

إستخدام نماذج **ARIMA** والشبكات
العصبية الاصطناعية فى التنبؤ بمؤشر
EGX30 سوق المال المصرى

بحث مقدم من

صفاء محمد على مصطفى

مدرس مساعد بقسم الإحصاء والرياضيات والتأمين

كلية التجارة – جامعة بورسعيد

ملخص البحث

يعد التنبؤ بمؤشرات الأسواق المالية من التقنيات المهمة في إتخاذ القرارات الإستثمارية حيث أنه يوفر الأدوات اللازمة لتحقيق الربح وتعظيمه أو لتفادي الخسارة المتوقعة .

يهدف هذا البحث إلى إستخدام أسلوب الشبكات العصبية كأحد التقنيات الحديثة وإستخدام منهجية **Box – Jenkins** متمثلة في نماذج **ARIMA** في التنبؤ بالقيم المستقبلية لمؤشر البورصة المصرية الرئيسي **EGX30** وأيضاً الدمج بين هذين الأسلوبين للحصول على أفضل النتائج الممكنة .

وإعتمد البحث على سلسلة بيانات يومية للمؤشر **EGX30** في الفترة من ٢٠١٤/٦/١ وحتى ٢٠١٧/٣/٢٩ باستثناء أيام الجمع والعطلات .

وخلص البحث إلى أن الدمج بين كلاً من أسلوب الشبكات العصبية ونماذج **ARIMA** يعطى أفضل نتائج تنبؤ وفقاً لمقاييس التنبؤ وأهمها معيار **MSE** .

Abstract

The Research Involves with an important sector in the Egyptian market which is locally and internationally influential, The Egyptian Stock Exchange where the stock exchange is considered as a mirror that reflects the state of the national economy

This Research aims to using Neural Networks method and ARIMA models and Combining the models of (ARIMA) with neural networks to Forecast the series of the Egyptian stock exchange main index (EGX30) using a Daily series From1/6/2014 to29/3/2017.

The Research proofs that Combining Neural Networks with ARIMA models was the best method to analyzing and forecasting the data series.

(١) مشكلة البحث :

تعتبر الأسواق المالية أداة هامة لعمليات التمويل المباشر للإقتصاد وذلك من خلال البورصة التي تمثل همزة الوصل بين المدخرين والمستثمرين لذا تظهر الحاجة إلى تتبع تطورات هذه الأسواق عن طريق دراسة مؤشراتنا ومحاولة التنبؤ بها بالإعتماد على مجموعة من الأساليب الكمية ، وعملية التنبؤ في السلاسل الزمنية تتأثر بشكل مباشر باختيار النموذج المناسب لبيانات السلسلة الزمنية والسلاسل الزمنية المالية تعاني من تقلبات كثيرة يجب أخذها في الإعتبار وهنا تظهر مشكلة البحث المتمثلة في محاولة معرفة أكثر النماذج ملائمة للتنبؤ بمؤشرات الأسواق المالية من بين نماذج **ARIMA** والشبكات العصبية .

(٢) أهداف البحث :

يهدف البحث إلى تحقيق مايلي :

- ١- استخدام نماذج **ARIMA** والشبكات العصبية الإصطناعية **ANN** في التنبؤ بسلسلة مؤشر **EGX30** .
- ٢- الدمج بين كلا من نماذج **ARIMA** والشبكات العصبية للتنبؤ بسلسلة المؤشر .
- ٣- المقارنة بين تنبؤات الطرق الثلاث المستخدمة في البحث في محاولة للتوصل إلى أفضلها .

(٣) أهمية البحث :

الأهمية العلمية : من حيث استخدام البحث لبعض أهم الأساليب الإحصائية المتقدمة وهي أسلوب الشبكات العصبية و نماذج **ARIMA** و برامج الحاسب الآلى المتخصصة مثل برنامج **EViews** وبرنامج **MATLAB** .

الأهمية العملية : من حيث مجال تطبيق البحث وهو قطاع هام بالسوق المصرى ومؤثر محلياً و دولياً ألا وهي البورصة المصرية حيث تعد بورصة الأوراق المالية المرآة التي تعكس حالة الإقتصاد القومى وقد تعددت الدراسات المتعلقة بالأسواق المالية عالمياً فيما ندرت عربياً.

(٤) فروض البحث :

يقوم البحث على الفروض التالية :

- ١- تعتبر الشبكات العصبية أفضل من نماذج ARIMA في التنبؤ بمؤشرات الأسواق المالية.
- ٢- الدمج بين الشبكات العصبية ونماذج ARIMA يؤدي إلى أفضل النتائج وأدق التنبؤات.

(٥) حدود البحث :

- المتغير محل البحث هو سلسلة بيانات المؤشر الرئيسي للبورصة المصرية (EGX30) و الذى يقيس أداء أعلى ثلاثين شركة من حيث السيولة والنشاط فقط.
- تم الإعتماد على البيانات التاريخية للبورصة المصرية وإستخدام سلسلة بيانات يومية للمؤشر (EGX30) فى الفترة من ٢٠١٤/٦/١ إلى ٢٠١٧/٣/٢٩ باستثناء ايام العطل أى باستخدام ٦٩٦ مشاهدة .

(٦) الدراسات السابقة :

• الخياط (٢٠٠٥)

" إستخدام الشبكات العصبية فى التكهّن بالسلسلة الزمنية لأستهلاك الطاقة الكهربائية فى مدينة الموصل "

هدفت هذه الدراسة إلى إستخدام أسلوب الشبكات العصبية فى التنبؤ بإستهلاك الطاقة الكهربائية فى مدينة الموصل ومقارنته بأسلوب Box-Jenkins كطريقة للتنبؤ وخلصت الدراسة إلى تفوق أسلوب الشبكات العصبية فى التنبؤ على أسلوب Box-Jenkins والمتمثل فى نموذج . ARIMA(2,1)

• جبارة (٢٠١٢)

" التنبؤ بالسلاسل الزمنية لمنسوب النيل الأزرق بإستخدام نماذج بوكس وجنكنز ونماذج الشبكات العصبية "

هدفت هذه الدراسة إلى معرفة مدى كفاءة نماذج (Box-Jenkins) والشبكات العصبية فى التنبؤ بالسلاسل الزمنية لمناسيب النيل والمفاضلة بين هذه النماذج .

