مقارنة بضعة مقدرات لاخطية لتقدير دالة الانحدار

أ. م. د. مناف يوسف حمود أ. م. مروان عبد الحميد عاشور
جامعة بغداد- كلية الادارة والاقتصاد- قسم الاحصاء

المستخلص

تركز هذا البحث على تقدير دالة الانحدار المتمثل بتطبيق عملي يشير الى كمية صادرات النفط السعودي كدالة بدلالة عدد حقول النفط المكتشفة فيها ،اذ تم اخذ السعودية ودراسة صادراتها كونها تعد من احدى كبريات الدول المصدرة للنفط ومن الدول المستقرة سياسيا وماليا وامنيا في منطقة الخليج العربي والشرق الاوسط .

ولغرض دراسة سلوك البيانات تبين ان السلوك الناتج لم يكن متبعا نمطا خطيا معروفا ولم تكن هنالك امكانية لمعرفة الاتجاه العام الناتج عن تلك الصادرات ،اذ لوحظ من خلال رسم البيانات وجود نمطا غير خطيا مما قاد هذا الى استخدام مقدرات لاخطية لامعلميه ،شبه معلميه فضلا عن مقدر يسمى بمقدر الشبكة العصبية الاصطناعية.

وبينت النتائج ان افضل مقدر اعطى تقديرا وافيا وكافيا وعبرعن الظاهرة المدروسة بشكل سليم هو مقدر الشبكة العصبية الاصطناعي الذي اثبت تفوقه على المقدرات اللامعلميه وشبه المعلميه المستخدمة في هذا البحث.

مفاتيح البحث: دوال الانحدار، المقدر الخطي الموضعي، المقدر شبه المعلمي المدمج، مقدر الشبكة العصبية الاصطناعية، الشبكات العصبية ذات الانتشار العكسي للخطأ

Comparing Several Nonlinear Estimators for Regression Function

Abstract

The aim of this paper is to estimate a nonlinear regression function of the Export of the crude oil Saudi (in Million Barrels) as a function of the number of discovered fields.

Through studying the behavior of the data we show that its behavior was not followed a linear pattern or can put it in a known form so far there was no possibility to see a general trend resulting from such exports.

We use different nonlinear estimators to estimate a regression function, Local linear estimator, Semi-parametric as well as an artificial neural network estimator (ANN).

The results proved that the (ANN) estimator is the best nonlinear estimator among the others in estimating the export of crude oil Saudi.

Key Words: regression function, Local Linear estimator, Combined semi-parametric estimator, Artificial neural network estimator (ANN), Back propagation Neural Nets.

1- المقدمة

يعرف تحليل الانحدار بأنه تحليل العلاقات بين المتغيرات وغالبا ما تكون العلاقات بين المتغيرات خطية، اذ ان صفة الخطية لها دورا مزدوجا فهي بالامكان ان تعبر عن وصف لحقيقة كون العلاقة بين متغير الاستجابة Y والمتغير التوضيحي X علاقة خطية كذلك تشير ايضا الى معالم انموذج الانحدار تكون خطية بغض النظر عن العلاقة بين X وY سواء اكانت تربيعية ،تكعيبية ... الخ.

ان تحليل الانحدار يبدأ عادة بالرسم البياني للبواقي مقابل المتغير التوضيحي X وكذلك مقابل القيمة التقديرية $\hat{Y}$ ،اذ ان هذه الرسوم البيانية تعد نقاط البداية للتحقق من فرضيات الانموذج او نقصها مثل ملائمة ام عدم ملائمة فرضية الخط المستقيم ،اختلال ثبات التباين (عدم تجانسه) ووجود قيم شاذة وكذلك في حالة كون الاخطا مرتبطة.

فضلا عن هذا وفي حالة كون الانموذج المستخدم معرف سواء اكان خطيا ام لاخطيا فأنه هنالك عدد من الاساليب والطرائق لحل تلك النماذج ،اذ في حالة كون الانموذج المستخدم هو انموذجا لاخطيا فأن احدى اساليب الحل لمثل هكذا انموذج تكون بعمل تحويل للمتغيرات كي يصبح بعدها الانموذج خطيا ويتم تقديره وفق طرائق التقدير الخطية.

لكن المشكلة التي قد تظهر عندما لاتكون هنالك صيغة معلميه معرفة بمعالم سواء اكانت المعالم خطية ام لاخطية او كانت المتغيرات التوضيحية خطية ام لاخطية ،عندها يتم اللجوء الى طرائق بديلة تستخدم مع النماذج التي لا تحتوي معالم وليس لها صيغة دالية محددة او معروفة،

ان هدف هذا البحث هو تقدير دوال الانحدار اللاخطية والمتمثلة ببيانات حقيقية تشير الى كمية الصادرات من النفط الخام السعودي كدالة بدلالة عدد الابار المكتشفة للسنوات من 1981-2000 وباستخدام مقدرات مختلفة منها مقدرلامعلمي متمثلا بمقدر الانحدار الخطي الموضعي ، واخر شبه معلمي وكذلك مقدر يدعى بمقدر الشبكة العصبية الاصطناعية .

2- المقدرات اللاخطية لدالة الانحدار:

2-1 مقدر الانحدار الخطي الموضعي

ان افتراض كون العلاقة بين كلا من متغير الاستجابة والمتغيرات التوضيحية هي علاقة خطية لا تخلو من عيوب ومشاكل اذ ان الافتراضات الخاصة بنماذج الانحدار المعلميه كأن تكون العلاقة خطية وتوزيع متغير الاستجابة هو توزيع طبيعي فضلا عن استقلالية المشاهدات غالبا ما تكون بعيدة عن الواقع مثال ذلك عندما تكون البيانات عبارة عن سلسلة زمنية تكون عندها الاخطاء غير مستقلة كذلك قد يكون التباين الشرطي لمتغير الاستجابة غير ثابت وان التوزيع الشرطي لهذا المتغير ربما لا يكون طبيعيا وقد يكون ملتويا او ثقيل الذيل Skewed or Heavy tailed .

