المستقبل في مصر، خاصة في الحالات التي تشهد تغييرات ديناميكية في أعداد الإصابات.
 - (٤) يمكن لنموذج MLP أن يكون جزءا من أنظمة التحنير المبكر لتوجيه القرارات المتعلقة بالإجراءات الاحترازية أو توزيع الموارد لصحية. فهو يمثل خطوة مهمة نحو تحسين التنبؤات لصحية واتخاذ القرارات المستنيرة خلال الأوبئة.
 - (٥) على الرغم من أن نموذج SIR قد يوفر فهما عاما لديناميكيات الوباء، إلا أنه يحتاج إلى تحسينات أو دمج مع نماذج أخرى مثل MLP لزيادة دقته.

٦ إجراء العديد من الدراسات باستخدام تقنيات الذكاء الاصطناعي والتعلم الآلي، والتي ستساعد العديد من الأشخاص في العديد من القطاعات وخاصة القطاع الصحي على الاستعداد الجيد واتخاذ الاحتياطات اللازمة لتقليل انتشار العديد من الأمراض والأوبئة.

٧ نوصي بضرورة تدريس النماذج الرياضية للأمراض المعدية وتقنيات الذكاء الاصطناعي في قسم الإحصاء حيث إن الذكاء الاصطناعي حاليا أثبت كفاءته مقارنة بالعديد من النماذج الرياضية والإحصائية؛ لذا يمكن لقسم الإحصاء الاستفادة من ذلك في تدريس لطلاب كلا المنهجين الأمر الذي يؤدي إلى ابتكار العديد من النماذج الممنجة.

٨ نظرا لطبيعة الديناميكية للأوبئة، فإننا دائما بحاجة الي نماذج مبتكرة؛ لعل تنبؤات قصيرة، ومتوسطة، وطويلة الأجل. ونماذج الذكاء الاصطناعي أثبتت جودتها، ولكننا دائما نطمع فيما هو أفضل، ولكن استخدام أساليب AI للتنبؤ لا يمكنه التقاط نمط تغير الوقت لانتشار الأمراض المعدية، ولعل هذه المشكلة فضل استخدام النماذج الممنجة أي دمج النماذج الرياضية بنماذج الذكاء الاصطناعي، كاستخدام نموذج SEIRVD – DL، التي يجمع بين النموذج الرياضي SEIRVD، وأحد نماذج التعلم العميق Deep Learning، وتقدير والتنبؤ بالمعلومات الوبائية في النموذج الرياضي باستخدام أحد نماذج DL مثل نموذج شبكة LSTM (long – short term memory network) بدلا من طرق التقدير التقليدية للتقدير وللتنبؤ مثل: OLS، MLE. فلقد وجد أنه منلب للتنبؤات قصيرة ومتوسطة وطويلة الأجل، وهذا أكثر عملية لصياغة لسياسات والتخطيط الاستراتيجي، وذلك باستخدام البرامج الإحصائية ولغات البرمجة، أهمها في هذا الصدد: R, MATLAB, Python

(Wang, X., Lin, X., & Dang, X., ٢٠٢٠)،(Liao, Zhifang, et al., ٢٠٢١).

٩ نوصي بإنشاء قنوات اتصال فعالة بين مراكز تحليل البيانات لصحية والهيئات الحكومية والمؤسسات لصحية، لتنفيذ توقعات الأبحاث المطبقة مما يضمن اتخاذ القرارات لصحية بناء على بيانات دقيقة وموثوقة، مما يقلل من التأخر في الاستجابة للأزمات لصحية.

١٠ مواصلة تطوير وتحديث نماذج التنبؤ؛ لضمان دقة أكبر وقدرة على التكيف مع المتغيرات ومرونتها في مواجهة التحديات المستقبلية.

المراجع

أولا: المراجع باللغة العربية:

(١) خضري، حنان (٢٠٢٣) ، "مقارنة الشبكات العصبية ذات الذاكرة قصيرة المدى المطولة ونماذج بوكس وجينكز"، المجلة العلمية للبحوث التجارية (جامعة المنوفية)، ٤٩(٢)، ٥١٧-٥٤٢.

(٢) سهاد علي شهيد التميمي، رواء صالح محمد (٢٠٢٠)، "النماذج الوبائية نو منظومة المعادلات التفاضلية المتعددة"، ١٢٦، ٣٥٨-٣٦٢، *Journal of Administration and Economics*

(٣) الشرقاوي، ولاء واخرون (٢٠٢١)، "نهاية الموجة الثالثة لفيروس كورونا في مصر" دراسة استكشافية (الإصدار الثالث)، لجهاز المركزي للتعبئة العامة والإحصاء في مصر.

(٤) لشوادفي، جمال أحمد & حجاج، عبد الوهاب السيد (٢٠١٣). "الذكاء الاصطناعي وتحليل لسلاسل الزمنية". المجلة العلمية لقطاع كليات التجارة بجامعة الأزهر (١)، ١٠، ٤٢-١.

ثانياً: المراجع باللغة الإنجليزية:

- ١) Mohajan, H. (٢٠٢٢). "Mathematical analysis of SIR model for COVID-١٩ transmission.", ١-١٨.
- ٢) Jakhar, M., Ahluwalia, P. K., & Kumar, A. (٢٠٢٠). COVID-١٩ epidemic forecast in different states of India using Sir model. Medrxiv, ٢٠٢٠-٠٥.
- ٣) Wang, X., Ren, H., & Wang, A. (٢٠٢٢). Smish: A novel activation function for deep learning methods. Electronics, ١١(٤), ٥٤٠.
- ٤) Paul, Amit, et al. "An introduction to Compartmental Model and its use in Stata." (٢٠٢٢).
- ٥) Katris, C. (٢٠٢١). A time series-based statistical approach for outbreak spread forecasting: Application of COVID-١٩ in Greece. Expert systems with applications, ١٦٦, ١١٤٠٧٧.
- ٦) Randolph, H. E., & Barreiro, L. B. (٢٠٢٠). Herd immunity: understanding COVID-١٩. Immunity, ٥٢(٥), ٧٣٧-٧٤١.
- ٧) Nguyen, M. M. (٢٠٢٤). Upper bounds on overshoot in SIR models with nonlinear incidence. npj Complexity, ١(١), ١١.
- ٨) Sebbagh, A., & Kechida, S. (٢٠٢٢). "EKF-SIRD model algorithm for predicting the coronavirus (COVID-١٩) spreading dynamics." Scientific Reports ١٢(١), ١٣٤١٥.
- ٩) Azeem, Muhammad, et al. (٢٠٢٣). "Neural Networks for the Detection of COVID-١٩ and Other Diseases: Prospects and Challenges." Bioengineering, ١٠(٧), ٨٥٠.
- ١٠) Cao, H., Wu, H., & Wang, X. (٢٠٢٠). Bifurcation analysis of a discrete SIR epidemic model with constant recovery. Advances in Difference Equations, ٢٠٢٠(١), ٤٩.

۱۱) Liao, Zhifang, et al. (۲۰۲۱). "SIRVD-DL: A COVID-۱۹ deep learning prediction model based on time-dependent SIRVD." *Computers in Biology and Medicine*, ۱۳۸, ۱۰۴۸۶۸.