وتم بناء نموذج $ARIMA(1,1,0)$ وقد تم التأكد من أن هذا النموذج جيد ويعطي تنبؤات دقيقة وقريبة من الواقع كما تم بناء نماذج الشبكات العصبية والذي تكونت بنيته المعمارية من ثلاث طبقات (طبقة مدخلات (2) ، طبقة خفية (5) وطبقة مخرجات (1)) وإتضح أن الشبكة الناتجة جيدة وأعطت تنبؤات دقيقة وقريبة من الواقع وخلصت الدراسة إلى كفاءة الأسلوبين في التعامل مع السلاسل الزمنية ولكن تفضل نماذج الشبكات على نماذج Box-Jenkins .

• رملى (٢٠١٣)

" المفاضلة بين أسلوب Box-Jenkins وأسلوب الشبكات العصبية الإصطناعية فى التنبؤ بحجم المبيعات فى المؤسسة الاقتصادية (دراسة حالة المؤسسة الجزائرية GIPLAIT)"

هدفت هذه الدراسة إلى تطبيق أسلوب السلاسل الزمنية من خلال نموذج Box-Jenkins وأسلوب الشبكات العصبية الإصطناعية فى التنبؤ بحجم المبيعات فى المؤسسة الاقتصادية GIPLAIT ومن ثم المقارنة بين النموذجين وخلصت الدراسة إلى كفاءة أسلوب الشبكات العصبية الإصطناعية فى التحليل والتنبؤ إذ أعطى أقل قيمة لجذر متوسط مربع الخطأ .

• تلمسانى (٢٠١٤)

" نمذجة قياسية لتطايير سعر الصرف الدينار الجزائرى بالنسبة لعملة الدولار الأمريكى باستعمال نماذج ARCH و ARIMA"

هدفت هذه الدراسة إلى بناء نموذج قياسي يسمح بدراسة تقلب سعر صرف الدينار الجزائرى مقابل الدولار الأمريكى والمقارنة بين نماذج ARIMA ونماذج ARCH فى التنبؤ بأسعار الصرف وخلصت الدراسة إلى عدم صلاحية نموذج ARIMA للتنبؤ وأن النموذج المقبول لتمثيل التباين الشرطى لسعر الصرف هو النموذج $GARCH(0,2)$.

• أبو عابدة (٢٠١٥)

" إستخدام الطرق الإحصائية فى التنبؤ بأسعار الذهب العالمية "

إستخدمت نماذج ARIMA وأسلوب الشبكات العصبية للتنبؤ بأسعار الذهب وخلصت إلى أن النموذج الأمثل هو نموذج $ARIMA(2,1,1)$ وذلك بعد أخذ الفروق الأولى للوغاريتم السلسلة لتثبيت التباين وأشارت النتائج إلى أن نموذج الشبكات العصبية هو النموذج الأفضل والذي تم الإعتماد عليه فى التنبؤ بالقيم المستقبلية لسلسلة أسعار الذهب العالمية .

(٧) الجانب النظرى للبحث :

(١-٧) منهجية Box-Jenkins : [٩, ١١]

وتعتمد فى تحليل السلاسل الزمنية على توظيف تاريخ السلسلة الزمنية وسلسلة الأخطاء للتوصل لوصف ملائم لنمط تغير البيانات من خلال أربع مراحل رئيسية تتمثل فيما يلي :

(١-١-٧) مرحلة التعرف على النموذج Model Identification

يتم فيها اختيار نموذج ملائم لوصف السلسلة الزمنية من بين مجموعة نماذج (ARIMA) و يقوم أسلوب Box-Jenkins فى على افتراض رئيسي وهو سكون السلسلة الزمنية

- التحقق من سكون السلسلة

من خلال رسم المنحنى الزمني للسلسلة الزمنية Time-Plot وفحص كلاً من دالة الارتباط الذاتي (ACF) Auto Correlation Function و دالة الارتباط الذاتي الجزئي Partial Auto Correlation Function (PACF) كما يمكن الحكم على إستقرار السلسلة الزمنية بإجراء بعض الإختبارات مثل إختبار ديكي فولر.

- النماذج المستخدمة فى أسلوب Box-Jenkins [٩, ٦, ١١]

- نموذج الإنحدار الذاتى (AR) Autoregressive Models :

وفيه يعتمد المتغير التابع y_t على القيم السابقة حتى الفترة P كالتالى:

$$y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \delta + \varepsilon_t$$

- ϕ_i : معاملات الإنحدار الذاتى و P : رتبة النموذج

- نموذج المتوسطات المتحركة (MA) Moving Average Model :

و يأخذ النموذج الصورة التالية:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

θ_j : معاملات نموذج المتوسطات المتحركة و q : رتبة النموذج

Autoregressive & Moving Average (ARMA p,q)

هو تركيبة من النموذجين السابقين و يعبر عن السلسلة الزمنية y_t كدالة خطية في كل من المتغيرات $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ والأخطاء $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots$ كالتالي :

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \delta + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

نموذج الإنحدار الذاتي والمتوسطات المتحركة التكاملية

Autoregressive Integrated Moving Average (ARIMA p,d,q)

النموذج يكون غير مستقر وإزالة عدم الإستقرار يتم أخذ الفروق بدرجات مختلفة حتى الوصول إلى سلسلة مستقرة وعندئذ تصبح السلسلة تكاملية من الدرجة d والصيغة العامة لهذا النموذج تأخذ الشكل التالي :

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \delta + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

وبعد الحصول على النماذج المقترحة لتمثيل السلسلة الزمنية يمكن عمل مفاضلة بينها باستخدام معيار **Ackaike (AIC)** والنموذج الأفضل هو الذي يأخذ أقل قيمة

(٢-١-٧) مرحلة تقدير النموذج Model Estimation

ويتم فيها تقدير معاملات النموذج المقترح الذي تم اختياره من نماذج **ARIMA** وتستخدم طريقة المربعات الصغرى **Least Squares Metho** لسهولةا والخصائص الجيدة التي تتمتع بها لأنها تعتمد على اختيار مقدرات المعالم التي تجعل مجموع مربعات الأخطاء اقل ما يمكن .

(٣-١-٧) مرحلة فحص النموذج Model Diagnostic Checking

وفيها يتم اختبار ملاءمة النموذج المقدر قبل التنبؤ حيث يتم التحقق من عشوائية البواقي الخاصة به و يعتبر النموذج نموذجاً جيداً إذا كانت الأخطاء عشوائية (متوسطها الصفر وتباينها

ثابت وغير مرتبطة زمنياً) وإذا كان النموذج ملائم يستخدم في المرحلة التالية (مرحلة التنبؤ) أما إذا كان النموذج غير ملائم يجب تحديد نموذج آخر وإعادة تقديره واختباره مره أخرى.