عندها يكون الانحدار اللامعلمي بديلا جيدا عن الانحدار الخطي المعلمي مما قد يعطي تقديرا اكثر دقة لدالة الانحدار.ان الانحدار اللامعلمي يفترض عدم تحديد الصيغة الخاصة بدالة الانحدار مسبقا على الرغم من ان هذا الافتراض له مساوئ كون بعض المتغيرات ذات الاهتمام والمطلوب دراستها يكون لها متغيرات مؤثرة عليها لكن ليس بالشرط ان هذه العلاقة بين تلك المتغيرات مع متغير الاستجابة هي علاقة خطية.

ان كثير من المقدرات اللامعلمية لدالة الانحدار تعتمد على عرض حزمة وان كبر هذه الحزمة ينتج عنه بالمقابل تباينا صغيرا وتحيزا كبيرا والعكس صحيح. وعلى الرغم من ان المقدرات اللامعلميه تكون متحيزة بصورة عامة لكنها تكون متسقة طالما ان دالة الانحدار للمجتمع تكون تمهيدية، لذلك فان الحاجة تكون بتقليص عرض الحزمة وجعلها مقاربة للصفر مع تزايد حجم العينة مما ينتج ان كلا من مربع التحيز والتباين للمقدر اللامعلمي لدلة الانحدار يقتربان للصفر، اي ان

$$bias\ (\hat{m}(x)) \xrightarrow[n\to\infty]{} 0$$
$$Var\ (\hat{m}(x)) \xrightarrow[n\to\infty]{} 0 \qquad (1)$$

مما ينتج ان:

$$MSE(\hat{m}(x)) \xrightarrow[n\to\infty]{} 0 \qquad (2)$$

ان كثير من المقدرات اللامعلمية والتي لها دورا في تحليل البيانات تعتمد على المعدلات الموضعية وتعود فكرة التموضع واستخدام المعدلات الموضعية الى كون دالة الانحدار هي دالة تمهيدية وان المشاهدات لقيم X حول النقطة ذات الاهتمام والمراد تقديرها x دون الحاجة الى المشاهدات الاخرى للمتغير X علما ان عملية التموضع يتم السيطرة عليها من خلال استخدام نافذة يعبر عنها بدالة تدعى بدالة النافذة **Window function** اذ تعمل هذه النافذة على الاحاطة فقط بالقيم القريبة والمؤثرة على المشاهدة ذات الاهتمام x.

وهكذا ولجميع المشاهدات للدالة يتم التقدير وفق نفس اسلوب التموضع بحيث يتم اخذ جميع النقاط المرتبة وذات الاهتمام وبأستخدام دالة نافذة لكل مشاهدة مما ينتج عنها تقدير لدالة الانحدار او منحنى الانحدار الكلي.

وهنالك نوعان لدوال النافذة ،يعتمد النوع الاول على استخدام نافذة ذات عرض ثابت يتمركز عند النقطة المراد تقديرها x، اما النوع الاخر فيكون من خلال تعديل عرض النافذة ليتضمن عدد ثابت من المشاهدات فقط كأن يكون m من المشاهدات فقط سوف تؤخذ بنظر الاعتبار ويدعى عندها المقدر بمقدر التجاور الاقرب m **(m-nearest neighbor)** للنقطة x .

ويعد مقدر اللب **Kernel** او ما يسمى بالمعدل الموضعي الموزون كتعميم لفكرة المعدل الموضعي وتستند فكرة هذا المقدر على اعطاء اوزان اكبر للمشاهدات القريبة من النقطة ذات الاهتمام المراد تقديرها x مع استخدام ما يسمى عرض حزمة او معلمة تمهيدية كبديل عن عرض النافذة المستخدم مع المعدل الموضعي.وهنالك عدة دوال لب تستخدم بشكل شائع لكن الدوال الاكثر استعمالا هما دالتي **Gaussian Kernel** و **Tricube Kernel** على التوالي :

$$K(Z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{Z^2}{2})$$
$$K(Z) = (1-|Z|^3)^3 \quad , \quad |Z| < 1 \qquad (3)$$

وقد قدم كلا من الباحث Fan (1993) وكذلك Fan and Gijbels (1992،1996)[14][15] مقدرا يعد من افضل المقدرات المستخدمة في تحليل الانحدار اللامعلمي وهو ما يسمى بالانحدار متعدد الحدود الخطي الموضعي ،اذ يعمل هذا المقدر على تصحيح بعض العيوب الموجودة في مقدرات اللب مع امكانية التعميم لحالة الانحدار المتعدد فضلا عن الانحدار اللامعلمي العام ويعمل هذا الاسلوب على تعميم مقدر اللب الى حالة موائمة متعدد الحدود عند النقطة x وباستخدام دوال لب موضعية $W = K\left(\frac{X_i - x}{h}\right)$

وباستخدام انحدار المربعات الصغرى الموزونة لموائمة المعادلة

$$Y = a + b_1(X_i - x) + b_2(X_i - x)^2 + ... + b_p(X_i - x)^p + e_i \qquad (4)$$

لتقليل مجموع مربعات البواقي الموزونة $\sum_{i=1}^{n} W_i e_i^2$ وان الحل الناتج عن المربعات الصغرى الموزونة يكون: $\left(\hat{Y} \mid x\right) = a$ ،علما ان قيمة p=1 سوف تشير الى المقدر الخطي الموضعي والذي تم اثبات افضليته على بقية المقدرات الاخرى سواء اكانت p=0 ام p=2 او p=3 وكما في Fan عام (1993) وكذلك Fan and Gijbels عامي (1992،1996) [14][15]

والمقدر الخطي الموضعي المستخدم في هذا البحث هو الاتي :

$$\hat{m}(x) = (nh)^{-1} \sum_{i=1}^{n} K(\frac{x - X_i}{h})[S_{n,2} - (x - X_i)S_{n,1}]Y_i \qquad (5)$$

اذ ان:

$$S_{n,l} = h^{-1} \sum_{i=1}^{n} K(\frac{x - X_i}{h})(x - X_i)^l \ , \qquad l = 1,2 \qquad (6)$$

2-2 المقدر المدمج شبه المعلمي:

تعود بدايات استخدام هذا المقدر الى عام (1987) اذ درس الباحثان Olkin and Spieglman [21] استخدام تقدير هجين خطي يدمج بين كلا من الصيغة المعلميه واللامعلميه لدالة الكثافة الاحتمالية.بعدها وبالتحديد عام 1992 درس الباحثان Burman and Chaudhuri [4][5] استخدام هذا الخليط بين المقدرات المعلميه واللامعلميه لتقدير دالة الانحدار بعدها توالت الاقتراحات الخاصة بهذا النوع من المقدرات من قبل عدد من الباحثين منهم Wooldridge عام 1992 [4] و Rahman and Gokhale and Ullah عام 1993 [22] والباحثة اسيل مسلم والباحث مناف يوسف عامي 2011 و 2012 على التوالي [4][8] .

ان الهدف الرئيس من استخدام هذه الطريقة يتمثل بمحاولة تمثيل الانموذج للبيانات بالشكل الصحيح او حتى قريب من الصحة ومحاولة الابتعاد عن عدم تمثيل المجتمع تمثيلا غير امثل .ان هذه الطريقة تعتمد على كون انموذج الانحدار من النوع الهجين بين كلا من انموذج معلمي ذو صيغة معروفة مجهولة المعالم وانموذج لامعلمي لدالة انحدار مجهولة الصيغة،اي ان انموذج الانحدار يكون:

$$Y_i = (1-a)\, m_1(B,x) + a\, m_2(x) \qquad (7)$$

اذ يشير $m_1(B,x)$ الى نموذج الانحدار المعلمي بمعالم مجهولة وصيغة معلومة ، في حين $m_2(x)$ فتشير الى دالة الانحدار اللامعلميه. اما a فتشير الى معلمة الدمج مع كون $0 < a < 1$ ولتقدير هذه الدالة استخدم الباحثان **Burman and Chaudhuri** كلا من المقدرين المعلمي واللامعلمي للتوصل الى صيغة المقدر المدمج شبه المعلمي:

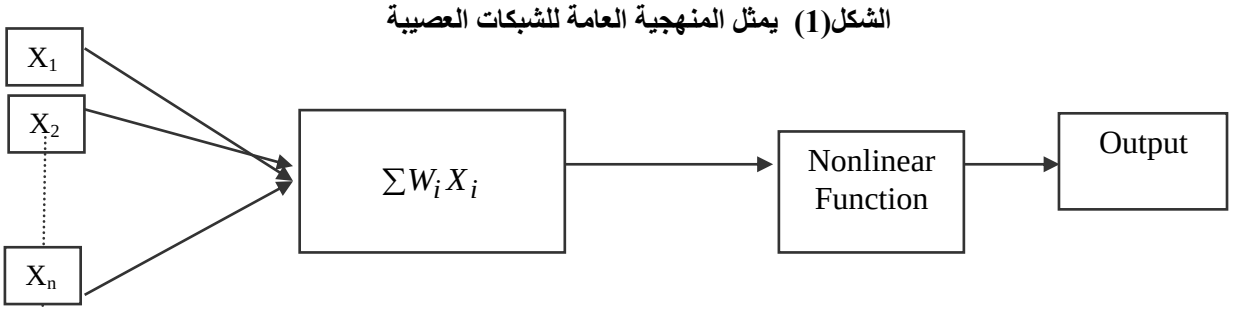
$$\hat{m}(X_i) = (1-\hat{a})\, m_1(\hat{B}, X_i) + \hat{a}\ \hat{m}_2(X_i) \qquad (8)$$

مع الاشارة الى ان $\hat{a}$ يشير الى مقدر المربعات الصغرى لمعلمة الدمج وعندما قيمة $\hat{a} = 0$ فأن المقدر (8) سوف يمثل المقدر المعلمي لدالة الانحدار ، اما عندما $\hat{a} = 1$ فأن المقدار (8) سوف يمثل المقدر اللامعلمي لدالة الانحدار.[9][16][22]

2-3 الشبكات العصيبية الاصطناعية : Artificial Neural Network (ANN)

تعد الشبكات العصبية الاصطناعية نظام معالجة معلومات تستند على نماذج رياضية بسيطة له مميزات أداء معينة بأسلوب يحاكي الشبكات العصبية البايولوجية " النظام العصبي ". وتعد إحدى انواع النماذج اللاخطية.وقد اقترحت الشبكة العصبية الاصطناعية،الممثلة بالمدرك الحسي "Perception" عام 1958 ،اذ يتكون المدرك الحسي من عصبونات أو عقد او وحدات تمثل المدخلات تحاكي الإشارات الداخلة الى الخلية العصبية البايولوجية وتمر الإشارات بين العقد "العصبونات" عبر خطوط ربط ويرفق كل خط ربط بوزن معين، و تضرب الإشارات الداخلة الى العقدة بهذه الاوزان وثم تجمع المدخلات الموزونة في العقد او العصبونات، وبعد ذلك يتم معالجة مخرجات كل عقدة بواسطة دالة غير خطية ذات عتبة معينة Threshold تعرف بدالة التنشيط Activation Function.والشكل الاتي يعرض المنهجية العامة للشبكات العصبية الاصطناعية.