Forecasting (٧-١-٤) مرحلة التنبؤ

يمثل التنبؤ المرحلة الرابعة والأخيرة من مراحل تحليل السلسلة الزمنية باستخدام أسلوب Box-Jenkins وفيها يتم التنبؤ بالقيم المستقبلية للسلسلة .

(٧-٢) الشبكات العصبية

انظمة تكنولوجية تقوم بتشغيل المعلومات من استقبال للمدخلات واجراء عمليات تشغيل ونتاج المخرجات اعتماداً على مجموعة من الدوال الرياضية التي تقوم بأدق الوظائف بأسلوب يحاكي الشبكات العصبية الطبيعية لدى الانسان فتساعد متخذ القرار على اتخاذ قرارات عالية الدقة .

(٧-٢-١) خطوات التحليل والتنبؤ باستخدام الشبكات العصبية : [٩, ١٢, ٤]

إن التنبؤ باستخدام الشبكات العصبية من الأساليب الحديثة التي لاقت إهتماماً واسعاً في مختلف المجالات واستخدمت بشكل واسع لكونها لا تحتاج إلى شروط صارمة ودقيقة للتنبؤ ويمكن تلخيص عمل الشبكة العصبية وفقاً لطريقة الإنتشار الخلفي في الخطوات التالية :

الخطوة الأولى : إختيار المتغيرات Variable Selection

يجب تحديد المتغيرات وإختيار المشاهدات بحيث تمثل المشكلة محل الدراسة تمثيلاً جيداً .

الخطوة الثانية : معالجة البيانات Data Processing

فيها يتم تحديد طريقة المعالجة من بين ثلاثة إختيارات متاحة للسلاسل الزمنية وهي :

- التنبؤ بقيم السلسلة اعتماداً على القيم السابقة لها وعلى القيم السابقة لسلسلة متغير خارجي آخر (NARX)

Nonlinear Autoregressive with External(exogenous)Input

- التنبؤ بقيم السلسلة اعتماداً على القيم السابقة لها فقط (NAR)

Nonlinear

Autoregressive

- التنبؤ بقيم السلسلة اعتماداً على القيم السابقة لمتغير آخر خارجي فقط (NIO)

Nonlinear

Input

Output

الخطوة الثالثة : تقسيم البيانات إلى مجموعات Divide Data into Sets

- ١- مجموعة التدريب Training Set مجموعة تعلم وتحديد نموذج البيانات
- ٢- مجموعة التدقيق Validation Set مجموعة الحكم على مهارة الشبكة
- ٣- مجموعة الإختبار Testing Set وهي مجموعة لإجراء إختبار نهائي للشبكة

الخطوة الرابعة : تصميم نموذج الشبكة العصبية Neural Network Design

عند تحديد نموذج الشبكة العصبية يجب إختيار مايلي :

- عدد النيورونز للإدخال والذي يساوي عدد المتغيرات المستقلة
- عدد الطبقات الخفية والذي يعتمد على التجربة ودرجة تعقيد المشكلة
- عدد النيورونز في الطبقة الخفية والذي يتحدد عن طريق التجربة
- نيورون الإخراج والذي عادة يساوي واحد
- تحديد درجة التأخر عند تعديل الأوزان

(٧-٢-٢) طريقة الإنتشار الخلفي Back Propagation : [٩, ١٦]

ترجع تسميتها الى طريقة التعلم التي تعتمد على مبدأ تصحيح الخطأ وتبدأ هذه الطريقة بايجاد الفرق بين المخرج المطلوب والفعلي وترجع بهذا الخطأ ارتدادياً من الطبقة الأخيرة الى الطبقات الخفية ثم أخيراً الى طبقة المدخلات وفي أثناء هذا الإرتداد يتم تغيير الأوزان في الاتجاه الذي يدفع بالخطأ الى النقصان وصولاً الى الصفر وتتم على مرحلتين كما يلي :

المرحلة الأولى : مرحلة الانتشار الأمامي :

وتبدأ بإعطاء أوزان عشوائية للترابط بين خلايا الشبكة ومد الشبكة بإحدى المدخلات المعدة للتدريب بطريقة التدريب الإشرافي وتنتقل البيانات من طبقة المدخلات للمعالجة في الطبقات الخفية وصولاً لتحديد مخرجات الشبكة ولا يحدث فيها أى تعديل للأوزان وتكون

$$v_j = \sum_{i=1}^n w_{ji}^h x_i \quad \text{مخرجات طبقة المدخلات هي المدخلات للطبقة الخفية كما يلي:}$$

$$z_j = f_j^h(v_j) = f_j^h(\sum_{i=1}^n w_{ji}^h x_i) \quad \text{أما مخرجاتها فتصبح :}$$

$$\text{net}^o = \sum_{j=1}^m w_{kj}^o z_j \quad \text{وتكون المدخلات لطبقة المخرجات كما يلي:}$$

$$y = f_k^o(\sum_{j=1}^m w_{kj}^o (f_j^h(\sum_{i=1}^n w_{ji}^h x_i))) \quad \text{أما مخرجاتها فتصبح :}$$

المرحلة الثانية : مرحلة الانتشار العكسي (الخلفي) :

هي مرحلة ضبط أوزان الشبكة ويطلق عليها الإنحدار التدريجي (Gradient descent) حيث يتم فيها إعادة إنتشار الإشارة من الخرج إلى الدخل بشكل عكسي ويتم ذلك ابتداء من طبقة المخرجات وتبدأ بمقارنة المخرجات الفعلية (y_k) مع المخرجات المطلوبة (t_k) و ذلك عند التكرار (n) و تحديد قيمة الخطأ كما يلي :

$$e_k(n) = t_k(n) - y_k(n)$$

وتكون دالة الخطأ كالتالى وذلك إتماداً على مجموع مربعات الخطأ :

$$E(n) = \frac{1}{p} \sum_{k=1}^p e_k^2(n)$$

وبعد حساب دالة الخطأ تبدأ الخطوة الثانية للانتشار الخلفي وينقل الخطأ للطبقات السابقة ويتم تصحيح الأوزان وتكون الأوزان الجديدة المنقولة من طبقة المخرجات للطبقة الخفية على

$$w_{kj}^o(n+1) = w_{kj}^o(n) + \eta \delta_k(n) z_j(n) \quad \text{الصورة :}$$

وبالمثل تكون الأوزان الجديدة المنقولة من الطبقة الخفية لطبقة المدخلات على الصورة :

$$w_{ji}^h(n+1) = w_{ji}^h(n) + \eta (f_j^h(v_j(n)) \sum_k \delta_k(n) w_{kj}^o(n))$$

ويتم تطبيق هذه الخطوات على كل مدخلات التدريب ولمرات عديدة حتى تصل الشبكة الى أقل خطأ ممكن وتصحح الأوزان بمايدفع بالخطأ للنقصان وصولاً للصفر وعندها تصبح الشبكة جاهزة.