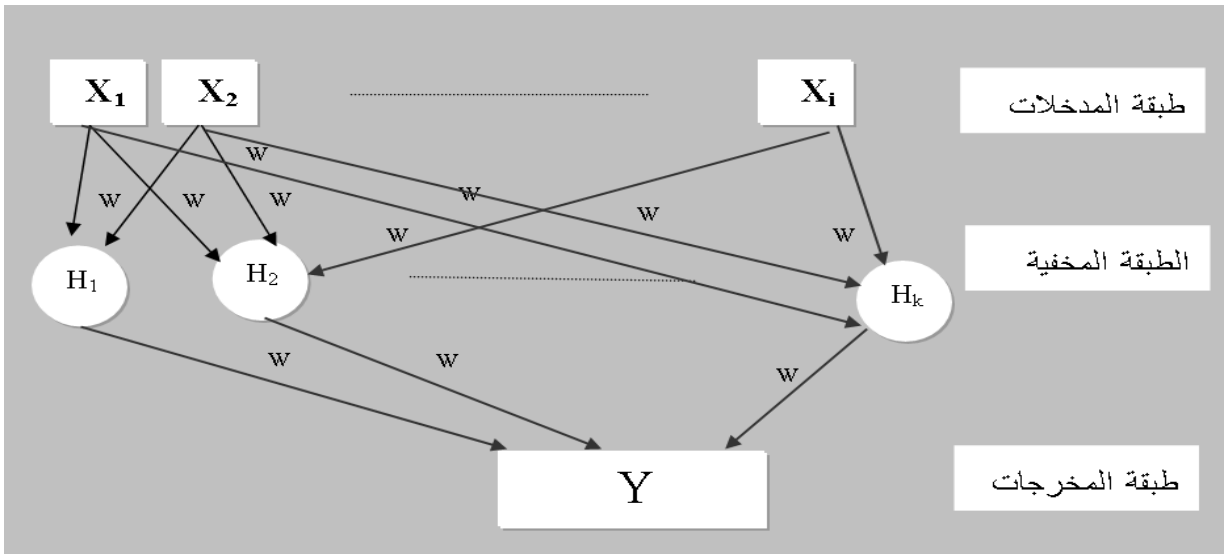
الشكل(1) يمثل المنهجية العامة للشبكات العصبية



2-3-1 البنية الهندسية للشبكات العصبية

تتكون الشبكة العصبية على الأقل من ثلاثة مستويات وكل مستوى يتضمن عدد من العقد ، اذ يمثل المستوى الأول المدخلات والمستوى الثاني يمثل المخرجات اما المستوى الأخر فيمثل الطبقة المخفية وتوجد هذه الطبقة ما بين مستوى المدخلات ومستوى المخرجات ويتضمن كل مستوى عددا من العقد وترتبط العقد في المستويات الثلاث بخطوط ربط ويرفق كل خط ربط بوزن معين والشكل الاتي يوضح البنية الهندسية للشبكات العصبية الاصطناعية.

شكل (3) يوضح البنية الهندسية للشبكات العصبية الاصطناعية



2-3-2 الشبكات العصبية ذات الانتشار العكسي للخطأ Back propagation Neural Nets

تأخر الاهتمام بالشبكات العصبية في السبعينيات بسبب محدودية استخدام الشبكات العصبية وحيدة الطبقة، ولكن اكتشاف فكرة الانتشار العكسي للخطأ لتدريب الشبكة متعددة الطبقات من قبل عدد من الباحثين في أواخر الثمانينات لعب الدور الأساس في إعادة بروز الشبكات العصبية كأداة لحل الكثير من المشاكل على نطاق واسع.

وبالرغم من ان الشبكات وحيدة الطبقة محدودة جدا في العمليات المناظرة "المقابلة" ألا انه باستطاعتها التعلم ، ولكن الشبكات العصبية متعددة الطبقة " بطبقة مخفية او اكثر "يمكن ان تتعلم أي عملية مقابلة مستمرة وذلك لدقة الاختيارية. ويمكن تعريف هذه الشبكة على أنها عبارة عن طريقة الانحدار التدريجي" gradient descent" وذلك لإيجاد القيمة الصغرى لمربع الخطأ الكلي لقيمة المخرجات المحسوبة من قبل الشبكة.وتستند الشبكة العصبية ذات الانتشار العكسي للخطأ على مفهوم تدريب الشبكة المعتمد على حجم الخطأ ، حيث تحدث الأوزان بين الطبقات بقدر مساهمة الوزن في تكوين الخطأ وتستمر الشبكة في تحديث الأوزان حتى الحصول على الأوزان ألمثلى التي تحقق افضل توفيق للنموذج.

2-3-3 خوارزمية الشبكة العصبية ذات الانتشار العكسي للخطأ

ان تدريب الشبكة العصبية باستخدام الانتشار العكسي يتضمن ثلاث مراحل هي:

- مرحلة الانتشار الأمامي للخطأ.
- مرحلة الانتشار الخلفي للخطأ.
- مرحلة توليف أوزان الشبكة.