الخطوة الخامسة : معيار التقييم Evaluation Criteria

المعيار المستخدم فى شبكة الإنتشار العكسى لتقييم الخطأ هو مجموع مربعات الخطأ MSE

الخطوة السادسة : تدريب الشبكة Neural Network Training

وفىها يتم تعليم الشبكة وإيجاد مجموعة الأوزان والتي تحدد أقل قيمة لمربع الخطأ

الخطوة السابعة : الإستكمال Implementation

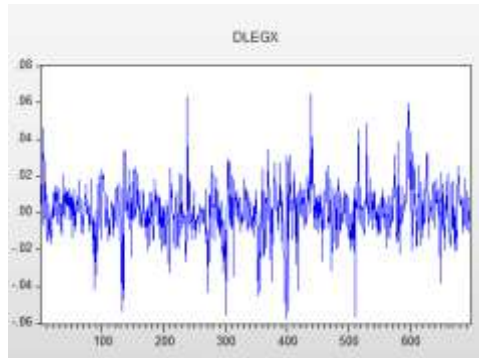
وهى من أهم الخطوات حيث تختبر الشبكة من حيث قدرة التكيف وإمكانية إعادة التدريب والوصول إلى أقل مربع خطأ عند تغير البيانات .

(٨) الجانب التطبيقى للبحث :

(٨-١) التحليل باستخدام نماذج ARIMA

(٨-١-١) رسم السلسلة الجديدة وإختبار إستقرارها

بعد رسم المنحنى الزمنى للسلسلة EGX (ملحق ١) إتضح عدم إستقرارها لذا تم أخذ الفروق الأولى للسلسلة لتثبيت الوسط وتطبيق التحويلة اللوغاريتمية لتثبيت التباين والشكل (١) يوضح المنحنى الزمنى للسلسلة الجديدة DLEGX ومن الرسم يتضح سكونها .



شكل (١) المنحنى الزمنى للسلسلة DLEGX

وقد تم التأكد من سكون السلسلة DLEGX باستخدام إختبار ديكى فولر المطور وكانت قيمة

$p > 0.05$ لذلك نرفض فرض العدم القائل بأن السلسلة غير ساكنة أي أنها ساكنة (ملحق ١)

(٨-١-٢) مرحلة التعرف على النموذج

تم إقتراح ثمانية من نماذج ARIMA (ملحق ١) و بمقارنة قيمة t-statistic لها تم ترشيح نموذجين فقط هما الأقرب لتمثيل السلسلة والجدول (٢) يوضح المقارنة بينهما من حيث معايير المفاضلة بين النماذج :

جدول (١) معايير المفاضلة بين نماذج ARIMA المرشحة لسلسلة DLEGX

model	AIC	SC	HQC
ARIMA(1,1,0)	-5.570887	-5.550798	-5.563108
ARIMA(0,1,1)	-5.569605	-5.549516	-5.561826

ومما سبق يتضح أن النموذج المناسب لوصف السلسلة الزمنية هو نموذج ARIMA(1,1,0)

(٨-١-٣) مرحلة تقدير النموذج

نموذج ARIMA(1,1,0) المقدر يكون في الصورة التالية :

Dependent Variable: DLEGX
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 10/18/17 Time: 11:24
Sample: 2 675
Included observations: 674
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000617	0.000782	0.806056	0.4188
AR(1)	0.248126	0.032027	7.747512	0.0000
SIGMSQ	0.000221	8.20E-08	26.94047	0.0000

R-squared	0.061644	Mean dependent var	0.000613
Adjusted R-squared	0.058847	S.D. dependent var	0.015355
S.E. of regression	0.014896	Akaike info criterion	-5.570887
Sum squared resid	0.148893	Schwarz criterion	-5.550798
Log likelihood	1880.389	Hannan-Quinn criter.	-5.563108
F-statistic	22.04034	Durbin-Watson stat	1.994821
Prob(F-statistic)	0.000000		

Inverted AR Roots .25

جدول (٢) نموذج ARIMA(1,1,0) المقدر

ويكون النموذج الأمثل على الصورة : $\hat{y}_t = 0.0006 + 0.2481 y_{t-1}$

(٨-١-٤) مرحلة إختبار بواقى النموذج

أولاً : التحقق من عشوائية بواقى النموذج وذلك وفقاً لما يلي :

(١) فحص دالتي الارتباط الذاتي والارتباط الذاتي الجزئي للبواقى

بالنظر إلى شكل كل من دالتي الارتباط الذاتي والارتباط الذاتي الجزئي لبواقي النموذج والموضح في جدول (٣) يتم التأكد من عشوائية البواقي حيث أن جميع القيم تقع داخل حدى الثقة مما يعنى أنه لا يوجد ارتباط ذاتى بين الأخطاء وبالتالي فإن النموذج ملائم .

جدول (٣) دالتي الارتباط الذاتي والارتباط الذاتي الجزئي لبواقي النموذج المقدر ARIMA(1,1,0)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1		0.002	0.002	0.0025	
2		-0.023	-0.023	0.3766	0.539
3		0.049	0.049	2.0135	0.366
4		0.060	0.059	4.4335	0.218
5		0.023	0.025	4.7860	0.310
6		-0.058	-0.058	6.9467	0.225
7		0.036	0.032	7.8528	0.249
8		-0.008	-0.017	7.8993	0.342
9		-0.053	-0.049	9.8093	0.279
10		0.024	0.026	10.193	0.335
11		0.024	0.022	10.601	0.389
12		-0.022	-0.020	10.935	0.449
13		-0.015	-0.006	11.094	0.521
14		-0.015	-0.021	11.253	0.580
15		-0.014	-0.022	11.398	0.656
16		-0.002	0.006	11.400	0.724
17		-0.029	-0.026	11.989	0.745
18		-0.015	-0.017	12.153	0.791
19		0.007	0.012	12.183	0.838
20		0.074	0.077	15.955	0.660

(٢) إختبار Ljung-box

بمقارنة قيمة Q المحسوبة وهي آخر قيمة فى العمود Q- stat فى جدول (٣) بالقيمة الجدولية لتوزيع مربع كاي والذي يعتمد عليه هذا الإختبار عند مستوى معنوية 0.05 ودرجات حريه 19 نجد أن $Q = 15.955 < \chi^2_{(19,0.05)} = 30.14$ ومنه نقبل فرض العدم الذى يقر بانعدام جميع معاملات الارتباط الذاتى أى أن سلسلة البواقي مستقرة ولا يوجد ارتباط ذاتى بين الأخطاء.