خلال مرحلة الانتشار الأمامي تنتشر إشارة المدخلات الى كل عقدة من عقد الطبقة المخفية ثم يتم حساب قيمة التنشيط لكل عقدة من عقد الطبقة المخفية لهذه الإشارة وبعدئذ ترسل هذه العقد إشاراتها الى كل عقدة من عقد طبقة المخرجات ثم يتم حساب قيمة التنشيط لكل عقدة من عقد طبقة المخرجات لتشكل استجابة الشبكة لعينة المدخلات المعطاة.

وخلال مرحلة التدريب تقوم كل عقدة في طبقة المخرجات بمقارنة تنشيطاتها المحسوبة مع قيمة المخرجات الفعلية لتحدد قيمة الخطأ الحاصل لتلك العقدة. واعتمادا على قيمة الخطأ يتم حساب معامل تصحيح الخطأ δ_K ،اذ يستخدم معامل تصحيح الخطأ δ_K لتوزيع الخطأ على عقد طبقة المخرجات لتتم أعادته الى كل عقدة في الطبقة السابقة وكذلك يستخدم هذا المعامل لتحديث الأوزان بين طبقة المخرجات والطبقة المخفية.

وبطريقة مشابهه يتم حساب معامل تصحيح الخطأ δ_j بالنسبة لكل عقدة من عقد الطبقة المخفية ويستخدم هذا العامل لتحديث الاوزان بين الطبقة المخفية وطبقة المدخلات وبعد تحديد كل عوامل تصحيح الخطأ δ يتم توليف الاوزان بالنسبة لجميع الطبقات في نفس اللحظة.

2-3-4 خوارزمية الشبكة العصبية ذات الانتشار العكسية للخطأ

يمكن تلخيص خوارزمية او منهجية عمل هذه الشبكة بالخطوات الاتية:

- توليد قيم أولية للأوزان " من إحدى التوزيعات الإحصائية".
- تستقبل كل عقدة في طبقة المدخلات إشارة دخلها ثم إرسالها الى جميع عقد الطبقة المخفية.

تجمع كل عقدة في الطبقة المخفية قيم و إشارات دخلها الموزونة وبموجب المعادلة الاتية.

$$\mathrm{hj} = \mathrm{f}(\sum \mathrm{W_{ij}} X_i - \theta_j) \qquad (9)$$

- تطبيق تابع التنشيط لتقدير مخرجات الطبقة المخفية ، وترسل قيم التنشيط الى جميع العقد في طبقة المخرجات.
- تجمع كل عقدة في طبقة المخرجات إشارات دخلها الموزونة وبموجب المعادلة الآتية.

$$\mathrm{Y}_k = \mathrm{f}(\sum \mathrm{W_{jk}} h_j - \theta_k) \qquad (10)$$

- تطبيق تابع التنشيط لتقدير مخرجات طبقة المخرجات.
- مقارنة مخرجات الشبكة العصبية مع القيم الحقيقية لتقدير الخطأ وبموجب المعادلة الاتية.

$$\delta_k = (t_\mathrm{k} - Y_k).f(v) \qquad (11)$$

ومن ثم حساب التغير في حجم الخطأ ΔW_{JK} وبموجب المعادلة الاتية.

$$\Delta W_{jk} = \alpha \cdot \delta_k \cdot H_j \qquad (12)$$

- تجمع كل عقدة في الطبقة المخفية إشارات المدخلات الموزونة الـ δ بموجب المعادلة الاتية.

$$\Delta_J = \sum \delta_k \; W_{jk} \qquad (13)$$

ومن ثم تضرب هذه القيمة بتابع التنشيط لحساب δ_j. بعدها يتم حساب التغير في حجم الخطأ ΔV_{ij} وبموجب المعادلة الاتية.

$$\Delta V_{ij} = \alpha \cdot \delta_j \cdot X_i \qquad (14)$$

- تحديث الأوزان كل عقدة في طبقة المخرجات وبموجب المعادلة الاتية.

$$W_{jk}(new) = W_{jk}(old) + \Delta W_{jk} \qquad (15)$$

ومن ثم تحديث للأوزان بالنسبة لكل عقدة في الطبقة المخفية وبموجب المعادلة الاتية

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij} \qquad (16)$$

- وتستمر الشبكة في تحديث الأوزان " أي عملية التعلم وتدريب " الى ان يتم الحصول على الأوزان المثلى ومن ثم الحصول على المخرجات المرغوبة بها أي التوصل الى افضل توفيق للنموذج قيد البحث.

2-4 المعايير الإحصائية للمفاضلة:

لغرض المقارنة بين النتائج تم اعتماد المعايير الإحصائية آلاتية:

- متوسط مربعات الخطأ (Mean square error) :

اذ ان

$$MSE = (1/n)\, \Sigma e^2_t(L) \qquad (17)$$

- معيار متوسط مطلق الخطا النسبي(Mean absolute percentage error) :

اذ ان

$$MAPE = \{(1/n)\ \Sigma\, e_t(L)/X_{t+L}\} \qquad (18)$$

3- الجانب التطبيقي

سيتم في هذا المبحث تقدير صادرات النفط للمملكة العربية السعودية كدالة بدلالة حقول النفط المكتشفة للسنوات من 1981-2000.

وقد تم اخذ السعودية كدراسة حالة ودراسة صادراتها كونها احدى اهم الدول المصدرة للنفط في الخليج العربي واحدى كبريات الدول المصدرة للنفط في العالم فضلا عن استقررها السياسي والمالي والامني في منطقة الخليج العربي والشرق الاوسط عموما.

3-1 الرموز المستخدمة في الشبكة

- X_i متغير توضيحي يشير الى عدد الحقول المكتشفة للنفط للمدة من 1981-2000 ويمثل ايضا المدخلات الى الشبكة العصبية.
- Y_i متغير استجابة يشير الى كمية الصادرات من النفط الخام السعودي للمدة من 1981-2000 ويمثل ايضا مخرجات الشبكة العصبية.
- W يشير الى الأوزان ما بين المستويات
- θ تمثل العتبة Threshold
- $f(.)$ تشير الى دالة التنشيط Function Activation. وتقدر دالة التنشيط بموجب المعادلة الاتية.

$$f(v) = \exp(v) \qquad (17)$$

وتنحصر قيمة الدالة $f(v)$ بين (1 ,1-) وتمثل V_k المخرجات التوليفية الخطية.

- α معدل التعلم.

ومن خلال استخدام من المقدر اللامعلمي لدالة الانحدار في المعادلة (5) والمقدر شبه المعلمي في المعادلة (8) فضلا عن مقدر الشبكة العصبونية الاصطناعية في المعادلات (9-16) اذ استخدمت شبكة البث (الانتشار) العكسي للخطأ BP ودوال tansig ,trainim, Purelin وبالمستويات الاتية:

1-عدد عقد طبقة المدخلات 1
2- عدد عقد الطبقة المخفية 25
3- عدد عقد طبقة المخرجات 1

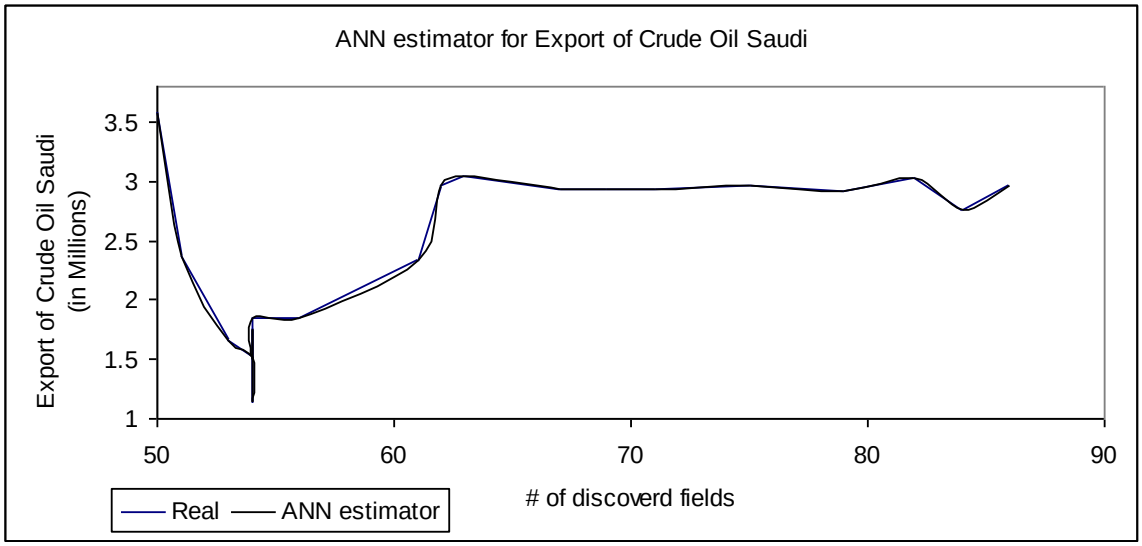
وقد تمت البرمجة باستخدام دوال شبكة العصبية للبث العكسي للخطأ في برنامج ماتلاب. وكانت النتائج كالاتي:

جدول رقم (1) يشير الى قيم متوسط مربعات الخطأ للمقدرات المستعملة

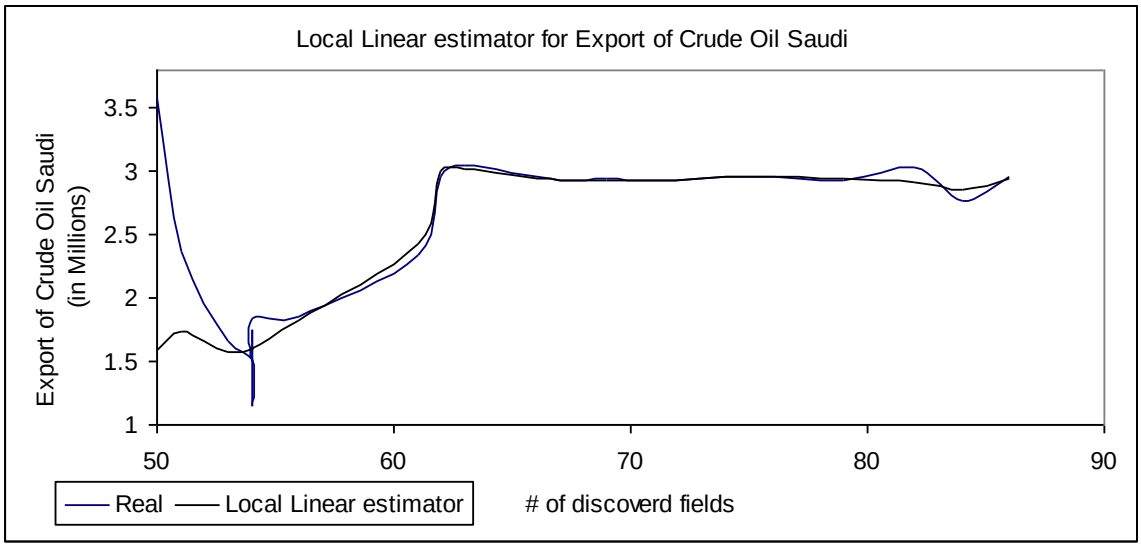
	المقدر اللامعلمي (المقدر الخطي لموضعي)	المقدر شبه المعلمي (المقدر المدمج)	مقدر الشبكة العصبية الاصطناعية
MSE	17742	31590	1.04847
MAPE	8.940	10.541	0.038

والاشكال (1،2،3) الاتية تشير الى المقدرات اللاخطية المستخدمة في البحث (مقدر الشبكة العصبية الاصطناعية، المقدر الخطي الموضعي والمقدر شبه المعلمي) مقارنة مع المنحنى الحقيقي لبيانات صادرت النفط الخام السعودي كمتغير تابع وكدالة لعدد حقول النفط المكتشفة كمتغير توضيحي.