ثانياً : إختبار (تجانس) ثبات تباين الأخطاء باستخدام إختبار أثر ARCH

بما أن قيمة الاحتمال P المناظرة ل NR^2 والموضحة فى جدول (٤) أكبر من 0.05 و بما أن قيمة NR^2 والتي تساوى 2.459 أقل من قيمة $\chi^2_{(0.05,1)}$ والتي تساوى 3.84 نقبل فرض العدم H_0 الذى ينص على تجانس الأخطاء أى أنه لا توجد مشكلة Heteroscedasticity

جدول (٤) نتائج أثر ARCH لبواقي النموذج المقدر ARIMA(1,1,0)

Heteroskedasticity Test: ARCH			
F-statistic	2.461325	Prob. F(1,671)	0.1172
Obs*R-squared	2.459639	Prob. Chi-Square(1)	0.1168

(٥-١-٨) مرحلة التنبؤ

جاءت قيم التنبؤ من المشاهدة ٦٧٦ وحتى المشاهدة ٦٩٦ وعددها ٢١ مشاهدة وفقاً لنموذج $ARIMA(1,1,0)$ كما هو موضح فى (ملحق ١) ووفقاً للنموذج المقدر $ARIMA(1,1,0)$ جاءت مقاييس التنبؤ كماهى موضحة فى جدول (٥)

جدول (٥) نتائج التنبؤ باستخدام نموذج $ARIMA(1,1,0)$

MSE	0.964
RMSE	0.855
MAE	0.822
MAPE	6.385

(٢-٨) التحليل باستخدام أسلوب الشبكات العصبية (ANN)

(١-٢-٨) تحديد البيانات (المتغيرات)

الخطوة الأولى فى تحليل الشبكات هى تحديد المتغيرات وهى هنا سلسلة EGX محل الدراسة وهى السلسلة كاملة لتدريب الشبكة وعددها ٦٩٦ مشاهدة وبعد الإنتهاء من التدريب سيتم إيجاد القيم المقدرة للجزء المراد التنبؤ به فى نهاية السلسلة وعدده ٢١ مشاهدة.

(٢-٢-٨) معالجة البيانات

حيث أنه تتوفر لدينا سلسلة زمنية واحدة فقط وهى المعنية بالدراسة ويراد توقع القيم المستقبلية للسلسلة الزمنية $Y(t)$ والذى يعتمد على القيم الماضية من هذه السلسلة فقط فإن هذا الشكل من أشكال الإنحدار الذاتى غير الخطى (NAR) $Non\ linear$ Autoregressive

(٣-٢-٨) تقسيم البيانات إلى المجاميع

فى هذه الخطوة سيتم تقسيم المدخلات عشوائياً إلى ثلاثة مجموعات :

المجموعة الأولى : التدريب Training

المستخدمة فى تدريب الشبكة وتحتوى على ٧٠% من بيانات السلسلة وعددها ٤٨٨ مشاهدة

• المجموعة الثانية : الفعالية Validation

المستخدمة للتحقق من تعميم الشبكة وتحتوى على ١٥% من البيانات وعددها ١٠٤ مشاهدة

• المجموعة الثالثة : الإختبار Testing

المستخدمة كإختبار مستقل للشبكة وتحتوى على ١٥% من البيانات وعددها ١٠٤ مشاهدة

(٨-٢-٤) بناء الشبكة العصبية

فى هذه الخطوة يتم تحديد نموذج أو بنية الشبكة العصبية كالتالى :

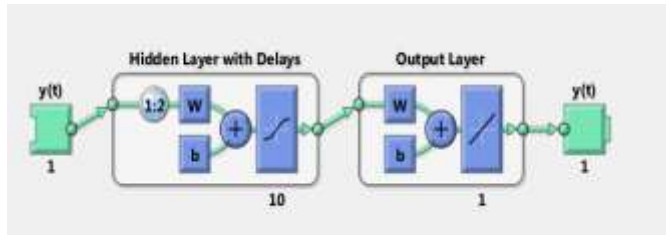
- عدد النيورونز الخاصة بالإدخال والذي يساوى عدد المتغيرات المستقلة والذي يساوى فى هذه الشبكة الواحد (١)

- عدد الطبقات الخفية والذي يعتمد على التجربة و يمكن التحكم فيه حسب نتائج الشبكة عند إعادة التدريب والتي حددت آلياً (إفتراضياً) بطبقة واحدة .

- عدد النيورونز فى الطبقة الخفية والذي يعتمد على التجربة و يمكن التحكم فيه حسب نتائج الشبكة عند إعادة التدريب والذي تم تحديده (١٢)

- النيورون الخاص بالإخراج والذي يساوى واحد (١)

- تحديد درجة التأخر فى حالة التغذية العكسية لتعديل وتحديث الأوزان فى الشبكة وقد تم تحديده آلياً (٢) والشكل (١) يوضح بناء الشبكة



شكل (٢) بنية الشبكة العصبية NN

(٥-٢-٨) تدريب الشبكة والتنفيذ

- يتم تدريب الشبكة تدريباً إشرافياً وباستخدام طريقة الانتشار الخلفي للخطأ من خلال الدالة (Trainlm) أو (Livenberg Marquardt Backprobagation) وهي الدالة الأسرع والأكثر استخداماً في الشبكات العصبية لتعليم وتدريب الشبكة وتتضمن محددات التدريب مثل معدل التعلم ودالة التنشيط المستخدمة والعدد الأقصى من الدورات وغيرها ويتم حساب مجموعة الأوزان بين النيورونز والتي تحدد أقل قيمة لمتوسط مربع الخطأ MSE (ملحق ٢).

- إختبار الشبكة من حيث قدرة التكيف وإمكانية إعادة التدريب للوصول إلى أقل مربع خطأ عند تغير البيانات وفي حالة عدم الحصول على النتائج المرضية يتم إعادة التدريب عدة مرات كما يمكن التغيير في بنية الشبكة حتى يتم الوصول للنتيجة المثلى والتي تضمن أقل قيمة لمتوسط مربع الخطأ.