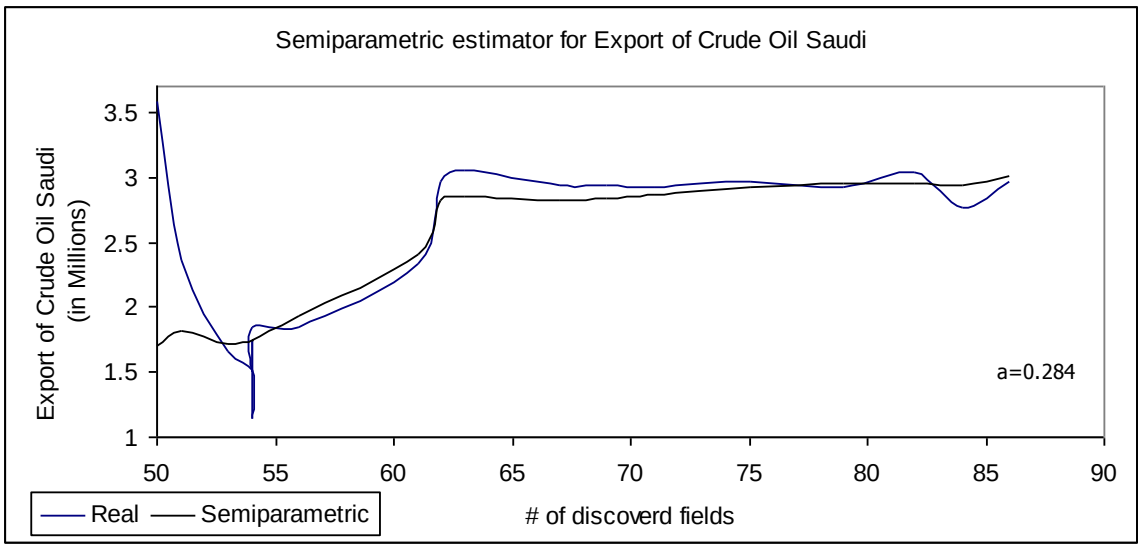
شكل (1)
يشير الى مقدر الشبكة العصبية الاصطناعية لصادرات النفط الخام السعودي كدالة بدلالة عدد حقول النفط المكتشفة مقارنة مع المنحنى الحقيقي للبيانات



شكل (2)
يشير الى المقدر الخطي الموضعي لصادرات النفط الخام السعودي كدالة بدلالة عدد حقول النفط المكتشفة مقارنة مع المنحنى الحقيقي للبيانات



شكل (3)
يشير الى المقدر شبه المعلمي لصادرات النفط الخام السعودي كدالة بدلالة عدد حقول النفط المكتشفة مقارنة مع المنحنى الحقيقي للبيانات



3-2 تحليل النتائج

من جدول (1) و الاشكال (3,2,1) على التوالي بينت النتائج والاشكال افضلية مقدر الشبكة العصبية الاصطناعية على كلا من لمقدر اللامعلمي الخطي الموضعي والمقدر شبه المعلمي في تقدير دالة الانحدار والمتمثلة بصادرات النفط الخام السعودي مستخدمين بذلك معيارين للخطأ هما:

1. معيار متوسط مربعات الخطأ(MSE) Mean square error.
2. معيار متوسط مطلق الخطا النسبي Mean absolute percentage error (MAPE)

ومن الجدير بالاشارة الى ان قيمة معلمة الدمج الخاصة بالمقدر شبه المعلمي والمقدرة باستخدام طريقة المربعات الصغرى كانت قيمتها هي a=0.284 ،ويلاحظ من الاشكال ان سلوك البيانات كان سلوكا غير خطيا في بداية الثمانينيات اذ كان هنالك تذبذبا في الصادرات وخاصة عام 1981 اذ وصلت الصادرات الى اعلى نسبه من بين الصادرات على مدى عشرون سنة تلاها انخفاضا كبيرا في السنوات اللاحقة وصل الى نصف صادرات عام 1981 تقريبا ويمكن عزو هذا الانخفاض الى تأثر الصادرات بالحرب العراقية-الايرانية في تلك الفترة وما اصابها من تلكوء وصول سفن الشحن الى الموانئ الخليجية .اما في السنوات التي تلت انتهاء الحرب العراقية-الايرانية لوحظ هنالك استقرارا في الصادرات وعلى مدى عقد من الزمن ويعزو هذا الاستقرار في الصادرات ايضا الى سيطرة منظمة الاوبك على سعر النفط واستقراره فضلا عن الزام الدول المنتجة للنفط بالابقاء على معدلات الانتاج والتصدير عند قيم مستقرة نسبيا.

4- الاستنتاجات والتوصيات

ان تمثيل الظواهر وفق نماذج محددة مسبقا خطية كانت او ما تسمى نماذج معلميه لا يفي بالغرض من الدراسة في اكثر التطبيقات العملية وهذا ماتم استنتاجه في هذا البحث اذ بينت النتائج افضلية المقدرات اللاخطية في تمثيل انموذج الدراسة علما ان المقدر الخطي المعلمي قد استخدم ضمنا ضمن المقدر شبه المعلمي مما اثر سلبا على التقديرات الناتجة لهذا المقدر الي بالنتيجة كان اقل المقدرات تمثيلا للظاهرة تحت الدراسة ولم يكن مقدرا مرضيا.

لذلك نوصي باستخدام المقدرات اللاخطية عموما ومقدر الشبكة العصبية الاصطناعية خصوصا بعد التحري عن البيانات ومعرفة سلوكها وبيان اتجاهها ويعود ذلك بسبب امتلاك الشبكات العصبية الاصطناعية خاصية التعلم والتكيف.

5- المصادر

اولا : المصادر العربية

1- الدليمي، محمد مناجد (1990) " تحليل الانحدار بالامثلة " مطابع التعليم العالي.
2- الناصر،عبد المجيد حمزة العبيدي،مروان عبد الحميد (2003)"لاستخدام الشبكات العصبية للتنبؤ المستقبلي بقيم السلاسل الزمنية غير منتظمة في طول الموسمية"،المجلة العراقية للعلوم الإحصائية- العدد5
3- محمد ، اسيل سمير (2005) " مقارنة بين طرائق تحليل وتنبؤ السلاسل الزمنية وتطبيقها على مبيعات الشركة العامة لتوزيع كهرباء بغداد" ،اطروحة دكتوراه - قسم الاحصاء - كلية الادارة والاقتصاد –جامعة بغداد.
4- عيسى، أسيل مسلم (2011) " مقارنة بعض المقدرات شبه المعلميه لتقدير دالة استهلاك الطاقة الكهربائية لمدينة بغداد"، رسالة ماجستير- قسم الاحصاء - كلية الادارة والاقتصاد – جامعة بغداد.
5- حمود، مناف يوسف، (2000) " مقارنة مقدرات Kernel اللامعلميه لتقدير دوال الانحدار ،رسالة ماجستير في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد .
6- حمود، مناف يوسف، (2005) " مقارنة المقدرات اللامعلميه لتقدير دوال الكثافة الاحتمالية " أطروحة دكتوراه في الإحصاء ،كلية الإدارة والاقتصاد، جامعة بغداد .
7- حمود، مناف يوسف (2010) " مقارنة مقدرات معلمة تمهيدية متغيرة باستخدام مقدر انحدار خطي موضعي مع تطبيق عملي " مجلة العلوم الاحصائية – العدد الثالث ، ص ص 43-63.
8- حمود، مناف يوسف وعيسى، أسيل مسلم (2012) " مقارنة بعض المقدرات شبه المعلميه لتقدير دالة الانحدار" بحث مقبول للنشر – مجلة العلوم الاقتصادية والادارية – جامعة بغداد.

ثانيا: المصادر الأجنبية

9- Ullah and H. D. Vinod (1993)“ General Nonparametric Regression Estimation and Testing in Econometrics” , in “Handbook of Statistics" (G. S. Maddala, C. R. Rao, and H. D. Vinod, Eds.), Vol. 11, pp. 85_117.
10- Bukhari, S. Adnan H.A. and Hanif, (2008) “Inflation Forecasting Using Artificial Neural Networks" (by Internet).
11- Daryl S. Paulson (2007) "Handbook of Regression & Modeling", Chapman & Hall/CRC.
12- Dong-Chul Prak (2011) "Prediction of Sunspot Series a Recurrent Neural Network”, IIE Publisher.
13- Edward, R.J (1997) "Neural Network Data Analysis Using Simulnet", Springer Verlag.
14- Fan. J and Gijbels .I (1992) “Variable Bandwidth and Local Linear Regression Smoothers" Ann. Stat. No-4.PP 2002-2036.
15- Fan. J and Gijbels .I (1996) “Local Polynomial Modeling and Its Applications" Chapman and Hall, London

16- Fan Y. and Ullah A. (1999) "Asymptotic Normality of a Combined Regression Estimator", Technical Report, University of Windsor.
17- Faraway, J. (1990) “Implementing Semiparametric Density Estimation “Statistics & Probability Letters Vol.-10- PP141-143.
18- Fox J. (2005) “Introduction to a nonparametric regression ‘’ lecture note, Dept. of Politics and international relations. McMaster University Canada.
19- Härdle, W. and Linton, O. (1994) “Applied Nonparametric Methods ", Cowles foundation for research in economics at Yale University.
20- Hardle, W. and Muller, M. (2000) “Multivariate and Semiparametric Kernel Regression", in. Schimek M. G (ed.), “Smoothing and Regression. Approaches, Computation and Application”, Wiley.
21- Olkin and Spieglman, C. H., (1987) “A semiparametric Approach to Density Estimation” JASA .Vol. 82, PP858_865.
22- Rahman, M., Gokhale, D.V and Ullah, A. (1997) "A Note on Combining Parametric and Non-parametric Regression”, Communications in Statistics - Simulation and Computation, Vol. 26, NO.2, PP 519 -529.
23- Stern. H.S. (1996)”Neural Networks in applied statistics” Techno metrics, 38(3):PP205-214.
24- Stetter Olav (2012)," Model-free reconstruction of neural network connectivity from calcium imaging signals", arXiv.
25- Ullah A. and Vinod,H.D. (1993) “ General nonparametric regression estimation and testing in econometrics” ,by Maddala, G.S. ,Rao,C.R. and Vinod H.D. “Handbook of statistics “ Vol.11,ch.4,PP 85- 115.
26- Xianggen Gan (2011),"The research of rainfall prediction model based on Matlab neural network", IIE Publisher.
27- Yan Bao-de (2001), “Perception and MATLAB”, Industrial Control Computer, Vol.14, NO.2, pp. 11-13.