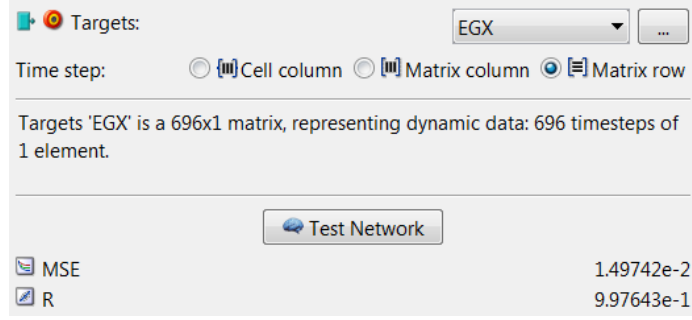
(٦-٢-٨) التنبؤ وفقاً للشبكة NN

تم الحصول على النتائج النهائية للشبكة كما هي موضحة في شكل (٣) وكانت مقاييس التنبؤ بقيم السلسلة EGX اعتماداً على النتائج النهائية للشبكة العصبية NN كما هي موضحة في جدول (٦)

جدول (٦) مقاييس التنبؤ باستخدام الشبكات العصبية NN

MSE	0.0149
RMSE	0.122
MAE	0.0723
MAPE	0.0017

شكل (٣) النتائج النهائية للشبكة



كما تم الحصول على قيم التنبؤات وفقاً للشبكة العصبية كما هي موضحة في (ملحق ٢)

(٣-٨) الدمج بين الشبكات العصبية و نموذج ARIMA

(١-٣-٨) إدخال البيانات

تم إدخال سلسلة البيانات في صورة سلسلتين الأولى وهي سلسلة المدخلات Input وهي سلسلة الأخطاء التي تم الحصول عليها من تحليل نموذج ARIMA وعددها ٦٩٥ مشاهدة والسلسلة الثانية هي سلسلة الهدف Target وهي بيانات السلسلة الأصلية المراد التنبؤ بها .

(٢-٣-٨) معالجة البيانات

حيث أنه تتوفر لدينا سلسلتين زمنييتين ويراد توقع القيم المستقبلية للسلسلة الزمنية $Y(t)$ والذي يعتمد على القيم الماضية من هذه السلسلة (الأصلية) بالإعتماد على سلسلة زمنية أخرى $X(t)$ وهي سلسلة الأخطاء فإن هذا الشكل يعد من أشكال التنبؤ الإنحدار الذاتي غير الخطى (NARX) مع متغير خارجي كما يوضح الشكل

Non linear Autoregressive With External Input

(٣-٣-٨) تقسيم البيانات إلى المجاميع

في هذه الخطوة سيتم تقسيم المدخلات عشوائياً إلى ثلاثة مجموعات الأولى ٧٠% للتدريب وعددها ٤٨٧ مشاهدة والثانية ١٥% وعددها ١٠٤ مشاهدة للفعالية والثالثة ١٥% وعددها ١٠٤ مشاهدة للاختبار .

ARIMA – NN العصبية (٤-٣-٨)

فى هذه الخطوة يتم تحديد نموذج أو بنية الشبكة العصبية ARIMA – NN كالتالى :

- عدد النيورونز الخاصة بالإدخال والذى يساوى عدد المتغيرات المستقلة والذى يساوى فى هذه الشبكة الواحد (٢)

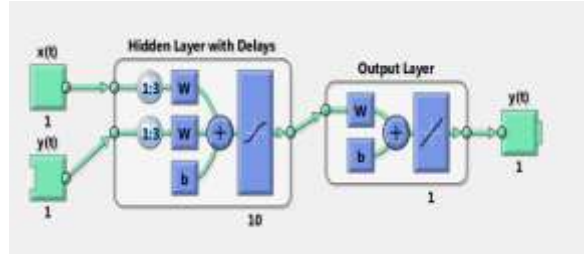
- عدد الطبقات الخفية والذى يعتمد على التجربة و يمكن التحكم فيه حسب نتائج الشبكة عند إعادة التدريب والتي حددت آلياً (إفتراضياً) بطبقة واحدة .

- عدد النيورونز فى الطبقة الخفية والذى يعتمد على التجربة و يمكن التحكم فيه حسب نتائج الشبكة عند إعادة التدريب والذى حدد آلياً (١٠)

- النيورون الخاص بالإخراج والذى يساوى واحد (١)

- تحديد درجة التأخر فى حالة التغذية العكسية لتعديل وتحديث الأوزان فى الشبكة وقد تم تحديده (٣) إلا أنه قد يتم تغيير هذا الرقم إذا كان أداء الشبكة غير مرضى وذلك عند إعادة التدريب

والشكل (٤) يوضح بناء الشبكة ARIMA – NN



شكل (٤) بنية الشبكة العصبية ARIMA – NN

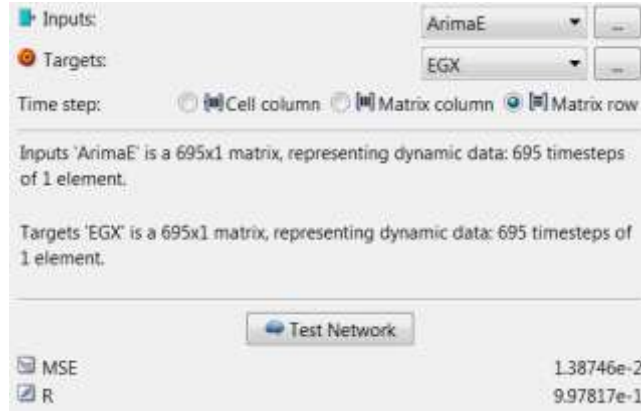
(٥-٣-٨) تدريب الشبكة

يتم تدريب الشبكة تدريباً إشرافياً وباستخدام طريقة الإنتشار الخلفى للخطأ

ARIMA – NN (١-٥-٣-٨) نتائج الشبكة

بعد عدة مرات من التدريب تم التوصل للنتيجة التالية كما هو موضح فى شكل (٥) حيث نجد

MSE=0.0138 للتنبؤ بقيمة معامل الارتباط بين القيم المستهدفة والمخرجات R = 0.99



شكل (٥) النتائج النهائية للشبكة ARIMA –NN

ARIMA –NN إختبار أخطاء الشبكة (٢-٥-٣-٨)

أولاً : إختبار تجانس أخطاء الشبكة

ومن الشكل (٦) يتضح عدم وجود مشكلة عدم تجانس الأخطاء حيث كانت نتيجة إختبار أرش

لأخطاء التنبؤ ARIMA –NN هي Logical 0

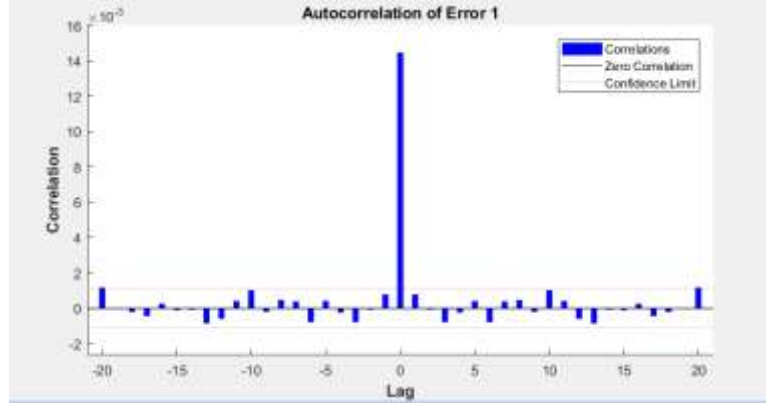
```
>> h=archtest(res)
h =
logical
0
```

شكل (٦) نتيجة إختبار ARCH لأخطاء الشبكة ARIMA –NN

ثانياً : إختبار الارتباط الذاتي للأخطاء

والشكل (٧) يوضح شكل دالة الارتباط الذاتي للأخطاء بعد إعادة تدريب الشبكة و من الملاحظ

أن كل الحدود تقع داخل مجال الثقة مما يدل على غياب الارتباط الذاتي بين الأخطاء



شكل (٧) دالة الارتباط الذاتي لأخطاء الشبكة ARIMA –NN

(٣-٥-٣-٨) التنبؤ:

جاءت قيم التنبؤ وفقاً للشبكة ARIMA –NN كما هو موضح في (ملحق ٢) كما تم الحصول على مقاييس التنبؤ بقيم السلسلة EGX اعتماداً على نتائج الشبكة ARIMA –NN كما هي موضحة في جدول (٧)

جدول (٧) مقاييس التنبؤ وفقاً للشبكة ARIMA –NN

MSE	0.0138
RMSE	0.117
MAE	0.0941
MAPE	0.0018

(٩) نتائج البحث:

بناءً على الدراسة التطبيقية في الفصل السابق توصلت الدراسة للنتائج التالية:

(١) لتحليل سلسلة بيانات مؤشر البورصة المصرية EGX 30 تم الاعتماد على منهجية Box-Jenkins وأخذ الفروق الأولى للسلسلة وتم ترشيح عدة نماذج ARIMA بدرجات مختلفة من (p,q) وتم تقدير النموذج الأفضل لتمثيل السلسلة وهو نموذج ARIMA (1,1,0) وباختبار أخطاء النموذج ثبت ملاءمته للسلسلة وتم التنبؤ.

(٢) باستخدام الشبكات العصبية الاصطناعية كطريقة حديثة من طرق الذكاء الاصطناعي قد تم بناء الشبكة ANN لتحليل السلسلة محل الدراسة ثم التنبؤ وأعطت نتائج مرضية بالمقارنة بنتائج نموذج ARIMA (1,1,0)

(٣) تم دمج نموذج أريما المقدر ونمذجة الأخطاء باستخدام الشبكات العصبية حيث تم تكوين الشبكة ARIMA - NN ثم التنبؤ .

(٤) تم مقارنة الطرق الثلاثة السابقة في التنبؤ بالسلسلة EGX كما هو موضح في الجدول (٨)

جدول (٨) المقارنة بين طرق التنبؤ بالسلسلة EGX

Model	ARIMA (1,1,0)	الشبكة NN	الشبكة ARIMA - NN
MSE	0.964	0.0149	0.0138
RMSE	0.855	0.122	0.117
MAE	0.822	0.0723	0.0941
MAPE	6.385	0.0017	0.0018

(٥) أفضل الطرق للتنبؤ هي باستخدام الدمج بين الشبكات العصبية ونموذج ARIMA حيث حققت أقل قيمة لمتوسط مربع الخطأ يليها طريقة الشبكات يليها نموذج ARIMA (1,1,0) .

(٦) إتضح ان أسلوب الشبكات العصبية الإصطناعية سواء بمفرده أو بالدمج مع اساليب أخرى هو الأفضل من حيث سهولة التحليل ودقة التنبؤ .

(١٠) التوصيات :

بناءً على ما توصل إليه البحث من نتائج فإنه يوصى بما يلي :

- استخدام الشبكات العصبية لدراسة سلاسل البيانات المالية والتي تتسم بالتقلبات ولما لها من خصائص تميزها عن باقي أنواع السلاسل الزمنية .
- استخدام طرق الدمج لأكثر من أسلوب إحصائي معاً لتحسين دقة التنبؤ .
- يراعى استخدام عينات كبيرة من البيانات عند استخدام الشبكات العصبية لأنه يحدث فقد بسيط لبعض البيانات أثناء عملية التدريب .

أولا : المراجع العربية

- ١- أبو عابدة ، ألفت فتحي سالم (٢٠١٥)
إستخدام الطرق الإحصائية فى التنبؤ بأسعار الذهب العالمية ، غزة ، جامعة الأزهر.
- ٢- الحسينى , قصى حبيب و الساعدى , أحمد عبد الأمير (٢٠٠٨)
مقدمة فى الشبكات العصبية الإصطناعية ، العراق ، بغداد، جامعة الإمام جعفر الصادق
- ٣- الخياط ، باسل و زكى ، عزة (٢٠٠٥)
إستخدام الشبكات العصبية فى التكهّن بالسلسلة الزمنية لأستهلاك الطاقة الكهربائية فى مدينة الموصل، العراق ، جامعة الموصل ، المجلة العراقية للعلوم الإحصائية ، العدد(٨)
- ٤- الشرقاوى , محمد على (١٩٩٦)
الذكاء الإصطناعى والشبكات العصبية، مصر ، الاسكندرية، مطابع المكتب المصرى الحديث.
- ٥- موقع البورصة المصرية <http://www.egx.com.eg/arabic/homepage.aspx>
- ٦- بن أحمد ، أحمد (٢٠٠٧) ،
النمذجة القياسية للإستهلاك الوطنى للطاقة الكهربائية فى الجزائر خلال الفترة (١٩٨٨-٢٠٠٧)، الجزائر ، جامعة الجزائر.
- ٧- تلمسانى ، حنان و زدون ، جمال (٢٠١٤)
نمذجة قياسية لتطير سعر الصرف الدينار الجزائرى بالنسبة لعملة الدولار الأمريكى بإستعمال نماذج ARIMA و ARCH ، الجزائر، تلمسان ، جامعة أبى بكر بلقايد .
- ٨- جبارة ، محمد جلال (٢٠١٢)
التنبؤ بالسلاسل الزمنية لمنسوب النيل الأزرق بإستخدام نماذج بوكس وجنكنزون نماذج الشبكات العصبية ، جامعة السودان للعلوم والتكنولوجيا .
- ٩- رملى ، محمد وآخرون (٢٠١٣)
المفاضلة بين أسلوب Box-Jenkins وأسلوب الشبكات العصبية الإصطناعية فى التنبؤ بحجم المبيعات فى المؤسسة الإقتصادية (دراسة حالة المؤسسة الجزائرية GIPLAIT)، جامعة سعيدة ، كلية العلوم الإقتصادية والتجارية.
- ١٠ شعراوى ، سمير مصطفى (٢٠٠٥) ،
مقدمة فى التحليل الحديث للسلاسل الزمنية ، مركز النشر العلمى ، السعودية ، جدة
جامعة الملك عبد العزيز

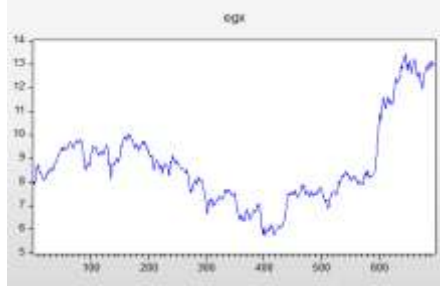
ثانيا : المراجع الاجنبية

- 11- Anderson, Dave and McNeill, George (1992)
Artificial Neural Networks Technology , Kaman
Sciences Corporation , Utica , New York , USA.
- 12- Box, George.E.P, Jenkins, Gwilym.M, Reinsel,
George.C(2008)
Time Series Analysis Forecasting and Control(fourth
edition)
A John Wiley & Sons Inc publication , New Jersey , USA.
- 13- B.Yengnanarayana(2005)
Artificial Neural Networks , Indian Institute of
Technology, Madras.
- 14- Colin, Fife(2000)
Artificial Neural Networks and Information
Theory, University of Paisley, Scotland .
- 15- Graupe, Daniel (2006).
Principles of Artificial Neural Networks , University of
Illinois , Chicago, USA.
- 16- Hamed , Rania Ahmed (2010) .
Enhancing the Efficiency of Forecasting Using Time
Series Models and Neural Networks , Port Said
University , Egypt.
- 17- Kriesel, David (2005).
A Brief Introduction to Neural Networks , University
of Bonn , Germany .
- 18- Yim , Juliana (2002)
A Comparison of Neural Networks With Time Series

**Models For Forecasting Returns On A Stock Market
Index ,RMIT University ,Australia .**

الملاحق

ملحق ١



المنحنى الزمني للسلسلة الزمنية EGX

إختبارات ديكي فولر المطور على السلسلة DLEGX

Exogenous	Test Critical Value (5%)	t-statistic	Prob
Intercept	-2.865488	-20.39022	0.0000
Intercept,Trend	-3.416306	-20.43230	0.0000
None	-1.616392	-20.37163	0.0000

قيم التنبؤ وفقاً لنموذج ARIMA (1,1,0)

Number	Date	Actual Value	Forecasting
1	1/3/2017	11.999	11.922
2	2/3/2017	12.31	11.923
3	5/3/2017	12.5	11.929
4	6/3/2017	12.623	11.936
5	7/3/2017	12.678	11.944
6	8/3/2017	12.735	11.951
7	9/3/2017	12.853	11.958
8	12/3/2017	12.92	11.966
9	13/3/2017	12.929	11.973
10	14/3/2017	12.79	11.981
11	15/3/2017	12.745	11.988
12	16/3/2017	12.984	11.995
13	19/3/2017	13.92	12.003
14	20/3/2017	13.23	12.010
15	21/3/2017	12.905	12.018
16	22/3/2017	12.879	12.025
17	23/3/2017	13.032	12.032
18	26/3/2017	13.108	12.040
19	27/3/2017	12.97	12.047
20	28/3/2017	12.983	12.055
21	29/3/2017	12.988	12.062

ملحق ٢

`trainlm` is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization.

`trainlm` is often the fastest backpropagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms.

`net.trainFcn = 'trainlm'` sets the network `trainFcn` property.

`[net,tr] = train(net,...)` trains the network with `trainlm`.

Training occurs according to `trainlm` training parameters, shown here with their default values:

<code>net.trainParam.epochs</code>	1000	Maximum number of epochs to train
<code>net.trainParam.goal</code>	0	Performance goal
<code>net.trainParam.max_fail</code>	6	Maximum validation failures
<code>net.trainParam.min_grad</code>	1e-7	Minimum performance gradient
<code>net.trainParam.mu</code>	0.001	Initial mu
<code>net.trainParam.mu_dec</code>	0.1	mu decrease factor
<code>net.trainParam.mu_inc</code>	10	mu increase factor
<code>net.trainParam.mu_max</code>	1e10	Maximum mu
<code>net.trainParam.show</code>	25	Epochs between displays (NaN for no displays)
<code>net.trainParam.showCommandLine</code>	false	Generate command-line output
<code>net.trainParam.showWindow</code>	true	Show training GUI
<code>net.trainParam.time</code>	inf	Maximum time to train in seconds

جدول قيم التنبؤ وفقاً لطريقتي الشبكات والدمج

Number	Date	Actual Value	Forecasting ARIMA –NN	Forecasting NN
1	1/3/2017	11.999	---	---
2	2/3/2017	12.31	---	---
3	5/3/2017	12.5	---	12.468
4	6/3/2017	12.623	---	12.557
5	7/3/2017	12.678	12.542	12.653
6	8/3/2017	12.735	12.708	12.690
7	9/3/2017	12.853	12.785	12.750
8	12/3/2017	12.92	12.936	12.882
9	13/3/2017	12.929	12.968	12.932
10	14/3/2017	12.79	12.951	12.927
11	15/3/2017	12.745	12.742	12.763
12	16/3/2017	12.984	12.745	12.737
13	19/3/2017	13.92	12.948	13.033
14	20/3/2017	13.23	13.159	13.091
15	21/3/2017	12.905	12.984	12.993
16	22/3/2017	12.879	12.879	12.878
17	23/3/2017	13.032	12.823	12.873
18	26/3/2017	13.108	12.972	13.052
19	27/3/2017	12.97	13.111	13.097
20	28/3/2017	12.983	12.935	12.933
21	29/3/2017	12.988	13.006	12.977